



## DDUSeg-Net as a Design of Convolutional Neural Network Architecture for Semantic Segmentation in Cervical Cancer

Rudiansyah<sup>1</sup>Anita Desiani<sup>2\*</sup>Dian Palupi Rini<sup>3</sup>Lucky Indra Kesuma<sup>4</sup>

<sup>1</sup>Doctoral Program in Engineering Faculty of Engineering Universitas Sriwijaya, Indralaya, Indonesia

<sup>2</sup>Department of Mathematics, Universitas Sriwijaya, Indralaya, Indonesia

<sup>3</sup>Department of Computer Science, Universitas Sriwijaya, Indralaya, Indonesia

<sup>4</sup>Department of Computer Science, Universitas Sjakhyakirti, Palembang, Indonesia

\* Corresponding author's Email: [anita\\_desiani@unsri.ac.id](mailto:anita_desiani@unsri.ac.id)

---

**Abstract:** Cervical cancer is a significant health issue for women and ranks fourth in the world among the most dangerous cancers. An automatic diagnostic system is needed for pap smears to assist medical experts in diagnosing cervical cancer. One of the automatic diagnosis systems in detecting cervical cancer is semantic segmentation. Convolutional Neural Networks (CNN), particularly the U-Net architecture, have been widely used for segmentation tasks in medical imaging. Although U-Net has demonstrated effectiveness, its performance on low-quality images is often suboptimal, with issues such as loss of fine details during the down-sampling process. This study combines image enhancement and Double Dropout USeg-Net (DDUSeg-Net). Image enhancement techniques are applied to pap-smear images to improve image quality such as Gamma Correction for enhanced contrast, and Median Filtering for reduced noise. The proposed DDUSeg-Net architecture builds on the U-Net model by incorporating two U-Net blocks for more detailed feature extraction. SegNet's pooling indices are added to preserve spatial information during the segmentation process. Additionally, dropout layers are introduced to prevent overfitting and reduce the model's overall complexity. The image enhancement results indicate that the Mean Squared Error (MSE), Peak Signal to Ratio (PNSR), and Structural Similarity Image Index (SSIM) are above 85%. The performance metrics for the DDUSeg-Net model obtained accuracy, precision, recall, and F1-score above 90%. This analysis used 2D pap-smear images from the Herlev dataset. Overall, the combination of image enhancement and DDUSegNet demonstrates strong robustness in the segmentation of pap-smear images, effectively balancing the detection of the intersection areas between the nucleus, cytoplasm, and background.

**Keywords:** Image enhancement, Segmentation, U-Net, Seg-Net.

---

### 1. Introduction

Cervical cancer is a significant health issue for women and ranks fourth in the world among the most dangerous cancers [1]. Detection of cervical cancer can be done through a pap smear test [2]. The pap-smear test examines cervical cells under a microscope to determine the potential of these cells to develop into cancer in the future [3]. However, the manual pap-smear examination is relatively time-consuming and subjective, making it prone to errors [4]. An automatic diagnostic system is needed for pap smears to assist medical experts in diagnosing

cervical cancer. One of the automatic diagnosis systems in detecting cervical cancer is image segmentation. Image segmentation is an image processing technique used to separate objects from the background for easier analysis. Segmentation in pap-smear images separates the nucleus and cytoplasm from other areas. The semantic segmentation methods are the *Convolutional Neural Network* (CNN).

CNN can be used in image-processing tasks such as image classification and image segmentation. CNN is capable of effectively extracting features from large datasets [5, 6]. One of the popular CNN architectures for segmentation is U-Net. The U-Net

consists of an encoder and a decoder connected by a bridge [7]. The encoder extracts features from the input image, while the decoder reconstructs the image size based on the features extracted by the encoder [8]. Desiani et al [9] applied U-Net for cervical cancer segmentation and obtained accuracy, sensitivity, specificity, and F1-Score below 80%. Arum et al [10] applied U-Net for cervical cancer segmentation and obtained accuracy, sensitivity, and IoU below 80%. Nazir et al [11] applied the U-Net for cervical cancer segmentation and obtained sensitivity above 90%. However, the IoU and specificity were below 80%. The U-Net is generally more effective with high-quality images, but pap-smear images sometimes have low quality, such as low contrast, non-uniform color, and noise [12, 13].

The low quality can be addressed through image enhancement techniques. Image enhancement methods include Gamma Correction and Median Filter. Gamma correction improves the contrast of dark images through non-linear pixel transformation [9]. Sangeeta [14] applied Gamma Correction with weight distribution and CLAHE on cervical segmentation and obtained an SSIM of 0.5 and a PSNR of 12 dB. The low SSIM and PSNR can be attributed to factors such as increased noise intensity due to image enhancement. One method to reduce noise in images is the median filter, which works by evaluating the light levels using the median value [15]. Shiney et al [16] applied a median filter without contrast enhancement in cervical cancer and obtained an SSIM value above 0.9. However, the PSNR was below 25 dB, which indicates that a combination of contrast enhancement and noise reduction is necessary to achieve a high-quality image. High-quality images can lead to more accurate and optimal segmentation. The U-Net applies down-sampling (reducing resolution) in the encoder to capture global features from the image. However, down-sampling can lead to the loss of important fine details, particularly in small objects or complex regions of the image, which are crucial for segmentation [17]. To overcome the limitations of U-Net, a new approach is needed to enhance the model's capacity to capture fine details in images, such as Double U-Net.

The Double U-Net is a development of the U-Net with the addition of a second U-Net block. The Double U-Net can capture more accurate and detailed features compared to a single U-Net, as the second U-Net block refines the results from the first block [18]. The Double U-Net has been applied in several studies, such as the segmentation of skin lesions, the segmentation of optic cups and optic discs, and the segmentation of brain metastases [19-21]. Double U-Net enhances segmentation accuracy but increases

model complexity. The high complexity arises from the double process in down-sampling and up-sampling, which demand more memory and computation time, potentially reducing efficiency, particularly with large or high-resolution datasets [22]. Additionally, the double down-sampling in Double U-Net can lead to loss of spatial information. The loss of spatial information, particularly in small objects or detailed areas, negatively impacts the final segmentation quality [23]. In addition, Double U-Net has a complex model that has the potential to overfit, especially on small datasets. Overfitting causes a decrease in the model's generalization ability on new data that has not been trained [24]. Another CNN architecture that can overcome the weaknesses of double U-Net is SegNet.

