IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁑ Impact Factor 8.102 ⁑ Peer-reviewed / Refereed journal ⁑ Vol. 12, Issue 6, June 2023

DOI: 10.17148/IJARCCE.2023.12675

Diabetic Retinopathy Detection and Multi Stage Classification using Deep Learning Models: A Quick Review

Ms. C. Saraswathy[1], Dr. S. Sarumathi[2], Ms. Sharmila Mathivanan[3], Mr. D. Poornakumar[4]

Associate Professor, Department of Electronics and Communication Engineering, K.S.Rangasamy College of Technology, Tamil Nadu, India[1]

Professor, Department of Artificial Intelligence and Data Science, K.S.Rangasamy College of Technology, Tamil Nadu, India[2]

Lecturer, Department of Computing (IT), FPT Greenwich University, Ho Chi Minh City, Vietnam[3]

Assistant Professor, Department of Electronics and Communication Engineering, K.S.Rangasamy College of Technology, Tamil Nadu, India[4]

Abstract: Diabetes is a disorder that causes an increase in blood glucose levels due to a lack of insulin and affects 425 million persons globally. Diabetes is the most common cause of retinopathy. The retina is the photosensitive tissue that lines the inside of the eye. Hyperglycemia (high blood sugar) can cause retinal vascular damage. Diabetic Retinopathy (DR) is a diabetic eye condition that causes the blood vessels of the retina to enlarge and leak fluids and blood. If left uncontrolled, it might cause partial or total blindness. The sustained eyesight can be treated, but it cannot be restored to its former state. The disease's prognosis worsens with age. This paper presents a detailed review of various retinopathy detection methods. A comparative study is conducted with their merits and demerits for identifying the challenges in those techniques and then this paper is concluded with suggestions of solutions for enhancing the efficiency of deep learning models.

Keywords: Diabetic Retinopathy, Principal Component Analysis, Convolutional Neural Network, Deep Learning.

I. INTRODUCTION

Diabetic retinopathy is a condition that may occur in people who have diabetes. It causes progressive damage to the retina, the light-sensitive lining at the back of the eye. It is a serious sight-threatening complication of diabetes. The Diabetic retinopathy is classified into five stages. If the retina is normal, it is classified as no DR. If there is a tiny bulges and a small amounts of leakage of blood vessel, it leads mild DR. If the retina swells, it leads to moderate DR. In severe DR stage, the blood vessels become even more blocked. In proliferate DR, a new blood vessels grow in the retinas and into the gel-like fluid that fills the eyes. CNN-based model DiaNet consists of a DenseNet-121 backbone and a few additional layers consisting of a pair of pooling layers followed by three composite layers each consisting of primarily a sequence of batch normalization (BN), dropout (Dr), linear (Lin), and a ReLU activation layer [1].

Automatic severity classification of DR based on DenseNet169 uses Gaussian filter for the removal of the black side borders and to analyze data distribution among different classes, Principle Component Analysis (PCA) was applied [2]. The features are filtered and extracted using the convolutional layer, the max-pooling layer, and the ave-pooling layer and the detection methods [3-5] based on VGG-16 and Spatial Pooling Layer (SPP) employed to overcome content loss and affects the accuracy and provides a fixed output. A computer-aided diagnostic (CAD) system is used to detect the non-proliferative diabetic retinopathy in optical coherence tomography images using CNN and SVM [6]. A novel convolutional neural network model with the Siamese-like architecture which is trained with a transfer learning technique, Inspection V3 can be customized to perform different image classification tasks. In addition to that Inception@4 model is formed by concatenating the features of the global average pooling layers from the four Inception-V3 models into one vector [7,8]. Diabetic retinopathy was also detected and measured using efficient convolutional neural networks and contrast limited adaptive histogram equalization(CLAHE) which proves Efficient-B5 is capable of producing results with better accuracy [9].

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁙ Impact Factor 8.102 ⁙ Peer-reviewed / Refereed journal ⁙ Vol. 12, Issue 6, June 2023
DOI: 10.17148/IJARCCE.2023.12675

