

原 著

AIは何をみて大腸 pT1b 癌を診断しているか？： Class Activation Mapping からみた検討

中島勇貴¹⁾ 根本大樹²⁾ 勝木伸一³⁾ 林 芳和⁴⁾ 愛澤正人²⁾
 歌野健一²⁾ 竹澤敬人⁴⁾ 相良裕一⁴⁾ 朱 欣⁵⁾ 澁川悟朗¹⁾
 山本博徳⁴⁾ 富樫一智²⁾

- 1) 福島県立医科大学会津医療センター 消化器内科,
- 2) 同 小腸・大腸・肛門科,
- 3) 小樽掖済会病院 消化器病センター,
- 4) 自治医科大学 消化器内科,
- 5) 会津大学 コンピュータ理工学部

要 旨

【目的】最近, Class Activation Mapping (CAM) という手法により人工知能 (AI) の関心領域 (ROI) を可視化できるようになった. 本研究では AI の大腸 T1b 癌診断における ROI の特徴を明らかとすることを目的とした.

【方法】大腸 pT1b 癌の通常内視鏡画像 114 病変 226 枚を使用した. CAM により, AI の ROI は赤色領域として示され, 大腸内視鏡医の ROI と比較した. 評価項目は, 内視鏡医との ROI 一致性, AI の ROI の性状とした. ROI 一致性の評価は, excellent : 75%以上が一致, poor : 25%未満, good : その中間とした.

【結果】全画像で CAM 画像が作成可能だった. ROI 一致性は, excellent 39%, good 34%, poor 27%であった. poor 例を除いた AI の ROI は, 多くが発赤部位であり (91%), 明らかに血液が付着した領域 (21%) やヒダ集中 (34%) は少なく, 隆起 (57%) や陥凹 (39%) が多かった.

【結論】ROI が一致する場合には内視鏡医との類似性がみられた. アノテーションにより ROI の不一致を克服することが, AI の学習効率の改善につながる可能性がある.

Key words 人工知能／早期大腸癌／大腸内視鏡検査／大腸内視鏡診断／関心領域／
Class Activation Mapping／読影試験

I 緒 言

近年, 治療内視鏡の進歩により大腸 cT1 癌に対する内視鏡治療の適応が広がっている^{1)~6)}. しかし

し, 粘膜下層に 1 mm 以上浸潤する大腸 T1b 癌はリンパ節転移のリスク因子となることから深達度診断の重要性は高い. したがって cT1 癌が疑われる場合は術前の拡大内視鏡観察や超音波内視鏡で適応を判断したうえで^{7)~12)}, 病変を一括切除し, 詳細な病理学的評価を行うことが求められる. このような病変は内視鏡的粘膜下層剝離術 (endoscopic submucosal dissection : ESD) の良い適応である. しかし, cT1b 癌と内視鏡切除で根治が望める cT1a 癌, cTis 癌は内視鏡所見が似ていることから, 診断が困難となることも少なくなく^{13), 14)}, 不十分な診断から分割切除が施行され, 正確な病理診断が損なわれる危険性も否めない. 日常内視

Gastroenterological Endoscopy 2021 ; 63 : 1232-40.

Yuki NAKAJIMA

Which Region Does Artificial Intelligence Look at to Predict T1b Colorectal Cancer? Analysis based on Class Activation Mapping.

別刷請求先 : 〒 969-3492 福島県会津若松市河東町谷沢字前田 21-2
 福島県立医科大学会津医療センター 消化器内科 中島勇貴

Table 1 CAM 画像を作成した pT1b 癌の特徴.

		病変, n (%)	画像, n (%)
大きさ	≤ 10mm	45 (39)	89 (39)
	11-20mm	53 (47)	105 (47)
	≥ 21mm	16 (14)	32 (14)
形態	0-I s	69 (61)	137 (61)
	0-II a	39 (34)	77 (34)
	0-II c	6 (5)	12 (5)
部位	近位結腸	42 (37)	84 (37)
	遠位結腸	37 (32)	72 (32)
	直腸	35 (31)	70 (31)
画像	近景	-	140 (62)
	遠景	-	86 (38)

CAM : class activation mapping¹¹⁾.

