THE AUGMENTATION DATA OF RETINE IMAGE FOR BLOOD VESSEL SEGMENTATION USING U-NET CONVOLUTIONAL NEURAL NETWORK METHOD

by Erwin Erwin

International Journal of Computational Intelligence and Applications

THE AUGMENTATION DATA OF RETINE IMAGE FOR BLOOD VESSEL SEGMENTATION USING U-NET CONVOLUTIONAL NEURAL NETWORK METHOD

ERWIN*
Department of Computer Engineering, University of Sriwijaya, Jalan Raya Palembang-Unsri KM 32 Indralaya, Indonesia.
erwin@unsri.ac.id

ASRI SAFMI
Department of Computer Engineering, University of Sriwijaya, Jalan Raya Palembang-Unsri KM 32 Indralaya, Indonesia
asrisafmi75@gmail.com

ANITA DESIANI
Department of Mathematics, University of Sriwijaya, Jalan Raya Palembang-Unsri KM 32 Indralaya, Indonesia
anita_desiani@unsri.ac.id

BAMBANG SUPRIHATIN
Department of Mathematics, University of Sriwijaya, Jalan Raya Palembang-Unsri KM 32 Indralaya, Indonesia
bambangs@unsri.ac.id

FATHONI
Department of Information System, University of Sriwijaya, Jalan Raya Palembang-Unsri KM 32 Indralaya, Indonesia
fathoni@unsri.ac.id

Received 22 May 2021
Revised 2 October 2021
Accepted 13 October 2021

Abstract. The retina is the most important part of the eye. Proper segmentation of retinal blood vessels can be the first step to detecting a disease. Segmentation and analysis of retinal blood vessels can assist medical personnel in detecting the severity of a disease. In this paper, segmentation is proposed using U-net architecture in the Convolutional Neural Network (CNN) method to train a semantic segmentation model on retinal blood vessels. U-Net architecture is one of the CNN Deep Learning methods that are most often used in the image segmentation stage. U-Net architecture is built on Fully Convolutional Network (FCN) and modified to produce better segmentation in medical images. Many methods can be used for the retinal image segmentation process, but some of them cannot get retinal vessel segmentation properly because some of the blood vessels are thin and not visible. In addition, the Contrast Limited Adaptive Histogram Equalization (CLAHE) method is used to increase the grayscale contrast and the Median Filter is used to obtain better image quality. Data augmentation is also used to maximize the number of datasets owned so that more. In this study, the dataset used was STARE which yielded performance values of accuracy, sensitivity, specificity, precision, and f1-score, respectively 97.64%, 78.18%, 99.20%, 88.77%, and 82.91. %.

Keywords: Augmentation; CLAHE; CNN; Diabetic Retinopathy; Median Filter; U-net.

* Correspondence Author

1. Introduction

The Retina is one of the most important parts of the eye. With proper feature extraction, it can be the first step for a disease to be detected quickly and effectively on a retinal image. Commonly detected parts of the retina include blood vessels, optic disk, macula, exudate, and hemorrhage. Of the several parts detected, blood vessels are one of the most important parts of the retinal image. The retinal blood vessels consist of arteries and arterioles that, when examined at certain intervals, can help diagnose a disease and can help medical personnel. Therefore, the proper extraction of blood vessels from the retinal image will help reduce dependence on the level of expertise of medical personnel and reduce error factors and shorten the time [1]. In ophthalmology, retinal fundus images play a role in the medical field by diagnosing diseases, for example, diabetes, glaucoma, cataracts, and others [2].

Retinal blood vessels are one of the important structures contained in the fundus retina. Vascular changes are closely associated with many systemic, metabolic, and hematology diseases. Retina vascular morphology can be used to identify and classify the severity of several diseases. A step such as segmentation and analyzing retina blood vessels can help medical personnel in detecting these changes. Blood vessel segmentation will produce information about the location of the blood vessels and also the variation in diameter of the existing blood vessels [3]. On the other hand, retina blood vessels can be extracted in several retinal imaging processes to estimate the status of vascular tissue in the retina. The final result obtained from the segmentation will be used as a parameter to ensure the performance result of the feature extraction.

In the process of detecting blood vessels on retinal images, sometimes there are still difficulties in the formation of a blood vessel that is not detected correctly. However, a method increasing contrast and improving retinal image quality is needed so that the results obtained are better, a method that is quite good in improving the quality of retinal images, namely the Contrast Limited Adaptive Histogram Equalization (CLAHE) method. CLAHE is used to correct problems in the form of low image contrast on a digital image, especially a medical image [4 - 9]. Then the median filter can also be selected as a method to improve the quality of the retinal image, in the median filter, the damaged pixel value will be replaced by the median of the window pixel value [5,10].

Vascular segmentation of retinal fundus image is one of the powerful methods for diagnosing vascular disease. Many methods can be used in the process of segmentation of retinal fundus image, but some of them cannot get a good image of retinal vessel segmentation in some thin blood vessels because they are not visible. Currently, deep learning has been widely used in the process of segmentating blood vessels in the retinal fundus and has archived several achievements in terms of classification, detection, and segmentation [2, 8,11]. Especially in the Convolutional Neural Network (CNN) algorithm that was selected for retinal fundus image segmentation. CNN is the most commonly used algorithm in computer vision to classify each pixel of the image. CNN showed impressive performance in completing segmentation tasks and still leaves room for improvement to produce more accurate segmentation [2, 6, 7, 12 - 15]. CNN is a development of Multilayer Perceptron (MLP) which is designed to process data in the form of images and is included in the Deep Neural Network because of its high network depth and is widely applied to image data. CNN is designed to recognize the visual pattern of an image pixel by going through pre-processing.

