RESEARCH ARTICLE

How to cite: Anita Desiani, Irmeilyana, Des Alwine Zayanti, Yadi Utama, Muhammad Arhami, Azhar K Affandi, Muhammad Aditya Sasongko, and Indri Ramayanti, "Data Augmentation and U-Net CNN for Accurate Nuclei Segmentation on Pap Smear Images", Journal of Electronics, Electromedical Engineering, and Medical Informatics, vol. 6, no. 3, pp. 264-275, July 2024.

Data Augmentation and U-Net CNN for Accurate Nuclei Segmentation on Pap Smear Images

Anita Desiani[1] , Irmeilyana[1] , Des Alwine Zayanti[1] , Yadi Utama[2] , Muhammad Arhami[3] , Azhar K Affandi[4] , Muhammad Aditya Sasongko[5] , and Indri Ramayanti[6]

[1] Department of Mathematics, Universitas Sriwijaya, Indralaya, Indonesia
[2] Informatics Department, Universitas Sriwijaya, Indaralaya, Indonesia
[3] Informatics Department, Politeknik Negeri Lhoseumawe, Aceh, Indonesia
[4] Informatics Department of Physics, Universitas Sriwijaya, Indralaya, Indonesia
[5] Department of Computer Engineering, Koc University, Istanbul, Turkey
[6] Department of Parasitology, Faculty of Medicine, Universitas Muhammadiyah, Palembang, Indonesia

Corresponding author: Anita Desiani (e-mail: anita_desiani@unsri.ac.id)

ABSTRACT The nuclei and cytoplasm can be detected through Pap smear images. To help women avoid cervical cancer, early detection of nuclei abnormalities can be known through image segmentation. U-Net CNN is known as an architecture commonly used for segmentation. U-Net CNN needs a large amount of data for training. The amount of pap smear images is limited. A small amount of data can cause overfitting and reduce the performance of U-Net. This study combines augmentation and segmentation, the augmentation in this study combines geometric transformation with flipping and rotation, and pixel-wise transform with Gamma Correction. The augmentation aims to generate new images that are more numerous and varied. The result of the proposed method is a pap smear image which only consists of two parts, the background and nuclei as the foreground. The performance evaluation of the combination of the augmentation method and U-Net CNN is accuracy, sensitivity, specificity, and F1 score. The application of augmentation using Flipping, rotation, and Gamma Correction can increase 10 times the amount of Pap Smear image data. The average results of accuracy, sensitivity, specificity, and F1 score of U-Net on augmented data are 93.75%. These results show that the combination of augmentation and U-Net CNN is excellent and robust at detecting nuclei in pap smear images. The proposed method can be developed to segment other cells such as cytoplasm and can be as a component for building an automatic cervical cancer detection.

INDEX TERMS Augmentation, Cervical Cancer, Images, Nuclei, Pap Smear, Segmentation

I. INTRODUCTION

Cervical cancer is the fourth most common type of cancer in women and is also one the most lethal. Cervical cancer is the second most common type of cancer in Indonesia, after breast cancer. In Indonesia, the incidence of new cases of cervical cancer in women is approximately 36,663 cases, with a mortality rate of 21,003 cases[1]. Early Cervical cancer detection can help prevent or delay the progression of cervical abnormalities to invasive cancer[2]. Cervical cancer can be detected earlier through Pap smear image screening. The Pap smear image is an image of the result of an examination used to detect cervical cancer in women. However, cancer screening on Pap smear images is hampered significantly by insufficient cell staining and overlapping cell images [3]. Manual segmentation of pap smear images is usually carried out by an expert. However, expert segmentation is not very

precise, is prone to differences in observations between experts, takes time due to various series of examinations, and is prone to errors due to human negligence such as fatigue[4]. It requires image segmentation software to assist medical practitioners in overcoming these obstacles. Segmentation is a technique for separating the object area from the background area of an image so that the object can be easily analyzed. Segmentation in cervical cancer is accomplished through the analysis of cervical cell datasets[5]. According to [6], nuclei segmentation outperforms cytoplasmic segmentation. The nuclei are the most critical feature in Pap smear image segmentation [7].

Numerous methods have been extensively used to segment Pap smear images, including [8], who used the Global and Local Graph Cuts method with sensitivity and F1-score values of 87% and 89%, respectively. However, this study did not

account for the performance of the other. [9] obtained an accuracy of 89.6%, specificity of 89.41%, sensitivity of 93%, and F1-scores of 72.1%, respectively, using the ENN-TLBO method. The sensitivity value in this study was still less than 70%, at 67%. [10] determined a sensitivity of 89.13% and a specificity of 89.31% using the Gaussian Mixture method. Other performance measures were not calculated in this study. [11][11] compared three methods for segmenting Pap smear images: global thresholding, Fuzzy C-means, and watershed segmentation. Their results indicated that the average accuracy was quite good at 79.3% and the average specificity was 82.3%. However, the average sensitivity value was less than 70%, at 65.3%. At the moment, deep learning methods have demonstrated superior performance in a variety of pattern recognition and imaging applications.[12] Convolutional Neural Network (CNN) is a widely used method for image segmentation and image classification in deep learning[13]. According to [14][14] CNN method in deep learning in biomedical images, segmentation had problems that are influenced by kernel size, if the kernel size was getting smaller, then the accuracy value was getting better. However, this will have an impact on the completion of such a long segmentation. In this study, the U-Net CNN method was able to work at the pixel level by utilizing the ReLU activation function to improve accuracy value and shorter completion time. The advantages of using U-Net architecture include U-Net being developed for segmenting medical images and extracting features that exist in the image. U-Net CNN produces excellent results in medical image segmentation because it trains on a large number of images and the objects it produces are more defined[15]. Otherwise, the overfitting can happen during training due to the small amount of data.[16].

Due to the large data requirement of U-Net CNN, it needs a technique to increase the amount of data. One of the techniques that can be applied is the augmentation technique. Augmentation is a technique to increase the number of data by applying some transformations to the existing data[17], [18]. The various transformations of augmentation that can be applied are geometric and pixel-wise transformations [19]. Flipping and rotation is a geometric transformation that can be applied to change the position of an image in several ways such as vertical and horizontal flipping or rotation by certain degrees. Flipping and rotation are easy to apply to get new images. [20] applied horizontal vertical flipping augmentation and rotation left and right to cervical cancer segmentation using the Inception Resnet architecture. Unfortunately, the curation and sensitivity obtained were only 81%. [21] applied rotation augmentation for nuclei segmentation with the R-CNN architecture but the sensitivity was only 72.5%. The use of augmentation with geometric transformation methods such as flipping and rotation is easy to do but has several disadvantages. Flipping has the limitation of changing geometry only based on horizontal and vertical positions. Rotation has a weakness, namely that if done carelessly it can cause the inclusion of other features that are not needed in the segmentation or classification process [22]. Another augmentation method is pixel-wise. Pixel-wise is a method of generating image variations by changing the pixel intensity,

both contrast and noise, in the original image. One method is Gamma Correction [19]. Gamma correction is the process of adjusting and controlling image contrast brightness to correct images that are too dark or too bright. [23] applied Gamma Correction to increase data on Respiratory classification and was able to increase sensitivity by 11.8%. [22] combines rotation and gamma correction for data augmentation in soil classification and can provide excellent average accuracy, sensitivity, and specificity performance of 94%.

Detection of cervical cancer on Pap smear images can be done by detecting abnormalities in the cytoplasm or nuclei abnormalities.[24]. U-Net CNN has been used for semantic segmentation of the nuclei and cytoplasm on Pap smear images, but unfortunately, the accuracy and sensitivity results obtained were still below 78%. [25]. [26] also conducted a study using the Multi-scale U-Net method of semantic segmentation of the nuclei and cytoplasm on Pap smear images and obtained sensitivity and F1-score below 90%, unfortunately, this study did not calculate accuracy and specificity. [27] combined U-Net with flipping, rotation, and Gaussian blur augmentation techniques for nuclei and cytoplasm segmentation but the sensitivity was only 89%. [28] applies U-Net CNN to segment nuclei and cytoplasm in pap smear images with 98% accuracy but it does not measure other performance. The nuclei are the most important element of the Pap smear image cell nuclei[29]. The presence of cytoplasm region was difficult to detect causing the performance of the learning model to decrease[30]. Previous studies have tried to perform semantic segmentation to get the cytoplasm and nuclei on Pap smear images but previous studies do not involve an augmentation process in the preprocessing.

