IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received December 23, 2021, accepted February 5, 2022, date of publication February 15, 2022, date of current version February 28, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3151644

A Deep Learning Approach for the Detection of Neovascularization in Fundus Images Using Transfer Learning

MICHAEL CHI SENG TANG[1], SOO SIANG TEOH[1], (Senior Member, IEEE), HAIDI IBRAHIM[1], (Senior Member, IEEE), AND ZUNAINA EMBONG[2]

[1]School of Electrical and Electronic Engineering, Engineering Campus, Universiti Sains Malaysia, Nibong Tebal 14300, Malaysia
[2]Department of Ophthalmology, School of Medical Sciences, Health Campus, Universiti Sains Malaysia, Kubang Kerian 16150, Malaysia

Corresponding author: Soo Siang Teoh (eeteoh@usm.my)

This work was supported by the Ministry of Higher Education Malaysia for Fundamental Research Grant Scheme with Project Code FRGS/1/2019/TK04/USM/02/7.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Ethics Committee of Universiti Sains Malaysia, Malaysia, under Protocol No. USM/JEPeM/20020118, and performed in line with the guidelines of the Declaration of Helsinki.

ABSTRACT Patients with diabetes are at risk of developing a retinal disorder called Proliferative Diabetic Retinopathy (PDR). One of the main characteristics of PDR is the development of neovascularization, a condition in which abnormal blood vessels are formed on the retina. This condition can cause blindness if it is not detected and treated early. Numerous studies have proposed different image processing techniques for detecting neovascularization in fundus images. However, because of its random growth pattern and small size, neovascularization remains challenging to detect. Hence, deep learning techniques are becoming more prevalent in neovascularization identification because of their ability to perform automatic feature extraction on objects with complex features. In this paper, a method of neovascularization detection based on transfer learning is proposed. The performance of the transfer learning method is investigated using four pre-trained Convolutional Neural Network (CNN) models, which include AlexNet, GoogLeNet, ResNet18, and ResNet50. In addition, an improved network based on the combination of ResNet18 and GoogLeNet is proposed. Evaluation on 1174 retinal image patches showed that the proposed network could achieve 91.57%, 85.69%, 97.44%, and 97.10% of accuracy, sensitivity, specificity, and precision, respectively. We demonstrated that the proposed method outperforms each individual CNN for neovascularization detection. It also shows better performance compared to another method that utilized deep learning models for feature extraction and Support Vector Machine (SVM) for classification.

INDEX TERMS Neovascularization detection, deep learning, convolutional neural networks, biomedical image processing, proliferative diabetic retinopathy.

I. INTRODUCTION

Diabetic Retinopathy (DR) is more prevalent in patients with long-term diabetes [1]. It is categorized into Non-proliferative DR (NPDR) and Proliferative DR (PDR). Patients with NPDR will have several clinical symptoms such as microaneurysms, hemorrhages, hard exudates, and cotton wool spots [2]. PDR is the advanced stage of DR, and it carries a significant risk of vision impairment [3]. This condition is caused by the development of small and irregular blood vessels in the retina, a process called neovascularization [4]. One of the primary causes of aberrant and fragile blood vessel growth is a lack of oxygen delivery in the blood vessels [5]. The newly formed vessels are delicate and can easily burst, resulting in retinal bleeding. If these new blood vessels are formed within the diameter of the optic disk, the condition is referred to as neovascularization at the optic disk (NVD). On the other hand, neovascularization elsewhere (NVE) refers to the new vessels forming one disk diameter away from the optic disk. Both NVD and NVE are equally

The associate editor coordinating the review of this manuscript and approving it for publication was Paulo Mendes.

blamed for vessel growth and vitreous hemorrhage, resulting in visual loss. Therefore, a referral to an ophthalmologist is necessary when neovascularization occurs, whether NVD or NVE.

PDR must be detected early to preserve the patient's vision. This can be accomplished by analyzing the patient's fundus image to detect blood vessels and identifying the newly formed vascular associated with neovascularization. Numerous techniques for segmenting blood vessels have been proposed [6]–[15], but detecting neovascularization remains difficult. The retinal vasculature is a visible circulatory system in the eye that provides valuable information about the body's microcirculation without the need for invasive procedures [16]. Effective computer-aided diagnosis algorithms may improve the accuracy and sensitivity of neovascularization identification during frequent follow-up visits or telemedicine consultations. If detection were more accurate, patients would be less likely to miss out on early and effective laser therapy. In comparison to microaneurysms, the shape and size of neovascularization vary, posing additional challenges and highlighting the importance of developing automated detection methods [17].

Multiple studies have demonstrated that image processing algorithms can automatically identify microaneurysms, hemorrhages, hard exudates, and cotton wool spots. However, research into detecting neovascularization is still in its infancy due to the difficulty of distinguishing between normal blood vessels and new blood vessels that have formed. Additionally, the number of labeled neovascularization images is limited, impeding the field's advancement. A comprehensive retinal image may be obtained using angiography-based techniques. However, due to the invasive nature of these procedures, they are generally not recommended, particularly for early-stage or routine diagnosis [18].

This paper proposed a deep learning approach for neovascularization detection based on transfer learning. A network based on the combination of ResNet18 and GoogLeNet is proposed. These two networks are combined using a depth concatenation layer. The performance of the combined network is compared to that of the original pre-trained networks, which include AlexNet, GoogleLeNet, ResNet18, and ResNet50. Additionally, we conducted experiments to evaluate the transfer learning results and determine the method's efficacy in detecting neovascularization. We demonstrated that the proposed network (ResNet18 + GoogLeNet combination) could outperform other pre-trained networks in detecting neovascularization through transfer learning.

This paper is organized as follows: Section II summarizes several previous studies on the detection of neovascularization. Section III explains the methodology and the proposed transfer learning method for classifying neovascularization. Section IV presents the evaluation results and performance comparison with other deep learning methods. Finally, a conclusion is given in Section V.

II. RELATED WORKS

Neovascularization lesions usually have complex features. They look like entangled tiny vessels and are challenging to detect because of their random pattern of growth. Furthermore, the blood vessel responsible for the lesion is typically as small as a single-pixel wide. Additionally, due to the scene's erratic lighting, the neovascularization becomes entangled with the background image. Typical image processing techniques used to recognize the complex neovascularization features are based on traditional machine learning and deep learning methods. While some researchers have achieved promising results in detecting neovascularization, their proposed methods continue to have some limitations.

For example, in a method proposed by Gandhimathi *et al.* [19], the blood vessels are segmented first using the Fuzzy C-means clustering technique. Then, neovascularization vessels are detected using morphological and threshold techniques. Their proposed method can identify whether a patient is at high risk of having neovascularization. However, the technique produced a very low specificity. On the other hand, Coelho *et al.* [20] introduced a neovascularization detection technique near the optic disk by measuring the angular spread of the Fourier power spectrum of the image's gradient magnitude. Based on the computed measures, they used a linear classifier to detect neovascularization at the optic disk. However, the neovascularization elsewhere is not investigated. Kar *et al.* [21] suggested that the vessel thickness can identify abnormal vessels, but false detection may occur when other tiny lesions are present inside a fundus image.

