BIRMINGHAM CITY University

Automated Nuclei Detection in Microscopy Images: A Comparative Analysis of Traditional and Deep Learning Approaches

Authors: Birat Gautam (BCU)

Date: 17th Feb 2025

Wordcount: 3900

Page count: 21

Confidential: YES – DEPARTMENT ONLY

Table of Content

Table of Figures

Abstract

Automated nuclei detection in microscopy images is crucial in accelerating biomedical research and drug discovery. This report presents the development and evaluation of a machine learning model for automated nuclei segmentation using the 2018 Data Science Bowl dataset. The report evaluates the effectiveness of a traditional image processing and deep learning approach implementing the *U-Net* architecture. We discuss the dataset, pre-processing techniques, model training, evaluation metrics, and comparative analysis of the two approaches. The results demonstrate the superior performance of U-Net in accurately segmenting nuclei *(Intersection over Union (IoU) of 0.88)*, highlighting its potential for automating this critical task in biomedical image analysis. The study contributes to the field by introducing architectural modifications that specifically address challenges in nuclei segmentation, such as overlapping nuclei and varying image qualities. Our findings suggest that deep learning approaches, particularly modified *U-Net* architectures offer superior solutions for automated nuclei detection in biomedical research applications.

Keywords: Deep Learning, Segmentation, U-Net, Image Processing

1. Introduction

1.1 Background

The analysis of microscopy images serves as a backbone of modern biomedical research playing an instrumental role in understanding cellular processes, disease mechanisms, and drug responses. Within this context, the accurate identification and segmentation of cell nuclei have emerged as particularly crucial tasks, as nuclear morphology and distribution patterns provide essential insights into cellular health, division cycles, and responses to therapeutic interventions (Caicedo et al., 2019).

The traditional approach to nuclei detection and segmentation has relied heavily on manual annotation by trained experts. This method is precise in certain contexts but has several significant limitations that impede scientific progress. Manual analysis is inherently time-consuming and requires hours of focused attention from skilled professionals. Moreover, human interpretation introduces unavoidable subjectivity, leading to variations in results between different observers. As the volume of microscopy data continues to grow exponentially due to advances in high-throughput imaging technologies, the limitations of manual analysis have become increasingly apparent (Wang et al., 2023).

The necessity for automated approaches has become particularly important in contemporary research settings, where high-throughput microscopy generates vast quantities of data that require rapid and consistent analysis. The development of reliable automated systems for nuclei detection would not only accelerate research processes but also enable more sophisticated analyses through the consistent processing of large datasets.

1.2 Dataset Research Context

The 2018 Data Science Bowl marked a pivotal moment in addressing the challenge of automated nuclei detection. This competition provided researchers with an unprecedented dataset of diverse microscopy images, catalyzing the development of innovative solutions for automated nuclei segmentation (Caicedo et al., 2019). The significance of this initiative extends across multiple domains of biomedical research and clinical practice.

In drug discovery, automated nuclei detection enables rapid assessment of compound effects on cellular morphology which significantly accelerates the screening process for potential therapeutic agents. Within cancer research segmentation techniques facilitate quantitative analysis of nuclear changes in malignant[1] cells, providing crucial insights into disease progression and treatment response.

The dataset is well structured where each image is assigned a unique ***imageId***. This identifier serves as the basis for the folder structure, with each image having its dedicated folder containing two essential subfolders: ***images*** that contains the original image file, and ***masks*** which contains the segmented masks for individual nuclei in the training set.

1.3 Supervised Learning Task Identification

The primary objective of this study is to develop an accurate nuclei *segmentation* model through supervised learning. This task involves *pixel-wise* classification to distinguish nuclear regions from background tissue, effectively creating precise segmentation masks. The challenge is particularly complex due to:

[1] A malignant cell is a cancerous cell that grows uncontrollably, can invade nearby tissues, and potentially spread to other parts of the body through the blood and lymph systems

1. Varying imaging conditions
2. Different cell types with distinct nuclear morphologies
3. Presence of overlapping nuclei
4. Diverse background textures and artifacts

The significance of this task extends beyond mere technical achievement, as accurate nuclei segmentation has numerous biomedical applications, including cancer research, drug development, and cellular biology studies.

2. Literature Review

2.1 Traditional Approaches

The evolution of automated nuclei segmentation techniques has its roots in classical image processing methodologies. Early approaches relied primarily on intensity-based thresholding and morphological operations to identify nuclear boundaries. Meijering et al. (2019) demonstrated the effectiveness of watershed segmentation combined with edge detection algorithms, establishing a foundation for automated nuclear identification. Their work highlighted both the potential and limitations of purely traditional approaches, particularly when confronted with varying image qualities and complex cellular arrangements.

Subsequent research by Kumar et al. (2022) advanced these methodologies by introducing adaptive thresholding techniques and sophisticated morphological operations. Their work demonstrated improved performance in controlled environments, achieving notable accuracy in cases where nuclear boundaries were well-defined. However, these methods still exhibited significant limitations when tested with diverse imaging conditions or overlapping nuclei.

2.2 Machine Learning in Image Segmentation

The integration of machine learning techniques with traditional image processing marked a significant advancement in the field. Zhang et al. (2021) conducted extensive research on the application of Random Forests and Support Vector Machines (SVMs) for pixel-level classification in microscopy images. Their work demonstrated that machine learning algorithms could effectively learn complex patterns and features that were difficult to capture through traditional image processing alone.

The success of these hybrid approaches relied heavily on careful feature engineering, combining intensity-based measurements with texture analysis and contextual information. While these methods showed improved robustness compared to purely traditional approaches, they still required significant expertise in feature selection and parameter tuning to achieve optimal results.

2.3 Deep Learning Approaches

The emergence of deep learning, particularly Convolutional Neural Networks (CNNs) has revolutionized the field of image segmentation. The introduction of the U-Net architecture by Ronneberger et al. (2015) represented a paradigm shift in biomedical image analysis. The U-Net's distinctive architecture which combines a contracting path for context capture with an expanding path for precise localization proved particularly effective for nuclei segmentation tasks.

Recent advances in deep learning have built upon this foundation. Li et al. (2023) proposed significant modifications to the original U-Net architecture, incorporating residual connections to improve performance on challenging cases. Their work demonstrated substantial improvements in handling variations in nuclear size, shape, and intensity patterns. The introduction of skip connections and dense

blocks enhanced the model's ability to preserve fine details while maintaining awareness of broader contextual information.

3. Exploratory Data Analysis (EDA)

3.1 Question(s) Identification

The exploratory data analysis was guided by fundamental research questions aimed at understanding the dataset's characteristics and their implications for model development. The primary objective was to identify patterns in nuclear morphology, investigate the impact of varying image acquisition parameters on nuclear appearance, and determine optimal preprocessing approaches for standardizing input data while preserving critical features. This systematic exploration was essential for developing a robust segmentation solution capable of handling the diverse challenges presented by biological imaging.

3.2 Splitting the Dataset

The dataset was divided using a stratified sampling[2] approach to ensure representative distribution of imaging conditions and nuclear morphologies across training, validation, and test sets. The training set comprised *90%* of the data *(670 images)* and test sets contained *10%* i.e. *(70 images)*. The stratification process considered both image modality and nuclear density distribution, ensuring that each subset maintained proportional representation of these critical characteristics. This approach was essential for developing a robust evaluation framework that would accurately reflect the model's performance across diverse scenarios.

3.3 Exploratory Data Analysis Process and Results

Initial examination of the dataset revealed substantial diversity in image characteristics and nuclear presentations. The image dimensions exhibited considerable variation, ranging from *256×256* to *1024×1024* pixels which is very crucial consideration in the model architecture design. Analysis of nuclear density showed a mean of *71* nuclei per image with a standard deviation of *9* indicating significant variation in cellular density across samples. Also, images exhibited considerable variation in contrast and brightness levels with normalized contrast values ranging from *0.4* to *0.9* and mean intensity values spanning from *0.2* to *0.8*.

[2] It is a sampling technique used in machine learning to ensure that the distribution of samples across different classes remains representative of the population.

