ECTI Transactions on Computer and Information Technology
Journal homepage: https://ph01.tci-thaijo.org/index.php/ecticit/
Published by the ECTI Association, Thailand, ISSN: 2286-9131

A Comprehensive Analysis of Diabetic Retinopathy Detection in Retinal Fundus Images Using Different Convolutional Neural Network

Smita Das[1], Madhusudhan Mishra[2] and Swanirbhar Majumder[3]

ABSTRACT

Diabetic Retinopathy (DR) harms the retinal tissue's blood vessels, causing them to leak fluid and results in permanent vision loss. Therefore, rigorous discussion and evaluation of the associated procedures and findings are necessary for DR detection. In this work, the Asia Pacific Tele-Ophthalmology Society 2019 Blindness Detection (APTOS 2019 BD) dataset and the High Resolution Fundus (HRF) datasets are utilized. The Contrast Limited Adaptive Histogram Equalization is employed in the preprocessing stage to enhance the quality of the fundus images. There are various popular edge detection algorithms utilized here like Robert, Sobel, Prewitt and Canny edge detector but our experimental findings have shown that the Canny edge detector performs better than its other counterparts. So, Canny has been used for the Segmentation of fundus images. Finally, the pretrained Convolutional Neural Networks namely MobileNet, InceptionV3, ResNet-50, VGG16 and VGG19 are proposed to detect DR in retinal fundus images. To evaluate the effectiveness of the proposed model, both datasets are divided into two parts, where one part is utilized for training the proposed model and another part is utilized for testing the proposed model. The performances of different techniques are evaluated based on the standard performance parameters like Sensitivity, Specificity, Precision, Accuracy and Receiver Operating Characteristic Curve etc.

Article information:

Keywords: Diabetic Retinopathy, Fundus Image, Preprocessing, Canny Edge Detector, Convolutional Neural Network

Article history:
Received: August 3, 2023
Revised: September 21, 2023
Accepted: October 26, 2023
Published: November 25, 2023 (Online)

DOI: 10.37936/ecti-cit.2023174.253699

1. INTRODUCTION

One of the most severe side effects of diabetes is diabetic retinopathy (DR), which damages the retina and results in blindness. It affects the blood vessels in the retinal tissue, which leads to fluid leakage and eventually irreversible vision loss [1]. It is a form of metabolic disease that manifests as a result of excessive blood glucose levels, which may impair insulin secretion and cause vision problems [2]. By 2030, it is predicted to affect 191 million people, making it the main factor in vision loss and permanent blindness in middle-aged and older people [1]. DR is diagnosed visually by looking at retinal images to find the three main pathological symptoms, namely exudates, hemorrhage and microaneurysm, that are most frequently present. The sign of the appearance of exudates on the retina has emerged as a crucial clinical marker for automated disease recognition and diagnosis. Figure 1 shows the entire retina image. The fundus image on the left side depicts the normal retina condition and the fundus image on the right side depicts a retina with diabetic retinopathy symptoms [3].

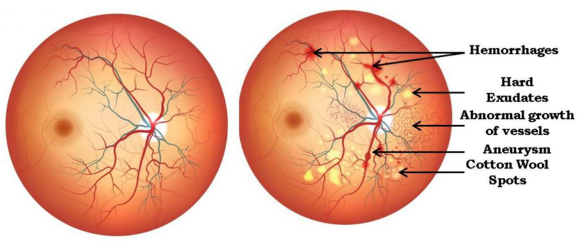


Fig.1: *Normal/Healthy Retina and Diabetic Retinopathy affected Retina.*

[1,3] The authors are with Department of Information Technology, Tripura University, Suryamaninagar, Tripura, India, E-mail: ersmitadas@gmail.com and swanirbhar@ieee.org
[2] The author is with Department of Electronics and Communication Engineering, NERIST (Deemed to be University), India, E-mail: ecmadhusudhan@gmail.com

Computer aided detection technique assists ophthalmologists in the detection of DR by automatically extracting the features of retinal fundus images. Preprocessing, Segmentation and classification are the three main steps in this Computer aided detection technique. In this paper, Contrast Limited Adaptive Histogram Equalization(CLAHE) is employed in the preprocessing stage to enhance the contrast between DR pathological signs and background [4]. Next, the images are segmented to differentiate between normal and abnormal cases. It involves separating the foreground from the background or grouping pixels together based on various colour or shape similarities. Canny edge detector has been used here for the Segmentation of fundus images.

Deep learning algorithms can be used in a real-time automated system to identify the early symptoms of DR disease. As a result, it is simple to reduce both the likelihood of human error and the effort required by the ophthalmologist. The classification of dark and bright lesions in the retinal fundus image is the sole basis for this diagnosis. So, pretrained CNN, namely MobileNet, InceptionV3, ResNet-50, VGG (Visual Graphics Group)16 and VGG19 have been proposed here for detecting DR in retinal fundus images that classify fundus images by automatically extracting the features. The Asia Pacific Tele-Ophthalmology Society 2019 Blindness Detection (APTOS-2019 BD) dataset [5] and the High Resolution Fundus (HRF) dataset [6] are utilized in this study. To evaluate the effectiveness of the proposed model, the datasets were divided into two parts: one for training the model and the other for testing the model. The performance of different techniques has been evaluated based on different performance parameters.

The paper is organized as follows: The next section presents related work on diabetic retinopathy detection. Section 3 illustrates the proposed methods. In section 4, the experiment with the results of the proposed algorithm is presented. Finally, a conclusion is presented in section 5.

2. RELATED WORKS

There are numerous challenges which interrupt the retinal blood vessel extraction process. A literature survey on this topic will provide a clear scenario of the existing techniques used in the detection of ophthalmologic diseases. It gives valuable resources toward the extraction or Segmentation of retinal images. Reveal statistical knowledge exists in the relevant research area. Increase statistical knowledge related to the research topic. A method to recognize microaneurysms, exudates and hemorrhages was employed by Khojasteh et al.[1] utilizing a probabilistic Convolutional Neural Network (CNN). Pretrained CNN was used by Mateen et al. [3] to identify exudates. Nneji et al. [4] identified DR using the weighted fusion deep learning based on Dual-channel Fundus scans. Ooi et al. [7] applied Contrast Limited Adaptive Histogram Equalization(CLAHE) on raw images, followed by the Canny edge detection technique for Segmentation. Wang et al. [8] used three kinds of region growing methods in order to segment the original image. Kusakunniran et al. [9] Proposed vessel segmentation based on Hybrid learning by combining supervised and instance learning steps. Li et al. [10] introduced an approach that makes use of teaching learning-based optimization, Support Vector Machines, and Deep CNN. CNN was utilized for the categorization of fundus images by Bulut et al. [11], where customized networks can be made by altering the output layers of CNN classifiers trained on large data sets like ImageNet and Coco. To discriminate between various vascular diseases and healthy controls on fundus images, Abitbol et al. [12] used a CNN classifier. Sungheetha et al. [13] designed an early detection and classification of DR by deep feature extraction based CNN. The study by Mushtaq et al. [14] investigates a deep learning methodology for the early identification of DR using the densely connected convolutional network DenseNet-169. Khalifa et al. [15] utilize AlexNet, Res-Net18, SqueezeNet, GoogleNet, VGG16, and VGG19 for the classification of APTOS data. Gupta et al. [16] proposed an Optimized hybrid machine learning approach for smartphone based diabetic retinopathy detection. Saranya et al. [17] do the Blood vessel segmentation in retinal fundus images for proliferative diabetic retinopathy screening using deep learning. Qureshi et al. [18] proposed an active deep learning method for the classification of fundus images. Bilal et al. [19] proposed an Artificial Intelligence based classification of DR using U-Net and deep learning. Sundaram et al. [20] used a hybrid segmentation approach for the extraction of retinal blood vessels. Albahli et al. [21] detected DR using Custom CNN to Segment the Lesions. Yazid et al.[22] proposed CNN for the detection. Akbar et al. [23] proposed a novel 3D-CNN based feature extraction based classification for diabetic retinopathy detection. Ayala et al. [24] proposed deep learning for improved detection of DR. Das et al. [25] proposed CNN for Feature Extraction and Classification for the detection of DR. Yasashvini et al. [26] presented CNN and Hybrid Deep CNN for Diabetic Retinopathy Classification. Devi et al. [27] proposed a deep transfer learning approach for the identification of diabetic retinopathy using data augmentation. Ismail et al. [28] applied Bayesian deep learning methods to analyze diabetic retinopathy disease. Datta et al. [29] proposed a review work on Hyper parameter tuning based gradient boosting algorithm for the detection of diabetic retinopathy. Wasekar et al. [30] analyzed supervised learning methodologies for the detection of exudates in diabetic retinopathy. This literature evaluation reveals the research's connected question's originality and significance. It highlights inconsistencies, including gaps in the literature, disagreements in earlier studies, and unanswered questions. While CNN-based approaches have achieved significant advancements in the diagnosis of DR, they still possess significant limitations, which are detailed below:

(i) Most of the present CNN models for DR diagnosis only focused on DR grading without determining the location of various DR lesions in the fundus image. They carried out end-to-end operations, which refer to a process in which the input image is fed directly into CNN, which subsequently outputs the image along with the DR severity level.

(ii) Modern CNN models require a larger collection of annotated samples for learning, which is an expensive and time-consuming task. In contrast, a CNN model that can intelligently learn from a small number of samples is more practical and necessary in today's ophthalmology for DR classification.

(iii) The developed CNN architectures failed to learn the complicated structures of DR lesions.

Making a case for the necessity of additional research using the concepts of existing literature, we plan new, fresh and original research ideas. So, the goal of the proposed approach is to solve the issue of low quality fundus images. The proposed approach was designed to help single models improve their shortcomings and minimize their defects by extracting the characteristics. With this approach, prediction variance is minimized while generalization error is reduced. Particularly when the dataset is noisy or small, deep learning algorithms may be prone to overfitting. A model is considered to have been overfitted when it becomes excessively specialized to the training set and will perform badly on new data. However, regularisation and data augmentations are widely employed to avoid overfitting. As a result, the objective of this study is to evaluate how well CNN models for DR detection might help ophthalmologists performance less stressfully and efficiently with the addition of preprocessing and Segmentation. The novelty of our proposed model includes better feature extraction through the use of fine-tuning approaches. In order to determine the quantity of CNN's accuracy when employing different deep learning models to recognize the DR, a study was conducted in this work.

3. THE PROPOSED METHODS

In order to fulfill our objectives, the Contrast Limited Adaptive Histogram Equalization (CLAHE) is employed here in the preprocessing stage to enhance the quality of fundus images, followed by Segmentation in order to obtain the regions of the fundus image that corresponds to the blood vessels and finally, CNN based image classification which assigns a label or class to an input image. So, this section is divided into four sub-sections. First Section 3.1 will describe details regarding the dataset used in this analysis. Then, Section 3.2 will introduce CLAHE, which is the basis of the Preprocessing approach. Next, Section 3.3 will introduce the Robert Operator, Sobel Operator, Prewitt Operator and Canny edge detector used for the Segmentation of the fundus images. In the last section 3.4, CNN namely MobileNet, InceptionV3, ResNet-50, VGG16 and VGG19 are proposed to detect DR in retinal fundus images. Figure 2 shows the basic Architecture of the proposed method.

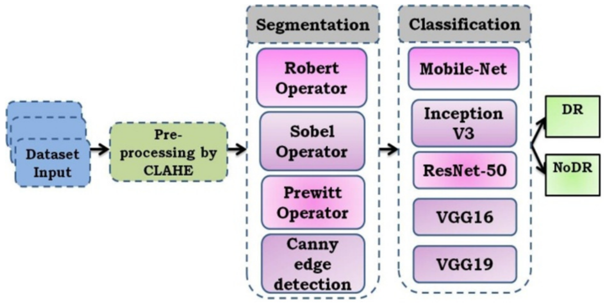


Fig.2: *Architecture of the proposed technique.*

3.1 Dataset

The Asia Pacific Tele-Ophthalmology Society 2019 Blindness Detection (APTOS-2019 BD) dataset [5], and the High Resolution Fundus (HRF) dataset [6] are utilized in this study. The APTOS dataset contains 3662 number of images of 2 classes, whereas the HRF dataset contains 30 images only. This APTOS fundus image dataset was compiled by India's Aravind Eye Hospital. A team of medical professionals evaluated and classified the samples collected using the International Clinical Diabetic Retinopathy Disease Severity Scale (ICDRSS). The APTOS dataset was divided into two parts: one for training the model and the other for testing the model. To evaluate the effectiveness of all the models, three experiments were conducted by each model with different dataset splits: 90% for training and 10% for testing, 80% for training and 20% for testing, and 70% for training and 30% for testing.

3.2 Preprocessing

In order to fix problems such as inadequate contrast, low quality and distortion in images caused by focusing and improper lighting, a preprocessing technique can be used to improve the quality of the images. This could improve the images suitability for accurate and efficient mining of valuable data [16]. During this preprocessing stage, to enhance the quality of fundus images, the CLAHE technique is used in this work [2,31]. Steps involved in applying CLAHE to preprocess fundus images include image acquisition, grayscale conversion, dividing the image into tiles, applying histogram equalization, contrast limiting, combining tiles, post-processing and finally, result assessment. Figure 3 shows the steps involved in the CLAHE technique. With high accuracy and contrast limiting, CLAHE works on discrete, smaller sections of the image, known as tiles, rather than the complete image.

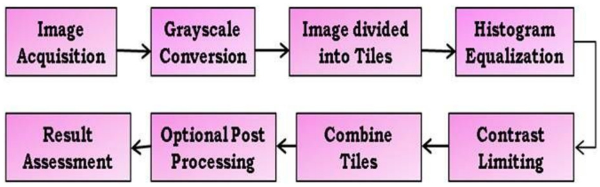


Fig.3: *Steps involved in CLAHE.*

To apply the CLAHE technique, the entire image is first divided into smaller sections or clips to create Adaptive Histogram Equalisation (AHE) clips. The histogram is then clipped at a predetermined value known as the clip limit by AHE, which helps limit amplification in CLAHE. This value determines the amount of noise in the histogram that should be reduced in order to improve contrast in the image. Figure 4 shows the raw image and outcomes of the CLAHE based image preprocessing, whereas Figure 4(a) presents the raw image, Figure 4(b) presents coloured processed image and Fig 4(c) presents the grayscale processed image. From Figure 4(a), 4(b) and 4(c), it can be seen that by applying the CLAHE based preprocessing on HRF image, there is a notable improvement in the quality of the original images, for example, the vessels and spots in the images become more identical and sharpened in the processed images [2, 31].

