AN IMPROVED EM-CNN DEEP NEURAL NETWORK FOR CLASSIFYING MILD-DIABETIC RETINOPATHY FROM NORMAL IMAGES

EBIN PM[1], D. P RANJANA[2]
[1]Research Scholar, Hindustan University, Department of CSE, Chennai, India
[2]Professor, Hindustan University, Department of CSE, Chennai, India
E-mail: [1]pmebin195@gmail.com, [2]pranjana@hindustanuniv.ac.in

ABSTRACT

One of the main issues middle-aged people have because of uncontrolled blood sugar is vision impairment, which is also known as diabetic retinopathy (DR). The early signs of diabetic retinopathy are minor abnormalities in the retinal capillaries known as microaneurysms and intraregional hemorrhage. Due to the lack of resources and skilled medical professionals, clinical diagnosis of diabetic retinopathy is delayed and challenging process, making early detection even more crucial to prevent the spread of the disease. Herein lies the value of an automated DR detection system to spot the early signs of DR. The research work uses a combination of two types of fundus images contrast-limited adaptive histogram equalization (CLAHE) and non-CLAHE. CLAHE is a technique used to enhance the contrast of an image by adjusting the intensity distribution, while non-CLAHE images are unaltered. Using a binary classification approach, the researchers in this publication constructed a new model called Experimental Minimal Convolutional Neural Network (EMCNN) model to categorize Mild-DR and No-DR fundus pictures. By training the EM-CNN model on both types of images, the researchers aim to improve the accuracy of DR classification. The result is compared with the model, which already exist. The model achieved 98%accuracy, which is better than existing models.

Keywords: *CLAHE, Convolutional Neural Networks, Deep Learning, Diabetic Retinopathy, EMCNN.*

1. INTRODUCTION

Significant portions of patients with diabetes who also have longer-lasting diabetes and less tightly managed blood sugar develop diabetic retinopathy (DR), an eye consequence. This often occurs in Type1 and Type2 diabetes problem. Usually, it only has moderate eyesight and affects both eyes, an issue at an early stage. Treatment does not stop DR, yet powerful in stopping additional eyesight loss. bodily evaluations OCT, pupil dilation, and the Visual Acuity Test are examples of tests Taking longer to complete and the delayed results result in therapy [1].

The fundus pictures of the eye may be viewed via ophthalmoscopy, and they offer Deep Learning crucial information regarding DR. Having diabetes mellitus for an extended length of time causes diabetic retinopathy, a disease that is extensively disseminated and ultimately leads to eyesight loss. To identify DR quickly and with less expense or therapy when the situation is severe, there have been several initiatives performed. Initial tests required a lot of effort and professional expertise, keeping in mind the feature extraction and detection method. The result was the creation of a fundus picture, which is far more intricate than a conventional one.

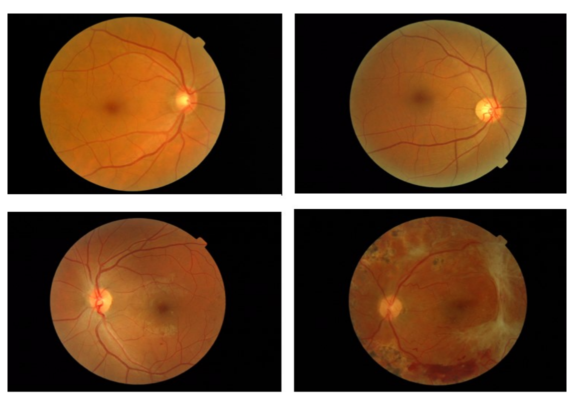

Fig. 1 Stages of Retinopathy: Top left (No DR) right top (mild DR) left bottom (Moderate DR) and Bottom right (Acute DR).

Depending on how it develops, various types of diabetic retinopathy exist like NPDR and PDR. NPDR stands for Non-Proliferative Diabetic Retinopathy and PDR stands for Proliferative

Diabetic Retinopathy. The early stage of DR, known as NPDR, can range in severity from mild to extremely severe. The severity of the DR is rated using colored pictures of the fundus. Micro aneurysms are a defining feature of mild NPDR. The last stage of retinopathy, known as PDR, is characterized by vitreous hemorrhage, retinal detachment, and neovascularization of the disc and retina. An automated method for diagnosing diabetic retinopathy in its early stages is required due to the lack of trained physicians and the laborious and drawn-out process of feature identification. Automated early stage detection is a useful tool for identifying and stopping additional vision loss. DME is a different ailment that indicates the severity of DR and is caused by the macula's fluid retention. ARIA classification methods are not seen to be sophisticated since they cannot distinguish minute differences between distinct classes of DR. Additionally, professional skill takes time. Eventually, computer-aided techniques were utilized for precise DR detection.

