International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Enhanced Diabetic Retinopathy Detection through Deep Learning Ensemble Models for Early Diagnosis

Dr. Deepak Dembla[1], Amita Meshram[2], Dr. Anooja A.[3]

Submitted: 12/12/2023 **Revised**: 24/01/2024 **Accepted**: 03/02/2024

Abstract: One of the most prevalent and primary causes of blindness associated with diabetes is diabetic retinopathy (DR). An early diagnosis of DR can stop the disease's progression. Vision impairment results from missed opportunities for diagnosis and treatment due to differences in the distribution of medical facilities. It is more effective and less expensive to classify and diagnose DR with better accuracy using neural network models. In order to enhance the functionality of DR classification models, This proposed model is an ensemble model that includes Transfer Learning algorithms that are Mobile Net, Xception, ResNet50V2, DenseNet201, DenseNet169, InceptionV3, and InceptionResNetV2 were added in this study for the detection of DR classification in fundus photographs. For that the APTOS and eyepacs dataset of a total of 5236 images are taken.This model taken 90% of images for training and 10% of images taken for testing on 30 epochs. The performance of DR classification can also be enhanced by adding extra 4 dense layers in each model and then ensemble that models. The output of the ensemble is applied for majority voting. The proposed ensemble model detected DR stages accurately and correctly. Hence proposed Ensemble model accuracy is 92.23%.

***Keywords**: Diabetic Retinopathy(DR); MobileNet; XceptionNet; ResNet50V2;DenseNet201; DenseNet169; InceptionV3.*

1. Introduction

Diabetes can induce (DR), a dangerous illness that causes retinal injury. There are five levels of DR severity: no DR, mild, moderate, severe, and proliferative DR. Individuals who have had diabetes for a longer period of time are more likely to acquire DR.

A competent ophthalmologist must inspect the fundus (the back of the eye) to diagnose DR. The condition increases during the proliferative DR (PDR) stage when aberrant new blood vessels grow in the retina. Early identification of diabetes is critical for identifying aberrant symptoms and initiating timely treatment, allowing diabetics to avoid problems connected with the illness. However, the diagnosis of diabetic retinopathy requires ophthalmologists to have professional clinical knowledge, expertise, and time. To determine the degree of the patient's DR, ophthalmologists often must undertake evidence gathered of the patient's retina and integrate the fundus retinal pictures captured by special tools. It will take a long time to do this.

Routine fundus examination screening is the most efficient way to identify early signs of DR in diabetic individuals. By regularly monitoring the fundus, healthcare professionals can identify any abnormalities and intervene early to prevent the progression of DR. Early diagnosis and timely treatment play a vital role in preserving vision and refining the overall quality of life for diabetic patients. Furthermore, the number of ophthalmologists in practice does not even come close to addressing the number of patients with diagnoses. As a result, the automated classification approach of diabetic retinopathy severity greatly improves the efficiency of DR diagnosis. This proposed model may detect the severity of DR illnesses more accurately by classifying them using seven transfer learning algorithms, and the output of those seven models is then ensembled and used to majority voting. The proposed DR classification model is more accurate and precise.

2. Related Work

The authors proposed using the CIE-Luv colour space, a Wiener filter to eliminate noise, a Caanny algorithm to generate edges and clustering to acquire the exudates aspirants, and normalization at each step to ensure that the data sharing aids in revealing the exudates candidates people in the retinal matter. Secondly, in terms of rating the DR severity level, XGBoost surpasses other ensemble approaches. The study's findings are as follows: First, CIE-Luv, CLAHE in many stages, the Wiener filter, the Canny algorithm, picture normalization at each level, and clustering approaches all contribute to the production of visually discernible exudate candidates. Using MESSIDOR and IDRiD datasets, XG-Boost then properly evaluated the DR severity level in multi-class scenarios. Second, their approaches outperform those of other researchers using the MESSIDOR dataset.[1]

1,2,3Department of Computer Application, JECRC: Jaipur Engineering College and Research Centre University, Jaipur, 303905, Rajasthan, India
2Department of Computer Science and Engineering, Yashwantrao Chavhan College of Engineering,Wanadongri, Nagpur, Maharashtra, India.

This research describes a novel cycle adaptive multi-target weighting network (CAMWNet) meant to mimic the natural retinal system of the human brain. In the context of diabetic retinopathy (DR) lesion classification, it overcomes the difficulty of class imbalance and improves segmentation performance. To address this, the authors create a flexible multi-target weighting approach designed specifically for DR lesion categorization. They proposed a unique joint loss function as a network augmentation approach based on the CAMWNet architecture. In terms of performance, the experimental findings confirm that the suggested technique outperforms current state-of-the-art technology.[2]

They employ a major part examination-based Deep neural network model with Grey Wolf Optimisation (GWO) computation to group the separated characteristics of the diabetic retinopathy dataset. GWO allows for the selection of appropriate boundaries for constructing the DNN model. They comprise conventional scaler normalization of the diabetic retinopathy dataset, dimensionality reduction with PCA, GWO selection of appropriate hyper borders, and lastly preprocessing of the dataset using a DNN model[3].

To extract retinal picture properties such as real vascular segments, the Faster RCNN was applied. The retinal visual data was pre-processed using fundamental image processing methods. A linked database comprising picture data from multiple publically available databases was utilized to train, test, and evaluate this suggested technique. With 92.81 percent sensitivity and 63.34 positive prediction values, our suggested approach recovered real vascular segments from the top initial layer of color retinal pictures [4].

They recommended for the use of deep learning (DL) and (CNNs) in the diagnosis of retinal diseases in order to detect, identify, and quantify aberrant traits. This method's effectiveness is increasing. This chapter provides an overview of the methods used to detect retinal abnormalities in retinal images associated with the most serious eye issues, such as network designs, post/pre-processing, and evaluation tests[5].

The need for extensive training to categorise images of serious incidents is greatly minimized. Their suggested segmentation approach, which uses the VGG-19 architecture, exhibits a substantial capability in spotting anomalies across 5000 validation pictures, with an excellent accuracy rate of 96%. Furthermore, the VGG-19 design successfully differentiates between different degrees of DR by analyzing the contour area to real contour arc length ratio. In upcoming investigations, the transfer learning principle will be used to detect neovascularization anomalies. [6]

The goal of this study was to use the EyePacs Diabetic Retinopathy Database to automatically quantify the degree of retinal impairment using three pre-trained deep neural networks. There are two types of neural categorization models. The first model divides retinal pictures into two categories: healthy and damaged. The second model categorises injured retinas, determining the extent of retinal damage. They randomly selected images from the dataset and separated them into training and testing data in a 4:1 ratio. The Resnet-50 network earned the greatest results for the first classification model's average classification accuracy, with an average classification accuracy of 92.64 percent on test data. Inceptionv3 achieved the best average classification in the second model[7].

VGG19 architecture-trained networks outperformed deeper architecture-trained networks. For the majority soft voting and stacking procedures, by implementing this network they found accuracies of 0.92 and 0.90. The best single data-type network and the hand-crafted feature network were both surpassed by both ensemble procedures. Grad-CAM has been demonstrated to identify diseased regions more precisely [8].

Using VGG-19 network segmentation based on deep learning, With 92 percent accuracy, 86.5 This method can differentiate the OC and OD with 93.5 percent sensitivity and 93.5 percent specificity.The suggested method, which has a high accuracy rate, may also be used to explore the most punctual phase of glaucoma. The performance of the classifier has also been evaluated.When compared to the VGG 19 RF classifier, the VGG 19 SVM classifier improves classification accuracy by 2%[9].

