



Early Screening and Diagnosis of Diabetic Retinopathy

Rachit Katiyar*

Department of Chemical Engineering, Computer Science Engineering Rajalakshmi Institute
of Technology, Chennai, India

*Corresponding Author's E-mail: revanth.s.b.2020.cse@ritchennai.edu.in

Received: 10-May-2024; Manuscript No: irjesti-24-134664; **Editor assigned:** 13-May-2024; Pre-QC No: irjesti-24-134664 (PQ); **Reviewed:** 27-May-2024; QC No: irjesti-24-134664; **Revised:** 09-Jan-2025; Manuscript No: irjesti-24-134664 (R); **Published:** 16-Jan-2025, DOI: 10.14303/2315-5663.2025.99

Abstract

This paper aims to review and propose some techniques for the early detection of diabetic retinopathy. Diabetic retinopathy is an eye complication that mostly happens in older age people which are affected with diabetes. Diabetes occurs when there is a high sugar level in our blood. When a person is affected with diabetes then there is a huge chance for diabetic retinopathy. In diabetic retinopathy, light sensitive tissue of the blood vessels at the back side of retina gets affected by diabetes. In reversible Vision loss, blindness may occur when necessary treatment has not been taken on time. Early detection of diabetic retinopathy may lead to stop these complications. Various deep learning models have been used to detect the diabetic retinopathy at early stage which are Vgg19, Inception, ResNet50, Vision transformer, AlexNet, DenseNet. In this paper, the author proposes a transfer learning model of CNN by using the Resnet50 architecture in which fully connected layer used to detect high level features. The preprocessing of the fundus images done through Contrast Limited Histogram Equalization (CLAHE), RGB to Grayscale conversion, RGB to HSI conversion, remove the noise by applying gaussian filtering mean and median filtering, and augmented the image through geometric transformation and colour transformation. Segmentation of the given model have been done by a U-Net architecture and multilayer thresholding and adaptive thresholding. The given model extracts the feature of different lesions which are microaneurysms, hemorrhage, exudates by feature extraction of retinal image by applying morphological operation, thresholding, clustering. The proposed model uses ResNet50 model for classification of different severity levels such as normal, early diabetic retinopathy, mild NPDR, moderate NPDR, severe NPDR, PDR and neovascularization. The proposed model has accuracy of 96.56% for successfully classification of different severity levels of DR which is much improved as compared to other deep learning models present. ODIR dataset have been used here where evaluation parameters are accuracy, precision, specificity, sensitivity and F1 score.

Keywords: Diabetic retinopathy, Preprocessing, CNN, Diabetes, VGG19, ResNet50, AlexNet, DenseNet, Vision transformer, CLAHE, Filtering, Transfer learning, Colour conversion

INTRODUCTION

Diabetes mellitus is a group of metabolic diseases that affects one out of 10 people worldwide with increasing age (Akram MU et al., 2013). If diabetes is not controlled, it can lead to visual impairment, the most severe form is diabetic retinopathy (Albahli S et al., 2021). Diabetic retinopathy affects the blood vessels of the light-sensitive tissue on the retina (Ali A et al., 2020). The number of patients suffering from diabetes has increased in the past

few years, making diabetic retinopathy a more and more serious problem (Aras RA et al., 2016). In some cases, diabetes patients don't even realize that they have diabetic retinopathy because in its early stage, the symptoms are not visible and only mild vision problems occur (Bala R et al., 2014). However, it can lead to more severe eye problems and result in blindness (Dutta S et al., 2018). The only thing that we can do is to prevent it from becoming worse and causing blindness (İncir R et al., 2024). So, the most effective way to fight diabetic

