

# 胃粘膜下腫瘍の画像診断に対する機械学習の適用

○石原里夏 半田久志 鎌田研 田中秀和 (近畿大学)

## Application of Machine Learning to Imaging Diagnosis of Gastric Submucosal Tumors

\*R. Ishihara, H. Handa, K. Kamata and H. Tanaka (Kindai University)

**Abstract**— In recent years, machine learning has been used to analyze and process medical images for anomaly detection and disease estimation. Gastric submucosal tumors have been diagnosed using ultrasound endoscopy. In addition to the pathological approach, imaging diagnosis, which is less burdensome for both physicians and patients, is increasingly being used. In this study, we applied machine learning to both CH-EUS and elastography, which are used for imaging diagnosis of gastric submucosal tumors. The sensitivity of CH-EUS was 92.9% and specificity was 76.9%, while elastography had a sensitivity of 100% and specificity of 50%.

**Key Words:** CH-EUS, エラストグラフィ, SiamMask, ResNet, 画像診断

### 1 はじめに

近年、画像化技術の発展により、組織を採取する病理学的アプローチから、医師、患者共に負担の少ない画像診断を用いる例が多くなっている。胃粘膜下腫瘍をはじめとする粘膜下腫瘍も、外科的切除を行う診断から、細胞の一部を採取し、病理学的に分析する手法に変わり、造影剤の発展により画像診断技術の有用性も示されている。他にも、組織の硬度を色情報でマッピングするエラストグラフィなどの取り入れも進められている。しかし、新たな検査法が提案されても症例が少ない上、描画法や解釈、画像解析法が複雑なため診断には医師の主観による差が生じる。そういった診断時の客観性を高める方法として、医用画像に対して機械学習を用いることで、自動診断を行う研究が行われている<sup>1)</sup>。

また、畳み込みニューラルネットワーク (convolutional neural network: CNN) の研究が進み、画像分類の分野において識別精度が向上している。CNN は層数を重ねることでより多くの表現を認識することが可能であると、経験的に知られている。しかし、単に層を追加したモデルは、恒等写像をうまく学習することができず、誤差の改善が行えない。この問題を解決する手法として、2015年に残差ネットワーク (residual network: ResNet) が提案された。ResNet では、shortcut connection を導入することにより、恒等写像の学習を容易にした。

本研究では、胃粘膜下腫瘍の診断における CH-EUS、エラストグラフィの2つの画像診断方法に対して、機械学習を用いることで、定量的・客観的な診断に貢献できるかを検討する。実験では、ResNet を用いた CH-EUS における画像分類問題と、SVM を用いたエラストグラフィにおける画像分類を行い、それぞれ性能評価を行う。また、医用画像には医療機器の一部や診断時に写る患者情報など、患部以外の情報が入っているため、そのままの画像を用いると患部以外の情報で診断を行うモデルになる可能性がある。そのため、事前処理として軽量かつトラッキングと同時にセグメンテーションも行うことができる SiamMask を用いて腫瘍部分のみの抽出を行い、セグメントの有用性の検証を行う。

### 2 胃粘膜下腫瘍

胃粘膜下腫瘍は胃粘膜の下 (胃の壁の中) に存在し、正常な粘膜に覆われており、胃の内腔になだらかに突出しているものの総称である<sup>2)</sup>。極めて発症が珍しく、外科的切除前の確定診断が困難なため、医学界でも負担の少ない診断方法について様々な研究が行われている。胃粘膜下腫瘍は従来、内視鏡的粘膜下層剥離術<sup>3)</sup> や免疫組織化学分析<sup>4)</sup> などによる病理学的研究が行われてきた。しかし、これらは医療機器の発展により最小限に病変を切除することが可能となったものの、切除した細胞を組織病理学的検査をもって確定診断を下さなければならぬという問題がある。これらは、過剰手術や再手術など生活の質の低下につながるいくつかの欠点があり、手術には時間を要する上、内視鏡技師、医師のテクニックも要する。免疫組織化学分析では手術前に組織学的診断を得る研究として吸引標本の組織病理学的評価で 91% と高い精度を示したが、臨床経過観察を要した<sup>4)</sup>。これに対して超音波内視鏡 (Endoscopic Ultrasonography: EUS) では画像のみで 78% の精度が見られ、EUS 画像診断の精度を高めることで医師、患者双方の負担軽減につながると考えられる。

