ISSN: 2633-4828 Vol. 5 No.4, December, 2023

International Journal of Applied Engineering & Technology

ATTENTION-ENHANCED RESIDUAL U-NET FOR NUCLEUS SEGMENTATION IN IMMUNOHISTOCHEMISTRY IMAGES

Stephy Benny[1] and Satishkumar L. Varma[2]
[1]Department of Computer Engineering, Pillai College of Engineering New Panvel, Navi Mumbai, India
[2]Department of Information Technology, Pillai College of Engineering, New Panvel, Navi Mumbai, India
[1]stephybenny241@gmail.com and [2]vsat2k@gmail.com
[1]Orcid Id: 0009-0006-8602-4442 and [2]Orcid Id 0000-0002-5203-8655

ABSTRACT
Histopathology image like Immunohistochemistry (IHC) stains helps diagnose aberrant cells in malignant tumours by identifying the antigens (proteins) present in a tissue cell. Manually inspecting the area of interest takes much time, and pathologists make subjective decisions based on their experience. Thus, this research proposes a model named Attention-Enhanced and Residual U-Net (AR-UNet) for cell prognosis to segment the protein-expressed nucleus in an aberrant cell. The suggested approach segments the nucleus semantically using deep learning techniques, which improves feature representation and localization ability by adding additional attention and residual blocks at the appropriate levels of U-Net architecture. These advantages include identifying spatial dependencies, emphasizing the nucleus region, and locating the details of the nucleus. The suggested AR-UNet outperformed other U-net versions, achieving an accuracy of 92% and a loss of 0.25 as compared to visual segmentation outcomes and performance metrics. Additionally, it expedited the laborious and prone-to-mistake process. The results show that the AR U-Net can help pathologists and researchers recognize and segment the nucleus in IHC for histopathological study and cancer diagnosis more precisely.

Keywords: Histopathology images, Immunohistochemistry (IHC), nucleus segmentation, attention, residual blocks, UNet

1. INTRODUCTION
Histopathology involves the analysis and examination of tissues at a microscopic level. Hematoxylin and Eosin (H&E) staining and Immunohistochemistry (IHC) are histopathological images that help classify and diagnose tumours H&E depicts tissue morphology, and IHC identifies specific protein markers. H&E stains act as fundamental stains to visualize cells, offering an overview of tissue structure, and IHC stains specifically target certain protein markers, providing a more focused examination of the presence of proteins in the tissue. Combining the two stains improves the overall understanding of cellular composition and protein location by researchers.

Histopathological images play a significant role in understanding the effects of diseases at the tissue level. Various histological subtypes of cancer can be identified through microscopic examination. Detection of tissue specimens and subsequent treatment can be facilitated through analysis of these microscopic images. Analysis of histological tissue is considered the gold standard method. In this process, a pathologist takes tissue samples from tumour regions, usually focusing on areas where malignancies are suspected. After that, these samples undergo processes such as cutting, staining, and fixing.

IHC stains are frequently used to improve visibility and analysis [1]. The immunohistochemistry (IHC) technique is used to detect and localize particular proteins in cells and tissues. A cell consists of a nucleus, membrane, and cytoplasm. Here in this paper, we try to segment the nucleus that expresses the protein.

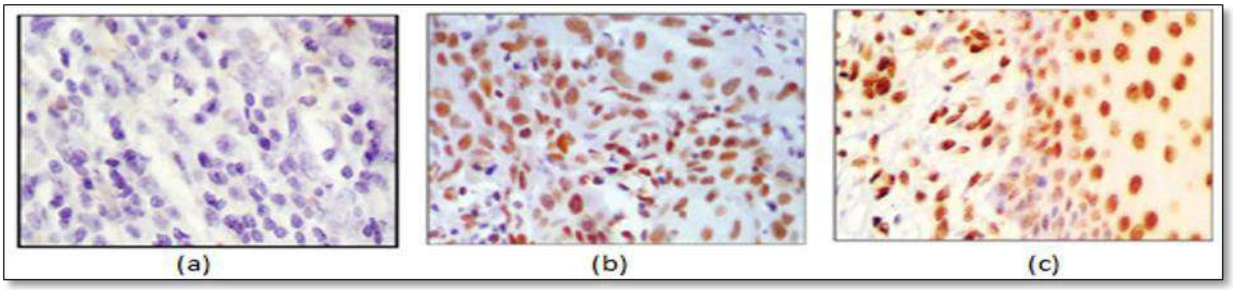


Figure 1: Representative images of IHC staining of a) negative staining, b) positive staining, c) positive staining,

Figure 1 illustrates three images obtained from different tissue samples where the presence of the relative marker is indicated by brown and blue stained regions. The brown stain shows a positive nucleus and the blue stain shows a negative nucleus. The positive and negative nuclei are distinguished by colour; segmenting the nucleus in an IHC stain is quite a challenging task. Many researchers have proposed algorithms for nucleus segmentation in histopathology images, but very little work is done on segmenting nuclei in IHC images due to the challenging nature of separating positive and negative nuclei based on colour.

With this motivation, Attention-Enhanced and Residual U-Net was proposed for nucleus segmentation in the IHC image to segment the brown stains. This extension of the UNet architecture incorporates both a residual unit and an attention mechanism, drawing inspiration from these techniques to improve segmentation accuracy. The architecture of the U-Net model has been modified to create the segmentation model. Instead of using the typical convolutional layers in each stage of the U-Net encoder-decoder network, a dual residual block is employed incorporating an attention block. This specialized segmentation model was applied for precise segmentation of cell nuclei. Attention blocks are added to the encoder and decoder pathways to focus on relevant regions by assigning higher weights to informative features, resulting in improved boundary, localization accuracy, and overall segmentation performance. Double residual blocks are integrated into the U-Net architecture to capture residual information for improved feature extraction, and residual shortcut networks help extract features and reduce image errors. A single channel mask is applied to the IHC image for multiclass segmentation. The multiclass segmentation applied for nucleus segmentation in the IHC image also contributed to the research. Nucleus segmentation in IHC images has contributed to research because of the challenging nature of separating the positive and negative nuclei based on colour.

The paper continues its structure as described below. In Related Works section, we provide a literature survey of U-Net, variants of U-Net, and nucleus segmentation in histopathology image. Section Methodology, we introduce a modified deep learning architecture called Attention-Enhanced and Residual U-Net for Nucleus Segmentation in IHC Image. In the Experimental Setup, the dataset, implementation details, experimental results and analyses, and contribution are presented. In the Conclusions section, we conclude by summarizing the findings and their implications and future research directions in this sector.

2. RELATED WORKS

Over the past few years, there has been a notable surge in research dedicated to the segmentation of nuclei in histopathology images. Researchers have dedicated efforts to creating diverse systems and technologies designed for segmenting nuclei in Hematoxylin and Eosin (H&E) staining as well as Immunohistochemistry (IHC) images. These research endeavors span a broad spectrum of deep learning architectures dedicated to nucleus segmentation. The use of deep learning is more widespread in biomedical images because of its superior performance, automated feature learning, robustness, variability, transfer learning, scalability, and generalization[2][3]. The organized presentation provides a structured overview that visually captures the wide range of research contributions. Each of these investigations presents distinctive perspectives and methodologies related to nucleus segmentation within the realm of deep learning. By categorizing these studies systematically, our objective is not only to offer a comprehensive perspective on the diverse strategies in use but also to create a

valuable reference for researchers and pathologists collaborating to improve the precision of histopathological studies and cancer diagnoses.