This study proposes a nuclei segmentation method in pap smear images by combining augmentation and U-Net CNN. Augmentation techniques are applied to overcome the problem of limited data for U-Net because U-Net has a deep network that requires a lot of data in the training process. The augmentation techniques used are flipping, rotation, and Gamma Correction so the resulting images are more varied. To overcome this problem, in flipping vertical and horizontal flipping is used, and in rotation, random degrees are limited to 0^0-20^0. This degree is limited so the new image does not lose its original characteristics and also avoids introducing new features to the image. Another augmentation method is Gamma correction. Gamma correction is applied to obtain new images that are more varied by changing the brightness intensity of the image due to the limited augmentation results from flipping and rotation. This study focuses on nuclei segmentation in pap smear images so that the segmented image will consist of two features, namely foreground and background. The foreground is the nuclei and the background is features that are not included in the observation or features other than the nuclei. Segmentation is carried out by applying U-Net CNN after the augmentation technique is applied. To measure the success of the proposed method, this study compares the performance results of accuracy, sensitivity, specificity, and F1 score on segmentation results using U-Net CNN from a dataset that has not been augmented with segmentation performance results from a dataset that has involved augmentation.

II.   MATERIALS AND METHODS
A. SOURCE DATA

This study uses available secondary data, collected and acquired by the University of Campinas, Brazil. This data set was also approved by CEP, Human Research Ethics Committee, Brazil. The Zenoda data were collected on six women selected on condition that they were not pre-menopausal, not pregnant, and aged between 25 and 40 years. This dataset provides 184 images of cervical cells from six non-premenopausal, non-pregnant women aged between 25 and 40 years. Each image is 80 pixels wide and 80 pixels high. The following example of a cervical cell nucleus image from the dataset can be seen in FIGURE 1. The number of images available in this dataset is very limited, only 184 images, so it needs to be augmented to be able to use it for segmentation using the U-Net CNN architecture.
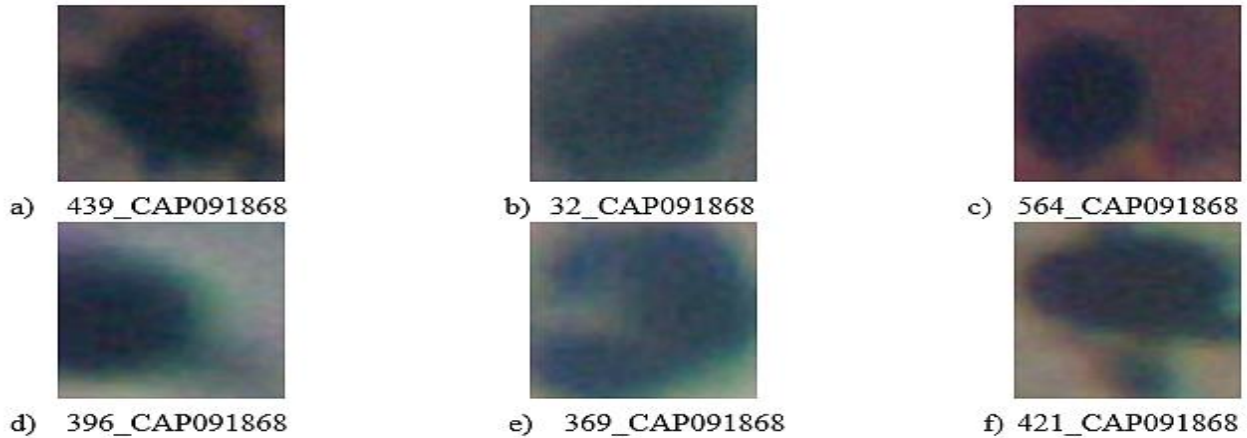

FIGURE 1. Six od Sample Original Image Cervical Cancer Nuclei on Pap Smear Image of Zenodo Dataset

B. AUGMENTATION

In this study, it is to meet the training data needs of the U-Net CNN, data augmentation was carried out on the input image by transforming the image using flipping, rotation, and Gamma Correction. The flipping techniques used are horizontal and vertical flips. The rotation technique is carried out by selecting degrees randomly between 0-20 degrees. Limiting the degree of rotation is to prevent the introduction of unnecessary new features in the image. Gamma correction is carried out by taking a random gamma value for each enlarged image to produce a more diverse image. Rotation and Gamma Correction were carried out 4 times for each image. The total data before augmentation was 184 images, but after augmentation, the amount of data was 1840 pap smear images. Equation (1) can be used to calculate the resulting image value after applying gamma correction [31].

$$G_{i,j} = 255 \left(\frac{h_{ij}}{255} \right)^{\gamma} \tag{1}$$

where $G_{i,j}$ is the result of the gamma correction image at the $(i,j)^{th}$ pixel, γ is the gamma correction value, and $h_{i,j}$ is the input image value at the $(i,j)^{th}$ pixel location.

Rotation and Gamma Correction were carried out 4 times for each image. The total data before augmentation was 184 images, but after augmentation, the amount of data was 1840 pap smear images. The illustration of data augmentation in this study can be seen in FIGURE 2.

C. U-NET CNN ARCHITECTURE

U-Net is a CNN architecture used in the medical world for image segmentation. U-Net was made up of an encoded line (left side) and a decoding line (right side). The encode path is used to preserve the image's context. To ensure proper localization, the decode path is used and the two paths are concatenated to create the letter U [32].

FIGURE 3 shows the U-Net CNN structure, where the light blue part presents the convolution layer with a kernel size 3×3 and the dark blue arrow presents batch normalization and ReLU. After that, a max-pooling 2×2 process is carried out to reduce the size of the convolution layer, then a transposed convolution 2×2 process is carried out to extract the image features along with merging two convolution layers, namely, concatenate until it reaches the original image size. Once it reaches its original size, it is convoluted with a 1×1 kernel. The U-Net CNN uses an encode-decode structure. As shown in Figure 3, the encode path on the left is used to get information from the image context by reducing the image size, while the decoding path on the right is used to get objects accurately in the segmentation process. The two paths on U-Net CNN are connected using a concatenate operation. The operations involved in the U-Net CNN layer are as follows:
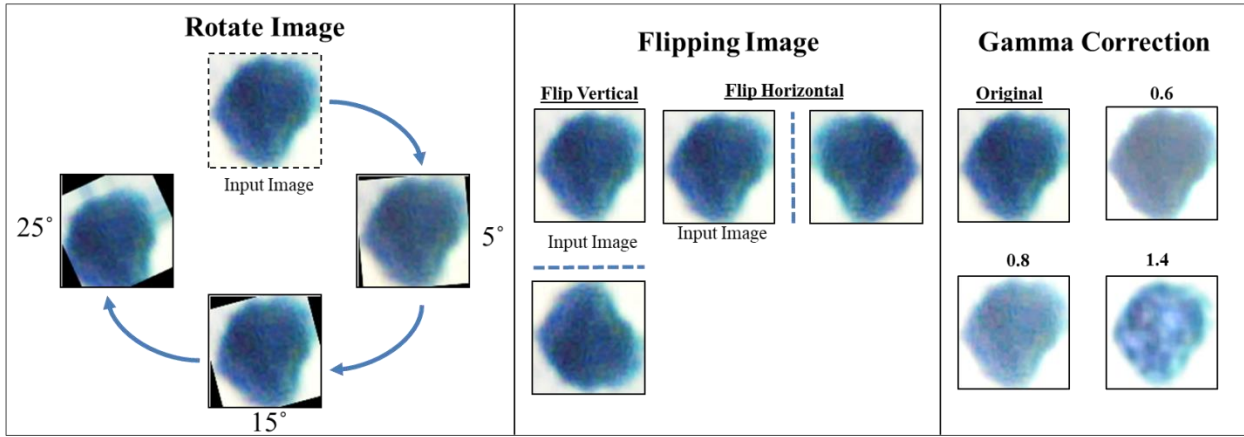
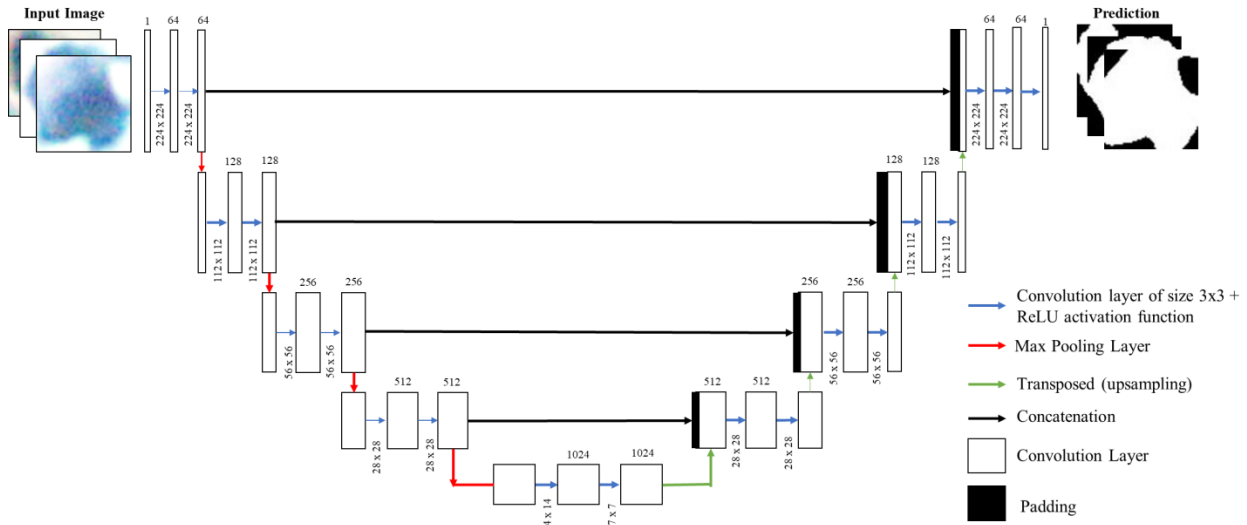**FIGURE 2.** Data augmentation using rotation, flipping, and correction