CNN is widely used in complex image classification tasks with many class objects with an impressive results. One of the most used architectures and produces good accuracy is the U-net architecture. U-net architecture is built on the Fully Convolutional Network (FCN) and modified in such a way as to produce better segmentation in medical images. Compared to FCN, U-net architecture is more symmetrical and the skip connection between the down-sampling line and the up-sampling line applies the combined operation rather than summation [2, 6, 12, 13]. Because the U-net architecture is a high capacity model, an augmentation stage is needed to add training data to avoid overfitting. Augmentation is a step used to study neural networks in the image when the available sample is limited [8].

In this study, we propose a new technique for retinal blood vessel segmentation by combining image enhancement and augmentation on the U-net architecture on CNN. In general, the steps proposed are to separate the Red channel, Green channel, and Blue channel. This color channel separation is done to produce an image with a grayscale. Then, the contrast and image quality were improved using CLAHE. CLAHE is better known as a block-based process to help solve the problem of peaks in the histogram that causes image noise. The purpose of using CLAHE is to obtain an image with an intensity level that is uniformly distributed across the intensity scale. CLAHE can produce great results for medical images. After that, the median filter will be used as a step to reduce noise that arises from the previous contrast enhancement results. The median filter is a non-linear digital filter technique, which is often used to remove the noise contained in an image. In the median filter, the defective pixel value will be replaced by the median of the window pixel value. Compared to other linear filters, the median filter performs very well in retinal image enhancement. In the next stage, we perform augmentation on the image. Augmentation is a process to increase the size and quality of a limited number of training datasets. Augmentation aims to increase diversity and modify images. Augmentation carried out in this study consisted of flip horizontal, flip vertical, and flip reverse vertical.

The most significant contributions produced in this study are:

- Augmentation of data on CNN method with U-net architecture has been applied to the process of segmenting blood vessels in retinal images.
- Image enhancement with CLAHE to increase contrast and median filter to remove noise is very important before the segmentation process is carried out.
- The combination of the enhancement and augmentation stages has increased the accuracy and specificity values by 1%-2% from Refs. 4, 8, 12, and 14.

2. Related Work

Currently, Deep Learning has been widely used in the research process to segment retinal blood vessels and has archived some great achievements in treating several diseases. For example, Shaohua Wan *et al.*[11] proposed that feature extraction using deep learning methods such as Alexnet, Vggnet, Googlenet, and Resnet are the newest CNN methods. By adding the normalization and augmentation methods to the preprocessing so that more datasets will be segmented and produce a more accurate fundus image. However, in carrying out the segmentation learning process it cannot be said to be good because the pre-process is used only a little.

CNN is designed to recognize the visual pattern of an image pixel by going through pre-processing[16]. CNN has several layers, including the convolutional layer, non-linearity layer, pooling layer, and fully connected layer. The convolutional layer and the fully connected layer have a parameter while the pooling layer and the non-linearity layer have no parameters. CNN has a very good performance in terms of machine learning, especially in image processing such as image classification, computer vision, and natural language processing[17]. Each CNN layer takes several inputs and converts them into outputs[18].

S. Dutta *et al.*[10] proposed an approach with several stages to start research, namely with the pre-processing stage, the image will be extracted to the RGB image will be taken and converted into grayscale to facilitate the process of removing noise and improving the image for more feature extraction by using 2 stages namely Median Filter and Edge Detection. Feature Extraction and Fuzzy C-Means are used for the filter classification process. Backpropagation NN, DNN, and CNN models are used in the post-processing and segmentation stages. However, in conducting training on the dataset, the Back Propagation NN model takes longer and gets lower accuracy than the DNN and CNN models.

Khan Bahadar Khan *et al.*[4] proposed techniques that they thought were efficient for segmentation of the retina, namely CLAHE and MISODATA. CLAHE was chosen in the preprocessing stage as a step to increase contrast in the image and be able to improve image quality well while MISODATA was used in the segmentation process. However, the performance and research result depends on the stage of cleaning the vessel image which is still considered to have noise. So that all blood vessels that have the same size as the existing noise are eliminated. The segmentation process using MISODATA in this study takes a long time compared to other methods.

Ibrahim Atli *et al.*[8] proposed vascular segmentation using a new model they call the Sine-Net that applied Up-sampling and Down-sampling to capture the features of thin and thick blood vessels. The segmentation image will go through the pre-processing stage using CLAHE and also Multiscale Top Hat Transform (MTHT). Furthermore, the retinal image will be augmented so that a limited number of datasets can be reproduced. But in some metrics its show a decrease that causes a loss.

J Dash *et al.*[9] proposed the extraction of blood vessels using Gamma Correction and CLAHE. Furthermore, image segmentation will be carried out using Local Adaptive Thresholding and also at the same time the noise cleaning stage using cleaning morphology. The disadvantage of this method is that in some cases pathological images, unable to handle connectivity, can lead to the inaccurate result of segmentation.

