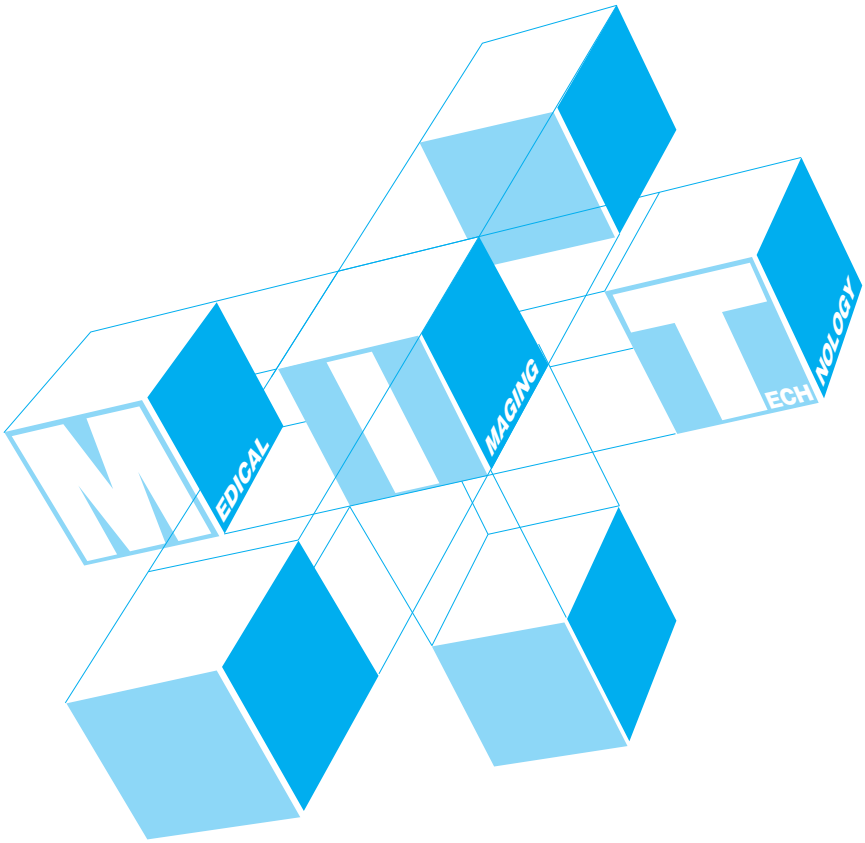
JAMIT Annual Meeting 2020

第39回日本医用画像工学会大会

会　期：2020 年 9 月 17 日(木)～9 月 19 日(土)
会　場：オンライン開催
大会長：湯浅 哲也(山形大学)



主　催　日本医用画像工学会
(The Japanese Society of Medical Imaging Technology)

後　援　一般社団法人　日本画像医療システム工業会
(JIRA:Japan Medical Imaging and Radiological Systems Industries Association)
一般社団法人　保健医療福祉情報システム工業会
(JAHIS:Japanese Association of Healthcare Information Systems Industry)
一般社団法人　電子情報技術産業協会
(JEITA:Japan Electronics and Information Technology Industries Association)

協　賛

医用画像情報学会	医用画像認知研究会
応用物理学会	可視化情報学会
画像電子学会	看護理工学会
三次元画像コンファレンス実行委員会	情報処理学会
電子情報通信学会	日本医学物理学会
日本医学放射線学会	日本医療情報学会
日本核医学会	日本核医学技術学会
日本画像医学会	日本コンピュータ外科学会
日本磁気共鳴医学会	日本写真学会
日本生体医工学会	日本超音波医学会
日本脳神経CI学会	日本放射線技術学会

学会長ご挨拶

第 39 回日本医用画像工学会大会は山形大学大学院教授 湯浅哲也大会長の運営のもとで，2020 年 9 月 17 日（木）から 19 日（土）の 3 日間，山形テルサにて開催されることになりました．一年以上前から準備をされてきた大会関係者や事務局の皆様，そして企業の業務委員会の皆様のご尽力に厚く御礼申し上げます．今年は，オリンピックが東京を中心に開催されるために，会期を 9 月に移動しての開催となっています．

日本医用画像工学会では学術大会の中で，AI 技術の医用画像工学への応用をこれまでさまざまな形で取り上げてきました．2016 年（千葉：羽石秀昭大会長）は教育講演という形でスタートし，2017 年（岐阜：藤田広志大会長）は教育講演とさまざまなセッションを，また，この大会から実際にコンピュータを使って講習を行うハンズオンセミナーを開始し，AI 技術を学ぼうとする研究者，企業の開発者，学生に対して実地訓練の場を提供してきました．そして，2018 年（筑波：工藤博幸大会長）以降は，AI 技術を大会のテーマに掲げ，特別講演，教育講演，シンポジウム，ハンズオンセミナーと展開してきました．昨年度の 2019 年（奈良：佐藤嘉伸大会長）には，AI 技術と画像診断・処理に重点をおいた企画が組まれ，今年の 2020 年の大会では「AI 画像：診断から治療へ」というテーマで，AI 技術と放射線治療の関係に踏み込んでいます．AI 技術は，もはや，一般的な技術になっている感がありますが，奥が深いので研究面，そして実用面での展開は限りがないように思います．

今年も，昨年に引き続き短時間の口頭発表（ティーザー）とポスター発表を組み合わせた発表形式になっており，すべての発表の概要を聞くことができるとともに，突っ込んだ議論がポスター会場でできるようになっています．また，特別講演で重粒子線治療の概要を知ることができるばかりでなく，実際の治療装置の見学ツアーも企画され，より理解が深まるものと思っています．この他にも，透過力の高いミューオンを用いたピラミッドのイメージングやオプティカル・コヒーレンス・トモグラフィ（OCT）による網膜イメージングなど，X 線やガンマ線を波動の主体としている医用画像処理の研究者に対しても目新しいテーマの講演を聴くことができます．さらに，政府が第 5 期科学技術基本計画の中で進めている Society5.0 の技術の医療面での社会実装に関する特別講演も用意され，盛りだくさんの内容となっています．この大会における皆様の活発な意見交換により，医用画像工学の研究が大きく進展することを期待しています．

日本医用画像工学会
会長　**尾川　浩一**（法政大学）

大会長ご挨拶

オンライン開催への移行について

すでに皆様に告知させていただいております第 39 回日本医用画像工学会（JAMIT）大会（2020 年 9 月 17 ～ 19 日：http://jamit2020.jamit.jp/）ですが，コロナ禍の影響のためオンライン開催へ移行することが常任幹事会（5/15 開催）で正式に決議されましたので，ご報告させていただきます．皆様を山形にお迎えすることはかなわなくなりましたが，これまでに用意したコンテンツはできるかぎり維持して開催する予定でおります．特別講演，シンポジウム，チュートリアル講演，ハンズオンセミナーはすべてオンラインにて当初の予定どおり実施させていただきます．また，MIT 誌査読付き論文および IJCARS（International Journal of Computer Assisted Radiology and Surgery）への投稿制度はそのまま維持いたします．

大会テーマは「AI 画像：診断から治療へ」です．AI の威力と影響力については今更言わずもがなでありましょうが，これまで JAMIT 大会では主に診断への応用が議論されてきました．そこで，本大会では，AI 画像処理の放射線治療への応用を一つのトピックにしようと考えております．ハードとソフトの調和により優れた治療システムが構築されるという観点に立ちプログラムを構成しました．まず，特別講演として放射線がん治療の権威である根本建二 先生を講師にお招きし，重粒子線がん治療の最前線についてご講演をお願いしています．次に，シンポジウムでは「粒子線治療と AI 画像技術」（オーガナイザー：臼井桂介 先生）というテーマで気鋭の先生方から AI を活用した治療支援技術に関するご講演をお願いしています．

もう一つのトピックとして，OCT（Optical Coherence Tomography）をシンポジウムのテーマにとりあげました．「OCT の技術的進歩と事業化、AI 適用、そして臨床応用」（オーガナイザー：秋葉正博 先生）というテーマで OCT の発明から現在までの進展を，OCT に関わる工学，医学，企業を代表する方々にご講演いただきます．

一方，医用画像処理とは少し距離をおきますが，「AI / IoT とヘルスケア」および「イメージング」をキーワードにお二方の先生をご招待させていただきました．お一方は，宮田裕章 先生です．報道番組の解説などでご存知の方も多いと思いますが，先生には，「Society 5.0 時代のヘルスケア」というテーマでご講演いただきます．ここでは，IoT や AI を軸に今後医療サービスがどのような変革を遂げ，どのような形態を取ってゆくべきかをご講演いただきます．感染症対策などに関するホットな話題についても聞けるのではないかと期待しております．もうお一方は，森島邦博 先生です．テレビ等でご存知の方も多いと思いますが，先生は，宇宙から降り注ぐミューオンを用いてピラミッドのレントゲン写真を撮像し，ピラミッド内部の未知の空間を発見されました．壮大なスケールのイメージングについてのお話しを聞いて，しばし太古に思いを馳せるというのはいかがでしょうか？

例年大好評を博しております深層学習ハンズオンは今大会もオンライン遠隔講義形式にて開催します．本年は，初級，中級，発展コースに分け，都合 5 回の開催を予定しております．事前受付制になっておりますので，ぜひ奮ってご応募ください．また，一般演題では，各演者に 2 分間の口頭発表をしていただく予定です．その後で演者と参加者が双方向でコミュニケイトできる場をご提供させていただきたいと考えております．他の学会での例を参考にしながら，オンライン開催の強みを十全に発揮させられるようにしてまいりたいと考えております．

コロナ禍により経済活動や社会活動のことごとくが停止・後退を強いられております．学術に関するアクティビティを維持するために学会の担う責任は非常に大きいものと考えます．かような状況下におきましても JAMIT 会員皆様が研究活動を促進できる機会をご提供できるよう最大限の努力をしてまいりたいと存じます．

第 39 回日本医用画像工学会大会
大会長　**湯浅　哲也**（山形大学）

参加者へのご案内

1. 会　期：2020 年 9 月 17 日(水)～9 月 19 日(土)

2. 開催形態：オンライン（ビデオ会議システム Zoom ウェビナー、ミーティングルーム使用）

3. 参加登録

1）オンライン登録：6 月 15 日(月)～9 月 19 日(土)

2）参加費

正会員（賛助会員含む）：13,000 円　　学生会員：5,000 円　　非会員：20,000 円

4. 総　会

日時：9 月 18 日(金)　12：40～13：40

5. 予稿集について

ダウンロード（PDF）形式で作成し、ダウンロードパスワードを参加者にお知らせいたします。

6. 発表者資格

一般演題の筆頭発表者は，日本医用画像工学会の会員に限ります．非会員の方は必ず 8 月 21 日(金)までに入会のお申込みをいただき，8 月 31 日(月)までに会費を納入してください．

7. 入会手続き及びお問い合わせ

日本医用画像工学会事務局（http://www.jamit.jp/）

〒104-0033　東京都中央区新川 1-5-19　6 階　（株）メイプロジェクト内

TEL: 03-6264-9071　FAX: 03-6264-8344　E-mail: jamit@may-pro.net

8. 利益相反（COI）について

筆頭発表者が，日本医用画像工学会における「利益相反の取扱いに関する規程」（http://www.jamit.jp/outline/agreement/rieki.html）において開示の対象となる場合，この規定に基づいて，発表者の申告した利益相反の状態が予稿集に開示されます．

9. 各種役員会

1）広報委員会　　9 月 17 日(木)　12：30～13：00

2）幹事会　　9 月 18 日(金)　11：45～12：35

3）編集委員会　　9 月 19 日(土)　11：45～12：30

10. 特別講演

特別講演１：9 月 18 日(金) 10：40～11：40

Society5.0 時代のヘルスケア

宮田　裕章（慶応義塾大学）

座　長：陣崎　雅弘（慶応義塾大学）

特別講演２：9 月 18 日(金) 13：00～13：50

山形大学医学部における重粒子線治療プロジェクト

根本　建二（山形大学）

座　長：尾川　浩一（法政大学／日本医用画像工学会 会長）

特別講演３：9 月 19 日(土) 10：40～11：40

宇宙線ピラミッドイメージング

森島　邦博（名古屋大学）

座　長：森　　健策（名古屋大学）

11. 第 10 回 JAMIT チュートリアル講演会（教育委員会企画）

（コニカミノルタ科学技術振興財団　JAMIT ハンズオンセミナー　連携企画）

「これから取り組む深層学習と研究の目標」

第 39 回日本医用画像工学会大会のプログラムの一環として「第 10 回日本医用画像工学会(JAMIT)チュートリアル講演会～これから取り組む深層学習と研究の目標～」を第 1 日目の午後に開催いたします．チュートリアルでは，医用画像工学に携わる若手研究者や最新の動向を得たい第一線の研究者を対象として，現在の研究に役立つテーマを専門の研究者が講演します．

日　時：9 月 17 日(木) 13：10～16：00

座　長：中田　典生（東京慈恵会医科大学）／小田　昌宏（名古屋大学）

講演１　基礎編：ディープラーニングの基礎～これから取り組む方へ～

李　　鎔範（新潟大学）

講演２　企業における技術展開：良い医療 AI への道

周　　暁妍／王　　禕楠（Infervision.Japan）

講演３　医師が取り組む AI 研究：腹部画像診断・IVR 診療における AI の潜在的臨床応用可能性

山田　　哲（信州大学）

12. コニカミノルタ科学技術振興財団　JAMIT ハンズオンセミナー

深層学習ハンズオンセミナー～ディープラーニングの標準的な手順を身につけよう～

（概要）深層学習の実行環境は，比較的簡単に構築できる．本ハンズオンでは，参加者は．事前に提供されるセットアップ資料に基づいて，自分自身のコンピュータに TensorFlow／Keras／Python による実行環境を構築し，その上で画像分類／領域分割／回帰／異常検知について，医用画像を利用した基本的な課題を実行する．また，必要に応じて DICOM ファイルの読み込みなどの前処理の方法や，構築したモデルの保存／読み込みといった実践的な方法についても触れる．そして，課題に応じた評価方法についても考え方と処理方法を実践的に習得する．

（本年の特徴）本ハンズオンは，コニカミノルタ科学技術振興財団助成事業として，初心者や初学

者である学生／若手研究者の支援を目的として実施する．学会期間中５回（各回９０分程度）のセミナーを実施する．５回の内容は以下の特徴がある．小さめのデータベースと取り組みやすい内容を用意する．

第１回　入門編：環境構築から自分の管理する画像の分類まで
第２回　画像分類：自分で構築するデータベースとその分類
第３回　領域分割：U-Net を利用した領域分割
第４回　回帰：画像から数値を予測
第５回　異常検知：AutoEncoder を利用した教師なし学習入門

（開催形態・会場）ウェブ開催される大会に合わせて，ハンズオンセミナーもウェブ上で開催します．事前に公開するセットアップ資料と題材に基づいて，セミナー開催までにご自身の PC に環境を構築してください．なお，オンライン参加のためには Zoom による接続が必要です．

（参加登録）ハンズオンの参加は，ハンズオン専用ウェブページから可能です．

http://micv.sakura.ne.jp/jamitho/

なお，JAMIT2020 大会への事前登録が必要です．JAMIT2020 大会の参加登録がない方は，ハンズオンに参加できません．JAMIT2020 大会参加者は，ハンズオンに無料で参加できます．

（登録後のスケジュール）９月　７日　セットアップ資料公開
９月１４日　コース資料の公開
９月１７～１９日　セミナー開催（オンライン）

（各回の予定：９０分）　００：００～００：０５　達成目標の確認
００：０５～００：１０　プログラム実行のデモ
００：１０～００：２０　自分の PC での実行確認
００：２０～００：３０　スクリプトの説明
００：３０～００：４０　結果の確認／評価方法
００：４０～００：５０　自分の PC での実行結果確認
００：５０～０１：００　問題点や限界の説明
０１：００～０１：１５　ご自身での改良，結果の確認
０１：１５～０１：２５　実習継続・質疑応答
０１：２５～０１：３０　全体まとめ

（運営スタッフ）　原　武史（岐阜大学）
李　鎔範（新潟大学）
中田典生（東京慈恵会医科大学）
小田昌宏（名古屋大学）

（協力）本セミナーは，コニカミノルタ科学技術振興財団の支援を得て実施しています．また，岐阜大学人工知能研究推進センター，東海国立大学機構医療健康データ統合研究教育拠点の協力で実施します．

13. シンポジウム

シンポジウム１：9 月 18 日(金)　15：20～16：50

粒子線治療と AI 画像技術

世話人／座長：臼井　桂介（順天堂大学）

講演１：診断から治療へ　-AI の積極的な臨床応用に向けて-

馬込　大貴（駒澤大学）

講演２：陽子線治療における画像誘導治療と応用

堀田　健二（国立がん研究センター東病院）

講演３：重粒子線治療の臨床運用における人工知能への期待と i-ROCK での取り組み

蓑原　伸一（神奈川県立がんセンター）

講演４：深層学習を用いた適応陽子線治療のための CBCT 散乱線補正と三次元陽子線線量分布計算法の開発

野村　友祐（北海道大学）

シンポジウム２：9 月 19 日(金)　12：40～14：10

OCT の技術的進歩と事業化、そして臨床応用の最前線

世話人／座長：秋葉　正博（トプコン）

講演１：座長のことば

秋葉　正博（トプコン）

講演２：第二世代 OCT の開発：基礎研究から市場へ

安野　嘉晃（筑波大学）

講演３：ダーウィンの荒波にもまれて…眼科用 OCT 事業化と更なる成長

福間　康文（トプコン）

講演４：眼科臨床における OCT の活用事例

中澤　　徹（東北大学）

講演５：AI 眼科診断・診療への応用に向けて

坂下　祐輔（ニデック）

講演６：血管内 OCT 技術と臨床での使われ方

平　　健二（ボストン・サイエンティフィックジャパン）

14. 特別企画

9 月 19 日（土）12:00－12:20

多種多様な情報が紐付いた大規模医療データベース

～次世代医療基盤法によって何が可能となったか～

小松　慶太（内閣官房　健康・医療戦略室）

座長：湯浅　哲也（山形大学／第 39 回日本医用画像工学会大会大会長）

ご 協 力 団 体

本大会は下記企業及び団体にご協力いただきました．心より感謝申し上げます．

後　　援：一般社団法人 日本画像医療システム工業会
一般社団法人 保健医療福祉情報システム工業会
一般社団法人 電子情報技術産業協会

助　　成：公益財団法人 コニカミノルタ科学技術振興財団

企 業 展 示：株式会社 HPC テック
キヤノンメディカルシステムズ株式会社

協 賛 金：エスペックテストシステム株式会社
株式会社トプコン
株式会社 PMOD

広 告 掲 載：医歯薬出版株式会社
株式会社インナービジョン
株式会社オーム社

協　　賛：医用画像情報学会
医用画像認知研究会
公益社団法人 応用物理学会
一般社団法人 可視化情報学会
一般社団法人 画像電子学会
看護理工学会
三次元画像コンファレンス実行委員会
一般社団法人 情報処理学会
一般社団法人 電子情報通信学会
公益社団法人 日本医学物理学会
公益社団法人 日本医学放射線学会
一般社団法人 日本医療情報学会
一般社団法人 日本核医学会
特定非営利活動法人 日本核医学技術学会
日本画像医学会
一般社団法人 日本コンピュータ外科学会
一般社団法人 日本磁気共鳴医学会
一般社団法人 日本写真学会
公益社団法人 日本生体医工学会
公益社団法人 日本超音波医学会
日本脳神経 CI 学会
公益社団法人 日本放射線技術学会

第１日目/9月17日(木)

	メインプログラム	ハンズオンセミナー	委員会	展示
			12:30-13:00 広報委員会	
13:00	13:00－13:10　開会式			展示 株式会社HPCテック
13:10	第10回JAMITチュートリアル講演会(教育委員会企画) コニカミノルタ科学技術振興財団、JAMITハンズオンセミナー連携企画 TL「これから取り組む深層学習と研究の目標」 13:10-14:00 TL1「基礎編：ディープラーニングの基礎～これから取り組む方へ～」 李　　鎔範 14:10-15:00 TL2「企業における技術展開：良い医療AIへの道」 周　　暁妍／王　　禕楠 15:10-16:00 TL3「医師が取り組むAI研究：腹部画像診断・IVR診療におけるAIの潜在的臨床応用可能性」 山田　　哲 座長：中田　典生／小田　昌宏			
16:00	休憩			
16:15	16:15-17:45　一般演題セッション１（前半：口頭発表　後半：質疑応答） P1「CAD」(P1-1~24) 座長：湯浅　哲也	16:15-17:45 ハンズオンセミナー1 HS1「入門編：環境構築から自分の管理する画像の分類まで」		

第2日目/9月18日(金)

	メインプログラム	ハンズオンセミナー	委員会	展示
9:00	9:00-10:30　一般演題セッション２（前半：口頭発表　後半：質疑応答） P2「医用画像処理」(P2-1~24) 座長：湯浅　哲也	9:00-10:30 ハンズオンセミナー2 HS2「画像分類：自分で構築するデータベースとその分類」		展示 株式会社HPCテック
10:30	休憩			
10:40	10:40－11:40　特別講演1 SL1「Society 5.0時代のヘルスケア」 宮田　裕章 座長：陣崎　雅弘			
11:40	休憩		11:45-12:35 幹事会	
12:40	12:40－13:40　総会			
13:40	休憩			
14:00	14:00－15:00　特別講演2 SL2「山形大学医学部における重粒子線治療プロジェクト」 根本　建二 座長：尾川　浩一			
15:00	休憩			
15:20	15:20-16:50　シンポジウム1 SY1「粒子線治療とAI画像技術」 馬込　大貴／堀田　健二／衰原　伸一／野村　友祐 座長：臼井　桂介			
16:50	休憩			
17:00	17:00-18:30　一般演題セッション３（前半：口頭発表　後半：質疑応答） P3「イメージング・再構成・画像生成」(P3-1~19) 座長：湯浅　哲也	17:00－18:30 ハンズオンセミナー3 HS3「領域分割：U-Netを利用した領域分割」		

第3日目/9月19日(土)

	メインプログラム	ハンズオンセミナー	委員会	展示
9:00	9:00-10:30　一般演題セッション４（前半：口頭発表　後半：質疑応答） P4「CT・核医学画像・MRI」(P4-1~20) 座長：湯浅　哲也	9:00-10:30 ハンズオンセミナー4 HS4「回帰：画像から数値を予測」		展示 株式会社HPCテック
10:30	休憩			
10:40	10:40-11:40　特別講演3 SL3「宇宙線ピラミッドイメージング」 森島　邦博 座長：森　　健策			
11:40	休憩		11:45-12:30 編集委員会	
12:00	12:00－12:20　特別企画 SE「多種多様な情報が紐付いた大規模医療データベース ～次世代医療基盤法によって何が可能となったか～」 小松　慶太 座長：湯浅　哲也			
12:20	休憩			
12:40	12:40－14:10　シンポジウム２ SY2「OCTの技術的進歩と事業化、そして臨床応用の最前線」 秋葉　正博／安野　嘉晃／福間　康文／中澤　　徹／坂下　祐輔／平　　健二 座長：秋葉　正博			
14:10	休憩			
14:20	14:20－15:50　一般演題セッション5（前半：口頭発表　後半：質疑応答） P5「CAS・生体情報・その他」(P5-1~24) 座長：湯浅　哲也	14:20－15:50 ハンズオンセミナー5 HS5「異常検知：AutoEncoderを利用した教師なし学習入門」		
15:55	15:55－16:00　閉会式			

特別講演

第2日　9月18日（金）

特別講演1

10:40 ～ 11:40　**SL1「Society5.0 時代のヘルスケア」**

座長：陣崎　雅弘（慶応義塾大学）

SL1　Society5.0 時代のヘルスケア
宮田　裕章（慶応義塾大学医学部医療政策・管理学教室　教授）

特別講演2

14:00 ～ 15:00　**SL2「山形大学医学部における重粒子線治療プロジェクト」**

座長：尾川　浩一（法政大学／ JAMIT 学会長）

SL2　山形大学医学部における重粒子線治療プロジェクト
根本　建二（山形大学　理事・副学長／山形大学医学部東日本重粒子センター　センター長）

第3日　9月19日（土）

特別講演3

10:40 ～ 11:40　**SL3「宇宙線ピラミッドイメージング」**

座長：森　　健策（名古屋大学）

SL3　宇宙線ピラミッドイメージング
森島　邦博（名古屋大学大学院理学研究科　特任助教／ JST さきがけ研究者）

特別企画

第3日　9月19日（土）

特別企画

12:00 ～ 12:20　**SE「多種多様な情報が紐付いた大規模医療データベース**
～次世代医療基盤法によって何が可能となったか～」

座長：湯浅　哲也
（山形大学／ JAMIT2020 大会長）

SE　多種多様な情報が紐付いた大規模医療データベース
～次世代医療基盤法によって何が可能となったか～
小松　慶太（内閣官房　健康・医療戦略室　主査　（併）内閣府，経済産業省）

第 10 回 JAMIT チュートリアル講演会

（教育委員会企画・コニカミノルタ科学技術振興財団　JAMIT ハンズオンセミナー連携企画）

第1日　9月17日（木）

第 10 回 JAMIT チュートリアル講演会

13:10 ～ 16:00　**TL「これから取り組む深層学習と研究の目標」**

座長：中田　典生（東京慈恵会医科大学）
座長：小田　昌宏（名古屋大学）

TL1　基礎編：ディープラーニングの基礎～これから取り組む方へ～
李　　鎔範（新潟大学）

TL2　企業における技術展開：良い医療 AI への道
周　　暁妍（株式会社 Infervision.Japan　代表取締役）
王　　禕楠（株式会社 Infervision.Japan　プロダクトマネージャー）

TL3　医師が取り組む AI 研究：腹部画像診断・IVR 診療における AI の潜在的臨床応用可能性
山田　　哲（信州大学）

シンポジウム

第２日　９月18日（金）

シンポジウム１

15:20 ～ 16:50　**SY1「粒子線治療と AI 画像技術」**

座長：臼井　桂介（順天堂大学）

SY1-1　診断から治療へ　-AI の積極的な臨床応用に向けて -
馬込　大貴（駒澤大学）

SY1-2　陽子線治療における画像誘導治療と応用
堀田　健二（国立がん研究センター東病院）

SY1-3　重粒子線治療の臨床運用における人工知能への期待と i-ROCK での取り組み
蓑原　伸一（神奈川県立がんセンター）

SY1-4　深層学習を用いた適応陽子線治療のための CBCT 散乱線補正と三次元陽子線線量分布計算法の開発
野村　友祐（北海道大学）

第３日　９月19日（土）

シンポジウム２

12:40 ～ 14:10　**SY ２「OCT の技術的進歩と事業化、そして臨床応用の最前線」**

座長：秋葉　正博（株式会社トプコン）

SY2-1　座長のことば
秋葉　正博（株式会社トプコン）

SY2-2　第二世代 OCT の開発：基礎研究から市場へ
安野　嘉晃（筑波大学）

SY2-3　ダーウィンの荒波にもまれて…眼科用 OCT 事業化と更なる成長
福間　康文（株式会社トプコン）

SY2-4　眼科臨床における OCT の活用事例
中澤　　徹（東北大学）

SY2-5　AI 眼科診断・診療への応用に向けて
坂下　祐輔（株式会社ニデック）

SY2-6　血管内 OCT 技術と臨床での使われ方
平　　健二（ボストン・サイエンティフィックジャパン株式会社）

コニカミノルタ科学技術振興財団
JAMIT ハンズオンセミナー

（運営スタッフ）原　　武史（岐阜大学／ JAMIT 教育委員会）
中田　典生（東京慈恵会医科大学）
小田　昌宏（名古屋大学）
李　　鎔範（新潟大学）

第 1 日　9 月 17 日（木）

16:15 ～ 17:45　**HS1「ディープラーニングの標準的な手順を身につけよう」**

HS1　入門編：環境構築から自分の管理する画像の分類まで

第 2 日　9 月 18 日（金）

ハンズオンセミナー 2

9:00 ～ 10:30　**HS2「ディープラーニングの標準的な手順を身につけよう」**

HS2　画像分類：自分で構築するデータベースとその分類

ハンズオンセミナー 3

17:00 ～ 18:30　**HS3「ディープラーニングの標準的な手順を身につけよう」**

HS3　領域分割：U-Net を利用した領域分割

第 3 日　9 月 19 日（土）

ハンズオンセミナー 4

9:00 ～ 10:30　**HS4「ディープラーニングの標準的な手順を身につけよう」**

HS4　回帰：画像から数値を予測

ハンズオンセミナー５

14:20 ～ 15:50　**HS5「ディープラーニングの標準的な手順を身につけよう」**

HS5　異常検知：Auto Encoder を利用した教師なし学習入門

第 1 日　9 月 17 日（木）

13:00 ～ 13:10　**開会式**

第 10 回 JAMIT チュートリアル講演会

13:10 ～ 16:00　**TL「これから取り組む深層学習と研究の目標」**

座長：中田　典生（東京慈恵会医科大学）
座長：小田　昌宏（名古屋大学）

TL1　基礎編 : ディープラーニングの基礎～これから取り組む方へ～
李　　鎔範（新潟大学）

TL2　企業における技術展開 : 良い医療 AI への道
周　　暁妍（株式会社 Infervision.Japan　代表取締役）
王　　禕楠（株式会社 Infervision.Japan　プロダクトマネージャー）

TL3　医師が取り組む AI 研究 : 腹部画像診断・IVR 診療における AI の潜在的臨床応用可能性
山田　　哲（信州大学）

第 1 日　9 月 17 日（木）

一般演題セッション 1

16:15 ～ 17:45　**P1「CAD」**

座長：湯浅　哲也（山形大学）

P1-1　単純腹部 CT 画像からの高精細な大腸領域検出法の検討
安倍　和弥（神奈川工科大学）

P1-2　特徴量選択手法およびシステム構成の最適化によるうつ病 CAD の高性能化
牧　　優太（神奈川工科大学）

P1-3　画像と音声情報を用いたうつ病 CAD の高性能化
和田　昇太（神奈川工科大学）

P1-4　敵対的生成ネットワークを用いて生成した硬性白斑画像についての検討
藤田　真穂（滋賀県立大学大学院工学研究科電子システム工学専攻）

P1-5　Modified Cycle-Consistent Adversarial Network を用いた病理組織画像における染色度合いの正規化を伴う病理組織分類
林　　大誠（立命館大学理工学部 電子情報工学科）

P1-6 SVM と時空間モデルを用いた嚥下時 X 線動画からの喉頭蓋の抽出に関する基礎的検討
杉田　絢美（筑波大学）

P1-7 U-Net を用いた X 線動画からの頸椎椎間板の抽出に関する基礎的検討
郡司絵莉華（筑波大学）

P1-8 ２種類の３次元畳み込みニューラルネットワークを用いた造影後 T1 強調画像における神経膠腫の全自動グレード分類パイプライン
山城　滉斗（藤田医科大学大学院保健学研究科）

P1-9 Mask R-CNN とテクスチャ解析による小児肺炎の検出
魚住　春日（藤田医科大学大学院 保健学研究科）

P1-10 V-Net を用いた Ai-CT に対する肝臓領域抽出手法の開発と肝臓の質量推定への応用
櫛部　光哉（山口大学大学院創成科学研究科）

P1-11 心臓超音波動画像における畳み込みニューラルネットワークを用いた心筋梗塞の自動分類手法の開発
村木　亮介（藤田医科大学大学院保健学研究科）

P1-12 体幹部 CT 画像における U-Net アーキテクチャによる体腔の自動認識
前田　泰良（愛知県立大学 情報科学部）

P1-13 機械学習に基づく食道・頭頸部癌早期発見の為の診断支援システムに関する検討
吉村裕一郎（千葉大学 フロンティア医工学センター）

P1-14 ディープラーニングに基づく３次元汎用半自動病変セグメンテーションの初期検討
野村　行弘（東京大学医学部附属病院コンピュータ画像診断学／予防医学講座）

P1-15 大腸内視鏡における穿孔の自動検出および位置推定に関する予備的検討
蒋　　凱（名古屋大学 大学院情報学研究科）

P1-16 機械学習によるコルポスコープ画像における白色上皮領域検出の基礎的検討
篠原　寿広（近畿大学生物理工学部生命情報工学科）

P1-17 One-class SVM による画素悪性度の集中性を用いた FDG-PET/CT 上の病変自動検出
田中　敦子（近畿大学大学院 生物理工学研究科）

P1-18 教師なし学習を利用した異常検知の PET 画像における腫瘍検出への応用
前田　健宏（岐阜大学大学院自然科学技術研究科知能理工学専攻）

P1-19 読影レポート解析を利用した医用画像データベースからのアノテーション付きデータセット作成に関する初期検討
林　雄一郎（名古屋大学大学院情報学研究科）

P1-20 極度に不均衡な頭部 CT データにおける転移学習を用いたクモ膜下出血自動検出の検討
魯　　仲陽（名古屋大学大学院情報学研究科）

P1-21 顎骨骨髄炎発症範囲推定のための頭部 CT・切除顎骨 CT の位置合わせ自動化
松岡　志宜（三重大学大学院工学研究科情報工学専攻）

P1-22 深層学習を用いた CT 画像解析による肺腺癌手術患者における術後再発予測
佐々木雄樹（新潟県立がんセンター新潟病院　中央放射線部）

P1-23 COVID-19 症例の定量評価のための CT 像からの肺野自動セグメンテーション
小田　昌宏（名古屋大学大学院情報学研究科）

P1-24　膀胱鏡検査画像を用いた間質性膀胱炎の分類に関する予備的検討
初　　韜（名古屋大学 大学院情報学研究科）

第1日　9月17日（木）

ハンズオンセミナー1

16:15 ～ 17:45　**HS1「ディープラーニングの標準的な手順を身につけよう」**

HS1　入門編：環境構築から自分の管理する画像の分類まで

第2日　9月18日（金）

一般演題セッション2

9:00 ～ 10:30　**P2「医用画像処理」**

座長：湯浅　哲也（山形大学）

P2-1　トポロジカル不変量を用いたレディオミクス特徴量に基づく肺癌予後推定法
二宮　健太（九州大学大学院　医学研究院保健学部門）

P2-2　マルチチャンネル化と CNN を用いた嚥下時 X 線動画からの頸椎椎間板の抽出：LeNet, AlexNet, GoogLeNet の性能比較
藤中　彩乃（筑波大学）

P2-3　取り下げ

P2-4　Automated whole-heart segmentation using U-Net in pediatric cardiac CT.
吉田　皓文（新潟医療福祉大学医療技術学部診療放射線学科）

P2-5　Deep Learning を用いた非球形状マーカの認識手法の開発
阿部　和貴（北海道大学大学院医理工学院）

P2-6　骨塩定量ファントムを用い臨床用 CT 画像から大腿骨骨密度を自動計測するシステムの開発
上村　圭亮（奈良先端科学技術大学院大学生体医用画像研究室）

P2-7　Conversion of H&E Stained to EVG Stained Histological Images using CycleGAN
ビシュワストンニ（情報通信系 , 東京工業大学）

P2-8　拡張型 / 肥大型心筋症病理画像の教師なし分類と評価
梁　　天恒（奈良先端科学技術大学院大学）

P2-9　Radiomics 画像特徴量による肺がん治療方針決定因子 PD-L1 発現予測の検討
松崎　和喜（日立製作所　ヘルスケアビジネスユニット）

P2-10 Dilated convolution を用いた FCN による腹腔鏡動画像からの血管領域抽出
盛満慎太郎（名古屋大学大学院情報学研究科）

P2-11 少数症例を用いた教師無し深層学習による病変検出画像特徴量の汎用的生成 – 胸部 CT を用いた実験的検証 –
牛房　和之（近畿大学大学院 生物理工学研究科）

P2-12 深層学習を用いた歯科インプラントの分類
吉位　和将（岐阜大学工学部電気電子・情報工学科）

P2-13 大域及び局所情報を用いた深層学習による出血領域自動セグメンテーション
山本　翔太（名古屋大学大学院情報学研究科）

P2-14 深層学習による大規模多様皮膚疾患画像データベースの分類
田中　雅也（東京農工大学大学院工学研究院）

P2-15 個別化運動器科学に向けた CT 画像からの下腿部・足部筋骨格の自動セグメンテーション
スーフィーマーゼン（奈良先端科学技術大学院大学）

P2-16 機械学習を用いたハイパースペクトルによる K-ras 遺伝子の識別
櫻井　千子（一般財団法人 未来科学研究所）

P2-17 機械学習を用いたハイパースペクトルによる肺癌の分類
江崎　雅隆（一般財団法人 未来科学研究所）

P2-18 広範囲の隣接関係を考慮したグラフニューラルネットワークを用いた腹部動脈血管名自動命名の検討
日比　裕太（名古屋大学大学院情報学研究科）

P2-19 歯科パノラマ X 線画像における分岐を含む SSD による 1 クラス検出と多クラス検出の結果の統合による歯列の認識手法
森下　拓水（岐阜大学大学院知能理工学専攻知能情報学領域）

P2-20 畳み込みニューラルネットワークによるクロスフェーズ CT 画像位置合わせ
胡　　涛（名古屋大学大学院情報学研究科）

P2-21 P5-24 に変更

P2-22 造影増強 MR 画像におけるベイズ U-Net に基づく不確実性推定を用いた肝臓及び脾臓のセグメンテーション精度予測
張　　彬（奈良先端科学技術大学院大学）

P2-23 皮弁壊死領域早期診断のための狭帯域光源・カラーカメラ連動システム設計
倉渕　瑶子（千葉大学大学院融合理工学府）

P2-24 画素毎の難易度に基づく注意機構を用いた敵対的学習によるセマンティックセグメンテーション
津田　大輝（名城大学）

ハンズオンセミナー２

9:00 ～ 10:30　**HS2「ディープラーニングの標準的な手順を身につけよう」**

HS2　画像分類：自分で構築するデータベースとその分類

第２日　９月18日（金）

特別講演１

10:40 ～ 11:40　**SL1「Society5.0 時代のヘルスケア」**

座長：陣崎　雅弘（慶応義塾大学）

SL1　Society5.0 時代のヘルスケア
宮田　裕章（慶応義塾大学医学部医療政策・管理学教室　教授）

12:40 ～ 13:40　**総会**

特別講演２

14:00 ～ 15:00　**SL2「山形大学医学部における重粒子線治療プロジェクト」**

座長：尾川　浩一（法政大学／ JAMIT 学会長）

SL2　山形大学医学部における重粒子線治療プロジェクト
根本　建二（山形大学　理事・副学長／山形大学医学部東日本重粒子センター　センター長）

シンポジウム１

15:20 ～ 16:50　**SY1「粒子線治療と AI 画像技術」**

座長：臼井　桂介（順天堂大学）

SY1-1　診断から治療へ　-AI の積極的な臨床応用に向けて -
馬込　大貴（駒澤大学）

SY1-2　陽子線治療における画像誘導治療と応用
堀田　健二（国立がん研究センター東病院）

SY1-3　重粒子線治療の臨床運用における人工知能への期待と i-ROCK での取り組み
蓑原　伸一（神奈川県立がんセンター）

SY1-4　深層学習を用いた適応陽子線治療のための CBCT 散乱線補正と三次元陽子線線量分布計算法の開発
野村　友祐（北海道大学）

一般演題セッション３

17:00 ～ 18:30　**P3「イメージング・再構成・画像生成」**

座長：湯浅　哲也（山形大学）

P3-1　3D 血管撮影のための少数方向トモシンセシスの画像再構成法
堀　　拳輔（杏林大学大学院保健学研究科）

P3-2　取り下げ

P3-3　畳み込みニューラルネットワークによる MR 及び PET 画像からの酸素摂取率画像の予測
松原　佳亮（秋田県立循環器・脳脊髄センター　脳血管研究センター放射線医学研究部）

P3-4　MR 画像の位相推定を不要とする深層学習型圧縮センシング再構成法の検討
大内　翔平（宇都宮大学大学院 工学研究科 システム創成工学専攻）

P3-5　少数投影方向マイクロ CT 像の補正手法の開発
岡本　尚之（千葉大学大学院融合理工学府）

P3-6　画像処理法を用いた冠動脈 CTA 石灰化プラーク除去法の開発と臨床評価
渋谷　幸喜（北海道科学大学　保健医療学部　診療放射線学科）

P3-7　位相拡散フーリエ法の信号－画像間学習による MR 画像の超解像
若槻　泰迪（宇都宮大学大学院 地域創生科学研究科 情報電気電子システム工学プログラム）

P3-8　位相拡散フーリエ変換映像法の導入による MR 深層学習再構成の高画質化
植松　　駿（宇都宮大学大学院 地域創生科学研究科
工農総合科学専攻 情報電気電子システム工学プログラム）

P3-9　数値シミュレーションを用いた非等方ボクセルに対応した粘弾性率分布推定手法の定量性評価
武田　将岳（千葉大学　大学院融合理工学府）

P3-10　U-Net による X 線透視血管造影像からの血管可視化技術の開発：時空間情報を利用した初期的検討
木村　竜誠（藤田医科大学大学院　保健学研究科）

P3-11　深層学習を用いた圧縮センシング CT 画像再構成の画質改善
工藤　博幸（筑波大学システム情報系）

P3-12　トポロジー制約を用いたスパースビュー CT 画像再構成の試み
千北　一期（筑波大学システム情報工学研究科）

P3-13　深層学習による圧縮センシング CT 画像再構成画質改善の新手法
数尾　　昴（筑波大学システム情報系）

P3-14　CT 画像再構成における混合インテリアエクステリア問題の解の一意性
藤井　克哉（筑波大学システム情報系）

P3-15　多方向 2 次元超解像を用いた臨床 CT 像の教師なし 3 次元超解像
鄭　　通（名古屋大学大学院情報学研究科）

P3-16　敵対的生成ネットワークによる屈折コントラスト CT のためのリングアーチファクト除去
黄　　卓然（名古屋大学大学院医学系研究科）

P3-17　深層学習を用いたヘマトキシリン・エオジン染色画像から Ki-67 免疫組織化学染色画像への変換
季　　存遠（東京工業大学　工学院　情報通信系）

P3-18　取り下げ

P3-19　マルチピンホール X 線蛍光 CT 画像再構成における EM-TV アルゴリズムの検討
金澤　道和（弘前大学 大学院理工学研究科）

第２日　９月18日（金）

ハンズオンセミナー３

17:00 ～ 18:30　**HS3「ディープラーニングの標準的な手順を身につけよう」**

HS3　領域分割：U-Net を利用した領域分割

第３日　９月19日（土）

一般演題セッション４

9:00 ～ 10:30　**P4「CT・核医学画像・MRI」**

座長：湯浅　哲也（山形大学）

P4-1　Swish を利用した Blind DnCNN とパラレルデノイジングへの応用
須貝　太郎（宇都宮大学　大学院地域創生科学研究科　工農総合科学専攻
情報電気システム工学プログラム）

P4-2　アーチファクト学習を導入した改良型 ISTA-Net による深層学習再構成
宮本　裕大（宇都宮大学 大学院 地域創生科学研究科
工農総合科学専攻 情報電気電子システム工学プログラム）

P4-3　放射光マイクロ CT におけるリングアーチファクト低減の検討
米山　明男（九州シンクロトロン光研究センター）

P4-4　MRI 高速撮像画像における微細構造を保存したノイズ低減
石原千鶴枝（株式会社 日立製作所 研究開発グループ）

P4-5　Support Vector Regression を用いた脳年齢推定と機種間 Harmonization
舞草　伯秀（国立研究開発法人　国立・精神神経医療研究センター）

P4-6　CNN を用いたイオフルパン SPECT からの特異的結合比の推定
久根崎遥加（山形大学大学院理工学研究科）

P4-7 C 型 Compton-PET ジオメトリの 3 次元イメージングシミュレーション
仁科　匠（千葉大学）

P4-8 深層学習による組織組成比推定を介した PET 吸収係数画像生成
小林　哲哉（株式会社 島津製作所 基盤技術研究所）

P4-9 WGI 試作機によるフルリングコンプトンイメージングの実証
田島　英朗（量子科学技術研究開発機構放射線医学総合研究所）

P4-10 MR 画像特徴量を用いた機械学習モデルによる前立腺癌患者の低悪性度群と高悪性度群への自動層別化
浦上　暉允（九州大学医学系学府保健学専攻）

P4-11 複数周波数同時加振 MR エラストグラフィによる撮像時間短縮の検討
久慈　陸（千葉大学）

P4-12 数値シミュレーションによる MR エラストグラフィに体動が与える影響の基礎的検討
小賀野　慎（千葉大学 大学院融合理工学府）

P4-13 機械学習を用いたコーンビーム CT の散乱線除去
村田　一心（法政大学理工学部応用情報工学科）

P4-14 デコンボリューション法による静止型ピンホール SPECT システムの空間分解能改善
大越　迪（法政大学大学院理工学研究科）

P4-15 At-211 内用療法おける ETCC を用いた画像再構成
山田　宙輝（法政大学大学院理工学研究科）

P4-16 PET/CT 画像および統計的画像を用いた悪性腫瘍の自動検出法
相原　拓実（岐阜大学工学部電気電子・情報工学科）

P4-17 手部 X 線画像を用いた骨髄浮腫画像分類の初期検討
潘　冬平（名古屋大学大学院情報学研究科）

P4-18 屈折コントラスト法間の画像描写能比較の試み
大塚　楠平（北海道科学大学大学院）

P4-19 1.5 T MRI における定量的磁化率マッピング（QSM）の画質評価
高橋　哲彦（群馬県立県民健康科学大学 診療放射線学部）

P4-20 深層学習を用いたマルチピンホール SPECT システムの画像再構成
紺野　宏彰（法政大学大学院理工学研究科）

第 3 日　9 月 19 日（土）

ハンズオンセミナー 4

9:00 ～ 10:30　**HS4「ディープラーニングの標準的な手順を身につけよう」**

HS4 回帰：画像から数値を予測

特別講演３

10:40 ～ 11:40　**SL3「宇宙線ピラミッドイメージング」**

座長：森　　健策（名古屋大学）

SL3　宇宙線ピラミッドイメージング
森島　邦博（名古屋大学大学院理学研究科　特任助教／ JST さきがけ研究者）

特別企画

12:00 ～ 12:20　**SE「多種多様な情報が紐付いた大規模医療データベース**
～次世代医療基盤法によって何が可能となったか～」

座長：湯浅　哲也
（山形大学／ JAMIT2020 大会長）

SE　多種多様な情報が紐付いた大規模医療データベース
～次世代医療基盤法によって何が可能となったか～
小松　慶太（内閣官房　健康・医療戦略室　主査　（併）内閣府，経済産業省）

シンポジウム２

12:40 ～ 14:10　**SY ２「OCT の技術的進歩と事業化、そして臨床応用の最前線」**

座長：秋葉　正博（株式会社トプコン）

SY2-1　座長のことば
秋葉　正博（株式会社トプコン）

SY2-2　第二世代 OCT の開発：基礎研究から市場へ
安野　嘉晃（筑波大学）

SY2-3　ダーウィンの荒波にもまれて…眼科用 OCT 事業化と更なる成長
福間　康文（株式会社トプコン）

SY2-4　眼科臨床における OCT の活用事例
中澤　　徹（東北大学）

SY2-5　AI 眼科診断・診療への応用に向けて
坂下　祐輔（株式会社ニデック）

SY2-6　血管内 OCT 技術と臨床での使われ方
平　　健二（ボストン・サイエンティフィックジャパン株式会社）

一般演題セッション5

14:20 ～ 15:50　**P5「CAS・生体情報・その他」**

座長：湯浅　哲也（山形大学）

P5-1　OCT を用いたマウス生殖機能評価に関する検討
渡部　裕輝（山形大学大学院理工学研究科）

P5-2　パルスレーザーダイオードを用いた光音響イメージングシステムの開発
甲田　梨緒（山形大学大学院理工学研究科）

P5-3　マウス脳 ex vivo イメージングのためのマルチピンホール型蛍光 X 線 CT における入射強度補正
保科　亮太（山形大学　大学院理工学研究科）

P5-4　細胞内構造の深層学習とレディオミクスに基づく癌細胞のラベルフリー分類
大手　希望（浜松ホトニクス株式会社）

P5-5　手術ロボット開発に向けた犬の避妊手術シミュレーション・データセット作成
早川　吉彦（北見工業大学工学部地域未来デザイン工学科機械知能・生体工学コース）

P5-6　Admissible SDF variation を用いたヒト胚子の脳の平均形状の推定
田邊　宥大（東京農工大学）

P5-7　内視鏡と連動した食道手術ナビゲーションシステムの開発
佐藤由季也（千葉大学大学院融合理工学府）

P5-8　多層構造物測定のための深さ超解像 OCT
椴山　　誉（株式会社トプコン）

P5-9　脳波信号への深層学習適用による標的刺激の推定
長島　和馬（山形大学大学院　理工学研究科）

P5-10　患者情報の併用による肺細胞診画像の良悪性鑑別 -4 種類の分類アルゴリズムによる性能比較 -
山田あゆみ（藤田医科大学　医療科学部）

P5-11　乳がん手術支援のための乳房追跡による MRI 画像重畳の検討
小口　貴也（千葉大学融合理工学府基幹工学専攻医工学コース）

P5-12　マルチタッチ LCD により直感的操作が可能な医用画像ディスプレイシステムの開発
増谷　佳孝（広島市立大学 大学院情報科学研究科）

P5-13　Image analysis for automatic detection of social interaction in drosophila: biased courtship behavior
渡邉　裕貴（山形大学大学院・理工学研究科・応用生命システム工学専攻）

P5-14　大腸内視鏡のための教師なし深度画像推定法における補助タスク検討
伊東　隼人（名古屋大学 大学院情報学研究科）

P5-15　位置特徴量とマルチスケール特徴量による胃壁マイクロ CT 像からの胃壁の層構造及び腫瘍の抽出
御手洗　翠（名古屋大学大学院情報学研究科）

P5-16　屈折コントラストＸ線ＣＴを用いた非浸潤性乳管癌と通常型乳管過形成の腺腔における３次元的な差異
谷口　華奈（名古屋大学大学院 医学系研究科）

P5-17　Visual SLAM を用いた内視鏡位置推定に関する基礎検討
上野　創史（大阪大学大学院情報科学研究科）

P5-18　大規模放射線読影レポートデータベースによる BERT モデルの事前学習とそれを用いた CT 画像の撮影目的の推定
本田　修平（奈良先端科学技術大学院大学）

P5-19　コンピュータシミュレーションで得られた最適ポータル位置のファントムによる検証
川上申之介（近畿大学工業高等専門学校）

P5-20　CT 画像から腹部臓器領域の自動抽出のための大規模データベース構築に向けた不確実性の推定に基づくアクティブラーニングの検証
政木　勇人（PSP 株式会社）

P5-21　空間情報を考慮した自己教師あり学習を用いたステレオ腹腔鏡画像からの深度推定
李　　文達（名古屋大学 大学院情報学研究科）

P5-22　CNN を用いた指の運動想起脳波の識別
加藤　　将（山形大学大学院）

P5-23　腹腔鏡手術動画像データベース構築に向けたリモートアノテーションツールのプロトタイプ開発
屠　　芸豪（名古屋大学 大学院情報学研究科）

P5-24　肺がん病理画像におけるベイズ U-Net に基づく不確実性を用いた細胞核自動セグメンテーション精度予測
スーフィーマーゼン（奈良先端科学技術大学院大学）

第３日　９月19日（土）

ハンズオンセミナー５

14:20 ～ 15:50　**HS5「ディープラーニングの標準的な手順を身につけよう」**

HS5　異常検知：Auto Encoder を利用した教師なし学習入門

15:55 ～ 16:00　**閉会式**

抄　録

(著者から提出された内容をそのまま掲載します)

第1日　9月17日(木)

第10回 JAMIT チュートリアル講演会

13:10 ～ 16:00　**TL「これから取り組む深層学習と研究の目標」**

座長：中田　典生（東京慈恵会医科大学）

座長：小田　昌宏（名古屋大学）

深層学習を利用した研究が一般的になってきた．JAMIT 大会でハンズオンを初めて行った 2017 年においては，Pubmed で検索される論文のうち，論文名に Deep learning を含む論文は 338 編であったが（企画中の 2016 年は 116 編），2019 年には 1777 編になった．2020 年 7 月現在ではすでに 1500 編あまりとなり，様々な問題を解決する手段として深層学習が活用される状況が反映されている．このような状況のもと，今回の JAMIT チュートリアルにおいても，大会期間中に会場で開催される「コニカミノルタ科学技術振興財団　JAMIT ハンズオンセミナー」と一部連携して，医用画像における深層学習の実際について集中的に講演を企画した．講演 1 では，これから深層学習の研究や勉強を改めて始めようとする学生や初学者に向けて，基礎的な内容を俯瞰し，様々な研究事例を紹介していただく．講演 2 では，その具体的な応用例として，COVID-19 の診断に関して AI 技術を応用した診断支援システムを開発した企業の開発者から，開発の経緯などを講演いただく．講演 3 では，医師が取り組む AI 研究として，臨床的な着眼点を元にした需要の掘り起こしや研究計画，そしてその実際について講演いただく．

TL1　基礎編：ディープラーニングの基礎～これから取り組む方へ～

○李　　鎔範

新潟大学

ディープラーニングモデルを構成する基本的な要素技術について概説するとともに，畳み込み，プーリング，事前学習などの用語について掻い摘んで解説する．また，ディープラーニングによってどのようなことができるのかに関連して，検出，領域分割，分類，推定，画質改善，画像生成などについて言及する．さらに，応用例として，医用画像分野におけるディープラーニングの研究事例をいくつか紹介る．最後に，ハンズオンセミナーの内容について簡単に紹介する．

TL2　企業における技術展開：良い医療 AI への道

○周　　暁妍[1]，王　　禕楠[2]

[1] 株式会社 Infervision Japan 代表取締役，[2] 株式会社 Infervision Japan プロダクトマネージャー

全てが始まる前に，自分に問いかけてみてください．医療業界は果たして，AI は必要なのか？医師の数が世界的に不足で，患者の数が増える一方である．更に，日本は少子高齢化などの問題を抱えている．少ない医師がより多くの患者の画像を扱うには，AI は良い選択に聞こえる．しかし，AI の力は限界がある．

良い医療 AI の三つの要素：

1. 特定の現実の医療問題に対処する
2. 設計から製品化までの流れは，医療機器の品質管理プロセスを準拠に行わなければならない
3. 長期的な計画性と協調性が必要

当社の製品，仕事環境，開発経緯，医療製品の開発プロセス（データ，チームワーク，フィードバック，承認申請）についても概説する．

TL3　医師が取り組む AI 研究：腹部画像診断・IVR 診療における AI の潜在的臨床応用可能性

○山田　　哲

信州大学

日々生み出され続ける膨大な情報と需要に対応しなくてはならない現代医療において，AI を用いた診療の質と効率の向上は大いに期待されている分野である．しかしながら真に役立つ AI システムを実現するためには，技術的な実現可能性の検討のみならず，臨床的な潜在的需要についての洞察が必要不可欠である．本講演では腹部放射線診断専門医および IVR 専門医として日常診療に携わる演者から見た，潜在的な AI の臨床応用可能性の高い放射線科診療領域について提案を行い，本領域への AI 応用推進の一助となれば幸いである．

一般演題セッション1

16:15 ～ 17:45　**P1「CAD」**

座長：湯浅　哲也（山形大学）

P1-1　単純腹部CT画像からの高精細な大腸領域検出法の検討

○安倍　和弥[1]，武尾　英哉[1]，永井　優一[2]，縄野　繁[3]
[1]神奈川工科大学，[2]国立がん研究センター東病院，[3]国際医療福祉大学三田病院

現在，医療の様々な分野においてコンピュータ画像診断支援（CAD：Computer Aided Diagnosis）の研究が行われている．大腸は周辺臓器との分離の難しさから単純CTを用いてのCAD開発はあまり行われてこなかった．最近では，内視鏡検査の高精度化やCTコロノグラフィー検査の普及などから大腸がんの発見などにCADの研究が行われている．しかし，内視鏡検査やCTコロノグラフィー検査は事前準備が必要であり，特にCTコロノグラフィーにおいては炭酸ガスの注入による大腸の拡張を行う点から患者に大きな負担がかかるため検査自体が行いにくいなどの問題もある．

そこで本研究では，腹部検診時に撮影する単純腹部CT画像から高精細に大腸領域を検出する手法について検討した．これにより，大腸に負担を与えるCTコロノグラフィー検査などを行わずにCAD開発に用いる画像も取得できる．また，検診画像からのがんの経時変化の評価などへの応用も可能となる．

P1-2　特徴量選択手法およびシステム構成の最適化によるうつ病CADの高性能化

○牧　　優太[1]，和田　昇太[1]，安倍　和弥[1]，武尾　英哉[1]，長井　優一[2]
[1]神奈川工科大学，[2]国立がん研究センター東病院

近年の精神病の診断は，医師の勤続年数等によって診断結果が左右されてしまうなど，科学的，客観的な評価による診断ができていないという問題があった．そのため，うつ病の客観的な評価を行うために，画像および音声工学技術を利用してうつ病である確率を算出し，うつ病であるかどうかの判定を行ううつ病CADの開発を行った．

従来のうつ病CADは視線方向と顔表情に着目し，それぞれの時系列データからうつ病判別を行う判別器を2つ作り，それらを組み合わせたものであった．そしてそれは約83%の正解率で判別することができた．

本研究はより判別精度を向上させるため，まずは音声信号を用いた判別器を追加した．次に，特徴量を用いている音声信号と視線方向の判別器に特徴量選択や次元削減を行った．最後に，これら判別器の組み合わせ手法であるアンサンブル学習の，Stacking構成の最適化をした．結果として，さらなる精度の向上が図られた．

P1-3　画像と音声情報を用いたうつ病CADの高性能化

○和田　昇太[1]，牧　　優太[1]，安倍　和弥[1]，武尾　英哉[1]，永井　優一[2]
[1]神奈川工科大学，[2]国立がん研究センター東病院

近年のうつ病などの精神疾患の診断は，医師の診断経験などに診断結果が左右され，科学的，客観的な評価による診断が行えていないという課題があった．そこで，画像および音声工学技術を利用してうつ病である確率を算出し，うつ病であるかどうかの判別を行う，うつ病CADを開発した．従来のうつ病CADは，顔表情や視線方向など患者の顔画像から得られる情報に着目して判別を行うものであり，それらの組合せによる判別精度は83%を実現した．本研究は，更なる精度向上のため，患者の音声情報に着目し判別器を作成した．うつ病患者の発声は健常者と比較して抑揚に乏しい，発話速度が遅いなどといった違いがみられることに注目し，それらを表す特徴量を作成しSVM判別器を作成した．その結果，音声判別モデル単体では75%の判別精度であったが，従来の顔表情と視線方向に基づく判別器と組合せてアンサンブル学習を行ったところ，88%の判別精度を得ることができ，システム全体としての判別精度を向上させることができた．さらに，上記の単純な波形信号からの特徴量に判別器に加え，発声した言葉の内容やキーワードによる判別器も開発し，これも含めて組み合わせた総合的な判別システムの検討も行っている．

P1-4　敵対的生成ネットワークを用いて生成した硬性白斑画像についての検討

○藤田　真穂[1]，畑中　裕司[2]，砂山　　渡[2]，村松千左子[3]，藤田　広志[4]

[1]滋賀県立大学大学院工学研究科電子システム工学専攻，[2]滋賀県立大学電子システム工学科，[3]滋賀大学データサイエンス学部，[4]岐阜大学工学部電気電子・情報工学科

糖尿病網膜症は中途失明の原因となる病であるが，早期発見と治療によって失明を防ぐことができる病でもある．近年は，眼底画像からの初期病変の検出に畳み込みニューラルネットワーク (CNN:Convolutional Neural Network) が成果をあげている．一方で，CNN を用いる際には病変データ数と非病変データ数の不均衡が課題となる．現在は，敵対的生成ネットワーク (GAN:Generative Adversarial Networks) を用いて病変データを生成し，上記の問題の解決を図る研究が行われている．本研究では，GAN を用いて初期病変の一つである硬性白斑の画像を生成した．また，生成データと実データ間の分布の差異について，複数の指標を用いて検証を行った．当日は生成したデータに関する検証結果と，生成したデータが CNN に与える影響について報告する予定である．

P1-5　Modified Cycle-Consistent Adversarial Network を用いた病理組織画像における染色度合いの正規化を伴う病理組織分類

○林　　大誠[1]，檜作　彰良[1]，中山　良平[1]，児玉　良典[2,3]，眞能　正幸[3]，吉岡　絵麻[4]，兼松　大介[4]，正札　智子[4]，金村　米博[4]

[1]立命館大学理工学部 電子情報工学科，[2]神戸大学大学院 医学研究科，[3]国立病院機構大阪医療センター 臨床検査科，[4]国立病院機構大阪医療センター 臨床研究センター

本研究の目的は，CycleGAN に新たな CNN を追加した Modified CycleGAN を構築することにより，病理組織画像の染色度合いを正規化し，病理組織の正常／異常を分類する手法を開発することである．実験試料は，HE 染色されたグリオーマ病理組織標本 53 例で，拡大率 20 倍でデジタイズした画像を用いた．これらの画像の異常組織から 21,677，正常組織から 7,892 の関心領域（ROI）を抽出した．また，染色度合いの基準となる病理組織画像を 1 例選択し，その画像から抽出した ROI を基準画像として用いた．Modified CycleGAN の学習は，入力した ROI を基準の染色度合いに近づける画像変換の学習と，ROI の正常／異常分類の学習を並行で行う．3 分割交差検証による分類の結果, 正答率 87.2%, 感度 91.9%, 特異度 73.0%, AUC 0.919 が得られ，提案手法の有用性が示唆された．

P1-6　SVM と時空間モデルを用いた嚥下時 X 線動画からの喉頭蓋の抽出に関する基礎的検討

○杉田　絢美[1]，藤中　彩乃[1,2]，目片幸二郎[3]，滝沢　穂高[1]，工藤　博幸[1]

[1]筑波大学，[2]日本学術振興会特別研究員，[3]神戸赤十字病院

嚥下障害のうち誤嚥は，喉頭蓋という軟骨組織が大きく関係している．喉頭蓋は気道の入り口にあり，気道に蓋をすることで嚥下物が気道に侵入することを防ぐ役割を持つが，その構造および動体解析は進んでいない．本研究では，SVM および喉頭蓋の時空間モデルを用いて嚥下時 X 線動画からの喉頭蓋を抽出する手法を提案する．まずピクセルベースの SVM を適用し，喉頭蓋の候補領域の抽出を行う．しかし，喉頭蓋が倒れて気道に蓋をするという動きは 1 秒以内の速い動きで，さらに造影剤が喉頭蓋に接触しているフレームでは喉頭蓋が確認できない．そこで，時空間モデルを作成し，SVM の識別結果を基に喉頭蓋の位置推定を行う．実際の X 線動画に本手法を適用した結果を示す．

P1-7　U-Net を用いた X 線動画からの頸椎椎間板の抽出に関する基礎的検討

○郡司絵莉華[1]，藤中　彩乃[1,2]，目片幸二郎[3]，滝沢　穂高[1]，工藤　博幸[1]

[1]筑波大学，[2]日本学術振興会特別研究員，[3]神戸赤十字病院

栄養低下や誤嚥を引き起こすことがある嚥下障害の詳細なメカニズムは未だに解明されていない．メカニズムを解明するために，頸部構造物の動態解析をし，嚥下障害との関係を調査する必要がある．その手がかりとしてまず椎間板の抽出を行う．本研究では二段階に分けて椎間板の抽出を行う．X 線動画の各フレームに，モルフォロジー演算と RANSAC による放物線あてはめを適用し，頸部領域の切り出しを行う．切り出した画像に対して，医用画像のセグメンテーションを目的に開発された U-Net を適用し，頸椎椎間板を抽出する手法を提案する．深層学習は一般的に学習データの数が多いほどより良いモデルを得られることが知られている．抽出部位の正解画像を作成する場合膨大な時間やコストが掛かる．そこで，少数のデータセットに対しても一定の精度を得られるように，アフィン変換を用いて回転・平行移動によるデータ拡張を行う．実際の X 線動画に適用した実験結果を示す．

P1-8 ２種類の３次元畳み込みニューラルネットワークを用いた造影後 T1 強調画像における神経膠腫の全自動グレード分類パイプライン

○山城　滉斗[1]，寺本　篤司[1]，齋藤　邦明[1]，藤田　広志[2]

[1] 藤田医科大学大学院保健学研究科，[2] 岐阜大学工学部

神経膠腫のグレードは治療方針や予後に大きな影響を与えるため，術前には MRI 検査によるグレード分類が行われるがその精度は高くない．そのため様々な画像認識技術を用いて精度向上を目指した研究が行われてきたが，手動で抽出した領域における二次元での解析に留まっている．またディープラーニングを医用画像に対して適用する場合，データの入手の制約等によりその適用が妨げられる．その解決のために NVIDIA 社は clara プロジェクトで様々な学習済みモデルを提供している．よって本研究では clara プロジェクトの３次元畳み込みニューラルネットワーク (3D CNN) を用いて脳腫瘍領域を抽出し，その領域に対して独自の 3D CNN でグレード II・III とグレード IV を全自動的に分類した．本手法に対して 46 個の腫瘍を用いて検証を行ったところ，総合正解率は 87.0％となり，良好な結果が得られた．

P1-9 Mask R-CNN とテクスチャ解析による小児肺炎の検出

○魚住　春日[1]，寺本　篤司[1]，日木あゆみ[2]，本元　　強[2]，河野　達夫[3]，齋藤　邦明[1]，藤田　広志[4]

[1] 藤田医科大学大学院 保健学研究科，[2] 茨城県立こども病院，[3] 東京都立小児総合医療センター 診療放射線科，[4] 岐阜大学 工学部

小児に高い感染のリスクがある肺炎の検査には胸部 X 線画像が用いられるが，医師ごとの診断にばらつきがあるため，小児肺炎の自動検出を試みた．小児胸部 X 線画像は，成長やポジショニングによって状態が変化しやすいため，画像全体を直接解析することは困難である．そこで本手法では，解析領域である肺野領域の抽出と，肺野内のテクスチャ解析の２段階による処理によって肺炎像の検出を行った．

肺野の抽出には物体検出と領域抽出処理が包含された Mask R-CNN を用いた．その後，肺野を 32 × 32pixels の小区域に分割し，テクスチャ特徴量を抽出した．識別器にてテクスチャ特徴量をもとに肺炎と正常に分類し，実画像上に肺炎の確率分布をマッピングした．

NIH が公開する胸部 X 線画像データベース ChestX-ray14 の画像を用いて学習を行い，東京都立小児総合医療センターで撮影された画像にて評価を行った．その結果，正常症例，肺炎症例共に約 70% の正解率となった．

P1-10 V-Net を用いた Ai-CT に対する肝臓領域抽出手法の開発と肝臓の質量推定への応用

○櫛部　光哉[1]，平野　　靖[1]，稲井　邦博[2]

[1] 山口大学大学院創成科学研究科，[2] 福井大学医学部

死亡時画像診断 (Autopsy imaging：Ai) とは，死後 CT 像を読影することによって，死因究明や死後経過時間の推定を行う手法である．臓器の質量は疾患や病態を反映するため，死因究明を行う上で有用な情報となる．一般的に，臓器の質量を測定するためには解剖を必要とする．しかし，日本では剖検率が極めて低いため正確な質量を得ることができない場合が多い．そこで本研究では，死後 CT 像における肝臓領域内の CT 値を用いて，解剖を行わずに肝臓の質量を推定する手法を開発した．CT 値から肝臓の質量を推定するためには，肝臓領域を抽出する必要がある．CT 像から肝臓領域を手作業で抽出するためには，多大な労力を要する．さらに得られた肝臓領域の抽出精度は安定しない．そこで本研究では，まず V-Net を用いて肝臓領域の抽出手法の開発を行った．24 症例に対して提案手法を適用した結果，正解肝臓領域と抽出結果との Dice 係数は 0.842 ± 0.069 であった．次に，抽出結果と CT 像を用いて肝臓領域内の CT 値から質量推定を行った．その結果，実際の質量と推定した質量の相関係数は 0.774 であった．

P1-11 心臓超音波動画像における畳み込みニューラルネットワークを用いた心筋梗塞の自動分類手法の開発

○村木　亮介[1]，木村　竜誠[1]，寺本　篤司[1]，山崎　歌音[1]，杉本　恵子[1]，杉本　邦彦[2]，山田　　晶[3]，渡邉　英一[3]，齋藤　邦明[1]，藤田　広志[4]

[1] 藤田医科大学大学院保健学研究科，[2] 藤田医科大学病院，[3] 藤田医科大学循環器内科，[4] 岐阜大学工学部

現在，心筋梗塞は生活習慣病が原因となり，突然死や慢性的な心不全を引き起こす可能性があり，早期発見が重要な疾患である．早期の検査法として超音波検査が挙げられる．しかし，心筋梗塞による異常壁運動は範囲や重症度により判別が困難であり，致死的な合併症につながるため高い正確性が必要とされる．そこで本研究では，心臓超音波動画像における畳み込みニューラルネットワーク（CNN）と従来の機械学習法を利用した心筋梗塞の自動分類手法の開発を目的とした．短軸像，左室長軸像の１心拍分にあたる動画像を入力し，CNN により特徴量を抽出した．特徴量より SVM，Neural Network，AdaBoost，Random Forest，Naïve Bayes の機械学習アルゴリズムを使用し，正常症例と心筋梗塞症例の分類を行った．結果は，１心拍分の動画像を使用したときの正解率は左室長軸像で 81.8％，短軸像で 82.8％となった．以上の結果より，本手法は正常心筋と心筋梗塞症例の分類に有効であることが示唆された．

P1-12　体幹部 CT 画像における U-Net アーキテクチャによる体腔の自動認識

○前田　泰良[1]，神谷　直希[1]，周　　向栄[2]，加藤　博基[3]，原　　武史[2]，三好　利治[4]，松尾　政之[3]，藤田　広志[2]

[1] 愛知県立大学 情報科学部，[2] 岐阜大学 工学部 電気電子・情報工学科，[3] 岐阜大学大学院 医学系研究科 放射線医学分野，[4] 岐阜大学医学部附属病院 放射線部

The body cavity is covered with bones and skeletal muscle and contains many organs. Therefore, the body cavity region is an important information in computer-aided diagnosis (CAD) systems. In particular, we are working on automatic recognition of skeletal muscles, so the body cavity, which is the boundary between muscles and other organ tissues, is an important information. In this study, we propose an automatic segmentation method of the body cavity using the U-Net architectures in torso CT images. We verify the segmentation accuracy by the hold-out method using 2D U-Net, 2D U-Net++ and 3D U-Net. Then, using the network with the best segmentation accuracy, we performed a 4-fold cross-validation to verify the validity of the result. Dice values were used to compare the results of segmentation with the ground truth of the torso CT images in 20 cases. The segmentation accuracies of 2D U-Net, 2D U-Net++ and 3D U-Net were 98.5%, 92.5% and 94.0%, respectively. Therefore, we performed a 4-fold cross-validation using 2D U-Net, and found that the Dice, Precision and recall values were 98.5%, 98.9%, and 98.3% on average, respectively. The segmentation accuracy in each vertebral section was evaluated, and the results showed that the thoracic, lumbar, and sacral vertebrae and caudal sections were 99.3%, 98.3% and 96.1%, respectively. In the proposed method, the segmentation of the body cavity in torso CT image was realized with high accuracy of average 98.5% using 2D U-Net. In the future, it is necessary to evaluate the automatic segmentation of the body cavity as the boundary of the muscle with the atrophic muscular disease.

P1-13　機械学習に基づく食道・頭頸部癌早期発見の為の診断支援システムに関する検討

○吉村裕一郎[1]，上里　昌也[2]，浦濱　竜馬[2]，加藤直菜美[2]，米本　昇平[2]，中口　俊哉[1]

[1] 千葉大学 フロンティア医工学センター，[2] 千葉大学医学部　先端応用外科学

本研究は食道内視鏡検査における，多発ヨード不染帯（まだら不染）を深層学習によって自動的に識別し，食道内や頭頸部における異時性・同時性癌の併発リスクを予測する手法について検討する．高リスク患者を，内視鏡検査時に選別するとともに，経過観察によって見逃しによる治療機械の損失を防ぐことを目的とする．

食道内視鏡検査において，ルゴール染色によってまだら不染を観察することが行われている．まだら不染帯は，前述のリスクと関連していることが指摘されているが，医師の主観に基づく観察では，併発リスクの予測精度は 39.5% である．そこで本研究では，内視鏡検査映像中から，高リスクシーンを自動的に抽出し，シーンを機械学習によって解析することで，併発リスクを予測する手法を提案する．

実験では，専門医が選択した画像を用いたモデル，臨床的知見による医師の精度と提案手法による比較を実施し，提案手法が最も高性能であることを確認した．

P1-14　ディープラーニングに基づく 3 次元汎用半自動病変セグメンテーションの初期検討

○野村　行弘[1]，花岡　昇平[2]，竹永　智美[1]，中尾　貴祐[1]，柴田　寿一[1]，三木聡一郎[1]，吉川　健啓[1]，渡谷　岳行[2]，林　　直人[1]，阿部　　修[2]

[1] 東京大学医学部附属病院コンピュータ画像診断学／予防医学講座，[2] 東京大学医学部附属病院放射線科

In the development of computer-aided detection software, a three-dimensional (3D) voxel labeling of lesion requires a significant effort for radiologists. To reduce the time required for 3D voxel labeling, we aimed to develop a generalized semiautomatic segmentation method based on deep learning. In the preliminary study, we investigated whether a single type of lesion can train a semiautomatic segmentation model for two types of lesions with different contrasts from the surroundings. We targeted lung nodules in chest CT images and liver lesions of hepatobiliary-phase images in Gd-EOB-DTPA-enhanced MR imaging. We input the 32 × 32 × 32 cubic volume of interest around the center of gravity of the lesion and the spherical mask determined by the maximum diameter of the lesion into a 3D FC-ResNet model to define the label of the lesion. We trained the model for each type of lesion with data augmentation, including random shifting and random monochrome inversion. The evaluation results showed that the Dice coefficients using the model trained on lung nodules were 0.681 ± 0.131 for the lung nodules, 0.685 ± 0.121 for the liver lesions. The Dice coefficients using the model trained on liver lesions were 0.471 ± 0.152 for the lung nodules, 0.757 ± 0.140 for the liver lesions. The model trained on lung nodules was able to label liver lesions as well as those of lung nodules well.

P1-15 大腸内視鏡における穿孔の自動検出および位置推定に関する予備的検討

○蒋　　凱[1]，伊東　隼人[1]，小田　昌宏[1]，奥村　大志[2]，森　　悠一[2]，三澤　将史[2]，林　　武雅[2]，工藤　進英[2]，森　　健策[1,3,4]

[1]名古屋大学 大学院情報学研究科，[2]昭和大学 横浜市北部病院消化器センター，[3]名古屋大学 情報基盤センター，[4]国立情報学研究所 医療ビッグデータ研究センター

Endoscopic submucosal dissection (ESD) is a minimally invasive treatment for early gastric cancer. In ESD, a physician directly removes the mucosa around the lesion under internal endoscopy by using the flush knife. However, the flush knife may accidentally pierce the colonic wall and generate a perforation on it. If doctors overlook a small perforation, patient may need emergency open surgery since perforation can easily cause peritonitis. Therefore, there is a potential demand for computer-aided diagnosis (CAD) system, which prevents an overlook of a perforation by detecting and localizing the perforation. In order to build a CAD system for ESD, we developed a model of perforation classification and localization for colonoscopy video by using YOLO-v3. We annotated 2,000 and 200 perforation images for training and test, respectively, from colonoscopy movies of totally 11 patients. Our model achieved 83% accuracy and 69.5% mean average precision for perforation classification and localization, respectively, of the test images.

P1-16 機械学習によるコルポスコープ画像における白色上皮領域検出の基礎的検討

○篠原　寿広[1]，天野　裕太[1]，村上　幸祐[2]，中迫　　昇[1]，松村　謙臣[2]

[1]近畿大学生物理工学部生命情報工学科，[2]近畿大学医学部産科婦人科学教室

子宮頸癌の診断において，コルポスコープとよばれる拡大鏡を用いて，目視による検査が行われている．本検査はコルポスコピーとよばれ，子宮頸癌の確定診断を行う組織診のための組織を採取する部位を決定する非常に重要な検査である．しかしながら，コルポスコピーの精度は，検査者の知識と経験に大きく依存し，とくに経験の浅い医師には正確な診断が難しい．そこで本研究では，機械学習を用いて，コルポスコピー支援を行い，安定して精度の高い診断の実現を目指す．本稿では，初期検討として，コルポスコープ画像から，子宮頸部の代表的な病変の一つである白色上皮の判別を行った．具体的には，コルポスコープ画像を小領域に分割し，CNN(Convolutional Neural Network) を用いて白色上皮かどうかを分類した．実際の 10 症例のコルポスコープ画像を使用し，交差検証により判別精度を検討した．

P1-17 One-class SVM による画素悪性度の集中性を用いた FDG-PET/CT 上の病変自動検出

○田中　敦子[1]，根本　充貴[2]，甲斐田勇人[3,4]，木村　裕一[1]，永岡　　隆[1]，牛房　和之[1]，山田　誉大[4]，花岡　宏平[4]，北島　一宏[5]，槌谷　達也[6]，石井　一成[3,4]

[1]近畿大学大学院 生物理工学研究科，[2]近畿大学 生物理工学部，[3]近畿大学 医学部 放射線医学教室 放射線診断学部門，[4]近畿大学 高度先端総合医療センター PET 分子イメージング部，[5]兵庫医科大学 放射線医学講座，[6]兵庫医科大学病院 放射線技術部

【目的】FDG-PET/CT 上の頚胸部領域における原発性・転移性病変の自動検出に関する初期検討として，病変候補点の自動検出法を提案し，検証する．【手法】提案法では，まず CT の閾値処理及びモロフォロジ処理によって左右肺を抽出し，抽出した肺の間を縦隔部として自動抽出する．頚部は，自動同定した下顎骨座標と首領域，左右肺領域を基準に抽出する．その後，正常な FDG-PET/CT で学習済の one-class SVM を用いて，対象領域内画素の異常度を計測する．得られた異常度画像の極大点を検出し，病変の候補点とする．提案法は，82 病変を含む FDG-PET/CT 有病 47 例で検証する．【結果】頚胸部の病変検出感度が 95.1% (79/82) となった．【結論】検出感度結果より提案法が有用であることが示唆された．今後の課題は，誤検出候補点の削減処理の追加と，異常検知特徴量の追加による性能向上である．

P1-18 教師なし学習を利用した異常検知の PET 画像における腫瘍検出への応用

○前田　健宏[1]，原　　武史[2]，片淵　哲朗[3]，周　　向栄[2]，松迫　正樹[4]，伊藤　　哲[5]，加藤　正也[5]，山田　　哲[6]，藤田　広志[2]

[1]岐阜大学大学院自然科学技術研究科知能理工学専攻，[2]岐阜大学工学部電気電子・情報工学科，[3]岐阜医療科学大学保健科学部，[4]聖路加国際病院放射線科，[5]大雄会病院第一放射線科，[6]信州大学医学部放射線科

Numerous studies have been reported on the detection of tumors in medical images such as Positron emission tomography (PET) and computed tomography (CT) using deep learning. Many of them used supervised deep learning on large data sets with label images for correct detection. However, the data set for training may be insufficient to train deep learning models due to the huge cost of creating label images. This study aims are utilization of PET images without label and development of a method for the detection of tumors using unsupervised learning. We propose a method of tumors detection based on anomaly detection with a deep learning model trained normal cases. A generation model based on GANs and AutoEncoder is used for detection. The model trained only normal PET images generates nearly normal images which are similar to training data. Therefore, the difference between images containing untrained anomalous accumulations and generated images is large. The abnormal regions are extracted by a dynamic thresholding on the differential images between inputs and outputs of the model for each cross-sections (axial, coronal and sagittal). Only areas where the results of all cross-sections match are candidates of tumors. Lastly, false positives are eliminated by using gradient of voxel values around extracted area. The detection performance was evaluated using 93 PET/CT images included 38 normal cases for a training dataset and 55 anomaly cases for a test dataset. The sensitivity was 76.5% with 95.4 FPs per case. The area of FPs was 938.3 voxels per case. These results indicate that unsupervised learning model trained normal cases without label contributed to automatic detection of tumors.

P1-19　読影レポート解析を利用した医用画像データベースからのアノテーション付きデータセット作成に関する初期検討

○林　雄一郎[1]，鈴村　悠輝[1]，岡崎　真治[1]，小田　昌宏[1]，森　　健策[1,2]

[1]名古屋大学大学院情報学研究科，[2]国立情報学研究所医療ビッグデータ研究センター

本稿では，読影レポート解析を利用した医用画像データベースからのアノテーション付きデータセット作成に関する初期検討について述べる．深層学習などの機械学習を用いた医用画像診断支援システムの開発では，アノテーション付きのデータセットを準備する必要がある．しかし，大量の医用画像に対するアノテーション作業は非常に労力を要する．そこで，画像データに付随する読影レポートを解析することでアノテーション付きデータセットを構築する手法を検討する．提案手法では，読影レポートから特定の疾患名を含むデータを取得する．次に読影レポートの 1 文中に特定の疾患名と否定表現が含まれるデータを除外することで特定の疾患を含むデータを取得する．国立情報学研究所の医療画像ビッグデータクラウド基盤のデータに提案手法を適用し，間質性肺炎のアノテーション付きデータセットの作成が可能であることを確認した．

P1-20　極度に不均衡な頭部 CT データにおける転移学習を用いたクモ膜下出血自動検出の検討

○魯　　仲陽[1]，小田　昌宏[1]，林　雄一郎[1]，胡　　涛[1]，伊東　隼人[1]，渡谷　岳行[2]，阿部　　修[2]，
橋本　正弘[3]，陣崎　雅弘[3]，森　　健策[1,4,5]

[1]名古屋大学大学院情報学研究科，[2]東京大学医学部附属病院放射線科，[3]慶応義塾大学医学部放射線科学教室，[4]名古屋大学情報基盤センター，[5]国立情報学研究所 医療ビッグデータ研究センター

Due to the blooming Artificial Intelligent technology, the image recognition field has achieved an unprecedently impressive breakthrough. Nevertheless, as the fuel to the engine, performances of AI models are heavily restricted by the quantity and quality of data. Typically, a huge, high-quality, and balanced dataset is essential, whereas considering the actual situation of emergency care, acquiring mess lesion cases is quite a difficult job. In this study, we expose a deep learning based Subarachnoid Hemorrhage (SAH) detection system from extremely imbalanced head computed tomography (CT) scans by using transfer learning technology. Our dataset contains 2,516 cases without Intracranial hemorrhage (ICH) and only 33 cases with SAH. Besides, a huge, high-quality public dataset (containing 3,549 cases SAH data), collected by the Radiological Society of North America (RSNA) organism is applied to improve the performance of our classifier by using transfer learning technology. Furthermore, some training strategies for improving classification performance, just like weighted middle slices (ICH usually happened) are utilized for this experiment.

According to the results of the experiment, our SAH detection system makes a significant improvement via transfer learning.

P1-21　顎骨骨髄炎発症範囲推定のための頭部 CT・切除顎骨 CT の位置合わせ自動化

○松岡　志宜[1]，盛田　健人[1]，武田　大介[2]，長谷川巧実[2]，若林　哲史[1]

[1]三重大学大学院工学研究科情報工学専攻，[2]神戸大学附属病院歯科口腔外科

近年，顎骨骨髄炎の治療法として顎骨切除が用いられているが，術前画像中から判別できる骨髄炎範囲が曖昧であるため手術計画の立案が容易ではなく，手術の長時間化等の問題が発生している．そこで，切除範囲の高精度な自動推定を可能とすることで，医師と患者の負担を軽減を行うことが求められている．そのため，術前頭部 CT 画像と切除顎骨 CT 画像と詳細な炎症範囲を診断できる病理画像を対応付ける必要があり，本研究では術前頭部 CT 画像と切除顎骨 CT 画像の位置合わせ手法を提案した．CT 値のみを用いた尤度関数では正確な位置合わせが困難であったため，本研究では 3 点のキーポイントを併用した尤度関数を用いる剛体位置合わせ行った．被験者 3 名のデータセットを用いた実験から，正規化相互相関の平均が 0.722 であり，目視評価からも位置合わせが正確にできていることを確認した．

P1-22　深層学習を用いた CT 画像解析による肺腺癌手術患者における術後再発予測

○佐々木雄樹[1,2]，李　　鎔範[2]，青木　　正[3]，関　　裕史[4]，古泉　直也[4]，尾崎　利郎[4]

[1]新潟県立がんセンター新潟病院　中央放射線部，[2]新潟大学大学院　保健学研究科　放射線技術科学分野，[3]新潟県立がんセンター新潟病院　呼吸器外科，[4]新潟県立がんセンター新潟病院　放射線診断科

目的：肺腺癌手術患者の術前 CT 画像を，深層学習を用いて解析することにより，術後再発の有無を予測する CAD を構築すること．

方法：肺腺癌摘出手術を行った 120 人（再発 60 人，無再発 60 人）の術前 CT 画像を使用した．CT 画像は腫瘍の充実成分径が最も大きいスライスを 1 枚選択し，腫瘍が入るように 72 × 72pixel（1pixel=0.68mm）のサイズに切り出し，入力画像とした．CNN には学習済み VGG19 モデルを使用した．性能評価法には 10-fold-cross-validation を採用した．

結果：提案法の再発評価における感度は 0.72（43/17），特異度は 0.70（42/60）であった．AUC 値は 0.76 であった．

結論：肺腺癌手術患者の術後再発の有無を術前 CT 画像から予測できる可能性が示唆された．

P1-23 COVID-19 症例の定量評価のための CT 像からの肺野自動セグメンテーション

○小田　昌宏[1]，林　雄一郎[1]，大竹　義人[2,3]，橋本　正弘[4]，明石　敏昭[5]，森　　健策[1,3]

[1] 名古屋大学大学院情報学研究科，[2] 奈良先端科学技術大学院大学先端科学技術研究科情報科学領域，[3] 国立情報学研究所医療ビッグデータ研究センター，[4] 應義塾大学医学部放射線科（診断），[5] 順天堂大学医学部放射線医学講座

本稿では，新型コロナウイルス感染症（COVID-19）症例の CT 像からの肺野セグメンテーション手法について述べる．2019 年 12 月に発生した COVID-19 は急速に世界中に広まり多数の患者が発生すると共に，日常生活に大きな打撃を与えている．COVID-19 を含むウイルス性肺炎の診断において胸部 CT 像が有用である．医師の診断を支援するために，CT 像からの肺野状態の定量評価手法の確立が急務である．本稿では Fully convolutional network を用いた CT 像からの肺野自動セグメンテーション手法を提案する．COVID-19 を含む疾患による肺野状態の定量評価には肺野内陰影の状態を把握することが重要である．そのため，提案手法ではすりガラス影や浸潤影を含む陰影領域とそれ以外の領域を識別してセグメンテーション可能とした．提案手法を COVID-19 症例の CT 像に適用し，肺野内の陰影領域等を良好にセグメンテーション可能であることを確認した．

P1-24 膀胱鏡検査画像を用いた間質性膀胱炎の分類に関する予備的検討

○初　　韜[1]，小田　昌宏[1]，古田　　昭[2]，山本　徳則[3]，森　　健策[1,4,5]

[1] 名古屋大学 大学院情報学研究科，[2] 東京慈恵会医科大学，[3] 名古屋大学医学系研究科，[4] 名古屋大学 情報基盤センター，[5] 国立情報学研究所 医療ビッグデータ研究センター

Interstitial cystitis (IC) is a chronic condition which nonspecific chronic inflammation of the bladder is observed, but the clear etiology of autoantibodies has not yet been identified. The condition is a part of a spectrum of diseases known as bladder pain syndrome (BPS). In actual clinical practice, bladder biopsy with suspicion of IC may occasionally reveal bladder intraepithelial carcinoma (CIS). On the other hand, surgery was suspected of CIS, but pathological results showed there is no cancer. Treatment is generally difficult and persistent symptoms severely compromise the patients' quality of life. To assist doctors to diagnose the including interstitial cystitis and bladder intraepithelial carcinoma through the cystoscopy images. We proposed a method which using a convolutional neural network and got some depth estimation results. We also used a technique called Gradient-weighted Class Activation Mapping (Grad-CAM) to produce visual explanations for decisions from CNN. By using the neural network, we achieved the AUC score of 0.694 on this classification task.

一般演題セッション2

9:00 ～ 10:30　**P2「医用画像処理」**

座長：湯浅　哲也（山形大学）

P2-1　トポロジカル不変量を用いたレディオミクス特徴量に基づく肺癌予後推定法

○二宮　健太[1]，有村　秀孝[2]

[1] 九州大学大学院　医学研究院保健学部門，[2] 九州大学大学院　医学系学府保健学専攻

We aimed to develop a novel topologically invariant radiomic features (TIF) for prognostic prediction in lung cancer patients. TIFs were calculated using the texture analysis of Betti number maps which represent topologically invariant characteristics on computed tomography images. The feasibility of TIFs were evaluated using statistically significant differences to compare the survival curves of high- and low-risk patients stratified based on the TIFs. Further, a deep learning(DL)-based approach was utilized to compare its feasibility to TIFs in prognostic prediction. The p-values between the survival curves for TIF and DL were 3.4 × 10^ − 5 and 1.2 × 10^ − 1, respectively. The proposed approach showed a promising potential to the prognostic prediction for lung cancer patients.

P2-2　マルチチャンネル化と CNN を用いた嚥下時 X 線動画からの頸椎椎間板の抽出：LeNet, AlexNet, GoogLeNet の性能比較

○藤中　彩乃[1,2]，目片幸二郎[3]，滝沢　穂高[1]，工藤　博幸[1]

[1] 筑波大学，[2] 日本学術振興会特別研究員，[3] 神戸赤十字病院

本研究では，マルチチャンネル化 (MC) と CNN を用いて，嚥下時 X 線動画 (videofluorography，VF) から椎間板を抽出する．濃淡画像である VF の各フレームに 21 種類の画像フィルタをそれぞれ適用し，同数の特徴画像を生成する．3 種類を選択し，カラー画像の RGB チャンネルに保存することでマルチチャンネル化画像 (MC 画像) を生成する．MC 画像にパッチベースの CNN を適用し，画素ベースの F 値で抽出精度を評価する．焼きなまし法を用いて抽出精度が最大となる 3 種類の特徴画像の組合せを求める．本研究では CNN として LeNet, AlexNet, GoogLeNet を用いて性能を比較した．3 つの CNN において, RGB チャンネルに同一の特徴画像を保存した際の最適な特徴画像は原画像であり，F 値は順に 0.562，0.604，0.598 となった．MC の最適な組合せは 3 つの CNN で異なり，F 値は順に 0.593，0.634，0.598 となった．今後より多くの CNN に MC を適用する．

P2-3　取り下げ

P2-4　Automated whole-heart segmentation using U-Net in pediatric cardiac CT.

○吉田　皓文[1,2]，李　　鎔範[2]，吉村　宣彦[3]，倉元　達矢[4]，長谷川　晃[1,5]，金沢　　勉[4]

[1] 新潟医療福祉大学医療技術学部診療放射線学科，[2] 新潟大学大学院保健学研究科放射線技術科学分野，[3] 新潟大学医歯学総合研究科放射線医学分野，[4] 新潟大学医歯学総合病院診療支援部放射線部門，[5] 金沢大学大学医薬保健学総合研究科保健学専攻

Complex cardiac malformations are extremely diverse. In emergency cardiac surgery, it is necessary to understand the complicated three-dimensional (3D) cardiac structure quickly and accurately before the operation. Tomographic images such as contrast-enhanced computed tomography (CT) provide a lot of precious information about cardiac morphology. However, a large quantity of CT images makes the understanding be difficult. To overcome this issue, 3D volume rendering images of whole heart are often generated and observed. In an emergency, the 3D images need to be generated in a short time.

We focused on U-Net, which is a deep learning method that is expected to automatically segment the anatomical structure as an elemental technology to support the rapid diagnosis of cardiac malformations. There are few reports that apply deep learning to regional segmentation of the pediatric cardiac system. In this study, we investigate the feasibility of pediatric whole heart segmentation using U-Net.

Approximately 7,000 cardiovascular-enhanced CT images were used for cases in which no morphological abnormality by radiologist was noted in the heart (N=20, age: 2 months to 13 years old, average 5 years old). The window setting was mediastinum condition (WL/WW=50/400HU). The cardiac regions including the atria, ventricles, great vessels (superior vena cava, pulmonary artery, ascending aorta), and soft tissues such as myocardium were manually extracted and used as teacher images of U-Net. The teacher images and the original CT images were converted into PNG format with 256x256 of matrix size and inputted to U-Net. The dice coefficients were obtained by the 10-fold cross-validation method, and the average dice coefficient was evaluated.

The average dice coefficient was 0.87. This result showed that U-Net provided good accuracy for the segmentation of pediatric whole heart. Deep learning like U-Net would become a useful elemental technology in computer-aided diagnosis of pediatric cardiology.

P2-5 Deep Learning を用いた非球形状マーカの認識手法の開発

○阿部　和貴[1]，宮崎　康一[2]，横川　航平[2]，高尾　聖心[3]，平田　雄一[4]，梅垣　菊男[2]，宮本　直樹[2,3]

[1]北海道大学大学院医理工学院，[2]北海道大学大学院工学研究院応用量子科学部門，[3]北海道大学病院医学物理部，[4]北海道大学アイソトープ総合センター

放射線治療中に得られるX線透視画像から，腫瘍近傍に留置された非球形状マーカを認識するため，Deep Learningを用いたマーカ認識手法を開発する．

非球形状マーカはX線透視角度によって画像上の形状が異なるため，テンプレートマッチングでは認識が困難である．本研究では，Deep Learningを用いてマーカ画像と背景画像を分類し，その分類スコアに基づいたマーカ位置決定を考えた．

学習用のマーカ形状パターンを増やすために，マーカ画像と背景画像をX線の減弱に基づいて合成し，マーカ画像を作成した．これらの画像をDeep Learningモデルに入力し，マーカの分類スコアを算出した．そして追跡用画像において座標毎にマーカ尤度を算出し，マーカ位置を決定した．

マーカ画像の分類精度は，検証用画像に対して99%の正解率で分類できた．また，追跡が成功した画像では2次元距離で2 pixel以下を達成した．

マーカ形状パターンを学習することで，形状変化にロバストなマーカ認識手法を開発した．

P2-6 骨塩定量ファントムを用い臨床用CT画像から大腿骨骨密度を自動計測するシステムの開発

○上村　圭亮[1,2]，大竹　義人[1]，スーフィマーゼン[1]，川崎　明宙[1]，高尾　正樹[3]，菅野　伸彦[2]，佐藤　嘉伸[1]

[1]奈良先端科学技術大学院大学生体医用画像研究室，[2]大阪大学運動器医工学治療学，[3]大阪大学整形外科

Preoperative assessment of femoral anatomy and femoral bone mineral density (BMD) is important in the treatment of hip diseases. While preoperative CT images are acquired routinely to evaluate the three-dimensional anatomy of the femur, dual energy X-ray absorptiometry (DXA) is recommended by the World Health Organization (WHO) when evaluating the BMD. There is clinical requirement to quantify the regional BMD from preoperative CT images as DXA is not available at all hospitals and as extra cost is needed when it is available.

Previous studies have reported the usefulness of quantitative CT (QCT) when measuring the BMD from clinical CT images. In QCT, a calibration phantom including hydroxyapatite with a known density (0 mg/cm3, 50 mg/cm3, 100 mg/cm3, 150 mg/cm3, 200 mg/cm3) is included in the field of view to convert the CT numbers (Hounsfield Units: HU) to mg/cm3. This conversion process is necessary to compare the results across patients as HU can be affected by several factors such as the type of CT scanners, and imaging protocol (e.g. tube voltage, slice thickness). Usually, researchers manually set regions of interest (ROI) to isolate the calibration phantom, measure the HU within the ROI, and adopt a linear regression model to convert the HU into BMD. However, this process is time consuming and is prone to manual selection error. Thus, we sought to use deep learning to automatically extract the calibration phantom from CT images and to develop a system that automatically calculates the BMD at the ROI. By developing this system, BMD can be automatically calculated in a large cohort, which will be useful in comparing the BMD data across subject groups and for analyses using a large database.

P2-7 Conversion of H&E Stained to EVG Stained Histological Images using CycleGAN

○ビシュワストンニ[1]，鈴木　裕之[2]，石川　雅浩[3]，小林　直樹[3]，小尾　高史[4]

[1]情報通信系，東京工業大学，[2]数理データ科学教育研究センター，群馬大学，[3]保健医療学部，埼玉医科大学，[4]科学技術創成研究院，東京工業大学

Microscopic examination of tissue is a very crucial part of disease diagnosis and prognosis for understanding tissue morphology and function. Unstained tissue provides very low contrast and less information. So, tissue needs to be stained with proper staining technique before observing with microscope. Staining refers to the method of using coloring reagent for improving tissue contrast and for revealing complex tissue morphology. Hematoxylin and Eosin (H&E) is the most common staining technique. Elastic fiber is a connective tissue fiber which is responsible for elasticity of the organ. Abnormality of elastic fiber in tissue is highly related to various kinds of diseases. But, elastic fiber cannot be identified with H&E staining because of its similar color and pattern to collagen fiber. To solve this problem, usually spatially consecutive pathological sections are stained with special staining technique such as Verhoeff's Van Gieson (EVG) staining which shows elastic fiber in deep blue and collagen fiber in orchid color thus makes them differentiable. But, EVG staining is very complex and costly. Computerized generation of EVG stained image can be a great savior. In this study, Generation of EVG stained image from H&E stained image using CycleGAN architecture has been addressed. With the generation of digitally stained images, this study facilitates pathologists to observe the effect of different stains on same specimen. Leveraging CycleGAN model, this work has been performed with unpaired images of training data where images from two different domains are of same location but need not to be exactly aligned and scale adjusted. The convenience of using unpaired training dataset helps to avoid the complexity of preparing exactly paired images. The quality of the generated image has been determined in terms of SSIM and PSNR. Both visual observation and quality metrics support the acceptability of the generated image for further pathological analysis.

P2-8 拡張型 / 肥大型心筋症病理画像の教師なし分類と評価

○梁　天恒[1]，義人　大竹[1]，尾上　健児[2]，SoufiMazen[1]，中野　知哉[2]，坂口　泰弘[2]，畠山　金太[3]，斎藤　能彦[2]，佐藤　嘉伸[1]

[1] 奈良先端科学技術大学院大学，[2] 奈良県立医科大学　循環器内科，[3] 国立循環器病研究センター　病理部

The diagnosis of dilated/hypertrophic cardiomyopathy (DCM/HCM) using whole slide images(WSI) is a laborious and time-consuming task due to the large matrix size and high resolution of the images. Several patch-based supervised learning approaches have been proposed to classify WSI images. However, given the patch diversity with respect to cellular contents, the assignment of a ground-truth label for each patch, rather than a global (i.e. WSI-wise) label, isessential as it affects the training reliability. This project aims at leveraging unsupervised learning approaches to quantitatively assess the patch diversity with respect to disease classes. The proposed approach uses t-distributed stochastic neighbor embedding (t-SNE) and K-meansclustering to assemble the patches into distinct clusters. In this study, 276 WSIs of 96 cardiomyopathy patients were used. The WSIs were cropped into patches with a matrix size of 256 × 256. Empty patches were excluded, thus 7423 patches were used in the analysis. The patches were grouped into 300 clusters using t-SNE with a perplexity of 80. Expert medical doctors were asked to classify the clusters into DCM, HCM or normal classes.

The patch diversity was assessed by evaluating the agreement among the experts on the classification of each cluster. Results have shown that 36 clusters (12%) had agreement by the 4 experts, 130 clusters (43.3%) had agreement by 3 experts, and 64 (21.3%) clusters had agreement by 2 experts. These results indicate the intrinsic diversity in the patch classifications, and thepotential of the proposed approach to obtain training datasets with a higher agreement.

P2-9 Radiomics 画像特徴量による肺がん治療方針決定因子 PD-L1 発現予測の検討

○松崎　和喜[1]，金子　幸生[1]，高橋　勲[1]，松元　祐司[2]

[1] 日立製作所　ヘルスケアビジネスユニット，[2] 国立がん研究センター中央病院内視鏡科

個別化医療の実現が求められる中，生検はがんの確定診断のみでなく，同時に取得される分子情報は治療方針や薬剤選択に必須となっている．また肺がん治療では免疫チェックポイント阻害剤を用いた新たな治療法が導入されている．この薬剤の効果予測には，programmed cell-death ligand 1 (PD-L1) タンパクの発現率が用いられるが，判定の閾値や必要な検体量等に課題が残る．本研究では，生検による取得情報を放射線診断画像の特徴量へ変換学習する画像情報の拡張技術により，生検と同等の介入判断の実現を目的とした．CT 画像，PET 画像，および診療録の患者情報を組合せた Radiomics 特徴量による PD-L1 の発現予測モデルを構築・評価し，発現有無の予測精度は AUC が 0.89 であった．また構築モデルの特徴量を分析し，モデルの説明性を改善した．以上から画像情報による PD-L1 発現予測の可能性が示された．

P2-10 Dilated convolution を用いた FCN による腹腔鏡動画像からの血管領域抽出

○盛満慎太郎[1]，山本　翔太[1]，北坂　孝幸[2]，林　雄一郎[1]，小田　昌宏[1]，伊藤　雅昭[3]，竹下　修由[3]，三澤　一成[4]，森　健策[1,5,6]

[1] 名古屋大学大学院情報学研究科，[2] 愛知工業大学情報科学部，[3] 国立がん研究センター東病院，[4] 愛知県がんセンター，[5] 名古屋大学情報基盤センター，[6] 国立情報学研究所医療ビッグデータ研究センター

本稿では，腹腔鏡手術動画像からの血管領域の自動抽出手法について述べる．腹腔鏡手術は難易度の高い手術であるため，腹腔鏡動画を自動解析することで手術を支援するシステムの開発が進められている．動画からの術具や臓器の自動認識は，コンピュータに手術工程を理解させ，術者の手技を適切に支援させることにつながる．手術工程の理解において，体内構造を表す血管の抽出は重要である．これまでの血管領域検出では U-Net をベースに，時系列情報を扱うための convolutional LSTM 層を追加して動画にした際の滑らかさを改善した．しかし，大域的な情報の不足により，推定領域の境界部分の抽出にばらつきがあった．そこで，本研究では境界部分の抽出精度を向上させるため，異なる dilation rate の畳み込み計算を並列に行う Dilation 層を追加した Dilated LSTM U-Net を用いて血管領域セグメンテーションを行う．実験の結果，境界部分の抽出精度が向上した．

P2-11 少数症例を用いた教師無し深層学習による病変検出画像特徴量の汎用的生成－胸部 CT を用いた実験的検証－

○牛房　和之[1]，根本　充貴[2]，木村　裕一[1]，永岡　隆[1]，山田　誉大[3]，林　直人[4]，田中　敦子[1]

[1] 近畿大学大学院 生物理工学研究科，[2] 近畿大学 生物理工学部，[3] 近畿大学高度先端総合医療センター PET 分子イメージング部，[4] 東京大学 医学部附属病院 22 世紀医療センター

本検討では，過去に提案した画像特徴量生成法について，学習データ量に対する頑健性を検証する．

提案法は正常データのみを学習した複数の Deep Convolutional Autoencoder (DCAE) を用いる．DCAE に入力する画像は，3 次元の肺局所パッチ画像を 2.5 次元変換した複数種の 3 チャネル画像とする．得られる特徴量を用いて，Adaboost 識別機アンサンブルによる肺結節識別処理を行う．本検討では，肺結節を含む胸部 300 例を学習した際の病変識別性能 (ANODE score) と，200 ～ 5 例を学習した際の識別性能を比較した．その結果，学習データ量を 100 例または 30 例以下とした際に 300 症例学習時の性能と統計的有意差が見られ，提案法は学習データ量が少ない場合にも適用できることが示唆された．

P2-12　深層学習を用いた歯科インプラントの分類

○吉位　和将[1]，助川信太郎[2]，原　　武史[1]，山下　勝督[3]，中野　敬介[4]，長塚　　仁[4]，古木　良彦[2]

[1]岐阜大学工学部電気電子・情報工学科，[2]香川県立中央病院歯科口腔外科，[3]ポリテクセンター香川，[4]岡山大学大学院医歯学総合研究科口腔病理病態学分野

インプラントには多くのブランドがあり，メーカーごとに独自に開発されている．患者のインプラントのメンテナンスをする際にはブランドを特定する必要があるが，一人に対して異なるブランドのインプラントが異なる歯科医によって埋入されることが頻繁にあり，歯科医の間で共有される情報がなく特定するのは困難である．歯科パノラマX線画像からは特定に必要な情報を得ることができるが，これを行うには歯科医の努力と経験が必要になる．本研究の目的は，深層学習を用いて歯科パノラマX線画像からインプラントの分類を行い，歯科医や患者の負担軽減に寄与することである．畳み込み層が6つのCNNと，VGG16とVGG19の転移学習及びファインチューニングモデルでの分類精度を評価した．5つのモデルのうちVGG16のファインチューニングが93.5%と最も高い精度を示し，インプラントの分類において深層学習が有用である可能性が示唆された．

P2-13　大域及び局所情報を用いた深層学習による出血領域自動セグメンテーション

○山本　翔太[1]，盛満慎太郎[1]，林　雄一郎[1]，北坂　孝幸[2]，小田　昌宏[1]，伊藤　雅昭[3]，竹下　修由[3]，森　　健策[1,4,5]

[1]名古屋大学大学院情報学研究科，[2]愛知工業大学情報科学部，[3]国立がん研究センター東病院，[4]名古屋大学情報基盤センター，[5]国立情報学研究所医療ビッグデータ研究センター

本稿では，腹腔鏡手術動画像に対して，深層学習を用い出血領域のセグメンテーションを自動的に行う手法について検討する．出血領域のセグメンテーション結果は，内視鏡手術の安全性を高める研究において利用価値が高い．しかし，手術映像中における出血領域を手動でセグメンテーションするのは困難であるため，効率的にセグメンテーションを行う手法が求められている．本研究では，深層学習を用いた出血領域の自動セグメンテーションを行う．出血領域は画像全体に対して小領域であるため，画像全体から出血領域を含む局所領域を検出するYOLOと，出血領域を切り出したデータで学習したU-Netを用意し，YOLOの推定結果にU-Netの推定結果を統合することでセグメンテーション結果を得る．実験により，腹腔鏡動画像における微小な出血領域の自動抽出には，局所領域に着目することが有効であるという知見が得られた．

P2-14　深層学習による大規模多様皮膚疾患画像データベースの分類

○田中　雅也[1]，斉藤　　篤[1]，志藤　光介[2]，藤澤　康弘[3]，山崎　研志[2]，藤本　　学[4]，村尾　晃平[5]，二宮洋一郎[5]，佐藤　真一[5]，清水　昭伸[1]

[1]東京農工大学大学院工学研究院，[2]東北大学大学院医学系研究科皮膚科学分野，[3]筑波大学医学医療系皮膚科，[4]大阪大学大学院医学系研究科皮膚科学，[5]国立情報学研究所医療ビッグデータ研究センター

Skin diseases have large variability in appearance as well as diversity in distribution of lesions throughout a human body, which makes the diagnosis difficult for non-dermatologist. In this paper, we present a deep learning based computer-aided diagnosis (CAD) system for photographic image classification of 59 skin diseases. The previous skin disease CAD systems were mainly developed to classify localized lesions, such as melanoma, with an image only presenting a skin lesion and taken by a standardized imaging condition. This study deals with various images of 59 skin diseases including skin tumors, inflammatory diseases, autoimmune diseases, etc., in which morphology and distribution of lesions in body parts are largely different each other. The proposed system allows input of a few to dozens of images for one patient taken by dermatologists using a standard optical camera with various imaging conditions including magnifications, angles, and lighting conditions. The system observes lesions from various viewpoints and effectively integrates features relating to skin lesions from multiple images across over body parts and from images with various imaging conditions as dermatologists do. This paper demonstrates the effectiveness of the system using nearly 90,000 images in the national database of Japanese Dermatological Association. We also discuss the consistency between the learned network and medical classification of skin diseases, and the limitation of the system.

P2-15　個別化運動器科学に向けたCT画像からの下腿部・足部筋骨格の自動セグメンテーション

○スーフィーマーゼン[1]，大竹　義人[1]，宮本　拓馬[2]，田中　康仁[2]，上村　圭亮[1]，高尾　正樹[3]，菅野　伸彦[3]，佐藤　嘉伸[1]

[1]奈良先端科学技術大学院大学，[2]奈良県立医科大学，[3]大阪大学大学院医学系研究科

我々は，CT・MRIをはじめとする医用画像から，筋肉を含む個人の筋骨格解剖を忠実に再現する3Dデータを，全自動かつ高速に構築することを目指している．これにより，従来の平均的な解剖モデルによるバイオメカニクス解析や，運動機能予測を被験者個別に行う事で，整形外科や歯科での治療計画やリハビリテーションプランの作成，スポーツ科学におけるトレーニング成果の予測などを可能にする．これまで，人工股関節手術の計画立案を目的として，骨盤・大腿骨および股関節周辺の筋肉を対象に，深層学習を用いたCT画像からの自動セグメンテーションシステムを構築し，多くのテストデータで実用的な精度を示す事を確認してきた．本研究では，下腿部および足部の筋骨格を対象として，Bayesian U-netを用いた自動セグメンテーションおよび精度の定量評価を行ったので報告する．

P2-16 機械学習を用いたハイパースペクトルによる K-ras 遺伝子の識別

○櫻井　千子[1]，市川　盛規[1]，吉谷　文孝[1]，伊藤　隆晃[1]，中矢　大輝[1]，佐鳥　　新[1]，野呂瀬朋子[2]，大池　信之[2]

[1] 一般財団法人 未来科学研究所，[2] 昭和大学藤が丘病院 認証病理診断科

ハイパースペクトルカメラは対象物から入射する電磁スペクトルの高い波長分解能で検出できる VNIR センサーであり，近年医療分野での応用が期待されている．本研究ではハイパースペクトルカメラを生体分野に応用し大腸癌の HE 標本を北海道衛星株式会社製の HSC1702 で撮影し遺伝子検査を実施することなく，K-ras 遺伝子の陽性，陰性の識別を実施する．

昭和大学藤が丘病院の大腸癌症例 37 例（陽性 20 例　陰性 17 例）の平均スペクトルデータに対し機械学習の一種であるランダムフォレストを用いることで 61％の識別精度が得られた．また，識別精度向上のため病理標本間での染色濃度の類似した検体（陽性 10 例　陰性 10 例）を選択し染色濃度のばらつきを減らし再度，機械学習を行ったところ識別精度が 77％に向上した．このことからハイパースペクトルカメラを病理組織診断に用いることにより K-ras 遺伝子の予測可能性が示唆された

P2-17 機械学習を用いたハイパースペクトルによる肺癌の分類

○江崎　雅隆[1]，神谷健太郎[1]，小倉　翔悟[1]，伊藤　隆晃[1]，中矢　大輝[1]，佐鳥　　新[1]，井上　雅文[2]

[1] 一般財団法人 未来科学研究所，[2] 日本独立行政法人 地域医療機能推進機構 東京新宿メディカルセンター 病理科

ハイパースペクトルカメラは対象物から入射する電磁スペクトルを高い波長分解能で検出できる VNIR センサーであり，近年医療分野での応用が期待されている．本研究ではハイパースペクトルカメラを生体分野に応用し，肺癌 HE 標本を北海道衛星株式会社製の HSC1702 で撮影，肺癌の扁平上皮癌と腺癌の機械学習による識別を試みた．

新宿メディカルセンターの肺癌症例 68 例（腺癌 42 例, 扁平上皮癌 26 例）のハイパースペクトルデータに対し, ランダムフォレスト等の学習器を組み合わせることによって一致率 84.4% の識別率を得られた．

そこでつぎに，細胞診検体を撮影し，細胞塊内のがん細胞の予測を試みた．結果として，病理医のあらかじめ指定した細胞塊に関して画像上で正しく識別出来ていることを確認した．このことから，ハイパースペクトルカメラの病理細胞診断における応用可能性が示唆された．

P2-18 広範囲の隣接関係を考慮したグラフニューラルネットワークを用いた腹部動脈血管名自動命名の検討

○日比　裕太[1]，林　雄一郎[1]，北坂　孝幸[2]，伊東　隼人[1]，小田　昌宏[1]，三澤　一成[3]，森　　健策[1,4,5]

[1] 名古屋大学大学院情報学研究科，[2] 愛知工業大学情報科学部，[3] 愛知県がんセンター，[4] 名古屋大学情報基盤センター，[5] 国立情報学研究所　医療ビッグデータ研究センター

本稿では，3 次元腹部 CT 像から抽出された腹部動脈領域に対してグラフニューラルネットワーク (Graph Neural Network, GNN) を用いた血管名自動命名についての検討を行ったので報告する．腹部動脈領域に対して機械学習を用いた血管名自動命名を行う手法はこれまでにもいくつか提案されてきた．また，近年ではグラフ構造に対する機械学習が盛んに行われており，その有用性が示されているため，GNN を用いた機械学習による血管名自

動命名について検討した．一般的な GNN では一つの隣接関係だけを学習することしかできないが，広い範囲の隣接関係を同時に使うことで高精度の自動命名が実現可能であると考えた．本稿では複数の隣接関係を用いて学習を行うことができる MixHop と一般的な GNN である Graph Convolutional Network を使用し，実験と比較を行った．実験の結果，MixHop は全症例のノードを用いた評価では 85.2%，症例ごとの評価では最高で 95.6% の精度で命名することができた．

P2-19 歯科パノラマ X 線画像における分岐を含む SSD による 1 クラス検出と多クラス検出の結果の統合による歯列の認識手法

○森下　拓水[1]，村松千左子[2]，周　　向栄[3]，髙橋　　龍[4]，林　　達郎[4]，西山　　航[5]，原　　武史[3]，有地　淑子[6]，有地榮一郎[6]，勝又　明敏[5]，藤田　広志[3]

[1] 岐阜大学大学院知能理工学専攻知能情報学領域，[2] 滋賀大学データサイエンス学部，[3] 岐阜大学工学部電気電子・情報工学科，[4] メディア株式会社，[5] 朝日大学歯学部口腔病態医療学講座歯科放射線学分野，[6] 愛知学院大学歯学部歯科放射線学講座

歯科パノラマ X 線画像は全国の 98% 以上の歯科医院で撮影されるが，画像には口内のすべての歯牙が描出され，歯科医師がすべての歯牙に正確な画像診断をするには多大な時間と労力を必要とする．本研究の目的は，深層学習を用いて歯科パノラマ X 線画像を解析し，歯科医師の診断補助に寄与することである．その初期段階として，100 症例を用いて画像内の歯牙の検出と歯種の分類を行った．Single Shot Multibox Detector (SSD) に分岐を与えたネットワークで，歯種を区別しない 1 クラスでの歯牙の検出結果と，上顎と下顎で区別した中切歯，側切歯，犬歯，第一小臼歯，第二小臼歯，第一大臼歯，第二大臼歯，第三大臼歯の 16 クラスでの歯牙の検出・分類結果を統合処理し，16 クラスでの検出・分類結果と比較した．結果，検出率が 2.0% 向上，誤検出数が 0.05 個増加した．また，本研究で歯種分類の指標としている歯牙別検出率（検出・分類に成功した歯牙の総数 *100/ データセット内の歯牙の総数）が 0.7% 向上した．

P2-20　畳み込みニューラルネットワークによるクロスフェーズ CT 画像位置合わせ

○胡　　涛[1]，小田　昌宏[1]，林　雄一郎[1]，魯　　仲陽[1]，隈丸加奈子[2]，青木　茂樹[2]，森　　健策[1,3,4]

[1] 名古屋大学大学院情報学研究科，[2] 順天堂大学大学院医学研究科，[3] 名古屋大学情報基盤センター，[4] 国立情報学研究所 医療ビッグデータ研究センター

Non-contrast and artery phase CT both play an important role in the diagnosis of many diseases. The registration from artery phase to non-contrast CT eliminates the displacement and deformation between the two phases, thus enabling convenient observation of vascular structures or lesions. However, the appearances of these regions can be greatly different between the two phases, sometimes leading to the failure of the cross-phase registration procedure. In this paper, a deep-learning based method has been proposed for improving the accuracy of cross-phase CT image registration. Different from traditional methods optimizing an overall objective function, our method utilizes a predefined region-of-interest and is able to pay more attention to the specified structures. The results show that the proposed approach achieves better performance compared with baseline method.

P2-21　P5-24 に変更

P2-22　造影増強 MR 画像におけるベイズ U-Net に基づく不確実性推定を用いた肝臓及び脾臓のセグメンテーション精度予測

○張　　彬[1]，スーフィーマーゼン[1]，大竹　義人[1]，堀　　雅敏[2]，富山　憲幸[3]，佐藤　嘉伸[1]

[1] 奈良先端科学技術大学院大学，[2] 神戸大学大学院医学研究科，[3] 大阪大学大学院医学系研究科

Bayesian U-Net has shown promising performance in image segmentation tasks. Particularly, in a previous study by our group, a Bayesian U-Net has shown possibility to segment hip and thigh muscles in CT images with a high accuracy based on epistemic (i.e. model-related) uncertainty estimation. In addition to segmenting the target structures, the estimated uncertainty demonstrated correlation with Dice coefficient. This allowed to use the uncertainty as a predictive measure for the segmentation accuracy, hence called ‘predictive uncertainty’. However, the capability of Bayesian U-Net to segment MR images has not been examined yet. In this study, we investigated the feasibility of Bayesian U-Net in segmenting the liver and spleen from MR images. A dataset consisting of 132 contrast-enhanced MR images was used. The Bayesian U-Net had an architecture of a conventional U-Net with a dropout layer added after each basic block. At testing, 10 Monte Carlo dropout samples were used for segmentation and uncertainty estimation. The performance of Bayesian U-Net was compared to that of the conventional U-Net trained with a weighted-cross-entropy loss function, in which larger weights were assigned at boundary regions between the liver and spleen. Segmentation accuracy was measured using Dice coefficient based on 2-fold cross-validation. The uncertainty was estimated as an average of the variance of softmax probabilities at the predicted label. Even though boundary regions were better segmented using conventional U-Net, Bayesian U-Net has yielded a higher overall accuracy. The correlation between Dice and the predictive uncertainty was -0.893 and -0.908 for the liver and spleen, respectively. This indicates the feasibility of the uncertainty estimated by Bayesian U-Net as a predictive measure of segmentation accuracy in MR images.

P2-23　皮弁壊死領域早期診断のための狭帯域光源・カラーカメラ連動システム設計

○倉渕　瑶子[1]，大西　　峻[2]，中野　和也[3]，秋田　新介[4]，羽石　秀昭[2]

[1] 千葉大学大学院融合理工学府，[2] 千葉大学フロンティア医工学センター，[3] 宮崎大学キャリアマネジメント推進機構テニュアトラック推進室，[4] 千葉大学医学部形成外科

血流のある皮膚・皮下組織や深部組織により構成される皮弁を用いた遊離皮弁移植術は重要な手技であるが，術後の血管吻合不良や血管閉塞に伴う壊死が発生する可能性がある．壊死の回避には早期の血管再吻合が必要だが，目視での早期発見は困難である．そこでわれわれは，壊死領域を早期に診断するための簡便な撮影システム構築を目指している．具体的には，適当な波長帯と強度を選択した LED 光源を 2 種類用意し，この光源の連続切替と RGB カメラ撮影により，6 次元の信号を取得し診断に利用する．本研究では基礎実験として，ラットの皮弁壊死モデルを作製し，ラット皮弁の血管結紮後早期の分光反射率を取得することで光源設計を行った．設計した光源を用いて壊死領域―正常領域間の分離度を評価したところ，従来の白色光下で取得した RGB 画像に比べて，提案方法において，より高い分離度が確認された．

P2-24　画素毎の難易度に基づく注意機構を用いた敵対的学習によるセマンティックセグメンテーション

○津田　大輝[1]，堀田　一弘[1]

[1] 名城大学

セマンティックセグメンテーションは入力された画像の全画素に対してラベル推定を行う処理である．セグメンテーションモデルを生成器として取り扱い，GAN のように識別器を加えて敵対的に学習を行うことによりセグメンテーションモデル単体よりも高精度な推定を行うことができる．我々はセグメンテーションモデルにとって識別が難しい画素を Attention Map として識別器にリークする方法を提案する．これにより，識別器が従来の学習よりも正確に真偽を識別し，それを基にセグメンテーションモデルは敵対的学習構造により識別器を騙すための学習するため，精度が向上すると期待される．評価実験では細胞画像のデータセットを用いて従来手法と提案手法の精度比較を行った．その結果，Intersection over Union を用いた精度評価において，提案手法が従来手法を上回ることを確認した．

特別講演１

10:40 ～ 11:40　**SL1「Society5.0 時代のヘルスケア」**

座長：陣崎　雅弘（慶応義塾大学）

SL1　Society5.0 時代のヘルスケア

○宮田　裕章
慶応義塾大学医学部医療政策・管理学教室　教授

AI がもたらす技術革新に加え，ICT や IoT，ビッグデータなど一連のイノベーションの大波を踏まえた社会変革は医療も大きく変えてきている．新しい社会では企業や行政，学会などの壁を越えてデータを活用し，新しい価値を共創することが不可欠の要素である．この時，疾患を有した後の病院での治療情報だけでなく，在宅医療や介護というその後のフォロー，健康な段階からの検診に加えた様々なライフログデータなど，様々な時系列データを扱うことが可能となる．今後は疾患の診断・治療を ICT や AI でサポートするだけでなく，より早期な段階からのエンパワーメントや治療後のサポートなどが重要となる．コロナショックにおいてもこうした技術は世界中で活用され，新しい社会における医療の役割が示されている．コロナショックの先にある社会の中で，どの様な医療が必要とされるのか？新しいヘルスケアはどのような未来につながっていくのか？現状と展望を概説する．

第２日　９月18日（金）

特別講演２

14:00 ～ 15:00　**SL2「山形大学医学部における重粒子線治療プロジェクト」**

座長：尾川　浩一（法政大学 /JAMIT 学会長）

SL2　山形大学医学部における重粒子線治療プロジェクト

○根本　建二
山形大学　理事・副学長／山形大学医学部東日本重粒子センター　センター長

重粒子線治療は放射線治療の一種で，炭素の原子核をシンクロトロンで加速し，がんに照射します．X 線が効きにくいがんにも有効で，副作用が少ない，短期間での治療が可能など，優れた特徴を持っています．一方で，巨大で高額な施設が必要で，世界では 12 施設（うち国内 6 施設）しかありません．山形大学医学部では，東芝と共同で超電導技術を用い小型化・高機能を実現する重粒子線治療装置の開発導入に取り組んできましたが，2021 年 3 月より前立腺がん，同年 9 月からはいろいろながんの治療が開始されます．以前は先進医療として扱われ，患者さんは約 300 万円の自己負担が求められていましたが，近年，保険適用の範囲も広がり，また，民間の保険でカバーされることが増えてきたため，ごく一般的な治療になりつつあります．今後，小型で導入が容易な山形大学モデルが，グローバルスタンダードな装置として世界に普及することが期待されています．

シンポジウム1

15:20 ～ 16:50　**SY1「粒子線治療と AI 画像技術」**

座長：臼井　桂介（順天堂大学）

放射線治療分野での医用画像の利用が一般化され始め，その流れは粒子線治療へと広がっている．粒子線治療は光子線と比較して，高い線量集中性が原理的に可能であることから，治療時の画像情報を有効的に取り扱うことが，高度な照射技術を併用した粒子線治療において非常に重要である．画像情報に対する人工知能（Artificial Intelligence: AI）の利用は幅広い研究があり，AI 画像技術の医療分野への応用が盛んに進行していることが実感される．AI による画像処理，セグメンテーションおよび予測分析といった技術は，放射線治療分野に新たな質を提供することができる．本シンポジウムでは，粒子線治療での医用画像の現状と治療精度向上へ向けた取り組みを紹介し，AI 技術の活用と今後期待される展望を議論する．本シンポジウムを通じて，画像診断のみならず放射線治療分野においても AI 技術の導入が求められていることをご理解いただければ幸甚である．

SY1-1　診断から治療へ　-AI の積極的な臨床応用に向けて -

○馬込　大貴
駒澤大学

人工知能（artificial intelligence: AI）の放射線治療への応用が進んでいる．放射線治療後の治療結果は，画像情報を含んだ多くの因子間の複雑な相互作用によって決定されるため，AI を用いることで治療結果を高精度に予測できる可能性がある．また，AI をどのように臨床応用するかは，アイディア次第で無限の可能性がある．本発表では，個別化医療の実現を目指し，より積極的に AI を活用する研究例を紹介する．

SY1-2　陽子線治療における画像誘導治療と応用

○堀田　健二
国立がん研究センター東病院

近年の陽子線治療では，治療室内 CT の利用により，軟組織の解剖学的構造を参照しての画像誘導が可能になった．一方で明確な飛程を持つという陽子線の特徴から，骨や腔など患者体内の解剖学的構造の変化により線量分布が変化しうる．我々は治療毎に適切な治療ビームが照射されることを確認するため，短時間で線量分布をシミュレート可能な画像誘導システムを開発した．さらに治療ビームを調整するレンジアダプティブ陽子線治療への発展を検討している．

SY1-3　重粒子線治療の臨床運用における人工知能への期待と i-ROCK での取り組み

○蓑原　伸一
神奈川県立がんセンター

i-ROCK では 2015 年 12 月に炭素線スキャニング法での治療を開始し，病院併設の重粒子線施設として年間約 500 名の治療（他に光子線で年約 1000 名）を行っている．各治療室は X 線撮影装置，in-room CT 等を有し，日々多くの画像データが発生する．一方，限られたマンパワーでこれらを系統的かつ効率的に解析評価し，患者臨床に短時間でフィードバックすることが求められ，AI 技術への期待は大きい．i-ROCK での産官学連携（横国大，東芝）による AI 利用に向けた取り組み等について紹介する．

SY1-4　深層学習を用いた適応陽子線治療のための CBCT 散乱線補正と三次元陽子線線量分布計算法の開発

○野村　友祐
北海道大学

リアルタイム適応陽子線治療のため CBCT プロジェクション画像の散乱線を補正する CNN と三次元陽子線線量分布を計算する 3D-CNN を開発した．散乱線補正 CNN は従来法より CBCT 画像をより正確に補正する．線量計算 3D-CNN は少ない陽子線スポットビームデータでも正確に陽子線線量分布を計算できる．本手法は個別に使用できる他，これらを組み合わせることでオンライン線量分布評価など様々な応用が期待される．

一般演題セッション3

17:00 ～ 18:30　**P3「イメージング・再構成・画像生成」**

座長：湯浅　哲也（山形大学）

P3-1　3D 血管撮影のための少数方向トモシンセシスの画像再構成法

○堀　　拳輔[1]，橋本　雄幸[1]
[1] 杏林大学大学院保健学研究科

Interventional radiology (IR) is known as minimally invasive procedure techniques under image guidance, and an x-ray projection image is mostly used. Computed tomography (CT) is recently beginning to be used for IR, however, the procedure and imaging cannot be simultaneously performed. Few projections tomosynthesis is a method that reconstructs tomogram using projection data acquired from few-views and limited-angle. In this method compared with CT, the apparatus is small and the tomogram enables to be reconstructed quickly, thus, it has possibility of obtaining the tomogram while performing procedure. The aim of this study is to improve the resolution of z-direction perpendicular to the detector plane and suppress the artifact caused by insufficient views and limited-angle. The algebraic reconstruction technique (ART) and the projection onto convex set (POCS) method were applied for tomosynthesis image reconstruction and compared by the reconstructed images. In the numerical simulation, the distance between the source plane and the detector plane was 100 cm, the detector element size was 0.2 × 0.2 mm^2, and ϕ 1.0-mm tube phantom arranged along y-direction was used. The source plane where the x-ray tubes were located was the parallel to the detector plane. The x-ray tubes were located using square arrangement that has great resolution in comparison to linear arrangement. In the images reconstructed by ART, the bow tie shaped pseudo shadow was appeared in xz-plane. The full width at half maximum (FWHM) for x- and z-direction was 1.12 mm and 3.03 mm, respectively. In the images reconstructed by POCS method, the pseudo shadow was suppressed and FWHM for x- and z-direction was improved to 1.02 mm and 2.65 mm, respectively. In conclusion, POCS method enabled to improve the image quality compared with ART for few projections tomosynthesis of 3D angiography.

P3-2　取り下げ

P3-3　畳み込みニューラルネットワークによる MR 及び PET 画像からの酸素摂取率画像の予測

○松原　佳亮[1]，茨木　正信[1]，篠原　祐樹[1]，高橋　規之[2]，豊嶋　英仁[1]，木下　俊文[1]
[1] 秋田県立循環器・脳脊髄センター　脳血管研究センター放射線医学研究部，[2] 福島県立医科大学　新医療系学部設置準備室

Oxygen extraction fraction (OEF) is an important biomarker to make a decision for treatments of stroke. However, acquisition of OEF map by positron emission tomography (PET) with oxygen-15 gas is not comfortable for patients because of long fixation time, invasive arterial sampling and radiation exposure. We aimed to predict OEF map from magnetic resonance (MR) and PET images by convolutional neural network, and demonstrate the effect of input data types in the prediction of OEF maps. Cerebral blood flow (CBF) and cerebral blood volume (CBV) maps, acquired from oxygen-15 PET, and routine MR images (T1, T2 and T2*-weighted images) for 120 patients with steno-occlusive disease were learned with U-Net. The MR and PET images acquired from other 25 patients were used as test data. We compared generated OEF maps and their structural similarity (SSIM) index in brain among models learned with only MR images, only PET images and both MR and PET images. Maps generated by the model learned with MR, CBF and CBV images were the most similar to the real OEF maps [SSIM: 0.585 ± 0.045]. The OEF maps generated with the model learned with only MR images had flat texture [SSIM: 0.419 ± 0.037]. Adding CBV maps as input slightly improved similarity: 0.567 ± 0.050 with the model learned with MR and CBF images. The contrast of OEF maps generated with the model learned with both MR and PET images was lower than the real OEF maps, whereas the contrast and texture of the generate OEF maps were similar to the real OEF maps. These results suggest U-Net learned with MR and PET images can predict qualitatively realistic OEF maps. Further improvement of quantitative accuracy to generate OEF maps is required.

P3-4　MR 画像の位相推定を不要とする深層学習型圧縮センシング再構成法の検討

○大内　翔平[1]，伊藤　聡志[1]

[1] 宇都宮大学大学院 工学研究科 システム創成工学専攻

MRI では，装置の磁場が厳密には不均一であること，組織間で磁化率が異なることなどが要因で，再生された MR 画像は実関数とはならず，一般に位相を含んだ複素画像となる．これまでに，CNN を利用した圧縮センシング再構成法が提案されているが，その多くは実関数を想定，もしくは緩やかな位相変化を推定する方法が多く，臨床画像のような複雑な位相変化を持つ画像の再構成を行った報告は少ない．本研究では，複素 MR 画像の実部と虚部をそれぞれ独立に計算できる信号収集法を採用することにより CNN によって独立に実部と虚部を再構成できる新たな再構成法を提案する．既存の位相変化を考慮した反復的再構成法および，CNN による再構成法との比較検討を行った結果，提案法では様々な位相変化を持つ MR 画像の再構成を良好に行うことができ，高い PSNR を得ることができたので報告する．

P3-5　少数投影方向マイクロ CT 像の補正手法の開発

○岡本　尚之[1]，大西　　峻[2]，羽石　秀昭[2]

[1] 千葉大学大学院融合理工学府，[2] 千葉大学フロンティア医工学センター

数ミクロンの空間分解能を有するマイクロ CT は，三次元ボリュームの非破壊イメージングが可能であり，病理診断への導入の有効性が示唆されている．しかし，高画質なマイクロ CT 像を得るためには，長時間撮影を要する．本研究は，病理診断応用に向けたマイクロ CT 像の画像補正手法の構築を目的とする．本稿では，投影方向数の削減に基づくマイクロ CT の高速撮影を検討する．提案手法では，少数投影方向とフルサンプリングのサイノグラム像をシミュレーションにより生成し，後者を教師とする Fully Convolutional Network を構築する．実験では，染色処理を実施したラット臓器を使用した．結果，補正前の CT 像に生じていたスポーク状のアーチファクトが低減し，血管構造が明瞭に描出されることを確認した．また，数値評価では，PSNR，SSIM ともに大幅な改善を示した．

P3-6　画像処理法を用いた冠動脈 CTA 石灰化プラーク除去法の開発と臨床評価

○渋谷　幸喜[1]，湯浅　哲也[2]，久根崎遥加[2]，工藤　和也[3]，東海林　綾[3]，藤原　敏弥[4]，柴原　　徹[4]，大村　知己[5]

[1] 北海道科学大学　保健医療学部　診療放射線学科，[2] 山形大学大学院工学部理工学研究科応用生命システム工学科，[3] 市立秋田総合病院放射線科，[4] 市立秋田総合病院循環器内科，[5] 秋田県立循環器・脳脊髄センター放射線科診療部

冠動脈に石灰化が付着している症例の冠動脈 CT 画像は，石灰化が原因で発生する高輝度陰影と blooming artifact が原因で診断不可能となる症例が多数存在するため，臨床現場で問題となっている．この問題を解決するため，画像処理技術を用い，中間強調再構成関数と中心係数 9，周辺係数 -1 に設定したラプラシアンフィルターを適用し 2 段階鮮鋭化を行った後，zero-cross 法を用い石灰化プラークの境界を抽出し石灰化除去を行う image processing algorithm for removing calcified plaque (iPARC) 法を開発し臨床評価を行った．臨床評価は，画像にブレが無くステントが挿入されていない石灰化付着 74 本の冠動脈を対象とし，経験年数 29 年と 23 年の循環器内科医師 2 名が行った．iPARC 法を用い石灰化除去を行うことにより，陽性予測的中率は 16.2% から 77.0% へ大幅に向上した．iPARC 法は，冠動脈 CT 石灰化プラーク付着症例新規診断画像として臨床使用が期待された．

Coronary CTA adhered calcified plaques with coronary arteries are problematic in clinical practice because there are many cases in which diagnosis is not possible due to high-intensity shadows caused by calcification and blooming artifacts. In order to solve this problem, we use image processing technology, apply the Laplacian type filter with the intermediate emphasis reprocessing function and the center coefficient of 9 and the marginal coefficient of -1 to perform two-step sharpening, and then use the zero-cross method. We developed an image processing algorithm for removing calcified plaque (iPARC), which extracts and removes calcified plaques, and conducted clinical evaluation. The clinical evaluation was carried out by two cardiologists with 29 years of experience and 23 years of experience, targeting 74 coronary arteries with calcified plaques without blurring in the images and without stent insertion. By using iPARC method to decalcify, positive predictive predictive value was significantly improved from 16.2% to 77.0%. The iPARC method is expected to be used clinically as a new diagnostic image for cases of coronary CTA adhered calcified plaques at with coronary arteries.

P3-7　位相拡散フーリエ法の信号－画像間学習による MR 画像の超解像

○若槻　泰迪[1]，植松　　駿[1]，伊藤　聡志[1]

[1] 宇都宮大学大学院 地域創生科学研究科 情報電気電子システム工学プログラム

位相拡散フーリエ変換法 (PSFT) の信号は，被写体が実関数との拘束条件を使用すると，撮像後の処理により信号の外挿が可能であり，高解像度化した画像を再生できる．従来，反復的手法により MR 信号を外挿し，再構成像の空間分解能改善を行ってきた．しかし，この方法では 20 回程度の反復処理を要する．本研究では Generic-ADMM-Net を利用し，位相拡散フーリエ法による再生像の空間分解能改善処理を検討する．シミュレーションでは信号取得幅を 1/2 に制限した PSFT 信号を深層学習に入力し，全信号による再生像を教師像として学習を行った．結果，雑音の増幅が抑えられ，かつ，折り返しアーティファクトの少ない画像が得られた．また，汎用のフーリエ変換法に，本方法を適用した場合よりも優れた分解能を有する再構成像を得ることができ，PSFT を利用することの優位性が示された．

P3-8　位相拡散フーリエ変換映像法の導入による MR 深層学習再構成の高画質化

○植松　駿[1]，伊藤　聡志[1]

[1] 宇都宮大学大学院 地域創生科学研究科 工農総合科学専攻 情報電気電子システム工学プログラム

深層学習を利用した圧縮センシングの画像再構成は，再構成時間の短縮化と画像の高画質化のいずれにも有効な方法として注目されている．MRI に圧縮センシングを応用する場合において非線形エンコーディングを行う位相拡散フーリエ変換映像法を使用すると，再構成像の画質を改善できることが報告されている．そこで，本研究では，交互方向乗数法 (ADMM) を CNN で実現した Generic-ADMM-Net を使用し，深層学習による画像再構成を試みた．画像再構成シミュレーションを行った結果，従来のフーリエ変換を基本とする MR 信号を入力とした場合よりも，画像の精鋭化，構造保存性，コントラスト保存性のいずれにも改善された画像が得られた．本研究により，位相拡散フーリエ映像法を利用した圧縮センシングは深層学習再構成によっても高画質化が図られる可能性が示された．

P3-9　数値シミュレーションを用いた非等方ボクセルに対応した粘弾性率分布推定手法の定量性評価

○武田　将岳[1]，藤原　宏志[2]，菅　　幹生[1,3]

[1] 千葉大学　大学院融合理工学府，[2] 京都大学　大学院情報学研究科，[3] 千葉大学 フロンティア医工学センター

【背景・目的】磁気共鳴エラストグラフィ（MRE）は，組織の粘弾性率を定量的に測定可能な非侵襲的手法である．撮像対象外部より振動を加え，内部に伝播する弾性波を画像化し，弾性波画像に対して逆問題を解くことで粘弾性率を推定する．ノイズに頑健な解析手法として３次元積分型公式（ITRF）がある．しかし，ハードウェアの制約などによって弾性波画像を非等方ボクセルでしか取得できない場合には，ITRF を適用できない．本研究では，非等方ボクセルに対応できるようにした ITRF（A-ITRF）の定量性を数値シミュレーションにより評価することを目的とした．【方法】均一弾性体を想定した非等方ボクセルの弾性波画像を作成し，A-ITRF のノイズや横波伝播方向と定量性の関係を評価した．【結果】実測で生じるノイズや横波伝播方向によらず，粘弾性率推定平均値の誤差は 0.2% 以内，変動係数は 1.1% 以内となった．【結語】A-ITRF は高い定量性で粘弾性率推定が可能であることが示唆された．

P3-10　U-Net による X 線透視血管造影像からの血管可視化技術の開発：時空間情報を利用した初期的検討

○木村　竜誠[1]，大野　智之[2]，寺本　篤司[1]，齋藤　邦明[1]，藤田　広志[3]

[1] 藤田医科大学大学院　保健学研究科，[2] 藤田医科大学ばんたね病院，[3] 岐阜大学　工学部

X 線透視血管造影像から血管のみを可視化する技術として，Digital Subtraction Angiography（DSA）がある．DSA では，造影剤投与後の画像から投与前の画像を差分することで，骨や器具などを除去し，血管のみを可視化する．しかし，撮像中の患者の動きにより，差分が正しく行われず，診断精度が低下する問題がある．そこで U-Net を用いることで，血管造影像のみから血管を自動で可視化する技術を新たに開発する．

本手法では，２D U-Net に時間軸を加えた 3D U-Net を使用する．注目フレーム及びその前後１フレームの血管造影像を U-Net に入力することで，血管のみが可視化された注目フレームの DSA 像が出力される．

検証には，藤田医科大学ばんたね病院にて撮影された，実際の頭部 DSA 像を使用した．出力結果から，本手法により不要な骨成分などが除去され，血管のみが良好に可視化されることが確認された．

P3-11　深層学習を用いた圧縮センシング CT 画像再構成の画質改善

○工藤　博幸[1]，森　　和希[1]，鈴木　朋浩[1]

[1] 筑波大学システム情報系

トータルバリエーション（TV）正則化などの圧縮センシングを用いた CT 画像再構成は，画像を階段状の関数で近似する原理に基づくため 1) Staircase Artifact が発生すること，2) 滑らかな濃度変化やテクスチャーが消失しやすいこと，の問題点がある．本論文では，深層学習を用いて圧縮センシングの画質性能を向上させる新手法を提案する．提案手法では，フィルタ補正逆投影（FBP）法と TV や非局所 TV に代表される圧縮センシングの再構成画像を，２入力１出力の画像変換を行う CNN に入力して，画質を改善した出力画像を得る．CNN の構造と圧縮センシングの正則化パラメータは，学習データを用いて二乗誤差が最小になるように自動決定する．スパースビュー CT と低線量 CT の状況を想定して，腹部 CT 画像のデータセットを用いたシミュレーション実験を行った．深層学習単独の場合には血管など学習データから学びにくい構造物の消失が随所に見られ，TV や非局所 TV では Staircase Artifact や滑らかな濃度変化の発生が見られた．これに対して，両者を組み合わせた提案手法では，各々の手法の欠点を補い劇的な画質改善が実現できた．

P3-12　トポロジー制約を用いたスパースビュー CT 画像再構成の試み

○千北　一期[1]，工藤　博幸[2]

[1] 筑波大学システム情報工学研究科，[2] 筑波大学システム情報系

スパースビュー CT 画像再構成は，投影方向数を通常 CT の数十分の一とすることで被曝量や撮影時間の削減を目指す技術として近年注目されている．一方で，再構成された画像には角度欠損によるアーティファクトが発生する等の課題も多くある．先行研究としてこれまで，圧縮センシング及び統計的逐次近似手法に基づく再構成法が多く検討されてきた．本研究では一般画像のセグメント抽出等に使われてきたトポロジー保存の方法に着目する．再構成対象画像の保存するべきトポロジー性質を先験情報として与え，これを定式化することで，逐次的再構成法の中でトポロジー保存を考慮する事のできる新たな正則化項の提案を行う．従来法である OS-SART-TV 法との比較のために，数値シミュレーション実験を異なるトポロジーを持つ複数の画像に対して行った．その結果，従来除去できなかったアーティファクトの除去が可能となり，画像品質向上に繋がることが示唆された．

P3-13　深層学習による圧縮センシング CT 画像再構成画質改善の新手法

○数尾　昴[1]，川又健太郎[1]，工藤　博幸[1]

[1] 筑波大学システム情報系

トータルバリエーション正則化などの圧縮センシングを用いた CT 画像再構成は，Staircase Artifact の発生および滑らかな濃度変化やテクスチャーが消失しやすいことの二点の問題点がある．深層学習単独の場合には学習データから学びにくい構造物の消失が問題点となっている．それぞれの問題に対し本論文では，両者を組み合わせて低線量 CT 及びスパースビュー CT に対する画質改善をさせる新手法を提案する．

提案手法では，フィルタ補正逆投影（FBP）法と TV や非局所 TV に代表される圧縮センシングの再構成画像を，複数入力 1 出力の画像変換を行う CNN に入力して，画質を改善した出力画像を得る．圧縮センシングの正則化パラメータは，視覚的に優れているものを数枚を決定する．スパースビュー CT と低線量 CT の状況を想定して，腹部 CT 画像のデータセットを用いたシミュレーション実験を行った結果，深層学習単独の手法と TV での画像再構成のそれぞれの欠点を補い画質改善を実現した．

P3-14　CT 画像再構成における混合インテリアエクステリア問題の解の一意性

○藤井　克哉[1]，金　鎔采[1]，工藤　博幸[1]

[1] 筑波大学システム情報系

不完全な投影データからの再構成法は，CT 画像再構成の分野のみならず応用数学の立場からも非常に研究されており，よく知られている問題として角度欠損問題，エクステリア問題，インテリア問題等がある． これらの欠損のある投影データから再構成した画像の正しさを保証する解の一意性を示すことは重要であり，近年までに上述の問題における解の一意性の理論は数学的に厳密に構成されている．例えば 2007 年頃に工藤らによって，関心領域内に面積が 0 でない先見領域があれば，インテリア問題における解の一意性が成立することが示された． 本論文では，これらの問題が混在する場合，特に投影データの外部と内部の欠損が同時に起きている混合インテリアエクステリア問題について考察する．この問題における解の一意性は，関心領域内に微小な面積を持つ円環領域が既知であれば成立することを示した．また低コントラスト画像である Shepp-Logan phantom と脳の CT データを用いてシミュレーション実験を行い，構築した理論を検証する．

P3-15　多方向 2 次元超解像を用いた臨床 CT 像の教師なし 3 次元超解像

○鄭　通[1]，小田　紘久[2]，小田　昌宏[1]，中村　彰太[2]，森　雅樹[3]，高畠　博嗣[4]，名取　博[5]，森　健策[1,6,7]

[1] 名古屋大学大学院情報学研究科，[2] 名古屋大学大学院医学系研究科，[3] 札幌厚生病院，[4] 札幌南三条病院，[5] 恵和会西岡病院，[6] 名古屋大学情報基盤センター，[7] 国立情報学研究所医療ビッグデータ研究センター

This paper proposes a novel, unsupervised super-resolution (SR) approach for performing the SR of a clinical CT volume into the resolution level of a micro CT (μ CT) using clinical CT- μ CT training dataset. The precise non-invasive diagnosis of lung cancer typically utilizes clinical CT data for making surgical plans. Due to the resolution limitations of clinical CT (500 × 500 × 500 μ m / voxel), it is difficult to obtain enough pathological information such as the invasion area at alveoli level. On the other hand, μ CT scanning allows the acquisition of volumes of lung specimens with much higher resolution (50 × 50 × 50 μ m / voxel or higher). Thus, SR of clinical CT volume may be helpful for diagnosis of lung cancer and surgery planning. An unsupervised SR method named SR-CycleGAN was proposed for performing SR of clinical CT images. However, this method could only perform SR of 2D clinical CT images in the axis plane: the input is a 2D clinical CT image, and corresponding output is a 2D SR image. For obtaining more pathological information from the SR result from 3D space, a method that could perform SR of 3D clinical CT volumes is desired. In this paper, we propose a SR method that could perform SR of 3D volume, by combining 2D SR results of clinical CT images in axial, coronal and sagittal plane respectively. 2D SR was performed by a modified SR-CycleGAN. Experimental results demonstrate that the proposed method can successfully perform SR of 3D clinical CT volumes of lung cancer patients to μ CT level resolution in 3D scale, while the conventional SR-CycleGAN could only perform SR of 2D clinical CT images.

P3-16 敵対的生成ネットワークによる屈折コントラストCTのためのリングアーチファクト除去

○黄　卓然[1]，砂口　尚輝[1]，島雄　大介[2]，市原　周[3]，湯浅　哲也[4]，安藤　正海[5]

[1]名古屋大学大学院医学系研究科，[2]北海道科学大学保健医療学部，[3]名古屋医療センター 臨床研究センター，[4]山形大学大学院理工学研究科，[5]総合科学研究機構

Refraction contrast X-ray CT (RCT) which represents the refractive distribution proportional to the electron density distribution in sample can obtain comparable tissue contrast as well as pathological section of biomedical soft tissue. The angle analyzer (AA) utilized in this method can modulate the X-ray refraction angle of 10-7 radian order in the sample to intensity. The surface of AA is mirror-polished for high accuracy angle analysis. If there are scratches or dusts on the crystal surface, they lead to angular estimation errors and result in ring artifacts on RCT. For absorption contrast X-ray CT, there are many attempts to remove RA mainly caused by non-uniform sensitivity of the detector, and they work to some extent in RCT. However, it is difficult to remove the RA completely because the reconstruction of RCT involves an integration process in the projection, with which the angular errors spread over the entire projection.

In this research, we proposed a machine learning method based on generative adversarial network (GAN) with physical properties of RCT in order to remove RA. Since the RCT projection is a first-order derivative of the radon transform of the physical quantity distribution of the sample, a Laplacian image of the physical quantity distribution can be obtained by reconstructing it after further differentiation. We incorporated the sparse property of Laplacian image that the sum of pixel values is close to 0 into a loss function of GAN. To demonstrate the effectiveness of this method, we applied the proposed and conventional methods to simulated data of RCT using numerical phantoms and evaluated them from root mean square error and structural similarity. In this result, the proposed method demonstrated to remove RA better than conventional methods. In addition, we obtained an excellent removal effect of RA when applying to actual imaging data of biological tissues.

P3-17 深層学習を用いたヘマトキシリン・エオジン染色画像からKi-67免疫組織化学染色画像への変換

○季　存遠[1]，上原　剛[2]，木村　文一[3]，大島　健吾[3]，山口　雅浩[1]

[1]東京工業大学　工学院　情報通信系，[2]信州大学医学部 病態解析診断学，[3]信州大学医学部　保健学科　病因・病態検査学

The whole slide image (WSI) is an important medium for modern pathology. Digitized microscopic images greatly facilitate the archiving, transferring, and evaluation of pathology specimens, making the diagnosis faster and more reliable.

In the diagnosis of cancers, tissues are commonly processed with the hematoxylin-eosin (H&E) stain, also named as general stain, to reveal the morphology, color, and texture features. Besides, the molecule level activity of the cells, which is utilized in the assessment of malignancy level or the classification of cancer subtypes, is important. Therefore, immunohistochemical (IHC) stain method is introduced to mark specific proteins in the cells. For instance, the Ki-67 protein is one of the popular biomarkers for evaluating the growth factor of a cell population. In practice, the pathologists evaluate the Ki-67 signals with IHC stain images to determine the proliferation activity of the tumor cells and rate the malignancy of cancer accordingly. However, IHC stain suffers from its complicated staining procedure and high cost compared with the HE stains.

In this paper, we propose a digital staining method that employs the U-Net to obtain the semantics of Ki-67 positive and negative nuclei and utilizes the color unmixing technique to yield stain colors. WSIs of well-differentiated endometrial adenocarcinoma of the uterine corpus from Shinshu University Hospital are used. We scanned the H&E stained specimens with a NanoZoomer scanner under 40x objective lens, then perform the IHC reaction for Ki67 in decolorized HE specimens. The classification accuracy of semantic areas reaches 82%, with a 48% precision and a 62% recall for the positive areas. To evaluate the quality of digital staining in terms of diagnostic metrics, we consider the positive rate of all nuclei, also known as labeling index. An end-to-end technique based on cycle-GAN is also tested for comparison.

P3-18 取り下げ

P3-19 マルチピンホールX線蛍光CT画像再構成におけるEM-TVアルゴリズムの検討

○金澤　道和[1]，保科　亮太[2]，佐々木洸介[1]，阿部　智大[1]，笹谷　典太[2]，砂口　尚輝[3]，河嶋　秀和[4]，兵藤　一行[5]，湯浅　哲也[2]，銭谷　勉[1]

[1]弘前大学 大学院理工学研究科，[2]山形大学 大学院理工学研究科，[3]名古屋大学 大学院医学系研究科，[4]京都薬科大学 放射性同位元素研究センター，[5]高エネルギー加速器研究機構 物質構造科学研究所

マルチピンホールX線蛍光CT（multi-pinhole x-ray fluorescence computed tomography; mp-XFCT）では，ある化合物に導入した非放射性核種を入射X線で励起し，発生する蛍光X線を多方向から計測することで，被写体内部の化合物の3次元分布を得る．しかし，生体撮像を行うためには，現行のmp-XFCTシステムでは測定の対象となる化合物に導入されるヨウ素濃度の検出限界が不十分であり，改善が必要である．そこで本研究では，画質改善による検出限界の向上を目的とし，トータルバリエーション（TV）を正則化項として組み込んだEM-TVアルゴリズムを用いた再構成法を提案する．実験では物理ファントムを用意し，その実測データを従来法であるOS-EM法およびEM-TVアルゴリズムで画像再構成を行い，比較した．EM-TVではノイズが抑制され，最小濃度である0.05 mg/mlのヨウ素領域も明瞭に描出された．また，コントラスト対ノイズ比（CNR）を用いて比較したところ，約1.3倍の改善が見られた．ヨウ素の検出限界もOS-EMで0.048 mg/mlであったのに対し，EM-TVでは0.036 mg/mlに改善することができた．

一般演題セッション４

9:00 ～ 10:30　**P4「CT・核医学画像・MRI」**

座長：湯浅　哲也（山形大学）

P4-1　Swish を利用した Blind DnCNN とパラレルデノイジングへの応用

○須貝　太郎[1]，高野　航平[1]，大内　翔平[1]，伊藤　聡志[1]

[1]宇都宮大学　大学院地域創生科学研究科　工農総合科学専攻　情報電気システム工学プログラム

深層学習を利用した雑音除去法は，従来の非線形フィルタを大幅に上回る特性を示し，大きな注目を集めている．本研究では，性能改善と実応用を考慮して Zhang らが提案した DnCNN に対し 3 点の改良を加えた．第 1 点と 2 点は，活性化関数に従来の ReLU ではなく Swish を応用したこと，および雑音量が未知であり，かつ，空間可変な雑音画像に適応しうるブラインド雑音除去を導入したことである．3 点目は，隣接するスライス画像を線形加算して複数の加算画像を作成し，高 S/N 化された加算画像をブラインド雑音除去し，その後に連立方程式を解くことにより雑音除去像を得る方法（パラレルデノイジング）の導入である．高 S/N の画像をデノイジングするので，画像劣化の抑制が期待できる．実験の結果，Swish の採用により ReLU を上回る PSNR が得られた．また，パラレルデノイジングの採用により画質劣化を抑制した雑音除去を行えることが示された．

P4-2　アーチファクト学習を導入した改良型 ISTA-Net による深層学習再構成

○宮本　裕大[1]，伊藤　聡志[1]

[1]宇都宮大学 大学院 地域創生科学研究科 工農総合科学専攻 情報電気電子システム工学プログラム

MRI の撮像短縮化において，圧縮センシングの利用が進んでいる．従来型の反復的再構成法は，多くの計算コストを要すること，スパース近似による画質の低下が課題である．

近年，深層学習を利用した再構成が注目されており，その中に従来型の反復的再構成をディープラーニングネットワークで構成した，ISTA-Net がある．ISTA-Net は，学習時間および再構成時間が比較的短い点で優れているが，アーチファクトの残留が目立つ問題がある．

本研究では，学習時間および再構成時間を維持しつつ画質の改善を図るために，ISTA 処理の後に，畳み込み処理および活性化関数から成る層をカスケード接続し，アーチファクトを検出・除去する方法を導入した．

健常ボランティア画像を使用して再構成シミュレーションを行った結果，アーチファクトの除去性能が向上し，再構成像の画質改善を図ることに成功した．

P4-3　放射光マイクロ CT におけるリングアーチファクト低減の検討

○米山　明男

九州シンクロトロン光研究センター

放射光マイクロ CT は，放射光の高輝度特性を利用してミクロンオーダーの高い空間分解能の非破壊 3 次元観察を短時間に行う撮像法であり，バイオメディカルから材料開発に至る幅広い分野で活用されている．本 CT において X 線画像検出器は，入射 X 線を可視光に変換する蛍光体，可視光用のレンズ系，及び可視光用 sCMOS カメラから一般には構成されている．蛍光体として CsI，GAGG，及び LuAG 等の結晶が使用されているが，結晶欠陥等が存在すると X 線強度と可視光強度が比例関係から外れ，再構成した断面像線には多数のリングアーチファクトが現れる．本発表では，放射光マイクロ CT の特徴である平行性を活用して同アーチファクトを低減する方法に加え，サイノグラム上におけるリングアーチファクトの特徴を捉え，フーリエ変換を利用したフィルターにより更に低減する方法について紹介する．

P4-4　MRI 高速撮像画像における微細構造を保存したノイズ低減

○石原千鶴枝[1]，金子　幸生[1]，白猪　　亨[1]，野口　喜実[1]，尾藤　良孝[2]，荻野　昌宏[1]

[1]株式会社 日立製作所 研究開発グループ，[2]株式会社 日立製作所 ヘルスケアビジネスユニット

MRI では撮像時間を短縮するために様々な高速撮像技術が開発されている．しかし，一般に高速撮像では SNR が低下するため，画質を担保した高速化が課題となる．本研究では，高速撮像技術の 1 つである Parallel Imaging(PI) における低画質画像を，SNR および微細構造の明瞭度の観点において，通常撮像と同等の画質に復元する畳み込みニューラルネットワーク (CNN) の開発を行った．通常撮像条件のボランティア撮像実験によって取得した k 空間信号から，PI 同等の間引きを行った低画質画像と，全てを使用した高画質画像を教師データとして学習させた．デノイズ処理のトレードオフとして生じる微細構造のぼやけ指標 (Blur) を定義し，本指標と SNR を用いて様々な復元系 CNN を評価した．その結果，間引き率 33%の場合においても，低 Blur を維持しながら SNR 向上率 50%以上を実現できる CNN を生成することができた．

P4-5 Support Vector Regression を用いた脳年齢推定と機種間 Harmonization

○舞草　伯秀[1]，ImanBeheshti[2]，曽根　大地[1,3]，松田　博史[1]

[1]国立研究開発法人　国立・精神神経医療研究センター，[2]Université Laval，[3]University College London

MRI から得られる脳体積分布から機械学習に基づく脳年齢推定は，アルツハイマー病や側頭葉てんかん等の脳変性の進行動態を表すバイオマーカーとして着目されている．しかしながら，MRI は磁場強度の違いや装置メーカー，装置モデルの違いにより撮像される画像品質が異なるため，得られた体積値に Measurement Bias が存在することが知られており，これが機械学習の汎化性を著しく減少させることが報告されている．本研究では，一般線形モデルと経験的ベイズ推定により機種間差を推定・補正する ComBat 法を脳体積値に適応し，脳年齢推定の汎化性が向上することを示した．

P4-6 CNN を用いたイオフルパン SPECT からの特異的結合比の推定

○久根崎遥加[1]，岩渕　　怜[1]，荒木　啓史[1]，姜　　時友[1]，湯浅　哲也[1]，松田　博史[2]

[1]山形大学大学院理工学研究科，[2]国立精神・神経医療研究センター

パーキンソン病の診断指標として用いられる特異的結合比 (SBR：Specific Binding Ratio) は，ドーパミントランスポーターに高い親和性をもつイオフルパンを造影剤とする SPECT 画像における線条体内の集積から算出される．しかし，SPECT の低い解像度では線条体の位置を特定することが困難な場合がある．本研究では，正確に線条体を特定した画像から Tossici-Bolt 法により算出した SBR を CNN に学習させ，SPECT 画像から回帰により SBR を推定する方法を提案する．臨床 SPECT 画像（CL: 192 症例，PD: 1186 症例）から算出した SBR を訓練データとして 3 層 CNN を学習させた後，テストデータ（CL: 21 症例，PD: 136 症例）に対して，CNN により算出された SBR に基づいて ROC 解析を行ったところ，0.94 の高い AUC を得た．

P4-7 C 型 Compton-PET ジオメトリの 3 次元イメージングシミュレーション

○仁科　　匠[1]，田島　英朗[2]，田久　創大[2]，錦戸　文彦[2]，菅　　幹生[1]，山谷　泰賀[2]

[1]千葉大学，[2]量研放医研

例えば MRI に後付けできる PET など，リングの一部を開放化した C 型 PET により PET の応用が広がる可能性がある．その一方で，測定データの欠損に起因した強いアーチファクトが再構成画像に発生してしまう．そこで我々は，開放部と対向する位置に散乱検出器を追加して，欠損情報をコンプトンカメラの原理により補う C 型 Compton-PET の開発を進めている．本研究では，C 型 Compton-PET ジオメトリのアーチファクト低減効果をモンテカルロシミュレーションによって検討した．具体的には，半径 20 cm，開放部の角度が 115 度の C 型 PET の内側に，散乱検出器を半径 15 cm で円弧状に配置したジオメトリを模擬した．視野中心に配置した半径 10 cm，体軸方向の長さ 15 cm の円柱ファントムの再構成画像からアーチファクトを定量的に評価した．結果として，視野領域全体における円柱内部の画素値の割合が増加したことからアーチファクト低減に有効であることが示された．

P4-8 深層学習による組織組成比推定を介した PET 吸収係数画像生成

○小林　哲哉[1]，繁木　結衣[1]，森田　尚孝[1]

[1]株式会社 島津製作所 基盤技術研究所

定量的 PET 再構成のためには，高精度な散乱補正・吸収補正が求められ，そのためには正確な吸収係数画像が必要となる．PET-CT 装置や PET-MR 装置では，CT 画像・MR 画像を処理することで吸収係数画像を生成する一方，トランスミッション線源を有さない PET 単独装置では PET データ自体から吸収係数画像を推定する必要がある．ここで，頭部の吸収係数画像は「軟部組織・骨・空洞・被写体外」の 4 つの組織組成比画像の線形結合でモデル化できるので，散乱補正・吸収補正をかけていない未補正 PET 画像から組織組成比画像を推定できれば，間接的に吸収係数画像を推定できると考えた．本研究では，未補正 PET 画像から組織組成比画像を推定するというアイディアおよびそれを実現する深層学習モデルを提案し，頭部 PET のシミュレーションデータを用いた原理検証実験（吸収係数画像の精度評価および PET 画像の定量性評価）の結果を報告する．

P4-9 WGI 試作機によるフルリングコンプトンイメージングの実証

○田島　英朗[1]，吉田　英治[1]，脇坂　秀克[1]，高橋美和子[1]，永津弘太郎[1]，辻　　厚至[1]，鎌田　　圭[2,3]，吉川　　彰[2,3]，パロディカティア[4]，山谷　泰賀[1]

[1]量子科学技術研究開発機構放射線医学総合研究所，[2]東北大学，[3]株式会社 C & A，[4]ミュンヘン大学

我々は，PET リングに散乱検出器リングを挿入し，PET とコンプトンイメージングを組みわせることで，様々なガンマ線をイメージングに用いること可能とする WGI(Whole Gamma Imaging) の開発を進めており，これまでに小動物サイズの WGI 試作機を開発した．WGI 試作機は，我々の知る限り世界初のフルリングコンプトンイメージング装置であり，本研究では，その有効性を実証することを目的とする．PET とコンプトン画像を同一装置で比較するため，陽電子と 909keV のガンマ線を同一分布から異なるタイミングで放出する 89Zr を測定対象に選択した．また，PET 及びコンプトン画像再構成法として，それぞれリストモード OSEM に検出器応答モデルと感度補正を組み込んだ．ファントム及びマウスを測定し，PET 画像と 909 keV のコンプトン画像をそれぞれ再構成した結果，PET に迫る高精細なコンプトン画像を得ることができた．

P4-10 MR 画像特徴量を用いた機械学習モデルによる前立腺癌患者の低悪性度群と高悪性度群への自動層別化

○浦上　暉允[1]，有村　秀孝[2]，木下　史生[3]，今田憲二郎[4]，小田　義直[3]，渡邊壽美子[2]

[1]九州大学医学系学府保健学専攻，[2]九州大学大学院　医学研究院保健学部門，[3]医学研究院形態機能病理学，[4]九州大学病院泌尿器・前立腺・腎臓・副腎外科

[Purpose].

Non-invasive prediction of prostate cancer grade group (GG) with magnetic resonance (MR) imaging without prostate biopsy would be greatly beneficial for patients. This study aimed to explore the potential of automatic stratification of prostate cancer patients into low-grade and high-grade groups using machine learning models with MR image features.

[Method]

MR images of 40 cancer regions corresponding to histopathological images of prostate cancer were selected and they were divided into two groups, low-grade tumors (≤GG2) and high-grade tumors (GG3≤), based on their grade groups. 4320 MR image features were consisted of original-images (16 types of images in 5 sequences)-based and two-dimensional wavelet decomposition-based image features. Significant features were chosen by using a least absolute shrinkage and selection operator (LASSO) algorithm, and they were used to construct 5 supervised machine learning models. Parameters of each machine learning model were optimized based on 5-fold cross validation test so that areas under receiver operating characteristic curves (AUCs) could be maximized.

[Results]

In 5 supervised machine learning model with four MR image features, support vector machine (SVM) achieved the best prediction (AUC = 0.965) for the stratification of prostate cancer patients.

[Conclusion]

This study showed the potential of automatic stratification of prostate cancer patients into low-grade and high-grade groups using the SVM with MR image features.

P4-11 複数周波数同時加振 MR エラストグラフィによる撮像時間短縮の検討

○久慈　　陸[1]，菅　　幹生[1,2]，岸本　理和[2]，野澤　真吾[2]，小畠　隆行[2]

[1]千葉大学，[2]量子科学技術研究開発機構　放射線医学総合研究所

MRE は，生体組織の粘弾性率分布を測定する手法である．複数周波数 MRE では組織粘弾性の周波数特性を取得できるが，撮像時間を要する問題がある．本研究では，40，60，90 Hz の 3 周波数で同時加振する場合において，シミュレーションによって取得位相数による撮像時間短縮効果と弾性率推定精度の関係を明らかにすることを目的とした．単一周波数ごとに 4 位相収集する従来法で 3 周波数を取得する場合と比較して，同時加振で位相数が 7, 8, 9 での撮像時間短縮効果はそれぞれ 44%，37%，30% となる．振幅対ノイズ比がそれぞれ 25，25，7 の弾性波画像を用いた検討では，同時加振で位相数が 7，8，9 での弾性率分布の平均値はいずれも真値と一致し，変動係数はそれぞれ 7%，5%，0%以下となり，位相数 9 では単一周波数加振 MRE よりも変動係数は小さくなった．以上の結果より，3 周波数同時加振 MRE の有効性が示唆された．

P4-12 数値シミュレーションによる MR エラストグラフィに体動が与える影響の基礎的検討

○小賀野　慎[1]，谷口　　陽[1,2]，西原　　崇[2]，瀧澤　将宏[2]，菅　　幹生[1,3]

[1]千葉大学 大学院融合理工学府，[2]（株）日立製作所ヘルスケアビジネスユニット，[3]千葉大学　フロンティア医工学センター

MR エラストグラフィは対象を加振し，MRI により取得した弾性波画像から逆問題を解くことで弾性率を非侵襲的に測定する手法である．撮像中の体動が粘弾性推定に与える影響はこれまで十分に検討されていない．本研究では，撮像中の体動が MRE 撮像に与える影響と，体動による影響を低減可能な撮像条件を数値シミュレーションにより明らかにすることを目的とした．シミュレーションでは，成人男性の腹部が 3 kPa で均一で，呼吸によって前後方向に変形することを想定し，振動子による変位の発生方向と MR 信号収集方向の組み合わせを検討した．腹部変位と横波変位の方向が平行で，phase-encoding 方向が平行と垂直の場合に弾性率の平均値と標準偏差は，3.0 ± 0.2 と 3.0 ± 0.1 となった．腹部変位と横波変位の方向が垂直で，phase-encoding 方向が平行と垂直の場合には，3.0 ± 0.7 と 2.8 ± 1.3 となった．以上の結果より，腹部変位と横波変位の方向が平行となるように振動子を配置することで体動の影響を低減できることが示唆された．

P4-13 機械学習を用いたコーンビーム CT の散乱線除去

○村田　一心[1]，橋本　航汰[1]，臼井　桂介[2]，尾川　浩一[1]

[1]法政大学理工学部応用情報工学科，[2]順天堂大学保健医療学部診療放射線学科

本研究では，コーンビーム CT 撮影で発生する散乱線を機械学習を用いて除去する方法を提案する．コーンビーム CT は検出器面に対し，２次元検出器アレイを用いるため，コリメータを用いた散乱線の除去が難しい．そのため，再構成画像の画質は劣化してしまう．しかし，その画質を改善することができれば，治療計画 CT にコーンビーム CT を使用できる可能性があり，極めて有用である．そこで本研究では，機械学習を用いた散乱線除去方法を提案する．学習データの生成には，モンテカルロ法による光子輸送シミュレーションを用いた．すなわち，散乱線有り無しの画像を作成し，それぞれ入力データ，教師データとした．画像の再構成には FDK 法を適用した．学習結果を適用した画像を評価したところ，散乱線を有意に除去できたことを確認した．この結果は，散乱線の除去に機械学習が有効であることを示す．

P4-14　デコンボリューション法による静止型ピンホール SPECT システムの空間分解能改善

○大越　迪[1]，村田　一心[2]，尾川　浩一[2]

[1]法政大学大学院理工学研究科，[2]法政大学理工学部

通常の SPECT システムでは，データを取得する際に検出器を回転させるため，データ収集に時間がかかり，放射性医薬品の動態情報を得ることは困難である．そこで近年，検出器を回転させない静止型ピンホール SPECT の研究が進められている．このシステムでは，検出器を回転させない代わりに，複数のピンホールを有するガンマカメラを複数台使用している．一般に SN 比を向上させるためにはピンホール径を大きくし，多くの光子を計測すればよいが同時に空間分解能の劣化が生ずる．この補正法として 7 rays 法なども提案されているが，種々の補正も絡むため十分な方法とは言えない．そこで本研究では，再構成空間の中心に対する点拡がり関数を仮定し，投影データに対してデコンボリューションを行い空間分解能の改善を試みた．再構成画像の画質を評価したところ，従来の方法より高い空間分解能が確認でき，PSNR 値も向上したので報告する．

P4-15　At-211 内用療法おける ETCC を用いた画像再構成

○山田　宙輝[1]，村田　一心[2]，谷森　達[3]，尾川　浩一[2]

[1]法政大学大学院理工学研究科，[2]法政大学理工学部，[3]京都大学大学院理学研究科

電子飛跡追跡型コンプトンカメラ (ETCC) は，コンプトン散乱時に発生する反跳電子の軌跡を計測できるコンプトンカメラであり，この軌跡情報から通常のコンプトンカメラでは得ることができない，光子が放出された方向の精度を向上することが可能である．このため，従来のコンプトンカメラと比較して少ないイベント数で画像を再構成できる可能性がある．本研究ではその有効性を検証するために，近年注目されている At-211 内用療法のイメージングを模したシミュレーションを行った．シミュレーションでは半径 10 cm の水球内に At-211 の点線源を設置し，再構成画像の画質を評価した．この結果，大きな腫瘍を見つけるのに十分な空間分解能が得られることがわかり，ETCC が At-211 治療のイメージングに有効であると考えられたので報告する．

P4-16　PET/CT 画像および統計的画像を用いた悪性腫瘍の自動検出法

○相原　拓実[1]，原　武史[1]，周　向栄[1]，藤田　広志[1]，伊藤　哲[2]，加藤　正也[2]，松迫　正樹[3]，片淵　哲郎[4]

[1]岐阜大学工学部電気電子・情報工学科，[2]大雄会病院第一放射線科，[3]聖路加国際病院放射線科，[4]岐阜医療科学大学保健科学部

PET images are useful for finding cancer and checking if the cancer has spread or not. In addition, it has been reported that Z-score images, which are statistical images, can distinguish the normal regions from the abnormal regions than the SUV (Standardized uptake value). The purpose of this study was to develop an automated detection system of tumors in PET/CT images using Z-score images and reducing the burden of doctors. Z-score image is created by comparing pixel values of normal model and target image, so I created Z-score images using TPS(Thin-Plate Spline) transformation for normal models. Many false positives (FPs) are contained in the candidate region. I classified true positives (TPs) and FPs using CNN (convolutional neural network) and deleted FPs. I used 63 abnormal cases as detection targets, and the number of tumors was 336 in the trunk of the body. The number of detected tumors was evaluated using TLG (Total Legion Glycolysis), which is an index of tumor volume and degree of glucose metabolism, and the center of gravity of the tumor area. As a result of experiment that inputs PET images and Z-score images to CNN, 80.4% (270/336) of tumors were extracted by this method, and 18.9 FPs remained per case.

P4-17　手部 X 線画像を用いた骨髄浮腫画像分類の初期検討

○潘　冬平[1]，小田　昌宏[1]，片山　耕[2]，森　健策[1,3,4]

[1]名古屋大学大学院情報学研究科，[2]片山整形外科リウマチ科クリニック，[3]名古屋大学情報基盤センター，[4]国立情報学研究所医療ビッグデータ研究センター

Bone marrow edema is a symptom of early rheumatoid arthritis, which is reported as the main incentive to joint destructions. To diagnose early rheumatoid arthritis, bone marrow edema has been detected through Magnetic Resonance Imaging (MRI). However, the cost of taking MRI images is quite high, which puts a heavy economic burden on the patient. In this work, we are exploring the possibility of recognizing bone marrow edema through X-ray images which are fast and economic. We used a convolutional neural network consisting of convolution blocks and fully connected layers to classify whether bone marrow edema occurs or not. We also used an interpretation technique called the Grad-CAM for visual explanations. By using the neural network, we achieved a recall rate of 84.0% and a precision of 72.4% on this binary classification task.

P4-18　屈折コントラスト法間の画像描写能比較の試み

○大塚　楠平[1]，砂口　尚輝[2]，島雄　大介[1,3]

[1]北海道科学大学大学院，[2]名古屋大学，[3]北海道科学大学

X線は物質中で吸収・散乱を起こすほか，わずかに屈折する．この屈折の程度を画像コントラストとする手法を屈折コントラスト法という．X線の屈折もスネルの法則に従い，物質間の屈折率の差と，異なる物質の境界へのX線の入射角という2つのパラメータに依存する．そのため，この両者を同時に評価する手法が望まれる．そこで我々は先行研究にて，自作ファントムを用いX線の入射角ごとにコントラストを得るために必要な入射角の変位量（Δ）を測定する，という評価法を提案した．本研究ではこの評価法を2種類の屈折コントラスト法に適用した．その結果，これら2種類の画像法のΔには2倍程度の差がある，と評価することができた．この評価法はファントムの投影像を1枚撮影するだけで行える簡便なものであり，各種の屈折コントラスト法の相対評価に有用であると考える．

P4-19　1.5 T MRI における定量的磁化率マッピング（QSM）の画質評価

○高橋　哲彦[1,2]，山路　菜月[1]，林　　則夫[1,2]，竹内　友一[2]，小倉　明夫[1,2]

[1]群馬県立県民健康科学大学 診療放射線学部，[2]群馬県立県民健康科学大学 大学院　診療放射線学研究科

診断画像の定量化や読影支援技術の開発や利用において，元画像の撮像条件と定量解析結果の関連を把握することは重要である．

本研究では，定量的磁化率マッピング (QSM) を汎用的な 1.5TMRI 装置で実用することを目標とし，QSM において，ファントムおよび健常ヒト脳の画質（定量値）と撮像パラメータの関係を数値計算と実測で比較検討した．

QSM 解析ソフトウエアは，日立製作所製を用いた．ヒト MRI 撮像については倫理審査の承認を得た．

被検体の体位は仰臥位，撮像断面はHF方向である．撮像シーケンスは3DGrEシーケンスとし，撮像条件はTRとTEを変えた．QSM 画像を作成し解析した結果，

① TE を変えても QSM 値に大きな変動はなかった．これは結果が定量的であることを示唆した．

② TE>30ms で QSM のノイズが低減した．撮像時間との兼ね合いから TE は最短値 (13.8ms) よりも長めに設定することが望ましい．

これらの傾向は，計算式と一致した．また，得られた QSM 値は 3 T の既報告値と整合した．

P4-20　深層学習を用いたマルチピンホール SPECT システムの画像再構成

○紺野　宏彰[1]，大越　　迪[1]，村田　一心[2]，尾川　浩一[2]

[1]法政大学大学院理工学研究科，[2]法政大学理工学部

本研究では，マルチピンホール SPECT システムの画像再構成における深層学習の有効性を検証した．静止型ピンホール SPECT システムはガンマカメラを回転させずに投影データを収集するため，放射性医薬品の動態情報を取得できる．その画像再構成には通常，最尤推定・期待値最大化法が用いられるが，必要な検出確率を厳密に求めるには現実的でない計算コストがかかる．その補正のため，7rays 法を用いて近似解を求めることが多いが，特にピンホール経が大きいときには画質の劣化が著しい．そこで本研究では，深層学習を用いた画像再構成手法を提案する．実際の医用画像を用いたシミュレーションを行ったところ，高い空間分解能の再構成画像を取得できることがわかった．これは同システムの画像再構成において深層学習が有効であることを示す．

特別講演3

10:40 ～ 11:40　**SL3「宇宙線ピラミッドイメージング」**

座長：森　　健策（名古屋大学）

SL3　宇宙線ピラミッドイメージング

○森島　邦博
名古屋大学大学院理学研究科 特任助教／ JST さきがけ研究者

宇宙線イメージングは，大気上層部で発生する宇宙線中に含まれる素粒子“ミューオン”が持つ極めて高い透過力を利用して巨大な物体内部を X 線レントゲン撮影のように可視化する新しい技術である．私は「原子核乾板」と呼ぶ銀塩写真フィルム型の 3 次元放射線飛跡検出器を用いた技術開発を行い，2015 年，福島第一原発 2 号機内部の炉心溶融の可視化に初めて成功した．2017 年には，4500 年前にエジプトに建造されたクフ王ピラミッドの内部を宇宙線イメージングにより調査した結果，その中心部に未知の巨大な空間を発見した（K.Morishima et al., Nature, 2017）．宇宙線イメージングの適用対象はこれらにとどまらず，社会インフラ構造物の老朽化検査や工業用プラントの内部診断，地質調査・資源探査などへの応用も期待される．講演では，ピラミッド内部の可視化研究を中心に技術と今後の展望について紹介する．

第3日　9月19日（土）

特別企画

12:00 ～ 12:20　**SE「多種多様な情報が紐付いた大規模医療データベース ～次世代医療基盤法によって何が可能となったか～」**

座長：湯浅　哲也
（山形大学／ JAMIT2020 大会長）

SE　多種多様な情報が紐付いた大規模医療データベース
～次世代医療基盤法によって何が可能となったか～

○小松　慶太
内閣官房　健康・医療戦略室　主査　（併）内閣府，経済産業省

医療分野の研究開発においてもビッグデータが活用されるようになって久しいが，これまで画像・検査値・診断名等の多種多様な情報を時系列で紐付けたデータベースを構築することは困難であり，利用者が限定されていることも多かった．

そこで，2018 年，医療分野の研究開発に資するための匿名加工医療情報に関する法律（通称「次世代医療基盤法」）が施行され，時代の要請に応えたデータベースを整備するための環境が整えられた．

昨年から今年にかけて，データベース整備の主体となる法人が複数認定され，近く，利用が開始される見込みとなった．

当該データベースがこれまでのものとどう違うのか，何が可能となったか，今後の展望も含めて明らかにする．

シンポジウム２

12:40 ～ 14:10　SY ２「OCT の技術的進歩と事業化、そして臨床応用の最前線」

座長：秋葉　正博（株式会社トプコン）

時計の針を戻すこと 30 年．光干渉断層計（OCT）の歴史は平成とともに始まりました．1990 年に山形大学米沢キャンパスで産声を上げた OCT が 30 年の時を経て山形の地に戻ってきました．OCT は Metrology として広く活用されてきた低コヒーレンス干渉を生体計測へと応用しました．古くて新しい技術でありますが，その使われ方を“ずらす”ことで新たな産業を産んだのです．今や OCT は診断に活用され，日本では保険適用されており，世界市場は 1,000 億円規模にまで成長しました．OCT はなぜ魔の川・死の谷・ダーウィンの海を渡り，実社会で活用されてきたのでしょうか．今回のシンポジウムではその鍵を握るキーとなる先生方に原理開発及び実用化までの苦労，眼科学におけるパラダイムシフトをご紹介頂くとともに，その将来像もご紹介いただきます．生誕 30 年といえば，実社会では期待の星として位置づけられる時です．現在の OCT は働き盛り．世界各地で人々の健康維持に役立つ装置として活用されております．

SY2-1　座長のことば

○秋葉　正博
株式会社トプコン

SY2-2　第二世代 OCT の開発：基礎研究から市場へ

○安野　嘉晃
筑波大学

第一世代の OCT は 1991 年前後に相次いでその基本原理が示され，眼科臨床検査機器として商業化された．一方，2000 年前後からスペクトル領域干渉計による形状計測の研究が始まる．演者もその研究者の 1 人であり，工業計測応用を目指して研究をすすめていた．一方，国内でも産学官グループが世界市場を目指した第一世代眼底 OCT の研究開発が進められていた．この二つの動きは合流し，第二世代の眼底 OCT の開発が始まり，のちに，3 次元眼底断層装置 3D-OCT 1000 として結実することになる．本講演ではスペクトル干渉原理の発見から製品の上市まで，研究と製品開発の近接に注目しながら概観する．

SY2-3　ダーウィンの荒波にもまれて…眼科用 OCT 事業化と更なる成長

○福間　康文
株式会社トプコン

OCT は日本の山形からの発明でもあるが，海外メーカーがいち早くタイムドメイン OCT を商品化し，臨床エビデンスを揃えて市場優位に立った．これに対し，日本企業は SD-OCT，SS-OCT いう新技術をもって市場参入を果たした．今 OCT は眼科医のみならず，海外ではオプトメトリー（アイケア専門医）にも普及し，AI との組合せでさらなる発展を遂げている．我々は産学連携により，山形で OCT の開発を始め，我が国の OCT の発展と共に歩み，今もこの山形の地で OCT を製造している．本講演では，OCT 市場開発の歴史と今後の展開を紹介する．

SY2-4　眼科臨床における OCT の活用事例

○中澤　　徹
東北大学

緑内障診療において，OCT による画像診断はその有用性の高さ故に広く普及している．OCT は，網膜や視神経の状態を層別に測定・評価することが可能で，緑内障の早期診断，進行判定，予後予測などに活用できる．また，画像・数値の形式でデータが得られる特性を活かし，AI と組み合わせることによって，機能測定などでの活用にも発展していくと考えられている．本講演では，眼科診療の現場における具体的な事例を紹介する．

SY2-5　AI 眼科診断・診療への応用に向けて

○坂下　祐輔
株式会社ニデック

近年の AI 技術による画像認識性能の向上に伴い，眼科診断において AI 技術による支援を行うことで，医師の負担を減らす試みが行われている．しかし，現在主流の人間が作成した正解に基づいて学習を行う AI は，人間の精度が 100% でないのと同様に完璧ではない．その中で，日々の診療業務を AI によりどう効率化していくか？
本講演では，AI をより使いやすく，より確かなものにするための技術を眼底画像と OCT 画像を例に紹介する．

SY2-6　血管内 OCT 技術と臨床での使われ方

○平　　健二
ボストン・サイエンティフィックジャパン株式会社

狭窄した冠動脈内にカテーテルを挿入し血管を広げる低侵襲的な治療法として経皮的冠動脈形成術 (PCI) が行われている．その際に血管内 OCT が用いられている．OCT は優れた空間分解能を持つため狭窄部周辺の内腔径計測やステント留置状態の詳細観察に活用され，またプラークや石灰化等の組織性状の確認などにも使用されている．臨床的知見の蓄積並びに近年の技術的進展により血管内 OCT の更なる発展も期待されている．本講演では血管内 OCT に用いられている技術と臨床での使われ方を紹介する．

第 3 日　9 月 19 日（土）

一般演題セッション５

14:20 ～ 15:50　**P5「CAS・生体情報・その他」**

座長：湯浅　哲也（山形大学）

P5-1　OCT を用いたマウス生殖機能評価に関する検討

○渡部　裕輝[1]，高倉　　啓[1]，黒谷　玲子[1]，阿部　宏之[1]
[1] 山形大学大学院理工学研究科

光干渉断層画像化法 (OCT，Optical Coherence Tomography) は，生体に対して非接触，非侵襲であり，高い空間分解能を有するイメージングシステムである．すでに眼科用医療機器として不可欠な存在となっている．また網膜内の血流を OCT の信号から抽出するドップラー OCT からの派生した OCT Angiography がよく利用されている．我々は，これらの OCT 技術を用いて，新たなに生殖分野での応用を検討している．卵巣内に存在する卵胞の大きさや数を把握することは，卵巣機能の評価する上で極めて重要である．我々は OCT を用いることで，超音波画像では観察できないマウス卵巣に存在する直径 50 m 程度の前胞状卵胞の非侵襲イメージングに成功した．また週齢の異なるマウス精巣の OCT 計測では，ドップラー OCT 画像に精巣網に明らかな違いが確認できた．

P5-2　パルスレーザーダイオードを用いた光音響イメージングシステムの開発

○甲田　梨緒[1]，渡部　裕輝[1]
[1] 山形大学大学院理工学研究科

光音響イメージングは，光学イメージングの分解能と超音波イメージングの測定深度を併せ持った画像化技術で，生体内部の情報を非侵襲に取得する方法として注目されている．
光音響イメージングは，ナノ秒パルス光を照射し，物質の光吸収によって発生した超音波を検出する．一般的に光源は Q スイッチパルスレーザーが使用されるが，装置が大きく高価である．そのため，近年ではレーザーダイオード (LD) や発光ダイオード (LED) によるコンパクト化および低コスト化の研究が行われている．そこで我々は，レーザーダイオードを用いた低コストな光音響イメージングシステムを開発した．このシステムでは，波長 905nm の LD 光をパルス幅 100ns，繰り返し周波数 1000Hz で照射し，中心周波数 5MHz，比帯域幅 61.6％の超音波プローブを用いて光音響画像を取得した．

P5-3 マウス脳 ex vivo イメージングのためのマルチピンホール型蛍光 X 線 CT における入射強度補正

○保科　亮太[1]，金澤　道和[2]，佐々木洸介[2]，笹谷　典太[1]，砂口　尚輝[3]，河嶋　秀和[4]，兵藤　一行[5]，
銭谷　勉[2]，湯浅　哲也[1]

[1] 山形大学　大学院理工学研究科，[2] 弘前大学　大学院理工学研究科，[3] 名古屋大学　大学院医学系研究科，[4] 京都薬科大学　放射性同位元素研究センター，[5] 高エネルギー加速器研究機構　物質構造科学研究所

マルチピンホール型蛍光 X 線 CT（multi-pinhole x-ray fluorescence computed tomography; mp-XFCT）は，入射 X 線により励起された元素から発生する蛍光 X 線を複数のピンホールを用いて効率的に計測することで，対象の 3 次元機能情報を取得する．現行 mp-XFCT システムには，入射 X 線強度分布が一様でないという問題があり，ビーム断面積より大きい対象を分割して撮像した場合，強度ムラの影響で再構成画像が不連続になる．本研究では，あらかじめ計測したビームプロファイルにより入射 X 線強度分布が一様になるように補正する手法を提案する．高エネルギー加速器研究機構 PFAR の偏向電磁石ビームライン ARNE-7A に，ヨウ素から発生する蛍光 X 線を検出する mp-XFCT システムを構築した．非放射性のヨウ素化合物 N-isopropyl-4-iodoamphetamine（iofetamine）を投与後に摘出したマウス脳を，全体を 3 分割して ex vivo 計測したところ，脳全体に分布するヨウ素を滑らかな 3 次元画像として再構成することができた．

P5-4 細胞内構造の深層学習とレディオミクスに基づく癌細胞のラベルフリー分類

○大手　希望[1]，橋本二三生[1]，山田　秀直[1]，山内　豊彦[1]，羽田　綾馬[2]，唐津　周[2]，菊池　寛利[2]，
岡崎　茂俊[1,2]，竹内　裕也[2]，上田　之雄[1]

[1] 浜松ホトニクス株式会社，[2] 浜松医科大学

ラベルフリーの細胞分類は，試薬による細胞の損傷を避けるため重要である．その方法の 1 つとして定量位相顕微鏡（QPM）を用いた細胞分類がある．これまで QPM を用いた細胞内構造に基づく白血球と癌細胞株の 2 クラス分類は検討されていたが，癌細胞株内の多クラス分類は検討されていなかった．そこで，本研究では QPM を用いた細胞内構造に基づく癌細胞株の多クラス分類を試みた．QPM によって撮影された 5 種類の癌細胞株の画像 250 枚を使用した．さらにレディオミクスによって輝度値，形状，テクスチャの特徴を抽出した．画像とレディオミック特徴を学習済みの畳み込みニューラルネットワークである VGG16 に入力しファインチューニングを行った．10 交差検証により精度，AUC を評価し，精度 80.5 %，AUC = 0.957 を得た．これらの結果から QPM を用いた細胞内構造に基づく癌細胞株のラベルフリー分類の可能性が示唆された．

P5-5 手術ロボット開発に向けた犬の避妊手術シミュレーション・データセット作成

○早川　吉彦[1]，久冨木　望[2]，イン　ゼン[2]，パティッタースティージャールワット[3]，長瀬　和弘[3]

[1] 北見工業大学工学部地域未来デザイン工学科機械知能・生体工学コース，[2] 北見工業大学情報システム工学科，[3] 北見工業大学大学院情報システム工学専攻

術者の動作・処置が定型化した手術では，自律型手術ロボットの開発に対してニーズが高い．今回は，小動物・犬を対象に定型的手術である避妊手術（卵巣子宮摘出手術）を採り上げ，シミュレーションモデルを用いて手術に関するデータセットの作成を行った．画像認識技術によって，手術ツールの認識と手術フェーズ・ワークフローの認識を行った．避妊手術モデルを使い，術中時間 15 から 25 分の手術ビデオを 42 本作製した．小動物・犬の手術を対象にデータセット作成を行った．手術ツールの認識では 15 本のビデオにアノテーションをつけた．物体検出を目的としたデータセットでは，長方形の枠線で物体の領域を囲む bounding-box アノテーションが使われる．そこで位置情報の必要性と精度向上のために bounding-box でアノテーションした．手術フェーズ・ワークフローの認識では，6 つのフェーズについて全ビデオにアノテーションした．

P5-6 Admissible SDF variation を用いたヒト胚子の脳の平均形状の推定

○田邊　宥大[1]，斉藤　篤[1]，高桑　徹也[2]，山田　重人[2]，清水　昭伸[1]

[1] 東京農工大学，[2] 京都大学

本報告では，符号付距離関数 (Signed Distance Function, SDF) によって表現された形状の平均を正確に推定する方法について報告する．SDF 同士の線形演算の結果が SDF である保証はなく，従来の SDF の相加平均によって推定した平均では微細な解剖構造が十分表現できない可能性がある．そこで提案手法では，1 次の図形モーメントと SDF 間の L2 ノルムを用いたエネルギー汎関数の最小化によって平均形状を推定する．エネルギー汎関数の最小化の際に admissible SDF variation を用いることで，推定された平均が SDF であることを保証する．最後に，提案手法を京都コレクションの胚子の脳実質の形状ラベルに適用して，従来の手法で推定した平均形状との比較により，提案手法の有用性について議論する．

P5-7　内視鏡と連動した食道手術ナビゲーションシステムの開発

○佐藤由季也[1]，上里　昌也[2]，吉村裕一郎[1]，米本　昇平[2]，浦濱　竜馬[2]，加藤直菜美[2]，中口　俊哉[1]
[1] 千葉大学大学院融合理工学府，[2] 千葉大学医学部　先端応用外科学

近年，消化管外科領域で管腔内視鏡下・腹腔鏡下手術が盛んに行われるなか，より患者の負担を軽減できる方法として 2004 年に胃などの管腔臓器壁を人工的に穿孔させ腹腔内に内視鏡を挿入する NOTES(Natural Orifice Translumenal Surgery) が提案されたが，壁外の把握が困難であることから通常手術として行われていない．本研究では NOTES を安全に実施するため管腔臓器壁外を把握するためのナビゲーションシステムを提案する．事前に撮影された CT 画像から食道壁外の臓器または組織を選択して仮想空間上の解剖モデルを構築する．術中の実内視鏡操作に連動して仮想空間上の内視鏡も移動させることで，内視鏡カメラから観察できる食道の壁外の状況を提示することを目指す．実内視鏡操作によるスコープの挿入深さは，患者の口元にカメラを設置して撮影したスコープの映像情報から得ることとした．本研究では深層学習に基づいてスコープ表面の模様を検出し，その移動量を画像計測により算出する．実験では食道を模した管と，その周囲に臓器を模したファントムを配置し，本システムが提示した仮想映像と実内視鏡が撮影した実験映像の一致度を検証した．実験の結果，映像上での一致度は 86% を示し，スコープの挿入の深さによらず安定した高い精度を実現した．

P5-8　多層構造物測定のための深さ超解像 OCT

○椴山　　誉[1,2,3]，佐々木芳彰[2]，吉峯　　功[2]，永野　繁憲[1]，湯浅　哲也[3]，大谷　知行[2]
[1] 株式会社トプコン，[2] 理化学研究所光量子工学研究センター，[3] 山形大学大学院理工学研究科

光干渉断層計 (OCT) は，非接触，低侵襲で人体器官の三次元画像を取得できることから，主に医療検査技術として応用研究が進んできた．また，OCT はその有用性の高さから医療用途のみならず産業向けの非破壊検査技術としても利用が検討されている．特に，タイムドメイン OCT と比較して SN 比で有利なフーリエドメイン OCT(FD-OCT) は，検査応用で主流な方式となっている．従来の FD-OCT の信号処理では，フーリエ解析を用いて測定した干渉信号を深さ方向の強度プロファイルに変換する．しかし，この従来解析法では光源の帯域幅によって深さ分解能が制限される．本研究では，多層構造物の検査を目的とした FD-OCT 測定における深さ分解能を改善するため，アニヒレーティングフィルタ法をベースにした超解像解析法を適用し，従来のフーリエ解析の約 1/3 の深さ分解能を達成した．

P5-9　脳波信号への深層学習適用による標的刺激の推定

○長島　和馬[1]，松村　泰志[2]，姜　　時友[1]，石川文之進[3]，加我　君孝[4]，深見　忠典[1]
[1] 山形大学大学院　理工学研究科，[2] パクテラ デジタルイノベーション事業開発部，[3] 報徳会宇都宮病院，[4] 東京医療センター・感覚器センター

【目的】深層学習を用いて , 聴覚オドボール課題で計測された単一の脳波応答が , 標的刺激と非標的刺激のいずれであるかの判別を行う .【方法】被験者 76 名に対して , オドボール課題を施行し ,10 ／ 20 国際電極配置法に基づく 19 電極において約 10000 応答を得た . これらの応答のうち 約 8000 応答を訓練データ , 約 2000 応答をテストデータとした . 深層学習において CNN と時間方向と空間方向へのフィルタリングを組み合わせた EEGnet を用い , 正答率の比較を行った .【結果】正答率は CNN で 80.0%,EEGnet で 85.0％であった . これは ,EEGnet がより脳波の処理に適したネットワーク構造となっているためであると考えられる .【結論】単一の脳波応答における分類において ,CNN よりも EEGnet でより高い正答率が得られた .

[Objectives] We used deep learning to determine whether a single EEG response measured in the auditory oddball paradigm is a target or a non-target stimulus. [Methods] The oddball paradigm was performed on 76 subjects, and about 10,000 responses were obtained at 19 electrodes based on the international 10/20 system.About 8,000 responses were used as training data, and about remaining 2,000 responses were used as test data.In the deep learning, the CNN and the EEGnet that combined filtering in the temporal direction and the spatial direction were used to compare the accuracies. [Results] The accuracies were 80.0% for CNN and 85.0% for EEGnet. It is considered that EEGnet has a network structure more suitable for processing EEG. [Conclusion] For classification in a single EEG response, EEGnet showed a higher accuracy than CNN.

P5-10　患者情報の併用による肺細胞診画像の良悪性鑑別 -4 種類の分類アルゴリズムによる性能比較 -

○山田あゆみ[1,2]，寺本　篤司[1]，桐山　諭和[3]，星　　雅人[1]，塚本　徹哉[3]，今泉　和良[3]，藤田　広志[4]，齋藤　邦明[1,2]
[1] 藤田医科大学　医療科学部，[2] 藤田医科大学大学院　保健学研究科，[3] 藤田医科大学　医学部，[4] 岐阜大学　工学部

肺細胞診における良悪性鑑別は，病理医や細胞検査士にとって煩雑で労力を必要とする作業である．ここで，実際の診療プロセスにおいて，術者は細胞の形状変化のみならず種々の患者情報を参照して判断することが一般的である．本研究では液状細胞診 (LBC) 画像と患者情報を用いた肺癌良悪性自動鑑別手法を開発し，基礎評価を行った．はじめに，Convolutional Neural Network のひとつである VGG-16 を用いて LBC 画像から 4096 個の画像特徴量を抽出した．次に，電子カルテから LBC 画像に対応する患者情報 (年齢，性別，臨床検査結果等) を 27 種類収集した．得られた画像特徴量と患者情報から特徴選択を行い，機械学習を用いて良悪性の分類を行った．45 症例の臨床データを用いて , 3 分割交差検証にて評価を行った結果 , 患者情報単体 , 画像情報単体での識別結果と比較して，それらを組み合わせて分類したときに識別精度が向上した．これらの結果から，提案手法の有用性が示唆された．

P5-11 乳がん手術支援のための乳房追跡による MRI 画像重畳の検討

○小口　貴也[1]，藤本　浩司[2]，吉村裕一郎[3]，林　　秀樹[3]，中口　俊哉[3]

[1] 千葉大学融合理工学府基幹工学専攻医工学コース，[2] 千葉大学医学部付属病院乳腺・甲状腺外科，[3] 千葉大学フロンティア医工学センター

乳がんの腫瘍切除範囲の設定は，MRI 画像上における腫瘍の位置を術者が頭の中で実際の患者体位へと置換して行われる．先行研究では腫瘍部を視覚化するために術前に手術体位で撮影した MRI 画像を患者の体表に投影する手法が提案された．しかし，この手法は投影像と体表の位置合わせは術者の経験に依る部分が大きい．この問題を克服するために，術前の MRI 画像から腫瘍の３次元モデルを作成し，タブレット端末のカメラ画像上で患者体表に重畳するシステムを考案した．提案システムでは最初 MRI 画像上とカメラ画像上からそれぞれ４点を対応点として手動で選択し，以後は選択点のトラッキングにより重畳のための位置合わせを自動で行う．その際，加速度センサから取得したタブレット端末の姿勢を考慮することでより正確な重畳を実現した．重畳精度を評価するためファントムを用いた実験を行った．この実験より，３次元モデルが十分な精度で重畳されていることが示唆された．

P5-12 マルチタッチ LCD により直感的操作が可能な医用画像ディスプレイシステムの開発

○増谷　佳孝[1]，出羽　貴之[2]，藤川　拓馬[2]，見澤　勇太[2]

[1] 広島市立大学 大学院情報科学研究科，[2] 広島市立大学 情報科学部

我々はマルチタッチ液晶ディスプレイを用いた医用画像表示システムの開発を行ってきた．主要な機能としてボリュームデータの任意断面像および投影像の表示があり，さらに１～１０点のタッチ数に応じた直感的な操作による断面の回転や拡大･縮小，投影の深さ変更などが可能で，さらに対話的なインタフェースにより精密な設定も可能となっている．本報告では主要機能の紹介に加え，タッチ数やアクションの認識に関する実験，および任意断面検索の迅速性を評価する実験について述べる．実験の結果，タッチ数およびアクションの認識に関する最適パラメタの決定を行い，任意断面を迅速に表示できることが示され，本システムの有効性が示唆された．将来的には，別システムとして開発してきたアノテーション機能を統合し，医用画像処理，特に機械学習によるセグメンテーションに対する強力な支援システムとなることを目指す．

P5-13 Image analysis for automatic detection of social interaction in drosophila: biased courtship behavior

○渡邉　裕貴[1]，成　　耆鉉[2]，松村　泰志[3]，姜　　時友[1]

[1] 山形大学大学院・理工学研究科・応用生命システム工学専攻，[2] 理化学研究所筑波地区・CPR・眞貝細胞記憶研究室，[3] パクテラ・デジタル・イノベーション事業開発部

Courtship behavior is an important life-event in order to mating and surviving the species. In drosophila known as fruit fly and typically used in genetic research field, courtship behavior from male to female has been examined, however, the ones from female to male is still unknown because it is weak and un-observable under normal condition. To capture the female courtship behavior, a free-moving field with male and female targets is designed. Each well has three individuals, one is a freely-moving female drosophila before eclosion, and the others are immobilized male and female as targets of courtship behavior of freely-moving female. Behaviors of a freely-moving female drosophila was recorded as movie and captured through image processing as follows: First, foreground, that is, drosophila was extracted from background. Second, thresholding and morphological filters were applied onto extracted the foreground binary images for noise reduction. Third, numbering and elimination were performed and central gravity was calculated in order to obtain a moving trajectory. After that, approaching of the freely-moving female drosophila to the immobilized target animals was quantified. As a result, it was revealed that frequency of approach to female or male showed significant biases. Present research shed light to courtship behavior of female drosophila.

P5-14 大腸内視鏡のための教師なし深度画像推定法における補助タスク検討

○伊東　隼人[1]，小田　昌宏[1]，堀田　欣一[2]，森　　悠一[3]，三澤　将史[3]，高畠　博嗣[4]，森　　雅樹[5]，名取　　博[6]，工藤　進英[3]，森　　健策[1,7,8]

[1] 名古屋大学 大学院情報学研究科，[2] 静岡県立静岡がんセンター 内視鏡科，[3] 昭和大学横浜市北部病院 消化器センター，[4] 札幌南三条病院，[5] 札幌厚生病院，[6] 恵和会西岡病院，[7] 情報基盤センター，[8] 国立情報学研究所 医療ビッグデータ研究センター

医用画像処理および医用画像理解においては３次元情報の抽出が非常に重要である．一方で，大腸内視鏡診断では２次元画像もしくはそれらの時系列集合しか得られない．大腸内視鏡検査におけるコンピュータ支援診断システムを構築する上で，単眼画像から３次元情報を抽出・推定する技術が求められている．本稿では教師なし深層学習を用いた深度推定の高精度化を目的に，深度推定における補助タスクの有用性を検討する．

P5-15 位置特徴量とマルチスケール特徴量による胃壁マイクロ CT 像からの胃壁の層構造及び腫瘍の抽出

○御手洗　翠[1]，小田　紘久[1]，杉野　貴明[1]，守谷　享泰[1]，伊東　隼人[1]，小田　昌宏[1]，小宮山琢真[2]，古川　和宏[2]，宮原　良二[2]，藤城　光弘[2]，森　雅樹[3]，高畠　博嗣[4]，名取　博[5]，森　健作[6,7]

[1] 名古屋大学大学院情報学研究科，[2] 名古屋大学大学院医学系研究科，[3] 札幌厚生病院，[4] 札幌南三条病院，[5] 恵和会西岡病院，[6] 名古屋大学情報基盤センター，[7] 国立情報学研究所医療ビッグデータ研究センター

本稿では，Spherical K-means (SpK) により抽出された複数スケールの特徴量と位置特徴量を用いた胃壁 μ CT 画像からの粘膜層，粘膜下層，筋層及び腫瘍の抽出手法について報告する．胃壁の μ CT 画像から解剖学的構造を抽出することで腫瘍の立体構造把握が可能である．従来手法では複数スケールの特徴抽出フィルタを SpK によって学習し，パッチから特徴抽出を行った．最後に得られた特徴量に基づいて SVM によって分類した．しかし，SpK により生成された特徴抽出フィルタは濃度値に基づく特徴量を抽出するため，粘膜層と筋層のようにコントラストが低い領域の特徴量が近しい値になる．本手法では SpK により得られる複数スケールの特徴量にパッチの位置特徴量を追加して分類する事によって，濃度値に基づく特徴量では分類できない領域の分類を可能にする．パッチの位置特徴量としてパッチの中心座標を用いる．本手法を胃壁 μ CT 像に適用した結果，良好な結果が得られた．

P5-16 屈折コントラスト X 線ＣＴを用いた非浸潤性乳管癌と通常型乳管過形成の腺腔における 3 次元的な差異

○谷口　華奈[1]，砂口　尚輝[1]，市原　周[2]，西村理恵子[2]，岩越　朱里[2]，島雄　大介[3]，湯浅　哲也[4]，安藤　正海[5]

[1] 名古屋大学大学院 医学系研究科，[2] 名古屋医療センター 病理診断科，[3] 北海道科学大学 保健医療学部，[4] 山形大学大学院 理工学研究科，[5] 総合科学研究機構

乳癌など乳腺疾患の病理診断は , 2 次元である組織切片の顕微鏡観察により行われている . 腺腔 (病変内の微小な穴) を含む組織構造の特徴は重要な指標であるが , 従来の手法では腺腔の 3 次元構造観察は難しく , それに関連する研究も少ない．本研究では , 腺腔の詳細な 3 次元構造を得るために，染色した組織像と同等のコントラストが得られる屈折コントラスト X 線 CT を用いて，非浸潤性乳管癌 (DCIS) と良性病変である通常型乳管過形成 (UDH) を含む乳腺組織を撮像した . 再構成像から腺腔領域を抽出し 3 次元解析を行った結果 , DCIS ではバブル状 , UDH では管状に近い腺腔構造が多数確認された . また , 腺腔密度や腺腔構造の経路長といった差異を数値で表す新しいパラメータを導入した .DCIS は転移能を有する浸潤癌に進行するリスクがあり , 腺腔の 3 次元構造の差異を詳細に解析することで浸潤リスクを予知できる可能性がある .

P5-17 Visual SLAM を用いた内視鏡位置推定に関する基礎検討

○上野　創史[1]，藤橋　卓也[1]，菊川　忠彦[2]，猿渡　俊介[1]，安藤英由樹[3]，雑賀　隆史[2]，渡辺　尚[1]

[1] 大阪大学大学院情報科学研究科，[2] 愛媛大学医学系研究科，[3] 大阪芸術大学

映像データを用いた内視鏡手術の遠隔指導は指導する熟練外科医・指導を受ける外科医間の物理的制約を緩和する一方，内視鏡および鉗子の細かな操作技術を熟練外科医が映像データのみから適切に指導することは困難である．内視鏡および鉗子にセンサを取り付けて 3 次元空間における内視鏡・鉗子の位置姿勢を取得する方法も提案されているが，内視鏡に対する新たなセンサの導入はコストの増大や内視鏡・鉗子操作に影響を及ぼす．新たなセンサを導入することなく内視鏡の位置姿勢を推定するため，本研究では内視鏡カメラで撮影した映像のみをもとにその位置・姿勢を取得する手法を提案する．本手法では，映像を入力としてカメラの位置姿勢を推定する Visual SLAM を用いて内視鏡の位置および姿勢をリアルタイムで取得する．単眼の内視鏡カメラを用いた実験評価から提案手法による内視鏡位置の推定精度は平均 3.6cm 程度の誤差に収まることを確認した．

P5-18 大規模放射線読影レポートデータベースによる BERT モデルの事前学習とそれを用いた CT 画像の撮影目的の推定

○本田　修平[1]，大竹　義人[1,3]，高尾　正樹[2]，荒牧　英治[1]，矢田竣太郎[1]，合田　憲人[3]，佐藤　真一[3]，橋本　正弘[4]，明石　敏昭[5]，菅野　伸彦[2]，佐藤　嘉伸[1]

[1] 奈良先端科学技術大学院大学，[2] 大阪大学，[3] 国立情報学研究所，[4] 慶應義塾大学，[5] 順天堂大学

我々の研究グループでは，CNN を用いた筋骨格解剖パラメータの自動計測システムを開発し，大規模 CT 画像データベースを用いて骨盤傾斜角・筋体積などの解剖学パラメータの年齢・性別による変化を解析してきた．本研究では，解析結果から得られる医学的知見を増強するため，各 CT 画像に紐づけされた放射線読影レポートを用いて，CT が撮影された撮影目的（疾患背景）の推定を試みた．本研究で用いた BERT モデルは，広く使われている Wikipedia を用いた pre-trained モデルではなく，約 32 万件の放射線読影レポートから成る大規模なコーパスで pre-training を行い，12 クラスの撮影目的分類タスクを 299 症例の専門家による手動分類データを用いて fine-tuning した．これにより，BiLSTM による従来手法より 11.3%，Wikipedia コーパスによる pre-training モデルより 4.0% の予測精度向上を達成した．さらに，モデルの解釈と予測精度向上のため，Attention 層の可視化と temperature scaling を組み込んだネットワークの較正にも取り組んだので報告する．

P5-19 コンピュータシミュレーションで得られた最適ポータル位置のファントムによる検証

○川上申之介[1]，福田　紀生[2]，西井　　孝[3]，大竹　義人[4]，佐藤　嘉伸[4]

[1]近畿大学工業高等専門学校，[2]脳情報通信融合研究センター，[3]大阪急性期・総合医療センター整形外科，[4]奈良先端科学技術大学院大学

損傷した股関節唇の修復で行われる低侵襲整形外科手術においては，皮膚に開けた直径数ミリのポータルと呼ばれる挿入口から股関節鏡を挿入して得られる視野を頼りに狭い空間での処置が行われる．股関節鏡の視野内に適切に股関節唇を捉えるためにはポータルの位置が重要であるが，代表的なポータル位置決定手法により決めた位置から挿入した場合に良い視認性が得られるかどうかは患者の股関節の構造による．そこで我々は股関節の３Ｄサーフェスモデル上にポータル位置を決めると自動的にその位置とそこからの挿入方向を網羅的に変えながら股関節鏡を挿入して画像を生成し，その患者の股関節構造に基づいた手術に最適なポータル位置を推定できるシミュレータを開発した．本報告では，ファントムから作成した３Ｄサーフェスモデルに対してシミュレータを実行し推定した最適ポータル位置から実際に股関節鏡を挿入し，得られた視認性の検証結果について報告する．

P5-20 CT 画像から腹部臓器領域の自動抽出のための大規模データベース構築に向けた不確実性の推定に基づくアクティブラーニングの検証

○政木　勇人[1]，大竹　義人[2]，スーフィーマーゼン[2]，堀　　雅敏[3]，大西　裕満[4]，富山　憲幸[4]，佐藤　嘉伸[2]

[1]PSP 株式会社，[2]奈良先端科学技術大学院大学，[3]神戸大学大学院医学研究科，[4]大阪大学大学院医学系研究科

堅牢な腹部臓器の自動抽出モデルを構築するためには，大規模なデータセットが必要であるが，アノテーションのコストが非常に大きく非常に困難である．それに対し我々のグループでは，下肢領域の CT 画像においてモデルの不確実性の推定に基づく下肢筋骨格自動抽出のためのアクティブラーニング手法を提案したが，シミュレーション環境下での評価にとどまっていた．本研究では，多時相撮影された数 1000 人規模の CT 画像に対して腹部臓器領域を自動抽出し，推定された不確実性の調査を行った．さらに提案されたアクティブラーニング手法の有用性を検証する実験をしたのでそれらの結果を報告する．

P5-21 空間情報を考慮した自己教師あり学習を用いたステレオ腹腔鏡画像からの深度推定

○李　　文達[1]，林　雄一郎[1]，小田　昌宏[1]，北坂　孝幸[2]，三澤　一成[3]，森　　健策[1,4,5]

[1]名古屋大学 大学院情報学研究科，[2]愛知工業大学情報科学部，[3]愛知県がんセンター，[4]名古屋大学 情報基盤センター，[5]国立情報学研究所 医療ビッグデータ研究センター

Depth estimation plays an important role in robotic surgical navigation systems. There are many studies on depth estimation from laparoscopic images. In recent years, self-supervised siamese network have been proposed for depth estimation from stereo laparoscopic images. In this method, some of spatial information in the image will be lost as a result of continuous convolution and pooling. In order to solve this problem, we add contextual encoding module and multi-scale space merging module to original method. We also change the basic module to U-Net for feature encoding and feature decoding. The context encoding module and multi-scale spatial merging module is formed by dense atrous convolution block and spatial pyramid pooling block. Finally, we output disparity predictions by fusing multi-scales decoder outputs in the decoder blocks. After comparing with the original method, the experimental results showed that the proposed method improved depth estimation results.

P5-22 CNN を用いた指の運動想起脳波の識別

○加藤　　将[1]，叶賀　　卓[2]，長島　和馬[1]，星野　貴之[2]，深見　忠典[1]

[1]山形大学大学院，[2]産業技術総合研究所

様々な動作の想起により生じる脱同期のパターンを脳波から識別する技術は，ブレインコンピュータインタフェース (BCI) への応用が期待されている．なぜなら，この技術を用いた BCI は使用できる場面・タイミングの自由度が高いためである．しかしながら，現存の識別技術は対象である脳波データが，大きい動作 (例えば肘関節の屈曲や肩関節の外転) の想起時に有効である．言い換えると，指の動きのような小さい動作想起時の脳波データを識別することは難しい．本稿では，ダイナミクスの小さい指動作想起時の脳波データからでも識別力の高い脳波特徴を表現し，高精度に識別できる技術を提案する．このために，畳み込みニューラルネットワーク (CNN) を使用する．結果として，CNN による識別精度は各指の屈曲動作の想起 (5 クラス) 時の脳波データを 59.8% で識別でき，一般的な空間フィルタとサポートベクタマシンの組み合わせの手法より高精度な精度を示した．

Machine learning-based techniques to discriminate electroencephalogram (EEG) patterns of desynchronization caused by motion imageries can be applied to brain—computer interfaces (BCIs). Motor imagery-based BCIs are not phase-locked: thus, they have a high degree of flexibility in terms of the scenarios and timings in which they can be used. However, traditional techniques are effective when the target EEG data is recorded during imaging large-scale motions (e.g., elbow flexion or shoulder abduction). In other words, it is still difficult to discriminate EEG data during small-scale motion such as finger motions. To improve the usability of motor imagery-based BCIs, we propose a technique to learn discriminative features from EEG data during imaging finger motions and to classify such features with high classification accuracy. For this purpose, we applied convolutional neural networks to EEG data of imaging each finger motion. The classification accuracy by proposed technique was 59.8% for the 5-class EEG data, which was higher than the traditional technique of combining spatial filters and support vector machines.

P5-23 腹腔鏡手術動画像データベース構築に向けたリモートアノテーションツールのプロトタイプ開発

○屠　芸豪[1]，伊東　隼人[1]，小澤　卓也[1]，小田　昌宏[1]，竹下　修由[2]，伊藤　雅昭[2]，森　健策[1,3,4]

[1]名古屋大学 大学院情報学研究科，[2]国立がん研究センター東病院 大腸外科，[3]名古屋大学 情報基盤センター，[4]国立情報学研究所 医療ビッグデータ研究センター

本稿では腹腔鏡手術動画像における術具や解剖構造をウェブベースでアノテーションするためのリモートアノテーションツールの開発について報告する．腹腔鏡手術動画像において術具や解剖構造のアノテーションデータを作成することにより，機械学習を利用した画像処理によって手技解析法や手術支援技術の開発が可能となる．先行研究として我々はアノテーションのための Windows アプリケーションを開発した．しかし大規模データベースの構築においては，病院へ専門的な機材やソフトウェアを導入するコストが高い．加えて，データの外部への持ち出しによる流出が懸念される．そこで，開発したソフトウェアをウェブブラウザ上で動作するように拡張することにより，病院へのシステム導入のコスト削減，web を介したアノテーション作業によって作業の効率化を図る．

P5-24 肺がん病理画像におけるベイズ U-Net に基づく不確実性を用いた細胞核自動セグメンテーション精度予測

○スーフィーマーゼン[1]，大竹　義人[1]，永島　慧法[2]，明神　大也[3]，大林　千穂[3]，佐藤　嘉伸[1]

[1]奈良先端科学技術大学院大学，[2]奈良県立医科大学 医学部，[3]奈良県立医科大学 病理診断学講座

Nucleus segmentation is indispensable in accelerating the processing pipeline of high-resolution pathological images, i.e. whole slide images (WSIs). Convolutional neural networks, such as U-Net, have demonstrated potential in nucleus segmentation tasks. However, given the high resolution of WSIs, creating ground-truth labels from the whole image is a non-trivial task. Therefore, a predictive measure of the segmentation accuracy in the absence of ground-truth is needed. In this study, we investigated the potential of Bayesian U-Net in predicting the segmentation accuracy of nuclei in pathological images based on uncertainty estimation. The basic idea was to model the uncertainty based on dropout training, and inferring the target labels using Monte Carlo dropout sampling at testing time. A Bayesian U-Net of 5-layers with dropout (ratio= 0.5) was used. The nucleus was segmented by averaging the softmax probabilities of the prediction in 10 samples at testing time, while the uncertainty was modeled as the area variation coefficient (AVC) of the segmented nucleus areas (in pixels) in the 10 samples. To evaluate the performance, leave-one-patient-out cross-validation was performed on patches (n= 20,069; matrix size= 128 ☒ 128 pixels) sampled from WSIs of 11 lung cancer patients. The segmentation accuracy was assessed using Dice coefficient. The correlation between the predictive uncertainty and Dice coefficient was investigated. A conventional (i.e. deterministic) U-Net was used as baseline. The accuracy of the Bayesian and baseline U-Nets were 0.909 ± 0.079 and 0.899 ± 0.101, respectively ($P < 0.001$; Mann-Whitney U-Test). The correlation coefficient between the AVC and Dice coefficient for tumor and non-tumor nuclei were -0.664 and -0.610, respectively. This showed the feasibility of Bayesian U-Net in predicting the segmentation accuracy of the nuclei in pathological images based on uncertainty estimation.

以降の予稿は，著者から提出された原稿を査読なしで，
そのまま掲載しています．

著者らが自らの内容を MEDICAL IMAGING TECHNOLOGY（MIT）誌へ投稿する場合，その原稿にこの PDF に類似した内容が含まれていても，二重投稿とは判断しません．

発行日：２０２０年９月１６日

発行者：湯浅哲也

製作所：トーヨー企画株式会社

単純腹部 CT 画像からの高精細な大腸領域検出法の検討

安倍　和弥[*1]　武尾　英哉[*1]　永井　優一[*2]　縄野　繁[*3]

要旨

現在，医療の様々な分野において画像支援診断（CAD : Computer Aided Diagnosis）の研究・開発が行われている．大腸は周辺臓器との分離の難しさから CT を用いての CAD 開発はあまり行われてこなかった．最近では，内視鏡検査の高精度化や CT コロノグラフィー検査の普及などから大腸がんの発見などに CAD の研究が行われている．しかし，内視鏡検査や CT コロノグラフィー検査は事前準備が必要であり，特に CT コロノグラフィーにおいては炭酸ガスの注入による大腸の拡張を行う点から一部の症例では大腸に負担がかかるため検査自体が行いにくいなどの問題もある．

本研究では，腹部検診時に撮影する単純腹部 CT 画像から高精細に大腸領域を検出する手法についての研究・開発を行った．これにより，大腸に負担を与える CT コロノグラフィー検査などを行わずに CAD 開発に用いる画像の取得が行える．本手法での検出結果と手動での検出結果の一致率は 71%となり，本手法での検出性能が確認できた．

キーワード：大腸，領域抽出，単純腹部 CT 画像，CAD

1. はじめに

近年，医療の現場では CT や MRI など撮影装置の高精細化により生成される医用画像の量が増大している．撮影される画像の枚数は患者一人あたり数百枚にもおよび，それを読影する医師の負担も急増している．そのため診断の補助を行うコンピュータ画像支援診断（Computer Aided Diagnosis/Detection : CAD）の開発が非常に重要視されている[1].

しかし，単純腹部 CT 画像を用いた大腸領域検出 CAD の開発は，周辺臓器や体表・体内の脂肪領域との癒着などからあまり行われてこなかった．そのため大腸検査においては，内視鏡を用いる検査が一般的であった．近年では腸内にガスを注入し，大腸を膨張させる形で撮影を行う CT コロノグラフィーも精密検査に用いられるようになった．しかし，内視鏡検査や CT コロノグラフィー検査は事前準備が必要であり，特に CT コロノグラフィーにおいては炭酸ガスの注入による大腸の拡張を行う点から一部の症例では大腸に負担がかかるため検査自体が行いにくいなどの問題もある．

本研究では，腹部検診時に撮影する単純腹部 CT 画像から高精細に大腸領域を検出する手法についての研究・開発を行う．これにより，大腸に負担を与える CT コロノグラフィー検査などを行わずに CAD 開発に用いる画像の取得が行える．これにより，従来違和感などがあってから検査を行うことの多い大腸の病変に対し，腹部検診など大腸を主目的としない検査において病変の早期発見の可能性が考えられる．

CT コロノグラフィーを用いた大腸の検出は，Wyatt らの手法[2]や Näppi らの手法[3]に代表される通りのガスにより膨張させた大腸壁を検出する手法が主に活用されている．しかし，本手法の利点は前述の通り腸内の洗浄やガス注入が不要でありながら大腸検出が行える簡便さにある．

*1 神奈川工科大学工学部電気電子情報工学科
〔〒243-0292 厚木市下荻野 1030〕
e-mail: abex0930@ele.kanagawa-it.ac.jp
*2 国立がん研究センター東病院
*3 国際医療福祉大学三田病院

大腸・小腸を除いた腹部臓器の検出については様々な手法が確立されている．例えば藤本らの手法[4]のような，目標とする腹部臓器の認識率が 90%を超える性能をもつものや，Park らの手法[5]や清水ら[6]の確率アトラスを用いた臓器モデルを用いて複数臓器を同時に検出できるものなどがあるが，大腸・小腸については単純 CT 画像からの検出はできていない．

本研究では，開発した多段階式の領域抽出手法にて困難である単純腹部 CT 画像からの大腸領域の抽出を目的とする．

本論文のポイントを以下に示す．

① 単純腹部 CT 画像からの大腸領域抽出の提案

② 多段階式領域抽出による精密な大腸領域抽出手法の開発

本論文では，2 章で単純腹部 CT 画像からの大腸領域抽出の手法について，3 章で提案手法での抽出結果を示す．4 章にて考察，5 章にて本研究のまとめについて述べる．

2. 単純腹部 CT 画像からの大腸領域抽出

2.1 大腸領域抽出の流れ

本研究における大腸領域抽出の流れを図 1 に示す．

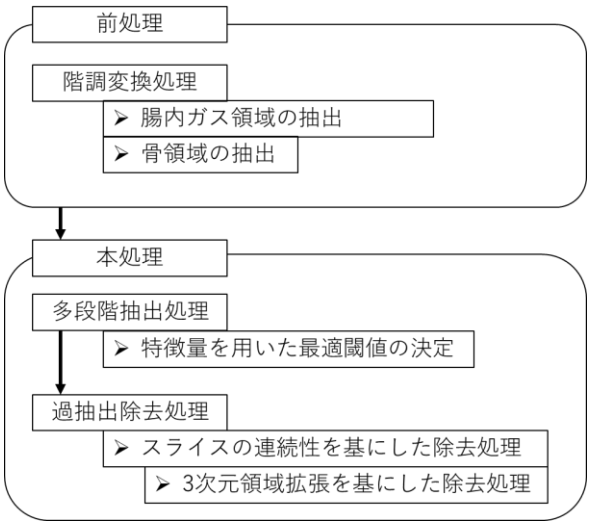


図 1 大腸領域抽出の流れ

2.2 前処理

前処理として，初めに階調変換処理を行い大腸の強調を行う．本研究では，ウィンドウレベルは腹部で一般的は 50HU，ウィンドウ幅は後のガス検出を考慮して 400HU として階調変換を行った．ウィンドウ幅を狭くすれば大腸や近しい CT 値の臓器が強調されるが，ガス領域と体内領域が等しく低い値に抑えられてしまい分離が困難となったため広いウィンドウ幅を設定した．

次に骨領域の検出と腸内のガス領域の検出を行う．

後の閾値を用いた多段階式抽出の際に背骨や骨盤および一部の肋骨近辺に誤抽出が発生しやすいため，あらかじめ骨領域の抽出を行い特徴量として利用する．図 2 に骨領域抽出の 1 例を示す．



図 2 骨領域の抽出画像

また，大腸内にはまばらにガスがたまっており，これを精密検出の際の特徴量の一つとして用いるためあらかじめ検出を行う．図 3 にガス領域抽出の 1 例を示す．

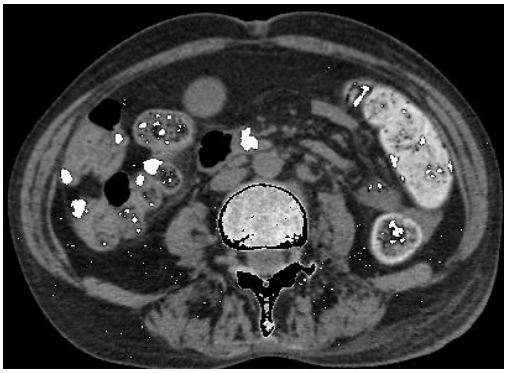
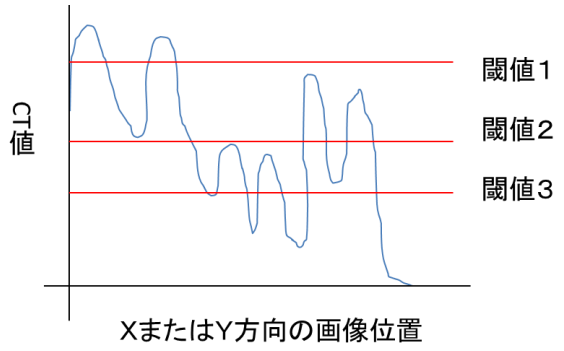


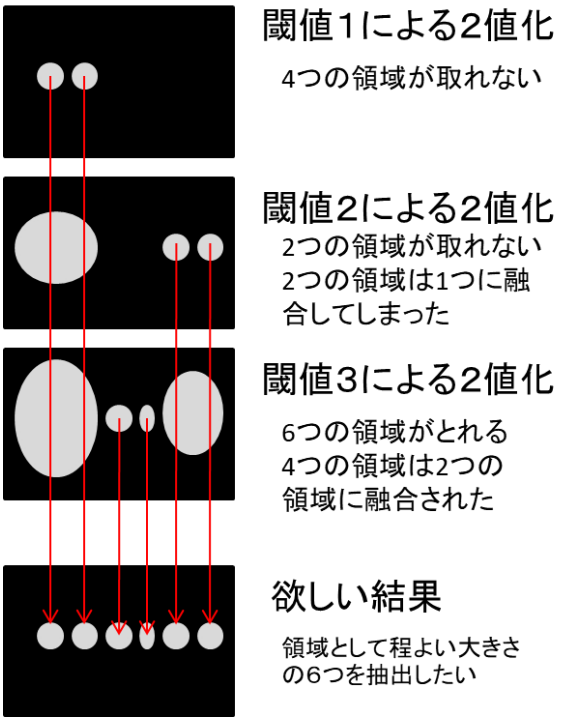
図 3 腸内ガス領域の抽出画像

2.3　多段階式領域抽出による大腸領域抽出手法

今回大腸抽出に用いる多段階式抽出手法は，本研究室の堂之前らが持つ特許の手法[7]を応用し，閾値を高い値から徐々に低い値に変動させながら 2 値化を行い設定した特徴を基に最適形状を検出する手法である．図 4 を基に概要を説明する．



(a)　画像の CT 値分布の例



(b)　閾値ごとの抽出結果と理想の抽出

図 4　段階的領域抽出の概要

図 4(a)のような CT 値を持った画像に対し単純に閾値を用いると図 4(b)の上 3 枚のようにうまく抽出が行えない．本当に抽出したい領域は図 4(b)の下画像である．そこで，図 5 に示すようにあらかじめ大きく設定した閾値を総当り的に減衰させながら抽出を行い，特徴量条件を満たすもののみを検出する．それにより癒着した領域を細かく分離して抽出することが可能となる．

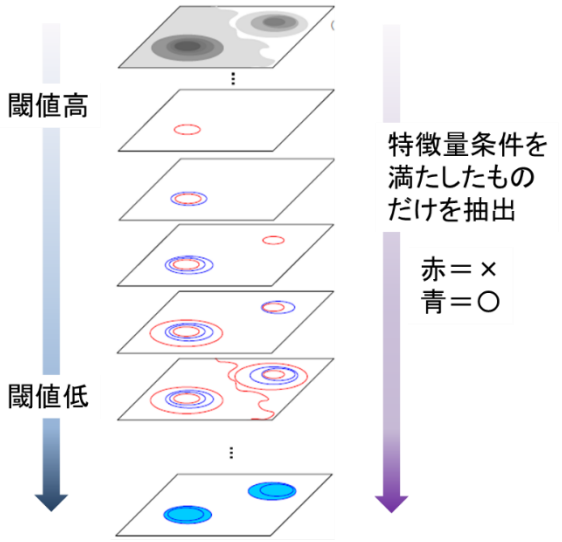


図 5　段階的抽出法

上記で行った階調変換により -350HU～450HU の幅にて階調変換を行った画像に対し，閾値の初期値を 90HU とし，10HU ずつ減算していき 0HU まで処理を行った．図 6(a)(b)に 90HU と 0HU の時の抽出画像を示す．



(a)　90HU　　(b)　0HU

図 6　多段階式抽出の中間画像

個々の抽出領域に対しシード点を個別に設定[8]，閾値変動による抽出領域のずれを抑制する．その各領域に対し，腸内ガス領域の抽出結果や骨領域の情報，円形度や形状特徴，重心といった特徴量を用いて最適閾値を決定する．ス

ライス上の複数個所に大腸が存在するような場合もあるため，最適閾値は抽出領域ごとに個別に設定する．これにより大腸領域の精密抽出が可能となる．

2.4　過抽出領域の除去

前項で述べた処理により大腸の抽出は行えるが，やはり過抽出も起こる．それに対し，スライスの連続性を基にした除去処理と，3 次元領域拡張による除去を行う．

まず，スライスの連続性での除去について述べる．前項での抽出処理では，あくまでもスライスごとの処理であるためスライス間での情報は引き継がれない．本手法では，対象スライスと前後 2 スライスの計 5 スライスより補間を行う．対象スライスに強く，離れるごとに弱く重みを与えて画像を合成，評価画像を作成する(図 7)．複数枚にて大腸として抽出されたものは画素値が大きくなり，まれにある脂肪領域などを過抽出したものは画素値が小さくなる．閾値を下回るものを除去することにより突発的に発生した過抽出が抑制できる．

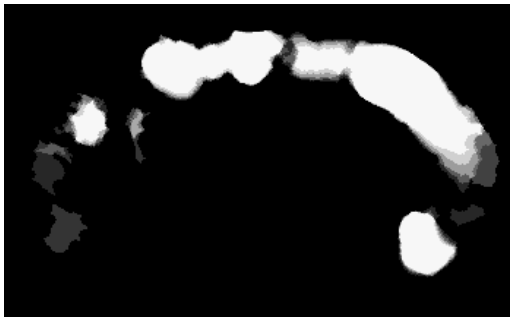
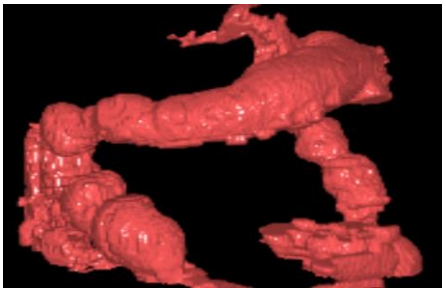


図 7　連続性を基にした除去の評価画像

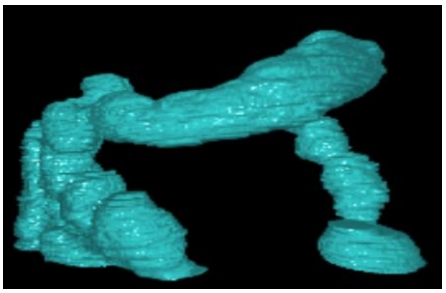
次に 3 次元領域拡張による除去を行う．先の連続性を用いる処理で突発的に発生した過抽出などは除去できるが，小腸の一部や連続性のある過抽出などは除去できないことがある．それらに対し 3 次元領域拡張による抽出を行う．3 次元では大腸はチューブ状に抽出されるため，チューブ状に連結しているもののみを抽出することにより不要部の除去が行える．

3. 本手法による検出結果

本手法での大腸領域の検出結果を図 8(a)に，手動での抽出結果を図 8(b)に示す．また一致度判定のために手動抽出の結果と合成した画像を図 9 に示す．



(a)　本処理での抽出結果



(b)　手動での抽出結果

図 8　大腸抽出結果の 3D 表示

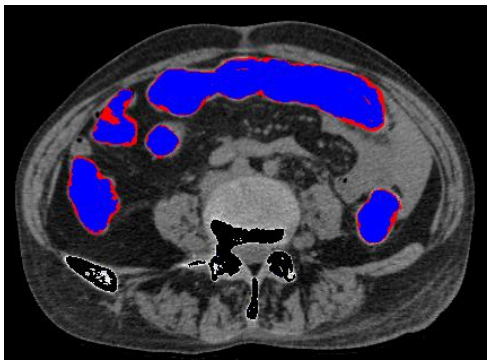


図 9　手動抽出との比較
(青が一致している箇所，赤が不一致の箇所)

図 9 の画像において，一致した領域は青，不一致の領域は赤で示している．手動検出に比べてやや抽出範囲が小さくはなったが，ほぼ一致した抽出結果となった．

検出結果の一致率は 71%となった．手動検出より辺縁が小さくとれる点，および手動では大きいガス領域も含めて抽出した点などが差に

なったと考えられる．

誤検出に関しては，体表部の脂肪領域と小腸，大腸が他臓器と癒着している部位での閾値設定の甘さなどによるものが見られる．

4. 考察

3 章の抽出結果の画像より，単純腹部 CT 画像より大腸領域の抽出が行えていることが確認できる．抽出した大腸領域の一致率は 71％となり高い抽出性能が確認できる．また，2.4 節で示した過抽出抑制により小腸や脂肪領域誤抽出が抑制でき，結果抽出性能が向上したと考えられる．しかし，スライス間の補間を利用した誤抽出除去処方により一部の大腸領域が削られてしまうものなどもあり，より高度な除去手法を検討する必要がある．また，3 次元領域拡張による除去処理では，連結できなかった極小の大腸領域を除去してしまうなどの課題もある．

本手法は，単純腹部 CT 画像から大腸の抽出を行っているが，今回使用した対象に大腸がんの症例はない．しかし，本手法で高い抽出結果が得られたことで，本手法は単純腹部 CT 画像を用いた大腸がん検出手法への応用などを期待できると考えられる．

5. まとめ

本論では，単純腹部 CT 画像を用いた大腸領域の抽出手法の提案を行った．多段階式の抽出手法により，周辺臓器とのコントラストに差が出にくい大腸を精密に抽出することが可能となった．本処理の検出結果と，手動での検出結果を比較したところ一致率は 71％となり，本手法での抽出性能が確認できた．

利益相反の有無

利益相反　なし

倫理規範の順守

本研究にて使用した患者の症例データは，すべての患者の初診時に撮影画像の研究利用も含めた「包括的合意」にサインを受け，また各病院にて倫理委員会を通したうえで提供を受けたものを使用している．

文　献

［1］ 土井邦雄：医用画像とコンピュータ支援診断 ～現状と将来の可能性～．映像情報メディア学会誌, 65(4), pp.427-431, 2011

［2］ Wyatt CL., Ge Y., Vining DJ.: Automatic segmentation of the colon for virtual colonoscopy. Comput Med Imaging Graph, Vol.24(l), pp.1−9, 2000

［3］ Näppi J1., Dachman AH., MacEneaney P., et al: Automated knowledge-guided segmentation of colonic walls for computerized detection of polyps in CT colonography. Journal of Computer Assisted Tomography, 26(4), pp.493-504, 2002

［4］ 藤本秀行， 顧力栩， 金子豊久：3 次元モフォロジー法による腹部臓器の認識．IEICE, Vol.J84−D2(5)，pp.843−850，2001

［5］ Park H., Bland PH., Meyer CR.: Construction of an abdominal probabilistic atlas and its application in segmentation．, IEEE Trans Med lmag, Vol.22(4), pp.483−492, 2003

［6］ Shimizu R,.Ohno T,. Ikegami H,. et al.: Multi−organ segmentation in three dimensional abdominal CT images. CARS2006, Vol.1, pp.76-78, 2006

［7］ 堂之前義文，武尾英哉：画像処理方法，装置及びプログラム．特開 2013-105245

［8］ 安倍和弥，武尾英哉，永井優一：CT 画像からの 3 次元心肺ボリューム比計測の自動化．Med Imag Tech， Vol.29(5)，pp.267-276，2011

Examination of the high-resolution large intestine domain detection method from a simple abdominal CT image

Kazuya ABE[*1], Hideya TAKEO[*1], Yuuichi NAGAI[*2], Shigeru NAWANO[*3]

*1 Kanagawa Institute of Technology

*2 National Cancer Center Hospital East

*3 International University of Health and Welfare, Mita Hospital

Abstract

At present, the research & development of computer-aided diagnosis (CAD) is being conducted in various medical fields. Because of the difficulty in separating the large intestine from peripheral organs, little development has been carried out in the area of CAD that employs CT. Recently, due to improved endoscopic precision and the widespread use of CT colonography, research has been underway into such areas as applying CAD to discover colorectal cancer. However, endoscopy and CT colonography require advance preparation. CT colonography in particular is difficult to conduct in some cases because of stress placed on the colon when the colon is inflated by injecting carbon dioxide gas. In this study, research & development was conducted into a method for detecting, at high resolution, the large-intestine region from plain abdominal CT images captured during an abdominal examination. This method makes it possible to acquire images for use in CAD development without necessitating colon-stressing CT colonography. The concordance rate of 71% between the detection results of this method and manual detection demonstrate this method's detection performance.

Key words: large intestine, region extraction, plain abdominal CT image, CAD

特徴量選択手法およびシステム構成の最適化による うつ病 CAD の高性能化

牧　優太[*1] 和田　昇太[*1] 安倍　和弥[*1] 武尾　英哉[*1] 永井　優一[*2]

要旨

我が国におけるうつ病患者は増加傾向にあるが，これは，受診者の増加が原因の一つとしてあげられる．この受診者の増加によって，過剰診断等の問題が増えており，より適切な診断が行われることが求められている．しかしながら，現在の診断は客観性のある指標を用いておらず，その診断の正確性が問われている．そこで我々はうつ病の客観的な診断のために，うつ病である確信度を算出し，医師の診断の参考にしてもらうシステムの開発を行っている．従来の判別精度より高い精度を出せるようになることを目標として，従来のプログラムの構成に対し，音声判別器の追加などの構成の最適化を行った．また，18 の特徴量選択手法を用いた特徴量選択や，特徴量の組み合わせの総当たりアルゴリズムを用いた最適モデルの抽出も行った．その結果，従来の判別精度である 83% より高い，男性と女性，男女混合のデータでそれぞれ 90%，95 %，100% の判別精度を得ることができた．

キーワード：うつ病，特徴量選択，最適化

1. はじめに

我が国におけるうつ病患者は増加傾向にある．その要因は複数あるが，一つとして受診者数の増加があげられる．そのようになったのは，以下のような結果が現れてきたからである．

1) 「うつ病は心の風邪で，誰にでもなりうる」といった風潮の広がりによる，精神科病院への受診の抵抗感の低下[1]
2) 国や民間の啓蒙活動によるうつ病の症状の認知

この受診者数の増加によって，過剰診断等の問題が見過ごせないものとなったため，より適切な診断を行えるようになることが求められている[2].

このうつ病等精神疾患の診断方法であるが，問診によって患者の症状及びその期間を聞き取り， DSM-5(精神障害の診断と統計マニュアル第 5 版) や ICD-10(国際疾病分類第 10 版)などの診断基準に照らし合わせ，診断を行うという方法をとっている[3]．しかしながら，この診断方法には，以下のような問題がある．

1) 患者自身の自覚症状の有無や自己申告に頼らざるを得ない
2) 診断には医師の主観が影響してしまう
3) 客観的な指標(バイオマーカー等)がない

上記の問題から，うつ病をはじめとした精神疾患は，客観性の欠けた方法で診断が行われているというのが現状である．

以上を踏まえ，我々はうつ病の客観的な診断のために，画像工学技術等を利用してうつ病である確信度を算出し，医師の診断の参考にしてもらうシステムの開発を行っている．

本研究のようにうつ病の症状等を定量化し，その数値をバイオマーカーとして用いることを試みる研究や，うつ病の判別を行う研究，重症度を推定する研究は国内外で行われている．例えば音声に関しては，篠原ら[4]は，ゼロ交差

*1 神奈川工科大学大学院　工学研究科 電気電子工学専攻
〔〒243-0292 厚木市下荻野 1030〕
e-mail: maki2019@ele.kanagawa-it.ac.jp
*2 国立がん研究センター東病院

率とハースト指数の関係性に基づく音声指標を用いて重症度測定を行い，その有効性を示唆している．四井ら[5]はうつ病の重症度の評価尺度の開発のため，音声の波形データを変換し，CNN でうつ病か健常者かどうかの判別を行うことをしている．医用画像に関しては，清水ら[6]は fMRI 画像に対し主に LASSO を用いた機械学習を行った結果，90%以上の精度でうつ病の判別を行えている．また，市川ら[7]も fMRI 画像を用いた複数の機械学習アルゴリズムによる判別で，84%の精度を得られている．一般画像では，Andrew ら[8]が Instagram に投稿した画像から，うつ傾向のある人物を約 70%の精度で識別している．本研究では，上記で示したような音声や画像等の情報単体で判別等を行うのではなく，複数の情報を用いて，総合的にうつ病である確信度の算出等を行う．

このうつ病である確信度を算出するシステムについて，従来の研究では 2 段 stacking 構成で行っていた．これは，視線方向と表情に着目し，それぞれのデータを単体の判別器で学習させ[9], その 2 つのモデルの出力を組み合わせる．そして，それを入力として複数の判別器で学習を行い，逐次選択型という方法などで，スコアの算出を行う方法である．その結果，約 83%の精度でうつ病患者と健常者の判別を行うことができていた [10][11].

本論文では，より判別精度を向上させる，より正確なうつ病である確信度を出せるようにするため行った，特徴量選択の手法や 2 段 stacking 構成の最適化などの方法とその結果について記す.

本論文では，第 2 章で本研究に使用している動画とラベル付けについて，第 3 章では精度向上のために用いる手法等を説明する．第 4 章では第 3 章の手法による結果を，第 5 章は第 4 章の結果からの考察をまとめている．第 6 章では現在の取り組みについて述べる．

表 1 動画サンプル

	Depression	Normal	Total
Men	10	10	20
Women	10	10	20
Mix	20	20	40

2. データ

本研究に使用する動画は，動画投稿サイトに投稿されていたものを使用している．使用する動画の条件としては，投稿者がカメラの正面に位置し，カメラに向かって話しかけているもののみとし，収集を行っている．収集した動画に対するうつ病患者と健常者のラベル付けは，動画内で自身がうつ病だと明言している動画の人物をうつ病患者，同様にうつ病だと明言していない動画の人物を健常者というように行った．結果として，表 1 のようにそれぞれうつ病患者 10 名, 健常者 10 名の計 20 名のデータを，男女別に集めることができ, サンプル数は 40 名となった.

本論文では，実験用データの作成にあたって，使用する動画の長さを 200 から 240 秒の間と一定にしている.

3. 手法

本論文で用いた，うつ病確信度算出，ひいてはうつ病判別システムに用いるプログラムの構成を図 1 に示す．全体的な構成は従来の 2 段 sacking 構成とほぼ同じであるが, 異なる点を以下に示す.

1) 音声を用いた判別器の追加
2) 特徴量選択手法のアンサンブル化モデルの実装
3) 特徴量の組み合わせの総当たりアルゴリズムの実装
4) 2 段 Stacking 構成の最適化

これらの詳細もしくは概要を本章 3,1，3.2，3.3，3.4 にて述べる.

