

尿路上皮細胞診システムにおける Semantic Segmentation の改善

Improvement of Semantic Segmentation on Urine Cytology Systems

橋本 明空[†] 大野 将樹[‡] 尾矢 剛志[§] 常山 幸一[§] 獅々堀 正幹[‡]
Akira Hashimoto Masaki Oono Takeshi Oya Koichi Tsuneyama Masami Shishibori

1. はじめに

尿路上皮細胞診システムでは、診断基準となる細胞核と細胞質へ面積比 (N/C 比) を算出するために、深層学習に基づく Semantic Segmentation モデルが導入されている。巨島ら[1]は U-net[2]と呼ばれる深層学習モデルの損失関数に円形度と呼ばれる閉領域の真円度合いの尺度を導入し、細胞診システムに特化した Semantic Segmentation モデルを提案した。このように近年は、深層学習を用いた尿路上皮細胞診システム[3]が開発されているが、精度低下に繋がる要因がいくつか確認されている。従来手法では細胞画像を背景、細胞質、細胞核の 3 クラスの領域に分割するために、3 クラスの特徴を同時に学習する必要がある。通常、分類問題では分類クラス数が少ないほど精度向上が望める。

そこで本研究では、細胞画像を、細胞核と背景、及び細胞全体と背景といった 2 パターンの 2 クラス分類を行い、2 パターンの分類結果を合成することで最終的な領域分割画像を作成する。ただし、単純に合成すると、細胞学上矛盾した結果 (例えば、細胞質の存在しない細胞核だけの領域) が作成されるため、2 パターンの分類結果に関連性を持たせる手法を提案する。本手法では、細胞核と背景を最初に分類し、元画像において細胞核と判定された領域の輝度値を低下させる修正を施した後、修正後の画像を細胞全体と背景に分類することで問題点を改善する。

2. 2 クラス分類による合成手法の領域分割

3 クラス分類による問題を改善するため、細胞画像を細胞核と背景、細胞全体と背景という 2 パターンの 2 クラス分類を実施し、最終的にそれらの分類結果を合成して 3 クラスの領域分割画像を生成する。まず、細胞核と背景の学習と分類を行った後、細胞全体と背景の学習と分類を実施する。そして細胞全体と背景の分類結果と、細胞核と背景の分類結果の細胞核部分が重なる箇所のピクセル値を変更することで、合成を実現している。

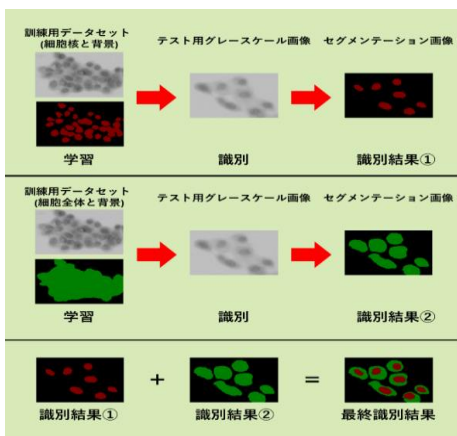


図 1 合成手法の流れ

3. 提案手法

合成手法では、細胞核だけが単体で細胞質に囲われずに存在するといった誤分類を誘発した。この誤分類は、細胞全体と背景の 2 クラス分類の段階で、細胞核領域の輝度値が高いサンプルでは、細胞核領域を背景と誤認識することに起因している。この問題点を解決するためには、細胞核の領域を背景ではなく細胞全体と認識させる必要がある。そこで、細胞核と背景の 2 クラスで細胞核と認識された領域に対して、現画像の同じ部分の輝度を低下させる手法を導入した。以下に核を強調したデータの作成手順を示す。まず、細胞核と背景の 2 クラス分類を実施して、その分類結果①にグレースケール化を施し、細胞核を抽出する。次に、先の結果を用いて入力画像の細胞核部分の輝度値を低下させることで、細胞核が強調された入力画像を生成する。この入力画像を用いて、細胞全体と背景の 2 クラス分類を実施し、この過程で得た分類結果②と分類結果①を合成して 3 クラスの領域分割画像を生成する。

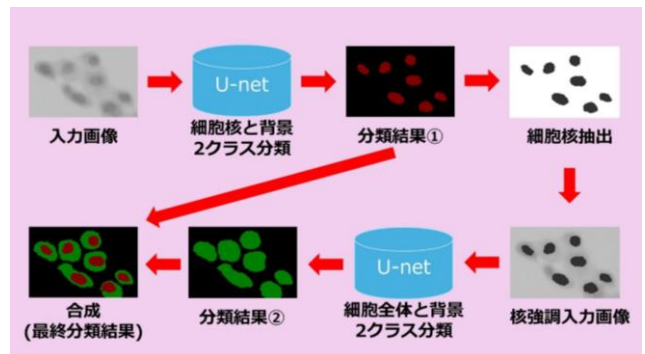


図 2 提案手法の流れ

識別の際に細胞核が強調された画像を用いるため、学習の際にも細胞核が強調された画像を使用する必要がある。学習用画像の細胞核の強調は細胞核と背景の学習用画像を用いる。以下に核を強調した学習用画像の作成手順を示す。

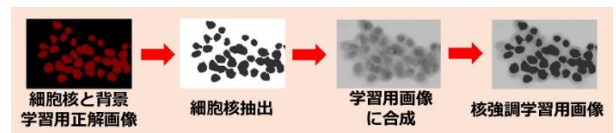


図 3 核強調学習用画像生成の流れ

[†]徳島大学大学院創成科学研究科, Graduate School of Sciences and Technology for innovation, Tokushima University

[‡]徳島大学大学院社会産業理工学研究部, Graduate School of Technology, Industrial and Social Science, Tokushima University