#### 2.1 CH-EUS

通常の EUS に超音波造影剤を加えることで腫瘍が造影され、腫瘍の種類を特定する手法である。造影 CT や MRI で使用される造影剤と比べて副作用もほとんどない<sup>5)</sup>。日本で現在使用できる超音波造影剤はソナゾイドという直径約  $3 \mu\text{m}$  のマイクロバブルである。これを静脈内に投与し、超音波で共振または崩壊させ、発生した二次性高調波を画像化したものが造影 EUS となる。鑑別のポイントは血流の量と、irregular vessel の有無である。良性腫瘍は血流が少ないのに対し、悪性腫瘍はとても血流が豊富で、悪性度が高くなるにつれ、irregular vessel が出てくる。CH-EUS では、血管が白く描写され、irregular vessel は白く写る腫瘍内部に一部分だけ黒く写る箇所として見受けられる。Fig. 1 に CH-EUS 検査から得られる画像を示す。

## 2.2 エラストグラフィ

エラストグラフィは、従来の超音波検査で分からなかったしこりや臓器の硬さを色で判別する検査である<sup>6)</sup>。癌性腫瘍は多くの場合、早期の段階から血管や細胞の密度の増加によりその硬さが増すことがわかっている。エラストグラフィは組織の硬さを組織硬度の軟を赤、中間を緑、硬を青として1~256の階調に映像化することで、組織が良性か悪性かを鑑別できる。指標の1つとしてGiovanniniらが色調パターンをFig. 3のように5段階に分類している<sup>7)</sup>。Score 1は軟シグナル(緑)が均一なもの、Score 2は軟シグナル不均一なもの、Score 3は硬シグナル(青)が均一で軟シグナル(緑)が縁取りしているもの、Score 4は硬シグナル均一、Score 5は硬シグナルが不均一で、Scoreが高くなるにつれ悪性度が高くなる。スコアの決め方は医学的経験則に基づいたものであり、胃粘膜下腫瘍への適用に関してはまだ確立しておらず、研究段階である。Fig. 2にエラストグラフィ検査から得られる画像を示す。

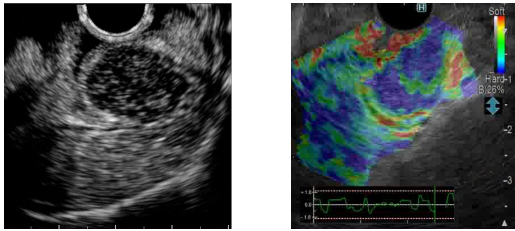


Fig. 1: CH-EUS

Fig. 2: elastography

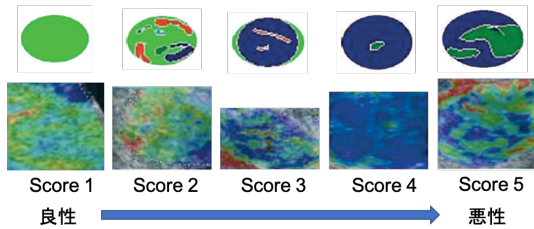


Fig. 3: Evaluation Score(Modified from quotation<sup>7)</sup>)

## 2.3 実験で用いるデータ

実験に用いる超音波内視鏡画像は、近畿大学附属病院にてCH-EUSは2015年6月~2020年2月、エラストグラフィは2017年2月~2020年2月に検査を受けた患者の動画データより切り出した画像である。また、患者は確定診断を受けており、分類の正解ラベルはこれに準拠する。

## 2.4 画像データの前処理

EUS検査では、画像自体に情報量が少なく、内視鏡先端部のエッジが強いため、そのまま機械学習を用いても腫瘍内部の特徴量をうけない可能性がある。また、エラストグラフィは医師により加圧箇所や加減が異なり、腫瘍以外の部分にも色がつく場合があるため、腫瘍部分のみのセグメンテーションを行う。EUS検査は、染影中に腫瘍箇所の位置や形状が大きく変化することがあるため、高速なトラッキングとセグメンテーションが可能なSiamMaskを用いる。そのための前処理として、患者1人分のEUS動画を0.1秒間隔で切り出す。

## 3 SiamMask

SiamMask<sup>8)</sup>は、Siamese Network系アプローチの1つであり、従来のSiamFC、SiamRPNにMaskを出力するブランチを追加し、オブジェクトトラッキングと半教師付きオブジェクトのセグメンテーションを同時に解く手法である。従来手法と異なり、フレームごとに二値セグメンテーションマスクを出力する。