In existing methods, the researchers are exploring various methods to extract features from CLAHE images, such as using convolutional neural networks (CNNs) or other machine learning techniques. These methods aim to automatically identify patterns and features in the images that are indicative of DR, such as the presence of microaneurysms, hemorrhages, or exudates. However, there are still challenges in accurately identifying these features, as there may be variability in the appearance of DR lesions, and these features may be subtle or difficult to distinguish from other structures in the image. Additionally, there is a need for large, diverse datasets to train these models and ensure their accuracy and generalizability. Still there is no study takes place to investigate and compare the effectiveness and accuracy of DR detection using a mixture of CLAHE and Non-CLAHE fundus images.

In this study, the researchers investigate the effectiveness of using a combination of CLAHE and non-CLAHE fundus images for DR classification. This approach can reduce the over-enhancement and texture loss associated with CLAHE and improve the overall accuracy of the classification. In addition to the use of several pre-processing techniques, including resizing, normalization, and augmentation, to improve the quality and quantity of the training data and reduce overfitting. The researchers provided details on the training and validation of the CNN model, including the use of cross-validation and early stopping to prevent overfitting and optimize the hyperparameters. The performance of the proposed approach is evaluated using several metrics, including classification accuracy, precision, recall, F1 score and AUC-ROC, and compare it with existing DR classification methods. The study pattern as follows.

To identify the mild symptoms of diabetic retinopathy and categorize mild DR and No DR categories from eye fundus pictures, we created a new model architecture called Experimental Minimal Convolution Neural Network (EMCNN). The model is tested and compared with existing models after being trained on a mixture of CLAHE and Non-CLAHE fundus pictures from No DR and Mild DR Datasets. The study involves the following process sequence.

- The researchers developed a new model called Experimental Minimum CNN (EMCNN) model to classify Mild and No-DR images with a good accuracy.
- The Model is trained with a jumble of fundus images, which contains CLAHE, and Non-CLAHE images.
- Finally, the model is tested with 30% images to prove the accuracy level of EMCNN.

2. RELATED WORK

The Cross-disease Attention Network (CANet), graded by DR and DME, and its internal interactions, was developed by the researchers [2]. The ISBI 2018 IDRiD challenge dataset and the Messidor dataset are used to assess the model. The disease-dependent attention module is used to teach the intrinsic links between diseases, whereas the illness-specific attention module is used to educate disease characteristics. These two modules are combined to create the features of a deep neural network. The researchers utilised ResNet50 to create the feature map, and they used the lowest resolution feature map with semantic information as the input for the two disease-specific attention modules. To comprehend the DR and DME disorders, the module teaches disease-specific characteristics. The disease-dependent attention module will then discover.

To identify region-specific characteristics of unbalanced DR data distributions, the researchers created unique CAB, stands for Category Attention

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Block [3]. A Global Attention Block (GAB) is also suggested to identify information about tiny lesions. By combining the CAB and GAB for DR grading, the CABNet is finally built. The suggested model is composed of four components: a classifier, a backbone, a GAB, and a CAB. In addition to MobileNet1.0, VGG16, ResNet 50, DenseNet 121, and Xception, other backbone networks are also used. Good accuracy of 0.8569 and a Kappa score of 0.8794 were provided by MobileNet 1.0. Three publicly accessible datasets—DDR, Messidor, and EyePACS—were used in the research. The input resolution was 512×512, and data augmentation was done to lessen overfitting. All modules are Adam-optimized and trained across 70 epochs.

The researchers presented the VGG-NiN model [4], a stack of multiclass classification models that uses the VGG16, Spatial Pyramid Pooling (SPP), and Network in Network (NiN). The SPP reduces the input picture size from 1024×1024 to 224×224 by being inserted in the first fully connected layer and the final Conv layer, which produces the output vector with a fixed size. In order to learn the nonlinear behavior of data, NiN is layered on top of the SPP layer. Using knowledge transfer method, the convolution layers of the VGG network are frozen, and the fully connected layer of the network is tweaked. 50 epochs with a batch size of 8 were used to train the model. The model produced a 95% micro AUC with 52% less parameters. The researchers presented an intuitive DR grading system on the fundus pictures without using any pre-processing methods [5]. The method used a two-stage fine-tuning procedure, the first of which involved embedding the diabetic retinopathy lesion structure in a pre-trained CNN model utilizing lesion ROIs. The completely linked layers are eliminated and replaced in the second stage with a Principal Component Analysis (PCA) layer, which aids in minimizing the overfitting issue. EyePACS and Messidor, two publicly available difficult data sets, are used. The ROIs around the lesion were extracted using one of two data augmentation techniques, and they were then resized to the same 64×64 size. Once again flip each ROI horizontally and rotate each ROI four times. The ResNet152 with Gradient Boosting and re-initialized CONV1.

