

# Enhanced Nucleus Segmentation with Sobel Edge Detection and Attention Gate with Modified UNET

Sravanthi Jakkula

Department of Computer Science and Engineering, National  
Institute of Technology, India  
sravs521@gmail.com

Raju Bhukya

Department of Computer Science and Engineering, National  
Institute of Technology, India  
raju@nitw.ac.in

**Abstract:** Automated segmentation of nuclei in Hematoxylin and Eosin (H&E) stained histopathology images plays a vital role in accurate cancer diagnosis and prognosis. These techniques enable detailed analysis of numerous nuclei in H&E images, providing both qualitative and quantitative insights. However, challenges arise when segmenting nuclei of varying sizes and with indistinct boundaries, which can undermine the reliability of segmentation outcomes. To address these issues, we propose a novel approach that integrates edge information, extracted from the input data, into the UNET architecture a well-established model for image segmentation. Our approach involves modifying the Attention Gate (AG) mechanism within the UNET to emphasize edge features during segmentation. This modification improves the precision of nucleus boundary delineation, particularly in cases with vague or overlapping boundaries, reducing segmentation errors and boosting overall accuracy.

**Keywords:** UNET, nucleus segmentation, sobel filter.

Received October 1, 2024; accepted May 21, 2025

<https://doi.org/10.34028/iajit/22/4/7>

## 1. Introduction

Histopathology images are integral to modern medical diagnostics, providing detailed insights into the cellular and structural composition of tissues. These images are generated through the microscopic examination of thin tissue sections stained with specialized dyes, enabling pathologists to observe cellular morphology, detect abnormalities, and diagnose diseases. In this context “nucleus segmentation” refers to the process of identifying and delineating nuclei within tissue samples. Nuclei are key indicators of cellular activity, offering essential information on cell proliferation, differentiation, and pathology. Accurate segmentation of nuclei is crucial for quantifying cellular characteristics, measuring biomarkers, and analyzing tissue morphology all of which are critical components of histopathological analysis.

However, nucleus segmentation in histopathology images is particularly challenging due to the complex and heterogeneous nature of tissues. Variations in staining intensity, tissue thickness, cellular density, and the presence of artifacts can hinder accurate segmentation. Manual segmentation methods, while traditionally used, are labor-intensive, time-consuming, and prone to subjectivity, making them unsuitable for large-scale analysis or time-sensitive clinical environments.

To overcome these challenges, automated image analysis techniques, especially the UNET architecture, have become widely adopted. UNET is designed specifically for biomedical image segmentation, featuring a contracting path for extracting context and

an expansive path for accurate localization. Its ability to retain fine details while capturing high-level features makes it particularly well-suited for nucleus segmentation in histopathology images. By incorporating skip connections and utilizing end-to-end learning, UNET provides a scalable and efficient solution for automating this critical task.

In this research, we explore the use of UNET for nucleus segmentation in histopathology images, focusing on the impact of preprocessing techniques to enhance segmentation performance. Through extensive experimentation and evaluation on benchmark datasets, including the Kasturba Medical College (KMC) liver and MoNuSeg datasets, we aim to contribute to the development of automated nucleus detection methods in histopathological analysis.

## 2. Literature Review

This study includes citations to several helpful studies that helped us arrive at our findings. The list below includes a few of them.

Lal *et al.* [8] presented a deep learning architecture called NucleiSegNet for segmenting nuclei in Hematoxylin and Eosin (H&E) stained liver cancer histopathology images. The model to tackle problems such as segmenting close nuclei and fluctuating nucleus shapes, the model integrates robust residual blocks and attention methods. It demonstrates its efficacy in nuclear segmentation tasks by outperforming current models in both quantitative and qualitative ways. Along with suggesting possible future uses and expansions of their work in histopathology image processing, Lal *et al.* [8]

also offer a fresh collection of annotated liver nuclei pictures.

Oktay *et al.* [12], introduce an Attention Gate (AG) model for medical imaging. This model is trained to focus on input structures of variable shapes and sizes, improving sensitivity and accuracy. By training with AGs, irrelevant regions in input images are suppressed while salient features are highlighted, eliminating the need for external localization modules. AG are also integrated with the existing UNET architectures. This approach demonstrates notable enhancements in tissue/organ identification and localization, particularly benefiting small organs like the pancreas.

Zhang *et al.* [16] used a deep residual UNET model to offer a novel approach for road extraction from aerial photos. Road segmentation in satellite photos is made more efficient with this method, which blends residual learning and UNET architecture. The efficiency of the method in remote sensing applications is demonstrated by its superior precision and recall compared to current state-of-the-art techniques.

