

AIによる血液細胞の画像分類の基礎から実線まで

◎佐藤 正一¹⁾
順天堂大学医療科学部¹⁾

人工知能 (AI: Artificial Intelligent) を利用した画像分類が近年注目され、臨床検査への応用が進んでいる。今の AI ブームは歴史的には第 3 次ブームであり、コンピュータのハードとソフトの進化により各種の業務に利用できる技術として認知されている。診療放射線の領域では、がん細胞の検出や肺癌の種類を分けるなどの方法が開発され、組織病理学的画像分野でも高精度でがん細胞を識別・分類可能という報告がある。

現在、特に注目されている画像解析は深層学習 (DL: Deep Learning) という手法を用いたもので、画像データに正解情報のタグを付ける教師あり学習法による機械学習に属する。DL の手法には様々な方法があり、基本的には脳の視覚野に関する脳科学の研究からヒントを得た、畳み込みニューラルネットワーク (CNN: Convolution Neural Network) を用いている。DL の画像解析工程は、対象とする画像を多数収集し、全ての画像に正解タグを付ける。次に正解タグを付けた画像を訓練データと評価データに分けて、訓練データを用いて学習を行い、評価データで学習の成果を評価する作業を繰り返すものである。今回、AI を用いた白血球分類方法の実践として、機械学習を実施する前段階である教師画像作成 (画像アノテーション) から画像学習、画像判定までの過程を紹介する。さらに、この技術を好中球アルカリフォスファターゼスコア (NAP スコア) 判定に適応した事例を紹介する。

はじめに、画像アノテーションは AI の精度を大きく左右する部分であり、様々なノウハウと人手と労力が必要な地道な作業である。今回使用した末梢血画像は、用手法で染色しスマートフォンを使用して撮影した画像の白血球を矩形として領域指定シタグ付けを行っている。つぎに、学習は物体検出と判定を行う Single Shot MultiBox Detector 法を用い、細胞判定は CNN を用いて実施した。末梢血白血球 6 種類の細胞画像分類では、的確に白血球を検出しており、各細胞 100 枚程度の少数であるものの、好中球以外はほぼ 100% の一致率を示した。一方、好中球の桿状核球と分葉核球は、好中球として判別はできるが、二者を分類するのは困難であった。この結果から、人間においても分類困難な細胞に関しては、AI においても困難である事がわかった。画像認識の可視化法の一つであるヒートマップを用いて学習状況を確認すると、細胞画像のどの部分を見て予測しているのかを確認できる。今回、10 回実施した学習結果では同じ細胞であっても画像の注目部分が異なっていた。機械学習であっても毎回同じ学習結果が得られるわけではなく、学習の初期値やパラメータ設定によって細胞を見るポイントが異なることは注目すべきところと考えられる。学習時間については、高速なコンピュータを使用して 2 時間ほどを要したが、学習後は約 1 秒程度で白血球を検出し分類することが可能である。機械学習も人間と同様、学習には非常に時間を要するが、学習してしまうと非常に高速に分類処理が可能である。骨髓像の白血球分類では、提示した細胞数が少ないことや骨髓芽球、前骨髓球、後骨髓球が加わることで、分類性能は低下した。細胞が成長する過程での変化を捉えるのは人間も機械も困難であることを再確認することができた。

NAP スコアへの応用では、0 型から V 型までの 6 グレードの内、一ランク誤差程度に収まることのできたことから、学習細胞数を増やし精度を上げると実用的なレベルまで引き上げられる可能性が示唆された。

血液像のように細胞の重積がない画像分類に関しては、健常人の血液像であれば比較的容易に分類できる。ただし、炎症性細胞や白血病のような異常細胞の検出となると相当数の細胞提示による学習が必要になる。また、私の実験では、提示する画像サイズがまちまちである場合には学習精度が大きく低下することも分かっており、データを集める際の標準化が必要である事も重要なポイントといえる。

連絡先 - 047-354-3311