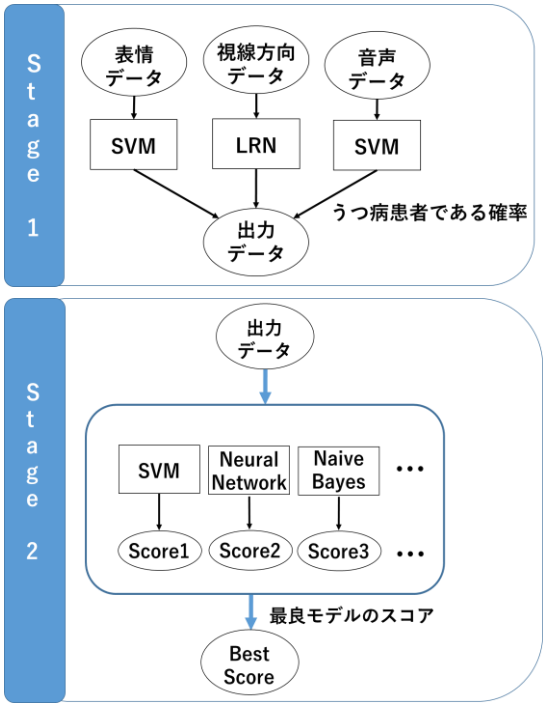


図 1 2 段 Stacking モデル

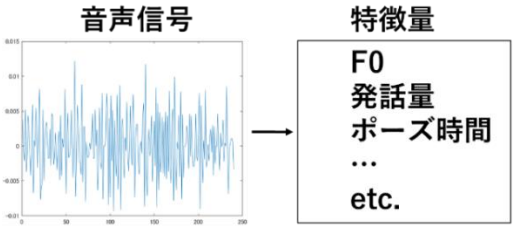


図 2 音声データからの特徴量作成

3.1 音声を用いた判別器の追加

和田ら[12]が作成した, 音声によるうつ病判別器を Stage1 に組み込んだ．これは，図 2 に示すように，音声信号から発話量などに着目して特徴量を作成し，それらを用いて線形 SVM で学習等を行ったものである．

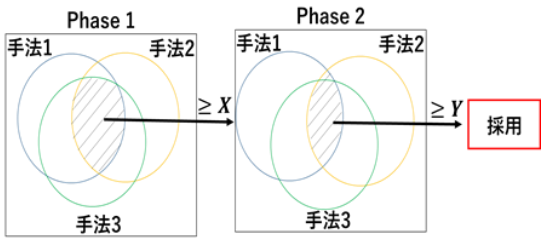


図 3 アンサンブル化モデル

表 2 各 Phase，方法での特徴量選択手法の数

Phase	Method	Total
1	Filter	8
	wrapper	1
	Embedded	6
2	Decision tree	3

3.2 特徴量選択のアンサンブル化モデルの実装

図 3 に今回使用した特徴量選択のアンサンブル化モデル(以降, アンサンブル化モデルと呼ぶ)の概要について記す. このアンサンブル化モデルは，複数の特徴量選択手法で，段階に分けて特徴量を選択する方法である．図に示した通り，複数の特徴量選択手法を用いて特徴量選択を行い，共通部分(厳密にいえば，手法毎の組み合わせによる共通部分も斜線部分に入る)にある特徴量を選択する．Phase1 では共通部分の特徴量の中で，X 回以上各手法で選択されれば採用とし，Phase2 に移行．Phase1 で採用された特徴量に対し複数の手法で特徴量選択を行い，Y 回以上で最終的な特徴量の採用としている．

表 2 に，各 Phase での，特徴量選択の手法を大まかなタイプに分けたときの総数を示す．表より, Phase1 では Filter 法で 8, wrapper 法で 1, Embedded 法で 6 の総計 15 種類の特徴量選択手法を用いている．Phase2 では，Decision tree そのものや，それを用いたランダムフォレストなどの 3 種類を用いている．表 3 および表 4 に Phase1 での，wrapper 法を除いた Filter 法と Embedded 法で用いた特徴量選択手法と，選択する際に目安とした値について示す．

表 3 Filter 法で用いた特徴量選択手法

method	parameter
分散分析	F値
χ 2統計量	p値 (−log（p）)
最小冗長性特徴選択	スコア
F検定	p値 (−log（p）)
ラプラシアンスコア	スコア
相関の高さを利用した特徴量選択	相関係数
近傍成分分析	重み
ReliefFアルゴリズム	重要度(重み)

表 4 Embedded 法で用いた特徴量選択手法

method	parameter
線形判別分析分類器	係数
ECOC(誤り訂正出力符号)	係数
線形回帰モデルのあてはめ	係数
リッジ回帰モデル	係数
ElasticNetを用いた正則化	係数
LASSO 正則化	係数

以下に Phase1 で使用した手法と選択の仕方について簡単に記す.

1) Filter 法

・分散分析

分散分析を使用して, 個々の特徴量の F 統計量を算出する．そして，その値が大きいものから選択する.

・χ^2 統計量

カイ 2 乗検定を使用して，個々の特徴量の検定統計量 p 値を算出する．そして，真数として−log()の式に入力し，そのスコアが大きいものから選択する.

・最小冗長性特徴選択

最小冗長性最大関連性（mRMR）アルゴリズム[13]を使用して, 各特徴量のスコア(重要度)を算出する．その値の大きいものから選択する.

・F 検定

応答(ラベル)をうつ病患者 1，健常者 0 というように数値化し(以降, 数値化と呼ぶ), それに対し F 検定を行う．各特徴量の値でグループ化された応答値が同じ平均の母集団から抽出されたという帰無仮説を、母集団の平均がすべて同じではないという対立仮説に対して検定を行い，p 値を得る．そして，真数として−log()の式に入力し，そのスコアが大きいものから選択する.

・ラプラシアンスコア

初めに数値化を行ってから，ラプラシアンスコア[14]を使用して, 各特徴量のスコア(重要度)を算出する．そして，そのスコアが大きいものから選択する.

・相関の高さを利用した特徴量選択

初めに数値化し，その応答値と各特徴量の間の相関係数を算出する．そして，相関の高い(正負は問わない)ものから選択する.

・近傍成分分析

正則化パラメーターの値を変えながら 5 分割交差検証法で近傍成分分析に学習を行い，最小の損失となる時の，最適な正則化パラメーターを見つける．そのパラメーターの状態での，学習済みモデルの特徴量の重みを抽出し，その重みが大きいほうから選択する.

・ReliefF アルゴリズム

k 最近傍分析(k=10)を ReliefF アルゴリズム[]に適応し，各特徴量の重みを算出する．そして，その重みが大きいほうから選択する.

2) wrapper 法

・前方特徴量選択

特徴量を 0 から，1 つ 1 つと特徴量を追加していく．基準となる誤分類率が減少しなくなるまで続けていき，追加を終了した時点での特徴量を，追加が早い順に選択する.

3) Embedded 法

・線形判別分析分類器

線形判別分析分類器の学習の際，5 分割交差検証損失が最小となるようなハイパーパラメーターの最適化を行い，その状態の時の各特徴量に対応する係数を求める．そして，その係数の大きい方から選択する.

・ECOC

ロジスティック回帰学習器を使用する線形分類モデルを基とした，ECOC(誤り訂正出力符号)モデルを用いる．5 分割交差誤差が最小となるような LASSO ペナルティ(L1 正則化項)を見つけ，その時の線形係数を求め，その値の大きい方から選択する.

・線形回帰モデル

ロジスティック回帰学習器を使用する線形分類モデルを用いる．LASSO ペナルティの値を変化させたときの，線形分類モデルの交差検証 AUC と使用された特徴量数(係数の非 0 の数)のグラフを作成する. そのグラフを図 4 に示す. この図から特徴量数と AUC の高さのバランス

がとれている LASSO ペナルティの値を選択し，その値の時の線形係数を求め，その値の大きい方から選択する．

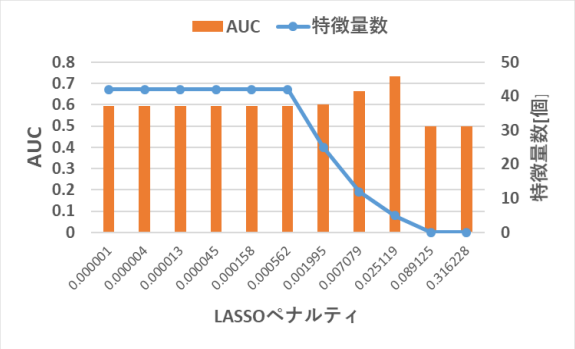


図 4 特徴量数と AUC の 2 軸グラフ

・リッジ回帰モデル

初めに数値化を行う．次にリッジ パラメーターを 5，スケーリングを 0 に設定して学習を行った時の，リッジ回帰モデルの係数を求める．そして，その係数の大きい方から順に選択する．

・ElasticNet を用いた正則化

初めに数値化を行う．次に 10 分割交差検証と L_1 正則化の重みを 0.75 と設定した時の Elastic Net 法を用いて，線形回帰モデルの係数を求める．そして，MSE と最小 MSE の差が 1 標準誤差以内になる最大の正則化係数の値を求め，その値の大きい方から選択する．

・LASSO 正則化

初めに数値化を行う．次に 10 分割交差検証を使用して LASSO 近似を求める．そして，MSE と最小 MSE の差が 1 標準誤差以内になる最大の正則化係数の値を求め，その値の大きい方から選択する．

表 5 Phase2 で用いた特徴量選択手法

method	parameter
分類木	重要度
分類木のアンサンブル	重要度
ランダム フォレスト	重要度

表 5 に Phase2 で用いた特徴量選択手法と選択する際に目安とした値について示す．以下に Phase2 で使用した手法，選択の仕方について簡単に記す．

・分類木

代理分岐がある分類木を使用し，CART 法を用いて学習を行う．そして，各特徴量の分割によるリスクの変化を合計したものを，ノード数で割ることで重要度を算出し，その重要度の大きい方から選択する．

・分類木のアンサンブル

ブースティングされた分類木のアンサンブルモデルを用いる． 初めに 150 本のブースティングされた分類木のアンサンブルに 5 分割交差検証で学習を行う．次に，アンサンブル内の木の本数が増加する度に誤分類率を算出させ，その誤分類率が最小になる最大分割数、木の本数および学習率を特定する．そして，それらを用いて学習を行い，分類木の時と同様に重要度を算出し，その値の大きい方から選択する．

・ランダムフォレスト

バギングされた 200 本の分類木のアンサンブルに学習をさせて，out-of-bag 分類誤差とモデル(使用された値)の誤差を求める．各決定木でその 2 つの差を求め，重要度である木あたりの平均を得る．そして，その値の大きい方から選択する．

以上である．本論文は各手法の簡単な説明にとどまっている．そのため，より詳しく知りたいのであれば次の資料を参考にしてほしい[15].

3.3 特徴量の組み合わせの総当たりアルゴリズムの実装

図 5 に示すような特徴量の組み合わせの総当たり(全組み合わせ)アルゴリズムを作成した．Stage1 の学習においては，このアルゴリズムを用いて学習を行い，その組み合わせ毎の判別精度の中で，最も判別精度の良かった時のスコア(うつ病患者である確信度)を Stage2 への入力に採用した．

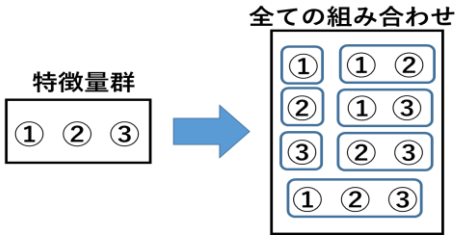


図 5 特徴量の全組み合わせ

3.4 Stacking 構成の最適化

従来の Stacking 構成における Stage2 での学習では，逐次選択型を用いていた．しかしながら，その効果は見込めなかったため，本論文では 60 モデルのうち最も高い判別精度の出したモデルとその精度を採用するようにした．
また，学習方法も 4 分割交差検証ではなく，1 個抜き交差検証(Leave-one-out cross-validation)に変更を行った．

3.5 比較

作成したアンサンブル化モデルと，ランダムフォレスト単体を以下の 2 点で比較した．

1) Stage1 での各方法による視線方向データと音声データに対する判別精度
2) Stage2 での判別精度

データは表 1 に示した男性 20 名と女性 20 名，合わせた 40 名のデータを基に作成した，視線方向データセットと音声データセットを用いる．このデータセットの 1 人当たりの特徴量の個数は，視線方向は 12 種類，音声は 42 種類である．

特徴量の選択については，アンサンブル化モデルは視線方向: X＝7，Y＝2，音声信号:X＝7，Y = 3，各 18 手法の特徴量選択数は入力される特徴量数の半分としている．このように設定した時に採用された特徴量とした．一方，ランダムフォレストにおいては 10 回特徴量選択を行ったうち，7 回以上選ばれたものとした．こちらも，アンサンブル化モデルと同様に 1 回の特徴量選択数は入力特徴量数の半分とした．

学習にあたってはどちらも線形 SVM を用い，1 個抜き交差検証で学習を行った．なお，今回は特徴量選択による判別精度の比のため，Stage1 での表情データのスコアはどちらも同様のものを使用しており，単体での判別精度はどのデータに対しても 65 % となっている．

4. 結果

表 6 にアンサンブル化モデルとランダムフォレストによって選択された特徴量を用いて学習した時の，各 SVM での判別精度を示す．この表は，同データに対して判別精度が高い方法の方は，赤色で塗りつぶしている．音声データに対しては一部アンサンブル化モデルの方が高めのものもあったが，全体的にはランダムフォレストの方が精度は高めであるという結果となった．

次に，表 7 に Stage2 での判別精度を示す．表より男性のデータに対してアンサンブル化モデルの方が若干精度は低いものの，他に対してはほぼ同値を示した．最終的に，どのデータに対しても 90%以上の判別精度を得られ，前回の判別精度である 83%を超え，判別精度の向上を行うことができた．

表 6 特徴量選択手法，データ別の判別精度

Metod	Data	Men	Women	Mix
Ensemble	Eye_direction	75.0	75.0	62.5
	Voice	75.0	93.8	75.0
Random	Eye_direction	100.0	81.3	75.0
	Voice	75.0	75.0	79.2

表 7 特徴量選択手法別の判別精度

	Men	Women	Mix
Ensemble	90.0	95.0	100.0
Random	100.0	95.0	100.0

表 8 各特徴量選択で採用された特徴量数

Data	Method	Men	Women	Mix
Eye	Ensemble	3	2	1
	Randomforest	7	1	6
Voice	Ensemble	6	7	8
	Randomforest	8	8	14

表 9 特徴量選択別最高判別精度時の特徴量数

Data	Method	Men	Women	Mix
Eye	Ensemble	2	1	1
	Randomforest	1	1	1
Voice	Ensemble	3	2	3
	Randomforest	3	3	5

5. 考察

表 8 に，アンサンブル化モデルとランダムフォレストで特徴量選択を行った時の，選択された特徴量の数を示す．また，表 9 に特徴量の組み合わせの総当たりアルゴリズムを適用して学習を行った時の，最高精度時に入力された特徴量の数を示す．

表 8 より全体的に見れば，相対的にランダムフォレストの方が選択される特徴量数が多い傾向にある．視線方向データに関していえば，アンサンブル化モデルは選択した特徴量数がランダムフォレストの方より少ない．しかしながら精度は 60 から 70 半ばでとどまっているので，判別に有効な特徴量を選べていないと考えられる．一方として音声データに関していえば，アンサンブル化モデルはより少ない特徴量でほぼ変わらないか高い判別精度を得られている．以上のことから，アンサンブル化モデルは特徴量を厳しく絞り込むので，比較的多数の特徴量に対しては有効である．その一方，少量の特徴量に対しては，X と Y の値を適切に設定する必要があると考えられる． また，表 9 より，最高精度となるときの特徴量の個数は，視線方向に関しては 1 つ 2 つと極々少数となっている．確かに最高精度となるものの，これは 1 つの値だけみて判別を行っていることに等しく，汎化性があるとはとてもいえない．そのため，最高精度を算出するこの特徴量組み合わせの総当たりアルゴリズムに，最低限の数の特徴量を組み合わせた中で，最高精度のモデル選択をできるようにする必要があると考えられる．

6. 現在の取り組み

本章では，3 章で述べたうつ病確信度算出(判別)システムへ適応予定の，行っている判別器作成等の取り組みについて記す．現在，以下のようなことに取り組んでいる．

1) 表情データの特徴量化
2) 動作速度に着目した判別器の作成

これらの概要を本章 6.1，6.2 にて述べる．

6.1 表情データの特徴量化

現在のところ，表情データの学習に関しては時系列データそのものを使用している．このデータそのものであっても単体モデルでは 65%ほどの判別精度を持つが，他のモデルと見比べるとやはりいくらか見劣りしてしまうのが現状である．そのため，より精度の向上の目的として，表情に関しても，特徴量ベースによる判別器の作成を行っている．表情の時系列データから作成した，特徴量の例を一つあげる．女性の表情データより，表情変化の回数に着目し，1 分あたりの表情変化の回数の平均をとったものをまとめたグラフを図 6 に示す図より，相対的に見て，健常者の方が値は大きいことがわかる．また，このグラフに 2.5 の閾値を設け，その値より上であれば健常者，それ以下はうつ病患者とする．そうすると健常者は 7/10，うつ病患者は 9/10 正しく判別されたことになり，その判別精度は 80 %である．この表情データの特徴量化は，現状の判別精度を超えることを目標に，分離性の高い特徴量を作成し SVM に学習を行わせて，評価をしていく．

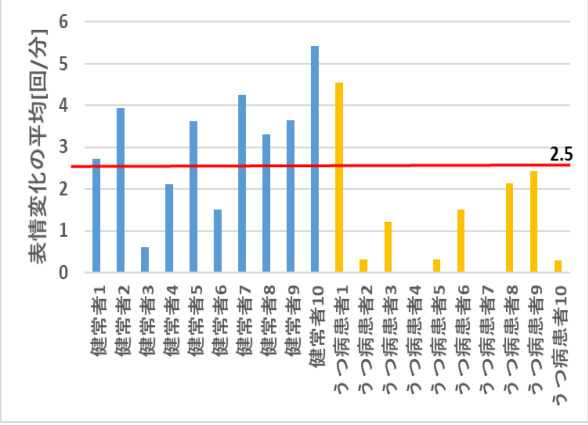


図 6 1 分あたりの表情変化回数の平均

6.2 動作速度に着目した判別器の作成

うつ病患者には，以下のような特徴がある．

1) 動作速度

動作が鈍くなり，ゆっくりとした動きになる．まるでロボットのように体の動きが遅くぎこちなくなる[16][17]

2) 落ち着きのなさ

強い不安、焦燥感からイライラして足踏みをしたり，ソワソワと落ち着きなく動き回ったりする[16]

上記の特徴に着目し，それらを用いた新しい判別器の作成を行っている．

図 7 に動作速度の算出方法を示す．上半身を検出する事前学習済みの分類器を用いて，30 毎秒フレームの画像の上半身を特定する．そして，特定した際の境界ボックスの左下の座標を利用し，フレーム間毎の移動速度を次の式で求める．

$$v = \frac{\sqrt{(x'-x)^2+(y'-y)^2}}{l} \qquad (1)$$

そして 1 秒ごとの平均値を求めまとめた．そのデータを用いて，特徴量の作成を行っている．表情と同様に，作成した特徴量の例を一つあげる．女性の動作速度の時系列データより，最大と最小の値の差をとったものをまとめたグラフを図 8 に示す. このグラフに 130000 と 450000 の閾値を設ける．前者はその値より下であれば，後者はその値より上であればうつ病患者とし，この 2 つの閾値の間であれば健常者とする．そうすると健常者は 8/10，うつ病患者は 7/10 正しく判別されたことになり，その判別精度は 75 % である．このように分離性の高い特徴量を作成し，それらで SVM に学習を行わせて，判別精度を評価していく．

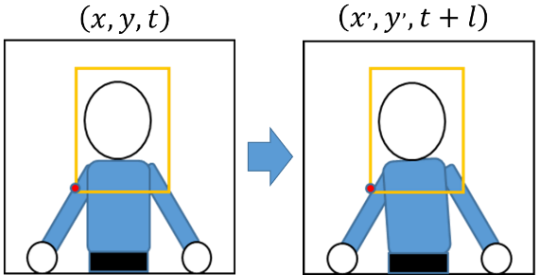


図 7 上半身の検出と座標，時間

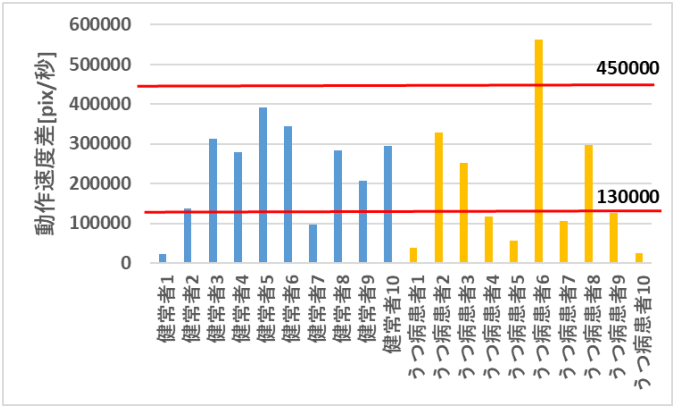


図 8 動作速度の最小値と最大値の差

7. まとめ

本論文では，アンサンブル化モデルという特徴量選択や従来の構成の最適化などを行った．その結果，前回の約 83%の判別精度を超える，90%以上の精度を得ることができた．

今後の取り組みとしては，上記の表情データの特徴量化や，動作速度に着目した判別器の追加の他に，汎化性の高い総当たりアルゴリズムに改善する，現在の構成に対し，段を追加するなどの更なる最適化を行う予定である．

利益相反の有無

利益相反　なし

文　献

［1］ 大坪天平：うつ病を診る上で注意すべきこと．女性心身医学 **20** 巻 3 号: 271-275, 2016

［2］ 日本うつ病学会：うつ病 Q&A，https://www.secretariat.ne.jp/jsmd/ippan/qa.html (2020 年 7 月 25 日閲覧)

［3］ 二村隆史：うつ病治療薬開発の難しさ．ファルマシア **53** 巻 7 号: 696-698, 2017

［4］ 篠原 修二，戸田 裕之，中村 光晃 他：ハースト指数とゼロ交差率に基づく音声指標によるうつ重症度測定．生体医工学 Annual**57** 巻 Abstract 号 S247_1, 2019

［5］ 四井美月，Kuo-ching Liang，廣原茉耶 他：音声データを用いたうつ病重症度

の深層学習解析．人工知能学会全国大会論文集 JSAI2018(0), 4C2OS27b02-4C2OS27b02, 2018

［6］ Shimizu Y, Yoshimoto J, Toki S, et al.: Toward Probabilistic Diagnosis and Understanding of Depression Based on Functional MRI Data Analysis with Logistic Group LASSO. PLOS ONE **10(5)**: e0123524, 2015

［7］ Ichikawa N, Lisi G, Yahata N, et al.: Primary functional brain connections associated with melancholic major depressive disorder and modulation by antidepressants. Scientific Reports **10**: 3542, 2020

［8］ [D] Andrew R, Christopher D: Instagram photos reveal predictive markers of depression. EPJ Data Science **6**, 15, 2017EPJ Data Science 2017

［9］ Maki Y, Abe K, Takeo H, et al.: Quantification of the diagnosis of depression through application of image recognition technology. Computer Assisted Radiology and Surgery 33rd International Congress and Exhibition, CARS2019, S181-S82, 2019

［10］ 牧優太，和田昇太，安倍和弥 他：画像認識技術によるうつ病診断の定量化．第 38 回日本医用画像工学会大会(JAMIT2019) 予稿集, OP1-17 , 2019

［11］ Maki Y, Wada S, Abe K, et al.: Developing High Performance CAD for Depression by Integrating Multiple Classifier Systems. Computer Assisted Radiology and Surgery 34rd International Congress and Exhibition, CARS2020, S204-S205, 2020

［12］ Wada S, Maki Y, Abe K, et al.: Developing High Performance CAD for Depression by Employing Image and Voice Information. Computer Assisted Radiology and Surgery 34rd International Congress and Exhibition, CARS2020, S203-S204, 2020

［13］ Chris D, Hanchuan P: Minimum redundancy feature selection from microarray gene expression data. Journal of Bioinformatics and Computational Biology. **3**, 2, 185–205, 2005

［14］ Xiaofei H, Deng C, Partha N: Laplacian Score for Feature Selection. NIPS'05: Proceedings of the 18th International Conference on Neural Information Processing Systems, 507–514, 2005

［15］ MathWorks：Introduction to Feature Selection, https://jp.mathworks.com/help/stats/feature-selection.html?lang=en(2020 年 7 月 29 日閲覧)

［16］ うつ病の情報・サポートサイト こころの陽だまり：うつ病症状の進行，https://www.cocoro-h.jp/untreated/overview/symptoms.html(2020 年 7 月 29 日閲覧)

［17］ 磐田こころのクリニック：うつ病について，https://www.iwatakokorono-cl.com/target_disease/depression/(2020 年 7 月 29 日閲覧)

Improving the performance of Depression CAD by using feature selection methods and optimizing system configuration

Yuta MAKI[*1], Shota WADA[*1], Kazuya ABE[*1], Hideya TAKEO[*1], Yuichi NAGAI[*2]

*1 Kanagawa Institute of Technology

*2 National Cancer Center Hospital East

The number of subjects with depression in Japan has been increasing and the number of medical examinees has also been increasing. Therefore, more appropriate diagnosis is required. However, the diagnosis of depression and other mental disorders has not been made using objective indicators such as biomarkers, and the accuracy of the diagnosis has been questioned. Therefore, we are developing a system for calculating the confidence of depression to help physicians to make a diagnosis. To achieve a higher accuracy than the conventional discrimination accuracy, we optimized the program's configuration by adding a voice classifier and using Leave-one-out cross-validation etc. Also, we performed feature selection using 18 feature selection methods and extraction of the optimal model using a brute force algorithm of feature combinations. As a result, the discrimination accuracies of 90 %, 95 %, and 100 % for male, female, and mixed-sex data were achieved, respectively, which is higher than the conventional accuracy of 83 %.

Key words: Depression, Feature Selection, Optimization

画像と音声情報を用いたうつ病 CAD の高性能化

和田　昇太[*1] 牧　優太[*1] 安倍　和弥[*1] 武尾　英哉[*1] 永井　優一[*2]

要旨

我が国におけるうつ病者は増加傾向にあり，2011 年には従来の 4 大疾病にうつ病を加え 5 大疾病とされた．このことからもうつ病者の確実な発見が望まれるが，現状の問診を主としたうつ病診断では，その客観的指標の欠如から診断結果の正確性を問う声もある．我々は患者の顔画像や音声情報に着目し，AI を用いてうつ病である確信度を算出することで客観的指標を提示し，医師の診断に役立ててもらうためのシステムを開発している．本稿では，そのシステムを構成する音声判別器の作成について述べる．患者の音声とうつ病者に見られる特徴から，全 42 個の特徴量を作成し SVM を用いて判別器の作成と評価を行ったところ，女性データに対しては 90%の精度が得られたが，男性データに対しては 55%程度の精度であった．そこでこの特徴量を 18 の特徴量選択アルゴリズムとその組合せに手法により選別し，再度判別器の作成と評価を行ったところ，男性データでも 90%の精度が得られた．

キーワード：うつ病，発話音声，機械学習，SVM，特徴量選択

1. はじめに

我が国において，うつ病患者は増加傾向にある[1]．また，平成 19 年から 21 年の警察庁統計における自殺の動機として各年うつ病が原因であるものが 4 割以上を占め，厚生労働省では自殺対策として，うつ病対策を重点的に行っている[2]．こうした背景もあり，2011 年には，従来の「がん」，「脳卒中」，「急性心筋梗塞」，「糖尿病」からなる 4 大疾病に，新たにうつ病を含む「精神疾患」が追加され 5 大疾病となり，国民の健康保持のため広範かつ継続的な医療の提供が必要な疾病であるとされた[3].

現在のうつ病診断は，DSM-5(精神障害の診断と統計マニュアル第 5 版) や ICD-10(国際疾病分類第 10 版)などの診断基準に照らし合わせ行われるものである．しかしながらこれが客観性を欠き，うつ病者の発見を阻害しているという指摘もあり，うつ病判別に有効なバイオマーカーについての研究が行われている[4].

例えば人間の発話音声を対象として分析し，バイオマーカーとして用いた研究として，和家ら[5]は，収録環境を統一したうえで様々な課題に沿って収録した音声を分析し，従来から検討されてきた音声のエネルギーなどの有効性を確認し，また時間内に想起された単語数といった新たなバイオマーカーを抽出して有効性を示した．しかしながら，これらの手法には，文章読み上げなど患者側の負担が考えられる．

我々もうつ病判別システムの開発を行ってきた[6][7][8]．これは AI 技術と画像工学技術を組み合わせうつ病の確信度を算出することで，医師に客観的な指標を提示し，診断の参考にしてもらうことを目的としたものである．本研究においては，新たに検討した，音声を用いたうつ病判別器について述べる，最終的に前述の画像工学技術を用いた判別器と統合

*1 神奈川工科大学大学院　工学研究科 電気電子工学専攻
〔〒243-0292 厚木市下荻野 1030〕
e-mail: wada2019@ele.kanagawa-it.ac.jp
*2 国立がん研究センター東病院

し，システム全体としての判別精度向上を実現する．発話音声については定型文読み上げなどを用いない自由発話の音声を用いて判別器の検討を行う．この 2 点が新規な点である．

本論文の構成としては，2 章で研究に用いたデータについてとその分析結果について整理し述べる．第 3 章では判別器の作成・評価方法とうつ病者に見られる特徴をグループ分けしそれぞれについて検討した特徴量の算出方法について述べる．4 章では，3 章の内容についての結果を示し，5 章ではその考察を述べる．6 章で現在の取り組みについて，7 章で総まとめとする．

2　データ

2.1　データの準備

本研究のデータは，動画投稿サイト YouTube(https://www.youtube.com/?gl=JP) よりうつ病者・健常者のそれぞれについて，男女各 10 名ずつの計 40 名の動画を収集した(表 1)．実際の問診を想定し動画の内容は，自己紹介など自分について自由に話しているものとした．動画からサンプリング周波数 44,100[Hz]，16[bit]で音声を抽出し，モノラルオーディオに変換したものから 240[s]に相当する 10,584,000[sample]を切り出し利用した．またうつ病者・健常者のラベル付けは，自身が動画内などでうつ病を罹患していると申告している者をうつ病者，そうでない者を健常者とした．

表 1．収集し研究に用いたデータの総数

	うつ病者	健常者	合計
男性	10	10	20
女性	10	10	20
合計	20	20	40

2-2　データの分析

前節で準備したデータを視覚・聴覚的に評価した．うつ病者と健常者では，

・発話量に乏しい
・発声量に乏しい
・声の抑揚に乏しい

などの特徴がみられた．

これらの観察結果と先行研究などの文献[5][9]などとを照らし合わせたうえで，うつ病者の発話音声に特有の特徴をグループ分けし以下表 2 にまとめた．

一例としてグループ 2 について，うつ病者と健常者の音声波形を以下図 1, 図 2 に示す．うつ病者は振幅が小さく声量が小さい傾向がみられた．

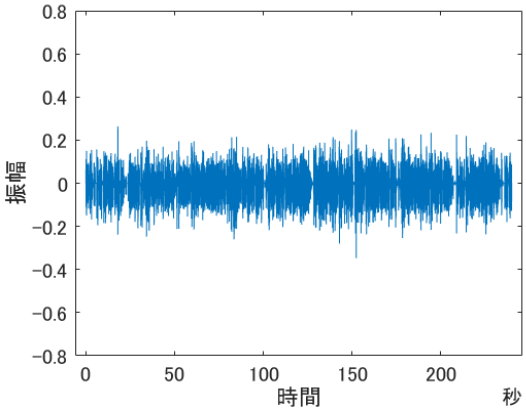


図 1．うつ病者の音声波形の例

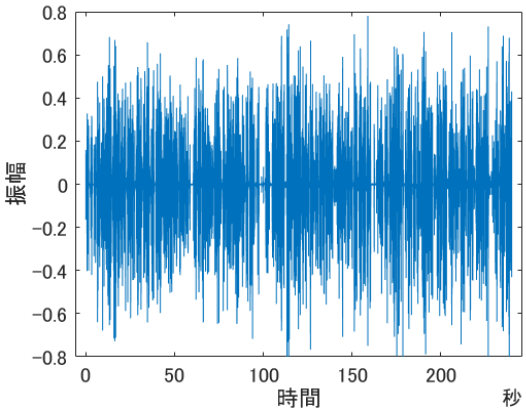


図 2．健常者の音声波形の例

表 2．うつ病患者の特徴

グループ	発話音声の特徴	うつ病患者の場合
1	発話速度	1. 口数が少ない
		2. 話し方が単調で途切れ途切れ
		3. 動作が遅い
		4. 思考速度が遅い
2	発声量	1. 声量が小さい
3	抑揚	1. 抑揚が乏しい
		2. 活気がない
		3 感情が動かない

3. 手法

判別器は SVM(Support Vector Machine：サポートベクターマシン)で作成する．それにあたり，表 2 にまとめたうつ病者の特徴をもとにグループごとに複数の特徴量を作成した．作成した判別器は，LOOCV(Leave-One-Out Cross-Validation：一個抜き交差検証)により，判別精度を式(1)として評価を行った．判別器は男女の性差を考慮するため，及び判別精度向上のための特徴量選択による次元圧縮をより有効的に行うために男女でデータを分けて，男性用判別器，女性用判別器の 2 つを作成・評価した．

$$\text{判別精度} = \frac{\text{真陽性+真陰性}}{\text{全体}} \quad (1)$$

3.1 特徴量の作成(グループ 1 の特徴定量化)

発話速度に関連した特徴を表 2 中グループ 1 とした．うつ病患者には表 2 中の 1-1.口数が少ないという傾向がある．また表 2 中 1-2.話し方が単調，1-3.動作が遅い，1-4.思考速度が遅いなどの特徴と相まってこの傾向を強めていることが考えられる．そこで，聴覚評価により，音声 240[s]に相当する，10,584,000[Sample]の発話・非発話区間をラベル付けし，全サンプルにおける発話サンプルの割合を特徴量とし定量化した．以下図 3, 図 4 の有色部分は，それぞれ音声 60[s]に相当する 2,646,000[Sample]中の有声部分である．一発話の継続時間や 1 秒以上のつまり(ポーズ)回数も計算し特徴量として用いた．

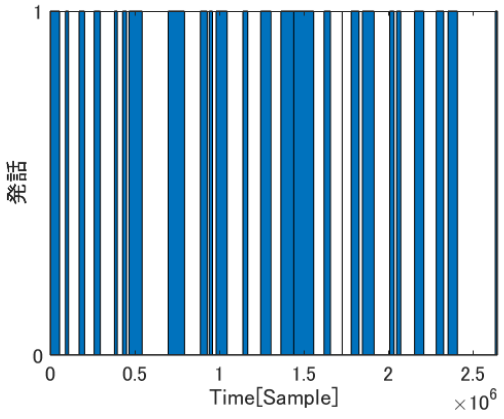


図 3. うつ病者の発話量

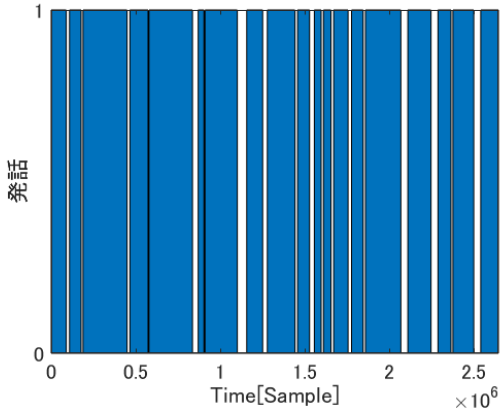


図 4. 健常者の発話量

3.2 特徴量の作成(グループ 2 の特徴定量化)

発声量について，健常者とうつ病者で差が見られた．具体的には，健常者の方が発声量が大きく，逆にうつ病者においては小さいといった具合である(図 1，図 2)．発声量を以下(2)式, (3)式の様に計算し特徴量として定量化した．音声波形をフーリエ変換したときのパワースペクトルの例を以下図 5, 図 6 に示す．音声波形の変動成分はうつ病者において小さい傾向にあった．

$$\textit{振幅・FFT：}RMS[Pa] = \sqrt{\frac{1}{t_2-t_1}\int_{t_1}^{t_2} x^2(t)dt} \quad (2)$$

$$\textit{信号パワー：}p[w/m^2] = \frac{1}{\rho cT}\int x^2(t)dt \quad (3)$$

ρ :空気の密度：$1.14[kg/m^3]$
c:空気中の音速：353[m/s]
T:音源の長さ：[s]
$x^2(t)$：音圧[Pa]

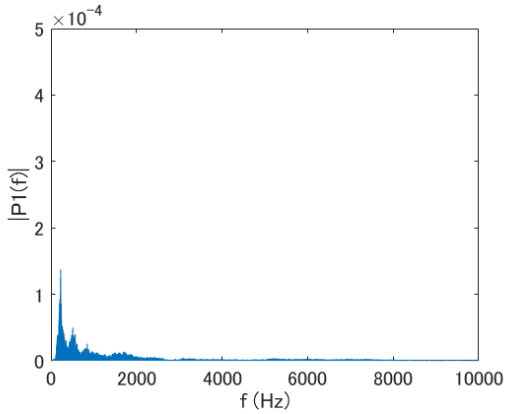


図 5. うつ病者の音声パワースペクトル例

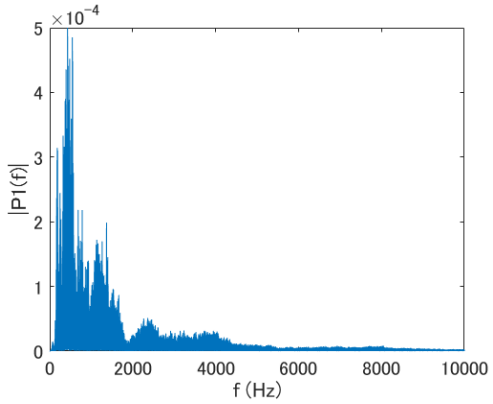


図 6. 健常者の音声パワースペクトル例

3.3 特徴量の作成(グループ 3 の特徴定量化)

うつ病者における発話音声の特徴として，抑揚の消失があげられる．表 2 中の 3-1.抑揚に乏しい，3-2.活気がないといった傾向をグループ 3 に分類し定量化した．特徴量には音声信号から wavesurfer[10]を用い抽出した基本周波数 F0 を用いた. 同ツールでは, RAPT(A Robust Algorithm for Pitch Tracking)というアルゴリズムを用いることで，精度の高い F0 抽出を可能にしている． 0.01[s]毎に抽出した F0 と，その変化量ΔF0 の時系列波形から，最大・最小値や平均値，分散などの基本統計量を計算し，特徴量として用いた．以下図 7 に抽出した F0 の時系列波形を示す.

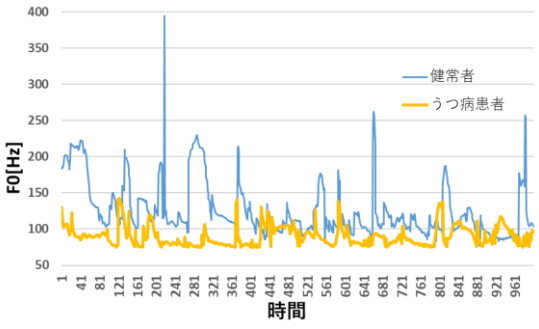


図 7. 抽出した F0 の時系列波形

3.4 特徴量選択

判別精度向上の為に本節迄の手法により作成した特徴量 42 個を選択・圧縮し，判別器を作成後評価した．以下にその手法を述べる.

3.4.1 選択手法①

入力された特徴量を表 3，表 4 に示す計 18 の手法により選択する．この過程は 2 段階に分けて行う．まず 1 段目では，表 3 に示す 15 の特徴量選択方法(アルゴリズム)を用いて選択を行う．これらの複数のアルゴリズムにより，7 回以上選ばれた特徴量 X を次段に入力する．2 段目では入力された特徴量 X を表 4 の 3 つのアルゴリズムを用いて再度選択する. これら一連の流れを手法①とし，概略図を以下図 7 に示す．表 3，表 4 中の各方法の詳細な説明については，次の資料を参考にされたい[11].

3.4.2 選択手法②

入力された複数の特徴量の組合せ全通りを入力として SVM による判別器作成と LOOCV による評価を行い，判別精度が最大になったときの特徴量を選択する. これを手法②とし，概略図を以下図 8 に示す.

3.4.3 各手法の組合せ

本項までに述べた手法①，②，及びランダムフォレストを組み合わせて特徴量選択を行い，最終的な判別器の精度を評価した．以下にその組合せについて列挙する.

1)ランダムフォレストのみ
2)ランダムフォレスト ＋ 手法②
3)手法①のみ
4)手法① ＋ 手法②

表 3. 1 段目の特徴量選択方法

段階	方法	
1	Filter	分散分析
		χ2乗統計量
		最小所調整特徴選択
		F検定
		ラプラシアンスコア
		相関
		近傍成分分析
		RelieF
	Wrapper	
	Embedded	線形判別分析分類器
		ECOC(誤り訂正出力符合)
		線形会期モデル
		リッジ回帰モデル
		ElasticNet正則化
		Lasso正則化

表 4. 2 段目の特徴量選択方法

段階	方法	
2	Decision tree	分類木
		分類木のアンサンブル
		ランダムフォレスト

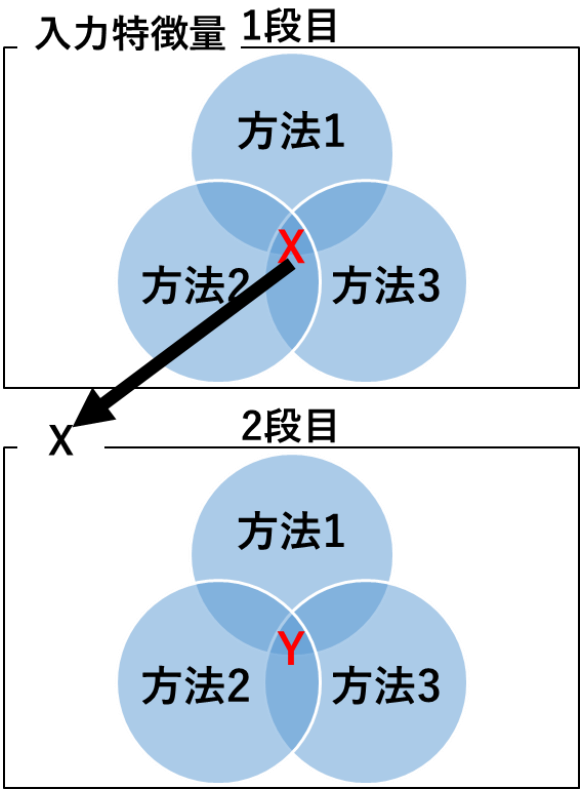


図 7. 手法①概略

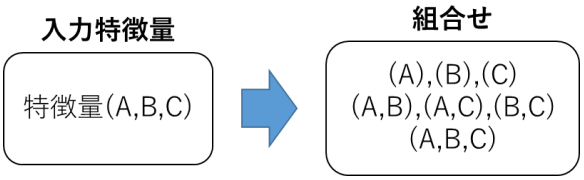


図 8. 手法②概略

4. 結果

3.1～3.3 節で述べた手法で算出した特徴量の一部分布を以下 4.1～4.3 節に示す．4.4～4.5 節では特徴量選択の結果について述べる．

4.1 グループ 1 の特徴量について

音声 240[s]相当の 10584000[Sample]における発話サンプルの割合を発話量と定義した．一例として発話量の最大値が 1，最小値が 0 となるように正規化したものをヒストグラムにし以下図 9 に示す．また発話中に 1 秒以上詰まる回数について，同様に正規化後ヒストグラムにしたものを以下図 10 に示す．

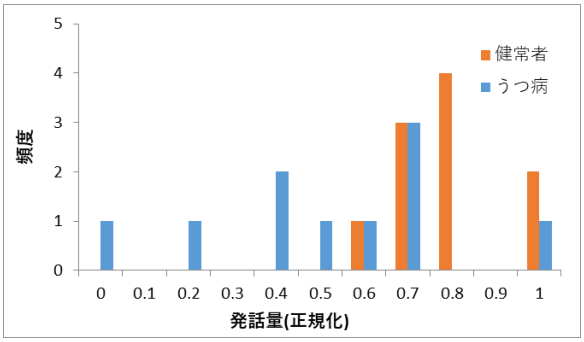


図 9. 発話量(男性)

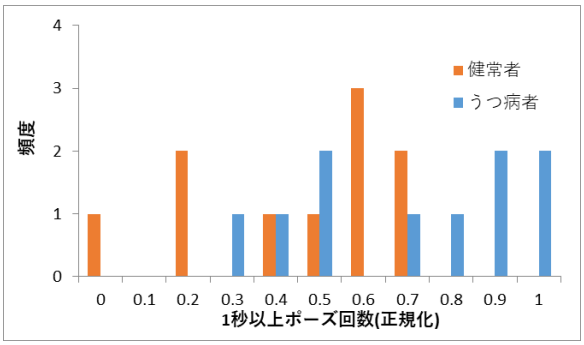


図 10. 1 秒以上ポーズ回数(男性)

4.2 グループ 2 特徴量について

3.2 において定義した式(2)より信号のパワーを計算した．これは表 2 中特徴グループ 2 の声量を定量化したものに相当する．一例として以下図 11 に示す．

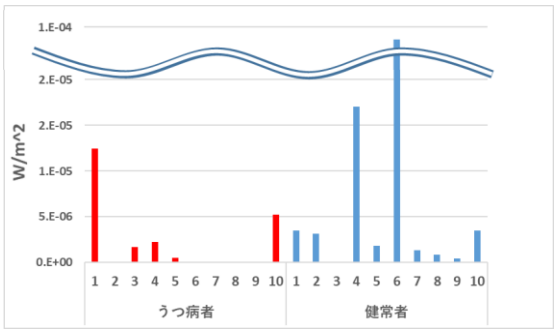


図 11. 信号パワー(女性)

4.3 グループ 3 の特徴量について

ΔF0 の最大値，及び分散について計算し，それらの最小値が 0，最大値が 1 となるように正規化したうえでヒストグラムにまとめたものを一例として以下図 12，図 13 に示す．これらは，特徴量グループ 3 の 1.抑揚の大き

さに相当する.

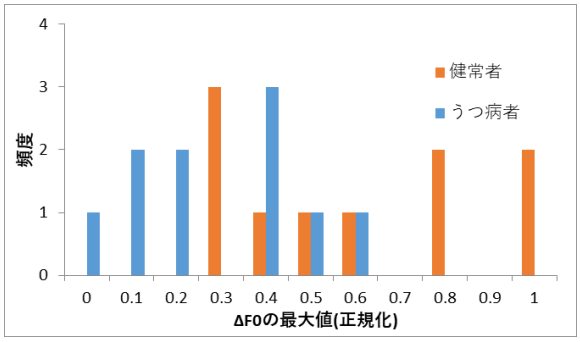


図 12. ΔF0 の最大値(女性)

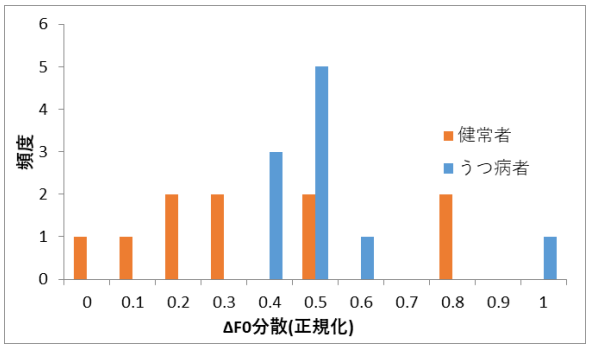


図 13. ΔF0 の分散(女性)

特徴グループ 1, 2, 3 の計 42 個の全ての特徴量を用いて, SVM で学習したときの判別精度は，男性判別器で 55%，女性判別器で 90%であった.

4.4 特徴量選択

3.4.3 項で述べた各手法の組合せ(1)～(4)により選択された特徴量の個数と，それを用いて SVM で学習・LOOCV で評価を行ったときの判別精度を以下表 5 に示す.

表 5. 選択された特徴量の個数と判別精度

各手法の組合せ	データ	選択された特徴量	判別精度
(1) ランダムフォレストのみ	男性	8個	70%
	女性	8個	90%
(2) ランダムフォレスト+手法②	男性	3個	80%
	女性	3個	90%
(3) 手法①のみ	男性	6個	75%
	女性	7個	75%
(4) 手法①＋手法②	男性	3個	90%
	女性	2個	90%

4.5 特徴量選択の有用性

特徴量 42 個を全て用いたときの判別器の判別精度と，手法の組合せ(4)により選択された特徴量を用いたときの判別精度を以下表 6 に示す．男性判別器について，35 ポイントの大幅な精度向上が見られた.

表 6. 本手法で得られた判別精度まとめ

	全42特徴量を使用	特徴量選択法(4)
男性	55%	90%
女性	90%	90%

5. 考察

表 6 より，特徴量選択法によって男性判別器，女性判別器でともに 90%の精度が得られた．しかしながら，表 5 より今回の特徴量選択手法では，判別精度を最大化することに重点を置き，最終的に 3 つ程度の特徴量で判別器を作成している．これは，精度を最大化できる代わりに，未知データに対する汎化性能が低下することが考えられる．そのため例えば，表 2 に示した特徴グループに対応した特徴量を，各グループから必ず複数個以上の特徴量が選ばれるようにするなどの方法で汎化性能を担保するなどの検討が必要であると考えられる. また全特徴量 42 個を用いたときの男性判別器の判別精度が著しく低いことからも，よりうつ病者の特徴をとらえた特徴量作成について引き続き検討することが最優先事項であると考えられる.

6. 今後の展望

今後の展望として，現在検討していることについて述べる.

6.1 特徴グループ 4 の設定

音声の評価を改めて行ったところ，うつ病者は健常者に比べて声がかすれているように聞こえる傾向が見られた．そのため，これらの特徴を，以下表 7 に示す特徴グループ 4 に設定した．音声合成システムにおいて，かすれ声を合成する際に非周期的な信号を加えることから，発話音声中に含まれる非周期成分の割合を，声分析変換合成システム”WORLD”[12]を用いて 5[ms]毎に計算した. 以下図 14，図 15 にその時系列波形を示す. またその平均値を計算することで特徴量としてかすれを定量化した. 健常者とうつ病者で,

平均値の分布をヒストグラムにしたものを以下図 16 に示す. 今後はこれらの特徴量も詳細に検討し，判別器作成に用いる予定である.

表 7. 特徴量グループ 4 の設定

グループ	発話音声の特徴	うつ病患者の場合
4	その他の特徴	1.悲観的
		2.劣等感や無力感
		3.声がかすれる
		4. 周囲や他者への興味関心の喪失

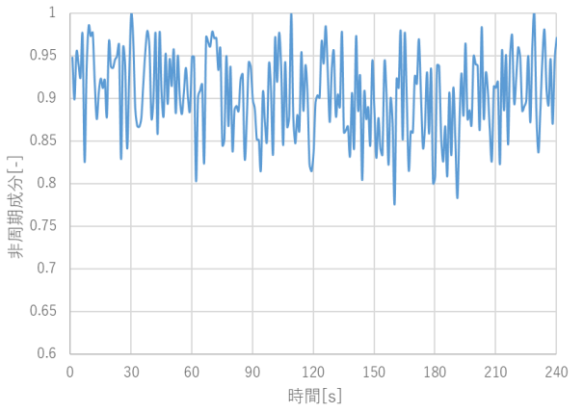


図 14. 非周期成分(うつ病者)

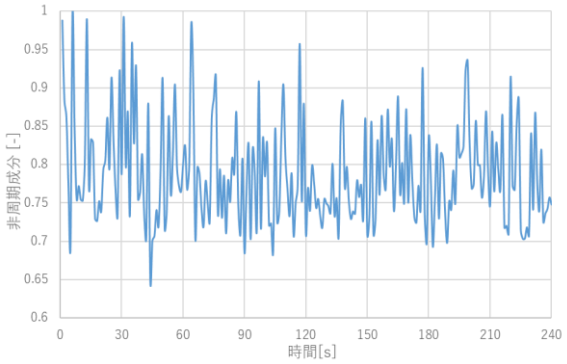


図 15. 非周期成分(健常者)

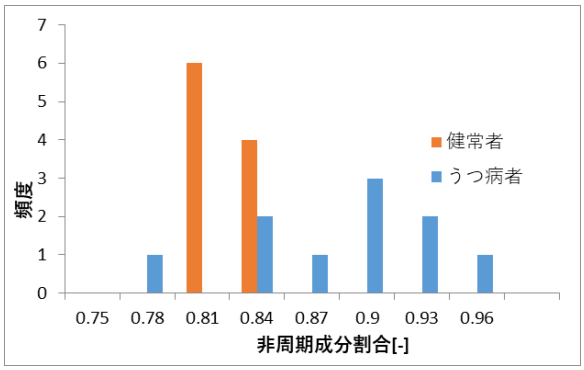


図 16. 非周期成分割合(女性)

6.2 発話内容に着目しうつ病判別器を作成

うつ病者の特徴として表 7 中 4-1.悲観的になる, 4-4. 他人への関心の低下があげられる. そのため，この特徴が発話内容に影響を及ぼす可能性を言語分析的観点から分析し，判別器を作成することを検討している．具体的には，発話音声の内容を全て文字起こししたうえで，例えばポジティブな単語には加点，逆にネガティブな単語には減点といったスコア付けを行い何か傾向が見られないか分析する. 現状は発話内容の文字起こしを行っており，以下図 17，図 18 は男女 8 人ずつの発話を全てひらがなで文字を起こしたときの文字数を示している．ここから前述の言語分析や，モーラ拍を定義し，グループ 1 の特徴量として加えるなどを検討する.

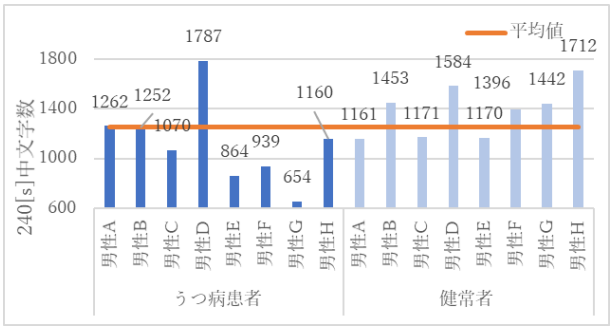


図 17. 男性音声の発話量

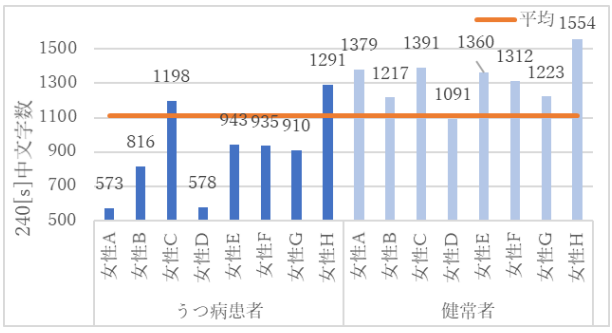


図 18. 女性音声の発話量

7. まとめ

我々が提案するうつ病判別システムに組み込むべく，うつ病の音声判別器の作成と高性能化について検討した．特徴量選択の後，作成・評価を行った判別器の判別精度は，男性データ，女性データともに 90%となった．今後は 6.1 節で述べたような，より判別に有意な特徴量の追加による判別精度の向上や，6.2 節で述べた新たな判別器の作成，特徴量選択

の最適化による判別器の汎化性能の向上についての検討が必要であると考えられる.

利益相反の有無

なし

文　献

［1］ 厚生労働省：うつ病対策について, https://www.mhlw.go.jp/kokoro/nation/dyp.html(2020 年 7 月 31 日閲覧)

［2］ 厚生労働省：自殺・うつ病等対策プロジェクトチームとりまとめについて, https://www.mhlw.go.jp/seisaku/2010/07/03.html(2020 年 7 月 31 日閲覧)

［3］ 厚生労働省:5 疾病・5 事業について, https://www.mhlw.go.jp/file/05-Shingikai-10801000-Iseikyoku-Soumuka/0000139231.pdf(2020 年 7 月 31 日閲覧)

［4］ 厚生労働省：研究内容と成果(うつ病に関する例), https://www.mhlw.go.jp/shingi/2009/08/dl/s0806-16b_0013.pdf(2020 年 7 月 31 日閲覧)

［5］ 和家 尚希, 鈴木 雅之, 長野 徹他:精神疾患診断補助に有効な発話課題と音声特徴に関する検討. 信学技報 SP2014-133(2015)

［6］ 牧優太 和田昇太 安倍和弥 他: 画像認識技術によるうつ病診断の定量化 第 38 回 日本医用画像工学会大会 (JAMIT2019) 予稿集 , OP1-17 , 2019

［7］ Maki Y, Wada S, Abe K, et al.: Developing High Performance CAD for Depression by Integrating Multiple Classifier Systems. Computer Assisted Radiology and Surgery 34rd International Congress and Exhibition, CARS2020,S204-S205, 2020

［8］ Wada S, Maki Y, Abe K, et al.: Developing High Performance CAD for Depression by Employing Image and Voice Information. Computer Assisted Radiology and Surgery 34rd International Congress and Exhibition, CARS2020, S203-S204, 2020

［9］ 忠井俊明 : ようこそ精神医学へ. ミネルヴァ書房, 2003

［10］ http://www.speech.kth.se/wavesurfer/

［11］ MathWorks Introduction to Feature Selection https://jp.mathworks.com/help/stats/feature-selection.html?lang=en(2020 年 7 月 31 日閲覧)

［12］ http://www.kki.yamanashi.ac.jp/~mmorise/world/

Developing High Performance CAD for Depression by Employing Image and Voice Information

Shota WADA*1, Yuta MAKI*1, Kazuya ABE*1, Hideya TAKEO*1, Yuichi NAGAI*2

*1 Kanagawa Institute of Technology

*2 National Cancer Center Hospital East

The number of depressed people is on the rise in Japan. In 2011, depression was added to the four major diseases, making it one of the five major diseases. This suggests the need for reliable detection of depressed people. Currently, the diagnosis of depression is mainly based on the interview. However, this lacks objective indicators, and some have questioned the accuracy of the diagnostic results. In order to solve this problem, we are developing an AI-based system that calculates the degree of certainty that a patient is depressed by focusing on facial images and voice information. The purpose of this study is to use this as an objective indicator to assist in the diagnosis of depression. This paper describes the construction of a voice discriminator that constitutes the system. A total of 42 features were created based on the patient's speech and the characteristics of the depressed person. The SVM was used to create and evaluate the discriminator, and the accuracy was 90% for the female data. However, the accuracy for male data was about 55%. Hence, these features were selected by 18 feature selection methods. The discriminator was made and evaluated again, and the accuracy was 80%, even for male data.

Key words: Depression, Voice information, Machine learning, SVM, Feature selection

敵対的生成ネットワークを用いて生成した硬性白斑画像についての検討

藤田　真穂[*1]　畑中　裕司[*2]　砂山　渡[*2]

村松　千左子[*3]　藤田　広志[*4]

要旨

糖尿病網膜症は中途失明の原因となる病であるが，早期発見によって失明を防ぐことができる．近年は眼底画像における初期病変検出に，畳み込みニューラルネットワーク (CNN:Convolutional Neural Network) が成果を挙げている．一方で，CNN を用いる際には，病変と非病変データ数の不均衡が課題となる．この課題に対する対応策として，敵対的生成ネットワーク (GAN:Generative Adversarial Networks) を用いて病変データ数を増加させる研究が行われている．本研究では多様性と安定した学習が特長の Cramer GAN を用いて，糖尿病網膜症の初期病変の一つである硬性白斑画像を生成した．Cramer GAN の採用理由は Unrolled GAN を用いて同様の実験を行った際，学習の不安定さと生成画像の多様性の欠如が課題となったためである．また，生成データと実データ分布の差異について，複数の特徴を用いて検証を行っている．さらに，生成データを用いて病変と非病変データ数の不均衡を改善することが，CNN を用いた硬性白斑の検出精度の向上に貢献するかを検証した．

キーワード : 糖尿病網膜症，畳み込みニューラルネットワーク，敵対的生成ネットワーク，不均衡データ，硬性白斑検出

1. はじめに

糖尿病網膜症は中途失明の原因となる病であるが，早期発見によって失明を防ぐことができる．近年は畳み込みニューラルネットワーク (CNN:Convolutional Neural Network) が，眼底画像からの糖尿病網膜症の初期病変検出に成果を挙げている［1］．しかし，機械学習による病変検出においては，得られる病変と非病変データの数が均衡でないことが障害となる．この課題に対して現在，敵対的生成ネットワーク (GAN:Generative Adversarial Networks) を用いて病変データを生成する研究が行われている．

本研究では多様性と安定した学習が特長とされる Cramer GAN［2］を用いて，糖尿病網膜症の初期病変の一つである硬性白斑画像を生成した．Cramer GAN の採用理由は Unrolled GAN を用いて同様の実験を行った際, 学習の不安定さと生成画像の多様性の欠如が課題となったためである．また，生成データと実データの分布の差異について，検証を行っている．さらに，生成データを用いてデータ数の不均衡を改善することが, CNN の性能向上に貢献するかについても検証を行った．

2. 実験

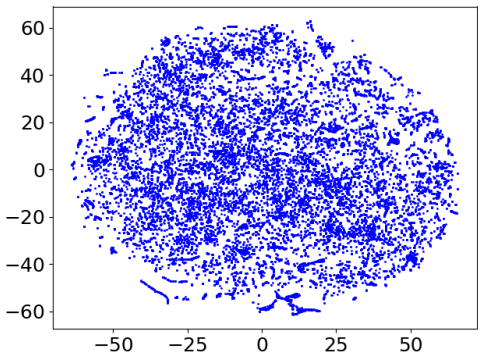
眼底画像のデータベースとして，Indian Diabetic Retinopathy Image Dataset［3］を用いた．

*1 滋賀県立大学大学院工学研究科電子システム工学専攻
〔〒522-0033 滋賀県彦根市八坂町 2500〕
e-mail: oi23mfujita@ec.usp.ac.jp
*2 滋賀県立大学工学部電子システム工学科
*3 滋賀大学データサイエンス学部
*4 岐阜大学工学部電気電子・情報工学科

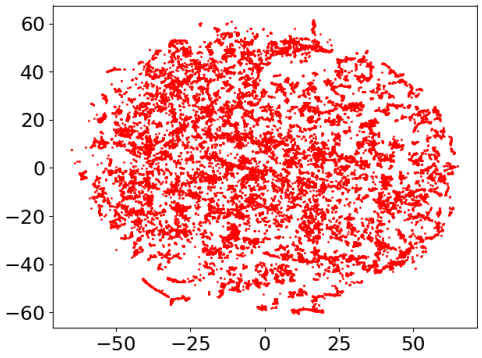
また, Cramer GAN の学習には 38 枚の画像から，無作為に硬性白斑の領域を抽出した 3,000 枚の実データを用いた．

図 1 に t-SNE を用いて描いた，実データと生成データ各 30,000 の分布図を示す．図 1 では，両データ間に大きな差異は見受けられない．

図 2 に各学習データを用いた CNN における Average Accuracy を示す．各学習データの非硬性白斑データは共通して，IDRiD から得た 30,000 枚の実データである．GAN に加えて，回転や複製によって，3,000 枚の硬性白斑画像から 27,000 枚の画像を作成した．また，図 2 に示した値は各条件において，10 回試行を行った際の平均値である．図 2 より GAN による生成データは不均衡データに対して Average Accuracy を約 3%向上させ，複製データよりも Average Accuracy 向上に貢献した．一方で，回転データ



(a) 実データ



(b) 生成データ

図 1　t-SNE による分布図

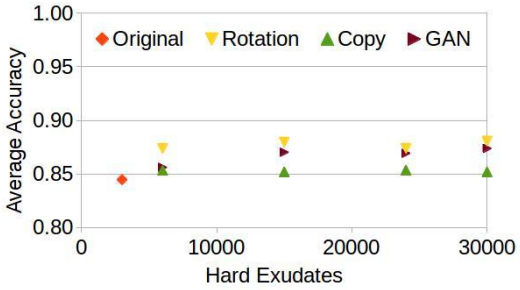


図 2 Average Accuracy と硬性白斑データ数の関係

と比較すると，生成データの貢献度は低いものとなった．

3. おわりに

生成した硬性白斑画像を用いてデータ数の不均衡を改善することで，硬性白斑検出を行う CNN の Average Accuracy は約 3%向上した．今後は，より CNN に寄与できるよう，生成データの質を高めていく．

謝辞

本研究の一部は JSPS 科研費 16K01415 および 26108005 の助成を受けたものです．

利益相反の有無

なし

文　献

［1］ Tan JH, Fujita H, Sivaprasad S, et al.: Automated Segmentation of Exudates, Haemorrhages, Microaneurysms using Single Convolutional Neural Network. Information Sciences **420**: 66-76, 2017

［2］ Bellemare M, Danihelka I, Mohamed S, et al.: The Ctamer Distance as a Solution to Biased Wasserestein Gradients. arXiv: 1705.10743v1, 2017

［3］ Porwal P, Pachade S, Kamble R, et al.: Indian Diabetic Retinopathy Image Datasets (IDRiD): A Database for Diabetic Retinopathy Screening Research. data, 2018

Investigation of Hard Exudates Image Generated by Generative Adversarial Networks

Maho FUJITA[*1], Yuji HATANAKA[*2], Wataru SUNAYAMA[*2], Chisako MURAMATSU[*3], Hiroshi FUJITA[*4]

*1 Division of Electronic Systems Engineering, Graduate School of Engineering, the University of Shiga Prefecture
*2 Department of Electronic Systems Engineering, School of Engineering, the University of Shiga Prefecture
*3 Faculty of Data Science, Shiga University
*4 Department of Electrical, Electronic & Computer Engineering, Faculty of Engineering, Gifu University

Early detection and treatment of diabetic retinopathy (DR) contribute to prevent blindness, although this disease is a main factor of blindness. Convolutional neural network (CNN) has been applied for early DR lesion detection with good results. However, it has the challenge that number of abnormal data are not balanced against normal ones. Generative adversarial networks (GAN) has been applied to solve the challenge by generating abnormal data. In this study, we augment hard exudates images, which is a one of the typical pathological features in early DR, by using Cramer GAN that has great diversity and stable learning. The reason why we use Cramer GAN is we faced problem of lacking diversity and unstable learning when researching on this theme with Unrolled GAN. We then investigate difference between generated and real data by some features. This paper discusses about the potentialities that using generated data, to redress the balance between abnormal and normal, contributes improvement of CNN which detects hard exudates.

Key words: Diabetic Retinopathy, Convolutional Neural Network, Generative Adversarial Networks, Imbalanced Data, Hard Exudates detection

Modified Cycle-Consistent Adversarial Network を用いた病理組織画像における染色度合いの正規化を伴う病理組織分類

林 大誠[*1]， 檜作 彰良[*1]， 中山 良平[*1]
児玉 良典[*2,3]，眞能 正幸[*3]，吉岡 絵麻[*4]
兼松 大介[*4]，正札 智子[*4]，金村 米博[*4]

要旨

本研究の目的は，CycleGAN に SqueezeNet を追加することにより，病理組織画像の染色度合いを正規化し，病理組織の正常／異常を分類する Modified CycleGAN を開発することである．実験試料は，グリオーマ病理組織画像 53 症例（学習用 36 例，評価用 17 例）から抽出した正常組織の 7,892 関心領域および異常組織の 21,667 関心領域である．Modified CycleGAN の学習では，入力 ROI の染色度合いを正規化する CycleGAN の学習と関心領域の正常／異常を分類する SqueezeNet の学習を交互に行う．Modified CycleGAN の ROC 曲線下の面積は 0.916 で，CycleGAN による事前の染色度合いの正規なしの SqueezeNet（0.875）および正規化ありの SqueezeNet（0.895）より有意に高い結果が得られ，その有用性が示唆された．

キーワード：病理組織画像，CycleGAN，正規化，組織分類

1. はじめに

近年，がん罹患者数の増加に伴い，病理組織診断件数も大幅に増大している．しかし，病理組織診断を実施する病理医は年々減少し，慢性的な病理医不足が問題となっている[1, 2].

この問題の解決策の一つとして，病理組織画像を対象としたコンピュータ支援診断（CAD: Computer Aided Diagnosis）に関する研究が進められてきた[3-7]．近年では，畳み込みニューラルネットワーク（CNN: Convolutional Neural Network）を用いた CAD システムの研究が活発に行われている[5-7]．Sairam らは，既存の CNN モデルである ResNet[8]を用いて，腎臓病理組織の画像特徴を自動抽出し，それらの画像特徴を DAG SVM（Directed Acyclic Graph Support Vector Machine）に入力することにより，腎細胞がん組織と正常組織を分類する手法を提案した[5]．Alom らは，様々な大きさの画像特徴抽出が可能な Inception モジュール，過去に抽出した特徴から新たな画像特徴を抽出する Recurrent convolution，さらに勾配欠損を防ぐ Residual モジュールを結合した IRRCNN（Inception Recurrent Residual Convolutional Neural Network）を提案し，脳病理組織画像の腫瘍と非腫瘍を高精度に分類できることを示した[6]．また，Bardou らは，脳病理組織画像から脳腫瘍の病種分類に適した画像特徴を抽出する CNN 構造を明らかにし，それらの画像特徴を用いた SVM による脳腫瘍の病種分類法を提案した[7].

CNN を分類問題に適用する際，データセット内の各クラスのデータ数の偏りや，クラス内の特徴分布のばらつきが，分類精度を低下させる要因となることが報告されている[9,10]．また，病理組織画像を対象とした CNN では，染色度合いのばらつきが分類精度の低下を招く大きな要因となっている[11,12].

深層学習を用いた画像生成や画像変換の一

*1 立命館大学理工学部電子情報工学科
〔〒525-8577 草津市野路東 1-1-1〕
e-mail: ri0063he@ed.ritsumei.ac.jp
*2 神戸大学大学院医学研究科
*3 国立病院機構大阪医療センター臨床検査科
*4 国立病院機構大阪医療センター臨床研究センター

つとして，敵対的生成ネットワーク（GAN: Generative Adversarial Nets）がある[13]. GAN は，Generator と Discriminator の 2 つのネットワークで構成され，Generator は学習データに類似する新しい画像を生成し，Discriminator は学習データと Generator により生成された画像を判別する．この GAN の派生形の一つとして，CycleGAN（Cycle-Consistent Adversarial Networks）がある[14]. CycleGAN は画像の色やスタイル変換に長けたモデルであり，2 つの GAN が互いの Generator で結合する構造をもつ. CycleGAN は，画像特性の異なる 2 つの画像群に対して，一方の群の画像を他方の群の画像特性に変換する順変換と，順変換した画像を元の画像特性に戻す逆変換を繰り返しながら学習を行う．この学習により，画像に含まれる構造を維持しながら，画像特性だけを変換することが可能となる．病理組織画像の染色度合いの正規化に CycleGAN を適用できるが，CycleGAN では組織分類を考慮した学習ができないことから，組織の正常／異常を分類するために重要な画像特徴が欠損する可能性がある．

そこで本研究では，病理組織画像の染色度合いを正規化する CycleGAN に正常／異常組織を分類する CNN を追加することにより，組織分類を考慮した正規化をしながら，組織分類を行う Modified CycleGAN を開発する．

2．実験試料

実験試料は，大阪医療センターで採取され，H&E（ヘマトキシリン・エオジン）染色された 53 症例のグリオーマ病理組織標本である．これらをバーチャルスライドスキャナ（NanoZoomer, 浜松ホトニクス，浜松市）により，20 倍率でデジタル化し用いた．すべての病理組織画像の異常組織領域に対し，経験豊富な病理専門医によりアノテーションが付与された．図 1 に，グリオーマ病理組織画像の例を示す．

アノテーションに基づき，224×224 画素の関心領域（ROI: Region of Interest）を，重複しないように 224 画素間隔で移動させ，ROI を抽出した．抽出した ROI のうち，病理組織を 90%以上



図 1 アノテーションが付与されたグリオーマ病理組織画像の例

含む正常 ROI 7,892 例および異常 ROI 21,667 例を提案手法の学習および評価に用いた．本研究では，53 症例のうち 36 症例から抽出された ROI を学習用データ（正常 ROI：3,956，異常 ROI：17,233），残り 17 症例から抽出された ROI を評価用データ（正常 ROI：3,936，異常 ROI：4,434）として用いた．また，学習用データの 2 割をモデル決定のための検証用データとした．

本研究は，大阪医療センターおよび立命館大学の倫理審査委員会において承認を受けている．

3．方法

1）Modified CycleGAN のネットワーク構造

Modified CycleGAN は，入力 ROI の染色度合いを正規化する CycleGAN と正規化された ROI を正常／異常組織に分類する分類器 CNN を結合したネットワーク構造である．図 2 に，Modified CycleGAN の概要図を示す．CycleGAN と分類器 CNN を統合的に学習させることで，正常／異常組織分類に有用な病理組織情報を保持した染色度合いの変換ができると期待する．

Modified CycleGAN に含まれる CycleGAN の Generator は，畳み込み層，Instance-Normalization 層，Leaky ReLU（Rectified Linear Unit）で構成される層群を 3 つ連ねた Encoder ブロック，Reflection Padding 層，畳み込み層で構成される層群を 6 つ連ねた Residual ブロック，そして転置畳み込み層，Instance-Normalization 層，Leaky ReLU で構成される層群を 3 つ連ねた Decoder

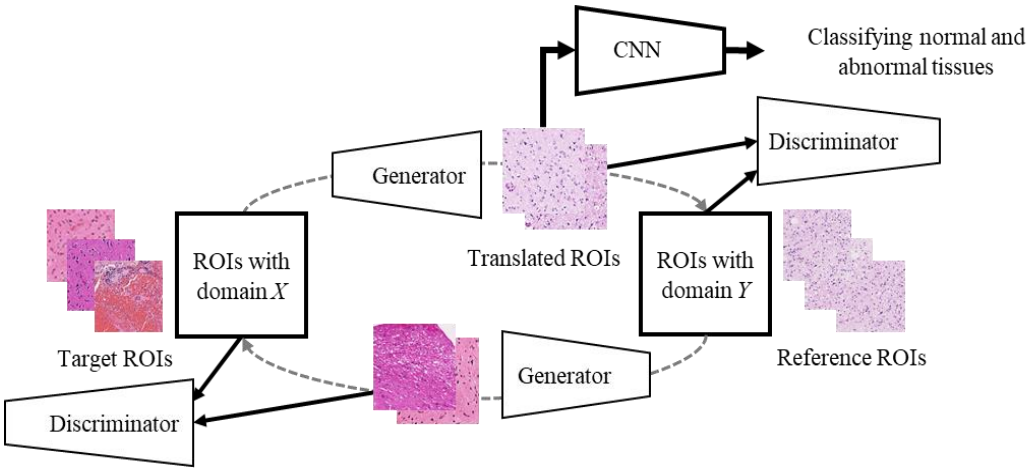


図 2 Modified CycleGAN の概要図

ブロックを順に結合した Encoder-Decoder 構造で構成される．一方，Discriminator は，畳み込み層，Instance-Normalization 層，Leaky ReLU で構成される層群を 5 つ連ねた構造である．

CycleGAN に追加する分類器 CNN のモデルを決定するために，検証用データを用いて，既存の CNN モデルによる正常／異常組織分類の精度比較を行った．ここでは，ImageNet[15]で事前学習された AlexNet[16]，GoogLeNet[17]，ResNet18[8]，MobileNet[18]，SqueezeNet[19]を用いた．そして，分類精度が最も高い CNN モデルを分類器 CNN として決定した．

2）Modified CycleGAN の損失関数

CycleGAN による画像変換と，分類器 CNN による正常／異常組織分類を統合的に学習させるため，Modified CycleGAN の損失関数を定義する必要がある．一般に CycleGAN の損失関数として，写像の変換 $G_Y: \text{domain } X \to Y, G_X: \text{domain } Y \to X$ を行う 2 つの Generator G_Y, G_X と Discriminator D_Y, D_X により定義された下記損失$\mathcal{L}$が用いられる．

$$\mathcal{L} = \mathcal{L}_{adv}(G_Y, D_Y, X, Y) + \mathcal{L}_{adv}(G_X, D_X, Y, X) + \lambda_{cycle}\mathcal{L}_{cycle}(G_X, G_Y, Y, X) + \lambda_{ident}\mathcal{L}_{ident}(G_X, G_Y, X, Y) \quad (1)$$

ここで，λ_{cycle}とλ_{ident}は重み係数であり，$\mathcal{L}_{adv}$，$\mathcal{L}_{cycle}$，$\mathcal{L}_{ident}$は，次式で与えられる．

$$\mathcal{L}_{adv}(G_Y, D_Y, X, Y) = \mathbb{E}_y[\log D_Y(y)] + \mathbb{E}_x[\log D_Y(G_Y(x))] \quad (2)$$

$$\mathcal{L}_{cycle}(G_X, G_Y, Y, X) = \mathbb{E}_x\|G_X(G_Y(x)) - x\| + \mathbb{E}_y\|G_Y(G_X(y)) - y\| \quad (3)$$

$$\mathcal{L}_{ident}(G_X, G_Y, X, Y) = \mathbb{E}_x\|G_X(x) - x\| + \mathbb{E}_y\|G_Y(y) - y\| \quad (4)$$

各ネットワークの入力として用いるx, yは，$x \in X, y \in Y$を満たす．Modified CycleGAN の損失関数$\mathcal{L}_{modified}$は，CycleGAN の損失関数$\mathcal{L}$に，分類器 CNN により正常／異常組織を分類した際の分類損失$\mathcal{L}_{classify}$を損失項として追加し，下式で定義した．

$$\mathcal{L}_{modified} = \mathcal{L}_{adv}(G_Y, D_Y, X, Y) + \mathcal{L}_{adv}(G_X, D_X, Y, X) + \lambda_{cycle}\mathcal{L}_{cycle}(G_X, G_Y, Y, X) + \lambda_{ident}\mathcal{L}_{ident}(G_X, G_Y, X, Y) + \lambda_{classify}\mathcal{L}_{classify}(G_Y, X) \quad (5)$$

ここで，$\mathcal{L}_{classify}$は Binary Cross Entropy Loss である．本研究では，重み係数$\lambda_{cycle}, \lambda_{idt}, \lambda_{classify}$をそれぞれ$10.0, 5.0, 0.5$ で与えた．

3）Modified CycleGAN の学習

Modified CycleGAN の学習では，まず，学習データから染色度合いの基準となる病理組織画像を 1 症例選択し，その基準画像から抽出された 209 ROI を Reference ROI して domain Y の画像群に設定した．本研究では，基準画像として，染色度合いが比較的薄い画像を選択した．そして，Reference ROI 以外の全ての ROI を Target 画像として domain X の画像群に設定し，Reference ROI の染色度合いに近づける正規化を行った．そして，分類器 CNN により，domain

X から domain Y へ変換した translated ROI を正常／異常組織に分類した．Modified CycleGAN の学習パラメータは，エポック数: 50，バッチサイズ: 1，初期学習率: 1.0×10^{-5}，オプティマイザ: Adam ($\beta_1 = 0.5,\ \beta_2 = 0.999$) とした．

4. 結果と考察

検証データに対する既存 CNN モデルの正常／異常組織分類の結果，SqueezeNet の Receiver Operating Characteristic（ROC）曲線下の面積（AUC: Area under the ROC curve）は 0.973 で，AlexNet（0.871, P < .001），GoogLeNet（0.879, P < .001），ResNet18（0.848, P < .001），MobileNet（0.756, P < .001）より有意に高かった．したがって，SqueezeNet を CycleGAN の分類器 CNN として決定した．SqueezeNet 以外の既存 CNN モデルは，細胞や細胞核のような細かな構造情報をプーリング処理の繰り返しにより，欠損してしまった可能性がある．一方, SqueezeNet は，Fire モジュールを導入することにより，プーリングによるダウンサンプリングの回数を可能な限り減らしているため，その欠損を回避できたと考える．

図 3 に，CycleGAN と Modified CycleGAN による染色度合いの正規化画像の比較を示す．CycleGAN，Modified CycleGAN により，Target ROI の染色度合いを Reference ROI に近づけることができることを確認した．また，CycleGAN と Modified CycleGAN による変換画像間に，視

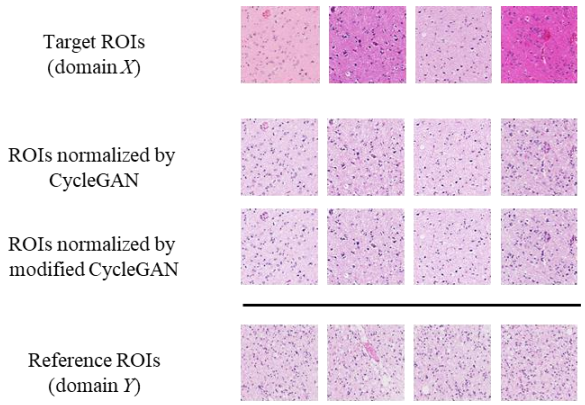


図 3 CycleGAN と Modified CycleGAN による正規化画像の比較

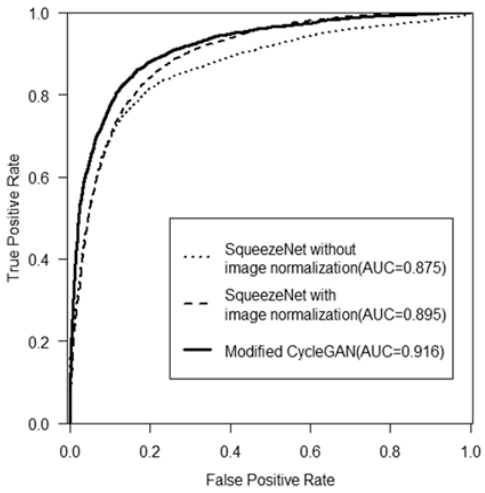


図 4 ROC 曲線による比較結果

覚的に大きな差異は見られなかったが，Modified CycleGAN は CycleGAN より細胞核のコントラストを向上する傾向があった．

図 4 に評価用データに対する Modified CycleGAN の ROC 曲線を示す．また，精度比較のため，CycleGAN による事前の染色度合いの正規なしの SqueezeNet（SqueezeNet without image normalization）および正規化ありの SqueezeNet（SqueezeNet with image normalization）の ROC 曲線も示す．また，Modified CycleGAN の AUC は 0.916 で, 事前の染色度合いの正規なしの SqueezeNet（0.875, P<.001）および正規化ありの SqueezeNet（0.895, P<.001）より有意に高い結果が得られた．事前の染色度合いの正規なしの SqueezeNet の AUC が，CycleGAN による事前の正規化により改善されたことから，染色度合いのばらつきよる分類精度の低下を回避できたと考える．また，事前の染色度合いの正規ありの SqueezeNet より Modified CycleGAN の AUC が有意に高かったことから，Modified CycleGAN よる画像変換は，正常／異常組織分類に重要な病理組織情報を保持した画像変換が可能であったと考える．

5. まとめ

本研究では，染色度合いを正規化する CycleGAN と正常組織と異常組織を分類する

SqueezeNet を組み合わせることにより，Modified CycleGAN を開発した．正常組織と異常組織を分類するための重要な画像特徴を損失することなく染色度合いを正規化できる Modified CycleGAN は，高い分類精度を達成した．

謝辞

本研究で使用した病理組織標本を提供して頂いた国立病院機構大阪医療センターの皆様に厚く感謝の意を表します．

利益相反の有無

なし

文　献

[1] 厚生労働省,: 平成 30 年（2018）人口動態統計（確定数）の概況. https://www.mhlw.go.jp/toukei/saikin/hw/jinkou/kakutei18/index.html, Accessed 27 July 2020.

[2] Cancer Registry and Statistics: Vital Statistics Japan. Cancer Information Service, National Cancer Center, https://ganjoho.jp/reg_stat/statistics/dl/index.html, Accessed 27 July 2020.

[3] 檜作彰良, 中山良平.: 病理組織画像解析の研究動向. 医用画像情報学会雑誌, 36(2), 53-58, 2019.

[4] Nahid, Abdullah-Al, Aaron Mikaelian, et al: Histopathological breast-image classification with restricted Boltzmann machine along with backpropagation. Biomedical Research, 2018.

[5] Tabibu, Sairam, P. K. Vinod, et al.: Pan-Renal Cell Carcinoma classification and survival prediction from histopathology images using deep learning. Scientific reports 9.1 (2019): 1-9, 2019.

[6] Alom, Md Zahangir, MS Nasrin, et al.: Breast cancer classification from histopathological images with inception recurrent residual convolutional neural network. Journal of digital imaging 32.4 (2019): 605-617, 2019.

[7] Bardou, Dalal, Kun Zhang, Sayed Mohammad Ahmad.: Classification of breast cancer based on histology images using convolutional neural networks. IEEE Access 6 (2018): 24680-24693, 2018.

[8] He, Kaiming, S Ren, et al.: Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition: 770-778, 2016.

[9] Passalis, Nikolaos, Kanniainen, et al.: Deep adaptive input normalization for price forecasting using limit order book data. arXiv preprint arXiv:1902.07892 ,2019.

[10] Bettinger Ch, Zimmermann HW, “I. Spectroscopic and physico-chemical properties of hematoxylin and hematein,” Histochem 95(1991): 279-288, 1991.

[11] Shaban, M. Tarek, C., Navab, et al.: Staingan: Stain style transfer for digital histological images. 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019): 953-956, IEEE, 2019.

[12] Zhou, Niyun, X Han, et al.: Enhanced Cycle-Consistent Generative Adversarial Network for Color Normalization of H&E Stained Images. International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham: 694-702, 2019.

[13] Radford, Alec, Chintala.: Unsupervised representation learning with deep convolutional generative adversarial networks. arXiv preprint arXiv:1511.06434, 2015.

[14] Zhu, Jun-Yan, P Isola, et al.: Unpaired image-to-image translation using cycle-consistent adversarial networks. Proceedings of the IEEE international conference on computer vision: 2223-2232, 2017.

[15] J. Deng, W. Dong, R. Socher, et al.: ImageNet: A Large-Scale Hierarchical Image Database. In CVPR09, 2009.

[16] Krizhevsky, Alex, Ilya Sutskever, et al.: Im

agenet classification with deep convolutiona l neural networks. Advances in neural info rmation processing systems: 1097-1105, 20 12.

[17] Szegedy, Christian, Y Jia, et al.: Going de eper with convolutions. Proceedings of the IEEE conference on computer vision and p attern recognition: 1-9, 2015.

[18] Sandler, Mark, M Zhu et al.: Mobilenetv2: Inverted residuals and linear bottlenecks. P roceedings of the IEEE conference on com puter vision and pattern recognition: 4510-4520, 2018.

[19] Iandola, Forrest N., MW Moskewicz, et al.: SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and< 0.5 MB model size. arXiv preprint arXiv:1602.07360, 2016.

Computerized Scheme for Classifying Normal and Abnormal Tissues in Pathological Images Using Modified Cycle-Consistent Adversarial Network

Daisei Hayashi[1], Akiyoshi Hizukuri[1], Ryohei Nakayama[1],
Yoshinori Kodama[2,3], Masayuki Mano[3], Ema Yoshioka[4],
Daisuke Kanematsu[4], Tomoko Shofuda[4], Yonehiro Kanemura[4]

[1] Graduate School of Science and Engineering, Ritsumeikan University, 1-1-1 Noji-higashi, Kusatsu, Shiga 525-8577, Japan
[2] Graduate School of Medicine, Kobe University, 1-1 Kusunoki, Chuo-Ku, Kobe, Hyogo 650-0017, Japan
[3] Clinical Laboratory, National Hospital Organization Osaka National Hospital, 2-1-14 Houenzaka, Chuo-Ku, Osaka, Osaka 540-0006 Japan
[4] Institute for Clinical Research, National Hospital Organization Osaka National Hospital, 2-1-14 Houenzaka, Chuo-Ku, Osaka, Osaka 540-0006 Japan

The purpose of this study was to develop a modified CycleGAN which normalizes staining intensity to be suitable for classifying tissues by adding a CNN for tissue classification to a CycleGAN for image normalization. Our database consisted of 53 brain pathological images digitized at the magnification ratio of 20 times. Those specimens were stained with Hematoxylin and Eosin. From those images, 21,677 abnormal and 7,892 normal regions of interest (ROIs) with 224×224 pixels were extracted. ROIs for 36 patients and for the remaining 17 patients were used as training dataset and test dataset, respectively. The modified CycleGAN consisted of an original CycleGAN for image normalization and a CNN based on a pre-trained SqueezeNet for tissue classification. The original CycleGAN and the CNN were trained alternately and repeatedly. Area under the ROC curve for the modified CycleGAN was 0.916, showing significantly greater than those for SqueezeNet without image normalization (0.875, $p<.001$) and SqueezeNet with image normalization by the original CycleGAN (0.895, $p<.001$). The modified CycleGAN which can normalize the staining intensity without the loss of important image features for classifying normal and abnormal tissues was able to yield high classification accuracy in pathological images.

Key words: Digital pathological images, CycleGAN, Normalization, Tissue classification

SVM と時空間モデルを用いた嚥下時 X 線動画からの喉頭蓋の抽出に関する基礎的検討

杉田　絢美[*1]　藤中　彩乃[*1,2]　目片　幸二郎[*3]

滝沢　穂高[*1]　工藤　博幸[*1]

要旨

嚥下障害のうち誤嚥は，喉頭蓋という軟骨組織が大きく関係している．喉頭蓋は気道の入り口にあり，気道に蓋をすることで嚥下物が気道に侵入することを防ぐ役割を持つが，その構造および動体解析は進んでいない．本研究では，SVM および喉頭蓋の時空間モデルを用いて嚥下時 X 線動画から喉頭蓋を抽出する手法を提案する．まずピクセルベースの SVM を適用し，喉頭蓋の候補領域の抽出を行う．喉頭蓋が気道に蓋をする動きは 1 秒以内と速く，さらに造影剤が喉頭蓋に接触しているフレームでは喉頭蓋が確認できない．そこで，喉頭蓋の時空間モデルを候補領域に当てはめることによって喉頭蓋の位置と姿勢を推定する．実際の X 線動画に本手法を適用した結果を示す．

キーワード：嚥下時 X 線動画，喉頭蓋，SVM，嚥下障害

1. はじめに

嚥下とは人が食べ物を飲み込むという動作のことである[1]．嚥下障害は，加齢による嚥下能力の低下のみでなく，脳梗塞や脳出血などの脳血管障害によっても引き起こされることがある[2]．また，頚椎の手術により嚥下障害を発症した症例も報告されている[3]．さらに，嚥下障害になると，食べ物を飲み込むことが困難になる嚥下困難や食物が気道に入る誤嚥などの症状が現れ，栄養失調や肺炎に陥ることがあるが，嚥下障害が発生するメカニズムは未だ解明されていない．特に誤嚥は，喉頭蓋という軟骨組織と大きく関係している．喉頭蓋は気道の入り口にあり，嚥下の際に気道に蓋をして嚥下物が気道へ侵入することを防ぐ役割がある．しかし，嚥下の際に喉頭の挙上が不十分な場合，喉頭蓋の後屈も不十分となり気道に蓋をすることができず誤嚥を引き起こしてしまう．また，嚥下運動は食道内の食物の有無にかかわらず約 1 秒間で運動が終了するが，嚥下圧が弱い場合，食道に食物が残ったまま嚥下運動が終了してしまうため誤嚥を引き起こすことがわかっている[4]．

そこで，臨床現場で一般的に使用されている嚥下時 X 線動画を用いて喉頭蓋の位置と姿勢を推定する手法を提案する．

2. 喉頭蓋近傍マスクの作成

Fujinaka ら[5]の提案した手法を基に作成した頸部マスク画像を用いて，SVM による喉頭蓋領域の抽出を行う．

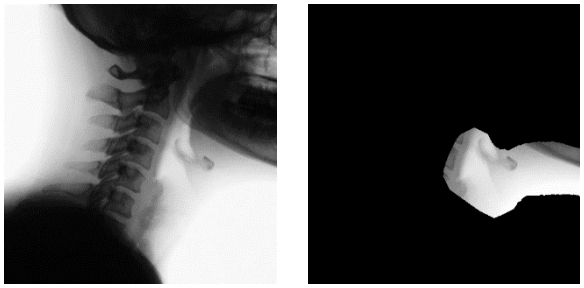
嚥下時 X 線動画の各スライスに判別分析法を適用し，面積の大きい２つの領域を背景領域とする．この背景領域の各重心を求め，その重心

*1 筑波大学
〔〒305-8573 つくば市天王台 1-1-1〕
*2 日本学術振興会特別研究員
*3 神戸赤十字病院

を通る直線を背景重心線とする．背景領域でない領域の距離変換を適用し，背景領域からの距離を算出する．画像の上辺から下辺に，距離値が最大となるように直線を引き，首中心線とする．背景領域の境界線の上部から，3 点をそれぞれランダムに選択し, RANSAC を用いた放物線の当てはめを行い，下顎の放物線とする．

背景重心線と首中心線，下顎の放物線で囲まれた領域に Opening 処理を 10 回と Closing 処理を 5 回適用し, 面積最大の領域をマスクとする．

図 1 はマスク画像作成を示しており，(a)はあるスライスの原画像, (b)は ROI の整形後の最終的に作成されたマスク画像を表している．



(a) (b)

図 1 マスク画像の抽出図.

3. ピクセルベースの SVM による喉頭蓋の粗領域の抽出

3.1 SVM の学習

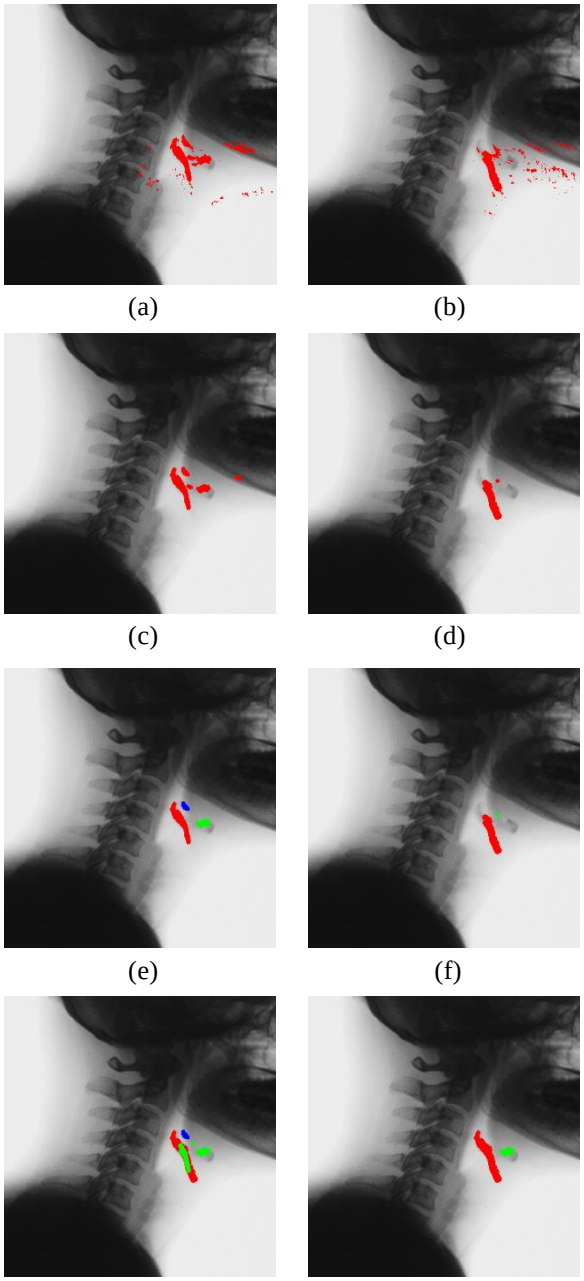
各症例で，喉頭蓋が気道に蓋をしていることが目視で確認できるフレームとその前後およそ 30 フレームの 3 つのフレームを選択し，喉頭蓋領域を手動で抽出する．嚥下時に喉頭蓋の上部は形を大きく変えるのに対して，下部は形がほぼ一定なことから，上部と下部に分けて学習する．喉頭蓋上下部それぞれの領域内の画素を中心とする$N \times N$ピクセルのパッチを設定し，濃度値の平均と分散がそれぞれ 0 と 1 になるように正規化し，ベクトル化したものを喉頭蓋の上下部クラスとする．非喉頭蓋クラスはマスク内かつ喉頭蓋領域に属さない画素をランダムに選択し, 正規化とベクトル化を行う. ただし，非喉頭蓋クラスのパッチ数と喉頭蓋クラスのパッチ数が同じになるようにし，喉頭蓋領域の近傍ピクセル（距離$N/2$ピクセル以内）は非喉頭蓋クラスの中心画素として選択しないようにする．このクラスを基に，喉頭蓋の上部とその他で学習と予測、喉頭蓋の下部とその他で学習と予測を行う．

3.2 SVM の予測

テスト症例の全フレームからマスクを抽出し，そのマスクの各画素にパッチを設定し，ベクトル化および正規化を行い，学習済みの SVM で各クラスの帰属確率を計算する．喉頭蓋クラスの帰属確率が 0.5 以上の画素を喉頭蓋の粗領域として抽出する．

4. 喉頭蓋領域の抽出精度の向上

SVM の識別結果には小さな誤抽出された領域が数多く残っている．この誤抽出領域を削除するために，Opening 処理を各フレームで行うことで，小領域の削除を行う．面積の大きい順の上位 3 領域のみ選択することより，喉頭蓋の領域は残したまま，誤抽出領域を減少させる．喉頭蓋の上部と下部の結果画像を重ね合わせる．余分な小領域が生成されてしまった場合は, 面積の大きい上位 2 領域を選択することで，喉頭蓋の候補領域を 2 領域まで限定した．図 2 は喉頭蓋領域の抽出精度の向上過程を示しており，(a)はあるスライスの喉頭蓋上部の抽出結果, (b)はあるスライスの喉頭蓋下部の抽出結果, (c)は(a)に Opening 処理を適用した画像, (d)は(b)に Opening 処理を適用した画像，(e)は(c)に 4 近傍ラベリング処理と面積の大きい順の上位 3 領域を抽出した画像, (f)は(d)に 4 近傍ラベリング処理と面積の大きい順の上位 3 領域を抽出した画像，(g)は(e)と(f)を重ね合わせた画像，(h)は(g)に 8 近傍ラベリング処理と面積の大きい順の上位 2 領域を適用した画像である．



(a) (b) (c) (d) (e) (f) (g) (h)

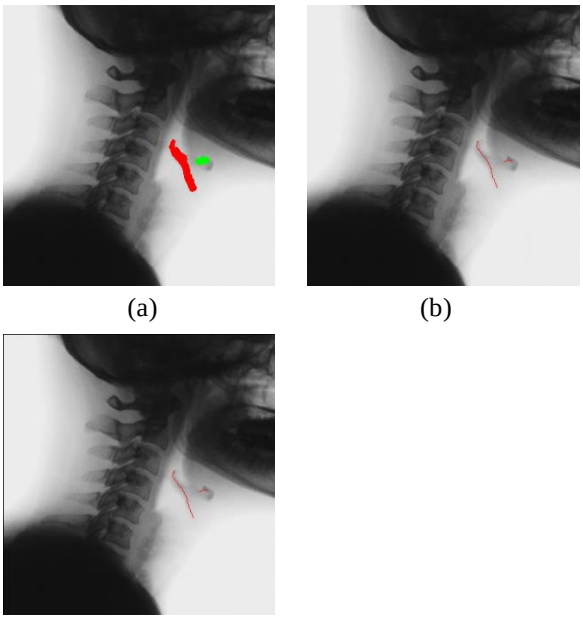
図 2 喉頭蓋領域の抽出精度の向上過程図.

5. 時空間モデルを用いた喉頭蓋領域の位置推定

画像上で喉頭蓋が確認できないフレームに対しても時空間モデルを用いて喉頭蓋の位置推定を行う.

5.1 前処理

4 章までの結果から得られる喉頭蓋の候補点はかなり多いため，候補点の数を限定させる必要がある．そこで，形状や結合関係を保持したまま喉頭蓋の候補点を減少させるためにヒルディッチの細線化処理[6]を行う．また，本研究で使用している嚥下時 X 線動画では，撮影時に被験者が動いている症例が多くある．そこで，栗田[7]の提案した先行手法で平行移動と回転を行い，0 フレーム目を基準として喉頭蓋の位置合わせを行う．図 3 は前処理過程を示しており，(a)はあるスライスの喉頭蓋の抽出精度向上処理後の画像, (b)は(a)にヒルディッチの細線化処理を適用した画像, (c)は(b)に平行移動と回転を適用した画像である.



(a) (b) (c)

図 3 前処理の過程図.

5.2 喉頭蓋の時空間モデル

本研究では，(x, y, t)の 3 次元で喉頭蓋の動きを表した物理モデルのことを時空間モデルと呼ぶ．xは x 座標，yは y 座標，tはフレーム数を表している．まず時空間モデルの作成方法は，喉頭蓋の上端の点 A と喉頭蓋の上部と下部を分ける点 B を結んだ線分 AB と点 B と喉頭蓋の下端の点 C を結んだ線分 BC の二つの線分で喉頭蓋を表現する．この 2 線分モデルを 10 フレームごとに作成し，線分 BC は移動しないと仮定した．また，目視で確認できないフレームについては前後の座標から線形補間した.

学習データの各症例全ての特徴を保持したまま 1 つの平均時空間モデルを作成するために，まず 0 フレーム目の喉頭蓋を上部と下部に分ける点を原点に平行移動させ，各点の平均座標を求めることで平均時空間モデルを作成する．

図 4 は作成した時空間モデルを示している．

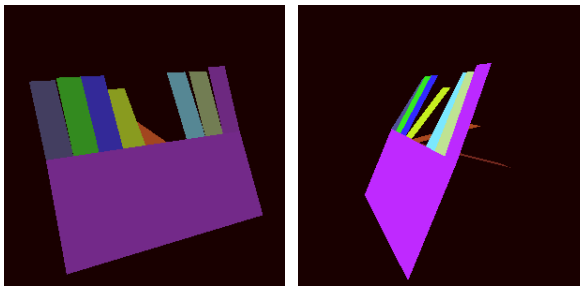


図 4 時空間モデル．

5.3　喉頭蓋の位置推定

作成した時空間モデルをテストデータで使用した症例に適用し，山下り法により位置推定を行う． 0 フレーム，70 フレームと喉頭蓋が気道に蓋をするフレームのそれぞれ喉頭蓋の上端と下端の 6 点の座標を SVM の識別結果から手動で求める．そして，これに対応する時空間モデルの 6 点との二乗距離の合計を求める．この二乗距離の合計をエネルギー関数とし，山下り法により最小化し，最適な時空間モデルを求めた．

6. 実験

6.1　実験条件

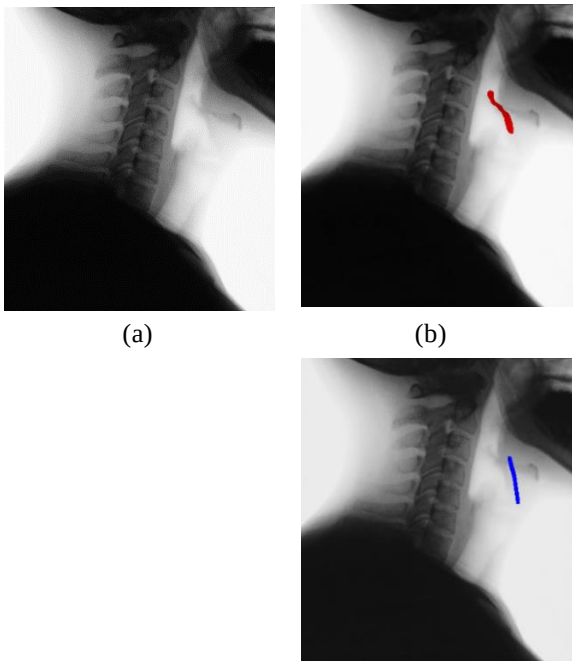
本研究では全 19 症例のうち，9 症例と 10 症例のセットに分け，交差検証を行った．

また，本研究では SVM のカーネル関数として放射基底関数(RBF)を使用する．SVM の学習の際に使用した ROI のサイズは 15×15 とし，学習のパラメータは$(C,\gamma) = (2^3, 2^{-8})$とした．これらのパラメータは実験的に決定した．

なお，本研究は神戸赤十字病院の倫理委員会の承認を受けて行われた．

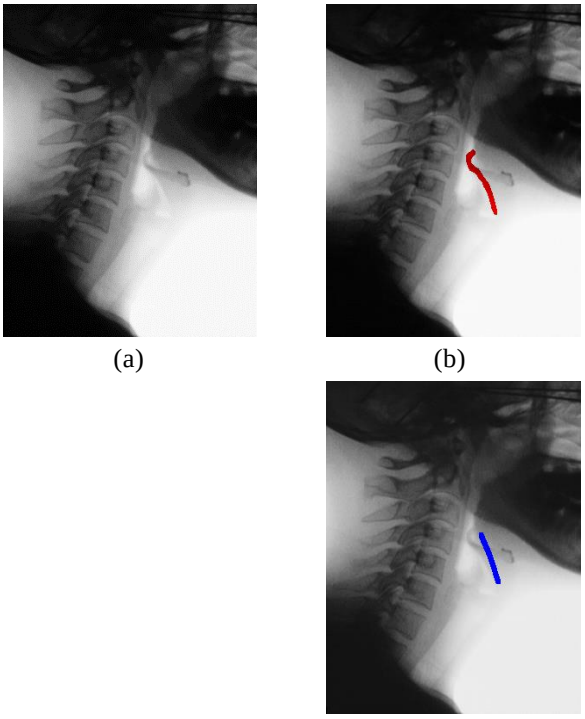
6.2　実験結果

図 5 は症例 No.13 の(a)は原画像，(b)は正解画像，(c)は結果画像であり，図 6 は症例 No.18 の 14 フレームの(a)は原画像，(b)は正解画像，(c)は結果画像であり， 図 7 は 44 フレームの(a)は原画像，(b)は正解画像，(c)は結果画像であり，図 8 は 70 フレームの(a)は 70 フレームの原画像，(b)は正解画像，(c)は結果画像である．青線は喉頭蓋を示している．



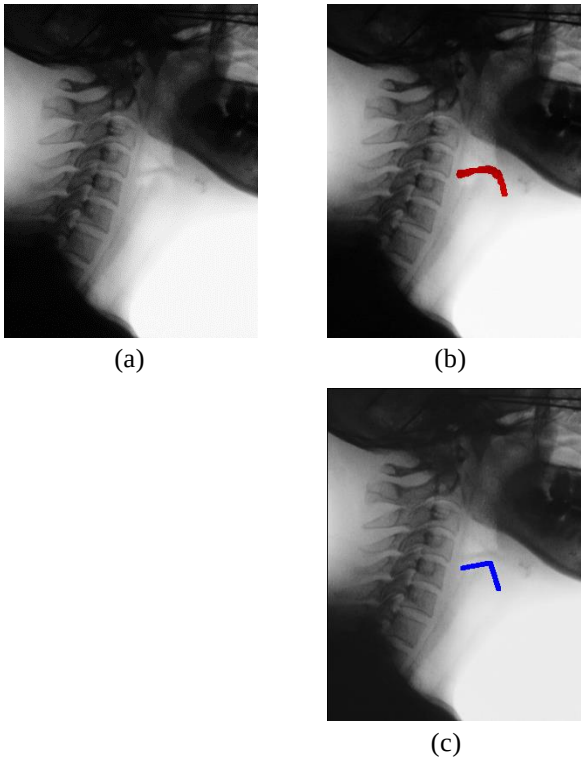
(a)　(b)　(c)

図 5 症例 No.13 の結果画像．



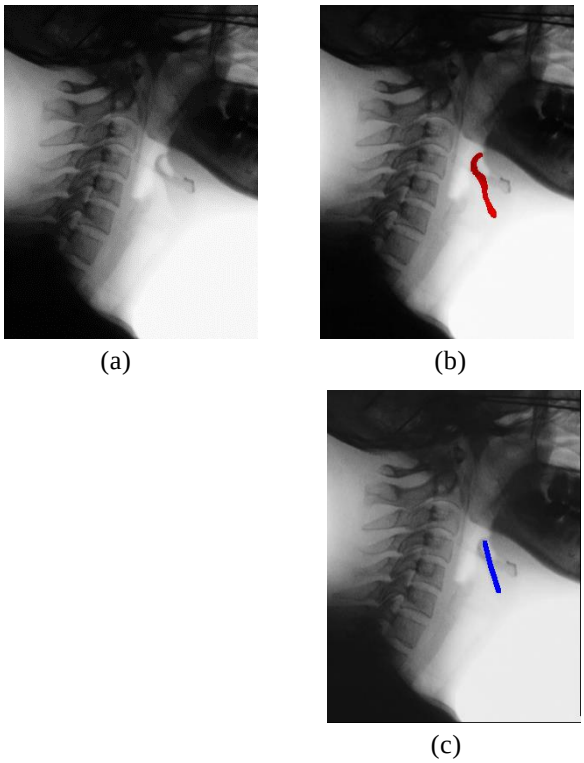
(a)　(b)　(c)

図 6 症例 No.18 の 14 フレームの結果画像．



(a) (b) (c)

図 7 症例 No.18 の 44 フレームの結果画像.



(a) (b) (c)

図 8 症例 No.18 の 70 フレームの結果画像.

7. 考察

症例 No.13 で時空間モデルと正解に大きなずれが発生した．喉頭蓋が気道に蓋をするフレームは，喉頭蓋が速い動きであるためインターレースが生じやすく SVM の識別精度が低い．そのため，現在の SVM の識別後の後処理では喉頭蓋領域が削除されてしまった．また，SVM の識別結果において一部分の喉頭蓋しか抽出できておらず，また顎領域が大きく抽出されてしまっているため，喉頭蓋の上部の領域が削除されてしまい，喉頭蓋が抽出できていない．このような誤差を減らすために, SVM の識別精度の向上，後処理の手法の改良が必要であると考えられる．

また，症例 No.18 は，概ね正しい位置で喉頭蓋の抽出ができている．しかし，現時点で喉頭蓋のモデルを 2 つの線分で表現しているが，実際の喉頭蓋の構造は時空間モデルよりも複雑であるため，時空間モデルの喉頭蓋の再現度の向上が必要であると考えられる．

8. まとめ

本研究では SVM による喉頭蓋領域の抽出と時空間モデルを用いた喉頭蓋の位置推定を行なった. 今後の課題は, SVM の識別精度の向上，後処理の改良，喉頭蓋の時空間モデルの自動当てはめと，再現度を向上させることなどがある．

謝辞

研究遂行にあたり貴重なご協力を賜った京都大学大学院医学研究科人間健康科学系専攻リハビリテーション科学コース 松林潤氏，岡山大学整形外科 瀧川朋亨氏，神戸赤十字病院整形外科 戸田一潔氏，神戸赤十字病院整形外科 伊藤康夫氏に深謝いたします．

本研究の一部は JSPS 科研費 20J10293 の助成を受けたものです．

利益相反の有無

「なし」

文 献

[1] 日本気管食道科学会：嚥下障害．http://www.kishoku.gr.jp/public/disea

se04.html
［2］ 藤雄一 : 脳血管障害症例の嚥下動体. 耳鼻と臨床，vol28,No.6,pp1126-1160,1982
［3］ 廣田隆一，高ノ原恭子，安江友世，馬場均，長谷斉，久育男 : 上喉頭神経内枝麻痺を伴った頚椎前方手術後嚥下障害の 2 例. 耳鼻と臨床， vol55, suppl.Vo.2, ppS142-S150,2009
［4］ 広戸幾一郎，小宮山荘太郎，渡辺宏 : 誤嚥の手術的治療. 耳鼻と臨床，vol.27,No.2,pp.265-271,1981
［5］ Ayano Fujinaka, Kojiro Mekata, Hotaka Takizawa, Hiroyuki Kudo: Automated Segmentation of Cervical Intervertebral Disks from Videofluorography Using a Convolutional Neural Network and its Performance Evaluation. In: Journal of Signal Processing Systems, Springer, Vol.9, No.3, pp.299-305, 2020
［6］ 田村秀行 : コンピュータ画像処理. オーム社
［7］ 栗田尚征：テンプレートマッチングを用いた頸部 X 線透視動画における基準座標系の設定に関する基礎的検討. 電子情報通信学会技術研究報告，IEICE-MICT2019-35,IEICE-MI2019-62, Vol.IEICE-119,No.263(MICT), No.264(MI), 2019

Preliminary study on extraction of epiglottises from videofluorography with SVM and spatiotemporal models

Ayami SUGITA[*1], Ayano FUJINAKA[*1,2], Kojirou MEKATA[*3], Hotaka TAKIZAWA[*1], Hiroyuki KUDO[*1]

*1 University of Tsukuba.

*2 Research Fellowship for Young Scientists, Japan Society for the Promotion of Science.

*3 Kode Red Cross Hospital.

In this paper, we proposed an extraction method of epiglottises in videofluorography (VF). First, the candidate regions of epiglottises are extracted by using pixel-based SVM. In general, epiglottises move fast, and therefore they cannot be observed in several frame images of VF. In order to predict the positions of such unobservable epiglottises, spatiotemporal models are introduced. The proposed method is applied to actual VF, and several experimental results are shown.

Key words: videofluorography, epiglottis, SVM, dysphagia

U-Net を用いた X 線動画からの頸椎椎間板の抽出に関する基礎的検討

郡司　絵莉華[*1]　藤中　彩乃[*1*2]　目片　幸二郎[*3]

滝沢　穂高[*4]　工藤　博幸[*4]

要旨

栄養低下や誤嚥を引き起こすことがある嚥下障害の詳細なメカニズムは未だに解明されていない．そこで本研究では頸部構造物を動態解析し，嚥下障害との関係を調査することを目指し，まず椎間板の抽出を行う． X 線動画の各フレームに，モルフォロジー演算と Random sample consensus (RANSAC)による放物線あてはめ処理などを適用し，頸部領域を抽出する．次に，医用画像のセグメンテーションを目的に開発された U-Net を適用し，頸椎椎間板を抽出する．一般に深層学習は一般的に学習データの数が多いほどより良いモデルを得られることが知られている．そこで，アフィン変換を用いて回転・平行移動によるデータ拡張を行う．実際の X 線動画に適用した実験結果を示す.

キーワード : X 線画像, 頸椎椎間板, U-Net，データ拡張, RANSAC

1. はじめに

嚥下障害は食べ物を飲み込むことが困難になることによる栄養低下や，むせることにより誤嚥性肺炎を引き起こすことがある[1]. 特に誤嚥性肺炎は，日本において高齢化に伴い増加する疾病の一つであり，厚生労働省の調査によると, 死因別死亡数の第 7 位である[2]. 嚥下障害の原因としては，脳血管障害等の疾病や頸部手術の副作用などであることが知られているが[3][4], そのメカニズムは未だに解明されていない．そこで，我々は頸椎構造物と嚥下障害との関係を調査するために嚥下時 X 線動画を用いて頸部構造物の抽出及び動態解析からの嚥下障害の解析を試みている．本報告では頚部椎間板の抽出を行う.

*1 筑波大学大学院システム情報工学研究群情報理工学位プログラム
〔〒305-8573 つくば市天王台 1-1-1〕
*2 日本学術振興会特別研究員
*3 神戸赤十字病院
*4 筑波大学システム情報系

2. 頸部領域の抽出

先行研究[5]で提案した次の (1) から (6) の手順に基づき，首中心線と頭部・肩部のおおよその境界線を特定する.

(1) 判別分析法による人体と背景分離
(2) Opening 処理による領域の整形
(3) 背景領域の選択
(4) 背景重心線の算出
(5) 距離変換を用いた首中心線の算出
(6) RANSAC による放物線当てはめ

図 1 は実際の症例に適用した結果を示している．緑色の直線は首中心線を示し，青色の放物線は頭部と肩部のおおよその境界線を示している．

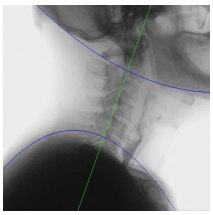


図 1　首中心線と頭･肩部のおおよその境界線

(5),(6) の結果，得られた首中心線と頭部・肩部の放物線の交点をそれぞれ p_{head} と $p_{shoulder}$ とする．p_{head} と $p_{shoulder}$ の中点を首中心点 p_{center} として求める (図 2). p_{center} を中心とする 260 × 320 の画像にクロップする．図 3 の赤線で囲われた枠内がクロップ領域となる．症例によってクロップ領域が画像外にはみ出ることがあるが，はみ出し領域は除外する．

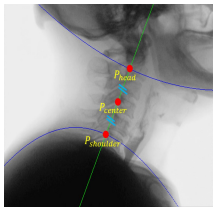


図 2 首中心点

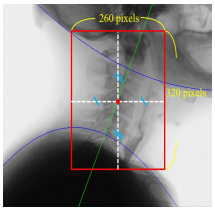


図 3 クロップ画像

3. U-Net

前節で得られた画像を，U-Net に適用して学習を行う．U-Net は 2015 年に Ronneberger らが医用画像のセグメンテーションのために開発したネットワークである[8]. U 字型の構造をしており，前半部分では畳み込みとプーリングにより縮小して特徴情報を学習し，後半部分は逆畳み込みと連結パスにより前半で失われた位置情報を持ちながら膨張して学習する．図 4 では，本研究で使用した U-Net の構造を示したものである．

4. データ拡張

U-Net は少数のデータセットに対しても学習ができるが，精度向上にはデータ拡張が有効

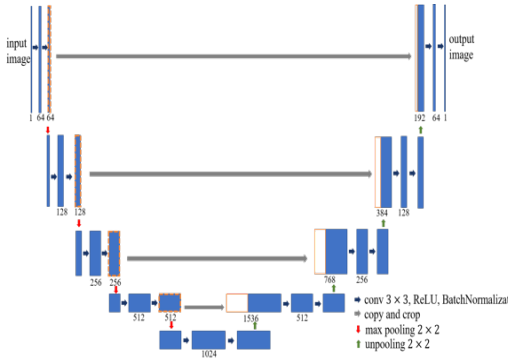


図 4　U-Net のアーキテクチャ

である．

データ拡張は，幾何学的変換やカーネルフィルタの適用といった画像処理を用いた手法が使用される．近年では，GANs[7] やニューラルネットワークによる画風変換が使用される手法もある[8]. データセットの症例ごとに考え得る変動として椎間板の位置と傾きがある．そこで，位置の変動に対応するために平行移動，傾きの変動に対応するために回転移動によるデータ拡張を行う．用いるデータセットでは図 5 のように，首の中心点を通り画像に対して垂直な直線 1 から角度 θ (≥ 0) の間に首中心線があるため, 0 度から θ 度の間でデータ拡張を行う．これは首の傾きに着目したものであるが，椎間板の傾きを考えたとき，左側に傾いているものもある．そこで，図 6 は椎間板の傾きに対応した範囲でデータ拡張を行う．

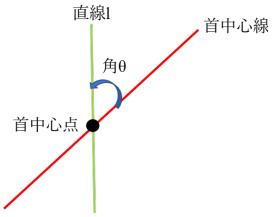


図 5 垂直方向までの回転の様子

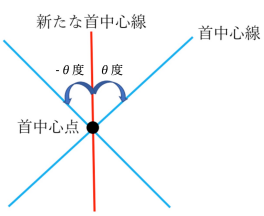


図 6 直線 l までの回転の様子

さらに，首中心点は計算によって求まるものであり，一定の位置にないため水平方向，垂直方向にそれぞれ平行移動してデータ拡張を行

い，位置の変動に対応させる．

5. 実験と結果

5.1 データセット

X 線動画 58 症例を実験に用いる．各症例から舌骨最大挙上フレームの椎間板の位置を特定し，椎間板を手動抽出し，正解データとする．U-Net による学習を行うにあたって，全 58 枚を約 4 : 1 : 1 の 割合で学習・検証・テストデータになるよう分けた．フレーム画像を 256 × 256 の同一サイズに変換した．なお，本研究は神戸赤十字病院の倫理委員会の承認を受けて行われた．

5.2 パラメータ

U-Net のパラメータは表 1 のように設定し，汎化誤差が最小となるエポックで得られた結果を採用する．データ拡張については，2.2 節において平行移動の最大ピクセル d_{max} を 30，回転の最大角度 θ_{max} を 30 度とする．

表 1 学習に関するパラメータ

コスト関数	バイナリクロスエントロピー
勾配降下法の最適化アルゴリズム	Adam
最大エポック数	10
バッチサイズ	2

5.3 結果

データ拡張の有効性を調べるために，U-Net への入力画像を次のような 6 種類に分けた．

実験 1 前処理なしの画像
実験 2 頸部領域のクロップ画像
実験 3 首の角度を垂直に揃えたクロップ画像
実験 4 実験 2 について，平行移動によるデータ拡張
実験 5 実験 3 について，回転移動によるデータ拡張
　5-A) 首中心点中心に時計回り回転
　5-B) 首中心点中心に反時計回り回転
実験 6 実験 4 と実験 5-A，実験 5-B を組み合わせる

次に，6 つの実験の実験結果画像を示す．ここでは，緑色が TP，青色が FN，赤色が FP を表す．



図 7 原画像

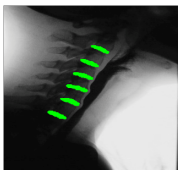


図 8 正解画像

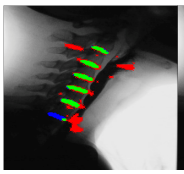


図 9 実験 1

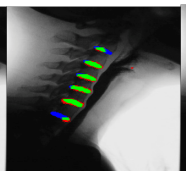


図 10 実験 2

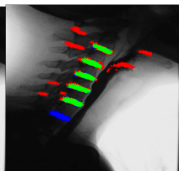


図 11 実験 3

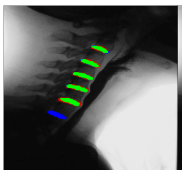


図 12 実験 4

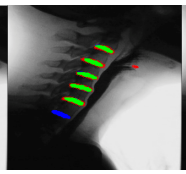


図 13 実験 5-A

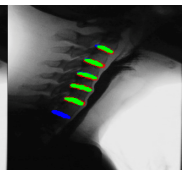


図 14 実験 5-B

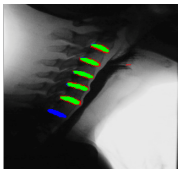


図 15 実験 6-A

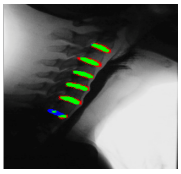


図 16 実験 6-B

6. 考察とまとめ

U-Net を用い，X 線透視画像からの頸椎椎間

板の抽出することができた．図 9 から図 11 と比べ，図 12 から図 16 の方が赤色の FN の領域が減っていることがわかる．このことから，データ拡張したことにより過剰検出が減ったと言える．また，平行移動と回転のみによるデータ拡張によって精度は向上したが，それらを組み合わせても精度は向上しなかった．特に抽出できなかった箇所として，肩部の隠されている部分があげられる．図 9 から図 16 の結果のどの場合においてもほとんど抽出されなかった．これは肩部にある椎間板は他に比べ，画素値が低いためであると考えられる．このため，濃淡値を調整したデータ拡張も精度向上に有効であると考えられる．今後の課題として，このような未検出を減らして更なる精度向上を目指す．

謝辞

研究遂行にあたり貴重なご協力を賜った京都大学大学院医学研究科人間健康科学系専攻リハビリテーション科学コース 松林潤氏，岡山大学整形外科 瀧川朋亨氏，神戸赤十字病院整形外科 戸田一潔氏，神戸赤十字病院整形外科 伊藤康夫氏に深謝いたします．

本研究の一部は JSPS 科研費 20J10293 の助成を受けたものです．

利益相反の有無

「なし」

文　献

［1］ 日本気管食道科学会「嚥下障害」 http://www.kishoku.gr.jp/public/disease04.html

［2］ 厚生労働省「平成 30 年（2018）人口動態統計月報年計（概数）の概況」 https://www.mhlw.go.jp/toukei/saikin/hw/jinkou/geppo/nengai18/dl/kekka30-190626.pdf

［3］ 平山三智子，西川光一「頚椎後方手術後に脳神経麻痺による嚥下障害を生じた 3 症例」日本臨床麻酔学会誌, Vol.32, No.7, pp.943-947, 2012

［4］ 林寿光ほか「頸椎手術後に悪化した嚥下障害例に対する手術治療経験」日本嚥下医学会，Vol.54，No.2，pp.152-156，2008

［5］ Fujinaka Ayano, Takizawa Hotaka, Mekata Kojiro & Kudo Hiro, “Automated Segmentation of Cervical Intervertebral Disks from Videofluorography Using a Convolutional Neural Network and its Performance Evaluation”, Journal of Signal Processing Systems/92(3)/pp.299-305, 2020-01

［6］ Olaf Ronneberger, Philipp Fischer, and Thomas Brox "UNet: Convolutional Networks for Biomedical Image Segmentation" Medical Image Computing and Computer Assisted Intervention 2015

［7］ Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio "Generative Adversarial Nets" Annual Conference on Neural Information Processing System, 2014

［8］ Jason Wang, Luis Perez "The Effectiveness of Data Augmentation in Image Classification using Deep Learning" arXivs-prints, Dec 2017

Preliminary study on extraction of cervical intervertebral disks from videofluorography using U-net

Erika GUNJI[*1], Ayano FUJINAKA[*1*2], Kojiro MEKATA[*3], Hotaka TAKIZAWA[*4], Hiroyuki KUDO[*4]

*1 Department of Computer Science, Graduate School of Systems and Information Engineering, University of Tsukuba

*2 Research Fellowship for Young Scientists, Japan Society for the Promotion of Science

*3 Kobe Red Cross Hospital

*4 Faculty of Engineering, Information and Systems, University of Tsukuba

Dysphagia may lead to malnutrition and aspiration, but its detailed mechanism remains unclear. To elucidate the mechanism, we aim to analyze the dynamics of cervical structures and investigate the relationship with dysphagia. In this report, cervical intervertebral discs are extracted from the frame images of videofluorography (VF). First, a neck region is also extracted based on morphology operation and RANSAC-based parabolic fitting. Then, U-Net, which was developed to segment medical images, is applied to a crop image. Generally, it is well known that deep learning generally produces better models with a larger number of training data. Therefore, affine-based data augmentation is applied to the images. The proposed method was applied to actual VF, and several experimental results are shown.

Key words: videofluorography, cervical intervertebral disks, U-Net, Data Augmentation, RANSAC

２種類の３次元畳み込みニューラルネットワークを用いた造影後 T1 強調画像における神経膠腫の全自動グレード分類パイプライン

山城　滉斗[*1]　寺本　篤司[*1]　齋藤　邦明[*1]　藤田　広志[*2]

要旨

神経膠腫の悪性度の評価に用いられるグレードは治療方針の決定などに大きく関わる．術前のグレードの分類では MRI が大きな役割を果たしており，その精度向上のために画像解析技術を用いた多くの研究が行われてきた．ただしその多くは手動で抽出した腫瘍領域に対しての２次元解析に留まっている．また，Deep learning では学習に多くのデータや高性能マシンなどを必要とし，医療への適用の障壁となっている．その解決のために NVIDIA 社は clara プロジェクトにて様々な学習済みの Deep learning モデルを提供している．そこで本研究では clara プロジェクトの３次元畳み込みニューラルネットワーク(3D CNN)を用いて造影後 T1 強調画像から脳腫瘍領域を抽出し，その領域に対して独自の 3D CNN を用いてグレード II・III とグレード IV を全自動的に分類することで，高精度なグレード分類パイプラインの作成を行った．この手法に対して検証を行ったところ総合正解率は 87.0%となり，高精度にグレード分類がなされた．

キーワード：脳腫瘍，CNN，MRI，グレード分類

1. はじめに

脳腫瘍の一種である神経膠腫では，悪性度の指標であるⅠ～Ⅳのグレードが診断や治療方針の決定に大きな影響を与える[1]．術前のグレード分類では MRI 検査が大きな役割を果たしているが，MRI 検査によるグレード分類には医師の知識や経験が重要であるため，煩雑で負担が大きい．また，術後に摘出した組織を用いる病理検査に比べて精度は低い．したがって，畳み込みニューラルネットワーク(CNN)に代表される Deep learning 技術を用いて手動で抽出した領域に対しての２次元での解析が行われてきた[2]．しかし，本来 MR 画像は３次元画像である上，脳腫瘍も多方向に進展する．また，手動での領域抽出では抽出した領域のばらつきがその後の分類処理の精度に影響を及ぼす．

ここで，セグメンテーション技術を用いて正確に腫瘍領域を抽出することができれば，腫瘍の解析の精度を高めることができる．しかし，３次元画像のセグメンテーションには大量のデータと高性能な GPU を搭載したコンピュータが必要になり，一般的な研究室レベルで処理を実行することが困難であった[3,4]．この課題を解決するため，NVIDIA 社は clara プロジェクトにて自社製の高性能マシンを用いて様々な学習済みモデルを作成し，公開している[5]．

そこで本研究では，clara プロジェクトのモデルと独自に作成したモデルの２種類の 3D CNN

*1　藤田医科大学大学院保健学研究科〔〒470-1192 愛知県豊明市沓掛町田楽ヶ窪 1-98〕
e-mail: 82020307@fujita-hu.ac.jp
*2 岐阜大学工学部〔〒501-1193 岐阜県岐阜市柳戸 1-1〕

を用いてグレード分類を全自動的に行う手法を開発し，その有用性の評価を行った．

2．方法

本手法では脳腫瘍領域抽出用 3D CNN を用いて造影後の T1 強調画像から腫瘍領域のみが抽出され，それをもとに腫瘍領域と周囲の脳実質を含む VOI が作成される．この VOI がグレード分類用 3D CNN へ入力されるとグレード II・III またはグレード IV の確率が出力され，確率が高い方のグレードと判定される．

2.1．使用画像

２種類の 3D CNN の学習や評価には脳腫瘍の領域抽出や生存予測の精度を競う大会である BraTS2018 で使用されたデータセットを使用する[6,7,8]．このデータセットには 243 症例の造影 T1 強調画像とそれに対応した腫瘍領域のラベル画像が含まれており，240x240x155 ボクセルに統一された上で骨や眼球などの脳実質以外の部分は除去されている．

2.1．脳腫瘍領域抽出

脳腫瘍領域抽出用 3D CNN には前述の clara プロジェクトの脳腫瘍セグメンテーションモデルを使用する[9]．このモデルは画像生成に用いる Variational autoencoder(VAE)を基本構造としており，入力された造影後の T1 強調画像から画像特徴を抽出するエンコーダ部と画像特徴から腫瘍領域のみの画像を出力するデコーダ部，そして学習時のネットワークの正規化に関わる VAE 部で構成されている．また，層数が増加しても誤差伝播を容易にするために Skip connection がエンコーダ部内とエンコーダ部-デコーダ部間に組み込まれている．

2.2．関心領域の切り出し

はじめに脳腫瘍領域抽出用 3D CNN からの出力をもとに腫瘍領域の外接矩形を設定し，その最大辺の長さを算出する．この長さの２倍の大きさを一辺とする立方体の関心領域（VOI）を作成する．これにより脳腫瘍領域とその周囲の脳実質が VOI に含まれた状態で抽出される．その後，サイズのばらつきをなくすために VOI は 64x64x64 ボクセルにリサイズされ，グレード分類に使用される．

2.3．グレード分類

グレード分類には独自に作成したグレード分類用 3D CNN を使用した．このモデルは ResNet を３次元に拡張したものであり，畳み込み層を 14 層設けており Skip connection により誤差伝播を容易にしている．入力された VOI から画像特徴が抽出され，その画像特徴は全結合層で統合されてグレード II・III またはグレード IV の確率が出力される．

3．検証

3.1．検証方法

本手法の有効性を評価するために未学習の 42 症例を用いて脳腫瘍領域抽出とグレード分類の精度を評価した．脳腫瘍領域の抽出精度の評価には２つの集合の一致度を定量的に表す Dice index を使用し，その算出には式(1)を用いた．

$$Dice\ index = \frac{|A \cap B| \times 2}{|A| + |B|} \quad (1)$$

3.2．結果および考察

脳腫瘍領域の抽出では全ての症例において腫瘍領域が抽出され，全症例の Dice index の平均値は 0.839 となった．高グレードの腫瘍の典型所見であるリング状造影効果がある腫瘍は高精度に抽出されたが，浮腫領域を含む腫瘍では抽出精度が低下する傾向にあった．

続けて，この結果をもとに作成した VOI に対してグレード分類を行ったところ，総合正解率は 87.0％となった．高グレードの腫瘍の典型所見であるリング状造影効果のある腫瘍はグレード IV と分類されやすい傾向にあったが，この所見が無いものもグレード IV と分類されていたことからリング状造影効果以外の特徴も分類に寄与していると考えられる．

4. 結論

本研究では２種類の 3D CNN を用いた造影後の T1 強調画像における神経膠腫のグレード分類手法を開発した．検証を行った結果，グレード分類が高精度に行われ本手法の有用性が確認された．

利益相反の有無

なし

文　献

［1］ Wu J, Qian Z, Tao, L, et al.: Resting state fMRI feature-based cerebral glioma grading by support vector machine. Int J CARS **10**: 1167–1174, 2015

［2］ Yang Y, Yan LF, Zhang X, et al.: Glioma Grading on Conventional MR Images: A Deep Learning Study With Transfer Learning. Frontiers in neuroscience vol 12: 804, 2018

［3］ Beluch W, Genewein T, Nürnberger A, et al.: The Power of Ensembles for Active Learning in Image Classification. CVPR 10.1109: 9368-9377, 2018

［4］ Nima T, Jae YS, Suryakanth RG, et al.: Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?. IEEE Transactions on Medical Imaging 35.5: 1299-1312, 2016

［5］ NVIDIA.: NVIDIA Clara. https://developer.nvidia.com/clara, (参照 2020-07-27)

［6］ Menze BH, Jakab A, Bauer S, et al.: The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS). IEEE Transactions on Medical Imaging 34.10: 1993-2024, 2015

［7］ Bakas S, Akbari H, Sotiras A, et al.: Advancing The Cancer Genome Atlas glioma MRI collections with expert segmentation labels and radiomic features. Sci Data 4: 170117, 2017

［8］ Bakas S, Reyes M, Jakab A, et al.: Identifying the Best Machine Learning Algorithms for Brain Tumor Segmentation, Progression Assessment, and Overall Survival Prediction in the BRATS Challenge. ArXiv: 1811.02629, 2018

［9］ Myronenko A, et al.: 3D MRI brain tumor segmentation using autoencoder regularization. ArXiv: 1810.11654, 2018

Fully automated grading pipeline of glioma for contrast enhanced MR images using two kind of 3D convolutional neural networks

Hiroto YAMASHIRO[*1], Atsushi TERAMOTO [*1], Kuniaki SAITO[*1], Hiroshi FUJITA[*2]

*1 Graduate School of Health Sciences, Fujita Health University

*2 Faculty of Engineering, Gifu University.

The grade used for evaluation of the malignant glioma is influenced to the decision of treatment policy. MRI plays a major role in preoperative grading, and many studies using image analysis techniques have been performed to improve its accuracy. However, most of them are performed using two-dimensional analysis and manually extracted tumor areas. Also, deep learning requires a lot of data and high-performance machines for training, which is a barrier to applying to medical. To solve this problem, NVIDIA provides various pre-trained Deep learning models in the clara project. Therefore, in this study, the brain tumor region is extracted from the post contrast T1-weighted image using the 3D convolutional neural network (3D CNN) of the clara project. In addition to that, using original 3D CNN designed by us, probability of grade II・III and grade IV are obtained from that region. In this way, grading pipeline with high accuracy is created. In order to evaluate the usefulness of our method, the pipeline was evaluated using 42 case. As a result, the overall correct answer rate was 87.0%, and the tumors were classified with high accuracy.

Key words: brain tumor, CNN, MRI, grading

Mask R-CNN とテクスチャ解析による小児肺炎の検出

魚住　春日[*1]　寺本　篤司[*1]　日木　あゆみ[*2]　本元　強[*2]

河野　達夫[*3]　齋藤 邦明[*1]　藤田　広志[*4]

要旨

小児に高い感染のリスクがある肺炎の検査には胸部 X 線画像が用いられるが，医師ごとの診断にばらつきがあるため，小児肺炎の自動検出を試みた．小児胸部 X 線画像は，成長やポジショニングによって状態が変化しやすいため，画像全体を直接解析することは困難である．そこで本手法では，解析領域である肺野領域の抽出と，肺野内のテクスチャ解析の 2 段階による処理によって肺炎像の検出を行った．

肺野の抽出には物体検出と領域抽出処理が包含された Mask R-CNN を用いた．その後，肺野を 32×32pixels の小区域に分割し，テクスチャ特徴量を抽出した．識別器にてテクスチャ特徴量をもとに肺炎と正常に分類し，実画像上に肺炎の確率分布をマッピングした．

NIH が公開する胸部 X 線画像データベース ChestX-ray8 の画像を用いて学習を行い，東京都立小児総合医療センターで撮影された画像にて評価を行った．その結果，正常症例，肺炎症例共に約 80%の正解率となった．

キーワード：Deep learning, テクスチャ解析, 肺炎, 小児, 胸部 X 線画像

1. はじめに

免疫力が低く，集団生活の中にいることが多い小児は，肺炎への感染や重症化のリスクが高く，WHO の調査によると 5 歳未満の小児の死因の 15%を肺炎が占めている[1]. 肺炎が疑われると胸部 X 線撮影が行われるが，小児はポジショニングが難しいうえに，成長に伴い肺野が大きく変化するため，図 1 に示す様な多様な画像となる．そのため医師は症状だけでなく，年齢など様々な条件を考慮したうえで診断を下しており，その負担は大きい[2]. 以上のことから小児胸部 X 線画像を用いた肺炎の診断支援が重要といえる．

肺炎検出に関する研究は古くから行われているが, 暉 和彦らによる手法[3]は, 間質性肺炎の好発部位である肋横隔付近にのみ，矩形の ROI を設置した限定的な処理であり，処理範囲が肺野のような非矩形領域に対して適した手法ではない．近年，深層学習による物体検出モデルを用いた手法も発表されている．M.K.Dharani らは Faster R-CNN を用いた手法[4]を，Shangjie Yao らは YOLOv3 を用いた手法[5]を発表している．深層学習を用いた手法はデータが大量にある場合に有効であり，データが少数である小児胸部 X 線画像に対する有効性は示されていない．

本手法では，胸部 X 線画像から，ポジショニングや成長などにより変動が生じる小児肺野を抽出し，抽出した肺野内を小区域ごとにテクスチャ解析を行うことで，小児肺炎の検出を行った．



(a)　(b)　(c)

図 1　成人と小児胸部 X 線画像の比較
(a)成人 (b)小児(1 歳) (c)小児(10 歳)

*1 藤田医科大学大学院保健学研究科
〔〒470-1192 愛知県豊明市沓掛町田楽ケ窪 1 番地 98〕
e-mail: teramoto@fujita-hu.ac.jp(寺本篤司)
*2 茨城県立こども病院 〔〒311-4145 茨城県水戸市双葉台 3 丁目 3−3 番地 1〕
*3 東京都立小児総合医療センター 〔〒183-8561 東京都府中市武蔵台 2 丁目 8−29〕
*4 岐阜大学 工学部 電気電子・情報工学科
〔〒501-1112 岐阜県岐阜市柳戸 1−1〕

2. 手法

手法の概要を図 2 に示す．本手法は，解析領域である肺野の抽出と肺野の内部を小区域ごとに分割し，テクスチャ特徴量を算出するという 2 段階の処理から成る．

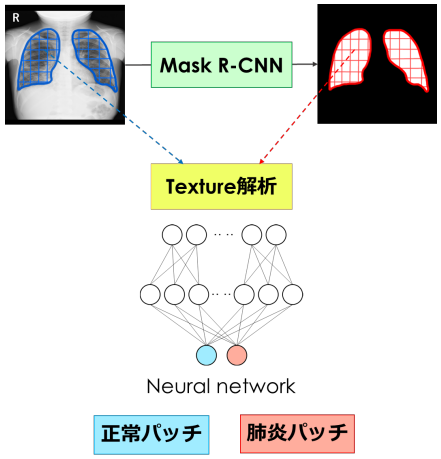


図 2 提案手法の概要

2-1. 肺野の抽出

小児胸部 X 線画像は，前述した様にポジショニングや成長による影響から，成人に比べて多様な像となる．確実に肺野を抽出するため，Deep learning 手法のひとつであり，検出した物体に対して外接矩形を設定する Object detection と，画素単位でクラス分類し，領域抽出を行う Semantic segmentation を内包している Mask R-CNN[6]を用いて，小児胸部 X 線画像における肺野の自動抽出を試みた[7]．本手法で用いた Mask R-CNN では，Microsoft COCO(MS COCO)[8]により学習された重みを初期重みとし，Object detection を担う Head 部の学習を行った後，独自に作成した学習データを用いてネットワーク全体をファインチューニングした．Mask R-CNN に入力する画像サイズは 512x512 とし、Head 部の学習係数は 0.001，エポック数は 50[epochs]，ファインチューニングの際の学習係数は 0.0001，エポック数は 150[epochs]とした．また，Head 部，ファインチューニング共に，バッチサイズは 1，最適化手法は SGD とした．学習された Mask R-CNN に未学習の小児胸部 X 線画像を入力することで，肺野を示す肺野マスクが得られる．肺野マスクの辺縁を補正するため，円形フィルタを用いたモフォロジー処理を行った．

2-2. テクスチャ特徴量の算出

本手法では，基礎統計量，GLCM，GLRLM，GLSZM，GLDM に基づく 88 種類のテクスチャ特徴量を算出した．

肺野の内部には肋骨や肺血管など炎症の他に様々な陰影が含まれる．そこで，肺野を 32×32[pixels]の小区域に分割して，テクスチャ特徴量を算出した．肺野の辺縁付近の非矩形部分は，肺野マスクの形状に沿ってテクスチャ特徴量を算出した．なお，テクスチャ特徴量の算出には Python のパッケージである Pyradiomics を用いた．

2-3. 分類

算出したテクスチャ特徴量を Neural network を用いて肺炎と正常の 2 つに分類した．学習係数は 0.001，エポック数は 100[epochs]，バッチサイズは 128 とした．また，損失関数に二値交差エントロピーを用い，最適化手法は Adam をとした．

3. 使用画像

3-1. Mask R-CNN の学習

Mask R-CNN の学習には NIH が公開している ChestX-ray8 データベース[8]を使用した．このデータベースから肺野を明確に定義することができる小児の胸部 X 線画像を 200 枚選択し，画像枚数の不足を補うため成人の胸部 X 線画像を 800 枚追加した．その後，抽出対象である肺野領域のみを一定のラベル値で塗りつぶし

たラベル画像を小児専門の放射線科医の指摘に基づき作成し，原画像とのペアを学習データとした．

3-2．テクスチャ特徴量の算出

ChestX-ray8 を基に作成され，RSNA2018 で行われた肺炎検出コンペティションで使用されたデータ[9]から，肺炎と診断され，チューブなどの医療装具が混入していない 418 枚の胸部 X 線画像と，正常と診断された 300 枚の胸部 X 線画像を選択した．選択した画像から算出したテクスチャ特徴量を Neural network の学習データとした．

3-3．テストデータ

Mask R-CNN による肺野抽出とテクスチャ解析による肺炎像の検出には，東京都立小児総合医療センターで撮影された正常症例 52 枚と肺炎症例 81 枚の小児胸部 X 線画像を用いた．

4．検証

4-1．検証方法

本手法の有効性を評価するために，肺野抽出精度と肺炎パッチの分類精度を評価した．

肺野抽出精度の評価には，正解画像との一致率を表す Jaccard index と Dice index を用い，混同行列と ROC 曲線を用いて，各パッチの分類精度を評価した．

4-2．結果及び考察

4-2-1．肺野抽出精度

テストに用いた小児胸部 X 線画像を Mask R-CNN に入力した結果を図 4 に示す．正解画像との一致率は Jaccard index = 0.92, Dice index = 0.96 となり，学習に用いたデータベースと異なるデータベースの画像を用いた場合でも高い抽出精度が確認された．

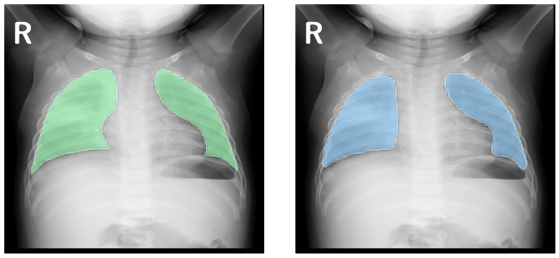


図 4　未学習小児胸部 X 線画像における肺野抽出結果（左：正解画像，右：検出結果）

4-2-2．肺炎パッチ分類精度

4-2-2-1．正常症例に対する結果

肺野内から小区域ごとのテクスチャ特徴量を算出し，Neural network によって肺炎と分類されたものが赤く示されるように可視化した．図 5 に検出結果を示す．ここでは正常症例に対する結果であるため，検出された箇所は偽陽性となる．正しく正常と分類できたパッチは 2806 個，誤って肺炎と分類したパッチは 473 個となり，86%の正解率が得られたが，多くの偽陽性が生じ，特に(b)の様な年齢が低い幼児において偽陽性が多く生じた．幼児の場合，低管電圧・ノングリッドで撮影するため，散乱線によってランダムノイズの多い 画像となることが原因であると考えられる．

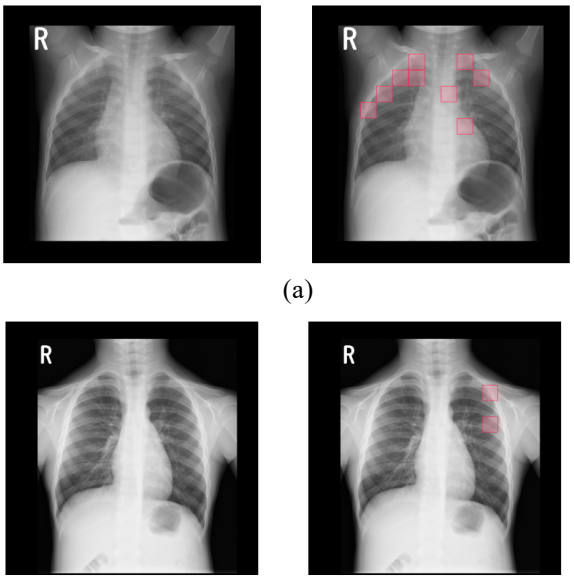


図 5　正常症例における検出結果
（左：入力画像，右：検出結果）

4-2-2-2．肺炎症例に対する結果

図 6 に検出結果，図 7 に ROC 曲線，表 4 に混同行列を示す．正解率は 77%となったが，肺炎症例においても多くの偽陽性・偽陰性が見られた．偽陽性は肺尖などの比較的画素値が高い箇所において多く確認された．

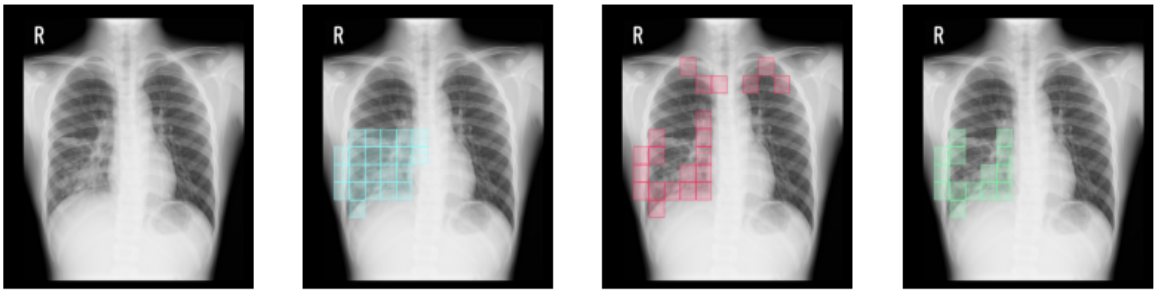


図 6　肺炎症例における検出結果（左から順に入力画像，正解画像，検出結果，真陽性を表す.）

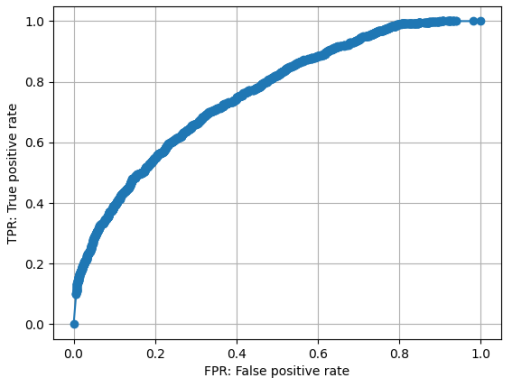


図 7　肺炎症例における ROC 曲線

表 4　肺炎症例における混同行列（単位：パッチ）

予想／正解	正解が肺炎	正解が正常
肺炎と予想	526	525
正常と予想	807	3915

5. まとめ

肺炎に対して高い感染と死亡のリスクがある小児の胸部 X 線画像を用いて，肺炎像の自動検出を行った．本手法は解析領域である肺野の抽出と肺野内のテクスチャ解析の 2 段階の処理に分かれ，肺野抽出には Mask R-CNN を用いた．正解画像と比較して 90%を超える高い抽出率が得られた　．肺野内のテクスチャ解析では，多くの肺炎症例で正しく検出されたが，正常症例，肺炎症例ともに偽陽性・偽陰性が多く生じた．テクスチャ特徴量が輝度に影響を受けやすく，肺尖などの辺縁は高輝度であるため偽陽性が生じたと考えられる．また，学習に用いたデータベースには，成人の画像が多く含まれ，テストに用いた小児の胸部 X 線画像とは肺野内の状態が大きく異なることも偽陽性・偽陰性の要因であると考えられる．

利益相反の有無

なし

文　献

［1］ World Health Organization: WHO: Pneumonia. 2019 https://www.who.int/news-room/fact-sheets/detail/pneumonia(accessed on23 July 2020)

［2］ 相田典子：小児科領域における医療技術を考える．映像情報メディカル 47/13 2015 年 12 月号: 1138-1142, 2015

［3］ 暉和彦，島田哲雄，児玉直樹，田中啓之，et al. :胸部単純 X 線写真による間質性肺炎自動診断支援システムの基礎研究. 長岡技術科学大学研究報告 **25**: 99-103, 2003

［4］ M.K.Dharani, S.Santhoshkumar, T.Balaji, B.Dhayaneshwar, et al.: Detection of Pneumonia from X-Ray Imaging using Faster RCNN. International Journal of Control and Automation**13**: 227-234, 2020

［5］ Yao. S, Chen. Y, Tian. X, Jiang. R, et al.: An Improved Algorithm for Detecting Pneumonia Based on YOLOv3. Applied Sciences**10**: 2020

［6］ Kaiming He, Georgia Gkioxari, Piotr

Dollar, et al.: Mask R-CNN. Proceedings of the IEEE international conference on computer vision: 2961-2969, 2017

［7］ 魚住春日, 松原尚輝, 寺本篤司, 日木あゆみ, et al.: Mask R-CNN を用いた小児胸部 X 線画像における肺領域の自動抽出. Medical Imaging Technology**38**, 126-131, 2020

［8］ Tsung-Yi Lin, Michael Maire, Serge Belongie, et al.: Microsoft coco: Common objects in context. European conference on computer vision**8693**: 740-755, 2014

［9］ Xiaosong Wang, Yifan Peng, Le Lu, et al.: Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. Proceedings of the IEEE conference on computer vision and pattern recognition: 2097-2106, 2017

［10］ Kaggle. Available online: https://www.kaggle.com/c/rsna-pneumonia-detection-challenge/data (accessed on23 July 2020)

Detection of Pneumonia Using Mask R-CNN and Texture Analysis

Haruka UOZUMI[*1], Atsushi TERAMOTO[*1], Ayumi NIKI[*2], Tsuyoshi HONMOTO[*2], Tatsuo KONO[*3], Kuniaki SAITO[*1], Hiroshi FUJITA[*4]

*1 Graduate School of Health Sciences, Fujita Health University

*2 Ibaraki Prefectural Children's Hospital

*3 Department of Radiology, Tokyo Metropolitan Child Medical Center

*4 Department of Electrical, Electronic & Computer Engineering, Gifu University

Children have a high risk of infection of pneumonia, and Chest X-ray images are used in pneumonia diagnosis. The diagnosis of pneumonia varies among physicians. We developed an automated detection scheme of pneumonia in pediatric chest X-ray. This method has two steps that is extraction of lung regions to obtain the analysis regions and detection pneumonia with texture analysis.

Mask R-CNN including object detection and semantic segmentation was used to extract the lung regions. The lung regions were divided into image patches of 32 × 32 pixels and texture features were extracted by subdivision. Texture features were classified them into pneumonia and normal using discriminator and the location of the pneumonia was mapped on chest X-ray images.

As for the training of network, Chest X-ray images of ChestX-ray8 database were used and the images acquired at Tokyo Metropolitan Children's Medical Center were used for testing. As a result of evaluation, the overall rate reached 80% for both normal and pneumonia cases.

Key words: Deep learning, Texture analysis, Pneumonia, Pediatrics, Chest X-ray images

V-Net を用いた Ai-CT に対する肝臓領域抽出手法の開発と肝臓の質量推定への応用

櫛部　光哉[*1]　平野　靖[*1]　木戸　尚治[*2]　木下　一之[*3]

稲井　邦博[*3]　法木　左近[*4]

要旨

死亡時画像診断(Autopsy imaging：Ai)とは，死後 CT 像を読影することによって，死因究明や死後経過時間の推定を行う手法である．臓器の質量は疾患や病態を反映するため，死因究明を行う上で有用な情報となる．一般的に，臓器の質量を測定するためには解剖を必要とする．しかし，日本では剖検率が極めて低いため正確な質量を得ることができない場合が多い．

本研究では，死後 CT 像における肝臓領域内の CT 値を用いて，解剖を行わずに肝臓の質量を推定する手法を提案する．CT 値から肝臓の質量を推定するためには，肝臓領域を抽出する必要がある．しかし，CT 像から肝臓領域を手作業で抽出するためには多大な労力を要する上に，得られた肝臓領域の抽出精度は安定しない．そこで本研究では，まず V-Net をベースとした手法を用いて肝臓領域の抽出し，次に抽出された肝臓領域内の CT 値を基に質量の推定を行った．

1 はじめに

死亡時画像診断(Autopsy imaging：Ai)とは，CT(Computed Tomography)装置などを用いて遺体を撮影し，死因究明や死後経過時間推定を行う検査手法であり，解剖を行わなくても遺体の内部情報を得られることから，近年実施される機会が多い．臓器の質量は疾患や病態を反映する[1]ため，死因究明において重要な情報を与える．一方で，臓器質量は解剖によってのみ得ることができる．しかし，日本の剖検率は総死亡者数の 4%と先進国の中でも極めて低い[2]．この原因として解剖医の不足と遺族の解剖に対する抵抗感が挙げられる．そこで本研究では物体の CT 値と密度の関係から質量を求める手法[3]を用いた，死後 CT 像中の肝臓領域の質量推定手法を提案する．

*1 山口大学大学院創成科学研究科電気電子情報系専攻

〔〒755-8611 宇部市常盤台 2-16-1〕

e-mail: a062vgu@yamaguchi-u.ac.jp

*2 大阪大学大学院医学系研究科

*3 福井大学医学部

*4 福井県立大学看護福祉学部

ここで，肝臓は臓器の中でも質量が極めて狭い範囲で維持されている[4]ため，質量が大きく変化している場合は疾患や病態を反映している可能性が高い．

肝臓の質量推定を行う際，肝臓領域を死後 CT 像から抽出する必要がある．そこで，本研究では V-Net[5]をベースにした手法を用いて肝臓領域を抽出した．

2　手法

2.1　手法の概要

本研究における手法のフローチャートを図 1 に示す．まず，前処理として実験に使用する CT 像と正解肝臓領域に対して肝臓の領域を含むように切り出しを行い，画像のスライス枚数の統一を行った．次に，死後 CT 像と正解肝臓領域のサイズを縮小した．それらの画像を用いて V-Net をベースとしたネットワークの学習を行い，肝臓領域の抽出を行った．さらに，抽出した肝臓領域を用いて質量推定を行った．



図 1 手法の流れ

2.2　前処理

計算コストの削減を目的に，CT 像に対して，肝臓領域が収まるように肝臓の中心スライスから上下 128 スライスを切り出し，画像サイズを$512 \times 512 \times 256$[voxel]に統一した．さらに$128 \times 128 \times 64$[voxel]となるように 1/64 に縮小した．

2.3　モデルの学習

本研究では V-Net をベースとした手法を用いて肝臓領域抽出を行った．図 2 に使用したネットワークモデルの概要を示す．本実験のパラメータ更新アルゴリズムには，Adaptive moment estimation(Adam[6])を用い，学習係数を1.0×10^{-5}，バッチサイズを 2 とした．

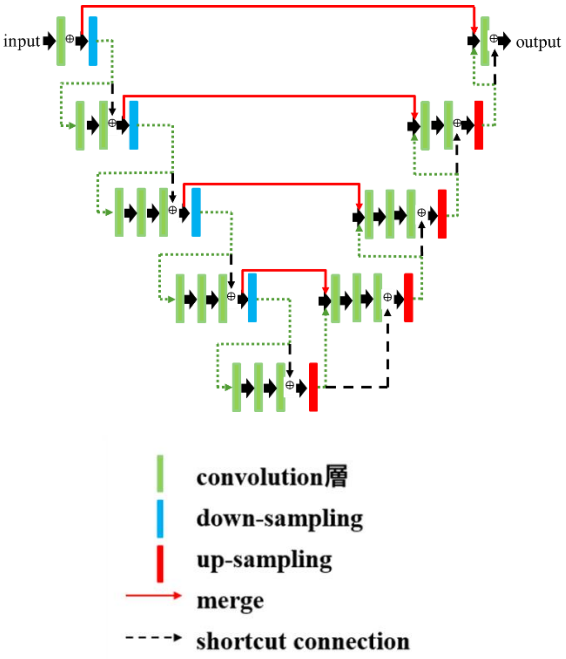


図 2　本研究で使用したネットワークモデルの概要

本研究で使用したネットワークモデルの各層におけるフィルタサイズ，ストライド，出力マップサイズおよび適用した活性化関数を表 1 に示す．なお，表中の X は出力される画像の横方向の画素数，Y は縦方向の画素数，Z はスライス枚数，および K はチャンネル数を表す．

表 1 本研究で使用したネットワークモデルの各層の構成

層種・名称	フィルタサイズ	ストライド	出力マップサイズ (X×Y×Z×K)	活性化関数
input			128 × 128 × 64 × 1	
Convolution1_1	5 × 5 × 5	1	128 × 128 × 64 × 16	PReLU
down-sampling1	2 × 2 × 2	2	64 × 64 × 32 × 32	PReLU
Convolution2_1	5 × 5 × 5	1	64 × 64 × 32 × 32	PReLU
Convolution2_2	5 × 5 × 5	1	64 × 64 × 32 × 32	PReLU
down-sampling2	2 × 2 × 2	2	32 × 32 × 32 × 64	PReLU
Convolution3_1	5 × 5 × 5	1	32 × 32 × 16 × 64	PReLU
Convolution3_2	5 × 5 × 5	1	32 × 32 × 16 × 64	PReLU
Convolution3_3	5 × 5 × 5	1	32 × 32 × 16 × 64	PReLU
down-sampling3	2 × 2 × 2	2	16 × 16 × 8 × 128	PReLU
Convolution4_1	5 × 5 × 5	1	16 × 16 × 8 × 128	PReLU
Convolution4_2	5 × 5 × 5	1	16 × 16 × 8 × 128	PReLU
Convolution4_3	5 × 5 × 5	1	16 × 16 × 8 × 128	PReLU
down-sampling4	2 × 2 × 2	2	8 × 8 × 4 × 256	PReLU
Convolution5_1	5 × 5 × 5	1	8 × 8 × 4 × 256	PReLU
Convolution5_2	5 × 5 × 5	1	8 × 8 × 4 × 256	PReLU
Convolution5_3	5 × 5 × 5	1	8 × 8 × 4 × 256	PReLU
up-sampling1	2 × 2 × 2	2	16 × 16 × 8 × 128	PReLU
Convolution6_1	5 × 5 × 5	1	16 × 16 × 8 × 128	PReLU
Convolution6_2	5 × 5 × 5	1	16 × 16 × 8 × 128	PReLU
Convolution6_3	5 × 5 × 5	1	16 × 16 × 8 × 128	PReLU
up-sampling2	2 × 2 × 2	2	32 × 32 × 16 × 64	PReLU
Convolution7_1	5 × 5 × 5	1	32 × 32 × 16 × 64	PReLU
Convolution7_2	5 × 5 × 5	1	32 × 32 × 16 × 64	PReLU
Convolution7_3	5 × 5 × 5	1	32 × 32 × 16 × 64	PReLU
up-sampling3	2 × 2 × 2	2	64 × 64 × 32 × 32	PReLU
Convolution8_1	5 × 5 × 5	1	64 × 64 × 32 × 32	PReLU
Convolution8_2	5 × 5 × 5	1	64 × 64 × 32 × 32	PReLU
up-sampling4	2 × 2 × 2	2	128 × 128 × 64 × 16	PReLU
Convolution9_1	5 × 5 × 5	1	128 × 128 × 64 × 16	PReLU
output			128 × 128 × 64 × 1	sigmoid

2.4 質量推定

提案手法で得られた肝臓領域と，肝臓領域内の平均 CT 値から，以下の推定式[3]を用いて質量を推定した．

$$\text{体積}[cm^3] = \text{領域内の voxel 数} \times 1\text{voxel 当たりの体積}[cm^3] \quad (1)$$

3 実験

3.1 実験データ

本研究では，福井大学医学部附属病院で撮影された死後 CT 像 24 症例を用いた．このうち，15 症例は肝臓の質量が既知であり，残りの 9 症例は未知である．また，医用画像工学の専門家の指導の下，手作業で作成した正解肝臓領域を抽出した．使用した CT 像の詳細を以下に示す．

・スライスサイズ　: 512×512[pixel]
・スライス数　: 180~571 枚
・スライス厚　:1.0~2.5[mm]
・スライス内の画素サイズ:0.626~0.841[mm]

3.2 肝臓領域抽出のためのデータセットの作成

24 症例の CT 像と正解肝臓領域を学習データとテストデータに分けた．質量が既知である 15 症例を 3 分割し，3 分割交差検証を行った．質量が未知である 9 症例は学習データに含めた．

3.3 評価方法

本研究では，領域抽出精度の評価方法として Dice 係数を用いた．Dice 係数は，出力肝臓領域を A とし, 正解肝臓領域を B とする場合, 式(2)で求められる．Dice 係数は 0.0 から 1.0 の範囲の値を取る．値が 1.0 に近づくほど A と B の領域は一致しているとされている．

$$D(A,\ B) = \frac{2 \times |A \cap B|}{|A| + |B|} \quad (2)$$

A：出力肝臓領域

B：正解肝臓領域

$|X|$：領域 X の画素数

質量推定の評価方法としては, 実際の質量と推定した質量の相関係数を用いた．

3.4 肝臓領域の抽出結果

提案手法を用いて CT 像中の肝臓領域を抽出した結果を示す．本研究では，学習を 500epoch まで行った結果，平均 Dice 係数は 0.842±0.069となった．

3.5 質量推定の結果

実際の質量と出力肝臓領域を用いて推定した質量の分布を図 3 に示す．得られた結果から実際の質量と推定した質量の相関係数を求めたところ 0.774 となった．また、絶対誤差と標準偏差も求めたところ 617.8±231.8[g]となった．なお，正解肝臓領域を用いた推定も行った．実際の質量と推定した質量の分布を図 4 に示す．得られた結果から実際の質量と推定した質量の相関係数を求めたところ 0.942 となった．また，平均絶対誤差と標準偏差は 782.5±194.5[g]となった．

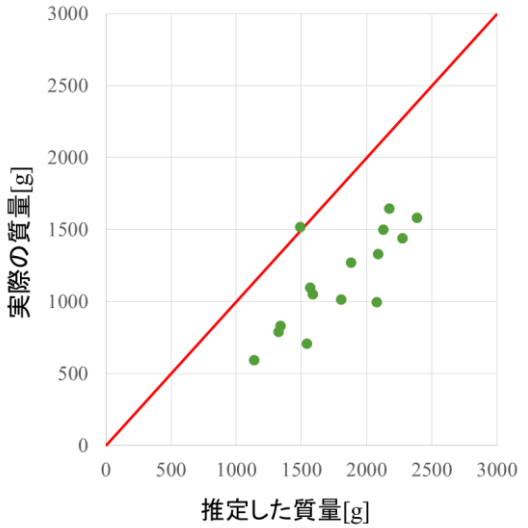


図 3　実際の質量と出力肝臓領域を使用して推定した質量の散布図

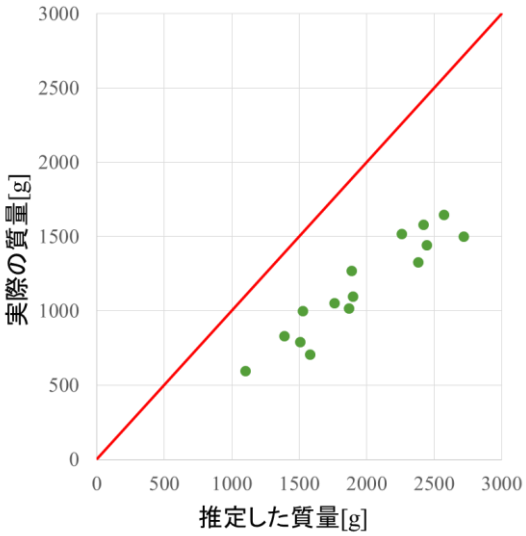


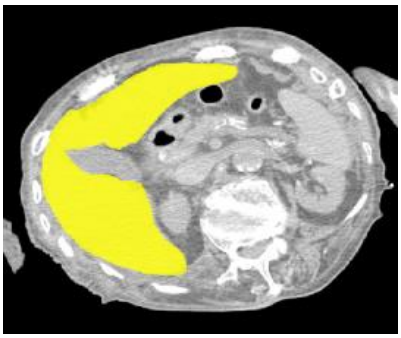
図 4　実際の質量と正解肝臓領域を使用して推定した質量の散布図

4　考察

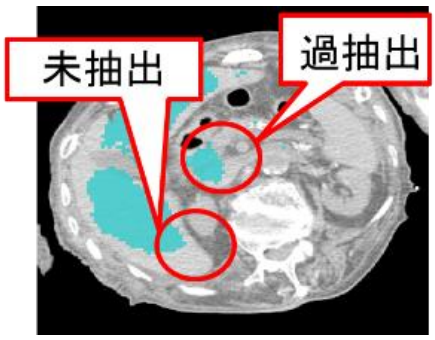
4.1　肝臓領域抽出結果に対する考察

提案手法を用いて領域抽出を行った結果，Dice 係数が最も低かった症例を図 5 に示す．この症例では，肝臓領域の一部が未抽出であったことや他の臓器を過抽出していることが確認できる．まず，未抽出が発生した理由として，肝臓はその大きさや形状が症例によって異なるため，適切に肝臓領域抽出を行うための学習に必要なデータのバリエーションが不足していたことが考えられる．また，図 5 に示した症例では，撮影時に人体が傾いていたことも影響したと考える．そこで，回転処理を用いた学習データの水増しや学習に用いる症例数の増やすことが対策として考えられる．また，過抽出になった理由として，肝臓と他の組織の CT 値が類似していることが考えられる．そこで，前処理としてガンマ変換などを行い，他の組織との CT 値のコントラストを強調することで過抽出が改善すると考えられる．

また，ネットワークモデルの改良を行うことも，過抽出や未抽出の問題への対策であると考える．モデルの具体的な改良策として，ハイパーパラメータ(ネットワークの構造を決定するパラメータ)自動最適化ツールの Optuna[7]を用いて，層数やチャネル数などを最適化することが挙げられる．



(a)　正解肝臓領域



(b) 抽出結果(Dice 係数：0688)

図 5 過抽出・未抽出の例

4.2 出力肝臓領域を使った質量推定の結果に対する考察

出力肝臓領域を使用した質量推定では正解肝臓領域を使用した場合よりも大きな誤差が生じていることが確認できる．原因として，前処理で行った CT 像に対する縮小処理による画素値の変化が関係していることが考えられる．そのため，縮小を行わずに領域抽出を行う手法を開発する必要もあると考える．具体的には，CT 像を 2 次元のスライスでごとに入力し， 2 次元の畳み込みを行い，各スライスに対する領域マップを作成した後に，それらを 3 次元画像に再構成することで，同時に GPU 内に保存しておくべき画像のデータ量や最適化すべきパラメータの数を削減することで GPU 内のメモリの使用量を減らす手法などが挙げられる．

4.3 正解肝臓領域を使った質量推定の結果に対する考察

図 4 から正解肝臓領域を使用した場合でも，症例ごとの誤差にばらつきが生じていることが確認できる．原因として，アーチファクトによる CT 値の変化が質量推定に用いる計算式(式(1))に影響を及ぼしたと考えられる．解決策として，アーチファクトと思われる CT 値に重みをかけることにより誤差を小さくすることが考えられる．

5 まとめ

本研究では，V-Net をベースとした手法を用いて死後 CT 像から肝臓領域抽出を行い，抽出した肝臓領域から求めた肝臓領域の体積と肝臓領域内の CT 値の平均を利用し，質量推定を行った．その結果，肝臓領域抽出では Dice 係数 0.842±0.069，肝臓領域抽出の出力結果を使用した質量推定では，実際の臓器の体積と推定された体積の相関係数として 0.774 が得られた．また，質量推定のリミテーションを確認するために行った正解肝臓領域を使用した質量推定の結果では，相関係数が 0.942 であった．ここで，領域抽出の精度が不十分だった原因として，学習データの不足などがあげられる．そのため，今後の課題として症例数を増やすことによって学習データのバリエーションを増やすことやデータの水増しをすることで抽出精度の向上が期待される．また，本研究では比較的高い相関関係は得られたが，どの症例も誤差が非常に大きいため，より精度の高い質量推定手法の開発も行っていく必要がある。

謝辞

本研究は日本学術振興会科学研究費補助金新学術領域研究(26108009)および基盤研究(B)(20H03908)の支援を受けた．

御助言及び御協力を頂きました本研究室の大学院生，学部生の皆様に心よりお礼申し上げます．

利益相反の有無

なし.

文　献

[1] 臓器重量について, 日本医事新報社, https://www.jmedj.co.jp/journal/paper/detail.php?id=4718 (2020/1/20)

[2] 藤宮龍也, “世界の検死制度と医療関連死”, 国立大学附属病院病理部長会議, 2007

[3] S.Nicola, F.Simone, et al.,“CT-based weight assessment of lung lobes: comparison with ex vivo measurements.”, Diagnostic and Interventional Radiology 19.5, pp.355.-359, 2013

[4] 肝重量の恒常性維持機構, 肝細胞研究会, http://hepato.umin.jp/kouryu/kouryu20.html (2020/1/20)

[5] M.Fausto, N.Nassir, et al.,“V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation”, 2016 Fourth International Conference on 3D Vision (3DV).IEEE, pp.565-571, 2016

[6] D. Kingma, J. Ba,“Adam: A Method for Stochastic Optimization”, International Conference on Learning Representations, 2014

[7] Optimize Your Optimization, Optuna, https://optuna.org/ (2020/1/20)

心臓超音波動画像における畳み込みニューラルネットワーク用いた心筋梗塞の自動分類手法の開発

村木　亮介[*1]　木村　竜誠[*1]　寺本　篤司[*1]　山崎　歌音[*1]

杉本　恵子[*1]　杉本　邦彦[*2]　山田　晶[*3]　渡邉　英一[*3]

齋藤　邦明[*1]　藤田　広志[*4]

要旨

現在、急性心筋梗塞は生活習慣病が原因となり、突然死や慢性的な心不全を引き起こす可能性があり、早期発見が重要な疾患である。急性心筋梗塞の検査法として超音波検査が挙げられる。しかし、所見である異常壁運動は範囲や重症度により判別が困難であり、致死的な合併症につながるため高い正確性が必要とされる。そこで本研究では、心臓超音波動画像における畳み込みニューラルネットワーク（CNN）と従来の機械学習法を利用した急性心筋梗塞の自動分類手法の開発を目的とした。短軸像、左室長軸像の 1 心拍分にあたる動画像を入力し、CNN により特徴量を抽出した。特徴量より Support vector machine、Artificial neural network、AdaBoost、Random forest、Naïve Bayes の機械学習アルゴリズムを使用し、正常症例と心筋梗塞症例の分類を行った。結果は、1 心拍分の動画像を使用したときの精度は左室長軸像で 81.8%、短軸像で 82.8%となった。以上の結果より、本手法は心筋梗塞の自動検出に有効であることが示唆された。

キーワード：ディープラーニング、心臓超音波動画像、心筋梗塞、機械学習

1. はじめに

国内の疾患別死亡率において，心疾患は悪性新生物に次いで 2 位とされ，突然死や心不全をはじめとする生死に直結した疾患が多く存在する[1]．特に虚血性心疾患である急性心筋梗塞は心臓突然死の原因の中で最も多く，慢性的な心不全や致死的な不整脈を引き起こす可能性を持つ．これらは患者の予後に大きく影響することから，心筋梗塞の早期診断は非常に重要である．

心筋梗塞診断時に利用される画像検査として、超音波検査、核医学検査、CT、MRI 等があるが、特に超音波検査は検査プロトコルの早期に利用される。心臓超音波検査の特徴として、心臓の形態や壁運動をリアルタイムに評価することが可能な点や、心筋梗塞の所見である異常壁運動の観察を行える点、同時に合併症の有無や機能評価を行える点が挙げられる。しかし、所見である異常壁運動は範囲や重症度により評価が困難な点や手技者の技量や熟練度によって認識精度に差が生じるといった問題点も挙げられる。上記の問題点より、心臓超音波動画像における壁運動の定量的な評価が求めら

*1 藤田医科大学大学院保健学研究科〔〒470-1192 豊明市沓掛町田楽ヶ窪 1 番地 98〕e-mail: teramoto@fujita-hu.ac.jp
*2 藤田医科大学病院
*3 藤田医科大学循環器内科
*4 岐阜大学工学部

れる。

現在、人工知能技術を利用した医用画像の定量的な評価やコンピュータ支援診断技術（CAD）等の研究が盛んに行われている．特に人工知能技術の一つであるディープラーニングを心臓超音波動画像へ応用し、断層像のセグメンテーション[2]や、左室駆出率などの心機能を予測・推定する研究[3]は多く存在する。しかし，ディープラーニングにより心臓超音波動画像を解析し，急性心筋梗塞の検出や分類を行った研究はない.

よって本研究は，人工知能技術の一つである畳み込みニューラルネットワーク（CNN）[4]を用いた心臓超音波動画像における急性心筋梗塞の自動検出手法の開発を目的とし，初期的な検討として機械学習法を用いた自動分類手法の開発を行った.

2. 手法

図 1 に示すような短軸像(乳頭筋レベル)、左室長軸像の 2 種類の断層像を入力し、機械学習法用いて正常像と急性心筋梗塞症例の分類を行った。超音波動画像から 1 心拍分にあたる動画像を入力し、CNN により約 12 万個の特徴量を抽出した。その後主成分分析を行い、特徴量を次元圧縮した。圧縮した特徴量に対して機械学習法である、Support vector machine、Artificial neural network、AdaBoost、Random forest、Naïve Bayes を使用し、正常像と急性心筋梗塞発症例の分類を行った。

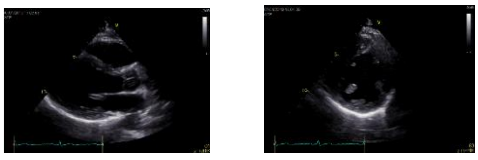


図1:超音波動画像
(左:左室長軸像 右:短軸像乳頭筋レベル)

3. 結果および考察

44 症例の 1 心拍分の動画像に対して分類を行い、leave one case out 法で総合分類精度を算出した。1 心拍分の動画像を使用したときの正解率は Naive Bayes を使用した場合に左室長軸像で 81.8%、短軸像で 82.1%となった。また、拡張末期、収縮中期、収縮末期にあたる 3 フレームで同様に特徴量を抽出し、総合分類精度を算出した。一方で、3 フレーム使用時の精度は、左室長軸像で 70.8%、短軸像で 72.9%となった。以上の結果より、1 心拍分全体の動画情報が心筋梗塞の分類に有用であることが示唆された。

利益相反の有無

なし.

文　献

［1］ 平成 30 年（2018）人口動態統計月報年計（概数）の状況：https://www.mhlw.go.jp/toukei/saikin/hw/jinkou/geppo/nengai18/dl/gaikyou30.pdf（2020/07/16）

［2］ Gao X, Li W, Loomes M, Lianyi W. A fused deep learning architecture for viewpoint classification of echocardiography. Inf Fusion. Volume36, pp103–113, July, 2017.

［3］ J. F. Silva, J. M. Silva, A. Guerra, et al.: "Ejection fraction classification in transthoracic echocardiography using a deep learning approach", Proc. IEEE 31st Int. Symp. Comput .-Based Med. Syst., pp. 123-128, 2018.

［4］ K. Simonyan, A. Zisserman, ”Very Deep Convolutional Networks for Large-Scale Image Recognition” arXiv:1409.1556 , 2014.

Development of Automated Classification Scheme for Myocardial Infarction Using Convolutional Neural Network in Echocardiography

Ryosuke MURAKI[*1], Ryusei KIMURA[*1], Atsushi TERAMOTO[*1]
Kanon YAMAZAKI[*1], Keiko SUGIMOTO[*1], Kunihiko SUGIMOTO[*2]
Akira YAMADA[*3], Eiichi WATANABE[*3], Kuniaki SAITO[*1]
Hiroshi FUJITA[*4]

[1)] Graduate School of Health Sciences, Fujita Health University
[2)] Fujita Health University Hospital
[3)] School of Medicine, Fujita Health University
[4)] Faculty of Engineering, Gifu University

The acute myocardial infarction is caused by lifestyle-related diseases and may cause sudden death or chronic heart failure, and early detection is essential. An ultrasound is mentioned as an examination of acute myocardial infarction. However, the abnormal wall motion which is one of major finding of myocardial infarction is difficult to distinguish depending on the range and the severity. Missing of this finding leads to a fatal complication, so high accuracy is required. Therefore, in this study, we aimed to develop an automated classification scheme of acute myocardial infarction using convolutional neural network (CNN) and conventional machine learning method in echocardiography. Ultrasound images of short-axis view and a long-axis view of the left ventricle for one cardiac cycle were given to the CNN and the features were extracted. And then, healthy cases and myocardial infarction cases were classified using conventional machine learning algorithms such as Support vector machine, Artificial neural network, AdaBoost, Random forest, Naïve Bayes. As a result, the accuracy when using one cardiac cycle images reached 81.8% in the left ventricular long-axis view and 82.8% in the short-axis view. These results indicate that the proposed scheme may be useful to detect the myocardial infarction in echocardiography.

Key words: Deep learning, echocardiography, myocardial infarction, machine learning

Template for IJCARS Special Issue of JAMIT2019 & 2020

Automated Segmentation of Body Cavity Based on U-Net Architectures in Torso CT Images

Taira MAEDA[*1], Naoki KAMIYA[*1], Xiangrong ZHOU[*2], Hiroki KATO[*3], Takeshi HARA[*2], Toshiharu MIYOSHI[*4], Masayuki MATSUO[*3], Hiroshi Fujita[*2]

Abstract

The body cavity region is positioned as a boundary region between skeletal muscle and other organ tissues in the surface layer. Previous studies have addressed the application of body cavities to the classification of subcutaneous adipose tissue. However, the body cavity as an inscribed surface of skeletal muscle in the surface layer has not been discussed. In this study, we propose a segmentation method for body cavity regions in torso CT images and compare its accuracy with three U-Net architectures: 2D U-Net, 2D UNet++ and 3D U-Net. In this paper, we first compare the segmentation accuracy between 2D architectures: U-Net and UNet++, and patch-based 3D architecture using hold-out method. Then, using the network with the best segmentation accuracy, we evaluate the accuracy of body cavity segmentation by 4-fold cross-validation. Dice values were used to compare the results of segmentation with the ground truth of the torso CT images in 20 cases. The segmentation accuracies of 2D U-Net, 2D UNet++ and 3D U-Net were 98.5%, 92.5% and 94.0%, respectively. Therefore, we performed a 4-foldcross-validation using 2D U-Net, and found that the Dice, Precision and recall values were 98.5%, 98.9%, and 98.3% on average, respectively. The segmentation accuracy in each vertebral section was evaluated, and the results showed that the thoracic, lumbar, and sacral vertebrae and caudal sections were99.3%, 98.3% and 96.1%, respectively. In the proposed method, the segmentation of the body cavity in torso CT image was realized with high accuracy of average 98.5% using 2D U-Net. In the future, it is necessary to verify whether the body cavity region obtained by the proposed method can be used as a pre- and post-processing for other computer-aided diagnosis (CAD) systems.

Keywords：body cavity, segmentation, U-Net, UNet++, 3D U-Net

1. Purpose

The body cavity is covered by bones and skeletal muscle and contains many organs. Therefore, the body cavity region provides important spatial and positional information in computer-aided diagnosis (CAD)

*1 School of Information Science and Technology, Aichi Prefectural University
〔1522-3 Ibaragabasama, Nagakute, Aichi 480-1198, Japan〕
e-mail: is171078@cis.aichi-pu.ac.jp
*2 Department of Electrical, Electronic and Computer Engineering, Faculty of Engineering, Gifu University.
*3 Department of Radiology, Graduate School of Medicine, Gifu University.
*4 Radiology Service, Gifu University Hospital.

systems. In the previous studies, the body cavity area is used to divide subcutaneous and visceral fat region [1]. In addition, extraction of the boundary to divide the body cavity into regions of the thoracic cavity, peritoneal cavity, and pelvic cavity from torso CT images [2] and segmentation of the thoracic cavity and peritoneal cavity using multi-resolution 3D U-Net [3] were performed. However, the body cavity as an inscribed surface of skeletal muscle in the surface layer has not been discussed. The body cavity region is positioned as a boundary region between skeletal muscle and other organ tissues in the surface layer. In this study, we propose a segmentation method for body cavity regions in torso CT images and compare its accuracy with three U-Net architectures: 2D U-Net, 2D UNet++ and 3D U-Net.

2. Method

In this study, we propose an automatic segmentation method of the body cavity in torso CT images using the U-Net architecture, which has produced good results in the segmentation task of medical images. We use U-Net and UNet++ as the 2D architecture, and patch-based 3D U-Net as the 3D architecture. UNet++ is an improved version of U-Net that can be represented in a nested U-Net and exceeds the accuracy of U-Net for some tasks [4].

The size of the input CT image is 512×512×802-1104[voxel]. At the training phase, data augmentation is performed by translation of [-25, +25] % of the matrix size, rotation of [-10, +10] deg, scale of [-35, +35] %, shear transform with the shear angle of [-π/8, +π/8] rad, and flipping in the right-left direction [5]. In the 2D architecture, all axial slices of 512 × 512 [pixel] are inputted, and in the 3D U-Net, the patch of 512×512×8 [voxel] is made, and it is inputted by sliding in the body axis direction. In the post processing, in order to eliminate the erroneous extraction of the region outside the body cavity, the three-dimensional labeling around 26 is applied to the segmentation result, and the region with the largest volume is made to be the body cavity.

It is defined as ground truth including the region including psoas major and iliac muscles which are deep muscles in the body cavity region. The ground truth is generated by manual extraction for all axial cross sections. In the evaluation experiment, concordance rate between ground truth and segmentation result is evaluated by Dice values for 20 cases of torso CT image. In this paper, we first compare the segmentation accuracy between 2D architectures: U-Net and UNet++, and patch-based 3D architecture using hold-out method. Then, using the network with the best segmentation accuracy, we evaluate the accuracy of body cavity segmentation by 4-fold cross-validation.

3. Results

Automatic body cavity segmentation was performed in 20 torso CT images, and the Dice values of U-Net and UNet++ were 98.5% and 92.5% for U-Net with 2D architecture, respectively, compared to 94.0% for U-Net with 3D architecture. Therefore, four-fold cross-validation was performed using 2D U-Net, which had the highest segmentation accuracy. The mean of Dice, Precision and Recall rate were 98.5%, 98.9% and 98.3%, respectively.

In order to pseudo-evaluate the segmentation accuracy in the thoracic, abdominal, and pelvic cavity regions, the mean Dice values were 99.3%, 98.3%, and 96.1% in the thoracic, lumbar, and sacrococcygeal cross-sectional regions, respectively. Figure 1 shows the segmentation result of the body cavity using each network superimposed on the original image for each vertebral body section. The recognition result showed the high value in order of thoracic vertebrae, lumbar vertebrae, sacrum and coccyx region, even if any network was used.

Figure 2 shows the segmentation results of body cavity by 2D U-Net in torso CT images and whole-body CT images. Here, the volume rendering was carried out simultaneously with the skeleton in order to show the anatomical positional relation and outline of the body cavity. Right figure shows the result of applying the proposed method to the whole-body CT image with the amyotrophic disease. Though it is a visual evaluation, the good segmentation result was obtained even in the image including limbs. We have shown that automatic body cavity recognition is possible even in CT images with different imaging areas, although differences in the imaging conditions need to be considered.

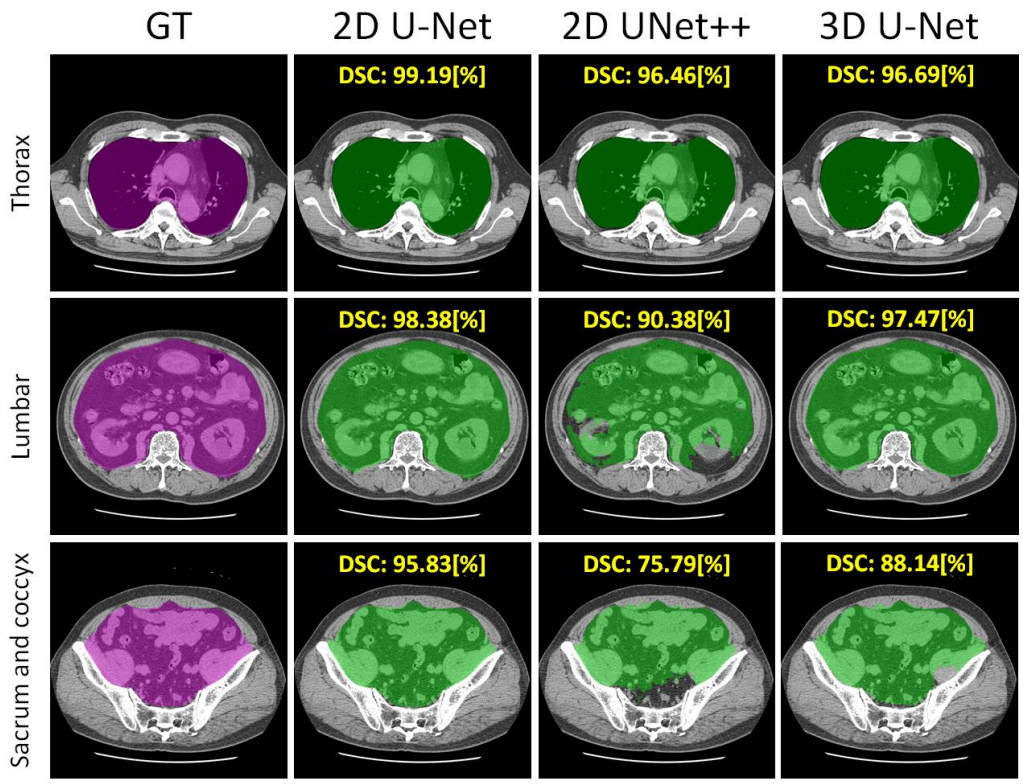


Fig. 1 Predicted results in each vertebral section.

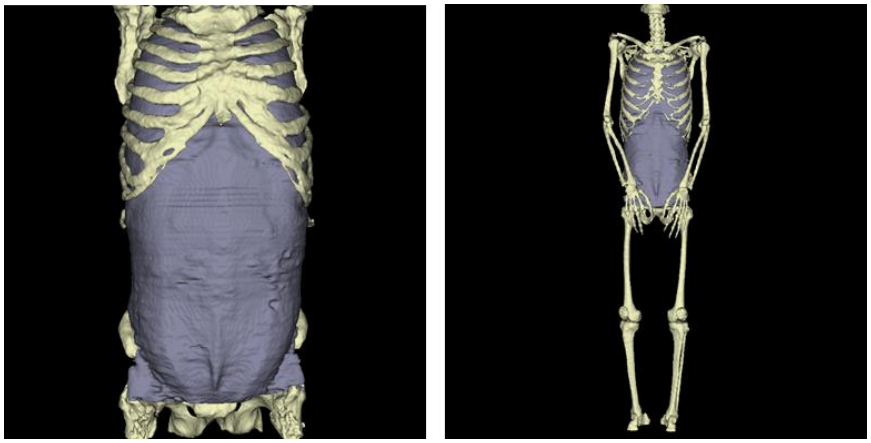


Fig. 2 Volume rendering of the segmentation result in torso and whole-body CT images

4. Conclusion

In this study, we proposed a segmentation method of body cavity region in torso CT images and compared its accuracy by three U-Net architectures: 2D U-Net, 2D UNet++, and 3D U-Net. We performed a 4-fold cross-validation using 20 cases, and the highest accuracy of body cavity segmentation was achieved using 2D U-Net, with a mean Dice coefficient of 98.5%. In the future, it is necessary to verify whether the body cavity region obtained by the proposed method can be used as a pre- and post-processing for other CAD systems. For example, there is a possibility of deletion of false extractions in the recognition of skeletal muscles around the body cavity. For this purpose, it is necessary to use not only dice values but also surface topography as the evaluation of recognition results.

Competing interests

The authors declare that they have no conflict of interest.

References

[1] Takahashi M, Takenaga T, Nomura Y et al: A preliminary study on creating a ground-truth image for deep learning-based body fat segmentation in MRI sing multi-atlas segmentation. Proceedings of Computer Assisted Radiology and Surgery 2020, S205, 2020

[2] Naramura I, Hayashi I, Zhou X et al: Automated recognition of body cavities in torso X-ray CT images. Med Imag Tech **26**: 261-271, 2008

[3] Koitka S, Kroll L, Malamutmann, E et al: Fully-automated body composition analysis in routine CT imaging using 3D semantic segmentation convolutional neural networks. arXiv preprint arXiv:2002.10776, 2020

[4] Zhou Z, Siddiquee MMR, Tajbakhsh N et al: UNet++: A nested U-Net architecture for medical image segmentation. Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, 3-11, 2018

[5] Hiasa Y, Otake Y, Takao M et al: Automated muscle segmentation from clinical CT using bayesian U-Net for personalized musculoskeletal modeling, IEEE Trans Med Imag **39**: 1030-1040, 2019

機械学習に基づく食道・頭頸部癌併発リスク予測の為の診断支援システムに関する検討

吉村　裕一郎[*1]　上里　昌也[*2]　浦濱　竜馬[*2]　加藤　直菜美[*2]

米本　昇平[*2]　中口　俊哉[*1]

要旨

食道内視鏡検査において，ルゴール染色によって多発ヨード不染（まだら不染）帯を観察することが行われている．まだら不染帯は，食道内や頭頸部領域における異時性・同時性癌の併発リスクと関連していることが指摘されているが，医師の主観に基づく観察のみでは，併発リスクの予測可能性は 39.5%である．そこで本研究では，内視鏡検査映像中から，高リスクシーンを自動的に選択し，選択されたシーンの画像を機械学習によって解析することで，重複・併発癌発生リスクを予測する手法を提案する．高リスク患者を内視鏡検査時に選別することで，経過観察によって治療機会の損失を防ぐことを目的とする．

実験では，千葉大学医学部附属病院で得られたデータを使用し，専門医監修の基，エキスパートの医師が，臨床的知見によって選択した画像による診断精度と，提案手法による比較を行った．提案手法の精度は 0.8(AUC:0.792)となり，最も高性能であることを確認した．

キーワード：内視鏡検査，多発ヨード不染，機械学習，画像認識

1. はじめに

食道は発生学的に外胚葉と呼ばれる組織を起源としており，皮膚，咽頭，喉頭，口腔，肺などに分化する．そのため食道癌は，他の臓器の癌と併発することが多々見られる[1]．しかし頭頸部領域は所謂“のど”の検査であり，狭い空間であることや，嘔吐反射の起こり易さに起因し，専門医においても見逃しを防ぎきる事は困難である．

そこで我々の研究グループでは，上部消化管内視鏡検査時に行うルゴール染色による多発ヨード不染帯（multiple lugol-voiding lesions; multiple LVL）に着目している．健常者の食道の粘膜上皮にはグリコーゲン顆粒が含まれており，ルゴール液の散布を行えば，ヨウ素デンプン反応により食道は褐色に染色される．しかし上皮が癌化するとグリコーゲン顆粒を含まなくなるために，褐色にならず不染と呼ばれる状態となる．この上皮が白くまだら状に広がる現象が“まだら不染”や“まだら食道”と呼ばれる状態であり，早期癌の検出に有用であるとされている[2]-[5]．

現在まだら不染は，白い模様の数や密度などによって定義され，それらの分布と早期食道癌や頭頸部癌との間に関連性が観られることが示されている．例えば堅田ら[5]は不染の数に着目し，一回の食道内視鏡検査時において，最も不染領域の数が多く観察されるシーンにおいて，その不染領域の個数が 10 個以上の場合，有意に頭頸部重複癌が存在すると指摘している．

そこで本研究では，まだら不染帯を画像処理技術によって自動的に検出することにより，同

*1 千葉大学フロンティア医工学センター
〔〒263-8522 千葉市稲毛区弥生町 1-33〕
e-mail: yysmr@chiba-u.jp
*2 千葉大学医学部　先端応用外科学
投稿受付：2020 年 8 月 5 日

時性，異時性の食道・頭頸部癌併発リスクの有無を予測することを試みる．

2．従来手法

世界に先駆け消化管内視鏡を開発した本邦では，現在においても機器の使用，研究，開発において先進的な位置にある．特に上部から挿入を行う内視鏡に関しては，診断や治療に関しても先駆け的な位置付けにある．本研究の基本部である早期食道癌発見の為のルゴール染色法は，日本においては 1974 年に実施され，現在もその早期癌発見への確実性から維持・継続されている．しかし本研究の背景である，食道粘膜におけるまだら不染の程度による，多発食道癌，頭頸部癌の発生リスクに関しての報告は，医師の主観に基づくものが多く，折角画像計測機器を用いながらも，定量的な報告があるとは言い難い．例えば文献 2 においては，Muto らは不染の大きさを 5mm 以上と定義しているが，画像中或いは食道内において，どのようにしてサイズを計測したかに関する言及はなく，実際には観察者の主観に基づいているものと考えられる．

また複数の文献において，まだら不染が多く認められる食道癌患者は，同時性・異時性食道癌が多いという言及がなされている．しかし同様の患者において，食道内多発癌や重複癌を併発（合併）し易いと指摘している文献は堅田らの研究グループのみであり，これらの知見に定量性を持たせることは有用であると考えられる．

そこで 2010 年 1 月から 2012 年 8 月の間に，千葉大学医学部附属病院にて内視鏡的粘膜下層剥離術（Endoscopic Submucosal Dissection; ESD）を施行した 50 症例を対象とし，堅田の手法に基づき，重複癌の予測性能を検証した．これらの症例データは 2018 年 12 月に電子カルテ上から取得したものであり，ESD 施行後，データ取得時までの重複・併発癌に関する情報や，堅田の基準に基づいた診断，ヨード不染帯の個数などが記されている．表 1 に，同データに基づいた，まだら不染の数，堅田基準と併発癌の有無について，それぞれ相関値を示す．使用データの内，重複・併発癌有りの症例は 36 例である．専門医が堅田の基準に基づき行った予測によると，重複・併発癌の予測性能は 0.48，感度 0.73 であり，文献 5 における，0.395（多発ヨード不染帯がある症例の内，同時性・異時性癌のある症例の割合）と比較しても，予測可能性を有することが示唆された．一方でヨード不染帯の個数や堅田の基準と併発癌の有無については，相関が無いことから，スクリーニングの段階にノウハウがある可能性が考えられる．

表 1.堅田[5]の基準に基づく指標と相関値

	ヨード不染数	堅田分類	重複・併発癌
ヨード不染数	1		
堅田分類	0.842	1	
重複・併発癌	0.110	0.071	1

よって提案手法では，従来法の知見に学び，まだら不染帯の多い高リスク画像を検査中に自動的に選び出す．検出したシーンに対して併発の有無を予測することで，食道・頭頸部癌併発リスク提示するシステムを構築することを考える．

3．提案手法

1）実験データの収集

本研究では，2 章で使用した 50 症例に加え，同様に 2012 年 9 月から 2013 年 5 月の間に，千葉大学医学部附属病院にて ESD を施行し，2019 年 7 月にデータを取得した 35 症例を合わせた，計 85 症例のデータを使用する．85 症例の内，35 症例に同時性・異時性癌が生じており，いずれの症例についても，治療日と共に，データ取得日までの間に生じた癌の有無やその箇所に関する記録を持つ．但し，新たに取得したデータに，専門医による堅田の基準に基づいた診断に関する記録はついていない．またいずれの症例も，治療前に内視鏡検査を行った際に撮影された，ほぼ全ての検査画像を含んでいる．図 1 に取得した画像の例を示す．本研究では検査中の画像の内，ルゴール染色後の画像を対象とし，使用する．

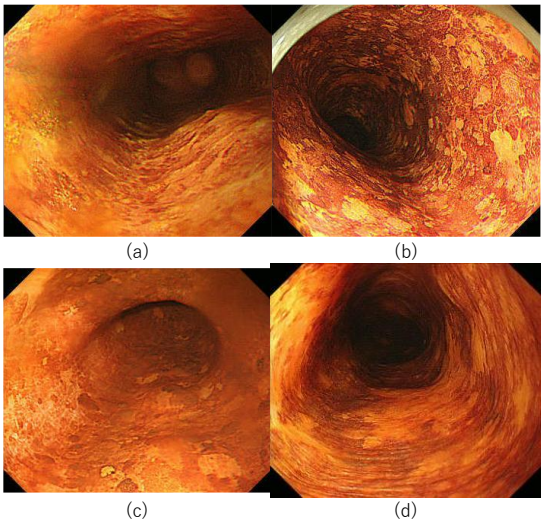


図 1 実験データ例

2）まだら不染帯の検出手法の検討

まだら不染帯の検出は，物体検出法に基づき，不染帯をそれぞれ識別し，その数や大きさについて評価する方法と，領域分割法に基づき，不染領域を pixel 単位で抽出する主に 2 つの手法が考えられる．ここで，後者を採用した場合，教師データとなるまだら不染帯のアノテーション作業において，使用している機器の差など，作業者の診断能以外の個人差が生じる[6]こと．また本研究において使用するデータは，異時性癌を含むことから，治療日より少なくとも 5 年以上が経過したデータを使用している．よって必ずしも同一の施設や装置により，取得されたデータではないことから，前者の手法を採用することとした．

3）不染帯検出に基づく高リスク画像選択

検査画像群中から高リスク画像を選択する．具体的には，1 回の検査中に撮影した画像群の中からルゴール染色済みの画像を取得する．まだら不染の個数を自動的にカウントし，最も不染領域数の多い画像を高リスク画像として 1 枚選択する．

収集した症例の画像中から，日本消化器内視鏡学会消化器内視鏡専門医による監修の基，エキスパートの医師 2 名（以下専門医）によって，食道の上部，中部，下部それぞれの領域において，最もまだら不染帯が多く確認できる画像を，

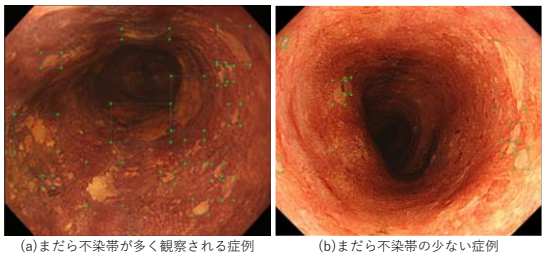


図 2 実験データ例

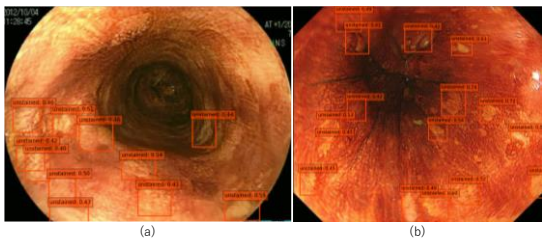


図 3 まだら領域検出結果例

それぞれ 1 枚ずつ選択する．まだら不染が確認できない場合は，観察位置はそのままに，腫瘍部のなるべく映らない画像を採用した．アノテーション画像の例を図 2 に示す．アノテーションは図 2 に示すように，原画像から患者情報などを含む，外枠部を除いたものを使用し，まだら不染帯それぞれを囲むように行われた．

まだら不染帯検出のための手法としては，M2Det[7]を使用する．Backbone は VGG-16 を使用し，学習回数は 100epochs，画像サイズは 300×300pixels，バッチサイズは 1 とする．ハイパーパラメータについては Optuna[8]を用いて最適化を行い，最適化手法には SGD が採用された．図 3 に検出結果例を示す．検出結果を確認したところ，人間によるアノテーションと比較して，より細かな不染帯についても検出が出来ていることを確認した．

提案手法においては，検出されたまだら不染帯の個数を指標とし，高リスク画像を選択する．1 検査中において，ルゴール染色後の画像の中から最も不染帯の数の高い画像を高リスク画像として選択した．

4）機械学習に基づく併発リスクの予測

前述のまだら不染帯検出法により選択された，高リスク画像を入力とし，重複・併発癌発生有無の予測を行う．機械学習法を使用し，入

表 2. 各手法の予測精度

	従来手法	提案手法
EfficientNet-B0	0.6	0.68
EfficientNet-B1	0.64	0.72
EfficientNet-B2	0.62	0.74
EfficientNet-B3	0.62	0.8
EfficientNet-B4	0.64	0.7
EfficientNet-B5	0.6	0.64
EfficientNet-B6	0.6	0.68
EfficientNet-B7	0.58	0.7
Ensemble	0.7	0.68
BOW	0.72	0.6

力画像に対して，患者情報より得られた，食道・頭頸部癌重複及び併発の有無を識別するモデルを作成した．

4. 実験と考察

まだら不染帯検出に基づく，高リスク画像選択手法の有効性について検証する．専門医によって従来の知見に基づいて選ばれた高リスク画像と，提案手法によって選ばれた高リスク画像を使用し，それぞれ機械学習法を用いてモデルを作成し，重複・併発癌予測性能を検証した．

実験には教師データとして 2,3 章で示した 85 症例の中から，治療日順に 2010 年から 2012 年までの 60 症例（重複・併発癌有り 25 例，無し 35 例）と，テストデータとして以降の 25 症例（重複・併発癌有り 9 例，無し 16 例）を使用する．

高リスク画像選択のためのまだら不染帯検出手法については，教師データ 60 症例の中から， 3 章と同様に，専門医によって 1 症例当たり 3 枚の画像を選択し，まだら不染帯をマーキングした教師データを作成する．教師データを 5 分割して学習を行い，学習に用いなかった症例のデータについては，専門医によって選択されていないルゴール染色済みの画像についても，まだら不染帯の検出を行う．これらを繰り返すことで，60 症例全ての高リスク画像を 1 枚ずつ選択する．選択された高リスク画像を教師データとして，機械学習法を使用し，重複・併発癌の発生有無を教師として学習を行った．ま

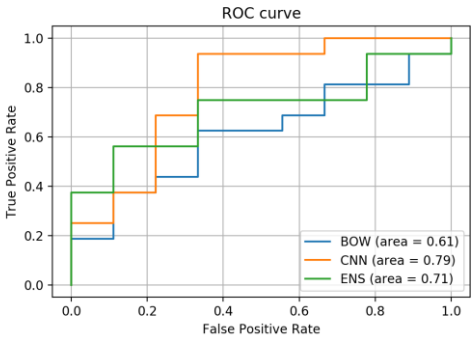
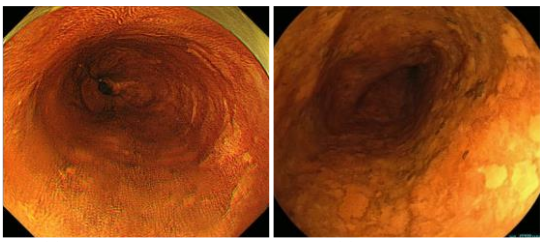


図 4 各手法の ROC と AUC



(a)異時性食道癌のある症例　(b)併発癌の無い症例

図 5 予測に失敗した症例の例

たテストデータの検証では，前述のモデルに対して，各症例のルゴール染色済みの画像を全て適用し，高リスク画像を 1 枚選択する．選択された画像を，機械学習法に適用することで，症例ごと癌の発生有無を予測する．予測結果を実際の結果と比較することにより，性能の評価を行った．

機械学習法としては， EfficientNet[9]，AKAZE[10]を用いた Bag of Visual Words，PCA で画像の次元数を圧縮し，複数のクラス分類手法（SVM，Random Forest，Logistic Regression，k-nearest neighbor，naïve bayes）の出力結果をアンサンブル学習させたものによって比較検証を行った．EfficientNet は ImageNet で pre-trained 済みのモデルに対して，最終層の出力を変更し，再学習の範囲はハイパーパラメータとして fine-tuning を行った．学習回数は 100epochs，画像サイズは 224×224pixels，バッチサイズは 16 とし，最適化手法には AdaBound[11]を使用した．

表 2 に提案手法に基づく，重複・併発癌の発生有無予測性能を示す．また図 4 に CNN，Bag of Words，アンサンブル学習を用いた場合の ROC を示す．表 2 より，提案手法を使用し，

EfficientNet-B3 を用いた際に，最も性能が高くなることを確認した. またこの時の感度は 0.875, 特異度は 0.67 であった．よって従来の知見や，実験に用いたデータを医師が診断した場合と比較し，提案手法を用いることにより，重複・併発癌予測性能が向上する事が示唆された．

図 5 に予測に失敗した画像例を示す．(a)は異時性食道癌が後に生じた症例であるが，提案手法では予測に失敗している．画像より，食道内にまだら不染帯は殆ど確認できないことが解る．物体検出手法を用いた段階においても，検出した不染領域の数は確信度 0.4 を閾値とした場合に 9 個であり，従来の診断基準を満たしていない．このように撮影時にはまだら不染が観察されなくとも，数年後に異時性癌が生じる症例も存在する．従って，同時性癌の生じた症例をより多く収集し，臨床的な解析を行うことや，異時性癌患者の経時変化を観察可能なデータを収集する必要があると考えられる．(b)の画像は重複・併発癌が生じていないが，提案手法ではリスク有りと判断された症例である．画像中に多くの不染領域が観察されるが，転移の生じていない症例について，今回の検証では食道・頭頸部癌以外の部位に関する治療歴や病歴に関しては考慮していないことから，その他の要因についても今後検証していく必要があると考えられる．

5. まとめ

本研究では，食道内視鏡検査におけるまだら不染帯を，画像処理技術によって自動的に検出することにより，同時性・異時性の食道・頭頸部癌重複・併発リスクの有無を予測することを試みた．初めに内視鏡検査中の画像から，専門医によって選ばれた画像群より，まだら不染帯を検出するモデルを作成する．次に作成したモデルを使用し，不染個数の多い領域を高リスク画像として選択する．選択された画像群を入力として機械学習を用いることにより，各症例に対して，重複・併発癌リスクの有無を予測する手法を提案した．実験によって提案手法の性能を確認したところ，従来の専門医の所見と比較し，正確な予測が行われていることを確認した．

今後の課題として，現在提案手法を用いたリアルタイムリスク提示ソフトウェアの開発が完了しているものの，現状では性能の検証が行うことが出来ないことが挙げられる．併発リスクの予測には，ESD 後少なくとも数年単位での追跡調査が必要であるが，通常の内視鏡の検査時の全動画像の保存は，特別な事情を除き，現在もまだ容量の都合から必ずしも行われてはいない．よって現在は，85 症例の検査時に撮影されたデータのみによって，モデルの作成と評価を行っていることから，今後は中長期的なスパンによりデータを収集し，提案手法の工学的，臨床的有効性を検証する．

利益相反の有無

なし

文　献

［1］ 辻直子，石黒信吾，真能正幸，他 : まだら食道の病理．胃と腸，vol.36, 8, 1049-1056, 2001

［2］ M Muto, S Hironaka, M Nakane, et al. : Association of multiple Lugol-voiding lesions with synchronous and metachronous esophageal squamous cell carcinoma in patients with head and neck cancer. Gastrointestinal Endoscopy, vol.56, 4,517-521,2002

［3］ 長井健悟，石原立，松井芙美，他 : まだら食道を背景とした食道癌の初期浸潤所見．胃と腸，vol.47,9,1419-1426,2012

［4］ 竹内学，橋本哲，小林正明，他 : 病変の形態からみた発育進展 初期病巣から粘膜癌までを中心に 表層拡大型癌の発育進展．胃と腸，vol.47,9,1410-1417,2012

［5］ 堅田親利，武藤学 : 背景因子からみた異時性多発食道癌の長期経過 多発ヨード不染帯との関連を中心に．胃と腸，vol.42,9,1355-1363,2007

［6］ S. A. A. Kohl, B. Romera-Paredes, C. Meyer, et al.:A Probabilistic U-Net for

Segmentation of Ambiguous Images. Advances in Neural Information Processing Systems 31, 6965-6975, 2018

［7］ Q. Zhao, T. Sheng, Y. Wang, et al.: M2Det: A Single-Shot Object Detector based on Multi-Level Feature Pyramid Network. AAAI 2019

［8］ T. Akiba, S. Sano, T. Yanase, et al.: Optuna: A Next-generation Hyperparameter Optimization Framework. KDD '19: Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining,2623–2631,2019

［9］ M. Tan, Q. V. Le: EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. ICML2019

［10］ P. F. Alcantarilla, J. Nuevo and A. Bartoli: Fast Explicit Diffusion for Accelerated Features in Nonlinear Scale Spaces. In British Machine Vision Conference, Bristol, 2013

［11］ L. Luo, Y. Xiong, Y. Liu et al. :Adaptive Gradient Methods with Dynamic Bound of Learning Rate. ICLR 2019

Diagnostic support system for early detection of esophageal and head and neck cancer by using machine learning

Yuichiro YOSHIMURA[*1], Masaya UESATO[*2], Ryuma URAHAMA[*2], Manami KATO[*2]
Shohei YONEMOTO[*2], Toshiya NAKAGUCHI[*1]

*1 Center for Frontier Medical Engineering, Chiba University
*2 Department of Frontier Surgery, Chiba University Graduate School of Medicine

In the esophageal endoscopy, multiple Lugol-voiding leisions (multiple LVL) have been observed by Lugol staining. Multiple LVL has been noted with a risk of coexisting heterogeneous and concurrent cancers in the esophagus, head and neck area. However, the accuracy of predicting the risk of these cancers is 39.5% based on physician's subjective. In this study, we propose a method for predicting the risk of metastasis by automatically selecting high-risk scenes in endoscopy images then analyzing images of the selected scenes by using machine learning. To screen high-risk patients at the time of endoscopy to prevent chance to receive a treatment. In the experiment with clinical data obtained from Chiba University Hospital, the diagnostic accuracy of predicting metastasis risk was 80%.

Key words: endoscopy, multiple Lugol-voiding leisions, machine learning, image inspection

Preliminary Study of Generalized Semiautomatic Segmentation for 3D Voxel Labeling of Lesions Based on Deep Learning

Yukihiro NOMURA[*1], Shouhei HANAOKA[*2], Tomomi TAKENAGA[*1], Takahiro NAKAO[*1], Hisaichi SHIBATA[*1], Soichiro MIKI[*1], Takeharu YOSHIKAWA[*1], Takeyuki WATADANI[*2], Naoto HAYASHI[*1], Osamu ABE[*2]

Abstract

In the development of computer-aided detection software, a three-dimensional (3D) voxel labeling of lesions requires significant radiologists' effort. To reduce the time required for the 3D voxel labeling, we aimed to develop a generalized semiautomatic segmentation method based on deep learning. In the preliminary study, we investigated whether a single type of lesions can train a semiautomatic segmentation model for two types of lesions with different contrasts from the surroundings. We targeted lung nodules in chest CT images and liver lesions of hepatobiliary-phase images in Gd-EOB-DTPA-enhanced MR imaging. We input the 32 × 32 × 32 isotropic volume of interest around the center of gravity of the lesion and the spherical mask determined by the maximum diameter of the lesion into a 3D FC-ResNet model to define the label of the lesion. We trained the model for each type of lesions with data augmentation, including random shifting and random monochrome inversion. The evaluation results showed that the Dice coefficients using the model trained on lung nodules were 0.728±0.089 for the lung nodules, 0.625±0.149 for the liver lesions. The Dice coefficients using the model trained on liver lesions were 0.489±0.232 for the lung nodules, 0.798±0.078 for the liver lesions. The model trained on lung nodules can label liver lesions as well as lung nodules.

Keywords：semiautomatic segmentation, deep learning

1. Purpose

In the development of computer-aided detection software, a three-dimensional (3D) voxel labeling of lesions requires significant radiologists' effort. To reduce the time required for the 3D voxel labeling, we aimed to develop a generalized semiautomatic segmentation method based on deep learning to implement into our web-based image database system, named CIRCUS DB [1]. In the preliminary study, we investigated whether a single type of lesions can train a semiautomatic segmentation model for two types of lesions with different contrasts from the surroundings.

2. Method

We targeted lung nodules in chest CT images and liver lesions of hepatobiliary-phase images in Gd-EOB-DTPA-enhanced MR imaging.

*1 Department of Computational Diagnostic Radiology and Preventive Medicine, The University of Tokyo Hospital
〔7-3-1 Hongo, Bunkyo-ku, Tokyo, 113-8655, Japan〕
e-mail: nomuray-tky@umin.ac.jp
*2 Department of Radiology, The University of Tokyo Hospital

1) Datasets

[Lung nodules in chest CT images] 311 cases of chest CT volumes with 400 lung nodules were collected from our hospital. The datasets were examined with a GE LightSpeed CT scanner. The original voxel size was 0.781 × 0.781 × 1.250 mm^3. Each case included at least one nodule, and the diameter of these nodules was ranged from 5 mm to 25 mm. For each nodule, a board-certified radiologist defined its ground truth by voxel-by-voxel painting. Among the 311 cases, 231 cases with 300 nodules were used for training, 38 cases with 50 nodules were used for validation, and the remaining 42 cases with 50 nodules were used for evaluation.

[Liver lesions of hepatobiliary-phase images in Gd-EOB-DTPA-enhanced MR imaging] 137 Gd-EOB-DTPA-enhanced MR images including 152 hepatocellular carcinomas (HCC), 88 metastatic liver tumors, six hemangiomas, and 146 cysts were used in this study. 3-Tesla MR scanners (Siemens Skyra, GE Signa HDxt) and 1.5-Tesla MR scanners (Siemens Avanto, GE Signa HDxt/EXCITE) with a T1-weighted 3D gradient echo sequence were employed for acquisition of MR images. The original voxel size was 0.7–1.3 × 0.7–1.3 × 2.5 or 3.0 mm^3 when the slice thickness was 5.0 or 3.0 mm. Each case included at least one lesion, and the diameter of these legions was ranged from 5 mm to 25 mm. For each lesion, radiological technologists defined its ground truth by voxel-by-voxel painting, and defined regions were checked by a board-certified radiologist. Among the 137 cases, 107 cases with 302 lesions were used for training, 14 with 42 lesions were used for validation, and the remaining 16 cases with 48 lesions were used for evaluation.

2) Training of deep learning models

The images were resampled using a trilinear interpolation to obtain the isotropic volume, for which the voxel size was 1.0 × 1.0 × 1.0 mm^3. For each lesion, the 32 × 32 × 32 cubic volume of interest (VOI) around the center of gravity of the lesion was extracted, and the voxel value of the VOI was normalized by subtracting the mean value of the VOI then dividing by the standard deviation. We input the VOI and the spherical mask determined by the maximum diameter of the lesion into a 3D FC-ResNet [2] model to define the label of the lesion. The model was implemented using PyTorch version 1.5.0. We trained the model for each type of lesions with data augmentation, including random shifting and random monochrome inversion.

In the training of each model, we carried out 200 trials of hyperparameter tuning with the asynchronous parallel Bayesian optimization algorithm [3]. The number of parallel workers was set to four. The tuned hyperparameters of the model were the parameter n of 3D FC-ResNet model (1-10), the batch size ($2b$, b = 2–32), the loss function (dice loss or focal Tversky loss [4]), two parameters of the rectified adaptive moment estimation (RAdam) [5] ($\alpha_0 = 10^{-2}$–10^{-6}, β_1 = 0.9–0.99), and the number of data augmentations for each lesion (1–5). This processing was carried out on the Reedbush-H supercomputer system at our institute, which consists of 120 computing nodes equipped with two Intel Xeon E5-2695v4 processors, a memory of 256 GB, and two NVIDIA Tesla P100 GPUs.

3. Results

The evaluation results showed that the Dice coefficients using the model trained on lung nodules were 0.728±0.089 for the lung nodules, 0.625±0.149 for the liver lesions. The Dice coefficients using the model trained on liver lesions were 0.489±0.232 for the lung nodules, 0.798±0.078 for the liver lesions. Figures

1 and 2 show the segmentation results of lung nodules and liver lesions, respectively.

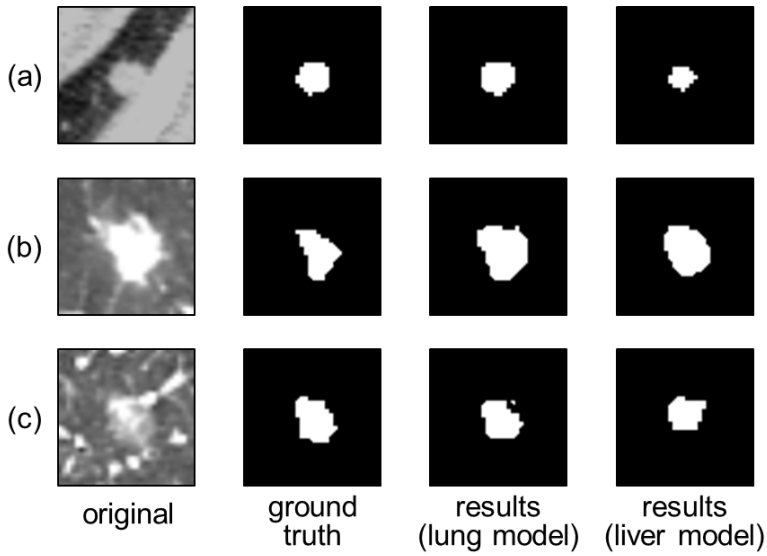


Fig. 1 Segmentation results of lung nodule. (a) 7 mm solid nodule, (b) 14 mm solid nodule, (c) 11 mm pure ground-glass nodule. First column: original image, second column: ground truth, third column: segmentation results using the model trained on lung nodules, fourth column: segmentation results using the model trained on liver lesions.

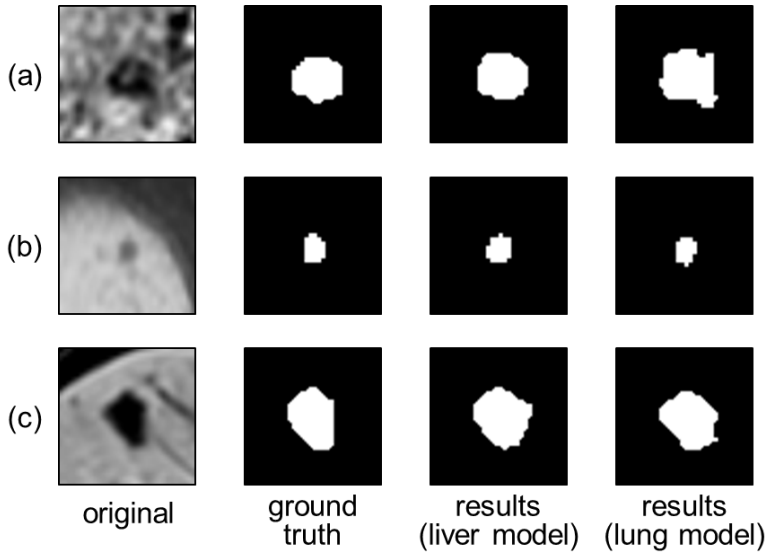


Fig. 2 Segmentation results of liver lesion. (a) 12 mm HCC, (b) 6 mm cyst, (c) 14 mm metastatic liver tumor. First column: original image, second column: ground truth, third column: segmentation results using the model trained on liver lesions, fourth column: segmentation results using the model trained on lung nodules.

4. Conclusion

We have investigated whether a single type of lesions can train a semiautomatic segmentation model for two types of lesions with different contrasts from the surroundings. According to our results, the model trained on lung nodules can label liver lesions as well as lung nodules.

Acknowledgement

This work was supported by Japan Society for the Promotion of Science (JSPS) KAKENHI Grant No. 18K12096. It was also supported by the Joint Usage/Research Center for Interdisciplinary Large-scale Information Infrastructures and High Performance Computing Infrastructure projects in Japan (Project IDs: jh190047-DAH and jh200042-DAH).

Conflict of interest

Yukihiro Nomura, Tomomi Takenaga, Takahiro Nakao, Hisaichi Shibata, Soichiro Miki, Takeharu Yoshikawa, and Naoto Hayashi are belong to an endowment department, sponsored by HIMEDIC Inc. and Siemens Healthcare K.K.; however, the status is irrelevant to the contents of the study. The remaining authors have no other conflicts of interest related to the present study.

References

[1] Nomura Y, Miki S, Hayashi N, et al.: Novel platform for development, training, and validation of computer-assisted detection/diagnosis software. Int J Comput Assist Radiol Surg **15**: 661-672, 2020

[2] Takenaga T, Hanaoka S, Nomura Y, et al.: Four-dimensional fully convolutional residual network-based liver segmentation in Gd-EOB-DTPA-enhanced MRI. Int J Comput Assist Radiol Surg **14**: 1259-1266, 2019

[3] Nomura Y, Sato I, Hanawa T, et al.: Development of training environment for deep learning with medical images on supercomputer system based on asynchronous parallel Bayesian optimization. J Supercomput **76**: 7315-7332, 2020

[4] Abraham N, Khan NM: A novel focal Tversky loss function with improved attention U-Net for lesion segmentation. Proc. of 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019): 683-687, 2019

[5] Liu L, Jiang H, He P, et al.: On the variance of the adaptive learning rate and beyond. arXiv: 1908.03265, 2019

Preliminary Study of Perforation Detection and Localization for Colonoscopy Video

Kai Jiang*[1], Hayato Itoh*[1], Masahiro Oda*[1], Taishi Okumura*[2]
Yuichi Mori*[2], Masashi Misawa*[2], Takemasa Hayashi*[2],
Shin-Ei Kudo*[2], Kensaku Mori*[1,3,4]

Abstract

Endoscopic submucosal dissection (ESD) is a minimally invasive treatment for early gastric cancer. In ESD, a physician directly removes the mucosa around the lesion under internal endoscopy by using the flush knife. However, the flush knife may accidentally pierce the colonic wall and generate a perforation on it. If physicians overlook a small perforation, patient may need emergency open surgery since perforation can easily cause peritonitis. Therefore, there is a potential demand for computer-aided diagnosis (CAD) system, which prevents an overlook of a perforation by detecting and localizing the perforation. In order to build a CAD system for ESD, we developed a model of perforation detection and localization for colonoscopic videos by using YOLO-v3. We annotated 2,000 and 200 perforation images for training and test, respectively, from colonoscopic videos of totally 11 patients. Our model achieved 0.835 accuracy and 0.746 mean average precision for perforation detection and localization, respectively, of the test images.

Keywords：Deep learning, YOLO-v3, Colonoscopic Video, Colon-perform localization

1. Introduction

Endoscopic submucosal dissection (ESD) is a minimally invasive treatment for early gastric cancer. In ESD, a physician directly removes the mucosa around the lesion under internal endoscopy by using the flush knife. However, the flush knife may accidentally pierce the colonic wall and generate a perforation on it. If physicians overlook a small perforation, patients may need emergency open surgery, since perforation can easily cause peritonitis [1]. Therefore, there is a potential demand for computer-aided diagnosis (CAD) system, which prevents an overlook of a perforation by detecting and localizing a perforation. Purpose of this work is to construct a model for perforation detection and localization in scenes of colonoscopic videos. To construct the model, we use a variant of You Only Look Once (YOLO-

*1 Graduate School of Informatics,
Nagoya University〔Furo-cho, Chikusa-ku, Nagoya 464-8601, Japan〕
e-mail: kjiang@mori.m.is.nagoya-u.ac.jp
*2 Digestive Disease Center, Showa University Northern Yokohama Hospital
*3 Information Technology Center, Nagoya University
*4 Research Center for Medical Bigdata, National Institute of Informatics

v3) [2], which can classify perforation scenes and detect location of perforation simultaneously. Compared with traditional localization methods, e.g. Faster R-CNN, RetinaNet and SSD513 [3-5], YOLO-v3 achieved faster processing and more accurate localization on small targets on an image [2]. This property is promising for real-time processing in a CAD. We experimentally evaluate a perforation of YOLO-v3 for detection and localization of a perforation in ESD as preliminary study.

2. Method

This section briefly introduces the original YOLO-v3 method [2], which is utilized as a perforation detection and localization model. YOLO-v3 is a supervised-learning object localization method for detection and localization of objects in a 2D image. To train YOLO-v3, we need images with annotated bounding boxes, which comprise of a center point $(x, y)^\top$, width W, height H and a class label L of a target object. We set L to be 0 and 1 for non-perforation and perforation classes, respectively. For predicted i-th object in a test image, a trained YOLO-v3 can predict a center point $(x_i, y_i)^\top$, width W, height H_i and a class probability $\mathrm{P}(L_i)$ of the detected object, simultaneously. YOLO-v3 comprises of a deep convolutional neural network by using residual units [6]. YOLO-v3 predicts bounding boxes of a target, and selected the best bounding box by non-maximum suppression [7] algorithm. To predict the size and location of a bounding box fast and precisely, YOLO-v3 selects a initial bounding box that comprise of a center point $(x_1, y_1)^\top$, width W_1, height H_1, and predicts biases $b_x,\ b_y,\ b_w,\ b_h$ of the initial bounding box, e.g. center point $(x_i, y_i)^\top$ of the i-th predicted bounding box in an image can be calculated by $x_i = x_1 + b_x,\ y_i = y_1 + b_y$, width W_i, height H_i of this bounding box can be calculated by $W_i = W_1 + b_w,\ H_i = H_1 + b_h$. Referred to the original YOLO-v3 work, we constructed a YOLO-v3 model, the detail settings of our model are the same to the original YOLO-v3.

3. Dataset Generation

We collected ESD videos of 11 patients in the digestive center of Showa University Yokohama Northern Hospital with IRB approval. Expert endoscopists annotated perforation-existing scenes of these videos. Based on this annotation, we extracted perforated and non-perforated images from colonoscopic videos by 30 and 1 fps, respectively. We resized all extracted images to 224×224 pixel with Lanczos interpolation. We splitted all resized images of 11 patients into training, validation, and test dataset without the duplication of patients. We annotated the bounding boxes that contain perforations in extracted images in training, validation and test datasets manually. For test dataset, we only annotated 150 perforation-existing images due to the huge amount of number of test images. We convert each bounding box to a txt file, which comprises of a center point $(x, y)^\top$, width W and height H of a bounding box, and class label $L = 1$ of perforation. Table 1 shows the number of extracted images and patients of training, validation, and test data. Figure 1 shows examples of training, validation and test data.

Table 1 Number of images of training, validation and test data

Dataset	Number of non-perforated images	Number of perforated images	Number of perforated cases
Training	0	1,814	7
Validation	0	385	2
Test (detection)	16,564	6,034	2
Test (localization)	0	150	2

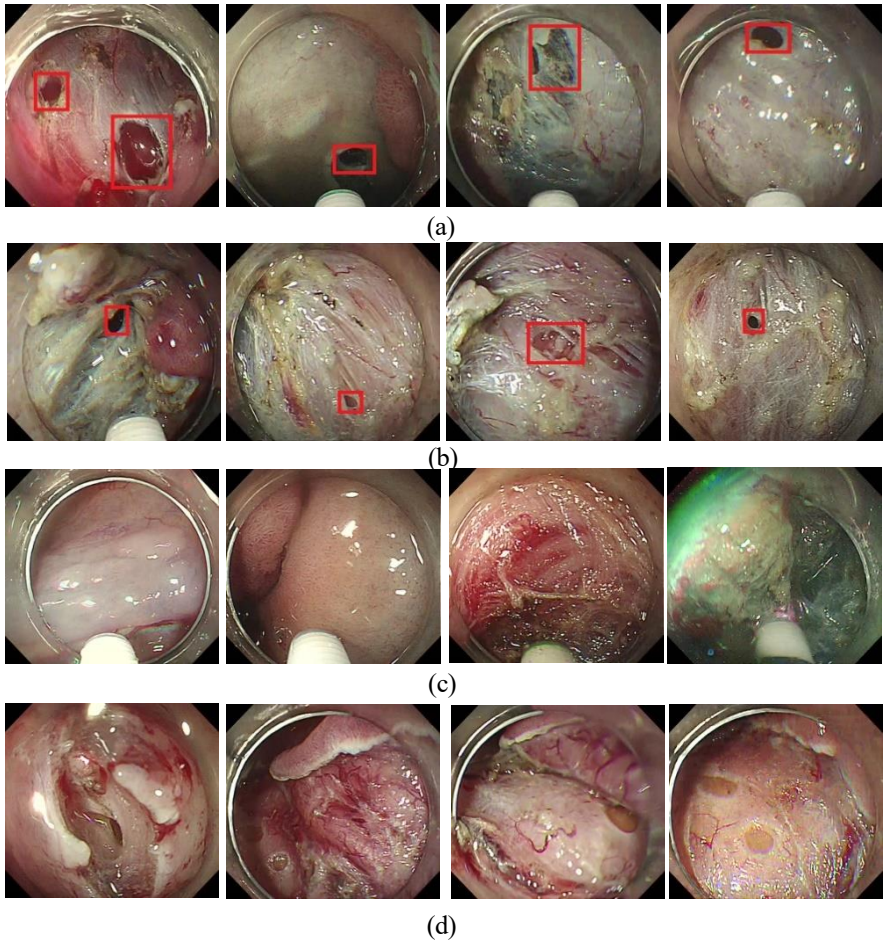


Fig. 1 Examples of training, validation and test data. (a) perforation-existing images in training dataset. (b) perforation-existing images in validation dataset. (c) no perforation images in test dataset. (d) perforation-existing images in test dataset. The red bounding boxes annotated perforations in images.

4. Experiments and Results

We evaluated the detection and localization accuracy of a trained YOLO-v3 model. We constructed a model with the Pytorch. We set minibatch size to be 64 to train the model for 300 epoch on NVIDIA Tesla V100 PCIe 32 GB with CUDA 10.0. We selected the weights that gave the best performance on images in validation dataset for evaluation. To evaluate the accuracy of trained YOLO-v3, we set a threshold t, and classify a test image as perforation-existing image when test class probability $\mathrm{P}(L) \geq t$. We defined proportion of incorrectly classified images in all no perforation images

as 1-specificity, proportion of correctly classified images among all perforation-existing images as sensitivity. We defined proportion of correctly classified images among all test images as accuracy. We set thresholds $t \in [0,1]$, and calculated sensitivity and 1-specificity values, respectively. Figure 2 (a) shows Receiver Operating Characteristic (ROC) curve drawed by these sensitivity and 1-specificity values. We used accuracy and Area Under the Curve (AUC) of ROC to evaluate detection accuracy of trained YOLO-v3. Our model achieved 0.835 accuracy and 0.942 AUC score.

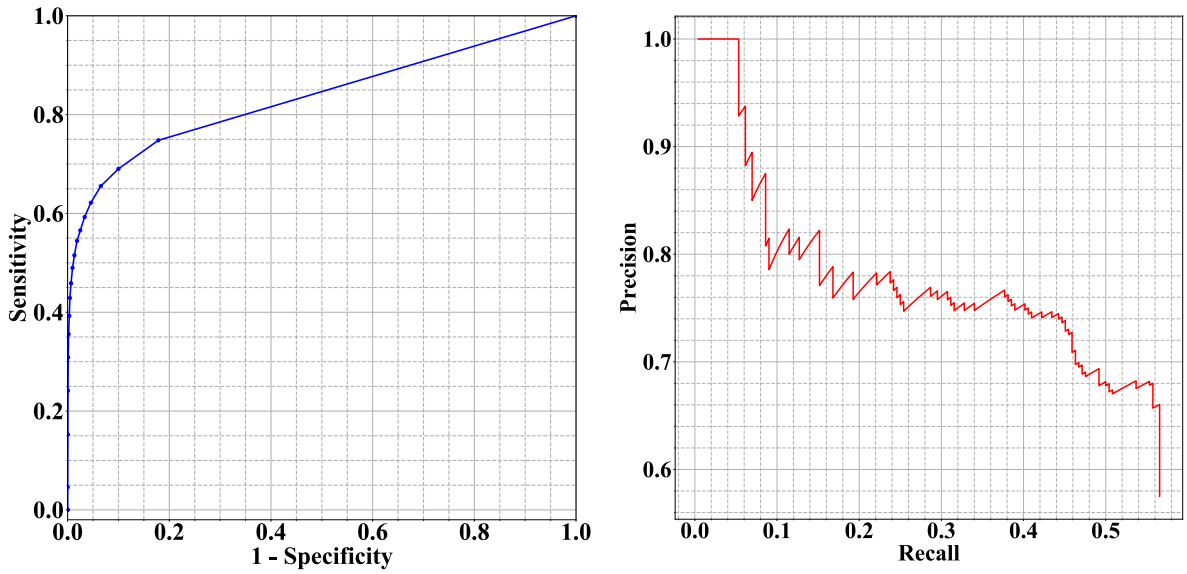
We used 150 annotated images in test dataset to evaluate the localization accuracy of trained YOLO-v3. To better evaluate localization accuracy, we defined Intersection over Union (IoU) between area A_1 and area A_2 by

$$IoU(A_1, A_2) = \frac{A_1 \cap A_2}{A_1 \cup A_2}. \quad (1)$$

We calculated IoU of each predicted bounding box and ground truth bounding box. We defined images that classified as perforation-existing images by trained YOLO-v3 as predicted positive images. We defined an image as a true positive image, when IoU between a predicted bounding box and a ground truth bounding box in this image is greater than the threshold. We defined proportion of number of true positive images with number of all predicted positive images as precision. We defined proportion of number of true positive images with number of all perforation-existing images as recall. We sorted 150 images by their classified probability from the largest to the smallest. We caculated 150 precision and recall value for these images, where the first precision and recall values only calculated by the first image, e.g. when the first image is a true positive image, precision and recall value are 1 and 0, respectively. The second precision and recall calculated by the first and the second image, then we gradually increase the number of calculated images, the last precision and recall value calculated by all 150 images. We defined the average value of these 150 precision values as average precison (AP). We defined the mean value of AP values of all classes as mean Average Precision (mAP). We used these precision and recall values to draw a mAP curve, then the mAP value can be represented by AUC of mAP curve. Figure 2 (b) shows mAP curve of localization results of 150 images predicted by trained YOLO-v3. Our model achieved 0.746 mAP in 150 annotated images when we set threshold to 0.6. Figure 3 compares a image predicted by trained YOLO-v3 and a ground truth image.

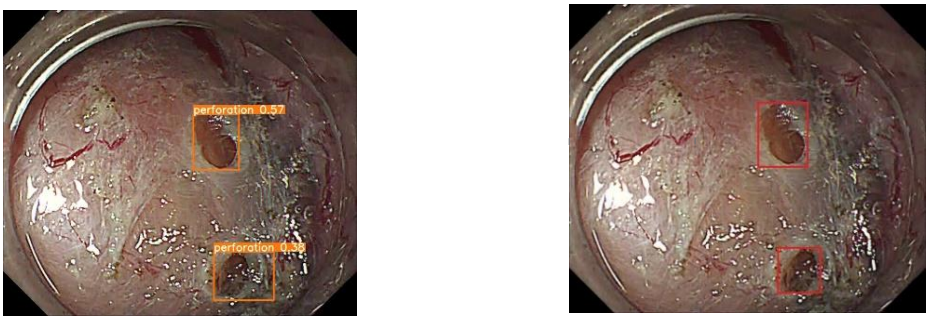
5. Discussion

Our constructed model achieved 0.835 accuracy and 0.746 mAP (mean average precision) in detecting and localizing perforations, respectively, in scenes of colonoscopic video. Figure 3 shows our constructed model localized perforations precisely in scenes of colonoscopic video. According to the ROC curve, we founded sensitivities are much higher than 1-specificities in the same threshold, which refers to that our constructed model performs high accuracy in detecting no perforation images than perforation-existing images. This result implies the insufficient size of training data. To increase sensitivity of detection, we have to collect perforation scenes more and more. Furthermore, as mentioned in Section 3, training and validation dataset only contain perforated images. To improve the detection specificity of our constructed model, we have to add non-perforated images in training and validation dataset.



(a) (b)

Fig. 2 Qualitative evaluations. (a) ROC curve in scene detection. (b) mAP curve in perforation localization.



(a) (b)

Fig. 3 Comparison of a predicted image and a ground truth image. (a) a predicted image. (b) a ground truth image corresponding to prediction image. In predicted image, we write predicted class of object and probability of this class on the top of the corresponding bounding box.

6. Conclusion

This work reports a preliminary study of perforation detection and localization for colonoscopic video to construct a CAD system for ESD. We extracted 24,947 images from 11 colonoscopic videos, and splitted them into training, validation and test dataset for experiment. We used YOLO-v3 to construct a model for detection and localization of a perforation. We evaluated this model by using AUC and mAP scores for detection and localization, respectively. The constructed model achieved 0.942 AUC score and 0.746 mAP score in detection and localization, respectively, for the limited-size samples. To acquire higher detection and localization accuracy, we are going to train the constructed model

with various loss function in the future.

Competing interests

None

Acknowledgement

Parts of this research was supported by MEXT·JSPS KAKENHI (26108006, 17H00867, 17K20099), the JSPS Bilateral International Collaboration Grants, the AMED (19hs0110006h0003) and the Hori Sciences & Arts Foundation.

References

［1］Saito Y, Uraoka T, Yamaguchi Y, et al.: A prospective, multicenter study of 1111 colorectal endoscopic submucosal dissections (with video), Gastrointestinal Endoscopy 2010, **72**(6):1217-1225, 2010

［2］Redmon J, Farhadi A: Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767, 2018

［3］Wang X, Shrivastava A, Gupta A: A-fast-rcnn: Hard positive generation via adversary for object detection. Computer Vision and Pattern Recognition, 2606-2615, 2017

［4］Lin T Y, Goyal P, Girshick R, et al.: Focal loss for dense object detection. International Conference on Computer Vision, 2980-2988, 2017

［5］Fu C Y, Liu W, Ranga A, et al.: Dssd: Deconvolutional single shot detector. arXiv preprint arXiv:1701, **06659**, 2017

［6］He K, Zhang X, Ren S, et al.: Deep residual learning for image recognition, Computer Vision and Pattern Recognition, 770-778, 2016

［7］Neubeck A, Van Gool L: Efficient non-maximum suppression, International Conference on Pattern Recognition, **3**, 850-855, 2006

機械学習によるコルポスコープ画像における白色上皮領域検出の基礎的検討

篠原 寿広[*1]　天野 裕太[*1]　村上 幸祐[*2]　中迫 昇[*1]　松村 謙臣[*2]

要旨

子宮頸癌の診断において，コルポスコープとよばれる拡大鏡を用いて，目視による検査が行われている．本検査はコルポスコピーとよばれ，子宮頸癌の確定診断を行う組織診のための組織を採取する部位を決定する非常に重要な検査である．しかしながら，コルポスコピーの精度は，検査者の知識と経験に大きく依存し，とくに経験の浅い医師には正確な診断が難しい．そこで本研究では，機械学習を用いて，コルポスコピー支援を行い，安定して精度の高い診断の実現を目指す．本稿では，初期検討として，コルポスコープ画像から，子宮頸部の代表的な病変の一つである白色上皮の判別を行った．具体的には，コルポスコープ画像を小領域に分割し，CNN (Convolutional Neural Network)を用いて白色上皮かどうかを判別した．実際のコルポスコープ画像 10 症例を使用し，交差検証により判別精度を検討した．

キーワード：コルポスコピー，子宮頸癌，白色上皮，診断支援，機械学習

1. はじめに

子宮頸癌はヒトパピローマウイルスの持続的な感染が原因となって発症する癌で，全世界では，年間約 570,000 人が罹患し，約 311,000 人が死亡しており，女性の癌による死亡原因の第 4 位となっている［1］．日本においても年間約 10,000 人が子宮頸癌に罹患し，約 3,000 人が死亡している［2, 3］.

子宮頸癌の診断において，コルポスコープとよばれる拡大鏡を用いて，酢酸で加工した子宮頸部を目視により観察する検査が行われている．本検査はコルポスコピーとよばれ，子宮頸癌の診断において重要な役割を担っている．しかしながら，コルポスコピーは検査者の知識と経験に強く依存し［4］，経験の浅い医師には正確な診断が難しい．また，性能の良いデジタルカメラ付きコルポスコープは非常に高価なため，いまだ直接コルポスコープを覗いて診断する古い装置を用いている施設が多いのが現状である．一方，現在，スマートフォンやタブレット端末には非常に高解像度のカメラが搭載されており，コルポスコピーの使用にも耐えうる解像度を有している［5］. そこで本研究では，スマートフォンやタブレット端末のような安価な装置を用いて，機械学習によるコルポスコピー支援システムを開発し，検査者の知識や経験に頼らない安定して精度の高い子宮頸癌診断の実現を目指す.

本稿では，その初期検討として，実際のコルポスコープ画像において，子宮頸部の代表的な病変の一つである白色上皮の判別を機械学習により行い，その判別精度を検討した.

