Research Report

Shahzada Muhammad Ali, Hamza Tariq

Classical Image-Based Cell Segmentation and Phenotype Classification Using Mask-Guided Feature Extraction

Abstract

Cell-level image analysis is a foundational task in biomedical imaging, enabling downstream applications such as phenotype classification, disease characterization, and quantitative pathology. While deep learning approaches dominate recent literature, classical image processing and feature-based learning remain valuable in low-data regimes, interpretable systems, and resource-constrained settings. In this work, we present a complete end-to-end pipeline for cell segmentation, feature extraction, and phenotype classification from microscopy images. Using mask-guided cropping, intensity-based feature extraction, and classical machine learning, we demonstrate a scalable framework capable of aggregating data across multiple patients while maintaining interpretability and evaluation rigor. The pipeline integrates segmentation validation via Intersection-over-Union (IoU) and classification analysis using sensitivity and specificity, providing a transparent baseline for biomedical image analysis.

1. Introduction

Automated analysis of microscopic cell images is a core problem in biomedical image processing, with applications ranging from hematology to cancer diagnosis. A typical workflow involves three tightly coupled stages: segmentation of individual cells, extraction of discriminative features, and classification of cell phenotypes.

Although deep learning methods have achieved impressive performance, they often require large annotated datasets and provide limited interpretability—both of which are problematic in clinical and academic settings. In contrast, classical pipelines allow fine-grained control over each stage, explicit error analysis, and clearer biological interpretation.

This work investigates the following question:

To what extent can a classical, interpretable image processing pipeline achieve reliable cell-level classification when segmentation and annotation metadata are available?

2. Problem Definition

Given:

- Microscopy images containing multiple cells
- Ground-truth segmentation masks and bounding box annotations
- Cell-level phenotype labels

The objectives are to:

1. Segment individual cells from full microscopy images
2. Extract biologically meaningful, mask-guided features
3. Train a supervised classifier for cell phenotype prediction
4. Evaluate segmentation and classification performance across multiple patients

3. Dataset and Assumptions

Each patient dataset consists of:

- A full microscopy image
- A corresponding segmentation mask
- A CSV file containing bounding boxes and phenotype labels
- Cropped single-cell images stored separately

The pipeline assumes:

- Segmentation masks are available at the image level
- Bounding boxes provide approximate cell localization
- Phenotype labels are categorical and mutually exclusive

4. Methodology

4.1 Mask-Guided Cell Extraction

Using bounding box coordinates from annotation CSV files, individual cell masks are cropped from full segmentation masks. This ensures:

- Pixel-level alignment between cell image and mask
- Elimination of background noise during feature extraction
- Consistency across patients

Each cropped mask is saved independently, enabling modular reuse in later experiments.

4.2 Feature Extraction

For each cell, features are extracted using the corresponding mask to isolate biologically relevant pixels.

The following features are computed from the green color channel:

- **Cell area** (number of foreground pixels)
- **Mean intensity**
- **Standard deviation of intensity**

These features were selected due to:

- Their biological relevance in fluorescence and histology imaging
- Robustness to limited sample sizes
- Interpretability for downstream analysis

Mask–image shape mismatches are explicitly detected and excluded, ensuring data integrity.

4.3 Classical Segmentation of Full Images

To validate segmentation quality independently of annotations, full images are segmented using:

- Grayscale conversion
- Otsu's adaptive thresholding
- Morphological filtering to remove small artifacts
- Connected-component labeling

Each connected region is treated as a candidate cell.

4.4 Segmentation Evaluation via IoU

Predicted cell regions are evaluated against ground-truth bounding boxes using Intersection-over-Union (IoU). A match is considered valid if IoU > 0.5.

This provides:

- Quantitative validation of segmentation quality
- An explicit link between pixel-level segmentation and annotation-level ground truth

4.5 Classification Model

A Random Forest classifier is trained on extracted features due to:

- Robustness to small datasets
- Ability to model non-linear feature interactions
- Built-in feature importance estimation

The dataset is split into training and testing subsets using stratified sampling where possible.

5. Multi-Patient Aggregation

To assess scalability and generalization, the pipeline is extended to aggregate data across:

- 5 patients
- Subsequently, all available patients (up to 28)

Features and labels are concatenated across patients, enabling:

- Cross-patient learning
- Analysis of class imbalance
- Robust performance estimation

6. Evaluation Metrics

Classification performance is evaluated using:

- Accuracy
- Confusion matrices
- Class-wise sensitivity (recall)
- Class-wise specificity

These metrics are particularly relevant in biomedical settings, where false positives and false negatives have asymmetric consequences.

7. Results and Analysis

The pipeline successfully:

- Extracts valid cell-level features across multiple patients
- Trains a stable classifier despite class imbalance
- Provides interpretable performance metrics
- Identifies feature importance trends (e.g., dominance of cell area or intensity statistics)

Visualization of segmentation results and class distributions further supports qualitative evaluation.

8. Discussion

This work demonstrates that classical image processing pipelines remain viable and informative, particularly when:

- Annotation metadata is available
- Dataset sizes are limited
- Interpretability is a priority

Rather than optimizing for peak accuracy, the system emphasizes transparency, modularity, and reproducibility—qualities often undervalued but essential in research contexts.

9. Limitations and Future Work

Limitations include:

- Reliance on handcrafted features
- Sensitivity to segmentation quality
- Absence of texture or shape descriptors

Future work could explore:

- Morphological and texture-based features
- Hybrid classical–deep learning pipelines
- Weakly supervised segmentation refinement
- Graph-based modeling of cell–cell interactions

10. GenAI Usage Disclosure

We employed ChatGPT to assist in rephrasing the report for improved clarity. All core content, including research design, data analysis, and result interpretation, was conducted without the aid of generative AI tools.