The goal of your study to processed the picture is encoded and then decoded for quality assessment during the segmentation step. Furthermore, the approach employs multiple instance learning (MIL) to extract features and categorize data. However, it should be noted that the technique needed additional development, which was seen as a serious disadvantage. To detect diabetic retinopathy (DR), the approach was tested using the CHASE dataset. Notable results were 96.62% accuracy, 95.31% sensitivity, 97.3% specificity precision, and a computational time of 0.99 seconds. These results outperformed other approaches, such as MLBNVD, GAB-DR, and DLA-DR, lowering the chance of vision loss. [10].

To properly realize the ramifications, the number of topographies must be reduced while simultaneously leveraging the organization's presence. As a consequence, they describe an RBV segmentation approach that uses KGMO-FCM to fine-tune and optimise the features. This procedure is used to discover pathology and allow the doctor to begin treating the patient. The ideal characteristics then vote out the organization's training, and the process is completed by organizing the pictures using a hybrid strategy such as CNN-BiLSTM with classifier, that SA for the powerful connection of pictures, as the fine obsessive retina [11].

In this study, the VGG-16 and ResNet-50 networks were compared in the classification of DR. Using VGG-16 and ResNet50, the analysed models exhibited accuracy of 25% and 70%, respectively. ResNet50 got the best results, indicating the architecture's effectiveness in image categorization, especially for systems with a high number of images. [12].

Computerized Diabetic Retinopathy Screening Devices are inexpensive; they shorten detection time, enabling eye specialists to detect retinal defects and providing early treatment to prevent potential complications.These DR detection technologies are essential for more accurate disease diagnosis. This research begins with a discussion of Diabetic Retinopathy, its many types, and symptoms[13].

They successfully built a Convolutional Neural Networks model that diagnoses diabetic retinopathy and other retinal disorders using a pre-trained VGG-16 framework. It also offers information on the condition's severity. They were successful in obtaining a model accuracy of 74.58 percent. This approach can be helpful in assisting doctors for a variety of reasons. This condition can be identified sooner[14].

They worked for two years on the construction of two independent Deep Learning frameworks targeted at forecasting the progression of diabetic retinopathy. These frameworks were carefully evaluated using two independent datasets: an internal validation set consisting of photos gathered from primarily Hispanic patients in the United States, and an external validation set obtained from Thailand. The Deep Learning methodology performed wonderfully on both datasets, both alone and when tailored to readily available hazard indicators. When used in conjunction with available danger factors, the visualisation enhanced over utilising the risk factors alone[15].

They proposed a deep learning-based feature extraction technique for early-stage glaucoma diagnosis. They train and evaluate their proposed model using retinal fundus images. The initial step is to preprocess these pictures, followed by segmentation to determine the area of interest (ROI). Then, to extract distinguishing attributes of the optic disc (OD) and optic cup (OC) present in the images, a combined feature descriptor that incorporates convolutional neural network (CNN), local binary patterns (LBP), histogram of oriented gradients (HOG), and accelerated robust features (SURF) is used. The LBP and SURF descriptors are used to extract texture structures, whereas HOG is used to extract low-level features. CNN can also compute high-level properties.[16].

Their research created a graded fusion-based DR identification method that can analyse CLAHE and CECED fundus pictures at the same time. These channels, as previously stated, are coupled to extract significant information from fundus pictures and reach better degrees of identification accuracy. To fully utilise the visual features acquired from the multiple channels, the weighted fusion approach was applied. The suggested WFDLN model tackles the issue of low-quality fundus images by merging the weighted characteristics supplied by the CLAHE and CECED pre-processing phases [17].

They acquired 3662 retinal images from Kaggle depicting various stages of diabetic retinopathy (DR). They used oversampling and balancing techniques to fix class imbalance, resulting in 5711 images. The collection of pictures was split into 4659 training photos and 10% validation images.They employed five deep learning models to categorise DR at various stages: CNN, EfficientNet, VGG16, MobileNet, and ResNet50. They intended to improve the accuracy of these models in identifying sickness by adjusting various parameters. They observed that MobileNet outperformed the other four of the five models. MobileNet's validation accuracy was 85.48% before oversampling and balancing, training accuracy was 93.42%, and testing accuracy was 81.81%[32].

Authors evaluates the performance of five deep learning models—CNN, Efficient Net, VGG 16, Mobile Net, and RESNET 50. Through systematic parameter adjustments for these models, the study aims to enhance accuracy in classifying DR at various stages. Among the total images, the findings highlight that Efficient Net achieved a training accuracy of 0.9342 and a testing accuracy of 0.8181, while RESNET 50 achieved accuracies of 0.9329 for the training dataset and 0.8116 for the testing dataset. Notably, Efficient Net and Res Net 50 demonstrated superior accuracy compared to the other three models, establishing them as effective performers in DR classification.[33]

In their experiments, the CNN model exhibited a classification accuracy of 71.34% for the training dataset and 71.41% for the testing dataset after 20 epochs. In comparison, the Efficient Net model, after the same number of epochs, demonstrated significantly improved accuracy, reaching 93.42% for the training dataset and 81.81% for the testing dataset. Despite the progress with Efficient Net, it's acknowledged that the overall accuracy is still constrained by the limited dataset.[34]

3. Methodology

3.1. Data Source

This study's data came from an open dataset of APTOS 2019 Diabetic Retinopathy Detection Blindness

Detection[18]. The datasets comprise 3662 retinal pictures from various situations. In addition, 1574 photos of the DR dataset (eyepacks) are captured. This model preprocessed this dataset 1024*1024 into 224*224 picture size and combined it with 3662 to generate 5236 retinal images, which are classified as shown in Table 1. A total of 5236 photos were utilized for training, with 4712 images used for testing.

Table 1: Images are divided into a category

Image Category	Count
Mild	1034
Moderate	1185
No_DR	1867
Proliferate	605
Severe	545
Total Images	5236

Because retinal pictures are collected from various places and with different equipment, there is a lot of noise in the image, which is unwanted, and something needs to be eliminated, numerous preprocessing procedures are necessary. For diabetics, the retinopathy associated with each image was assessed on a scale of 0 to 4, as follows:

0- No DR

1- Mild

2- Moderate

3- Proliferative

4- Severe

No DR denotes the absence of any DR-related symptoms in the retinal picture. Level 1 for Mild DR implies the existence of a range of features in the retinal image, such as hemorrhages, hard exudates, and macular edema. The amount of extensive hemorrhages is indicated by Level 2 for Moderate DR. Proliferative DR is defined by the presence of hemorrhages in all four quadrants, whereas Severe DR is defined by Pre Retinal Haemorrhages[30].

3.2. Image Classification Using a Deep Learning Model

This step must be done in order to facilitate the operation. The Retinal Images divides them into many groups. Feature Extraction is conducted after pre-processing. Several picture understanding alternatives are recovered at this step, including Optic Distance, vessel, scientific Entropy Exudates space, and a chemical appraisal of these possibilities. The Deep Learning algorithmic program is used to choose and categorise the picture alternatives. Alternative area devices come in a number of designs as a consequence, both the accuracy of each class and the class's total accuracy improve. In the proposed model, Mobile Net, Xception, ResNet50V2, DenseNet201, DenseNet169, InceptionV3, and InceptionResNetV2 were all employed.

Mobile Net

Figure 1 shows MobileNets rely on two important hyperparameters to achieve a balance of accuracy and efficiency. Breaking down convolution kernels is the core idea of MobileNet. The traditional convolution procedure is separated into depth-wise and point-wise convolution. Depth-wise separable convolution and a 1x1 convolution kernel are used to achieve this. In this method, depth-wise convolution layers' outputs. In contrast to the traditional convolutional filter, which merges inputs to generate new outputs, depth-wise separable convolution divides inputs into two layers—one for filtering and the other for merging.—a typical convolutional filter mixes the inputs into a new set of outputs. Extra thick layers have been added to this suggested model, as illustrated in section B2 to increase the model's accuracy. Section B2 receives the output (weights) from Section B1 and performs the operation. [25].

model's accuracy. Section B2 receives the output (weights) from Section B1 and performs the operation. [25].

