*Sensors and Materials*, Vol. 35, No. 4 (2023) 1433–1448
MYU Tokyo



S & M 3258

# Image Enhancement and 27 Pretrained Convolutional Neural Network Models for Diabetic Retinopathy Grading

Isoon Kanjanasurat,[1] Thanavit Anuwongpinit,[2*] and Boonchana Purahong[2]

[1]College of Computing, Khon Kaen University,
123 Vidhayavibaj Building, Mitraparp Road, Muang District, Khon Kaen 40002, Thailand
[2]School of Engineering, King Mongkut's Institute of Technology Ladkrabang,
1 Chalong Krung, 1 Alley, Lat Krabang, Bangkok 10520, Thailand

(Received August 15, 2022; accepted February 7, 2023)

***Keywords:*** diabetic retinopathy, grading, classification, deep learning

Diabetic retinopathy (DR) affects the retina's blood vessels and causes vision loss. Fundus images are used to diagnose DR, which is a lengthy process because experienced clinicians must accurately diagnose the disease and identify microlesions early to prevent blindness. Computer vision can be used for retinal image classification. The APTOS dataset contains 5990 normal, moderate, mild, proliferate, and severe retinal images. In this study, we proposed a convolutional neural network (CNN) ensemble for DR fundus grading. Each image channel was enhanced by contrast-limited adaptive histogram equalization (CLAHE) and gamma correction and then fed to 27 pretrained CNN models for one-time training to examine the DR grading. The results showed that MobileNet's green channel with the CLAHE technique is sufficiently fast and accurate for disease classification. The grading retinal images had an accuracy of 96.95%, a precision of 96.17%, a sensitivity of 97.80%, an F1 score of 96.98%, and a specificity of 97.75%. In addition, the proposed method improves the speed and robustness of retinal DR grading.

## 1. Introduction

There are 2.2 billion people in the world with some form of vision impairment, according to the *World Report on Vision* presented by the World Health Organization (WHO) in October 2019, and at least one billion of these cases could have been prevented through timely screening.(1) Patients with diabetes are more likely to experience vision loss or blindness as a direct result of diabetic retinopathy (DR).(2) Such patients often experience DR as a consequence of having high levels of glucose in their blood. DR occurs through the formation of abnormal blood vessels (leaks or swelling) at the posterior pole (back) of the eye, which damages the retina. If DR is not detected early, it can lead to permanent vision impairment. In terms of its clinical presentation, the two distinguishable stages of DR are proliferative diabetic retinopathy (PDR) and non-proliferative diabetic retinopathy (NPDR).(3) PDR, the more advanced stage of DR, can cause severe vision loss in patients if new aberrant arteries bleed into the vitreous

*Corresponding author: e-mail: thanavit.an@kmitl.ac.th
https://doi.org/10.18494/SAM4084

ISSN 0914-4935 

humor or if there is tractional retinal detachment. NPDR is associated with the early stages of DR and characterized by increased vascular permeability and capillary occlusion in the retinal vasculature. NPDR is classified into three stages: mild, moderate, and severe. Table 1 outlines the clinical characteristics used to grade the severity of DR. As clinical features, microaneurysms are small red or pink dots, which are round areas of swelling in the blood vessels of the eye. Dot and blot hemorrhages are small red or pink dots, or round or irregularly shaped areas of bleeding in the retina. Hard exudates are deposits of fatty material that accumulate in the retina, which may appear as yellow or white spots on the retina. Cotton wool spots are areas of damaged nerve tissue in the retina, which may appear as white or pale-yellow areas on the retina. In this study, we propose a method with five classifications: normal, mild DR, moderate DR, severe DR, and PDR. Figure 1 depicts fundus images corresponding to these five classifications.

The imaging of the fundus is a frequently used method for grading DR in the general population. Manual fundus image screening takes a significant amount of time and is highly dependent on the ophthalmologist's expertise.[4] Computer-aided medical image processing is a promising method for performing mass screening with prompt and accurate diagnoses. Various

Table 1
Clinical features of DR.

| Grade | Clinical features | Category |
|---|---|---|
| 0 | No signs observed | Healthy |
| 1 | Microaneurysms present in one of the four quadrants | Mild DR |
| 2 | Microaneurysms, dot and blot hemorrhages, hard exudates, cotton wool spots | Moderate DR |
| 3 | Intraretinal microvascular abnormalities, definite venous beading, intraretinal hemorrhages | Severe DR |
| 4 | Vitreous/preretinal hemorrhage | PDR |

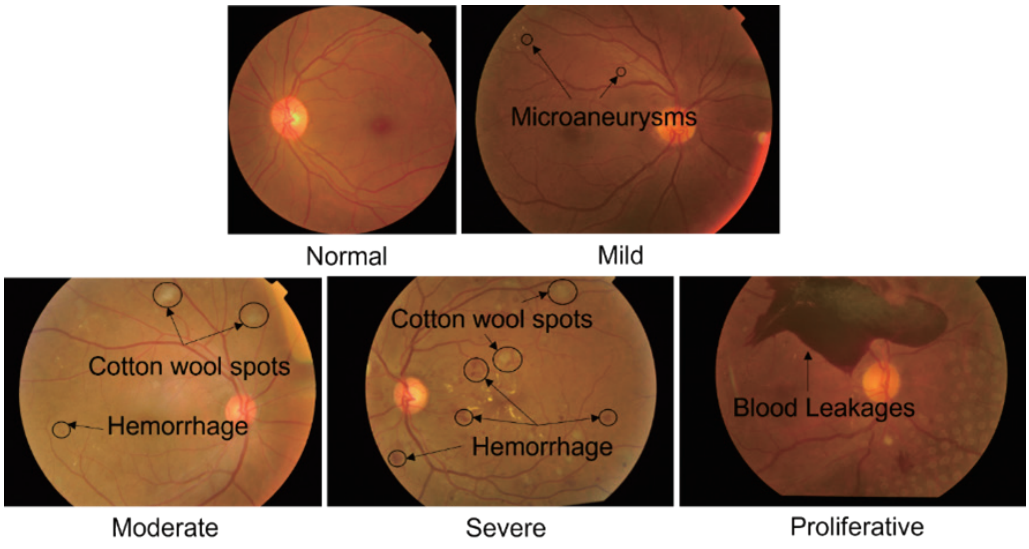



Fig. 1. (Color online) Grading of DR.

computer-aided diagnostic systems based on conventional image processing methods have been reported. Image preprocessing can be performed in different ways, each of which can achieve high performance with less calculation time and improved output.

