IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Deep Convolution Neural Network based Multi-Retinal Disease Diagnosis and Classification Model

Dr. S. Megala, Assistant Professor/ Programmer, Annamalai University, Annamalai Nagar.

Dr. T.S. Subashini, Professor, Dept. of Computer Science, Annamalai University, Annamalai Nagar.

Dr. N. Rajkumar, Assistant Professor, Manbumigu Puratchi Thalaivar M.G.R. Government Arts & Science College, Kattumannar Koil.

Abstract

Due to the excessive availability of the fundus cameras, large amount of fundus images are being generated. The fundus images can be utilized by the ophthalmologists to automatically diagnose the retinal diseases. At the same time, the development of artificial intelligence (AI) techniques, especially deep learning (DL) models finds useful for object detection in several application areas, including medical imaging. In this view, this paper presents an automated DL based multi-retinal disease diagnosis and classification model. The presented technique comprises three processes, such as pre-processing, feature extraction, and classification. For feature extraction purposes, AlexNet and Residual Network (ResNet) models are employed. Besides, convolutional neural network (CNN) and deep neural network (DNN) models are applied to classify the retinal fundus images into different retinal diseases. The incorporation of DL models for feature extraction and classification resulted to enhanced diagnostic performance. A series of experiments were conducted on benchmark dataset and the experimental results are examined under different aspects. The experimental values exhibited that the ResNet-CNN model has achieved better performance with the accuracy of 98.84%, precision of 98.84%, recall of 98.17%, and F1-score of 98.48%.

Keywords: Multi-retinal disease, Retinal fundus image, Deep learning, Feature extraction

1. Introduction

Clinical imaging is significant in the medical diagnosis process and unique treatment for eye infections [1]. It offers high-resolution data about the anatomic and functional modifications. Recently, the imaging models were deployed progressively along with the therapeutic enhancements. Therefore, under the improvement of imaging model, comprehension and administration of eye disease is highly complicated because of massive images and findings recorded for patients and hypotheses. The traditional diagnostic approaches are relied on physicians' experience and knowledge which results in maximum rate of misdiagnosis and unwanted clinical data [2]. The medical analysis and therapeutics immediately requires smart devices for managing clinical data significantly. Artificial Intelligence (AI) was employed in various medical applications. Specifically, collaborations among clinical imaging and AI strategies have ensured importance in medical disease diagnosis.

The major objective of this study is to examine the clinical images where retinal images are main factor for ophthalmologists to examine the different types of eye diseases [3]. Automatic examination of digital retina image is feasible to assist ophthalmologist for screening the eye disease. Retinal eye image classification is one of the significant applications in computer vision issues with extensive range of applications in clinical sector. The awareness of retinal image is significant for ophthalmologists for assessing eye infections like glaucoma and hypertension, when the disease is left untreated results in vision issues and loss of eyesight. Traditionally, studies have developed for learning vascular irregularities of retinal images guides the physician to offer primary analysis and treatment.

Retina of human eyes are light-sensitive tissues which are significant and essential for vision power. Obviously, retinal blood vessels are impacted by microvascular state of a brain [4] results in damaged retinal eye for representing systemic microvascular damage related various infections like hypertension and diabetes. In past decades, classifications of retinal images have accomplished an extensive study that tends in massive reports. Even though various models have been presented, examination of retinal eye images are assumed as a challenging processing issues. These challenges have fundal variability in each user. Also, researchers manufactured automated prediction of retinal disease and Diabetic Retinopathy (DR).

AI represents the application of computer science which reflects human cognitive function [5]. ML is defined as sub class of AI which is applied for learning data and predictions are performed; where the operations are categorized as supervised and unsupervised learning. Under the former model, a machine undergoes training with input data obtained by humans for predicting the required output. Therefore, it is time-consuming as it needs acceptable data which has to be labelled with human intervention while in unsupervised learning, a machine is provided with input data that is not labelled explicitly. It helps in finding the structures and patterns from massive collection of objects without manual power. Traditional machine learning (ML) schemes are Decision Tree (DT), naive Bayes (NB), radial basis function (RBF), Random Forest (RF), Support Vector Machine (SVM), k-nearest neighbor (KNN). In spite of gaining best performance with tiny datasets, ML system makes vulnerable for reaching the issues like convergence and

overfitting as manual features selection process degrades the application. Followed by, the effective method is DL, which reflects the process of human brain under the application of various layers of Artificial Neural Networks (ANN) which produces automatic predictions from input data.

This paper develops a new automated DL based multi-retinal disease diagnosis and classification model. The presented techniqueinvolves three processes, such as pre-processing, feature extraction, and classification. Here, AlexNet and Residual Network (ResNet) models are utilized as feature extractors. Moreover, convolutional neural network (CNN) and deep neural network (DNN) models are used to classify the retinal fundus images into different retinal diseases. A detailed experimentation takes place to validate the diagnostic performance of the presented technique and the results are examined under different aspects.

2. Literature Review

Tan et al. [6] applied a deep CNN for predicting AMD with 10-fold cross-validation principle. Gulshan et al. [7] used DL for automated prediction of DR and DME in retinal fundus images on EyePACS-1 dataset as well as Messidor-2 dataset, correspondingly. Even though these approaches are suitable in gaining better results, actual images are applied for training CNN from scratch which requires massive training data as well as processing time for accomplishing maximum classification accuracy. Lu et al. [8] deployed a novel intelligent model according to DL for automated prediction and distinguish multi-categorical irregularities from optical coherence tomography (OCT) images. Multiple images from similar eye is composed of different partition; thus the performance of model has been biased. Li et al. [9] applied Deep Transfer Learning (DTL) approach relied on visual geometry group 16 (VGG-16) system for classifying AMD and DME in retinal OCT images. Karri et al. [10] defined a TL framework on Inception network for significant identification of retinal pathologies applied for retinal OCT images. Then, gradient diminishing is prone to have training process.

Fauw et al. [11] implied a DL structure for examining the OCT images and develops a referral suggestion, where the performance has exceeded for specialists from sight-threatening retinal infections was depicted. Even though the function has enhanced from segmentation method and classification model instances, training time-complexity has been maximized for classification network that depends upon the network segmentation. Fang et al. [12] introduced a lesion-aware CNN (LACNN) technology for classifying retinal OCT images, and employed retinal lesions predicted by using Lesion Detection Network (LDN) inside OCT images which helps in CNN to concentrate on discriminative data from local lesion-related sites for accomplishing accurate classification. [13] projected an ensemble method of multi-scale convolutional mixture of expert (MCME) for classifying normal retina, dry AMD and macular edema, producing the maximum precision rate and AUC with MCME with 4 scale-based experts.

3. The Propsoed Model

The basic operational principle involved in the presented technique is provided in Fig. 1. The figure shown that the input retinal images are initiallypre-processed to raise the image quality and to achieve better decision making. Next, the AlexNet and ResNet models are employed to the pre-processed image for extracting the feature vectors. Finally, the extracted feature vectors are fed into the CNN and DNN classifiers to determine the proper retinal disease class labels.

3.1. Image Pre-processing

At this stage, the input retinal image is initially resized into a default image size of 480*640 pixels. Firstly, the noise exist in the resized image is discarded using median filtering (MF) technique. The MF is a familiar non-linear filtering technique. It can be represented by

$$\hat{f}(x,y) = \text{MEDIAN}_{(k,l)\in S}\big(f(k,l)\big) \qquad ()$$

where f and $\hat{f}$ represents the actual and restored images,and S defines the adjacent pixels of $f(x,y)$. Based on the sliding window mask, several versions of MF can be designed. In the 3x3 MF, 3 x 3 sliding window is used in which the intermediate pixel is restored by the median value of the nine pixels of the window. It operates quite well under minimum noise density. Once the noise removed, the contrast limited adaptive histogram equalization (CLAHE) approach is employed for improving the contrast level of the input image. CLAHE is applied to improve the contrast level of the retinal images[14].