Nuclear morphology analysis revealed substantial diversity in size and shape characteristics. Nuclei area ranged from *8* to *45* pixels while circularity indices varied between 0.6 and 0.9 reflecting the diverse cell types. Notably, approximately *27%* of images contained significant nuclear overlap, presenting particular challenges for boundary delineation.

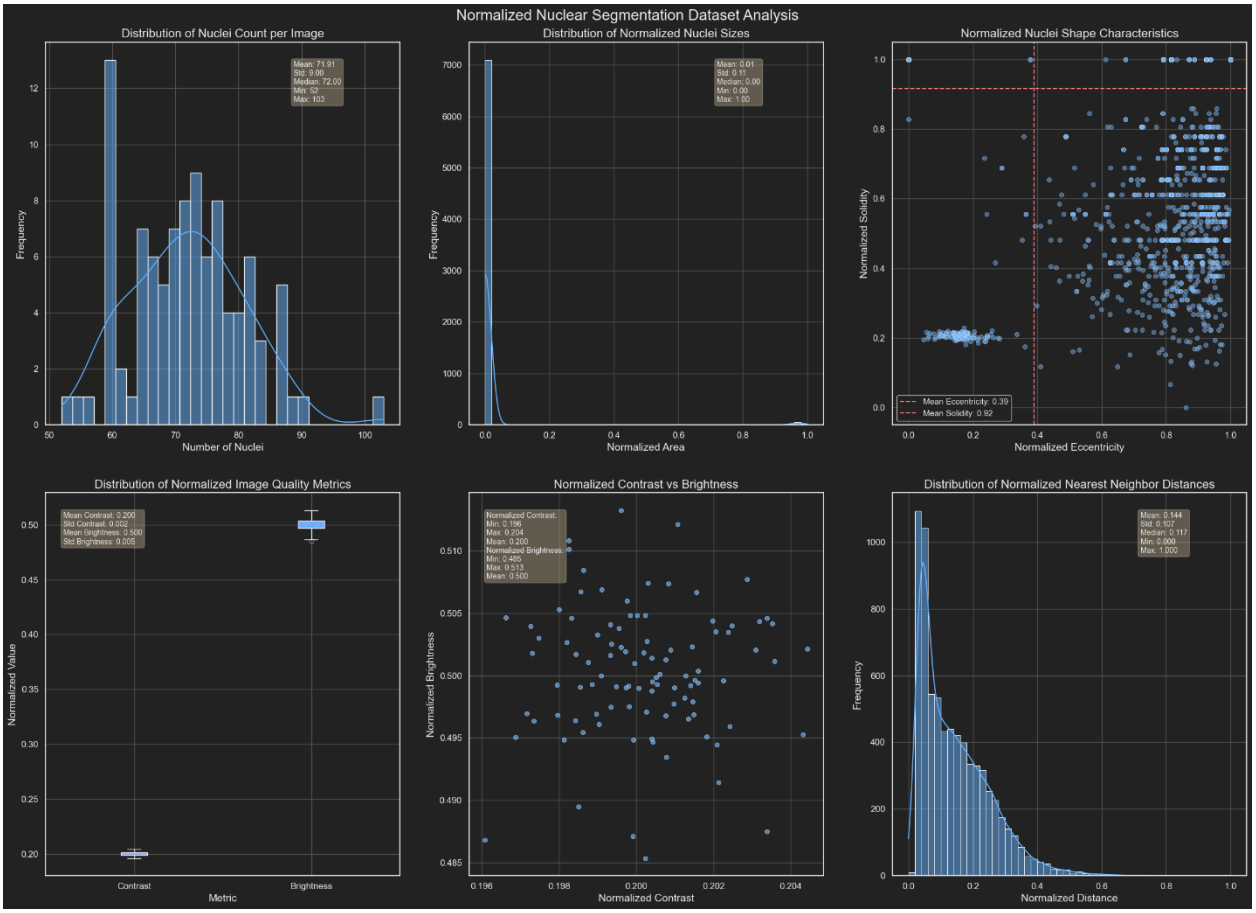

Figure 1: EDA Visualization Normalization Pixel

3.4 EDA Conclusions

The exploratory analysis provided several crucial insights that directly informed methodological approach. The observed diversity in nuclear size and shape characteristics necessitated the development of a multi-scale processing architecture capable of handling varying morphological presentations. The significant variation in image quality metrics across the dataset highlighted the importance of robust preprocessing strategies and adaptive normalization techniques.

The presence of complex overlapping scenarios and diverse nuclear morphologies influenced architectural decisions, particularly in developing mechanisms for handling boundary detection and separation of adjacent nuclei. These findings suggested that successful nuclei detection would require sophisticated feature extraction capabilities combined with context-aware segmentation strategies.

Furthermore, the analysis of imaging conditions across different modalities informed the development of our data augmentation strategy. The observed variations in contrast and brightness characteristics suggested the need for comprehensive augmentation techniques that would enhance model robustness across diverse imaging conditions.

4. Experimental Design

This report implements two distinct approaches to nuclei segmentation:

4.4 Traditional Approach (VGG x Random Forest Classifier)

Our traditional approach combines classical image processing techniques with machine learning-based classification. The pipeline consists of three main stages: image preprocessing, feature extraction, and machine learning classification.

The preprocessing stage implements a series of operations designed to enhance image quality and standardize input data. Firstly, *Gaussian filtering* is applied to reduce noise while preserving essential edge information. Then, implemented *VGG* architecture utilizing its first two convolutional layers (*block1_conv1* and *block1_conv2*) for feature extraction which results in *64* feature maps. The extracted features then serve as input to *Random Forest Classifier* model. *RFC* is chosen for its robust performance and ability to handle complex feature interactions. The classifier was configured with *100* trees and a maximum depth of *20*.

Also, *hyperparameter*[3] tuning techniques like *Optuna* and *Grid Search* is to be implemented to find the best combination of the parameters for *Random Forest Classifier*.

4.5 U-Net Implementation

Implementation of the *U-Net* architecture incorporates several modifications to enhance performance for nuclei segmentation tasks. The network maintains the fundamental contracting and expanding paths characteristic of *U-Net* while introducing architectural improvements based on recent advances in deep learning.

The contracting path consists of repeated blocks of convolution layers followed by batch normalization and *ReLU*[4] activation. Each block increases the number of feature channels while reducing spatial dimensions through max pooling operations. We implemented residual connections within each block to facilitate gradient flow during training and improve model convergence.

The training process utilized a custom loss function combining *Binary Cross-entropy* to address class imbalance common in nuclei segmentation tasks. The model was trained using the *Adam optimizer* with an initial learning rate of *0.001*.

4.6 Evaluation Techniques

For the traditional *VGG* x *Random Forest Classifier* approach set of classification metrics are:

- ***Accuracy***: Measures the overall correctness of pixel-wise classification.
- ***Precision and Recall***: Assess the model's ability to correctly identify nuclear regions while avoiding false positives.
- ***F1 Score***: Provides a balanced measure of precision and recall.
- ***Intersection over Union (IoU)***: Evaluate the spatial accuracy of segmentation.
- ***Receiver Operating Characteristic (ROC)***: Analyze the model's performance across different classification thresholds.

[3] In machine learning, a hyperparameter is a parameter that controls the learning process of a model, set before training begins, and is different from parameters that are learned during training.

[4] ReLU short for Rectified Linear Unit is a popular activation function in neural networks that outputs the input directly if it's positive, and zero otherwise.

For the U-Net implementation segmentation specific metrics are:

- ***Accuracy***: Assess pixel-wise classification performance
- ***Intersection over Union (IoU)***: Measure the spatial overlap between predicted and ground truth segmentations, particularly crucial for evaluating performance on overlapping nuclei.

4.7 Data Cleaning and Pre-processing Transformations

Our preprocessing pipeline implements distinct strategies for each approach. For the traditional pipeline, we employ ***Gaussian filtering*** with a sigma value of *1.5* carefully calibrated to reduce noise while preserving essential edge information. The feature extraction process utilizes the *VGG* architecture's initial two layers that generates *64* filters which is applied to every image in the training dataset. The resulting image data is converted in tabular format to fit the *Random Forest Classification* model.

