

子宮内膜細胞診判定プロトタイプモデルに資する AIの研究

Research on AI That Contributes to a Prototype Model for Endometrial Cytology Diagnosis

今枝俊輔・システム分科会・中央大学

The shortage of pathologists has become a serious issue, highlighting the importance of reducing their workload and ensuring diagnostic quality in regional hospitals. The use of AI for diagnostic support has been increasingly reported. However, in endometrial cytology, identifying malignant cells remains challenging due to the morphological diversity of endometrial carcinoma, and research on AI-based diagnosis in this field is still limited. Additionally, in pathology-related AI applications, deep learning-based diagnostic systems often suffer from low explainability. This study aims to develop an explainable classifier for endometrial cytology by focusing on the uneven distribution of nuclear chromatin, a malignant feature of endometrial carcinoma. We demonstrate that this uneven chromatin distribution is reflected in the shape of the histogram of the green channel in cytological images. Furthermore, we propose a method to quantify these histogram shape irregularities and use them as feature descriptors. To enhance explainability, we introduce two classification approaches: a decision tree-based method and a rule-based method.

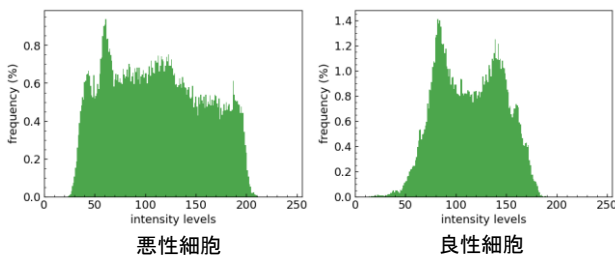
研究概要

病理医不足が深刻化し、病理医の業務負担軽減や地域病院における診断の質の確保の重要性が明らかになり、診断支援AIの活用が報告されつつある中、子宮体癌は細胞形態の多様性から悪性細胞の識別が難しく、子宮内膜細胞診のAI研究報告が少ない。また、病理領域でのAI活用において、深層学習をもちいた診断システムは、説明可能性の低さが課題であるとされている。

本研究では、説明可能な子宮内膜細胞診判定器の開発を目的として、子宮体癌の悪性所見である核クロマチンの不均等分布に着目し、緑色チャンネルのヒストグラムの形状に、核クロマチンが不均等に分布した際の特徴が表れることを示した。また、それらのヒストグラム凹凸形状特徴を定量化し、特徴量として記述する方法を提案した。さらに、この特徴量を使用する説明可能性の高いモデルとして、決定木による判定方法とルールベースによる判定方法の二つを提案した。

提案手法

下図は、悪性細胞と良性細胞の緑色チャンネルのヒストグラムである。悪性初見の一つである、核クロマチンの不均等分布により悪性細胞のヒストグラムには横幅の広い凸形状が存在し、明確な細胞境界により良性細胞のヒストグラムには凹形状が存在することが分かる。この形状をヒストグラムの形状を、移動平均やメディアンフィルタ、正規化などを用いて強調する。これにより、形状を安定化させ、外れ値などに対する形状特徴量のロバスト性を高める。強調後のヒストグラムの形状に対して、連続する3つの極値の組を凹凸形状として定量化する。ここで、中央の点が極大値の場合は「凸形状」、極小値の場合は「凹形状」となる。各凹凸形状に対し、位置、高さ、幅を定義し、凹凸計6つの特徴量を定義する。これらの特徴量は、ヒストグラムの局所的な形状情報を簡潔かつ定量的に記述するものであり、この凹凸形状特徴量を使用することで、悪性所見にもとづいた説明可能性の高いモデルを作成することができる。



評価

ヒストグラムへの変換および凹凸形状特徴への変換それぞれに対して、多層パーセプトロンを使用して良悪性判定性能を定量評価し、ヒストグラムへの変換後にはF1-Scoreは約0.823に保持され、凹凸形状特徴への変換後にはF1-Scoreは約0.758に保持されることを示した。

提案した形状特徴量を使用した説明可能性の高い良悪性判定モデルとして、決定木による判定と、ルールベースによる判定を提案した。決定木による判定においてF1-Scoreは約0.718を示した。また、ルールベースによる判定では、領域四分木をもちいた再帰的なアルゴリズムを使用し、画像全体に対する判定を可視化することで定性評価をおこない、誤った判定がなされた箇所について判定理由の説明が可能であることを示した。

今後の展望

本研究では、子宮内膜細胞診判定プロトタイプモデルとして使用可能な説明可能性の高いモデルを構築することを目的とし、悪性所見の一つである核クロマチンの不均等分布を定量化して判定モデルの入力として使用するために、緑色チャンネルのヒストグラムから凹凸形状特徴量を計算する手法を提案した。

本研究において示した凹凸形状特徴を使用したモデルは決定木とルールベースによる単純な判定である。

また、凹凸形状特徴量内にはF1-Score約0.758を導く情報があることが示された。

これは、病理のドメイン知識をさらに活用することにより、より性能の高いモデルを構築できる可能性を示している。