retinopathy is by preventing it with early detection and curing the disease to prevent it from becoming worse (Jabbar MK et al., 2022). But early detection of diabetic retinopathy is difficult. Early detection of diabetic retinopathy can be done by looking at the digital image of the retina, by identifying micro aneurysms, hemorrhage and exudates which are the early symptoms of diabetic retinopathy (Kandel I et al., 2020). However, manual detection of these symptoms is time-consuming and requires the skill of experts in the field of ophthalmology (Karki SS et al., 2021). Therefore, a framework to automatically detect these symptoms is necessary to help by using the early detection of Dr. Sugihara and Zhang developed a system with fovea localization, optic disc removal, and image enhancement as a preprocessing stage to automatically detect hard exudates (Li Q et al., 2019). With fovea localization, the system can search for light reflections on the retina that work as the center of the fovea area (Lin CL et al., 2023). Hard exudates mostly occur in the area of the fovea, so with the knowledge of the center of the fovea from the location of the light reflection, the search area for hard exudates becomes more focused. The optic disc often becomes noise when detecting exudates, and hard exudates can also occur everywhere around the optic disc, so it is better to remove a big area around the optic disc when looking for hard exudates (Aatila M et al., 2021). After locating the fovea and removing the optic disc, the image of the retina

can be processed to decrease the development of hard exudates using image enhancement (Nahiduzzaman M et al., 2021). After the preprocessing stage, hard can be detected using thresholding and have an accuracy of 85% to 91%. Detection of all exudates is necessary for the early detection of DR because exudates are some early symptoms of diabetic retinopathy. In the next study, Sugihara and Zhang developed a method to detect all the symptoms using mathematical morphology and achieved greater accuracy than the previous study (Saeed F et al., 2021). An automatic detection system is necessary, but detection using digital image processing still has difficulty in achieving good accuracy for exudate detection (Saranya P, 2022). This difficulty can be seen from the overall results of exudate detection in some research methods and various databases (Shaukat N et al., 2023). A method using automatic detection is identification by extracting image features of exudates, giving these tasks to an interpretation method or intelligent system, and then the system makes a classification whether an extracted image feature is an exudate or non-exudate (Akram MU et al., 2014). Image processing of retinal fundus images is the primary source of research with an overall accuracy of 75.5% to 97.0%. However, research using a method of image processing with digital retinal photographs and then giving feature extraction to a diagnostic workstation has better results (**Figure 1**).

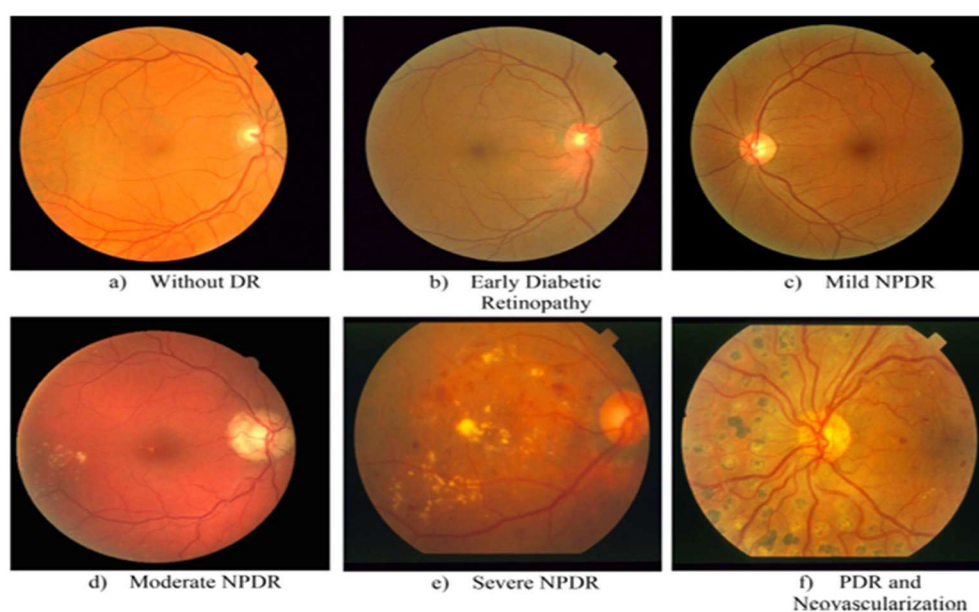


Figure 1. Steps of increasing DR.