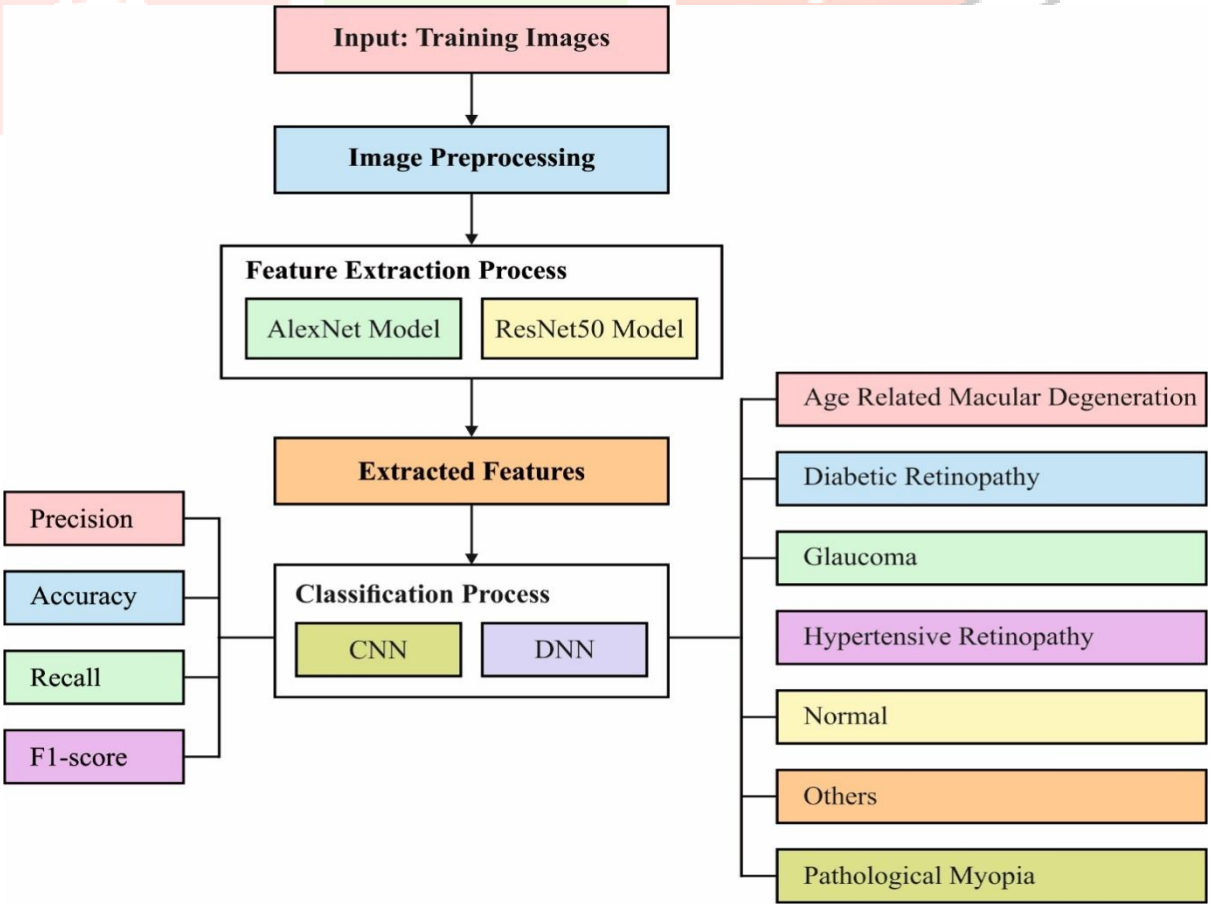


Fig. 1. Block diagram of the present model

3.2. DL based Feature Extractor

A Convolutional Neural Network (ConvNet/CNN) is evolved from DL approach which considers the input as an image and assign significance (learnable weights and biases) for different aspects which is capable to discriminate from one another. The pre-processing is essential in ConvNet, which is lower than alternate classifiers. In case of primitive models, the filters are hand-crafted with adequate training and ConvNets are capable of learning these characteristics. The structure of ConvNet is analogous for connectivity pattern of the neurons in the human brain and developed by the company of Visual Cortex. Unique neurons respond to stimuli when the restricted region of visual application as named as Receptive Field. A ConvNet is suitable in capturing Spatial and Temporal dependencies using related filters. The structure computes a better fitting to image dataset because of reduced parameters applied and reusability of weights. Additionally, the system is trained for learning the sophistication of an image. The structure of CNN is shown in Fig. 2.

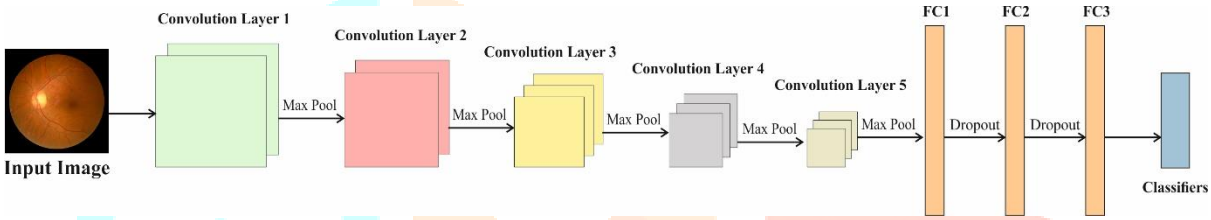


Fig. 2. Structure of CNN

Convolution Layer

The major performance of a convolutional layer is to predict local conjunctions of features from existing layer and mapping the appearance of a feature map. Consequently, convolution in NN is applied in an image is classified as perceptron's, developing local receptive fields and composed of perceptron's in feature maps

Pooling Layer

As same as Convolutional Layer, Pooling layer is applicable to reduce the spatial size of Convolved Feature. It is also applied to reduce the processing energy needed for computing a data by dimensionality reduction. Moreover, it is helpful in filtering dominant features that are rotational as well as positional invariant, hence retain the efficiency of a method. Pooling layers are classified as 2 modules namely, Max and Average Pooling. Max. Pooling offers higher value from the image enclosed by Kernel while average Pooling offers the average values from part of image concealed by the Kernel.

Fully Connected (FC) Layer

The inclusion of a FC layer is a better way to learn non-linear combinations of high-level features as depicted by the result of convolutional layer. The FC layer is learning a feasibly non-linear function in a space. In this study, two DL based feature extractors are utilized to derive the feature vectors from the preprocessed image.

3.2.2. AlexNet model

AlexNet has gained maximum attention from massive developers which is implanted with distinct advantages when compared with traditional approaches. AlexNet is defined as a large scale network with numerous parameters and neurons. The parameters are trained by Krizhevsky [15] with various enhancements. Initially, activation function was applied to offer nonlinearity. The classical activation functions are logistic function, tanh function, arctan function, and so on. However, in deep models, the above defined functions intends to implement the gradient diminishing issues as a gradient is a maximum value only if the input is smaller in range of 0. In order to resolve these issues, a novel activation function has been employed as rectified linear unit (ReLU) which is illustrated as,

$$\mathrm{ReLU}(x) = \max(x, 0) \tag{1}$$

The gradient of ReLU is allocated as 1 when the input is greater than 0. It is ensured that deep networks with ReLU is employed as activation function is converged robust when compared with tanh unit. This simulation is contributed for training. Followed by, dropout has been applied for eliminating the over-fitting issues. Generally, it is applied in Fully Connected (FC) layers. In dropout, portion of neurons are trained in all rounds. Dropout promotes a neuron to be operated with neighbouring one, that limits the joint adaptation from neurons and maximize the generalization process. A network is segregated into various sub-networks using a dropout. Although a single sub-network is applied, then it is over-fitted for greater extent, however it shares the same-loss function. The simulation outcome of complete system is defined as sub-networks. Hence, dropout also enhances the efficiency. Convolution is defined as a model used for signal analysis. The image M in size of (m, n), the convolution is depicted as,

$$C(m, n) = (M * w)(m, n) = \sum_{k} \sum_{l} M(m - k, n - l) w(k, l) \tag{2}$$

where v ∨ refers the convolution kernel in size of (k, l) . Convolution provides a solution for a method learning features from images and parameter sharing limits the model complexity. Pooling facilitates a feature reduction approach. Pooling is considered with a group of neighbouring pixels in a feature map and provides a value for depiction by certain principles.

The feature map of 4×4, the max pooling produces a max value for all 2×2 block, which decreases the feature dimension substantial manner.Cross channel normalization comes under a local normalization model which enhanced the generalization. Then, feature maps are normalized in prior to fed to consecutive layers. This approach is referenced in real-time neurons. The FC layers are used in classification process. The neurons in FC layers are connected directly. The activation function in these layers are named as softmax, which is represented as

$$\mathrm{softmax}(x)_i = \frac{\exp(x_i)}{\sum_{j=1}^{n} \exp(x_j)} \tag{3}$$

Softmax is constrained with the output from (0,1), which proves the activation of neurons.Here, various other models are applied in training AlexNet, such as overlapping pooling and training multiple GPUs. The classification accuracy of AlexNet has surpasses optimal approaches on standard datasets with massive contribution. This structure is also employed for retinal disease classification, as shown in Fig. 3.

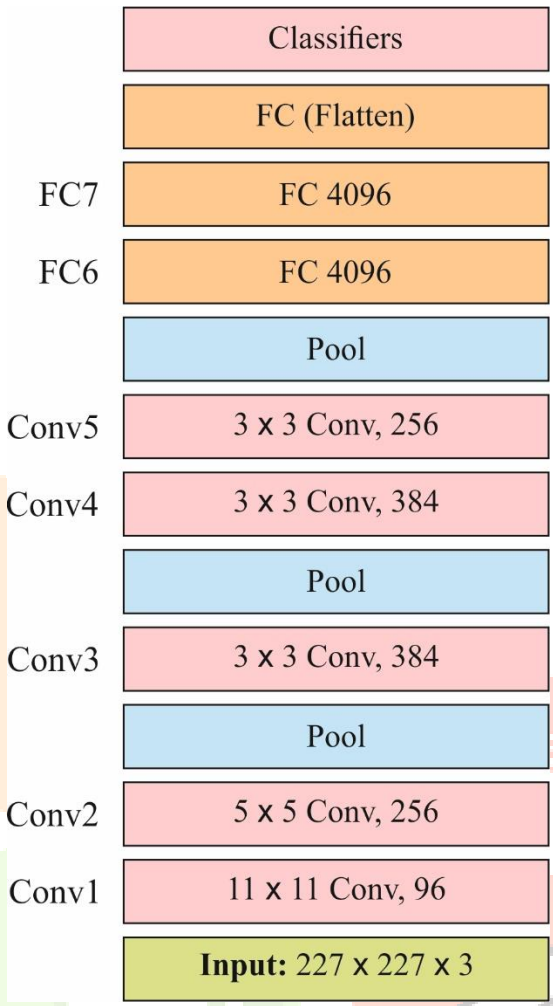


Fig. 3. AlexNet architecture

3.2.3. ResNet50 model

The residual neural network (ResNet) is composed of a residual learning mechanism. The basic idea of ResNet is to use the skip link to allow an input into the layer comprehensively for mitigating the data flow and gradient vanishing problem. The ResNet structure is effective, which can considerably speed up the training of the network and increase the accuracy. The ResNet is depicted as [16]:

$$y = F(x, \{W_j\}) + x \tag{4}$$

where x and y means the input and output vectors of applied layers, W_i refers the weight present in weight matrix, and F showcases the residual mapping to be known. For instance, 2 layers (Fig. 3), where the residual mapping function is given below:

$$F = W_2 f(W_1 x + b_1) + b_2 \tag{5}$$

$$f(x) = \max(0, x) \tag{6}$$

where f implies the ReLU activation function, and W_1 , W_2, b_1 , and b_2 are weights and biases of 1st layer and 2nd layer, correspondingly. The structure of ResNet-50 is shown in Fig. 4.