Ronneberger *et al.* [13] demonstrated how well UNET performs in a variety of segmentation tasks, such as cell segmentation in light microscopy images. UNET performs faster and more accurately than earlier techniques, particularly when applied to difficult 2D transmitted light datasets. For more study and use in biomedical image segmentation tasks, Ronneberger *et al.* [13] additionally offer trained networks, implementation information, and supplemental materials.

Szegedy *et al.* [14] introduced the GoogLeNet architecture and the Inception module, which helped make deep neural networks more efficient and accurate. Instead of just making networks deeper or wider, it used different sized filters in parallel to capture more features with fewer computations. This innovation was key in advancing deep learning in image recognition tasks.

Long *et al.* [9] introduced Fully Convolutional Networks (FCNs), a new way to do semantic segmentation, which means labeling every pixel in an image. Instead of using fully connected layers, it uses only convolutional layers so the network can take any size image and output a segmented version. It was one of the first deep learning models to do pixel-level prediction directly and accurately.

He *et al.* [5] suggested Residual Networks (ResNet), a deep learning model that solves the problem of training very deep neural networks. It uses “skip connections” to let the network learn the difference (residual) between layers instead of learning everything from scratch. This made training easier and allowed the creation of very deep networks with much better accuracy.

Yu *et al.* [15] present a Discriminative Feature Network (DFN) that improves how well a model can tell different objects apart in semantic segmentation. It focuses on making features more distinct

(discriminative) and enhances both the detail and the overall structure in segmented images. The method helps the model better separate objects, especially when they have similar appearances.

Milletari *et al.* [10] introduced V-Net, a deep learning model designed to segment 3D medical images like MRI scans. Unlike 2D models, V-Net works directly on volumetric 3D data and uses convolutional layers to process the whole volume at once. It also uses a special Dice loss function to better handle class imbalance, which is common in medical imaging.

Karthek *et al.* [6] introduced a new feature extraction method called Local Triangular Patterns (LTP) for facial expression recognition. It captures detailed facial texture information by focusing on triangular patterns in image regions. This handcrafted method helps improve the accuracy of recognizing different facial expressions, especially in challenging lighting or noise conditions.

Gugulothu and Bhukya [4] present a deep learning method optimized using a hybrid algorithm called Coot-Lion to predict how fast point mutations happen in COVID-19 genomes. By analyzing genetic data, the model aims to detect and predict virus mutations more accurately. The hybrid optimization improves model performance and prediction accuracy.

Dasari and Bhukya [3] proposed a deep learning model to predict new viral genomes with a focus on explainability making the model’s decisions easier to understand. It helps researchers not only detect unknown viruses from genome data but also understand why the model made a certain prediction. This is important for building trust and aiding scientific discovery in virus research.

Bhukya *et al.* [2] present a hybrid deep learning model with attention mechanisms to accurately find transcription factor binding sites important regions in DNA where proteins attach to control gene activity. The attention layer helps the model focus on the most important parts of the DNA sequence, improving both prediction accuracy and interpretability.

The MoNuSeg dataset in [11] is a collection of histopathology images with manually labeled nuclei used for training and testing nucleus segmentation models. It’s widely used in medical image analysis challenges to help improve algorithms that detect and separate cell nuclei in tissue samples.

Karthek *et al.* [7] proposes a new method for recognizing facial expressions by analyzing texture patterns in images using symbolic features. The technique captures subtle facial changes more effectively, helping machines understand emotions from faces with improved accuracy.

Bhukya [1] Suggested a new way to guess gene activity by first shrinking large sets of gene data using deep autoencoders. Then, a type of neural network Multi-Layer Perceptron (MLP) is used to make predictions from that smaller set. This method works

better because it finds hidden patterns and makes the data easier to handle.

### 3. Proposed Work and Methodology

Although the UNET architecture has achieved considerable success in image segmentation, particularly in non-medical fields, its application in medical imaging remains less explored, presenting opportunities for further refinement. One key limitation we identified is its inadequate performance in accurately delineating nucleus boundaries, leading to errors, especially at the edges. To overcome this issue, our approach involves preprocessing input images to extract edge maps, which are then integrated into the AG of the UNET architecture. Figure 1 illustrates the complete methodology of the proposed nucleus segmentation approach, showing the step-by-step process from input histopathological images to segmented outputs. It integrates key techniques such as Sobel edge detection for boundary enhancement, AG mechanisms for feature refinement, and a modified UNET architecture for accurate segmentation. The figure provides a structured visualization of how these components work together to improve segmentation performance.

#### 3.1. Dataset

We trained our model using two datasets: the KMC liver dataset and the MoNuSeg dataset. The KMC liver dataset comprises 194 training images and 4 test images, each with a resolution of  $512 \times 512$  pixels. The MoNuSeg dataset includes 30 training images with annotations for 22,000 nucleus boundary segmentations, and 7 test images containing 7,000 nucleus boundaries.

