

別添 4

II. 分担研究報告

(1) 研究分担者報告書

希少がんである肉腫の病理診断への AI 応用に関する研究

研究分担者 高澤 豊 がん研究会がん研究所 病理診断科副部長

研究要旨:「病理デジタル画像データの深層学習・人工知能による病理画像認識診断支援ツールの開発研究」では希少がんの診断支援が重要なテーマである。がん研究会有明病院では、希少がんのうち肉腫症例が多く、これらの症例を活用することを目的として、データベースの作成、デジタル画像ファイルの蓄積、AI を用いた診断ツールの開発を行った。がん研有明病院に蓄積された約 23,000 例の骨および軟部腫瘍を WHO 分類 (2013) に基づいて、亜型を含めて約 250 の組織型に分類し、各組織型の代表例を抽出し、組織像の確認したのち、典型的な部分の組織プレパラートをスキャナーで読み込み、デジタル画像を取得し、ファイルサーバ上に保存した(2,5000 症例)。希少がんのうちで頻度の高い脂肪性腫瘍、特に高分化型脂肪腫様脂肪肉腫の診断、について AI による診断ツールの開発を試みた。深層学習のための注釈支援ツールなどのソフトウェア開発および基礎実験を行い、AI による診断に必要な注釈の種類、データ量などを評価し、高分化型脂肪腫様脂肪肉腫と脂肪腫の鑑別がほぼ 100%可能な AI 診断ツールを開発した。

A. 研究目的

「病理デジタル画像データの深層学習・人工知能 (以下 AI) による病理画像認識診断支援ツールの開発研究」では、術中迅速病理診断と希少がんの診断支援が最も重要なテーマである。希少がんは病理診断自体が難しいもの、言い換えると、診断者間の不一致が少なくないという問題点が存在する。がん研究会有明病院では、希少がんのうちとりわけ肉腫症例が多く集まっており、これらの症例を活用することによって「AI 活用による病理診断支援ツール開発」が期待される。

B. 研究方法

1. 希少がんのデータベース整備

がん研有明病院に蓄積された約 23,000 例の骨および軟部腫瘍を WHO 分類 (2013)

に基づいて再分類し、希少がんの抽出を行う。必要な場合には、免疫組織化学的検索、遺伝子検索、Fusion Panel 解析等を追加し、診断を確認する。

2. 希少がんの病理組織のデジタル画像化
データベースから各疾患の代表例を抽出し、組織像の確認したのち、症例ごとに典型的な部分を含むプレパラートを選択し、スキャナーで読み込み、デジタル画像をサーバー上に保存する。

(倫理面への配慮)データベースは全てがん研有明病院の症例であり、院内で研究担当者のみが閲覧可能なシステム上に作成している。デジタル画像としてアーカイブ化する症例は、検体の研究や教育使用に関する患者本人の承諾を得られている症例であり、デジタル画像については個人情報を含まない。遺伝子検索は腫瘍組織の体細

胞変異に限って行い、通常の診断過程で検索されるものを解析する。

3. AI を用いた希少がんの診断ツールの開発

脂肪性腫瘍の中で、良性の脂肪腫 10 症例と悪性の高分化型脂肪腫様脂肪肉腫の 15 症例を用いて、教師データ作成のためのソフトウェア開発を行い、深層学習の様々なモデルやアルゴリズムを用いて AI による深層学習を行い、適合率（正診率）を評価し、教師データの必要データ量や今後の課題を検討する。

C. 研究成果

1. 希少がんのデータベース整備

がん研における骨軟部腫瘍データベースの整備を継続した。希少がんの多い骨軟部腫瘍は分子病理学的知見により、分類、疾患名が変化している。データベースは、現在最も標準的な分類である WHO 分類（2013）による約 250 の組織型、年齢、性別、既往歴、家族歴、腫瘍の部位、腫瘍の大きさ、肉眼的特徴、組織像の特徴、腫瘍の免疫組織化学的形質、遺伝子検索の結果を含む。今年度はデータベースの整備を継続し、蓄積された約 23,000 症例のうち約 7000 症例のデータベースを完成させることができた。また、Fusion Panel 解析を追加し、確定困難であった症例の診断を確認することができた。