Sinan *et al.* developed a transfer convolutional neural network (CNN) to diagnose retinal diseases. This network supplanted a contrast-limited adaptive histogram filter with a ResNet50 CNN that had been fine-tuned for optimal performance.[5] Kanjanasurat *et al.* demonstrated that contrast-limited adaptive histogram equalization (CLAHE) enhances blood vessel visibility in DR images when used in conjunction with Delaunay triangulation and optic disc retinal vascular patterns for personal identification.[6] By preprocessing data with an IterNet model capable of achieving high accuracy, Tenghongsakul *et al.* improved the quality of blood vessel extraction from the retinal green channel.[7]

In the preprocessing of CNN-based automated screening, Yuping *et al.* employed the image enhancement provided by the Caffe framework and achieved high performance.[8] By combining adaptive histogram equalization and morphological shape, Prateek *et al.* improved the detection of exudates in DR, using a cellphone to analyze their shape.[9]

Many CNN models have received considerable attention from researchers. Using the MESSIDOR database, Rancisco *et al.* demonstrated the application of deep residual transfer learning to the autonomous diagnosis of DR using various convolutional architectures.[10] Gayathri *et al.* used a lightweight CNN to classify DR and achieved high accuracy for both binary and multiclass classifications.[11] Xiang *et al.* graded images of DR by employing a lesion-attention pyramid network,[12] which generates a lesion-activation map that provides consistent lesion locations for clinical diagnosis. Teresa *et al.* used deep learning (DL) to analyze fundus images of the eye to determine the presence of DR.[13] In most cases, the maps focused on locations that are significant for diagnosis. The results of these studies demonstrate the significant potential of a second-opinion approach in determining the severity of DR. Sraddha *et al.* successfully classified DR by employing a DL architecture based on the segmented fundus image characteristics taken from the DIARETDB1 dataset.[14] This method outperformed more traditional methods. Zhan *et al.* used a coarse-to-fine classification for DR grading in conjunction with a CNN to identify five classes of DR severity,[15] significantly improving the classification of DR. AbdelMaksoud *et al.* used a new machine-learning-based classification system for DR grading to detect the early signs of DR. They used the new ML-CAD system on four public datasets: DRIVE, STARE, IDRiD, and MESSIDOR.[16] Rory *et al.* integrated gradients and DL algorithms to assist the grading of DR.[17] This increased the accuracy of grading but also increased the time required for grading. De la Torre *et al.* tackled the problem of grading DR with an innovative method of pixelwise score propagation.[18] They assigned a score for each point of the fundus image. Hassan *et al.* presented a cascaded decoupled CNN for lesion-assisted identification and grading on OCT scans.[19] This network compared two distinct modules for grading retinopathy using lesion-assisted grading under clinical standards.

In this study, we investigated the classification of diabetic retinal images using 27 pretrained CNN models and various image enhancement techniques. The findings make two significant contributions. First, we review how various methods of image enhancement can increase the

accuracy in each channel of the retinal image. Second, we compare different pretrained CNN models to find a classification algorithm that produces the most accurate results in the shortest computing time.

The following is the structure of this paper. The datasets are described in Sect. 2.1. Image enhancement is discussed in Sect. 2.2 and the CNNs are described in Sect. 2.3. The findings are presented in Sect. 3. A conclusion is given in Sect. 4.

## 2. Materials and Methods

The block diagram in Fig. 2 shows the methodology of our research. An individual DR database from a previous publication.[7] was analyzed in this study. The experiments involved evaluating each channel of the original DR images using two improved techniques for classifying DR into five grades: none, mild, moderate, severe, and proliferative. The approach used 27 pretrained CNNs in a single training session, as well as the classification of DR into two categories (no DR and DR). In addition to developing various performance matrices for quantifying network performance, we determined the elapsed time per image for the best-performing network in each image enhancement classification technique.

The specifications of the study include information on the datasets and methods for improving strategies. In addition, performance matrices were used to assess the efficacy of various strategies. The following is a description of the parameters of the investigation, including information about the datasets, improvement strategies, methods used, and performance matrices.

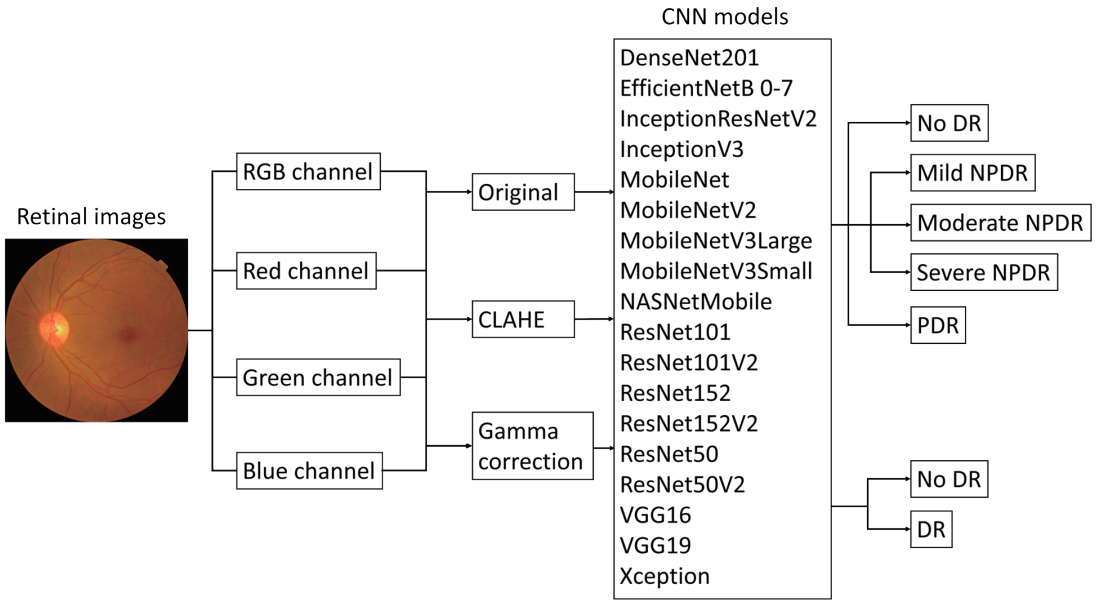



Fig. 2. (Color online) Overall process proposed in this study.

### 2.1 Datasets

The Kaggle APTOS 2019 Blindness Detection (APTOS2019) dataset was created specifically for the Asia Pacific Tele-Ophthalmology Society's 2019 Blindness Detection competition. This dataset, which is publicly available, was used to evaluate the DR grading in this study. It is a large collection of retinal images obtained under various imaging conditions. The collection was compiled to develop a machine learning model that could detect blindness without the need for medical screening. This dataset consists of 3662 color retinal images obtained from a variety of subjects in rural India. A clinician evaluated each image for DR and assigned it a severity score between 0 and 4 (1–mild, 2–moderate, 3–severe, 4–multiple). No PDR was present. As with any real-world dataset, both the images and the labels contain noise. The images may contain artifacts, be out of focus, or have inappropriate exposure. The photographs were collected over an extended period from a variety of medical facilities using a variety of camera types.

### 2.2 Image enhancement techniques

Image enhancement is the process of highlighting important details during image processing and reducing or removing irrelevant details to improve image quality. The objective is to make images such as those in Fig. 3 more useful for a particular task. In this study, CLAHE and gamma correction were employed. This subsection describes these image enhancement techniques.