FIGURE 3. U-Net CNN Architecture Model

1) CONVOLUTION

The U-Net training process will pass through the convolution blocks, each of which consists of a convolution layer with a 3×3 kernel size, ReLU, and max pooling, and ends by passing a dropout layer on the third convolution block. The convolution layer can be processed using Equation 2 [32].

$$G_q^p = (D_p * k_q) + b_q, \qquad (2)$$

where * is a convolution operation. To calculate the results of each entry matrix $D_p * K_q$ can be obtained using Equation (3).

$$g_{i,j} = \left(\sum_{u=0}^{n-1} \sum_{v=0}^{n-1} (d_{u+i,v+j} \times k_{+1,j+1}) \right) + b_q, \qquad (3)$$

2) BATCH NORMALIZATION

Each layer in CNN architecture is dependent on the previous layers so large fluctuations in mean and input variance in the previous layers result in significant changes in parameters in the following layer. This issue can be addressed by utilizing batch normalization. Calculating batch normalization required the average (μ) and variance (σ^2), both of which are defined in the following Equation (4) and (5)[33].

$$E[G] = \mu_j = \frac{1}{m} \sum_{i=1}^{m} g_{ij} \qquad (4)$$

$$var(G) = \sigma_j^2 = \frac{1}{m} \sum_{i=1}^{m} (g_{ij} - \mu_j)^2 \qquad (5)$$

Where G is feature map's output matrix, i is row matrix, j is column matrix, n is number of mini-batch samples, m is number pixels in a single mini-batch, and g_{ij} is the matrix's pixel value. For normalization, it is used Equation (6).

$$\hat{G} = \frac{g_{ij} - \mu_j}{\sqrt{\sigma_j^2 + \varepsilon}}, \qquad (6)$$

σ_j is added to the denominator for numerical stability and is typically a small constant of the order 10^{-8} (do not divide by zero if $\sigma_j = 0$).

3) ACTIVATION FUNCTION

This activation function operates by accepting a value and performing a mathematical operation on it. The ReLU

Journal of Electronics, Electromedical Engineering, and Medical Informatics
Multidisciplinary: Rapid Review: Open Access Journal

Vol. 6, No. 3, July 2024, pp: 264-275; eISSN: 2656-8632

Equation can be written mathematically as Equation (7)[34]. The sigmoid function converted the input value feature map to a value between 0 and 1. Equation (8) is defined as a sigmoid function[35] where x is feature map's pixel value and $R(\text{x})$ is ReLU activation function to the pixel value x.

$$R(G) = max(g, 0) \qquad (7)$$

$$\sigma(x) = \frac{1}{1+e^{-g}} \qquad (8)$$

4) MAX-POOLING
Max pooling takes the maximum pixel value in the pooling region and ignores all other values, making it the pooling region's output. Max-pooling is defined mathematically as Equation (9) [36]. Where y is output pooling area, i is row index pooling area, j is column index pooling area, v is row index feature map after pooling, w is column index feature map after pooling.

$$y_{vw} = max\,(g_{ij}) \qquad (9)$$

5) TRANSPOSED CONVOLUTION
The transposed convolution operation generates an up sampling feature map output by randomizing several feature maps obtained by applying multiple convolution operations to the feature map input. Equations (10) and (11) illustrate transposed convolution operation[37]:

$$F_n = A_{ij} \circledast k \qquad (10)$$

$$F_{out} = F_1 \oplus F_2 \oplus \ldots \oplus F_n \qquad (11)$$

where $\circledast$ is convolution operation, $\oplus$ is periodic randomization and combination, i is row, j is column, F is an intermediate feature map generated, n is number pixels, A_{ij} is pixel value feature maps, k is a kernel, and F_{out} is output transposed convolution.

6) CONCATENATE LAYER
Concatenate layer is a technique for combining two CNN layers into a single. In this study, a low-layer feature map is the result of convolution performed before pooling in the encoding section. A high-layer feature map layer is the result of the decoding section's transposed convolution[38]. In this study, the low-layer feature map is a ReLu layer and the high-layer feature map is a transpose convolution layer where each color is a feature map matrix that is combined using concatenate as illustrated in FIGURE 3. In FIGURE 4, it can be seen that the dimensions of the low-layer feature map and the high-layer feature map had the same height (h) and width (w) with each number of low-layer matrices. The dimension size (channel) of the feature map layer (c_1) and high-layer feature map (c_2) are the same so that the concatenated results of the two layers produce feature maps concat dimensions, namely $h \times w \times (c_1 + c_2)$.
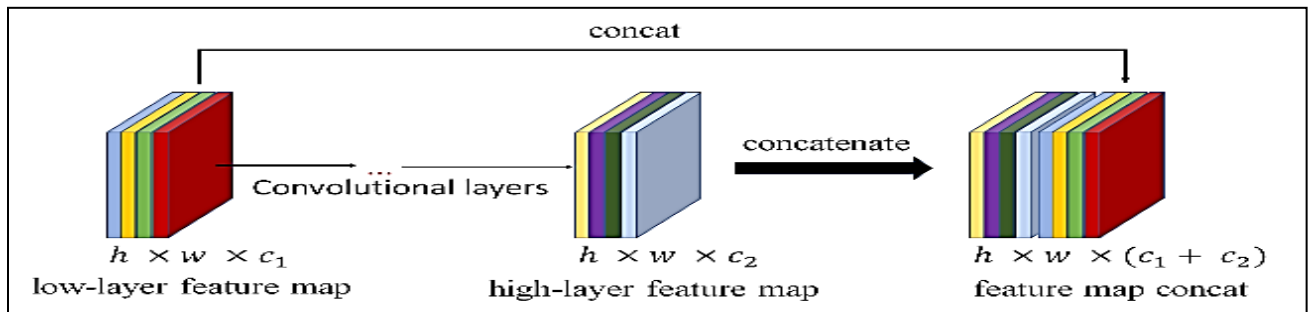

FIGURE 4. Illustration transposed convolution with 2×2 kernel matrix

7) TRAINING STAGE
At this stage, the goal is to train the data to recognize image features and patterns using the U-Net architectural model. In this study, the total input image is 1840 data. The data will be divided into 3 groups, namely 1208 images for training data, 294 for validation data, and 368 images for testing data. The testing data is used in the testing stage. A binary loss function is used to determine the magnitude of the error or loss to assess the model's performance in the training stage. Binary cross-entropy is defined in Equation (12)[39]. Where, n is the number of pixels, i is the pixel row, j is the pixel column, t_{ij} is the ground-truth, and o_{ij} is a predicted pixel.

$$L = -\frac{1}{n^2}\left[\sum_{i=1}^{n}\sum_{j=1}^{n}\left(t_{ij}\,log(o_{ij}) + (1 - t_{ij})\,log(1 - o_{ij})\right)\right] \qquad (12)$$

8) TESTING STAGE
The learning weights generated by the model at the training stage were implemented on 346 separate testing data. The data is to be segmented using the best weights in the training stage. The result of the application of these weights is a binary image

containing two features, namely nuclei and background. The result images should be compared to ground truth to evaluate the performances of U-Net CNN. The performances are accuracy, sensitivity, specificity, and F1-score.