The researchers suggested a CNN with a two-fold feature augmentation capacity that offers additional feature-level generalization [6]. Utilizing a small data set from the Republic of Korea's Kyung Hee University Medical Centre (KHUMC), the model is verified. The model uses ResNet's residual blocks as its primary feature extractor. Weight-sharing convolution kernels are the first stage of the two-stage feature augmentation, and RCA stands for Reverse Cross Attention stream is the second step. Three components make up the RCA stream itself. Self-Contest Aggregation (SCA), Pairwise Reverse Attention (PRA), and Multi-Level Fusion make up the first, second, and third, respectively (MLF). All of the photographs have been resized to various sizes, including 224×224, 448×448 and 600×600. RCA is assessed using the Messidor dataset, which reaches

To enhance the Proliferative Diabetic Retinopathy (PDR) picture collection, the researchers developed the neovessel (NV) production technique [7]. Neovessels are brand-new blood vessels since DR phases are so severe. The model training dataset can be enhanced by using the DR labelled dataset produced by the method. According to the study, the newly created NV is only placed in the OD or the adjacent region, where it was set to 0.15, elsewhere it was 0.40, and on both sites, it was 0.45. 25% of the image's size was chosen as the ROI region. On the basis of Otsu thresholding on the green channel of the picture, the vessel segmentation is carried out. The optic disc segmentation was carried out using the UOLO framework, which is proven. The accuracy of the model is impacted by the various color aberrations and irrelevant illuminations in the dataset of eye pictures that is currently accessible. A novel image-processing framework along with stacking deep learning approach for pinpointing were proposed by the researcher [8]. The grey world color constancy method is used to enhance picture quality and brightness normalization. Three distinct CNNs are used to prepare the stack generalization model. The dataset utilized is EyePACS, and data augmentation techniques like horizontal and vertical flips, width and height shifts, fill mode, and zoom range are employed to enhance the dataset. For binary classification as well as for multiclass classification, the stack model provided test accuracy results 87.45 percent.

To analyses the existence of micro aneurysms in the fundus pictures of the eyes, the researchers devised a model [9]. A method known as the semantic segmentation algorithm is utilized to identify the existence of micro aneurysms and categorize the fundus pictures as normal and damaged. In order to diagnose Non Proliferative Diabetic Retinopathy (NPDR), the researchers used deep convolutional neural networks and fundus

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

pictures of the subjects' eyes as input. Ieee-Dataport.org is where the dataset was obtained. Enhancing the boundary line of the dark lesions during the first stage of pre-processing aids in separating it from the backdrop. The next stage involves using an ideal wideband pass filter to improve the contrast of exudates removing the optical disc, removing the artery, and a potential lesion. The researchers suggested an ensemble model that combines five deep convolutional neural networks (CNNs), namely Inception-Version3, Resnet-50, Dense121, Dense169, as well as Xception, in order to identify all stages of DR [10]. The model utilized a dataset from Kaggle with images that were each 38882951 in size. To balance the dataset, the input picture is enlarged to 786512 and up- and down-sampling procedures are performed. Training (64%), testing (20%), and validation (16%) make up the dataset. Adaptive Moment Estimation with Nesterov-accelerated is used to update the learning parameters. NVIDIA Tesla K40 GPU was used to train the model, and accuracy of 80.8% was attained. Class 0 provided the highest recall (0.97), while class 1 provided the lowest recall (0.54). With 113 samples, class 4 earned an AUC of 0.97.

The researchers put forth the Hyperparameter Tuning Inception version-4 (HPTI-v4) model, which can identify as well as categorize the various phases of diabetic retinopathy [11]. The MESSIDOR dataset is used in the experiment, and the CLAHE technique was used to increase contrast. Using a histogram-based segmentation model, the pre-processed pictures are divided into segments, and the characteristics of the segmented images are extracted using HPTI-v4. By employing CLAHE for pre-processing, undesired noise, amplification is prevented. The experiment is run using 500 epochs, a learning rate of 0.001, and momentum of 0.09. By using Bayesian optimization, the best collection of hyperparameters is selected. The HPTI-v4 reported 99.4 percent accuracy, 98.8 percent sensitivity, and 99.6 percent specificity. In this study, the researchers suggested a model that avoids overfitting while assessing the model's architecture by utilizing an Auto Encoder (AE) and VGG network (VGG AE) [12]. In an auto encoder network, both the encoder and the decoder have been trained to convert the input picture to a latent representation. The VGG network is a hit with the encoder network. EyePACs datasets are utilized for training, and after pre-processing, the pictures are reduced in size to 256×256. The data was enhanced using rotation and flipping techniques. ReLU activation function provides the best and most accurate results when compared to sigmoid, tanh, and other activation functions, according to the researchers. Test accuracy is 76.27% according to the suggested design VGG AE. The researcher's technique for categorizing the phases of diabetic retinopathy is transfer learning using residual networks [13]. The dataset is collected from the Kaggle dataset, downsized to 512×512, and with severity labels ranging from 0 to 4. ReLU is employed as the activation function on top of ResNet-50, the pretrained architecture used in the experiment. The model was executed using Tesla V100 GPU cards on a DGX server. 89.5 percent accuracy, 96.0 percent specificity, as well as 57.0 percent sensitivity were provided by the network. The model was then contrasted with others, including AlexNet, VGG19, Inception V3, and ResNet18. The model using ResNet 50 performed more accurately, specifically, and sensitively.