*1 近畿大学生物理工学部生命情報工学科
〔〒649-6493 和歌山県紀の川市西三谷 930〕
e-mail: sinohara@waka.kindai.ac.jp
*2 近畿大学医学部産科婦人科学教室

2. 子宮頸癌診断

子宮頸癌診断は，まず問診，視診，細胞診を

行って異常所見の有無を調べる．そこで何らかの異常が見られたなら，コルポスコピーを行い，子宮頸部を詳しく観察する．子宮頸癌の前段階の病変として知られている異形成を疑う所見が観察されたなら，その部位を含め，子宮頸部の組織を複数箇所採取し，採取した組織を顕微鏡で観察する組織診により，子宮頸癌の確定診断を行う．コルポスコピーは組織を採取する侵襲的な検査であることから，患者の負担を減らすために，正確な診断が必要である．

異形成には白色調や周辺との境界の明瞭さなど，多様な所見に合わせて名前が異なり，また同じ異形成においても，その程度により区別されるため，コルポスコピーにおいて正確に診断するには豊富な知識と経験が必要である．異形成は軽度の症状であればそのほとんどが消失するが，高度な異形成はその 20～30%が癌へと進行する．

3. 機械学習による白色上皮の判別

本研究では Convolution Neural Network (CNN) とよばれる機械学習を用いて，子宮頸部の代表的な病変の一つである白色上皮を判別する．CNN を用いるにあたり，はじめにコルポスコープ画像を小領域に分割し，医師による診断データを元に分割した画像の各画素にラベル付けを行ない，それらを学習用画像とする．判別対象のコルポスコープ画像 (以下，テスト用画像とよぶ)も同様に分割し，それらを CNN により判別できるかを検討した．本研究で用いるラベルは，白色上皮とそれ以外の 2 種類である．コルポスコープ画像を分割する際，組織の境界に位置する分割画像は学習用画像として適さないため，白色上皮もしくはそれ以外の領域が分割画像全体の 70%以上を占めている分割画像のみ学習用画像として採用した．

本研究で用いる CNN の構造は VGG (Visual Geometry Group)16［6］に基づく．VGG16 は畳み込み層が 13 層，全結合層が 2 層および最終出力層の合計 16 層の深いネットワークであり，その判別性能と汎用性が高いことが知られている．

4. 白色上皮判別実験

近畿大学病院においてコルポスコピーを実施した 20 代～40 代の患者のコルポスコープ画像 10 症例を対象に実験を行った．本研究は近畿大学医学部倫理委員会および同大学生物理工学部生命倫理委員会によって承認され，コルポスコープ画像の使用はすべての患者より書面による同意を得ている．また，使用したコルポスコープ画像はデジタルカメラ (E-PL6, OLYMPUS)により撮影した 4,806×3,456 pixels の RGB 画像である．

実験に使用した計算機環境を表 1 に示す．また，画像枚数や CNN の学習パラメータなどの実験条件を表 2 に記す．

表 1 計算機環境.

CPU	Intel Core i7-4770K
メインメモリ	16GB
GPU	NVIDIA Quadro K5000
OS	Ubuntu 16.04 LTS
ソフトウェア	Chainer

表 2 実験条件.

分割画像サイズ (pixels)	128×128
白色上皮学習用画像数 (枚)	7,475
白色上皮以外の学習用画像数 (枚)	32,188
白色上皮テスト用画像数 (枚)	643
白色上皮以外のテスト用画像数 (枚)	9,077
学習回数 (回)	500
ミニバッチサイズ	5
最適化手法	慣性項付き確率的勾配降下法
学習率	0.001
慣性項	0.9

コルポスコープ画像の分割サイズは実験的に 128×128 pixels とし，分割する際，学習用画像は，分割領域を縦横それぞれ 32 pixel ずつずらして分割した．すなわち，隣り合う分割画像

の 3/4 は重なっている. 一方, テスト用画像は, 重なりなく画像を分割した. なお, 学習用画像は回転などによる加増は行っていない. 10 症例のうち, 学習用画像を作成するために 9 症例を使用し, テスト用画像を作成するために残り 1 症例を使用して, 交差検証により判別性能を検討した.

実験結果の一例を図 1 に示す. 黄色が真陽性, 青色が偽陰性, 緑色が偽陽性を表している.

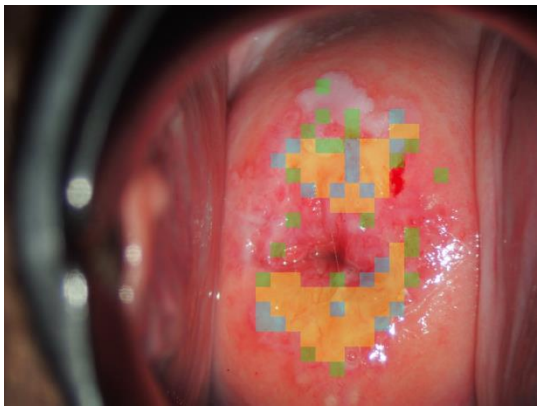


図 1 白色上皮判別結果の一例
(黄色：真陽性，青色：偽陰性，緑色：偽陽性).

表 3 に平均の判別性能を示す. 一連の実験には約 48 時間要した.

表 3 白色上皮の平均判別性能.

感度 (%)	67.2±27.8
適合率 (%)	54.5±29.7
特異度 (%)	96.1±2.6

特異度は良い結果得られたが，比較的判別が容易な白色上皮以外の領域が，白色上皮の領域に比べ，かなり大きいことによる. 感度，適合率は，良い結果が得られた症例もあれば，その逆もあり，全体的にはあまり良い結果が得られず，標準偏差が大きくなった. 子宮頸部の色は個人差があることに加え，撮影条件にも左右され，コルポスコープ画像にバリエーションがあるため，このような結果となった. 特に，白色上皮を正しく判別できなかった偽陰性の部分は正常上皮との境界や光の反射で色飛びした部分である. また，誤って白色上皮と判別した偽陽性の部分は，子宮頸部の色や撮影条件により白っぽく撮影され，白色上皮と似たような色となってしまったために誤判別された.

これらの問題を解決するため，学習用画像を増やすことがまず考えられるが，子宮頸部の個人差，撮影条件によるコルポスコープ画像のバリエーションを補正または分類する前処理や，誤って判別された部分を修正する後処理を追加することにより判別性能の向上を図ることも期待できる.

5. まとめ

本稿では，機械学習を用いたコルポスコピー支援による，安定して精度の高い診断の実現を目指し，その初期検討として，コルポスコープ画像から，子宮頸部の代表的な病変の一つである白色上皮の判別を行った. 具体的には，コルポスコープ画像を小領域に分割し，CNN (Convolutional Neural Network) を用いて白色上皮かどうかを判別した. 実際のコルポスコープ画像 10 症例を用いて白色上皮判別実験を行った結果，平均判別性能は，感度 67.2%，適合率 54.5%，特異度 96.1%となった.

今後は判別性能向上のため，学習用データの加増や，前処理および後処理を考えるとともに，VGG 以外の深層学習の適用を試みる. また，白色上皮以外の異形成の判別や，異形成の軽度，高度などの程度の判別についても検討を加える予定である.

謝辞

コルポスコープ画像をご提供してくださった近畿大学医学部産科婦人科学教室の皆様に深謝いたします.

利益相反の有無

なし.

文　献

［1］ Bray F, Ferlay J, Soerjomataram I, et al.: Global cancer statistics 2018: GLOBOCAN

estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 68(6): 394-424, 2018.

［2］ 厚生労働省: 平成 29 年全国がん登録罹患数・率報告, 2017.

［3］ 国立がん研究センターがん情報サービス: がん登録・統計(人口動態統計), 2018.

［4］ Xue P, Ng M.T.A and Qiao Y: The challenges of colposcopy for cervical cancer screening in LMICs and solutions by artificial intelligence. BMC Med 18, 169 2020.

［5］ Xue Z, Novetsky A, Einstein M, et al.: A demonstration of automated visual evaluation (AVE) of cervical images taken with a smartphone camera. International Journal of Cancer, 2020.

［6］ Karen S and Andrew Z: Very Deep Convolutional Networks for Large-Scale Image Recognition. Proceedings of International Conference on Learning Representations, 2015.

Basic consideration on region detection of acetowhite epithelium in colposcopic image by using machine learning

Toshihiro SHINOHARA[*1], Yuta AMANO[*1], Kosuke MURAKAMI[*1], Noboru NAKASAKO[*1], Noriomi MATSUMURA[*2]

*1 Department of Computational Systems Biology, Faculty of Biology-Oriented Science & Technology, Kindai University

*2 Department of Obstetrics & Gynecology, Faculty of Medicine Kindai University

Colposcopy is an important inspection for determining the place to obtain cells of the cervix for tissue diagnosis of cervical cancer. This is carried out by visual inspection with a magnifying lens called colposcope. This inspection strongly depends on the knowledge and experience of inspectors, so that it is often difficult for inexperienced gynecologists to accurately diagnose lesions on the cervix. Thus, this study aims at the realization of stable high accurate diagnosis of the colposcopy by assisting the inspector with a machine learning. In this paper, regions of the acetowhite epithelium, one of the representative cervical lesions, in colposcopic image was detected by using the machine learning as a basic consideration of this aim. In the proposed method, the colposcopic images divided into small patches are classified into the acetowhite epithelium and the other by using a convolutional neural network, one of the machine learnings. The classification accuracy of the actual colposcopic images of 10 cases was discussed.

Key words: Colposcopy, Cervical cancer, Acetowhite epithelium, CAD, Machine learning

One-class SVM による画素悪性度の集中性を用いた FDG-PET/CT 上の病変自動検出

田中　敦子[*1]　根本　充貴[*2]　甲斐田　勇人[*3*4]　木村　裕一[*1]

永岡　隆[*1]　牛房　和之[*1]　山田　誉大[*4]　花岡　宏平[*4]

北島　一宏[*5]　槌谷　達也[*6]　石井　一成[*3*4]

要旨

【目的】FDG-PET/CT 上の頚胸部領域における原発性・転移性病変の自動検出に関する初期検討として，one-class SVM による病変画素強調処理に基づく病変候補点の自動検出法を提案し，検証する．【手法】提案法では，まず CT の閾値処理及びモロフォロジ処理によって左右肺を抽出し，抽出した肺の間を縦隔部として自動抽出する．頚部は，自動同定した下顎骨座標と首領域，左右肺領域を基準に抽出する．その後，正常な FDG-PET/CT で学習済の one-class SVM を用いて，対象領域内画素の異常度を計測する．得られた異常度画像の極大点を検出し，病変の候補点とする．提案法は，82 病変を含む FDG-PET/CT 有病 47 例で検証する．【結果】頚胸部の病変検出感度が 95.1%(79/82)となった．【結論】検出感度結果より提案法が有用であることが示唆された．今後の課題は，誤検出候補点の削減処理の追加と，異常検知特徴量の追加による性能向上である．

キーワード：コンピュータ診断支援，FDG-PET/CT，one-class SVM，異常検知

1. はじめに

近年の画像撮像機器の高性能化に伴い，患者一人当たりの画像枚数は膨大となっている．そのため，大量の画像を読影する読影医にかかる負担は莫大となっている．これより，診断精度の低下やばらつきが懸念されている．そこで読影医の負担軽減のための画像診断支援（CAD）システムの開発が行われている[1]．本研究では，FDG-PET/CT（以下 PET/CT）の CAD システムに着目する．PET/CT は，全身の病変に対しての読影が必要となる．臨床的には有用であるが，全身が対象になるために読影にかかる負担が大きい．

PET/CT の CAD システムに関する先行研究[2,3]では，その大多数が単一の病変や単一の臓器に焦点が当てられてきている．寺本らの研究[2]では，対象を肺結節とし，畳み込みニューラルネットワーク（CNN）由来の特徴量などを用いた病変多段階識別によって病変を検出している．越野らの研究[3]では，2 種類の CNN を使用して骨転移候補検出及び，

*1 近畿大学大学院生物理工学研究科
〔〒649-6493 紀の川市西三谷 930〕
e-mail: 1933730006u@waka.kindai.ac.jp
*2 近畿大学生物理工学部
*3 近畿大学医学部放射線医学教室放射線診断学部門
*4 近畿大学高度先端総合医療センター PET 分子イメージング部
*5 兵庫医科大学放射線医学講座
*6 兵庫医科大学病院放射線技術部

骨転移候補を識別している．一方で，複数の病変を対象とした研究もある．箕浦らの研究[4]では，CT からの対象領域抽出結果と PET からの薬剤高集積部分抽出結果の重ね合わせによって乳腺腫瘍及び腋窩リンパ転移病変を検出している．これらの先行研究では，有病症例を機械学習した識別処理を用いることで病変を検出する手法となっている．

我々はこれまで，有病症例データを学習に用いない異常検知処理に基づく PET/CT 上の頚胸部領域の肺がんや食道がんといった原発性病変と，リンパ節転移性病変の自動検出 CAD システムについて検討を行ってきた．提案法は収集が比較的容易である正常データのみを学習に用いた手法であるため，システム開発が容易などの利点がある．最近では one-class SVM（以下 OCSVM）を用いた画素異常度計測による頚胸部領域の病変強調の有用性を実験的に確認した[5]．そこで本研究では，OCSVM により得た画素異常度画像上の極大点を病変候補点として検出する処理の有効性を実験的に検証する．過去検討[5]より，画素悪性度は病変尤度として利用でき，病変領域では局所的に高い画素悪性度を示すことが分かっている．よって，画素悪性度の極大点は病変候補点として扱うことができると考えた．本研究では，病変候補点の自動検出法を提案し，臨床 PET/CT を用いて手法の有用性を検証する．

2. 手法

2.1 実験試料

本実験では，近畿大学病院（以下近大病院）にて撮像された全身 FDG-PET/CT データ 47 症例及び，兵庫医科大学病院（以下兵庫医大病院）にて撮像された同データ 40 症例を使用する．うち，正常症例は近大病院で 20 症例，兵庫医大病院で 20 症例であり，有病症例は近大病院で 27 症例，兵庫医大病院で 20 症例（肺領域 29 病変，頚部・縦隔領域 53 病変の計 82 病変を含む）であった．撮像機器は近大病院が GE Discovery PET/CT 710 scanner であり，兵庫医大病院が GE Discovery IQ である．施設ごとの PET/CT の解像度については以下，表 1 に示す．

表 1 施設ごとの PET/CT 解像度．

	近大病院	兵庫医大病院
CT [mm]	0.977×0.977×3.27 1.37×1.37×3.27	0.977×0.977× 3.26 1.17×1.17×3.26
PET [mm]	2.60×2.60×3.27 3.65×3.65×3.27	2.60×2.60×3.26 3.13×3.13×3.26

2.2 提案手法

提案する手法を図 1 のフローチャートに示す．本提案法では，1）対象領域の自動抽出，2）OCSVM による病変尤度計測，3）病変尤度画像を用いた病変候補点自動抽出の 3 段階の処理を行う．

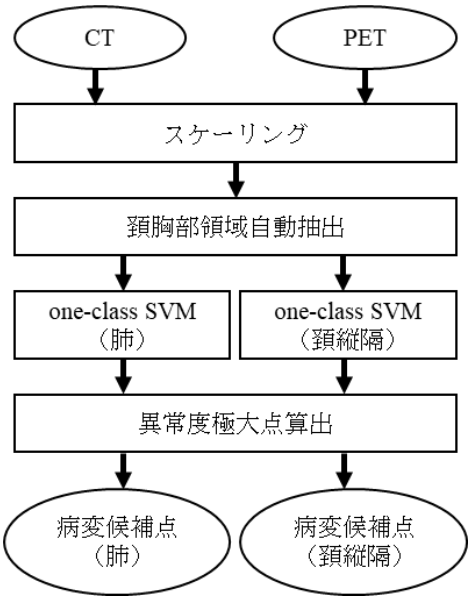


図 1 頚胸部病変の候補点検出処理．

1） 対象領域自動抽出

使用する CT 及び PET の解像度を合わせるため，まずスケーリング処理を行う．近大症例，兵庫医大症例ともに CT と PET を 2.4mm 解像度に等方化する．補間法にはバイキュービック補間を使用する．この補間法を使用する理由としては，線形補間よりもより明瞭に補間を行うことができるから

である．次に CT 値の閾値処理，モロフォロジ処理等により左右肺の自動抽出を行う．抽出した左右肺間の領域を縦隔領域として，それに連結する頚部領域とを合わせて頚部・縦隔領域として抽出する．頚部・縦隔領域の上端は歯科治療等によるメタルアーチファクトの影響回避のため，下顎骨下端とする．また頚部・縦隔領域の下端は，左右肺の下端面を補間して推定した横隔膜面までとする．横隔膜面推定にはバイハーモニックカーネルを使用した放射基底関数に基づく局面推定[6]を使用する．

2） OCSVM による病変尤度計測

自動抽出した左右肺領域，頚部・縦隔領域の 2 領域のそれぞれで OCSVM による任意 voxel ごとの異常度の計測を行う．

OCSVM の学習には正常 PET/CT 計 40 症例を使用する．OCSVM は肺領域と頚部・縦隔領域で別のものを学習，利用する．OCSVM が使用する特徴量は，肺領域では CT 値，SUV，胸壁からの距離の 3 特徴量，頚部・縦隔領域では，CT 値と SUV の 2 特徴量とする．特徴量を変える理由は，臓器ごとに病変の画像パターンや考慮すべき周辺臓器が異なるためである．

OCSVM に使用したパラメータは ν と γ の 2 種類であり，それぞれパラメータチューニングによって最適パラメータを決定している． ν は以下のソフトマージン最大化の式(1)で示される[7].

$$\min_{w\in F,\boldsymbol{\xi}\in\mathbb{R}^{\ell},\rho\in\mathbb{R}} \frac{1}{2}\|w\|^2+\frac{1}{\nu\ell}\sum_i \xi_i-\rho \qquad (1)$$

また γ は OCSVM で使用する RBF カーネルの式に示される．RBF カーネルの式は以下式(2)である[8].

$$K(x_i,x_j)=\exp(-\gamma\|x_i-x_j\|)^2 \qquad (2)$$

使用したパラメータとその値を表 2 に示す.

表 2 使用パラメータとパラメータ値.

パラメータ	肺	頚部・縦隔
ν	0.125	0.25
γ	0.25	0.0625

3） 病変候補点自動検出

異常度画像を σ =1.0, 2.0 の 2 種類のカーネルでガウシアン平滑化を行う．2 種類の平滑化異常度画像から極大点を検出し，両結果を統合して病変候補点とする．統合方法としては，まず σ =2.0 の極大点を中心とした周辺 26 近傍に σ =1.0 の極大点が存在するか探索する．存在した場合には，極大点部分の異常画素値とその周辺 26 近傍の異常画素平均との比をとる．算出した比の大きい方を残し，比の小さい方は削除するという方法で両 σ の極大点結果を統合する．

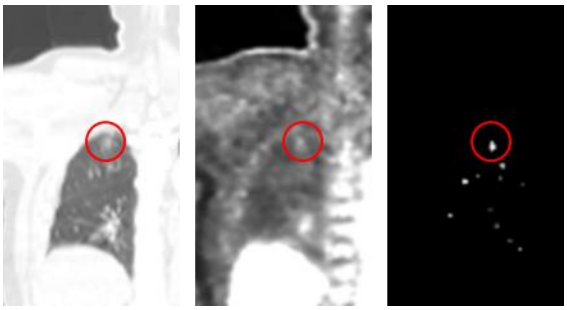
3. 結果

候補点検出結果と平均偽陽性（FP）数を以下表 3 に示す．ただし，頚胸部とは，肺領域と頚部・縦隔領域を合わせた領域とする．

表 3 領域ごとの候補点検出結果.

領域	感度 [%]	平均 FP 数[/症例]
肺	100	280.3
頚部・縦隔	92.5	412.0
頚胸部	95.1	324.2

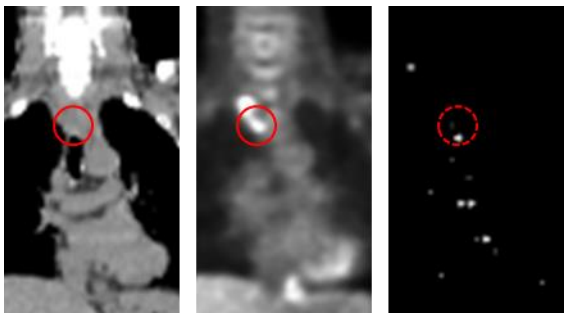
候補点検出成功症例を図 2 に，候補点検出失敗症例を図 3 に示す. (a)は CT, (b)は PET, (c)は候補点分布画像を示す．ただし候補点分布画像は見やすさを考慮し，6 近傍膨張処理を施した．



(a)CT (b)PET (c)候補点画像

図 2 候補点検出成功症例． 病変は右肺のすりガラス状結節であり，画像中の〇でそれぞれ示

す.



(a)CT (b)PET (c)候補点画像

図 3 候補点検出失敗症例．病変は気管分岐部リンパ節であり，画像中の実線の○でそれぞれ示す．点線の○は候補点が検出できていない病変部分を示す.

4. 考察

図 2 より病変部分の候補点が検出できていることから本提案が有用であることが示唆された.

頚部・縦隔領域にて候補点が検出できなった病変は，上部上縦隔リンパ節転移病変，食道傍リンパ節転移病変 2 病変，気管分岐部リンパ節転移病変の 4 病変であった．図 3 のような候補点検出失敗症例の要因として，頚部・縦隔領域の抽出不足，病変部分の SUV が淡いために OCSVM による異常度計測で病変部分が異常とみなされない，ということが考えられる．病変サイズが小さい，または病変部分の濃淡が淡い場合などに関してはパッチベースで画素の統計的特徴量の抽出を行うことで異常判定の向上が期待できるのではないかと考える.

表 3 より，全領域において高感度検出の結果が確認できた．一方で，頚部・縦隔領域における平均 FP 数は，肺領域，頚胸部領域と比較して高い数値が得られた．この原因としては，縦隔領域に存在する心臓や，頚部領域の甲状腺付近などが挙げられる．特に心臓では高 SUV となることが多いためである. 肺領域での FP の多くは肺血管に多く，これは肺内の空気領域に比べ，肺血管で CT 値が高くなるために発生すると考えられる．これらの候補点 FP 削減のためには，CT のテクスチャ情報や，FDG の生理的高集積部分の情報を加味した特徴量を加えることが必要だと考える. 特に FP が多い心臓に関しては，心筋からの距離特徴量を加える予定である.

5. 結論

One-class SVM により得た異常度画像から病変候補点の自動検出法を提案し，臨床 PET/CT を用いて手法の有用性を検証した．頚胸部領域にて感度 95.1%（79/82 病変），平均 FP 数 324.3 が得られた. 高い検出感度が得られたため，提案法が有用であることが示唆された．今後の課題は，候補点 FP の削減処理の追加と，異常検知特徴量の追加による性能向上である.

謝辞

本研究の遂行において，医用画像ビューワ VOLUME-ONE の使用，及び画像処理ライブラリの使用をご快諾いただいた，広島市立大学の増谷佳孝教授に深く感謝申し上げます．本研究の一部は，JSPS 科研費 17K17680 および 17H05284，近畿大学学内研究助成金 SR06 により行われた.

利益相反の有無

利益相反なし.

文 献

［1］ Ronald A. C: Computer aided detection (CAD): an overview. Cancer Imaging, **5(1)**: 17-19, 2005

［2］ Teramoto A, Fujita H, Yamamuro O, et al: Automated detection of pulmonary nodules in PET/CT images: Ensemble false-positive reduction using a convolutional neural network technique. Medical Physics **43(6)**, 2821-2827, 2016

［3］ Koshino S, N. V. Apeldoorn, A. Choppin, et al: AI Pipeline System for Detection

of Bone Metastases on PET-CT. 105th RSNA2019, McCormick Please, United States of America, 2019

［4］ 箕浦葉月，寺本篤司，高橋克彰 他: PET/CT 画像を用いた乳房領域の抽出と乳腺腫瘍および腋窩転移の自動検出の試み. Medical Imaging Technology **35(3)**: 158-166, 2017

［5］ Tanaka A, Nemoto M, Kaida H, et al: Automatic detection of cervical and thoracic lesions on FDGPET/CT by organ specific one-class SVMs. CARS 2020—Computer Assisted Radiology and Surgery Proceedings of the 34th International Congress and Exhibition: S208-S209, 2020

［6］ 増谷佳孝，木村文彦，佐久間一郎: 境界形状の特徴抽出および動径基底関数による形状再構成に基づく X 線 CT 像における肝臓領域の自動抽出と形状モデリング．コンピュータ支援画像診断学会論文誌 **8(1-2)**: 10-17 2004.

［7］ Bernhard S, Robert W, Alex S et al: Support Vector Method for Novelty Detection. Proceedings of the 12th International Conference on Neural Information Processing Systems, 582-588, 1999

［8］ Han S, Cao Q, Han M: Parameter selection in SVM with RBF kernel function. World Automation Congress 2012, 24-28, 2012

Automated detection of lesion on FDG-PET/CT with the centrality of voxels malignancy by one-class SVM

Atsuko TANAKA[*1], Mitsutaka NEMOTO[*2], Hayato KAIDA[*3*4], Yuichi KIMURA[*1], Takashi NAGAOKA[*1], Kazuyuki USHIFUSA[*1], Takahiro YAMADA[*4], Kohei HANAOKA[*4], Kazuhiro KITAJIMA[*5], Tatsuya TSUCHITANI[*6], Kazunari ISHII[*3*4]

*1 Graduate School of Biology-Oriented Science and Technology, Kindai University
*2 Faculty of Biology-Oriented Science and Technology, Kindai University
*3 Department of Radiology, Kindai University Faculty of Medicine
*4 Division of Positron Emission Tomography Institute of Advanced Clinical Medicine, Kindai University
*5 Department of Radiology, Hyogo College of Medicine
*6 Department of Radiology, Hyogo College of Medicine Hospital

【Purpose】As a pilot study to detect primary and metastatic lesions in the neck and chest region on FDG-PET/CT, we propose and verify a lesion candidate detection from lesion enhancement volume estimated by one-class SVM.
【Method】At first, the lung regions are extracted by the thresholding and the morphology process on CT. Secondly, a soft tissue region sandwiched between the extracted lungs is defined as the mediastinum. The cervical region is identified based on the mandible position, neck region, and bilateral lung positions. At each voxel within the lung regions and the cervical mediastinum region, a voxel abnormality is estimated using the region-specific one-class SVM trained by normal FDG-PET/CT previously. The local maximum points on the abnormality volume are detected as the lesion candidate points. The proposed method is validated in 47 cases of FDG-PET/CT with 82 lesions.
【Results】Sensitivity for neck and chest lesion detection was 95.1% (79/82). 【Conclusion】The results of detection sensitivity suggest that the proposed method is useful. The future task is to improve the performance by adding processing to reduce false detection candidate points and adding anomaly detection features.

Key words: Computer-aided-diagnosis, FDG-PET/CT, one-class SVM, anomaly detection

Application of unsupervised learning model for anomaly detection to PET images

Takehiro MAEDA[*1], Takeshi HARA[*2], Tetsuro KATAFUCHI[*3], Xiangrong ZHOU[*2], Masaki MATSUSAKO[*4], Satoshi ITO[*5], Masaya KATO[*5], Akira YAMADA[*6], Hiroshi FUJITA[*2]

Abstract

Numerous studies have been reported on the detection of tumors in medical images such as Positron emission tomography (PET) and computed tomography (CT) using deep learning. Many of them used supervised deep learning on large data sets with label images for correct detection. However, the data set for training may be insufficient to train deep learning models due to the huge cost of creating label images. This study aims are utilization of PET images without label and development of a method for the detection of tumors using unsupervised learning. We propose a method of tumors detection based on anomaly detection with a deep learning model trained normal cases. A generation model based on GANs and AutoEncoder is used for detection. The model trained only normal PET images generates nearly normal images which are similar to training data. Therefore, the difference between images containing untrained anomalous accumulations and generated images is large. The abnormal regions are extracted by a dynamic thresholding on the differential images between inputs and outputs of the model for each cross-sections (axial, coronal and sagittal). Only areas where the results of all cross-sections match are candidates of tumors. Lastly, false positives are eliminated by using gradient of voxel values around extracted area. The detection performance was evaluated using 93 PET/CT images included 38 normal cases for a training dataset and 55 anomaly cases for a test dataset. The sensitivity was 76.5% with 95.4 FPs per case. The area of FPs was 938.3 voxels per case. These results indicate that unsupervised learning model trained normal cases without label contributed to automatic detection of tumors.

Keywords：PET, CAD, GANs, AutoEncoder

---

*1 Department of Intelligence Science and Engineering, Graduate School of Natural Science and Technology, Gifu University

〔1-1 Yanagido, Gifu, Gifu 501-1193, Japan〕

e-mail: tmaeda@fjt.info.gifu-u.ac.jp

*2 Department of Electrical, Electronic and Computer Engineering, Faculty of Engineering, Gifu University

*3 Faculty of Health Science, Gifu University of Medical Science

*4 Department of Radiology, St. Luke's International Hospital

*5 Department of Radiology, Daiyukai General Hospital

*6 Department of Radiology, Shinshu University School of Medicine

1. Purpose

The disease rate of cancer in Japan is a tendency to increase[1]. Early detection of cancer and medical treatment is important to reduce the number of deaths due to cancer. FDG-PET is one of the most effective Inspections of cancer screening to identify the location of cancer lesions and predict treatment effects. The diagnosis of PET images, which are volume data consisting of many tomographic images, is extremely sensitive. With the increase in the number of examinations, the burden on diagnostic radiologists has increased[2]. Various studies have been reported owing to the demand for computer aided diagnostic systems[3]. Recent studies on cancer detection from medical images, such as positron emission tomography (PET) and computed tomography (CT), often use deep learning methods. Many of these studies use supervised learning on large datasets containing label images for accurate detection. However, the training dataset may not be sufficient to train a deep learning model due to the huge cost of creating label images. The purpose of this study is a development of cancer detection method with unsupervised learning using label-less PET images. We propose a method for tumor detection based on anomaly detection utilizing normal cases.

2. Method

1) Unsupervised learning model

The proposed method generates a normal image that is similar to the input image and detect cancer from the difference image. OCGAN is a proposed Generative Adversarial Networks (GANs) model for anomaly detection of images other than the class used for training[4]. Anomaly images in this study are slice images containing anomalous accumulations in FDG-PET. We used an OCGAN-based deep learning model combining AutoEncoder and GANs cancer detection. The model trained only normal PET images has weights to generate normal-like images. Therefore, normal-like images are output even if the input contains lesions because the model don't have weights to generate anomaly images. we used 3 cross sections (Axial, Coronal and Sagittal) slice images resized to 128×128[pixel] as inputs.

2) Anomaly detection

When slice image which contain anomaly area is input to the model, the output image is a normal-like slice image converted from the anomalous accumulations to normal. In addition, the area of anomalous accumulation is a small compared to other normal areas in a whole body. A region of the anomalous accumulations locally has a larger difference than the surrounding region in a difference image between the input and output images. Therefore, detection of locally high pixel value in the difference images leads to extract anomaly regions.

$$T(x,y) = \frac{1}{n}\sum_{i=0}^{w}\sum_{j=0}^{h}\left|x_{i,j} - y_{i,j}\right| + \alpha\sqrt{\frac{1}{n}\sum_{i=0}^{w}\sum_{j=0}^{h}\left(x_{i,j} - y_{i,j}\right)^2} \quad (1)$$

Where x and y are respectively the input image and the output image of the model. w and h are width and height of input images. α is a constant for adjusting the threshold value. The difference images are binarized with the value calculated by Eq. (1). The same process is performed for all slices of the 3 cross sections. It is necessary to include information on multiple cross sections for accurate detection. Finally segmentation results are only the areas where the results of 3 cross sections match. Fig.1 shows flow of segmentation to get the results .

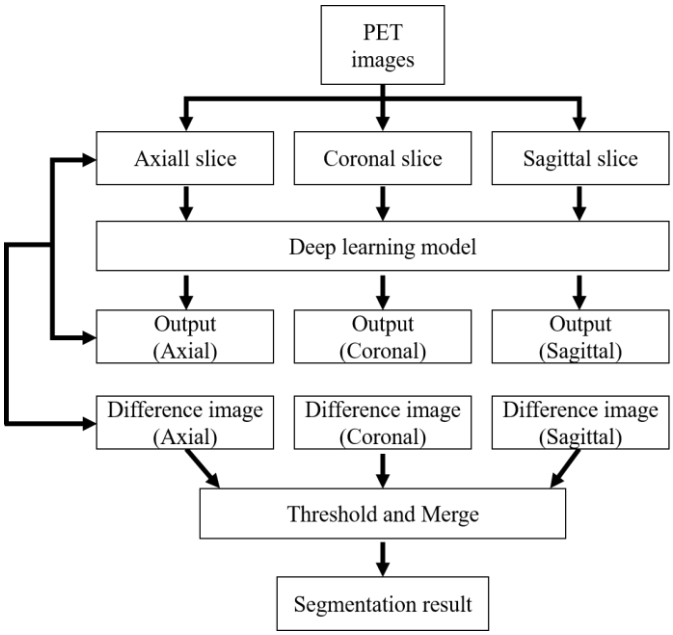


Fig. 1 Flowchart of segmentation

3. Results

The detection performance was evaluated using 93 PET images included 38 normal cases for a training dataset and 55 anomaly cases for a test dataset. The size of PET images is $144 \times 144 \times 213 - 276$[voxel]. Their pixel size and slice thickness are 4.00[mm].

The sensitivity was 76.5% with 95.4 false positives per case. The area of false positives was 938.3 voxels per case. 2 examples of the segmentation results are shown in Fig.2.

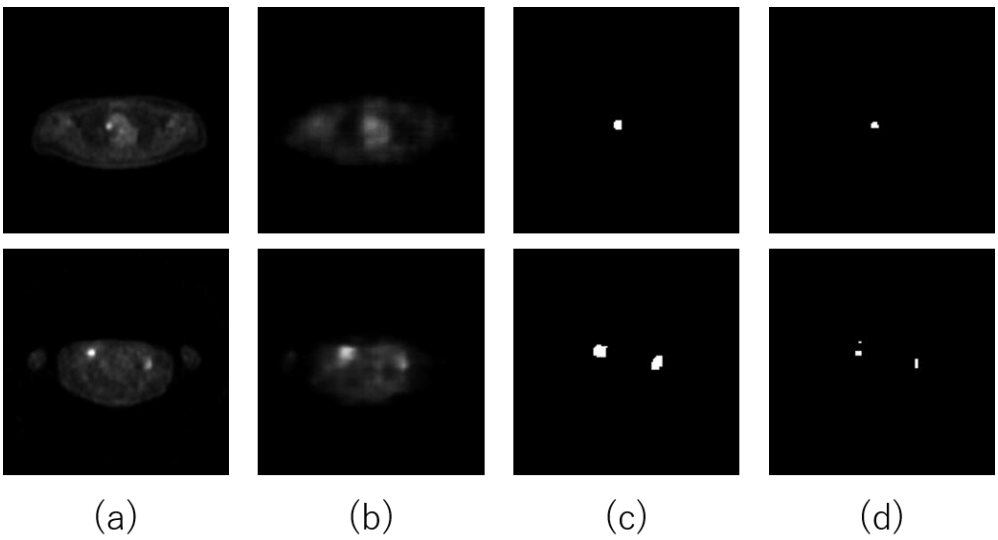


Fig. 2 Segmentation results on PET images.
(a) Axial slice. (b) Generated images. (c) Ground-truth. (d) Segmentation results

4. Conclusion

In this paper, GANs model is introduced to detect anomalous accumulation in PET images. The experimental results show that the unsupervised learning model, which trained only label-less normal cases, contributes to the automatic extract anomaly regions.

Competing interests

The authors have no conflicts of interest directly relevant to the content of this article.

Acknowledgement

Parts of this research were supported by MEXT/JSPS KAKENHI (26108005, 18K12102).

References

[1] Katanoda K, Hori M, Matsuda T et al: An updated report on the trends in cancer incidence and mortality in Japan, 1958–2013. Japanese Journal of Clinical Oncology. Vol. 45. Issue 4. 2015. pp390-401

[2] Nisiyama Y, Kinuya S, Kato T et al: Nuclear medicine practice in Japan: a report of the eighth nationwide survey in 2017. Annals of Nuclear Medicine 33. 2019. pp725-732

[3] Fujita H: AI-based computer-aided diagnosis (AI-CAD): the latest review to read first. Radiological Physics and Technology 13. 2020. pp6-19

[4] Pramuditha Perera, Ramesh Nallapati, Bing Xiang: OCGAN: One-Class Novelty Detection Using GANs With Constrained Latent Representations. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2019. pp2898-2906

読影レポート解析を利用した医用画像データベースからのアノテーション付きデータセット作成に関する初期検討

林　雄一郎[*1]　鈴村　悠輝[*1]　岡崎　真治[*1]

小田　昌宏[*1]　森　健策[*1, 2, 3]

要旨

本稿では，読影レポート解析を利用した医用画像データベースからのアノテーション付きデータセット作成に関する初期検討について述べる．深層学習などの機械学習を用いた医用画像診断支援システムの開発では，アノテーション付きのデータセットを準備する必要がある．しかし，大量の医用画像に対するアノテーション作業は非常に労力を要する．そこで，画像データに付随する読影レポートを解析することでアノテーション付きデータセットを構築する手法を検討する．提案手法では，読影レポートから特定の疾患名を含むデータを取得する．次に読影レポートの 1 文中に特定の疾患名と否定表現が含まれるデータを除外することで特定の疾患を含むデータを取得する．国立情報学研究所の医療画像ビッグデータクラウド基盤のデータに提案手法を適用し，間質性肺炎のアノテーション付きデータセットの作成が可能であることを確認した．

キーワード：読影レポート，自然言語処理，ビッグデータ，機械学習，コンピュータ支援診断

1. はじめに

CT 画像をはじめとした医用画像をコンピュータによって解析することで医師の読影を支援する，コンピュータ支援診断（Computer Aided Diagnosis System: CAD）システムの研究が行われてきた［1, 2］．近年は人工知能技術の発展により，深層学習などの機械学習を利用した CAD システムの研究開発が盛んになっている［3, 4］．機械学習を用いた CAD システムの開発では，学習を行うためにアノテーション付きのデータセットを準備する必要がある．しかし，大量の医用画像に対して手動でアノテーションを付与する作業は非常に労力を要する．そのため，手動でのアノテーション作業を減らすことができれば CAD システムの研究開発にとって非常に有用である．

医療現場において CT 画像などの医用画像を読影した医師は，異常所見や診断を読影レポートに記載する．近年は画像データに加えてこのような読影レポートを含む大規模な医用画像データベースの構築が進められている［5-7］．読影レポートに含まれる異常所見や診断の情報は，アノテーション情報の宝庫であると考えられるため，本研究では読影レポートに着目する．しかし，読影レポートは自由なフォーマットで記載されており，単純な

*1 名古屋大学大学院情報学研究科
〔〒464-8601 名古屋市千種区不老町〕
e-mail: yhayashi@mori.m.is.nagoya-u.ac.jp
*2 名古屋大学情報基盤センター
*3 国立情報学研究所医療ビッグデータ研究センター

疾患名の検索のみでは対象疾患のデータを取得することは難しい．さらに，CAD システムの開発では，対象の疾患を持つ画像に加えて対象の臓器に異常がない画像も必要になることが多い．一方，読影レポートには異常所見がない臓器についての情報が記載されないことも多いため，対象の臓器に異常がない画像を単純な検索のみで取得することは困難である．そこで本研究では，読影レポート特有の記述に着目してレポート中の文章を解析することで，医用画像データベースから特定の疾患を持った画像と異常がない（正常）画像を取得する方法を検討したので報告する．

2. 提案手法

提案手法では，間質性肺炎を対象疾患として，医用画像データベースから，間質性肺炎とアノテーションが付いた画像と，肺が正常（正常肺）とアノテーションが付いた画像からなるアノテーション付きデータセットを構築する．まず，医用画像データベースから読影レポートに対象の疾患名を含む症例を検索して候補症例を取得する．その後，読影レポート特有の記述を考慮したルールベース処理により症例を絞り込むことでデータセットを構築する．間質性肺炎と正常肺それぞれのデータセットの作成手法について下記に述べる．

1）間質性肺炎データセット

データベースから読影レポートに「間質性肺炎」を含む症例を取得する．取得した症例の中には「間質性肺炎を疑う所見は指摘できません」などの間質性肺炎の存在を否定する記述の症例も含まれる．そこで次に，読影レポートの一文中に「間質性肺炎」と「消失」，「間質性肺炎」と「消退」，「間質性肺炎」と「ない」，「間質性肺炎」と「なし」，「間質性肺炎」と「ません」などの間質性肺炎を否定する表現が含まれる症例を除外する．ここで除外された症例の中には，「間質性肺炎は前回 CT から著変ありません」などの，前回の検査で指摘された間質性肺炎が今回の検査でも変化せずに存在している症例も含まれているため，前の処理で除外された症例の中で「間質性肺炎」と「著変なし」のような間質性肺炎を否定していない記述を持つ症例を追加する．

2）正常肺データセット

前述のように対象臓器が正常な画像のデータセットの作成は難しい．そこで，正常肺のデータセット作成では，肺転移に着目する．肺以外の臓器にがんが発生した場合，通常は肺にがんの転移が無いかを調べる．読影レポートに肺転移が無いことが記載されている画像を取得すれば肺が正常である可能性が高い．まず，データベースから読影レポート中に「転移」を含む症例を取得する．次に取得した症例の中で，読影レポートの一文中に「肺」と「転移」と「ない」，「肺」と「転移」と「ません」，「肺」と「転移」と「なし」などの肺に転移が無いことを示す表現を持つ症例を選択する．最後に，選択した症例の中で，「肺炎」，「肺気腫」，「無気肺」，「肺水腫」などの単語を含む肺に異常がある症例や，一文中に「肺転移」と「著変なし」などの肺転移を否定していない記述を持つ症例や，「肺」と「術後」，「肺」と「切除」などの手術後であることを示す記述を持つ症例を除外する．

3. 結果と考察

提案手法を国立情報学研究所が構築している医用画像ビッグデータクラウド基盤［6, 7］上の放射線画像データベースに適用し，間質性肺炎と正常肺のデータセットを構築した．このデータベースには，CT 画像とその読影レポートが格納されている．間質性肺炎のデータセット作成では，画像データベースから読影レポートに「間質性肺炎」を含む症例を検索して 7,646 症例のデータを取得した．その後，ルールベース処理を行い，最終的に 6,440 症例の間質性肺炎データセットを構築した．正常肺のデータセット作成では，画像データベースから読影レポートに「転移」を含む症例を検索して 90,214 症例を取得した．取得した症例に対してルールベース処理を行

前回ＣＴ(2016/06/13)と比較しました。両肺末梢優位に網状影を認め、一部に蜂巣肺を認めます。UIP パターンの間質性肺炎と思われます。前回と著変ありません。明らかな悪性所見は認めません。左腎嚢胞(+)

(a) 間質性肺炎症例と判定された読影レポート

右肺尖部に石灰化を伴う索状影を認め、陳旧性炎症性変化と思われます。間質性肺炎の所見はありません。その他に両肺に特記すべき異常はありません。有意なリンパ節腫大を認めません。胸水を認めません。

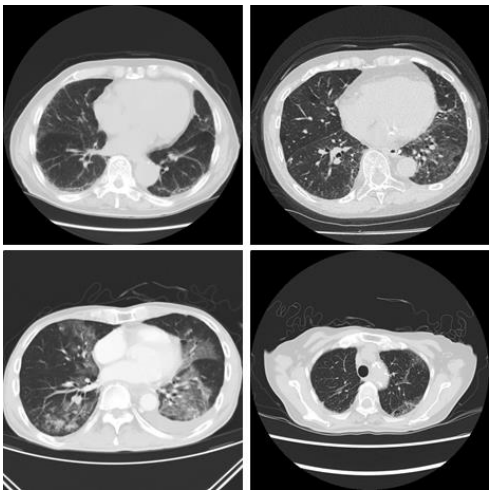
(b) 間質性肺炎症例から除外された読影レポート

図 1 読影レポートの例

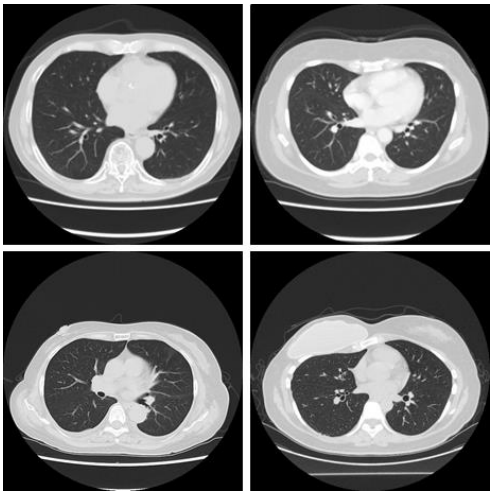
い，最終的に 9,952 症例の正常肺データセットを作成した．読影レポートの例を図 1 に示す．図 1 の 2 つの読影レポートには「間質性肺炎」が含まれており，これらの症例は間質性肺炎データセットの候補症例である．図 1 (a)の症例は，読影レポートに間質性肺炎を否定する表現が含まれていないため，ルールベース処理で候補症例から除外されず，間質性肺炎データセットの症例となった．一方，図 1 (b)の症例は，読影レポートの一文中に「間質性肺炎」と「ません」が含まれており，間質性肺炎が否定されていることから，ルールベース処理で間質性肺炎データセットの候補症例から除外された．構築した間質性肺炎と正常肺のデータセットに含まれる CT 画像の例を図 2 に示す．このように提案手法では，読影レポートを解析することで医用画像データベースから間質性肺炎と正常肺の CT 画像のデータセットを構築することが可能であった．

4. むすび

本稿では，読影レポート解析を利用した医用画像データベースからアノテーション付きデータセット作成手法の検討について述べた．提案手法を国立情報学研究所が構築している



(a) 間質性肺炎



(b) 正常肺

図 2 間質性肺炎データセットと正常肺データセットに含まれる CT 画像例

医用画像ビッグデータクラウド基盤上の放射線画像データベースに適用し，間質性肺炎と正常肺のデータセットを構築することが可能であることを確認した．今後の課題としては，他の疾患のデータセット構築や深層学習を用いた読影レポート解析手法の検討などが挙げられる．

謝辞

日頃ご討論いただく名古屋大学森研究室諸氏に感謝する．本研究の一部は AMED の課題番号 JP20lk1010036 ，JSPS 科研費 JP17H00867，堀科学芸術振興財団の研究助成事業によった．

利益相反の有無

なし

文　献

［1］ 鳥脇純一郎:X 線像のコンピュータ支援診断－研究動向と課題. 電子情報通信学会論文誌 D-II **J83-D-II**: 3-26, 2000

［2］ 小畑秀文:多次元医用画像の知的診断支援システム. Medical Imaging Technology **26**: 157-161, 2008

［3］ Suzuki K: Surgery of deep learning applications to medical image analysis. Medical Imaging Technology **35**: 212-226, 2017

［4］ 藤田広志:診断を支援する. 藤田広志編:2020-2021 年版はじめての医用画像ディープラーニング－基礎・応用・事例－. オーム社, 東京. 2020, pp.124-138

［5］ 橋本正弘:画像診断ナショナルデータベース (Japan Medical Imaging Database: J-MID). 藤田広志編:2020-2021 年版はじめての医用画像ディープラーニング－基礎・応用・事例－. オーム社, 東京. 2020, pp.152-154

［6］ 佐藤真一, 合田憲人, 村尾晃平 他:AMED プロジェクトの概要:クラウド基盤と AI 画像解析. Medical Imaging Technology **37**: 67-71, 2019

［7］ Murao K, Ninomiya Y, Han C, et al: Cloud platform for deep learning-based CAD via collaboration between Japanese medical societies and institutes of informatics. Proc SPIE **11318**: 113180T, 2020

A preliminary study on annotated dataset generation from medical imaging database using analysis of radiology reports

Yuichiro HAYASHI[*1], Yuki SUZUMURA[*1], Shinji OKAZAKI[*1],
Masahiro ODA[*1], Kensaku MORI[*1, 2, 3]

*1 Graduate School of Informatics, Nagoya University
*2 Information Technology Center, Nagoya University
*3 Research Center for Medical Bigdata, National Institute of Informatics

In this paper, we describe a preliminary study on annotated dataset generation from medical imaging database using analysis of radiology reports. The annotated dataset is required in development of computer aided diagnosis system using machine leaning such as deep leaning. However, annotation for a large number of medical images is hard task. We study a method for generating the annotated dataset by analyzing radiology report attached to medical image data. The proposed method extracts data which include specific disease name in radiology report from medical imaging database. To obtain specific disease dataset, we exclude data which contain the specific disease name and negative expressions in one sentence of the radiology reports. We applied the proposed method to medical imaging database in medical imaging bigdata cloud platform build by National Institute Informatics. The results showed that the proposed method e could generate the annotated dataset of interstitial pneumonia.

Key words: Radiological report, Natural Language Processing, Bigdata, Machine leaning, Computer aided diagnosis

A study on Subarachnoid Hemorrhage automatic detection utilized Transfer Learning on extremely imbalanced brain CT datasets

Zhongyang LU[*1], Masahiro ODA[*1], Yuichiro HAYASHI[*1],
Tao HU[*1], Hayato ITO[*1], Takeyuki WATADANI[*2], Osamu ABE[*2],
Masahiro HASHIMOTO[*3], Masahiro JINZAKI[*3], Kensaku MORI[*1,4,5]

Abstract

In this study, we utilized transfer learning methods to deal with the Subarachnoid Hemorrhage detection problem. Due to the blooming Artificial Intelligent technology, the image recognition field has achieved an unprecedently impressive breakthrough. Nevertheless, as the fuel to the engine, performances of AI models are heavily restricted by the quantity and quality of data. Typically, a huge, high-quality, and balanced dataset is essential, whereas considering the actual situation of emergency care, acquiring mess lesion cases is quite a difficult job. In this study, we expose a deep learning based Subarachnoid Hemorrhage (SAH) detection system from extremely imbalanced head computed tomography (CT) scans by using transfer learning technology. Our dataset contains 2,519 cases without Intracranial hemorrhage (ICH) and only 33 cases with SAH. Besides, a huge, high-quality public dataset (containing 3,549 cases SAH data), collected by the Radiological Society of North America (RSNA) institution is applied to improve the performance of our classifier by using transfer learning technology. Furthermore, some training strategies for improving classification performance, just like weighted middle slices (ICH usually happened) are utilized for this experiment.
According to the results of the experiment, our SAH detection system makes a significant improvement via transfer learning. The F-measure achieve around 33% improvement with fine tuning based transfer learning strategy.

Keywords：Transfer Learning, Imbalanced classification, Subarachnoid Hemorrhage detection, CT scan

1. Introduction

Subarachnoid Hemorrhage (SAH) is a category of intracranial hemorrhage (ICH), caused by acute hemorrhage in the subarachnoid space. With the native of high mortality and extremely easily misdiagnosed as other ICH diseases (**Fig. 1** shows examples of no SAH and SAH), SAH diagnosis confuses clinical residents for a long time. Making a general survey for the research of computer-aided based SAH detection, previous works have made spectacular

*1 Graduate School of Informatics, Nagoya University, Aichi, Japan〔Furo-cho, Chikusa-ku, Nagoya, 464-8601, Japan〕
e-mail: zlu@mori.m.is.nagoya-u.ac.jp
*2 Department of Radiology, The University of Tokyo Hospital, Tokyo, Japan
*3 Department of Radiology, Keio University School of Medicine, Tokyo, Japan
*4 Information Technology Center, Nagoya University, Nagoya, Aichi, Japan

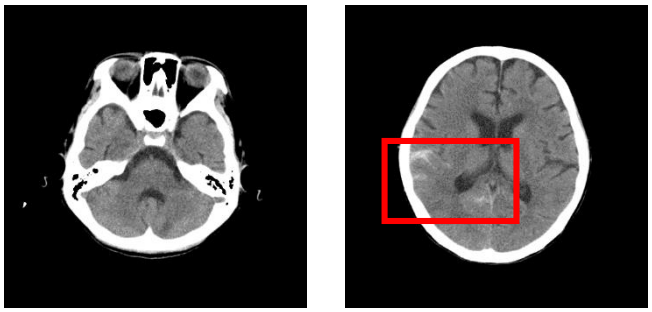


Fig. 1 Example of no-SAH CT and SAH CT,
where the gray arachnoid texture in the red bounding box of the right picture is SAH regions.

achievements [1-3], especially during the past decade, due to the bloomed deep learning technology.

Whereas, the need for numerous data and sufficient well-annotated labels greatly diminish the confidence and attraction of deep learning based diagnosis systems utilized into clinical practice. Collecting and annotating large-scale medical data is not merely an onerous time-consuming drudgery, but also immensely test the professional level of neuroradiologists. Moreover, actually, in clinical practices, cases with SAH lesions are centesimal or millesimal of the cases without hemorrhage cases, which means that experimental samples should be extremely imbalanced for mimicking the clinical scenes. Even though the previous studies mentioned their dataset is in the imbalanced level [1][2], none of their studies has an extremely imbalanced dataset as real-life the clinical situation.

To meet the requirement of needing enormous data and have a better simulation for clinical scenes, in this study, we utilized transfer learning technology for our extremely imbalanced brain CT dataset in subarachnoid hemorrhage CT detection. Our dataset consists of 2,552 cases of brain CT scans corrected by the National Institute of Informatics (NII), including 2,519 cases no hemorrhage cases, and only 33 cases with SAH, which simply restore the clinical scenes. Additionally, in order to obtain variously latent patterns, and learn more high-level features, a mass, high-quality dataset (containing 3,549 cases SAH data), collected by the Radiological Society of North America (RSNA) institution is applied as a source feature domain to discover latent relationships between samples, improving the performance of our classifier. Contributions of this paper are as follows:

(1) We utilized the RSNA data as a source feature domain, solving the problem of lacking SAH patterns, and makes our SAH classifier has a significant improvement on the NII dataset.

(2) Some training strategies like the Stochastic Gradient Descent with Warm Restarts (SGDR) strategy are applied to improve the performance.

2. Methodology

Overview. Transfer Learning is considered as an efficient and effective approach to solve the insufficient data problem. We applied one method of Transfer Learning strategies, named fine-tuning for our experiment. Generally speaking, we retrained the network trained on the RSNA source domain, transferring these latent feature patterns, and applied them to learn latent information of the NII data more effectively. In addition, a stochastic gradient descent method, named SGDR was employed to obtain a better classifier.

Transfer Learning Strategy. Broadly speaking, deep learning-based transfer learning can be categorized into four categories: Instances-based, Mapping-based, Network-based, and Adversarial-based strategy [4]. Among these four kinds of approaches, the Network-based transfer learning strategy is mainly divided into two kinds, Fine-tuning based transfer learning and Convolutional layers as feature encoders based transfer learning. Fine-tuning based transfer learning (fine-tuning, in short) refers to reuse the structure of the trained network and retrain its weights on the source domain (in our experiment which means the RSNA dataset). On the other hand, Convolutional layers as feature encoders based transfer learning (Convs encoders, in short) literally refer to utilize the both the structure and weights of whole convolutional layers as a feature encoder on the target domain (in our experiment which means the NII dataset) directly, and just the weights of fully connection layers are changed.

In our study, both of the two Network-based transfer learning approaches were used in our works, respectively. In the fine-tuning based transfer learning method, we trained our classifier with the RSNA dataset, and we took advantage of their weights as initial values, retraining the model on the target domain (refers to the NII dataset). It suggests that the target model required the 'experience', the latent feature patterns, derived from the source domain, and use the 'experience' to optimize the capacity of the target classifier about discovering the relationship between samples in the target problem. And in the Convs encoders in our studies, due to the differences between the source domain and the target domain, we not only let the model change the weights of fully connection layers, but also unfrozen the weights of two convolutional blocks for making them trainable in order to achieve better performance. The more details will be disclosed in the experiments section.

Warm Restarts Gradient Descent. With decreasing the learning rate makes classifiers acquire the optimal solution, which has become the common sense for the deep learning technique. Whereas, extremely small learning rates cause classifiers to converge quite slowly, even caught in local minimums. Here we employ the Stochastic Gradient Descent with Warm Restarts (SGDR) strategy [5] to let learning rate strategy have periodic variations like a part of cosine function, setting the learning rate of each epoch using cosine annealing schedule. The learning rate γ_t updating method is as

$$\gamma_t = \gamma_{min} + \frac{1}{2}(\gamma_{max} - \gamma_{min})\left(1 + \cos\left(\frac{T_{cur}}{T_i}\pi\right)\right), \tag{1}$$

where γ_{min} and γ_{max} indicate the range of the learning rate in our experiments. T_i means the training epoch number since training start, and T_{cur} is the current epoch number since the end of the last learning rate variation period.

Weighted Binary Cross Entropy. As mentioned earlier, we face the imbalanced problem in both the RSNA and the NII dataset, and imbalanced issues can significantly degrade the performance of classifiers. As a solution, we weighted the Binary Cross Entropy loss function by using the ratios of hemorrhage and no-hemorrhage instances. The weighted Binary Cross Entropy (W-BCE) loss function is defined as

$$L_{W-BCE}(\boldsymbol{x}, \mathbf{y}) = -\frac{1}{B}\sum_{s=1}^{B} w_p(\boldsymbol{y}_s \log p(\boldsymbol{x}_s)) + w_n(1 - \boldsymbol{y}_s)\log(1 - p(\boldsymbol{x}_s)), \tag{2}$$

$$w_p = \frac{N_B}{P_B + N_B} \quad w_n = \frac{P_B}{P_B + N_B},$$

where B means the batchsize of the experiments, $\mathbf{x}$ and $\mathbf{y}$ represent the CT images and their labels in one batch, and $p(\boldsymbol{x}_s)$ is the predicted probability of s-th CT imge in one batch. P_B and N_B mean the numbers of SAH CT scans and

no-SAH CT scans in one batch.

3. Experiments

Datasets introduction and partition. Firstly, 19,530 cases brain CT collected by RSNA, including 8,003 cases with ICH and 11,527 no hemorrhage cases, were applied as the source domain for Transfer Learning. In the 8,003 ICH cases, every CT slice is labeled with six kinds of labels, including five kinds of subtypes of ICH covering subarachnoid hemorrhage (SAH), subdural hemorrhage (SDH), epidural hemorrhage (EDH), intraparenchymal hemorrhage (IPH), and intraventricular hemorrhage (IVH), meanwhile we use the label 'any' to describes one slice has any kind of bleed subtype mentioned above or not. In the NII data, we just have 2,519 cases no hemorrhage brain CT and 33 cases brain CT labeled with SAH or not by slice. Our purpose is to take advantage of the 'experiments' from the RSNA data to improve the performance of SAH detection on the NII data. We used 1,632 cases (including 20 cases SAH data) as training dataset, 409 cases (including 6 SAH cases), and 511 cases (including 7 SAH cases) as validation and test data respectively.

Data augmentation strategy. Half of the CT slices in training data are randomly selected to be data augmented in every training period, including rotation (around ±10 degree), in addition vertical flip and horizontal flip were also used in data augmentation.

Network training and quantitative evaluation. DenseNet-121 was selected as the backbone for our experiments. In order to verify the effectiveness of transfer learning, we used the fine-tuning and Convs encoders respectively, simultaneously trained on the NII data without transfer learnings was employed as the comparison experiment. TheFor making our experiments more reliable, F-measure, accuracy score, recall score, and precision score are used as quantitative evaluation standards.

4. Results and discussion

Table 1 shows the quantitative evaluation results of fine-tuning, Convs encoders, and the experiment trained model on the NII data directly (no transfer learning), respectively. The results suggest that fine-tuning achieves significant improvement, compared with training DenseNet-121 directly on the NII data (in other words, no transfer learning). Whether accuracy, F-measure, recall score, or precision score had better performances, especially in F-measure had above 33% improvement after fine-tuning. It means the classifier learned the useful feature patterns from the RSNA dataset. On the other hand, although with the Convs encoders, the performances of recall score made impressive progress, the other evaluation standards, particularly precision score degraded greatly. The reason that resulted in this situation is considered that the feature domain in the RSNA data is far more complicated than the NII data. There are five kinds of ICH subtypes in the RSNA data, which means using the Convs encoders directly is not suitable, and cause too much misdiagnosis. On the contrary, fine-tuning does not merely learn the latent feature patterns from the RSNA data, but also make the experiments suitable for the target problem.

Table 1 Quantitative evaluation results of experiments.

Experiments	Accuracy	F-measure	Recall	Precision
No transfer learning	0.982	0.301	0.515	0.212
Fine-tuning	**0.994**	**0.638**	**0.721**	**0.577**
Convs encoders	0.923	0.150	0.927	0.08

5. Conclusions

In our studies, we utilized relatively abundant the RSNA data as the source domain to improve the performance of SAH detection in an extremely imbalanced the NII dataset by transfer learning technology. We utilized two kinds of transfer learning approach, and through the fine-tuning based transfer learning method, the performance of SAH detection had a significant improvement in the NII dataset, demonstrating the remarkable effectiveness of transfer learning technology.

As the results showed, even though the performance make a great process by transfer learning methods, the F-measures are not high enough. For fixing the problem, some methods deal with imbalanced problem and the spatial information between the CT scans will be considered in the future.

Competing interests

None.

Acknowledgement

Thank for the help and advices from the members of Mori laboratory and AMED NII team. A part of this research was supported by the AMED Grant Numbers JP19lk1010036, JP20lk1010036, and MEXT/JSPS 26108006, 17H00867.

References

［1］Chilamkurthy S, Ghosh R, Tanamala S, et al.: Deep learning algorithms for detection of critical findings in head CT scans: a retrospective study. The Lancet **392**: 2388-2396, 2018

［2］Lee H, Yune S, Mansouri M, et al.: An explainable deep-learning algorithm for the detection of acute intracranial haemorrhage from small datasets. Nature Biomedical Engineering **3**: 173-182, 2019

［3］Prevedello L M, Erdal B S, Ryu J L, et al.: Automated critical test findings identification and online notification system using artificial intelligence in imaging. Radiology **285**: 923-931, 2017

［4］Tan C, Sun F, Kong T, et al.: A survey on deep transfer learning. International conference on artificial neural networks. Springer, Cham, 270-279, 2018

［5］Loshchilov I, Hutter F. SGDR: Stochastic gradient descent with warm restarts. Learning **10**, 2016

顎骨骨髄炎発症範囲推定のための 頭部CT・切除顎骨CTの位置合わせ自動化

松岡 志宜 * 盛田 健人 * 武田 大介 † 長谷川 巧実 † 若林 哲史 *

要旨

近年，顎骨骨髄炎の治療法として顎骨切除が用いられているが，術前画像中から判別できる骨髄炎範囲が曖昧であるため手術計画の立案が容易ではなく，手術の長時間化等の問題が発生している．そこで，切除範囲の高精度な自動推定を可能とすることで，医師と患者の負担を軽減を行うことが求められている．そのため，術前頭部CT画像と切除顎骨CT画像と詳細な炎症範囲を診断できる病理画像を対応付ける必要があり，本研究では術前頭部CT画像と切除顎骨CT画像の位置合わせ手法を提案した．CT値のみを用いた尤度関数では正確な位置合わせが困難であったため，本研究では3点のキーポイントを併用した尤度関数を用いる剛体位置合わせを行った．被験者5名のデータセットを用いた実験から，正規化相互相関の平均が0.798であり，目視評価からも位置合わせが正確にできていることを確認した．

キーワード：顎骨骨髄炎，コンピュータ支援診断，CT画像，病理画像，位置合わせ

1 はじめに

口腔外科領域において最も難治性の疾患である顎骨骨髄炎の発症により，顎骨が壊死する．治療法として，これまでは抗菌薬治療が行われてきたが，長期に渡る疼痛・排膿や病的骨折などで患者に多大な苦痛を与えるため治療に難渋している [1]．そのため，近年では治療期間を短縮可能とする顎骨の切除が提案されている [2]．顎骨切除術にあたり，術前での切除範囲の決定は術前画像と医師の経験により行われるため，多大な労力を要する．また実際の術中所見と異なることが多く，手術の長時間化や再発による再手術などが課題である．そこで，切除範囲の高精度な自動推定が可能になれば，先に挙げた課題を解決し，医師と患者の負担を軽減することが可能になる．

そこで本研究では，術前頭部CT画像から炎症範囲を高精度に自動推定するために，術前頭部CT画像と顎骨切除術によって得られた切除顎骨の断面を撮影した病理画像を使用する．病理画像から詳細な炎症範囲を診断可能であるため，術前頭部CT画像と対応付けることで術前頭部CT画像中の詳細な炎症範囲を決定することができる．これにより，炎症範囲とCT画像との関連性を明らかにすることで，将来的に術前画像を用いた炎症範囲の自動推定を行うことが可能になる．

前述の研究の中で，術前頭部CT画像と切除顎骨CT画像の位置合わせは必要不可欠である．これまでCT画像やMRI画像を用いた位置合わせの研究はこれまでなされてきている [3] が，顎骨を対象にした研究や，術前画像と切除術で得られた骨の一部分を撮影した画像の位置合わせを行う研究は，いまだ十分になされていない．そのため，本文では術前頭部CT画像と切除顎骨CT画像の位置合わせ法を提案する．

2 データセット

本研究では，神戸大学付属病院で撮影された5名分の術前CT画像および切除顎骨CT画像を使用する．術前の頭部を撮影したものを術前頭部CT画像（図1），顎骨切除術により切除された顎骨を撮影したものを切除顎骨CT画像（図2）として取得した．切除顎骨CT画像については，病理画像取得のために分断する位置を示す切り込みを入れた後に撮影を行った．また，本研究で使用するデータセットの詳細を表1，2に示す．

* 三重大学大学院 工学研究科 情報工学専攻
† 神戸大学附属病院歯科口腔外科

表 1: 術前 CT 画像データセット

被験者	画像サイズ $(voxel)$	分解能 (mm^3)
#1	$512 \times 512 \times 401$	$0.351 \times 0.351 \times 0.3$
#2	$512 \times 512 \times 282$	$0.391 \times 0.391 \times 1.0$
#3	$512 \times 512 \times 531$	$0.351 \times 0.351 \times 0.3$
#4	$512 \times 512 \times 401$	$0.352 \times 0.352 \times 0.3$
#5	$512 \times 512 \times 299$	$0.352 \times 0.352 \times 0.4$

表 2: 切除顎骨 CT 画像データセット

被験者	画像サイズ $(voxel)$	分解能 (mm^3)
#1	$512 \times 512 \times 171$	$0.196 \times 0.196 \times 0.6$
#2	$512 \times 512 \times 445$	$0.196 \times 0.196 \times 0.25$
#3	$512 \times 512 \times 241$	$0.351 \times 0.351 \times 0.3$
#4	$512 \times 512 \times 411$	$0.352 \times 0.352 \times 0.3$
#5	$512 \times 512 \times 141$	$0.195 \times 0.195 \times 0.3$

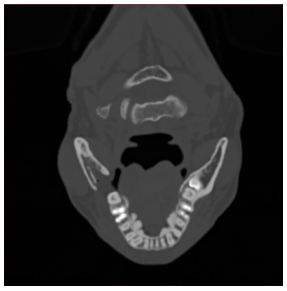


図 1: 術前 CT 画像

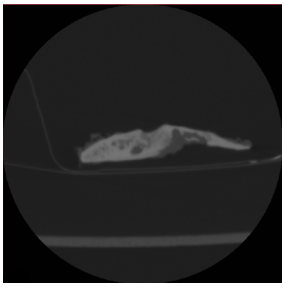


図 2: 切除顎骨 CT 画像

3 提案手法

切除顎骨 CT 画像は，外科的手術により術前 CT 画像中の骨の一部を切除することにより得られた画像であるため，術前 CT 画像の一部と切除顎骨 CT 画像を一致させる部分位置合わせである．部分位置合わせにおいて，画像全体での CT 値の分布の一致度のみの評価では位置合わせが困難であることから，提案手法では術前 CT 画像・切除顎骨 CT 画像のそれぞれの対応する位置から手作業で取得したキーポイントを位置合わせに用いる．キーポイントとして，各画像中で 3 点のキーポイントが一直線にならずそれぞれ位置が離れるように，喉側と口側とその中間から取得した．

キーポイントのみを用いた位置合わせでは，キーポイントの取得誤差が位置合わせ精度に大きな影響を与えるため，本研究では 3.1 節，3.2 節の通り，キーポイントの一致度と画像 CT 値分布の一致度を考慮した二段階の位置合わせ法を提案する．

本研究で提案する位置合わせは，同一患者の CT 画像同士を一致させる剛体位置合わせであるため，3 次元回転行列 $\mathbf{R}$，3 次元並進ベクトル $\mathbf{t}$ で構成される剛体変換行列 $\mathbf{M}$ を用いて切除顎骨 CT 画像を回転・平行移動する．

$$\mathbf{M} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ 0 & 1 \end{bmatrix} \tag{1}$$

また，各軸周りの回転を表す回転行列 $\mathbf{R}_x$，$\mathbf{R}_y$，$\mathbf{R}_z$ で表現される回転行列 $\mathbf{R}$，並進ベクトル $\mathbf{t}$ はそれぞれ次式であらわされるため，

$$\mathbf{R} = \mathbf{R}_x \mathbf{R}_y \mathbf{R}_z \tag{2}$$

$$\mathbf{R}_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix} \tag{3}$$

$$\mathbf{R}_y = \begin{bmatrix} \cos\phi & 0 & \sin\phi \\ 0 & 1 & 0 \\ -\sin\phi & 0 & \cos\phi \end{bmatrix} \tag{4}$$

$$\mathbf{R}_z = \begin{bmatrix} \cos\rho & -\sin\rho & 0 \\ \sin\rho & \cos\rho & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{5}$$

$$\mathbf{t} = (x, y, z) \tag{6}$$

3 次元並進パラメータ (x, y, z) と 3 次元回転パラメータ (θ, ϕ, ρ) の計 6 つのパラメータを，堅実に解のある範囲を狭めていくことができるネルダーミード法を用いて最適化する．また，本文ではパラメータ $(x, y, z, \theta, \phi, \rho)$ の探索範囲をそれぞれ $[x_1, x_2]$, $[y_1, y_2]$, $[z_1, z_2]$, $[\theta_1, \theta_2]$, $[\phi_1, \phi_2]$, $[\rho_1, \rho_2]$ とした時，ネルダーミード法で用いられるシンプレックスの各頂点を下式のように初期化する．

$$S_1 = (x_1, y_1, z_1, \theta_1, \phi_1, \rho_1) \tag{7}$$
$$S_2 = (x_2, y_1, z_1, \theta_1, \phi_1, \rho_1) \tag{8}$$
$$S_3 = (x_2, y_2, z_1, \theta_1, \phi_1, \rho_1) \tag{9}$$
$$S_4 = (x_2, y_2, z_2, \theta_1, \phi_1, \rho_1) \tag{10}$$
$$S_5 = (x_2, y_2, z_2, \theta_2, \phi_1, \rho_1) \tag{11}$$
$$S_6 = (x_2, y_2, z_2, \theta_2, \phi_2, \rho_1) \tag{12}$$
$$S_7 = \frac{S_1 + (x_2, y_2, z_2, \theta_2, \phi_2, \rho_2)}{2} \tag{13}$$

3.1 キーポイント位置合わせ

キーポイントのみを用いた位置合わせでは，平行移動パラメータ (x, y, z) の探索範囲は，画像の x，y，z 軸方向のボクセル数 X，Y，Z を用いて $[-X, X]$，$[-Y, Y]$，$[-Z, Z]$ とする．また，回転パラメータの探索範囲 (θ, ϕ, ρ) はすべて $[-180°, 180°]$ としてシンプレックスの各頂点を初期化する．次に，以下の尤度関数 (14) を最小化するように一段階目の最適化を行う．

$$f = \|\mathbf{p}_0 - \mathbf{q}_0\| + \|\mathbf{p}_1 - \mathbf{q}_1\| + \|\mathbf{p}_2 - \mathbf{q}_2\| \quad (14)$$

ただし，$\mathbf{p}_i$ は術前頭部側の i 番目のキーポイント座標，$\mathbf{q}_i$ は切除顎骨側の i 番目のキーポイント座標とする．

3.2 キーポイントと画像一致度を併用した位置合わせ

ここで，平行移動パラメータ x，y，z の探索範囲をそれぞれ $[x^{'} - sX, x^{'} + sX]$，$[y^{'} - sY, y^{'} + sY]$，$[z^{'} - sZ, z^{'} + sZ]$，回転パラメータ θ，ϕ，ρ の探索範囲をそれぞれ $[\theta^{'} - t, \theta^{'} + t]$，$[\phi^{'} - t, \phi^{'} + t]$，$[\rho^{'} - t, \rho^{'} + t]$ としてシンプレックスの各頂点を (7)〜(13) 式により初期化する．ただし $x^{'}$，$y^{'}$，$z^{'}$, $\theta^{'}$，$\phi^{'}$，$\rho^{'}$ はキーポイント位置合わせにて算出されたパラメータ，X, Y, Z はそれぞれの画像のサイズ．s は探索範囲を狭めるための重み (0.06 で固定)，t は位置合わせ角度の範囲 (60° で固定) とする．

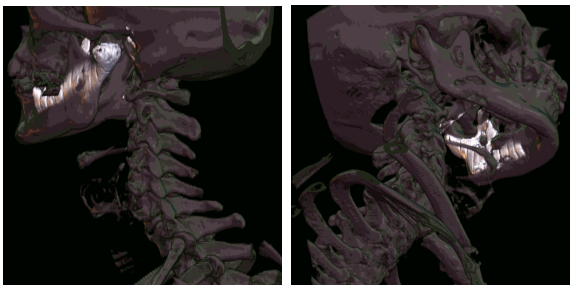
正規化相互相関 (Normalized Cross Correlation;NCC) と尤度関数 (14) を組み合わせた以下の尤度関数 (15) を最小化するように二段階目の最適化を行う．尤度関数の計算には切除顎骨 CT 画像中の顎骨部分と，その顎骨に対応する術前 CT 画像の一部を使用する．

$$g = w\frac{f}{f_0} - (1 - w)\frac{\mathbf{A} \cdot \mathbf{B}^{'}}{\sqrt{\|\mathbf{A}\|\|\mathbf{B}^{'}\|}} \quad (15)$$

ただし，w は重み (0.1 で固定)，f_0 はキーポイント位置合わせで算出した f の最小値，$\mathbf{A}$ は術前頭部の CT 値，$\mathbf{B}^{'}$ は術前頭部 CT 画像に対応する切除顎骨での CT 値とする．

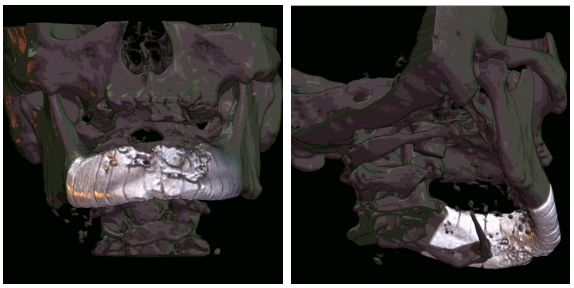
表 3: 各患者ごとの NCC 及び尤度関数 (式 (15)) の値

被験者	NCC	尤度関数値
#1	0.778	-0.773
#2	0.642	-0.589
#3	0.780	-0.692
#4	0.867	-0.838
#5	0.923	-0.881



(a) 左側　　(b) 右側

図 3: 位置合わせ結果（被験者#2）



(a) 左側　　(b) 右側

図 4: 位置合わせ結果（被験者#5）

4 実験結果

実験では，算出された剛体変換行列を用いて，術前 CT 画像と切除顎骨 CT 画像の正規化相互相関，尤度関数 (15) の値を用いて一致度を評価する．

表 3 に位置合わせ終了後の結果を示す．また，図 3a〜4b に位置合わせした画像を示す． 被験者#4，#5 は適切に位置合わせできており，NCC も 0.8 を超えているため十分な結果を得られた．被験者#1 と#3 もほぼ適切に位置合わせできているにも関わらず，NCC が 0.8 を下回っている．これは，切除

した顎骨が術後に空気中の水分を吸ってしまうなどして，同じ位置でも術前 CT 画像と切除顎骨 CT 画像とで CT 値に差異が生じてしまうことが原因であると考えられる．また，同データにおいて NCC と (15) 式の関係が逆になっているが，これは患者番号 0003 のキーポイントに誤差があるため尤度関数 (15) の値が低くなっていると考えられる．被験者 #2 の NCC が#1 と#3 と比較して低いのは，切除した顎骨の頭上側の一部が変形してしまっており，適切に位置が重ならないためであると考えられる．

5　まとめと今後の課題

術前 CT 画像と切除顎骨 CT 画像の位置合わせのために提案手法を提案し，精度の評価を行った．提案手法で位置合わせを行ったところ，高精度に位置合わせが実行できた．今後は切除した顎骨の変形を考慮に入れた位置合わせを行えるように手法を改良する必要がある．また，今後の課題として病理画像と切除顎骨 CT 画像の位置合わせに取り組み，病理画像と術前 CT 画像の位置合わせの手法を提案する．

謝辞

本研究の一部は JSPS 科研費若手研究 20K20221 の助成を受けて行った．

利益相反の有無

本研究に関し利益相反はありません．

参考文献

[1] 島原政司ら, “ビスフォスフォネート投与と関連性があると考えられた顎骨骨髄炎ならびに顎骨壊死に関する調査,” 日本口腔外科学会雑誌, Vol.53, No.10, pp.594–602, 2007.

[2] S. Hayashida, S. Soutome, S. Yanamoto, S. Fujita, T. Hasegawa, T. Komori, Y. Kojima, H. Miyamoto, Y. Shibuya, N. Ueda, T. Kirita, H. Nakahara, M. Shinohara, M. Umeda, “Evaluation of the Treatment Strategies for Medication-Related Osteonecrosis of the Jaws (MRONJ) and the Factors Affecting Treatment Outcome: A Multicenter Retrospective Study with Propensity Score Matching Analysis,” Journal of Bone and Mineral Research, Vol.30, No.10, pp.2022–2029, 2017.

[3] 水田忍，浦山慎一，レザ. A. ゾロフィ，宇山親雄, “3 次元脳地図と医用頭部画像の自動非線形位置合わせ,” Medical Imaging Technology，Vol.16, No.3，pp.175–183, 1998.

Automatic image registration of head CT and resected jaw bone CT for the Computer-aided diagnosis of osteoradionecrosis of the jaws

Yukinobu MATSUOKA[*1], Kento MORITA[*1], Daisuke TAKEDA[*2],
Takumi HASEGAWA[*2] and Tetsushi WAKABAYASHI[*1],

*1 Department of Information Engineering, Graduate School of Engineering, Mie University, Japan
*2 Department of Oral and Maxillofacial Surgery, Graduate School of Medicine, Kobe University, Japan

Recently, surgical excision of the jawbone has been used as a treatment for osteoradionecrosis of the jaws. The precise surgical planning is a main difficulty of this surgery because the osteoradionecrosis region is not clearly appeared in the CT image. Therefore, surgeons require a method to accurately estimate the osteoradionecrosis region in pre-operative CT. Because pathologist can easily annotate the osteoradionecrosis region in the pathological image taken from the resected bone, our research tries to associate the pathological image with the pre-operative head CT image. This paper proposes a method to calculate the resected bone CT image alignment with the pre-operative head CT image. The CT value based likelihood is not suitable to partial registration, therefore, we propose a likelihood function using CT value and manually extracted three key points. Experiments using a dataset of five subjects showed that the mean of the normalized cross-correlation was 0.798, and we confirmed that the alignment was accurate from visual evaluation.

Key words: Osteoradionecrosis of the jaws, Computer-aided diagnosis, Computed tomography image, Image registration, Optimization

深層学習を用いたCT画像解析による肺腺癌手術患者における術後再発予測

佐々木　雄樹[*1,2]　李　鎔範[*2]　青木　正[*3]

関　裕史[*4]　古泉　直也[*4]　尾崎　利郎[*4]

要旨

肺癌の根幹となる治療方法は手術である．しかし，手術を行ってもある一定の割合で再発を来す患者がいる．治療開始前に術後再発の有無を予測できれば，患者個人に適した治療を提供できる可能性がある．そこで本研究は，深層学習を用いて肺腺癌手術患者の術前 CT 画像から，術後再発の有無を予測するコンピュータ支援診断システムを提案した．提案法による分類結果は，感度 0.72，特異度 0.70，AUC 値 0.75 であった．この結果から，肺腺癌手術患者の術後再発の有無を術前 CT 画像から予測できる可能性が示唆された．

キーワード：肺腺癌，術後再発予測，CT，深層学習，CAD

1. はじめに

肺癌は日本人の部位別がん死亡数第 1 位の疾患であり，その中でも肺腺癌は発生頻度の高い分類である[1]．手術治療は肺癌の根幹となる治療方法[2]だが，手術を行ってもある一定の割合で再発を来す肺癌患者が存在する[3]．このような肺癌患者を術前から予測できれば，手術を含めた集学的治療を行うことで治療成績の向上が見込まれ，再発を来さないと予測できれば術後補助化学療法を行う必要がなくなる可能性がある．したがって，治療開始前に術後再発の有無を診断することは，患者個人に適した治療を提供することに繋がり，肺癌手術患者の治療成績向上のために必要性の高い医療技術と言える．

そこで本研究は，深層学習を用いて肺腺癌手術患者の術前 CT 画像から，術後再発の有無を予測評価するコンピュータ支援診断システム（Computer-Aided Diagnosis：CAD）の構築を試みた．

2. 方法

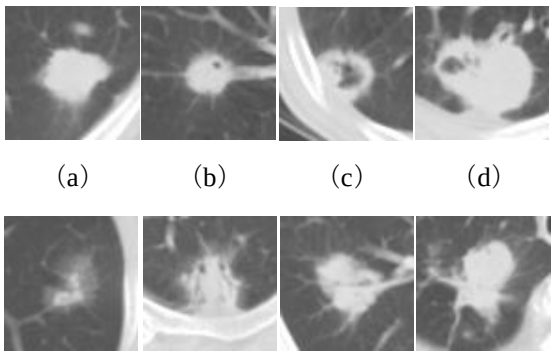
本研究では，深層畳み込みニューラルネットワーク（Deep Convolutional Neural Network：DCNN）の一種である VGG-19[4]を用いた転移学習により，術後再発の有無を自動分類する CAD を構築した．DCNN の学習および分類を行うための入力画像には，術

*1 新潟県立がんセンター新潟病院　中央放射線部
〔〒951-8566 新潟市中央区川岸町 2 丁目 15 番地 3〕
e-mail: ysasaki@niigata-cc.jp
*2 新潟大学大学院保健学研究科　放射線技術科学分野
*3 新潟県立がんセンター新潟病院　呼吸器外科
*4 新潟県立がんセンター新潟病院　放射線診断科

前 CT 画像の腫瘍部分を切り出した画像を用いた．

2.1 入力画像

本研究は新潟県立がんセンター新潟病院の倫理審査委員会の承認を得た．肺腺癌の摘出手術を行った 120 症例（再発 60 例，無再発 60 例）の術前造影 CT 画像を使用した．画像条件はスライス厚を 5mm，Window Width = 1800, Window Level = -600 に設定した．腫瘍の充実成分径が最も大きい横断像を 1 枚選択し，腫瘍が入るように 72×72 pixel（1pixel = 0.68mm）のサイズに切り出し，JPEG 形式に変換した画像を入力画像とした．入力画像例を図 1 に示す.



（a） （b） （c） （d）

（e） （f） （g） （h）

図 1 入力画像例

（a-d）再発症例画像，（e-h）無再発症例画像

2.2 DCNN への学習

ImageNet[5]の画像データベースによって事前学習済みの VGG-19 を用いた転移学習を行った．VGG-19 の畳み込み層における学習済みの重みは更新せずそのまま使用し，全結合層を新しく置き換え，追加した全結合層にのみ学習を行った.

実施環境として，Windows 10，Python 3.6, Tensorflow ver 1.13.1，Keras ver 2.3.1 を使用した．DCNN の学習パラメータとして，損失関数はバイナリークロスエントロピー，学習最適化アルゴリズムには Adam を用いた．学習の繰り返し回数である epoch 数は 50 とした.

3. 結果と考察

構築した CAD の分類精度の検証方法には，10-fold-cross-validation を採用した．分類結果は，感度 0.72（43/60），特異度 0.70（42/60）であった．ROC（Receiver Operating Characteristic）解析における AUC（Area Under the Curve）値は 0.75 であった．提案法を用いることで，肺腺癌手術患者の術後再発の有無を術前 CT 画像から予測できる可能性が示唆された.

4. まとめ

深層学習を用いて術前 CT 画像から肺腺癌手術患者の術後再発を予測する CAD を提案した．提案法は非侵襲的に術後再発の有無を予測することが可能であり，手術患者個人に適した術後療法や経過観察に有用であると考える.

利益相反の有無

なし

文 献

［1］ 国立がん研究センターがん情報サービス / がん登録・統計，https://ganjoho.jp/public/index.html.

［2］ 日本肺癌学会, 肺癌診療ガイドライン 2019 年版

［3］ Noriyoshi S, Prognosis of Lung Cancer Patients in Japan According to Data from the Japanese Joint Committee of Lung Cancer Registry, Respir Investig, 52(6):317-21, 2014

［4］ Karen Simonyan, Andrew Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition. International Conference on Learning Representation, 2015

［5］ IMAGENET，http://image-net.org/

Prediction of Postoperative Recurrence in Patients with Lung Adenocarcinoma by CT Image Analysis Using Deep Learning

Yuki SASAKI[*1,2], Yongbum LEE[*2], Tadashi AOKI[*3], Hiroshi SEKI[*4], Naoya KOIZUMI[*4], Toshiro OZAKI[*4]

*1 Division of Central Radiology, Niigata Cancer Center Hospital

*2 Department of Radiological Technology, Graduate School of Health Sciences, Niigata University

*3 Department of Thoracic Surgery, Niigata Cancer Center Hospital

*4 Department of Radiology, Niigata Cancer Center Hospital

The surgery is main treatment method of lung cancer. However, some patients have a recurrence of cancer even after surgery. If the postoperative recurrence can be predicted before the start of treatment, it may be possible to provide suitable treatment for each patient. Therefore, this paper proposed a computer-aided diagnosis system that predicted postoperative recurrence with preoperative CT image of lung adenocarcinoma surgery patients by using deep learning. The experimental result showed that the sensitivity for the detection of the recurrence patients was 0.72 and specificity was 0.70. The area under the curve of the ROC analysis was 0.75. These results indicate that the proposed system may predict postoperative recurrence in patients with lung adenocarcinoma surgery with preoperative CT image.

Key words: Lung Adenocarcinoma, Prediction of Postoperative Recurrence, CT, Deep Learning, CAD

COVID-19 症例の定量評価のための CT 像からの肺野自動セグメンテーション

小田　昌宏[*1]　林　雄一郎[*1]　大竹　義人[*2,3]

橋本　正弘[*4]　明石　敏昭[*5]　森　健策[*1,3]

要旨

本稿では，新型コロナウイルス感染症（COVID-19）症例の CT 像からの肺野内陰影領域等のセグメンテーション手法について述べる．2019 年 12 月に発生した COVID-19 は急速に世界中に広まり多数の患者が発生すると共に，日常生活に大きな打撃を与えている．COVID-19 を含むウイルス性肺炎の診断において胸部 CT 像が有用である．医師の診断を支援するために，CT 像からの肺野状態の定量評価手法の確立が急務である．本稿では fully convolutional network を用いた CT 像からの肺野自動セグメンテーション手法を提案する．COVID-19 を含む疾患による肺野状態の定量評価には肺野内陰影の状態を把握することが重要である．そのため，提案手法ではすりガラス影や浸潤影を含む陰影領域とそれ以外の領域（正常領域と呼ぶ）を識別してセグメンテーション可能とした．提案手法を COVID-19 症例の CT 像に適用したところ，正常領域の DICE 係数は 0.906，陰影領域は 0.721 となった．

キーワード：COVID-19，肺野，セグメンテーション，深層学習

1. はじめに

2019 年 12 月に発生した新型コロナウイルス感染症（COVID-19）は急速に世界に広まり，多数の患者が発生している．2020 年 8 月上旬における世界の累計感染者数は約 2000 万人，死亡者数は約 74 万人となっている[1]．本邦の累計感染者数は約 5 万人，死亡者数は約千人であり[2]，さらに増加を続けている．COVID-19 感染者に対し適切な治療を行い，他者への感染を防ぐ上で，正確な診断方法が求められる．COVID-19 の検査では RT-PCR（reverse transcriptase polymerase chain reaction）が多く用いられるが，その Sensitivity は 42%から 71%と低い[3]．これに対し CT 像による COVID-19 検査の Sensitivity は 97%と高い[4]．日本は人口あたりの CT 装置保有数が高いため，COVID-19 検査において CT 像を活用できると考えられる．CT 像を用いた医師の診断を支援し，判断の定量化を目指す上で，コンピュータによる肺野状態の定量評価手法の確立が求められる．COVID-19 による肺炎を含むウイルス性肺炎では肺野内にすりガラス影や浸潤影が発生するため[3]，これらを含めた肺野状態の定量評価が必要である．

本稿では，fully convolutional network（FCN）を用いた CT 像からの肺野領域自動セグメンテーション手法を提案する．セグメンテーション対象は，すり

*1 名古屋大学大学院情報学研究科
〔〒464-8601 愛知県名古屋市千種区不老町〕
e-mail: moda@mori.m.is.nagoya-u.ac.jp
*2 奈良先端科学技術大学院大学
先端科学技術研究科
*3 国立情報学研究所
医療ビッグデータ研究センター
*4 慶応義塾大学医学部放射線科
*5 順天堂大学医学部放射線医学講座
投稿受付：2020 年 8 月 5 日

ガラス影や浸潤影を含む「陰影領域」とそれ以外の領域（便宜上「正常領域」と呼ぶ）の 2 領域である．

2．手法

2.1 FCN への入力画像

FCN に入力する画像は CT 像から取り出した Axial 画像（2D 画像）である．FCN に入力する際の画像サイズは 256×256 画素とした．2D 画像処理を採用した理由は，現時点で多数の 3D の正解領域を準備できておらず多数のパラメータを含む 3D FCN の学習が困難であること，診断に用いる CT 像に 1.0～5.0 mm まで様々なスライス厚のものが含まれ，体軸方向の情報量のばらつきを統一的に扱うことが困難だからである．

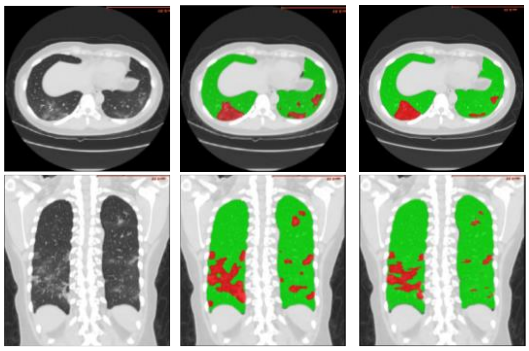
正解領域は正常領域，陰影領域，体外領域を手動セグメンテーションしたものを用いる．

2.2 FCN の構造

FCN の構造は U-Net[5]をベースとし，様々な陰影や形状のパターンを持つ正常及び陰影領域をセグメンテーションするため，いくつかの修正を加えた．まず全ての convolution 層の後に batch normalization を追加した．また，様々な空間スケール情報を FCN のエンコーダ及び bottleneck block での特徴変換において利用するため，FCN のエンコーダ部分に dense pooling connection（DPN）[6]を導入した．Pooling における空間情報欠落を軽減するため，max pooling と average pooling を併用する mixed pooling を使用した．さらに，convolution において空間的に広範囲の情報を利用するため，エンコーダに含まれる 2 つの convolution 層を dilated convolution block で置き換えた．dilated convolution block 内では複数の dilation rate の dilated convolution[7] を行い，その結果を concatenate で統合する．

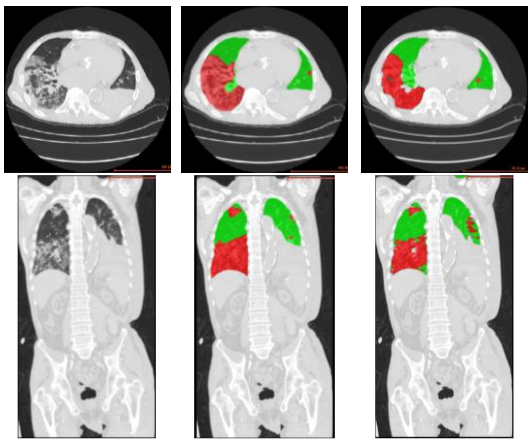
2.3 セグメンテーション処理

提案手法によりセグメンテーションを行う上で，学習データを用いて FCN を学習し，その後推定用 CT 像の Axial 画像を学習済み FCN に入力してセグメンテーション結果を得る．



(a) (b) (c)

図 1 1.25 mm 厚胸部 CT 像での結果．緑色が正常領域，赤色が陰影領域である．上側が Axial，下側が Coronal スライス．(a)CT 像，(b)正解領域，(c)セグメンテーション結果．

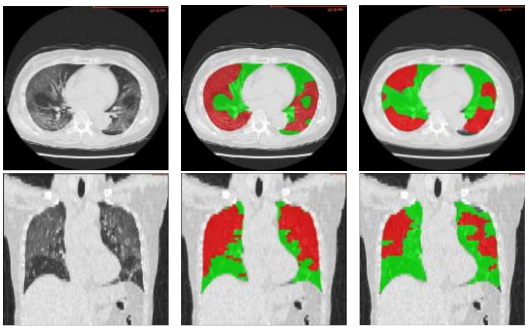


(a) (b) (c)

図 2 1.00 mm 厚の胸部から腹部まで含む CT 像での結果．緑色が正常領域，赤色が陰影領域である．上側が Axial，下側が Coronal スライス．(a)CT 像，(b)正解領域，(c)セグメンテーション結果．

3．実験及び考察

本邦の複数の医療機関で撮影された，COVID-19 感染例の CT 像 20 例を用いて実験を行った．CT 像の仕様は，画像サイズ 512×512 pixels，スライス枚数 56～722 枚，画素間隔 0.63～0.78 mm，スライス厚 1.00～5.00 mm である．FCN 学習時のミニバッチ数は 30，epoch 数は 75 とした．損失関数は Generalized dice loss を使用した[8]．正解領域は



(a) (b) (c)

図 3 5.00 mm 厚胸部 CT 像での結果．緑色が正常領域，赤色が陰影領域である．上側が Axial，下側が Coronal スライス．(a)CT 像，(b)正解領域，(c)セグメンテーション結果．

表 1 提案 FCN と U-Net を用いた場合のセグメンテーション結果の DICE 係数．

セグメンテーション手法	DICE 係数	
	正常領域	陰影領域
提案 FCN	**0.906**	**0.721**
U-Net[5]	0.897	0.708

医師によって確認されたものである．セグメンテーション結果は 10 分割交差検定により評価を行い，評価指標として正常領域と陰影領域それぞれの DICE 係数を用いた．

提案手法によるセグメンテーション結果を図 1-3 に示す．提案 FCN と U-Net[5]を用いた場合のセグメンテーション結果の DICE 係数の比較を表 1 に示す．正常領域及び陰影領域の両方において提案 FCN は高い精度を達成した．

陰影領域が少ない図 1 の症例，多い図 2,3 の症例それぞれにおいて，正常及び陰影領域を高い精度でセグメンテーションすることができた．実験に用いたデータセットには図 1,3 のように胸部のみを含む CT 像，図 2 のように胸部及び腹部を含む CT 像が混在しているが，撮影範囲の違いによる影響を大きく受けずにセグメンテーションが可能である．また，CT 像のスライス厚は 1.00～5.00 mm の範囲の様々なもの（図 2 と図 3 のように）を混在させ学習及び評価に用いた．提案手法は Axial スライス上でセグメンテーションを行うため，スライス厚の違いによる影響を受けずにセグメンテーションが可能である．

提案手法を COVID-19 の診断支援において用いる場合，様々な撮影条件で撮影された CT 像に対応可能であることが重要である．実験結果に示したように，提案手法は撮影範囲やスライス厚の変化に頑健なセグメンテーションを行うことができる．

提案手法は肺野内の正常領域を高い精度でセグメンテーションできているものの，陰影領域の F 値が低い．図 2 では高濃度値の陰影領域，図 3 では淡いすりガラス影の一部を誤って正常領域としている．これは学習データ数が少ないために陰影パターンを十分に学習できず，見落としたと考えられる．より多くの学習データの作成と，data augmentation 等を用いて多くの陰影パターンを学習させることが必要である．

4. むすび

本稿では COVID-19 症例の CT 像からの肺野内の正常領域及び陰影領域の自動セグメンテーション手法を提案した．セグメンテーションには dense pooling connection，dilated convolution を用いた U-Net を使用した．実験では正常領域のセグメンテーション結果の DICE 係数が 0.906，陰影領域が 0.721 となり，正常領域を高精度にセグメンテーションすることが可能であった．今後の課題として，学習データ追加等による陰影パターンの学習を通した陰影領域のセグメンテーション精度向上が挙げられる．

謝辞

本研究の一部は，JSPS 科研費 17H00867，17K20099，AMED 研究費「医療ビッグデータ利活用を促進するクラウド基盤・AI 画像解析に関する研究」JP19lk1010036，JP20lk1010036，JSPS 二国間交流事業，公益財団法人 堀科学芸術振興財団，公益財団法人 人工知能研究振興財団によった．

利益相反の有無

なし

文　献

［1］ Coronavirus Update, https://www.worldometers.info/coronavirus/（2020 年 8 月 11 日アクセス）

［2］ 厚生労働省 国内の発生状況など, https://www.mhlw.go.jp/stf/covid-19/kokunainohasseijoukyou.html（2020 年 8 月 11 日アクセス）

［3］ Simpson S, Kay FU, Abbara S, et al.: Radiological Society of North America Expert Consensus Statement on Reporting Chest CT Findings Related to COVID-19. Endorsed by the Society of Thoracic Radiology, the American College of Radiology, and RSNA. Radiology: Cardiothoracic Imaging **2**(2): 2020

［4］ Ai T, Yang Z, Hou H, et al.: Correlation of Chest CT and RT-PCR Testing for Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases. Radiology **296**(2): 2020

［5］ Ronneberger O, Fischer P, Brox T: U-Net: Convolutional Networks for Biomedical Image Segmentation. MICCAI **9351**: 234-241, 2015

［6］ Playout C, Duval R, Farida C: A Multitask Learning Architecture for Simultaneous Segmentation of Bright and Red Lesions in Fundus Images. MICCAI **11071**: 101-108, 2018

［7］ Yu F, Koltun V: Multi-scale Context Aggregation by Dilated Convolutions. ICLR, 2016

［8］ Sudre CH, Li W, Vercauteren T, et al.: Generalized Dice Overlap as a Deep Learning Loss Function for Highly Unbalanced Segmentations. DLMIA, ML-CDS **10553**: 240-248, 2017

Automated Segmentation Method of Lung Region from CT Volumes for Quantitative Analysis of COVID-19 Cases

Masahiro ODA[*1], Yuichiro HAYASHI[*1], Yoshito OTAKE[*2,3],
Masahiro HASHIMOTO[*4], Toshiaki AKASHI[*5], Kensaku MORI[*1,3]

*1 Graduate School of Informatics, Nagoya University
*2 Graduate School of Science and Technology,
Nara Institute of Science and Technology
*3 Research Center for Medical Bigdata, National Institute of Informatics
*4 Department of Radiology, Keio University School of Medicine
*5 Department of Radiology, Juntendo University

This paper proposes a segmentation method of the lung including shadow regions from CT volumes of novel coronavirus disease 2019 (COVID-19) cases. COVID-19 is spreading over the world causing large number of infected patients. Our lifestyle has significantly changed to diminish the incidence of infection. In diagnosis of viral pneumonia including COVID-19, CT volume is quite useful. A quantitative analysis method of condition of the lung from CT volumes is required to assist diagnosis of COVID-19 pneumonia cases. We propose a segmentation method of the lung from CT volumes using a fully convolutional network. In the diagnosis of COVID-19 cases, quantitative analysis of condition of normal and shadow regions in the lung is important. Our method discriminates normal region and shadow region (including ground-glass opacity and consolidation) in the lung to segment them separately. We applied the proposed method to CT volumes of COVID-19 cases. DICE coefficients of the segmentation results of normal and shadow regions were 0.906 and 0.721, respectively.

Key words: COVID-19, Lung region, Segmentation, Deep learning

Preliminary Study on Classification of Interstitial Cystitis Using Cystoscopy Images

Tao CHU[*1], Masahiro ODA[*1], Akira FURUTA[*2],
Tokunori YAMAMOTO[*3], and Kensaku MORI[*1,4,5]

Abstract

Interstitial cystitis (IC) is a chronic condition which nonspecific chronic inflammation of the bladder is observed, however, the clear etiology of autoantibodies has not yet been identified. In actual clinical practice, bladder biopsy with suspicion of IC may occasionally reveal bladder intraepithelial carcinoma (CIS). On the other hand, the surgery with suspicion of CIS may also occasionally reveal there is no cancer. Treatment is generally difficult and persistent symptoms severely compromise the patients' quality of life. In this paper, we proposed a method using a convolutional neural network (CNN) to classify the interstitial cystitis and bladder intraepithelial carcinoma through the cystoscopy images, and got some depth estimation results. We also used a technique called Gradient-weighted Class Activation Mapping (Grad-CAM) to produce visual explanations for decisions from CNN. By using the CNN, we achieved the AUC score of 0.694 on this classification task.

Keywords：Cystoscopy images, CNN, Interstitial cystitis, Classification

1. Introduction

Interstitial cystitis (IC) is a chronic condition which nonspecific chronic inflammation of the bladder is observed, however, the clear etiology of autoantibodies has not yet been identified. Major clinical symptoms of IC include frequent urination (about 91%) and urgency (about 62%) [1]. Among severely ill patients (mainly IC patients), the extremely urination sensation regardless of day or night, and the quality of sleeping is markedly disturbed by constant urinary urgency. It is a disease in which the quality of life is extremely impaired, such as the inability to concentrate on work and housework.

At present, there are no diagnostic markers for IC, thus, according to the cystoscopy images to confirm Hanna-type lesions (necessary inspection) and pathological examination of biopsy specimens (auxiliary inspection) are crucial.

---

*1 Graduate School of Informatics, Nagoya University〔Furo-cho, Chikusa-ku, Nagoya, 464-8601 JAPAN〕
e-mail: tchu@mori.m.is.nagoya-u.ac.jp
*2 The Jikei University School of Medicine
*3 Nagoya University Graduate School of Medicine
*4 Information Technology Center, Nagoya University
*5 Research Center for Medical Bigdata, National Institute of Informatics

Besides, in actual clinical practice, bladder biopsy with suspicion of IC may occasionally reveal bladder intraepithelial carcinoma (CIS), on the other hand, surgery was suspected of CIS, however, pathological results showed there is no cancer. The Hanna type lesion is a lesion in which epithelial detachment of the bladder and capillaries are caused by chronic inflammation of unknown cause. On the other hand, CIS is a non-elevated superficial bladder cancer and is accompanied by angiogenesis, so it is difficult to distinguish them visually. Since IC is a rare disease, it is pretty hard to even for urologists to diagnose. It takes an average of 36.5 months to be diagnosed with IC, and the quality of life of patients is significantly impaired during that period [1]. Therefore, it is crucial to find a new way to assist doctors in diagnosing IC early.

We are attempting to utilize the deep learning to help doctors to classify cystoscopy images, and the final purpose of this research is to develop a software for assisting cystoscopy image diagnosis with artificial intelligence, and to contribute to the early detection and early treatment of IC.

2. Method

CNN (Convolutional Neural Networks) is a kind of deep learning method mainly used in the field of image recognition [2]. In this paper, we used a 2D-CNN to classify the cystoscopy images. In our experiment, the convolutional neural network contains three convolution blocks and three fully connected layers. Each convolution block consists of a 2D convolutional layer, a leaky relu activation layer, a maxpooling layer, and a batch normalization layer. The last layer is the softmax which can output the probability of the image category. The structure is shown as the Fig. 1.

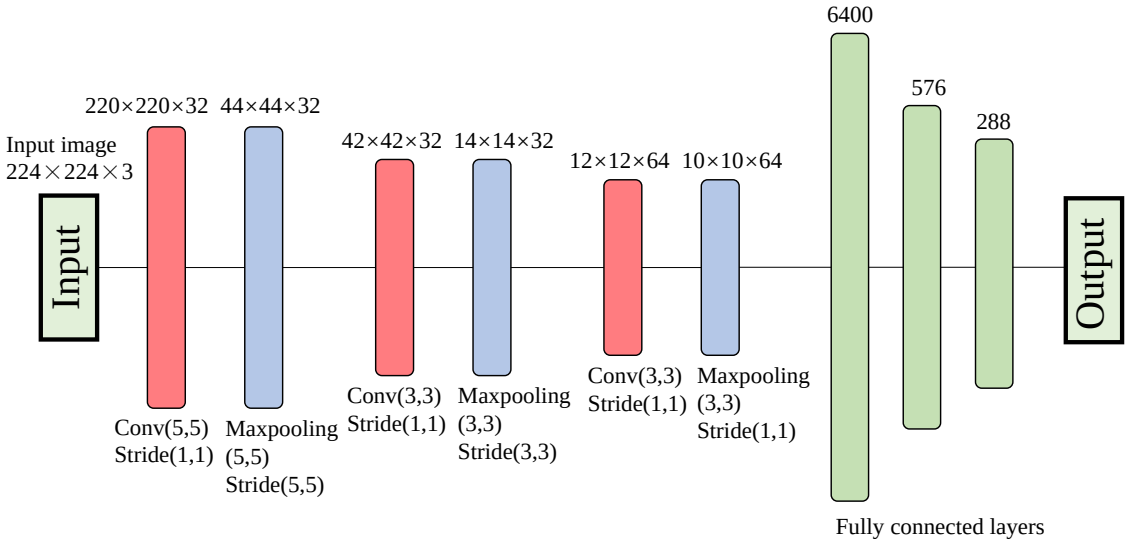


Fig. 1 The structure of the CNN. Conv is the convolutional layer, maxpooling is the maxpooling layer. The numbers after the conv, maxpooling represent the size of the convolution kernel. The numbers behind the stride indicate the stride width in the vertical and horizontal directions. The numbers on the top represent the length, width, and number of channels of the image, and the numbers on the top of the fully connected layers are the number of units in the layer.

In our experiment, we used the Gradient-weighted Class Activation Mapping (Grad-CAM) to analyze the trained CNN model [3]. Grad-CAM uses the weight of trained CNN model to determine the region on the input image that contributed to the classification of the particular class, and the form of output is the heat map.

3. Experiments and results

The original dataset consists of 176 cystoscopy images including the 135 IC images and 41 CIS images, and the sizes of the images are different from each other. Since the amount of dataset is crucial for the deep learning, we need to find a way to preprocess the original dataset to expand it. In this experiment, as is shown in the Fig. 2 we rescaled the size of the images, rotated the images every ten degrees, and we cropped the images into the size of 224×224 pixels according to the center of the rescaled and rotated dataset. In this way, we could obtain the dataset which could expand the original dataset by 36 times.

Since we obtained the expanded dataset, we were able to start our experiment. We divided each category of dataset (IC or CIS) into training, validation and test sets. We used the training and validation sets to train the CNN and obtained a trained model, then we utilized the test set to test the model. The classification result of test set was set as true positive (TP), true negative (TN), false positive (FP) and false negative (FN) according to the correspondence between the correct label and the predicted label. True positive is an image with correct label CIS and predictive label with CIS, true negative is image with correct label IC and predictive label with IC, false positive is image with correct label IC and predictive label with CIS, false negative is correct answer images with CIS label and predictive label with IC.

The evaluation indexes such as accuracy, sensitivity and specificity calculated from the number of test result of these images were used in our experiment. Accuracy, sensitivity and specificity are defined as

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}}, \tag{1}$$

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP+FP}}, \tag{2}$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN+FN}}, \tag{3}$$

where TP, FP, TP and FP are the numbers of the true positive, false positive, true negative and false negative images in a classification result, respectively.

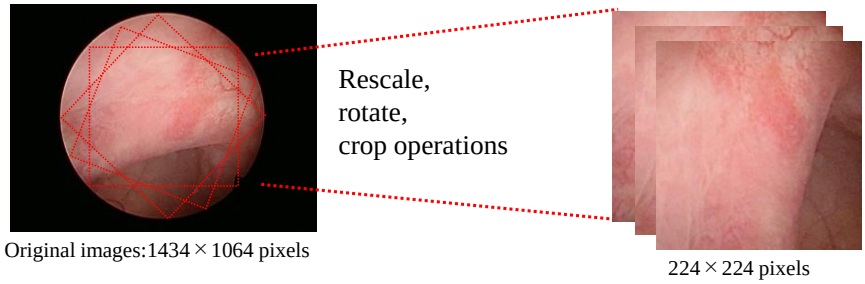


Fig. 2 Data preprocessing (Expanded to 36 times).

Table 1 Result of the accuracy, sensitivity and specificity.

Accuracy	Sensitivity	Specificity
0.568	0.761	0.738

The results we got is shown in the Table 1. We also generated an ROC curve to evaluate whether the model we got could distinguish two types of dataset well. As is shown in the Fig. 3, we got the ROC curve and the AUC score. The AUC score equals 0.694.

To visualize the result, we used the Grad-CAM which utilizes the weight of trained CNN model to determine the region on the input image that contributed to the classification of the particular class, the output form of Grad-CAM is the heat map. The results are shown as the Figs. 4-7. The red regions represent the high score of the focus pixels for a category, the orange and yellow regions represent a lower score than red regions, and the blue regions represent a lower score than orange and yellow regions. According to the results, in our experiments, when the features two types of images are pretty obvious, the model could distinguish them well.

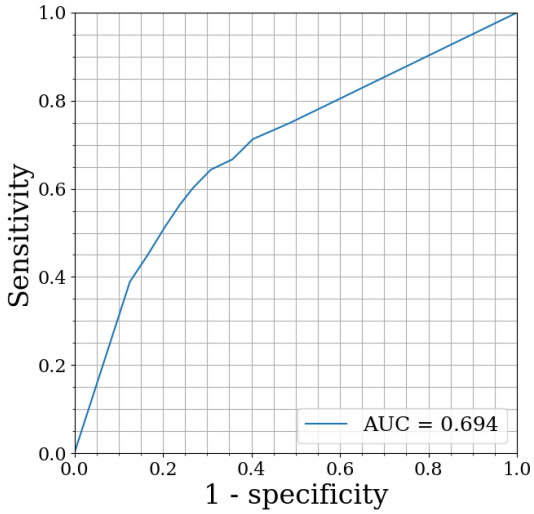


Fig. 3 ROC curve. The horizontal axis is (1-specificity), and the vertical axis is sensitivity.

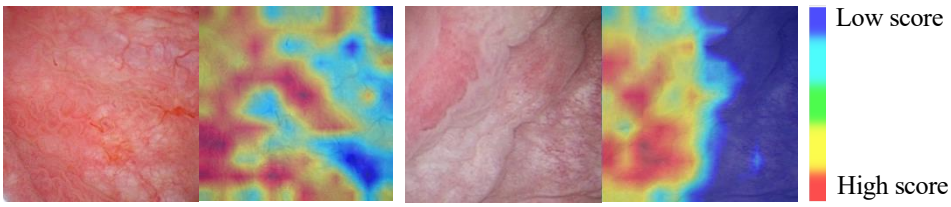


Fig. 4 The result of visualizing the region where the TP images of the test data, the correct label of images is CIS and predictive label is CIS.

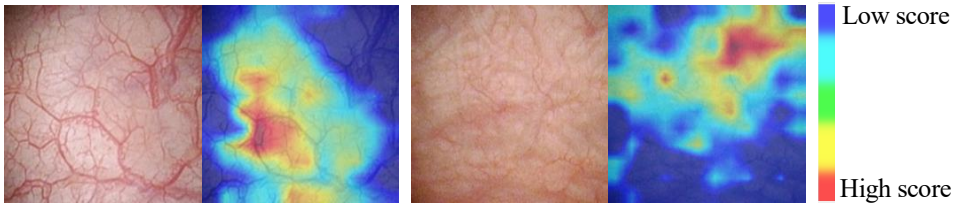


Fig. 5 The result of visualizing the region where the TN images of the test data, the correct label of images is IC and predictive label is IC.

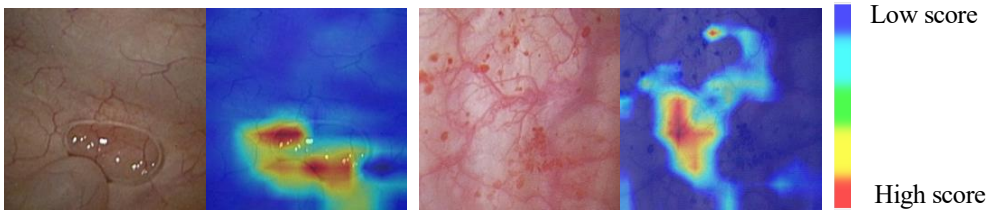


Fig. 6 The result of visualizing the region where the FP images of the test data, the correct label of images is IC and predictive label is CIS.

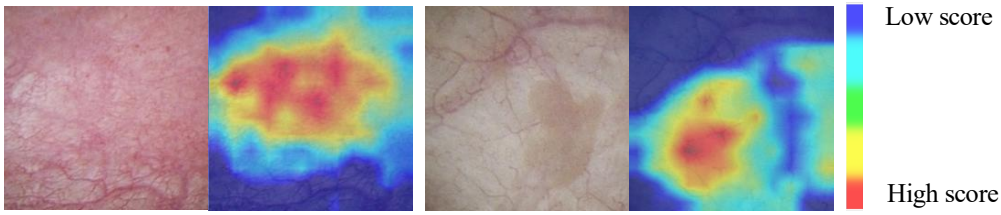


Fig. 7 The result of visualizing the region where the FN images of the test data, the correct label of images is CIS and predictive label is IC.

4. Conclusions

In our experiment, we used the CNN to classify the cystoscopy images, and achieved the AUC score of 0.694 on this classification task. We used the Grad-CAM to evaluate the models we got, and we can acknowledge that the trained models have the ability to determine the region on the input images. We obtained the preliminary classification result of the cystoscopy images, which tells us that it is possible to use deep learning to classify the cystoscopy images and then achieve the purpose of medical assistance.

Future work includes increment of the amount of data and development of a software to assist doctors in early detection and early treatment of IC.

Competing interests

None.

Acknowledgement

Parts of this research was supported by MEXT·JSPS KAKENHI (26108006, 17H00867, 17K20099), the JSPS Bilateral

International Collaboration Grants, the AMED (19hs0110006h0003) and the Hori Sciences & Arts Foundation.

References

［1］Homma Y, Akiyama Y, et al.: Clinical Guideline for Interstitial Cystitis/Bladder Pain Syndrome, 1-5, 2019

［2］Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift, arXiv preprint arXiv:1502.03167, 2015

［3］Selvaraju R R, Cogswell M, Das A, et al.: Grad-cam: Visual explanations from deep networks via gradient-based localization, proceedings of the IEEE international conference on computer vision, 618-626, 2017

A novel topologically invariant radiomic features for prognostic prediction in lung cancer patients

Kenta NINOMIYA[*1], Hidetaka ARIMURA[*2]

Abstract

We aimed to develop a novel topologically invariant radiomic features (TIF) for prognostic prediction in lung cancer patients. TIFs were calculated using the texture analysis of Betti number maps which represent topologically invariant characteristics on computed tomography images. The feasibility of TIFs was evaluated using statistically significant differences to compare the survival curves of high- and low-risk patients stratified based on the TIFs. Further, a deep learning (DL)-based model was built to compare its feasibility to TIFs in prognostic prediction. The p-values between the survival curves for TIF and DL were 3.4×10^{-5} and 1.2×10^{-1}, respectively. The proposed approach showed a promising potential to the prognostic prediction for lung cancer patients.

Keywords：Radiomics, Topologically invariant tumor characteristics, Cox proportional hazard model, Lung cancer

1. Introduction

Several radiomics studies have attempted to evaluate lung cancer prognosis by characterizing the tumor traits [1–3]. Prognosis prediction using a deep learning (DL) was slightly more accurate than that employing conventional radiomic features, although there was not always a significant performance improvement in two-year survival prediction among lung cancer patients ($p > 0.05$, permutation test) [4]. These conventional methods could not adeqately interpret underlying prognostic tumor traits of lung cancer that are present in medical images.

The topologically invariant Betti numbers, which indicate the number of n-dimensional holes in an object, have been succesfully utilized in the medical image analysis such as detection or classification of prognostic cancer legions in pathological, magnetic resonance or computed tomography (CT) images [5–8]. The zero- and one-dimensional Betti numbers, which can be calculated from two-dimensional images, indicate the number of connected components (0-dimensional holes; b0) and holes (1-dimensional holes; b1) in an object. However, Betti numbers have not been utilized for the evaluation of prognosis in lung cancer captured on the CT images.

We aimed to develop a novel topologically invariant radiomic features (TIFs) based on the Betti numbers for prognostic prediction in lung cancer patients. Further, we compared the feasibility of the TIFs with a DL model.

---

*1 Department of Health Sciences, Graduate School of Medical Sciences, Kyushu University〔3-1-1 Maidashi, Higashi-ku, Fukuoka, 812-8582, Japan〕

e-mail: knt.eat.24splash@gmail.com

*2 Sciences Department of Health Sciences, Faculty of Medical Sciences, Kyushu University

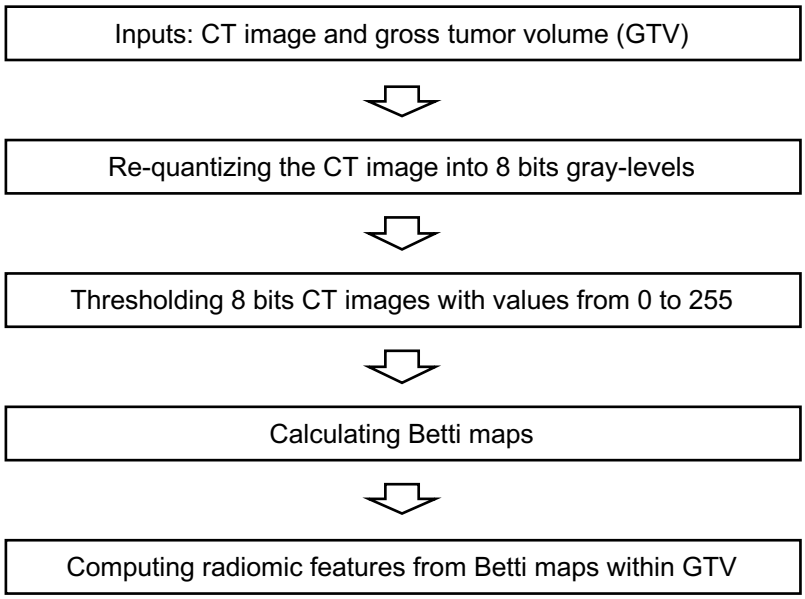


Figure 1. Overall scheme for the calculation of the topologically invariant radiomic features.

2. Materials and frameworks

A total of 205 CT images of lung cancer patients (male: 147; female: 58; age range: 43–92 y; subject to complete follow up: 135; censored: 70; stages I: 38 (18.5%); II: 22 (10.7%); IIIA: 52 (25.4%); IIIB: 92 (44.9%); unavailable stage data: 1 (0.5%)), who underwent either radiotherapy or radio-chemotherapy, were selected from the publicly available non-small-cell lung carcinoma (NSCLC)-Radiomics database of The Cancer Imaging Archive. The gross tumor volumes (GTVs) were defined based on fused ^{18}F-fluor-odeoxyglucose positron emission tomography (FDGPET)-CT images. The image features were extracted using an axial plane of each CT image that contained the maximum GTV area.

Figure 1 shows the overall scheme for the calculation of TIFs. A total of 13,824 TIFs were calculated within the GTV from 54 features based on the histogram and 4 texture characterization matrices, i.e., the gray level co-occurrence matrix, the gray level run-length matrix, the gray level size-zone matrix, and the neighborhood gray-tone difference matrix, for 256 Betti maps. These maps were obtained by calculating the ratio of one- to zero-dimensional Betti numbers (b1/b0) on binary images within local regions of interest (ROIs) as determined by a 7×7 kernel based on a convolutional computation procedure. The b0 and b1 values were computed by counting connected components and holes, respectively, with eight-connectivity within local ROIs. The binary images were derived from eight-bit CT images by thresholding the images with a value from 0 to 255. Eight-bit CT images were obtained based on a look-up table ranging 0 to 255 corresponding to a range of −1000 to 1500 Hounsfield units.

The DL model based on AlexNet pre-trained using natural images from the ImageNet database was retrained using training datasets, including CT images of patients whose survival times (training data) were known. We used eight-bit CT images and survival times as training data for transfer learning. The DL model was trained using the Adam algorithm with 100 epochs. Data augmentation including rotation, flipping, translation, scaling, and shearing, was applied during the training procedure.

The prognostic potentials of the TIFs were evaluated using statistically significant differences (p-values < 0.05, log-rank test) to compare two survival curves (Kaplan-Meier curves) of high- and low-risk patients, which were stratified with medians of radiomic scores (rad-scores) of signatures constructed by using an elastic-net-regularized Cox proportional hazard model (CPHM). A rad-score was computed for each patient based on an inner product of constituent features in the signature and respective coefficients determined using the CPHM.

3. Results and Discussion

Figure 2 depicts the Kaplan–Meier curves for the training and validation datasets, where patients were stratified using the medians of the rad-scores or survival times for the training dataset. For the training dataset, the p-values [hazard ratio (HR), confidence interval (CI)] between the two survival curves for the TIF and the DL were 6.7×10^{-6} (HR: 0.41, 95% CI: 0.26–0.65) and 1.1×10^{-3} (HR: 0.51, 95% CI: 0.33–0.78), respectively. For the validation dataset, the p-values (HRs, CIs) between the two survival curves for the TIF and the DL were 3.4×10^{-5} (HR: 0.32, 95% CI: 0.16–0.62)and 1.2×10^{-1} (HR: 0.63, 95% CI: 0.35–1.1).

In this study, two-dimensional computation procedure was applied for the creation of the Betti maps. However, the tumor traits associated with prognosis may reside throughout the volume of a tumor [9]. We will work on the three-dimensional computation of TIFs in the future work.

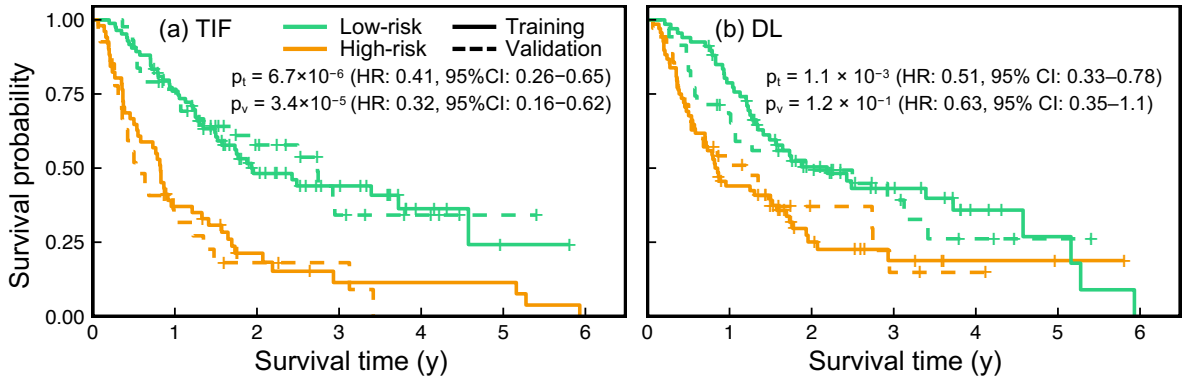


Figure 2. Kaplan–Meier curves for (a) the topologically invariant radiomic features (TIF) based on the Betti numbers and (b) the deep learning (DL) in the training and validation datasets. p_t and p_v indicates the p-values from training and validation procedures, respectively.

4. Conclusions

We have developed a novel topologically invariant radiomic features based on Betti numbers for prognostic prediction in lung cancer patients. The proposed approach showed a promising potential to the prognostic prediction.

Competing interests

The authors do not have any conflict of interests to disclose.

Acknowledgement

This study was partially supported by a grant from Center for Clinical and Translational Research of Kyushu University Hospital and JSPS KAKENHI Grant Number 20K08084. The authors would like to express their appreciation to all of the members of the Arimura laboratory who provided precious comments and advices for this study.

References

1. Lambin P, Rios-Velazquez E, Leijenaar R, et al (2012) Radiomics: Extracting more information from medical images using advanced feature analysis. Eur J Cancer 48:441–446. https://doi.org/10.1016/j.ejca.2011.11.036
2. Aerts HJWL, Velazquez ER, Leijenaar RTH, et al (2014) Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun. https://doi.org/10.1038/ncomms5006
3. Arimura H, Soufi M, Kamezawa H, et al (2018) Radiomics with artificial intelligence for precision medicine in radiation therapy. J Radiat Res 1–8. https://doi.org/10.11821/dlxb201802008
4. Hosny A, Parmar C, Coroller TP, et al (2018) Deep learning for lung cancer prognostication: A retrospective multi-cohort radiomics study. PLOS Med 15:e1002711. https://doi.org/10.1371/journal.pmed.1002711
5. Adcock A, Rubin D, Carlsson G (2014) Classification of hepatic lesions using the matching metric. Comput Vis Image Underst 121:36–42. https://doi.org/10.1016/j.cviu.2013.10.014
6. Nakane K, Takiyama A, Mori S, Matsuura N (2015) Homology-based method for detecting regions of interest in colonic digital images. Diagn Pathol 10:1–5. https://doi.org/10.1186/s13000-015-0244-x
7. Oyama A, Hiraoka Y, Obayashi I, et al (2019) Hepatic tumor classification using texture and topology analysis of non-contrast-enhanced three-dimensional T1-weighted MR images with a radiomics approach. Sci Rep 9:8764. https://doi.org/10.1038/s41598-019-45283-z
8. Kadoya N, Tanaka S, Kajikawa T, et al (2020) Homology-based radiomic features for prediction of the prognosis of lung cancer based on CT-based radiomics. Med Phys. https://doi.org/10.1002/mp.14104
9. Gerlinger M, Rowan AJ, Horswell S, et al (2012) Intratumor Heterogeneity and Branched Evolution Revealed by Multiregion Sequencing. N Engl J Med 366:883–892. https://doi.org/10.1016/j.eururo.2013.04.025

マルチチャンネル化と CNN を用いた嚥下時 X 線動画からの頸椎椎間板の抽出：LeNet, AlexNet, GoogLeNet の性能比較

藤中　彩乃[*1, *2]　目片　幸二郎[*3]　滝沢　穂高[*1]　工藤　博幸[*1]

要旨

本報告では，マルチチャンネル化 (MC) と CNN を用いて，嚥下時 X 線動画 (videofluorography，VF) から椎間板を抽出することを目指す．濃淡画像である VF の各フレームに 21 種類の画像フィルタをそれぞれ適用し，同数の特徴画像を生成する．3 種類を選択し，カラー画像の RGB チャンネルに保存することでマルチチャンネル化画像 (MC 画像) を生成する．MC 画像にパッチベースの CNN を適用し，画素ベースの F 値で抽出精度を評価する．焼きなまし法を用いて抽出精度が最大となる 3 種類の特徴画像の組合せを求める．本報告では CNN として LeNet，AlexNet，GoogLeNet を用いて性能を比較した．3 つの CNN において，RGB チャンネルに同一の特徴画像を保存した際の最適な特徴画像は原画像であり，F 値は順に 0.562, 0.604, 0.598 となった．MC の最適な組合せは 3 つの CNN で異なり，F 値は順に 0.593，0.634，0.598 となった．

キーワード：嚥下時 X 線動画，椎間板，マルチチャンネル化，CNN，焼きなまし法

1. はじめに

嚥下は軟口蓋，咽頭，喉頭，喉頭蓋などあらゆる頸部構造物や筋肉，支配神経の協調によって成り立ち，そのステージは認知期，口腔準備期，口腔期，咽頭期，食道期に分けられる．そのいずれが障害されても嚥下障害を来たしうるが，嚥下障害の原因は大まかに，口腔癌，前縦靭帯骨化症，変形性脊椎症などの静的障害 (器質的障害) と，脳血管障害，パーキンソン病，強皮症などの動的障害 (機能的障害) に分けられる．嚥下障害では流涎，誤嚥，湿性嗄声，体重減少などあらゆる症状を引き起こし [1]，患者の生活の質や生命予後に影響を与える．したがって，嚥下障害の早期発見，再発予防が必要となる．

*1 筑波大学
〔〒305-8573 つくば市天王台 1-1-1〕
*2 日本学術振興会特別研究員
*3 神戸赤十字病院
投稿受付：　年　月　日

本報告では，頸椎疾患に焦点を当てて嚥下のメカニズムの解明を目指す．正常嚥下群と頸椎固定術術前患者群の実験参加者の嚥下時 X 線動画 (videofluorography, VF) を用いて，マルチチャンネル化 (MC) [2]，畳み込みニューラルネットワーク (CNN), 焼きなまし法に基づく特徴選択によって頸椎椎間板領域を抽出する．3 つの CNN を用いて性能を比較した．

2. 方法

VF の各フレームは 8 bit の濃淡画像である．各フレーム画像に，Sobel filter，Gaussian filter，ヒストグラム平坦化など N 種類の画像フィルタを別個に適用し，N 種類の特徴画像を生成する．重複を許して 3 種類を選択し，マルチチャンネル化画像 (MC 画像) と呼ばれるカラー画像の red (R), green (G), blue (B) チャンネルにそれぞれ保存する．MC 画像に対してパッチベースの CNN を適用し，画素単位の F 値を求める．3 種類の特徴画像の組合せを変化させ，最適な組合せを探索する．

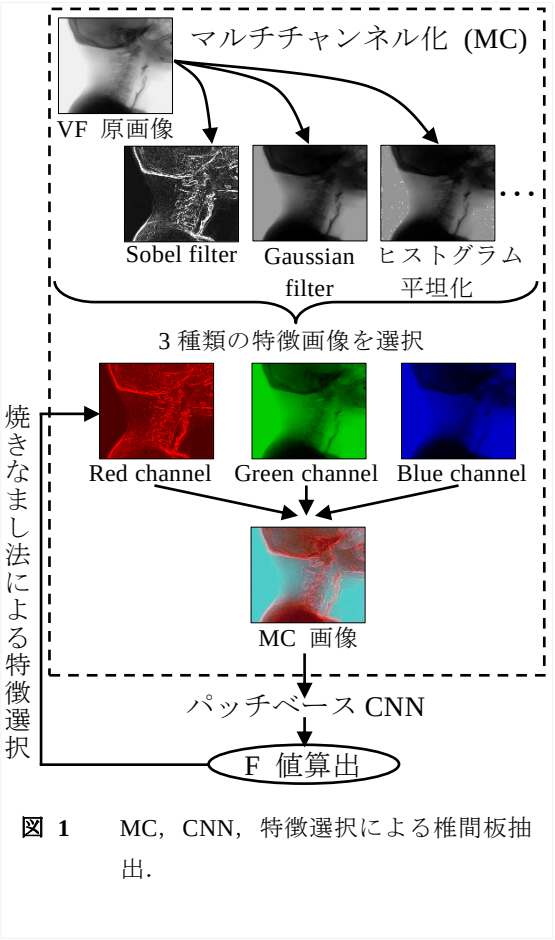


図 1 MC，CNN，特徴選択による椎間板抽出．

医用画像処理における特徴選択の手法として [3, 4, 5] などが報告されているが，本手法では焼きなまし法を用いる．手法の流れを図 1 に示す．

特徴画像を f_i (ただし $i = 1, 2, \ldots, N$) と定義し，RGB チャンネルに保存する特徴画像の組合せを (f_i, f_j, f_k) と定義する．1 – F 値で表されるエネルギー関数を $e(f_i, f_j, f_k)$ (ただし $0 \leqq e(f_i, f_j, f_k) \leqq 1$) と定義する．MC 画像の RGB チャンネルに同一の特徴画像を保存し CNN を適用する実験を全ての特徴画像に対して行い，F 値の降順に順位付けをする．(f_1, f_1, f_1) を焼きなましの初期状態とし，$e(f_1, f_1, f_1)$ を計算した後，RGB チャンネルに保存する特徴画像のうち一枚をランダムに選択し，[-step, +step] の範囲で動かし，$e'(f_{i'}, f_{j'}, f_{k'})$ を計算する．遷移確率 p を (1) 式で求める．

$$p = \exp\left(\frac{e-e'}{T}\right) \qquad (1)$$

e’ ＜ e または p が 0 以上 1 未満の乱数より大きい場合，変更後の特徴画像を保存する．これを温度 T が収束条件を満たすまで繰り返し，最もエネルギー関数の小さい特徴画像の組合せを最適解とする．

3. 実験

本報告では，患者群 19 症例と正常群 39 症例の合計 58 症例の VF を用いて実験を行った．患者群 13 症例と正常群 26 症例を学習症例に割り当て，患者群 6 症例と正常群 13 症例をテスト症例に割り当てた．使用した画像フィルタの数は $N = 21$ 種類である．画像フィルタの一覧と英数字 3 文字の略称を表 1 に示す．LeNet，AlexNet，GoogLeNet の 3 つの CNN を用いて実験し，性能を比較した．焼きなましの初期温度を T = 10000 とし，冷却係数 cool を CNN の学習時間の関係で順に 0.99，0.9，0.8 とし，終了条件を T ≦ 0.0001 とした．Step = 2 とした．なお，本研究は神戸赤十字病院の倫理委員会の承認を受けて行われた．

表 1 使用した画像フィルタとその略称．

画像フィルタ	略称
原画像	GRY
平均化フィルタ	MEA
ガウシアンフィルタ	GAU
メディアンフィルタ	MED
K 最近傍平滑化フィルタ	KNN
濃度正規化	NML
ヒストグラム平均化	HGE
X 方向のソーベルフィルタ	SBX
Y 方向のソーベルフィルタ	SBY
勾配方向のソーベルフィルタ	SBM
角度方向のソーベルフィルタ	SBD
バイラテラルフィルタ	BLT
ラプラシアンフィルタ	LPL
モルフォロジー演算の Top-hat フィルタ	TOP
指数対数型トーンカーブ	ETC
S 字トーンカーブ	STC
ソラリゼーション	SOL
エンボス	EMB

Local binary pattern 処理	LBP
ポスタライゼーション	POS
全画素値を 0 にするフィルタ	ZSP

4. 結果

RGB チャンネルに同一の特徴画像を保存して F 値を比較した際, 3 種類の CNN 全てにおいて最適な特徴画像は原画像であり，F 値は順に 0.562，0.604，0.598 となった．MC の最適化実験において, 3 種類の CNN における最適な特徴画像の組合せはそれぞれ (SBM, TOP, SBM)，(SOL，GRY，SBY)，(STC，BLT，GRY) であり, F 値は順に 0.593, 0.634, 0.598 となった．

図 2 に正常症例の原画像と，原画像に 3 つの CNN を適用した結果, 3 つの CNN に MC を導入した結果を示す．MC の特徴画像の組合せとして，それぞれの CNN で求まった最適解を用いる. 緑色画素が TP, 赤色画素が FP, 青色画素が FN を表している. F 値はそれぞれ，0.582，0.630，0.609，0.639，0.665，0.562 であった．

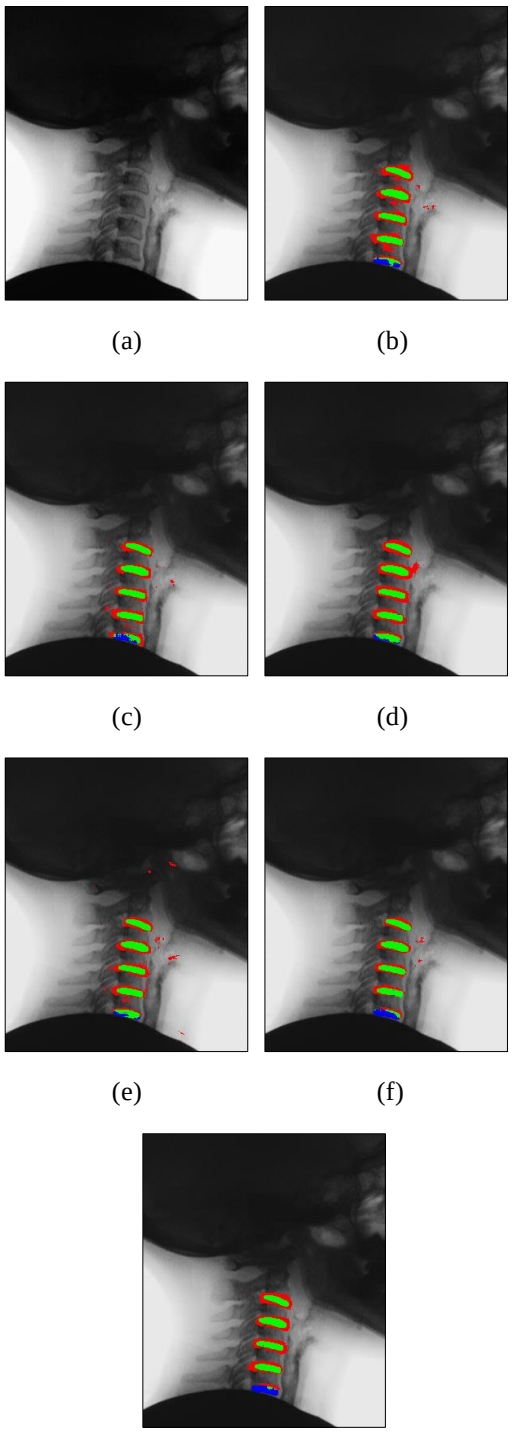
図 3 に患者症例の結果画像を示す．F 値はそれぞれ，0.390，0.423，0.408，0.445，0.341，0.510 であった．

5. 考察

前処理として MC を導入することで，原画像をそのまま入力した場合に比べて，精度が同等以上に向上することがわかった．また，非線形フィルタが使われた際に精度が向上することもわかった．GoogLeNet ではほぼ同等の精度になった．これは層数が多く MC の効果が出にくかったためか，実験に使用した CNN プラットフォームの仕様からパッチを 10.6 倍に拡大しなければならず，パッチ画像の局所的な特徴が失われたためと考えられる．今後は，U-Net などより多くの種類のネットワークを用いて MC の効果を検証する．

6. まとめ

本報告では，MC，CNN，特徴選択によっ



(a) (b) (c) (d) (e) (f) (g)

図 2 正常症例の結果画像．(a) 原画像．(b, c, d) 原画像に LeNet，AlexNet，GoogLeNet をそれぞれ適用した結果．(e, f, g) 最適な MC 画像に LeNet，AlexNet，GoogLeNet をそれぞれ適用した結果.

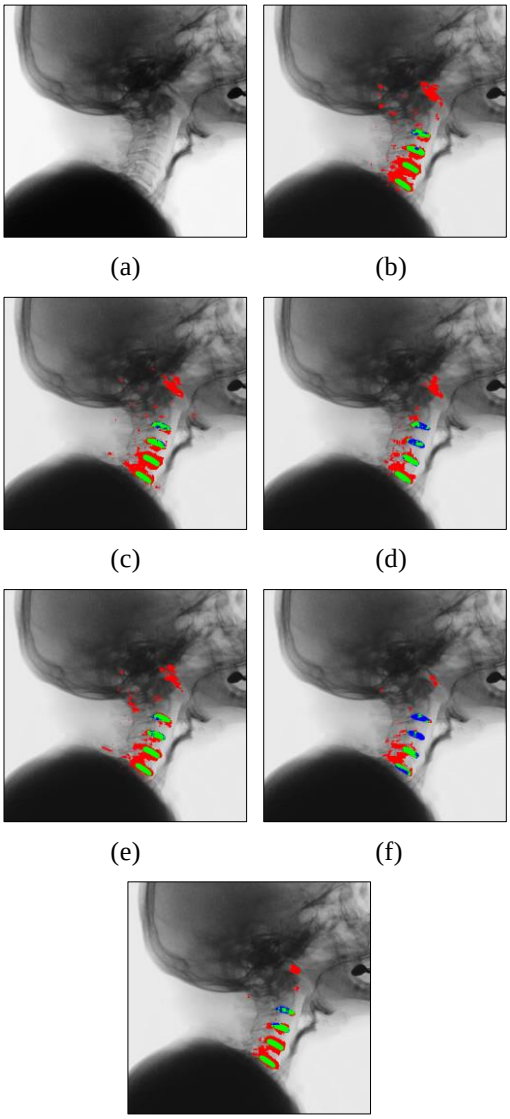


図 3　患者症例の結果画像．(a) 原画像．(b, c, d) 原画像に LeNet，AlexNet，GoogLeNet をそれぞれ適用した結果．(e, f, g) 最適な MC 画像に LeNet，AlexNet，GoogLeNet をそれぞれ適用した結果．

て VF から椎間板を抽出する手法を提案した．LeNet，AlexNet，GoogLeNet における画素単位の F 値を比較した．MC を導入することによって，LeNet と AlexNet では F 値が 3％ 程度上昇したが，GoogLeNet における同程度の精度であった．

謝辞

研究遂行にあたり貴重なご協力を賜った京都大学大学院医学研究科人間健康科学系専攻リハビリテーション科学コース 松林潤氏，岡山大学整形外科 瀧川朋亨氏，神戸赤十字病院整形外科 戸田一潔氏，神戸赤十字病院整形外科 伊藤康夫氏に深謝いたします．

本研究の一部は JSPS 科研費 20J10293 の助成を受けたものです．

利益相反の有無

なし

文　献

［1］ 谷口洋：嚥下障害，診られますか？．羊土社，東京，2015，pp13-35

［2］ Ayano F, Kojiro M, Hotaka T, et al.: Segmentation of cervical intervertebral disks in videofluorography by CNN, multi-channelization and feature selection. International Journal of Computer Assisted Radiology and Surgery **2020**: 1-8, 2020

［3］ Amin K, Abbas K, Thanh N, et al.: Medical image analysis using wavelet transform and deep belief networks. Expert Systems with Applications **86**: 190-198, 2017

［4］ Jothi G, Hannah I H.: Hybrid Tolerance Rough Set–Firefly based supervised feature selection for MRI brain tumor image classification. Applied Soft Computing **46**: 639-651, 2016

［5］ 根本充貴，清水昭伸，萩原義裕 他：多数の特徴量からの特徴選択による乳房 X 線像上の腫瘤影判別精度の改善と高速な特徴選択法の提案. 電子情報通信学会論文誌 **D 88.2**：416-426, 2005

Segmentation of cervical intervertebral disks in videofluorography with multi channelization and CNN: comparison of LeNet, AlexNet and GoogLeNet

Ayano Fujinaka[*1, *2], Kojiro Mekata[*3], Hotaka Takizawa[*1], Hiroyuki Kudo[*1]

*1 University of Tsukuba

*2 Research Fellowship for Young Scientists, Japan Society for the Promotion of Science

*3 Kobe Red Cross Hospital

In this study, cervical intervertebral disks (CIDs) are segmented from videofluorography (VF) by use of convolutional neural network (CNN) and our *multi-channelization* (MC) method. The frame images of VF are grayscale images. Twenty-one kinds of image filters are individually applied to each frame image to obtain the same number of *feature images*. Three feature images are selected and are input into the RGB channels of a color image, which is called a *multi-channelized (MC) image*. Patch-based CNN is applied to the MC images and pixel-based F-measures are calculated to evaluate the segmentation accuracy. The optimal combination of the three feature images is obtained by the simulated annealing method with pixel-based F-measures. LeNet, AlexNet, and GoogLeNet are used as CNN and their segmentation accuracy are compared with one another. When input images for CNNs are composed of the same feature images, the best F-measures are 0.562, 0.604 and 0.598, respectively. On the other hand, when input images are generated by MC, the best F-measures are 0.593, 0.634 and 0.598, respectively.

Key words: Videofluorography, Cervical intervertebral disks, Multi-channelization, CNN, Simulated annealing

Automated whole-heart segmentation using U-Net in pediatric cardiac CT

Akifumi YOSHIDA[*1,2], Yongbum LEE[*2], Norihiko YOSHIMURA[*3], Tatsuya KURAMOTO[*4], Akira HASEGAWA[*1, 5], Tsutomu KANAZAWA[*4]

Abstract

This study investigated the feasibility of pediatric whole heart segmentation using U-Net. The contrast-enhanced cardiovascular CT images (N=20, age: 2 months to 13 years old, average 5.3 years old) were used for cases in which no morphological abnormality was noted in the heart. The cardiac regions including the atria, ventricles, great vessels (superior vena cava, pulmonary artery, ascending aorta), and soft tissues including myocardium were manually annotated and converted into PNG format with 256x256 of matrix size, and inputted to U-Net. Dice coefficients were obtained by the 10-fold cross-validation method for the investigation of overall segmentation accuracy. Moreover, Dice coefficients for each age category (< 1 year, 1 - < 5 years, 5 - < 10 years, and 10 - < 15 years) were also evaluated. In order to obtain the Dice coefficient for each age group, we calculated the Dice coefficients for each of the 20 cases using the leave-one-out method and obtained the average Dice coefficients. The average Dice coefficient obtained by the 10-fold cross-validation method was 0.935. Analysis of variance showed no significant differences in the Dice coefficient between age categories. Deep learning like U-Net would become a useful elemental technology in heart segmentation of pediatric cardiology.

Keywords：pediatric cardiology, heart segmentation, computed tomography, deep learning

1. Purpose

Complex cardiac malformations are extremely diverse. In emergency cardiac surgery, it is necessary to understand the complicated three-dimensional (3D) cardiac structure quickly and accurately before the operation. Tomographic

*1 Department of Radiological Technology, Faculty of Medical Technology, Niigata University of Health and Welfare〔Shimamichou 1398, Kita-ku, Niigata, 950-3198, Japan〕

e-mail: 44da.af@gmail.com

*2 Department of Radiological Technology, Graduate school of health science, Niigata University.

*3 Department of Radiology and Radiation Oncology, Graduate School of Medical and Dental Sciences, Niigata University

*4 Department of Radiological Technology, Niigata University Medical and Dental Hospital

*5 Graduate school of Medical Sciences, Kanazawa University

Receive: August 5, 2020

images such as contrast-enhanced computed tomography (CT) provide a lot of precious information about cardiac morphology. However, a large quantity of CT images makes the understanding be difficult. To overcome this issue, 3D volume rendering images of whole heart are often generated and observed. In an emergency, the 3D images need to be generated in a short time.

We focused on U-Net, which is a deep learning method that is expected to automatically segment the anatomical structure as an elemental technology to support the rapid diagnosis of cardiac malformations. Although there have been reports on whole heart segmentation of adult cardiac system using deep learning [1, 2], there are few reports that apply deep learning to regional segmentation of the pediatric cardiac system. In this study, we investigated the feasibility of pediatric whole heart segmentation using U-Net.

2. Method

This study was approved by the Research Ethics Committee of Niigata University (No. 2018-0098). Approximately 7,700 contrast-enhanced cardiovascular CT images (N=20, age: 2 months to 13 years old, average 5.3 years old) were collected for this study. These CT images had been taken at Niigata University Medical and Dental Hospital between 2013 and 2018, and were cases in which no morphological abnormality by radiologist was noted in the heart. Ten cases were scanned without electro-cardiogram (ECG)-gating technique and only a single cardiac phase CT data was used. The remaining 10 cases scanned using ECG-gating technique and had diastolic and systolic CT data. The window setting of original CT images were mediastinum condition (WL/WW=50/400HU). The cardiac regions including the atria, ventricles, great vessels (superior vena cava, pulmonary artery, ascending aorta), and soft tissues such as myocardium were manually annotated using an image processing workstation and output as DICOM data. The DICOM data were converted binarized images and used as teacher images. The teacher images and the original CT images were converted into PNG format with 256x256 of matrix size and inputted to U-Net. The U-Net was implemented in TensorFlow, using an NVIDIA GeForceRTX 2060SUPER GPU which has 8 GB memory. Dice coefficients were obtained by the 10-fold cross-validation method for the investigation of overall segmentation accuracy. Moreover, Dice coefficients for each age category (< 1 year, 1 - < 5 years, 5 - < 10 years, and 10 - < 15 years) were also evaluated. A general numerical computing software MATLAB R2020a (The Math Works, Inc.) was used for statistical analysis and performed one-way analysis of variance (ANOVA). Four age categories were set according to the National diagnostic reference levels in Japan (2020). The Dice coefficients for each of the 20 cases using the leave-one-out method were calculated and the average Dice coefficients for each age category were obtained.

3. Results

The average dice coefficient by the 10-fold cross-validation method was 0.935. Figure 1 shows the example of original images, teacher images, and output images (Dice coefficient: 0.957). The agreement between the teacher image used and the estimated image output by U-Net was qualitatively good. The mean dice coefficients for each age

category are shown in the Table 1. The shape of the pediatric heart is not simply a miniature of the adult heart, and the shape of the heart varies with age, even in normal cases. However, ANOVA showed that there were no significant differences in the Dice coefficients of whole heart segmentation at the any age categories ($F = 2.57$, $P > 0.09$). The results showed that U-Net provided good accuracy for the segmentation of pediatric whole heart.

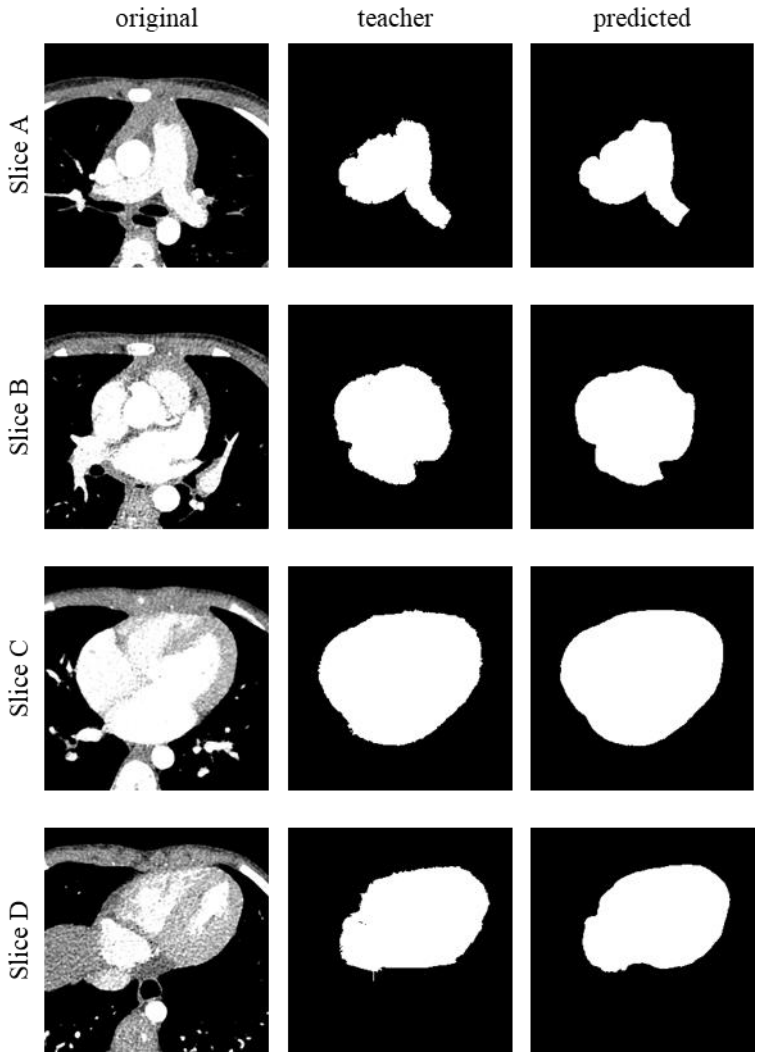


Fig. 1 Example of original images, teacher images, and output images.

Table 1 Comparing the accuracy (Dice coefficient) of the whole heart segmentation with U-Net in CT images for four pediatric age categories.

age category	Dice coefficient
0 - <1 y	0.952
1 - <5 y	0.948
5 - <10 y	0.960
10 - <15 y	0.875

4. Conclusions

The whole heart segmentation was performed from pediatric contrast-enhanced CT images using U-Net. Deep learning like U-Net would become a useful elemental technology in heart segmentation of pediatric cardiology.

Competing interests

The authors have no conflicts.

Acknowledgement

This paper was supported by a Grant-in-Aid for Research Activity Start-up and Challenging Research from the Niigata University of Health and Welfare, 2019.

References

［1］Xu Z,Wu Z, Feng J.: CFUN: combining faster R-CNN and U-net network for efficient whole heart segmentation. arxiv, 2018. abs/1812.04914. Available online at: http://arxiv.org/abs/1812.04914 (accessed August 2, 2020).

［2］Ye C, Wang W, Zhang S, Wang K. Multi-depth fusion network for whole-heart CT image segmentation. IEEE Access. (2019) 7:23421–9. doi: 10.1109/ACCESS.2019.2899635

動体追跡放射線治療における Deep Learning を用いた X 線透視画像上の非球形マーカ位置の認識

阿部　和貴[*1]　宮崎　康一[*2]　横川　航平[*2]　高尾　聖心[*3]

平田　雄一[*4]　梅垣　菊男[*2]　宮本　直樹[*2,3]

要旨

動体追跡放射線治療では形状が球形のマーカと非球形のマーカの両方が用いられるが，後者の場合は X 線の透視角度により透視画像上の形状が変化するため，透視角度毎にテンプレート画像を作成する必要がある．そこで本研究では，透視角度毎のテンプレート画像が不要な非球形マーカ位置の認識手法を開発した．初めに，CNN を用いて非球形マーカが写るマーカ画像と写らない背景画像の分類モデルを作成した．特に，本研究では学習に使えるマーカ画像数の制限上，X 線の減弱に基づいた画像合成を用いた画像水増し手法を取り入れた．次に，学習した CNN を用いて透視画像を走査しながら逐次マーカと背景を分類することでマーカ尤度マップを作成し，マーカ尤度の重心位置をマーカの認識位置とした．ファントム実験から取得した画像に対して認識精度を求めた結果，従来法であるテンプレートマッチングによる認識位置誤差は 4.9±3.1 pixel に対し，提案手法では 3.5±1.7 pixel であった．

キーワード：動体追跡放射線治療，畳み込みニューラルネットワーク，深層学習，マーカ検出

1. はじめに

肺や肝臓等の呼吸性移動を伴う腫瘍を治療する方法のひとつとして動体追跡放射線治療［1］がある．この治療方法は腫瘍近傍に留置した金属製のマーカを X 線透視によってリアルタイムに認識し，マーカの位置に応じて照射タイミングを制御する方法である．治療中のマーカ位置は，マーカ単体で X 線透視した画像をテンプレートとしたテンプレートマッチングにより認識する．

従来，動体追跡放射線治療において使用されるマーカには球形マーカと非球形マーカがある．非球形マーカは組織の留置位置から移動しにくいという利点がある一方で，X 線の透視角度によって透視画像上での形状が変化することから，透視角度毎にテンプレート画像が必要という欠点がある．また今後，非球形マーカの認識は 2 軸 4D-CBCT［2］や VMAT［3］といった新規技術でも必要となると考えられる．これらのケースでは撮像中，治療中に X 線の透視角度が連続的に変化することから，1 枚のテンプレート画像で非球形マーカを認識した場合は精度が低下すると考えられる．

*1 北海道大学大学院医理工学院
放射線医学物理学分野
〔〒060-8628 札幌市北区北 13 条西 8 丁目〕
e-mail: kazuki-abe@eis.hokudai.ac.jp
*2 北海道大学大学院工学研究院
応用量子科学部門
*3 北海道大学病院医学物理部
*4 北海道大学アイソトープ総合センター

この課題を解決するため，本研究では多数のテンプレート画像を必要とせず，透視角度の変化に依らない非球形マーカの認識手法を開発することを目的とした．この認識手法は，Deep Learning を用いた画像認識に，本研究独自の画像合成手法を組み合わせたものである．

2. 手法

2.1 認識手法の概要

非球形マーカ位置の認識手法の概要を図 1 に示す．

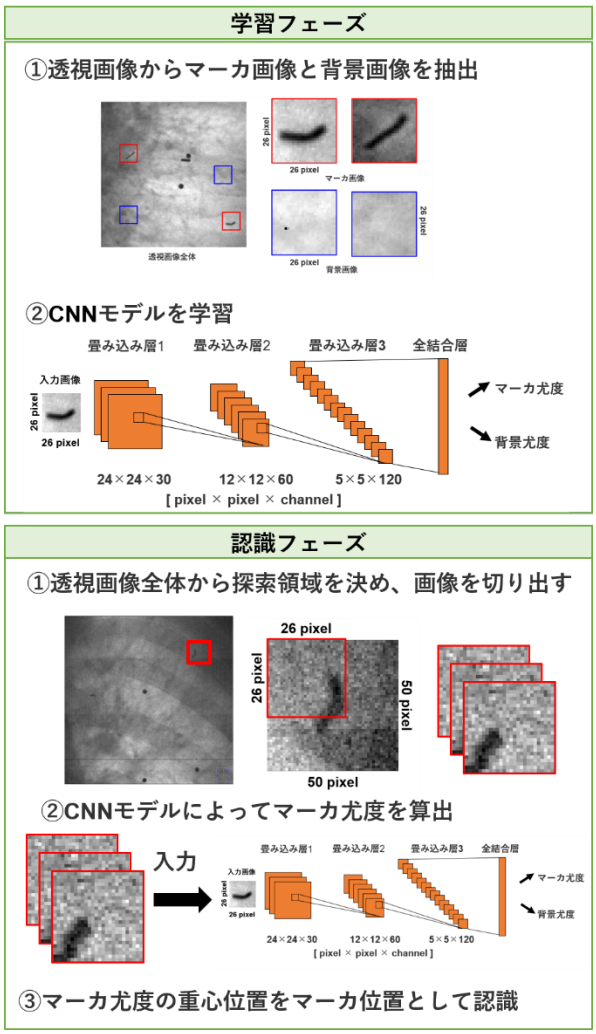


図 1 非球形マーカ位置の認識手法の概要

認識手法は学習フェーズと認識フェーズに分かれる．学習フェーズでは，X 線透視画像から非球形マーカが写ったマーカ画像と，非球形マーカが写っていない背景画像をそれぞれ取得する．その後，CNN によりマーカ画像と背景画像の分類モデルを作成する．

認識フェーズでは，X 線透視画像全体からマーカ近傍の領域を探索領域としてあらかじめ手入力し，その内部を走査しながら逐次的に画像を切り出す．それぞれの切り出した画像を分類モデルに入力し，マーカ尤度マップを算出する．図 2 にマーカ尤度マップの 1 例を示す．最終的に得られたマーカ尤度マップから重心を求め，マーカ位置として認識する．探索領域は 50×50 pixel とした．また逐次的に切り出す際の画像サイズは 26×26 pixel とした．

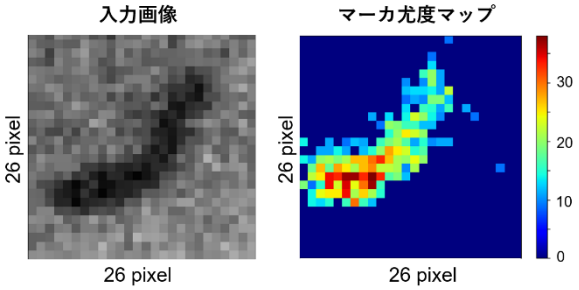


図 2 入力画像に対するマーカ尤度マップ

2.2 CNN モデルの学習・モデル

マーカ画像は他の一般画像認識で用いられる画像群と比べ単純な形状であることと，マーカ認識にかかる時間を短縮することから，本研究ではシンプルな CNN モデル［4］を用いた．3 層の畳み込み層と 2 層の Max pooling 層ののち，1 層の全結合層を設定した．26×26 pixel のグレー画像を入力とし，全結合層からマーカ尤度と背景尤度の 2 つの値を出力する．活性化関数には ReLU を用いた．学習はバッチサイズ 128 で最適化は Momentum SGD を用い、50 epoch 実施した．

2.3 学習画像の取得

臨床下で撮影されたマーカ画像を再現するために，本研究では人体ファントムを使った撮影実験により画像を取得した．

非球形マーカを埋め込んだ胸部ファントムを回転ステージ上に乗せ，ステージを連続的

に回転させながら 2 方向からの X 線透視撮影を実施した．撮影体系を図 3 に示す．

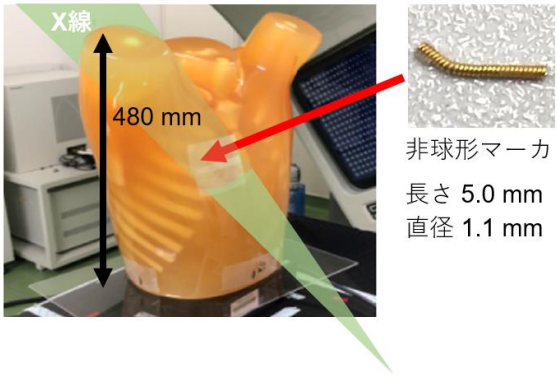


図 3 ファントム実験の撮影体系

この実験によって取得した透視画像から CNN モデルの学習に使用するマーカ画像と背景画像を抽出した．画像サイズは全て 26×26 pixel として用意した．画像のピクセルサイズは 0.39 mm/pixel である．

2.4 学習画像の水増し方法

実験から取得したマーカ画像が少なかったため，学習画像の水増し方法を検討した．

まずはマーカ画像と背景画像をそれぞれランダムな角度で回転させることで，画像パターンを増やした．また X 線撮影装置とマーカとの距離により透視画像上のマーカの大きさが変化することを再現するため，画像の拡大・縮小をランダムに行った．これらの処理により，マーカ画像と背景画像をそれぞれ 10 万枚作成した．

これらの画像は背景パターンが共通であるため，X 線の減弱に基づいたマーカ画像と背景画像の合成により，背景パターンを水増しした［5］．

非球形マーカを透過したマーカ画像の画素値I_1と胸部ファントムだけを透過した背景画像の画素値$I_2(x,y)$はそれぞれ式(1)，(2)により求まる．ただし，$I_0(x,y)$は物体がない場合の画素値，μ_Mとμ_fはそれぞれマーカと胸部ファントムの線減弱係数 [m^{-1}]を，T_MとT_fはそれぞれマーカと胸部ファントムの厚さ [m]を示す．

$$I_1(x,y) = I_0(x,y)\exp(-\mu_M T_M) \quad (1)$$

$$I_2(x,y) = I_0(x,y)\exp(-\mu_f T_f) \quad (2)$$

このように，マーカ画像と背景画像の画素値の違いは通過する物質の違いと，μとTの違いによって生じる．この現象を利用した画像合成の方法を式(3)と式(4)に示す．ただし，$S(x,y)$は合成画像の画素値，$I_M(x,y)$はマーカ画像の画素値，$I_{BG}(x,y)$は背景画像の画素値，$\overline{I_{BG}}$はマーカ画像の背景パターンの平均画素値，αは合成画像におけるマーカのコントラストを調整する係数である．本研究ではαに加算するβの値として 5 を用いた．

$$S(x,y) = \frac{I_M(x,y) \times I_{BG}(x,y)}{\overline{I_{BG}}} \times \alpha \quad (3)$$

$$\alpha = \frac{I_M(x,y)}{\overline{I_{BG}}} + \beta \quad (4)$$

この方法により，マーカ画像と背景画像の組み合わせを変えることで，様々なマーカのパターンを増やすことができる．

作成したデータセットに対して平均が 0，標準偏差が 1 となるように正規化を適用した．

2.5 評価方法

認識位置精度を評価するため，前立腺ファントムに非球形マーカを設置し，ファントムを回転しながら X 線透視画像を取得した．そして X 線透視画像上のマーカ位置を 300 フレーム認識し，認識位置と正解位置との 2 点間距離を評価した．マーカの正解位置はあらかじめ手動で決定したものを用いた．また比較として従来法であるテンプレートマッチングによる認識を実施した．テンプレート画像には初期フレームのマーカ画像を用いた．

3. 結果

CNN モデルを用いて非球形マーカを追跡した結果を表 1 に示す．

表 1 2 点間距離の結果
値は全て平均値±標準偏差として示す.

マーカ認識手法	CNN モデル	テンプレートマッチング
2 点間距離 [pixel]	3.5±1.7	4.9±3.1

提案手法を使用したマーカ認識手法では 2 点間距離は 3.5±1.7 pixel であったのに対し，従来手法であるテンプレートマッチングの場合は 2 点間距離 4.9±3.1 pixel であった.

次に，フレーム毎の 2 点間距離の推移を図 4 に示す．フレームが進む毎に画像上のマーカ形状が変化するため，テンプレートマッチングでは 2 点間距離が増加し，認識に失敗するフレームがあった．一方で，提案手法では最後まで認識することができた.

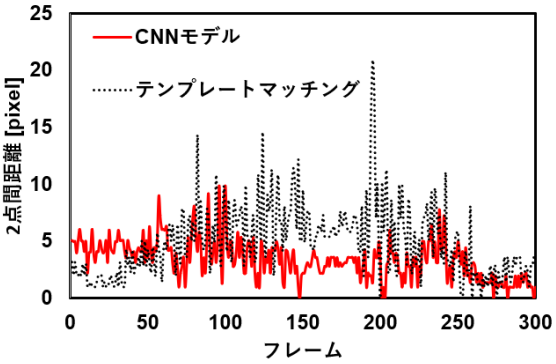


図 4 フレーム毎の 2 点間距離の推移

4. 考察

CNN モデルを用いて非球形マーカを認識した結果，2 点間距離は 3.5±1.7 pixel だったが，マーカが骨に重なったフレームでは 2 点間距離は大きくなった．今後，さらに認識位置精度を向上させるために，骨に重なったマーカ画像を学習させる必要があると考える．また，画像合成をする際のマーカ画像と背景画像の組み合わせや，αに加算するβを適切に選択することで，マーカ画像の背景パターンをさらに増やしていく必要があると考える.

5. まとめ

本研究は，元々少ないマーカ画像のパターンを増やすため，マーカ画像を画像合成によって作成したデータセットを用いて CNN モデルを作成した．また，CNN モデルを用いたマーカの認識位置精度をテンプレートマッチングの認識位置精度と比較した．300 フレーム分マーカを認識した結果，CNN モデルは X 線透視角度によって形状が変化する非球形マーカを 300 フレーム認識できた.

利益相反の有無

なし

文献

［1］ Shirato H, Shimizu S, Shimizu T, et al.: Real-time tumor-tracking radiotherapy *Lancet* 353 1331–2, 1999

［2］ Fujii T, Matsuura T, Takao S, et al.: Software Development for 4D-CBCT Research of Real-Time-Image Gated Spot Scanning Proton Therapy. Fifty-eighth annual meeting of the american association of physicists in medicine, Washington, DC, 2016, pp3389

［3］ Ruijiang L, Edward M, Bin H, et al.: Evaluation of the geometric accuracy of surrogate-based gated VMAT using intrafraction kilovoltage x-ray images Oncol Biol Phys. 72:996–1001,2008

［4］ Adam M, Paul J. K, Jeremy T.B, et al.: A deep learning framework for automatic detection of arbitrarily shaped fiducial markers in intrafraction fluoroscopic images Medical Physics 46:2291, 2019

［5］ Miyamoto N, Ishikawa M, Bengua G, et al.: Optimization of fluoroscopy parameters using pattern matching prediction in the real-time tumor tracking radiotherapy system Phys. Med. Biol 56 4803–4813, 2011

A deep learning based recognition method for arbitrarily shaped markers in fluoroscopic images

Kazuki Abe[*1], Koichi Miyazaki[*2], Kouhei Yokokawa[*2], Seishin Takao[*3]
Yuichi Hirata[*4], Kikuo Umegaki[*2], Naoki Miyamoto[*2,3]
*1 Graduate School of Biomedical Science and Engineering, Hokkaido University
*2 Division of Quantum Science and Engineering, Faculty of Engineering, Hokkaido University
*3 Department of Medical Physics, Hokkaido University hospital
*4 Central Institute of Isotope Science, Hokkaido University

In real-time tumor-tracking radiotherapy, spherical markers or non-spherical markers are used, but it is necessary to create template images for each fluoroscopy angle because the shape of the non-spherical marker changes depending on the X-ray fluoroscopy angle. In this study, we developed a method for recognizing non-spherical markers that does not require a template image for each angle of view. At first, we used CNN to create classification models of marker images (the marker is visible on fluoroscopic image) and background images (the marker is invisible on fluoroscopic image). We proposed a new image-augmentation-method based on X-ray attenuation information to synthesize marker images and background images because we cannot prepare a lot of marker images. Secondly, we applied our trained CNN model to generate a marker likelihood map by scanning a fluoroscopic image and sequentially separate marker and background. Next, we defined centroid position of marker likelihood map as marker recognition position. Finally, we conducted our proposal method and conventional template matching method to compare each recognition accuracy for images obtained by phantom experiment. As a result, the accuracy were 3.5±1.7 pixel and 4.9±3.1 pixel in our method and conventional template matching method, respectively.

Key words: Real-time tumor-tracking radiotherapy, Convolution Neural Network, deep learning, fiducial marker detection

著者紹介



阿部　和貴（あべ　かずき）
2019 年弘前大学医学部保健学科卒. 令和 3 年北海道大学大学院医理工学院博士前期課程修了予定. 現在，動体追跡放射線治療における Deep Learning を用いた金マーカ認識の研究に従事.
日本医用画像工学会会員

Automated segmentation of an intensity calibration phantom to quantify bone mineral density of the femur from clinical CT images

Keisuke UEMURA[*1, 2], Yoshito OTAKE[*1], Akihiro KAWASAKI[*1], Mazen SOUFI[*1], Masaki TAKAO[*3], Nobuhiko SUGANO[*2], Yoshinobu SATO[*1]

Abstract

Quantitative computed tomography (QCT) is often used to quantify bone mineral density (BMD) from CT images acquired in clinic. In QCT, an intensity calibration phantom with a known density needs to be included in the CT images to convert the CT numbers (Hounsfield Units: HU) into BMD. Usually, researchers manually select the region of the calibration phantom in the image, measure the HU within the phantom, generate a liner regression model between HU and density (mg/cm^3), and finally calculate the BMD at the region of interest. However, this process requires manual effort and is therefore time-consuming and is also prone to selection error. In this study, we developed a method that uses Bayesian U-Net to automatically segment the calibration phantom from CT images and create a regression model that converts the HU into BMD. Two-hundred and forty cases of CT images acquired with a calibration phantom at two institutes were analyzed herein. Manual segmentation was done for forty cases and was used as a training dataset for the Bayesian U-Net. These forty cases were later used for the assessment of accuracy by a 4-fold cross validation. After training, two-hundred cases that were not included in the training dataset were tested. Bayesian U-Net was able to successfully segment the calibration phantom and the correlation coefficients between HU and density (mg/cm^3) ranged between 0.9978 and 1 (all $p<0.001$), indicating an excellent correlation. In the 4-fold cross validation, Dice coefficient was 0.97 and the Average Symmetric surface Distance was 0.16mm, both indicating excellent accuracy. In conclusion, Bayesian U-Net applied in this study successfully established an automated system to segment a calibration phantom from clinical CT images. Using this system, HU can be automatically converted into BMD, which would be useful when analyzing QCT in a large cohort and in multicenter studies.

Keywords：bone mineral density, calibration phantom, deep neural network, regression model

---

*1 Imaging-based Computational Biomedicine Lab, Nara Institute of Science and Technology
〔8916-5 Takayama-cho, Ikoma, Nara 630-0192, Japan〕
e-mail: keisuke-uemura@is.naist.jp

*2 Department of Orthopaedic Medical Engineering, Osaka University Graduate School of Medicine

*3 Department of Orthopaedics, Osaka University Graduate School of Medicine

1. Purpose

Preoperative assessment of femoral anatomy and femoral bone mineral density (BMD) is important in the treatment of hip diseases. Often, three-dimensional anatomy of the femur is evaluated using preoperative CT images. On the other hand, for the evaluation of BMD, dual energy X-ray absorptiometry (DXA) is recommended by the World Health Organization (WHO) and is therefore utilized in clinic. However, DXA is not available at all hospitals and requires extra cost for examination. Further, analysis using DXA is two-dimensional and includes both cortical and cancellous bone that should be differentiated when understanding the effects of disease and treatment upon the femur [1].

To this end, previous studies have reported the usefulness of quantitative CT (QCT) when measuring the three-dimensional distribution of BMD from clinical CT images [1]. Generally, in QCT, an intensity calibration phantom that contains hydroxyapatite ($Ca_{10}(PO_4)_6(OH)_2$) or dipotassium hydrogenphosphate (K_2HPO_4) is included in the field of view to convert the CT numbers (Hounsfield Units: HU) to density (i.e. mg/cm^3). This conversion process is necessary to compare the results across patients as HU can be affected by several factors such as the type of CT scanners, imaging protocol (e.g. tube voltage, slice thickness), and reconstruction protocol (e.g. convolution kernel) [3, 4]. Usually, researchers manually set the regions of interest (ROI) to isolate the calibration phantom, measure the HU within the ROI, and adopt a linear regression model to convert the HU into BMD [2]. However, this process is time consuming and is prone to manual selection error. If HU can be automatically converted into BMD, it would aid researchers to compare the QCT data across subject groups and would be useful for analyses using a large database. Thus, we sought to use deep neural network to automatically extract the calibration phantom from CT images and to develop a system that automatically develops a regression model to convert the HU into BMD.

2. Method

CT images acquired at two hospitals (defined here as hospital A and B) for preoperative planning of hip surgery were used in this retrospective study that was approved by the Institutional Review Board. At both hospitals, CT images were acquired with an intensity calibration phantom (B-MAS200, FujiRebio, Tokyo, Japan) that is made from urethane foam and contains hydroxyapatite with a known density (0 mg/cm^3, 50 mg/cm^3, 100 mg/cm^3, 150 mg/cm^3, and 200 mg/cm^3) at specific regions (Fig. 1A). The CT manufacturer, model, and imaging protocol (slice thickness, kVP, and convolution kernel) in hospitals A and B were: General Electric, Optima CT660, 1.25mm, 120kV, and standard (soft tissue) for hospital A, and Toshiba, Activion16, 1mm or 2mm, 120kV, and FC30 (bone) for hospital B.

To create a training dataset that is used as an input for deep neural network, calibration phantom was manually segmented from the CT images for 40 cases (20 cases each for both hospitals) using Synapse Vincent (Fujifilm, Tokyo, Japan) (Fig. 1B). Due to the material property of the calibration phantom, the phantom can deform with patient's weight. Thus, while the same calibration phantom was used throughout the study, rigid registration was not employed. Instead, the Bayesian U-Net [5] was employed for segmentation in this study. Validation of the Bayesian U-Net was performed using an independent test dataset of 200 cases (100 cases for each hospital) (Fig. 1C-F). For post-processing, the segmented phantom was filled and eroded to exclude small isolated islands and to exclude the regions at the boundaries that can affect the measurement of the HU. To evaluate the outcome of the segmentation, a

linear regression model between the HU and density (mg/cm^3) was adopted for each test case. The variability of the regression slope and the correlation coefficients were calculated. Further, for the assessment of the accuracy, 4-fold cross validation was performed for the 40 cases that were manually segmented (i.e. training dataset). Similarities of the segmentation between manual and Bayesian U-Net were compared using the Dice coefficient and Average Symmetric surface Distance (ASD). Further, differences in the HU within the phantom was compared between manual and Bayesian U-Net segmentation.

3. Results

The mean correlation coefficient of the regression between HU and BMD was 0.9996 (range: 0.9978-1) for hospital A and was 0.9999 (range: 0.9992-1) for hospital B, both indicating excellent relationships. The mean ± standard deviation (SD) slope of the regression model was 0.843 ± 0.020 for hospital A and 0.743 ± 0.032 for hospital B. When BMD was calculated by altering the HU from -100 to 700, the SD of the BMD was small (<20) for both hospitals, which tended to expand as the HU increased (Fig. 2). In the 4-fold cross validation, mean Dice coefficient was 0.97 and the ASD was 0.16mm. The mean absolute difference in the HU within the phantom region was 0.23 between manual segmentation and the segmentation using the Bayesian U-Net.

4. Conclusion

Our method that utilizes Bayesian U-Net was able to accurately segment the calibration phantom from the CT images acquired at two different hospitals. As less than one HU difference is expected from the manual segmentation that have been performed, it seems that this automatic method can be applied to assess the BMD in a large CT dataset and be used to compare QCT data acquired in different hospitals. Our next step is to develop a system that automatically isolates the ROI such that the BMD at the ROI can be full-automatically quantified from QCT images and to use the system in multicenter QCT studies.

Competing interests

The authors declare noting to disclose.

Acknowledgement

This study was supported by the Japan Society for the Promotion of Science (JSPS) Grants-in-Aid for Scientific Research (KAKENHI) Numbers 19H01176 and 20H04550.

References

[1] Adams JE: Quantitative computed tomography. European journal of radiology 71(3): 415, 2009

[2] Uemura K, Takao M, Otake Y, et al: The distribution of bone mineral density in the femoral heads of unstable intertrochanteric fractures. J Orthop Surg (Hong Kong). 26(2):2309499018778325, 2018

[3] Eggermont F, Derikx LC, Free J, et al: Effect of different CT scanners and settings on femoral failure loads calculated by finite element models. Journal of orthopaedic research 36(8): 2288, 2018

[4] Giambini H, Dragomir-Daescu D, Huddleston PM, et al: The Effect of Quantitative Computed Tomography

Acquisition Protocols on Bone Mineral Density Estimation. Journal of biomechanical engineering 137(11): 114502, 2015

[5] Hiasa Y, Otake Y, Takao M, et al: Automated Muscle Segmentation from Clinical CT Using Bayesian U-Net for Personalized Musculoskeletal Modeling. IEEE transactions on medical imaging 39(4): 1030, 2020

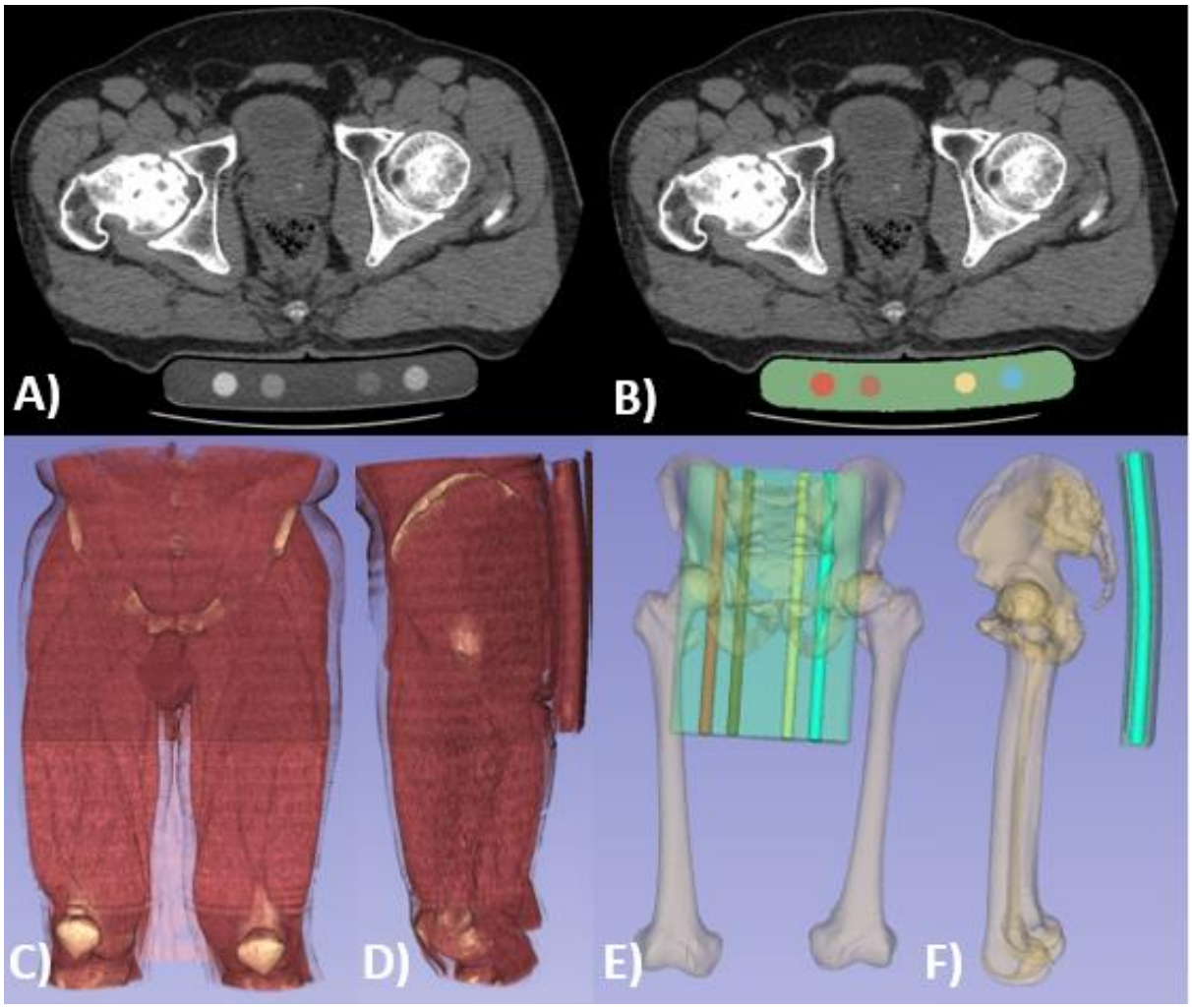


Fig. 1 An intensity calibration phantom included in the field of view of the CT images (A) was segmented manually (B). Using the Bayesian U-Net, phantom in CT images (C and D) were automatically segmented (E and F). Regions indicated in green, yellow, brown, cyan, and red in B), E), and F) indicate the regions of the calibration phantom that represent 0 mg/cm^3, 50 mg/cm^3, 100 mg/cm^3, 150 mg/cm^3, and 200 mg/cm^3, respectively. C) and D) represent the volume rendering of the CT images on the anteroposterior view (C) and the lateral-medial view (D).

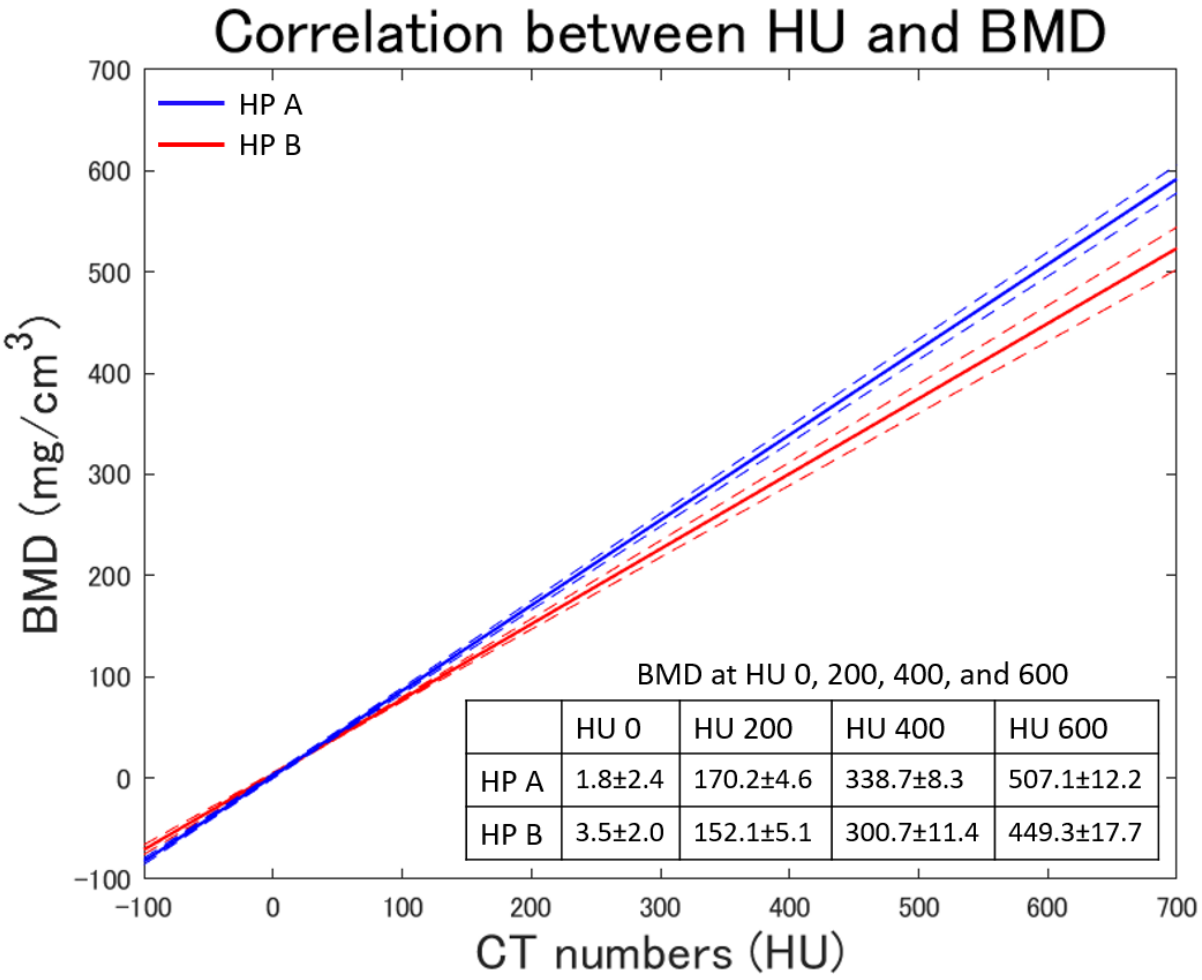


Fig. 2 Results of the regression model for hospitals A (blue) and B (red). Solid line indicates the mean of each model while the dotted lines indicate the ± standard deviations (SD). Small table inserted in the right lower corner indicates the mean (± SD) BMD calculated at each HU. For example, when the HU is 200, the BMD would be calculated as 170.2±4.6 mg/cm^3 in HP A and 152.2±5.1 in HP B. HP: hospital

Conversion of H&E Stained to EVG Stained Histological Images using CycleGAN

Tanwi BISWAS[*1], Hiroyuki SUZUKI[*2], Masahiro ISHIKAWA[*3], Naoki KOBAYASHI[*3], Takashi OBI[*4]

Abstract

This study involves generation of EVG stained pathological image from H&E stained image. Though H&E staining is the most commonly applied staining technique, it cannot identify elastic fiber because of its similar color and pattern to collagen fiber. In conventional pathology, to solve this problem, special staining technique such as Verhoeff's Van Gieson (EVG) staining is applied on the spatially consecutive pathological section which shows elastic and collagen fiber in different color thus makes them easily distinguishable. But process of EVG staining is very complex, costly and time consuming. This problem has been tried to be solved in this study by providing a CycleGAN based EVG stained image generation technique that reduces both time consumption and cost. CycleGAN is a deep learning framework which has been deployed and trained with unpaired images of training data for converting H&E stained image to its EVG stained counterpart. With the generation of digitally stained image, this study facilitates the pathologists to observe the effect of different stains on same tissue specimen. The quality of the generated image has been determined using SSIM, PSNR and MSE.

Keywords：pathological image analysis, CycleGAN, staining conversion, image to image translation

1. Introduction

Microscopic examination of tissue is a very crucial part of disease diagnosis and prognosis for understanding tissue morphology and function. For microscopic analysis, tissue section of very thin layer and high quality is needed to be prepared and then it is stained with coloring reagents which reveals complex tissue morphology to observer. Hematoxylin and Eosin (H&E) staining is widely used standard staining technique which is applied routinely on all types of tissue slides. H&E staining method stains nuclei in blue color and fibers and cytoplasm in pink color. Elastic fiber is a connective tissue fiber which is responsible for elasticity of the organ. The degradation of elastic fibers is one of the important histopathological features of disease analysis [1]. Different studies show that abnormality of elastic fiber in tissue is highly related to various kinds of diseases which are either inherited or acquired [2-3]. But, H&E stained image shows elastic fiber and collagen fiber with similar color (pink) and pattern. So, appearance of elastic

*1 Department of Information and Communications Engineering, Tokyo Institute of Technology〔4259-R2-60, Nagatsuta-cho, Midori-ku, Yokohama-shi, Kanagawa 226-8503, Japan〕
e-mail: biswas.t.aa@m.titech.ac.jp
*2 Center for Mathematics and Data Science, Gunma University.
*3 Faculty of Health and Medical Care, Saitama Medical University
*4 Institute of Innovative Research, Tokyo Institute of Technology

and collagen fiber cannot be distinguished using this H&E staining technique. That's why in conventional pathology, spatially consecutive tissue slide needs to be stained with special staining technique such as EVG. EVG stain shows elastic fiber in deep blue color and collagen fiber in orchid color and thus makes it easily distinguishable. This study provides a technique of computerized generation of EVG stained image from its H&E stained counterpart utilizing CycleGAN architecture. To observe the efficiency of the proposed method, H&E and EVG stained images of human pathological pancreatic Tissue MicoArrays (TMAs) have been used in this research.

2. Material and Method

In this study, H&E stained images have been converted to EVG stained image leveraging CycleGAN architecture [4] which has been trained with unpaired images of training data consisting of images from two different domains: H&E stained and EVG stained. Unpaired means images from H&E stained and EVG stained are not exactly aligned and scale adjusted. The model consists of two generators ($G_{H\text{-}E}$ and $G_{E\text{-}H}$) and two discriminators (D_{HE} and D_{EVG}). Generators are based on U-Net architecture and discriminators are based on PatchGAN architecture. The functionality of CycleGAN can be described by completing a cycle: $G_{H\text{-}E}$ takes H&E stained image as input and converts it to EVG stained image and $G_{E\text{-}H}$ takes this generated EVG stained image as input and converts it back to original H&E stained domain. The cycle repeats in the same way while starting with EVG stained image as first input. The loss between the input image and reconstructed image is cycle consistency loss. CycleGAN network has been trained considering 3 types of losses: Adversarial loss (eqn 1 and 2, $G_{E\text{-}H}$ and D_{HE} can be calculated similarly), cycle consistency loss and identity loss (eqn 3 and 4) have been calculated in term of mean absolute error. Adversarial loss has been calculated in term of mean squared error and others are calculated in term of mean absolute error.

$$\mathcal{L}_G(G_{H-E}, HE) = \mathbb{E}_{HE\sim p_{data}(HE)}\left[\left(D_{EVG}\left(G_{H-E}(HE)\right) - 1\right)^2\right] \tag{1}$$

$$\mathcal{L}_D(G_{H-E}, D_{EVG}, HE, EVG) = \mathbb{E}_{EVG\sim p_{data}(EVG)}[(D_{EVG}(EVG) - 1)^2] + \mathbb{E}_{HE\sim P_{data}(HE)}[D_{EVG}(G_{H-E}(HE))^2] \tag{2}$$

$$\mathcal{L}_{cycle}(G_{H-E}, G_{E-H}) = \mathbb{E}_{HE\sim p_{data}(HE)}\left[\left\|G_{E-H}\left(G_{H-E}(HE)\right) - HE\right\|\right] + \mathbb{E}_{EVG\sim p_{data}(EVG)}\left[\left\|G_{H-E}\left(G_{E-H}(EVG)\right) - EVG\right\|\right] \tag{3}$$

$$\mathcal{L}_{Identity}(G_{H-E}, G_{E-H}) = \mathbb{E}_{HE\sim p_{data}(HE)}[\|G_{E-H}(HE) - HE\|] + \mathbb{E}_{EVG\sim p_{data}(EVG)}[\|G_{H-E}(EVG) - EVG\|] \tag{4}$$

Image augmentation has been performed to prepare the training dataset. The training dataset contains images of H&E stained and EVG stained images of size 128 × 128 and consisting of total 1800 training images in each domain. The network has been trained using Adam optimizer with learning rate 0.0002 and number of image batch used is 6.

3. Result

Ground truth EVG stained images have been prepared by performing image registration between input H&E stained image and original EVG stained image. Fig. 1represents the result of this experiment where (a) input H&E stained image (b) Converted EVG stained image and (c) Ground truth EVG stained image. Table 1 represents the quality of the generated EVG stained images which have been shown in Fig. 1 in terms of SSIM, PSNR and MSE.

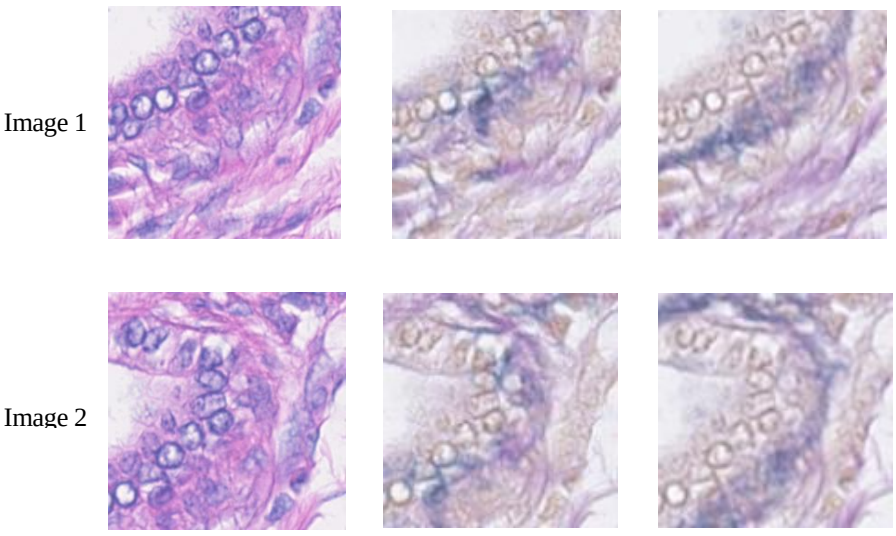


Fig. 1 Visual analysis of experimental result: (a) Input H&E (b) Generated EVG (c) Ground truth EVG stained image

Table 1 Quality of the generated EVG stained image.

Images	SSIM	PSNR	MSE
Image 1	0.6735	22.8470 dB	351.2346
Image 2	0.7510	23.2342 dB	330.5153

3. Conclusions

In this research, CycleGAN architecture has been deployed to convert H&E stained image to its EVG stained counterpart. Both visual quality and result of quality metrics of the generated image supports the acceptability of using the digitally generated EVG stained images for further pathological analysis. Generating EVG stained images from its same slide of H&E stained images facilitates the pathologists to observe the effect of different staining on the same tissue specimen.

Competing interests

In this work, the authors have no competing interest to declare.

References

［1］ Jing G, Qing-Song L, Zhen-Yong Z, et al.: The role of elastic fibers in pathogenesis of conjunctivochalasis. Int J Ophthalmol **10(9):** 1465–1473, 2017

［2］Hiroshi W: Role of Elastic Fibers on Cardiovascular Disease. Journal of Health Science **57:** 449-457, 2011

［3］Lakiotaki E, Sakellariou T, Evangelou K, et al.: Vascular and ductal elastotic change in pancreatic cancer, Acta Pathologica, Microbiologica et Immunologica Scandinavica. John Wiley and Son, **124:** 181–187, 2015

［4］Jun-Yan Z, Taesung P, Phillip I, et al. : Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks, ICCV, 2017

Unsupervised Classification of Dilated/Hypertrophic Cardiomyopathy in Pathological Images: Quantitative Analysis of Patch Diversity

Tianheng LIANG[1], Yoshito OTAKE[1], Kenji ONOUE[2],

Mazen SOUFI[1], Tomoya NAKANO[2], Yasuhiro SAKAGUCHI[2],

Kinta HATAKEYAMA[3], Yoshihiko SAITO[2], Yoshinobu SATO[*1]

Abstract

The diagnosis of dilated/hypertrophic cardiomyopathy (DCM/HCM) using pathological images requires expertise in the field but the number of expert doctors is limited. Also it is a time-consuming task, highly depends on pathologists' experience and knowledge. Several patch-based supervised learning approaches have been proposed to classify pathological images. However, given the patch diversity with respect to cellular contents, the assignment of a ground-truth label for each patch (i.e., patch-wise label), rather than a global (i.e. image-wise) label, is essential as it affects the training reliability. This project aims at leveraging unsupervised learning approaches to quantitatively assess the patch diversity with respect to disease classes. The proposed approach uses t-distributed stochastic neighbor embedding (t-SNE) and K-means clustering to assemble the patches into distinct clusters. In this study, 276 images from 96 cardiomyopathy patients were used. The images were split into patches with a matrix size of 256×256. Those patches including only the background were excluded, thus 7423 patches were used in the analysis. The patches were grouped into 300 clusters using t-SNE with a perplexity of 80. Four expert medical doctors were asked to classify the clusters into DCM, HCM or normal classes. The patch diversity was assessed by evaluating the agreement among the experts on the classification of each cluster. Results have shown that 36 clusters (12%) had agreement by all four experts, 130 clusters (43.3%) had agreement by three experts, and 64 (21.3%) clusters had agreement by two experts. These results indicate the intrinsic diversity in the patch classifications.

Keywords：Cardiomyopathy, Pathological image, Unsupervised learning

1. Introduction

Cardiomyopathy is a group of diseases that affect the heart muscle. Early on there may be few or no symptoms. As the disease worsens, shortness of breath, general fatigue, and edema of the legs may gradually occur, due to the progression of heart failure. Those affected are at an increased risk of sudden cardiac death [1]. There are many types of cardiomyopathy, the major two types are hypertrophic cardiomyopathy (HCM) and dilated cardiomyopathy (DCM).

---

*1 Nara Institute of Science and Technology〔8916 - 5 Takayamacho, Ikoma, Nara, Japan 630-0192〕

e-mail: Yoshinobu Sato：yoshi@is.naist.jp

2 Cardiovascular Medicine, Nara Medical University

3 Department of Pathology, National Cerebral and Cardiovascular Center

Clinically, HCM is characterized by an asymmetric thickening of the ventricular walls and a decrease in the ventricular chamber size. In addition, fibrosis of the myocardium and cardiomyocyte disarray are associated with the disease. Approximately two-thirds of HCM patients are known to develop primarily due to sarcomere-related gene mutations. DCM also results from single missense mutations in not only the same sarcomeric genes but also nuclear membrane, intercalated disk, sarcoplasmic reticulum genes, leads to ventricular chamber dilation and reduced systolic function, rather than hypertrophy. [2] Some patients with cardiomyopathy do not realize the illness, and sudden cardiac death could be the first sign of the symptom. Therefore, the early diagnosis is necessary. Compared to many diagnostic methods such as ECG, MRI or Echocardiography, the cardiac biopsy is the most invasive but the most accurate method, especially in classifying types of cardiomyopathy. However, training of pathologists requires a long-term education and clinical experience, and even for well-trained pathologists, it is difficult to classify types of cardiomyopathy. Thus, development of automated classification method using deep learning is considered to be a useful tool in supporting pathologists in their work and training. Such approaches have been tested on several pathologies, such as colon cancer [3]; however, up to our knowledge, the classification of cardiomyopathy in pathological images has not been addressed before.

Because of the diversity of annotations of the expert pathologists and the distribution of different types of local image features in one image, it is difficult to obtain a ground-truth label to perform an accurate supervised learning. In this paper, to obtain more accurate patch labels, we use an unsupervised learning method to classify patches from HCM, DCM and Normal class images, assess the patch diversity with respect to disease classes. In addition, we evaluated results from four experts' classification to analyze agreement/disagreement of the diagnosis by human experts.

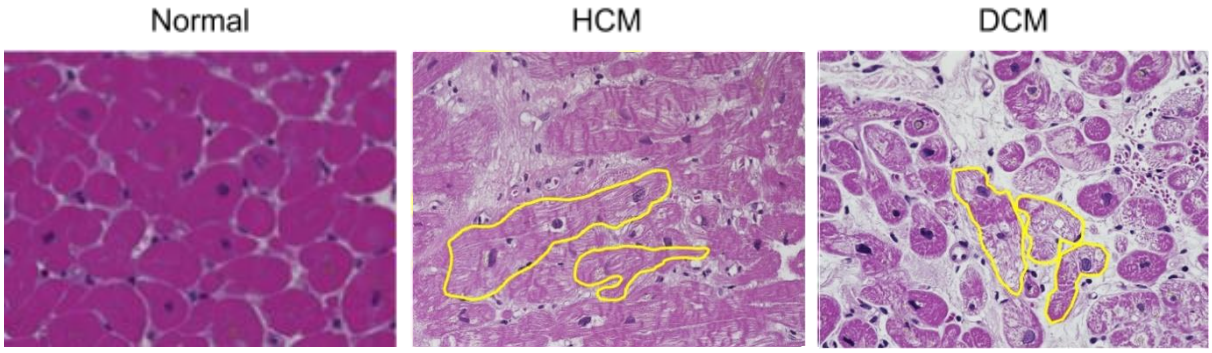


Figure 1: Representative pathological images of 2 major types of cardiomyopathy and normal cardiac cell. The parts that are especially distinct for each disease type are delineated by yellow lines. In HCM, each cardiomyocyte tends to be hypertrophied, the alignment is disordered and sometimes fibrosis grows between cells. In DCM, beside interstitial fibrosis, the density of the sarcomeric fiber is reduced in the cardiomyocyte.

2. Method

The data set used in this study consisted of 276 images from 96 cardiomyopathy patients, including 147 DCM cases, 99 HCM cases and 30 Normal cases. The resolution of the image was 1920×1440 pixels. Firstly, we split the images into patches, after several trials for testing performance with different patch sizes, we selected the size of 256×256 pixels,

resulting in 35 patches for each image.

Next, we excluded patches including only background (outside the tissue region). T-distributed stochastic neighbor embedding (t-SNE) [4] and k-means clustering were used to assemble the patches into distinct clusters. T-SNE is a method for dimensionality reduction which maps an image patch to a point in a 2-dimensional (2D) space. k-means clustering algorithm grouped nearest points into a cluster based on Euclidean distance in the 2D space. The number of clusters k was set to 300. The results are shown in Figure 2.

The remaining patches were re-processed by t-SNE and K-means clustering to generate 300 clusters. The clusters were sent to four expert pathologists for annotation. Experts were asked to label each cluster into either HCM, DCM or normal. Experts only looked at patches in each cluster independently from the surrounding patches in the original image. Finally, we compared the four experts' results and analyzed their agreements on each cluster and labeled patches on the images with their cluster annotation. In addition, we assessed the agreement in the annotation of the four experts by estimating a voting percentage of the patches in each image.

3. Results

In the first step, 9660 patches were obtained from 276 images. However, most of the image contains white space and boundaries, thus, only 7423 patches were selected for further processing and other patches were excluded from experiment. Some examples of excluding process are shown in Figure 2. Same method was then processed in next operation to generate 300 clusters.

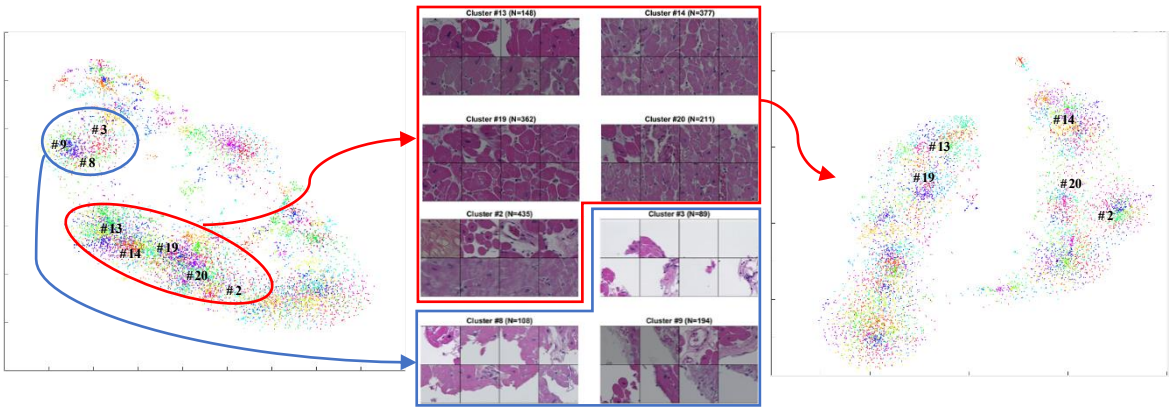


Figure 2: T-SNE maps and representative clusters. T-SNE map of 9660 patches was visualized on the right side, 9660 points marked with 300 different colors for different clusters, the representative cluster number were overlaid. Eight representative clusters were selected from left-side t-SNE map. Three 'white' patches' clusters were selected from top left corner of t-SNE map, 5 other clusters were randomly selected from the map. On the right side, t-SNE maps shows 7423 patches distributed into 300 clusters, in this map, 'white' patches were excluded.

Based on the four experts' annotation, only 36 clusters (12%) had agreement by the four experts, 130 clusters (43.3%) had agreement by 3 experts, and 64 (21.3%) clusters had agreement by 2 experts. Clustered patches are shown in Fig. 3.

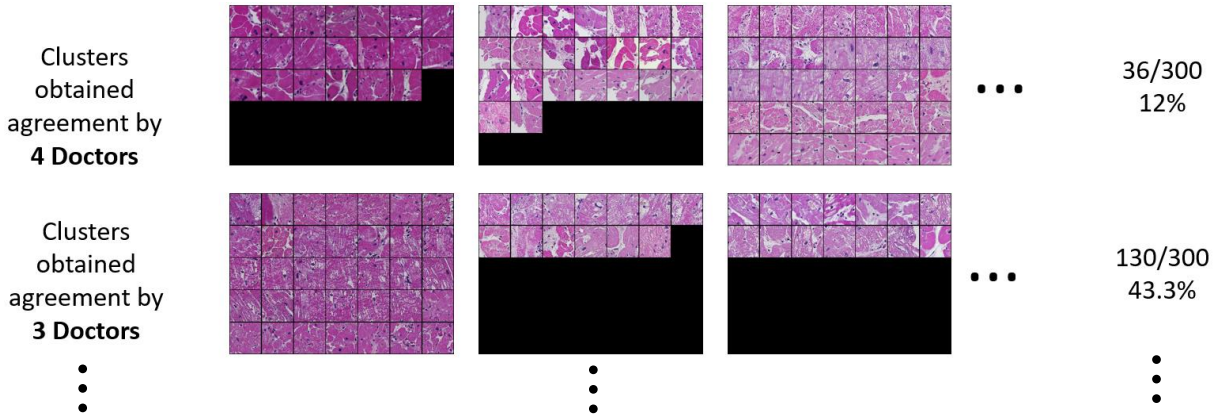


Figure 3: Cluster view of 300 patches. Six example clusters were shown in this figure, top 3 clusters obtained agreement by 4 experts, while the lower 3 clusters obtained agreement by 3 experts.

We transferred the annotation into image-wise view, thereby calculating the percentage of correctly labeled patches in a single image. As shown in Fig. 4, an example image was marked into different classes based on cluster annotation, the percentage of major class was then calculated to assign the image class. This calculation has been processed through all images.

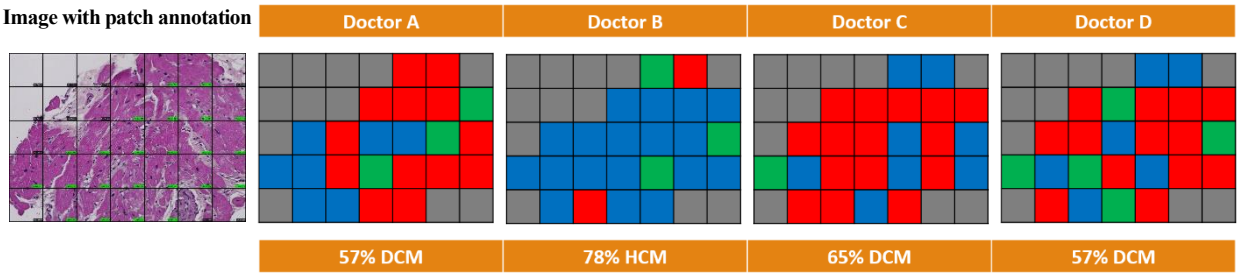


Figure 4: 4 Experts annotation on an example image. Left side is the original image after splitting into patches. Classes of each patch are shown on the right side, red - DCM, blue - HCM, green - Normal.

Beside image-wise comparison, visualizations of 3 classes in t-SNE map were also made. We obtained the most agreed annotation from each cluster, labeled those patches and plotted its location on the same t-SNE map. In this way, the patch diversity before and after cluster labeling are clearer to identify. The results were shown in Fig. 5.

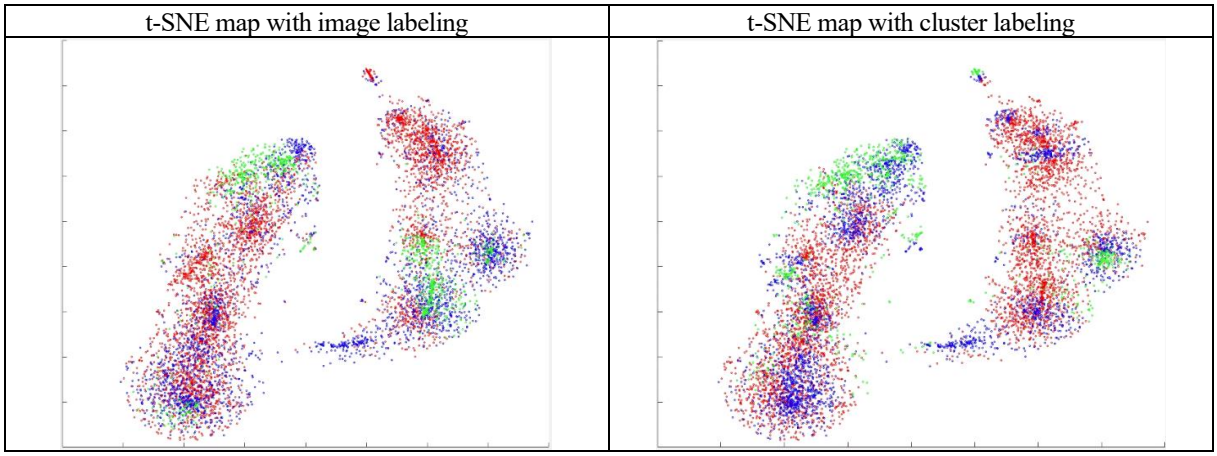


Figure 5: T-SNE map comparison of image labeling and cluster labeling, DCM (red), HCM (blue), Normal (green). 7423 patches distributed in t-SNE maps, left map shows patch labeling from original image, right side map shows labeling from cluster.

In addition, the results by the four experts were concluded into a confusion matrix (Table 1). The matrix concludes the annotation results in comparison with true image class. Furthermore, the percentages of accurate classifications from experts were also calculated.

Table 1: Confusion matrix of experts' annotation accuracy. Numbers in the matrix represent the images, percentage on the top left corner represent the overall accuracy of an expert.

	Doctor A				Doctor B				Doctor C				Doctor D			
True Class	57%	HCM	DCM	NOR	50%	HCM	DCM	NOR	52%	HCM	DCM	NOR	40%	HCM	DCM	NOR
	HCM	31	67	1	HCM	56	37	6	HCM	34	47	18	HCM	17	72	10
	DCM	28	119	0	DCM	62	75	10	DCM	36	104	7	DCM	43	81	23
	NOR	3	20	7	NOR	19	3	8	NOR	6	16	8	NOR	9	10	11

4. Discussion and Conclusion

T-SNE and K-means clustering have shown good performance for preprocessing (i.e. excluding white patches) as well as grouping clusters (i.e. elucidating the patch diversity). However, we found that the accuracy of four experts ranged from 40% to 57%. There were two possible reasons for this discrepancy, one is that the experts were labeling based on cluster-wise patches, which is different from their common clinical experience. In addition, some global features, which scale-up to the whole image cannot be identified in small patches. On the other side, some typical features may not only belong to one class, such as fibrosis that appear in both HCM and DCM cells, and this was also a difficult point to address by the pathologists.

By checking patches' distribution on t-SNE map in Fig. 5, we can recognize that the normal patches (green points) are more concentrated by image labeling method. This is because the normal class in our dataset is the smallest case, which could merge into many other clusters, thus increasing the difficulty of annotation by experts. In contrast, DCM (red points) and HCM (blue points) are more concentrated in cluster labeling, while in the image labeling, they tend to be mixed together, such as the reformation on top right corner of t-SNE map. These changes show that the features of different classes were more accurately separated.

The diversity of experts' agreements reflects to a certain extent the hardness of pathological image diagnostic on the patch level. In supervised deep learning-based approaches [5], building accurate training data set is a process of the most importance. This paper gives an idea that obtaining more accurate annotation from unsupervised cluster-wise views might assist in building such a dataset. In the future work, the integration of cluster-wise annotations into supervised learning schemes, with relevant comparisons, will be investigated.

Competing interests

The authors have no competing interests.

Acknowledgement

This research was supported by Japan Agency for Medical Research and Development (AMED) Number 20ek0109466h0001.

5. References

[1] Frenneaux, Michael P. "Assessing the risk of sudden cardiac death in a patient with hypertrophic cardiomyopathy." Heart (British Cardiac Society) vol. 90,5 (2004): 570-5. doi:10.1136/hrt.2003.020529

[2] Spudich, James A. "Hypertrophic and dilated cardiomyopathy: four decades of basic research on muscle lead to potential therapeutic approaches to these devastating genetic diseases." Biophysical journal vol. 106,6 (2014): 1236-49. doi:10.1016/j.bpj.2014.02.011

[3] Sari, Can Taylan, and Cigdem Gunduz-Demir. "Unsupervised feature extraction via deep learning for histopathological classification of colon tissue images." IEEE transactions on medical imaging 38.5 (2018): 1139-1149.

[4] Maaten, Laurens van der, and Geoffrey Hinton. "Visualizing data using t-SNE." Journal of machine learning research 9.Nov (2008): 2579-2605.

[5] Peikari, Mohammad, et al. "A cluster-then-label semi-supervised learning approach for pathology image classification." Scientific reports 8.1 (2018): 1-13.

Radiomics 画像特徴量による肺がん治療方針決定因子 PD-L1 発現予測の検討

松崎　和喜[*1]　金子　幸生[*1]　高橋　勲[*1]　松元　祐司[*2]

要旨

個別化医療の実現が求められる中，生検はがんの確定診断のみでなく，同時に取得される分子情報は治療方針や薬剤選択に必須となっている．また肺がん治療では免疫チェックポイント阻害剤を用いた新たな治療法が導入されている．この薬剤の効果予測には，programmed cell-death ligand 1(PD-L1)タンパクの発現率が用いられるが，判定の閾値や必要な検体量等に課題が残る．本研究では，生検による取得情報を放射線診断画像の特徴量へ変換学習する画像情報の拡張技術により，生検と同等の介入判断の実現を目的とした．CT 画像，PET 画像，および診療録の患者情報を組合せた Radiomics 特徴量による PD-L1 の発現予測モデルを構築・評価し，発現有無の予測精度は AUC が 0.89 であった．また構築モデルの特徴量を分析し，モデルの説明性を改善した．以上から画像情報による PD-L1 発現予測の可能性が示された．

キーワード：放射線画像，Radiomics，肺がん，治療効果予測，PD-L1

1. はじめに

医用画像診断領域では取得した画像から，より多くの情報を抽出し診断に利用する試みとして Radiomics[1]が注目されている．Radiomics は放射線画像の持つ多系統の情報をデータとして統合的に解析することで，より画像診断の効率と精度を高め，病理診断や分子・遺伝子情報等の臨床情報および他の検査情報を有効に利用した診断，治療方針の決定，予後予測，予後追跡等と結び付けようとする試みである．一方，肺がんは全がん種の中で世界的に最も高い死因である[2]. 従来, 肺がんと診断された場合，手術，放射線治療，化学療法から適切と考えられる方針が選択され治療が行われてきた．しかし近年，新たな免疫療法，特に免疫チェックポイントを標的とした治療法が確立され，その注目が増している. これらは Programmed cell death 1（PD-1）および PD-Ligand 1（PD-L1）に対する治療であり，治療に用いられる免疫チェックポイント阻害剤は，多くの種類が開発され，既に実臨床に用いられている．これら薬剤の効果予測および適応には PD-L1 の発現率が用いられることが多いが，様々な臨床試験で用いられた抗体やカットオフ値が異なり，抗体間での対比可能性については異論が残る. また PD-L1 発現判定では，取得した検体において一定の腫瘍細胞数を含む組織片が必要なことなど，解決すべき課題が少なくない．本研究では，放射線画像からの特徴量や，医師の診療録情報など，様々な情報を用いた Radiomics アプローチによる PD-L1 発現の予測機能を開発，放射線画像による PD-L1 発現予測の可能性を検討し，免疫チェックポイント阻害剤の治療方針決定支援を目的とする．

*1 （株）日立製作所　ヘルスケアビジネスユニット
〔〒185-8601 国分寺市東恋ヶ窪 1-280〕
e-mail: kazuki.matsuzaki.gv@hitachi.com
*2 国立がん研究センター中央病院 内視鏡科，呼吸器内科

2. 方法

1) 解析臨床データ

本研究の解析は，末梢肺病変に対して気管支鏡を施行し，原発性肺がんと診断された患者のうち 22C3 抗体を用いた免疫染色による PD-L1 発現（Tumor Proportion Score：TPS[%]）の結果が判明している患者 106 例を用いた．対象患者において放射線画像データとして造影 CT 画像および FDG-PET 画像，また患者および診療に関わる情報として診療録データを収集した．診療録データは，患者背景として年齢，性別，喫煙歴，画像所見として病変の境界（明瞭，不明瞭），病変の辺縁（不整，平滑），病変の性状（solid，part-solid），気管支透亮像（あり，なし），空洞（あり，なし），そして病理所見として PD-L1 発現（TPS）を用いた．本検討では TPS が 1%以上を発現あり，1%未満を発現なし，とした．

2) Radiomics 解析を用いた PD-L1 発現予測

放射線画像から Radiomics 特徴量を算出するため，画像内の病変領域抽出には，半自動領域抽出アルゴリズムを適用し，必要に応じて手動にて抽出領域を修正した．

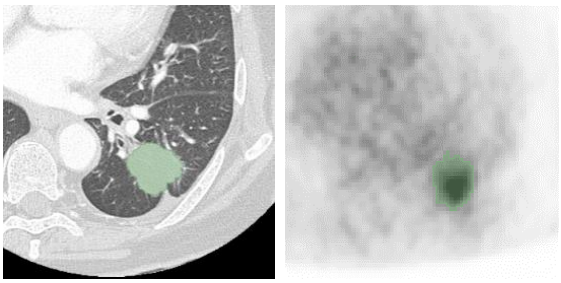


図 1 病変領域抽出例
（左：CT 画像，右：FDG-PET 画像）

また抽出した病変領域から Radiomics 画像特徴量として Shape, First Order, gray level size zone matrix (GLSZM)，gray level run-length matrix (GLRLM)，neighborhood gray tone difference matrix (NGTDM)，gray level co-occurrence matrix (GLCM)を算出した．これら画像特徴量と，診療録データから抽出した患者情報（年齢，性別，喫煙指数）の Radiomics 特徴量を説明変数，PD-L1 発現有無を目的変数とし，分類アルゴリズムとしてニューラルネットワークを用いた発現予測モデルを構築した．

最後に予測モデルの精度評価として，前述した CT 画像，PET 画像，および診療録データから成る臨床データ 106 症例のうち，86 症例をモデル学習用，20 症例をテスト用に分割して使用した．2 値分類でパフォーマンス計測に一般的な ROC（Receiver Operating Characteristic）曲線，および AUC（Area Under the Curve）を用いた．

3） 予測モデルの画像特徴量解析

Radiomics 解析に用いた画像特徴量と，診療録データから抽出した画像所見の関係を検討するため，説明変数を Radiomics 特徴量，目的変数を画像所見の各項目とした教師あり学習を実施し，画像所見の各項目を予測するモデルを構築し，関連する画像特徴量を抽出した．さらに，PD-L1 発現，および画像所見に関連する特徴量を解析し，共通する特徴量が存在するかを検討した．

3. 結果

1) PD-L1 発現予測モデルの精度評価

CT, PET の放射線画像から算出した画像特徴量と，診療録の患者情報（年齢，性別，喫煙指数）の各組合せを説明変数として構築した予測モデルの精度評価結果について以下に述べる．画像特徴量と患者情報を組合せて分類した際の ROC 曲線を図 2 に，ROC 曲線下の面積であり予測モデル精度を表す AUC 値を表 1 に示す．まず放射線画像からの特徴量のみの予測では，CT, PET を単独で用いた場合には，AUC が 0.75 から 0.76 であり，差はみられなかったが，CT

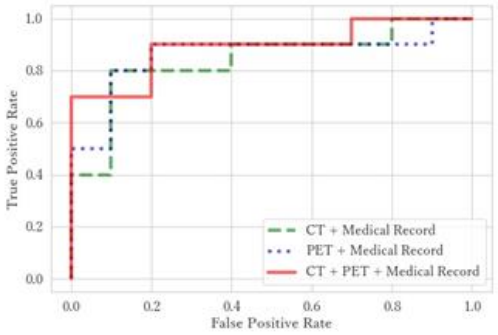


図 2 PD-L1 発現予測モデル評価結果

と PET 両方の画像特徴量を用いた場合では AUC が 0.87 と向上した．一方，画像特徴量に患者情報の特徴量を組合せた場合の精度は，画像特徴量単独と比較し精度が向上し，CT，PET に患者情報を組合せた場合，AUC が 0.89 で最も高い精度であった．

表 1 発現予測モデル評価結果（AUC）

項目	特徴量	AUC
画像のみ	CT	0.75
	PET	0.76
	CT + PET	0.87
画像 + 患者情報	CT + 患者情報	0.84
	PET + 患者情報	0.86
	CT + PET + 患者情報	0.89

2) 予測モデルの特徴量解析

本研究では医学知識である画像所見と関連の高い画像特徴量と，PD-L1 発現予測モデルに関連のある画像特徴量とを照合し，予測結果と同時に画像特徴量と画像所見の関連を提示することで，説明性が向上させる可能性を検討した．Radiomics 解析による PD-L1 発現予測モデルから抽出した発現予測に関連する画像特徴量と，別途 Radiomics 解析により診療録の画像所見と関連性の高い画像特徴量とを比較した結果を表 2 に示す．PD-L1 発現に高い関連がある特徴量と，画像所見に関連の高い特徴量には，共通する複数の特徴量が存在することが明らかとなった．テクスチャ特徴量 GLRLM の RunEntropy や，GLCM の SumEntoropy は，PD-L1 発現と腫瘍内への気管支有無を表す気管支透亮像と，また形状特徴量の体表面積は，PD-L1 発現と病変の境界と一対一に対応していた．一方，GLCM 特徴量の Imc（Informational Measure of Correlation）は，PD-L1 発現と病変の辺縁や性状の複数の画像所見と関連が見られた．

表 2 PD-L1 発現予測と読影所見の関連特徴量

発現予測に関連する特徴量	読影所見
Glrlm_RunEntropy	気管支透亮像
Glcm_SumEntoropy	気管支透亮像
Glcm_Imc1	辺縁 / 性状
Shape_SurfaceArea	境界
Ngtdm_Busyness	辺縁
Glcm_Imc2	辺縁 / 性状

4. まとめ

放射線画像から算出される画像特徴量と，診療録から得られる患者情報を Radiomics 特徴量とした PD-L1 発現予測モデルを構築，精度評価を実施し，AUC が 0.89 という結果を得た．また発現予測と，画像所見に関連のある画像特徴量を比較し，それらに共通する特徴量が存在することを明らかにした．以上から，従来生検により取得していた分子情報を，放射線画像の画像特徴量を用いて予測する可能性を見出し，また画像特徴量を医学知識（画像所見）と関連させ，予測モデルの説明性が向上する可能性が示された．

利益相反の有無

本研究の発表者である松崎和喜，金子幸生，高橋勲は，（株）日立製作所に属する社員である．

倫理審査の状況

本研究計画は，国立がん研究センター中央病院倫理委員会，および日立グループ倫理審査委員会で承認済みである．

文　献

［1］ P. Lambin, et al.: Radiomics: Extracting more information from medical images using advanced feature analysis”, European Journal of Cancer, 48, 441-446, 2012

［2］ RL. Siegel, et al.,“Cancer statistics, 2016”, CA Cancer J Clin. 66:7-30, 2016

A prediction method of PD-L1 expression for therapeutic decision support of lung cancer using radiomics imaging features

Kazuki MATSUZAKI[*1], Yukio KANEKO[*1], Isao TAKAHASHI[*1], Yuji MATSUMOTO[*2]

*1 Hitachi Ltd., Healthcare business unit

*2 National cancer center hospital, Department of endoscopy, Respiratory Endoscopy Division

Since the discovery of immune checkpoint proteins in recent years, there has been special interests in developing antibodies that block programmed cell death 1 (PD-1) or PD-ligand 1 (PD-L1) for a subset of cancer patients. PD-L1 signaling represents a valuable therapeutic target for cancer immunotherapy, while currently the information can be obtained only by biopsy with some difficulties. In this study, we investigated the association between CT- and/or PET-based image radiomic features and PD-L1 expression. A total of 106 radiomic features coming from 7 different feature categories were extracted from segmented volumes of each tumor in CT or PET images. Clinical and patient background information were obtained from medical record. Using PD-L1 expression label in medical record, PD-L1 expression prediction models based on image with patient information (CT-, PET-, CT and PET- with patient background) were built, and resulted in an AUC of 0.84, 0.86, and 0.89, respectively. From the examination of radiomics features in the PD-L1 expression predictive model and the diagnostic image report model, there found common radiomic features in the two models. In conclusion, it was suggested that our models using radiomics approach has good potential to predict PD-L1 expression for immunotherapy.

Key words: Radiological image, Radiomics, Lung cancer, Therapeutic decision support, PD-L1

Dilated convolution を用いた FCN による 腹腔鏡動画像からの血管領域抽出

盛満　慎太郎[*1]　山本　翔太[*1]　北坂　孝幸[*2]　林　雄一郎[*1]

小田　昌宏[*1]　竹下　修由[*3]　伊藤　雅昭[*3]

三澤　一成[*4]　森　健策[*1,5,6]

要旨

本稿では，腹腔鏡下手術動画像からの血管領域の自動抽出手法について述べる．腹腔鏡下手術は難易度の高い手術であるため，腹腔鏡動画を自動解析することで手術を支援するシステムの開発が進められている．動画からの術具や臓器の自動認識は，コンピュータに手術工程を理解させ，術者の手技を適切に支援させることにつながる．手術工程の理解において，体内構造を表す血管の抽出は重要である．これまでの血管領域抽出では U-Net をベースに，時系列情報を扱うための convolutional LSTM 層を追加して動画にした際の滑らかさを改善した．しかし，大域的な情報の不足により，推定領域の境界部分の抽出にばらつきがあった．そこで，本研究では境界部分の抽出精度を向上させるため，異なる dilation rate の畳み込み演算を行う dilated convolution 層を追加した Dilated LSTM U-Net を用いて血管領域セグメンテーションを行う．実験の結果，境界部分の抽出精度が 2.3% 向上した．

キーワード： Dilated convolution，腹腔鏡動画，血管領域抽出

1. はじめに

現在，がん治療等において低侵襲手術の 1 つである腹腔鏡下手術が広く用いられている．腹腔鏡下手術は患者への負担が少ない一方，鉗子の動きの制限や術中の視野の狭さから難易度の高い手術となっている[1]．そこで腹腔鏡動画に映る術具や臓器をコンピュータが自動的に認識し，術者の適切な手術を支援する研究が行われている[2,3]．認識対象の中でも体内構造をよく表す血管を認識することは，外科医にとって術野の適切な理解につながる．そのため本研究では血管領域抽出の初期目標として腹腔鏡下手術で切断対象となることの多い下腸間膜動脈（IMA）を対象にセグメンテーションを行う．

*1 名古屋大学大学院情報学研究科
〔〒464-8601 名古屋市千種区不老町 IB 電子情報館南棟 4 階 466 号室〕
e-mail: smorimi@mori.m.is.nagoya-u.ac.jp
*2 愛知工業大学情報科学部
*3 国立がん研究センター東病院
*4 愛知県がんセンター
*5 名古屋大学情報基盤センター
*6 国立情報学研究所医療ビッグデータ研究センター

2. 提案手法

提案手法では，従来手法[4]における LSTM U-Net のネットワーク構造を拡張し，dilated convolution 層を追加して血管領域抽出を行った．

従来の血管領域抽出では，医用画像処理におけるセグメンテーションに主に使われる U-Net [5] に時系列情報を扱うための convolutional LSTM 層を追加した LSTM U-Net [4] を用いてセグメンテーションを行っていた．そこでは推定結果の境界部分が粗く抽出されており，精度

を下げているという問題点があった．これは基本的な U-Net 構造だけでは対象領域の境界部分の特徴を正確にとらえきれていなかったことが原因と考えられた．そこで本手法ではより広域的な特徴を学習することで境界部分を正確にとらえるため dilated convolution [6] を用いることを検討した．

Dilated convolution 層では，dilation rate の異なる畳み込み演算を行う．図 1 に示すように dilation rate の値を変えると注目画素と畳み込むピクセルとの距離を変えることができる．そのため pooling などによって画像を縮小することなく広域的な情報を取り出すことができる．提案手法では畳み込むピクセルの間隔を変えた畳み込みを並列に行い，様々な解像度での特徴を扱うことで，今まで正確にとらえることのできなかった部分を詳細にセグメンテーションすることを可能とした．Dilated convolution 層では 1×1 の畳み込み，global average pooling，dilation rate が 2, 4, 8 の畳み込みを並列に行い，その出力を連結（concatenate）して次の層への入力とした．

モデルは LSTM U-Net のエンコーダーとデコーダーの間に Dilated convolution 層を加えた Dilated LSTM U-Net を用いた．学習では，1280×720 [pixels] および 1920×1080 [pixels] の画像を 256×256 [pixels] に縮小してモデルに入力した．その後出力された画像を元の大きさに戻して，正解ラベルとの Dice loss を計算し，Adam によって最適化を行った．評価では，データ数と症例数がグループ間で同程度になるように 5 分割して交差検定を行い，適合率，再現率，F 値で評価した．

3. 実験ならびに結果

実験では，腹腔鏡下 S 状結腸切除術または直腸癌に対する腹腔鏡下前方切除術の動画 37 症例，2566 枚分を用いた．正解データは医師の協力のもと，手動で作成した．

実験の結果として評価用画像に対する評価値を表 1 に示す．また ground truth を元画像に重畳表示した画像を図 2 (a)，従来手法[4]による抽出結果を重畳表示した画像を図 2 (b)，提案手法による抽出結果を重畳表示した画像を図 2 (c) に示す．表 1 を見ると F 値，適合率，再現率のそれぞれで従来手法より良くなっていることが分かる．

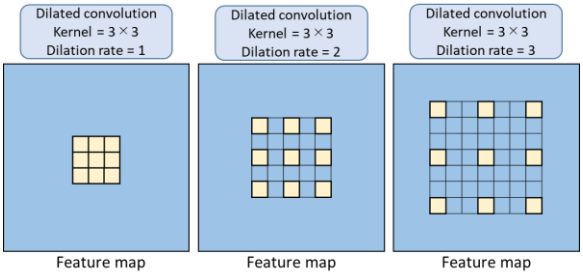
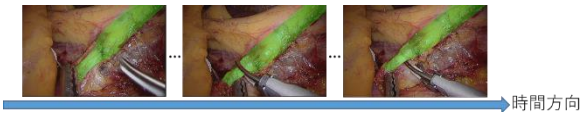


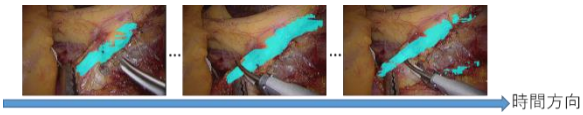
図 1 Dilation rate の異なる畳み込み演算の概要図．Dilation rate を変えると注目画素と畳み込むピクセルとの距離を変えることができる．

表 1 評価用データに対する精度評価．

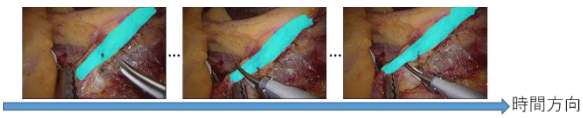
	F 値 (%)	適合率 (%)	再現率 (%)
従来手法[4]	47.6	55.7	49.0
提案手法	49.9	60.3	50.6



(a) ground truth



(b) 従来手法による抽出結果



(c) 提案手法による抽出結果

図 2 推定結果の一部．(a) は Ground truth を重畳表示した画像，(b) は従来手法（LSTM U-Net）による抽出結果，(c) は提案手法（Dilated LSTM U-Net）による抽出結果である．緑色は正解ラベル，青色は推定結果を表す．

4. 考察

図 2 を見ると従来手法では粗かった境界部分が滑らかに検出されていることが分かる．これは dilated convolution 層にて様々な解像度の特徴を学習したことにより，モデルが血管領域とそれ以外の部分をより判別できるようになったためであると考えられる．しかし，血管とカメラとの距離が変わると正しく推定できない場合があり，シーンによって精度にばらつきがある．そこで，多くのシーンに対応するためマルチスケールなモデルを使う，より様々なシーンを再現するデータ拡張を行うなど改良の余地がある．

5. むすび

本研究では，dilated convolution を用いた手法で腹腔鏡動画からの血管領域抽出を行った．実験の結果，境界部分の正確なセグメンテーションが可能となり，F 値が 2.3%上昇した．さらなる精度向上のためにはより多くのシーンに対応できるモデルの検討やデータ拡張などを行う必要がある．

謝辞

本研究の一部は JSPS 科研費 17H00867, 17K20099，26108006，AMED 課題番号 18he1802002, JP19lk1010036, 二国間交流事業，堀科学芸術振興財団の支援を受けた．

利益相反の有無　なし

文　献

［1］ 佐久間一郎 :低侵襲手術ロボットシステム．バイオメカニズム学会誌, **32**, 3, 147-152, 2008

［2］ Lee EJ, Plishker W, Liu X, et al: Segmentation of Surgical Instruments in Laparoscopic Videos: Training Dataset Generation and Deep Learning-based Framework. SPIE Medical Imaging 2019, **10951**:109511T, 2019

［3］ Gibson E, Robu MR, Thompson S, et al: Deep residual networks for automatic segmentation of laparoscopic videos of the liver. SPIE Medical Imaging 2017, **10135**: 101351M, 2017

［4］ 盛満慎太郎，山本翔太，小澤卓也，他.: LSTM を用いた腹腔鏡画像からの血管領域抽出．日本コンピュータ外科学会誌　第 28 回日本コンピュータ外科学会大会特集号　**21**:281-282, 2019

［5］ Ronneberger O, Fischer P, Brox T : U-Net: Convolutional Networks for Biomedical Image Segmentation. MICCAI 2015, **9351**: 234-241, 2015

［6］ Yu F, Koltun V: Multi-scale context aggregation by dilated convolutions. ICLR 2016

Segmentation of Blood Vessel Regions from Laparoscopic Videos Using FCN with Dilated Convolution

Shintaro MORIMITSU[*1], Shota YAMAMOTO[*1], Takayuki KITASAKA [*2], Yuichiro HAYASHI [*1], Masahiro ODA [*1], Masaaki ITO [*3], Nobuyoshi TAKESHITA [*3], Kazunari MISAWA [4], Kensaku MORI [*1,5,6]

*1 Graduate School of Informatics, Nagoya University
*2 School of Information Science, Aichi Institute of Technology
*3 National Cancer Center Hospital East
*4 Aichi Cancer Center Hospital
*5 Information Technology Center, Nagoya University
*6 Research Center for Medical Bigdata, National Institute of Informatics

In this paper, we propose an automatic segmentation method of blood vessel regions from laparoscopic surgery videos. Since laparoscopic surgery is one of highly difficult operations, development of a system that supports surgeons by automatically analyzing laparoscopic videos is underway. Automatic segmentation of surgical tools and organs from videos enables computers to understand surgical process and also enables appropriate support of the surgeon's procedure. Segmentation of the blood vessels is important for understanding surgical process. In the conventional blood vessel segmentation methods, the network architecture that was build based on U-Net and a convolutional LSTM layer for handling time series information were used to get smooth segmentation results along time. However, due to the lack of global information, the segmentation results of the boundary parts was not good. Therefore, in this study, in order to improve the segmentation accuracy of the boundary regions, we perform the blood vessel segmentation using the Dilated LSTM U-Net which has a dilated convolution layer that performs convolution calculations with different dilation rates. As a result of the experiment, the segmentation accuracy of the boundary part was improved by 2.3%.

