[1]Ebin P. M.

[2]P. Ranjana

Comparative Analysis of Early Diabetic Retinopathy Detection: Enhanced Minimal CNN vs VGG Architecture on CLAHE- Preprocessed Retinal Images

Journal of Electrical Systems

Abstract: - Ocular impairment is one of the prominent problems affecting middle-aged individuals due to uncontrolled blood sugar levels, commonly known as Diabetic Retinopathy (DR). The small abnormalities in the retinal capillaries, called microaneurysms and intra retinal bleeding, are the initial symptoms of Diabetic Retinopathy. Clinically recognizing diabetic retinal disease is a time-consuming and difficult process due to limitations in resources and experienced doctors. Early detection is crucial in avoiding the progression of Diabetic Retinopathy, highlighting the importance of an automated DR detection method to identify symptoms in its early stages. In this paper, researchers developed an unfamiliar framework known as Enhanced Minimal Convolutional Neural Network (EMCNN) to classify Mild-DR and No-DR ophthalmic photos using a binary classification process. The proposed new model EMCNN is compared with the migration learning method using the existing framework VGG16 and VGG19 in terms of precision and effectiveness. Before being sent across the network, the fundus pictures underwent preprocessing using the Contrast Limited Adaptive Histogram Equalization (CLAHE) tactic. EMCNN is an experimental model that enjoys a minimum number of layers and batch normalization to minimize the training effort. The EMCNN model achieved 94.89% accuracy using 3100 image dataset which is a remarkable improvement when compared with VGG architectures since the VGG architecture is trained with millions of images.

Keywords: CLAHE, Deep Learning, Diabetic Retinopathy, EMCNN, VGG

I.INTRODUCTION

There exist many reasons for visual impairment and one of the major reason is diabetes. Diabetic Retinopathy (DR) is an eye problem that may occur owing to long-term diabetes and inadequate blood sugar management. DR makes damage in the blood vessels of the retina and black specks, fuzziness, and changes in vision are the main symptoms and finally the loss of vision. DR affects both eyes and may lead to the permanent blindness. The treatments of the Diabetic Retinopathy can prevent the further vision loss only. Non-Proliferative Diabetic Retinopathy (NPDR) and Proliferative Diabetic Retinopathy (PDR) are the twain substantial categories into which diabetic retinopathy stages can be separated. The main symptoms at the Non-Proliferative stage are exudates and microaneurysms, while at the Proliferative stage; the blood vessel walls leak fluid and blood. Proliferative is the severe type and leads to grow new and abnormal blood vessels, leak protein and lipids in the vitreous. Nowadays, physical exams like the Visual Acuity Test (VAT), Pupil Dilation, and Optical Coherence Tomography (OCT) are accessible, but they are time-consuming and uncomfortable for patients. The interior surface of the eye can be examined by Ophthalmoscopy, which cover retina, optic disc, macula, fovea and blood vessels. The ophthalmic photos of the eye as shown in figure 1 can give important information related to the Diabetic Retinopathy.

[1] [1]Department of CSE, Hindustan University, Padur, Chennai, India

ORCID ID: 0000-0001-8302-796X

[2]Department of CSE, Hindustan University, Padur, Chennai, India

ORCID ID: 0000-0003-4680-4998

* Corresponding Author Email: pmebin74@gmail.com

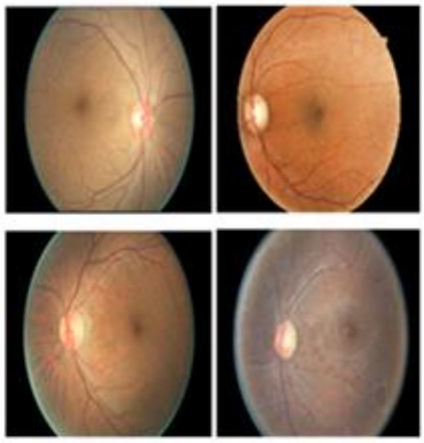


Figure. 1. No-DR (Top) Mild DR (Bottom)

Many approaches are raised by the big data and Artificial Intelligence (AI) to address the emergence and spread of DR. Artificial Intelligence was majorly used in detecting many health issues and plays paramount role in tackling DR disease to an extent. The field of AI helps a lot to cluster the disease areas and to have a proper monitoring of the cases. It also helps to predict future outbreaks based on existing cases and to analyses the mortality risks. Disease management studies the trend of the disease by allocating resources, doing proper training, maintaining proper records and finally doing pattern recognition. The field of computer science can contribute efficient methods and models for early-stage diabetes-related retinopathy detection

through Deep Learning mechanism. Artificial Intelligence (AI) can helps medical diagnosis in a large extent and can give accurate prediction in an automated way. Deep Neural networks inspired from brain can learn the patterns from the images, which can be used for disease identification in medical field. When considering Deep Learning, there is a mechanism called Transfer Learning, which can provide more accurate models in less epochs. The top layer of the pretrained model can be fine-tuned according to our need and the model can be developed. In two-class classification, the model can recognize the features like exudates and hemorrhages in the fundus images. The images can be downloaded from Kaggle, Messidor that are publically available.

1.1. Paper Contribution

This work involves the following processes.

1. To proposed an Enhanced Minimal Convolution Neural Network (EMCNN) to detect the mild diabetic retinopathy symptoms and classify mild DR and No DR varieties from the ophthalmic images of the eye.

2. To utilize EyePACS (Kaggle) publically available dataset and apply CLAHE algorithm before passing the dataset to the proposed EMCNN

3. To calculate the mean and variance of features with a batch size of 30, using batch normalization (mini batch) approach.

4. To compare the performance with two divergent deep learning frameworks, EMCNN and VGG, in the context of detecting diabetic retinopathy and their respective abilities to accurately identify early symptoms of the disease from retinal images.

5. The accuracy, Precision, Recall, and F1 score of the EMCNN are compared with VGG16 and VGG19 models after training it on the CLAHE dataset. The effectiveness of the suggested classification model was further examined using a ROC curve.

II.BACKGROUND STUDY

Recently different research papers are dealing Deep learning for the identification of diabetic retinopathy, transfer-learning methods, and produced some interesting results. Some authors are dealing with transfer learning methods and they all are used publically available datasets like kaggle and Messidor. Some of the recent works with fruitful results are given. Table 1 shows the existing body of work recap.