II. DIFFERENT DIABETIC RETINOPATHY DETECTION METHODS

The author proposed a Deep Learning based architecture to diagnose diabetes using a relatively small-sized dataset, develop a multi-stage convolutional neural network (CNN)-based model DiaNet that takes a retinal image as input. The proposed approach uses two datasets to achieve state-of-the-art in detecting diabetes from retinal images i.e. the QBB retina-image dataset and the EyePACS dataset. A multiple pre-processing and data augmentation techniques are applied to increase the robustness. In pre-processing stage, first extracted the circular region from the image and then resized it. In data augmentation, the random horizontal flip, brightness and contrast perturbation are applied. On retinal images from the QBB and the EyePACS dataset, the same pre-processing and augmentation processes are executed. DenseNet 121 with a random initialization is the top-left model (M0). M1 is created by training M0 on the ImageNet dataset and is capable of classifying images. We get M2, a model with the ability to understand retinal images, by adding a few extra layers and honing it on the EyePACS retinal image dataset. Finally, our suggested model (M3) for diabetes detection from retinal pictures is created by optimising M2 using the QBB dataset. The DenseNet 121 backbone serves as the foundation for DiaNet's proposed architecture, which also includes a pair of pooling layers, three composite layers, each of which essentially comprises of batch normalisation (BN), dropout (Dr), linear (Lin), and a ReLU activation layer. One neuron with the expected label (diabetic/non-diabetic) is present in the final layer. For quantitative performance reporting, mean accuracy, sensitivity, specificity, precision, F1 score and AUC ROC are used. DiaNet fine-tuned on the EyePACS and the QBB dataset performed best with an 84.47% mean accuracy. The highest F1 score of 84.71 and 79.47 from DiaNet and QBBNet, respectively. The highest accuracy of 84.47 and 80.65 from DiaNet and QBBNet were achieved respectively. The highest AUC for the DiaNet and QBBNet was 84.4 and 83.10, respectively[1].

In paper[2], the authors suggested the DenseNet169 encoder for a visual embedding construction. In addition to that, Convolutional Block Attention Module (CBAM) is placed at the top of the encoder to reinforce the discriminative power. The main aim of DR detection is to classify the disease severity based on the probability value produced that an image is located in one of the five clusters: No DR, Mild, Moderate, Severe and Proliferate DR. Approximately, over 13,000 datasets were collected from APTOS dataset but had access to use only 3662 images. The retinal images are passed into the pre-processing stage which undergoes quality enhancement of the image. Gaussian filter is applied for the removal of the black side borders and provided the clarity of visual biomarkers. By using Bilinear interpolation, the image was resized to 256×256. The datasets produced severe class imbalance belonging to normal, mild, moderate, severe, and proliferative DR. Therefore, to analyze data distribution among different classes, Principle Component Analysis (PCA) followed by t-Distributed Stochastic Neighbor Embedding (t-SNE) algorithm was applied. To reduce overfitting and improvise the model's generalizability, random horizontal, vertical flipping, and rotation were applied. DenseNet is proposed as the backbone network which is used as a feature extractor for the input image and CBAM is used for the refinement of the features. Four 1×1 convolution layer, four 3×3 convolution layer and one 3×3 max pooling layer make up the DenseNet integrated with the Convolutional Block Attention Module. For pooling, Transition Block (TB) was integrated, consisting of batch normalization, 1×1 convolution, and 2×2 average pooling. The different positions were tried for CBAM in the modified DenseNet and the best performance of enhancing the model's representational power without increasing the complexity was achieved by positioning CBAM on top of the convolutional encoder. Therefore, our model established low training time (9 seconds/epoch) and relatively high inference speed (1.166 seconds/32 images). The network has achieved 97% accuracy on the binary classification task and 82% accuracy for severity grading. The significance of the proposed network is that it classifies the severity level of DR efficiently while reducing the time and space complexity.

The authors used Deep Convolutional Neural Network to detect Microaneurysms and Hard Exudates of Diabetic Retinopathy[3]. This model is proposed to increase the efficiency of Feature Extraction. The average pooling layer can improve the detection of microaneurysms and the detection of hard exudates can be done using max pooling. The performance of the classification is highly dependent on the feature extraction algorithms and hence the symmetric convolutional network is used to enhance the model's capacity for the location and classification of targets. The datasets are collected from DIARETDB1 which comprises of 89 colour fundus images of size 1500×1152 containing digital images with exudates, fragile exudates microaneurysms and haemorrhages. The pre-processing stage consists of three stages such as feature filtering, feature extraction, and classification stage in which the image is classified into red, blue and green channels. When compared with the red and blue channels, the green channel is selected as input for the network because it consists of more information and can display different kinds of lesions. Following that, the features are filtered and extracted using the convolutional layer, the max-pooling layer, and the ave-pooling layer. After that, the SoftMax classifier is used to classify targets. The proposed method shows better accuracy, sensitivity and specificity when compared to other methods. Therefore, the proposed method has achieved an accuracy of 93.2% using max pooling and 93.6% using average pooling.