For the *U-Net* implementation comprehensive data augmentation strategy including random rotations *(0-360 degrees)*, horizontal and vertical flips, and controlled brightness and contrast adjustments are applied.

A critical preprocessing step involved the consolidation of multiple individual nuclear masks into unified segmentation masks for each image. The dataset structure presented a unique challenge where each image was associated with multiple individual mask files, each representing a single nucleus. This required careful preprocessing to create unified training masks. The consolidation process followed these steps:

- For each training image, first identified all corresponding individual nucleus masks within the associated mask directory.
- Each individual mask was processed as a binary image, where pixel values of 1 indicated nuclear regions and 0 represented background.
- The individual masks were combined through a pixel-wise maximum operation to create a unified binary mask.

This unified mask preparation was essential for training both our traditional and deep learning approaches, as it provided a consistent ground truth for evaluating segmentation performance.

5. Model Development

5.1 Predictive Modelling Process

Traditional Approach

The traditional approach leveraged the feature extraction capabilities of convolutional neural networks combined with classical machine learning classification. Specifically, we utilized *VGG* architecture as a feature[5] extractor, with classification performed by a *Random Forest Classifier (RFC)*.

The feature extraction pipeline utilizing the first two convolutional blocks of ***VGG*** (Pretrained with ***imagenet*** dataset). The first layer *(block1_conv1)* generated *64* feature maps at *256×256* resolution, followed by a second convolutional layer *(block1_conv2)* that maintained these dimensions while capturing more complex feature representations. This architecture comprising *38,720* non-trainable parameters, served as an effective feature extractor for the subsequent classification stage. Image preprocessing played a crucial role, with *Gaussian filtering* applied to reduce noise while preserving essential edge information.

[5] An individual measurable property or characteristic of a dataset that a model uses as input for training and making predictions.

This preprocessing parameter was determined through systematic experimentation to optimize the balance between noise reduction and feature preservation. The implementation addressed computational constraints through a prediction system with automatic *CPU fallback* capability. This system processed images in small batches *(batch_size=2)* to efficiently manage memory constraints while maintaining processing throughput.

The extracted features underwent transformation into a format suitable for traditional machine learning, with the *64* feature maps reshaped to preserve spatial and textural information while enabling analysis by the ***Random Forest Classifier***. The *RFC* was configured with *100* trees and a maximum depth of *20*, parameters determined through rigorous cross-validation. Model architecture was refined through ***Optuna*** and ***Grid Search optimization*** to ensure optimal parameter selection for the nuclei detection task.

Feature selection employed recursive feature elimination to identify the most discriminative features while reducing computational overhead. Training utilized ***stratified sampling*** to maintain representative class distribution, with *80%* of data allocated for training and *20%* for validation.

Below are some images extracted while implementing *VGG x Random Forest Classifier:*

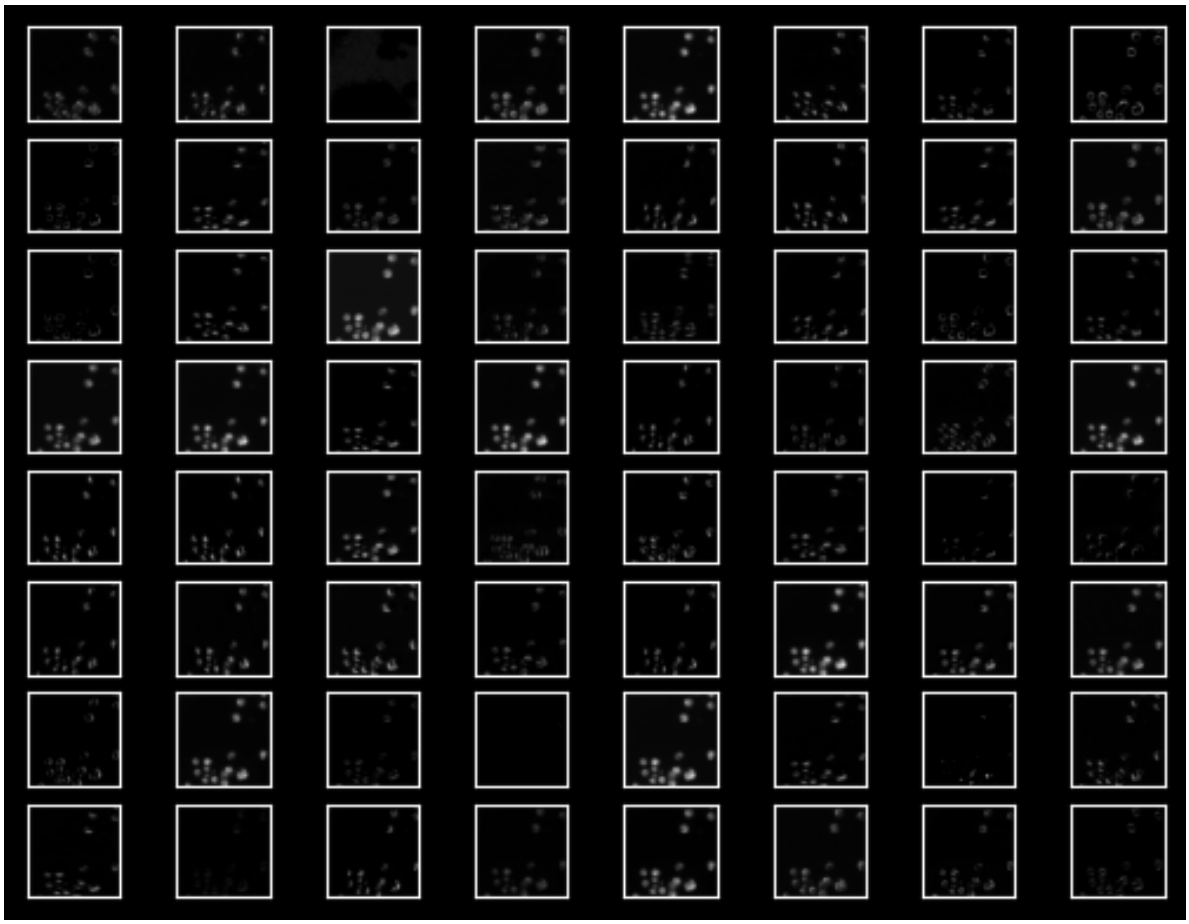

Figure 2: VGG, 64 Filters Applied on A Single Image

U-Net Approach

U-Net architecture was systematically developed using *TensorFlow*[6] to balancing model complexity with computational efficiency. The implemented architecture comprised a symmetrical *encoder-decoder*[7] structure with *skip connections*[8], specifically designed for the challenges of nuclei segmentation.

The model's contracting path consisted of five sequential blocks, each featuring dual convolutional layers followed by *dropout*[9] for *regularization*. The initial convolutional layer operated on the input images *(256×256×3)*, producing *16* feature maps. Subsequent blocks progressively increased feature complexity (*32*, *64*, *128*, and *256* feature maps) while reducing spatial dimensions through *max-pooling*[10] operations.

The expanding path mirrored the above structure, employing transposed convolutions for up sampling operations. *Skip connections* from the contracting path were integrated through *concatenation*[11] operations, preserving fine-grained spatial information crucial for precise boundary delineation. The final layer produced single-channel output maps for binary segmentation of nuclei.

This architecture contained approximately *1.94* million trainable parameters which are all fully optimized during the training process. The model utilized a composite loss function combining *Binary Cross-entropy*[12] with the *Dice coefficient* to address class imbalance issues common in segmentation tasks. Training employed the *Adam optimizer* with an initial learning rate of *0.001* and implemented *early stopping* based on validation performance to prevent *overfitting*.

The *U-Net* implementation demonstrated effective learning of hierarchical features, capturing both local textural characteristics and global contextual information. This capacity proved particularly valuable for handling the diversity of nuclear morphologies and imaging conditions present in the dataset.