S. S. Karki et al [1] Diabetic Retinopathy severity experiment was conducted using an ensemble method with Efficient Nets B1 (256×256), B2 (224×224), B3 (256×256) and B5 (256×256). The model was combined with Different EfficientNet models and finally it achieved 92% quadratic kappa score. The datasets were used is EyePACS plus APTOS. In the pre-processing, radius reduction method was used to eliminate the imperfections around the circular border. The model was trained on 20 epoch using Tesla K80 GPU. The EfficientNet B3 performed better on the test set as compared to others.

R. S. Rajkumar et al [2] prepared the model using the advantage of Transfer Learning with ResNet-50 to reduce the vanishing gradient problem. The model was used with 35,000 images resized to 512×512, and an image thresholding method was applied to separate the object from the foreground pixels and background pixels. Tesla V100 GPU was used to execute the framework and a veracity of 89.4%, specificity of 97% and sensitivity of 57% was achieved. The suggested model fared well when compared to other models, according to the comparison.

The researchers [3] are proposed a hardware compatible Binary Convolutional Neural Network (BCNN) framework to reduce the memory consumption and speed up the execution. The convolution and dense layers are binarized. The binarized activations and biases reduces the number of parameters, which reduces the memory occupancy. The BCNN functions as a stabilization term to manage overfitting without the usage of Dropout layers. The dataset were used is EYEPACS and the experiments were carried out in NVDIA GPU system. The model is compared with others and gain 37.50% of downturn in memory utilization, 49.34% gain in run time. The researchers [4][29] proposed a method using the combination of VGG network and auto encoder, which prevents the overfitting while training the model. VGG network is used for transfer learning purpose. Auto encoder encodes the input image and VGG nets inspire it. Data augmentation and batch normalization was applied in the pre-processing stage. ReLU activation function was used in all the layers of auto encoder and classifier. The training was done in 100 epochs. The optimizer is used as Adam with learning rate 0.001 and parameters 0.9 and 0.999. The auto encoder is trained for the first 50 epochs and for the next 30 epochs; the auto encoder and classifier are jointly trained on the reconstruction loss and classification loss. Finally, in the last 20 epochs the encoder weights are freeze and train only the classifier weights on the classification loss. For the training, the environment used as GPU based virtual machine on Google Colab.

To classify the DR stages, the researchers [5][28] used a deep learning approach using more than 3700 images. The images were collected from AECS and sectioned into 5 groups: unaffected images, incipient Nonproliferative, intermediate Nonproliferative, critical Nonproliferative, and proliferative. The dataset is imbalanced so that 85% data is put into service for training and 15% is put into service for testing. The dataset were unstable because of various resolutions. OpenCV converted the images to Gray scale using weighted method. Laplacian method was used to drop the blurred images. The image is then resized to 256×256 and cropped it circularly o remove the background. Augmentation is applied to increase the diversity in data sets. The researchers used Dense Net as their CNN model in which the pooled information from the preceding layer is distributed to each subsequent layer. The model serving as the Adam optimizer, the loss function was binary cross entropy, and the measure was accuracy. The model obtained the training accuracy 96.98% and validation accuracy 96.69%. Early stopping method is provided to prevent overfitting which is happened after 11th epoch. The researchers also performed fivefold cross validation got accuracy 83.6% and kappa score of 88.4%. The training setup was in Linux environment with ASPEED Technology. The researchers [6][25] proposed a CNN ensemble based framework to classify and detect Diabetic Retinopathy stages via colour fundus images. The ensemble method combines several machine learning models and the researchers combines Xception, Inceptionv3, Resnet50, Dense121 and Dense169. Kaggle dataset were used and some pre-processing techniques like resizing, up sampling flipping and down sampling. 64 percent data is put into service for training, 20% for testing and 16% for validation were used. Accuracy of 80.8% recall, 51.5% specificity, 86.72% precision, and a 53.74% F1 score were achieved using the model that was trained on the NVIDIA Tesla k40.

Table 1. Summary of Existing Models

Research	**Methodology & Model Used**	**Dataset Used**	**Results/ Accuracy obtained**
S. S. Karki et al[1]	Ensemble method with Efficient Nets B1, B2, B3, and B5. Radius reduction method to reduce irregularities along the circular border	EyePACS plus APTOS	92% quadratic kappa score, EfficientNet B3 performed best
R. S. Rajkumar et al [2]	Migration Learning with ResNet-50. Using thresholding, distinguish objects from foreground and background pixels in an image.	Kaggle dataset	accuracy of 89.4%, specificity of 97%, and sensitivity of 57%
Kolla,M.&T,V.[3]	Hardware- compatible Binary Convolutional Neural Network (BCNN) model. Binarized convolution and dense layers, binarized activations and biases to reduce number of parameters	Kaggle dataset	37.50% reduction in memory usage, 49.34% gain in run time
N.Barhate, S.Bhave et al [4]	Combination of VGG network and auto encoder. Application of batch normalization and data augmentation, ReLU activation function used in all layers	Kaggle dataset	96.98% accuracy for training and 96.69% for validation
A.Singh and W.Kim [5]	Dense Net. OpenCV to convert images to grayscale, Laplacian method to drop blurred images, augmentation applied to increase diversity	Kaggle dataset	With fivefold cross validation, accuracy was 83.6% and kappa score was 88.4%.
S.Qummar et al.[6]	CNN ensemble based framework with Xception, Inceptionv3, Resnet50, Dense121 and Dense169. Resizing, up-sampling, flipping, and down-sampling	Kaggle dataset	Recall of 51.5%, specificity of 86.72, precision of 63.85%, and F1 score of 53.74 percent make up the accuracy of 80.8%.
R. N. Lazuardi et al [7]	EfficientNet-B4 and EfficientNet-B5. To enhance contrast and image centre cropping, use the Contrast Limited Adaptive Histogram Equalisation (CLAHE) technique.	Kaggle dataset	Kappa score of 0.7922, F1 score of 0.8269, and accuracy of 83.87% for EfficientNet-B4

