

Deep Learning Empowered Diabetic Retinopathy Detection and Classification using Retinal Fundus Images

^{1*} Sril Lxmi Kuna, ^{2,3} Dr. A.V. Krishna Prasad

^{1*}Research Scholar, Department of CSE, Koneru Laxmaiah Educational Foundation, Vaddeswaram, A.P, India &

²Research Supervisor, Department of CSE, Koneru Laxmaiah Educational Foundation, Vaddeswaram, A.P, India &

³ Associate Professor in IT Department, Maturi Venkata Subba Rao Engineering College, Hyderabad

Corresponding author: srilaxmiravi2016@gmail.com

Abstract—Diabetic Retinopathy (DR) is a commonly occurring disease among diabetic patients that affects retina lesions and vision. Since DR is irreversible, an earlier diagnosis of DR can considerably decrease the risk of vision loss. Manual detection and classification of DR from retinal fundus images is time-consuming, expensive, and prone to errors, contrasting to CAD models. In recent times, DL models have become a familiar topic in several applications, particularly medical image classification. With this motivation, this paper presents new deep learning-empowered diabetic retinopathy detection and classification (DL-DRDC) model. The DL-DRDC technique aims to recognize and categorize different grades of DR using retinal fundus images. The proposed model involves the Contrast Limited Adaptive Histogram Equalization (CLAHE) technique as a pre-processing stage, which is used to enhance the contrast of the fundus images and improve the low contrast of medical images. Besides, the CLAHE is applied to the L channel of the retina images that have higher contrast. In addition, a deep learning-based Efficient Net-based feature extractor is used to generate feature vectors from pre-processed images. Moreover, a deep neural network (DNN) is used as a classifier model to allocate proper DR stages. An extensive set of experimental analyses takes place using a benchmark MESSIDOR dataset and the results are examined in terms of different evaluation parameters. The simulation values highlighted the better DR diagnostic efficiency of the DL-DRDC technique over the recent techniques.

Keywords: Diabetic Retinopathy, Retinal fundus images, Deep learning, CLAHE, Classification.

I. INTRODUCTION

People who have diabetes may develop diabetic retinopathy, an eye disorder that can lead to blindness and visual loss. Diabetes develops over time and gradually damages the retina of the human eye when the pancreas fails to release adequate insulin. A patient's eyesight begins to deteriorate as the condition worsens, which causes diabetic retinopathy. Diabetes-related elevated blood sugar results in diabetic retinopathy. The retina's blood vessels are impacted. DR is a major disease created by diabetes. It damages the blood vessels of the retina, for the person having type-I/II diabetes. The continuous rise in the blood sugar level harms the blood vessels of the retina. The increased glucose level mostly wounds the blood vessel and causes blood leakages in the eye and thus weakening the vision [1, 2]. Optical coherence tomography (OCT), Slit lamp biomicroscopy, fundus photography, and fundus fluorescein angiography (FFA) are the few eye testing methods utilized for diagnosing the disease in its earlier phase by themselves. The existence of magnitude and their abnormalities determine the seriousness of the diseases. The automatic approaches for detecting DR are time-saving and costly and are highly effective compared to manual

diagnoses [3]. A manual diagnosis is liable to misdiagnoses and needs additional efforts compared to automated techniques. The DL is a division of ML methods that includes a hierarchical layer of nonlinear handling phases for unsupervised feature learning and pattern categorization. DL can be employed for medical image analyses including segmentation, detection, registration, classification, and retrieval of the images. In recent times, DL is broadly utilized in DR diagnostic process. It could effectively learn the features of input data while a lot of heterogeneous sources are incorporated. It has several DL-based approaches like convolution neural networks (CNNs), restricted Boltzmann Machines, sparse coding, and auto encoder [4].

The efficiency of these techniques rises if the amount of training data increases because of the increased learned features different from ML approaches [5]. As well, DL techniques did not need handcrafted feature extraction. CNNs are extensively utilized compared to other approaches in medical image analyses, and it is very efficient [6]. The CNN framework consists of fully connected layers (FC), convolution layers (CL), and pooling layers (PL). The number of filters, layers, and size of the CNN differ based on the

researcher's vision. All the layers in the CNN framework perform an important part. In the CL, various filters convolve an image for extracting the features. Usually, PL follows the CL for reducing the dimension of feature maps [7]. It has various methods for PL however max and average PL are mostly adapted. An FC layer is a compact feature for describing the entire input image. The softmax (SM) activation function is the most utilized classification function. It contains distinct available pre-trained CNN frameworks on ImageNet datasets like ResNet, AlexNet, and Inception-v3. Few researchers namely TL, pre-trained frameworks to accelerate training whereas other researchers create their individual CNN from scrape for classification [8]. The TL methods of the pre-trained model include training all layers of the pre-trained model/fine-tuning multiple layers/fine-tuning the last FC layer. In general, the procedure utilized for detecting and classifying the DR images starts by gathering the dataset and employing the required pre-process for improving and enhancing the images [9]. Later, it is fed into the DL approach for extracting the features and for classifying the images.

This paper develops a new DL-empowered DR detection and classification (DL-DRDC) technique to recognize and categorize different grades of DR using retinal fundus images. The proposed model involves Contrast Limited Adaptive Histogram Equalization (CLAHE) technique as a preprocessing stage. The usage of the CLAHE technique improves the quality of the retinal fundus image and thereby improves the classification results. Moreover, DL-based EfficientNet-based feature and deep neural network (DNN) based classifier model is employed to allocate proper DR stages. A comprehensive set of simulations was carried out on a benchmark MESSIDOR dataset and the outcomes are assessed in terms of varying performance parameters.

The major contribution of this research work is manifested below:

- This study introduces the DL-DRDC novel model, which uses deep learning to identify and categorize diabetic retinopathy. The goal of the DL-DRDC approach is to identify and classify various degrees of DR utilizing retinal fundus pictures.
- To assign suitable DR phases, a new deep neural network (DNN) is employed as a classifier model.

The rest of this essay is organized as follows: The discussion of Section II is literature works undergone in the Diabetic Retinopathy Detection and Classification. Section III explains the Proposed DR Model. Section IV describes the findings obtained using the projected model, and Section V concludes this research.

II. LITERATURE REVIEW

In 2022, Haideret *al.*[15] have proposed employing retinal fundus images, artificial intelligence-based computers can assist in the diagnosis of glaucoma. This study precise pixel-by-pixel segmentation of the OC and OD, we suggest two networks: the detachable linked segmented network (SLS-Net) and the separable linked segmentation residual network (SLSR-Net). A large global feature map can be preserved in SLS-Net and SLSR-Net, which improves the performance of OC and OD segments by reducing the loss of spatial information.

In 2021, Hemalakshmi *et al.* [16] have discussed the classification of a retinal fundus picture using MS-DRLBP characteristics and a CNN-RBF classifier. To address these issues, the research presented here suggests an effective classification paradigm for retinal fundus image recognition. Additionally, classification was performed using a hybrid Convolution Neural Network (CNN) and Radial Basis Function (RBF) model, which divides the retinal fundus photographs into four classes: DR, AMD, CNV, and Normal (NR).