近位側結腸 : 盲腸から横行結腸.

遠位側結腸 : 下行結腸から S 状結腸.

鏡診療の中でこのような深達度診断が重要となる早期癌を通常内視鏡画像から簡便に診断し, 深達度を推測できる技術があればその有用性は高い.

人工知能 (AI) を内視鏡的深達度診断に応用する研究も始まってはいるが, pT1b 癌の AI 診断の報告は未だ少ない現状にある^{15)~17)}. 最近, われわれが行った先行研究¹⁸⁾において, 通常内視鏡画像を用いた AI による大腸 pT1b 癌の診断精度は比較的高い結果が得られた (感度 80%, 特異度 87%, 正診率 84%). 更なる診断精度の向上が求められるが, AI が学習する過程はブラックボックスとされ, 少なからず精度向上の妨げとなっていた. 近年, Class Activation Mapping (CAM) という手法の開発により, AI の画像認識モデルにおいて画像認識に寄与したとされる領域を可視化することが可能となった¹⁹⁾. これは AI が画像を認識する過程で, 診断の根拠として用いた部分を, その重要度に応じて色分けして出力する手法である. すなわち CAM を用いることで, AI が何をみて診断しているかを認識できるようになった. CAM は元来, AI が効率良く学習するために開発されたものであるが, CAM を用いて AI の関心領域 (Region of Interest, ROI) を知ることにより, 内視鏡医の診断能力の向上につながる可能性も考えられる. しかし, 著者らの知る限り, CAM を用いた ROI の報告は英文誌を含めてもない. 本研究では, AI の ROI の特徴を明らかとすることを主たる目的とし, 内視鏡専門医との類似性についても検討を加

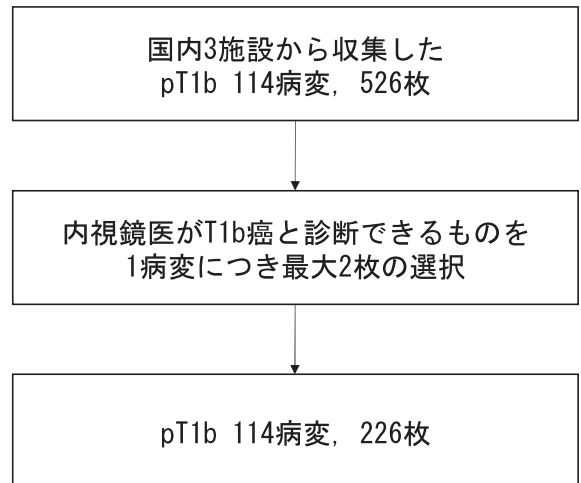


Figure 1 CAM 画像作成に使用した内視鏡画像の選定.

えた.

II 方法

研究デザイン

本研究は, 大腸 pT1b 癌の AI 診断に関する post hoc 研究である. CAM 画像を作成することにより, AI の内視鏡診断における ROI を明らかにすることを目的とした. なお, 先行研究 (福島県立医科大学倫理委員会登録番号 2952)¹⁸⁾ では, 福島県立医科大学会津医療センター及び自治医科大学から収集した内視鏡画像を用いて, AI の学習を行い, 小樽掖済会病院の内視鏡画像により AI の診断能を評価した.

使用した内視鏡画像

CAM を作成する内視鏡画像は, 先行研究のために国内 3 施設から集積したものをを用いた. 教師データ及び試験データとして用いた内視鏡画像のうち, 組織学的に T1b 癌と判断された 114 病変 (0-I s 69, 0-II a 39, 0-II c 6; 最大径 16.5 ± 13.4 mm) の内視鏡画像 226 枚 (近景 140 枚, 遠景 86 枚) を使用した (Table 1). それぞれの画像において T1b 癌と診断できるものに限定し, 画質も考慮して 1 病変につき最大 2 枚の内視鏡画像を選択した (Figure 1). 最終的に, 1 枚選択したものが 2 病変, 2 枚選択したものが 114 病変であり, 2 枚選択した場合はそれぞれ別病変として取り扱った.