2.1 UNET

There are many articles related to segmentation in biomedical images[4]. The most popular deep learning architecture in U-NET[5]. The U-Net architecture is a fully convolutional encoder-decoder design that incorporates skip connections, which is an advancement of the Fully Convolutional Network (FCN). The architecture includes a contracting path (for downsampling) and an expanding path (for upsampling). In the expanding phase, the upsampled feature maps undergo concatenation with their corresponding feature maps from the contracting phase. This concatenation process enables U-Net to retain spatial information and as the architecture goes deeper, it captures progressively more intricate features [6]. However, this also results in numerous redundant low-level feature extractions, given the inadequate representation of the feature in the initial layers.

To improve performance, variants of U-Net architecture were developed [5] like UNet++ [7] for the examination and integration of multimodal learning in supporting clinical decision-making, U2Net [8] for detecting salient objects, attention U-Net [9] for segmenting pancreas, ResNet-a [10] used for remote sensing data, R2UNet [11], Nested U-Net [12], UNet3+ [13], MD-UNet [14], Multi-level dilated residual [15], ATTransUNet [16] for medical image segmentation. Several architectures have been proposed for histopathology images by customizing the UNET framework, as observed in the MobileNetV2 based U-net Model [17], ADS_UNet [18] for segmenting histopathology images, DenseRes-Unet [19] for segmenting overlapped or clustered nuclei from multi organ histopathology images, Hybrid-attention nested UNet [20] , ASPPU-Net [21] , RGSB-UNet [22], MIU-Net[23], efficientunet++[24], RIC-Unet[22] for nucleus segmentation, PPC-UNet[25] for whole slide image segmentation of colorectal histopathology, Alter-AttUNet model[26] for colon cancer histopathological image segmentation, , Attention-Guided deep atrous-residual U-Net [27] for automated gland segmentation in colon histopathology images, Su-net and du-net [28] for tumour segmentation in histopathology images, MultiResUNet [29] for multimodal biomedical image segmentation and DETisSeg: [30] , HookNet [31], MDA-unet [14], GA-UNet [32], YAMU [33] for semantic segmentation of histopathology image.

2.2 Nucleus Segmentation Using Modified UNET

Ahmed et al.[34] proposed a new U-net architecture that combines spatial channels and attention while incorporating ResNet blocks in the encoder layers. This U-net baseline was created to preserve both coarse and fine features while effectively addressing tissue variability issues. It has provided segmentation performance after rigorous testing on three benchmark datasets. However, it is significant to recognize that the model's exceptional performance on benchmark datasets may have limitations when dealing with previously unseen data or datasets from various domains.

In a similar vein, Iqra et al. [19] proposed the DenseRes-U-net model, which uses dense blocks in the final layers of the U-net encoder to highlight relevant features from previous stages. Their nucleus segmentation model was thoroughly validated on four publicly available datasets, yielding promising results. The DenseRes-U-net model may improve segmentation performance by focusing on relevant features from previous layers. Nevertheless, the researchers emphasized the importance of taking into account the increased complexity, which comes with potential trade-offs in terms of computational cost and memory requirements.

He and colleagues [20] proposed a hybrid-attention nested U-net (Han-Net), which is made up of two modules: a hybrid attention-nested U-shaped network and an attention block. It effectively segments the boundaries of nuclei that are both complex and diverse, as well as small and dense. The limitation about light weight is stated in the conclusion, which says that unimportant weights or filters in the network need to be taken out or set to zero in order to reduce the number of parameters, and, by extension, the size of the model and the amount of processing it needs.

Lal et al. [35] proposed a three-block architecture that includes a robust residual block, a bottleneck block, and an attention decoder block. Finding the best hyper parameter settings can be time-consuming and requires a lot of trial and error.

The SU-nets proposed by Kong et al. [36] are made up of four parallel backbone nets linked together by an attention generation model. The use of different loss functions for different networks may introduce inconsistencies or challenges in balancing the contributions of each network during training.

Zeng et al. [22] propose a U-net-based neural network, RIC-U-net (Residual Inception-Channel Attention-U-net), for nuclei segmentation. Although the RIC-U-net is claimed to be cost-effective enough to assist doctors in more accurately diagnosing histology images, the model's generalisation to previously unseen data or datasets from different domains is not addressed. Table 1 shows a detailed study of nucleus segmentation in H&E image.

Table 1: Nucleus segmentation using U-Net

Literature	Architecture	Description and Observations
Zeng *et al.*, 2019	RIC-U-Net [22]	• Residual inception-channel attention-U-Net for nuclei segmentation
Kong *et al.*, 2020	SU-Nets [36]	• Parallel backbone nets with attention generation model
He *et al.*, 2021	Hybrid-Attention Nested U-Net[20]	• It consists of modules for segmenting complex and diverse nuclei boundaries
Lal *et al.*, 2021	NucleiSegNet [35]	• Incorporates robust, bottleneck, and attention decoder blocks
Kiran *et al.*, 2022	DenseRes-U-Net [19]	• Utilizes dense blocks to highlight relevant features
Li *et al.*, 2022	Residual-Attention UNet++ [37]	• Residual blocks with attention mechanism and nested U-Net
Le Dinh *et al.*, 2022	Nested Unet with EfficientNet Encoder [38]	• Nested UNet with efficient net encoder
Ahmad *et al.*, 2023	ResNet-based U-Net [34]	• Combines spatial channels, attention, and ResNet blocks for preserving features
Chowdary *et al.*, 2023	Residual SE-UNet [39]	• Residual blocks with squeeze and excitation blocks

2.3 Nucleus Segmentation in IHC Images

IHC-Net, a fully convolutional neural network proposed by Mahanta et al. [40], was created for automated nuclear segmentation in breast pathology. To precisely segment nuclei in medical images, the architecture uses a stack of six encoders and five decoder blocks with skip connections. For feature extraction and downsampling, encoder blocks use convolution, batch normalization, ReLU activation, and max-pooling layers. Concatenation, convolution, batch normalization, unpooling layers for upsampling, and reconstruction are all techniques used in decoder blocks. This method offers a comprehensive solution to evaluate breast pathology while addressing the problems of nuclear segmentation with an accuracy of 94.82%. A modified U-Net architecture designed specifically for segmenting IHC images was introduced [41] by adding a convolutional layer with a 32x32 pixel receptive field. With this change, the segmentation accuracy is improved, especially for IHC. Even if the architecture is tailored for IHC image processing, the fundamental U-Net structure is still there. This method demonstrates how deep learning architectures can be tailored to meet domain-specific needs. Her2Net [1] is a sophisticated framework for the semantic segmentation and classification of cell membranes and nuclei in the evaluation of breast cancer. By maintaining the cellular and textural characteristics, this architecture introduces a revolutionary trapezoidal LSTM connection topology (TLSTM) structure, improving performance. Multiple layers, including convolution, max-pooling, spatial pyramid pooling, upsampling, and TLSTM layers, are used in

the convolution and deconvolution processes of Her2Net. This innovative layout ensures effective training while optimizing accuracy. The architecture's use of TLSTMs is notable as a cutting-edge strategy for preserving structural integrity during image processing, achieving 98.33% accuracy.

The ColorAE: U-Net methodology [42] is a prime example of group methods in medical image analysis. This approach combines the benefits of U-Net, a cell-type identification architecture, and ColorAE, a color deconvolution methodology to predict pixel-level colour composition. The ensemble method accomplishes thorough cell-type identification without explicit colour deconvolution by utilizing both ColorAE's color composition prediction and U-Net's ability to identify cell types. This method demonstrates how combining several architectural characteristics can lead to increased accuracy in challenging medical image processing jobs. Inspired by SegNet and U-Net, this architecture employs an encoder-decoder structure along with a scoring layer. By capturing essential patterns and reconstructing segmented images, HscoreNet automates scoring, reducing the manual burden and potentially improving diagnostic accuracy with 95.87% precision [43]. Table 2 shows a detailed study of nucleus segmentation in IHC image.