An ensemble method experiment on the severity of diabetic retinopathy was carried out utilizing the Efficient Nets B1 (256256), B2 (224224), B3 (256256), and B5 (256256). The model eventually obtained a 92% quadratic kappa score after being integrated with other EfficientNet models [14]. EyePACS and APTOS were the datasets utilized. The imperfections around the circular border were reduced during pre-processing using the radius reduction approach. The Tesla K80 GPU was used to train the model for 20 epoch, and the EfficientNet B3 outperformed the competition on the test set. To decrease memory use and speed up execution, researchers presented a hardware-friendly Binary Convolutional Neural Network (BCNN) model [15]. Binarization is applied to thick layers and convolution. The number of parameters is decreased due to binarized activations and biases, which also results in a decrease in memory occupancy. Without using Dropout layers, the BCNN acts as a regularizer to control overfitting. The studies were carried out on an NVDIA GPU machine, and the dataset utilized was EYEPACS. When the model is compared to others, it uses 37.50 percent less memory and runs at a 49.34 percent faster pace.

The researchers employed a deep learning strategy with more than 3700 photos to categorize the DR phases [16]. Mild-Non-proliferative, Moderate-Non-proliferative, Severe-Non-proliferative, and Proliferative are the five categories into which the photos from AECS are gathered and categorized.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Due to the imbalance in the dataset, 85 percent of the data are used for training and 15 percent for testing. The different resolutions caused the dataset to be unstable. Using a weighted strategy, OpenCV transformed the photos to grayscale. The photos that were blurry were removed using the Laplacian approach. The image is then circularly cropped to eliminate the backdrop and downsized to 256256. To make data sets more diverse, augmentation is used. As their CNN model, the researchers employed DenseNet, in which each layer received the aggregate knowledge of the layer before it. The model employed the Adam optimizer, utilized Accuracy as the metric, and used Binary Cross Entropy as the loss function. Model's accuracy in both training and validation was 96.98 percent. A strategy of early halting is offered to stop overfitting from occurring after the 11th epoch. The accuracy and kappa scores for the researchers' 5-fold cross validation were 83.6% and 88.4%, respectively. The training environment used Linux and ASPEED Technology. Summary of Related work is given below in a table format.

Table 1: Summary of Related work

SN	Paper Title	Works Done	Identifications
1	CANet: Cross-disease Attention Network for Joint Diabetic Retinopathy and Diabetic Macular Edema Grading	Disease-dependent attention module and illness-specific attention module is created for feature detection. Utilized ResNet50 to create the feature map	The disease-dependent attention module performs well. The paper not concentrating classification of mild-DR and No-DR conditions.
2	CABNet: Category Attention Block for Imbalanced Diabetic Retinopathy Grading	CAB and GAB for DR grading and the CABNet is finally built.	Accuracy 0.8569 .Kappa score 0.8794. Any type of image preprocessing method is not used and described.
3	Diabetic Retinopathy Detection Using VGG-NIN a Deep Learning Architecture	A stack of multiclass classification models that uses VGG16 architecture	The model produced a 95% micro AUC. No pre-processing methods are used in the input images.
4	Diabetic Retinopathy Diagnosis From Fundus Images Using Stacked Generalization of Deep model	Stack generalization model with grey world colour constancy method is used to improve the image quality.	Accuracy of 87.45% is achieved. Less accuracy when compared with other existing models.
5	Diabetic Retinopathy Detection Using Prognosis of Micro aneurysm and Early Diagnosis System for Non-Proliferative Diabetic Retinopathy Based on Deep Learning Algorithms	The semantic segmentation algorithm is used to detect micro aneurysms.	Accuracy 80.8% achieved. Less number of images (113) are used for training multi class classification CNN.
6	Reducing Overfitting in Diabetic Retinopathy Detection using Transfer Learning	VGG network with Auto Encoder (AE) used.	Accuracy is 76.27%. Trained with less number of images. No pre-processing.

Journal of Theoretical and Applied Information Technology
15th June 2023. Vol.101. No 11