#### 3.2. Preprocessing

Histopathology images often exhibit variations in color and intensity due to differences in staining protocols, tissue characteristics, and imaging conditions associated with H&E staining. These inconsistencies create challenges for nucleus segmentation, as it becomes difficult to differentiate between true biological structures and staining artifacts. To address this, normalization is applied to the dataset as a preliminary step. Additionally, the manual annotation of nuclei in tissue samples usually results in a relatively small dataset, which can hinder model generalization and increase the risk of overfitting. To overcome these limitations, dataset augmentation techniques such as random rotation, cropping, blurring, and horizontal/vertical flipping are employed to expand the dataset and improve model performance. Figure 2 illustrates an intermediate stage in the process depicted in Figure 1. The H&E stained tissue image (Figure 2-a)) represents the original input, whereas the Sobel edge map (Figure 2-b)) is a processed feature extracted using

edge detection techniques.

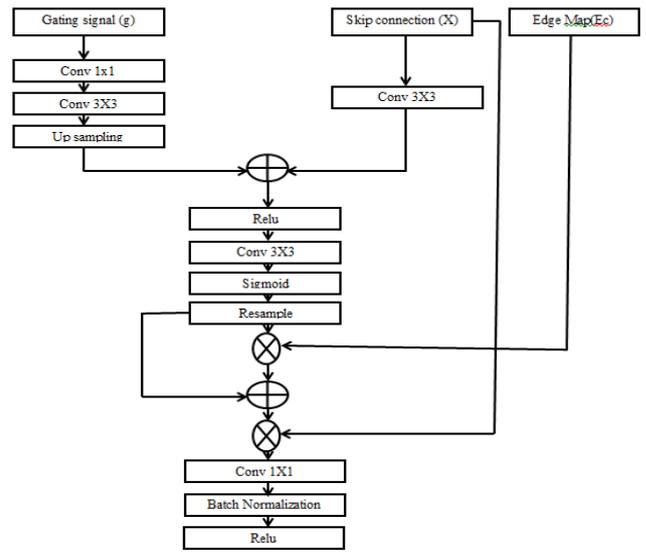
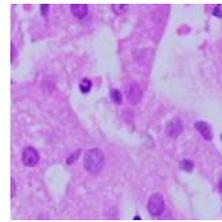
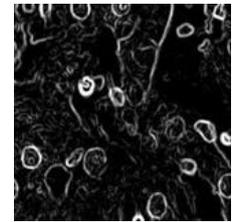


Figure 1. Attention gate combining edge map features.



a) H & E stained tissue image.



b) Edge map obtained after applying Sobel edge detection.

Figure 2. Input tissue image and corresponding edge map.

#### 3.3. Edge Detection

Edge maps are extracted from the input image using the Sobel filter, which is defined by two convolution kernels:  $G_x$  and  $G_y$ , representing the horizontal and vertical gradients, respectively. These kernels are applied to the image through convolution to calculate the gradient magnitude and direction at each pixel.

The Sobel kernels are defined as:

$$G_x = \begin{bmatrix} 1 & 0 & 1 \\ -2 & 0 & 2 \\ 1 & 0 & 1 \end{bmatrix}, G_y = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

The convolution between the Sobel kernels and the image is performed as:

$$I_x = I * G_x, I_y = I * G_y \quad (1)$$

where  $I_x$  and  $I_y$  are the convolved images representing the horizontal and vertical gradients, respectively. These gradient images are then combined to compute the gradient magnitude  $M$  at each pixel location using the following equation:

$$M = \sqrt{I_x^2 + I_y^2} \quad (2)$$

Next, we apply a threshold to the gradient magnitude image to convert it into a binary image. Pixels with gradient magnitudes exceeding a specified threshold are categorized as edge pixels, while those below the

threshold are classified as non- edge pixels. Through experimentation, we determined that a threshold value of 0.2 yielded optimal outcomes.

### 3.4. UNET Architecture

- 1) Encoder: the encoder in the UNET model, adapted from the NucleiSegNet paper, consists of four residual blocks and three down sampling layers. Each residual block is composed of a  $3 \times 3$  convolution layer, followed by a separable convolution layer, and another  $3 \times 3$  convolution layer. All convolution operations are followed by batch normalization and ReLU activation. The residual blocks are repeated four times, with down sampling layers inserted between them, having the output image size at each stage. After each residual block, the number of channels is doubled to retain critical information. This encoder effectively captures intricate image details necessary for accurate segmentation in the decoder.
- 2) Bottleneck layer: the bottleneck layer compresses the input data, retaining only the essential information needed for accurate reconstruction. In NucleiSegNet, this layer comprises three convolutional layers with  $3 \times 3$  kernels, each followed by batch normalization and ReLU activation. This setup enables the encoding of global information about significant regions in H&E stained images, allowing the attention block to filter out irrelevant background areas effectively.
- 3) Decoder: the decoder reconstructs the segmentation map using features extracted by the encoder. The compressed features from the bottleneck layer are passed through four convolution blocks, each containing two convolutional layers followed by batch normalization. After each convolution block, an attention block is applied. The output image size is doubled, and the number of channels is halved after each convolution, allowing for a gradual reconstruction of the segmentation map.
- 4) Novel AG: in the attention block, the high-level features from the bottleneck block and the low-level features from the encoder block are combined to train the model's attention to task-relevant regions. However, we found that the majority of segmentation problems happen at the mask's boundaries. In order to enhance this, we implemented a new strategy.