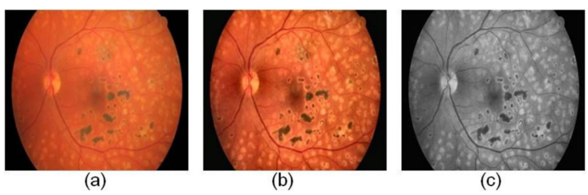


Fig.4: *Raw image and outcomes of the CLAHE based image preprocessing (a) Raw Image before preprocessing, (b) Coloured CLAHE'd image after preprocessing, (c) Grayscale CLAHE'd image after preprocessing.*

3.3 Segmentation

Image segmentation is a technique for breaking up a digital image into smaller groups called image segments, which reduces the complexity of the image and makes each segment more easily processed or analyzed. Segmentation involves applying a filter to the image, classifying pixels as being on or off edges based on the filter's output, and assigning all pixels that are not separated by edges to the same category. There are various popular edge detection algorithms utilized here like Robert, Sobel, Prewitt and Canny edge detector. In order to determine which segmentation technique can precisely segment the blood vessels of fundus images, an exploratory study on Robert, Sobel, Prewitt and Canny edge detector was carried out in this research work. Robert operator is computationally less expensive but it is not effective in detecting subtleties or handling noisy images. In fundus image segmentation, it missed fine details and produced a lot of false positives in the presence of noise. The Sobel operator is more effective at capturing edges with varying orientations and is less sensitive to noise as compared to the Robert, but still struggles with complex structures and noisy images. Prewitt can be used as an alternative to Sobel, but it does not offer a significant advantage in most cases. Canny involves multiple stages including Gaussian smoothing, gradient calculation, nonmaximum suppression, double thresholding and edge tracking by hysteresis and is able to detect edges accurately by suppressing noise effectively and providing continuous and well connected edge contours. It can handle complex structures and noise images better than simpler operators like Robert, Sobel or Prewitt. So, the Canny edge detector has been used here for Segmentation of fundus images. Figure 5 shows CLAHE'd image and outcomes of all the segmentation approach, whereas Figure 5(a) shows CLAHE'd image before Segmentation, Figure 5(b) shows output image by Robert operator, Figure 5(c) shows output image by Sobel operator, Figure 5(d) shows output image by Prewitt operator and Figure 5(e) shows output image by Canny edge detector.

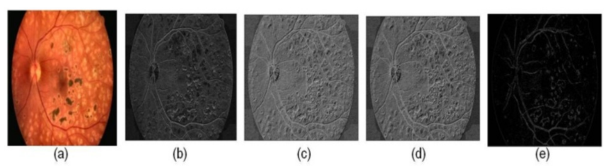


Fig.5: *CLAHE'd image and Outcomes of all the segmentation approaches (a) CLAHE'd image before Segmentation, (b)Output image after Segmentation by Robert Operator, (c) Output image after Segmentation by Sobel Operator, (d) Output image after Segmentation by Prewitt operator and (e) Output image after Segmentation by Canny edge detector.*

Implementation of Canny Edge Detection: The Canny edge detection algorithm consists of 5 steps. Figure 6 shows all the steps to implement it.

i. Noise reduction/ Smoothing: In this experiment, a Gaussian Kernel of 5x5 is used with the image convolution technique to achieve smoothing. The equation for a 2D Gaussian function has the form:

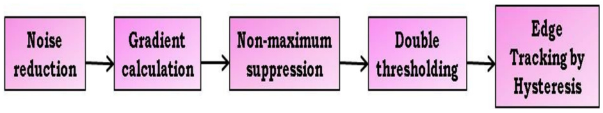


Fig.6: *Flowchart of Canny Edge Detection.*

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{1}$$

Where σ is the variance of the Gaussian function, $x = x$ coordinate value, $y = y$ coordinate value and $\pi =$ Mathematical Constant Pi.

ii. Gradient calculation/ Finding Gradients: The derivatives I_x and I_y with respect to x and y are determined after the image has been smoothed. It can be implemented by convolving I with Sobel kernels K_x and K_y, respectively:

$$K_x = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix}, K_y = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix}$$

Then, the magnitude G and the slope θ of the gradient are calculated as follows:

$$G = \sqrt{I_x^2 + I_y^2}, \tag{2}$$

$$\theta(x,y) = \arctan\left(\frac{I_y}{I_x}\right) \tag{3}$$

Where I_x and I_y are the gradients in the x and y directions respectively.

iii. Non-maximum suppression: During this step, every point on the gradient intensity matrix is examined by the algorithm, which then determines the pixels with the highest values in the direction of the east. Only the edge pixels that have local gradient maxima are kept. This procedure produces an image with sharper edges while retaining all of the important edge information.

iv. Double thresholding: The double thresholding stage separates pixels into three groups: strong, weak, and irrelevant. Strong pixels are those in the image with edge strengths above the high threshold, whereas irrelevant pixels are those with edge strengths below the low threshold. Week pixels are those with edge strength in between these two criteria.

v. Edge Tracking by Hysteresis: The hysteresis technique converts low-intensity pixels to high-intensity pixels only if at least one of the surrounding pixels has already been classified as a high-intensity pixel based on the threshold results.

Figure 7 shows the healthy image and the resulting image of the Canny edge detection technique, whereas Figure 7(a) shows a Normal/Healthy image, Figure 7(b) shows Smoothed Image, Figure 7(c) shows a Gradient magnitude, Figure 7(d) shows edges after non-maximum suppression, Figure 7(e) shows Double thresholding and Figure 7(f) shows Edge tracking by hysteresis.

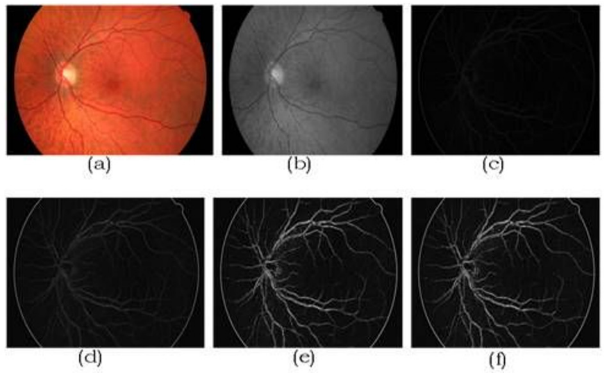


Fig.7: *Normal/Healthy image and resulted image of Canny edge detection based Segmentation (a) Normal/Healthy image before Canny based Segmentation, (b) Smoothed Image of Canny based Segmentation, (c) Output image after applying Gradient magnitude, (d) Output image after nonmaximum suppression, (e) Output image after Double thresholding, (f) Output image after Edge tracking by hysteresis.*

Figure 8 shows DR effected image and the resulting image of the Canny edge detection technique, whereas, Figure 8(a) shows DR effected image, Figure 8(b) shows the Smoothed Image, Figure 8(c) shows the Gradient magnitude, Figure 8(d) shows edges after non maximum suppression, Figure 8(e) shows Double thresholding and Figure 8(f) shows Edge tracking by hysteresis.