A paper by Lee *et al.* [22] introduced an automated neovascularization detection system using statistical texture analysis (STA), high order spectrum analysis (HOS), and fractal analysis (FA) with reasonable accuracy. However, their proposed system cannot grade the severity of the disease. Saranya *et al.* [23] segmented the blood vessels from fundus images using the Fuzzy C Means Clustering technique. The features based on shape, brightness, position, and contrast are then extracted from the segmented images. These features are then used to classify the segmented images as normal or abnormal using K-Nearest Neighbour. However, this technique is incapable of determining the location of the abnormal vessels. It is only capable of determining whether a fundus image is normal or abnormal.

Goatman *et al.* [24] described a method for detecting neovascularization lesions on the optic disk. They extracted 15 neovascularization features using watershed lines and ridge strength measurement and trained a Support Vector Machine (SVM) to identify neovascularization vessels in the optic disk. However, this method is only designed to detect neovascularization on the optic disk (NVD). Detection of neovascularization vessels outside the optic disk field (NVE) was not investigated.

Convolutional neural networks (CNN) are widely used for medical image classification due to the advancement of deep

FIGURE 1. The flowchart of the methodology.

learning in computer vision technology. Setiawan *et al.* [25] recently published a study that used several pre-trained convolutional neural networks to extract the image features of neovascularization and perform classification using SVM. They managed to demonstrate that the feature extraction method through deep learning models can yield favorable results. Carrillo-Gomez *et al.* [26] described a method for detecting NVD using a deep learning algorithm. They evaluated several neural networks for their ability to detect NVD. DenseNet-161 and Efficientnet-B7 are two of these networks. Their experiment demonstrated that both of these networks are capable of detecting NVD with high accuracy and sensitivity. Abu Hassan *et al.* [27] published a paper in which they developed a CNN for detecting PDR in fundus images. Their CNN achieves an accuracy of 73.81%, a sensitivity of 76%, and a specificity of 69%, respectively. A deep learning-based semantic segmentation technique has also been applied for neovascularization detection. In [28], a semantic segmentation convolutional neural network is used to detect the position of neovascularization in fundus images.

While several deep learning methods have been proposed for neovascularization detection, a method based on transfer learning remains unexplored. Transfer learning is a technique in which an already-trained deep neural network is adapted to detect a new object class. To our knowledge, transfer learning has not been thoroughly investigated to identify neovascularization. This paper assesses the performance of the transfer learning approach using several pre-trained CNN for detecting neovascularization. Additionally, a method for improving the transfer learning results based on the combination of two pre-trained networks is proposed.

III. METHODOLOGY

Fig. 1 shows the flowchart of the overall methodology of this study. First, a set of fundus images with neovascularization is collected. The images are pre-processed and divided into patches suitable for network training. The detail of the data preparation is given in Subsection A. Next, several pre-trained CNN were evaluated for neovascularization detection based on transfer learning. The pre-trained

networks include AlexNet [29], GoogLeNet [30], ResNet18, and ResNet50 [31]. The implementation and training of the networks are explained in Subsection B. Subsequently, in Subsection C, the proposed method based on the combination of ResNet18 and GoogLeNet using transfer learning is presented. The networks are trained using the prepared dataset, and their performance was evaluated using several metrics. The results are then compared to another deep learning method based on CNN feature extraction and SVM classification as proposed by Setiawan *et al.* [25]. The metrics and the performance comparison are explained in Subsection D.

A. DATA PREPARATION

The fundus image datasets used in this study are Messidor [32], Diaretdb0 [33], and a private dataset collected from the Department of Ophthalmology at Universiti Sains Malaysia's Hospital. Ethical approval was obtained from the Univesiti Sains Malaysia's Ethics Committee to collect and use the patients' fundus images. A total of fifty-three images with neovascularization were collected for use in this study. The images are in RGB color format with resolutions of 1488 × 2240 pixels (Messidor), 1152 × 1500 pixels (Diaretdb0), and 2000 × 2368 pixels (private dataset).

Next, the ground truth images for neovascularization were created with the help of an ophthalmologist. Fig. 2(a) shows a raw fundus image that contains neovascularization. The raw fundus image must be enhanced first to improve the visibility of blood vessel structure and facilitate the ophthalmologist to identify the neovascularization regions on the images. This is done by extracting the green channel from the RGB fundus image. The green channel is used because it shows the blood vessels more clearly [34], [35]. The green channel image is then enhanced using Contrast Limited Adaptive Histogram Equalization (CLAHE) [36]. Fig. 2(b) shows an image that has been pre-processed using green channel extraction and CLAHE. Following that, the ophthalmologist labeled regions of neovascularization, as shown in Fig. 2(c). Next, the labels are used to generate the ground truth image, as shown in Fig. 2(d). The ground truth images will be used as the reference for determining whether an area within an image contains neovascularization.

For the networks' training, the 3-channel images are used without any pre-processing. This is to allow the networks to learn all the available features in the raw images. First, the fundus images are split into smaller patches so that they can be processed by the networks. Due to the difference in image resolution, images from the different datasets are split into a different number of patches. For Diaretdb0, each fundus image is divided into 30 patches, while for the Messidor and the self-collected neovascularization images, each image is split into 60 and 80 patches, respectively. A total of 2980 image patches were obtained from the 53 fundus images. In the first layer, GoogLeNet, ResNet18, ResNet50, and the proposed network (ResNet18 + GoogLeNet) use an input image resolution of 224 × 224 pixels. On the other hand, the first layer of AlexNet needs a minimum input size of 227 × 227 pixels. Due to the different image sizes required, two sets of data with the appropriate sizes are created by resizing the image patches.

There are 571 image patches containing neovascularization (Neo), while 2409 image patches do not contain neovascularization (NotNeo). This imbalanced data may affect the network's performance. To achieve a balance between Neo and NotNeo, image augmentation was performed on the 571 Neo image patches. This is done by rotating the image patches by 90 degrees three times, resulting in an additional 1713 Neo patches (571 extra Neo image patches per rotation). This resulted in a total of 2284 Neo image patches. As a result, the total image patches become 4693 with more evenly distributed image data (2284 Neo image patches and 2409 NotNeo image patches). Fifty percent of these patches are used as a training set, 25 percent for a validation set, and 25 percent for a testing set. Fig. 3 illustrates several image patches that were used to train and test the deep learning models.

B. MODIFICATION ON THE PRE-TRAINED CNN FOR NEOVASCULARIZATION DETECTION

In this study, four popular pre-trained networks (AlexNet, GoogLeNet, ResNet18, and ResNet50) are modified to train on the prepared dataset for neovascularization detection. The four CNNs were pre-trained on images from ImageNet [37]. Using transfer learning, these networks can be retrained with a new dataset to adapt the network to detect new objects such as neovascularization. The advantage of transfer learning is that it can avoid training a machine learning model from scratch, which will take a long time and requires many training images. In this study, transfer learning is used to detect neovascularization to circumvent this lengthy training process and limited training images. These pre-trained networks will no longer require weight adjustment from scratch. Instead, it can use the pre-trained weights obtained from the ImageNet's training and adjust them to fit the neovascularization dataset's images. This method can save a lot of time while still achieving excellent results.

To begin, all the pre-trained networks (AlexNet, GoogLeNet, ResNet18, and ResNet50) are modified to allow for training using the prepared images to perform two-class classification. This is accomplished by removing the final fully connected layer and replacing it with a new one that contains two defined classes: neovascularization (Neo) and non-neovascularization (NotNeo).