9) EVALUATION
Evaluation is used to measure the results segmentation of nuclei cells and ground truth the entire cell to evaluate segmentation. Performance measures that are commonly used to measure a segmentation model are sensitivity, specificity, and accuracy[40]. Specificity and sensitivity refer to the algorithm's ability to distinguish between segmented and non-segmented areas. Accuracy is a metric that indicates how closely predicted[39]. In addition, another commonly used model performance measure is the F1-score. F1-score is a weighted average specificity and sensitivity measurement.[41] These performance measures are based on a confusion matrix. The confusion matrix consists of TP (True Positive), FP False Positive), TN (True Negative), and FN (False Negative)[42]. All the steps carried out in the proposed method can be seen

Journal of Electronics, Electromedical Engineering, and Medical Informatics
Multidisciplinary: Rapid Review: Open Access Journal

Vol. 6, No. 3, July 2024, pp: 264-275; eISSN: 2656-8632

as a flowchart in Fig. 4. This study evaluates the performance including accuracy (ACC), sensitivity (SN), specificity (SP), and F1-score (F1). These performances are defined in Equations (13), (14), (15), (16) [25].

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (13)$$

$$Sensitivity = \frac{TP}{TP + FN} \times 100\% \quad (14)$$

$$Spesitifity = \frac{TN}{TN + FP} \times 100\% \quad (15)$$

$$F1 - Score = \frac{2TP}{2TP + FP + FN} \times 100\% \quad (16)$$

III. RESULT

A. IMPLEMENTATION OF DATA AUGMENTATION ON IMAGE INPUT

Cervical cell images from the Pap smear dataset with a total of 184 images are used as image input, and after augmenting the data using rotation transformation, vertical flip, and horizontal flip. The total data is 2160 images. Some examples of augmented images can be seen in TABLE 1.
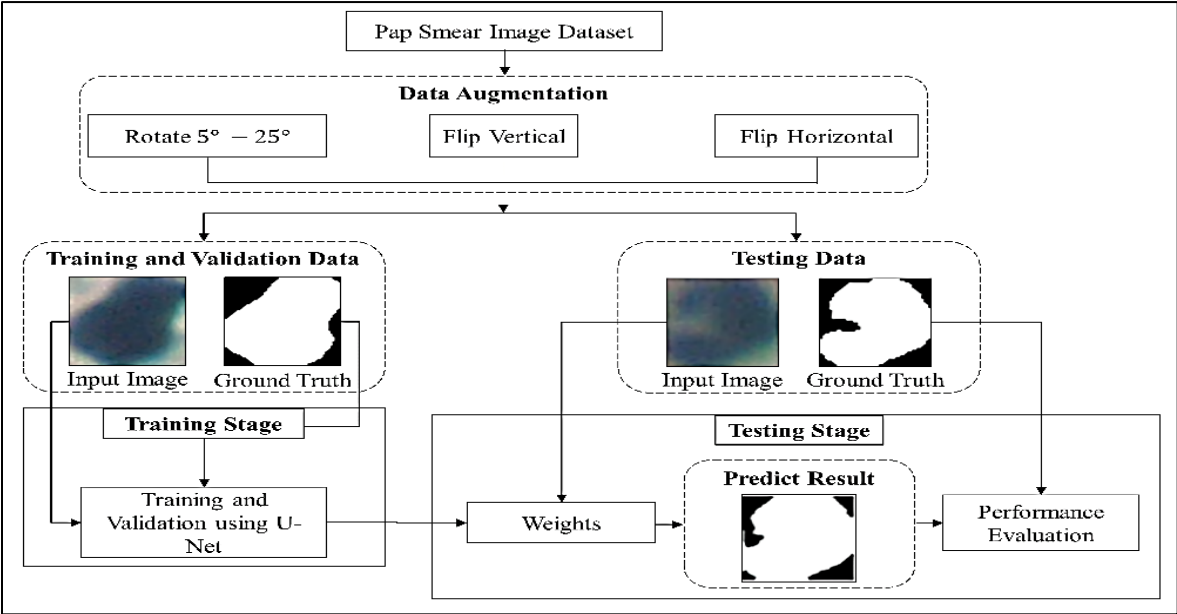

FIGURE 5. Research flowchart for segmentation of cervical cancer nuclei

TABLE 1
Augmentation of cervical cell image on pap smear dataset

No	File Name	Rotation 5°	Rotation 15°	Rotation 25°	GammaC 0.6	Gamma C 1.4	Flip Vertical	Flip Horizontal
1.	45_CAP091868							
2.	56_CAP091868							
3.	65_CAP091868							

B. IMPLEMENTATION OF U-NET ARCHITECTURE-BASED CNN MODEL

U-Net CNN was applied to Zenoda's original pap smear dataset and a new dataset which was the result of applying flipping, rotation, and Gamma Correction augmentation to pap smear images. For training on U-Net CNN, 100 epochs with 4 batch sizes in each epoch are taken for training. The resulting weights for each epoch are directly applied to the validation data to measure model performance in the training stage. To measure the error at each epoch, binary cross entropy is used. FIGURE 6 contains precision and recall graphs on the original dataset (a) and the new dataset (b). The blue line in the graph depicts the results on the training data and the yellow line represents the results on the validation data. The precision graphs on the training data and validation data have a large gap even though both graphs continue to

increase until the 100th epoch. However, a gap that is too large indicates overfitting in U-Net during training. Overfitting is a condition where a model or architecture is good at recognizing patterns in training data but is unable to work well in recognizing patterns in data that has never been trained by a U-Net CNN. The recall results in Figure 6a show that the recall patterns in the training data and testing data have no gaps and have the same pattern. Unfortunately, at the 20th epoch, the recall results continued to decline until the 100th epoch. the precision results of the graph in Figure 6b on the training data and validation data for the new dataset have almost the same pattern and reach a good fit condition until the 100th epoch. Fig. 6b shows the recall results have a close pattern at the 20th epoch. and reached a good fit condition up to the 100th epoch. The good fit condition shows that the model works excellently to recognize patterns in the testing data or data that has never been used. The recall graph in Figure 6b experiences an increase from epoch 4 and is stable until epoch 100.

FIGURE 7 shows the loss and accuracy results of the U-Net CNN on the original data (FIGURE 7a) and the loss and accuracy results on the new dataset (FIGURE 7b). FIGURE 7 shows the loss and accuracy results of the U-Net CNN for the original data (FIGURE 7a) and the loss and accuracy results on the new dataset (FIGURE 7b). The line graph in blue shows the loss and accuracy results of the training data, while the line graph in yellow shows the loss and accuracy results of the validation data. In FIGURE 7a, it is seen that the loss in the training and validation data from the original

dataset experienced a large gap at the 20th epoch. The accuracy results in FIGURE 7a on the training data and validation data also show a large gap between the accuracy of the training data and the accuracy of the data. testing. The loss and accuracy results of U-Net on the original dataset show that U-Net's performance is still poor because it has overfitting.