SegNet is a network architecture specifically designed for image segmentation. The SegNet uses the VGG16 backbone, which consists of deeper convolutional layers [25]. VGG16 can extract more complex and detailed features, enabling the model to recognize various patterns and characteristics in images [26]. SegNet can preserve spatial information by using a max-pooling index during down-sampling, which ensures crucial spatial details are reconstructed correctly during up-sampling [27]. Teixeira et. al [28] applied the SegNet for cervical segmentation and obtained accuracy above 90%. However, this study obtained sensitivity below 70%. Yu [29] applied the SegNet for epithelial segmentation in cervical images and obtained accuracy and specificity above 90%. However, this study obtained sensitivity below 75%. The pooling indices in max pooling can retain spatial information in Double U-Net. Additionally, the large number of layers in Double U-Net increases the number of parameters, which can lead to overfitting [30]. One method that can reduce the number of parameters and prevent overfitting is Dropout [31]. Dropout deactivates neurons randomly during training, reducing parameters and helping prevent overfitting [32], [33]. Rasheed et al [34] applied Dropout in U-Net for cervical cancer segmentation and obtained accuracy and F1-Score above 90%. Nazir et al [11] applied Dropout on U-Net in cervical cancer segmentation and obtained sensitivity above 90%, but the specificity and IoU were still below 80%.

This study not only focuses on cervical cancer cell segmentation but also combines image enhancement methods and new CNN architectures. The image enhancement process includes gamma correction and a median filter. The gamma correction is used to enhance contrast and improve pixel intensity distribution in pap-smear images, while the median filter is used to reduce noise in pap-smear

images. In the segmentation stage, this proposed method developed a new architecture called DDUSeg-Net, which is an enhancement of the traditional U-Net. The DDUSeg-Net combines two U-Net blocks where each U-Net block is replaced with SegNet and dropout layers are added to every encoder and decoder layer. The DDUSegNet is designed to improve segmentation performance and maintain spatial information by using pooling indices. The dropout layers within the encoder and decoder help reduce the number of model parameters, optimizing the architecture for greater efficiency. By combining image enhancement and DDUSeg-Net architecture, the proposed method is expected to provide more accurate and valid results in identifying changes in cervical cancer cells. The image enhancement evaluation is measured based on *Mean Squared Error (MSE)*, *Peak Signal-to-Noise Ratio (PSNR)*, and *Structural Similarity Index (SSIM)*, while model performance evaluation is measured by calculating accuracy, sensitivity, specificity, F1-Score, Intersection over Union IoU, and G-mean. The proposed method is expected to help obtain the features needed by medical personnel to diagnose cervical cancer so that the diagnosis results are more valid and accurate. This paper is structured as follows: section 2 describes the proposed method, which consists of data description, image enhancement, DDUSeg-Net architecture, and evaluation. Section 3 describes the research results and discussions. Section 4 concludes this paper.

## 2. Research method

The study proposes a method to combine image enhancement with a new CNN architecture. The image enhancement stage is performed using gamma correction and median filter. The image enhancement aims to optimize image quality by adjusting brightness levels and reducing noise. The new architecture is named DDUSegNet, which is specifically designed to obtain important features in pap smears to diagnose cervical cancer. The workflow of this study is illustrated in Figure 1.

Based on Figure 1, the input image for the proposed method uses a pap-smear image. In the image enhancement using Gamma Correction and median filter. The dataset will be divided into training data, validation data, and testing data. The training data is used to train the DDUSeg-Net model, while the validation data is used to assess the performance of the model during training. After completing the training stage and saving the results, the process moves to the testing stage to evaluate the model's segmentation performance in segmentation cervical cancer. The effectiveness of image enhancement will be assessed using MSE, PSNR, and SSIM, while the model performance will be evaluated based on accuracy, sensitivity, specificity, F1-score, *Intersection over Union (IoU)*, and *Geometric Mean (G-Means)*.

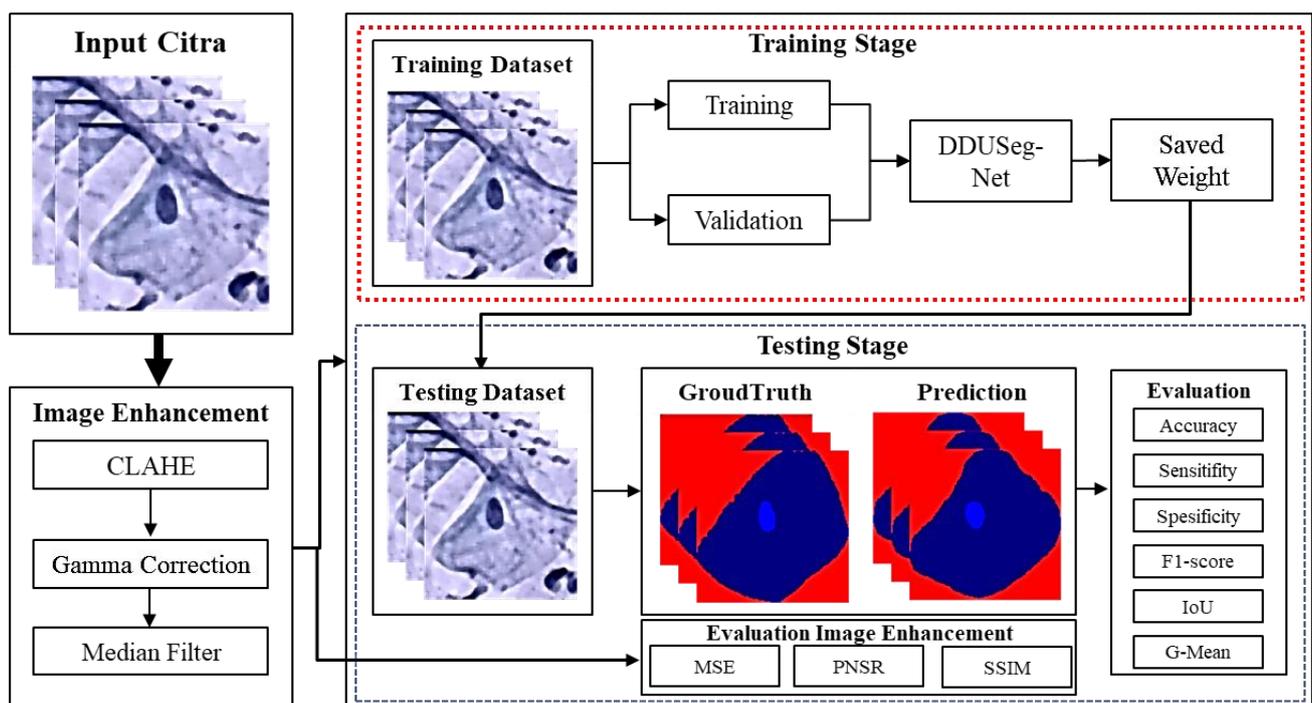


Figure. 1 Workflow of the proposed method for cervical cancer segmentation

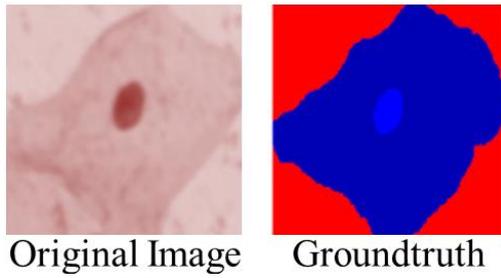


Figure. 2 The dataset of pap-smear image

## 2.1 Data description

The study used publicly accessible pap-smear images from the Herlev dataset [38]. The dataset consists of 917 images in 2D format (JPG), each sized  $150 \times 140$  pixels. The 2D pap smear images are designed to capture cellular details that are important for the diagnosis of cervical cancer. The dataset is illustrated in Figure 2.