N. S, S. S, M. J and S. C [13]	DiaNet Model (DNM) with Gabor filter	APTOS dataset from Kaggle	Accuracy obtained 90.02%
A. V. Kumar and A. S. Babu [14]	Densenet-169, ConvLSTM, Dense-LSTM	MESSIDOR dataset	Densenet-169: 94% ConvLSTM: 99% Dense-LSTM: 83%
P. Hatode, M. M. Edinburgh and M. Jha [15]	ResNet-50	Kaggle dataset	Accuracy obtained 91.60%
M. T. Al-Antary and Y. Arafa [16]	MSA-Net with deep CNNs	EyePACS, APTOS datasets	EyePACS: 77%, APTOS: 82%
Y. S. Boral and S. S. Thorat [17]	Neural network with SVM	Kaggle dataset	Accuracy obtained98.885%
Islam, Sheikh Muhammad Saiful [18]	CNN with data augmentation	Kaggle dataset	Sensitivity: 98%, Specificity: 94%
K. C. Pathak, R. B. Shah et al.[19]	Various techniques including DCNN	Kaggle dataset	DCNN: 96.5%
J. De La Torre, A. Valls and D. Puig [20]	Interpretable deep learning classifier	Kaggle dataset	Sensitivity > 90%
Gangwar, Akhilesh Kumar, and Vadlamani Ravi [21]	Migration learning and deep learning	Messidor-1 dataset	Test accuracy: 73.33%
Gao, Zhentao, et al.[22]	Deep CNN models	DIARETDB0 and DIARETDB1 datasets	Accuracy: 88.72%
Liu, Tieyuan, et al .[23]	Symmetric convolutional neural networks	DIARETDB1 dataset	Accuracy: 93.6%
Nasir, Nida, et al.[24]	CNN for microaneurysms detection	IDRID dataset	Sensitivity: 94%, Accuracy: 97%

R. N. Lazuardi et al [7] used EfficientNet-B4 and EfficientNet-B5 to train and detect the Diabetic Retinopathy on Kaggle data sets. Using the Contrast Limited Adaptive Histogram Equalisation (CLAHE) technique, the pictures are pre-processed to enhance the contrast in photos and image central cropping. The images were standardized in to256×256 and 512×512 pixels. 20% of the photos were utilised for testing, while the remaining 80% were used for training. For the initial training 256×256 pixels images were used and for progressive training 512×512 pixels images were used. In total, one model is trained for 100 epochs. Dataset is highly imbalanced so that the evaluation matric used in this research is kappa score. For EfficientNet-B4 the kappa score 0.7922, F1 score 0.8269 and Accuracy 83.87% obtained. For Efficient-B5 the kappa score 0.7931, F1 score 0.8265 and Accuracy 83.89% obtained.

The researchers [8][27] did DR classification with eleven pretrained deep learning models using OCT images, selected best one and optimized. The optimization process decreases the training time and maintain high classification accuracy. The database contains four CNV, DME, DRUSEN and NORMAL images. Pre-processing consists of the steps of removing white pixels from the image, image segmentation into 3 levels, converting segmented pictures into color pictures and subtracting the green channel, image conversion into binary, the removal of tiny items, and the identification of ROI. During the training period, validation was performed every five iterations to avoid overfitting problems. Out of eleven pre trained deep learning model, DenseNet-201 obtained 97%accuracy, 99% specificity and 97% precision. The experiments were conducted in MATLAB. The optimized DenseNet-201 is used as the pattern identifier and trained an ANN to conduct the diagnosis of Diabetic Retinopathy. The proposed model (DenseNet-201 + ANN) obtained 98% accuracy, 99% specificity and 98% precision. Chetoui M et al [9] used EfficientNET-B7 CNN model and fine-tuned to detect Diabetic Retinopathy. The two publically available datasets EyePACS and APTOS2019 were used. The researchers added Global Average Pooling (GAP) to improve the accuracy and reduce overfitting. After that a dense layer of 1024 and Dropout of 25% was added. Softmax gave the probability prediction score later.

In the pre-processing stage, images are resized in to 224×224 pixels, the experiment were carried out for 200 epochs with batch size 64. For Referable Diabetic Retinopathy (RDR), the model obtained 91% sensitivity, 98% specificity and 98% AUC on EyePACS dataset. For APTOS datasets, 97% specificity, 91% sensitivity and 96% AUC was scored. For VTDR, the model obtained 99%AUC, 98%sensitivity and 93% specificity on EyePACS dataset. For APTOS datasets, 92% specificity, 99% sensitivity and 99% AUC was scored. The researchers [10] are focused to fewer the number of learnable parameters to improve the accuracy for better classification. EyePACS dataset is the main source of dataset, which is downloaded from the kaggle, and the pre-processing such as image resizing, min-max normalization and augmentation are performed. To produce a steady size output vector and to avoid the problem of resolution reduction, a Spatial Pyramid Pooling Layer (SSP) is fixed amid the end convolution layer and foremost fully connected layer. The data is nonlinear and to deal the non-linear behavior the researchers add NiN, which is an assortment of micro networks, termed mlpconv layer on the top of the SSP layer. The parametric Relu (PRelu) function helps to overcome the overfitting problem. The framework was trained on the NVIDIA Tesla K40 and obtained an average recall of 55.6, precision of 67, specificity of 91, and F1 score of 59.6. Additionally, it achieved a micro AUC of 95.0 and a macro AUC of 84.0. To hasten the discovery of diseases, the researchers [11] proposed a model to segment the capillaries abnormalities present in the picture and to localize and detect the object. The retinal images were taken from Kaggle competition APTOS 2019 and 2015 with total 23302 images ranging from 0 to 4 classifications including No DR, Incipient, Intermediate, Profound and Proliferative. In the data-augmentation the researchers used flipping method, improves the brightness and contrast. The image size is reduced to 256×256 and Gaussian blur subtraction method was exercised to enhance the portrayal to get more fine details from the image. For training, the researchers used ResNet50, DenseNet12, DenseNet169 and DenseNet256. The model is trained with 10 epochs and the optimization technique used is Adam. The best model DenseNet121 got kappa score of 80.02.The researchers implemented a website and the user can upload retina image and predict the result.

K, Harihanth et al [12][26] created the model, which is trained with green channel of the image with highly imbalanced dataset and the ensemble method with five CNN. The image channels (Red, Green and Blue) are distinctly passed through the framework and the result is analyzed. The EyePACS images that are resized to 512 ×512 were used for training purpose. To maintain the shape of 512x512x3, the channels of the images are separated and then stacked. Data augmentation such as vertical flipping and horizontal flipping was introduced to reduce overfitting. The effectiveness indicators was precision 70%, Recall 51% and F1 score 56% was obtained and the model was compared with another work shows better performance in current model. Red channel, blue channel and green channel confusion matrix are compared and it is noticed that the accuracy of the model using red channel was 0.7489, Blue channel was 0.7386 and green channel was 0.8185.