Key words: Dilated convolution, Laparoscopic surgery videos, blood vessel segmentation

少数症例を用いた教師無し深層学習による病変検出画像特徴量の汎用的生成
－胸部 CT を用いた実験的検証－

牛房　和之[*1]　根本　充貴[*2]　木村　裕一[*1]　永岡　隆[*1]

山田　誉大[*3]　田中　敦子[*1]　林　直人[*4]

要旨

病変検出システムにおいて画像特徴量は重要な要素であることから，我々はこれまでに画像特徴量の自動生成法について研究してきた．本検討では，過去に提案した画像特徴量生成法について，学習データ量に対する頑健性を検証する．検証実験には，胸部 CT 上の肺結節画像パッチの認識問題を適用する．既提案の特徴量生成法では，正常データのみを学習した複数の Deep Convolutional Autoencoder (DCAE) を用いる．DCAE に入力する画像は，3 次元の肺局所パッチ画像を 2.5 次元変換した 3 チャネル画像である．4 種類の 2.5 次元画像生成法を用いて，それぞれの DCAE への入力データを作成する．本検討では，胸部 CT 300 例を学習した DCAE による特徴量セットと，より少ないデータで学習した DCAE による特徴量セットとの病変識別性能を比較した．識別処理には，共通して ada-boosted アンサンブルを用いる．検定による識別性能の比較の結果で，学習データ量を 5 例とした際の特徴量と 300 症例学習時の特徴量の性能との間に統計的有意差は見られなかった．このことから，提案法は学習データ量が少ない場合にも有用な特徴量を生成できることが示唆された．

キーワード：画像特徴量，　深層学習，　教師なし学習，　異常検知，　Convolutional Autoencoder

1. 背景・目的

コンピュータ検出支援(CADe)システムのような病変検出支援システムにおいて，画像特徴量はシステムの性能を大きく左右する重要な要素である[1]．特徴量は対象病変や画像モダリティの種類によって最適なものが異なる．最適な特徴量の選択と定義には専門的な知識と豊富な臨床経験が必要である．最近では，深層学習による特徴量の自動生成手法が注目を集めており，数多くの研究成果が報告されている[2]．深層学習を用いる手法では，学習データセットに基づいて特徴量が自動的に生成されるため，病変や画像モダリティの種類に合わせた特徴量を用いることができる．一般的に，深層学習では学習データが大量に必要といわれるが，病変を含む医用画像データは収集困難である．特に，罹患率の低い疾患などを対象とする場合は，大量のデータを収集することは現実的に困難である．

*1 近畿大学大学院 生物理工学研究科
〔〒649-6493 紀の川市西三谷 930〕
e-mail: k.ushifusa@gmail.com
*2 近畿大学 生物理工学部
*3 近畿大学高度先端総合医療センター PET 分子イメージング部
*4 東京大学 医学部附属病院 22 世紀医療センター

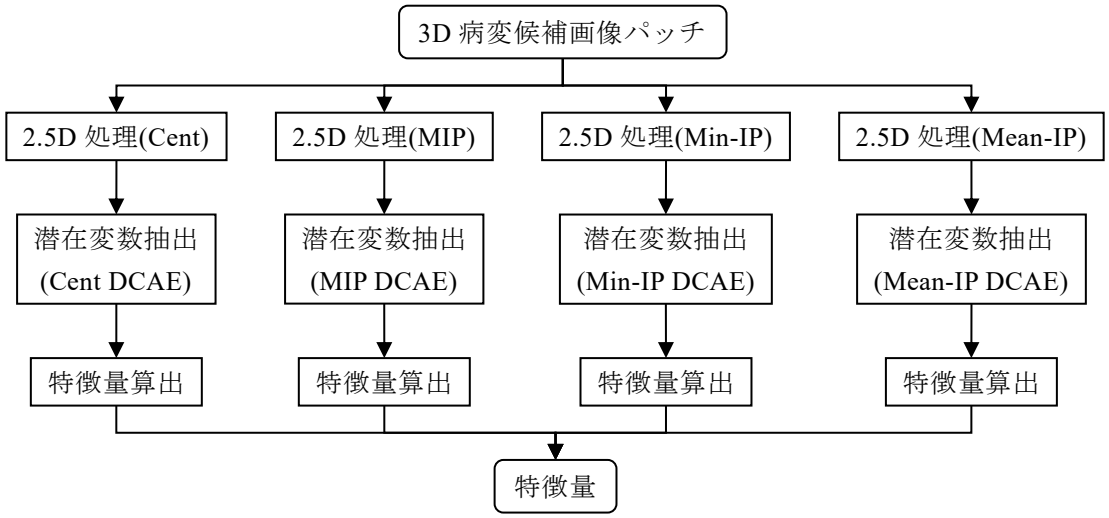


図 1 Deep convolutional autoencoder を用いた特徴量抽出と抽出特徴量を用いた識別

表 1 Deep convolutional autoencoder **の構造**

層の名称	カーネルサイズ	ストライド	出力サイズ
Input	ー	ー	32 × 32 × 3
Convolution(Conv)	3 × 3	1	32 × 32 × 6
Max Pooling(Pool)	2 × 2	1	16 × 16 × 6
Conv	3 × 3	1	16 × 16 × 9
Pool	2 × 2	1	8 × 8 × 9
Conv	3 × 3	1	8 × 8 × 12
Full connection	ー	ー	n

そのため我々はこれまでに，病変データと比べて収集が容易な正常データのみを深層学習する手法について検討してきた[3]. 本検討では，様々な規模の胸部 CT データセットを学習し，取得した生成特徴量の病変識別性能を確認する．そして学習データの規模を変更した際に病変識別性能における統計的有意差がみられるかを確認することで，提案法の学習データ数に対する頑健性を評価する．

2. 手法

1) 特徴量生成法

提案する特徴量生成法は，局所画像パッチからの画像特徴量抽出器を，正常画像パッチデータのみから学習する手法である．正常データのみ学習した 4 種の Deep convolutional autoencoder (DCAE)[4]を用いて得られる局所画像パッチの潜在変数ベクトルに基づき，複数の特徴量を計測する．正常データのみ学習した DCAE を使い，DCAE 未学習クラスである病変データから潜在変数を抽出する場合，学習済みの正常データから潜在変数を抽出する場合に比べて情報の損失が大きくなり，潜在変数の差異が発生することが期待される．

提案手法のフローチャートを図 1 に示す．提案手法ではまず，病変候補領域から抽出された局所 3 次元(3D)画像パッチを 4 種の投影法で 2.5D の局所画像パッチ[5]に変換する．次にそれぞれの 2.5D 画像パッチを対応する DCAE に入力し，潜在変数を抽出する．そして，その潜在変数から特徴量を算出する．最後にそれぞれの DCAE から抽出された特徴量を入力された 3D 画像パッチの特徴量セットとする．

入力される 3D 画像パッチの大きさは，32×32×32 voxels である．2.5D 画像は 3D 画像を少数次元に圧縮表現したものであり，少数データを用いた DCAE の学習の収束性向上が期待される．本手法では 2.5D 画像パッチは，3D 画像パッチを変換した，32×32 pixels の 3 チャンネル画像とする．さらに，様々な対象病変や画像

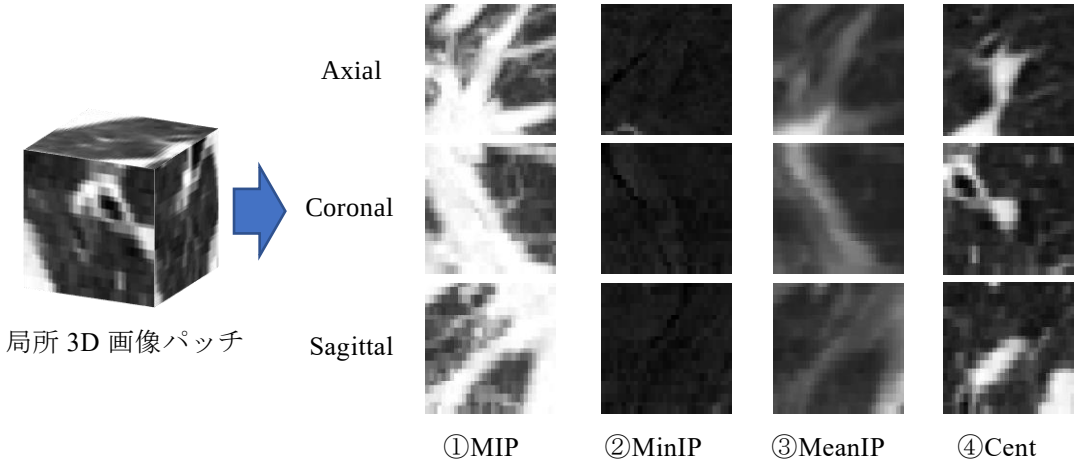


図 2 4 種の 2．5D 処理

モダリティの種類に対応するため，図 2 のような 3D 画像パッチの axial，coronal，sagitttal 断面への①最大値投影 : MIP(Max intensity projection)，②最小値投影 : Min-IP，③平均値投影 Mean-IP，および④3D 画像パッチの中心座標を通る axial，coronal，sagittal 断面の抽出 : Cent (Center Inteisity Extraction)の 4 種類の 2.5D 処理を利用する．

DCAE はこれらの 4 種の 2.5D 画像パッチに対応するため，4 種の DCAE を用意する．提案手法における DCAE の構造を表 1 に示す．この DCAE は VGG16[6]や Google Net[7]などと比べると簡易な構造であり，少数データを用いた DCAE の学習の収束性向上が期待される．表 1 で”n”として示されている潜在変数の次元数は，4 種の DCAE ごとの入力データの主成分分析における累積寄与率が 90%となる次元数とする．それぞれの DCAE の学習時は，潜在変数から画像を復元する Decoder もセットで学習する．Decoder の構造は，DCAE と同様のものを入出力反転させたものとする．DCAE および Decoder 学習における損失関数は，入力画像と潜在変数からの復元画像とのクロスエントロピー誤差とする．学習のミニバッチサイズは 256，エポックの数は 500 とし，学習終了条件の決定用データセットの損失が最も小さい時の学習結果を採用する．これらは実験的に決定した．

生成する特徴量は，DCAE により抽出される潜在変数，入力画像パッチと潜在変数からの復元画像パッチのピクセル値間の平均二乗誤差，潜在変数空間上での学習正常データセットの分布とのマハラノビス距離，および入力画像パッチと復元画像パッチ間の差分(L1 ノルム)画像の 9 種類のピクセル値統計量(最大，最小，平均，歪度，尖度，標準偏差，第 1 /第 2 /第 3 四分位)である．

2)　自動生成特徴量を用いた胸部 CT 上肺結節の認識法

前述の提案法で得た特徴量の性能検証のため，胸部 CT 上の肺結節検出 CADe システムで得られる肺動脈領域画像パッチの病変認識処理へ本手法で生成する特徴量を適用する．

肺血管領域画像パッチの抽出処理の内容は，以下の通りである．

①　0.78 mm 等方解像度に変換したボリュームデータに対する 2 値化処理等による動脈領域の自動抽出[8]

②　自動抽出した血管領域に対し，ストライド幅 16 voxels で 32×32×32 voxels の 3D 局所矩形領域をサンプリング

得られた病変候補パッチは，図 1 に示す通り，4 つの DCAE により抽出された潜在変数をもとに得られる特徴量を用いて識別される．

識別器には，学習時に特徴量の自動選択を同時に行う ada-boosted アンサンブル[9]を採用する．アンサンブルの学習には病変パッチと正常

パッチの数の不均等補正のため Cost sensitive 学習[10]を採用する．また，弱識別器は 100 個の決定木とする．ada-boost の学習過程では，得られた特徴量セットから肺結節認識に適した特徴量のみが自動選択される．前述の手法で得られる特徴量は，様々な病変や画像モダリティに用いることを前提とした多様なものであるため，より効果的な識別を実現するためには特徴量の選択が必要と考えた．

3)　特徴量生成法の学習データ数に対する特徴量識別性能の比較検証法

上述の特徴量生成法の学習データ数に対する頑健性を評価するため，学習データ数を変化させた際の病変識別性能を実験的に評価する．具体的には，300 症例を学習データに用いて自動生成される特徴量セットでの肺結節識別性能と，学習データを 200～5 症例に 9 段階で変化させた際にそれぞれ自動生成される特徴量セットでの肺結節識別性能を比較する．比較学習用の 200～5 症例の学習データは，300 症例の中からの無作為抽出により取得する．本検討では，3 回の無作為抽出および DCAE の学習・評価を行い，それらの平均性能を得る．性能検証する学習データの症例数と，それらの症例数から得られる局所画像パッチ数平均および標準偏差を表 2 に示す．

本検討では，ada-boosted アンサンブルの学習データは，特徴量生成器の学習症例数によらず，300 症例から得られたものを共通して用いる．その理由は，特徴量生成法の学習データ数の変化による生成特徴量への影響のみを評価結果に反映させるためである．

肺結節識別性能の評価は FROC 曲線を用いて行い，特に ANODE score [11]を性能指標として用いる．ANODE score とは 1 症例当たりの False Positive (FP) 数が 1/8，1/4，1/2，1，2，4，8 の時の感度の平均値であり，値が 1 に近ければ近いほど識別性能が高いことを示す．

また，300 症例学習で得た特徴量と 200 症例以下での学習で得た特徴量との ANODE score の有意差検定には，u 検定を用いる．

3．試料・実験方法

検証実験には，東京大学医学部附属病院で撮影され，放射線診断医による病変領域の画素単位での指摘が付けられている 450 例の胸部 CT を用いる．CT の解像度は $0.78\times0.78\times1.25\ mm^3$ である．

肺結節認識性能の検証には，150 例を用いる．残りの 300 例全て，あるいはその一部を用いて特徴量生成器の学習および識別器アンサンブルの学習を行う．300 例での特徴量生成器学習と性能評価は 3 回試行し，それ以下の症例数での特徴量生成器学習と評価は 9 回行う．

表 2 特徴量生成器の学習症例数と，それぞれの症例数から得られる学習パッチ数

症例数	学習パッチ数 (平均±標準偏差)
5	14368±2304
10	27429±2796
30	81832±4937
50	136105±4208
70	190897±4421
90	245372±6065
100	274362±7857
150	411232±10055
200	549049±11400
300	817849±20630

4．結果・考察

図 3 に DCAE の学習に 5 ~ 300 症例用いた場合の ANODE score [11]を箱ひげ図で示す. また，図 4 に DCAE の学習に 5 ~ 300 症例用いた場合の病変識別 FROC 曲線を示す. これらを見ると，DCAE の学習データ数を 5 症例まで減少させた場合でも 300 症例学習した場合の性能と明確な差異は無い(図 3，表 3)．実際，DCAE の学習に 300 症例を用いた場合の ANODE score を基準とした，5 ~ 200 症例学習時との ANODE score の差についての u 検定による有意差検定の結果，DCAE の学習データ数を 5 症例まで減少させた場合においても 300 症例学習時との有意差はみられなかった．この結果は，DCAE に入力する 3D 画像を 2.5D にしたことや DCAE の構造が比

較的単純な構造であることから，学習データが少数であっても，DCAE が最適に学習されたためと考えられる．さらに DCAE の学習データ数に依らず，常に 300 症例用いて学習する ada-boosted アンサンブルの識別性能が高いことも有意差がみられなかった要因と考えられる．

一方，DCAE の学習に 300 症例用いた場合と 5 ~ 200 症例用いた場合では識別性能差がみられることが分かる(図 4)．これは，試行数が少ないため，300 症例学習時において識別性能が低くなる場合や，5~200 症例学習時において識別性能が高くなる場合の結果が反映されていないためである．

これらの結果から，今後さらなる検証を重ねる必要はあるものの，本手法は学習データが少量であっても適用可能であることが示唆された．

5. 結論

本検討では過去に提案した画像特徴量生成法について，学習データ量に対する頑健性を実験的に検証した．

本手法における特徴量生成器の学習データセットの規模を変化させた場合の病変識別性能を比較した結果，提案法は，小規模なデータセットに対して有用な特徴量を自動生成することを確認した．今後は DCAE 学習時の損失関数などを変更することで，より正常データと病変データに差異が生じるような特徴量生成手法を検討する．

謝辞

本研究を行うにあたりご討論頂いている近畿大学生物理工学部医用工学科・根本研究室の学生諸氏，近畿大学大学院生物理工学部システム生命工学科・木村研究室および永岡研究室の諸氏に感謝します．そして，貴重な臨床データをご提供頂いた，東京大学医学部附属病院放射線科，同 22 世紀医療センターコンピュータ画像診断学／予防医学講座（CDRPM）の先生方に深く感謝します．

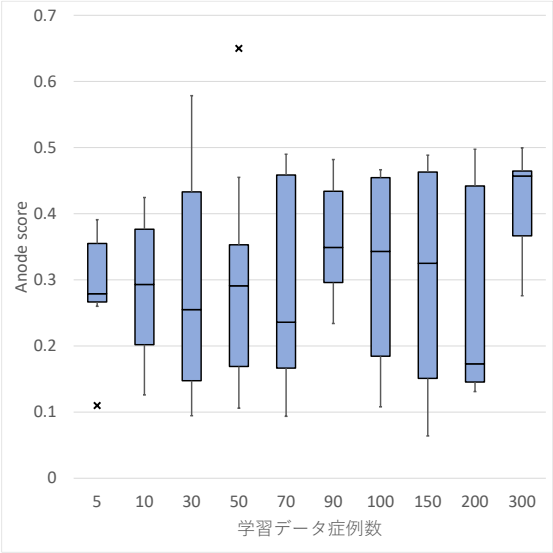


図 3 各学習データ数に対する ANODE score

表 3 各学習データ数に対する ANODE score

症例数	ANODE score (平均±標準偏差)
5	0.288 ± 0.0769
10	0.290 ± 0.0960
30	0.282 ± 0.144
50	0.290 ± 0.152
70	0.297 ± 0.148
90	0.357 ± 0.0709
100	0.330 ± 0.145
150	0.314 ± 0.166
200	0.278 ± 0.171
300	0.402 ± 0.0891

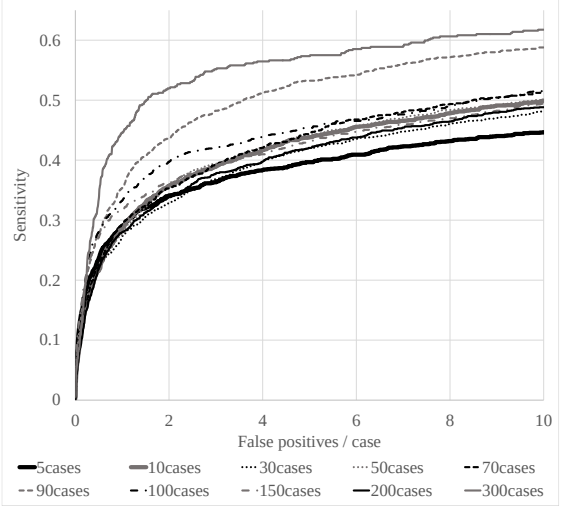


図 4 各学習データ数に対する FROC 曲線

利益相反の有無

なし

文　献

［1］ B. van Ginneken, C. M. Schaefer-Prokop, M. Prokop: Computer-aided diagnosis: how to move from the laboratory to the clinic, Radiology vol.261 no.3: 719-732, 2011

［2］ G. Litjens, T. Kooi, B. E. Bejnordi, et al.: A survey on deep learning in medical image analysis, Medical Image Analysis vol.42: 60-88, 2017

［3］ K. Ushifusa, M. Nemoto, Y. Kimura, et al.: A generalized image feature generation based on unsupervised deep learning with small scale normal dataset, Proc. CARS 2020 Computer Assisted Radiology and Surgery - 34th International Congress and Exhibition: 210-212

［4］ D. Sato, S. Hanaoka, Y. Nomura, et al.: A primitive study on unsupervised anomaly detection with an autoencoder in emergency head CT volumes, Proc. SPIE vol. 10575 Medical Imaging 2018 Computer-Aided Diagnosis: 105751P-1-10575P-6

［5］ H. R. Roth, L. Lu, A. Seff, K. M. Cherry, et al.: A New 2.5D Representation for Lymph Node Detection using Random Sets of Deep Convolutional Neural Network Observations, Proc. Medical image computing and computer-assisted intervention – MICCAI 2014: 520-527

［6］ K. Simonyan, A. Zisserman: Very Deep Convolutional Networks For Large-Scale Image Recognition, arXiv 1409 1556v6: 1-14, 2015

［7］ C. Szegedy, W. Liu, Y. Jia, et al.: Going deeper with convolutions, Proc. 2015 IEEE Conference on Computer Vision and Pattern Recognition: 1-9

［8］ Y. Nomura, M. Nemoto, Y. Masutani, et al.: Reduction of false positives at vessel bifurcations in computerized detection of lung nodules, Journal of Biomedical Graphics and Computing vol.4 no.3: 36-46, 2014

［9］ P. Viola, M. Jones: Rapid Object Detection using a Boosted Cascade of Simple Features, Proc. 2001 IEEE Computer Vision and Pattern Recognition: I-511-I-518

［10］ Y. Sun, M. S. Kamel, A. K.C. Wong et al.: Cost-sensitive boosting for classification of imbalanced data, Pattern Recognition vol.40 no.12: 3358‐3378, 2007

［11］ B. Ginneken, S. G.Armato 3rd, B. de Hoop: Comparing and combining algorithms for computer-aided detection of pulmonary nodules in computed tomography scans: The ANODE09 study, Medical Image Analysis vol.14 no.6: 707-722, 2010

A generalized generation of image feature for lesion detection by unsupervised deep learning with few cases - Experimental verification with chest CT -

Kazuyuki USHIFUSA[*1], Mitsutaka NEMOTO[*2], Yuichi KIMURA[*1], Takashi NAGAOKA[*1], Takahiro YAMADA[*3], Atsuko TANAKA[*1], Naoto Hayashi[*4]

*1 Graduate School of Biology-Oriented Science and Technology, Kindai University

*2 Faculty of Biology-Oriented Science and Technology, Kindai University

*3 Division of Positron Emission Tomography Institute of Advanced Clinical Medicine, Kindai University

*4 Department of Computational Diagnostic Radiology and Preventive Medicine, 22nd Century Medical and Research Center, the University of Tokyo Hospital

We have developed an image feature generation method since the image feature is one of the critical factors in the computer-aided detection system. In this study, we evaluate the robustness of the proposed method against the amount of the training dataset. The evaluation subject in this study is the classification of lung nodule patches from chest CT. The proposed method uses multiple deep convolutional autoencoders (DCAEs) that trained only normal data. An input image to the DCAE is a three-channeled 2.5-dimensional (2.5D) image from a three-dimensional local image patch. Four kinds of 2.5D image extraction methods are employed to obtain the input data for each DCAE. The performance of lung nodule classification with a feature set trained by a smaller number of chest CT cases is compared to the performance with a feature set trained by 300 CT cases. The ada-boosted classifier ensemble is employed as the standard classifier. As a result of the u-test for the classification performances, there was no statistically significant difference between the performances of the generated feature set by the training with 300 chest CT cases and the feature set by the training with 5 cases. It was shown that the proposed method could be generated useful image feature set by the training with a small amount of data.

Key words: Image feature, Deep learning, Unsupervised learning, Anomaly detection, Convolutional Autoencoder

深層学習を用いた歯科インプラントの分類

吉位　和将[*1]　助川　信太郎[*2]　原　武史[*1]　山下　勝督[*3]

中野　敬介[*4]　長塚　仁[*4]　古木　良彦[*2]

要旨

インプラントには多くのブランドがあり，メーカーごとに独自に開発されている．患者のインプラントのメンテナンスをする際にはブランドを特定する必要があるが，一人に対して異なるブランドのインプラントが異なる歯科医によって埋入されることが頻繁にあり，歯科医の間で共有される情報がなく特定するのは困難である．歯科パノラマ X 線画像からは特定に必要な情報を得ることができるが，これを行うには歯科医の努力と経験が必要になる．本研究の目的は，深層学習を用いて歯科パノラマ X 線画像からインプラントの分類を行い，歯科医や患者の負担軽減に寄与することである．畳み込み層が 6 つの CNN と，VGG16 と VGG19 の転移学習及びファインチューニングモデルでの分類精度を評価した．５つのモデルのうち VGG16 のファインチューニングが 94%と最も高い精度を示し，インプラントの分類において深層学習が有用である可能性が示唆された．

キーワード：歯科パノラマ X 線画像，深層学習，インプラント

1. はじめに

近年，歯科インプラント治療は歯を失った際の治療として急速に需要が高まっており，多くの患者の生活を改善させている．インプラントにはさまざまなブランドがあり[1]，メーカーごとに独自に開発されている．インプラントは定期的にメンテナンスが必要であり[2]，メンテナンスの際に使用する工具はブランドごとに異なるためブランドを特定する必要がある．しかし，患者一人に対して異なるブランドのインプラントが異なる歯科医によって埋入されることが頻繁にあり，歯科医の間で患者情報の共有がないため特定するのは困難である．

歯科医院や口腔外科で撮影されている歯科パノラマ X 線画像（図１）からは，ブランドの特定に必要な情報を得ることができる[3]．しかし，ブランドの特定には歯科医の努力と経験が必要であり，歯科医の大きな負担となっている．そのため，歯科パノラマX線画像からインプラントの分類を自動で行うことができれば，歯科医の負担を軽減することが可能である．本研究では，深層学習を用いて歯科パノラマX線画像からインプラントの分類を行い，その精度を評価した．

*1 岐阜大学工学部電気電子・情報工学科〔〒501-1194 岐阜県岐阜市柳戸１－１〕
e-mail: yoshii@fjt.info.gifu-u.ac.jp
*2 香川県立中央病院歯科口腔外科
*3 ポリテクセンター香川
*4 岡山大学大学院医歯学総合研究科口腔病理病態学分野

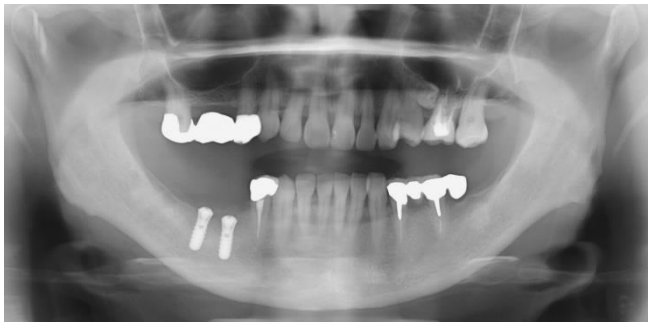


図 1 歯科パノラマ X 線画像

2. 方法

本研究では，インプラントの分類のために深層学習を用いた．ネットワークには畳み込み層 6 つの CNN，VGG16，VGG19 を使用し[4]，VGG16 と VGG19 については転移学習とファインチューニングを行った．転移学習では事前学習済みのモデルをそのまま使用し，ファインチューニングでは事前学習済みモデルの一部を再学習させた．これら 5 つのモデルを使用し，分類精度を評価した．

使用する画像は歯科パノラマ X 線画像のインプラントが写っている部分を切り出した画像 8840 枚であり，分類するインプラントのブランドは以下の 11 種類である．画像の例を表 1 に示す．

•Full OSSEOTITE 4.0: Full OSSEOTITE Tapered Certain (Zimmer Biomet, Florida, USA), diameter of 4 mm; lengths of 8.5, 10, 11, and 11.5 mm.
•Astra EV 4.2: Astra Tech Implant System OsseoSpeed EV (Dentsply IH AB, Molndal, Sweden), diameter of 4.2 mm; lengths of 9 and 11 mm.
•Astra TX 4.0: Astra Tech Implant System OsseoSpeed TX (Dentsply IH AB, Molndal, Sweden), diameter of 4 mm; lengths of 8, 9, and 11 mm.
•Astra TX 4.5: Astra Tech Implant System OsseoSpeed TX (Dentsply IH AB, Molndal, Sweden), diameter of 4.5 mm; lengths of 9 and 11 mm.
•Astra MicroThread 4.0: Astra Tech Implant System MicroThread, (Dentsply IH AB, Molndal, Sweden), diameter of 4 mm; lengths of 8, 9, and 11 mm.
•Astra MicroThread 4.5: Astra Tech Implant System MicroThread, (Dentsply IH AB, Molndal, Sweden), diameter of 4.5 mm; lengths of 9 and 11 mm.
•Brånemark Mk III 4.0: Brånemark System Mk III TiUnite (Nobelbiocare, Göteborg, Sweden), diameter of 4 mm; lengths of 8.5, 10, and 11.5 mm.
•FINESIA 4.2: FINESIA BL HA TP (KYOCERA Co., Kyoto, Japan), diameter of 4.2 mm; lengths of 8 and 10 mm
•Replace Select Tapered 4.3: Replace Select Tapered (Nobelbiocare, Göteborg, Sweden), diameter of 4.3 mm; lengths of 8, 10, and 11.5 mm.
•Nobel Replace CC 4.3: NobelReplace Conical Connection, (Nobelbiocare, Göteborg, Sweden), diameter of 4.3 mm; lengths of 8, 10, and 11.5 mm.
•Straumann Tissue 4.1: Standard Plus Implant Tissue Level implants (Straumann Group, Basei, Switzerland), diameter of 4.1 mm; lengths of 8 and 10 mm.

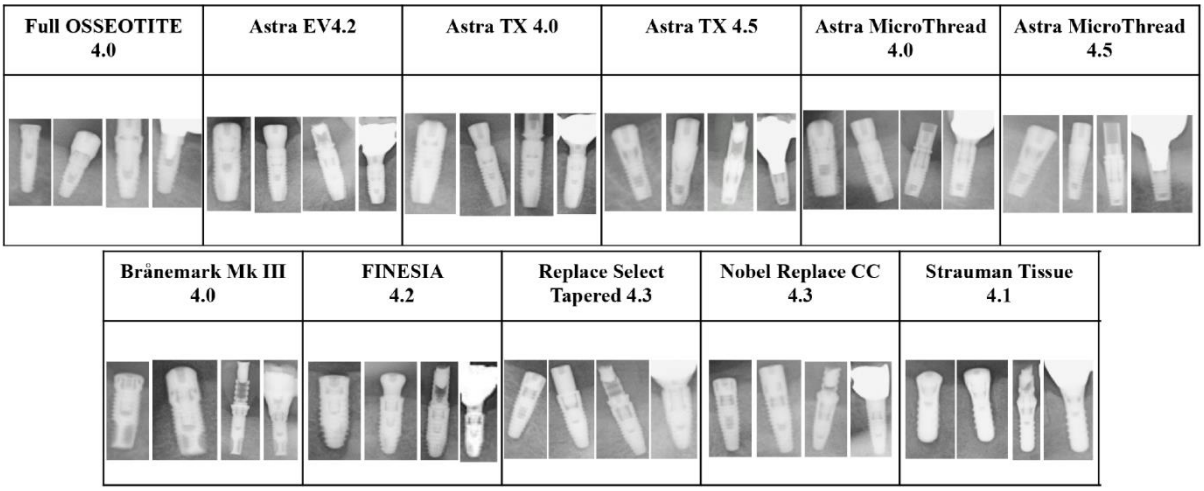


表 1 インプラントの画像の例

学習データについては回転, 拡大縮小をランダムに行う Data Augmentation を施すことで過学習を抑制している.

3. 結果

ネットワークの学習に 7072 枚，テストデータには 1768 枚の画像を用いた．ネットワークの学習に使用するデータ 7072 枚のうち 5304 枚を訓練用データ，1768 枚を検証用データとして 4fold の交差検証を行った.

分類精度の評価には Accuracy，Precision，Recall，F1 Score を使用した．それぞれの値の算出方法は以下の通りである．TP は真陽性，FP は偽陽性，TN は真陰性，FN は偽陰性である.

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (1)$$

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

$$F1\ Score = \frac{2\times(Recall\times Precision)}{Recall+Precision} \quad (4)$$

モデル	Accuracy	Precision	Recall	F1 Score
CNN	0.8143	0.8008	0.7361	0.7614
VGG16 転移学習	0.9089	0.8962	0.8826	0.8869
VGG16 ファインチューニング	0.9417	0.9304	0.9213	0.9248
VGG19 転移学習	0.8746	0.8643	0.8428	0.8501
VGG19 ファインチューニング	0.9403	0.9234	0.9192	0.9199

表 2 各モデルの結果

表 2 に各モデルの結果を示す. 5 つのモデルの中で VGG16 のファインチューニングが最も優れた結果となった.

4. まとめ

本研究では，深層学習を用いた歯科パノラマX線画像におけるインプラントの分類を行った．精度は VGG16 のファインチューニングが最も高く 94.17%だった．本研究により，インプラントの分類において深層学習が有用である可能性が示唆された．

利益相反の有無

なし

文　献

［1］ Jokstad A, Braegger U, Brunski J.B, et al.: Quality of dental implants*. Int. Dent. J. 53(S6P2), 2003, 409–443.

［2］ Lee K.-Y, Shin K.S, Jung, J.-H, et al.: Clinical study on screw loosening in dental implant prostheses: A 6-year retrospective study. J. Korean Assoc. Oral Maxillofac. Surg. 46(2), 2020, 133–142.

［3］ Molander B.: Panoramic radiography in dental diagnostics. Swed. Dent. J. Suppl. 119, 1996, 1–26.

［4］ Simonyan K, Zisserman A: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014).

Template for Medical Imaging Technology

Kazumasa YOSHII[*1], Shintaro SUKEGAWA[*2], Takeshi HARA[*1], Katsusuke YAMASHITA[*3], Keisuke NAKANO[*4], Hitoshi NAGATSUKA[*4], Yoshihiko FURUKI[*2]

*1 Department of Electrical, Electronic and Computer Engineering, Faculty of Engineering, Gifu University
*2 Department of Oral and Maxillofacial Surgery, Kagawa Prefectural Central Hospital
*3Polytechnic Center Kagawa
*4Department of Oral Pathology and Medicine, Okayama University Graduate School of Medicine, Dentistry and Pharmaceutical Sciences

There are many brands of dental implants, and each manufacturer has developed its own brand. It is necessary to identify the brand when maintaining a patient's implants. However, different brands of implants are frequently placed by different dentists for a patient, and it is difficult to identify them because information about a patient isn’t shared among dentists. Dental panoramic radiographs can provide the necessary information for identification, but this requires the effort and experience of the dentist. The purpose of this study is to use deep learning to classify implants from dental panoramic radiographs to help reduce the burden on dentists and patients. This study assesses the accuracy of using digital panoramic X-ray radiograph images for classification and clarification of dental implant brands via CNN with six convolutional layers, the transfer learning and fine-tuning models of VGG16 and VGG19. Among the five models, the fine-tuning of VGG16 showed the highest accuracy of 94%. This result suggests that deep learning may be useful in implant classification.

Key words: Dental panoramic radiographs, Deep learning, Dental implant

大域及び局所情報を用いた深層学習による出血領域自動セグメンテーション

山本　翔太[*1]　盛満　慎太郎[*1]　林　雄一郎[*1]　北坂　孝幸[*2]

小田　昌宏[*1]　伊藤　雅昭[*3]　竹下　修由[*3]　森　健策[*1,4,5]

要旨

本稿では，腹腔鏡下手術動画像に対して，深層学習を用い出血領域のセグメンテーションを自動的に行う手法について検討する．出血領域のセグメンテーション結果は内視鏡手術の安全性を高める研究において利用価値が高い．しかし，手術映像中における出血領域を手動でセグメンテーションするのは困難であるため，効率的にセグメンテーションを行う手法が求められている．従来手法として，画像中の注目領域が異なる複数の U-Net を用いて出血領域の自動抽出を行うものがある．しかし，微小な出血領域の抽出が困難であるという問題点があった．本研究では，画像全体から出血領域を含む局所領域を検出する YOLOv3 と，局所的な出血領域抽出を行う U-Net を使用する．YOLOv3 が推定した位置を U-Net で出血領域推定することによりセグメンテーション結果を得る．実験により，提案手法による腹腔鏡下手術動画からの出血領域自動抽出の F 値が 57.98%を示した．

キーワード：深層学習，内視鏡手術動画，セグメンテーション，U-Net，出血領域

1. はじめに

腹腔鏡下手術は，低侵襲手術であるため患者への負担が小さいという利点がある一方，視野が制限されることや手術器具の操作が鉗子挿入口により拘束されることから，高度な技術が要求される難易度の高い手術である[1]．手術手技は医師の経験や感覚が大きく寄与しており，暗黙知とされている部分があるため，手術中の状況を認識するシステムの開発が望まれている[2]．日本内視鏡外科学会の技術認定評価項目には，評価対象として出血に関する項目が設定されている[3]．そこで，手術で重要とされる出血に着目し，腹腔鏡下手術映像中から出血領域を自動的に抽出することで，手術手技の定量化や客観性の向上，また，解析結果を基に新たな手術支援システムへの応用が期待される[4]．山本らは，画像全体から大まかな抽出を行う広域U-Net[5]と，画像全体の一部を切り出した画像から詳細な抽出を行う局所U-Netを用いて，それぞれの抽出結果をOR演算により合成することで出血領域自動抽出を行った[6]．しかし，画像全体を縮小するため局所領域の抽出が困難であるという問題点があった．そのため，本稿では，局所領域を検出するためのYOLOv3[7]と局所領域から出血を詳細に抽出するためのU-Netの2つのモデルを用いて出血領域の自動抽出を行う．

*1 名古屋大学大学院情報学研究科
〔〒464-8601 名古屋市千種区不老町 IB 電子情報館南棟 4 階 466 号室〕
e-mail: syama@mori.m.is.nagoya-u.ac.jp
*2 愛知工業大学情報科学部
*3 国立がん研究センター東病院
*4 名古屋大学情報基盤センター
*5 国立情報学研究所医療ビッグデータ研究センター

2. 出血領域自動抽出手法

本手法では，画像中に出現する局所的な出血の検出と，出血領域の詳細な抽出の処理を順に

行う．はじめに腹腔鏡下手術映像中の任意のフレーム画像から YOLOv3 を用いて出血領域を含む局所領域の検出を行う．次に検出結果と対応するフレーム画像の局所領域を U-Net へ入力し，出血領域の抽出を行う．そして，U-Net の抽出結果を YOLOv3 で検出した位置へ統合することで自動抽出結果を得られる．

学習には，出血が出現する手術動画中の一部のフレーム画像と，それに対応する出血領域のラベル画像を用いる．YOLOv3 の学習では，1280×720 pixels または 720×480 pixels のフレーム画像を 416×416 pixels へ縮小した画像と，ラベル画像から取得した出血領域のバウンディングボックスを用いる． U-Net の学習では，出血領域のバウンディングボックスを切り出したフレーム画像とラベル画像をそれぞれ 256×256 pixels へリサイズ処理し学習用画像として用いる．

推定処理の流れを図 1 に示す．はじめに対象画像を 416×416 pixels へ縮小した画像を YOLOv3 へ入力し，出血領域のバウンディングボックスの推定を行う．次に，検出されたバウンディングボックスと対応する領域を対象画像から切り出し, 256×256 pixels へリサイズ処理を行う．リサイズ後の画像を U-Net に入力し出血領域の詳細な抽出を行う．U-Net から得られた結果をリサイズ前の大きさへ拡大し，対象画像の元の位置へ統合することで出血領域自動抽出結果を得る．

3. 実験と結果

実験では，従来手法[6]と提案手法の抽出精度の比較実験を行った．13 症例分の腹腔鏡手術動画から出血領域を手動でラベル付けしたフレーム画像 753 枚を学習用画像, 83 枚をバリデーション用画像として使用し，245 枚で推定を行った．各手法による出血領域自動抽出結果は図 2 の通りとなった．また，それぞれの再現率，適合率，F 値を表 1 に示す．

4. 考察

図 2 より，提案手法により詳細に出血領域が抽出されていることが分かる．このことから，提案手法は出血領域自動抽出について有効であると考えられる．

表 1 より，提案手法の F 値が 57.98%を示した．従来手法と比較して再現率・適合率がともに高い値を示していることから，過抽出を抑えつつ，出血領域をより多く抽出できていると考えられる．

しかし，従来手法と同様に，再現率が適合率より低い傾向を示しており，出血領域の抽出が不十分であることが分かる．そのため，より多く抽出するための工夫が必要である．

5. むすび

本稿では，YOLOv3 と U-Net を用いた腹腔鏡手術動画からの出血領域自動抽出手法についての検討を行い，有効性の確認を行った．今後は抽出精度の向上を図るとともに，自動抽出結果に基づく手術支援技術への応用を目指す．

謝辞

本研究の一部は JSPS 科研費 17H00867, 17K20099, 26108006, 26560255, 二国間交流事業, 堀科学芸術振興財団, AMED の課題番号 19he2102001h0701,20he2102001h0702, JP19lk1010036, JP 20lk1010036 の支援を受けた．

利益相反の有無　なし

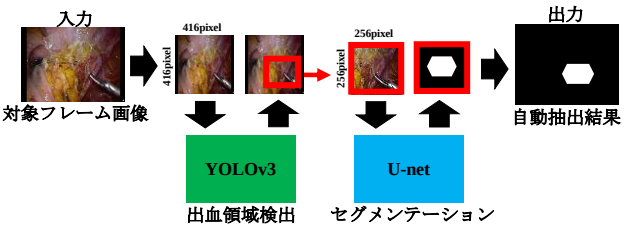


図 1 出血領域の自動抽出手法．対象画像を 416×416 pixels へ縮小した画像を YOLOv3 へ入力し，出血領域のバウンディングボックスを推定する．対象画像から推定された位置を切り出し，256×256 pixels へリサイズする．リサイズされた切り出し画像を U-Net へ入力し，出血領域とその周辺の抽出結果を得る．U-Net から得られた結果を対象画像の元の位置へ統合することで出血領域自動抽出結果を得る．

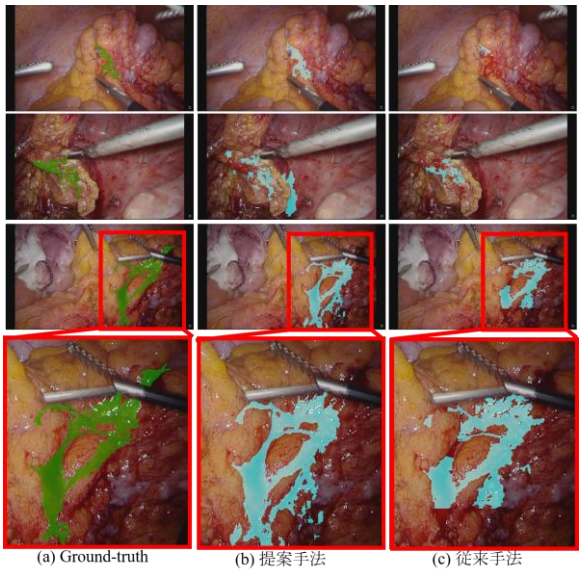


図 2 各手法の自動抽出結果と対象画像の重畳表示とその拡大図．緑の領域は出血領域の Ground-truth を表し，青の領域は抽出結果を表す．

表 1 出血領域の自動抽出精度．各手法の再現率，適合率，F 値を表す．

	再現率(%)	適合率(%)	F 値(%)
従来手法	41.19	56.08	48.64
提案手法	**50.32**	**68.38**	**57.98**

文　献

［1］ 長谷川博俊，渡邊昌彦，西堀英樹，他：腹腔鏡下大腸全摘術．消化器外科 **27**: 655-667, 2004

［2］ 佐久間一郎：低侵襲手術ロボットシステム．バイオメカニズム学会誌 **32**: 147-152, 2008

［3］ Kimura T, Mori T, Kitajima M: Skill Accreditation System for Laparoscopic Gastroentrologic Surgeons in Japan. MITAT **19**: 41-46, 2010

［4］ 伊藤雅昭：内視鏡手術のデータベース化と応用展開．日本コンピュータ外科学会誌 **21**: 211, 2019

［5］ Ronneberger O, Fischer P, Brox T: U-Net:Convolutional Networks for Biomedical Image Segmentation. MICCAI **9351**, 234-241, 2015

［6］ 山本翔太，小田紘久，林雄一郎，他：深層学習を用いた腹腔鏡手術動画像の出血領域自動セグメンテーション．第 38 回日本医用画像工学会大会予稿: 503-505，2019

［7］ Redmon J, Farhadi A: YOLOv3: An Incremental Improvement. CVPR, 2018

Automatic Segmentation of Bleeding Area by Deep Learning Using Global And Local Information

Shota YAMAMOTO[*1], Shintaro MORIMITSU[*1], Yuichiro HAYASHI[*1], Takayuki KITASAKA[*2], Masahiro ODA[*1], Masaaki ITO[*3], Nobuyoshi TAKESHITA[*3], and Kensaku MORI[*1,4,5]

*1 Graduate School of Informatics, Nagoya University
*2 Faculty of Information Science, Aichi Institute of Technology
*3 National Cancer Center Hospital East
*4 Information Technology Center, Nagoya University
*5 Research Center for Medical Big Data,
National Institute of Informatics

We discuss a method for automated segmentation of bleeding area using deep learning from laparoscopic surgical movie in this paper. Segmentation results of bleeding area are valuable for research to enhance the safety of endoscopic surgery. However, since manual segmentation of bleeding area from surgical movies is difficult, an automatic bleeding area segmentation method is required. Conventionally, U-Net with different regions of interest in images was used for automatic segmentation of the bleeding area. However, it was difficult to segment small bleeding areas. In this study, we use two deep learning models, YOLOv3 that detects local areas including the bleeding area from the entire image and U-Net that segments bleeding areas from local images. The segmentation result is obtained by estimating the bleeding area using U-Net at the position estimated by YOLOv3. In our experiments, F-measure of the automatic segmentation of bleeding areas from laparoscopic surgery videos by the proposed method was 57.98%.

Key words: Deep learning, Laparoscopic Surgical Movie, Segmentation, U-Net, Bleeding Area

Classification of large-scale image database of various skin diseases using deep learning

Masaya Tanaka[*1], Atsushi Saito[*1], Kosuke Shido[*2], Yasuhiro Fujisawa[*3], Kenshi Yamasaki[*2], Manabu Fujimoto[*4], Kohei Murao[*5], Youichirou Ninomiya[*5], Shin'ichi Satoh[*5], Akinobu Shimizu[*1]

Abstract

Skin diseases vary significantly in appearance and diversity in the distribution of lesions over the human body; thus, diagnosis is difficult for non-dermatologists. In this paper, we present a deep learning-based computer-aided diagnosis (CAD) system for the photographic image classification of 59 skin diseases. The previous skin disease CAD systems were mainly developed to classify localized lesions, such as melanoma, using an image presenting only a skin lesion and captured under a standardized imaging condition. In this study, we deal with various images of 59 skin diseases, including skin tumors, inflammatory diseases, and autoimmune diseases, in which the morphology and distribution of lesions on body parts largely differ from each other. The proposed system allows the input of a few to dozens of images for one patient taken by dermatologists using a standard optical camera under various imaging conditions, including changes in the magnification, angle, and lighting. Similar to dermatologists, the proposed system observes lesions from various viewpoints and effectively integrates features relating to skin lesions from multiple images of body parts and from images captured under various imaging conditions. The effectiveness of the system is demonstrated using approximately 90,000 images from the national database of the Japanese Dermatological Association. We also discuss the consistency between the learned network and the medical classification of skin diseases as well as the limitations of the system.

Keywords：skin disease classification, computer-aided diagnosis, multi-view, ensemble

1. Purpose

The appearances of skin diseases widely differ. For example, a malignant melanoma, which is a typical malignant tumor, is a localized lesion to a limited extent. In contrast, bullous pemphigoid, a typical autoimmune disease, covers a large area on the body surface. There are more than 2000 skin diseases that vary in their morphology and appearance, not only among different diseases, but also within the same disease[1],

*1 Institute of Engineering, Tokyo University of Agriculture and Technology
e-mail: simiz@cc.tuat.ac.jp
*2 Department of Dermatology, Tohoku University Graduate School of Medicine
*3 Department of Dermatology, Faculty of Medicine, University of Tsukuba
*4 Department of Dermatology, Osaka University Graduate School of Medicine
*5 Research Centre for Medical Bigdata, National Institute of Informatics

making diagnosis by non-specialists difficult. Therefore, there is a need to develop a computer-aided diagnosis (CAD) for the photographic images of various types of skin diseases.

Several CAD systems have been developed for skin diseases, and these have shown comparable or even higher accuracies than those of dermatologists[2,3,4]. However, most of these studies focused on localized diseases and required the use of carefully adjusted imaging conditions or special imaging equipment, such as a dermatoscope. A CAD system that is truly helpful for non-specialists is expected to allow multiple input images of various types of skin diseases, including both localized and diffused lesions, captured using standard optical cameras that generate images under various conditions, including different magnifications, angles, lighting, and background noises.

This paper presents a deep-learning-based CAD system for 59 diseases, including both localized and diffuse diseases. All images were taken by dermatologists using standard optical cameras, and the number of images differed between patients. To effectively integrate the information of multiple images with various imaging conditions and different numbers of images, we introduce an ensemble technique. We discuss the performance of the system and the consistency between the deep network and medical classification.

2. Method

ResNet18[5] was used for classification, the model of which has 17 trainable convolutional layers followed by one fully connected layer, and the weights were pre-trained using the ILSVRC[6] dataset. We slightly modified the network such that the size of the input layer was 448 px × 448 px and the output layer had 59 nodes corresponding to the 59 skin diseases. The cross entropy loss and Adam[7] optimizer were used for training, and the model was trained by 200 epochs with a learning rate of 10^{-5}. As a spatial normalization of an input image, the shorter side of the image was resized to 512 px by an affine transformation, while maintaining the aspect ratio. Subsequently, 448 px × 448 px sub-images were randomly cropped from the resized images. Furthermore, flip or rotation operations were randomly applied as augmentation in the training step. We adopted an ensemble technique that simply averages the confidence values over multiple images of a patient.

3. Results

We used images of 59 diseases from the national database of the Japanese Dermatological Association. The number of diseases was decided by considering the minimum number of cases in a disease class. The images were captured using 85 camera models, such as “Canon EOS 10D” and “NIKON D5000,” at 12 medical institutions in Japan and classified into one of the 59 skin diseases. A taxonomy tree (8 classes in the first level, 13 classes in the second level, and 59 classes in the third level) was also designed for this study according to clinical appearance. A total of 72,095 images from 11,456 patients were used as the training set, and 17,613 images from 2855 patients were used as validation and test sets.

Figure 1 shows the relationship between the classification accuracy and number of images used for the ensemble that were randomly selected from images of a patient in the test set. Note that if the number of images in the subset was less than a given number, random sampling without replacement was performed to meet the condition regarding the number. “ALL” in the figure shows the results of the average confidence values over all images of a patient. The error bars show the maximum and minimum accuracies by iterating the random selection of images 50 times.

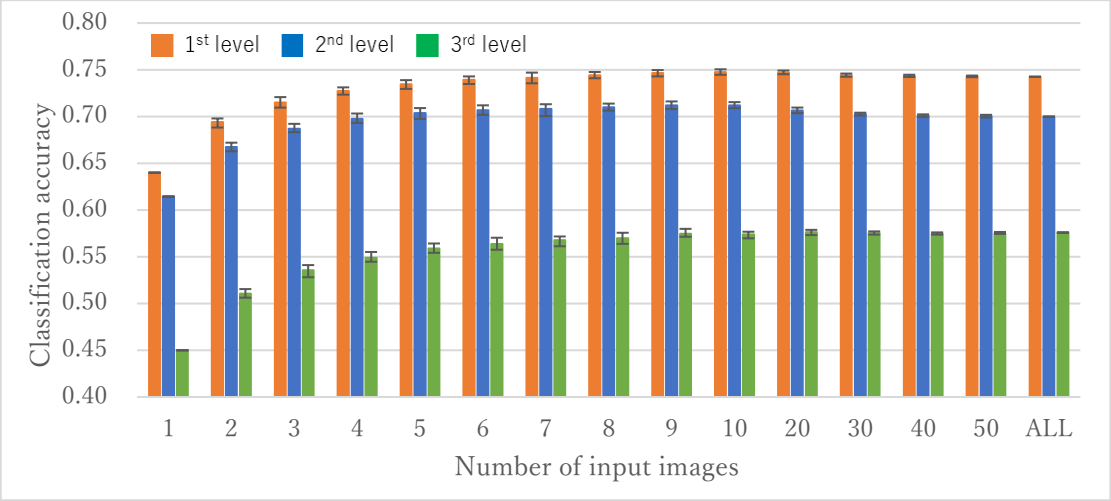


Fig. 1 Relationship between classification accuracy and number of images used in the proposed system

Figure 1 shows that the accuracy significantly improved at all levels of the taxonomy tree by an ensemble of multiple images of a patient. It was also confirmed that the accuracies saturated around ten images and were comparable to those of "ALL," in which all images of a patient are used. The differences between the accuracies of "ALL" and those using an image were 10.3 points at the 1st level, 8.6 points at the 2nd level, and 12.6 points at the 3rd level.

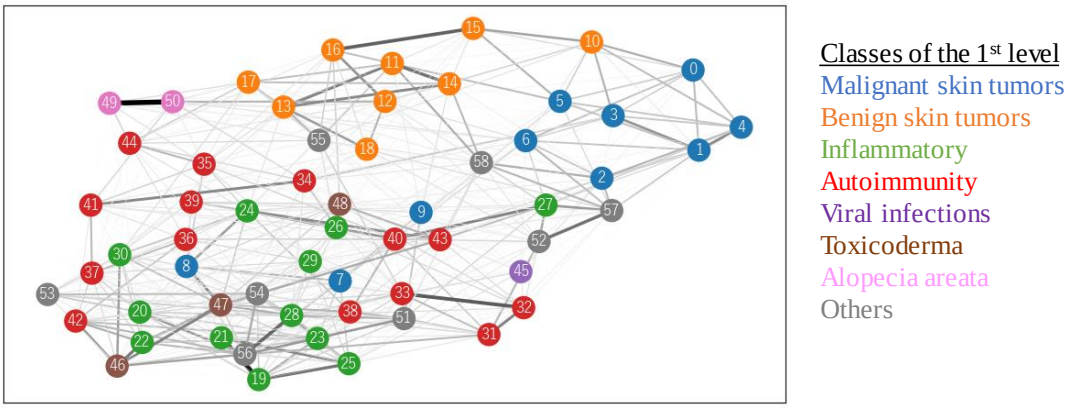


Fig. 2 Similarity of weights at fully connected layer between all possible pairs of 59 classes. The graph was constructed by the method in [8]. The thicker lines indicate a higher positive cross-correlation coefficient between the weights of paired classes. Numerals in nodes indicate the predicted classes at the 3rd level, and the colors of the nodes represent the predicted classes at the 1st level.

As observed from the graph in Figure 2, skin diseases classified into the same class at the 1st level are closer together, which is consistent with the fact that most skin diseases in the same class at 1st level are similar in appearance. Such consistency with medical classification is expected to make the CAD system more reliable and explainable. In contrast, we should note some exceptions. For example, lymphocytic malignant skin tumors (nos. 7, 8, and 9) are far from the other malignant skin tumors shown by the blue nodes (nos. 0–6) because of the difference in appearance.

4. Conclusion

This paper presented a deep-learning-based CAD system for 59 skin diseases. Input images were captured using standard optical cameras, and the number of images differed between patients. A national database comprising approximately 90,000 images was used for training, validating, and testing the system. It was confirmed from the experiment that the classification accuracy was improved by the proposed technique, which effectively integrates information of multiple images taken under various conditions. The consistency between the trained network and medical classification was confirmed by analyzing the weights at the fully connected layer of the network.

Competing interests

The authors declare no conflicts of interest associated with this manuscript.

Acknowledgement

This research is in part supported by the ICT infrastructure establishment and implementation of artificial intelligence for clinical and medical research from Japan Agency for Medical Research and development, AMED, JP20lk1010036.

References

[1] Griffiths, C., Barker, J., Bleiker, et al.: Rook's Textbook of Dermatology, 9th ed., John Wiley & Sons, 2016

[2] Esteva, A., Kuprel, B., Novoa, et al. : Dermatologist-level classification of skin cancer with deep neural networks. Nature **542**: 115-118, 2017

[3] Fujisawa, Y., Otomo, Y., Ogata, Y., et al.: Deep‐learning‐based, computer‐aided classifier developed with a small dataset of clinical images surpasses board‐certified dermatologists in skin tumour diagnosis. British Journal of Dermatology **180**(2): 373-381, 2019

[4] Codella, N., Rotemberg, V., Tschandl, P., et al.: Skin Lesion Analysis Toward Melanoma Detection 2018: A Challenge Hosted by the International Skin Imaging Collaboration (ISIC). arXiv:1902.03368, 2019

[5] He, K., Zhang, X., Ren, S., et al.: Deep Residual Learning for Image Recognition. In: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016

[6] Russakovsky, O., Deng, Jia., Su, H., et al.: ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision **115**: 211–252, 2015

[7] Kingma, D. and Ba, J. : Adam: A method for stochastic optimization," In ICML, 2015

[8] Fruchterman, T.M.J. and Reingold, E.M..: Graph drawing by force-directed placement. Software – Practice and Experience **21**: 1129-1164, 1991

Automated segmentation of musculoskeletal structures from CT image of lower leg: toward personalized musculoskeletal sciences

Mazen SOUFI[*1], Yoshito OTAKE[*1], Takuma MIYAMOTO[*2], Yasuhito TANAKA[*2], Keisuke UEMURA[*1,3], Masaki TAKAO[*3], Nobuhiko SUGANO[*3], Yoshinobu SATO[*1]

Abstract

In this research, we aim at reconstructing the high fidelity three-dimensional (3D) structure of the musculoskeletal anatomy from CT and MR images. Based on this reconstruction, instead of using conventional generic musculoskeletal models, patient-specific models can be obtained, which will make it possible to build personalized plans and predict outcomes in sports trainings, orthopedic surgery and rehabilitation. In our previous study, our group developed an automated segmentation approach based on deep learning, so-called Bayesian U-Net, which has shown high accuracy in segmenting musculoskeletal structures at the vicinity of the hip joint in CT images, aiming at applications in hip joint surgical planning. In this research, we aim at extending our work to segment the lower leg structures including muscles and bones from the knee to the foot. We report our preliminary results on 8 CT images and future directions.

Keywords：Bayesian U-Net, Muscle segmentation, Uncertainty

1. Introduction

The development of patient-specific plans in orthopedic surgeries, sports trainings, orthopedic surgery and rehabilitation, besides the assessment of their future outcomes, highly depend on the 3D geometry of the musculoskeletal structures. These plans are usually built by adjusting generic models derived from medical images, such as magnetic resonance (MR) images, to approximate the patient's structures based on physical measurements on the patient's body [1–3]. However, this approach lacks the high fidelity representation of the patient's geometry, which may affect the accuracy of the reconstructed models. In the same time, given the complicated structure of the musculoskeletal system, the manual segmentation of the muscles and bones becomes a challenging task due to the excessively required time and efforts.

In order to automatically generate the patient-specific high fidelity models, several automated segmentation approaches of muscles and bones from CT and MR images have been developed. Recently, an automated leg muscle segmentation algorithm using a deep convolution neural network (CNN) was presented in [4–7]. The advent of deep

1 Division of Information Science, Nara Institute of Science and Technology

8916-5, Takayama-cho, Ikoma, Nara 630-0192, Japan

* email: msoufi@is.naist.jp

2 Nara Medical University, 840, Shijocho, Kashihara, Nara, Japan 634-8521

3 Graduate School of Medicine, Osaka University, 2-2 Yamadaoka, Suita, Osaka, Japan 565-0871

learning, particularly convolutional neural networks (CNN), such as U-Net, made it possible to obtain high accuracies [5–6]. The U-Net structure, including multi-scale encoding and decoding paths with skip connections, enabled capturing the detailed structures of individual muscles. However, as the ground-truth is an essential requirement to assess the segmentation accuracy, it becomes difficult to assess the accuracy for images without ground-truth, which might limit the clinical applicability.

Recently, our group has addressed this problem through the development of a Bayesian U-Net, which produces uncertainty measurement of the segmentation obtained by the network based on dropout sampling [7]. Applied to muscles and bones at the vicinity of the hip joint, a good correlation was obtained between the measured uncertainty and segmentation accuracy. This has shown the possibility to use the uncertainty as a predictive measure for the segmentation accuracy in the absence of ground-truth labels. In this study, we aim apply the Bayesian U-Net for the automatic segmentation of lower leg bones and muscles in CT images.

2. Methods

In this study, a Bayesian U-Net was used to segment the bones and muscles from the CT images. The network architecture follows that of a 2D U-Net. The difference lies in adding dropout layers, which randomly set a portion of the model's weights equal to zero, before each max pooling layer and after each up-convolution [7]. The output labels are inferred at testing time by taking T samples and calculating the average of output probabilities by the softmax function for each segmented structure. For more details on Bayesian U-Net, the reader is referred to Ref. [7].

3. Experiments

Ten clinical CT images were used in this study. The images were selected so that they are balanced with respect to patient age group and gender. The study was approved by Research Ethics Committee of Nara Medical University (study No. 2647). The target structures were the tibia and fibula bones and the 4 compartments, i.e. anterior, lateral, superficial posterior and deep posterior compartments, from knee-to-foot region. In this preliminary study, the ground-truth labels were manually segmented by engineering students and have not been verified by a medical expert. Figure 1 shows an example from the dataset. The matrix size was 512×512, the in-plane pixel size was varied among images, and slice thickness was 0.725 mm.

The Bayesian U-Net was trained for 10^5 iterations, with a batch size of 2 and a learning rate of 0.0001 using an Adam optimizer. At the training, the images were normalized within a window width of 300 HU and a window level of 40, and image augmentation, including horizontal/vertical flipping and affine transformations, was applied. The transformations comprised a rotation of [-10°, 10°], horizontal/vertical translation of [-25, 25] pixels, a shear of an angle within $[-\pi/8, \pi/8]$ and a scaling of a ratio within [65%, 135%]. At the testing, 10 Monte Carlo dropout samples for each slice were used to infer the output label. The training/testing were performed in a leave-one-patient-out cross-validation, where in each fold, 8 images were used for training, 1 image for validation and 1 image for testing. The experiments were performed using a cluster of 5 CPUs equipped with 5 GPUs (3×NVIDIA GeForce GTX 1080ti, 2×Titan RTX).

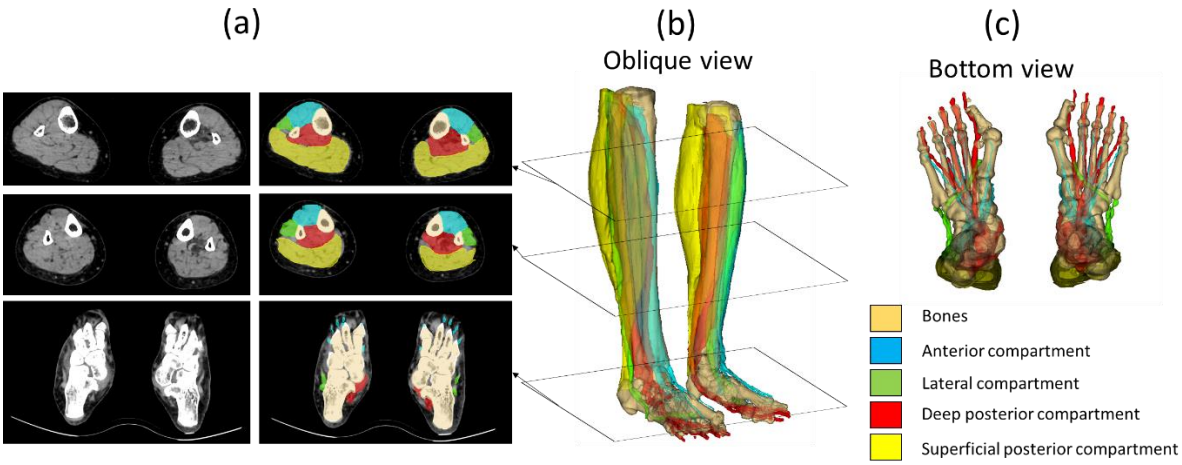


Fig. 1 Ground-truth labels of the muscles and bones in the lower leg. (a) Raw and overlaid CT images (WW: 300, WL: 40). (b) Surface rendering in oblique view. (c) Surface rendering in bottom view.

4. Results

Segmentation accuracy

Figure 1 shows the segmentation accuracy using Bayesian U-Net for the 10 CT images. The bone structures had the largest Dice of 0.979±0.002 and lowest ASD of 0.369±0.078. This aligns with the previous study [7], due to the large contrast between the bones (high density) and surrounding tissues. Among the muscle compartments, the deep posterior compartment had the smallest Dice of 0.800±0.034, whereas the superior superficial compartment had the largest Dice of 0.910±0.020.

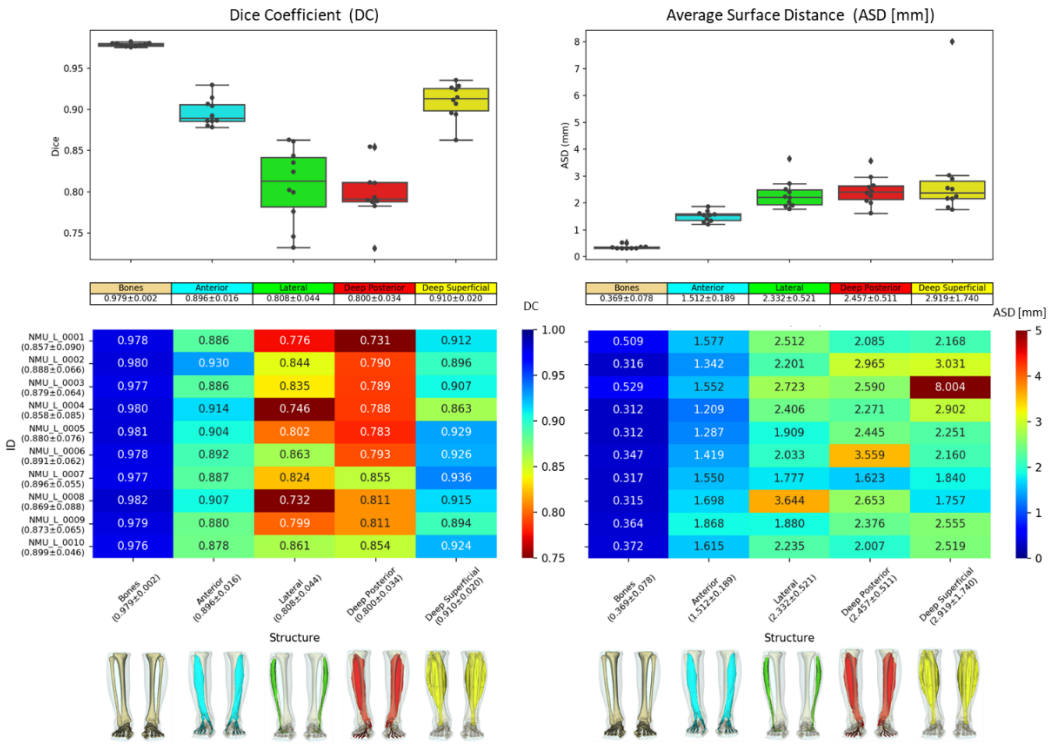


Fig. 2 Segmentation accuracy represented by Dice coefficient and average surface distance (ASD).

5. Discussion and Conclusions

In this paper, we reported preliminary results on the segmentation of lower leg bones and muscles in CT images using a Bayesian U-Net. Even though preliminary, the results show the possibility to use Bayesian U-Net to segment individual muscles from lower leg muscles. The deep posterior compartment has shown the lowest accuracy among the muscles. Indeed, this compartment had notably a low boundary contrast compared with the superficial posterior compartment, which may have led to human-induced errors in those labels. In addition to verifying the ground-truth labels by medical experts, future work will include analysis of the prediction uncertainty of the target structures and correlation with the segmentation accuracy.

Competing interests

The authors have no competing interests.

Acknowledgements

This research was supported by Japan Society for the Promotion of Science (JSPS) Grants-in-Aid for Scientific Research (KAKENHI) Number 19H01176 / 20H04550.

References

[1] J. D. Webb, S. S. Blemker, and S. L. Delp. 3d finite element models of shoulder muscles for computing lines of actions and moment arms. *Comp. Meth. Biomech. Biomed. Eng*., 17(8):829–837, 2014.

[2] S. S. Blemker and S. L. Delp. Three-dimensional representation of complex muscle architectures and geometries. *Annals of Biomedical Engineering*, Vol. 33, No. 5, pp. 661–673, 2005.

[3] C. Pons, B. Borotikar, M. Garetier, V. Burdin, D. B. Salem, M. Lampereur, and S. Brochard. Quantifying skeletal muscle volume and shape in humans using MRI: A systematic review of validity and reliability. *PLoS One,* 13(11): e0207847, 2018.

[4] F. Yokota, Y. Otake, M. Takao, T. Ogawa, T. Okada, N. Sugano, et al. Automated muscle segmentation from CT images of the hip and thigh using a hierarchical multi-atlas method. *Int J Comput Assist Radiol Surg*, 13(7):977–986, 2018.

[5] S. Dabiri, K. Popuri, E. M. C. Feliciano, B. J. Caan, V. E. Baracos, M. F. Beg. Muscle segmentation in axial computed tomography (CT) images at the lumbar (L3) and thoracic (T4) levels for body composition analysis. *Computerized Medical Imaging and Graphics*, Volume 75, Pages 47-55. Elsevier, 2019.

[6] N. Kamiya, A. Oshima, X. Zhou, H. Kato, T. Hara, T. Miyoshi, M. Matsuo, H. Fujita. Surface muscle segmentation using 3D U-Net based on selective voxel patch generation in whole-body CT images. Applied Sciences, vol 10, 4477, 2020.

[7] Y. Hiasa, Y. Otake, M. Takao, T. Ogawa, N. Sugano and Y. Sato. Automated Muscle Segmentation from Clinical CT Using Bayesian U-Net for Personalized Musculoskeletal Modeling. *IEEE Transactions on Medical Imaging*, vol. 39, no. 4, pp. 1030-1040, 2020.

機械学習を用いたハイパースペクトルによるK-ras遺伝子の識別

櫻井　千子[*1]　市川　盛規[*1]　吉谷　文孝[*1]　伊藤　隆晃[*1]

中矢　大輝[*1]　佐鳥　新[*1]　野呂瀬　朋子[*2]　大池　信之[*2]

要旨

ハイパースペクトルカメラは対象物から入射する電磁スペクトルの高い波長分解能で検出できる VNIR センサーであり，近年医療分野での応用が期待されている．本研究ではハイパースペクトルカメラを生体分野に応用し大腸癌の HE 標本を北海道衛星株式会社製の HSC1702 で撮影し遺伝子検査を実施することなく，K-ras 遺伝子の陽性，陰性の識別を実施する．

昭和大学藤が丘病院の大腸癌症例 37 例（陽性 20 例　陰性 17 例）の平均スペクトルデータに対し機械学習の一種であるランダムフォレストを用いることで 61%の識別精度が得られた．また，識別精度向上のため病理標本間での染色濃度の類似した検体（陽性 10 例　陰性 10 例）を選択し染色濃度のばらつきを減らし再度，機械学習を行ったところ識別精度が 77%に向上した．このことからハイパースペクトルカメラを病理組織診断に用いることにより K-ras 遺伝子の予測可能性が示唆された．

キーワード：ハイパースペクトルカメラ，大腸癌，ランダムフォレスト

1. はじめに

ハイパースペクトルカメラ（HSC）は、物質の色合い（スペクトル）情報と空間情報を同時に取得できる宇宙産業由来のセンサーである．近年医療分野での応用が期待されており人の目では識別することのできない物質の特定や分類に適している。

また、HSC は北海道衛星株式会社により開発された HSC1702 を使用している。本研究では大腸癌の病理標本を HSC で撮影し遺伝子検査を実施することなく、K-ras 遺伝子の陽性、陰性の識別を実施する。

なお、K-ras 遺伝子とは細胞増殖を促進するシグナルを細胞内で伝達するという役割を持つ遺伝子であり、K-ras 遺伝子の陽性・陰性は最適な化学療法を選択する際に重要な情報である。

*1 未来科学研究所　〒107-0052 東京都港区赤坂二丁目 10 番 14 号

*2 昭和大学藤が丘病院 認証病理診断科

2. ハイパースペクトルカメラを用いたK-ras遺伝子の識別

1）解析対象

昭和大学藤が丘病院より提供していただいた大腸癌症例 37 例（陽性 20 例, 400 細胞　陰性 17 例，340 細胞）のヘマトキシリン・エオジン染色（HE 染色）を行った病理組織標本を, 光学顕微鏡に取り付けられた HSC を使用して撮影したハイパースペクトルデータを使用する。なお，ハイパースペクトルデータは

350nm から 1050nm の反射強度を 5nm ごとに 141 バンド使用している。

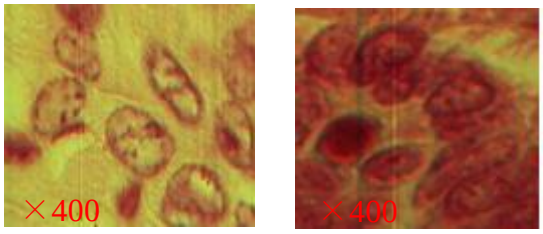


Fig.1 大腸癌標本の RGB 画像*400(左:陽性 右:陰性)

2）解析方法

2.1）ランダムフォレストを用いた K-ras 遺伝子の陽性陰性の識別

HSC で撮影された大腸癌の細胞核の平均スペクトルデータを取得し、K-ras 遺伝子の陽性，陰性のデータをトレーニングデータとテストデータ（7：3）に分割した。その後、機械学習手法の一種であるランダムフォレストを用いて解析を行い、混同行列、再現率、識別精度を評価した。

ランダムフォレストとは、アンサンブル学習法の 1 種であり、複数の決定木学習の予測値を使用し、多数決を行うことで最終的な目的変数を予測する。

2.2）識別精度向上のための解析

ランダムフォレストでは、決定木を作成するときの分岐による情報利得に基づいて，不純度の平均的な減少量として特徴量の重要度が評価できる。

識別に有意である波長域を特定するため、検体ごとのスペクトル強度の差に着目し陽性、陰性のスペクトル全体で一番分散が大きい波長領域において平均スペクトル強度上位 10 検体と平均スペクトル強度下位 10 検体を選択し、平均スペクトル強度上位 10 検体、下位 10 検体、それぞれにランダムフォレストを用いて解析を行った。

また、平均スペクトル強度上位 10 検体にランダムフォレストを用いて得られた特徴量重要度の高い波長帯と低い波長帯を選択し、正規化植生指標 NDVI を用いて算出した陽性、陰性の NDVI 値を、t 検定を使用し有意差の有無を確認した。

その後、平均スペクトル強度下位 10 検体を 400nm から 900nm の範囲を 100nm ごとに解析にかけ，解析にかける範囲によって生じる識別精度の変化を確かめた。

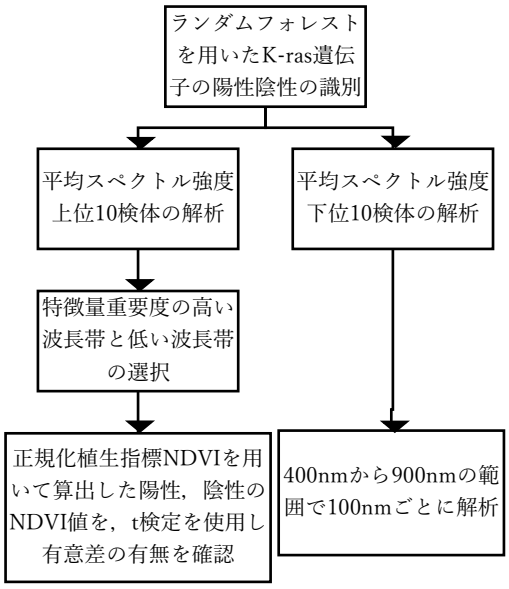


Fig.2 フローチャート

3）解析結果

3.1）ランダムフォレストを用いた K-ras 遺伝子の陽性陰性の識別

ランダムフォレストの解析の結果、トレーニングデータに対する正解率は 100%、テストデータに対する正解率は 61%であり、混同行列を Table.1 再現率、適合率を Table.2 に示す。

Table.1 混同行列（セル内は細胞数）

		予測結果	
		陽性	陰性
正解	陽性	89	26
	陰性	58	46

Table.2 再現率，適合率

	陽性	陰性
再現率（感度）	75%	44%
適合率（精度）	61%	61%

これらの結果より，K-ras 遺伝子の陽性，陰性のランダムフォレストを用いた識別は 60 数パーセントの低い予測率であった、この結果は 500nm から 650nm の間の波長帯に影響

する染色濃度によるスペクトルのばらつきが最も影響していると考えられる。

よって，染色濃度が類似したスペクトルをグループ分けするため、次の識別精度の向上のための解析では分散を使用した。

3.2）識別精度向上のための解析

大腸癌の細胞核の平均スペクトルデータの分散を取り一番ばらつきが大きい波長帯は620nm の波長帯であった。

また、620nm の波長帯で平均スペクトル強度上位 10 検体と平均スペクトル強度下位 10 検体に分けそれぞれにランダムフォレストを用いて解析を行った。

解析の結果、平均スペクトル強度上位 10 検体はトレーニングデータに対する正解率 100%、テストデータに対する正解率 76%であった、混同行列を Table.3 再現率、適合率を Table.4 に示す。

Table.3 混同行列（セル内は細胞数）

		予測結果	
		陽性	陰性
正解	陽性	44	10
	陰性	19	47

Table.4 再現率，適合率

	陽性	陰性
再現率（感度）	81%	71%
適合率（精度）	70%	82%

平均スペクトル強度下位 10 検体の解析の結果は、トレーニングデータに対する正解率 100%、テストデータに対する正解率 65%であった、混同行列を Table.5 再現率、適合率を Table.6 に示す。

Table.5 混同行列（セル内は細胞数）

		予測結果	
		陽性	陰性
正解	陽性	13	39
	陰性	3	65

Table.6 再現率，適合率

	陽性	陰性
再現率（感度）	25%	95%
適合率（精度）	81%	62%

平均スペクトル強度上位 10 検体のランダムフォレスト解析による特徴量重要度と波長帯の順位を Table.7 に示す。

Table.7 特徴重要度と波長帯の順位

順位	波長帯	特徴量重要度
1	580	0.0188
2	560	0.0185
3	585	0.0170
4	575	0.0159
…	…	…
134	720	0.00267
135	780	0.00216
136	775	0.00202
137	765	0.00174
…	…	…
141	600	0.00140

特徴量重要度の上位には 500nm~600nm の波長帯が、下位には 700nm~800nm の波長帯が多く見受けられる。よって，500nm~600nm と 700nm~800nm の波長域を選択し正規化植生指標 NDVI によって値を算出その後陽性、陰性の NDVI 値に t 検定を使用し有意差の有無を確認した。結果として、t 値は 2.43E-11 で有意差があると確認することができた。

また、平均スペクトル強度下位 10 検体を 400nm から 900nm の範囲で 100nm ごとずつランダムフォレストによる解析を行った。正解率を Table.8 に示す。

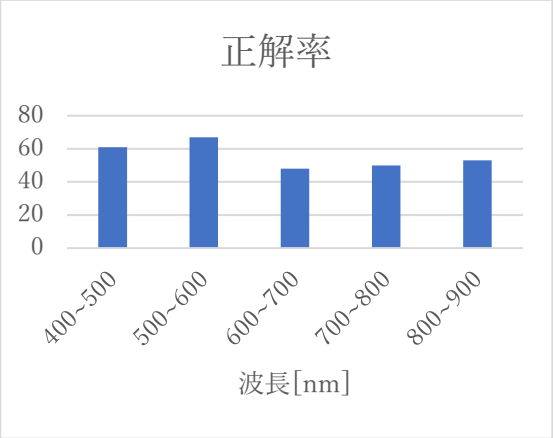


Fig.3 100nm ごとによる正解率

500nm から 600nm が正解率 67%で最も正解率が高く 600nm から 700nm が正解率 48%で最も正解率が低かった。

よって、ランダムフォレストの平均スペク

トル強度上位 10 検体で得た特徴量重要度の大きい波長帯と下位 10 検体での識別率が高い波長帯が一致した。

このことから、K-ras 遺伝子の陽性、陰性の識別において、500nm から 600nm は識別に有用な波長域である可能性が示された。

3. まとめ

本研究では、ハイパースペクトルカメラを用いて計測したハイパースペクトルデータをもとに、機械学習の一種であるランダムフォレストを用いて大腸癌の K-ras 遺伝子の陽性、陰性の識別を行った。

識別精度向上のために染色濃度の類似した検体を選択し、機械学習を行ったところ識別精度が向上した。

平均スペクトル強度上位 10 検体をランダムフォレスト解析にかけると、特徴量重要度の大きい波長帯は 500nm から 600nm であり、小さい波長帯は 700nm から 800nm であった。

平均スペクトル強度下位 10 検体を解析にかけると, 500nm から 600nm が最も正解率が高く、600nm から 700nm が最も正解率が低いという結果を得た、また、特徴量重要度大きい波長帯と識別率の高い波長帯が一致した。

よって、今後の展望としてノイズとなる染色のばらつきや 500nm から 600nm の波長帯を強調することで、識別精度を向上させることができると考える。

利益相反の有無

なし.

文　献

[1]Daiki Nakaya,“Machine Learning Classification of Colorectal Cancer Using Hyperspectral Images,”Journal of the color science association of japan, Vol. 41(2017), No.3

[2 山口雅弘:“マルチスペクトル病理画像解析技術“, 映像情報メディア学会誌,Vol.69,No.5(2015)

K-ras genes using machine learning and hyperspectral analysis

Chiko SAKURAI[*1], Seiki ICHIKAWA[*1], Fumitaka YOSHIGAI[*1], Takaaki ITO[*1], Daiki NAKAYA[*1], Shin SATORI[*1], Tomoko NOROSE[*2], Nobuyuki OIKE[*2]

*1 Future Science Institute

*2 Showa University Fujigaoka Hospital Diagnostic Pathology

A hyperspectral camera is a VNIR sensor capable of detecting the electromagnetic spectrum incident from an object with high wavelength resolution, and recently it is expected to be applied to the medical field. In this study, we apply the hyperspectral camera to the field of biology to identify positive and negative K-ras gene specimens of colorectal cancer by using the HSC1702 of Hokkaido Satellite Co.

A random forest, a type of machine learning technique, was used for the average spectral data of 37 colorectal cancer cases (20 positive and 17 negative) at the Showa University Fujigaoka Hospital, and the discrimination accuracy of 61% was achieved. To improve the discrimination accuracy, we selected specimens with similar staining concentrations between pathological specimens (10 positive and 10 negative specimens) to reduce the variation in staining concentrations, and performed machine learning again to improve the discrimination accuracy to 77%. This suggests the predictability of the K-ras gene by using a hyperspectral camera for histopathological diagnosis

Translated with www.DeepL.com/Translator (free version)

Key words: A hyperspectral camera, Colorectal cancer, A random forest

機械学習を用いたハイパースペクトルによる肺癌の分類

Classification of Lung Cancer Using Machine Learning with Hyperspectral imaging

江崎雅隆[*1]　神谷健太郎[*1]　小倉翔悟[*1]　伊藤隆晃[*1]

中矢大輝[*1]　佐鳥新[*1]　井上雅文[*2]

要旨

ハイパースペクトルカメラは対象物から入射する電磁スペクトルを高い波長分解能で検出できる VNIR センサーであり,近年医療分野での応用が期待されている。本研究ではハイパースペクトルカメラを生体分野に応用し、肺癌 HE 標本を北海道衛星株式会社製の HSC1702 で撮影、肺癌の扁平上皮癌と腺癌の機械学習による識別を試みた.

新宿メディカルセンターの肺癌症例 68 例（腺癌 42 例、扁平上皮癌 26 例）のハイパースペクトルデータに対し、ランダムフォレスト等の学習器を組み合わせることによって一致率 84.4%の識別率を得られた.

そこでつぎに、細胞診検体を撮影し、細胞塊内のがん細胞の予測を試みた。結果として、病理医のあらかじめ指定した細胞塊に関して画像上で正しく識別出来ていることを確認した.このことから、ハイパースペクトルカメラの病理細胞診断における応用可能性が示唆された.

キーワード：ハイパースペクトル、肺癌、腺癌 擦過細胞診検体、機械学習

1. はじめに

ハイパースペクトルカメラは対象物から入射する電磁スペクトルを高い波長分解能で検出できる VNIR センサーであり，近年医療分野での応用が期待されている．[1] 本研究ではハイパースペクトルカメラを生体分野に応用し，東京新宿メディカルセンター病理科から提供していただいた肺癌 HE 標本を,通常の RGB バンドではなく，141 バンドの北海道衛星株式会社製 HSC1702 を使用して 400 倍の倍率で撮影し，肺癌の扁平上皮癌と腺癌の機械学習による識別を試みた．その際に複数の機械学習手法を組み合わせて，ハイパースペクトルカメラを用いた新たな分類器の可能性について評価した．また，同施設から提供していただいた腺癌と扁平上皮癌の擦過細胞診検体に対してそれぞれ HSC1702 で撮影した倍率 20 のハイパースペクトルデータを使用した．今回はデータ量の都合上，そのうちの 1 ブロックを用いて教師なし学習によるクラスタリング解析を実施し，識別の可能性を評価した．

*1 一般財団法人未来科学研究所

*2 日本独立行政法人 地域医療機能推進機構 東京新宿メディカルセンター 病理科

2. 肺癌 HE 標本を使用した扁平上皮癌と腺癌の機械学習による識別

1）解析対象

解析には新宿メディカルセンターから提供していただいたヘマトキシリン・エオジン染色（HE 染色）を行った肺癌症例 68 例（腺癌 42 例，扁平上皮癌 26 例）の標本を，顕微鏡に取り

付けられた HSC1702 で撮影したハイパースペクトルデータを使用した．また，それらのデータは 350nm から 1050nm の反射強度を 5nm ごとに 141 バンドを使用した．また比較検証のために 15 バンドのマルチスペクトルデータも使用した．

2）解析方法

癌細胞は，腫大した細胞核により細胞質が観察困難なため、解析には細胞核をそれぞれ手動で抽出し，その平均スペクトルを使用した．

本研究では，2 種類の肺癌の分類において，サポートベクターマシン（SVM），ランダムフォレスト（RF），ロジスティック回帰分析（LR），K 近傍法（KNN）,Light GBMClassifier(LGBM)[2]の５つの機械学習モデルを使用した．

また，他の 3 つのモデル（SVM， RF，LGBM）の確率的予測値を集約したデータセットである 決定木分類器（DTC）を採用し，クロスバリデーション法を用いてモデルの性能を調べた． その際，まず前述した 5 つのモデル（SVM，RF，LGBM，LR，KNN）で識別を第一に行い，それらから得られた予測結果を用いて，DTC でさらに識別をした場合の識別率を評価した．

3）解析結果

マルチスペクトルデータ（15band）とハイパースペクトルデータ（141band）のそれぞれの機械学習モデルによる一致率（Accuracy）の結果を Table.1 に示す．

Table.1 一致率

	15band	141band
SVC(rbf)	0.7481	0.8049
RF	0.7457	0.837
LightGBM	0.6469	0.7333
LR	0.6988	0.7531
KNN	0.6988	0.7556
DTC3	0.7531	0.8444

以上の結果より，すべての機械学習モデルにおいてハイパースペクトルデータがマルチスペクトルデータの識別率を上回る傾向がみられた．このことから，ハイパースペクトルカメラが有効であることがうかがえる．また，高い識別率を示したのは SVC，RF，DTC の 3 つであり，そのうち最も高かったのが DTC の 84.4%である．このことから、複数の機械学習モデルを組み合わせて，ハイパースペクトルデータを用いた分類が有効である可能性があると考えられる。DTC から得られる予測結果の混同行列を Table.2 に示す．

Table.2 混同行列:DTC

DTC		予測結果	
		adeno	scc
正解	adeno	109	52
	scc	11	233

3．腺癌と扁平上皮癌の擦過細胞診検体の識別

本研究では，腺癌と扁平上皮癌の擦過細胞診検体を撮影したハイパースペクトルデータをクラスタリングし，スクリーニングが可能であるかの基礎検討のための解析を行った．また，擦過細胞診とは病変部を綿棒やブラシ，ヘラ等で擦って細胞を採取した細胞の腫瘍性病変を疑う所見を調べるものである．

1）解析対象と方法

提供していただいた腺癌（adeno）と扁平上皮癌（SCC）を 200 倍の倍率で HSC が取り付けられた光学顕微鏡でスライドガラスをスライドさせつつ撮影し，得られた全体のハイパースペクトル画像を使用した．今回は，画像のデータ量が膨大なため，マークされている箇所を含む 1 ブロック（解像度 1720×8000）を用いてそれぞれ教師なし学習であるクラスタリングでスペクトルにおいて性質の近いグループに分類した．クラスタリングは教師データを必要とせず，より客観的な分類が可能であるが，クラスタ数（分類数）のみ指定する必要がある．クラスタ数は４つと６つのものを使用し，分類され

たスペクトルで平均した波形を評価する．

2）腺癌の解析結果

ハイパースペクトルデータを RGB に変換した画像を Fig.1，クラスタ数 4 の画像を Fig.2，クラスタ数 6 の画像を Fig.3，クラスタ数 4 のクラスタ別平均スペクトルのグラフを Fig.4 として以下に示す．

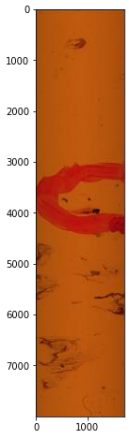


Fig. 1

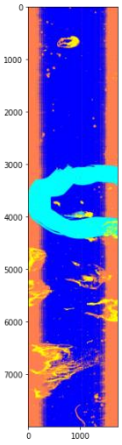


Fig. 2

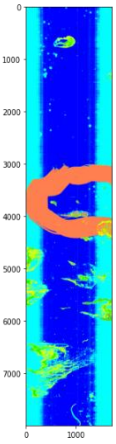


Fig. 3

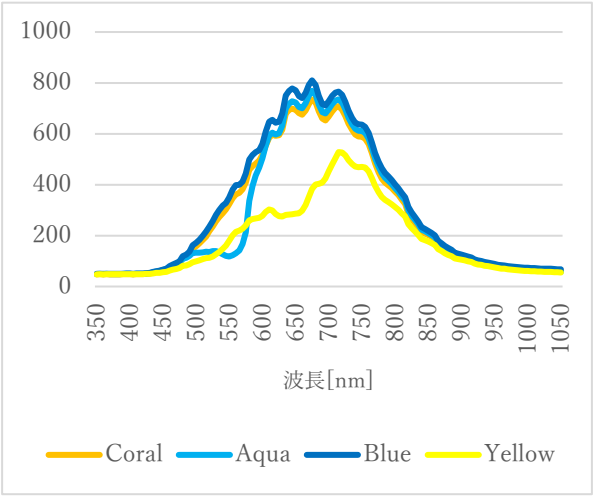


Fig. 4 Adeno_4 クラスタ別平均スペクトル

図と平均スペクトルから Blue，Coral は光源，Aqua はマーカー，Yellow が組織の吸収スペクトルであり，625nm～655nm においてスペクトルの強い吸収が見られる．以下にクラスタ数 6 のクラスタ別平均スペクトルのグラフを Fig.5 として示す．

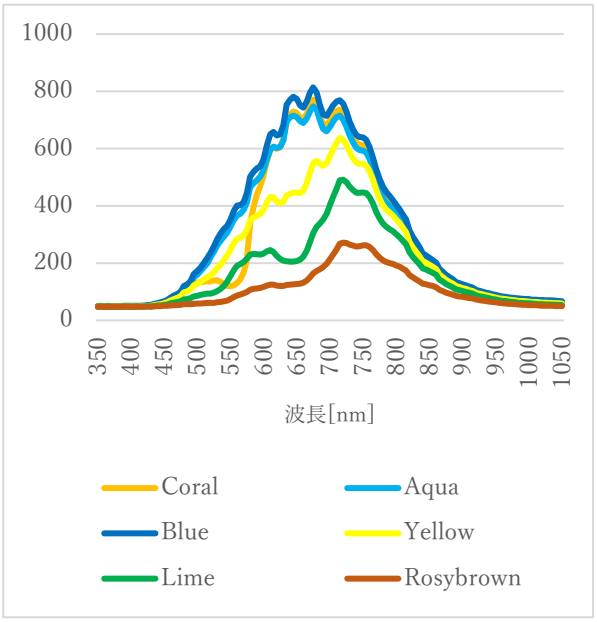


Fig. 5 adeno_6 クラスタ別平均スペクトル

Yellow，Lime，Rosybrown が組織のスペクトルであり，特に Lime において 630nm～655nm に強い吸収スペクトルがあることがわかった．クラスタ数の増加により，より細かい分類が可能となることが評価された．

3）扁平上癌の解析結果

ハイパースペクトルデータを RGB に変換した画像を Fig.6，クラスタ数 4 の画像を Fig.7，クラスタ数 6 の画像を Fig.8，クラスタ数 4 のクラスタ別平均スペクトルのグラフを Fig.9 として以下に示す．

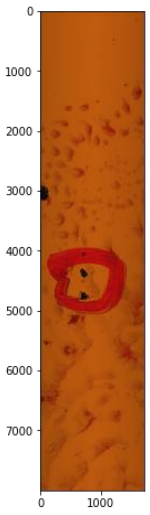


Fig. 6

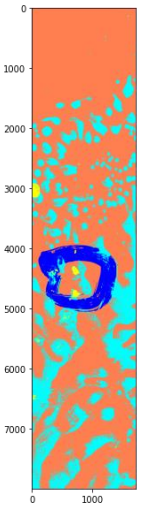


Fig. 7

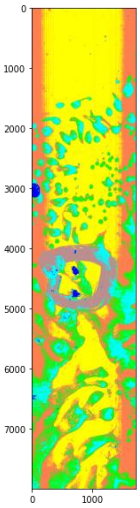


Fig. 8

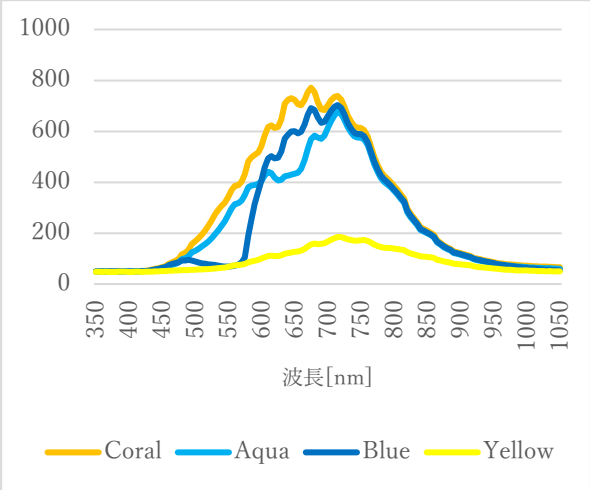


Fig. 9 SCC_4 クラスタ別平均スペクトル

Yellow，Aqua が組織部分のスペクトルであり、Aqua において 630~655nm に強い吸収スペクトルが見られる．クラスタ数 6 のクラスタ別平均スペクトルのグラフを Fig.10 として以下に示す．

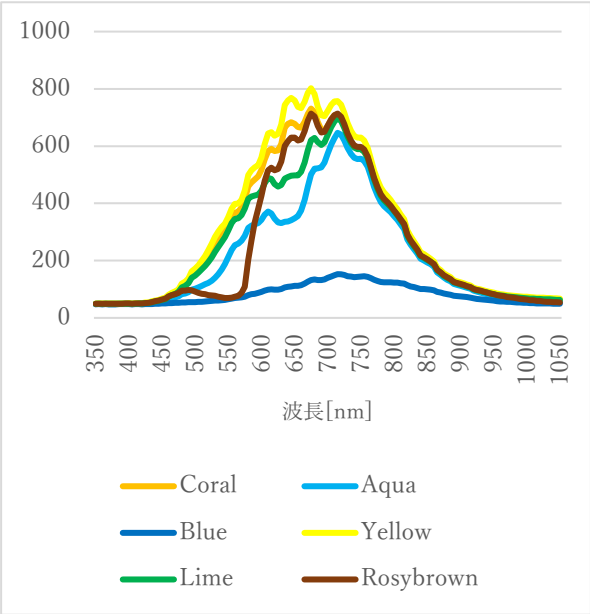


Fig. 10 SCC_5 クラスタ別平均スペクトル

Blue，Aqua，Lime が組織部位のスペクトルである。Aqua，Lime において 630nm~655nm に強い吸収スペクトルが見られる．

4. まとめ

本研究では，ハイパースペクトルカメラを用いて計測したデータをもとに，腺癌と扁平上皮癌の分類とそれらの擦過細胞診検体の画像から得たデータから癌の識別を行った．HE 標本を用いた検体の複数の機械学習による識別では，ハイパースペクトルデータを用いた新しい分類器を作成することが有効である可能性があると考えられる．今回のモデルは実用的には十分ではなかったが，より多くのデータを追加することでモデルが洗練され，機械学習法を実用的な診断に利用になることが期待される．

また，擦過細胞診検体を撮影しクラスタリング解析を行ったところ，目視で明らかにとらえられる色の違いはクラスタリングにより識別がされており，スクリーニングに使用できる可能性があることがわかった．

利益相反の有無

なし

文　献

[1]Daiki Nakaya,“Machine Learning Classification of Colorectal Cancer Using Hyperspectral Images,”Journal of the color science association of japan, Vol. 41(2017), No.3

[2] Guolin Ke1etal.:LightGBM: A Highly Efficient Gradient Boosting Decision Tree(NIPS2017)

Classification of Lung Cancer Using Machine Learning with Hyperspectral imaging

Masataka ESAKI[*1], Kentaro KAMIYA [*1], Shogo OGURA[*1], Takaaki ITOU[*1],
Nakaya DAIKI[*1], Shin SATORI[*1], Masafumi INOUE[*2]

*1 Future Science Institute
*2 Tokyo Shinjuku Medical Center

A hyperspectral camera is a VNIR sensor that can detect the electromagnetic spectrum incident from an object with high wavelength resolution, and is expected to be applied to the medical field. In this study, we applied the hyperspectral camera to the biomedical field and tried to discriminate between squamous cell carcinoma and adenocarcinoma of lung cancer with machine learning by using the HSC1702 manufactured of Hokkaido Satellite Co. A discrimination rate of 84.4% was obtained by combining hyperspectral data from 68 lung cancer cases (42 adenocarcinomas and 26 squamous cell carcinomas) at Shinjuku Medical Center with several machine learning models such as random forests.
And then we attempted to take cytological specimens and try to predict the cancer cells in the cell masses. As a result, we confirmed that the cytological specimens were correctly identified on the images as to the cell masses specified in advance by the pathologist. These results suggest that the hyperspectral camera can be used for pathological cytological diagnosis.

Key words: Hyperspectral, lung cancer, adenocarcinoma scraping cytology specimen, machine learning machine

広範囲の隣接関係を考慮した グラフニューラルネットワークを用いた 腹部動脈血管名自動命名の検討

日比　裕太[*1]　林　雄一郎[*1]　北坂　孝幸[*2]　伊東　隼人[*1]

小田　昌宏[*1]　三澤　一成[*3]　森　健策[*1,4,5]

要旨

本稿では，3 次元腹部 CT 像から抽出された腹部動脈領域に対して Graph Neural Network（GNN）を用いた血管名自動命名についての検討を行ったので報告する．腹部動脈領域に対して機械学習を用いた血管名自動命名を行う手法はこれまでにもいくつか提案されてきた．また，近年ではグラフ構造に対する機械学習が盛んに行われており，その有用性が示されているため，GNN を用いた機械学習による血管名自動命名について検討した．一般的な GNN では一つの隣接関係だけを学習することしかできないが，広い範囲の隣接関係を同時に使うことで高精度の自動命名が実現可能であると考えた．本稿では範囲の異なる複数の隣接関係を用いて学習を行うことができる MixHop と一般的な GNN である Graph Convolutional Network を使用し，CT 像 100 症例に対して 10 分割交差検定による実験と比較を行った．実験の結果，MixHop は全症例のノードを用いた評価では 85.2%，データセットごとの評価では最高で 88.2%の精度で命名することができた．

キーワード：血管，グラフニューラルネットワーク，3 次元 CT 像，解剖学的名称認識，血管構造解析

1. はじめに

近年広く行われている腹腔鏡手術では，低侵襲であるが視野が狭く，個人差が大きな血管構造を正確に把握することは難しい．そのため術前に 3 次元 CT 画像から血管構造を把握することが重要である．これまでにも腹部動脈領域に対して自動命名を行う手法がいくつか提案されてきた．鉄村らは血管領域を木構造と捉え，条件付確率場を用いた機械学習により自動命名を行った［1］．本稿では血管木構造をグラフと捉え，グラフニューラルネットワーク(Graph Neural Network, GNN)を用いた血管名自動命名手法について検討する. GNN を使った既存の研究として既存研究[2]などがある．また，グラフ上のノードの特徴量はより広い範囲を集約したものを利用することでネットワークの表現能力が上がると考えられるため，広範囲の隣接関係を同時に扱うことができる GNN と 1 つの隣接関係のみ学習する一般的な GNN の 2 つの手法を用いて実験及び比較を行う．

2. 手法

本手法では，はじめに従来手法［1］と同様に，3 次元腹部 CT 像から腹部動脈領域の木構造の構築を行い，血管の直径や長さ，血管以外の臓器との位置関係を特徴量として抽出する．

学習と推定には一般的な GNN である vanilla GCN [3]と複数の隣接関係を扱うことができる Mixhop [4]を用い，その結果を比較する．一般的

*1 名古屋大学大学院情報学研究科
〔〒464-8601 名古屋市千種区不老町〕
e-mail: yhibi@mori.m.is.nagoya-u.ac.jp
*2 愛知工業大学情報科学部
*3 愛知県がんセンター
*4 名古屋大学情報基盤センター
*5 国立情報学研究所　医療ビッグデータ研究センター

なグラフ畳み込みでは，グラフ上の注目ノードとその隣接ノード（1-hop）の特徴量を集約し，フィルタをかけることで次の層のノードの特徴量とする．MixHop では 3 つの隣接関係（0-hop, 1-hop, 2-hop）にあるノードの特徴量を集約し，それらを concatenate することで次の層への出力とする．本稿では，血管木構造を重みなし無向グラフと捉え各手法を適用する．このグラフでは，ノードを各血管枝に，ノードの特徴量を血管枝が持つ特徴量に，ノード間の接続関係を血管枝の接続関係に対応させる．入力には，ノードの接続関係と各ノードの特徴量及び学習用の正解ラベルデータを使用する．ネットワークは各ノードについてその推定ラベルを出力することで自動命名を行う．

3. 実験と結果

本稿では，腹腔鏡手術を行う際に重要となる 22 種類の腹部動脈を主な自動対応付けの対象とする．学習と推定には vanilla GCN と MixHop を使用し，ともに学習率は 0.01，結果は 1000epoch 学習後のものとする．自動命名には 3 次元腹部 CT 像から抽出した血管領域画像 100 症例を対象に 10 分割交差検定法を用いて実験を行った．精度の評価には F 値を使用した．実験の結果，MixHop の精度が vanilla GCN を全体で約 5 ポイント上回り，vanilla GCN では命名精度 0%になってしまった血管もあるが，MixHop にはなかった．各手法による自動命名結果を表 1 と表 2 に，自動命名結果の例を図 1 に示す．図 1(a)では，命名失敗血管が多く，正解データと比べてラベル数が少なくなっている．(b)ではおおむね命名に成功してはいるものの，正解データには存在しない PHA が命名されているほか，左脚部分に胃の付近の血管である RGEA が命名されている．

4. 考察

vanilla GCN では 1-hop の隣接関係のみを用いて学習していることから，注目ノードの付近に多く存在する血管の特徴量が必然的に多く集約されたと考える．そのために数が少ない血管の特徴を学習できず，命名に失敗した一方で，より遠い 2-hop の隣接関係を加えて学習した MixHop ではより多くのノードの特徴量の集約が行えたので数が少ない血管の特徴も学習できたと考える．この実験の結果から，血管の自動命名において GNN を用いる際には隣接関係の広さは重要であるといえる．しかしながら，RGEA が脚の血管に命名されるというような誤りもあった．このことから，広い範囲の特徴を集約することによって特徴量が平均化されやすくなり，位置としては離れた血管が誤って命名される原因になったと考える．

5. むすび

本稿では，広範囲の隣接関係を考慮した GNN を用いて 3 次元腹部 CT 像から抽出した動脈領域に対する自動命名手法の検討を行うと同時に一般的な GNN との比較を行った．その結果，広い隣接関係を用いた手法の方が高精度の命名に成功したことから，隣接関係の広さは GNN による血管名自動命名において重要であることが示された．

謝辞

日頃から熱心に御討論頂く名古屋大学森研究室の諸氏に感謝する．本研究の一部は科研費 17H00867，17K20099，26560255，26108006，AMED 19lk1010036, JP20lk1010036, 堀科学芸術振興財団，JSPS 二国間交流事業によった．

利益相反の有無

なし．

文　献

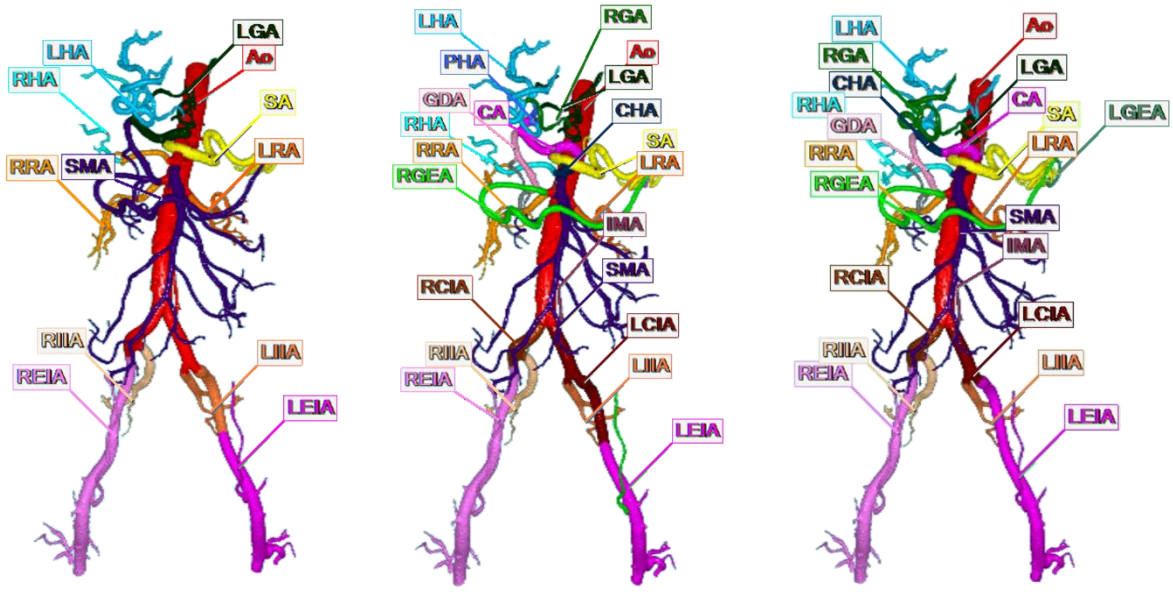
［1］ 鉄村悠介，林雄一郎，小田昌宏 他：機械学習を用いた腹部動脈血管名自動命名におけるデータ拡張法の適用に関する検討．信学技報 **118** : 191-196, 2019

［2］ 日比 裕太， 林 雄一郎， 北坂 孝幸 他：グラフ畳み込みニューラルネットワークによる腹部動脈血管名自動命名におけるデータ拡張による精度改善．JSCAS: **21**,226-227, 2019

［3］ Thomas N. Kipf and Max Welling: Semi-Supervised Classification with Graph Convolutional Networks. ICLR, 2017

［4］ Sami Abu-El-Haija, Bryan Perozzi, Amol Kapoor, et.al: Mixhop: Higher-order graph

convolutional architectures via sparsified neighborhood mixing. ICML, 2019



(a) vanilla GCN (b) MixHop (c) 正解データ

図 1 血管名自動命名の例（同じ色は同じラベルを示している.）

表 2 MixHop と vanilla GCN による各データセットの自動命名精度の比較

（すべてのデータセットにおいて MixHop が vanilla GCN を上回っている.）

データ セット	1	2	3	4	5	6	7	8	9	10	全体
MixHop (%)	**82.4**	**83.1**	**84.3**	**86.8**	**86.4**	**86.7**	**83.2**	**84.0**	**85.2**	**88.2**	**85.2**
vanilla GCN (%)	75.1	76.4	78.7	80.6	78.9	81.1	78.4	81.6	79.5	81.8	79.5

表 2 MixHop と vanilla GCN による各血管の自動命名精度の比較

（vanilla GCN では 9 種の血管の命名精度が 0%になっている.）

血管名	Ao	CA	CHA	PHA	LHA	RHA	GDA	LGA	RGA	LGEA	RGEA
MixHop (%)	**95.8**	**91.1**	**81.8**	**60.3**	**68.4**	**81.6**	**80.4**	64.8	**43.2**	**59.9**	**50.6**
vanilla GCN(%)	67.4	0.0	0.0	0.0	62.0	64.5	0.0	**66.5**	0.0	0.0	0.0
血管名	SA	SMA	IMA	LRA	RRA	LCIA	RCIA	LEIA	REIA	LIIA	RIIA
MixHop (%)	86.4	**86.5**	**35.6**	91.6	92.3	**70.5**	**70.3**	92.1	92.6	84.9	88.3
vanilla GCN(%)	**87.3**	84.1	0.0	**95.6**	**96.1**	0.0	0.0	**94.6**	**94.7**	**89.3**	**91.0**

Study on Automated Anatomical Labeling of Abdominal Arteries Using Graph Neural Network Considering Wide Range of Adjacencies

Yuta HIBI[*1], Yuichiro HAYASHI [*1], Takayuki KITASAKA[*2],
Hayato ITOH[*1], Masahiro ODA[*1], Kazunari MISAWA[*3] and Kensaku MORI[*1, 4, 5]

*1 Graduate School of Informatics, Nagoya University
*2 Faculty of Information Science, Aichi Institute of Technology
*3 Aichi Cancer Center Hospital
*4 Information Technology Center, Nagoya University
*5 Research Center for Medical Bigdata, National Institute of Informatics

In this paper, we report an automated anatomical labeling method using graph neural networks for abdominal artery regions extracted from 3D abdominal CT images. Several automated blood vessel labeling techniques for the abdominal artery region using machine learning have been proposed. In recent years, machine learning for graph structures has been actively studied, and its usefulness has been shown. Among them, automated anatomical labeling by machine learning using Graph Neural Network (GNN) was examined. Conventional GNN learn, only one adjacency relation on a graph structure. We thought that high precision automated anatomical labeling could be realized by using a wide range of adjacency relations at the same time. In this paper, we compared the MixHop, which can perform learning using multiple adjacency relations, and the Graph Convolutional Network, which is a conventional GNN. In the experiment, 10-fold cross-validation was performed on 100 CT images. As a result of the experiment, the MixHop achieved an accuracy of 85.2% in the evaluation using the nodes in all cases and a maximum accuracy of 88.2% in the evaluation for each dataset.

Key words: Blood vessel, Graph neural network, CT volume, Anatomical names recognition, Blood vessel structures analysis

歯科パノラマX線画像における分岐を含むSSDによる1クラス検出と多クラス検出の結果の統合による歯列の認識手法

森下 拓水[*1] 村松 千左子[*2] 周 向栄[*3] 髙橋 龍[*4] 林 達郎[*4] 西山 航[*5]
原 武史[*3] 有地 淑子[*6] 有地 榮一郎[*6] 勝又 明敏[*5] 藤田 広志[*3]

要旨

本研究の目的は，深層学習を用いて歯科パノラマX線画像を解析し，歯科医師の診断補助に寄与することである．その初期段階として，100 症例を用いて画像内の歯牙の検出と歯種の分類を行った．Single Shot Multibox Detector (SSD) に分岐を与えたネットワークを提案し，歯種を区別しない1クラス検出と，16クラス（上顎と下顎で区別した中切歯，側切歯，犬歯，第一小臼歯，第二小臼歯，第一大臼歯，第二大臼歯，第三大臼歯）検出を行った．また，これらの結果を用いて統合処理などの後処理を行った．結果，検出率98.8%, 誤検出数0.33個/画像, 歯種32分類についての分類率92.4%となり, 提案手法の有効性が示された．

キーワード：歯科パノラマX線画像，深層学習，コンピュータ支援診断，Single Shot Multibox Detector (SSD)

1. はじめに

歯科パノラマX線撮影装置は日本の歯科医院全体の98%が保有しており，歯科領域において最も普及している撮影装置である．しかし，歯科パノラマX線画像には口内の全ての歯牙についての情報と，歯周病などの歯牙以外の病変の情報が描出されるため，歯科医師が全ての歯牙に関して正確な画像診断をするには多大な時間と労力が必要となる．本研究の目的は，歯科パノラマX線画像から読影に必要な情報を自動抽出し，歯科医師の診断の補助に寄与することである．本稿では，歯科パノラマX線画像内の歯牙の検出と，検出された歯牙の歯種の分類を行った．

Attention Branch Network[1]では，attention map と物体検出をマルチタスク学習し，attention map を特徴マップに合わせることで，自然画像において認識性能を向上させた．医用画像への適用は少ないが，ネットワークの分岐によるマルチタスク学習と，分岐の出力を特徴マップに合わせることは，情報量を増やすという観点から，様々な症例が存在する医用画像に有用であると考えた．本研究では，attention map ではなく32の各歯牙の有無を表す配列を作成して特徴マップに掛け合わせる．Single Shot Multibox Detector (SSD)[2]をベースに提案モデルを構築し，SSDと比較した．

*1 岐阜大学大学院自然科学技術研究科知能理工学専攻
〔〒501-1194 岐阜県岐阜市柳戸 1-1〕
e-mail: tmori@fjt.info.gifu-u.ac.jp
*2 滋賀大学データサイエンス学部
*3 岐阜大学工学部電気電子・情報工学科
*4 メディア株式会社
*5 朝日大学歯学部口腔病態医療学講座歯科放射線学分野
*6 愛知学院大学歯学部歯科放射線学講座

2. 方法

2. 1. 提案モデル

歯種の分類問題において，対象歯牙の周囲の歯牙は重要な情報である．そこで提案モデルでは，SSD で得られる特徴マップを用いて 32 の各歯牙の有無を 0，1 で表現する 1 次元の配列を作成し，それを特徴マップに掛け合わせる．これにより，全ての歯牙の有無の情報を加味した特徴マップからの検出・分類を行う（図 1）．作成する配列について，本来は歯科医師による歯科所見の記録を意味するため非常に簡略化した概念ではあるが，本研究ではこれを Dental chart と呼称する．Dental chart には正解を与え，誤差関数に 2 値クロスエントロピー（Binary Cross Entropy：BCE）を用いて学習を行う．

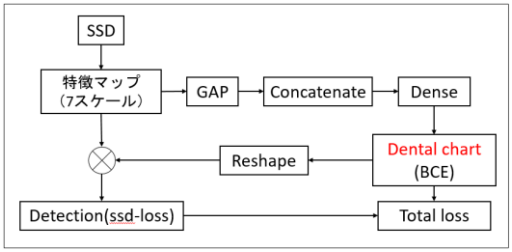


図 1 提案モデル

2. 2. 後処理

2. 2. 1. 概要

SSD と提案モデルによって検出された 2D-Bounding Box（以下：検出 box）に対して，統合処理，検出した中切歯を基準とした歯種のさらなる分類（歯種 32 分類），確信度が高い検出 box を基準に歯牙の並びに関する一般的な位置関係を利用したナンバリングの修正を行う．これらの処理には，検出 box の座標，クラス，確信度を用いる．

2. 2. 2. 統合処理

1 クラス検出（クラスは「歯牙」のみ）と，16 クラス検出（クラスは，上下顎を区別した「中切歯，側切歯，犬歯，第一小臼歯，第二小臼歯，第一大臼歯，第二大臼歯，第三大臼歯」）の結果を統合する．本研究では，対象の歯牙の周辺部の特徴を考慮した多クラス検出による歯種分類を行った．しかし，16 クラス検出よりも 1 クラス検出の方が高い検出性能を示したため，高い検出性能を維持しながら歯種分類を行うために本処理を行う．

まず，両方の検出 box を全て採用して重複している検出 box を削除する．ある 2 つの検出 box の一致率（IoU）が 0.6 以上のとき，これらは重複していると見なし，異なるネットワーク出力の場合は 16 クラス検出の検出 box を残して 1 クラス検出の検出 box を削除する．それぞれのネットワークでの重複検出 box の場合は検出 box の確信度が高い方を残す．

2. 3. 試料画像

朝日大学で撮影された歯科パノラマ X 線画像 100 症例を用いる．画像の詳細は，撮影装置：QRmaster-P，空間分解能：0.1 mm，画像サイズ：3024×1573 pixels である．深層学習には，歯科パノラマ X 線画像から歯列範囲を抽出し，長辺が 512 pixels になるように縮小した後，正方形になるようにパディングした，512×512 pixels の画像を入力する．

2. 4. 評価方法

歯科パノラマ X 線画像 100 症例のうち，75 症例を学習データ，25 症例をテストデータとして 4 fold の交差検証を行う．

歯牙の検出に関する評価方法として，検出 box を成功（正解 box と IoU が 0.6 以上の検出 box）と誤検出（成功でない検出 box）に分ける．また，多クラス検出の検出・分類性能の指標に，検出率（歯種の予測を問わず検出された歯牙の総数*100/データセット内の歯牙の総数），歯牙別検出率（正しい歯種で検出された歯牙の総数*100/データセット内の歯牙の総数），分類率（歯牙別検出率*100/検出率）を用いる．

3. 結果

3. 1. SSD と提案モデルの比較

検出性能を表 1 に示す．また，16 クラス検

出の歯牙別検出率は，SSD が 89.7%，提案モデルが 90.8%であった．提案モデルによる Dental chart の正解率は, 1 クラス検出で 90.4%, 16 クラスで 90.6%であった．いずれも交差検証の平均値である．

表 1 検出性能

		検出率 [%]	誤検出数 [個/画像]
1 クラス	SSD	98.1	0.14
	提案モデル	97.6	0.27
16 クラス	SSD	96.1	0.33
	提案モデル	97.1	0.30

3. 2. 統合処理を含む全ての処理後の結果

検出性能を考慮し，SSD による 1 クラス検出と, 提案モデルによる 16 クラス検出の結果を統合した．全ての後処理を行った結果（歯種 32 分類)，検出率が 98.8%，画像当たりの誤検出数が 0.33 個，歯牙別検出率が 91.3%，分類率が 92.4%となった．後処理後の最終出力の結果例を図 2 に示す．この症例では全ての歯牙を正しく検出・分類している．

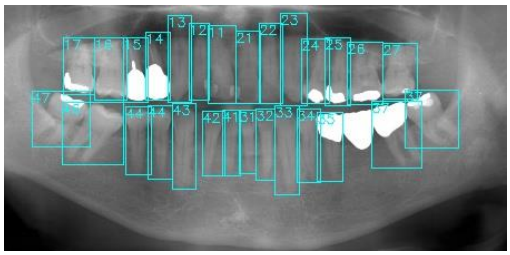


図 2 最終出力例

4. 考察

4. 1. 提案モデル

SSD と提案モデルを比較して, 16 クラス検出では検出率と歯牙別検出率について提案モデルが SSD を上回った．1 クラス検出と 16 クラス検出で Dental chart の予測性能はほぼ同等だが, 16 クラス検出でのみ性能が向上したため，Dental chart を予測して特徴マップに掛け合わせることはクラスの予測に寄与したと考える． Dental chart の予測精度には改善の余地があるため，より高い精度で Dental chart を予測するモデルや, Dental chart の精度と検出・分類精度の関係を今後検討する予定である．

4. 2. 統合処理

統合により，検出率が 1.7%向上，画像あたりの誤検出数が 0.03 個増加，歯牙別検出率が 0.5%向上した．これらが全て増加することは処理内容から予想できるが，検出率の向上量と誤検出数の増加量を比較した場合，検出率の向上量が大きく，検出性能の向上に有効であると考えている．

5. まとめ

SSD に分岐を与えた提案モデルにより，16 クラス検出について SSD より高い検出率と歯牙別検出率を得た．また統合処理などの後処理により，高い検出率を維持した歯種 32 分類を行った．今後は使用する症例数の拡張を行い，より多くのデータを用いて高い検出･分類性能と汎用性を持った手法を目指し，歯科医師の画像診断の補助への貢献を行う．

謝辞

本研究の一部は, JSPS 科研費 JP19K10347, 及びアイテック株式会社から共同研究（プロジェクト NO：501001258）助成を受けた．本研究を行うにあたり，有益なご助言を頂いた皆様に感謝の意を表します．

利益相反の有無

無し

文　献

［1］ Fukui H，Hirakawa T，Yamashita T，et al.: Attention branch network: Learning of attention mechanism for visual explanation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition: 10705-10714, 2019

［2］ Liu W，Anguelov D，Erhan D，et al.: Sdd: Single shot multibox detector. In European conference on computer vision: 21-37, 2016

Tooth recognition by integration of single-class and multi-class detection using branched SSD in dental panoramic radiographs

Takumi MORISHITA[*1], Chisako MURAMATSU[*2], Xiangrong ZHOU[*3], Ryo TAKAHASHI[*4], Tatsuro HAYASHI[*4], Wataru NISHIYAMA[*5], Takeshi HARA[*3], Yoshiko ARIJI[*6], Eiichiro ARIJI[*6], Akitoshi KATSUMATA[*5], Hiroshi FUJITA[*3]

*1 Department of Intelligence Science and Engineering, Graduate School of Natural Science and Technology, Gifu University

*2 Faculty of Data Science, Shiga University

*3 Department of Electrical, Electronic and Computer Engineering, Faculty of Engineering, Gifu University

*4 Media Co., Ltd

*5 Department of Oral Radiology, School of Dentistry, Asahi University

*6 Department of Oral and Maxillofacial Radiology, School of Dentistry, Aichi-Gakuin University

The purpose of this study is to analyze dental panoramic radiographs using deep learning, and to contribute to the diagnosis by dentists. As the initial stage, we detected teeth in images and classified their tooth types. We proposed single-shot multibox detector (SSD) networks with branching for single-class detection without distinguishing the tooth type and 16-class (Center incisor, lateral incisor, canine, first premolar, second premolar, first molar, second molar, third molar, distinguished by upper and lower jaws) detection. In addition, post-processing, such as integration processing, was performed using these results. As a result, the detection rate of the teeth was 98.8%, the number of false detections was 0.33 per image, and the classification rate was 92.4% for 32 tooth types. Therefore, the effectiveness of the proposed method was shown.

Key words: Dental panoramic radiographs, Deep Learning, CAD, Single Shot Multibox Detector

Cross-phase CT Image Registration Using Convolutional Neural Network

Tao HU[*1], Masahiro ODA[*1], Yuichiro HAYASHI[*1], Zhongyang LU[*1], Kanako Kunishishima KUMAMARU[*2], Shigeki AOKI[*2], Kensaku MORI[*1,3]

Abstract

This paper proposes a deep-learning method to improve the accuracy of cross-phase CT image registration. Artery phase CT (ACT) and non-contrast CT (NCT) both play an important role in the diagnosis of many diseases. The registration from ACT to NCT eliminates the displacement and deformation between the two phases, thus enabling convenient observation of vascular structures and lesions. However, in ACT images the intensities of these regions vary from small to large scales, and the appearances between ACT and NCT can be greatly different, sometimes leading to the failure of the cross-phase registration procedure. Different from traditional methods optimizing an overall objective function, our method utilizes a multi-resolution optimization design to look into both global and regional structures. The results show that the proposed approach acquires better performance compared with baseline method.

Keywords：CT, image registration, convolutional neural network

1. Introduction

Artery contrast computed tomography (ACT) has been widely used for the diagnosis and treatment of various diseases related to blood vessel. To get ACT, contrast agent is supposed to be injected for the imaging. However, the introduction of contrast agent might cause potential damage to the patients' health. To prevent the possible accident caused by contrast agent injection, previous study proposed to generate ACT from non-contrast CT (NCT) using deep learning techniques [1].

In clinical practices, it is very difficult to acquire the ACT and NCT scans simultaneously, and irregular movements can also be inevitable during scanning. As a result, the registration of ACT and NCT is necessary to effectively restore the misaligned structures or organs. This preprocessing procedure reduces the disruption introduced by deformations and helps ensure that reliable and accurate modality translation can be learned.

Traditional registration methods such as mutual information based algorithms [2,3] can be regarded as

*1 Graduate School of Informatics, Nagoya University, Aichi, Japan
〔Furo-cho, chikusa-ku, Nagoya, 464-8601, Japan〕
e-mail: thu@mori.m.is.nagoya-u.ac.jp
*2 Department of Radiology, Juntendo University
*3 Research Center for Medical Bigdata, National Institute of Informatics

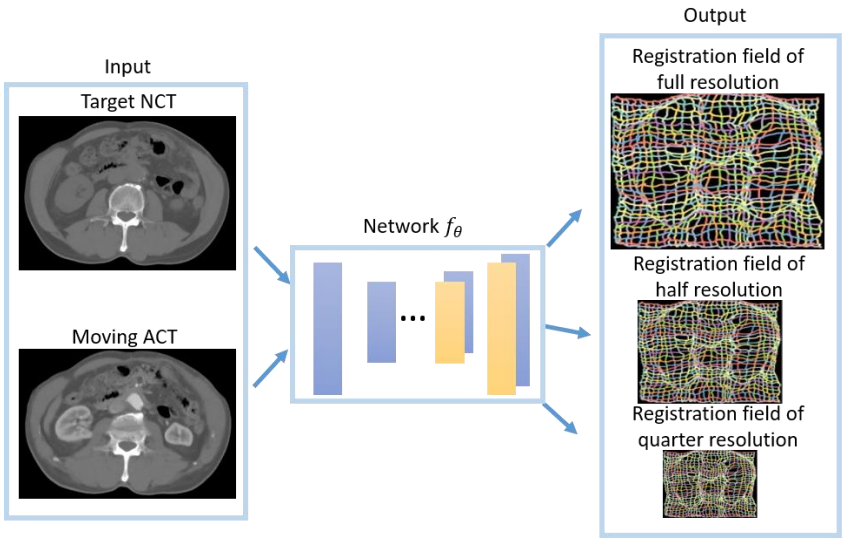


Fig.1 Proposed CNN architecture. Input is the pair of moving ACT and Target NCT. Output is the registration fields of three resolution level: full, half and quarter.

optimization procedures with prior constraints. The procedure aims at generating a voxel-to-voxel mapping able to align regions with similar anatomy structures. However, the computation burden of pairwise optimization problem can be very heavy due to the high resolution of input images, and the time-consuming methods might limit the applications in practice. To address the problem, this paper adopts an unsupervised registration method using convolutional neural network (CNN) described in [4] for ACT to NCT image registration. The proposed CNN is very convenient for the implementations on graphics processing unit (GPU), thus reducing the runtime of registration.

The proposed CNN takes two 2-D input CT slices and generates the corresponding transformation field to map the pixels in a moving slice to the given fixed slice. By minimizing a predefined energy function such as structural similarity based metric iteratively, the network is able to estimate a precise deformation field for input ACT and NCT. In addition, no manual generated deformation field is needed for ground truth, and this method can be trained in unsupervised way.

2. Method

For a given input pair of moving ACT denoted by $\mathbf{I_m}$ and fixed NCT denoted by $\mathbf{I_f}$, the goal of our network is to estimate an optimal deformation field $\boldsymbol{\phi}$ to wrap $\mathbf{I_m}$ so that the predefined loss value can be minimized. Specifically, let f_θ be the proposed CNN with parameters θ, then the loss value L during training can be calculated as:

$$L = L_s(\mathbf{I_m} \circ \boldsymbol{\phi}, \mathbf{I_f}; \theta) + \lambda C(\boldsymbol{\phi}) \quad (1)$$

where L_s represents the loss value based on similarity metric between deformed image $\mathbf{I_m} \circ \boldsymbol{\phi}$ and $\mathbf{I_f}$,

function C represents the constraints on $\boldsymbol{\phi}$, and λ is a positive constant. In order to combine global and local features, L_s is computed as the linear combination of the similarity loss values over three resolution levels:

$$L_s = \sum_{i=1}^{3} \alpha_i L_i(\mathbf{I_{mi}} \circ \boldsymbol{\phi_i}, \mathbf{I_{fi}}; \theta) \quad (2)$$

where $\mathbf{I_{mi}}$, $\mathbf{I_{fi}}$ respectively stands for the moving image and target image after i-level's downsampling, and $\boldsymbol{\phi_i}$ is the estimated deformation field of the corresponding registration task.

In this paper, normalized cross-correlation coefficient [5] is used for calculating L_i. Since we assume that the desired deformation field tends to be smooth in all directions, function C is computed with the gradients of $\boldsymbol{\phi}$:

$$C(\boldsymbol{\phi}) = \sum_p \|\nabla \boldsymbol{\phi}(p)\|^2 \quad (3)$$

where p is the pixel in every location.

The network architecture is shown in Fig.1. The majority of the architecture is U-Net [6] with depth of 5. Deconvolution layer rather than interpolation is used for upsampling operations. The input $\mathbf{I_m}$ and $\mathbf{I_f}$ are 2-D CT slices with resolution 512×512. The output of the network includes three deformation fields corresponding to the three registration tasks of different resolution levels. And all the convolutional layers are 2-D in the present study. Additionally, spatial transformation layer [4] is used to convert the feature map of final convolution block to $\boldsymbol{\phi}$, then with equation (1) we can backpropagate errors during training. In practice, for an unseen ACT and NCT image pair, the network is trained for 500 iterations, and the final estimated deformation field $\boldsymbol{\phi}$ is obtained for registration.

3. Experiments

In this paper, we explore the performance of the proposed network on 2-D cross-phase CT registration. Abdomen CT slices from 4 cases (36 slices) with both NCT and ACT are utilized to verify the effeteness of the proposed method. The reconstructed images had between 0.4 and 0.5 mm in-plane voxel sizes and 0.45/0.90 mm slice spacing/thickness.

For the training of the proposed network, we set the learning rate to 0.002, iteration number to 500, and λ to 0.01 to encourage a smooth deformation field. The linear weights α_1, α_2 and α_3 was set to 1, 1/4 and 1/8 respectively. No pre-training procedure was conducted, and the network was trained from the scratch. Additionally, the Insight ToolKit (ITK) registration framework was used as baseline method [7]. The optimization metric was set as mutual information, and B spline transform was used for generating deformation field.

4. Results

Quantitative evaluation was conducted to compare the performances of the two methods. Structural similarity index (SSIM) [8] between the transformed ACT slice and the target NCT slice was calculated as

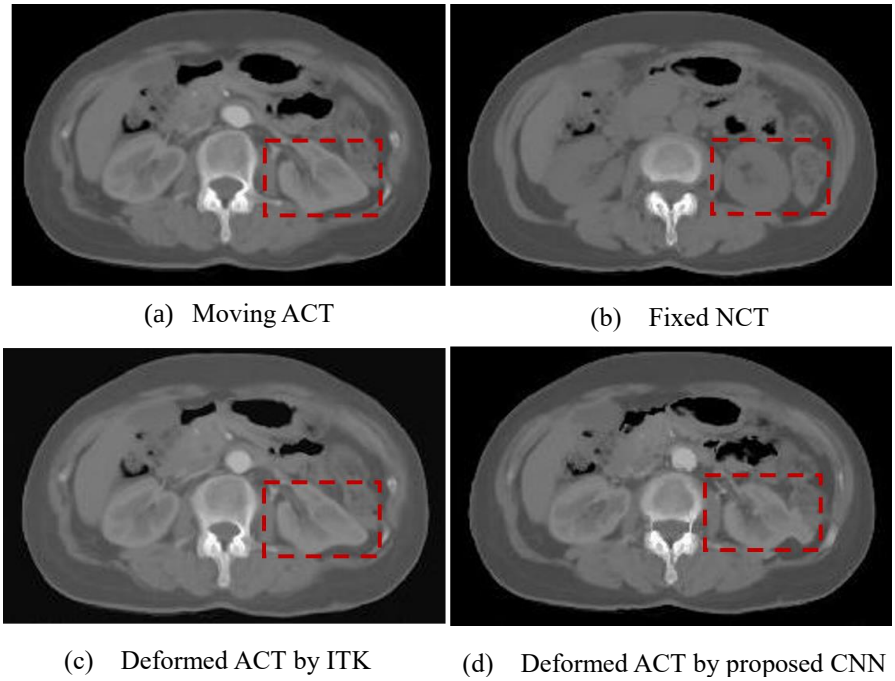


Fig. 2 Registration results by proposed CNN and ITK method. The red box shows the region of left kidney. (a): Moving ACT. (b): Fixed NCT. (c): Deformed ACT by ITK. (d): Deformed ACT by proposed CNN.

evaluation metric. Results from Tab. 1 showed that compared with baseline method, the proposed network was able to achieve higher structural similarity between the registered image and target image.

Fig. 2 presented the qualitative registration results of the proposed method and baseline method. It could be observed that compared with baseline method, the contour of the left kidney (in red box) registered by the proposed method was more similar to that of the target NCT slice. Although part of the nearby tissues was not registered as well as the larger organs, the global similarity could be improved by the proposed registration method.

5. Discussion and conclusions

In this paper, we compare the performances of an unsupervised CNN model and traditional method for cross-phase CT registration. The results show that the investigated CNN model outperforms with less distortion in organ regions. In the cases of image pairs with large appearance differences between ACT and NCT, more regular deformation field can be estimated by CNN method. Furthermore, compared with traditional method, one promising advantage of CNN method is the decreased runtime consumed for high resolution medical image registration, and this improvement may be more significant in 3-D medical image analysis.

Additionally, with proper tuning procedures, this CNN model can be easily extended to the learning model for a wide variety type of medical image modalities or anatomy structures, which will benefit the

Table 1 Mean SSIM between deformed ACT and target compared to a method from ITK

	ITK	Proposed
SSIM	0.7077	**0.7157**

further research such as CT modality translation and lesion segmentation. Also, to improve the registration accuracy and model robustness, our future study will investigate the sensitivity of the CNN model to the setting of smooth constraint hyper-parameter λ.

Competing interests

None.

Acknowledgement

Thank for the help and advices from the members of Mori laboratory and AMED NII team. A part of this research was supported by 26108006, 17K20099, the AMED Grant Numbers JP19lk1010036, JP20lk1010036, and JSPS KAKENHI 17H00867.

References

[1] Oda M, Kumamaru K, Aoki S, et al: Non-contrast to contrasted abdominal CT volume regression using fully convolutional network. Int. Journal of Computer Assisted Radiology and Surgery, vol. 14, Supplemental Issue, pp. s103-s104 (2019)

[2] Maes F, Vandermeulen D, and Suetens P: Medical image registration using mutual information. Proceedings of the IEEE 91.10, pp.1699-1722 (2003).

[3] Klein S, Staring M, and Pluim J: Evaluation of optimization methods for nonrigid medical image registration using mutual information and B-splines. IEEE Trans Imag Proc 16.12, pp.2879-2890 (2007).

[4] Balakrishnan, G, Zhao A, Sabuncu M, et al: Voxelmorph: a learning framework for deformable medical image registration. IEEE Trans Med Imag 38.8, pp.1788-1800 (2019).

[5] Avants, B, Epstein C, Grossman M, et al: Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Med Imag Anal 12.1, pp.26-41 (2008).

[6] Ronneberger, O, Fischer P and Brox T: U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention, pp. 234-241 (2015).

[7] Avants B, Tustison, N, Song G, et al: A unified image registration framework for ITK. In International Workshop on Biomedical Image Registration, pp. 266-275 (2012).

[8] Wang Z, Bovik A, Sheikh H, et al: Image quality assessment: from error visibility to structural similarity. IEEE Trans Imag Proc 13.4, pp.600-612 (2004).

Prediction of Segmentation Accuracy of Liver and Spleen in Contrast-Enhanced MR Images Using Uncertainty Estimated from Bayesian U-Net

Bin ZHANG[1], Mazen SOUFI[1], Yoshito OTAKE[1], Masatoshi HORI[2], Noriyuki TOMIYAMA[3], Yoshinobu SATO*[1]

Abstract

Bayesian U-Net has shown promising performance in image segmentation tasks. Particularly, in a previous study by our group, a Bayesian U-Net has shown possibility to segment hip and thigh muscles in CT images with a high accuracy based on epistemic (i.e. model-related) uncertainty estimation. In addition to segmenting the target structures, the estimated uncertainty demonstrated correlation with Dice coefficient. This allowed to use the uncertainty as a predictive measure for the segmentation accuracy, hence called 'predictive uncertainty'. However, the capability of Bayesian U-Net to segment MR images has not been examined yet. In this study, we investigated the feasibility of Bayesian U-Net in segmenting the liver and spleen from MR images. A dataset consisting of 132 contrast-enhanced MR images was used. The Bayesian U-Net had an architecture of a conventional U-Net with a dropout layer added after each basic block. At testing, 10 Monte Carlo dropout samples were used for segmentation and uncertainty estimation. Segmentation accuracy was measured using Dice coefficient based on 2-fold cross-validation. The uncertainty was estimated as an average of the variance of softmax probabilities at the predicted label. Bayesian U-Net has yielded a high overall accuracy. The correlation between Dice and the predictive uncertainty was -0.933 and -0.789 for the liver and spleen, respectively. This indicates the feasibility of the uncertainty estimated by Bayesian U-Net as a predictive measure of segmentation accuracy in MR images.

Keywords：Bayesian U-Net, Segmentation accuracy prediction, Uncertainty

1. Introduction

Segmentation is a process to extract structures or organs of interest from images, which is essential in various image-guided diagnostic and therapeutic procedures. Bayesian U-Net has recently shown promising performance in image segmentation tasks. Particularly, in a previous study by our group, Bayesian U-net has demonstrated the possibility of

1 Division of Information Science, Nara Institute of Science and Technology, 8916-5, Takayama-cho, Ikoma, Nara 630-0192, Japan

* Email: yoshi@is.naist.jp

2 Department of Radiology, Kobe University Graduate School of Medicine, 7-5-1, Kusunokicho, Chuo, Kobe, Hyogo 650-0017, Japan

3 Department of Diagnostic and Interventional Radiology, Osaka University Graduate School of Medicine D1, 2-2 Yamadaoka, Suita, Osaka, 565-0871, Japan

hip and thigh muscles segmentation from CT images with a high accuracy based on epistemic uncertainty estimation [1]. In addition to predicting the target structure, the estimated uncertainty showed very high correlation with Dice coefficient. Thus, uncertainty could be used as an indicator of segmentation accuracy. However, the capability of Bayesian U-Net to segment magnetic resonance (MR) images has not been examined yet.

In this study, we investigated the feasibility of Bayesian U-Net in predicting the segmentation accuracy of liver and spleen in contrast-enhanced MR images. The segmentation of these organs could be helpful for the automated diagnosis of liver diseases using image analysis approaches, such as statistical shape analysis [2, 3].

2. Methods

The Bayesian U-Net has an architecture of a conventional U-Net with the dropout layer added before each max pooling layer and after each up-convolution layer [1]. This architecture is similar to the one proposed for the Bayesian SegNet network proposed by Kendall et al. [5]. The liver and spleen MR images were first segmented by Bayesian U-Net. In the inference process, model uncertainty was estimated based on the variance of probabilistic softmax output of multiple Monte Carlo dropout samples. The output label was computed by finding the label corresponding with the maximum of the average of the output softmax probabilities.

3. Experiments

In this study, 132 contrast-enhanced MR images were used. The matrix size was 512×512, and the slice thickness was 4.0 mm. Each image was normalized by its mean and standard deviation. The Bayesian U-Net was trained with a dropout ratio of 0.5 for 5×10^4 iterations, batch size of 4, and 10 Monte Carlo dropout samples were used at the test time. At the training, data augmentation including flipping and affine transformations was applied. The transformations comprised a rotation of [5°, 5°], horizontal/vertical translation of [5, 5] pixels, a shear of an angle within [π/16, π/16] and a scaling of a ratio within [95%, 105%]. The segmentation accuracy was assessed by using Dice coefficient. Structure-wise uncertainty was computed as the average of the estimated pixel-wise variance in the estimated labels. The predictive capability of the Bayesian U-Net for the segmentation accuracy was assessed by computing Pearson correlation coefficient between the Dice coefficient and structure-wise uncertainty. The experiments were performed using a graphics card NVIDIA GeForce GTX 1080ti.

4. Results

4.1 Correlation between uncertainty and Dice

Figure 1 shows joint histograms of the accuracy and uncertainty. The correlation between Dice and uncertainty was -0.933 and -0.789 for the liver and spleen, respectively. This suggests that the uncertainty can be used as a predictive measure for segmentation accuracy in MR images. Figure 2 shows representative results, i.e. cases with minimum, median and maximum Dice for the liver and spleen by Bayesian U-Net. It could be observed by comparing ground truth with predicted labels that the segmentation failures corresponded with locations having a large uncertainty.

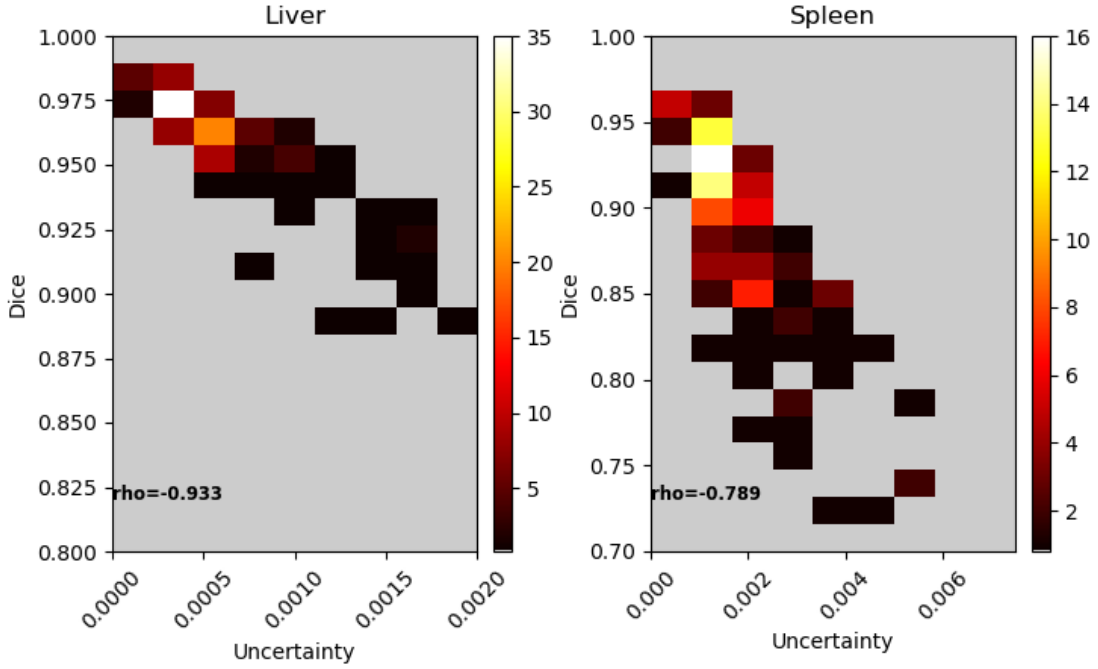


Fig. 1 Joint histograms between the uncertainty and segmentation accuracy produced by Bayesian U-Net. High correlations (liver: -0.933; spleen: -0.789) were obtained.

4.2 Segmentation accuracy

Figure 3 shows the segmentation accuracy using Bayesian U-Net. The Bayesian U-Net showed high Dice in both the liver and spleen, which indicated the capability of Bayesian U-Net in the segmentation of MR image. The failure case of liver with Dice value 0.33 and spleen with Dice value 0.27 were caused by a low contrast inside the organ, as shown in Fig.4.

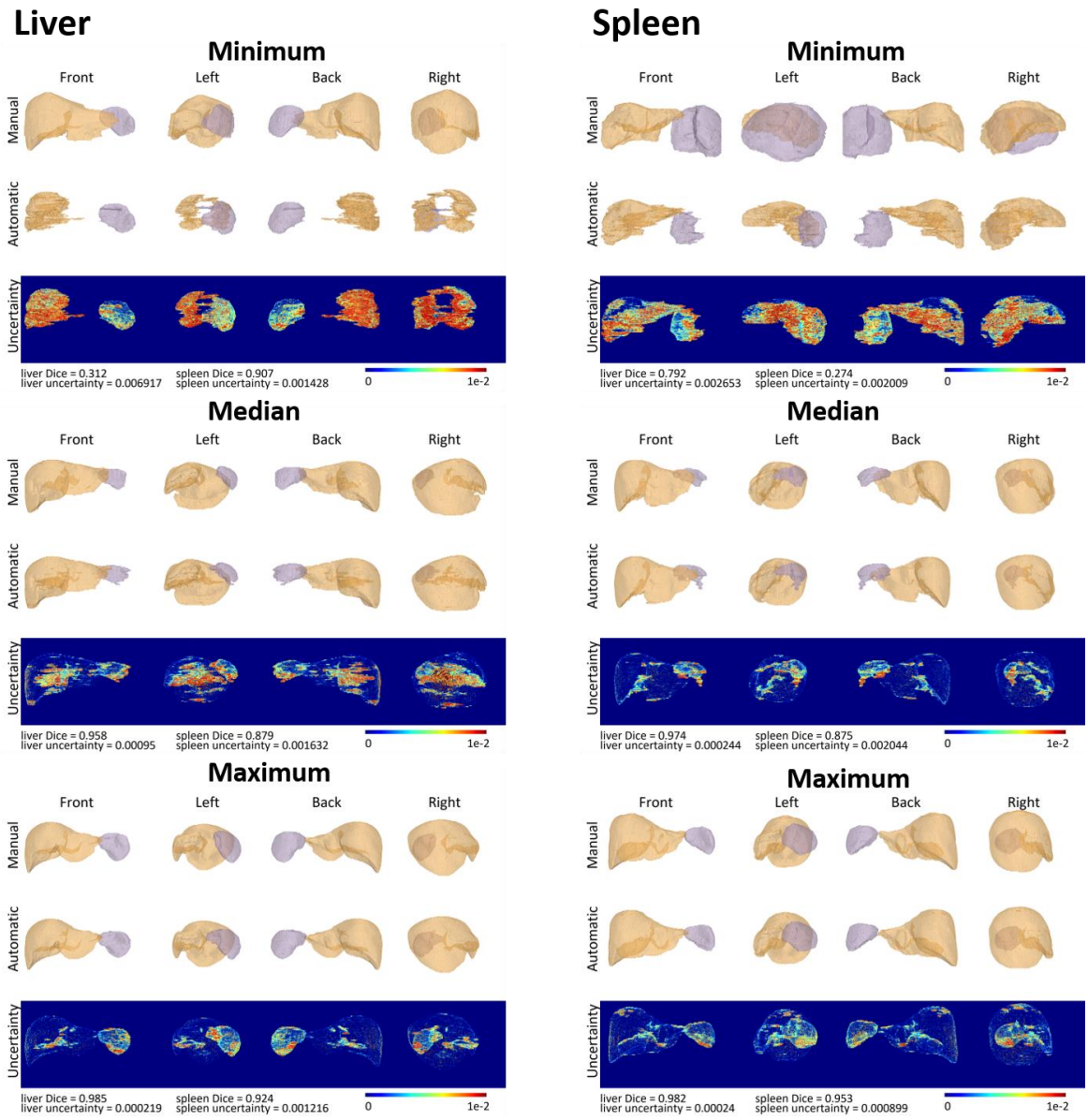


Fig. 2 Representative results for segmentation and uncertainty estimation by Bayesian U-Net.

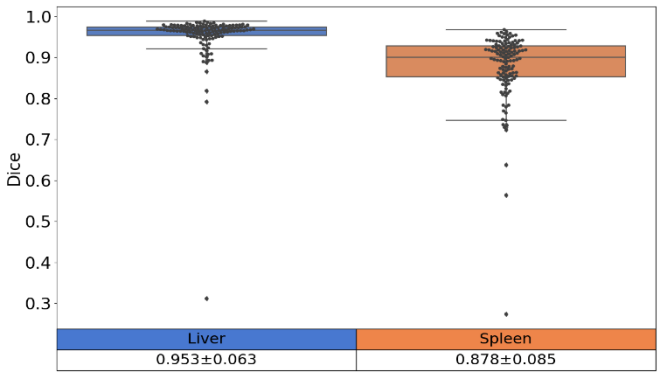


Fig. 3 Dice coefficient of the segmentations obtained using Bayesian U-Net.



Fig. 4 Representative cases with low (Dice < 0.5) and high (Dice > 0.95) segmentation accuracy per organ. (a) Low segmentation accuracy; (b-c) High segmentation accuracy.

5. Discussion and Conclusions

In this paper, we demonstrated the feasibility of the uncertainty estimated by using a Bayesian U-Nets as a predictive measure for the segmentation accuracy of liver and spleen in MR images. Moreover, it also indicated that the uncertainty generated by Bayesian U-Net is capable of indicating the locations with low segmentation accuracy. Similar to the study by Hiasa et al, an active-learning approach can be implemented to improve the segmentation accuracy of Bayesian U-Net while reducing the manual annotation cost. This will be a major point in our future study.

Competing interests

The authors have no competing interests.

Acknowledgement

This research was supported by Japan Society for the Promotion of Science (JSPS) Grants-in-Aid for Scientific Research (KAKENHI) Number 19K20711.

References

［1］ Hiasa Y., Otake Y, Takao M., et al.: Automated muscle segmentation from clinical CT using Bayesian U-net for personalized musculoskeletal Modeling. IEEE Trans on Med Imag. 39.4:1030–1040 (2019).

［2］ Hori M, Okada T, Higashiura K, et al.: Quantitative imaging: quantification of liver shape on CT using the statistical shape model to evaluate hepatic fibrosis. Acad Radiol 22(3):303–309, 2015

［3］ Soufi M, Otake Y, Hori M, et al.: Liver shape analysis using partial least squares regression-based statistical shape model: application for understanding and staging of liver fibrosis, Int J Comput Assist Radiol Surg. 14(12):2083–2093 (2019).

［4］ Ronneberger O., Fischer P., and Brox T.: U-net: Convolutional networks for biomedical image segmentation." International Conference on Medical image computing and computer-assisted intervention (MICCAI2015). Springer, Cham, 2015.

［5］ Kendall A, Badrinarayanan V, Cipolla R. Bayesian SegNet: Model Uncertainty in Deep Convolutional Encoder-Decoder Architectures for Scene Understanding. arXiv:1511.02680

皮弁壊死領域早期診断のための狭帯域光源・カラーカメラ連動システム設計

倉渕　瑶子[*1]　中野　和也[*2]　大西　峻[*3]　秋田　新介[*4]

羽石　秀昭[*3]

要旨

血流のある皮膚・皮下組織や深部組織により構成される皮弁を用いた遊離皮弁移植術は重要な手技であるが，術後の血管吻合不良や血管閉塞に伴う壊死が発生する可能性がある．壊死の回避には早期の血管再吻合が必要だが，目視での早期発見は困難である．そこでわれわれは，壊死領域を早期に診断するための簡便な撮影システム構築を目指している．具体的には，適当な波長帯と強度を選択した LED 光源を 2 種類用意し，この光源の連続切替と RGB カメラ撮影により，6 次元の信号を取得し診断に利用する．本研究では基礎実験として， ラットの皮弁壊死モデルを作製し，ラット皮弁の血管結紮後早期の分光反射率を取得することで光源設計を行った．得られた光源スペクトルを用いて，設計した光源下と従来光源下で得られる RGB カメラ画像をシミュレーションにより作成し，壊死領域の抽出を行った．従来の白色光下の RGB 画像に比べて，提案システム光源の RGB 画像では，より高い感度で壊死領域の抽出を実現した．

キーワード：マルチバンドイメージング，Light-emitting Diode，診断支援システム，皮弁，壊死

1. はじめに

血流のある皮膚・皮下組織や深部組織により構成される皮弁は，疾患や外傷により欠損した部位の再建に用いられる．皮弁移植術は血管吻合を伴うため，吻合不良や血管の閉塞によって術後に皮弁壊死が発生する場合がある．皮弁壊死の回避のためには，皮弁血流状態悪化の早期発見による再手術が必要であるが，血流状態の把握には僅かな皮弁色調の変化の観察が必要となるため困難である．この観察には分光的な計測が有効である[1].

本研究では皮弁色調変化の早期発見を目的とした簡便な光源－RGB カメラ連動システムの構築を目指す．提案システムは，2 種類の光源の連続切替と RGB カメラ撮影により，6 次元信号の取得を実現する．本検討では，ラットを用いた皮弁壊死モデルを作製し，分光反射率を計測した．計測データを用いて提案システムの光源設計を行い，シミュレーションによる有効性検証を行った．

2. データ取得

皮弁表層の血流状態悪化の分光的特徴を観察するため，ラット(SLC:Wistar,オス)を用いた皮弁壊死モデル[2]を作製した．本検討では皮弁栄養血管のうち動脈を結紮した阻血モデルを使用した．継続的な分光反射率取得を

*1 千葉大学大学院融合理工学府基幹工学専攻医工学コース
〔〒263-8522 千葉市稲毛区弥生町 1-33〕
e-mail: y_kurabuchi@chiba-u.jp
*2 宮崎大学 キャリアマネジメント推進機構テニュアトラック推進室
*3 千葉大学フロンティア医工学センター
*4 千葉大学 医学部 形成外科

目的として作製したチャンバーに皮弁を固定したうえで(図 1(a))，ハイパースペクトルカメラによる撮影を行った．図 1(b)にモデル作製後 3 時間経過時の取得分光反射率を示す．

3. 提案システム設計

生着領域と壊死領域の分光反射率(図 1(b))をもとに，提案システムに用いる光源スペクトルの設計を行った．光源スペクトルは LED 光源の重みづけ合成により定義した．LED 光源スペクトルは事前に作製した分光分布可変照明[3]に搭載された 13 種類の LED を，RGB カメラ感度は Grasshopper3 (FIR 社製)のものを用いた(図 2 (a))．光源スペクトルは生着・壊死領域間の分離度が最大となるものを粒子群最適化により算出した．分離度はフィッシャーの線形判別により定義した．図 2(b)に取得光源スペクトルを示す．取得光源スペクトルは RGB センサごとに独立したピークを持ち，また生着・壊死領域の差異の大きな位置と合致していることを確認した．

4. 評価

各光源下の取得画像をシミュレーションにより生成し(図 3)，線形 SVM による壊死領域検出を行った．壊死領域の正解値は，モデル作製後 3 日経過時の画像より手動で選択した．壊死領域検出結果を図 4 に，各光源下での感度，特異度，精度を表 1 に示す．提案システムでは従来光源下と比較し，全ての評価値が優れていたことから，その有効性を確認した．

5. おわりに

皮弁壊死の早期発見を目的とした光源－RGB カメラ連動システムの光源スペクトル設計を行った．設計光源を用いることで，従来光源下より正確に壊死領域を抽出可能であることを確認した．今後は提案システムの実機実装と手術環境下での評価実験を行う．

利益相反の有無

なし

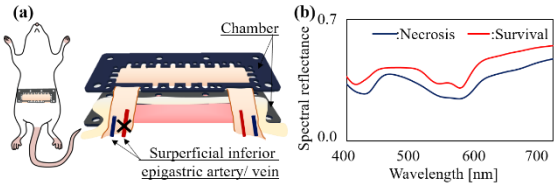


図 1 作製皮弁モデル(a)皮弁固定チャンバー (b) 取得分光反射率

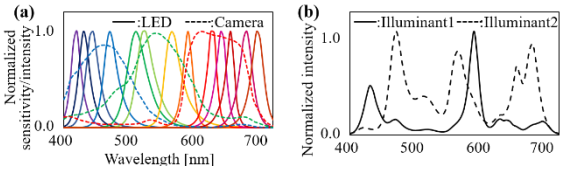


図 2 システム分光特性(a)LED スペクトルとカメラ感度(b)設計スペクトル



図 3 シミュレーションによる取得画像 (a)従来光源下(b)設計光源 1 下(b)設計光源 2 下

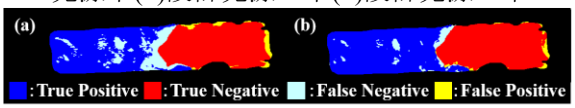


図 4 壊死領域検出 (a)従来光源下(b)設計光源下

表 1 評価値による壊死領域検出比較

Illuminant	Sensitivity [%]	Specificity [%]	Accuracy [%]
Conventional	77.5	93.9	93.6
Designed	92.0	96.4	96.7

文 献

[1] Zhu C., Chen S., Chui C. H. K.,et al. : Early detection and differentiation of venous and arterial occlusion in skin flaps using visible diffuse reflectance spectroscopy and autofluorescence spectroscopy. Biomed Opt Express **7**(2), 570-580, 2016

[2] Matsumoto N., Aoki M., Nakao J., et al.: Experimental rat skin flap model that distinguishes between venous congestion and arterial ischemia: the reverse U-shaped bipedicled superficial inferior epigastric artery and venous system flap. J. Plast. Reconstr. Aesthet. Surg, **139**(1): 79e-84e, 2017

[3] Kurabuchi Y., Nakano K., Ohnishi T., et al.: Optimization of surgical illuminant spectra for organ microstructure visualization. IEEE Access **7**: 70733-70741, 2019

Design of a system combining narrow band illuminants and color camera for early diagnosis of skin flap necrosis

Yoko KURABUCHI[*1], Kazuya NAKANO[*2], Takashi OHNISHI[*3], Shinsuke AKITA[*4], Hideaki HANEISHI[*3]

*1 Graduate School of Science and Engineering, Chiba University

*2 Institute for Tenure Track Promotion, University of Miyazaki

*3 Center for Frontier Medical Engineering, Chiba University

*4 Department of Plastic, Reconstructive, and Aesthetic Surgery, Chiba University

In plastic surgery, free tissue graft procedures are an essential part of fixing tissue lack injuries or diseases. However, tissue necrosis caused by vascular occlusions or poor vascular anastomosis is a severe problem for prognosis. Follow-up surgery in the early stage of skin flap necrosis is essential for salvage the tissue. In this paper, we designed a simplified system combining a two-channel narrow-band illuminant and a color camera to capture six-channel spectral signals for supporting the early diagnosis of skin flap necrosis. The two-channel illuminant was designed as the highest separation spectra between normal-necrosis regions, which were composed of 13 kinds of commercially available light-emitting diode (LED) spectra. We measured spectral reflectance of the early stage of necrosis of the rat model for designing illuminant spectra for the proposed system. A prototyped skin flap chamber was used for stable measurements. We evaluated the effectiveness of the proposed system by extracting the early necrosis region using linear support vector machine. Simulated RGB image signals under designed-illuminants, and a conventional illuminant were used for evaluation experiments. Sensitivity under the designed-illuminants (92.0 %) was higher than that under a conventional illuminant (77.5 %).

Key words: Multiband imaging, Light-emitting Diode, Surgery assistant system, Free tissue graft, Necrosis.

画素毎の難易度に基づく注意機構を用いた敵対的学習による セマンティックセグメンテーション

津田　大輝 *1　堀田　一弘 *1

要旨

セマンティックセグメンテーションは入力された画像の全画素に対してラベル推定を行う処理である．セグメンテーションモデルを生成器として取り扱い，GAN のように識別器を加えて敵対的に学習を行うことによりセグメンテーションモデル単体よりも高精度な推定を行うことができる．我々はセグメンテーションモデルにとって識別が難しい画素を Attention Map として識別器にリークする Adversarial Pixel-wise Difficulty Attention Network を提案する．これにより，識別器が従来の学習よりも正確に真偽を識別し，それを基にセグメンテーションモデルは敵対的学習構造により識別器を騙すための学習するため，精度が向上すると期待される．評価実験では細胞画像と眼底画像のデータセットを用いて従来手法と提案手法の精度比較を行った．その結果，Intersection over Union と F 値を用いた精度評価において，提案手法が従来手法を上回ることを確認した．

キーワード：セマンティックセグメンテーション，注意機構，画素難易度，敵対的学習，深層学習

1. はじめに

セマンティックセグメンテーションは画像内の画素に対してクラス推定を行う基本的なタスクの一つである．特に医用画像や細胞画像ではヒューマンエラーを防止し，アルゴリズムによる自動処理により人間が判断する時間の効率化や客観的な判断の用いる際に重要である．しかし，これらの画像では自動運転のデータセットなどと比較して教師付き画像が少なく，決まった法則性がないため難しいタスクである．また，多数の教師付き画像を作成するためには専門家によるラベリングが必要であり，これには多大な労力と時間とコストがかかる．そのため少数の学習画像で画素毎の識別能力を強化する必要がある．

セマンティックセグメンテーションでは FCN[1] の登場以来，Convolutional Neural Network(CNN)[2] を用いた手法が主流となっている．特にエンコーダに ResNet[3] や DenseNet[4] を用いた非常に深いネットワーク構造を有するモデルが主流となっている．また，これらに Attention Machanism を用いて特徴マップ間や画素間で注目するべきと判断した箇所を強調し，学習により深い特徴抽出と文脈情報の統合を行いセグメンテーションにおける精度を大きく向上させた．しかし，これらの手法は学習パラメータ数が増大し，計算コストが非常に大きくなるという問題がある．この問題には軽いネットワークにおいて重要な箇所を重点的に学習することが必要である．

画像生成タスクで一般的な Generative Adversarial Network (GAN) は生成器の画像生成と識別

*1名城大学 理工学研究科 電気電子工学専攻
〔〒 468-8502 愛知県名古屋市天白区塩釜口 1-501〕
e-mail: 193427019@ccalumni.meijo-u.ac.jp, kazuhotta@meijo-u.ac.jp

器の生成されたものか本物かを判断する敵対的学習により本物らしい画像を生成することができる．これを Image-to-Image に拡張した pix2pix[5] はセグメンテーションに適用しても効果的である．我々はこの関係性に注目し，生成器が画像変換する際に難しいと判断した情報を識別器へリークすることにより，識別器が本物の画像と生成された画像の違いを明確に判断する根拠とすることができると考えた．我々は生成器からセグメンテーション難易度の高い画素を強調する Attention Map を作成し，生成器と識別器で用いる Adversarial Pixel-wise Difficulty Attention Network(APDA-Net) を提案する．

本稿の実験では 2 種類の細胞画像 [6, 7] と 3 種類の眼底画像 [8, 9, 10] のデータセットにおいて Mean IoU と F 値などによる評価を行った．その結果，我々の提案手法である APDA-Net が従来手法よりも高精度であることを確認した．また，いくつかの Ablation Study を行い，提案手法の有効性を示した．

本論文の構成は以下の通りである．まず，2 節で関連研究について述べる．次に，3 節で提案手法を示し，4 節で実験結果を示す．最後に結論として，まとめと今後の課題を 5 節で述べる．

2. 関連研究

2.1. セマンティックセグメンテーション

オリジナルの FCN[1] は畳み込み処理とプーリング処理により，徐々にダウンサンプリングを行い，最終的に低解像度の高次元特徴マップを作成する．この特徴マップはセマンティック情報を有しているが，細かい情報は失われてしまい，物体の境界付近の予測が不正確になる．高解像度の重要な特徴マップを得るために SegNet[11]，U-Net[12] はエンコーダにより得られた様々なコンテキスト情報をデコーダで組み合わせることにより，空間情報を徐々に抽出するように設計されている．SegNet[11] は VGG16[13] を基にしており，Max Pooling により選択されたインデックスをデコーダにコピーすることでエンコーダの空間情報を利用している．U-Net[12] は医用画像のセグメンテーションのために提案された最も有名なセグメンテーションモデルの一つである．これはエンコーダとデコーダ間でスキップコネクションにより接続することにより空間情報が補完され，各画素により正確にクラスラベルを割り当てることができる．

2.2. 敵対的学習

敵対的学習は画像生成手法である Generative Adversarial Network(GAN)[14] によ ri 提案された手法である．GAN では画像生成をはじめとして画像変換 [5, 15]，ドメイン適応 [16, 17] など多くのタスクにおいて成功を収めている．敵対的学習では生成器と識別器が互いに競い合うように学習を行うことにより，生成の精度を高めることができる．また，非敵対的ネットワークよりもセグメンテーションに関して優れた結果を示している [18]．敵対的学習構造を利用して効率的に学習を行うことを提案した手法として [19] がある．この論文は識別器がどのような特徴に注目しているかという情報を生成器に教えることにより，正解画像に近い画像を生成することを目的としている．

2.3. Attention Mechanism

Attention Mechanism は人間の注意機構を機械学習に応用した技術であり，画像認識分野では入力に対して学習に重要な空間またはチャンネルを強調する．CNN に ResNet[3] のように残差構造の Attention Mechanism を導入した Residual Attention Network[20] や特徴マップをチャンネル毎に適応的に強調する Attention Mechanism を導入した Squeeze-and-Excitation Network(SE-Net)[21]，CNN が判断した根拠をヒートマップとして可視化する Class Activation Mapping (CAM)[22] を Attention Map として利用する Attention Branch Network[23] などがある．また，自然言語分野の Attention Mechanism のみで翻訳タスクを行う Transformer [24] では同じテンソルを使用する Self-Attention と 2 つのテンソルを使用する Source-Target-Attention がある．この Self-Attention を応用して，画像の特徴マップにおけるピクセル間の類似度を学習させるネットワークがいくつか提案されている [25, 26]．

3. 提案手法

セマンティックセグメンテーションは出力層でクラス数へと圧縮し，Softmax 関数を適用して各クラスの確率値に変換する．この際，クラスの確率値が均等な画素は識別が難しいため，識別器が正解画像と生成画像の違いを判断する根拠として生成器の難しいと判断した画素情報を明示的に与える

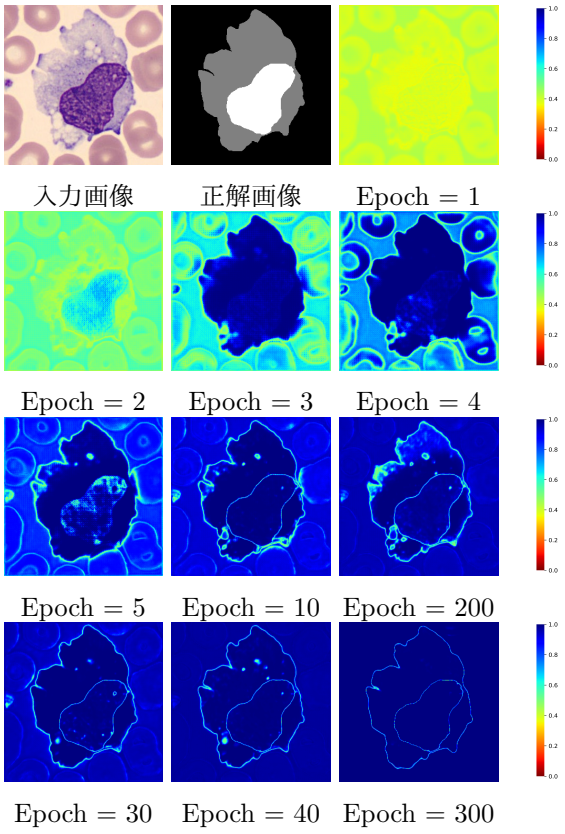


図 1　WBC[6] での確率マップの可視化

ことにより正確な識別ができると考えられる．図 1 に白血球細胞 [6] の 3 クラスの分類を行う画像において，入力画像および正解画像（黒で示した赤血球を含む背景，灰色が白血球の細胞質，白色が細胞核），U-Net[12] の出力層に Softmax 関数を適用後，確率の最大値を画素毎にヒートマップで可視化した確率マップを示す．確率マップは確率値が 1 に近づくほど青色に，0 に近づくほど赤色になるヒートマップを示している．図 1 では 3 クラス分類問題のため，最も難易度が高い場合には 3 クラスそれぞれに約 33%と出力する．そのため，学習を始めたばかりの 1epoch 目では多くの画素において約 40%と出力されていることが分かる．学習が進んだ epoch でも確率の最大値が約 50%の画素は他の画素と比較して識別が難しいことが示されている．一方，確率の最大値が 100%に近い青色の画素は学習が進んでも変化しない画素が多い．これらのことから確率の最大値が高い画素は識別が簡単で学習回数を重ねても分類結果が変わらないため，学習時の重要性は相対的に低いと考えられる．また，確率の最大値が相対的に低い約 50%の画素では物体の外形を識別するための情報が多く集まっていることから，学習時の重要性は相対的に高いと考えれられる．セグメンテーション精度を向上させるには，このような識別が高難易度の画素を正確に学習させることが重要である．

我々の提案する Adversarial Pixel-wise Difficulty Attention Network(APDA-Net) では，医用画像や細胞画像のセマンティックセグメンテーションにおいて高精度を示し，計算コストが軽いため標準的に利用される U-Net[12] を生成器のベースとし，識別器には 6 層の畳み込み層を持つ CNN を使用する，図 2 に提案手法である APDA-Net の詳細なモデルを示す．APDA-Net では U-Net のデコーダ部分に Pixel-wise Difficulty Attention Module(PDA Module)[27] を参考にしたものを採用した．PDA Module はクラス分類の難易度に注目した Attention Mechanism であり，識別が難しい画素を強調する Attention Map を生成し，掛け合わせることによりネットワークが難易度を考慮した学習が可能になる．

図 3 に PDA Module のネットワーク構造を示す．PDA Module ではクラス分類が高難易度の画素を確率の最大値に基づいて強調するためにデコーダの各解像度でセグメンテーションを行う．デコーダからの特徴マップに対して最初に 1×1 畳み込み処理によりチャンネル数次元を半分に圧縮し，Batch Normalization と ReLU 関数で正規化を行う．再度，1×1 畳み込み処理により出力したいクラス数まで次元を圧縮し，Softmax 関数でクラス毎の確率値を求めることができる．この確率マップから画素毎の最大値を抽出する．これらの確率の最大値を 1 から引くことにより難しい画素に対して高い値が付与された Attention Map を作成することができる．しかし，この Attention Map の取り得る値を x とした時の範囲はクラス数に依存して $0 \leq x \leq (N-1)/N$ となってしまう，これにより，画素毎の強調と抑制が十分になされないことが考えられるため，$0 \leq x \leq 1$ の範囲になるように正規化を行う．正規化を行った Attention Map とデコーダからの特徴マップで乗算し，Residual な加算を行うことにより次の解像度のデコーダへの入力特徴マップとする．また，Softmax 関数を適用する前のクラス数までチャンネル次元を圧縮した特

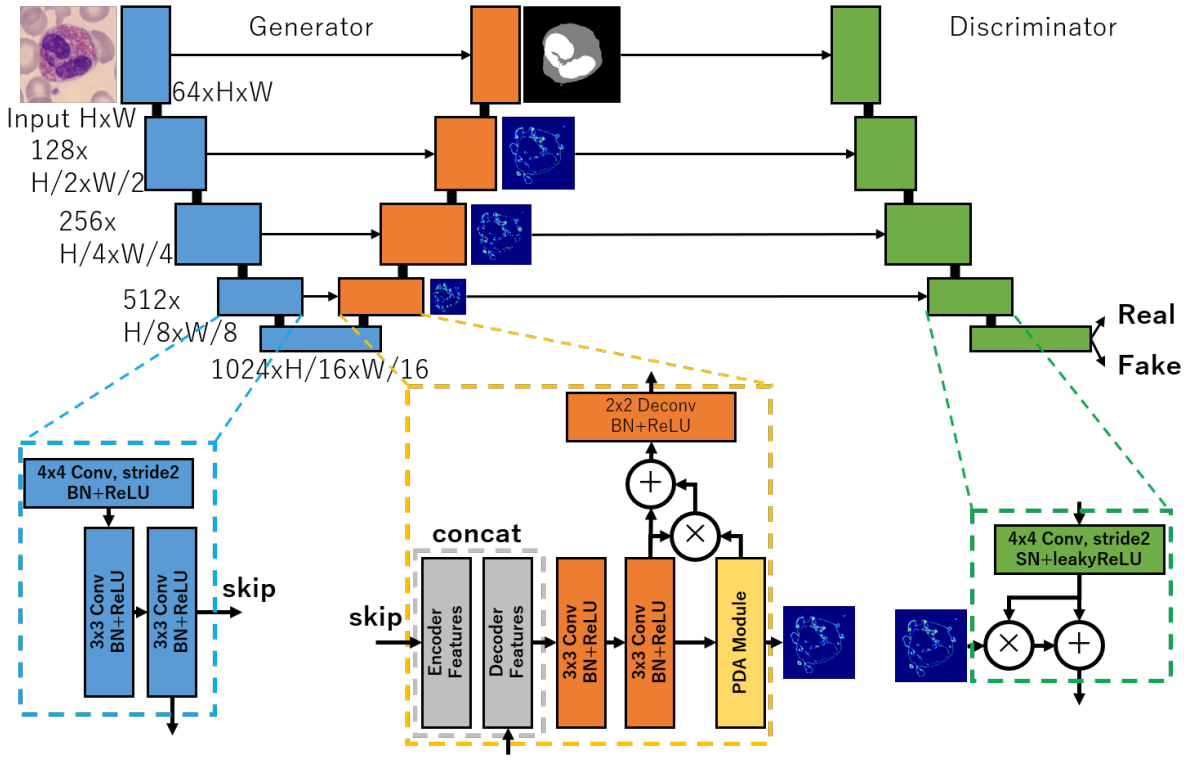


図 2　Adversarial Pixel-wise Difficulty Attention Network(APDA-Net) のネットワーク構造

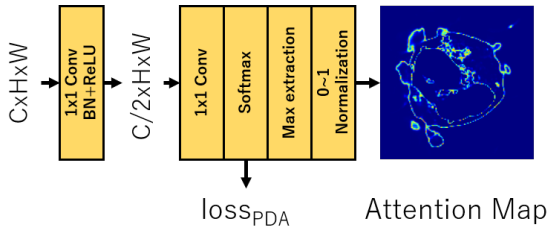


図 3　Pixel-wise Difficulty Attention Module(PDA Module) のネットワーク構造

徴マップは元の画像サイズまで Upsampling して損失関数を算出する．元画像サイズで損失関数を算出することで低解像度の特徴マップでも最終的な出力に近い学習を行い，PDA Module は画素の難易度に注目した Attention Map を作成できるようになる．また，損失関数は Softmax cross-entropy を使用した．

敵対的学習を用いている APDA-Net の識別器の目的として生成画像と正解画像との違いを学習し，判断した真偽を生成器の学習に用いることにより生成器単体の場合から精度の向上を行うことである．敵対的学習を用いている APDA-Net の識別器は生成画像と正解画像との違いの特徴量を集約して真偽を判定することにより，生成器単体の場合よりも精度の向上した結果の出力を目的としている．従来の識別器が判断する根拠となる生成画像と正解画像との違いは学習過程においてボトムアップ的に獲得していくため，重要性が低い意図しない画素が判断する根拠になってしまうことが考えられる．そのため，生成器の PDA Module で作成した Attention Map を識別器にリークすることにより，生成画像の難しい画素の情報を基に生成画像と正解画像の違いを判断する根拠として利用できる考えた．

APDA-Net の識別器では各解像度の PDA Module によって作成された Attention Map を利用するためにネットワークの設計をした．識別器はエンコーダ構造をしており，生成画像または正解画像とペアになる入力画像に対して初めに 3×3 畳み込み処理で特徴を抽出する．次に，各解像度において 4×4 のカーネルでストライド数 2 の畳み込み処理を行うことにより Downsampling を行う．また，この畳み込み処理では GAN の学習安定化手

法である Spectral Normalization[28] を使用することにより識別器の学習を安定化させる．得られた特徴マップに Leaky ReLU 関数を適用し，生成器に含まれる PDA Module で作成した同じ解像度の Attention Map と乗算と Residual な加算を行うことにより生成器が重要視して学習をしている情報を識別器にリークすることができる．最後に最終層で 1×1 畳み込み処理により真偽を予測する計算を行う．また，この際に pix2pix[5] で提案されている Patch GAN を参考に入力画像の 16×16 画素を Patch として識別器が真偽を判断し，損失関数を計算して平均を計算する．これにより画像 1 つに対しての大局的な判断でなく，切り出した Patch での局所的な判断を基に生成器の学習を向上させる．

我々は標準的なクラス分類の損失関数と敵対的学習構造を用いた損失関数で APDA-Net を学習させる．生成器の学習には用いた U-Net の最終出力では従来のセマンティックセグメンテーションと同様に Softmax Cross-entropy を用いてロス ($loss_{CE}$) を算出する．生成器に含まれる 3 つの各解像度での PDA Module においても Softmax Cross-entropy を用いてそれぞれのロス ($loss_{pda1}, loss_{pda2}, loss_{pda3}$) を算出する．さらに，生成器が識別器を騙すことができたかどうかの Adversarial Loss を用いることにより，生成器単体のロス $loss_{CE}$ だけでは判断できない本物らしさを考慮して生成器を学習することができる．Adversarial Loss は以下の式になる．

$$\begin{aligned}\pounds_{adv}(G,D) = \mathbb{E}_{x,y\sim p_{data}(x,y)}\left[logD(x,y)\right] \\ +\mathbb{E}_{x\sim p_{data}(x)}\left[log(1-D(x,G(x)))\right] \quad (1)\end{aligned}$$

式の入力は 2 つあり，x は入力画像，y は正解画像を表す．また，p_{data} は学習データの分布を示し，p_{data} からデータをサンプリングする過程を $x, y \sim p_{data}(x, y)$ とする．$G(\cdot)$ は生成器が出力するセグメンテーション結果，$D(\cdot)$ は識別器が出力する本物か偽物か判断した結果を表す．この式を基にして以下の Min-Max の目的関数で敵対的学習の最適化が行われる．

$$arg\min_{G}\max_{D}\pounds_{adv}(G,D) \quad (2)$$

生成器 G が出力するセグメンテーション画像 $G(x)$ が識別器を騙すように $\pounds_{adv}$ を最小化し，識別器 D は生成器の出力 $G(x)$ を偽，正解画像 y を真と判断できるように $\pounds_{adv}$ を最大化する．この中でも生成器に関する Adversarial Loss は以下のようになる．

$$\pounds_{adv}(G) = \mathbb{E}_{x\sim p_{data}(x)}\left[log(1-D(x,G(x)))\right] \quad (3)$$

これらの損失関数を合わせて最終的な生成器の損失関数は以下のようになる．

$$\begin{aligned}\pounds_{total} = \pounds_{CE} + \pounds_{pda1} + \pounds_{pda2} + \\ \pounds_{pda3} + \lambda_{adv}\pounds_{adv}(G) \quad (4)\end{aligned}$$

ここで，λ_{adv} はハイパーパラメータであり，pix2pix[5] をはじめとして識別器が強くなりすぎないように弱く設定されている．本稿においても $\lambda_{adv} = 0.1$ として実験を行った．また，4.3. 節では値を変更した場合と比較する．

4. 評価実験

本節では評価実験の結果を示す．本実験では 2 つの細胞画像データセットと 3 つの眼底画像から血管領域抽出のデータセットを用いて評価する．本実験の評価指標は Intersection over Union(IoU)，F 値，Precision，Recall を採用した．IoU は各クラスに対して画素毎に重なり率を算出するので，各クラスを厳密に検出を行いたい細胞画像のセマンティックセグメンテーションにおいて適している．また，F 値は検出したい物体が検出できているか評価したい眼底画像の血管領域抽出において適している．また，Precision は適合率を示しており，正解と予測した結果の正確性を示す．また，Recall は再現率を示し，正解の取りこぼしの少なさを示す評価指標である．IoU，F 値，Precision，Recall は次のような式で表すことができる．

$$IoU = \frac{TP}{TP+FP+FN} \quad (5)$$

$$\begin{aligned}F\text{ 値} &= \frac{TP}{TP+\frac{1}{2}(FP+FN)} \\ &= \frac{2(Precision\cdot Recall)}{Precision+Recall} \quad (6)\end{aligned}$$

$$Precision = \frac{TP}{TP+FP} \quad (7)$$

$$Recall = \frac{TP}{TP+FN} \quad (8)$$

ここで TP は True Positive，FP は False Positive，FN は False Negative を表す．

本実験では Optimizer に Adam を採用した．学習係数は $\alpha = 10^{-3}$，$\beta_1 = 0.9$，$\beta_2 = 0.999$，$\varepsilon = 10^{-8}$

とした．本実験に使用した GPU は GeForce RTX 2080 Ti を 1 枚で学習を行った．実験方法はいずれのデータセットにおいてもデータ枚数が少なく，適切に汎化性能を評価するために評価画像を固定し，学習と検証を入れ替える 5 分割交差検証により実験を行った．また，検証画像に対して細胞画像は Mean IoU が，眼底画像の血管領域抽出は F 値が最も高い結果を示した学習時の生成器のモデルを評価時に採用した．実験結果に対する評価指標の標準偏差を確認するために同様の実験を初期値を変更して 3 回行い，合計 15 回の平均値により評価を行った．

4.1. 節では今回使用するデータセットについて述べ，4.2. 節では評価実験の結果を示す．最後に提案手法が効果的であることを示す Ablation studies を 4.3. 節に示す．

4.1. データセット

本節では実験に用いるデータセットについて説明する．それぞれのデータセットでの学習，検証，評価をどのように定義し，画像サイズの正規化や学習時のバッチサイズについて説明する．

WBC White Blood Cell(WBC)[6] は白血球の細胞核，細胞質，赤血球を含むその他背景の 3 クラスを含む 300 × 300 画素の 100 枚のデータセットである．本実験ではバイリニア補間法により 302×302 画素にリサイズ処理を行い，学習を 64 枚，検証を 16 枚，評価を 20 枚として分類し，バッチサイズは 8 で学習を行った．

ショウジョウバエの細胞画像 ショウジョウバエの細胞画像 [7] は細胞膜，細胞質，ミトコンドリア，シナプスの 4 クラスを含む全 20 枚のデータセットである．画像サイズは 1024 × 1024 画素である．本実験では 12 枚を学習画像，3 枚を検証，5 枚を評価として分類し，元画像を重なりのない 256 × 256 画素の 16 等分にし，学習 192 枚，検証 32 枚，評価 48 枚とした．また，バッチサイズを 12 として学習を行う．

DRIVE DRIVE Dataset[8] は眼底画像から血管領域とその他背景の 2 クラス分類である．565 × 584 画素の 40 枚で構成されており，公式に学習セット 20 枚と評価セット 20 枚で分けられている．本実験では 512 × 512 画素にバイリニア補間法でリサイズ処理を行い，学習セットから学習 16 枚，検証 4 枚に分け，バッチサイズを 2 で学習を行う．

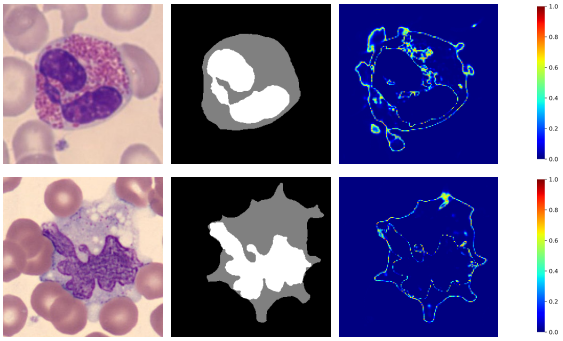


図 4　PDA Module の Attention Map の可視化

STARE STARE Dataset[9] は DRIVE と同様の眼底画像からの血管領域抽出であり，画像サイズは 700 × 605 画素の 20 枚で構成されている．本実験では学習 12 枚，検証 3 枚，評価 5 枚に分け，バイリニア補間法で 512 × 512 画素にリサイズ処理を行い，バッチサイズ 2 で学習を行う．

CHASE_DB1 CHASE_DB1[10] は 14 人の左右の眼底画像 28 枚からの血管領域抽出であり，画像サイズは 999 × 960 画素のデータセットである．本実験では学習 16 枚，検証 4 枚，評価 8 枚に分け，バイリニア補間法で 512 × 512 画素にリサイズ処理を行い，バッチサイズ 2 で学習を行う．

4.2. 実験結果

WBC の評価画像での PDA Module で作成した Attention Map の可視化を図 4 に示す．それぞれの Attention Map から特にクラスが隣接するエッジ部分に対して強調されていることが分かる．このように高難易度の画素が強調されることにより，生成器の学習においてはエッジ情報を重視し，識別器ではエッジ情報から真偽を判断することができていることが分かる．

表 1 に IoU，表 2 に Precision，Recall，F 値での WBC の評価実験結果を示す．表中の赤字は最も精度の高いことを表し，青い太字は 2 番目に精度が良いことを表している．表 1 より我々の提案した APDA-Net が従来手法と比較して精度が向上していることが分かる．特に敵対的学習と識別器へのリークを行わない従来の PDA Module からすべてのクラスにおいて IoU による精度向上が確

表 1　WBC[6] に対する精度の比較結果

手法	細胞質 [%]	細胞核 [%]	背景 [%]	Mean IoU[%]
U-Net[12]	69.95 ± 7.89	89.31 ± 2.06	94.42 ± 2.49	84.56 ± 3.86
pix2pix[5]	71.62 ± 5.56	89.45 ± 1.88	95.21 ± 1.74	85.43 ± 2.70
SAGAN[26]	**74.32** ± 4.92	**89.94** ± 1.37	**95.87** ± 1.50	**86.71** ± 2.35
PDA Module[27]	72.80 ± 10.13	88.40 ± 4.19	95.64 ± 3.23	85.62 ± 5.45
APDA-Net(Ours)	**76.39** ± 6.86	**89.64** ± 2.82	**96.68** ± 1.68	**87.57** ± 3.34

表 2　WBC[6] に対する精度の比較結果

Precition Recall F 値	細胞質 [%]	細胞核 [%]	背景 [%]
U-Net[12]	76.30 ± 8.51	**92.47** ± 2.81	**99.28** ± 0.30
	89.26 ± 2.56	96.42 ± 2.29	95.07 ± 2.62
	82.06 ± 5.56	94.34 ± 1.16	97.11 ± 1.32
pix2pix[5]	79.66 ± 6.72	92.38 ± 1.99	99.03 ± 0.48
	87.82 ± 3.16	96.64 ± 2.54	96.13 ± 2.01
	83.34 ± 3.83	94.42 ± 1.06	97.54 ± 0.92
SAGAN[26]	81.74 ± 6.21	**92.81** ± 1.55	**99.21** ± 0.20
	89.27 ± 2.26	96.71 ± 1.87	**96.61** ± 1.60
	85.17 ± 3.29	**94.70** ± 0.77	**97.89** ± 0.79
PDA Module[27]	**81.79** ± 10.79	90.80 ± 5.23	99.23 ± 0.38
	86.88 ± 6.15	**97.25** ± 1.72	96.37 ± 3.43
	83.82 ± 7.47	93.79 ± 2.44	97.74 ± 1.75
APDA-Net(Ours)	**85.62** ± 6.68	91.95 ± 3.58	99.11 ± 0.36
	87.65 ± 4.74	**97.34** ± 1.25	**97.53** ± 1.65
	86.44 ± 4.47	**94.51** ± 1.60	**98.31** ± 0.87

認できる．表 2 においても APDA-Net は細胞質の Precision が最も高く，正確性が向上していることから高難易度画素を重視して学習していることを示している．また，細胞核と背景の Recall が向上していることから従来手法では細胞質と判断していた画素が細胞核と背景にクラスを割り当てて取りこぼしが少なくなっていることを示している．Precision と Recall の調和平均である F 値で比較しても IoU と同様に APDA-Net の精度が従来手法と比較して良い結果を出力していることが確認できる．図 5 に WBC での出力結果を示す．セグメンテーション結果において白色は細胞核，灰色は細胞質，黒色は赤血球を含む背景を示す．従来手法では背景の赤血球に反応してしまい過剰に検出した細胞質の領域が APDA-Net では正解画像に近い出力になっていることが分かる．また，細胞核も従来手法では細胞質と混同してしまい，正確な領域を示すことが出来ていなかった．しかし，APDA-Net では細かな細胞核の形状を捉えることに成功し，良好な結果を出力している．

表 3 にショウジョウバエの細胞画像での IoU の実験結果を示す．ショウジョウバエのデータセットにおいても APDA-Net が従来手法と比較しても良い結果を示していることが分かる．特に他のクラスと比較して検出しづらいシナプスの精度が最も向上していることから高難易度画素の学習が効果的であることが分かる．

表 4 に DRIVE，STARE，CHASE_DB1 での血

表 3　ショウジョウバエの細胞画像 [7] に対する精度の比較結果

手法	細胞膜 [%]	ミトコンドリア [%]	シナプス [%]	細胞質 [%]	Mean IoU[%]
U-Net[12]	74.24 ± 1.76	71.01 ± 5.88	43.08 ± 3.33	92.03 ± 0.41	70.09 ± 2.57
pix2pix[5]	75.17 ± 1.01	74.80 ± 2.08	42.89 ± 3.08	92.24 ± 0.18	71.28 ± 1.10
SAGAN[26]	**76.64** ± 0.58	**78.86** ± 1.38	43.33 ± 2.98	**92.70** ± 0.13	**72.88** ± 0.86
PDA Module[27]	75.19 ± 0.89	74.50 ± 2.12	**46.11** ± 2.53	92.10 ± 0.26	71.98 ± 1.02
APDA-Net(Ours)	**76.48** ± 0.89	**76.71** ± 1.69	**46.40** ± 3.40	**92.52** ± 0.13	**73.03** ± 1.13

表 4　眼底画像 [8, 9, 10] に対する精度の比較結果

Dataset	手法	Precition[%]	Recall[%]	F 値 [%]	IoU[%]
DRIVE[8]	U-Net[12]	87.38 ± 3.72	**79.65** ± 3.14	**83.20** ± 0.82	**71.24** ± 1.20
	pix2pix[5]	**87.77** ± 2.97	78.43 ± 3.61	82.72 ± 1.41	70.56 ± 2.03
	SAGAN[26]	84.10 ± 4.51	71.62 ± 3.36	77.20 ± 1.86	62.91 ± 2.47
	PDA Module[27]	86.02 ± 3.27	**80.03** ± 2.78	82.82 ± 1.22	70.70 ± 1.75
	APDA-Net(Ours)	**87.55** ± 2.18	79.55 ± 2.52	**83.30** ± 0.91	**71.39** ± 1.33
STARE[9]	U-Net[12]	**86.99** ± 2.43	55.17 ± 2.36	67.46 ± 1.54	50.92 ± 1.75
	pix2pix[5]	**87.64** ± 1.79	54.06 ± 2.84	66.83 ± 2.46	50.24 ± 2.73
	SAGAN[26]	84.35 ± 2.42	43.61 ± 2.86	57.43 ± 2.56	40.32 ± 2.49
	PDA Module[27]	86.56 ± 2.11	**55.81** ± 4.98	**67.70** ± 3.40	**51.28** ± 3.94
	APDA-Net(Ours)	86.83 ± 1.72	**57.27** ± 4.89	**68.87** ± 3.39	**52.63** ± 3.93
CHASE_DB1[10]	U-Net[12]	77.41 ± 1.13	73.47 ± 3.49	75.34 ± 1.80	60.47 ± 2.28
	pix2pix[5]	**77.80** ± 1.41	73.30 ± 2.95	75.44 ± 1.38	60.58 ± 1.77
	SAGAN[26]	**77.60** ± 3.73	64.91 ± 4.90	70.55 ± 3.47	54.61 ± 4.06
	PDA Module[27]	76.66 ± 2.41	**75.25** ± 2.69	**75.90** ± 1.69	**61.19** ± 2.17
	APDA-Net(Ours)	76.23 ± 1.92	**76.57** ± 2.60	**76.35** ± 1.40	**61.77** ± 1.80

管領域抽出の実験結果を示す．いずれのデータセットにおいても APDA-Net が F 値と IoU 共に最も高い精度を出力していることが分かる．STARE と CHASE_DB1 において血管の Recall が向上しており，血管領域の取りこぼしが少なくなっており，多くの血管領域を抽出したいこれらのデータセットにおいて良い結果を示していると言える．また，細胞画像では良い精度を出力していた SAGAN[26] がこれらのデータセットにおいては最も悪い精度であることから APDA-Net のほうが汎用性に優れていることが分かる．

これらのデータセットでの出力画像を図 6 に示す．ショウジョウバエの細胞画像では 16 分割した画像から元の 1 枚になるようにつなぎ合わせて図に示した．ショウジョウバエの細胞画像のセグメンテーション結果において赤色は細胞膜，緑色はミトコンドリア，青色はシナプス，黒色は細胞質を示す．黄色の枠を見ると，U-Net と PDA Module では上手く捉えることが出来なかったミトコンドリアを APDA-Net では検出していることが分かる．また，紫色の枠では従来手法でミトコンドリアと識別してしまっているが，APDA-Net では細胞膜と間違えることなく出力していることが分かる．白色の枠では検出が難しいシナプスを APDA-Net だけが捉えていることが分かる．眼底画像の結果で DRIVE は APDA-Net の出力画像はノイズのような出力部分が減っていることが分かる，特に STARE と CHASE_DB1 において APDA-Net が血管領域の取りこぼしが少ないため，血管が途切れることなく繋がって出力されている箇所が多くなっていることが分かる．

表 5　機構の一部を削除した精度の比較結果

手法	細胞質 [%]	細胞核 [%]	背景 [%]	Mean IoU[%]
Attention 正規化なし	74.25 ± 6.95	**90.04** ± 1.66	95.80 ± 2.20	**86.70** ± 3.32
生成器の Attention なし	**74.26** ± 8.84	89.41 ± 2.59	**95.98** ± 2.64	86.55 ± 4.23
識別器の Attention なし	73.91 ± 8.04	**90.14** ± 2.16	95.57 ± 2.53	86.54 ± 3.66
APDA-Net(Ours)	**76.39** ± 6.86	89.64 ± 2.82	**96.68** ± 1.68	**87.57** ± 3.34

表 6　識別器への接続方法の違いによる精度の比較結果

手法	細胞質 [%]	細胞核 [%]	背景 [%]	Mean IoU[%]
Concat	69.07 ± 4.41	88.63 ± 1.06	94.42 ± 1.45	84.04 ± 1.99
Add	70.86 ± 6.91	88.84 ± 2.47	94.99 ± 2.15	84.90 ± 3.38
1×1Conv	71.72 ± 6.54	**89.51** ± 1.48	95.03 ± 2.14	85.42 ± 3.19
SE-Net[21]	**73.34** ± 6.80	89.66 ± 1.07	**95.50** ± 2.39	**86.17** ± 3.19
Light Attention[29]	70.45 ± 4.82	89.46 ± 1.65	94.75 ± 1.49	84.89 ± 2.40
Self-Attention[24]	71.35 ± 6.13	88.62 ± 2.27	95.29 ± 1.89	85.08 ± 3.13
Source-Target-Attention[24]	62.84 ± 8.25	88.48 ± 1.62	91.84 ± 3.76	81.05 ± 4.09
APDA-Net(Ours)	**76.39** ± 6.86	**89.64** ± 2.82	**96.68** ± 1.68	**87.57** ± 3.34

4.3. Ablation Study

我々の提案手法の機構が効果的であることを確認するために一部の機構を削除した際の実験結果を表 5 に示す．Attention Map を $0 \leq x \leq 1$ の範囲になるように正規化する機構と Attention Map を生成器内もしくは識別器内で使用しない手法と比較するとどの機構を削除した際でも我々の APDA-Net から精度が低下していることが分かる．これらの結果から提案した機構が精度向上に寄与していることが確認できる．

生成器から識別器へのリークする際の接続方法による精度への影響を実験で確認した．Attention Mechanism を用いない手法として U-Net のように特徴マップを結合させる手法，単純に足し合わせる手法と 1×1 畳み込み処理を行ってから足し合わせる手法と比較する．そして，Attention Mechanism を用いる手法として特徴マップのチャンネル間での抑制と強調を行う SE-Net[21] を生成器の特徴マップに適用し，識別器の特徴マップに足し合わせる手法と 3×3 畳み込み処理によって特徴マップの空間とチャンネルにおいて抑制と強調を行う Light Attention[29] を生成器の特徴マップに適用し，識別器の特徴マップに足し合わせる手法，生成器の特徴マップから Attention Map を作成して識別器に足し合わせる Self-Attention[24]，生成器と識別器の関連性から Attention Map を作成する Source-Target-Attention[24] と比較する．表 6 に接続方法の違いによる実験結果を示す．表 6 から PDA Module による識別器へのリークがもっとも精度が高いことが分かる．これは他の手法と比較して我々の提案手法が最も効果的に識別器に情報を伝達できることを示している．

さらに，識別器からの損失関数である Adversarial Loss のハイパーパラメータ λ_{adv} の精度への影響を確認するために実験を行った．表 7 に実験結果を示す．ハイパーパラメータを低くした際には精度は大きく変わらないものの標準偏差が大きくなっていることが確認できる．また，Adversarial Loss は他の $\pounds_{CE}$ などと同様の重みを与えた際には識別器の影響が大きくなることで精度が落ちることが分かる．

5. まとめ

本論文では，セマンティックセグメンテーションにおいて画素毎の難易度に注目した Attention Machanism と識別器に高難易度の画素をリークする敵対的学習を行う APDA-Net を提案した．APDA-Net により難しい画素に注目した学習を行うことでセグメンテーション精度の向上を実現した．

表 7　λ_{adv} の値を変更した精度の比較結果

手法	細胞質 [%]	細胞核 [%]	背景 [%]	Mean IoU[%]
$\lambda_{adv} = 0.01$	**74.35** ± 10.28	**90.03** ± 3.61	**95.70** ± 2.40	**86.69** ± 5.17
$\lambda_{adv} = 1$	72.36 ± 8.81	89.01 ± 2.75	95.47 ± 2.45	85.61 ± 4.39
$\lambda_{adv} = 0.1$(Ours)	**76.39** ± 6.86	**89.64** ± 2.82	**96.68** ± 1.68	**87.57** ± 3.34

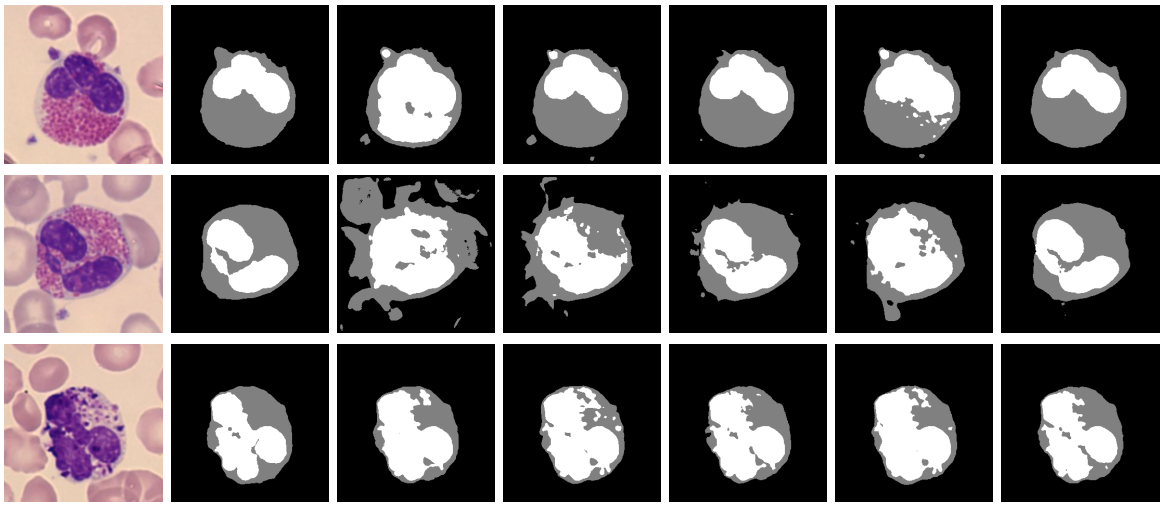


図 5　WBC[6] でのセグメンテーション結果

今後の展望として高難易度の画素と周辺の関連するクラスとの文脈情報を利用した学習方法を提案手法に導入し，セグメンテーション精度向上を目指したい．

利益相反の有無

なし

参考文献

[1] Long J, Shelhamer E, Darrell T: Fully Convolutional Networks for Semantic Segmentation, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3431–3440, 2015.

[2] LeCun Y, Bottou L, Bengio Y, et al.: Gradient-Based Learning Applied to Document Recognition, Proceedings of the IEEE, Vol. 86, No. 11, pp. 2278–2324, 1998.

[3] He K, Zhang X, Ren S, et al.: Deep Residual Learning for Image Recognition, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770–778, 2016.

[4] Huang G, Liu Z, Van Der Maaten L, et al.: Densely Connected Convolutional Networks, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 4700–4708, 2017.

[5] Isola P, Zhu J.Y, Zhou T, et al.: Image-to-Image Translation with Conditional Adversarial Networks, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1125–1134, 2017.

[6] Zheng X, Wang Y, Wang G, et al.: Fast and robust segmentation of white blood cell images by self-supervised learning, Micron, Vol. 107, pp. 55–71, 2018.

[7] Gerhard S, Funke J, Martel J, et al.: Segmented anisotropic ssTEM dataset of neural tissue, 11 2013.

[8] Staal J, Abràmoff M.D, Niemeijer M, et al.: Ridge-based vessel segmentation in color im-

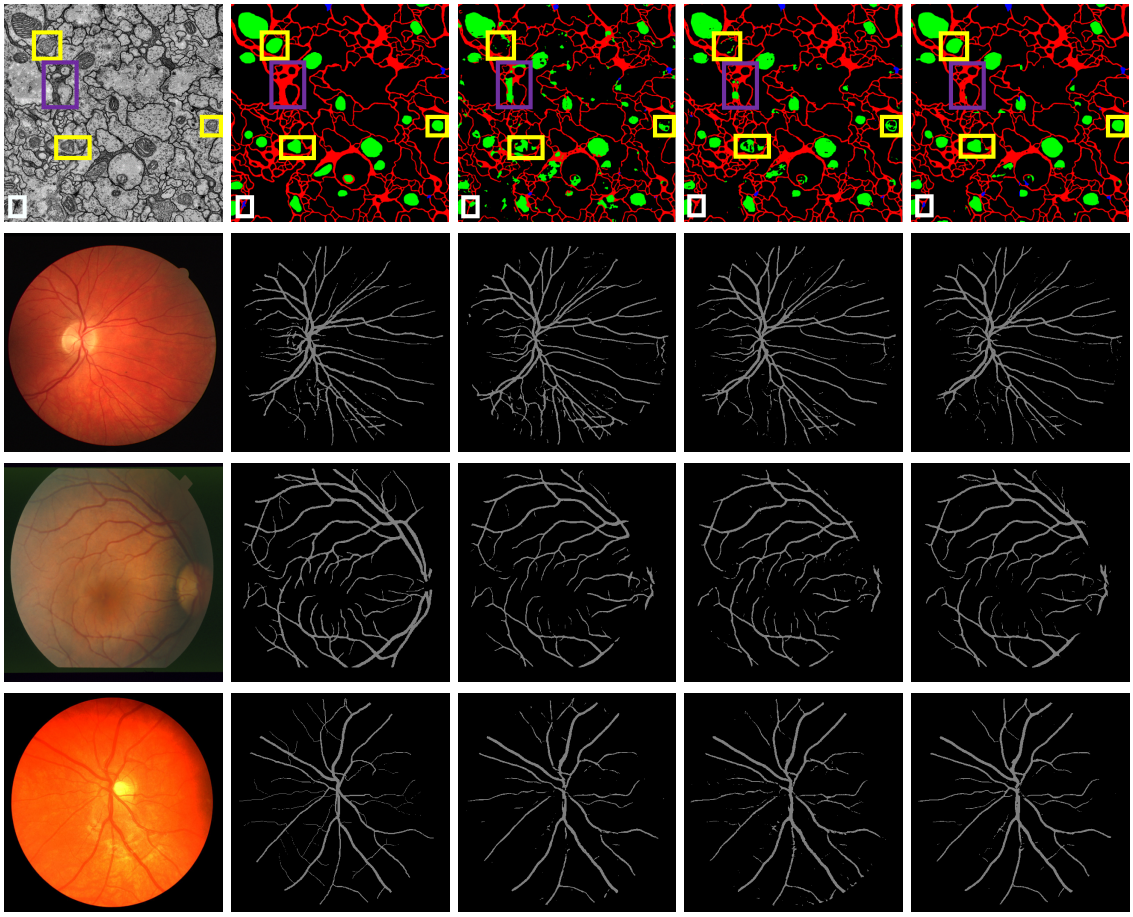


図 6　セグメンテーション結果（上からショウジョウバエの細胞画像 [7]，DRIVE[8]，STARE[9]，CHASE_DB1[10]）

ages of the retina, IEEE transactions on medical imaging, Vol. 23, No. 4, pp. 501–509, 2004.

[9] Hoover A, Kouznetsova V, Goldbaum M: Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response, IEEE Transactions on Medical imaging, Vol. 19, No. 3, pp. 203–210, 2000.

[10] Owen C.G, Rudnicka A.R, Mullen R, et al.: Measuring retinal vessel tortuosity in 10-year-old children: validation of the computer-assisted image analysis of the retina (CAIAR) program, Investigative ophthalmology & visual science, Vol. 50, No. 5, pp. 2004–2010, 2009.

[11] Badrinarayanan V, Kendall A, Cipolla R: SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 39, No. 12, pp. 2481–2495, 2017.

[12] Ronneberger O, Fischer P, Brox T: U-Net: Convolutional Networks for Biomedical Image Segmentation, International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 234–241. Springer, 2015.

[13] Simonyan K, Zisserman A: Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv:1409.1556, 2014.

[14] Goodfellow I, Pouget-Abadie J, Mirza M,

et al.: Generative Adversarial Nets, Advances in Neural Information Processing Systems, pp. 2672–2680, 2014.

[15] Taigman Y, Polyak A, Wolf L: Unsupervised Cross-Domain Image Generation, In International Conference on Learning Representations, 2017.

[16] Choi J, Kim T, Kim C: Self-Ensembling with GAN-based Data Augmentation for Domain Adaptation in Semantic Segmentation, Proceedings of the IEEE International Conference on Computer Vision, pp. 6830–6840, 2019.

[17] Hoffman J, Tzeng E, Park T, et al.: CyCADA: Cycle-Consistent Adversarial Domain Adaptation, International Conference on Machine Learning, pp. 1994–2003, 2017.

[18] Luc P, Couprie C, Chintala S, et al.: Semantic Segmentation Using Adversarial Networks, Advances in Neural Information Processing Systems Workshops, 2016.

[19] Sawada K, Hotta K, Imanishi A, et al.: Segmentation of Cell Images by Leaking the Information of Discriminator to Generator, International Conference of the IEEE Engineering in Medicine and Biology Society, 2018.

[20] Wang F, Jiang M, Qian C, et al.: Residual Attention Network for Image Classification, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3156–3164, 2017.

[21] Hu J, Shen L, Sun G: Squeeze-and-Excitation Networks, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 7132–7141, 2018.

[22] Zhou B, Khosla A, Lapedriza A, et al.: Learning Deep Features for Discriminative Localization, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2921–2929, 2016.

[23] Fukui H, Hirakawa T, Yamashita T, et al.: Attention Branch Network: Learning of Attention Mechanism for Visual Explanation, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 10705–10714, 2019.

[24] Vaswani A, Shazeer N, Parmar N, et al.: Attention Is All You Need, Advances in Neural Information Processing Systems, pp. 5998–6008, 2017.

[25] Fu J, Liu J, Tian H, et al.: Dual Attention Network for Scene Segmentation, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3146–3154, 2019.

[26] Zhang H, Goodfellow I, Metaxas D, et al.: Self-Attention Generative Adversarial Networks, International Conference on Machine Learning, pp. 7354–7363, 2019.

[27] Matsuzuki D, Hotta K: Cell Image Segmentation Using Attention Module at Each Layer, Proceedings of The 20th International Conference on Systems Biology, 2019.

[28] Miyato T, Kataoka T, Koyama M, et al.: Spectral Normalization for Generative Adversarial Networks, In International Conference on Learning Representations, 2018.

[29] Hiramatsu Y, Hotta K: Semantic Segmentation using Light Attention Mechanism, Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, pp. 622–625, 2020.

Semantic Segmentation by Adversarial Training using Difficult Pixels Attention Mechanism

Hiroki TSUDA[*1], Kazuhiro HOTTA[*1]

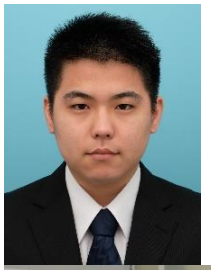
*1 Meijo University

Semantic segmentation is to assign class label to all pixels in an input image. By using a segmentation model as a generator and adding a discriminator such as GANs, we can estimate the segmentation result more accurately than a segmentation model alone. We propose an Adversarial Pixel-wise Difficulty Attention Network that leaks difficult to identify pixels to the discriminator as an attention map. Leakage from the generator allows the discriminator to judge whether the generated image or ground truth accurately in comparison with conventional methods, and on that basis the segmentation model will learn to fool the discriminator through adversarial training. In experiments, we compared the accuracy of the proposed method with that of the conventional method using cell and fundus image datasets. We confirmed that the proposed method outperformed the conventional method on intersection over union and F-measure.

Key words: Semantic Segmentation, Attention Mechanism, Pixel Difficulty, Adversarial Training, Deep Learning

著者紹介

和文でも英文でも可。



津田　大輝（つだ　ひろき）
2019年名城大学・理工学部・電気電子工学科卒業．2021年同大大学院博士前期課程・理工学研究科電気電子工学専攻修了予定．現在，敵対的学習と注意機構を用いたセマンティックセグメンテーションの研究に従事．



堀田　一弘（ほった　かずひろ）
1997年埼玉大学工学部卒．2002年同大学院博士後期課程修了．博士(工学)．1999年～2002 年日本学術振興会特別研究員(DCI)．2002 年電気通信大学情報通信工学科助手，2007年同大学助教，2010年から名城大学理工学部電気電子工学科准教授．2018年から同大学教授．2012年メリーランド大学Visiting Scholar．パターン認識，コンピュータビジョンの研究に従事．電子情報通信学会，情報処理学会，IEEE Computer Society各会員．

Reconstruction method of few projections tomosynthesis for 3D angiography

Kensuke HORI[*1], Takahisa KOIKE[*1], Kiichi TADANO[*1], Takeyuki HASHIMOTO[*1]

Abstract

Interventional radiology (IR) has widely known as a technique used in various disease of the brain, and x-ray 2D projection images are mostly used. Few projections tomosynthesis is a method that reconstructs tomogram using projection data acquired from few-views and limited-angle. In this method, the scanning is very quick and it has possibility of obtaining the tomogram of how the contrast agent is flowing. The aim of this study is to improve the resolution of y-direction perpendicular to the detector plane and suppress the artifact caused by insufficient views and limited-angle. In the numerical simulation, the distance between the source plane and the detector plane was 1000 mm, the detector element size and matrix size were 1.0×1.0 mm^2 and 512×512 pixels, and φ6.0-mm tube phantom arranged along z-direction was used (Fig. 1). The algebraic reconstruction technique (ART) and the projection onto convex set (POCS) method were applied for tomosynthesis image reconstruction and compared by the reconstructed images. In static imaging, the bow tie shaped pseudo shadow was appeared in x-y plane images reconstructed by ART (Fig. 2). The full width at half maximum (FWHM) for x- and y-direction was 5.5 mm and 16 mm, respectively. In the images reconstructed by POCS method, the pseudo shadow was suppressed and FWHM for x- and y-direction was improved to 2.6 mm and 7.7 mm, respectively. In dynamic imaging, the state of contrast agent flowing in tube was able to be depicted and POCS method was better than ART as in static imaging. In conclusion, POCS method enabled to improve the image quality compared with ART for few projections tomosynthesis of 3D angiography.

Keywords：Few projections tomosynthesis, 3D angiography, flow imaging, numerical simulation

1. Purpose

Interventional radiology (IR) has widely known as a technique used in various disease of the brain, and x-ray 2D projection images are mostly used [1, 2]. Tomosynthesis is a method that reconstructs tomogram using projection data acquired from limited-angle [3]. It is known that the quality of diagnosis can be improved by tomosynthesis in many medical fields. We have developed and reported few projections tomosynthesis that acquired only four projections [4, 5]. In this method, the scanning is very quick and it has possibility of obtaining the tomogram of how the contrast agent is flowing. However, the image quality is distorted by limitation when the data is acquired. The limitation is the projection angle and the projection number that can be imaged while blood is flowing. The aim of this study is to improve the

*1 Kyorin University Graduate School of Health Sciences
〔5-4-1 Shimorenjaku, Mitaka, Tokyo 181-8612, Japan〕
e-mail: hori1911n@std.kyorin-u.ac.jp

resolution of direction that is perpendicular to the detector plane and suppress the artifact caused by insufficient views and limited-angle.

2. Method

The numerical simulation of few projections tomosynthesis for 3D angiography was performed to quantitatively evaluate the image quality. The geometry of numerical simulation was shown in Fig. 1. The static and dynamic scan was simulated. The mean velocity of brain blood flow was reported to about 55 cm/sec [6] and that of contrast agent was assumed as the same. The four x-ray sources were arranged in a circle. The frame rate of tomosynthesis scanning was 15 frame/sec, because the time resolution of scanning was limited by response of detector that can be realized by using a commercial flat panel detector. The distance between the source and the detector, pixel size and matrix size of detector are 1000 mm, 1.0×1.0 mm^2 and 512×512 pixels, respectively. The algebraic reconstruction technique (ART) and projection onto convex set (POCS) were used for the tomosynthesis reconstruction.

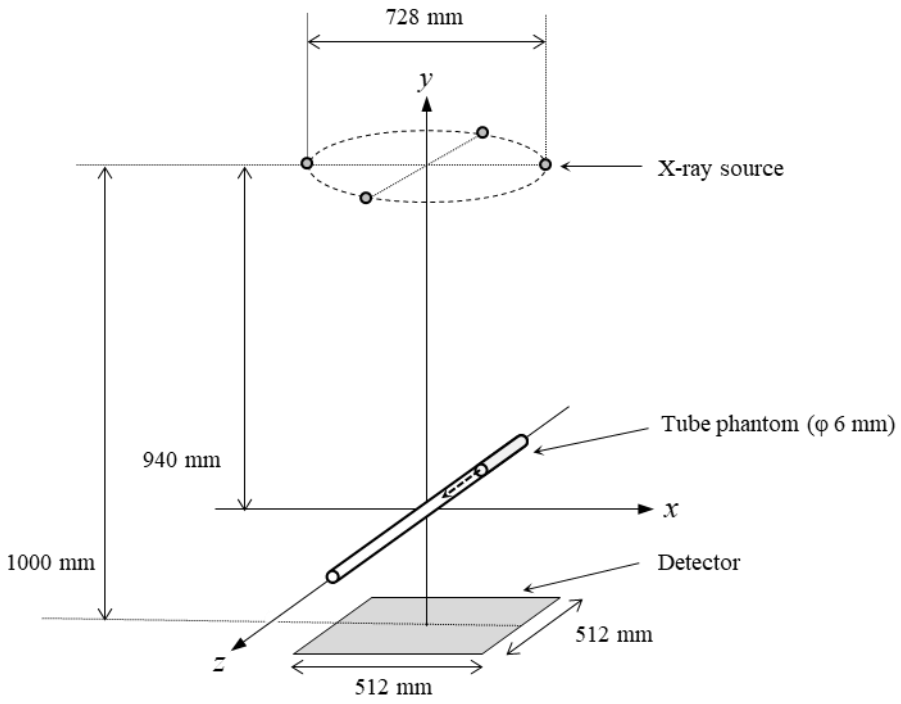


Fig. 1 The geometry of numerical simulation.

The tomosynthesis from four projections for 3D angiography was performed. The contrast agent was flowed to the z direction in the tube phantom in static (0 cm/sec) and dynamic (55 cm/sec) conditions.

3. Results

The images reconstructed from the static condition were shown in Fig. 2(a). In the images reconstructed by ART, the bow tie shaped pseudo shadow was appeared in xy-plane. The full width at half maximum (FWHM) for x- and y-direction were 5.5 mm and 16 mm, respectively. In the images reconstructed by POCS, the pseudo shadow was suppressed and FWHM for x- and y-direction were improved to 2.6 mm

and 7.7 mm, respectively.

The images reconstructed from the dynamic condition were shown in Fig. 2(b). The state of contrast agent flowing in tube was able to be depicted. The time resolution was sufficient to reconstructed tomogram of contrast agent flowing. In the images reconstructed by ART, specific artifacts were appeared in the area, where projection number was insufficient in Fig. 2(ii) and (iii). In the images reconstructed by POCS, the specific artifacts were suppressed. In addition, the volume rendering image was able to be constructed from reconstructed images.

4. Conclusion

The few projections tomosynthesis was able to realize the short scanning time that is sufficient to reconstruct the tomogram of contrast agent flowing. In the ART method, the bow tie shaped pseudo shadow was appeared in the reconstruction images, however, in the POCS method, the shadow was suppressed. The POCS method was considered as useful for the reconstruction of few projections tomosynthesis angiography.

References

[1] Etard C, Bigand E, Salvat C et al: Patient dose in interventional radiology: a multicentre study of the most frequent procedures in France. Eur Radiol **27**: 4281–4290 2017

[2] Kai D, Bo W, Mo C et al: Realization of real-time x-ray stereoscopic vision during interventional procedures. Sci Rep **8**: 15852, 2018

[3] Dobbins JT III and Godfrey DJ: Digital x-ray tomosynthesis: current state of the art and clinical potential. Phys. Med. Biol. **48**: R65 2003

[4] Hori K, Dong J, Okamoto H et al: Preliminary study to detect pulmonary nodules by tomosynthesis using a small number of projections. Transactions of the Japanese Society for Medical and Biological Engineering **56**: 16-21 2018

[5] Hori K, Dong J, Okamoto H et al: Optimum total scan angle for four-projection chest digital tomosynthesis to detect lung nodules. Transactions of the Japanese Society for Medical and Biological Engineering **56**: 260-264 2018

[6] Hennerici M, Rautenberg W, Sitzer G et al: Transcranial doppler ultrasound for the assessment of intracranial arterial flow velocity-Part I. Examination technique and normal values. Surg Neurol **27**: 439-448 1987

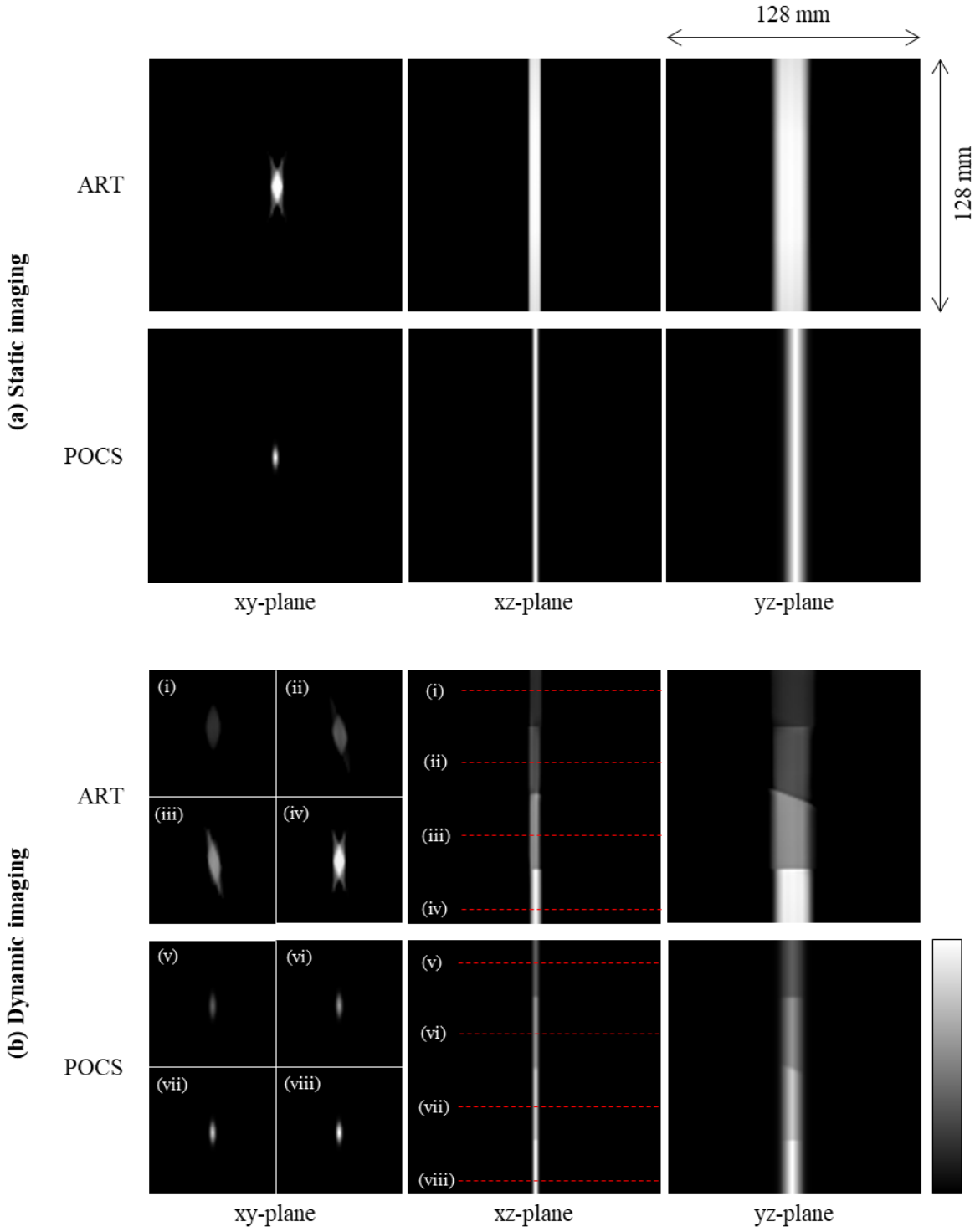


Fig. 2 The images reconstructed by ART and POCS.

The contrast agent was static (a) and dynamic condition (b). The red dotted line represents the slice location of the images shown in x-y plane. (i)-(viii) are the images with the center of the images cropped.

Prediction of oxygen extraction fraction map by convolutional neural network learned with MR and PET images

Keisuke MATSUBARA[*1], Masanobu IBARAKI[*1], Yuki SHINOHARA[*1],
Noriyuki TAKAHASHI[*2], Hideto TOYOSHIMA[*1],
Toshibumi KINOSHITA[*1]

Abstract

Purpose: Oxygen extraction fraction (OEF) is an important biomarker to make a decision for treatments of steno-occlusive cerebrovascular disorder [1]–[3]. Positron emission tomography (PET) with oxygen-15 gas (^{15}O PET) is a gold-standard method to acquire quantitative OEF map [1], [4]. However, ^{15}O PET scans are not comfortable for patients because of long fixation time, invasive arterial sampling and radiation exposure.

We aimed to predict OEF map from magnetic resonance (MR) and PET images by convolutional neural network. To optimize the model to predict the OEF map, we demonstrated the effect of combinations of PET and MR images as input data in the prediction of OEF maps.

Methods: Cerebral blood flow on rest (CBF) and stressed with acetazolamide (sCBF), cerebral blood volume (CBV) and OEF maps were acquired by ^{15}O PET scans for 138 patients with chronic steno-occlusive disease. Routine MR images, T1, T2 and T2*-weighted images, were also acquired from the 138 patients. The PET and MR images for 113 patients were trained with U-Net [5]. The OEF maps were regarded as target to predict. Data augmentation with horizontal flip and rotation were performed. The training of 219 minibatches with 8 datasets were iterated with 100 epochs. Optimization of the weights were implemented with Adam algorithm [6]. Initial learning rate was set to 0.0002.

The prediction performance of the learned model were evaluated by five-fold cross validation. We compared predicted OEF maps, their structural similarity index (SSIM) and inter-class correlation (ICC) of ipsi-/contra-lateral ratio of OEF (rOEF) among the models learned with single image and combinations of PET and MR images. The model with the best performance in the cross validation was tested with the data for other 25 patients.

Results: The model learned with MR images and CBF map resulted in the best ICC [0.506 ± 0.154] whereas the model learned with MR images, CBF, sCBF and CBV maps resulted in the best SSIM [0.592 ± 0.012], as shown in Figure 1. The OEF maps predicted with the model learned with MRI and CBF [SSIM: 0.568 ± 0.014] were structurally similar to the model learned with full input data. We selected the model learned with MRI and CBF for the following test because correspondence of rOEF is more clinically important than structural similarity of OEF maps.

---

*1 Department of Radiology and Nuclear Medicine, Research Institute for Brain and Blood Vessels, Akita Cerebrospinal and Cardiovascular Center〔SenshuKubota-machi 6-10, Akita 010-0874, Japan〕
e-mail: matsubara@akita-noken.jp

*2 Preparing section for New Faculty of Medical Science, Fukushima Medical University

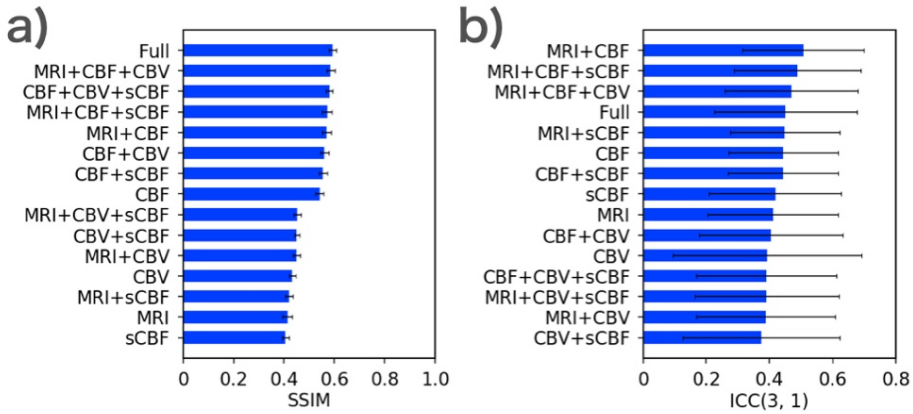


Figure 1 Average values of a) SSIM and b) ICC for each model learned with different combination of input data in the cross validation. Error bars indicate 95% confidence intervals among cross validation datasets. 'Full' indicates the model learned with all MR and PET (CBF, CBV and sCBF) images.

SSIM and ICC for OEF predicted for the test data with the model learned with all training data were 0.563 and 0.614, respectively. The contrast of predicted OEF maps was similar and but low relative to the real OEF maps, as shown in the upper image on Figure 2. Underestimation of rOEF was observed in the test data. In some cases, the learned model failed to predict quantitatively and structurally accurate OEF maps, as shown in the lower image on Figure 2.

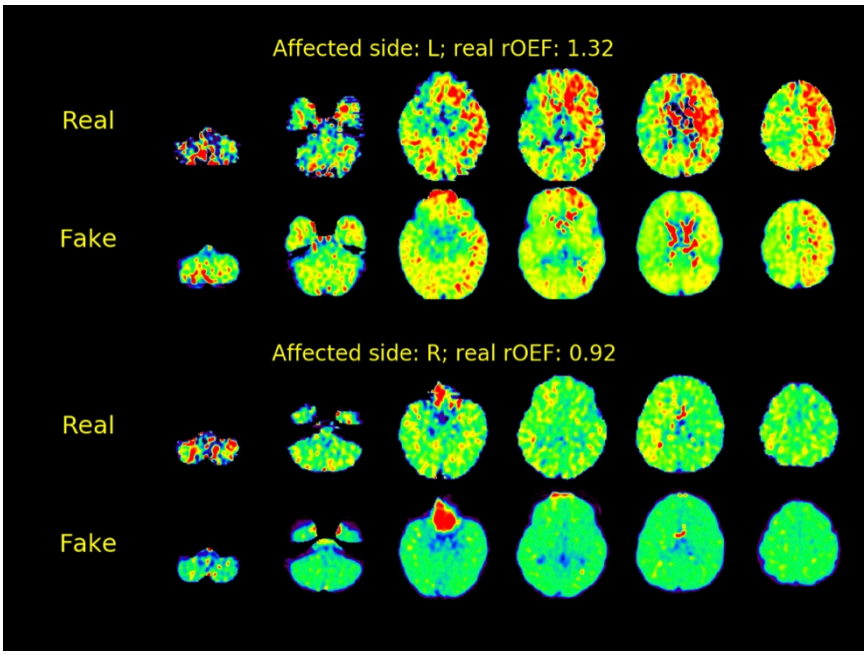


Figure 2 OEF maps successfully-predicted (upper) and failed to predict (lower) on test data.

Conclusion: These results suggest U-Net learned with routine MR images and CBF maps can predict qualitatively realistic OEF maps. Further improvement of quantitative accuracy to predict OEF maps is required.

Competing interests

The authors declare that they have no conflicts of interest.

References

[1] Okazawa H and T Kudo: Clinical impact of hemodynamic parameter measurement for cerebrovascular disease using positron emission tomography and 15O-labeled tracers. Ann Nucl Med **23**: 217–227, 2009

[2] Powers WJ, RL Grubb and ME Raichle: Physiological responses to focal cerebral ischemia in humans. Ann Neurol **16**: 546–552, 1984

[3] Yamauchi H, *et al.*: Evidence of misery perfusion and risk for recurrent stroke in major cerebral arterial occlusive diseases from PET. Journal of Neurology, Neurosurgery & Psychiatry **61**: 18–25, 1996

[4] Mintun MA, *et al.*: A quantitative model for the in vivo assessment of drug binding sites with positron emission tomography. Ann Neurol **15**: 217–227, 1984

[5] Ronneberger O, P Fischer and T Brox: U-Net: Convolutional Networks for Biomedical Image Segmentation. 2015

[6] Kingma D and J Ba: Adam: A Method for Stochastic Optimization. arXiv:14126980 [cs] 2014

MR 画像の位相推定を不要とする 深層学習型圧縮センシング再構成法の検討

大内 翔平[*1]，伊藤 聡志[*1]

要旨

MRI では，装置の磁場が厳密には不均一であること，組織間で磁化率が異なることなどが要因で，再生された MR 画像は実関数とはならず，一般に位相を含んだ複素画像となる．これまでに，CNN を利用した圧縮センシング再構成法が提案されているが，その多くは実関数を想定，もしくは緩やかな位相変化を推定する方法が多く，臨床画像のような複雑な位相変化を持つ画像の再構成を行った報告は少ない．本研究では，位相変化を含む現実的な画像再構成を目的とし，複素 MR 画像の実部と虚部をそれぞれ独立に計算できる信号収集パターンを採用することにより単一の CNN を利用して独立に実部と虚部を再構成できる新たな再構成法を提案する．既存の位相変化を考慮した反復的再構成法および，他の CNN 再構成法との比較検討を行った結果，提案法では様々な位相変化を持つ MR 画像の再構成を良好に行うことができ，信号収集率が低い場合に他の方法を上回る PSNR を得ることができた．

キーワード： 深層学習，圧縮センシング，位相画像

1. はじめに

MRI(Magnetic Resonance Imaging)では，生体に非侵襲かつ高品質な撮像が可能である．しかし，撮像に時間を要する問題があり，これまでに様々な撮像時間の高速法が提案されている．なかでも圧縮センシング(Compressed Sensing : CS)[1][2]の応用[3]による撮像時間の高速化は，ハードウェアの改修を必要とせず，撮像シーケンスとソフトウェアの改修のみで撮像時間を短縮できるため，実用化が進んでいる．圧縮センシングによる画像再構成（以下，CS-MRI とする）は，スパース性を満たした画像に対して，数理的反復処理による L1-L2 ノルムの最小化を行うことで実現する．そのため，再構成が長時間化し，再構成像に人工的様相が発生する問題がある．これに対し，近年では CNN(Convolutional Neural Network)を利用した再構成法（以下，CNN-CS とする）が多数提案されており[4]~[7]，前述の CS-MRI の問題を解決できる可能性があることから，大きな注目を集めている．

一方で，MRI では，装置の静磁場が厳密には均一とならないこと，撮像対象の組織間で磁化率が異なることが要因となり，励起された水素原子核のスピンには位相差が生じる．そのため，実際の MR 画像は実関数とはならず，位相分布を持った複素画像となる．複素画像の CS-MRI としては，Lustig が，アンダーサンプリングで得られた画像から位相歪みを求め，位相を補正しながら画像再構成を行う手法を提案している[3]．しかし，空間的な位相変化を含む場合，画像のスピン密度分布と位相分布の 2 つの量を求める必要があり，問題はより困難なものとなる．上記の問

*1 宇都宮大学大学院 工学研究科
システム創成工学専攻
[〒321-8585 栃木県宇都宮市陽東 7-1-2]
e-mail: ultuchiboy@gmail.com

題から，単純な実関数画像と比較して，複素画像では再構成像の品質が低下する．また，位相歪みに対応した CNN-CS も提案されているが, Lee らの手法では，スピン密度分布（強度画像）と位相分布画像を CNN で別個に推定する必要があり[8]，Yang らの手法では，CNN の複雑性から，学習に時間を要する問題がある[9]．これらの背景により，MR 画像の CS 再構成法では，位相変化のない実関数画像を想定するものや，緩やかな位相変化を推定する方法が多く，臨床機で発生しうる複雑な位相変化を持つ画像の再構成は，報告が少ない．

本研究では，複素画像の実部像と虚部像をそれぞれ実関数化できるような信号収集パターン[10]を採用し，それぞれの像を単一の CNN で独立に再構成する新たな再構成法を提案する．本手法によれば，複素画像の実部と虚部を独立に再構成できるため，位相分布を求めるための計測や推定が不要になる．その結果として，様々な位相歪みに対して頑健な再構成が期待できる．また，CNN も単一で済み，再構成系と学習の過程が簡略化される．さらに，実部と虚部の学習を行えることから学習枚数は 2 倍に増やすことができる．本研究では，提案する新たな再構成法について他の方法と比較を行い，提案法の特徴を明らかにすることを目的とする．

2. 位相歪みに頑健な CNN-CS

2.1 複素 MR 画像の実関数化処理

MR 信号$\mathrm{s}(\mathbf{k})$は，緩和項を無視すると，式(1)で表される．

$$\begin{aligned}\mathrm{s}(\mathbf{k}) &= \int \rho(\mathbf{x})e^{-j\phi(\mathbf{x})}e^{-j(\mathbf{k}\cdot\mathbf{x})}d\mathbf{x} \\ &= \mathrm{F}\left[\rho(\mathbf{x})e^{-j\phi(\mathbf{x})}\right]\end{aligned} \tag{1}$$

式(1)中の関数および変数は次の通りとする．

- $\mathbf{k}$　k 空間ベクトル
- $\mathbf{x}$　空間ベクトル
- $\rho(\mathbf{x})$　原子核のスピン密度分布
- $\phi(\mathbf{x})$　位相分布
- F　フーリエ変換

$$\begin{aligned}&\mathrm{F}\{\mathrm{Re}[\rho(\mathbf{x})e^{-j\phi(\mathbf{x})}]\} \\ &= {}^{1}\!/_{2}\,\mathrm{F}\{\rho(\mathbf{x})e^{-j\phi(\mathbf{x})} + \rho(\boldsymbol{x})e^{j\phi(\mathbf{x})}\} \\ &= {}^{1}\!/_{2}\,\{\mathrm{s}(\mathbf{k}) + \mathrm{s}(-\mathbf{k})^{*}\}\end{aligned} \tag{2}$$

$$\begin{aligned}&\mathrm{F}\{\mathrm{Im}[\rho(\mathbf{x})e^{-j\phi(\mathbf{x})}]\} \\ &= -{}^{j}\!/_{2}\,\mathrm{F}\{\rho(\mathbf{x})e^{-j\phi(\mathbf{x})} - \rho(\boldsymbol{x})e^{j\phi(\mathbf{x})}\} \\ &= -{}^{j}\!/_{2}\,\{\mathrm{s}(\mathbf{k}) - \mathrm{s}(-\mathbf{k})^{*}\}\end{aligned} \tag{3}$$

$$\begin{aligned}\mathrm{s}(-\mathbf{k})^{*} &= \int \rho(\mathbf{x})e^{j\phi(\mathbf{x})}e^{-j(\mathbf{k}\cdot\mathbf{x})}d\mathbf{x} \\ &= \mathrm{F}\left[\rho(\mathbf{x})e^{j\phi(\mathbf{x})}\right]\end{aligned} \tag{4}$$

このとき, 信号$\mathrm{s}(\mathbf{k})$のフーリエ変換によって得られる複素画像$\rho(\mathbf{x})e^{-j\phi(\mathbf{x})}$に対して，式(2)，(3)による実関数化を行うことにより，複素画像から実部像と虚部像を生成する．式(2)，(3)では，式(4)の関係を利用する．式(4)は，MR 画像$\rho(\mathbf{x})e^{-j\phi(\mathbf{x})}$の複素共役画像である$\rho(\boldsymbol{x})e^{j\phi(\mathbf{x})}$の k 空間は，元画像$\rho(\mathbf{x})e^{-j\phi(\mathbf{x})}$の k 空間の複素共役かつ，原点対称であることを示している．

2.2 実関数化処理の CS 再構成への応用

式(2)，(3)による実関数化処理では, k 空間において原点対称となる要素同士の加減算を伴う．実関数化処理の CS への応用にあたっては，間引き信号が原点対称の関係を満たす必要がある．本研究では図 1 のように, k 空間が原点対象となるカルテシアン間引きを使用する．

2.3 実関数化処理を伴う CNN 再構成

上記の手順により，目標画像（フルデータ像）と間引き再生像（ゼロフィル再構成像）の，実部

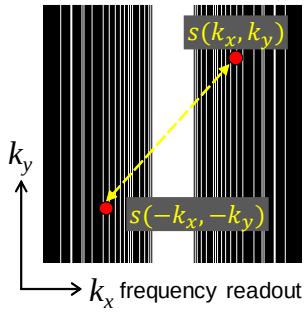


図 1. 原点対称を満たすサンプリングパターン

像と虚部像をそれぞれ生成する．ゼロフィル再構成像と目標画像の関係を CNN で学習することで，画像空間上で再構成を行う CNN を構築する．この際，本手法で生成された実部像と虚部像に対して，それぞれ独立に CNN を構築して学習する必要はなく，入力が 1 チャンネルの単一の CNN で，実部像と虚部像を混合したデータセットでの学習が可能である．

学習済みの CNN による再構成は，実部像と虚部像で独立して行い，得られた再構成像の複素画像化を行うことで完了する（図 2）．本手法による再構成は，以下の利点を有する．

- 実部・虚部像を混合したデータセットで学習するため，学習枚数は実質的に倍増する．
- 位相歪みの推定・補正を伴わないため，複雑な位相歪みに対して頑健である．
- 再構成像の高品質化に寄与する，実関数の拘束条件が適用できる．

2.4 Data Consistency

本研究では，CNN 再構成像の k 空間（図 2 の③と④の画像のフーリエ変換空間）の対応箇所を，図 2 の①と②の収集信号が存在する点で更新する Data Consistency を行って，再構成像の高品質化を図る．Data Consistency は，実関数化した画像の k 空間を用いて，実部と虚部の再構成像を複素数化する直前に行う．

2.5 CNN アーキテクチャ

本研究では，Han らが提案した Dual Frame U-Net[11]をベースとした図 3 の CNN を，位相歪みに頑健な再構成に使用した．入力層および中間層は，畳み込み(Conv)，Batch Normalization(BN)，ReLU で構成されており，Pooling と Unpooling により特徴マップの拡大と縮小を行う．出力層は Conv のみで構成される．エンコーダ側とデコーダ側の対応する層の間には，エンコーダの出力をデコーダに連結するためのスキップ接続（図 3 中の黄緑の矢印）がある．Dual Frame U-Net では，上記のスキップ接続に加えて図 3 中の紫の矢印で示したスキップ接続を設けることで，生成画像の高品質化を図っている．

3. 再構成シミュレーションの条件と結果

3.1 学習およびテスト用画像

CNN の学習には，キヤノンメディカルシステムズ社製の Vantage Titan 3T で撮像された T1 強調像を計 25 枚使用した．

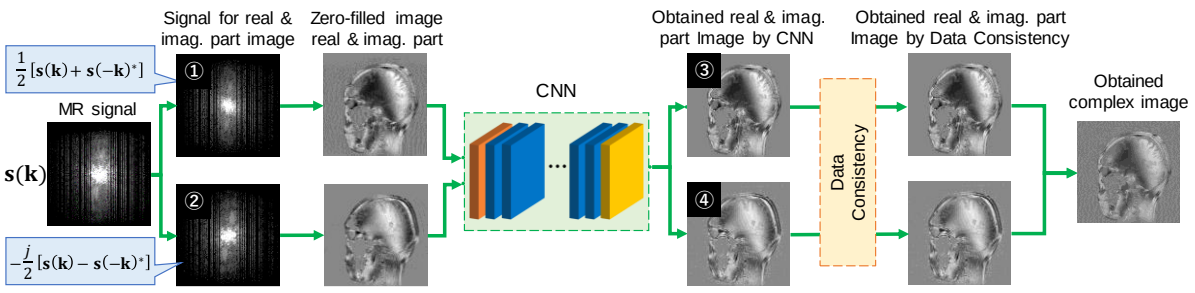


図 2. 位相歪みに頑健な CNN-CS による再構成処理の手順

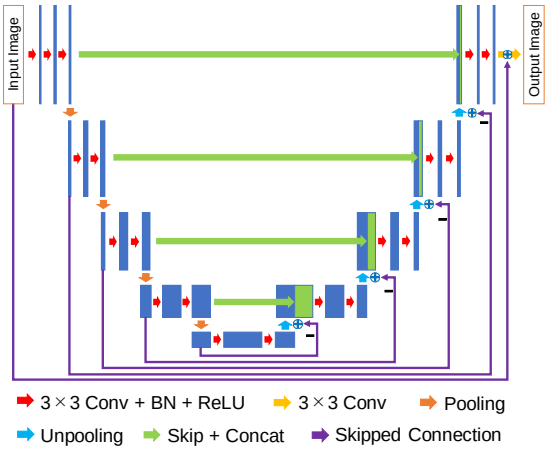


図 3. Dual Frame U-Net のアーキテクチャ

3.2 CNN の学習

25 枚の学習用画像から，実部と虚部のフルデータ像とゼロフィル再構成像をそれぞれ作成し，これらのペアを計 50 組用意した．この 50 組のペアを 1 つのデータセットとして，単一の Dual Frame U-Net に与え，ゼロフィル再構成像とフルデータ像の関係を学習させた．学習時の適用条件と，計算機環境をそれぞれ表 1，表 2 に示す．CNN の学習には MATLAB と MatConvNet[12]を使用した．200 エポックの学習には約 1 時間を要した．

表 1. CNN の学習条件

Optimizer	Adam
Learning Rate	$1 \times 10^{-3} \sim 1 \times 10^{-5}$
Loss function	画像空間の MSE
Epoch	200

表 2. 計算機環境

CPU	Intel Core i7-9700K
RAM	DDR4-2666 64GB
GPU	NVIDIA GeForce RTX 2080ti
Software	MATLAB R2017b CUDA toolkit 10.2 cuDnn 7.6.5 MatConvNet 1.0-beta25

3.3 比較手法

提案する位相歪みに頑健な再構成法との比較のため，Zhao らが提案した位相歪みを考慮する CS-MRI[13], Lee らが提案した強度画像と位相分布をそれぞれ独立の CNN で学習する CNN-CS（以降，Mag&Phase とする）による再構成法[8]および, Yang らが提案した ADMM-CSNet[9]による再構成を行った．Mag&Phase では，Lee らの学習方法を参考に，以下の 2 ステップの学習を行った．

1. 強度画像用の CNN に対する 2500 枚の実関数 T1 強調像を用いた学習(Pre-train)
2. 強度画像用と位相分布用の CNN に対して 25 枚の複素画像を用いた学習（2 つの CNN をそれぞれ独立に学習）

ステップ 1 では, IXI dataset[14]を使用して 200 エポックの学習を，ステップ 2 では 3.1 節の画像を使用して 100 エポックの学習を行った. また，ADMM-CSNet は，CS-MRI の一種である ADMM[15]の処理を CNN で実装した手法である．本研究では，ADMM-CSNet で 25 枚の複素画像を用いて 300 エポックの学習を行った．

再構成像の品質評価には PSNR(Peak Signal-to-Noise Ratio)を使用した．

3.4 サンプリングパターン

本研究では，提案法及び比較手法による再構成に，信号収集量が 20, 30, 40%かつ，原点対称となるサンプリングパターンを使用した．

3.5 再構成シミュレーション結果

各種の再構成法によってテスト画像の再構成を行った．得られた再構成像の PSNR を図 4 に示す．また，収集信号量 30%における再構成像を図 5 に，再構成像から生成した位相分布図を図 6 に示す．図 5 において，(a)から(f)は再構成像の強度画像を，その右下には，(a)中に枠で示した箇所

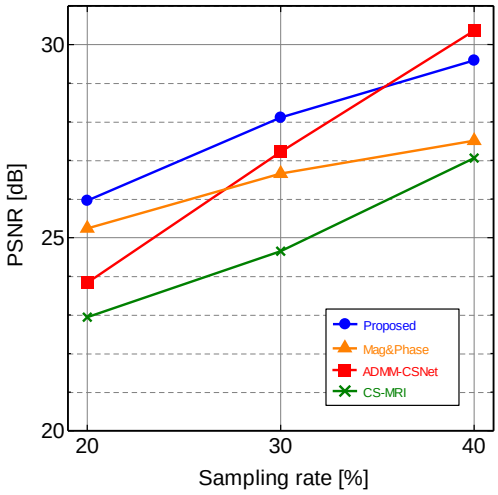


図 4. 再構成像の PSNR の平均値

を拡大した画像を示している．また，(g)から(j)は，それぞれ(b)から(e)の再構成像とフルデータ像の誤差画像であり，下部のカラーバーに従って着色している．図 6 の(a)から(f)の右下には，(a)中に枠で示した箇所を拡大した画像を示している．

図 4 より，提案法は他手法と比較して，低信号量において高い PSNR を示した．一方で，ADMM-CSNet は信号量の増加に比例して画像品質が向上し，40%において最も良好な性能を示した．

図 5 に示した再構成像の強度画像において，(f)のゼロフィル再構成像と(b)から(e)の再構成像を比較すると，(d)の CS-MRI ではアーチファクトの残留が目立ったが，その他の手法では良好に除去できた．また，(c)の Mag&Phase では，被写体の構造は良好に復元されたものの，再構成像が平滑化される傾向にあり，(e)の ADMM-CSNet では，小脳付近をはじめとした細部の構造を失う傾向にあった．一方の提案法では，アーチファクトの除去性能，画像の鮮鋭度において他手法に比べて優れており，拡大画像で示したように，細部の構造も良好に復元できた．

図 6 に示した再構成像の位相分布において，(b)のゼロフィル再構成像では，注目箇所のように位相分布の乱れが発生する場合があった．各種の再構成法を適用した結果，(c)の提案法と(f)の ADMM-CSNet では位相分布の乱れを除去できたが，(d)の Mag&Phase では乱れが発生した箇所を含む，位相の変化が激しい領域の推定が良好に行えず，(e)の CS-MRI では，乱れが残った．また，(a)の拡大箇所では，周囲の組織よりも位相が僅かに異なる領域が確認できる．この箇所は(b)より，間引きによって消失しているが，(c)の提案法では，他手法と比較して最も良好に復元が行えた．

位相分布データの応用例として温度分布の測定画像の再構成を試みた．MRI を利用し，位相データから温度分布の計測が可能である．本検討では Zhao らが作成した，計 4 つのガウス分布状のホットスポットが設けられている腹部の T2 強調像[8]に，信号量が 40%となるサンプリングパターンを適用し，CNN で再構成を行った．結果を図 7 に示す．図 7 において，(a)から(c)は強度画像を，(d)から(f)はそれぞれ(a)から(c)の位相分布図を示している．提案法によってフルデータを利用した場合に近い温度計測を行うことができた．

4. 考察

図 6(d)より Mag&Phase では，位相が急激に変化する領域で，正しい推定が行えなかった．Mag&Phase では，強度画像と位相分布のそれぞれを CNN で学習して再構成を行う．このうち，強度画像に関しては位相分布の影響を受けないため，良好な再構成が行えたが，一方の位相分布では，変化が緩やかな領域は良好な推定が行えたものの，位相は不連続関数であるため位相変化が急激かつ大きい領域に関しては，推定が困難となる場合があり，乱れが発生した．また CS-MRI と ADMM-CSNet では，再構成像にアーチファクトが残っていたり，細部の構造が消失したりする傾向があったため，再構成像から生成し

た位相分布にも乱れが残留したり，小さな位相差の復元が行えない場合があった．一方の提案法では，位相分布を用いずに再構成が可能なため，急激な位相の変化を伴う画像でも，その影響を受けずに再構成が行えた．また，実関数の拘束条件により，高品質な再構成を行えた点が，小さな位相差の復元に寄与したと考えられる．この結果から，提案法は強度画像に加え，位相分布の推定にも優れており，磁化率強調画像やQSM(Quantitative Parameter Mapping)のような，画像の生成に正確な位相分布が求められる手法に対しても応用可能であることが示唆された．一方で，図 7 に示した温度分布計測画像に関しては，同様の画像が収集できなかったため類似画像を学習モデルに含めることができる，3.1 節の画像で学習済みの CNN モデルを使用して再構成を行った．学習用画像と図 7 の画像では，特徴が大きく異なっているため，再構成像には僅かにアーチファクトが残留する結果となった．汎化性能の向上のためには，様々な特徴を持つ画像を学習に取り込むことが必要である．

5. まとめ

本研究では，空間的に位相分布を持ち複素数となる MR 画像の実部と虚部を，それぞれ実関数化できる信号間引き法の採用により，位相分布推定が不要かつ，実部と虚部を CNN で独立に再構成できる新たな手法を提案した．シミュレーションの結果，提案法では複雑な位相分布に対しても頑健かつ高品質な再構成を行うことができた．

今後は，学習枚数の増加，様々な位相分布を持つ画像を用いた再構成シミュレーションや他手法との特徴の比較および，QSM などへの応用について検討する予定である．

謝 辞

本研究の一部は，科学研究費補助金(19K04423)および柏森情報科学振興財団の助成により実施された．また，本研究に使用した IXI dataset を提供する Information eXtraction from Images プロジェクトおよび，MR 画像をご提供いただいた，キヤノンメディカルシステムズに感謝の意を表します．

利益相反の有無

なし

文 献

[1] Donoho DL: Compressed sensing. IEEE Transactions on Information Theory **52** : 1289-1306, 2006

[2] Candès EJ, Wakin MB: An Introduction To Compressive Sampling. IEEE Signal Processing Magazine **25** : 21-30, 2008

[3] Lustig M, Donoho DL, Pauly JM: Sparse MRI: The application of compressed sensing for rapid MR imaging. Magnetic Resonance in Medicine **58** : 1182-1195, 2007

[4] Wang S, Su Z, Ying L, et al. : Accelerating magnetic resonance imaging via deep learning. IEEE 13th International Symposium on Biomedical Imaging, Prague, Czech Republic 2016, 514–517

[5] Hyun CM, Kim HP, Lee SM, et al. : Deep learning for undersampled MRI reconstruction. Physics in Medicine and Biology 2018, 63

[6] Guang Y, Yu S, Dong H: DAGAN: Deep De-Aliasing Generative Adversarial Networks for Fast Compressed Sensing MRI Reconstruction. IEEE Transactions on Medical Imaging **37** : 1310-1321, 2018

[7] Zhu B, Liu JZ, Cauley SF, et al. : Image reconstruction by domain-transform manifold learning. Nature **555** : 487-492, 2018

[8] Lee D, Yoo J, Tak S, et al. : Deep Residual Learning for Accelerated MRI Using Magnitude and Phase Networks. IEEE Transaction on Biomedical Engineering 6**5** : 1985-1995, 2018

[9] Yang H, Sun J, Li H, et al. : ADMM-CSNet: A Deep Learning Approach for Image Compressive Sensing. IEEE Transactions on Pattern Analysis and Machine Intelligence **42** : 521-538, 2020

[10] Ito S, Yamada Y : Compressed Sensing for Magnetic Resonance Images with Phase Variations. IEEE International Conference on Audio, Speech, and Signal Processing, Florence, Italy 2014, BISP-P7.1, 6648-6651

[11] Han Y, Ye JC : Framing U-Net via Deep Convolutional Framelets : Application to Sparse-view CT. IEEE Transactions on Medical Imaging **37** : 1418-1429, 2018

[12] Vedaldi A, Lenc K : MatConvNet: Convolutional Neural Networks for MATLAB. Proceedings of the 23rd ACM international conference on Multimedia, Brisbane, Australia 2015, 689-692

[13] Zhao F, Noll DC, Nielsen JF, et al. : Separate Magnitude and Phase Regularization via Compressed Sensing. IEEE Transactions on Medical Imaging **31** : 1713 - 1723, 2012

[14] IXI dataset. https://brain-development.org/ixi-dataset/

[15] Liu Y, Cai JF, Zhan Z, et al. : Balanced sparse model for tight frames in compressed sensing magnetic resonance imaging. PLoS ONE 2015, 10:e0119584.