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⋇ Impact Factor 8.102 ⋇ Peer-reviewed / Refereed journal ⋇ Vol. 12, Issue 6, June 2023
DOI: 10.17148/IJARCCE.2023.12675

The authors proposed a accurate detection of non-proliferative diabetic retinopathy in optical coherence tomography images be achieved using deep learning including convolutional neural networks (CNNs). In this study[4], a computer-aided diagnostic (CAD) system using CNNs is proposed for the early diagnosis of non-proliferative DR (NPDR). For the optical coherence tomography (OCT) imaging modality, the proposed system is created. OCT is superior than fundus imaging because it offers quantitative evaluations, can capture depth, is less expensive, and enables human bias-free monitoring of changes when compared to fundus imaging. The technique begins by utilising unsupervised learning to approximately separate the 12 layers of the retina and localise the fovea. The system consist of pre-processing stage that comprises patch extraction and alignment, rough segmentation of the retinal layers, and fovea detection; a CNN-based feature extraction stage; and a classification stage. Data used for training and testing of the CAD system were obtained using a clinical OCT scanner, Cirrus HD-OCT 5000 (Carl Zeiss Meditec, Dublin, California). The foveal coordinate localization applies the Median filter to remove impulse noise. Pre-processing also eliminates irrelevant information. Therefore, it increases the speed and efficiency of the system. The AlexNet CNN is used with an input patch of size 227×227×3. The patches provided in the pre-processing stage are sent to a convolutional neural network (CNN) from which features are extracted and sent to a support vector machine (SVM) to perform classification. Pre-training of the CNN used for transfer learning was performed on a subset of the large ImageNet database containing 1.2 million real images and 1000 object categories. In order to reach highest accuracy, this paper results found to be 94 %.

The authors of the research[5] used the VGG-NiN model to classify the five stages of Diabetic Retinopathy. VGG-16, Spatial Pooling Layer (SPP) and network-in-network are stacked to form a VGG-NiN model. A total of 88,702 images were obtained from the EyePacs dataset, out of which 35,126 are labelled images and 53,575 images are not labelled. Only labelled images are used in this study. In the pre-processing stage, the original size of the image which is of the size 3888×2592 is resized to 1349×1024 which in turn helps to avoid the loss of features from the image. The resized images are then cropped to a fixed size of 1024×1024. The distribution of the dataset into training, testing and validation is 64%, 16% and 20% respectively. Image augmentation is done by using the Keras data Generator which automatically augments the data. VGG16 requires RGB image size of 224×224. The cropped image which is the size 1024×1024 which is once again reduced to 224×224 to match the VGG16 input size. This may lead to content loss and affects the accuracy. The Spatial Pooling Layer (SPP) is thus employed to overcome these problems and provides a fixed output. The image is then passed to a series of convolution layers containing filters of 3×3 receptive fields which is then followed by a block of three fully connected layers. The parametric relu (PRelu) reduces the model overfitting. The softmax layer at the last is used as an activation for classification and Categorical cross-entropy is applied as loss function. The Area Under the Curve (AUC) of the proposed model is 0.95 and the model has also provided lower computational results.

The author[6] presented a hybrid learning model named Diabetic Retinopathy classification by analySing retinal Images (DRISTI) which is composed of VGG16 and capsule network to classify the stages of Diabetic Retinopathy. In this paper, two types of classification of DR are performed such as two stage and five stage classification. In two stage classification, the images are classified as DR and Normal. The five stage classification includes the classified images as No DR, Mild DR, Moderate DR, Severe DR and Proliferate DR. The datasets are obtained from IDRID, DIARETDB1, DIARETDB0, STARE, DRIVE, APTOS and MESSIDOR-2 datasets in which 65% are training images and 35% are testing images. MESSIDOR-2 dataset is used for five class classification testing and rest of the datasets are used for two class classification. 29101 images and 21024 images are augmented for five class and two class classification respectively by changing the rotation range, width and height shift range. The VGG16 model outperforms in two class classification when compared with ResNet50, Inception3, Xception and Capsule Network models. Similarly, for five class classification, the Xception network outperforms the other models. The Xception network is based on the depth wise separable convolution layers. The VGG16 model has an input size of 224×224 during training, The filters with a small receptive size of 3×3 with a stride of 1 pixel is used when the images are passed through various network layers. The spatial pooling is carried out by five max-pooling layers. A stack of fully connected layers are followed by three fully connected layers. The hybrid model results in overall validation accuracy of 82.06% and testing accuracy of 75.81% for five class and 96.24% of validation accuracy and 95.55% of testing accuracy for two class classification. To demonstrate the efficiency of DRISTI, cross-dataset and mixed dataset testing is also performed.

The author proposed an Automated Diabetic Retinopathy Detection Based on a novel convolutional neural network model with the Siamese-like architecture which is trained with a transfer learning technique[7]. Fundus image data set gathered from public resources is pre-processed and augmented before being used to train our model. The image dataset used is obtained from the website of Kaggle diabetic retinopathy competition provided by EyePACS. The processing method is referred to the algorithm proposed by Graham. The image data set has the drawback of being too small for a deep learning model to accurately tackle a medical image recognition challenge. To enhance the proposed model's ability to generalise, additional image augmentation processes are added to the pre-processing steps and applied to the data set. The deep learning model sends the left and right fundus images into the Siamese-like blocks after accepting the left and right fundus