III.SUGGESTED APPROACH

1.2. Dataset Cleansing and Preparation

Image enhancement is a fundamental aspect of image processing, aiming to improve the perceptual quality and interpretability of images. One common challenge is dealing with images that exhibit uneven illumination and contrast variations across different regions. The Contrast Limited Adaptive Histogram Equalization (CLAHE) strategy addresses this issue by locally adjusting the histogram of an image to enhance details while avoiding over amplification of noise. The CLAHE algorithm operates on the principle of adaptive histogram equalization. It dissect a portrayal into non-overlapping blocks or tiles, and within each tile, the histogram equalization is performed. However, to prevent excessive contrast enhancement, CLAHE introduces a contrast limiting mechanism. We divide the input image into smaller non-overlapping tiles and calculated the histogram of intensity values for each individual tile. The histogram represents the distribution of pixel intensities in that specific tile. After that we applied clipping to the histogram bins in each tile and limits the intensity values that stretched during the histogram equalization process. The clip limit 8 gave a decent result on our kaggle dataset. We took dimensions of each contextial region to be a square of 32X32 .The excessive stretching of very high or very low intensity values is limited to prevent noise amplification.

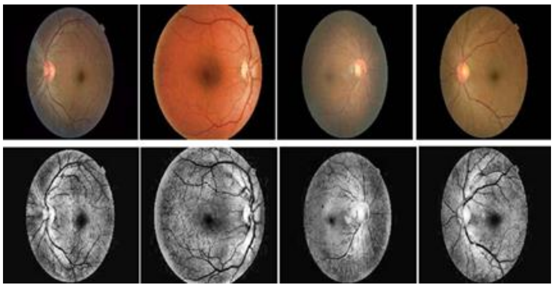


Figure. 2. Fundus image and its corresponding CLAHE image

The cumulative distribution function (CDF) for each tile based on the clipped histogram was calculated and normalized the calculated CDF. After enhancing each tile separately, we combined them to reconstruct the enhanced image. Data sets are available openly and for this analytical intent, the data set is downloaded from Kaggle in to the local system. The data set is refined to remove the dark images, blur images, and combine both images to make a proper number of deserts. Datasets and its proper ratio is very much important while training and testing a model. A sample fundus image from kaggle is converted to CLAHE as shown in figure 2.

In the pre-processing stage, the data set is passed to an algorithm termed as Contrast Limited Adaptive Histogram Equalization (CLAHE) which takes care the over amplification of the picture contrast. The working principle behind the CLAHE is that it operate on very small regions (tiles) in the images. CLAHE is the advanced version of AHE, which is commonly used to improve the image contrast but over amplify the noise. To resolve this problem CLAHE exist in path. Tile generation, histogram equalization, and bilinear interpolation are the basic three parts in CLAHE. The clip limit β is given by the equation (1). 1550 CLAHE No DR and 1550 CLAHE Mild DR pictures are used for the model training purpose. The CLAHE pre-processing stage and the output obtained is shown in figure 3.

$$\beta = mn \ \ 1+\propto 100Smax-1 \qquad (1)$$

Where m is the number of pixels and n is the gray levels in each region. $\propto$ is the clipping factor and Smax is the slope of transformation function.

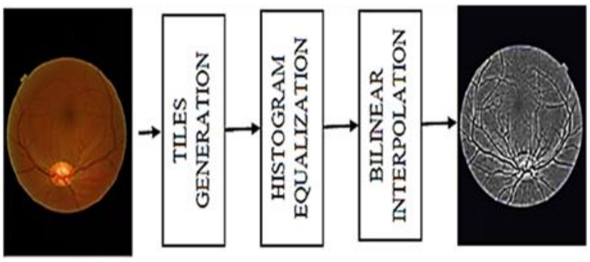


Figure.3. CLAHE Image Processing Step

3.2. Proposed Architecture EMCNN

The Enhanced Minimal Convolutional Neural Network (EMCNN) enjoys the sequential architecture. The patterns initially have a convolution1 layer of 64 filters of size 3×3 and it accepted the input picture shape of 224×224×1. The initial convolution1 layer uses channel first format. Followed by convolution1 layer ReLu activation function, maxpooling with pool size (2, 2) and drop out layer is used. Three stack are added in a sequence style containing Convolution, Maxpool and Dropout. The terminal dense layer (Fully connected layer) having 2 neurons and softmax is the activation function. The architecture employs ‘categorical cross entropy’ as the loss function; ‘rmsprop’ as the optimizer and ‘accuracy’ is the metrics. The model architecture is shown in figure 4 and has total 603330 trainable parameters.

Model: "sequential"

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 64, 224, 224)	640
activation (Activation)	(None, 64, 224, 224)	0
max_pooling2d (MaxPooling2D)	(None, 32, 112, 224)	0
dropout (Dropout)	(None, 32, 112, 224)	0
conv2d_1 (Conv2D)	(None, 10, 37, 128)	258176
activation_1 (Activation)	(None, 10, 37, 128)	0
max_pooling2d_1 (MaxPooling2	(None, 5, 19, 128)	0
dropout_1 (Dropout)	(None, 5, 19, 128)	0
conv2d_2 (Conv2D)	(None, 1, 6, 256)	295168
activation_2 (Activation)	(None, 1, 6, 256)	0
max_pooling2d_2 (MaxPooling2	(None, 1, 3, 256)	0
dropout_2 (Dropout)	(None, 1, 3, 256)	0
flatten (Flatten)	(None, 768)	0
dense (Dense)	(None, 64)	49216
activation_3 (Activation)	(None, 64)	0
dropout_3 (Dropout)	(None, 64)	0
dense_1 (Dense)	(None, 2)	130
activation_4 (Activation)	(None, 2)	0

Total params: 603,330
Trainable params: 603,330
Non-trainable params: 0

Figure.4. **EMCNN Model Architecture at Implementation Level**

The EMCNN is trained with 3100 images. Figure 5 displays the block diagram in detail. The input picture size was 224×224×1. The model is implemented with the help of the GPU Nvidia Geforce GTX graphics processor, Jupyter notebook and Tensor flow. Early stopping mechanism is used while training time and batch size 32. The training is done with 20 epochs and acquired good result when compared to the earlier implementation. The training, validation loss and accuracy is illustrated in the image 6. In the graph, the X-axis demonstrates the duration of epochs and Y-axis depicts accuracy. The blue line indicates training phase accuracy and the green line indicates vindication time accuracy. Similarly, in the case of training loss and validation loss, After 5 epochs, training accuracy increases and training loss decreases gradually. On 20 epochs, the model gave the training accuracy 0.9799 and the validation accuracy 0.92.