To gradually extract informative features, edge maps are first generated from the input data and processed via four residual blocks. The goal of this step is to train the model to identify pertinent edges for better attention management. To create  $E_c$ , the resulting edge maps are scaled to fall between 0 and 1. Each residual block's output s then routed to an AG. Here, the skip connection  $x$  from the matching encoder is inserted after the gating signal  $g$  from the next lower layer containing high level features has undergone.

$$X_c = \sigma(U(H_{1 \times 1}(H_{3 \times 3}(g)))) + H_{3 \times 3}(x) \quad (4)$$

$$X_c = X_c E_c + X_c$$

Here  $H_{y \times y}$  means standard convolutional operation with kernel size  $y$ ,  $U$  means un sampling and  $\sigma$  means sigmoid function. This operation assigns higher importance to pixels with greater values in  $E_c$ , signifying nuclei boundaries. Finally, the gating signal  $g$  is scaled with  $X_c$  which diminishes attention to non-essential features, aiding in precise segmentation of nuclei boundaries while minimizing focus on irrelevant areas.

## 4. Experiment Results

### 4.1. Training

During the model training process, we utilized two datasets: the KMC liver dataset and the MoNuSeg dataset. Both datasets were split into distinct sets for training, validation, and testing. The KMC dataset as divided in a 1:6:24 ratio for testing, validation and training respectively. While the MoNuSeg dataset followed a 1:1:5 ratio. The input images were H&E stained and corresponding outputs included ground truth annotations.

Training was conducted using the Adam  $W$  optimizer. After experimenting with different learning rates, we found that 0.001 worked well as the starting value. To reduce the model's complexity, we divided input images into patches of size  $256 \times 256$ , which helped lower the number of parameters. For the test dataset, we predicted the segmentation mask for each patch, reassembled the patches and then evaluated the final scores.

### 4.2. Evaluation Metrics

For evaluating the performance of the model, we use the following metrics. Here  $t_n$ ,  $f_n$ ,  $f_p$ , and  $t_p$ , represent true-negative, false-negative, false-positive, and true-positive with respect to the pixels of the predicted segmentation mask.

- 1) Dice score: the Dice score assesses the overlap between actual segmentations and predicted segmentations in image segmentation tasks. It focuses on true positives ( $t_p$ ), measuring the agreement between positive predictions and ground truth. True negatives ( $t_n$ ) aren't directly considered in its calculation.

$$Dicescore = \frac{2t_p}{2t_p + f_p + f_n} \quad (5)$$

- 2) IOU score:  $IOU$  measures the similarity between actual segmentation and predicted segmentations by comparing their intersection to their union. It considers true positives ( $t_p$ ) but doesn't explicitly account for true negatives ( $t_n$ ).

$$IOU \text{ Score} = \frac{t_p}{t_p + f_p + f_n} \quad (6)$$

- 3) Accuracy: *Accuracy* evaluates overall classification correctness by considering both true positives ( $t_p$ ) and true negatives ( $t_n$ ). It measures the correctness of the predicted samples of the total number of samples.

$$Accuracy = \frac{t_p + t_n}{t_p + t_n + f_p + f_n} \quad (7)$$

- 4) Precision: *Precision* assesses the correctness of positive predictions by focusing on minimizing false positives. It is the division of true positive to addition of true positive and false positives. It emphasizes the model's ability to avoid incorrectly labeling negative samples as positive, thus reducing false positives.

$$Precision = \frac{t_p}{t_p + f_p} \quad (8)$$

### 4.3. Results

In Figure 3, we compare the predicted masks of the models. Figure 3-a) shows the original H&E tissue image. Figure 3-b) shows the true mask or ground truth. Figure 3-c) shows the mask predicted by the existing model, and Figure 3-d) shows the mask predicted by the proposed model.

