

共同企画

共同企画4

AIを用いた画像解析および画像診断－医療画像関連6学会によるAMED事業の成果と課題－

2021年11月19日(金) 14:10～16:10 |会場(4号館3階431+432)

[2-I-2-04] 日本病理学会における病理診断支援 AI開発 － Japan Pathology AI Diagnostics Project (JP-AID)－ AI for support of pathological diagnosis developed by Japanese Society of Pathology - Japan Pathology AI Diagnostics Project (JP-AID) -

*阿部 浩幸^{1,2}、吉澤 明彦^{1,3}、深山 正久^{1,2,4}、北川 昌伸^{1,5}（1. 日本病理学会、2. 東京大学、3. 京都大学、4. 国保旭中央病院、5. 東京医科歯科大学）

*Hiroyuki Abe^{1,2}, Akihiko Yoshizawa^{1,3}, Masashi Fukayama^{1,2,4}, Masanobu Kitagawa^{1,5}（1. The Japanese Society of Pathology, 2. The University of Tokyo, 3. Kyoto University, 4. Asahi General Hospital, 5. Tokyo Medical and Dental University）

キーワード：Artificial intelligence, Pathology, Whole slide image, Gastric biopsy, Segmentation

日本では慢性的な病理医の不足が続いており、病院内に勤務する病理医が1名しかいない、いわゆる「一人病理医」が少なくない。日本病理学会では日本医療研究開発機構(AMED)の支援を受け、国立情報学研究所(NII)と連携し、主に一人病理医の診断ダブルチェックに役立つような病理診断支援 AIの開発に取り組んだ。まず日常病理診断業務で最も頻度の高いと考えられる胃生検を対象に、癌の有無及び存在範囲の描出を行う AIを開発した。病理デジタル画像(whole slide image, WSI)は超高解像度の画像であり、通常の GPUで処理するには画像を細かいパッチに分割する必要があるが、AIにパッチレベルでの学習と判定をさせるだけでは弱拡大の情報が生かせず、偽陽性が多くなる問題が生じた。そこでパッチから抽出した特徴量の元画像における分布も考慮した学習手法(multi-stage semantic segmentation for pathology, MSP法)を開発し、機械学習に適用した。その結果、癌/非腫瘍の判定において感度92.8%、特異度94.6%、ROC曲線の AUC値0.975という実用レベルの AIモデルを作成できた。このモデルは、異なる10施設から収集した病理画像でも検証され、AUC値の平均0.969±0.020と高い性能を発揮できることが示された。また本事業では、開発した AIモデルを福島・徳島の地域遠隔病理診断ネットワークで実装実験を行い、その有用性を確認した。MSP法は胃生検だけでなく大腸生検でも有用であり、非腫瘍/良性腺腫/癌の3段階分類で病理診断との一致率92%を達成した。JP-AID事業では他にも子宮頸部生検(異形成・癌の検出)、胃生検(腸上皮化生の評価)、核分裂像計数、腫瘍細胞比率判定、肺腺癌の間質浸潤判定、乳腺乳管内病変の良悪判定、の各課題でも多施設からのアノテーション収集と AIモデルの作成・検証を実施した。本事業で作成した AIモデルが病理診断業務の負担軽減につながる事が期待される。

日本病理学会における病理診断支援 AI 開発 - Japan Pathology AI Diagnostics Project (JP-AID) -

阿部浩幸^{*1,2}、吉澤明彦^{*1,3}、深山正久^{*1,2,4}
北川昌伸^{*1,5}

*1 日本病理学会、*2 東京大学、*3 京都大学
*4 国保旭中央病院、*5 東京医科歯科大学

AI for support of pathological diagnosis developed by Japanese Society of Pathology - Japan Pathology AI Diagnostics Project (JP-AID) -

Hiroyuki Abe^{*1,2}, Akihiko Yoshizawa^{*1,3}, Masashi Fukayama^{*1,2,4},
Masanobu Kitagawa^{*1,5}

*1 The Japanese Society of Pathology, *2 The University of Tokyo, *3 Kyoto University,
*4 Asahi General Hospital, *5 Tokyo Medical and Dental University

In Japan, the number of pathologists is small and a lot of pathologists are working alone at hospital without colleagues providing secondary reviews for pathological diagnoses. The Japanese Society of Pathology (JSP) collaborated with National Institute of Informatics to develop AI models to support pathologists. There are eight projects of AI development, one of which is detection of tumor from the whole slide images of gastric biopsy. We prepared 996 cases, 4550 tissues of gastric biopsy with annotations of pathologists for training dataset, and developed AI model for tumor detection. Additionally, we have developed a new algorithm for AI training, which is named as “multi-stage semantic segmentation for pathology (MSP)”. MSP can utilize information from low-power view as well as images of small patches (high-power view) by learning the distribution of feature values extracted from each patch. As a result, accuracy of 94.4% and AUC of 0.975 were achieved. In addition, the AI model was validated with samples collected from 10 institutions, and MSP showed stable and high performance compared to the conventional patch-based method. The developed AI model was also implemented in the networks for remote pathological diagnosis in Fukushima and Tokushima prefectures. The AI model is thought to be useful for double-check of pathologists’ diagnoses.