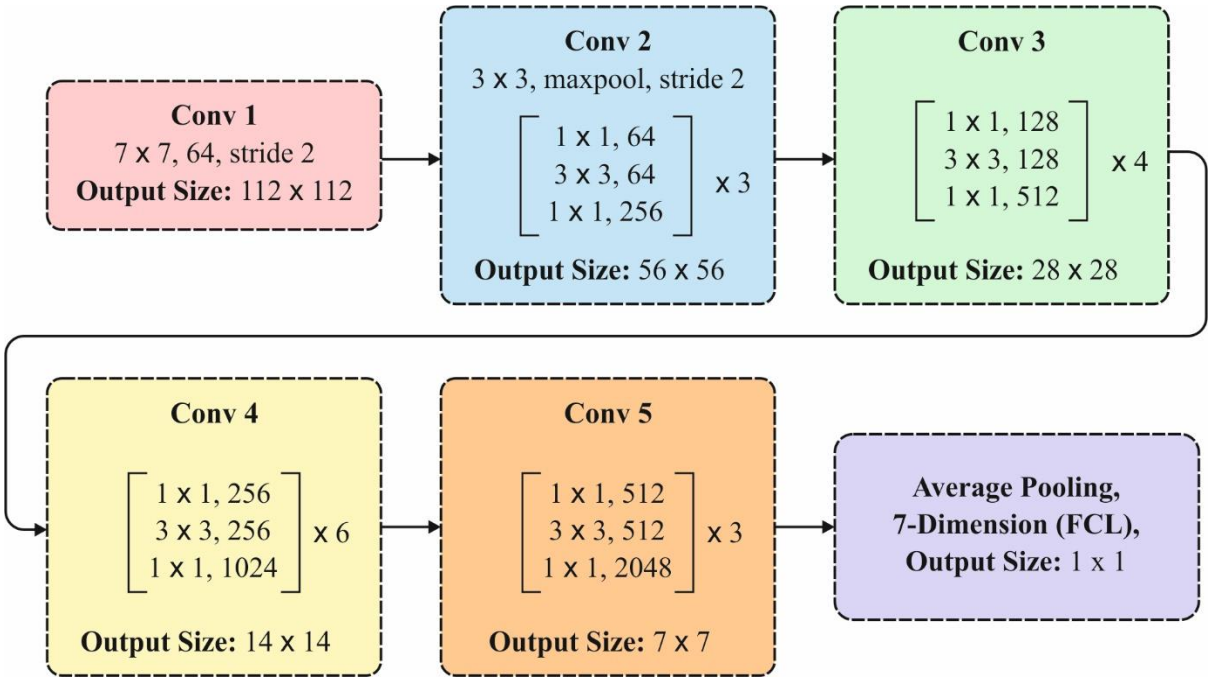


Fig. 4. Layered Structure of ResNet50 model

3.2. DL based Image Classification

After the extraction of features, two DL models namely CNN and DNN are utilized to classify the retinal images into different set of retinal diseases.

3.3.1. CNN Model

In general, autoencoder (AE) is defined as Neural Network (NN) which intends to generate the result of developing model's input [17]. The AE model has accomplished remarkable success in dimensionality reduction model. The basic principle of a hidden layer is facilitated as bottle-neck with few nodes when compared to alternate layers. Moreover, hidden layer is applied for implying the significant attributes of an image with minimum of data. Under the application of image inputs, AE transforms the unstructured information as feature vectors which are computed using ML frameworks like k-means clustering technology.

Encode

A CNN-related AE is divided as 2 principle phases namely Encoding and Interpreting, as given below.

$$O_m(\mathrm{i},j) = a\left(\sum_{d=1}^{\mathrm{D}}\sum_{u=-2k-1}^{2k+1}\sum_{v=-2k-1}^{2k+1} F_{m_d}^{(1)}(\mathrm{u},\mathrm{v})I_d(i-\mathrm{u},\mathrm{j}-v)\right) \quad (7)$$

$$O_m(\mathrm{i},j) = a\left(\sum_{d=1}^{\mathrm{D}}\sum_{u=-2k-1}^{2k+1}\sum_{v=-2k-1}^{2k+1} F_{m_d}^{(1)}(\mathrm{u},\mathrm{v})I_d(i-\mathrm{u},\mathrm{j}-v)\right) \quad (8)$$
$$m = 1,\dots,n$$

where $F \in \left\{F_1^{(1)}, F_2^{(1)}, F_n^{(1)},\right\}$ denotes a convolutional filter, with convolution between input volume as illustrated by $I = \{I_1, \ldots I_D\}$ where it represents the input by the combination of non-linear functions:

$$z_m = O_m = a\left(I * F_m^{(1)} + b_m^{(1)}\right) m = 1,, m \tag{9}$$

where $b_m^{(1)}$ means the bias, and No. of zeros for mapping the input in which: $\dim(\mathrm{I}) = \dim\left(\mathrm{decode}(\mathrm{encode}(\mathrm{I}))\right)$. Consequently, the encoding convolution is similar to:

$$\begin{aligned} O_w = O_h = {} & (I_w + 2(2k+1) - 2) - (2k+1) + 1 \\ & = I_w + (2k+1) - 1 \end{aligned} \tag{10}$$

Decode

The decoding convolution step generates n feature maps $z_{m=1,\ldots,n}$. The redeveloped final outcomes $\hat{I}$ is a result of convolution among volume of feature maps $Z = \{z_{i=1}\}^n$ and convolutional filters volume $F^{(2)}$.

$$\tilde{I} = a\left(Z * F_m^{(2)} + b^{(2)}\right) \tag{11}$$

$$O_w = O_h = (I_w + (2k+1) - 1) - (2k+1) + 1 = I_w = I_h \tag{12}$$

where Eq. (5) refers the decoding convolution with I dimensions. Thus, input's dimensions are similar to final dimensions.

3.3.2. DNN Model

The DL model employed in this study is defined as a feed-forward ANN, which are trained with Stochastic Gradient Descent (SGD) by applying back propagation (BP). In this model, various layers of hidden units have been employed among the inputs and outputs of a method [18].A hidden unit, j, generally applies a logistic function β as it is relevant to hyperbolic tangent applied in a function with equipped derivative for mapping the outputs in $\mathrm{y_j}$ overall input from $\mathrm{x_j}$:

$$\mathrm{y_i} = \beta(\mathrm{x_j}) = \frac{1}{1 + \mathrm{e}^{-\mathrm{xj}}} \tag{13}$$

In case of multiclass classification, a problem of hyper-spectral image classification, output unit j transforms the overall input, $\mathrm{x_j}$, as class probability, $\mathrm{P_j}$, by applying a normalized exponential function termed as "softmax:

$$\mathrm{P_j} = \frac{\exp(\mathrm{X_j})}{\Sigma_{\mathrm{h}} \exp(\mathrm{X_h})} \tag{14}$$

where h means the index over all classes. DNNs are differentially trained using backpropagating derivatives of cost function which estimates the discrepancy among the desired results and original simulation outcome as generated for a training case [$\mathcal{BB}$]. While applying the softmax output function, natural cost function C is a cross-entropy among target probabilities d and softmax outputs, P:

$$\mathrm{C} = -\sum_{\mathrm{i}} \mathrm{d_j} \ln \mathrm{P_j} \tag{15}$$

where a target probability, by considering values of 1 or 0, are supervised data offered for training DNN method.

4. Performance Validation

The presented technique is implemented in a PC i5-8600k processor, GeForce 1050Ti, 4GB RAM, 16GB OS Storage and 250GB SSD File Storage. To simulate the presented technique, Python 3.6.5 tool is utilized with few packages namely tensorflow, keras, numpy, pickle, matplotlib, sklearn, pillow, and opencv-python. The experimental results analysis of the presented technique is evaluated on the benchmark dataset, which comprises images under seven class labels such as AgeRelatedMacularDegeneration (AMD), DR, Glaucoma, HypertensiveRetinopathy (HR), Normal, Others, and PathologicalMyopia (PM). All the classes have a set of 200 images excluding the HR, which contains 92 images. The information relevant to the dataset is given in Table 1 and the test images are given in Fig. 5.

Table 1 Dataset Descriptions

Class Names	**No. of Images**
Age_Related_Macular_Degeneration	200
Diabetic_Retinopathy	200
Glaucoma	200
Hypertensive_Retinopathy	92
Normal	200
Others	200
Pathological_Myopia	200