In 2020, Kouroupi *et al.*[17] have examined Principles and implementations of artificial intelligence-assisted diabetic retinopathy diagnosis using digital fundus pictures in the National Health Service. This study argues that automated diagnostics have the potential to change how healthcare providers use cutting-edge technology.

In 2020, Dharmana and Aiswaraya [18] have developed Utilizing Blob Detection, pre-diagnosis of diabetic retinal disease. The method presented in this work proposes an efficient feature extraction technique centered on blob detection, followed by machine learning methodology categorization of various phases of diabetic retinopathy.

In 2020, Hassan *et al.*[19] have recommended Fundus Pictures Using Convolution Neural Networks for Proliferative Diabetic Retinopathy Identification. The three goals of this study were to perform two pre-processing methods on the data set, which involved resizing and normalizing the fundus images; establish a deep learning operations and maintenance Artificial Intelligence (AI) network of feature extraction algorithm for identification of PDR on the fundus images; and evaluate the output classification of the network, which took accuracy, sensitivity, and specificity into account.

In 2021, Freeda *et al.*[20] have proposed a Disease Recognition from Human Fundus Images Using Image Processing Techniques. An illness diagnosis and classification strategy based on image processing and deep learning techniques is suggested by our study system. Our system is PC-based; therefore, it may be used even in far-off places. The patient must submit the eye's fundus image, which is used as input by the program.

In 2020, Prem and Umesh [21] have discussed a Deep Learning Classification of Exudates for Diagnosis of Diabetic Retinopathy. The research suggests an approach that leverages features such the local binary pattern (LBP) and wavelet transform approximations coefficient matrix to target DR identification based on exudates.

In 2021, Shanthinet *al.*[22] have examined a threshold, describing the characteristics multi-layer analysis with a neural network convolution to identify diabetic retinopathy. This work introduces a threshold segmentation-based DR detection technique. Through pixel-based segmentation, this approach is adept at separating the foreground from the background of the given retinal image.

In 2020, Shoba and Therese [23] have recommended using morphological operations and the finite element approach, glaucoma disease can be identified in fundus pictures. The current study has been designed to analyse blood flow in physiologically plausible retinal arteries and venous networks using a sophisticated computational model. Compared to existing methods, the suggested method has the higher precision of 94.86% for the reliable estimation of glaucomatous illness.

In 2021, Vanakovaryanet *al.*[24] have developed a self-sufficient glaucoma detection utilizing IOT and classification using a fundus image of the eye obtained using the CNN technique. A unique methodology for glaucoma location using the edge strategy for fractal examination is provided in light of the fact that the fractal assessment is used to identify the component of unexpected articles. The computerized image preparation method used by CNN to organize and identify the normal and glaucoma-related pictures is carried out using IOT.

III. THE PROPOSED DR DIAGNOSIS MODEL

In this paper, a new DL-DRDC technique is derived for DR diagnosis using retinal fundus images as shown in Figure 1. The DL-DRDC technique encompasses three different stages of operations such as pre-processing, feature extraction, and classification. Primarily, the CLAHE technique is employed for the enhancement of contrast levels in the retinal fundus images. In addition, the EfficientNet technique extracts a set of feature vectors from the preprocessed images. Finally, the DNN classifier allocates proper class labels to the applied retinal fundus images. The detailed processes involved in the three modules are offered in the subsequent subsections.

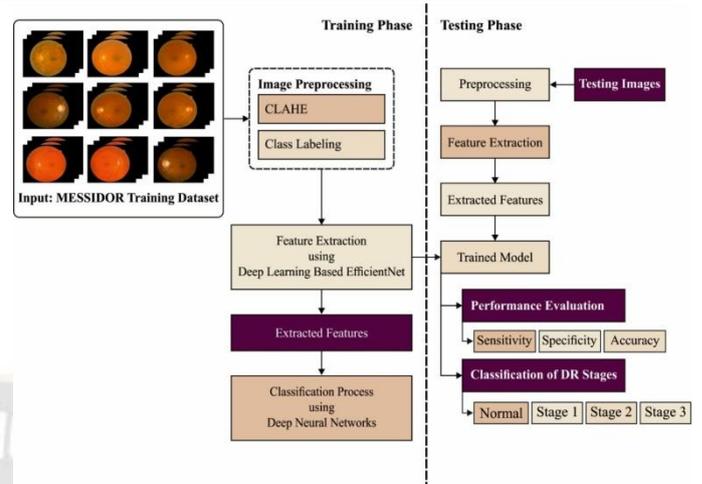


Figure 1: Overall working process of the DL-DRDC model

A. Image Pre-processing: CLAHE Technique

Initially the collected raw images pre-processed to remove the noise and to enhance the quality of the image. In this research work, the pre-processing is accomplished via Gaussian filtering and CLAHE technique.

Gaussian noise: White noise or normal noise is another name for Gaussian noise, which is normally distributed. It is an additive noise that arises from a variety of sources [28]. Because this noise only adds to the signal rather than integrating it, it is a cumulative process. The standard deviation is determined by Gaussian Noise, in which the normal (bell) is shaped noise. Gaussian Noise Distribution is the probability density function P of a Gaussian Random Variable z .

$$P_k(m) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(m-\mu)^2}{2\sigma^2}} \quad (1)$$

Where,

m - Grey Level Gaussian Distribution noise, μ - The Mean Value

σ - The Standard Deviation

Salt and pepper (Binary Noise): Salt and pepper noise is caused by failed camera sensor cells, synchronization errors during image digitization or transmission, or memory cell failure, among other things[29]. This type of noise is sometimes referred to as impulsive noise, shot noise, or binary noise. An image's presence is indicated by a scattering of white or black pixels, or both, in the image. The pepper noise and salt noise have intensity values of 0 and 255, respectively, in the 8 bits/pixels image. It can be stated as follows:

$$P(g) = \begin{cases} M_a \\ M_b \\ 0 \end{cases} \quad (2)$$

In the equation above, $g = a$ and $g = b$, respectively. g - The amount of salt and pepper noise in the image is defined by M_a and M_b .

Speckle noise: The speckle noise is a multiplicative noise type. During the capture and transmission of this random undesired signal, it is multiplied into different important signals [30]. Radar and medical pictures are the most common places where speckle noise can be seen. An exponential random variable can be made out of the product of the mean value. Image texture is how speckle noise is commonly defined based on medical surveys. The speckle noise is represented by the following:

$$K(i, j) = x(i, j) * n(i, j) + \varepsilon(i, j) \quad (3)$$

Where $x(i, j)$ is the practical image, $n(i, j)$ is the speckle noise multiplicative factor, and $\varepsilon(i, j)$ is the speckle noise additive factor. The axial and neighbouring keys of an image sample are denoted by i and j , respectively. Speckle noise is a serious issue in radar applications.