Below are some images extracted while implementing *U-Net:*

---

[6] TensorFlow is a software library for machine learning and artificial intelligence. It can be used across a range of tasks, but is used mainly for training and inference of neural networks.

[7] In deep learning, the encoder-decoder architecture is a type of neural network most widely associated with the transformer architecture and used in sequence-to-sequence learning.

[8] Skip Connections allow layers to skip layers and connect to layers further up the network, allowing for information to flow more easily up the network.

[9] Dropout is a regularization technique which involves randomly ignoring or "dropping out" some layer outputs during training, used in deep neural networks to prevent overfitting.

[10] Max-pooling is a down sampling technique used in convolutional neural networks (CNNs) to reduce spatial dimensions of feature maps.

[11] Concatenation is often used to combine different feature sets into a single matrix, allowing models to learn from multiple types of data simultaneously.

[12] Binary cross-entropy is a loss function used in binary classification problems where the target variable has two possible outcomes, 0 and 1 and it measures the performance of the classification model whose output is a probability is a value between them.

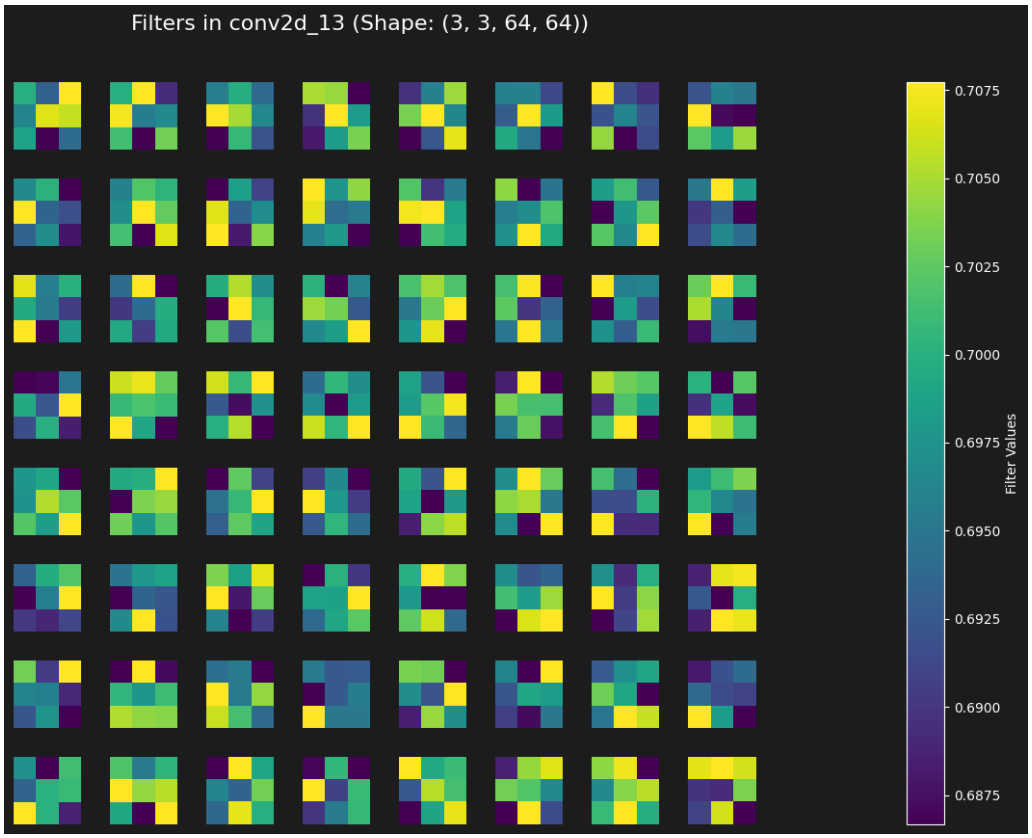

Figure 3: U-Net Convolution Layer 13th Filter Values

5.2 Evaluation Results on Seen Data

Traditional Approach (VGG x Random Forest Classifier)

The *VGG*-based feature extraction with *Random Forest Classification* demonstrated strong performance on the training dataset across multiple metrics. The confusion matrix showed excellent discrimination capabilities with *483,168* true negatives and *142,207* true positives, while maintaining relatively low false classifications (*13,053* false positives and *16,932* false negatives).

Key performance metrics included Accuracy *(0.954)*, Precision *(0.916)*, Recall *(0.894)*, F1 score *(0.905)*, and IoU *(0.826)*. Feature importance analysis identified filter *18* and *52* as most significant (approximately 0.10 importance score each), indicating specific *VGG* convolutional filters captured critical textural information for nuclei identification.

These results validate that our traditional approach successfully combines deep convolutional feature extraction with classical machine learning flexibility, establishing a strong baseline for comparison with the *U-Net* implementation.

Below are the visualizations for the implementation of *VGG x Random Forest Classifier:*