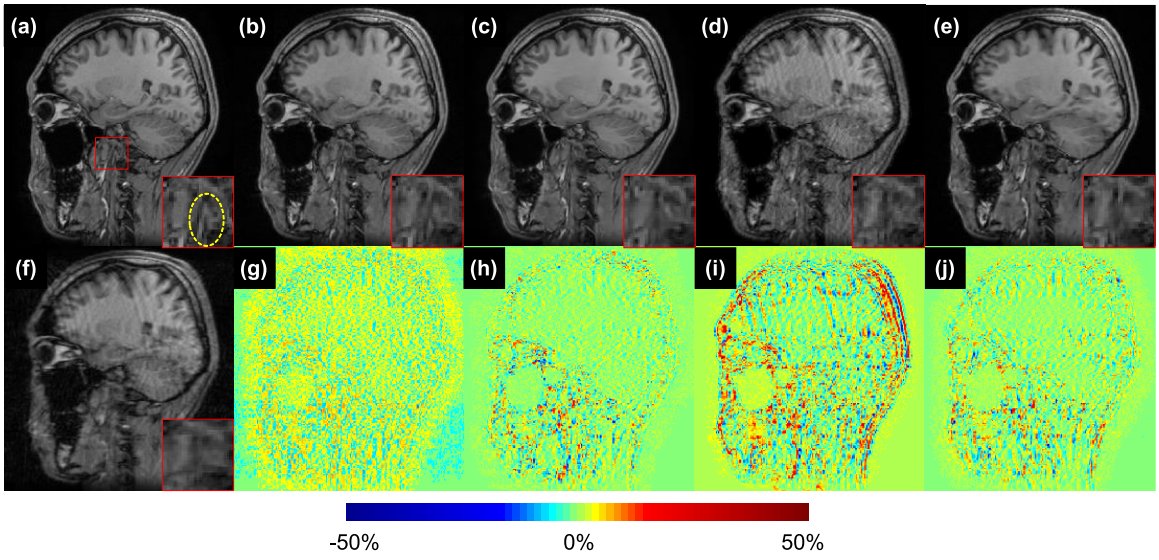


図 5. 複素 MR 画像の再構成結果，(a) フルデータ像，(b) 提案法による再構成像，(c) Mag&Phase による再構成像，(d) CS-MRI による再構成像，(e) ADMM-CSNet による再構成像，(f) ゼロフィル再構成像，(g) - (j) 誤差画像，(a) - (f) の右下の画像:赤枠領域の拡大画像

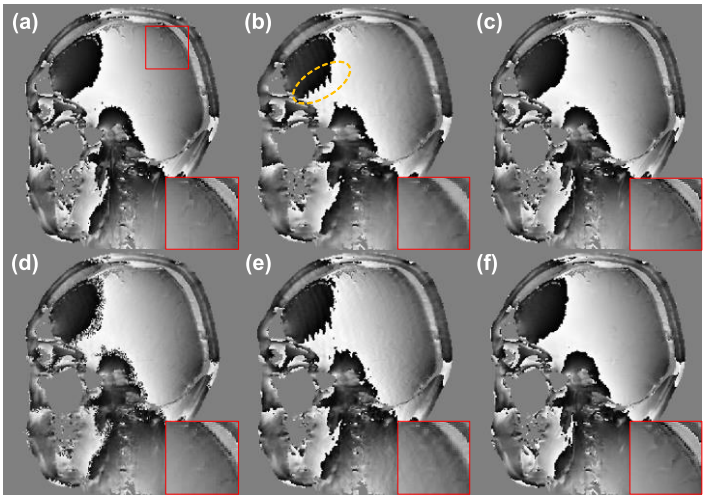


図 6. 再構成像の位相分布比較，(a) フルデータ像，(b) ゼロフィル再構成像，(c) 提案法による再構成像，(d) Mag&Phase による再構成像，(e) CS-MRI による再構成像，(f) ADMM-CSNet による再構成像

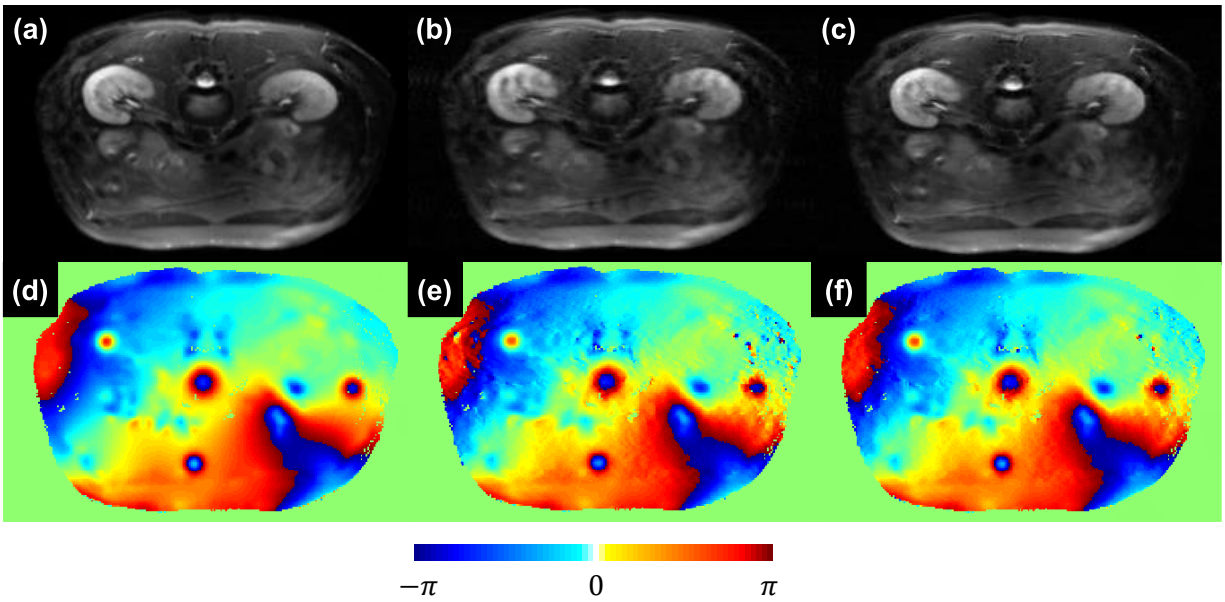


図 7. 位相分布の温度分布計測への応用，(a) フルデータ像，(b) ゼロフィル再構成像，(c) 提案法による再構成像，(d)，(e)，(f)は(a)，(b),，(c)の各位相分布（カラーバーに従って着色）

Deep Learning based MR Image Reconstruction Robust to Spatial Phase Variation

Shohei OUCHI[*1], Satoshi ITO[*1]

*1 Graduate School of Engineering, Utsunomiya University

MR images have phase variation due to inhomogeneities of static field strength and the difference of magnetic susceptibility, therefore image reconstruction must consider phase on the image as well as its magnitude. Even many Deep Learning based reconstruction methods are proposed, most of them are assumed real value images or relaxed phase variation. In this study, we proposed Deep Learning based reconstruction robust to phase variation. According to our method, real and imaginary part image of complex MR image is reconstructed respectively by applying symmetrical under-sampling pattern. The simulation experiments showed that our method allows high-quality reconstruction compared to other CS iterative method and Deep Learning based method corresponding to phase variation.

Key words: Deep Learning, Compressed Sensing, Phase variation

少数投影方向マイクロ CT 像の補正手法の開発

岡本　尚之[*1]　大西　峻[*2]　羽石　秀昭[*2]

要旨

数ミクロンの空間分解能を有するマイクロ CT は, 三次元ボリュームの非破壊イメージングが可能であり, 病理診断への導入の有効性が示唆されている．しかし，高画質なマイクロ CT 像を得るためには，長時間撮影を要する．本研究は，病理診断応用に向けたマイクロ CT 像の画質補正手法の構築を目的とする．本稿では，投影方向数の削減に基づくマイクロ CT の高速撮影を検討する．提案手法では，少数投影方向とフルサンプリングのサイノグラム像をシミュレーションにより生成し，後者を教師とする Fully Convolutional Network を構築する．実験では，公開データによる有効性評価後，ラット臓器に対する転移学習を実施した．結果，補正前の CT 像に生じていたスポーク状のアーチファクトが低減し，数値評価において PSNR，SSIM 共に大幅な改善を示した．

キーワード：マイクロ CT，病理診断，画質補正，サイノグラム，Fully Convolutional Network

1. はじめに

腫瘍の確定診断は，摘出・薄切した病理標本を顕微鏡下で観察することが一般的である．しかし，病理標本の作成工程は煩雑であり，加えて対象標本の断面のみが観察可能であることから, 検査効率の向上が求められている．数ミクロンの空間分解能を有するマイクロ CT は，三次元ボリュームの非破壊イメージングが可能であり，病理検査の問題点を克服できる可能性が示唆されている．しかし，生体を対象とした撮影の場合，高画質を担保するためには撮影に長時間を要するといった課題がある［1］．本発表では，短時間撮影実現に向けた少数投影方向数の CT 像に対する補正手法を提案する．

*1 千葉大学大学院融合理工学府基幹工学専攻医工学コース
〔〒263-8522 千葉市稲毛区弥生町 1-33〕
e-mail: t_okamoto@chiba-u.jp
*2 千葉大学フロンティア医工学センター

2. 手法

少数投影方向の CT 像(LR：low resolution)からフルサンプリングの CT 像(HR: high resolution)を推定した補正像(SR: super resolution)を取得する．**図 1(a)**に提案手法の概略図を示す．

サイノグラムは，各投影位置から得られる投影データを配置したものである．検出器サイズを固定した場合，投影方向数が増加するごとにサイノグラムの画像サイズは行方向の

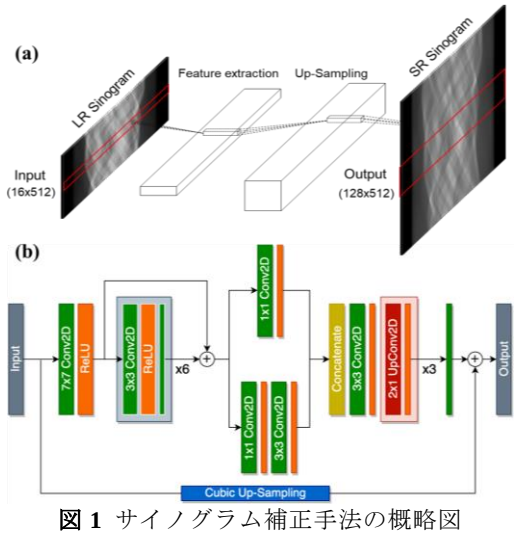


図 1 サイノグラム補正手法の概略図

みに増大する．よって提案手法では，サイノグラムの行方向の拡張を考慮したネットワーク(**図 1(b)**)を構築する．学習では，サイノグラム像を行方向のみに限定したパッチ画像に分割し, LR と HR をペアとする教師あり学習を実施する．

3. 実験・結果

二種の実験により手法の有効性を評価した．

A) 公開データ(クルミのマイクロ CT 像［2］)による学習・適用

B) A より取得した学習済みモデルによるラット臓器(肝臓・腎臓)への適用

使用したサイノグラムはシミュレーションにより生成した．LR は投影数 128，HR は投影数 1024 であり，本手法により，8 倍の投影数削減を目指す．

サイノグラムの推定結果を**図 2** に示す．推定結果 SR は LR と比較して線構造が連結し，微細な内部構造も復元されている．

次に，パッチ化されたサイノグラム像を連結し，フィルタ補正逆投影法を適用した結果を**図 3** に示す．**図 3(a)**より，スポーク状のアーチファクトが除去され，LR では不鮮明である外殻構造のエッジが SR では保存された．**図 3(b)**においても，アーチファクト除去の有効性は確認できるが，一方で臓器の微細な内部構造が平坦化された．これは，クルミと軟部組織のサイノグラムの特性の違いによるものであり，学習データの変更により解消可能であると考える．

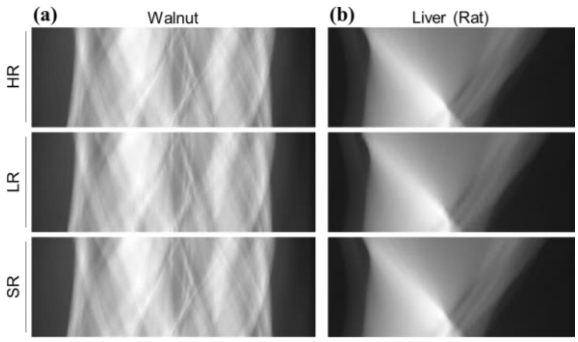


図 2 サイノグラム推定結果 (掲載の都合上，左右に存在する画素値 0 の領域をトリミング済)，LR は SR との比較のために行方向の Cubic 補間を実施．

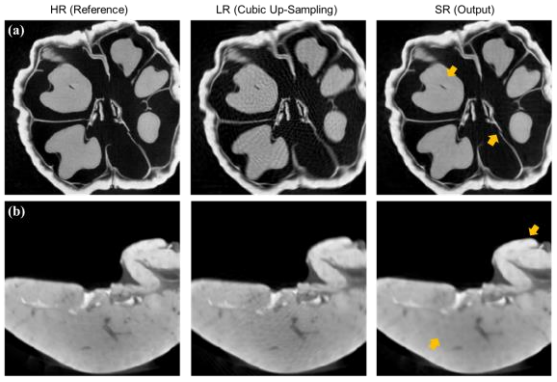


図 3 再構成結果. LR は Cubic 補間によるサイノグラムの伸長を実施後，再構成像を取得．

また，**表 1** に示す数値評価では，両実験において大幅な改善が確認された．特に学習データとして利用したクルミに対する効果が顕著であり, PSNR では 10 [dB]の向上を示した．

表 1 各データにおける数値評価結果

	Walnut		Kidney		Liver	
	LR	SR	LR	SR	LR	SR
PSNR	25.15	**35.41**	31.90	**39.00**	32.12	**44.38**
SSIM	0.92	**0.98**	0.97	**0.99**	0.98	**0.99**

4. まとめ

サイノグラム像に対する画質補正手法を構築した．結果，CT 像上においてアーチファクトの低減を確認し，数値評価において大幅な改善を示した．

謝辞

本研究の一部は，文部科学省科学研究費補助金基盤研究(A)(課題番号 : 19H01172)により行われた．

利益相反の有無

なし

文　献

［1］ Teplov A, Tabata K, Fu X, et al.: Development of Standard Operating Procedure (SOP) of Micro-computed tomography (micro-CT) in Pathology. Diag Pathology **5(1)**, 2019

［2］ Der Sarkissian H, Lucka F, van Eijnatten, M et al.: A cone-beam X-ray computed tomography data collection designed for machine learning. Sci Data **6**: 215 2019

Development of denoising method for sparse-view micro-CT images

Takayuki OKAMOTO[*1], Takashi OHNISHI[*2], and Hideaki HANEISHI[*2]

*1 Graduate School of Science and Engineering, Chiba University

*2 Center for Frontier Medical Engineering, Chiba University

Micro-computed tomography (Micro-CT) enables us to create non-destructive imaging of three dimensional volumes with microscopic level resolution. Therefore, it has been suggested that this new modality is effective for pathological diagnosis. However, it requires a long scanning duration to acquire high quality images. The purpose of this study is to develop an image quality correction method of micro-CT images for pathological diagnosis. In this report, we described a denoising method for sparse-view micro-CT imaging. We constructed a fully convolutional network using sparse-view and full-sampling sinogram images. Experiments were conducted transfer learning on rat organs after evaluation of efficacy using public data. The results showed that the spoke-like streak artifacts were suppressed and both PSNR and SSIM metrics showed significant improvement.

Key words: Micro-CT, Pathological diagnosis, Image quality correction, Sinogram, Fully Convolutional Network

画像処理法を用いた冠動脈 CTA 石灰化プラーク除去法の開発と臨床評価

渋谷　幸喜[*1]　湯浅　哲也[*2]　藤原　敏弥[*3]　柴原　徹[*3]

工藤　和也[*4]　東海林　綾[*4]　久根崎　遥加[*2]　大村　知己[*5]

要旨

冠動脈に石灰化が付着している症例の冠動脈 CT 画像は，石灰化が原因で発生する高輝度陰影と blooming artifact が原因で診断不可能となる症例が多数存在するため問題となっている．この問題を解決するため，中間強調再構成関数と中心係数 9，周辺係数-1 に設定したラプラシアンフィルターを適用し 2 段階鮮鋭化を行った後，zero-cross 法を用い石灰化プラークの境界を抽出し石灰化除去を行う image processing algorithm for removing calcified plaque (iPARC)法を開発し臨床評価を行った．臨床評価は，画像にブレが無くステントが挿入されていない石灰化付着 74 本の冠動脈を対象とし，経験年数 29 年と 23 年の循環器内科医師 2 名が行った．iPARC 法を用い石灰化除去を行うことにより，陽性予測的中率は 12.2%から 77.0%へ大幅に向上した．

キーワード：冠動脈 CT，石灰化除去，画像処理技術

1.　はじめに

Japan circulation society (JSC : 日本循環器学会)の本邦における循環器疾患の診療実態調査の推移では，coronary angiography (CAG : 冠動脈造影検査) の検査総数は 2014 年度に 501,665 件，2018 年度に 498,597 万件と漸減であったのに対し，coronary computed tomography angiography (CCTA : 冠動脈 CTA) は 2014 年度の 413,495 件から 2018 年度には 481,150 件へと大幅に増加している[1].

*1 北海道科学大学診療放射線学科〒006-8585 札幌市手稲区前田 7 条 15 丁目 4-1l: shibuya-k@hus.ac.jp，*2 山形大学大学院工学部応用生命システム工学科，*3 市立秋田総合病院循環器内科，*4 市立秋田総合病院放射線科，*5 秋田県立循環器・脳脊髄センター放射線科診療部

CCTA 件数増加の理由は，血管内腔と石灰化プラークの両者が評価可能である[2,3]，検査費用が安価である (CCTA 約 50,000 円、CAG 約 300,000 円) ，静脈注射だけの低侵襲性検査のため被検者の心身の負担が少ない，通院で行うことが可能なため、時間および介護の負担が少ない，被曝低減技術の進歩により従来と比較し大幅な X 線量の低減が可能等において優れているため，現在，冠動脈疾患あるいはそれを疑われる患者の管理において不可欠な検査となっている[4-10].

2.　本研究の目的

CCTA 画像において，冠動脈に石灰化プラークが付着している症例では，石灰化プラークの主成分であるカルシウムの X 線不透過性に起因する blooming artifact[11]が発生し，このartifact にともなう辺縁不鮮明高輝度拡大陰影が CCTA による血管内腔の評価・診断を妨げて

いる[12-17]．2009 年に Society of Cardiovascular Computed Tomography (SCCT)によって策定された Guidelines for Performance of Coronary Computed Tomographic Angiography[18] では Agatston スコアが 600 以上の場合 CCTA を用いて血管内腔狭窄の評価・診断はすべきではないと勧告が出された．翌 2010 年に出された Appropriate Use for Cardiac Computed Tomography[16]では，Agatston スコアが 400 以上の場合に CCTA による狭窄評価は困難と勧告が出された．しかし，Agatston スコアが 400 以下の症例でも，maximum intensity projection (MIP : 最大値投影法) と straight view と cross section view の複数の画像を観察しても読影不可能な症例が多数存在するため，石灰化プラーク付着症例の CCTA 適応基準は確立されていない[19].

本研究の目的は，CCTA 画像から画像処理技術を用い石灰化プラークを高精度に除去し，冠動脈内腔定量評価可能な画像再構築法の開発し，冠動脈狭窄病変の詳細把握においてもっともすぐれた検査法である CAG[20]と比較評価を行い，臨床有効可能性を示すことである．

3. 方法

3. 1 CCTA 撮影条件と Agatston スコア計測法

CT 装置は 320 列 MSCT Aquilion ONE (キャノンメディカル) を用いた．CCTA 撮影条件は，0.5×320 列 Non Helical Volume Scan，心臓用 Auto Exposure Control: 0.5mm 厚 Real Prep 撮影時 Standard Division 30，管電圧 125kV，管電流 300-550mA, スキャン時間 0.35sec/rot, Scan Field of View (FOV) 320mm に設定しスキャンを行った．画像再構成条件は，Adaptive Iterative Dose Reduction 3D (AIDR3D : 逐次近似応用再構成法)Blending 比率 50%に設定し，Display FOV 150~200mm, 512×512matrix, スライス厚 0.5mm, 間隔 0.25mm に設定し再構成を行った. Agatston スコアは文献[16]に従い，スライス厚 3.0mm, 間隔 3.0mm で計測を行った.

3. 2 画像処理法を用いた CCTA 石灰化プラーク除去法の開発

骨などの X 線不透過物質が原因で Blooming Artifact が発生すると，隣接臓器の CT 値が変動することが知られている[21]. 人体臓器の CT 値はおおよそ-1,000～60HU の範囲に, 骨と石灰化プラークの CT 値はおおよそ 130～3,000HU の範囲に, CCTA の冠動脈の CT 値はおおよそ 300～500HU の範囲に分布する．そのため，CCTA 画像での最大 CT 値差は，4,000HU になる場合があり, この CT 値差が原因で Blooming Artifact が発生する．Blooming Artifact は，辺縁不鮮明高輝度拡大陰影として描出され, Fig 1 に示すようにウインドー幅とウインドーレベルの設定値の違いにより冠動脈内腔の見え方は大きく異なるため，診断・評価の妨げになっている．

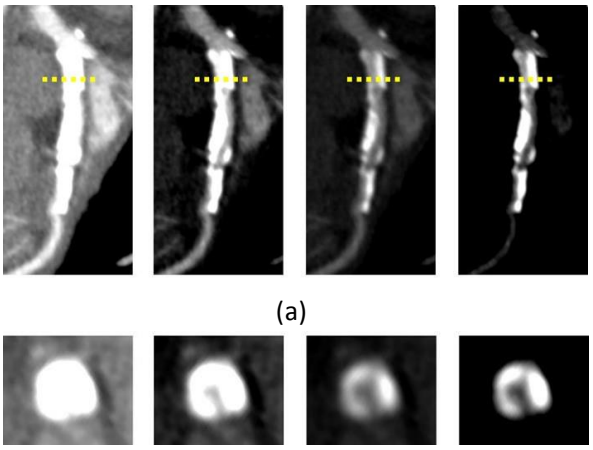


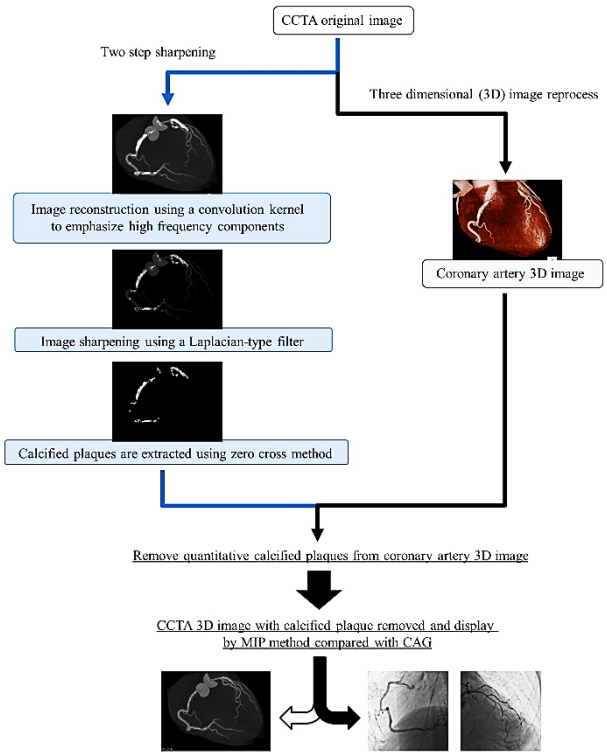
Fig.1 Difference in appearance of calcification and vessel lumen caused by difference between set window width and window level
(a) curved planar reprocessing image,
(b) axial image of yellow line

この問題を解決するために，2 段階鮮鋭化と 2 次微分フィルターを用い石灰化プラーク除去のための image processing algorithm for removing calcified plaque (iPARC) 法を開発した. iPARC 法フローチャートを Fig 2 に示す．iPARC 法は，はじめに，画像再構成過程の filter back projection 法再構成関数を中間強調に設定し畳み込み演算処理を行った．次に，式(1)を用い中心係数 9，周辺係数-1 に設定したラプラシアンタイプフィルターを適用し石灰化プラークの

鮮鋭化を行った．最後に，zero cross 法を用い石灰化プラークと造影冠動脈と周囲臓器の境界を検出し石灰化プラーク抽出を行った後，volume data から抽出した石灰化プラークを差分し冠動脈のみを MIP 表示する方法である．

$$g(x,y) = Lf(x,y) - \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) f(x,y) \qquad (1)$$

Fig.2 iPARC method image processing flowchart



and clinical image compared with CAG

3．3　対象患者

対象患者は，2012 年 11 月から 2013 年 7 月まで日本海総合病院で CCTA を行った 444 症例のうち，後日 CAG を行った石灰化プラーク付着 48 症例の右冠動脈 29 本，左冠動脈 45 本を対象とした．男性 36 症例 (平均年齢 70.6 歳，最高齢 83 歳，最低齢 55 歳)，女性 12 症例 (平均年齢 75.3 歳，最高齢 89 歳，最低齢 58 歳) であった．Agatston スコアは，右冠動脈平均 400.44 (最高 2025.00，最低 0.0) ，左冠動脈平均 579.30 (最高 3100.00，最低 27.06)，全冠動脈平均 971.9(最高 3507.4，最低 27.1)であった．CCTA 実施後，1 ヵ月以内に CAG を実施のは 30 症例，2 ヵ月以内に CAG を実施のは 14 症例，3 ヵ月以内に CAG を実施のは 2 症例，3 ヵ月以後に CAG を実施のは 2 症例であった．

4.　臨床評価方法

4．1　CAG を模範解答とした iPARC 法石灰化プラーク除去画像臨床有効性

CCTA における経皮的冠動脈形成術等の血行再建術の実施判断は，75%以上の狭窄病変有無が指標となっている[14,22-24]. CCTA による狭窄率評価は臨床現場で日常的に行われているが，CAG は AHA 分類の 75%を有意狭窄とする判定基準に加え，エッジディテクション法やビデオデンシトメトリー法などの定量的計測[25]が行われているため，CCTA 狭窄率確定のためには依然として CAG は必須の検査である[20]. 今回の iPARC 法石灰化プラーク除去画像臨床評価は，図 3 の方法を用い，CAG による AHA（America Heart Associ-ation）分類[26,27]に従い，CAG を模範解答とし，経験年数 29 年と 23 年の循環器内科医師 2 名が行った．はじめに，事前に CAG の確定診断情報を一切与えず，CAG 画像とほぼ同方向からの投影から作成した石灰化プラーク除去前 MIP 画像で読影を行った．次に，石灰化プラークの障害陰影が原因で読影を行えない場合は，iPARC 法石灰化除去後 MIP 画像で追加読影を行った．最後に CAG を模範解答とし，全ての読影結果を以下の (A) ～ (D) の 4 分類に分けて評価を行った．

(A) 石灰化プラーク除去後 iPARC 法 MIP 画像で狭窄ありと評価され，CAG でも同等の狭窄ありと評価した場合，iPARC 法で石灰化プラークは適切に除去されている(図 3(a))．この場合を true positive (TP)とした．

(B) 石灰化プラーク除去後 iPARC 法 MIP 画像で狭窄なしと評価され，CAG でも狭窄なしと評価した場合，iPARC 法で石灰化プラークは適切に除去されている(図 3(b))．この場合を true negative (TN)とした．

(C) 石灰化プラーク除去後 iPARC 法 MIP 画像で狭窄ありと評価され，CAG では狭窄なしと評価した場合，iPARC 法で石灰化プラークを過大除去したことにより誤診された (図 3(c))．この

場合を falsep positive (FP)とした.
(D) 石灰化プラーク除去後 iPARC 法 MIP 画像で狭窄なしと評価され, CAG では狭窄ありと評価した場合，iPARC 法で石灰化プラークを過少除去したことにより誤診された (図 3 (d)). この場合を false negative (FN)とした.

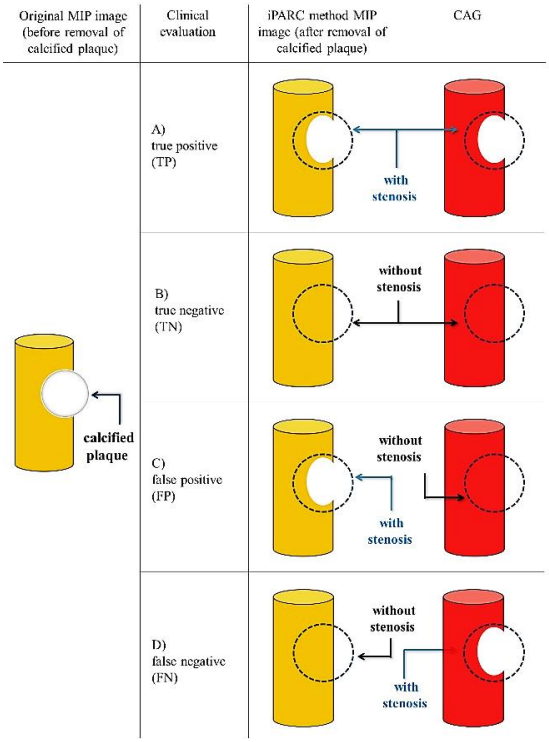


Fig.3 Calcified plaque removal evaluation parameter and schematic diagram

4．2　Agatston スコア

Agatston スコアを，0-200, 201-400, 401- 600, 601-800, 801-1000, 1001-1200, 1201‐1400, 1401-1600, 1601‐1800, 1801‐2000, 2001 以上の 11 グループに分類し評価を行った.

4．3　CAG を模範解答とした iPARC 法石灰化プラーク除去画像臨床評価の感度・特異度・陽性的中率・陰性的中率

CAG を模範解答とした iPARC 法石灰化プラーク除去画像臨床有効性(A) ～ (D)の 4 分類の評価結果を用い，Table 1 (a), (b) の式に従い，石灰化除去前 MIP 画像と iPARC 法石灰化除去後 MIP 画像の感度・特異度・陽性的中率・陰性的中率の算出を行った[28-30].

Table 1 Clinical evaluation of iPARC method compared with CAG as correct diagnosis

(a) accuracy

		CAG	
		with stenosis	without stenosis
CCTA	with stenosis	true positive	false negative
	without stenosis	false positive	true negative

(b) sensitivity, specificity, positive predictive value and negative predictive value

		CAG		
		with stenosis	without stenosis	total
CCTA	with stenosis	a	b	a+b=m (+)
	without stenosis	c	d	c+d=m (-)
	total	a+c=n (+)	b+d=n (-)	a+b+c+d
	Sensitivity	a:n (+)		
	Specificity	d:n (-)		
	Positive Predictive Value	a:m (+)		
	Negative Predictive Value	d:m (-)		

5.　結果

5．1　CAG を模範解答とした iPARC 法石灰化プラーク除去画像臨床有効性

石灰化プラーク除去前 MIP 画像は, 右冠動脈 29 本のうち, CAG の読影結果と true positive あるいは true negative の相関関係を示し臨床有効性ありと判断したのは 2 本 (6.9%)，相関関係が無く臨床有効性なしと判断したのは 27 本(93.1%) であった. 臨床有効性なしと判断した 27 本 (93.1%) は全て false positive であり，原因は石灰化プラークの過大評価であった. 左冠動脈 45 本のうち，CAG の読影結果と true positive あるいは true negative の相関関係を示し臨床有効性ありと判断したのは 10 本 (22.2%)，相関関係が無く臨床有効性なしと判断したのは 35 (77.8%) 本であった. 臨床有効性なしと判断した 35 本は全て false positive であり，原因は石灰化プラークの過大評価であった. iPARC 法石灰化プラーク除去後 MIP 画像は, 右冠動脈 29 本のうち，CAG の読影結果と true positive あるいは true negative の相関関係を示し臨床有効性ありと判断したのは 23 (79.3%) 本，相関関係が無く臨床有効性なしと判断したのは 6 (20.7%) 本であった. 臨床有効性なしと判断した 6 (20.7%) 本の内訳は，blooming artifact が 1 本，over removal が 5 本であり，全て false positive であった. 左冠動脈 45 本のうち，CAG の読影結果と true positive あるいは true negative の相関関係を示し臨床有効性ありと判断したのは 34 本 (75.6%)，相関関係が無く臨床有効性なしと判断したのは 11 本

(24.4%) であった．臨床有効性なしと判断した 11 本の内訳は，blooming artifact が 3 本，over removal が 8 本であり，全て false positive であった．

5．2　Agatston スコア

Table2 は，左＋右冠動脈 74 本に対し，Agatston スコアを，0-200,201-400 401-600,601-800,801-1000,1001-1200,1201 - 1400,1401-1600,1601 - 1800,1801-2000,2001 以上の 11 グループに分類し評価を行った．CAG を模範解答とした石灰化プラーク除去前 MIP 画像と iPARC 法石灰化プラーク除去後 MIP 画像の臨床評価における，臨床有効性ありと判断した左＋右両冠動脈の血管数と比率である．iPARC 法を用い石灰化プラーク除去を行うことにより，0～200 は 5.9%から 88.2%，201～400 は 5.6%から 72.2%，401～600 は 8.3%から 91.7%，601～800 は 11.1%から 66.7%，801～1000 は 50.0%から 66.7%，1001～1200 は 20.0%から 100%へと，大幅に臨床有効性が向上した．

Table 2 Clinical evaluation of original MIP image, iPARC method MIP image compared with CAG as correct diagnosis of each Agatston score

		Original MIP image		iPARC method MIP image	
Agatston score	Number of total arteries	Number of TP and TN arteries	Percentage of TP and TN	Number of TP and TN arteries	Percentage of TP and TN
0～200	17	1	5.9	15	88.2
201～400	18	1	5.6	13	72.2
401～600	12	1	8.3	11	91.7
601～800	9	1	11.1	6	66.7
801～1,000	6	3	50.0	4	66.7
1,001～1,200	5	1	20.0	5	100.0
1,201～1,400	1	0	0.0	1	100.0
1,401～1,600	2	0	0.0	0	0.0
1,601～1,800	1	0	0.0	1	100.0
1,801～2,000	1	1	100.0	0	0.0
2,001～	2	0	0.0	1	50.0
total	74	9	12.2	57	77.0

5．3　CAG を模範解答とした iPARC 法石灰化プラーク除去画像臨床評価の感度・特異度・陽性的中率・陰性的中率

感度は，石灰化除去前 MIP 画像と iPARC 法石灰化除去後 MIP 画像ともに 100%であった．特異度は，石灰化石灰化除去前 MIP 画像が石灰化プラーク過大評価のため読影不可能と判断し全症例 CAG を行っているため 0%，iPARC 法石灰化除去後 MIP 画像が 73%であった．陽性的中率は，石灰化石灰化除去前 MIP 画像が 15%，iPARC 法石灰化除去後 MIP 画像が 39%であった．陰性的中率は，石灰化石灰化除去前 MIP 画像が石灰化プラーク過大評価のため読影不可能と判断し全症例 CAG を行っているため計算不能であり，iPARC 法石灰化除去後 MIP 画像は 100%であった．

6.　考察

6．1　CAG を模範解答とした iPARC 法石灰化プラーク除去画像臨床評価

冠動脈狭窄率を測定することは，疾患の重症度の評価と血行再建を含めた治療方針を決定する上で必須なため，従来は CAG 画像のみで行われていた．近年の CT 装置 hard ware および soft ware の進歩に伴い[31,32]，冠動脈狭窄率の評価精度は飛躍的に向上し，多くの施設で臨床使用されている[20,33-39]．CAG を模範解答とした iPARC 法石灰化プラーク除去画像臨床評価において，石灰化プラーク除去前 MIP 画像は，左＋右冠動脈 74 本のうち臨床有効性ありは 16.2% (12 本)，臨床有効性なしは 83.8% (62 本)であった．臨床有効性なしと判断した 83.8%の症例は全て false positive であり，原因は石灰化プラークの過大評価であった． iPARC 法石灰化プラーク除去後 MIP 画像は，左＋右冠動脈 74 本のうち臨床有効性ありは 77.0% (57 本)，臨床有効性なしは 23.0% (17 本)であった．臨床有効性なしと判断した 23.0%の症例は全て false positive であり，原因は blooming artifact が 23.5% (4 本)，over removal が 76.5% (13 本)であった．石灰化プラーク除去法を臨床使用する上で絶対禁忌事項である false negative は，存在しなかった．このことは，iPARC 法を用いた石灰化プラーク除去 MIP 画像が，CAG 画像と等価の診断能力を有する可能性とともに，病変見落としがないことを示唆している．

CAG は，冠動脈内腔に直接造影剤を注入し周囲臓器との X 線吸収差を投影画像として表示するため，石灰化プラークが冠動脈壁に付着して

いても血管内腔状態を詳細に観察することが可能な検査法である． iPARC 法が CAG と良好な相関関係にあることは，周囲臓器より大幅に高い CT 値を示す石灰化プラークと石灰化プラークのX線不透過性が原因で発生する blooming artifact[11]の不鮮明高輝度拡大陰影の三次元的形状を，適切に除去可能な処理法であることを示している．しかし，iPARC 法の zero cross point を用いた境界決定は，隣接 Pixel 間の CT 値差を強調し境界鮮鋭化後，操作者の手動操作で行われているため，主観的決定と恣意的操作の可能性を有している．これらの可能性を排除し，客観性と再現性を有する定量的境界を決定するためには，石灰化プラーク個々の CT 値分布を詳細に解析した後に，しきい値を自動決定する必要である．このことは石灰化プラーク除去精度向上とともに，臨床応用に際し操作者が短時間で簡便に操作するために必要な改善点と考える．

6．2　Agatston スコア

Agatston スコアを，0～200，201～400，401～600，601～800，801～1000，1000 以上の 6 グループに分類し，CAG を模範解答とした石灰化プラーク除去前 MIP 画像と iPARC 法石灰化プラーク除去後 MIP 画像の臨床評価における臨床有効性ありと判断した比率は， iPARC 法を用い石灰化プラーク除去を行うことにより，各グループとも 10%台から 70～80%台に向上し，全体では 16.3%から 77.0%に向上した．

true positive と true negative の評価は iPARC 法石灰化プラーク除去後 MIP 画像が CAG と同等の診断能力を有しているため追加検査必要なしを意味し，false fositive の評価は狭窄率を過大評価しているため CAG の確認追加検査必要ありを意味している．Agatston スコア 600 を超える症例であっても CAG と良好に相関する症例は多数存在した．しかし，今回の評価は，単一施設の後ろ向き調査 (Retrospective survey) のため，臨床有効性ありと判断した比率は予測的中率を示しているにすぎない．そのため，今回の評価だけで Agatston スコア Cut off 値引き上げを言及することは不可能である．今後は，無作為抽出多施設試験等の追加試験と被検者の症状と状態を考慮した専門医の総合判断を行うことにより，iPARC 法臨床使用時の血管内腔評価可能な Agatston スコア Cut off 値引き上げが可能になると考える．

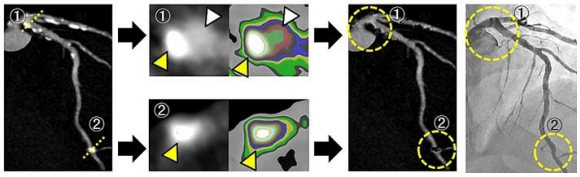
6．3　CAG を模範解答とした石灰化プラーク除去前 MIP 画像と iPARC 法石灰化プラーク除去後 MIP 画像の感度・特異度・陽性的中率・陰性的中率

iPARC 法は，感度と陰性的中率は 100%であったが，陽性適中率 39%であった．この値は，表 8 のメタ解析平均陽性適中率の半分の精度である．陽性適中率は，陽性になった症例の中で本当に病気に罹患している割合を示しており，有病率が低いほど陽性的中率は低下することが知られている[28-30]．今回の評価は，75%以上狭窄率を有する血管は 75 本中 11 本 (14.9%) と低かったことが iPARC 法陽性適中率低下の原因と考える．

7.　iPARC 法抽出限界

Fig.5 ①は，iPARC 法分離抽出可能な石灰化プラークである．石灰化プラークの cross section view は，石灰化プラークを白色～灰色に，冠動脈を灰色～黒色に表示している．cross section view を一定の CT 値毎に同色で表示した CT 値カラーマップも，石灰化プラークと冠動脈を双峰性ピークを有する異なる色で描出している．オリジナル画像で石灰化プラークと冠動脈が解像されている場合，iPARC 法は両者の濃度差を検出し客観的分離抽出が可能となる．これに対し，Fig.5 ②は，iPARC 法分離抽出不可能な石灰化プラークである．iPARC 法は，再構成関数と鮮鋭化フィルターの二段階処理を行い冠動脈と石灰化プラーク CT 値差を顕在化後 zero cross points を検出し，両者を分離抽出している．Fig.5②示すように，石灰化プラークから発生した blooming artifact にともなう辺縁不鮮明高輝度陰影が，CT 装置の持つ最大画像分解能力(over all resolution)の上限を超え，オリジナル画像で石灰化プラークと冠動脈が分離抽出されていない画像は，iPARC 法を用いても両者の分

離抽出は不可能である.



befor calcified plaques removal　cross section view　after calcified plaques removal　CAG

Fig.5 Limitation of iPARC method

8.　結論

iPARC 法を用い石灰化プラーク除去を行うことにより，臨床有効性は 16.2%から 77.0% に向上し，CAG との相関関係は Agatston スコアの大小にかかわらず 16.3%から 77.0%に向上した. iPARC 法陽性適中率は，石灰化プラーク非付着症例 78%に対し 39%であった. iPARC 法を用いることにより CAG と良好に相関関係し石灰化プラーク症例の臨床有効性は向上するため，CCTA 石灰化プラーク付着症例新規診断画像として臨床使用が期待された.

9.　謝辞

研究遂行にあたり貴重なご助言とご協力を賜った共同研究者の皆様方に深謝いたします.

10.　利益相反の有無

本研究に利益相反はありません.

11.　参考文献

［1］　The Japanese Registry Of All cardiac and vascular Diseases (JROAD): Annual Report2017 http://www.j-circ.or.jp/jittai_chosa/.(accessed March 28, 2020)

［2］　日本循環器学会　循環器疾患診療実態調査　2017 年(2018 年度実施・公表)循環器疾患診療実態調査　報告書　Web 版 http://www.j-circ.or.jp/jittai_chosa/.(accessed March 28, 2020)

［3］　Schulman-Marcus J, Lin FY, Gransar H, et al. Coronary revascularization vs. medical therapy following coronary-computed tomographic angiography in patients with low-, intermediate- and highrisk coronary artery disease: results from the CONFIRM long-term registry. Eur Heart J Cardi-ovasc Imaging 2017; 18: 841-848. PMID:28329294

［4］　Motoyama S, Ito H, Sarai M, et al. Plaque Characterization by Coronary Computed Tomog-raphy Angiography and the Likelihood of Acute Coronary Events in Mid-Term Follow-Up. J Am Coll Cardiol 2015; 66: 337-346. PMID: 26205589

［5］　北川　覚也, 佐久間　肇: 4. Beyond 64 – slice MDCT 4-2. 320 列 CT, 日獨医報, 第 55 号 第 1 号 43 – 49, (2010)

［6］　市川　勝弘: 誌上講座 CT における三次元画像再構成, 日本放射線技術学会雑誌, Vol. 68 No. 12 Dec 1609 – 1675, (2012)

［7］　Abbara S, Blanke P, Maroules CD, et al. SCCT guidelines for the performance and acquisition of coronary computed tomographic angiography:A report of the society of Cardiovascular Com-puted Tomography Guidelines Committee: Endorsed by the North American Society for Car-diovascular Imaging (NASCI). J Cardiovasc Comput Tomogr 2016; 10: 435-449. PMID: 27780758

［8］　Douglas PS, De Bruyne B, Pontone G, et al. PLATFORM Investigators. 1-Year Outcomes of FFRCT-Guided Care in Patients With Suspected Coronary Disease: The PLATFORM Study. J Am Coll Cardiol 2016; 68: 435-445. PMID: 27470449

［9］　Feuchtner G, Kerber J, Burghard P, et al. The high-risk criteria low-attenuation plaque <60 HU and the napkin-ring sign are the most powerful predictors of MACE: a long-term follow-up study. Eur Heart J Cardiovasc Imaging 2017; 18: 772-779. PMID: 27502292

［10］　Nørgaard BL, Hjort J, Gaur S, et al.

Clinical Use of Coronary CTA-Derived FFR for Deci-sion-Making in Stable CAD. JACC Cardiovasc Imaging 2017; 10: 541-550. PMID: 27085447

［11］ Curzen NP, Nolan J, Zaman AG, et al. Does the Routine Availability of CT-Derived FFR In-fluence Management of Patients With Stable Chest Pain Compared to CT Angiography Alone?: The FFRCT RIPCORD Study. JACC Cardiovasc Imaging 2016; 9: 1188-1194.PMID: 27568119

［12］ Gilbert L. Raff, Aiden Abidov, Stephan Achenbach, et al.: SCCT guidelines for the interpretation and reporting of coronary computed tomographic an-giography. Journal of Cardiovascular Computed Tomography, 3, 122-136, (2009)

［13］ 井田 義宏: マルチスライス CT による心臓診断技術, 日本放射線技術学会雑誌 , 第 62 巻 第 6 号 804 – 810, (2006)

［14］ 近藤 武, 高瀬 真一, 安野 泰史, 元山 貞子: マルチスライス CT による冠動脈評価, THE JOURNAL of JAPANEASE COLLEGE of ANGIOLOGY, Vol. 46, 157 – 164, (2006)

［15］ Agatston A. S., Janowitz W. R., Hildner F. J., et al.: Quantification of coronary artery calcium using ultrafast computed tomography., J Am Coll Cardiology Mar 15;15(4), 827 – 832, (1990)

［16］ Suhny Abbara, Armin Arbab-Zadeh, Tracy Q. Callister, et al.: SCCT guidelines for performance of coronary computed tomographic angiography: A report of the Society of Cardiovascular Computed Tomography Guidelines Committee, (2009)

［17］ Pamela S. Douglas, Mario J. Garcia, David E. Haines, et al. : ACCH/SCCT/ACR/AHA/ASF/ASNC/NASCI/ SCAI/SCMR 2010 Appropriate Use Criteria for Cardiac Computed Tomography, (2010)

［18］ 船橋 伸禎 : 欧州心臓学会より 2012 年に発表された 6 つの診療ガイドラインと CT の役割 , INNERVISION , (27, 11), 74 – 79 , (2012)

［19］ Suhny Abbara, Armin Arbab-Zadeh, Tracy Q. Callister, Milind Y. Desai, Wilfred Mamuya, et al.: SCCT guidelines for performance of coronary computed tomographic angiography: A report of the Society of Cardiovascular Computed Tomography Guidelines Committee, (2009)

［20］ 船橋 伸禎 : 冠動脈CT血管造影における冠動脈狭窄の検出と再定義 (前篇) , IN-NERVISION , (28,7), 82 – 85 , (2013)

［21］ 山岸 正和, 玉木 長良 : 慢性冠動脈疾患診断ガイドライン（2018 年改訂版）, https://www.j-circ.or.jp/cms/wp-content/uploads/2020/02/JCS2018_yamagishi_tamaki.pdf

［22］ 笠井 俊文, 小川 敬壽 : 診療画像機器学, (株) オーム社 , 199 – 202 ,(2006)

［23］ White CW, Wright CB, Doty DB, et al. Does visual interpretation of the coronary arteriogram predict the physiologic importance of a coronary stenosis? N Engl J Med 1984; 310: 819-824.

［24］ Little WC, Constantinescu M, Applegate RJ, et al. Can coronary angiography predict the site of a subsequent myocardial infarction in patients with mild-to-moderate coronary artery disease? Cir-culation 1988; 78: 1157-1166.

［25］ Falk E, Shah PK, Fuster V. Coronary plaque disruption. Circulation 1995; 92: 657-671.

［26］ Brown BG, Bolson E, Frimer M. Quantitative coronary arteriography : estimation of di-mensions, hemodynamic resistance, and atheroma mass of coronary artery lesions using the arte-rio-gram and digital computation. Circulation 1977; 55: 329-337.

［27］ WG Austen, JE Edwards, RL Frye, et al.: A reporting system on patients evaluated for

coronary artery disease. Report of the Ad Hoc Committee for Grading of Coronary Artery Disease, Council on Cardi-ovascular Surgery, American Heart Association .,Circulation. ,51:5-40, (1975)

［28］ 川口 直人,城戸 輝仁,望月 輝一 : 冠動脈 CT, 臨床放射線,Vol.58,No.7,933-943,(2013)

［29］ 感度 Wikipedia : http://ja.wikipedia.org/wiki/%E6%84%9F%E5%BA%A6

［30］ いちばんやさしい医療統計 陽性的中率と陰性的中率 特徴と違いとは？ : https://best-biostatistics.com/contingency/ppv-npv.html

［31］ 統計 WEB 1-1. 検査精度 : https://bellcurve.jp/statistics/course/14802.html

［32］ 安野 泰史 : 0.4 秒スキャン MSCT の有用性 心臓への応用, INNERVISION 別冊付録 , (18.5), 4 – 10 , (2003)

［33］ 新沼 廣幸 : MSCTA による心臓弁膜症の評価 : 心エコーとの比較, 第 54 回日本心臓病学会学術集会・東芝共催セミナー講演集 , 2 – 7 , (2007)

［34］ 田中 良一, 吉岡 邦浩 : 石灰化サブトラクション技術を用いた冠動脈 CT の実際 , INNERVISION , (29.7), 118 – 119. (2013)

［35］ Min JK, Shaw LJ, Berman DS. : The present state of coronary computed tomography angi-ography a process in evaluation., J. Am. Coll. Cardiol. 55, 957 – 965, (2010)

［36］ Gerald M. Burma, Peter S. Fail, et al. : 心臓 CT はどのように有効か？ 医師たちの率直な意見, INNERVISION 別冊付録 , (21.8) 7 – 10. (2006)

［37］ Jagat Narula, Renu Virmani : 弱性プラークの性状評価, INNERVISION 別冊付録, (21.8), 11 – 13. (2006)

［38］ 藤井 隆 : MDCT の有用性 - 64 列 MDCT の一歩進んだ活用方法, INNERVISION , (23.2), 4 – 11. (2008)

［39］ マルチモダリティによる Cardiac Imaging の戦略と選択 –非侵襲的な検査としての CT・MRI・ Fusion imaging の有用性, INNERVISION , (24.4) , 4 – 5. (2009)

Development Image-Processing Based Method to Remove Calcified Plaques from Coronary Computed Tomography Angiography: Methodology and Clinical Evaluation

PhD Koki Shibuya[*1]; PhD Tetsuya Yuasa[*2], MD PhD Toshiya Fujiwara[*3], MD PhD Toru Shibahara[*3], Kazuya Kudo[*4], Aya Shoji[*4],;Haruka Kunezaki[*2], PhD Tomomi Omura[*5],

*1 Department of Radiological Technology, Faculty of Health Sciences

*2 Graduate School of Sience and Engneering, Yamagata University

*3 Department of Cardiovascular Medicine

*4 Department of radiology, Akita City Hospital

*5 Department of radiology, Akita Cerebrospinal and Cardiovascular Center

Abstract:

Background: Blooming artifacts and beam hardening caused by calcified plaque adhering to coro-nary arteries have provided an obstacle to vascular lumen diagnosis on coronary computed tomog-raphy angiography. Development and clinical evaluation of image processing-based calcified plaque removal methods is essential to provide excellent characteristics of coronary computed tomography angiography (CCTA) to patients.

Method and Result: An image processing-based calcified plaque removal algorithm was developed using two-step sharpening and a secondary differential filter. This algorithm was named the image processing algorithm for removing calcified plaques (iPARC). This study evaluated 74 arteries in 48 patients who underwent electro-cardiogram (ECG) -gate CT angiography using MDCT. Accuracy of the iPARC method was compared with findings from coronary angiography (CAG) as the correct diagnosis. Two cardiologists evaluated the iPARC method as clinically effective in 57 arteries (77.0%), and not clinically effective in 17 arteries (23%). Clinical effectiveness for 11 groups classified by Agatston score (0–1200) improved significantly from 16.2% to 77.0%.

Conclusion: The iPARC method can predict the stenosis rate with comparable accuracy to CAG within the Agatston score range of 0–1,200. This method is expected to be used clinically as a new diagnostic imaging sequence for cases showing calcified plaque on CCTA.

Key words:

coronary computed tomography angiography; calcified plaque removal; image processing-based method

位相拡散フーリエ法の信号-画像間学習による MR 画像の超解像

若槻　泰迪[*1]　植松　駿[*1]　伊藤　聡志[*1]

要旨

位相拡散フーリエ変換法(PSFT)の信号は，被写体が実関数との拘束条件を使用すると，撮像後の数値処理により信号の外挿が可能であり，高解像度化した画像を再生できる．これまで反復的手法により MR 信号を外挿し，再生像の空間分解能改善を行ってきた．しかし，この方法では分解能の改善は画像中央で小さく，また，画像の再生には 20 回程度の反復処理を要する．本研究では深層学習再構成である Generic-ADMM-Net を利用し，PSFT 信号の信号外挿と空間分解能の改善を図った．信号取得幅を 1/2 に制限した PSFT 信号を深層学習に入力し，全信号による再生像を教師像として学習を行った結果，折り返しアーティファクトが少なく，高分解能化された再生像を得ることができた．また，信号に雑音が含まれる場合には雑音の少ない画像を再生するという反復的解法にはない特徴が示された．

キーワード：位相拡散，MRI，深層学習，超解像

1. はじめに

核磁気共鳴映像法(MRI)において画像の高分解能化は，画像診断の精度に直結する重要なテーマである．汎用法であるフーリエ変換映像法を使用する場合，信号のデータ数が与えられたときに画像の分解能は画像空間での折り返しを避けるようサンプリング定理に従い決定される．高分解能画像を得る為，位相エンコード方向のデータ数を増やすと一般的に撮像時間が長くなるが，本研究では撮像時間を延長することなく，撮像後に再生像を高分解能化する方法について検討を行う．

位相拡散フーリエ変換法［1］［2］は，信号のダイナミックレンジを低減するために提案された方法であり，フーリエ変換映像法のパルスシーケンスの位相エンコードに同期して，被写体のスピンに二次の位相変調を与えることにより，すべてのスピンが同相になる時相を排除し，大振幅の信号の出現を押さえることができる．この二次の位相を与える処理は信号ダイナミック低減以外に，エイリアスレス再構成，単一信号パラレルイメージング，エコープラナー法との併用による高画質化など，多くの優れた特性が明らかになっている［3］［4］．伊藤らは位相拡散フーリエ変換法の信号に対し，Gerchberg の超解像アルゴリズム［5］と類似の反復アルゴリズムを利用し，信号を外挿する方法により事後的に再生像分解能改善を行う方法を提案した［6］．この方法は少数の観測信号から画像を復元する点で圧縮センシングに似ているが，信号は間引きをせずに信号空間の収集領域を制限する点，および画像再生に要する計算コストは圧縮センシングに比べ少なく済む点で異なる．この方法は 10~20 回程度の反復処理を要し，また，信号の打ち切りによるエイリアスが残る場合があった．

近年，医用画像に深層学習を用いる研究が注目されている．深層学習を，MR 画像の超解像や圧縮センシング MRI(CS-MRI)［7］に応用すると，従来の反復的手法よりも高品質な画像を再生できることが報告されている［8］［9］．中でも，交互方向乗数法（Alternating Direction Method of Multipliers: ADMM）の計算ステップをモデルとして深層学習の層に組み込んだ

*1 宇都宮大学大学院 地域創生科学研究科 工農総合科学専攻
〔〒321-8585 宇都宮市陽東 7-1-2〕
e-mail:mc196878@cc.utsunomiya-u.ac.jp

Generic-ADMM-Net は，間引き信号を入力とし良好な再構成像が得られている［10］.

本研究では，位相拡散フーリエ変換法の信号を Generic-ADMM-Net を使用して外挿し，超解像的に画像拡大する新たな方法を提案する．深層学習の利用により，エイリアスや雑音の抑制，高分解能化，再構成時間の短縮化が期待できる．本研究では，Generic-ADMM-Net を使用した提案法を従来型の反復的再生像と比較し，深層学習を利用する方法の特徴を明らかにすることを目的とする．

2. 位相拡散フーリエ変換法

2.1 原理

位相拡散フーリエ変換法(phase-scrambling Fourier transform imaging: PSFT)は，位相エンコードに同期して二次関数状の磁界を一定の期間印加するか，あるいは線形勾配磁界を印加のもと周波数変調された RF 励起パルスを照射するなどの方法により，被写体内のスピンの位相を拡散させ，読み出し用の勾配磁界によりエコー信号を読み出す方法である．信号式を式(1)に示す．

$$v(\gamma G_x t_x, \gamma G_y t_y) = \iint_{-\infty}^{\infty} \{\rho(x,y) e^{-j\gamma b\tau(x^2+y^2)}\} \times e^{-j(\gamma G_x t_x x + \gamma G_y t_y y)} dxdy \quad (1)$$

ここで，$\rho(x,y)$は xy 平面のスピン密度分布，b は 2 次関数状磁界の係数，τ はその印加時間，G_x と G_y はそれぞれ x 方向と y 方向の線形勾配磁界の係数である．また，t_x，t_y をそれぞれ G_x と G_y の印加時間，γ を核磁気回転比とした．式(1)は $k_x = \gamma G_x t_x$，$k_y = \gamma G_y t_y$ と変数変換することにより，式(2)のように位相変調を施したスピン密度関数のフーリエ変換式になる．

$$v(k_x, k_y) = \iint_{-\infty}^{\infty} \{\rho(x,y) e^{-j\gamma b\tau(x^2+y^2)}\} \times e^{-j(k_x x + k_y y)} dxdy \quad (2)$$

PSFT 信号からの画像再生は，式(2)の信号を逆フーリエ変換した後に，式(3)に示す与えた二次の位相変調効果の逆位相関数を乗じればよい．

$$\rho(x,y) = e^{j\gamma b\tau(x^2+y^2)} F^{-1}[v(k_x, k_y)] \quad (3)$$

2.2 信号の外挿

被写体関数が実関数であるとの拘束条件を与えた場合に，信号を外挿する従来法の反復処理手順を図 1 に示す．反復アルゴリズムには，Gerchberg の超解像アルゴリズム［5］と類似の反復アルゴリズムを利用する．

1) (a)の PSFT 信号(N 点)の外側にゼロデータを充填し，サンプリング長を N から $2N$ に広げる．
2) PSFT 信号から(b)に示す再生像を得るために，信号を逆フーリエ変換し，施した位相変調の逆位相を乗算する．
3) (c)で得られた像に実関数の拘束条件を適用し，実部を再生像とする．この再生像に対して位相変調を与え，フーリエ変換によって(e)の PSFT 信号を計算合成する．
4) この PSFT 信号の中央 N 点分は(a)で収集した真値の信号で置換する．
5) (f)の PSFT 信号を逆フーリエ変換し，(b)の更新された再生像を得ることができ，以降 2)か

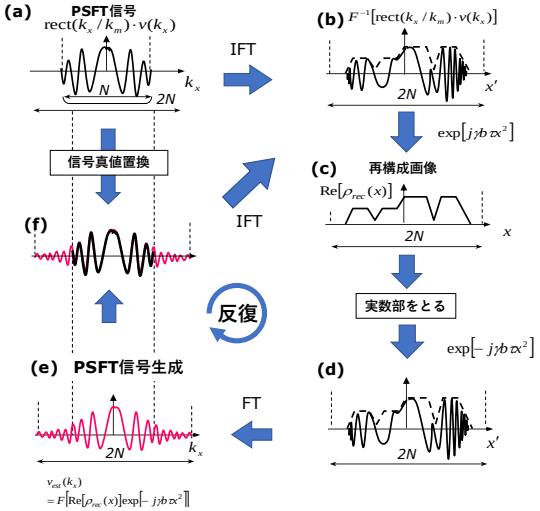


図 1 PSFT 信号の反復手法による信号の外挿概略図：(a) PSFT エコー信号，(b) 信号(a)の逆フーリエ変換，(c) 2 次位相を復調した再生像，(d) (c)実数部への二次位相変調，(e) (d)のフーリエ変換，(f) (e)中心データの真値置換

ら 5)までの処理を繰り返すことにより，PSFT 信号帯域はゼロデータを充填した空間に外挿され，空間分解能を改善した画像を得ることができる．信号帯域の外挿は，ハーフフーリエ法で帯域制限された信号が，実関数の拘束条件で外挿される過程と同じ原理である．ただし，外挿には反復処理を要する．

3. Generic-ADMM-Net

3.1 交互方向乗数法(ADMM)

拘束条件のもとで与えられる条件を満足する解を求める方法に交互方向乗数法(alternating direction method of multipliers: ADMM)がある．

ここで， x を MR 再構成像，y を制限された k 空間とする．信号に制限を与える関数を P，フーリエ変換を F とする．一般的な CS-MRI モデルは式(4)の最適化問題を求めることで再構成像を推定する．D_l は離散コサイン変換などの変換行列，λ_l は正則化パラメータ，$g(\cdot)$は正則化関数， L はフィルタの数を示す．

$$\hat{x} = \underset{x}{\operatorname{argmin}} \left\{ \frac{1}{2} \|PFx - y\|_2^2 + \sum_{l=1}^{L} \lambda_l g(D_l x) \right\} \quad (4)$$

ADMM は拡張ラグランジュ乗数法および変数分離により，式を 3 つの最適化問題を解く式(5)へと変形する．

$$\begin{cases} \underset{x}{\operatorname{argmin}} \dfrac{1}{2} \|PFx - y\|_2^2 + \displaystyle\sum_{l=1}^{L} \frac{\rho}{2} \|x + \beta - z\|_2^2, \\ \underset{z}{\operatorname{argmin}} \displaystyle\sum_{l=1}^{L} \lambda_l g(D_l z) + \frac{\rho}{2} \|x + \beta - z\|_2^2, \\ \underset{\beta}{\operatorname{argmin}} \displaystyle\sum_{l=1}^{L} \langle \beta_l,\ x - z \rangle . \end{cases} \quad (5)$$

式(6)は，式(5)に示された最適化問題を解く反復式で，n 回の反復に応じて解を求める．

I は $N \times N$ の単位行列を示す．$z^{(n,k)}$では関数の最小値を探索する勾配法の一種である最急降下法が用いられており，k は最急降下法の反復回数である．$\mu_1 = 1 - l_r$，$\mu_2 = l_r \rho$，$\tilde{\lambda} = l_r \lambda_r$，$l_r$ は最急降下法のステップサイズ，$\mathcal{H}(\cdot)$は正則化関数の勾配に対する非線形変換である．

$$\begin{cases} x^{(n)} = F^{\mathrm{T}}(P^{\mathrm{T}}P + \rho I)^{-1} \\ \qquad \left[P^{\mathrm{T}} y + \rho F\left(z^{(n-1)} - \beta^{(n-1)}\right)\right], \\ z^{(n,\mathrm{k})} = \mu_1 z^{(n,k-1)} + \mu_2 \left(x^{(n)} + \beta^{(n-1)}\right) \\ \qquad - \displaystyle\sum_{l=1}^{L} \tilde{\lambda}_l D_l^{\mathrm{T}} \mathcal{H}\left(D_l z^{(n,k-1)}\right), \\ \beta^{(n)} = \beta^{(n-1)} + \tilde{\eta}\left(x^{(n)} - z^{(n)}\right). \end{cases} \quad (6)$$

3.2 ネットワーク構造

ADMM では，最適な変換行列 D_l や非線形変換$\mathcal{H}(\cdot)$を選択することは困難なうえ，異なる信号収集方法に対してパラメータを調整する必要がある．Generic-ADMM-Net はこれらをネットワークで学習し，CS-MRI の最適解を求める．Generic-ADMM-Net の構造を図 2 に示す．

このネットワーク構造は，ADMM アルゴリズムの各処理に対応した層と処理間のデータの流れに対応した矢印で表される．ADMM アルゴリズムの n 回目の反復は図 2 の n 番目のステージに対応しており，式(4)に示す $x^{(n)}$，$z^{(n,k)}$，$\beta^{(n,k)}$ の処理にそれぞれ対応した再構成層($X^{(n)}$)，非線形変換層($Z^{(n)}$)，乗数更新層($M^{(n)}$)より構成される．Generic-ADMM-Net では，間引き MR 信号空間を入力とし，各ステージを経て再構成 MR 画像を出力する．本研究では，CS-MRI における信号制限関数 P を図 3 に示す帯域制限フィルタ，MR 信号を PSFT 信号とすることで，ネットワークへの入力を信号取得幅が 1/2 に制限された PSFT 信号とし，画像再生を検討した．

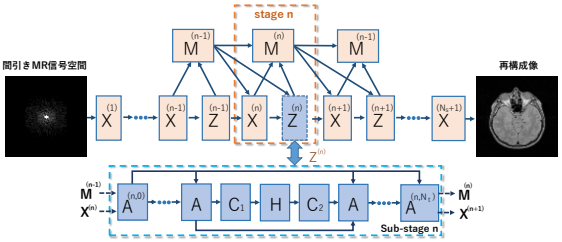


図 2 Generic-ADMM-Net の構造

二次の位相変調を式(7)に示す S とすると，PSFT 信号を入力とした Generic-ADMM-Net では，再構成層($X^{(n)}$)の順伝播式，逆伝播式はそれぞれ式(8)，(9)で与えられる．S*は S の複素共役である．

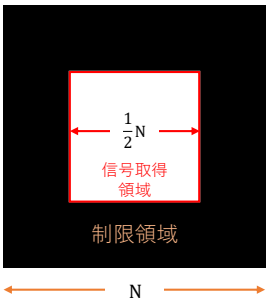


図 3 帯域制限フィルタ

$$S = e^{-j\gamma b\tau\left(x^2+y^2\right)} \tag{7}$$

$$x^{(n)} = S^*F^{\mathrm{T}}(P^{\mathrm{T}}P + \rho I)^{-1} \times \left[P^{\mathrm{T}}y + \rho FS\left(z^{(n-1)} - \beta^{(n-1)}\right)\right] \tag{8}$$

$$\frac{\partial E}{\partial \rho^{(n)}} = \frac{\partial E}{\partial x^{(n)T}} S^*F^{\mathrm{T}}\left(P^{\mathrm{T}}P + \rho^{(n)}I\right)^{-1} \times \{F\left[S\left(z^{(n-1)} - \beta^{(n-1)}\right)\right] - \left(P^{\mathrm{T}}P + \rho^{(n)}I\right)^{-1} \times [P^{\mathrm{T}}y + \rho^{(n)}F\left(S\left(z^{(n-1)} - \beta^{(n-1)}\right)\right)]\} \tag{9}$$

4. ネットワークによる学習

本研究では MR 画像から数値的に合成した PSFT 信号に対し，Generic-ADMM-Net を利用した画像再生(Net-PSFT と称する)を行い，従来の反復手法を利用した画像再生(Iter-PSFT と称する)，および，汎用のフーリエ変換法に本方法を適用した場合(Net -FT と称する)と比較を行った．学習とテストに使用した MR 画像には IXI Dataset［11］に含まれる頭部 T1 強調像であり，Philips 製 1.5 テスラ MRI により撮像された画像である．ネットワークを利用した再生実験では，120 枚の画像を学習に 100 枚，テストに 20 枚の割合で分割し，ホールドアウト法による検証を行った．信号データと画像データマトリクスは，256×256 とし，信号は中央の 128×128 領域を収集する．再生像の評価にはピーク信号対雑音比(Peak Signal-to-Noise Ratio)と SSIM (Structural SIMilarity index)［12］を使用した．学習と再構成に使用した計算機環境を表 1 に示す．学習には GPU を用い，表 2 に示す学習条件でおよそ 44 時間を要した．

表 1 計算機環境

CPU	Intel Core i7-8700 (3.2GHz)
RAM	32GB
GPU	NVIDIA GeForce GTX 1070 Ti
Software	MATLAB R2017b CUDA Toolkit 9.0 / cuDNN 7.0.5

表 2 学習条件

畳み込みフィルタサイズ	5×5
畳み込みフィルタ数	128
ステージ数	10
パラメータ更新手法	L-BFGS-B［13］

5. 画像再生シミュレーション

5.1 数値シミュレーション

提案法の空間分解能改善程度を調べるために，数値モデルを使用したシミュレーションを実施した．図 4 に結果を示す．図 4(a)は全信号を利用した再生像，(b)は信号取得幅が 1/2 に制限された PSFT 信号のゼロフィル再生像，(c)は従来法である Iter-PSFT 再生像，(c)は提案法の Net -PSFT 再生像である．各図上段には再生像中央部を，下段には(a)中の A-A’線上のプロファイルを示す．比較のため，プロファイル図中に振幅 100 および 25 となる位置に破線を記している．(c) Iter-PSFT 再生像では小孔部の谷の振幅が像空間の中心領域で低下している．(d) Net -PSFT 再生像で中心付近の谷の落ち込みが Iter-PSFT よりもやや大きく，鮮鋭度が改善されていることがわかる．

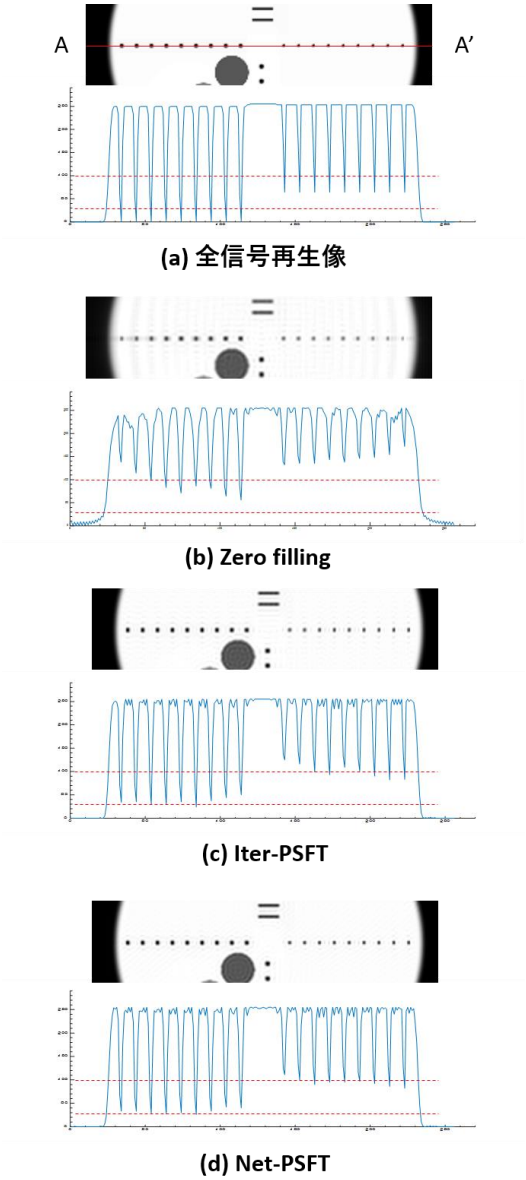


図 4　プロファイルによる分解能評価：(a) 全信号再生像，(b) ゼロフィル再生像，(c) Iter-PSFT による再生像，(d) Net-PSFT による再生像

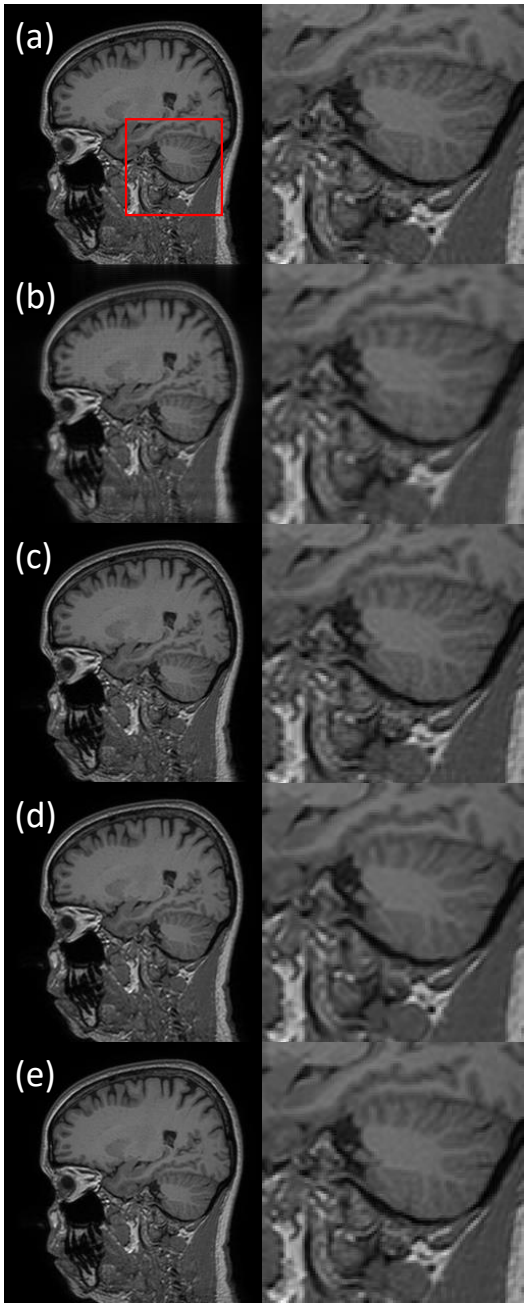


図 5　画像再生結果:(a) 全信号利用再生像，(b) ゼロフィル再生像, (c) Iter-PSFT 再生像，(d) Net-FT 再生像，(e) Net-PSFT 再生像

5.2 MR 画像シミュレーション

MR 画像に対し再生シミュレーションを行った．再生像を図 5 に，再生像の PSNR および SSIM を表 3 に示す．図 5(b)に示す信号取得幅が 1/2 に制限された PSFT 信号の再生像は，信号の打ち切りによる折り返しアーティファクトが見られるが, Iter-PSFT による再生像(c)および Net -PSFT による再生像(e)はアーティファクトが少なく，良好に画像再生を行えていることが見てとれる. さらに, Net -PSFT では Iter-PSFT と比較して周辺領域では同等の鮮鋭度をもち

つつ，画像の中央領域でより構造保存性が改善され，図 4 の鮮鋭度復元評価と符合する結果が得られた．

次に雑音の混入された条件における再生像への影響を調べるため，MR 画像の最大振幅に対し，雑音の標準偏差が 2.5%となるように調整した白色ガウス雑音を PSFT 信号に与え，画像再生を行った．Net -PSFT では，雑音を加えた PSFT 信号を入力し，雑音を付与してない全信号による再生像を教師像として学習を行った．

雑音量が 2.5%の場合の再生像を図 6 に，再生像の PSNR および SSIM を表 4 に示す．Iter-PSFT では雑音を含めて拡大されるため，帯域制限により補間された雑音が重畳しているが，Net -PSFT では雑音を抑えた画像が得られた．表 4 より，Net-PSFT では，PSNR が大きく改善された画像が得られた．画像再生に要した時間は 1 枚当たり約 1.27 秒と，短時間で再生可能であった．

表 3　再生像評価結果

	PSNR[dB]	SSIM
Zero-fill	28.93	0.9365
Iter-PSFT	35.44	0.9613
Net-FT	33.98	0.9711
Net-PSFT	**36.41**	**0.9833**

表 4　2.5%雑音信号の評価結果

	PSNR[dB]	SSIM
Zero-fill	27.61	0.8980
Iter-PSFT	31.74	0.9145
Net-PSFT	**33.69**	**0.9645**

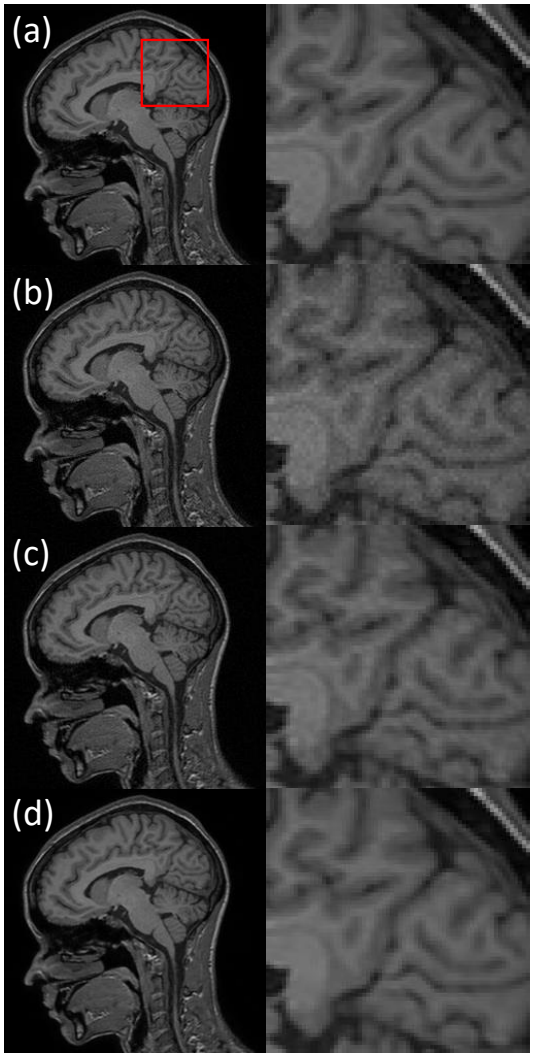


図 6　雑音(2.5%)を含む場合の画像再生結果：(a) 雑音付加無し全信号利用再生像，(b) 2.5%雑音を含む全信号利用再生像，(c) Iter-PSFT 再生像，(d) Net-PSFT 再生像

5. 考察

図 4 のプロファイルおよび図 5 の再生像より，Net-PSFT では従来の反復手法である Iter-PSFT に比べ，画像の中央領域において分解能がやや改善されていることが分かる．これは，信号外挿による分解能改善効果に深層学習ネットワークによる超解像効果が加えられたことに依るものと考える．また，汎用のフーリエ変換法を用いた Net-FT には見られない分解能改善効果が Net-PSFT に現れたのは，汎用の FT 信号には帯域の外挿効果が期待できないためである．

Iter-PSFT では，分解能は中心からの距離に比例して改善される一方で，S/N 比は距離の平方根に応じて低下する性質がある［6］．それに対し，Net-PSFT では，雑音が抑制された画像を再生できた．これは，雑音を含まない再生像を教師像としたために雑音を抑制するようなネットワークが構築されたことに起因すると考える．帯域拡大による超解像と雑音除去を両立し

た画像再生を行うことができるのは深層学習再構成の大きな特徴であると言える.

6. まとめ

位相拡散フーリエ変換法の信号は, 信号-画像間の深層学習再構成により外挿し，外挿信号から分解能改善効果を含んだ画像再生が可能であることが示された．また，雑音に対して頑健であり，雑音を抑制した画像を生成できる可能性が示された．今後の課題は，さらなる高分解能化と空間的に位相をもつ画像への適用があげられる.

謝辞

本研究の一部は，科学研究費補助金(19K04423)および柏森情報科学振興財団の助成により実施された. また, 本研究で利用した IXI Dataset を提供する Information eXtraction from Images プロジェクトに感謝の意を表します.

利益相反の有無

なし

文　献

［1］ Maudsley AA: Dynamic range improvement in NMR imaging using phase scrambling. Journal of Magnetic Resonance **76**: 287-305, 1969

［2］ Wedeen VJ, Chao YS, Ackerman JL: Dynamic range compression in MRI by means of nonlinear gradient pulse. Magnetic Resonance in Medicine **6**: 287-295, 1988

［3］ 伊藤聡志，中村俊，劉娜 他: 位相拡散フーリエ変換映像法を用いた MR アンチエイリアス画像再構成．電子情報通信学会論文誌 **J90-D**: 1149-1157, 2007

［4］ 伊藤聡志, 山田芳文: 位相拡散フーリエ変換法を使用した単一信号 PILS．日本磁気共鳴医学会論文誌 **31**: 20-31, 2011

［5］ Gerchberg RW: Super-resolution through error energy reduction. Optica Acta **21**:709-720, 1974

［6］ 伊藤聡志，劉娜，山田芳文 : 位相拡散フーリエ変換法と超解像法的手法による MRI 測定時間の短縮．日本磁気共鳴医学会誌 **28**: 142-153, 2008

［7］ Donoho DL: Compressed sensing. IEEE Transactions on Information Theory, **52**, no.4, pp.1289-1306, 2006

［8］ Pham CH, Docournau A, Fablet R, et al.: Brain MRI super-resolution using deep 3D convolutional networks. 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), pp. 197-200, 2017

［9］ Yang Y, Sun J, Li H, et al.: Deep Admm-Net for Compressive Sensing MRI. Advances in Neural Information Processing Systems, pp.10-18, 2016

［10］ Yang Y, Sun J, Li H, et al.: ADMM-CSNet: A Deep Learning Approach for Image Compressive Sensing. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018

［11］ IXI Dataset. https://brain-development.org/ixi-dataset/

［12］ Wang Z, Bovik AC, Sheikh HR, et al.: Image quality assessment: from error visibility to structural similarity, IEEE Transactions on image processing, **13.4**: pp.600-612, 2004

［13］ Zhu C, Byrd RH, Lu P, et al.: Algorithm 778: L-BFGS-B: Fortran subroutines for large-scale bound-constrained optimization. ACM Transactions on Mathematical Software 23.4 (1997): 550-560

Super-resolution of MR images using Signal-image Domain Learning in Phase Scrambling Fourier Transform Imaging

Hiromichi WAKATSUKI[*1], Shun UEMATSU[*1], Satoshi ITO[*1]

*1 Graduate School of Engineering, Utsunomiya University

Spatial resolution of images in phase-scrambling Fourier transform imaging can be improved by post-processing signal band extrapolation. Iterative reconstruction process has been applied to expand the signal band so far. However, the improvement of spatial resolution is small in the central area on image space and takes several time to image processing. In our research, super-resolution using deep convolutional neural network Generic-ADMM-Net is applied to expand the signal band and improve the resultant spatial resolution of images. Simulation studies using the half-width signal showed that spatial resolution was improved with reduced aliasing artifacts. It was also shown that reduce noise images are reconstructed when noisy signals were input to the network, which is a new feature not found in iterative reconstruction.

Key words: Phase Scrambling, MRI, Deep Learning, Super resolution

位相拡散フーリエ変換映像法の導入による MR 深層学習再構成の高画質化

植松 駿[*1], 伊藤 聡志[*1]

要旨

深層学習を利用した圧縮センシングの画像再構成は，再構成時間の短縮化と画像の高画質化のいずれにも有効な方法として注目されている．MRI に圧縮センシングを応用する場合において非線形勾配磁場を利用する位相拡散フーリエ変換映像法を使用すると，再構成像の画質を改善できることが報告されている．そこで，本研究では，交互方向乗数法(ADMM)を CNN で実現した Generic-ADMM-Net を使用し，深層学習による画像再構成を試みた．画像再構成シミュレーションを行った結果，従来のフーリエ変換を基本とする場合，および反復的再構成を使用した場合よりも，画像の精鋭度，構造保存性，コントラスト保存性のいずれにも改善された画像が得られた．本研究により，位相拡散フーリエ映像法を利用した圧縮センシングは深層学習再構成によっても高画質化が図られる可能性が示された．

キーワード : 圧縮センシング，位相拡散, 交互方向乗数法(ADMM), 深層学習, MRI

1. はじめに

核磁気共鳴現象を利用した生体映像法(Magnetic Resonance Imaging : MRI)では, 緩和による信号の回復を待つ時間を設けるため一般的に信号の収集時間が長く X 線 CT に比べ撮像に多くの時間を要する．撮像時間を短縮する手法として多くの方法が提案されているが, 圧縮センシング(Compressed Sensing : CS)[1]を MRI に応用する圧縮センシング MRI (CS-MRI)[2]は, 数理的に撮像時間短縮を可能とする新たな試みであり，さまざまな研究と応用が試みられている．

CS-MRI は, L1 ノルムを含んだ評価関数の最小化問題を解く形で画像再構成が行われる．多くの解法が提案されているが, 収束保証を伴う変数分割アルゴリズムである交互方向乗数法(Alternating Direction Method of Multipliers : ADMM)[3]は, L1 ノルム最小化問題の解法として一般的に知られる反復的ソフト閾値法に比べて良質な画像を再生できると報告されている[4]. しかしながら, 再構成の性能と時間に影響を与える最適な再構成パラメータを決めるのは必ずしも容易ではなく，これらを経験的に与える場合が多い．

近年, CS-MRI の再構成に深層学習を利用して行う研究が注目されている．深層学習を用いた画像再構成は, 従来の反復的解法よりも高速, かつ高画質であることが報告されており[5]~[8], 圧縮センシングの画像再構成への応用に期待が高まっている．

フーリエ変換映像法の信号（以降, FT 信号と称する）を間引きした信号から良質な画像を再構成する深層学習を利用した再構成法に Generic-ADMM-Net[6]がある．この方法は, ADMM の反復処理の各手順を畳込みニューラルネットワー

*1 宇都宮大学大学院 地域創生科学研究科
工農総合科学専攻
情報電気電子システム工学プログラム
[〒321-8585 栃木県宇都宮市陽東 7-1-2]
e-mail: mc196809@cc.utsunomiya-u.ac.jp

ク（CNN）のレイヤで構成する. Generic-ADMM-Net は, 反復再構成の手順をモデルとしているが, 各反復ステップにおける重みや係数を学習により調整できる柔軟な構造となっている. そのため, ADMM の反復再構成に比べて高品質かつ高速な画像再構成が期待できる.

近年, 汎用のフーリエ変換映像法に非線形勾配磁界を導入すると, フーリエ変換映像法にない性質を持つことがあることが報告されている[9]~[11]. 一例として, 位相拡散フーリエ変換映像法[12][13]の信号は, 被写体のフーリエ変換空間の信号であるが, 一方で被写体関数の回折波に似た信号式であることから画像空間に位置するとみなすことができ, この性質に注目した研究が増えつつある. CS-MRI の高画質化の一手法として, MRI の撮像法として汎用のフーリエ変換映像法に 2 次の位相変調項を追加した位相拡散フーリエ変換映像法の利用がある. 位相拡散フーリエ変換映像法に含まれる 2 次の位相変調は信号間引きを行った場合に再生像の誤差を拡散する効果があり, 圧縮センシングに利用するとフーリエ変換映像法よりも高画質な再構成像が得られることが報告されている[14][15]. また, 位相拡散フーリエ変換映像法では, 等間隔で間引かれた信号から画像再構成を行う方法も報告されている[16]. 本研究では, 位相拡散フーリエ変換映像法の信号を Generic-ADMM-Net により深層学習再構成を行い, 再構成像のさらなる画質改善について検討を行う. 汎用の FT 信号を入力とする深層学習再構成と比較を行い, 提案法の有用性を検討することを目的とする.

2. Generic-ADMM-Net

2.1. 交互方向乗数法 (ADMM)

圧縮センシングとは, 信号がスパース性を満たすとき, 標本化定理を満たさない少数の信号から元の信号を再構成する技術である. この手法を MRI に応用すると, FT 信号を間引くことにより, 撮像時間を短縮することができる. xを求めたい画像, yを間引き信号, 間引きを行う関数をU, フーリエ変換をF, スパース性を促す関数をΨとするとき, 画像再構成は式(1)の最適化問題を解くことにより行われる. ここで, λは正則化パラメータである.

$$\underset{x}{\operatorname{argmin}}\left\{\frac{1}{2}\|UFx-y\|_2^2+\lambda\|\Psi x\|_1\right\} \quad (1)$$

交互方向乗数法(ADMM)は, 二つの関数を含む最適化問題に対して, 新たな変数と制約条件を加え, 拡張ラグランジュ法を利用する方法である[3]. ADMM では式(1)において新たな変数zと制約条件を加えた式(2)の最適化問題を解く. ここで, $g(\cdot)$は L1 ノルムをより一般化した正則化関数である.

$$\underset{x,z}{\operatorname{argmin}}\left\{\frac{1}{2}\|UFx-y\|_2^2+\sum_{l=1}^{L}\lambda_l g(\Psi_l z)\right\} \quad (2)$$
$$s.t. \quad z=x$$

式(2)に拡張ラグランジュ法を適用すると式(3)が得られる.

$$L_{aug}(x,z,h)=\frac{1}{2}\|UFx-y\|_2^2+\sum_{l=1}^{L}\lambda_l g(\Psi_l z) +(h)^{\mathrm{T}}(x-z)+\frac{\mu}{2}\|x-z\|_2^2 \quad (3)$$

ここで, μを罰金項の係数とすると, 新たに導入した変数zと元の変数x, およびラグランジュ未定乗数 h は式(4)のように更新される.

$$\begin{cases} x^{(n)}=\underset{x}{\operatorname{argmin}}\ L_{aug}\left(x^{(n-1)},z^{(n-1)},h^{(n-1)}\right) \\ z^{(n)}=\underset{z}{\operatorname{argmin}}\ L_{aug}\left(x^{(n)},z^{(n-1)},h^{(n-1)}\right) \\ h^{(n)}=h^{(n-1)}+\mu\left(x^{(n)}-z^{(n)}\right) \end{cases} \quad (4)$$

ここで, nは反復回数を示す. このように変数を分割して最適化を行う方法が ADMM である.

2.2. Generic-ADMM-Net の構造

本研究で使用する Generic-ADMM-Net [6]は上記に示した ADMM から構成される. ADMM の式(4)から式(5)が導かれる. ここで, ηは更新係数, $\beta = h/\mu$としている.

$$\begin{cases} X^{(n)}: x^{(n)} = F^{\mathrm{T}}(U^{\mathrm{T}}U + \mu I)^{-1} \\ \qquad \{U^{\mathrm{T}}y + \mu F\left(z^{(n-1)} - \beta^{(n-1)}\right)\} \\ Z^{(n)}: z^{(n)} = \tilde{S}\left(x^{(n)} + \beta^{(n-1)}\right) \\ M^{(n)}: \beta^{(n)} = \beta^{(n-1)} + \eta\left(x^{(n)} - z^{(n)}\right) \end{cases} \tag{5}$$

このとき, $Z^{(n)}$に関する式に含まれる$\tilde{S}(\cdot)$は, Ψ_l, λ_lで決められる非線形演算子である. この式を近似的に解く手法の一つとして, 式(6) に示す最急降下法を用いる. ここで, kは$Z^{(n)}$式内で行われる反復回数に相当し, l_rは最急降下法のステップ幅, $\rho_1 = 1 - l_r\mu$, $\rho_2 = l_r\mu$, $\tilde{\lambda_l} = l_r\mu$, $H(\cdot)$は勾配に対する非線形変換を示す. また, $k = 0$においては $z^{(n,0)} = x^{(n)} + \beta^{(n-1)}$ とする.

$$\begin{aligned} z^{(n,k)} &= z^{(n,k-1)} - l_r\{\mu z^{(n,k-1)} - \mu\left(x^{(n)} + \beta^{(n-1)}\right) \\ &\qquad + \sum_{l=1}^{L} \lambda_l \Psi_l^{\mathrm{T}} H(\Psi_l z^{(n,k-1)})\} \\ &= \rho_1 z^{(n,k-1)} + \rho_2\left(x^{(n)} + \beta^{(n-1)}\right) \\ &\qquad - \sum_{l=1}^{L} \tilde{\lambda}_l \Psi_l^{\mathrm{T}} H\left(\Psi_l z^{(n,k-1)}\right) \end{aligned} \tag{6}$$

上記の式(6)は$c_1^{(n,k)}$, $c_2^{(n,k)}$, $h^{(n,k)}$, 畳込み演算子 D を用いて以下の式で表される.

$$\begin{cases} C_1^{(n,k)}: c_1^{(n,k)} = Dz^{(n,k-1)} \\ H^{(n,k)}: h^{(n,k)} = H\left(c_1^{(n,k)}\right) \\ C_2^{(n,k)}: c_2^{(n,k)} = Dh^{(n,k)} \\ A^{(n,k)}: z^{(n,k)} = \rho_1 z^{(n,k-1)} + \rho_2\left(x^{(n)} + \beta^{(n-1)}\right) - c_2^{(n,k)} \end{cases} \tag{7}$$

Generic-ADMM-Net の構造を図 1 に示す. ADMM のn回目の反復はn番目のステージに対応し, そのステージには再構成層$\left(X^{(n)}\right)$, $Z^{(n)}$層, 乗数更新層$\left(M^{(n)}\right)$からなる層で構成される. さらに$Z^{(n)}$層は, 加算層$\left(A^{(n,k)}\right)$, 畳込み層$\left(C_1^{(n,k)}, C_2^{(n,k)}\right)$, 非線形変換層$\left(H^{(n,k)}\right)$からなる層

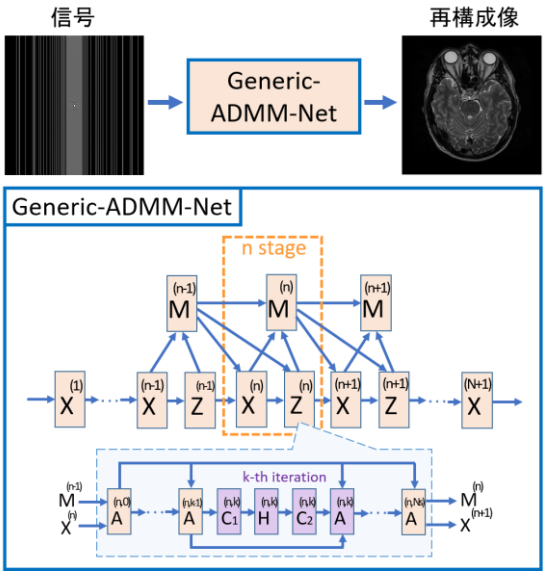


図 1　Generic-ADMM-Net の構造

により構成されており, 式(7)は図 1 の紫部に対応する. 先行研究[6]では, この Generic-ADMM-Net の入力を, 間引いた FT 信号とし, 出力を再構成された MR 画像としていた. そこで本研究では, 位相拡散フーリエ変換映像法の信号を入力とし, 出力を再構成 MR 画像とした場合の画像再構成について検討を行った.

3. 位相拡散フーリエ変換映像法

位相拡散フーリエ映像変換法(Phase Scrambling Fourier Transform imaging : PSFT)は, 位相エンコード方向勾配磁界に同期して 2 次関数状の磁界を一定期間印加するか, あるいは線形勾配磁界を印加のもと周波数変調された RF 励起パルスを照射するなどの方法により, 被写体内のスピンの位相を規則的に拡散してからフーリエ変換映像法の撮像シーケンスを実行する手法である[12][13]. $\rho(x,y)$をxy平面のスピン密度分布, bを 2 次関数上磁界の係数, τをその印加時間, γを核磁気回転比とすると, PSFT 信号は式(8)で与えられる. ここで $k_x = \gamma g_x t_x$, $k_y = \gamma g_y t_y$としており, g_xとg_yはそれぞれx方向とy方向の線形勾

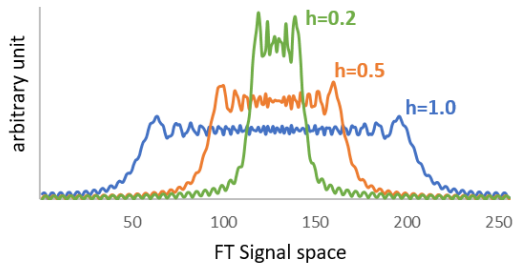


図 2　位相拡散係数hと信号拡がりの関係

配磁界の係数, t_xを読み出し用勾配磁界g_xの反転時間を考慮しない印加時間, t_yをg_yの印加時間とする.

$$v(k_x,k_y)=\iint_{-\infty}^{\infty}\{\rho(x,y)e^{-j\gamma b\tau(x^2+y^2)}\}e^{-j(k_xx+k_yy)}dxdy \tag{8}$$

PSFT 信号からの画像再構成には, 式(8)の信号を逆フーリエ変換した後, 与えた 2 次の位相変調項の逆位相を乗じればよい.

$$\rho(x,y)=e^{j\gamma b\tau(x^2+y^2)}F^{-1}v(k_x,k_y) \tag{9}$$

ここで, 撮像パラメータ$\gamma b\tau$は, 画像空間全体で標本化定理を満足する値 $\overline{\gamma b\tau}$ （$\pi/N\Delta x^2$: Nはデータ数, Δxは画素幅）を基準として使用し, 新たに位相拡散係数 h を導入して $h\overline{\gamma b\tau}$として表現することにする[16]. 256 点のデータ列の中央に 200 点の幅をもつ矩形密度分布を置き, 位相拡散係数hを 0.2, 0.5, 1.0 とした場合の PSFT 信号の分布を図 2 に示す. h が大きいほど信号が広い領域に分布する形になる. ここで, h が 0 の場合はフーリエ変換映像法に相当する.

汎用的なフーリエ変換映像法において信号をランダムに間引き収集すると, この間引きにより生じるアーティファクトはランダム雑音状になり, 除去しやすくなる. 一方 PSFT 法では 2 次関数状の位相変調と再構成における復調操作により, 信号間引きに伴う誤差は画像空間と信号空間で拡散する. このため, PSFT 法では信号空間の間引きは必ずしもランダムである必要はなくなる[16].

4.　PSFT 信号を入力とした Generic-ADMM-Net

PSFT 信号を入力とする Generic-ADMM-Net を構成した. 間引きされた PSFT 信号を入力とした場合の Generic-ADMM-Net の$X^{(n)}$層の処理を式(10)に変更する.

$$X'^{(n)}:x^{(n)}=Q^{\mathrm{T}}(U^{\mathrm{T}}U+\mu I)^{-1}\left\{U^{\mathrm{T}}s+\mu Q\left(z^{(n-1)}-\beta^{(n-1)}\right)\right\} \tag{10}$$

ここで, sを間引きされた PSFT 信号, Qを画像空間から PSFT 信号空間への変換（式(8)）, Q^{T}をその逆変換（式(9)）とする. この式(10)は, 従来の$X^{(n)}$層のフーリエ変換FをQ, 逆フーリエ変換F^{T}をQ^{T}, 間引き FT 信号yを間引き PSFT 信号sに変更したものとなる. また, この式は順伝播の処理変更であるため, 逆伝播においても同様の処理変更を行う. 本研究では, 上記のように処理を変更したネットワークを用いて, 間引きされた PSFT 信号を入力とした場合の深層学習再構成について検討した.

5.　画像再構成シミュレーション

本研究で用いる Generic-ADMM-Net における学習条件は, エポック数を 300, ステージ数nを 10, 反復回数kを 1, 畳込みフィルタのサイズと数を 5×5, 128 とし, パラメータ更新手法は Adam[17]とした. MR 画像において, 画像上の空間的な位相変化が小さく, 位相を高い精度で推定できる場合, 位相補正処理を行うことで被写体は実関数とみなせる. そこで今回は, 深層学習再構成時に被写体が実関数という拘束条件を使用する場合で画像再構成を実施した. 学習とテストに使用した MR 画像は, 一般的に公開され, 使用が許可されている IXI Dataset[18]に含まれる,

表 1　計算機環境

CPU	Intel Core i7-8700 (3.2GHz)
GPU	NVIDIA GeForce GTX1070Ti
RAM	32GB (DDR4-2400)
Software	MATLAB R2017b CUDA Toolkit 9.0 / cuDNN 7.0.5

256×256 画素の頭部の MR 画像を使用した. IXI Dataset には, 被験者 600 人分の MR 画像が含まれており, 本研究ではその中の T_2 強調像を使用した. Generic-ADMM-Net に入力する信号は, フーリエ変換映像法と位相拡散フーリエ変換映像法の信号を MR 画像より数値的に合成し, 信号収集点分布に従い間引き処理を行った信号を使用した. 間引き処理では, 臨床で使用されることの多いカルテシアン座標系を用いて検討した. 本研究で使用した計算機環境を表 1 に示す. 学習には GPU を用いた. 再構成像の評価方法には, ピーク信号対雑音(Peak Signal-to-Noise Ratio : PNSR), および両画像間で対応する局所領域内での輝度値やコントラスト, 構造の相似性を評価する SSIM(Structural SIMilarity index)[19]を使用した.

信号収集点分布を図 3 に示す. ここで, 信号収集点分布は中央 30 点を連続的に収集し, 信号収集量を 33%としたランダム間引きとしている. $N = 256$, $\Delta x = 0.1$ cm とするとき, $\overline{\gamma b \tau} = 1.227$ rad/cm^2 であり, h を 0.1 から 1.0 まで 0.1 刻みで変えて求められた PSNR と SSIM を図 4 に示す. それぞれ, 25 種類の MR 画像を用いて学習を行った. ミニバッチサイズを 5 とした. 学習にはそれぞれ約 7.5 時間を要した. テストには, 学習で用いた画像と重複の無い 20 種類の MR 画像を用いており, 図 4 はテスト画像の PSNR, SSIM の平均値を示している. 図 4 の結果から, 位相拡散係数 h を大きくするほど良好な結果を示しており, h を 1.0 としたときに最も高い PSNR と

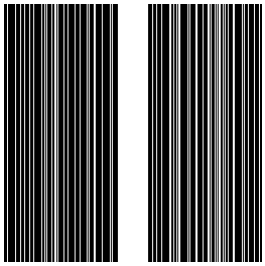


図 3　信号収集点分布 (信号収集量 : 33%, 中央 30 点を連続的に収集)

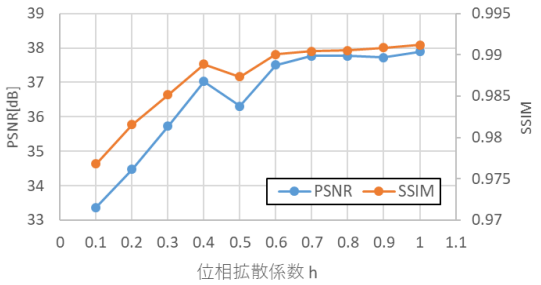


図 4　位相拡散係数hと PSNR, SSIM の関係

SSIM を記録した. これは, 信号収集量を変えた場合も同様であった. よって, 本研究では位相拡散係数 h として 1.0 を使用した.

FT 信号を入力とした場合の画像再構成と比較を行った. このとき, 学習には 100 種類, テストには 20 種類の MR 画像を用いた. また, ミニバッチサイズを 20 とした. 学習に要する時間は, FT 信号を入力とした学習は約 29 時間, PSFT 信号を入力とした場合は約 30 時間を要した. さらに, 従来の反復的な手法として ADMM を利用した反復的解法である C-SALSA-B[20]を使用した再構成を行った. ここで, C-SALSA-B(以降, 反復再構成と称する)ではスパース化関数にはウェーブレット変換(Daubechies, N=3)を適用した. PSFT 信号, FT 信号を入力とした Generic-ADMM-Net, PSFT 信号を入力とした反復再構成による PSNR と SSIM を表 2 に示す. この結果より, PSFT 信号を入力とした Generic-ADMM-Net の方が高い PSNR と SSIM が得られた. 図 5 に再構成像の例を示す. なお, コントラストをわかりやすくするための調整を行っている. 図 5 における拡大画

像から, 反復再構成による再構成像にはアーティファクトが残留した. PSFT 信号と FT 信号を入力とした Generic-ADMM-Net 同士の比較では, 赤枠で示す部分のように, PSFT 信号を入力とした場合の方が目標とする画像の特徴を多く有している.

画像 1 枚当たりの再構成に要する時間を比較すると, Generic-ADMM-Net において, 表 2 での FT 信号を入力とした場合の再構成時間は CPU で約 2.77 秒（GPU で約 0.98 秒）, PSFT 信号を入力とした場合は CPU で約 2.81 秒（GPU で約 1.02 秒）と同程度の再構成時間であった. また, 表 3 における反復再構成の再構成時間は CPU で約 10.66 秒であることから, Generic-ADMM-Net を用いた方が短時間で画像再構成を行えることが示された.

図 6 の中央を連続的に収集しない場合の信号収集点分布に対して画像再構成シミュレーションを実施した. ここで, 図 6(a)は全体をランダムに間引いており, 信号収集量を 33%としている. 図 6 (b)は信号収集比を 1/3(33%)として等間隔に間引いている. また, 位相拡散係数 h は 1.0 とした. 学習・テスト画像は, 図 3 の信号収集点分布を使用した場合と同様に 100 種類の学習画像, 20 種類のテスト画像を使用した. 間引きされた PSFT 信号を入力とした場合の PSNR と SSIM の平均を表 3 に示す. 図 3 の信号収集点分布を用いた場合よりも, 高い PSNR と SSIM が得られた. 図 6 の信号収集点分布 2 種の比較においては, 等間隔間引きを用いた方が高い値が得られた. 図 7 に図 6 の信号収集点分布を用いた場合の再構成像の例を示す. 図 7 の拡大画像から,ランダム間引きよりも等間隔間引きを用いた場合の方が赤矢印で示した部分のコントラストの表現が比較的良好であった.

表 2　図 3 の信号収集点分布を用いた場合の PSNR と SSIM

再構成法	PSNR[dB]	SSIM
PSFT-Net	39.01	0.9927
FT-Net	32.48	0.9718
PSFT-反復再構成	35.09	0.9833

※表中は[(入力信号) - (再構成手法)]と示している. ここで Net は Generic-ADMM-Net の略である.

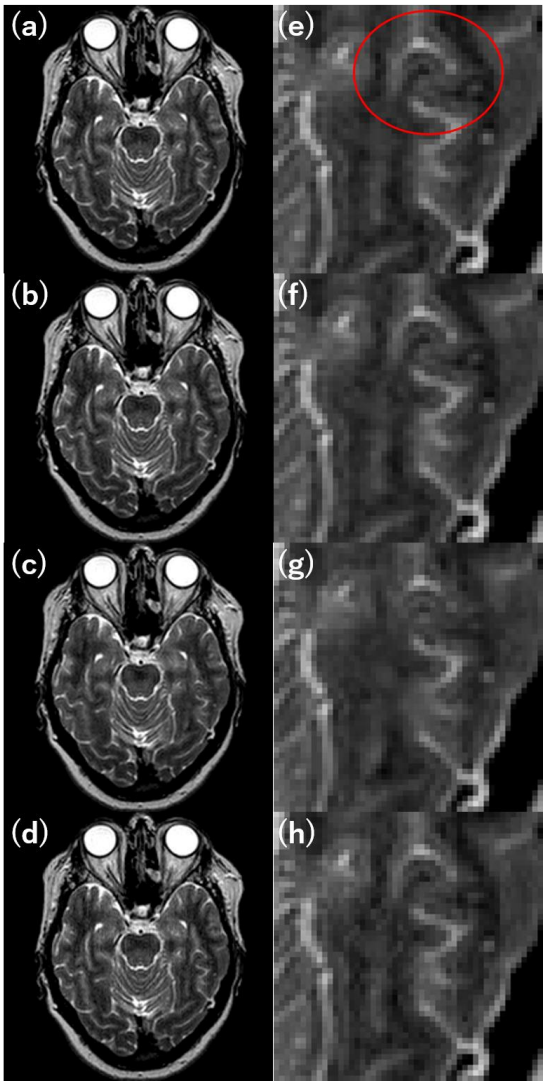
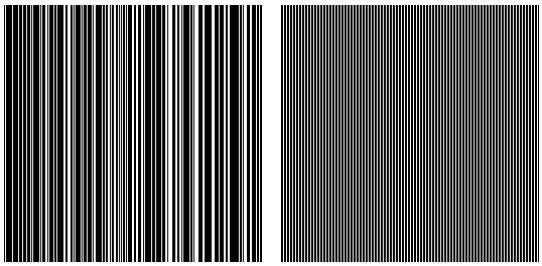


図 5　図 3 の信号収集点分布を用いた場合の再構成像の比較 : (a) 目標画像, (b) PSFT 信号を入力とした Generic-ADMM-Net, (c) FT 信号を入力とした Generic-ADMM-Net, (d) PSFT 信号を入力とした反復再構成, (e)～(h) それぞれ(a)～(d)の拡大画像



(a) ランダム間引き　　(b) 等間隔間引き

図 6　中央を連続的に収集しない場合の信号収集点分布 (信号収集量 : 33%)

6. 考察

図 4 より位相拡散係数 h の値を 0.5 とした場合は, h =0.4 よりも PSNR と SSIM ともに減少したが, hの値を大きくするほど PSNR と SSIM 共に高い値となる傾向があった. また, 表 2 より FT 信号を入力した場合よりも, PSFT 信号を入力とした場合の方が高い PSNR と SSIM が得られた. PSFT 法に含まれる 2 次の位相変調は信号間引きを行った場合に再構成像の誤差を拡散する効果があるが, これは Generic-ADMM-Net による深層学習再構成においても有効であると考える.

図 6 の中央を連続的に収集しない信号収集点分布を使用した場合では, 表 3 に示すように図 6 の信号収集点分布を使用した方が, 表 2 に比べ高い PSNR と SSIM が得られた. 一般的に, FT 信号は信号エネルギーが信号空間の中央部に集中する性質があるが, PSFT 法の信号は図 2 で示すように, hの値を大きくすると信号空間の広い領域に分布する形となる. よって, Generic-ADMM-Net においては, 局所的な信号を有した信号空間よりも, 分散され広い領域に分布した信号を, 図 6 の信号収集点分布のように一様に間引いた方が信号の推定がしやすく, かつ, 比較的高周波成分を多く収集しているため, 図 7 に示すように画像の細部構造を復元できた画像を再構成できたと考える.

図 6(a)のランダム間引きと(b)の等間隔間引き

表 3　図 6 の信号収集点分布を用いた場合の PSNR と SSIM

信号収集	PSNR[dB]	SSIM
ランダム間引き	39.68	0.9935
等間隔間引き	44.73	0.9978

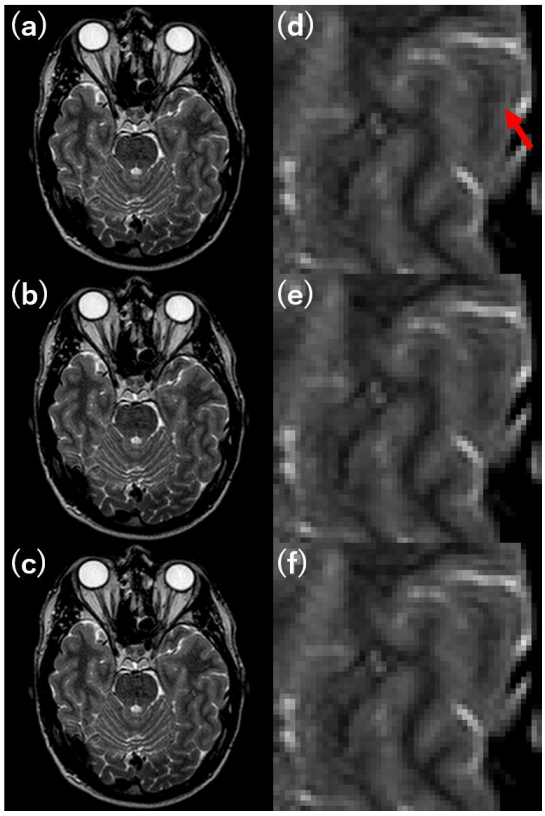


図 7　図 6 の信号収集点分布を用いた場合の再構成像の比較 : (a) 目標画像, (b) ランダム間引き, (c) 等間隔間引き, (d)～(f) それぞれ(a)～(c)の拡大画像

の比較では, 表 3 より等間隔間引きを使用した方が高い PSNR と SSIM を記録した. PSFT 法は, 等間隔間引きを使用した場合において, 2 次関数状の位相変化は, 再構成誤差を拡散する効果があり, アーティファクトのコヒーレント性の低下に寄与する. さらに, ランダム間引きでは, 信号の収集密度が局所的に低下することがあるが, 等間隔間引きでは信号空間の低域から高域まで等密度で信号を収集できるので, 再構成に関わる誤差を軽減できたと考える. 以上の 2 点より,

等間隔間引きを使用した方がコントラストや構造的な特徴の復元性能が高くなると考える.

本法は被写体が実関数であるという拘束条件を使用した場合で検討した. この拘束条件を使用しない位相を含む画像の場合は, アーティファクトの除去性能が低下する. これに対し, 等間隔で間引いた PSFT 信号からの画像再構成を行う研究[16]では, PSFT 信号をフレネル変換信号に変形し, その空間で線形補間を行った信号から初期画像を生成し, 反復的な処理を行う方法により画像再生に成功している. この手法を今回の深層学習再構成に利用すれば, 実関数という拘束条件を使用しない場合においても画像再構成を行える可能性がある. 位相画像の再構成については, 今後の課題としたい.

7. まとめ

本研究では, 位相拡散フーリエ変換映像法を利用した, Generic-ADMM-Net による深層学習再構成についての検討を行った. 汎用のフーリエ変換映像法による信号を入力する方法に比べ, 画質の精鋭度, 構造保存性, コントラスト保存性の点でも優れた画像が得られた. また, ランダム間引きに比べ等間隔信号間引きを使用した方が高い PSNR と SSIM が得られることが示された. 以上から, 位相拡散フーリエ映像法を利用した圧縮センシングは深層学習再構成によっても高画質化が図られることが示され, そして, 反復再構成よりもさらに高画質化される可能性が示された.

謝 辞

本研究の一部は, 科学研究費助成金 (19K04423), および柏森情報科学振興財団の助成により実施された. また, 本研究を遂行する際に利用した IXI Dataset を提供する Information eXtraction from Images プロジェクトに感謝の意を表します.

利益相反の有無

なし

文 献

[1] Donoho DL: Compressed sensing. IEEE Transactions on Information Theory **52**: 1289-1306, 2006

[2] Lustig M, Donoho DL, Santos JM, et al.: Compressed Sensing MRI. IEEE Signal Processing Magazine **25**: 72-82, 2008

[3] Glowinski R and Marrocco A: Sur l'approximation, par éléments finis d'ordre un, et la résolution, par pénalisation-dualité, d'une classe de problèms de Dirichlet non linéares. ESAIM: Mathematical Modeling and Numerical Analysis-Modélsation Mathématique et Analyse Numérique, 9: 41-76, 1975

[4] Wang Y, Yin W, Zeng J: Global Convergence of ADMM in Nonconvex Nonsmooth Optimization. arXive:1511.06324, 2015

[5] Wang S, Su Z, Ying L, et al.: Accelerating magnetic resonance imaging via deep learning. IEEE 13th International Symposium on Biomedical Imaging, Prague, 2016, pp514–517

[6] Yang Y, Sun J, Li H, et al.: ADMM-Net: A Deep Learning Approach for Compressive Sensing MRI. arXiv:1705.06869, 2017

[7] Han YS, Lee D, Yoo J, et al.: Accelerated Projection Reconstruction MR imaging using Deep Residual Learning. 25th Annual Meeting of International Society for Magnetic Resonance in Medicine, Hawaii, O690, 2017

[8] Yang Y, Sun J, Li H, et al.: ADMM-CSNet: A Deep Learning Approach for Image Compressive Sensing. IEEE Transactions on Pattern Analysis and Machine Intelligence **42**: 512-538, 2018

[9] Hennig J, Welz AM, Schultz G, et al.: Parallel imaging in non-bijective, curvilinear magnetic field gradients: a concept study. MAGMA **21**: 5-14, 2008

[10] Ito S, Y Yamada: Alias-free image reconstruction using Fresnel transform in the phase-scrambling Fourier imaging technique. Magnetic Resonance in Medicine **60**: 422-430, 2008

[11] Stockmann JP, Ciris PA, Galiana G, et al.: O-Space Imaging: Highly Efficient Parallel Imaging

Using Second-Order Nonlinear Fields as Encoding Gradients with No Phase Encoding. Magnetic Resonance in Medicine **64**: 447-456, 2010

[12] Maudsley AA: Dynamic range improvement in NMR imaging using phase scrambling. Journal of Magnetic Resonance **76**: 287-305, 1988

[13] Wedeen VJ, Chao YS, Ackerman JL: Dynamic range compression in MRI by means of a nonlinear gradient pulse. Magnetic Resonance in Medicine **6**: 287-295, 1988

[14] Ito S, Yasaka S, Yamada Y: MR Image Reconstruction of a Regularly Undersampled Signal using Quadratic Phase Scrambling. IEEE International Conference on Image Processing (ICIP2015), TEC-P26.2, Quebec City, Canada, 2015

[15] Wiaux Y, Puy G, Gruetter R, et al.: Spread Spectrum for Compressed Sensing Techniques in Magnetic Resonance Imaging. IEEE Intern ational Symposium on Biomedical Imaging, 756-759, Rotterdam, The Netherlands, 2010

[16] 八坂俊吾, 伊藤聡志: 位相拡散フーリエ変換映像法を利用した等間隔間引きを基本とするMR圧縮センシング. 電子情報通信学会論文誌 **J101-D**: 27-35, 2018

[17] Kingma DP, Ba J: Adam: A Method for Stochastic Optimization. arXiv:1412.6980, 2017

[18] IXI Dataset. https://brain-development.org/ixi-dataset/

[19] Wang Z, Bovik AC, Sheikh HR, et al.: Image quality assessment: from error visibility to structural similarity. IEEE Transactions on Image Processing **13**: 600-612, 2004

[20] Liu Y, Cai JF, Zhan Z, et al.: Balanced Sparse Model for Tight Frames in Compressed Sensing Magnetic Resonance Imaging. PLoS ONE **10**: 1-19, 2015

Image Quality Improvements by adopting MR Phase Scrambling Fourier Transform Imaging in Deep Learning Reconstruction

Shun UEMATSU[*1], Satoshi ITO[*1]

*1 Graduate School of Regional Development and Creativity, Utsunomiya University

Deep learning has received much attention because of its excellent performance in image quality and reconstruction time. It was reported that image quality can be improved in MR compressed sensing (CS) by using the phase scrambling Fourier transform imaging (PSFT) that use nonlinear phase encoding. In this paper, an image reconstruction using Generic-ADMM-Net as a CNN image reconstruction was examined. Simulation studies showed that sharpness, structural similarity and image contrast were all improved compared to standard Fourier transform based CS. These studies indicate that PSFT has the possibility to reconstruct higher quality images in deep learning image reconstruction as well as iterative reconstruction.

Key words: Compressed Sensing, Phase Scrambling, ADMM, Deep Learning, MRI

数値シミュレーションを用いた非等方ボクセルに対応した粘弾性率分布推定手法の定量性評価

武田　将岳*1　藤原　宏志*2　菅　幹生*1,3

要旨

磁気共鳴エラストグラフィ（magnetic resonance elastography: MRE）は，組織の粘弾性率を定量的に測定可能な非侵襲的手法である．撮像対象外部より振動を加え，内部に伝播する弾性波を画像化し，弾性波画像に対して逆問題を解くことで粘弾性率を推定する．ノイズに頑健な解析手法として 3 次元積分型公式（integral type reconstruction formula: ITRF）がある．ハードウェアの制約などによって弾性波画像を非等方ボクセルでしか取得できない場合には，等方ボクセルを仮定した ITRF を適用できない．本研究では，非等方ボクセルに対応できるように改良した ITRF（anisotropic ITRF: A-ITRF）の定量性を数値シミュレーションにより評価することを目的とした. 均一弾性体を想定した非等方ボクセルの弾性波画像を作成し, A-ITRF のノイズや横波伝播方向と定量性の関係を評価した．実測と同程度のノイズや横波伝播方向によらず，粘弾性率推定誤差は 0.1%以下，変動係数は 1.0%以下となった．非等方ボクセルデータに対しても A-ITRF の利用により高い定量性で粘弾性率が推定可能であることが示唆された.

キーワード：磁気共鳴エラストグラフィ，粘弾性率分布推定手法，3 次元積分型公式，非等方ボクセル

1. はじめに

疾患や機能障害と生体内組織の硬さには相関があり，腫瘍や肝線維症などの病変部は正常組織に比べて硬くなることが知られている[1]．このことから，生体内組織の粘弾性率を定量的に測定することは診断において有用である. 核磁気共鳴画像装置（magnetic resonance imaging: MRI）を用いた非侵襲的かつ定量的な粘弾性率測定法に MR elastography（MRE）がある[2]. MRE は加振装置を用いて体表から振動を与え，測定対象に伝播する弾性波を画像化し，逆問題を解くことで粘弾性率分布を推定することができる.

解析手法に 3 次元積分型公式（integral type reconstruction formula: ITRF）がある[3,4]. ITRF は，ノイズに頑健であるが，等方ボクセルを仮定しているため，ハードウェアの制約などにより，非等方ボクセルで撮像した弾性波画像には適用できない．そこで，非等方ボクセルに対応した ITRF（anisotropic ITRF: A-ITRF）を開発した.

本研究では，実測データを模擬した数値シミュレーションによる弾性波画像に対して A-ITRF を適用し, 粘弾性率推定の定量性を評価することを目的とした.

2. 弾性率分布推定手法

2.1 ITRF

Voigt 型線形粘弾性体を想定した運動方程式に対して，物質の等方性と非圧縮性を仮定すると式(1a)(1b)が導かれる.

$$-\rho\omega^2\phi - \mu\Delta\phi - \omega\eta\Delta\psi = 0 \qquad (1a)$$

$$-\rho\omega^2\psi - \mu\Delta\psi + \omega\eta\Delta\phi = 0 \qquad (1b)$$

ここで，μは剛性率，ηは粘性，ρは物質密度，ωは外部加振の角周波数，ϕとψは位相が$\pi/2$異なる弾性波画像を表し，貯蔵弾性率$G' = \mu$,

*1 千葉大学 大学院融合理工学府
〔〒263-8522 千葉市稲毛区弥生町 1-33〕
e-mail: affa7027@chiba-u.jp
*2 京都大学 大学院情報学研究科
*3 千葉大学　フロンティア医工学センター

損失弾性率$G'' = \omega\eta$で求められる．

ノイズの影響を低減するため，式(2)を満たすテスト関数と式(3)に示す Green の公式を適用することで，弾性波画像への微分を避けた 3 次元積分型公式(4)を導くことができる．

$$v|_{\partial D} = \frac{\partial v}{\partial n_{\partial D}} = 0 \tag{2}$$

D はテスト関数が定義される領域を表す．

$$\int_D \Delta f(x)v(x)dx = \int_D f(x)\Delta v(x)dx \tag{3}$$

ここで

$$\Phi = \int_D \phi(\Delta v)dx \quad \Phi_0 = \int_D \phi v dx$$

$$\Psi = \int_D \psi(\Delta v)dx \quad \Psi_0 = \int_D \psi v dx$$

とすると，

$$\begin{pmatrix}\mu\\ \eta\end{pmatrix} = \frac{-\rho\omega}{\Phi^2+\Psi^2}\begin{pmatrix}\omega\Phi & \omega\Psi\\ \Psi & -\Phi\end{pmatrix}\begin{pmatrix}\Phi_0\\ \Psi_0\end{pmatrix} \tag{4}$$

ここでは，式(2)を満たす滑らかなテスト関数vとして，式(5)に示す三角関数型を用いた．

$$v(r) = \begin{cases}1 + cos\pi\dfrac{r}{R} & r < R\\ 0 & r \geq R\end{cases} \tag{5}$$

ただし，3 次元座標軸は$x = (x_1, x_2, x_3)$，$D = \{{x_1}^2 + {x_2}^2 + {x_3}^2 < R^2\}$である．

2.2 A-ITRF

3 次元の弾性波画像で位相差が$\pi/2$のものを$\vec{\phi} = (\phi_1, \phi_2, \phi_3)$，$\vec{\psi} = (\psi_1, \psi_2, \psi_3)$とする．これらが満たす 3 次元の Voigt 型粘弾性方程式に対して，式(2)を満たすテスト関数を用いて Green の公式を適用した 3 次元積分型公式を式(6)(7)に示す．

$$\begin{aligned}-\rho\omega^2\int_D \phi_i v dx = {}& \mu\left\{\int_D \left(\phi_i\Delta v + \vec{\phi}\cdot\nabla\frac{\partial v}{\partial x_i}\right)dx\right\}\\ &+ \lambda\left\{\int_D \vec{\phi}\cdot\nabla\frac{\partial v}{\partial x_i}dx\right\}\\ &+ \omega\eta\left\{\int_D \left(\psi_i\Delta v + \vec{\psi}\cdot\nabla\frac{\partial v}{\partial x_i}\right)dx\right\}\\ &+ \omega\zeta\left\{\int_D \vec{\psi}\cdot\nabla\frac{\partial v}{\partial x_i}dx\right\}\end{aligned} \tag{6}$$

$$\begin{aligned}-\rho\omega^2\int_D \psi_i v dx = {}& \mu\left\{\int_D \left(\psi_i\Delta v + \vec{\psi}\cdot\nabla\frac{\partial v}{\partial x_i}\right)dx\right\}\\ &+ \lambda\left\{\int_D \vec{\psi}\cdot\nabla\frac{\partial v}{\partial x_i}dx\right\}\\ &- \omega\eta\left\{\int_D \left(\phi_i\Delta v + \vec{\phi}\cdot\nabla\frac{\partial v}{\partial x_i}\right)dx\right\}\\ &- \omega\zeta\left\{\int_D \vec{\phi}\cdot\nabla\frac{\partial v}{\partial x_i}dx\right\}\end{aligned} \tag{7}$$

ただし，$i = 1,2,3$である．λはラメ係数，ζは体積粘性率である．ここで，

$$\Delta v = \frac{\partial^2 v}{\partial x_1^2} + \frac{\partial^2 v}{\partial x_2^2} + \frac{\partial^2 v}{\partial x_3^2},\ v_{,i} = \frac{\partial v}{\partial x_i},$$

$$\nabla v_{,i} = \left(\frac{\partial^2 v}{\partial x_i \partial x_1}, \frac{\partial^2 v}{\partial x_i \partial x_2}, \frac{\partial^2 v}{\partial x_i \partial x_3}\right)$$

さらに，

$$\Phi_0^{(i)} = \int_D \phi_i v dx \qquad \Psi_0^{(i)} = \int_D \psi v dx$$

$$\Phi^{(i)} = \int_D \phi_i \Delta v dx \qquad \Psi^{(i)} = \int_D \psi_i \Delta v dx$$

$$\Phi_i = \int_D \phi \cdot \nabla v_{,i}\, dx \quad \Psi_i = \int_D \psi \cdot \nabla v_{,i}\, dx$$

とすると，式(6)(7)は連立一次方程式

$$\begin{pmatrix}\Phi^{(1)}+\Phi_1 & \Phi_1 & \omega\left(\Psi^{(1)}+\Psi_1\right) & \omega\Psi_1\\ \Phi^{(2)}+\Phi_2 & \Phi_2 & \omega\left(\Psi^{(2)}+\Psi_2\right) & \omega\Psi_2\\ \Phi^{(3)}+\Phi_3 & \Phi_3 & \omega\left(\Psi^{(3)}+\Psi_3\right) & \omega\Psi_3\\ \Psi^{(1)}+\Psi_1 & \Psi_1 & -\omega\left(\Phi^{(1)}+\Phi_1\right) & -\omega\Phi_1\\ \Psi^{(2)}+\Psi_2 & \Psi_2 & -\omega\left(\Phi^{(2)}+\Phi_2\right) & -\omega\Phi_2\\ \Psi^{(3)}+\Psi_3 & \Psi_3 & -\omega\left(\Phi^{(3)}+\Phi_3\right) & -\omega\Phi_3\end{pmatrix}\begin{pmatrix}\mu\\ \lambda\\ \eta\\ \zeta\end{pmatrix} = -\rho\omega^2\begin{pmatrix}\Phi_0^1\\ \Phi_0^2\\ \Phi_0^3\\ \Psi_0^1\\ \Psi_0^2\\ \Psi_0^3\end{pmatrix} \tag{8}$$

と表すことができる．式(8)は，未知数 4 個，方程式 6 個の過剰決定系であるため，最小二乗法を用いて解を求めることができる．本論文ではこれを 3 次元積分型公式と呼ぶ．特に領域 D での積分を非等ボクセルに分割して適用する場合が A-ITRF に相当する．

2.3 局所波長推定手法（local frequency estimation: LFE）

LFE は，フーリエ変換を用いた周波数成分解析に窓関数を適用することで局所的な周波数解析ができる[5]．窓関数wを用いたフーリエ変換を式(9)に示す．

$$G(x,u)=\int_{-\infty}^{\infty} w(x-x')f(x')e^{-iyt'}dx' \qquad (9)$$

また，測定対象に粘性がないことを仮定すると剛性率は周波数f，波長λと式(10)の関係があり[6]，局所的な周波数解析により求めた波数$2\pi/\lambda$より，弾性と粘性が合わさった硬度を求めることができる．ただし，硬度から弾性と粘性に分離して求めることはできない．

$$\mu=\rho(f\lambda)^2 \qquad (10)$$

3. 方法

数値シミュレーションでは，均一弾性体（3 kPa，8 kPa）に対して，周波数 62.5 Hz での加振により平面波が伝播することを仮定して弾性波画像を生成した．振幅とノイズは実測データと同程度に設定し，減衰はないものとした．また，横波の進行方向の角度を$\pi/8$，$\pi/4$，$3\pi/8$の 3 方向で検討した．画像サイズ128 mm × 128 mm × 45 mm，等方ボクセルの場合のボクセルサイズは1 mm × 1 mm × 1 mm, 非等方ボクセルの場合のボクセルサイズは1 mm × 2 mm × 3 mmとした．解析結果の比較に，等方ボクセルのみ適用可能な ITRF，非等方ボクセルも適用可能な LFE を用いた．

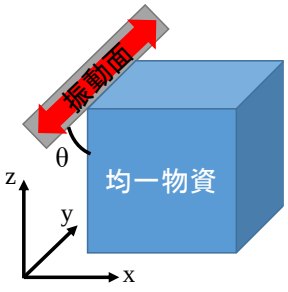


図 1　数値シミュレーション概念図

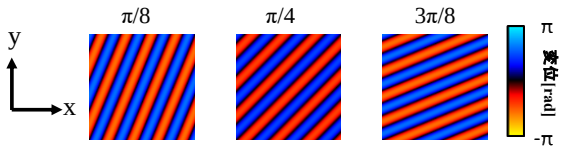


図 2　弾性波画像

4. 結果

ノイズを付加した等方ボクセルデータに対して，A-ITRF と ITRF，LFE を適用した場合の推定剛性率の誤差と変動係数をそれぞれ表 1 と表 2 に示す．同様にノイズを付加した非等方ボクセルデータに対して A-ITRF と LFE を適用した結果を表 3 と表 4 に示す．A-ITRF は ITRF と同程度の推定精度を有していることを確認し，A-ITRF は，ボクセルの形状，伝搬方向に関係なく，推定誤差は 0.1%以下，変動係数は 1.0%以下となった．

表 1 推定剛性率の誤差（等方ボクセル）

真値	推定手法	π /8	π /4	3 π /8
3 kPa	A-ITRF	0.0%	0.0%	0.0%
	ITRF	0.0%	0.0%	0.0%
	LFE	-1.0%	-0.5%	-0.7%
8 kPa	A-ITRF	-0.0%	-0.0%	0.1%
	ITRF	-0.0%	0.1%	0.0%
	LFE	-10.3%	-8.6%	-10.3%

表 2 推定剛性率の変動係数（等方ボクセル）

真値	推定手法	π /8	π /4	3 π /8
3 kPa	A-ITRF	0.2%	0.3%	0.2%
	ITRF	0.2%	0.2%	0.2%
	LFE	2.2%	2.0%	2.2%
8 kPa	A-ITRF	0.3%	0.3%	0.3%
	ITRF	0.3%	0.2%	0.2%
	LFE	6.3%	6.6%	6.2%

表 3 推定剛性率の誤差（非等方ボクセル）

真値	推定手法	π /8	π /4	3 π /8
3 kPa	A-ITRF	-0.1%	-0.1%	0.1%
	LFE	0.1%	-0.0%	-0.1%
8 kPa	A-ITRF	0.1%	-0.1%	0.1%
	LFE	-0.8%	-0.6%	-0.5%

表 4 推定剛性率の変動係数（非等方ボクセル）

真値	推定手法	π /8	π /4	3 π /8
3 kPa	A-ITRF	0.7%	0.7%	0.6%
	LFE	0.8%	0.8%	1.1%
8 kPa	A-ITRF	1.0%	0.7%	0.7%
	LFE	1.8%	1.4%	1.4%

5. 考察と今後の課題

剛性率についてノイズや伝搬方向によらず，A-ITRF は高い定量性を有していることが示

唆された．A-ITRF では，損失弾性率について推定することも可能である．そのため，粘性を考慮した数値シミュレーションを行い，損失弾性率について定量性を検討する必要がある．

6. まとめ

本研究では，非等方ボクセルに対応した粘弾性率分布推定手法である A-ITRF を開発し，数値シミュレーションを用いて定量性を評価した．A-ITRF はボクセルの形状，弾性波の伝播方向によらず高い定量性を示した．

利益相反の有無

なし

文　献

［1］ 菅幹生，他，MR elastography 法　MRI を用いた定量的触診，映像情報 Medical，MRI 臨時増刊号，Vol.35，No.15，p.96-100，2003

［2］ A. Manduca et.al.，Magnetic resonance elastography: Non-invasive mapping of tissue elasticity, Medical Image Analysis, 5, 237-254, 2001

［3］ 藤原宏志，“MRE にもとづく弾性係数決定逆問題の数値計算の問題点,” 日本応用数理学会, pp. 14-15, 2007

［4］ 藤原宏志，“積分型公式による MRE での剛性率の再構成,” 第 57 回理論応用力学講演会論文集，pp. 427-428, 2008

［5］ H. Knutsson, et.al., "Local multiscale frequency and bandwidth estimation," Proceedings of 1st International Conference on Image Processing, vol.1, pp. 36-40, 1994

［6］ 荒木力，エラストグラフィ徹底解説 生体の硬さを画像化する，秀潤社，2011

U-Net による X 線透視血管造影像からの血管可視化技術の開発：時空間情報を利用した初期的検討

木村　竜誠[*1]　大野　智之[*2]　寺本　篤司[*1]

齋藤　邦明[*1]　藤田　広志[*3]

要旨

X 線透視血管造影像から血管のみを可視化する技術として，Digital Subtraction Angiography（DSA）がある．DSA では，造影剤投与後の画像から投与前の画像を差分することで，骨や器具などを除去し，血管のみを可視化する．しかし，撮像中の患者の動きにより，差分が正しく行われず，診断精度が低下する問題がある．そこで U-Net を用いることで，血管造影像のみから血管を自動で可視化する技術を新たに開発する．本手法では，時空間情報を取り入れた U-Net を使用する．注目フレーム及びその前後 1 フレームの血管造影像を U-Net に入力することで，血管のみが可視化された注目フレームの DSA 像が出力される．
検証には,藤田医科大学ばんたね病院にて撮影された，実際の頭部 DSA 像を使用した．出力結果から，本手法により不要な骨成分などが除去され，血管のみが良好に可視化されることが確認された．

キーワード：DSA，時空間情報， U-Net, 血管造影，ミスレジストレーション

1. はじめに

現代医学において， digital subtraction angiography(DSA)は X 線画像から血管を可視化する強力な技術である[1-2]．この技術は脳血栓症や冠状動脈性心臓病などの血管疾患の診断や治療に重要な役割を果たす. DSA では，予め造影剤を投与する数秒前に画像を撮影しておき，造影剤投与後の画像からリアルタイムで差分する[3]．これにより，骨などの不要な成分が除去され, 血管のみが可視化される．しかし，差分する造影剤注入前後の画像には時間差がある. そのため, 患者の動きにより，造影剤注入前後の画像の位置に不整合が生じると，差分が正しく行われない．実際には，呼吸，嚥下，眼球運動，筋肉運動などの患者の運動は避けられない．これが原因でモーションアーチファクトが発生し，診断および治療精度の低下を引き起こす．この問題を解決するために，造影剤注入前後の画像の位置合わせの研究が多く行われている[4-7]．しかし多くの場合はモーションアーチファクトを完全に回避することはできず，現実的な解決には至っていない．そこで本研究では画像認識の分野において優れた精度を示す deep learning に注目した．我々は，差分を行わずに擬似的に DSA 像を得る手法を新たに“Virtual DSA”と定義し，これまでに deep learning の一種である U-Net[8]を用いて，Virtual DSA の開発を検討してきた．その検証の結果，画像差分なしに血管を可視化するこ

*1 藤田医科大学大学院保健学研究科〔〒470-1192 愛知県豊明市沓掛町田楽ヶ窪 1-98〕e-mail: teramoto @fujita-hu.ac.jp
*2 藤田医科大学ばんたね病院〔〒454-8509 愛知県名古屋市中川区尾頭橋三丁目 6 番 10 号〕
*3 岐阜大学工学部〔〒501-1193 岐阜県岐阜市柳戸 1-1〕

とができ，モーションアーチファクトを抑えることができた．しかしながら，血管造影検査では、血管を流れる造影剤により血管が可視化される。そのため、時間により血管の形状が異なり、時相によっては血管の可視化に不具合が生じるという問題があった。また挿入された金属器具を血管と誤認識することがあるという課題があった。そこで我々は、時間により，撮像される造影血管の変化が他の部位より大きいことに注目した。本研究では、時空間情報を利用することで、血管可視化技術の精度向上を目的とする

2. 方法

本手法の概要を図１に示す．本手法では U-Net に領域分割とパッチベース処理を加えた領域分割型パッチベース U-Net を使用する．造影剤投与後像を領域分割型パッチベース U-Net に入力することで，血管のみが可視化されるよう処理を行い，Virtual DSA 像を生成する．また入力は，時空間情報を利用するために，時間的に連続な 3 枚の造影剤投与後像が入力される．以下に本手法の詳細を述べる．

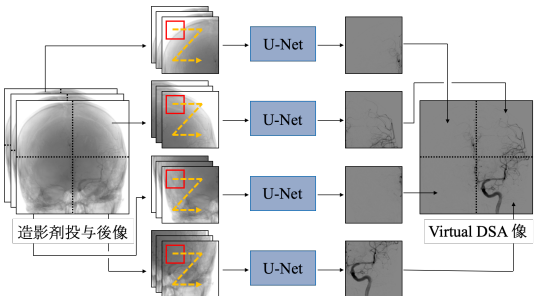


図１　提案手法の概要

2.1 ネットワーク構造

U-net モデルの概略図を図２に示す．U-Net はエンコーダー部とデコーダー部で構成されている．エンコーダー部のプーリングによりダウンサンプリングされた特徴マップを，デコーダー部でアップサンプリングする．Contracting path によりエンコーダー部とデコーダー部が連結されており，局所的特徴と全体的位置情報の両立が可能である．なお，本研究では 5 層構造の U-Net を使用した．また，U-Net の入力サイズは 128×128×3 である．

2.2 画像入力

入力画像として造影剤投与後像を領域分割型パッチベース U-Net に入力する．本手法では，時間による造影血管の変化など，時空間情報を利用するために，目的フレームの造影剤投与後像とその前後１フレームの合計３フレームの造影剤投与後像を入力する．

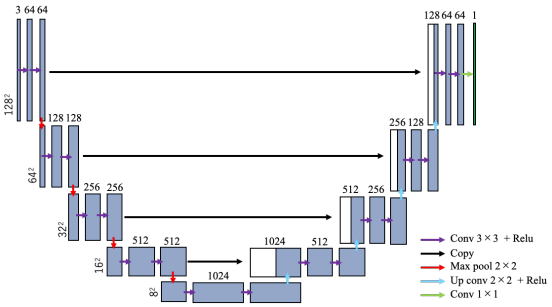


図２　U-Net

2.3 領域分割

本手法では，入力画像を垂直方向，水平方向の計４つの領域に分割して，それぞれに適した別々のネットワークで処理を行う．頭部血管造影像は，眼窩部や造影側の大局的な位置ごとに，血管の数や太さなどの特性が異なる．そのため，領域ごとに処理特性の異なる U-Net を使用する．

2.4 パッチベースの入力

本手法において，U-Net は小領域ごとに血管の可視化を行う．U-Net の処理は一般的に画像全体に適用される．しかし，頭部血管造影像は，眼窩や鼻腔などの構造が複雑で，X 線減弱の差が大きい．そのため，これらの構造と重なる血管の一部が不明瞭に出力されることがある．そこで小領域ごとに処理することで，より詳細な血管抽出を行う．時間的に連続した３フレームの造影剤投与後像から取り出した，注目領域近傍 128×128 画素の画像パッチを U-Net に入力する．

2.5　画像出力

入力された３フレーム分の画像パッチに対して U-Net の各層で処理を行う．入力の目的フレームに対して血管のみが可視化された注目領域の画像パッチが出力される．この処理をスライディングウィンドウ方式で画像全体に対して行う．出力された複数の画像パッチを結合することで，目的フレームの Virtual DSA 像が生成される．

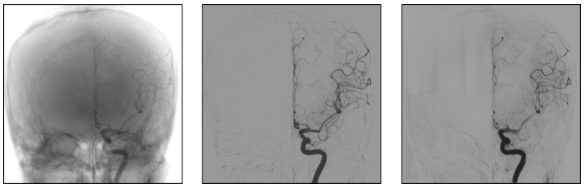
3．検証

3.1 検証方法

本手法の有効性を確認するために，７症例の頭部血管造影像の画像データを使用して血管抽出精度の評価を行なった．画像データは藤田医科大学ばんたね病院にて撮影された．撮影装置として Infinix Celeve-i INFX-8000V (Canon Medical Systems)を使用した. 撮像条件は，管電圧 89 kV，管電流 125 mA である．FOV は約 130 mm である. 造影剤注入条件は，カテーテル位置は内頸動脈で，注入速度が 4 ml/sec ，注入量が 6 ml である．血管抽出精度の評価には，出力画像と理想画像の正規化平均二乗誤差(NMSE)を算出した．

3.2 結果及び考察

出力結果の例を図 3 に示す．本手法により生成された virtual DSA 像は正確に血管が抽出されており, アーチファクトが少ない. NMSE は 22.9%であり，良好に血管が抽出されていることが確認された.また, 時空間情報を利用していないネットワークと精度を比較した結果，NMSE は 1.7%向上しており，時空間情報を利用することにより精度が向上することが確認された．



(a)入力画像　(b)理想画像　(c)出力画像

図３　出力結果

4．まとめ

本研究では，造影剤投与後像から画像差分なしで DSA 像を生成する Virtual DSA 像を提案した．領域分割型パッチベース U-Net により，骨などの不要な成分が除去され，血管のみが可視化された．また，時空間情報を利用することで，血管可視化の精度向上が見られ，本手法の有効性が示唆された．

利益相反の有無

なし

文　献

［1］ Mistretta CA, Crummy AB, Strother CM (1981) Digital angiography: a perspective. Radiology 139(2):273–276.

［2］ Brody WR (1982) Digital subtraction angiography. IEEE Sci29(3):1176–1180.

［3］ Takahashi M, Sato N, Fukui K, Kohrogi Y, Yamashita Y, Shinzato J, Higashida Y (1986) Hybrid digital subtraction angiography: Initial clinical experience. Computerized Radiology 10(4):147–154.

［4］ Oung H, Smith AM (1984) Real time motion detection in digital subtractive angiography. Proc. SPIE 0515, Medical Images and Icons 515:336–339.

［5］ Meijering EH, Zuiderveld KJ, Viergever MA (1999) Image registration for digital subtraction angiography. Int J Comput Vision 31:227–246.

［6］ Buzug TM, Weese J (1998) Image registration for DSA quality enhancement. Comput Med Imaging Graph 22(2):103–113

［7］ Lee S, Jeon CH, Sunwoo L, Oh DY, Lee KJ (2019) Phase-Based Nonrigid Deformation for Digital Subtraction Angiography.

［8］ Ronneberger O, Fischer P, Brox T (2015) U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical Image Computing and Computer-Assisted Intervention 2015 9351:234–241.

Development of Blood Vessel Visualization Technique from X-ray Fluoroscopic Angiogram by U-Net: Preliminary Study Using Spatiotemporal Information

Ryusei KIMURA[*1], Tomoyuki OHNO[*2], Atsushi TERAMOTO[*1], Kuniaki SAITO[*1],
Hiroshi FUJITA[*3]

*1 Graduate School of Health Sciences, Fujita Health University
*2 Fujita Health University Bantane Hospital
*3 Faculty of Engineering, Gifu University

Digital subtraction angiography (DSA) is a powerful technique for visualizing blood vessels from X-ray images. For image acquisition in DSA, pre-contrast images are subtracted from post-contrast images. However, the subtraction images obtained with this technique suffer from artifacts caused by patient motion. To avoid these artifacts, we developed a new technique for automatically visualizing blood vessels from only angiographic images by using U-Net. In this method, U-Net with spatiotemporal information is used. In the proposed method, DSA images only containing the extracted blood vessels were generated by inputting a series of post-contrast images. A series of post-contrast images included the target frame and one frame before and after it. For evaluation, an actual head DSA image taken at Fujita Health University Bantane Hospital was used. As a results, DSA images in which blood vessels were visualized were obtained by proposed method. These results indicate that the proposed method can visualize blood vessels from a series of post-contrast image.

Key words: Digital subtraction angiography, Spatiotemporal Information, U-Net, Angiography, Misregistration

深層学習を用いた圧縮センシング CT 画像再構成の画質改善

工藤　博幸[*1]　森　和希[*1]

要旨

トータルバリエーション（TV）正則化などの圧縮センシングを用いた CT 画像再構成は，画像を階段状の関数で近似する原理に基づくため 1) Staircase Artifact が発生すること，2) 滑らかな濃度変化やテクスチャーが消失しやすいこと，の問題点がある．本研究では，深層学習を用いて圧縮センシングの画質性能を向上させる新手法を提案する．提案手法では，フィルタ補正逆投影（FBP）法と TV や非局所 TV に代表される圧縮センシングの再構成画像を，2 入力 1 出力の画像変換を行う CNN に入力して，画質を改善した出力画像を得る．CNN の構造と圧縮センシングの正則化パラメータは，学習データを用いて 2 乗誤差が最小になるように自動決定する．スパースビューCT と低線量 CT の状況を想定して，腹部 CT 画像のデータセットを用いたシミュレーション実験を行った．深層学習単独の場合には血管など学習データから学びにくい構造物の消失が随所に見られ，TV や非局所 TV では Staircase Artifact や滑らかな濃度変化の発生が見られた．これに対して，両者を組み合わせた提案手法では，各々の手法の欠点を補い劇的な画質改善が実現できた．

キーワード： CT，画像再構成，圧縮センシング，深層学習，畳み込みニューラルネットワーク

1. はじめに

現状において，スパースビューCT や低線量 CT の画像再構成に最も有効な手法は圧縮センシング（Compressed Sensing, CS）だといわれているが，2016 年以降深層学習を用いた CT 画像再構成の研究が精力的に行われ論文数が世界的に急増している [1-8]．深層学習を画像再構成に用いる場合, まずフィルタ補正逆投影（FBP）法など既存の画像再構成法を用いてストリークアーティファクトや雑音で劣化した画像 **z** を求め，**y** を画像変換を行う畳み込みニューラルネットワーク（Convolutional Neural Network, CNN）に入力して最終的な再構成画像 **x** を求める [1,2]．即ち，CNN をアーティファクトや雑音を除去するポストフィルタとして使用する．そして, CNN の学習は多数の正解画像と劣化画像のペア（**x**,**y**）からなる学習データを用いて行う．深層学習が CS の性能を上回るかどうかは現状では分かっていないが，問題点として学習データに含まれていない構造物や個人差が大きく学習データから学びにくい構造物（例えば、病変や低コントラストの血管など）を消すように動作することが知られている．一方, CS では学習データを用いず各々の測定データに合致するように忠実に画像復元を行うため，構造物を消すタイプの画質劣化は比較的起こりにくいことが知られる．

*1 筑波大学システム情報系
〔〒305-8573 つくば市天王台 1-1-1〕
e-mail: kudo@cs.tsukuba.ac.jp
投稿受付：2020 年 8 月 5 日

2. 提案手法

本研究では, CS と深層学習を巧妙に組み合わせ更なる画質改善を実現する新しい画像再構成の数学的枠組みを提案する．**図 1** に提案手法の原理を示す．提案手法は，手法 1 とその簡易版である手法 2 の 2 つの手法から構成される．手法 1 では, 通常の FBP 法の再構成画像 **y** とトータルバリエーション（TV）や非局所 TV など既存の CS を用いた再構成画像 **z** を求め，これらを U-Net など画像変換を行う CNN に入力して再構成画像 **x** を求める．入力画像が 2 枚であ

るため，2 入力 1 出力のマルチチャンネル構造の CNN を用いる．そして，CNN のパラメータ（重み **w** 及びバイアス **b**）と CS 画像再構成の正則化パラメータ β の値は，学習データから 2 乗誤差が最小になるように自動決定する．手法 2 は，CNN の処理を簡単化するために，手法 1 において FBP 法の再構成画像を入力せず，CS 画像 1 枚のみを入力するようにした簡易版である．大雑把には，提案手法は CS 再構成画像の画質を学習データを利用して深層学習で更に改善することを実現した手法，といえる．

3. シミュレーション実験

提案手法の性能を評価するシミュレーション実験を，医療用腹部 CT の実画像データセットを用いて行った. 想定した CT 撮影の状況は，スパースビューCT と低線量 CT の両者である．学習データは、58 症例の 3D 画像から肝臓が大きく写る 5 スライスを切り出した合計 290 枚の画像，評価に用いるテストデータは 16 症例の 3D 画像から肝臓が大きく写る 5 スライスを切り出した合計 80 枚の画像とした．CNN の入力画像を生成する CS 画像再構成法としては，TV と非局所的 TV を用い，評価関数最小化の反復解法は筆者らが開発した高速に真の解に収束する「改良型 OS-SART-TV 法」と呼ばれる手法を用いた．また，CNN としては U-Net，学習の損失関数は 2 乗誤差, 学習アルゴリズムは Batch Normalization を組み込んだ Adam というごく一般的な深層学習手法を用いた．

スペースの制約から，投影データの測定方向数が 64 方向のスパースビューCT の場合の手法 1 と手法 2 による再構成画像を，**図 2** に 1 サンプルのみまとめて示す（MSE（Mean Squares Error）の値と SSIM の値も画像中に示す）．深層学習単独の再構成画像には上述のように血管など構造物の消滅が随所に見られ，TV の画像は構造物の消滅はないが Staircase アーティファクトや滑らかな濃度変化の消滅が見られる．これに対して，TV+深層及び非局種 TV+深層の画像は CS と深層学習の長所を上手く組み合わせ，数値評価でも視覚評価でも劇的な画質改善を実現できている．他のテスト画像についても同様な画質改善効果が得られた．**図 3** に，全テストデータの再構成画像の PSNR 値と SSIM 値の平均値をグラフにまとめたものを示す．手法 1 と手法 2 の比較を行った結果、大半の例について手法 1 の方が若干画質改善効果に優れ, CS 画像に加えて FBP 画像を入力することが有効であることが分かった．また，低線量 CT の場合もほぼ同様な画質改善が実現できた．

なお，本結果の再現性を検証するため，2 人で独立して同様な実験を行ったが同じ傾向の結果が得られ再現性は高いと考えている．

4. まとめ

本研究では，CS 再構成画像を CNN に入力して更なる画質改善を実現する，新しい CT 画像再構成の数学的枠組みを提案した．シミュレーション実験の結果, CS 単独や深層学習単独の場合と比較して, 比較的単純な CS と CNN の組み合わせ方で各々の手法の欠点を補い劇的な画質改善が実現できた．

最後に，本手法の性能を更に改善した手法も JAMIT 2020 大会で発表している [8].

謝辞

本研究は, JST-ERATO（JPMJER1403）及び JST-CREST（JPMJCR1765）の助成を受けたものです.

利益相反の有無

なし.

文　献

［1］ K.H.Jin, M.T.McCann, E.Froustey et al, Deep Convolutional Neural Network for Inverse Problems in Imaging, IEEE Trans Image Process, 26, 4509-4522, 2017.

［2］ H.Chen, Y.Zhang, W.Zhang et al, Low-dose CT via Convolutional Neural Network, Biomed Opt Express, 8, 679-694, 2017.

［3］ B.Zhu, J.Z.Liu, S.F.Cauley et al, Image Reconstruction by Domain-Transform Manifold Learning, Nature, 555, 487–492, 2018.

［4］ D.Wu, K.Kim, G.El Fakhri et al, Iterative Low-dose CT Reconstruction with Priors Trained by Artificial Neural Network, IEEE Trans Med Imaging, 36, 2479-2486, 2017.

［5］ C.Tang, J.Li, L.Wang et al, Unpaired Low-Dose CT Denoising Network Based on Cycle-Consistent Generative Adversarial Network with Prior Image Information, Comput Math Methods Med, 2019, Article ID 8639825, 2019.

［6］ Q.Yang, P.Yan, Y.Zhang et al, Low-Dose CT Image Denoising Using a Generative Adversarial Network with Wasserstein Distance and Perceptual Loss, IEEE Trans Med Imaging, 37, 1348-1357, 2018.

［7］ 森和希，工藤博幸，千北一期，畳み込みニューラルネットワークを用いた低線量 CT 画像再構成法の多時相 CT イメージングへの拡張，第 38 回日本医用画像工学会大会（JAMIT 2019）予稿集, Paper No OP2-6, 2019.

［8］ 数尾昴，川又健太郎，工藤博幸，深層学習による圧縮センシング CT 画像再構成画質改善の新手法，第 39 回日本医用画像工学会大会（JAMIT 2019）予稿集, 2020（本大会発表予定）

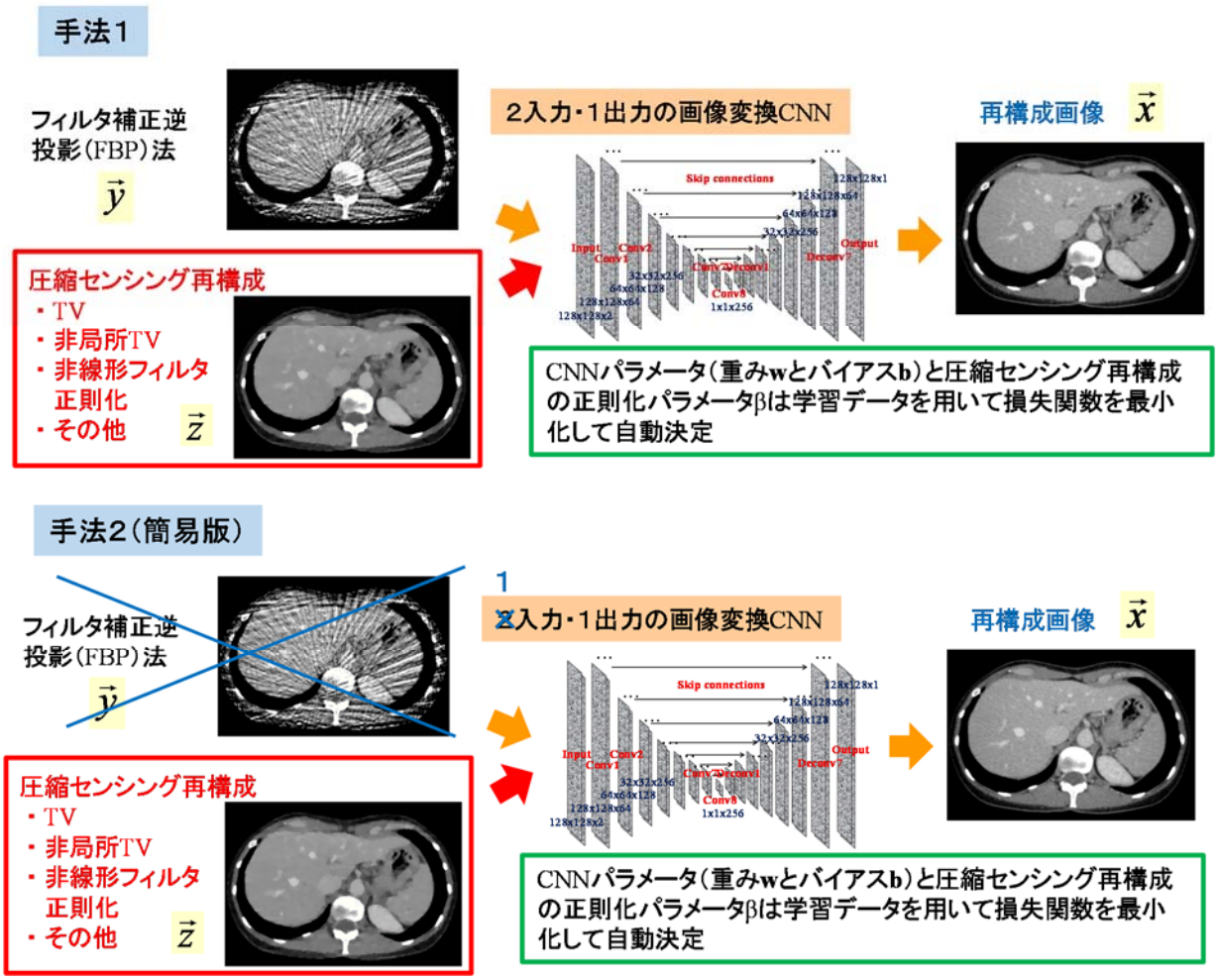


図 1 提案手法（手法１と手法２）の原理

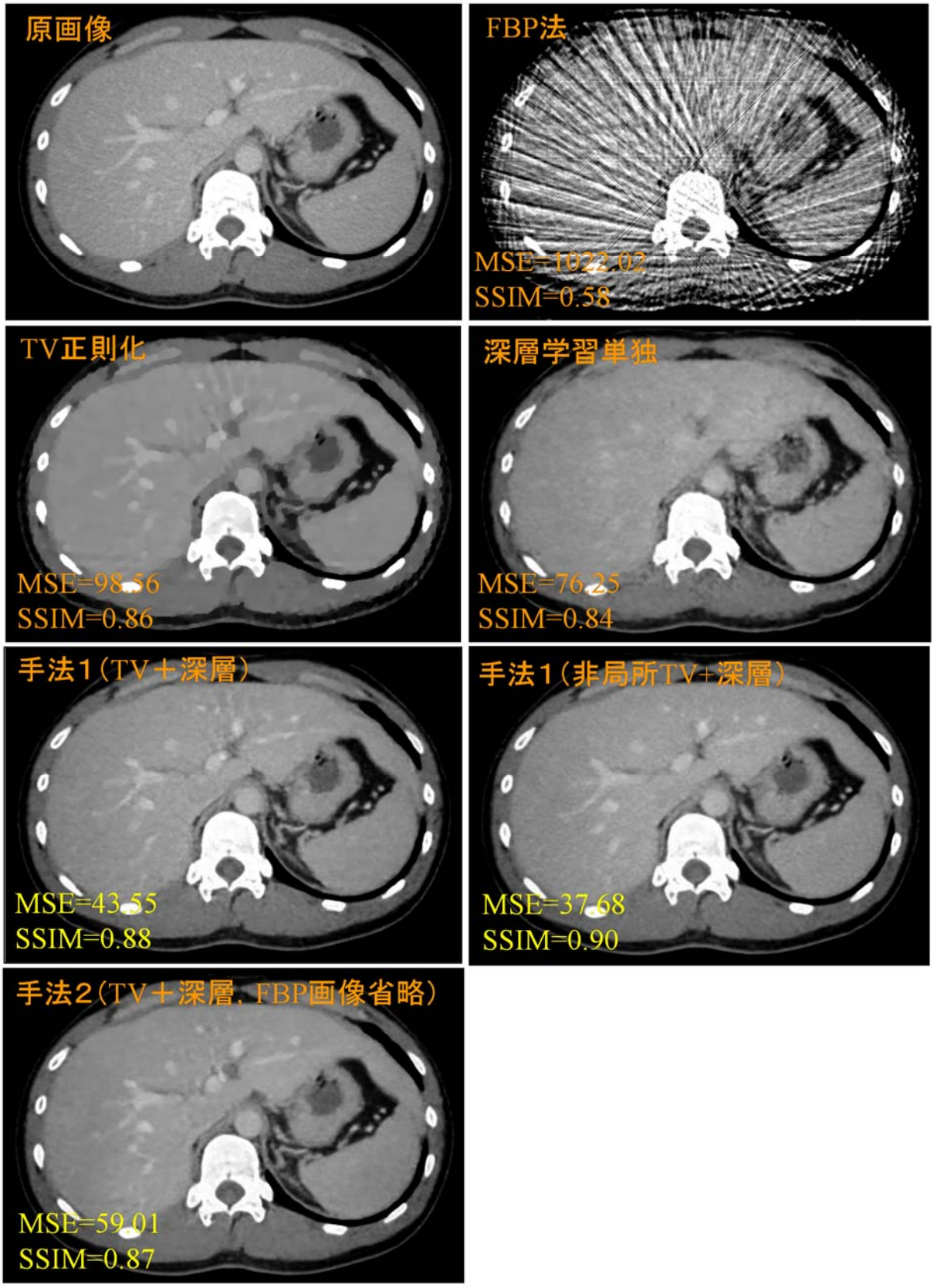
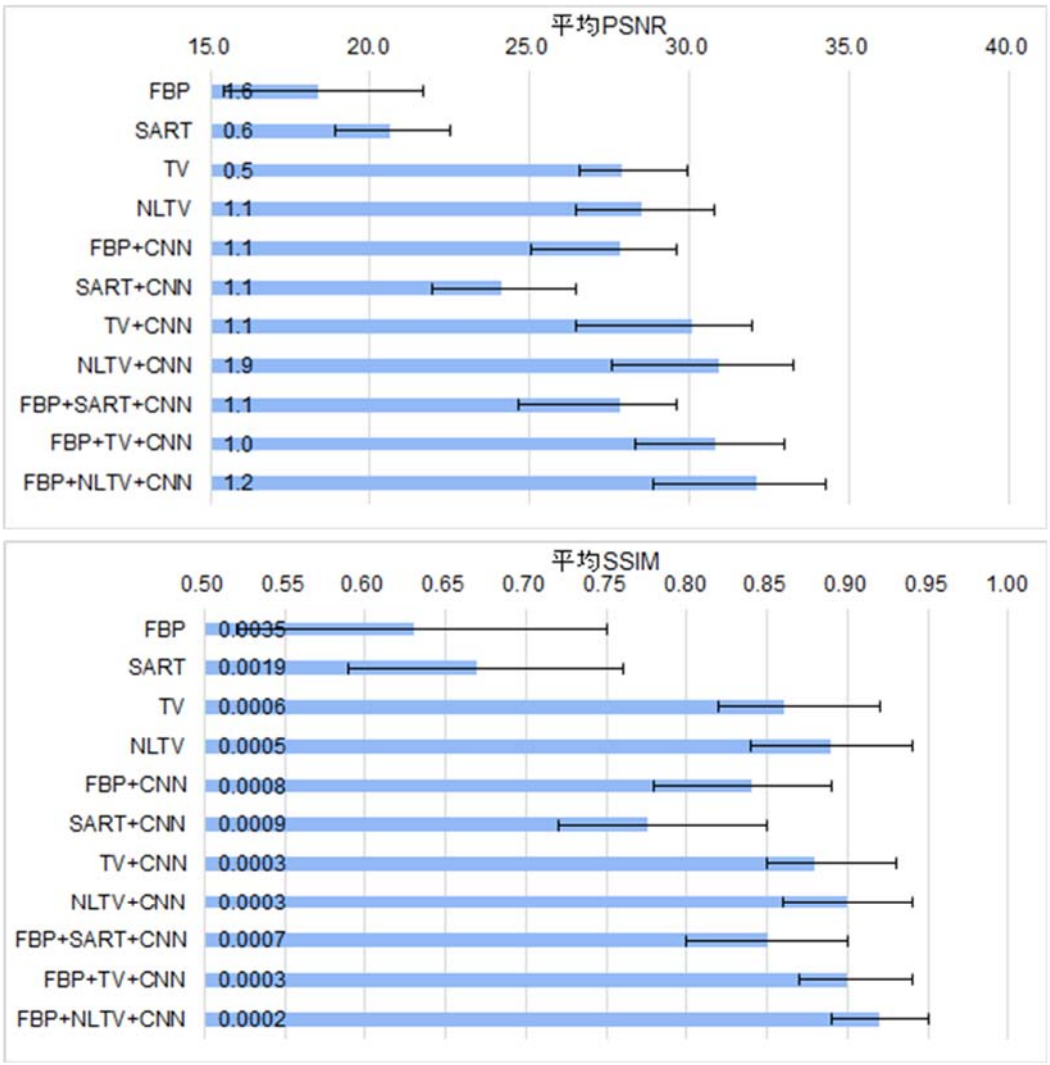


図 2 スパースビューCT（投影データ方向数 64）の場合の再構成画像例



FBP -> FBP法, TV -> TV再構成, NLTV -> 非局所TV再構成, FBP +CNN -> 深層学習単独
TV+CNN -> TV入力の手法2, NLTV+CNN -> 非局所TV入力の手法2,
FBP+TV+CNN -> FBPとTV入力の手法1, FBP+NLTV+CNN -> FBPと非局所TV入力の手法1

図 3 スパースビューCT（投影データ方向数 64）再構成画像の PSNR 値と SSIM 値の全テストデータの平均値（グラフの黒実線はエラーバー，左側の数値は分散の数値）

Improving Image Quality in Compressed Sensing CT Image Reconstruction Using Deep Learning

Hiroyuki KUDO[*1], Kazuki MORI[*1]

*1 University of Tsukuba

Compressed Sensing (CS) image reconstruction in CT suffers from the drawbacks such as 1) appearance of staircase artifacts and 2) loss in image textures and smooth intensity changes. These drawbacks stem from the fact that CS is based on approximating the image by a piecewise-constant function. In this research, we propose a method to improve image quality in CS using deep learning. In the proposed method, FBP reconstructed image and CS (TV or Nonlocal TV) reconstructed image are inputted into a CNN with two input channels and single output channel, and the final reconstructed image is obtained by the output of CNN. Parameters (weight and bias) of CNN together with a regularization parameter of CS are estimated by minimizing the average least-squares loss function by using learning data, i.e. a set of triplet of degraded FBP and CS reconstructions and answer image. We have applied the proposed method to sparse-view CT and low-dose CT image reconstruction using a dataset of abdominal CT image. The standard deep learning reconstruction often suffered from loss in structures such as blood vessels in liver, whereas the standard CS reconstruction suffered from the staircase artifacts and loss in smooth intensity changes. On the other hand, the proposed method succeeded in generating images with much higher quality by compensating for each other's weakness.

Key words: CT, Image Reconstruction, Compressed Sensing, Deep Learning, Convolutional Neural Network

トポロジー制約を用いた スパースビューCT 画像再構成の試み

千北　一期*1　工藤　博幸*2

要旨

スパースビューCT 画像再構成は，投影方向数を通常 CT の数十分の一とすることで被曝量や撮影時間の削減を目指す技術として近年注目されている．一方で，再構成された画像には角度欠損によるアーティファクトが発生する等の課題も多くある．先行研究としてこれまで，圧縮センシング及び統計的逐次近似手法に基づく再構成法が多く検討されてきた．本研究では一般画像のセグメント抽出等に使われてきたトポロジー保存の方法に着目する．再構成対象画像の保存するべきトポロジー性質を先験情報として与え，これを定式化することで，逐次的再構成法の中でトポロジー保存を考慮する事のできる新たな正則化項の提案を行う．従来法との比較のために，数値シミュレーション実験を異なるトポロジーを持つ複数の画像に対して行った．その結果，従来除去できなかったアーティファクトの除去が可能となり，画像品質向上に繋がることが示唆された．

キーワード：スパースビューCT，画像再構成，トポロジー拘束条件

1. はじめに

コンピュータ断層撮影 (Computed Tomography: CT) は低被ばく化が求められ，少数方向の投影データから再構成を行うスパースビューCT や，照射する X 線の線量を小さくする低線量 CT などの手法が提案されている. [1] これまで CT 画像再構成の問題に対して，逐次近似画像再構成法や圧縮センシングを組み合わせた方法が取られてきた．特に定式化において正則化項にどのような特性をもたせるかが重視されている．一例としては正則化項に対してトータルバリエーション (Total Variation : TV)法を用いるものがある．この方法は CT 画像に対して有効であることが示されている．[2]

本稿では，通常の CT 画像再構成での投影方向 1000 から 2000 程度に対して， 8 方向以下の少数方向スパースビューCT を考える．対象とする物体は 0 または 1 の二種のラベルによってラベル付けされる離散トモグラフィーとし，その対象画像のトポロジー制約は維持されるという先験情報を正則化項に適用することを試みる．

*1 筑波大学システム情報工学研究科
〔〒305-8573 つくば市天王台 1-1-1〕
e-mail: chigita@imagelab.cs.tsukuba.ac.jp
*2 筑波大学システム情報系
〔〒305-8573 つくば市天王台 1-1-1〕
投稿受付：2020 年 8 月 5 日

2. 提案手法

2.1 概要

対象の再構成画像は 2 値に分類できる離散トモグラフィーを仮定する．そのため，画像には予めラベル$\boldsymbol{L_1}$及び，$\boldsymbol{L_2}$を振り分ける．

この時これらが構成するトポロジー制約は

保存されると仮定する．トポロジー制約が保存されるには2.2で述べるトポロジーチェック［3］を行う必要がある．これは，$\boldsymbol{L_2}$の連結成分と$\boldsymbol{L_1}$の連結成分の個数，また孔の個数は変化しないことと同値である．この制約条件は再構成を行う反復の過程でも常に満たされる．そのため，$\boldsymbol{L_2}$のの孔が2つ以上に分離したり，孔の状態を満たさなくなったりすることはない．これを定式化し正則化項として再構成を行う．

2.2 トポロジーチェック

トポロジーチェックとは対象画像のあるピクセルのラベルを操作した時，元の画像と同じトポロジー構造を保持するかどうかを検査する手法である．以下の手順によって行われる．

[Step 1] :検査したいピクセルを中心に3 x 3の window からなる領域を取り出す．取り出した領域を表すベクトルを$\boldsymbol{w}$とする．$\boldsymbol{w}$は**図 1** のように中心から時計回りに走査したものを並べている．$\boldsymbol{w_0}$が変更可能性を調べるピクセルである．

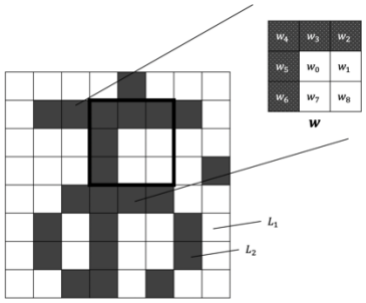


図 1: トポロジーチェック領域抽出例

[Step 2] : $\boldsymbol{w}$の各ピクセルを反転する．つまり，L_1はL_2に，L_2はL_1とする．ここでは反転済みのベクトルを$\tilde{\boldsymbol{w}}$と表す

[Step3] ： 以下の式(1)を計算する．ただし$w_9 = w_1$とする．このときcが1であれば変更可能であり，0 のときにはトポロジー構造を破壊する操作である．

$$c = \sum_{i=0}^{3} \tilde{\boldsymbol{w}}_{2i+1} - \sum_{i=0}^{3} \prod_{j=1}^{3} \tilde{\boldsymbol{w}}_{2i+j} \qquad (1)$$

2.3 トポロジー拘束条件つき正則化項

トポロジー制約を常に満たす方向にのみ更新をかけるような正則化を考える．この項はTV 正則化に追加して定式化する．制約を常に満たすためには反復計算の一時保存として前回の反復結果画像を保存し，その差分に対してトポロジーチェックを行うことで実現する．

2.4 定式化

以下の式(2)のように定式化する．この時，$\boldsymbol{A}$はシステム行列，$\boldsymbol{x}$は再構成画像，$\boldsymbol{b}$は投影データ，βはデータ項と正則化項の重みを制御するパラメータである．

$$\min_{\boldsymbol{x}} \|\boldsymbol{Ax} - \boldsymbol{b}\|_2^2 + \beta(\|\boldsymbol{x}\|_{TV} + \|\boldsymbol{x}\|_{topology}) \quad (2)$$

これをデータ項は SART 法をベースとしたローアクション型の近接スプリッティング[4] によって解く．トポロジー正則化項は2.3に示した方針によって解く．これを交互に繰り返し反復することで解の収束を目指す．

3. シミュレーション実験

提案手法の効果を評価するために，いくつかのトポロジーパターンを持つ人工的に作成した二値画像にて評価実験を行った．比較対象の手法として OS-SART-TV 法を用いた．スパースビューCT の制約から投影方向は 8 方向とした．計測ノイズを表現するため投影データ作成時にはポアソンノイズを付加した．また，PSNR によって画像品質を定量的に比較する．

結果は**図 2** のようになった．これを定性的に評価すると従来法ではノイズの影響を受けたピクセルに一部間違ったラベルによって復元しているが，提案法ではトポロジー拘束を満たすように動作させるため特に赤枠では顕著にノイズ除去の効果も現れている．また，2種類のトポロジー条件をもつ画像においてもその構造を変化させることなく問題なく動作しているといえる．

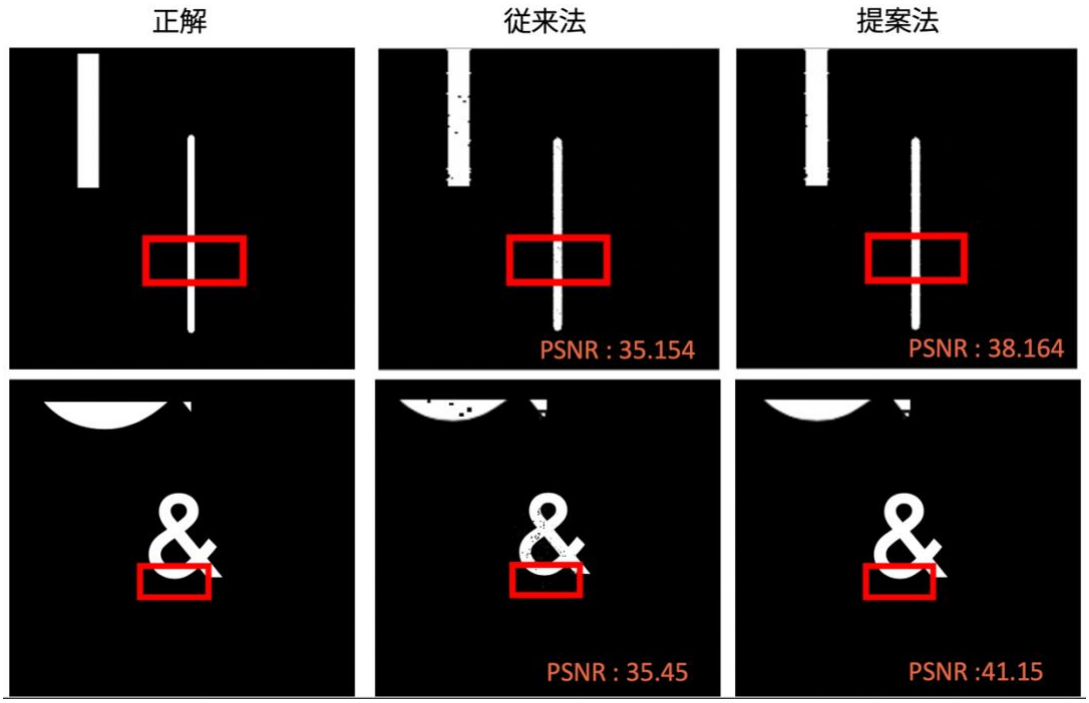


図２： 結果画像

4. まとめ

本稿では，トポロジー拘束条件付き正則化項を導入し定式化を行った．またシミュレーション実験によってその効果を評価した．その結果，画像再構成においてノイズ除去の効果があり，再構成画像の品質向上に効果がある事がわかった．今後の展望としてラベルの多値化，連続画像へのアルゴリズムの応用などが挙げられる．

利益相反の有無

無

文　献

［1］ 工藤 博幸."低被曝 CT における画像再構成法-統計的画像再構成，逐次近似画像再構成,圧縮センシングの基礎-".Medical Image Technology Vol. 32 No.4 (2014)

［2］ L.Ritschl, F. Bergrner and C. Fleischmann et al. "Improved total variation-based CT image reconstruction applied to clinical data". Phys. Med. Biol.56(2011) 1545-1561

［3］ 阿南 泰三，工藤 博幸，斎藤 恒雄．"画像の領域抽出のための確率的領域膨張収縮モデル". 電子情報通信学会論文誌 D. Vol. J80-D2 No. 1 pp. 156-165 (1997)

［4］ 山嵜　深，根本拓也，高木景太，et al. "ローアクション型トータルバリエーション正則化法と CT 画像再構成への実用" 電子情報通信学会技術研究報告 MI no. 482. pp.199-204 (2015)

A method for sparse-view CT reconstruction using topology-constraint

Kazuki CHIGTA*1, Hiroyuki KUDO*2

*1 Graduate School of Systems and Information Engineering, University of Tsukuba

*2 Faculty of Engineering, Information and Systems, University of Tsukuba

Recently, sparse-view CT is focused on a technique to reduce the amount of radiation exposure and imaging time by reducing the number of projection directions. However, there are many problems in the reconstructed images, such as the occurrence of artifacts due to angular defects. Previous studies have examined many methods of reconstruction based on compressed sensing and statistical sequential approximation. In this study, we focus on the methods of topology constraint that have been used for segment extraction of general images. We propose a new regularization term that can take into account the topology constraint in the sequential reconstruction method by giving a priori information on the topological properties of the image to be reconstructed and formulating it. In order to compare with the conventional method, numerical simulation experiments were performed on several images with different topologies. As a result, it was suggested that the removal of artifacts that could not be removed by the conventional method is possible, and it leads to the improvement of image quality.

Key words: sparse-view CT, image reconstruction, digital topology

深層学習による
圧縮センシング CT 画像再構成画質改善の新手法

数尾 昴[*1]　工藤 博幸[*1]　川又 健太郎[*1]

トータルバリエーション正則化などの圧縮センシングを用いた CT 画像再構成は,Staircase Artifact の発生および滑らかな濃度変化やテクスチャーが消失しやすいことの二点の問題点がある.深層学習単独の場合には学習データから学びにくい構造物の消失が問題点となっている.それぞれの問題に対し本論文では,両者を組み合わせて低線量 CT 及びスパースビューCT に対する画質改善をさせる新手法を提案する.
提案手法では,フィルタ補正逆投影（FBP）法と TV や非局所 TV に代表される圧縮センシングの再構成画像を複数入力 1 出力の画像変換を行う CNN に入力して,画質を改善した出力画像を得る.圧縮センシングの正則化パラメータは,視覚的に優れているもの数枚を決定する.スパースビューCT と低線量 CT の状況を想定して,腹部 CT 画像のデータセットを用いたシミュレーション実験を行った結果,深層学習単独の手法と TV での画像再構成のそれぞれの欠点を補い画質改善を実現した.

キーワード：画像再構成,低線量 CT,スパースビューCT,深層学習,畳み込みニューラルネットワーク

1. はじめに

Computed Tomography(CT) 法は多方向から X 線を照射し臓器の吸収した線量を計測し投影データを得て,再構成することにより臓器を破壊することなく臓器の断面図を得ることができる技術である. CT は現代医療にとって欠かせないものとなっているが, CT の撮像に伴う撮影対象の被曝量による健康被害が問題視されている.被曝量を減らす方法としては, 大きく 2 つあり X 線管の電流を落として投影データ測定を行う低線量 CT, または撮影方向数を削減して投影データ測定を行うスパースビューCT を行うことにより被曝量を減らすことができるが,低線量 CT は雑音が多いデータを生成してしまい,スパースビューCT はアーティファクトと呼ばれる実際の臓器とは異なる形を含むデータを生成してしまう.この様に低被曝化と画像の質の向上はトレードオフの関係にあるため低被曝でありながら画質の向上を目指すには復元技術の向上が必要となり, 近年精力的に研究が行われてきている.これまでトータルバリエーション正則化項を用いた圧縮センシングやニューラルネットワークによる雑音除去手法やアーティファクトを除去する手法が提案されている. [1-2]
圧縮センシングでは Staircase アーティファクトの発生や滑らかな濃度変化の消失が起きてしまい,深層学習では学習データでは学びにくい構造の消失や本来存在しない細かい構造を出力してしまう問題がある.
提案手法では低線量 CT,スパースビューCT において畳み込みニューラルネットワークを拡張し複数のハイパーパラメータのトータルバリエーション正則化項を用いた圧縮センシングを施した画像を入力画像とするネットワークを用い,ノイズやアーティファクトを含む腹部 CT 画像の画質の改善をする手法を示す.

*1 筑波大学大学院システム情報工学研究科コンピュータサイエンス専攻
〔〒305-8573 つくば市天王台 1-1-1〕
e-mail: kazuo@imagelab.tsukuba.ac.jp
投稿受付：2020 年 8 月 5 日

2. 提案手法

提案手法では劣化画像に複数のトータルバリエーション正則化項のパラメータ β を用いた圧縮センシング再構成画像を一つのインプ

ットとして畳み込みニューラルネットワークに学習させることで従来手法より質の高いフィルタを得る.

森らによって提案された深層学習を用いた圧縮センシング CT 画像再構成の画質改善[9]では CS 画像再構成の正則化パラメータを学習データから二乗誤差が最小となるものを一つ選択していたが,パラメータを複数選択することで各パラメータにおける正解画像に近くなるための情報を選択し出力を得ることができる.

また提案手法では複数のトータルバリエーション正則化を行うという性質上,計算時間が増えてしまうため,逐次近似法などは用いていない.

2.2 ネットワーク構造

畳み込みニューラルネットワークは本研究では特徴量抽出を目的としたエンコーダ部分と抽出された特徴量から解像度を戻していく.

デコーダ部分からなる完全畳み込みニューラルネットワークの一つである Unet[3]を採用した.エンコーダ部分では畳み込みはパディングをして大きさを変えず 3x3 カーネルで畳み込みをする.プーリング層でマックスプーリングに通し,特徴量を圧縮している.

デコーダ部分ではアップサンプリングで得た特徴を大きくしてから畳み込みを繰り返し行なっている.損失関数は畳み込みニューラルネットワークのみを用いた従来手法と同じく出力画像と正解画像の平均二乗誤差を用い,式のようになる

$$Loss = MSE(w, b, \beta_1, \dots, \beta_n)$$
$$= \frac{1}{N}\sum_{i=1}^{N}\left(t_i - CNN(w, b, \beta_1, \dots, \beta_n)\right)^2 \qquad (1)$$

2.3 データセット

100 人の患者の腹部 CT 画像(512x512)を 5 枚ずつ取り出し 500 枚の腹部 CT のうち 400 枚を学習画像とし 100 枚をテスト画像とした.

劣化画像は正解画像から作成し以下のようになる.

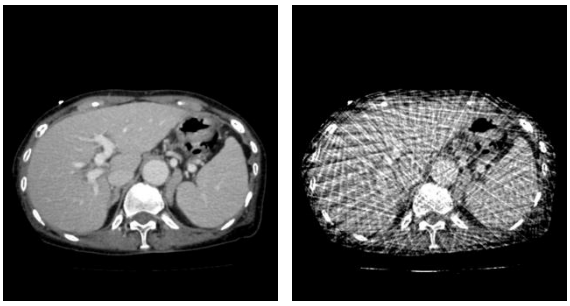


図 2 元画像と劣化画像例

2.4 評価指標

定量的な評価指標として PSNR, SSIM を用いる.また定性的にも評価する.各指標の式は次のようになる.

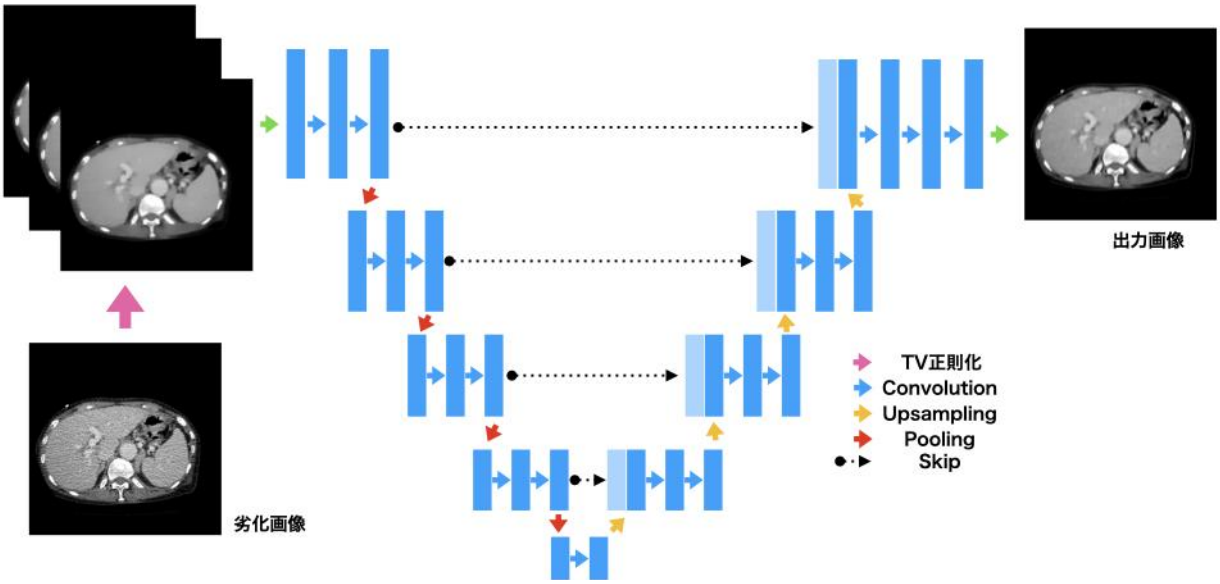


図 1 ネットワーク構造

$$PSNR = 10 \cdot \log_{10} \frac{MAX^2}{\mathrm{MSE}} \quad (2)$$

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_j - t_i)^2 \quad (3)$$

$$SSIM = \frac{(2\mu_y\mu_t + C_1)(2\sigma_{yt} + C_3)}{(\mu_y^2 + \mu_t^2 + C_2)(\sigma_y^2 + \sigma_t^2 + C_3)} \quad (4)$$

3 実験

損失関数には劣化画像と非劣化画像の平均二乗誤差を用い,最適化法には Adam を用いて100epoch,学習率は 0.001 を初期値に設定し15epoch ごとに 0.8 倍をした.実装には TensorFlow,計算には GPU(GeForce RTX 280)を用いた.

従来手法(TV＋CNN)[9]では,TV 正則化のパラメータは二乗誤差が最小になるものを選択し,提案手法のパラメータは視覚的に優れている TV 正則化パラメータを等間隔で 5 つを選択した.

出力例として低線量 CT における各手法の再構成画像を図に示す.TV の画像は Staircase アーティファクトが見られ,CNN 手法は画像中部における構造が消失していることがわかる.TV＋CNN と提案手法を視覚的に比較すると画像左側の血管の再現において提案手法が優れていることがわかる.複数のパラメータを入力することにより,血管など細かい構造の特徴量をとらえやすくなり,従来手法よりきれいに再現できていると考えられる.また深層学習による画像再構成における本来ない構造を出力してしまう問題を提案手法は解決している.これは他の低線量 CT 画像においても,スパースビューCT においても同じ傾向にあった.

TV 正則化パラメータをより増やすことで画質の改善を図ったが大きな変化はなく,適切な数を選択する必要があることがわかった.

3.1 低線量 CT

光子数 25000 の低線量 CT における再構成画像の平均 PSNR 及び SSIM は表のようになった. 数値評価上では大幅な改善は見られないものの,実際の画像を視覚で確認すると細かい構造において従来手法より改善が見られた.

手法	PSNR	SSIM
劣化画像	14.27	0.797
TV	18.81	0.811
CNN	20.08	0.814
TV+CNN	20.55	0.834
提案手法	21.05	0.847

3.2 スパースビューCT

64 方向スパースビューCT における再構成画像の平均 PSNR 及び SSIM は以下のようになった.スパースビューCT に対しては数値比較,視覚による比較において改善はみられるものの低線量 CT と比較すると大きな改善は見られなった.

手法	PSNR	SSIM
劣化画像	13.27	0.713
TV	15.09	0.788
CNN	18.02	0.801
TV+CNN	18.24	0.811
提案手法	18.49	0.818

4. まとめ

本研究では,シミュレーションにより作成した劣化画像において複数のパラメータのトータルバリエーション正則化を用いることで従来手法における問題であった Staircase アーティファクトや構造の消失に有効な手法であることが分かった.

今後,パラメータの選別や逐次近似法を組み込むことでより鮮明な画像の出力を追求していく予定である.

利益相反の有無

なし

文　献

［1］ K.H.Jin, M.T.McCann, E.Froustey et al, Deep Convolutional Neural Network for Inverse Problems in Imaging, IEEE Trans Image Process, 26, 4509-4522, 2017.

［2］ H.Chen, Y.Zhang, W.Zhang et al, Low-dose CT via Convolutional Neural Network, Biomed Opt Express, 8, 679-694, 2017.

［3］ Ronneberger O, Fischer P, Brox T：U-Net: convolutional networks for biomedical

image segmentation. International Conference on MICCAI, 2015

［4］ Long J, Shelhamer E, Darrell T: Fully Convolutional Networks for Semantic segmentation. International Conference on CVPR, 2015

［5］ Ioffe S, Szegedy C: Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. International Conference on Machine Learning, 2015

［6］ Kingma D, Ba Jimmy: Adam: A Method for Stochastic Optimization. International Conference on Learning Representations, 2015

［7］ Wang Z et al: Image Quality Assessment: From Error Visibility to Structual Similarity. ［8］ IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL 13, NO.4, 2004

［9］ Buades A, Coll B, Morel J: A non-local algorithm for image denoising. In Proc IEEE Conf Comp Vis Patt Recogn 2: 60-65, 2005

［10］ 森和希, 工藤博幸, 千北一期, 畳み込みニューラルネットワークを用いた低線量 CT 画像再構成法の多時相 CT イメージングへの拡張, 第 38 回日本医用画像工学会大会（JAMIT 2019）予稿集, Paper No OP2-6, 2019.

［11］ 深層学習を用いた圧縮センシング CT 画像再構成の画質改善,第 39 回日本医用画像工学会大家(JAMIT2020)予稿集,2020(本大会発表予定)

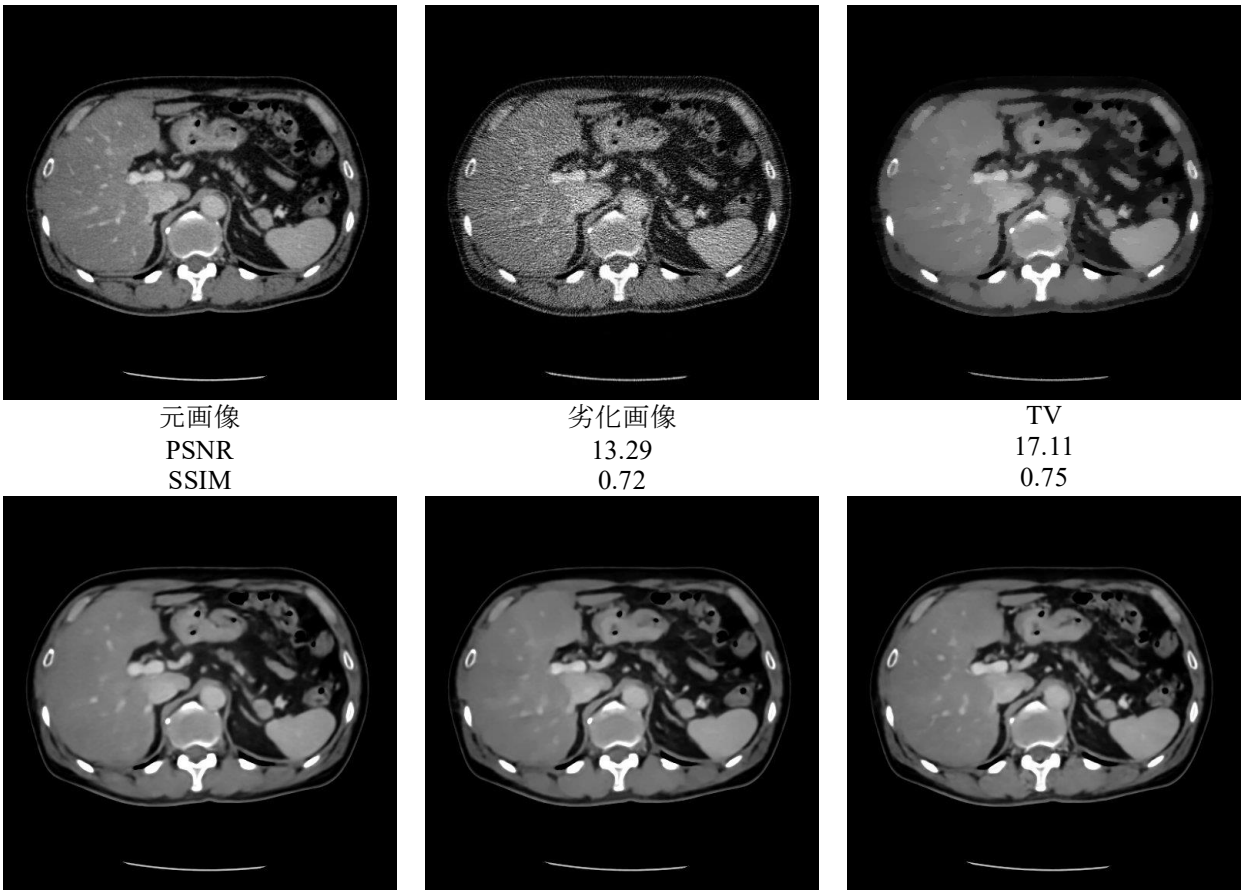


図 3 低線量 CT(光子数 25000)における画像再構成例

A New Method for Improvement of Image Quality in Compressed Sensing CT Image Reconstruction Using Deep Learning

Subaru Kazuo[*1], Hiroyuki Kudo[*1], Kentaro Kawamata[*1]

*1 University of Tsukuba

There are two problems in the reconstruction of CT images using compressed sensing such as total variation regularization: the occurrence of staircase artifacts and the loss of smooth density changes and textures. In the case of deep learning alone, the problem is the loss of structures that are difficult to learn from the training data. In this paper, we propose a new method to improve the image quality of low-dose CT and sparse view CT by combining both methods.In the proposed method, the filter-corrected back projection (FBP) method and the reconstruction images of TV and non-local TV are input to a CNN which is capable of multiple inputs and one output, and the output image quality is improved. The regularization parameter of the compressed sensing system is determined for a few visually superior images. Experimental simulations are performed using a dataset of abdominal CT images for both sparse-view and low-dose CT, and the improvement of image quality is achieved by compensating for the shortcomings of both deep learning and TV image reconstruction methods.

Key words: Image reconstruction, Low-dose CT, Sparse view CT, Deep learning, Convolutional neural network

CT 画像再構成における混合インテリアエクステリア問題の解の一意性

藤井　克哉[*1]　金　鎔采[*1]　工藤　博幸[*1]

要旨

不完全なデータからの再構成法は，CT 画像再構成の分野で非常に重要な研究であり，一般に不完全な投影データからの再構成の問題は，インテリア問題，エクステリア問題等が良く知られている．近年までにこれらの問題における解の一意性と安定性の理論は各問題に対して構築されている．本稿では，エクステリア問題の解の一意性に用いられる Cormack 型の再生公式を念頭に置き，これらの問題が混在する場合の解の一意性を提案する．また得られた一意性の検証として，modified Shepp-Logan ファントムと胸部 CT 画像を用いてシミュレーション実験を行った．

キーワード：CT, インテリア CT, エクステリア CT, 画像再構成

1. はじめに

X 線コンピューター断層撮影(CT)は，さまざまな方向から X 線を照射することにより対象物の断層画像を取得する手法である．近年，CT への応用では，制限された投影データからの再構成問題に関心が集まっており，そこで数学的に問題となるのは，解の一意性と安定性である．本稿では一意性の問題に焦点を当てる．一般に，一意性とは再構成した結果の正確性を保証する．不完全投影データの問題はよく知られた不良設定問題となるため一意性を示す必要がある．

制限された投影データからの再構成におけるよく知られている問題には，内部の問題（インテリア問題），外部の問題(エクステリア問題)，制限された角度の問題(Limited Angle problem)などが挙げられる．本稿では，インテリア問題とエクステリア問題を扱う．インテリア問題とエクステリア問題は，それぞれオブジェクトの内側と外側にあるときに関心領域（ROI）を通過する投影データのみに基づく再構成方法である(図 1)．エクステリア問題は物体の内部構造が X 線を完全に吸収する場合の再構成に利用される．また，インテリア問題は内部関心領域を通る投影のみを測定することで患者の X 線被曝を低減できるため，医用画像への応用が期待できる．

以下では，これら 2 つの問題の一意性の結果を簡単に紹介する．エクステリア問題における解の一意性が成り立つことは, Cormack の再生公式によって導出されるサポート定理として知られている[1, 2]．一方で，インテリア問題の解の一意性は成り立たないことが長い間知られていた[2]．しかし 2004 年頃から、インテリア問題の一意性の問題は劇的に変化し，再構成する ROI の形状に応じて解が一意性を持つことがわかってきた．特に, ROI がオブジェクト内に完全に含まれている場合，ROI の小さな領域が既知であれば一意性を満たすことが（[3-5]を参照）．しかしながら，これらの不完全投影データを組み合わせると，解の一意性が成立するかどうかは自明ではない． そこで，本稿では両側欠損（外部と内部）のある投影データからの画像再構成における解の一意性を提案する．

*1 筑波大学システム情報系
〔〒305-8573 つくば市天王台 1-1-1〕
e-mail: fujii@imagelab.cs.tsukuba.ac.jp
投稿受付：2020 年 8 月 5 日

2. 提案手法

前節において，それぞれ内部欠損と外部欠損のある不完全投影データからの再構成における一意性は成り立つことを見た．では，投影データが内部と外部に同時に欠損がある場合はどうだろうか．このような問題を本稿では，インテリアエクステリア問題と呼ぶことにする．この問題の解の一意性を導く鍵は図 2 のような円環先見領域が既知であるという条件である．

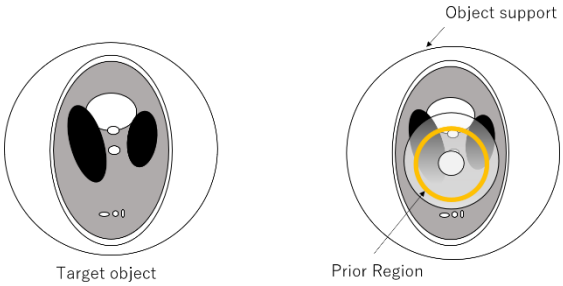


図 2　インテリアエクステリア問題における先見領域

インテリア問題は，内部の投影データがあるため，有限逆ヒルベルト変換と先見情報を用いることで解析接続により ROI の解の一意性を導いている[3]．インテリアエクステリア問題の場合は，ベースとなる数式を以下の Cormack の再生公式を用いて図 2 の先見情報と組み合わせ，解析接続を行うことで円環領域の ROI の解の一意性を示すことができる．

[Cormack の再生公式]

$$f_\ell(p) = 2\int_p^1 \hat{f}_\ell{}'(r)\frac{T_\ell\left(\frac{r}{p}\right)}{\sqrt{p^2-r^2}}dr \qquad (1)$$

ここで，$\hat{f}$が投影，T_ℓは第一種チェビシェフ多項式である．

3. 数値シミュレーション

提案手法を数値シミュレーション実験により確認する．データセットは医療用腹部 CT 画像 512x512pixels と modified Shepp-Logan ファントム 256x256pixels を用いた．ROI はともに円環領域であるとする．また投影方向数は 1024 方向，不完全投影データは内部と外部の欠損が起きているサイノグラムとする．このとき，円環領域の ROI は再構成できていることが確認できる（図 3, 4 参照）．

4. まとめ

本研究では，以前より提案されていたインテリア問題とエクステリア問題を組み合わせ，より投影データの制限の多いインテリアエクステリア問題の解の一意性を提案した．また数値シミュレーションにより提案した手法の有用性を示すことができた．

謝辞

本研究は，JSTCREST（JPMJCR1765）の助成を受けたものです．

利益相反の有無

なし．

文　献

［1］ A. M. Cormack: Representations of a function by its line integrals, with some radiological applications. J Appl Phys 34:2722–7, 1963.

［2］ F. Natterer: The Mathematics of Computerized Tomography (second edition), SIAM (Classics in Applied Mathematics, Vol32), Philadelphia, PA, 2001.

［3］ H. Kudo, M. Courdurier, et al: Tiny A Priori Knowledge Solves the Interior Problem in Computed Tomography. Phys Med Biol. 53:2207-31, 2008.

［4］ Y. Ye, H. Yu, et al: A general local reconstruction approach based on a truncated Hilbert transform. Int.J.Biomed.Imaging. **2007**:Article ID 63634, 2007.

［5］ 工藤博幸:第 2 章　投影からの画像再構成．日本医用画像工学会編：医用画像工学ハンドブック．国際文献社. 東京. 2012, pp39-84.

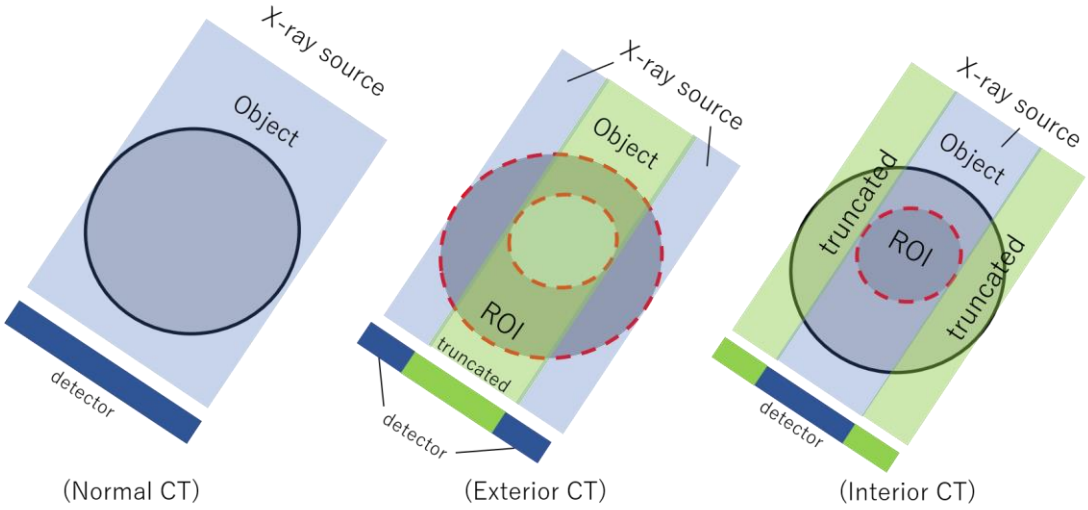


図 1 インテリア CT とエクステリア CT

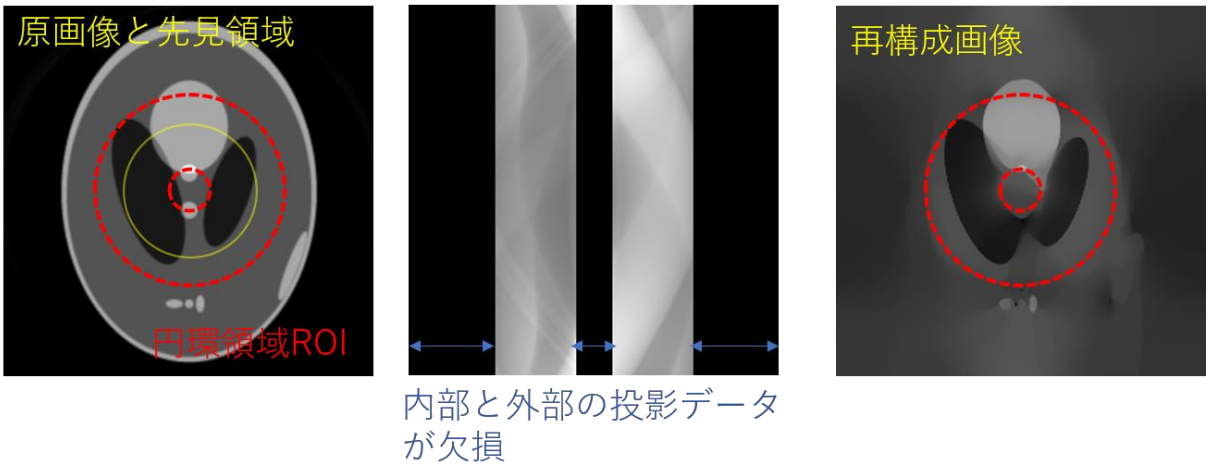


図 3 modified Shepp-Logan ファントム

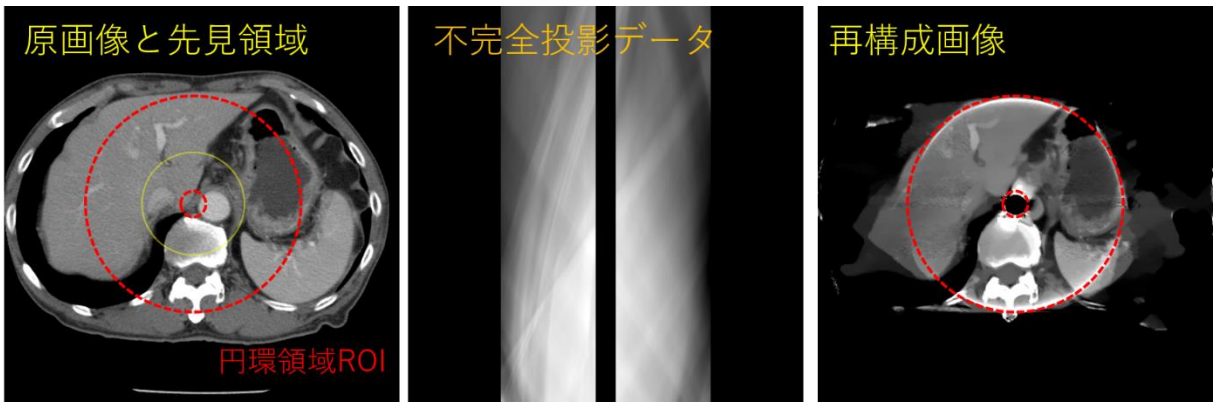


図 4 医療用腹部 CT 画像

Uniqueness of Solutions to Mixed Interior and Exterior Problem in CT Image Reconstruction

Katsuya FUJII[*1], Yongchae Kim[*1], Hiroyuki Kudo[*1]

*1 University of Tsukuba

The Reconstruction method from incomplete data is a very important research in the field of CT image reconstruction. In general, reconstruction problems from incomplete projection data are well known for the limited angle problem, the exterior problem, and the interior problem. Until recently, the theory of solution uniqueness and stability has been constructed for each problem. In this paper, we consider the uniqueness of the solution when these problems are mixed. In order to verify the obtained uniqueness results, a simulation experiment was performed using a modified Shepp-Logan phantom and an abdominal CT image.

Key words: CT, Image Reconstruction, Interior CT, Exterior CT

Unsupervised 3D Super-resolution of Clinical CT Volumes by Utilizing Multi-axis 2D Super-resolution

Tong ZHENG[*1], Hirohisa ODA[*1], Masahiro ODA[*1],
Shota NAKAMURA[*2], Masaki MORI[*3],
Hirotsugu TAKABATAKE[*4], Hiroshi NATORI[*5],
and Kensaku MORI[*1,6,7]

Abstract

This paper proposes a novel, unsupervised super-resolution (SR) approach for performing the SR of a clinical CT volume into the resolution level of a micro CT (μCT) using clinical CT-μCT training dataset. The precise non-invasive diagnosis of lung cancer typically utilizes clinical CT data for making surgical plans. Due to the resolution limitations of clinical CT (about 500 μm×500 μm×500 μm / voxel), it is difficult to obtain enough pathological information such as the cancer invasion area at alveoli level. On the other hand, μCT scanning allows the acquisition of volumes of lung specimens with much higher resolution (50 μm×50 μm×50 μm / voxel or higher). Thus, SR of clinical CT volume may be helpful for diagnosis of lung cancer. An unsupervised SR method named SR-CycleGAN was proposed for performing SR of clinical CT images. However, this method could only perform SR of 2D clinical CT images in the axial plane: the input is a 2D clinical CT image, and corresponding output is a 2D SR image. For obtaining more pathological and spatial information from the SR result from 3D space, a method that could perform SR of 3D clinical CT volumes is desired. In this paper, we propose an SR method that could perform SR of 3D clinical CT volume, by combining 2D SR results of clinical CT images in axial, coronal and sagittal plane, respectively. 2D SR of each axis was performed by an SR-CycleGAN. Experimental results demonstrate that the proposed method can successfully perform SR of 3D clinical CT volumes of lung cancer patients to μCT level resolution in 3D scale, while the previous SR-CycleGAN could only perform SR of 2D clinical CT images.

Keywords：Unsupervised super-resolution, fine anatomical structure analysis, inter-modality conversion

1. Introduction

Lung cancer is now the most common cancer among men worldwide [1]. Pre-surgery non-invasive diagnosis of lung

*1 Graduate School of Informatics, Nagoya University
〔Furo-cho, chikusa-ku, Nagoya, 464-8601, Japan〕
e-mail: tzheng@mori.m.is.nagoya-u.ac.jp
*2 Nagoya University Graduate School of Medicine
*3 Sapporo-Kosei General Hospital
*4 Sapporo Minami-sanjo Hospital
*5 Keiwakai Nishioka Hospital
*6 Information Technology Center, Nagoya University
*7 Research Center for Medical Big Data, National Institute of Informatics

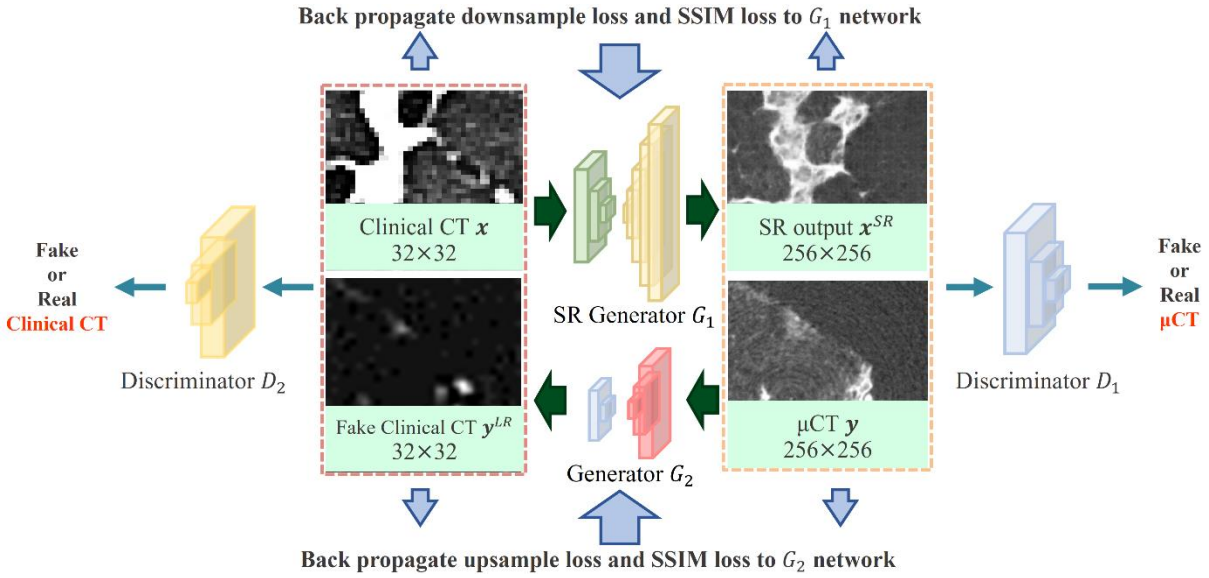


Fig. 1 Structure of SR-CycleGAN [2]. For training SR-CycleGAN, we use clinical CT images of size 32×32 pixels and μCT images of 256×256 pixels. SR generator G_1 is the generator that could generate SR image from a given clinical CT image. Generator G_2 is a generator that could generate clinical CT-like images from μCT images. For inference, we input clinical CT images into SR generator G_1 to obtain SR images.

cancer mainly depends on clinical computed-tomography (clinical CT) volumetric images (volumes). However, it is infeasible to observe enough pathological information such as the invasion of cancer due to low resolution (about 500 μm×500 μm×500 μm / voxel) of clinical CT volumes. The micro-focus X-ray CT (μCT) is another CT modality. Isotropic resolution of μCT volumes is usually 50 μm×50 μm×50 μm / voxel or higher. However, μCT scanner cannot perform in-vivo scan. A method named SR-CycleGAN [2] for performing SR of clinical CT volume into μCT level was proposed.

SR-CycleGAN is an unsupervised SR method, which does not require aligned pairs of low-resolution (LR) and high-resolution (HR) images for training. However, this method could only perform SR on 2D clinical CT images, meaning the input is a $n \times n$ pixels 2D clinical CT image (n is a positive integer) and corresponding output is a $8n \times 8n$ μCT-like 2D SR image. Since clinical CT images and μCT are both scanned with 3D scanning devices, and both are 3D volumes with size of $x \times y \times z$ voxels (x, y, and z are positive integers), we could not obtain enough anatomical information such as the 3D structure of bronchus and invasive cancer from 2D SR images. For obtaining anatomical information at 3D scale such as bronchiole's shape and the invasion condition of tumor, an SR method that could perform 3D SR of clinical CT volumes is desired. In this paper, we modify SR–CycleGAN to an unsupervised 3D SR method by combining multi-axis 2D SR result, which could perform SR of clinical CT volumes into μCT level.

2. Method

In prior to inference, training of the network is required using clinical CT and μCT volumes. Input clinical CT patches are of size 32×32 pixels, cropped from the axial, coronal and sagittal planes in CT volumes, respectively. Input μCT

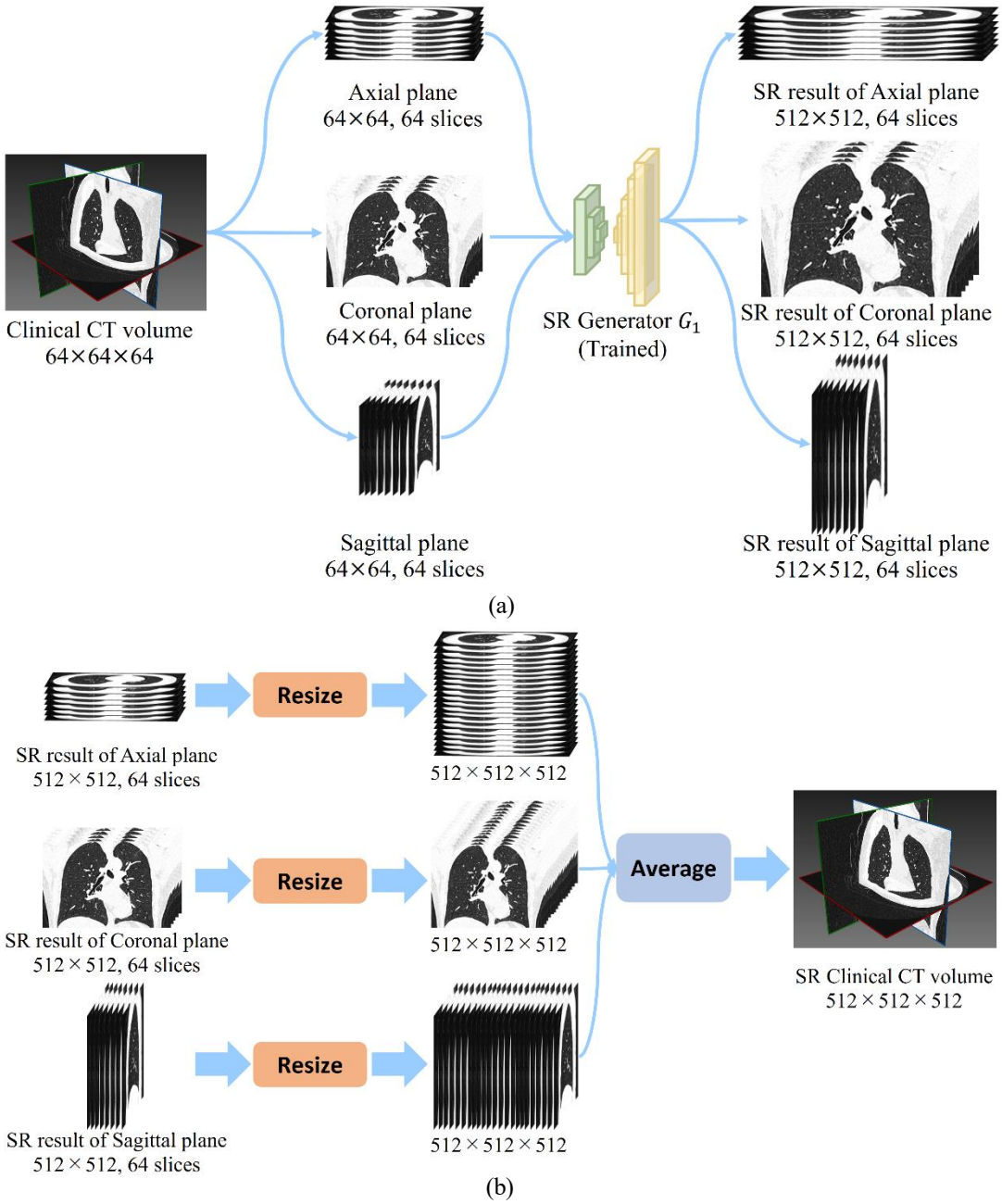


Fig. 2 Workflow of proposed method. (a) Performing multi-axis 2D SR of a given clinical CT volume. Utilizing the trained generator G_1 of SR-CycleGAN, we input images of axial, coronal and sagittal planes into the G_1. Correspondingly, we obtain SR images of each plane. (b) Combining multi-axis 2D SR result to 3D SR clinical CT volume. We utilize tricubic interpolation [3] to upscale the SR result with a scale of 8. For instance, input volume is of size of 512×512×64 voxels, we then obtain a volume of 512×512×512 voxels after tricubic interpolation. After tricubic interpolation of SR results of axial, coronal and sagittal planes, we sum the result by each voxel. Finally, we obtain SR result of 512×512×512 voxels, comparing to the size about input clinical CT volume (64×64×64 voxels).

patches are of 256×256 pixels, cropped from each plane of μCT volumes. In our study, scale of original μCT volumes is at least 8-times larger than that of the clinical CT volumes. Because of this, we consider 8-times SR to be the most

proper. Input of our networks are 2D patches cropped from the clinical volumes. Corresponding output SR result size is 256×256 pixels. Note that we input clinical CT patches of the axial, sagittal and coronal planes into the same trained generator G_1, and obtain corresponding SR results of axial, sagittal and coronal planes, respectively. As the final step, we combine the SR results from axial, coronal and sagittal planes to output an SR result.

1) Overview of the previous method: SR-CycleGAN

SR-CycleGAN [2] is a method to perform SR of clinical CT images into μCT level. SR-CycleGAN is summarized as follows: (1) a fully convolutional network (FCN)-based generator in CycleGAN performs SR of LR images into HR images. Another FCN-based generator shrinks HR images into LR images; (2) a convolutional neural network (CNN)-based discriminator classifies a given image into a real HR image or an SR result from LR image. Another CNN-based discriminator classifies a given image into a real LR image or shrunk from an HR image; (3) additional loss terms named downsample loss, upsample loss and SSIM loss [2] are added to enhance the stability of training SR-CycleGAN. The structure of SR-CycleGAN is shown as in Fig. 1.

However, as mentioned in the introduction section, the SR-CycleGAN could only perform SR of 2D clinical CT images. For observing anatomical information from 3D scale, based on SR-CycleGAN, we propose a method that could perform SR of a given clinical CT volume by combining multi-axis 2D SR result.

2) 3D SR of clinical CT volumes by combining multi-plane 2D super-resolution result

Based on SR-CycleGAN, we perform 3D SR of clinical CT volumes by combining multi-axis SR results. This is conducted by the following steps: (1) we train an SR-CycleGAN with 2D clinical CT and μCT patches cropped form axial, coronal and sagittal planes; (2) with the trained SR-CycleGAN, we input 2D clinical CT images from axial, coronal and sagittal planes to obtain SR results of each plane; (3) with SR results of clinical CT images about axial, coronal and sagittal planes, we combine SR results of each plane to obtain the final SR result. The approach about how to combine the result will be illustrated in the next paragraph.

3) Combine multi-plane 2D SR results to 3D SR results

As explained in the previous paragraph, we obtain SR results about axial, coronal and sagittal planes of a given clinical CT volume. For instance, with a given clinical CT volume of 64×64×64 voxels, we obtain an SR result from this clinical CT volume as the following steps: (1) we perform SR of clinical CT volume using 2D SR-CycleGAN in axial plane. The SR result is a volume of 512×512×64 voxels; (2) similarly, we perform SRs in the coronal and sagittal planes. SR results are volumes of 64×512×512 and 512×64×512 voxels, respectively; (3) we combine SR results of each plane. For SR result from the axial plane, we use tricubic interpolation, to interpolate the SR result from 512×512×64 voxels into a volume of 512×512×512 voxels. For SR results from the coronal and sagittal planes, similarly we use tricubic interpolation to interpolate these SR results into volumes of 512×512×512 voxels. At last, we compute the average of the three interpolated SR result of each plane as:

$$SR = \frac{SR_A + SR_C + SR_S}{3}, \quad (1)$$

where SR is the 3D SR result (as the final result), and SR_A, SR_C and SR_S are SR results (resized by tricubic interpolation) of axial, coronal and sagittal planes, respectively. The whole workflow is shown in Fig. 2. The SR result will be illustrated in the next section.

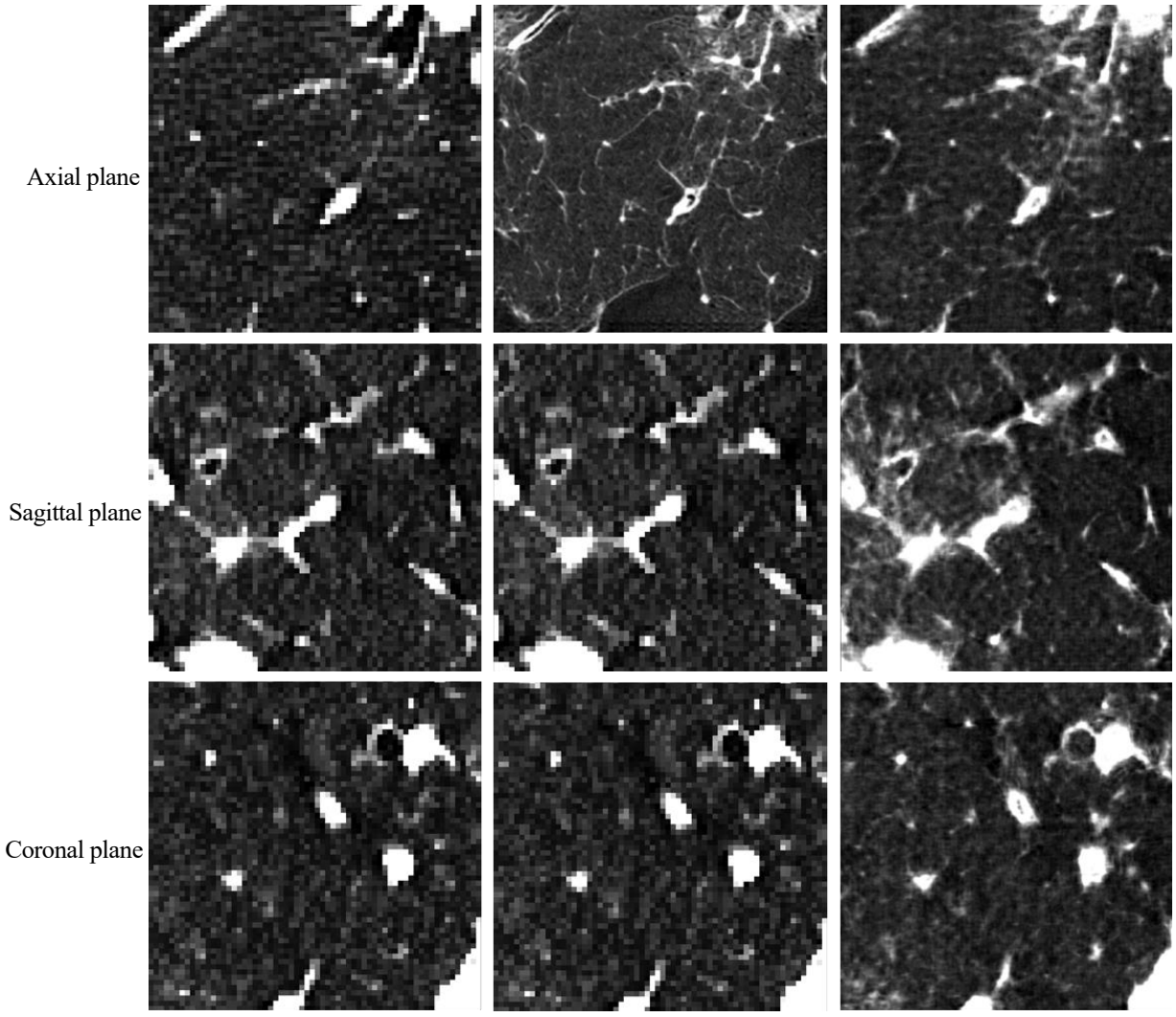


Fig. 3 SR result of proposed method and the previous method (SR-CycleGAN). (1) Original CT image; (2) SR result of SR-CycleGAN [2]; (3) SR result of our method. Top, middle and bottom rows are images cropped from axial, sagittal and coronal planes respectively. Proposed method could successfully perform SR of 3D clinical CT volume, while the previous method (SR-CycleGAN) could only peform SR of 2D clinical CT images in axial plane.

3. Experiments and results

We trained the proposed method on five clinical CT volumes and their five corresponding μCT volumes. We evaluated the proposed approach qualitatively on one clinical CT volume. The clinical CT volumes were scanned by a clinical CT scanner (SOMATOM Definition Flash, Siemens Inc., Munich, Germany). The resolution of the clinical CT volume was 0.625 mm×0.625 mm×0.6 mm / voxel. The μCT volumes were scanned by a μCT scanner (inspeXio SMX-90CT Plus, Shimadzu Inc., Kyoto, Japan). The lung cancer specimens from the studied patients were fixed by Heitzman's method [4] and scanned with isotropic resolutions in the range of 42 μm×42 μm×42 μm ~ 52 μm×52 μm×52 μm / voxel.

In the training phase, we extracted 2000 patches of size 32×32 pixels randomly from the lung region of each volume. Since we have five clinical CT volume for training, the total number of patches was 10000. We also extracted 2000 patches of size 256×256 pixels randomly from each μCT volume. We performed 8-times SR from clinical CT volumes to μCT-level, which means the SR result was 8-times larger by length, width and height. All programs were implemented using PyTorch on NVIDIA GeForce GTX TITAN X (12 GB memory).

SR results of our proposed method were compared to original clinical CT images, as shown in Fig. 3. We could obtain more details from SR results than the original images. Lung anatomies, such as the bronchus looks more clearly than the original image. Compared to the previous method (SR-CycleGAN), SR-CycleGAN could only perform SR of clinical CT images in the axial plane; while proposed method successfully performed SR of 3D clinical CT volume in axial, sagittal and coronal planes.

4. Discussion

Quantitative evaluation of proposed method remains a problem. In conventional SR methods, quantitative evaluation is usually conducted by comparing SR and original image pairs. However, it is infeasible to obtain such pairs between clinical CT and μCT volumes, as also mentioned in Introduction section. We consider solving this by making paired clinical CT – μCT images, by scanning clinical CT image of lung specimen using clinical CT scanner.

The proposed method could perform SR of clinical CT volumes, while the previous method (SR-CycleGAN) could only perform SR of 2D clinical CT images. On the other hand, as shown in column (a) in Fig. 3, arteries in the SR result of SR-CycleGAN have some holes, while proposed method avoided this.

5. Conclusion and future work

We proposed an SR method that could perform SR of 3D clinical CT volumes, based on combining 2D SR results about axial, coronal and sagittal planes of a given clinical CT volume. Proposed method overcomes the problem of performing 3D SR of clinical CT volumes by combining multi-axis 2D SR result. We demonstrated proposed method is useful for the SR of lung clinical volumes into μCT scale.

Future work includes quantitative evaluation of the proposed methods. Since it is infeasible to obtain paired HR (μCT) and LR (clinical CT) data, we could not evaluate the result quantitatively directly. We could downsample a μCT volume to clinical CT scale, and evaluate how well the proposed method could rescale the downsampled μCT volume to its original scale.

Competing interests

None

Acknowledgement

Parts of this research was supported by MEXT/JSPS KAKENHI (26108006, 17H00867, 17K20099), the JSPS Bilateral International Collaboration Grants, the AMED (JP19lk1010036 and JP20lk1010036) and the Hori Sciences

&Arts Foundation.

References

［1］Rafiemanesh H, Mehtarpour M, Khani F, et al.: Epidemiology, incidence and mortality of lung cancer and their relationshipwith the development index in the world, Journal of Thoracic Disease **8**(6): 1094-1102, 2016

［2］Zheng T, Oda H, Moriya T, et al.: Multi-modality super-resolution loss for GAN-based super-resolution of clinical CT images using micro CT image database, Proc. SPIE 11313, Medical Imaging 2020: Image Processing **1131305-1-7**, 2020

［3］Lekien F, Marsden J: Tricubic interpolation in three dimensions. International Journal for Numerical Methods in Engineering **63**(3): 455-471, 2005

［4］Heitzman E: The lung: radiologic-pathologic correlations, C.V. Mosby Company, Maryland Heights, 1993.

多方向 2 次元超解像を用いた臨床 CT 像の教師なし 3 次元超解像

鄭 通[*1], 小田 紘久[*1], 小田 昌宏[*1], 中村 彰太[*2],
森 雅樹[*3], 高畠博嗣[*4], 名取 博[*5], 森 健策[*1,6,7]

*1 名古屋大学大学院情報学研究科
*2 名古屋大学大学院医学系研究科
*3 札幌厚生病院
*4 札幌南三条病院
*5 恵和会西岡病院
*6 名古屋大学情報基盤センター
*7 国立情報学研究所医療ビッグデータ研究センター

本稿では, 臨床用 3 次元 CT 像の教師なし超解像手法を提案する. 肺がん症例の臨床 CT 像から腫瘍の浸潤状況など疾患に関する情報を取得するため, 肺臨床 CT に超解像を適用し, μCT レベルの解像度を得る手法が求められている. 従来手法である SR-CycleGAN は, 対応関係のある低解像度と高解像度の画像ペアからの学習が不要の超解像手法であるが, Axial 断面の 2 次元臨床 CT 像の超解像のみを行っており, 3 次元的な構造に不自然な部分があった. 臨床 CT 像は通常 3 次元画像であり, 多断面から詳細な病理情報を得るには、臨床 CT 像に 3 次元の超解像を適用し, μCT レベルの解像度を得る手法が求められている. 我々は臨床 CT 像の Axial, sagittal, coronal それぞれの断面に SR-CycleGAN を用いた 2 次元超解像を適用し, それぞれの断面の 2 次元超解像結果を統合することによって 3 次元臨床 CT 像の超解像を実現した. 実験の結果, 3 次元臨床 CT 像の μCT レベルへの超解像が可能であった.

キーワード：教師なし超解像, 微細解剖学構造分析, モダリティ変換

Ring artifact removal for refraction-contrast CT via generative adversarial network

Zhuoran HUANG[*1], Naoki SUNAGUCHI[*1], Daisuke SHIMAO[*2], Shu ICHIHARA[*3], Tetsuya YUASA[*4], Masami ANDO[*5]

Abstract

Refraction-contrast X-ray CT (RCT), which provides a high-contrast image of biomedical soft tissues, has recently been used in pathology and micro anatomy research. One of the characteristics of RCT is that the projection obtained by RCT system represents the first differentiation of the object's radon transformation. Therefore, the reconstruction of RCT involves an additional integration process of the projection before conventional back-projection. However, this process causes the noise generated by the detector or refraction angle estimation errors to propagate throughout the projection image, which leads to increase ring artifacts (RA) on the reconstructed image. Since various methods for removing RA have been developed in the conventional absorption-based CT field, it is difficult to eliminate the RA increased by integration as RCT.

In this research, we propose RA removal method based on conditional generative adversarial network (cGAN) incorporated the sparse property of Laplacian CT image. Laplacian image of RCT are easily obtained by reconstructing the differential projection. To demonstrate the effectiveness of this method, we applied the proposed and conventional methods to simulation data of numerical phantoms and compared them by the root mean square error and the structural similarity. In the result, the proposed method demonstrated to remove RA better than conventional methods.

Keywords：refraction-contrast CT, ring artifact, machine learning, GAN

1. Purpose

Refraction-contrast X-ray CT (RCT) is an imaging method for constructing density-distribution images by detecting the refraction angle of X-rays with an analyzer crystal, taking advantage of the fact that X-rays are refracted according to the density difference when they pass through a sample [1]. However, if there are scratches or dusts on the crystal surface, they lead to angular estimation errors and result in ring artifacts (RA), which also appears in absorption contrast X-ray CT (ACT) caused by non-uniform sensitivity of the detector. There are lots of RA removal methods that applied to ACT, include the image filter method proposed by Prell [2], the RTV-based iterative method proposed by Liang [3], and machine

*1 Graduate School of Medicine, Nagoya University
〔1-1-20 Daiko-Minami, Higashi-ku, Nagoya City, Aichi Prefecture, 461-8673, Japan〕
e-mail: huang.zhuo.ran@j.mbox.nagoya-u.ac.jp
*2 Faculty of Health Sciences, Hokkaido University of Science
*3 Clinical Research Center, Nagoya Medical Center
*4 Faculty of Science, Yamagata University
*5 Comprehensive Research Organization for Science and Society

learning methods such as generative advertiser network [4] can also be used in image denoising. Regardless of whether it is a non-machine learning method or a machine learning method, the results are not satisfactory when applied to RCT because the reconstruction of RCT involves an integration process in the projection, with which the angular errors spread over the entire projection. In this study, we proposed a machine learning method approach to remove ring artifacts in RCT based on a kind of conditional generative adversarial network (cGAN) proposed by Isola et al [5] in 2017.

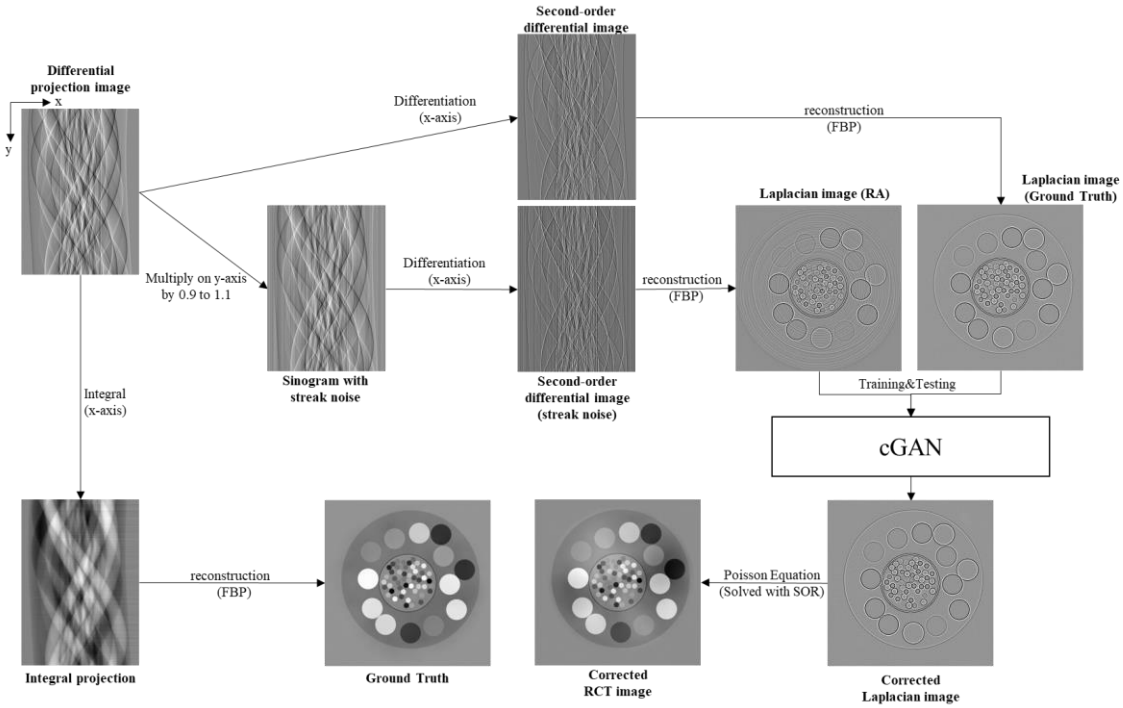


Fig. 1 Flowchart of training process for ring artifact removal

2. Method

The training process for RA removal with cGAN is shown in Fig. 1. First, we prepared 3000 sets of numerical phantoms containing random circles with different refractive index δ, radius, and center. Considering the case where the sample is ring-shaped, a ring-shaped structure with randomly varying radii is added to the center of the phantom. Each phantom is measured using a refraction X-ray CT imaging simulator based on X-ray beam tracking developed by our previous research [6], where the number of projections is 720 from 0 to 180 degrees. The sinogram obtained from the simulation is the first derivative of the line integral of δ as shown in an example of upper left corner of Fig. 1, where horizontal and vertical directions show the array of detector elements and the angle of incidence of X-rays to the sample, respectively. RCT is generally reconstructed by applying FBP after integrating the sinogram in the horizontal direction. In this method, however, Laplacian RCT obtained by differentiating instead of integrating is used in order to take advantage of a benefit of Laplacian RCT described below. The training with cGAN proceeds using sets of this ground truth Laplacian RCT and a Laplacian RCT with RA reconstructed from a sinogram containing noise that causes ring artifacts. As for loss function of cGAN, we add the sum of the pixel values of the generated CT in addition to the general L2 norm of the

generated and ground truth CTs, which takes advantage of the fact that the sum of the pixel values of the Laplacian RCT is theoretically zero. This regularization prevents excessive flattening of pixel values such as total variation regularization and improves the denoising effect in areas where RA is dense, such as central of an image. The network was implemented in Python3.6 and TensorFlow. The GPU for training is Nvidia GTX 2080Ti and RAM is 256 GB. In the process of training, we set the training rate = 0.0002 and iterations = 200,000.

After obtaining the Laplacian RCT with the artifacts removed, we can convert the Laplacian RCT to the normal RCT, thus δ image, by solving the Poisson equation with successive over relaxation (SOR) method. To evaluate the effectiveness of this method, we compared the results of the proposed method, image filtering-based method by Prell et al [2] and iterative method by Liang et al [3] using 100 images not used for training. As an evaluation method, we used the root mean squared error (RMSE) and the structural similarity (SSIM) between the images to which each method was applied and the ground truth image.

3. Results

The RMSE and SSIM were 0.2358 and 0.9228 in the Prell method, 0.2588 and 0.9279 in Liang's Iterative method, and 0.2043 and 0.9438 in the proposed method, respectively. The proposed method was the best on both indicators.

