

深層学習を用いた画像による微生物の判別システムの開発

長岡工業高等専門学校 非会員 ○新部陽向
非会員 土田勝範 齊藤那生, 正会員 川上周司

1. はじめに

自然界には非常に多様で未知の微生物種が混在しており, 単一の種を同定することやその共生関係, 相互関係を理解することは困難である. また, 微生物種の同定作業は時間と労力も伴うことから迅速な微生物種同定技術が求められる.

近年, 画像解析や分類の分野で成果を上げている深層学習を用いたモニタリング技術が数多く報告されているが, 微生物種への同定作業へ適用された事例は未だ少ない. 我々は深層学習を用いた画像解析の技術を微生物の同定作業に応用することで, より手軽で効率的に微生物を判別する技術を開発できるのではないかと考えた. 本研究では, 画像解析により微生物種の同定を行う深層学習モデルを開発することを目的とした.

2. 方法

2.1 データセットの作成

対象とする微生物は, *Pseudomonas*, *Acinetobacter olevatorans* strain DR1, *Acinetobacter tjernbergiae* strain DSM 14971, *Acinetobacter geminorum* strain J00019 の 4 種とした. また, 微生物の微細な表面構造を撮影する為, 走査電子顕微鏡(Scanning Electron Microscope: SEM, 日本電子株式会社, JSM-6060LV)を使用した. 微生物の形態観察に適した SEM 試料台に微生物を含む懸濁液を滴下し, フィルター越しに水分のみを吸引した. 当該試料は絶縁体の為, SEM に試料を設置する前に白金でイオンパッタを行い, 導電性を付与した. 画像の撮影後, 得られたデータ中の微生物の座標を示すアノテーション加工を label Img を用いて施した. 画像をデータセットとして作成し, 教師データ各種 100 枚のうち, 訓練データ 70 枚, 検証データ 20 枚, テストデータ 10 枚とした.

2.2 深層学習モデルの作成

本研究では, 深層学習モデルとして物体検出アルゴリズムである You only look once (YOLO)と Convolutional Neural Network (CNN) の 2 パターンを採用した. YOLO は, 画像中の物体検出と分類を同時に行うことができ, その他の物体検出システムの中でも比較的高速処理が可能である. また, CNN を用いた学習では, 予め大量の画像データで学習されている AlexNet¹⁾モデルを使用し, 転移学習を行った.

2.3 深層学習モデルの性能評価

深層学習モデルの学習が終了した後, 未知データを用いて, 深層学習モデルが正確に対象を検出でき, 高い信頼度で識別しているかを評価した. さらに, CNN を用いた解析では, 正解率からの評価に加え深層学習モデルが画像のどこを評価し, 分類を行っているかを観察するために Gradient weighted Class Activation Mapping (Grad-CAM)²⁾という手法を使用した.

3. 結果および考察

3.1 YOLO による微生物種の同定の検討

YOLO を用いた解析では, 2 組の異なる種同士の判別と異なる属同士の判別を行い計 6 組の解析を行った. ここで, 閾値は 0.20 とし, 20%未満の信頼度によって検出された結果は表示しなかった. 検出結果を図 1, 図 2 に示す. 異なる属同士の判別では, *Acinetobacter* 属は比較的多くの個体を判別する事ができた. 特に個体の輪郭が鮮明で形状を捉え易いものは信頼度が 60%~80%であった. これは検出した個体の中でも高い値を示した. しかしながら, 組み合わせによっては検出すらできていないものもあった. *Pseudomonas* はいくつか検出しているものの信頼度が 20%~30% と低く, 誤検出も見られた.