Based on Figure 2, the original image shows noise and some unidentified features. However, the ground truth image is segmented which separates Nucleus, Cytoplasm, and Background. Groundtruth serves as a reference for evaluating the segmentation model by giving each cell component a different color.

## 2.2 Image enhancement

Image enhancement is the first step in improving pap-smear image quality before segmentation [35]. In the study, the image enhancement stage includes quality enhancement, contrast improvement, and noise reduction. This stage addresses challenges such as low contrast and noisy images, making the images more suitable for accurate analysis.

### 2.2.1. Contrast limited adaptive histogram equalization

Contrast Limited Adaptive Histogram Equalization (CLAHE) is a method used to enhance contrast in images. CLAHE can distribute intensities and adjust the original shadow intensities of the image [36]. The brightness level of the image contrast is set using the clip limit.

### 2.2.2. Gamma correction

Gamma correction is a method for adjusting image brightness by using a gamma function. Gamma correction classifies pixel intensities into bright and dark areas.

### 2.2.3. Median filter

The median filter is a method used to remove noise from images. The median filter removes noise

by sorting pixels within a kernel and replacing the central pixel with the median value.

## 2.3 Evaluation image enhancement

In the study, the effectiveness of the image enhancement is evaluated using the MSE, PSNR, and SSIM. The MSE measures the average squared error between the original image and the image after enhancement. The lower MSE indicates a better method. MSE is calculated by Eq. (1) [37].

$$MSE = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n (a_{ij} - m_{ij})^2 \quad (1)$$

$m$  is image height.  $n$  is the image width.  $a_{ij}$  is the pixel value of the original image at  $(i,j)$ .  $m_{ij}$  is the pixel value after median filtering.

PSNR estimates image quality by comparing the processed image to the original. PSNR is calculated by Eq. (2) [37].

$$PSNR = 10 \log_{10} \left( \frac{s^2}{MSE} \right) \quad (2)$$

PSNR is the peak signal-to-noise ratio.  $s$  is the highest image range.

SSIM measures the similarity between two images. SSIM compares the original image to the distorted image using Eq. (3) [37].

$$SSIM(i, j) = \frac{(2\mu_i\mu_j) + (\sigma_{ij}^2)}{(\mu_i^2 + \mu_j^2)(\sigma_i^2 + \sigma_j^2)} \quad (3)$$

$\mu_i$  is the mean of the original image.  $\mu_j$  is the mean of the processed image.  $\sigma_i$  is the variance of the original image.  $\sigma_j$  is the variance of the processed image.  $\sigma_{ij}^2$  is the covariance between the original and processed images.

## 2.4 Architecture modification

The DDUSeg-Net architecture consists of two U-Net blocks combined into a single model. Each U-Net block is modified with pooling Indices in the max pooling operation and a dropout at the end of each encoder and decoder block. The DDUSegNet consists of various layers, including the convolutional layer, pooling layer, and softmax activation function. The convolution layer process can be calculated by Eq. (4).

$$t_{ij} = \sum_{u=0}^{m-1} \sum_{v=0}^{n-1} m_{u+is, v+js} \times k_{u,v} + b_q \quad (4)$$

$i = 0, 1, \dots, (n - 1)$  and  $j = 0, 1, \dots, (n - 1)$ .  $t_{ij}$  is the entry of the convolution result matrix of the  $i$ -th and  $j$ -th,  $m_{u+is, v+js}$  is the entry of the input matrix of the  $u + is$  row and  $v + js$  column.  $k_{u,v}$  is the entry of  $u$ -th kernel matrix and  $v$ -th column.  $n$  is kernel height and  $b_q$  is the bias on  $q$ -th filter with  $q = 0, 1, \dots, (n - 1)$ .

In the proposed method, the pooling layer uses **max-pooling** for downsampling and **pooling indices with unpooling** for upsampling to restore spatial information. The max pooling reduces the spatial dimensions by selecting the maximum value within each defined region. The max pooling process can be calculated by Eq. (5).

$$Y_{(p,q)} = \max (T_{i,j}) \tag{5}$$

$Y_{(i,j)}$  is the pooled value from the input feature map  $T_{i,j}$ . The max pooling operation captures important features while making the model more efficient. During the max-pooling process, **pooling indices** are recorded, which represent the positions of maximum values, thereby preserving spatial details [9].

In the upsampling stage, the unpooling employs the recorded pooling indices to accurately place the maximum values back in their original positions. This process ensures precise spatial reconstruction. The upsampling stage can be calculated by Eq. (6).

$$Y'_{(i,j)} = \begin{cases} Y_{(p,q)} & \text{if } (i,j) = \text{pooling index of } (p,q) \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

$(i,j)$  refers to the original position in features before the max pooling stage.

The softmax activation function is one of the non-linear activation functions that functions for multiclass classification. The results of the softmax activation function can be seen in Eq. (7).

$$a(t_{ij}) = \frac{e^{r(t_{ij})}}{\sum_{m=1}^n e^{s_m}} \tag{7}$$

$a(t_{ij})$  is the probability value for the  $i$ -th row and  $j$ -th column of the softmax function.  $r(t_{ij})$  is the input value of the ReLU activation function result. Categorical cross-entropy is a loss function used in semantic segmentation tasks. Categorical cross entropy can be calculated using Eq. (8).

$$L = - \sum_n^m y_i \cdot \log a(t_{ij}) \tag{8}$$

$y_i$  is the probability of the segmentation label at the  $i$ -th label.  $a(t_{ij})$  is the input value of the softmax function result.

### 2.5 Training and testing

In the training stage, the model is trained using preprocessed data. The training configuration includes setting the number of epochs to 30 epochs, the batch size to 32, and the optimizer using Adam with a learning rate of 0.0001. In the testing stage, the model uses the best weights obtained from the training process to evaluate its performance on the testing data. The prediction results are compared to the ground truth. This comparison obtained a confusion matrix. The confusion matrix is essential for evaluating model performance.