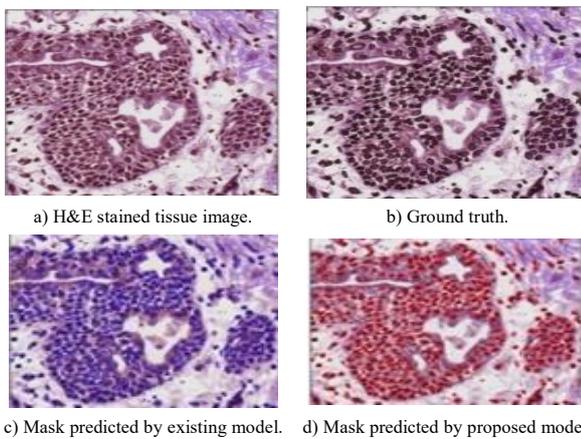


Figure 3. Comparison of ground truth and predicted segmentation masks.

Table 1. Performance metrics for deep learning models on KMC liver dataset.

Model	Dice	IOU	Accuracy	Precision
Existing	85.26	74.38	84.00	98.27
Proposed	87.69	78.50	88.59	98.84

Table 1 presents a comparison of the performance metrics of the existing and proposed models on the KMC liver dataset, focusing on Dice score, IOU score, accuracy, and precision. The proposed model consistently outperforms the existing model across all metrics, indicating improved segmentation quality. The Dice score, which measures the overlap between predicted and actual segmentations, is higher for the proposed model (86.69) compared to the existing model (85.26). Similarly, the IOU score, which measures

similarity by comparing the intersection to the union, is also higher for the proposed model (76.60) compared to the existing model (74.38). The proposed model also demonstrates better overall accuracy, achieving a score of 86.58 versus 84.00 for the existing model. Additionally, the precision is slightly improved with the proposed model (98.46) compared to the existing model (98.27). These improvements highlight the effectiveness of the proposed model in achieving better nuclei segmentation, leading to enhanced accuracy and reliability in the segmentation process for the KMC liver dataset.

Table 2. Performance metrics for deep learning models on MoNuSeg dataset.

Model	Dice	IOU	Accuracy	Precision
Existing	81.65	69.05	82.60	93.24
Proposed	82.56	70.76	86.73	94.85

Table 2 compares the performance of the existing and proposed models on the MoNuSeg dataset using the metrics Dice score, IOU score, accuracy, and precision. The proposed model achieves slightly better results compared to the existing model across all metrics, indicating improved segmentation performance. Specifically, the proposed model has a Dice score of 82.13, which is higher than the existing model's score of 81.65. The IOU score for the proposed model is 69.76, compared to 69.05 for the existing model. The proposed model also demonstrates better accuracy, achieving a score of 85.83, while the existing model scores 82.60. In terms of precision, the proposed model shows a slight improvement, with a score of 93.85 compared to 93.24 for the existing model. These results highlight the effectiveness of the proposed model in enhancing segmentation accuracy and reliability on the MoNuSeg dataset.

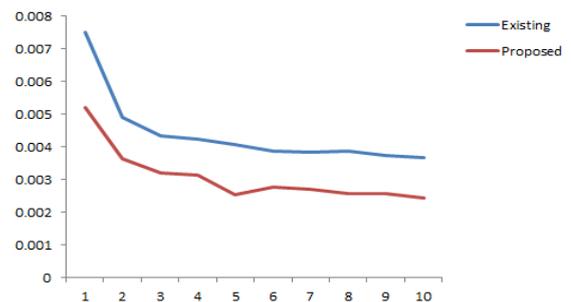


Figure 4. Training loss progression over epochs for KMC liver dataset.

Table 3. Training loss progression for KMC liver dataset.

Iteration	Existing	Proposed
5	0.007516	0.005220
10	0.004923	0.003633
15	0.004336	0.003221
20	0.004256	0.003145
25	0.004060	0.002529
30	0.003881	0.002761
35	0.003847	0.002722
40	0.003858	0.002573
45	0.003732	0.002578
50	0.003665	0.002430

From Figure 4, we compare the training loss of the existing model with that of our proposed model on the KMC liver dataset. As observed in Table 3 the training loss progression for both the existing and proposed models on the KMC liver dataset over 50 iterations as shown in above figure. The proposed model consistently achieves lower training loss than the existing model, starting with a value of 0.005220 at iteration 5 and converging to a final value of 0.002430 at iteration 50. In comparison, the existing model starts with a loss of 0.007516 and converges to 0.003665 at iteration 50. This indicates that the proposed model learns more effectively, with better convergence behavior and lower loss throughout the training process.