For instance, in ResNet18, a fully connected layer exists in the architecture's rear part. The network's original fully connected layer is omitted. This is because this fully connected layer was previously trained to classify multiple classes. It cannot be used to perform binary classification. A new fully connected layer is created with only two output classes. This new fully connected layer is then used in place of the previous one in ResNet18. Through this approach, the ResNet18 has been modified to perform binary classification.

The four modified pre-trained models are then trained using the two prepared datasets (227 × 227 pixels for AlexNet

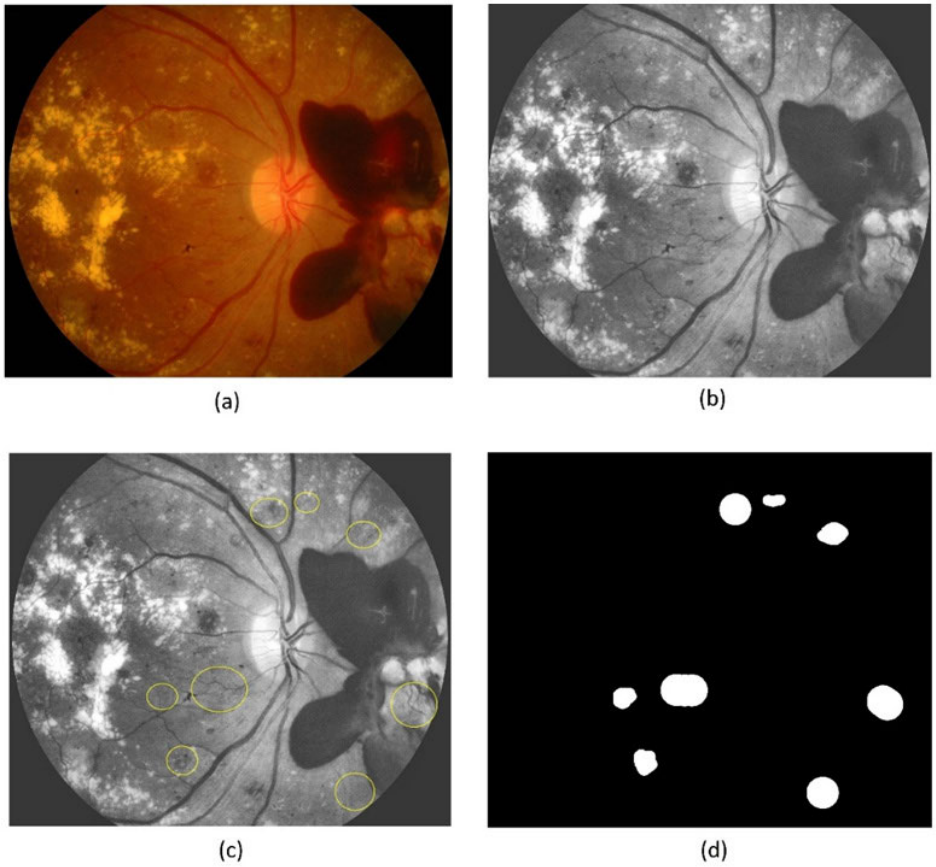


FIGURE 2. The procedures for image pre-processing and the labeling of ground truth. (a) A raw fundus image. (b) The image after green channel extraction and contrast enhancement. (c) The labeled image. The yellow circles are the neovascularization regions labeled by the ophthalmologist. (d) The generated ground truth image.

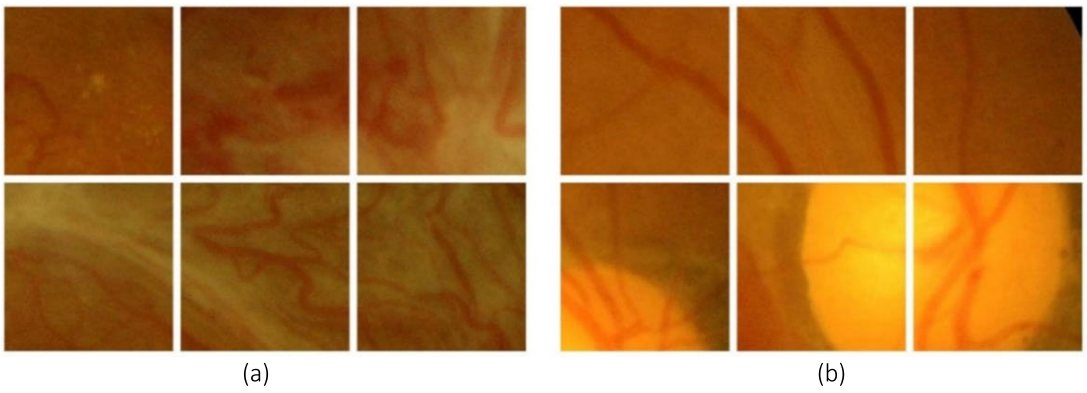


FIGURE 3. Samples of cropped image patches used in the networks' training. (a) Positive samples - image patches that contain neovascularization. (b) Negative samples - image patches without neovascularization.

and 224 × 224 pixels for the other networks). The optimal values of mini-batch size and learning rates for each network are also determined to ensure that the networks learn the neovascularization features to the best of their ability. Fig. 4 depicts a flowchart of the pre-trained networks' transfer learning process.

C. THE PROPOSED NETWORK

A network based on the combination of Resnet18 and GoogLeNet is proposed. This network combines the ResNet18 and GoogLeNet architectures, allowing for the detection of neovascularization using the outputs of both networks. A depth concatenation layer is used to join the two

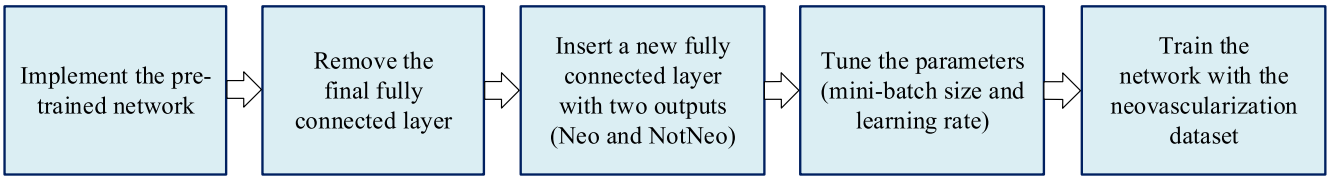


FIGURE 4. The flowchart of the pre-trained networks' transfer learning process.

networks. ResNet18 and GoogLeNet were chosen as they are smaller networks. By combining two smaller networks, computation power can be conserved, as larger networks have more parameters to train. Moreover, ResNet18 and GoogLeNet were combined because they both requires input images of the same size (224×224). Unlike AlexNet, despite its small network architecture, it requires input images with a resolution of 227×227, making it unsuitable for combining with ResNet18 or GoogLeNet. Other networks with similar input sizes, such as ResNet50, are too large. When ResNet50 is combined with GoogLeNet or ResNet18, a massive network is created that requires a lot of computation power to train. As a result, the only networks suitable for combination are ResNet18 and GoogLeNet. Fig. 5 depicts the proposed network's transfer learning approach.