Background

Diabetic retinopathy is a severe health condition that can cause blindness to patients (Zhou D et al., 2021). It is a slow progressing disease creating damage to retinal blood vessels, leading to diabetic macular edema and retinal ischemia. DMR is swelling in the macula and is the most common cause of visual loss in diabetes (Zhou Y et al., 2020). A long duration of diabetes is a major risk factor

for DMR. Retinal ischemia contributes to the development of new vessels and is the most common cause of severe visual loss in diabetic retinopathy. Prompt management can prevent or limit the loss of vision from DMR and retinal ischemia. There are effective treatments for diabetic retinopathy, such as laser therapy, to support an improving global detection and prevention of vision loss. Retinal photocoagulation has been shown to reduce moderate to severe visual loss by at least 50% using argon

green, krypton red, and diode laser. But no more than 5% of patients with proliferative diabetic retinopathy have severe visual loss compared with a third who develop the condition, indicating that many are developing retinal ischemia. Existing treatment for DMR or retinal ischemia involves low risk and less sight-threatening forms of photocoagulation with partial or full-thickness retinal destruction and no more than halting the advancement of the disease. Step one in the development of DMR and retinal ischemia treatment is classification, allowing identification of different levels of the disease to guide an appropriate intervention. This has delivered a widespread understanding that significant development of evidence-based treatment will need to be multi-centred and will require efficient evidence-based identification and selection of patients.

LITERATURE REVIEW

In this paper, the author proposed a Transformer model for classification of DR with the help of invasive fundus retinopathy. The author proposed a novel transformer model to obtain superior performance and accuracy with existing CNN models. In this model 224 x 224 size high resolution images are being used, which are captured by fundus cameras. For the segmentation of high-resolution images 224 x 224 Transfer learning model with Multiple Instance Learning (TMIL) were taken in use. To extract features from each image, vision transformer and GICB (Global Instance Computing Block) used to calculate inter instance feature. This model can load pretrained weights of transformer without affecting the performance of given model while using high resolution images. APTOS and MESSIDOR-1 dataset have been used to evaluate the classification report by using TMIL and reduced the interface time by 62%. The major drawback of this model is that exists a false negative result of 30%.

In this paper, the author proposed an extreme learning approach to classify diabetic retinopathy which occurs due to diabetes. To overcome diabetes at an early stage, the author used a hybrid CNN model for binary and multiclass DR classification with the help of Messidor-2 and Aptos-2019 datasets. At preprocessing stage, Ben graham approach have been used here after that to enhance brightness and contrast, CLAHE technique have been used here with lowering noise and distinguishing more features from existing fundus images. Here hybrid CNN singular value decomposition model has been use as it has lesser input feature for classifier to easily classify the features. Thus, the author proposed an ELM model for minimization of training time cost. This model produces an acc and recall score of 99.73% and 100% respectively for two class classification and 98.09% and 96.26% for multiclass classification respectively with the use of APTOS and Messidor-2 dataset which performed well than other existing CNN models.

In this proposed model, the author used a deep learning model which employed to automatic classification of diabetic retinopathy due to outstanding performance

over fundus images. This model needs a heavy amount of data which is unavailable in case of diabetic retinopathy and overfitting is also unavailable. This model employs a two-stage transfer learning method by developing a computer aided system with the help of pre-trained convolutional neural network model, region of interest ROI lesions and being used from annotated fundus images. In this model, high level features are captured by fully connected layer. By replacing FC layer with PCA and use it in unsupervised manner, discriminant feature was extracted. These steps tried to reduce model complexity and overfitting. After this a gradient boosting layer is used as classification layer. The proposed model has been tested over two datasets which are eyePACS and Messidor. The author also aided ResNET-152 with re-initialized CONV1 and GB layer ResNET-GB for best performance.

In this proposed model different methods were employed to categorize for diabetic retinopathy at early stage. The author uses Ada-boost technique and different interpretable deep learning models by combining to create the model. The output of the developed method can be given by weighted graph activation map that explains the regions of our eyes that are affected by diabetic retinopathy. To generate distinct transformation for each picture the author used expanded module and employs eight image modification techniques at preprocessing stage to classify diabetic retinopathy. for random alteration, and picture has been captured and classified using this model. By implementing many integrated models that ensemble the capability of all InceptionV3, ResNet152, and Inception-Resnet-V2, to minimize biasness in the model the author used the Ada-boost method. The CAMs of the individual models as well as the integrated model have been taken, together with the DR classified scores. Compared to previous deep learning models, this paper produces robust classification results which may result in more resilience and superior performance.