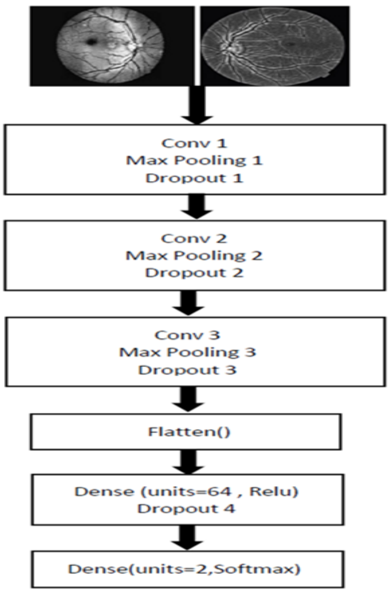


Figure.5. EMCNN Architectural Block Diagram (Flow of Layers)

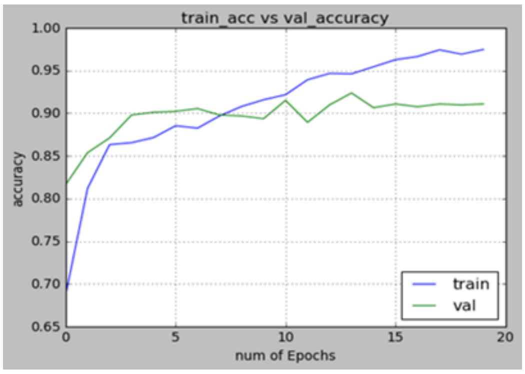


Figure.6. EMCNN Model Training, validation accuracy

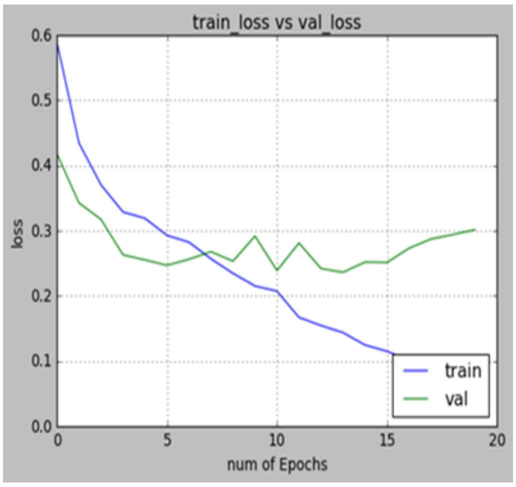


Figure.7. EMCNN Model Training, validation loss

3.3. Migration Learning with VGG Architectures

Migration learning is a formidable approach in the field of deep learning where a pre-trained neural network, often trained on a large and diverse dataset, is fine-tuned for a precise chore or dataset of interest. It has been a game-changer in many computer vision applications, incorporating medical image analysis, and it offers several advantages, such as reduced training time, improved convergence, and the ability to achieve good performance with limited data. In our research work migration learning likely played a significant role to find the performance of the deep learning outlines VGG16 and VGG19 on the same dataset used to train EMCNN. Migration learning allows us to leverage the knowledge acquired by models on the datasets. By fine-tuning these models on our dataset specific to diabetic retinopathy, they can effectively adapt the pre-trained features to the target domain, mitigating the need for an enormous amount of annotated data. When fine-tuned on diabetic retinopathy data, VGG can adapt their representations to focus on the relevant disease-specific features while retaining the ability to recognize general image patterns. By comparing the performance of these three architectures, the researchers could gain insights into which one is better suited for the task of early symptom identification in diabetic retinopathy. The numbers in VGG16 and VGG19 denote the total number of weight layers. VGG16 has 16 weight layers, embracing 13 convolutional layers and 3 fully connected layers, while VGG19 has 19 weight layers, consisting of 16 convolutional layers and 3 fully connected layers.

The researchers freeze some initial layers of the pretrained models. The earlier layers capture low-level features like edges and textures, which are generally applicable to most image analysis tasks. These layers are usually frozen to retain the knowledge. We replaced the final classification layer of the pretrained models with a new output layer as shown in the figure 8 and 9. The new output layer should have the identical quantity of nodes as the number of classes in your diabetic retinopathy dataset. This step ensures that the model's output matches the specific task. During training, the model adjusts its parameters to minimize the chosen loss function. This process fine-tunes the pretrained features for the specific task. Regularization techniques like dropout were applied to inhibit overfitting, and early stopping was implemented to prevent overfitting. We split our dataset into training and validation sets to keep an eye on the model's results during training. After fine tuning our VGG16 fine-tuned model have 6439362 trainable parameters, 14714688 non trainable parameters and VGG19 fine-tuned model have total parameters of 26463746 shown.

```
flatten (Flatten)          (None, 25088)        0
dense (Dense)              (None, 256)          6422784
dropout (Dropout)          (None, 256)          0
dense_1 (Dense)            (None, 64)           16448
dropout_1 (Dropout)        (None, 64)           0
dense_2 (Dense)            (None, 2)            130
=================================================================
Total params: 21,154,050
Trainable params: 6,439,362
Non-trainable params: 14,714,688
```

Figure .8. VGG16 architecture with newly added layers

```
flatten (Flatten)          (None, 25088)        0
dense (Dense)              (None, 256)          6422784
dropout (Dropout)          (None, 256)          0
dense_1 (Dense)            (None, 64)           16448
dropout_1 (Dropout)        (None, 64)           0
dense_2 (Dense)            (None, 2)            130
=================================================================
Total params: 26,463,746
Trainable params: 6,439,362
Non-trainable params: 20,024,384
```

Figure .9. VGG19 architecture with newly added layers

3.4. EMCNN Frame Work

The model is trained with 3100 fundus CLAHE images out of which 1550 are No-DR images and 1550 are Mild-DR images. The researchers followed a 70% to 30% split for choosing training and testing images. The recommended schema EMCNN is tested with 930 fundus CLAHE images, out of 930 pictures, 406 pictures are recognized as True Positive and 441 images are recognized as True Negative. The outcomes are shown in the form of Confusion Matrix, a table accompanied by four different combinations of genuine and predicted values, figure 10. The model shows additional aspects like precision, recall and F1 score. The ratio of true positives to false positives + true positives represents accuracy. It can be given as Precision = TP/ (FP+TP). Recall is the correlation of true positive to the total number of actual positive samples. It can be given as Recall= TP/ (FN+TP). F1 score is the weighted average of the given model's precision and recall. The F1 score can be represented by F1 = two*(precision*recall)/ (precision+ recall). The EMCNN gave the values for mild DR, recall is 0.89, F1 score is 0.91, and precision is 0.92. The value for No-DR, the precision 0.90, recall 0.93 and F1 score 0.91. The values obtained are shown in the figure 11.