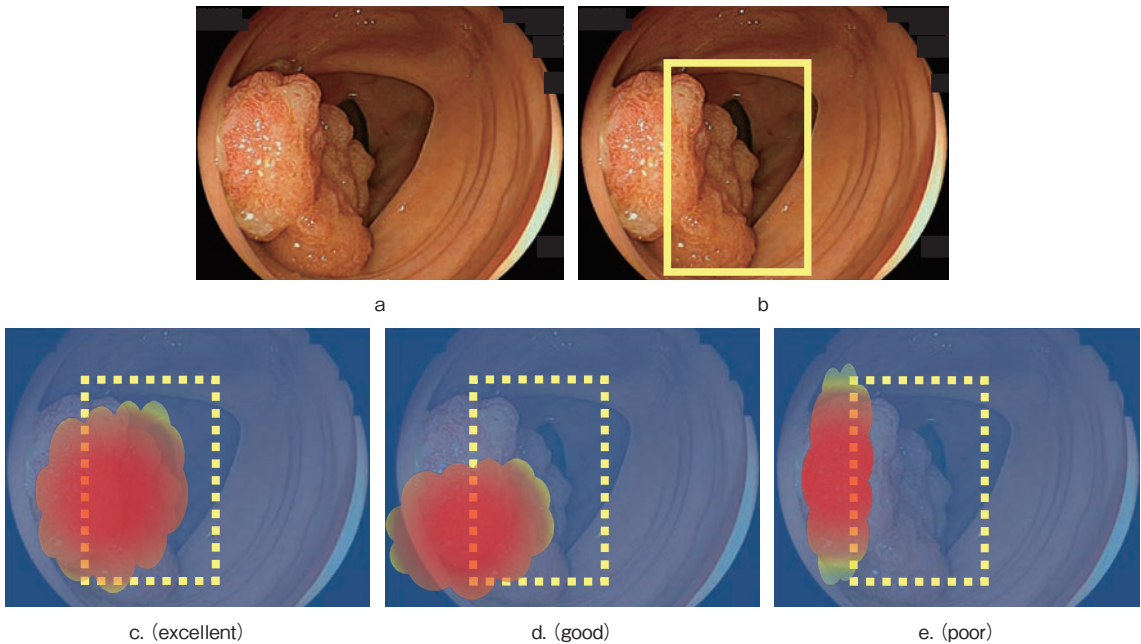


Figure 2 ROI 一致性的評価基準.

a : CAM 画像作成に用いる内視鏡画像.

b : a における内視鏡医の ROI (黄色四角).

c~e : CAM 画像. 赤い領域が AI の ROI. 黄色点線の四角が内視鏡医の ROI.

c : excellent. AI と内視鏡医の ROI 一致性が 75% 以上.

d : good. AI と内視鏡医の ROI 一致性が 25-74%.

e : poor. AI と内視鏡医の ROI 一致性が 25% 未満.

CAM 画像の作成

CAM 画像の作成は、会津大学コンピュータ理工学部の共同研究グループが行った。先行研究¹³⁾において AI の学習には ResNet50²⁰⁾ を使用した。ResNet50 とは、画像認識に用いられる畳み込みニューラルネットワークの一つである。ResNet50 を用いた画像認識の過程で重み付けられた領域をヒートマップ化することで CAM 画像を作成した。作成された CAM 画像では AI の ROI は赤色領域として示された。

内視鏡医による ROI の評価

内視鏡医の ROI は、従来から報告されている陥凹や表面不整、潰瘍、びらん、血液付着、ヒダ集中などの cT1b と診断可能な内視鏡所見⁹⁾ を有する領域を画像毎に大腸内視鏡専門医 2 名が合議制で、AI の ROI 結果を blind として決定した。なお、大腸内視鏡専門医 2 名 (DN, YN) はそれぞれ 12 年目、8 年目の消化器病学会/消化器内視鏡学会の専門医であるばかりでなく、日常的に大腸内視