In this paper, the author suggests a multi-cell design that progressively raises the input image resolution and the deep neural network's depth, which extends training time and raises classification accuracy. Large local receptive fields are used to identify late-stage diseases because retinal images have extremely high resolution and small diseased tissue can only be recognized with large resolution images. To predict the label while anticipating the connection between images at different phases, author propose a Multi-task learning technique that combines regression and classification based on the experiments conducted on the collected dataset, our method ranks fourth among all state present in the art system with a Kappa score of 0.841 on the testing set. Moreover, our general framework for multicell multitask convolutional neural networks facilitates seamless integration with other deep neural network architectures.

In this paper the author developed an automated knowledge model to classify the key antecedent of

diabetic retinopathy. In this proposed model the author used three types of models which are Back Propagation Neural Network (BNN), CNN, Deep Neural Network (DNN). While testing stage with CPU trained neural network give very low accuracy and precision because of only one hidden layer. Whereas deep learning models perform well with given training data. The deep learning model has the capability to distinguish the feature as blood vessels, exudates, fluid drip, hemorrhage and micro aneurysms into different classes. This model has the capability to return weight which is very helpful to check severity level in patient eyes. The major challenge for this study to extract each feature for accurate threshold class. A threshold weighted fuzzy C - mean algorithm has been used here for classification. BNN and DNN model are trained for statistical image data. By hyper parameter adjustment, DNN provide us best accuracy, in CNN with the use of VGGNET model, it provides 72.5% accuracy whereas DNN provide 89.6% accuracy for training and 86.3% for training.

In this paper, the author developed a unique hybrid classifier-based approach to identifying retinal lesions. At the preprocessing stage, the author tried to extract the lesion, feature set formulation and classification. In preprocessing stage system eliminates background pixel, remove noise, and extract blood vessels optic disk from human retinal image. Lesion detection unit detected using filter banks. Every region can have different types of lesions. In this proposed model a feature set is developed using several descriptor intensity and statistics and shape which help us to add further categorization. to increase the classification accuracy of model the author proposed a hybrid classifier that is connected to the M-Medoid based modelling technique combining with Gaussian model as ensemble. To classification of severity the author uses fundus images and their performance metrics which include sensitivity, accuracy, specificity and receives operating characteristics curve. In this model provide an average accuracy of 98.52% and 94.98% for different datasets for image level and pixel level respectively.

In this proposed model, the author studied to detect blood vessels and classify haemorrhage and different phase of diabetic retinopathy into normal, NPDR, moderate diabetic retinopathy. To identify and measure the blood vessels and haemorrhage in the human retinal image provide a base for the categorization of various stages of diabetic retinopathy. With the help of contrast enhancement of blood vessels and the surrounding backdrops, segmentation has been done in retinal image to identify potential haemorrhage victim. The author used bounding box and density analysis to classify the various phases of diabetic retinopathy random forest algorithm is used according to eye illness on a specific location and circumference of image blood vessels and haemorrhage. The developed model results the accuracy evaluation of successfully classified output as 90% for normal and 87.5% of moderate and NPDR patient for DR classification.

In this paper, the author applies different machine learning classification algorithms to features extracted from various preprocessing filter such as gaussian filter, contrast enhancement and CLAHE such as optic disk diameter, exudates, micro aneurysms, and image level to make a person is affected with disease or not. To predict and classify the occurrence of diabetic retinopathy, different machine learning algorithms has been used for decision making such as decision trees, ada-boost, naive bayes, random forest, and SVM.

In this paper, the authors develop a new feature extraction model using some modification in Xception architecture to detect diabetic retinopathy. This model based on deep learning aggregation which combines different convolutional neural network layer of Xception as multi-level feature. These extracted multi-level feature are trained with multilayer perceptron for diabetic retinopathy severity check. The performance of developed model compared with four deep learning model feature extraction which include inceptionV3, mobileNET, ResNet and Xception architecture. To make work more efficient, hyper parameter tuning and transfer learning has been taken into consideration. The developed model gave an accuracy with APTOS 2019 dataset of 83.09%, sensitivity 88.24% and specificity 87%.

