

主論文の要旨

**Artificial intelligence-based diagnosis of  
upper gastrointestinal subepithelial lesions  
on endoscopic ultrasonography images**

超音波内視鏡における人工知能を用いた上部消化管上皮腫瘍の  
診断能の検討

名古屋大学大学院医学系研究科 総合医学専攻  
病態内科学講座 消化器内科学分野

(指導：川嶋 啓揮 教授)

平井 恵子

## 【緒言】

超音波内視鏡 (endoscopic ultrasonography; EUS) は上皮下腫瘍 (subepithelial lesion; SEL) の鑑別診断に有用だが、すべての病変に対して正確に鑑別を行うことは容易ではない。gastrointestinal stromal tumor (GIST) は消化管 SEL の中で最多であり、malignant potential を有するため、GIST と他の良性の SEL との鑑別は特に重要である。

EUS による SEL の鑑別診断の正診率は 45.5–66.7% と高くないため、EUS-guided fine-needle aspiration biopsy (EUS-FNAB) や粘膜切開生検などの組織採取法が推奨される。既報では、EUS-FNAB の正診率は 62.0–93.4% と幅広く、2cm 未満の小病変では低い傾向であり、粘膜切開生検の正診率は 85.2–94.3% であった。しかし、いずれの方法も侵襲的で術者の技量を要するため、高精度で非侵襲的な診断法が望まれる。

一方、近年 deep learning を用いた人工知能 (artificial intelligence; AI) は医療分野において目覚ましく発展している。EUS における AI を用いた SEL 診断の正診率は 79.2–90.0% と報告されている。しかし、既報では全 dataset の症例数は 300 例未満、参加施設は単施設または 4 施設未満、鑑別は GIST と非 GIST の 2 分類のみと限られていた。本研究は多施設の EUS 画像を用いて SEL を多クラスに分類する AI を構築し、その有用性を検討することを目的とした。

## 【対象及び方法】

2005 年 1 月から 2020 年 12 月までに、当院および 11 参加施設で病理学的に診断された上部消化管 SEL (GIST、平滑筋腫、神経鞘腫、神経内分泌腫瘍、異所性膵)、および病理学的診断はないが臨床・画像所見に基づいて異所性膵と診断された病変のうち、EUS 画像が収集可能であった症例を対象とした。病理学的に診断された SEL の EUS 画像は症例数に基づきランダムに training、validation、test dataset に分割した。deep learning のアルゴリズムは PyTorch を用いて EfficientNetV2-L を元に構築し、ImageNet21k で学習済みのモデルの重みを用いて convolutional neural network 全層の fine-tuning を行った。EUS 画像は病理結果に基づいてラベリングを行い、deep learning アルゴリズムの入力サイズに合わせて 224×224 pixel の正方形にトリミングした。作成した dataset は deep learning のアルゴリズムに学習させた。他の SEL よりデータの少ない神経鞘腫と異所性膵の画像の学習過程にはそれぞれ deep convolutional generative adversarial network (DCGAN) と非ラベル化データを利用した半教師あり学習を用いた。training の効率化のため、Distributed-Data-Parallel、multi-node、multi-process、スーパーコンピュータ「不老」を用いた。過学習防止のため、data augmentation として回転、水平・垂直反転、random erasing、RandAugment を用いた。最適化アルゴリズムには SGD with momentum を用い、学習パラメータはバッチサイズ 288、2,000 エポック、学習率 0.004 とし、cosine annealing により減衰した。training 後の AI は各 SEL の probability score の合計が 1 となるように score を出力した。画像毎の診断は最高の probability score を示す SEL、症例毎の診断は各症例における全画像の予測値の最頻値とした。また、expert、non-expert の各 2 名が test dataset に含まれた各症例の全画像に基づいて症