Yinling Cheng et al.[6] proposed the U-net method for the segmentation of blood vessels in the fundus retinal. The performance result of the proposed method produces a good performance but there are still deficienci8 in the slightly lower accuracy value. Tahsin Mostafiz et al.[12] proposed two efficient methods for the segmentation of blood vessels using a Fuzzy Classifier and U-net autoencoder combined with a residual block. With this method, the accuracy obtained is quite high. But on the part of precision and sensitivity are still low.

The U-Net architecture was introduced by Ronneberger et al, in 2015 to assist in the biomedical segmentation process[13]. In the U-Net architecture, there is a decoder and encoder section. U-Net architecture is built on Fully Convolutional Network (FCN) and modified to produce better segmentation in medical images. Compared to FCN, the U-net architecture is more symmetrical and the skip connection between the downsampling path and the upsampling path implements the merge operation. This skip connection aims to provide global information during the upsampling process.

Augmentation is a very important step for studying neural networks in images when the available samples are limited. The U-net architecture is a high-capacity model, therefore an augmentation stage is needed to add training data to avoid overfitting[19].

3. Proposed method

The segmentation of blood vessels is an interesting thing to study. In this study, the proposed method for segmentation of retinal vessels here is the U-net architecture of the CNN model. The step to be carried out consists of two stages, the first step is to improve the image quality by converting the retinal image to each RGB color channel and taking the extraction of the green channel which is then followed by increasing the contras and quality of the image using the Contrast Limited Adaptive Histogram Equalization (CLAHE). To reduce the noise resulting from contrast enhancement of the retinal image, a Median Filter is used for the cleaning stage. Next is the phase to increase the limited number of datasets by using rotation augmentation to obtain dataset results from various points of view. The last step that will be done is to segment the retinal blood vessels as a result of augmentation.

The following stages of the proposed method are illustrated in Figure 1.

3.1. *Pre-processing*

The step taken before extracting blood vessels is the preparation of retinal image data that will be used. The dataset to be used is retrieved from the STARE database[20]. After preparing the retinal image as input, the next steps that will be carried out are changing the color, increasing contrast, cleaning noise, and increasing the number of datasets so that later performance with high accuracy will be produced. Here are the steps being taken in the preprocessing of the retinal image.

3.1.1. *Green Channel Separation*

Color channel separation was conducted to produce the image with grayscale with the aim of the later image that has been separated color channel to be more simple and will get the result of the blood vessel image that was clean from the background image[21]. The output image from the result of this color channel separation will consist of red (R), green (G), Blue (B).

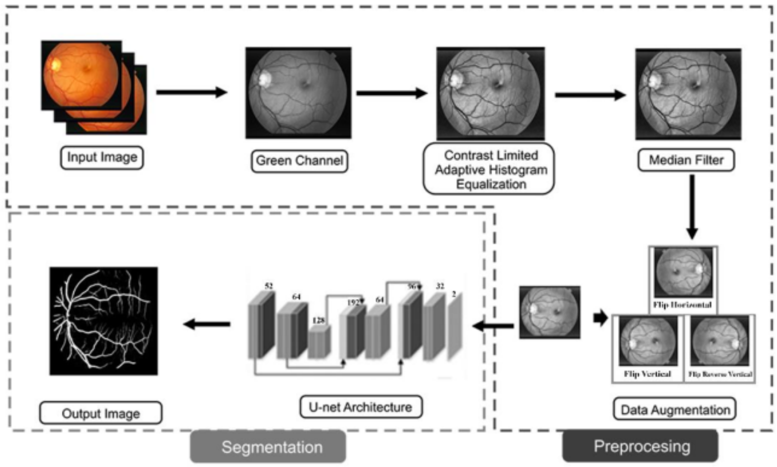


Fig.1. Flow chart of the proposed method for data augmentation on vascular retinal image segmentation

The separation step is done by separating the original image into three color channels, namely the red channel, the green channel, and the blue channel. Furthermore, the retinal blood vessel extraction process was carried out for each of the red, green, and blue color channels. The output image on the red channel appears to produce a lighter image while on the blue channel the resulting image tends to be dark in color. The green color channel was chosen because the resulting image has good contrast and quality. Vessel objects are more visible in the green color channel than in the blue and red color channels.

3.1.2. *Contrast Limited Adaptive Histogram Equalization (CLAHE)*

CLAHE was developed as an imaging process in medical terms which later will give good results in improving the quality and contrast of an image. Increasing contras can be done by using the Adaptive Histogram Equalization (AHE) and Histogram Equalization (HE) method[4]. After doing the green channel separation stage, the next step is to improve the contrast using CLAHE. In the retinal image that have been generated from the previous stages, there are still difficulties in differentiating the contrast between blood vessels and non-vessels. CLAHE was chosen as a method to increase the contrast of the retinal image so that the parts that have a low contrast improvement method from the previous method, AHE with aim of increasing contrast and minimizing noise in the image. CLAHE limits the amplification by cutting the limits of the histogram values. The values of the histogram that is also called the clip limit.

CLAHE has two parameters, namely Block Size (BS) and Clip Limit (CL). These parameters control the image quality. Determination of BS and CL is done heuristic. CLAHE is an equalization histogram that focuses on grayscale transformations on contrast enhancement. CLAHE itself is a method that is formulated based on dividing the image into several non-overlapping parts which are divided into equal sizes.

3.1.3. *Median Filter*

The median filter is a non-linear digital filter technique, which is often used to removed noise in an image. A median filter was chosen as a method to reduce noise in the image while maintaining existing details such as image edges for better feature extraction. The median filter works by calculating the middle value of the calculated result with the pixels in the image being considered [5].