2. 希少がんの病理組織のデジタル画像化

データベースから各疾患例を抽出し、組織像の確認ができたものについて、症例ごとに最も典型的な部分の組織プレパラート（数枚）を選択した。過去の症例で染色の劣化が高度なものについては、標本を再作成した。今年度新たにデジタル画像化し

たプレパラートは 1500 症例分である。デジタル画像は院内のファイルサーバ上に、個人情報を除いた形で保存した。

3. AI を用いた希少がんの診断ツールの開発

(a) 教師データ作成

脂肪性腫瘍のうち、良性の脂肪腫の 10 症例と高分化型脂肪腫様脂肪肉腫の 15 症例を用いた。教師データを効率的に作成するため、デジタルデータから AI の深層学習に用いるためのデータを作成するためのツールを開発した（ソフトウェア名「Annon」（© 2019 メタデータ株式会社）。元のデジタル画像から教師データ作成のための画像の切り抜きと注釈付けの機能を持つ「Annon」を用いて、「注釈」の付け方を検討し、「良性細胞」、「悪性細胞（脂肪芽細胞あるいは異型間質細胞）」、「良悪不明な細胞」の 3 群の「注釈」付けが効率的かつ実用的な教師データであることを確認した。

(b) 深層学習の手法とアルゴリズム

教師データのデータセットとしてランダムな 100, 200, 300...1000 画像を用いた検討により、脂肪腫と高分化型脂肪腫様脂肪肉腫の鑑別補助 AI が、3 群の検出で適合率(mean average precision) 0.5 を達成するためには、1500 画像以上のデータセットが必要であることが示された(繰り返し=50 の場合)。学習手法の比較検討では、一段階検出の代表的アルゴリズムである SSD と二段階検出の代表的アルゴリズムである FasterRCNN を比較検討し、適合率は FasterRCNN が高いことを確認した。テストデータを用いた検証により、今回作成した AI 診断ツールを用いた個々の細胞の評価の結果から、腫瘍全体の組織診

断、すなわち、高分化型脂肪腫様脂肪肉腫と脂肪腫の鑑別は 100%可能であることが確認された。

D．考察

AI を病理組織学的診断に応用するための研究が世界各国で行われており、臨床応用の段階にあるものも知られている。個々の希少がんの発症率は 1%に満たないが、希少がん全体ではがんの約 20%を占める“希少でない”がんである。にもかかわらず、希少がんに対する AI の病理組織学的診断への応用は未開拓の領域である。希少がんの一つである脂肪性腫瘍についての我々の研究では、AI による正診率に最も影響を与えるのは「注釈(annotation)」の付け方であり、その基礎実験の重要性を確認した。

また、現在最も標準的な分類である WHO 分類を用いて組織型を決定する過程で、分類不能な症例が相当数含まれていた。それらを AI でどのように扱い、どう学習させるのかは今後の課題である。腫瘍の診断に用いられる遺伝子などの検索手法は日々進歩しており、バイオインフォマティクスと組織形態学的特徴の解析に AI を利用することによって新しい疾患概念が生まれる可能性がある。

E．結論

今年度は、がん研有明病院に蓄積された 23000 症例の骨軟部腫瘍のデータベース整備を継続した。AI で学習させるためのデジタル画像もアーカイブ化も進めた。AI を用いた希少がんの診断ツールで効率的に教師データ作成するためのソフトウエ

ア開発を行った。基礎的実験を行い、脂肪性腫瘍のうち高分化型脂肪腫様脂肪肉腫と脂肪腫の鑑別が可能な AI 診断ツールを開発した。

F．健康危険情報（総括参照）

G．研究発表

1. 論文発表：特になし
2. 学会発表：日本病理学会第 108 回総会「希少がんである肉腫の病理診断への病理デジタル画像・人工知能技術の応用」

H．知的財産権の出願・登録状況

1. 特許取得：特になし
2. 実用新案登録：特になし
3. その他：特になし