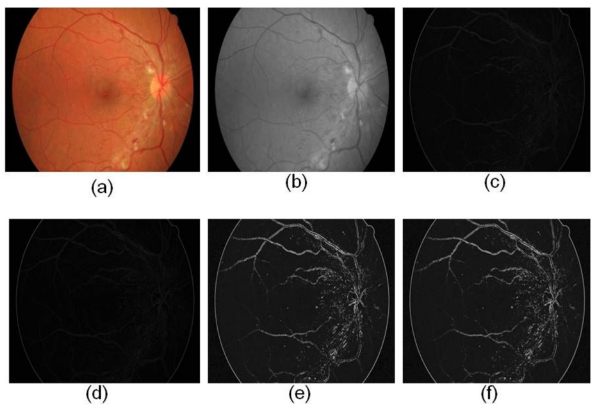


Fig.8: *DR effected image and resulted image of Canny edge detection based Segmentation (a) DR effected image before Canny based Segmentation, (b) Smoothed Image of Canny based Segmentation, (c) Output image after applying Gradient magnitude, (d) Output image after nonmaximum suppression, (e) Output image after Double thresholding, (f) Output image after Edge tracking by hysteresis.*

3.4 Classification by Convolutional Neural Networks (CNN):

Image classification is the task of assigning a label or class to an input image. Convolutional neural networks are a subclass of neural networks that are mostly employed in voice and image recognition applications. With no loss of information, its integrated convolutional layer lowers the high dimensionality of images. CNNs are therefore very well suited for this use case. Five popular models of CNN architectures MobileNet, InceptionV3, ResNet-50, VGG16 and VGG19 have been used for the classification of retinal fundus images to detect DR in retinal fundus images. The choice of MobileNet, InceptionV3, ResNet-50, VGG16, or VGG19 for retinal fundus image classification is only for their strong performance in computer vision tasks, their adaptability through transfer learning, their availability in the deep learning community, their architecture diversity and model size.

MobileNet: It is a novel type of CNN model which consists of i) One standard Convolutional Layer of 3x3 convolutions along with Rectified Linear Unit(ReLU) and batch normalization, ii) One 3x3 Depthwise separable Convolutional Layer and one 1x1 Convolutional Layer with batch normalization and ReLU, iii) Width Multiplier and iv) Resolution Multiplier.

Inception-v3: The Factorized 7×7 convolutions, Label Smoothing, and an auxiliary classifier are the main components of the Inception-v3 CNN architecture. The model's larger Convolutions were factorized into smaller Convolutions to make it better.

ResNet-50: ResNet-50 (version of a ResNet model) contains One Average Pool layer, One Max-Pool layer and 48 Convolution layers. The ResNet-50 model specifically has 5 phases, each containing a residual block. There are three layers in each residual block, each with 1x1 and 3x3 convolution layers.

VGG16: The VGG is a common, multi-layer deep CNN. VGG is possible with transfer learning, in which the parameters are changed for greater precision. This model improved the state-of-the-art setups by increasing the depth utilizing architecture with very small (3x3) convolution filters. It consists of a large number of layers, consisting mainly of 3x3 convolutional filters and pooling layers.

VGG19: A VGG model called VGG19 consists of 19 layers in total (1 SoftMax layer, 5 MaxPool layers, 16 convolution layers and three fully connected layers). The VGG19 model's principle is identical to that of the VGG16, with the exception that it supports 19 layers. The model VGG19 is an upgrade over VGG16.

The presented CNN models classify fundas images by automatically extracting features. To conduct a thorough analysis of the performance of DR detection using deep learning models, it is important to take into account various factors, including accuracy, specificity, sensitivity, precision, ROC curve, confusion matrix etc. Our analysis evaluates the performance of the model using parameters such as sensitivity, specificity, precision, accuracy and ROC curve.

4. RESULTS AND DISCUSSION

Our objective is to conduct a comprehensive analysis of diabetic retinopathy detection in retinal fundus images using various pretrained CNN. CLAHE is used to improve the visibility of edges and other details in fundus images. It is seen that after applying the CLAHE based preprocessing, the quality of the raw images is significantly improved; for instance, the vessels and spots are sharpened in the processed images. In this case, a clip limit of 2.0 was utilized with a tile grid size of 8x8. However, we also investigated the use of other clip limits and tile grid sizes. The experimental results indicate that a clip limit of 2.0 with a tile grid size of 8x8 performs the best in this scenario. After applying the CLAHE algorithm, the resulting images were further processed using edge detection techniques to segment the image into distinct regions. There are various popular edge detection algorithms utilized here like Robert, Sobel, Prewitt and Canny edge detector but our experimental findings have shown that the Canny edge detector performs better than its other counterparts. When using the Canny edge detection algorithm, the gradient magnitude image is tested by two threshold values to produce the output edges. In the case of segmenting APTOS images, the values of low threshold and high threshold were set to 220 and 255, respectively, after trying different combinations of threshold values. This means that any pixel with a gradient magnitude value of 255 or higher is considered to be an edge pixel, while pixels with a gradient magnitude value below 220 are discarded as non-edge pixels. Pixels with a gradient magnitude value between 220 and 255 are retained as weak edge pixels and may be included in the final edge map if they are connected to strong edges.

Finally, we proposed five models of deep learning architectures, MobileNet, InceptionV3, ResNet-50, VGG16 and VGG19 to detect DR in retinal fundus images. These models are pre-trained on large image datasets such as ImageNet and are able to recognize a wide range of visual patterns and features. Different performance parameters such as sensitivity, specificity, precision, accuracy and ROC curve were used to evaluate the performance of the different models. The performance of the models was found to fluctuate across repeated experiments, which can be due to a range of factors, such as the quality of the training data, the choice of hyperparameters and the random initialization of the model weights. To account for this variability, multiple experiments are often conducted and the results are averaged to obtain a more reliable estimate of the model's performance. In this experiment, each of the models was trained for 10 epochs with a batch size of 16. The APTOS dataset was divided into two parts: one for training the model and the other for testing the model. To evaluate the effectiveness of the models, three experiments were conducted by each model with different dataset splits: 90% for training and 10% for testing, 80% for training and 20% for testing, and 70% for training and 30% for testing. Overall, the use of pre-trained CNNs for retinal fundus image classification is a promising approach, and the choice of model and dataset split can have a significant impact on the performance of the model. Further experimentation and fine-tuning of the models may be necessary to optimize their performance for specific applications.