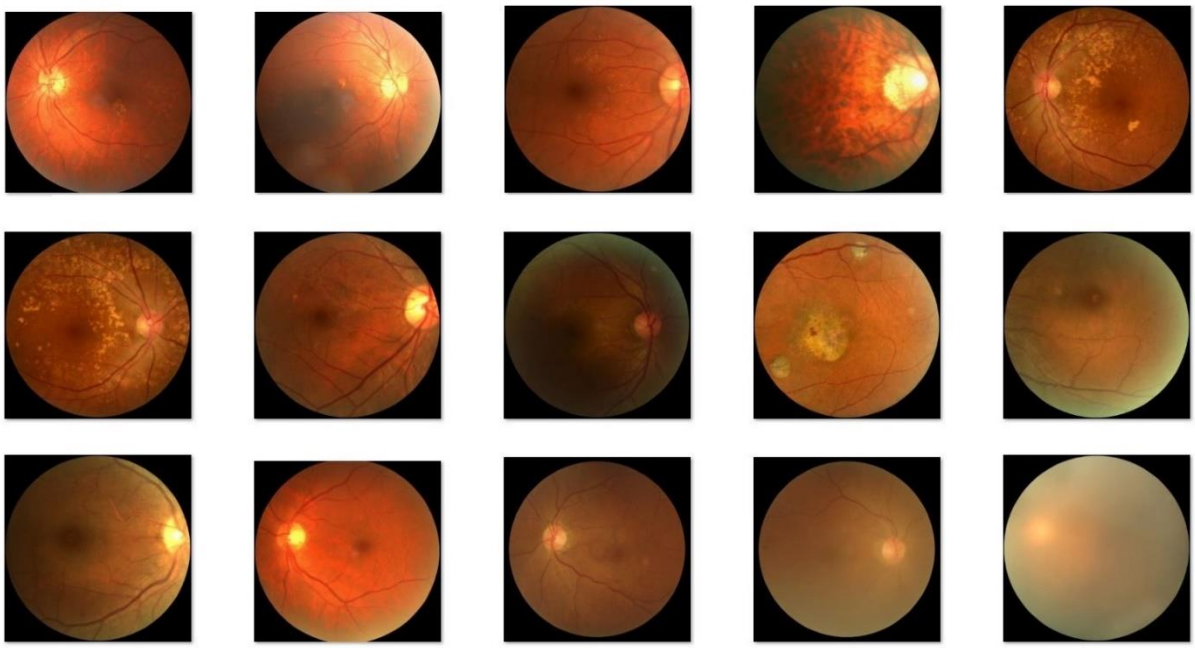


Fig. 5. Sample Images

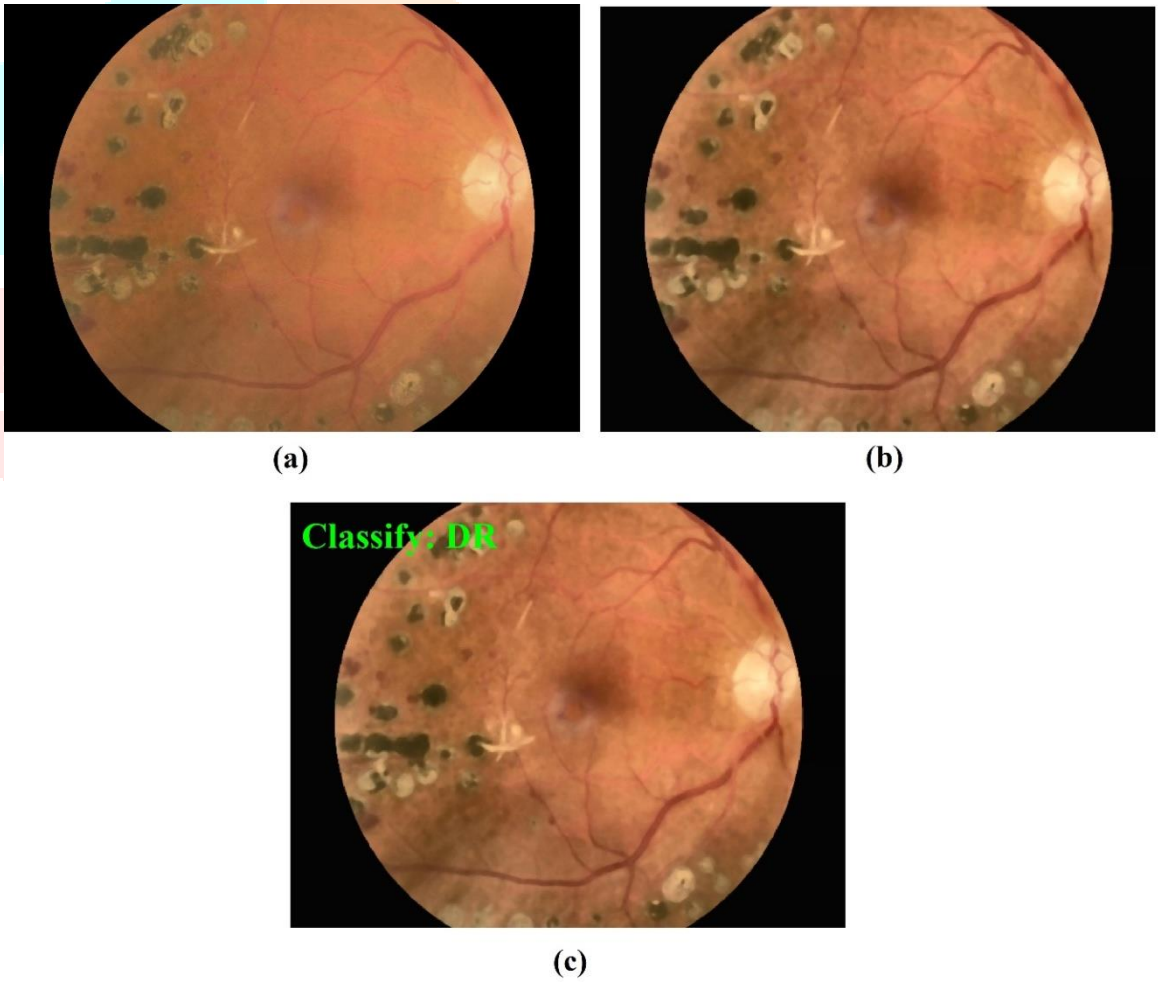
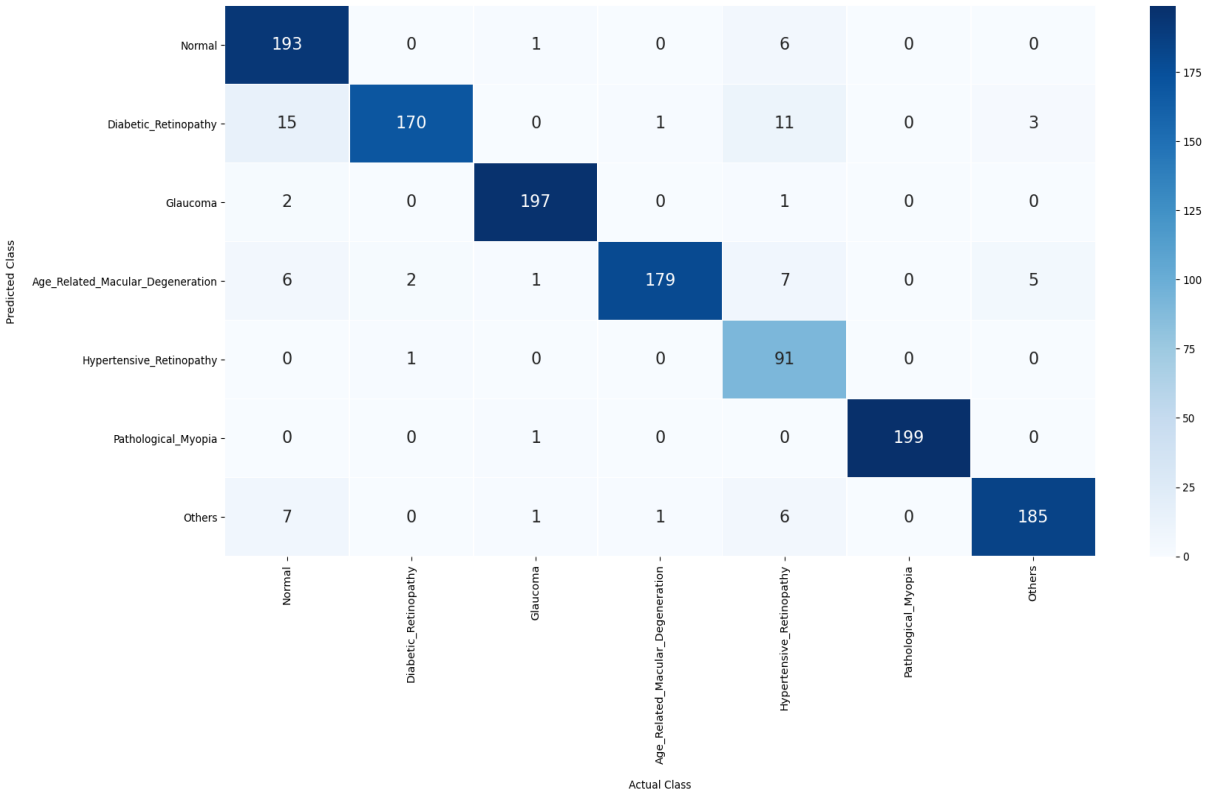


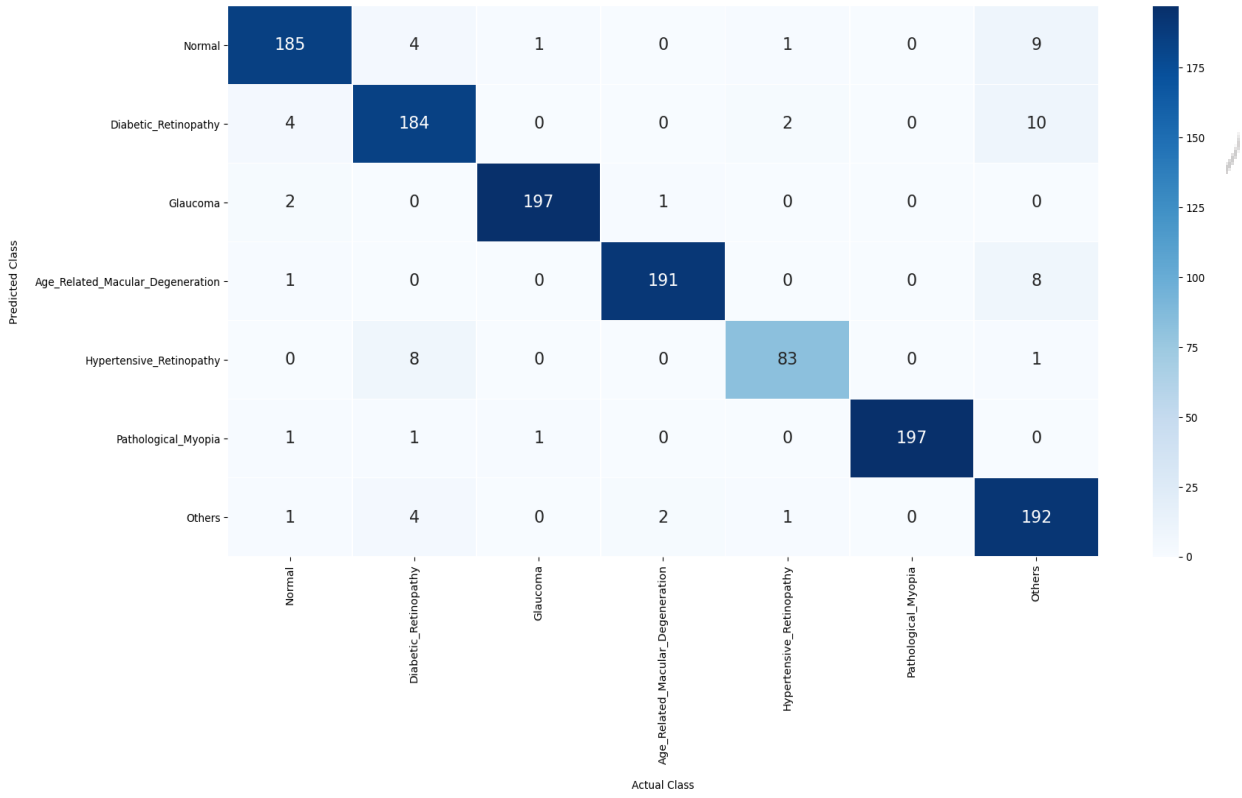
Fig. 6. a) Original Images b) Preprocessed Images c) Classified Image

Fig. 6shows the visualization results analysis of the presented technique. Fig. 1a illustrates the actual retinal image and the preprocessed image is depicted Fig. 1b. The figure represents that the image quality is suggestivelyimprovedusing preprocessing techniques.