Table 2: Nucleus Segmentation In IHC Image

Literature	Architecture	Purpose	Key Features
Saha and Chakraborty, 2018)	Her2Net [1]	Semantic Segmentation and Classification	• Trapezoidal LSTM connection topology (TLSTM) structure Convolution, max-pooling, spatial pyramid pooling, and TLSTM layers in the convolution and deconvolution process • Preserves cellular and textural characteristics
Fassler *et al.*, 2020	ColorAE: U-Net [42]	Ensemble Method for Cell Type Identification	• Combines U-Net for cell type identification and ColorAE for prediction of pixel-level color composition prediction • Thorough cell type identification without explicit color deconvolution • Leverages benefits of both architectures
Saha *et al.*, 2020	HscoreNet [43]	Estrogen and Progesterone Scoring in Breast IHC Images	• Encoder-decoder structure with a scoring layer • Captures essential patterns and reconstructs segmented images • Automates scoring, reduces the manual burden and enhances diagnostic accuracy
Berezsky *et al.*, 2021	Modified U-Net [41]	Immunohistochemistry Image Segmentation	• U-Net with additional convolutional layer (32x32 receptive field • Improved segmentation accuracy for immunohistochemistry images
Mahanta *et al.*, 2021	IHC-Net [40]	Automated Nuclear Segmentation in Breast Pathology	• Stack of encoder and decoder blocks with skip connections • Utilizes convolution, batch normalization, ReLU activation, and max-pooling for feature extraction • Concatenation, convolution, batch normalization, unpooling for upsampling and reconstruction.

In the context of nucleus segmentation in immunohistochemistry (IHC) images, using U-Net variants demonstrates advantages in segmentation, yet researchers must consider trade-offs like increased complexity, memory usage, and model interpretability when selecting the most suitable approach for specific tasks and resources available. Numerous research gaps have been identified, each offering different challenges and opportunities for improvement.

- The first among these is the constrained generalization observed in current U-Net variants, necessitating a thorough investigation into the factors limiting their adaptability across varied datasets and image characteristics.
- Additionally, concerns related to the computational cost and memory requirements of these models require exploration to enhance efficiency while preserving segmentation accuracy. Addressing the challenge of optimizing U-Net variants through weight-pruning techniques is another significant gap, aiming to simplify model complexity. The imperative need for automating hyper-parameter tuning processes arises due to the time-consuming nature of the current manual trial-and-error approach. Moreover, the incorporation of different loss functions across networks within U-Net variants introduces potential inconsistencies and challenges, warranting research to maintain a cohesive approach.

Research initiatives targeting these gaps hold the promise of advancing the field, resulting in U-Net variants that are more resilient, efficient, and adaptable for nucleus segmentation in IHC images. Given an image stained with IHC, our objective is to segment the nucleus to determine whether it expresses protein or not (brown colour present or not). More precisely, when presented with an image denoted I(x, y), the objective of a semantic segmentation model is to assign a specific semantic label to each pixel within the image, along with an associated labelled mask. Thus we focus on developing an architecture for nucleus segmentation that combines the U-net framework[44] with attention [9] and residual blocks[45] which excels in performance, successfully outperforming other methods.

3. METHODOLOGY

3.1 Proposed Architecture

A novel segmentation architecture utilizing U-Net was introduced for the precise segmentation of nuclei in IHC images. The structure of the network, as depicted in Figure 2 incorporates a distinctive approach in which a double residual block replaces the conventional convolutional layers at every stage of the U-Net encoder-decoder network, and attention blocks are added in the encoder and decoder pathways to focus on relevant regions by assigning higher weights to informative features, resulting in improved boundary, localization accuracy, and overall segmentation performance.

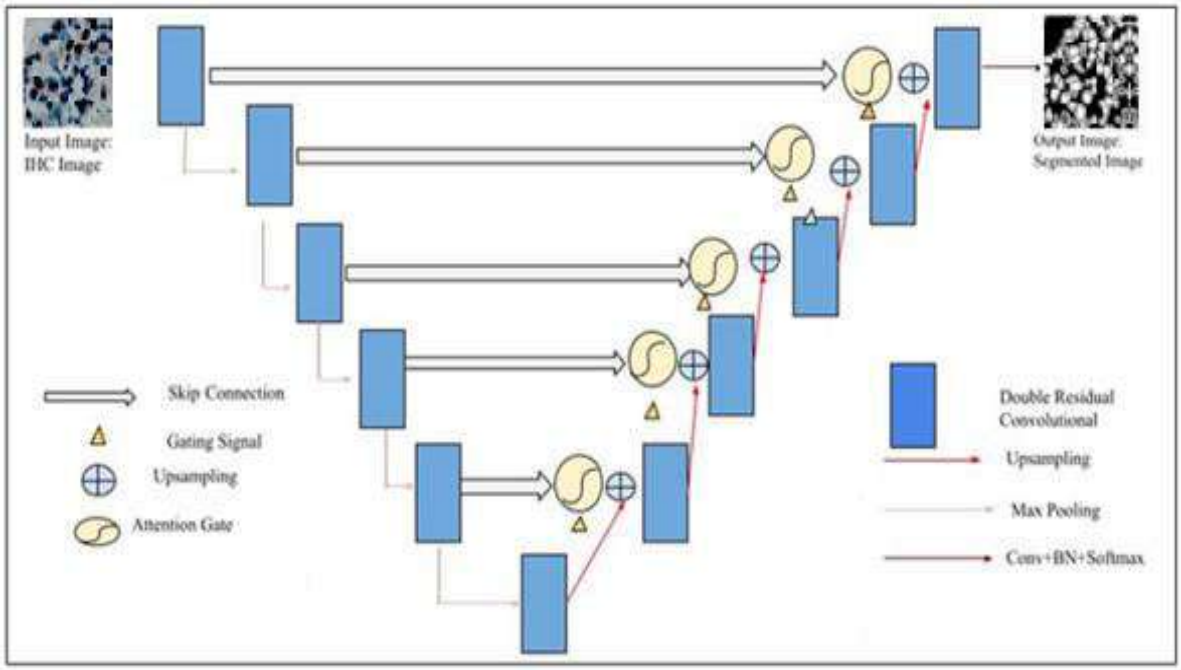


Figure 2: Architecture of the proposed AR-UNet framework with attention and residual blocks for improved nuclei segmentation.

The proposed work in this study comprises ten levels of architecture and follows an encoder-decoder structure with skip connections, facilitating effective image segmentation. The encoder pathway is responsible for reducing the input image resolution to capture high-level features, while the decoder pathway increases the resolution to generate a segmentation mask. The encoding block incorporates six double residual modules, with five downsampling operations applied after each double residual block to extract high-level semantic information. Notably, a stride of 2 is applied to the first convolutional layer in each double residual encoding module, downsampling the feature map by half to preserve positional information. Similarly, the decoder path consists of five double residual modules, where the feature map from the corresponding encoding path is concatenated with the upsampled feature map from the previous module. Throughout the encoder-decoder pathway, attention blocks are strategically inserted at various levels to emphasize relevant regions by assigning higher weights to informative features. Following the last encoding module, a 1×1 convolutional layer and a softmax activation layer are used to project the desired segmented image.

3.2 Double Residual Blocks

Double residual blocks are integrated into the U-Net architecture to capture residual information for improved feature extraction. Residual blocks use skip connections to preserve low-level features across the network, allowing the model to capture fine-grained nuclei details. The concept conveyed by the double residual block involves fusing the input features with the features acquired through two or three convolution operations, addressing the issue of model degradation by determining the output mode of fusion.