Figure 9 illustrates the loss and accuracy of all models on the APTOS dataset split of 90% for training and 10% for testing, whereas Figure 9(a) shows training loss per epoch, Figure 9(b) shows training accuracy per epoch, Figure 9(c) shows validation loss per epoch and Figure 9(d) shows validation accuracy per epoch. Figure 10 illustrates the loss and accuracy of all models on the APTOS dataset split of 80% for training and 20% for testing, whereas Figure 10(a) shows training loss per epoch, Figure 10(b) shows training accuracy per epoch, Figure 10(c) shows validation loss per epoch and Figure 10(d) shows validation accuracy per epoch. Figure 11 illustrates the loss and accuracy of all models on the APTOS dataset split of 70% for training and 30% for testing, whereas Figure 11(a) shows training loss per epoch, Figure 11(b) shows training accuracy per epoch, Figure 11(c) shows validation loss per epoch and Figure 11(d) shows validation accuracy per epoch. The evolution of the loss function during the training of a model can be seen in a graph of training loss per epoch. The loss function is a measurement of the model's effectiveness at a specific task. In contrast, A CNN's validation loss per epoch graph shows how the model's loss on a different validation dataset changes during the course of training. During each epoch of training, the validation loss represents how well the model is performing on fresh and untested data. A CNN's training accuracy per epochs graph illustrates how the accuracy of the model on the training data changes throughout the period of training. The accuracy is the percentage of samples in the training dataset that were correctly categorized at each epoch. In contrast, a CNN graph of validation accuracy per epoch graph shows how the model's performance on a different validation dataset evolves throughout the period of training. The x-axis of the graph represents the epochs, which refers to the number of times the model has iterated through the entire dataset during training, whereas the y-axis represents the training loss, training accuracy, validation loss and validation accuracy respectively.

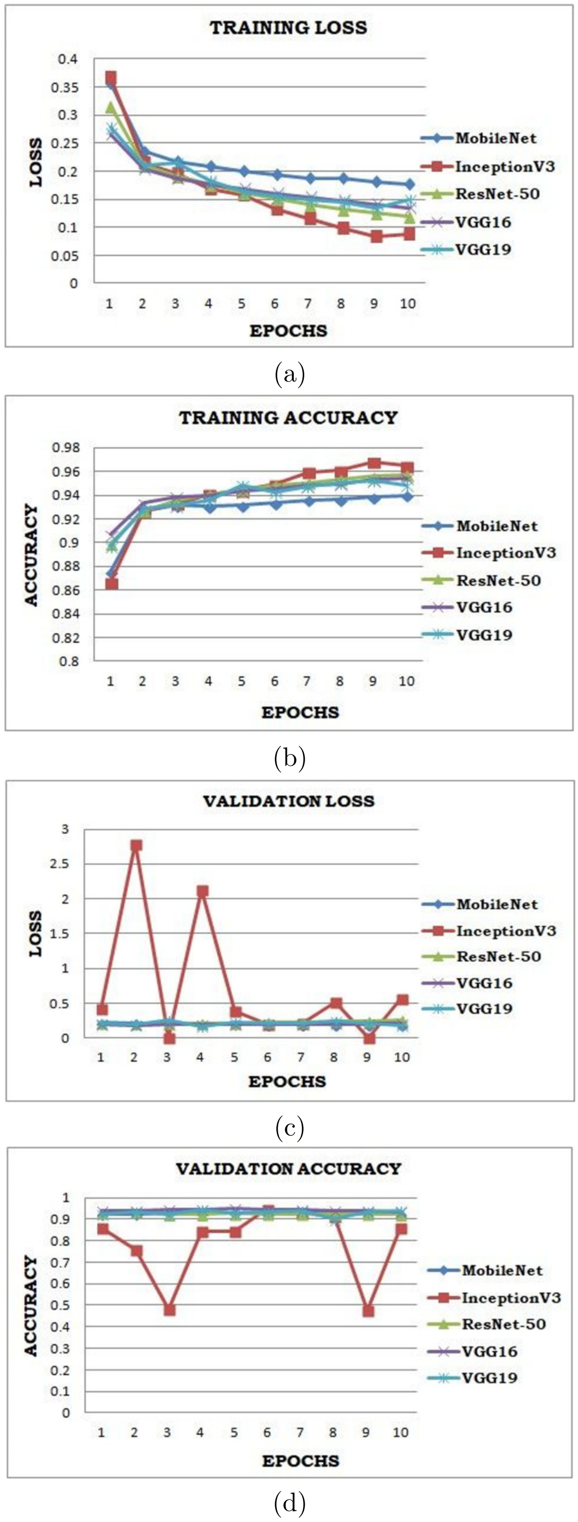
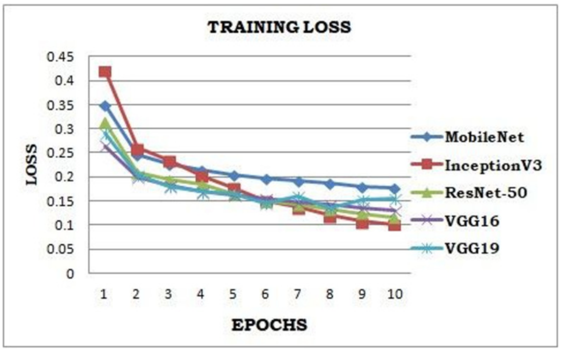
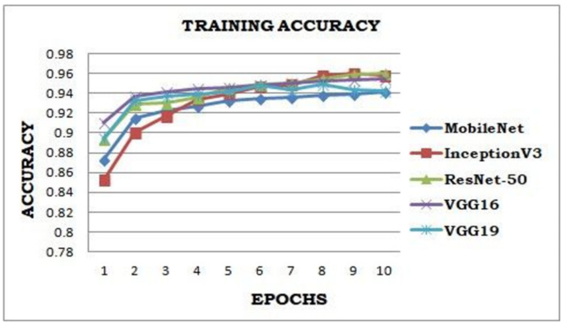


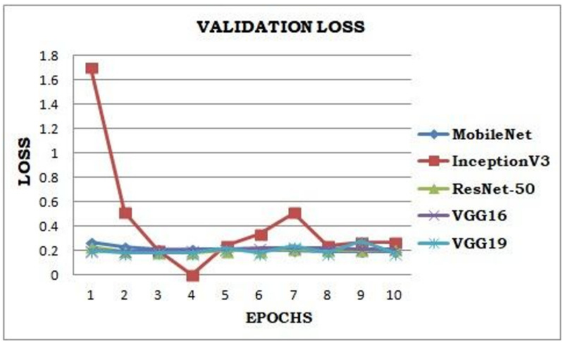
Fig.9: *The loss and accuracy of all models on the APTOS dataset split of 90% for training and 10% for testing, (a) shows training loss per epoch, (b) shows training accuracy per epoch, (c) shows validation loss per epoch and (d) shows validation accuracy per epoch.*



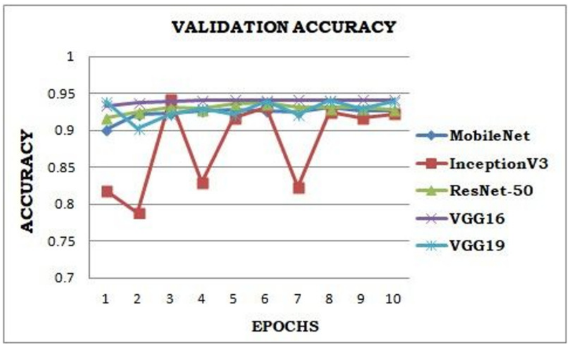
(a)



(b)

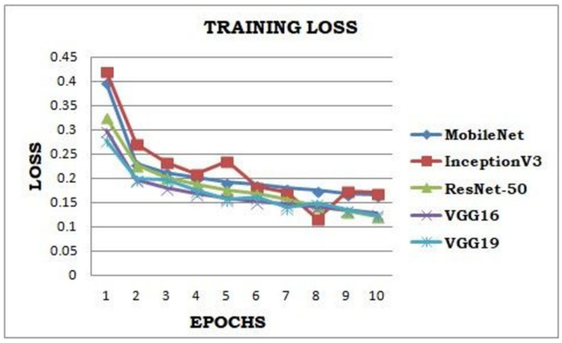


(c)