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁘ Impact Factor 8.102 ⁘ Peer-reviewed / Refereed journal ⁘ Vol. 12, Issue 6, June 2023
DOI: 10.17148/IJARCCE.2023.12675

images as inputs. The information from the two eyes is collected into a fully connected layer, and finally the model outputs the diagnosis result of each eye, respectively. By adopting the transfer learning method, Inception V3 can be customized to perform different image classification tasks. The original Rectified Linear Unit (ReLU) activation function used in Inception V3 is replaced by the Leaky ReLU function. The leaky rate is set to be 0.1. Design of a binocular network with a Siamese structure capable of synchronously receiving binocular fundus images, collecting their features, and creating prediction of each eye. By loading weights of Inception V3 blocks pre-trained on the ImageNet dataset, the model will have a better initialization of weights before running gradient optimization. The optimizer used in this work is called Adam. The proposed binocular model for RDR detection, as well as a similar model for the original five-level DR classification task, are trained on a single server with NVIDIA GeForce GTX1080TI graphics cards. The evaluation result shows that the proposed binocular model achieves high performance with an AUC of 0.951 and a sensitivity of 82.2% in the high-specificity operating point and a specificity of 70.7% in the high-sensitivity operating point.

The author proposed an Automatic Diabetic Retinopathy Diagnosis Using Adaptive Fine-Tuned Convolutional Neural Network. Transfer learning was employed to overcome the overfitting problem[8]. VGGNet, ResNet, and DPN, trained on ImageNet dataset are usually employed for transfer learning. A pre-trained CNN took images of the fundus as input and gave a normal or DR grade as output. It consisted of CONV layers fine-tuned using ROIs extracted from fundus images, an FC layer created using PCA, and a classifier layer. The model was evaluated by employing EyePACS and Messidor because both have image level annotations for DR grading. The EyePACS dataset has five classes: normal, mild DR, moderate DR, severe DR, and proliferative DR. Messidor has four classes: normal and three DR grades based on the number of MAs and HEs. E-optha contains two subsets of color fundus images for EX and MA. To avoid overfitting, the training examples were extended so that the inner structures of the ROIs reflect the actual structures of the fundus images. VGG19, ResNet152, and DPN107, pre-trained on the ImageNet dataset. The VGGNet and ResNet architecture achieved an excellent accuracy on ImageNet. The dual path network (DPN) is based on the philosophy of ResNet. In the first stage the low-level layers are fine-tuned using extracted ROIs. In the second stage the FC layers are removed and added a new FC layer using the PCA technique. First-layer CONV filters can be directly re-initialized using lesion regions of ROIs extracted from fundus images. After the first reinitialization of the CONV layer, the pre-trained model was fine-tuned using the extracted ROIs to match the lesion and normal structures of the retinal fundus images. An adaptive max pooling layer was introduced before the FC layers to avoid problems. For fine-tuning, the cross-entropy loss, Adam optimizer with 0.001 learning rate, 3000 batch size, and 21 epochs are used. For the classification layer, three tree-based classifiers are considered, namely decision tree (DT), random forests (RF) and gradient boosting (GB). Accuracy (ACC), area under ROC curve (AUC), sensitivity (SE), specificity (SP), precision (PR), recall (RC), F1-score, and kappa were used to evaluate the performance of DR diagnosis systems. The nonparametric Mann–Whitney–Wilcoxon (WMW) test is to verify whether the difference among ResNet152, VGG19, and DPN107 was statically significant. ResNet152 with re-initialized CONV1 and GB layer (ResNetGB) achieved the best results.

Using the Inception V3 and Inception@4 model, the authors suggested to detect and diagnose the Diabetic Retinopathy. The datasets used in this work is 4476 images which are moderate in size. The data pre-processing stage includes Size Normalization, Shape Normalization and Colour Normalization. To overcome the shortage of datasets, data augmentation techniques such as Augmentator software package is used which flips the image horizontally, vertically and randomly rotates, zooms in or out and distorts the image. In this paper[9], Inception V3 is used as the base model. The different kernel sizes are combined with the convolution layers of the Inception A/B/C/D/E blocks which forms the network structure of Inception V3 network. The Inception@4 model is formed by concatenating the features of the global average pooling layers from the four Inception-V3 models into one vector and fed to a softmax output layer. Transfer learning is another technique which is used to mitigate the data insufficiency problem. When compared to Resnet-18, Resnet-101 and VGG-19, the suggested Inception V3 and Inception@4 model has produced a classification accuracy of 88.35% and 88.72% respectively. The models are deployed on a cloud computing platform for diagnosis, then the output is displayed in the web page and the users from different hospitals are provided with user interface via network. The users can upload the images in the system and obtain the results of the diagnosis. The images uploaded by the users are once again analysed by the ophthalmologists. The diagnose accuracy results between the system and the ophthalmologists has reached 91.8% which demonstrates the reliability of the system.