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

3. PROPOSED MODEL

Openly accessible data sets are obtained from Kaggle and Messidor into the local system for this research purpose. The data set is cleaned to eliminate the shadowed photos, blur the photographs, and merge the two images to create the appropriate amount of deserts. The right ratio of datasets is crucial for developing and testing a model. Figure 2 displays several sample pre-processed fundus photos from Kaggle. The data set is sent to an algorithm known as Contrast.Limited.Adaptive.Histogram Equalization (CLAHE) during the pre-processing stage, which takes care of the excessive amplification of the picture contrast. The CLAHE operates on extremely small areas (tiles) in the photos as its guiding concept. The improved variant of AHE, known as CLAHE, is used to increase picture contrast but excessively amplifies noise. CLAHE is present in the way to tackle this issue. The core components of CLAHE are bilinear interpolation, histogram equalization, and tile creation. For the aim of creating the model, 3100 CLAHE No DR and 3100 CLAHE Mild DR images, a combination of Kaggle and Messidor datasets are employed. Sequential architecture benefits the Experimental Minimal Convolutional Neural Network (EMCNN). The model begins with a convolution layer of 64 neurons and accepts an input picture with the dimensions 224×224×1. Next, ReLu activation function, maxpooling with a pool size of 2×2, and drop out layer are employed. There are three new stacks, including Convolution, Maxpool, and Dropout. The activation function is located in the last fully connected layer (dense layer), which is a fully linked layer with two neurons. The model employs "accuracy" as the metrics, "rmsprop" as the optimizer, and "categorical cross entropy" as the loss function. Figure 3 depicts the model architecture, which has 603330 trainable parameters. Any picture may be expressed as the two-dimensional (2D) function f(x,y). Here x and y are spatial (plane) coordinates, and the amplitude of f at any given pair of coordinates (x,y) is referred to as the intensity or grey level of the provided image at that location. The resultant picture is considered as a digital picture when the intensity values of f and the coordinates (x, y) are all discrete, finite numbers. Only a small number of pixels, each of which has a unique position and value, make up a digital image. Images with strong contrast and vivid colors are a requirement for interpreting genuine scenarios in the field of digital picture applications, such as the identification and classification of underwater dam fractures and multitarget spotting in challenging environments. As opposed to photographs with lower contrast levels, those with higher contrast levels often exhibit a greater degree of color scale disparity. In photography, light is essential for creating photos of high quality. In contrast to weak light, which results in an image that is too dark to be seen, vigorous light gives an image a decolorized look. In these two instances, the photos' low contrast makes it impossible to see out their intricate textures. Histogram Equalization (HE), a traditional method of enhancing contrast, performs well in common pictures like portraits of people or scenes from nature. This technique disperses the most common intensity values to boost contrast over the whole picture. However, in highly homogenous areas, it experiences noise amplification. Adaptive histogram equalization (AHE) is a local histogram equalization that has been extended from HE. AHE is contingent on HE, where the adaptive approach creates each sub-histogram image's in order to disperse the pictures' brightness values. AHE is therefore suited for enhancing local contrast and bringing forth additional details in a picture. Some AHE algorithms have made significant advancements in reducing noise and boosting contrast. The enhancing impact on hot items rather than background can be improved by the hybrid cumulative histogram equalization (HCHE).

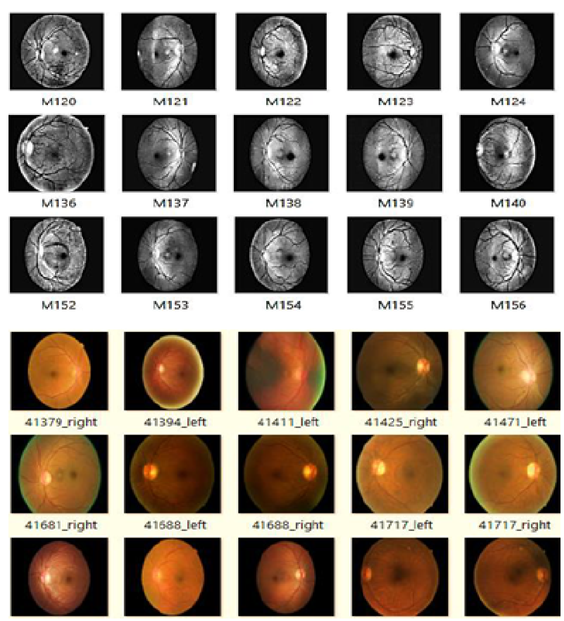

Fig. 2 CLAHE and Non-CLAHE images used for training

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

The over-enhancement issue may be resolved, and the feature loss issue in the dark areas of the image can be reduced, using the gap correction histogram equalization. However, because the global histogram equalization amplifies noise in more homogenous places, the issue still exists. To address this dilemma, contrast limited adaptive histogram equalization (CLAHE) was advised. The eminent block-based processing CLAHE can solve the problem of excessive noise amplification in the homogeneous zone of the picture by using conventional histogram equalization. The CLAHE method varies from normal HE in that it redistributes the brightness values of the picture by operating on tiny parts of the image, known as tiles, and computing several histograms, each corresponding to a different segment of the picture. The CLAHE enhancement method may be used in a variety of color spaces, including RGB, YIQ, HSI, and others. According to the RGB color model, a color space is described in terms of its RGB components. These three elements are pictures of monochromatic intensity. Therefore, whether pictures are recorded by a colour video camera or presented on a color monitor screen, the Red-Green-Blue model is the best tool for color generations. CLAHE may be used on each of the three components of the RGB color model separately. The R, G, and B separate components can be combined to create a full-color RGB picture. The RGB color space is ideal for displaying color images, but because of the strong association between these three elements, analysis and processing of images cannot be done in this space. To increase our dataset, data augmentation is required. The amount of data we provide the neural network during training will determine how accurate the model is. Data augmentation techniques are available in the Python library. For flipping the photographs horizontally and vertically, the "ImageDataGenerator" package, which is the most helpful, is utilized. The offered rotation is 45 degrees, with a zoom range of 0.2. Every image has noise in it. The noises in the image can be reduced using noise reduction techniques. During training and testing, rescaling is done before sending the pictures to the network.

```
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
```