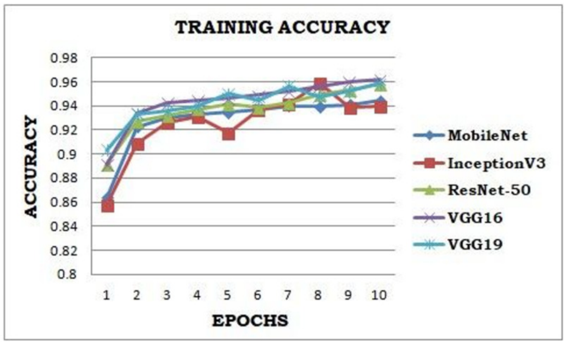


(d)

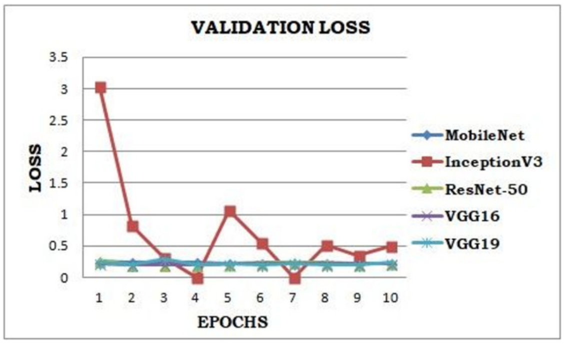
Fig.10: *The loss and accuracy of all models on the APTOS dataset split of 80% for training and 20% for testing, (a) shows training loss per epoch, (b) shows training accuracy per epoch, (c) shows validation loss per epoch and (d) shows validation accuracy per epoch.*



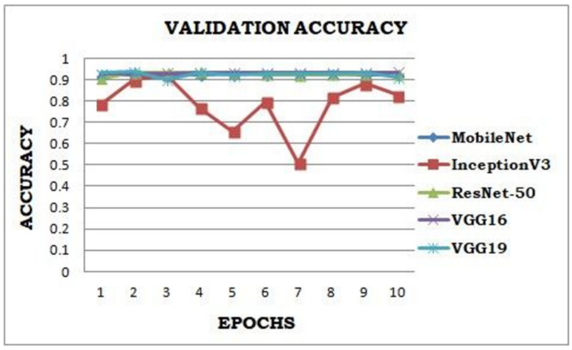
(a)



(b)



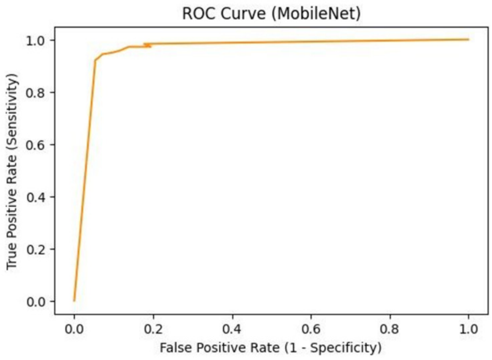
(c)



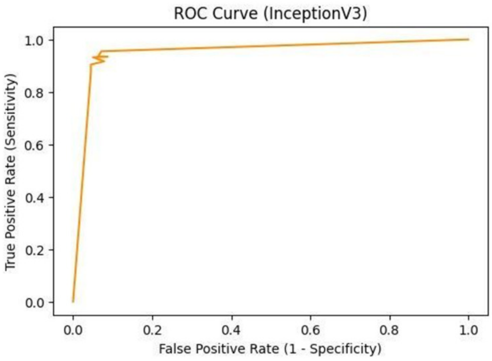
(d)

Fig.11: *The loss and accuracy of all models on the APTOS dataset split of 70% for training and 30% for testing, (a) shows training loss per epoch, (b) shows training accuracy per epoch, (c) shows validation loss per epoch and (d) shows validation accuracy per epoch.*

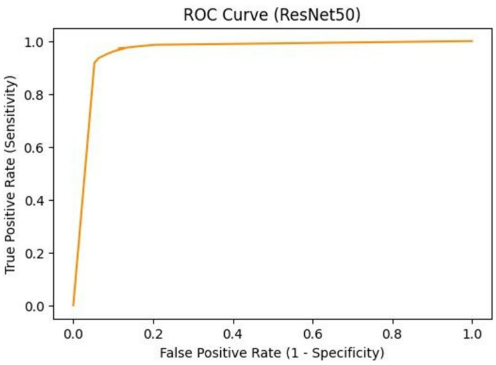
In deep learning, Keras offers an estimate of the accuracy while the model is being trained. The threshold used by Keras for this estimation is 0.5. This is so because the decision boundary between the two classes is 0.5. In order to understand the trade-offs between the different metrics and select the threshold that best satisfies the particular requirements of the problem, it is necessary to evaluate the model's performance across a range of threshold values and visualize the results using appropriate performance. This analysis shows different results for different threshold values ranging from 0.1 to 0.9. Figure 12 shows the ROC curve. Where Fig 12(a) shows the ROC curve of the MobileNet model, Fig 12(b) shows the ROC curve of the InceptionV3 model, Fig 12(c) shows the ROC curve of the ResNet50 model, Fig 12(d) shows the ROC curve of the VGG16 model and Fig 12(e) shows the ROC curve of the VGG19 model.



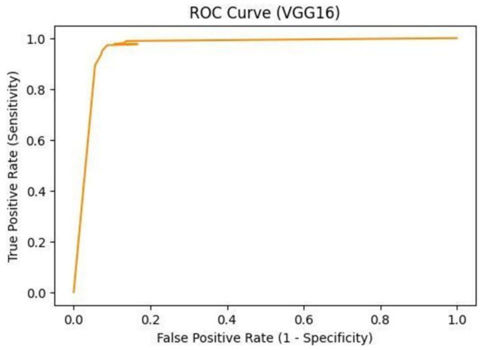
(a)



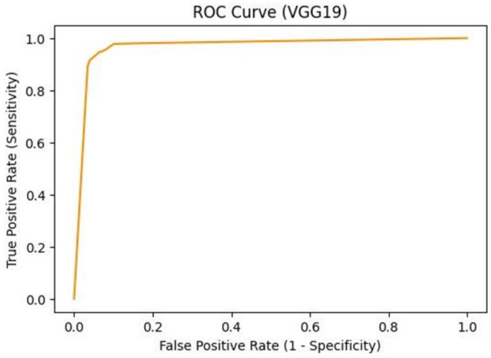
(b)



(c)



(d)



(e)

Fig.12: *ROC curve of all models. (a) shows the ROC curve of the MobileNet model, (b) shows the ROC curve of the InceptionV3 model, (c) shows the ROC curve of the ResNet50 model, (d) shows the ROC curve of the VGG16 model and (e) shows the ROC curve of the VGG19 model.*

However, here is a general comparison [32, 33] of the performance parameter of the models: MobileNet has been shown to have an accuracy of upto 0.9358, whereas InceptionV3 has been shown to perform very well on DR detection tasks, with an accuracy of upto 0.9431. ResNet-50 is another deep learning model that has been widely used for DR detection and has achieved accuracy upto 0.9373 in this analysis. VGG16 and VGG19 have also shown good accuracy upto 0.9413 and 0.9503 respectively. The model with an accuracy of 0.9503 suggests that it is the best-performing model and could have been trained on a larger dataset with more complex architecture or optimized using advanced techniques. In summary, MobileNet, InceptionV3, ResNet-50, VGG16 and VGG19 are generally considered to be highly accurate models for DR detection.