The authors proposed an automated detection and monitoring of Diabetic Retinopathy in this study[10]. In order to improve the quality of the image and to equalize the intensities of the image uniformly, Contrast Limited Adaptive Histogram Equalization (CLAHE) is proposed as the pre-processing step. Following that, for the classification of the images such as No DR, Mild DR, Moderate DR, Severe DR and Proliferate DR, Efficient-B5 architecture is being proposed. The datasets used for this project are obtained from Messidor-2 and IDRiD. The Messidor-2 dataset consists of 1,748 retinal colour images of 874 subjects and the IDRiD test set includes 516 images. The CLAHE method is very effective because it helps in increasing the image contrast effect which helps to detect and classify DR more efficiently when compare to other methods. This method is used for the amplification of vessels in the retinal fundus images which is beneficial in diagnosis of DR. The two main parameters for CLAHE are Block Size (BS) and Clip Limit (CL) which is used for controlling the quality of the image. Therefore by this method, the vessels in the retinal image are displayed better and it is easy for the network to learn. When compared to other popular CNNs, the EfficientNet is advantageous in reducing the number of parameters and FLOPS and thus increasing the speed and accuracy. Firstly, the Efficient-B0 network was applied and the three dimensions such as Depth, Width and Resolution was scaled. Although it produces uniform scaling, it has resulted in less accuracy and efficiency. After performing experiments with EfficientNets-B0 to B4, it is observed that these networks reduce image resolution as a result of which the useful information of the image is lost. Therefore, Efficient-B5 network proves to be capable of producing results with better accuracy. This network is pre-trained on the ImageNet dataset which helps to save time and acquire better accuracy with the limited number of datasets without performing data augmentation. When the model is trained on a mixture of two datasets Messidor-2 and IDRiD and evaluation is performed on the Messidor dataset. The Area Under the Curve (AUC) is enhanced from 0.936 to 0.945 and achieved sensitivity of 92%.Also, for the further evaluation of the performance of the model, it is trained on a mixture of two datasets, Messidor-2 and Messidor, and evaluated on the IDRiD dataset. The AUC has been enhanced from 0.796 to 0.932 and obtained sensitivity of 93%.

The authors proposed a deep learning ensemble approach for Diabetic Retinopathy Detection in this study[11]. It uses end-to-end deep ensemble networks to detect all stages of DR, even the moderate stage. Kaggle dataset of retinal images are used for input data, which include 35126 colour fundus pictures, each 3888 x 2951 pixels in size. It includes pictures from five classes based on the degree of diabetic retinopathy (DR) namely no DR, normal, mild, moderate, severe and Proliferative DR. Deep network classification bias results from training on imbalanced data. In the first preprocessing step, the input image is resized. The dataset is split into three segments: training, testing, and validation sets, with respective ratios of 64%, 20%, and 16%. The validation set is utilised during training to assess and minimise over-fitting. Ensemble method is a meta-algorithms that combine multiple machine learning techniques into a single predictive model. The proposed approach ensembles the five deep CNN models Resnet50, Inceptionv3, Xception, Dense - 121, Dense169. The proposed model is quantitatively evaluated by accuracy, sensitivity, specificity, precision, F1 measures, AUC (Area Under the Curve) and ROC (Receiver Operating Characteristics) as performance metrics. The higher the AUC score, the better the model, and vice versa. To show the effect of the imbalanced dataset the three different datasets are used : i) Imbalanced ii) Up Sample and iii) Down Sample Dataset. The parameters such as Recall, Precision, Specificity, F1-score, and ROC-curve are used to provide unbiased results. The achieved accuracy, recall, specificity, precision, and F1-score are 80.8%, 51.5%, 86.72%, 63.85% and 53.74% respectively and also classified the ROC. A CNN ensemble-based framework to detect and classify the DR's different stages in color fundus images were proposed. The largest publicly available dataset of fundus images (Kaggle dataset) are used to train and evaluate the model.

The authors[12] use EfficientB3 CNN architecture to detect and classify the five stages of Diabetic Retinopathy. The input image size has been chosen as 300×300×3. Following that, 2502 images were collected from APTOS dataset and were divided into training and validation datasets. The retinal images were preprocessed using Gaussian blur and circle cropping methods. To reduce the model's training time, transfer learning approach was used. A global average pooling layer and a dense layer with sigmoid activation were placed after the EfficientNetB3 model. With a set of specified scaling coefficients, the EfficientNetB3 scaling algorithm consistently scales network width, depth, and resolution. In the training set, the performance of the proposed model was 99.18% categorically accurate, 99.19% precise, 99.31% recall, and 99.25 F1- score and in the validation set, it was 75.68% categorically accurate, 76.31% precise, 75.00% recall, and 75.95 F1- score.