局所画像 $z$ をそれより大きい探索画像 $x$ と比較し、密な類似性マップを得る。 $z$ は対象物体を中心に $w \times h$ でcropされ、 $x$ は最後に推定された対象の位置を中心に、それより大きいcropを適用されている。2つの入力は同じCNN( $f_\theta$ )で処理され、相互に相関のある2つの特徴量マップを出力する。その後、局所画像の特徴量マップをカーネルとした2次元畳み込みを、探索対象画像の特徴量マップに対して行うことで類似性マップを得る。

$$g_\theta(z, x) = f_\theta(z) * f_\theta(x) \quad (1)$$

類似性マップの各空間要素(式1での左辺)をresponse of a candidate window(RoW)と呼ぶ。 $g_\theta^n(z, x)$ は $z$ と $x$ 内の $n$ 番目の候補ウィンドウとの類似度を出力している。

各RoWについて、シンプルな2レイヤニューラルネットワーク $h_\phi$ で構成されるピクセルごとにsigmoidを適用したあと、出力をしきい値0.5で二値化する。 $\phi$ は学習可能パラメータ、 $m_n$ が $n$ 番目のRoWに対応する予測されたマスクであるとする。

$$m_n = h_\phi(g_\theta^n(z, x)) \quad (2)$$

式2からマスク予測はセグメントすべき $x$ とターゲット $z$ 両方の関数であり、 $z$ はセグメンテーション処理をガイドする参照として使用できる。つまり、別の画像が与えられれば、ネットワークは異なるセグメンテーションマスクを出力できる。

## 4 ResNet

深層学習モデルの一つに、画像処理などのタスクにて高い性能を挙げているCNNがある。しかし、CNNのアーキテクチャでの多層化にあたり劣化問題による精度の低下が見られた。ResNetは、劣化問題を解決するために考案されたCNNアーキテクチャである。ResNetは複数のResBlockで構成されている。ResBlockにはShortcut connectionが導入されており、ResBlock内部の非線形関数 $F(x)$ はResBlockの出力と入力の残差を学習するように設計されている。求める特徴マップを $H(x)$ とした時、畳み込み演算によって得られる特徴マップ $F(x)$ は式3となるように学習する。

$$F(x) = H(x) - x \quad (3)$$

## 5 評価方法

### 5.1 Leave-one-out cross validation 法

モデルの性能比較として、Leave-one-out cross validation法がある。はじめに全データから1つのサンプルを除き、残りのデータでモデルを作成する。作成したモデルを用いてあらかじめ除いた1つのサンプルを予測し、同様の操作をデータの数だけ行い、それらの誤差の平均をとる。

本研究では Leave-one-out cross validation 法を参考に、患者 1 人の動画から得られた画像を 1 塊とする。まず、全患者 55 人の中から推定を行う 1 人を選ぶ。Fig. 4 のように選んだ 1 人を除く患者の動画から得られた画像を学習データとし、分類器を作成する。そしてあらかじめ除いた 1 人の動画から得られた画像全てに対し、作成した分類器を用いて推定を行い、得られた推定結果の平均を最終推定結果とする。この方法を全患者 55 人分繰り返す。

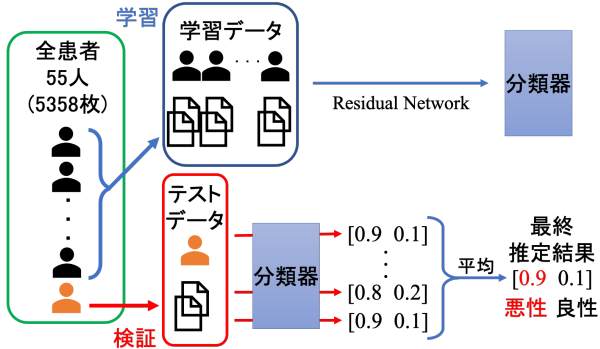


Fig. 4: Evaluation Method

## 5.2 評価指標

本研究で扱う評価指標は、感度 (sensitivity), 特異度 (specificity), 正診率 (accuracy) である。Fig. 5 は 2 クラス分類における混合行列である。

		検査結果	
		陽性	陰性
群	疾患群	True Positive	False Negative
	正常群	False Positive	True Negative

Fig. 5: mixed matrix

感度は、疾患群が陽性になる確率を示しており、次式で求められる。

$$Sensitivity = \frac{TP}{TP + FN} \quad (4)$$

特異度は、正常群が陰性になる確率を示しており、次式で求められる。

$$Specificity = \frac{TN}{TN + FP} \quad (5)$$

正診率は、正しく診断する確率を示しており、次式で求められる。

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (6)$$

## 6 実験

### 6.1 実験準備

本研究では CH-EUS では患者 55 人 (悪性: 42 人, 良性: 13 人), エラストグラフィでは患者 21 人 (悪性: 17 人, 良性: 4 人) を対象とする。実験の準備として、SiamMask による腫瘍部分の抽出を行う。最初の 1 フレームから手動で関心領域をバウンディングボックス