Keywords: Artificial intelligence, Pathology, Whole slide image, Gastric biopsy, Segmentation.

1. 緒論

日本では慢性的な病理医の不足が続いている。病理専門医は全国に 2620 名(2020 年 11 月現在)であり、人口当たりの専門医数では米国の 1/3 以下にとどまっている。400 床以上の一般病院 706 施設のうち 31.7%にあたる 224 病院で常勤病理医が不在(2017 年データ)であり、病理学会認定・登録施設 856 のうち 43.6% (368 施設) (2016 年データ)では常勤病理医が一人しか勤務していない(日本病理学会施設年報調査)。多くの病理医が同僚のダブルチェックを受けることもできない過酷な環境下で勤務せざるを得ない状況におかれている(いわゆる「一人病理医」問題)。

近年、機械学習に基づく人工知能の発達が著しく、特に畳み込みニューラルネットワークは画像認識の分野で大きな成果を示し、医療分野への応用が期待されている。病理診断の分野では放射線画像や内視鏡画像と比較し、画像のデジタル化において遅れを取っていたが、近年はスライド全体をデジタル化する whole slide image (WSI)スキャナーを各社が開発しており、病理画像デジタル化の技術的障壁は低くなってきた。

以上のような背景を踏まえ、日本病理学会では日本医療研究開発機構(AMED)の支援を受け、国立情報学研究所(NII)と連携し、様々なタスクに対応する病理診断支援 AI の開発に取り組んだ。具体的には「胃生検からの腫瘍検出」、「大腸生検からの腫瘍検出」(東京大学分担課題)、「子宮頸部生検からの異形成・癌の検出」(京都大学分担課題)、「乳腺生

検における核分裂像の計数」(秋田大学分担課題)、「肺腺癌浸潤の判定」(筑波大学分担課題)、「胃生検における慢性胃炎評価」(信州大学分担課題)、「肺癌における腫瘍細胞比率推定」(奈良県立医科大学分担課題)、「乳腺乳管内病変の良悪判定」(広島大学分担課題)の 8 課題について、それぞれ NII に所属する情報系研究者と開発チームを結成して取り組んだ。本発表では特に日常病理診断における業務量の負荷が多いと考えられる「胃生検からの腫瘍検出」病理診断支援 AI モデル開発と、胃生検 AI モデルを遠隔病理診断ネットワークに実装した社会実験の成果について発表する。

2. 目的

日常病理診断業務で最も頻度の高いと考えられる胃生検を対象に、病理画像から癌の有無の判定及び存在範囲の描出を行う AI モデルを作成する。また胃生検 AI モデルを福島県・徳島県に敷設した遠隔病理診断ネットワークに実装し、その有用性について検討する。

3. 方法

3.1 データセット

学習データとして東京大学医学部附属病院で施行された、腫瘍を含む胃生検 996 症例、4550 組織片の HE 染色病理組織標本を用いた。浜松ホトニクス社の WSI スキャナー(Nanozoomer)を用いてデジタル化し、病理医が各組織片に対して腫瘍及び非腫瘍の領域を特定して囲み、学習データ

セットを作成して機械学習に提供した。

AI モデルの検証用データとして、上記の学習データとは異なる東京大学医学部附属病院の胃生検連続 999 症例、2685 組織片を用意し、AI に腫瘍・非腫瘍を判定させた。また AI モデルの多施設での検証のため、異なる 10 施設から 2000 症例、3450 組織片の胃生検 WSI を収集し、検証に用いた。

3.2 機械学習

病理デジタル画像(whole slide image, WSI)は超高解像度の画像であり、通常の GPU で処理するには画像を細かいパッチに分割する必要がある。本研究では WSI を 225×225 pixels のパッチに分割し、個々のパッチに腫瘍/非腫瘍のラベルを付与した。畳み込みニューラルネットワークには GoogLe Net (Google Inc.)を採用し、事前に一般画像で学習させた Image Net (Stanford University)を用いてパッチ画像を学習させた。検証データの判定も、個々のパッチごとに行った (Patch-based, PB 法)。