The median filter is carried out in steps, namely the CLAHE image is converted into pixel values in the form of an ordered vector. Next, the center value is calculated in the specified mask environment and then replaces the center value from the calculated results with pixels in the selected image.

3.1.4. *Data Augmentation*

Augmentation is the process of increasing the size and quality of a limited number of training datasets. Augmentation is a very important step for studying neural networks in images when the available samples are limited. Because the U-net architecture is a high-capacity model, an augmentation stage is needed to add training data to avoid overfitting. The augmentation performed on the median filter retinal image is rotational augmentation which aims to increase the diversity of retinal image data. The augmentation technique is performed using a transformation matrix. Suppose (x, y) is the coordinate position of the image pixel and (x', y') is the new position of the pixel augmented with Eqs. (1), (2), and (3) for horizontal flip, vertical flip, and vertical flip, respectively.

$$flip\ horizontal = \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -x \\ y \end{bmatrix} \quad (1)$$

$$flip\ vertical = \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ -y \end{bmatrix} \quad (2)$$

$$flip\ reverse\ vertical = \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -x \\ -y \end{bmatrix} \quad (3)$$

The augmentation stages were carried out by flip horizontally, flip vertically, and flip reverse vertically for the input image used, namely 20 median filter images and 20 ground truth images. In horizontal flip, the inputted image is the image resulting from the median filter, it is flipped to the right and then saved and produces 20 horizontal images. Then on the vertical flip, the image will be flipped upwards and then saved and produces 20 vertical images. Next, flip reverse vertical with the image will be flipped up first and then flipped to the right, after that the image is saved in the form of 20 flip reverse vertical images. The number of augmented datasets will produce 60 images and the total amount of data to be used is 80 images consisting of flip vertical, flip horizontal, flip reverse vertical, and median filter images.

3.1.5. *U-net Architecture*

U-net Architecture is one of CNN's deep learning methods that is most often used in the image segmentation stage. The U-net architecture was introduced by Ref. 22 to assist in the biomedical segmentation process[13]. The U-net architecture which is divided into 3 parts, namely the contracting/down-sampling route, bottleneck, and expanding/up-sampling path

Contracting or down-sampling consist of 4 blocks and each block is composed of a 3 x 3 convolution layer plus an activation function and 2 x 2 max pooling. It should be noted that the feature maps will be doubled for each pooling, starting with 64 feature maps in the first block, 128 for the second block, and so on. Contracting aims to capture the context of the input image to the segment it. This contextual information is then transferred to the up-sampling path using a skip connection. The bottleneck is the part that between the up-sampling and down-sampling path. The bottleneck is constructed from only two convolutional layers (with batch normalization).

Expanding or up-sampling has the aim of allowing proper localization combined with contextual information from the contracting path consisting of 4 blocks and each block consisting of a deconvolution layer with two strides, merging with a feature map that is cut from the contracting path, and a 3 x 3 convolution layer plus an activation function (with batch normalization).

The U-net training process is carried out on sub-images of the previously processed image. Each patch with dimensions of 256 x 256, is obtained by randomly selecting its center in the entire data. Some of the patches outside the Field of View (FoV) will be selected, in this way, the U-net gains the knowledge to distinguish the FoV boundaries of the veins. The neural network architecture is derived from the U-net architecture. The loss function used is binary cross entropy and adam is used for optimization. The activation function in the hidden layer uses a Rectifier Linear Unit (ReLU) with a learning rate of 0.00001. The training was conducted with 500 epochs and 32 batch sizes.

After obtaining the results of segmentation using U-net architecture, the next step is to compare the results of segmentation between segmentation using augmentation and non-augmentation with the same parameters.

3.1.6. *Evaluation and Loss Function*

The next step is an important calculation are made to determine the suitability of the image based on the existing ground truth of the proposed method so that it is known how good the result of the research has been. The method used is matching based on the ground truth dataset STARE with a confusion matrix. Measurement parameters to be searched for are accuracy, sensitivity, specificity, precision, and f1-score using the following Eqs. (4), (5), (6), (7), and (8)

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP} \quad (4)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (5)$$

$$Specificity = \frac{TN}{TN + FP} \quad (6)$$

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

$$F1\text{-}score = \frac{2TP}{2TP + FP + FN} \quad (8)$$

Where:

•TP (True Positive), namely the number of positive images that are matched and classified according to the dataset correctly by the system.

•TN (True Negative), is the number of negative images that are matched and classified according to the dataset correctly by the system.

•FN (False Negative), namely the number of negative images that are matched based on the wrong dataset by the system.

•FP (False Positive), namely the number of positive images that are matched based on the wrong dataset by the system.

The loss function is very important for machine learning algorithms. This function measures the distance between the model predictor and the target value. In most cases, we will have a set of data and targets to train the algorithm on. The loss function compares the target value with the predicted value and assigns a numerical distance between them.

The loss function is used to measure the accuracy of CNN results. The loss function shows the difference between the expected value and the predicted value. The loss function that is often used for classification problems is cross-entropy. Loss function cross-entropy consists of 2 types, namely binary cross-entropy, which is usually used for classification problems between 2 classes, while categorical cross-entropy is usually used in multi-class classification problems. In general, Eq. (9) of this loss function (L) is as follows.

$$L = -\sum_i^m y_i \log p_i(x) \quad (9)$$

where m is the number of nodes in the output layer, y is the vector of the expected values, and p is the probability value for each node predicted (x) by the U-net.