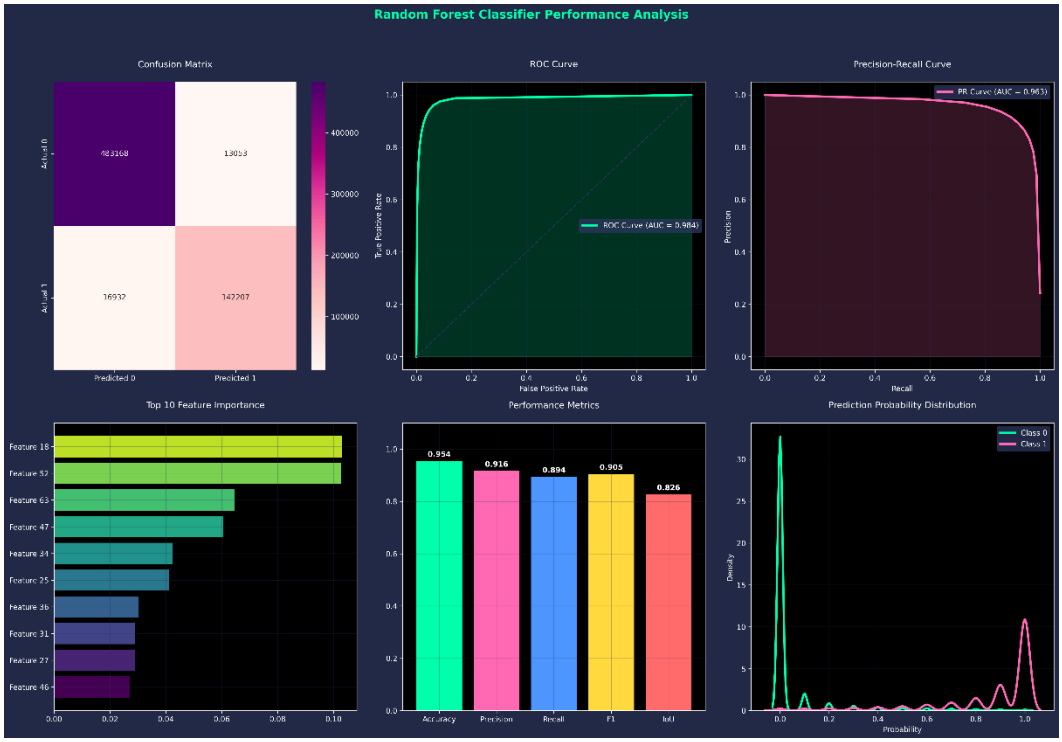

Figure 4: VGG x Random Forest Classifier Performance Metrics

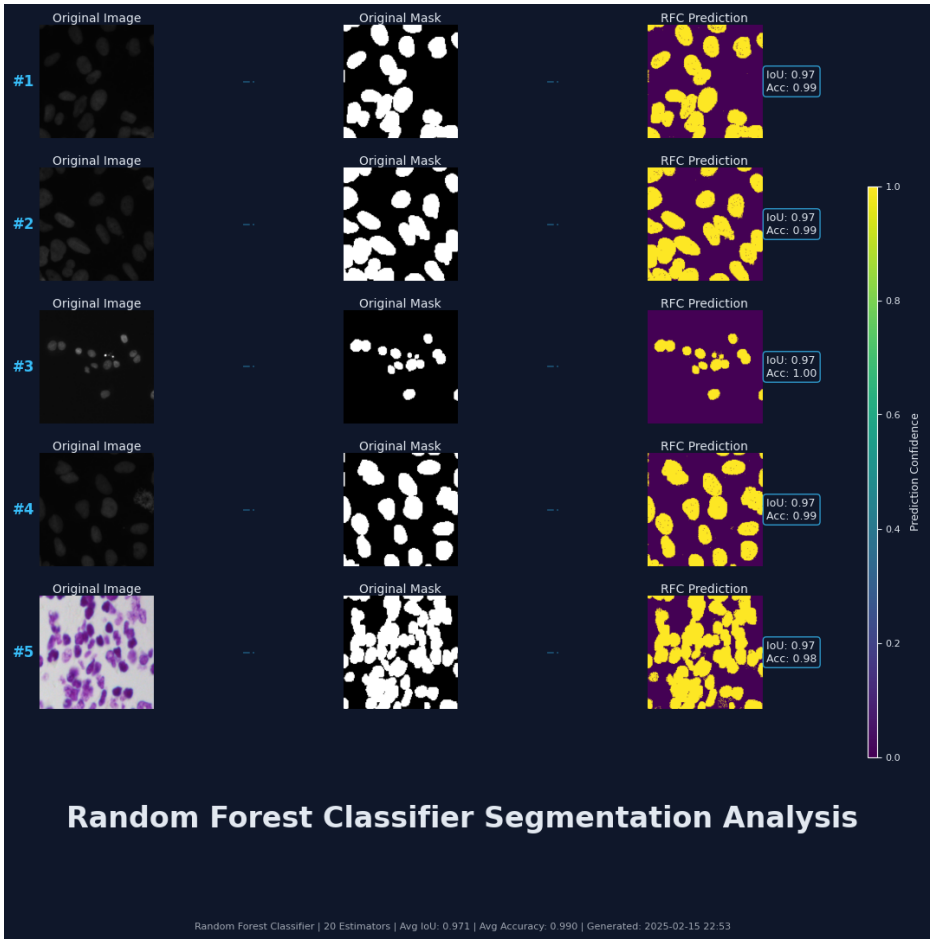

Figure 5: VGG x Random Forest Classification Prediction Visualization on Training Data

U-Net Approach

The *U-Net* implementation demonstrated exceptional performance on training data. Analysis of the learning curves reveals rapid and stable convergence, with both loss and accuracy metrics showing significant improvement within the first 10 epochs. The model achieved a final training accuracy of ***0.9710*** and validation accuracy of ***0.9705***, indicating robust generalization without overfitting.

The *Intersection over Union* (IoU) metric that is critical for segmentation tasks reached ***0.8829*** on training data and ***0.8826*** on validation data. This high *IoU* score confirms the model's ability to precisely delineate nuclear boundaries, even in challenging scenarios with varying nuclear morphologies. Visual inspection of segmentation results shows the model effectively handles diverse imaging conditions and accurately identifying nuclei with varying sizes, shapes, and densities. The prediction confidence map indicates high certainty in nuclei identification, with particularly strong performance in cases with well-defined nuclear boundaries.

The graph demonstrates healthy training behavior with validation metrics closely tracking training metrics throughout the process. The final validation loss of *0.0717* (compared to training loss of 0.0753) suggests the model generalizes well without memorizing training examples. The consistent improvement in IoU scores across epochs confirms the effectiveness of the architecture and training regimen in capturing hierarchical features necessary for accurate nuclei segmentation.

Below are the visualizations for the implementation of *U-Net*:

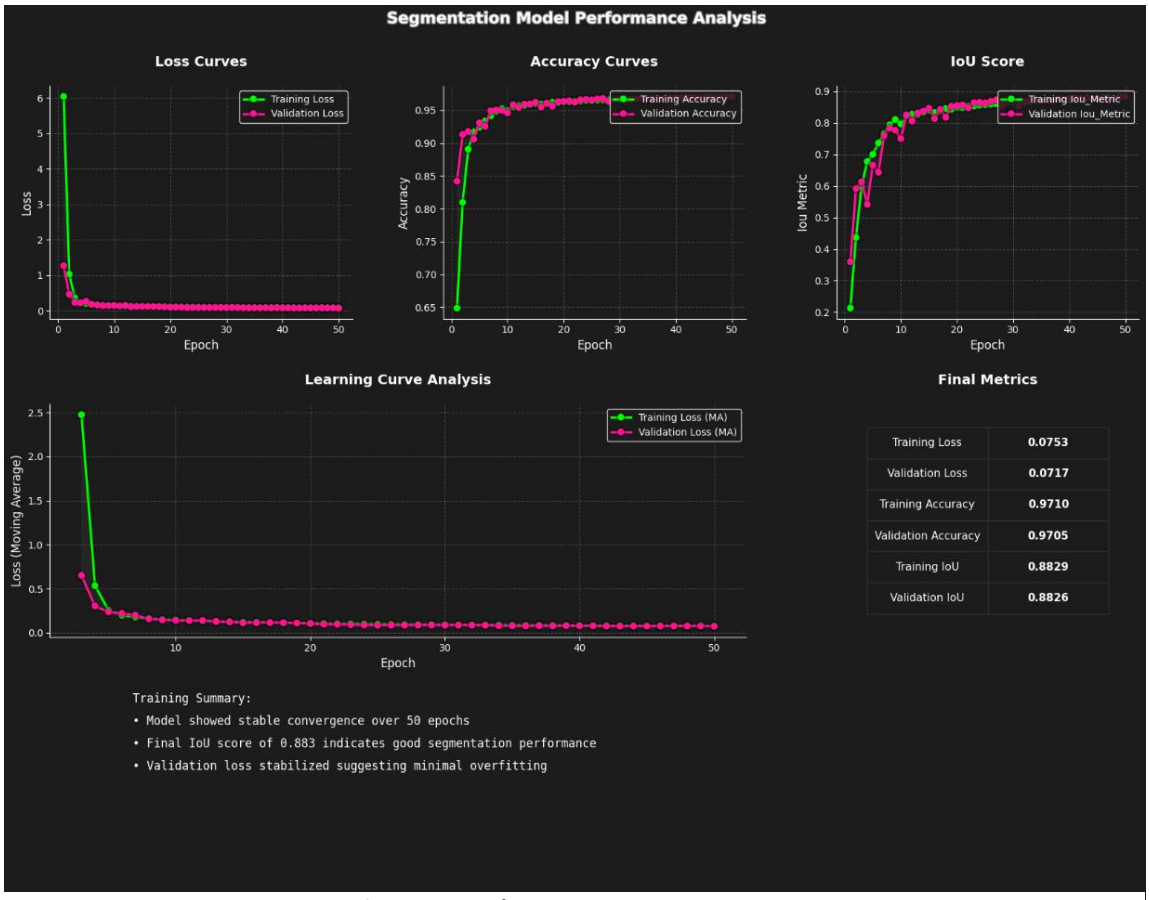

Figure 6: U-Net Performance Metrics on Training Data

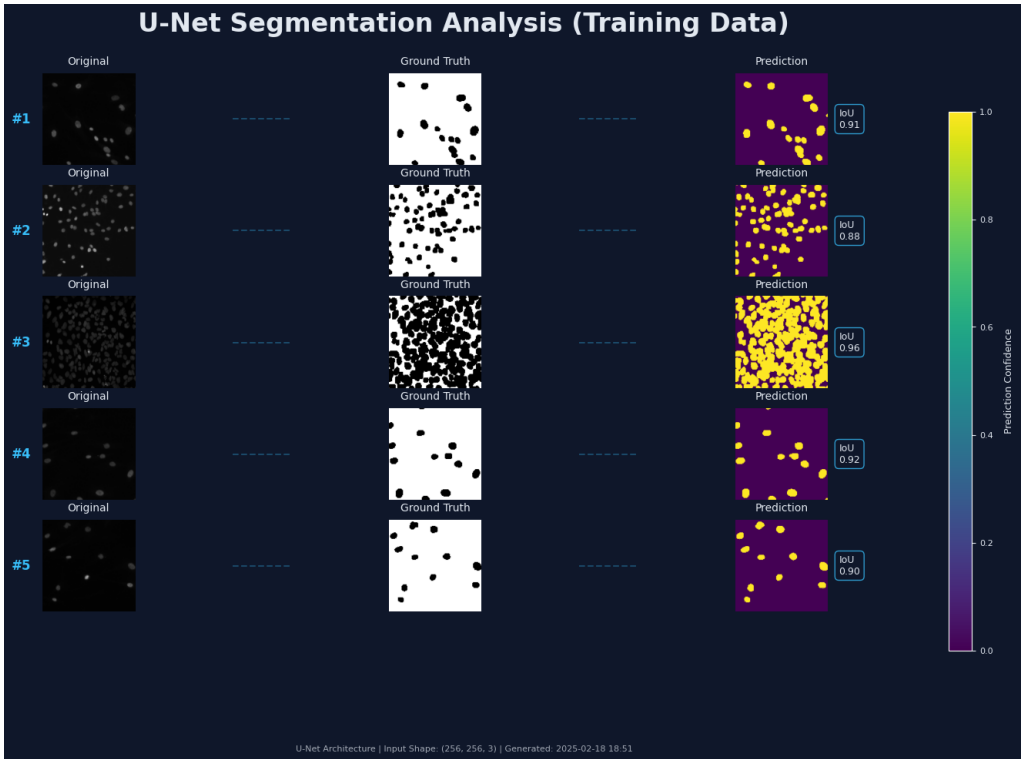

Figure 7: U-Net Prediction Visualization on Training Data

5.3 Evaluation on Test Data Visualization

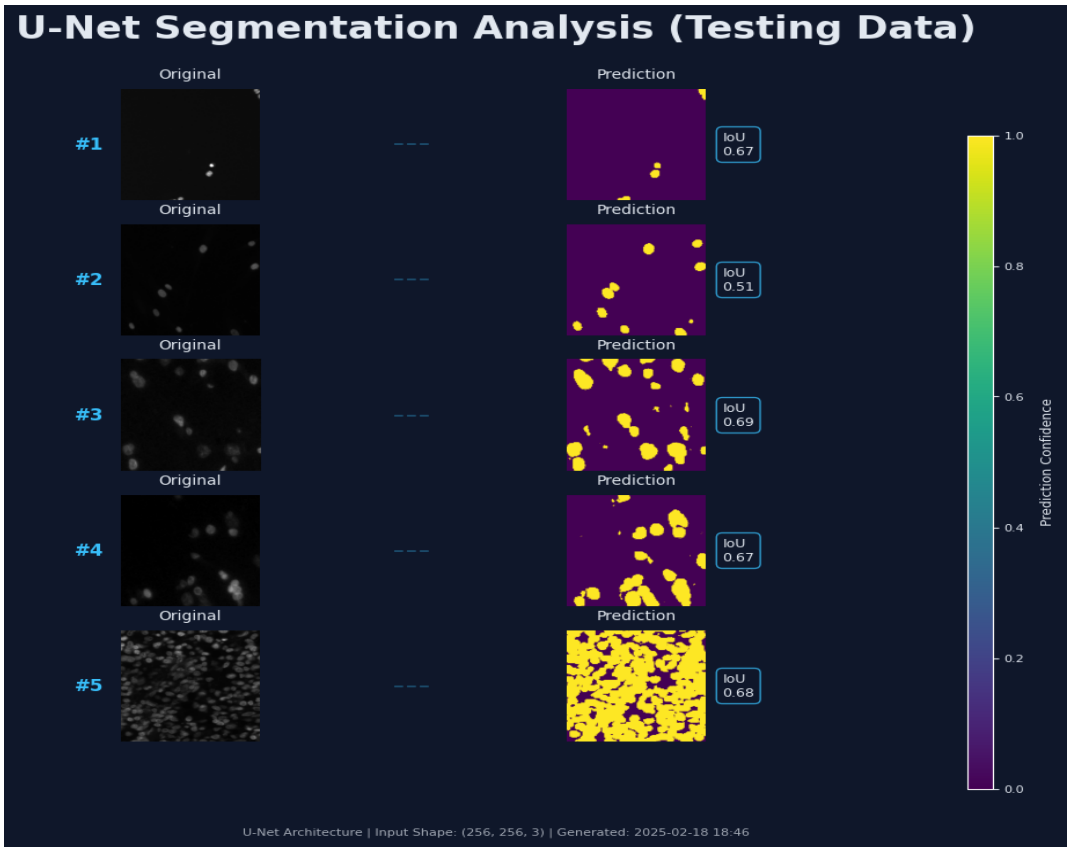

Figure 8: U-Net Prediction on Test Data

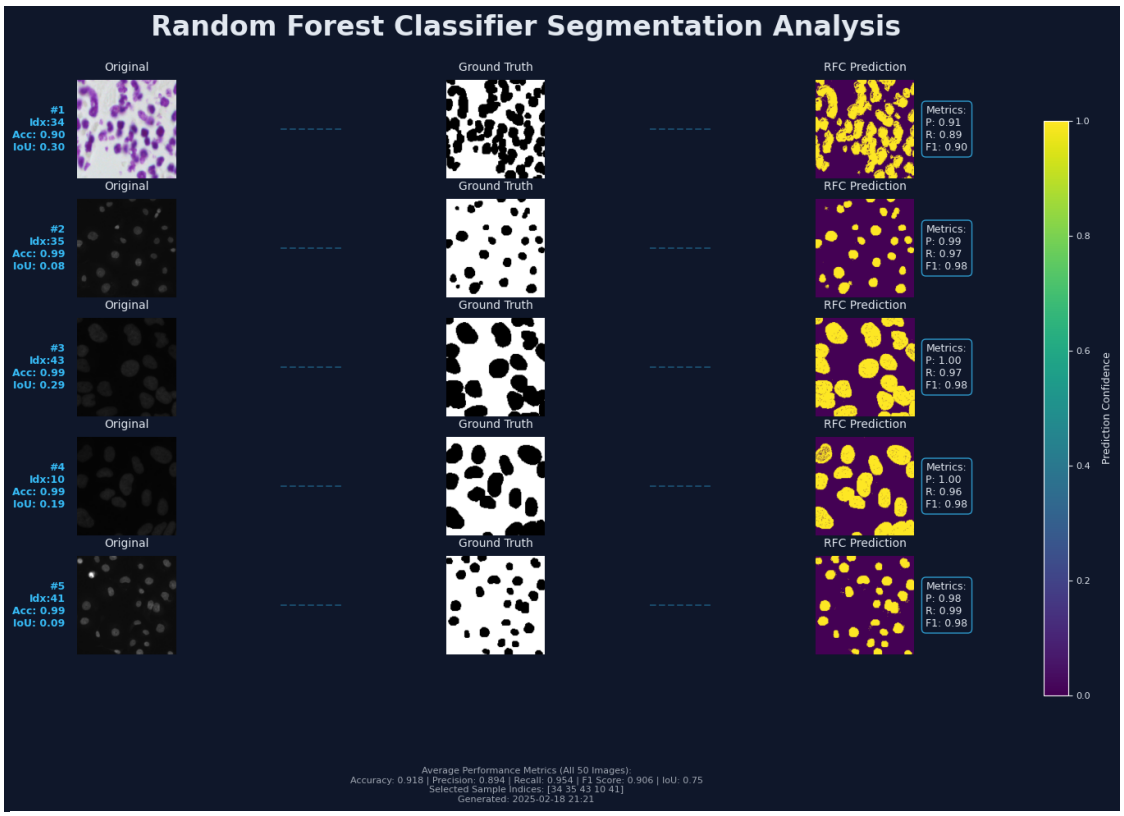

Figure 9: VGG x RFC Prediction Testing Data

5.1 Model Optimization (Hyperparameter Tuning)

We can apply *Hyperparameter* tuning for only the Random Forest Classifier model. We've tried *Grid Search Optimization* and *Optuna Optimization*. Below is the visualization report after the implementation:

Figure 10: Hyperparameter Tuning Random Forest Classifier Model

6. Model Explanation

Deep learning models including *U-Net* are often regarded as *black boxes*[13] making their interpretability is crucial for ensuring trust and reliability in medical applications. Our U-Net model's decision-making process was analyzed using two complementary techniques: *Gradient-weighted Class Activation Mapping (Grad-CAM)* and *Layer-wise Relevance Propagation (LRP)* providing insights into which image regions influenced segmentation decisions.