PB 法はシンプルな学習方法であるが、AI にパッチレベルでの学習と判定をさせるだけでは、WSI 全体の分布に関する情報(病理医が低倍率レンズで観察した時に得る情報に相当する)を生かせない問題が考えられた。そのため本事業では、個々のパッチから抽出する特徴量だけでなく、特徴量の元画像における分布も考慮した学習手法(multi-stage semantic segmentation for pathology, MSP 法)を開発し、適用した。¹⁾

PB 法、MSP 法それぞれで作成した AI モデルを、東京大学医学部附属病院(学習データ提供施設と同施設)及びその他の 10 施設の検証データに適用し、腫瘍/非腫瘍を判定させ、感度・特異度・病理医診断との一致率・AUC 値を算出した。

3.3 病理診断ネットワークへの胃生検 AI 実装

JP-AID 事業では、特に病理医が少ないとされている福島県と徳島県において、地域基幹病院(福島県立医科大学病院及び徳島大学病院)と地域の病院とをセキュアな回線をつなぎ、病理画像を送受信して遠隔病理診断を行う地域ネットワークを整備した。この地域ネットワークに上記の AI モデルを実装し、地域の病院から基幹病院に胃生検病理画像を送信する際に AI 判定結果の画像を付与する実装実験を行った。^{2,3)}(病理診断への介入を避けるため、AI 判定結果は診断後に参照した。)

4. 結果

作成した AI モデルで東京大学医学部附属病院の検証データセットを判定させたところ、PB 法では腫瘍/非腫瘍の判定において感度 91.6%、特異度 96.0%、病理医診断との一致率 95.4%、ROC 曲線の AUC 値 0.969 となった。一方、MSP 法では感度 92.8%、特異度 94.6%、病理医診断との一致率 94.4%、ROC 曲線の AUC 値 0.975(図1)となり、両者の性能は概ね同等と考えられた。但し MSP 法ではパッチごとの腫瘍/非腫瘍の判定を行う閾値(DegA)が 0.4 の時が感度・特異度が最良となったのに対し、PB 法では一つの組織片内において少数のパッチが偽陽性になってしまうケースが稀でなく、閾値を 0.99 と極めて高い値に設定する必要が生じた。⁴⁾

次に東京大学医学部附属病院以外の 10 施設より収集した胃生検画像を用いて AI モデルを検証した。その結果、AUC 値の平均は PB 法で 0.906 ± 0.045 、MSP 法で 0.969 ± 0.020 となり、MSP 法の方が PB 法よりも多施設における汎用性が高い手法であることが示された($P < 0.001$)(図2)。⁴⁾

開発した AI モデルを用い、地域の遠隔病理診断ネットワー

クで実装実験が行われた。遠隔ネットワークでは依頼元病院より送信された WSI を医師が観察し診断する。WSI 送信時にクラウド上の AI モデルが判定を行い、結果をヒートマップとして WSI に重ね合わせ表示させることができるシステムを構築した。肉芽組織を癌と誤認したり、少量の印環細胞癌を見落とししたりした症例があったものの、90%以上の症例で AI モデルは概ね正しく腫瘍・非腫瘍を判別できていた。

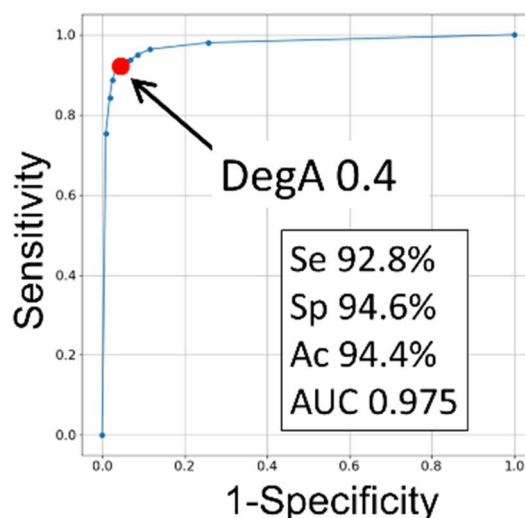


図1 癌・非癌の胃生検検体に対し胃生検 AI モデル (MSP 法)を適用した場合の ROC 曲線
0-1 の範囲で動く cut off 値(DegA)を 0.4 に設定した場合に最良の感度・特異度が得られた