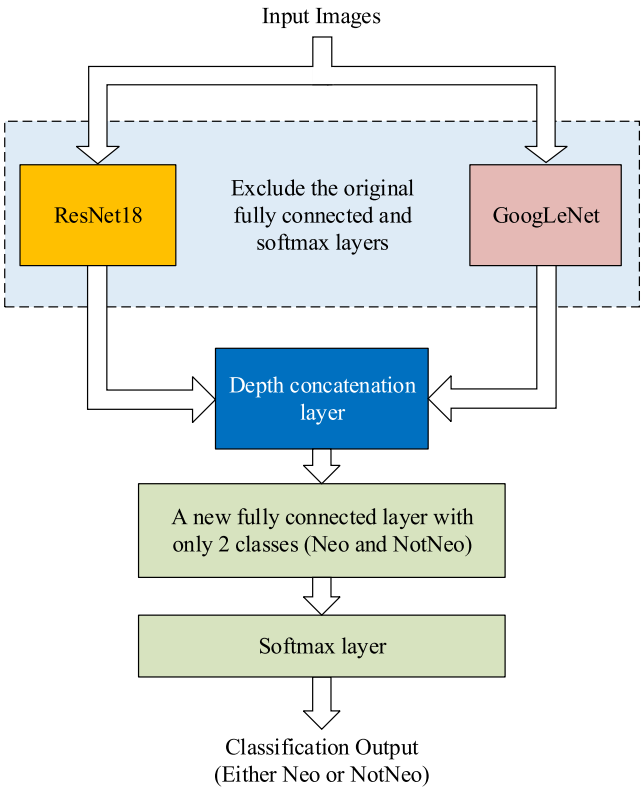


FIGURE 5. The proposed transfer learning approach based on the combination of ResNet18 and GoogLeNet.

The layer before the classification layer for Resnet18 is a Global Average Pooling layer with size $1 \times 1 \times 512$, while for GoogLeNet, it is a Dropout Layer with size $1 \times 1 \times 1024$. In the proposed network, the ResNet18's Global Average Pooling output is combined with the output of GoogLeNet's Dropout Layer using a depth concatenation layer. This results in an output with a size of $1 \times 1 \times 1536$. This output is connected to a newly created fully connected layer with only two outputs representing two new classes (Neo and NotNeo). This combined network is then trained to recognize the features of neovascularization using the prepared dataset.

The learning rate and mini-batch size for the network training are adjusted until their optimal values are discovered. The network is then trained using both the training and validation sets. During training, the network makes predictions on the validation set during each mini batch (a subset of all training image patches used in one iteration) and measures the error produced. The cross-entropy loss function is used to calculate this error. The measured error is then used to fine-tune the weights to obtain the best possible prediction. Stochastic Gradient Descent with Momentum (SGDM) is used as an optimizer to accelerate the process of determining the optimal weights. After training, the performance of the networks is evaluated using the testing set.

D. PERFORMANCE METRICS

The proposed network is compared to other pre-trained networks (AlexNet, GoogLeNet, ResNet18, and ResNet50) that also employ transfer learning. Additionally, the transfer learning results for all networks are compared to Setiawan *et al.*'s [25] method that used CNN for feature extraction and SVM for classification. In implementing Setiawan's method, we first combined the training and validation sets and fed them into a pre-trained network for feature extraction. After extracting the features, an SVM classifier was trained using the extracted features. The trained classifier was then used to classify the images in the testing set for neovascularization detection. This procedure was repeated for all the pre-trained convolutional neural networks.

The metrics used to evaluate the models' performance are accuracy, sensitivity, specificity, and precision. These performance metrics are calculated by comparing the classified patches to their ground truths. For instance, if an image patch is classified as having neovascularization when the ground truth indicates that it does not, this is a false positive detection. When a model classifies an image patch as having no neovascularization, but the ground truth indicates that it does, this is referred to as a false negative detection. True positive

and true negative detections indicate that a model classified an image patch correctly. When a specific model is used to classify the images in the testing set, the number of true positives, true negatives, false positives, and false negatives is determined. These values are then used to calculate the performance metrics, which quantify a particular model's accuracy, sensitivity, specificity, and precision in detecting neovascularization in the testing set. Accuracy is the number of correctly classified cases divided by the total number of instances. The following is the accuracy equation:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

True positive (TP) is the patches that have been correctly categorized as Neo. The patches that are accurately categorized as NotNeo are referred to as true negative (TN). The NotNeo patches that are mistakenly detected as Neo are labeled as false positives (FP). The Neo patches that are wrongly categorized as NotNeo are labeled as false negative (FN).

Aside from that, sensitivity can be used to evaluate the performance of a suggested algorithm. It measures the propensity of accurately categorized cases. The following is the equation for sensitivity:

$$Sensitivity = \frac{TP}{TP + FN} \tag{2}$$

Specificity is another useful performance parameter. It assesses the likelihood of accurately categorized negative events. The following is the specificity equation:

$$Specificity = \frac{TN}{TN + FP} \tag{3}$$

Precision is determined by comparing the number of correctly identified positive samples to the total number of detected positive samples. It is a metric that indicates how accurate a model is in classifying a positive sample. The equation for precision is as follows:

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

IV. RESULTS AND DISCUSSIONS

This section is divided into three subsections. Subsection A discusses the networks' parameter tuning. Then, Subsection B presents the evaluation results based on the testing set. Finally, Subsection C discusses the performance comparison.

A. PARAMETER TUNING

The learning rate and mini-batch size are tuned for each network to determine their optimum values so that the networks perform the best at identifying Neo and NotNeo image patches.

1) LEARNING RATE

For backpropagation learning, the learning rate is an essential parameter that controls the update step of learnable weights [38]. When the learning rate is too fast, gradient descent may increase rather than decrease the training error. On the other hand, using a slow learning rate may result in sluggish training and consistently high training errors. As a result, determining the optimal learning rate is critical for optimizing the search for the smallest point of loss in backpropagation learning.

We followed the method given in [39] to determine the optimum learning rate by starting with a larger rate and decreasing it by 0.1 until the optimum learning rate was discovered. The initial learning rate was set to 0.1 in this experiment, and the mini-batch size was set to 32. Then, the learning rate was decreased from 0.1 to 0.01, 0.001, 0.0001, and 0.00001 to see which one of them performs the best. This is done to determine the optimal learning rate. This process is applied to all the convolutional neural networks (the proposed network and the four pre-trained networks) to determine the optimal learning rates for each of them. Fig. 6 illustrates the accuracy obtained (from the validation set) with various learning rates for each pre-trained network.

According to Fig. 6, the accuracy of all pre-trained networks is greatest when the learning rate is 0.001. This demonstrates that the optimal learning rate is 0.001. This optimum learning rate was determined after testing the trained networks on the validation set. The validation set is balanced in terms of classes. In other words, the validation set contains an equal number of Neo and NotNeo images. Thus, the highest accuracy obtained from the validation set (when a learning rate of 0.001 is used) indicates that the networks did the best at differentiating Neo and NotNeo image patches.

2) MINI-BATCH SIZE

While training the pre-trained networks, the training set will be divided into smaller mini batches. These mini batches include a restricted number of training samples. The mini-batch size controls the accuracy of the error gradient estimation during network training. The error gradient is used to update the weights of the networks, and the process is repeated. According to the study in [40], using an overly large mini-batch size may significantly degrade the trained network's quality due to a lack of generalization capacity, causing it to converge to a sharp minimum. Consequently, the optimum mini-batch size must be established to optimize the convergence rate and stability of the network training [38].