Fig. 3 Proposed EMCNN model architecture

4. RESULTS AND DISCUSSION

Combining CLAHE and non-CLAHE images can provide complementary information that can improve the accuracy of feature extraction. Advanced image processing techniques, such as deep learning models, can be trained using both CLAHE and non-CLAHE images to extract relevant features from the images. By using both CLAHE and non-CLAHE, images in training the models, the ensemble can capture a wider range of features and patterns that may be indicative of DR. By generating synthetic images that combine CLAHE and non-CLAHE images, the model can learn to recognize and classify DR in a variety of image types and conditions. 1860 fundus CLAHE, Non-CLAHE pictures are used to evaluate the suggested model EMCNN; 468 of the 1860 images are classified as TP (True Positive) and 1369 as TN (True Negative). Figure 4 and 5 Confusion Matrix, a table with four alternative combinations of actual and anticipated values, serves as a visual representation of the findings. Other metrics along with accuracy, recall, and F1 score are displayed by the model. The proportion of true positives to false positives + true positives is known as accuracy. Precision = TP / (FP + TP) is one way to express it. Recall is the proportion of real positive samples to

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

all samples that are actually positive. The method is Recall=TP/(FN+TP). The weighted average of the accuracy and recall for the specified model is the F1 score. The method F1 = two*(precision*recall)/(precision + recall) may be used to get the F1 score. The accuracy., recall., and F1 score. values for moderate DR were obtained from the EMCNN. Precision., Recall., and F1 score. for No-DR are each 0.98, 1.00, and 0.99 respectively. The values obtained are displayed in figure 6.

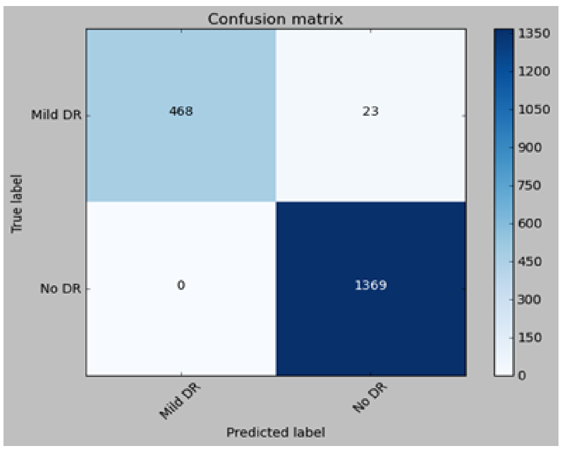

Fig.4. Confusion Matrix obtained from EMCNN Model

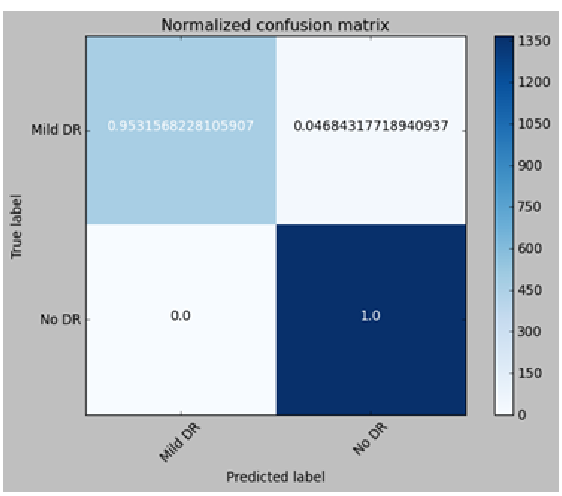

Fig.5. Normalized Confusion Matrix obtained from EMCNN Model

The Keras deep learning framework was used to create the model. The model is run locally on a PC with an Nvidia Geforce GTX 1650-Ti 4GB GPU. Application uses the Anaconda framework and Jupyter notebook. The dataset was pre-processed and enhanced before training to improve performance.

```
              precision    recall  f1-score   support

     Mild DR       1.00      0.95      0.98       491
       No DR       0.98      1.00      0.99      1369

    accuracy                           0.99      1860
   macro avg       0.99      0.98      0.98      1860
weighted avg       0.99      0.99      0.99      1860

[[ 468   23]
 [   0 1369]]
Confusion matrix, without normalization
[[ 468   23]
 [   0 1369]]
Normalized confusion matrix
[[0.95 0.05]
 [0.   1.  ]]
```

Fig.6. the Precision value, Recall value, and F1 score value obtained from EMCNN

The Receiver.Operating.Characteristics. (ROC) curve is another indicator used to evaluate the effectiveness of the model. In ROC curve presentations, the parameters.True Positive (TP) and False.Positive (FP) are employed (FP). Draws using TP and FP are conceivable. Figure 7 shows the graph, and the area under the curve for the Mild and No DR classes is 1.00. The suggested model performed better than other existing models.