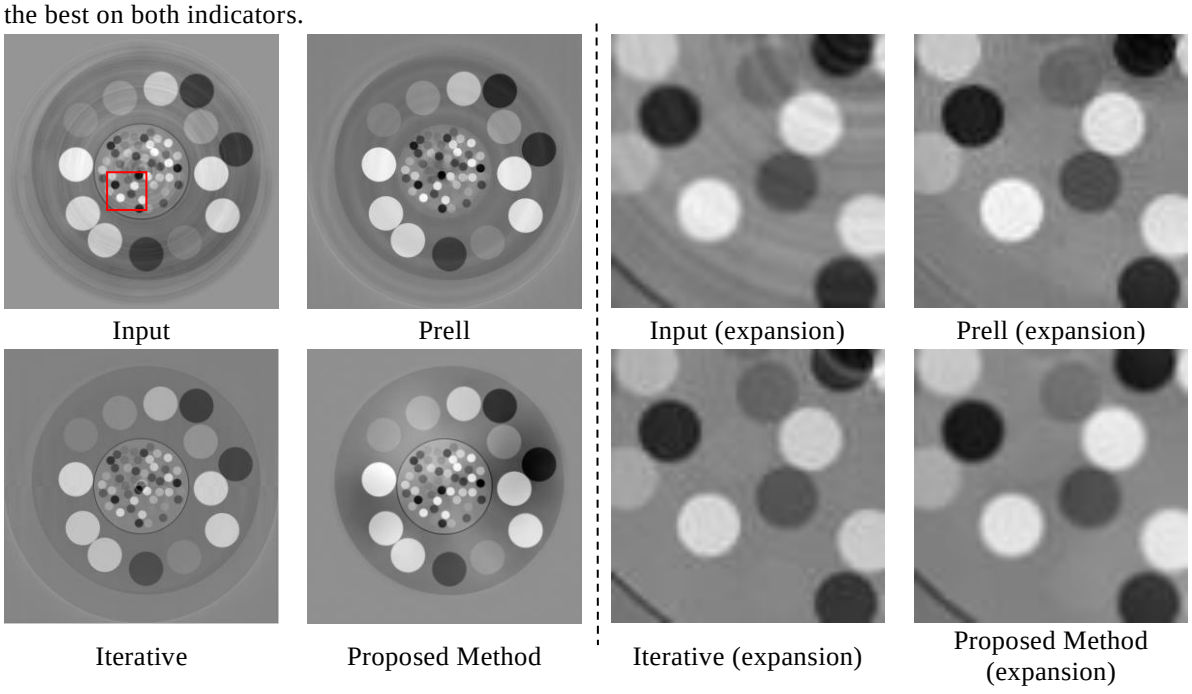


Fig. 2 Results of different methods on simulated data. The 4 images on the right side are the images on the left side on the same place zoomed in.

By observing the images of RA removal results of these methods, as shown in Fig. 2, the Prell method has a large amount of RA remain and a ring-shaped structure can be mistakenly removed. The Iterative method is better than Prell method, but it doesn't completely remove the RA. The proposed method succeeded in completely eliminating the RA and a ring-shaped structure not RA remains clearly. However, the only disadvantage is the overall uniformity of the contrast observed, which appears when converting to δ image using the SOR method due to a small error in the Laplacian RCT.

4. Conclusion

In this study, we proposed a machine learning based ring artifact removal method that involves the physical properties of RCT where the sum of pixel value of the Laplacian RCT is theoretically zero. The proposed method shows more effective in removing RA than other methods. However, the contrast uniformity that appears on the results of proposed method needs to be solved. On the other hand, what is important is whether or not it is effective in actual biological tissues, and we plan to apply this method to various biological samples in the future.

Competing interests

The authors declare that they have no conflict of interest.

References

[1] Sunaguchi N, Yuasa T, Huo Q et al: X-ray refraction-contrast computed tomography images using dark-field imaging optics. Applied Physics Letters **97**, 2010

[2] Prell D, Kyriakou Y, Kalender W A: Comparison of ring artifact correction methods for flat-detector CT. Physics in Medicine & Biology **54**: 3881-3895, 2009

[3] Liang X, Zhang Z, Niu T et al: Iterative image-domain ring artifact removal in cone-beam CT. Physics in Medicine & Biology **62**: 5276-5292, 2017

[4] Goodfellow I J, Pouget-Abadie J, Mirza M et al: Generative adversarial nets. In Proceedings of NIPS, 2672-2680, 2014

[5] Isola P, Zhu J Y, Zhou T et al: Image-to-Image translation with Conditional Adversarial Networks. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR): 1125-1134, 2017

[6] Sunaguchi N, Yuasa T, Sun F et al: Limited view reconstruction for differential phase-contrast computed tomography. Optics Express **23**: 9717-9729, 2015

Transformation from Hematoxylin-Eosin Stain to Ki-67 Immunohistochemistry Stain Image Using Deep Learning

Cunyuan JI[*1 †], Takeshi UEHARA[*2], Fumikazu KIMURA[*3], Kengo OSHIMA[*3], Masahiro YAMAGUCHI[*1]

Abstract

Whole slide imaging (WSI) system produces high resolution digital images from pathology specimens, which is the key to modern digital pathology. Digitized microscopic images greatly facilitate the archiving, transferring, and evaluation of pathology specimens, making the diagnosis faster and more reliable. In the diagnosis of cancers, tissues are commonly processed with the hematoxylin-eosin (H&E) stain, also known as general stain, to reveal the morphology, color, and texture features. Besides, the molecule level activity of the cells, which is utilized in the assessment of malignancy or the classification of cancer subtypes, is important. Therefore, immunohistochemical (IHC) stain method is introduced to mark specific proteins in the cells. For instance, the Ki-67 protein is one of the popular biomarkers for evaluating the growth factor of a cell population. In practice, the pathologists evaluate the Ki-67 signals with IHC stain images to determine the proliferation activity of the tumor cells and rate the malignancy of cancer accordingly. However, IHC stain suffers from its complicated staining procedure and high cost compared with the HE stains. In this paper, we propose a digital staining method that employs the U-Net to obtain the semantics of Ki-67 positive and negative nuclei and utilizes the color unmixing technique to yield stain colors. WSIs of well-differentiated endometrial adenocarcinoma of the uterine corpus from Shinshu University Hospital were used. We scanned the H&E stained specimens with a NanoZoomer scanner under 40x objective lens, then perform the IHC reaction for Ki67 in decolorized HE specimens. The classification accuracy of semantic areas reaches 82%, with a 48% precision and a 62% recall for the positive areas. To evaluate the quality of digital staining in terms of diagnostic metrics, we consider the positive rate of all nuclei, also known as labeling index. An end-to-end technique based on cycle-GAN is also tested for comparison.

Keywords：whole slide image, digital stain, semantic segmentation, deep learning

1. Introduction

In today's pathology diagnosis, whole slide image plays a crucial role, where the pathological specimens taken from patients are archived with microscopic scanners as high-resolution images. With WSI as a digital medium, it is possible to analyze the specimen automatically with image processing software and save the labor of pathologists. To emphasize biology markers for certain diseases, the conventional way is to apply physical stain using chemicals, while digital staining methods recognize different regions in the WSIs with image processing algorithms and color the image

*1 Department of Information and Communications Engineering, Graduate School of Engineering, Tokyo Institute of Technology〔Nagatsuta 4259-G2-28, Yokohama, Kanagawa 226-8503, Japan〕
ji.c.ac@m.titech.ac.jp

*2 Department of Laboratory Medicine, School of Medicine, Shinshu University

*3 Department of Biomedical Laboratory Sciences, School of Health Sciences, Shinshu University

† Corresponding author.

accordingly. The common routine for physical staining is the H&E stain, which reveals the morphology and texture features of cells. However, to probe the expression of certain proteins, IHC stain, which is relatively expensive and laborious, is necessary in many cases. If we may find a method to produce the IHC digital stain image from its H&E counterpart, then the additional cost of physical stain would be saved.

In this paper, we discuss the digital IHC staining of Ki-67 protein, which appears in the $G1$, S, $G2$, and M phases of a cell cycle, while the cells in the quiescence phase do not express it. Therefore, in the diagnosis of cancers, the pathologists quantitate the Ki-67 signals to determine the proliferation activity of the tumor cells and rate the malignancy of cancer accordingly[1]. Kimura et al showed that the Ki-67 positive nuclei are predictable from the HE stain by analyzing the color, texture, and morphology features[2]. Therefore, we used a deep learning method to predict the Ki-67 positive regions in a specimen and utilized the color-unmixing[3] technique to manipulate the colors.

2. Method

The schematic of our method is shown in fig. 1. Input to the system is an H&E stained image, and the output is the Hematoxylin-DAB staining of the same specimen. We used a U-Net[4] to segment the Ki-67 positive nuclei that should be stained with DAB. To make the digital staining natural, we multiplied the output mask with the original H channel, which is obtained using the color unmixing technique, as the prediction of the DAB stain component. This DAB component is then remixed with the input hematoxylin component to generate the digital staining. Fig. 2 shows the training scheme of the U-Net. We used the binarized DAB component unmixed from the manually stained IHC as the label of Ki-67 positive nuclei. The loss function is the summation of binary cross entropy and Jaccard loss.

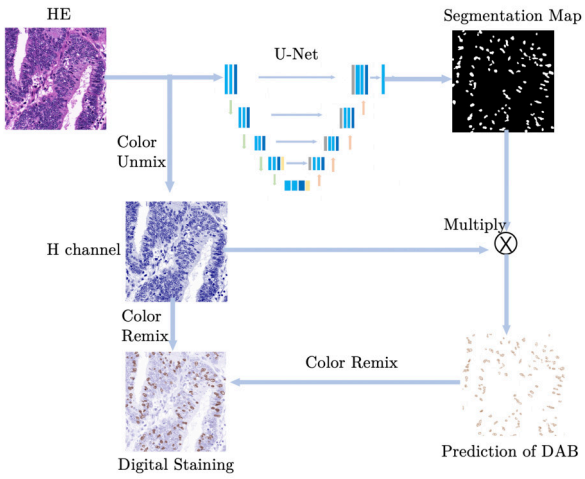


Fig. 1 Schematic of digital staining

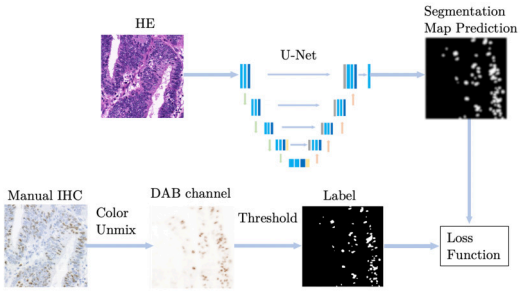


Fig. 2 U-Net training scheme

3. Result

All the images in our experiment are from two pairs of H&E and IHC WSIs of endometrial adenocarcinoma specimens of the uterine corpus. The acquisition of all the specimens in this research was approved by the Committee for Medical Ethics of Shinshu University School of Medicine. Slides were scanned with a Hamamatsu NanoZoomer scanner at 40x magnification. The WSIs were cropped into 256×256 patches, where 14480 patches were used for training and 1280 patches were for validation. Using the configuration explained in section 2, we trained the U-Net for 200 epochs. The intersection over union (IOU) of predicted positive nuclei regions converges to a high level of over 80% on the training set. The pixel-wise receiver operating characteristic (ROC) of the validation set is shown in fig. 3, where the area under curve is 0.917. As is shown in table 1, the pixel-wise accuracy on validation set reaches 0.95. The average validation IOU is 0.326.

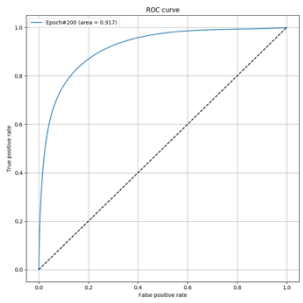


Fig. 3 ROC in the validation

To evaluate the performance of our method by quantitating the nuclei approximately rather than pixels, we quantified the connected regions in the label that overlap with the prediction without partitioning the nuclei. The evaluation result is shown in table 2, proposed model gives an 82% overall accuracy on the validation set. For the positive nuclei, the precision and recall are 0.48 and 0.62. To check the agreement of the digital and manual stain, we also used ImmunoRatio[5] to calculate the area ratio of Ki-67 positive regions. A paired sample t test on seventeen 1024×1024 patches in the validation set shows that the area ratios derived from the digital stain and the manual stain have a +2.29% bias, which indicates that the false positive rate needs to be alleviated. We will present the agreement evaluation result of labeling index[6], defined as the ratio of positive nuclei number and all nuclei number in the tumor tissue, on the conference.

Fig. 4 shows an example of our digital staining result. For comparison, in fig. 5, we also present the result from a popular end-to-end generative architecture, the Cycle GAN[7], [8]. This method preserves the textures and colors better, while our method has more precise semantics on the positive nuclei.

Table 1. Pixel-wise classification result

	Precision	Recall	F1-score
False	0.98	0.97	0.97
True	0.44	0.56	0.49
Accuracy	0.95		

Table 2. Classification result of connected regions

	Precision	Recall	F1-score
False	0.92	0.86	0.88
True	0.48	0.62	0.54
Accuracy	0.82		

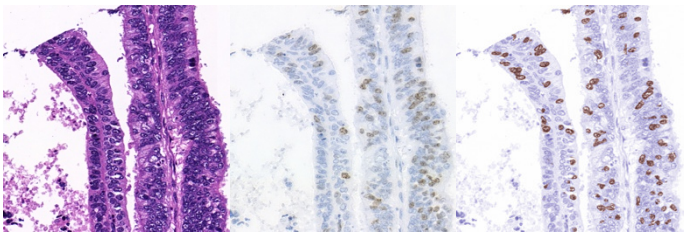


Fig. 4 From left to right: (1) manual H&E, (2) manual IHC, (3) digital stain

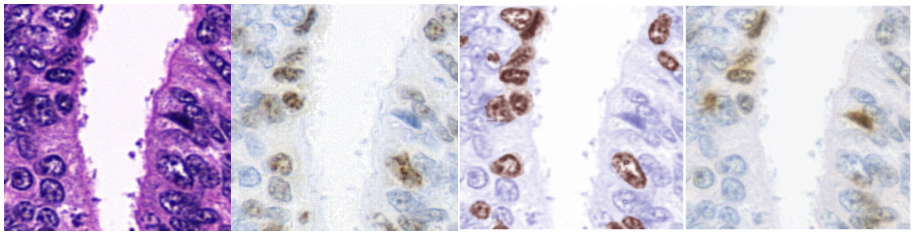


Fig. 5 From left to right: (1) manual H&E, (2) manual IHC, (3) proposed and (4) Cycle GAN

4. Conclusion

We show a method to perform digital IHC staining from HE stained images of endometrial adenocarcinoma. From the results above, we may conclude that there is a noticeable correlation between the metrics derived from the physical stain and our digital stain. Additionally, the results indicate that the nuclei in different proliferation statuses are distinguishable in the HE stained specimen without marking the protein signals. Interpretation of the significant features for recognizing the nuclei would be a valuable research direction in the future.

Competing interests

None.

References

[1] T. Scholzen and J. Gerdes, “The Ki-67 protein: From the known and the unknown,” J. Cell. Physiol., vol. 182, no. 3, pp. 311–322, 2000.

[2] F. Kimura et al., “Detection of Ki67 Expression by Analyzing Texture of Hematoxylin-and-Eosin--Stained Images, the Effectiveness of Signal Intensity, and Co-occurrence Matrix Features,” Anal. Quant. Cytopathol. Histopathol., vol. 40, no. 1, pp. 9–19, 2018.

[3] A. C. Ruifrok, D. A. Johnston, et al., “Quantification of histochemical staining by color deconvolution,” Anal. Quant. Cytol. Histol., vol. 23, no. 4, pp. 291–299, 2001.

[4] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” pp. 1–8, 2015.

[5] V. J. Tuominen, S. Ruotoistenmäki, A. Viitanen, et al., “ImmunoRatio: A publicly available web application for quantitative image analysis of estrogen receptor (ER), progesterone receptor (PR), and Ki-67,” Breast Cancer Res., vol. 12, no. 4, 2010.

[6] A. Y. Kwon et al., “Practical approaches to automated digital image analysis of Ki-67 labeling index in 997 breast carcinomas and causes of discordance with visual assessment,” PLoS One, vol. 14, no. 2, pp. 1–13, 2019.

[7] J. Y. Zhu, T. Park, P. Isola, et al., “Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks,” Proc. IEEE Int. Conf. Comput. Vis., vol. 2017-Octob, pp. 2242–2251, 2017.

[8] Z. Xu, C. F. Moro, B. Bozóky, et al., “GAN-based Virtual Re-Staining: A Promising Solution for Whole Slide Image Analysis,” pp. 1–16, 2019, [Online]. Available: http://arxiv.org/abs/1901.04059.

マルチピンホール X 線蛍光 CT 画像再構成における EM-TV アルゴリズムの検討

金澤　道和[*1]　保科　亮太[*2]　佐々木　洸介[*1]　阿部　智大[*1]

笹谷　典太[*2]　砂口　尚輝[*3]　河嶋　秀和[*4]　兵藤　一行[*5]

湯浅　哲也[*2]　銭谷　勉[*1]

要旨

マルチピンホール X 線蛍光 CT（multi-pinhole x-ray fluorescence computed tomography; mp-XFCT）では，ある化合物に導入した非放射性核種を入射 X 線で励起し，発生する蛍光 X 線を多方向から計測することで，被写体内部の化合物の 3 次元分布を得る．しかし，現行の mp-XFCT システムで生体撮像を行うためには, 測定の対象となる化合物に導入されるヨウ素濃度の検出限界が不十分であり, 改善が必要である. そこで本研究では, 画質改善による検出限界の向上を目的とし，トータルバリエーション（TV）を正則化項として組み込んだ EM-TV アルゴリズムを用いた再構成法を提案する. 評価実験では物理ファントムを用意し, その実測データを従来法である OS-EM 法および EM-TV アルゴリズムで画像再構成し，再構成画像を比較した．EM-TV ではノイズが抑制されたことにより，最小濃度である 0.05 mg/ml のヨウ素領域が明瞭に描出された．また，コントラスト対ノイズ比（CNR）は約 1.3 倍向上し，ヨウ素の検出限界は OS-EM で 0.049 mg/ml であったのに対し，EM-TV では 0.036 mg/ml に改善した.

キーワード : X 線蛍光 CT，マルチピンホール，画像再構成，トータルバリエーション

*1 弘前大学大学院理工学研究科
〔〒036-8561 青森県弘前市文京町 3〕
e-mail: **ms19507@eit.hirosaki-u.ac.jp**
*2 山形大学大学院理工学研究科
*3 名古屋大学大学院医学系研究科
*4 京都薬科大学放射性同位元素研究センター
*5 高エネルギー加速器研究機構物質構造科学研究所

1. はじめに

XFCT （ X-ray Fluorecence Computed Tomography ） は ， SPECT （ Single Photon Emission Computed Tomography ） や PET （Positoron Emission Tomography）といった放射型 CT に分類され，非放射性核種を外部から照射する X 線で励起し，脱励起の際に発生する蛍光 X 線を多方向から計測することで，被写体内部の化合物の 3 次元分布を得る．SPECT や PET と比較した XFCT の特色とし

て，機能情報の高感度な検出，1mm 以下の高い空間分解能，高い定量性，非放射性造影剤による撮像が挙げられる．そのような特色から，小動物 in vivo イメージングに適し，新しい高分解能分子イメージングモダリティとして期待されている[1-4]．従来の FXCT では，ペンシルビーム入射によるポイントスキャン方式を採用していたため，検出効率が悪く，膨大な計測時間が必要であった．この問題を解決するため，Sunaguchi らはペンシルビーム入射を被写体の大きさに合わせたボリュームビーム入射に変更し 2 次元検出器と組み合わせると共に，被写体内に分布する非放射性核種（本研究ではヨウ素）から発生する蛍光 X 線を，ピンホールを用いてコリメートすることで 3 次元データを一挙に収集することができるようになった[5]．さらに，シングルピンホールをマルチピンホールに変更することで感度が大幅に改善され，より短時間での撮像が可能となった[6]．しかしながら，生体撮像を行うには薬理効果が生じない 0.05 mg/ml 以下のヨウ素を使用する必要があり，視覚的な診断を考慮すると，検出限界はそれよりもさらに低い濃度である必要があるが，現行のマルチピンホール XFCT システムでは不十分である．

近年，圧縮センシングと呼ばれる情報処理技術を応用した画質改善が注目され[7, 8]，推定画像の全変動（TV: Total Variation）[9]を評価関数に組み込んだ手法が医用工学分野における画像再構成に応用されている．放射型 CT の場合，ML-EM（Maximum Likelihood - Expectation Maximization）法と TV 最小化を組み合わせた EM-TV（Expectation Maximization - Total Variation）アルゴリズムが Swatzky らによって提案されている[10]．

本研究では，EM-TV アルゴリズムを用いた画像再構成による画質改善および，画質改善による検出限界の向上を目的とした．物理ファントムを対象としてマルチピンホール XFCT 撮像を行い，得られた投影データを従来法である OS-EM（Ordered Subsets - Expectation Maximization）法および EM-TV アルゴリズムで画像再構成した．得た再構成画像のコントラストノイズ比（CNR: Contrast Noise Ratio）および検出限界（LOD: Limit of Detection）を算出し，結果を比較した．

2. 方法

2-1. マルチピンホール XFCT 撮像

ピンホール XFCT システム[5]の構成図を図 1 に示す．

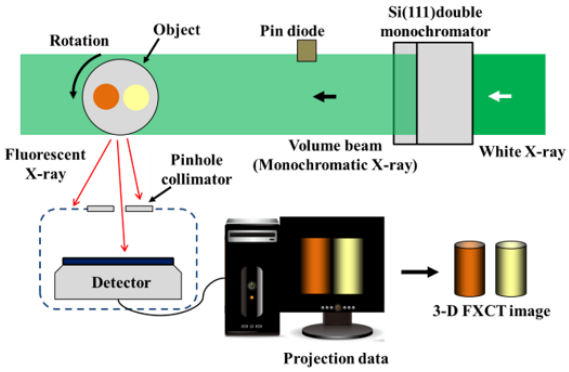


図 1 ピンホール XFCT システムの構成図．

マルチピンホール XFCT システムを，高エネルギー加速器研究機構 PF-AR の偏向電磁石ビームライン AR-NE7A に撮像システムを構築した．加速器からの白色 X 線は水平な偏光を持ち，Si(111)からなる 2 結晶モノクロメータにより単色 X 線化され，25～55keV の範囲の X 線をエネルギーとして扱うことができる．単色化された X 線は X 線スリットによって横 35 mm，縦 5 mm の立体ビームに成形され，X 線エネルギーを 30 keV とした場合，被写体直前における X 線強度は約 10^8 ～ 10^9 photo/mm^2/s となる．本実験では，X 線の入射エネルギーをヨウ素の K 吸収端エネルギー 33.17 keV の直上 33.2 keV に設定した．また，ピンホールの直径はそれぞれ 0.1 mm とし，被写体中心からピンホールおよび検出面までの距離は，それぞれ 29.1 mm および 38.2 mm とした．1 投影の撮像時間を 180 s とし，投影数を 180 とした．さらに，散乱線の低減のためにマルチピンホールコリメータの前面に厚さ 0.2 mm のスズ薄膜フィルタを設置した[11]．

本実験で使用した物理ファントムの概要を図 2 に示す. ファントムの直径は 10 mm とし, ヨウ素領域の孔の直径を 3 mm とした. 6 つの孔に濃度の異なるヨウ素溶液を封入した。

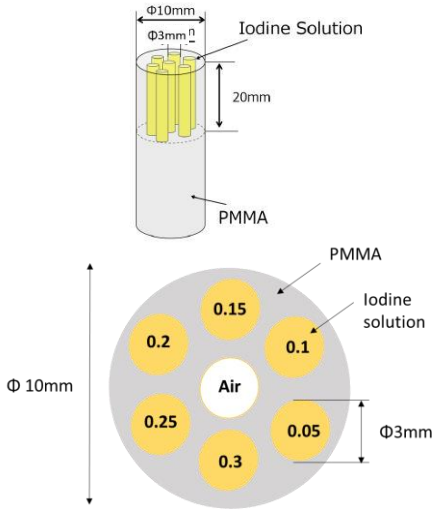


図 2 物理ファントムの概要. 下図中の数値はヨウ素濃度[mg/ml]を示す.

2-2. EM-TV アルゴリズムによる画像再構成

図 3 に, マルチピンホール XFCT の幾何学を示す. 被写体中の非放射性核種は X 線が照射されると, 光電効果による励起によって蛍光 X 線を放射する. ボクセルから発生した蛍光 X 線は, 入射 X 線に垂直な y 軸方向に配置された複数のピンホールを通り, その直線状にある 2 次元検出面上で検出され, 投影データを得る. 検出器上にある pi はピクセルで検出された光子数, 被写体にある Vj はボクセルのヨウ素濃度を表している.

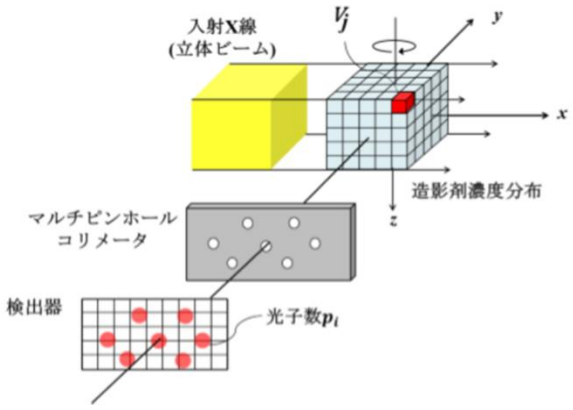


図 3 マルチピンホール XFCT の幾何学.

EM-TV アルゴリズムは, Sawatzky らによって提案された, 従来法である ML-EM 法と TV 最小化アルゴリズムを組み合わせた手法である[10]. 放射型 CT に分類される XFCT では, 観測データがポアソン分布に従うため, 以下の式を目的関数とする最小化問題を解くことにより再構成画像を得ることができる.

$$\min_{V\geq 0}\left\{\sum_{i=1}^{I}\left\{\sum_{j=1}^{J}a_{ij}V_j - p_i\log\left(\sum_{j=1}^{J}a_{ij}V_j\right)\right\} + \lambda\sum_{j}\left|(\nabla V)_j\right|\right\} \tag{1}$$

ここで, j は再構成画素の通し番号, J は画素の総数, i は検出器画素の通し番号, I は検出器の総数, a_{ij} は画素 j から放出されたガンマ線が検出器 i で検出される確率（システムマトリクス）, p_i は計測により得た投影データ, V は再構成画像を表す. また, 式(1)の第一項はポアソン分布に関する負号付き対数尤度, 第二項は TV 項であり, λ は 2 つの項を調整するパラメータとなる. 本実験ではλの値を経験的に設定した. 式(1)は, 凸制約最適化問題に帰着するため, 最適性条件より

$$V_j^{k+\frac{1}{2}} = \frac{V_j^k}{\sum_{i=1}^{I}a_{ij}}\sum_{i=1}^{I}a_{ij}\left(\frac{p_i}{\sum_{t=1}^{J}a_{it}V_t^k}\right) \tag{2}$$

$$V_j^{k+1} = V_j^{k+\frac{1}{2}} - \frac{\lambda}{\sum_{i=1}^{I}a_{ij}}V_j^k div\left(\frac{\nabla V}{|\nabla V|}\right)_j \tag{3}$$

という 2 つの式で表すことができる. 式(2)を EM step, 式(3)を TV step と呼ぶ. 式(2)は, ML-EM 法と同様の式である. 式(3)は, ROF モデル[9]の data fidelity 項を$\sum_{i=1}^{I}a_{ij}/V_j^k$で重み付けした形とみなすことができる. 式(2)と式(3)を交互に繰り返し計算し, 解を求める.

イタレーション回数について OS-EM では 2 回, EM-TV では 5 回の再構成画像で評価を行った.

2-3. 性能評価指標

本実験では，OS-EM および EM-TV での再構成画像より，CNR および LOD を算出し，それぞれ比較した．

コントラストノイズ比（CNR: Contrast Noise Ratio）は，設定した関心領域（ROI: Region of Interest）内が均一であれば，標準偏差は小さくなり，その ROI 内に別の輝度値を持つ領域が存在するとき，その画素値の差を反映するように ROI 内の標準偏差が変動する，という性質を利用した低コントラスト分解能の評価指標である．CNR は以下のように定義される．

$$CNR = \frac{ROI_M - ROI_B}{SD_B} \tag{4}$$

ここで，ROI_M はヨウ素領域に設定した ROI 内の画素値の平均，ROI_B はバックグラウンド領域に設定した ROI 内の画素値の平均，SD_B はバックグラウンド領域に設定した ROI 内の画素値の標準偏差を意味する．本実験では，異なる濃度のヨウ素領域それぞれに ROI を設定し，それぞれの CNR 値を算出，比較した．

検出限界（LOD: Limit of Detection）は分析化学などの分野において，物質がない状態のシステムノイズ（バックグラウンドノイズ）と検出対象のシグナルが区別できる最低量として使用されている指標である．本研究では，バックグラウンド領域と区別できる最小の造影剤濃度として，Currie によって提案された以下の式で定義する[12]．

$$LOD = X_b + 3.29\sigma_b \tag{5}$$

X_b，σ_bはそれぞれ，バックグラウンド領域の平均値，バックグラウンド領域の標準偏差を示す．LOD の算出について，バックグラウンド領域に複数の ROI を設定，それぞれで LOD を計算し，その平均を最終的な LOD 値とした．

3. 結果

図 4 に，OS-EM，EM-TV でそれぞれ再構成した画像を示す．

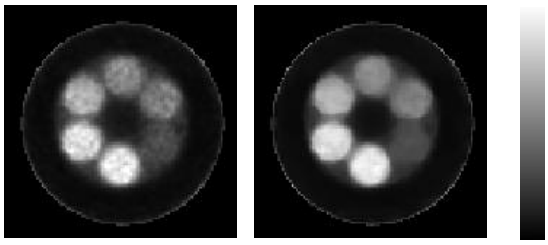


図 4 XFCT 再構成画像．左図は OS-EM による再構成画像．右図は EM-TV による再構成画像．

また，個々のヨウ素領域に設定した ROI の平均画素値を実際のヨウ素濃度に対してプロットしたそれぞれの結果を図 5，図 6 に示す．

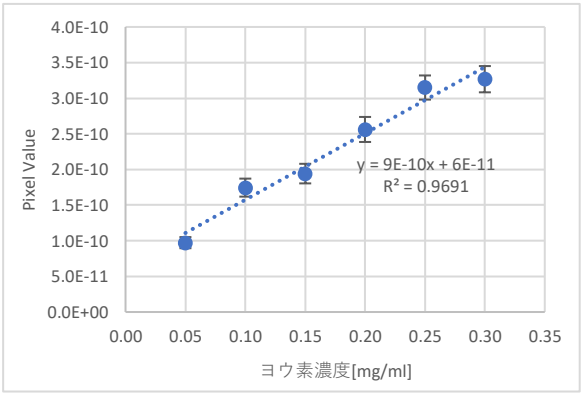


図 5 OS-EM による再構成画像の画素値と実際のヨウ素濃度の関係

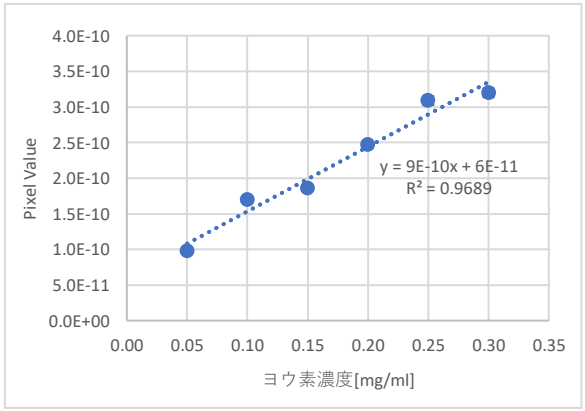


図 6 EM-TV による再構成画像の画素値と実際のヨウ素濃度の関係

さらに，表 1 にそれぞれのヨウ素濃度における CNR を示す．表 2 にそれぞれの再構成画像の LOD を示す．

表 1 各再構成画像のヨウ素濃度ごとの CNR.

ヨウ素濃度[mg/ml]	CNR	
	OSEM	EMTV
0.30	26.02	33.37
0.25	24.89	31.98
0.20	19.20	24.04
0.15	13.21	16.27
0.10	11.31	14.23
0.05	3.85	4.99

表 2 各再構成画像の LOD.

再構成法	LOD (iodine) [mg/ml]
OSEM	0.049
EMTV	0.036

4. 考察

図 4 より，OS-EM による再構成画像の 0.05mg/ml のヨウ素領域はバックグラウンドノイズに埋もれてしまい，その境界は視覚的に不明瞭であるが, EM-TV による再構成画像では境界が視覚的に明瞭ではっきりしている．表 1 を見ると，すべてのヨウ素濃度の ROI において，OS-EM よりも EM-TV による再構成画像の CNR が高くなっている．特に，最小濃度である 0.05mg/ml の ROI の CNR は，OS-EM に比べ, EM-TV が 1.3 倍向上している．以上より, EM-TV による再構成画像の画質向上が示された．また，表 2 より，OS-EM による再構成画像の LOD が 0.049mg/ml であるのに対し，EM-TV による再構成画像の LOD は 0.036mg/ml となり，検出限界に関しても改善されていることが確認できた．

図 5，図 6 より，OS-EM，EM-TV どちらの再構成画像においても，画素値と実際のヨウ素濃度に相関があり，定量性が保たれていることが確認できた．しかし，散乱線があるため，切片は 0 にはならない．また，OS-EM に比べ，EM-TV による再構成画像は，すべてのヨウ素濃度の ROI において標準偏差が小さく，ヨウ素領域のノイズが抑制されていることが分かる．

今後の課題として，散乱線補正によるさらなる検出限界の改善を挙げる．今回の実験では，散乱線補正を行っていないが，現在のマルチピンホール XFCT システムで使用する 2 次元検出器はエネルギー分解能が低く，かつマルチピンホールコリメータの配置の都合上，散乱線の影響が大きい．そのため，散乱線の補正が必要であるが，エネルギー分解能がない検出器で散乱線補正を行うには，散乱線の観測過程のモデル化や透過 CT 像が必要であった．そこで，Sasaya らは造影剤元素の K 吸収端の直上と直下のエネルギーで取得した 2 種類の投影データから散乱線補正を行う独自の ML-EM 再構成アルゴリズムを開発した [13]．それによって，シングルピンホールコリメータを用いた XFCT 撮像系において，散乱線補正および検出限界の改善に成功している．今後，その散乱線補正法を，マルチピンホールコリメータを用いた XFCT 撮像系に適応し，さらに EM-TV アルゴリズムと組み合わせることで，さらなる検出限界の改善を期待する．

5. まとめ

本研究では，マルチピンホール X 線蛍光 CT の画質改善および検出限界の向上を目的として, EM-TV アルゴリズムを組み込んだ画像再構成法を提案した．物理ファントム実験の結果，再構成画像のノイズが抑制され、ヨウ素領域が明瞭に描出できた．また，従来の OS-EM 法に対して、コントラストは 1.3 倍改善し、検出限界も 0.049 mg/ml から 0.036 mg/ml に改善した．

謝辞

本研究はJSPS科研費 JP17H04116の助成を受けたものです.

利益相反の有無

なし

文　献

［1］ Lwin TT, Takeda T, Wu J et al.: Preliminary quantitative analysis of myocardial fatty acid metabolism from fluorescent X-ray computed tomography imaging. J Synchrotron Radiat 14: 158-162, 2007

［2］ Lwin TT, Takeda T, Wu J et al.: Visualization of age-dependent myocardial metabolic impairment in cardiomyopathic model hamster obtained by fluorescent X-ray computed tomography using I-127BMIPP. J Synchrotron Radiat 15: 528-531, 2008

［3］ Takeda T, Wu J, Lwin TT et al.: X-ray fluorescent CT imaging of cerebral uptake of stable-iodine perfusion agent iodoamphetamine analog IMP in mice,. J Synchrotron Radiat 16: 57-62, 2009

［4］ Huo Q, Yuasa T, Akatsuka T et al.: Sheet-beam geometry for *in vivo* fluorescent x-ray computed tomography: proof-of-concept experiment in molecular imaging. Opt. Lett. 33: 2494-2496 2008

［5］ Sunaguchi N, Yuasa T, Hyodo K et al.: Fluorescent x-ray computed tomography using the pinhole effect for biomedical applications. Opt. Commun. 297: 210-214, 2013

［6］ Sasaya T, Sunaguchi N, Hyodo K et al.: Multi-pinhole fluorescent x-ray computed tomography for molecular imaging. Sci. Rep. 7: 5742, 2017

［7］ 篠原広行，小畠隆行，橋本雄幸 : 圧縮センシングによる少数投影からの画像再構成．断層映像研究会雑誌 **40**: 31-42, 2014

［8］ 工藤博幸 : 低被曝 CT における画像再構成法―統計的画像再構成, 逐次近似画像再構成，圧縮センシング―. Med Imag Tech **32**: 239-248, 2014

［9］ Rudin L, Osher S, Fatemi E: Nonlinear total variation based noise removal algorithms, Physica D 60: 259-268, 1992

［10］ Sawatzky A, Brune C, Wbbeling F, et al.: Accurate EM-TV algorithm in PET with low SNR. 2008 IEEE Nuclear Science Symposium Conference Record: 5133-5137, 2008

［11］ 砂口尚輝, 湯浅哲也, 兵藤一行 他: 散乱線低減のための Sn 箔フィルタを用いたピンホール蛍光 X 線 CT. 第 61 回応用物理学会春季学術講演会講演予稿集: 02 076, 2014

［12］ Currie LA, Limits for qualitative detection and quantitative determination. Anal. Chem. 40: 586-593, 1968

［13］ Sasaya T, Sunaguchi N, Lwin TT et al.: Dual-energy fluorescent x-ray computed tomography system with a pinhole design: Use of K-edge discontinuity for scatter correction. Sci. Rep. 7: 44143, 2017

EM-TV algorithm for multi-pinhole X-ray fluorescence CT image reconstruction

Michikazu KANAZAWA[*1], Ryota Hoshina[*2], Kosuke SASAKI[*1], Tomohiro ABE[*1], Tenta SASAYA[*2], Naoki SUNAGUCHI[*3], Hidekazu KAWASHIMA[*4], Kazuyuki HYODO[*5], Tetsuya YUASA[*2], Tsutomu ZENIYA[*1],

*1 Hirosaki University
*2 Yamagata University
*3 Nagoya University
*4 Kyoto Pharmaceutical University
*5 High Energy Accelerator Research Organization (KEK)

In multi-pinhole x-ray fluorescence computed tomography (mp-XFCT), nonradioactive nuclides introduced into a compound are excited by incident X-rays and the generated fluorescent X-rays are measured from multiple directions. By reconstructing the measured projection data, mp-XFCT provides the three-dimensional distribution of the compound inside the subject. However, in order to perform biological imaging with the current mp-XFCT system, the detection limit of the iodine concentration introduced into the compound to be measured is insufficient, and improvement is necessary. In this study, we propose a reconstruction method using the EM-TV algorithm that incorporates the total variation (TV) as a regularization term for improving the detection limit and the image quality. In the evaluation experiment, a physical phantom was prepared, the measured projection data were reconstructed by the conventional OS-EM method and EM-TV algorithm, and the reconstructed images were compared. In the EM-TV, the noise was suppressed, so the iodine region at the minimum concentration of 0.05 mg/ml was clearly depicted. The contrast-to-noise ratio (CNR) was improved about 1.3 times, and the detection limit of iodine was 0.049 mg/ml in OS-EM, whereas it was improved to 0.036 mg/ml in EM-TV.

Key words: X-ray fluorescence computed tomography, Multi-pinhole, Image reconstruction, Total variation

Swish を利用した Blind DnCNN とパラレルデノイジングへの応用

須貝 太郎 [*1)]，高野 航平 [*2)]，大内 翔平 [*2)]，伊藤 聡志 [*1,2)]

要旨

MR 画像の品質改善により読影を容易にすることを目的とし，Zhang らが提案した DnCNN に対し 3 点の改良を加えた．第 1 点は，活性化関数に Swish を応用，第 2 点は，雑音量が未知，かつ，空間可変な雑音画像に適応しうるブラインド雑音除去の導入，第 3 点は，隣接するスライス画像を線形加算して高 S/N 画像を作成し，ブラインド雑音除去後に連立方程式を解くパラレルデノイジングの導入である．高 S/N 画像をデノイジングするため画像劣化の抑制が期待でき，線形加算の逆変換により画像を分離するので，ぼけも少ない．実験の結果，Swish の採用とパラレルデノイジングの導入により画質劣化を抑制し，従来法を上回る雑音除去性能が得られることが示された．

キーワード: Swish, ブラインド雑音除去, 深層学習

1 はじめに

磁気共鳴映像法 (Magnetic Resonance Imaging: MRI) は，電離放射線を生じる X 線を使用しないため安全性が高く，多くの医療分野で使用されている．MRI を構成する主磁界コイルの磁界強度が大きくなったこと，および，信号受信と伝送技術の進歩により，高分解能かつ高 S/N 比の画像が得られるようになった．一方で，パラレルイメージングや機能的イメージングなどの新たな撮像法が使用されることがあるため，依然として低 S/N 比の画像が再生されることがある．そのため，MR 画像の雑音を除去する要望は高いものがある．MR 画像の雑音を除去する方法として今日までに多くの方法が提案されてきた．しかし，生体内の構造を保存する条件と，高い雑音除去性能の条件を両立することは難しく，検討の余地が残されている．

*1 宇都宮大学 大学院地域創生科学研究科
工農総合科学専攻
*2 宇都宮大学 大学院工学研究科
［〒321-8585 栃木県宇都宮市陽東 7-1-2］
e-mail: mc206833@cc.utsunomiya-u.ac.jp

自然画像用に提案された雑音除去法は，しばしば医用画像の雑音除去問題に応用されてきた．近年，雑音除去のあらたな潮流として，微分方程式を解く形式の方法が提案され，それらの代表的な方法に全変動フィルタ [1]，一般化全変動フィルタ [2]，非等方拡散フィルタ [3]，非局所平均フィルタ (Non-local mean: NLM)[4]，ブロックマッチ 3D フィルタ [5]，そして核ノルム最小フィルタ (weighted nuclear norm minimization : WNNM)[6] などがあり，画像の構造を保持しつつ優れた雑音除去性能を持つことで知られている．

近年，雑音除去問題への深層学習の導入が注目されており，Zhang らは，He らが提案する残差学習法 [7] を用いた DnCNN(Denoising Convolutional Neural Network)[8] により，自然画像の雑音除去で良好な結果が得られたと報告している．伊藤は，DnCNN を MR 画像に応用し，BM3D や WNNM よりも優れた雑音除去性能を有すること，および T1,T2 などの強調画像は，学習時も同種の画像を使用した方が良いことを示した [9]．

本研究では，DnCNN を 3 通りの方法により改良し，性能改善を行うことを目的とする．1 つ目は，活性化関数に Swish[10] を使用することである．これまで多くの CNN の活性化関数には，ReLU が使用されてきた．Zhang らが提案する DnCNN も ReLU を使用している．これに対し，Ramachandran らは，新しい活性化関数として Swish を提案し，画像分類や機械翻訳など多くの分野に応用した結果，平均的に他法を上回る性能が得られることを報告している．そこで本研究では，Swish を DnCNN に導入し，その有効性を検証した．2 つ目は，ブラインド雑音除去の応用である．DnCNN は，ブラインド雑音に対する有効性を示しているが，単一画像内で雑音量が変化する場合の検討はなされていない．本研究では，SENSE 法 [11] などにみられるような画像内で S/N が変化する場合を想定した検討を行った．3 つ目は，パラレルデノイジングの導入である．一般に系列の異なる雑音を加算平均すると信号対雑音比 (S/N 比) は向上する．雑音除去処理において雑音画像の S/N 比が高いほど，雑音処理による画像劣化やぼけは低減される傾向がある．そこで，本研究では異なる重み系列によって加算平均した画像を雑音除去したあとに，連立方程式を解くことで元の画像に復元する手法 (Parallelized Blind Image Denoising : ParBID) を考えた．

2　DnCNN の改良と ParBID

本研究で使用される基本的なニューラルネットワークは，Zhang らが提案する DnCNN である．DnCNN の構成を図 1 に示す．このネットワークは，入力層，中間層 (隠れ層)，出力層を合わせた 17 の層によって構成されている．そのうちの入力層では畳み込み (Conv) と活性化関数 (AF) が適用される．中間層では畳み込みと活性化関数の間にバッチ正規化 (BN) が追加される．出力層では畳み込みのみが適用される．畳み込みと活性化関数は入力データの特徴を抽出する役割を，バッチ正規化は学習効率を高める役割をもつ．DnCNN は残差学習法を用いている．残差学習法は，雑音重畳画像から雑音成分のみを推定することで，雑音除去画像を直接推定する従来法よりも高精度な雑音除去を実現することができる．なお，損失関数には，推定した雑音成分と真の雑音成分の平均二乗誤差を用いる．

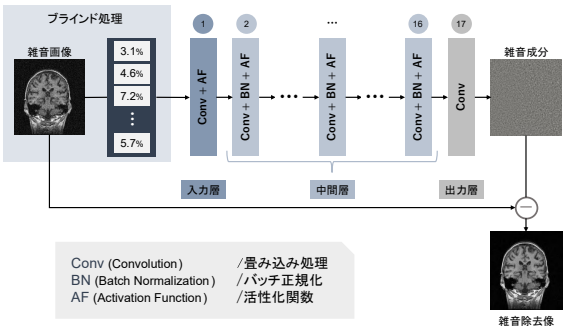


図 1　DnCNN の構成

本研究では DnCNN に対して，活性化関数の変更，ブラインド雑音除去の応用，パラレルデノイジングの 3 点による改良を図った．

2.1　Swish の導入

ReLU 関数は，多くのニューラルネットワークにおいて，活性化関数として広く使用されている．Zhang らが提案する DnCNN においても ReLU が使用されている．本研究では，雑音除去性能の向上を図るため，Ramachandran らが提案する Swish 関数を導入する．Swish は，式 (1) に示すシグモイド関数と入力 x の乗算として定義される (式 (2))．ReLU と類似した形状をしているため，関数を置き換えることが容易であり，入力が負の領域でも微分値をもち，また，負の入力に対するバンプ状の出力によって，雑音除去性能の向上が期待できる．また，β の値を正の定数に設定することで，バンプ状の入出力特性を変化させることができる．Swish の入出力特性を図 2 に示す．

$$\sigma(x) = \frac{1}{1+\exp(-\beta x)} \tag{1}$$

$$h(x) = x\,\sigma(x) \tag{2}$$

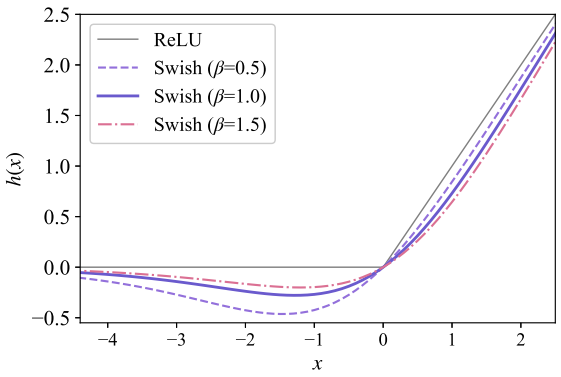


図 2　Swish の入出力特性

2.2　ブラインド雑音除去

一般的に MR 画像に含まれる雑音量は未知 (ブラインド) である．これに対応するために DnCNN では，雑音量が不明な画像に対しても雑音除去が行えるように拡張できる (本方法を B-DnCNN と称する)[8]．一方で，MR 画像は SENSE のように画像空間内で雑音量が変化する場合があるが，これまで，単一画像内で雑音量が変化する場合の検討は行われていない．そこで，雑音量が変化する画像に対する B-DnCNN の雑音除去特性の検討を行う．具体的には，異なるレベルの雑音が含まれた画像群を小さなパッチに分割し，雑音量がランダムなパッチ群からバッチを構成し，このバッチごとに学習を行った．

2.3　パラレルデノイジング (ParBID)

雑音がゼロを平均としてランダムに変動し，その統計が画像上で独立している場合，画像信号を加算すると信号は枚数に比例して増加するのに対して，雑音は加算の枚数の平方根に比例して増加する．したがって，加算された画像の S/N 比は理論的には加算枚数の平方根に比例して改善する．本研究では，マルチスライスによって得られた隣接する複数のスライス画像には分布に相似性がある点に注目し，複数のスライス画像を重み付け加算により合成したのち，S/N 比が高まった合成画像の雑音除去を行う．画像合成により画像にぼけが生じるが，異なる重みにより加算された複数の合成画像を使用して，ぼけを取り除くことが可能である．本研究では，S/N 比が高まった加算画像を雑音除去し，雑音除去画像を逆問題を解く手法によりスライス画像に分離する並列型雑音除去法 (Parallelized Blind Image Denoising : ParBID) を提案する．ParBID は以下の 3 ステップによる雑音除去を実行する．その流れを図 3 に示す．

Step-1. 画像の重み付き加算

隣接する複数のスライス画像系列 $\mathbf{R}(r_1, r_2, ...r_n)$ に重み行列 $\mathbf{A}(a_1, a_2, ...a_n)$ を与えて線形結合することで，S/N が改善した像 $\mathbf{I}(i_1, i_2, ...i_n)$ を得る．

$$\mathbf{I}^{\mathrm{T}} = \mathbf{A}\mathbf{R}^{\mathrm{T}} \tag{3}$$

Step-2. DnCNN によるブラインド雑音除去

$\mathbf{I}$ の各要素となる線形加算画像 $I_k(k = 1, 2, ...n)$ に雑音除去を適用することで，雑音除去像 $\mathbf{D}(d_1, d_2, ...d_n)$ を得る．

$$d_s = \text{B-DnCNN}(i_s) \tag{4}$$

Step-3. 連立方程式を解くことによる画像分解

線形結合された雑音除去像 $\mathbf{D}$ を式 (3) の逆変換式である式 (5) により分離する．この処理によって，線形加算によりぼけた画像からスライスごとの画像に分離され，鮮鋭な画像が復元される．

$$\mathbf{P}^{\mathrm{T}} = (\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{D}^{\mathrm{T}} \tag{5}$$

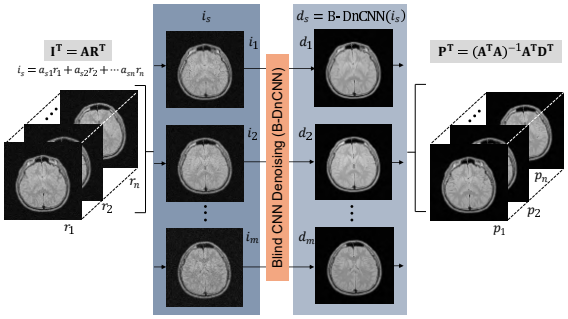


図 3　ParBID の流れ

3　DnCNN の学習

IXI データセット [12] に含まれるプロトン密度強調像 (PDWI)500 枚，および T2 強調像 (T2WI)500 枚を合わせた計 1000 枚の MR 画像のフーリエ変換信号に対して，標準偏差が画像の最大輝度値の 2.5%~7.5% になるような雑音を与え，その再構成された雑音画像と雑音無し画像との対を学習用データセットとして，30 エポックの学習を行った．このときの実行環境を表 1 に示す．また，B-DnCNN における学習設定を表 2 に示す．学習のオプティマイザには，Adam[13] の改善手法として開発された Eve[14] を使用した．学習には GPU(NVIDIA GeForce GTX1080Ti) を使用して，約 20 時間を要した．実行環境として MATLAB の深層学習ライブラリである MatConvNet[15] を用いた．

表 1　実行環境

CPU	Intel Core i7-7700 3.60GHz
GPU	NVIDIA GeForce GTX1080Ti
メモリ	32GB
コンパイラ	Visual Studio 2017 MATLAB R2017b CUDA 9.0 cuDnn 7.0.5

表 2　学習設定

学習する雑音量	Blind(2.5〜7.5%)
オプティマイザ	Eve
パッチサイズ	35×35
ストライド	10
バッチサイズ	128
エポック数	30
学習率	1〜20 エポック / 1.0×10^{-3} 21〜30 エポック / 1.0×10^{-4}
データ拡張	スケール変更，回転，反転
学習時間	約 20 時間 (GPU 使用)

4　雑音除去実験

本研究の雑音除去実験に使用する画像はすべて，キヤノンメディカル製 Vantage Titan によって被験者の同意を得て撮像された MR 画像である．基本的には，高 S/N 比を有する MR 画像として，3.0T コイルで撮像された画像を使用する．臨床応用を想定した検討の場合にのみ，1.5T コイルで撮像された画像を使用する．撮像条件を表 3 に示す．雑音除去実験における実行環境は，表 1 に示した学習時の環境と同じである．

表 3　撮像条件

キヤノンメディカル製 Vantage Titan		
	3.0T	1.5T
TR [ms]	3,500	50
TE [ms]	352	40
フリップ角 [°]	90	20
スライス厚 [mm]	1.2	1.5
スライス間隔 [mm]	1.2	1.5
マトリクスサイズ	256×256	256×256

4.1　活性化関数の検討

活性化関数による MR 画像の雑音除去性能を比較するため，表 2 の学習設定は変更せずに，活性化関数のみを ReLU，Swish($\beta = 1.0$)，Swish($\beta = 1.5$) と変更して学習した B-DnCNN を用意した．テスト画像 50 枚に対して，画像空間における雑音の標準偏差が最大輝度値の 2.5%，5.0%，7.5% となるように MR 信号に雑音を重畳し，再構成画像に対して各活性化関数を用いて学習した B-DnCNN による雑音除去実験を行った．50 枚の PSNR(Peak Signal-to-Noise Ratio) の平均値を表 4 に示す．表 4 より，どの標準偏差の雑音に対しても，Swish は ReLU よりも高い PSNR を示した．また，Swish の係数 β について，$\beta = 1.0$ よりも $\beta = 1.5$ のほうが高い PSNR を示した．以後の雑音除去実験では，B-DnCNN の活性化関数に Swish($\beta = 1.5$) を用いる．

表 4　各活性化関数の PSNR

	雑音標準偏差 σ		
	2.5%	5.0%	7.5%
ReLU	36.79	32.61	30.32
Swish(β=1.0)	36.83	32.65	30.36
Swish(β=1.5)	**36.84**	**32.67**	**30.37**

4.2　画像空間内不均一雑音の除去

ブラインド雑音除去の応用として，画像空間内で雑音量が変化する MR 画像に対する雑音除去性能を評価した．本検討の場合は，高 S/N 比を有する MR 画像 100 枚に対して，図 4 に示すように画像空間内での不均一分布を想定したガウス型白色雑音を重畳し，それらの画像に対して学習済みの B-DnCNN を用いた雑音除去を行った．

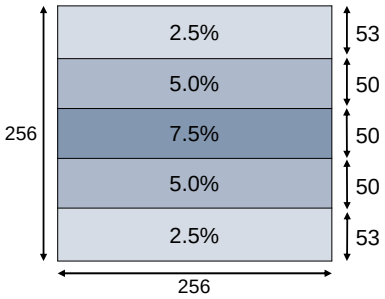


図 4　空間内で不均一に分布した雑音

雑音除去結果の例として，図 5 に雑音重畳前の高 S/N 比画像，図 4 の不均一雑音を重畳した画像，B-DnCNN による雑音除去像，除去された雑音を示す．また，比較のために，5.0% の雑音だけを学習した DnCNN，および 5.0% を想定して実行した BM3D の雑音除去像も示す． なお，細部の構造を確認しやすくするために，各画像に同条件のコントラスト調整を行っている．DnCNN(5.0%) および BM3D は，5.0% 部分の雑音は除去できているが，2.5% 部分の雑音に対しては画像の平滑化がみられ，7.5% 部分の雑音は残留している．一方で，B-DnCNN はどの雑音量に対しても高い雑音除去性能を示した．また，

(iv) 差画像から除去された雑音量は，図 4 に応じており，空間的な雑音量の変化に応じた雑音除去が行えていることがわかる．B-DnCNN は，MR 画像空間内で不均一に分布した雑音に対しても，高い雑音除去性能を有することが示された．

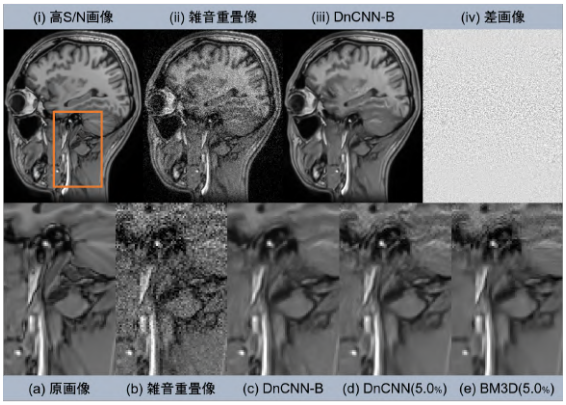


図 5　不均一に分布した雑音の除去像比較

4.3　隣接したスライス画像の雑音除去

キヤノンメディカル製 Vantage Titan 3.0T で撮像された MR 画像のうち，隣接したスライス画像 50 種類に対して，画像空間の雑音の標準偏差が最大輝度値の 2.5%，5.0% となるよう MR 信号に雑音を重畳し，再構成画像に対して学習済み B-DnCNN による雑音除去実験を行った．このとき，2 枚ごとの並列型雑音除去 (Par-2)，3 枚ごとの並列型雑音除去 (Par-3)，重み付き加算を行わないブラインド雑音除去 (Single) それぞれの雑音除去性能を比較した．比較には，実際の雑音除去像，PSNR および SSIM(Structual SIMlarity index)[16] を用いた．なお，Par-3 において，3 枚の連続したスライス画像のうち両端の画像を (end)，中央の画像を (mid) と称する．Par-2 の重み係数は $0.4, 0.6$，Par-3 は $0.2, 0.3, 0.5$ を使用した．

$$\begin{cases} \mathrm{I}_1 = 0.4\,\mathrm{r}_1 + 0.6\,\mathrm{r}_2 & (6) \\ \mathrm{I}_2 = 0.6\,\mathrm{r}_1 + 0.4\,\mathrm{r}_2 & (7) \end{cases}$$

$$\begin{cases} \mathrm{I}_1 = 0.2\,\mathrm{r}_1 + 0.3\,\mathrm{r}_2 + 0.5\,\mathrm{r}_3 & (8) \\ \mathrm{I}_2 = 0.3\,\mathrm{r}_1 + 0.5\,\mathrm{r}_2 + 0.2\,\mathrm{r}_3 & (9) \\ \mathrm{I}_3 = 0.5\,\mathrm{r}_1 + 0.2\,\mathrm{r}_2 + 0.3\,\mathrm{r}_3 & (10) \end{cases}$$

表 5 に示した PSNR と SSIM の評価結果は，いずれの雑音量においても，Par-2 および Par-3 は，Single よりも高い PSNR，SSIM を示した．パラレルデノイジングの中でも，Par-3(mid) が最も高い評価値を示した．また，雑音量が少ない画像のほうが改善率が大きかった．雑音除去像の例として，雑音量 2.5% のときの雑音付加前の高 S/N 画像と雑音重畳画像，雑音除去像を図 6 に示す．なお，細部の構造を確認しやすくするために，各画像に同条件のコントラスト調整を行っている．Par-2 および Par-3 は，Single よりも平滑化が小さく，判然としたコントラストの画像が得られた．以上のことから，隣接したスライス画像の雑音除去において，ParBID が高い雑音除去性能を有することが示された．

表 5　各雑音除去法における PSNR [dB] と SSIM

	2.5% noise		5.0% noise	
	PSNR	SSIM	PSNR	SSIM
Single	35.45	0.972	31.57	**0.937**
Par-2	35.70	0.972	31.60	0.935
Par-3 (end)	35.72	0.972	31.59	0.934
Par-3 (mid)	**36.05**	**0.974**	**31.72**	**0.937**

4.4　臨床画像への応用

ParBID の実応用を検証するため，臨床で得られた雑音が含まれる MR 画像に対して ParBID を適用し，雑音除去性能を評価した．キヤノンメディカル製 Vantage Titan 1.5T で撮像された雑音を含んだ MR 画像に対して，2 枚ごとの並列型雑音除去 (Par-2)，3 枚ごとの並列型雑音除去 (Par-3)，重み付き加算を行わないブラインド雑音除去 (Single) それぞれの雑音除去性能を比較する．図 7 に雑音を含んだ原画像，Single，Par-2，Par-3(mid) による雑音除去像を示す．Par-3 では，Single に比べて鮮鋭度が高く，かつ，コントラストの高い画像が得られた．このことから，ParBID の臨床で得られた MR 画像に対する有効性が示唆された．

5　考察

線形結合された画像の S/N 比は，理論的には加算枚数の平方根に比例して改善するため，隣接したスライス画像を 2 枚用いた Par-2 よりも，3 枚用いた Par-3 のほうが高い雑音除去性能を示したと考えられる．また，ParBID は，各スライス間の雑音がランダムで相関関係のないときに最も優れた性能を

発揮する．MR 画像に含まれる雑音は，撮像方法やその他の要因によって，スライス画像間の雑音に相関関係が生じる場合がある．この場合は，線形加算画像の S/N 比がランダム雑音の場合に比べて低下するため，提案法の有効性はやや低下するものと考えられる．3 枚の中で両端に位置するスライス画像 (end) よりも，中央のスライス (mid) が高い評価値を示したのは，画像の類似性によるものである．中央のスライス画像は両方向に 1 つのスライス間隔しかないのに対して，両端のスライス画像間には 2 つ分のスライス間隔があるため，画像の類似性が低下し，雑音除去性能が低下したものと考えられる．それでも，Par-3(end) は Par-2 と同等の PSNR および SSIM を示していたため，ParBID に最適なスライスは 3 枚であると推測できる．ただし，最適なスライス枚数はスライス間隔によって異なると考えられる．スライス間隔が小さい場合には，隣接したスライス画像の類似性が高まるため，3 枚より多い枚数での ParBID のほうが，雑音除去性能が向上する可能性がある．また，ParBID の原理は，B-DnCNN だけでなく，BM3D や WNNM など，ほかの雑音除去法にも有効である．ただし，線形結合後における雑音量の推定は困難であるため，本方法のようにブラインド雑音処理の適用が適していると考える．

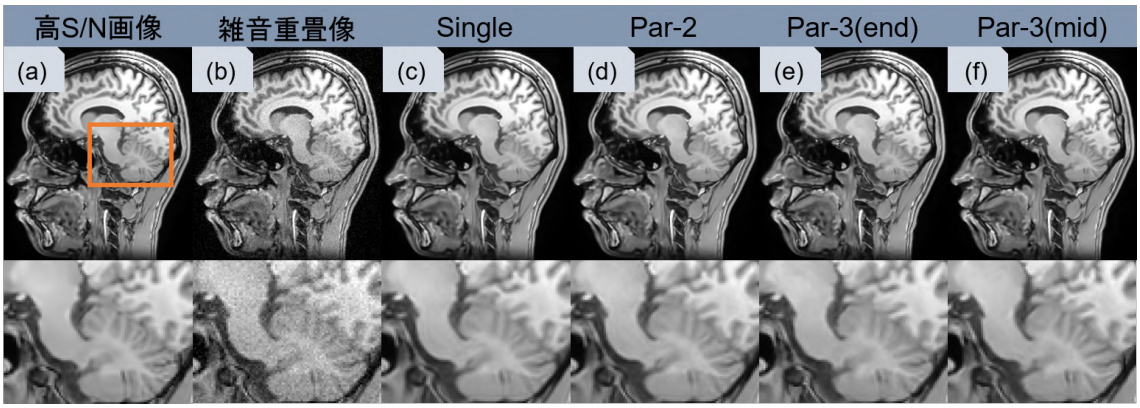


図 6　隣接したスライス画像の雑音除去像比較

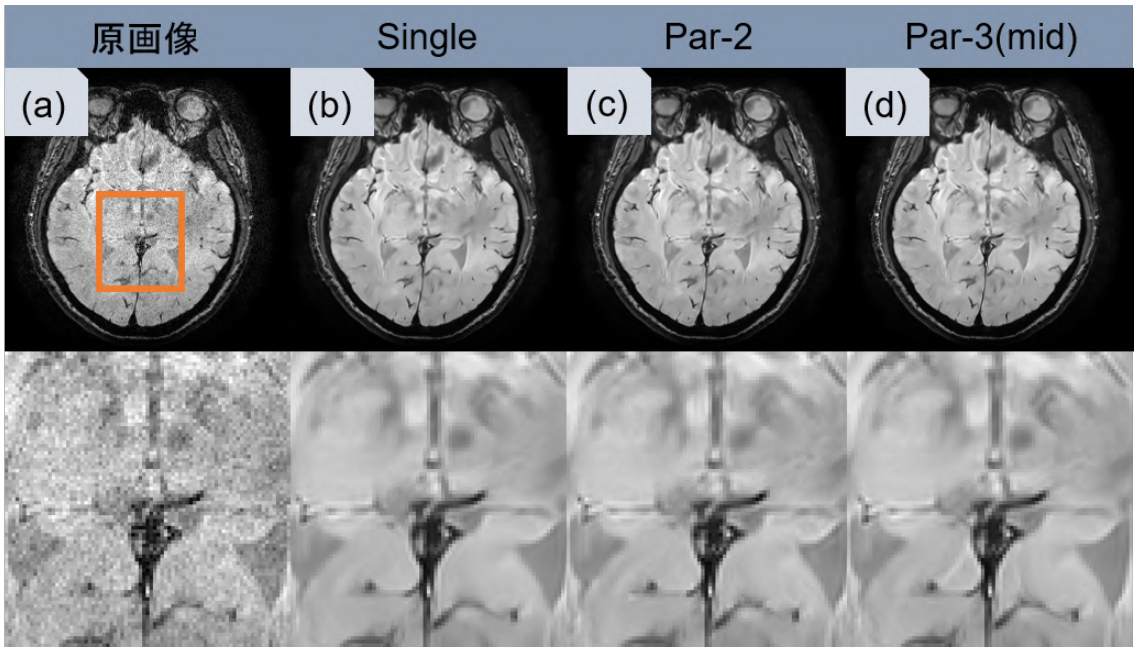


図 7　臨床画像の雑音除去像比較

6　まとめ

本研究では，3 通りの方法により DnCNN の性能改善を図った．第 1 に活性化関数には ReLU に替わり Swish を導入した点，第 2 にブラインド雑音除去の導入，第 3 にパラレルデノイジングの導入である．これらによって，雑音除去像の PSNR および SSIM を改善することができた．また，画像はコントラストの構造保存性に優れる可能性が示された．

謝辞

本研究の一部は，科学研究費助成金 (19K04423) および柏森情報科学振興財団の助成により実施された．また，MR 画像の提供を受けたキヤノンメディカルシステムズ，および本研究を遂行する際に利用した IXI データセットを提供する Information eXtraction from Images プロジェクトに感謝の意を表します．

利益相反の有無

なし．

文　献

[1] Rudin LI, Osher S, Fatemi E. Nonlinear total variation based noise removal algorithms. Physica D 1992; 60: 259–268.

[2] Bredies K, Kunisch K, Pock T. Total generalized variation. SIAM J Imaging Sci 2010; 3: 492–526.

[3] Gerig G, Kubler O, Kikinis R et al. Nonlinear Anisotropic Filtering of MRI data. IEEE Trans. Medical Imaging 1992; 11: 221–232.

[4] Buades A, Coll B, Morel JM. A Non-Local Algorithm for Image Denoising, Proc. IEEE Conf. on Computer Vision and Pattern Recognition, San Diego, USA, 2005; 2: 60–65.

[5] Dabov K, Foi A, Katkovnik V et al. Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering. IEEE Trans. Image Processing 2007; 16: 2080–2095.

[6] Gu S, Xie Q, Meng D et al. Weighted Nuclear Norm Minimization and Its Applications to Low Level Vision. International Journal of Computer Vision 2017; 121: 183–208.

[7] He K, Zhang X, Ren S et al. Deep residual learning for image recognition. IEEE Conf. on Computer Vision and Pattern Recognition, Las Vegas, USA 2016; 770–778.

[8] Zhang K, Zuo W, Chen Y et al. Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising. IEEE Trans. Image Processing 2017; 26: 3142–3155.

[9] 伊藤 聡志, 深層学習を用いた MR 画像のノイズ除去に関する検討, 第 45 回日本磁気共鳴医学会大会講演抄録集, O2-107,2017, p.272.

[10] Ramachandran P, Zoph B, Le QV. Searching for activation functions. CoRR 2017, arXiv:1710.05941.

[11] Pruessmann KP, Weiger M, Scheidegger MB et al. SENSE: Sensitivity encoding for fast MRI, Magn Reson Med 1999; 42: 952–962.

[12] Imperial College London, https://brain-development.org/ixi-dataset/ (Accessed April 20 2020).

[13] Kingma DP, Ba J. Adam: A Method for Stochastic Optimization. International Conf. on Learning Representations, San Diego, USA, 2015, arXiv:1412.6980.

[14] Hayashi H, Koushik J, Neubig G, Eve:A Gradient Based Optimization Method with Locally and Globally Adaptive Learning Rates, arXiv:1611.01505ver3, 2018.

[15] The MatConvNet Team. MatConvNet: CNNs for MATLAB. http://www.vlfeat.org/matconvnet/ (Accessed April 20 2020).

[16] Wang Z, Bovik AC, Sheikh HR et al. Image Quality Assessment: From Error Visibility to Structural Similarity. IEEE Trans. Image Processing 2004; 13: 600–612.

Blind DnCNN using Swish and Application to Parallel Image Denoising

Taro SUGAI[*1)], Kohei TAKANO[*2)], Shohei OUCHI[*2)], Satoshi ITO[*1,2)]

*1 Graduate School of Regional Development and Creativity, Utsunomiya University
*2 Graduate School of Engineering, Utsunomiya University

To improve the performance of the denoising convolutional neural network (DnCNN), a refinement involving an activation function and parallelized blind image denoising (ParBID) were examined. Swish can improve the denoising performance of the DnCNN, and the denoising performance was further improved by the ParBID. Blind denoising using Deep Convolutional Neural Network(CNN) was examined. We created single CNN trained with the range of the noise levels as 2.5∼7.5% for blind denoising. It was shown that high denoising performances was obtained using single CNN for blind noise MR images.

Keyword: Swish, Blind Denoising, Deep Learning

アーチファクト学習を導入した
改良型 ISTA-Net による深層学習再構成

宮本 裕大[*1]，伊藤 聡志[*1]

要旨

MRI の撮像時間を短縮する圧縮センシングの再構成法として，深層学習を利用する方法が大きな注目を集めている．Zhang らは，L1 ノルム最適化の手法である反復的ソフトしきい値法をネットワークで模擬した ISTA-Net を提案し，画素の欠落がある画像の修復に応用した．この方法を MRI の圧縮センシング再構成に応用すると，アーチファクトが残留する場合があった．本研究では，アーチファクトの除去性能を改善するために，ネットワークにおいて，畳み込み処理と活性化関数からなる層をカスケード接続することにより，アーチファクト推定を可能とする新たなネットワークを構成した．カルテシアン座標系による 2 次元撮像シミュレーションの結果，アーチファクト除去効果が向上し，再構成像の PSNR と SSIM をともに改善することができた．

キーワード： ISTA，圧縮センシング，再構成，深層学習

1. はじめに

医療現場で MRI(Magnetic Resonance Imaging)による画像を診断に使用する際，高品質な画像が求められる．しかしながら，高品質な画像を得るためには，信号対雑音比を高める必要があり，一般に撮像に多くの時間を要する．

MRI の撮像時間を短縮化するための手法として，圧縮センシング(Compressed Sensing：CS)[1]の利用が進んでいる．CS では，サンプリング定理を満足しない少数の信号から画像を再構成することができるため，MRI へ応用することができれば，信号収集時間の短縮につながる(CS-MRI)[2]．CS は，元画像を推定する方法として，画像にスパース性があることを仮定し，L2 ノルム拘束条件を付与した L1 ノルム最小化問題を解く．このとき，反復的ソフトしきい値法(Iterative Shrinkage soft Thresholding Algorithm：ISTA)[3]や交互方向乗数法(Alternating Direction Method of Multipliers：ADMM)[4]などの反復を伴うアルゴリズムを使用する．これらの反復的解法は，数式的な反復処理を基本として再構成を行うため，多くの計算コストを要すること，スパース化関数に依存した画質の変化および，スパース近似により画質が低下することなどの課題がある．

近年，これらの課題を解決する方法として，深層学習(Deep Learning：DL)を利用する画像再構成が注目されている[5-7]．DL では学習の時間を要することが課題であるものの，再構成には学習したパラメータを用いるため，短時間で行うことが可能である．また，煩雑な再構成パラメータの選択，スパース化関数の選択を必要とせず，反復再構成法を上回る良質な再構成像が得られ

*1 宇都宮大学大学院 地域創生科学研究科
工農総合科学専攻
情報電気電子システム工学プログラム
[〒321-8585 栃木県宇都宮市陽東 7-1-2]
e-mail: mc206864@cc.utsunomiya-u.ac.jp

ることから，大きな期待が寄せられている[8]. DL を使用した画像再構成法の一つに, ISTA を畳み込みニューラルネットワークで模した ISTA-Net[9][10]がある. ISTA-Net は, CS において重要となるスパース化関数，しきい値など，従来型の反復再構成では経験的に設定するパラメータを学習によって求めることができる. ISTA-Net は，画素値が欠落した画像を入力とし，全画素値が与えられる画像を出力する画像－画像間学習ネットワークである．本研究ではこれに改良を加えて，信号－画像間学習とし, CS-MRI の再構成法に応用した．具体的には，入力は画像ではなく，間引きされた MR 信号とし，出力を再構成画像とする．また，損失関数において，目標画像と出力画像の L2 ノルムに加え，画像空間のアーチファクトに対応するために，新たにアーチファクトを学習する項を導入した(本法を ISTA-Net-MR とする).

本研究では, ISTA-Net-MR を，臨床で一般的に使用されるカルテシアン座標系による間引き信号に応用した. 2 次元撮像における再構成シミュレーションを実施し，アーチファクト学習を導入した DL による画質改善効果について検証を行った.

2. ISTA-Net-MR の構成

2.1. ISTA-Net

x を求める画像, Φ は観測行列, y を観測信号とするとき，以下の式(1)の関係がある.

$$y = \Phi x \qquad (1)$$

圧縮センシングでは，サンプリング定理を満足しない観測信号 y から x を求めようとするときに，式(2)に示される L2 ノルム拘束条件のもとで L1 ノルム最小化問題を解く.

$$\underset{x}{\operatorname{argmin}} \left\{ \frac{1}{2} \|\Phi x - y\|_2^2 + \lambda \|\psi x\|_1 \right\} \qquad (2)$$

x が推定した再構成像, ψ をスパース化関数, λ は正則化パラメータとする.

ISTA[3]は，式(2)に対し，メジャライザー最小化およびソフトしきい値処理を反復的に行うことで最小化するアルゴリズムである. ISTA-Net では，この反復を式(3)(4)のような DL ネットワークに書き換え，画像を再構成する. $Soft$ はソフトしきい値処理，$\theta^{(k)}$ はしきい値, $\psi^{(k)}$ はスパース化関数, $\tilde{\psi}^{(k)}$ は逆スパース化関数, $\varepsilon^{(k)}$ はステップサイズとする.

$$\begin{cases} r^{(k)} = x^{(k-1)} - \varepsilon^{(k)} \Phi^{\mathrm{T}} \left(\Phi x^{(k-1)} - y \right) & (3) \\ x^{(k)} = x^{(k-1)} - \tilde{\psi}^{(k)} \left(Soft\left(\psi^{(k)} \left(r^{(k)} \right), \theta^{(k)} \right) \right) & (4) \end{cases}$$

圧縮センシングでは，観測画像 x にスパース性を仮定している．反復再構成では，経験的にスパース化関数を与えるが, ISTA-Net ではスパース化関数 $\psi^{(k)}$ および逆スパース化関数 $\tilde{\psi}^{(k)}$ は，損失関数が最小になるように学習される．また，しきい値 $\theta^{(k)}$ およびステップサイズ $\varepsilon^{(k)}$ も学習される.

2.2. ISTA-Net-MR

ISTA-Net を MRI へと応用する方法では，入力が画像のフーリエ変換信号であることから，フーリエ変換を組み込む必要がある．そこで，式(5)が解くべき問題となる.

$$y = \mathrm{UF}x \qquad (5)$$

ここで, U は信号収集点分布として MR 信号の間引きに使用する．F はフーリエ変換である.

DL によって再構成を行う場合のアルゴリズムは，式(6)(7)となる．F^{T} は逆フーリエ変換に相当する.

$$\begin{cases} r^{(k)} = x^{(k-1)} - \varepsilon^{(k)}\mathrm{F}^{\mathrm{T}}\left(\mathrm{UF}x^{(k-1)} - y\right) & (6) \\ x^{(k)} = x^{(k-1)} - \tilde{\psi}^{(k)}\left(Soft\left(\psi^{(k)}\left(r^{(k)}\right), \theta^{(k)}\right)\right) & (7) \end{cases}$$

入力は間引いた MR 信号であり，ゼロフィル再構成像を初期画像 $x^{(0)}$ とする．$r^{(k)}$ は，k 回目の再構成像である $x^{(k)}$ の生成に使用する．図 1 に ISTA-Net-MR のネットワーク構成を示す．$r^{(k)}$ および $x^{(k)}$ は，図 1 内の $r^{(k)}$ および $x^{(k)}$ と対応している．スパース化関数 $\psi^{(k)}$ と，その逆変換 $\tilde{\psi}^{(k)}$ は，畳み込み処理 Conv と活性化関数 ReLU を組み合わせた Conv+ReLU+Conv によって学習する．図 1 における，スパース化関数 $\psi^{(k)}$ の直前および逆スパース化関数 $\tilde{\psi}^{(k)}$ の直後の Conv では，特徴量の抽出を行う．式(7)のソフトしきい値処理 $Soft$ は，図 1 の $Soft$ に対応する．

2.3. アーチファクト学習の導入

ISTA-Net を MRI に応用する場合，信号の間引き率が大きく，かつ，ISTA における反復処理の回数に相当するフェイズが少ない場合にアーチファクトが現れることがある．提案法とする ISTA-Net-MR では，このアーチファクトを除去するため，ISTA 処理のあとに，畳み込み処理および活性化関数(Conv+ReLU)からなる層をカスケード接続し，最後に特徴を抽出する畳み込み処理 Conv を行うことで，アーチファクトの学習を行う(図 1 の朱色部分)．そして，ISTA 処理後の再構成像から推定されたアーチファクトを除去することにより，再構成像の品質改善を図る．

3. 画像再構成シミュレーション

3.1. 学習データセットとテスト画像

ISTA-Net-MR の学習では，間引き処理を行った MR 信号と，全信号から再構成された MR 画像の組を用いる．学習とテスト用の画像は，IXI dataset[11]に含まれる，ボランティアの頭部プロトン密度強調像(PDW 像) を使用し，91 枚を学習に，学習には使用していない 30 枚をテストに使用した．信号の収集法は，臨床で最も使用例のあるカルテシアン座標系とし，間引きの方法には，位相エンコード方向ランダム間引きを用いた．信号の収集比は，k 空間の全信号データの 30%であり，k 空間の中央領域 60 行は，信号エネルギーが大きいため，間引きを行わない．図 2 に，信号収集点分布を示す．

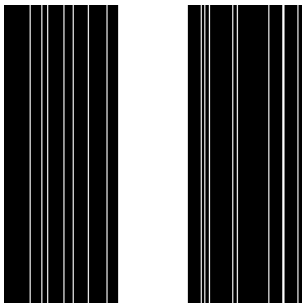


図 2 信号収集点分布
(30%：横方向が位相エンコード方向)

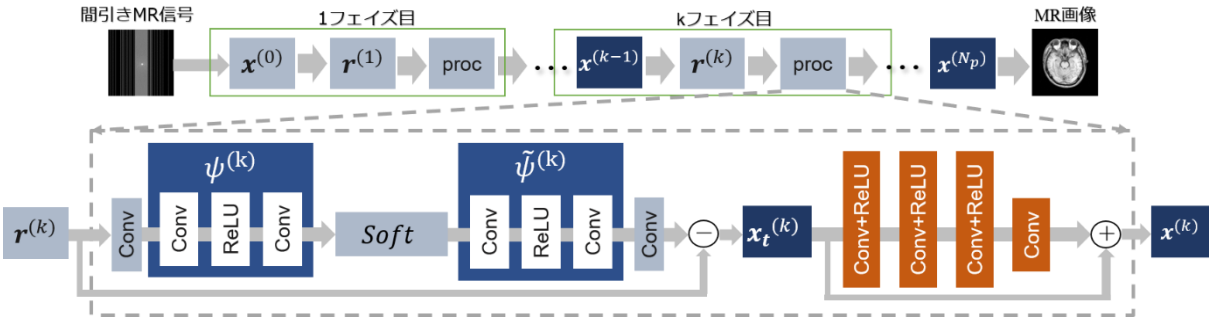


図 1 ISTA-Net-MR のネットワーク構成

3.2. 学習条件

本研究における再構成のシミュレーションでは, フェイズは5, エポック数は5,000とし, アーチファクトの推定に用いるConv+ReLUからなる層は, 3 層とした. 表 1 に本研究におけるシミュレーションの条件を示す. また, 表 2 に計算機環境を示す. 学習および再構成には CPU および GPU を使用した.

表 1 シミュレーションにおける条件

学習画像	91 枚
テスト用画像	30 枚
エポック	5,000
学習率	5.0×10^{-5}
アーチファクト推定層数	3
学習時間	約 7 時間

表 2 計算機環境

CPU	Intel Core i7-7700 (3.60GHz)
GPU	NVIDIA GeForce RTX2080 SUPER
RAM	32GB
Software	Python 3.7.3 Tensorflow-gpu 1.13.1 CUDA Toolkit 10.0 / cuDNN 7.6.0

3.3. 画像再構成

図 3 に, 再構成結果を示す. 表 3 に, 画質評価法として用いた PSNR(Peak Signal-to-Noise Ratio) および SSIM(Structural SIMilarity index)[12]の値を示す.

使用した再構成法は, 反復的解法である ISTA, 同条件で学習を行った従来型のアーチファクト学習を導入しない ISTA-Net, アーチファクト学習を導入した ISTA-Net-MR である. (a), (e), (i)は目標とするフルデータ再生像と拡大図であり, (b), (f), (j)は従前の反復的解法である ISTA による再構成像と拡大図, (c), (g), (k)は従来型の ISTA-Net による再構成像と拡大図, (d), (h), (l)は提案法である ISTA-Net-MR による再構成像と拡大図である.

図 3 より, 従前の反復的解法である ISTA と, ISTA-Net-MR の再構成像を比較すると, ISTA では強く平滑化され, 細部の構造が失われている部分が見られるが, ISTA-Net-MR では過度な平滑化は見られない. また, ISTA-Net では, 細部にアーチファクトが残存しており, フルデータ再生像にはない構造が描出されていることに対して ISTA-Net-MR では, ISTA-Net において描出されているアーチファクトが見られない. また, 細部の構造がより鮮鋭に再現されている. ISTA-Net-MR は, アーチファクトが少なく, 鮮鋭に再構成ができているが, 依然として平滑化が見られる.

表 3 より, PSNR と SSIM のいずれの評価法においても, ISTA-Net-MR では高い値が得られた.

表 3 PSNR と SSIM の評価結果

	ISTA	ISTA-Net	ISTA-Net-MR (提案法)
PSNR[dB]	23.69	26.07	**27.01**
SSIM	0.822	0.881	**0.895**

4. 考察

図 3 の再構成結果より, ISTA-Net では描出されているアーチファクトが, ISTA-Net-MR では低減されていることが確認できた. 畳み込み処理および活性化関数からなる層(Conv+ReLU)を, ネットワークに組み込むことにより, アーチファクトの推定が可能になったためであると考えられる. 本研究において, Conv+ReLU は 3 層としているが, 層の数を増やすことで, より柔軟な学習が可能になり, アーチファクトを推定する精度が向上すると考えられる.

ISTA-Net-MR においてアーチファクトの除去性能が向上したが, 細部の構造において依然として平滑化が見られた. これは信号収集点分布において, 高周波成分の信号収集量が少ないことに起因する. 信号における高周波成分は, 画像の分解能に影響し, 高周波成分の収集を少なくする場合, 画像にぼやけが生じる. この問題は, 信号における高周波成分の収集量を増やすことで改善が見込まれる.

また, 学習を必要とする ISTA-Net および ISTA-Net-MR いずれの場合においても, 学習枚数が 91 枚であり, 十分な学習が行えていない可能性がある.

5. まとめ

ISTA-Net は, 信号の間引き率が大きく, フェイズが少ない場合に, アーチファクトが発生しやすくなる課題がある. 本研究では, アーチファクト学習型 ISTA-Net を構成し, 再構成実験を行った. 結果, アーチファクト除去性能の改善に成功し, アーチファクト学習の導入が有効であることが示された. 今後は, ネットワークの改良や, 学習画像の増加および, 信号収集点分布の検討を行う予定である.

謝 辞

本研究の一部は,科学研究費補助金(19K04423)および柏森情報科学振興財団の助成により実施された. また, 本研究を遂行する際, IXI dataset を利用した. Information eXtraction from Images プロジェクトに感謝の意を表します.

利益相反の有無

なし

文 献

[1] Donoho DL : Compressed sensing. IEEE Transactions on Information Theory 52 : pp.1289-1306, 2006

[2] Lustig M, Donoho DL, Santos JM, et al.: Compressed Sensing MRI. IEEE Signal Processing Magazine, vol.25, no.2, pp.72-82, 2008

[3] Lustig M, Donoho DL, Pauly JM: Sparse MRI: The application of compressed sensing for rapid MR imaging. Magnetic Resonance in Medicine, 58 : pp.1182-1195, 2007

[4] Boyd S, Parikh N, Chu E, et al.: Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers, Foundation and Trends in Machine Learning, vol.3, no.1, pp.1-122, 2010

[5] Han Y, Ye JC : k-Space Deep Learning for Accelerated MRI, arXiv:1805.03779v1

[6] Lee D, Yoo J, Tak S ,et al.: Deep Residual Learning for Accelerated MRI Using Magnitude and Phase Networks. IEEE Transactions on Biomedical Engineering, 65 : pp.1985-1995, 2018

[7] Zhu B, Liu JZ, Cauley SF ,et al. : Image reconstruction by domain-transform manifold learning. Nature, 555 : pp.487-492, 2018

[8] Yang Y, Sun J, Li H, Xu Z : Deep Admm-Net for Compressive Sensing MRI, Advances in Neural Information Processing Systems, pp.10-18, 2016

[9] Zhang J ,et al.: ISTA-Net Interpretable OptimizationInspired Deep Network for Image Compressive Sensing : CVPR pp.1828-1837, 2018

[10] 宮本裕大, 伊藤聡志 : ISTA-Net を応用した MR 圧縮センシングの深層学習再構成. IEICE : p.77, 2019

[11] IXI dataset : https://brain-development.org/ixi-dataset/

[12] Wang Z, Bovik AC, Sheikh HR, Simoncelli EP : Image quality assessment from error visibility to structural similarity : IEEE Transactions on Image Processing, vol.13, no.4, pp. 600-612, 2004

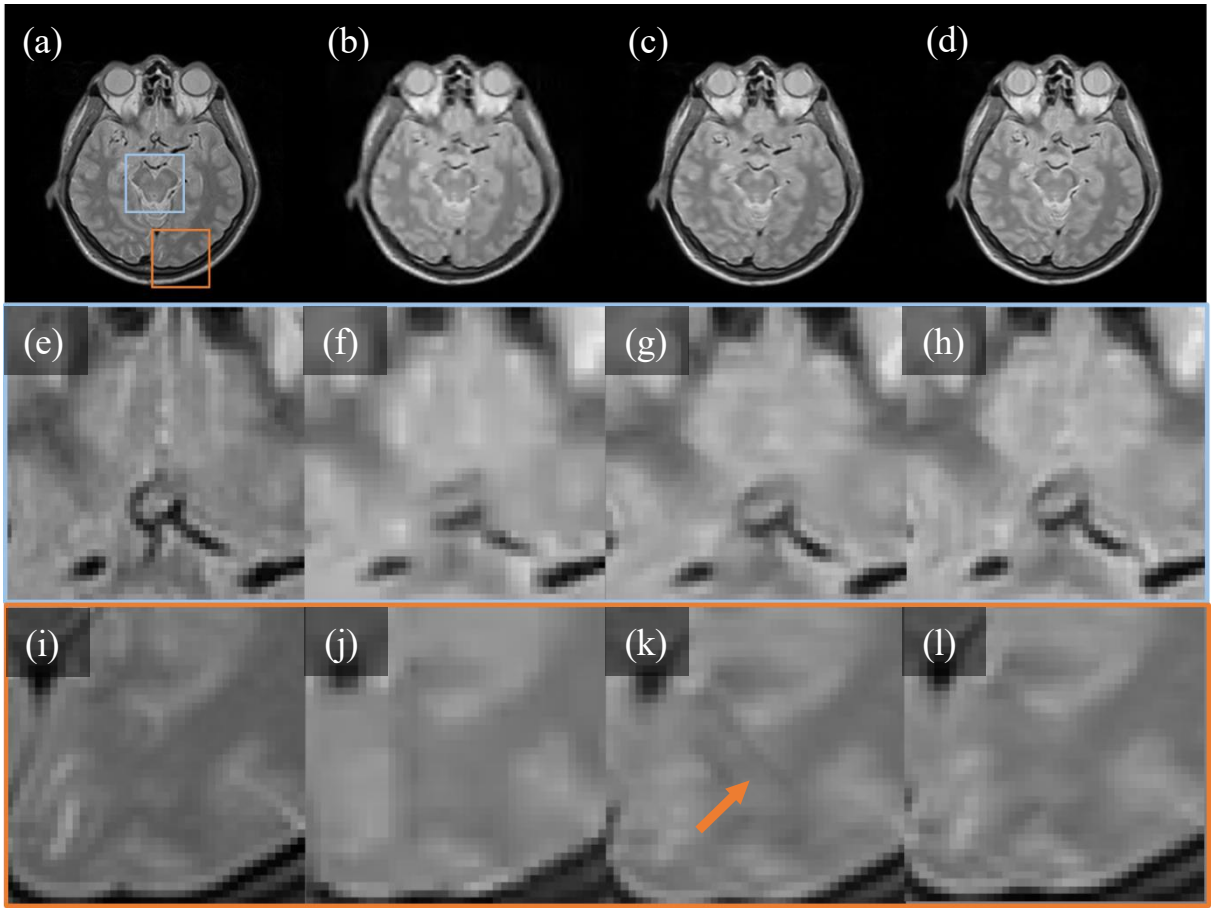


図 3 再構成像 ：

(a) フルデータ再生像, (b) ISTA, (c) ISTA-Net, (d) ISTA-Net-MR(提案法),

(e)~(h) (a)~(d)の青枠部拡大, (i)~(l) (a)~(d)橙枠部拡大

Deep Learning Reconstruction using Refined ISTA-Net Introducing Artifact Learning

Yuta MIYAMOTO[*1], Satoshi ITO[*1]

*1 Utsunomiya University, Graduate School of Regional Development and Creativity, Division of Engineering and Agriculture

Recently, deep learning approaches have achieved success in compressed sensing image reconstruction. Zhang proposed a deep learning network named ISTA-Net that imitates Iterative Soft Threshold Algorithm (ISTA) solving L1-norm problem. When ISTA-Net was applied to compressed sensing image reconstruction problem, aliasing artifacts due to sub-sampling in k-space sometimes remain on the reconstructed images. In this study, a new deep learning approach that connects convolution layer and an activation function to estimate the aliasing artifacts is proposed and demonstrated. Simulation experiments reconstructing two-dimensional images showed that aliasing artifacts were much more removed and the resultant PSNRs and SSIMs are improved compared to Zhang's network.

Key words: ISTA, Compressed Sensing, Reconstruction, Deep Learning

Ring artifact reduction method using 2-directional Fourier transform filter

Akio YONEYAMA[*1, 3], Thet Thet Lwin[2], and Rika BABA[3]

Abstract

Novel 2-dictional Fourie transform filter for sinogram was developed to reduce ring artifacts in sectional images of X-ray micro computed tomography. Vertical stripe lines in sinogram generate ring artifact, and therefore the new filter was designed to suppress the lines by combination of vertical high-pass filter and horizontal low-pass filter. Feasibility study of the new filter using a CT data set obtained by synchrotron radiation based micro CT was performed, and the fine sectional images reduced ring artifact were obtained.

Keywords：CT, ring artifact, reduction, Fourier transform

1. Purpose

Synchrotron radiation (SR) based micro computed tomography (SR-μCT) is a powerful tool for non-destructive observation of sample with μm-order spatial resolution, and widely used in many fields such as biomedical, material, and geological science. X-ray imager with high-spatial resolution was indispensable for observations, and an indirect type imager composed of a scintillator converting the incident X-ray to visible light, a relay lens system, and a visible light camera was commonly used [1]. Crystal scintillator such as CsI and LuAG sometimes contain crystal defects while its high light outputs, and the visible light output was not proportional to the incident X-ray intensity in the defect area. As a result of the defects, ring artifacts greatly reducing the image quality were generated in a reconstructed sectional image. To suppress the ring artifact, we developed a novel filter for sinogram combined with vertical high- and horizontal low-pass filter.

2. Method

X-ray beam in SR-μCT can be assumed as a parallel beam because of the small X-ray source and the long distance from the X-ray source to samples, and therefore the sectional images can be reconstructed separately from each sinogram. Defects of scintillator generate vertical stripe lines in the sinogram as shown in Fig. 1, so ring artifacts in sectional images can be reduced by removing the vertical lines.

---

*1 SAGA Light Source
〔8-7 Yayoigaoka, Tosu, Saga 841-0005, Japan〕
e-mail: yoneyama@saga-ls.jp
2 Kitasato University
3 Hitachi Ltd.

Vertical lines in the sinogram can be removed simply by using a vertical high-pass filter. However, low-frequency signal was also reduced simultaneously, and CT values in the reconstructed sectional image were differ from original CT values. On the other hand, horizontal low-pass filter maintains the low-frequency signal and reduces high-frequency noise. Therefore, a sinogram reduced the vertical lines and maintaining the low-frequency signal can be obtained by addition of the sinograms of filtered by vertical high-pass filter and horizontal low-pass filter as shown in Fig. 1.

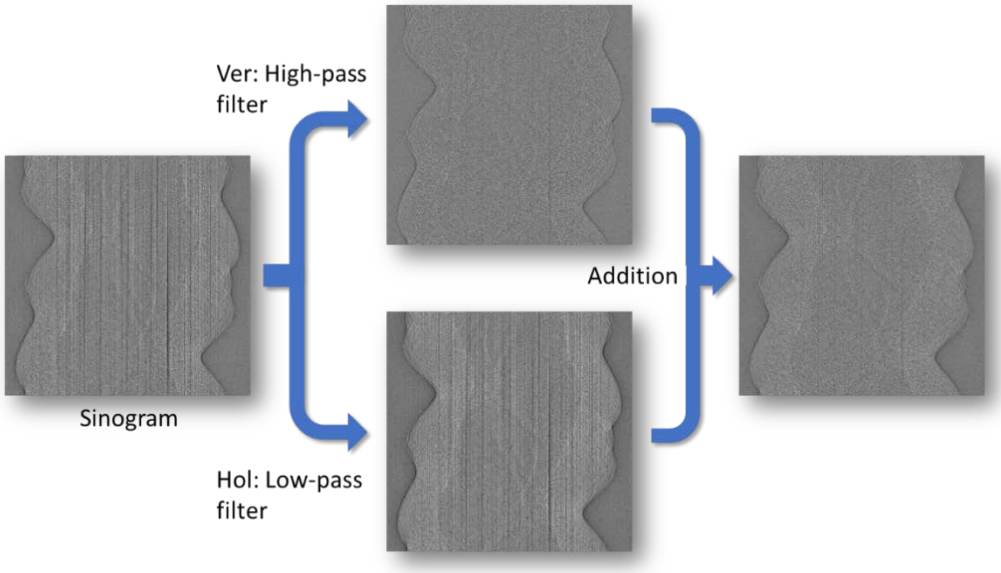


Fig. 1 Newly developed filter for sinogram was composed of vertical high-pass filter and horizontal low-pass filter.

3. Results

Feasibility study of the newly developed filter was performed using a CT data set of a rat bone obtained by SR-μCT at SAGA Light Source in Tosu, Japan. SR of 10-keV was used for the observation, and each projection image was captured with 10 s exposure time. The projection number was set at 1000 for 360 degree. Figure 2 shows the whole reconstructed sectional image (left), and the magnified images indicated as a red rectangle in the left image without the filter (a), with the vertical high-pass filter (b), with the horizontal low-pass filter (c), and with the newly developed 2-directional filter (d). The cut-off frequency and the degree of both filters were set at 5 and 1, respectively. Note that, each high- and low-pass filter was applied to sinograms after Shepp-Logan filter.

Strong ring artifacts caused by defects of the scintillation of the X-ray camera were seen in Fig 2 (a) and image quality was reduced. On the other hand, ring artifacts were almost removed in Fig. 2 (b), but the CT value was lower than original CT values because of the decreasing of the low-frequency signal by the low-pass filter. High frequency noise was suppressed in Fig. 2(c) while the strong ring artifacts was remained. In Fig. 2 (d), ring artifacts were removed, the CT values were similar to the original values, and the high-frequency noise was suppressed. In addition, small structures in the bone can be depicted, and therefore the spatial resolution was kept the same level of the original image.

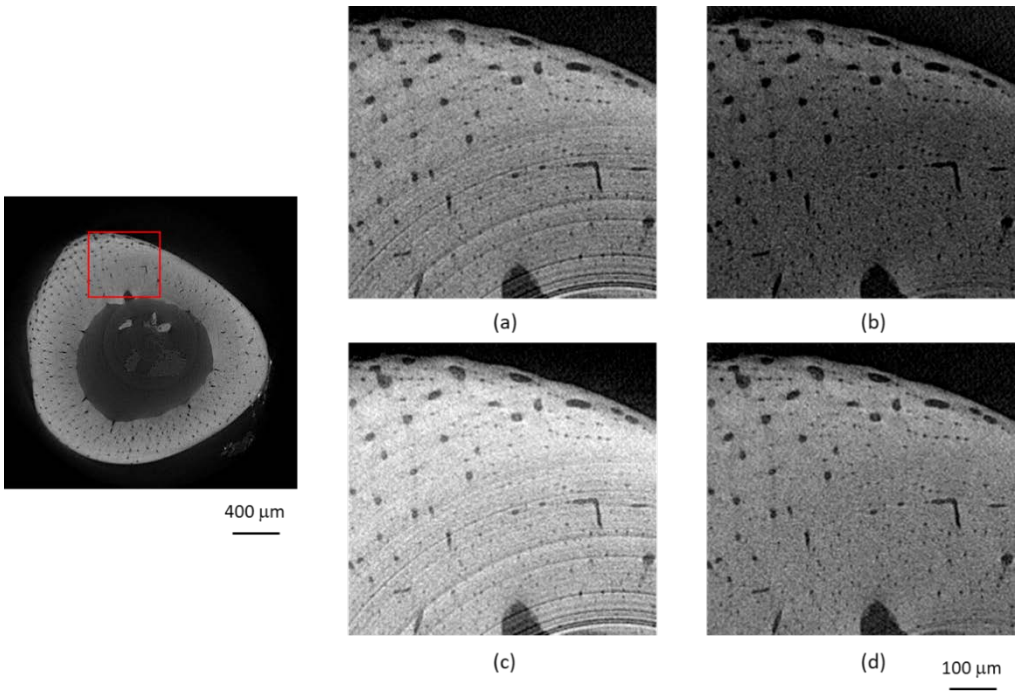


Fig. 2 Reconstructed sectional image of rat bone (left), magnified image without filter (a), with vertical high-pass filter (b), with horizontal low-pass filter (c), and with newly developed (2-directional) filter (d).

4. Conclusion

Novel 2-dictional Fourie transform filter for sinogram was developed to reduce ring artifacts in the sectional images of SR-µCT. Feasibility study of the new filter using a data set obtained by SR-µCT was performed, and the fine sectional image removed ring artifact was obtained. As the next step, we are planning to apply the filter to various CT data sets obtained by SR-µCT at SAGA Light Source for the optimization of the parameter of the new filters.

Competing interests

Akio Yoneyama and Rika Baba are employed by Hitachi Ltd.

Acknowledgement

The biological experiment was approved by the President of Kitasato University through the judgment of the Animal Care and Use Committee of Kitasato University (approval no. 14-02).

References

[1] Yoneyama A, Baba R, Hyodo K et al: Development of high-resolution x-ray CT system using parallel beam geometry. AIP Conference Proceedings, 1696, 020007 (2016).

MRI 高速撮像画像における微細構造を保存したノイズ低減

石原　千鶴枝[*1]，金子 幸生[*1]，白猪　亨[*1]，野口　喜実[*1]，

尾藤　良孝[*2]，荻野　昌宏[*1]

要旨

MRI では撮像時間を短縮するために様々な高速撮像技術が開発されている．しかし，SNR が低下するために画質を担保した高速化が課題となっている．本研究では，3 倍高速撮像の Parallel Imaging（PI）における低画質画像を，SNR および微細構造の明瞭度の観点において，通常撮像と同等の画質に復元する畳み込みニューラルネットワーク（CNN）の開発を行った．ボランティア撮像実験によって取得した k 空間信号から，PI 同等の間引きを行った低画質画像と，フルサンプリング時の通常撮像画像の学習データセットを作成した．更に，ノイズ特徴を表す g-factor マップに基づいてデータを分割し，それぞれによる複数の学習済 CNN を推測に用いた．定義した微細構造のぼやけ度（Blur）と SNR を用いて，提案方式およびその他の方式と比較評価した結果，低 Blur を維持しながら SNR 向上率 70%以上となる性能を確認した．

キーワード：MRI，高速撮像，CNN，デノイズ，微細構造

1．はじめに

MRI（Magnetic Resonance Imaging: 核磁気共鳴撮像装置）は，頭部をはじめとして様々な部位の検査に用いられている．MRI では，例えば頭部ルーチン検査においては 15～20 分程度，長い撮像時間を要するため，k 空間データを特定の間隔で間引いて撮像する Parallel Imaging [1][2]など，撮像時間を短縮するための様々な高速撮像技術が開発されている．しかし，高速撮像下では画像の SNR（Signal-to-Noise Ratio）が低下するため，画質を担保した高速化が課題である．

これまで，我々の研究グループでは，CNN を利用して，高速撮像時の画像を通常撮像時の画像（フルサンプリング画像）相当に復元する手法について検討を行ってきた．同一撮像空間内において，血管などの微細構造や脳実質などの様々な特徴のある画像を同時に復元するために，複数の CNN を使った multi-adaptive 法（MA-CNNR : Multi-Adaptive CNN Reconstruction）を開発した[3][4]．本方式では，学習データを，画像の特徴に沿ってサブセットに分類し，各サブデータセットでそれぞれ学習を行った CNN を作成する．推測時には，画像の特徴に沿って適用する CNN を切り替えることで，単一の CNN よりも高性能な画質改善効果が得られる．

本報告では，MRI 検査スループット向上を目標に，SENSE 法[1]に基づく Parallel Imaging による 3 倍高速撮像画像から，フルサンプリング画像を予測する MA-CNNR を検討する．今回は，Parallel Imaging 時に用いる g-factor マップを利用して，学習データセットをサブセットに分ける方法を用いる．g-factor マップは撮像時の感度

*1 ㈱日立製作所研究開発グループ
〔〒185-8601　東京都国分寺市東恋ヶ窪 1-280〕
e-mail: chizue.ishihara.zn@hitachi.com
*2 ㈱日立製作所ヘルスケアビジネスユニット

情報を反映した分布図で，撮像空間のノイズ量の特徴を示している．このため，ノイズ特徴の類似したサブデータセットによる学習が可能となり，生成される CNN はノイズ量応じた適切なデノイズ性能になると考えられる．更に本報告では，脳実質における SNR と，微細構造の明瞭度について評価する指標を用意し，MA-CNNR とその他の手法を定量的に比較評価する．

第 2 章では研究方法として，学習データ，提案する g-factor マップを活用した MA-CNNR の手法の説明，および手法の評価方法について述べる．第 3 章では結果と考察，第 4 章でまとめを述べる．

2．方法

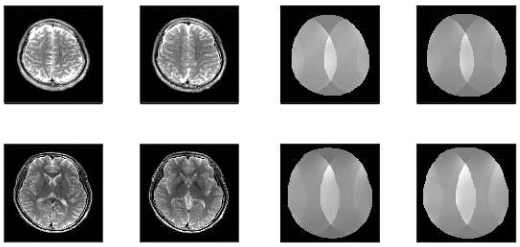
2-1．学習データ

Parallel Imaging（PI 法）では，k 空間上の信号を特定の間隔で間引いて複数受信コイルで同時に計測する．その後，計測した各受信コイルの k 空間データと，受信コイルの感度分布の違いを利用して画像再構成する．本報告では，健常ボランティアによる撮像実験を行い，実験で得られた k 空間データを 1/3 に間引いた信号と，すべての信号，それぞれで再構成した T2 強調画像（PI 画像および Full sampling 画像）を学習データセットとした．

撮像実験は，日立製の 3T MRI 装置を用いて頭部 T2 強調撮像ルーチンプロトコルによって行った．また，ボランティア撮像データは（株）日立製作所研究開発グループで定める倫理審査基準に則り審査され，すべての被験者からインフォームド・コンセントを得た上で取得された．

2-2．g-factor を活用した MA-CNNR

始めに，PI 法における g-factor マップについて説明する．PI 法による再構成画像の SNR は，受信コイルの感度に依存して，同一撮像空間内で異なる値を持つ特徴がある．受信コイルの感度は，PI 法の間引き数を上げるほど撮像空間内において差が生じる．g-factor（geometry factor：各ボクセルの折り返しを元に戻す際の正確さを表す因子）マップは，その受信コイルの感度を反映した撮像空間内の分布図である．図 1 に k 空間データを 1/3 に間引いた PI 法処理下の MRI 画像と，g-factor マップの例を示す．受信コイルから離れる深部の領域で特に g-factor 値が高く，感度としては悪くなり SNR が低下する．本報告ではこの g-factor マップに基づく感度情報を，画像のノイズ特徴量として MA-CNNR に活用する．



(a) MRI images　　(b) g-factor map

図 1　MRI 画像と g-factor マップ

表 1　学習データ数

	Train data	Validation data
All	88,248	10,839
Subset 1 (≦ boundary)	47,960	4,745
Subset 2 (> boundary)	40,288	6,094

次に，データセットについて説明する．まず，実験で取得した学習データセットを，微小サイズのパッチに切り出す．画像 1 枚当たり 512×512 サイズの画素サイズに対し，32×32 サイズのパッチとする．次に，各パッチの領域における g-factor 値の平均値を同時に取得し，その値に沿ってパッチデータをサブセットに分類する．今回，g-factor 値の境界値（boundary）を 1.239 とした 2 種類のデータに分ける．g-factor 値が境界値より低いパッチデータはサブセット 1，高いパッチデータはサブセット 2 とする．最終的に用いた学習データ数を表 1 にまとめる．

最後に，学習および高画質化処理について説明する．g-factor 値に基づいて分類した 2 つの

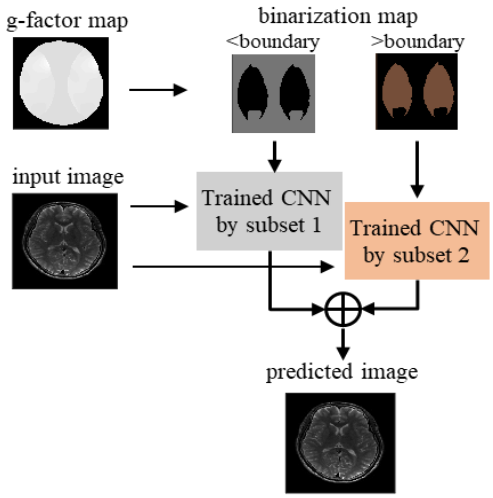


図 2　予測画像生成処理

サブセットデータそれぞれを用いて CNN の学習を行う．作成した 2 つの学習済 CNN を使って，次のような手順で高画質化処理（予測画像の生成）を行う．図 2 に処理の概念図を示す．まず，高画質化処理を適用する画像とその撮像空間における g-factor マップ画像を用意する．画像に対して，2 つの学習済 CNN を適用し 2 種類の出力画像を生成する．g-factor マップ画像からは，学習データの 2 つのサブセットを作る際に用いた g-factor の境界値を下回る領域と，上回る領域のマスク画像を生成する．それぞれを，2 つの出力画像と乗算し，最後に加算処理を行うことで，画像の領域ごとのノイズ量に応じた適切なデノイズ処理が適用される．

なお，本報告の学習および推測処理では，NVIDIA 社製の GeForce RTX 2080 GPU 使用し，Python Keras/Tensorflow のソフトウェアを用いた．また，ネットワークの重みパラメータの最適化（optimizer）には Adam，loss 関数は MSE（mean squared error）を用いた．

2-3.　評価方法

提案手法の評価は，テスト用の PI 画像を入力とした予測処理結果におけるデノイズ効果および微細構造の明瞭度の観点で行う．

まず，デノイズ効果の評価として，SNR を算出する．脳実質における輝度が平坦な領域から，任意の大きさの ROI（Region of interest）を 32 箇所設定し，各 ROI において

$$SNR = \frac{m(I)}{\sigma(I)} \qquad (1)$$

を算出する．m，σはそれぞれ平均値，標準偏差を表し，Iは ROI 内の各画素の信号強度を示す．全 ROI の SNR 平均値と，Full sampling 画像から最も SNR が低下した上位 3 つの ROI の平均値を算出し，各平均値の PI 画像を基準とした変化率（ratio）を評価する定量値として用いる．

次に，微細構造の明瞭度の評価として，画像のぼやけ度（Blur）を算出する．まず，淡い濃淡の微細構造画像に着目し，その画像の Full sampling 画像との差分をとる．PI 画像では，ノイズが重畳するため差分画像における輝度分布は白色雑音のノイズに特徴的な正規分布に従う．一方，それをデノイズ処理すると，同時に微細構造のエッジがなまり，差分画像に構造情報が出現する．この差分画像の輝度分布は，偏りのある形状をした正規分布になると考えられる．この特徴を利用し，Blur は次のように算出する．まず，微細構造画像の領域における Full sampling 画像との差分を取る．その差分値が正規分布に従うと仮定した場合の，分布形状の左右の偏り度（歪度，skewness）を次の式に従って求める：

$$\text{skewness} = \frac{\sum(x-\mu)^3}{N\sigma^3} \qquad (2)$$

x は分布の各データ，μはそれらの平均値，N はデータ数，σはデータの標準偏差である．歪度が 0 であれば分布形状の左右対称を表し，0 から離れると分布が歪むことを意味する．この skewness を評価する Blur の定量値として用いる．

以上に示す画質の定量評価を，次の各手法について行う：ガウス関数型ローパスフィルタによる CNN を用いないデノイズ処理（LPF），SRCNN[5]と Residual[6]方式を組み合わせた CNN によるデノイズ処理（CNN type1），および CNN type1 に層を増やして skip connection[7]を加えたデノイズ処理（CNN type2）とする．CNN type1 および CNN type2 それぞれについて，単

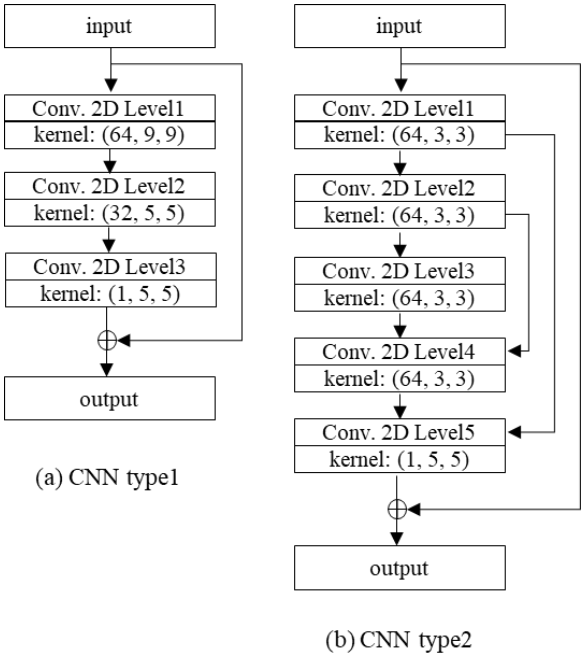


図 3　ネットワーク構成

一の CNN を用いる場合（1-CNNR）と，提案する MA-CNNR 方式を適用する場合（MA-CNNR）とで比較する．CNN type1 および type2 のネットワーク構成は図 3 に記載する．

3. 結果および考察

図 4 に結果画像の 1 例を示す．(a)はあるスライスの MRI 画像で，(b)は(a)に示す ROI の微細構造の拡大図である．(c)は Full sampling 画像との差分画像である．表 2 に各画像における各種定量評価結果をまとめる．Blur は図 4(c)に示す微細構造の差分画像から算出した．

図 4(b)から，LPF ではノイズは低減しているものの，微細構造が不明瞭になっている．一方，CNN においては微細構造の明瞭度は改善しており，MA-CNNR 方式でより明瞭な画像へ復元している．図 4(c)に示す差分画像からも，LPF では構造情報が残っており，MA-CNNR 方式でほとんど残っていないことが分かる．表 2 のぼやけ度を示す Blur からも，MA-CNNR 方式でぼやけが低減していることが分かる．

表 2 に示す，すべての ROI（all ROIs）における SNR ratio では，いずれのデノイズ処理において，Full sampling 相当以上（70%以上）SNR を

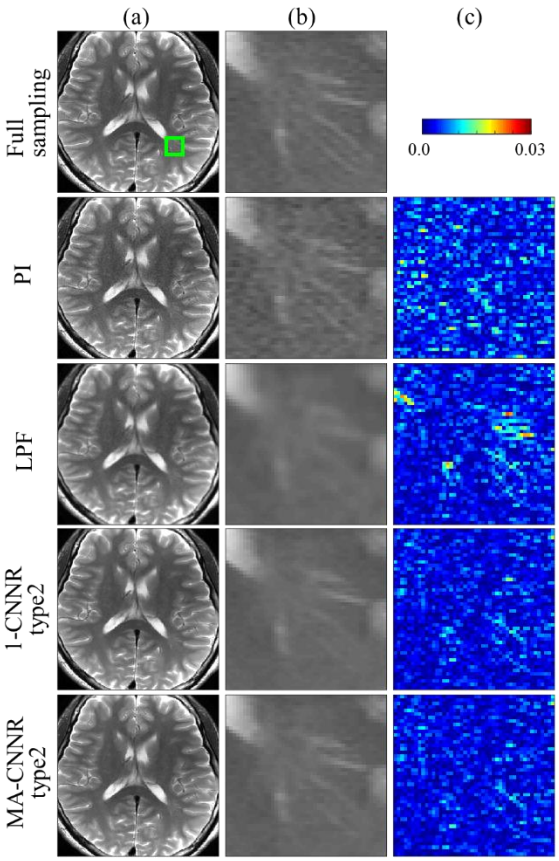


図 4　結果画像

表 2　画質評価結果

		SNR ratio		Blur
		all ROIs	low-SNR ROIs	
Full sampling		1.7	2.3	
PI		1.0	1.0	0.08
LPF		2.1	2.0	0.54
1-CNNR	Type1	2.1	1.8	0.13
	Type2	2.2	2.0	0.18
MA-CNNR	Type1	2.2	2.0	0.12
	Type2	2.2	2.1	0.08

改善している．1-CNNR と MA-CNNR で比較すると，Type1 および Type2 それぞれについて，all ROIs における SNR ratio は同等レベルであるものの，最も SNR が低下した上位 3 つの ROI（low-SNR ROI）においては，より高い SNR の向上を示している．

以上の結果について考察する．表 2 に示す通

り，MA-CNNR 方式を適用することで，Blur は低く維持しつつ， low-SNR ROIs においてより高い SNR 改善を示した．MA-CNNR 方式では，データを g-factor 値に沿って分割することで，ノイズの性質が類似したデータ学習による推定能力を向上させている. 特に SNR の低い領域での復元性能が向上し，構造情報を維持したまま SNR の改善が可能になったと考えられる. また, CNN のネットワーク構成においては, Type2 の方が高い性能を示した．層を追加した高性能な CNN との組み合わせにより，復元性能が向上したものと考えられる．一方で，low-SNR ROIs における SNR ratio は，いずれのデノイズ処理も Full sampling 相当まで到達しなかった. 低 SNR の領域では, 更に高性能なデノイズ処理が必要となるが考えられる．これにおいては，MA-CNNR 方式におけるデータ分割数の追加や, 層を追加した高性能な CNN の組み合わせを用いることで，対応できる可能性は高い.

4. まとめ

MRI 検査スループット向上を目的として，Parallel Imaging による 3 倍高速撮像画像を通常撮像相当の画像に復元する CNN の開発を行った. 撮像空間のノイズ分布の特徴を表す g-factor マップに基づいてデータを 2 分割し，ノイズの性質が類似したデータ学習によって推定能力を向上させる, MA-CNNR 方式を提案した. SNR および微細構造の明瞭度の観点で画質を評価した結果, 微細構造の明瞭度を維持しつつ, 70% 以上 SNR を向上することを確認した．特に，高速撮像によって SNR が低下する ROI において，より高い向上効果を有する．今後，データ分割数の追加や，層を追加した高性能な CNN の組み合わせを用いることで，更なる復元性能の向上が期待できる.

利益相反の有無

著者は株式会社日立製作所の社員である.

文　献

［1］ Pruessmann KP, Weiger M, Scheidegger MB, et al.: SENSE: Sensitivity Encoding for Fast MRI. Magnetic Resonance in Medicine **42**:952-962, 1999

［2］ Griswold MA, Jakob PM, Heidemann RM, et al.: Generalized autocalibrating partially parallel acquisitions (GRAPPA). Magnetic Resonance in Medicine **47**:1202-1210, 2002

［3］ 野口喜実, 荻野昌宏 : クラスタリングに基づく CNN 選択による MRI 高画質化の検討，OP2-8，JAMIT，2018

［4］ 金子幸生，野口喜実，尾藤良孝　他 : マルチ CNN による MRI 画像におけるノイズ低減，OP4-5，JAMIT，2019

［5］ Chao Dong, Chen Change Loy, Kaiming He, et al.: Learning a Deep Convolutional Network for Image Super-Resolution, ECCV, 2014

［6］ Jiwon Kim, Jung Known Lee, Kyoung Mu Lee: Accurate Image Super-Resolution Using Very Deep Convolutional Networks, arXiv: 1511.04587, 2015

［7］ Xiao-Jiao Mao, Chunhua Shen, Yu-Bin Yang: Image Restoration Using Convolutional Auto-encoders with Symmetric Skip Connections, arXiv: 1606.08921, 2016

Noise Reduction while Preserving Fine Structure for Fast Scan MRI

Chizue ISHIHARA[*1], Yukio KANEKO[*1] Tohru SHIRAI[*1], Yoshimi NOGUCHI[*1],
Yoshitaka BITO[*2], Masahiro OGINO[*1]

*1 Hitachi, Ltd. Research & Development Group
*2 Hitachi, Ltd. Healthcare Business Unit

For the scan time reduction in MRI (Magnetic Resonance Imaging), many fast scanning techniques have been developing. Those techniques, however, have an issue to present low SNR. Thus, the fast scanning techniques while preserving high image quality are necessary. In this paper, we proposed a denoising CNN (Convolutional Neural Network) which can restore the image quality from 3-times fast scan imaging to full-sample imaging, in terms of SNR and high-definition of fine structure. We firstly acquired k-space data from the volunteer study by MRI system. By using the data, we generated training data set of the under-sample image (by Parallel Imaging) and full-sample image. The training data were separated to multiple data set according to g-factor map which presents feature of noise in the scanning area of Parallel Imaging. In predict process, we used multiple CNNs which have been trained by the multiple data for each. Blur (newly defined index as a low definition of fine structure) and SNR were evaluated in predicted images, and image restoring performance was compared with that for other denoising filters. As a result, we confirmed that the proposed method presents the performance of more than 70 % SNR enhancement while preserving low Blur.

Key words: MRI, fast scan, CNN, denoise, fine structure

Support Vector Regression を用いた脳年齢推定と機種間 Harmonization

舞草　伯秀[*1]　Iman Beheshti[*2]　曽根　大地[*3]　松田　博史[*1]

要旨

MRI から得られる脳体積(ベイズ推定による事後確率)分布から、機械学習法を用いた回帰モデルに基づく脳年齢推定は、アルツハイマー病や側頭葉てんかん等の脳変性の進行動態を表すバイオマーカーとして着目されている。しかしながら、MRI は磁場強度の違い等により撮像された画像品質が異なるため、得られた解析値に Measurement Bias が存在することが知られており、これが機械学習の汎化性を著しく減少させることが知られている。本研究では、一般線形モデルを経験的ベイズ推定により拡張した ComBat 法を脳体積値に適応し、脳年齢推定の汎化性が向上することを示した。

キーワード：MRI，SVR, Machine learning, Brain Age Estimation, Neurodegenerative disorder

1. はじめに

アルツハイマー病に代表される認知症や側頭葉性てんかんなどに代表される脳変性疾患の診断は、通常、神経心理バッテリーなどの問診・質問紙による脳機能評価により行われる。しかしながら、神経心理バッテリーによる評価では、被検者のコンディション、測定者の習熟度の問題などにより再現性や客観性に課題を有していることが知られている。脳変性疾患の画像診断では、灰白質と呼ばれる大脳皮質が、変性によりその体積が減衰あるいは変化することが知られており、T1 強調 MRI などを用いて、灰白質の体積減少を定量的にとらえられれば、脳変性疾患の進行動態をとらえる、再現性・客観性の高い代理バイオマーカーになりうることが期待されており、これまでその有効性に関する報告が多数なされてきた[1-6]。

MRI を用いた脳体積の分布から Support Vector Regression :SVR などの機械学習法を用いて脳年齢を推定するいわゆる Brain Age Estimation :BAE は、脳灰白質における海馬や偏桃体などの解剖学的関心領域に基づく体積減少などの指標に比べて、単一の評価指標として用いることができるため、臨床応用に適した手法として注目されている。特に Beheshti らは BAE を用いてアルツハイマー病の識別[7]、さらに Sone らは BAE を用いて側頭葉性てんかんの識別が行えることを報告している[8]。

一方 BAE の推定モデル構築には 1000 例以上と多数の MRI 健常データを必要とし、推定モデルの構築に多大なコストがかかることが課題である。しかし MRI は撮像する機種により灰白質のコントラスト、画像信号の均一性、

*1 国立研究開発法人　国立精神・神経医療研究センター　脳病態統合イメージングセンター
〔〒187-8553 東京都小平市小川東町 4-1-1〕
e-mail: maikusa@ncnp.go.jp
*2 Université Laval
*3 University College London
投稿受付：2020 年 8 月 5 日

Signal to Noise Ratio などが異なるため、解析結果に機種間差(Measurement Bias)が存在する。したがって機種間差が存在したままで推定モデルを構築した場合、異なる機種に推定モデルを適応できないという問題が発生する。

そこで本研究では、BAE の臨床応用での有効性を向上させるために、機種間を取り除く ComBat 法[9-11]による機種間差の補正(Harmonization)を適応し、脳年齢の推定モデルを構築したデータセットとは別の検証データセットを用いて脳年齢の推定を行い、推定精度および汎化性が向上していることを検証する。

2 手法

2.1 トレーニングデータセット

本研究では、国立精神・神経医療研究センター病院で撮像された 20 歳以上の健常ボランティアの 3D T1 強調 MRI を用いて SVR による脳年齢推定モデルを構築する。MRI はベンダーが異なる 2 つの機種により撮像されており、撮像パラメータは以下の通りである。

・プロトコル1: 3.0 T MR装置(Philips Medical System社製): repetition time (TR)/echo time (TE), 7.18 ms/3.46 ms; flip angle, 10°; number of excitations (NEX), 1; 0.68 × 0.68 mm^2 in plane resolution; 0.6-mm effective slice thickness with no gap; 300 slices; matrix, 384 × 384; field of view (FOV), 26.1 × 26.1 cm.

・プロトコル2: 3.0 T MR装置 (モデル名:Verio, Siemens社製): TR/TE, 1800 ms/2.25 ms; flip angle, 9°; NEX, 1; 0.87 × 0.87 mm^2 in plane resolution; 0.8-mm effective slice thickness with no gap; 224 slices; matrix, 320 × 280; FOV, 25 × 25 cm.

プロトコル1では、798名の被検者を撮像し、プロコル2では398名の被検者の撮像を行った。画像は放射線読影医が読影を行い、他の器質的疾患がないことを確認し、認知症スクリーニング検査(Mini-Mental State Examinaton: MMSE)により、認知機能低下が発生していないことを確認した。

すべての被検者はインフォームドコンセントに同意し署名を得た。

2.2 検証データセット

2.1 節のデータセット用いて構築した脳年齢推定モデルを用いて、異なる検証データセットにより脳年齢推定を行う。検証データセットとして、Japanese Alzheimer's Disease Neuroimaging Initiative :J-ADNI の健常データセットを用いる。J-ADNI は本邦における 38 の臨床施設で行われたアルツハイマー病進行動態解明のための、60 歳以上の高齢者を対象とした多施設大大規模画像研究である。J-ADNI における健常者の定義は以下のとおりである。

・MMSE が 24 点から 30 点の範囲であること

・Clinical Dementia Rating (CDR)が 0 であること

・うつ病の既往がないこと

・記憶障害がないこと

本研究では、上記の組み入れ基準を満たす 146 名の健常者を対象とした。また本研究で用いる J-ADNI の 3D T1 強調 MRI はすべて 1.5T で撮像されており、ベンダーは Philips Medical System 社, Siemens 社, GE 社の撮像装置が用いられた。どの被検者がどの装置で撮像されたかの対応表については、J-ADNI のデータベースより取得した。

本研究は、国立精神・神経医療研究センターの倫理委員会にて承認を得た。

2.3 画像処理

3D T1 強調 MRI に対して、(Wellcome Trust Centre for Neuroimaging, London, UK; www.fl.ion.ucl.ac.uk/spm/)により、ベイズ推定にて信号不均一補正と、白質・灰白質の事後確率マップへの Segmentation を行う。さらに SPM の DARTEL 法に MNI 標準脳空間へ解剖学的標準化を行い、DARTEL の非線形変換式より Jacobian 行列を求め、事後確率マップの modulation を行った後、半値幅 4 mm の

Gaussian kernel を用いて平滑化を行った。これによりベイズ推定による事後確率は、Voxel-wise 脳体積値として扱うことができる。次に Franke らの手法[12]に従い、各事後確率マップを 、8 mm の等方 Voxel に resample を行い、全脳マスク内の事後確率マップを抽出し SVR への入力 vector とした。

2.4 ComBat Harmonization

トレーニングデータセットおよび検証データセット内の機種間差を軽減するため、入力 vector に対して ComBat 法を適応する。ComBat による機種間差モデルは式(1)で表される。

$$y(i,j,v) = a(v) + \boldsymbol{X}^T(i,j)\boldsymbol{\beta}(v) + \gamma(i,v) + \delta(i,v)\varepsilon(i,j,v) \quad (1)$$

ここで、$y(i,j,v)$は、i 番目の装置で撮像された、j 番目の被験者のv番目の voxel で観測された事後確率である。$\boldsymbol{\beta}$は事後確率に交絡する共変量であり $\boldsymbol{X}$はその計画行列である。γ 及びδは i 番目の装置の measurement bias の定数項と係数項、εは平均 0 の誤差項である。経験的ベイズ推定により、measurement bias の推定値 γ^*及びδ^*を求め、harmonize された観測値は式(2)で表され

$$y_{(i,j,v)}^{combat} = \frac{y_{(i,j,v)} - \hat{a}_{(v)} - \boldsymbol{X}_{(i,j)}\hat{\boldsymbol{\beta}}_{(v)} - \gamma^*_{(i,v)}}{\delta^*_{iv}} + \hat{a}_{(v)} + \boldsymbol{X}_{(i,j)}\hat{\boldsymbol{\beta}}_{(v)} \quad (2)$$

となる。

2.5 SVR による脳年齢推定および統計解析

脳年齢推定のため、IBSVM toolbox (http://www.csie.ntu.edu.tw/~cjlin/libsvm/) の nu-support vector regression (nu-SVR)モデルを用いる。LIBSVM のパラメータは、C = 1, ν = 0.5 を設定した。事後確率マップからの入力は高次元ベクトルとなるため、オーバーフィッティングの軽減と次元削減のため主成分分析を行った。ここでは、第 100 主成分までを SVR への入力として用いる。作成した回帰モデルを評価するため、トレーニングデータセットに対して、10-fold 交差検証を行った。その後、構築した回帰モデルを検証データセットに適応し、回帰モデルの汎化性を検証した。モデルの評価には、推定年齢と実年齢の平均絶対誤差を用いた。はじめに対応のある t 検定を行い、harmonization 前後のトレーニングデータセットおよび検証データセットの平均絶対誤差を比較した。次に harmonization 後のトレーニングデータセットと harmonization 前後の検証データセットの平均絶対誤差を対応のない t 検定を行い、トレーニングデータセットと検証データセットの差を比較した。この時トレーニングデータセットの平均絶対誤差は、検証データセットと年齢範囲を同一にするため 60 歳以上の結果を対象とした。

3 結果

10-fold 交差検証によるトレーニングデータセットおよび検証データセットの実年齢と推定脳年齢の散布図を図 1 に示す。図 1(a)は harmonization 前、図 1(b)は ComBat による harmonization 後の事後確率マップを入力とした結果である。図 1(a)に示す、harmonization 前の平均絶対誤差は、トレーニングデータセットにおいて 4.99 歳、検証データセットでは 7.64 歳であった。一方、ComBat による harmonization を行った後では、トレーニングデータセットでは 4.90 歳、検証データセットでは 5.82 歳であった。

Harmonization 前後の平均絶対誤差を比較したところ、検証データセットでは、平均 1.82 歳の有意な改善(p<.0001)が認められた。一方トレーニングデータセットでは、平均 0.09 歳向上したが有意差は認められなかった(p=.415)。60 歳以上のトレーニングデータセットと検証データセットの平均絶対誤差の比較を行った。harmonization 前ではトレーニングデータセットと検証データセット間に有意差 (p<.0001) が認められた。一方 harmonization 後では、トレーニングデータセットと検証データセットの間に有意差は認められなかった(p=.391)。

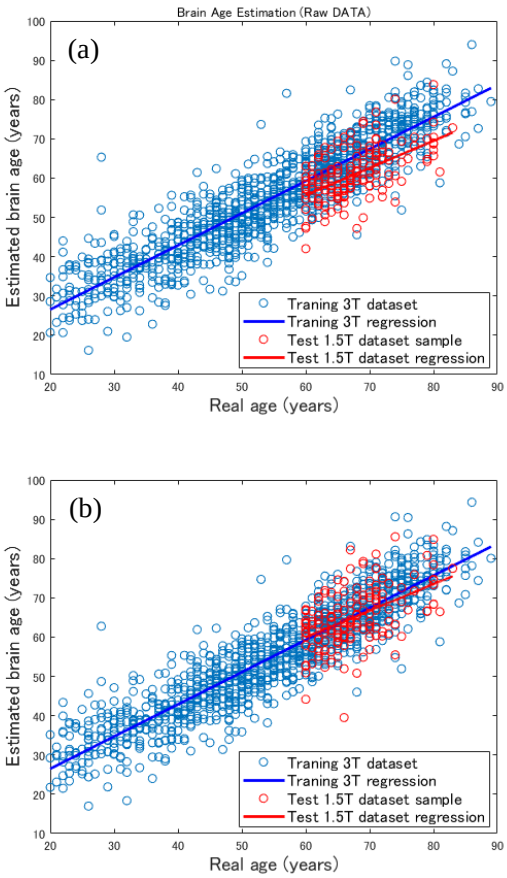


図 1 SVR による脳年齢推定結果　(a) Harmonization 前, (b) ComBat 法による Harmonization 後

4. 考察・まとめ

本研究では、T1 強調 MRI から得られる脳白質･灰白質の事後確率マップから SVR を用いた脳年齢推定行い、推定精度の観点からの汎化性について ComBat 法による harmonization の効果について検証を行った。

SVR のトレーニングデータセットに関して、harmonization による脳年齢推定の精度向上は認められなかった。トレーニングデータセットは 2 つの異なるベンダーの装置による撮像であったが、装置の磁場強度がどちらも 3 T であった。Da ら、異なる磁場強度は T1 強調 MRI を用いた構造解析に影響があるが、同一の磁場強度については大きな影響が見られないと報告しており[13]、同様の結果が得られたものと考えられる。

一方、検証データセットにおける脳年齢の推定精度は、harmonization 前ではトレーニングデータセットと有意に異なっており、異なる磁場強度間では構築した機械学習の回帰モデルが適応できず汎化性を有していないことを示した。

ComBat 法による harmonization は、検証データセットにおける脳年齢の推定精度を有意に改善し、かつトレーニングデータセットとの精度との差をなくした。このことは、作成した機械学習モデルが異なる磁場強度間であっても適応可能であることを示しており、汎化性の向上を意味している。

現在先進国における 3 T MRI 装置の保有率は年々上昇しているが、これまで蓄積された多くの MRI は 1.5T MRI 装置である。MRI を用いた疾患の進行動態や、その発生機序の解明、さらには機械学習による診断・バイオマーカーモデルの構築には、多数の症例・健常例のデータが必要となる。本研究の結果、harmonization による汎化性の向上により、異なる磁場強度間においても、機械学習による解析が可能であるということを示しており、蓄積された過去の医学的レガシィを活用できるという観点から、本研究の成果は意義があるものであると考える。

謝辞

本研究のトレーニングデータセットの撮像にご尽力いただいた国立精神・神経神経医療研究センター病院・放射線診療部部長佐藤典子先生、医長　木村有喜男先生に謝意を示す。検証データセットに用いた J-ADNI データベースの撮像・構築にご尽力いただいた関係各位に謝意を示す。

利益相反の有無

なし

文 献

［1］ Nestor SM, Rupsingh R, Bartha R, et al.: Ventricular enlargement as a possible measure of Alzheimer’s disease progression validated using the Alzheimer’s disease neuroimaging initiative database. Brain **131**, 2443-2454, 2008

［2］ Schuff N, Jack CR, Weiner MW, et al.: MRI of hippocampal volume loss in early Alzheimer’s disease in relation to ApoE genotype and biomarkers. Brain **132**, 1067-1077, 2009

［3］ Hua X, Weiner MW, Thompson PM, et al.: Mapping Alzheimer’s disease progression in 1309 MRI scans: Power estimates for different inter-scan intervals. Neuroimage **51**, 63-75, 2010

［4］ Leung KK, Fox NC, Ourselin S, et al.: Robust atrophy rate measurement in Alzheimer’s disease using multi-site serial MRI: Tissue-specific intensity normalization and parameter selection. Neuroimage **50**, 516-523, 2010

［5］ Leung KK,Fox NC, Ourselin S, et al.: Automated cross-sectional and longitudinal hippocampal volume measurement in mild cognitive impairment and Alzheimer’s disease. Neuroimage **51**, 1345-1359, 2010

［6］ Aisen PS, Petersen RC, Weiner MW et al.: Clinical core of the Alzheimer’s disease neuroimaging initiative: Progress and plans. Alzheimers Dement **6**, 239-246, 2010

［7］ Beheshti I, Maikusa N, Matsuda H: The association between “Brain-Age Score” (BAS) and traditional neuro-psychological screening tools in Alzheimer's disease. Brain and Behavior **8**(8), 1-14, 2018

［8］ Sone D, Maikusa N, Matsuda H et al.: Neuroimaging-based brain-age prediction in diverse forms of epilepsy: a signature of psychosis and beyond. Molecular Psychiatry, 1-10, 2019

［9］ Johnson W E, Cheng L, Rabinovic A: Adjusting batch effects in microarray expression data using empirical Bayes methods. Biostatistics **8**(1), 118–127, 2007

［10］ Fortin JP, Parker D, Shinohara RT et al.: Harmonization of multi-site diffusion tensor imaging data. NeuroImage **161**, 149-170, 2017

［11］ Fortin JP, Cullen N, Shinohara RT et al.: NeuroImage Harmonization of cortical thickness measurements across scanners and sites. NeuroImage **167**, 104–120, 2018

［12］ Franke K, Ziegler G, Gaser C et al.: Estimating the age of healthy subjects from T1-weighted MRI scans using kernel methods: exploring the influence of various parameters. Neuroimage **50**(3), 883-892, 2010

［13］ Da M, Karteek P, Mirza FB, et al.: Quantitative assessment of field strength, total intracranial volume, sex, and age effects on the goodness of harmonization for volumetric analysis on the ADNI database. Hum Brain Mapp **40**(5), 1507-1527, 2019

Brain Age Estimation using Support Vector Regression and Harmonization Across Scanner and Site

Norihide Maikusa[*1], Iman BEHESHTI[*2], Daichi SONE, Hiroshi MATSUDA

*1 National Center of Neurology and Psychiatry

*2 Université Laval

*3 University College London

Brain age estimation based on machine learning from MRI images has been attracting attention as a biomarker of the progression of brain degeneration in such as Alzheimer's disease and temporal lobe epilepsy. However, it is known that measurement bias has been existed in the volume due to differences in image quality of MRI obtained from different magnetic field strength, manufacturer and model, which has been reported to significantly reduce the generalizability of machine learning. In this study, we used ComBat harmonization method, which uses a general linear model and empirical Bayesian estimation to correct for differences between scanner and site, which improved the generalizability and accuracy of brain age estimation across different scanner and site.

Key words: MRI, SVR, Machine learning, Brain Age Estimation, Neurodegenerative disorder

CNN を用いたイオフルパン SPECT からの特異的結合比の推定

久根崎　遥加[*1]　岩渕　怜[*1]　荒木　啓史[*1]

姜　時友[*1]　湯浅　哲也[*1]　松田　博史[*2]

要旨

パーキンソン病（PD: Parkinson's Disease）の診断指標として用いられる特異的結合比（SBR : Specific Binding Ratio）は，ドーパミントランスポーターに高い親和性をもつイオフルパンを造影剤とする SPECT 画像における線条体内の集積から算出される．しかし，SPECT の低い解像度では線条体の位置を特定することが困難な場合がある．本研究では，正確に線条体を特定した画像から Tossici-Bolt 法により算出した SBR 値を学習させた CNN（Convolutional Neural Network）によって，SPECT 画像から回帰により SBR 値を推定する方法を提案する．臨床 SPECT 画像（Control: 192 症例，PD: 1,186 症例）およびそれらから算出した SBR を訓練データとして 3 層 CNN を学習させた後，テストデータ（Control: 21 症例，PD: 136 症例）に対して，CNN により予測された SBR に基づいて ROC（Receiver Operating Characterisic analysis）解析を行ったところ，0.94 の高い AUC（Area Under Curve）を得た．

キーワード：パーキンソン病，特異的結合比，SPECT，CNN，Tossici-Bolt 法

1. はじめに

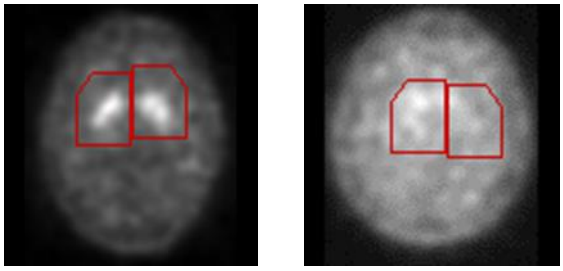
パーキンソン病（PD: Parkinson's Disease）は，黒質線条体のドーパミン神経細胞が変性する運動失調疾患であり，その神経終末に存在するドーパミントランスポーター（DAT: Dopamine transporter）密度が低下していることが知られている[1]．その DAT に高い親和性を有するイオフルパン（^{123}I-FP-CIT）を PD 患者に投与し，SPECT（Single Photon Emission Computed Tomography）で画像化すれば，DAT の脳内分布を確認でき，PD 患者のドーパミン神経の変性や脱落を評価できる．

現在，PD は視覚評価だけでなく定量評価を行うことで診断能の向上を図っている．その定量評価に用いられるのが，特異的結合比（SBR: Specific Binding Ratio）[1,2] であり，

$$SBR = \frac{Cs}{Cr} \qquad (1)$$

と定義される．ここで，Cs は線条体の特異的集積に起因する SPECT カウント濃度，Cr は非特異的集積による参照領域の SPECT カウント濃度を示す．すなわち，SBR とは，イオフルパン投与後平衡状態に至った時点での特異的集積と非特異的集積のカウント比である．SBR の算出には，SPECT 画像に抽出される線条体を含むような VOI（Volume of Interest）を設定する必要がある．現在の医療現場では，VOI 設定は放射線技師が目視で決めることが一般的であるが，VOI 設定が作業者の主観に委ねられるため，SBR の客観性は損なわれる．先行研究では，客観性を確保するため，SPECT 画像の画素値から画像処理によって線条体を含む VOI を自動設定し，Tossici-Bolt 法により SBR を算出する全

*1 山形大学大学院理工学研究科応用生命システム工学専攻
〔〒992-8510 米沢市城南 4-3-16〕
*2 国立精神・神経医療研究センター



(a) 成功例　　(b) 失敗例

図 1　VOI 設定結果

自動化処理を開発した. しかし, 1,573 データ中, 38 データで VOI の自動設定に失敗した. 図 1 に VOI 設定の成功例及び失敗例を示す. 失敗症例は, すべて PD 患者であった. 図 1(b)のように, 失敗症例では，PD の進行により線条体へのイオフルパンの集積が乏しく，線条体領域を正確に捉えることができず，VOI の設定に失敗し，SBR が算出できなかったと考えられる．MR（Magnetic Resonance）または CT 画像を SPECT 画像に重ね合わせ，線条体に対応する領域を特定することで高精度に SBR を算出できるが, 臨床検査では患者への負担軽減のため, SPECT 画像のみ撮影される場合が多い．PD が進行した場合，イオフルパン集積が乏しいため，SPECT 画像のみからは正確な SBR を得ることは難しい. 一方, 近年注目される Deep Learning を用いれば，データ数が確保できれば，ロバストな処理が期待できるため，線条体が不鮮明な場合にも SBR が算出できることが期待される. 本研究では，Deep Learning の手法の 1 つである CNN（Convolutional Neural Network）[3-5] を用いて，SPECT 画像から回帰により SBR 値を算出する方法を提案し，臨床データに適用することで有効性を検証する.

2. 処理手法

2.1　使用したデータセット

本研究で用いた臨床画像データは，USC（University of Southern California）の PPMI（Parkinson's Progression Markers Initiative）で公開されているイオフルパン SPECT 画像を用いた．使用データの SPECT 画像に関する情報を表 1 に示す.

表 1　SPECT 画像情報

画像サイズ [pixel (D×H×W)]	91×109×91
ボクセル解像度 [mm^3]	2×2×2
データ数	1,535 （Control: 213，PD: 1,322）

表 2　学習用データ

	Train data		Test data
	Original	Augmentation	
Control	192	960	21
PD	1,186	5,930	136

先行研究で作製した自動 VOI 設定処理を適用した後，処理結果を目視で確認し，VOI 内に線条体が含まれていると判断できる 1,535 症例のデータセットを学習データおよびテストデータとして用いた．これらのデータに対して，それぞれ Tossici-Bolt 法により SBR 値を求めた（表 2）．これらのデータから，学習データとして，Control 192 例，PD 1,186 例をランダムに抽出した．さらに，学習データは水増しにより表 2 のように増大させた．水増し方法は，実際にあり得ない変化を与えてはいけないため，本研究では，± 3 pixel の範囲でランダムに平行移動させることで作成した．残りのデータをテストデータ（Control 21 例，PD 136 例）としてモデル評価に用いた.

入力画像は 3 次元画像のためデータ量が大きくなることから，画像を縮小した．その際，$n = 4$ の条件でバイキュービック法を用いた．結果として，91×109×91 の画像は 91×27×23 の画像に縮小される．画像縮小後は，画素値を 0 から 1 にスケーリングし, ネットワークに入力した.

表 3　CNN 構造

Layer	Filter size	Stride	Output size	Activation
Input			(91, 27, 23, 1)	
Convolution	(3, 3, 3)	(1, 1, 1)	(91, 27, 23, 32)	ReLU
Max Pooling	(2, 2, 2)	(2, 2, 2)	(45, 13, 11, 32)	
Convolution	(3, 3, 3)	(1, 1, 1)	(45, 13, 11, 64)	ReLU
Convolution	(3, 3, 3)	(1, 1, 1)	(45, 13, 11, 64)	ReLU
Max Pooling	(2, 2, 2)	(2, 2, 2)	(22, 6, 5, 64)	
FC			42240	ReLU
FC			512	ReLU
FC			2	ReLU

2.2　CNN

モデルへの入力を SPECT 画像，正解ラベルを左右それぞれの線条体における SBR 値とし学習を行った．本研究で用いた CNN の構造を表 3 に示す．畳み込み層 3 層，全結合層 3 層の計 6 層とした．全ての畳み込み層では，フィルターサイズを 3×3×3 とし，ストライドは 1×1×1 とした．プーリング層は．2×2×2 サイズのフィルターとし，ストライドは 2×2×2 とした．表 3 の Output size は（奥行き，高さ，横幅，チャンネル）で表される．各層において，活性化関数は ReLU を用いた．出力層では，左右線条体それぞれの SBR 値を出力する．

CNN のパラメータを表 4 に示す. 損失関数として利用した Huber 損失関数は，誤差を$E(\boldsymbol{x})$とし，モデルから出力された値を$f(\boldsymbol{x})$，モデルから出力された値に対応する正解データを $\boldsymbol{y}$ とすると，式(2)で表される．本研究では，$\boldsymbol{x}$ を SPECT 画像，$\boldsymbol{y}$ を SBR 左右値，δ を 1.0 とした．

$$E(\boldsymbol{x}) = \begin{cases} \frac{1}{2}|f(\boldsymbol{x}) - \boldsymbol{y}|^2 & \cdots |f(\boldsymbol{x}) - \boldsymbol{y}| \geq \delta \\ \delta|f(\boldsymbol{x}) - \boldsymbol{y}| - \frac{\delta}{2} & \cdots other \end{cases} \quad (2)$$

本研究で使用した計算機のスペックを表 5 に示す.

表 4　CNN パラメータ

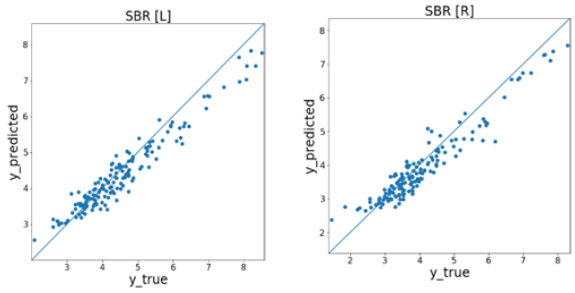
Loss Function	Huber Loss
Optimizer	RMSprop
Batch Size	5
Epoch	15

表 5　計算機スペック

CPU	Intel(R)Core(TM)i7-8700 CPU 3.40GHz
OS	Windows 10
メモリ	8.00GB
開発環境	Jupyter Lab

3. 実験および結果

表 2 で示したテストデータを，学習させた CNN に入力して予測した SBR 値を出力 $\hat{y}_i$ として得る．一方，従来の自動 VOI 設定処理により得られた結果から導かれる SBR 値を y_i とする．横軸に y_i，縦軸に $\hat{y}_i$ として作成した yyplot を図 2 に示す．図 2(a)および(b)はそれぞれ左右線条体における SBR 値に対応する．図 2 より, 直線 $y = x$ のまわりに予測値が集積しているため，概ね良好に予測ができているといえる．しかし，図 2 において SBR 値が 6.0 より大きい領域では, 点群が直線 $y = x$ から下方へ逸脱する傾向が見られる．これは，学習に用いたデータにおける PD 群の症例数が Control 群に対して約 6 倍多かったことに起因するものと考えられる．



(a) SBR(L)算出結果　(b) SBR(R)算出結果

図 2　SBR 値の yyplot

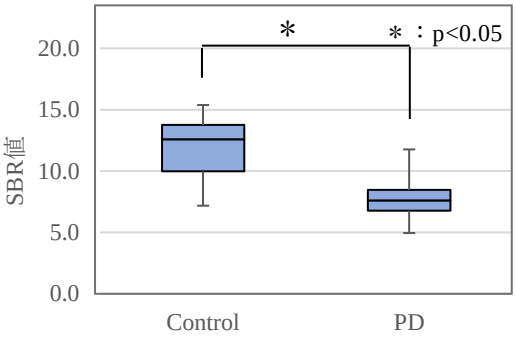


図 3　予測 SBR 値の分布

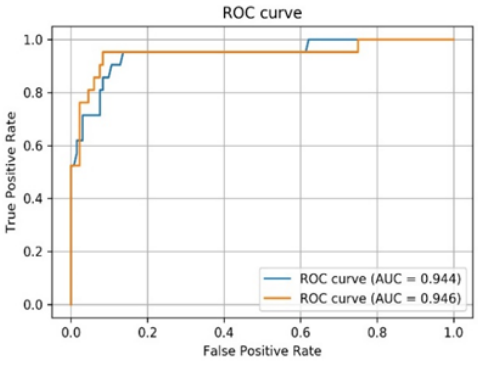


図 4　ROC 曲線

表 6　SBR 値予測精度評価

	RMSE	MAE	RMSE/MAE
L	0.3766	0.2940	1.2810
R	0.3870	0.3054	1.2671

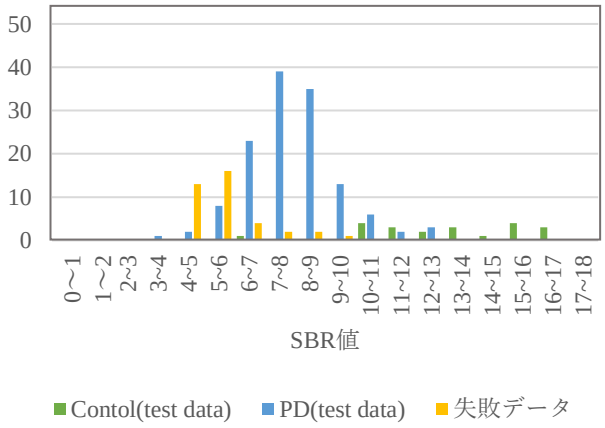


図 5　予測 SBR 値の分布

図 2 の yyplot を RMSE（root mean squared error）と MAE（mean absolute error）を用いて定量的に評価する：

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \qquad (3)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \qquad (4)$$

前者は二乗誤差であるために外れ値が含まれると大きな値を取るのに対して，後者は絶対値誤差であるために外れ値の影響を受けにくい．誤差 $y_i - \hat{y}_i$ が平均値 0，標準偏差 σ の正規分布に従う場合，MAE に対する RMSE の比 $RMSE/MAE$ は $\sqrt{\pi/2} \approx 1.253$ となる．良いモデルが構築できたとき，モデルはデータの大まかな特徴を表現し，正規分布に従うようなノイズが誤差として残ると考えられる．この場合，$RMSE/MAE$ は 1.253 に近い値を取ることが期待される．大きく予測を外すデータが多くなると分子の RMSE が相対的に大きくなり，一方，誤差が一様分布に近い場合，RMSE が相対的に小さくなる．すなわち，$RMSE/MAE$ は，前者では1.253よりも大きな値を取り，後者の場合 1.253 よりも小さな値を取る．図 2 で得られたデータに対する$RMSE/MAE$ を表 6 に示す．左側および右側線条体に対して，それぞれ 1.281

および 1.267 であり，若干大きめの値を取るものの，1.253 に十分近い値を取ることより，本モデルは良好に SBR 値を予測していると考えられる．

次に，Control 群と PD 群のそれぞれにおける予測 SBR 値の分布を図 3 に示す．まず，両者の分散に差があるかを F 検定（$p < 0.05$）したところ，有意差が確認された．次に，両群に対して，不等分散の場合の t 検定（$p < 0.05$）を実施したところ，有意差が認められた．したがって，予測された SBR 値を用いて PD の診断が可能であることが示唆された．さらに，これらの分布を用いて ROC（Receiver Operating Characteristic analysis）曲線を作成した（図 4）．青線が本研究，オレンジ線が先行研究による ROC 曲線である．両者はほぼ同じ ROC 曲線を描くことがわかる．AUC（Area Under Curve）を求めた結果，両者約 0.94 を得た．以上の結果から，CNN は先行手法である画素値情報に基づいた方法と同精度で SBR 値を予測できることが示された．

1. 節で述べたように，先行研究による SBR 値の算出では，38 データで SBR 算出に失敗した．これらの失敗症例に提案手法を適用し，SBR 値を予測した．失敗症例から予測された SBR 値を，テストデータの Control 群及び PD 群から予測された SBR 値を比較した（図 5）．緑色はテストデータの Control 群，青色はテストデータの PD 群の分布を示す．黄色は，従来手法で算出に失敗したデータ群に対して，提案手法を適用して予測した SBR 値の分布である．従来手法で算出に失敗したデータはすべて PD あり，しかも PD が進行し線条体へのイオフルパンの集積は乏しかった．図 5 において PD 群は全体的に左側に分布しており，Control 群は右側に分布している．また図より，先行手法での失敗データ群は PD 群側に分布していることがわかる．さらに，先行手法での失敗データ群は PD 群の中においても左側に分布している．すなわち，失敗したデータ群の SBR 値は PD 群の中でも比較的低い値を示している．この結果は SPECT 画像を目視での評価と矛盾はせず，妥当なものであった．したがって，本手法は PD 群において線条体がはっきりと描出されていないデータに対しても適用できることが示唆された．

4. まとめ

先行研究では，SPECT 画像の画素値情報から線条体を含む VOI を自動的に設定し，その後 Tossici-Bolt 法によって SBR 値を算出した．しかし，PD 群のデータではイオフルパンの線条体への集積量が少ないため，線条体の形状を画素値情報だけから推定するのが困難な場合が少なからずあった．本研究では，先行研究で良好な結果を得たデータを学習データとして用いて CNN を学習させ，SBR 値に回帰させる手法を提案した．先行研究で成功したデータ群に対しては，本手法と先行手法とは同等の精度で SBR 値を算出することができた．また，先行手法で失敗したデータ群に対して本手法を適用して予測された SBR 値は，他の結果と矛盾しない妥当な値であった．本手法は線条体が判別できない画質の劣化した画像に対しても SBR 値を算出できることが示唆された．

利益相反の有無

なし

文　献

［1］ イオフルパン診療ガイドライン第 2 版，日本核医学会日本脳神経核医学研究会（2017）

［2］ L. Tossici-Bolt, ”Quantification of[^{123}I]FP-CIT SPECT brain images:an accurate technique for measurement of the specific binding ratio,” *European Journal of Nuclear Medicine and Molecular Imaging,* vol. 33, no. 12, pp. 1492-1499 (2006)

［3］ R. Vaillant, C. Monrocq, Y. LeCun, “Original approach for the localisation of objects in images,” *Proc. Vision, Image, and Signal Processing*, vol. 141, *pp*. 245–250 (1994)

［4］ D. Simard, P. Y. Steinkraus, J. C. Platt, “Best practices for convolutional neural networks,” *Proc. Document Analysis and Recognition*, pp. 958–963 (2003)

［5］ S. Lawrence, C. L. Giles, A. C. Tsoi, A. D. Back, “Face recognition: a convolutional neural-network approach,” *IEEE Trans. Neural Networks*, vol. 8, pp. 98–113 (1997)

C 型 Compton-PET ジオメトリの 3 次元イメージングシミュレーション

仁科　匠[*1]　田島　英朗[*2]　田久　創大[*2]

錦戸　文彦[*2]　菅　幹生[*1,2,3]　山谷　泰賀[*2]

要旨

例えば MRI に後付けできる PET など，リングの一部を開放化した C 型 PET により PET の応用が広がる可能性がある．その一方で，測定データの欠損に起因した強いアーチファクトが再構成画像に発生してしまう．そこで我々は，開放部と対向する位置に散乱検出器を追加して，欠損情報をコンプトンカメラの原理により補う C 型 Compton-PET の開発を進めている．本研究では，C 型 Compton-PET のアーチファクト低減効果をモンテカルロ・シミュレーションによって検討した．具体的には，半径 20 cm，開放部の角度が 115 度の C 型 PET の内側に，散乱検出器を半径 15 cm で円弧状に配置したジオメトリを模擬した．視野中心に配置した半径 10 cm，体軸方向の長さ 15 cm の円柱ファントムを再構成し，定量的に評価した結果，視野領域全体における円柱内部の画素値の割合が増加したことからアーチファクト低減に有効であることが示された．

キーワード：PET，コンプトンカメラ，画像再構成，モンテカルロ・シミュレーション，PET/MRI

1. はじめに

量子科学技術研究開発機構放射線医学総合研究所では MRI (Magnetic Resonance Imaging) 用コイル（RF (Radio Frequency) コイル）と PET (Positron Emission Tomography) 検出器を一体化することで，既存の MRI を手軽に PET/MRI へのアップグレード可能にする「Add-on PET」の開発を進めている[1]．しかし，シンチレータに加え，光検出器や電子回路をシールドボックスの内側に積層する必要があるため PET リングの厚みは少なくとも 5 cm 程度以上となってしまう．そのため，フルリング型の PET の検出器では，図 1(a)のように寝台が干渉してしまうため，MRI 検査の体勢のまま被験者にセットアップすることができない．開発された試作機では，検出器を寝台上に載せて撮像を行っているため，セットアップから撮像に至るまで頭部を持ち上げる必要があり，被験者にも負担がかかるなど実用上困難な課題が多い．そこで，リングの一部を開放化した C 型 PET を検討した[2][3]．C 型 PET の開放部を下にして配置することで，図 1(b)のように寝台上にいる被験者に負担をかけずに撮像を行うことができるため，Add-on PET の利便性が向上する．また C 型 PET はリ

*1 千葉大学大学院融合理工学府基幹工学専攻医工学コース
〔〒263-8522 千葉市稲毛区弥生町 1-33〕
e-mail: tnishina@chiba-u.jp
*2 量子科学技術研究開発機構放射線医学総合研究所
*3 千葉大学フロンティア医工学センター
e-mail: tashima.hideaki@qst.go.jp
投稿受付：2020 年 8 月 5 日

ングの体軸方向のみならず，開放部からも患者へアクセスすることができる．そのため，全身型に応用することができれば，腫瘍を観察しながらがん治療ができるようになるなど PET の応用を広げられる可能性がある．

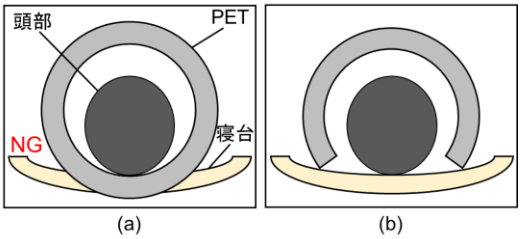


図 1 Add-on PET の現状と展望．(a) フルリング型の撮像，(b) 部分リング型の撮像．

その一方で，C 型 PET の開放部に向かうγ線対は片側しか検出できないため，同時計数線を取得することができない．そのため，測定データの欠損に起因した強いアーチファクトが再構成画像に発生してしまう．

そこでγ線イメージング手法の一種であるコンプトンカメラに注目した．コンプトンカメラは散乱検出器でコンプトン散乱し，かつ吸収検出器で光電吸収したイベント（コンプトンイベント）を検出し，それぞれ測定されたエネルギーE_1とE_2からγ線の入射角度を計算することで，入射方向をコンプトンコーンと呼ばれる円錐面で制限する[4][5]. 複数のコンプトンイベントを収集することで，シングルγ線が発生した位置の推定が可能となる．

我々は，図 2 のように C 型 PET を吸収検出器とみなし，その開放部と対向する位置に散乱検出器を追加して，欠損情報をコンプトンカメラの原理により補う C 型 Compton-PET [6]の開発を進めている．これは PET とコンプトンカメラを組み合わせた新しいイメージング技術「WGI (Whole Gamma Imaging)」[7]を応用したものである． PET・コンプトン両方の観測データを組み合わせて画像再構成を行うことでアーチファクトの低減が期待できる．本研究では，C 型 Compton-PET ジオメトリのアーチファクト低減効果をモンテカルロ・シミュレーションによって検討した．

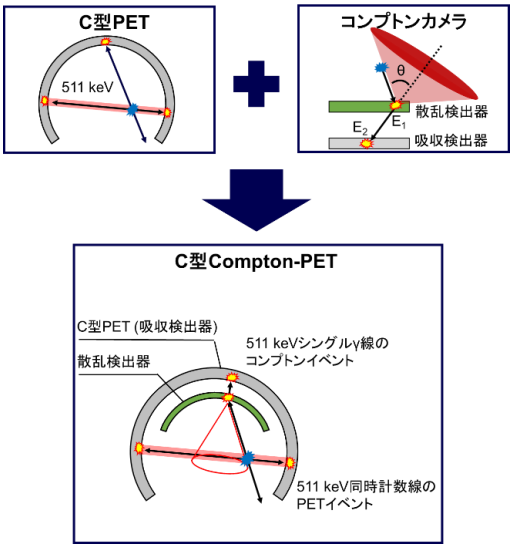


図 2 C 型 Compton-PET の概要．

2. 方法

1）シミュレーションジオメトリ

Geant4 Application [8]を用いて，図 3 のように半径 200 mm，開放部の角度が 115.2 度の C 型 PET の内側に，散乱検出器を半径 150 mm で円弧状に配置した C 型 Compton-PET ジオメトリを模擬した．ジオメトリの主なパラメータを表 1 に示す．C 型 PET の検出器は WGI で使われている検出器[7]を参考に GSOZ とし，エネルギー分解能は実際の測定データから設定した．また，シンチレータの深さ情報を知ることが出来る 4 層 DOI 検出器[9]とし，円弧状に検出器を 1 列あたり 17 個，そして体軸方向に 4 列配置した．

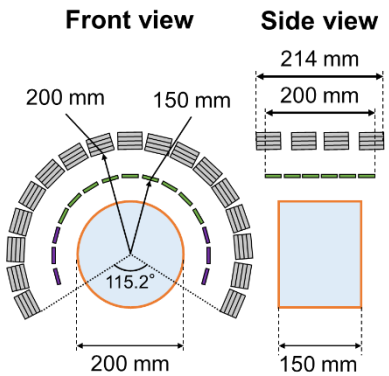


図 3 シミュレーションジオメトリ．

表 1 ジオメトリの主なパラメータ

	散乱検出器			吸収検出器
素材 (密度 [g/cm^3])	GAGG (6.63)	Si (2.32)		GSOZ (6.71)
エネルギー分解能 @511 keV	8.9%	2.5%		13.7%
素子サイズ [mm^3]	1.0×1.0×6.0			2.85×2.85 ×7.5
結晶数	30×30×1			16×16×4
検出器数	9×6	9×6	15×6	17×4
リング半径 [mm]	150			200
中心角 [°]	90.0			90.0
ステップ角 [°]	15.0			14.4

散乱検出器の素材は GAGG と Si の 2 種類を検討した．GAGG は現在 WGI の散乱検出器として使われている素材である．一方で Si は高価ではあるが，優れたエネルギー分解能と相互作用におけるコンプトン散乱の寄与が大きいことから，散乱検出器に適した素材として検討した．GAGG のエネルギー分解能は実際の測定データから[7]，Si は WGI 用に向けて開発中の Si チップ検出器のデータを参考に設定した[10]．散乱検出器のジオメトリは，欠損イベントとなるγ線を検出できるように，円弧状に 1 列あたり 9 個，体軸方向に 6 列配置した．また，Si 散乱検出器の場合，吸収検出器の角度に近づけることでより広い角度範囲のγ線をコンプトンイベントとして検出できるように，1 列あたり 15 個，体軸方向に 6 列配置したジオメトリについてもシミュレーションを行った．また，比較のために C 型 PET のみの場合でも同様にシミュレーションを行った．

撮像対象として，視野中心に配置した半径 100 mm，体軸方向の長さ 150 mm の円柱ファントムを設定した．円柱ファントムから一様な確率でγ線対が放出されるようにし，検出器と相互作用を起こしてエネルギーを付与したイベントを記録した．そしてγ線対の発生を 1.35×10^9 回繰り返した．なお，円柱ファントムの減弱の影響はないものとした．

2）同時計数イベント抽出

モンテカルロ・シミュレーションでは，散乱検出器（S）と吸収検出器（A）でエネルギーを付与したすべてのイベントを出力した。PET 及びコンプトンの同時計数イベント抽出のためのイベント選択フローを図 4 に示す．散乱検出器のエネルギーの検出下限を考慮して GAGG は 50 keV，Si は 30 keV 以下のデータを切り捨てた．PET イベントは 3 パターン（A-A, A-S, S-S）の同時計数を抽出した後，γ線が光電吸収されたイベントを弁別するためにエネルギーウィンドウを適用した．コンプトンイベントの同時計数は，散乱検出器数が 1 列あたり 9 個の場合は，散乱検出器と吸収検出器で相互作用が 1 回ずつ発生したイベントを抽出した．一方で，一列あたり 15 個の場合はγ線対の片方のみもしくは両方共がコンプトンイベントとして検出される場合があるとして，散乱検出器と吸収検出器での相互作用発生回数がそれぞれ 1 回または 2 回のイベントから S-A のペアを複数作成してコンプトンイベントの候補とした．その際，隣同士の検出器となるデータは含めないようにした．S-A のペアを抽出した後，図 5 のように同時計数を抽出する検出器のペアに制限をかけることでγ線対の片方でコンプトン散乱，もう片方で光電吸収を起こしているイベントなどをカットした．また 1 列あたりの散乱検出器数が 15 個のジオメトリの場合は，この制限によって複数のペアの中から正しい確率の高い同時計数のイベントを抽出することが出来る．その後，エネルギーウィンドウを合計エネルギー，散乱検出器のみのエネルギーでそれぞれ別々に適用した．GAGG は後方散乱の影響を考慮し，計測するエネルギーの上限を 150 keV

に設定した．また散乱検出器のエネルギー分解能を考慮し，GAGG の場合は合計エネルギーのエネルギーウィンドウを Si より広く設定した．PET イベントとコンプトンイベントの感度は，抽出された同時計数イベント数を発生させたγ線対数で割ることで算出した．

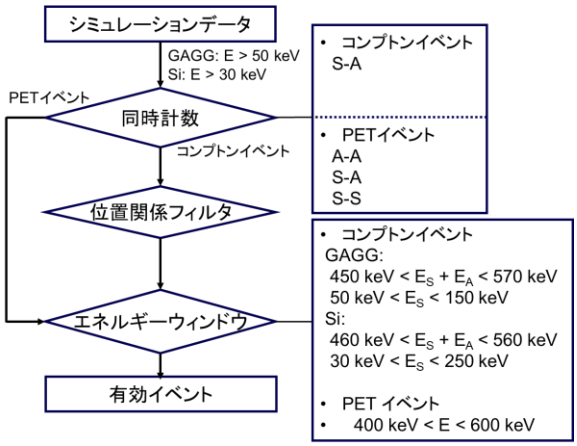


図 4 イベント選択フロー．

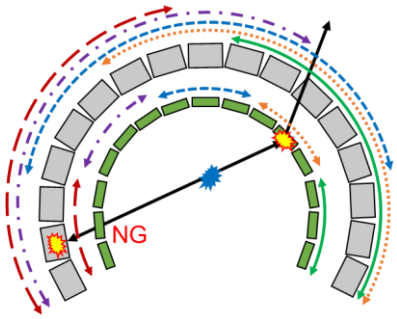


図 5 位置関係フィルタ．上記のようなイベントは感度に含めないように同時計数を同じ矢印のイベントに制限した．

3）画像再構成法と評価方法

PET イベントとコンプトンイベントを組み合わせたハイブリッド画像再構成法を適用した[11]．この方法は逐次近似型画像再構成法である 3D LM-OSEM (List-Mode Ordered Subset Expectation Maximization) 法を基にしている．具体的には，逐次近似における現在の画像から，サブ反復ごとに PET イベントとコンプトンイベントのサブセットをそれぞれ別々に用いて更新画像を作成し，2 つの画像の平均を更新画像とした．画像再構成におけるパラメータを表 2 に示す．

表 2 イメージングパラメータ

反復回数	10
サブセット数	8
ボクセルサイズ [mm^3]	2.0×2.0×2.0
FOV [mm]	Radius: 150 Axial length: 200

C 型 PET，3 種類の C 型 Compton-PET の再構成画像におけるアーチファクトを定量的に評価するため，均一性とリカバリー係数の 2 つの指標を定義した．

まず，円柱線源の上部・下部にあたる 3 次元領域を ROI (Region Of Interest)とし，それぞれ ROI1，ROI2 とした．各 ROI の画素値の合計値を算出し，SUM_{ROI1}，SUM_{ROI2} と定義した．その際，円筒形の ROI を設定するために Bilinear 法を使って再構成画像のボクセルサイズを 1.0×1.0×1.0 mm^3 に変換した．バックグラウンドを含む FOV (Field Of View) でも画素値の合計を算出し，SUM_{FOV} と定義した．ジオメトリの構造から SUM_{ROI2} が SUM_{ROI1} より低くなることを予想し，上部と下部の均一性を式(1)で評価した．

$$\text{均一性 (\%)} = \left(\frac{SUM_{ROI2}}{SUM_{ROI1}}\right) \times 100 \qquad (1)$$

更に，FOV 全体の画素値に対する円柱内部の画素値の割合をリカバリー係数として，式(2)で評価した．

$$\text{リカバリー係数 (\%)} = \left(\frac{SUM_{ROI1}+SUM_{ROI2}}{SUM_{FOV}}\right) \times 100 \qquad (2)$$

3．結果

C 型 PET と 3 種類の C 型 Compton-PET の再構成画像を図 6，感度評価結果を表 3，定量評価結果を図 7 に示す． C 型 PET の再構成画像には，開放部による測定データの欠損に起因したアーチファクトが円柱下部の左右の領域に生じていた (図 6(a))．その一方で，散乱検出器数が 9 個，6 列の C 型 Compton-PET は C 型 PET より正確に円柱を再現できた(図 6(b)，(c))．ま

た素材が Si の場合，C 型 PET より PET 感度が低下した(表 3)．一方で GAGG の場合，PET 感度は増加した．合計感度では Si，GAGG どちらの場合も C 型 PET の PET 感度を上回った．

散乱検出器数が 15 個，6 列のジオメトリは 9 個，6 列のジオメトリより円柱を再現できた(図 6(d))．PET 感度は C 型 PET より低下したが，コンプトン感度は大幅に増加したため合計感度は最も高かった(表 3)．また定量評価結果から，C 型 Compton-PET は C 型 PET より均一性，リカバリー係数ともに増加した(図 7)．

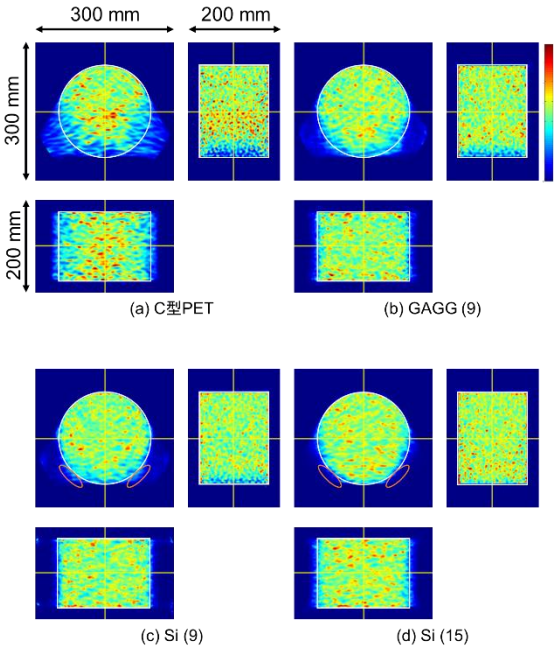


図 6 C 型 PET と GAGG 散乱検出器，Si 散乱検出器を追加した C 型 Compton-PET の画像再構成の結果．白線で囲まれた部分は円柱線源内部を示す．(a) C 型 PET，(b) GAGG，一列あたりの検出器数 9 個，(c) Si，一列あたりの検出器数 9 個，(d) Si，一列あたりの検出器数 15 個．

表 3 感度評価結果

	C 型 PET	GAGG (9)	Si (9)	Si (15)
PET 感度	2.87%	2.91%	2.74%	2.52%
コンプトン感度		0.73%	0.80%	1.47%
合計感度	2.87%	3.64%	3.54%	3.99%

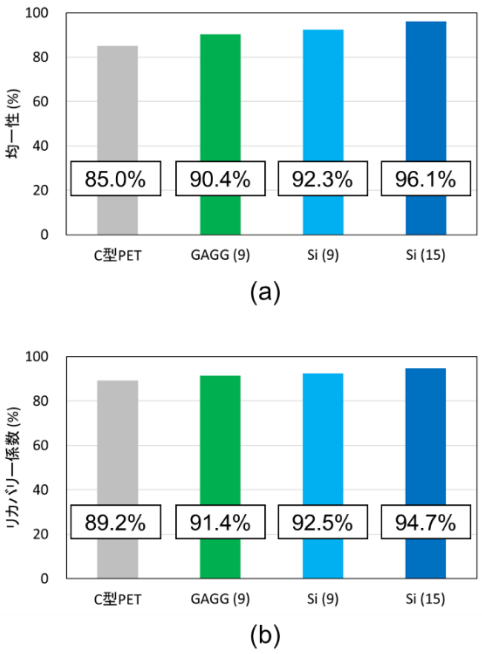


図 7 定量評価結果. (a) 均一性, (b) リカバリー係数.

4. 考察

C 型 Compton-PET では欠損イベントをコンプトンイベントとして補うことで，C 型 PET の円柱下部に生じたアーチファクトを低減できた．図 6(b)，(c)と図 7 から，散乱検出器の素材が Si の方が GAGG よりアーチファクトを低減できていた．これは Si のエネルギー分解能が GAGG より優れているため，検出したエネルギーの誤差が小さくなり，円柱下部の領域の推定に必要なコンプトンコーンの誤差を抑えられたためと考えられる．また表 3 から，GAGG の場合に PET 感度が C 型 PET より増加したが，Si の場合は減少した．GAGG は Si より密度も高く，相互作用における光電吸収の寄与が大きいため散乱検出器での光電吸収のイベントが増えたことがあげられる．更に C 型 PET よりもリング径が小さく立体角が大きいため，PET 感度が C 型 PET より増加したと考えられる．Si は相互作用におけるコンプトン散乱の寄与が大きいため PET イベントで検出できたγ線対の片方を散乱したためと考えられる．

図 6(c)，(d)と図 7 から，横方向のコンプトンイベントを取り入れることはアーチファクト低減に効果的であった．コンプトンイベントの

データから，散乱検出器を頂点とする線積分の情報が取得できる[12]とした場合，検出器が再構成部分を囲み様々な角度からデータを取得することで角度欠損の問題が緩和されると考えられる.1列あたりの検出器数が9個の場合，オレンジ色に囲んだ領域に対して斜め 45 度方向の投影データを取得するには不十分であったが, 15 個にして散乱検出器がカバーする角度を増加させることでイメージングに十分な投影データを取得できたことが考えられる.

今回のシミュレーションではジオメトリの条件を揃えるために,散乱検出器の厚みを 6 mm に設定した．定量評価の結果，Si の方が GAGG よりアーチファクトが低減したため，一列あたりの検出器数が 15 個のジオメトリで追加検討を行った．しかし，厚さ 6 mm の Si 検出器の作成は困難であるため，実際にある程度の厚みを実現するためには，検出器を何層か積層する必要がある．その場合，深さ情報を取得できるためイメージング精度の更なる向上が期待できる. また今回 GAGG のアーチファクト低減結果は Si よりも劣っていたが，6 mm 以上の厚みをもたせることは容易であるため，GAGG の適切な厚みを感度の面から決定し，散乱検出器のジオメトリを素材の特性に合わせて最適化した場合についても検討する必要がある.

また，本研究では C 型 Compton-PET のアーチファクト低減効果のみ評価を行ったが，今後，点線源を使ったシミュレーションを行い，位置による分解能依存を検討する予定である.

5. 結論

Geant4 を用いたイメージングシミュレーションによって，C 型 PET に散乱検出器を追加した C 型 Compton-PET のアーチファクト低減効果を検討した．その結果，C 型 Compton-PET のジオメトリは, C 型 PET における測定データの欠損をコンプトンイベントとして補うことができ，アーチファクト低減に有効であることが示された.

謝辞

本研究は JSPS 科研費 20K12683 の助成を受けたものです.

利益相反の有無

なし

文　献

［1］ Nishikido F, Fujiwara M, Tashima H, et al.: Development of a full-ring “add-on PET” prototype: A head coil with DOI-PET detectors for integrated PET/MRI. Nucl Instr Meth Phy Res A**863**: 55-61, 2017

［2］ Furuta M, Kitamura K, Ohi J, et al.: Basic Evaluation of a C-Shaped Breast PET Scanner. In IEEE Nucl Sci Symp Conf Rec, Orlando, 2009, M05-1

［3］ Kim H I, Chung Y H, Lee K, et al.: Preliminary results of a prototype C-shaped PET designed for an in-beam PET system. Nucl Instr Meth Phys Res Sect A **822**: 57-62, 2016

［4］ 村山秀雄 : 医学物理学教科書シリーズ : 核医学物理学. 国際文献社, 東京, 2017, pp280-283

［5］ Takeda S, Aono H, Okuyama S, et al.: Experimental results of the gamma-Ray imaging capablity with a Si/CdTe semiconductor Compton camera. IEEE Transactions on Nuclear Science **56**: 783-790, 2019

［6］ 田島英朗, 山谷泰賀: コンプトン検出器を組み合わせた部分リング PET ジオメトリの提案. 第 38 回日本医用画像工学会大会予稿集, 奈良, 2019, pp221-224

［7］ Yoshida E, Tashima H, Nagatsu K, et al.: Whole gamma imaging: a new concept of PET combined with Compton imaging. Physics in Medicine & Biology **65**: 125013, 2020

［8］ Geant4：Overview. https://geant4.web.cern.ch/ (2020 年 7 月 29 日閲覧)

［9］ Hirano Y, Nitta M, Inadama N, et al.: Performance evaluation of a depth-of-interaction detector by use of position-sensitive PMT with a super-bialkali photocathode. Radiol Phys Technol **7**: 57-66, 2014

［10］ Nishikido F, Uenomachi M, Zhihong Z, et al.: Development of silicon detectors of whole gamma imager for low energy gamma ray measurement, IEEE NSS&MIC, Sydney, 2018, M-22-3

［11］ Tashima H, Yoshida E, Wakizaka H, et al.: Whole gamma imaging of 89Zr-injected mouse: comparison of PET and Compton imaging. IEEE NSS MIC, Manchester, 2019, MIC-19-02

［12］ Smith B: Line-reconstruction from Compton cameras: data sets and a camera design. Optical Engineering **50**: 053204, 2011

3D Imaging Simulation of a C-shape Compton-PET geometry

Takumi NISHINA[*1], Hideaki TASHIMA[*2], Sodai TAKYU[*2],
Fumihiko NISHIKIDO[*2], Mikio SUGA[*1], Taiga YAMAYA[*2],

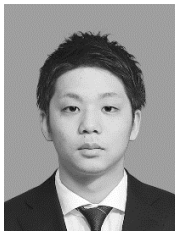
*1 Chiba University

*2 National Institute of Radiological Science, National Institute for Quantum and Radiological Science and Technology

C-shape PET, in which a part of the ring is open, may expand the application of PET such as clinical insert for an existing MRI system. However, the measurable projection data of C-shape PET are truncated, and strong artifacts are generated in the reconstructed images. Therefore, we have been developing C-shape Compton-PET which is C-shape PET with a scatterer insert on the opposite side of the open part, that can compensate for the truncation by applying the Compton camera technique. In this work, we conducted a 3D imaging simulation using the Geant4 Monte Carlo toolkit to demonstrate C-shape Compton-PET. We modeled C-shape Compton-PET geometry, in which scatterer (15 cm inner diameter) was arranged in an arc and inserted into a C-shape PET (20 cm inner diameter and the opening angle of 115 degrees). We simulated a uniform cylindrical phantom with a radius of 10 cm and a height of 15 cm and the reconstructed image of that was evaluated quantitatively. As a result, the percentage of pixel values in the cylinder against the entire FOV was increased by the scatterer insert, and the cylinder shape was reproduced more accurately than the C-shape PET. In conclusion, we demonstrated that C-shape Compton-PET is effective for reducing image artifacts.

Key words: PET, Compton camera, Image reconstruction, Monte Carlo simulation, PET/MRI

著者紹介



仁科　匠（にしな　たくみ）
2020 年千葉大・工学部・メディカルシステム工学科卒．現在，同大大学院修士課程･融合理工学府･基幹工学専攻･医工学コース在学中．日本医学物理学会，日本医用画像工学会学生会員．



田島　英朗（たしま　ひであき）
2009 年東京工業大学総合理工学研究科博士課程修了．博士（工学）．同大学像情報工学研究施設研究員，放射線医学総合研究所（放医研）博士研究員，日本学術振興会特別研究員 PD を経て，放医研研究員，現在は主任研究員として放医研独自の試作装置等含む核医学装置における画像再構成法の研究開発に従事．日本医用画像工学会，応用物理学会等会員．



田久　創大（たきゅう　そうだい）
2010 年千葉大学理学部物理学科卒. 2015 年東北大学大学院博士後期課程修了. 博士（医工学）. 2015 年から 2016 年まで東北大学大学院工学研究科特任助教. 2016 年から現在まで量子科学技術研究開発機構放射線医学研究所にて博士研究員として PET やコンプトンカメラの放射線検出器開発に従事.



錦戸　文彦（にしきど　ふみひこ）
早稲田大学理工学部応用物理学科，卒業. 同大学院理工学研究科にて博士(工学)取得後，早稲田大学理工学研究センター助手を経て，現在量子科学研究開発機構･放射線医学総合研究所･先進核医学基盤研究部に研究員として在籍.



菅　幹生（すが　みきお）
1997年3月京都大学大学院人間・環境学研究科博士後期課程退学．2003年3月同大学院論文博士取得．現在，千葉大学フロンティア医工学センター准教授．京大博士（人間・環境学）．MRIおよびPET装置の研究開発を中心とした生体医工学の研究に従事．日本磁気共鳴医学会代議員，ISMRM，IEEE，日本医用画像工学会，日本生体医工学会等会員．



山谷　泰賀（やまや　たいが）
2000年博士（工学）取得。ポスドク・助手を経て、2004年放射線医学総合研究所研究員、2009年よりグループリーダー（現職）。病気で困らない未来を一日でも早くをモットーに、独自アイディアを具現化する研究を推進。千葉大フロンティア医工学センター客員教授として学生研究指導にも従事。

深層学習による組織組成比推定を介した PET 吸収係数画像生成

小林 哲哉[*1]　繁木 結衣[*1]　森田 尚孝[*1]

要旨

定量的 PET 再構成のためには，高精度な散乱補正・吸収補正が求められ，そのためには正確な吸収係数画像が必要となる．PET-CT 装置や PET-MR 装置では，CT 画像・MR 画像を処理することで吸収係数画像を生成する一方，トランスミッション線源を有さない PET 単独装置では PET データ自体から吸収係数画像を推定する必要がある．ここで，頭部の吸収係数画像は「軟部組織・骨・空洞・被写体外」の 4 つの組織組成比画像の線形結合でモデル化できるので，散乱補正・吸収補正をかけていない未補正 PET 画像から組織組成比画像を推定できれば，間接的に吸収係数画像を推定できると考えた．本研究では，未補正 PET 画像から組織組成比画像を推定するというアイディアおよびそれを実現する深層学習モデルを提案し，頭部 PET のシミュレーションデータを用いた原理検証実験（吸収係数画像と PET 画像の精度評価）の結果を報告する．

キーワード：PET，吸収係数画像，吸収補正，深層学習，セマンティックソフトセグメンテーション

1. はじめに

PET (positron emission tomography)は，生体に投与した放射性薬剤の時空間分布（放射能分布）を，消滅放射線の同時計数データに基づいて推定するトモグラフィー技術である．消滅放射線は生体組織と相互作用するため，同時計数データは散乱線を含み，不均一に減弱されている．そのため，定量的な PET 画像を得るためには，散乱線分布と透過率分布の両方を高精度に推定し補正すること（散乱補正，吸収補正）が求められ，そのためには，正確な吸収係数画像が必要となる．

PET-CT 装置や PET-MR 装置といった一体型装置では，形態画像（CT 画像・MR 画像）を処理することで吸収係数画像を生成する．また，トランスミッション線源を有する PET 単独装置では，透過率分布を実測する．一方，トランスミッション線源を有さない PET 単独装置では，同時計数データ自体から吸収係数画像を推定する必要がある．仮に推定が可能となれば，吸収係数画像の生成に伴う患者負担（追加の被ばく・撮像時間）を無くすことができる．すなわち，一体型装置よりも非常に安価な単独装置で，形態画像を必要とせずに，定量性と患者負担低減を両立した PET 検査が可能となる．

飛行時間差計測型 PET (TOF-PET) の場合，同時計数データから吸収係数画像の投影データを半定量的に復元できることが 2012 年に Defrise らによって証明され[1]，以後，様々なアルゴリズムが提案されているが，適用可能な被写体が数学的に限定される，同時計数データの統計誤差に弱い，散乱線分布をはじめとした補正データの系統誤差に弱い，といった問題を抱えている[2]．また，MR 画像に基づく吸収補正も課題を抱えている[3].

*1 株式会社島津製作所 基盤技術研究所
〔〒619-0237 京都府相楽郡精華町光台 3-9-4〕
e-mail: t_kobaya@shimadzu.co.jp

このような背景もあり，近年，深層学習を応用した吸収係数画像推定が精力的に研究されており，同時計数データのみを用いた手法も複数提案されている[4-6]．筆者の知る限り，全ての既存手法は，散乱補正・吸収補正をかけていない PET 画像（以下，未補正 PET 画像）を畳み込みニューラルネットワークで直接的に吸収係数画像に変換するアプローチを採用している．

本研究では，未補正 PET 画像から「組織組成比画像」を推定し，間接的に吸収係数画像を推定するというアイディア及びそれを実現する深層学習モデルを提案する．組織組成比画像とは，1 つのボクセルに混在する複数の物質の割合（0 以上 1 以下の値）を画素値とするマルチチャンネル画像（チャンネル数＝物質数）であり，チャンネル方向の画素値の和は 1 となる．「組織組成比」の英語表記は技術分野によって様々であるが，composition fraction, mixture ratio, membership ratio 等がそれに相当する．ここで，頭部の吸収係数画像は，大まかに，軟部組織，骨，空洞，被写体外（空気）という 4 つの組織で構成されていると見なすことができ，幸いに，ヒトでの各組織の平均的な吸収係数（線減弱係数）は既知である．結局，図 1 に示すように，頭部の吸収係数画像は，吸収係数を係数とし，組織組成比画像を基底画像とした線形結合でモデル化できるため，組織組成比を推定できれば，吸収係数画像を生成できる．既存手法[4-6]では，吸収係数の値が適正範囲を逸脱する可能性があるが，提案手法ではその可能性はない．図 2 に，本研究が見据える，組織組成比推定を含む PET 画像生成処理のフローを示す．なお，組織組成比推定は，ソフトセグメンテーション（ファジーセグメンテーション）という古典的な画像処理問題に属し，今日的な言葉で表現すれば，セマンティックソフトセグメンテーション[7]であり，問題設定として特に目新しくはなく，応用として新しい．

第 2 章～第 5 章では，それぞれ，組織組成比推定モデルの学習，頭部 PET のシミュレーションデータを用いた原理検証実験（吸収係数画像と PET 画像の精度評価），考察，結論を述べる．

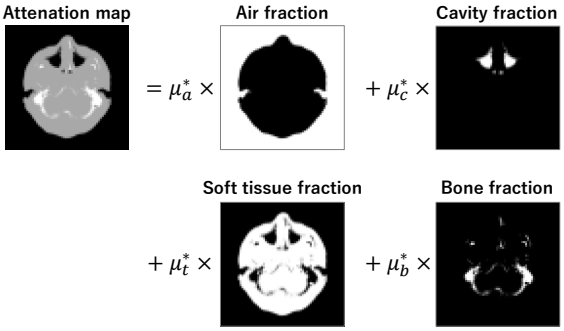


図 1 組織組成比画像の線形結合による吸収係数画像のモデル化（鼻腔が写ったスライスで例示）．

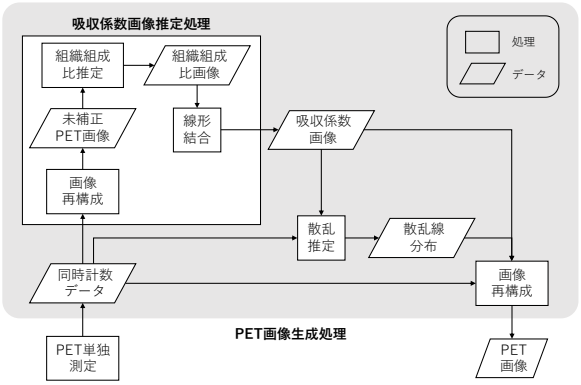


図 2 組織組成比推定を介した吸収係数画像生成に基づく PET 画像生成処理のフロー

2. 組織組成比推定モデルの学習

1）ニューラルネットワーク構成

ニューラルネットワークは U-net [8]を改変した構成であり，入力は 2 次元の未補正 PET 画像（64×64，4[mm]ボクセル），出力は 4 チャンネルの組織組成比画像（64×64×4）である．入力画像の画素値は，スライス毎に 0 以上 1 以下の値に正規化した．組織組成比には「0 以上 1 以下」，「チャンネル方向の和が 1」という制約があるので，それを自動的に満足する softmax 関数を出力層の活性化関数とした．

2）学習データ生成

学習データは，3 次元頭部ラベル画像のオ

ープンデータベース BrainWeb [9]（症例数：21 症例, 領域ラベル数 : 12, ボクセルサイズ：0.5[mm]等方）を改変して利用し，それに基づいたシミュレーション計算により作成した．具体的には，フルオロデオキシグルコース(FDG)を模擬した放射能濃度（灰白質：白質= 4:1）及び平均的な線減弱係数を各ラベルに割り当てて，放射能分布画像及び吸収係数画像を作成し，自作の 3 次元 PET 計測シミュレータに，装置パラメータ等と併せて入力して，ノイズなしの即発計数サイノグラムと偶発計数サイノグラムを計算した．即発計数に占める散乱計数および偶発計数の割合（以下，散乱割合, 偶発割合）はそれぞれ 25[%]とした．即発計数 32, 64, 128, 256, 512[M]に相当するポアソンノイズを付加し，3-D OSEM 法[10]により未補正 PET 画像（偶発補正あり）を再構成した．ポストフィルタはかけていない．一方，組織組成比画像は，元のラベルにない空洞を加えた 13 の領域ラベルを，軟部組織・骨・空洞・被写体外の 4 つの領域ラベルに統合した後に，4[mm]ボクセルの条件で作成した．21 症例のうち 12 症例を学習データ用，6 例を検証データ用，3 例を評価データ用に用いた．図 3 に学習データに含まれる未補正 PET 画像及び組織組成比画像の一例を示す．

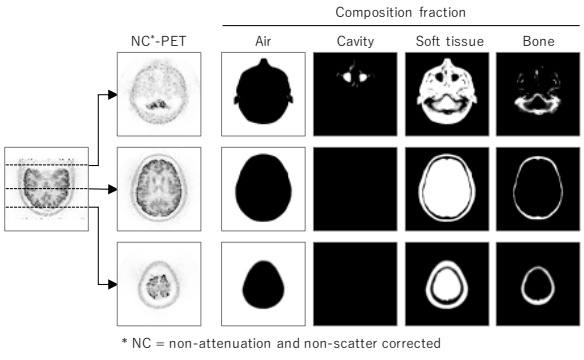


図 3 学習データ（症例 ID = 04）の一例

3）学習条件

各種の学習条件は，事前検討（省略）を経て次のように設定した．損失関数：Cross Entropy，最適化手法 : Adam（学習率 0.001），エポック数：1500，データ拡張：左右反転，拡大縮小(±5%), 水平・垂直シフト(±12mm). なお，ノイズ量の異なる画像は混ぜて学習せずに，特定のノイズ量に特化した 5 つのモデル（以下，ベースモデル）を構築した．

3.4 節で示すとおり，ベースモデルを局所的な高集積部位を含む画像（学習データには含まれない画像）に適用すると，組織組成比の推定精度が大きく低下した(図 10, 図 11)．この問題に対処するために，未補正 PET 画像（画素値範囲[0,1]）の画素値を 1/n 倍した画像(n は 2～9 の整数)も学習データに加えて，ベースモデルを初期値として追加学習した（最大 500 エポック)．このモデルを追加学習モデルと呼ぶ．

3. 原理検証実験

1）実験のねらい

ねらいは 2 つある．ひとつ目は，未補正 PET 画像から組織組成比を推定するというアプローチが深層学習によって実現できる見込みがあるか検証すること，二つ目は極めて限定的な条件で学習したモデルの汎化性能を調べることである．頭部 PET 画像の多様性は，概ね「放射能分布の多様性」と「画像化システムの多様性」の組合せで捉えることができ（表 1），本実験では，例外的な放射能分布パターンは評価対象から外し，また，装置条件・画像化条件は固定して，残りの多様性を生む要因（以下，変動軸）について評価した．つまり，学習データに含まれない変動を与えた画像を用いて，モデルの汎化性能（入力画像の変化に対する不変性・追従性）を調べた．

表 1 頭部 PET 画像の多様性

放射能分布の多様性	・頭部・脳の解剖学形状 ・大域的な放射能分布パターン ・局所的な放射能分布パターン ・例外的な放射能分布パターン
画像化システムの多様性	・頭部・脳の位置・姿勢 ・装置条件 ・同時計数データの品質 ・画像化条件

表 2 変動パラメータの範囲. 太字:学習時の条件. 下線：図 4 の画像の条件.

変動軸	変動させた範囲
(a)	8:1, **4:1**, 2:1, 1:1, 1:2, 1:4, 1:8
(b)	[x] 0.85, 0.90, **1.00**, 1.10, 1.15 [y] 0.85, 0.90, **1.00**, 1.10, 1.15
(c)	0.125, 0.250, 0.500, **1.000**, 2.000, 4.000, 8.000
(d)	0.70, 0.85, **1.00**, 1.15, 1.30
(e)	**1.0**, 2.0, 4.0, 8.0
(f)	位置：(x, y) = (40, 0), (0, 40), (-40, 0), (0, -40) [mm] 倍率：1.0, 2.0, 4.0, 8.0
(g)	[x] -16, -8, **0**, 8, 16 [mm] [y] -16, -8, **0**, 8, 16 [mm]
(h)	-20, -10, **0**, 10, 20 [degree]
(i)	**32**, **64**, **128**, **256**, **512** [M]
(j)	Scatter: 15, **25**, 35 [%] Random: 15, **25**, 35 [%]

2）評価データ生成

生成方法は学習データと同じである．本実験で考慮した PET 画像の変動軸は,

(a) 薬剤［灰白質と白質の集積比］
(b) 頭部の大きさ［倍率］
(c) 皮膚集積量［画素値倍率］
(d) 灰白質集積量［画素値倍率］
(e) 線条体集積［画素値倍率］
(f) 腫瘍集積［位置・画素値倍率］
(g) 頭部の位置［平行移動量］
(h) 頭部の姿勢［回転角度］
(i) 統計ノイズ［即発計数］
(j) 散乱割合・偶発割合［同左］

の 10 個の軸である. 角括弧内は変動させた数値パラメータであり，表 2 に具体的な数値範囲を示した．また，図 4 に画像例を示す．白質集積型（灰白質集積＜白質集積）の放射能分布は，アミロイドイメージング薬剤に相当する．薬剤，統計ノイズ，散乱割合・偶発割合が異なる画像は，再構成計算を実施して生成したが，それ以外の変動軸は，計算時間を短縮するために，学習データと同じ条件で作成した未補正 PET 画像から画像処理により作成した．なお，複数の変動を組合わせた画像での評価も考えられるが，組合せが膨大になり，実験結果の解釈が難しくなるため，今回の原理検証実験では実施を見送った．

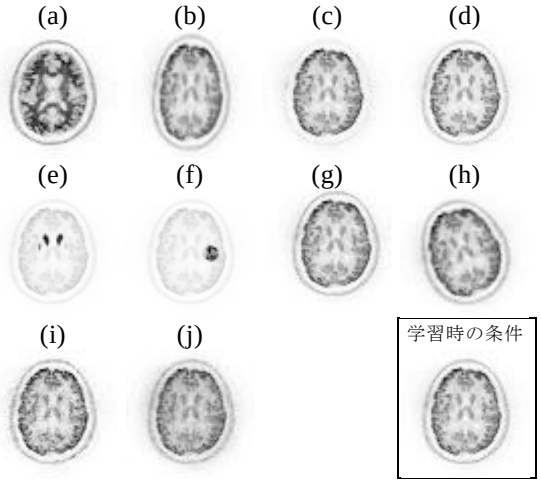


図 4 評価データ（症例 ID=54）の画像例（表 2 の下線の条件）

3）評価方法

PET 画像の定量性を評価した. 具体的には, 推定した組織組成比画像の線形結合により吸収係数画像を求め，それを用いて補正処理をかけた画像を再構成し，真の画像（真の吸収係数画像及び散乱分布を用いて補正した画像）に対する相対誤差を算出した．そして，脳領域（頭部全体ではない）を関心領域として，相対誤差が±5[%]以下の画素割合（以下，正解画素割合と呼ぶ）を算出した．再構成条件は, 未補正 PET 画像の再構成条件と同一である．なお，頭部の形状・位置・姿勢が変動する条件(b), (g), (h)では，元の状態に戻した吸収係数画像を用いて再構成処理を実施した.

4）実験結果

学習データと同じ条件で生成した評価データに追加学習モデル（即発計数 128M）を適用した結果(組織組成比画像, 吸収係数画像)を図 5 に示す．吸収係数画像のスライス毎のピーク信号対雑音比(PSNR)を図 6 に示す．また，図 5 の吸収係数画像を用いて再構成した画像及び相対誤差画像を図 7 に示す．なお，評価用の症例数は 3 例あるが，結果の傾向は同様であったので，本節ではその内 1 例（患者 ID=54）の結果を示す.

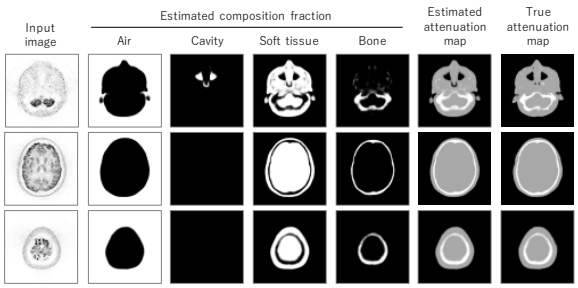


図 5 学習データと同じ条件で生成した評価データに追加学習モデル（即発計数 128M）を適用した結果（組織組成比画像，吸収係数画像）

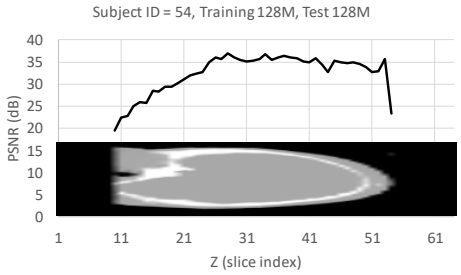


図 6 図 5 の吸収係数画像の PSNR

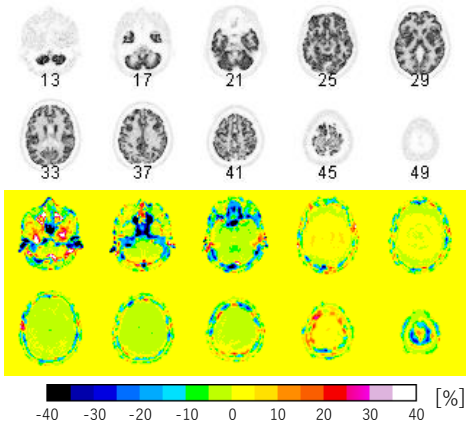


図 7 図 5 の吸収係数画像を用いて再構成した画像及び相対誤差画像

図 4 の画像に追加学習モデル（即発計数 128M）を適用した結果（吸収係数画像，相対誤差画像）を図 8，図 9 に示す．局所的な高集積部位（線条体，腫瘍）のある画像にベースモデル（即発計数 128M）を適用した結果（吸収係数画像）を図 10，図 11 に示す．学習時と推論時のノイズ量が異なる場合の結果（吸収係数画像）を図 12 に示す．異なる散乱割合・偶発割合，異なる皮膚集積量の画像に追加学習モデル（即発計数 128M）を適用した場合の正解画素割合を表 3 に示す．上記以外の変動軸については，検討したパラメータの範囲（表 2）では，正解画素割合が 80～90% 以上となったので，具体的な数値は省略する．つまり，散乱割合・偶発割合，皮膚集積量，統計ノイズ以外の変動においては，組織組成比の推定誤差が，最終的な PET 画像の定量性に大きく影響しなかった．ただし，ほぼ完璧な組織組成比推定ができたわけではない．

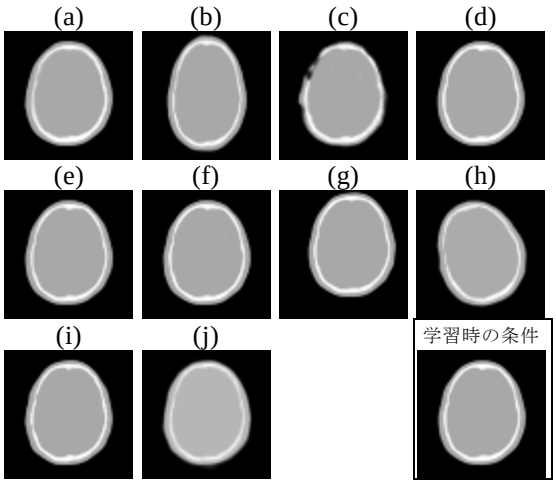


図 8 図 4 の画像に追加学習モデル（即発計数 128M）を適用した結果（吸収係数画像）

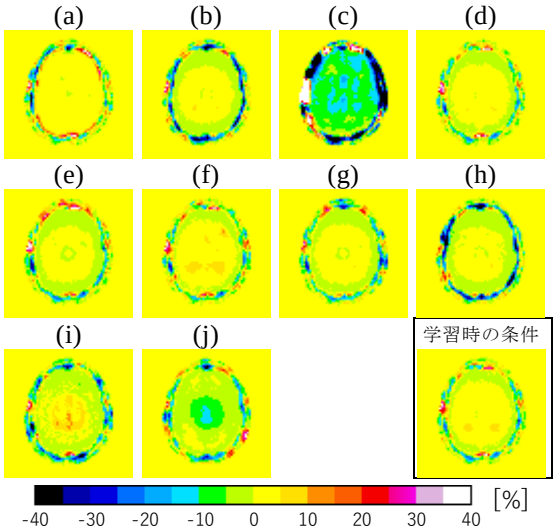


図 9 図 4 の画像に追加学習モデル（即発計数 128M）を適用した結果（相対誤差画像）

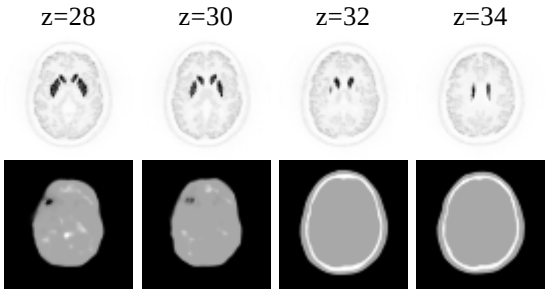


図 10 局所的な高集積部位（線条体）のある画像にベースモデル（即発計数 128M）を適用した結果（吸収係数画像）

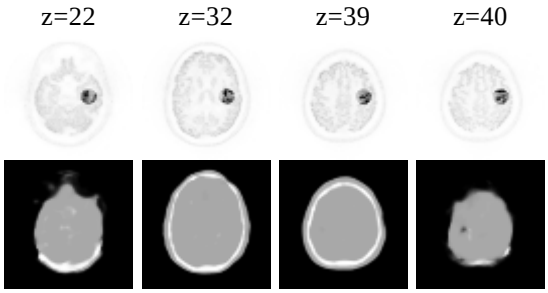


図 11 局所的な高集積部位（腫瘍）のある画像にベースモデル（即発計数 128M）を適用した結果（吸収係数画像）

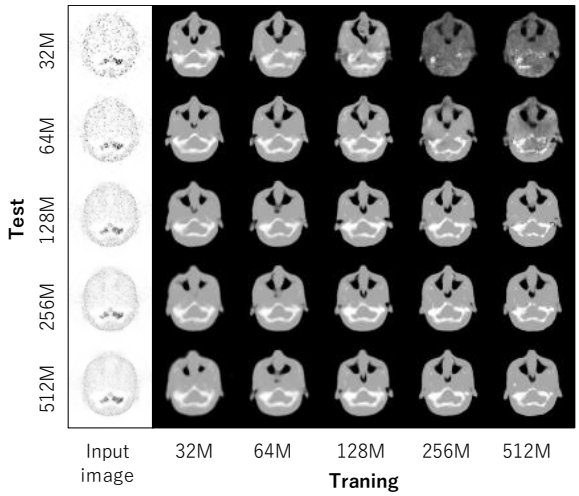


図 12 学習時と推論時のノイズ量が異なる場合の結果（吸収係数画像）

表 3 異なる散乱割合・偶発割合(a)，異なる皮膚集積量(b)の画像に追加学習モデル（即発計数 128M）を適用した場合の正解画素割合[%].

(a) Scatter ratio, random ratio

[%]		Random ratio [%]		
		15	25	35
Scatter ratio [%]	15	92.9	95.7	95.6
	25	97.8	96.4	88.3
	35	92.0	77.2	58.6

(b) Skin uptake

Scale	0.125	0.25	0.50	1.0	2.0	4.0	8.0
[%]	7.5	35.0	93.4	96.4	95.2	89.0	73.8

4. 考察

図 5 は，深層学習によって組織組成比推定が可能であることを示しており，原理検証として前向きな結果が得られたといえる．未補正 PET 画像には，被写体外は画素値がゼロに近い，骨領域は被写体の辺縁付近に存在する，骨領域の画素値は他の領域よりも小さい，脳領域は他の領域よりも画素値が大きい，空洞（鼻腔）は首側で鼻の付近に存在する，といった特徴が存在するので，正確なところは分からないが，ニューラルネットワークがそれらの特徴を獲得し，組織組成比推定に繋げることができたと推察される．ただし，組織組成比及び吸収係数画像の推定精度には，スライス依存性が認められ，骨の構造が複雑な頸部側のスライスほど推定精度が低い（図 6）．

学習時の条件よりも「散乱割合・偶発割合がともに大きい」，「皮膚集積量が極端に少ない」，「統計ノイズが大きい」場合において PET 画像の定量性が顕著に低下した．散乱割合・偶発割合については，それらが小さいと骨の吸収係数が過大推定され，反対に大きい場合は過小推定される（図 8：条件(j)）という傾向が見られた．これはモデルが骨領域の画素値の絶対値を特徴として獲得してしまったためと考えられる．骨の吸収係数の過大推定は，PET 画像への影響は小さかった（表 3(a)）．皮膚集積量については，集積量が低下するほど人間の目でも頭部の輪郭を判定しにくくなるので，推定精度が低下したことはリーズナブ

ルな結果と考えている（図 8:条件(c), 表 3(b)）．統計ノイズについては，ノイズが大きい画像で学習したモデルは，内部に何らかのデノイジング機能を獲得していると考えられるが，ノイズが小さい画像で学習したモデルにはその機能がない又は弱いと考えられるため，ノイズが大きい画像を入力するとノイズ量の違いに対応できず推定精度が低下したと推察される（図 12）．ノイズが大きい画像で学習したモデルは，入力画像のノイズの変動に対して比較的ロバストであったことからも，リーズナブルな推察と考えている．上記以外の変動は，PET 定量性に顕著に影響しなかった．薬剤分布の変化に対してある程度対応できたことは意外であった（図 8:条件(a)）．ただし，現実世界の多様な PET 画像に対してもロバストに動作するかは全くの未知であり，今回のシミュレーションデータでも大なり小なり影響があったことを踏まえると楽観はできない．放射性薬剤は，灰白質集積型・白質集積型・均一集積型に大別できるので，それぞれに特化したモデルを準備しておくことが無難と考えられる．

最後に，追加学習の効果について述べる．特異的な高集積部位がある場合，当該部位以外の画素値は相対的に低下するため，モデルが画素値の絶対値を特徴として獲得した場合，そのような画像に対しては汎化性能が低下する（最悪，まったくおかしな推定結果を出力する）ことが予想され，図 10，図 11 はそれが顕在化した結果と考えている．入力画像の画素値の絶対値に対する依存性を弱めることをねらいに，前述したように，画像全体を定数倍した画像を含めて追加学習を行い，線条体や腫瘍といった高集積部位を含む画像において推定精度を大幅に改善できた（図 8(e)(f)）．線条体の集積量を変動させた画像を作成することは比較的容易であるが，腫瘍の位置・形状は千差万別であるため，網羅的な学習データの作成は現実的でない．今回の実験では，円形の腫瘍しか評価できていないが，腫瘍が写った画像を学習データに含めずに，それらに対応できたことは，実用化に向けた重要な一歩と考えている．

5. 結論

本研究では，深層学習に基づく吸収係数画像推定の新しいアプローチとして，組織組成比推定を介した間接的手法を提案し，原理検証実験を実施し，前向きな結果を得た．また，学習データに含まれない変動をもった画像を用いて，組織組成比推定モデルの汎化性能を調査した．その結果，散乱割合・偶発割合，皮膚集積量，ノイズ量といった変動に弱いことが判明した．

今後は，シミュレーションデータのみで学習した現状のモデルを実データに適用して課題を抽出し，本研究で明らかになった課題と併せて，モデルの改良を進める予定である．

最後に，PET は CT・MR と比較してオープンデータが圧倒的に少ないため，PET 分野で AI 技術を実用化する上では，高速・高精度なシミュレーション技術が必要不可欠と考えている．本研究に関して言えば，未補正 PET 画像のオープンデータベースが構築される可能性は低く，加えて頭部 MR 画像の高精度なセグメンテーション技術をもたない限り頭部ラベル画像を作成できず，教師データとなる組織組成比画像を作成できないので，頭部ラベル画像のオープンデータ[9]及び PET 計測シミュレーション技術をフル活用して，完全シミュレーションベースで実データに対応した組織組成比推定モデルの開発を目指している．

利益相反の有無

なし

文　献

［1］ Defrise M, Rezaei A, Nuyts J: Time-of-flight PET data determine the attenuation sinogram up to a constant. Phys Med Biol **57**: 885, 2012.

［2］ Nuyts J, Rezaei A, Defrise M: The validation problem of joint emission/

transmission reconstruction from TOF-PET projections. IEEE Trans Rad Plasma Med Sci **2**: 273-278, 2018.

[3] Beyer T, Weigert M, Quick H.H, et al.: MR-based attenuation correction for torso-PET/MR imaging: pitfalls in mapping MR to CT data. EJNMMI **35**: 1142-1146, 2008.

[4] Liu F, Jang H, Kijowski R, et al.: A deep learning approach for 18 F-FDG PET attenuation correction. EJNMMI physics **5**: 1-15, 2018.

[5] Hashimoto F, Ito M, Ote K, et al.: Emis2Trans: attenuation correction for brain PET with many types of PET ligands using convolutional neural networks. 2019 IEEE Nucl Sci Symp & Med Imag Conf.

[6] Dong X, Wang T, Lei Y, et al: Synthetic CT generation from non-attenuation corrected PET images for whole-body PET imaging. Phys Med Biol **64**: 215016, 2019

[7] Aksoy Y, Oh T.H, Paris S, et al.: Semantic soft segmentation. ACM Trans Graphics **37**: 1-13, 2018.

[8] Ronneberger O, Fischer P, Brox T: U-net: Convolutional networks for biomedical image segmentation. International Conference on MICCAI, 234-241, 2015.

[9] Aubert-Broche B, Griffin M, Pike G, et al. Twenty new digital brain phantoms for creation of validation image data bases. IEEE Trans Med Imag **25**: 1410-1416, 2006.

[10] Hudson H.M, Larkin R.S: Accelerated image reconstruction using ordered subsets of projection data. IEEE Trans Med Imag **13**: 601-609, 1994.

PET attenuation map generation via deep learning-based tissue composition estimation

Tetsuya KOBAYASHI[*1], Yui SHIGEKI[*1], Hisanori MORITA[*1]

*1 Technology Research Laboratory, Shimadzu Corporation

Accurate scatter and attenuation corrections are essential for obtaining quantitative PET reconstruction, and accurate patient attenuation map is crucial for achieving that. In integrated PET-CT and PET-MR scanners, attenuation map can be generated from CT/MR image, however, in stand-alone PET scanner with no transmission source, attenuation map must be generated from measured PET data itself. Here, attenuation map of human head can be modeled by the linear combination of four composition maps of soft tissue, bone, cavity and outside-subject (i.e., air), then attenuation map can be generated indirectly if the composition maps are estimated from the non-scatter and non-attenuation corrected PET image. In this study, we propose the conceptual idea of estimating composition maps of a subject from non-corrected PET image and a deep learning model for its realization, and presents the results of the proof-of-concept study for evaluating the accuracy of attenuation map and PET image using simulated PET data of human head.

Key words: PET, Attenuation map, Attenuation correction, Deep learning, Semantic soft segmentation

WGI 試作機によるフルリングコンプトンイメージングの実証

田島　英朗[*1]　吉田　英治[*1]　脇坂　秀克[*1]　高橋　美和子[*1]

永津　弘太郎[*1]　辻　厚至[*1]　鎌田　圭[*2,3]　吉川　彰[*2,3]

Katia Parodi[*4]　山谷　泰賀[*1]

要旨

我々は，PET リングに散乱検出器リングを挿入し，PET とコンプトンイメージングを組みわせることで，様々なガンマ線をイメージングに用いること可能とする WGI (Whole Gamma Imaging) の開発を進めており，これまでに小動物サイズの WGI 試作機を開発した．WGI 試作機は，我々の知る限り世界初のフルリングコンプトンイメージング装置であり，本研究では，その有効性を実証することを目的とする．PET とコンプトン画像を同一装置で比較するため，陽電子と 909keV のガンマ線を同一分布から放出する ^{89}Zr を測定対象に選択した．また，PET 及びコンプトン画像再構成法として，それぞれリストモード OSEM に検出器応答モデルと感度補正を組み込んだ．ファントム及びマウスを測定し，PET 画像と 909 keV のコンプトン画像をそれぞれ再構成した結果，PET に迫る高精細なコンプトン画像を得ることができた．

キーワード：WGI，コンプトンイメージング，^{89}Zr，核医学，PET

1．はじめに

コンプトンイメージングまたはコンプトンカメラは，宇宙線計測の分野で発展，実用化され [1-4]，コリメータ不要の核医学装置としても医学生物学応用を目指した研究開発が行われてきた [5-8]．これまでに，小動物やヒトのイメージングの試みがなされているが [9-12]，分解能や定量性など画質の面で，PET（Positron Emission Tomography）はもちろんのこと，SPECT（Single Photon Emission Computed Tomography）にも遠く及ばないのが現状であった．一方，我々は PET リングに散乱検出器リングを挿入し，PET とコンプトンイメージングを組み合わせることで，様々なガンマ線をイメージングに用いることを可能とする WGI (Whole Gamma Imaging) の開発を進めており，これまでに小動物サイズの試作機を開発した [13]．PET リングはコンプトンイメージングにおいて，吸収検出器の役割を担う．我々の知る限り，WGI 試作機は世界初のフルリング型のコンプトンイメージング装置である．本研究では，

*1 量子科学技術研究開発機構
放射線医学総合研究所
〔〒263-8555 千葉市稲毛区穴川 4-9-1〕
e-mail: tashima.hideaki@qst.go.jp
*2 東北大学
*3 株式会社 C&A
*4 Ludwig-Maximilians-Universität München

WGI コンプトンイメージングを，すでに確立したモダリティである PET と比較するため，ジルコニウム-89 (^{89}Zr)に着目した．^{89}Zr は陽電子と 909keV のガンマ線を秒オーダーの時間差で放出するため，時間識別とエネルギー識別により PET イベントとコンプトンイベントをそれぞれ抽出することができる．そして，それらは同一の分布から発生しているため，別々に再構成した画像を直接比較することが可能である．

2．方法

1）WGI 試作機

図 1 に WGI 試作機の構成を示す．WGI 試作機の吸収検出器リング（Absorber ring）は Whole-body DOI PET [14]と同じであり，4 層 DOI (Depth-of-interaction) 検出器 160 個を 1 リング 40 個，4 リング配置した構成となっている．4 層 DOI 検出器は，2.8×2.8×7.5mm^3 の GSOZ シンチレータを 16×16×4 に配列し，64ch の光電子増倍管に光学接続した構造となっている．吸収検出器リングに挿入した散乱検出器リング（Scatterer ring）の検出器は，光出力の高いシンチレータである GAGG (0.9×0.9×6.0mm^3)の 24×24 配列を，8×8 配列の MPPC（Multi-pixel photon counter）に光学接続した構造となっており，1 リング 10 個，2 リングに配置した構成とした．散乱検出器と吸収検出器のエネルギー分解能はそれぞれ 17%と 14%であった[13]．データ収集回路の積分時間と GAGG の発光減衰時間の相性が良くなかったため，期待していたよりも散乱検出器のエネルギー分解能が低かった．

WGI 試作機のデータ収集ではシングルデータをすべて保存し，同時計数処理やエネルギーウィンドウ処理はオフラインのソフトウェアで行った．PET イベントは吸収検出器同士のイベントの他，散乱検出器同士や，散乱検出器－吸収検出器の組み合わせもすべて利用した．エネルギーウィンドウは散乱検出器で 350-650 keV，吸収検出器で 400-600 keV とした．コンプトンイベントの抽出では，散乱検出器－吸収検出器間の同時計数イベントのみ抽出し，エネルギーウィンドウを散乱検出器 50-350 keV，吸収検出器 550-850 keV とした．

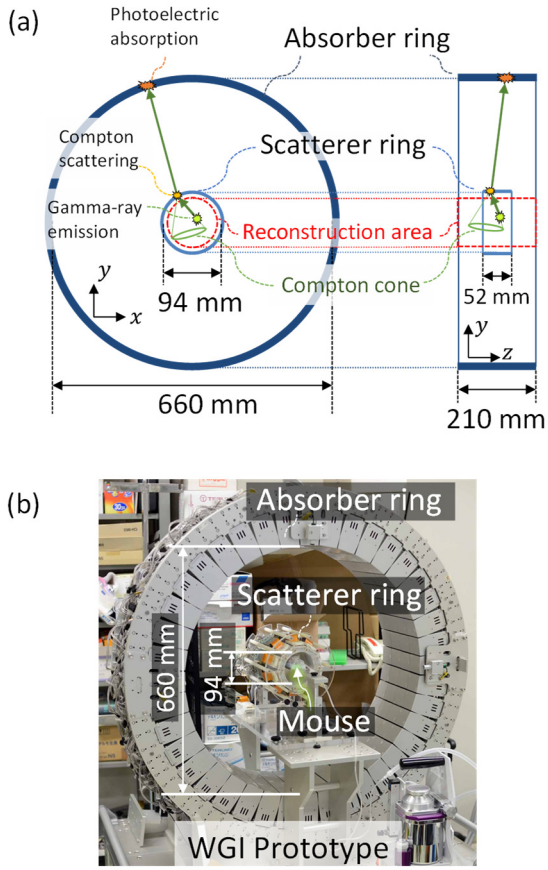


図 1 WGI 試作機の散乱・吸収検出器リング構成(a)と測定時の装置外観(b).

画像再構成法としては，定量性を確保するために必要な感度補正と偶発同時計数補正，高分解能化のために必要な検出器応答関数モデルを組み込んだリストモード OSEM（Ordered subset expectation maximization）法を開発した．また，感度補正用として，^{89}Zr 水溶液を中空の円筒状に分布させることが可能な中空円筒ファントムを作成し，PET とコンプトンイメージング共通で用いた．偶発同時計数の補正は，遅延同時計数の時間窓をソフトウェア同時計数法に組み込むことで行った．検出器応答関数モデルとしては，エネルギーの揺らぎの度合いが角度の揺らぎの度合いに線型的に変換できると仮定して，コンプトンコーンの表面に角度依存のガウス関数によるぼけを与えた．ボクセル

サイズは 1×1×1mm^3 とし，サブセット数 16，反復回数は収束性の違いから PET の場合には 10，コンプトンイメージングの場合には 50 とした．

2）^{89}Zr ファントム・動物実験

WGI 試作機を用いて，円筒ファントムとマウスの測定を行った．測定時間はそれぞれ 1 時間とした．円筒ファントムは内径 38mm，長さ 60mm で，10.2MBq の ^{89}Zr 溶液を封入した．マウス実験では，測定の 23 時間前に，9.8MBq のシュウ酸ジルコニウム-89(^{89}Zr)を投与した．円筒ファントムの再構成画像の評価として，均一性を ROI（Region of interest）値の変動係数（COV: Coefficient of variation）として求めた．その際，体軸方向の位置として 5mm ステップ 9 か所，そして各体軸位置の横断面内に 7 か所ずつ，直径 10mm，長さ 5mm の円筒型の ROI を合計 63 個設定した．マウス画像の評価は視覚的に行った．

3．結果

図 2 に円筒ファントム，図 3 にマウスの再構成像を示す．PET の場合，PET リングが体軸方向に十分長く，測定対象全体がリング内に入るため，全体が均一な品質で再構成された．一方，コンプトンイメージングの場合，散乱検出器の外側の領域も画像化ができている．特に散乱検出器リングの内側の領域で高い画質が得られているものの，その外側の領域では画質の劣化と円筒ファントムのノズル部分が縮むなどの歪みが発生していることが観察された．円筒ファントムの均一性評価のための ROI はすべて散乱検出器リング内の領域に位置していた．定量評価の結果，円筒ファントムの COV は PET 画像で 3.7%，コンプトン画像で 4.8%であり，PET に近い均一性が得られた．ノイズが多かったのは，コンプトンイメージングの感度が PET よりも大幅に低く， PET イベント 190 M カウント，コンプトンイベント 30.7 M カウントとであったためと考えられる．また，反復回数についても経験的に定めており，ノイズとコントラスト・分解能のトレードオフより最適化する必要がある．

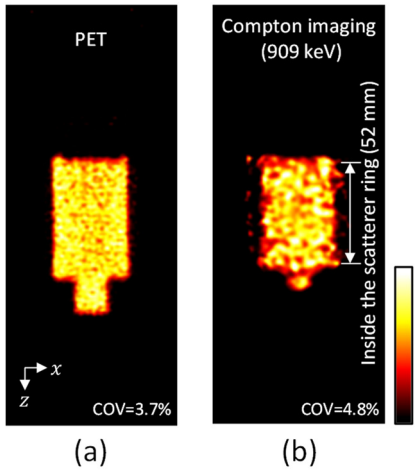


図 2 ^{89}Zr 円筒ファントムの PET(a)とコンプトンイメージング(b)の再構成画像（中心スライス）．

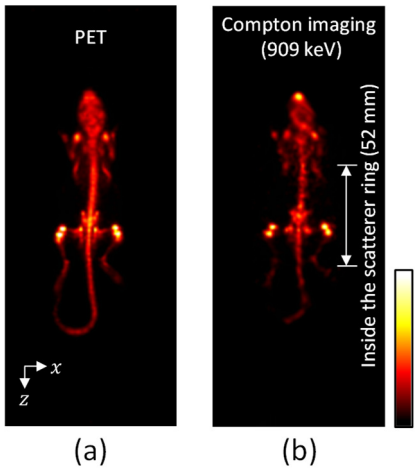


図 3 シュウ酸ジルコニウム-89 (^{89}Zr) 投与マウスの PET(a)とコンプトンイメージング(b)の再構成画像（MIP：Maximum Intensity Projection）．

マウス実験の結果，PET とコンプトンイメージング両方の画像はよく一致しており，骨の構造が鮮明に描出された．しかしながら，マウスの頭蓋骨に注目すると，PET 画像では眼窩と下顎骨を特定可能であったが，コンプトン画像では特定することが困難であった．これらは散乱検出器リングの外側に位置していた．一方，散乱検出器リングの内側に位置していた膝の部分の左右 2 つずつの高い集積は両画像とも鮮明

に分離できていることが観察できた．コンプトンイメージングでは，散乱検出器の外側の領域の画質はやや劣化するものの，散乱検出器リングの内側では PET に迫る高い画質を得ることができた．

4．考察

WGI 試作機の散乱検出器は，これまでにコンプトンイメージングに用いられてきたものと比べてあまり良くないエネルギー分解能であるにも関わらず，今回得られた PET に迫る高精細な ^{89}Zr コンプトン画像は，これまでのコンプトンイメージングの歴史上革新的なものであった．この理由としては，まず用いたガンマ線のエネルギーが 909keV と，これまでに対象にされてきた核種と比べ高いものであったことがあげられる．次に，画像再構成法として，感度補正，偶発同時計数補正，及び検出器応答関数モデルを組み込んだことで，検出器の素子ごとのばらつきなどを高精度に補正できたことが考えられる．特に，感度補正についてはこれまでの研究開発では幾何学的に求めた係数や，モンテカルロ法によって求めた係数を用いているものがほとんどで，PET 手法を転用し，実際の測定系で校正用のファントムを測定して行ったのはこれまでにない試みであったと考えられる．そして，フルリング型の装置構成にしたことで，これまでのシステムと比べて十分な角度範囲のコンプトンイベントが取得でき，投影データの実質的な完全性が特に散乱検出器の内側で達成できたためであると考えられる．また，従来研究で行われている単一の検出器を回転させて様々な角度方向のデータを取得する回転型のシステムと比較した場合，検出器の数に比例した大幅な感度向上があると考えられ，統計的ノイズを低減できたことが鮮明な画像を得られた要因の一つと考えられる．今後，様々な再構成条件や対象エネルギーの検討を行い，WGI コンプトンイメージングのより詳細な特性を明らかにしていく予定である．

5．まとめ

フルリング型コンプトンイメージングシステムである WGI 試作機を用いて ^{89}Zr ファントムとマウスの測定を行い，PET と比較した結果，各種補正を組み込んだコンプトン画像再構成法により，PET に近い均一性と，PET に迫るかつてない高精細なマウスのコンプトン画像を得ることができた．

謝辞

本研究は JSPS 科研費（16H02641, 18K19949, 19KK0280, 20K12683, 20H00671），及び QST 戦略的理事長ファンド（国際リサーチイニシアティブ，創生的研究，萌芽的研究）の助成を受けたものです．

利益相反の有無

なし

文　献

［1］ Pinkau K: Die Messung solarer und atmosphärischer Neutronen. Zeitschrift für Naturforsch A **21**: 2100–2101, 1966

［2］ White R S: An experiment to measure neutrons from the sun. Bulletin of American Physics Society **13**: 714, 1968

［3］ Schönfelder V, Hirner A and Schneider K: A telescope for soft gamma ray astronomy. Nucl Instr Meth **107**: 385–394, 1973

［4］ Schönfelder V, Diehl R, Lichti G G, et al.: The imaging compton telescope comptel on the gamma ray observatory. IEEE Trans Nucl Sci **31**: 766–70, 1984

［5］ Todd R W, Nightingale J M and Everett D B: A proposed γ camera. Nature **251**: 132–134, 1974

［6］ Singh M and Brechner R R: Experimental test-object study of electronically collimated SPECT. J Nucl Med **31**: 178–86, 1990

［7］ LeBlanc J W, Clinthorne N H, Hua C H, et al.: C-SPRINT: A prototype compton

camera system for low energy gamma ray imaging. IEEE Trans Nucl Sci **45**: 943–9, 1998

［8］ Rogers W L, Clinthorne N H and Bolozdynya A: Compton Cameras for Nuclear Medicine Imaging. In Wernick M and Aarsvold J N eds.: Emission Tomography: The Fundamentals of PET and SPECT, Academic Press, 2004, pp383–419

［9］ Kabuki S, Kimura H, Amano H, et al.: Electron-tracking Compton gamma-ray camera for small animal and phantom imaging. Nucl Instr Meth Phys Res Sect A **623**: 606–7, 2010

［10］ Takeda S, Odaka H, Ishikawa S, et al.: Demonstration of in-vivo multi-probe tracker based on a Si/CdTe semiconductor Compton camera. IEEE Trans Nucl Sci **59**: 70–6, 2012

［11］ Kishimoto A, Kataoka J, Taya T, et al.: 2017 First demonstration of multi-color 3-D in vivo imaging using ultra-compact Compton camera. Sci Rep **7**: 2110, 2017

［12］ Nakano T, Sakai M, Torikai K, et al.: Imaging of ^{99m}Tc-DMSA and ^{18}F-FDG in humans using a Si/CdTe Compton camera. Phys Med Biol **65**: 05LT01, 2020

［13］ Yoshida E, Tashima H, Nagatsu K, et al.: Whole gamma imaging: a new concept of PET combined with Compton imaging. Phys Med Biol **65**: 125013, 2020

［14］ Akamatsu G, Tashima H, Iwao Y, et al.: Performance evaluation of a whole-body prototype PET scanner with four-layer DOI detectors. Phys Med Biol **64**: 095014, 2019

Demonstration of full-ring Compton imaging with a WGI prototype

Hideaki TASHIMA[*1], Eiji Yoshida[*1], Hidekatsu WAKIZAKA[*1], Miwako TAKAHASHI[*1], Kotaro NAGATSU[*1], Atsushi TSUJI[*1], Kei KAMADA[*2,3], Akira YOSHIKAWA[*2,3], Katia PARODI[*4], Taiga YAMAYA[*1]

*1 National Institute of Radiological Sciences, QST

*2 Tohoku University

*3 C&A Corporation

*4 Ludwig-Maximilians-Universität München

We have developed a prototype whole gamma imaging (WGI) system for small animal experiments. WGI is a combination of PET and Compton imaging by inserting a scatterer detector ring into a PET ring for utilizing various gamma rays for imaging. To the best of our knowledge, the WGI prototype is the world's first full-ring Compton imaging system. In this study, we aimed at demonstrating the effectiveness of the full-ring system. We selected ^{89}Zr, which emits both positrons and 909-keV gamma rays from the same distribution, for the imaging target to compare PET and Compton images on the same system. We developed PET and Compton image reconstruction methods each incorporating detector response function modeling and normalization into the list-mode OSEM algorithm. We conducted phantom and mouse measurements, and PET and 909-keV Compton images were reconstructed. As a result, we obtained high-definition Compton images approaching PET.

Key words: WGI, Compton imaging, ^{89}Zr, Nuclear medicine, PET

Automatic stratification of prostate cancer patients into low- and high-grade groups using machine learning models with MR image features

Akimasa URAKAMI[*1], Hidetaka ARIMURA[*2], Fumio KINOSHITA[*3], Kenta NINOMIYA[*1], Sumiko WATANABE[*2], Kenjiro IMADA[*4], Yoshinao ODA[*3]

Abstract

Non-invasive prediction of prostate cancer grade group (GG) with magnetic resonance (MR) imaging without prostate biopsy would be greatly beneficial for patients. This study aimed to explore the potential of automatic stratification of prostate cancer patients into low- and high-grade groups using machine learning models with MR image features. MR images of 40 cancer regions corresponding to histopathological images of prostate cancer were selected and they were divided into two groups, low-grade tumors (≤GG 2) and high-grade tumors (GG 3≤), based on their grade groups. 4,320 MR image features were derived from original images (16 types of images in 5 sequences) and two-dimensional wavelet decomposition-based images. Significant features were chosen by using a least absolute shrinkage and selection operator (LASSO) algorithm, and they were used to construct 7 supervised machine learning models. Parameters of each machine learning model were optimized based on five-fold cross validation test so that areas under receiver operating characteristic curves (AUCs) could be maximized. Among 7 supervised machine learning models with four MR image features, support vector machine (SVM) achieved the best prediction (AUC = 0.965) for the stratification of prostate cancer patients. This study showed the potential of automatic stratification of prostate cancer patients into low- and high-grade groups using the SVM with MR image features.

Keywords：prostate cancer, grade group, MR image features, LASSO, SVM

1. Purpose

Prostate cancer is the most common cancer in men, with almost 1.3 million new cases per year worldwide [1].One of the most important indicators of a definitive diagnosis of prostate cancer is grade groups (GGs) 1-5, and accurate stratification is necessary, especially for GG 2 and GG 3, because of the significant differences in distant metastasis and survival rates [2,3]. GGs are determined by invasive prostate biopsy, which cause lower urinary tract symptoms, erectile dysfunction and infection [4,5]. In contrast, magnetic resonance (MR) imaging is widely used in clinical practice as a non-invasive examination for identification, staging and biopsy targeting of patients with prostate cancer, and it has the potential to predict GGs [6,7]. Hence, prediction of lower and higher GGs using MR images without prostate biopsy would be greatly beneficial to patients for better quality of life.

This study aimed to explore the potential of automatic stratification of prostate cancer patients into low- and high-grade groups using machine learning models with MR image features.

2. Methods

Figure 1 shows an overall scheme of this study. we used isovoxel MR images (pixel size: 0.3906 mm, matrix size: 512×512) of 40 cancer regions corresponding to histopathological images of prostate cancer and they were divided into two groups, low-grade tumors (≤ GG 2) and high-grade tumors (GG 3 ≤), based on their grade groups.

Figure 2 shows an example of calculation of MR image features for each sequence. 4,320 MR image features were derived from original images for 5 sequences and their corresponding wavelet decomposition-based images based on histograms, gray-level co-occurrence matrices (GLCMs), gray-level run-length matrices (GLRLMs), gray-level size-zone matrices (GLSZMs) and neighbor gray-tone difference matrices (NGTDMs). The original images were acquired by using 5 sequences of T1 weighted image (WI), T2WI, Diffusion weighted image (DWI), apparent diffusion coefficient (ADC) map and dynamic contrast enhanced (DEC) T1WI.

Figure 3 shows the MR image features selected by least absolute shrinkage and selection operator (LASSO). Four features with the highest absolute values of the coefficient were considered significant features that were used to construct 7 supervised machine learning models [support vector machine (SVM) based on Gaussian, linear and polynomial kernel, logistic regression (LR), random forest (RF), k-nearest neighbor (KNN) and Adaboost]. Areas under the receiver operating characteristic curve (AUC) were used to optimize parameters and to evaluate the performances of 7 proposed models based on five-fold cross validation test.

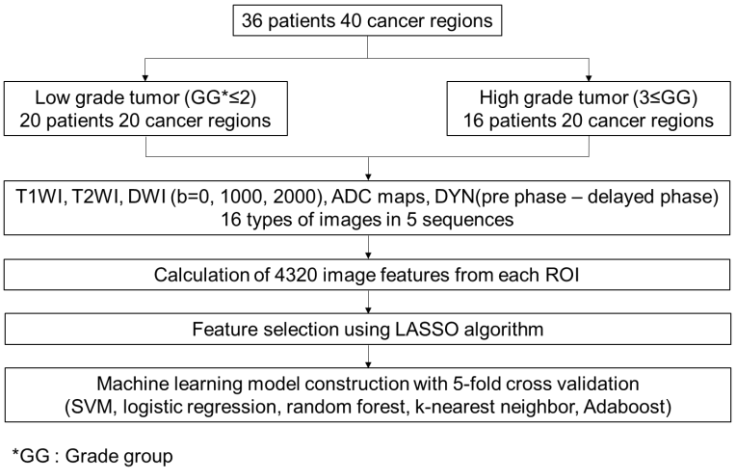


Fig. 1 An overall scheme used in this study

*1 Department of Health Sciences, Graduate School of Medical Sciences, Kyushu University〔3-1-1 Maidashi, Higashi-ku, Fukuoka, 812-8582, Japan〕 e-mail: akimasa19980107@gmail.com

*2 Department of Health Sciences, Faculty of Medical Sciences, Kyushu University

*3 Department of Anatomic Pathology, Graduate School of Medical Sciences, Kyushu University

*4 Department of Urology, Prostate, Kidney, Adrenal Surgery, Kyushu University Hospital

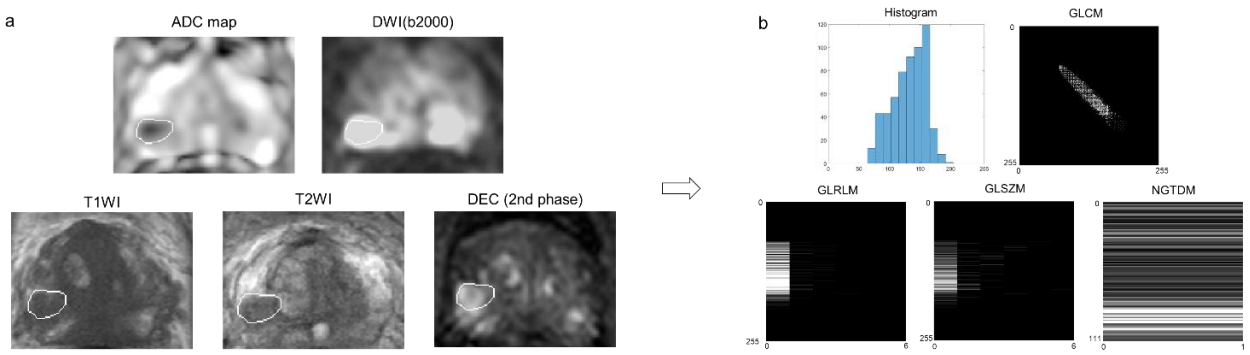


Fig. 2 An example of calculation of MR image features for each sequence. a: Segmentation of tumor region of interest on each sequence image (white). b: An example of histogram, GLCM, GLRLM, GLSZM and NGTDM derived from ADC map.

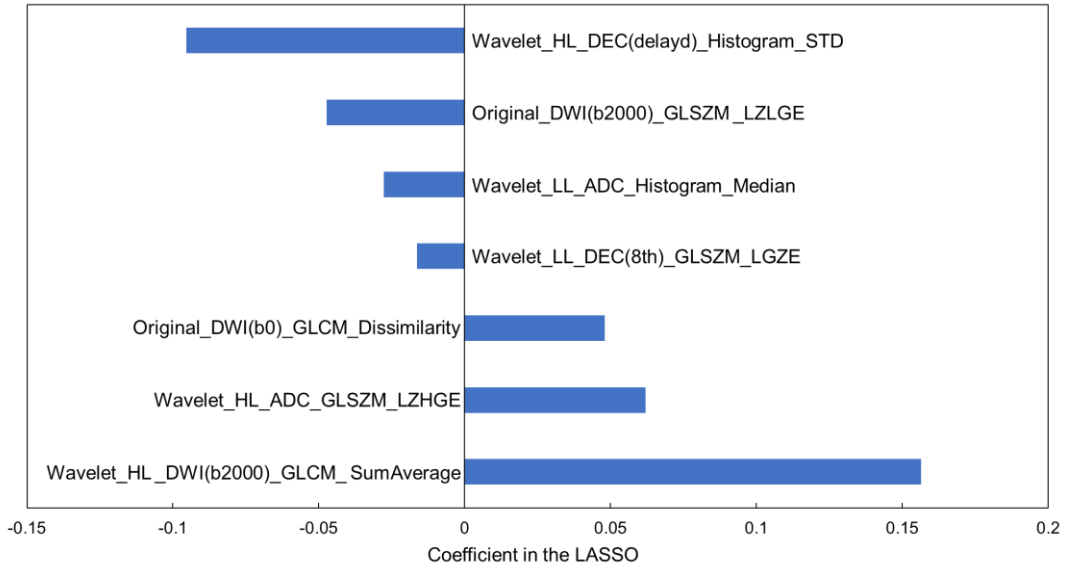


Fig. 3 MR image features selected by LASSO. H: high-pass filter, L: low-pass filter

3. Results

Table 1 shows predictive performances of 7 machine learning models. Among 7 supervised machine learning models with four significant features, the SVM model achieved the best prediction (AUC = 0.965) for the stratification of prostate cancer patients.

In this study, there were one false positive case and three false negative cases. Different wavelet filters and shape features may improve the misclassification [8,9]. We will work on new feature extraction in the future work.

Table 1 Predictive performance of 7 machine learning models

Machine learning model	AUC	Accuracy	Sensitivity	Specificity
Gaussian SVM	0.965	0.899	0.843	0.950
Linear SVM	0.965	0.899	0.843	0.950
Polynomial SVM	0.965	0.874	0.833	0.900
LR	0.953	0.932	0.960	0.900
RF	0.963	0.925	0.900	0.950
AdaBoost	0.928	0.853	0.860	0.850
KNN	0.910	0.855	0.763	0.950

4. Conclusion

This study showed the potential of automatic stratification of prostate cancer patients into low- and high-grade groups using the SVM with MR image features.

Competing interests

The authors do not have any conflict of interests to disclose.

Acknowledgement

The authors would like to express our sincere thanks to all members in the Arimura laboratory for their valuable comments for this study.

References

[1] Freddie B, Jacques F, Isabelle S, et al.: Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA CANCER J CLIN **68**: 394-424, 2018

[2] Epstein J, Lars E, Mahul B, et al.: The 2014 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma: definition of grading patterns and proposal for a new grading system. Am J Surg Pathol **40**: 244-252, 2016.

[3] Theresa YC, Alan WP, Patrick CW, et al.: Prognostic significance of Gleason score 3+4 versus Gleason score 4+3 tumor at radical prostatectomy. ELSEVIER SCIENCE INC. **56**, 823-827, 2000

[4] Franck B, Sandra M, Philippe B, et al.: Prosbiotate: A Multicenter, Prospective Analysis of Infections Complications after Prostate Biopsy. THE JORNAL OF UROLOGY. **193**: 145-150, 2015.