Experiments were conducted to determine the optimal mini-batch size for each of the pre-trained networks. The learning rate was set to its optimal value of 0.001. Fig. 7 illustrates the accuracy obtained from the validation set when the networks were trained using mini-batch sizes of 4, 8, 16, and 32. The results show that the optimal mini-batch size for AlexNet, ResNet18, ResNet50, and the proposed network is 32, while the optimal mini-batch size for GoogLeNet is 16.

Using the optimal mini-batch size and learning rate for each network training will produce the best results. Thus, the same parameters will be used in the subsequent evaluation.

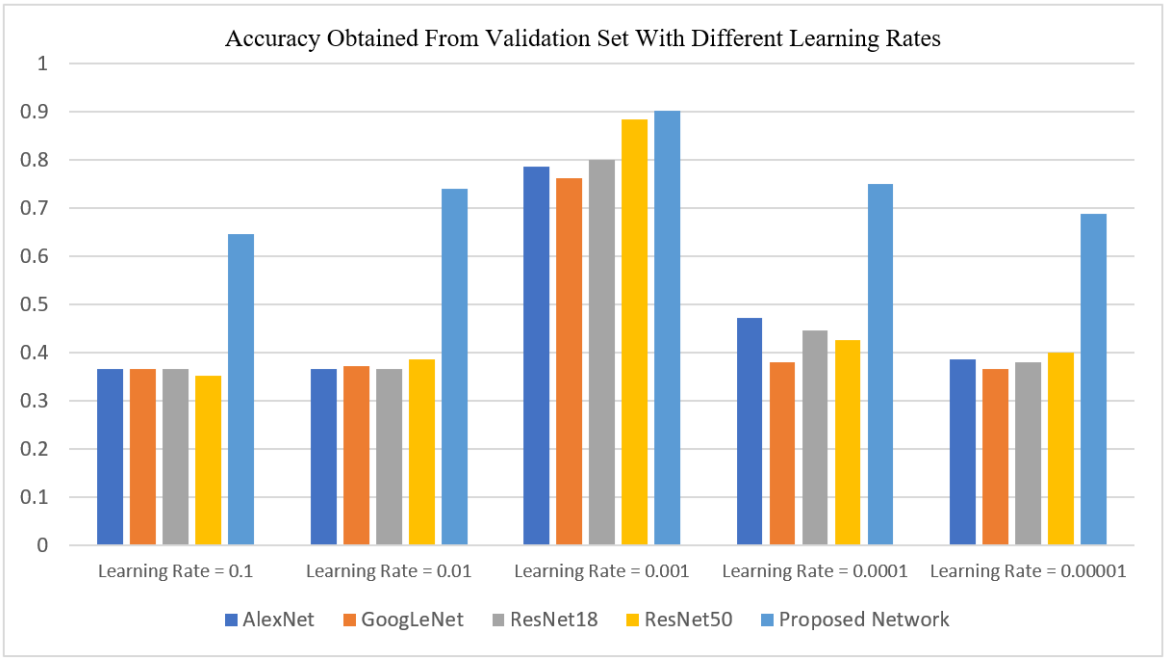


FIGURE 6. The accuracy obtained from the validation set with various learning rates for each pre-trained network.

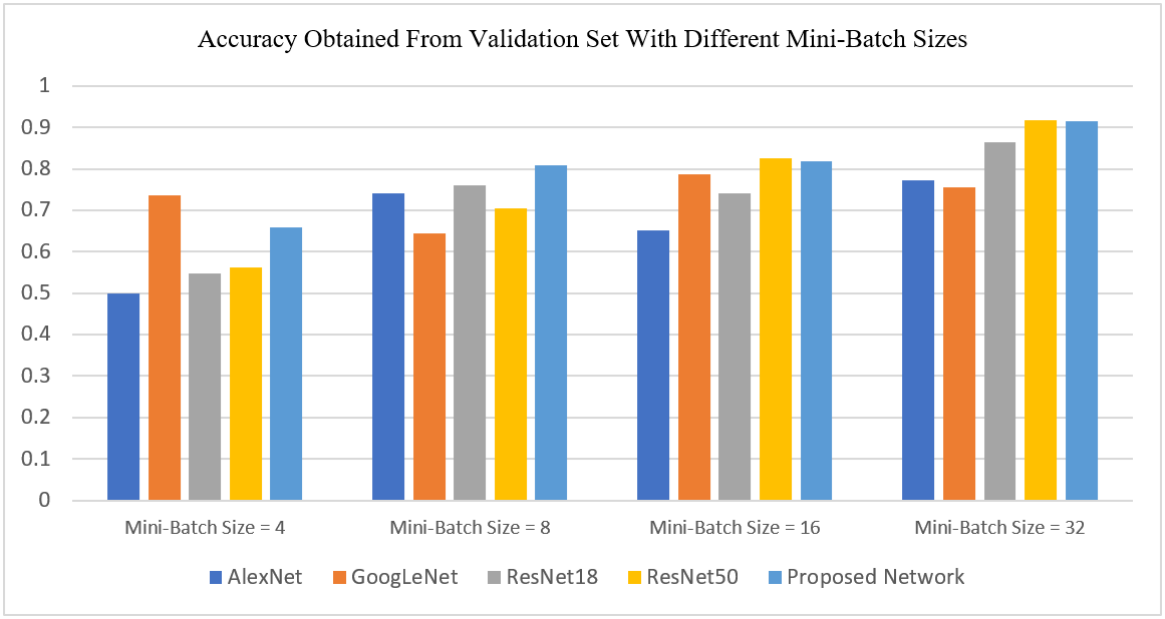


FIGURE 7. The accuracy obtained from the validation set with various mini-batch sizes for each pre-trained network.

B. CLASSIFICATION RESULTS

The performance of the networks trained using the optimum parameters was then evaluated using the testing set. The testing set includes images that have never been viewed previously by the network. The trained convolutional neural networks were used to classify the testing set images into Neo or NotNeo. The number of true positives, true negatives, false positives, and false negatives was determined by comparing each classified image patch to its ground truth.

Fig. 8 illustrates some of the classified image patches. Fig. 8(a) and (b) show examples of correctly classified positive and negative images, respectively. Neovascularization is present in Fig. 8(c), whereas it is absent in Fig. 8(d). However, the network misclassified these two groups of image patches. Due to the image patch's low contrast in