#### 2.2.1 Contrast-limited adaptive histogram equalization

Histogram equalization (HE) can be upgraded to a more sophisticated adaptive histogram equalization (AHE), which equalizes the histograms of small sections and increases the contrast of each region separately. Thus, the contrast and edges in each section of an image are modified in accordance with the local distribution of pixel intensities, rather than the actual image itself. However, this can increase the noise of the image. CLAHE prevents the contrast from being amplified to an undesirable degree. CLAHE is designed to work on discrete parts of an image, which it refers to as tiles, rather than the entire picture. After the so-called fictitious boundaries have been removed, the adjoining tiles are blended by bilinear interpolation.

#### 2.2.2 Gamma correction

During image normalization, linear operations on individual pixels, such as scalar multiplication, addition, and subtraction, are carried out. These operations are performed on the pixels themselves. The pixels in the source image are subjected to a nonlinear operation known as gamma correction. Alternating the pixel values is one way that gamma correction can help improve the quality of an image. This is achieved by using the projection relationship between the pixel value and the gamma value, which is determined using an internal map.

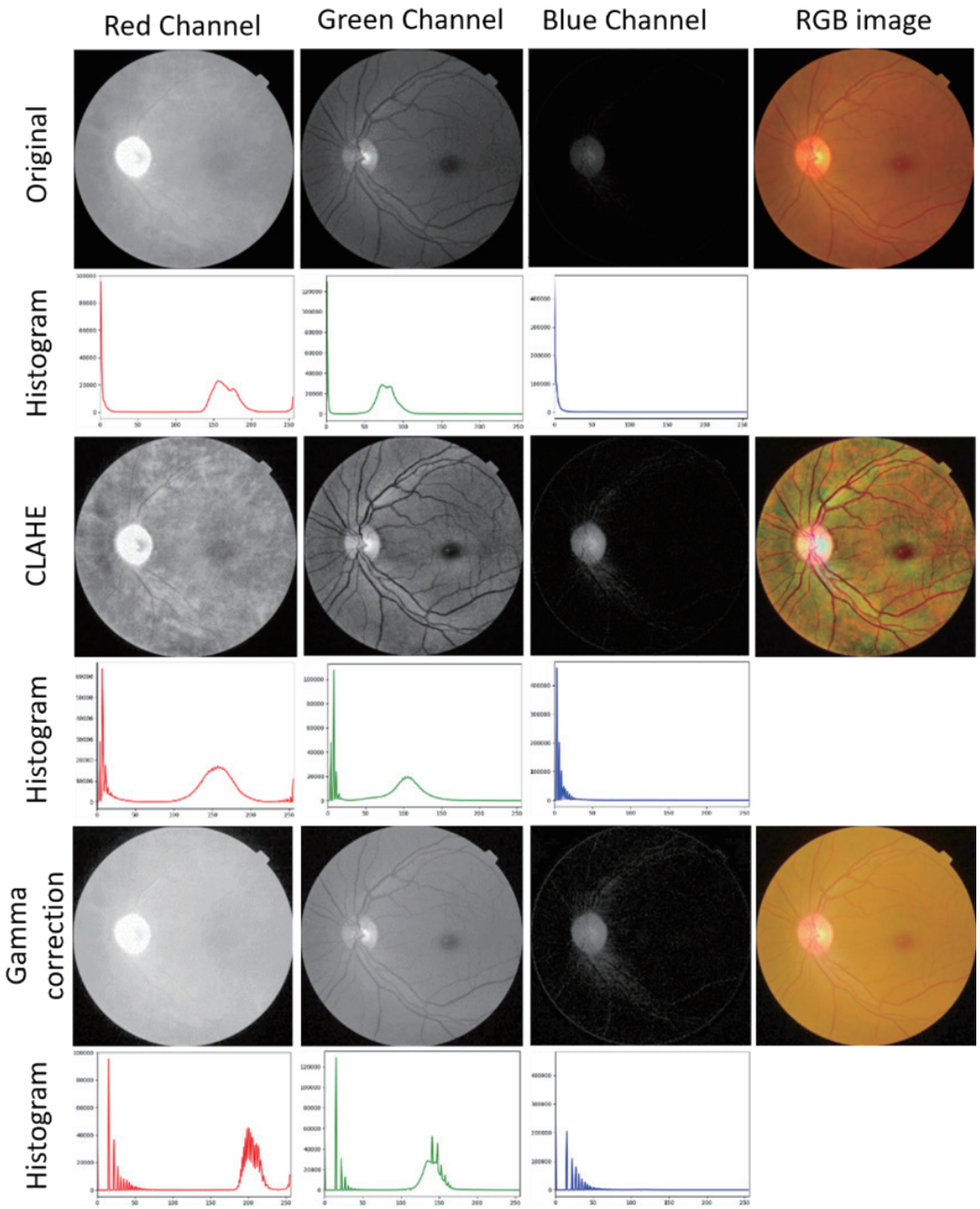



Fig. 3. (Color online) Image enhancement technique and histogram.

## 2.3 Convolutional neural networks

Deep learning (DL) is an approach in machine learning that uses artificial neural networks and is modeled after the architecture of the human brain. DL also refers to methods that can automatically learn the mathematical representation of hidden and intrinsic relations within data.

It extracts the spatial and temporal characteristics of an image in an automated fashion using neural networks. DenseNet121, DenseNet169, DenseNet201, EfficientNetB0 to EfficientNetB7, Inception ResNetV2, InceptionV3, MobileNet, MobileNetV2, MobileNetV3Large, MobileNetV3Small, NASNetMobile, ResNet101, ResNet101V2, ResNet152, ResNet152V2, ResNet50, ResNet50V2, VGG16, VGG19, and Xception were among the 27 pretrained DL CNNs used for the evaluation in this study.

DenseNet121, DenseNet169, and DenseNet201 are CNNs developed by Huang *et al.* in 2017.[20] These models use dense connectivity, meaning that each layer in the network is connected to all the previous layers, which allows the network to learn more efficiently and reduces the number of parameters needed to train a model. DenseNet121 has 121 layers and approximately 8 million parameters, while DenseNet169 has 169 layers and approximately 14 million parameters. DenseNet201 has 201 layers and approximately 20 million parameters. These models have high accuracy and efficiency and are often used in applications such as image classification and object detection.

EfficientNetB0 to EfficientNetB7 are CNN models developed by Tan and Le in 2019.[21] These models use a combination of efficient architecture design and model scaling to improve their accuracy and efficiency. They use techniques such as depthwise separable convolutions and weight-sharing to reduce the number of parameters needed to train them. EfficientNetB0 has 6.4 million parameters, while EfficientNetB7 has 66.4 million parameters.

Inception ResNetV2 is a CNN model developed by Szegedy *et al.* in 2016.[22] It combines the Inception architecture, which uses a combination of different filters and pooling operations to capture complex patterns in the data, with the Residual Network (ResNet) architecture, which uses skip connections to improve the flow of information through the network. Inception ResNetV2 has approximately 55 million parameters and has high accuracy and efficiency.

