

Malignant Cell Detection Using Deep Learning and Segmentation Techniques

Neeraj Dandamudi

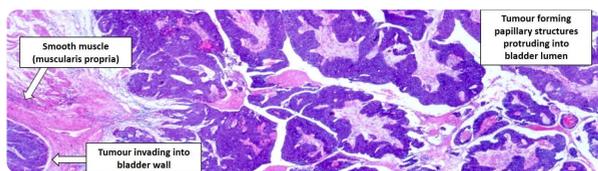
Emerging Diagnostic and Investigative Technologies, Department of Pathology, Dartmouth Hitchcock Medical Center

ABSTRACT

- Bladder cancer is a leading cause of cancer-related morbidity, with cytology screening playing a vital role in detection.
- Manual nuclear-to-cytoplasmic (N/C) ratio estimation is labor-intensive, subjective, and prone to variability.
- This study compares deep learning and traditional segmentation approaches for malignant urothelial cell detection.
- We evaluate segmentation accuracy and N/C ratio agreement with pathologist-provided ground truth, and assess clinical validity for malignancy prediction.

INTRODUCTION

Figure 1: Urothelial Carcinoma Closeup



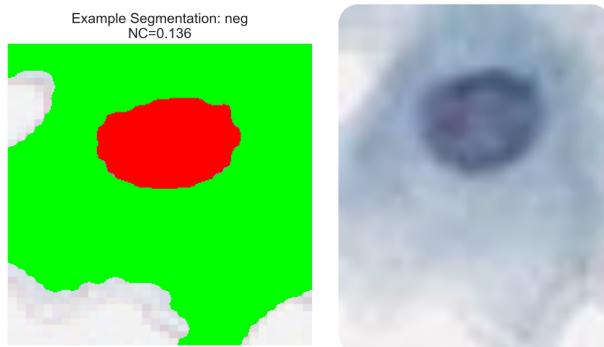
- Bladder cancer:** Sixth most common cancer in the United States, with high recurrence rates and reliance on invasive, costly, and subjective screening methods such as urine cytology and cystoscopy.
- Nuclear-to-cytoplasmic (N/C) ratio:** An established cytological feature where increased nuclear size relative to cytoplasm often indicates malignancy. Serves as an objective, quantitative measure that can complement pathologist assessment.

Current challenges:

- Manual estimation of N/C ratio is time-consuming, variable between observers, and prone to inconsistency.
- Traditional segmentation methods are computationally efficient but struggle with boundary precision in complex or noisy images.
- Deep learning approaches produce more accurate masks but require more computation and training data.

Figure 2 & 3: Whole slide image + segmentation

- Medical imaging context:** Cytology images (128x128 px single-cell crops) from urine specimens provide the spatial resolution necessary to delineate nuclear and cytoplasmic regions. Accurate segmentation is critical for reliable N/C ratio measurement.



METHODS

- Goal:** Evaluate the effectiveness of traditional and deep learning-based segmentation methods in accurately identifying malignant urothelial cells from cytology whole slide images (WSIs) of urine specimens.
- Experimental Design:**
 - WSIs collected from bladder cancer cytology cases, stained using standard cytopathology protocols.
 - Regions of interest (ROIs) manually annotated for nuclei and cytoplasm by cytopathologists.
 - Annotations used to generate ground truth segmentation masks for nucleus and cytoplasm.
 - Dataset split into training, validation, and test sets for evaluation.
 - Multiple segmentation methods compared, including traditional image processing and deep learning approaches. Comparison is done based on segmentation performance and accurate Nuclear-to-Cytoplasmic (N/C) ratio calculation.
 - Calculated N/C ratios for each segmented cell are tested as a quantitative malignancy indicator using one of the selected segmentation techniques.
- Segmentation Methods:**
 - Traditional Approaches:**
 - Intensity Thresholding – segmented nuclei based on pixel intensity ranges.
 - K-Means Clustering – unsupervised clustering in color space to separate nuclei and cytoplasm regions.
 - Three Random Forest classifiers with/and without optimization, and GLCM/Multiscale Features.
 - Deep Learning approaches:**
 - U-Net and Feature Pyramid Network (FPN) architectures with ResNet-34 encoder backbones.