### 2.6 Evaluation model performance

The parameters used to assess the success of segmentation performance, metrics such as accuracy, sensitivity, specificity, f1-score, IoU, and G-Means are used. The accuracy measures the overall correctness of the model. The sensitivity evaluates the model's ability to identify true positives. The specificity evaluates the model's ability to identify true negatives. The IoU assesses the accuracy of the intersection area between predicted and ground truth. F1-score evaluates the balance between sensitivity and specificity when the segmented image closely matches the ground truth. The performance of the proposed method will be compared with previous research results.

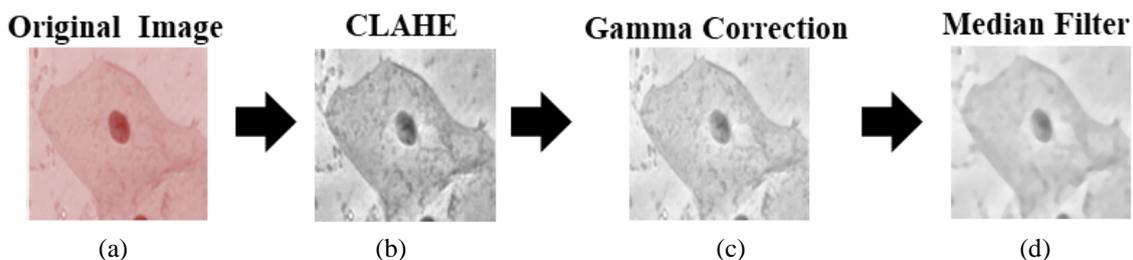


Figure. 3 Example of the results image enhancement stage on pap-smear images (a)Original Image (b)CLAHE (c)Gamma Correction ( $\gamma = 0.8$ ) (d)Median Filter ( $k = 4$ )

Table 1. PSNR and SSIM Results of Image Enhancement

No	File Name	MSE	PSNR	SSIM
1	148...4986-001.BMP	126.3	27.1	0.92
2	148...5504-001.BMP	54.2	30.8	0.88
3	148...5585-001.BMP	115.1	27.5	0.89
⋮	⋮	⋮	⋮	⋮
5502	2095...6517-001.BMP	98.4	27.2	0.86
Average		101.1	28.2	0.81

Table 2. Comparison of PSNR and SSIM Results with Other Studies

No	Image Enhancement	MSE	PSNR	SSIM
1	Gamma Correction, Weight Distribution, and CLAHE [14]	-	12	0.5
2	CLAHE and Gaussian Filter [38]	-	20.01	<b>0.90</b>
3	CLAHE [39]	34.50	21.63	-
4	<b>Proposed Method</b>	<b>101.1</b>	<b>28.2</b>	0.81

### 3. Result and discussion

#### 3.1 Image enhancement stage

The pap-smear image used in this study was obtained from the Herlev dataset with a dimension size of  $150 \times 140$  pixels. The pap-smear image was image enhanced using the gamma correction and median filter. The image enhancement stage was carried out 6 times, consisting of 3 times for gamma correction and 3 times for the median filter. An example of the results of the image enhancement stage on pap-smear images can be seen in Figure 3.

Based on Figure 3, the image enhancement results obtained 5502 from 6 times image enhancement where the original image was 917 images. The gamma correction used gamma ( $\gamma$ ) of 0.6, 0.8, and

1.2. The median filter used kernel ( $k$ ) of 3, 5, and 7. Figure 3 illustrates that the original image used CLAHE for contrast improvement, followed by gamma correction for brightness adjustment and median filtering to reduce noise. The effectiveness of the image enhancement stage was evaluated using MSE, PSNR, and SSIM. The PSNR and SSIM results for the 5502 images are summarized in Table 1.

Based on Table 1, the average MSE is 101.1 which shows the difference in image pixels resulting from image quality enhancement is relatively small compared to the original image. The average PSNR value obtained is 28.2 which shows that image quality enhancement is good enough, but there is still noise. The average SSIM obtained 0.92, indicating that the quality improvement result has a structure similar to the original image. The effectiveness of the proposed method in image enhancement is compared with the image enhancement result in other studies which can be seen in Table 2.

Based on Table 2 shows the comparison of image enhancement results with other studies. [14] used the Gamma Correction, Weight Distribution, and CLAHE methods to obtain a PSNR of 13dB and SSIM of 0.6, indicating a significant difference between the enhanced image and the original image. Yang *et al* [38] used CLAHE and Gaussian Filter methods to obtain a PSNR of 20.01 and SSIM of 0.90. The result of [38] indicated good quality improvement in the enhanced images compared to the original images. Nahrawi *et al* [39] obtained an MSE of 34.50 and a PSNR of 21.63, which were better than [14] but still below the proposed method. The proposed method obtains an excellent MSE of 101.1 and PSNR of 28.2. However, SSIM in the proposed method was not as optimal as [38]. The difference was not significant and the higher MSE indicated excellent improvement in image quality.

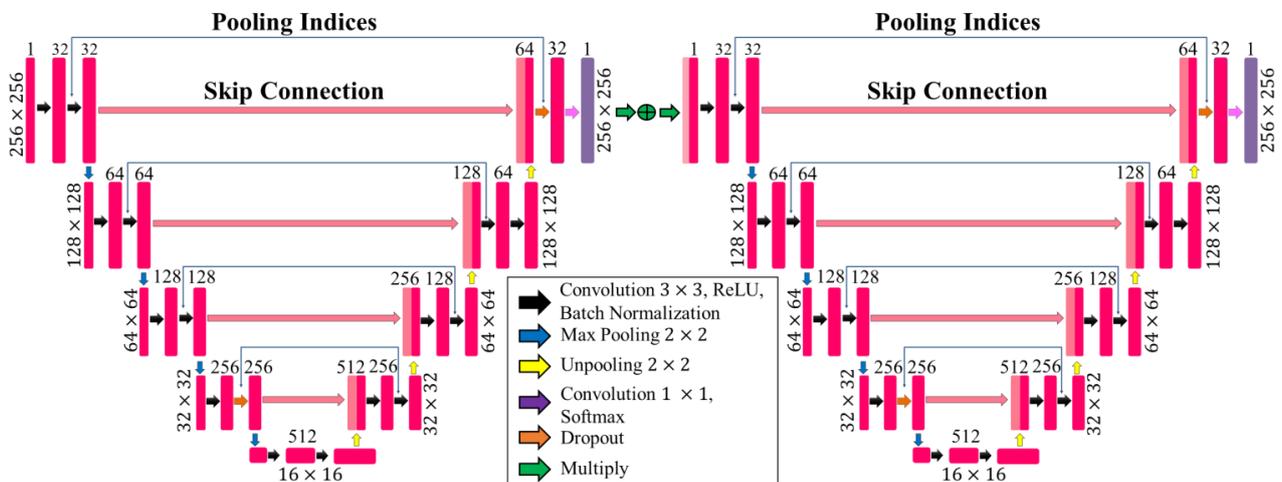


Figure. 4 Illustration of DDUSeg-Net architecture for cervical cancer segmentation