Poisson Disturbance: Shot noise is another name for photon noise. It is based on the Poisson distribution, which differs significantly from Gaussian noise in most cases. Because this is also a form of additive noise [31]. Different pixels' sounds are unrelated to one another. Poisson noise is commonly found in radiography images. Dark current leakage is treated by Poisson noise in the image sensor, resulting in a noise type known as "dark shot noise."

The CLAHE pre-processing approach is effective in removing Gaussian noise and to a lesser extent, speckle noise. In the first stage, the CLAHE technique is employed as a pre-processing tool to increase the contrast level of the retinal fundus images. In adaptive histogram equalization, every histogram of a sub-image reallocates the brightness values of the image which enhances local contrast and produces further details. But, this procedure might improve noise when it is utilized on noisy images like medical images. A prolonged way of adaptive histogram equalization is called CLAHE. The CLAHE method allocates image brightness and extremities noise amplification. To attain this, CLAHE splits an image into smaller overlapping areas. Later, contrast improvement is operating every tile using histogram equalization. Clipping limits are utilized for overcoming noise amplification issues by cutting histogram height beforehand and calculating the increasing distribution function [22]. The histogram structure of every tile is reorganized based on distribution. The CLAHE technique contains various stages.

1. Split the original image into $M \times N$ areas.
2. Evaluate a histogram for every area according to image grayscale levels.
3. Evaluate a contrast limit histogram for every area:
 - 3.1. Utilize the number of grayscale levels N_{gray} in the area and amount of pixels in the dimension $X \& Y$ for calculating the average amount of pixels.

$$N_{avg} = (N_X \times N_Y) / N_{gray} \quad (4)$$

Here N_X and N_Y are the X and Y images in N areas.

3.2. Assume that N_{CL} represents the original clip limit CL , and N_{clip} denotes the normalized CL in a range of zero and one. Clip the pixel when the number of pixels is higher compared to N_{CL}

Evaluate the overall amount of trimmed pixels $N_{\Sigma clip}$ and the average of outstanding pixels

$$N_{avggray} = N_{\Sigma clip} = N_{gray} \quad (5)$$

Assume that $H_{area}(i)$ denotes the actual histogram and $H_{region_clip}(i)$ represents the trimmed histogram of every area at the i th grayscale level, and it is represented as:

When $(H_{area}(i) > N_{CL})$

OR

$(H_{area}(i) + N_{avggray}) > N_{CL}$ then

$$H_{region_clip}(i) = N_{CL}$$

Otherwise

$$H_{(region_clip)}(i) = H_{area}(i) + N_{CL} \quad (6)$$

Assume that N_{remain} denotes the residual amount of trimmed pixels, $Step$ denotes a positive integer i.e., higher/equivalent to one, reallocate the residual pixels by:

$$Step = N_{gray} = N_{remain} \quad (7)$$

Growth from minimal to the maximal grayscale level through the prior step. When the amount of pixels in the grayscale level is lesser compared to N_{CL} , allocate 1 pixel to the grayscale level.

When the entire pixels have not been allocated before termination, evaluated a novel step and begin a new searching iteration. The end of the residual pixels was allocated.

Improve the intensity value in every area by a transform (for example Rayleigh).

Generate a transfer function by converting the trimmed histogram to an increasing likelihood $P_{input}(i)$. Assume y_{min} denotes the lower bound of pixel value, and α indicates the Rayleigh scaling variable, and it is expressed as follows:

$$y(i) = y_{\min} + \sqrt{2\alpha^2 \ln\left(\frac{1}{1 - P_{input}(i)}\right)}$$

$$p(y(i)) = \frac{(y(i) - y_{\min})}{\alpha^2} \cdot \exp\left(\frac{(y(i) - y_{\min})^2}{2\alpha^2}\right) \text{ for } y(i) < y_{\min} \quad (8)$$

6. Decrease abruptly changing effect by the linear contrast stretch.

$$y(i) = \frac{x(i) - x_{\min}}{x_{\max} - x_{\min}} \quad (9)$$

Remove edge artifacts by evaluating the grayscale level assignments of pixels through bi-linear interpolation.

Figure 2 visualizes some of the sample's preprocessed images. Figure 2a, c displays the original DR images and the preprocessed versions are shown in Figure 2b, d. The figure has shown that the DR image quality is considerably improved and the lesion area is visible.

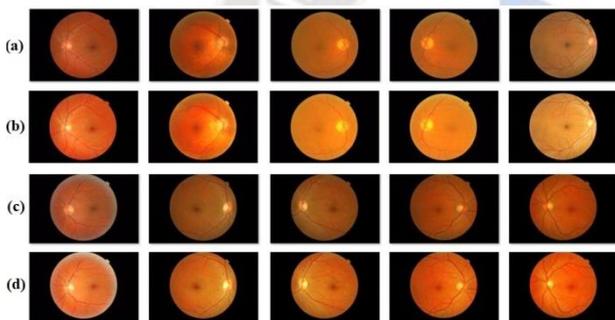


Figure 2. Preprocessed images (a), (c)-Input images; (b), (d)-Output images

B. Feature Extraction: EfficientNet Technique: At this stage, the contrast improved pre-processed image is passed into the different layers of the EfficientNet model to derive feature vectors. A CNN is a type of trainable FFNN. The common CNN comprises CL, the input layer, the FC layer, the output layer, and PL. As displayed in Figure 3, the initial layer of a general CC is the input layer, i.e., typically followed by a framework that integrates PL and multiple convolutions.

Designers must select a strong CNN as the initial component to execute feature extraction on each fundus picture and produce a small feature vector representation in order to train a powerful feature extractor. The majority of CNN's current development is dependent on fixed development resources. The network will keep growing if there is enough processing power. EfficientNet consistently scaled all depth, breadth, and resolution dimensions using simple, effective composite coefficients. Additionally, EfficientNet outperformed state-of-the-art accuracy on ImageNet and five widely used transfer learning datasets with orders of magnitude fewer parameters

and floating point operations per second (FLOPS). This article developed a multi-label classification model for fundus pictures due to EfficientNet's excellent efficiency. Utilising EfficientNet to extract features. At the same time, we also used some excellent CNNs as feature extractors for comparative research.