Models

Machine learning is a subset of AI, ML is the ability of any machine to work like human, think like human, perform every competition task to inspire from human. There are various machine learning techniques to perform a major role to complete almost every task such as image processing, image classification, progression product management, IOT traffic detection, self- driving car, text classification and medical diagnostic. Throughout previous years AI and machine learning provides most feasible solution in medical field and healthcare. ML algorithm can work well with them. The data which are stored to work on healthcare, there are very large amount of data which must be identified medical professional and doctor news different method for storing and processing it the given stored data but the recent era of AIML can perform any task on a store data for classification based on all data collected from previously diseased person. ML algorithm can provide different solution which are much efficient and accurate as compared to manual method. Let us discuss some of ML algorithm which gives us some useful information for the detection and diagnosis of DR.

There are many architectures or models for CNN such as Vgg, GoogleNet or Inception, ResNet, Xception, Vision transformer, mobileNet, efficientNet, alexNet, DenseNet, MobileNet, RetinaNet.

DenseNet is a CNN model which computes dense, multiclass features from the objects which are CNN based classifier. In this model the system can share a high precision and significant work on those regions which are overlapping to be classified, prohibitively slow training and slow runtime of such type of detectors. It is feed

forward network where next layer receives input from

the output generated by past layer (Figure 2).

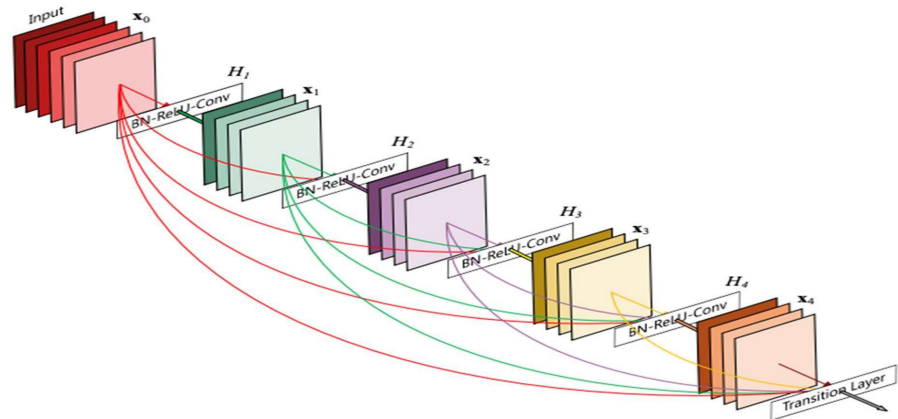


Figure 2. Architecture of DenseNet.

Vgg stands for visual geometry group. This architecture is a well-known architecture used by CNN. This architecture is well known for their simplicity and their uniform architecture. This architecture contains mainly CNN layers followed by Max pooling layers. There are two

most popular variants of this model which are VGG 16 this network consists of 16 hidden layers which include 13 CNN layers and 3 fully connected layers. As name VGG19 it contains 19 hidden layers which include 16 CNN layers and 3 fully connected layers (Figure 3).

