

博士の学位論文審査結果の要旨

申請者氏名 熊頭 勇太

横浜市立大学大学院医学研究科 外科治療学

審査員

主査 横浜市立大学大学院医学研究科 神経解剖学 主任教授 船越 健悟  
副査 横浜市立大学大学院医学研究科 肝胆膵消化器病学 主任教授 中島 淳  
副査 横浜市立大学大学院医学研究科 消化器・腫瘍外科学 准教授 石部 敦士

## 博士の学位論文審査結果の要旨

### Automated segmentation by deep learning of loose connective tissue fibers to define safe dissection planes in robot-assisted gastrectomy

(ロボット胃癌手術における安全な剥離層を構成する疎性結合組織の深層学習を用いた自動認識)

審査冒頭で以下のように学位研究の要旨が説明された。

#### 1. 序論

手術内視鏡と手術支援ロボットによる技術革新は外科医の目と手の支援を可能にした。しかし手術合併症は依然として克服できていない課題である。ヒューマンエラーに関連した手術合併症のおよそ3割は、手術中の誤認識が原因であるとされている。外科医の手術時の認知能力は、外科医自身の経験や身体的、精神的状況に左右される。新たな技術革新によって外科医の認識、すなわち頭脳を支援することができれば、手術合併症の減少に貢献できる可能性がある。本研究は、手術中の複雑で認識が困難な解剖学的構造を認識するために、医用画像解析技術のディープラーニングを用いて検討した。手術中の安全な剥離層を構成する疎性結合組織を、AIは外科医が認める高い精度で自動的に推論し、外科医が安全な剥離層を可視化できるAIモデルを開発することを目的とした。

#### 2. 実験材料と方法

1. AIモデルの開発方法：ロボット支援下胃癌手術の20例の手術動画を編集し、疎性結合組織が明瞭に映る静止画を1800枚抽出した。それぞれの静止画に対して、外科医が疎性結合組織の領域を20000箇所以上アノテーション（情報付け）し、教師画像を作成した。U-netをもとに構築した、深層学習アルゴリズムを用いて原画像と教師画像をAIに学習させ、AIモデルを作成した。

##### 2. AIモデルの評価方法

学習データに含まれる画像とは異なる手術動画10例から20フレームの評価画像を抽出してAIモデルの評価を行った。評価方法は下記の2つの方法に分けて解析した。

##### 2.1 コンピュータ測定による定量的画像評価

評価画像に対して外科医が手動でアノテーションした正解画像とAI推論画像を比較して、画像中の画素の一致率を評価した。評価指標は、セグメンテーション技術を用いた深層学習アルゴリズムの評価に頻用される Recall（感度）と Dice 係数（構造物の面積の類似率）を用いた。

##### 2.2 外科医の目視による定性評価

20名の外科医が、原画像とAI推論画像を並列した画像を確認し、質問用紙を用いて下記の質問に回答した。

質問 1. (Sensitivity score:感度 スコア) 本来結合組織である領域を、AIが正確に認識して

いるか？

回答 1

0(Fail): 0-19%

1(Poor): 20-39%

2(Fair): 40-59%

3(Good): 60-79%

4(Excellent): 80-100%

質問 2. (Misrecognition score:誤認識 スコア) 本来結合組織でない領域を, AI が誤って認識していないか？

回答 2.

0(Excellent) :誤認箇所はない

1(Good): 1 箇所

2(Fair): 2 箇所

3(Poor): 3 箇所

4(Fail): 4 箇所以上

### 3. 結果

定量評価では, 平均 Recall は 0.606, 平均 Dice 係数は 0.549 であった. 外科医の定性評価では, 4 点満点の評価スコアで感度 score が 3.52 点 (範囲:2.45-3.95), 誤認識スコアが 0.14(範囲:0-0.7)であった.

感度スコアの回答の 78%が最高得点の 4 点 (80-100%) であり, 外科医は AI が推論した疎性結合組織領域に高い評価を行なった. また, 誤認識スコアの平均値は 0.14 (範囲 0-0.7) と低く, 偽陽性はほとんど認められなかったが, 一部の画像はガーゼの繊維や手術鉗子先端の細かい溝, 脂肪や血液の表面の細かいハレーションなどの疎性結合組織と類似した外見的特徴を誤認識した. また, 感度スコア (外科医の評価) と Recall (コンピュータ評価) には, 強い相関があることが明らかになった(相関係数 0.733).

外科医が手動でアノテーションした正解画像と AI 推論画像に不一致がほとんど認められず, 肉眼的にはほぼ同様の領域を示していた画像の Dice 係数は 0.642 であった.

### 4. 考察

AI モデルの性能は定量的に実証された. しかし, 定量評価は正解画像の不確実性と適切な水準が不明瞭で臨床的解釈が困難という弱点があるため, 外科医の目視による評価を併せて行い, 感度は高く, 偽陽性は少ないという高い評価を得た. 手術中の安全な剥離層を構成する疎性結合組織を AI は外科医が高く評価する精度で推論が可能であった. 本技術は外科医の安全な剥離層の認識を支援し, 手術合併症の減少に貢献できる可能性がある.