[§]徳島大学大学院医歯薬学研究部, Graduate School of Biomedical Sciences, Tokushima University

4. 評価実験

徳島大学大学院医歯薬学研究部分子病理学研究室の常山幸一教授、尾矢剛志准教授らから提供された尿細胞画像を使用した。また学習に使用する教師画像は上記先生方に確認済みのデータを用いた。また、データ 136 枚をランダムに学習用 108 枚、評価用 28 枚に振り分けたものを 5 パターン作成して、5 回交差検定により検証した。学習用画像に対しては、上下左右反転、ごま塩ノイズ、ガンマ補正、コントラスト調整、グレースケール濃淡調整による 6 種類のデータ拡張により画像枚数を 64 倍にした。学習時のバッチサイズは 48 で統一した。評価指標は、再現率(Recall)、適合率(Precision)、F 値(F-measure)を用いる。各項目の算出式を以下に示す。

$$Recall_k = \frac{\text{クラス } K \text{ と予測された画素数中のクラス } K \text{ の正解画素数}}{\text{クラス } K \text{ のすべての正解画素数}}$$

$$macro - Recall = \frac{\sum_{k=1}^3 Recall_k}{3}$$

$$Precision_k = \frac{\text{クラス } K \text{ と予測された画素数中のクラス } K \text{ の正解画素数}}{\text{クラス } K \text{ であると予測された画素数}}$$

$$macro - Precision = \frac{\sum_{k=1}^3 Precision_k}{3}$$

$$F - measure_k = \frac{2 \times Recall_k \times Precision_k}{Recall_k + Precision_k}$$

$$macro - F - measure = \frac{\sum_{k=1}^3 F - measure_k}{3}$$

4.1 実験結果

従来手法と合成手法、提案手法における再現率(Recall)、適合率(Precision)、F 値(F-measure)を比較した。表 1 に交差検定実験の評価値の平均を示し、図 4 にセグメンテーションの実行結果例を示す。

表 1 交差検定の評価値の平均

	従来	合成	提案手法
Recall (%)	82.7	83.0	84.8
Precision (%)	82.9	82.9	85.2
F-measure (%)	80.6	80.9	83.5

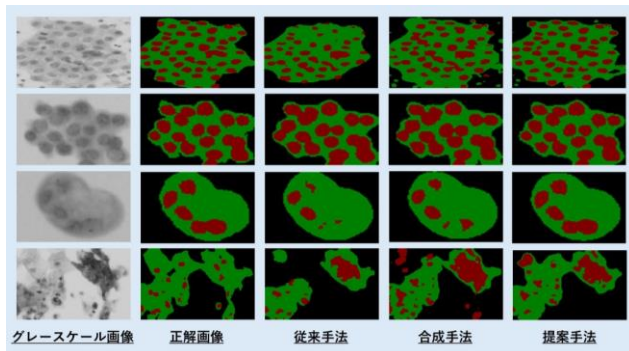


図 4 セグメンテーション実行結果例

4.2 考察

従来手法と合成手法を比較すると、ほとんど精度向上は確認できなかった。これは細胞核と背景、細胞全体と背景の分類を個別に実施しているため、細胞核と背景の分類結果が、細胞全体と背景の分類の際に考慮されていないことが起因していると考えられる。これにより、図 4 のセグメンテーション実行結果例の 4 行目に示す通り、合成手法では細胞核領域が単独で(細胞質に囲われずに)分類されるケースがあり、精度低下を招いた。

次に従来手法と提案手法を比較すると、精度の向上が確認できた。このことから細胞核と背景の分類結果と、細胞全体と背景の分類結果の間で、関連性を持たせたことで、より教師画像に近い画像を生成できるようになったと考えられる。実際に、図 4 のセグメンテーション実行結果例の 3 行目に示す通り、従来手法よりも正解画像に近い画像が生成できていることが確認できる。一方で、誤った細胞核の識別がなされた場合も、細胞核と背景の識別で用いる入力用画像にその特徴が合成されるため、識別の際に障害になってしまうことがあった。その結果、正解画像に細胞核が無い部分にも、細胞核を認識してしまい、誤った識別がなされることがあった。

5. まとめ

本稿では、尿路上皮細胞画像を対象にした領域分割モデルにおいて、3 クラス分類の問題点に着目し、2 パターンの 2 クラス分類結果を合成するアプローチをベースとして、領域分割画像を作成した。また、2 クラス分類において、独立に作成した分類結果を合成する際に生じる問題点を明らかにし、その問題点を解決するために、各分類結果に関連を持たせて合成する手法を提案した。評価実験の結果、従来手法や合成手法と比較して、提案手法の精度が向上し、有効性が示された。しかし、細胞核と背景の識別において誤った細胞核の識別がなされた場合も、学習用画像にその特徴が合成されるため、識別の際に障害になってしまうという新たな課題も見つかった。また、本実験は良性細胞に対してのみ実施しているため、悪性細胞についての精度は不明である。今後は、細胞核と背景の 2 クラス識別精度の向上施策の立案と、悪性細胞に対する精度の算出を実施し、問題点を明らかにする予定である。

参考文献

- [1] 巨島弘基他, “細胞診断システムにおける円形度に着目した領域分割の改善”, 2023 年電気学会電子・情報・システム部門大会, PS1-10, 2023.
- [2] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, “U-net: Convolutional networks for biomedical image segmentation”, International Conference on Medical Image Computing and Computer-Assisted Intervention, pp.234-241, 2015.
- [3] Louis J. Vaickus, Arief A. Suriawinata, Jason W. Wei, and Xiaoying Liu, “Automating the Paris System for Urine Cytopathology—A Hybrid Deep-Learning and Morphometric Approach”, Cancer Cytopathology, pp.98-115, 2019.