InceptionV3 is a CNN model developed by Szegedy *et al.* in 2015.[23] It uses the Inception architecture with a few modifications, such as the use of batch normalization and more filters in some layers. It has approximately 23.9 million parameters and has high accuracy and efficiency. It is often used in applications such as image classification and object detection.

MobileNet, MobileNetV2, MobileNetV3Large, and MobileNetV3Small are CNN models developed by Google.[24] These models use a combination of depthwise separable and pointwise convolutions to reduce the number of parameters needed to train a model. MobileNet has approximately 4.2 million parameters, while MobileNetV2 has approximately 3.4 million parameters. MobileNetV3Large and MobileNetV3Small have approximately 9.9 and 2.5 million parameters, respectively. These models are designed to be efficient and run well on mobile devices and are often used in applications such as image classification and object detection.

NASNetMobile is a CNN model developed by Google in 2018.[25] It is designed to be efficient and accurate and has approximately 5.3 million parameters.

ResNet101, ResNet101V2, ResNet152, ResNet152V2, ResNet50, and ResNet50V2 are CNN models developed by He *et al.* in 2015.[26] These models use the ResNet architecture, which uses skip connections to improve the flow of information through the network. ResNet101 has approximately 44.6 million parameters, while ResNet101V2 has approximately 44.5 million

parameters. ResNet152 has approximately 60.2 million parameters, while ResNet152V2 has approximately 60.1 million parameters. ResNet50 has approximately 25.6 million parameters.

VGG16 and VGG19 are CNNs developed by Simonyan and Zisserman in 2014.[27] These models have high accuracy and are often used in applications such as image classification and object detection. VGG16 has approximately 138 million parameters, while VGG19 has approximately 143 million parameters. These models use a series of convolutional layers with small 3 × 3 filters, followed by max pooling layers, to extract features from the input data.

Xception is a CNN model developed by Chollet in 2016.[28] It has high accuracy and efficiency, and is often used in applications such as image classification and object detection. Xception has approximately 22 million parameters and uses a combination of depthwise separable and pointwise convolutions to reduce the number of parameters needed to train a model. It also uses a special type of skip connection called an "excitation" connection, which helps to improve the flow of information through the network.

### 2.4 Evaluation

Accuracy, sensitivity, specificity, precision, and F1 score were the five performance metrics used to evaluate the classification performance of the various CNNs. These metrics are defined in Eqs. (1)–(5). The networks were compared using performance metrics specific to each class, as well as an overall accuracy metric. This was necessary because each class contained a different number of images.

$$Accuracy\ (ACC) = (TP + TN) / (TP + FN + FP + TN) \tag{1}$$

$$Sensitivity\ (SE) = TP / (TP + FN) \tag{2}$$

$$Specificity\ (SP) = TN / (FP + TN) \tag{3}$$

$$Precision\ (P) = TP / (TP + FP) \tag{4}$$

$$F1\ score\ (F) = 2 * TP / (2 * TP + FN + FP) \tag{5}$$

We used the terms true positive (*TP*), true negative (*TN*), false positive (*FP*), and false negative (*FN*) to indicate the images correctly identified as having diabetic disease, the healthy retina images correctly identified as normal, the images incorrectly identified as diabetic, and the unhealthy retina images that were incorrectly identified as healthy, respectively. We compared the performance characteristics of classification networks by analyzing their elapsed time per image, which is the amount of time required for each network to classify an input image into each category.

## 3. Results

The initial step of the experiment consisted of enhancing the quality of the original images. Following this, the dataset was divided into a learning set of 70% and a test set of 30%. Then, the learning dataset was used to instruct the model and determine which model provided the best overall performance. After obtaining the model with the highest performance, the test dataset was used for evaluation. Figure 4 is a block diagram of the experimental setup.

The aim of the experiments was to determine how DR patients should be compared with normal patients. Figure 5 shows a comparison of the accuracies of the models in classifying the images by using two image enhancement methods. The initial part of the experiment consisted of assessing how well each model performed. Only one session was required to learn the model.

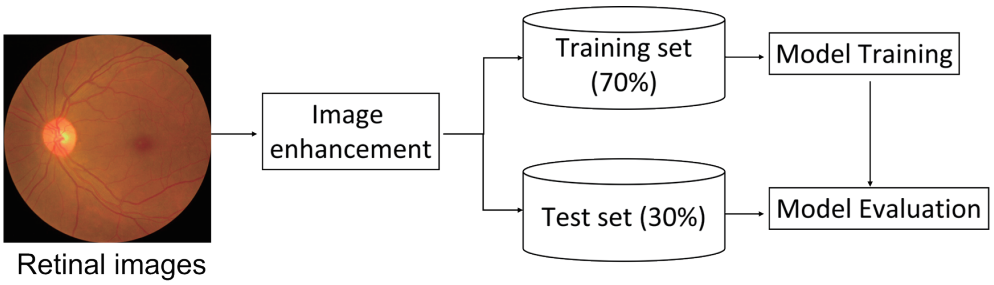



Fig. 4. (Color online) Block diagram of the experimental setup.

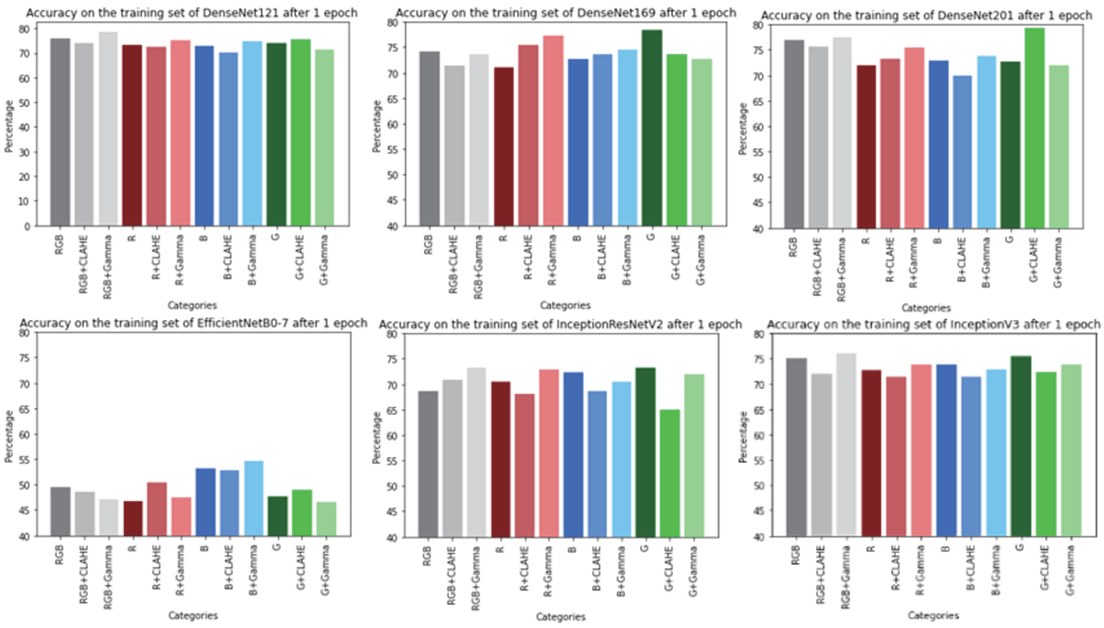



Fig. 5. (Color online) Accuracies of models for classification between the five classes of DR images by using two image enhancement methods.