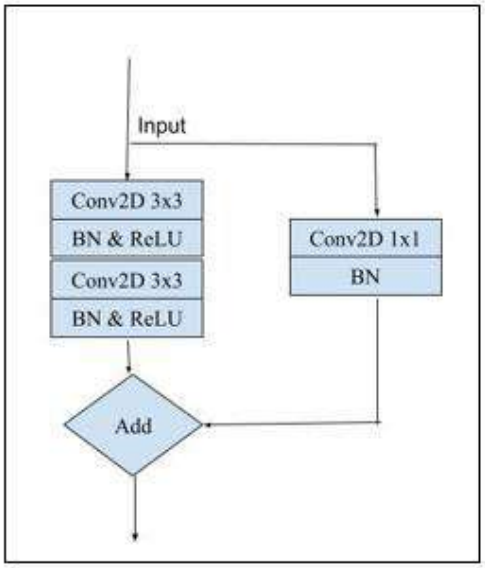


Figure 3: Double Residual Blocks :- Convolutional 2D, Batch Normalization, Activation-ReLU -> Convolutional 2D, Batch Normalization, Activation-ReLU -> Add :Convolutional 2D, Batch Normalization

Figure 3 shows the block diagram of double residual blocks, representing the input (X) to the blocks and the output (Y) following a two-layer convolution operation. This relationship can be formulated as follows:

$$Y = F(X, W_{1,} W_{2)} + W_X X \quad (1)$$

W_1 and W_2 serve as the weight parameters for the convolution layer, while W_x functions as the weight parameter governing the transformation from input X to output.

These blocks alleviate the problem of vanishing gradients and improve gradient flow during training, resulting in better segmentation results.

3.3 Attention Blocks

An attention mechanism is built into the U-Net to improve feature representation. At various levels of the encoder-decoder pathway, attention blocks are inserted to focus on relevant regions by assigning higher weights to informative features. This attention mechanism improves nuclei boundary, localization accuracy, and overall segmentation performance. The additive attention gate (AG), is a critical component of the model's attention mechanism. By selectively highlighting informative spatial regions while suppressing less relevant ones, the AG plays a critical role in improving segmentation accuracy.

Attention mechanisms can be used in either hard or soft attention mode. In this work, soft attention assigns weights to pixels based on relevance, actively suppressing activations in irrelevant regions during training. Given feature map F^l at layer l (l = 1, ..., L), compute the attention weights W^l using a convolutional operation:

$$W^l = \text{Conv2D}(F^l) \quad (2)$$

Apply the softmax function to normalize the attention weights:

$$(A^l) = \text{Softmax}(W^l) \quad (3)$$

Multiply the feature map by the attention weights to obtain the attended feature map F^l_{att}:

$$F_{att}^{l} = F^{l} * A^{l} \quad (4)$$

The additive attention gate (AG) combines activations and contextual information from different scales to compute attention coefficients that highlight relevant spatial regions in the input features. Using these attention coefficients, the AG then scales the features, allowing the model to selectively emphasize significant regions while suppressing less important ones. The proposed model can better focus on informative regions and improve the localization accuracy of nuclei boundaries during the segmentation process by incorporating the AG into the attention mechanism. As a result, the segmentation of the histopathological image becomes more precise and accurate.

Furthermore, at appropriate levels, the attention and residual blocks are integrated into the U-Net. Attention blocks detect spatial dependencies and highlight relevant regions, whereas residual blocks detect fine nuclei details. The proposed AR-UNet aims to achieve superior segmentation accuracy compared to traditional U-Net models by combining the strengths of attention and residual blocks with the U-Net framework. This integration improves the model's ability to focus on important features and capture residual information, resulting in more precise and accurate histopathological image segmentation results.

4. EXPERIMENTAL SETUP

4.1 Dataset

DeepLIIF [41], a public dataset consists of 598 Ki67 IHC images with a resolution of 512 x 512 and a magnification of 40x. These images were obtained from slides of bladder carcinoma. Ki67 expression is a critical marker related to tumour cell proliferation and growth that is widely used in routine pathological investigations.

4.2 Implementation Details

The training data set consists of the IHC input and the respective mask image from the DeepLIIF dataset. Various transformations were used to improve the data set and increase image diversity, resulting in 3000 augmented images. The augmentation pipeline seeks to simulate variation in tissue morphology and staining, thereby improving generalizability by capturing diverse image variations[46]. The protein-expressed nucleus(positive) is represented in red in the segmentation mask, while the nucleus that does not express protein is represented in blue (negative), as shown in Figure 4.

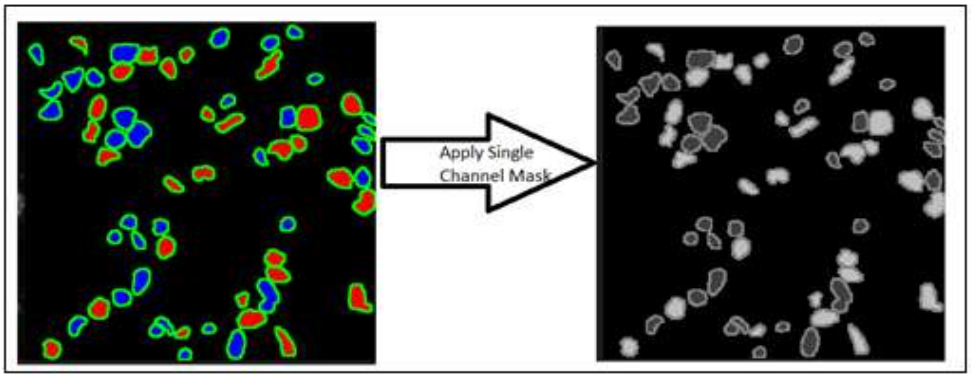


Figure 4: Single-channel mask applied to IHC masked image for multiclass nucleus segmentation

The model is trained with IHC images as input (Figure 5a) and the corresponding single-channel images as ground truth labels (Figure 5b). The model learns to map the input IHC images to the corresponding segmentation masks during the training process, effectively segmenting different classes or regions in the image.

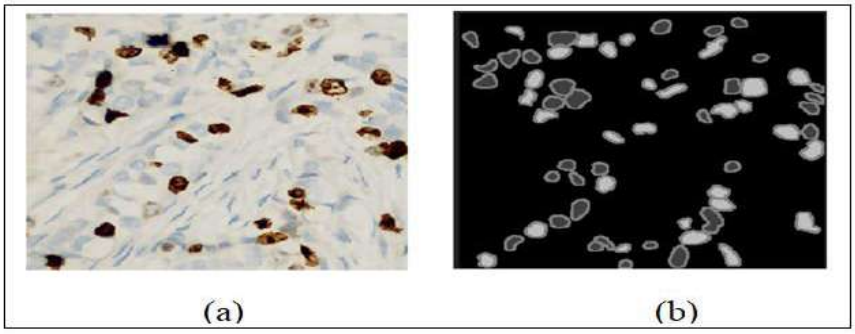


Figure 5: The proposed AR-UNet is trained by (a) original image and (b) ground truth image.