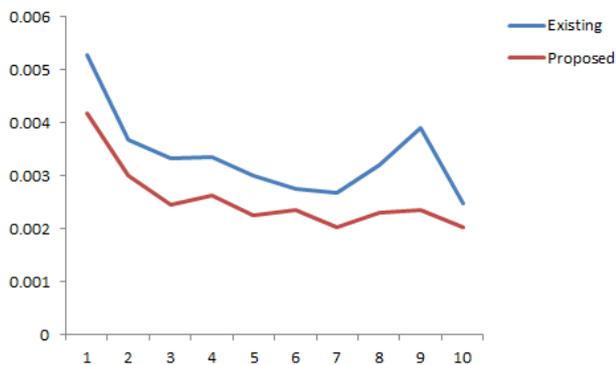


Figure 5. Validation loss progression over epochs for KMC liver dataset.

Table 4. Validation loss progression for KMC liver dataset.

Iteration	Existing	Proposed
5	0.005292	0.004184
10	0.003673	0.003016
15	0.003332	0.002456
20	0.003347	0.002623
25	0.003009	0.002266
30	0.002744	0.002344
35	0.002670	0.002034
40	0.003203	0.002298
45	0.003900	0.002345
50	0.002480	0.002033

From the above Figure 5, we compare the validation loss of the existing model with that of our proposed model on the KMC liver dataset. As observed in Table 4 the validation loss progression of both the existing and proposed models on the KMC liver dataset over 50 iterations. The proposed model consistently outperforms the existing model in terms of validation loss, starting with a lower loss value of 0.004184 at iteration 5 compared to 0.005292 for the existing model. As the iterations progress, the proposed model's validation loss continues to decrease, ultimately converging at 0.002033 by iteration 50, whereas the existing model converges at a higher loss of 0.002480. This demonstrates that the proposed model generalizes better to the validation set, achieving a more stable and lower validation loss, indicating superior performance and less over fitting compared to the existing model.

In Figure 6, we compare the F1-score of the existing model with that of our proposed model on the KMC

liver dataset. The F1-score progression for existing and proposed models on the KMC liver dataset over 50 iterations. The proposed model starts with an F1-score of 0.817854 and improves to 0.848766 while the existing model reaches 0.844240 showed in Table 5. This indicates that the proposed model consistently outperforms the existing one in accurately identifying nuclei boundaries.

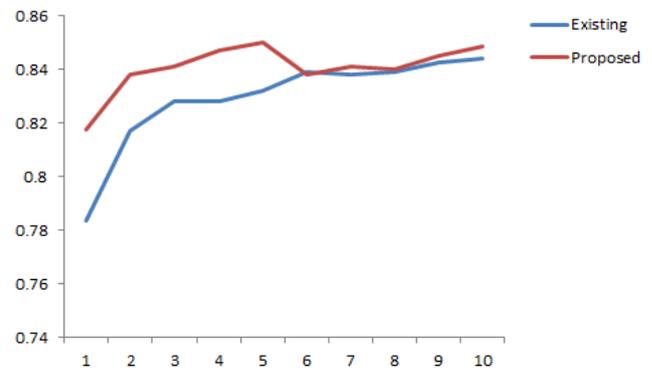


Figure 6. F1-score progression over epochs for KMC liver dataset.

Table 5. F1-score progression for KMC liver dataset.

Iteration	Existing	Proposed
5	0.017645	0.016718
10	0.015675	0.015553
15	0.014921	0.014456
20	0.014598	0.013554
25	0.013995	0.013244
30	0.013990	0.012908
35	0.013715	0.012789
40	0.013459	0.012567
45	0.013600	0.012456
50	0.013317	0.012234

In Figure 7, we compare the training loss of the existing model with that of our proposed model on the MoNuSeg dataset. The proposed model consistently achieves lower training loss, starting at 0.016718 and converging to 0.012234, while the existing model begins at 0.017645 and ends at 0.013317 showed in Table 6. This indicates that the proposed model learns more effectively, resulting in improved training performance. Table 6 presents the training loss progression for the existing and proposed models on the MoNuSeg dataset over 50 iterations.

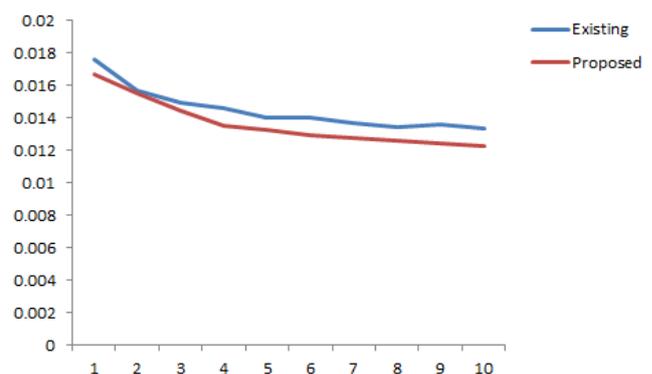


Figure 7. Training loss progression over epochs for MoNuSeg dataset.

Table 6. Training loss progression for MoNuSeg dataset.