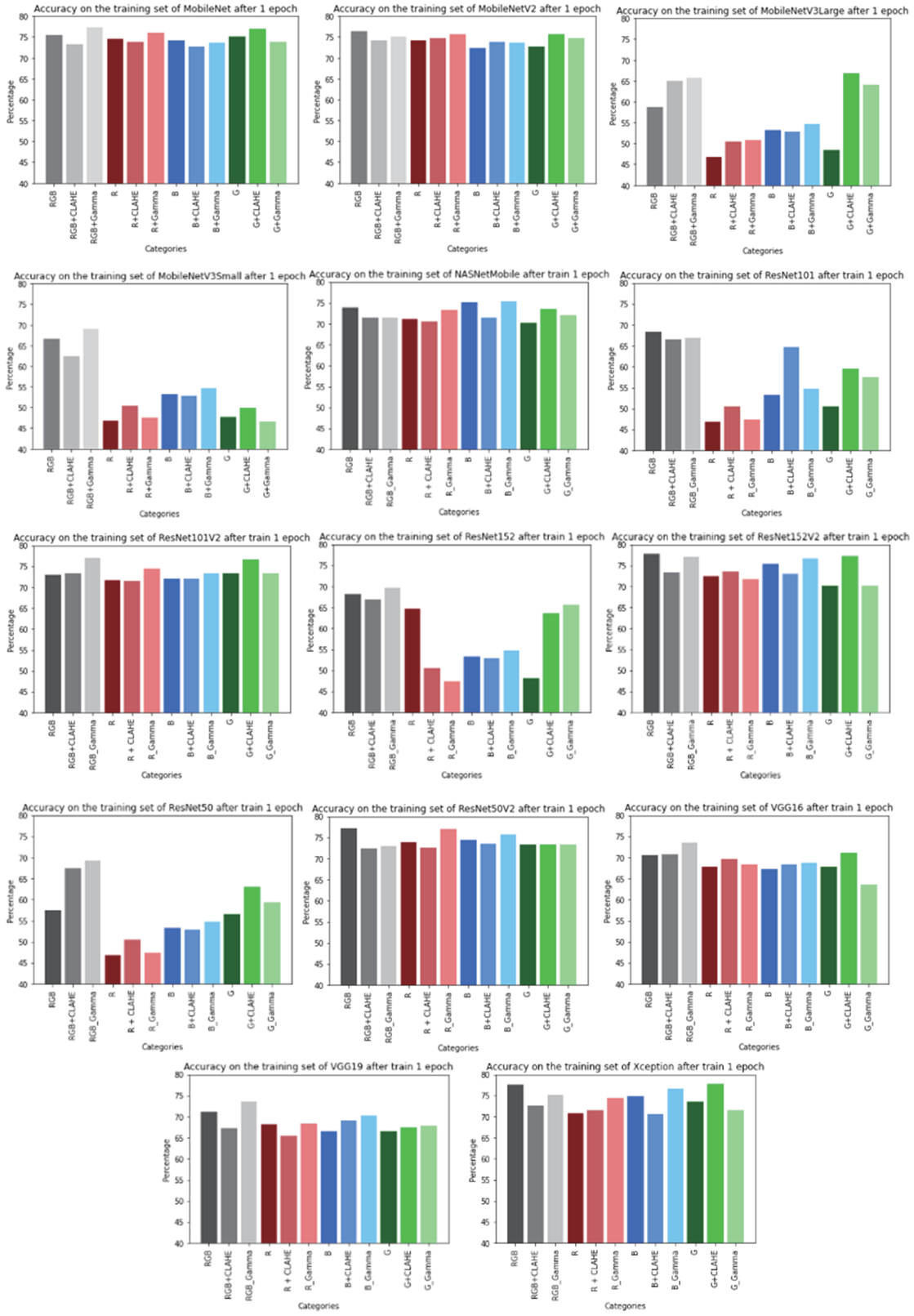



Fig. 5. (Color online) (Continued) Accuracies of models for classification between the five classes of DR images by using two image enhancement methods.

Each chart provides specific information regarding the accuracy of each method and the color space of the image, which is broken down into 12 categories: RGB, RGB with CLAHE, RGB with Gamma, R channel, R channel with CLAHE, R channel with Gamma, B channel, B channel with CLAHE, B channel with Gamma, G channel, G channel with CLAHE, and G channel with Gamma. The following are the findings for each model and technique. When used in conjunction with DenseNet201, the green channel image and the CLAHE technique achieved the highest accuracy of 79.33%, as shown in Fig. 5, but the processing of this model took a considerable amount of time.

On the other hand, the minimum amount of time needed to calculate MobileNetV3Small was 38.68 s (Fig. 6), but the accuracy was below 75%. Selecting models and improving image quality involve making trade-offs between accuracy and processing time. MobileNet and CLAHE were combined to create the model used in this investigation, which had an accuracy of 74.9% and a training time of 111.41 s per epoch.

Comparing the accuracies of the models in each improvement technique for each color channel of a DR image, in most cases we found that the gamma technique was more accurate than the CLAHE technique for both color and red channel images. This is because the gamma technique uses a gamma matrix instead of a CLAHE matrix to calculate its improvements. However, for the green channel, the CLAHE technique was more accurate. The green channel with CLAHE also had a significantly higher accuracy for color images in some models, despite having no significant technical bearing. Most of the technical improvements demonstrated an increase in accuracy, as seen in the results and performance comparison in Table 2.[29–38]

Figure 7 shows the outcomes of using the MobileNet model on a test set organized according to the severity of the disease after the model had been trained for a total of 50 epochs. Normal subjects were identified with a precision of 96%. Among those with DR, the precision was 50% for mild DR, 68% for moderate DR, 43% for severe DR, and 59% for proliferative DR. Figure 8 illustrates an example of classification into the five grades of DR with MobileNet.