FIGURE 7b shows the loss and accuracy results of U-Net training on the dataset with augmentation (new dataset). The loss and accuracy graphs show the same pattern for both training data and testing data. The loss in FIGURE 7b decreased significantly at the 20th epoch and continued to decrease until the 100th epoch. The graph in FIGURE 7b shows that accuracy in the training data and validation data increased at the 20th epoch and continued to increase until the 100th epoch. The results of this graph show that U-Net has very good performance and does not experience overfitting when training on datasets that have undergone augmentation. The loss and accuracy graph in FIGURE 7b is in a good fit condition, namely that U-Net can recognize or predict patterns both in data that has been trained and data that has never been encountered before. The results show that the combination of augmentation techniques and U-Net CNN works excellent in nuclei segmentation of pap smear images. It means the proposed method can recognize nuclei in the pap smear image, even though the image has never been used in the training stage. The result of the training is the best weights from the training data. The best weights are applied but testing data with different data from training data.
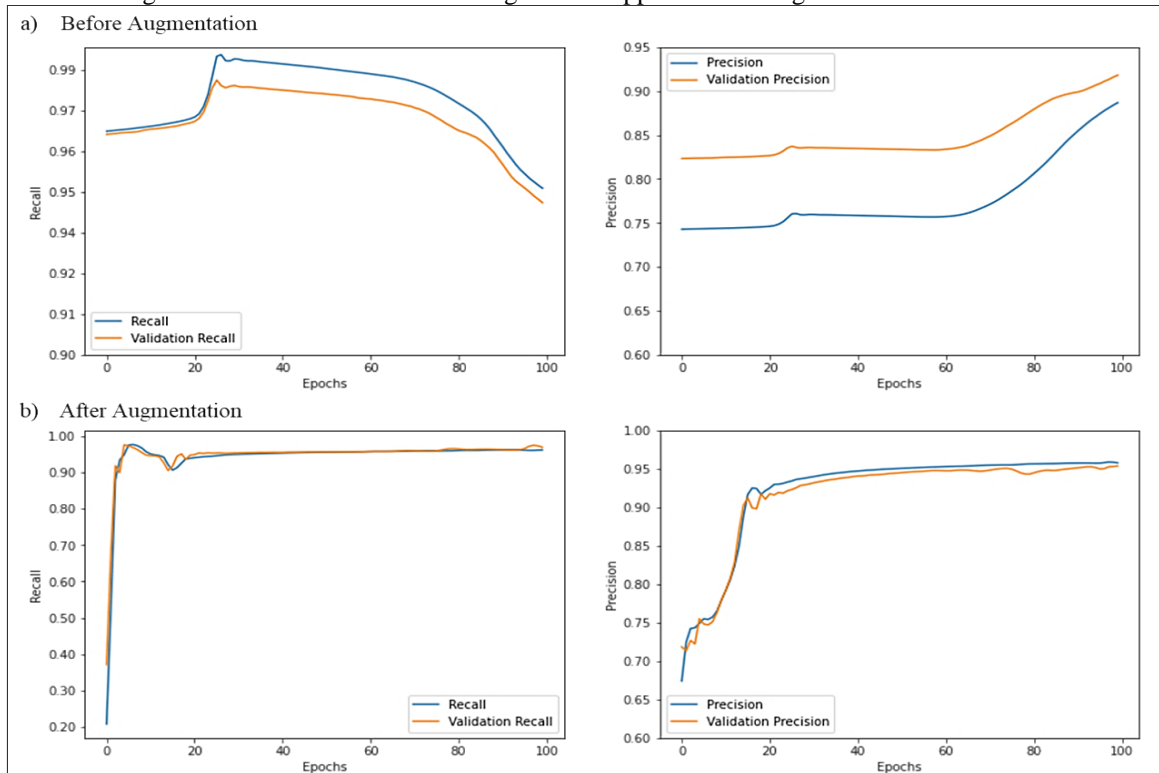

FIGURE 6. Graphs Model of recall and precision for training data and validation data (a) Before augmentation (b) After augmentation
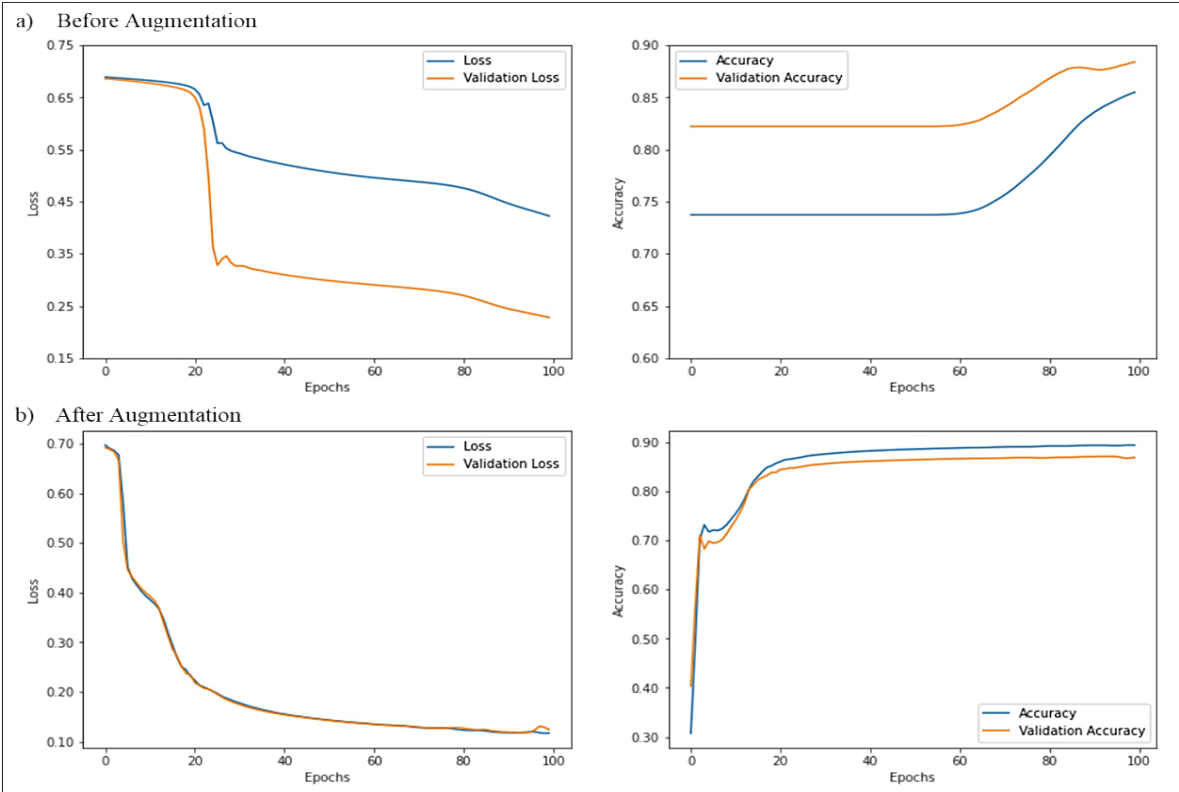
FIGURE 7. Graphs of loss and accuracy for training data and validation (a) Before augmentation (b) After augmentation
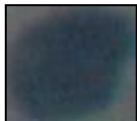
C. TESTING DATA

The data is tested at this stage using images that had not been studied by training data. The model weight values are used in the data testing process. The process's result is the binary segmentation of nuclei images. TABLE 2 shows the results of the data testing process for five examples of the 368 images original images, ground truth, and binary segmentation results. The results of the binary segmentation result obtained are True Positive (TP) of 281,282, indicating that the number of nuclei pixels from segmentation ground truth detected correctly is 281,282. False Positive (FP) is 45,360. It indicated that the number of pixels segmented against ground truth is not correctly detected at 45,360. Furthermore, a False Negative (FN) of 32,087, indicates that the number of non-nuclei pixels generated by segmentation on ground truth is not correctly detected at 32,087. True Negative (TN) is 1,055,671. It indicated that the number of non-nuclei pixels segmented against ground truth is correctly detected at 1,055,671. To determine a good threshold for the model, it used ROC (Receiver Operating Characteristics) parameters on segmentation results TPR (True Positive Rate) and FPR (False Positive Rate). The graph below depicts the relationship between TPR and FPR. FIGURE 8 is based on the ROC results obtained during the testing process.

FIGURE 8 shows the ROC results on the original dataset (8a) and ROC on the new dataset (8b). FIGURE 8 shows the calculation of the area under the blue line on the ROC graph with the Area Under Curve (AUC). The AUC for the original data set was 0.9439. AUC on the new dataset was able to increase the AUC results by 0.9847. These AUC results show that Combining augmentation and U-Net can effectively differentiate positive (core) and negative (background) classes. Based on the AUC results, the threshold value of the architectural model, which is greater than 0.9, indicated that the quality of the proposed method is excellent for segmenting the nuclei in pap smear images.