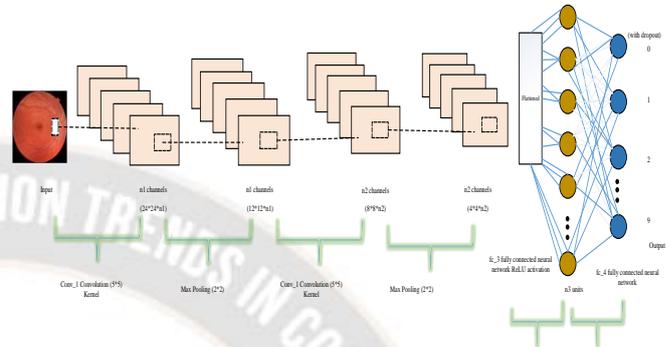


Figure 3: Structure of CNN

The final layer is the FC layer, and the SM classifiers are utilized for identifying the stages of diabetic retinopathy. It has 2 phases in finding the stages with the help of CNN: classification and feature learning. The initial phase comprises CL, FC, and PL, and the final phase has an SM classifier. The main task of the activation function is to create the feature map of the output that contains non-linear relations. Also, it has 3 types of saturated non-linear functions, tanh, sigmoid, and unsaturated non-linear function ReLU and softsign. Commonly, ReLU is utilized as an activation function in CNN because of the benefits of the speed of the unsaturated nonlinear functions if the training gradient decreases. The process of ReLU is given in Eq. (10), whereas x denotes the input of the activation function [23].

$$f(x) = \begin{cases} 0, & \text{if } x < 0, \\ x, & \text{if } x \geq 0. \end{cases} \quad (10)$$

As a sampling procedure, the PL is utilized for screening the features in the perceptive area and extracting the illustrative features in the field that could efficiently decrease the output feature scale, therefore it reduces the number of variables needed by the module and maintains translation invariance. The maximal pool technique is adopted in this work. The final layer of the CNN is the FC layer, i.e., in charge of summarizing the features learned using the CNN. Lastly, the feature is categorized with the help of an SM classifier. Assume the length of the output vector and input vector denotes M & N , correspondingly, the amount of variables in the FC layer is evaluated by:

$$Q = M \times N + N. \quad (11)$$

Compared to the 2 states of polymorphic problem, the prolonged SM regression method is often a preference for classification. The training set comprises label samples $\{(x^{(1)}, y^{(1)}), \dots, (x^{(k)}, y^{(k)})\}$, whereas the value of label y is zero/one and the input feature is $x^{(i)} \in R^{n+1}$. They consider the LR function that is expressed by:

$$h_{\theta}(x) = \frac{1}{1+e^{-\theta^T x}} \tag{12}$$

Whereas θ denotes the model variable of the minimal loss function $J(\theta)$ afterward training.

$$J(\theta) = -\frac{1}{k} \left[\sum_{i=1}^k y^{(i)} \log h_{\theta}(x^{(i)}) + (1 - y^{(i)}) \log (1 - h_{\theta}(x^{(i)})) \right] \tag{13}$$

For multistate classifier problems, they consider n corresponding labels and n conditions in the SM classification. The training set comprises label samples $\{(x^{(1)}, y^{(1)}), \dots, (x^{(k)}, y^{(k)})\}$, whereas the label is $y^{(i)} \in \{1, 2, \dots, n\}$. For a provided training sample x with n classes, the existence likelihood of state i is $p(y = j|x)$, and the output of SM regression is given below:

$$h_{\theta}(x) = \frac{1}{\sum_{j=1}^n e^{x\theta_j}} \begin{bmatrix} P(y = 1|x; \theta) \\ P(y = 2|x; \theta) \\ \dots \\ P(y = n|x; \theta) \end{bmatrix} = \begin{bmatrix} e^{x\theta_1} \\ e^{x\theta_2} \\ \dots \\ e^{x\theta_n} \end{bmatrix} \tag{14}$$

In which $\theta_1, \theta_2, \dots, \theta_k$ denote the model's parameter, $(1/\sum_{j=1}^n e^{x\theta_j})$ plays the part of normalization. Therefore, the loss function $J(\theta)$ is given by:

$$J'(\theta) = -\frac{1}{k} \left[\sum_{i=1}^k \sum_{j=1}^n \{y^{(i)} = j\} \log \frac{e^{x\theta_j}}{\sum_{j=1}^n e^{x\theta_j}} \right] \tag{15}$$

Whereas $\theta_1, \theta_2, \dots, \theta_k$ represent the model's variable of the minimal loss function $J(\theta)$ afterward training to attain the SM classifier.

EfficientNet is a set of classifiers presented currently in 2019 and depends on Compound Scaling and AutoML. The compound scaling technique is utilized for scaling up this baseline to attain EfficientNet B1 to B7. As well, AutoML is utilized for developing a mobile size baseline network (EfficientNet-B0). The Compound Scaling approach measures uniformly each dimension of resolution, depth, and width by a

simple but very efficient compound coefficient. The depth of layers needs to rise by 20%, the image resolution by 15%, and the width by 10% for keeping things effective, although improving the execution and increasing the accuracy. Gamma, Alpha, and beta are the scaling multipliers for resolution, depth, and width, correspondingly. They are attained by a grid search. Phi is a user-specific coefficient. It is a real number that handles assets. The Depth, weight, and resolution based on Phi equations are given below:

$$Depth : d = \alpha^{\phi}, \tag{16}$$

$$Width : w = \beta^{\phi}, \tag{17}$$

$$Resolution : r = \gamma^{\phi}, \tag{18}$$

$$while: \alpha, \beta^2, \gamma^2 \approx 2; \alpha \geq 1, \beta \geq 1 \text{ and } \gamma \geq 1. \tag{19}$$

The efficientNet-B0 framework is a mobile-sized framework containing 11M trainable variables. Its framework is shown in Figure 4, whereas each row is an individual phase i in the network. Each phase i is characterized using an input resolution size $\langle \hat{H}_i, \hat{W}_i \rangle$, an output channels size \hat{C}_i and amount of layers \hat{L}_i .

It utilizes 7 reversed residual blocks. Excitation and Squeeze blocks are utilized with the swish activation function. EfficientNet utilizes seven MBConv blocks. Each MBConv block takes 2 inputs. The initial one is data, and the next one is block arguments. The data is obtained from the final layer. The block argument is a group of attributes to be utilized inside MBConv blocks, such as input filters, squeeze ratio, output filters, expansion ratio, and so on. The expansion stage aims to extend the layer to create a wider [24]. The depth-wise convolution stage employs a depth-wise convolution by the kernel size stated in the block arguments. The excitation and Squeeze stage extracts the global features with the help of the global average PL. Later, it squeezes the number of channels with the help of the squeeze ratio. The Output stage employs a convolutional function by the output filters stated in the block arguments.

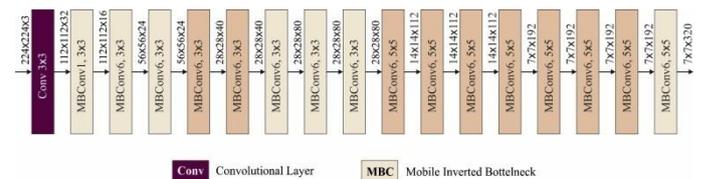


Figure 4: Framework of EfficientNet.

Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 32, 32, 32)	896
conv2d_2 (Conv2D)	(None, 30, 30, 32)	9248
max_pooling2d_1 (MaxPooling2D)	(None, 15, 15, 32)	0
dropout_1 (Dropout)	(None, 15, 15, 32)	0
conv2d_3 (Conv2D)	(None, 15, 15, 64)	18496
conv2d_4 (Conv2D)	(None, 13, 13, 64)	36928
max_pooling2d_2 (MaxPooling2D)	(None, 6, 6, 64)	0
dropout_2 (Dropout)	(None, 6, 6, 64)	0
conv2d_5 (Conv2D)	(None, 6, 6, 64)	36928
conv2d_6 (Conv2D)	(None, 4, 4, 64)	36928
max_pooling2d_3 (MaxPooling2D)	(None, 2, 2, 64)	0
dropout_3 (Dropout)	(None, 2, 2, 64)	0
flatten_1 (Flatten)	(None, 256)	0
dense_1 (Dense)	(None, 512)	131584
dropout_4 (Dropout)	(None, 512)	0
dense_2 (Dense)	(None, 10)	5130
Total params: 276,138		
Trainable params: 276,138		
Non-trainable params: 0		

Figure 5: Partial model description.

C. Image Classification:

DNN Model: Finally, the DNN model receives the feature vectors of the retinal fundus images as input and computes the predictive class score for each image. Essentially, the architecture of DNN consists of a hidden layer, an input layer, and an output layer. By taking into account the efforts of preference weight fitness, the DNN is intended with 2 hidden layers for accurately learning the mapping relations among the output and input data. In the training stage, the DNN iteratively upgrades the weight of nodes in the hidden layer. Because of the rise in the training iteration, the NN repeatedly fits the labeled training data's decision boundary [25]. To improve the training speed of DNN and the classification accuracy, 2 hidden layers are created. In the hidden layer, the overall amount of nodes is estimated in Eq. (20).

$$n = \sqrt{a + b} + c \quad (20)$$

Where the input and output, and hidden layer counts are defined by $a, b, \text{ and } c$. To enable the nonlinear fitness capability, a sigmoid activation function is included in the hidden layer of DNN.

IV. PERFORMANCE VALIDATION

A. Implementation of Data:

The proposed model is simulated on Python 3.6.5 tool. The experimental validation of the presented model takes place using the MESSIDOR dataset [26]. The dataset holds a set of 1200 retinal fundus images with four class labels namely

healthy, DR stages 1-3 respectively. The amount of data collected, 60% of the infor

mation has been used for training purpose and the rest 40% has been utilized for testing purpose. Moreover, the results are inspected in terms of different performance metrics. Figure. 6 showcases the sample test images.

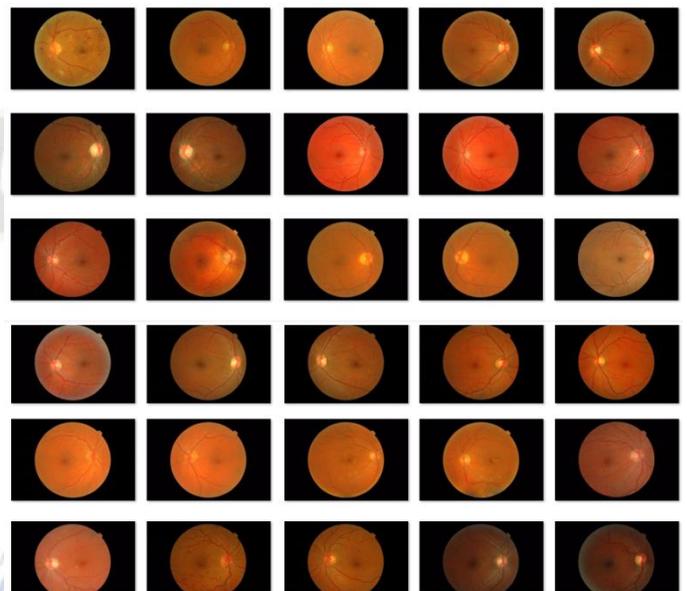


Figure 6: Sample images

B. Results Analysis

Figure7 exhibits the set of confusion matrices produced by the DL-DRDC technique on the applied dataset. The proposed DL-DRDC technique has effectually classified the different stages of DR under different runs. For instance, with run-1, the DL-DRDC technique has classified a set of 543 instances into normal, 150 instances into stage-1, 245 instances into stage-2, and 251 instances into stage-3. Likewise, with run-2, the DL-DRDC method has classified a set of 544 instances into normal, 150 instances into stage-1, 247 instances into stage-2, and 251 instances into stage-3. Concurrently, with run-3, the DL-DRDC manner has classified a set of 544 instances into normal, 151 instances into stage-1, 245 instances into stage-2, and 252 instances into stage-3. Simultaneously, with run-4, the DL-DRDC algorithm has classified a set of 543 instances into normal, 151 instances into stage-1, 246 instances into stage-2, and 252 instances into stage-3. Lastly, with run-5, the DL-DRDC methodology has classified a set of 545 instances into normal, 152 instances into stage-1, 246 instances into stage-2, and 252 instances into stage-3.

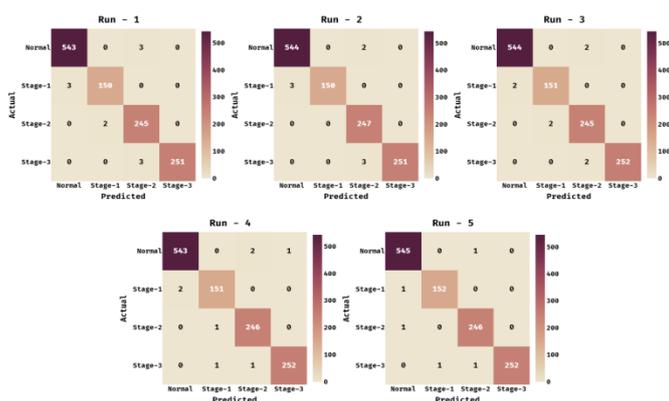


Figure 7: Confusion matrix analysis of DL-DRDC model

Figure 8 displays the ROC analysis of the DL-DRDC technique under different runs of execution. The figure demonstrated that the DL-DRDC technique has accomplished improved performance with the maximum ROC values under all runs.

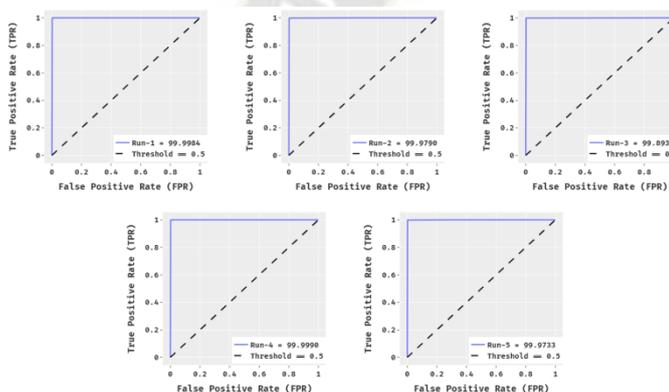


Figure 8: ROC analysis of DL-DRDC model

A detailed classification results analysis of the DL-DRDC technique under varying runs of execution is displayed in Table 1. The experimental results demonstrated that the DL-DRDC technique has shown proficient results in terms of different measures. Under execution run-1, the DL-DRDC technique has classified the normal class with an accuracy of 0.995. In the meantime, the DL-DRDC manner has classified the stage-1 class with an accuracy of 0.996, At the same time, the DL-DRDC methodology has classified the stage-2 class with an accuracy of 0.993. Concurrently, the DL-DRDC approach has classified the stage-3 class with an accuracy of 0.998.