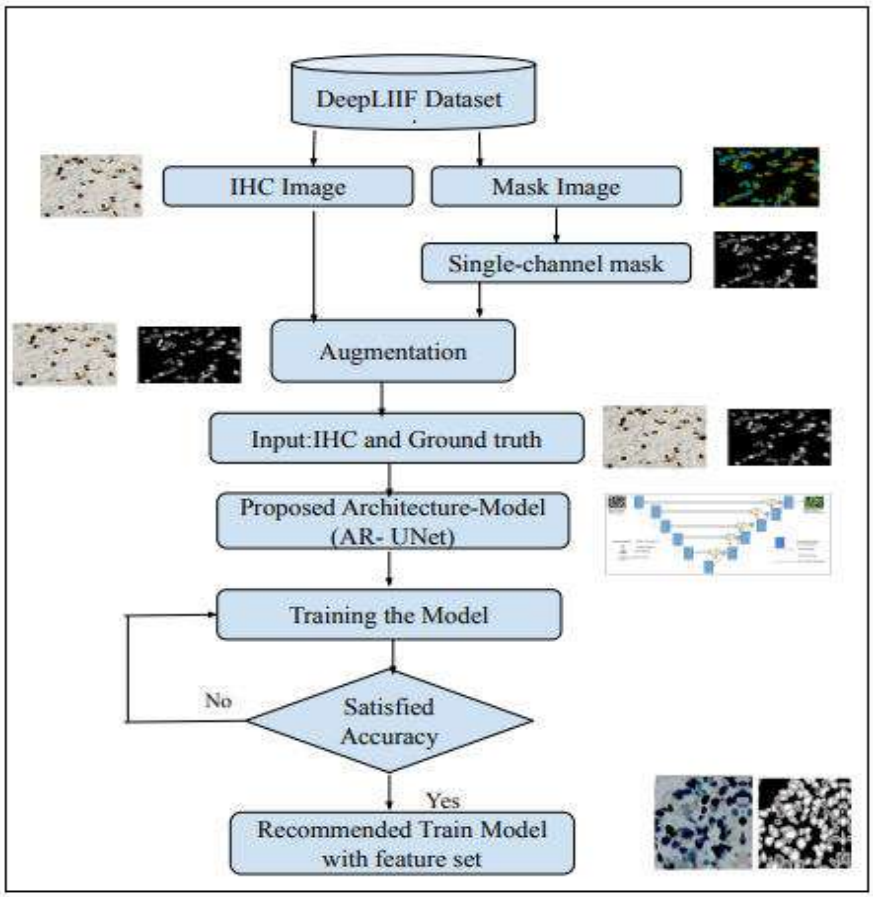


Figure 6: Implementation Steps for Proposed AR-UNet: Integrated double residual blocks and attention blocks in U-Net for nucleus segmentation of IHC image

Implementation Steps for training the model are described in Figure 6. The data set is used for tasks that require multiclass segmentation. The data set begins with IHC (immunohistochemistry) images as input. The ground truth of the IHC images is then processed to generate single-channel masks, each pixel in the mask representing a different class or category in the image. Figure 4 shows the single-channel mask. Pixel-wise mapping is performed to convert the original multiclass segmentation problem into a single mask, assigning a unique value to each class or category. The single-channel mask is the result that condenses multiclass information into a single channel, making it suitable for training the segmentation model.

4.3 Experimental Results and Analyses

The proposed Attention-Enhanced and Residual U-Net (AR-UNet) was evaluated and compared to other variants of U-Net, including basic U-Net, U-Net with ResNet backbone, and Attention U-Net in IHC-stained tissues from the test data set. Semantic segmentation models were efficiently evaluated with the help of precision, accuracy, precision, recall, F1 score, IoU, and loss[47].

Precision (Pr) measures the accuracy of the positive predictions made by the model

$$\mathrm{Pr}(\%) = \frac{TP}{TP+FP} \times 100 \quad (5)$$

Recall (Re) quantifies the model's ability to identify all positive instances among the actual positives correctly

$$Re(\%) = \frac{TP}{TP+FN} \times 100 \quad (6)$$

Where,

TP (True Positives): The number of pixels that are correctly predicted as positive.

FP (False Positives): The number of pixels that are predicted as positive but are actually negative in the ground truth.

FN (False Negatives): The number of pixels that are actually positive in the ground truth but are predicted as negative.

F1 score is a balanced metric, considering both false positives and false negatives.

$$F1\ score(\%) = 2 \times \left(\frac{\mathrm{Pr} \times Re}{\mathrm{Pr} + Re}\right) \times 100 \quad (7)$$

Accuracy is defined as the ratio of correctly predicted instances to the total number of instances.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (8)$$

Where,

TN (True Negative): The number of pixels that are correctly predicted as negative.

Intersection over Union (*IoU*) is a common metric used for evaluating the performance of segmentation models. *IoU* measures the overlap between the predicted segmentation and the ground truth segmentation.

$$IoU = \frac{TP}{TP+FN+FP} \quad (9)$$

Loss functions quantify the error between predicted and true values during model training. The chosen loss function, such as the combination of Dice Loss and Categorical Focal Loss guides the model towards optimal parameter values. Minimizing the loss function is the primary objective during training.

Let $dice\_loss(y_{true}, y_{pred})$ be the Dice Loss, and $\mathrm{focal\_loss}(y_{true}, y_{pred})$ be the Categorical Focal Loss. Also, let class_weights = $[w1, w2, w3, w4]$ represent the class weights for the Dice Loss. These weights are associated with the weight for the first, second, third, and fourth class in our segmentation problem.

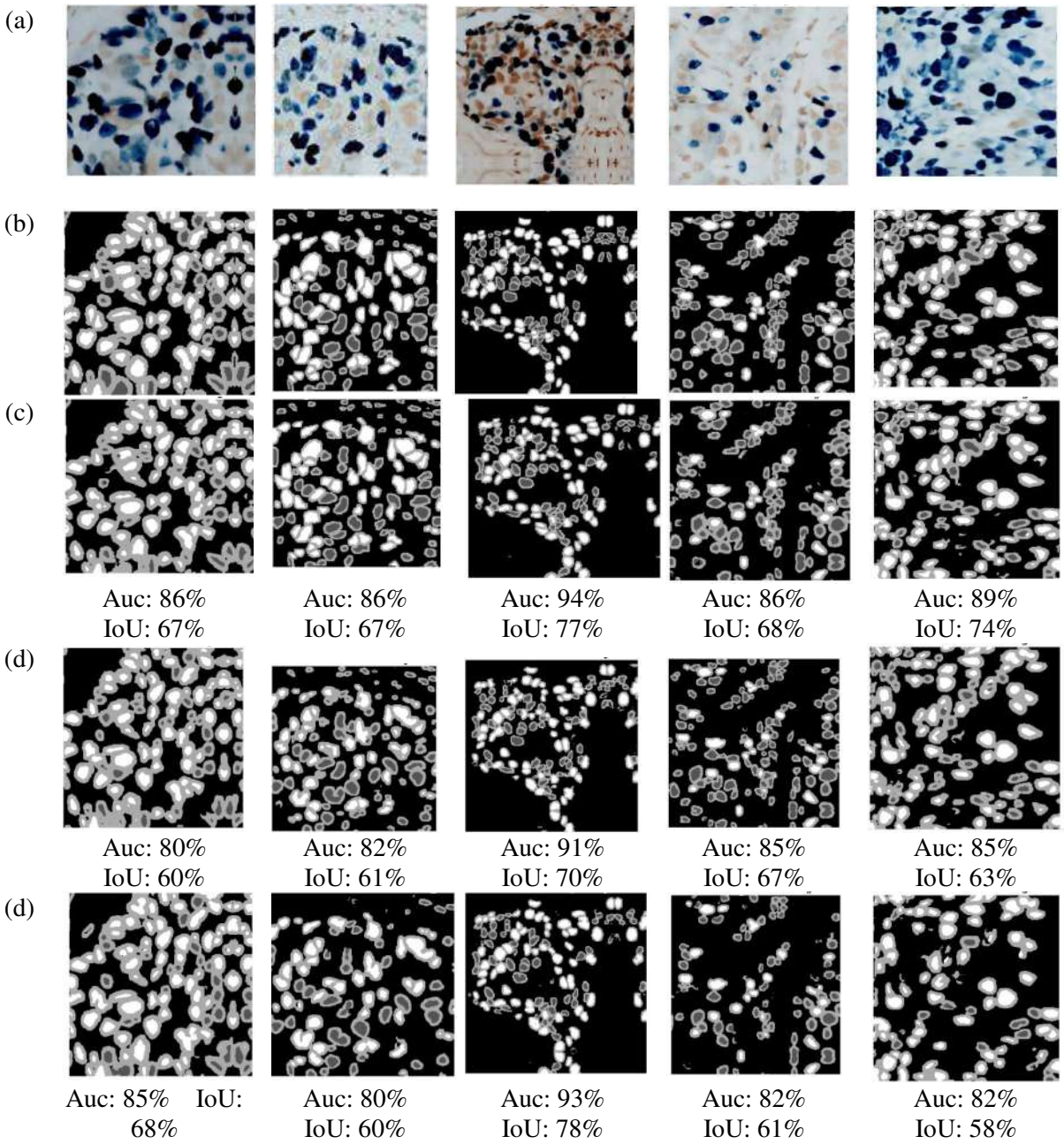
$$total\_loss = dice\_loss(y_{true}, y_{pred}, \mathrm{class\_weights}) + \mathrm{focal\_loss}(y_{true}, y_{pred}) \quad (10)$$