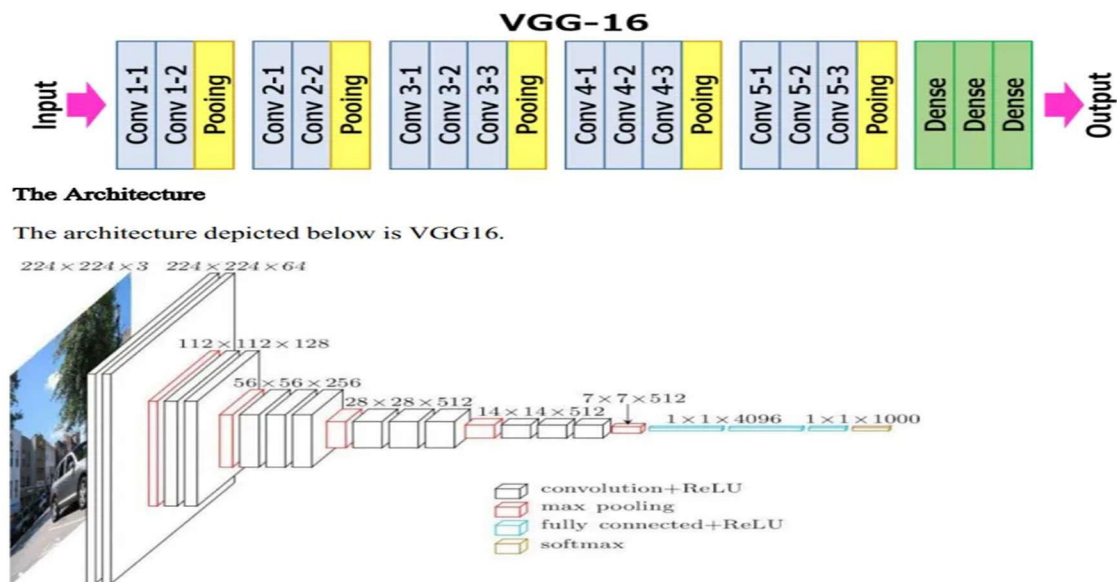


Figure 3. Architecture of VGG16.

GoogleNet or inception also known as InceptionV1. It is a convolutional neural network CNN architecture developed by Google in 2014. This network contains 22 hidden layers. This network trained on ImageNet classifies images which has capability to categorise 1000 images such as mouse, pencil, keyboard and many more

categories. The network trained on places365 dataset, much similar to the network trained with ImageNet but this classifies image into 365 different place categories. This model is based on many different very small convolutions to reduce the number of parameters (Figure 4).

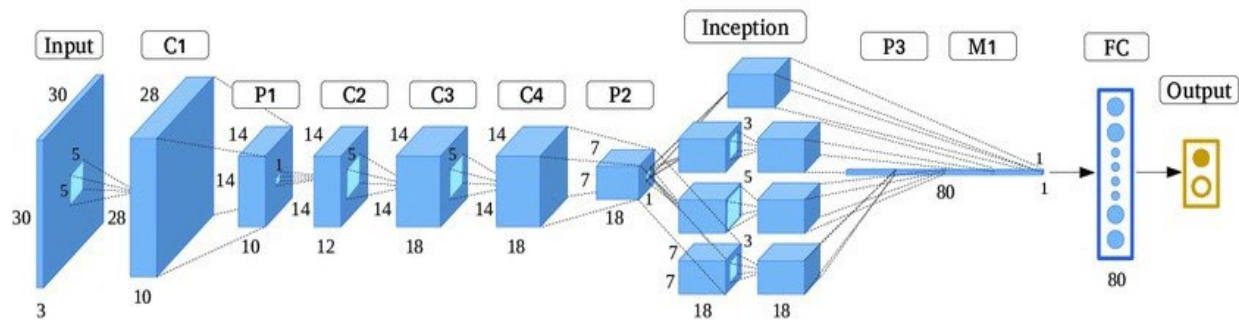


Figure 4. Architecture of GoogleNet.

ResNet stands for residual network. It is a groundbreaking CNN model introduced by Kaiming et al. Resnet given the concept of residual learning which is stakes the problem of diminishing gradient in very deep neural network. It uses the residual block where the input dual layer is mulched with output layer throughout skip connection.

Resnet architecture mainly consist of convolutional layer, residual block fully connected layer and down sampling for classification. This shortcut connection allows network to map the residual and result in deep architecture and try to improve performance on computer vision tasks like object and image classification (Figure 5).

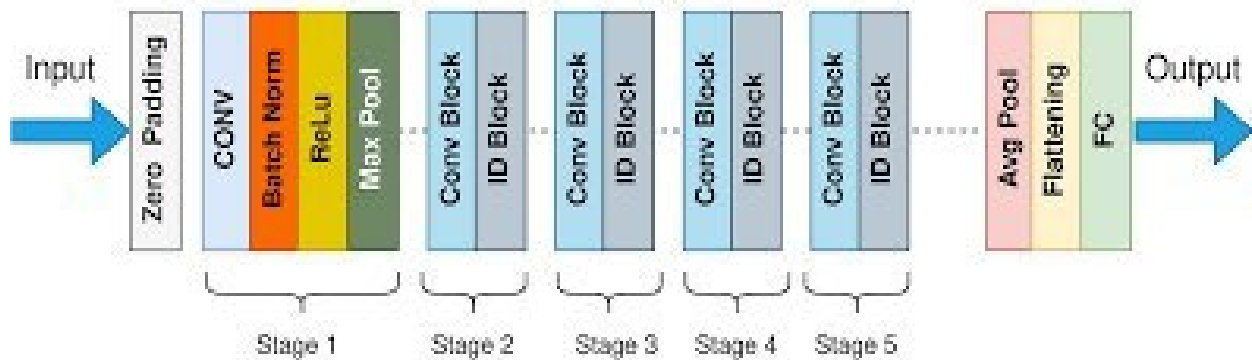


Figure 5. Architecture of ResNet50.

