CNN による少数教師データからの 早期胃がん領域の検出*

Applying Convolutional Neural Networks for Automatic Detection of Early Gastric Cancer with Limited Endoscopic Images

竹本智子** 堀 圭介*** 坂井良匡† 西村将臣† 池松弘朗††
矢野友規*** 横田秀夫**

Satoko TAKEMOTO, Keisuke HORI, Yoshimasa SAKAI, Masaomi NISHIMURA, Hiroaki IKEMATSU,
Tomonori YANO and Hideo YOKOTA

Key words

endoscopic diagnosis, early gastric cancer, convolutional neural networks, lesion detection, segmentation

1. は じ め に

がんは昭和 56 年以降、一貫して日本人の死因順位の 1 位であり、社会の高齢化にともないその総数は年々増加している¹⁾。一方で、がんの診断・治療技術は年々進歩しており、早期に発見して適切な治療を施すことによって、生存率が大きく上昇する。しかし、胃がんは早期に自覚症状が出にくく、進行しても胃炎や胃潰瘍と症状が似ているため、診断されたときには既かなりの進行ステージであることが多い。胃がんの発見時期はその後の生存率に大きく影響するとされ、発見が早期で病変が限局している状態で治療した場合の 5 年生存率は 95.9% であるのに対し、胃がん診断全体での 5 年生存率は約 65% と著しく低下する²⁾。すなわち、自覚症状が出る前に、検診等によって胃がんを早期に発見することが重要である。

内視鏡診断技術は、胃がんの早期発見に有効な検査手段の一つである。内視鏡とは、先端にカメラがついた細く柔軟なチューブを指し、検診者や患者の口または鼻、肛門から挿入することで、消化管内部を直接観察することができる。図 1 に検査の様子と撮影動画のキャプチャ画像および内視鏡装置例を示す。検診などで多く行われている胃部エックス線検査と比較すると診断精度が高く、被ばくがないメリットがある。一方で、診断精度は医師の経験や能力に大きく依存する問題がある。早期の胃がんは色や形態的特徴に乏しく、消化器内視鏡の専門医であっても発見が難しいためである。

そこで最近、内視鏡診断に機械学習を導入して胃がんを自動検出し、医師の診断を支援しようとする取り組みが始

まっている。世界的な動向としては、大腸内視鏡を対象とした画像診断では既に数多くの研究成果が報告されており⁴⁾⁵⁾、日本でも方法論の基礎的検討のみならず、既に医薬品医療機器等法に基づく承認を取得した臨床現場での実用化まで進められようとしている⁶⁾。ただし、大腸内視鏡で進められている取り組みを、早期胃がんに適用することは容易ではない。機械学習は大量の学習用データを必要とすることが一般的だが、早期胃がんは病変と正常粘膜との境界が不明瞭であり、胃炎などの炎症との区別も難しい。そのため、消化器内視鏡専門医であっても学習用データの作成には多くの時間を要する。また、画像内の病変の有無だけでなく、領域情報を得るためには内視鏡検査後の病理検査との併用が必要になるなど、質の良い学習用データの作成には大変な手間がかかる。少数の学習用データから精度良く早期胃がんを自動検出する試みは、チャレンジングな課題となっている。

本解説では、著者らの研究グループが行った畳み込みニューラルネットワーク⁷⁾ (Convolutional Neural Network: CNN) を用いた早期胃がんの自動検出の取り組みを紹介する。CNN は主に画像認識で高い性能を発揮する機械学

図 1 内視鏡画像検査の様子と内視鏡の例
内視鏡装置はオリンパスホームページ³⁾より転載。画像左の検査画面ははめ込み合成

*原稿受付 令和元年 7 月 1 日

**正 会 員 国立研究開発法人理化学研究所 (埼玉県和光市広沢 2-1)

***国立研究開発法人国立がん研究センター東病院 (千葉県柏市柏の葉 6-5-1)

†国立研究開発法人理化学研究所 (埼玉県和光市広沢 2-1)

††国立研究開発法人国立がん研究センター 先端医療開発センター (千葉県柏市柏の葉 6-5-1)