Iteration	Existing	Proposed
5	0.783715	0.817854
10	0.817152	0.838346
15	0.828144	0.841027
20	0.827895	0.847224
25	0.832375	0.850276
30	0.838988	0.838118
35	0.838260	0.841200
40	0.838884	0.839956
45	0.842489	0.844985
50	0.844240	0.848766

In Figure 8, we compare the validation loss of the existing model with that of our proposed model on the MoNuSeg dataset. The validation loss of our model is higher in the starting iterations, as shown in Table 7. The proposed model starts with a training loss of 0.027786 and decreases to 0.016243, while the existing model shows a smaller decrease from 0.020152 to 0.016824. This indicates that the proposed model learns more effectively, resulting in a more significant reduction in training loss.

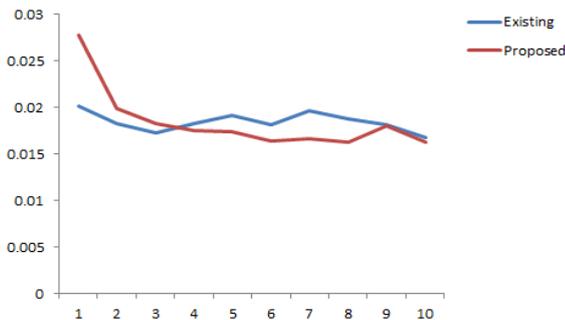


Figure 8. Validation loss progression over epochs for MoNuSeg dataset.

Table 7. Validation loss progression for MoNuSeg dataset.

Iteration	Existing	Proposed
5	0.020152	0.027786
10	0.018317	0.019876
15	0.017239	0.018332
20	0.018278	0.017545
25	0.019173	0.017432
30	0.018103	0.016420
35	0.019677	0.016654
40	0.018749	0.016332
45	0.018186	0.017996
50	0.016824	0.016243

In Figure 9, we compare the F1-score of the existing model with that of our proposed model on the MoNuSeg dataset. Table 8 displays the F1-score progression for the existing and proposed models on the MoNuSeg dataset over 50 iterations. The proposed model starts with an F1-score of 0.801967 at iteration 5 and increases to 0.826980 by iteration 50, whereas the existing model improves from 0.794418 to 0.821524 during the same period. This consistent increase in the proposed model's F1-score highlights its superior performance in accurately identifying nuclei boundaries. Overall, the results indicate that the proposed model demonstrates better learning efficacy and effectiveness in the segmentation task.

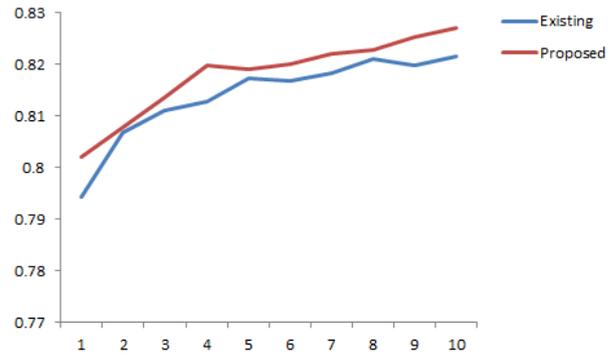


Figure 9. F1-score progression over epochs for MoNuSeg dataset.

Table 8. F1-score progression for MoNuSeg dataset.

Iteration	Existing	Proposed
5	0.794418	0.801967
10	0.806771	0.807878
15	0.811101	0.813567
20	0.812827	0.819890
25	0.817243	0.818954
30	0.816816	0.819976
35	0.818373	0.821987
40	0.821007	0.822765
45	0.819720	0.825432
50	0.821524	0.826980

## 5. Conclusions

In conclusion, our research explores the well-established field of UNET models, aiming to introduce a novel approach to image segmentation that enhances traditional deep learning methods. Our technique involves extracting edge information from input images before feeding them into the UNET model, allowing the attention mechanism to better preserve crucial details. By testing our model on two datasets, we evaluated its performance based on training loss, IOU scores, Dice scores, accuracy, and precision, demonstrating the effectiveness of our method.

## References

- [1] Bhukya R., "Encoding Gene Expression Using Deep Autoencoders for Expression Inference," *The International Arab Journal of Information Technology*, vol. 18, no. 5, pp. 625-633, 2021. <https://doi.org/10.34028/iajit/18/5/1>
- [2] Bhukya R., Kumari A., Dasari C., and Amilpur S., "An Attention-Based Hybrid Deep Neural Networks for Accurate Identification of Transcription Factor Binding Sites," *Neural Computing and Applications*, vol. 34, no. 21, pp. 19051-19060, 2022. <https://doi.org/10.1007/s00521-022-07502-z>
- [3] Dasari C. and Bhukya R., "Explainable Deep Neural Networks for Novel Viral Genome Prediction," *Applied Intelligence*, vol. 52, no. 3, pp. 3002-3017, 2022. <https://link.springer.com/article/10.1007/s10489-021-02572-3>
- [4] Gugulothu P. and Bhukya R., "Coot-Lion