鏡検査を行い、毎週行われる大腸内視鏡検査に関する院内グループカンファランスに参加し、少なくとも 500 例以上の早期大腸癌に対する内視鏡診断の経験を有する。

AI と内視鏡医の比較

評価項目は、内視鏡医との ROI 一致性、色調、凹凸、血液の付着、ヒダ集中とした。ROI 一致性の評価は、excellent : AI の ROI の 75% 以上が内視鏡医の ROI に含まれる、poor : 25% 未満が含まれる、good : その中間とした (Figure 2)。本研究では、上記の記述統計を主たる目的とし、統計学的検定は行っていない。

III 結 果

226 枚の全画像で CAM 画像が作成可能だった。CAM 画像から得られた AI と内視鏡医との ROI 一致性は、excellent 27%、good 34%、poor 40% であり、excellent と good を合わせた一致例は 60% 程度であった (Table 2)。病変の内視鏡画像を、

Table 2 AI と内視鏡医間 ROI の一致性.

		excellent	good	poor
全体	n (%)	58 (26)	77 (34)	91 (40)
大きさ	≤ 10mm, n (%)	30 (34)	28 (31)	31 (35)
	11-20mm, n (%)	22 (21)	35 (33)	48 (46)
	≥ 21mm, n (%)	6 (19)	14 (44)	12 (37)
形態	0-I s, n (%)	32 (23)	56 (41)	49 (36)
	0-II a, n (%)	22 (29)	20 (26)	35 (45)
	0-II c, n (%)	4 (33)	1 (9)	7 (58)
部位	近位結腸, n (%)	28 (33)	20 (24)	36 (43)
	遠位結腸, n (%)	13 (18)	32 (44)	27 (38)
	直腸, n (%)	17 (24)	25 (36)	28 (40)
画像	近景, n (%)	25 (18)	55 (39)	60 (43)
	遠景, n (%)	33 (38)	22 (26)	31 (36)

AI : artificial intelligence ; ROI : region of interest.

excellent : AI と内視鏡医の ROI 一致性が 75% 以上.

good : AI と内視鏡医の ROI 一致性が 25-74%.

poor : AI と内視鏡医の ROI 一致性が 25% 未満.

近位結腸 : 盲腸から横行結腸.

遠位結腸 : 下行結腸から S 状結腸.

Table 3 AI と内視鏡医で一致した ROI* の特徴.

色調, n (%)	赤色調	125 (93)
	非赤色調	10 (7)
凹凸, n (%)	隆起	75 (56)
	平坦	7 (5)
	陥凹	53 (39)
血液付着, n (%)	あり	29 (21)
	なし	106 (79)
ヒダ集中, n (%)	あり	51 (38)
	なし	84 (62)

*excellent と good であった 135 病変を対象とした.

ROI : region of interest.

1 ヒダ以内の近接像を近景画像, 2 ヒダ以上離れた画像を遠景画像と定義して検討すると, 遠景画像で excellent が増加する傾向がみられた. しかし, 病変の大きさ・形態・部位別の解析では特徴的な所見は認められなかった. poor 例を除いた AI の ROI は, 多くが赤色調を呈したが, 明らかに血液が付着した領域は少なかった (Table 3). 凹凸は, 隆起 56%, 陥凹 39%, 平坦 5% であった. ヒダ集中は 34% に認められた.

poor 例では ROI が病変外に存在するものは 60 例 (66%) があった (Table 4). 一方, 内視鏡医との ROI の一致性が評価基準を満たさないが ROI が病変内に存在するものや, 内視鏡医の ROI

Table 4 不一致 (poor) 例における AI の ROI : 病変との関係.

病変内, n (%)	31 (34)
病変外, n (%)	60 (66)

AI : artificial intelligence ; ROI : region of interest.

内に存在するが 25% の一致を満たさないものが 31 例 (34%) あった.

症例を 2 例提示する. 症例 1 は, 径 10mm, 肉眼型 0-I s の pT1b 病変である. 発赤が目立ち, 形態はやや分葉状であるが, 隆起部の陥凹は目立たず, ヒダ集中は認めない. AI の ROI は, 発赤調, 隆起部であった (Figure 3). 症例 2 は, 径 10mm, 肉眼型 0-I s の pT1b 病変である. 発赤調で頂部の陥凹が目立つ. AI の ROI は, 発赤調, 陥凹を含めた隆起部であった (Figure 4).