### 5. 現在の研究状況と将来展望

疎性結合組織のみならず神経, 血管, 膵臓, 尿管, 出血, 手術器具と幅広い構造物を認識する AI モデルを構築した. また, 対応する手術領域や手術内視鏡システムの拡大, リアルタイムに使用できる解析速度を達成するための改良も行なっている.近い将来に臨床導入される日を目指し, 研究開発を今後も進める.

論文要旨の説明に続き、以下の質疑応答がなされた。

副査：中島教授

Q1. 切除部位を適切にアシストするようなシステムの開発はないのか？

A1. 現状の研究開発は、構造物を認識するAIシステムであり、外科医が見ているものをサポートするにとどまる。その一歩先、どこを切れば良いかというものを指示するシステムを実装する場合は、技術的に非常に難しく、医療機器の薬事承認を得る上で要求水準が高いという2つの課題がある。構造物を認識するデジタル情報をデータベースとして活用できれば、例えば疎性結合組織の量によって、切るタイミングを知らせるなどのナビゲーションは将来的に可能と考える。

Q2. AIの診断が間違った時の訴訟リスクはどうなっているか？

A2. 現在、日本で上市されている診断AI機器において、AIが最終判断を行う機器は存在しない。医師が最終責任を負うものとなっている。

Q3. 裁判になった場合はどのように評価されるか？基準はあるか？

A3. ガイドラインのようなものは、現状は存在しないと思う。私たちも取扱説明書や添付文書を作成する上で弁護士に相談を行い、「AIはミスするものであり最終判断は医師が行う」と明記すべきと伺っている。それ以上は、訴訟になった際に裁判所が判断することになると思う。

副査：石部准教授

Q4. ロボットの時代になって、新しくロボット始める上で教育的にはかなり有用と思う。教育の面で有用性を示した論文はあるのか？

A4. 現在は非医療機器で教育用途に限り、医療機関に試用してもらっているが、教育効果を上げたという論文はまだない。現場からは、教育が行いやすくなるという声は頂戴している。

Q5. Deep learnnigの中でU-netというソフトがあるということか？

A5. Deep learningの中でもどのように学習させるか決めるアルゴリズムがあり、その一つがU-netである。ソフトウェアではない。

Q6. コンピュータの性能によって変わるものか？スーパーコンピュータより量子コンピュータの方が精度はよくなると聞いた。

A6. 学習する速度は速くなるが、精度は変わらないと思っていたが、不勉強であり今後の調査課題とする。

Q7. 臨床のアウトカムはないのか？

A7. 特定臨床研究は資金が必要であり、医療機器承認を得てから臨床アウトカムを研究したい。

Q8. 構造物を二つ同時に推論する仕組みがあったが、二つで十分か？

A8. 三つ以上を同時に推論させることは理論的には可能であるが、逆に見にくくなる可能性があると考えている。

主査：船越教授

Q9. 疎性結合組織を認識させるというのは、どういう視覚の手がかりで、どういう画像上の形態の手がかりで手術者は認識しているか？AIも同様か？

A9. 例えば、脂肪と血管の間を術野展開するような場面で、白く、光沢ある、透明感のある線維組織を疎性結合組織として認識している。つまり、手術の場面と構造物の外見的特徴を手がかりに人間は認識している。同様にAIもパラメーターとして認識することが可能である。

Q10. ロボットに触覚情報はない、視覚情報だけで判断するのはリスクが高いと思う。将来的には触覚情報を取り入れることは、研究の方向性としては異なると思うが、そういう考えはあるか？

A10. 研究の構想にはないが、触覚情報もあればより良いものになると考える。

ロボットアームなどの情報も必要となり、手術支援ロボットを開発する大きな会社と共同研究になるかもしれないが、検討していきたい。

Q11. 臓器間の疎性結合組織を発表していたが、臓器の中にある疎性結合組織も認識できるか？その場合、疎性結合組織は新たに学習させる必要があるか。

A11. 例えば、ESDの粘膜剥離もある程度も認識できると考えている。精度は消化器外科手術の疎性結合組織を認識する精度より下がると思うので、新たに学習させる必要がある。

Q12. 画像認識技術で3つの手法を提示しているが、今回はsegmentationを用いている。何故か？疎性結合組織のみならず、神経や血管もsegmentationを用いているか？外科手術のモデルではsegmentation以外は必要ないのか？

A12. 全てsegmentationを使用した。その理由は、研究コンセプトが「手術支援はどのようなものが良いか」ということから着想したためである。2020年頃の他の研究では、胆嚢管を検出するためにdetectionを用いた研究もあった。しかし、手術の場合は、見つけたという情報よりも、見つけた上でどこを切るべきか、ということが重要であり、Detectionでは十分にその役割を果たすことができないため、全てsegmentationを用いている。ただし、AIモデルの精度を向上する上で、臓器によって難易度は差がある。

その他にもいくつかの質疑が行われたが、いずれも適切な回答がなされた。本研究はロボット支援下胃癌手術において、疎性結合組織剥離層を推論するAIモデルの開発と、その有用性を明らかにしたもので、臨床的に大変意義ある、将来性のある研究と評価された。以上より博士（医学）の学位授与に値すると判断された。