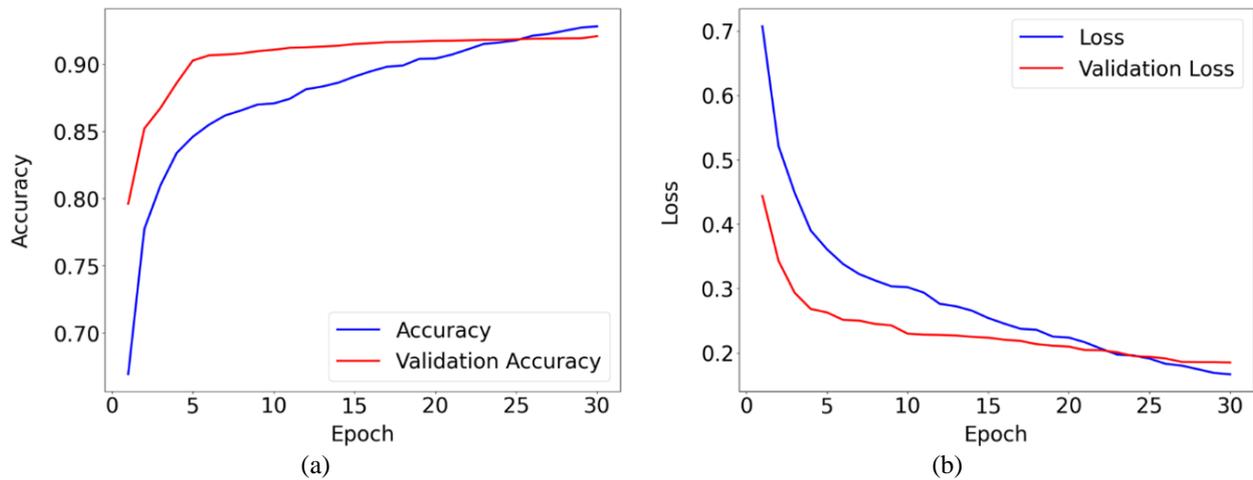


Figure. 5 Training results of the DDUSeg-Net architecture for cervical cancer segmentation (a)accuracy (b)loss

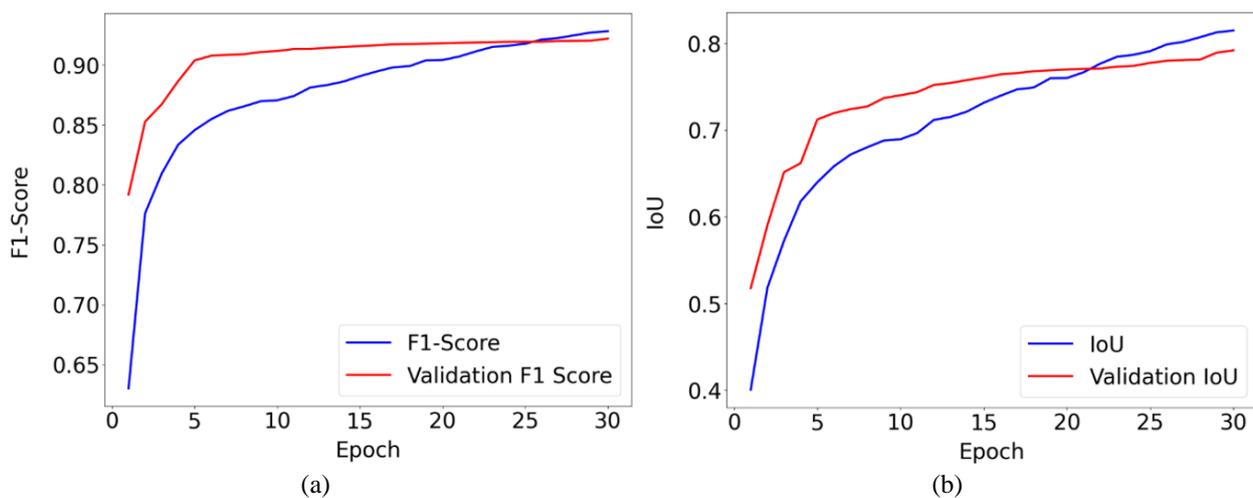


Figure. 6 Training results of the DDUSeg-Net architecture for cervical cancer segmentation (a)f1-score (b)IoU

### 3.2 DDUSeg-Net architecture

The DDUSeg-Net is an architecture consisting of two U-Net blocks combined into a single model. Each U-Net in DDUSeg-Net has been modified with pooling indices on the max pooling operation and adding dropout at the end of each encoder and decoder block. The architecture of DDUSeg-Net is shown in Figure 4.

Based on Figure 4, The DDUSeg-Net consists of two U-Net blocks, each with an encoder, bridge, and decoder. The encoder has 4 layers with convolution, ReLU activation, batch normalization, and max pooling with Pooling Indices. The first block uses a  $3 \times 3$  kernel with 32 filters, and subsequent blocks use 64, 128, and 256 filters, with Dropout before pooling in the fourth block. The bridge features two convolution layers, ReLU activation, and batch normalization. The decoder has four blocks, each including unpooling with indices, two convolution

layers, ReLU activation, and batch normalization, with 256, 128, and 64 filters in the first three blocks, and Dropout plus  $1 \times 1$  convolution layers in the final block, ending with a softmax activation function.

### 3.3 Training stage

In this study, the data was split into 80% for training and 20% for testing. From the training set, 20% was set aside for validation to monitor learning. The best weights were saved for testing. The accuracy and loss were measured during training and validation, as shown in Figure 5.

Based on Figure 5 (a) it can be seen that the accuracy of the graph on the training data (blue line) and data validation (red line) continues to increase and is stable until the last epoch. In the first epoch, the accuracy of the training data is 66.9% which continues to increase gradually to 92.8%. For the data validation in the first epoch, the accuracy value is 79.6% and continues to increase gradually to 92%. In

Figure 5 (b) the loss graph continues to decrease until the last epoch. In the first epoch, the loss value on the training data (blue line) is obtained at 0.7072 then the loss value decreases until the last epoch of 0.1669. In data validation (red line) the loss value also decreases and is stable until the last epoch. In data validation, the loss value in the first epoch is obtained at 0.4438 and decreases until the last epoch of 0.1852. The increase in accuracy value and decrease in loss value at the training stage indicate that the model has good performance in learning and recognizing features in both training data and data validation. In addition to the accuracy and loss, the f1-score and IoU are also measured at the training stage. The results of the f1-score and IoU graphs are shown in Figure 6.