Where

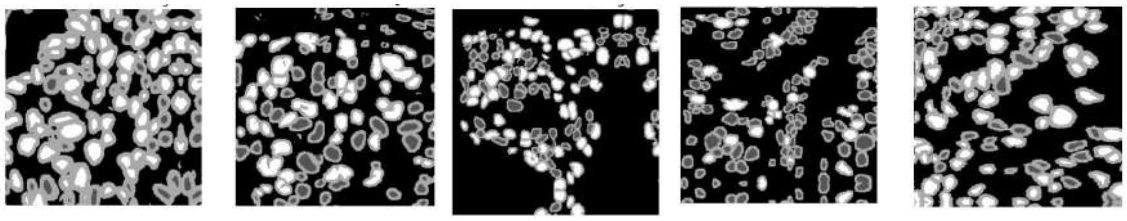
y_{true} typically refers to the actual values

y_{pred} refers to the predicted values

In Figure 7, the outcomes of our proposed model, the Attention-Enhanced and Residual U-Net (AR-UNet), are presented and compared with other models. The attention mechanism is used to highlight important regions, and residual blocks are used to capture fine nuclei details, resulting in improved feature representation and localization. AR-UNet can accurately highlight nucleus boundaries and distinguish between positive and negative regions by leveraging the attention mechanism. Furthermore, because of the integration of residual blocks, the model captures fine-grained details of the nuclei, further improving the segmentation accuracy.



(e)



Auc: 86% IoU: 69% | Auc: 84% IoU: 68% | Auc: 93% IoU: 79% | Auc: 87% IoU: 71% | Auc: 87% IoU: 70%

Figure 7: Results on several samples. Each row presents: (a) input image, (b) ground truth- white: positive nucleus, grey: negative nucleus, border of nucleus, background (c) U-Net [6] (d) ResNet U-Net [48] (e) Attention U-Net [9], (f) Proposed AR-UNet

Table 3 presents a comparative analysis of the proposed AR-UNet with other variants based on evaluation metrics. The AR U-Net is the one to win.

Table 3: Comparison of Proposed RA U-Net and U-Net based on performance metrics

Method	**Metrics**					
	Precision (%)	Recall (%)	F1-score (%)	IOU (%)	Accuracy (%)	Loss
U-Net	85	85	80	68	85%	0.47
UNet- ResNet	87	87	83	71	87	0.34
Attention U-Net	91	91	89	80	91	0.26
Proposed AR-UNet	**92**	**92**	**89**	**81**	**92**	**0.25**

The AR-UNet correctly identified the nuclei and relevant structures in the histopathological image with an accuracy of 92% with a low loss value of 0.25. The F1 score of 89% demonstrated balanced performance, striking a strong balance between precision and recall with 92%, which measures the model's ability to capture the majority of true positive instances. The IoU of 81% indicated a close alignment between the AR-UNet's predicted segmentation mask and the ground truth, accurately capturing the boundaries of nuclei. AR U-Net receives an F1 score of 89%.

The outstanding performance of the AR-UNet highlights its potential as a powerful tool in histopathological image analysis. Its accurate segmentation of the nuclei and relevant structures can greatly aid pathologists and researchers in diagnosing and analyzing tissue morphology tumour cell characteristics and proliferation markers in histopathological images. The promising results of AR-UNet offer enormous opportunities to advance medical research and improve clinical practices in the field of digital pathology.

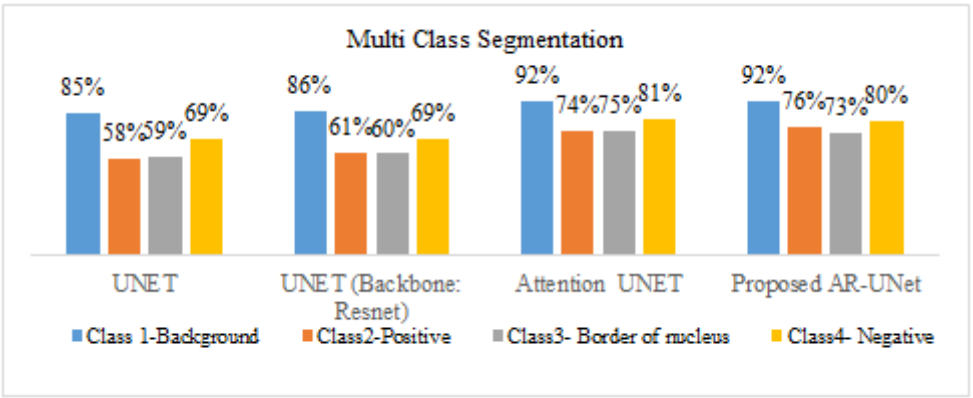


Figure 8: Comparison of the graphical plot of AR-UNet with other U-Net variants in terms of labels of multitask segmentations (classes)

In Figure 8 & 9, gives a comprehensive comparative evaluation is performed to compare the proposed Attention-Enhanced and Residual U-Net (AR U-Net) with other U-Net variants in terms of multitask segmentations or class labels. The evaluation involves four classes: Class 1 (Background), Class 2 (Positive), Class 3 (Border of Nucleus), and Class 4 (Negative). The comprehensive comparative evaluation highlights the strengths of the AR U-Net proposed in multitask segmentation or class-wise precision for histopathological image segmentation. The model demonstrates superior accuracy performance. Our research identifies current trends and innovations within the realm of segmentation. This encompasses the incorporation of deep learning U-Net architecture, attention, and residual blocks for nucleus segmentation in IHC images.

