

修士学位論文要約（令和 4 年 3 月）

深層学習による胃 X 線画像を用いたピロリ菌感染鑑別に関する研究

田口 佳那

指導教員：杉田 典大

A Study on Differentiation of *H. pylori* Infection Using Gastric X-ray Images by Deep Learning

Kana TAGUCHI

Supervisor: Norihiro SUGITA

In this study, we developed a deep learning-based method for diagnosis of *H. pylori* infection using gastric X-ray images. To quantitatively evaluate the diagnosis performance, a comparison with radiological technologists was conducted. We used two CNNs, VGG16 and Inception_v3, and transfer learning to train and test the deep learning models. We then calculated the Receiver Operating Characteristic curve and the Area Under the Curve (AUC). The resulting AUC values were 0.87 for VGG16 and 0.94 for Inception_v3. We used Grad-CAM to visualize the classification basis of the CNN. The visualization results were generally consistent with the areas that doctors' interpretation guideline focuses on. We used Intersection over Union (IoU) and hit rate as indices for quantitative evaluation. The IoUs and hit rates between the visualization results and the technologists' averaging area were lower than those between the technologists. This implies that clinical basis of CNNs can still be improved further.

1. はじめに

ピロリ菌は胃の粘膜に棲みつく細菌であり除菌しない限り存在し続ける。感染すると胃炎が起り胃がんのリスクを高める原因となる。血液検査の精度は 96.0%¹⁾と高いものの、自主的に検査を受けない限り感染の有無を知ることはできない。ピロリ菌除菌後は血液検査で偽陽性となる場合があるが画像検査は除菌後でも正確な検査が可能であり、胃 X 線画像からの診断が求められる。また現在胃 X 線検査は日常診察で行われなくなり読影医の不足が深刻化している。よってコンピュータ支援診断 (computer-aided diagnosis: CAD) システムの必要性が高まっている。先行研究¹²⁾では画像中の Region of interest (ROI) を抜き出したり細かいパッチ状にしたりするなどの処理を加えた画像を入力している。医療現場での ROI の指定は医師の手間の増加に繋がる。また、これらのシステムはブラックボックスであり、臨床で信頼できるか判断できないことに加え、誤診や誤判断に対する説明ができない問題がある。そこで本研究では、胃 X 線画像全体を入力とするピロリ菌感染診断 CAD システムを考案し、Gradient weighted Class Activation Mapping (Grad-CAM)³⁾による診断根拠提示及び提示された根拠の定量的な評価について検討する。

2. ピロリ菌感染の鑑別

畳み込み層とプーリング層を備えた畳み込みニューラルネットワーク (Convolutional Neural Network, CNN) は画像処理に適しており、本研究では代表的な CNN

である VGG16⁴⁾と Inception_v3⁵⁾を採用し、転移学習を行った。画像分類の標準的データセットである ImageNet⁶⁾により事前学習したモデルを用い、学習済みの重みを固定して、新たに出力を感染と非感染の 2 つにするためのネットワークを最終層として追加した。精度向上のため、学習データに対してランダムな幾何学処理を行ってデータを拡張した。1 症例につき撮影体位の異なる 8 枚の画像で構成された、宮城県対がん協会から提供された画素数 1024×1024 の胃 X 線画像を、256×256 にリサイズして使用した。陽性/陰性は各 100 症例で、学習用 70 症例、検証用 30 症例に分割し、データ拡張により枚数を 5 倍に増加させた。評価には Receiver Operating Characteristic (ROC) 曲線及び曲線下面積 (Area Under the Curve, AUC) を使用した。ROC 曲線は基準となる確率以上のものを陽性とみなすことでモデルの精度を把握する曲線である。縦軸は

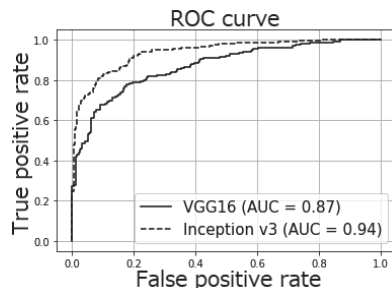


図 1 胃 X 線画像を用いたピロリ菌感染鑑別の ROC 曲線

真陽性率、横軸は偽陽性率である。AUC は、最大値の1に近いほど良い。AUC の値は VGG16 で 0.87、Inception_v3 で 0.94 であった(図1)。

3. 診断根拠の可視化による妥当性検証

Grad-CAM は CNN の特徴量マップから出力値間の逆伝播時の勾配を利用して、入力画像中で着目している部分を判断根拠として可視化する手法の一つである。図2(右)は、分類への影響度に応じて色分けした heatmap であるが、印刷の都合上とくに影響が大きい部分の輪郭のみ示している。