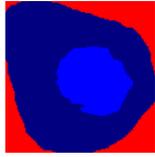
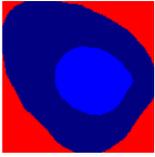
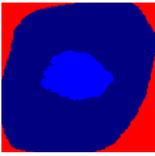
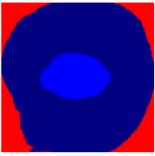
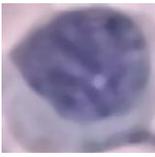
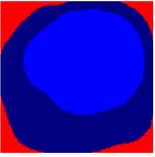
Based on Figure 6 (a) shows the F1-Score value on the training data (blue line) and validation data (red line) shows a significant and stable increase until the last epoch. The F1-Score value of the training data in the last epoch obtained of 92.8% and the validation data was 92.2%. The F1-score graph shows that the model training process has a very good level of precision and recall. In Figure 6 (b) the IoU on the training data (blue color) and validation data (red color) also show an increase from each epoch. The IoU on the last epoch of the training data was 81.5% and the validation data was 79.2%. The IoU obtained on the validation data is still below 80%, indicating that it is not optimal for detecting the cytoplasm, nucleus, and background.

### 3.4 Testing stage

In the testing stage, testing was carried out using testing data to see the success rate of the DDUSeg-Net architecture in segmenting cervical cancer in pap smear images. At this stage, the best weights obtained from the training results were used for segmenting cervical cancer in the testing data. A comparison of the results of the testing stage and ground truth can be seen in Table 3.

Based on Table 3 shows the results of segmentation and ground truth at the segmentation stage using the DDUSeg-Net architecture on pap smear images. At the segmentation stage, the prediction results produced are similar to the ground truth. The segmentation stage using the DDUSeg-Net architecture can recognize features in the image. The model performance evaluation value is obtained by comparing the segmentation results at the testing stage with the ground truth using the confusion matrix table. Performance evaluation is measured based on the values of accuracy, sensitivity, specificity, F1-Score, and IoU. The test results on the testing data can be seen in Figure 7.

Table 3. Comparison of segmentation results and ground truth on the Herlev dataset

No	Original Image	Ground Truth	Prediction
1			
2			
3			

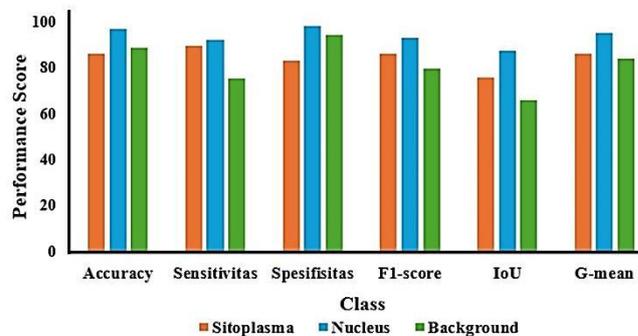


Figure. 7 The comparison of DDUSegNet performance for cervical cancer segmentation

Based on Figure 7 shows the evaluation value based on the accuracy, sensitivity, specificity, F1-Score, IoU, and G-mean for each label. The nucleus label is excellent in all five evaluations compared to other labels. The values obtained for the nucleus label are accuracy of 97.08%, sensitivity of 92.24%, specificity of 98.46%, F1-Score, 93.34%, Iou 87.52%, and G-mean 95.30%. The cytoplasm label obtained good accuracy, sensitivity, specificity, F1-Score, and G-mean of 86.27%, 89.64%, 83.06%, 88.43%, 86.29%, while the IoU value obtained a sufficient value of 76.11%. In the Background label, very good accuracy and specificity were obtained, namely 90.71% and 91.93%, the sensitivity of F1-Score, and G-Mean obtained good values, namely 85.75%, 86.46%, and 88.79%, while IoU obtained a sufficient value of 76.59%. Another model performance evaluation was measured the Receiver Operating Characteristic (ROC) and Area Under the Curve (AUC). The ROC and AUC in cervical cancer segmentation can be seen in Figure 8.

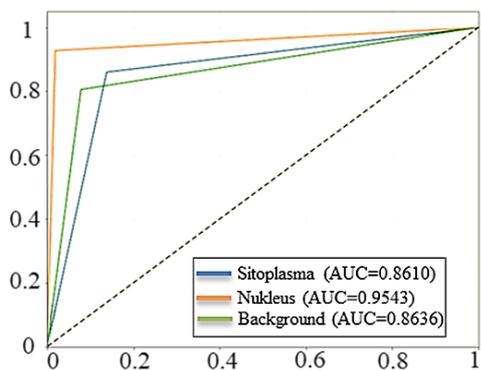


Figure. 8 The ROC graph on cervical cancer segmentation using DDUseg-Net architecture

Based on Figure 8, the highest AUC value is obtained on the nucleus label, which is above 0.9, which means that the model performance is very good in distinguishing the nucleus label pixels from other label pixels. On the cytoplasm and background labels, the AUC value is good, which is above 0.85. This shows that the model performance results are good in distinguishing the cytoplasm and background pixels from other label pixels.

### 3.5 Analysis and discussion

This study combines the image enhancement stages (median filter, CLAHE, and gamma correction) with the segmentation stage using the DDUseg-Net architecture. To determine the level of success of the proposed method, a comparison of performance results with several other studies is carried out as shown in Table 4.

Based on Table 4, the comparison of other studies using the same dataset on pap smear images. [40] and [41] did not apply image enhancement stages.

Kurnianingsih *et al* [41] applied Mask-RCNN for cervical cancer segmentation with 3 classes and obtained a sensitivity of 94%, but a specificity below 90%. *Desiani et al* [9] applied U-Net with normalization, CLAHE, and gamma correction obtained accuracy, sensitivity, and specificity below 80%. Although the sensitivity value of the proposed method is lower than [40]. The sensitivity of the proposed method is excellent, which is above 85%. The sensitivity above 85% indicated predicts the nucleus and cytoplasm. The accuracy of the proposed method is excellent compared to other studies, which is above 90%. The accuracy above 90% indicates the DDUsegNet demonstrates robust performance in correctly predicting the images. The specificity of the proposed method is above 90%. The specificity above 90% indicated that the DDUseg-Net model is excellent for predicting background. The comparison of f1-Score, IoU, and G-mean with other studies can be seen in Table 5. Based on Table 5, the comparison of other studies results using the same dataset on pap smear images. The f1-Score of the proposed method is more excellent than other studies, which is above 85%. The f1-score above 85% indicated that the DDUseg-Net has a balanced precision and recall. Sabeena *et al* [40] applied PA-Trans architecture and obtained IoU better than the proposed method. The IoU of the proposed method is 76%, indicating that the DDUseg-Net has a good enough overlap between the prediction results and the ground truth. The G-Mean of the proposed method was obtained above 85%, indicating that the segmentation performance has a robust prediction between the nucleus, cytoplasm, and background. The results in Table 4 and Table 5 show that the DDUseg-Net is excellent and robust for cervical cancer segmentation in pap smear images.