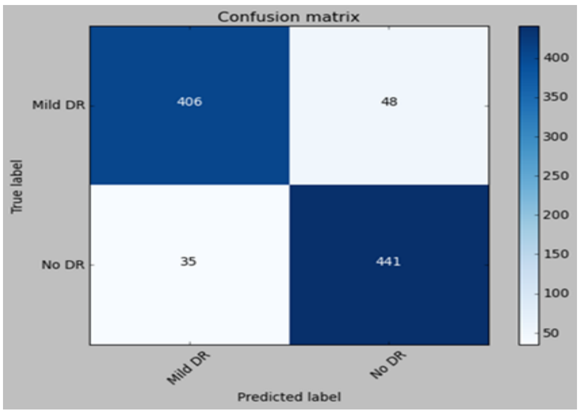


Figure.10. Confusion Matrix from EMCNN Model

Receiver Operating Characteristics (ROC) curve is subsequent parameter to define the architecture's performance. ROC graph plots using pair of variables namely True Positive (TP) and False Positive (FP). It can be drawn TP against FP. The graph is shown in figure 12 where, the area under the curve is 0.97 for incipient and No DR classification means the proposed classifier is a strong and excellent one.

```
              precision    recall  f1-score   support

     Mild DR       0.92      0.89      0.91       454
       No DR       0.90      0.93      0.91       476

    accuracy                           0.91       930
   macro avg       0.91      0.91      0.91       930
weighted avg       0.91      0.91      0.91       930

[[406  48]
 [ 35 441]]
Confusion matrix, without normalization
[[406  48]
 [ 35 441]]
```

Figure.11. Precision Recall and F1 score of EMCNN

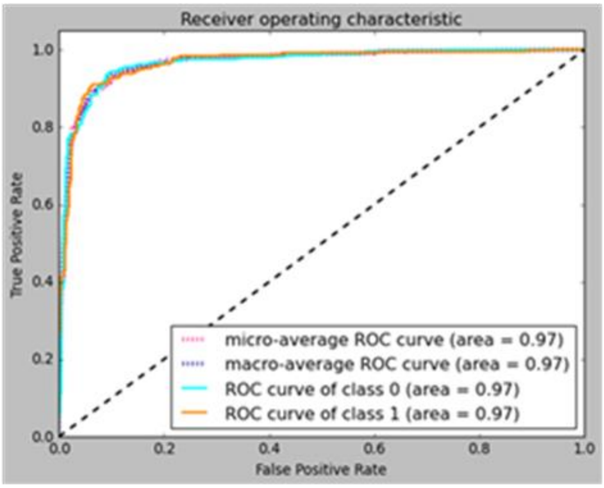


Figure.12. ROC Curve obtained from EMCNN

The table 2 shows the model performance comparison with recently developed models. The Ensemble Approach [6] method has accuracy of 0.80, precision 0.63, recall 0.51 and F1 score 0.53 using Kaggle dataset. The EfficientNet [7] has obtained the accuracy of 0.83 and F1 score 0.82 using the dataset Kaggle. The VGG-NIN [10] obtained the accuracy 95% and Individual channel training [12] model obtained 81%. With 0.97 accuracy, 0.92 precision, 0.93 recall, and 0.91 F1 score computed using Kaggle datasets, the table 2 analysis manifest that the proposed framework outperformed other current models.

Table 2. Proposed protocol parameter consideration

Model	Accuracy	Precision	Recall	F1 Score	Dataset Used
Ensemble Approach [6]	0.80	0.63	0.51	0.53	Kaggle
EfficientNet [7]	0.83	-	-	0.82	Kaggle
VGG-NIN [10]	0.95	0.67	0.55	0.59	Kaggle
Individual channel training [12]	0.81	0.70	0.51	0.56	Kaggle
EMCNN	**0.94**	**0.92**	**0.93**	**0.91**	**Kaggle**

3.5. VGG Frame Work

The method for detecting diabetic retinopathy follows two architectures namely VGG16 and VGG19. The proposed system follows the proper fine tuning of VGG architectures and the comparison is carried out to find the best fine-tuned model. The researchers used the same CLAHE fundus images to train the frameworks. The table II exhibits the outcomes obtained from the fine-tuned models. VGG16 procured a precision of 99%, recall of 94%, F1 score of 96% for mild DR case and 94%, 99%, 97% for No DR case respectively. The overall test loss 0.097 and 96.5% accuracy obtained from fine-tuned VGG16 model. VGG19 procured a precision of 97%, recall of 95%, F1 score of 96% for mild DR case and 96%, 97%, 96% for No DR case respectively. The overall test loss 0.166 and 96.1% accuracy obtained from fine-tuned VGG19 model. Out of 930 test images VGG16 identified 426 as True positive and 472 as True negative. VGG19 identified 433 True positives and 461 True negatives. The figure 13 shows the generated confusion matrix of both VGG16 and VGG19. The obtained precision- recall-F1 score measures screen shot is shown in figure 14. ROC curve is the parameter to justify the research work which is shown in figure15. VGG16 obtained micro-average ROC curve area 1.00 and VGG19 scored micro-average ROC curve area 0.99.handled by implementing flat routing and gradient-based routing [28].

It is observed that already some of the mechanisms, like traditional energy harvesting and the piezoelectric nano-generator concept of energy harvesting, have been introduced to improve the performance of energy-oriented multipath routing, but there are still some challenges, like the high installation charge and low energy quantity [29] [30]. For this reason, these mechanisms need to be reviewed. We have noticed that restricted energy