Table 1 DR Classification result analysis of DL-DRDC technique

No. of Runs	Classes	Sensi	Speci	Preci	Accu	F-Score
Run-1	Normal	0.995	0.995	0.995	0.995	0.995
	Stage-1	0.980	0.998	0.987	0.996	0.984
	Stage-2	0.992	0.994	0.976	0.993	0.984
	Stage-3	0.988	1.000	1.000	0.998	0.994

	Average	0.989	0.997	0.989	0.995	0.989
Run-2	Normal	0.996	0.995	0.995	0.996	0.995
	Stage-1	0.980	1.000	1.000	0.998	0.990
	Stage-2	1.000	0.995	0.980	0.996	0.990
	Stage-3	0.988	1.000	1.000	0.998	0.994
	Average	0.991	0.998	0.994	0.997	0.992
Run-3	Normal	0.996	0.997	0.996	0.997	0.996
	Stage-1	0.987	0.998	0.987	0.997	0.987
	Stage-2	0.992	0.996	0.984	0.995	0.988
	Stage-3	0.992	1.000	1.000	0.998	0.996
	Average	0.992	0.998	0.992	0.997	0.992
Run-4	Normal	0.995	0.997	0.996	0.996	0.995
	Stage-1	0.987	0.998	0.987	0.997	0.987
	Stage-2	0.996	0.997	0.988	0.997	0.992
	Stage-3	0.992	0.999	0.996	0.998	0.994
	Average	0.992	0.998	0.992	0.997	0.992
Run-5	Normal	0.998	0.997	0.996	0.998	0.997
	Stage-1	0.994	0.999	0.994	0.998	0.994
	Stage-2	0.996	0.998	0.992	0.998	0.994
	Stage-3	0.992	1.000	1.000	0.998	0.996
	Average	0.995	0.998	0.995	0.998	0.995

An average classification result of the DL-DRDC technique is investigated under different runs in Table 2 and Figure 9. The results exhibited that the DL-DRDC technique has gained effectual DR diagnostic outcomes under distinct runs. For instance, with run-1, the DL-DRDC technique has obtained an increased sensitivity of 0.989, specificity of 0.997, the precision of 0.989, accuracy of 0.995, and F-score of 0.989. Next to that, with run-3, the DL-DRDC manner has resulted in an improved sensitivity of 0.992, specificity of 0.998, the precision of 0.992, accuracy of 0.997, and F-score of 0.992. Eventually, with run-5, the DL-DRDC algorithm attained an enhanced sensitivity of 0.995, specificity of 0.998, precision of 0.992, accuracy of 0.997, and F-score of 0.992.

Table 2 Average Analysis of DL-DRDC technique under different runs

No. of Runs	Sens.	Spec.	Prec.	Accu.	F-Score
Run-1	0.989	0.997	0.989	0.995	0.989
Run-2	0.991	0.998	0.994	0.997	0.992
Run-3	0.992	0.998	0.992	0.997	0.992
Run-4	0.992	0.998	0.992	0.997	0.992
Run-5	0.995	0.998	0.995	0.998	0.995
Average	0.992	0.998	0.992	0.997	0.992

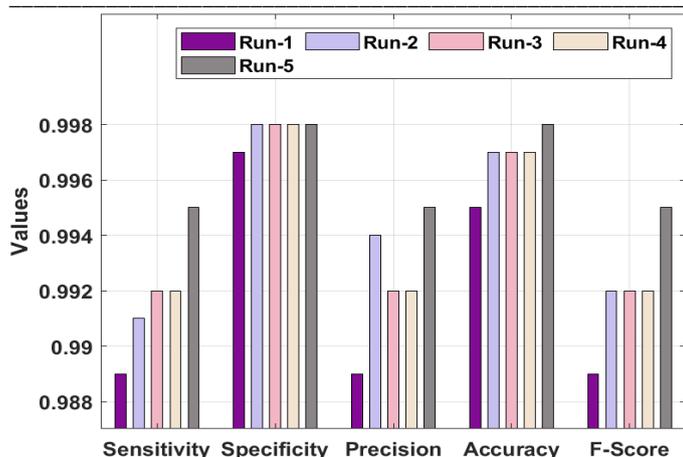


Figure9. Result analysis of DL-DRDC model under distinct runs To showcase the superior result of the DL-DRDC approach, a series of simulations take place in Table 3 [27, 28]. Figure10 assesses the sensitivity of the DL-DRDC technique with other existing techniques. The figure exhibited that the A-Net, VNet-s, and VNet-19 models have obtained lower diagnostic outcomes with a sensitivity of 0.8127, 0.8647, and 0.8931. Also, the VNet-16 and Modified A-Net models have shown reasonably nearer sensitivity of 0.9235 and 0.9791. Concurrently, the DNN-MSO, DL-Inception v4, and CNN-SDL techniques have accomplished near optimal sensitivity of 0.9791, 0.9847, and 0.9854 respectively. However, the DL-DRDC technique has outperformed the other ones with a higher sensitivity of 0.992.

Table 3 Performance Analysis of Various Models with Proposed DL-DRDC Method

Methods	Sensitivity	Specificity	Accuracy
DL-DRDC	0.9920	0.9980	0.9970
DL-InceptionV4	0.9847	0.9959	0.9937
CNN-SDL	0.9854	0.9938	0.9928
DNN-MSO	0.9791	0.9947	0.9912
Modified A-Net	0.9235	0.9745	0.9600
A-Net	0.8127	0.9407	0.8975
VNet-s	0.8647	0.9743	0.9568
VNet-16	0.9078	0.9432	0.9317
VNet-19	0.8931	0.9649	0.9373

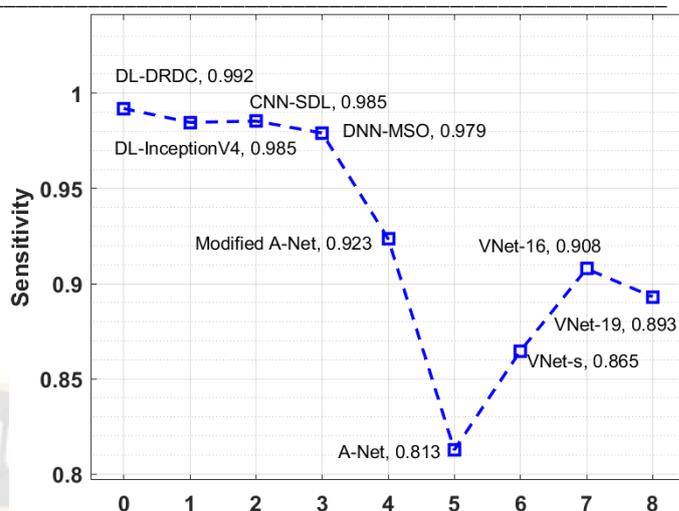


Figure10: Sensitivity analysis of DL-DRDC model with existing techniques

Figure 11 evaluates the specificity of the DL-DRDC method with other existing approaches. The figure portrayed that the A-Net, VNet-16, and VNet-19 techniques have achieved minimal diagnostic outcomes with the specificity of 0.9407, 0.9432, and 0.9649 correspondingly. Besides, the VNet-s and Modified A-Net methods have outperformed reasonably closer specificity of 0.9743 and 0.9745 respectively. Along with that, the CNN-SDL, DNN-MSO, and DL-Inception v4 systems have accomplished near optimal specificity of 0.9938, 0.9947, and 0.9959 correspondingly. Eventually, the DL-DRDC methodology exhibited the other ones with a maximum specificity of 0.998