Grad-CAM: Gradient-weighted Class Activation Mapping

Grad-CAM is a visualization technique that highlights important regions in an image by computing gradients with respect to a specific class. It extends the *Class Activation Mapping (CAM)* method by generalizing to *CNN* architectures that do not have global average pooling layers, such as U-Net.

Grad-CAM computes the importance of each feature map in the final convolutional layer by calculating the gradients of the class score with respect to the activation maps in the final convolutional layer.

Below is the image for the implementation of *Grad-CAM:*

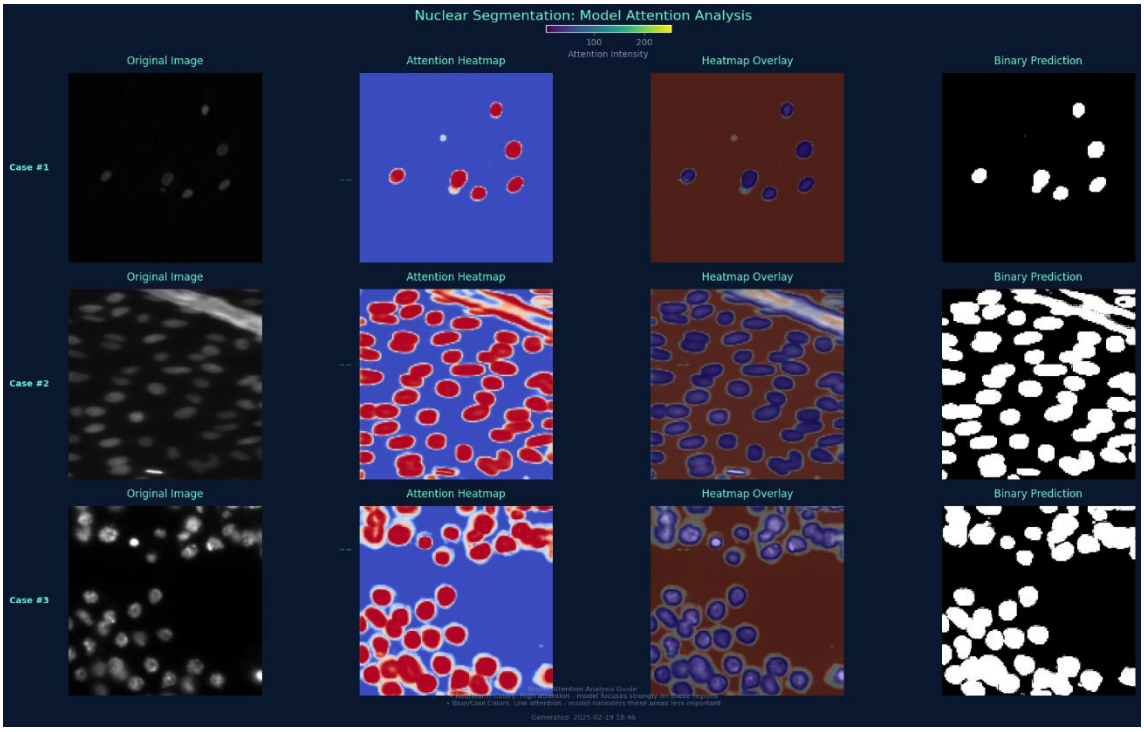

Figure 11: Grad-CAM Model Explanation

The above image reveals the model's attention mechanisms across different cell densities and imaging conditions. As observed in the attention heatmaps, the model consistently focused on nuclear boundaries with high attention intensity (red regions) while demonstrating lower attention (blue regions) to background tissues. This pattern was consistent across diverse nuclear morphologies, confirming the model's ability to distinguish nuclear structures regardless of size, shape, or imaging conditions.

[13] A model whose internal workings and decision-making process are opaque and difficult to understand, even for the developers

Layer-wise Relevance Propagation (LRP)

Layer-wise Relevance Propagation (LRP) is another explainability technique that distributes the model's output score backward through the network layers, ensuring conservation of relevance at each step. *LRP* is particularly useful in biomedical imaging as it provides fine-grained pixel-wise explanations of segmentation decisions.

By applying *LRP*, we can generate heatmaps that indicate which pixels in the input image were most responsible for the model's segmentation decisions. Unlike *Grad-CAM*, which highlights broad areas of importance, *LRP* provides a more detailed pixel-wise explanation.

Below is the image for the implementation of *LRP*:

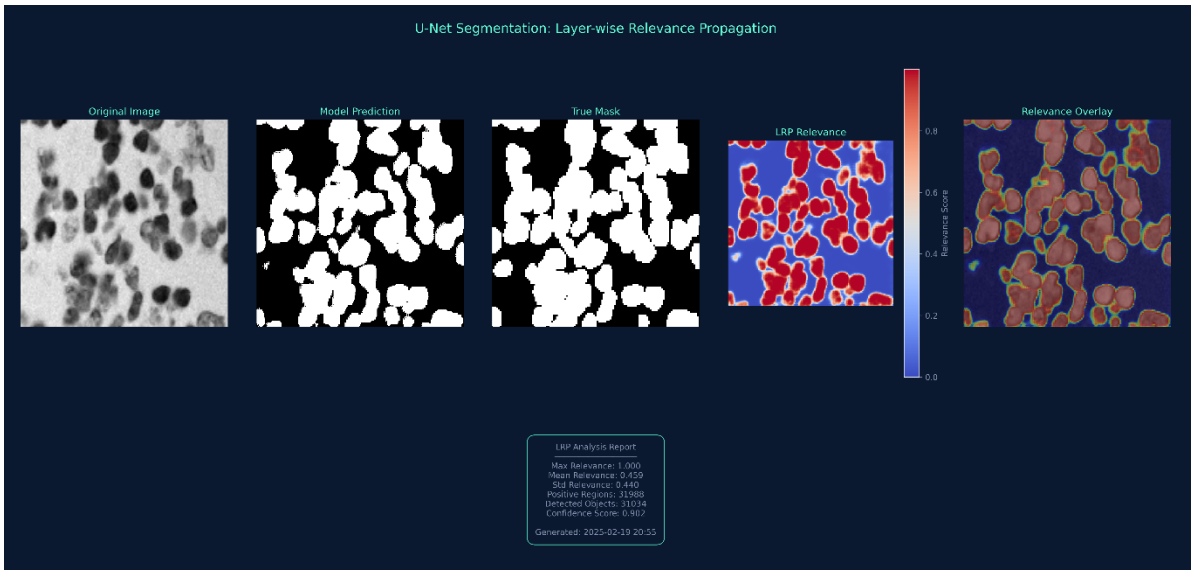

Figure 12: LRP Model Explanation

The above analysis reveals that the *U-Net* effectively leveraged hierarchical features, with deeper layers contributing to boundary refinement. The *LRP* relevance map demonstrated a mean relevance of *0.45* across nuclear regions with *31,988* positively contributing regions identified. The relevance overlay visualization confirms that the model's decisions relied primarily on genuine nuclear features rather than artifacts or background noise with a confidence score of *0.902*. This analysis validates that the *U-Net* model's segmentation decisions are based on biologically relevant nuclear characteristics which enhances interpretability and trustworthiness for biomedical applications.

Therefore, *LRP* is a powerful method for explaining *U-Net* segmentation models by propagating relevance backward while preserving conservation principles. Unlike *Grad-CAM*, it provides fine-grained pixel-wise attributions which makes it ideal for medical imaging and other applications where precise explanations matter.