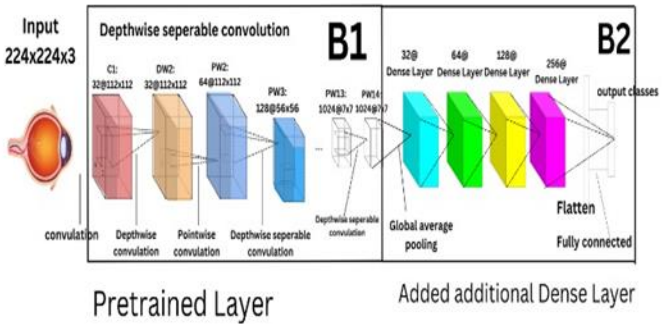

Fig. 1. Mobile Net

XceptionNet

Figure 2 shows the network, is composed of architecture comprises of 36 convolutional layers divided over 14 modules (excluding the fully linked layer at the end). Except for the start and terminal modules, each module is surrounded by linear residual linkages. This approach includes a succession of depth-separable convolutional layers that build spatial links by applying depth-wise convolutions to input channels individually. This xception model is shown in Section B1. [27]. Extra dense layers were added to this suggested model, as illustrated in section B2, to increase the model's accuracy. Section B2 receives the output (weights) of section B1 and performs the operation.

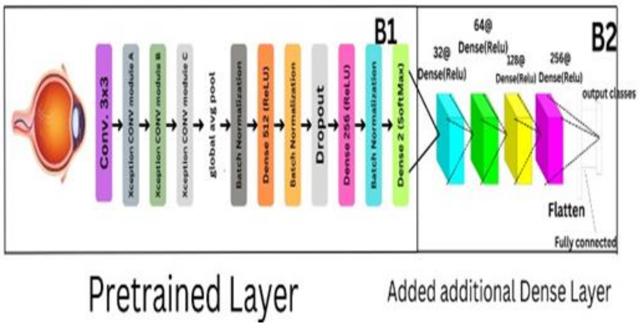

Fig 2.XceptionNet

ResNet50V2

Figure 3 depicts ResNet50V2's "50" representing the total number of network layers. It is made up of a deep stack of residually linked convolutional layers that allow for the training of deep networks without running into the disappearing gradient problem.

The usage of residual connections is the key innovation of the ResNet design. ResNet50V2 allowing the network to transmit gradients as it is trained. This helps to resolve the degradation problem that happens when the depth of deep networks begins to impair their performance. Because of the residual connections, the network may learn residual mappings, making it easier to train and optimize very deep architectures. Extra dense layers were added to this suggested model, as illustrated in section B2, to increase the model's accuracy. Section B2 receives the output (weights) of section B1 and performs the operation.

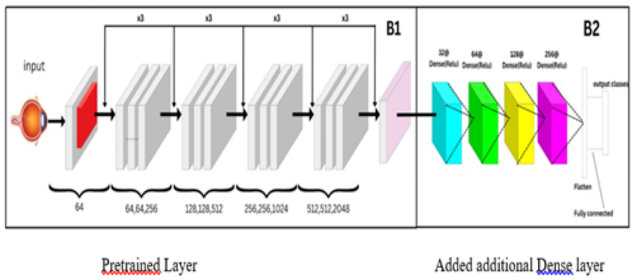

Fig.3. ResNet50V2[28]

DenseNet201

Figure 4 shows the total amount of network layers, which is denoted by the number 201 in DenseNet201. It is composed of many dense blocks, each of which has several convolutional layers, bottleneck layers, and batch normalisation. Convolutional and pooling methods are employed in transition layers that connect dense blocks to reduce spatial dimensions and the number of feature maps.

When compared to traditional designs, DenseNet201's use of dense connections enables exceptionally effective parameter sharing, resulting in a reduction in the number of parameters. Because of its parameter efficiency, DenseNet201 is computationally efficient, which also assists in more accurate deep network training. Extra dense layers were added to this suggested model, as illustrated in section B2, to increase the model's accuracy. Section B2 receives the output (weights) of section B1 and performs the operation.

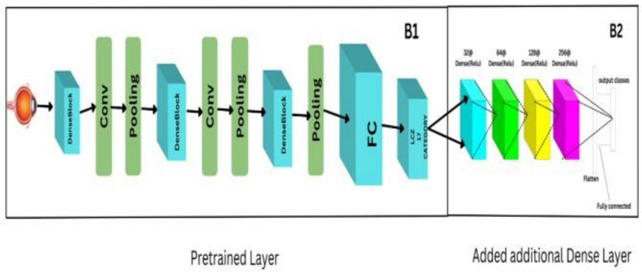

Fig 4. DenseNet201[29]

DenseNet169

Figure 5 shows Convolutional layers, max pool layers, thick layers (completely connected layers), and transition layers are all part of the architecture. The top layer is activated with SoftMax, while the rest of the pattern is activated with ReLU. Convolutional layers extract picture characteristics, whereas maxpool layers reduce input dimensionality. Following the flatten layer, the fully connected layers [30] function as an artificial neural network with a single array input from the flatten layer. The architecture is depicted in Figure 5, which contains convolutional layers, max pool layers, thick layers (completely connected layers), and transition layers. The top layer is activated with SoftMax, while the rest of the pattern is activated with ReLU. Extra dense layers were added to this proposed model, as illustrated in section B2, to increase the model's accuracy. Section B2 receives the output (weights) of section B1 and performs the operation.

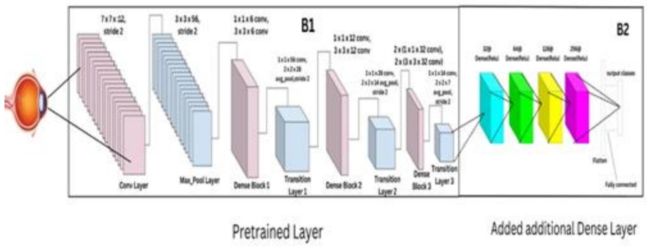

Fig5. DenseNet169[30]

InceptionV3

Figure 6 shows this. The major goal of the Inception v3 architecture was to improve the efficacy and precision of image recognition models. It achieved this through a number of important advancements, notably the extensive use of 1x1 convolutional layers, which reduces the network's computational cost. Inception v3 also employs other techniques, such as batch normalisation, which speeds up training and helps to reduce overfitting, and the inclusion of auxiliary classifiers, which are employed during training to facilitate gradient propagation across the network. For image interpretation tasks, the Inception v3 model has proven to be successful and versatile. Its strong performance on benchmark datasets has led to widespread use in both academic and commercial applications Extra dense layers were added to this suggested model, as illustrated in section B2, to increase the model's accuracy. Section B2 receives the output (weights) from Section B1 and performs the operation[31].