[5] Marco B, Hashim A, Robert N, et al.: Complications After Systematic, Random, and Image-guided Prostate Biopsy. Eur Urol. **71**: 353-365, 2017

[6] Virendra K, Girdhar SB, Rajeev K et al.: Multiparametric (mp) MRI of prostate cancer. Progress in Nuclear Magnetic Resonance Spectroscopy. **105**: 23-40, 2018

[7] Radtke JP, Schwab C, Wolf MB, et al.: Multiparametric magnetic resonance imaging (MRI) and MRI-transrectal ultrasound fusion biopsy for index tumor detection: correlation with radical prostatectomy specimen. Eur Urol **70**: 846-853, 2016

[8] Jussi T, Ileana MP, Parisa M, et al.: Radiomics and machine learning of multisequence multiparametric prostate MRI: Towards

improved non-invasive prostate cancer characterization. PloS ONE **14**: 1-23, 2019

[9] Liu B, Cheng J, Guo DJ, et al.: Prediction of prostate cancer aggressiveness with a combination of radiomics and machine learning-based analysis of dynamic contrast-enhanced MRI. Clinical Radiology **74**: 896.e1-896.e8, 2019

複数周波数同時加振 MR エラストグラフィによる撮像時間短縮の検討

久慈　陸[*1]　菅　幹生[*1,2]　岸本　理和[*2]　野澤　真吾[*2]　小畠　隆行[*2]

要旨

MRE は，生体組織の粘弾性率分布を測定する手法である．複数周波数 MRE では組織粘弾性の周波数特性を取得できるが，撮像時間を要する問題がある．本研究では，40, 60, 90 Hz の 3 周波数で同時加振する場合において，取得位相数による撮像時間短縮効果と弾性率推定精度の関係を計算機シミュレーションによって明らかにすることを目的とした．単一周波数ごとに 4 位相収集する従来法で 3 周波数を取得する場合と比較して，同時加振で位相数が 7, 8, 9 での撮像時間短縮効果はそれぞれ 44%, 37%, 30%となる．振幅対ノイズ比がそれぞれ 25, 25, 7 の弾性波画像を用いた検討では，同時加振で位相数が 7，8，9 での弾性率分布の平均値はいずれも真値と一致し，最大の変動係数はそれぞれ 7%, 5%, 0%以下となり，位相数 9 では加振周波数によらず単一周波数加振 MRE の変動係数よりも小さくなった．以上の結果より，3 周波数同時加振 MRE の有効性が示唆された．

キーワード：MRI，MR エラストグラフィ，粘弾性，撮像時間短縮

1. はじめに

MRI を用いて粘弾性率を非侵襲的かつ定量的に測定する手法として magnetic resonance elastography (MRE)がある[1]．MRE では，MRI 装置と同期して測定対象に外部から特定周波数の正弦波振動を加えて測定対象内部に伝播する弾性波を MRE パルスシーケンスによって画像化し，粘弾性率推定手法を適用することで粘弾性率分布を求めることができる．粘弾性物体には測定に用いる外部振動の周波数により粘弾性率が変化する周波数特性がある．複数周波数加振での MRE 測定による生体組織の周波数特性評価は疾患の鑑別診断への利用が期待されている[2]．複数周波数加振 MRE において単一周波数の加振による MRE を繰り返すことは，取得する周波数の数に応じて撮像時間が長くなり被検者の負担となるため，撮像時間の短縮が望まれる．撮像時間の短縮のために，複数の周波数成分を含む合成波で加振する，複数周波数同時加振 MRE を提案し, 2 周波数の同時加振 MRE についての有効性を確認している[3]．周波数特性を示す粘弾性モデルを推定するには最低でも 3 周波数が必要となる．本研究では，3 周波数で同時加振する場合において，取得位相数による撮像時間短縮効果と弾性率推定精度の関係を計算機シミュレーションによって明らかにすることを目的とした．

2. 方法

2.1 弾性率算出法

弾性波画像に対して波動方程式に基づいた弾性率推定式を適用することで，弾性率分布画像を

*1 千葉大学
〔〒263-8522 千葉市稲毛区弥生町 1-33〕
e-mail: afpa7035@chiba-u.jp
*2 量子科学技術研究開発機構　放射線医学総合研究所　〒263-8555 千葉県千葉市稲毛区穴川 4-9-1

算出できる．本研究では弾性率推定式として式(1)に示す 3 次元積分型公式[4]を使用した．

$$G' = -\rho\omega^2 \frac{\int_D \phi(\Delta v)dxdydz \int_D \phi v dxdydz + \int_D \psi(\Delta v)dxdydz \int_D \psi v dxdydz}{\left\{\int_D \phi(\Delta v)dxdydz\right\}^2 + \left\{\int_D \psi(\Delta v)dxdydz\right\}^2} \quad (1)$$

ρ：密度，ω：加振角周波数，
ϕ，ψ：互いに位相が 90° 異なる弾性波の変位，
$v(r)$：式(2)を満たすテスト関数，D：領域．

$$v|_{\partial D} = \frac{\partial v}{\partial n}\bigg|_{\partial D} = 0 \quad (2)$$

式(3)は上記条件を満たす関数の一例である．

$$v(r) = \begin{cases} 1+\cos\pi\dfrac{r}{R} & r < R \\ 0 & r \geq R \end{cases} \quad (3)$$

ただし，$D=\{x^2+y^2+z^2<R^2\}$，$r=|x-x_0|$．

2.2 周波数分離方法

弾性率算出法に適用するために，複数周波数成分を含む合成波を単一周波数成分の弾性波へ分離する必要がある．初期位相が異なる弾性波画像の画素値と合成波を表す式(4)を利用して最小二乗法により単一周波数成分の弾性波画像に分離した．

$$A_1 sin(2\pi f_1 t + B_1) + A_2 sin(2\pi f_2 t + B_2) + A_3 sin(2\pi f_3 t + B_3) + E \quad (4)$$

f_1~f_3：周波数，A_1 ~A_3：f_1~f_3の振幅，
B_1~B_3：f_1~f_3の初期位相，E：誤差，t：時間

2.3 撮像時間

MRE の撮像時間は式(5)で表される．

$$TR \times 3 \times \{1 + \textit{取得初期位相数} \times 2\} \quad (5)$$

式(4)に含まれる未知数が 7 個であるため，最低 7 位相の弾性波画像が必要となる．また，単一周波数 MRE より撮像時間が短くなるのは 9 位相であるため，本研究では取得する初期位相数を 7 から 9 とした．単一周波数ごとに 4 位相収集する従来法で 3 周波数を取得する場合と比較して，同時加振で位相数が 7, 8, 9 での撮像時間短縮効果はそれぞれ 44%, 37%, 30%となる．

2.4 計算機シミュレーション

計算機シミュレーションのパラメータを表 1 に，概要を図 1 に示す．

表 1．計算機シミュレーションパラメータ

Parameter	Unit	Value		
Vibration frequency	[Hz]	40	60	90
Matrix size		64×64		
Slice number		15		
Pixel size	[mm^3]	3×3×3		
Phase number		7, 8, 9		
Expected storage modulus	[kPa]	3.0		

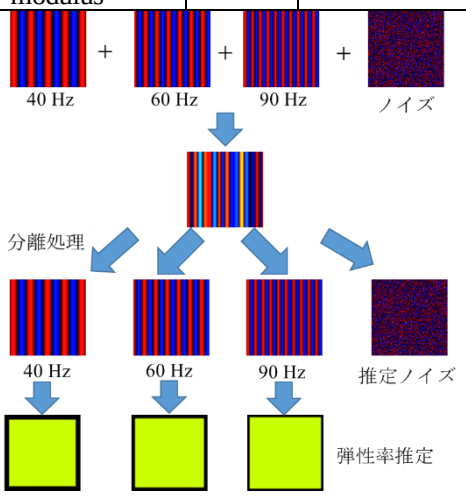


図 1．計算機シミュレーション概要

40, 60, 90 Hz での単一周波数の弾性波画像と MRI 測定系によるノイズを足し合わせることで合成波画像を作成した．実測の単一周波数加振 MRE と同じ Amplitude to noise ratio(ANR)となるように弾性波振幅とノイズレベルを設定した(表 2)．初期位相数が 7 から 9 の合成波画像に対して周波数分離処理を適用し，分離した各周波数の弾性波画像から算出した弾性率分布と単一周波数加振 MRE で得られた弾性率分布，ノイズを含まない弾性波画像を利用した場合の弾性率分布を比較した．

表 2．各周波数の弾性波パラメータ

周波数 [Hz]	ANR	弾性波振幅	ノイズ SD
40	25	10	0.4
60	25	10	0.4
90	7	2.8	0.4

3. 結果

計算機シミュレーションによって得られた弾性率分布を図 2 に，弾性率分布から求めた CV を図 3 に示す．振幅対ノイズ比がそれぞれ 25, 25, 7 の弾性波画像を用いた検討では，同時加振で位相数

が 7，8，9 での弾性率分布の平均値はいずれも真値と一致し，最大の変動係数はそれぞれ 7%，5%，0%以下となった．位相数 9 では加振周波数によらず単一周波数加振 MRE の変動係数よりも小さくなった．これらの結果から，周波数分離処理によるノイズ除去の効果が認められた．

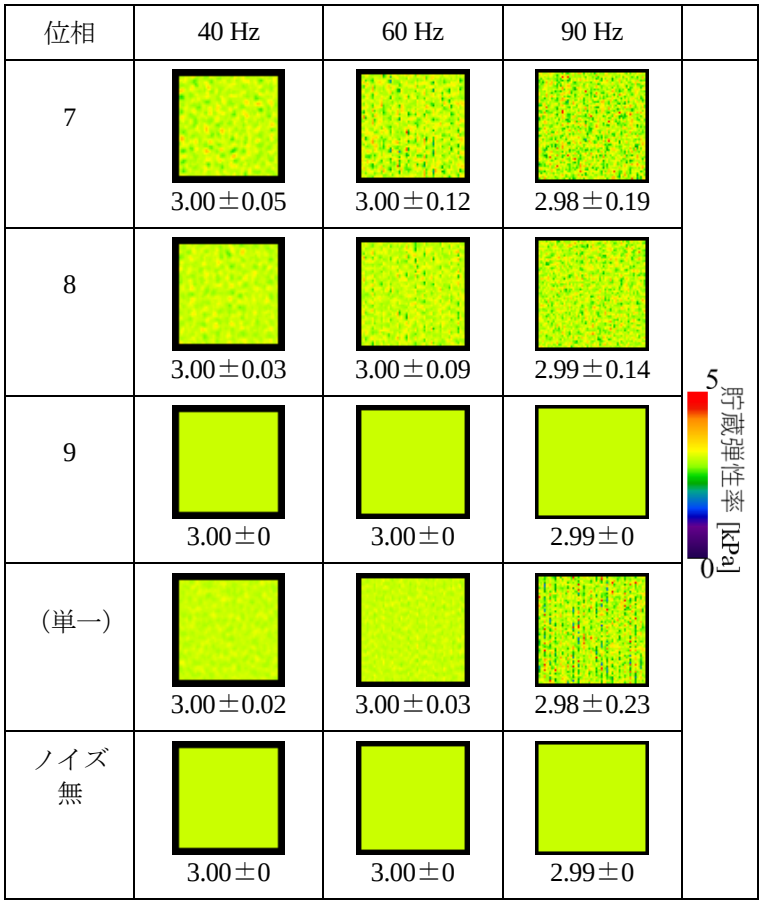


図 2. 弾性率分布

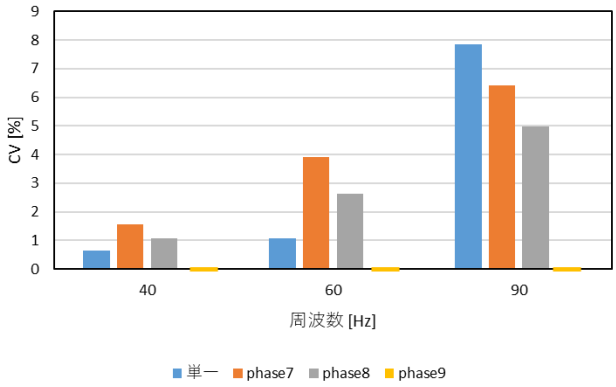


図 3. CV

4. まとめ

計算機シミュレーションにより 3 周波数同時加振 MRE について，取得位相数による撮像時間短縮効果と弾性率推定精度の関係を検討した．単一周波数ごとに 4 位相収集する従来法で 3 周波数を取得する場合と比較して，同時加振で位相数が 7，8，9 での撮像時間短縮効果はそれぞれ 44%，37%，30%となる．実測の MRE での振幅対ノイズ比を参考にした弾性波画像を用いた検討では，同時加振での位相数によらず弾性率分布の平均値は真値と一致し，周波数分離処理のノイズ除去効果によって位相数 9 では加振周波数によらず単一周波数加振 MRE よりも変動係数は小さくなった．以上の結果より，3 周波数同時加振 MRE の有効性が示唆された．

利益相反の有無

なし

文　献

［1］ Muthupillai R, Lomas D J, Rossman P J, et al.:Science **269**, 1995, pp.1854-1857.

［2］ Wuerfel J, Paul F, Beierbach B, et al.: MR-elastography reveals degradation of tissue integrity in multiple sclerosis. NeuroImage **49**, 2010, pp.2520−2525.

［3］ 大橋幸司，菅幹生，石井孝樹 他：複数周波数 MR エラストグラフィにおける撮像時間短縮のための撮像パラメータの検討. 電子情報通信学会技術研究報告 **118(412)**, 2019, pp.239-240.

［4］ 藤原宏志, 積分型公式による MRE での剛性率の再構成．理論応用力学講演会論文集，**57**, 2008, pp.427-428.

Study on reduction of imaging time by simultaneous multi-frequency MR elastography

Riku KUJI[*1], Mikio SUGA[*1,2], Riwa KISHIMOTO[*2], Shingo NOZAWA[*2], Takayuki OBATA[*2]

*1 Chiba University

*2 QST, National Institute of Radiological Sciences

Magnetic resonance elastography (MRE) is a non-invasive imaging technique that measures the tissue viscoelasticity by generating shear waves, imaging their propagation using MRI, and processing to generate a viscoelasticity map. The multi-frequency MRE can acquire the frequency characteristics of tissue viscoelasticity, but it has a problem of long imaging time. This study aims to clarify the relationship between the effect of shortening the imaging time by the number of acquired phases of the wave image and the estimation accuracy of the viscoelasticity in simultaneous excitation at three frequencies of 40, 60, and 90 Hz.

Compared with where three frequencies are acquired by the conventional method, the effect of shortening the imaging time when the number of phases is 7 to 9 by simultaneous excitation is 44%, 37%, and 30%.

In the simulation, the amplitude-to-noise ratios of shear wave images at three frequencies were set to with reference to experimental data. The mean values of the storage modulus distributions all agreed with the true values, and the coefficient of variation was less than 7%, 5% and 0%. At phase number is 9, the coefficient of variation is smaller than the conventional method. These results suggest the effectiveness of 3-frequency simultaneous excitation MRE.

Key words: MRI, MR Elastography, viscoelasticity, shortening the imaging time

数値シミュレーションによる MR エラストグラフィに体動が与える影響の基礎的検討

小賀野　慎[*1]　谷口　陽[*1, 2]　西原　崇[*2]　瀧澤　将宏[*2]　菅　幹生[*1, 3]

要旨

MR エラストグラフィ(MRE)は，撮像対象を加振しながら MRI で取得した弾性波画像から，逆問題を解くことで弾性率を非侵襲的に測定する手法である．MRE では，撮像中の体動が粘弾性率推定に与える影響はこれまで十分には検討されていない．本研究では，体動が MRE に与える影響と，体動による影響を低減可能な撮像条件を数値シミュレーションにより明らかにすることを目的とした．シミュレーションでは，成人男性の腹部(弾性率 3 kPa 均一)が呼吸によって前後方向に変形することを想定し，振動子の設置位置(横波の変位方向)と MR 信号収集方向(phase-encoding 方向)の組み合わせを検討した．腹部変位に対して横波変位の方向が平行で phase-encoding 方向が平行と垂直の場合，弾性率はそれぞれ 2.84±1.5 と 2.72±0.4 となった．また，腹部変位と横波変位の方向が垂直の場合にはそれぞれ 2.97±0.3 と 2.98±0.1 となった．以上の結果より，腹部変位に対して横波変位方向と phase encoding 方向がともに垂直となるようにすることで体動の影響を低減できることが示唆された．

キーワード：MRI，MR エラストグラフィ，体動，数値シミュレーション

1. はじめに

磁気共鳴画像装置（magnetic resonance imaging: MRI）を利用して，生体内の粘弾性率を非侵襲的に評価する手法として magnetic resonance elastography（MRE）[1]がある．MRE は MRI 装置内で外部加振装置を用いて撮像対象内部に弾性波を伝播させ，MRI パルスシーケンスに変位エンコード傾斜磁場（motion encoding gradient: MEG）を付加した MRE パルスシーケンスを用いて撮像することにより，弾性波を画像化する．得られた弾性波画像から逆問題を解くことで粘弾性率を推定する．MRI では撮像中の被検者の動きによって画像にモーションアーチファクトが生じることがある．MRE では，モーションアーチファクトを含む弾性波画像に粘弾性率推定式を適用すると，粘弾性率推定値に誤差が生じることが考えられる．しかし，撮像中の体動が粘弾性率推定に与える影響はこれまで十分には検討されていない．本研究では，数値シミュレーションを用いて，呼吸による腹部領域の体動が MRE 撮像に与える影響を調査し，体動による影響を低減可能な撮像条件を明らかにすることを目的とした．

*1 千葉大学　大学院融合理工学府
〔〒263-8522 千葉市稲毛区弥生町 1-33〕
e-mail: afta7014@chiba-u.jp
*2 株式会社日立製作所　ヘルスケアビジネスユニット
*3 千葉大学　フロンティア医工学センター

2. 方法

• 2. 1 数値シミュレーションの概要

数値シミュレーションのフローチャートを図 1 に示す．呼吸による腹部変位に伴うモーションアーチファクトを再現するため，あらかじめ取得しておいた腹囲の時系列計測データを用いて，各 phase encoding step における腹部変位を MRI の撮像パラメータである repetition time（TR）から求める．成人男性の腹部横断面を模擬した腹部数値モデルは，プロトン密度が均一の楕円形状とし，背中側を固定して前後に膨張収縮させた．

これを強度画像とする．また，想定貯蔵弾性率と加振周波数から計算した波長の進行波を位相画像とする．これらの強度画像と位相画像から複素画像を作成してフーリエ変換し，周波数空間（k空間）に 1 行ずつ充填していく．充填された k 空間画像を逆フーリエ変換して，強度画像と位相画像を生成する．

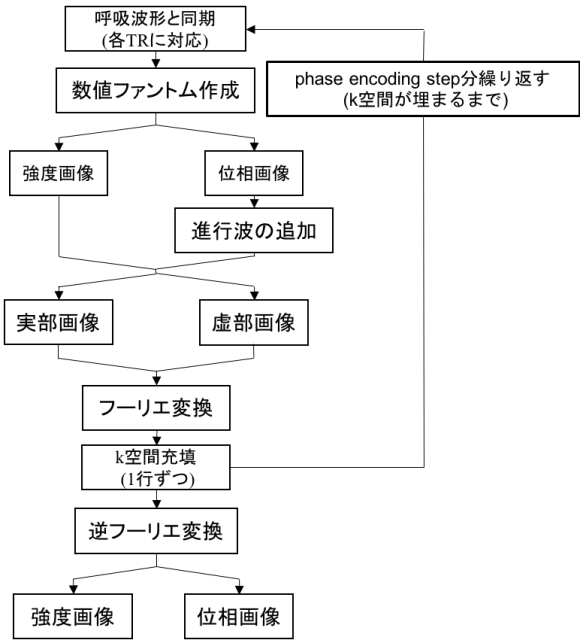


図 1　数値シミュレーションの流れ

- **2．2 数値シミュレーション設定**

腹部数値モデルは，貯蔵弾性率 3 kPa，長軸 30 cm，短軸 20 cm とし，呼吸によって腹壁が体の前後に±1.5cm 動くと仮定した．これを図 2 に示す．腹部数値モデル内部には，貯蔵弾性率 7.5 kPa，直径 5.5 cm の円形の内包物を配置した．今回，腹壁の動きによる影響のみを評価するために，内包物は移動と変形をしないと仮定した．弾性率評価のため，2 つある内包物の外周から 6 pixel 内側の領域を合わせて内包物の関心領域とし，腹部数値モデル外周から 6 pixel 内側の領域で，内包領域の外周から 6 pixel 外側の領域を背景の関心領域とした．

体動の影響を低減可能な撮像条件を明らかにするために，振動子による変位の発生方向と MR 信号収集方向の適切な組み合わせを検討した．具体的には，振動子による変位の発生方向が左右（振動子背面配置）と前後（振動子側面配置）の場合，MR 信号収集の phase-encoding 方向が左右と前後の場合の組み合わせの 4 通りとした（表1）．また，シミュレーションに用いた MRI のパルスシーケンスは RF-spoiled GRE-MRE，撮像パラメータは，TR: 480 ms，FOV: 768×768 mm^2，matrix: 256×256，slice number: 15 枚，slice thickness: 3.0 mm，vibration frequency: 62.5 Hz とした．粘弾性率分布の算出には，積分型粘弾性率分布推定手法（integral type reconstruction formula: ITRF[2]）を用いた．

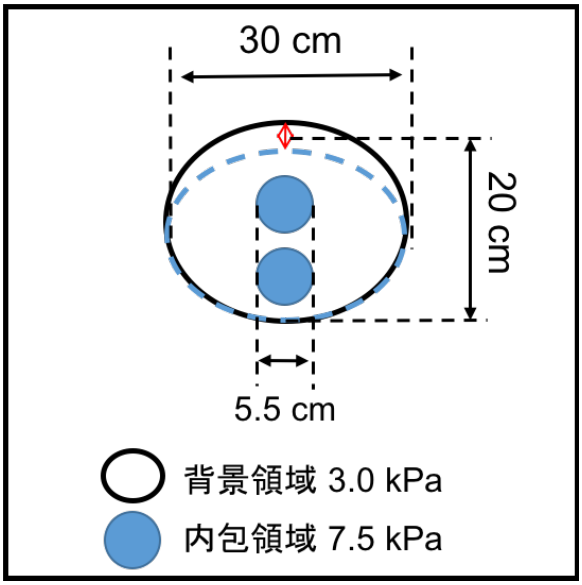


図 2　硬い内包物を配置した腹部数値モデル

表 1　シミュレーション組み合わせ

	phase encoding 前後	phase encoding 左右
横波変位 左右	(a)	(b)
横波変位 前後	(c)	(d)

比較のための腹部の動きがない場合の弾性波画像と貯蔵弾性率分布を図 3 に，貯蔵弾性率のヒストグラムを図 4 に示す．

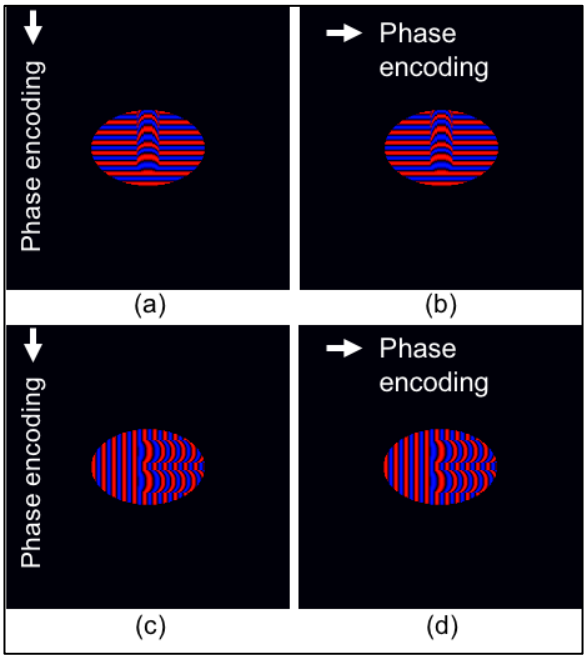


図 3-1 弾性波画像（体動なし）

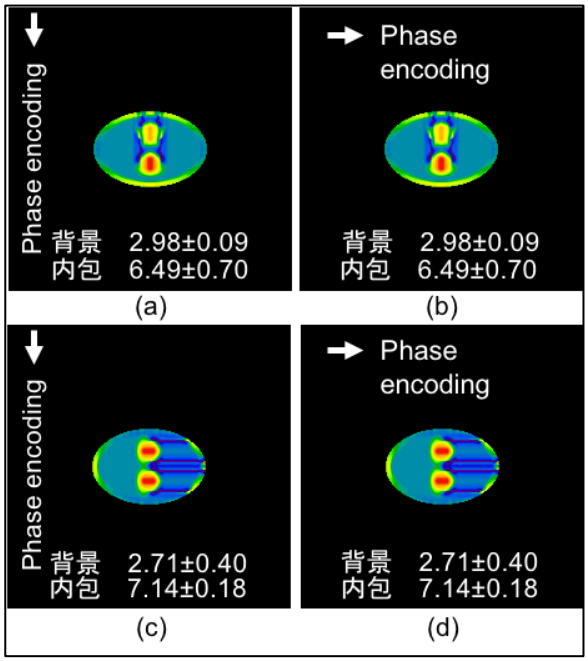


図 3-2 貯蔵弾性率分布（体動なし）

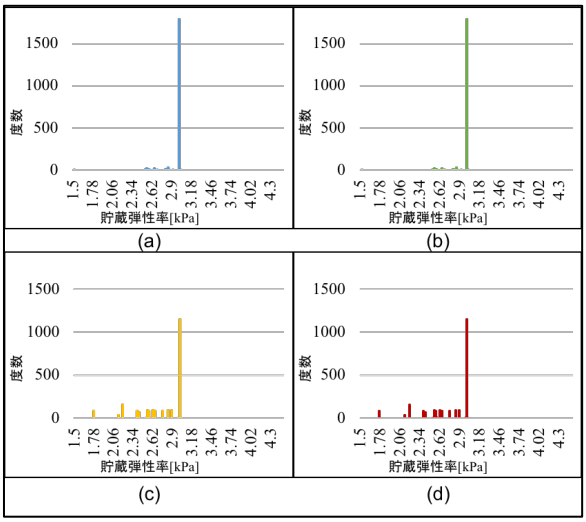


図 4-1 背景領域における貯蔵弾性率のヒストグラム（体動なし）

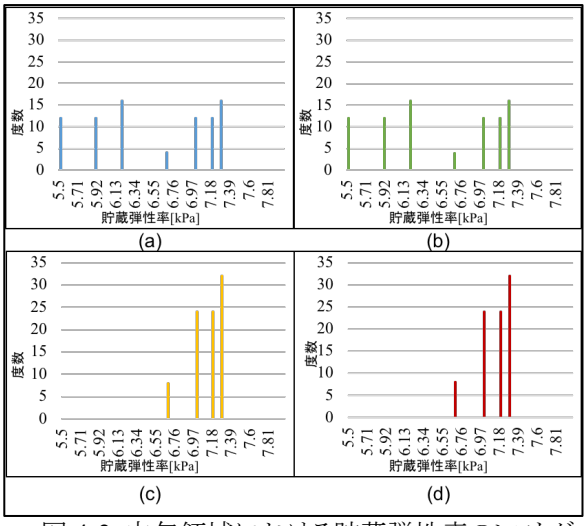


図 4-2 内包領域における貯蔵弾性率のヒストグラム（体動なし）

3. 結果

3．1 弾性波画像と貯蔵弾性率分布

腹部の動きがある場合の弾性波画像と貯蔵弾性率分布を図 5 に示す．図中の数値は背景と内包物の関心領域における貯蔵弾性率の平均値と標準偏差である．また，矢印は phase encoding 方向を表している．

phase encoding 方向が左右方向の場合，貯蔵弾性率は背景，内包とも体動のない場合（図 3-2）とほとんど変化しなかった．一方，phase encoding 方向が前後方向の場合には，弾性波画像に生

じたモーションアーチファクトによって誤差が生じ，特に側面加振の背景領域において体動がない場合との違いが大きくなった（図 5-2(c)）．

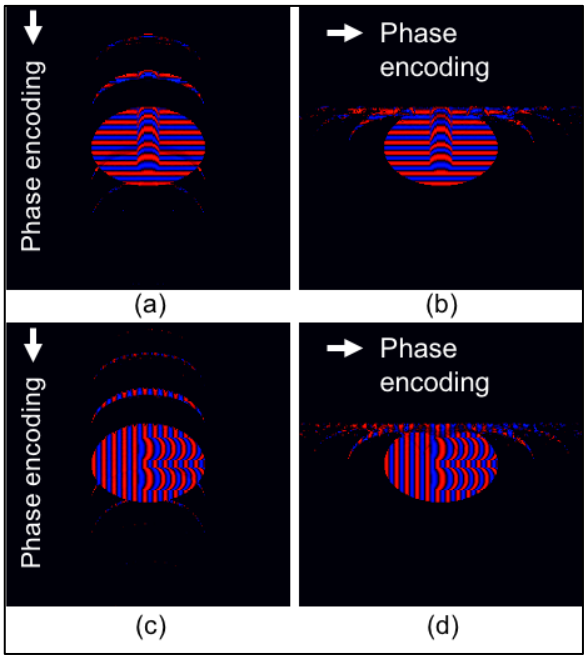


図 5-1 弾性波画像

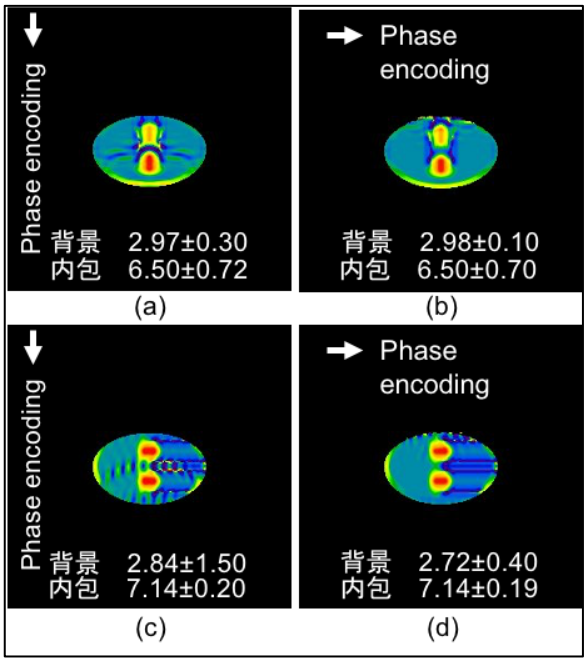


図 5-2 貯蔵弾性率分布

- **3．2 背景領域のヒストグラム**

背景領域における貯蔵弾性率のヒストグラムを図 6 に示す． 背景領域においては，(a)と(c)に比べて，(b)と(d)の phase encoding が左右方向の場合にばらつきが少なく，背面加振でより少なかった．

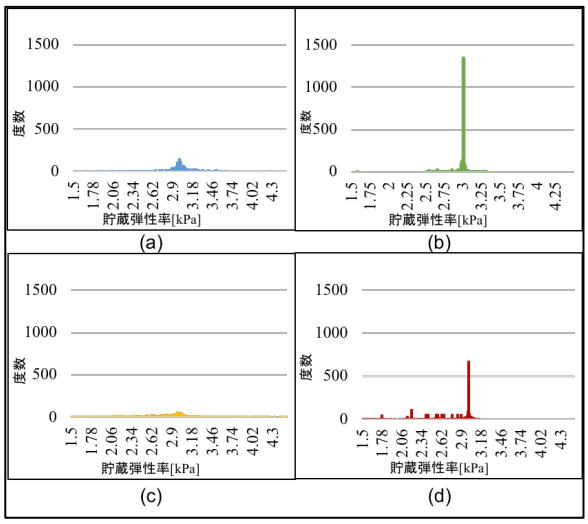


図 6 背景領域における貯蔵弾性率のヒストグラム

- **3．3 内包領域のヒストグラム**

内包領域における貯蔵弾性率のヒストグラムを図 7 に示す．内包領域でも背景領域と同様に phase encoding が左右方向の場合にばらつきが少なくなった．

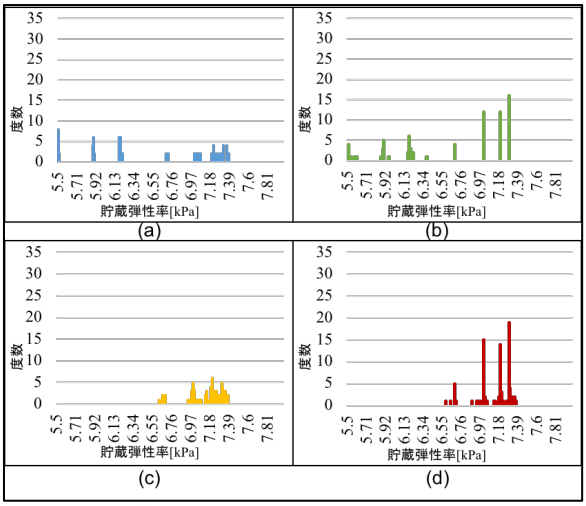


図 7 内包領域における貯蔵弾性率のヒストグラム

4．考察

腹部の動きがない条件でも，横波変位が左右

方向（背面加振）の場合に内包物の弾性率推定誤差が大きくなった．これは，背面側（画像下側）の内包領域を通過した弾性波が腹部側の内包領域に伝播しているためであり，誤差の大きさは加振方向ではなく，加振方向と内包物の位置関係に依存する．腹部の動きがある条件では，phase encoding が上下方向の場合にはモーションアーチファクトが重なる領域が多いため，phase encoding が左右方向の場合よりも弾性率の推定誤差が大きくなっていると考えられる．

5. まとめ

撮像中の腹壁の動きがMRE撮像に与える影響と，体動による影響を低減可能な撮像条件を検討するために数値シミュレーションを実施した．横波の変位方向とphase encodingの組み合わせを検討した結果より，phase encoding と横波変位を体動と垂直となる左右方向にすることで，体動の影響を抑えられることが示唆された．

利益相反の有無

なし

文　献

［1］ Muthupillai R，et al.，Science，**269**，pp．1854-1857，1995．

［2］ 藤原宏志，”積分型公式による MRE での剛性率の再構成”, 第 57 回理論応用力学講演会論文集，2008

機械学習を用いたコーンビーム CT の散乱線除去

村田一心[*1], 橋本航汰[*1], 臼井桂介[*2], 尾川浩一[*1]

要旨

本研究では、コーンビーム CT 撮影で発生する散乱線を、機械学習を用いて除去する方法を提案する. コーンビーム CT は検出器面に対し、２次元検出器アレイを用いるため、コリメータを用いた散乱線の除去が難しい. そのため、再構成画像の画質は劣化してしまう. しかし、その画質を改善することができれば、治療計画 CT にコーンビーム CT を使用できる可能性があり、極めて有用である. そこで本研究では、機械学習を用いた散乱線除去性能を検証する. 学習データの生成には、モンテカルロ法による光子輸送シミュレーションを用いた. すなわち、散乱線有り無しの画像を作成し、それぞれ入力データ、教師データとした. 異なる条件下での散乱線除去性能を調べるため、トレーニングデータおよびテストデータを取得する際、異なる電圧を用いた. 画像の再構成には FDK 法を適用した. 学習結果を適用した画像を評価したところ、散乱線を有意に除去できたことを確認した. この結果は、散乱線の除去に機械学習が有効であることを示す.

キーワード：コーンビーム CT, 散乱線, 機械学習

1. はじめに

コーンビーム CT は画像誘導放射線治療に使われており、治療計画 CT への応用も期待されている. これは、コーンビーム CT が一度のスキャンで 3 次元断層画像を取得できるため、モーションアーチファクトや被ばく量を低減できるためである. しかし、コーンビーム CT は 2 次元検出器アレイを用いるため、コリメータによる散乱線の除去が難しい. そのため、空間分解能の低下や、定量における大きな不定性が問題となる.

そこでこれまで、コーンビーム CT の散乱線を除去する研究が多く行われてきた. モンテカルロ・シミュレーション[1]やカーネルを用いた解析的方法[2], 機械学習を用いた方法[3] が開発されている. 中でも機械学習を用いた方法は、精度や速度の面で非常に優れており、今後の発展が期待されている.

機械学習を用いた方法として、Nomura らは、投影データに対して散乱線除去を適用した. その際、学習に人体ではないファントムを適用してもなお高精度な散乱線除去が可能であることを示した. さらに彼らは、フルファンでの学習結果に対し、ハーフファンのデータを用いた転移学習を行うことで、ハーフファンで取得した画像の散乱線除去も効率的に行える事を示した.

しかし彼らのシミュレーションでは、トレーニングとテストに使用したデータの取得条件は基本的に同一であり、わずかに条件を変えた時の散乱線除去性能は不明瞭である.

そこで本研究では、トレーニングデータとテストデータとで、取得条件を変えた時の散乱線除去性能を調査することにした.

2. シミュレーション

本研究では、実際の CT 画像を使用した散乱線除去シミュレーションを行った. 使用したファントムは胸部 X 線 CT 画像の CT 値から骨と軟部組織に分割したものである. その際、減衰係数はそれぞれカルシウムと水で近似した. こ

*1 法政大学理工学部
〔〒185-8584 東京都小金井市梶野町 3-7-2〕
e-mail: murata.kazumi.204@gmail.com
*2 順天堂大学保健医療学部

のようにして得られたファントムに対し、モンテカルロ法による光子輸送シミュレーションを行い、プライマリ光子と散乱線の強度を推定した. シミュレーションのジオメトリは、線源とファントムとの間を 32 cm, 線源と検出器との間を 44 cm とし、1 度ずつ 360 方向の投影データを取得した. かけた管電圧は 90, 100, 130 kV とし、一方向あたり 10 億個の光子を放射させた. ピクセルサイズは 0.2 cm×0.2 cm とし、1 ピクセルあたりの光子数は約 16000 個であった. 投影データに対して FDK 法を適用し、プライマリ光子のみ、および散乱線を含む再構成画像を生成した. 再構成画像のサイズは 128 × 128×128 voxel, ピクセルサイズは検出器と同様 0.2 cm×0.2 cm×0.2 cm とした.

このようにして取得したデータを使い、scatter-net [3]に学習させた. このネットワークは U-net と Res-net を組み合わせたもので、最初に画像を圧縮しつつ特徴量を抽出し、入力の畳み込み結果との差分を出力する. 画像圧縮の際にも Res-net を組み込み、効率よく学習できるようにしている. ここで、複数の層を一つのブロックとしてまとめている. ブロックには 4 種類あり、主に畳み込みを行うコンバートブロック、差分を計算するレズブロック、画像を縮小するダウンブロックに、画像を拡大するアップブロックである. 本研究では、最初にコンバートブロックとレズブロックを一つずつ入れ、その後ダウンブロックを 2 個入れて画像を圧縮し、アップブロックを 2 個入れて画像を復元させるようにした. アップブロックの直後には、ダウンブロック後の同じサイズの画像も同時に入力する事で、効率よく学習させるようにした. ネットワークの構造は基本的に先行研究[3]と同様だが、画像サイズが小さいことを考慮し、アップ/ダウンブロックの個数は 2 個ずつ減らした.

学習はプライマリ光子のみで再構成した画像を教師、散乱線込みの画像を入力、損失関数は Mean Absolute Error (MAE)、最適化法は Adam として行った. エポック数は 1000, バッチサイズは 8 とした. 画像の端を避けるため、上部と下部の画像は使わないこととし、合計 68 枚の画像を用いて学習を行った. ここで、シミュレーション条件による散乱線除去性能への影響を見るため、学習には管電圧 100 kV のデータのみを使用した.

このようにして学習したネットワークを使い、散乱線を除去し、プライマリ光子のみで再構成した画像との比較を行った.

3. 結果

図 1 に再構成画像の例を示す. 一見すると両者に違いは見えないが、その差分 (図 2) はどの管電圧の画像においても、領域によって画素値が大きく異なることがわかる. MAE は管電圧 100, 90, 130 kV の画像でそれぞれ、0.0039, 0.0043, 0.0031 cm^{-1} であった.

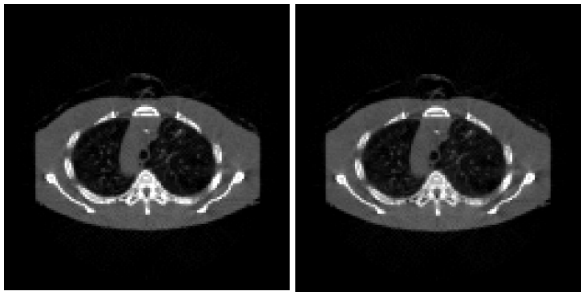


図 1 再構成画像. 左はプライマリ光子のみ、右は散乱線を含む.

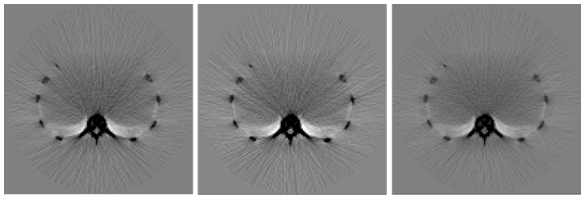


図 2 散乱線込みとプライマリ光子のみの再構成画像との差分. かけた管電圧は左から順に 100kV, 90kV, 130kV. 表示範囲は-0.05 ~ 0.05 cm^{-1}.

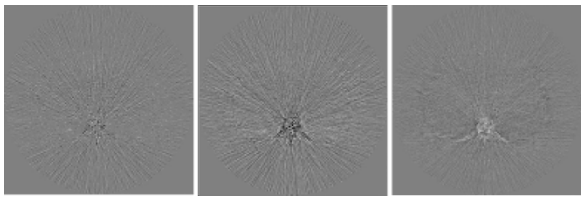


図 3 散乱線除去後の画像とプライマリ光子のみで再構成した画像との差分. かけた管電圧は左から順に 100, 90, 130 kV. 表示範囲は-0.05 ~ 0.05 cm^{-1}.

一方、ネットワークを通して散乱線を除去した後の画像とプライマリ光子だけを用いた再構成画像との差分 (図 3) を見ると、どの管電圧においても散乱線を綺麗に除去できていることがわかる. MAE は各管電圧の画像でそれぞれ、0.0013, 0.0022, 0.0016 cm^{-1} となり、2~3 倍画質が改善されたことを意味する.

しかし、学習に用いた管電圧 100 kV の画像に対してネットワークを適用した時が最も散乱線の除去性能が高かった. これは汎用性の高い散乱除去ネットワークを構築するには、幅広い条件の画像を学習に用いる重要性を示唆する. そのような学習には当然データ量と計算コストの問題が発生する. そのため将来研究では、散乱線除去性能と学習に使うデータの取得条件との関係を詳しく調査する必要がある.

4. まとめ

本研究では、機械学習を用いたコーンビーム CT の散乱線除去シミュレーションを行った. その結果、学習に用いたデータと異なる管電圧で取得された画像に対しても綺麗に散乱線を除去することができた. しかし、学習時と条件を変えることで散乱線の除去性能は劣化することがわかった. 将来研究では、異なる条件下での散乱線除去性能劣化を防ぐ方法を考案する必要がある.

謝辞

本研究は JSPS 科研費 20K20232 の助成を受けたものです.

利益相反の有無

なし

文　献

［1］ Xu Y., Bai T., Yan H., et al: A practical Cone-beam CT Scatter Correction Method with Optimized Monte Carlo Simulations for Image-Guided Radiation Therapy. Phys. Med. Biol., 60 (9), pp.3567-3587, 2015

［2］ Sun M, Star-Lack JM. Improved scatter correction using adaptive scatter kernel superposition. Phys Med Biol. 55, 6695–6720, 2010

［3］ Nomura Y., Xu Q., Shirato H., et al. : Projection-domain scatter correction for cone beam computed tomography using a residual convolutional neural network. Med. Phys. **46** (7), 2019

デコンボリューション法による
静止型ピンホール SPECT システムの空間分解能改善

大越　迪[*1]　村田　一心[*2]　尾川　浩一[*2]

要旨

従来の SPECT システムでは, データを取得する際に検出器を回転させるため, データ収集に時間がかかり, 放射性医薬品の動態情報を得ることが困難である．そこで近年，検出器を回転させない静止型ピンホール SPECT システムの開発が進められている．このシステムでは，検出器を回転させない代わりに，複数のピンホールを有するコリメータを用いることで，足りなくなった投影方向数を補っている．しかしこの時，ピンホール径が有限の大きさを持つため，空間分解能の劣化が生じてしまう．その補正方法として，従来 7rays 法が用いられてきたが，特にピンホール径が大きい場合，空間分解能の十分な補正はできなかった．そこで本研究では，再構成中心に対する点拡がり関数を推定し，得られた投影データに対してデコンボリューションを行い，空間分解能の改善を試みた．再構成画像の画質評価を行った結果，従来の方法よりも定性的、定量的に高い空間分解能が確認された．

キーワード：静止型 SPECT システム，ピンホールコリメータ，デコンボリューション，点拡がり関数

1. はじめに

SPECT は体内に投与した放射性物質を含む薬剤から放出される放射線を検出することで，その放射性同位元素の集積を画像化する装置である．X 線 CT などと異なり臓器の形だけでなく，臓器の機能情報も確認できるという利点がある．従来の SPECT システムでは，検出器を回転させることでデータの取得を行っているが，データ収集に時間がかかるため放射性医薬品の動態情報を得ることが困難という欠点がある．そこで近年，検出器を回転させない静止型ピンホール SPECT システム[1][2]の開発が進められている．このシステムでは，検出器を回転させない代わりに，複数のピンホールを有するガンマカメラを使用することで，投影方向を増加させるとともに，感度を改善している．しかしこのようなピンホールデータ収集システムでは，ピンホール径が有限の大きさを持つことによる，空間分解能劣化の影響を強く受ける．このような劣化を補正する方法として，従来 7rays 法[3]などが提案されてきた．しかしこの方法は，本来無数の方向から到来する光子をわずか 7 本の投影線で近似するため，特にピンホール径が大きい場合に空間分解能の劣化を十分に補正することができなかった．

そこで本研究では，逆畳み込み法を用いた空間分解能改善手法を提案する．SPECT システムにおける投影データは，ピンホール径に応じた広がりを持つ．そのため，点拡がり関数 (point spread function; PSF) の逆畳み込みを行うことで，ピンホール径が無限小の場合の投影画像に近似することができる．したがって，ピンホール径が有限であることによる空間分解能の劣化を防ぎ，高画質な画像を取得できると期待さ

*1 法政大学大学院理工学研究科応用情報工学専攻
〔〒184-8584 東京都小金井市梶野町 3-7-2〕
e-mail: michi.okoshi.6u@stu.hosei.ac.jp
*2 法政大学理工学部応用情報工学科

れる． 本研究では、その有用性を検証するため，静止型 SPECT システムにおける画像再構成シミュレーションを行った．その際，簡単のためコリメータ内のピンホールは中心の一つだけとした．

2. シミュレーション

本研究では，シングルピンホール SPECT を用いて，PSF による逆畳み込みを用いた補正方法の有用性を示すためにシミュレーションを行った．本シミュレーションでは，ターゲットである脳を模擬したファントムの周囲に検出器を 18 個，20 度ずつ配置した静止型 SPECT システムを用いた．また，物体中心からコリメータ中央面までの距離を 25 cm，コリメータ中央面から検出器までの距離を 7.5 cm とし，検出器のサイズは 512×256 pixels，ピクセルサイズを 0.08×0.08 cm^2とした．ピンホールは，コリメータ中心に一つ配置した．その際のピンホール径は，本研究における提案手法の効果を検証するため，大きめの 0.5 cm とした．この時の開口角は，先行研究[2]に合わせ 24 度，形状はナイフエッジとした．

これらの条件の下，モンテカルロ法[4]による光子輸送シミュレーションを行った．本シミュレーションの目的は，ピンホール径に由来する空間分解能劣化に対する，本手法による改善効果を明らかにすることにある．そこで簡単のため，検出器の固有分解能は考慮せず，ターゲットによる散乱も計算には入れなかった．すなわち理想的な検出器を想定し，散乱線は完璧に除去できたものとした．光子の入射方向による感度の劣化は，モンテカルロ法を用いた平均立体角推定により補正した．

このようにして得られた投影データは，ピンホール径に応じた広がりを持っている．これに対して PSF の逆畳み込みを適用するのが本研究における提案手法である．ターゲット中心から放射された X 線の PSF は，入射角に応じた立体角の補正をしているため，矩形関数で表される(図 1)．各ピンホールの視野内では，PSF は一定とみなし，投影データに対してフーリエ空間で逆畳み込みを行った．この時，高周波成分の増幅を防ぐため，ナイキスト周波数の 1/2 以上の周波数成分は除外するようにバターワースフィルタを設計，適用した．

3. 結果

図 2 に，逆畳み込み適用前後の投影画像を示す．逆畳み込みを適用したことで，細部の構造が見やすくなったことがわかる．こうして得られた投影画像に対し，最尤推定―期待値最大化法 (ML-EM 法 [4][5])を適用し，画像再構成を行った．その際，ピンホール径は無限小として計算した．また，比較のため逆畳み込みを適用していない投影画像に 7 rays 法を適用して再構成した．

図 3 に得られた投影データ(図 2(上))を用いて 7rays 法による補正を行った結果と，逆畳み込み後の投影データ(図 2(下))を用いて本提案手法による補正を行った結果を示す．図 3 を見ると，本提案手法による再構成画像の方がより細かい構造まで再構成できていることがわかる．また両者の PSNR 値はそれぞれ，10.6 と 11.6 [dB]となった．このことからも，本提案手法の方がより正確に画像を再構成できていることがわかる．

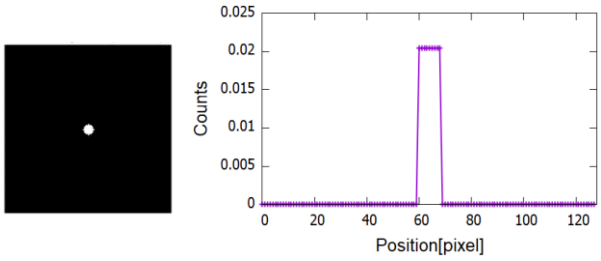


図 1 推定された点拡がり関数とその中心の横方向プロファイル．本シミュレーションにおけるジオメトリから解析的に推定した．

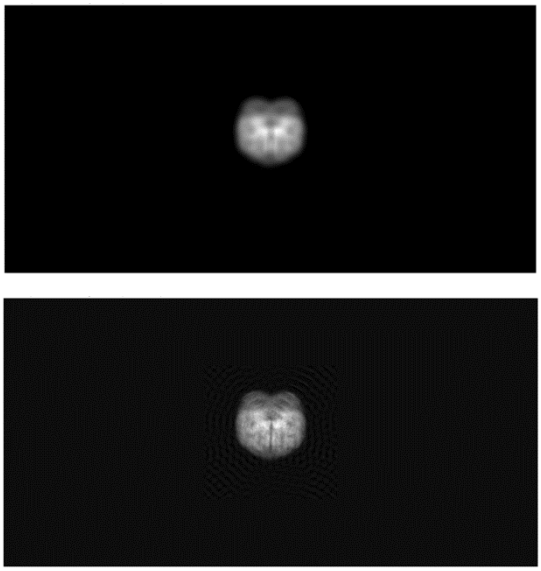


図 2 シミュレーションで得られた投影データ(上)とPSF による逆畳み込みを適用した投影データ(下). 逆畳み込み後の方が, 細部の構造が鮮明になっていることが分かる.

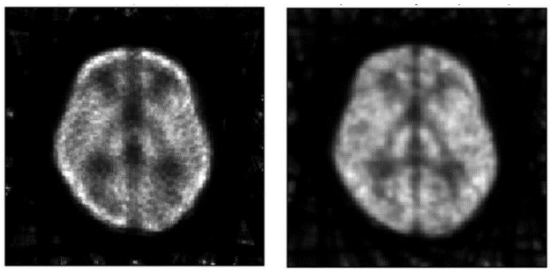


図 3 再構成画像の比較. 左が従来の 7 rays 法、右が本研究で提案する逆畳み込み法による再構成画像. 従来法と比べ, 提案手法では細部の構造まで再構成できていることがわかる.

4. 考察

図 3 の再構成結果の比較より，検出器中心における点拡がり関数を推定し，それを用いて補正を行った結果, 従来の補正方法である 7rays 法による結果よりも，画像的にも PSNR 値的にも向上した．これは，点拡がり関数を用いることで，ピンホール径の大きさによるぼけを正しく推定でき，従来の補正方法である 7rays 法の近似による補正方法より精度が高くなったためだと考えられる.

したがってこの結果は，逆畳み込み法が静止型 SPECT システムにおける空間分解能の改善に有効であることを示す. しかし，本研究では簡単のため，ピンホールは中心の一つのみしか用いておらず, 視野内で PSF は一定とみなした. しかし，コリメータ内に複数のピンホールがある場合，中心以外のピンホールにおける PSF は視野内で大きく異なることが予想される. そこで将来研究では，複数のピンホールを持つシステムにおいて，場所ごとに異なる PSF を適用する手法を開発する予定である.

5. まとめ

本研究では，静止型 SPECT システムにおける投影画像に対して PSF の逆畳み込み法を適用し，空間分解能を改善する手法を提案した．理想的な条件でシミュレーションを行い，7 rays 法による再構成画像と比較した結果，定性的にも定量的にも本手法の有用性が示された.

利益相反の有無

なし

文　献

［1］ H. Kubota，et al.，“Multi-pinhole SPECT system with a triple head gamma camera，” IEEE Nuclear Science Symposium, Medical Imaging Conference and Room-Temperature Semiconductor Detector Workshop(NSS/MIC/RTSD), DOI:10.1109/NSSMIC.2016.8069609，2016

［2］ Y. Fujishiro，et al.，“List Mode Image Reconstruction With a Multi-pinhole Triple Head SPECT System，”IEEE Nuclear Science Symposium and Medical Imaging Conference Proceedings(NSS/MIC) , DOI: 10.1109/NSSMIC.2018.8824476，2018

［3］ A. Andreyev，et al.，“Pinhole SPECT Reconstruction Using Blobs and Resolution Recovery，”IEEE Transaction on Nuclear Science，vol. 53，No. 5，pp.2719-2728, DOI:10.1109/TNS.2006.878004，2006.

［4］ F. P. Jansen，et al. ，“Uniformity correction

using non-uniform floods，”IEEE Nuclear Science Symposium and Medical Imaging Conference，pp.2314-2318，DOI:10.1109/NSSMIC.2010.5874197，2010

［5］ 橋本　雄幸，小島　広行，”C 言語による画像再構成の基礎,”医療科学社, 2006.

lmprovement of the spatial resolution of a static pinhole SPECT system with a deconvolution method

Michi Okoshi[*1], Kazumi MURATA[*2], Koichi OGAWA[*2]

*1 Graduate School of Science and Engineering, Hosei University

*2 Faculty of Science and Engineering, Hosei University.

As a conventional SPECT system rotates detectors to acquire projection data, it takes long time and cannot perform a dynamic study of an organ. In contrast, a multi-detector SPECT system with multi-pinhole collimator requires no detector rotation in data acquisition. In this pinhole SPECT system, the quality of images depends on a pinhole size, i.e. a large pinhole size leads to high sensitivity but low spatial resolution. To improve the spatial resolution of reconstructed images, a 7-rays method is sometimes effective, but it is not enough method to improve it. The purpose of our study is to improve the spatial resolution of reconstructed images acquired with a multi-pinhole SPECT system. To improve the spatial resolution, we used a deconvolution method. The deconvolution kernel used was a point spread function(PSF) of a point source located at the center of the imaging area. For simplicity, we used a single PSF measured with a pinhole at the center of the detector. Reconstructed images were evaluated with a peak signal to noise ratio. The results showed that our proposed method could improve the spatial resolution of reconstructed images compared with that corrected with the 7-rays method.

Key words: static SPECT system, pinhole collimator, deconvolution, point spread function

At-211 内用療法おける ETCC を用いた画像再構成

山田　宙輝[*1]　村田　一心[*2]　谷森　達[*3]　尾川　浩一[*2]

要旨

電子飛跡追跡型コンプトンカメラ (ETCC) は，コンプトン散乱時に発生する反跳電子の軌跡を計測できるため,通常のコンプトンカメラでは得ることができない，光子が放出された方向の精度を向上できる．このため，従来のコンプトンカメラと比較して少ないイベント数で画像を再構成できる可能性がある．本研究ではその有効性を検証するために，近年注目されている ^{211}At 内用療法のイメージングを模したシミュレーションを行った．シミュレーションでは半径 10 cm の水球内に ^{211}At の点線源を設置し，再構成画像の画質を評価した．この結果，大きな腫瘍を見つけるのに十分な空間分解能が得られることがわかった．これは，ETCC が ^{211}At 治療のイメージングに有効であることを示す.

キーワード：コンプトンカメラ，画像再構成，^{211}At 内用療法

1. はじめに

ETCC (Electron Tracking Compton Camera) は，コンプトン散乱角の他に反跳電子の情報を用いる新しいコンプトンカメラである[1〜3]．従来型のコンプトンカメラでは散乱角の情報のみを利用するため，1 つの光子では円錐曲面状にしか線源位置を推定できず，画像再構成には多くの光子が必要となる．一方 ETCC は，コンプトン散乱時の電子の軌跡情報を加えることで，原理的には飛来方向を一意に特定できる．そのため，より少ない光子数での画像再構成が期待できる．

近年, ^{211}At を用いた内部療法が注目されている.^{211}At 内部療法は，α 線を利用した放射線治療である．飛程距離が短いα線を用いることで正常細胞への影響を低減し，腫瘍部位に集中的に攻撃できる．そのため，従来の放射線療法よりも高い治療効果が期待できる．

本研究では, ^{211}At 内部療法のイメージングにおける ETCC の適用可能性を検討した．^{211}At の壊変では，α線と同時γ線 (687, 570, 897 keV)も放出される[4]．そのため，これらのγ線を ETCC で検出すれば，患者の体内における ^{211}At の分布を映像化できると考えられる．本研究では, ^{211}At 分布の画像再構成シミュレーションを行い, ETCC による画像再構成能力を調査した．

2. シミュレーション方法

本研究では ETCC を用いた ^{211}At 内部療法のイメージングの原理実証を行うため，図 1 に示す簡単なジオメトリでシミュレーションを行った．中心に ^{211}At の点線源を置き，その周囲は半径 10 cm 水球で覆った．水球に隣接して２台の ETCC を直交に配置した. ETCC は電子飛跡を検出するガス検出器 (detector 1) および散乱γ線を検出するシンチレーション検出器 (detector 2) から構成されている．ガス検出器は 20×20×20 cm^3 の領域に 3 気圧の CF_4 ガスを充満させたもの，シンチレーション検出器は GSO を用いた 20×20 cm^2 のアレイとした[3].

データの取得はモンテカルロ法による光子輸送シミュレーションで行った．^{211}At から放射されるγ線の相互作用を計算し，ガス検出器内でコンプトン散乱した光子の座標とエネルギ

*1 法政大学大学院理工学研究科応用情報工学専攻
〔〒184-8584 小金井市梶野町 3-7-2〕
e-mail: hiroki.yamada.7z@stu.hosei.ac.jp
*2 法政大学理工学部
*3 京都大学大学院理工学研究科

ー，およびそれに伴う反跳電子の軌跡を測定した．これらの測定の不定性に起因する入射光子の到来方向の不定性は，コンプトン散乱角 θ の誤差である Angular Resolution Measure (ARM)とコンプトン平面における ϕ 方向の誤差である Scatter Plane Deviation (SPD)の実測値[3]を用いて考慮した．これらはコーシー分布に従い，その FWHM はそれぞれ 5°，102°(80 keV 未満は 74°)と測定されている[3]．線源の強度は 1.0 MBq，データ収集時間は 10 分とした．

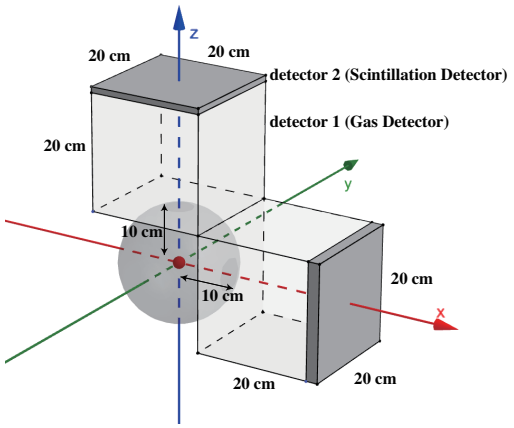


図 1 シミュレーションジオメトリ．原点の赤い丸は ^{211}At の線源を，その周囲の灰色の領域は半径 10 cm の水球を示す．水球に隣接して 2 台の ETCC を配置した．

画像再構成には，高い画質を得るため，検出確率と Visibility を計算に入れた List-Mode ML-EM 法を用いた．ML-EM 法は投影と逆投影を繰り返すことで画像を再構成する手法である．List-Mode ではイベントごとに投影・逆投影を行う．その際，ETCC によって一意に決定された到来方向に逆投影することもできる．しかし，推定した到来方向の不定性の確率分布を考慮した逆投影を行うことで，より精度の高い画像再構成が可能になる．本研究では，検出確率の有無による再構成画像の比較も行った．

Visibility とは，各イベントによる逆投影空間と再構成空間との重複領域である(図 2)．再構成空間の領域サイズは有限のため，到来方向の不定性により領域の端に逆投影した場合，Visibility は極端に小さな値となり (図 2 右)，画像再構成における大きな不定性になる．そこで，検出確率に Visibility を組み込むことで，より高画質な再構成画像が得られると考えられる．そこで本研究では，Visibility の有無による再構成画像の画質も比較した．

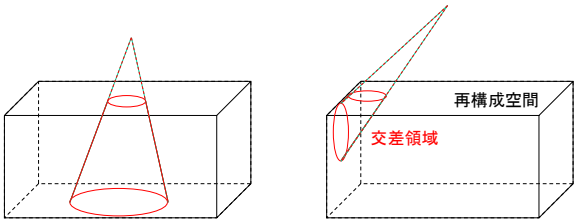


図 2 Visibility の概念図．再構成空間の端を通るコーン状の逆投影空間では，Visibility が小さくなる．これは画像再構成に大きな系統誤差を与える．

3．結果

本シミュレーションを行った結果，1037 個の光子が検出された．これらのイベントを用いて画像再構成を行った結果を図 3 および 4 に示す．

図 3 は 検出確率の有無による再構成画像の比較を示す．検出確率を用いない場合では，中心の線源位置すら全く確認できない (図 3 左)．一方，検出確率を用いることで，^{211}At 分布が中心に集中していることが見て取れる (図 3 右)．

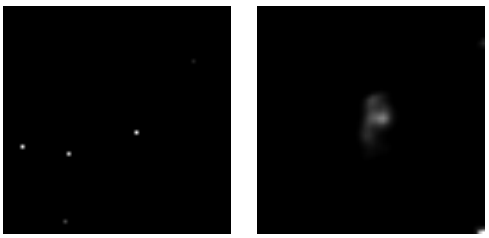


図 3 検出確率の有無による再構成画像の比較．左：検出確率なし．右：検出確率あり．

図 4 は検出確率を考慮し，Visibility の有無による再構成画像の比較を示す．Visibility を考慮しなかった場合 (図 4 左) は，画像の端に外れ値が生成されてしまった．一方，Visibility を考慮することで，そのような外れ値は出現せず，中心集中した ^{211}At の分布を再現することができた．両者の中心の横方向プロファイル (図 5)

を見ると, Visibility を計算に入れることで, 空間分解能も 12 mm から 9 mm (FWHM) に改善されたことがわかる.これは大きな腫瘍であれば判別できる程度の空間分解能である.

これらの比較より, 画像再構成に光子の到来方向の不定性に伴う検出確率および Visibility を考慮することで, ETCC による ^{211}At 分布を再構成できることがわかった. ETCC による到来方向の不定性は ARM が 5°, SPD が 102° と大きい. しかし本研究の結果は, それらを適切に取り扱うことで, 線源の分布が適切な位置に収束することを示している.

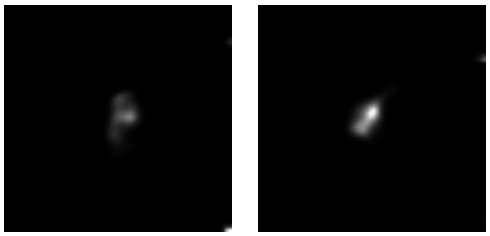


図 4 Visibility の有無による再構成画像の比較. 左 : Visibility なし. 右 : Visibility あり. ここで検出確率は両者とも考慮に入れている.

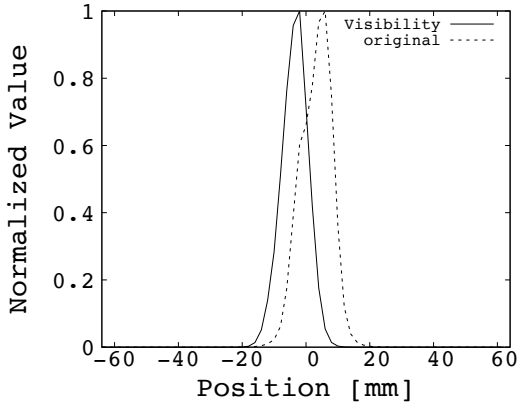


図 5 Visibility の有無による再構成画像の中心の横方向プロファイル. 破線 : Visibility なし. 実線 : Visibility あり.

4. まとめ

本研究では, ETCC を用いた ^{211}At 内部療法のイメージングの可能性を検討した. 画像再構成シミュレーションを行い, その際光子の到来方向推定に伴う不定性も考慮した結果, 1037 個という少数のイベントでも線源の分布を再構成できることがわかった. 再構成された線源分布の FWHM は 9 mm と, 大きな腫瘍を判別するには十分な空間分解能であった.これらの結果は ETCC が ^{211}At 内部療法のイメージングに有効であることを示唆する.

利益相反の有無

なし.

文　献

［1］ M. Inagaki, K. Ogawa, T. Tanimori et al.: Image Reconstruction With an Electron Tracking Compton Camera. IEEE Nuclear Science Symposium and Medical Imaging Conference, 2019

［2］ T. Tanimori, Y. Mizumura, A. Takada, et al.: Establishment of Imaging Spectroscopy of Nuclear Gamma-Rays based on Geometrical Optics. Scientific Reports 7, 41511, 2017

［3］ T. Tanimori, H. Kubo, A. Takada, et al.: An electron-tracking Compton telescope for a survey of the deep universe by MeV gamma rays. Astrophysical Journal, vol. 810, no, 1, id. 28, 2015

［4］ Y. Nagao, M. Yamaguchi, S. Watanabe, et al.: Astatine-211 imaging by a Compton camera for targeted radiotherapy.; Appl. Radiat. Isot., vol. 139, 2018

Automated detection of malignant tumors on PET/CT images and statistical images

Takumi AIHARA[*1], Takeshi HARA[*1], Xiangrong ZHOU[*1], Hiroshi FUJITA[*1], Satoshi ITO[*2] , Masaya KATO[*2], Masaki MATSUSAKO[*3],Tetsuro KATAFUCHI[*4]

*1 Department of Electrical, Electronic and Computer Engineering, Faculty of Engineering, Gifu University[1-1 Yanagido, Gifu, Gifu 501-1193, Japan]e-mail: aihara@fjt.info.gifu-u.ac.jp

*2 Department of Radiology, Daiyukai General Hospital

*3 Department of Radiology, St. Luke's International Hospital

*4 Faculty of Health Science, Gifu University of Medical Science

Abstract

PET images are useful for finding cancer and checking if the cancer has spread or not. In addition, it has been reported that Z-score images, which are statistical images, can distinguish the normal regions from the abnormal regions than the SUV (Standardized uptake value). The purpose of this study was to develop an automated detection system of tumors in PET/CT images using Z-score images and reducing the burden of doctors. Z-score image is created by comparing pixel values of normal model and target image, so I created Z-score images using TPS (Thin-Plate Spline) transformation for normal models. Many false positives (FPs) are contained in the candidate region. I classified true positives (TPs) and FPs using CNN (convolutional neural network) and deleted FPs. I used 63 abnormal cases as detection targets, and the number of tumors was 336 in the trunk of the body. The number of detected tumors was evaluated using TLG (Total Legion Glycolysis), which is an index of tumor volume and degree of glucose metabolism, and the center of gravity of the tumor area. As a result of experiment that inputs PET images and Z-score images to CNN, 80.4% (270/336) of tumors were extracted by this method, and 18.9 FPs remained per case.

Keywords：PET Z-score CNN

1. Purpose

Many medical images are taken in the medical field, and it is a heavy burden for the doctor to read all images. Although PET images are useful for detecting malignant tumors and detecting systemic metastases, but currently only methods for specific tumors are used. Therefore, in this study, we will develop an automatic detection system for tumors of the trunk using the image statistical analysis method in PET/CT images.

2. Method

2.1 Creating human body region images and organ region images

First, in order to add anatomical information to PET images, we created human body region images from CT images. We extracted the regions that the pixel value is larger than 700 from CT images and labeled. The label with the largest area in the Axial section is judged as the human body region, so we deleted other labels. After that, we deleted the head, arms, and legs labels from the human body region.

Second, we created organ region images. The organ region was semi-automatically extracted using GraphCut. The organs to be extracted were the left and right lungs, heart, liver, spleen, left and right kidneys, and bladder. The organ region images were resized according to the size of PET images and aligned position with the standard image using the centroid of the organ region. Also, considering the Physical difference between men and women, we prepared different images for men and women.

2.2 Anatomical standardization

Z-score images need to be anatomically standardized in order to compare pixel values of the normal model and the target image. Therefore, we set the Landmark (LM) for anatomical standardization to human body region images and organ region images. we used Thin-plate spline (TPS) method [3] for anatomical standardization. TPS method is a method of non-rigid transformation by setting the LM before movement and after movement. In this study, target images were transformed using LM of the standard image.

2.3 Creating Z-score images

First, we created an average image and a standard deviation image from normal cases standardized by TPS transformation. There are large differences in physiological accumulation among normal cases, so we used trimmed mean to create normal model.

Second, we used TPS method for normal model using abnormal case as the target case and calculated Z-score of each pixel using equation (1).

$$\mathrm{Z-score(x,y,z)} = \frac{SUV(x,y,z) - mean(x,y,z)}{standerd\ deviation(x,y,z)} \qquad (1)$$

As seen Fig.1, the Z-score image shows less physiological accumulation at the heart and bladder in abnormal cases.

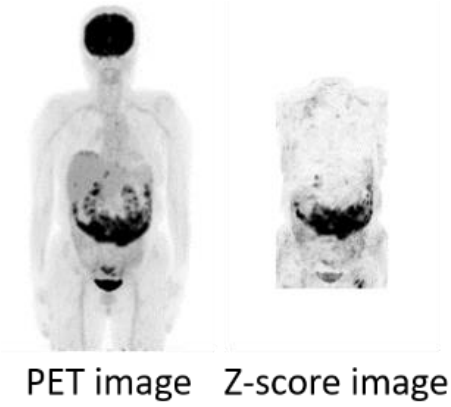


Fig. 1 PET image and Z-score image.

2.4 Automatic detection tumor regions

In order to determine the candidate region of malignant tumors, we created a histogram from Z-score images and decide the threshold. The volume of the abnormal region in the body is smaller than that of the normal region, so it is considered that the maximum frequency class in the histogram is the normal region. Therefore, we set the threshold of Z-score to half the maximum frequency. Also, we set the threshold of Standardized Uptake Value (SUV), which is a semi-quantitative index of PET images, to 1.5 for all cases. This is because the SUV is generally judged as malignant when it is 2 to 2.5 or higher, but we decided low accumulation tumors as target. We set the region where SUV and Z-score were higher than the threshold as the candidate region. Also, we consider labels with a volume of 5 voxels or less as noise and deleted.

2.5 Delete false positives

There are many false positives in tumor labels. In this study, there is a variation in the volumes and shapes of tumors because we targeted tumors of the trunk. Therefore, it is difficult to delete false positives like conventional methods [1][2]. In this study, we classified true positives and false positives using the convolutional neural network (CNN) and deleted false positives.

3. Results

We used total lesion glycolysis (TLG), which is an index that reflects the size and spreads of tumors, and the center of gravity for evaluation. In order to calculate TLG, we calculated the Metabolic volume (MV) and the average value of SUV (SUV mean). TLG is calculated by equation (2).

$$TLG = MV \times SUVmean \qquad (2)$$

We decided that the tumor label that TLG ratio to the Gold standard (GS) was 10% to 200%, and the center of gravity was within the GS was the detection was successful. Also, the tumor label with a TLG ratio of 90% to 110% was successfully extracted and the labels that don't overlap with GS were as false positives. As a result of experiment that inputs PET images and Z-score images to CNN, 80.4% (270/336) of tumors were extracted by this method, and 18.9 FPs remained per case. Table 1 shows the number of successful detections and the number of successful extractions. The detection rate was high due to low physiological accumulation in the lung, but false positives were increased in the liver due to high physiological accumulation.

Table 1 the number of successful detections and extractions.

	The number of detection targets	The number of successful detections	The number of successful extractions
Right lung	31	30	19
Left lung	31	25	18
Liver	37	26	12
Abdomen	127	96	75
others	110	93	63
total	336	270	187

4. Conclusion

In this study, we used the image statistical analysis method to automatically detect tumors from PET/CT images. We used anatomical standardization to create Z-score images and determined the threshold of the tumors from the Z-score histogram. As a result of experiment that inputs PET images and Z-score images to CNN, 80.4% (270/336) of tumors were extracted by this method, and 18.9 FPs remained per case. We succeeded to delete 47.6 false positives per case using, but 11 true positives were falsely deleted. We could reduce false positives with similar detection accuracy using Z-score images to detect tumors in PET images.

Competing interests

There is no applicable item of competing interests.

References

[1] 寺本篤司:PET/CT 画像を用いたコンピュータ支援診断技術. 医用画像情報学会雑誌, vol.31, no.2, pp.vi-ix, 2014.

[2] A. Kumar, M. Fulham, D. Feng, et al.: Co-Learning Feature Fusion Maps from PET-CT Images of Lung Cancer. IEEE Transaction on Medical Imaging, vol.39, no.1, pp.204-217, 2019.

[3] F. L. Bookstein, thin-plate splines and the decomposition of deformations, IEEE Transaction on Pattern Analysis and Machine Intelligence, vol.11, no.6, pp.567-585, 1989.

Preliminary Study on Classification of Hands' Bone Marrow Edema Using X-ray Images

Dongping PAN[*1], Masahiro ODA[*1], Kou KATAYAMA[*2], Takanobu Okubo[*2], and Kensaku MORI[*1,3,4]

Abstract

Bone marrow edema is a symptom of early rheumatoid arthritis, which is reported as the main incentive to joint destructions. To diagnose early rheumatoid arthritis, bone marrow edema has been detected through Magnetic Resonance Imaging (MRI). However, the cost of taking MRI images is quite high, which puts a heavy economic burden on the patient. In this work, we are exploring the possibility of recognizing bone marrow edema through X-ray images which are fast and economic. We used a convolutional neural network consisting of convolution blocks and fully connected layers to classify whether bone marrow edema occurs or not. We also used an interpretation technique called the Grad-CAM for visual explanations. By using the neural network, we achieved a recall rate of 87% and a precision of 63% on this binary classification task.

Keywords：bone marrow edema, rheumatoid arthritis, CNN, classification

1. Introduction

Bone marrow edema (BME) is an inflammatory lesion in bone. For rheumatoid arthritis (RA) patients, bone marrow edema indicates the possibility of future joint destructions. Timely diagnosis and early treatment of BME is important for every rheumatism patient [1]. However, we can only recognize BME from Magnetic Resonance Imaging (MRI), which is very expensive for patients. Normally a confirmed patient should receive regular MRIs to check the condition of BME until the patient gets well. It becomes attractive to find a cheaper alternative to MRI. Jun Fukae et al. converted clinical information to a 2D image and used a convolutional neural network to classify RA or nonRA [2]. They are pioneers to use a convolutional neural network (CNN) for RA diagonosis, but they haven't used any information from medical images. From X-ray images, swelling in soft tissue can usually be oberseved from BME cases, which may help diagonosis. In this paper, we attempt to utilize CNN to classify X-ray images into BME or non-BME. Although it is unclear whether there is any evidence of bone marrow edema in X-ray images, we try to use machine learning based

*1 Graduate School of Informatics, Nagoya University〔Furou-cho, Chikusa-ku, Nagoya, 464-8601, Japan〕
e-mail:dpan@ mori.m.is.nagoya-u.ac.jp
*2 Katayama Orthopedic Rheumatology Clinic
*3 Information Technology Center, Nagoya University
*4 Research Center of Medical Bigdata, National Institute of Informatics

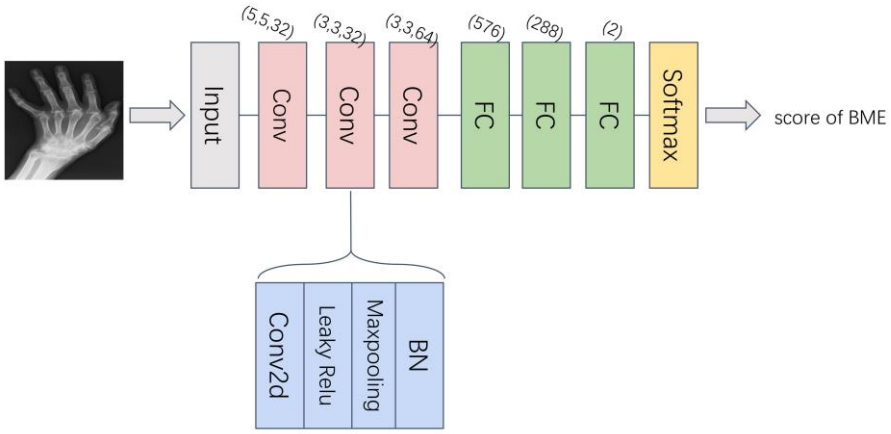


Fig. 1 The convolutional neural network design. A red block "Conv" means a convolutional block. It contains a 2D convolution layer, a leaky relu activation function, a maxpooling layer and a batch normalization layer. The numbers above each "Conv" block are (kernal size, kernal size, kernal number). A green block "FC" is a fully connection layer. The number above it is (neuron number). After the last Softmax layer, the output becomes the probability of BME which ranges from 0 to 1.

methods to explore its feasibility. Our goal is to develop an algorithm to predict whether there is BME or not from X-ray images and thus estimate the necessity of taking MRI for further diagnosis.

In this paper, we will show the convolutional neural network design, data preprocessing, assessment criteria in the methods section. Details of data and result of expeiments are introduced in experiment section. Visualization will be discussed in the visualization and discussion section.

2. Methods

In this work, we use a neural network consisted of convolutional layers and fully connected layers to classify X-ray images. The network design is demonstrated in Fig. 1. The input of the network is an image with a resolution of 224×224 pixels. The original images are saved in JPEG format and have a size of about 800×450 pixels. Because the sizes of the images are not aligned, the images are resized to 224×224 pixels. The convolutional neural network contains three convolution blocks and three fully connected layers. Each convolution block consists of a 2D convolutional layer, a leaky relu activation layer, a maxpooling layer, and a batch normalization layer [3]. The last layer is the softmax layer which can output the probability of each category. In this paper, the output is the socre of BME.

The neural network was trained for 200 epochs with an Adma optimizer. The loss function we used was binary crossentropy. We applied a weighted effect on the loss function, which allowed the minor category to play a more influential role in the training process. To be specific, if the number of BME samples is n times the number of non-BME samples, the loss from non-BME samples will be multiplied by the factor of n.



Fig. 2 Example of X-ray images. The left two images are BME cases. The right two images are non-BME case. Swelling in soft tissue and unclear margin of bone cortex are usually seen in early BME cases.

3. Experiment

The original data consists of 577 X-ray images of hands. 450 of them are free of BME and the remaining 127 images are containing BME. As we introduced previously, because BME is typically found with an MRI, the characteristic of BME in X-ray images is hard to recognize. Thus a medical doctor used the corresponding MRI diagnosis results from corresponding MRIs to label the X-ray images. In Fig. 2, there are four examples of BME and non-BME images. The left two images are BME cases. The right two images are non-BME case.

Regarding data split, 104 images including 79 non-BME images and 25 BME images are used as a hold-out test set. The rest of the images (473 images) are used as training data and validation data. Five fold cross-validation is used for these 473images. For each fold, there are about 378 images including 297 non-BME images and 81 BME images in the training set. There are about 95 images including 74 non-BME images and 21 BME images in the validation set. In order to fully utilize every image and unify the distribution of the training set and the validation set, the ratio of non-BME and BME is controlled to be the same which is about 3.66:1.

The augmentation of this dataset including flips and rotations. The data generator function from Keras is used so that every image selected will flip at a chance. The augmentation of flip includes a horizontal flip and a vertical flip. The augmentation of rotations is handled offline. We rotate each image for 5 and 10 degrees clockwise, and also counterclockwise. Through rotation augmentation, the data set is enlarged to five times its original number of images.

We used ROC curve, precision, and recall rate to evaluate the performance of our model. Since the dataset is imbalanced with the ratio of 3.66:1, it is meaningless to use overall accuracy to evaluate the model performance. By predicting every image as non-BME, we can get an overall accuracy of 78.5% which is misleading. Instead of overall accuracy, ROC curve, precision, and recall rate were applied to evaluate the performance of our neural network.

Here is a brief introduction to these evaluation criteria. Previously we mentioned that the final output is a score of BME which means the possibility of BME. For this binary classification task, we will select a threshold. If the output score exceeds the threshold, the image will be regarded as BME. To choose a threshold value, we calculated the f-

		Predicted label	
		BME	Non-BME
Real label	BME	True positive	False negative
	Non-BME	False positive	True negative

Fig. 3 Illustration of the confusion matrix. True positive/negative mean that the positive/negative prediction is correct. Similarly False positive/negative mean that the positive/negative prediction is wrong.

measure for thresholds of 0.1 to 0.9. We chose 0.4 whose f-measure is highest as the threshold of our model. To grasp the classification performance of every threshold, ROC curve is broadly used. ROC curve uses false positive rate as x axis and true positive rate as y axis. For every threshold, we can generate a confusion matrix in Fig. 3 and further calculate the true positive rate (TPR) as and false positive rate (FPR) as

$$\mathrm{TPR} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}, \tag{1}$$

$$\mathrm{FPR} = \frac{\mathrm{FP}}{\mathrm{FP} + \mathrm{TN}}, \tag{2}$$

where TP/TN stand for the number of true positive/negative which mean that the positive/negative prediction is correct. Similarly FP/FN stand for the number of false positive/negative which mean that the positive/negative prediction is wrong. These definitions are shown in Fig. 3.

Every threshold will be corresponding to a point in the ROC curve. While ROC curve indicates the overall capability of the network at every threshold, precision and recall rate at an optimal threshold can show the best performance of this

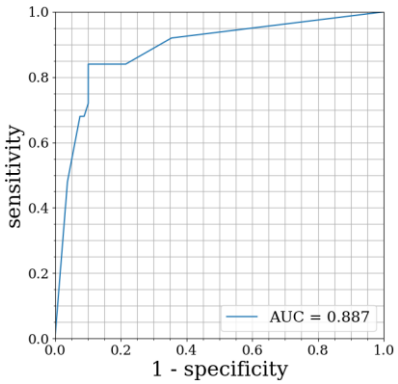


Fig. 4 ROC curve of fold five

Table 1 Precision and recall results (threshold = 0.4)

	Precision	Recall	AUC
Fold1	0.64	0.80	0.89
Fold2	0.62	0.92	0.91
Fold3	0.59	0.92	0.88
Fold4	0.59	0.88	0.89
Fold5	0.72	0.84	0.89
Average	0.63	0.87	0.89

model. The recall rate is the same as the true positive rate and precision is calculated as

$$\text{Precision} = \frac{\text{TP}}{\text{FP} + \text{TP}}, \tag{3}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \tag{4}$$

The five folds showed similar performance on the hold-out test set. There is a ROC curve of fold five showed in Fig. 4. AUC is the area under the ROC curve. As the result in Table 1, AUC which indicates the general performance of this model, ranged from 0.88 to 0.91. The average precision was 63% and the average recall rate was 87%.

In this experiment, the initialization seed will greatly influence the final result. For example, AUC can be reduced to 0.73 from 0.89 because of a different initialization seed. It perhaps results from the shortage of data, which can easily make the neural network drop into a local minimum.

4. Visualization and discussion

We also utilized Grad-cam [4] to visualize the result. The result of Grad-cam shows the importance of each part to the final prediction. Red regoins corresponds to high score for a category. In Fig. 5, red regions show a high score for BME. Yellow regions show a lower score than red regions. Green and blue regions show a lower score than yellow regions.

From the result in Fig. 5, we can find that there are a lot of red regions on the finger joints on the left image. It is a typical look for most true negative Grad-cam results. It can be interpreted that there are some characteristics on the finger joints showing this hand is a non-BME case. When the model predicts positive, no matter it's true or false, most red region is not concentrating on the hand itself. Some red region falls on the background and others on the radius and ulna.

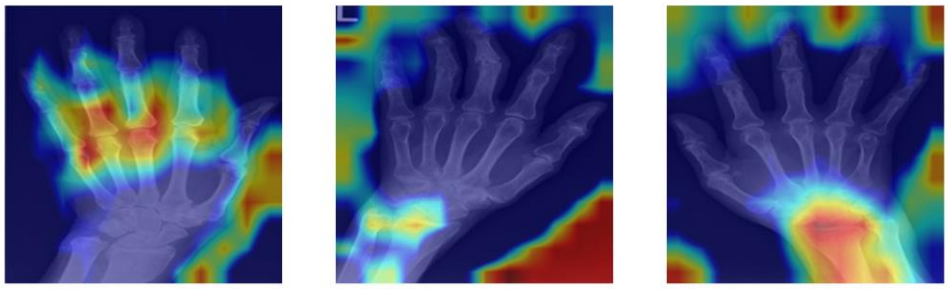


Fig. 5 Result of Grad-cam. Numbers in the parenthesises are the possibilities of BME. The middle case is unexpected because red region is the surrounding of the hand. The left and right cases may indicate the evidence for prediction.

The positive predictions seem to be based on areas of the radius and ulna.

5. Conclusions

It will be significant if we can judge BME from X-ray images. Although there is no medical evidence of the relation between X-ray images and BME, we used machine learning to exploring their relation. Using X-ray images and MRI-based labels, we design a binary classification task between BME and non-BME. The preliminary result is much better than a random guess. According to this result, there should be a certain difference between BME and non-BME images. If it's the characteristic of BME that domains this difference, our classification algorithm will be feasible for BME. Our future work is to justify the evidence of the predictions and improve performance.

Competing interests

None

Acknowledgement

Part of this research was supported by the MEXT/JSPS KAKENHI (26108006).

References

[1] Monti S, Montecucco C, Bugatti S, et al. Rheumatoid arthritis treatment: the earlier the better to prevent joint damage. RMD open, **1**(Suppl 1): e000057, 2015.

[2] Fukae J, Isobe M, Hattori T, et al. convolutional neural network for classification of two-dimensional array images generated from clinical information may support diagnosis of rheumatoid arthritis. Scientific reports, **10**: 1-7, 2020.

[3] Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167, 2015.

[4] Selvaraju R R, Cogswell M, Das A, et al. Grad-cam: Visual explanations from deep networks via gradient-based localization. Proceedings of the IEEE international conference on computer vision. 2017: 618-626.

屈折コントラスト法間の画像描写能比較の試み

大塚　楠平[*1]　砂口　尚輝[*2]　島雄　大介[*1, 3]

要旨

X 線は物質中で吸収・散乱を起こすほか，わずかに屈折する．この屈折の程度を画像コントラストとする手法を屈折コントラスト法という．X 線の屈折もスネルの法則に従い，物質間の屈折率の差と，異なる物質の境界へのX線の入射角という 2 つのパラメータに依存する．そのため，この両者を同時に評価する手法が望まれる．そこで我々は先行研究にて，自作ファントムを用い X 線の入射角ごとにコントラストを得るために必要な入射角の変位量（Δ）を測定する，という評価法を提案した．本研究ではこの評価法を 2 種類の屈折コントラスト法に適用した．その結果，これら 2 種類の画像法のΔには 2 倍程度の差がある，と評価することができた．この評価法はファントムの投影像を 1 枚撮影するだけで行える簡便なものであり，各種の屈折コントラスト法の相対評価に有用であると考える．

キーワード：X 線，屈折コントラスト，シンクロトロン，画像評価

1. はじめに

X 線は物質に入射した際，その内部で吸収や散乱といった相互作用を起こすほか，わずかに屈折する．この X 線の屈折の程度を画像コントラストとするイメージング法を屈折コントラスト法という．X 線の屈折は可視光などと同様に，ある物質から屈折率の異なる物質へ入射した際に起こり，その屈折の程度は以下のスネルの法則に従う．

$$n_2 = \frac{\sin\theta_1}{\sin\theta_2} n_1 \quad (1)$$

$n_{1,2}$:各物質の相対屈折率

$\theta_{1,2}$：入射角および屈折角

*1 北海道科学大学大学院保健医療学研究科医療技術学専攻

〔〒006-8585 札幌市手稲区前田 7 条 15 丁目 4-1〕

e-mail: 9193301@hus.ac.jp

*2 名古屋大学大学院医学系研究科総合保健学専攻

*3 北海道科学大学保健医療学部診療放射線学科

投稿受付：2020 年 8 月 5 日

つまり，屈折コントラスト法の画像コントラストは物質間の屈折率の差（≒密度差）と，物質境界への X 線の入射角という 2 つのパラメータに依存することになる．そのため屈折コントラスト法の画像評価はこの両者を同時に評価可能なものが望ましいが，その手法はいまだ確立されていない．そこで我々はアクリル製の自作ファントムを用いた簡便な評価法を提案するとともに，実際に 2 種類の屈折コントラスト法に適用し，比較したため結果を報告する．

2. 撮影法と評価法

1）撮影法

本研究では各種ある屈折コントラスト法のうち，X 線回折強調法（X-ray Diffraction Enhanced Imaging：DEI）[1]と X 線暗視野法（X-ray Dark Field Imaging：DFI）[2]の 2 種類を評価した．この 2 種類の光学系は似通っているが，屈折の程度を X 線強度へと変換するアナライザーの種類が異なる．図 1, 2 にそれぞれの概要図を示し，各光学系の大まかな概要を説明する．これらは共に光源として高い輝度と平行度を持つシンクロトロン放射光を

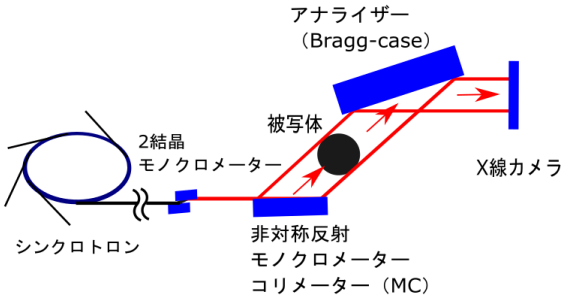


図 1 DEI 光学系の模式図

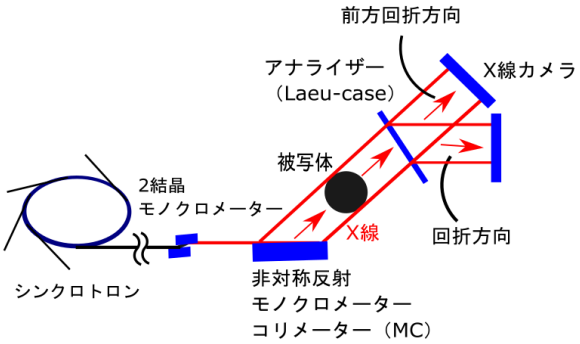


図 2 DFI 光学系の模式図

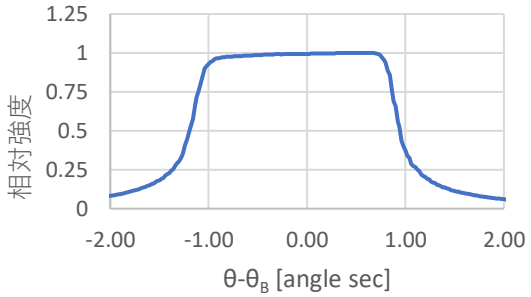


図 3 DEI において得られた回折強度曲線

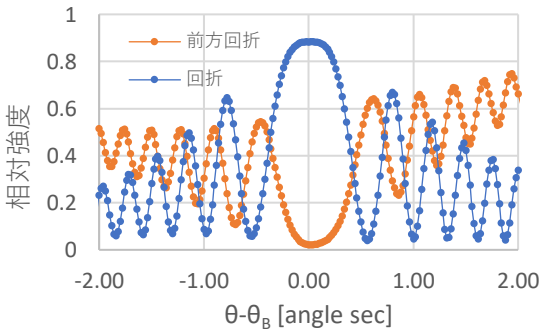


図 4 DFI において得られた回折強度曲線

用いている．放射光は元々白色光であるが，2 結晶モノクロメーターによって単色化される．この時点の単色 X 線ビーム横幅は数 mm 程度と非常に細いが，非対称モノクロメーターコリメーター（MC）に入射させることで，さらに平行度を高めつつビームサイズを数 cm まで拡大することができる．拡大された X 線は被写体を透過したのち，アナライザーに入射し，変換された回折強度分布を X 線カメラで測定する．

各手法で用いるアナライザーはどちらも Si 単結晶だが，DEI ではある程度厚みをもった結晶による Bragg 反射（Bragg-case）を利用し，DFI ではごく薄い結晶による透過（Laue-case）を利用する．Bragg-case ではアナライザーの表面を入射 X 線に対してブラッグ角 θ_B 近傍で配置することで X 線ビームが $2\theta_B$ 方向へ回折し，Laue-case ではアナライザー結晶表面に対し垂直に設定された回折面を，入射 X 線ビームに対し θ_B 近傍で配置することで $2\theta_B$ 方向への回折ビームと透過方向の前方回折ビームに分割される．今回用いたアナライザーはそれぞれ DEI では Si（220）結晶，DFI では厚さ 356 μm の Si（111）結晶であり，MC は各種法のアナライザーと同じ回折面に設定された Si 結晶を用いた．DEI と DFI それぞれにおけるアナライザー透過後の X 線回折強度分布は図 3，4 のようになり，これらの曲線の勾配を利用することで，被写体で生じたわずかな屈折を X 線強度に変換し，最終的に画像コントラストとすることができる．

2）撮影実験

撮影は高エネルギー加速器研究機構の放射光施設（BL-14B）に構築された光学系にて行い，入射 X 線エネルギーは 19.8 keV とした．

本研究ではアクリル製の自作ファントムを評価に用いた．このファントムはアクリル製の水槽内中央に同じくアクリル製の円柱を配置したものである．アクリルの密度 d は 1.11 g/cm^3, 円柱の直径は 10 mm である．撮影は この水槽内に液体を入れた状態で行うが，その際，液体の種類を変更することでアクリル

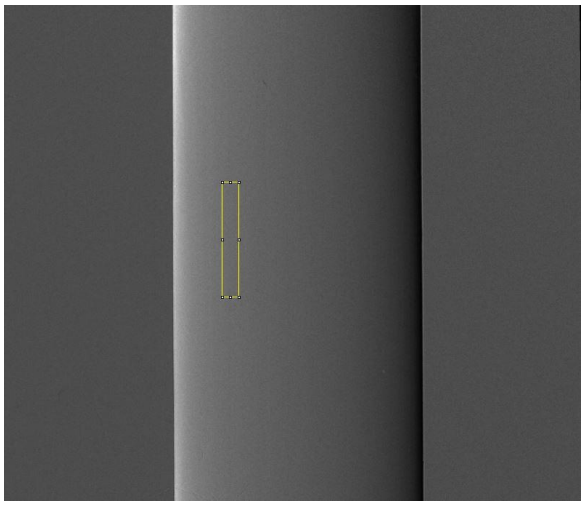


図 5 アクリル製自作ファントムの屈折コントラスト像（評価画像）と ROI の配置例

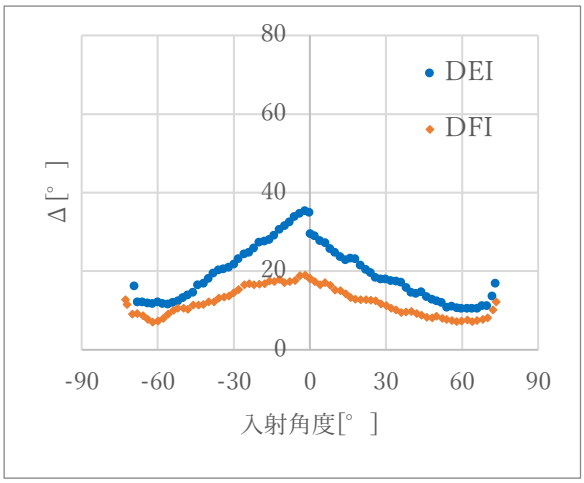


図 6 DEI と DFI の比較（水の場合）

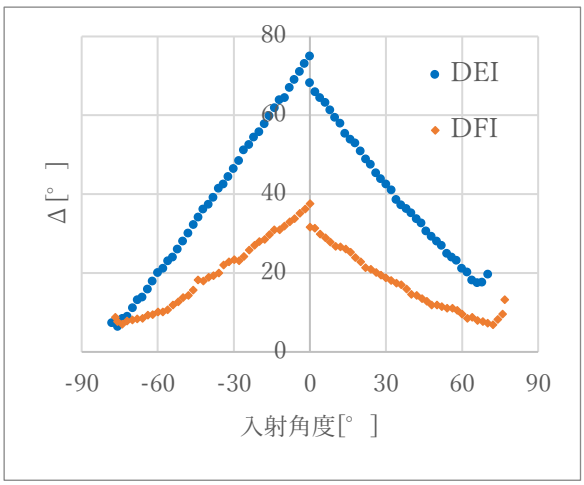


図 7 DEI と DFI の比較（エチレングリコールの場合）

と液体の密度差，つまり相対屈折率の差を任意に設定できる．本研究では水（d=1.00 g/cm^3）とエチレングリコール（d=1.18 g/cm^3）の 2 種類の液体を用いた．また，このファントムは円柱部分を撮影することで X 線のアクリルへの入射角を連続的に変化させることができ，入射角ごとにコントラストを観察できる．

3）評価法

撮影した画像の処理および解析は画像解析ソフト Image J を用いて行った．ファントムの円柱部分を撮影した画像を，液体のみの部分を撮影した画像で除算することで評価画像を作成した．この評価画像の円柱部分に対し矩形の ROI（縦 200 pixel×横 20 pixel）を設定した．その様子を図 5 に示す．設定した ROI を円柱の端から端へ 1 画素ごとに走査し，それぞれの位置での ROI 内部の画素値の平均値および標準偏差を測定した．この時，ある位置 A（入射角 θ_A，平均画素値 V_A，標準偏差 SD_A）と，ある位置 B（θ_B，V_B）を比較した際，以下の式（2）を満たす場合を“A に対し B でコントラストが得られた”と定義した．

$$|V_A - V_B| > 2SD_A \quad (2)$$

さらに入射角ごとに，コントラストを得るのに必要な入射角の変位量 Δ

$$\Delta = |\theta_A - \theta_B| \quad (3)$$

を算出し，グラフ化した．

3. 結果と考察

水槽内を満たす液体が水の場合とエチレングリコールの場合，それぞれにおける DEI と DFI での Δ の測定結果を図 6，7 に示す．どちらにおいても DEI のグラフは DFI のものより上に来ており，Δ の値は 2 倍程度となっている．このことから，同じ屈折コントラスト法の中でも DFI より DEI の方がコントラストがつきにくいことがわかる．

また，液体の違いに着目すると，どちらの撮像法においても，アクリルとの密度差の小さいエチレングリコールの方が Δ の値が大きくなっている様子が示されており，密度差の違いによるコントラストのつき方の変化につ

いても評価できていると言える．

グラフの両端においてΔが大きくなっているのは，円柱の辺縁に行くにつれ入射角の差のステップが小さくできなくなることに起因すると考えられる．今回Δをある2つの画素間の入射角の差として定義したため，ある画素において取りうる最小のΔの値は，その画素に隣接する画素との入射角の差となる．本来その差より小さい入射角度変位でコントラストが得られていたとしてもその角度を評価できず，Δを過大評価してしまったのだと考えられる．そのため入射角度の大きい場合は評価できない．しかし，屈折コントラスト法において議論されるのは，入射角度変位が小さくコントラストがつきにくい部分であるため，実際に評価を行う場合問題にならないと考える．

4. まとめ

本評価法は各種屈折コントラスト法の撮影装置においてファントムを配置し，投影像を1枚撮影するだけで行える簡便な方法であり，屈折率の差と入射角の両者を同時に評価可能である．この手法は入射角度の大きい辺縁部分の評価には使用できないものの，各種屈折コントラスト法間の画像描写能の相対評価には十分に有用であると考える．

利益相反の有無

なし

文　献

［1］ Thomlinson W, Chapman D, Zhong Z, et al.: Diffraction Enhanced X-ray Imaging. Phys Med Biol **42**: 2015-2025, 1997

［2］ Ando M, Maksimenko A, Sugiyama H, et al.: Simple X-ray dark-and bright-field imaging using achromatic Laue optics.Jpn J Appl Phys 2002;41:L1016–8.

Trial of Comparing Refraction-Contrast Imaging Methods on Depiction Ability

Nampei OTSUKA[*1], Naoki SUNAGUCHI[*2], Daisuke SHIMAO[*1,3]

*1 Hokkaido University of Science Graduate School

*2 Nagoya University

*3 Hokkaido University of Science

There are such interactions of X-rays with substances as absorption, scattering and refraction. Imaging techniques based on the refraction of X-rays are called refraction-contrast imaging. The refraction of X-rays follows Snell`s law, so the image contrast depends on two parameters; one is difference of refractive indices between substances, the other is angle of incidence to boundary of substances. Therefore, a method to evaluate image contrast due to the differences of refractive indices and those of the incident angles at the same time is desirable. Thus, we proposed an evaluation method that compares "necessary displacement of an incident angle to gain contrast at each angle of incidence" in previous research. In this study, we apply this evaluation method for two kinds of refraction-contrast imaging. As a result, we could evaluate the difference of the depiction ability between the two imaging methods quantitatively. This evaluation method will be an easy and useful tool in relative evaluation on refraction-contrast imaging.

（151 words）

Key words: Template, Format　(maximum　5 words)

X-rays, refraction-contrast imaging, synchrotron, image evaluation

1.5 T MRI における定量的磁化率マッピング（QSM）の画質評価

高橋 哲彦[*1*2] 山路 菜月[*1] 林 則夫[*1*2] 竹内 友一[*2] 小倉 明夫[*1*2]

要旨

診断画像の定量化や読影支援技術の開発や利用において，元画像の撮像条件と定量解析結果の関連を把握することは重要である．本研究では，定量的磁化率マッピング(QSM)を汎用的な 1.5TMRI 装置で実用することを目標とし，QSM において，ファントムおよび健常ヒト脳の画質(信号値・ノイズ)と撮像パラメータの関係を数値計算と実測で比較検討した．

キーワード：MRI，QSM，磁化率，画質

1. はじめに

診断画像の定量化や読影支援技術の開発や利用において，元画像の撮像条件と定量解析結果の関連を把握することは重要である．
本研究では，定量的磁化率マッピング(quantitative susceptibility mapping: QSM)を汎用的な 1.5T MRI (magnetic resonance imaging)装置で実用することを目標とし, QSM において, ファントムおよび健常ヒト脳の画質(信号値･ノイズ)と撮像パラメータの関係を数値計算と実測で比較検討した．

2. 方法

・QSM 画像

複素 MRI 画像(絶対値，位相)から QSM を得る計算は，least squares estimation with adaptive edge-preserving filtering (LSE with AEFP) 法を用いた. この方法は, regularization を用いずに edge preserving filter を用いて微細構造を保持する．[1]

・数値計算

QSM の信号値は，被検体の磁化率[ppb](定量値)で，一方，ノイズは撮像シーケンスの *TR*,*TE*, 被検体の *T1*,*T2* 値から得られる．
位相信号の SN 比(SNR_θ)は，周波数シフト f，エコー時間 t，元画像の信号値 S_0，元画像のノイズ σ_0 のとき[2]，(1)式で示される．

$$SNR_\theta = \frac{2\pi f t S_0}{\sigma_0} \quad (1).$$

ここで単位時間当たりの位相$2\pi f$は，

$$2\pi f = \gamma B_0 \left[\chi(r) * \left(\frac{1}{4\pi}\cdot\frac{3\cos^2\theta - 1}{r^3}\right)\right] \quad (2)$$

(*は convolution 演算, r は位置)
であり[3]，磁化率 $\chi(r)$に比例する. (1)式から，

$$SNR_\theta = \frac{2\pi f}{\sigma_0 / S_0 t} \quad (3).$$

*1 群馬県立県民健康科学大学　診療放射線学部
〔〒371-0052 前橋市上沖町 323-1〕
e-mail: ttakahashi@gchs.ac.jp
*2 群馬県立県民健康科学大学大学院 診療放射線学研究科
投稿受付：2020 年 8 月 5 日

したがって位相画像のノイズ σ_θ (∝QSM のノイズ)は,

$$\sigma_\theta \propto \frac{1}{S_0 t} \qquad (4).$$

ここで, t= *TE*.

$$S_0 = \frac{\exp\left(-\frac{TE}{T2}\right)\left(1-\exp\left(-\frac{TR}{T1}\right)\right)\sin\alpha}{1-\cos\alpha\left(\exp\left(-\frac{TR}{T1}\right)\right)}.$$

・実験

MRI 画像の取得は，1.5T MRI 装置(Philips 社製)を用いた．QSM 解析ソフトウエアは，日立製作所製ソフトウエアを用いた．ファントムは，円柱の液体ファントム(*T1*=355 ms, *T2*＝309 ms)を用いた．ヒト被検体は 3 例の健常例 (20 歳代)で, 学内で倫理審査の承認を得た．撮像体位は仰臥位，撮像断面は AXI (スライスエンコード方向:HF)である．撮像シーケンスは 3DGrE (Flow compensation 付)とし，撮像条件は *TR* と *TE* を変えた. 得られた複素画像を使って，QSM 画像を作成した．画像解析は, ImageJ を用い ROI の信号値と標準偏差を求めた.

3. 結果

QSM ノイズの TE 依存性の計算結果(図 1)から *TE* を短くするとノイズが増えた(*TR* shortest, *TR* fix)．また *T2* が短くなると *TE* が延長したときに再びノイズが増加した.

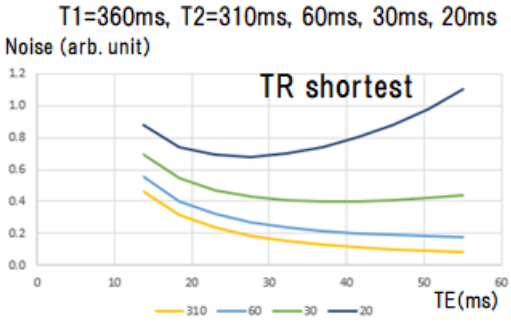


図 1 計算結果

図 2 は QSM ノイズの計算値と実測平均値を重ね合わせて表示した(実測値を *TE*=35ms でスケーリング)．図から，計算と実測の傾向が一致していると分かった.

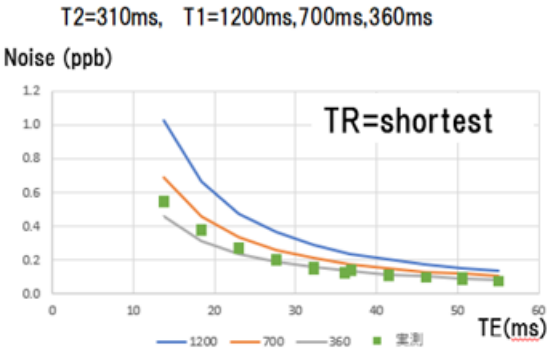


図 2 計算結果と実測値

図 3 は，ヒト QSM 画像例である。表 1 は，3 例の信号平均値である。淡蒼球>尾状核＞被殻，の関係は，3T での報告値[4]と一致した.

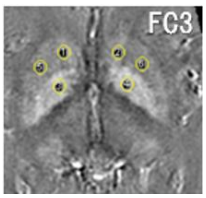


表 1 ヒト磁化率(実測値)

	平均値	分散	文献値[3]
尾状核 CN	31.55	2.29	32
淡蒼球 GP	108.50	4.42	75
被殻 PUT	21.96	3.92	20

図 3 ヒト撮像結果

4. まとめ

MRI の定量画像である QSM を 1.5T MRI 装置で実用することを目標とし， QSM において，ファントムおよびヒト脳の磁化率とノイズと撮像パラメータの関係を実測し，ノイズについては計算と，磁化率については 3T 先行事例と矛盾がないと確認した.

謝辞　研究用 QSM 解析ソフトウエアを貸与いただいた(株)日立製作所に感謝致します.

利益相反の有無　なし

文　献

[1] Shirai, T, et.al.: *Proc. ISMRM*, p3319, 2015

[2] Wu, B, et.al.: Neuroimage **59**, pp297-305, 2012

[3] Buch, S, et.al.: Magn Reason Med. **73**, pp 2185-2194 (2015)

[4] Gong NJ, et.al.: NMR Biomed. **28**, pp 1267-1274, 2015

Image Quality Evaluation of Quantitative Susceptibility Map (QSM) on 1.5T MRI

Tetsuhiko TAKAHASHI[*1*2], Natsuki YAMAJI[*1], Norio HAYASHI[*1*2], Tomokazu TAKEUCHI[*2], Akio OGURA[*1*2]

*1 Gunma Prefectural College of Health Sciences, School of Radiological Technology

*2 Gunma Prefectural College of Health Sciences, Graduate School of Radiological Technology

Usage of diagnostic quantitative images and/or development of diagnostic imaging assist technology seeks adequate understanding of relation between original imaging parameters and obtained quantitative images. In this study, we aim that quantitative susceptibility mapping (QSM) on 1.5T MRI apparatus, widely used in routine clinical situation, becomes into diagnostic utilization. And we evaluate imaging parameter and QSM image quality (susceptibility and noise) by using numerical calculation and experimental study of phantom and of brain imaging with healthy volunteers.

Key words: MRI, QSM, susceptibility, image quality

深層学習を用いたマルチピンホール SPECT システムの画像再構成

紺野　宏彰[*1]　大越　迪[*1]　村田　一心[*2]　尾川　浩一[*2]

要旨

本研究では、マルチピンホール SPECT システムの画像再構成における深層学習の有効性を検証した．静止型マルチピンホール SPECT システムはガンマカメラを回転させずに投影データを収集するため、放射性医薬品の動態情報を取得できる．その画像再構成には通常、最尤推定・期待値最大化法が用いられるが、必要な検出確率を厳密に求めるには現実的でない計算コストがかかる．その補正のため、7 rays 法を用いて近似解を求めることが多いが、特にピンホール径が大きいときには画質の劣化が著しい．そこで本研究では、深層学習を用いた画像再構成手法を提案する．実際の医用画像を用いたシミュレーションを行ったところ、高い空間分解能の再構成画像を取得できることがわかった．これは同システムの画像再構成において深層学習が有効であることを示す．

キーワード：SPECT，画像再構成，深層学習

1. はじめに

単一光子放出コンピュータ断層撮影（SPECT）システムとは，現代の核医学における代表的なシステムであり，放射性医薬品を用いた断層画像を取得できる．一般的な SPECT システムは患者の周囲に検出器を回転させてデータを収集する．しかしこの方法はデータ収集に時間がかかるため，患者の負担が大きくなるだけでなく，モーションアーチファクトが発生しやすい．また，動体情報も取得できない．そこで近年，検出器を固定してデータ収集を行う静止型 SPECT システムが開発されている．同システムでは，マルチピンホールコリメータを使用する．すなわち，コリメータに複数のピンホールを配置することで、一つの検出器で複数の投影方向の情報を取得する．

しかし静止型 SPECT システムでは，ピンホール径が有限の大きさを持つことに起因する空間分解能劣化の影響を強く受ける．これは原理的には，最尤推定・期待値最大化（MLEM）法[1]による画像再構成の際，厳密な検出確率を適用することで改善できるが，それには現実的でない計算コストがかかる．一方，近似解を推定する 7 rays 法[2]では，特にピンホール径が大きい時に十分な補正を行うことができない．

そこで本研究では，深層学習を用いた画像再構成方法を提案する．深層学習は近年，医用画像の雑音除去や超解像，画像再構成などにおいて，視覚的・数値的に画質の高い画像を取得できる観点から注目されている[3][4]．しかし，マ

*1 法政大学大学院理工学研究科
〔〒184-8584 小金井市梶野町 3-7-2〕
e-mail: hiroaki.konno.4z@stu.hosei.ac.jp
*2 法政大学理工学部
〔〒184-8584 小金井市梶野町 3-7-2〕

ルチピンホール SPECT システムの画像再構成に深層学習手法を適用した例は，著者らの知る限りでは存在しない．

そこで本研究では，深層学習を用いたマルチピンホール SPECT システムの画像再構成手法を提案する．投影画像を入力すると再構成画像を出力する CNN ベースの学習ネットワークを構築した．このネットワークの有効性を検証するため，実際の医用画像を用いた画像再構成シミュレーションを行ったところ，高い空間分解能の再構成画像を取得できることがわかった．この実験結果は同システムの画像再構成において深層学習が有効であることを示している．

2. シミュレーション方法

シミュレーションには、The Cancer Imaging Archive (TCIA) から取得した画像を用いた (図 1 左)．本研究は SPECT の画像再構成を目的としているが，より複雑な画像の再構成性能の調査をするため，ここでは CT 画像を適用した．取得した画像を 64×64[pixel]にリサイズし, これらの画像に対し、先行研究[5]で用いた静止型 SPECT システムに合わせたジオメトリで投影画像を取得した．ターゲット中心からコリメータまでの距離は 25 cm，コリメータから検出器までの距離は 7.5 cm とした．検出器は 6 個とし，各コリメータは 3 つの水平に並んだピンホールを有する．本研究では簡単のため，2 次元画像での投影データを取得した (図 1 右).

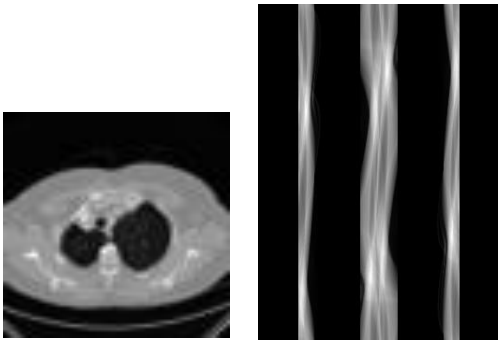


図 1 原画像及び，原画像から作成したサイノグラムの例．左：原画像，右：サイノグラム

このようにして得られた投影画像を再構成するため，図 2 に示すようなネットワークを構築した．ネットワークは畳み込み層 (Conv)，活性化関数 (ReLU)，Batch Normalization (BN)で構成されており，畳み込み層のカーネルサイズは 3×3× 64 で統一し，全体の層数は 20 とした．細かい画像サイズの調整は，畳み込み第 2 層と第 3 層のストライドを(2,2)として圧縮し，第 4 層から第 16 層までのゼロパディングを省くことで行った．学習には 100 枚ずつの原画像と投影画像を使用し, 90 枚を学習に, 10 枚をテストデータとして用いた．バッチサイズは 16，学習回数は 2000, パラメータ更新には Adam を用いた．

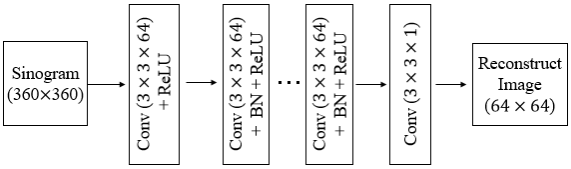


図 2 ネットワーク構成図．投影画像 (Sinogram)を入力し，畳み込み層を通して再構成画像を出力する．各層は畳み込み(Conv)とバッチ規格化 (BN)，活性化関数(ReLU)で構成されている．

3. シミュレーション結果

本手法により再構成した画像の例を図 3 に示す. 得られた再構成画像 (右)は元画像 (左)と比べ，遜色ない画質を有し，高い空間分解能も維持していることがわかる．テストデータ 10 枚における再構成画像の原画像に対する RMSE の平均値は 0.0068 であり, 定量的にも高画質な再構成画像を取得できたと言える．

4. 結論

本研究では，マルチピンホール SPECT システムにおける深層学習を用いた画像再構成手法を提案した．同手法による画像再構成シミュレーションを行った結果，定性的・定量的に高画質な再構成画像を取得することができた．これは同システムの画像再構成において深層学習が有効であることを示す．

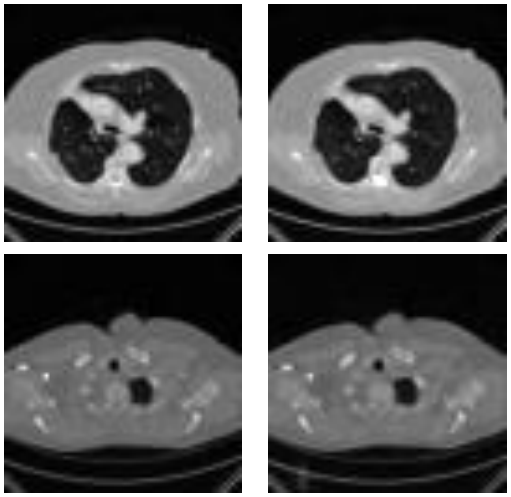


図 3 原画像 (左) と再構成画像 (左) の比較例.

利益相反の有無

なし

文　献

［1］ Shepp L. A, Vardi Y, Maximum likelihood reconstruction for emission tomography. IEEE Trans. Med. Imag., 1, pp. 113-122, 1982

［2］ Andriy Andreyev, et al.: Pinhole SPECT reconstruction using blobs and resolution recovery, IEEE Trans. Nucl. Sci., vol. 53, no. 5, 2719-2728, 2006

［3］ H. Gupta, K. H. Jin, H. Q. Nguyen, M. T. McCann et al., "CNN-Based Projected Gradient Descent for Consistent CT Image Reconstruction," in IEEE Trans. Med. Imag., vol. 37, no. 6, pp. 1440-1453, 2018.

［4］ C. Chrysostomou, L. Koutsantonis, C. Lemesios and C. N. Papanicolas, "SPECT Imaging Reconstruction Method Based on Deep Convolutional Neural Network," 2019 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), Manchester, United Kingdom, 2019, pp. 1-4

［5］ Y. Fujishiro , et al., “List Mode Image Reconstruction With a Multi-pinhole Triple Head SPECT System, ”IEEE Nuclear Science Symposium and Medical Imaging Conference Proceedings(NSS/MIC), 2018

Multi-pinhole SPECT system Reconstruction Method Based on Deep Learning

Hiroaki KONNO[*1], Michi OKOSHI[*1], Kazumi MURATA[*2], Koichi OGAWA[*2]

*1 Graduate School of Science and Engineering, Hosei University

*2 Faculty of Science and Engineering, Hosei University

A multi-detector SPECT system equipped with multi-pinhole collimators has great advantages compared with a conventional one owing to a data acquisition without rotating a gamma camera. It reduces the data acquisition time and obtains dynamic information of tracers. However, it has a problem in the image reconstruction; reconstructed images are highly degraded when the size of a pinhole is large. This is caused by the exact detection probability could not be calculated and a correction method such as a 7-rays method could not compensate the influence of the large pinhole size. Hence, a new image reconstruction method is required for the case with a large pinhole collimator. So, we propose a deep-learning based image-reconstruction method for a multi-pinhole static SPECT system. Our network successfully reconstructed the images and the spatial resolution was significantly higher than those with a conventional method. Our results indicate that the machine-learning based method is effective for the image reconstruction of a multi-pinhole static SPECT system.

Key words: SPECT, Image Reconstruction, Deep Learning

OCT を用いたマウス生殖機能評価に関する検討

渡部 裕輝[*1]，高倉 啓[*1]，黒谷 玲子[*1]，阿部 宏之[*1]

要旨

光干渉断層画像化法(OCT，Optical Coherence Tomography)は，生体に対して非接触，非侵襲であり，高い空間分解能を有するイメージングシステムである．すでに眼科用医療機器として不可欠な存在となっている．また網膜内の血流を OCT の信号から抽出するドップラー OCT からの派生した OCT Angiography がよく利用されている．我々は，これらの OCT 技術を用いて，新たなに生殖分野での応用を検討している．卵巣内に存在する卵胞の大きさや数を把握することは，卵巣機能の評価する上で極めて重要である．我々は OCT を用いることで，超音波画像では観察できないマウス卵巣に存在する直径 50 μm 程度の前胞状卵胞の非侵襲イメージングに成功した．また週齢の異なるマウス精巣の OCT 計測では，ドップラー OCT 画像に精巣網に明らかな違いが確認できた．

キーワード：光干渉断層画像化法，ドップラー OCT，生殖機能

1. はじめに

光コヒーレンストモグラフィ(OCT, Optical Coherence Tomography)は，十数μm の高い空間分解能を有し，生体などの高散乱体の断層画像を非侵襲に計測することができ，様々な応用研究が進められている．すでに眼科用医療機器として不可欠な存在となっている．また網膜内の血流を OCT の信号から抽出するドップラー OCT からの派生した OCT Angiography がよく利用されている．

我々は，これらの OCT 技術を用いて，新たなに生殖分野での応用を検討している．哺乳動物の卵巣は，発達段階の異なる多数の卵胞および卵母細胞をもっており，卵胞数は胎児期に最も多く個体の成長に伴って減少することから，卵胞の数は卵巣機能評価の重要な指標となる．現在，卵胞の解析は超音波画像診断(エコー)により行われているが，超音波画像より空間分解能が高い OCT によりマウス卵巣の画像観察を行った[1]．またマウス精巣の OCT 計測では，精巣上体へ精子を運ぶ輸出管の根元付近にある精巣内の構造である精巣網の観察を行い，週齢による違いを評価した．

*1 山形大学理工学研究科
〔〒992-8510 米沢市城南 4-3-16〕
e-mail: ywata@yz.yamagata-u.ac.jp

2. OCT システム

使用した OCT システムの概略を図 1 に示す．中心波長 1322 nm，半値全幅 102 nm の SLD 光源を使用し，スペクトル干渉信号は，InGaAs ラインカメラ(1024 pixels, 91 kHz)で検出した．このシステムの深さ測定範囲は 2.9 mm，深さ方向分解能は 7.5 μm，横方向分解能は 21.3 μm である．

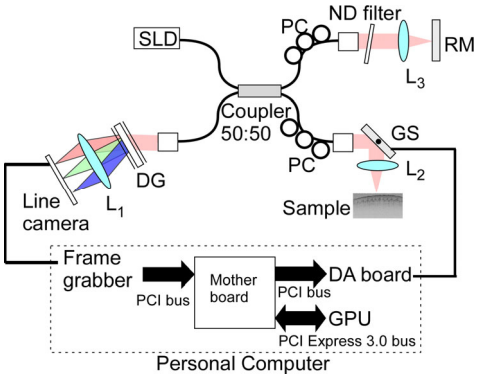


図 1 OCT システムの概要

また生体内の動きを可視化するため，同一箇所で OCT 画像を 2 枚以上取得し，差の二乗を計算することで，動いている部分を強調し

た画像を抽出した．以後，この画像をドップラー OCT 画像と呼ぶことにする．

3．測定結果

卵巣は B6C3F1 系雌マウス(25.5 日齢)から採取した. 採取した卵巣をリン酸緩衝液(PBS)で湿潤状態を保ち，OCT により卵巣の画像観察を行った．なお動物実験は山形大学動物実験委員会の承認のもと実施した．卵胞の発育過程は原始卵胞(primordial follicle)，一次卵胞(primary follicle)，二次卵胞(secondary follicle)，前胞状卵胞(preantral follicle)，および胞状卵胞(antral follicle)に分類される．胞状卵胞は卵母細胞に隣接する流体が充填された卵胞腔を形成する．マウス卵巣での卵胞の大きさは，原始卵胞(直径約 20 μm)，1 次卵胞(直径 50～130 μm)，2 次卵胞(直径 290～310 μm)，胞状卵胞(直径 400～500 μm)へと発育を続ける．

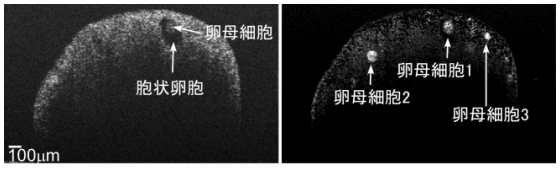


図 2　卵巣の OCT 画像(左)　ドップラー OCT 画像(右)　画像化領域：2.2mm × 1.0mm (横× 深さ)

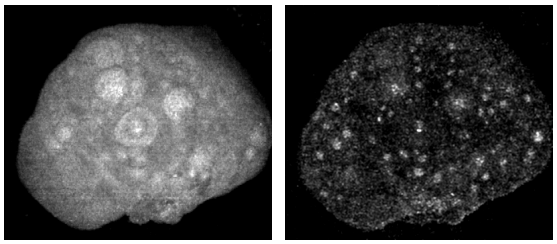


図 3　深さ方向へ平均した卵巣の OCT 画像(左)　最大値を投影したドップラー OCT 画像(右)

図 2 は OCT 画像とドップラー OCT 画像でサイズは 2.2mm × 1.0mm (横× 深さ)である．OCT 画像では，卵胞腔を形成する胞状卵胞がはっきりと確認できた. ドップラー OCT 画像では，卵母細胞の動きがあり，その部分を強調した画像が得られた．図 3 は 3 次元測定したボリュームデータから深さ方向へ平均した OCT 画像と最大値を投影したドップラー OCT 画像である．これらの測定結果から，大きい胞状卵胞の直径は約 300 μm であり，直径 50 μm 程度の前胞状卵胞とみられる構造が多数観察された

精巣は，C57BL/6N 系雄マウス(7 週齢)から採取し，精巣網付近で画像観察を行った．図 4 は深さ方向へ平均した OCT 画像と最大値を投影したドップラー OCT 画像で, 測定範囲は 7.0mm × 4.0mm である．ここで画像の右側が精巣輸出管である. ドップラー OCT 画像では液体の流れがあるところが強調されるので，輸出管や精巣網内にその様子が確認できた. 4 週齢のマウスでは輸出管で同様の様子が確認できたが，精巣網では確認できなかった．

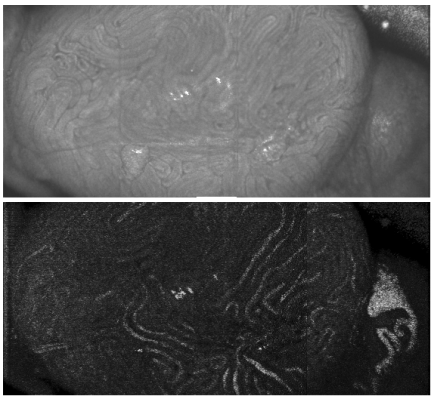


図 4　深さ方向へ平均した精巣網の OCT 画像(上)　最大値を投影したドップラー OCT 画像(下)

4．まとめ

OCT 技術を用いて，マウス生殖機能評価の検討を行った．卵巣，精巣の OCT 計測において，動いている部分を可視化することが発育過程を評価するのに有効であることがわかった．

利益相反の有無

なし

文　献

［1］　Y. Watanabe, K. Takakura, R. Kurotani, and H. Abe :Optical coherence tomography imaging for analysis of follicular development in ovarian tissue Applied Optics **54**, 6111-6115, 2015

パルスレーザーダイオードを用いた光音響イメージングシステムの開発

甲田　梨緒*1　渡部　裕輝*1

要旨

光音響イメージングは，光学イメージングの分解能と超音波イメージングの測定深度を併せ持った画像化技術で，生体内部の情報を非侵襲に取得する方法として注目されている．一般的に光源はQスイッチパルスレーザーが使用されるが，装置が大きく高価である．そのため，近年ではレーザーダイオード(LD)や発光ダイオード(LED)によるコンパクト化および低コスト化の研究が行われている．そこで我々は，レーザーダイオードを用いた低コストな光音響イメージングシステムを開発した．このシステムでは，波長905nmのLD光をパルス幅143ns，繰り返し周波数700Hzで照射し，中心周波数5MHz，比帯域幅61.6%の超音波プローブを用いて光音響画像を取得した．

キーワード：光音響イメージング，レーザーダイオード，低コスト化

1. はじめに

光音響イメージング(Photoacoustic（PA) imaging)は，光学イメージングの分解能と超音波イメージングの測定深度を併せ持った画像化技術で，生体内部の情報を非侵襲に取得する方法として注目されている．その原理は，ナノ秒パルス光を照射し，物質の光吸収によって発生した超音波を検出する[1]. 一般的に光源はQスイッチパルスレーザーが使用されるが，装置が大きく高価である．そのため，近年ではレーザーダイオード(LD)や発光ダイオード(LED)によるコンパクト化および低コスト化の研究が行われている[2].

そこで我々は，レーザーダイオードを用いた低コストな光音響イメージングシステムを開発した．このシステムでは，2軸自動ステージによるサンプル走査またはガルバノスキャナによるレーザー走査を選択できるように設計し，取得した画像を評価した．

*1 山形大学大学院理工学研究科応用生命システム工学専攻
〔〒992-8510 山形県米沢市城南4-3-16〕
e-mail: thh61146@st.yamagata-u.ac.jp

2. 光音響イメージングシステム

開発した光音響イメージングシステムの概略を図1に示す．繰り返し周波数700Hz，デューティ比0.01%の方形波を用いてレーザーダイオード(LD)(波長905nm，SPL UL90AT08)からパルス光を出力した．このときパルス幅は143nsである．出射した光は，コリメートされ，ガルバノスキャナ(GS)を通り，対物レンズ(L, f=35mm)により集光されてサンプルに照射する．サンプルの光吸収によって発生した超音波は，サンプルに対して45°に配置した振動子径5mmの超音波プローブ(中心周波数5MHz，比帯域幅61.6%)により計測する．検出した超音波は，広帯域アンプにより増幅され，USBオシロスコープ(Analog Discovery2, Digilent)でPCに取り込んだ. サンプリング周波数100MHz,

512 点で測定したので，水中での音速を 1500[m/s]とすると，観測可能な深さ範囲は 7.68mm である．また USB オシロスコープのアナログ出力によりガルバノスキャナを制御した．

取得した信号はヒルベルト変換により画像化した．この際，画像をリアルタイムに表示するために GPU(Graphics processing Unit)を用いて並列処理を行った．

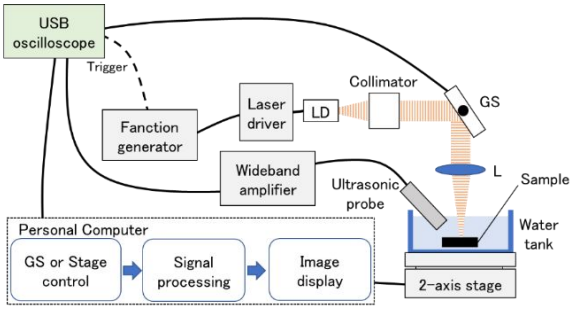
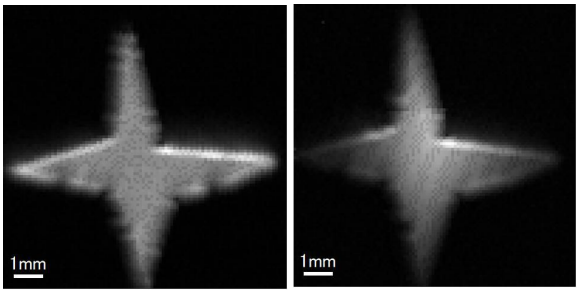


図 1 光音響イメージングシステムの概略図．LD, laser diode; GS, galvano scanner; L, objective lens.

3. 撮像した画像の評価

光吸収のサンプルとして，厚さ 0.2mm の十字に切った黒いビニルテープを用いた．撮像範囲は 10mm×10mm，間隔は 0.1mm とし，100×100 点で計測した．画像は振幅の最大値を表示している．

サンプル走査により撮像した PA 画像を図 2(a)に示す．ステージ移動の際の水面の揺れが無くなった後に測定したため，撮像には約 3 時間掛かった．レーザー走査により撮像した画像を図 2(b)に示す．撮像時間は約 15 秒であった．レーザー走査による撮像では照射する光が撮像範囲内を移動するため，超音波プローブの中心軸から離れるほど振幅が小さくなることが分かった．これは振動子径 5mm のプローブでレーザー走査範囲 10mm を撮像したためで，より大きな振動子径のものを用いることで改善が期待できる．



(a) (b)

図 2 光音響画像 (a)サンプル走査による撮像. (b)レーザー走査による撮像

4. まとめ

本研究では，レーザーダイオードを用いた低コストな光音響イメージングシステムを開発し，サンプル走査とレーザー走査による PA 画像の評価を行った．サンプル走査の利点は，プローブの大きさに関わらず撮像範囲すべてが一定の信号強度で撮像できることである．しかし，ステージの動作速度により撮像に時間がかかることや，振動による水の界面の揺れなどの欠点がある．これに対し，レーザー走査方では短時間での撮像が可能であるが，プローブの中心から離れるにつれて信号強度の低下がみられた．今後は，プローブの中心軸外の振幅の改善や，光の集束を改善して横方向分解能を向上させることなどが課題としてあげられる．

利益相反の有無

なし

文　献

［1］ Xu M and Wang LV: Photoacoustic imaging in biomedicine: Review of scientific instruments 77: 041101, 2006

［2］ Hongtao Z , Tingyang D et al.: Review of Low-Cost Photoacoustic Sensing and Imaging Based on Laser Diode and Light-Emitting Diode: Sensors 2018, 18, 2264

マウス脳 ex vivo イメージングのための マルチピンホール型蛍光 X 線 CT における入射強度補正

保科 亮太[*1]　金澤 道和[*2]　佐々木 洸介[*2]　笹谷 典太[*1]
砂口 尚輝[*3]　河嶋 秀和[*4]　兵藤 一行[*5]　銭谷 勉[*2]　湯浅 哲也[*1]

要旨

マルチピンホール型蛍光 X 線 CT（multi-pinhole x-ray fluorescence computed tomography; mp-XFCT）は，入射 X 線により励起された元素から発生する蛍光 X 線を複数のピンホールを用いて効率的に計測することで，対象の 3 次元機能情報を取得する．現行 mp-XFCT システムには，入射 X 線強度分布が一様でないという問題があり，ビーム断面積より大きい対象を分割して撮像した場合，非一様性の影響で再構成画像が不連続になる．本研究では，あらかじめ計測したビームプロファイルにより入射 X 線強度分布を補正する手法を提案する．非放射性のヨウ素化合物 N-isopropyl-4-iodoamphetamine（iofetamine）を投与後に摘出したマウス脳の 3 次元再構成を試みた．脳全体は入射ビーム断面積で覆いつくすことができないため，全体を 3 分割して高さを変えながら 3 回 ex vivo 計測した．本手法を適用することで，多少の不連続性は残るものの，脳全体に分布するヨウ素をシームレスな 3 次元画像として再構成することができた．

キーワード：蛍光 X 線，CT 画像再構成，ビーム強度補正，分子イメージング，機能情報

1.　はじめに

近年，遺伝子工学の技術の向上に伴い，遺伝子操作による疾患モデル小動物を用いた，疾患の進展過程や原因の究明，診断・治療法の確立，治療薬開発における前臨床研究が盛んに行われている．この際，小動物用の SPECT（Single Photon Emission Computed Tomography）や PET（Positron Emission Tomography）などの核医学イメージング法により，機能情報を評価することが一般的であるが，臓器や器官のサイズに対して，空間分解能が不十分であることや，造影剤として放射性同位体を投与する必要があることなどの制約がある．

このような背景のもと，PET・SPECT に代わる新しい分子イメージングモダリティとして，蛍光 X 線 CT（X-ray Fluorescence Computed Tomography: XFCT）が開発されてきた[1]-[16]．XFCT は，高感度な微量元素分析手法である蛍光 X 線分析技術と，着目する物理量の空間分布図を得るための画像再構成技術を組み合わせた断層撮像法である．XFCT の特色は機能情報の高感度な検出に加え，1 [mm] 以下の高い空間分解能，高い定量性，非放射性の元素濃度分布の撮像が挙げられる．我々の研究グループは，効率的なデータ収集を実現するために，立体ビームを使用したマルチピンホール型 XFCT（multi-pinhole X-ray Fluorescence Computed Tomography: mp-XFCT）を開発し，Sunaguchi[10]や Sasaya[16]らにより，mp-XFCT の 3 次元画像再構成に成功している．しかし，現行の mp-XFCT システムは，入射 X 線強度分布が一様でないために，ビーム断面積より大きい対象を分割して撮像した場合，強度不均一性の影響で再構成画像が不連続になる．本研究では，あらかじめ計測したビームプロファイルを用いて入射 X 線強度分布を補正する手法を提案する．KEK AR-NE7 に構築した mp-XFCT システムを用いて非放射性のヨウ素化合物 N-isopropyl-4-iodoamphetamine（iofetamine）を投与後に摘出したマウス脳の 3 次元再構成を試みた．脳全体は入射ビーム断面積を覆いつくすことができないため，全体を 3 分割して高さを変えながら 3 回 ex vivo 計測した．本手法を適用することにより，多少の不連続性は残るものの，脳全体に分布するヨウ素をシームレスな 3 次元画像として再構成することができた．

2.　方法

2.1　mp-XFCT 撮像原理

図 1 に mp-XFCT 撮像システムのジオメトリを示す．被写体内部の元素は X 線が照射される

*1 山形大学大学院理工学研究科応用生命システム工学専攻
〔〒992-8510 米沢市城南 4 丁目 3-16〕
e-mail: tms20907@st.yamagata-u.ac.jp
*2 弘前大学大学院理工学研究科
*3 名古屋大学大学院医学系研究科
*4 京都薬科大学放射線同位元素研究センター
*5 高エネルギー加速器研究機構物質構造科学研究所

と，光電効果による励起とそれに続く脱励起によって，蛍光 X 線を放射する．発生した蛍光 X 線は，コンプトン散乱の影響を防ぐために入射 X 線に対し垂直に設置したピンホールを通り，検出面上の円形領域内で観測される．離散空間における被写体内の j 番目ボクセル v_j $(1 \le j \le J;\ J$は総ボクセル数$)$の造影剤濃度 q_j[g/mm^3] と，検出器Dで観測される総 X 線光子数 p_i[photons] との関係は，

$$p_i = \sum_j C_{ij} q_j \qquad (1)$$

と表される．投影行列C_{ij}は，

$$C_{ij} = N\,\omega\,\mu_{ph}\,t\,V\left(\frac{\delta_j}{4\pi}\right) \times \exp\left(-\sum_{m \in T^j} \mu_m^I\, l_m\right) \exp\left(-\sum_{n \in U^{ij}} \mu_n^F\, l_n\right) \qquad (2)$$

と表される[10]．ここで，N[photons / mm^2 / s], ω, μ_{ph} [mm^2/g], t [s], V [mm^3], δ [steradian] は，それぞれ，被写体直前で入射X線強度，蛍光 X 線効率，造影剤元素の光電吸収係数，1 投影あたりの露光時間，ボクセルの体積，v_j から通過するピンホールを見込んだ立体角を示す．また，μ_m^I，μ_n^F は，それぞれ入射 X 線エネルギー，蛍光 X 線エネルギーでの被写体の線減弱係数を示しており，既知の値とする．入射 X 線がv_j に到達するまでに通過したボクセル集合 T^j，V_j から発生した蛍光 X 線が D に到達するまでに通過するボクセル集合を U^{ij}とし，投影線がボクセルを横切る長さl_m，l_nの計算には Siddon のアルゴリズム[17]を用いた．式(1)に対して統計学的逐次近似手法である OS-EM （Ordered Subsets Expectation Maximization）法[18]を用いることで造影剤濃度 q_j を推定する．

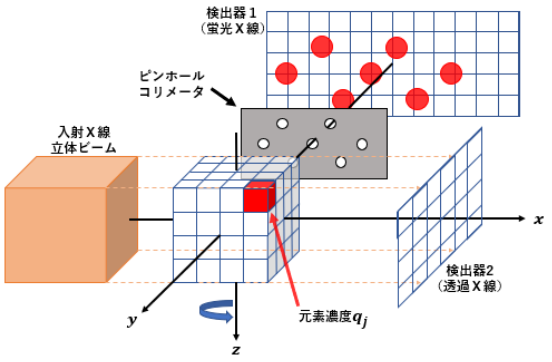


図 1　mp-XFCT の撮像ジオメトリ

2.2　ビーム強度補正

放射光ビームの強度分布は二次元ガウス分布を呈する．すなわち，ビームの中心での強度が最も強く，中心から離れるにしたがって弱くなる．従来は，入射ビーム断面は一様な強度分布として再構成をしてきたため，同一濃度のヨウ素分布はビーム中心近傍では周辺部よりも比較的高く推定される傾向があった．図 2 にビームライン AR-NE7A におけるビームプロファイルを示す．一般的には，2 次元ガウス分布を取るが，本実験で用いたビームラインでは，水平方向の標準偏差は大きくほぼ一様分布と見なせる．本研究では，図 2 のように水平方向および鉛直方向のビームプロファイルは，それぞれガウス分布および一様分布であると近似する．

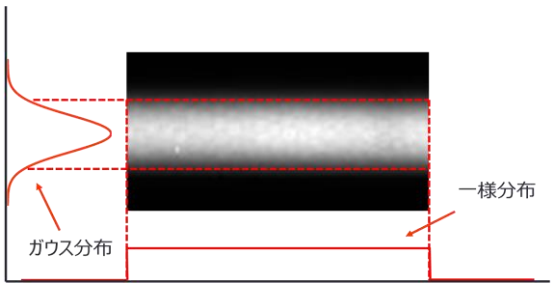


図 2　ビーム強度分布

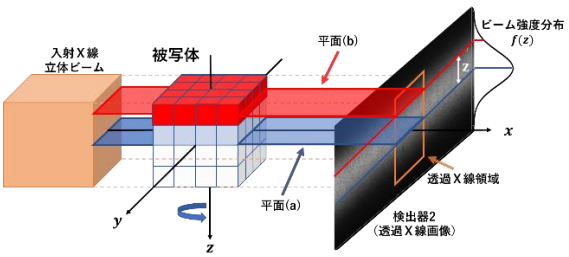


図 3　透過X線画像のジオメトリ

入射 X 線ビームの分布を画像再構成に反映させるために，入射ビームの強度分布画像を取得する．被写体下流に透過 X 線用の検出器として，Photonic Science 社製 X-ray FDI VHR 16M（ピクセル数 : 2436 ×1624、ピクセルサイズ : 7.4 [μm^2]）を設置した（図 1 : 検出器 2）．

図 3 は入射X線ビームと離散化再構成空間との位置関係を示す．撮像システムを構築する際，再構成空間の原点がビームプロファイルで最大値を取る水平面上にあるように注意深く装置を配置した．すなわち，3 次元再構成画像の中央スライスは，鉛直方向のビームプロファイルが最大値を示す高さと一致する．

図中の青および赤の水平なシート状ビームを

それぞれ平面(a)および平面(b)と呼称する．平面(a)は再構成空間の原点を含み，平面(a)と平面(b)の距離をzとする．ビームプロファイルは，図 2 のように水平方向に一様分布であると仮定したので，平面(a)と検出面との交線上では一定値（ビームプロファイル最大値）を取る．同様に，平面(b)と検出面との交線上でも一定値を取る．このとき，鉛直方向はガウス分布であるので，平面(b)の強度は平面(a)の強度から以下のように表現できる．すなわち，鉛直方向の分布を $f(z) = \exp(-z^2/\sigma^2)$ とすれば，平面(b)での入射強度は $Nf(z)$ と表せる．ここで，σ は鉛直方向の広がりを表すパラメータであり，計測データよりフィッティングにより求める．したがって，式(2)の右辺において，平面(b)内にあるボクセルに対応する C_{ij} については，入射強度 N を $Nf(z)$ で置き換える．他の水平面についても同様の考察により，C_{ij} を変換する．なお，係数行列の入射ビーム強度補正をするにあたり，蛍光X線検出器と透過X線検出器の pixel サイズを考慮する必要がある．検出器 2 と再構成ボクセルの一辺の長さはそれぞれ 7.4 [μm]，172 [μm]であることを考慮した．

3. 実験

3.1 撮像条件と再構成条件

提案手法の有用性を示すために，高エネルギー加速器研究機構 PF-AR（加速電圧 5.0 [GeV]）の偏向電磁石ビームライン AR-NE7A に構築した撮像システムで撮像実験を行った．入射エネルギーは，ヨウ素の K 吸収端エネルギーである 33.17 [keV]の直上となる 33.4 [keV]に設定した．被写体直前における入射ビーム中央の単位時間・単位面積あたりの光子数 N は，約 1.0×10^9 [photons/mm^2/s] であった．被写体の回転中心からピンホールコリメータ，ピンホールコリメータから検出面までのそれぞれの距離は，29.1 [mm], 38.2 [mm] であった（図 1）．また，散乱光を抑制するために，0.2 [mm] 厚の Sn 箔を検出器 1 の検出面前に貼付した．2° ステップ 180 枚，各ステップでの露光時間は 180 秒で投影を取得した．

画像再構成法は OS-EM 法を用いて，サブセット数 10，繰り返し回数は 10 回とした．ボクセルサイズは 0.172×0.172 [mm^2] である．また，被写体内での減衰補正に使用する 33.4 [keV]（入射 X 線エネルギー）および 28.3 [keV]（ヨウ素 Kα 蛍光 X 線エネルギー）における水の線減弱係数の値 μ_m^I，μ_n^F はそれぞれ 0.0316 [mm^{-1}] および 0.0385 [mm^{-1}] とした．また，μ_{ph}, ω はそれぞれ 3.5×10^3 [mm^2/g], 0.83 とした．

3.2 被写体

被写体として物理ファントムと生体試料を用いて撮像実験を行った．

3.2-1 物理ファントム

物理ファントムには直径 10 [mm], 高さ 40 [mm] の PMMA 製円筒に直径 3 [mm]，深さ 20 [mm]の孔を 7 つ加工し，濃度が異なるヨウ素溶液（0.05～0.30 [mg/ml]）を図 4 のように封入した．図中の数字はヨウ素濃度 [mg/ml] を示す．再構成画像サイズは 70×70×36 である．従来手法と提案手法について，再構成画像，3D-VR 画像を比較し評価を行った．

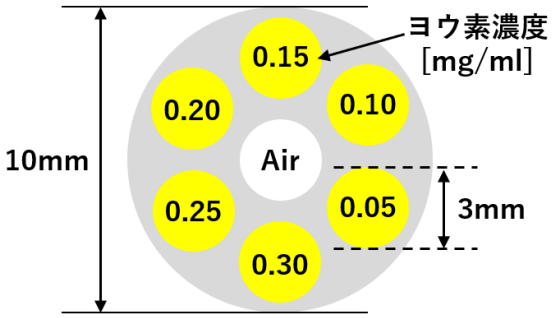


図 4 物理ファントム

3.2-2 生体試料

マウスに非放射性ヨウ素化合物 N-isopropyl-4-iodoamphetamine（iofetamine）を投与後，脳を摘出しホルマリンにより固定した．本サンプルは京都薬科大学の動物実験倫理指針に準拠して処置された．マウス脳を図 5 に示すように 4.5 [mm]間隔で全体を 3 分割して ex vivo 計測した．3 次元平滑化処理としてσ = 1.5 [pixel] （172 [μm] × 1.5 = 258 [μm]）の 3 次元ガウシアンフィルタを用いた．再構成画像から 3D-VR 画像の作成し，従来手法と提案手法の比較を行った．

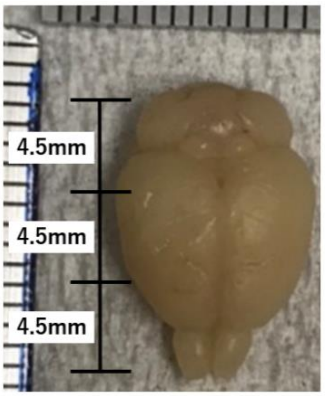


図 5 摘出したマウス脳

4. 結果

4.1 物理ファントム

図 6 は従来法と提案法の再構成結果であり，それぞれの上から 8，18，28 番目のスライスの断層像を示している．ここで，18 スライスがビーム強度最大の断層である．従来法の再構成では，18 スライスに比べ，8 および 28 スライスのコントラストが著しく低い．提案法ではコントラストが回復しており，高さ方向に比較的均一なコントラストになっている．一方，8 および 28 スライスの画質は 18 スライスに比べ低い．これは，検出される蛍光X線光子数が少ないため，再構成画像の S/N が低下することによる．

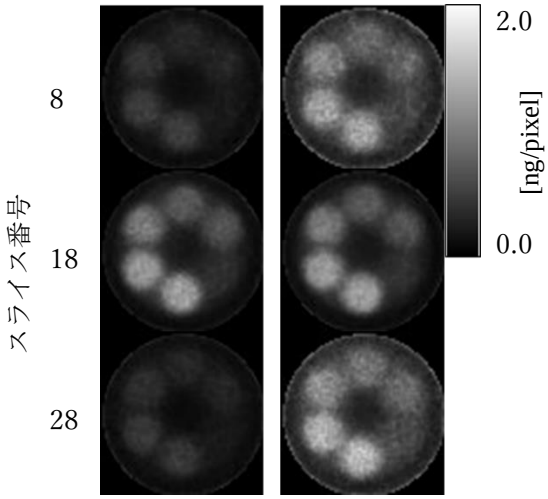


図 6 再構成画像結果

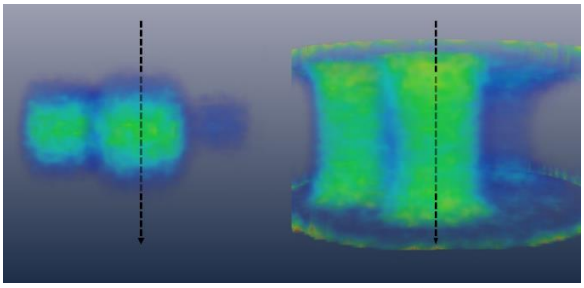
得られた再構成画像から作成した，3D-VR 画像を図 7(a)に示す．閾値は同一に設定した．従来法では，上部および下部でコントラストが消失しているが，提案法では良好に描出されている．一方，提案法で上部および下部で円柱の半径が中央部に比べ大きくなる傾向がある．これは，再構成画像の S/N が低下したことにより，ヨウ素領域のエッジが鈍ったことによるものと推測している．

図 7(a)のヨウ素濃度 0.3 [mg/ml] 領域内に示した破線上における推定ヨウ素濃度を図 7(b)に示す．縦軸および横軸は，それぞれスライス番号および推定されたヨウ素濃度値である．提案法では，上部と下部におけるコントラストが，比較的一様に回復している．

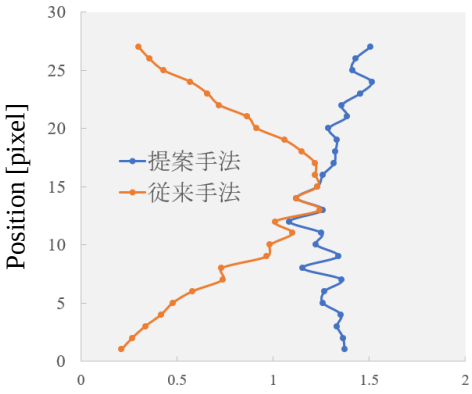
それぞれのヨウ素濃度領域における，輝度値のばらつきを評価するために，変動係数(CV : Coefficient of Variation)を求めた．

$$\mathrm{CV} = \frac{\sigma}{m} \qquad (3)$$

ここで，σ，m はそれぞれ標準偏差，平均値である．各ヨウ素濃度における領域における CV を図 8 に示した．提案法では，従来法に比べ，ばらつきが 2〜5 倍程度の改善された．



(a)



(b)

図 7 **(a)** 3D-VR 画像，**(b)** 0.3 mg/ml 領域のヨウ素分布の均一性

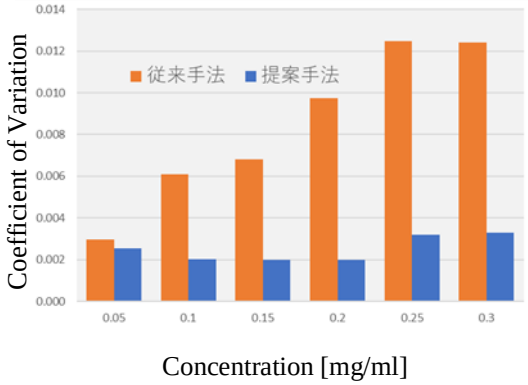


図 8 ヨウ素濃度領域における変動係数

4.2 生体試料

図 9 にマウス脳に最も顕著にヨウ素の集積が見られたスライスの再構成画像を示す（3 次元再構成画像の上から 6.88 [mm]に相当するスライス）．両画像は最も蓄積領域が視認しやすいようにコントラストを調整しているため，両画像における画素値の比較は意味がない．従来法と提案法の再構成画像の濃度分布には大きな差はなく，いずれも視床部へのヨウ素の蓄積が確認できる．

また，図 10 にはマウス脳全体の 3D-VR 画像を示す．図 10 の赤い線は図 9 のスライス位置を示す．閾値は両者で同一である．従来法では，継ぎ目近傍でコントラストが消失している．一方，提案手法では，従来法に比べてシームレスに表示されている．しかし，不連続性は完全には補正されておらず，今後の改良を要する．

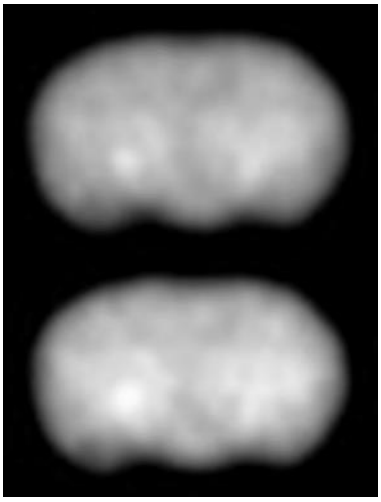


図 9 マウス脳の再構成画像

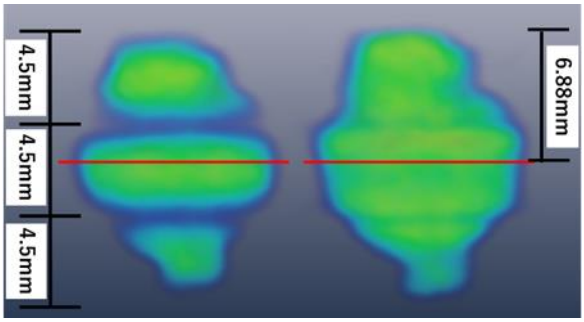


図 10 マウス脳の 3D-VR 画像

5. 総括

本研究では，mp-XFCT における，入射ビーム強度分布に起因するコントラスト不均一性の改善方法の提案と検証を行った．さらに，入射ビーム断面積よりも大きい生体試料を分割撮像したデータに提案手法を適用して，その有効性を確認した．

入射ビーム強度補正項を再構成アルゴリズムに組み込むことで，より現実的な実験条件をモデルに反映させることが可能になった．結果より，濃度分布不均一の改善が確認できた．定量性の向上に関しては，各ヨウ素領域での濃度分布のばらつきが約 2〜5 倍改善することが確認できた．しかしながら，強度分布画像の中心座標の取得や測定機器間のピクセルサイズの調整は手動で行っているため，操作する者により差が生じる可能性が考えられる．今後は，これらも含め自動化を進めてゆく予定である．

提案手法を用いて再構成することにより，分割撮像をしたマウス脳の全体を観測する 3 次元再構成が可能となったが，完全には不均一性を補正するには至らなかった．今後は，実用に耐えうるよう，データ取得方式および再構成アルゴリズムにさらなる改良を加えてゆきたい．

利益相反の有無 なし

文 献

［1］ J.P. Hogan, R.A. Gonsalves, A.S. Krieger: Fluorescent computer tomography: a model for correction of X-ray absorption, IEEE Trans. Nucl. Sci., vol.38, no.6, pp.1721-1727 (1991).

［2］ T. Takeda, M. Akiba, T. Yuasa et al: Fluorescent x-ray computed tomography with synchrotron radiation using fan collimator, Proc. SPIE, vol.2708, pp.685-695 (1996).

［3］ T. Takeda, T. Yuasa, A. Hoshino et al: Fluorescent x-ray computed tomography to visualize specific material distribution, Proc. SPIE, vol.3149, pp.160-172 (1997).

［4］ T. Yuasa, M. Akiba, T. Takeda et al: Reconstruction method for fluorescent X-ray computed tomography by least-squares method using singular value decomposition, IEEE Trans. Nucl. Sci., vol.44, no.1, pp.54-62 (1997).

［5］ Thet-Thet-Lwin, T. Takeda, J. Wu et al: Preliminary quantitative analysis of myocardial fatty acid metabolism from fluorescent X-ray computed tomography imaging, J. Synchrotron Radiat., vol.14, no.1, pp.158-162 (2007).

［6］ Thet-Thet-Lwin, T. Takeda, J. Wu et al: Visualization of age-dependent cardiomyopathic model hamster obtained by fluorescent X-ray

computed tomography using I127-BMIPP, J Synchrotron Rad., vol.15, no.5, pp.528-531 (2008).
[7] T. Takeda, J. Wu, Thet-Thet-Lwin, Q. Huo, et al: X-ray fluorescent CT imaging of cerebral uptake of stable-iodine perfusion agent iodoamphetamine analog IMP in mice, J. Synchrotron Rad., vol.16, no.1, pp.57-62 (2009).
[8] Q. Huo, T. Yuasa, T. Akatsuka et al: Sheet-beam geometry for in vivo fluorescent x-ray computed tomography : proof-of-concept experiment in molecular imaging, Opt. Lett., vol.33, no.21, pp.2494-2496 (2008).
[9] Q. Huo, H. Sato, T. Yuasa et al: First experimental result with fluorescent X-ray CT based on sheet-beam geometry, X-Ray Spectrom., vol.38, no.5, pp.439-445 (2009).
[10] N. Sunaguchi, T. Yuasa, K. Hyodo et al: Fluorescent x-ray computed tomography using the pinhole effect for biomedical applications, Opt. Comm., vol.297, no.15, pp.210-215 (2013).
[11] G. L. Brownell: Theory of isotope scanning, Medical isotope scanning, vol.43, no.45, pp.1-11 (1959).
[12] S. K. Cheong, B. L. Jones, A. K. Siddiqi et al: X-ray fluorescence computed tomography (XFCT) imaging of gold nanoparticle-loaded objects using 110 kVp x-rays, Phys. Med. Biol., vol.55, no.3, pp.647-662 (2010).
[13] Y. Kuang, G. Pratx, M. Bazalova et al: First Demonstration of Multiplexed X-Ray Fluorescence Computed Tomography (XFCT) Imaging, IEEE Trans. Med. Imag., vol.32, no.2, pp.262-267 (2013).
[14] D.P. Clark, K. Ghaghada, E.J. Moding et al: In vivo characterization of tumor vasculature using iodine and gold nanoparticles and dual energy micro-CT, Phys. Med. Biol., vol.58, no.6, pp.1683-1704 (2013).
[15] H.M. Hertz, J.C Larsson, U. Lundström: Laboratory x-ray fluorescence tomography for high-resolution nanoparticle bio-imaging, Opt. Lett.,vol.39, no.9, pp.2790-2793 (2014).
[16] T. Sasaya, T. Yuasa, K.Hyodo, et al: Multi-pinhole fluorescent x-ray computed tomography for molecular imaging, Sci Rep., vol. 7, no. 5742 (2017)
[17] R.L. Siddon: Fast calculation of the exact radiological path for a three dimensional CT array, Med. Phys., vol.12, no.2, pp.252-255 (1985)
[18] H.M. Hudson, R.S. Larkin: Accelerated image reconstruction using ordered subsets of projection data, IEEE Trans. Med. Imag., vol.13, no.4, pp601-609, (1994).

細胞内構造の深層学習とレディオミクスに基づく癌細胞のラベルフリー分類

大手　希望[*1]　橋本　二三生[*1]　山田　秀直[*1]　山内　豊彦[*1]　羽田　綾馬[*2]
廣津　周[*2]　菊池　寛利[*2]　岡崎　茂俊[*3]　竹内　裕也[*2]　上田　之雄[*1]

要旨

ラベルフリーの細胞分類は，試薬による細胞の損傷を避けるため重要である．その方法の 1 つとして定量位相顕微鏡（QPM）を用いた細胞分類がある．これまで QPM を用いた細胞内構造に基づく白血球と癌細胞株の 2 クラス分類は検討されていたが, 癌細胞株内の多クラス分類は検討されていなかった. そこで, 本研究では QPM を用いた細胞内構造に基づく癌細胞株の多クラス分類を試みた．QPM によって撮影された 5 種類の癌細胞株の画像 250 枚を使用した．さらにレディオミクスによって輝度値，形状，テクスチャの特徴を抽出した．画像とレディオミック特徴を学習済みの畳み込みニューラルネットワークである VGG16 に入力しファインチューニングを行った. 10 交差検証により正答率, AUC を評価し, 正答率 80.5 %, AUC = 0.96 を得た．これらの結果から QPM を用いた細胞内構造に基づく癌細胞株のラベルフリー分類の可能性が示唆された.

キーワード : 細胞分類，深層学習，畳み込みニューラルネットワーク，定量位相顕微鏡，レディオミクス

1. はじめに

原発腫瘍組織または転移腫瘍組織から遊離し血液中に浸潤した癌細胞は，血中循環腫瘍細胞（circulating tumor cell; CTC）と呼ばれる．この CTC は固形癌患者の抹消血液中に極微量に存在し転移に関わるとされ，近年研究が盛んに行われている. 乳がん患者においては, 全血 7.5 mL 中に CTC が 5 個未満であれば 1 年後の死亡率が 19 %であり，CTC が 5 個以上であれば 1 年後の死亡率が 53 %であったと, 報告されている［1］．一方，抹消血液中の有核細胞はその殆どが白血球であるため，白血球と癌細胞とを識別することが重要となる．

*1 浜松ホトニクス株式会社中央研究所
〔〒434-8601 浜松市浜北区平口 5000〕
e-mail: kibou@crl.hpk.co.jp
*2 浜松医科大学医学部
*3 浜松医科大学光尖端医学教育研究センター

これらを識別する方法として，細胞の外形状だけでなく，内部構造が反映されている形態情報と機械学習による自動判別手法が提案されている［2］．ここで，CTC をいくつかの種類に細分類できれば，CTC の数だけではなく種類に基づいた質的診断が可能となり，個別化医療の実現が期待できる．

放射線医学における腫瘍組織の質的診断法としてレディオミクスがある［3，4］．レディオミクスは腫瘍組織の内部構造に着目し，輝度値，形状，テクスチャなどの特徴量を網羅的に抽出し，バイオマーカーとする手法である. これを CTC に応用した例はまだ報告されていない．さらに，畳み込みニューラルネットワーク(CNN)により画像から抽出された深層特徴とレディオミック特徴を結合しマルチモーダル学習を行うハイブリッドモデルが, CNN，レディオミクス両者の性能を改善するものとして注目されている［3］.

そこで本研究では, CNN とレディオミクスを組み合わせた CTC 細分類法を確立するこ

とを目的とし，樹立癌細胞株を対象としてその性能を評価することで，CTC 細分類の実現可能性について検討を行った．

2. 資料と方法

1）データ

定量位相顕微鏡によって撮影された 5 種類の樹立癌細胞株の画像 50 枚，合計 250 枚を対象とした［2］．画素数は 49×49 である．

2）提案手法

図 1 に提案手法の構成を示す．提案手法では，CNN によって抽出された画像特徴量と，レディオミック特徴量を全結合層にて結合するマルチモーダル学習を行うことで，樹立癌細胞株の細分類を行う．本研究では，CNN の構造として VGG16 を使用した． また，レディオミック特徴の抽出には PyRadiomics フレームワークを利用し，107 個の特徴を抽出した［4］．

3）検証方法

単純な VGG16 とレディオミック特徴を追加した VGG16 の性能を比較した．10 交差検証を 3 回繰り返し，正解率の平均，標準偏差を求めた．また 3 回分の結果からプール法により ROC 曲線を作成し，ROC 曲線下面積(AUC)を求めた．

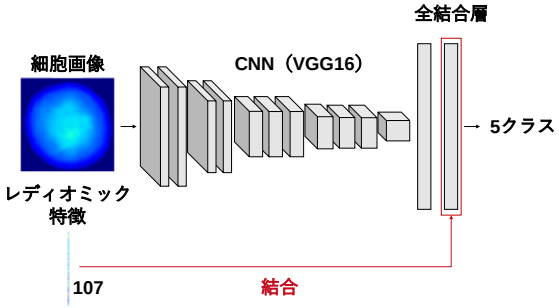


図 1 提案手法の構成．VGG16 の最後の隠れ層にレディオミック特徴 107 個を結合．

3. 結果と考察

表 1 に単純な VGG16 とレディオミック特徴を追加した VGG16 による樹立癌細胞株の細分類結果を示す．レディオミック特徴を追加した VGG16 では，単純な VGG16 に比べて，正解率が 6%，AUC が 0.02 改善した．提案手法において，正答率 80.5%，AUC 0.96 がえられた．これらの結果からレディオミクスは CTC にも有用であることが示唆された．

表 1 VGG16 とレディオミック特徴を追加した VGG16 による樹立癌細胞株の細分類結果

指標	VGG16	VGG16 + Radiomics
正解率(%)	74.5±0.4	80.5±1.0
AUC	0.94	0.96

4. まとめ

本研究では，深層学習とレディオミクスを組み合わせた CTC 細分類法を提案した．提案手法では，樹立癌細胞株の細分類において 80%を超える正答率がえられ，CTC 細分類の実現可能性が示唆された．

謝辞 本研究は AMED の課題番号 JP20he2202006 の支援を受けた．

利益相反の有無 なし

文 献

［1］Cristofanilli, M et al: Circulating tumor cells, disease progression, and survival in metastatic breast cancer. N Engl J Med **351**: 781-791, 2004

［2］Ozaki, Y et al: Label-free classification of cells based on supervised machine learning of subcellular structures. PLoS One **14**: e0211347, 2019

［3］Afshar, P et al: From hand-crafted to deep learning-based cancer radiomics: challenges and opportunities. IEEE Signal Processing Magazine **36**: 132-160, 2019

［4］Van Griethuysen, J J M et al: Computational radiomics system to decode the radiographic phenotype. Cancer Research **77**: e104-e107, 2017

手術ロボット開発に向けた犬の避妊手術シミュレーション・データセット作成

早川 吉彦[*1] 久冨木 望[*2] Yin Ran[*2] Pathitta Suteecharuwat[*3] 長瀬 和弘[*3]

要旨

術者の動作・処置が定型化した手術では，自律型手術ロボットの開発に対してニーズが高い．今回は，小動物・犬を対象に定型的手術である避妊手術（卵巣子宮摘出手術）を採り上げ，シミュレーションモデルを用いて手術に関するデータセットの作成を行った．画像認識技術によって，手術ツールの認識と手術フェーズ・ワークフローの認識を行った．避妊手術モデルを使い，術中時間 15 から 25 分の手術ビデオを 42 本作製した．小動物・犬の手術を対象にデータセット作成を行った．手術ツールの認識では 15 本のビデオにアノテーションをつけた．物体検出を目的としたデータセットでは，長方形の枠線で物体の領域を囲む bounding-box アノテーションが使われる．そこで位置情報の必要性と精度向上のために bounding-box でアノテーションした．手術フェーズ・ワークフローの認識では，6 つのフェーズについて全ビデオにアノテーションした．

キーワード : Surgical operation work-flow, Annotation, Surgical instrument, Canine spay surgery, Image recognition

1. はじめに

「定型手術」のように動作が定型化している手術は，医師の減少する地方都市では自律型手術ロボットとして技術的にもニーズが高い．しかし，研究開発でデータ作成を行う際に患者同意を得る必要がありデータ収集におけるハードルが高い. そこで, 比較的同意を得やすい小動物「犬」を対象に最初のステップとして定型手術である避妊手術のシミュレーションモデルを用いたデータセットの作成を行った. 本研究では画像認識技術によって「手術ツールの認識」「手術フェーズ・ワークフローの認識」を目的とする.

*1 北見工業大学工学部地域未来デザイン工学科機械知能・生体工学コース
[090-8507 北海道北見市公園町 165 番地]
e-mail:hayakayo@mail.kitami-it.ac.jp
*2 北見工業大学工学部情報システム工学科
*3 北見工業大学大学院情報システム工学専攻
投稿受付：2020 年 8 月 5 日

2. 方法

1) 犬の避妊手術シミュレーションデータセット(C3S データセット)の撮影

使用した機材は，ミラーレス一眼レフカメラ OM-D E-M1 MarkII（OLYMPUS），モデルは，3B サイエンティフィック(ドイツ ハンブルグ)社製の犬の避妊手術シミュレーション模型，交換用皮膚，生殖器，膀胱セットである．撮影条件は，フレームレート 24fps、画素数 1080×720（42 本目は 1920×1080），ビデオ 1 本の長さ 15～25 分（術中時間）である．

術者の視点から術部の距離は約 25cm 程あるため，カメラの位置からモデルまでの距離も同様に設定した．また，実際の手術では多様な照明状態が存在しているため，今回は 3 種類の照明状態(1.やや暗い 2.やや明るい 3.明るい)で調整して撮影を行った（図 1）.

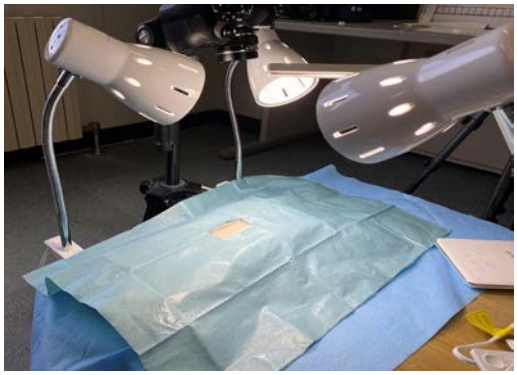


図 1：術中の照明状態(3. 明るい)

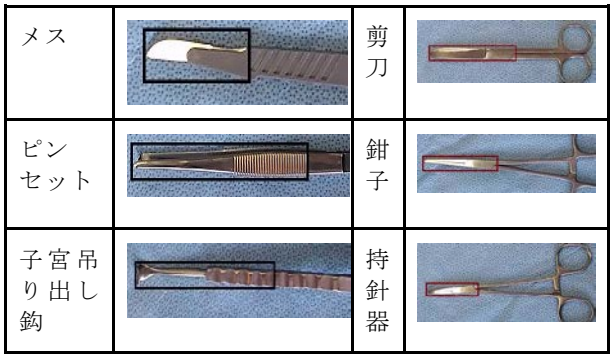


図 2：ツールのアノテーション領域

2) 避妊手術（卵巣子宮摘出手術）の流れ

流れは，次のようになる．

(1) メスを用いて腹部正中皮膚を切皮し，剪刀を用いて頭尾側方向へ切開を拡大する．

(2) 子宮吊り出し鈎で卵巣を露出させる．

(3) 鉗子をかけて 1 つ目の卵巣茎の結紮処理をし切断する．

(4) 2 つ目の卵巣茎の処理も同様に行う．

(5) 子宮頸部に鉗子をかけて結紮処理をした後に子宮断端を処理する．

(6) 閉腹処理をする．

3) ツール認識のデータセット

ツール認識のデータセットでは，1-15 番のビデオにアノテーションをつけた．クラス分類のためのデータセットでは 1/0 で物体の存在を注釈するバイナリー・アノテーションが使われ，検出のためのデータセットでは，長方形の枠線で物体領域を囲む bounding-box アノテーションが使われる．本研究では「ツール認識の精度向上」と手術ロボット開発のために「位置情報」が必要であるため，bounding-box でアノテーションした．24fps で撮影したが，フレームとフレームが類似していると学習価値が低いため，ビデオを 1fps に down-sampling してツールアノテーションをした．

外科手術のツールは手で把持されている状態が多く，ツール全体のアノテーションすることが難しいため，ツールの一部分だけアノテーションをした(図 2)．

剪刀と持針器と鉗子は真横から見た状態が非常に類似していて，人間の眼でも判別できないという課題が存在する．そこで，縫合針を挟んでいる持針器にアノテーションする際は，縫合針と一緒に bounding-box をつけることで持針器の区別ができるように改善した．このようにアノテーションすることで，持針器の画像がぼやけている場合の認識精度を高めることもできると考えられる．縫合針も使われたが，1. 縫合針のサイズが小さい，2. 角度により見える場合が少ない，3. 持針器で挟まれる場合が多い．この 3 つの理由で，個別に bounding-box のアノテーションはしなかった．

ツールのアノテーション領域が 25%以下の場合，あるいは画像の不鮮明であるなど人間の眼で判断できない場合はアノテーションしなかった．認識が難しい object に対しては difficult flag を「1」，そうではないなら「0」にする．アノテーション情報は xml ファイルで保存した．各 bounding-box は，name flag でツール名を保存した．

結果として，ツール認識でアノテーションされたフレームは 16,327 枚になっており，8 割以上のフレームは複数のツールが存在している．深層学習での認識はより理想な精度に達成できるといえる．

4) フェーズ認識のデータセット

フェーズ認識のデータセットではツール認識のビデオ 15 本に加えて，合計 42 本のビデオにアノテーションをつけた．アノテー

ション情報は xml ファイルで保存した．

フェーズ認識のデータセットの中の 1-15 番のビデオはツール認識のデータセットと同じなので，フェーズ認識はツールの情報を加えて学習することもできる．ツールの出現はフェーズと相関性があり，フェーズ認識を行う場合はツールの検出情報と RGB ビデオの情報からフェーズの認識を行う方が精度を上げることができる．

3. 結果と考察

シミュレーションモデルを用いることと多様な照明状態を設定することで，実際の手術により近い手術環境を再現し撮影することができた．

1）ツール認識のデータセット

16,327 枚の画像を bounding-box でアノテーションした．低侵襲手術の先行研究により，各手術ツールの bounding-box アノテーションは 300-1,000 枚以上でより良い精度に達成したが，開腹手術のツール認識には以下の問題が想定される：

(1) 低侵襲手術と比べて，開腹手術では形が似ているツールが多い．

(2) 開腹手術のツールは,「角度が変わる」「遮られる」「移動により不鮮明に映る」状況が多い．

(3) ツールは全てステンレス製で，鏡面反射により色が変わることで，ツール自体とその影の区別が難しい．

2）フェーズ認識のデータセット

フェーズ認識のデータセットは，42 本のビデオをアノテーションし，各フェーズについては最初のフレームの番号と最後のフレームの番号でアノテーションをつけた．

4. 結論

本研究の成果として，犬の避妊手術シミュレーションモデルを用いたツール認識とフェーズ認識のデータセットが作成できた．今後，このデータセットを用いたニューラルネットワークの設計とトレーニングは手術ロボット開発に向けた次へのステップとなる．

先行研究により，これからのニューラルネットワークの設計は LSTM と RCNN を使用する．ツール認識の精度向上のため，16 番目以降のビデオをアノテーションしてデータセットを拡張する，あるいは物体トラッキング技術を併用することを検討している．フェーズ認識について，少量なデータセットでより理想な認識精度の達成に至るため，ツール検出情報とビデオ情報を加えて認識することを考える．

本研究での撮影は固定カメラを使用したが，実際の手術を撮影する場合はスマートカメラで術者の視点からの撮影も検討している．その際，術者の視点が移動することでツールとフェーズ認識を行う前に術中画像と非術中画像を分類するニューラルネットワークの設計が必要である．

利益相反の有無

なし。

文　献

[1] Jin, A., et al. “Tool detection and operative skill assessment in surgical videos using region-based convolutional neural networks.” 2018 IEEE Winter Conference on Applications of Computer Vision (WACV). IEEE, 2018.

[2] Namazi, B., Sankaranarayanan, G., Devarajan, V. “Automatic Detection of Surgical Phases in Laparoscopic Videos.” 2018.

[3] 竹内啓，獣医外科手術，講談社, 1994

[4] 秋吉秀保，スキルアップシリーズ Target28 -若手獣医師のための実践プログラム- 不妊手術，インターズ, 2016 [DVD]

[5] Countryside Clara, “Canine Spay” Youtube <https://www.youtube.com/watch?v=aUOczV4TCuM>

[6] sheltervet, “Dog Ovariohysterectomy (Spay)” Youtube<https://www.youtube.com/watch?v=rEyZQyeFJFc>

Surgical simulation dataset for surgical robot development:

A canine spay surgery as an example

Yoshihiko HAYAKAWA[*1], Nozomu KUBUKI[*2], Ran YIN[*2], Pathitta SUTEECHARUWAT[*3],

Kazuhiro NAGASE[*3]

*1 Dept. of Engineering on Intelligent Machines & Biomechanics, School of Regional Innovation & Social Design Engineering, Faculty of Engineering, Kitami Institute of Technology

*2 Dept. of Computer Science, Kitami Institute of Technology

*3 Course of Computer Science, Graduate School, Kitami Institute of Technology

Development of autonomous surgical robotic machines for standard surgical procedures is under requisition rather than master-slave machines. Ovariohysterectomy is the common sterilization procedure for small animals, namely cats and dogs, etc. We tried the simulation of the canine spay surgery, ovariohysterectomy, and made the data set of the surgery. The automatic image recognition of tools (surgical instruments) and surgical operation (work-flow) phases were carried out. Using a simulation model of the canine spay surgery, 42 video movies (24 fps), in the range of 15-25min length, were recorded during every surgical procedure. The annotation of surgical tools was carried out in the first 15 video movies (1 fps) by means of the bounding-box annotation. Two or more tools are annotated on more than 80 percent in total 16,327 frames. This method is fitting for the precise object detection having the individual locational data. The annotation of work-flow phases was carried out in all movies. Six consecutive phases in the canine spay surgery were recognized. Since selected tools are used in specific phases, correlations between tools and phases are helpful for the automatic phase recognition. The annotation data-set of tools and work-flow phases were obtained using the simulation of ovariohysterectomy. The next step is the designing of a neural network to recognize these tools and surgical phases. (COI: Nothing to disclose)

Key words: Surgical operation work-flow, Annotation, Surgical instrument, Canine spay surgery, Image recognition

著者紹介

早川　吉彦　(はやかわ　よしひこ)
北見工業大学工学部地域未来デザイン工学科機械知能・生体工学コース．2011 年医用画像情報学会内田論文賞受賞．日本歯科放射線学会監事・代議員・医療情報委員．DICOM WG22 委員．電子情報通信学会, 日本医学物理学会, IADMFR 各会員．CARS 国際会議組織委員. IJCARS & OOOO journal 編集委員.

久冨木　望 (くぶき　のぞむ)
北見工業大学情報システム工学科

Yin Ran (イン ゼン)
北見工業大学情報システム工学科

Pathitta Suteecharuwat (パティッター スティージャールワット)
北見工業大学大学院情報システム工学専攻

長瀬和弘 (ながせ　かずひろ)
北見工業大学大学院情報システム工学専攻

Admissible SDF variationを用いたヒト胚子の脳の平均形状の推定

田邊 宥大[*1]，斉藤 篤[*1]，高桑 徹也[*2]，山田 重人[*2]，清水 昭伸[*1]

要旨

本報告では，符号付距離関数(Signed Distance Function, SDF)によって表現された形状の平均を正確に推定する方法について報告する．SDF 同士の線形演算の結果が SDF である保証はなく，従来の SDF の相加平均によって推定した平均では微細な解剖構造が十分表現できない可能性がある．そこで提案手法では，1 次の図形モーメントと SDF 間の L2 ノルムを用いたエネルギー汎関数の最小化によって平均形状を推定する．エネルギー汎関数の最小化の際に admissible SDF variation を用いることで，推定された平均が SDF であることを保証する．最後に，提案手法を京都コレクションの胚子の脳実質の形状ラベルに適用して，従来の手法で推定した平均形状との比較により，提案手法の有用性について議論する．

キーワード：ヒト胚子，統計的形状モデル，平均の推定，エネルギー汎関数，admissible SDF variation

1. はじめに

ヒトの発生過程において，妊娠 3~9 週は胚子期と呼ばれ，器官を形成するための重要な時期であり，さまざまな異常発生を起こす危険性のある臨界期でもある[1]．また，新生児の死亡原因の 25%が先天異常であると言われており[2]，形態異常については出生前に診断することが可能である．微細な胚子の形態の診断に，正常群を対象とした統計モデルの構築を応用する可能性が検討されている．

これまでに，構築されたモデルの例として，Level Set Distribution Model (LSDM) を用いたヒト胚子の脳を対象としたモデル[3]や，ヒト胚子の複数器官を対象としたモデルの例[4]がある．しかし，[3]や[4]では統計モデルの構築に必要な平均を Level Set(LS)関数同士の線形演算を用いて推定している．LS 関数として用いられる符号付き距離関数(Signed Distance Function，SDF)同士の線形演算の結果が SDF である保証はない．従って，平均形状が不自然な形状になってしまうことがある．

本稿では，1 次の図形モーメントと SDF 間の L2 ノルムを用いたエネルギー汎関数の最小化による平均形状の推定方法について提案する．また，提案手法を京都コレクションのヒト胚子の脳実質の形状ラベルに適用し，従来の手法で推定した平均形状と比較した結果についても報告する．

2. 平均形状の推定

提案手法は，N個のラベルから，式(1)で示したエネルギー汎関数の最小化によって，それらの平均形状を推定する．

$$F(\phi_t) = \beta_1 \sum_{i=1}^{N} F_m(\phi_t, \phi_i) + \beta_2 \sum_{i=1}^{N} F_n(\phi_t, \phi_i) \quad (1)$$

ここで，ϕ_tはt回目の更新後の形状に対応する SDF，$\phi_i (i = 1, \cdots, N)$は$i$番目のデータに対応する SDF，$\beta_1, \beta_2$は重みである．$F_m(\phi, \phi_i)$は 1 次の図形モーメント，すなわち，ラベルの重心間の距離を評価する項である(式(2))．

$$F_m(\phi_t, \phi_i) = \left\| \frac{1}{\int_\Omega H_\epsilon(-\phi_t(\boldsymbol{x}))d\boldsymbol{x}} \int_\Omega \boldsymbol{x} H_\epsilon(-\phi_t(\boldsymbol{x}))d\boldsymbol{x} - \frac{1}{\int_\Omega H_\epsilon(-\phi_i(\boldsymbol{x}))d\boldsymbol{x}} \int_\Omega \boldsymbol{x} H_\epsilon(-\phi_i(\boldsymbol{x}))d\boldsymbol{x} \right\|_2^2 \quad (2)$$

Ωは画像空間，$\boldsymbol{x}$はΩ上の任意の点，$H_\epsilon(x)$は$\epsilon >$

*1 東京農工大学大学院工学研究院
[〒184-8588 東京都小金井市中町 2-24-16]
e-mail: s207087s@st.go.tuat.ac.jp
*2 京都大学 [〒606-8507 京都府京都市左京区聖護院川原町 53]

0をパラメータとするヘビサイド関数の近似式である(式(3)).

$$H_\epsilon(x)=\begin{cases}0(x<-\epsilon)\\1(x>\epsilon)\\\frac{1}{2}\left(1+\frac{x}{\epsilon}+\frac{1}{\pi}\sin\left(\frac{\pi x}{\epsilon}\right)\right)(|x|\le\epsilon)\end{cases} \quad (3)$$

式(1)の$F_n(\phi_t,\phi_i)$はSDF間のL2ノルムを用いた(式(4)).

$$F_n(\phi_t)=\|\phi_t(\boldsymbol{x})-\phi_i(\boldsymbol{x}-\boldsymbol{t_i})\|_2^2 \quad (4)$$

ここで，$\boldsymbol{t_i}$はϕ_iに対応する形状の重心からϕ_tに対応する形状の重心に向かうベクトルである．

式(1)の最小化には勾配法を用いるが，更新後のϕ_tがSDFであることを保証するためにadmissible SDF variation[5]を利用する(式(5)).

$$\delta\phi(x)=\sum_{j=1}^{n}h_j^x\sum_{y\in\Omega}h_j^y\nabla_{\phi_t}F\left(\phi_t(\boldsymbol{y})\right) \quad (5)$$

$\boldsymbol{x},\boldsymbol{y}$は画像空間Ω上の任意の点，$\nabla_{\phi_t}F\left(\phi_t(\boldsymbol{y})\right)$は式(1)の$\phi_t$に関する勾配，$n$は$\phi_t=0$となる点の数，$h_j^x$は$\sum_{j=1}^{n}h_j^x=1$を満たす重みである(式(6)).

$$h_j^x=\frac{\exp\left(-\alpha*d\left(\Gamma(s_j),\Gamma(s_x)\right)\right)}{\sum_{k=1}^{n}\exp\left(-\alpha*d\left(\Gamma(s_k),\Gamma(s_x)\right)\right)} \quad (6)$$

$\Gamma(s_j),\Gamma(s_k)$は$\phi_t=0$となる点，$\Gamma(s_x)$は画像空間上の任意の点$\boldsymbol{x}$からϕ_tに対応する形状に向けておろした垂線の足，$d(\cdot,\cdot)$は2点間のユークリッド距離，$\alpha>0$はパラメータである．

3. 実験条件と実験結果

実験試料には，京都大学が所蔵するヒト胎児サンプル群である京都コレクション[6]の一部の，Carnegie stage; CS 23におけるヒト胚子の脳ラベル画像10個体を使用した．ヒト胚子の画像は，2.35[T]，ボア径40[cm]の横型超伝導磁石の超並列型MR microscopeを用い，100[ms]の繰り返し時間と 10-16[ms]のエコー時間を有するT1強調スピンエコーシーケンスにより撮影されている[7]．その画像から3次元の脳領域ラベルを作成し[8]，Sagittal面で面積最大の2次元ラベル(連結成分数=1)を抽出して利用した(図1).

提案法により計算した平均形状が，狙い通りSDFとなっているか否かを評価するために，式(7)のSDFらしさE_{SDF}を評価した．この値が小さいほど，SDFらしいということを意味する．

$$E_{SDF}=\sum_{x\in\Omega}(1-|\nabla_x\phi|)^2 \quad (7)$$



図1 胚子の2次元ラベル画像

パラメータを$\beta_1=10^5,\beta_2=1,\alpha=0.01,\epsilon=1$として推定した平均形状と評価値，従来行っていたSDF同士の線形演算による結果をそれぞれ，表1と図2に示す.

表1 推定した平均のSDFらしさ

データ(平均)	提案	線形演算
2.49×10^2	1.42×10^2	1.12×10^3



(a) 提案　　(b) 線形演算

図2 推定した平均形状

表1より，提案手法によって推定された平均は線形演算によって推定された平均よりもSDFらしいことがわかる．また，図2より，従来の線形演算によって推定された平均形状では消失している橋屈などの微細な構造が，提案手法による推定では表現できていることが確認できた．

4. まとめ

本稿では，1次の図形モーメントとSDF間のL2ノルムを用いたエネルギー汎関数の最小化による平均形状の推定方法を提案した．また，本手法により推定した平均は，従来の線形演算によって推定した平均よりもSDFらしく，重要な形状を表現可能であることを確認した．今後は，admissible SDF variationを用いたSDFの共分散行列の推定を行い，統計的形状モデルの構築と3次元ラベルへの拡張を目指す.

謝辞

日頃から熱心にご討論いただく東京農工大学清水昭伸研究室の各位に感謝いたします.

利益相反の有無

なし

文　献

［1］ R. F. Gasser, “ATLAS OF HUMAN EMBRYOS,” Haper and Row, New York, 1975.

［2］ B.S. de Bakker, K.H. de Jong, J. Hagoort et al., “An interactive three-dimensional digital atlas and quantitative of human development,” Science, Vol.354, No.6315, 2016.

［3］ K. Kasahara, A. Saito, T. Takakuwa et. al, “A Spatiotemporal Statistical Shape Model of the Brain Surface during Human Embryonic Development”, Advanced Biomedical Engineering, Vol.7, pp.146-155, 2018

［4］ A. Saito, M. Tsujikawa, T. Takakuwa et. al, “Level set distribution model of nested structures using logarithmic transformation”, Medical Image Analysis, Vol. 56, pp.1-10, 2019

［5］ Siqi Chen, Guillaume Charpiat, Richard J. Radke, “Converting Level Set Gradients to Shape Gradients”, Computer Vision – ECCV 2010, Vol.6315, pp.715-728, 2010

［6］ Y. Yamaguchi, S. Yamada, “The Kyoto collection of human embryos and fetuses: history and recent advancements in modern methods,” Cells Tissues Organs 205, pp.314–319 , 2018

［7］ Y. Matsuda, S. Ono, Y. otake et al., “Imaging of a large collection of human embryo using a super-parallel MR microscope” Magnetic Resonance in Medical Sciences, Vol.6, pp.139-146, 2007.

［8］ N. Shiraishi, A. Katayama, T. Nakashima et al., “Morphology and morphometry of the human embryonic brain: A three-dimensional analysis,” NeuroImage, Vol.115, pp.96-103, 2015.

Estimation of a mean shape of embryonic human brains using admissible SDF variation

Yudai Tanabe[*1], Atsushi Saito[*1], Tetsuya Takakuwa[*2], Shigehito Yamada[*2], Akinobu Shimizu[*1]

*1 Tokyo University of Agriculture and Technology

*2 Kyoto University.

In this paper, we report a method for accurately estimating the average of shapes represented by the Signed Distance Function (SDF).Because the result of a linear operation between SDFs is not guaranteed to be SDF, the mean shape may not represent the fine anatomical structures. The proposed method estimates the average shape by minimizing the energy functional using the L2 norm between the first order figure moments and the SDF. By using admissible SDF variation in minimizing the energy function, we ensure that the estimated mean is SDF. Finally, we apply the proposed method to the brain label of embryos in the Kyoto collection and discuss the usefulness of the proposed method in comparison with the average shape estimated by conventional methods.

Key words: human embryo, statistical shape model, estimation of mean, energy functional, admissible SDF variation

内視鏡と連動した食道手術ナビゲーションシステムの開発

佐藤　由季也[*1]　上里　昌也[*2]　吉村　裕一郎[*1]

米本 昇平[*2]　浦濱 竜馬[*2]　加藤 直菜美[*2]　中口　俊哉[*1]

要旨

本研究では，非侵襲的な手術である NOTES(Natural Orifice Translumenal Surgery)を安全に実施するためのナビゲーションシステムを提案する．事前に撮影された CT 画像から食道壁外の臓器または組織を選択して仮想空間上の解剖モデルを構築する．術中の実内視鏡操作に連動して仮想空間上の内視鏡も移動させることで，内視鏡カメラから観察できる食道の壁外の状況を提示することを目指す．実内視鏡操作によるスコープの挿入深さは，患者の口元にカメラを設置して撮影したスコープの映像情報から得ることとした．本研究では深層学習に基づいてスコープ表面の模様を検出し，その移動量を画像計測により算出する．実験では食道を模した管と，その周囲に臓器を模したファントムを配置し，本システムが提示した仮想映像と実内視鏡が撮影した実験映像の一致度を検証した．実験の結果，映像上での一致度は 86%を示し，スコープの挿入の深さによらず安定した高い精度を実現した．

キーワード：NOTES，仮想内視鏡，食道手術支援，SegNet

1. はじめに

近年，世界的にも消化管外科領域で管腔内視鏡下・腹腔鏡下手術が盛んに行われている．また，その低侵襲性が高く評価されている．しかし，内視鏡，腹腔鏡それぞれが消化管内腔側と腹腔側のみからのアプローチであり，その境界である消化管壁を超えての手術は通常行われない．

2004 年 A.N.Kalloo[1]らにより自然開口経腔的内視鏡手術（NOTES：Natural Orifice Translumenal Endoscopic Surgery）が報告された・胃や膣などの人体に元から備わっている管腔臓器の壁を突き抜けて軟性内視鏡を腹腔内に挿入し，体表に傷創を作らずに胆嚢や虫垂に対して施術を行う新しい腹腔鏡手術である．しかし，NOTES は現時点で通常手術として行われず，その原因となる問題が 2 つある．1 つは正常管腔臓器の壁を人工的に穿孔させて内視鏡を目的部位へアプローチする手技のため，最終的に穿孔部位を完全に閉鎖する技術が必要であること[2]．もう 1 つは正常管腔臓器の壁を人工的に穿孔させる際に，その壁外に何があるのかを管腔内から透視できるかが重要となる．本研究は後者の課題に着目した．

本システムでは事前に撮影した患者の CT 画像から仮想空間上で 3 次元モデルを作成し，仮想空間上でも同様に内視鏡を食道へ挿入する．仮想内視鏡は実内視鏡の動きに伴って移動しており，実内視鏡と同じ視野で食道内外の解剖構造を提示する．仮想表示においては食道壁を半透明にすることで，食道外の臓器・組織を観察することが出来る．本システム使用時には二つのディスプレイを使用し，メインディスプレイには通常通り実内視鏡の映像を提示し，サブディスプレイには仮想内視鏡の映像を提示する．二つの映像を比較することで実内視鏡の位置把握を容易にすることが出来る．

*1 千葉大学院融合理工学府
[〒 3863-8522　千 葉 市 稲 毛 区　1-33]
sato.yukiya@chiba-u.jp
*2 千葉大学医学部　先端応用外科学

2. 提案システム

術前の準備手順として，患者の CT 撮影を事前に行い， CT 画像から仮想空間上に表示する臓器モデルを作成する．術中においては，まず実内視鏡と仮想内視鏡の初期位置の同期を施術医によって手動で行う．初期位置設定後は実内視鏡の相対的な移動量（挿入距離）をリアルタイムに計測し，仮想内視鏡を連動して移動させる．仮想内視鏡視点の CT 解剖情報を 3 次元コンピュータグラフィクスにより生成し，モニタに提示する．

2.1 3D 臓器モデルの作成

ここで「投影」とはサブディスプレイへの表示を表す．初めに投影対象とする臓器の 3 次元モデルの作成を行う．関心臓器の 3 次元モデルの作成には患者の CT 画像を用いる．投影対象とする 3 次元モデルの作成には， 3 次元画像解析ソフトウェア「3DSlicer」を用いてセグメンテーションを行う．

上記で作成した 3 次元臓器モデルを統合させ，仮想空間に取り込む．作成したモデルの編集にはフリー3DCG ソフトである「Blender」を使用し，食道の 3 次元臓器モデルのアルファ値を 30%に調整し，位置合わせを患者の CT 画像を基に手動で行っている． また，食道は術中に食道内に内視鏡を挿入しやすくするためにガスを注入する必要があるため，食道形状が術前と術中で異なる．食道モデルはこれを解決するために独自で開発した術中の食道を CT 画像からスライスごとに手動で中心点を選ぶことで術中の食道を再現するソフトウェアを用いた，位置合わせは事前に 3DSlicer で食道の始点も同時に出力し，その位置を参照することで食道モデルを適切な位置に配置することができる．また，食道モデルと重複した食道外の組織は削除した．

取り込み先の仮想空間の作成には 3DCG 開発プラットフォームである「Unity」を使用した．仮想空間に構築した 3 次元臓器モデルを図 1 に示す．Unity では C#言語によるスクリプトを記述することによって仮想空間内に内視鏡の配置や，移動などの処理をシミュレーションすることが出来る．仮想内視鏡視点の映像例を図 2 に示す．食道壁を透過させることで壁外の臓器配置を把握することができる．

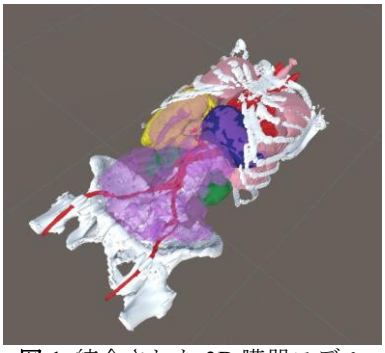


図 1 統合された 3D 臓器モデル

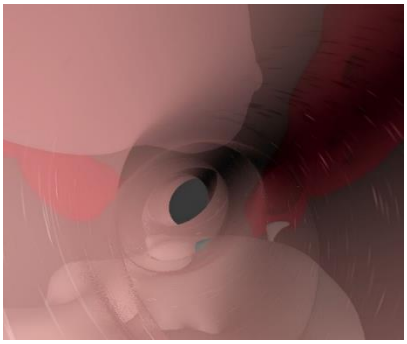


図 2 仮想内視鏡の映像例

2.2 術中処理

つぎに内視鏡挿入時の処理について述べる．仮想内視鏡と実内視鏡の位置の同期を行う際に初期位置は術者が手動で同期を行う必要がある．

初期位置設定において両内視鏡が対応をとるための目印が求められる．本システムでは目印として実際に手術にも解剖学的な目印として用いられる声帯を利用した．

初期位置の同期を行った後は，仮想内視鏡と実内視鏡の動きを連動させる必要がある．食道は他の臓器と比較して枝分かれの無い単純な構造をしており，実内視鏡の移動もほとんど直進に限られるという特徴がある．



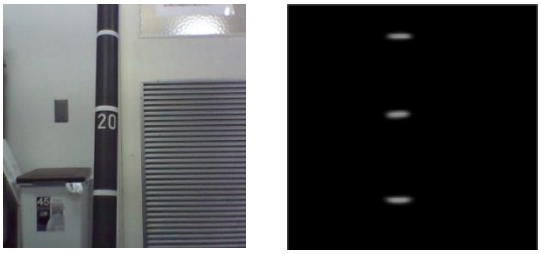
図 3 外部カメラによる内視鏡撮影例



図 4 内視鏡軟性部上の白帯

よって本システム使用時は実内視鏡は極力首振り運動をせず，食道の前方方向を常に直視した状態で挿入することとし，実内視鏡の挿入距離に関する情報を取得することが出来れば仮想内視鏡も同様に仮想食道モデルの中心線上を直進することで両内視鏡の視野を一致させることが出来る．

実内視鏡の挿入距離を測定するには，実内視鏡の軟性部上の等間隔の白帯を使用し，外部カメラを患者の口腔付近に配置し内視鏡が映像中で垂直になるようにする．配置例を図 3 に示す．内視鏡上の目盛りの例を図 3 に示す．映像中からの白帯を検出する手法として，深層学習モデルの一つであるエンコーダとデコーダをもつオートエンコーダ[3]構造を利用した SegNet[4]をベースにしたモデルによって，動画



(a) (b)

図 6 白帯検出結果例

フレーム間の白帯位置を解析し，実内視鏡の挿入距離を計測する．

2.3 白帯の自動検出

内視鏡軟性部上の白帯を検出するために SegNet をベースにしたセマンティックセグメンテーション用途のモデルを利用した．実際のモデルは当システムのリアルタイム性を増すためにコンボリューション層とアップサンプリング層を一層ずつ減らした．使用したアーキテクチャを図 5 に示す．

学習時には実際に用いる外部カメラで内視鏡軟性部を撮影した画像を 150 枚と，汎用性向上のために手術室の画像背景に内視鏡軟性部の画像を前景として人工合成した画像 150 枚，総計 300 枚の画像をデータセットとして利用した．学習は損失関数として binary cross entropy, 勾配化法として adadelta，学習回数は 100 エポック，バッチサイズ 4, GPU は NVIDIA GeForce GTX970 を使用した．また過学習を防ぐために Dropout[5]を各コンボリューション層に 0.2 の割合で適用した. 白帯検出結果例を図 6 に示す. 図 6(a)で入力画像を示し，図 6(b)にセグメンテーション結果を示す．

3. 性能評価

3.1 評価方法

提案システムを評価するために実験ファン

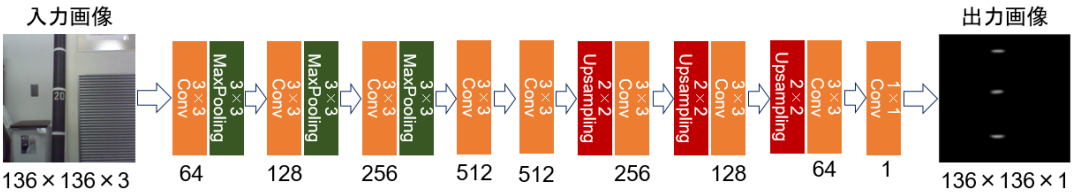
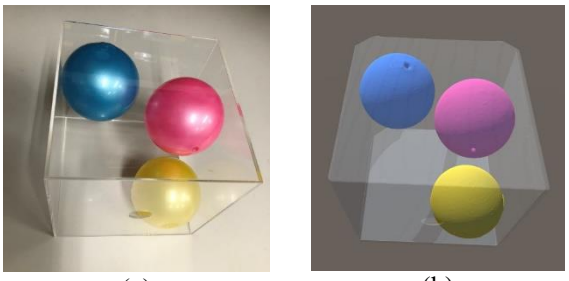


図 5 使用したアーキテクチャ



(a) (b)

図 7 実験用ファントム

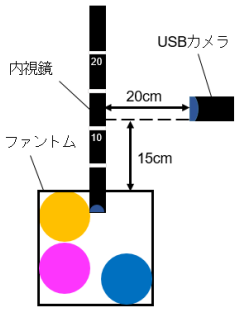


図 8 実験環境

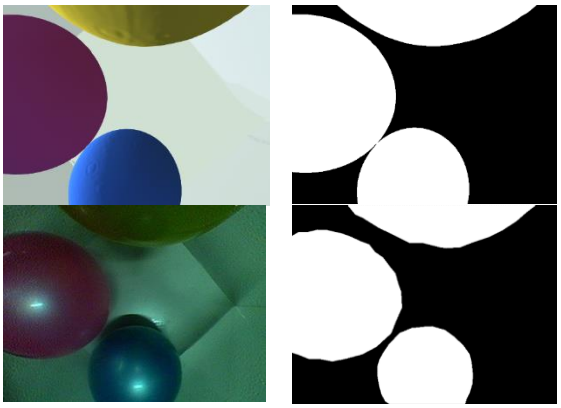


図 9 仮想内視鏡(上部)と実内視鏡(下部)の撮影像と模擬臓器の領域

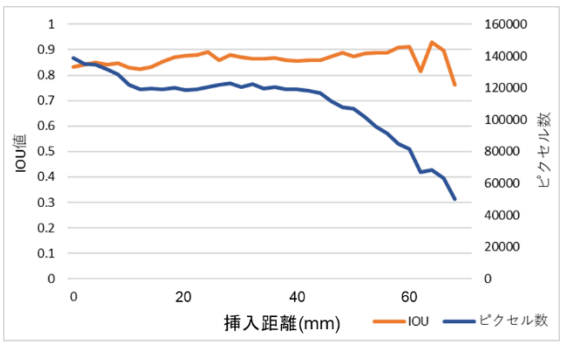


図 10 挿入距離と IOU 値

トムを作成した．ゴムボールを適当な場所に配置し，それを投影したい臓器とみなす．ファントムの上部には食道の入口を模した直径 20mm の穴があり，そこから経口内視鏡を挿入する．当ファントムを CT 撮影し 3 次元モデルを作成する．実験用ファントムの実物と 3 次元モデルを図 7 に示す．図 7(a)に実際の実験用ファントムを示し，図 7(b)に実際の実験用ファントムの CT 画像から作成した 3 次元モデルを示す．図 8(a)のように実内視鏡ではゴムボールが写った映像が取得され，本システムの出力結果として図 7(b)のように仮想内視鏡においても同様な映像が取得される．挿入できる最大距離である 70mm まで 2mm 刻みで両内視鏡から画像をキャプチャし，同じ挿入距離における画像同士のゴムボールの見え方を比較する．見え方の比較は両内視鏡の画像においてゴムボールの領域のセグメンテーションを領域拡張法もしくは手動によって行い，両画像の IOU 値を評価指標とした．

3.2 実験環境

実験環境を図 8 に示す．実験にはオリンパス社製の経口内視鏡 EG-450HR を使用した．本実験は基礎実験の位置づけにより実内視鏡が常に垂直に挿入されるよう固定し，2mm ずつ実内視鏡を挿入してた．さらに実内視鏡の軟性部上の白帯を観測できる位置に USB カメラを設置した．

3.3 結果と考察

実内視鏡と仮想内視鏡によって取得される画像と領域拡張法を施した後に手動で調整した模擬臓器領域の例を図 9 に示す．各挿入距離の模擬臓器領域画像間の IOU 値の遷移と実内視鏡の撮影画像における模擬臓器領域のピクセル数を図 10 に示す．IOU 値の平均値は 0.862 となった．

挿入距離に伴う誤差の影響によって IOU 値の低下を予期していたが，最大挿入距離まで大きな変動が無かったことから仮想内視鏡，実内視鏡がともに連動した映像を取得できていることが考えられる．IOU 値が最大挿入距離付近で大きく増減しているが，模擬臓器領域のピク

セル数が最大挿入距離付近では大きく減少し，それが IOU 値の増減への影響が大きくなることが原因として考えられる．

4. まとめ

本研究では，内視鏡操作と連動したナビゲーションシステムを提案した．本システムの流れは二段階に分かれており術前に投影対象の 3 次元臓器モデルの作成，術中は仮想内視鏡と実内視鏡の連動を行う．実験用ファントムを作成して本システムの評価を行った結果，システムが出力する仮想内視鏡視野の映像は実内視鏡視野の映像に高い精度で一致していることを確認した．

現行のシステムでは内視鏡の向きが映像中で垂直でなければいけないが，実用性向上のためにも映像中の向きによらない映像計測技術まで発展させる必要がある．他にも，映像からカメラの移動と回転を推定するビジュアルオドメトリの技術と組み合わせて使うことでより当システムの発展に寄与できるのではないかと考えている．

利益相反の有無

なし

文　献

［1］ Flexible transgastric peritoneoscopy: a novel approach to diagnostic and therapeutic interventions in the peritoneal cavity. Kalloo AN, Singh VK, Jagannath SB, Niiyama H, Hill SL, Vaughn CA, Magee CA, Kantsevoy SV.Gastrointest Endosc. 2004 Jul;60(1):114-7.PMID: 15229442

［2］ NOTES が拓く内視鏡治療の将来(総説) 田尻 久雄(東京慈恵会医科大学 内視鏡科), 北野 正剛, 炭山 和毅, 池田 圭一, 安田 一弘 Gastroenterological Endoscopy (0387-1207)52 巻 12 号 Page3259-3266(2010.12)

［3］ G. E. Hinton et al. "Reducing the Dimensionality of Data with Neural Networks" Science Vol. 313, Issue 5786, pp. 504-507, 2006

［4］ Alex Kendall et al. "SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation" CVPR, 2015

［5］ N. Srivastava et al. “Dropout: A simple way to prevent neural networks from overfitting” The Journal of Machine Learning Research pp.1929-1958 2014

Development of an esophageal surgery navigation system linked to an endoscope

Yukiya SATO[*1], Toshiya NAKAGUCHI[*1], Masaya UESATO[*2], Yuichiro YOSHIMURA[*1]
Ryuma URAHAMA[*2], Manami KATO[*2], Shohei YONEMOTO[*2]

*1 Graduate School of Science and Engineering Chiba University
*2 Department of Frontier Surgery, Chiba University Graduate School of Medicine.

We propose a navigation system for safely performing Natural Orifice Translumenal Surgery (NOTES), a non-invasive surgery. We aim to present the situation outside the wall of the esophagus in the virtual space by moving the endoscope in the virtual space in conjunction with the actual endoscope operation during surgery. The insertion distance of the scope through the actual endoscopic operation was obtained from the video information of the scope captured by a camera installed at the mouth of the patient. In this study, we detect patterns on the surface of a scope based on deep learning and calculate the amount of movement by image measurement. In the experiments, a tube that imitates the esophagus and a phantom that imitates an organ were placed around the tube, and the agreement between the virtual images presented by this system and the experimental images taken by the real endoscope was verified. In our experiments, a tube that mimics the esophagus and a phantom that mimics an organ were placed around the tube, and the coincidence between the virtual image presented in this system and the experimental image taken by the actual endoscope was verified, and the coincidence was 86%.

Key words: NOTES, Virtual Endoscope, Esophageal Surgery Support, SegNet

多層構造物測定のための深さ超解像 OCT

椴山　誉[*1, *2, *3]　佐々木　芳彰[*2]　吉峯　功[*2]　永野　繁憲[*1]

湯浅　哲也[*3]　大谷　知行[*2]

要旨

光干渉断層計（OCT）は，医療用途だけでなく産業向けの検査技術としても研究開発が行われている．検査応用で主流となってるフーリエドメイン OCT（FD-OCT）の信号処理では，逆高速フーリエ変換（IFFT）を用いて測定した干渉信号を深さプロファイルに変換する．しかし，この従来解析法では光源の帯域幅と中心波長によって深さ方向の分解能が制限される．本研究では，FD-OCT の一種であるスウェプトソース OCT（SS-OCT）を用いた多層構造物の測定における深さ分解能を改善するため，アニヒレーティングフィルタ法をベースにした深さ超解像解析法を適用し，分解能の改善効果を検証した．プラスチック製の段差サンプルを用いて測定を行った結果，厚さ 1 mm の表裏面を分離することが出来た．これは，従来の IFFT 解析の 1/3 の深さ分解能を達成したことに相当する．また，厚さ測定の再現性は 46 μm であった．

キーワード：光干渉断層計，超解像，深さ分解能，多層構造物

1. はじめに

光干渉断層計（OCT）は非接触かつ低侵襲で人体器官の三次元画像を取得できることから，主に医療検査技術として応用研究が進んできた. また，OCT はその有用性の高さから医療用途のみならず産業向けの非破壊検査技術としても利用が検討されている．特にタイムドメイン OCT と比較して SN 比で有利なフーリエドメイン OCT（FD-OCT）は，検査応用で主流な方式となっている[1,2]．従来の FD-OCT の信号処理では，逆高速フーリエ変換（IFFT）を用いて測定した干渉信号（インターフェログラム）を深さプロファイルに再構成する．しかし，この従来解析法では光源の帯域幅と中心波長によって深さ方向の分解能が制限される．

本研究では，FD-OCT の一種であるスウェプトソース OCT（SS-OCT）を用いた多層構造物の測定における深さ分解能を改善するため，アニヒレーティングフィルタ（AF）法をベースにした深さ超解像解析法を適用した. その結果を従来の IFFT 解析による結果との比較し，深さ分解能の改善効果について議論する．

*1 ㈱トプコン

〔〒174-8580 東京都板橋区蓮沼町 75-1〕

e-mail: hmomiyama@topcon.com

*2 国立研究開発法人理化学研究所光量子工学研究センター（RAP）

*3 山形大学大学院理工学研究科バイオ工学専攻

投稿受付：2020 年 7 月 13 日

2. SS-OCT システムの構成

周波数範囲 75–110 GHz の可変周波数光源とマイケルソン干渉計を用いて図 1 に示すような SS-OCT システムを構築した．光源の周波数掃引

は GPIB 経由で PC から制御した．光源から出射されたビームをシリコンビームスプリッタ（Si-BS, 抵抗値 8-12 kΩ･cm, 分配比 50:50）で二つの光路に分割し，片方はリファレンス側のミラーに，もう一方はサンプルに照射し，各反射光によって生じるインターフェログラムをカロリーメータ（PM4, Virginia Diodes Inc.製）で検出した．検出した信号は DAQ を介して PC に保存した．また，サンプルは 3 軸の自動駆動ステージに取り付けて，PC からスキャンの制御を行った．

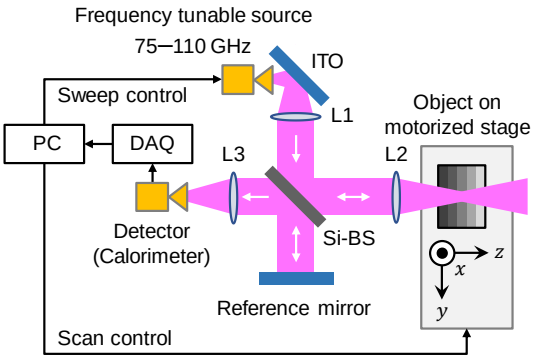


図 1　SS-OCT システムの概略図．（L1–3: 凸レンズ，ITO: 酸化インジウムスズの膜付きガラス，Si-BS: シリコンビームスプリッタ，DAQ: データ収集システム）

3. 深さ超解像解像解析法

OCT における深さ分解能は，撮像特性を決定する重要なパラメータの一つである．従来、SS-OCT の信号処理では，光源の周波数掃引によって得られるインターフェログラムを IFFT することで深さプロファイルに再構成する（A-scan）．光源のスペクトルがガウス関数の場合，深さ分解能Δz は以下の式で与えられる[3,4].

$$\Delta z = \frac{2\ln 2}{\pi}\frac{{\lambda_0}^2}{n\Delta\lambda} \tag{1}$$

ここで，Δλは光源の帯域幅（半値全幅，FWHM），λ_0は中心波長，n は測定対象物の屈折率であり，これら三つの要素によって深さ分解能が決定する．

本研究では深さ分解能を改善するため， AF 法とノイズ除去法を組み合わせた深さ超解像解析方法（DSRA）を適用した[5,6]．ノイズ除去処理は，大きく二種類の方法に分けられる．一つはバンドパスフィルタ（BPF）であり，多重反射や外乱光の干渉によって生じる周期的なノイズを除去するために使用した．このようなノイズは深さプロファイルにおいて，サンプルの位置以外の領域にもピークとして現れる．そのため，サンプル位置付近のみをパスするフィルタリング処理を行った後，FFT でインターフェログラムに再変換することで，測定対象物周辺の深さ情報を抽出し，ノイズを除去した．もう一つは反復の特異値分解（SVD）であり，ランダムな加算性ホワイトガウシアンノイズを除去するために使用した．ノイズを含むインターフェログラム$\tilde{I}_\kappa$から以下に示すような$(K - N + 1)\times N$ のテプリッツ行列を生成する．

$$\tilde{\mathbf{I}} = \begin{pmatrix} \tilde{I}_{N-1} & \tilde{I}_{N-2} & \cdots & \tilde{I}_1 & \tilde{I}_0 \\ \tilde{I}_N & \tilde{I}_{N-1} & \cdots & \tilde{I}_2 & \tilde{I}_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \tilde{I}_{K-2} & \tilde{I}_{K-3} & \cdots & \tilde{I}_{K-N} & \tilde{I}_{K-N-1} \\ \tilde{I}_{K-1} & \tilde{I}_{K-2} & \cdots & \tilde{I}_{K-N+1} & I_{K-N} \end{pmatrix} \tag{2}$$

ここで，κは波数番号，K は波数の数，N は分解する要素数であり，$\tilde{\mathbf{I}}$の SVD は以下のように書ける．

$$\tilde{\mathbf{I}} = \mathbf{USV}^T \tag{3}$$

ここで，**U** と **V** はそれぞれ$(K-N+1)\times N$ と $N\times N$ ユニタリ行列であり，**S** は以下に示す特異値を含むの対角行列である．

$$\mathbf{S} = \begin{pmatrix} \sigma_1 & & & 0 \\ & \sigma_2 & & \\ & & \ddots & \\ 0 & & & \sigma_N \end{pmatrix}, \tag{4}$$

$$\sigma_1 \geq \sigma_N \geq \cdots \geq \sigma_{N-1} \geq \sigma_N$$

得られた特異値から信号成分とみなせる R 個の上位要素のみを保持し，ノイズ成分である残りの要素をゼロフィリングして，**S**’を生成する．そして，**S**’から$\tilde{\mathbf{I}}'$を生成するという処理を繰り返してノイ

ズ成分を除去した．また，以下の式から評価値 E を算出し，E が閾値以下になることを停止条件とした．

$$E = \frac{\sigma_{R+1}}{\sigma_R} \tag{5}$$

ここで，$0 < R < N$ である．これらのノイズ除去により，ノイズレベルが高いデータの解析における AF 法の推定性能向上が期待できる．一般的に AF 法は，少ないデータ数からスパースなオリジナル信号を復元するために使用される．SS-OCT のデータ解析においては，対象物の境界面数を仮定することで，インターフェログラムから深さプロファイルを再構成できる．いま区間的に屈折率が一定な多層構造サンプルを，SS-OCT で測定した場合を考える．サンプルの減衰，多重反射および内部の界面同士の干渉を無視できると仮定すると，測定したインターフェログラム I_κは以下のようにモデル化される．

$$I_\kappa = \sum_{l=1}^{L} (A_l \gamma_l^\kappa + A_l^* \gamma_l^{*\kappa}), \tag{6}$$

$$A_l = a_l e^{i2\pi b_l k_{\min}},$$

$$\gamma_l = e^{i2\pi b_l \Delta k}$$

ここで，$a_l = (n_{l+1} - n_l)/(n_{l+1} + n_l)$は各界面の反射係数, $b_l = \left[\sum_{p=1}^{l} (z_p - z_{p-1})\, n_p\right]$は光学的距離の和，$L$ は界面の数，z_p は幾何学的距離，n_p は各層の屈折率，k は光源の波数，Δk は波数間隔，$k_{\min}$ は最小波数を表す．γ_lとγ_l^*を根とする多項式が

$$\prod_{l=1}^{L} (t^{-1} - \gamma_l^{-1})(t^{-1} - \gamma_l^{*-1}) = \sum_{j=0}^{2L} p_j t^{-j} \tag{7}$$

インターフェログラムと同じ根をもつとき，係数 p_jとインターフェログラムの畳み込みは0となる．

$$\sum_{j=0}^{2L} p_j I_{\kappa - j} = 0 \tag{8}$$

この性質によって，インターフェログラムから係数 p_jが得られる．さらに，多項式(7)の根を求めることで $2L$ 個のγ_lとγ_l^*を得る．b_lは複素数γ_lの偏角φ_lから次式で得られる．

$$b_l = \frac{\varphi_l}{2\pi \Delta k} \tag{9}$$

また, A_l と A_l^*は測定したインターフェログラムと信号モデルに関する線形最小二乗問題を解くことによって得られる．得られたパラメータ A_l, A_l^*, γ_lおよびγ_l^*から式(6)を用いて各界面とリファレンスミラーの干渉によるインターフェログラムを再構成し，中心周波数における振幅値を得る．

4. 従来法における深さ分解能

従来の IFFT 解析における深さ分解能を確認するため，サンプルとして平面ミラーを用いて A-scan を行った．A-scan では，75–110 GHz の範囲を約 60 MHz 間隔で周波数掃引し，その時の強度を測定することでインターフェログラムを取得した．サンプルミラーは焦点位置付近から±50 mm の範囲を 5 mm 間隔で移動させ各距離で繰り返し 10 回測定を行った．前処理として，リファレンススペクトルによる背景除去処理とガウス関数（λ_0 = 3.24 mm, $\Delta\lambda$ = 0.619 mm）によるアポダイゼーションを行った．この場合の深さ分解能の理論値は，式(1)から空気中（n = 1）で Δz=7.49 mm である．また，深さプロファイルのデータ間隔を小さくするため, 測定された 585 点のデータを 16384 点になるようにゼロフィリングした．その後，IFFT によって深さプロファイルに再構成した．再構成された各深さプロファイルをガウス関数でフィッティングすることで，深さ分解能（FWHM）を求めた．図 2(a)は，サンプルミラーが焦点位置における深さプロファイルの測定値とフィッティングの結果であり，ガウス関数に近い形状が得られた．そして，全データのフィッティング結果から深さ分解能の平均値と標準偏差を求めると 7.65±0.02 mm（n = 1）であり，理論値 7.49 mm に対して+2.1%とほぼ一致する．図 2(b)はアポダイゼ

ーションなし（Boxcar 窓）の場合の深さプロファイルであり，深さ分解能は 5.09±0.03mm（n = 1）であった．ガウス関数によるアポダイゼーションを行った結果と比べてサイドローブが大きくなるが，高い分解能が得られる．また，図 4 はガウス関数でフィッティングしたピーク位置の平均値とミラーの移動量の関係を表しており，ゲイン誤差 0.19%とほぼ線形な特性が得られた．

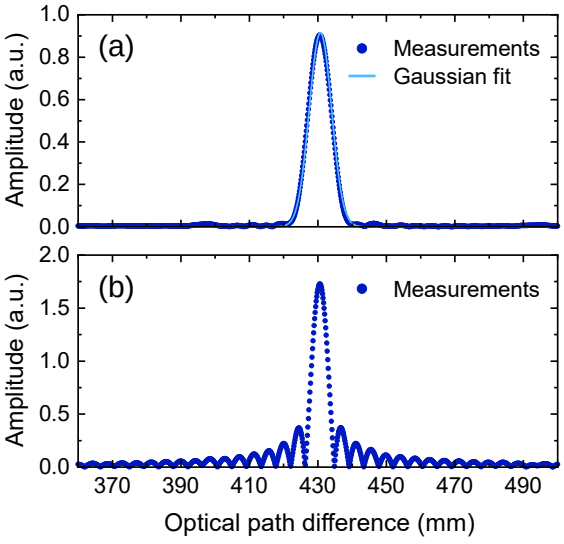


図 2　(a)ガウス関数でアポダイゼーションした場合と(b)アポダイゼーションなしの場合の IFFT 解析によるサンプルミラーの深さプロファイル．

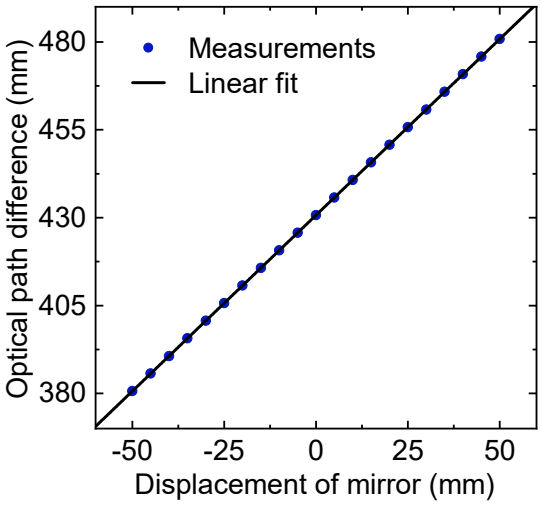


図 3　サンプルミラーを深さ方向に移動した場合のリニアリティ．

5. 実験方法

DSRA の深さ分解能を評価するため，対象物として図 4 に示すような段差構造を持つプラスチック製のサンプルを用いた．段差（厚さ）の設計値は 1, 2, 3, 5, 10 mm の 5 種類であり，材質は高密度ポリエチレン（HDPE，比重 0.96）を使用した．リファレンスとなる各段差の実際の厚みを確認するため，接触式三次元測定機（CMM, CRYSTA-Apex S574, Mitutoyo 社製）を用いてサンプルの段差付近 80×180 mm の範囲を 5 mm 間隔で測定した．その結果を図 5 に示す．各段の中心付近 25×50 mm の厚みの平均値と標準偏差は 0.88±0.03，1.89±0.03，2.86±0.02，4.82±0.01，9.91±0.04 mm であった．また，サンプルの屈折率は，設計厚さ 10 mm を A-scan して得られた光学的な厚み 15.3 mm と CMM で測定した実際の厚み 9.91 mm から算出し，n = 1.54 であった．この結果は，これまでの報告とほぼ一致する結果であり[7]，この場合のサンプル中におけるアポダイゼーションなしでの IFFT 解析による深さ分解能は 3.29mm となる．

SS-OCT システムを用いて，この段差サンプルの各段差の中心付近を繰り返し 10 回 A-scan した．

得られたインターフェログラムのノイズを除去するため, 前述の二種類のノイズ除去を行った. BPF では，サンプルの設置位置付近 320–538 mm をパス範囲とした．そして，SVD の条件として分解する要素数 $N = 200$, 残す要素数はサンプル面数の二倍の $R = 4$，閾値 $E = 1\times10^{-8}$ を使用した．また，AF 法の条件として面数 $L = 2$ を使用した.

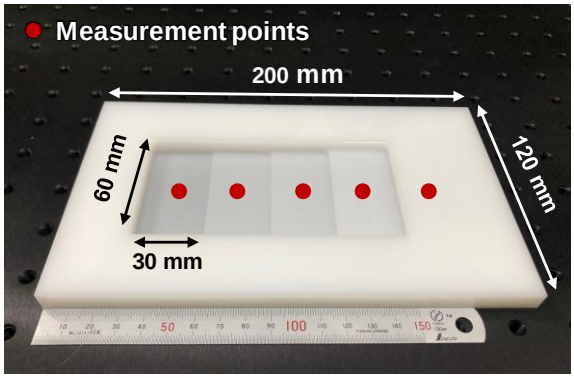


図 4　段差サンプルの外観および寸法.

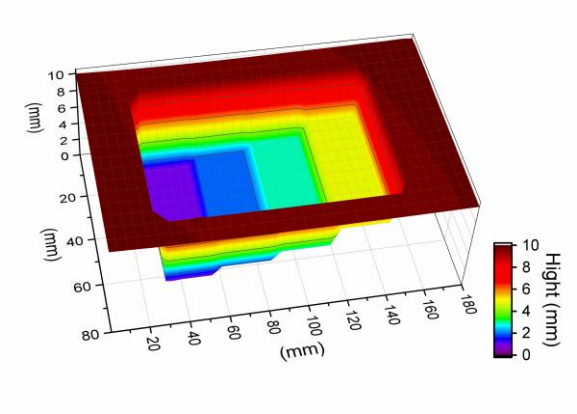


図 5　CMM で計測した段差サンプルの形状.

6. 結果及び考察

各段差におけるノイズ除去前とノイズ除去後のインターフェログラムを図 6 に示す．この結果から，ノイズ除去処理を用いることでサンプルの隣接する二つ界面によって生じるビート信号をクリアに抽出できていることが確認できる．このノイズ除去されたインターフェログラムを AF 法によって深さプロファイルに再構成し，アポダイゼーションなしでの IFFT 解析結果と比較した（図 7）．その結果，DSRA では厚さ 1 mm の表面と裏面を分離することが出来た．この結果は，従来の IFFT 解析の 1/3 の高い分解能を達成したことに相当する．図 7 に示したように，DSRA によって再構成される深さプロファイルはデルタ関数列として得られるため, 分解能が無限に小さいように見える．しかし，実際には DSRA の深さ分解能は，繰り返し測定した場合のピーク位置の再現性で制限されると考えられる．DSRA の分解能のポテンシャルを評価するため，厚さ 1 mm における厚さの再現性を確認した．結果，再現性は 46 μm（FWHM）であり，これは従来の IFFT 解析の 1/72 の非常に高い分解能に相当する．一方で，IFFT の結果に注目すると, IFFT 解析の分解能は 3.29 mm にも関わらず，3 mm 以下の厚みを分離できているかのような二つのピークが確認できた．これは，超解像領域で生じるフリンジの影響と考えられる[8]．そこで，どちらの結果がより正確であるか確認するため，表裏面の二つのピークから厚みに換算し，CMM で測定したリファレンスの厚みと比較した．その結果を図 8 に示す．IFFT によって得られた厚みは，3 mm 以下で値がほぼ変わっていなかった（slope = 0.05）．一方で，DSRA によって得られた厚みは，3 mm 以下でもほぼリニアな変化が確認できた（slope = 0.92）．従って，超解像領域において DSRA の方が IFFT より正確な測定が可能といえる.

以上のように DSRA による深さ分解能の改善効果を実証したが，実際の測定においては対象物の境界面数情報が不十分な場合もある．そのような場合において，統計学的な手法を用いて最適なモデルを推定することが必要であり、その方法を検討することが今後の研究課題である.

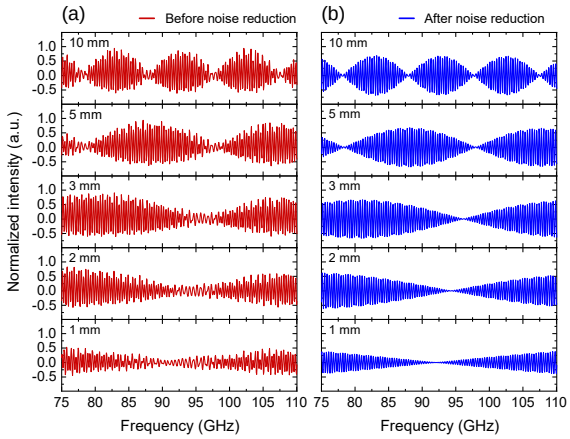


図 6　段差サンプルの各段差（厚さ 1, 2, 3, 5, 10 mm）における(a)ノイズ除去前と(b)ノイズ除去後のインターフェログラム.

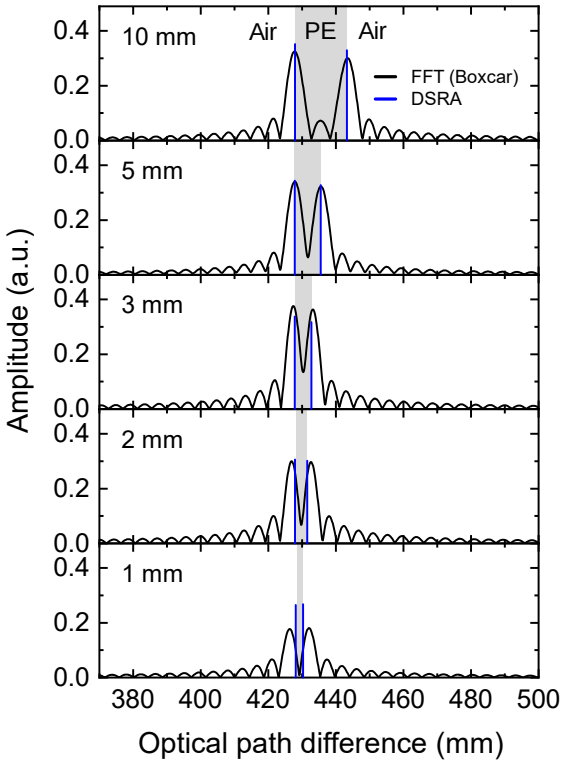


図 7　段差サンプルの各段差（厚さ 1, 2, 3, 5, 10 mm）における深さプロファイル.

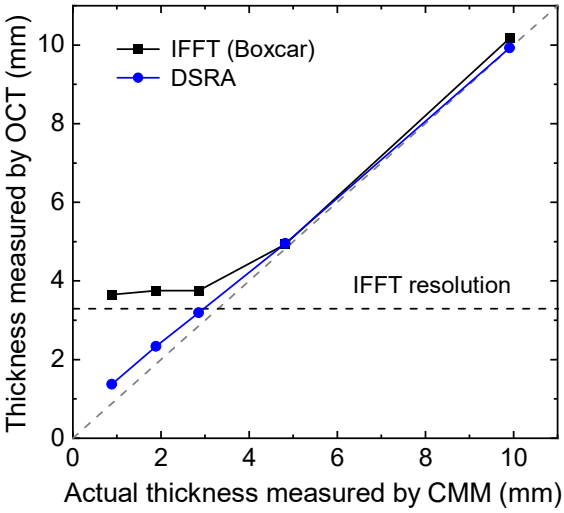


図 8　IFFT と DSRA の厚みの比較. 黒の点線は IFFT による深さ分解能（3.29 mm）.

7. まとめ

本研究では，多層構造物の測定を目的とした周波数 75–110 GHz の SS-OCT システムを構築し，信号処理に深さ超解像解析法である DSRA を適用した．本システムの深さ分解能を評価するため，プラスチック製の段差サンプルを用いた測定を行い，厚さ 1 mm のプラスチックサンプルの表裏面を分離できることを確認した．これは，従来の IFFT 解析による分解能の 1/3 の高い分解能を達成したことに相当する．本研究結果は，SS-OCT による多層構造物測定の性能向上に寄与すると考えられる.

謝辞

本研究を遂行するにあたり有益なご助言を賜った緑川克美先生，南出泰亜先生（理化学研究所），時実悠先生（徳島大学），福間康文博士，森口祥聖博士（トプコン）ならびにテラヘルツイメージング研究チームの皆様に深く感謝いたします．本研究は理研トプコン共同研究の一環として実施したものです.

利益相反の有無

なし.

文　献

［1］ R. Leitgeb, C. K. Hitzenberger, and A. F. Fercher: Performance of fourier domain vs time domain optical coherence tomography. Opt. Express **11**(8), 889–894, 2003

［2］ N. Nassif, B. Cense, B. Hyle Park, et al.: In vivo human retinal imaging by ultrahigh-speed spectral domain optical coherence tomography. Opt. Lett. **29**(5), 480–482, 2004

［3］ W. Drexler and J. G. Fujimoto: Optical Coherence Tomography: Technology and Applications, 2nd ed., Springer, 2015

［4］ S. Yun, G. Tearney, J. de Boer, et al.: High-speed optical frequency-domain imaging. Opt. Express **11**(22), 2953–2963, 2003

［5］ H. Momiyama, Y. Sasaki, I. Yoshimine, et al.: Improvement of the depth resolution of swept-source THz-OCT for non-destructive inspection. Opt. Express **28**(8), 12279–12293, 2020

［6］ M. Vetterli, P. Marziliano, and T. Blu: Sampling of signals with finite rate of innovation. IEEE Trans. Signal Process. **50**(6), 1417–1428, 2002

［7］ R. Piesiewicz, C. Jansen, S. Wietzke, et al.: Properties of Building and Plastic Materials in the THz Range. Int. J. Infrared Millimeter Waves **28**(5), 363–371, 2007

［8］ V. Aranha dos Santos, L. Schmetterer, G. J. Triggs, et al.: Super-resolved thickness maps of thin film phantoms and in vivo visualization of tear film lipid layer using OCT. Biomed. Opt. Express **7**(7), 2650–2670, 2016

脳波信号への深層学習適用による標的刺激の推定

長島　和馬[*1]　松村　泰志[*2]　姜　時友[*1]　石川　文之進[*3]

加我　君孝[*4]　深見　忠典[*1]

要旨

【目的】ブレインコンピュータインターフェイスの研究において，多数の選択肢の中からユーザが意図する一つ（標的刺激）を特定する方法の開発は極めて重要である．そこで，本研究では深層学習を用いて,聴覚オドボール課題で計測された単一の脳波応答が,標的刺激と非標的刺激のいずれであるかの判別を行う.【方法】被験者 76 名に対して,オドボール課題を施行し,10 ／ 20 国際電極配置法に基づく 19 電極において約 10000 応答を得た.これらの応答のうち 約 8000 応答を訓練データ,約 2000 応答をテストデータとした.深層学習において CNN と時間方向と空間方向へのフィルタリングを組み合わせた EEGnet を用い,正答率の比較を行った.【結果】正答率は CNN で 80.0%,EEGnet で 85.0%であった.【結論】単一の脳波応答における分類において,CNN よりも EEGnet でより高い正答率が得られた.

キーワード：オドボール課題, 標的刺激, 脳波分類, CNN, EEGnet

1. はじめに

1.1 研究背景

脳波は自発的に出現する脳波(自発脳波)に加え,筋電図や眼球運動といったアーチファクトが多く含まれるため,前処理なしの計測信号を解析するのは困難である.そのため,標的刺激により得られる多数回の応答に対して加算平均法を用いてアーチファクトを取り除く方法が一般的であるが,計測回数が多くなるにつれて計測時間が長くなるという問題がある.

1.2 研究目的

本研究では,深層学習を用いて,オドボール課題遂行時における標的と非標的刺激に対する応答の分類を多数の応答に対する加算平均を用いることなく単一応答で行うことを目的とする.

1.3 脳波

脳波(Electroencephalogram：EEG)には,自発脳波と誘発脳波の 2 種類があり,自発脳波は生きている間常に生じる脳波である.誘発脳波は音や光の刺激に反応して生じる脳波であり,事象関連電位(Event Related Potentials：ERP)も誘発脳波である． ERP は,「客観的に定義できる」事象に「時間的に」関連した電位のことであり,視覚,聴覚などで生じる感覚刺激に対して出現する.ERP の測定には,オドボール課題が広く用いられている[1].

ERP の中でも，P300 は,標的刺激呈示後約 300[msec]に出現する陽性の電位変化で，人間の判断や認知に関連している電位として知られている．よって，P300 を検出することで,刺激に対して注意を向けているかどうかを客観的に判断することができるが,P300 は被験者のモチベーションに大きく影響されるという特

*1 山形大学大学院　理工学研究科
〔〒992-8510 米沢市城南 4-3-16〕
e-mail: tcd58317@st.yamagata-u.ac.jp
*2 パクテラ デジタルイノベーション事業開発部
*3 報徳会宇都宮病院
*4 東京医療センター・感覚器センター

徴がある[2].

2. 手法

深層学習の手法として,本研究では CNN(畳み込みニューラルネットワーク)と,時間方向・空間方向へのフィルタリングを組み合わせた EEGnet[3]を用いた.

本研究で用いる脳波は,オドボール課題の標的および非標的刺激に対する脳波応答であり,表 1 に示す 2 種類の音を刺激として用いる.標的刺激(低頻度刺激)と非標的刺激(高頻度刺激)は 1 セットあたり 2 : 8 の割合で呈示する.
脳波計測は 2 セット行い,標的刺激呈示時にボタンを押すタスクを被験者に与える.

表 1 本実験における条件.

	標的刺激	非標的刺激
刺激の種類	2000Hz(高音)	1000Hz(低音)
タスク	ボタン押し	なし

これらの脳波データは,報徳会宇都宮病院で健常例に対して計測された. この脳波データに関しては,宇都宮病院倫理委員会の承認を得ている(倫委 第 18-04 号).今回,76 名の被験者に対しオドボール課題を施行し, 10 ／ 20 国際電極配置法に基づく 19 電極において約 10000 応答を得た.そのうち約 8000 応答を訓練データ,約 2000 応答をテストデータとした.

学習には, CNN(畳み込みニューラルネットワーク)と EEGnet を用い, 性能比較を行った. 前者は, 畳み込み層,プーリング層,全結合層からなるネットワークで,本研究でははじめに畳み込み層を 2 回適用し,プーリング層を適用した.さらにもう一度畳み込み層とプーリング層を適用し,最後に全結合層を適用している.今回,刺激呈示後 1[sec]間の脳波(19 電極)を用い,2 次元で表現したものを入力波形とした(図 1).一方,後者は, 基本的には CNN と同じ構成となっているが,畳み込み層で時間方向と空間方向へのフィルタリングが行われているという点で CNN と異なっている.入力波形は CNN と同様とした.

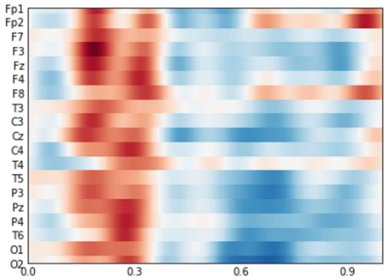


図 1 入力波形の例(縦軸:電極, 横軸:脳波).

3. 結果

それぞれのネットワークで学習を行った後,テストデータによる正答率を求めた.CNN で 80.0%,EEGnet で 85.0％であった.これは,EEGnet がより脳波の処理に適したネットワーク構造となっているためであると考えられる.

4. 結論

単一の脳波応答における分類において,CNN よりも EEGnet でより高い正答率が得られた.この結果から,深層学習を用いることでオドボール課題遂行時における標的と非標的刺激に対する応答の分類を多数の応答に対する加算平均を用いることなく単一応答で行うことができることが明らかになった.

利益相反の有無

無し

文 献

［1］ 入戸野宏：事象関連電位(ERP)と認知活動:工学心理学での利用を例に. 日本行動科学学会誌 **42**: 25-35, 2003

［2］ 加賀 佳美, 相原 正男：P300 基礎. 臨床神経生理学 **41**: 80-85, 2013

［3］ Vernon J Lawhern et al.,: EEGNet: a compact convolutional neural network for EEG-based brain‑computer interfaces. Journal of Neural Engineering **15**: 056013, 2018

Estimation of target stimulus by applying deep learning to EEG

Kazuma NAGASHIMA[*1], Taishi MATSUMURA[*2], Siu KANG[*1], Bunnoshin ISHIKAWA[*3], Kimitaka KAGA[*4], Tadanori FUKAMI[*1]

*1 Yamagata University Graduate School of Science and Engineering
*2 Pactera Digital Innovation Business Development
*3 Hotokukai Utsunomiya Hospital
*4 Tokyo Medical Center

[Objectives] In the study of brain computer interface(BCI), it is extremely important to develop a method for identifying the one (target stimulus) that the user intends to input from among many choices. Therefore, we used deep learning to determine whether a single EEG response measured in the auditory oddball paradigm is due to a target or a non-target stimulus. [Methods] The oddball paradigm was performed on 76 subjects, and about 10,000 responses were obtained at 19 electrodes based on the international 10/20 system. About 8,000 responses were used as training data, and about remaining 2,000 responses were used as test data. In the deep learning, the CNN and the EEGnet that combined filtering in the temporal direction and the spatial direction were used to compare the accuracies. [Results] The accuracies were 80.0% for CNN and 85.0% for EEGnet. [Conclusion] For classification in a single EEG response, EEGnet showed a higher accuracy than CNN.

Key words: oddball paradigm, target stimulus, EEG classification, CNN, EEGnet

患者情報の併用による肺細胞診画像の良悪性鑑別
－ 4 種類の分類アルゴリズムによる性能比較 －

山田　あゆみ[*1,2]　寺本　篤司[*1]　桐山　諭和[*3]　星　雅人[*1]
塚本　徹哉[*3]　今泉　和良[*3]　藤田　広志[*4]　齋藤　邦明[*1,2]

要旨
肺細胞診における良悪性鑑別は，病理医や細胞検査士にとって煩雑で労力を必要とする作業である．ここで，実際の診療プロセスにおいて，術者は細胞の形状変化のみならず種々の患者情報を参照して判断することが一般的である．本研究では液状細胞診画像と患者情報を用いた肺癌良悪性自動鑑別手法を開発し，基礎評価を行った．はじめに，Convolutional Neural Network のひとつである VGG-16 を用いて細胞診画像から 4096 個の画像特徴量を抽出した．次に，電子カルテから細胞診画像に対応する患者情報(年齢，性別，臨床検査結果等)を 27 種類収集した．得られた画像特徴量と患者情報から特徴選択を行い，機械学習を用いて良悪性の分類を行った．45 症例の臨床データを用いて，3 分割交差検証にて評価を行った結果，患者情報単体，画像情報単体での識別結果と比較して，それらを組み合わせて分類したときに識別精度が向上した．これらの結果から，提案手法の有用性が示唆された．
キーワード：肺癌，液状細胞診，良悪性鑑別，電子カルテ

1．はじめに

肺癌は世界的にみても癌死亡の主要な原因となっている[1]．肺癌の病理診断において，病理医の診断の前に細胞検査士(スクリーナー)が異常の疑われる細胞を拾い上げるため，最初の段階で偽陰性を減らすことは重要である．これまでに，細胞診画像を用いた肺癌良悪性自動鑑別手法が提案されているが[2-3]，臨床応用にはさらなる識別精度の向上が課題となっている．

ここで本研究では，肺癌診療の中で収集される多種多様な患者情報に注目する．近年，画像と電子カルテ情報などの非画像情報を組み合わせたアプローチが多数報告されている[4-6]．筆者らは，以前細胞診画像と患者情報を組み合わせた肺癌組織型の自動分類手法を報告した[7]．しかしながら，細胞診画像と患者情報を組み合わせた肺癌の良悪性自動分類についてはこれまでに報告がない．本研究では，細胞診画像と患者情報を併用した肺癌良悪性自動鑑別手法を提案することを目的とした．

*1 藤田医科大学医療科学部〔〒470-1192 愛知県豊明市沓掛町田楽ヶ窪 1-98〕
e-mail: ayumi926@fujita-hu.ac.jp
*2 藤田医科大学大学院保健学研究科
*3 藤田医科大学医学部
*4 岐阜大学工学部

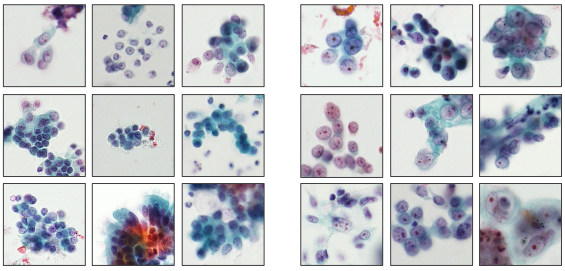
2．使用データ

藤田医科大学病院にて気管支鏡生検が施行された 45 症例(良性 : 18 症例，悪性 : 27 症例)の肺癌患者の細胞診画像および患者情報を収集した．また，すべての症例において，組織学的病理診断および免疫組織化学的診断による最終診断の情報がある．なお，本研究は施

設内倫理審査委員会により承認を得ている(HM18-352).

1）液状細胞診画像

気管支鏡生検にて肺細胞を収集し，SurePath 法を用いて液状細胞診標本を作製した．その後，Papanicolaou 染色法にて染色を行った．40 倍の対物レンズを有する顕微鏡(BX51, オリンパス)に取り付けられたデジタルスチルカメラ(DP70, オリンパス)を使用し，良性 307 枚，悪性 309 枚の細胞像を JPEG 形式にて収集した．初期マトリクスサイズは，2040 x 1536 pixel であり，すべての画像を 296 x 296 pixel に切り出したのち 256 x 256 pixel にリサイズした(図 1).