However, it is important to note that not all models should be judged on their accuracy alone. The performance of the model is further evaluated by additional parameters like specificity, sensitivity, precision etc and the final model selection may also be influenced by other aspects like computational effectiveness and ease of implementation. The better a model performs, in general, the higher its accuracy,

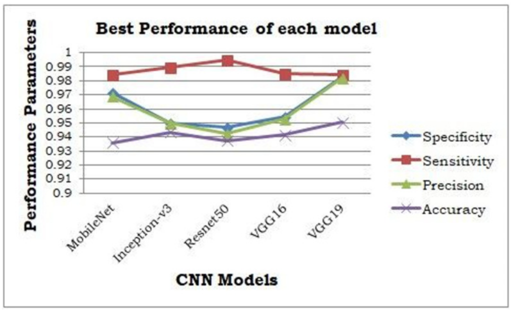


Fig.13: *Performance of all the proposed models.*

precision, sensitivity, and specificity. Figure 13 depicts details of the best specificity, sensitivity, precision and accuracy achieved by the model. It is noted that all the models MobileNet, InceptionV3, ResNet-50, VGG16 and VGG19 show precision upto 0.9687, 0.9497, 0.9421, 0.9526 and 0.9818 respectively, which indicates that all models are performing well for DR detection in retinal fundus images. However, there are some variations in their precision. Due to a simpler architecture or insufficient training data, the model with precision values of 0.9421 and 0.9497 can have marginally lower precision than other models. This indicates that certain non-DR images may be labeled as DR by the model. Models with a precision of 0.9687 and 0.9526 perform moderately. Models with a precision of 0.9818 might have performed better with better hyperparameters. This indicates that it is the best-performing model since more DR images are being accurately identified by the models, while fewer non-DR images are being categorized as DR. All the models MobileNet, InceptionV3, ResNet-50, VGG16 and VGG19 show sensitivity upto 0.9843, 0.9895, 0.9947, 0.9850 and 0.9843 respectively, which indicates that all models are performing well for DR detection in retinal fundus images. This means that the model is correctly identifying more DR images, and fewer DR images are being classified as non-DR. All the models show specificity upto 0.9714, 0.9497, 0.9470, 0.9542 and 0.9828 respectively, which indicates that all models are performing well in this analysis and are able to correctly identify a high proportion of negative cases. However, there are some differences in their specificity. Specifically, the model with a specificity value of 0.9828 is the most accurate in correctly identifying negative cases, while the model with a specificity value of 0.9470 is the least accurate among the five models.

Based on the experimental comparison, we can conclude which deep learning model is best suited for DR detection. The comparison can also help identify areas for future research and optimization to improve the performance of deep learning models for DR detection. The deep architecture of VGG19 is distinguished by a large number of trainable parameters and a uniform filter size. In contrast to larger filters, the network can collect more detailed features by the repetitive stacking of convolutional layers and by using smaller filters to maintain spatial information. All of these features make VGG19 slightly better than other models in the classification of APTOS Data.

5. CONCLUSIONS

In this analysis, CLAHE is used to improve the visibility of edges and other details in digital images in the preprocessing stage. After applying the CLAHE algorithm, the resulting images were further processed using Canny edge detection techniques for Segmentation of fundus images. Finally, pretrained Convolutional Neural Network, namely MobileNet, InceptionV3, ResNet-50, VGG16 and VGG19 is proposed to detect DR in retinal fundus images. To evaluate the effectiveness of the model, the performance of different techniques is evaluated based on different performance parameters like Sensitivity, Specificity, Precision, Accuracy and ROC curve. Overall, the experiment provides valuable insights into the performance of pre-trained CNNs in detecting DR, which could have important implications for the development of automated screening systems for this condition. Future research will concentrate on hybrid-based strategies. The method will be examined on online databases and local databases collected from IGM Hospital, Agartala, India. Different performance metrics, such as sensitivity, specificity, accuracy and ROC Curve will be utilized to assess the suggested approach.

References

[1] P. Khojasteh, B. Aliahmad and D. K. Kumar, "Fundus images analysis using deep features for detection of exudates, hemorrhages and microaneurysms," *BMC Ophthalmology*, vol. 18, no. 1, Nov. 2018.

[2] S. Gnanamurthy and V. K. Kaliappan, "Cloud-based onboard prediction and diagnosis of diabetic retinopathy," *Concurrency and Computation: Practice and Experience*, vol. 33, no. 24, Jun. 2021.

[3] M. Mateen, J. Wen, N. Nasrullah, S. Sun and S. Hayat, "Exudate Detection for Diabetic Retinopathy Using Pretrained Convolutional Neural Networks," *Complexity*, vol. 2020, pp. 1–11, Apr. 2020.

[4] G. U. Nneji, J. Cai, J. Deng, H. N. Monday, M. A. Hossin and S. Nahar, "Identification of Diabetic Retinopathy Using Weighted Fusion Deep Learning Based on Dual-Channel Fundus Scans," *Diagnostics*, vol. 12, no. 2, p. 540, Feb. 2022.

[5] "APTOS-2019 dataset," www.kaggle.com. `https://www.kaggle.com/datasets/mariaherrerot/aptos2019` (accessed January 16, 2023).

[6] "High-Resolution Fundus (HRF) Image Database," www5.cs.fau.de. `https://www5.cs.fau.de/research/data/fundus-images` (accessed January 16, 2023).

[7] A. Z. H. Ooi *et al.*, "Interactive Blood Vessel Segmentation from Retinal Fundus Image Based on Canny Edge Detector," *Sensors*, vol. 21, no. 19, p. 6380, Sep. 2021.

[8] H. -r. Wang, J. -l. Yang, H. -j. Sun, D. Chen and X. -l. Liu, "An Improved Region Growing Method for Medical Image Selection and Evaluation Based on Canny Edge Detection," *2011 International Conference on Management and Service Science*, Wuhan, China, pp. 1-4, 2011.

[9] W. Kusakunniran, P. Charoenpanich, P. Samunyanoraset, S. Suksai, S. Karnjanapreechakorn, Q. Wu and J. Zhang, "Hybrid Learning of Vessel Segmentation in Retinal Images," *ECTI Transactions on Computer and Information Technology (ECTI-CIT)*, vol. 15, no.1, pp. 1–11, 2020.

[10] Y.H. Li, N.N. Yeh, S.J. Chen and Y.C. Chung, "Computer-Assisted Diagnosis for Diabetic Retinopathy Based on Fundus Images Using Deep Convolutional Neural Network," *Mobile Information Systems*, vol. 2019, pp. 1–14, Jan. 2019.

[11] B. Bulut, V. Kalın, B. B. Güneş and R. Khazhin, "Deep Learning Approach For Detection Of Retinal Abnormalities Based On Color Fundus Images," *2020 Innovations in Intelligent Systems and Applications Conference (ASYU)*, Istanbul, Turkey, pp. 1-6, 2020.