- Optimized Deep Learning Algorithm for COVID-19 Point Mutation Rate Prediction Using Genome Sequences,” *Computer Methods in Biomechanics and Biomedical Engineering*, vol. 27, no. 11, pp. 1410-1429, 2023. <https://doi.org/10.1080/10255842.2023.2244109>
- [5] He K., Zhang X., Ren S., and Sun J., “Deep Residual Learning for Image Recognition,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, pp. 770-778, 2016. DOI:10.1109/CVPR.2016.90
- [6] Kartheek M., Prasad M., and Bhukya R., “Local Triangular Patterns: Novel Handcrafted Feature Descriptors for Facial Expression Recognition,” *International Journal of Biometrics*, vol. 15, no. 2, pp. 194-211, 2023. <https://doi.org/10.1504/ijbm.2023.129224>
- [7] Kartheek M., Prasad M., and Bhukya R., “Texture Based Feature Extraction Using Symbol Patterns for Facial Expression Recognition,” *Cognitive Neurodynamics*, vol. 18, pp. 317-335, 2024. <https://doi.org/10.1007/s11571-022-09824-z>
- [8] Lal S., Das D., Alabhya K., Kanfode A., Kumar A., and Kini J., “NucleiSegNet: Robust Deep Learning Architecture for the Nuclei Segmentation of Liver Cancer Histopathology Images,” *Computers in Biology and Medicine*, vol. 128, pp. 104075, 2021. <https://doi.org/10.1016/j.combiomed.2020.104075>
- [9] Long J., Shelhamer E., and Darrell T., “Fully Convolutional Networks for Semantic Segmentation,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Boston, pp. 3431-3440, 2015. DOI:10.1109/CVPR.2015.7298965
- [10] Milletari F., Navab N., and Ahmadi S., “V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation,” in *Proceedings of the 4<sup>th</sup> International Conference on 3D Vision*, Stanford, pp. 565-571, 2016. DOI:10.1109/3DV.2016.79
- [11] MoNuSeg Dataset, MoNuSeg-Grand Challenge, <https://monuseg.grand-challenge.org/Data/>, Last Visited, 2024.
- [12] Oktay O., Schlemper J., Le Folgoc L., Lee M., et al., “Attention U-Net: Learning Where to Look for the Pancreas,” *arXiv Preprint*, vol. arXiv:1804.03999v3, pp. 1-10, 2018. <https://arxiv.org/abs/1804.03999>
- [13] Ronneberger O., Fischer P., and Brox T., “UNET: Convolutional Networks for Biomedical Image Segmentation,” in *Proceeding of the 18<sup>th</sup> International Conference on Medical Image Computing and Computer-Assisted Intervention*, Munich, pp. 234-241, 2015. [https://link.springer.com/chapter/10.1007/978-3-319-24574-4\\_28](https://link.springer.com/chapter/10.1007/978-3-319-24574-4_28)
- [14] Szegedy C., Liu W., Jia Y., and Sermanet P., et al., “Going Deeper with Convolutions,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Boston, pp. 1-9, 2015. DOI:10.1109/CVPR.2015.7298594
- [15] Yu C., Wang J., Peng C., Gao C., Yu G., and Sang N., “Learning a Discriminative Feature Network for Semantic Segmentation,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Utah, pp. 1857-1866, 2018. DOI:10.1109/CVPR.2018.00199
- [16] Zhang Z., Liu Q., and Wang Y., “Road Extraction by Deep Residual U-Net,” *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 5, pp. 749-753, 2018. DOI:10.1109/LGRS.2018.2802944



**Sravanthi Jakkula** received her B.Tech degree in Computer Science and Engineering From JITS, Karimnagar, affiliated to JNTUH University, Hyderabad, India 2013 and M.Tech degree in Computer Networks and Information Security from JITS, Karimnagar, affiliated to JNTUH University, Hyderabad, India 2015. Currently, Pursuing PhD in Computer Science and Engineering from NIT, Warangal, India. Her research interests are Bioinformatics, Deep learning, Machine Learning.



**Raju Bhukya** has received his B.Tech in Computer Science and Engineering from Nagarjuna University in the year 2003, M.Tech degree in Computer Science and Engineering from Andhra University in the year 2005 and P.hD in Computer Science and Engineering from National Institute of Technology (NIT) Warangal in the year 2014. He is currently working as an Assistant Professor in the Department of Computer Science and Engineering in National Institute of Technology, Warangal, Telangana, India. He is currently working in the areas of Bio-Informatics and Data Mining.