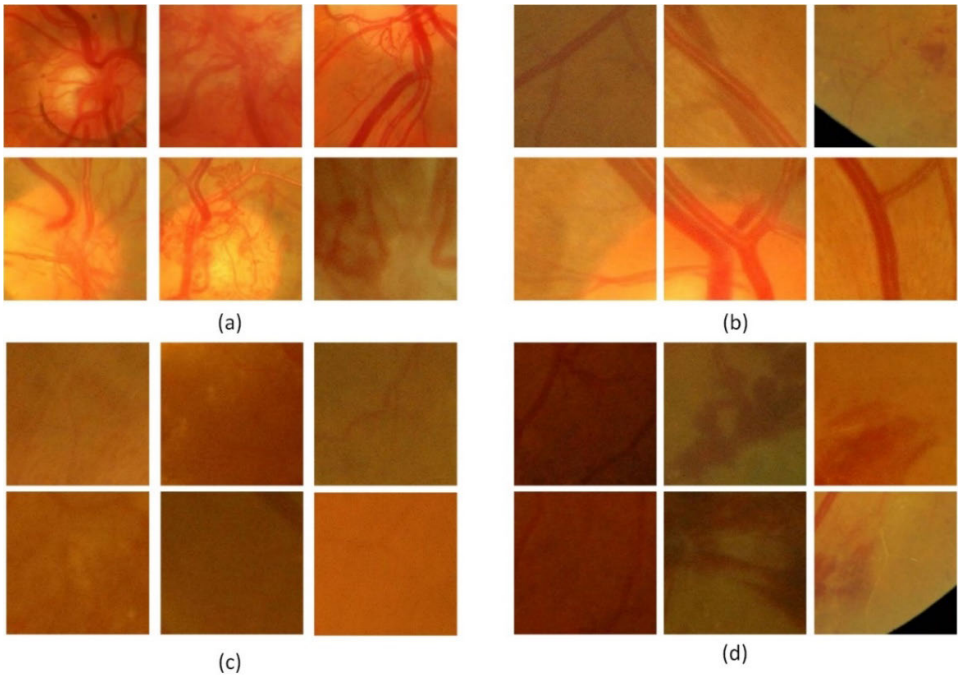


FIGURE 8. Examples of classification output from the transfer learning method (a) True positive. (b) True negative. (c) False negative. (d) False positive.

TABLE 1. Comparison of feature extraction and transfer learning performance tested on different pre-trained networks. The best results are highlighted in bold.

Pre-trained Models	Accuracy		Sensitivity		Specificity		Precision	
	Transfer Learning	Feature Extraction + SVM	Transfer Learning	Feature Extraction + SVM	Transfer Learning	Feature Extraction + SVM	Transfer Learning	Feature Extraction + SVM
AlexNet	0.7913	0.6533	0.8143	0.7019	0.7683	0.6048	0.7785	0.6398
GoogLeNet	0.7649	0.6337	0.6491	0.7155	0.8807	0.5520	0.8448	0.6149
ResNet18	0.8842	0.6908	0.8228	0.7649	0.9455	0.6167	0.9379	0.6662
ResNet50	0.8271	0.7274	0.7138	0.7632	0.9404	0.6917	0.9229	0.7122
Proposed Method	**0.9157**		**0.8569**		**0.9744**		**0.9710**	

Fig. 8(c), misclassification of positive patches occurred. Additionally, the images appear to be blurry. As a result, the neovascularization features become obscure, leading to wrong classifications.

In Fig. 8(d), false positives occurred because certain image patches resemble the features of neovascularization. These image patches contain various types of lesions that resemble neovascularization features, including hemorrhages and microaneurysms. Therefore, the networks may occasionally misinterpret these objects as neovascularization. After calculating the total number of true positives, false positives, and false negatives in the testing set, accuracy, sensitivity, specificity, and precision were determined.

C. PERFORMANCE COMPARISON

The performance of the proposed network is compared to those of the pre-trained networks. The results are also compared to the feature extraction and SVM classification method proposed by Setiawan *et al.* [25]. To ensure a fair comparison, the classifier in [25] was trained and tested on our dataset and evaluated using the same performance metrics. The pre-trained models used in the implementation are also the same as those used in the transfer learning method, including AlexNet, GoogLeNet, ResNet18, and ResNet50. Table 1 presents the results of transfer learning based on the individual pre-trained CNN, feature extraction + SVM methods, and the proposed method.

In general, the transfer learning approach outperformed feature extraction and SVM classification. For all the pre-trained networks, the transfer learning method has higher accuracy, specificity, and precision. This is because the transfer learning process retrained the entire pre-trained network, adjusting all the weights in the network to fit for neovascularization features detection. By contrast, the feature

extraction + SVM classification method extracts features using the original pre-trained networks and train the features on an SVM to detect neovascularization. Since the features extracted from the original pre-trained networks were not optimized for neovascularization detection, this method shows inferior performance compared to the transfer learning method.

The proposed method, which is based on the combination of ResNet18 and GoogLeNet, yields the highest results. This is because neovascularization lesions are identified using feature maps from both networks (ResNet18 and GoogLeNet). This has improved the results because more kernels are used to extract and learn neovascularization features, thereby increasing the detection accuracy of neovascularization.

ResNet50 outperforms the proposed model in the validation set when the optimum mini-batch size of 32 is used (see Fig. 7). However, when the same mini-batch size is used, ResNet50 produces lower accuracy than the proposed model in the testing set. This demonstrates that when the neovascularization dataset is used for training, ResNet50 is prone to overfitting.

The Receiver Operating Characteristic (ROC) curve and the area under the ROC curve (AUC) are used to analyze the networks' performance to determine which network is the most effective at classifying Neo and NotNeo patches via transfer learning. ROC depicts the diagnostic capabilities of a binary classifier system when its discrimination threshold is adjusted, while AUC can be used to summarize the classifier's ability to differentiate classes. The ROC plots for each network are shown in Fig. 9. It can be observed that the proposed network gives the best performance while AlexNet shows the worst results.

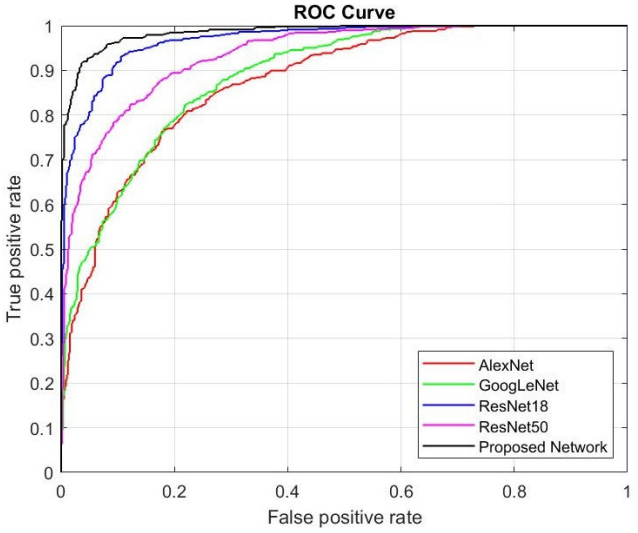


FIGURE 9. ROC curves to compare the performance of different networks for neovascularization detection using transfer learning.

The AUC of AlexNet, GoogLeNet, ResNet18, ResNet50, and the proposed network are 0.8737, 0.8864, 0.9685, 0.9345, and 0.9855, respectively. These results show that the proposed network with the biggest AUC is the best network for classifying Neo and NotNeo image patches. This demonstrates that the combination of ResNet18 and GoogLeNet outperforms the original pre-trained networks in terms of neovascularization detection via transfer learning.

The benefit of the transfer learning approach is that it requires less effort to implement. Furthermore, the model training does not begin with a random weight. Instead, pre-trained weights that have been previously trained are used as the starting point before being tuned to find the optimum weights for detecting neovascularization features. Thus, applying the transfer learning method will circumvent the lengthy training process and the problem of limited training images.