Binary cross entropy compares each of the predicated probabilities to actual class output which can be either 0 or 1. It then calculates the score that penalizes the probabilities based on the distance from the expected value. In the case of segmentation with 2 classes, namely foreground and background, binary cross-entropy is a reasonable choice. In general, this algorithm is to minimize the negative log-likelihood of the dataset ($f(x)$), which is a direct measure of the prediction performance of the model using Eq. (10).

$$\begin{aligned} f(x) &= \frac{1}{m}\sum_x (target(x) - predictor(x))^2 \\ &= \frac{1}{m}\sum_x (target(x) - max(0, \sum_i^x w_i x_i + b))^2 \end{aligned} \quad (10)$$

where w is the weighted value and b is the bias value.

4. Experiments and Result

We have performed retinal image analysis and segmentation using the extracted green channel. Where the color channel separation process is divided into 3 color channels with different values for each color channel. In Figure 2.(a) as input from the initial image, Figure 2.(b) is the extraction result from the Red channel, Figure 2.(c) is the extraction result from the Green channel, Figure 2.(d) is the result of the extraction of the Blue channel. In the red channel result, the resulting output image has a color contrast that is too bright, causing undetectable parts of the blood vessels. Then in the green channel result, the resulting output image is very good and makes parts of the blood vessels in the retinal image visible. Furthermore, in the blue channel, the resulting output image has too dark a contrast and cause not all of the veins to be visible and it can be concluded that the green channel can be selected because the resulting output image is very clear compared to the other 2 color channels.

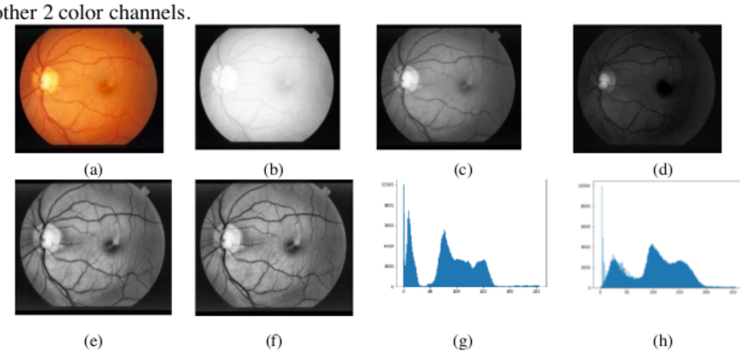


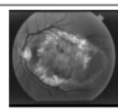
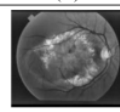
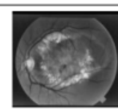
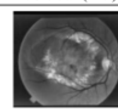
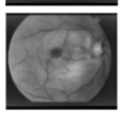
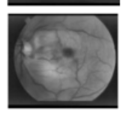
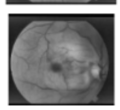
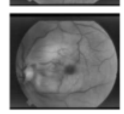
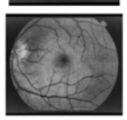
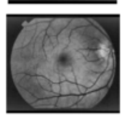
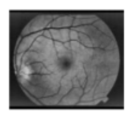
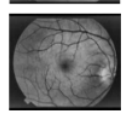
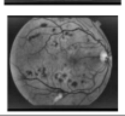
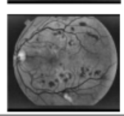
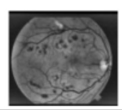
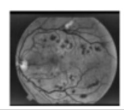
Fig.2. (a) Original Image, (b) Red Channel, (c) Green Channel, (d) Blue channel, (e) CLAHE image, (f) Image After Median Filter, (g) Histogram of the green channel and (h) Histogram of the CLAHE Image

In Figure 2. (g) is the result of the histogram before the increase in contrast and image quality, Figure 2. (e) Result of improved contrast and vascular quality in retinal images, Figure 2.(h) is a histogram of the result of increasing contrast and image quality. Then by using the CLAHE method each pixel will be increased on the contrast of each part will be strengthened and also fix the parts that should be used so that the output image histogram parameter will produce a clearer image with blood vessels as in Figures 2.(e) and (h).

The next step is cleaning the CLAHE image from noise using the median filter. At this stage, the median filter will clean the small pixel that is obtained from the result of the contrast enhancement using the previous CLAHE. Figure 2.(f) is the image resulting from the median filter. As the result of contrast enhancement using CLAHE in Figure 2.(e), the resulting output image is proven to be good, but the image still produces a lot of noise. To reduce the amount of noise, the median filter was chosen because it can increase the clarity of an image and reduce noise so that the resulting image is smoother by replacing the damaged pixel value and fixing it as in Figure 2.(f). The median filter process cleans the small pixels generated from CLAHE. The median filter works by evaluating the brightness level of a pixel and determining which pixel to increase its brightness.

After enhancing and repairing the retinal image, the next step is augmentation. Augmentation was carried out on the STARE dataset from Adam Hoover's ground truth result. Augmentation used rational augmentation consisting of flip horizontal, flip vertical, and flip reverse vertical. Where in the image augmentation dataset is the result of retinal image repair using the median filter method consisting of 20 image data, then rotated vertically to produce 20 new data, then the median filter dataset is rotated horizontally to produce 20 new data and the last dataset is rotated reverse vertical to produce 20 new datasets with the total number of augmentation result in the median filter dataset which is 80 data and can be seen in Table 1. The dataset of Adam Hoover ground truth consists of 20 black and white image data, then rotated vertically to produce 20 new datasets, then rotate back horizontally to get 20 horizontal datasets, and finally rotated reverse vertically to get 20 datasets with the total number of ground truth datasets, namely amounting to 80 data.