図2 内視鏡画像と病変領域のアノテーション例

入力画像のサイズは 1408×1056 pixel, 24 ビットフルカラー

習法の一つであり、近年は生物科学や医用画像などさまざまな分野に応用されている。本取り組みは、少ないデータから学習させた CNN が内視鏡画像から精度良く早期胃癌を自動検出できることに加え、画像内の病変の有無だけでなく病変領域を提示できることを目指したフィジビリティスタディである。

2. 早期胃癌領域検出フレームワーク

著者らの研究グループでは、少数の内視鏡画像に病変に関する正解情報を与えたデータを用いて効率的に CNN を学習させることで、早期胃癌領域の検出を可能とする病変自動検出フレームワークを構築した。用いたのは、内視鏡検査で撮像した静止画像に対し、消化管内視鏡科の専門医が早期胃癌領域を手作業でアノテーションした 100 枚の「がん画像」と、画像中にがん病変が存在しない 100 枚の「正常画像」である。図 2 に静止画像とそれに対応する胃癌領域のアノテーション例を示す。研究グループでは、内視鏡的粘膜下層剝離術⁸⁾と呼ばれる内視鏡を使った治療手技によって病変を胃粘膜から切り取り、その後の病理検査により「がん」が確定した領域についてのみ、遡及的に画像を精査し正確なアノテーション情報を獲得している。獲得した計 200 枚の画像に対し、病変領域を含む画像と含まない局所画像をそれぞれランダムに切り出し、約 2 万枚の画像を生成した。その後、画像に移動、回転、拡大、縮小などの操作によるデータ拡張を行うことで、最終的には計約 36 万枚からなる「がん」および「正常」画像データ群を生成した。

次に、画像データ群を用いて CNN を早期胃癌の検出に学習する際、CNN の重みパラメータを一から学習するのではなく、他のデータセットを使用した既存の学習済みモデルを用いる。これは「転移学習」や「ファインチューニング⁹⁾」と呼ばれ、重みパラメータの一部や全体を、現在の学習用データセットを用いて再学習するものである。一から学習するよりも、高速かつ高精度に学習が進む。著者らは ImageNet データベース¹⁰⁾を使用した事前学習済み CNN モデルの一つである GoogLeNet¹¹⁾に対し、上述の約 36 万枚の画像データ群を用いて早期胃癌用の学習済みモデルを生成した。事前学習済みの CNN モデルは

ほかにも多く存在するため、解きたい問題によって適切に選択し、場合によってはネットワークそのものをカスタマイズすることも必要である。

著者らの学習済みモデルは、新たな内視鏡画像の入力に対して「がん」らしさを表す尤度を予測する。このとき、著者らは元画像をそのまま CNN への入力とするのではなく、画像をブロック分割し、ブロックごとに求めた尤度を入力画像に反映した。この理由は、ブロックごとの予測結果を入力画像に反映させることで、画像中の病変の有無だけでなく、病変領域が存在する領域をヒートマップとして提示できることにある。以上に基づく早期胃癌の自動検出フレームワークを図 3 に示す。

3. 精度検証と検出結果

著者らが対象とした早期胃癌は、日本胃癌学会が定めた肉眼型分類¹²⁾のうち、明らかな腫瘤状の隆起が認められる隆起型 (Type 0-I)、明らかな隆起や陥凹は認められないが低い隆起が認められる表面隆起型 (Type 0-II a)、わずかに粘膜の陥凹が認められる表面陥凹型 (Type 0-II c) の三つである。特に Type 0-II c は、早期胃癌の中でも発見が難しいとされる。図 4 に対象とした胃癌の肉眼型分類と、対応する内視鏡画像の例を示す。これらの 3 タイプの早期胃癌からなる約 36 万枚の画像を用いた学習にかかる時間は、学習回数にもよるが、高性能 GPU 搭載のマシン (Intel Xeon (2.4 GHz, 128 GB memory), NVIDIA GeForce GTX 1080 (8 GB×2 memory)) で約 10 時間であった。