[12] E. Abitbol, A. Miere, J.B. Excoffier, CJ Mehanna, F. Amoroso, S Kerr, M Ortala and E.H. Souied, "Deep learning-based classification of retinal vascular diseases using ultra-widefield colour fundus photographs," *BMJ Open Ophthalmology*, vol. 7, no. 1, pp. e000924, Feb. 2022.

[13] A. Sungheetha and R. Sharma R, "Design an Early Detection and Classification for Diabetic Retinopathy by Deep Feature Extraction based Convolution Neural Network," *Journal of Trends in Computer Science and Smart Technology*, June 2021, vol. 3, no. 2, pp. 81–94, Jul. 2021, DOI: https://doi.org/10.36548/jtcsst.2021.2.002.

[14] G. Mushtaq and F. Siddiqui, "Detection of diabetic retinopathy using deep learning methodology," *IOP Conference Series: Materials Science and Engineering*, vol. 1070, pp. 012049, Feb. 2021.

[15] N. Khalifa, M. Loey, M. Taha, and H. Mohamed, "Deep Transfer Learning Models for Medical Diabetic Retinopathy Detection," *Acta Informatica Medica*, vol. 27, no. 5, pp. 327, 2019.

[16] S. Gupta, S. Thakur, and A. Gupta, "Optimized hybrid machine learning approach for smartphone based diabetic retinopathy detection," *Multimedia Tools and Applications*, vol. 81, pp. 14475-14501, Feb. 2022.

[17] P. Saranya, S. Prabakaran, R. Kumar and E. Das, "Blood vessel segmentation in retinal fundus images for proliferative diabetic retinopathy screening using deep learning," *The Visual Compute: International Journal of Computer Science*, vol. 38, pp.977-992, Jan. 2021.

[18] I. Qureshi, J. Ma, and Q. Abbas, "Diabetic retinopathy detection and stage classification in eye fundus images using active deep learning," *Multimedia Tools and Applications*, vol. 80, no. 8, pp. 11691–11721, Jan. 2021.

[19] A. Bilal, L. Zhu, A. Deng, H. Lu, and N. Wu, "AI-Based Automatic Detection and Classification of Diabetic Retinopathy Using U-Net and Deep Learning," *Symmetry*, vol. 14, no. 7, pp. 1427, Jul. 2022.

[20] R. Sundaram, R. KS, P. Jayaraman, and V. B, "Extraction of Blood Vessels in Fundus Images of Retina through Hybrid Segmentation Approach," *Mathematics*, vol. 7, no. 2, pp. 169, Feb. 2019.

[21] S. Albahli and G. Nabi Ahmad Hassan Yar, "Detection of Diabetic Retinopathy Using Custom CNN to Segment the Lesions," *Intelligent Automation & Soft Computing*, vol. 33, no. 2, pp. 837–853, 2022.

[22] R. K. Yazid and S. Samsuryadi, "Detection of Diabetic Retinopathy Using Convolutional Neural Network (CNN)," *Computer Engineering and Applications Journal*, vol. 11, no. 3, pp. 203–213, Oct. 2022.

[23] S. Akbar and D. Midhunchakkaravarthy, "A novel 3D-CNN based feature extraction based classification for diabetic retinopathy(DR) detection," *Journal of Mechanics of Continua and Mathematical Sciences*, vol.15, no. 2, p.103-117, Feb. 2020.

[24] A. Ayala, T. Ortiz Figueroa, B. Fernandes and F. Cruz, "Diabetic Retinopathy Improved Detection Using Deep Learning," *Applied Sciences*, vol. 11, no. 24, pp. 11970, Dec. 2021.

[25] D. Das, S.K. Biswas and S. Bandyopadhyay, "Detection of Diabetic Retinopathy using Convolutional Neural Networks for Feature Extraction and Classification (DRFEC)," *Multimedia Tools and Applications*, pp. 1-59, Oct. 2022.

[26] Y. R., V. Raja Sarobin M., R. Panjanathan, G. J. S., and J. A. L., "Diabetic Retinopathy Classification Using CNN and Hybrid Deep Convolutional Neural Networks," *Symmetry*, vol. 14, no. 9, p. 1932, Sep. 2022.

[27] Y.S. Devi and S.P. Kumar, "A deep transfer learning approach for identification of diabetic retinopathy using data augmentation," *IAES International Journal of Artificial Intelligence*, vol. 11, no. 4, pp. 1287-1296. Dec. 2022.

[28] H.R. Ismail and M.M. Hassan, "Bayesian deep learning methods applied to diabetic retinopathy disease: a review," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 30, no. 2, pp. 1167-1177, May 2023.

[29] P. Datta, P. Das and A. Kumar, "Hyperparameter tuning based gradient boosting algorithm for detection of diabetic retinopathy: an analytical review," *Bulletin of Electrical Engineering and Informatics*, vol. 11 , no. 2, pp. 814-824, Apr. 2022.

[30] U.W. Wasekar and R.K Bathla, "A review on supervised learning methodologies for detection of exudates in diabetic retinopathy," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 23, no. 2, pp. 837-846, Aug. 2021.

[31] Md. R. Islam *et al.*, "Applying supervised contrastive learning for the detection of diabetic retinopathy and its severity levels from fundus images," *Computers in Biology and Medicine*, vol. 146, pp. 105602–105602, May. 2022.

[32] A.-O.Asia *et al.*, "Detection of Diabetic Retinopathy in Retinal Fundus Images Using CNN Classification Models," *Electronics*, vol. 11, no. 17, pp. 2740, 2022.

[33] E. AbdelMaksoud, S. Barakat and M. Elmogy, "A computer-aided diagnosis system for detecting various diabetic retinopathy grades based on a hybrid deep learning technique," *Medical and Biological Engineering & Computing*, vol. 60, no. 7, pp. 2015–2038, 2022.

Smita Das received her B.Tech degree in Computer Science and Engineering from National Institute of Technology, Agartala, Tripura, India in 2006, M.Tech in Computer Science and Engineering from Tripura University (A Central University), India in 2008. Currently, she is working as an Assistant Professor in the Department of Computer Science, MBB College, Tripura, India. She is pursuing Ph.D degree from Information Technology Department of Tripura University (A Central University), Tripura, India. Her research interests include medical image processing.

Madhusudhan Mishra has been working as an Associate Professor in the Department of Electronics and Communication Engineering, North Eastern Regional Institute of Science and Technology (NERIST), Nirjuli, Arunachal Pradesh since 2019. He received his B.Tech. Degree in Electronics and Communication Engineering from NERIST (2004), M. Tech in Signal Processing from IIT Guwahati (2011), and holds a Ph.D. Degree from the Department of Electrical Engineering, IIT Kharagpur, India (2020). His research areas focus on Biomedical Signal Processing and Machine Learning. With more than 18 years of experience in research, teaching, designing, and leading programs, he co-edited a research book and published his work in different journals, conferences, and professional press of repute.

Swanirbhar Majumder received his Ph.D, M.Tech, and B.Tech from Jadavpur University, in West Bengal; University of Calcutta, West Bengal; and North Eastern Regional Institute of Science and Technology (NERIST), under North Eastern Hill University (NEHU), Shillong in India respectively. Currently, since 2020 he has been working as a Professor of Information Technology at Tripura University (A Central University). His research interests include Signal Processing, Medical Image processing.