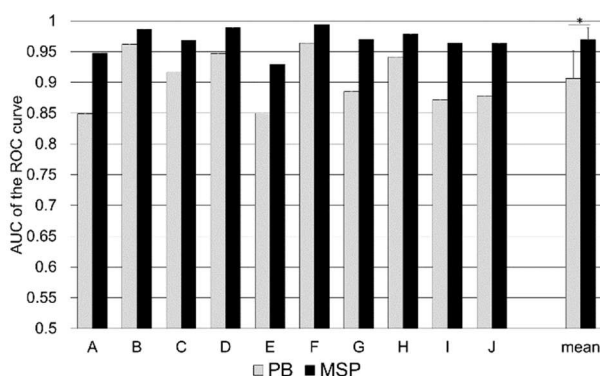


図2 多施設での PB 法、MSP 法の胃生検 AI モデル検証
PB 法より MSP 法の方が 10 施設全てで AUC 値が高く、施設を越えて汎用性の高いモデルであることが示された。

5. 考察

我々は病理専門医によるアノテーション付き画像を用いて胃生検から腫瘍を検出する AI モデルを作成した。従来法 (PB 法)では細かく分割したパッチ単位で腫瘍・非腫瘍を判定するため、一つの組織片の中で少数のパッチが腫瘍と判断されただけで組織片全体が腫瘍と判定され、偽陽性が出やすい傾向があった。我々が新規に開発した MSP 法を用いた場合、パッチ単位の情報だけでなく、パッチの特徴量の元の組織片内での分布も含めた学習と判定を行うことができるため、偽陽性を抑えることに成功し、高い特異度を達成した。特に

多施設での検証において MSP 法は PB 法に比した優位性を示し、優れた手法と考えられた。

今回開発した AI モデルは感度・特異度・病理医との一致率いずれも 90%を超える性能を示している。すなわち、胃生検病理診断のダブルチェックにこの AI モデルを用いた場合、全診断のうち 10%未満の AI と病理医の判定が不一致だった症例に限り、もう一度見直せばよいことになり、実用に耐えるレベルのダブルチェックを提供することができ、一人病理医にとって有用性の高い AI と考えられた。

なお現状の感度・特異度では、病理医を介さない AI 単独で行う病理診断は不可能である。また今後、学習データの増加や AI 技術の発展に伴い、より高い性能の AI が開発されたとしても、AI が単独で病理診断を行うことは不可能と考えている。第一に、病理診断は医行為であり、医師が責任をもって行う必要がある。また、現在の AI は特定の作業(癌の有無の判定等)のみを実施するアルゴリズムだが、実際の病理診断では癌の有無だけでなく、様々な炎症性疾患の評価や、稀な腫瘍の診断も行わなくてはならない。病理医の行う複雑なプロセスである病理診断を、AI が全て代替することは不可能と考えている。⁵⁾病理診断において AI は病理医をサポートする重要な技術と考えているが、AI を病理診断でどのように活用していくかは今後の課題である。日本病理学会は現在、AI ガイドラインの策定を進めており、近日中に公開される予定である。⁶⁾

6. 結論

日本病理学会では病理診断のダブルチェックに利用可能なレベルの胃生検 AI モデルを開発した。他にも 7 課題の AI モデル開発が進行中であり、JP-AID 事業から病理医の日常診断に役立つ様々な AI が誕生することが期待される。

7. 謝辞

JP-AID 事業に対し多大な支援を頂きました日本医療研究開発機構(AMED)の方々、国立情報学研究所(NII)の情報系研究者の皆様、AI モデル開発に参加された先生方、貴重な病理画像を提供して下さいました各施設病理部門の先生方に感謝申し上げます。

参考文献

- 1) Takahama S, Yusuke Kurose, Yusuke Mukuta, et al. Multi-stage pathological image classification using semantic segmentation. Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) 2019; 10702-10711.
- 2) 橋本優子. 福島県での病理診断ネットワーク体制. 病理と臨床 2021;39;775-782.
- 3) 上原久典, 佐竹宣法. 徳島県での病理診断ネットワーク体制. 病理と臨床 2021;39;783-788.
- 4) 阿部浩幸, 黒瀬優介, 原田達也, 牛久哲男, 深山正久. 胃生検組織診断での AI. 病理と臨床 2021;39;789-792.
- 5) 日本病理学会. ステートメント: 人工知能 AI と病理医について: 日本病理学会, 2021. [<https://www.pathology.or.jp/ippan/AI-statement.html>] (cited 2021-Aug-19)]
- 6) 吉澤明彦, 酒井康弘, 白石泰三. 病理 AI ガイドラインについて. 病理と臨床 2021;39;807-813.