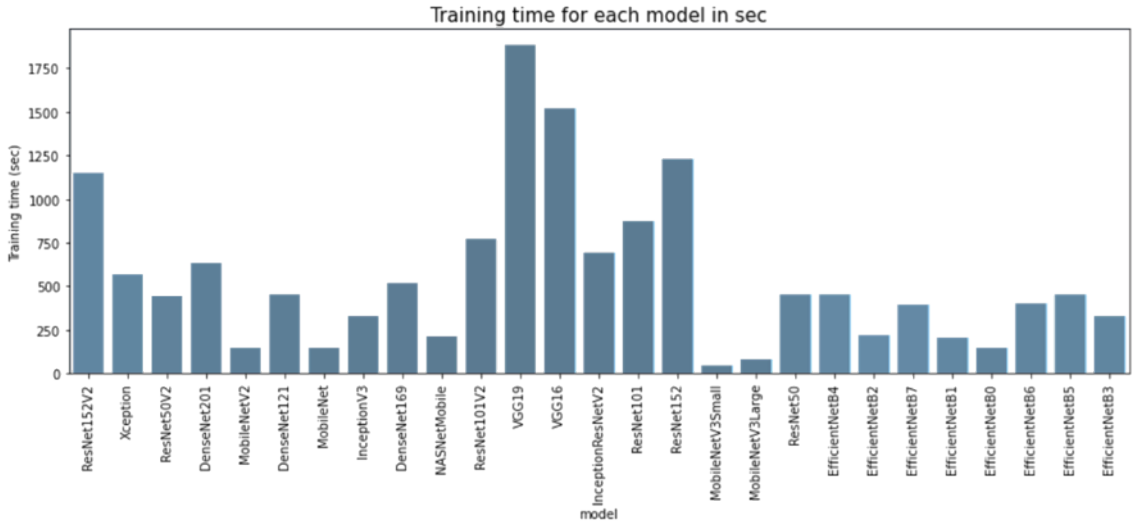



Fig. 6. (Color online) Training time for each model.

Table 2
Comparison of performance characteristics of the DR grading.

| Author (year) | Dataset (class) | Architecture | ACC (%) | SEN (%) | SP (%) | P (%) | F1 (%) |
|---|---|---|---|---|---|---|---|
| Gabriel *et al.*[29] (2017) | EyePACs (2) | VGG16 | 83.68 | 57.47 | 93.65 | — | — |
| Sraddha *et al.*[30] (2021) | DIARETDB1(2) | Custom | 98.7 | 99.6 | 98.2 | — | — |
| Ramzi *et al.*[31] (2021) | APTOS 2019(2) | ResNet | 96.35 | — | — | — | — |
| Zhan *et al.*[32] (2020) | EyePACs (2) | CF-DRNet | 83.10 | 53.99 | 91.22 | — | — |
| Ghosh *et al.*[37] (2017) | EyePACs (2) | Custom | 95 | — | — | — | — |
| Shaohua *et al.*[33] (2018) | EyePACs (5) | VGGNet | 95.68 | 86.47 | 97.43 | — | — |
| Zhitao *et al.*[34] (2021) | EyePACs (5) | SE-MIDNet | 88.24 | 99.43 | 97.6 | — | — |
| Mehedi *et al.*[35] (2022) | APTOS 2019 (5) | Custom | 92.49 | 92 | — | 93 | 93 |
| Yusaku *et al.*[36] (2020) | EyePACs (5) | Custom | 48.8 | 81.5 | 71.9 | — | — |
| Ghosh *et al.*[37] (2017) | EyePACs (5) | CNN | 85 | 95 | 88.2 | — | — |
| Shanthi *et al.*[38] (2019) | MESSIDOR (5) | Custom | 96.6 | 95.3 | 97.3 | 95.3 | — |
| Our proposed | APTOS 2019 (5) | MobileNet | 96.95 | 97.80 | 97.75 | 96.17 | 96.98 |

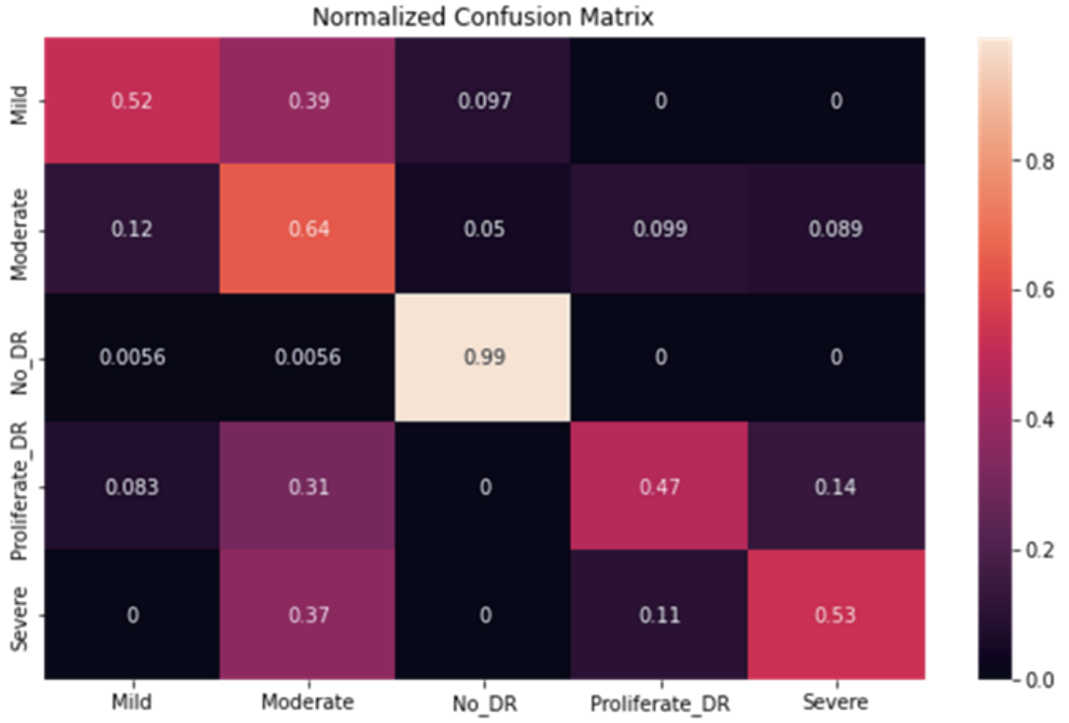



Fig. 7. (Color online) Confusion matrix of MobileNet with five classifications.

In addition, in the case that only one disease is considered from the model that came before it, Fig. 9 depicts the confusion matrix for MobileNet, which included two classifications. When evaluating retinal images, we obtained an accuracy of 96.95%, a precision of 96.17%, a sensitivity of 97.80%, a specificity of 96.98%, a specificity of 97.75%, and an F1 score of 96.98%.