V. CONCLUSION

This paper presented a transfer learning approach for detecting neovascularization. A network that is based on the combination of ResNet18 and GoogLeNet is proposed. The performance of four pre-trained convolutional neural networks, which are AlexNet, ResNet18, ResNet50, and GoogLeNet, was also investigated for neovascularization detection through transfer learning. Experiment results based on the performance metrics and ROC plots show that the proposed network outperformed all these networks. This is because more features can be extracted by combining two pre-trained networks, resulting in more accurate detection of neovascularization. The results are also compared to another deep learning approach, which uses the pre-trained CNN for feature extraction and SVM for classification. Evaluation results show that the transfer learning approach yields superior performance. This paper demonstrated that applying transfer learning on the combined features of two pre-training networks can effectively detect neovascularization on fundus images. The work contributed toward the automatic detection of neovascularization, which is an important topic for the diagnosis of proliferative diabetic retinopathy.

REFERENCES

[1] P. Scanlon, S. Aldington, and I. Stratton, "Epidemiological issues in diabetic retinopathy," *Middle East Afr. J. Ophthalmol.*, vol. 20, no. 4, p. 293, 2013, doi: 10.4103/0974-9233.120007.

[2] S. D. Candrilli, K. L. Davis, H. J. Kan, M. A. Lucero, and M. D. Rousculp, "Prevalence and the associated burden of illness of symptoms of diabetic peripheral neuropathy and diabetic retinopathy," *J. Diabetes Complications*, vol. 21, no. 5, pp. 306–314, Sep. 2007, doi: 10.1016/j.jdiacomp.2006.08.002.

[3] I. Pearce, R. Simó, M. Lövestam-Adrian, D. T. Wong, and M. Evans, "Association between diabetic eye disease and other complications of diabetes: Implications for care. A systematic review," *Diabetes, Obesity Metabolism*, vol. 21, no. 3, pp. 467–478, Mar. 2019, doi: 10.1111/dom.13550.

[4] S. Loukovaara, E. Gucciardo, P. Repo, H. Vihinen, J. Lohi, E. Jokitalo, P. Salven, and K. Lehti, "Indications of lymphatic endothelial differentiation and endothelial progenitor cell activation in the pathology of proliferative diabetic retinopathy," *Acta Ophthalmol.*, vol. 93, no. 6, pp. 512–523, Sep. 2015, doi: 10.1111/aos.12741.

[5] N. D. Roy and A. Biswas, "Early detection of proliferative diabetic retinopathy in neovascularization at the disc by observing retinal vascular structure," in *Proceedings of the International Conference on ISMAC in Computational Vision and Bio-Engineering* (Lecture Notes in Computational Vision and Biomechanics) vol. 30. Cham, Switzerland: Springer, 2019, pp. 1441–1450.

[6] S. Aslani and H. Sarnel, "A new supervised retinal vessel segmentation method based on robust hybrid features," *Biomed. Signal Process. Control*, vol. 30, pp. 1–12, Sep. 2016, doi: 10.1016/j.bspc.2016.05.006.

[7] D. Youssef and N. H. Solouma, "Accurate detection of blood vessels improves the detection of exudates in color fundus images," *Comput. Methods Programs Biomed.*, vol. 108, no. 3, pp. 1052–1061, Dec. 2012, doi: 10.1016/j.cmpb.2012.06.006.

[8] D. N. H. Thanh, D. Sergey, V. B. Surya Prasath, and N. H. Hai, "Blood vessels segmentation method for retinal fundus images based on adaptive principal curvature and image derivative operators," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. W12, pp. 211–218, May 2019, doi: 10.5194/isprs-archives-XLII-2-W12-211-2019.

[9] A. K. Shukla, R. K. Pandey, and R. B. Pachori, "A fractional filter based efficient algorithm for retinal blood vessel segmentation," *Biomed. Signal Process. Control*, vol. 59, May 2020, Art. no. 101883, doi: 10.1016/j.bspc.2020.101883.

[10] E. Imani, M. Javidi, and H.-R. Pourreza, "Improvement of retinal blood vessel detection using morphological component analysis," *Comput Methods Programs Biomed.*, vol. 118, no. 3, pp. 263–279, Mar. 2015, doi: 10.1016/j.cmpb.2015.01.004.

[11] M. Merlin and B. Shan, "Robust and efficient segmentation of blood vessel in retinal images using gray-level textures features and fuzzy SVM," *Biomed. Pharmacol. J.*, vol. 8, no. 2, pp. 1111–1120, Dec. 2015, doi: 10.13005/bpj/865.

[12] S. W. Franklin and S. E. Rajan, "Computerized screening of diabetic retinopathy employing blood vessel segmentation in retinal images," *Biocybernetics Biomed. Eng.*, vol. 34, no. 2, pp. 117–124, 2014, doi: 10.1016/j.bbe.2014.01.004.

[13] T. J. Jebaseeli, C. A. Deva Durai, and J. D. Peter, "Retinal blood vessel segmentation from diabetic retinopathy images using tandem PCNN model and deep learning based SVM," *Optik*, vol. 199, Dec. 2019, Art. no. 163328, doi: 10.1016/j.ijleo.2019.163328.

[14] C. Kirbas and F. Quek, "A review of vessel extraction techniques and algorithms," *ACM Comput. Surv.*, vol. 36, no. 2, pp. 81–121, Jun. 2004, doi: 10.1145/1031120.1031121.

[15] M. C. Seng Tang and S. S. Teoh, "Blood vessel segmentation in fundus images using Hessian matrix for diabetic retinopathy detection," in *Proc. 11th IEEE Annu. Inf. Technol., Electron. Mobile Commun. Conf. (IEMCON)*, Nov. 2020, pp. 0728–0733, doi: 10.1109/IEMCON51383.2020.9284931.

[16] G. Liew, J. J. Wang, P. Mitchell, and T. Y. Wong, "Retinal vascular imaging: A new tool in microvascular disease research," *Circulat., Cardiovascular Imag.*, vol. 1, no. 2, pp. 156–161, Sep. 2008, doi: 10.1161/CIRCIMAGING.108.784876.

[17] R. Gargeya and T. Leng, "Automated identification of diabetic retinopathy using deep learning," *Ophthalmology*, vol. 124, no. 7, pp. 962–969, 2017, doi: 10.1016/j.ophtha.2017.02.008.

[18] N. Hiraoka, E. Allen, I. J. Apel, M. R. Gyetko, and S. J. Weiss, "Matrix metalloproteinases regulate neovascularization by acting as pericellular fibrinolysins," *Cell*, vol. 95, no. 3, pp. 365–377, Oct. 1998, doi: 10.1016/S0092-8674(00)81768-7.

[19] S. Gandhimathi and K. Pillai, "Detection of neovascularization in proliferative diabetic retinopathy fundus images," *Int. Arab J. Inf. Technol.*, vol. 15, no. 6, pp. 1000–1009, 2018.

[20] D. F. G. Coelho, R. M. Rangayyan, and V. S. Dimitrov, "Detection of neovascularization near the optic disk due to diabetic retinopathy," in *Proc. 24th Eur. Signal Process. Conf. (EUSIPCO)*, Aug. 2016, pp. 2040–2044, doi: 10.1109/EUSIPCO.2016.7760607.

[21] S. S. Kar and S. P. Maity, "Detection of neovascularization in retinal images using mutual information maximization," *Comput. Electr. Eng.*, vol. 62, pp. 194–208, Aug. 2017, doi: 10.1016/j.compeleceng.2017.05.012.