で指定し、次フレームから指定領域のトラッキングを行い、セグメンテーション部分を画像として切り出す。CH-EUS では医師の診断基準である造影剤投与後の 20 秒間の画像を、エラストグラフィでは加圧時の色がついたフレームの画像を用いる。

領域抽出妥当性の検証のため、CH-EUS の各患者の動画 1s, 10s, 20s 時点の SiamMask によってマスキングしたものと、医師がマスキングした画像を比較する。Fig. 6, 7 に SiamMask によって領域抽出したものと、医師がマスキングしたものをそれぞれ腫瘍部分の範囲を白、それ以外の背景を黒として塗りつぶした画像を示す。セグメント範囲一致率には Dice 係数 (F 値) を用いる。追跡できている画像の F 値は平均 96% であった。

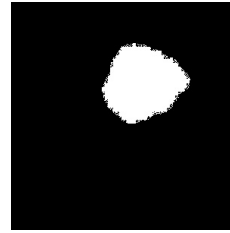


Fig. 6: Tumor image with region extraction using SiamMask



Fig. 7: Tumor image with region extraction by a doctor

SiamMask により得た画像枚数は、CH-EUS では 5358 枚 (悪性: 3847 枚, 良性: 1511 枚), エラストグラフィでは 4285 枚 (悪性: 3657 枚, 良性: 628 枚) であった。これらの画像を  $256 \times 256$  ピクセルに成形したものにラベルを付加し、データセットとする。

### 6.2 実験 1 : CH-EUS

ResNet を用いて学習・推定を行う。悪性, 良性それぞれ 1000 枚ずつ、計 2000 枚を学習データとし、学習は 20 epoch 行う。ResNet の出力は softmax を用いており、悪性・良性の推定割合が出る。カットオフは 50%とした。

確定診断結果を知らない状態で、動画のみから所見診断を行う blind reading, SiamMask のセグメントを用いない画像での診断, SiamMask のセグメントを用いた画像での診断の 3 手法で比較を行う。blind reading は医師の方 2 名によって下された結果とする。結果を Table 1 に示す。Table 1 より、SiamMask によるセグメントを用いることで感度・特異度共に精度の向上が見られることがわかる。

Table 1: Evaluation results of CH-EUS

	感度 [%]	特異度 [%]	正診率 [%]
blind reading	90.5	69.2	85.5
セグメントなし	84.3	63.6	79.0
セグメントあり	92.9	76.9	89.1

また、医師による blind reading と本手法共に推定を間違えた例は全患者 55 名中 2 名のみであり、診断を掛け合わせる事でより良い診断精度が期待できる。

本手法, 医師共に診断を間違えた症例の 1 つを造影剤投与後 20 秒時点の腫瘍画像を Fig. 8 に、造影剤投与後 50 秒時点の腫瘍画像を Fig. 9 に示す。腫瘍箇所は赤線で囲んでいる部分である。この症例は、良性腫

瘍の特徴として見られる造影効果の低さが時間経過により徐々に認められる例だったため、今回の検討時間である 20 秒間では推定が難しい症例であった。この問題に対して、検討時間を 60 秒に延長することで精度の向上が期待できる。

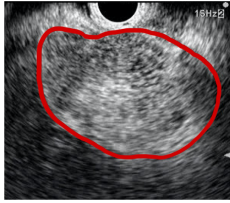


Fig. 8: 20 seconds after administering the contrast medium

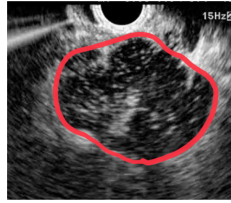


Fig. 9: 50 seconds after administering the contrast medium

### 6.3 実験 2 : エラストグラフィ

造影時の画像のみを用いるため、患者一人につき取得できる枚数に差がある。そのため、患者 1 人に対してそれぞれ 5 枚ランダムで取ったものを学習・検証に用いる。悪性と良性の患者数に極端に偏りがある点、色情報 RGB の 3 次元情報が診断指標となる点から、サポートベクターマシン (SVM) を用いる。

結果は、セグメントなしでは感度 100%、特異度 0% に対し、セグメントありでは感度 100%、特異度 50% となった。

Fig. 10 に腫瘍部分の Blue と Green の割合を説明変数として RBF カーネルを用いたものを示す。悪性患者のデータを赤丸、良性患者データを青四角とする。図より、悪性腫瘍は青色の割合が大きく、良性腫瘍は緑色の割合が大きいという指標のもと、分類できている事がわかる。