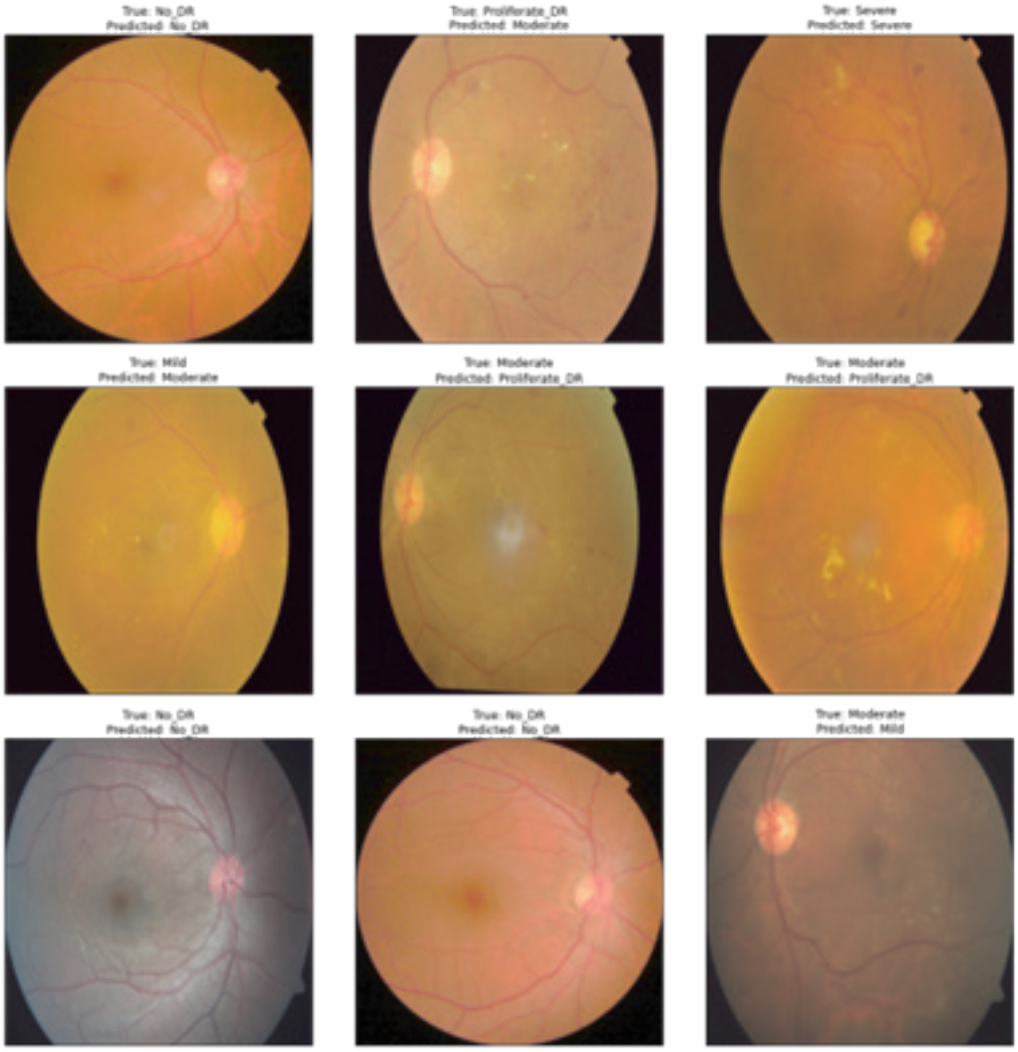



Fig. 8. (Color online) Example of classification into five DR grades with MobileNet.

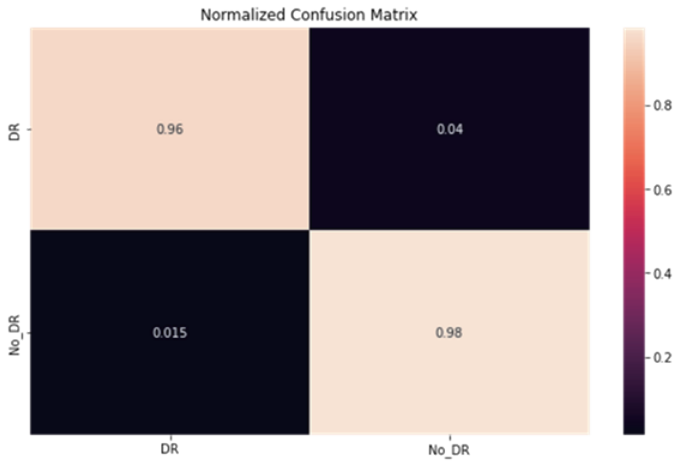



Fig. 9. (Color online) Confusion matrix of MobileNet with two classifications.

## 4. Conclusion

We proposed a method of grading DR using two distinct image enhancement techniques for each image channel and 27 pretrained CNN models for a single epoch. Using the Kaggle APTOS 2019 Blindness Detection dataset, performance was categorized as normal, moderate, mild, proliferate, or severe. The MobileNet model with CLAHE enhancement and the green channel demonstrated the best performance. The five-class DR grading system achieved high accuracy for normal images. For the two-class evaluation, the F1 score was 96.98%, the sensitivity was 97.80%, and the accuracy was 96.17%. Even though the CNN model with the enhancement technique has made more precise management and therapy possible, ophthalmologists must further improve its performance, interpretability, and reliability. In the future, the trained model in this study will be embedded into a controller or mobile phone application connected to a microscope for DR grading.

## References

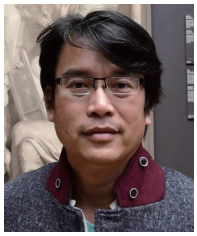

1 World Report on Vision: https://www.who.int/publications/i/item/9789241516570 (accessed July 2022).
2 M. M. Nentwich and M. W. Ulbig: World J. Diabetes **6** (2015) 489. https://doi.org/10.4239/wjd.v6.i3.489
3 K. Viswanath and D. D. McGavin: Community Eye Health **16** (2003) 21. https://europepmc.org/article/pmc/pmc1705856
4 G. Lim, V. Bellemo, Y. Xie, X. Q. Lee, M. Yip, and D. Ting: Eye Vision **7** (2020) 1. https://doi.org/10.1186/s40662-020-00182-7
5 S. M. S. Sinan, T. Tian-Swee, M. A. Wan, and H. W. H. Joyce: S. Y. S. ICT Express **8** (2022) 142. https://doi.org/10.1016/j.icte.2021.05.002
6 I. Kanjanasurat, B. Purahong, H. Aoyama, C. Benjangkaprasert, and C. Pintavirooj: Int. J. Innovative Comput. Inf. Control. **16** (2020) 879. https://doi.org/10.24507/ijicic.16.03.879
7 K. Tenghongsakul, I. Kanjanasurat, B. Purahong, and A. Lasakul: Int. Computer Science and Engineering Conf. (IEEE, 2020) 1–5. https://doi.org/10.1109/ICSEC51790.2020.9375423
8 J. Yuping, W. Huiqun, and D. Jiancheng: Proc. 1st Int. Conf. Medical and Health Informatics (ACM, 2017) 90–94. https://doi.org/10.1145/3107514.3107523
9 P. Prateek, J. Shubham, B. Neelakshi, and M. Anant: Int. Conf. Pervasive Computing Technologies for Healthcare and Workshops (ACM, 2013) 176–179. https://doi.org/10.4108/icst.pervasivehealth.2013.252093
10 J. M. Rancisco, O. Andrés, R. Javier, M. G. Juan, and C. Ricardo: Neurocomputing **452** (2021) 424. https://doi.org/10.1016/j.neucom.2020.04.148
11 S. Gayathri, P. Varun, P. Gopi, and P. Palanisamy: Biomed. Signal Process. Control. **62** (2020) 1. https://doi.org/10.1016/j.bspc.2020.102115
12 L. Xiang, J. Yuchen, Z. Jiusi, L. Minglei, L. Hao, and Y. Shen: Artif. Intell. Med. **126** (2022) 102259. https://doi.org/10.1016/j.artmed.2022.102259
13 A. Teresa, A. Guilherme, M. Luís, P. Susana, M. Carolina, C. Ângela, M. Ana, and C. Aurélio: Med. Image Anal. **63** (2020) 101715. https://doi.org/10.1016/j.media.2020.101715
14 D. Sraddha, K. Krity, M. Suchetha, R. Rajiv, and D. D. Edwin: Biomed. Signal Process. Control. **68** (2021) 102600. https://doi.org/10.1016/j.bspc.2021.102600
15 W. Zhan, S. Gonglei, C. Yang, S. Fei, C. Xinjian, C. Gouenou, Y. Jian, L. Limin, and L. Shuo: Artif. Intell. Med. **108** (2020) 101936. https://doi.org/10.1016/j.artmed.2020.101936
16 E. AbdelMaksoud, S. Barakat, and M. Elmogy: Comput. Biol. Med. **126** (2020) 104039. https://doi.org/10.1016/j.compbiomed.2020.104039
17 S. Rory, T. Ankur, R. Ehsan, B. Katy, C. David, H. Naama, K. Jonathan, N. Arunachalam, R. Zahra, R. W. Derek, X. Shawn, B. Scott, J. Anthony, S. Michael, S. Jesse, B. S. Arjun, S. C. Greg, P. Lily, and R. W. Dale: Ophthalmology **126** (2019) 552. https://doi.org/10.1016/j.ophtha.2018.11.016
18 J. de la Torre, A. Valls, and D. Puig: Neurocomputing **396** (2020) 465. https://doi.org/10.1016/j.neucom.2018.07.102