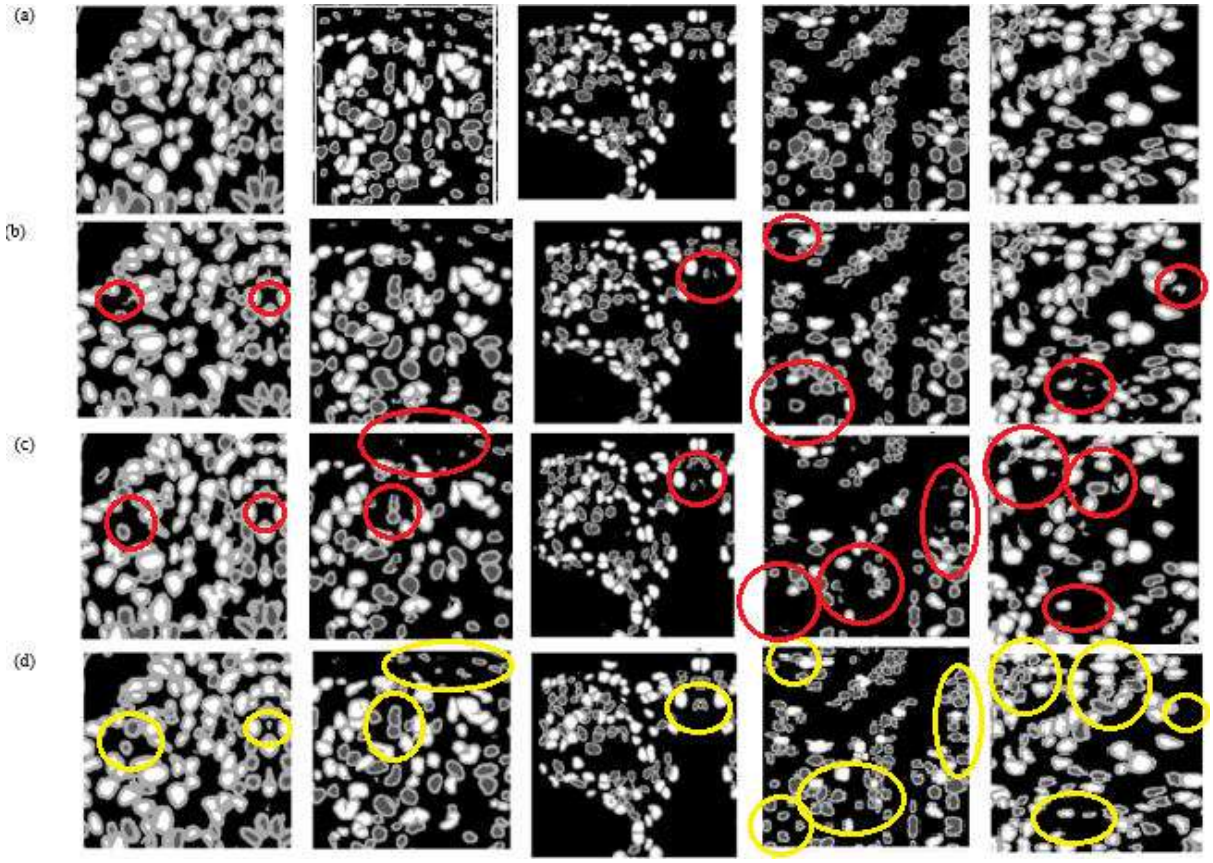


Figure 9: Results on five samples with positive, negative nucleus ,border of nucleus and background class. In each row we present: (a) ground truth and (b) predicted images of U-Net (c) UNet- ResNet (d) Attention U-Net (e) Proposed AR-UNet. Red circle shows pixels misclassified while yellow circle shows correctly classified label

5. CONCLUSION

This paper presents AR U-Net, which was designed and developed for nucleus segmentation in IHC images. The proposed architecture incorporates attention mechanisms and residual blocks into the U-Net architecture. Attention mechanisms dynamically highlight important features and suppress irrelevant ones while using residual blocks, helping to learn more features that often contain critical information for segmentation, resulting in improved feature representation and localization. Extensive tests on the DeepLIIF testing dataset revealed that AR U-Net performed admirably, with an F1 score of 89%, IOU of 81%, an accuracy of 92%, and a negligible loss of 0.25. Visual segmentation results and performance metrics were compared with other U-Net variants, such as U-Net, U-NET with a ResNet backbone, and Attention U-Net, validating the proposed AR U-Net's superiority in accurately identifying and segmenting nuclei in histopathological images.

There are several possibilities for the future scope of this work. After segmentation, we intend to measure protein content within the nucleus by employing an optimal method. The proposed AR U-Net expands the field of histopathological image analysis and has the potential to significantly contribute to the accurate identification and segmentation of nuclei in histopathological images to help pathologists and researchers in cancer diagnosis and analysis.

Ethics approval and consent to participate

The research conducted in this study involving participants was reviewed and approved by the International Journal of Computers and Applications. All participants provided informed consent before participating in the study.

Funding

This study did not receive any funding in any way.

Data Availability Statement

Not applicable.

Consent for publication

All authors have provided their consent for the publication of this paper.

Declaration of Conflict of Interests

The author(s) declared that they have no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

REFERENCES

[1] Saha M, Chakraborty C. Her2Net: A deep framework for semantic segmentation and classification of cell membranes and nuclei in breast cancer evaluation. IEEE Transactions on Image Processing. 2018;27(5):2189–2200.

[2] Hegde RB, Prasad K, Hebbar H, Singh BMK. Comparison of traditional image processing and deep learning approaches for classification of white blood cells in peripheral blood smear images. Biocybernetics and Biomedical Engineering. 2019;39(2):382–392.

[3] Karri M, Annavarapu CSR, Mallik S, Zhao Z, Acharya UR. Multi-class nucleus detection and classification using deep convolutional neural network with enhanced high dimensional dissimilarity translation model on cervical cells. Biocybernetics and Biomedical Engineering. 2022;42(3):797–814.

[4] Vetriselvi D, Thenmozhi R. A Systematic Literature Review on Deep Learning Based Medical Image Segmentation. International Journal of Intelligent Systems and Applications in Engineering. 2023;11(5s):519–526.

[5] Benny S, Varma SL. Semantic segmentation in immunohistochemistry breast cancer image using deep learning. In: 2021 International Conference on Advances in Computing, Communication, and Control (ICAC3). IEEE; 2021. p. 1–3.

[6] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18. Springer; 2015. p. 234–241.

[7] Zhou Z, Rahman Siddiquee MM, Tajbakhsh N, Liang J. Unet++: A nested u-net architecture for medical image segmentation. In: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, P. Springer; 2018. p. 3–11.

[8] Qin X, Zhang Z, Huang C, Dehghan M, Zaiane OR, Jagersand M. U2-Net: Going deeper with nested U-structure for salient object detection. Pattern recognition. 2020;106:107404.

[9] Oktay O, Schlemper J, Folgoc L Le, Lee M, Heinrich M, Misawa K, Mori K, McDonagh S, Hammerla NY, Kainz B. Attention u-net: Learning where to look for the pancreas. arXiv preprint arXiv:1804.03999. 2018.

[10] Diakogiannis FI, Waldner F, Caccetta P, Wu C. ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data. ISPRS Journal of Photogrammetry and Remote Sensing. 2020;162:94–114.

[11] Alom MZ, Hasan M, Yakopcic C, Taha TM, Asari VK. Recurrent residual convolutional neural network based on u-net (r2u-net) for medical image segmentation. arXiv preprint arXiv:1802.06955. 2018.

[12] Li Z, Zhang H, Li Z, Ren Z. applied sciences Residual-Attention UNet ++ : A Nested Residual-Attention U-Net for Medical Image Segmentation. 2022.

[13] Huang H, Lin L, Tong R, Hu H, Zhang Q, Iwamoto Y, Han X, Chen YW, Wu J. UNet 3+: A Full-Scale Connected UNet for Medical Image Segmentation. ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings. 2020;2020-May(ii):1055–1059. doi:10.1109/ICASSP40776.2020.9053405

[14] Amer A, Lambrou T, Ye X. MDA-unet: a multi-scale dilated attention U-net for medical image segmentation. Applied Sciences. 2022;12(7):3676.

[15] Gudhe NR, Behravan H, Sudah M, Okuma H, Vanninen R, Kosma V-M, Mannermaa A. Multi-level dilated residual network for biomedical image segmentation. Scientific Reports. 2021;11(1):14105.

[16] Li X, Pang S, Zhang R, Zhu J, Fu X, Tian Y, Gao J. ATTransUNet: An enhanced hybrid transformer architecture for ultrasound and histopathology image segmentation. Computers in Biology and Medicine. 2023;152:106365.

[17] Kanadath A, Jothi JAA, Urolagin S. Histopathology image segmentation using MobileNetV2 based U-net model. In: 2021 International Conference on Intelligent Technologies (CONIT). IEEE; 2021. p. 1–8.