Using the Low-Dimensional Spatial-Spectral Matrix Representation, the authors[13] proposed to segment lesion in multispectral images for the detection of Diabetic Retinopathy. By utilizing both the spatial and spectral properties of MSI retinal images, a learning-based strategy is proposed in this paper. 50 sequences of MSI images were obtained from Annidis RHA (Annidis Health Corporation, Ottawa, Canada). Out of 50 sequences, 40 sequences are unhealthy images and 10 sequences are healthy images. The pixel size of the image is 2048×2048. The colour fundus images are obtained from MESSIDOR dataset. The image is resized to 800×600. The sampled pixels are extracted from the MSI image sequence. Using the Local Binary Pattern (LBP) operator, the sampled pixels are converted into LBP featured images. Then, local LBP image patches are vectorized to produce LBP feature vectors. The ultimate arrangement of the LBP feature vectors is into a single matrix. The obtained output is compared with the manual annotation of the ophthalmologists. The accuracy when compared with the ophthalmologist1 is 0.911 and with ophthalmologist2 is 0.903.

The author[14] presented an ensemble of deep convolutional neural network (CNN) models utilising fundus images to accurately detect and grade DR. The datasets are collected from Diabetic Retinopathy Database (DIARETDB), Structured Analysis of Retina (STARE), Retinopathy Online Challenge (ROC) and Kaggle. At the first stage, each input image is

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 6, June 2023
DOI: 10.17148/IJARCCE.2023.12675

split into four patches and sent to pre-trained CNN models (InceptionV3, Xception) for training. Prior knowledge is derived from the essential features in the shallow-dense layers of CNN models. The model is assisted in learning the important data from DR images by the merging of shallow and dense layer features. The fused probability vectors of the four patches are utilised to train a classifier based on an artificial neural network in the second stage. In the third stage, the outcomes of various CNN models are combined to produce the final result. The total classification accuracy of grading diabetic retinopathy is increased by using an ensemble strategy of multiple-stage deep learning model. With fivefold cross-validation, this ensemble classifier achieves a classification accuracy of 96.2%.

To diagnose different stages of Diabetic Retinopathy, the author[15] built a new DIY smartphone enabled camera. Preprocessing is first carried out using Contrast Limited Adaptive Histogram Equalization (CLAHE) and green channel transformation. Additionally, the segmentation process begins with the segmentation of the optic disc by the watershed transform and the segmentation of abnormalities (exudates, microaneurysms, haemorrhages, and IRMA) by the Triplet halfbandfilter bank (THFB). Then, using the Haralick and Anisotropic Dual Tree Complex Wavelet Transform (ADTCWT) techniques, the various features are extracted. The finest features are selected from the mined features by the life choice-based optimizer (LCBO) algorithm. In order to categorise the severity level as mild DR, severe DR, normal, moderate DR, and proliferative DR, the selected features are then applied to the optimised hybrid ML (machine learning) classifier with the combination of NN and DCNN (Deep Convolutional Neural Network). The proposed method outperformed other schemes with accuracy levels of 99% for the EyePacs dataset and 98.9% for the APOTS 2019 dataset.

III. COMPARISON OF DIFFERENT DIABETIC RETINOPATHY DETECTION TECHNIQUES

This section examines various stages of Diabetic Retinopathy detection and classification methods. The author's name, publication year, methods used, precision of detection and classification, dataset size, disease kind, and parameters are all taken into consideration in this survey. Table 1 summarises the different stages of Diabetic Retinopathy detection using deep learning techniques.

TABLE I **COMPARISON OF DIFFERENT DIABETIC RETINOPATHY DETECTION TECHNIQUES**

S.No.	Author	Year	Methodology	Dataset	Image count	Parameter
1.	Mohammad Tariqul Islam et al.	2020	Dianet based CNN model	EyePACS and QBB datasets	1852 images from QBB dataset and 80000 images from EyePACS dataset	achieves an accuracy of 84%
2.	Mohamed M. Farag et al.	2021	Gaussian filter, Bilinear interpolation, Linear classification, DenseNet169 and Convolutional Block Attention Module (CBAM)	APTOS dataset	13000 images	achieved 97% accuracy on the binary classification task and 82% accuracy for severity grading
3.	Tieyuan Liu et al.	2022	Deep symmetric Convolutional Neural Network	DIARETD B1 dataset	Trained and tested on the public database DIARETDB1 which contain 89 fundus colour images	achieved an accuracy of 93.2% using max pooling and 93.6% using average pooling
4.	Mohammed Ghazal et al.	2020	Computer-Aided Diagnostic(CAD) system for early diagnosis of NPDR, Median filter to remove impulse noise, Transfer learning,	Optical Coherence Tomography (OCT) dataset	83416 training images and 968 test images	Overall accuracy of 94% was achieved