Figs. 7-8exhibits the confusion matrices produced by AlexNet-DNN, AlexNet-CNN, ResNet50-DNN, and ResNet50-CNN models during simulation. Fig. 7a illustrates the confusion matrix of the AlexNet-DNN model, indicating that it has categorized 193 images intoNormal_class, 170 images into DR class, 197 images into Glaucoma, 179 images into AMD class, 91 images into HR, 199 images into PM, and 185 images into other class. Fig. 7b outlines the confusion matrix generated by the AlexNet-CNN model. The figure delineate that the AlexNet-CNN model has proficiently classified a set of 185 images intoNormal_class, 184 images into DR class, 197 images into Glaucoma, 191 images into AMD class, 83 images into HR, 197 images into PM, and 192 images into other class.



(a)



(b)

Fig.7. Confusion Matrix a) AlexNet-DNN b) AlexNet-CNN

Fig. 8adelivers the confusion matrix formed at the execution time of the ResNet50-DNN model. From the figure, it is noticed that the ResNet50-DNN model has competently classified a total of 188 images intoNormal_class, 193 images into DR class, 198 images into Glaucoma, 195 images into AMD class, 92 images into HR, 200 images into PM, and 197 images into other class.

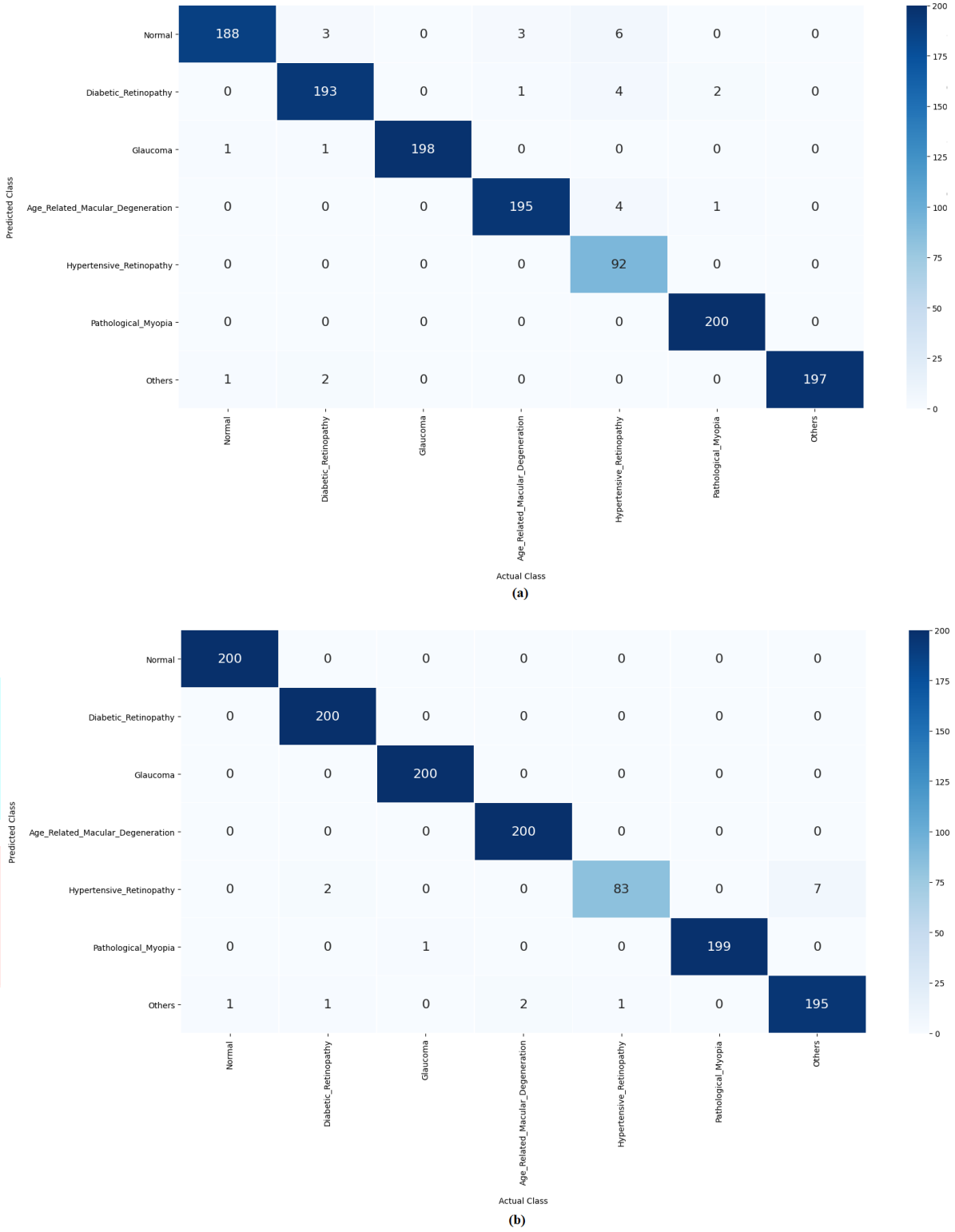
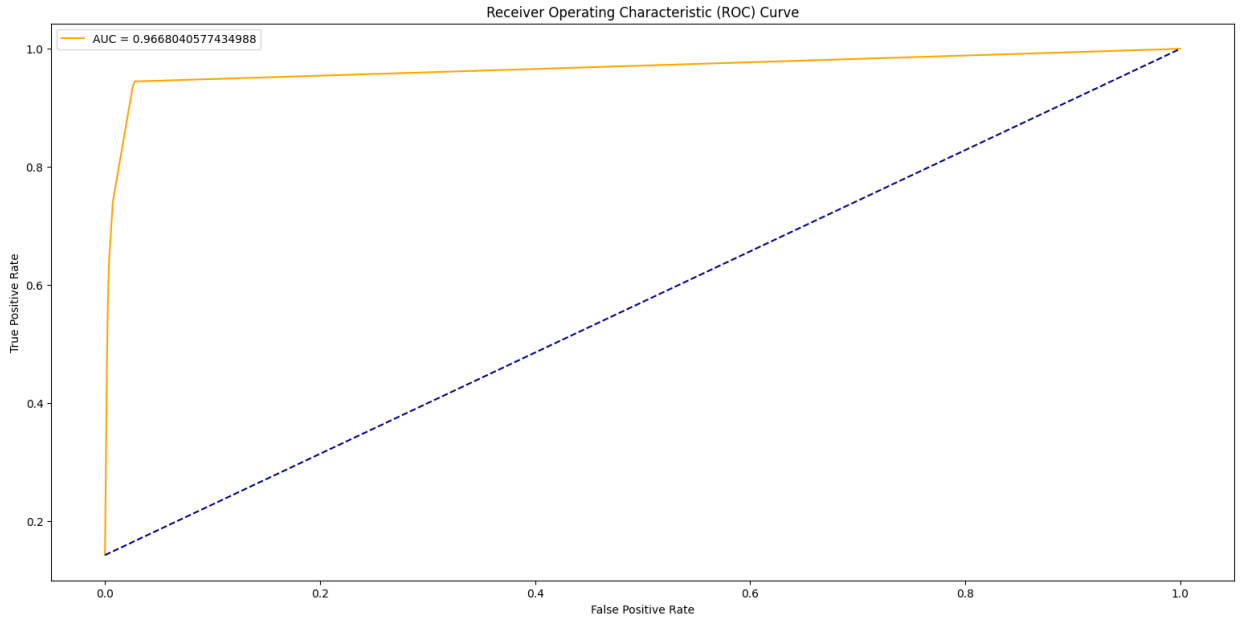
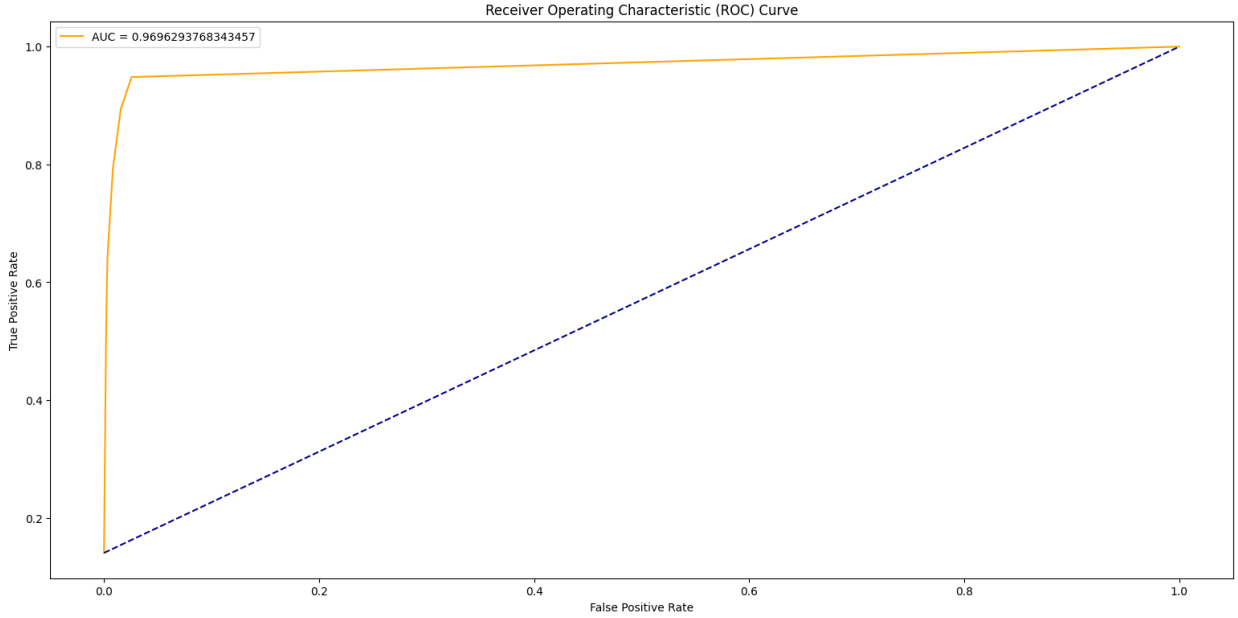


Fig. 8. Confusion Matrix a) ResNet50-DNN b) ResNet50-CNN

Lastly, Fig. 8b presents the confusion matrix of the ResNet50-CNN model. From the figure, it can be observed that the ResNet50-DNN model has proficiently classified a set of 200 images into Normal_class, 200 images into DR class, 200 images into Glaucoma, 200 images into AMD class, 83 images into HR, 199 images into PM, and 195 images into other class.