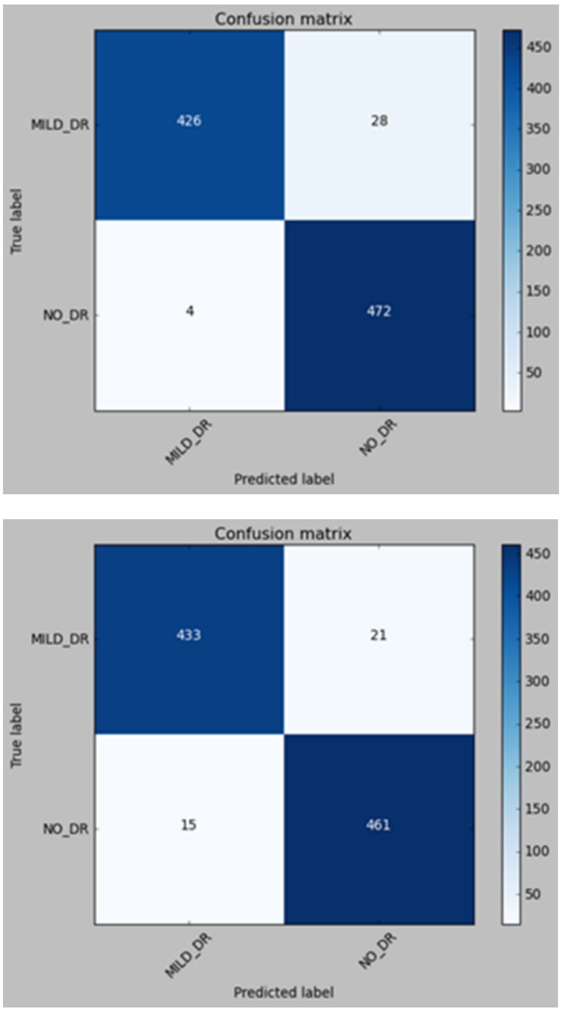


Figure.13. Confusion matrix obtained from VGG16 and VGG19 respectively

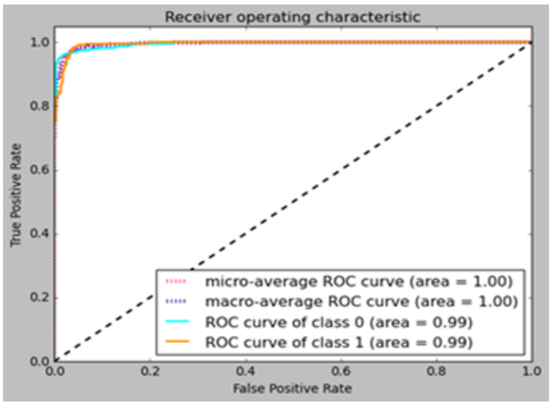
```
              precision    recall  f1-score   support

     MILD_DR       0.99      0.94      0.96       454
       NO_DR       0.94      0.99      0.97       476

    accuracy                           0.97       930
   macro avg       0.97      0.96      0.97       930
weighted avg       0.97      0.97      0.97       930

[[426  28]
 [  4 472]]
Confusion matrix, without normalization
[[426  28]
 [  4 472]]
Normalized confusion matrix
[[0.94 0.06]
 [0.01 0.99]]
```

```
              precision    recall  f1-score   support

     MILD_DR       0.97      0.95      0.96       454
       NO_DR       0.96      0.97      0.96       476

    accuracy                           0.96       930
   macro avg       0.96      0.96      0.96       930
weighted avg       0.96      0.96      0.96       930

[[433  21]
 [ 15 461]]
Confusion matrix, without normalization
[[433  21]
 [ 15 461]]
Normalized confusion matrix
[[0.95 0.05]
 [0.03 0.97]]
```

Figure.14. Confusion matrix obtained from VGG16 and VGG19 respectively

Table 3. Performance Comparison of Fine-Tuned Models with EMCNN

		Precision	**Recall**	**F1 Score**	**Test Loss**	**Test Accuracy**
VGG 16	Mild DR	0.99	0.94	0.96	0.097	**0.965**
	No DR	0.94	0.99	0.97		
VGG 19	Mild DR	0.97	0.95	0.96	0.166	**0.961**
	No DR	0.96	0.97	0.96		
EMCNN	Mild DR	0.92	0.83	0.91	0.233	**0.948**
	No DR	0.90	0.93	0.91		



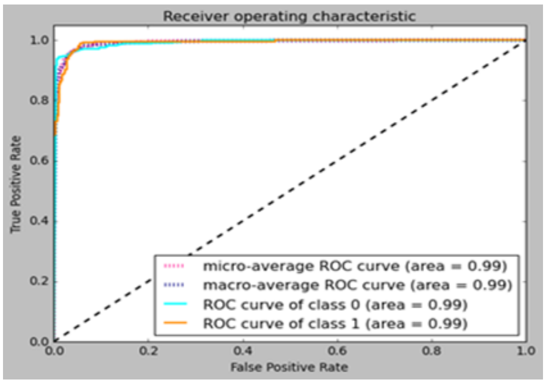


Figure15. **ROC Curve obtained from VGG16 and VGG19 respectively**

IV.CONCLUSION

The work is focused on creating an Enhanced Minimal CNN (EMCNN) model to classify Diabetic Retinopathy (DR) stages and detect mild symptoms of DR. The model is created with minimum number of convolution, pooling layers and an experiment is conducted to detect mild DR symptoms and classify Mild DR, No-DR images with minimum effort. Proposed model is trained with fundus images, which are pre-processed with CLAHE algorithm to enhance the contrast and reduce the noise. The image sets are resized in to 224×224 before passing to the EMCNN. The proposed EMCNN model gave the accuracy 94.89% and F1 score of 0.91. The model is evaluated against other models and the difference gives excitement. The model employs 3100 photos, which are restricted in number, so that the model may be trained with additional images in the future to enhance accuracy and other parameters. When compared with fine-tuned models of VGG architectures, the migration learning approach obtains 96% accuracy since the models are pre trained with millions of images and already learnt the features. The proposed model EMCNN introduced by the researchers is trained with 3100 fundus CLAHE pre-processed images and achieved the accuracy near to fine tuning method with limited number of dataset. In future the research will extended with training the EMCNN using more number of CLAHE images and comparison with the pre-defined architecture called Inception V3. The researchers are interested to apply distinct pre-processing methods in fundus portraits and find out exemplary result pertaining to accuracy.

Conflicts of interest

No competing interests are relevant to this paper.

REFERENCES

[1] S. S. Karki and P. Kulkarni, "Diabetic Retinopathy Classification using a Combination of EfficientNets," 2021 International Conference on Emerging Smart Computing and Informatics (ESCI), 2021, pp. 68-72, doi: 10.1109/ESCI50559.2021.9397035.

[2] R. S. Rajkumar, T. Jagathishkumar, D. Ragul and A. G. Selvarani, "Transfer Learning Approach for Diabetic Retinopathy Detection using Residual Network," 2021 6th International Conference on Inventive Computation Technologies (ICICT), 2021, pp. 1189-1193, doi: 10.1109/ICICT50816.2021.9358468.

[3] Kolla, M., & T, V. (2021). Efficient Classification of Diabetic Retinopathy using Binary CNN. 2021 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE), 244-247.

[4] N. Barhate, S. Bhave, R. Bhise, R. G. Sutar and D. C. Karia, "Reducing Overfitting in Diabetic Retinopathy Detection using Transfer Learning," 2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA), 2020, pp. 298-301, doi: 10.1109/ICCCA49541.2020.9250772.