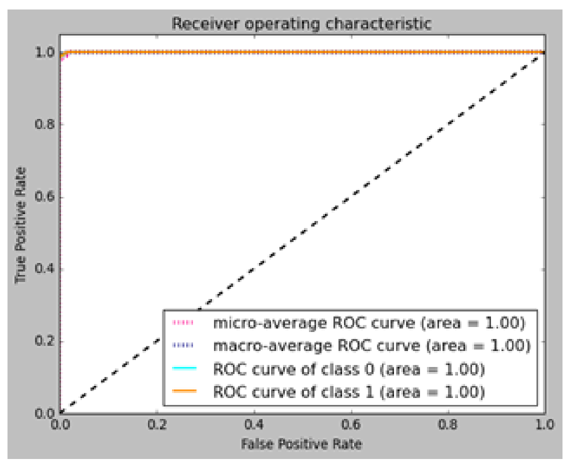

Fig.7. ROC Curve obtained from EMCNN

Figure 8 and 9 displays the train-validation accuracy and loss in 20 epochs. The training deficit is a measure of how suitably a deep learning model matches the training data. It estimates the model's glitch on the training set, in other words. The dataset used to first train the model is known as the training set. The training deficit is estimated based on the sum of fallacy for each sample in the training set. The training deficit is measured at the conclusion of each batch. To demonstrate this, a

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

training deficit curve is frequently used. On the other hand, a statistic known as validation deficit that is used to assess a deep learning model's fulfilment on the validation set. A portion of the dataset designated to examine the model's accuracy is called the validation set. The validation loss is computed by accumulating the fallacy for one by one sample in the validation set, much like with the training loss. The validation loss is also assessed at the climax of each session. This enables us to opt whether the model need additional tuning or adjustment. To achieve this, we frequently sketch a learning curve for the validation loss.

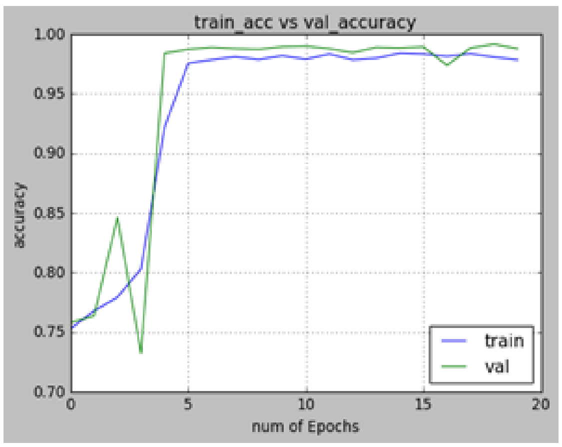

Fig.8. Train, Validation accuracy obtained from EMCNN

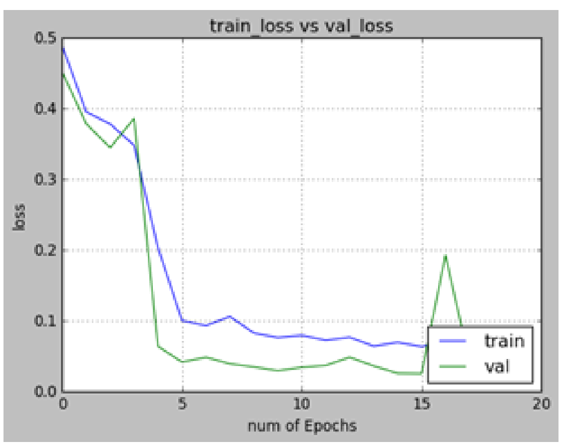

Fig.9. Train, Validation loss obtained from EMCNN

The suggested model is compared with different existing models and the result of comparison is shown in Table 2. The Model VGG-NIN used Kaggle dataset for the training and testing motive and 95 percent accuracy achieved. The model also shows 67% precision 55% recall and 59% F1 score. The model created through individual channel training, gave the accuracy of 81% from kaggle dataset and 70% precision 51% recall 56% F1 score also obtained. The model used Ensemble approach obtained 80%accuracy 63% precision 51% recall 53% F1 score using Kaggle dataset. The proposed model Experimental Minimal Convolution Neural Network (EMCNN) obtained the accuracy 98%; precision 98% recall 95% and F1 score 99% using a mix of Kaggle and Messidor data set.