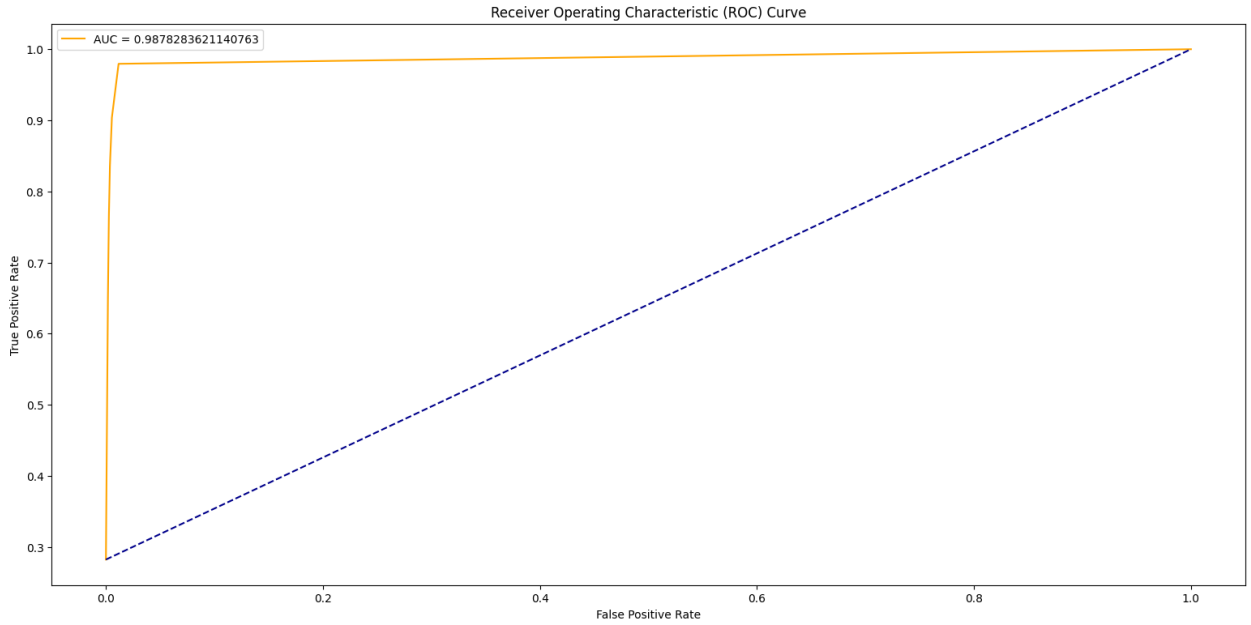
(a)



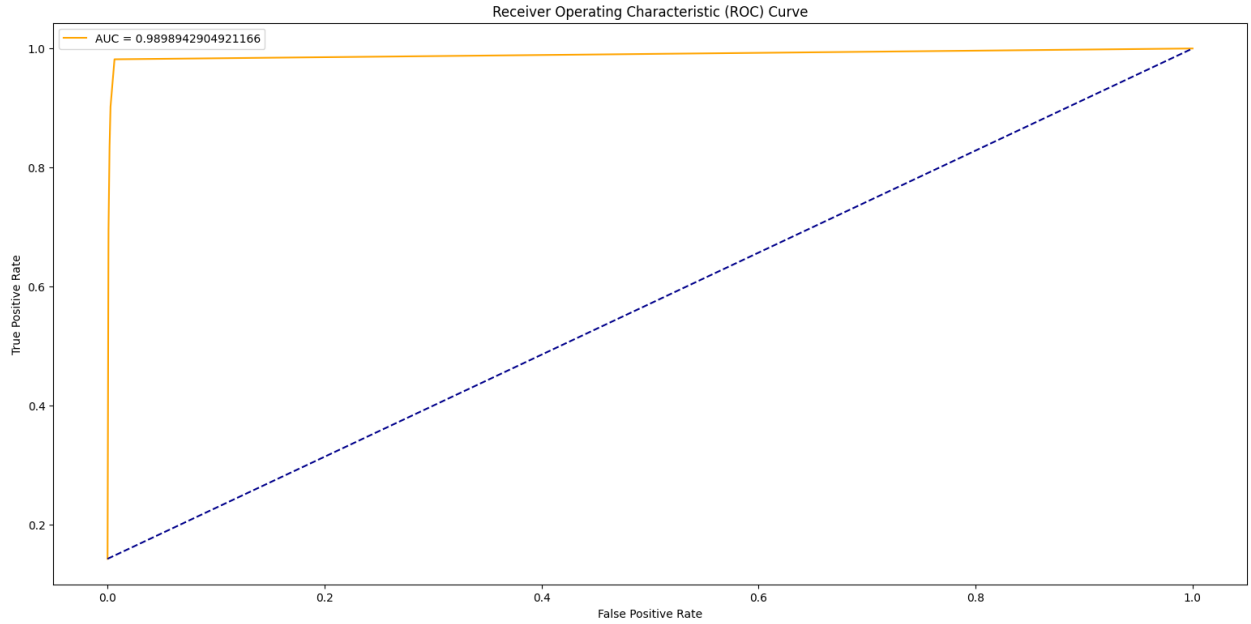
(b)

Fig. 9.ROC Analysis a) AlexNet-DNN b) AlexNet-CNN

Fig. 9expresses the ROC analysis of the two presented techniques namely AlexNet-DNN and AlexNet-CNN. The figure revealed that the AlexNet-DNN model has attainedoperative multi-retinal diagnostic outcome with the high AUC of 0.9668. Similarly, the figure also described that the AlexNet-CNN model has led to a certainly higher AUC value of 0.9696.



(a)



(b)

Fig. 10.ROC Analysis a) ResNet50-DNN b) ResNet50-CNN

Fig. 10showcases the ROC analysis of the two presented techniques such as ResNet50-DNN and ResNet50-CNN. The figure shown that the ResNet50-DNN model has reachedto efficientmulti-retinal diagnostic outcome with the high AUC of 0.9878. Similarly, the figure also conveyed that the ResNet50 model has led to a somewhat higher AUC value of 0.9898. Among the all proposed models, the ResNet50-CNN model has displayed a supreme AUC value over the other presented techniques.

Table 2 Result Analysis of Proposed Methods in terms of Different Measures

Methods	Accuracy	Precision	Recall	F1-Score
AlexNet-DNN	93.96	93.17	94.34	93.37
AlexNet-CNN	95.12	95.29	94.75	94.97
ResNet50-DNN	97.76	97.03	97.93	97.40
ResNet50-CNN	98.84	98.84	98.17	98.48

Table 2 and Figs. 11-12summaries the diagnostic outcome of the presented techniques in the classification of multi-retinal diseases. The experimental results notified that the AlexNet-DNNtechnique has found to be a least performer over other proposed methods with the accuracy of 93.96%, precision of 93.17%, recall of 94.34%, and F1-score of 93.37% correspondingly. Simultaneously, the AlexNet-CNN model has led to a moderate outcome with the accuracy of 95.12%, precision of 95.29%, recall of 94.75%, and F1-score of 94.97% correspondingly. After that, the ResNet50-DNN model has outpaced the previous two models by accomplishing a high accuracy of 97.76%, precision of 97.03%, recall of 97.93%, and F1-score of 97.40 % correspondingly. But the presented ResNet50-CNN model has gained superior performance with the accuracy of 98.84%, precision of 98.84%, recall of 98.17%, and F1-score of 98.48% correspondingly.