19 B. Hassan, S. Qin, T. Hassan, M. U. Akram, R. Ahmed, and N. Werghi: Biomed. Signal Process. Control. **70** (2021) 103030. https://doi.org/10.1016/j.bspc.2021.103030
20 G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger: IEEE Conf. Computer Vision and Pattern Recognition (IEEE, 2016) 4700–4708. https://doi.org/10.48550/arxiv.1608.06993
21 M. Tan and Q. Le: Int. Conf. Machine Learning (arXiv, 2019) 6105–6114. https://doi.org/10.48550/arXiv.1905.11946
22 C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi: The Thirty-First AAAI Conf. Artificial Intelligence (AAAI, 2017) 4278–4284 https://doi.org/10.48550/arxiv.1602.07261
23 C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna: IEEE Conf. Computer Vision and Pattern Recognition (IEEE, 2016) 2818–2826, https://doi.org/10.1109/CVPR.2016.308
24 A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam: arXiv preprint (arXiv, 2017) 1. https://doi.org/10.48550/arxiv.1704.04861
25 B. Zoph, V. Vasudevan, J. Shlens, and Q. V. Le: IEEE Conf. Computer Vision and Pattern Recognition (IEEE, 2018) 8679–8710. https://doi.org/10.48550/arxiv.1707.07012
26 K. He, X. Zhang, S. Ren, and J. Sun: IEEE Conf. Computer Vision and Pattern Recognition (IEEE, 2015) 770–778. https://doi.org/10.48550/arxiv.1512.03385
27 K. Simonyan and A. Zisserman: Int. Conf. Learning Representations (arXiv, 2015) 1. https://doi.org/10.48550/arxiv.1409.1556
28 F. Chollet: IEEE Conf. Computer Vision and Pattern Recognition (IEEE, 2017) 1800–1807. https://doi.org/10.1109/CVPR.2017.195
29 G. Gabriel, G. Jhair, M. Antoni, L. Jorge, and D. C. Christian: Int. Conf. Artificial Neural Networks (Springer, 2017) 635–642. https://doi.org/10.1007/978-3-319-68612-772
30 D. Sraddha, K. Krity, M. Suchetha, R. Rajiv, and D. D. Edwin: Biomed. Signal Process. Control. **68** (2021) 102600. https://doi.org/10.1016/j.bspc.2021.102600
31 A. Ramzi, M. Kahlil, and M. Novi: Procedia Comput. Sci. **179** (2021) 88. https://doi.org/10.1016/j.procs.2020.12.012
32 W. Zhan, S. Gonglei, C. Yang, S. Fei, C. Xinjian, C. Gouenou, Y. Jian, L. Limin, and L. Shuo: Artif. Intell. Med. **108** (2020) 105302. https://doi.org/10.1016/j.artmed.2020.101936
33 W. Shaohua, L. Yan, and Z. Yin: Comput. Electr. Eng. **72** (2018) 274. https://doi.org/10.1016/j.compeleceng.2018.07.042
34 X. Zhitao, Z. Yaxin, W. Jun, and Z. Xinxin: Int. Conf. Computing and Artificial Intelligence (ACM, 2021) 92–98. https://doi.org/10.1145/3467707.3467720
35 M. Mehedi, F. A. Mohammed, and Z. Yin: Multimedia Comput. Commun. Appl. **18** (ACM, 2022) 1–16. https://doi.org/10.1145/3470976
36 K. Yusaku, O. Nobuhiro, M. Kanato, O. Yoshiko, T. Kazuo, and K. Toshihide: Intell.-Based Med. **3**–**4** (2020) 1. https://doi.org/10.1016/j.ibmed.2020.100024
37 R. Ghosh, K. Ghosh, and S. Maitra: Int. Conf. Signal Processing and Integrated Networks (IEEE, 2017) 550–554. https://doi.org/10.1109/SPIN.2017.8050011
38 T. Shanthi and R. S. Sabeenian: Comput. Electr. Eng. **76** (2019) 56. https://doi.org/10.1016/j.compeleceng.2019.03.004


## About the Authors

**Isoon Kanjanasurat** received his B.S. degree from King Mongkut's Institute of Technology Ladkrabang, Thailand, in 2014 and his M.S. and Ph.D. degrees from King Mongkut's Institute of Technology Ladkrabang in 2016 and 2020, respectively. Currently, he works as a lecturer on the Computer Science Program at College of Computing, Khon Kaen University, Khon Kaen, Thailand. His research interests are in image processing and ML.
(isoonkan@kku.ac.th)

**Thanavit Anuwongpinit** received his B.S. degree from King Mongkut's Institute of Technology Ladkrabang, Thailand, in 2014 and his M.S. and Ph.D. degrees from King Mongkut's Institute of Technology Ladkrabang in 2016 and 2020, respectively. He works as a lecturer on the IoT System and Information Engineering Program at School of Engineering, King Mongkut's Institute of Technology Ladkrabang. His research interests are in IoT and embedded systems. (thanavit.an@kmitl.ac.th)



**Boonchana Purahong** received his B.S. degree from King Mongkut's Institute of Technology Ladkrabang, Thailand, in 1994 and his M.S. and Ph.D. degrees from King Mongkut's Institute of Technology Ladkrabang in 2006 and 2000, respectively. He works as a lecturer on the IoT System and Information Engineering Program at School of Engineering, King Mongkut's Institute of Technology Ladkrabang, His research interests are in embedded systems and image processing. (boonchana.pu@kmitl.ac.th)