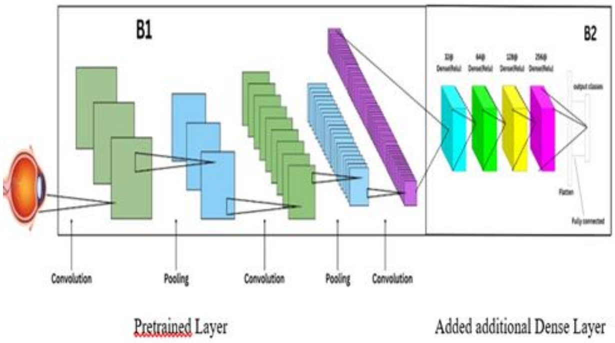

Fig.6. InceptionV3[32]

InceptionResNetV2

Figure 7 shows the "Inception" design.ResNetV2 extends the Inception v3 model with residual connections obtained from the ResNet architecture. Because of residual connections, the network may learn residual mappings, which improves gradient flow during training and addresses the vanishing gradient problem. This allows for more accurate training of very deep neural networks. InceptionResNetV2 employs the concept of residual blocks, which comprise both Inception modules and residual connections. The Inception modules extract multi-scale properties from images using varying filter sizes, and the residual connections allow gradients to propagate and help in deeper learning. Extra thick layers were added to this suggested model, as illustrated in section B2, to increase the model's accuracy. Section B2 receives the output (weights) of Section B1. Extra dense layers were added to this suggested model, as illustrated in section B2, to increase the model's accuracy. Section B2 receives the output (weights) of Section B1 and performs the operation.

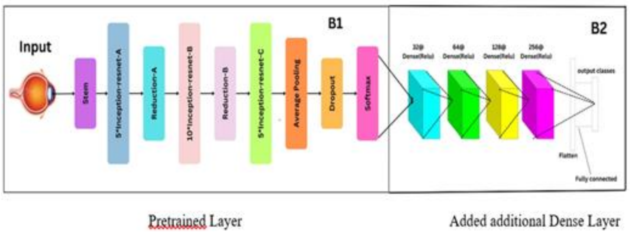

Fig7. InceptionResNetV2[32]

3.3. Proposed Deep Learning Ensembled Model For Image Classification

Retinal pictures are applied to pre-trained seven models, including MobileNet, Xception, ResnetV2, DenseNet201, DenseNet169, Inception, and inceptionResnetV2, and the output of these models is ensembled and majority voting is done to determine accurate and correct different classifications, as shown in figure 8.

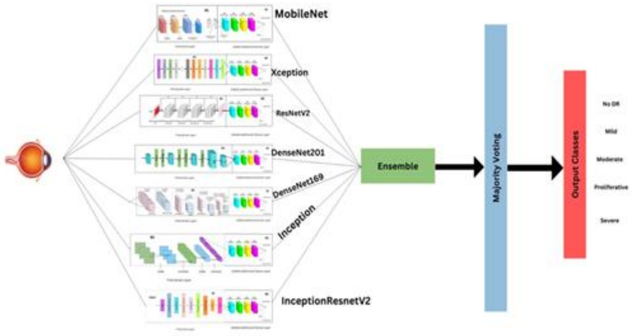

Fig. 8. Proposed Ensemble Model

As shown in Figure 9, a series of pretrained retinal pictures is fed into several DL models such as Mobile Net, Xception, ResNet50V2, DenseNet201, DenseNet169, InceptionV3, and InceptionResNetV2.To improve accuracy, the layers of this suggested model were fine-tuned. After that, categorise these models and ensemble their output using majority voting to attain accuracy. The output of these is then ensembled, and majority voting is used to obtain the final result by categorization. The proposed approach improved detection accuracy when tested on 10% retinal samples.

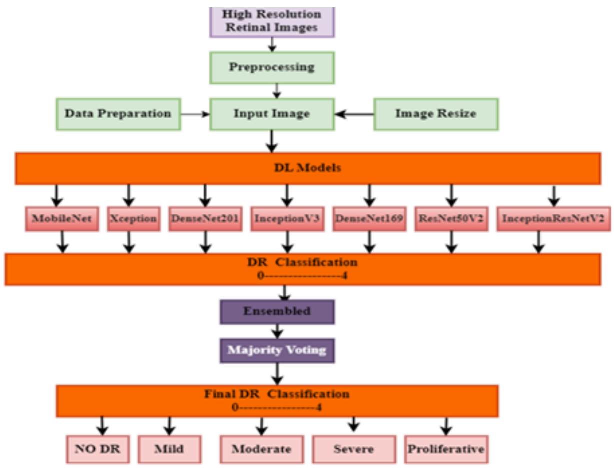

Fig.: 9. Block diagram of Proposed Ensemble Model.

Algorithm : Proposed Algorithm

To classify raw retinal images into different categories (No DR, Low, Medium, High, Proliferative), follow these steps:

1. Extract and read all retinal images.

2. Label the images into categories (0 - No DR, 1 - Low, 2 - Medium, 3 - High, 4 - Proliferative).
3. Split the retinal images into training and testing datasets.
4. Set the batch size to 125 for training.
5. Import the pre-trained models (MobileNet, Xception, ResNet50V2, DenseNet201, DenseNet169, InceptionV3, InceptionResNetV2).
6. For each epoch from 1 to 5, do the following: a. Divide the training dataset into small batches (xi, yi) => (X_test, Y_test, X_train, Y_train).
7. End the loop for epochs.
8. For each test image x in X_test, do the following: a. Pass x through each pre-trained model. b. Ensemble the outputs of all the models.
9. End the loop for test images.
10. Take majority voting on the ensemble outputs.
11. The final output is the predicted category for each retinal image.

The algorithm shows the proposed model in detail. All retinal images from the dataset are applied to these models (X,Y), where X the set of N images each of size 224 x 224, and Y contains equivalent labels. Divide the training set(Xtrain, Ytrain) into small batches, each size of 125 In the case of testing, the proposed model utilise layering to integrate the outcomes of all individual models and provide a uniform output to forecast the class label of the unseen sample. The ensemble method improves performance by combining the advantages of distinct models. Figure 9. shows the proposed ensemble model ensemble in action. It gives the final output.

Performance Evaluation Metrics

1. Confusion Metrix
2. F1Score
3. Model accuracy
4. Model loss
5. Precision
6. Recall

Table 2: Confusion Matrix

		ACTUAL	
PREDICTION		Negative	Positive
	Negative	True Negative	False Negative
	Positive	False Positive	True Positive

The size of the matrix in Confusion Matrix is 2X2 for binary classification with actual values on one axis and predicted values on the other. Within the confusion matrix, the following terms are being measured: True Positive (TP), True Negative (TN), False Negative (FN), and False Positive (FP).Percentage of actually Positive from Out of all the positive foretold.

Accuracy=(TP+TN)/(TP+TN+FP+FN)

Error Rate=(FP+FN)/(TP+TN+FP+FN)

Precision=TP/(TP+FP)

Recall= TP/(TP+FN)

4. Result and Discussion

To diagnose Diabetic Retinopathy in its early stages, the proposed model employed 5236 retinal pictures for training, testing, and validation. Again, the model utilized 10% of the total photos for testing. For DR classification, the suggested model included a total of seven transfer learning methods. To identify DR correctly, the proposed methodology constructed an ensemble model that included majority voting. MobileNet, Xception, ResNet50V2, DenseNet201, DenseNet169, InceptionV3, and InceptionResNetV2 algorithms were employed in the suggested model.

Mobile net training and validated accuracy is 97.31% and 8.78%, respectively, while precision is 87% and 83%, F1-Score training and validated is 84% and 81%, Recall training and validated is 84% and 82%, and Loss training and validated is 0.8%, according to table3. Mobile net training and validated accuracy for ResNet50V2 are 91% and 75%, respectively. For the similarly stated model, the accuracy for training and validation is 73% and 70%, F1-Score training and validation is 73% and 70.23%, recall training and validation is 74% and 72%, and loss training and validation is 0.22 and 1.2%. MobileNet's testing accuracy is 84%, whereas ResNet50V2's is 75%.