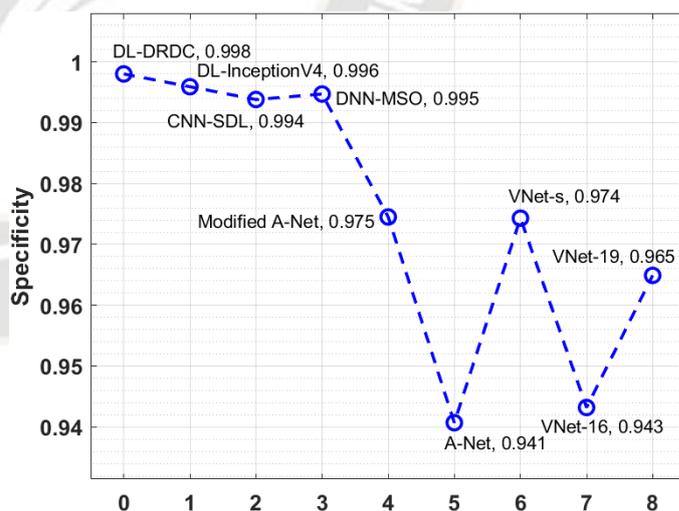


Figure 11. Specificity analysis of DL-DRDC model with existing techniques

Figure 12 measure the accuracy of the DL-DRDC manner with other recent methods. The figure displayed that the A-Net, VNet-16, and VNet-19 approaches have gained minimum diagnostic outcomes with the accuracy of 0.8975, 0.9317, and

0.9373 correspondingly. Also, the VNet-s and Modified A-Net methods have exhibited moderately closer accuracy of 0.9568 and 0.9600 correspondingly. Simultaneously, the DNN-MSO, CNN-SDL, and DL-Inception v4 manners have accomplished near optimum accuracy of 0.9912, 0.9928, and 0.9937 correspondingly. Finally, the DL-DRDC algorithm has showcased the other ones with a superior accuracy of 0.997.

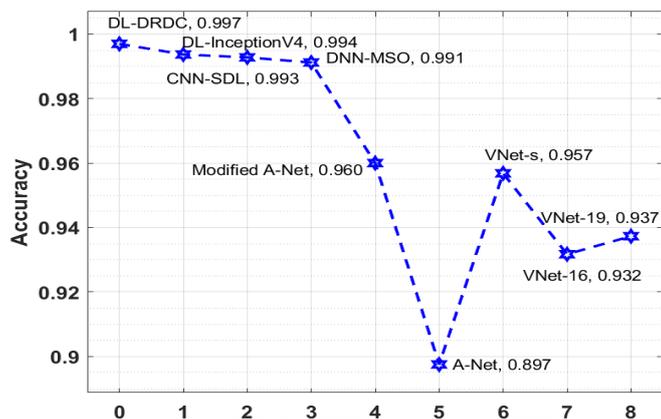


Figure 12: A comparison study of the DL-DRDC technique in terms of accuracy

There are around 60 million parameters in AlexNet, A-Net, VNet-16, VNet-19, and ResNet-152, more computations are required. But with the proposed framework i.e. efficient net need 40 million parameters and fewer computations. It takes about the same amount of time to train an AlexNet, A-Net, VNet-16, VNet-19, and ResNet-152 as it does to train Inception. The proposed model needs 10 times less memory is required, but the results are more accurate (about 9 percent)

V. CONCLUSION

In this study, a new DL-DRDC technique is designed to accomplish optimum solutions to classify the DR persons depending upon the severity levels of the disease. The DL models find it useful to accurately and rapidly produce diagnostic outcomes. The EfficientNet with DNN model can be used to recognize and categorize the fundus images, which assists the ophthalmologist to a maximum level in eliminating vision loss owing to DR. Besides, the use of the CLAHE technique aids to increase the contrast level of the retinal fundus image and thereby the diagnostic performance can be improved. The probability score for every predictive class is determined using the DNN model and the class obtaining the highest score is selected as the predicted class. The validation process of the DL-DRDC technique takes place using the MESSIDOR dataset and 10-fold cross-validation techniques are employed. The experimental outcomes pointed out the better DR diagnostic efficiency of the DL-DRDC technique over the recent state of art DR diagnostic models. Therefore, the DL-DRDC technique can be employed as a proper tool to

diagnose DR in real-time scenarios. As a part of the future scope, the DR diagnostic results can be further enhanced by advanced DL-based segmentation and hyper parameter optimization techniques.

REFERENCES

- [1] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Commun ACM* 2017;60(May (6)):84–90.
- [2] Hajeb Mohammad Alipour S, Rabbani H, Akhlaghi MR. Diabetic retinopathy grading by digital curvelet transform. *Comput Math Methods Med* 2012;2012:1–11.
- [3] Alyoubi, W.L., Shalash, W.M. and Abulkhair, M.F., 2020. Diabetic retinopathy detection through deep learning techniques: A review. *Informatics in Medicine Unlocked*, 20, p.100377.
- [4] Guo Y, Liu Y, Oerlemans A, Lao S, Wu S, Lew MS. Deep learning for visual understanding: a review. *Neurocomputing* 2016;187:27–48.
- [5] R. Kamalraj, M. Ranjith Kumar, V. Chandra ShekharRao, RohitAnand, Harinder Singh, "Interpretable filter based convolutional neural network (IF-CNN) for glucose prediction and classification using PD-SS algorithm", *Measurement*, Vol.183,2021, <https://doi.org/10.1016/j.measurement.2021.109804>.
- [6] Satpathy S, DebbarmaSwapana, SenguptaAditya S.C, BhattacharyyaBidyut K.D, "Design a FPGA, fuzzy based, insolent method for prediction of multi-diseases in rural area", *Journal of Intelligent & Fuzzy Systems*, vol. 37, no. 5, pp. 7039-7046, 2019. doi: 10.3233/JIFS-181577.
- [7] Kuna, Sri Laxmi and Prasad, Dr. A. V. Krishna, Deep Learning Models for Classification of Diabetic Retinopathy Color Fundus Images (September 14, 2022). Available at SSRN: <https://ssrn.com/abstract=4218649> or <http://dx.doi.org/10.2139/ssrn.4218649>
- [8] Jayasankar, U., Thirumal, V. and Ponnurangam, D., 2018. A New Objective Image Quality Assessment Metric: For Color and Grayscale Images. *3D Research*, 9(3), pp.1-16.
- [9] Paulraj, D 2020, 'A gradient boosted decision tree-based sentiment classification of twitter data', *International Journal of Wavelets, Multiresolution and Information Processing*, vol. 18, no. 4, pp. 205027 1-21. DOI: <https://doi.org/10.1142/S0219691320500277>
- [10] SambitSatpathy, Sanchali Das, SwapanDebbarma, "A new healthcare diagnosis system using an IoT-based fuzzy classifier with FPGA", *Journal of Supercomputing* vol. 76, no. 8, pp. 5849–5861, 2020. doi: 10.1007/s11227-019-03013-2.
- [11] dos Santos, J.C.M., Carrijo, G.A., dos Santos Cardoso, C.D.F., Ferreira, J.C., Sousa, P.M. and Patrocínio, A.C., 2020. Fundus image quality enhancement for blood vessel detection via a neural network using CLAHE and Wiener filter. *Research on Biomedical Engineering*, pp.1-13.
- [12] Si, L., Xiong, X., Wang, Z. and Tan, C., 2020. A deep convolutional neural network model for intelligent