Table 4. Performance comparison of DDUseg-Net architecture for cervical cancer segmentation

Segmentation Method	Image Enhancement Method	Acc (%)	Sen (%)	Spe (%)
FC DenseNet56 [40]	None	-	<b>87.67</b>	-
Mask-RCNN [41]	None	89,8	72,5	<b>94,3</b>
U-Net [9]	Normalization, CLAHE, Gamma Correction	77	72	71
<b>DDUseg-Net (Proposed)</b>	<b>Median Filter, CLAHE, Gamma Correction</b>	<b>90,71</b>	85,75	91,93

Table 5. Comparison of F1-Score, IoU, and G-Mean with Other Studies

Segmentation Method	Image Enhancement Method	F1-Score (%)	IoU (%)	G-Mean (%)
FC DenseNet56 [40]	None	-	<b>87.65</b>	-
PATrans [42]	None	-	89.7	-
U-Net [9]	Normalization, CLAHE, Gamma Correction	69	-	-
<b>DDUseg-Net (Proposed)</b>	<b>Median Filter, CLAHE, Gamma Correction</b>	<b>86.46</b>	76.59	<b>88.79</b>

#### 4. Conclusion

The proposed method combined cervical cancer segmentation with image enhancement methods. The image enhancement stage used CLAHE, Gamma Correction, and Median Filter. The MSE of the proposed method is above 100, indicating a relatively small difference in the enhanced image compared to the original image. The PSNR of the proposed method is above 25. The PSNR above 25 shows that image quality improvement has good enough results, but there is still noise. The SSIM of the proposed method indicates that the enhanced image retains the structure and details similar to the original image. In the segmentation stage, the DDUSeg-Net model demonstrates excellent and robust performance in cervical cancer segmentation with all metrics approaching 90%. The accuracy, sensitivity, specificity, F1-Score, and G-Mean of the proposed method show effectiveness in segmenting the relevant features in pap smear images. In the future, the proposed method should focus on improving PSNR in image quality and increasing IoU in segmentation.

#### Conflicts of Interest

The authors declare no conflict of interest.

#### Author Contributions

We certify that all authors contributed to this study. Paper conceptualization and methodology, Rudiansyah, Anita Desiani, Dian Palupi Rini, and Lucky Indra Kesuma; software and validation; formal analysis and investigation; resources; data curation and writing-original draft preparation; writing-review editing and visualization; supervision; and project administration.

#### Acknowledgments

The study/publication of this article was funded by DIPA of the Public Service Agency of Universitas Sriwijaya 2024. Decree Number 0459/E5/PG.02.00/2024, June on 2024. Under Contract Agreement Number 090/E5/PG.02.00.PL/2024, June on 2024.

#### References

- [1] N. Sompawong *et al.*, “Automated Pap Smear Cervical Cancer Screening Using Deep Learning”, *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2019.
- [2] V. Karra, P. S. Ramulu, Rajanikar, and B. Sudhakar, “Cervical pap smear study and its utility in cervical cancer detection and prevention”, *Indian J. Obstet. Gynecol. Res.*, Vol. 8, 2021.
- [3] W. A. Mustafa, N. A. Alias, M. A. Jamlos, S. Ismail, and H. Alquran, “A Recent Systematic Review of Cervical Cancer Diagnosis: Detection and Classification”, *J. Adv. Res. Appl. Sci. Eng. Technol.*, Vol. 1, 2021.
- [4] Y.-M. Lee, B. Lee, N.-H. Cho, and J. H. Park, “Beyond the Microscope: A Technological Overture for Cervical Cancer Detection”, *diagnostics*, Vol. 13, 2023.
- [5] A. Botalb, “Contrasting Convolutional Neural Network (CNN) with Multi-Layer Perceptron (MLP) for Big Data Analysis”, In: *Proc. of International Conference on Intelligent and Advanced System (ICIAS)*, Kuala Lumpur, Malaysia: IEEE, 2018.
- [6] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, “Convolutional neural networks: an overview and application in radiology”, *Insights Imaging*, Vol. 9, 2018.
- [7] A. Desiani, Erwin, B. Suprihatin, F. Efriliyanti, M. Arhami, and E. Setyaningsih, “VG-DropNet a Robust Architecture for Blood Vessels Segmentation on Retinal Image”, *IEEE Access*, Vol. 10, 2022.
- [8] I. A. Kazerouni, G. Dooly, and D. Toal, “Ghost-UNet: An Asymmetric Encoder-Decoder Architecture for Semantic Segmentation From Scratch”, *IEEE Access*, Vol. 9, 2021.
- [9] A. Desiani, M. Erwin, B. Suprihatin, S. Yahdin, A. I. Putri, and F. R. Husein, “Bi-path Architecture of CNN Segmentation and Classification Method for Cervical Cancer Disorders Based on Pap-smear Images”, *Int. J. Comput. Sci.*, Vol. 48, 2021.
- [10] A. W. Arum, S. Nurmaini, D. P. Rini, P. Agustiansyah, and M. N. Rachmatullah, “Segmentation of Squamous Columnar Junction on VIA Images using U-Net Architecture”, *Comput. Eng. Appl. J.*, Vol. 10, 2021.
- [11] N. Nazir, A. Sarwar, B. S. Saini, and R. Shams, “A Robust Deep Learning Approach for Accurate Segmentation of Cytoplasm and Nucleus in Noisy Pap Smear Images”, *Computation*, Vol. 11, 2023.
- [12] Y. Song, E. Tan, X. Jiang, S. Member, J. Cheng, and D. Ni, “Accurate Cervical Cell Segmentation from Overlapping Clumps in Pap Smear Images”, *Transaction Med. Imaging*, Vol. 36, 2017.
- [13] M. K. Ghoben and L. A. N. Muhammed, “Exploring the Impact of Image Quality on