RESULTS

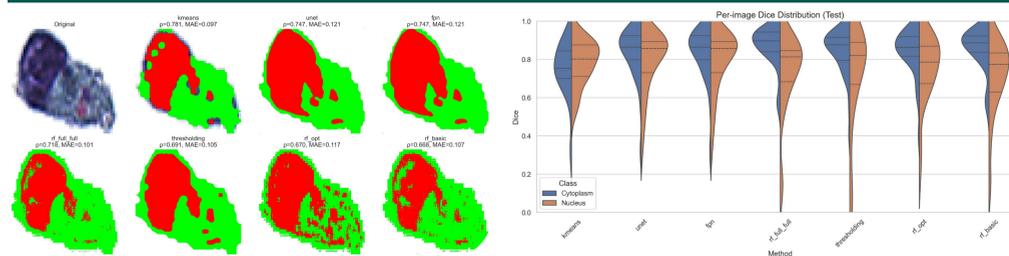


Figure 4 & 5: Segmentation quality comparison and violin chart of mask quality over the dataset

Segmentation Accuracy:

- Deep learning methods achieved the highest Dice and IoU scores (U-Net mIoU ≈ 0.77151), producing visually cleaner masks with accurate borders.
- Traditional methods underperformed in segmentation metrics but preserved size-based nuclear features more faithfully.
- In comparison to traditional methods, excluding K-means, deep learning methods also produce fewer low quality masks over the entire dataset.

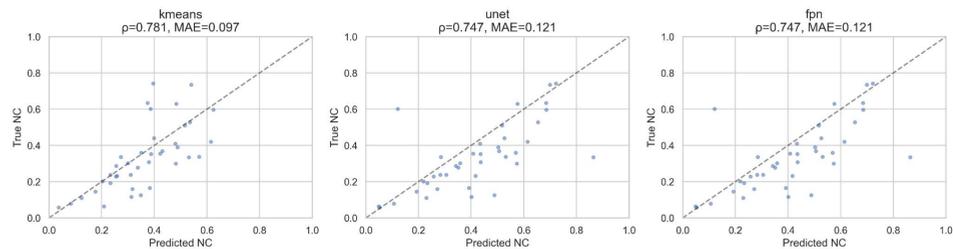


Figure 6: Scatterplots of the N/C predicted vs true correlation for top 3 performing segmentations.

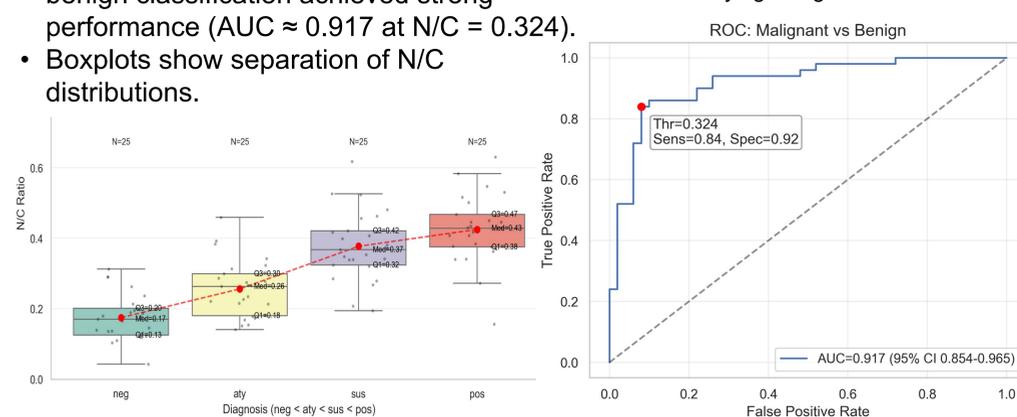
N/C Ratio Agreement:

- K-means achieved the highest correlation with ground truth N/C ratios (Spearman ≈ 0.781), despite lower segmentation scores.
- Deep learning models slightly overestimated nuclei boundaries, introducing N/C bias. This results in a higher average mean error than K-means.

Clinical Validity:

- Using K-means N/C ratios, malignant vs. benign classification achieved strong performance (AUC ≈ 0.917 at N/C = 0.324).
- Boxplots show separation of N/C distributions.

Figure 7 & 8: ROC Curve and Box Plot showcasing N/C ratio performance at identifying malignant cells



CONCLUSION

Potential for Clinical Impact:

- Demonstrates an approach to quantify the (N/C) ratio from cytology images using both deep learning and traditional segmentation methods.
- Identifies K-means clustering as a strong performer for clinical malignancy prediction despite lower segmentation metrics, achieving AUC ≈ 0.917 .
- Suggests that automated N/C ratio estimation can complement pathologist assessment and reduce subjectivity in bladder cancer screening.

Limitations:

- Traditional methods lack boundary precision in complex cellular images.
- N/C overestimation in deep learning outputs may biased predictions.
- Study limited to a single dataset. Dataset size limited to 300 images.

Future Directions:

- Explore deep learning models to overcome N/C accuracy issues as they are promising due to their potency at creating segmentation masks of high quality.
- Expand evaluation across larger, multi-institutional cytology datasets to test generalizability.
- Investigate hybrid models combining boundary precision of deep learning with size fidelity of traditional methods.
- Optimize deep learning architectures to reduce N/C bias without losing segmentation quality.
- Integrate automated segmentation into end-to-end clinical workflows for real-time malignancy screening.