4. DISCUSSION

The accuracy (ACC), sensitivity (SS), specificity (SP), and F1-score (F1) of the model are determined by the results of cervical cancer nuclei segmentation. To determine the model's success rate, compare the performance results obtained with previous studies with different datasets, as shown in TABLE 3. TABLE 3 is used to see the performance of U-Net combined with augmentation in performing nuclei segmentation in pap smear images.

Journal of Electronics, Electromedical Engineering, and Medical Informatics
Multidisciplinary: Rapid Review: Open Access Journal

Vol. 6, No. 3, July 2024, pp: 264-275; eISSN: 2656-8632

TABLE 2
The example of the binary segmentation results and ground truth on test data on new pap smear images dataset

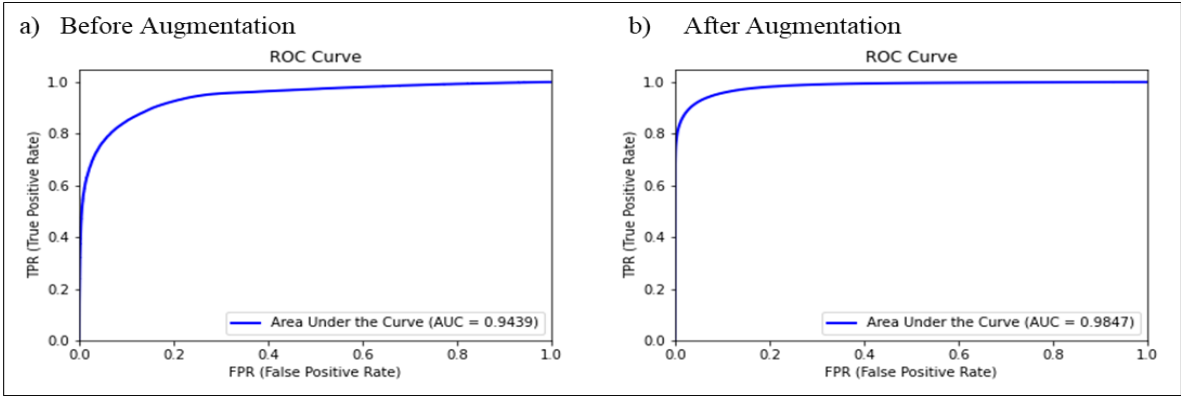| No	File Name	Original Image	Ground Truth	Segmentation Results
1	439_CAP091868			
2	32_CAP091868			
3	564_CAP091868			
4	369_CAP091868			
5	421_CAP091868			

FIGURE 8. Graph the ROC curve segmentation results on test data of nuclei segmentation a) Before augmentation b) After augmentation

TABLE 3 is a comparison of the results of the performance evaluation of the augmentation and U-Net with several other studies. In Table 3 several studies focus on the segmentation of the nuclei on pap smear images but with different datasets. Based on Table 3, the results of the accuracy, sensitivity, and F1-score of combination of augmentation and U-Net CNN are higher than the methods in other studies with results above 90%. These results indicate that the combination of augmentation and U-Net has excellent performance in nuclei segmentation on pap smear images. The specificity obtained for the proposed method is lower than [43]. However, it does not mean that the proposed method has poor performance because specificity indicates the method's ability to separate features that are not needed in the image (background), while sensitivity is used to measure the performance of the method

in obtaining nuclei features (foreground). The sensitivity in this study was much higher than the specificity. It means that the combination of augmentation and U-Net CNN has a more accurate ability to obtain the required nuclei features on pap smear images.

To see the effect of the augmentation techniques used on the U-Net CNN, the augmented dataset will be applied in several segmentation models. The results of applying several methods to the new dataset were compared with the results of U-Net CNN to see the success of U-Net CNN in segmenting nuclei in Pap Smear images. The comparison results can be seen in TABLE 4.

Journal of Electronics, Electromedical Engineering, and Medical Informatics
Multidisciplinary: Rapid Review: Open Access Journal

Vol. 6, No. 3, July 2024, pp: 264-275; eISSN: 2656-8632

TABLE 3
Comparison of the methods used with other research methods

Method	Data Set	ACC (%)	SN (%)	SP (%)	F1 (%)
Global and Local Graph Cuts[8]	Shenzhen Hospital dataset	85	87	85	89
ACM-AHP[9]	DICOM Pap Smear dataset	89.6	93	89.41	72.1
Hierarchical median narrow band level [44]	EDF Image dataset	-	90	90	92
Mean-Shif Algorithm[29]	Personal dataset	92.9	94.25	93.45	-
Bi-path CNN[25]	Herlev Dataset	77	72	71	69
Adaptive Shape[26]	ISBI Dataset	-	87.34	-	89.51
FCM[45]	ISBI	-	91.5	91.8	91.6
Super pixel[43]	Fiji dataset	-	89.5	95.6	92.6
Proposed method	Zenoda dataset	93	95	93	94

TABLE 4
The comparison of performance results of several methods for nuclei segmentation on dataset pap smear image after augmentation

No	Method	Accuracy (%)	Sensitivity (%)	Specificity (%)	F1-Score (%)
1	Otsu Threshold [46]	88	77	71	87
2	Global Threshold [47]	87	74	80	87
3	Sauvola Threshold [48]	83	71	76	75
4	Niblack Threshold [49]	85	72	72	71
5	Canny Edge Detection [50]	84	70	75	74
6	Proposed Method	93	95	93	94

TABLE 4 is a comparison of the performance results of the proposed method with several other methods on dataset augmentation results for core segmentation in pap smear images. TABLE 4 contains the results of core segmentation with various methods on the same dataset. The dataset used is the Zenodo dataset which has been augmented. Based on Table 4, the accuracy, sensitivity, specificity, and F1-score results of the augmentation and U-Net CNN are the highest compared to the performance results of the results in this study. The performance result that shows the success of the method for nuclei segmentation on pap smears can be seen from the sensitivity result. The highest sensitivity results were obtained by U-Net CNN. The research results show that the combination of augmentation using flip, rotation, and Gamma Correction techniques can significantly improve the performance of U-Net CNN. The combination of augmentation and U-Net CNN is excellent in nuclei segmentation in Pap smear images.

Based on these performance results, it can be concluded that the results of the resulting image of segmentation and ground truth have a great similarity. The proposed method is still limited to one feature segmentation, while the detection of cervical cancer requires not only the nuclei but also other cells such as the cytoplasm. In addition, this study only reaches the segmentation stage and does not reach the classification stage. In future studies, this study can be expanded for other features such as cytoplasmic

segmentation and continued for the classification stage so the results can be applied in building an automatic cervical cancer detection.

IV.   CONCLUSION

Based on research findings and discussion, it can be concluded that applying augmentation to small amounts of data can improve U-Net performance results. Combining flipping, rotation, and Gamma Correction as augmentation techniques can meet the training data needs of U-Net CNN. The performance results of the U-Net CNN on augmented data in the core binary segmentation process of Pap smear images can improve the performance results of U-Net on original data. The average performance of accuracy, sensitivity, specificity, and F1-score is above 90%. These results mean that U-Net has excellent performance in separating nuclei as foreground and other features as background. The results obtained also show that augmentation techniques can influence the performance of CNN architectures. The combination of augmentation techniques and U-Net CNN provides excellent and powerful performance in performing core segmentation on Pap Smear images. In further research, methods can be developed to detect other cells such as cytoplasm in pap smear images with other algorithms such as YOLO, RCNN, and others. For the classification stage, this method can be applied to develop automatic cervical cancer detection.

REFERENCES

[1]   World Health Organization, "Indonesia Source GLOBOCAN 2020," *Int. Agency Res. Cancer*, 2021.

[2]   Ernawati, D. Oktaviana, Mantasia, R. A. Yusuf, and Sumarmi, "The Effect of Health Education Based on the Health Belief Model about Pap Smear Test on Women in Rural District Indonesia," *Medico-legal Updat.*, vol. 21, no. 2, pp. 7–12, 2021.

[3]   H. Bandyopadhyay and M. Nasipuri, "Segmentation of Pap Smear Images for Cervical Cancer Detection," *2020 IEEE Calcutta Conf. CALCON 2020 - Proc.*, pp. 30–33, 2020, doi: 10.1109/CALCON49167.2020.9106484.