[18] Yang Y, Dasmahapatra S, Mahmoodi S. ADS_UNet: A nested UNet for histopathology image segmentation. Expert Systems with Applications. 2023;226:120128.

[19] Kiran I, Raza B, Ijaz A, Khan MA. DenseRes-Unet: Segmentation of overlapped/clustered nuclei from multi organ histopathology images. Computers in Biology and Medicine. 2022;143:105267.

[20] He H, Zhang C, Chen J, Geng R, Chen L, Liang Y, Lu Y, Wu J, Xu Y. A hybrid-attention nested UNet for nuclear segmentation in histopathological images. Frontiers in Molecular Biosciences. 2021;8:614174.

[21] Chanchal AK, Lal S, Kini J. High-resolution deep transferred ASPPU-Net for nuclei segmentation of histopathology images. International journal of computer assisted radiology and surgery. 2021;16:2159–2175.

[22] Zeng Z, Xie W, Zhang Y, Lu Y. RIC-Unet: An improved neural network based on Unet for nuclei segmentation in histology images. Ieee Access. 2019;7:21420–21428.

[23] Li J, Li X. MIU-Net: MIX-Attention and Inception U-Net for Histopathology Image Nuclei Segmentation. Applied Sciences. 2023;13(8):4842.

[24] Dinh T Le, Kwon S-G, Lee S-H, Kwon K-R. Breast tumor cell nuclei segmentation in histopathology images using efficientunet++ and multi-organ transfer learning. Journal of Korea Multimedia Society. 2021;24(8):1000–1011.

[25] Wang Y, Li X, Liu J, Shen L, Sun K, Wang S. A Contrastive Learning-based PPC-UNet for Colorectal Histopathology Whole Slide Image Segmentation. In: 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). IEEE; 2021. p. 2072–2079.

[26] Hamida A Ben, Devanne M, Weber J, Truntzer C, Derangère V, Ghiringhelli F, Forestier G, Wemmert C. Weakly Supervised Learning using Attention gates for colon cancer histopathological image segmentation. Artificial Intelligence in Medicine. 2022;133:102407.

[27] Dabass M, Vashisth S, Vig R. Attention-Guided deep atrous-residual U-Net architecture for automated gland segmentation in colon histopathology images. Informatics in Medicine Unlocked. 2021;27:100784.

[28] Li Y, Xu Z, Wang Y, Zhou H, Zhang Q. Su-net and du-net fusion for tumour segmentation in histopathology images. In: 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI). IEEE; 2020. p. 461–465.

[29] Ibtehaz N, Rahman MS. MultiResUNet: Rethinking the U-Net architecture for multimodal biomedical image segmentation. Neural networks. 2020;121:74–87.

[30] He P, Qu A, Xiao S, Ding M. DETisSeg: A dual-encoder network for tissue semantic segmentation of histopathology image. Biomedical Signal Processing and Control. 2024;87:105544.

[31] Van Rijthoven M, Balkenhol M, Siliņa K, Van Der Laak J, Ciompi F. HookNet: Multi-resolution convolutional neural networks for semantic segmentation in histopathology whole-slide images. Medical image analysis. 2021;68:101890.

[32] Kaur A, Kaur L, Singh A. GA-UNet: UNet-based framework for segmentation of 2D and 3D medical images applicable on heterogeneous datasets. Neural Computing and Applications. 2021;33(21):14991–15025.

[33] Samanta P, Singhal N. YAMU: yet another modified U-Net architecture for semantic segmentation. In: International Conference on Medical Imaging with Deep Learning. PMLR; 2022. p. 1019–1033.

[34] Ahmad I, Xia Y, Cui H, Islam ZU. DAN-NucNet: A dual attention based framework for nuclei segmentation in cancer histology images under wild clinical conditions. Expert Systems with Applications. 2023;213:118945.

[35] Lal S, Das D, Alabhya K, Kanfade A, Kumar A, Kini J. NucleiSegNet: Robust deep learning architecture for the nuclei segmentation of liver cancer histopathology images. Computers in Biology and Medicine. 2021;128:104075.

[36] Kong Y, Genchev GZ, Wang X, Zhao H, Lu H. Nuclear segmentation in histopathological images using two-stage stacked U-nets with attention mechanism. Frontiers in Bioengineering and Biotechnology. 2020;8:573866.

[37] Li Z, Zhang H, Li Z, Ren Z. Residual-Attention UNet++: A Nested Residual-Attention U-Net for Medical Image Segmentation. Applied Sciences. 2022;12(14):7149.

[38] Le Dinh T, Lee S-H, Kwon S-G, Kwon K-R. Cell Nuclei Segmentation in Cryonuseg dataset using Nested Unet with EfficientNet Encoder. In: 2022 International Conference on Electronics, Information, and Communication (ICEIC). IEEE; 2022. p. 1–4.

[39] Chowdary GJ, Yogarajah P. Nucleus segmentation and classification using residual SE-UNet and feature concatenation approach incervical cytopathology cell images. Technology in Cancer Research & Treatment. 2023;22:15330338221134832.

[40] Mahanta LB, Hussain E, Das N, Kakoti L, Chowdhury M. IHC-Net: A fully convolutional neural network for automated nuclear segmentation and ensemble classification for Allred scoring in breast pathology. Applied Soft Computing. 2021;103. doi:10.1016/j.asoc.2021.107136

[41] Berezsky O, Pitsun O, Derysh B, Pazdriy I, Melnyk G, Batko Y. Automatic segmentation of immunohistochemical images based on U-net architecture. In: 2021 IEEE 16th International Conference on Computer Sciences and Information Technologies (CSIT). Vol. 1. IEEE; 2021. p. 29–32.

[42] Fassler DJ, Abousamra S, Gupta R, Chen C, Zhao M, Paredes D, Batool SA, Knudsen BS, Escobar-Hoyos L, Shroyer KR, et al. Deep learning-based image analysis methods for brightfield-acquired multiplex immunohistochemistry images. Diagnostic Pathology. 2020;15(1). doi:10.1186/s13000-020-01003-0

[43] Saha M, Arun I, Ahmed R, Chatterjee S, Chakraborty C. HscoreNet: A Deep network for estrogen and progesterone scoring using breast IHC images. Pattern Recognition. 2020;102:107200.

[44] Ronneberger, Olaf, Philipp Fischer and TB. U-Net: Convolutional Networks for Biomedical Image Segmentation. IEEE Access. 2021;9:16591–16603. doi:10.1109/ACCESS.2021.3053408

[45] Diakogiannis FI, Waldner F, Caccetta P, Wu C. ResUNet-a: A deep learning framework for semantic segmentation of remotely sensed data. ISPRS Journal of Photogrammetry and Remote Sensing. 2020;162(December 2019):94–114. https://doi.org/10.1016/j.isprsjprs.2020.01.013. doi:10.1016/j.isprsjprs.2020.01.013

[46] Paquin F, Rivnay J, Salleo A, Stingelin N, Silva C. Multi-phase semicrystalline microstructures drive exciton dissociation in neat plastic semiconductors. J. Mater. Chem. C. 2015;3:10715–10722. http://xlink.rsc.org/?DOI=C5TC02043C. doi:10.1039/b000000x

[47] Reddy CRP, Srinagesh DA. International Journal of INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING Deep Learning Algorithms to Detect Human Pancreatic Cancer from MRI Scan Images. Original Research Paper International Journal of Intelligent Systems and Applications in Engineering IJISAE. 2023;2023(6s):584–591. http://crdd.osdd.net/raghava/pcmdb/

[48] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 2016;2016-Decem:770–778. doi:10.1109/CVPR.2016.90