According to Table 3, the training and verified accuracy of Xception are 97.95% and 90.33%, respectively. While precision training and validated accuracy are 90% and 86%, respectively, F1-Score training and validated accuracy are 90% and 87%, recall training and verified accuracy are 90% and 85%, and loss training and validated accuracy are 0.05% and 0.53%, respectively. The training and confirmed accuracy rates for DenseNet201 are 95.80% and 85.23%, respectively. Similarly, training and validation accuracy is 85% and 80%, F1-Score training and validation accuracy is 84% and 79%, recall training and validation accuracy is 85% and 81.23%, and loss training and validation accuracy is 0.12% and 0.58%.Xception's testing accuracy is 86%, whereas DenseNet201's is 85%.

Table 3:MobileNet , ResNet50V2, Xception, DenseNet201

Metrics	**Mobile Net**		**ResNet50V2**		**Xception**		**DenseNet201**	
	Train	**Validate**	**Train**	**Validate**	**Train**	**Validate**	**Train**	**Validate**
Accuracy	97.31	87	91	73	97.95	90.33	95.8	85.23
Precision	87	83	73	70	90	86	85	80
F1-Score	84	81	73	70.23	90	87	84	79
Recall	84	82	74	72	90	85	85	81.23
Loss	0.8	0.8	0.22	1.2	0.05	0.53	0.12	0.58

Figures 10 and 11 show the MobileNet Accuracy and Confusion Matrix, which identifies DR images. The MobileNet model obtained 85% detection accuracy, and the confusion matrix specifies the detection accuracy of No DR as 97%, Mild as 91%, Moderate as 55%, Severe as 81%, and Proliferative as 91%. The blue line indicates training accuracy, whereas the orange line indicates testing accuracy.

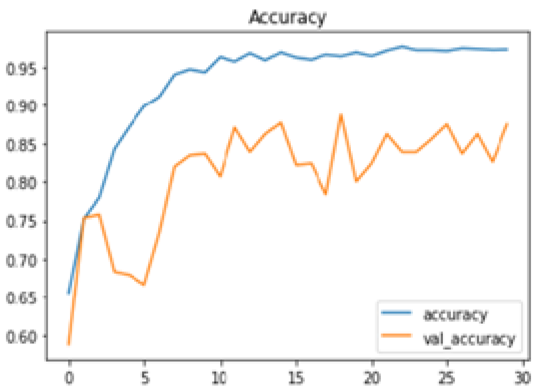

Fig.10. Accuracy of Mobile Net

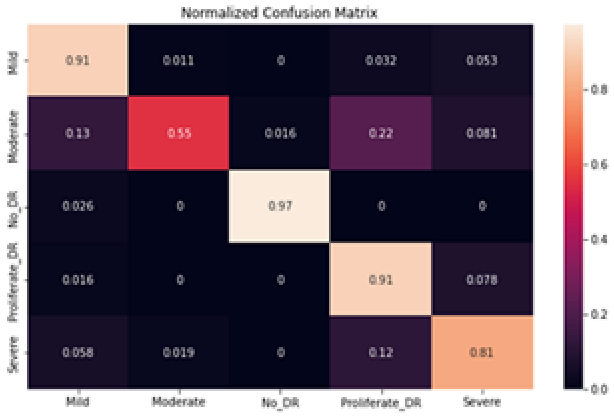

Fig.11. Confusion Matrix

figures 12 and 13 show the accuracy and confusion matrix of ResNet50V2, which classifies DR images. The ResNet50V2 model obtained 75% detection accuracy, while the confusion matrix specifies the detection accuracy of No DR as 98%, Mild as 85%, Moderate as 64%, Severe as 4%, and Proliferative as 36%.In the graph, The blue line indicates training accuracy, whereas the orange line indicates testing accuracy.

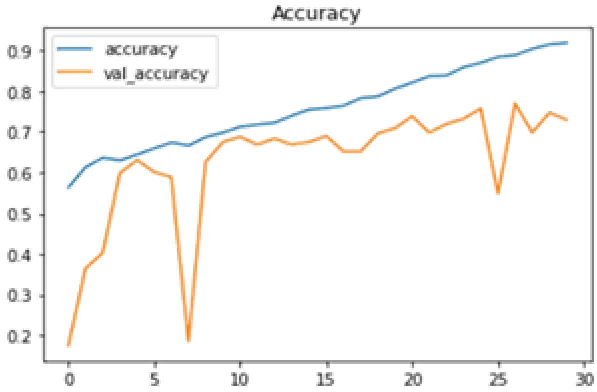

Fig. 12. Accuracy for ResNet50V2

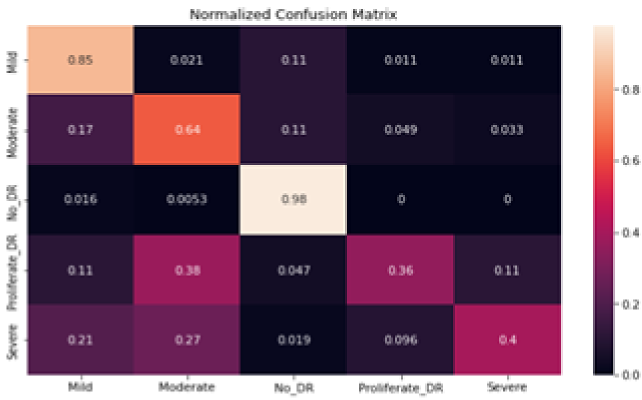

Fig. 13. Confusion Matrix

Figures 14 and 15 shows the Accuracy and Confusion Matrix of xception, which identifies DR pictures. The xception model obtained 90% detection accuracy, while the confusion matrix specifies the detection accuracy of No DR as 98%, Mild as 91%, Moderate as 85%, Severe as 83%, and Proliferative as 78%.In the graph The blue line indicates training accuracy, whereas the orange line indicates testing accuracy.

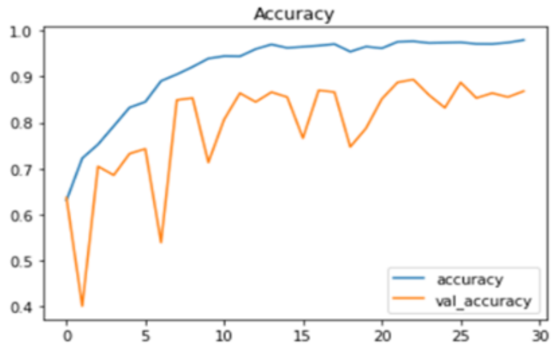

Fig. 14. Accuracy for Xception

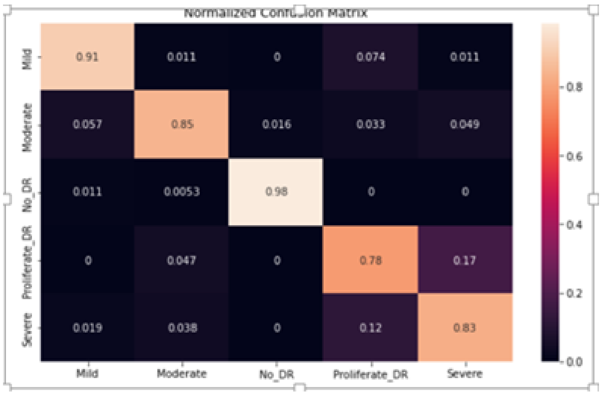

Fig. 15. Confusion Matrix of XceptionNet

Figures 16 and 17 shows the DenseNet201 Accuracy and Confusion Matrix, which is used to classify DR images. The DenseNet201 model achieved 85% detection accuracy, while the confusion matrix showed No DR detection accuracy of 97%, Mild detection accuracy of 89%, Moderate detection accuracy of 76%, Severe detection accuracy of 54%, and Proliferative detection accuracy of 81%.The blue line in the graph shows training accuracy, whereas the orange line indicate testing accuracy.