IV 考 察

近年, AI による内視鏡診断の研究が進み, 高い診断能力が報告されている. 今後, 実用化に向けて更なる診断精度の向上が求められるが, AI の学習や診断の過程はブラックボックスとされており, 学習過程の最適化を図ることは困難であった. しかし, 今回用いた CAM は, AI の画像認識モデルにおいて, 認識に寄与したとされる領域を可視化



Figure 3 症例 1.

径 10mm の 0-I s. 3-a は元画像. 3-b は CAM 画像. AI の ROI は, 発赤調, 隆起部である.

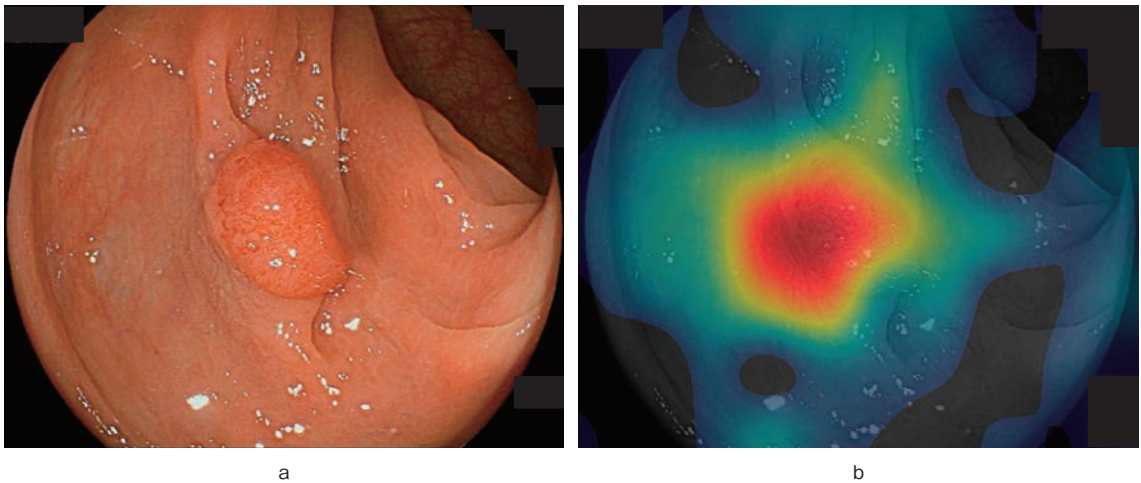


Figure 4 症例 2.

径 10mm の 0-I s. 4-a は元画像. 4-b は CAM 画像. AI の ROI は, 発赤調, 陥凹を含めた隆起部が含まれている.

できるため, AI が診断に用いた ROI を知ることが可能であり, AI の診断精度の向上の一助となる可能性がある.

われわれは pT1b 癌の内視鏡画像から CAM 画像を作成することが可能であった. 今回 CAM 画像作成には, 今回先行研究で用いた白色光非拡大画像のみを使用しており, 特殊光や拡大観察画像は用いていない. この理由を説明する. 画像強調内視鏡や拡大内視鏡での観察や超音波内視鏡による深達度診断法は, きわめて有用であることは誰もが認めるところであるが, 正確に診断するため

には専門的な知識・技術を要し, この国際化が進んでいない現状がある. 今後, グローバルな早期大腸癌診療の展望を考えたときに, 内視鏡医なら誰もが用いる白色光非拡大画像での AI 診断が可能となれば, 国際的なインパクトは計り知れないと著者らは考えているので, 敢えて白色光非拡大画像のみを使用した.