学習済みモデルが新たな入力に対して正しい推論ができるかどうかは、注意深く検証する必要がある。ここでは、著者らが行った精度検証について、その一部を紹介する。まずは、前章で述べた方法と同様に、内視鏡画像からランダムに切り出した「がん」、「正常」の局所画像を各々約 5000 枚用意し、それらを学習済みモデルに入力し、出力した尤度から求められる入力画像のクラス (ここでは、「がん」か「正常」) を正解情報と照合した。表 1 は、画像のクラス判定について真陽性 (True Positive)、疑陽性 (False Positive)、真陰性 (True Negative)、疑陰性 (False Negative) の画像数を示している。それらの結果から感度 (Sensitivity)、特異度 (Specificity)、陽性的中率 (PPV)、陰性的中率 (NPV)、精度 (Accuracy) を求めたところ、すべてにおいて 80% 以上の精度を得た。

また、図 3 のフレームワークに従って求めた領域検出の結果に対し、前章で述べたブロックごとに正解情報との一致を検証した。検証に使用した画像は、学習に用いていない内視鏡画像約 800 枚である。結果として、すべての検証画像中に存在する「がん」ブロックのうちの 73.3%、「正常」ブロックのうちの 98.8% について、正しく推論が行われていた。図 5 に肉眼型分類のタイプ別の検出成功例を示す。図中の正解領域は、専門医が手作業で病変領域を示したものである。ここに示した例だけでなく、多くの入力

図 3 早期胃がん領域の自動検出フレームワーク

図 4 早期胃がんの肉眼型分類

表 1 学習済みモデルの精度検証

		Prediction		
		Positive	Negative	
Ground truth	Positive (cancer)	TP 3,723	FN 930	Sensitivity 80.0%
	Negative (normal)	FP 262	TN 4,735	Specificity 94.8%
		PPV 93.4%	NPV 83.6%	Accuracy 87.6%

画像に対して、病変の大きさや形状によらず、専門医が示した病変領域とよく一致した領域を自動検出できていた。

一方、図 6 (a) に示すように、内視鏡装置付近に存在する病変を見落とすなど、検出に失敗することもある。原因の一つとしては、内視鏡装置の影などにより病変領域が周囲と比較して低輝度であることなどが考えられる。また、図 6 (b) の例では正常領域をがんの見誤ってしまう過検出が生じている。専門医によると、病変領域の周囲に炎症が生じていたことで、正常領域であってもわずかに表面が隆起しているとのことから、CNN はこのわずかな隆起をがんの見誤って判断した可能性がある。著者らは、これらの見落としや過検出を減らす方法として、学習モデルそのものの見直しだけでなく、病変が他と比べて低輝度であるケースや、胃の炎症が多数存在するケースなど、学習用データに取って間違えやすい画像を入れていくことなどによって、予測精度の向上を図っている。

(a) Type 0-I に対する検出結果

(b) Type 0-IIa に対する検出結果

(c) Type 0-IIc に対する検出結果

図 5 タイプ別の早期胃がんの自動検出例
入力画像 (左)、正解画像 (中央)、自動検出結果 (右)

4. 今後の展望

本解説では、著者らの研究グループが取り組んでいる早期胃がん領域の自動検出を紹介した。自動検出の実現は、将来的には医師の診断を支援し、診断精度向上の可能性や、医師の診断能の個人差をフォローできる可能性がある。ただし、今回紹介した方法はフィジビリティスタディであり、少数の学習用データからでも早期胃がん領域の自動検出が可能であることを示せたまでである。将来の臨床応用に向けて検討しなくてはならない点はまだ多く、例