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 12, Issue 6, June 2023

DOI: 10.17148/IJARCCE.2023.12675

			Support Vector Machine(SVM) for classification, AlexNet			
5.	Zubair Khan et al.	2021	Keras date generator for agumentation. VGG16, Spacial Pyramid Pooling layer (SPP) and Network-in-Network (NiN) are combined to form VGG-NiN architecture	EyePACS dataset	A total of 88702 images were obtained, out of which 35126 are labelled images and 53575 images are not labelled	Area Under the Curve (AUC) of the proposed model is 0.95
6.	Gaurav Kumar et al.	2021	DRISTI is a hybrid Deep Learning model composed of VGG16 and Capsule network	IDRID, DIARETD B1, DIARETD B0, STARE, DRIVE, APTOS and MESSIDO R-2 datasets	30% images from each dataset were taken and splitted into 65% training and 35% testing images	validation accuracy of 82.06% and testing accuracy of 75.81% for five class and 96.24% of validation accuracy and 95.55% of testing accuracy for two class classification.
7.	Xianglong Zeng et al.	2019	Binocular Siamese-Like CNN, Inception V3 for classification	EyePACS dataset	28104 training images and 7024 test images	AUC of 0.951 and a sensitivity of 82.2% in the high-specificity operating point and a specificity of 70.7% in the high-sensitivity operating point
8.	Fahman Saeed et al.	2021	Adaptive Fine-Tuned Convolutional Neural Network, Transfer learning to overcome the overfitting problem	EyePACS and Messidor datasets	Out of 88702 colour retinal fundus images, 35126 images are used for training and 53576 images are used for testing	achieved above 95% SP, SE, kappa, AUC and ACC
9.	Zhentao Gao et al.	2018	Inception V3 is a base model and the characteristics of the global average pooling layers are concatenated to create the Inception@4 model. Transfer learning is used to overcome data insufficiency problem	DIARETD B0/B1, MESSIDO R, DRIVE, STARE, REVIEW, Kaggle and E-ophtha datasets	A total of 4476 images	The accuracy of the Inception V3 and Inception@4 models is 88.35% and 88.72% respectively is higher when compared to ResNet-18, ResNet101 and VGG19

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁙ Impact Factor 8.102 ⁙ Peer-reviewed / Refereed journal ⁙ Vol. 12, Issue 6, June 2023
DOI: 10.17148/IJARCCE.2023.12675

10.	Asra Momeni Pour et al.	2020	Contrast Limited Adaptive Histogram Equalization (CLAHE), Efficient-B5 architecture	Messidor-2 and IDRiD datasets	1748 and 516 retinal images from Messidor-2 and IDRiD respectively	The AUC of the Messidor and the IDRiD are 0.936 to 0.945 and 0.796 to 0.932 with sensitivity of 92% and 93% respectively
11.	Sehrish Qummar et al.	2019	The Ensemble model is composed of Resnet50, Inceptionv3, Xception, Dense - 121, Dense169	Kaggle dataset	35126 colour fundus images which is divided into 65% training, 20% testing and 16% validation sets	Accuracy, recall, specificity, precision, and F1-score were all obtained, with respective values of 80.8%, 51.5%, 86.72%, 63.85%, and 53.74%.
12.	Praveen B et al.	2022	Gaussian blur and circle cropping methods	APTOS 2019 dataset	2502 images	99.18% categorical accuracy in training set and 75.68% categorical accuracy in validation set
13.	Yunlong He et al.	2020	Low-Dimensional Spatial-Spectral Matrix Representation, Local Binary Pattern(LBP) Operator	MESSIDOR dataset	Out of 50 sequences 40 sequences are un healthy images and 10 sequences are healthy images	The accuracy of the obtained output when compared with the results of the ophthalmologist 1&2 are 0.911 and 0.903 respectively
14.	V. Deepa et al.	2022	Inception V3 and Xception models are used for training, SVM is used to classify the features	DIARETDB, STARE, ROC and Kaggle datasets		overall accuracy is 96.2%
15.	Shubhi Gupta et al.	2022	Contrast Limited Adaptive Histogram Equalization (CLAHE) and green channel transformation as preprocessing stage. Segmentation of optic disc and abnormalities are performed using Watershed Transform and Triplet Half Band Filter bank respectively. Feature extraction is done using Haralick and anisotropic Dual Tree Complex wavelet Transform (ADTCWT). Hybrid classifier is composed of NN and DCNN architectures	EyePACS and APTOS datasets	35126 images and 3648 fundus images were collected from EyePACS and APTOS dataset. The datasets are divided into 70% training and 30% testing images.	The classification accuracy of EyePACS and APTOS datasets are 99% and 98.9% respectively

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⪤ Impact Factor 8.102 ⪤ Peer-reviewed / Refereed journal ⪤ Vol. 12, Issue 6, June 2023
DOI: 10.17148/IJARCCE.2023.12675

IV. CONCLUSION

Diabetic retinopathy, a prominent side effect of diabetes mellitus, involves slow retinal degeneration and can potentially result in blindness. Automated screening approaches expedite patient care by lowering the time required for diagnosis. This saves time and money for ophthalmologists. The type of lesions that occur on the retina determines the stage of DR. The most recent deep learning-based automated approaches for identifying and categorising diabetic retinopathy were studied in this paper. CNN has been used in the majority of research to categorise and detect DR images because to its effectiveness. Furthermore, the benefits and drawbacks of various techniques are explored in order to offer future directions for improving the performance of deep learning-based automated techniques. Furthermore, this work has discussed the practical ways that can be used to identify and categorise diabetic retinopathy using deep learning.
.