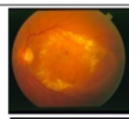
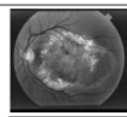
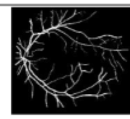
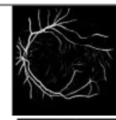
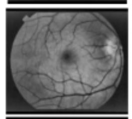
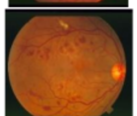
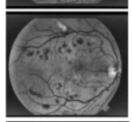
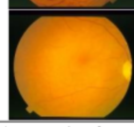
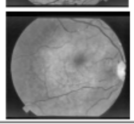
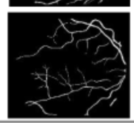
Table 1. Image Augmentation of Median Filter Result

No.	File Name	Median Filter Result	Flip Horizontal (H)	Flip Vertical (V)	Flip Reverse Vertical (RV)
1.	Im0002				
2.	Im0004				
3.	Im0081				
4.	Im0139				

The next stage is a segmentation of blood vessels from the result of previous processing. After passing through the data augmentation stage, the current dataset is 80 for image enhancement and 80 for ground truth. Furthermore, the retinal image segmentation process has previously been divide into 80% training data and 20% testing data. The segmentation result of the U-net architecture can be seen in Table 2.

Table 2. U-net Segmentation Result from STARE Dataset

No.	File Name	Original Image	Pre-processing	Ground Truth Adam Hoover	Segmentation result

1.	im0002				
2.	H_im0081				
3.	V_im0139				
4.	RV_im0319				

After obtaining the result of segmentation using the U-net architecture, the next step is to conduct an evaluation. Following are the result of retinal image segmentation using U-net architecture with epoch 500, batch size 32, and the loss function used is binary cross-entropy. In Figures 3, it can be seen that the accuracy and lost values in the model show pretty good results.

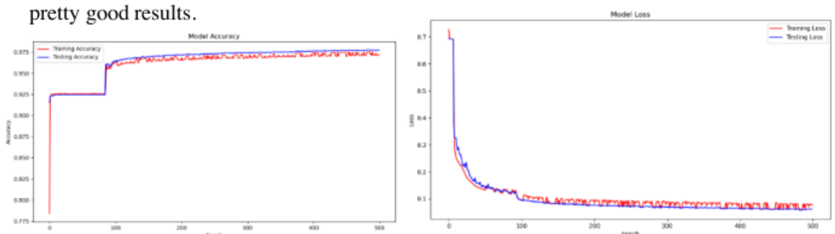


Fig.3. a. Graph of Training and Testing Process Accuracy b. Loss Graph of the training and testing process

In Figure 3.a, the epoch used is 500 and it can be seen that the increment value for training data and testing data increases regularly and there is a significant increase at epoch 90 toward 100. The result obtained on the accuracy model is from epoch 100 to 500 shows a very good result. In Figure 3.b, the loss model used is binary cross-entropy with a fair good loss result, indicated by the average results of training and testing data that drop regularly and no value increases too significantly.

Neural network training is performed on patches of preprocessed images. Each patch with the dimensions of 48x48 is obtained with the center chosen at random. Patches that are partially or completely outside the FOV are selected, in this way the neural network will learn how to distinguish the boundary of the FOV from blood vessels. A set of 190,000 patches was obtained by randomly extracting 9,500 patches on each image. Although patches overlap, i.e. different patches may contain the same portion of the original image, no further data addition is performed. Of the first 90% of the dataset used for training 171,000 patches, while the last 10% used for validation 19,000 patches.

To improve performance, the vein probability of each pixel is obtained by averaging some predictions. With a height and width of 5 pixels, several consecutive overlapping patches were extracted in each test image. Then, for each pixel, the vascular probability is obtained by averaging the probabilities on all the predicted patches covering the pixel.

In this Convolutional Neural Network, the ReLU (Rectified Linear Unit) activation function is used. The activation function is used to determine whether the result of the network calculation is true or false. The ReLU activation function belongs to the category of non-linear activation function. ReLU is the most frequently used function in almost all deep learning. Computing using the ReLU activation function is more efficient and faster than the sigmoid activation function. The sigmoid activation function is more appropriate when used for models that predict probability as the output, because the output is 0 or 1. However, the output is not zero centered, making it difficult to optimize. The exponential on the sigmoid activation function can also be said to be compute expensive, expensive in resources or long in computation.

Based on the result of these parameters accuracy, sensitivity, specificity, precision, and f1-score of the segmentation of retinal vessel using the STARE dataset are obtained:

Table 3. Parameter of Accuracy, sensitivity, specificity, precision, and f1-score

Methods	Accuracy	Sensitivity	Specificity	Precision	F1-score
Augmentation	97.64%	78.18%	99.20%	88.77%	82.91%
Non Augmentation	95.98%	65.73%	98.97%	86.09%	73.86%

Based on Table 3, it can be seen that evaluation of the result parameter from the STARE dataset using Adam Hoover's ground truth includes accuracy, sensitivity, specificity, precision, and f1-score. The value obtained from each result parameter is 97.64% for accuracy, 78.18% for sensitivity, 99.20% for specificity, 88.77% for precision, and 82.91% for F1-score. From the result that has been obtained, it can be seen that the calculated result for the accuracy and specificity parameters of the proposed method produces a relatively high value because the resulting input image can be detected properly by the system. Then the precision and f1-score get a pretty good results with quite high results. The value that still needs to be improved is the sensitivity value which is still relatively low.