本研究の結果から, AI と内視鏡医の ROI 一致性は excellent と good を合わせて 61% と一致しているものが多く, 一致している場合は, 隆起部や陥凹部を ROI としており, 内視鏡医の ROI と酷似

していた。一方、poor 例が 39% と必ずしも内視鏡医と一致していない症例もあった。これらは異なる ROI により大腸 pT1 癌を診断している可能性が考えられた。

近年の AI 技術の進歩は目覚ましいものがあり、日常生活で実用化されているものも少なくない。多くの場合、AI がいかに人間の行う動作や判断に近づき、ヒューマンエラーを防止できるかを目的としており、AI を用いた内視鏡診断も同様の理念で研究が進んでいる。既報では、AI を用いた大腸 pT1b 癌診断の正診率は 81.2% と報告されており¹⁰⁾、内視鏡医の正診率⁹⁾ には及んでいないのが現状である。そのため、まずは内視鏡医と同等の診断能力に近づくために、AI の ROI を内視鏡医に近づけることは有効な方法論と考えられる。今回の結果で、ROI 不一致例が約 40% を占めていたことから、AI の学習段階で pT1b と診断可能な内視鏡医の ROI を学習させることができれば、AI の診断精度向上につながる可能性がある。そのため、内視鏡医の ROI を四角などで囲ってタグ付けするアノテーションという作業を加えることにより、AI の学習効率が改善し、診断精度が向上することが期待される。

一方、ROI 不一致例については、内視鏡医と異なる ROI で T1b 癌を診断している可能性があり、その ROI を診断価値のないものと位置付けるのは時期尚早である。特に ROI 不一致例の 66% が病変外を ROI としていたことから、病変外の何らかの要素を診断根拠としていた可能性も考えられる (Table 4)。今後そのような観点から、現在の内視鏡医の T1b 診断における ROI 以外に精度の高い診断基準が明らかとなる可能性もあり、その所見が今回の ROI 不一致例に含まれているかもしれない。その場合は、内視鏡医の ROI をアノテーションするのではなく、より多くの画像を教師データとして用いることが診断精度向上に有用となるだろう。現時点で、アノテーションと教師データ数の増量による学習効率の改善、診断精度上昇などの比較はできておらず、今後の研究課題としたい。

本研究の強みは、多施設共同研究であるため、比較的多く症例を集積し、解析できたことである。本研究では 114 例の T1b 癌を集積できたことが、今回のような一定の結論につながったと考えられ

る。一方、本研究にはいくつかの問題も潜んでいる。第 1 に、画像選択の段階で画質の低いものを除外した対象としたために、研究結果は限定的である可能性がある。第 2 には、ROI の性状を 2 名の大腸内視鏡専門医による合議制で評価していることである。そのため、評価の妥当性を客観的に検証することができておらず、将来的に多くの専門医による再検証が必要になるのかもしれない。第 3 には、ROI が一致していない症例が 1/3 程度存在し、この原因が完全には解明できていない点がある。今後、CAM に熟知した工学者との更なる議論が必要と考えられる。第 4 には、本研究において組織型や深達度別の解析ができていないことである。特に深達度については、今回の研究に用いられた内視鏡画像のほとんどが内視鏡治療適応となった病変であったため、cTis や cT1a の診断で内視鏡治療が行われた病変が多く含まれており、典型的な pT1b の所見が多くなかった可能性がある。今後、初回治療として外科切除が施行された pT1b 病変も含めた検討ができれば、AI の診断傾向が明瞭となる可能性がある。

本研究は、T1b 癌の拾い上げ (存在診断) に寄与できる側面もあると考えられる。著者らは、スクリーニング内視鏡検査中に T1b 癌に遭遇した際、AI が red flag などにより注意喚起できるような AI システムを構築したいと考えている。本研究がそのような AI システムを構築するための契機となることにも期待している。

V 結 論

多施設から収集した pT1b 画像を用いて CAM 画像を作成し、AI の ROI について検討した。AI と内視鏡医の ROI は概ね一致しており、その場合の AI の ROI は内視鏡医と類似性があると考えられた。また T1b 癌と診断可能な領域をアノテーションすることにより ROI の不一致を克服することが、AI の学習効率の改善につながる可能性があると考えられた。

本論文内容に関連する著者の利益相反：朱欣 (日本学術振興会 (2018 年 -2020 年度科研費基礎研究 C 研究課題番号 18K11532 18K08010))、山本博徳 (富士フイルムメディカル株式会社)