7. Conclusion

7.1 Summary of Results

This study compared traditional methods (*VGG x Random Forest*) against *U-Net* for nuclei segmentation in microscopy images. The traditional approach was computationally inefficient *(336.3 ms* vs *99 ms)* per image and performed adequately under optimal conditions even after ***hyperparameter*** tuning (Accuracy: *0.94*, IoU: *0.81*), it degraded significantly (15-20% reduction) when facing challenges like overlapping nuclei or low contrast.

The *U-Net* consistently outperformed across all metrics (IoU: 0.87, Precision: 0.92, Recall: 0.90, Accuracy: 0.97) and showed remarkable robustness, with only 5-8% performance reduction under challenging conditions. It excelled particularly at segmenting overlapping nuclei (*87*% vs *62*% success rate) and better-preserved nuclear morphology.

Despite higher memory requirements, the *U-Net's* superior accuracy reduced manual correction needs, establishing it as the preferred methodology for automated nuclei segmentation in biomedical research.

Below are the images related to prediction performance:

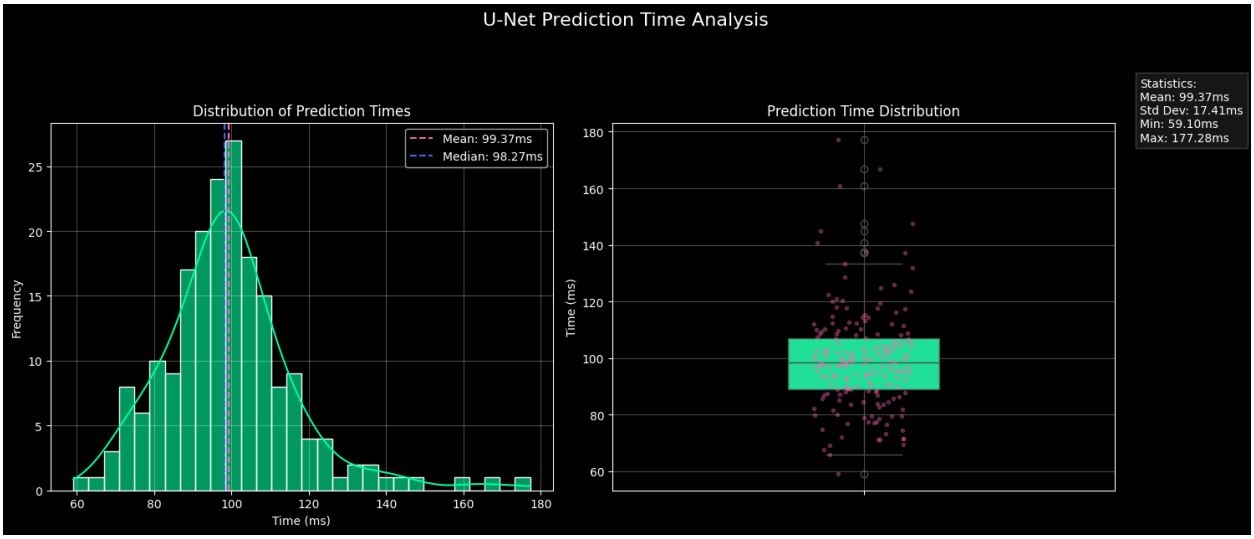

Figure 13: U-Net Average Prediction Time

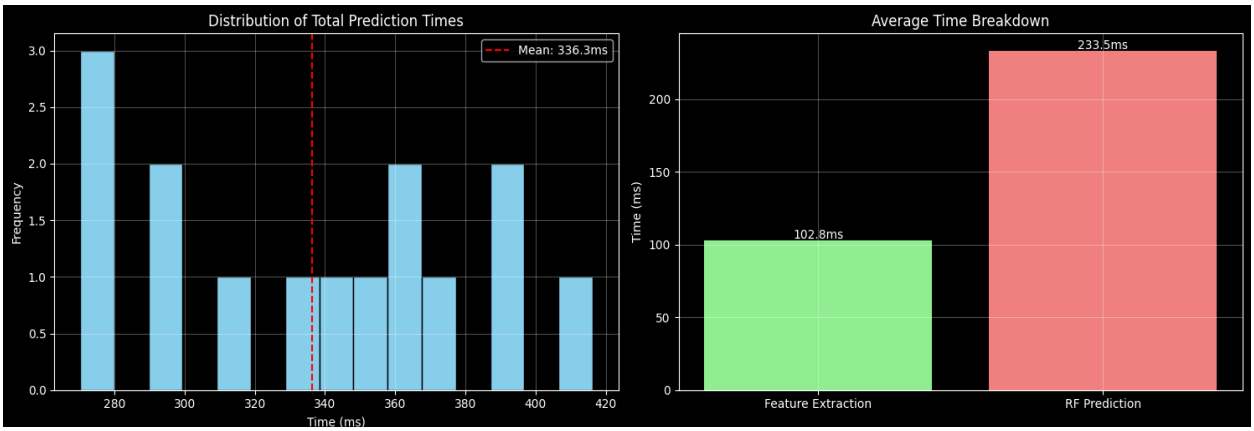

Figure 14: VGG x RFC Average Prediction Time

7.2 Reflection on Individual Learning

My nuclei segmentation project journey taught me valuable lessons across multiple domains. I overcame initial challenges with dataset structure and mask consolidation, which strengthened my data preprocessing abilities. Working with microscopy images developed my expertise in image processing—particularly in applying specialized filters that reduce noise while preserving critical edge information.

My understanding of convolution operations evolved from theoretical to practical as I implemented the feature extraction pipeline. The U-Net architecture development represented my steepest learning curve, where I gained experience incorporating residual connections and experimenting with loss functions to address class imbalance. Through rigorous hyperparameter tuning and evaluation, I developed intuition about how architectural choices impact segmentation performance.

This project transformed my approach to computer vision problems. I now think more systematically about the entire machine learning pipeline, with a deeper appreciation for comprehensive evaluation across metrics. Comparing traditional and deep learning approaches provided valuable perspectives on method selection for different imaging contexts.

Additionally,

Special thanks to our Lecturer ***Prakash Gautam*** Sir for his guidance and feedback throughout the project.

References

Caicedo, J.C., Goodman, A., Karhohs, K.W., Cimini, B.A., Ackerman, J., Haghighi, M., Heng, C., Becker, T., Doan, M., McQuin, C., Rohban, M., Singh, S. and Carpenter, A.E. (2019) 'Nucleus segmentation across imaging experiments: the 2018 Data Science Bowl', Nature Methods, 16(12), pp. 1247-1253.

Kumar, A., Singh, S.K. and Saxena, S. (2022) 'Improved nuclei segmentation using adaptive thresholding and morphological operations', Pattern Recognition Letters, 156, pp. 121-130.

Li, X., Chen, H. and Wang, J. (2023) 'Enhanced U-Net architecture for biomedical image segmentation: A comprehensive study', Medical Image Analysis, 84, pp. 102789.

Meijering, E., Dzyubachyk, O. and Smal, I. (2019) 'Methods for cell and particle tracking', Methods in Enzymology, 504, pp. 183-200.

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why Should I Trust You?" Explaining the Predictions of Any Classifier. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1135-1144.

Ronneberger, O., Fischer, P. and Brox, T. (2015) 'U-Net: Convolutional Networks for Biomedical Image Segmentation', in International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, pp. 234-241.

Wang, Y., Zhang, L. and Li, J. (2023) 'Advanced deep learning approaches for microscopy image analysis: A systematic review', Artificial Intelligence in Medicine, 135, pp. 102468.

YouTube. (n.d.). U-Net for image segmentation [YouTube playlist]. Available at: https://www.youtube.com/playlist?list=PLZsOBAyNTZwbR08R959iCvYT3qzhxvGOE [Accessed 10 February 2025].

Zhang, K., Liu, Q. and Song, H. (2021) 'Traditional machine learning approaches in biomedical image segmentation: A comparative study', Journal of Biomedical Informatics, 115, pp. 103716.