The application of augmentation improves segmentation performance for accuracy, sensitivity, specificity, precision, and F1-score by 1.63%, 12.45%, 0.23%, 2.7%, and 9.06%, respectively, compared to the results of non-augmentation segmentation.

The results of comparisons made by previous studies can be seen in Table 4. The proposed method outperforms the methods proposed by other researchers in the parameters of accuracy, specificity, precision, and F1-score. The following is a comparison table between the previously proposed method and other methods used by previous researchers.

Table 4. Comparison of Segmentation result from the proposed method with the previous method

Method	Accuracy	Sensitivity	Specificity	Precision	F1-score
K.B. Khan et al[4]	95.7%	74.5%	97.4%	-	-
T.A. Soomro et al[14]	96.8%	84.8%	98.6%	-	-
T. Mostafiz et al[12]	95.37%	55.82%	98.62%	86.22%	-
I. Atli et al[8]	96.89%	79.87%	98.54%	-	-
Proposed Method	**97.64%**	**78.18%**	**99.20%**	**88.77%**	**82.91%**

The proposed approach will be further developed in the future, to increase the value of the blood vessel segmentation process evaluation outcomes. The results of the measurement of accuracy and specificity parameters get a high value compared to other researchers. However, the sensitivity is still low because there is a ratio between the retinal blood vessel pixels that are classified correctly and the ground truth is not good and needs to be improved.

5. Conclusion

In this study, the blood vessel segmentation of the retinal image was carried out using U-net architecture assisted by the data augmentation method using the STARE dataset and the result obtained were 97.64% accuracy, 78.28% sensitivity, 99.20% specificity, 88.77% precision, and f1-score 82.91%. These results indicate that the method that has been proposed produces more advantages than the method proposed by the other researchers. The combined use of the Contras Limited Adaptive Histogram Equalization (CLAHE) and Median Filter produces a good output image in terms of enhancing the contrast and quality of the blood vessels.

The augmentation method is very helpful in the pre-processing stage before performing segmentation using U-net architecture. Because with augmentation, the training process can be maximized and avoid overfitting. The result shown, it explains that the U-net architecture with the proposed method can be used in the medical world, especially in the field of blood vessel segmentation. For future work, it is hoped that the result of the parameters of sensitivity can be increased so that later the result obtained will be better.

Acknowledgments

This article is partly supported by the Directorate of Research and Community Service, the Directorate General of Strengthening Research and Development and the Rector of the University of Sriwijaya.

References

1. S. Shahbeig and M. Helfroush. A Novel and Efficient Method to Extract Blood Vessels from Retinal Images. *Bull la Société R des Sci Liège*. 2016;85:139-151. doi:10.25518/0037-9565.5759
2. X. Gao, Y. Cai, C. Qiu, and Y. Cui. Retinal blood vessel segmentation based on the Gaussian matched filter and U-net. *Proc - 2017 10th Int Congr Image Signal Process Biomed Eng Informatics, CISP-BMEI 2017*. 2018;2018-Janua:1-5. doi:10.1109/CISP-BMEI.2017.8302199
3. N. Eladawi, M. Elmogy, O. Helmy , et al. Automatic blood vessels segmentation based on different retinal maps from OCTA scans. *Comput Biol Med*. 2017;89(August):150-161. doi:10.1016/j.compbiomed.2017.08.008
4. K. B. Khan, A.A. Khaliq, M. Shahid, and S. Khan. An efficient technique for retinal vessel segmentation and denoising using modified isodata and CLAHE. *IIUM Eng J*. 2016;17(2):31-46. doi:10.31436/iiumej.v17i2.611
5. J. Dash and N. Bhoi. A thresholding based technique to extract retinal blood vessels from fundus images. *Futur Comput Informatics J*. 2017;2(2):103-109. doi:10.1016/j.fcij.2017.10.001
6. Y. Cheng, M. Ma, L. Zhang, C.J. Jin, L. Ma, and Y. Zhou. Retinal blood vessel segmentation based on densely connected U-Net. *Math Biosci Eng*. 2020;17(4):3088-3108. doi:10.3934/MBE.2020175
7. T.A. Soomro, A.J. Afifi, A. Ali Shah, et al. Impact of Image Enhancement Technique on CNN

Model for Retinal Blood Vessels Segmentation. *IEEE Access*. 2019;7:158183-158197. doi:10.1109/ACCESS.2019.2950228