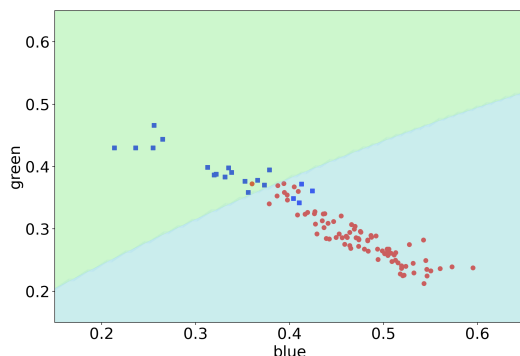


Fig. 10: SVM results

また、SiamMask によるセグメント有無の比較のために、セグメントなしを Fig. 11 に、セグメントありを Fig. 12 にそれぞれ RGB を軸とした 3 次元グラフを示す。

Fig. 12 より、悪性腫瘍データはまとまっているのに対し、良性腫瘍データには広がりがある。これは、本研究で用いた良性腫瘍 4 例のなかに、腫瘍全体が染色できていないものが 2 例あり、映し出された腫瘍の一部分の影響を強く受けているためと考えられる。

エラストグラフィ検査に関しては総患者数が少ない上、きれいに腫瘍箇所を描画する技術を要求されるた

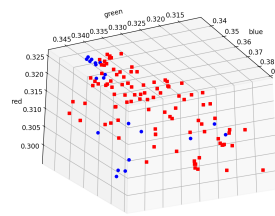


Fig. 11: No segment using SiamMask

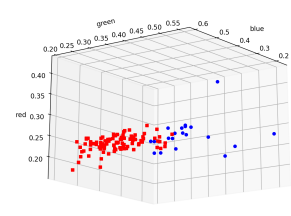


Fig. 12: Segmented using SiamMask

め、利用できるデータが少ない。今後検査数を増やすことで、異常検知手法の検討などが可能になると考えられる。

## 7 おわりに

本研究では、胃粘膜下腫瘍診断において、SiamMask による関心領域の抽出と、そこから得られる画像を利用した機械学習による所見疾患推定を行った。実験の結果、CH-EUS 検査に関しては ResNet による分類を行い、感度 92%、特異度 76% と高い精度を示した。エラストグラフィ検査に関しては SVM による色情報での分類を行い、感度 100%、特異度 50% という精度を得た。また、医師の診断指標におおよそ適した形での分類が可能であることを示した。

本研究では、SiamMask を用いた領域抽出画像のなかで、正しく追跡できている画像のみを用いたが、検査方法、患者、撮影時の医師の技術によって追跡の難度は変わる。診断に十分な画像枚数を得るためにも、さらなる検出率の向上が求められる。

## 参考文献

- 1) 山田真善, 上條憲一, 齋藤豊 : 形態情報定量化を基盤とする人工知能システムを活用した大腸がんおよび前がん病変発見のためのリアルタイム内視鏡画像自動解析システムの開発, 日本消化器病学会雑誌, 114(suppl-2), 498-498 (2017)
- 2) <https://medicalnote.jp/diseases/胃粘膜下腫瘍>
- 3) Shu Hoteya, Toshiro Iizuka, Daisuke Kikuchi, and Naohisa Yahagi : Endoscopic submucosal dissection for gastric submucosal tumor, endoscopic sub-tumoral dissection, Digestive Endoscopy, 266-269 (2009)
- 4) Nobuhiro Ando, Hidemi Goto, Yasumasa Niwa, Yoshihiro Hirooka, Naoki Ohmiya, Tetsuo Nagasaka, and Tetsuo Hayakawa : The diagnosis of GI stromal tumors with EUS-guided fine needle aspiration with immunohistochemical analysis, Gastrointestinal Endoscopy, 55, 37-43 (2002)
- 5) 北野 雅之, 工藤 正俊 : 造影 EUS のコツ, 日本消化器内視鏡学会雑誌, 53(1), 76-86 (2011)
- 6) 糸川 文英, 糸井 隆夫, 森安 史典, 土田 明彦, 長尾 俊孝, 大城 久 : EUS Elastography の有用性と可能性, MEDIX, 60, 11-16 (2014)
- 7) M Giovannini, L C Hookey, E Bories, C Pesenti, G Monges, and J R Delperio: Endoscopic ultrasound elastography: the first step towards virtual biopsy? Preliminary results in 49 patients, Endoscopy, 38, 344-348 (2006)
- 8) Qiang Wang, Li Zhang, Luca Bertinetto, Weiming Hu, and Philip H.S. Torr : Fast Online Object Tracking and Segmentation: A Unifying Approach, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 1328-1338 (2019)