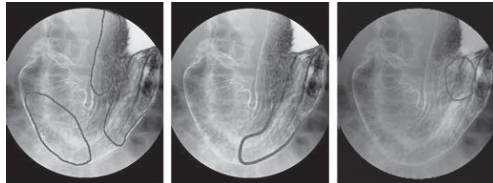


図2 診療放射線技師 A の着目部分(左)と診療放射線技師 B の着目部分(中央)、Grad-CAM 適用による heatmap 画像(右)

医療画像の専門職である診療放射線技師2名が胃X線画像診断時に着目する部分の輪郭を作成し、Grad-CAM により可視化されたシステムの着目領域と比較した。図2に比較例を示す。ピロリ感染の読影指針⁷⁾では、粘膜部分やひだの状態に着目しており、可視化結果の heatmap はX線画像を解釈する際に重視する部位と概ね一致したと言える。定量的評価の指標として、Intersection over Union (IoU) と hit 率を用いた。IoU とは領域の和集合に対して2つの領域がどれくらい重なっているかを表す指標のことである。hit 率は領域が互いに重なっている場合を hit とし、重なっていない場合は Not hit とし、画像ごとに計算する数値である。診療放射線技師の着目領域と Grad-CAM の結果で計算した IoU と hit 率の比較を、感染例と非感染例に分けて、それぞれ表1、表2に示す。

表1. 感染例の IoU と hit 率の比較

	IoU 平均値	IoU 最小値	IoU 平均値	hit 率
診療放射線技師①と診療放射線技師②	0.38	0.00012	0.74	100%
診療放射線技師①と Grad-CAM 可視化結果	0.021	0	0.089	50%
診療放射線技師②と Grad-CAM 可視化結果	0.016	0	0.072	50%

表2. 非感染例の IoU と hit 率の比較

	IoU 平均値	IoU 最小値	IoU 平均値	hit 率
診療放射線技師①と診療放射線技師②	0.41	0.02	0.72	100%
診療放射線技師①と Grad-CAM 可視化結果	0.12	0	0.23	90%
診療放射線技師②と Grad-CAM 可視化結果	0.15	0	0.32	90%

感染例の場合も非感染例の場合も、IoU の平均値は技師間で40%前後である。したがって、専門的な知識を持った技師が診断をした場合でも、その根拠となる着目領域を定量的に評価するとずれが生じていることが分かる。hit 率に関しては感

染例でも非感染例でも100%であり、同様な領域に着目していることが示された。一方で、技師と Grad-CAM 可視化結果の数値は非感染例で90%と高い結果となったことから、着目している部分が完全に外れているのではなく、一致している部分はあるものの重なっている領域が狭いと言える。感染例の技師と Grad-CAM 可視化結果の IoU 平均値は、非感染例の場合と比較して低い数値となった。hit 率に関しても非感染例では90%であるのに対して感染例では50%にとどまったことから、感染例の方が技師の着目部分と一致していない場合が多いと言える。

4. まとめ

本稿では胃X線画像全体からピロリ菌感染の診断を行うCADシステムを考案し、システムの着目箇所を Grad-CAM を用いて可視化した。また、可視化結果を IoU と hit 率という指標を用いて診療放射線技師の着目領域と比較し、その妥当性を検証した。

文献

- 1) Togo, Ren, et al. "Detection of gastritis by a deep convolutional neural network from double-contrast upper gastrointestinal barium X-ray radiography." *Journal of gastroenterology* 54.4 (2019): 321-329.
- 2) 重松綾, 中島寛隆, and 八巻悟郎. "Deep Learning と胃 X 線画像を用いた Helicobacter pylori 感染診断." *日本消化器がん検診学会雑誌* 57.5 (2019): 687-694.
- 3) Selvaraju, Ramprasaath R., et al. "Grad cam: Visual explanations from deep networks via gradient-based localization." *Proceedings of the IEEE international conference on computer vision*. 2017.
- 4) Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large scale image recognition." *arXiv preprint arXiv:1409.1556* (2014).
- 5) Szegedy, Christian, et al. "Rethinking the inception architecture for computer vision." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
- 6) Deng, Jia, et al. "Imagenet: A large-scale hierarchical image database." *2009 IEEE conference on computer vision and pattern recognition*. Ieee, 2009.
- 7) Ishihara, Kenta, Takahiro Ogawa, and Miki Haseyama. "Helicobacter Pylori infection detection from gastric X-ray images based on feature fusion and decision fusion." *Computers in biology and medicine* 84 (2017): 69-78.