文 献

1. Tanaka S, Oka S, Chayama K. Colorectal endoscopic submucosal dissection : present status and future perspective, including its differentiation from endoscopic mucosal resection. *J Gastroenterol* 2008 ; 43 : 641-51.
2. Tanaka S, Terasaki M, Kanao H et al. Current status and future perspectives of endoscopic submucosal dissection for colorectal tumors. *Dig Endosc* 2014 ; 24 : S73-9.
3. Tanaka S, Asayama N, Shigita K et al. Towards safer and appropriate application of endoscopic submucosal dissection for T1 colorectal carcinoma as total excisional biopsy : Future perspectives. *Dig Endosc* 2015 ; 27 : 216-22.
4. 岡 志郎, 田中信治, 田丸弓弦ほか. 大腸 T1(SM) 癌に対する内視鏡的摘除の現況と将来展望. *日本消化器病学会雑誌* 2016 ; 58 : 1311-23.
5. Tanaka S, Kashida H, Saito Y et al. Japan Gastroenterological Endoscopy Society guidelines for colorectal endoscopic submucosal dissection/endoscopic mucosal resection. *Dig Endosc* 2019 ; 32 : 219-39.
6. Hashiguchi Y, Muro K, Saito Y et al. Japanese Society for Cancer of the Colon and Rectum. Japanese Society for Cancer of the Colon and Rectum (JSCCR) guidelines 2019 for the treatment of colorectal cancer. *Int J Clin Oncol* 2020 ; 25 : 1-42.
7. Kudo S, Tamura S, Nakajima T et al. Diagnosis of colorectal tumorous lesions by magnifying endoscopy. *Gastrointest Endosc* 1996 ; 44 : 8-14.
8. Matsuda T, Fujii T, Saito Y et al. Efficacy of the invasive/non-invasive pattern by magnifying chromoendoscopy to estimate the depth of invasion of early colorectal neoplasms. *Am J Gastroenterol* 2008 ; 103 : 2700-6.
9. Mukae M, Kobayashi K, Sada M et al. Diagnostic performance of EUS for evaluating the invasion depth of early colorectal cancers. *Gastrointest Endosc* 2015 ; 81 : 682-90.
10. Gall TM, Markar SR, Jackson D et al. Mini-probe ultrasonography for the staging of colon cancer : a systematic review and meta-analysis. *Colorectal Dis* 2014 ; 16 : O1-8.
11. Sano Y, Tanaka S, Kudo SE et al. Narrow-band imaging (NBI) magnifying endoscopic classification of colorectal tumors proposed by the Japan NBI Expert Team. *Dig Endosc* 2016 ; 28 : 526-33.
12. Kobayashi S, Yamada M, Takamaru H et al. Diagnostic yield of the Japan NBI Expert Team (JNET) classification for endoscopic diagnosis of superficial colorectal neoplasms in a large-scale clinical practice database. *United European Gastroenterol J* 2019 ; 7 : 914-23.
13. Saitoh Y, Obara T, Watari J et al. Invasion depth diagnosis of depressed type early colorectal cancers by combined use of videoendoscopy and chromoendoscopy. *Gastrointest Endosc* 1998 ; 48 : 362-70.
14. Horie H, Togashi K, Kawamura YJ et al. Colonoscopic stigmata of 1 mm or deeper submucosal invasion in colorectal cancer. *Dis Colon Rectum* 2008 ; 51 : 1529-34.
15. Ito N, Kawahira H, Nakashima H et al. Endoscopic diagnostic support system for cT1b colorectal cancer using deep learning. *Oncology* 2019 ; 96 : 44-50.
16. Lui T, Wong K, Mak L et al. Endoscopic prediction of deeply submucosal invasive carcinoma with use of artificial intelligence. *Endosc Int Open* 2019 ; 7 : E514-20.
17. Tokunaga M, Matsumura T, Nankinzan R et al. Computer-aided diagnosis system using only white-light endoscopy for the prediction of invasion depth in colorectal cancer. *Gastrointest Endosc* 2021 ; 93 : 647-53.
18. Nakajima Y, Zhu X, Nemoto D et al. Diagnostic performance of artificial intelligence to identify deeply invasive colorectal cancer on non-magnified plain endoscopic images. *Endosc Int Open* 2020 (in press).
19. Zhou B, Khosla A, Lapedriza A et al. Learning Deep features for discriminative localization. 2016 IEEE conference on computer vision and pattern recognition. Publication Date : 12 Dec 2016 ; DOI : 10.1109/CVPR.2016.319.
20. Kaiming H, Zhang X, Ren S et al. "Deep residual learning for image recognition." 2016 IEEE conference on computer vision and pattern recognition. Publication Date : 12 December 2016 ; DOI : 10.1109/cvpr.2016.90.