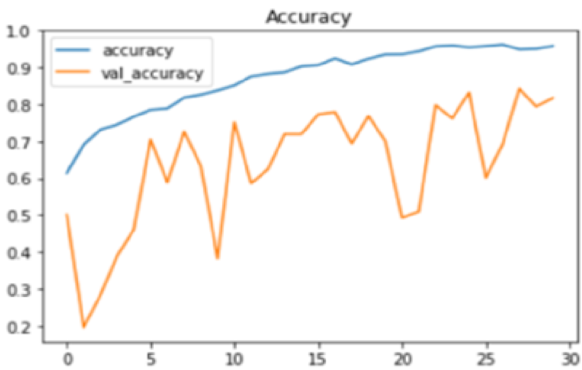

Fig. 16. DenseNet201

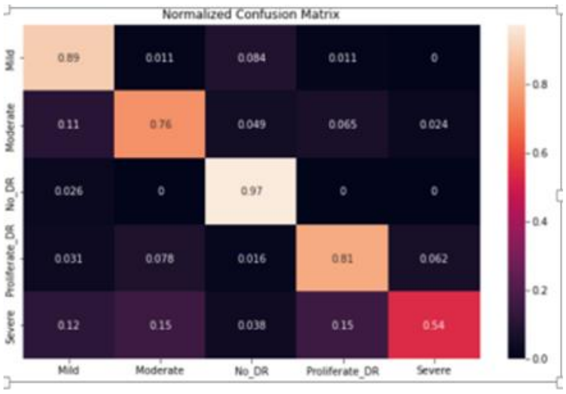

Fig. 17. Confusion Matrix

Figures 18 and 19 shows the Inception v3 accuracy and confusion matrix, which recognises DR images. The detection accuracy of the Inception v3 model was 87.23%, while the confusion matrix stipulates the detection accuracy of No DR as 98%, Mild as 94%, Moderate as 74%, Severe as 69%, and Proliferative as 75%.The blue line in the graph shows training accuracy, whereas the orange line indicate testing accuracy.

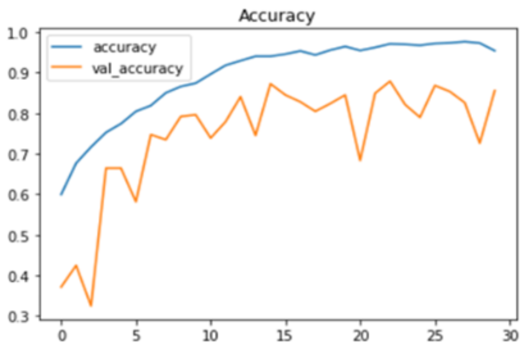

Fig. 18: Inception V3

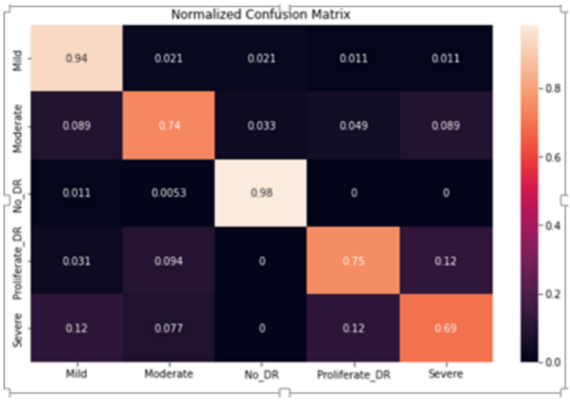

Fig.:19 Confusion Matrix

Figures 20 and 21 shows the accuracy and confusion matrix of DenseNet169, which classifies DR pictures.

DenseNet169model obtained 87.23% detection accuracy, and the confusion matrix classifies detection accuracy of No DR as 97%, Mild as 92%, Moderate as 72%, Severe as 65%, and Proliferative as 91%. In the graph, the blue line indicates training accuracy while the orange line indicates testing accuracy in the graph.

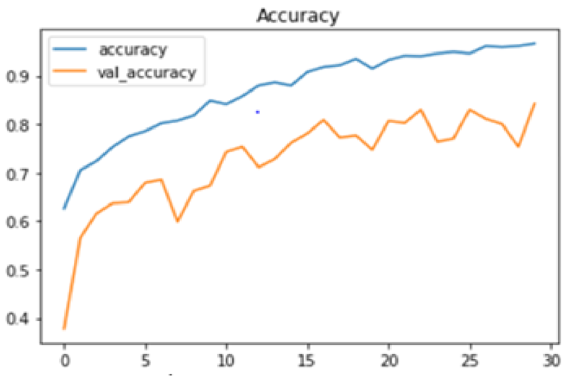

Fig 20: DenseNet 169

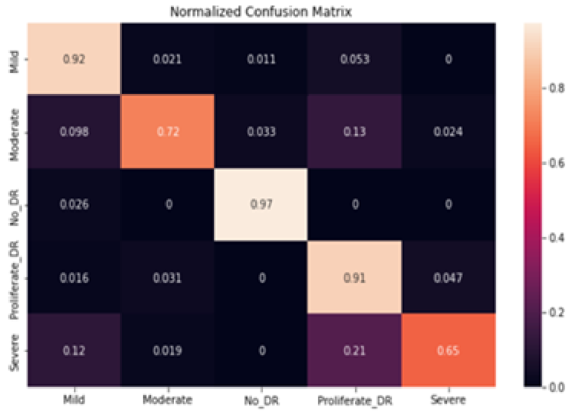

Fig 21. DenseNet 169 Confusion Matrix

Figures 22 and 23 shows the Accuracy and Confusion Matrix of InceptionResNetV2, which identifies DR pictures. The InceptionResNetV2 model obtained 90% detection accuracy, while the confusion matrix indicates the detection accuracy of No DR as 98%, Mild as 96%, Moderate as 76%, Severe as 83%, and Proliferative as 81%.In the graph, The blue line indicates training accuracy, whereas the orange line indicates testing accuracy.

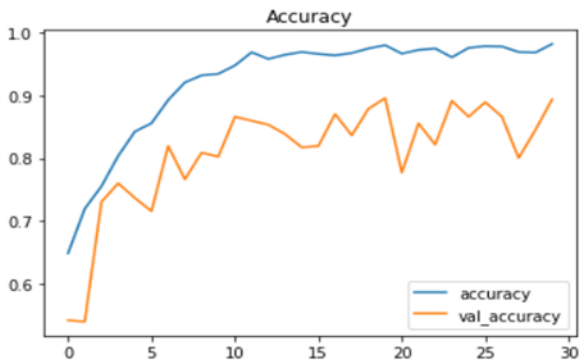

Fig 22. : InceptionResnetV2

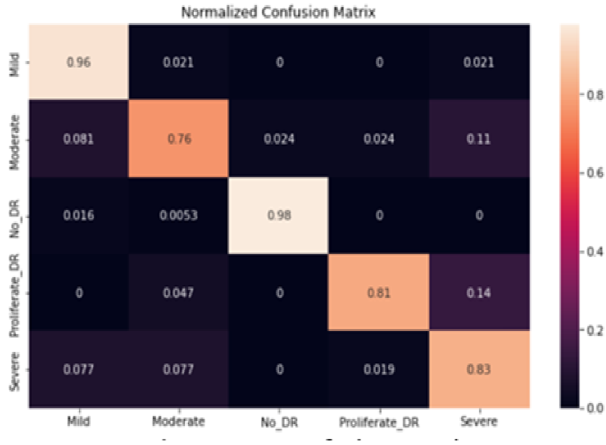

Fig 23 : InceptionResnetV2 Confusion Matrix

According to Table 4 InceptionV3 training and validated accuracy is 95.43% and 85.23%, respectively, while precision training and validated accuracy is 86% and 82%, F1-Score training and validated is 86% and 81%, Recall training and validated is 86% and 80%, and Loss training and validated is 0.13% and 0.55%.