Table 2. Model Comparison table

Model	Dataset Used	Accuracy	Precision	Recall	F1 Score
VGG-NIN [4]	Kaggle	0.95	0.67	0.55	0.59
Individual Channel Training [17]	Kaggle	0.81	0.70	0.51	0.56
Ensemble Approach [10]	Kaggle	0.80	0.63	0.51	0.53
Efficient Net [18]	Kaggle	0.83	-	-	0.82
EMCNN (Proposed)	Kaggle+Messidor	0.98	0.98	0.95	0.99

5. CONCLUSION AND FUTURE WORK

The use of convolutional neural networks (CNN) for image classification has shown promising results in recent years. In this paper, the researchers propose a new approach that uses a minimum CNN model to classify diabetic retinopathy. Using datasets from Kaggle and Messidor, the researchers suggested an Experimental Minimal Convolution Neural Network model to forecast and classify moderate DR symptoms. The dataset has undergone CLAHE pre-processing, and the MCNN model is trained using a mixture of CLAHE and identical Non-CLAHE pictures. The accuracy obtained is 98% and F1 score is 99%. The EMCNN model performs better when we compare with other architectures and the comparison is given in Table 1. The experimental results presented in the paper suggest that the proposed approach can achieve high accuracy in classifying diabetic retinopathy. The authors compare the performance of their approach to other state-of-the-art methods, and the results show that their approach outperforms these methods. The experimental results demonstrate the effectiveness of the approach and suggest that it could be used in clinical settings to aid in the early detection and diagnosis of diabetic retinopathy.

Researchers are interested in testing the current architecture with SVM linear combination design in upcoming projects. Combining CLAHE and non-CLAHE fundus images can provide a more comprehensive view of the retina, allowing for improved feature extraction and more accurate DR detection. By leveraging advanced image processing techniques, ensemble models, data augmentation, and transfer learning, we can continue to improve the accuracy and efficiency of DR detection, leading to earlier diagnosis and treatment of this serious eye disease.

REFERENCES:

[1] P. M. Ebin and P. Ranjana, "An Approach Using Transfer Learning to Disclose Diabetic Retinopathy in Early Stage," 2020 International Conference on Futuristic Technologies in Control Systems & Renewable Energy (ICFCR), pp. 1-4, doi:10.1109/ICFCR50903.2020.9249988,2020.

[2] Li, X., Hu, X., Yu, L., Zhu, L., Fu, C. -W., & Heng, P. -A. (2020). CANet: Cross-disease Attention Network for Joint Diabetic Retinopathy and Diabetic Macular Edema Grading. IEEE Transactions on Medical Imaging, 39(5), 1483-1493. doi:10.1109/tmi.2019.2951844

[3] He, T. Li, N. Li, K. Wang and H. Fu, "CABNet: Category Attention Block for Imbalanced Diabetic Retinopathy Grading," in IEEE Transactions on Medical Imaging, vol. 40, no. 1, pp. 143-153, Jan. 2021, doi: 10.1109/TMI.2020.3023463.

[4] Z. Khan et al., "Diabetic Retinopathy Detection Using VGG-NIN a Deep Learning Architecture," in IEEE Access, vol. 9, pp. 61408-61416, 2021, doi: 10.1109/ACCESS.2021.3074422.

[5] F. Saeed, M. Hussain and H. A. Aboalsamh, "Automatic Diabetic Retinopathy Diagnosis Using Adaptive Fine-Tuned Convolutional Neural Network," in IEEE Access, vol. 9, pp. 41344-41359, 2021, doi: 10.1109/ACCESS.2021.3065273.

[6] C. -H. Hua et al., "Convolutional Network With Twofold Feature Augmentation for Diabetic

Retinopathy Recognition From Multi-Modal Images," in IEEE Journal of Biomedical and Health Informatics, vol. 25, no. 7, pp. 2686-2697, July 2021, doi: 10.1109/JBHI.2020.3041848.

[7] T. Araújo et al., "Data Augmentation for Improving Proliferative Diabetic Retinopathy Detection in Eye Fundus Images," in IEEE Access, vol. 8, pp. 182462-182474, 2020, doi: 10.1109/ACCESS.2020.3028960.

[8] H. Kaushik, D. Singh, M. Kaur, H. Alshazly, A. Zaguia and H. Hamam, "Diabetic Retinopathy Diagnosis From Fundus Images Using Stacked Generalization of Deep Models," in IEEE Access, vol. 9, pp. 108276-108292, 2021, doi: 10.1109/ACCESS.2021.3101142.

[9] L. Qiao, Y. Zhu and H. Zhou, "Diabetic Retinopathy Detection Using Prognosis of Microaneurysm and Early Diagnosis System for Non-Proliferative Diabetic Retinopathy Based on Deep Learning Algorithms," in IEEE Access, vol. 8, pp. 104292-104302, 2020, doi: 10.1109/ACCESS.2020.2993937.

[10] 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.

[11] K. Shankar, Y. Zhang, Y. Liu, L. Wu and C. -H. Chen, "Hyperparameter Tuning Deep Learning for Diabetic Retinopathy Fundus Image Classification," in IEEE Access, vol. 8, pp. 118164-118173, 2020, doi: 10.1109/ACCESS.2020.3005152.

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] 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.

[18] 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.