(a) 良性細胞　　(b) 悪性細胞

図 1 収集した細胞診画像の例.

2）患者情報

次に，電子カルテより細胞診画像に対応する患者情報(年齢，性別，喫煙歴，血液検査結果等)を 27 種類収集した．血液検査結果は，細胞採取日と最も近い日付の検査データを収集した．収集項目の詳細を表 1 に示す.

表 1 患者情報の収集項目.

収集情報	収集項目
患者基本情報	年齢,性別,喫煙状況,ブリンクマン係数,喫煙開始年齢,喫煙年数,喫煙量(本/日),飲酒状況
血液検査情報	WBC, RBC, Hb, Hct, PLT, 分葉核球，リンパ球，単球，プロトロンビン時間，プロトロンビン比，プロトロンビン時間(INR)，APTT，フィブリノーゲン，総蛋白，Alb，CRP，LD(LDH)，BUN，UA，CRE，eGFRcreat

3．方法

はじめに，畳み込みニューラルネットワーク(Convolutional neural network: CNN)を用いて，細胞診画像から良悪性に関する画像特徴量を抽出する．その後，得られた画像特徴量とその画像に対応する患者情報を統合して識別器に入力し，肺細胞の良悪性鑑別を行う.

1）画像特徴量抽出

本研究では，CNN を特徴抽出器として利用する．CNN のモデルとして，ImageNet のデータセットで事前学習した VGG-16 を使用した(図 2)[8-9]．全結合層から 4096 次元のベクトルに圧縮された特徴量を抽出した．学習のフレームワークには Keras + Tensorflow を使用し，演算には，NVIDIA 社製の GEFORCE GTX 1080(メモリ: 8GB)を使用した.

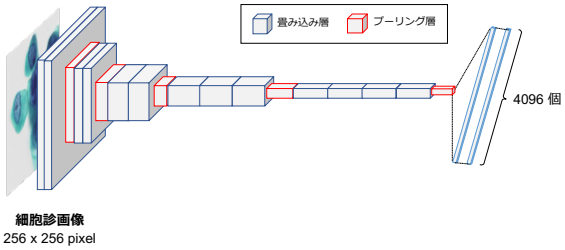


図 2 細胞診画像からの画像特徴量抽出器として使用したネットワークの概要.

2）患者情報の前処理

2. 2)にて収集した患者情報に対し，前処理を行った．欠落データは，全体の平均値で補間した．連続変数に対して，平均値が 0，分散が 1 となるように正規化処理を行い，カテゴリ変数に対してダミー変数化を行った.

3）特徴選択

3. 1)および 2)にて得られた特徴量から，情報獲得量(Information gain)を算出し，分類に有効な特徴量のランキングを作成した．情報獲得量は，決定木アルゴリズム ID3 で使用され

る古典的な特徴量の良さを示す基準である(式(1))[10].

$$Ent(D) = -\sum_{y \in Y} P(y|D) log P(y|D) \qquad (1)$$

本報告では初期検討として，特徴量の総数が 20 個となるように，患者情報・画像特徴量から使用する特徴量の数を決定した．入力特徴量の組み合わせを表 2 に示す.

表 2 検証に使用する入力特徴量の個数の組み合わせパターン.

	患者情報	画像情報
I	20	-
II	-	20
III	10	10
IV	5	15

4）機械学習による識別処理

3. 3)で選択された画像特徴量および，患者情報を機械学習に基づく分類アルゴリズムに入力し，良性/悪性を分類する．分類アルゴリズムとして，単純ベイズ分類器(Naïve Bayes classifier: NB)，ランダムフォレスト(Random forest: RF)，サポートベクタマシン(Support vector machine: SVM)，人工ニューラルネットワーク(Artificial neural network: ANN)を使用し，識別性能を比較・評価した．識別処理には，データマイニングソフトである Orange を使用した.

4. 結果

はじめに，入力特徴パターン I〜IV について，分類アルゴリズムごとの識別精度を比較した．検証で使用した分類アルゴリズムのパラメータを表 3 に示す.

表 2 に示した入力特徴量の組み合わせにおける分類結果を比較すると，最も良い精度を導いた分類アルゴリズムは，入力特徴量パターン I と IV，II，III のとき，それぞれ SVM，NB，ANN であった.

これらの条件における分類結果の詳細を表 4 に示す．ROC 曲線から算出した AUC は，入力特徴量パターン I，II，III，IV のときにそれぞれ 0.795，0.818，0.856，0.920 となった.

表 3 使用した分類アルゴリズムの各種パラメータ.

	パラメータ名
RF	木の数：25, 55
SVM	カーネル：Linear, Polynomial Cost：1.00, Gamma:0.50
ANN	ネットワーク構造(各層のノード数) ：20-128-2, 20-32-8-2, 20-32-128-2 学習回数：20,80,100 活性化関数：tanh, ReLU 最適化アルゴリズム：Adam 学習率：0.01

表 4 提案手法による分類結果の詳細.

入力特徴	識別器	識別率[%]			AUC
		良性	悪性	総合	
I	SVM	63.8	79.3	71.6	0.795
II	NB	69.4	76.1	72.7	0.818
III	ANN	82.7	70.6	76.6	0.856
IV	SVM	85.3	80.9	83.1	0.920

5. 考察

検証結果より，患者情報あるいは画像情報単体と比較して,それらを組み合わせたとき,識別精度が向上することが確認された．特に画像情報のみを使用した入力特徴量 II と，患者情報 5 個+画像情報 15 個を使用した入力特徴量 IV を比較すると,総合識別率は 10.4%向上した.

患者情報のうち有効であった特徴量には，Hct(ヘマトクリット)，フィブリノーゲン，APTT(活性化部分トロンボプラスチン時間)，年齢などが含まれていた．肺癌の発症は年齢とともに増加することや，担癌患者は凝固亢進状態にあるため，これらのマーカーが特徴量選択において選択されたと考えられる[11-12]．これらの結果から，患者の全身状態

や血液凝固の変化に応じた情報が補足されることで，肺癌の良悪性鑑別精度が改善されることが示された．

6. まとめ

本論文では，患者情報を併用した細胞診画像における肺細胞の良悪性自動鑑別手法を提案した．検証結果から，患者情報 5 個+画像情報 15 個を組み合わせて使用し，SVM にて分類を行ったとき最も高い識別精度が得られた．

今後の課題として，他施設の症例を含めたより多くの症例での有効性の評価に関する検討を行う必要がある．

利益相反の有無

なし

文　献

［1］ American Cancer Society, Cancer Facts and Figure, 2019

［2］ Teramoto A, Yamada A, Kiriyama Y, et al.: Automated classification of benign and malignant cells from lung cytological images using deep convolutional neural network. Inform Med Unlocked **16**: 100205, 2019

［3］ Teramoto A, Tsukamoto T, Yamada A, et al.: Deep learning approach to classification of lung cytological images: Two-step training using actual and synthesized images by progressive growing of generative adversarial networks. PLoS ONE **15**: No.3, e0229951, 2020

［4］ Jiachen W, Riqiang G, Yuankai H, et al.: Lung Cancer Detection using Co-learning from Chest CT Images and Clinical Demographics. Proceedings of the SPIE Int Soc Opt Eng., San Diego, California, United States, 2019 Mar, 10949

［5］ Young J, Modat M, Cardoso M.J, et al.: Alzheimer's Disease Neuroimaging Initiative. Accurate multimodal probabilistic prediction of conversion to Alzheimer's disease in patients with mild cognitive impairment. Neuroimage Clin **19**: 735–745, 2013

［6］ Patel M.J, Andreescu C, Price J.C, et al.: Aizenstein, H.J. Machine learning approaches for integrating clinical and imaging features in late-life depression classification and response prediction. Int J Geriatr Psych **30**: 1056–1067, 2015

［7］ 山田あゆみ，寺本篤司，桐山諭和他：液状細胞診画像と患者情報の併用による肺癌組織型自動分類の基礎的検討. Med Imag Tech **37**: 230-234, 2019

［8］ http://www.image-net.org, (2020 年 5 月 20 日閲覧).

［9］ Simonyan K, Zisserman A: Very deep convolutional networks for large-scale image recognition. A conference paper at ICLR, San Diego, pp.1-14, 2015

［10］ Quinlan J.R, “Induction of Decision Trees”, Machine Learning **1**: 81-106, 1986

［11］ Falanga A, Rickles F.R.: Pathophysiology of the Thrombophilic State in the Cancer Patient, Semin Thromb Hemost **25** 173-82, 1999

［12］ Dipasco P.J, Misra S, Koniaris LG, et al.: Pathogenesis of the hypercoagulable state associated with malignancy. J Surg Oncol **104** 316, 2011

Automated Classification of Benign and Malignant Cells Using Lung Cytological Images and Patient Information -Comparison of Performance by Four Types of Classification Algorithms-

Ayumi YAMADA[*1,2], Atsushi TERAMOTO[*1], Yuka KIRIYAMA[*3], Masato HOSHI[*1],
Tetsuya TSUKAMOTO[*3], Kazuyoshi IMAIZUMI[*3], Hiroshi FUJITA[*4], Kuniaki SAITO[*1,2]

*1 School of Medical Sciences, Fujita Health University
*2 Graduate School of Health Sciences, Fujita Health University
*3 School of Medicine, Fujita Health University
*4 Faculty of Engineering, Gifu University

In lung cytology, it is hard task for pathologists and cytotechnologist to distinguish between benign and malignant cells. The clinician generally diagnoses based on variety of patient information as well as changes in cell shape in the actual clinical process. In this study, we aimed to develop an automated classification method of benign and malignant cells using lung liquid-based cytology images and patient information. First, 4096 image features were extracted from LBC images using VGG-16, one of the convolutional neural networks. Subsequently, 27 types of patient information (age, gender, smoking status, laboratory test results etc.) were collected from electrical medical record (EMR). Feature selection was performed on the obtained image features and patient information, and benign/malignant classification was conducted using machine learning techniques. In the experiments, the proposed method was applied to 45 cases (benign; 18, malignant; 27) and evaluated via 3-fold cross-validation. Compared with classification results using patient information alone or image information alone, the classification accuracy was improved by combining them. These results indicate that the proposed method may be useful for benign/malignant cells classification of lung cancer.

Key words: Lung cancer, Liquid-based cytology, Classification, Electronic medical record

乳がん手術支援のための
乳房追跡による MRI 画像重畳の検討

小口　貴也[*1]　藤本　浩司[*2]　吉村　裕一郎[*3]

林　秀樹[*3]　中口　俊哉[*3]

要旨

乳がんの術前計画における腫瘍切除範囲の設定は，MRI 画像上における腫瘍の位置を術者が頭の中で実際の患者体位へと置換して行われる．先行研究では腫瘍部を視覚化するために術前に手術体位で撮影した MRI 画像を患者の体表に投影する手法が提案された．しかし，この手法は投影像と体表の位置合わせを手動で行うため術者の経験に依る部分が大きい. そこで, 術前の MRI 画像から腫瘍の 3 次元モデルを作成し，タブレット端末のカメラ画像上で患者体表に重畳するシステムを提案した．本システムは最初 MRI 画像上とカメラ画像上からそれぞれ 4 点を対応点として手動で選択し，以後は選択点のトラッキングにより重畳のための位置合わせを自動で行う．その際，加速度センサから取得したタブレット端末の姿勢を考慮することでより重畳精度を向上した．重畳精度を評価するためファントムを用いた実験を行った．この実験より，3 次元モデルが十分な精度で重畳されていることが示唆された．

キーワード：手術ナビゲーション，AR 画像重畳，乳房 MRI

1. はじめに

乳がんは，乳房内の組織から発生する悪性腫瘍であり，国内の女性において罹患数が最も多い癌である．罹患数は年々増加しており，年間 6 万人以上が乳がんと診断されている．また，死亡者数も年々増加傾向にあり，年間約 1 万 3000 人を超えている [1]．手術療法は現在でも最も確実な局所療法であり，大きく分けて腫瘍のみを切除する乳房温存手術と乳房全体を切除する乳房切除術がある．乳房温存手術は整容性に優れた手術であるが，腫瘍が残存するリスクがあるため，適切な腫瘍の局在診断が重要となる．診断にはマンモグラフィ検査，エコー検査，MRI 検査が用いられる．マンモグラフィ検査とは，乳房専用のレントゲン検査であり，乳房を透明な薄い板で挟み，上下・左右の 2 方向から撮影する．この方法は乳房全体像を簡便に捉えられる利点がある一方，乳房圧迫による変形がおこり，手術時と病変の位置関係が変化するという欠点がある．エコー検査は非侵襲的であり，腫

*1 千葉大学融合理工学府基幹工学専攻
[〒263-8522 千葉市稲毛区弥生町 1-33]
e-mail: oguchi.takaya@chiba-u.jp
*2 千葉大学医学部附属病院乳腺・甲状腺外科
[〒260-8677 千葉県千葉市中央区亥鼻 1-8-1]
*3 千葉大学フロンティア医工学センター
[〒263-8522 千葉市稲毛区弥生町 1-33]

瘤性病変の特定には有用であるが，非浸潤性乳管癌を代表とする非腫瘤性病変の広がり診断に関してはMRIに劣る．

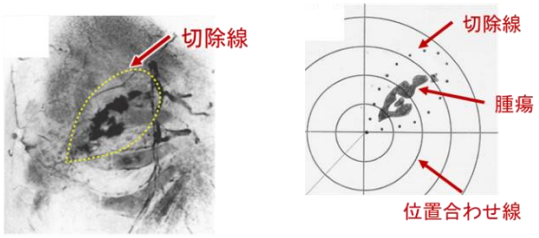
非腫瘤性病変に対する従来の乳がん手術は，マンモグラフィにおける石灰化病変を対象にフックワイヤを用いて行われてきた．術前にマンモグラフィの画像を基にマーキングを行い，手術時にマーキングされた箇所を切除する手法である．しかし，撮影時と手術時で乳房の形状が異なるため腫瘍の位置ずれが発生する可能性がある．位置ずれにより癌細胞が取り残される恐れがあり，その場合は再切除あるいは乳房切除が必要となる． また，石灰化を認めず，そもそもマンモグラフィでは病変が特定できない例も多く存在する．そのため，新たな手術方法として術前に患者の仰臥位MRI撮影を行い，撮影した画像を術前の患者の体表に投影し，切除領域の決定をサポートする手法が提案された[2]．この手法は，MRI 画像を手術時と同じ仰臥位で撮影し投影しているため，乳房が変形せず適切な切除領域を決定することが可能である．しかし，この方法は投影像と体表の位置合わせを手動で行うため術者の経験に依る部分が大きい，投影の準備が煩雑，投影器材が大型でありシステムの導入が困難，といった欠点がある．

本研究では，MRI 画像と体表を半自動で位置合わせが行える小型の乳がん手術支援システムの構築を行う．システムの小型化ならびに普及に適した構成を考慮して，身近なデバイスであるタブレット端末のみを使用する．タブレット端末のカメラ画像にMRI画像を重畳することで，切除領域の決定を支援する．

2. 先行研究

先行研究では，仰臥位MRI画像を体表に投影し，切除領域を決定する方法が提案された．まず，術前に手術時と同じ仰臥位でMRI撮影を行い，それをコロナル平面で重ね合わせ一枚の画像を作成する．次に，作成した画像を紙に印刷（図 1(a)）し，透明シートを重ねて腫瘍と切除線を手動で描画する（図 1(b)）．透明シートには，乳頭から一定距離を示す円と乳頭から上下/左右に線が描画されている．この線は位置合わせ線といい，投影時に体表との位置合わせに使用する．また，切除線は腫瘍の位置から医師の経験により決定される．その後，手術時に，描画された透明シートに真上から光を照射し，その影を患者の体表に投影する．それから，位置合わせ線を参考に医師が経験的に透明シートの位置を調整し，投影した影と体表との位置合わせを行う（図 2）．その後，影となった腫瘍と切除線を基に直接体表に描画し，この範囲にマーキングを行う．これには粘調性の青色色素を使用し，描画した切除線から胸壁の手前まで垂直に注入する．医師はこれに従って腫瘍の除去を行う（図 3）．

この手法はMRI画像を手術時と同じ仰臥位で撮影し投影しているため，乳房の変形が無く，正確な切除領域の決定を可能とした．マンモグラフィを使用していた従来の手法と比較して，腫瘍の残存が減り，再切除を行う確率が大幅に低くなることも示され手術時間の短縮も確認された．しかし，本手法は投影後の位置合わせが手動であるため準備が煩雑であること，位置の合わせを医師の経験に基づいて行っているため正確ではない可能性があるといった問題点がある．さらに，透明シートに光を照射する装置も大型であり，システムの導入が困難である．また， MRI画像を真上から投影しているため，仰臥位の患者に対して真上から胸骨の手前までの垂直な切除に限定される．そのため，腫瘍以外の部位も切除されてしまう可能性がある．



(a) 印刷した MRI 画像　　**(b)** 透明シート

図 1 透明シートへの描画 [2]



図 2 仰臥位 MRI 画像を術前の患者の体表に投影[2]

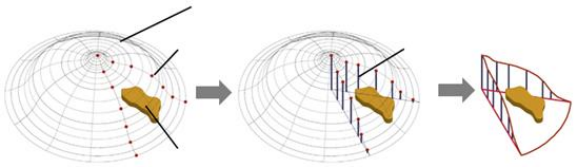


図 3 投影時の切除の流れ [2]

3. 提案システム

3.1 システムの構成

本研究では，タブレット端末のカメラ画像に腫瘍の 3 次元モデルを重畳するシステムを提案した．本システムを用いた腫瘍切除の流れを図 4 に示す．本システムの使用準備として事前に患者の仰臥位 MRI を撮影し，撮影画像から重畳する腫瘍の 3 次元モデルの作成を行う．

重畳する際，医師は 3 次元モデル上で患者の体表との対応がとれる 4 点の座標を取得する．また，対応をとる 4 点を患者の体表にもペンでマーキングを行う．マーキングによりカメラ画像から対応点の座標を取得することが可能となる．なお，乳頭や生検の跡など周辺と明らかにコントラストが異なっている点を使用する場合はマーキングを行う必要はない．マーキング後，タブレット端末でシステムを起動し体表の撮影を行う．撮影の際は，仰臥位の患者に対して真上から水平状態で撮影する．撮影開始後にカメラ画像上で体表の対応点をクリックすることで MRI 画像上の対応点との位置合わせが自動で行われ腫瘍の 3 次元モデルが重畳される．医師はカメラ画像上に重畳された腫瘍から切除領域を決定し，腫瘍領域と切除領域を患者の体表にペンで描画する．その後，システムを終了し，体表に描画した切除領域の切除を行う．3 次元モデルの重畳により，任意の位置から撮影することが可能であるため，腫瘍を最小限の領域で切除することができる．

重畳システムの流れを図 5 に示す．システム使用中はリアルタイムの重畳を可能とするため，カメラ画像上の対応点のトラッキングを行う．トラッキングには Kanade-Lucas-Tomasi Tracker(KLT Tracker) [3]を使用した．これは，オプティカルフローを用いてカメラ画像上の特徴量を検出する方法であり，高速かつ高精度な演算が可能となる．

位置合わせは，カメラキャリブレーションで取得した内部パラメータと歪係数を用いて，外部パラメータを求める．初期位置合わせでは，仰臥位の患者に対して真上から撮影することで正確な外部パラメータの取得が可能となった．そのため，初期位置のみタブレット端末の姿勢推定を行い，水平状態を促すシステムを提案した．タブレット端末に内蔵されている加速度センサを用いて傾きの検出を行い（図 6），画面上に傾いている方向に矢印が表示される機能を導入した．許容されるタブレット端末の傾きをピッチ角-3～+3 度，ロール角-3～+3 度と設定した．

使用者がタブレット端末映像を見ながら切除領域を体表に描画する際，対応点が隠れる可能性がある．そのため，切除領域を表示させるときはカメラ画像上の対応点座標を固定するフリーズを導入

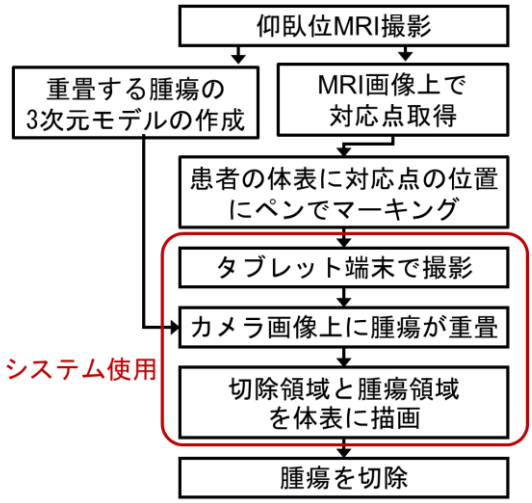


図 4 提案システムを使用した切除の流れ

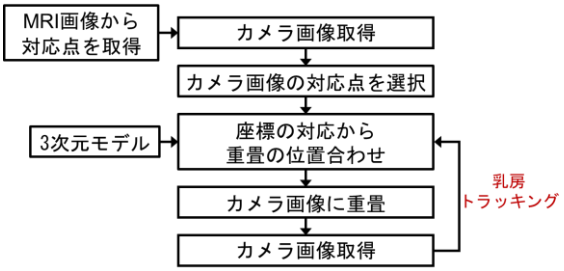


図 5 システムの流れ

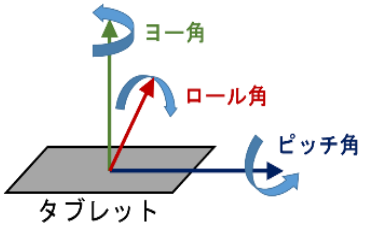


図 6 カメラの姿勢検出

した．しかし，フリーズする際はタブレット端末の位置を固定する必要がある．

3.2 乳房トラッキング

カメラ画像上の対応点は乳頭や生検の跡など目視での確認が可能な点であり，使用者がシステム使用時に選択することで取得する．そして，この対応点をトラッキングすることで，リアルタイムの重畳を可能とした．トラッキングには KLT Tracker を使用した．

本研究では，トラッキングの安定化のため，カメラ画像上で選択された対応点の一点に対して近傍の 3 つの特徴点をトラッキングした．トラッキングしている特徴点が 1 つでもロストした場合は，前フレームの対応点から近傍の特徴点 3 点を再度取得する．また，トラッキングに適した特徴点の重心と選択した対応点の位置がずれる場合があるため，初期フレームの特徴点と対応点の位置関係から現フレームにおける対応点の位置推定を行った．このアルゴリズムを以下に示す（図 7）．

1. 初期フレームf_0においてコーナー値が最も高い特徴点p_{00}から 2 番目に高い特徴点p_{10}へのベクトルv_0と，3 点の特徴点の重心p_{w0}と対応点p_{t0}間のベクトルv_{t0}を式(1)から取得する．

$$v_0 = p_{10} - p_{00} \ , \ v_{t0} = p_{t0} - p_{w0} \qquad (1)$$

2. 現フレームf_tにおいて，v_0と同じ特徴点間$(p_{0t} - p_{1t})$のベクトルv_tを取得する．
3. 図 8 に示すように，v_0, v_tのなす角$\Delta\theta_t$を算出する．これは特徴点p_{0t}を原点として式(2)からθ_0, θ_tを求め，式(3)に示すように差分を求めることで得られる．また，式(4)を用いてベクトル長の比λ_tも計算する．

$$\theta_0 = \tan^{-1}\left(\frac{x_0}{y_0}\right), \theta_t = \tan^{-1}\left(\frac{x_t}{y_t}\right) \qquad (2)$$

$$\Delta\theta_t = \theta_t - \theta_0 \qquad (3)$$

$$\lambda_t = \left|\frac{v_t}{v_0}\right| \qquad (4)$$

4. 図 9 に示すように，現フレームの特徴点の重心位置p_{wt}とv_{t0}から$\Delta\theta_t, \lambda_t$を用いて式(5)から対応点を推定する．

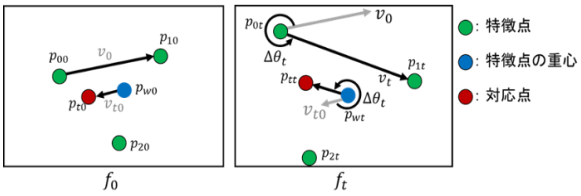


図 7 対応点と特徴点の位置関係

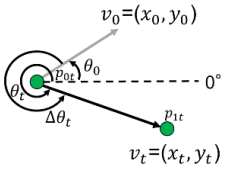


図 8 $\Delta\theta_t$の算出

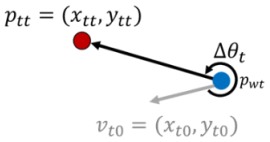


図 9 p_{tt}の算出

$$\begin{bmatrix} x_{tt} \\ y_{tt} \end{bmatrix} = \begin{bmatrix} \cos(\Delta\theta_t) & -\sin(\Delta\theta_t) \\ \sin(\Delta\theta_t) & \cos(\Delta\theta_t) \end{bmatrix} \begin{bmatrix} x_{t0} \\ y_{to} \end{bmatrix} \lambda_t + p_{wt} \quad (5)$$

3.3 3 次元モデルの重畳

重畳準備として，カメラキャリブレーションを行い内部パラメータと歪み係数の導出を行った．キャリブレーションは Zhang の手法を用いた [4].

本システムでは判別式を用いて 4 点の対応点から算出した外部パラメータの精度を評価した．判別式導出のために，初期位置合わせ時に加速度センサを用いてタブレット端末の姿勢を水平状態に制限した．水平状態にすることで取得した外部パラメータの回転ベクトルから算出したオイラー角の一つが一定の値になり，これを用いて判別を行った．

外部パラメータの精度の判別方法の手順を以下に示す．

1. 取得した外部パラメータの回転ベクトルからロドリゲス変換により回転行列を得る．
2. 式(6)を用いて，回転行列R_0をオイラー角$(\theta_x, \theta_y, \theta_z)$に変換する．

$$\begin{aligned} \theta_x &= \tan^{-1}\left(\frac{R_{32}}{R_{33}}\right) \\ \theta_y &= \sin^{-1}(-R_{31}) \\ \theta_z &= \tan^{-1}\left(\frac{R_{21}}{R_{11}}\right) \end{aligned} \quad (6)$$

3. 算出したオイラー角θ_yが-80 ∓ 10度となっていれば外部パラメータが十分な精度であると判断する．

十分な精度と判断されれば次のカメラフレームの初期近似値として利用され，最適化される．精度が不足している場合は，次のフレームで初期近似値を用いずに再度精度の評価を行う．これにより，タブレット端末のあらゆる姿勢で正確な外部パラメータが取得可能となるため，タブレット端末の姿勢制限をなくすことができる．

最後に，得られた正確な外部パラメータを使用し，3 次元モデルの重畳を行った．以上のように，初期位置合わせのみタブレット端末の姿勢を制限することで自動位置合わせを実現した．

4. 重畳精度評価実験

4.1 実験準備

4.1.1 乳房ファントムの作成

重畳準備として乳房ファントムを作成した．まず 3D プリンタで用いてファントムの型を印刷した．日本人女性の乳房の直径は平均 100～120mm [5]であるため，型も同様のサイズ（直径 110mm）とした．次に，ポリウレタン製のファントムを型に流し込み固めた．この際疑似腫瘍として，直方体のゴム材(47×7×22 mm)を埋めた（図 10）．またトラッキングにおいて乳頭と肌のコントラストが特徴点となるため，乳頭の画像をファントムに貼付することでコントラストにリアリティを出した．最後に，ファントムを発泡スチロールの土台に接着し，生検の跡，剣状突起，胸骨と仮定した箇所に少量のファントムを貼付した（図 11）．これにより，MRI 画像上で対応点の取得が可能となる．また，同様の部分にペンで黒く印をつけることで，カメラ画像上でも対応点の取得を可能とした．



図 10 疑似腫瘍

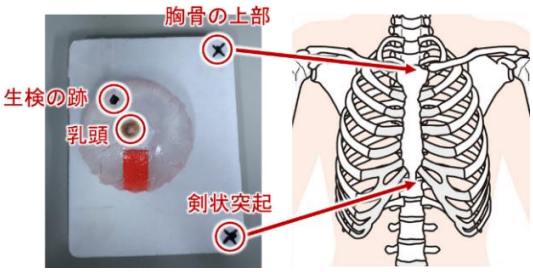


図 11 ファントムと対応点

4.1.2 MRI 画像の撮影

作成した乳房ファントムの MRI 撮影を行った．撮影条件は，コロナル平面，スライス厚 2.5mm，T1 強調とした．T1 強調は，造影剤のコントラストが明瞭となるため [6]，造影剤を使用する乳がんの検査で使用されている．撮影後，MRI 画像から対応点座標の取得を行った（図 12）．

4.1.3 3 次元モデルの作成

3D Slicer ソフトウェア[7]を用いて MRI 画像からファントム形状の抽出と 3 次元再構成（図 13(a)）を行った後，腫瘍部のセグメンテーションを行い，重畳モデル（図 13(b)）を構築した．

4.2 実験概要

3 次元モデルの重畳精度を評価するための 2 つの実験を行った．重畳の精度は，カメラ画像上の腫瘍領域を正解領域とし，重畳した腫瘍領域を推定領域として IoU(Intersection over Union)値の算出を行った．IoU 値は正解領域と推定領域において，領域の共通部分と領域の和集合の比率で表される．腫瘍領域は，腫瘍の色領域から抽出した．また，本実験で使用したタブレット端末の仕様を表 1 に示す．

4.2.1 移動実験

本実験では水平にしたタブレット端末をファントムの真上から撮影しながら移動させた．移動前，移動後，移動後のずれた対応点の修正後の 3 回で精度を測定した．タブレット端末はファントムに対して垂直方向と水平方向に移動させた．垂直方

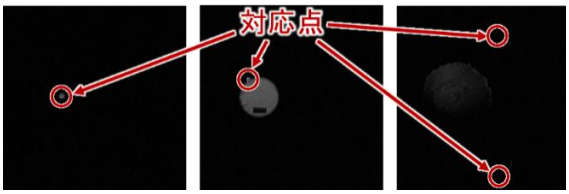
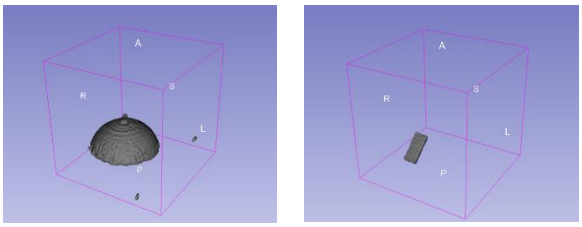


図 12 撮影した MRI 画像

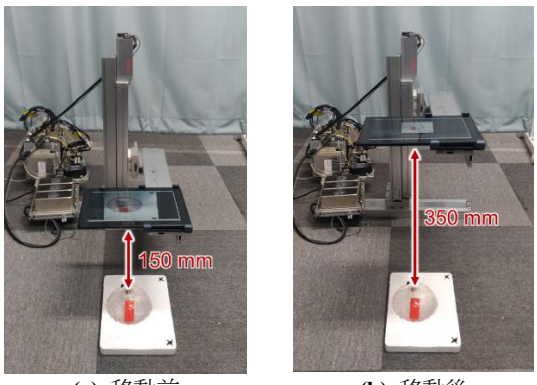


(a) 3 次元再構成　**(b)** 腫瘍部の 3 次元モデル

図 13 乳房ファントムの 3 次元モデル

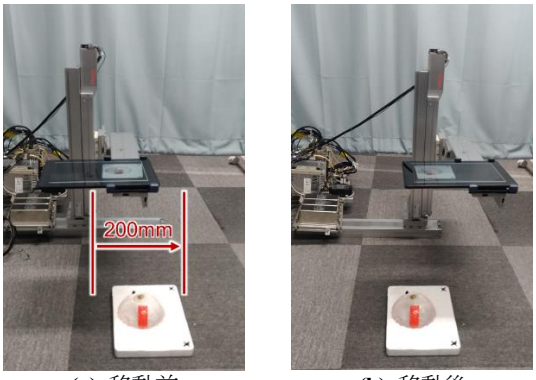
表 1 タブレット端末の仕様

CPU	Intel Core i5-4300 2.5GHz
メモリ	4.00 GB
カメラ解像度	640×480 pixel



(a) 移動前　**(b)** 移動後

図 14 垂直移動の様子

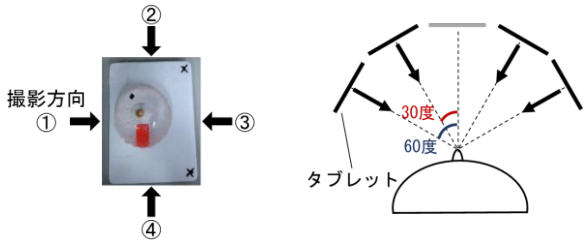


(a) 移動前　　**(b)** 移動後

図 14 水平移動の様子



図 15 傾斜実験の様子



(a) 撮影方向　　**(b)** 撮影角度

図 16 傾斜実験の撮影方向と角度

向では，乳頭から高さが 150mm～350mm にかけて（図 14），水平方向では高さ 300mm において 200mm 横（図 15）に移動させた．移動速度は電動スライダを用いて 100mm/s に制御した．また，本実験は MRI 画像の重畳と 3 次元モデルの重畳を対象とした．

4.2.2 傾斜実験

本実験ではタブレット端末を傾けた際の精度を評価した．3 次元モデルの重畳を対象とし，図 15 に示すようにタブレット端末でファントムを 4 方向から撮影した．さらに，各方向において水平状態から 30 度と 60 度それぞれ傾けた際の精度を算出した．また，本実験は 3 次元モデルの重畳を対象とした．実験 2 の様子を図 16 に示す．

4.3 結果と考察

4.3.1 移動実験

重畳の様子を図 17 に示す．また，本実験で算出した IoU 値を表 2 に示す．

表 2 より，垂直に移動した際は十分な精度が得られたが，水平移動では精度が低下した．原因として，ファントム表面の反射や屈折により，カメラ画像から正確な腫瘍領域の抽出に失敗したことが考えられる．また，移動後の対応点の修正をすることでさらに精度が上がることが示された．

表 2 水平時の重畳精度

		移動前	移動後	修正後
移動方向	垂直	0.82	0.81	0.9
	水平	0.72	0.65	0.73

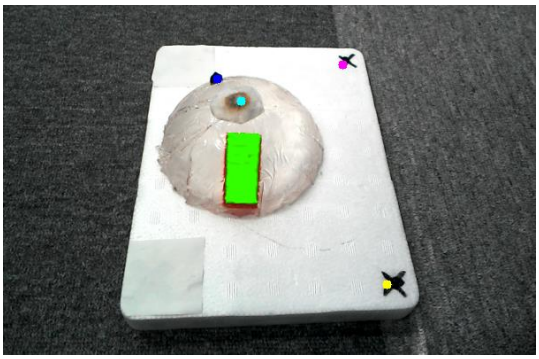


図 17 3 次元モデル重畳の様子

表 3 傾斜時の重畳精度

		撮影方向			
		①	②	③	④
撮影角	30度	0.66	0.75	0.67	0.82
	60度	0.68	—	0.66	0.79

4.3.2 傾斜実験

本実験で算出した IoU 値を表 3 に示す.

表 3 より撮影方向④において十分な精度が得られた．しかし，他の撮影方向からでは精度が低い．これはファントム表面の散乱やカメラ画像から正確な腫瘍領域が得られていないことが原因であると考えられる．また，撮影方向②において 60 度傾けた際は，乳頭により腫瘍部が隠れためカメラ画像上からの抽出が行えなかったことが原因である．

5. 結論

本研究では，非浸潤性乳管がんの手術において，術前に撮影した MRI 画像をタブレット端末の画面に重畳することで腫瘍位置を確認するシステムを構築した．本システムでは乳房トラッキングを用いて位置合わせを行い，腫瘍の 3 次元モデルの重畳を行った．位置合わせでは 4 つの対応点から外部パラメータを導出した.

また，タブレット端末を水平方向，垂直方向に移動させた際と，傾けた際で重畳精度を算出した．結果として，十分な精度で重畳されている可能性があることが示唆された.

今後は，反射や屈折などがないファントムを作り，重畳精度の算出を行う．また，MRI 画像の腫瘍部のみではなく撮影した MRI 画像をそのまま重畳する方法を検討する.

利益相反の有無

なし

文　献

［1］ "乳がん：ESMO/ACF 患者手引きシリーズ ESMO 診療ガイドライン," [Online]. Available: http://www.jsco.or.jp/guide/user_data/upload/File/breast.pdf. [Accessed 8 1 2020].

［2］ J. Y. S. Masahiro Sakakibara, "MRI-Guided Quadrantectomy in Patients with Ductal Carcinoma in Situ Detected Preoperatively by Mammoguraphic Calcifications," Elsevier Inc, 2014.

［3］ C. T. a. T. Kanade, Detection and Tracking of Point Features, Carnegie Mellon University: Technical Report CMU-CS-91-105, 1991.

［4］ Z. Zhang, "A Flexible New Technique for Camera Calibration," IEEE Transaction on Pattern Analysis and Machine Intelligence, 2000.

［5］ X. X. ZHU, Experimental Study on the Design of Brassieres to Prevent Breast Vibration, Department of Clothing Environment Graduate School of Life Environment Reserch Bunka Gakuen University Tokyo,Japan, 2017.

［6］ 白岩美咲, "乳腺 MRI を用いた画像診断の流れ," 異様画像情報学会雑誌, 2012.

［7］ "3Dslicer," 株式会社計算力学研究センタ, [Online].Available:https://www.slicer.org/wiki/Main_Page. [Accessed 15 1 2020].

マルチタッチ LCD により直感的操作が可能な医用画像ディスプレイシステムの開発

増谷 佳孝[*1]　出羽 貴之[*2]　藤川 拓馬[*2]　見澤 勇太[*2]

要旨

我々はマルチタッチ液晶ディスプレイを用いた画像表示システムの開発を行ってきた．主要な機能としてボリュームデータの任意断面像および投影像の表示があり，さらに 1〜10 点のタッチ数に応じた直感的な操作による断面の回転や拡大・縮小，投影の深さ変更などが可能で，さらに対話的なインタフェースにより精密な設定も可能となっている．本報告では主要機能の紹介に加え，タッチ数の認識に関する実験，および任意断面検索の効率を評価する実験について述べる．

キーワード：マルチタッチ LCD，医用画像ディスプレイ，ボリュームデータ

1. はじめに

臨床現場における読影作業あるいは手術計画などのためのボリュームデータの観察[1]において，３次元表示が一般的になって久しいが，２次元すなわち断面表示におけるウィンドウ調整による僅かなコントラストの差異の知覚が正常・異常構造の境界把握に重要な役割を果たすことがある．よって，直交３断面に加えて任意断面の表示は重要な役割を果たす．一般にこれらの画像表示は専用の医療用ワークステーションにて実装されている各種機能により実現されるが，マウス操作やそれに準ずるタッチ操作ではターゲットとなる任意断面を自由に設定して観察することはいまだ困難である．我々は，マルチタッチ液晶ディスプレイ（以下 LCD）を用いてボリュームデータの任意断面像を自由に設定し，同時に投影像の表示を行う医用画像表示システムの開発を行ってきた．本報告では同システムの記述に加え，タッチ数の認識に関する実験，および任意断面検索の効率性を評価する実験について述べる．

*1 広島市立大学 大学院情報科学研究科
〔〒731-3194 広島市安佐南区大塚東 3-4-1〕
e-mail: masutani@hiroshima-cu.ac.jp
*2 広島市立大学 情報科学部

2. システムの記述

本システムでは LCD にタッチした指の本数に応じた動作モード（表１）により観察条件を変化させる．スマートフォンなどにおける一般的なタッチ操作，および先行研究[2]と同様の操作（タッチ数が 1 または 2）に加え，タッチ位置に応じた任意の回転軸の設定による画像の回転（図 1），タッチした位置により範囲が決定される部分投影表示などが可能である．前者は血管や気管支，脊椎の観察に，後者は腫瘍などの観察に有効である．

タッチ数の認識については，意図した数のタッチまでに過渡的に少ないタッチ数の状態が存在するため，タッチ位置の一定以上の変化があるまではタッチ数を確定しない．その閾値を決定する実験，および本システムを使用することによる任意断面表示の効率性評価に関する実験結果を次節に示す．

表 1 指の本数に対応する動作（モード）

タッチ数	モード
1	断面内の平行移動
2	断面内の拡大･縮小と回転
3	奥行き方向への断面移動
4	断面の回転（図 1）
5	明るさ・コントラスト調整
6	部分投影表示
10	リセット

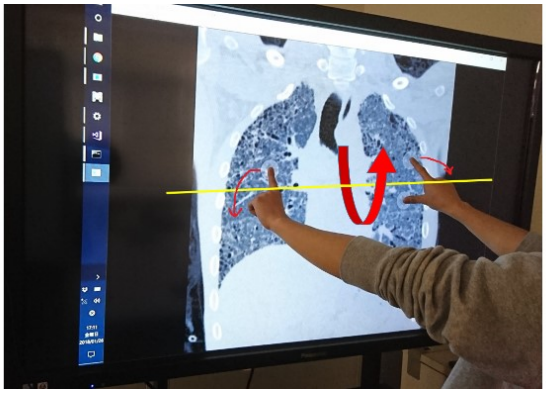


図 1　4 本の指による断面の回転操作
両手の各タッチ位置の中点を結ぶ直線を軸に回転

3. 実験

（実験 1）3 人の被験者で，各モードに対して，意図したタッチ数になるまでの途中のタッチ位置の移動量を 10 回ずつ計測した（図 2）．移動量の平均は 8.33 画素，標準偏差は 26.03 画素であった．したがって，モードの確定の閾値をこの結果に基づき，10 画素に調整した．
（実験 2）従来の操作（マウスによる回転・平行移動）と比較するため，5 名の被験者がボリュームデータ内の 3 つの球を同一断面上に表示するのに要した時間を計測したところ，本研究のインタフェースが，より効率的であることが示唆された（図 3）．

4. まとめ

本稿では，我々が開発中のマルチタッチ LCD による画像表示システムについて述べた．将来的には，対話的セグメンテーション機能[3]を統合する予定である．

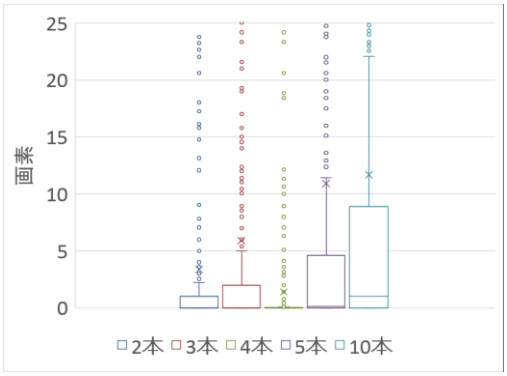


図 2　各モードにおいて最後の指が触れるまでに他の指が移動した量

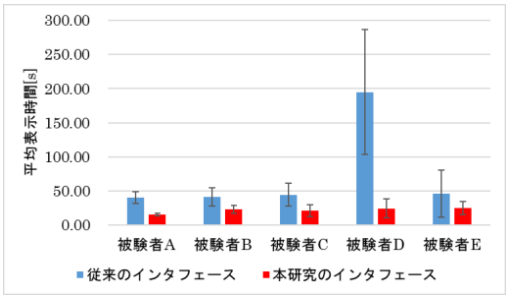


図 3　ボリュームデータ内の 3 つの球を表示するのに要した時間の計測結果（5 回の平均）

謝辞

本研究の成果の一部は，JST CREST (JPMJCR15D1)の助成によるものです．

利益相反の有無：なし

文　献

［1］ B. Preim, and C. Botha, Visual Computing in Medicine. Theory Algorithms and Applications, Morgan Kauffman (Elsevier), 2014

［2］ S. Bay, P. Brauner, T. Gossler, M. Ziefle, Intuitive Gestures on Multi-touch Displays for Reading Radiological Images, proc. Int. Conf. HIMI2013, p22-31, 2013

［3］ 松本尚，富岡義樹，青山正人，増谷佳孝，あいまいな境界情報の入力に基づく対話的臓器領域抽出法，医用画像連合フォーラム予稿集（信学技報 114 巻 482 号）pp.233-236, 2015

Development of Intuitive Medical Image Display System with Multi-Touch LCD

Yoshitaka Masutani[*1], Takayuki IZUWA[*2], Takuma FUJIKAWA[*2], Yuta MISAWA[*2]

*1 Hiroshima City University Graduate School of Information Sciences

*2 Hiroshima City University Faculty of Information Sciences

We have been developing a medical image display system using a multi-touch liquid crystal display. The main function is to display an arbitrary cross-sectional image and projection image of volume data, and it is possible to rotate and enlarge/reduce the cross-section and change the projection depth by intuitive operation according to the number of touches from 1 to 10 points. In addition, the interactive interface also enables precise settings. In this report, in addition to the introduction of the main functions, we describe an experiment on recognition of the number of touches and actions, and an experiment to evaluate the speed of arbitrary section retrieval. As a result of experiments, it was shown that the optimum parameters for the number of touches and the recognition of actions were determined, and arbitrary sections could be displayed quickly, suggesting the effectiveness of this system. In the future, we aim to integrate the annotation function developed as a separate system into a powerful support system for medical image processing, especially segmentation by machine learning.

Key words: Multi-touch LCD, Medical Image Display, Volume Data

著者紹介

増谷　佳孝（ますたに　よしたか）
1997 年東大・大学院・工学系・精密機械工学専攻博士課程了．ハンブルク大医用数学・情報処理研究所，シカゴ大病院放射線科，東大病院放射線科などを経て 2018 年現在，広島市立大学大学院医用情報科学専攻・教授．2003 年日本医用画像工学会論文賞および奨励賞受賞．博士（工学）・博士（医学）．拡散 MRI を中心とした多次元の医用画像の処理・解析の研究に従事．医用画像工学会，電子情報通信学会，磁気共鳴医学会，IEEE など，各会員．

出羽 貴之（いずわ　たかゆき）
2020 年広島市立大学情報科学部医用情報科学科卒

藤川 拓馬（ふじかわ　たくま）
2019 年広島市立大学情報科学部医用情報科学科卒

見澤 勇太（みさわ　ゆうた）
2018 年広島市立大学情報科学部医用情報科学科卒

Image analysis for automatic detection of social interaction in drosophila: biased courtship behavior

渡邊　裕貴[*1]　成 耆鉉[*2]　松村 泰志[*3]　姜 時友[*1]

要旨

求愛行動は種の保存にとって重要なライフイベントである．遺伝子研究に置いて頻繁に用いられるショウジョウバエに置いても例外ではなく，オスからメスへの求愛行動が頻繁に調べられており，様々な知見が得られている．しかし，その逆であるメスからオスへの求愛行動は，その頻度や性質においてオスからのものに比べて弱く，通常の条件下においては観測そのものが困難であり，依然よくわかっていない．本研究では，ショウジョウバエにおけるメスの求愛行動を捉えるために，メスとオスを含んだフィールドを用意した．各ウェルには以下のような 3 つの個体を含めた．一つ目は，主たる行動解析の対象である標的個体として，自由行動下のメスを羽化目前の蛹の状態で中央に配置した．その他に，被接近対象として，不動化処理を施したオス・メスをそれぞれ中央個体の左右両側に配置した．このようなウェルをチェンバー内に複数配置し，動画を撮影した．この動画に対して，個体抽出や二値化処理，モルフォロジー処理などを行うことによって，標的個体の求愛行動を定量化した．その結果，従来オスからメスへの求愛行動に埋没して明らかになっていなかったメスの求愛行動を明らかにした．本研究は，ショウジョウバエ社会的相互作用と関連した遺伝子研究への貢献が期待される．

キーワード：High-throughput screening, Biological data, Social interaction, Behavioral analysis, Image processing.

1. はじめに

1.1 背景

本研究の解析対象となっているショウジョウバエは，ショウジョウバエ科に属するハエの総称であり，卵期，幼虫期，蛹期，成虫期と成長する完全変態昆虫である．ショウジョウバエは個体サイズが小さく，1 世代にかかる期間が 2 週間程度と世代交代が早いため，飼育にかかる時間的・空間的コストが低いことだけでなく，染色体数が性染色体 1 対，常染色体 3 対の計 4 対と少ないこと，実験生物として長い歴史を持ち，遺伝子的知見や実験技術が蓄積されていることなどから，遺伝子研究におけるモデル生物として広く用いられている．

また，ショウジョウバエに限らず，求愛行動は，種を存続させ，交尾させるための重要な社会行動である．ショウジョウバエを用いた遺伝子研究の分野では，活発なオスからメスへの求愛行動が頻繁に研究されているが，反対のメスからオスへの求愛行動は相対的に不明瞭で，通常の条件下では観察すらできないことが多く，依然明らかになっていない．

1.2 目的

本研究では，ショウジョウバエのメスからオスへの求愛行動に関する新たな知見を得ることを目的として，ショウジョウバエの多個体同時撮像動画から，ショウジョウバエ個体間の社会的相互作用の一つである求愛行動の自動検

*1 山形大学大学院・理工学研究科・応用生命システム工学専攻
〔〒992-8510 米沢市城南 4-3-16〕
e-mail: siu@yz.yamagata-u.ac.jp
*2 理化学研究所筑波地区・CPR・眞貝細胞記憶研究室
*3 パクテラ・デジタル・イノベーション事業開発部

出システム構築を目指す．以下に記すような，オスとメスを含んだフィールドを設計し，撮影された動画に画像処理手法を適用することによって，今まで明らかになってメスの求愛行動の可視化を試みた．

2. 方法

2.1 動画について

各ウェルには以下のような 3 匹の個体を配置した．行動解析の対象となる自由行動可能なメスのショウジョウバエを，羽化前の蛹の状態で中央に配置した．

その他に，不動化処理を行った成虫のオス・メス一匹ずつを左右に配置した．

これら二匹が，標的個体である中央メスの被接近対象であり，標的個体の羽化後を開始時点とし，その接近行動を定量化した．このようなウェルを 6 個チェンバー内に配置し，動画として撮影した．

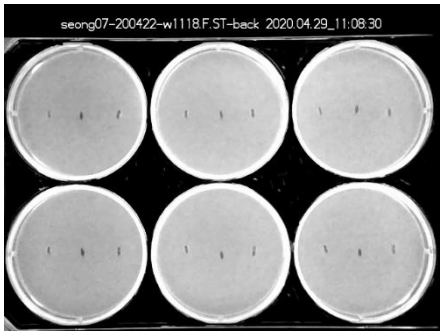


図 1 全体動画

ショウジョウバエの飼育および動画の撮影はつくば理化学研究所にて行われた．

2.2 解析アルゴリズム

解析アルゴリズムを図 2 に示す．まず，自由に動くメスのショウジョウバエの動きを捉えるために，この個体の領域抽出を行った．抽出した二値画像に，閾値処理とモルフォロジー変換を行い，ノイズを除去した．その後，自由行動個体の移動軌跡を得るために重心を計算し，この個体が不動化個体に対して行った接近行動を様々な指標を用いて定量化した．

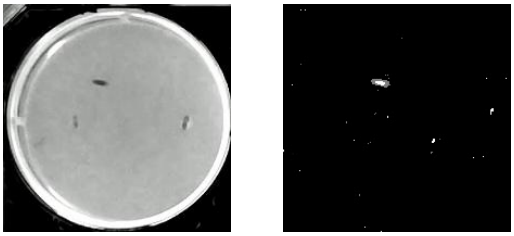
以下に，各ステップの詳細を記す．



図 2 解析フロー

2.3 個体抽出

元動画(図 3 a)に対し，混合正規分布を基にした前景・背景の領域分割アルゴリズムを適用し，自由行動個体（メス）のショウジョウバエの抽出を行った．この処理によって，背景オブジェクトを作成し，それ以降の画像を読み込む度に前景領域のマスク画像を取得することができる．



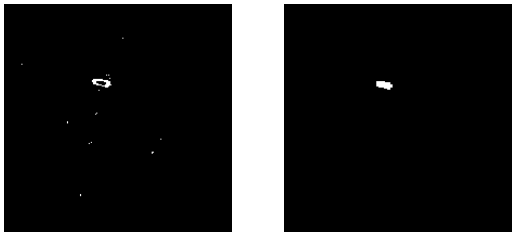
(a)元動画　　(b)抽出後

図 3 個体抽出

2.4 ノイズ除去

個体領域抽出後も残存するノイズを除去するために，閾値処理(図 4 a)とモルフォロジー変換(図 4 b)を行った．閾値は，抽出したショウジョウバエが灰色の領域となっているため下限と上限を設定し，白いノイズを除去するとともに，二値化を行った．次に，モルフォロジー変換を行った．これは，注目画素内に 1 画素でも白の画素が含まれれば領域内を全て白に置き換える膨張処理と注

目画素内に含まれる全ての画素が白であれば白のまま，そうでなければ領域内を黒に置き換える収縮処理がある．同じ回数分の膨張収縮を繰り返す closing 処理によって，個体領域抽出や閾値処理の際に生じてしまった個体の欠損箇所を埋め，かつ同じ回数分の収縮膨張を繰り返す opening 処理を行うことによって閾値処理だけでは除去できなかった細かいノイズを除去した．



(a)閾値処理　　(b)モルフォロジー変換

図 4 閾値処理およびモルフォロジー変換

2.5 重心の計算と接近の定量化

ノイズ除去後の画像に対して，オブジェクトをラベリングすることで抽出個体とそれ以外のオブジェクトとして区別し，個体以外のオブジェクトを除外するとともに，自由行動個体の重心を算出した．そして，2 匹の不動化個体との距離をそれぞれ計算し，自由行動個体と不動化個体との距離や，不動化個体の周囲一定半径内に進入した回数などをカウントすることによって，自由行動個体の接近行動を定量化した．

3. 結果

今回は標的個体の滞在分布，左右の個体との距離のバイアス，左右の個体に一定距離以内への接近回数，の 3 種類の指標を用いて定量化を行った．

図 5 に標的個体の滞在分布を示す．標的個体はオス個体と自身の蛹付近での滞在が多いことが分かった．

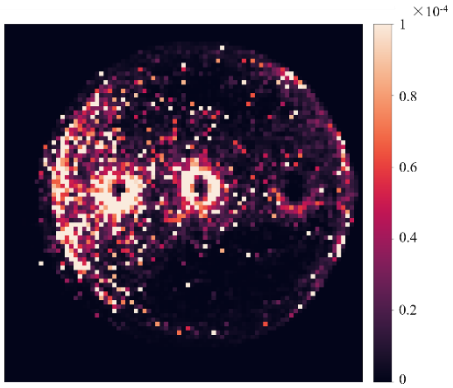


図 5 標的個体の滞在分布

図 6 に個体間距離のバイアスを示す．標的個体は左のオス個体よりも右のメス個体との距離が大きく，オス選考性を持つことが示された．

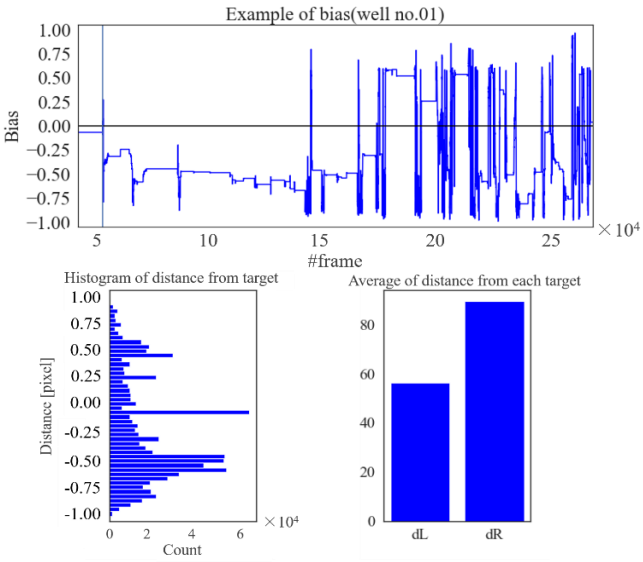


図 6 個体間距離のバイアス

図 7 に一定距離以内への接近回数を示す．左右個体に対し，画像上で 15[px]，20[px]，25[px]を半径とする円内への進入した回数を

カウントとした．接近回数においても，その半径に関わらず，標的個体はオス選好性を示した．

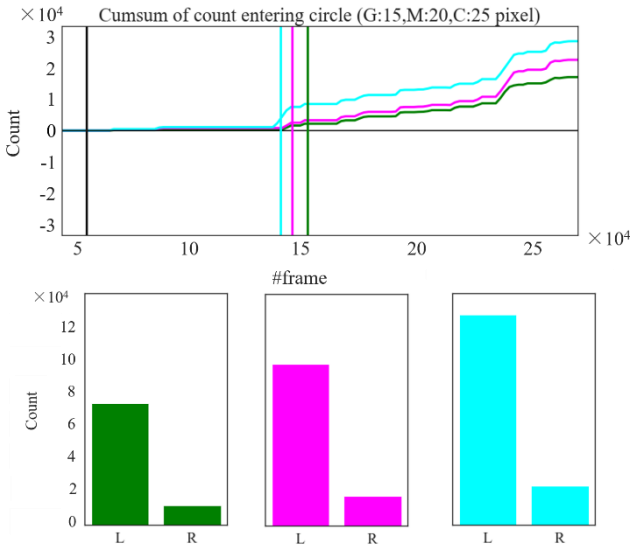


図 7 一定距離内への接近回数

4. まとめ

本研究では，求愛行動のターゲット個体を不動化するという特殊な状況を作り出すことで，従来用いられるような通常条件下では埋没してしまい観測できなかった，メスからオスへの求愛行動の定量化を試みた．その結果，滞在時間や距離バイアス，接近回数の 3 種の指標全てにおいて，オス選好性が見られた．また，予備的な結果として，自由行動下にあるメスは，オスだけでなく，自身の蛹にも誘引され，この付近での滞在確率が高いことがわかった．

本研究によって，メス個体の求愛行動に関する傾向が明らかになった．今後は，メスのこのようなオス選好性が集団においてどの程度見られるのかを定量化するとともに，羽化後のメスが初めてオスへの誘引行動を行うタイミングを検出するアルゴリズムの構築を目指す．

謝辞

本研究は，AMED (Japan Agency for Medical Research and Development, JP18gm1110001)の支援を受けたものです．

利益相反の有無

なし

文　献

［1］ Seong KH, Matsumura T, Shimada-Niwa Y, Niwa R, Kang S. (2020) “The Drosophila Individual Activity Monitoring and Detection System (DIAMonDs)”, *bioRxiv*, doi: https://doi.org/10.1101/2020.06.05.136085

Image analysis for automatic detection of social interaction in drosophila: biased courtship behavior

Hiroki WATANABE[*1], Ki-Hyeon SEONG[*2], Taishi MATSUMURA[*3] Siu KANG[*1]

*1 Yamagata University

*2 RIKEN Tsukuba Institute

*3 Pactera Technology JAPAN

Courtship behavior is an important life-event in order to mating and surviving the species. In drosophila known as fruit fly and typically used in genetic research field, courtship behavior from male to female has been examined, however, the ones from female to male is still unknown because it is weak and un-observable under normal condition. To capture the female courtship behavior, a free-moving field with male and female targets is designed. Each well has three individuals, one is a freely-moving female drosophila before eclosion, and the others are immobilized male and female as targets of courtship behavior of freely-moving female. Behaviors of a freely-moving female drosophila was recorded as movie and captured through image processing as follows: First, foreground, that is, drosophila was extracted from background. Second, thresholding and morphological filters were applied onto extracted the foreground binary images for noise reduction. Third, numbering and elimination were performed and central gravity was calculated in order to obtain a moving trajectory. After that, approaching of the freely-moving female drosophila to the immobilized target animals was quantified. As a result, it was revealed that frequency of approach to female or male showed significant biases. Present research shed light to courtship behavior of female drosophila.

Key words: High-throughput screening, Biological data, Social interaction, Behavioral analysis, Image processing.

著者紹介



渡邊　裕貴（わたなべ　ひろき）
2020 年 山形大学応用生命システム工学科卒．現在，山形大学大学院・理工学研究科・応用生命システム工学専攻・博士前期課程所属



成 耆鉉（そん　きひょん）
1994 年　朝鮮大・理・生物学科卒
2001 年　都立大大学院博士課程・理・生物科学専攻了
現在，理化学研究所開拓研究本部　眞貝細胞記憶研究室（筑波）協力研究員.





松村　泰志（まつむら　たいし）
2014 年 山形大システム創成工学科卒.
2019 年 同大学院博士後期課程修了．神経回路シミュレーション，深層学習，機械学習を用いた画像解析の研究に従事し，工学博士取得

姜 時友（かん　しう）
2001 年　明治大学大学院・理工学研究科・生物物学専攻・修了
2004 年　玉川大学大学院・工学研究科・生産開発工学専攻修了・工学博士
2003-2005 年　日本学術振興会特別研究員
2005-2010 年　理化学研究所・脳科学総合研究センター・研究員
2010 年より山形大学大学院・理工学研究科・助教

大腸内視鏡のための教師なし深度画像推定法における補助タスク検討

伊東 隼人 [*1], 小田 昌宏 [*1], 森 悠一 [*2], 三澤 将史 [*2], 工藤 進英 [*2], 堀田欣一 [*3], 高畠 博嗣 [*4], 森 雅樹 [*5], 名取博 [*6], 森 健策 [*1,7,8]

要旨

医用画像理解においては人体に関する 3 次元情報の抽出が非常に重要である．一方で，一般的な大腸内視鏡では 2 次元画像もしくはそれらの時系列集合しか得られない．大腸内視鏡診断におけるコンピュータ支援診断システムを構築する上で，単眼画像から大腸形状の 3 次元情報を抽出・推定する技術が求められている．先行研究として，多視点幾何学に基づくカメラ運動推定を視差推定の補助タスクとして加えた深度推定方法が提案されている．しかし，あくまで相対的な深度情報しか得られない上にその推定精度が不十分であると報告されている．本研究では深度推定の高精度化を目的に，拡散反射モデルを深度推定に組み込むための補助タスクを提案し，先行研究との比較実験的を行った．実験結果より提案した補助タスクが大腸腸壁のテクスチャに起因する誤推定を抑制することを示し，深度推定の高精度化を達成した．

キーワード: 大腸内視鏡検査, 深層学習，深度推定，補助タスク，拡散反射

1. はじめに

大腸内視鏡診断をサポートするコンピュータ支援診断システムを構築するためには高精度な大腸の 3 次元形状情報の推定・抽出が不可欠である．ポリープサイズの識別 [1]，大腸内視鏡診断中の見落とし領域の検出 [2]，そしてナビゲーションに向けた地図作成と自己位置推定 [3] などでは大腸内視鏡画像からの深度推定が特に重要となる．

一般的な大腸内視鏡では深度情報を捉えることはできないため，何らかの仮定に基づいて 2 次元画像のみから深度情報を推定する必要がある．視差推定と多視点幾何学を利用した教師なし深度推定手法 [4, 5] が提案されているものの，一般画像と大腸内視鏡では問題設定が大きく異なり，推定精度は未だ不十分である [1]．一方で，大腸内視鏡画像に対してはドメイン変換を利用した深度推定手法が提案されている [6]．文献 [6] では CT ボリュームデータから生成した仮想大腸内視鏡画像と深度画像のペアを用いた教師あり学習を利用し，仮想大腸内視鏡画像から深度画像を生成するドメイン変換によって深度推定を達成している．しかし，仮想大腸内視鏡画像はテクスチャや鏡面反射成分が存在しない拡散反射に基づく深度推定に理想的な画像となっており，実際の大腸内視鏡画像に対して最適化された手法ではない．

そこで本稿では拡散反射による輝度分布の推定という補助タスクをドメイン変換に基づく深度推定に導入することで，大腸内視鏡画像における深度推定の高精度化を図る．目的以外の処理（タスク）を補助タスクとして深層学習に組み込むことにより，目的のタスクに有用な情報の抽出がなされ，目的のタスクの精度向上が可能であるとの報告がある [7]．また先行研究により，画像の輝度分布から 3 次元形状を推定する際には拡散反射モデルが有効であることが判明している [8]．さらに，深度推定は大腸内視鏡画像から腸壁の形状を推定しているとも考えることができる．したがって，大腸内視鏡画像における拡散反射による輝度分布推定は，深層学習に基づく深度推定というタスクの精度向上に関して補助タスクとなり得ると仮定できる．補助タスクを導入した提案手法と先行研究の深度推定結果の比較を行い，補助タスクの妥当性を検討する．

[*1] 名古屋大学大学院情報学研究科
[*2] 昭和大学横浜市北部病院 消化器センター
[*3] 静岡がんセンター 内視鏡科
[*4] 札幌南三条病院 呼吸器内科
[*5] 札幌厚生病院 呼吸器内科
[*6] 恵和会西岡病院 呼吸器内科
[*7] 名古屋大学 情報基盤センター
[*8] 国立情報学研究所 医療ビッグデータ研究センター

2. 手法

1) 視差と多視点幾何学に基づく深度推定

時系列の順序構造をもった N_i 枚の連続した画像集合 $\{X_1, X_2, \ldots, X_{N_i}\}$ が N 個与えられているとする．ある集合のうちの 1 枚 X_{t_i} に対し，重複しない画像を $X_{s_i}, 1 \leq s_i \leq N_i, s_i \neq t_i$ とおく．X_{t_i} 上の画素の斉次座標を $\boldsymbol{p}_{t_i}$，X_{t_i} 上のある点 $\boldsymbol{p}_{t_i}$ に対して推定した深度を $\hat{D}_{t_i}(\boldsymbol{p}_{t_i})$，そして X_{t_i} と X_{s_i} の座標系間のカメラ運動を表す回転と平行移動による外部パラメータ行列 $\hat{\boldsymbol{T}}_{t_i \to s_i}$ とする．ここで剪断係数 s，光学中心 $(c_x, c_y)^\top$，そして画素単位の焦点距離 f_x, f_y が既知である内部パラメータ行列

$$\boldsymbol{K} = \begin{pmatrix} f_x & s & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{pmatrix} \tag{1}$$

に対し，X_{s_i} 上の座標 $\boldsymbol{p}_{s_i}$ と X_{t_i} 上の座標 $\boldsymbol{p}_{t_i}$ の関係を

$$\boldsymbol{p}_{s_i} \sim \hat{\boldsymbol{p}}_{s_i} = \boldsymbol{K}\hat{\boldsymbol{T}}_{t_i \to s_i}\hat{D}_{t_i}(\boldsymbol{p}_{t_i})\boldsymbol{K}^{-1}\boldsymbol{p}_{t_i} \tag{2}$$

と表す．座標系の変換 $\hat{X}_{s_i}(\boldsymbol{p}_{t_i}) = \mathcal{X}_{s_i}(\hat{\boldsymbol{p}}_{s_i})$，そして X_{t_i} の画像領域 Ω に対し，教師なし学習として損失

$$\mathcal{L} = \sum_{i=1}^{N} \sum_{s_i} \int_{\Omega} |X_{t_i}(\boldsymbol{p}) - \hat{X}_{s_i}(\boldsymbol{p})| d\boldsymbol{p} \tag{3}$$

を最小にする深度 $\hat{D}_{t_i}(\boldsymbol{p})$ および外部パラメータ行列 $\hat{\boldsymbol{T}}_{t_i \to s_i}$ の推定モデルをそれぞれ構築することで，深度推定を達成する．ここで $X_{t_i}(\boldsymbol{p})$ および $\hat{X}_{s_i}(\boldsymbol{p})$ はそれぞれ X_{t_i} および $\hat{X}_{s_i}$ の座標 $\boldsymbol{p}$ での画素値を表す．

Zhou らは深度推定および外部パラメータ行列の推定に DispNet および Pose Estimation Network をそれぞれ利用し，離散化した式 (3) に正則化を加えてマルチスケール化した上で，これらのネットワークを交互に学習させる教師なし単眼深度推定手法，SfMLearner を提案した [4]．ただし，得られる深度画像はあくまで画像内における被写体への相対的な距離を表すのみであり，物理量は伴わない．Zhou らの手法では，外部パラメータ行列の推定，つまりカメラ運動の推定を視差推定の補助タスクとして組み合わせることによって深度推定を実現した．

2) 拡散反射モデルに基づく深度推定

拡散反射に基づく輝度値の推定を補助タスクとして加えた教師なしドメイン変換による深度推定法を提案する．画素数が $H \times W$ pixels である大腸内視鏡画像 $X_i \in \mathbb{R}^{H \times W \times 3}$ の集合を $\{X_i\}_{i=1}^{n_i}$ とし，そのドメインを $\mathcal{R}_{RGB}$ とおく．CT ボリューム画像から生成した単一点光源と拡散反射に基づく輝度値分布を想定した仮想大腸内視鏡画像と仮想深度画像をまとめた 4 チャンネル画像 $Y_j \in \mathbb{R}^{H \times W \times 4}$ の集合を $\{Y_j\}_{j=1}^{n_j}$ とし，そのドメインを $\mathcal{V}_{RGBD}$ とおく．図 1 は 2 つのドメインにおける画像例を示す．

これら 2 つのドメインに対して，$G : \mathcal{R}_{RGB} \to \mathcal{V}_{RGBD}$ ならびに $F : \mathcal{V}_{RGBD} \to G : \mathcal{R}_{RGB}$ というドメイン変換を考える．さらに，$\{X_i\}_{i=1}^{n_i}$ と $\{F(Y_j)\}_{j=1}^{n_j}$，ならびに $\{Y_j\}_{j=1}^{n_j}$ と $\{G(X_i)\}_{i=1}^{n_i}$ をそれぞれ識別する敵対性識別器を $D_{\mathcal{V}}$ ならびに $D_{\mathcal{R}}$ とする．そして CycleGAN [9] の目的関数を

$$\begin{aligned} &\mathcal{L}(G, F, D_{\mathcal{V}}, D_{\mathcal{R}}) \\ &\quad = \mathcal{L}_{\mathrm{G}}(G, D_{\mathcal{V}}, \mathcal{R}_{RGB}, \mathcal{V}_{RGBD}) \\ &\qquad + \mathcal{L}_{\mathrm{G}}(F, D_{\mathcal{R}}, \mathcal{V}_{RGBD}, \mathcal{R}_{RGB}) \\ &\qquad + \lambda \mathcal{L}_{\mathrm{cyc}}(G, F), \end{aligned} \tag{4}$$

とし，RGB3 チャンネル大腸内視鏡画像と深度画像を 4 チャンネル目に加えた仮想大腸内視鏡画像の間のドメイン変換 F, G を求めるように拡張する．目的関数に対して次式の min max 問題

$$G^*, F^* = \arg\min_{G,F} \max_{D_{\mathcal{V}}, D_{\mathcal{R}}} \mathcal{L}(G, F, D_{\mathcal{V}}, D_{\mathcal{R}}) \tag{5}$$

を解くことにより，拡散反射に基づいた輝度画像と深度画像を 4 チャンネル画像として出力する関数 G を得る．F, G は入出力のチャンネル数以外は文献 [9] と同じ U-Net 構造とした．このとき，拡散反射に基づく輝度画像の生成が深度推定における補助タスクとなる．加えて，CT ボリュームデータから生成した深度画像は cm 単位の物理量を伴うため，推定深度も物理量を伴う．G に大腸内視鏡画像を入力し，出力される 4 チャンネル RGB-D 画像の 4 チャンネル目のみを抽出することにより，推定深度画像を得る．

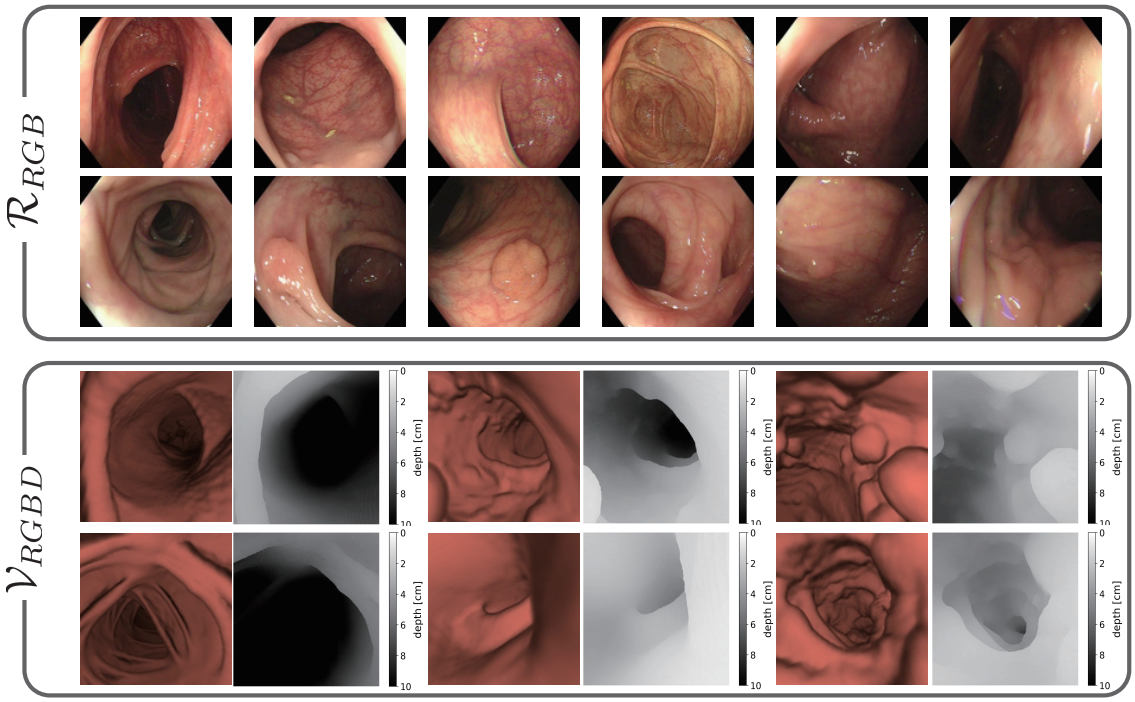


図 1 大腸内視鏡画像のドメイン $\mathcal{R}_{RGB}$ と仮想内視鏡画像ならびに仮想深度画像のドメイン $\mathcal{V}_{RGBD}$. 仮想画像は CT ボリュームデータから生成するため，深度画像は物理量 [cm] を伴う.

3. 実験

提案手法の妥当性を検証するためにドメイン変換 $\mathcal{R}_{RGB} \rightarrow \mathcal{V}_{D}$ という補助タスク無しの深度推定法，ドメイン変換 $\mathcal{R}_{RGB} \rightarrow \mathcal{V}_{RGBD}$ という拡散反射画像の生成を伴う提案手法，そして視差推定と多視点幾何学に基づく深度推定法のひとつである SfMLeaner，これら 3 つの深度推定結果の比較を行った.

補助タスク無しの深度推定法と提案手法の学習では 9,189 枚の大腸内視鏡画像，CT 画像から生成された 8,085 枚の仮想大腸内視鏡画像と仮想深度画像を利用した．大腸内視鏡画像は患者 30 名分の大腸内視鏡検査動画像から 5 fps で抽出した．仮想大腸内視鏡画像および深度画像は 7 名の大腸 CT ボリュームデータから NewVES [10] を用いて生成した．仮想大腸内視鏡画像は拡散反射のみを想定し，仮想深度画像は仮想内視鏡から 0-10cm を 256 階調濃淡値として生成した.

視差推定と多視点幾何学に基づく深度推定法の学習では，患者 30 名分の大腸内視鏡検査動画像から 30 fps で抽出した 33,254 枚の大腸内視鏡画像における連続した 3 画像の 34,349 組を用いて学習した．ここで 5 fps とした場合は，画像間のベースラインが大きすぎるためか深度推定が破綻することが予備実験より明らかになっている．内部パラメータ行列はキャリブレーションより求めた.

学習データに含まれない患者 5 名分の内視鏡動画像より抽出した静止画像に対し，3 つの学習モデルを用いて深度推定を行った．図 2 は，推定した深度値のそれぞれの逆数を表す disparity map [5] を示す．さらに，図 3 は提案手法を用いて大腸内視鏡画像から生成した，拡散反射に基づく輝度画像と深度画像を示す.

4. 考察

図 2 の比較結果は，提案手法による滑らかかつ立体感を表した深度推定の達成を示す．SfM-Learner による深度推定では大まかな相対距離は得られるものの，ところどころ推定値が欠損しており物体の形状も捉えられておらず，空間的な連続性も失われている.

大腸内視鏡画像間の対応点探索は識別的な特徴点が得られないことから難易度が高く，SfM-Learner でも外部パラメータ行列の推定精度に

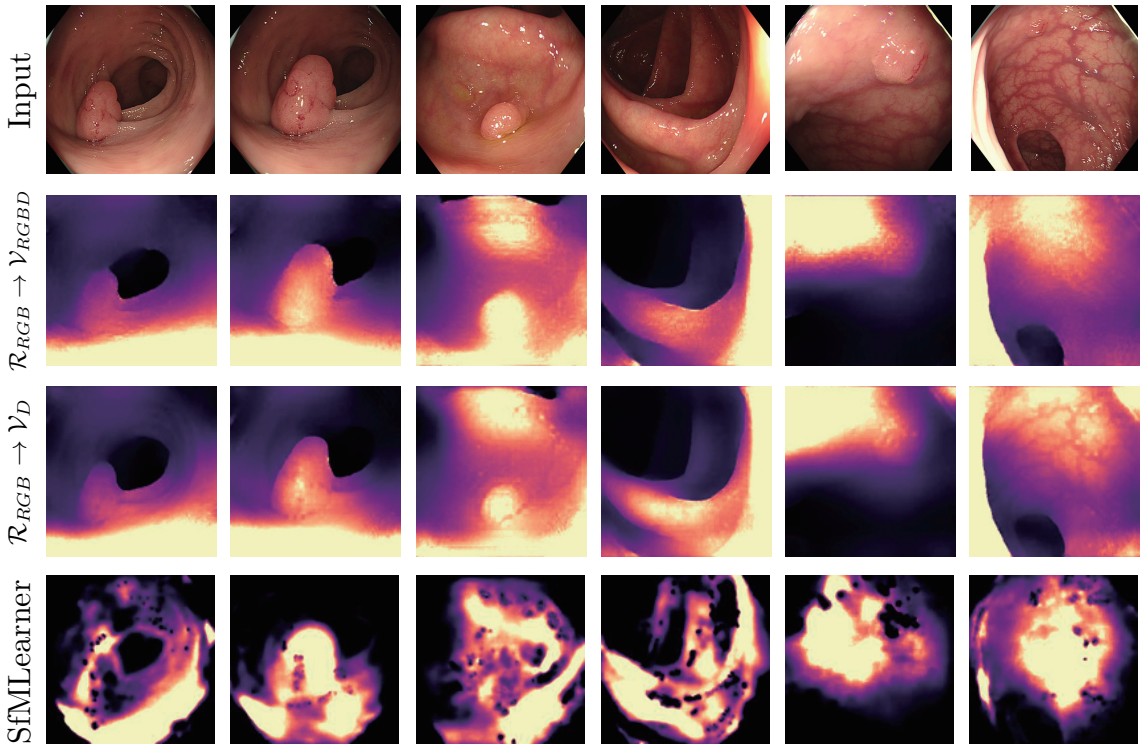


図 2 3 つの深度推定モデルにおける disparity map の比較．最上段の行が入力画像を表す．2-4 段目の行がそれぞれ提案手法，内視鏡画像と仮想深度画像のドメイン変換，そして SfMLearner による深度推定結果の diaparity map を示す．明るい暖色のピクセルが近距離であることを示す．

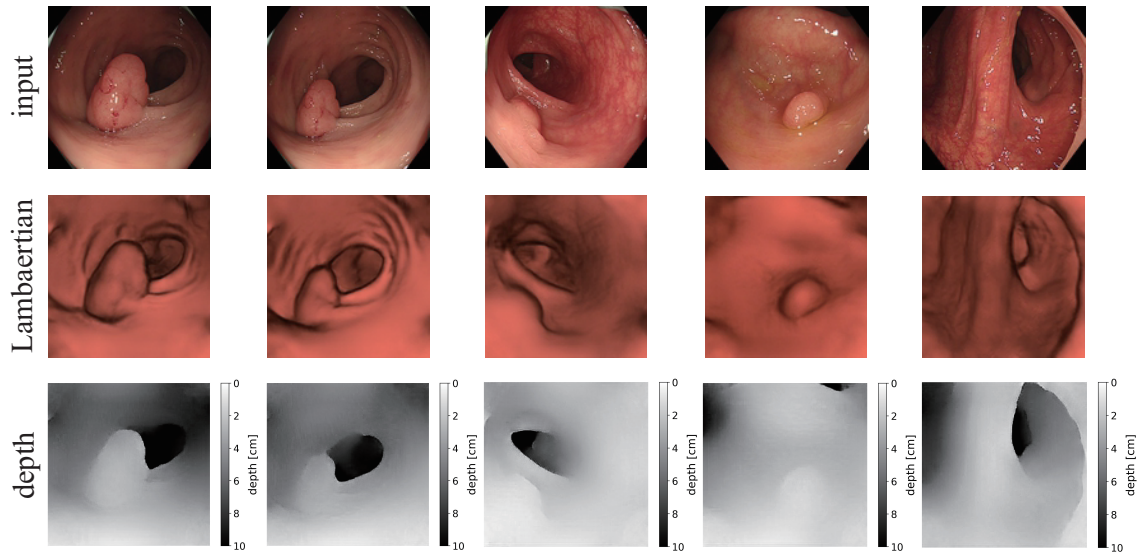


図 3 提案手法によって生成された拡散反射に基づく輝度画像と深度画像．

問題があるのではないかと予想される．さらに，鏡面反射などの影響も確認された．SfMLearnerでは拡散反射のみを入力画像に仮定しているものの，モデルの学習段階で拡散反射を考慮したパラメータ更新は行われていないことが原因と予想される．

提案手法は拡散反射に基づく輝度分布推定を介して，モデルの学習段階で拡散反射を想定したパラメータ更新を行う．図 2 の左から 3 列目や最右列目における比較が顕著に差を表しているが，補助タスク無しのドメイン変換による深度推定では血管のような腸壁のテクスチャが深度推定に大きく影響を与えているものの，提案手法ではテクスチャの影響なく滑らかな深度推定を達成している．図 3 の結果も同様の傾向を示す．

5. まとめ

拡散反射に基づく輝度分布推定を補助タスクとして導入した深度推定手法を提案し，補助タスク無しの深度推定法そして視差推定と多視点幾何学に基づく深度推定手法との比較検討を行った．実験結果より，提案手法に導入した拡散反射に基づいた輝度分布の推定がテクスチャの影響を軽減し，比較手法よりも大腸腸壁の形状をより正しく捉えた深度推定を達成していることを確認した．この深度推定の高精度化の達成より，導入した拡散反射に基づく輝度分布の推定が深度推定の高精度化に寄与する補助タスクであることを明らかにした．

謝辞 本研究の一部は JSPS/MEXT 科研費 (No. 26108006, 17H00867), 堀科学芸術振興財団そして AMED（No. 18hk0102034h0103, 19hs0110006h0003, 19lk1010036h0001) の支援を受けた。

利益相反 なし

文献

[1] Itoh H, *et al.*: Towards Automated Colonoscopy Diagnosis: Binary Polyp Size Estimation via Unsupervised Depth Learning, Proc. MICCAI, LNCS **11071**: 611–619, 2018.

[2] Ma R, *et al.*: Real-Time 3D Reconstruction of Colonoscopic Surfaces for Determining Missing Regions, Proc. MICCAI, LNCS **11768**: 573–582, 2019.

[3] Chen RJ, *et al.*: SLAM Endoscopy Enhanced by Adversarial Depth Prediction, Proc. KDD'19 Workshop on Applied Data Science for Healthcare, 2019.

[4] Zhou T, *et al.*: Unsupervised Learning of Depth and Ego-Motion from Video, Proc. CVPR: 6612-6619, 2017.

[5] Godard C, *et al.*: Digging Into Self-Supervised Monocular Depth Estimation, Proc. ICCV: 3827-3827, 2019.

[6] Rau A, *et al.*: Implicit Domain Adaptation with Conditional Generative Adversarial networks for depth prediction in endoscopy, International Journal of Computer Assisted Radiology and Surgery, **14**:1167-1176, 2019.

[7] Jaderberg M, *et al.* Reinforcement Learning with Unsupervised Auxiliary Tasks. Proc. ICLR: 2017.

[8] Prados E and Faugeras O: Shape From Shading, *Handbook of Mathematical Models in Computer Vision*: 357-388, Springer, 2006.

[9] Zhu J, *et al.*: Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks, Proc. CVPR: 2242-2251, 2017.

[10] Mori K, Suenaga Y and Toriwaki J: Fast Software-based Volume Rendering using Multimedia Instructions on PC Platforms and Its Application to Virtual Endoscopy, Proc. SPIE Medical Imaging, **5031**: 111-122, 2003.

Auxiliary tasks in unsupervised depth estimation for a colonoscope

Hayato Itoh[*1], Masahiro Oda[*1], Yuichi Mori[*2], Masashi Misawa[*2], Shin-Ei Kudo[*2], Kinnichi Hotta[*3], Hirotsugu Takabatake[*4], Masaki Mori[*5], Hiroshi Natori[*6], Kensaku Mori[*1,7,8]

[*]1 Graduate School of Informatics, Nagoya University
[*]2 Digestive Disease Center, Showa University Northern Yokohama Hospital
[*]3 Division of Endoscopy, Shizuoka Cancer Center
[*]4 Department of Respiratory Medicine, Sapporo-Minami-Sanjo Hospital
[*]5 Department of Respiratory Medicine, Sapporo-Kosei General Hospital
[*]6 Department of Respiratory Medicine, Keiwakai Nishioka Hospital
[*]7 Information Technology Center, Nagoya University
[*]8 Research Center for Medical Bigdata, National Institute of Informatics

Extraction of information about three-dsimensional structure of a human body is an essential task in medical image understanding. However, a colonoscope offers only a single two-dimensional image or time-series two-dimensional images. Extraction and estimation of three-dimensional structure of a colon from a single-shot image is required for a construction of computer-aided-diagnosis systems for colonoscopy. Previous works proposed depth-estimation methods based on disparity estimation with auxiliary task of a camera-motion estimation. These methods achieve an inaccurate estimation of relative depth information on an image. In this work, towards precise depth estimation, we propose Lambertian-reflection model to depth estimation, and experimentally compare its performance with the previous work. The experimental results show that the proposed auxiliary task reduces depth-estimation error caused by texture of a colon wall, and achieves accurate depth estimation.

Keywords: Colonoscopy, deep learning, depth estimation, auxiliary task, Lambertian refrection

位置特徴量とマルチスケール特徴量による胃壁マイクロ CT 像からの胃壁の層構造及び腫瘍の抽出

御手洗　翠[*1]　小田　紘久[*1]　杉野　貴明[*1]　守谷　享泰[*1]

伊東　隼人[*1]　小田　昌宏[*1]　小宮山　琢真[*2]　古川　和宏[*2]

宮原　良二[*2]　藤城　光弘[*2]　森　雅樹[*3]

高畠　博嗣[*4]　名取　博[*5]　森　健策[*1,6,7]

要旨

本稿では，Spherical K-means (SpK)により抽出された複数スケールの特徴量と位置特徴量を用いた胃壁 μ CT 画像からの粘膜層，粘膜下層，筋層及び腫瘍の抽出手法について報告する．胃壁の μ CT 画像から解剖学的構造を抽出することで腫瘍の立体構造把握が可能である．従来手法では複数スケールの特徴抽出フィルタを SpK によって学習し，パッチから特徴抽出を行った．その後得られた特徴量に基づいて SVM によって分類を行っていた．しかし，SpK により生成された特徴抽出フィルタは濃度値に基づく特徴量を抽出するため，粘膜層と筋層のように濃度値が類似した領域同士の特徴量が近しい値になる．本手法では SpK により得られる複数スケールの特徴量にパッチの位置特徴量を追加して分類する事によって，濃度値に基づく特徴量では分類できない領域の分類を可能にする．パッチの位置特徴量としてパッチの中心座標を用いる．本手法を胃壁 μ CT 像に適用した結果, 粘膜層, 粘膜下層, 筋層及び腫瘍の抽出に対する DICE 係数は 68.3%, 63.2%, 82.2%, 69.1%であった．　特に従来手法と比べ，腫瘍の抽出に対する DICE 係数は 26.1%向上した．

キーワード：マイクロ CT，セグメンテーション，胃壁

1. はじめに

胃壁のマイクロ CT 画像からは胃壁の層構造及び腫瘍などの解剖構造を 3 次元的に観察できるため，胃壁構造の立体把握が可能になる．しかし，各組織の濃度値が類似していることから，目視による層構造及び腫瘍の判別は困難である．そのため，各組織を半自動抽出する手法について検討した．従来手法[1]では Spherical K-means (SpK) によるマルチスケール特徴抽出を行い，抽出した特徴量を SVM に入力として与え，ラベルの割り当てを行った．しかし，腫瘍によって層構造が乱れている症例において，抽出精度低下がみられた．これは，SpK により抽出される特徴量が濃度値に依存するものであり，粘膜層と筋層のように濃度値が類似した領域同士の識別には有効ではないことが原因として考えられる．そこで胃壁の層構造の位置関係を識別に利用するために，位置特徴量を追加

*1 名古屋大学大学院情報学研究科
〔〒464-8601　名古屋市千種区不老町〕
e-mail: mmitarai@mori.m.is.nagoya-u.ac.jp
*2 名古屋大学大学院医学系研究科
*3 札幌厚生病院
*4 札幌南三条病院
*5 恵和会西岡病院
*6 名古屋大学情報基盤センター
*7 国立情報学研究所医療ビッグデータ研究センター

で利用することを検討したので報告する．

2. 手法

2.1 概要

入力は胃壁 3 次元マイクロ CT 画像（入力画像）と胃壁の粘膜層，粘膜下層，筋層及び腫瘍をアノテーションした 1 スライスの正解ラベル画像とする．入力画像の正解ラベルが存在する領域から抽出した特徴量を用いて SVM の訓練を行う．次に入力画像全体から切り出した局所領域の特徴量を訓練済みの SVM に入力し，各領域のラベルの割り当てを行う．本手法は(1) SpK による特徴抽出フィルタバンク生成，(2) フィルタバンクによる特徴量及び位置特徴量を用いた SVM の訓練，(3) SVM による分類の 3 項目からなる．また，図 1 に示すように胃壁の深さ方向を x 軸方向，胃壁の層に対して水平方向を y 軸方向，xy 平面に直交する方向を z 軸方向とする．

2.2 SpK による特徴抽出フィルタバンク生成

入力画像から 3×3×3 画素のパッチを複数切り出す．パッチ内の濃度値を 1 次元ベクトルとしたものを特徴ベクトルとする．次に SpK を用いて特徴ベクトル集合の K 個のクラスタ重心を算出する．各クラスタ重心ベクトルが線形フィルタの重みベクトルとして利用可能であるため，各クラスタの重心ベクトルを特徴抽出フィルタとして扱う[2]．次に 5×5×5 画素及び 7×7×7 画素のパッチを切り出し同様の処理を適用することで各スケールの特徴抽出フィルタを生成する．異なるサイズのパッチによって得られる特徴抽出フィルタをまとめてマルチスケールフィルタバンクとする．

2.3 フィルタバンクによる特徴量及び位置特徴量を用いた SVM の訓練

入力画像の正解ラベルが存在する画素を中心として，41×41×15 画素の領域を切り出す．この切り出す領域が非等方であるのは，胃壁構造の抽出には xy 平面上の層構造に関する画像特徴が識別に有効であるからである．切り出した領域内において，フィルタバンク内の各フィルタを適用し特徴抽出を行う．あるスケールのフィルタから得られる領域内での特徴ベクトルの集合から sum pooling により 8 個の特徴ベクトルを算出，連結することで，$8K$ 次元の特徴ベクトルを得る．これを切り出した領域に対して 3 スケールに渡って行い，得られた特徴ベクトルを連結することで $24K$ 次元特徴ベクトルとして得る．得られた特徴ベクトルに，切り出した領域の xy 平面上での中心座標（x, y）を追加する．抽出された特徴ベクトルと切り出した領域の中心における正解ラベルを用いて SVM を訓練する．

2.4 SVM による分類

入力画像全体から 41×41×15 画素の領域を切り出し，2.3 節と同様に特徴抽出を行う．特徴ベクトルを SVM に入力，分類を行い，その結果を対応する画素におけるラベルとして扱い，各領域に分類結果を割り当てる．

3. 実験と結果

胃癌患者から摘出した胃壁のマイクロ CT 像 5 画像に本手法を適用した．学習用の正解ラベルは各画像につき，xy 平面をアノテーションした 1 スライスである．提案手法では各スケールそれぞれ 10000 個のパッチを用いて SpK のクラスタ重心を算出した．K=75 とした．1500 個の 41×41×15 画素の領域を用いて SVM を訓練した．評価用スライスにおける粘膜層，粘膜下層，筋層及び腫瘍の抽出結果について DICE 係数で評価を行った．抽出結果を図 2 と表 1 に示す．

4. 考察

提案する手法ではマイクロ CT 像からマルチスケール特徴量と位置特徴量を用いた抽出手法を提案した．表 1 より従来手法と比べ，各構造の抽出精度向上を確認した．特に表 1 から分かるように従来手法に比べ腫瘍の抽出精度は大きく向上した．図 2 では従来手法に比べ，腫瘍領域及び筋層領域内部の誤ったラベル割り当ての大幅な減少が確認できる．これは従来の濃度値に基づく特徴量では粘膜層，筋層及び腫瘍はコントラストが低いことから識別が困難

だったが，位置特徴量によって識別可能になったためであると考えられる．これは位置特徴量を利用することで，胃壁の層の位置関係のような解剖学的構造の特徴を識別に用いることができたためと考える．

5. むすび

本稿では，SpK による特徴抽出と位置特徴量を用いて胃壁マイクロ CT 像から各構造を抽出する手法を提案した．提案手法は従来手法に比べ，各組織の抽出精度の向上を確認した．今後は抽出精度向上のため，各組織の形状に関する特徴量を利用することを検討する．

謝辞　本研究の一部は堀科学芸術振興財団，JSPS 二国間交流事業ならびに JSPS 科研費 (26108006，17H00867，17K20099) によった．
利益相反の有無　なし

文　献

［1］　御手洗翠，小田紘久，小宮山琢磨 他:マルチスケール特徴抽出による胃壁マイクロ CT 像からの解剖学的構造セグメンテーション手法．日本コンピュータ外科学会誌 第 28 回日本コンピュータ外科学会大会特集号 **21**: 273-274, 2019

［2］　Moriya T, Roth HR, Nakamura S, et al.: Unsupervised pathology image segmentation using representation learning with spherical k-means. SPIE Medical Imaging 2018 Digital Pathology **10581**: 1058111, 2018

表 1 提案手法による各組織の DICE 係数

	粘膜層	粘膜下層	筋層	腫瘍
従来手法[1]	54.1	47.6	77.8	43.0
提案手法	**68.3**	**63.2**	**82.2**	**69.1**

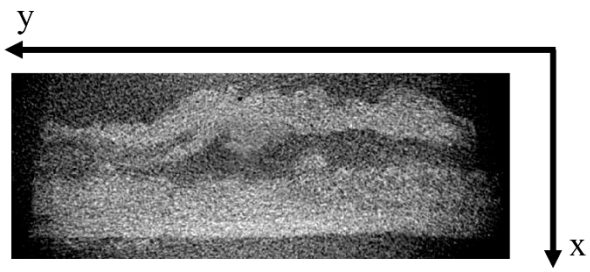


図 1 x 軸及び y 軸方向の定義．層構造の深さ方向を x 軸，層に対して水平方向を y 軸とする．

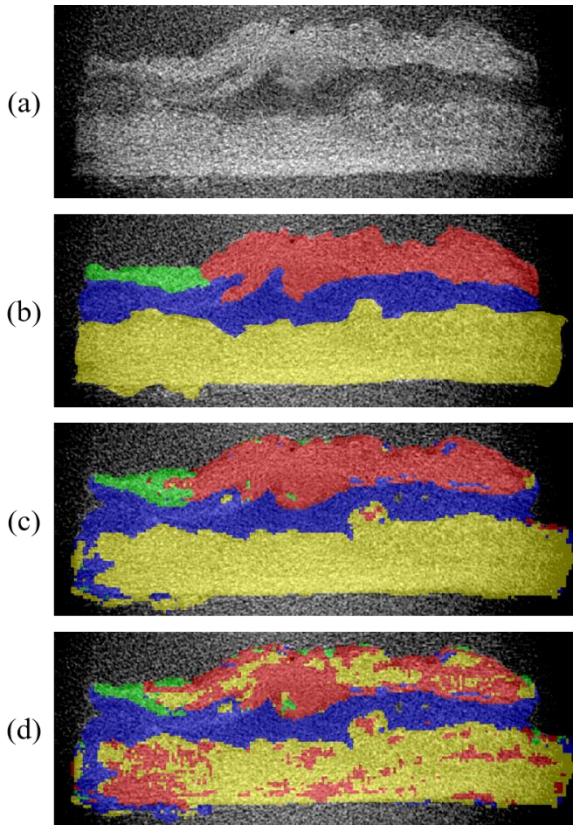


図 2 胃壁マイクロ CT 像の例．粘膜層は緑色，粘膜下層が青色，筋層が黄色，腫瘍が赤色のラベルである．(a) 入力画像．(b) 正解ラベル．(c) 提案手法による抽出結果．(d) 従来手法[1]による抽出結果．

Segmentation of Layer Structure of the Stomach Wall and Tumor on MicroCT Volumes using Spatial and Multi-scale Features

Midori MITARAI [*1], Hirohisa ODA[*1], Takaaki SUGINO [*1], Takayasu MORIYA [*1],
Hayato ITOH [*1], Masahiro ODA [*1], Takuma KOMIYAMA [*2], Kazuhiro FURUKAWA[*2],
Ryouji MIYAHARA[*2], Mitsuhiro FUJISHIRO [*2], Masaki MORI[*3],
Hirotsugu TAKABATAKE [*4], Hiroshi NATORI [*5] and Kensaku MORI [*1,6,7]

*1 Graduate School of Informatics, Nagoya University
*2 Nagoya University Graduate School of Medicine
*3 Sapporo-Kosei General Hospital
*4 Sapporo Minami-sanjo Hospital
*5 Keiwakai Nishioka Hospital
*6 Nagoya University Information Technology Center
*7 Research Center for Medical Bigdata, National Institute of Informatics

This paper reports a segmentation method of the mucosa layer, submucosa, muscle layer and tumor from stomach wall μCT volumes using spatial feature and multi-scale features extracted by Spherical K-means (SpK). The 3D structure of the tumor can be understood by extracting the anatomical structure from the μCT volumes of the stomach wall. In the conventional method, feature extraction filters of multi-scale were learned by the SpK, and feature extraction was performed from patches. Classification was performed by the SVM based on the obtained multi-scale features. However, since the feature extraction filter generated by the SpK extracts features based on intensity values, the features in regions with similar intensity values, such as the mucosal layer and the muscle layer, have close values. In the proposed method, the spatial features of patches are added to the multi-scale features obtained by the SpK for classification, which enables the classification of regions that cannot be classified by the features based on intensity values. The center coordinates of the patches are used as the spatial features of the patch. Experimental results showed that the proposed method can segment mucosa layer, submucosa, muscle layer and tumor from stomach wall μCT volumes with DICE coefficients of 68.3%, 63.2%, 82.2% and 69.1%. Especially, compared with the conventional method, the DICE coefficient of the tumor was improved by 26.1%.

Key words: Micro CT, Segmentation, Stomach wall

屈折コントラストX線CTを用いた非浸潤性乳管癌と通常型乳管過形成の腺腔における3次元的な差異

谷口　華奈[*1]　砂口　尚輝[*1]　市原　周[*2]　西村　理恵子[*2]
岩越　朱里[*2]　島雄　大介[*3]　湯浅　哲也[*4]　安藤　正海[*5]

要旨

乳癌など乳腺疾患の病理診断は，組織切片の顕微鏡観察により 2 次元的に行われる．その際，乳管内部に存在する腺腔の形状は診断のために重要な指標となる．この腺腔は 3 次元的に観察することで 2 次元平面では得られない診断に有用な様々な情報が得られる．しかし，組織切片の 3 次元化には膨大な労力を要し診断で利用されることは少ない．また，腺腔の 3 次元形状を報じる研究もこの労力のために非常に少ない症例で議論している．本研究では，腺腔観察の新たなアプローチとして，染色組織像と同等のコントラストが得られる屈折コントラスト X 線 CT を用いて，非浸潤性乳管癌(DCIS)と良性病変である通常型乳管過形成の各組織内で生じる腺腔を 3 次元的に観察する．得られた 3 次元像から腺腔を抽出し腺腔の密度や経路長を測定した結果，各組織間に明確な差異が確認された．これらのパラメータは DCIS や良性乳管内増殖性病変を鑑別する上で重要な指標になり得る．

キーワード：屈折コントラスト X 線 CT，非浸潤性乳管癌，通常型乳管過形成，篩状，3 次元可視化

1. はじめに

国立がん研究センターの最新がん統計によると，乳癌は現在，女性の罹患数の中で最も多い部位の癌であり，死亡数は大腸，肺，膵臓，胃に次いで第 5 位である [1]．乳癌の多くは乳管から発生する乳管癌であり，中でも癌細胞が乳管系にとどまり間質への浸潤がみられないものを非浸潤性乳管癌（Ductal carcinoma in situ; DCIS）という．病理診断学では，篩状型（cribriform type）DCIS と良性乳管内増殖性病変で鑑別の難しさが指摘されている．通常，篩状型 DCIS の腺腔は正円形に近く，大きさが揃っており，多孔性を示す．一方，良性の乳管内増殖性病変である通常型乳管過形成（Usual ductal hyperplasia; UDH）の腺腔は不規則な非円形で，スリット状の構造を示すことが多いとされる[2]．しかし，組織切片の切削方向によっては UDH の腺腔も正円形になる可能性があり，2 次元平面では両者の区別が付き辛い．

1985 年，Ohuchi らは DCIS と乳管過形成の腺腔構造の 3 次元再構築を行い，DCIS は散在した球状，乳管過形成は連続したネットワーク状

*1 名古屋大学大学院医学系研究科
〔〒461-8673 名古屋市東区大幸南 1-1-20〕
e-mail:taniguchi.kana@f.mbox.nagoya-u.ac.jp
*2 名古屋医療センター病理診断科
*3 北海道科学大学保健医療学部
*4 山形大学大学院理工学研究科
*5 総合科学研究機構
投稿受付：2020 年 8 月 5 日

であることを報告している[3]．この研究により，2 次元の病理観察では理解されていなかった腺腔の新たな 3 次元構造が明らかになり，3 次元的特徴が良悪性鑑別に利用できることが示された．2012 年，Norton らによって DCIS で生じた篩状構造の 3 次元再構成が行われ，篩状構造にはチューブ状とバブル状の 2 種類の 3 次元構造が存在することが報告された[4]．他にも 3 次元観察を行った研究は複数存在するが，すべて 2 次元平面である病理切片を積み重ねて 3 次元像を再構成している．この 3 次元化には膨大な数の病理切片を必要とし多大な労力を要するため，これまでの報告はすべて症例数が少ない．

我々は腺腔構造の新たな 3 次元可視化手法として，屈折コントラスト X 線 CT を使用する[5]．この方式は，染色した組織像と同等のコントラストでの撮像が可能であり，病理診断学への応用に向けた研究が進められている[6]．屈折コントラスト CT を使用する利点として非破壊・未染色で短時間に 3 次元像を得ることができることが挙げられる．また，病理像を用いる手法と比較し，ズレが生じる可能性のあるスライス間接続の過程を省略できることや，様々な画像処理を利用できることから，この手法は多くの症例を用いる研究に適していると考えられる．

本研究では，屈折コントラスト X 線 CT を用いて，DCIS と UDH を含む 2 種類の乳腺組織試料を撮像し，腺腔構造の 3 次元可視化および解析を行う．また，DCIS と UDH における 3 次元的な腺腔構造の差異を表現する新しい指標を提案する．

2. 撮像手法・試料および解析手法

図 1 は茨城県つくば市の高エネルギー加速器研究機構フォトンファクトリーBL14B ビームラインに構築された屈折コントラスト CT 撮像システムの概念図である．このシステムは試料の前後に非対称ブラッグケースモノクロコリメーター結晶（MC）とラウエケース角度アナライザー結晶（LAA）を配置する．加速器から発生した放射光 X 線は 2 結晶モノクロメーターにより単色化され，MC で回析した後，試料に到達する．試料を伝搬する過程で X 線は試料の電子密度に従い微小に屈折する．その角度偏差は LAA で X 線の強度情報に変換された後，X 線カメラで測定される．試料の回転毎に測定を繰り返すことで CT の再構成に必要な投影セットが得られる．X 線エネルギーは 19.8 keV, MC と LAA の各結晶の回折面は Si(111)であり，ピクセルサイズは 14.8 μm である．投影数は 1250 枚で，再構成アルゴリズムには逐次的再構成法[7]が採用された．

試料は乳癌と診断された日本人女性から摘出された DCIS と UDH を含む 2 種類の乳腺組織ブロックである．これらは 10%緩衝ホルマリン液による固定後，エタノールに浸して保存されており，撮像の 5〜6 時間前に，エタノールを除去するために蒸留水に移される．撮影の際，試料内のコントラストを向上させるために，試料は水を入れた水槽の中で回転する．本研究は，名古屋大学および名古屋医療センターの生命倫理審査で承認が得られている．

1）腺腔構造の 3 次元可視化

CT 画像から腺腔領域を抽出するために，画像処理ソフト Image J [8]の Drawing Tools が使用された．腺腔領域が手動で抽出された後，観察のためにボリュームレンダリング像が作成された．3 次元像のスライス厚は 14.8 μm で，DCIS と UDH のボリュームサイズはそれぞれ 184×166×33 μm^3, 218×235×35 μm^3 である．

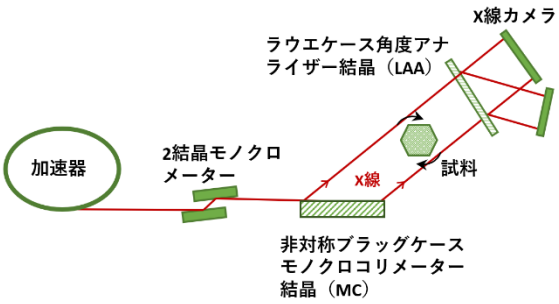


図 1. 屈折コントラスト CT 撮像システム．

2）腺腔構造の 3 次元解析

CT ボリュームから，(1) 腺腔の個数 (2) 腺腔体積 (3) 腺腔密度 (4) 腺腔構造の経路長が測定された．

腺腔の個数は，観察範囲内の独立した腺腔の個数を数えたものである．腺腔体積は，各腺腔の体積の平均である．腺腔密度は，単位体積あたりの腺腔の個数である．腺腔構造の経路長は，各腺腔に細線化処理を施した後の各線の長さの平均である．

3. 結果

図 2 は，DCIS と UDH それぞれの屈折コントラスト CT 像（左）と CT 像から腺腔領域を抽出し 2 値化処理を行った 2 値化像（右）である．CT 像の乳管内に存在する数 10μm 程度の黒い領域が腺腔である．DCIS では円に近い粒状の腺腔であり，UDH では線状で不規則な形態の腺腔であることが確認された．

1）腺腔構造の 3 次元可視化

図 3 に DCIS と UDH それぞれの腺腔の 3 次元像を示す．DCIS では，球状に近い腺腔が独立して存在しているが，UDH では，連続性のある腺腔が多く確認された．

2）腺腔構造の 3 次元解析

腺腔の個数は，DCIS で 541 個，UDH で 352 個であった．各体積の腺腔の割合を**図 4** で示す．このグラフは，20×10^3 μm^3 未満の体積の腺腔が，DCIS では全腺腔の 26.8%，UDH では 13.9%，20×10^3 μm^3 以上 50×10^3 μm^3 未満の体積の腺腔が，DCIS では全腺腔の 33.8%，UDH では 31.5% を占めていることを表す．平均腺腔体積は，DCIS で 95.0×10^3 μm^3，UDH で 178.8×10^3 μm^3 である．腺腔密度は DCIS で 165.6 個/mm^3，UDH で 60.6 個/mm^3 であった．各経路長の腺腔の割合を**図 5** で示す．このグラフは，50 μm 未満の経路長を持つ腺腔が，DCIS では全腺腔の 43.1%，UDH では 21.6%，50 μm 以上 100μm 未満の経路長を持つ腺腔が，DCIS では全腺腔の 30.1%，UDH では 26.7%を占めていることを表す．平均腺腔経路長は，DCIS で 113.6 μm，UDH で 223.4μm であった．

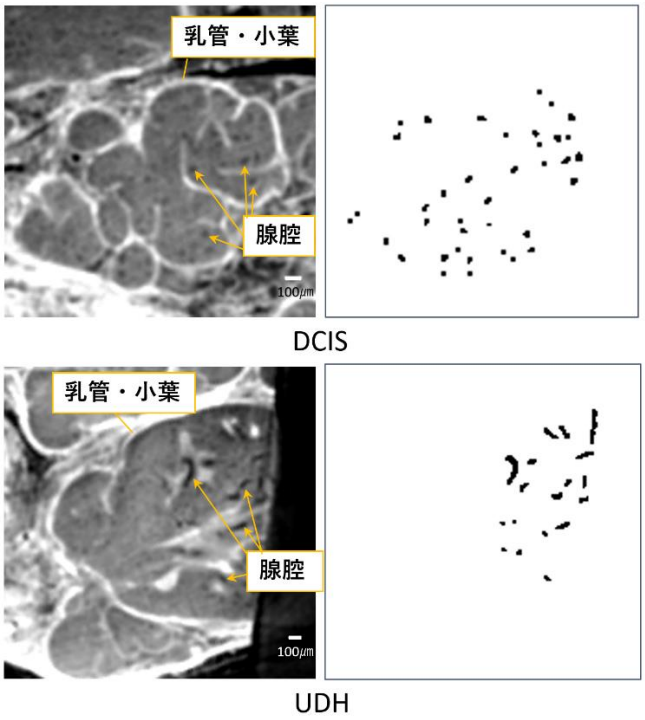


図 2. 屈折コントラスト CT 像(左)と 2 値化像(右).

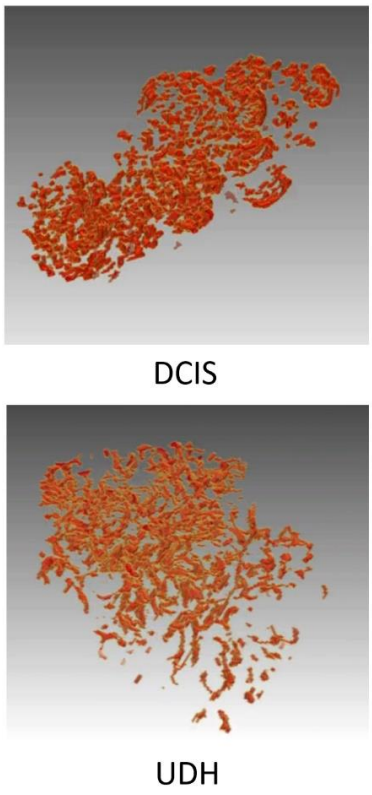


図 3. 3 次元可視化像.

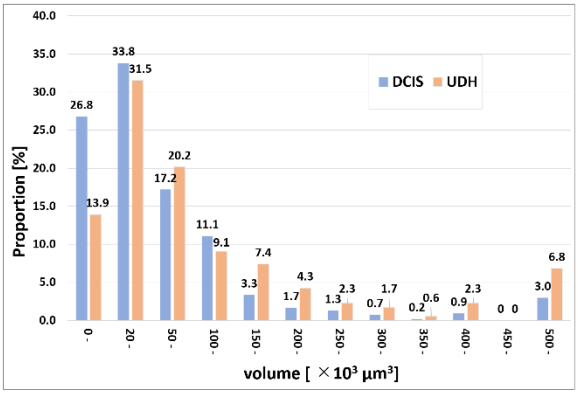


図 4. 各体積の腺腔の割合.

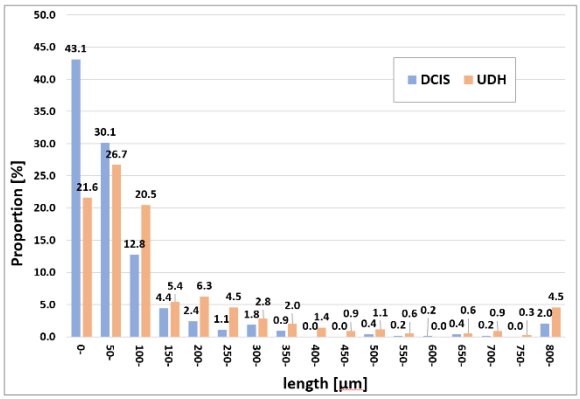


図 5. 各経路長の腺腔の割合.

4. 議論

DCIS や UDH などの乳管内増殖病変は，その後の治療が適切に決定されるために，正確に診断される必要がある．DCIS と UDH の鑑別は，核の形態や増殖パターンなどから総合的に判断され，特に増殖細胞の集塊内に生じる二次的空間（腺腔）の形態が重要な手がかりとなる[2].

図 3 の 3 次元可視化像より DCIS と UDH の腺腔構造を比較すると，DCIS は，球状の腺腔が多数散在しており，多孔性構造であることが分かる．これは 2 次元病理像の篩状構造に相等すると考えられる．一方 UDH では，連続性のある腺腔が多く確認できる．これは 2 次元病理像でみられる複雑な腺腔パターンやスリット状の構造と一致していると考えられる．DCIS ではバブル状, UDH では管状の腺腔が多くみられ，2 つの疾患の腺腔構造に差異があることが確認された．また， DCIS と UDH の構造の差異を個数，腺腔体積，腺腔密度，経路長で示すことにより，DCIS は UDH と比較して，体積が小さく経路長が短い腺腔の存在確率が高いことが分かった（図 4 および図 5）．これらの 3 次元的な特徴は，DCIS と良性の乳管内増殖病変の鑑別において重要な指標になり得る．

5. まとめ

本研究では，DCIS と UDH の腺腔構造を可視化し，DCIS ではバブル状, UDH では管状を形成することを確認した．また, 3 次元構造の差異を表すパラメータとして腺腔の個数，腺腔体積，腺腔密度，腺腔構造の経路長を導入し，UDH と比較して DCIS は小さく短い腺腔が単位体積あたり多数存在していることを示した．これらの 3 次元情報は，乳管内増殖病変の診断精度の向上に寄与できると考える．今後，屈折コントラスト CT を用いてさらに症例数を増やし本結果の信頼性を高めるとともに，新たな組織学的情報の取得を目指したい．

謝辞

本研究の試料を準備していただいた名古屋医療センター病理診断科の皆様に感謝申し上げます．また, 撮像システムで使用された LAA 結晶は KEK の PF 結晶加工室を利用して杉山弘助教および笹谷典太氏によって作成されました．ここで感謝を申し上げます。また，貴重な X 線 CCD カメラを貸していただいた KEK の平野馨一准教授および兵藤一行教授に感謝申し上げます．本研究は JSPS 科研費 16K01369, 16K08654, 22591353, 15H01129, 26286079, 18K13765，平成 29 年度国立病院機構共同臨床研究 H29-NHO(癌般)-01, H29-NHO(多共)-02，知の拠点あいち重点プロジェクト第 3 期および KEK 放射光実験課題 2016G0625 の支援を受けています．

利益相反の有無

なし．

文　献

［1］ 国立がん研究センター：最新がん統計. https://ganjoho.jp/reg_stat/statistics/stat/su

mmary.html（2020 年 7 月 27 日閲覧）
［2］ 市原　周：新版 乳腺病理学－細胞・組織・画像. 名古屋大学出版会，名古屋，2013，pp.63-87
［3］ Ohuchi N，Abe R，Takahashi T et al.：Three-dimensional atypical structure in intraductal carcinoma differentiating from papilloma and papillomatosis of the breast. Breast Cancer Research and Treatment 5: 57-65 ，1985
［4］ Norton KA，Namazi S，Barnard N et al.：Automated Reconstruction Algorithm for Identification of 3D Architectures of Cribriform Ductal Carcinoma In Situ. PLoS ONE 7:e44011, 2012
［5］ Sunaguchi N, Yuasa T, Huo Q et al.：X-ray refraction-contrast computed tomography images using dark-field imaging optics. Appl Phys Lett 107:253701，2015
［6］ 砂口 尚輝，島雄 大介，市原 周 他：X線暗視野法に基づく屈折コントラストＸ線 CT を用いたヒト乳頭の三次元可視化および解析．Medical Imaging Technology 38:1，2020
［7］ Sunaguchi N, Yuasa T, Gupta R, et al.：An efficient reconstruction algorithm for differential phase-contrast tomographic images from a limited number of views. Appl Phys Lett 107:253701, 2015
［8］ アメリカ国立衛生研究所（NIH）: ImageJ. https://imagej.nih.gov/ij/

Three-dimensional differences between ductal carcinoma in situ and usual ductal hyperplasia tissues using the refraction-contrast computed tomography

Kana TANIGUCHI[*1], Naoki SUNAGUCHI[*1], Shu ICHIHARA[*2], Rieko NISHIMURA[*2], Akari IWAKOSHI[*2], Daisuke SHIMAO[*3], Tetsuya YUASA[*4], Masami ANDO[*5]

*1 Nagoya University
*2 Nagoya Medical Center
*3 Hokkaido University of Science
*4 Yamagata University
*5 Comprehensive Research Organization for Science and Society (CROSS)

The pathological diagnosis of breast cancer and other breast diseases is made in two dimensions by microscopic observation of tissue sections. The shape of the glandular lumina within ducts is an important indicator for diagnosis. By observing lumina in three dimensions, various information useful for diagnosis that cannot be obtained in a two-

dimensional plane can be obtained. However, three-dimensional visualization of tissue sections requires a great deal of labor and is not used in ordinary diagnostics. Also, because of this, it's been discussed in very few cases in studies that report 3-D observations. In this study, we use refraction-contrast X-ray CT, which provides the same contrast as stained tissue images, as a new approach to observe lumina in three dimensions in each tissue of Ductal carcinoma in situ (DCIS) and benign lesion, Usual ductal hyperplasia (UDH). The luminal space was extracted from the 3D image and the density and pathway length of lumina were measured, and clear differences were observed in each tissue. These parameters may be important in differentiating DCIS and benign intraductal proliferative lesions.
Key words: Refractive Contrast X-ray CT, Ductal carcinoma in situ, Usual ductal hyperplasia, cribriform pattern, 3D Visualization

著者紹介



谷口　華奈（たにぐち　かな）

2020 年 藤田医科大学医療科学部放射線学科卒業．現在，名古屋大学大学院医学系研究科総合保健学専攻博士前期過程．日本医用画像工学会学生会員．

Visual SLAM を用いた内視鏡位置推定に関する基礎検討

上野　創史[*1]　藤橋　卓也[*1]　菊川　忠彦[*2]　猿渡　俊介[*1]

安藤　英由樹[*3]　雑賀　隆史[*2]　渡辺　尚[*1]

要旨

映像データを用いた内視鏡手術の遠隔指導は指導する熟練外科医・指導を受ける外科医間の物理的制約を緩和する．一方で，内視鏡手術において必要となる内視鏡および鉗子の細かな操作技術を熟練外科医が映像データのみから適切に指導することは困難である．内視鏡手術を対象とした遠隔指導において，外科医間での円滑なやりとりを実現するためには内視鏡から得られた 2 次元映像を元にして再構成した手術視野内の立体情報を共有することが望ましい．本研究では手術視野内の立体情報を取得する 1 検討として，カメラ映像を入力としてカメラ位置姿勢と環境地図の推定を行う Visual SLAM (Simultaneous Localization And Mapping) を用いた内視鏡位置推定手法を提案する．本手法では，内視鏡で撮影した映像に含まれる特徴点を抽出するとともに内視鏡の位置および姿勢をリアルタイムで取得する．単眼の内視鏡カメラを用いた実験評価から提案手法による内視鏡位置の推定精度は平均 3.6cm 程度の誤差に収まることを確認した．

キーワード：遠隔指導，内視鏡位置推定，Visual SLAM

1. はじめに

手術時の身体的負担の軽減や入院期間の短縮につながることから，昨今ではあらゆる外科領域で従来の開腹手術に代わって内視鏡手術が行われる機会が多くなっている［1］．内視鏡手術は外科的手術を小さく抑えて病変を切除できるため，開腹手術と比較して侵襲性が少なく，手術後の患者の臥床期間を短縮することにつながる．また，手術に要する時間も 1〜2 時間程度で済むことから執刀医や助手，麻酔科医などのスタッフの負担軽減にもつながる．

内視鏡手術では，患者の腹部に穴を数ヶ所開けてそれぞれから内視鏡や鉗子などの手術器具を患者の体内に挿入する．執刀医は内視鏡先端に取り付けられたカメラ（以下，内視鏡カメラ）の映像をもとに鉗子を操作して患者の体内の病変を切除する．このとき，執刀医は内視鏡カメラを動かして映像を自由に拡大/縮小することで病変を詳しく観察できる．

一般的に，内視鏡手術の執刀医には高度な技術が要求される．開腹手術は患者の体内を実際に観察する一方，内視鏡手術では内視鏡カメラから得られた二次元映像をもとにして執刀することから立体感覚の知覚に制限がかかる［2］．また，手術に用いる鉗子は細長い特殊なものを使用するため，操作性や触圧覚が制限され，手術の困難性は高くなる．

一方で，少子高齢化が問題視されている現在では，現場を離れる内視鏡手術に長けた熟練外科医の数［3］が増加している．内視鏡手術に長けた高齢の外科医が現場を離れる前に，若手外科医に高度な専門技術を伝承することが医療現場において喫緊の課題となっている．若手外科医が内視鏡手術に関する技術を習得するためには，まず若手外科医がトレーニング用の教材を用いて各自で十分に練習したあと，実際の手術に相当する環境下で熟練外科医から直接指導を受ける方法［4］が取られている．しかしながら，この方法による技術伝承では都市部の医療機関に属する若手外科医と地方の医療機関に属する若手外科医との間

*1 大阪大学大学院情報科学研究科
〔〒565-0871 吹田市山田丘 1-5〕
e-mail: ueno.soushi@ist.osaka-u.ac.jp
*2 愛媛大学大学院医学系研究科
*3 大阪芸術大学

で技術格差を招く．内視鏡手術に卓越した多くの熟練外科医は都市部の総合病院・大学病院に所属している．地方医療機関に所属する若手外科医が熟練外科医から直接指導を受けるためには，双方の都合を調整したうえで熟練外科医あるいは若手外科医が医療機関まで訪問する必要がある．

訪問にともなう外科医の負担を軽減するために，映像や音声を用いた遠隔指導が検討されている．遠隔指導では別々の医療機関に属する熟練外科医と若手外科医の一方もしくは双方が実施している内視鏡手術中の映像とその音声をネットワークを通じてやり取りする．遠隔指導は医療機関への移動にともなう負担を軽減できる一方，手術映像から得られる情報は限定的であるため，熟練外科医が経験の少ない若手外科医に分かるように指導を行うのは困難となる．例えば，熟練外科医が受信する内視鏡映像は若手外科医が内視鏡を動かして得られた映像であるため，内視鏡を動かしていない熟練外科医が受信した内視鏡映像から映像内の立体情報を推察することは困難である．この問題を解決する方法として，立体映像を生成して熟練外科医に伝送する方法や手術支援ロボット Da Vinci で利用可能なデュアルコンソールを用いて両外科医が同時に手術に参加する方法が挙げられるが，現状では価格的コストに起因して導入が困難である．

上述のとおり，内視鏡手術を対象とした遠隔指導において外科医間で円滑なやりとりを実現するためには，2 次元の内視鏡映像を元にして手術視野内の立体情報を再構成するとともに外科医間で立体情報を共有することが望まれている．本研究では，2 次元の内視鏡映像から手術視野内の立体情報を取得する 1 検討として Visual SLAM (Simultaneous Localization And Mapping)を用いた内視鏡位置推定手法を提案する．Visual SLAM とは，カメラ映像を入力としてカメラ位置姿勢と環境地図を作成する技術である．本研究では，推定した内視鏡の位置姿勢を利用することを考える．推定した内視鏡の位置姿勢は映像内に映った臓器の位置姿勢推定を可能にするとともに，映像上に臓器の 3D ファントムや指示内容を補助情報として重畳表示することも可能となる．単眼カメラおよび複眼カメラを用いた実験評価を通して，Visual SLAM による内視鏡位置推定手法の実現可能性について議論する．

2. 内視鏡 SLAM

2.1. 全体像

図 1 に本研究で想定する内視鏡カメラ映像を用いた遠隔指導の流れを示す．遠隔指導においては，内視鏡手術に長けた熟練外科医と不慣れな若手外科医が異なる医療機関に属するものとした．医療機関同士はネットワークを介して接続されているものとした．若手外科医は自身が属する医療機関にて手術を実施する．手術中の様子は内視鏡カメラでリアルタイムで撮影され，その映像と音声はネットワークを通じて遠隔の熟練外科医に伝送する．熟練外科医は受信映像をもとにして若手外科医の執刀に対して内視鏡や鉗子の操作方法にあげられる誤りを見つけ出し，その情報をネットワークを通じて若手外科医に返送する．熟練外科医の指導をもとにして若手外科医は内視鏡や鉗子の操作を修正して適切な手術処置を実践的に学ぶことができる．

図 2 に提案システムである内視鏡 SLAM の全体図を示す．外科医は内視鏡カメラを通して患者の体内の病変を撮影した映像を取得する．取得した映像を入力として Visual SLAM を用いて内視鏡の位置および姿勢を推定する．ここで，内視鏡の位置情報および姿勢情報とは内視鏡位置$X = (x, y, z)^{\mathrm{T}}$および内視鏡姿勢である四元数$q = q_0 + q_1\mathrm{i} + q_2\mathrm{j} + q_3\mathrm{k}$を意味する．四元数［5］$q$は複素数を用いてカメラ姿勢を表す概念である．内視鏡カメラの位置姿勢を取得した後，提案システムでは得られたカメラの位置情報・姿勢情報から 3D モデルの位置および姿勢を変換する．変換した 3D モデルは術者の手術視野内に内視鏡カメラ映像とともに表示する．

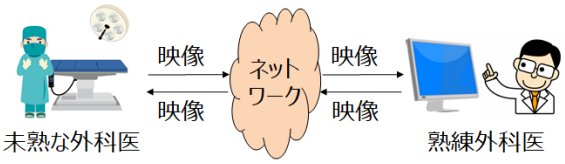


図 1. 想定する遠隔指導の流れ

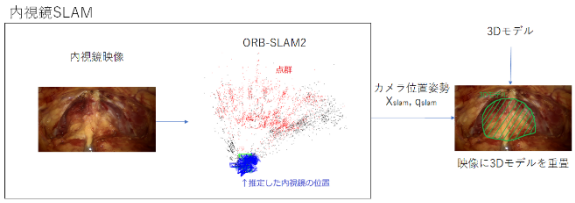


図 2. 提案システムの全体像

2.2. Visual SLAM を用いた内視鏡位置推定

内視鏡カメラ映像から内視鏡の位置および姿勢を推定するために，本研究では Visual SLAM(Simultaneous Localization And Mapping)［6］の一手法である ORB-SLAM2［7］を用いる．ORB-SLAM2 では内視鏡カメラ映像を構成する各ビデオフレームから得られる ORB 特徴量をもとにして内視鏡カメラの位置姿勢を推定する．具体的には連続するビデオフレーム間の対応する ORB 特徴量の位置の変化から現在のカメラ位置を推定する．ORB 特徴量とは FAST (Features from Accelerated Segment Test)による特徴点検出と BRIEF (Binary Robust Independent Elementary Features) による特徴量記述子を組み合わせたカメラの移動，回転，ズームに対して耐性を持つ特徴量である．ORB-SLAM2 では得られた ORB 特徴量をもとにしてトラッキング・ローカルマッピング・ループクローズの 3 処理を並行に実行することでカメラ位置姿勢を推定する．

本稿では取得した内視鏡カメラ映像から内視鏡カメラの位置姿勢を推定するトラッキング処理について述べる．トラッキング処理は主に ORB 特徴量の取得，カメラ位置姿勢の推定から構成される．ORB-SLAM2 では内視鏡カメラから新たなビデオフレームを取得すると，ビデオフレーム内に含まれる ORB 特徴量を取得する．現在のビデオフレームから取得した ORB 特徴量と直前のビデオフレームに含まれる ORB 特徴量との間でマッチングする ORB 特徴量の数を取得する．現在のビデオフレームと直前のビデオフレームで得られた ORB 特徴量がマッチングするとき，対応する特徴点は 2 フレーム間で同じ点を表している可能性が高い．2 ビデオフレーム間で十分な数の ORB 特徴量がマッチングできたときは直前のビデオフレームから得られたカメラ姿勢と等速直線運動モデルを元にして現在のカメラ位置姿勢を推定する．一方で，2 ビデオフレーム間でマッチングする ORB 特徴量が少ないときは，直前のビデオフレームから現在のビデオフレームの間に内視鏡の姿勢が変化したものと判断する．このとき，これまでのビデオフレームから得られた ORB 特徴量の集合であるキーフレームを用いてカメラ姿勢の推定を試みる．現在のビデオフレームと対応するキーフレームが存在するかは Bag of Words (BOW) を用いてそれぞれの ORB 特徴量を BOW 特徴量に変換して現在のビデオフレームとキーフレーム間の類似度を取得する．現在のビデオフレームと類似するキーフレームが存在するときはキーフレーム内に含まれる ORB 特徴量に対して RANSAC (Random Sample Consensus) と PnP (Perspective-n-Point) を組み合わせることで反復的にカメラ位置姿勢を推定する．具体的には，キーフレームに含まれる ORB 特徴量をランダムに 4 点選出するとともに取り出した 4 点に対応するカメラ位置姿勢を PnP アルゴリズムから繰り返し推定する．このとき，推定したカメラ位置姿勢と直前に推定したカメラ位置姿勢との間の誤差が十分小さくなったとき，得られたカメラ位置姿勢を内視鏡の位置姿勢としてみなす．

内視鏡カメラ映像から取得した内視鏡の位置姿勢はトラッキング処理開始時の内視鏡位置姿勢を原点とするワールド座標系をもとに得られる．ここで，遠隔指導における座標系とワールド座標系とがずれている場合，ワー

ルド座標系に従った熟練外科医による指導が遠隔の外科医に対して意図しない内視鏡操作を指導することにつながる．ワールド座標系を遠隔指導上での座標系に変換するために内視鏡 SLAM では四元数を用いる．

四元数 q は次式のとおり定義する［5］．

$$q = q_0 + q_1\mathrm{i} + q_2\mathrm{j} + q_3\mathrm{k} \quad (1)$$

ここで，$\mathrm{i},\mathrm{j},\mathrm{k}$は虚数単位，$q_0, q_1, q_2, q_3$は実数値である．また，共役四元数$\bar{q}$は次式のとおり定義する．

$$\bar{q} = q_0 - q_1\mathrm{i} - q_2\mathrm{j} - q_3\mathrm{k} \quad (2)$$

式 (1) における実部 q_0 をスカラ部，$(q_1, q_2, q_3)^T = q_1\mathrm{i} + q_2\mathrm{j} + q_3\mathrm{k}$をベクトル部と呼ぶ．ここで，ある軸を右ねじ方向に角度Ωだけ回転する操作を考える．この回転に対応する四元数は次のとおり定まる．

$$q = \cos(\Omega/2) + I\sin(\Omega/2) \quad (3)$$

$\mathrm{I} = (I_X, Y, I_Z)^{\mathrm{T}}$は 3 軸の単位ベクトルである．ある座標系における任意の位置情報$v = (v_X, v_Y, v_Z)^{\mathrm{T}}$が得られているとき，四元数で$q$と表される回転によって移動した後のベクトルは次のとおり定まる．

$$\hat{v} = qv\bar{q} \quad (4)$$

同様に，ある座標系におけるカメラ姿勢$p = p_0 + p_1\mathrm{i} + p_2\mathrm{j} + p_3\mathrm{k}$が得られているとき，四元数で$q$と表される回転をした後の姿勢は次のとおり定まる．

$$\hat{p} = qp \quad (5)$$

提案手法はトラッキング処理を開始すると，トラッキング処理開始時点における内視鏡の位置および姿勢を原点として内視鏡の位置情報v・姿勢情報qを推定する．ここで，トラッキング処理開始時刻を$t = 0$としたとき，遠隔指導における座標系（以降，指導座標系）での内視鏡位置情報をv_{MED}，内視鏡姿勢情報をq_{MED}とする．提案手法における座標系から指導座標系に内視鏡姿勢情報を変換するための並進$v_{m\to s}$は$t = 0$のときのv_{MED}によって定義され，回転$q_{m\to s}$は$t = 0$のときのq_{MED}によって定義される．$q_{m\to s}$および$v_{m\to s}$は既知の値であるとする．各時刻tにおいて提案手法が内視鏡位置v_{SLAM}，内視鏡姿勢q_{SLAM}を推定したとき，指導座標系における内視鏡の位置情報v_{MED}および姿勢情報q_{MED}は次のとおり求められる．

$$v_{MED} = q_{m\to s}v_{SLAM}\bar{q}_{m\to s} + v_{m\to s} \quad (6)$$

$$q_{MED} = q_{m\to s} \cdot q_{SLAM} \quad (7)$$

より具体的には，指導座標系における内視鏡の位置情報および姿勢情報は，提案手法での座標系から指導座標系への軸の回転と原点の並行移動を組み合わせて推定することができる．

3. 性能評価

単眼の内視鏡カメラおよび手術映像を用いた実験評価から，提案手法による内視鏡の位置姿勢推定精度を明らかにする．

3.1. 実験環境

撮影機器: 単眼の内視鏡カメラとして USB マイクロスコープ Supereyes B005+を用いた．内視鏡カメラの解像度は 640x480 画素，フレームレートは 30fps，画角は約 34 度である．

Ground Truth とするカメラ位置姿勢: 内視鏡 SLAM による内視鏡の位置・姿勢推定精度を評価するために，Supereyes 上部に POLHEMUS 社の G4 WIRELESS MOTION TRACKER （以降，ポヒマスセンサ）を固定して正解データとなる内視鏡の位置情報・姿勢情報を取得した．ポヒマスセンサは位置姿勢を 1 秒間あたり 120 回取得する．ポヒマスセンサの座標系は立方体のソースが基準となっており，ソースは固定されている．ポヒマスセンサによる位置情報・姿勢情報の精度は二乗平均平方根で 1m あたり 0.20cm である．また，ポヒマスセンサ上の座標系において，内視鏡位置は$\mathrm{v} = (0,0,1.2)^{\mathrm{T}}$ [cm]であった．

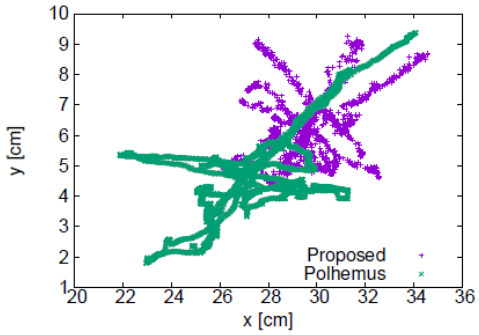
撮影環境: 手術映像中の 1 画像をラパロトレーニングバインダーに設置して Supereyes B005+で撮影するとともにその位置および姿勢を内視鏡 SLAM から推定した．図 3 に実験評価に使用した画像を示す．



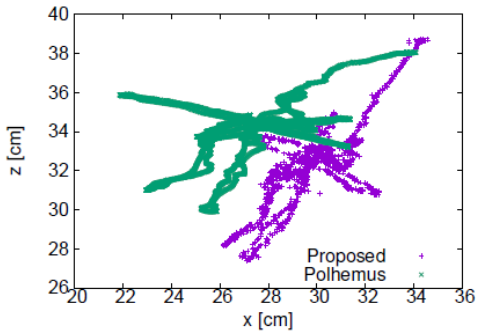
図 3. 模擬内臓環境
ラパロトレーニングバインダー

3.2. 内視鏡 SLAM による内視鏡位置・姿勢推定精度

図 4 に内視鏡 SLAM による内視鏡位置推定精度を示す．図 4(a)における横軸は x 軸方向の内視鏡位置，縦軸は y 軸方向の内視鏡位置を示す．緑色の線はポヒマスセンサから取得した正解データとなる内視鏡位置の軌跡，紫色の線は内視鏡 SLAM が取得した内視鏡位置v_{MED}の軌跡である．図 4(b)における横軸は x 軸方向の内視鏡位置，縦軸は z 軸方向の内視鏡位置を示し，各色の線に対応するデータは図 4(a)と同一である．また，図 4(c)には各軸における内視鏡 SLAM による内視鏡位置の推定誤差量を示す．なお，ポヒマスセンサおよび内視鏡 SLAM から得られるカメラ位置姿勢はサンプリング間隔が同一でないため，誤差量として aymmetric mean square error (AMSE)［8］を用いた．実験結果より，1) 内視鏡 SLAM が推定した軌跡はずれが含まれているものの，その概形は類似していること，2) 内視鏡 SLAM から得られた内視鏡位置とポヒマスセンサが取得した内視鏡位置との誤差は平均 3.65cm 程度となることが分かった．ただし内視鏡 SLAM による位置推定誤差は数 cm 程度で済む一方，遠隔指導ではわずかな誤差が誤った操作の学習につながることから数 mm 以下の推定誤差が必要であることがこれまで分かっているため，より高い推定精度を達成するアルゴリズムの設計が必要である．



(a) xy 平面上の内視鏡位置



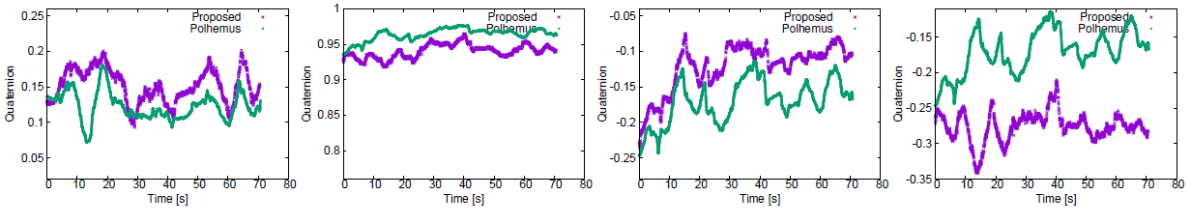
(b) xz 平面上の内視鏡位置

	x	y	z	距離
平均誤差	2.40	1.67	1.83	3.65
最大誤差	6.80	4.17	3.97	7.76

(c) 正解データとなる内視鏡位置と推定した内視鏡位置との誤差量 (単位は cm)

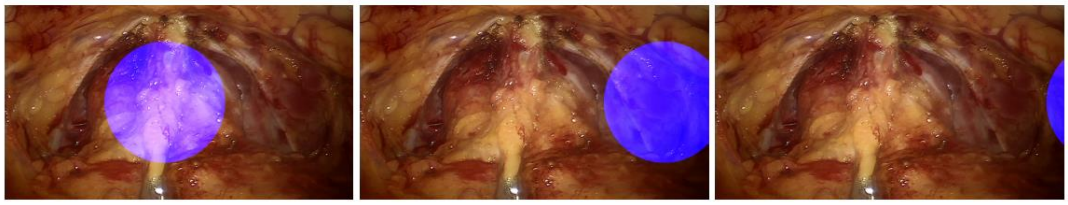
図 4. 内視鏡 SLAM による内視鏡位置情報の推定精度

同様に，図 5 に内視鏡 SLAM による内視鏡姿勢推定精度を示す．図 5(a-d)はそれぞれ内視鏡姿勢を示す四元数q_0, q_1, q_2, q_3を示す．緑色の線はポヒマスセンサから得られた正解データとなる内視鏡姿勢情報，紫色の線は内視鏡 SLAM から推定した内視鏡姿勢情報である．実験結果より，内視鏡 SLAM から推定した四元数の各値と正解データとなる四元数の各値がそれぞれ近似する値を取っていることが分かる．各時刻における正解データの四元数と内視鏡 SLAM で推定した四元数とのノルムを取得した結果，平均 0.0651，最大 0.121 となることがわかった．



(a) q_0 (b) q_1 (c) q_2 (d) q_3

図 5. 内視鏡 SLAM による内視鏡姿勢情報の推定精度



(a) 正解位置 (b) 向き情報の平均誤差量にしたがう重畳位置 (c) 向き情報の最大誤差量にしたがう重畳位置

図 6. 3D モデル重畳時における影響

3.3. 内視鏡姿勢推定精度に関する発展評価

内視鏡手術では，患者腹壁に内視鏡を通したあと，患者腹壁を支点として前後左右上下に内視鏡カメラを操作することで執刀医は患者の体内を観察する．これは内視鏡 SLAM から推定した内視鏡姿勢情報が遠隔指導時においてより大きな影響をもたらすことを示唆している．本節では内視鏡視野上に事前取得した 3D モデルを重畳したとき，内視鏡 SLAM によるカメラ姿勢推定誤差が重畳する 3D モデルにもたらす影響について評価した．

重畳する 3D モデルへの影響を評価するために，内視鏡 SLAM が推定した内視鏡の向きと実際の内視鏡の向きとの誤差δを定義する．ここで，内視鏡は正面方向(z 方向)を向いている．この向きは，四元数のベクトル部$q = (0,0,1)^T$に相当する．時刻tにおいてポヒマスセンサが取得した内視鏡姿勢をq_{GT}，内視鏡 SLAM が推定した内視鏡姿勢をq_{MED}とすると，q_zに対する内視鏡の向きはそれぞれ次式から得られる．

$$o_{MED} = q_{MED} q_z, \quad o_{GT} = q_{GT} q_z \tag{8}$$

ここで，正解データとなる内視鏡の向き情報o_{GT}と内視鏡 SLAM が推定した内視鏡の向き情報o_{MED}との誤差δはベクトルの内積に対する逆余弦関数から次式のとおり得られる．

$$\delta = \cos^{-1}(o_{MED} \cdot o_{GT}) \tag{9}$$

図 5 に示したポヒマスセンサおよび内視鏡 SLAM から得られたカメラ姿勢情報に対して上述の式を適用した結果，向き情報の平均誤差は 13.98 度，最大誤差は 21.79 度となることが分かった．

最後に，得られた向き情報の誤差量をもとに，手術映像上に 3D モデルに相当するオブジェクトを重畳表示したときの画像を示して姿勢情報の誤差が執刀医の視野にもたらす影響を評価した．実際の内視鏡手術における内視鏡と病変との間の距離は約 1〜10cm である．本評価では，内視鏡カメラが最も離れている a = 10cm を想定した．なお，向き情報の誤差がδであるとき，視野中心は$a\tan(\delta)$だけずれることとなる．また，簡単化のため，視野中心がずれる方向は右方向とした．

図 6 に正解となる位置に 3D モデルを重畳したとき，向き情報の平均誤差量だけずれた位置に 3D モデルを重畳したとき，向き情報の最大誤差量だけずれた位置に 3D モデルを重畳したときの画像を示す．なお，本評価では対象とした画像の中心部にある球状の器官に対応する 3D モデルを重畳することを想定

した．図 6(a)と図 6(b)を見比べると 3D モデルの重畳位置に無視できない程度の誤差が生じていることがわかる．また，図 6(c)においては 3D モデルが画面からほとんど見えなくなるほど離れた位置に重畳する結果となった．執刀医の視野上に3Dモデルを重畳する場合，内視鏡 SLAM によるカメラ姿勢推定精度は未だ不十分であることが推察できる．

内視鏡 SLAM による姿勢推定精度が低下した原因として体内を対象とした画像情報の特徴が考えられる．文献［9］によれば，体内を対象とした画像情報はテクスチャが均一であるため，各画像から特徴量・特徴点を抽出するのが困難であることが示されている．本実験で用いた模擬内臓環境においても同様の傾向が見られた．より具体的には，体内を対象とした画像情報からロバストな特徴点を取得できず，内視鏡 SLAM による内視鏡の位置・姿勢推定が困難になるとともに推定精度が低下するケースが見られた．

4. おわりに

本稿では遠隔指導での使用を想定して内視鏡 SLAM を用いた内視鏡カメラの位置姿勢推定手法を提案した．実験結果から，内視鏡 SLAM は数 cm 程度の誤差を達成しうるものの，最大誤差として医療現場で利用するには無視できない誤差が発生することが示された．今後の課題として，内視鏡映像に対応したアルゴリズムの再設計を通して内視鏡位置・姿勢の推定精度を改善すること，取得した点群情報から他の立体情報を外科医へのフィードバックする手法の検討が挙げられる．

謝辞

本研究は JSPS 科研費 (JP19H01101, JP20K19783) の支援の下で行った．

利益相反の有無

なし

文　献

［1］ 日本内視鏡外科学会:領域別の内視鏡手術総症例数の推移について. 日本内視鏡外科学会雑誌: 728-729, 2018

［2］ 山下実，植村宗則，松岡紀之 他：内視鏡外科手術における縫合技術評価法の有用性に関する研究. 日本コンピュータ外科学会誌: 15-22, 2015

［3］ 厚生労働省 医師需給分科会：医師の需給推定について．https://www.mhlw.go.jp/file/05-Shingikai-10801000-Iseikyoku-Soumuka/0000203368.pdf

［4］ 徳安達士，元土肥久美子 : 鉗子の操作感覚を体得する内視鏡外科手術訓練装置の開発. 福岡工業大学情報学研究所所報: 1-4, 2017

［5］ 金谷健一：3 次元回転 -パラメータ計算とリー代数による最適化- ．共立出版，東京，2019, pp27-30

［6］ Taketomi T, Uchiyama H, Ikeda S: Visual SLAM algorithms: a survey from 2010 to 2016, IPSJ Transactions on Computer Vision and Applications, 2017

［7］ Mur-Artal R, Tardos, J D: Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras. IEEE Transactions on Robotics: 1255-1262, 2017

［8］ Fujihashi T, Koike-Akino T, Watanabe T, et al.: HoloCast: Graph Signal Processing for Graceful Point Cloud Delivery. ICC 2019-2019 IEEE International Conference on Communications: 1-7, 2019

［9］ Qiu L, Ren H: Endoscope navigation and 3D reconstruction of oral cavity by visual SLAM with mitigated data scarcity. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops: 2197-2204, 2018

Preliminary Discussion on Endoscope Pose Estimation Using Visual SLAM

Soushi UENO[*1], Takuya FUJIHASHI[*1], Tadahiko KIKUKAWA[*2], Shunsuke SARUWATARI[*1], Hideyuki ANDO[*3], Takashi SAIKA[*2], Takashi WATANABE[*1]

*1 Graduate School of Information Science and Technology, Osaka University

*2 Graduate School of Medicine, Ehime University

*3 Osaka University of Arts

Remote lecture of endoscopic surgery watching surgical video alleviates physical constraints between experienced the surgeon and the unskilled surgeon. On the other hand, it is difficult for experienced surgeons to properly instruct the detailed operation techniques of endoscopes and forceps that are necessary for endoscopic surgery based only on video data. In order to realize smooth communication between surgeons in remote lecture for endoscopic surgery, it is desirable to share stereoscopic information in the surgical field reconstructed based on 2D images obtained from the endoscope. In this paper, we propose a method for estimating the position of an endoscope by using Visual SLAM (Simultaneous Localization And Mapping), which calculates a camera position and generates an environment map using camera images as input. In this method, the feature points included in the image photographed by the endoscope are extracted, and the position and attitude of the endoscope are calculated in real time. From the experimental evaluation using monocular endoscope camera, it was confirmed that the estimation accuracy of endoscope position by the proposed method was within the error of about 3.6 cm in average.

Key words: Remote lecture, Endoscope Pose Estimation, Visual SLAM

Pre-training of the BERT model using a large-scale radiology report database for estimation of purpose of CT scanning

Shuhei HONDA[*1], Yoshito OTAKE[*1,2], Masaki TAKAO[*3], Eiji ARAMAKI[*1], Shuntaro YADA[*1], Kento Aida[*2], Shinichi SATO[*2], Masahiro HASHIMOTO[*4], Toshiaki AKASHI[*5], Nobuhiko SUGANO[*3], Yoshinobu SATO[*1]

Abstract

Our group has built a large-scale database consisting of CT images and radiology reports associated with each CT scan. In order to obtain knowledge related to anatomical variation from this database, we developed an automatic measurement system for musculoskeletal anatomical parameters using CNN and have analyzed changes in anatomical parameters such as pelvic tilt angle and muscle volume with age and gender. In this study, we attempted to estimate the objective of CT imaging (i.e., disease background) by using the radiological reading reports associated with each CT image to augment the medical knowledge obtained from the analysis results. The BERT model used in this study was pre-trained with a corpus of approximately 320,000 radiological reading reports and compared with the widely used Wikipedia corpus-based pre-training model. We created a criteria of classification of imaging purpose consisting of twelve classes and 299 reports were manually classified by an expert, which were used for fine-tuning of the pre-trained BERT model. The radiology report pre-trained BERT model improved the prediction accuracy by 18.1% over the conventional method using LSTM and by 7.7% over the Wikipedia corpus-based pre-training model. Finally, attention map was visualized to understand the rationale for the classification of the model, which demonstrated a reasonable agreement with expert's judgement for correctly classified reports.

Keywords：BERT, NLP, radiology report, medical image

1. Purpose

We have been analyzing parameters related to musculoskeletal anatomy such as the variation in shape and density of bones and muscles in a large-scale medical image database (>10K patients) using CNN to acquire new knowledge for understanding of posture change with aging or for guidance in orthopedic surgical planning. However, the anonymized large-scale database lacks detailed attribute of each patient, which limited the interpretation of our analysis results. In this study, we extracted information related to the patient's disease background from a radiological report associated with each CT image using a deep neural network (DNN) -based natural language processing (NLP).

*1 Nara Institute and Science and Technology
〔8916-5 Takayama, Ikoma, Nara 630-0192, Japan〕
e-mail: {honda.shuhei.hq3, otake, yoshi}@is.nasit.jp
*2 National Institute of Informatics
〔2-1-1 Hitotubashi, Chiyoda, Tokyo 101-8430, Japan〕
*3 Osaka University
〔2-2 Yamadaoka, Suita, Osaka 565-0871, Japan〕
*4 Keio University
〔35 Shinanomachi, Shinjuku, Tokyo 160-8582, Japan〕
*5 Juntendo University
〔2-1-1 Hongo, Bunkyo, Tokyo 113-8421, Japan〕

2. Method

One of the state-of-the-art NLP models called bidirectional encoder representations from transformers (BERT) [1] pre-trained in radiology reports collected at collaborating institutions was applied to the radiology report written in Japanese associated with each CT image in order to extract information about the purpose of the scan, which would strongly correlate with the patient attribute. After several preliminary trials, we established a criteria of classification of imaging purpose consisting of twelve classes, cancer, lower limb joint disease, respiratory disease, trauma, aortic disease, heart disease, digestive system disease, liver complaint, spinal disease, renal urological disease, non-trauma exact examination, other. The pre-training of BERT involved byte-pair encoding of about 320,000 radiology reports into 32217 tokens with sentencepiece and 30 epochs of training over about 30 days. Fine-tuning and evaluation of BERT was performed using 299 reports manually classified into 12 disease classes by an expert orthopedic surgeon. 5-fold cross validation was performed.

As a comparison with previously proposed methods, we implemented the following two algorithms: 1) A simple model composed of embedding and unidirectional long short-term memory (LSTM), 2) pre-trained BERT model from Japanese Wikipedia [2].

Finally, we visualized the Attention Map of the fine-tuned model to understand the rationale for the classification of the model, and checked whether it was consistent with the expert's rationale

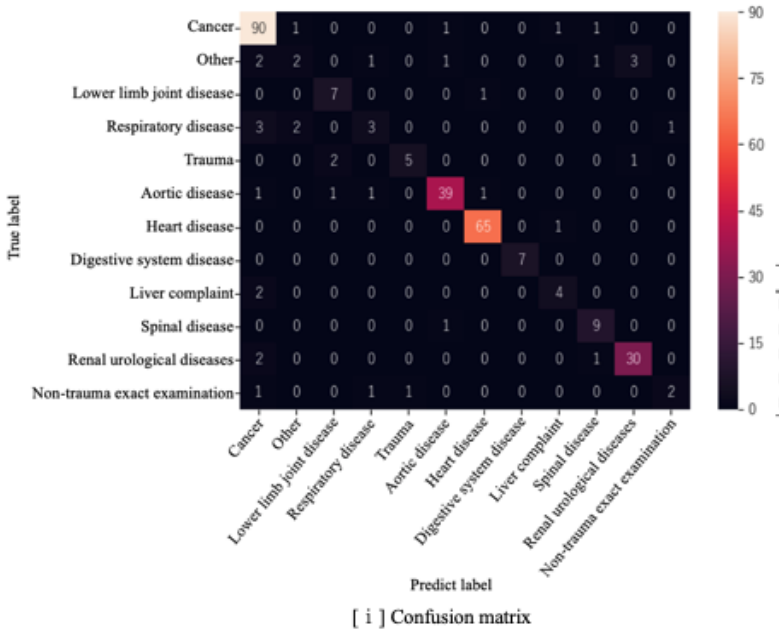
3. Results

Fig. 1 showed the confusion matrix of the results predicted by the proposed BERT-large model and the results of comparative study. The accuracy was 88.0%. The confusion matrix showed low accuracy for diseases with the total number of samples less than 10, namely: non-trauma exact examination (N=5), 40.0 %, respiratory disease (N=9), 33.3 %. Increase of the samples for fine-tuning would help reduce these misclassifications and improve the accuracy.

The proposed model showed the best performance compared with other experimental results, suggesting that BERT was generalizable to the radiological report analysis effectively using the pre-training by a large corpus from radiology reports.

Table 1 showed which words (tokens) the model payed attention to relative to each other when classifying the radiology reports. The cases shown in Table 1 are selected based on the criteria of: 1) successful with radiological-report-pre-training and failed with Wikipedia-pre-training, and 2) closest to median softmax probability. The English translation was performed using DeepL [§ 1]. First, we discuss our model (radiology reports). In the representative example of classifying the report as cancer, the focus was on "Post left lung surgery. Post left mammary surgery. No obvious metastases are evident", while in the example of classifying as heart disease, it was evident that the focus was on the "The bilateral coronary inlets are preserved and, within the scope of assessment, no specific postoperative findings of concern can be noted around the aortic valve rings". Based on the observation of these representative cases, the expert surgeon confirmed that BERT payed more attention to certain words that doctors generally pay attention to in the classification task. Next, we discuss the model as it is pre-trained in Japanese wikipedia. The Japanese Wikipedia pre-trained model is not able to grasp the sentence structure of the findings because it only focuses on the word level. Therefore, it is assumed that the table example

was misclassified due to a lack of judgment. We think there are two reasons for this difference between the two models. First, the names of the organs that could be important keywords are unknown in Japanese Wikipedia pre-trained model, which reduced the classification accuracy. Second, because the data for fine-tuning in this study was small (299 cases), it is more advantageous to acquire knowledge from the same domain pre-training dataset, resulted in higher generalizability.



[i] Confusion matrix

[ii] Evaluation results

Method	Accuracy
LSTM	69.9
BERT-base (Japanese Wikipedia)	78.7
BERT-large (Japanese Wikipedia)	80.3
BERT-base (Radiology reports) [Ours]	87.3
BERT-large(Radiology reports) [Ours]	**88.0**

Fig. 1 Confusion matrix and evaluation result of the tested model configurations. Results of 5-fold cross validation of the proposed model using 299 radiological reports manually classified into 12 disease classes by an expert surgeon.

Table 1 Visualization of attention map, representing the relative importance of the model for classifying reports. The more attention-grabbing words are highlighted in red[§1]. The [UNK] in the table is unknown word because it did not exist in the dictionary at the time of pre-training.

true label / predict label	corpus for pre-training	finding
cancer / cancer	radiology reports	前回 CT(2017/07/25)と比較しました。左肺術後。左乳腺術後。明らかな転移は認めません。肝 S8 に 1cm 大と 2cm 大の腫瘤を認めますが、前回と著変ありません。血管腫等の良性病変が疑われます。肝嚢胞 Compared to the previous CT (07/25/2017). Post left lung surgery. Post left mammary surgery. No obvious metastases are evident. A 1 cm and 2 cm large mass in liver S8 is observed, which is not markedly different from the previous one. Benign lesions such as hemangioma are suspected. Liver cyst
cancer / other	Japanese Wikipedia	前回CT[UNK][UNK][UNK][UNK][UNK][UNK][UNK]と比較しました。左##肺術##後。左乳##腺術##後。明らかな転移は認めません。肝[UNK][UNK]に[UNK][UNK]大と[UNK][UNK]大の[UNK]を認め.1 ますが、前回と著変ありません。血管##腫等の良##性病変が疑われます。肝嚢##胞 Compared to the previous CT [UNK] [UNK] [UNK] [UNK] [UNK] [UNK] [UNK]. Post left lung surgery. Post left mammary surgery. No obvious metastases are evident. A [UNK] [UNK] and [UNK] [UNK] large [UNK] in liver S8 is observed, which is not markedly different from the previous one. Benign lesions such as hemangioma are suspected. Liver cyst
heart disease / heart disease	radiology reports	両側冠動脈入口部は保たれており、評価可能範囲内で大動脈弁輪周囲に術後として特に問題となる所見を指摘できません。以下、著変なし。左腎動脈狭窄(+)腹腔動脈に口径不整(+)甲状腺両葉に低吸収域(+)変性側弯(+) The bilateral coronary inlets are preserved and, within the scope of assessment, no specific postoperative findings of concern can be noted around the aortic valve rings. No significant changes are noted below. Left renal artery stenosis (+) left renal artery stenosis (+) abdominal artery caliber irregularity (+) hypoabsorptive area in both thyroid lobes (+) degenerative scoliosis (+)
heart disease / aortic disease	Japanese Wikipedia	両側冠動脈入口部は保たれており、評価可能範囲内で大動脈弁輪周囲に術##後として特に問題となる所見を指摘できません。以下、著変なし。.1 左腎##動##脈[UNK][UNK] [UNK] [UNK] [UNK] 動脈に口径不##整[UNK] [UNK] [UNK]甲状腺両葉に低吸収域[UNK] [UNK] [UNK]変##性[UNK] [UNK] [UNK] The bilateral coronary inlets are preserved and, within the scope of assessment, no specific postoperative findings of concern can be noted around the aortic valve rings. No significant changes are noted below. Left renal artery stenosis [UNK] [UNK] [UNK] left renal artery stenosis [UNK] [UNK] [UNK] [UNK] artery caliber irregularity [UNK] [UNK] [UNK] hypoabsorptive area in both thyroid lobes [UNK] [UNK] [UNK] degenerative scoliosis [UNK] [UNK] [UNK]

§1 ALL reports used in this study are written in Japanese. The findings written in English in table 1 have been translated by DeepL (https://www.deepl.com/translator).

4. Conclusion

A DNN model for analyzing radiological report associated with CT images was proposed in order to extract information related to the purpose of the CT scanning which is, to some extent, correlated with patient's disease background. To the best of our knowledge, this is the first study that constructed a pre-trained DNN model using a large corpus (about 320,000) of radiology reports written in Japanese. We predicted the purpose of CT scanning with the accuracy higher than the Japanese Wikipedia pre-trained model, which is widely used in general purpose NLP but lacks medical knowledge. We expect the models pre-trained in this study to be used in a variety of applications especially for organizing large scale medical image database associated with radiology reports.

One of our future work includes an active learning framework to automatically increase the training data to improve classification accuracy. We would also like to explore new research areas integrating medical image analysis and NLP.

Competing interests

The authors indicated no conflicts of interest.

Acknowledgement

This study was supported by Japan Agency for Medical Research and Development (AMED) under Grant Number JP20lk1010036. The radiological report and CT image database used in this study was a part of Japan Medical Image Database (J-MID) constructed by Japan Radiology Society (JRS).

References

[1] J. Devlin, M. W. Chang, K. Lee, K. Toutanova (2018). "BERT: Pre-training of deep bidirectional transformers for language understanding." arXiv, preprint, arXiv:1810.04805.

[2] T. Shibata, D. Kawahara, S. Kurohashi (2019). "Improvement of accuracy of Japanese parsing by BERT." The Twenty-fifth Annual Meeting of the Association for Natural Language Processing, pp.205-208 [published in Japanese].

Verification of the Optimal Portal Search Simulator using Phantom

Shinnosuke KAWAKAMI[*1], Norio FUKUDA[*2], Takashi NISHII[*3], Yoshito OTAKE[*4], and Yoshinobu SATO[*4]

Abstract

Arthroscopic labral repair difficulty depends on the field of view by the arthroscope inserted through a few millimeters skin incision called a portal. The field of view of the arthroscope is different depending on the position of the portal and the insertion angle therefrom. A typical portal position is geometrically determined using anatomical landmarks of a hip. However, it is not necessarily appropriate portal that the insertion of the arthroscope from that position will give the best view. Therefore, we decided a typical portal using anatomical landmarks, then shifted its position on the simulator, and greedy inserted a virtual hip arthroscope from it, and from the obtained image of the hip lip We devised a method to check the visibility. We performed automated visibility measurement using 3D surface models of a phantom model on the hip arthroscope simulator. As a result, we estimated an optimal portal position of a Phantom. We report the results of verification by the hip arthroscope inserted from the optimal portal that the simulator calculated, by inserting the hip arthroscope into the phantom model while positioning with Polaris.

Keywords：preoperative planning, hip arthroscopy, medical image

1. Purpose

Manuscript should be written based on this format if MS-WORD is used. Departmental and institutional affiliations of all authors should be written in the left-bottom of the first page. It is convenient to use “text box” as this template does.

2. Method

A 3D surface model of the hip joint and hip lips is created with a scanner from the phantom model. The

*1 Kindai University Technical College
〔7-1 Kasugaoka, Nabari, Mie, Japan〕
e-mail: kawakami@ktc.ac.jp
*2 Center for Information and Neural Networks
〔1-4 Yamadaoka, Suita City, Osaka, Japan〕
*3 Department of Orthopaedic Surgery, Osaka General Medical Center
〔3-1-56 Mandaihigashi, Sumiyoshi-ku Osaka-shi, Osaka, Japan〕
*4 Graduate School of Information Science, Nara Institute of science and technology
〔8916-5 Takayama, Ikoma, Nara, Japan〕

3D model is read on the simulator, the virtual hip arthroscope is automatically and comprehensively inserted from the portal position shifted from the typical portal position which is the portal position defined by the textbook is defined using anatomical landmarks[1][2], and the most visible portal position for labral tear is calculated using the index value we propose. By reproducing the portal position calculated by the simulator on the phantom model and inserting an actual hip arthroscope from it to verify the visibility, it is verified that the results calculated by the simulator are consistent. In order to verify the visibility of the hip joint with a simulator, only when the proportion of the femoral head reflection in the simulator image showing the hip joint is 5%, 10%, 15%, 20%, 25%.

Arthroscopic images of labral tear are obtained by the virtual camera placed in the tip of the virtual arthroscope. Firstly, the femoral head is pulled 10 or 15 millimeters toward inferior and rotate 5 degree around inferior-superior axis toward anterior to ensure the workspace of the joint cavity based on an advice by the expert surgeon. A typical portal position is determined using landmarks on the 3D models. Inserted arthroscope can swing at the portal as a fulcrum. Finally, the arthroscope is inserted with various insertion direction at shifted portal positions by the following procedure.

1) to color the surface of labral tear at angular intervals of 30 degree which is corresponding to 0-12 o'clock
2) establish several positions which are shifted 5 millimeters from typical portal position toward inferior or superior and anterior or posterior direction which are numbered.
3) arthroscope direction is set to the target points within the space between articular cartilage and femoral head sphere, then the insertion angle from anterior axis and from superior axis determined.
4) arthroscope is inserted through the shifted portal positions to target point with rendering and rotating around the line of sight direction as a rotating axis at angular intervals of 45 degree till reaching the target or colliding with femoral head or articular cartilage
5) try next target point
6) arthroscope insertion to all target is finished, portal position is shifted from inferior to superior or anterior to posterior
7) iterate on 3) to 6) till the end of rendering at final portal position

3. Results

We use the 3D surface model of the phantom model loaded by the simulator and insert the virtual hip arthroscope comprehensively from the portal position shifted from the typical portal position, and the optimal portal that gives the best visibility. I was able to calculate the position.

4. Conclusion

At present, an experiment is underway to verify the simulator results by inserting an actual hip arthroscope into the phantom for this verification.

Competing interests

The authors declare no conflicts of interest associated with this manuscript.

Acknowledgement

References

[1] J.W. Thomas Byrd: Surgical Technique: Supine Position. Operative Techniques in Orthopaedics Vol.15, issue 3: 204-217, 2005

[2] William J. Robertson, M.D., and Bryan T. Kelly, M.D.: The Safe Zone for Hip Arthroscopy: A Cadaveric Assessment of Central, Peripheral, and Lateral Compartment Portal Placement. The Journal of Arthroscopic and Related Surgery, Vol.24, No9: 1019-1026, 2008

Bayesian U-Net を用いた造影 CT からの腹部臓器の自動抽出における不確実性の有用性の検証

政木　勇人[*1,2]　大竹　義人[*2]　スーフィー　マーゼン[*2]　堀　雅敏[*3]

大西　裕満[*4]　富山　憲幸[*4]　佐藤　嘉伸[*2]

要旨

近年の深層学習を用いた CT 画像からの腹部臓器の自動抽出手法は精度の向上と計算時間の短縮を実現し臨床応用を現実的にしたが，手動トレースなしの状態で抽出結果の精度を予め見積もることは難しかった．我々のグループでは，下肢筋骨格自動抽出において，予測の不確実性を推定する Bayesian U-Net を用いて CT 画像からの不確実性と抽出精度の関係を調査し，不確実性から抽出精度を予測できる可能性を示唆した．本研究では，CT 画像からの腹部臓器自動抽出における Bayesian U-Net から得られる不確実性の有用性の検証を目的とする．本稿では，49 症例(画像セット数 97)の腹部造影 CT 画像と 13 臓器のラベル付きデータセットを用いて Bayesian U-Net から得られる不確実性と抽出精度の間の関係を調査し抽出精度の予測精度を調査したので，その結果について報告する．

キーワード：ベイジアン深層学習，畳み込みニューラルネットワーク，造影 CT 画像，腹部臓器自動抽出，能動学習

1. はじめに

CT 画像からの腹部臓器領域の自動抽出は，コンピュータによる診断支援システムを構築するための基盤技術である．近年，CNN (Convolutional Neural Network) などの深層学習を用いた手法[1-3]が精度の向上と計算時間の短縮を実現し成果を挙げている．しかし，抽出結果の信頼性の予測は行われていない．

Gal ら[4]は MC dropout(Monte Carlo dropout)を深層学習のネットワーク構造に組み込むことで予測結果の不確実性を推定できることを示した．この手法を用いて，我々は U-Net に MC dropout を統合した Bayesian U-Net[5]を構築し，CT 画像からの下肢筋骨格自動抽出での抽出精度と不確実性の関係を調査し，ピクセル単位での不確実性に加え，臓器全体の抽出精度(Dice 係数)と高い相関を示す不確実性の指標を提案した．腹部臓器の自動抽出においても不確実性と抽出精度の関係を調査することは重要であるが，そうした研究は我々の知る限り存在しない．また，[5]では学習データが 20 症例のみ，と少なかったため不確実性指標と実際の Dice 係数の線形回帰により求めた予測 Dice 係数の信頼性については更なる検証が必要と考えられた．そこで本研究では，腹部臓器自動抽出において Bayesian U-Net から得られる不確実性の有用性の検証を目的とし，比較的大規模な(97 画像セット)ラベル付き 2 時相造影 CT 画像データセットを用いて不確実性と抽出精度の関係を調べ，不確実性を用いた抽出精度予測実験を行った．

*1 PSP 株式会社
*2 奈良先端科学技術大学院大学
〔〒630-0192 生駒市高山町 8916-5〕
e-mail: masaki.yuto.mq7@is.naist.jp
*3 神戸大学大学院医学研究科
*4 大阪大学大学院医学系研究科
投稿受付：2020 年 8 月 5 日

2. 腹部臓器の自動抽出

本研究では,腹部臓器の自動抽出には Hiasa ら[5]が提案した 2 次元画像を入力・出力とする 2D Bayesian U-Net を用いる．Bayesian U-Net は，医用画像処理に広く用いられている U-Net[6]に，Gal らが提案した MC dropout [4]による推論を統合することで，抽出結果に加えその予測の不確実性を出力する．MC dropout で得られる不確実性は，epistemic uncertainty と呼ばれ，学習データが十分に存在すると低下するとされるモデルの不確実性である [7]．Bayesian U-Net は，推論時に MC dropout を適用し,推論を複数回行って得られる MC sample (Monte Carlo sample)の予測分散(Predictive variance:PV)によって画像のピクセル単位の不確実性を算出する．抽出対象の各臓器における不確実性指標には，Hiasa らが提案した PSV(Predictive structure wise variance)[5]と呼ばれる予測した各臓器領域内の PV の平均値を用いる．推論時は，CT 画像をスライスごとに処理し 3 次元に再構成し予測ラベル及び不確実性画像を得る．

3. データセット

本研究では,上腹部が撮影された 49 症例の早期動脈相及び後期動脈相の 2 時相造影 CT 画像のラベル付きデータセットを用いた．データセットの内訳を表 1 に示す．代表的なデータセットの例を図 1 に示す． 1 症例のみ後期動脈相のみであったため，CT 画像セットの数は 97 であった．マトリックスサイズは,すべての CT 画像で 512×512 であった．撮影装置は,スライス厚 2.5mm の CT 画像は 4-slice Light Speed QX/i (GE Healthcare)，スライス厚 0.625mm の CT 画像は 64-slice Light Speed VCT (GE Healthcare)であった．

対象臓器は，肝臓，脾臓，膵臓，左腎，右腎，胆のう，大動脈，下大静脈，心臓，胃，門脈，左肺野，右肺野の 13 臓器であり，手動トレースによってラベル画像を作成した．

表 1 データセットの内訳

時相	CT 画像の数	スライス厚 [mm]
早期動脈相	10	2.5
	38	0.625
後期動脈相	10	2.5
	39	0.625

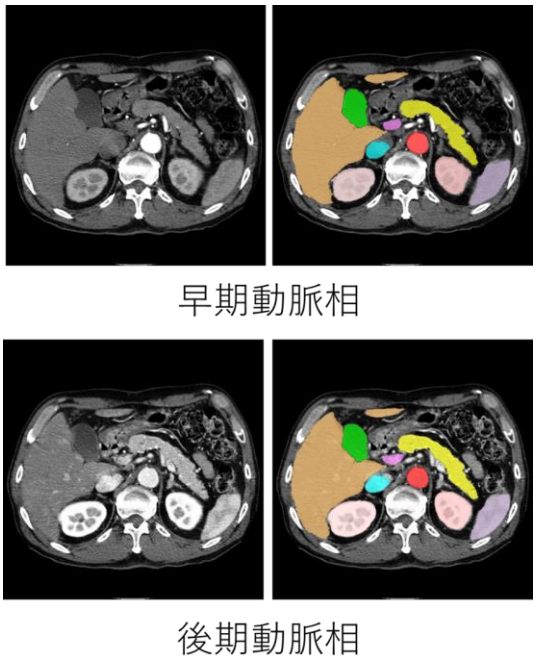


図 1 データセットの代表的な例

4. 実験

本研究では，Bayesian U-Net による腹部臓器の自動抽出性能を評価し，不確実性と抽出精度の関係に基づき不確実性から抽出精度を予測する実験を行った．ラベル付きデータセットを用いた 3-fold 交差検証により行った．

1） 評価指標

自動抽出の精度の評価指標として，Dice 係数を使用した．

不確実性の指標として，PV と PSV を使用した．ピクセル単位で PV を求め，空間分布(PV マップ)を可視化し，誤差の空間分布(誤差マップ)と視覚的に比較し妥当性を評価した．臓器ごとの PSV と Dice の関係を求める手法として線形回帰を用いた．線形回帰モデルの学習及び推論は，各 fold のテストデータ内で Leave-one-out 交差検証することによっ

て行った．実際の Dice 係数と予測した Dice 係数の絶対値誤差で評価した．

2） 実験条件

512×512 のアキシャル断面の CT スライスを入力画像とし，前処理として[-150,+350]HU の信号値を[0, +1]に線形変換した．Bayesian U-Net の構造は，Encoder 部の convolution 層数は 4 としフィルタ数は 32 とした．Dropout 層は，各 Max pooling 層及び Up convolution 層の後に追加した．出力層の活性化関数は Softmax を用い，重みの初期化は He らの手法[8]を用いた．損失関数は，Softmax cross entropy 損失を用いた．バッチサイズを 3 とし，学習係数を 1.0e-4 とした Adam[9]を用い，反復回数を 1.0e+5 として最適化を行った．データ水増しは，移動量(-25%～+25%)の平行移動，せん断角(-π/8～+π/8)，回転角(-10°～+10°)の回転を行った．推論における MC sample の数は 15 とし，その平均を計算して最終的な抽出結果を得た．なお，本稿では体領域の抽出といった前処理や最大連結成分の抽出といった後処理を行わず，Bayesian U-Net の出力結果をそのまま評価に使用した．また，学習データセットにおけるトレーニングデータセットと検証データセットに分離する割合は，0.9:0.1 としランダムに行った．

テストデータに近い画像が学習データに混入してしまうのを防ぐため，同一患者の複数時相データが複数のグループにまたがらないようにグループ分けした．

5. 結果

図 1 に各臓器の Dice 係数値の箱ひげ図を示す．肝臓，脾臓，左腎，右腎，大動脈，心臓，胃，左肺野，右肺野の Dice 係数が 0.9 以上となった．また，各臓器における Dice 係数と PSV の関係を図 2 に示す．相関係数は，大動脈が一番小さかったがすべての臓器で負の相関がみられた．平均 Dice 係数が高かった肝臓と平均 Dice 係数が低かった門脈について，PSV から算出した予測 Dice 係数と実際の Dice 係数の関係を図 4 に示す．全臓器における予測 Dice 係数と実際の Dice 係数の絶対値誤差の分布を図 5 の箱ひげ図に示す．肝臓の予測 Dice 係数の絶対値誤差が一番小さく，その値は 0.005 であった．予測 Dice 係数の絶対値誤差が一番大きかったのは門脈であり，その値は 0.071 であった．

データセット内で抽出精度の合計が高かった例，中間付近の例，低かった例をそれぞれ図 6，図 7，図 8 に示す．エラーマップの値が大きい領域に PV マップの大きい値が存在する傾向を確認した．

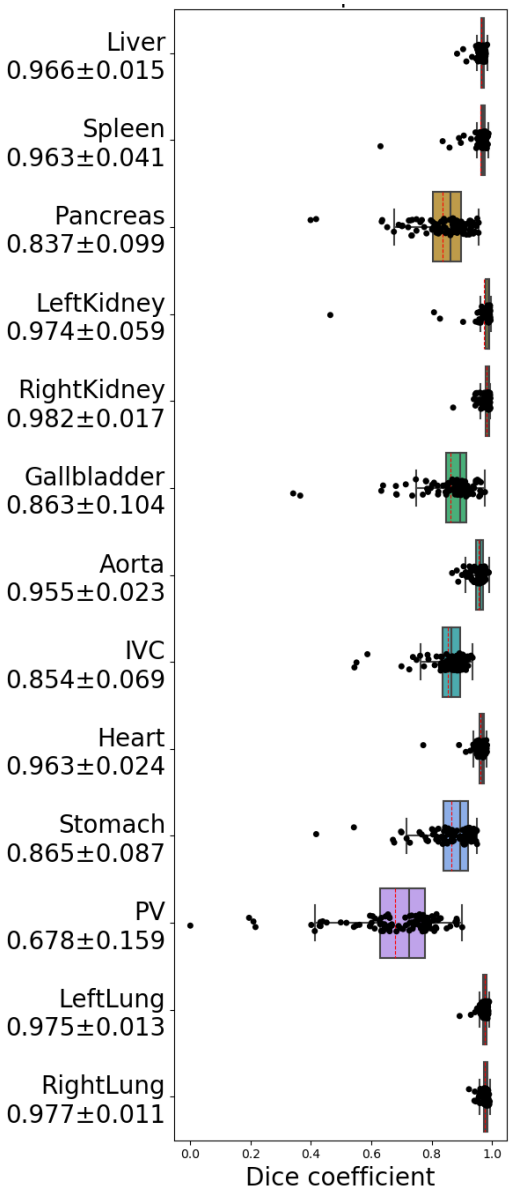


図 2 各臓器の Dice 係数

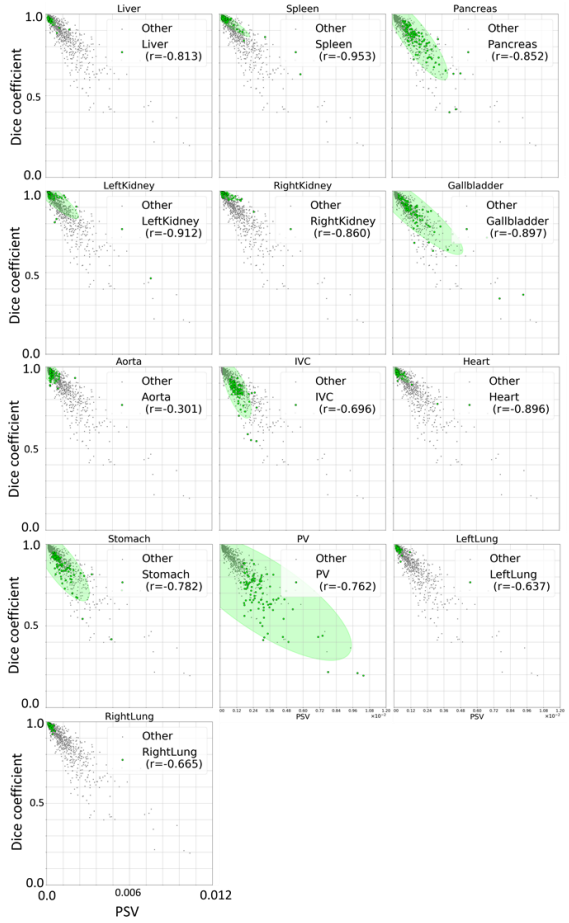


図 3 各臓器における PSV と Dice 係数の関係
楕円は，95%信頼楕円を示す.

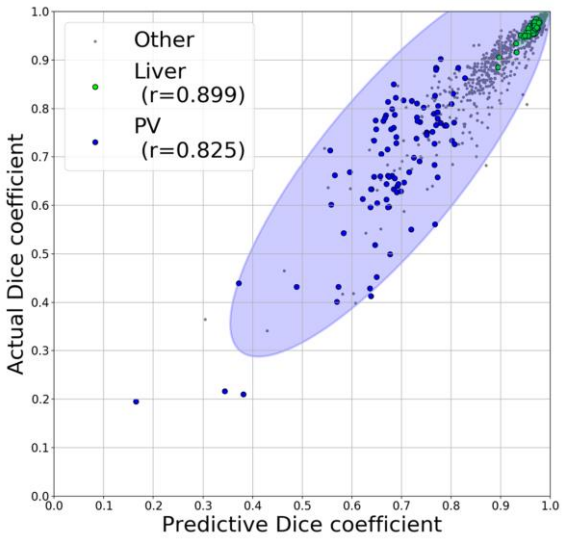


図 4 肝臓及び門脈の実際の Dice 係数と PSV より算出した予測 Dice 係数の関係
楕円は，95%信頼楕円を示す.

6. 考察

図 2 から，Bayesian U-Net により，肝臓，脾臓，左腎，右腎，大動脈，心臓，左肺野，右肺野は，平均 Dice 係数 0.90 を超え，良好な抽出結果が得られた．しかし，膵臓，胆のう，下大静脈，胃，門脈は，領域が比較的小さくかつコントラストなどのバリエーションも大きいため，精度が低下したと考えられる．



図 5 実際の Dice 係数と予測 Dice 係数の絶対値誤差

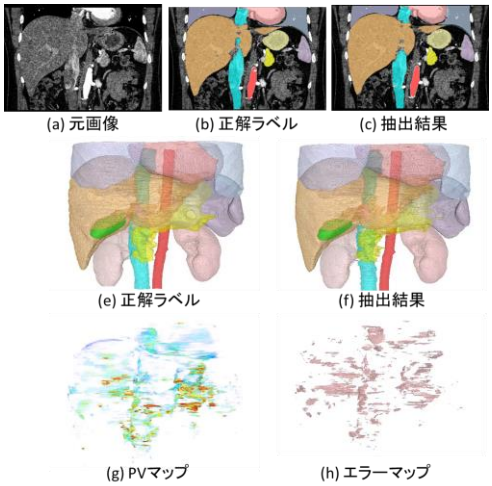


図 6 Dice 係数の合計が高かった例

PV マップ及びエラーマップの可視化の際には，どの臓器においても辺縁に必ず生ずるランダムな誤差の影響を除くため，正解臓器領域に対してカーネルサイズ 3 の膨張処理と伸縮処理の差分で生成した臓器辺縁領域マスクを用いて，臓器辺縁領域の輝度値を 0 で埋める処理を施した．

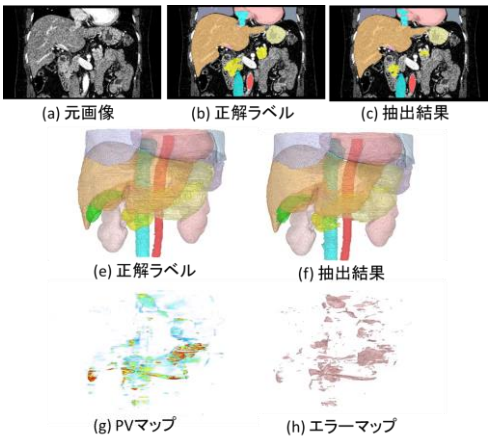


図 7 Dice 係数の合計が中間付近の例

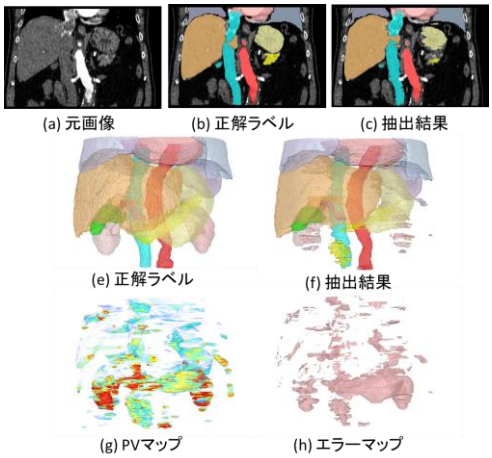


図 8 Dice 係数の合計が低かった例

Dice 係数と PSV の関係は，図 3 より Hiasa ら[5]が筋骨格自動抽出で報告しているのと同様に，負の強い相関関係を確認した．PSV の増加による Dice 係数の減少の傾向が 95%信頼楕円によって示され，PSV から実際の Dice 係数を見積もることができる可能性を示した．

実際に PSV から Dice 係数の予測値を計算し散布図として示した図 4 より，平均 Dice 係数が高い肝臓よりも平均 Dice 係数が低い門脈の方が，95%信頼楕円が大きく分布がばらついていた．しかし，平均 Dice 係数が低い臓器であっても，Dice 係数の予測値と実際の値の間には相関関係が見られ，抽出精度の傾向はある程度捉えられると考えられる．

図 5 より，各臓器の実際の Dice 係数と予測された Dice 係数の間の平均絶対誤差において，実際に得られた平均 Dice 係数が 0.9 以上の臓器の肝臓，脾臓，左腎，右腎，大動脈，心臓，左肺野，右肺野は，0.02 以下であり良好な予測が可能であることが示唆された．実際に得られた平均 Dice 係数が 0.9 より小さかった臓器の膵臓，胆のう，下大静脈，門脈は，95%信頼楕円が大きく分布のばらつきが大きいため予測精度のばらつきも大きかった．

図 6，図 7，図 8 から，エラーマップの値が大きい領域に PV マップの大きい値が存在することを確認し，不確実性と誤差との関連性を視覚的に確認した．このことは，能動学習において，学習データの修正を効率的に行う指針として重要であると考えられる．不確実性の高い領域を優先して修正することによって効率的に大規模データを構築できる可能性がある．

ただし，今回用いた CT 画像は比較的コントラストが近い動脈相の造影 CT 画像であり，今後は非造影や門脈相など様々な撮影条件での検証が必要である．

7. まとめ

本研究では，腹部臓器自動抽出における Bayesian U-Net の不確実性の有用性の検証を

目的とし，ラベル付き 2 時相造影 CT 画像データセットを用いて不確実性と抽出精度の関係の調査及び抽出精度の予測実験を行った．結果として Bayesian U-Net は，肝臓，脾臓，左右腎，大動脈，心臓，左右肺野において平均 Dice 係数 0.9 以上の抽出精度を示した．また，Bayesian U-Net から得られた不確実性は，抽出精度とよく相関しており，線形回帰モデルを用いると高い精度で実際の Dice 係数を予測した．

今後は，Bayesian U-Net で得られる不確実性を非造影や門脈相など様々な撮影条件で検証する．さらに，不確実性を用いたアクティブラーニングを行い，大規模データセットの効率的な構築手法を研究する．

謝辞

本研究は JSPS 科研費 19H01176 の助成を受けた．

利益相反の有無

なし．

文　献

［1］ Zhou X, Takayama R, Wang S, et al.: Deep learning of the sectional appearances of 3d ct images for anatomical structure segmentation based on an fcn voting method. Medical Physics, 44, 2017

［2］ Roth H, Oda H, Zhou X, et al.: An application of cascaded 3D fully convolutional networks for medical image segmentation. Comput Med Imag Grap, 66, 2018

［3］ Gibson E, Giganti F, Hu Y, et al. : Automatic Multi-Organ Segmentation on Abdominal CT With Dense V-Networks, IEEE Trans Med Imag 37: 1822-1834, 2018

［4］ Gal Y, Ghahramani Z, et al.: Dropout as a Bayesian approximation: Representing model uncertainty in deep learning. 33rd International Conference on Machine Learning 3, 1651–1660, 2016

［5］ Hiasa Y, Otake Y, Takao M, et al.: Automated Muscle Segmentation from Clinical CT using Bayesian U-Net for Personalized Musculoskeletal Modeling. IEEE Trans on Med Imag 39, 4, 1030-1040, 2019

［6］ Ronneberger O, Fischer P, Brox T, U-net: Convolutional networks for biomedical image segmentation. Lecture Notes in Computer Science, 9351, 234–241, 2015

［7］ Kendall A, Gal Y, et al.: "What uncertainties do we need in Bayesian deep learning for computer vision?". Advances in Neural Information Processing Systems 30, 5575–5585, 2017

［8］ He K, Zhang X, Ren s: Delving deep into rectifiers: Surpassing human-level performance on imagenet classification, IEEE International Conference on Computer Vision 1502, 1026–1034, 2015

［9］ Kingma D, Ba J, : Adam: A Method for Stochastic Optimization. International Conference on Learning Representations, 1-15, 2015

Usefulness of uncertainty estimation in automatic segmentation of abdominal organs in CT images using Bayesian U-Net : verification study

Yuto MASAKI[*1,*2], Yoshito OTAKE[*2], Mazen SOUFI[*2], Masatoshi HORI[*3], Hiromitsu ONISHI[*4], Noriyuki TOMIYAMA[*4], Yoshinobu SATO[*2]

*1 PSP corporation
*2 NARA Institute of Science and Technology
*3 Kobe University Graduate School of Medicine
*4 Osaka University Graduate School of Medicine

In recent years, the use of deep learning for automatic segmentation of abdominal organs from CT images has improved the segmentation accuracy, shortened the calculation time and made the clinical application more practical. However, it was difficult to estimate the accuracy of segmentation results in the absence of the manual (ground-truth) segmentations. Our group has recently investigated the relationship between the segmentation accuracy and prediction uncertainty based on Bayesian U-Net, and demonstrated the possibility of predicting the accuracy of automatic segmentation of musculoskeletal structures from CT images. The purpose of this study is to verify the usefulness of the Bayesian U-Net in automatic segmentation of abdominal organs from CT images. In this paper, we applied Bayesian U-Net to two phase contrast enhanced CT images of 48 cases (97 images) with manual segmentation of 13 organs, and report the correlation between the segmentation accuracy and prediction uncertainty.

Key words: Bayesian deep learning, Convolutional neural network, Contrast enhanced CT image, Multi organ segmentation, Active learning

Spatial Information Considered Self-Supervised Depth Estimation Based on Image Pairs from Stereo Laparoscope

Wenda LI[*1], Yuichiro HAYASHI[*1], Masahiro ODA[*1],
Takayuki KITASAKA[*2], Kazunari MISAWA[*3],
and Kensaku MORI[*1,4,5]

Abstract

This paper proposes a novel self-supervised depth estimation method considering spatial information of laparoscopic image pairs. Depth estimation plays an important role in robotic surgical navigation systems. There are many studies on depth estimation from laparoscopic images. In recent years, self-supervised siamese network has been proposed for depth estimation from stereo laparoscopic images. In this method, some of spatial information in the image will be lost as a result of continuous convolution and pooling. In order to reduce the loss of the spatial information in depth estimation, we add contextual encoding module and multi-scale space merging module to the original method. We also change the basic module to U-Net for feature encoding and feature decoding. The context encoding module and multi-scale spatial merging module is formed by dense atrous convolution block and spatial pyramid pooling block. Finally, we output disparity predictions by fusing multi-scales decoder outputs in the decoder blocks. After comparing with the original method, the experimental results showed the proposed method has about 8.9% improvement in SSIM and about 21% improvement in PSNR for stereo image pairs.

Keywords：Depth estimation, Laparoscopic surgery, Abdominal surgery, Self-supervised learning

1. Introduction

Minimally invasive robot-assisted laparoscopic surgery is useful due to small wounds and short recovery time compared with traditional surgery[1]. Depth estimation from laparoscopic images plays an important role in surgical assistance system for robotic surgical platform[2]. Recently deep learning has been developed to predict the whole depth maps or disparity maps[3,4]. However, a common limitation of CNN-based depth estimation[5] is that too many continuous convolution operations to learn abstract feature representations may reduce feature resolution. Although these operations is quite useful for classification or object detection tasks, it often hinders the prediction tasks such as depth estimation that require detailed spatial information[4,5]. To solve the above mentioned problems about depth

*1 Graduate School of Informatics, Nagoya University〔Furou-cho, Chikusa-ku, Nagoya 464-8601, Japan〕
e-mail: wdli@mori.m.is.nagoya-u.ac.jp
*2 Faculty of Information Science, Aichi Institute of Technology
*3 Aichi Cancer Center Hospital
*4 Information Technology Center, Nagoya University
*5 Research Center of Medical Bigdata, National Institute of Informatics

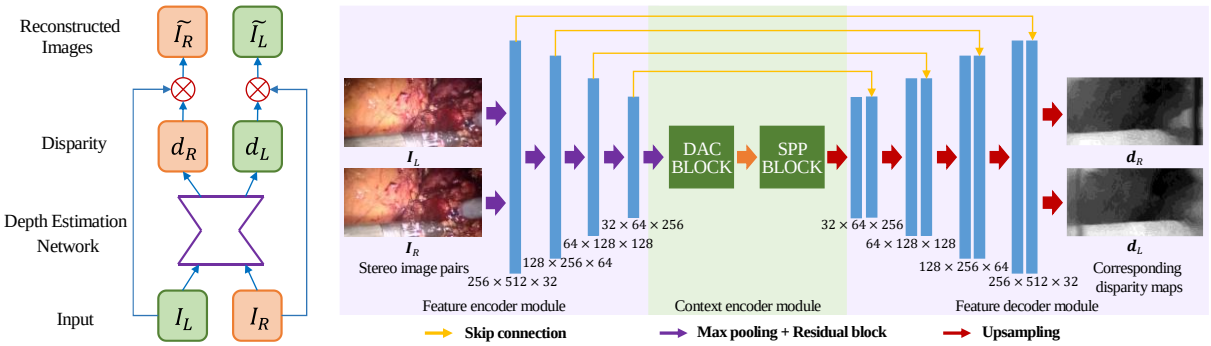


Fig. 1 Self-supervised deep learning network for depth estimation. (a) Self-supervised strategy of backward mapping (b) Diagram of our proposed depth estimation network. The blue block contains two 2D convolution layers with a maxpooling layer in feature encoder module and a deconvolution layer in feature decoder module. The numbers under blue blocks are the output dimension of them (channel×height×weight).

estimation, we proposed a novel self-supervised learning framework considered spatial information. We add the contextual encoding module named dense atrous convolution (DAC)[5] and multi-scale space merging module named spatial pyramid pooling (SPP)[6] to the previous method[4]. Furthermore, we output disparity predictions by fusing multi-scales decoder outputs in the decoder blocks inspired by[7].

2. Method

The proposed method predicts disparity maps based on image pairs from stereo laparoscope using self-supervised siamese network. We introduce U-Net with DAC module and SPP module as depth estimation network, and a different loss function to the previous method[4].

1) Depth estimation based on disparity prediction

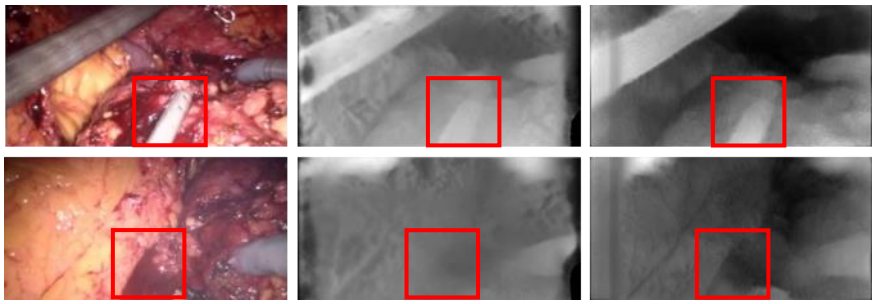
Following the self-supervised strategy of the previous method[4], the overview of this strategy for depth estimation is shown in Fig. 1(a). At training time, we input the rectified stereo image pairs as $\boldsymbol{I}_L$ and $\boldsymbol{I}_R$, corresponding to the left and right color images from stereo, into the proposed siamese network. Based on the left disparity map d_L and right image $\boldsymbol{I}_R$, this network reconstructs the left image represented as $\widetilde{\boldsymbol{I}_L}$. The right reconstructed image will also be predicted as $\widetilde{\boldsymbol{I}_R}$ in a similar way. Here, d_L and d_R means disparity maps that can be converted to depth maps by using the intrinsic parameters of the stereo cameras.

2) Depth Estimation Network

The proposed network consists of three major parts: the feature encoder module, the context extractor module, and the feature decoder module, as shown in Fig. 1(b). In the feature encoder module, our depth estimation network is based on the general U-Net architecture[8] where each block of encoder contains two convolution layers and one max pooling layer. Meanwhile, due to the loss of the spatial information of input image as the network depth continues to deepen. Inspired by [5] and [6], we introduce DAC and SPP to U-Net architecture called context extractor module, as shown in Fig. 1(b). DAC block is used to encode the high-level feature maps with four cascade branches that employs different receptive fields. SPP is also adopted to compress features into four scales where the low-dimensional feature maps are upsampled to the same size as input feature map. At last, in order to prevent local minimum trap in the training loss, we

Table 1 Comparison of results between the previous and the proposed method by SSIM and PSNR.

Methods	Previous Method[4]	Proposed Method
Mean SSIM (Left)	0.7884	0.8887
Mean SSIM (Right)	0.8433	0.8887
Mean PSNR (Left)	14.8683	19.8296
Mean PSNR (Right)	17.7534	19.8569



(a) Original images (b) Previous method (c) Proposed method

Fig. 2 Example and comparison of previous method[4] and proposed method. (a) original left image of stereo image pairs, (b) disparity maps estimated by the previous method, and (c) disparity maps estimated by the proposed method. Red boxes indicate the results of the previous method and the proposed method in the same area.

use the multi-scale estimation in the feature decoder module, which is similar to[7]. We output disparity maps at four different scales and upsample disparity maps to the original output size and they are also used to calculate loss.

3) Loss function

We adopt the multi-scale loss function which is a sum of different contributions computed at each scale.

3. Experimental results and discussion

The proposed method was implemented in Pytorch. We used the stereoscopic image pairs dataset from the Hamlyn Centre Laparoscopic/Endoscopic Video Datasets[4]. This dataset includes 30,816 rectified image pairs collected in partial nephrectomy using da Vinci robotic assisted surgery with stereo-laparoscope. The image pairs were resized to 256×512 pixels. We compared the performance of the proposed method with the previous method[4]. We trained on rectified image pairs from stereo laparoscope without any ground truth depth. Therefore, the criterion used in evaluation calculated based on the comparison between the reconstructed images predicted as $\widetilde{\boldsymbol{I}_L}$ and input original images predicted as $\boldsymbol{I}_L$ using the Structural Similarity (SSIM)[9] and Peak Signal-to-noise Ratio (PSNR).

The mean SSIM and PSNR of 7,167 stereo image pairs were shown in Table 1. Our experimental results showed that the proposed method has about 8.9% improvement in SSIM and about 21% improvement in PSNR for stereoscopic image pairs, comparing with the previous method[4]. Meanwhile, the spatial information of the input image was retained after continuous convolution operations, as shown red boxes in Fig. 2.

4. Conclusion

In this work, we proposed a novel self-supervised depth estimation method from stereo laparoscopic images. This method tried to keep the spatial information after a series of continuous convolution operations for depth estimation. Compared with the previous method, we added the DAC module and the SPP module as deep learning framework for

depth estimation. Furthermore, we introduced a classic multi-scale loss function to predict disparity maps at training time. Experimental result showed that our method improved the accuracy of the disparity map and retained the spatial information of the input image compared with the previous method as mentioned before. In this paper, our method is based on 2D images and it would be possible that proposed method is applied to 3D data in future work.

Competing interests

None.

Acknowledgement

Part of this research was supported by the MEXT/JSPS KAKENHI (17H00867, 26108006).

References

［1］Corcione F, Esposito C, Cuccurullo D, et al.: Advantages and limits of robot-assisted laparoscopic surgery: preliminary experience. Surgical Endoscopy and Other Interventional Techniques **19**(1): 117-119, 2005

［2］Lena M, Mountney P, Bartoli A, et al.: Optical techniques for 3D surface reconstruction in computer-assisted laparoscopic surgery. Medical image analysis **17**(8): 974-996, 2013

［3］ Nikolaus M, Ilg E, Hausser P, et al.: A large dataset to train convolutional networks for disparity, optical flow, and scene flow estimation. In Proceedings of the IEEE conference on computer vision and pattern recognition: 4040-4048, 2016

［4］ Ye M, Johns E, Handa A, et al.: Self-supervised siamese learning on stereo image pairs for depth estimation in robotic surgery. Hamlyn Symposium on Medical Robotics, 2017

［5］ Gu Z, Cheng J, Fu H, et al.: Ce-net: Context encoder network for 2D medical image segmentation. IEEE transactions on medical imaging **38**(10): 2281-2292, 2019

［6］ He K, Zhang X, Ren S, et al.: Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE transactions on pattern analysis and machine intelligence **37**(9): 1904-1916, 2015

［7］ Clément G, Aodha O, Brostow G: Unsupervised monocular depth estimation with left-right consistency. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition: 270-279, 2017

［8］ Olaf R, Fischer P, Brox T: U-net: Convolutional networks for biomedical image segmentation. International Conference on Medical image computing and computer-assisted intervention. Springer, Cham: 234-241, 2015

［9］ Wang Z, Bovik A, Sheikh H, et al.: Image quality assessment: from error visibility to structural similarity. IEEE transactions on image processing **13**(4): 600-612, 2004

空間情報を考慮した自己教師あり学習を用いたステレオ腹腔鏡画像からの深度推定

李　文達[*1], 林　雄一郎[*1], 小田 昌宏[*1], 北坂 孝幸[*2], 三澤 一成[*3], 森 健策[*1,4,5]

*1 名古屋大学大学院情報学研究科

*2 愛知工業大学情報科学部

*3 愛知県がんセンター

*4 名古屋大学情報基盤センター

*5 国立情報学研究所医療ビッグデータ研究センター

深度推定はロボット手術ナビゲーションシステムで重要な役割を果たす．そのため，腹腔鏡画像に基づく深度推定に関する多くの研究が行われている．近年，腹腔鏡画像からの深度推定において，自己教師あり学習を用いた手法が提案されている．この手法では，多くの畳み込みとプーリングにより画像内の情報が失われ，深度推定の精度が低下する問題があった．この問題を解決するために，深度推定のネットワークに U-Net を使用し，さらに U-Net に context encoding modul と multi-scale space merging module を追加する．これらは dense atrous convolution block と spatial pyramid pooling block によって構成されるものである．最後に，U-Net のデコーダーブロックのマルチスケール出力を融合することにより視差画像を出力する. 実験では, 提案手法は従来手法に比べてステレオ画像ペアに対する SSIM で約 8.9%改善し，PSNR で約 21%改善することを示した.

キーワード：深度推定, 腹腔鏡手術, 腹部手術, 自己教師あり学習

CNN を用いた指の運動想起脳波の識別

加藤　将[*1]　叶賀　卓[*2]　長島　和馬[*1]　星野　貴行[*2]　深見　忠典[*1]

要旨

身体動作の想起により生じる脳神経活動の脱同期パターンを脳波から識別する技術は，使用できる場面・タイミングの自由度が高いため，ブレインコンピュータインタフェース(BCI)への応用が期待されている．現存の機械学習による識別技術は大きな動作(例えば肘関節の屈曲や肩関節の外転)の想起時の脳波データに有効である. 言い換えると, 指の動きのような小さな動作想起時の脳波データを識別することは難しい．本稿では，指動作想起時の脳波データからでも識別力の高い脳波特徴を表現し，高精度に識別できる技術を検討する．このために，畳み込みニューラルネットワーク(CNN)を使用する．結果として，CNN による識別精度は各指の屈曲動作の想起時の 5 クラス脳波データを 59.3%で識別でき，一般的なニューラルネットワーク（NN）やサポートベクタマシン（SVM）より高精度な識別精度を示した．

キーワード：運動想起，指運動，畳み込みニューラルネットワーク

1. はじめに

運動想起は特定の身体部分を動かすイメージを行うことであり，脳神経活動の脱同期を誘発させる．この脳神経活動の変化は運動野周辺から計測された脳波の α および β (場合によっては γ や θ)周波数帯域で観測できるため, 運動想起はブレインコンピュータインタフェース(BCI)のパラダイムでよく用いられている. 事象関連電位のような他の BCI パラダイムと比較すると，任意の時間に入力を行えることから，使用する場面・タイミングの自由度が高いという利点がある．また，想起動作の種類を増やすことでコマンド数を増やすことができる．

想起の対象は，遠位にある身体部位の小さな動作の方が種類が豊富なため，コマンド数の増加に有利である．しかしながら，動作のダイナミクスが小さい場合，観測される脳波データの動作間の差異(特徴)も結果として小さくなるため，脳波での識別が困難となる．このため，ある程度の大きさの動作(例えば肘関節の屈曲や肩関節の外転)を想起対象とすることが一般的である．つまり，コマンド数の多い BCI の実現可能性がある一方で，実際にコマンド数を増やすのが難しいというのが運動想起ベース BCI の現状である．小さな動作に対して有効な特徴抽出・識別法を調査することは，この問題を打破し，運動想起ベース BCI の利便性を向上させることに寄与できる．

脳波の識別は様々な機械学習ベースの識別器が用いられるが，ニューラルネットワーク，特に畳み込みニューラルネットワーク(CNN)を用いた研究が盛んに行われている[1]． 時系列脳波データを時間周波数情報に変換し入力する場合 [2]や，EEGNet [3]のように，複数チャネルの時系列脳波データをそのまま入力する場合が提案されており，それぞれ高い識別精度が報告されている．

本研究では CNN とサポートベクターマシン（SVM），ニューラルネットワーク（NN）における指動作想起時の脳波データの識別精度の比較を行うことで，ダイナミクスの小さな運動想起ベース BCI に有効な識別法を検討する．また，それぞれの手法において(i) 時間領

*1 山形大学大学院理工学研究科情報科学専攻
〔〒992-8510 山形県米沢市城南 4-3-16〕
e-mail: trs53564@st.yamagata-u.ac.jp
*2 国立研究開発法人産業技術総合研究所 人工知能研究センター
〔〒135-0064 東京都江東区青海 2-4-7〕

域における脳波情報（時系列脳波データ），(ii) 周波数領域におけるパワースペクトル，(iii) 周波数領域における複素スペクトル，の 3 種類の脳波特徴との相性を調査することで，識別精度の向上に必要な要素を検討する．

2. 脳波データセット

本研究ではオープンデータセットである 5F データセット [4]を使用した． このデータセットには 8 名分のデータが含めれているが，含まれるセッション数が 1 から 4 までと異なっていており、各セッションでも試行数が異なっていたため，Subject A の 1 回目のセッションのデータのみを解析に使用した．

データは EEG-1200 JE-921 EEG system (Nihon Kohden, Japan)および脳波キャップ(Electro-Cap International, USA)により 200 Hz もしくは 1000 Hz で計測されている．解析対象のデータは 200 Hz で計測されている．公開されている時点で前処理として 0.53－70 Hz のバンドパスフィルタおよび 50 Hz のノッチフィルタが掛けられている．また，電極配置は，国際 10/20 法に従う 19 チャンネル(Fp1, Fp2, F3, F4, C3, C4, P3, P4, O1, O2, F7, F8, T3, T4, T5, T6, Fz, Cz, Pz)であり、基準電極は左および右耳朶(A1, A2)である．

このデータセットのタスク内容は，片手の各指の屈曲運動想起である．どれかの指を屈曲させる想像を 1 試行とし，それを 959 試行行う．最初の 1.5 から 2.5 秒間は想起をせず，そのあとに 1 秒間に各指に対応する番号(1 から 5)がランダムに提示されるので，実験参加者はその対象動作を想起した．300 試行(約 15 分)ごとに 2 分の休憩時間を設けられた．

3. 解析

1）時系列データ

解析に使用する時系列データは各試行でのタスク開始時(オンセット)から 2 秒間の 400 サンプル点（2 s × 200 Hz）である．また，使用するチャネルは 21 チャネルである．したがって，解析で使用する初期データ形状は（959 trials × 21 ch × 400 samples）である．

2）特徴データおよび前処理

前処理は識別器の種類および特徴によって異なる．そこで，SVM・NN と CNN で別々のセクションとして記述する．また，入力として用いる特徴は，(i) 時系列脳波, (ii) パワースペクトル, (iii) 複素スペクトル，の 3 種類である．

2.a) SVM・NN：

SVM と NN は 1 次元ベクトルが入力となるため，同じデータ形状を扱う．

時系列脳波を特徴とする場合は，40 Hz までの情報を使用するために 40 Hz のローパスフィルタを適用後，サンプリング周波数が 80 Hz になるようダウンサンプリングした．したがって，データ形状は 959×3360 次元（21 ch × 160 samples）である．

パワースペクトルを特徴とする場合は，直流成分を含めた 15 Hz 未満の情報を使用するために，各チャネルでフーリエ変換を窓長 2 s で行った．したがって，データ形状は 959×630 次元（21 ch × 30 frequency bins）となる．

複素スペクトルを使用する場合は，パワースペクトルと同じ設定とした．実数部と虚数部が存在するため，データ形状は 959×1239 次元（21 ch × (30 + 29) frequency bins）となる．

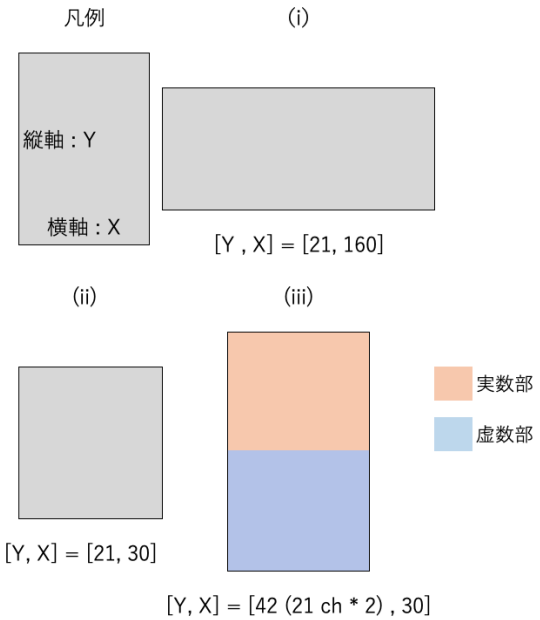


図 1 畳み込みニューラルネットワークにおける

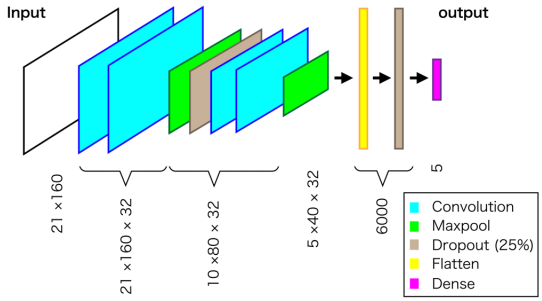


図 2 畳み込みニューラルネットワークの構成
(図の数値は(i)の画像サイズで示している)

2.b) CNN：
必要なデータ処理は SVM・NN と同じで，基本的には SVM・NN の特徴データの次元を 2 次元化している（図 1）.

3）学習パラメータ
SVM は先行研究 [5]から RBF カーネルを選択した．また SVM のハイパーパラメータである C および γ の値も先行研究を参考に決定した（$C=10$, $\gamma=0.001$）．NN は 3 層ニューラルネットワークとし，中間層を 4000，出力層を 5 クラス分類から 5 とした. CNN のネットワーク構成は図 2 の通りである.

4）評価方法
全手法の評価にあたっては, ネステッド 10 分割交差検証法を使用した．具体的には，データの全体を最初にランダムに 10 分割し, その 9 割をさらにランダムに 10 分割し，内側の 9 割を学習、1 割を検証，外側の 1 割をテストデータとした．この設定で得られたテストデータに対する識別精度に基づき精度を評価する.

4. 結果および考察

CNN と SVM，NN における 3 種類の特徴量: (i) 時間領域における脳波情報（時系列脳波データ）, (ii) 周波数領域におけるパワースペクトル，(iii) 周波数領域における複素スペクトル，の識別精度を表 1 に示す.

どの手法においてもパワースペクトルより時系列脳波データと複素スペクトルの方が高い精度を提示した．中でも CNN における時系列脳波データは最も識別精度が高く 59.3%であった.

表 1 各識別手法および特徴量における識別精度 (%)

	SVM	*NN*	*CNN*
(i)	46.6	49.4	59.3
(ii)	31.8	31.1	29.6
(iii)	52.8	48.8	32.8

解析に使用したデータは 1 施行につき 1 回の想起情報を含んでいる．このことから時系列脳波と複素スペクトルは識別に有益と考えられる位相情報を有しているために精度が高くなり，位相情報を持っていないパワースペクトルは低い識別精度を提示したと考えられる.

CNN と複素スペクトルの組み合わせによる識別精度が他の識別器(SVM と NN)に比べて低かった原因としては，入力形式に問題がある可能性がある．複素数部を実数部の隣に並べるのではなく，別レイヤーにすることで精度が向上すると考えられる.

5. 結論および今後の課題

本研究では CNN とサポートベクターマシン（SVM），ニューラルネットワーク（NN）における指動作想起時の脳波データの識別精度の比較を行うことで，ダイナミクスの小さな運動想起ベース BCI に有効な識別法を検討した．また，それぞれの手法において(i) 時間領域における脳波情報（時系列脳波データ），(ii) 周波数領域におけるパワースペクトル，(iii) 周波数領域における複素スペクトル，の 3 種類の脳波特徴との相性を調査することで，識別精度の向上に必要な要素を検討した.

結果から指動作想起のような小さな動作において位相情報が識別に重要な要素である可能性が示唆された．ただし，ユーザ内において運動想起の脱同期の位相は疲労や集中によって変化することが知られているため，より詳細な解析が必要である.

入力の形状やニューラルネットの構成によ

って識別精度が向上することが考えられるため，今後はさらに有効な入力形式・ニューラルネットの構成を検討していく予定である．

利益相反の有無

なし．

文　献

［1］ Cecotti H, Graeser A: Convolutional neural network with embedded Fourier transform for EEG classification. IEEE International Conference on Pattern Recognition: 1-4, 2008

［2］ Lai D, Zhang X, Ma K, et al.: Automated detection of high frequency oscillations in intracranial EEG using the combination of short-time energy and convolutional neural networks. IEEE Access **7**: 82501-82511, 2019

［3］ Lawhern J, Solon J, Waytowich R, et al.: EEGNet: A compact convolutional neural network for EEG-based brain–computer interfaces. J. Neural Eng. **15**(5): 056013, 2018

［4］ Kaya M, Binli K, Ozbay E, et al.: A large electroencephalographic motor imagery dataset for electroencephalographic brain computer interfaces. Scientific data **5**: 180211, 2018

［5］ Kato M, Kanoga S, Hoshino T, et al.: Motor imagery classification of finger motions using multiclass CSP. 42nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, in press, 2020

腹腔鏡手術動画像データベース構築に向けた リモートアノテーションツールのプロトタイプ開発

屠　芸豪[*1]　伊東　隼人[*1]　小澤　卓也[*1]　小田　昌宏[*1]

竹下　修由[*2]　伊藤　雅昭[*2]　森　健策[*1,3,4]

要旨

本稿では腹腔鏡手術動画像における術具や解剖構造をウェブブラウザを介してアノテーションするためのリモートアノテーションツールの開発について報告する．腹腔鏡手術動画像において術具や解剖構造のアノテーションデータを作成することにより，機械学習を利用した画像処理によって手技解析法や手術支援技術の開発が可能となる．先行研究として我々は動画像に向けたアノテーションソフトウェア NuVAT を開発した．しかし大規模データベースの構築においては，病院へ専門的な機材やソフトウェアを導入するコストが高い．加えて，データの外部への持ち出しによる流出が懸念される．そこで，開発したソフトウェアをウェブブラウザ上で動作するように拡張することによって，病院へのシステム導入のコスト削減ならびに web を介したアノテーション作業によって作業の効率化を図る．

キーワード：アノテーションツール，機械学習，セグメンテーション，腹腔鏡手術動画像解析，手術支援

1. 背景と目的

腹腔鏡下手術は低侵襲かつ入院期間が短いという利点を伴った外科手術であるために普及しているものの，手術そのものの難易度の高さや術者の育成といった課題がある[1-3]．この課題に対し，腹腔鏡手術動画像を機械学習等を用いてコンピュータで自動的に解析し，術者育成をサポートする取り組みが多く行われている[4,5]．しかし，機械学習モデルの構築には大量のアノテーションデータが必要である．特に大規模な腹腔鏡動画像データベースを構築するためには，複数施設へ専門的な機材やソフトウェアを導入するコストが高いと考える．そこで，大規模アノテーションに向けた処理基盤の構築が望まれている．

関連研究として，NuVAT[6]，TSLAB[7]，PLUTO[8]，MITK[9]，CVAT[10]，そして labelme[11]というソフトウェアの開発が報告されている．PLUTO と MITK は 3D 医用画像用のツールである．TSLAB は 2D 画像用の半自動アノテーションツールである．CVAT と labelme は動画像中の人や車を矩形でアノテーションするツールであるが，複雑な形状には向いていない．NuVAT は医用動画像向けのアノテーションツールである．

大規模アノテーションに向けて，インターネットを介したウェブアプリケーション利用による病院へのシステム導入のコスト軽減が期待できる．さらには，多人数の作業の効率

*1 名古屋大学大学院情報学研究科
〔〒464-8601 名古屋市千種区不老町〕
e-mail: ytu@mori.m.is.nagoya-u.ac.jp
*2 国立がん研究センター東病院大腸外科
*3 名古屋大学情報基盤センター
*4 国立情報学研究所医療ビッグデータ研究センター

化を図ることができると考える．本研究では NuVAT のユーザーインターフェースを改善し，サーバ上に蓄積された動画像に対してウェブブラウザからのアノテーションを可能とする NuVAT-Web（Nagoya University Video Annotation Tool-Web）のプロトタイプ開発を行った．

2. NuVAT-Web の概要

図 1 は NuVAT-Web の動作環境を示す．NuVAT-Web はクライアントサイドとサーバサイドによって構成される．図 2 はアノテーション作業画面を示す．クライアントサイドにてユーザーのパソコンからウェブブラウザに接続し，アノテーション作業を行う画面を表示する．表 1 に NuVAT-Web を実装するためのソフトウェアを示す．本研究において開発した NuVAT-Web は以下の特徴がある．

(1) リモートワークによる作業：インターネットを介してウェブブラウザ上でのアノテーション作業が可能．病院へのシステム導入が容易になり，導入コストや負担の削減が期待できる．

(2) 多様な描画ツール：画素単位で効率的にかつ容易に腹腔鏡動画像のアノテーション作業が可能．これを実現するために，複数種類の描画ツールや補助機能を実装した．表 2 に各描画ツールを説明する．作業画面の右側にある「Draw Tools」を通じて選択することができる．

(3) 複数種類の動画像フォーマット対応：動画像を読み込む際に，mp4，mov，AVI，mpg 等の複数の動画像フォーマットに対応する．

(4) 利用しやすいユーザーインターフェース：

①画像表示：1 枚フレームの画像をできるだけ大きく表示．②表示フレームの選択とアノテーションラベルのコピー：表示画像の下に複数のボタンを設置．フレームの遷移とスキップ，フレームラベルのコピー及び描画内容の保存を実行することができる．ボタンをクリックすれば,対応する機能が実行される．

表 1 NuVAT-Web を実装するソフトウェア

クライアント サーバシステム	ソフトウェア
クライアント サイド	アノテーション作業 画面の作成： HTML, CSS, JavaScript
	画像処理ライブラリ： OpenCV.js
サーバ サイド	実行環境： Node.js
	画像処理ライブラリ： opencv4nodejs

表 2 NuVAT-Web に実装した描画ツール

描画ツール	概要説明
Pencil	フリーハンド描画
Line	一時的な補助線を引きながら直線を描画
Polygon	複数の頂点を指定し，ポリゴンを描画
Closed Area Filling	閉領域の塗りつぶし
Eraser	描画内容を選択して消去
CanvasClear	描画した全内容を消去

表 3 実験に用いたサーバの仕様

構成要素	仕様
Server	Virtual server
OS	Debian 10.3
CPU	Intel(R) Xeon(R) CPU E5-2620 v4 @ 2.10GHz
Memory	4GB
Storage	1016GB

フレームラベルのコピーでは，現在のフレームのアノテーション結果を前/後のフレームにコピーする機能が実行される.これにより，アノテーション作業の手間を軽減することができると考える．各フレームの間に多少ずれがある場合，描画ツールを使用し，細かく修正することができる．③ツールプロパティの設定:「Tools Property」を通じて描画ツールのプロパティ（線と点の太さ）を直接調整することができる．④ラベルの追加や色の設定:「Annotation Pallet」を通じてラベルの追加やラベル色の調整が可能である．

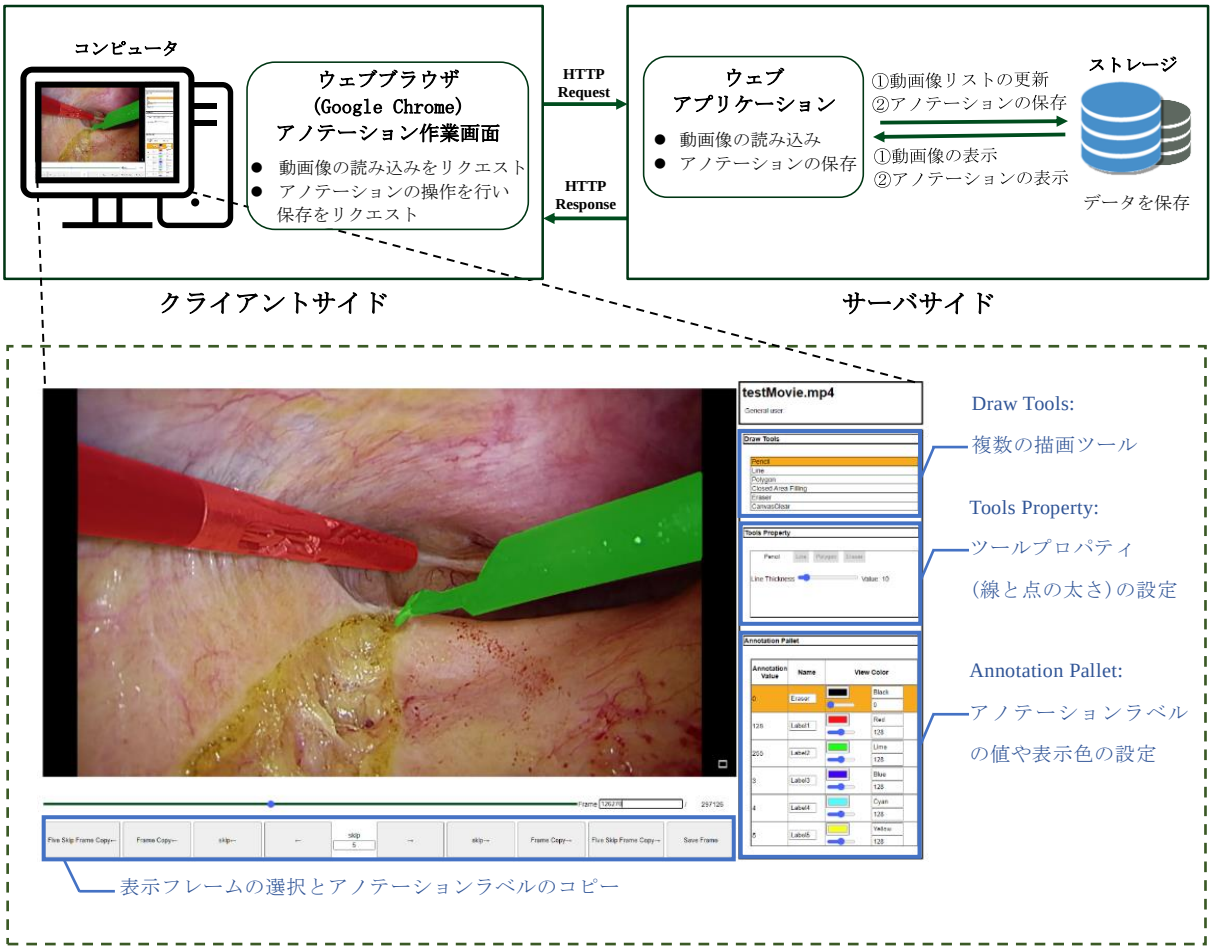


図 1 NuVAT-Web の動作環境

3. 実験と結果

NuVAT-Web と NuVAT の実行時間の比較を行った．今回はローカルネットワーク内で動作を行い，サーバの仕様を表 3 に示す．実験環境を示す．①データを読み込んでアノテーションが可能になるまでの時間は NuVAT-Web は約 2 秒，NuVAT は約 3 秒であった．②1 つのフレームからほかのフレームに遷移する時間は NuVAT-Web と NuVAT は両方とも約 1 秒であった．これらより NuVAT-Web はアノテーション作業の効率を落とさない高速な動作が可能であることがわかる．なお，計測時間に動画像のアップロード時間は含まない．

4. むすび

本研究ではサーバ上に蓄積された動画像に対してウェブブラウザからのアノテーションを可能とする NuVAT-Web を開発した．ウェブブラウザのみを必要とする NuVAT-Web により，病院へ専門的な機材やソフトウェアを導入するコストの削減が期待できる上に，リモートワークによるアノテーション作業が可能となる．さらに，複数種類の描画ツールや補助機能によって，アノテーション作業の効率化を図ることができると考える．今後，動画像に特化した補助機能の追加ならびにデータベースとの連携を実装し，内視鏡動画処理基盤の開発に繋げていきたい．

謝辞

本研究の一部は AMED（20he2102001h0702）の支援を受けたものである．

利益相反の有無　なし

文　献

［1］ 池田篤，渡邊昌彦 : 消化器外科 5 臨時増刊号 アトラスで学ぶ達人の手術．第 34 巻第 6 号．へるす出版，東京，2011，pp830-838.

［2］ 山本聖一郎，赤須孝之，藤田伸　他：消化器外科 5 臨時増刊号 鏡視下手術のための局所解剖アトラス．第 30 巻第 6 号，へるす出版，東京，2007，pp838-848.

［3］ 河村裕，小西文雄 : 消化器外科 5 臨時増刊号 鏡視下手術のすべて．第 27 巻第 6 号，へるす出版，東京，2004，pp855-860.

［4］ Shvets A, Alexander R, Alexandr A, et al.: Automatic Instrument Segmentation in Robot-Assisted Surgery using Deep Learning. 2018 17th IEEE International Conference on Machine Learning and Applications, DOI: 10.1109/ICMLA. 2018.00100: 624-628，2018.

［5］ 小澤卓也，林雄一郎，小田紘久　他：仮想腹腔鏡画像生成と深層学習による腹腔鏡画像からの術具領域セグメンテーション, 電子情報通信学会技術研究報告(MI)，**MI2019-94-119**(399)，129-134，2020.

［6］ 小澤卓也，小田紘久，伊東隼人　他：機械学習による内視鏡動画インスタンスセグメンテーションのための手動アノテーションツールの開発，第 37 回日本医用画像工学会大会予稿集，94-97，2018.

［7］ Carlos C, Yanez EM, Garca N: Tool for Semiautomatic Labeling of Moving Objects in Video Sequences: TSLAB. Sensors, **15**(7): 15159-15178, 2015.

［8］ 二村幸孝，出口大輔，北坂孝幸 他：PLUTO：医用画像診断支援共通プラットフォーム．Medical Imaging Technology，**26**(3)，187-191，2008.

［9］ Nolden M, Zelzer S, Seitel A et al: The Medical Imaging Interaction Toolkit: challenges and advances. International Journal of Computer Assisted Radiology and Surgery, **8**(4): 607-620, 2013.

［10］ Computer Vision Annotation Tool (CVAT): https://github.com/opencv/cvat (2020 年 7 月 21 日アクセス)

［11］ labelme: https://github.com/wkentaro/labelme (2020 年 7 月 21 日アクセス)

Prototype Development of a Remote Annotation Tool for Laparoscopic Video Database Construction

Yunhao TU[*1], Hayato ITOH[*1], Takuya OZAWA[*1], Masahiro ODA[*1],
Nobuyoshi TAKESHITA[*2], Masaaki ITO[*2], Mori KENSAKU[*1,3,4]

*1 Graduate School of Informatics, Nagoya University
*2 Department of Colorectal Surgery, National Cancer Center Hospital East
*3 Information Technology Center, Nagoya University
*4 Research Center for Medical Bigdata, National Institute of Informatics

We report a prototype of remote-annotation tool for web-based annotation of surgical instruments and anatomical structures in laparoscopic surgery videos. Large-scale laparoscopic-surgery videos with annotation labels of surgical instruments and anatomical structures enable us to develop machine-learning-based methods for surgical procedure analysis and computer-assisted surgery. As a previous work, we developed an annotation software: NuVAT for laparoscopic videos. However, for a construction of large-scale database, setting an annotation system in hospital is cumbersome. Furthermore, data carry among several institutions and data management, which prevents leakage of personal information, are tiresome. Therefore, we extend the NuVAT to web application, which we can use via web browser for remote annotation. This remote-annotation tool can reduce the cost of system constructions in hospitals and can improve an efficiency of data annotation.

Key words: Annotation tool, Machine learning, Segmentation, Laparoscopic surgery image analysis, Surgery support

Uncertainty-based prediction of nucleus segmentation accuracy in pathological images of lung cancer using Bayesian U-Net

Mazen SOUFI*[1], Yoshito OTAKE[1], Akinori NAGASHIMA[2], Tomoya MYOJIN[3], Chiho OHBAYASHI[3], Yoshinobu SATO[1]

Abstract

Nucleus segmentation is indispensable in accelerating the processing pipeline of high-resolution pathological images, i.e. whole slide images (WSIs). Convolutional neural networks, such as U-Net, have demonstrated potential in nucleus segmentation tasks. However, given the high resolution of WSIs, creating ground-truth labels from the whole image is a high-cost task, besides its difficulty due to overlapping structures and imaging conditions. Therefore, a predictive measure of the automatic segmentation accuracy in the absence of ground-truth is needed to facilitate the process, for example by using active learning. In this study, we investigated the potential of Bayesian U-Net in predicting the segmentation accuracy of nuclei in pathological images based on uncertainty estimation. The basic idea was to model the uncertainty by inferring the target labels using Monte Carlo dropout sampling at test time and estimating the variance of prediction probabilities. A Bayesian U-Net with dropout layers (dropout ratio= 0.5) was used. The segmentation mask was calculated based on the mean softmax probabilities of the prediction in 10 samples at test time, while the uncertainty was modeled as the area variation coefficient (AVC) of the segmented nucleus areas (in pixels) in the 10 samples. To evaluate the performance, leave-one-patient-out cross-validation was performed on patches (n= 20,069; matrix size= 128×128 pixels) sampled from WSIs of 11 lung cancer patients. The segmentation accuracy was assessed using Dice coefficient. The correlation between the predictive uncertainty and Dice coefficient was investigated. A conventional (i.e. deterministic) U-Net was used as baseline. The accuracy of the Bayesian and baseline U-Nets were 0.910±0.079 and 0.902±0.092, respectively ($P < 0.001$; Mann-Whitney U-Test). The correlation coefficient between the AVC and Dice coefficient for tumor and non-tumor nuclei were -0.664 and -0.610, respectively. This showed the feasibility of Bayesian U-Net in predicting the segmentation accuracy of the nuclei in pathological images based on uncertainty estimation.

Keywords： Pathological image, Bayesian U-Net

1. Introduction

Nucleus segmentation is indispensable in accelerating the processing pipeline of high-resolution pathological images, i.e. whole slide images (WSIs) [1, 2]. Convolutional neural networks, such as U-Net [3], have demonstrated potential in nucleus segmentation tasks. However, given the high resolution of WSIs, creating ground-truth labels from the whole image is a non-trivial task. Therefore, a predictive measure of the segmentation accuracy in the absence of ground-truth is needed. In this study, we investigated the potential of Bayesian U-Net [4] in predicting the segmentation accuracy of nuclei in pathological images based on uncertainty estimation. The basic idea was to model the uncertainty based on

1 Nara Institute of Science and Technology〔8916 - 5 Takayamacho, Ikoma, Nara, Japan 630-0192〕

*e-mail: Mazen Soufi：msoufi@is.naist.jp

2 School of Medicine, Nara Medical University

3 Department of Diagnostic Pathology, Nara Medical University

dropout training, and inferring the target labels using Monte Carlo dropout sampling at testing time.

2. Method

A Bayesian U-Net of 5-layers depth with dropout (ratio= 0.5) was used. The nucleus was segmented by averaging the softmax probabilities of the prediction in 10 samples at test time, while the uncertainty was modeled as the area variation coefficient (AVC) of the segmented nucleus areas (in pixels) in the samples. To evaluate the performance, leave-one-patient-out cross-validation was performed on patches (n= 20,069; matrix size= 128×128 pixels) sampled from WSIs of 11 lung cancer patients. The segmentation accuracy was assessed using Dice coefficient (DC). The correlation between the predictive uncertainty and Dice coefficient was investigated. A conventional (i.e. deterministic) U-Net was used as baseline.

3. Results

Figure 1 shows the segmentation accuracy of the annotated nuclei in the WSIs. The average DC of the Bayesian and baseline U-Nets were 0.910±0.079 and 0.902±0.092, respectively ($P < 0.05$; Mann-Whitney U-Test).

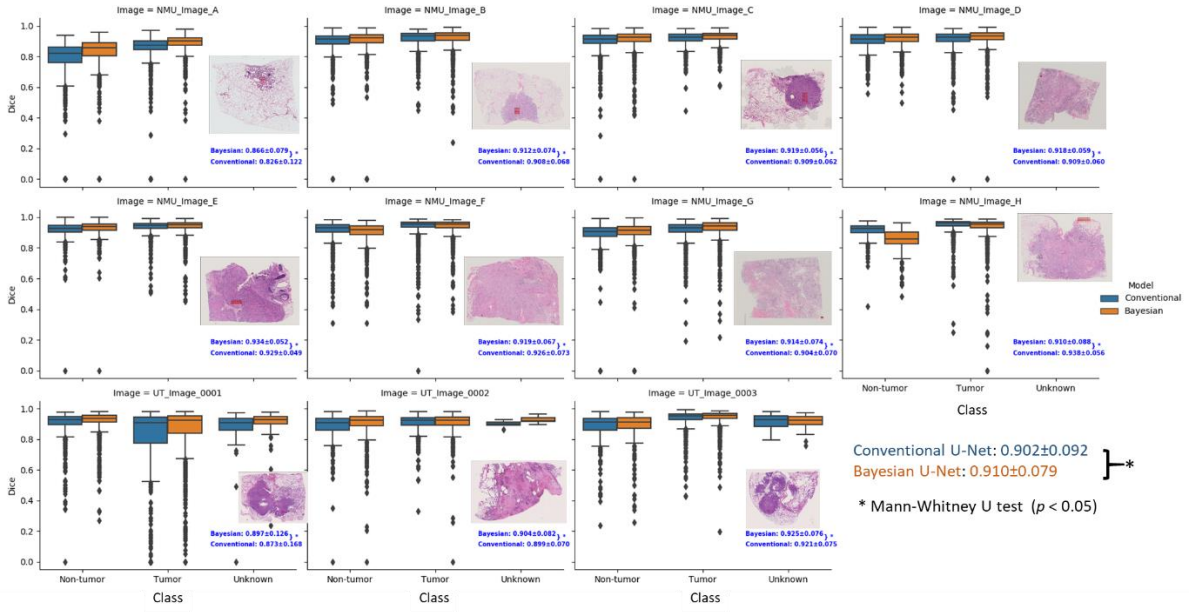


Figure 1: Segmentation accuracy (represented by Dice coefficient) for the annotated nuclei in 11 WSIs.

Figure 2 shows representative segmentation results with joint histograms indicating the correlation between the uncertainty (AVC) and segmentation accuracy (Dice). The correlation coefficient between the AVC and Dice coefficient for tumor and non-tumor nuclei were -0.664 and -0.610, respectively. Though some images had an overall strong correlation (e.g. -0.730 in image #3), the average values show a moderate correlation (<0.7) between the AVC and segmentation accuracy.

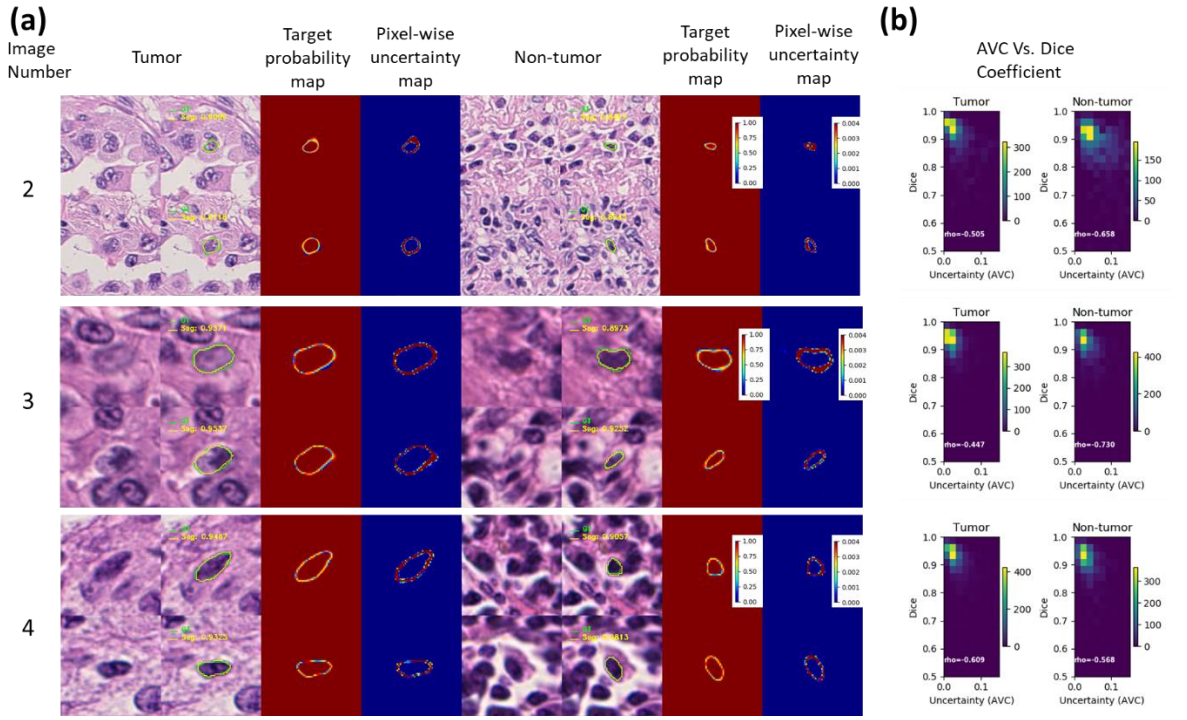


Figure 2: Representative figures for segmentation results. (**a**) Patches with segmented nuclei and pixel-wise uncertainty map. (**b**) Joint histograms showing relationship between uncertainty (area variation coefficient (AVC)) and segmentation accuracy (Dice coefficient).

4. Discussion and Conclusion

This study addressed the segmentation of nuclei in non-tumor and lung tumor cells in WSIs. By using dropout sampling via Bayesian U-Net, we obtained a predictive measure of the segmentation accuracy of the nuclei in tumor and non-tumor cells. We showed the feasibility of Bayesian U-Net in improving and predicting the segmentation accuracy of the nuclei in pathological images based on uncertainty estimation.

Competing interests

The authors have no competing interests.

Acknowledgement

This research was supported by Japan Agency for Medical Research and Development (AMED) Number JP20lk1010036.

5. References

[1] Graham S, Dang V Q, Raza SEA, et al.: Hover-Net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images. Med Image Anal. 2019 Dec;58:101563.

[2] Fraz MM., Khurram SA, Graham S.: FABnet: feature attention-based network for simultaneous segmentation of

microvessels and nerves in routine histology images of oral cancer Neural Computing and Applications Vol. 32 (2020), pp.: 9915–9928

［3］ Ronneberger O., Fischer P., and Brox T.: U-net: Convolutional networks for biomedical image segmentation." International Conference on Medical image computing and computer-assisted intervention (MICCAI2015). Springer, Cham, 2015.

［4］ Hiasa Y., Otake Y, Takao M., et al.: Automated muscle segmentation from clinical CT using Bayesian U-net for personalized musculoskeletal Modeling. IEEE Transactions on Medical Imaging 39.4 (2019): 1030-1040.

第 39 回日本医用画像工学会大会（JAMIT2020）
大会役員委員会名簿

五十音順・敬称略

大会長	
湯浅　哲也	山形大学
大会長補佐	
深見　忠典	山形大学
業務委員長	
大沢　博之	キヤノンメディカルシステムズ株式会社
チュートリアル委員長	
原　　武史	岐阜大学
企業連携委員長	
中田　典生	東京慈恵会医科大学
プログラム委員会	
委員長	
根本　充貴	近畿大学
委　員	
有村　秀孝	九州大学
石田　隆行	大阪大学
伊藤　聡志	宇都宮大学
彌冨　　仁	法政大学
上村　幸司	国立循環器病研究センター
内山　良一	熊本大学
大城　　理	大阪大学
小田　昌宏	名古屋大学
小尾　高史	東京工業大学
河田　佳樹	徳島大学
北坂　孝幸	愛知工業大学
北村　圭司	株式会社 島津製作所
木戸　尚治	大阪大学
金　　亨燮	九州工業大学
木村　裕一	近畿大学
工藤　博幸	筑波大学
佐藤　嘉伸	奈良先端科学技術大学院大学
椎名　　毅	京都大学
清水　昭伸	東京農工大学
周　　向栄	岐阜大学
庄野　　逸	電気通信大学
陣崎　雅弘	慶應義塾大学
菅　　幹生	千葉大学
杉本　直三	京都大学
滝沢　穂高	筑波大学
田中　利恵	金沢大学
陳　　延偉	立命館大学
中口　俊哉	千葉大学
野村　行弘	東京大学
畑中　裕司	滋賀県立大学
花岡　昇平	東京大学
羽石　秀昭	千葉大学
平野　　靖	山口大学
本谷　秀堅	名古屋工業大学
増谷　佳孝	広島市立大学
宮崎　　靖	株式会社 日立製作所
目加田慶人	中京大学
森　　健策	名古屋大学
山谷　泰賀	量子科学技術研究開発機構 放射線医学総合研究所

日本医用画像工学会（JAMIT）
役　員　名　簿
（令和２年６月22日現在）

五十音順・敬称略

	氏名	所属
会　　長	尾川　浩一	法政大学
副 会 長	工藤　博幸	筑波大学
	佐藤　嘉伸	奈良先端科学技術大学院大学
常任幹事	伊藤　聡志	宇都宮大学
	大沢　博之	キヤノンメディカルシステムズ
	小尾　高史	東京工業大学
	北坂　孝幸	愛知工業大学
	中田　典生	東京慈恵会医科大学
	原　　武史	岐阜大学
	花岡　昇平	東京大学
	増谷　佳孝	広島市立大学
	宮崎　　靖	日立製作所
	目加田慶人	中京大学
幹　　事	安藤　　裕	JCHO 埼玉メディカルセンター
	石垣　武男	名古屋城北放射線科クリニック
	石田　隆行	大阪大学
	上村　幸司	国立循環器病研究センター
	大松　広伸	網走刑務所
	岡田　知久	京都大学
	小田　昌宏	名古屋大学
	掛川　　誠	中央システム技研
	笠井　　聡	コニカミノルタ
	川岸　将実	キヤノン
	河田　佳樹	徳島大学
	木戸　尚治	大阪大学
	黒木　嘉典	新村病院、博愛会 HPN 東京
	陣崎　雅弘	慶応義塾大学
	杉本　直三	京都大学
	高橋　哲彦	群馬県立県民健康科学大学
	滝沢　穂高	筑波大学
	武田　　徹	筑波記念病院
	田中　利恵	金沢大学
	陳　　延偉	立命館大学
	土居　篤博	富士フイルム
	縄野　　繁	国際医療福祉大学
	野崎　太希	聖路加国際病院
	長谷川純一	中京大学
	畑中　裕司	滋賀県立大学
	羽石　秀昭	千葉大学
	平野　　靖	山口大学
	藤田　広志	岐阜大学
	細羽　　実	京都医療科学大学
	森　　健策	名古屋大学
	森　　雅樹	札幌厚生病院
	山谷　泰賀	量子科学技術研究開発機構放射線医学総合研究所
	湯浅　哲也	山形大学
	吉川　悦次	浜松ホトニクス
監　　事	篠川　　毅	島津製作所
	清水　昭伸	東京農工大学