Accuracy is 96.65% and 81.23% for DenseNet169 training and validation, precision is 88% and 80%, F1-Score training and validation is 86% and 81%, recall training and validation is 98.23% and 89.38%, and loss training and validation is 0.11% and 0.51%. For InceptionResNetV2 training and validation, accuracy is 96.65% and 81.23%, precision is 90% and 85%, and F1-Score training and validation is 89% and 84%. Recall training and validation are 90% and 81%, respectively, whereas loss training and validation are 0.046% and 48%.

InceptionV3 testing accuracy is 76.57%, DenseNet169 testing accuracy is 84.54%, and InceptionResNetV2 testing accuracy is 89%.

Table 4: Inceptionv3, InceptionResNetV2 and DenseNet169

Metrics	InceptionV3		DenseNet169		InceptionResNetV2	
	Train	Validate	Train	Validate	Train	Validate
Accuracy	95.43	85.23	96.65	81.23	98.23	89.38
Precision	86	82	88	80	90	85
F1-Score	86	81	86	81	89	84
Recall	86	80	86	80	90	81
Loss	0.13	0.55	0.11	0.51	0.046	0.48

Table 5 shows the accuracy of many models. Model VGG 16 has a testing accuracy of 73.04%[19], ResNet-18 has a testing accuracy of 67.14%[19], and DenseNet-121 has a testing accuracy of 72.95%[19]. The testing accuracy attained by Densenet 169 is 78.12%[20]. InceptionV3 has a testing accuracy of 77%[20], MobileNet has a testing accuracy of 77.7%[20], InceptionResNet has a testing accuracy of 75.68%[20], and ResNet50 has a testing accuracy of 67.78%. [20].All existing models achieved these accuracies individually.

Table 5: Comparison between existing individual models with proposed deep learning models

Models	Accuracy	
	Train	Test
VGG16 ,2022[19]	78.07%	73.04%
ResNet-18, 2022 [19]	78.44%	67.14%
DenseNet-121, 2022 [19]	91.11%	72.95%
DenseNet 169, 2022 [19]	70.73%	78.12%
InceptionV3, 2022 [19]	70.30%	77%
MobileNet, 2022 [20]	71.95%	77.70%
InceptionResNet V2, 2022 [20]	70.26%	75.68%
ResNet50, 2022 [20]	63.15%	67.78%
Proposed MobileNet Model	97.31	84.78
Proposed ResNet50V2 Model	91%	75%
Proposed Xception Model	97.95%	90.33%
Proposed Inception v3 Model	95.43%	84.23%
Proposed DenseNet201 Model	95.8	85.23
Proposed InceptionResNetV2	98.23%	90%
Proposed DenseNet169Model	96.65	87.23

Table 6 shows Hybrid model achieved testing accuracy is 90.60% [21]. EDLDR achieved testing accuracy is 86.08% [22]. Hybrid model achieved testing accuracy is 86.34%[23]. Hybrid model achieved testing accuracy is 89.29%[24].The proposed Ensemble model achieved accuracy more than the previous model is 92.23%.

Table 6: Comparison between an existing model with Proposed model

Models	Accuracy	
	Train	Test
Hybrid model ,2022[21]	92.50%	90.60%
EDLDR,2023[22]	NA	86.08%
Hybrid model ,2020[23]	NA	86.34%
Hybrid model,2022[24]	NA	89.29%
Proposed Ensemble Model	NA	**92.23%**

5. Conclusion

Diabetic retinopathy (DR) is a significant diabetic condition that damages the blood vessels in the retinal tissue of the eye. Traditional techniques of DR diagnosis have been tedious, costly, and time-consuming, with low accuracy. To address these issues, a suggested model for early and reliable detection of DR is developed seven deep learning models: MobileNet, Xception, ResNet50V2, DenseNet201, DenseNet169, InceptionV3, and InceptionResNetV2.

Each of the seven deep learning models is extended by an extra dense layer of 32, 64, 128, and 256 neurons in the proposed model. The outputs of these seven models are then ensembled, and DR is detected via majority voting MobileNet obtained an accuracy of 84%, XceptionNet obtained an accuracy of 90.33%, ResNet50V2 obtained an accuracy of 75%, DeneNet201 obtained an accuracy of 85.23%, DeneNet169 obtained an accuracy of 87.23%, ResNet101V2 obtained an accuracy of 66%, ResNet152V2 obtained an accuracy of 64%, InceptionV3 obtained an accuracy of 84.23%, and InceptionResNetV2 obtained an accuracy of 90%.In comparison to the individual models, the proposed ensemble model achieves 92.23% testing accuracy. The addition of an additional dense layer in each of the seven deep-learning models enabled this improvement. The model is trained and tested on a dataset of 5236 retinal images, with 30 epochs used during training.

The proposed ensemble model provided accurate and reliable classification results with higher accuracy by performing majority voting on the outputs of the seven models. This breakthrough has the potential to transform diabetic retinopathy diagnosis, allowing for earlier identification and more successful treatment, ultimately preventing visual damage and blindness in diabetic patients.

Acknowledgements

This research was supported by [JECRC University]. We thank our guide from [JECRC University] who provided insight and expertise that greatly assisted the research. We thank [Prof. Deepak Dembla] for assistance with [technique, methodology], and [Prof. Deepak Dembla, Dean,JECRC] for comments that greatly improved the manuscript.

Author contributions

Prof. Deepak Dembla1: Conceptualization, Methodology, Software, Field study **Amita Meshram2:** Data curation, Writing-Original draft preparation, Software, Validation., Field study **Anooja.A3:** Visualization, Investigation, Writing-Reviewing and Editing.

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] F. Alzami, R. Aria Megantara, A. Zainul Fanani, P. Nurtantio Andono, and M. Arief Soeleman, “Exudates Detection for Multiclass Diabetic Retinopathy Grade Detection using Ensemble,” 2020.

[2] L. Wang, Z. Chen, M. Wang, T. Wang, W. Zhu, and X. Chen, “Cycle adaptive multi-target weighting network for automated diabetic retinopathy segmentation,” in Proceedings - International Symposium on Biomedical Imaging, Apr. 2021, vol. 2021-April, pp. 1141–1144. doi: 10.1109/ISBI48211.2021.9433917.

[3] T. R. Gadekallu, N. Khare, S. Bhattacharya, S. Singh, P. K. R. Maddikunta, and G. Srivastava, “Deep neural networks to predict diabetic retinopathy,” Journal of Ambient Intelligence and Humanized Computing, 2020, doi: 10.1007/s12652-020-01963-7.

[4] A. G. Priya Henry and A. Jude, “Convolutional neural-network-based classification of retinal images with different combinations of filtering techniques,” Open Computer Science, vol. 11, no. 1, pp. 480–490, Jan. 2021, doi: 10.1515/comp-2020-0177.

[5] W. L. Alyoubi, W. M. Shalash, and M. F. Abulkhair, “Diabetic retinopathy detection through deep learning techniques: A review,” Informatics in Medicine Unlocked, vol. 20. Elsevier Ltd, Jan. 01, 2020. doi: 10.1016/j.imu.2020.100377.