Journal of Electronics, Electromedical Engineering, and Medical Informatics
Multidisciplinary: Rapid Review: Open Access Journal

Vol. 6, No. 3, July 2024, pp: 264-275; eISSN: 2656-8632

[4] D. Somasundaram, S. Gnanasaravanan, and N. Madian, "Automatic segmentation of nuclei from pap smear cell images: A step toward cervical cancer screening," *Int. J. Imaging Syst. Technol.*, vol. 30, no. 4, pp. 1209–1219, 2020, doi: 10.1002/ima.22444.

[5] M. Nikolic, E. Tuba, and M. Tuba, "Edge Detection in Medical Ultrasound Images Using Adjusted Canny Edge Detection Algorithm," *IEEE*, 2016.

[6] A. Taneja, P. Ranjan, and A. Ujlayan, "Automated cell nuclei segmentation in overlapping cervical images using deep learning model," *Proc. 2018 Int. Conf. Image Process. Comput. Vision, Pattern Recognition, IPCV 2018*, pp. 165–172, 2018.

[7] Y. Chen, J. Chen, D. Wei, Y. Li, and Y. Zheng, *Multiscale Multimodal Medical Imaging*, vol. 11977. 2020.

[8] L. Zhang *et al.*, "Segmentation of cytoplasm and nuclei of abnormal cells in cervical cytology using global and local graph cuts," *Comput. Med. Imaging Graph.*, vol. 38, no. 5, pp. 369–380, 2014, doi: 10.1016/j.compmedimag.2014.02.001.

[9] S. Geetha and S. Suganya, "Automatic Detection of Cervical Cancer Using UKF and ACM-AHP," *Int. J. Adv. Res. Eng. Technol.*, vol. 11, no. 8, pp. 285–295, 2020, doi: 10.34218/IJARET.11.8.2020.028.

[10] J. Liu, L. Li, and L. Wang, "Acetowhite region segmentation in uterine cervix images using a registered ratio image," *Comput. Biol. Med.*, vol. 93, pp. 47–55, 2018, doi: 10.1016/j.compbiomed.2017.12.009.

[11] S. Sivagami, P. Chitra, G. S. R. Kailash, and S. R. Muralidharan, "UNet Architecture Based Dental Panoramic Image Segmentation," *2020 Int. Conf. Wirel. Commun. Signal Process. Networking, WiSPNET 2020*, pp. 187–191, 2020, doi: 10.1109/WiSPNET48689.2020.9198370.

[12] J. Schmidhuber, "Deep Learning in neural networks: An overview," *Neural Networks*, vol. 61, pp. 85–117, 2015, doi: 10.1016/j.neunet.2014.09.003.

[13] W. Liu *et al.*, "CVM-Cervix: A hybrid cervical Pap-smear image classification framework using CNN, visual transformer and multilayer perceptron," *Pattern Recognit.*, vol. 130, p. 108829, 2022, doi: https://doi.org/10.1016/j.patcog.2022.108829.

[14] O. Ronneberger, P. Fischer, and T. Brox, "U-net: convolutional networks for biomedical image segmentation," in *International Conference on Medical Image Computing and Computer-Assited Intervention (MICCAI) 2015*, 2015, vol. 9351, pp. 234–241, doi: 10.1007/978-3-319-24574-4_28.

[15] V. Khryashchev, A. Lebedev, O. Stepanova, and A. Srednyakova, "Using Convolutional Neural Networks in the Problem of Cell Nuclei Segmentation on Histological Images," *Springer Nat. Switz. AG 2019*, pp. 149–161, 2019, doi: 10.1007/978-3-030-12072-6.

[16] S. Albawi, T. A. M. Mohammed, and S. Alzawi, "Layers of a Convolutional Neural Network," in *2017 International Conference on Engineering and Technology (ICET)*, 2017, pp. 1–6.

[17] M. Sonogashira, M. Shonai, and M. Iiyama, "High-resolution bathymetry by deep-learning-based image superresolution," *PLoS One*, vol. 15, no. 7, pp. 1–19, 2020, doi: 10.1371/journal.pone.0235487.

[18] W. Chen, B. Yang, J. Li, and J. Wang, "An approach to detecting diabetic retinopathy based on integrated shallow convolutional neural networks," *IEEE Access*, vol. 8, pp. 178552–178562, 2020, doi: 10.1109/ACCESS.2020.3027794.

[19] J. F. R. Rochac, L. Liang, N. Zhang, and T. Oladunni, "A Gaussian Data Augmentation Technique on Highly Dimensional, Limited Labeled Data for Multiclass Classification Using Deep Learning," *10th Int. Conf. Intell. Control Inf. Process. ICICIP 2019*, pp. 145–151, 2019, doi: 10.1109/ICICIP47338.2019.9012197.

[20] S. Dash, P. K. Sethy, and S. K. Behera, "Cervical Transformation Zone Segmentation and Classification based on Improved Inception-ResNet-V2 Using Colposcopy Images," *Cancer Inform.*, vol. 22, p. 11769351231161476, 2023, doi: 10.1177/11769351231161477.

[21] N. Sompawong *et al.*, "Automated Pap Smear Cervical Cancer Screening Using Deep Learning," *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, pp. 7044–7048, 2019, doi: 10.1109/EMBC.2019.8856369.

[22] S. I. Maiyanti *et al.*, "Rotation-Gamma Correction Augmentation on Cnn-Dense Block for Soil Image Classification," *Appl. Comput. Sci.*, vol. 19, no. 3, pp. 96–115, 2023, doi: 10.35784/acs-2023-27.

[23] A.-Y. Chang *et al.*, "GaP-Aug: Gamma Patch-Wise Correction Augmentation Method for Respiratory Sound Classification," in *ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2024, pp. 551–555, doi: 10.1109/ICASSP48485.2024.10447967.

[24] E. Hussain, L. B. Mahanta, C. R. Das, and R. K. Talukdar, "A comprehensive study on the multi-class cervical cancer diagnostic prediction on pap smear images using a fusion-based decision from ensemble deep convolutional neural network," *Tissue Cell*, vol. 65, no. 101347, pp. 1–8, 2020, doi: 10.1016/j.tice.2020.101347.

[25] A. Desiani, 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," *IAENG Int. J. Comput. Sci.*, vol. 48, no. 3, pp. 1–10, 2021.

[26] Y. Song, L. Zhu, J. Qin, B. Lei, B. Sheng, and K. S. Choi, "Segmentation of Overlapping Cytoplasm in Cervical Smear Images via Adaptive Shape Priors Extracted from Contour Fragments," *IEEE Trans. Med. Imaging*, vol. 38, no. 12, pp. 2849–2862, 2019, doi: 10.1109/TMI.2019.2915633.

[27] R. Rudiansyah, L. Iryani, L. I. Kesuma, P. Sari, and A. Alamsyah, "Combination of Image Enhancement and U-Net Architecture for Cervical Cell Semantic Segmentation," *JITE (Journal Informatics Telecommun. Eng.*, vol. 7, no. July, pp. 308–316, 2023.

[28] D. S. Wita, "Image Segmentation of Normal Pap Smear Thinprep using U-Net with Mobilenetv2 Encoder," *J. Med. Informatics Technol.*, vol. 1, no. 2, pp. 31–35, 2023, doi: 10.37034/medinftech.v1i2.6.

[29] P. Wang, L. Wang, Y. Li, Q. Song, S. Lv, and X. Hu, "Automatic cell nuclei segmentation and classification of cervical Pap smear images," *Biomed. Signal Process. Control*, vol. 48, pp. 93–103, 2019, doi: 10.1016/j.bspc.2018.09.008.

[30] A. Tareef *et al.*, "Automatic segmentation of overlapping cervical smear cells based on local distinctive features and guided shape deformation," *Neurocomputing*, vol. 221, pp. 94–107, 2017, doi: 10.1016/j.neucom.2016.09.070.

[31] Erwin, H. K. Putra, B. Suprihatin, and F. Ramadhini, "A Hybrid CLAHE-GAMMA Adjustment and Densely Connected U-NET for Retinal Blood Vessel Segmentation using Augmentation Data," *Eng. Lett.*, vol. 30, no. 2, pp. 485–493, 2022.

[32] 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, no. 16, pp. 1–27, 2023, doi: 10.1142/S0218001423570161.

[33] S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," *Proc. 32nd Int. Conf. Mach. Learn. Lille, Fr. 2015.*, vol. 37, 2015.