[5] A. Singh and W. Kim, "Detection of diabetic blindness with Deep-Learning," 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2020, pp. 2440-2447, doi: 10.1109/BIBM49941.2020.9313392.

[6] S. Qummar et al., "A Deep Learning Ensemble Approach for Diabetic Retinopathy Detection," in IEEE Access, vol. 7, pp. 150530-150539, 2019, doi: 10.1109/ACCESS.2019.2947484.

[7] R. N. Lazuardi, N. Abiwinanda, T. H. Suryawan, M. Hanif and A. Handayani, "Automatic Diabetic Retinopathy Classification with EfficientNet," 2020 IEEE REGION 10 CONFERENCE (TENCON), 2020, pp. 756-760, doi: 10.1109/TENCON50793.2020.9293941.

[8] K. T. Islam, S. Wijewickrema and S. O'Leary, "Identifying Diabetic Retinopathy from OCT Images using Deep Transfer Learning with Artificial Neural Networks," 2019 IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS), 2019, pp. 281-286, doi: 10.1109/CBMS.2019.00066.

[9] Chetoui M, Akhloufi MA. Explainable Diabetic Retinopathy using EfficientNET. Annu Int Conf IEEE Eng Med Biol Soc. 2020 Jul;2020:1966-1969. doi: 10.1109/EMBC44109.2020.9175664. PMID: 33018388

[10] Khan, Z., Khan, F. G., Khan, A., Rehman, Z. U., Shah, S., Qummar, S., Ali, F., & Pack, S. (2021). Diabetic Retinopathy Detection Using VGG-NIN a Deep Learning Architecture. IEEE Access, 9, 61408-61416. [9409084]. https://doi.org/10.1109/ACCESS.2021.3074422

[11] Lands, A., Kottarathil, A.J., Biju, A., Jacob, E.M., & Thomas, S. (2020). Implementation of deep learning based algorithms for diabetic retinopathy classification from fundus images. 2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184), 1028-1032.

[12] K, Harihanth and Karthikeyan B. “Diabetic Retinopathy Detection using ensemble deep Learning and Individual Channel Training.” 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS) (2020): 1042-1049.

[13] N. S, S. S, M. J and S. C, "An Automated Detection and Multi-stage classification of Diabetic Retinopathy using Convolutional Neural Networks," 2023 2nd International Conference on Vision Towards Emerging Trends in Communication and Networking Technologies (ViTECoN), Vellore, India, 2023, pp. 1-5, doi: 10.1109/ViTECoN58111.2023.10157960.

[14] A. V. Kumar and A. S. Babu, "Diabetic Retinopathy Detection using Deep Learning Methodology," 2022 IEEE 3rd Global Conference for Advancement in Technology (GCAT), Bangalore, India, 2022, pp. 1-6, doi: 10.1109/GCAT55367.2022.9971891.

[15] P. Hatode, M. M. Edinburgh and M. Jha, "Evolution and Testimony of Deep Learning Algorithm for Diabetic Retinopathy Detection," 2022 5th International Conference on Advances in Science and Technology (ICAST), Mumbai, India, 2022, pp. 122-126, doi: 10.1109/ICAST55766.2022.10039538.

[16] M. T. Al-Antary and Y. Arafa, "Multi-Scale Attention Network for Diabetic Retinopathy Classification," in IEEE Access, vol. 9, pp. 54190-54200, 2021, doi: 10.1109/ACCESS.2021.3070685.

[17] Y. S. Boral and S. S. Thorat, "Classification of Diabetic Retinopathy based on Hybrid Neural Network," 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), 2021, pp. 1354-1358, doi:10.1109/ICCMC51019.2021.9418224.

[18] Islam, Sheikh Muhammad Saiful; Hasan, Md Mahedi; Abdullah, Sohaib. “Deep Learning based Early Detection and Grading of Diabetic Retinopathy Using Retinal Fundus Images”.12/26/2018, Database: arXiv.

[19] K. C. Pathak, R. B. Shah, R. R. Tharakan, B. N. Patel and D. C. Jariwala, "Diabetic Retinopathy Diagnosis and Categorization using Deep Learning - A Review," 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), 2021, pp. 1063- 1069, doi: 10.1109/ICICCS51141.2021.9432312

[20] J. De La Torre, A. Valls and D. Puig, “A deep learning interpretable classifier for diabetic Retinopathy”. Neurocomputing, volume 396, 5th July 2020, Elsevier.

[21] Gangwar, Akhilesh Kumar, and Vadlamani Ravi, “Diabetic retinopathy detection using migration learning and deep learning,” Evolution in Computational Intelligence, Springer, Singapore, 2021, pp. 679-689.

[22] Gao, Zhentao, et al., “Diagnosis of diabetic retinopathy using deep neural networks,” IEEE Access 7, 2018, pp-3360-3370.

[23] 23. Liu, Tieyuan, et al., “A Novel Diabetic Retinopathy Detection Approach Based on Deep Symmetric Convolutional Neural Network,” IEEE Access 9, 2021, pp. 160552-160558.

[24] Nasir, Nida, et al., “Deep DR: Detection of Diabetic Retinopathy using a Convolutional Neural Network,” Advances in Science and Engineering Technology International Conferences (ASET), IEEE, 2022.

[25] Vetrithangam, D., Shruti, P., Arunadevi, b., Himabindu, R, kumar, p. N., kumar, a. R., ... & Arnet Zitha, d. R. (2023). Optimum feature selection based breast cancer prediction using modified logistic regression model. journal of theoretical and applied information technology, 101(8).

[26] Indira, V., Umar, s., pant, b., & goyal, m. K. (2022). Discriminating the pneumonia-positive images from covid-19-positive images using an integrated convolutional neural network. Mathematical problems in engineering, 2022.

[27] Vetrithangam, D., Senthilkumar, V., Neha, A., Naresh, P., & Kumar, m. S. (2022). Coronary artery disease prediction based on optimal feature selection using improved artificial neural network with meta-heuristic algorithm. Journal of theoretical and applied information technology, 100(24).

[28] Das, P., Vetrithangam, D., Krishnan, g. H., Kumar, A., & Reddy, r. V. K. (2023). Deep learning techniques revolutionizing biomedical applications: arrhythmia detection, cardiac sensed signals, and cell-free synthetic biology. In applications of synthetic biology in health, energy, and environment (pp. 68-91). Igi global.

[29] Vetrithangam, D., Satve, P. P., kumar, J. R. R., Anitha, P., Vidhya, S., & Saini, A. K. (2023). Prediction of pneumonia disease from x-ray images using a modified resnet152v2 deep learning model. Journal of theoretical and applied information technology, 101(17).