例毎に診断した。

## 【結果】

病理学的に診断された SEL631 例 16,110 枚および病理学的診断のない異所性膵 33 例 3,179 枚の画像を収集した。その中の神経鞘腫 14 例 441 枚の画像からは DCGAN を用いて 40,000 枚を自動作成し、3,000 枚を手動で抽出し、training data として用いた。ランダムに抽出した SEL122 例 2,875 枚の画像を test dataset とした (Figure 1)。患者背景について、training、validation、test dataset の間で年齢、性別、病変部位、病変径、病理組織型、組織採取方法に有意差はなかった (Table 1)。

AI は出力画像において、病変を緑の四角でマーキングし、画像毎の診断の予測値、probability score、症例毎の診断の予測値を表示した (Figure 2a、2b)。

AI の SEL の 5 分類 (GIST、平滑筋腫、神経鞘腫、神経内分泌腫瘍、異所性膵) における正診率は 86.1% であり、全内視鏡医よりも有意に高かった (正診率 27.0–68.0%、 $P < 0.001$ ) (Table 2)。AI と expert とともにカテゴリー毎の感度は GIST で高く (AI 98.8%、expert 63.5–77.6%)、神経鞘腫で低かった (AI 45.5%、expert 0%)。AI と expert とともに誤診例では神経鞘腫を GIST と判断した症例が最も多かった。AI の GIST に対するカテゴリー毎の感度は全内視鏡医より高かった (AI 98.8%、内視鏡医 25.9–77.6%)。

AI の GIST と非 GIST の鑑別における感度、特異度、正診率はそれぞれ 98.8%、67.6%、89.3% であり、感度と正診率は全内視鏡医よりも有意に高かった (感度 25.9–77.6%、 $P < 0.001$ 、正診率 44.3–72.1%、 $P < 0.001$ ) (Table 3)。AI の特異度は expert と比較して同等以上であった (AI 67.6%、expert 56.8–67.6%)。Non-expert の特異度 (75.7–86.5%) は AI と expert より高かったが、感度 (25.9–54.1%) と正診率 (44.3–60.7%) は AI と expert より低かった。

AI の GIST/神経鞘腫とその他の SEL の鑑別における感度、特異度、正診率はそれぞれ 100.0%、76.9%、95.1% であり、感度と正診率は全内視鏡医よりも有意に高く (感度 35.4–84.4%、 $P < 0.001$ 、正診率 42.6–82.8%、expert  $P = 0.003$ 、その他内視鏡医  $P < 0.001$ )、特異度は expert (76.9%) と同等で non-expert (50.0–69.2%) より高かった (Table 4)。

## 【考察】

本研究では多施設、多機種 of 超音波内視鏡スコープ・装置から得られた SEL の EUS 画像を多クラスに分類する AI を構築した。AI の GIST と非 GIST の鑑別における正診率 89.3% は、既報と比較し同等以上であった。AI の 5 分類における正診率は 86.0% を示し、臨床応用の可能性が示唆された。

内視鏡医の SEL 鑑別の正診率は 5 分類、GIST と非 GIST の鑑別ともに既報と同等であった。本研究では SEL の診断能は概して AI、expert、non-expert の順に高い結果となり、AI は特に non-expert にとって有用と考えられた。

本研究ではいくつかの限界がある。第一に、真の病理診断と外科的または内視鏡的切除を施行されなかった SEL の生検検体の病理診断、またはラベルなしデータとして

使用された異所性膵の画像診断は一致しない可能性がある。第二に、後ろ向きに収集された静止画を対象としたため、選択バイアスの可能性がある。第三に、**test dataset**に含まれる非 **GIST** の症例数は診断能の評価には不十分であった可能性がある。今後は静止画と動画を用いた大規模多施設前向き研究により、リアルタイムでの **AI** の診断能の検証や実臨床で使用可能な **graphical user interface** の開発が必要である。

#### **【結論】**

**AI** は上部消化管 **SEL** の **EUS** 画像診断において **expert** よりも高い診断能を示した。**AI** は内視鏡医による **SEL** 診断の精度向上に寄与できる可能性がある。