- Convolutional Neural Networks: A Study on Noise, Blur, and Contrast”, In: *Proc. of International Conference of Computer Science and Information Technology (ICOSNIKOM)*, Binjai, Indonesia, 2023.
- [14] K. Sangeeta and S. Kisan, “Brightness Preserving Medical Image Contrast Enhancement Using Entropy-Based Thresholding and Adaptive Gamma Correction with Weighted Distribution”, *Springer*, 2023.
- [15] S. Ramesh, S. Sasikala, and N. Paramanandham, “Segmentation and classification of brain tumors using modified median noise filter and deep learning approaches”, *Multimed. Tools Appl.*, Vol. 80, 2021.
- [16] T. S. S. Shiney, R. J. Rose, T. S. S. Shiney, and R. J. Rose, “Deep Auto Encoder Based Extreme Learning System for Automatic Segmentation of Cervical Cells”, *IETE J. Res.*, Vol. 69, 2023.
- [17] X. Huang, J. Chen, M. Chen, L. Chen, and Y. Wan, “TDD-UNet: Transformer with double decoder UNet for COVID-19 lesions segmentation”, *Comput. Biol. Med.*, Vol. 151, 2022.
- [18] D. D. Pham, G. Dovletov, S. Warwas, S. Landgraeber, M. Jäger, and J. Pauli, “Deep segmentation refinement with result-dependent learning: A double U-net for hip joint segmentation in MRI”, *Springer Vieweg*, 2019.
- [19] Y. Cao, A. Vassantachart, and J. . Ye, “Automatic Detection and Segmentation of Multiple Brain Metastases on MR Images Using Simultaneous Optimized Double-UNET Architecture”, *Int. J. Radiat. Oncol. Phys.*, Vol. 108, 2020.
- [20] D. Jha, M. A. Riegler, and D. Johansen, “DoubleU-Net: A Deep Convolutional Neural Network for Medical Image Segmentation”, *2020 IEEE 33rd Int. Symp. Comput. Med. Syst.*, No. 1, 2020.
- [21] R. Chen, X. Pei, J. Ma, and Q. Liu, “Multi-Attention Base Double Unet for Optic Cup and Disc Segmentation”, 2022.
- [22] N. Thompson, K. Lee, K. Greenewald, and G. F. Manso, “The Computational Limits of Deep Learning”, *LIMITS '21: Workshop on Computing within Limits*, York, 2021.
- [23] Y. Kwon, S. Bae, D. Chung, and M. Lim, “Semantic Segmentation by Using Down-Sampling and Subpixel Convolution: DSSC-UNet”, *Comput. Mater. Contin.*, Vol. 75, 2023.
- [24] X. Hu, J. Bian, W. Liu, and J. Pei, “Measuring Model Complexity of Neural Networks with Curve Activation Functions”, In: *Proc. of the 26th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, New York, 2020.
- [25] C. Zhang, W. Lu, J. Wu, C. Ni, and H. Wang, “SegNet Network Architecture for Deep Learning Image Segmentation and Its Integrated Applications and Prospects”, *Acad. J. Sci. Technol.*, Vol. 9, No. 2, 2024.
- [26] S. Montaha *et al.*, “BreastNet18: A High Accuracy Fine-Tuned VGG16 Model Evaluated Using Ablation Study for Diagnosing Breast Cancer from Enhanced Mammography Images”, *Biology (Basel)*, Vol. 10, 2021.
- [27] S. Almotairi, G. Kareem, M. Aouf, B. Almutairi, and M. A. M. Salem, “Liver tumor segmentation in CT scans using modified segnet”, *Sensors*, Vol. 20, 2020.
- [28] J. B. A. Teixeira *et al.*, “Segmentation of cervical nuclei using convolutional neural network for conventional cytology”, *Comput. Methods Biomech. Biomed. Eng. Imaging Vis.*, Vol. 11, 2023.
- [29] H. Yu *et al.*, “Segmentation of the cervical lesion region in colposcopic images based on deep learning”, *Front. Oncol.*, Vol. 12, 2022.
- [30] L. Chen, P. Bentley, K. Mori, K. Misawa, M. Fujiwara, and D. Rueckert, “DRINet for Medical Image Segmentation”, *IEEE Trans. Med. Imaging*, Vol. 37, No. 11, 2018.
- [31] I. Salehin and D.-K. Kang, “A Review on Dropout Regularization Approaches for Deep Neural Networks within the Scholarly Domain”, *Electronics*, Vol. 12, No. 14, 2023.
- [32] T. A. Soomro *et al.*, “Deep Learning Models for Retinal Blood Vessels Segmentation: A Review”, *IEEE Access*, Vol. 7, 2019.
- [33] Y. Li *et al.*, “A Survey on Dropout Methods and Experimental Verification in Recommendation”, *IEEE Trans. Knowl. Data Eng.*, Vol. 35, No. 7, 2023.
- [34] A. Rasheed, S. H. Shirazi, A. I. Umar, M. Shahzad, W. Yousaf, and Z. Khan, “Cervical cell’s nucleus segmentation through an improved UNet architecture”, *PLoS One*, Vol. 18, Oct. 2023.
- [35] A. Desiani *et al.*, “Denoised Non-Local Means with BDDU-Net Architecture for Robust Retinal Blood Vessel Segmentation”, *Int. J. Pattern Recognit. Artif. Intell.*, Vol. 37, 2023.
- [36] A. desiani, E. Erwin, B. Suprihatin, Y. Wahyudi, E. S. Cahyono, and M. Arhami, “BVU-Net: A U-Net Modification by VGG-Batch Normalization for Retinal Blood Vessel Segmentation”, *Int. J. Intell. Eng. Syst.*, Vol. 15, Dec. 2022.
- [37] E. Erwin, “Improving Retinal Image Quality

- Using the Contrast Stretching, Histogram Equalization, and CLAHE Methods with Median Filters”, *Int. J. Image, Graph. Signal Process.*, Vol. 12, Apr. 2020.
- [38] B. Yang, H. Zhao, L. Cao, H. Liu, N. Wang, and H. Li, “Retinal image enhancement with artifact reduction and structure retention”, *Pattern Recognit.*, Vol. 133, 2023.
- [39] N. Nahrawi, W. A. Mustafa, S. Nurul, and A. Mohd, “Color Contrast Enhancement on Pap Smear Images Using Statistical Analysis”, *Intell. Autom. Soft Comput.*, Vol. 30, 2021.
- [40] K. Sabeena, C. A. Gopakumar, R. Thampi, and A. S. P. G. Scholar, “An Improved Deep Convolutional Model for Segmentation of Nucleus and Cytoplasm from Pap Stained Cell Images”, *2020 6th Int. Conf. Adv. Comput. Commun. Syst.*, 2020.
- [41] Kurnianingsih, L. E. Nugroho, A. S. Prabuwo, and T. Mantroro, “Segmentation and Classification of Cervical Cells Using Deep Learning”, *IEEE Access*, Vol. 7, 2019.
- [42] H. Hu, J. Zhang, T. Yang, Q. Hu, Y. Yu, and Q. Huang, “PATrans : Pixel-Adaptive Transformer for edge segmentation of cervical nuclei on small-scale datasets”, *Comput. Biol. Med.*, Vol. 168, 2024.