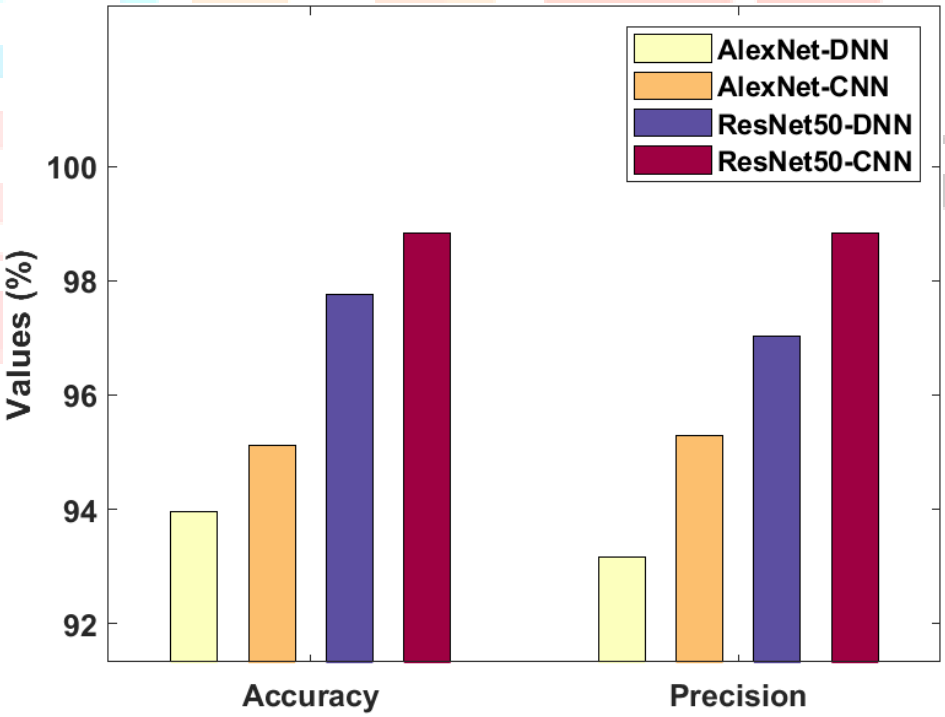


Fig. 11. Accuracy and Precision analysis of proposed model

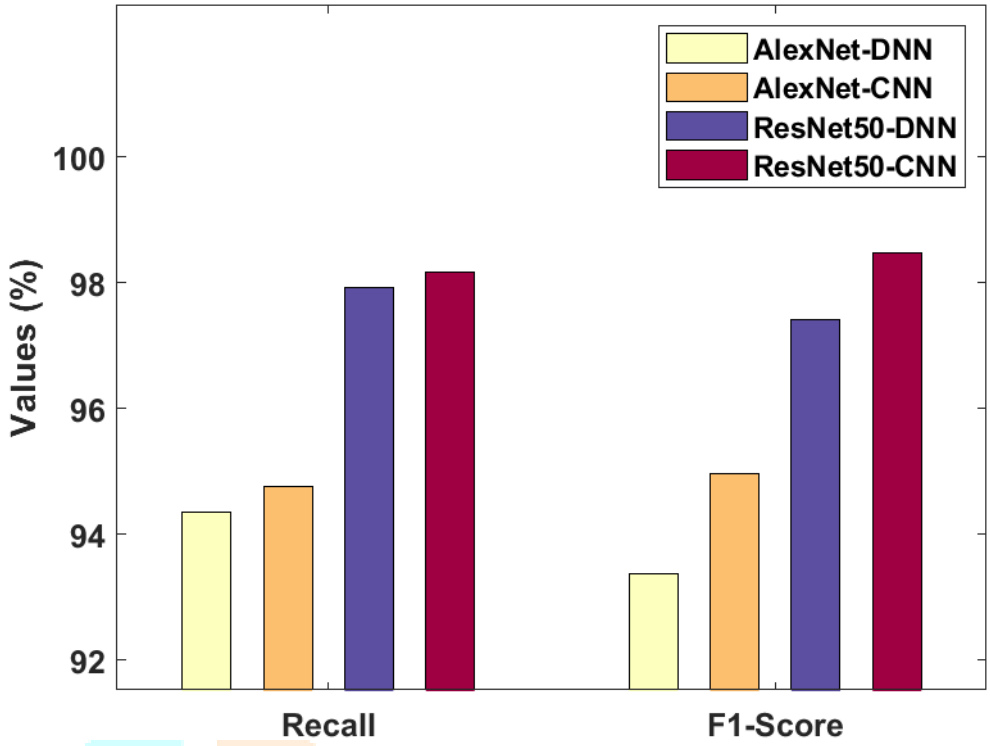


Fig. 12.F-score and F1-score analysis of proposed model

Table 3 offers a detailed comparative study of the results obtained by the presented and existing techniques [19-24]. From the table, it can be seen that the presented techniques have shown better results compared to the other methods.

Table 3 Result Analysis of Existing with Proposed Method in terms of Different Measures

Methods	**Accuracy**	**Precision**	**Recall**	**F1-Score**
ResNet50-CNN	98.84	98.84	98.17	98.48
ResNet50-DNN	97.76	97.03	97.93	97.40
AlexNet-CNN	95.12	95.29	94.75	94.97
AlexNet-DNN	93.96	93.17	94.34	93.37
VGG-16	87.63	89.65	87.12	85.30
DRDL	91.73	92.43	79.32	85.21
CNN-DL	80.93	81.23	80.45	80.96
MLP	81.54	82.14	81.32	81.78
RBF	83.08	83.86	83.03	83.17
SVM	81.54	82.07	81.42	81.68
ANN	75.00	75.90	74.87	75.13
NB	66.90	67.16	78.30	74.72

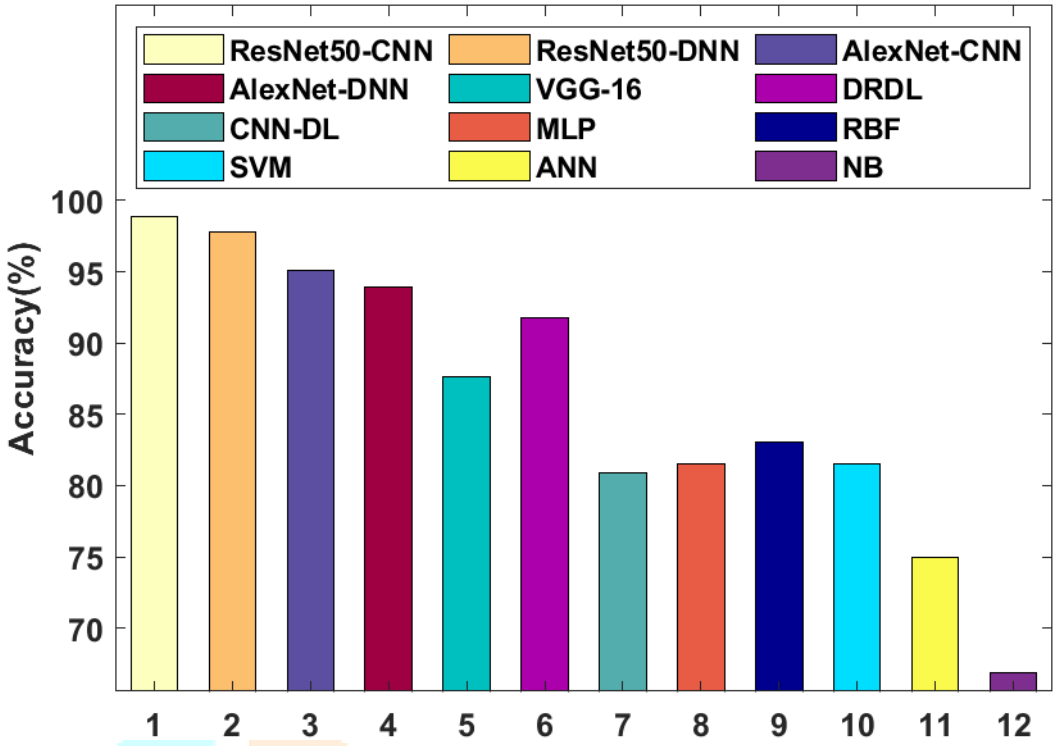


Fig. 13. Comparative results analysis interms of accuracy

Fig. 13examines the comparison study of the presented and existing techniques such as DRDL, CNN-DL, multilayer perceptron, radial basis function, ANN, and Naive Bayes (NB) with respect to accuracy. The obtained results notified that the NB technique has achieved worse outcome with the minimum accuracy of 66.9%. Eventually, the ANN model reached to a certainly better accuracy of 75%. Next, the CNN-DL model has controlled to exhibit certainly considerable performance with the better accuracy of 80.93%. Moreover, the MLP model has demonstrated moderate performance with the accuracy of 81.54% whereas the SVM model has also accomplished an accuracy of 81.54%. Afterward, the RBF model has achieved to a reasonable accuracy of 83.08%. In line with, the VGG-16 and DRDL models have appeared competitive with the accuracy of 87.63% and 91.73%. However, the presented techniques have surpassed the existing methods with the maximum accuracy values of 98.84%, 97.76%, 95.12%, and 93.96%.

Fig. 14 implies the comparative analysis of the proposed and previous models with respect to precision and recall. The figures portrayed that the NB method has attained insignificant classifier results with minimum precision of 67.16% and recall of 78.30%. Additionally, the ANN model has attempted to enhance the results by attaining a moderate precision of 75.90% and recall of 74.87%. Likewise, the CNN-DL scheme has tried to show certainly acceptable performance with the better precision of 81.23% and recall of 80.45%. In line with this, the SVM model has illustrated considerable function with the precision of 82.07% and recall of 81.42% whereas the MLP model has also attained a precision of 82.14% and recall of 81.32%. Then, the RBF framework has gained an acceptable precision of 83.86% and recall of 83.03%. Similarly, the VGG-16 and DRDL methodologies have demonstrated better result with precision and recall of 89.65%, 92.43% and 87.12%, 79.32%. However, the presented approaches have outperformed the conventional techniques with the maximum precision and recall values of 98.84%, 97.03%, 95.29%, and

93.17% and 98.17%, 97.93%, 94.75%, and 94.34%. From the available proposed techniques, a highest precision and recall of 98.84% and 98.17% has been provided by the ResNet50-CNN model.