[6] V. Sudha and T. R. Ganeshbabu, “A convolutional neural network classifier VGG-19 architecture for lesion detection and grading in diabetic retinopathy based on deep learning,” Computers, Materials and Continua, vol. 66, no. 1, pp. 827–842, 2021, doi: 10.32604/cmc.2020.012008.

[7] J. G. K. L. J. P. ´ cov ˇ a Slavom´ır Kajan, “Detection of Diabetic Retinopathy Using Pretrained Deep Neural Networks,” 978-1-7281-4381-1/20/$31.00 © 2020 IEEE.

[8] S. K. J. L. Z. M. T. Y. S. W. D. M. M. F. B. E. V. N. and M. V. S. Morgan Heisler1, “Ensemble Deep Learning for Diabetic Retinopathy Detection Using Optical Coherence Tomography Angiography,” . Trans Vis Sci Tech. 2020;9(2):20, https://doi.org/10.1167/tvst.9.2.20, vol. Vol. 9, no. Special Issue, 2022.

[9] J. Raja, P. Shanmugam, and R. Pitchai, “An Automated Early Detection of Glaucoma using Support Vector Machine Based Visual Geometry Group 19 (VGG-19) Convolutional Neural Network,” Wireless Personal Communications, vol. 118, no. 1, pp. 523–534, May 2021, doi: 10.1007/s11277-020-08029-z.

[10] M. H. Mahmoud, S. Alamery, H. Fouad, A. Altinawi, and A. E. Youssef, “An automatic detection system of diabetic retinopathy using a hybrid inductive machine learning algorithm,” Personal and Ubiquitous Computing, 2021, doi: 10.1007/s00779-020-01519-8.

[11] N. S. Murthy and B. Arunadevi, “An effective technique for diabetic retinopathy using hybrid machine learning technique,” Statistical Methods in Medical Research, vol. 30, no. 4, pp. 1042–1056, Apr. 2021, doi: 10.1177/0962280220983541.

[12] L. M. H. H. and K. A. AATILA Mustapha1, “Diabetic Retinopathy Classification Using ResNet50 and VGG-16 Pretrained Networks,” :International Journal of Computer Engineering and Data Science (ISSN:2737-8543) Volume 1– Issue 1, 20/07/2021, 2021.

[13] S. Valarmathi and R. Vijayabhanu, “A Survey on Diabetic Retinopathy Disease Detection and Classification using Deep Learning Techniques,” Mar. 2021. doi: 10.1109/ICBSII51839.2021.9445163.

[14] A. Deshpande and J. Pardhi, “Automated detection of Diabetic Retinopathy using VGG-16 architecture,” International Research Journal of Engineering and Technology, 2021, [Online]. Available: www.irjet.net

[15] A. Bora et al., “Predicting the risk of developing diabetic retinopathy using deep learning,” The Lancet Digital Health, vol. 3, no. 1, pp. e10–e19, Jan. 2021, doi: 10.1016/S2589-7500(20)30250-8.

[16] M. E. Hoque et al., “A deep learning approach for retinal image feature extraction,” Pertanika Journal of Science and Technology, vol. 29, no. 4, pp. 2543–2563, Oct. 2021, doi: 10.47836/PJST.29.4.17.

[17] G. U. Nneji, J. Cai, J. Deng, H. N. Monday, M. A. Hossin, and S. Nahar, “Identification of Diabetic Retinopathy Using Weighted Fusion Deep Learning Based on Dual-Channel Fundus Scans,” Diagnostics, vol. 12, no. 2, Feb. 2022, doi: 10.3390/diagnostics12020540.

[18] G. Mushtaq and F. Siddiqui, “Detection of diabetic retinopathy using deep learning methodology,” IOP Conference Series: Materials Science and Engineering, vol. 1070, no. 1, p. 012049, Feb. 2021, doi: 10.1088/1757-899x/1070/1/012049.

[19] Hanan Saleh Alghamdi,” Towards Explainable Deep Neural Networks for the Automatic Detection of Diabetic Retinopathy” Appl. Sci. 2022, 12, 9435. https://doi.org/10.3390/ app12199435

[20] Ch. Usha Kumari et al.," Deep Learning Based Detection of Diabetic Retinopathy using Retinal Fundus Images," 2022 Third International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT)

[21] Brahami Menaouer et al.," Diabetic Retinopathy Classifcation Using Hybrid Deep Learning Approach" SN Computer Science (2022) 3:357 https://doi.org/10.1007/s42979-022-01240-8

[22] Sambit S. Mondal et,al.," EDLDR: An Ensemble Deep Learning Technique for Detection and Classification of Diabetic Retinopathy", Diagnostics 2023, 13, 124. https:// doi.org/10.3390/diagnostics13010124

[23] Hao Liu et al.," Hybrid Model Structure for Diabetic Retinopathy Classification", Hindawi Journal of Healthcare Engineering Volume 2020, Article ID 8840174, 9 pages https://doi.org/10.1155/2020/8840174

[24] Muhammad Mohsin Butt et al.," Diabetic Retinopathy Detection from Fundus Images of the Eye Using Hybrid Deep Learning Features", Diagnostics 2022, 12, 1607. https://doi.org/10.3390/ diagnostics12071607.

[25] Wei Wang et al.," A New Image Classification Approach via Improved MobileNet Models with Local Receptive Field Expansion in Shallow Layers", Hindawi Computational Intelligence and Neuroscience Volume 2020, Article ID 8817849, 10 pages https://doi.org/10.1155/2020/8817849.

[26] Sohaib Asif et al.,"Improving Effectiveness of Different Deep Transfer Learning-Based Models for Detecting Brain Tumors From MR Images",IEEE Access volume 10 ,2022.

[27] Moctar Abdoul Latif Sawadogo et al."PTSD in the Wild: A Video Database for Studying Post-Traumatic Stress Disorder Recognition in Unconstrained Environments", arXiv:2209.14085v1 [cs.HC] 28 Sep 2022.

[28] Hemant Kumar et al.,"Transfer Learning and Supervised Machine Learning Approach for Detection of Skin Cancer: Performance Analysis and Comparison", Drugs and Cell Therapies in Hematology (ISSN: 2281-4876) Volume 10 Issue 1 (2021).

[29] Adarsh Vulli et al.,"Fine-Tuned DenseNet-169 for Breast Cancer Metastasis Prediction Using FastAI and 1-Cycle Policy", Sensors 2022, 22, 2988. https://doi.org/10.3390/s22082988.

[30] Christian Szegedy et al.,Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning", arXiv:1602.07261v2 [cs.CV] 23 Aug 2016.

[31] N. Dong et al.,"Inception v3 based cervical cell classification combined with artificially extracted features" Applied Soft Computing Volume 93, August 2020, 106311.

[32] Meshram, A., Dembla, D. (2023). Multistage Classification of Retinal Images for Prediction of Diabetic Retinopathy-Based Deep Learning Model. In: Rathore, V.S., Piuri, V., Babo, R., Ferreira, M.C. (eds) Emerging Trends in Expert Applications and Security. ICETEAS 2023. Lecture Notes in Networks and Systems, vol 682. Springer, Singapore. https://doi.org/10.1007/978-981-99-1946-8_20

[33] Meshram, A., Dembla, D. "Mcbm: Implementation Of Multiclass And Transfer Learning Algorithm Based On Deep Learning Model For Early Detection Of Diabetic Retinopathy" Asean Engineering Journal, 2023, 13(3), Pp. 107–116.

[34] Meshram, A., Dembla, D., Anooja, A.," Development And Analysis Of Deep Learning Model Based On Multiclass Classification Of Retinal Image For Early Detection Of Diabetic Retinopathy" Asean Engineering Journal, 2023, 13(3), Pp. 89–97