8. I. Atli and O.S. Gedik. Sine-Net: A fully convolutional deep learning architecture for retinal blood vessel segmentation. *Eng Sci Technol an Int J*. 2021;24(2):271-283. doi:10.1016/j.jestch.2020.07.008
9. J. Dash and N. Bhoi. An Unsupervised Approach for Extraction of Blood Vessels from Fundus Images. *J Digit Imaging*. 2018;31(6):857-868. doi:10.1007/s10278-018-0059-x
10. S. Dutta, B. C. S. Manideep, S. M. Basha, R. D. Caytiles, and N.C.S. N Iyengar. Classification of diabetic retinopathy images by using deep learning models. *Int J Grid Distrib Comput*. 2018;11(1):89-106. doi:10.14257/ijgdc.2018.11.1.09
11. S. Wan, Y. Liang, and Y. Zhang. Deep convolutional neural networks for diabetic retinopathy detection by image classification. *Comput Electr Eng*. 2018;72:274-282. doi:10.1016/j.compeleceng.2018.07.042
12. T. Mostafiz, I. Jarin, S. A. Fattah, and C. Shahnaz. Retinal Blood Vessel Segmentation Using Residual Block Incorporated U-Net Architecture and Fuzzy Inference System. *2018 IEEE Int WIE Conf Electr Comput Eng WIECON-ECE 2018*. 2018;(December):106-109. doi:10.1109/WIECON-ECE.2018.8783182
13. N. Sambyal, P. Saini, R. Syal, and V. Gupta. Modified U-Net architecture for semantic segmentation of diabetic retinopathy images. *Biocybern Biomed Eng*. 2020;40(3):1094-1109. doi:10.1016/j.bbe.2020.05.006
14. T. A. Soomro, A. J. Afifi, J. Gao, O. Hellwich, L. Zheng, and M. Paul. Strided fully convolutional neural network for boosting the sensitivity of retinal blood vessels segmentation. *Expert Syst Appl*. 2019;134:36-52. doi:10.1016/j.eswa.2019.05.029
15. S. Vitale, G. Ferraioli, and V. Pascazio. Multi-objective cnn based algorithm for sar despeckling. *arXiv*. Published online 2020:1-13.
16. S. Wang, Y. Yin, G. Cao, B. Wei, Y. Zheng, and G. Yang, "Hierarchical retinal blood vessel segmentation based on feature and ensemble learning," *Neurocomputing*, vol. 149, no. PB, pp. 708–717, 2015, doi: 10.1016/j.neucom.2014.07.059.
17. S. Albawi, T. A. M. Mohammed, and S. Alzawi, "Layers of a Convolutional Neural Network," *Ieee*, 2017.
18. A. Dasgupta and S. Singh, "a Fully Convolutional Neural Network Based Structured Prediction," *2017 IEEE 14th Int. Symp. Biomed. Imaging (ISBI 2017)*, pp. 248–251, 2017.
19. L. Luo, D. Chen, and D. Xue, "Retinal blood vessels semantic segmentation method based on modified U-Net," *Proc. 30th Chinese Control Decis. Conf. CCDC 2018*, pp. 1892–1895, 2018, doi: 10.1109/CCDC.2018.8407435.
20. A. Hoover, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," *IEEE Trans. Med. Imaging*, vol. 19, no. 3, pp. 203–210, 2000, doi: 10.1109/42.845178.
21. A. Vyas, S. Yu, and J. Paik. *Fundamentals of Digital Image Processing*.; 2018. doi:10.1007/978-981-10-7272-7_1
22. Ronneberger O., Fischer P., Brox T. "U-Net: Convolutional Networks for Biomedical Image Segmentation". In: Navab N., Hornegger J., Wells W., Frangi A. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science, 2015, vol 9351. Springer, Cham. https://doi.org/10.1007/978-3-319-24574-4_28

THE AUGMENTATION DATA OF RETINE IMAGE FOR BLOOD VESSEL SEGMENTATION USING U-NET CONVOLUTIONAL NEURAL NETWORK METHOD

ORIGINALITY REPORT

13%	10%	12%	5%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	researchbank.swinburne.edu.au Internet Source	3%
2	Erwin, Heranti Reza Damayanti. "Supervised Retinal Vessel Segmentation Based Average Filter and Iterative Self Organizing Data Analysis Technique", International Journal of Computational Intelligence and Applications, 2020 Publication	1%
3	Nabila Eladawi, Mohammed Elmogy, Omar Helmy, Ahmed Aboelfetouh, Alaa Riad, Harpal Sandhu, Shlomit Schaal, Ayman El-Baz. "Automatic blood vessels segmentation based on different retinal maps from OCTA scans", Computers in Biology and Medicine, 2017 Publication	1%
4	"Proceedings of the International Conference on ISMAC in Computational Vision and Bio-	1%

Engineering 2018 (ISMAC-CVB)", Springer Science and Business Media LLC, 2019
Publication

5	github.com Internet Source	1 %
6	medium.com Internet Source	1 %
7	towardsdatascience.com Internet Source	1 %
8	jist.ir Internet Source	1 %
9	www.analyticsvidhya.com Internet Source	1 %
10	"Next Generation of Internet of Things", Springer Science and Business Media LLC, 2021 Publication	1 %
11	Imania Ayu Anjani, Yulinda Rizky Pratiwi, S. Norfa Bagas Nurhuda. "Implementation of Deep Learning Using Convolutional Neural Network Algorithm for Classification Rose Flower", Journal of Physics: Conference Series, 2021 Publication	1 %
12	R. GeethaRamani, Lakshmi Balasubramanian. "Retinal blood vessel segmentation employing	1 %

image processing and data mining techniques for computerized retinal image analysis", Biocybernetics and Biomedical Engineering, 2016

Publication

Exclude quotes	On	Exclude matches	< 1%
Exclude bibliography	On		