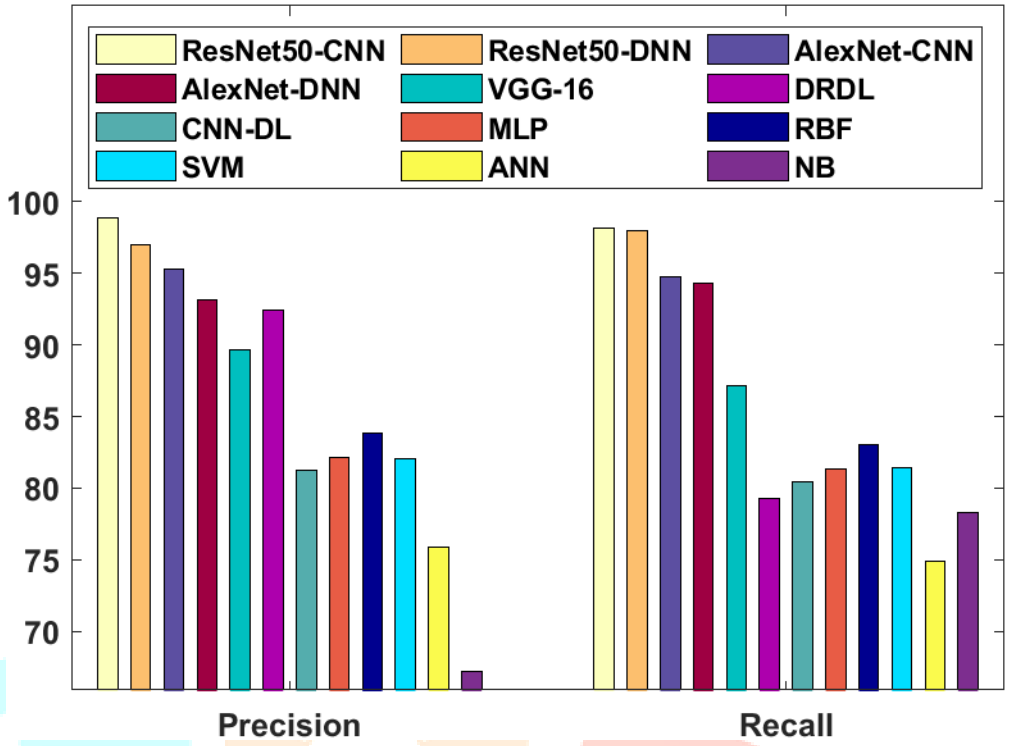


Fig. 14. Comparative results analysis interms of precision and recall

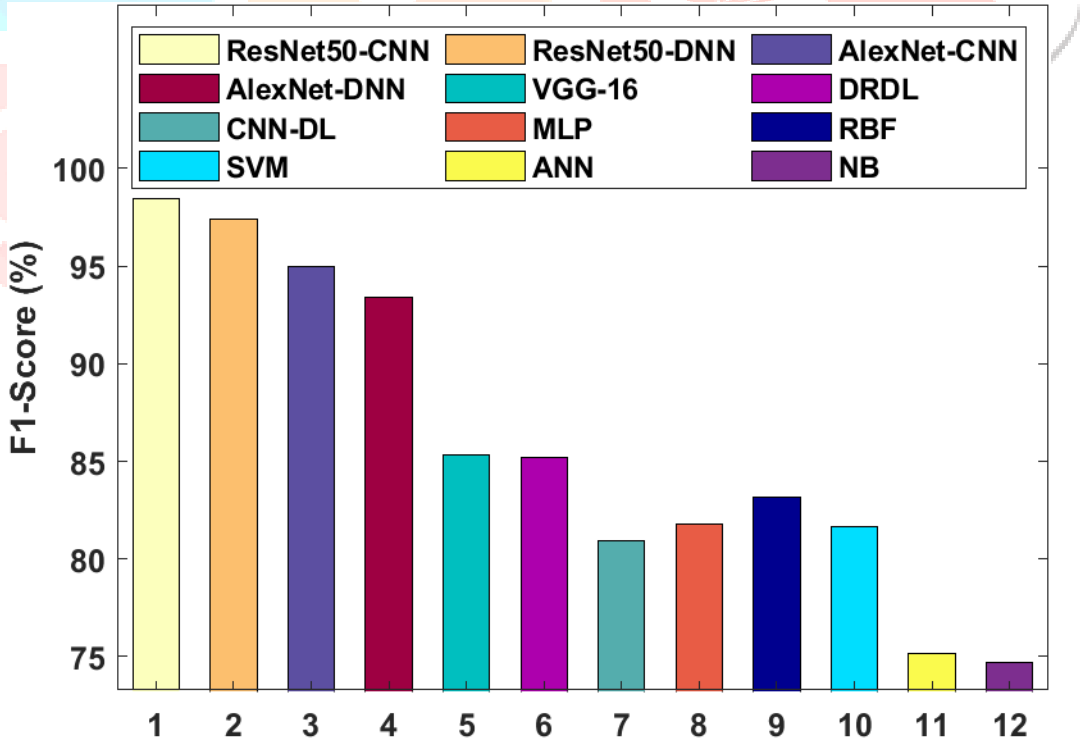


Fig. 15. Comparative results analysis interms of F1-score

Fig. 1 showcases the comparative investigation of the newly developed and former methods by means of F1-score. The figures implied that the NB framework has attained ineffective classifier results with the minimal F1-score of 74.72%. In addition, the ANN scheme has managed to maximize the results by gaining acceptable F1-score of 75.13%. Likewise, the CNN-DL model has tried to represent considerable

function with the better F1-score of 80.96%. Along with that, the SVM model has illustrated acceptable performance with the F1-score of 81.68% while the MLP model has also achieved an F1-score of 81.78%. Next, the RBF model has attained a moderate F1-score of 83.17%. Similarly, the VGG-16 and DRDL frameworks have implied better result with F1-score of 85.30% and 85.21%. However, the presented models have performed well than the traditional approaches with the optimal F1-score values of 98.48%, 97.40%, 94.97%, and 93.37%. From the available projected models, a maximum F1-score of 98.48% has been generated by the ResNet50-CNN model.

5. Conclusion

This paper develops a new automated DL based multi-retinal disease diagnosis and classification model. The presented technique involves three processes, such as pre-processing, feature extraction, and classification. Firstly, the input retinal images are pre-processed to raise the image quality and to achieve better decision making. Next, the AlexNet and ResNet models are employed to the pre-processed image for extracting the feature vectors. Finally, the extracted feature vectors are fed into the CNN and DNN classifiers to determine the proper retinal disease class labels. The utilization of DL models for feature extraction and classification resulted to enhanced diagnostic performance. Extensive set of simulations were carried out to ensure the superior performance of the presented technique. The attained experimental values indicated that the ResNet-CNN model has achieved better performance with the accuracy of 98.84%, precision of 98.84%, recall of 98.17%, and F1-score of 98.48%. As a part of future extension, the hyperparameters of the DL models are tuned by the use of metaheuristic algorithms.

References

[1] Bernardes R, Serranho P, Lobo C. Digital ocular fundus imaging: a review. Ophthalmologica. 2011;226(4):161–81.

[2] Murdoch TB, Detsky AS. The inevitable application of big data to health care. JAMA. 2013;309(13):1351–2.

[3] T.Y. Wong, R. McIntosh, Hypertensive retinopathy signs as risk indicators of cardiovascular morbidity and mortality, Br. Med. Bull. 73 (1) (2005) 57–70.

[4] N. Patton, T. Aslam, T. MacGillivray, A. Pattie, I.J. Deary, B. Dhillon, Retinal vascular image analysis as a potential screening tool for cerebrovascular disease: a rationale based on homology between cerebral and retinal microvasculatures, J. Anat. 206 (4) (2005) 319–348.

[5] Russell S, Bohannon J. Artificial intelligence. Fears of an AI pioneer Science. 2015;349(6245):252.

[6] J. H. Tan, S. V. Bhandary, S. Sivaprasad, Y. Hagiwara, A. Bagchi, U. Raghavendra, A. K. Rao, B. Raju, N. S. Shetty, A. Gertych, K. C. Chua, and U. R. Acharya, “Age-related macular degeneration detection using deep convolutional neural network,” Future Gener. Comp. Sy. 87, 127–135 (2018).

[7] V. Gulshan, L. Peng, M. Coram, M. C. Stumpe, D. Wu, A. Narayanaswamy, S. Venugopalan, K. Widner, T. Madams, J. Cuadros, R. Kim, R. Raman, P. C. Nelson, J. L. Mega, and D. R. Webster,

"Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," JAMA 316(22), 2402–2410 (2016).

[8] W. Lu, Y. Tong, Y. Yu, Y. Xing, C. Chen, and Y. Shen, "Deep learning-based automated classification of multi-categorical abnormalities from optical coherence tomography images," Trans. Vis. Sci. Techn. 7(6), 1–10 (2018).

[9] F. Li, H. Chen, Z. Liu, X. Zhang, and Z. Wu, "Fully automated detection of retinal disorders by image-based deep learning," Graefe's Arch. Clin. Exp. 257(3), 495–505 (2019).

[10] S. P. K. Karri, D. Chakraborty, and J. Chatterjee, "Transfer learning based classification of optical coherence tomography images with diabetic macular edema and dry age-related macular degeneration," Biomed. Opt. Express 8(2), 579–592 (2017).

[11] J. D. Fauw, et al., "Clinically applicable deep learning for diagnosis and referral in retinal disease," Nat. Med. 24(9), 1342–1350 (2018).

[12] L. Fang, C. Wang, S. Li, H. Rabbani, X. Chen, and Z. Liu, "Attention to lesion: Lesion-aware convolutional neural network for retinal optical coherence tomography image classification," IEEE Trans. Med. Imaging 38(8), 1959–1970 (2019).

[13] R. Rasti, H. Rabbani, A. Mehridehnavi, and F. Hajizadeh, "Macular OCT classification using a multi-scale convolutional neural network ensemble," IEEE T. Med. Imaging 37(4), 1024–1034 (2018).

[14] Jadoon, M.M., Zhang, Q., Haq, I.U., Butt, S. and Jadoon, A., 2017. Three-class mammogram classification based on descriptive CNN features. *BioMed research international*, *2017*.

[15] Lu, S., Lu, Z. and Zhang, Y.D., 2019. Pathological brain detection based on AlexNet and transfer learning. *Journal of computational science*, *30*, pp.41-47.

[16] Li, Z., Lin, Y., Elofsson, A. and Yao, Y., 2020. Protein Contact Map Prediction Based on ResNet and DenseNet. *BioMed Research International*, *2020*.

[17] Kowsari, K., Sali, R., Ehsan, L., Adorno, W., Ali, A., Moore, S., Amadi, B., Kelly, P., Syed, S. and Brown, D., 2020. HMIC: Hierarchical Medical Image Classification, A Deep Learning Approach. *Information*, *11*(6), p.318.

[18] Shadman Roodposhti, M., Aryal, J., Lucieer, A. and Bryan, B.A., 2019. Uncertainty assessment of hyperspectral image classification: Deep learning vs. random forest. *Entropy*, *21*(1), p.78.

[19] BambangKrismonoTriwijoyo, Boy SubirosaSabarguna , Widodo Budiharto , Edi Abdurachman, Deep learning approach for classification of eye diseases based on color fundus images, 2020, Diabetes and Fundus OCT, Elsevier, pp. 25-57.

[20] Zheng, Y., Hijazi, M.H.A. and Coenen, F., 2012. Automated "disease/no disease" grading of age-related macular degeneration by an image mining approach. Investigative ophthalmology & visual science, 53(13), pp.8310-8318.

[21] García, M., López, M.I., Álvarez, D. and Hornero, R., 2010. Assessment of four neural network based classifiers to automatically detect red lesions in retinal images. Medical engineering & physics, 32(10), pp.1085-1093.

[22] J. Pradeepkandhasamy, S. Balamurali, M. Arun, S. Gokulnath, Automated Recognition of Diabetic Retinopathy using Machine Learning Techniques, 2019, International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8 Issue-4S2.

[23] Arunkumar, R. and Karthigaikumar, P., 2017. Multi-retinal disease classification by reduced deep learning features. Neural Computing and Applications, 28(2), pp.329-334.

[24] Tong, Y., Lu, W., Yu, Y. and Shen, Y., 2020. Application of machine learning in ophthalmic imaging modalities. Eye and Vision, 7, pp.1-15.