論文受付 2020年6月10日

同 受理 2021年1月20日

WHICH REGION DOES ARTIFICIAL INTELLIGENCE LOOK AT TO PREDICT T1B COLORECTAL CANCER? ANALYSIS BASED ON CLASS ACTIVATION MAPPING

Yuki NAKAJIMA¹⁾, Daiki NEMOTO²⁾, Shinichi KATSUKI³⁾,
Yoshikazu HAYASHI⁴⁾, Masato AIZAWA²⁾, Kenichi UTANO²⁾,
Takahito TAKEZAWA⁴⁾, Yuichi SAGARA⁴⁾, Xin ZHU⁵⁾,
Goro SHIBUKAWA¹⁾, Hironori YAMAMOTO⁴⁾ AND Kazutomo TOGASHI²⁾

1) *Department of Gastroenterology, Aizu Medical Center, Fukushima Medical University.*

2) *Department of Coloproctology, Aizu Medical Center, Fukushima Medical University.*

3) *Department of Gastroenterology, Otaru Ekisaikai Hospital.*

4) *Department of Gastroenterology, Jichi Medical University.*

5) *Biomedical Information Technology, the University of Aizu.*

Background and Aims: We previously showed that the performance of artificial intelligence (AI) in diagnosing deep ($\geq 1\text{mm}$) submucosally invasive (T1b) colorectal cancer was relatively good after training with non-magnified white light images (sensitivity 80%, specificity 87%). However, the “region of interest” (ROI) within the image that was responsible for the AI diagnosis is a black box. Recently, the class activation mapping (CAM) technique has been developed, enabling identification of the ROI within the image that AI utilized for making the diagnosis. In this study, we aimed to investigate features of the ROI selected by AI using CAM and clarify the similarities and differences between AI and endoscopists.

Methods: We selected endoscopic digital images that were used for training or validation of our AI system in our previous study, comprising histologically proven T1b colorectal cancers ($n = 114$, 0-I s 69, 0-II a 39, 0-II c 6; maximum diameter $16.5 \pm 13.4\text{mm}$). The application of CAM was limited to a maximum of two images per lesion from which T1b cancer was diagnosed. The CAM images were generated on ResNet50, and the ROI defined by AI was depicted in red. Two expert colonoscopists rated characteristics of the ROI following discussion. The outcome measures were concordance of the ROI defined by AI with the ROI defined by expert endoscopists, and endoscopic features analyzed by AI, including color (red or non-red), surface morphology (depressed, flat, protruding), presence of bleeding, and fold convergence. Concordance of the ROI defined by AI and the ROI that was defined by expert endoscopists was rated by concordance area and classified into excellent ($\geq 75\%$), fair ($\geq 25\% < 75\%$) and poor ($< 25\%$).

Results: CAM images were successfully generated for all 226 images. The level of concordance between the ROI defined by AI and the ROI that was defined by expert endoscopists was excellent in 39%, fair in 34% and poor in 27%. In images showing poor concordance, the ROI defined by AI was distant from the T1b cancer. After excluding lesions with poor concordance, the vast majority (91%) of the ROI defined by AI was concordant with the ROI containing endoscopists' identification of red color, and a small proportion (21%) of the ROI defined by AI revealed bleeding. Among the lesions detected by AI, the surface morphology was depressed in 39%, flat in 5% and protruding in 57%. Fold convergence was observed in 34% of the ROI defined by AI.

Conclusions: Most of the ROIs identified by AI were concordant with ROIs defined by experienced endoscopists, although AI may diagnose T1b colorectal cancer using different features of the ROI. Since a quarter of ROIs were present within normal mucosa, annotation of the image that the image should be reviewed by expert endoscopists may improve the diagnostic accuracy of AI for T1b colorectal cancer.