- discrimination between coal and rocks in coal mining face. *mathematical Problems in Engineering*, 2020.
- [13] Ben Jabra, M., Koubaa, A., Benjdira, B., Ammar, A. and Hamam, H., 2021. COVID-19 Diagnosis in Chest X-rays Using Deep Learning and Majority Voting. *Applied Sciences*, 11(6), p.2884.
- [14] Ramesh, S. and Vydeki, D., 2020. Recognition and classification of paddy leaf diseases using Optimized Deep Neural network with Jaya algorithm. *Information processing in agriculture*, 7(2), pp.249-260.
- [15] Haider, A., Arsalan, M., Lee, M.B., Owais, M., Mahmood, T., Sultan, H. and Park, K.R., 2022. Artificial Intelligence-based computer-aided diagnosis of glaucoma using retinal fundus images. *Expert Systems with Applications*, 207, p.117968.
- [16] Hemalakshmi, G.R., Santhi, D., Mani, V.R.S., Geetha, A. and Prakash, N.B., 2021. Classification of retinal fundus image using MS-DRLBP features and CNN-RBF classifier. *Journal of Ambient Intelligence and Humanized Computing*, 12(9), pp.8747-8762.
- [17] Kouroupis, M., Korfiatis, N. and Cornford, J., 2020. Artificial intelligence-assisted detection of diabetic retinopathy on digital fundus images: concepts and applications in the National Health Service. *Innovation in Health Informatics*, pp.261-278.
- [18] Dharmana, M.M. and Aiswarya, M.S., 2020, July. Pre-diagnosis of diabetic retinopathy using blob detection. In *2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA)* (pp. 98-101). IEEE.
- [19] Hassan, H.A., Yaakob, M., Ismail, S., Abd Rahman, J., Rusni, I.M., Zabidi, A., Yassin, I.M., Tahir, N.M. and Shafie, S.M., 2020, February. Detection of proliferative diabetic retinopathy in fundus images using convolution neural network. In *IOP Conference Series: Materials Science and Engineering* (Vol. 769, No. 1, p. 012029). IOP Publishing.
- [20] Freeda, J., Priyanka, K., Priyanka, S. and Sakthipriya, V.R., Detection of Disease from Human Fundus Images using Image Processing Approach.
- [21] Prem, S.S. and Umesh, A.C., 2020, October. Classification of exudates for diabetic retinopathy prediction using machine learning. In *2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA)* (pp. 357-362). IEEE.
- [22] Shanthini, A., Manogaran, G., Vadivu, G., Kottilingam, K., Nithyakani, P. and Fancy, C., 2021. Threshold segmentation based multi-layer analysis for detecting diabetic retinopathy using convolution neural network. *Journal of Ambient Intelligence and Humanized Computing*, pp.1-15.
- [23] Shoba, S.G. and Therese, A.B., 2020. Detection of glaucoma disease in fundus images based on morphological operation and finite element method. *Biomedical Signal Processing and Control*, 62, p.101986.
- Vanakovarayan, S., Kumar, R.S., Chinnapparaj, S. and Kumar, S.S., 2021. Classification using fundus image of eye by CNN method and self-sufficient glaucoma detection using IOT. *Materials Today: Proceedings*, 37, pp.2812-2817.
- [24] S.Kannadhasan and R.Nagarajan, Performance Analysis of Circular Shape Microstrip Antenna for Wireless Communication System, AICTE Sponsored International E-Conference on Computing and Communication Systems for a Fourth Industrial Revolution (AICERA 2020), Amal Jyothi College of Engineering, Kerala, 14-16 December 2020, Published for IOP Conference Series: Materials Science and Engineering, Vol No: 1085, 2021, doi:10.1088/1757-899X/1085/1/012013
- [24] Michael David Abr'amoff, Yiyue Lou, Erginay Ali, Clarida Warren, Amelon Ryan, James C. Folk, MeindertNiemeijer, Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning, *Invest. Ophthalmol. Vis. Sci.* 57 (13) (oct 2016) 5200, <https://doi.org/10.1167/iov.16-19964>. ISSN 1552-5783..
- [25] E. Perumal, M. Elhoseny and P. T. Nguyen, "An iot-cloud based intelligent computer-aided diagnosis of diabetic retinopathy stage classification using deep learning approach," *Computers, Materials & Continua*, vol. 66, no.2, pp. 1665-1680, 2021.
- [26] Sri Laxmi Kuna , Dr. A.V. Krishna Prasad , "An Efficient Meta-Heuristic-Feature Fusion Model using Deep Neuro-Fuzzy Classifier" *International Journal of Advanced Computer Science and Applications*, Vol. 13, No. 11, 2022
- [27] L. Liu, Y. Long, P. W. Fieguth, S. Lao and G. Zhao, "BRINT: Binary Rotation Invariant and Noise Tolerant Texture Classification," in *IEEE Transactions on Image Processing*, vol. 23, no. 7, pp. 3071-3084, July 2014, doi: 10.1109/TIP.2014.2325777.
- [28] Y. Zhou et al., "Speckle Noise Reduction for OCT Images Based on Image Style Transfer and Conditional GAN," in *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 1, pp. 139-150, Jan. 2022, doi: 10.1109/JBHI.2021.3074852.
- [29] Z. Yan, Y. Chen, M. Zhang and H. Lv, "Finite-Time Annular Domain Bounded Control of Itô-Type Stochastic Systems With Wiener and Poisson Random Disturbance," in *IEEE Access*, vol. 9, pp. 17284-17295, 2021, doi: 10.1109/ACCESS.2021.3053352.