(a) 未検出の例

(b) 過検出の例

図6 誤検出が生じた例

例えば自動検出が目指すべき精度を、臨床現場で求められるニーズや適用範囲の観点から検討することもそれに含まれる。著者らが用いているCNNについても、ネットワークの構造や学習方法など、高精度化の観点から検討しなくてはならない点が多々ある。また本来、内視鏡検査ではリアルタイム検出が望まれていることから、高速化も重要課題である。以上のように問題山積であるが、悲観的な要素ばかりではない。本稿で紹介した画像例には、専門医にとっても病変の検出が容易ではないものが多数含まれていた。それにもかかわらずCNNは病変の有無だけでなく、存在領域まで精度良く検出している。このことは、熟練の専門医が観察力と経験によって捉える病変のわずかな画像特徴を、今回紹介した簡便な方法であっても、高速かつ的確に捉えられた結果であると推測できる。今後、各方面の専門家らの積極的な参入によって、臨床現場で医師の診断を支援する日が近くなることを期待したい。

参考文献

- 1) 厚生労働省：平成29年人口動態統計
- 2) 国立がん研究センター：全国がん罹患モニタリング集計
- 3) https://www.olympus.co.jp/technology/technology/luceraelite/?page=technology_technology
- 4) G. Urban, P. Tripathi, T. Alkayali, M. Mittal, F. Jalali, W. Karnes and P. Baldi: Deep Learning Localizes and Identifies Polyps in Real Time with 96% Accuracy in Screening Colonoscopy, *Gastroenterology*, **155**, 4 (2018) 1069.
- 5) N. Tajbakhsh, S.R. Gurudu, J. Liang: Automated Polyp Detection in Colonoscopy Videos Using Shape and Context Information, *IEEE Trans. on Medical Imaging*, **35**, 2 (2016) 630.
- 6) 国立研究開発法人日本医療研究開発機構プレスリリース（平成30年12月10日）。
- 7) Y. LeCun, L. Bottou, Y. Bengio and P. Haffner: Gradient-Based Learning Applied to Document Recognition, *Proc. of the IEEE*, **86**, 11 (1998) 2278.
- 8) T. Gotoda, H. Yamamoto and R.M. Soetikno: Endoscopic submucosal dissection of early gastric cancer, *J. Gastroenterol.*, **41**, 10 (2006) 929.
- 9) N. Tajbakhsh, J.Y. Shin; S.R. Gurudu, R.T. Hurst, C.B. Kendall, M.B. Gotway and J. Liang: Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning? *IEEE Trans. Medical Imaging*, **35**, 5 (2016) 1299.

- 10) <http://www.image-net.org/>
- 11) W.L. Szegedy, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke and A. Rabinovich: Going deeper with convolutions. In *Proc. of CVPR*, (2015) 1.
- 12) 日本胃癌学会（編）：胃癌取り扱い規約，14版，金原出版，(2010)。

竹本智子

2006年埼玉大学大学院理工学研究科博士課程修了（博士（工学））。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 画像情報処理研究チーム 研究員。主に生物科学データの画像処理、画像解析に従事。

堀 圭介

国立研究開発法人国立がん研究センター東病院消化管内視鏡科医員。2012年学位（医学博士）取得（岡山大学）。専門領域は消化器内視鏡、臨床腫瘍学。

坂井良匡

2019年電気通信大学大学院情報システム学研究科博士前期課程修了（修士（工学））。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 画像情報処理研究チーム テクニカルスタッフ。機械学習を用いた画像解析などに従事。

西村将臣

2004年埼玉大学大学院理工学研究科博士前期課程修了（修士（工学））。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画像情報処理研究チーム テクニカルスタッフ。三次元や四次元の観察画像を対象にした画像解析などに従事。

池松弘朗

国立研究開発法人国立がん研究センター 先端医療開発センター内視鏡機器開発分野分科長、兼国立がん研究センター東病院 消化管内視鏡科医長。2013年学位（医学博士）取得（獨協医科大学大学院）。専門領域は消化器内視鏡。

矢野友規

国立研究開発法人国立がん研究センター東病院消化管内視鏡科科長、兼内視鏡センター長。2017年学位（医学博士）取得（京都大学）。専門領域は消化器内視鏡、臨床腫瘍学。

横田秀夫

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 画像情報処理研究チーム チームリーダー。1999年学位（博士（工学））取得（東京大学）。計測装置開発、バイオシミュレーション、画像解析などに従事。