また, 異なる種同士の判別においては, 画像に

よって微生物の形態が不鮮明なものもあり、それらに関してはほとんど検出できなかった。また、検出できた中では種を間違える誤検出が多く見られた。

3.2 CNNによる解析

個々の微生物の特徴だけでなく画像全体の特徴から学習を行うためにCNNによる解析を行った。解析の結果、テストデータに対する正解率96.2%と高い値を示した。不正解であった3.8%のデータは、形状の似た異なる種同士のものであった。形状が楕円状と棒状で区別できる異なる属同士では、全て正確に判別することができていた。

次に Grad-CAM を用いて深層学習が情報を抽出している領域について調査した(図3)。図からCNNモデルは背景と物体を区別しているだけでなく、個体の形状が鮮明な箇所から重点的に情報を抽出している事が示唆された。

4. 考察

本研究の結果から、CNNを用いて解析を行った方がYOLOを用いた解析よりも認識精度が向上した。これは、データセットの構造的な違いが原因だと考える。YOLOのデータセットは、画像内の微生物個々にアノテーションを施し、それぞれにラベルを付けている。一方、CNNのデータセットは、画像全体に1つのラベルを割り当てている。従って、YOLOは画像内の特定の要素を個別に認識することができるが、複数の要素間の関連性を把握するのは難しい。それに対し、CNNは複数の要素間の関連性を把握することができる。この差が認識精度に影響を与えたと考えている。YOLOを用いた解析結果において認識性能が低下した要因の一つとして、試料の作成方法も挙げられる。試料作成の過程でナノパーコレーターを使用すると懸濁液を濾過するプロセスが発生し、これによって微生物が密集してしまった。密集した微生物はアノテーション加工が困難である。認識精度を向上させるためには懸濁液の濃度を低くするほか、別の試料作成方法を検討する必要があると考えられる。

また、解析全体を通して使用した画像の不明瞭な部分は信頼度が低く、検出されにくいことから

画質が認識精度に影響することが明らかになった。

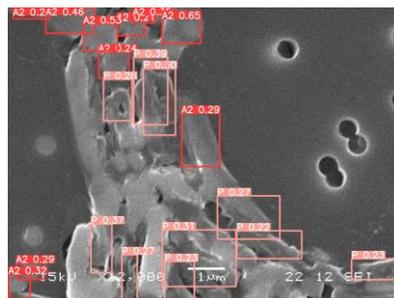


図1 YOLOを用いた異なる属同士の判別結果の例

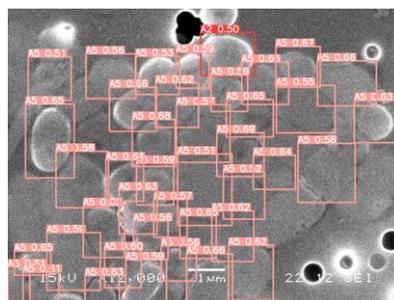


図2 YOLOを用いた異なる種同士の判別結果の例

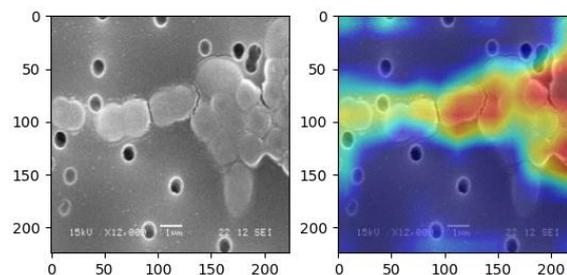


図3 Grad-CAM解析の結果。

左: 用いた教師データ画像, 右: ヒートマップ画像。

5. まとめ

本研究では画像解析により微生物の判別を行うシステムを検討した。CNNを用いた転移学習による解析では高い正解率を得たが、YOLOを用いた解析では更なる改善が必要だと明らかになった。今後の研究では、引き続き微生物判別システムの開発に向け、認識精度の向上を目指す。

参考文献

- 1) Krizhevsky *et al.*, *Advances in Neural Information Processing Systems*, Vol. 25, 2012.
- 2) Selvaraju, R.R., Cogswell, M. *et al.* "Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization", arXiv,(revised 3 Dec 2019)