[22] J. Lee, B. C. Y. Zee, and Q. Li, "Detection of neovascularization based on fractal and texture analysis with interaction effects in diabetic retinopathy," *PLoS ONE*, vol. 8, no. 12, Dec. 2013, Art. no. e75699, doi: 10.1371/journal.pone.0075699.

[23] K. Saranya, B. Ramasubramanian, and S. K. Mohideen, "A novel approach for the detection of new vessels in the retinal images for screening diabetic retinopathy," in *Proc. Int. Conf. Commun. Signal Process.*, Apr. 2012, pp. 57–61, doi: 10.1109/ICCSP.2012.6208394.

[24] K. A. Goatman, A. D. Fleming, S. Philip, G. J. Williams, J. A. Olson, and P. F. Sharp, "Detection of new vessels on the optic disc using retinal photographs," *IEEE Trans. Med. Imag.*, vol. 30, no. 4, pp. 972–979, Apr. 2011, doi: 10.1109/TMI.2010.2099236.

[25] W. Setiawan, M. I. Utoyo, and R. Rulaningtyas, "Classification of neovascularization using convolutional neural network model," *TELKOMNIKA Telecommun. Comput. Electron. Control*, vol. 17, no. 1, p. 463, Feb. 2019, doi: 10.12928/telkomnika.v17i1.11604.

[26] C. Carrillo-Gomez, M. Nakano, A. Gonzalez-H.Leon, J. C. Romo-Aguas, H. Quiroz-Mercado, and O. Lopez-Garcia, "Neovascularization detection on optic disc region using deep learning," in *Pattern Recognition* (Lecture Notes in Computer Science), vol. 12725. Cham, Switzerland: Springer, 2021, pp. 111–120.

[27] H. A. Hassan, M. Yaakob, S. Ismail, J. A. Rahman, I. M. Rusni, A. Zabidi, I. M. Yassin, N. M. Tahir, and S. M. Shafie, "Detection of proliferative diabetic retinopathy in fundus images using convolution neural network," in *Proc. IOP Conf. Mater. Sci. Eng.*, vol. 769, no. 1, Jun. 2020, Art. no. 012029, doi: 10.1088/1757-899X/769/1/012029.

[28] M. C. S. Tang, S. S. Teoh, H. Ibrahim, and Z. Embong, "Neovascularization detection and localization in fundus images using deep learning," *Sensors*, vol. 21, no. 16, p. 5327, Aug. 2021, doi: 10.3390/s21165327.

[29] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017, doi: 10.1145/3065386.

[30] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2015, pp. 1–9, doi: 10.1109/CVPR.2015.7298594.

[31] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.

[32] E. Decencière, X. Zhang, G. Cazuguel, B. Lay, B. Cochener, C. Trone, P. Gain, R. Ordonez, P. Massin, A. Erginay, and B. Charton, "Feedback on a publicly distributed image database: The messidor database," *Image Anal. Stereol.*, vol. 33, no. 3, pp. 231–234, Aug. 2014, doi: 10.5566/ias.1155.

[33] T. Kauppi, V. Kalesnykiene, J. Kamarainen, L. Lensu, and I. Sorri, "Diaretdb0: Evaluation database and methodology for diabetic retinopathy algorithms," Mach. Vis. Pattern Recognit. Res. Group, Lappeenranta Univ. Technol. Lappeenranta, Finland, Tech. Rep., 2006, pp. 1–17.

[34] K. Huang and M. Yan, "A region based algorithm for vessel detection in retinal images," in *Medical Image Computing and Computer-Assisted Intervention* (Lecture Notes in Computer Science), vol. 4190. Berlin, Germany: Springer, 2006, pp. 645–653.

[35] E. Daniel and J. Anitha, "Optimum green plane masking for the contrast enhancement of retinal images using enhanced genetic algorithm," *Optik*, vol. 126, no. 18, pp. 1726–1730, Sep. 2015, doi: 10.1016/j.ijleo.2015.05.027.

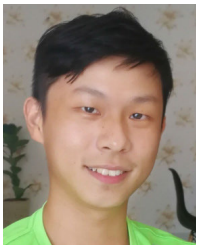
[36] S. M. Pizer, E. P. Amburn, J. D. Austin, R. Cromartie, A. Geselowitz, T. Greer, B. ter Haar Romeny, J. B. Zimmerman, and K. Zuiderveld, "Adaptive histogram equalization and its variations," *Comput. Vis., Graph., Image Process.*, vol. 39, no. 3, pp. 355–368, 1987, doi: 10.1016/S0734-189X(87)80186-X.

[37] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 248–255, doi: 10.1109/CVPR.2009.5206848.

[38] S. Khan, H. Rahmani, S. A. A. Shah, and M. Bennamoun, "A guide to convolutional neural networks for computer vision," *Synth. Lect. Comput. Vis.*, vol. 8, no. 1, pp. 1–207, Feb. 2018, doi: 10.2200/S00822ED1V01Y201712COV015.

[39] Y. Bengio, "Practical recommendations for gradient-based training of deep architectures," in *Neural Networks: Tricks of the Trade* (Lecture Notes in Computer Science), vol. 7700. Berlin, Germany: Springer, 2012, pp. 437–478.

[40] N. S. Keskar, D. Mudigere, J. Nocedal, M. Smelyanskiy, and P. T. P. Tang, "On large-batch training for deep learning: Generalization gap and sharp minima," in *Proc. 5th Int. Conf. Learn. Represent. (ICLR)*, Sep. 2016, pp. 1–16.



MICHAEL CHI SENG TANG received the B.Eng. degree in electrical and electronics engineering from University Malaysia Pahang, in 2019. He is currently pursuing the Ph.D. degree in electrical and electronics engineering with the University of Science, Malaysia. His current research interests include image processing and deep learning applications in the biomedical field.

SOO SIANG TEOH (Senior Member, IEEE) received the B.Eng. degree in electronic engineering from University Putra Malaysia, in 1993, the M.S. degree in digital electronics from the University of Manchester, U.K., in 1995, and the Ph.D. degree in computer engineering from the University of Western Australia, in 2012. He is currently a Senior Lecturer with the School of Electrical and Electronic Engineering, Universiti Sains Malaysia. His research interests include image processing and machine learning for industrial and biomedical applications.

HAIDI IBRAHIM (Senior Member, IEEE) received the B.Eng. degree in electrical and electronic engineering from Universiti Sains Malaysia, Malaysia, and the Ph.D. degree in image processing from the Centre for Vision, Speech and Signal Processing (CVSSP), University of Surrey, U.K., in 2005. His research interests include digital image and signal processing and analysis.

ZUNAINA EMBONG received the Medical degree from University Kebangsaan Malaysia, Malaysia, in 1990, and the Master of Medicine (M.Med.) degree in ophthalmology from Universiti Sains Malaysia, Malaysia, in 2003. She underwent both Medical Retina and Surgical Retina Training at Westmead Hospital, Sydney, Australia, in 2005 and 2009, respectively. She underwent further training in Uveitis under the International Council of Ophthalmology (ICO) at Yamaguchi University, Japan, in 2014. She is currently a Senior Lecturer with the School of Medical Sciences, Universiti Sains Malaysia. Her research interests include retina diseases and retinal imaging especially diabetic retinopathy.