[34] S. Hijazi, R. Kumar, and C. Rowen, "Using convolutional neural networks for image recognition," Cadence Design Systems Inc.: San Jose, CA, USA, 2015, pp. 1–12.

[35] S. Sharma, S. Sharma, and A. Anidhya, "Activation Functions in Neural Networks," *Int. J. Eng. Appl. Sci. Technol.*, vol. 4, no. 12, pp. 310–316, 2018.

[36] Z. Tong, K. Aihara, and G. Tanaka, *A hybrid pooling method for convolutional neural networks*. 2016.

[37] H. Gao, H. Yuan, Z. Wang, and S. Ji, "Pixel Transposed Convolutional Networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 5, pp. 1218–1227, 2019, doi: 10.1109/TPAMI.2019.2893965.

[38] C. Du, Y. Wang, C. Wang, C. Shi, and B. Xiao, "Selective feature connection mechanism: Concatenating multi-layer CNN features with a feature selector," *Pattern Recognit. Lett.*, vol. 129, pp. 108–114, 2020, doi: 10.1016/j.patrec.2019.11.015.

[39] T. Chankong, N. Theera-Umpon, and S. Auephanwiriyakul, "Automatic cervical cell segmentation and classification in Pap smears," *Comput. Methods Programs Biomed.*, vol. 113, no. 2, pp. 539–556, 2014, doi: 10.1016/j.cmpb.2013.12.012.

[40] A. Desiani, S. Yahdin, A. Kartikasari, and Irmeilyana, "Handling the imbalanced data with missing value elimination SMOTE in the classification of the relevance education background with graduates employment," *IAES Int. J. Artif. Intell.*, vol. 10, no. 2, pp. 346–354, 2021, doi: 10.11591/ijai.v10.i2.pp346-354.

[41] O. Oriola, "A Stacked Generalization Ensemble Approach for Improved Intrusion Detection," vol. 18, no. 5, pp. 62–67, 2020.

[42] A. Desiani, N. R. Dewi, A. N. Fauza, N. Rachmatullah, M. Arhami, and M. Nawawi, "Handling missing data using combination of

deletion technique, mean, mode and artificial neural network imputation for heart disease dataset," *Sci. Technol. Indones.*, vol. 6, no. 4, pp. 303–312, 2021, doi: 10.26554/sti.2021.6.4.303-312.

[43] D. Ushizima, A. G. C. Bianchi, and C. M. Carneiro, "Segmentation of subcellular compartments combining superpixel representation with Voronoi diagrams," *Off. Energy Res. U.S. Dep. Energy, under Contract Number DE-AC02-05CH11231*, pp. 1–3, 2015.

[44] A. M. Braga *et al.*, "Hierarchical median narrow band for level set segmentation of cervical cell nuclei," *Meas. J. Int. Meas. Confed.*, vol. 176, no. March, p. 109232, 2021, doi: 10.1016/j.measurement.2021.109232.

[45] R. Saha, M. Bajger, and G. Lee, "Spatial Shape Constrained Fuzzy C-Means (FCM) Clustering for Nucleus Segmentation in Pap Smear Images," *2016 Int. Conf. Digit. Image Comput. Tech. Appl. DICTA 2016*, pp. 1–8, 2016, doi: 10.1109/DICTA.2016.7797086.

[46] P. Yang, W. Song, X. Zhao, R. Z. Aheng, and L. Qingge, "An improved Otsu threshold segmentation algorithm Wei Song *, Xiaobing Zhao and Rui Zheng Letu Qingge," *Int. J. Comput. Sci. Eng.*, vol. 22, no. 1, pp. 146–153, 2020.

[47] N. Senthilkumaran and S. Vaithegi, "Image Segmentation by Using Thresholding Techniques fpr Medical images," *Comput. Sci. Eng. An Int. J.*, vol. 6, no. 1, pp. 1–13, 2016, doi: 10.5121/cseij.2016.6101.

[48] M. Kiran, I. Ahmed, N. Khan, and A. G. Reddy, "Chest X-ray segmentation using Sauvola thresholding and Gaussian derivatives responses," *J. Ambient Intell. Humaniz. Comput.*, vol. 10, no. 10, pp. 4179–4195, 2019, doi: 10.1007/s12652-019-01281-7.

[49] O. A. Samorodova and A. V Samorodov, "Fast implementation of the Niblack binarization algorithm for microscope image segmentation," *Pattern Recognit. Image Anal.*, vol. 26, no. 3, pp. 548–551, 2016, doi: 10.1134/S1054661816030020.

[50] Z. Xu, X. Baojie, and W. Guoxin, "Canny edge detection based on Open CV," in *2017 13th IEEE International Conference on Electronic Measurement & Instruments (ICEMI)*, 2017, pp. 53–56, doi: 10.1109/ICEMI.2017.8265710.

AUTHORS BIOGRAPHY

Anita Desiani. In 2020, She is currently working on a project for her Doctoral Program at Mathematics and Natural Science Faculty, Universitas Sriwijaya. She received mathematics bachelor from Universitas Sriwijaya in 2000, a magister degree in Computer Science from Universitas Gadjah Mada in 2003, and a doctoral degree in Mathematics and science department in 2022. In 2004, she joined as a lecturer at Mathematics Department at Universitas Sriwijaya until now. Her current research interests include the field of data mining, image processing, pattern recognition and computer vision, and artificial intelligence.

Irmeilyana was born on May 17, 1974. She earned her Bachelor of Science degree in Mathematics from Institut Pertanian Bogor in 1997, followed by a Master of Science degree in Mathematics from Institut Teknologi Bandung in 1999. In 2005, she joined Universitas Sriwijaya, as a lecturer, she now works in the Faculty of Mathematics and Science at Universitas sriwijaya.

Des Alwine Zayanti was born on December 4, 1970. She obtained her bachelor's degree from Sriwijaya University in 1994. Then, she pursued her master's degree at Institut Teknologi Bandung in 2005. In 2004, she joined Universitas Sriwijaya as a lecturer. Currently, she holds a position in the Faculty of Mathematics and Science at Universitas Sriwijaya

Yadi Utama was born on November 15 1979. He obtained a bachelor's degree from Gadjah Mada University in 2004. Then, a master's degree at Gadjah Mada University in 2006. In 2007, he joined Sriwijaya University as a lecturer. He now works in the Information Systems at Sriwijaya University.

Muhammad Arhami completed his bachelor's degree at Universitas Syiah Kuala in 1998, followed by a master's degree at Universitas Gadjah Mada in 2004. Currently, he is affiliated with Politeknik Negeri Lhokseumawe, specializing in Informatics Engineering. current research interests include machine learning, artificial intelligence, data mining, natural language processing, and expert systems. Arhami's dedication to advancing these fields contributes significantly to research and education in technology and computer science.

Azhar Kholiq Affandi, born on September 15, 1961. He graduated with a bachelor's degree from Institut Teknologi Bandung in 1988. He further pursued his studies and obtained a Master of Science degree from Institut Teknologi Bandung in 1993, followed by a Doctor of Engineering degree from Saitama University in 2008. In 2002, he commenced his career as a lecturer at Universitas Sriwijaya. In 2002, he joined Sriwijaya University as a lecturer. Now, he works in Department of Physics at Sriwijaya University.

Muhammad Aditya Sasongko Muhammad Aditya Sasongko. His Research interests are in performance analysis techniques for heterogeneous computing systems, and benchmarking for hardware components. He received a bachelor degree in Computer Education and Instructioal Technology from Middle East Technical University in 2012, a magister degree in Computer Engineering from Marmara University in 2017, and a doctoral degree in Computer science and engineering from Koç University in 2022. Since 2022, He has been working as a postdoctoral researcher at the Computer Engineering Department of Koç University until now.

Indri Ramayanti was born in Palembang, Indonesia, in 1983. She received his Bachelor of Biology from Universitas Sriwijaya, Indonesia, in 2005, and an M.Sc degree in Tropical Medicine from Universitas Gadjah Mada (UGM), Yogyakarta, Indonesia, in 2008. In 2009, she joined Universitas Muhammadiyah Palembang, as a lecturer, she now works in the Faculty of Medicine at Universitas Muhammadiyah Palembang. Then, in 2022, she received his environmental science doctorate at Sriwijaya University Postgraduate Program. Her current research interests include parasitology.