REFERENCES

[1] Mohammad Tariqul Islam, Hamada R. H. Al-Absi, Essam A. Ruagh, and Tanvir Alam "DiaNet: A Deep Learning Based Architecture to Diagnose Diabetes Using Retinal Images Only(2020)", Volume 9, pp. 15687-15695
[2] Mohamed M. Farag, Mariam Fouad, and Amr T. Abdel-Hamid "Automatic Severity Classification of Diabetic Retinopathy Based on DenseNet and Convolutional Block Attention Module(2022)", Volume 10, pp. 38299-38308
[3] Tieyuan Liu, Yi Chen , Hongjie Shen, Rupeng Zhou, Meng Zhang, Tonglai Liu, And Jin Liu "A Novel Diabetic Retinopathy Detection Approach Based on Deep Symmetric Convolutional Neural Network(2021)", Volume 9, pp. 160552-160558
[4] Mohammed Ghazal, Samr Samir Ali, Ali H. Mahmoud, Ahmed M. Shalaby, And Ayman El-Baz "Accurate Detection of Non-Proliferative Diabetic Retinopathy in Optical Coherence Tomography Images Using Convolutional Neural Networks(2020)", Volume 8, pp. 34387-34397
[5] Zubair Khan, Fiaz Gul Khan, Ahmad Khan, Zia Ur Rehman, Sajid Shah , Sehrish Qummar , Farman Ali and Sangheon Pack "Diabetic Retinopathy Detection Using VGG-NIN a Deep Learning Architecture(2021)", Volume 9,pp. 61408-61416
[6] Gaurav Kumar, Shraban Chatterjee, Chiranjoy Chattopadhyay "DRISTI: a hybrid deep neural network for diabetic retinopathy diagnosis(2021)", 15:1679-1686
[7] Xianglong Zeng, Haiquan Chen, Yuan Luo and Wenbin Ye "Automated Diabetic Retinopathy Detection Based on Binocular Siamese-Like Convolutional Neural Network(2019)", Volume 7, pp. 30744-30753
[8] Fahman Saeed , Muhammad Hussain and Hatim A. Aboalsamh "Automatic Diabetic Retinopathy Diagnosis Using Adaptive Fine-Tuned Convolutional Neural Network(2021)", Volume 9, pp. 41344-41359
[9] Zhentao Gao, Jie Li, Jixiang Guo1, Yuanyuan Chen 1,Zhang Yi 1, and Jie Zhong "Diagnosis of Diabetic Retinopathy Using Deep Neural Networks(2018)", Volume 7, pp. 3360-3370
[10] Asra Momeni Pour, Hadi Seyedarabi, Seyed Hassan Abbasi Jahromi, And Alireza Javadzadeh "Automatic Detection and Monitoring of Diabetic Retinopathy Using Efficient Convolutional Neural Networks and Contrast Limited Adaptive Histogram Equalization(2020)", VOLUME 8, pp. 136668-136673
[11] Sehrish Qummar, Fiaz Gul Khan , Sajid Shah ,Ahmad Khan, Shahaboddin Shamshirband, Zia Ur Rehman, Iftikhar Ahmed Khan and Waqas Jadoon "A Deep Learning Ensemble Approach for Diabetic Retinopathy Detection(2019)", Volume 7, pp. 150530-150539
[12] Praveen B, Guduru Mahesh Babu, Parvatareddy Ramakrishnareddy, Mani A "Diabetic Retinopathy Detection using Pre-trained EfficientNetB3 Model(2022)", ISSN: 2278-0181, Volume 11 issue 06, June, pp. 250-255
[13] Yunlong He , Wanzhen Jiao, Yunfeng Shi, Jian Lian , Bojun Zhao, Wei Zou, Yuemin Zhu, and Yuanjie Zheng "Segmenting Diabetic Retinopathy Lesions in Multispectral Images Using Low-Dimensional Spatial-Spectral Matrix Representation(2020)", Volume 24, pp. 495-502
[14] V. Deepa, C. Sathish Kumar, Thomas Cherian "Ensemble of multi-stage deep convolutional neural networks for automated grading of diabetic retinopathy using image patches(2022)", pp. 6255-6265
[15] Shubhi Gupta, Sanjeev Thakur, Ashutosh Gupta "Optimized hybrid machine learning approach for smartphone based diabetic retinopathy detection(2022)", pp. 14475-14501