Xception is deep CNN architecture that involve depth separable convolutions. It is developed by Google. Google proposed a deep cnn inspired by inception. By replacing inception model with depth wise convolution separators

this developed approach enables a more efficient operation which result in a more effective and efficient system for generating, classifying and analysing the data (Figure 6).

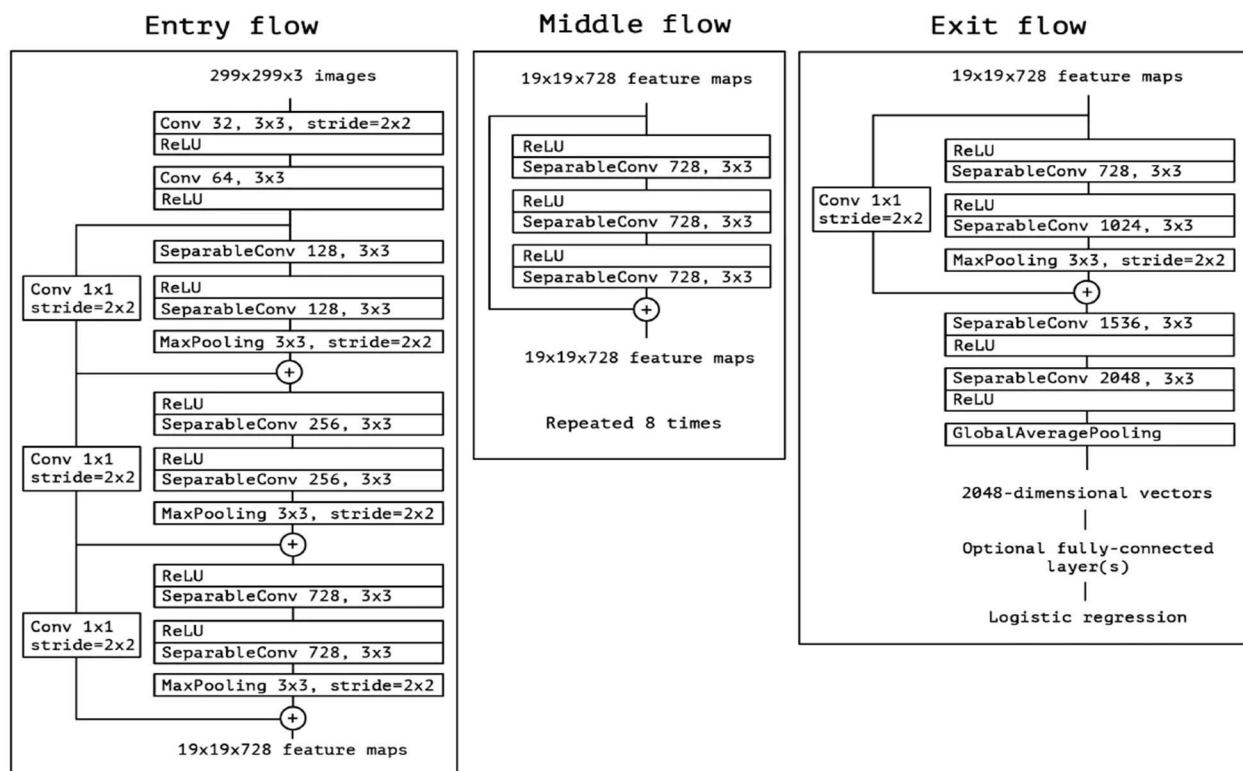


Figure 6. Architecture of Xception.

Mobile Net: A neural network architecture called Mobile Net, which has been developed by google to use on mobile and embedded devices. It is ideal for resource-constrained models like smartphones and tablets

because of its lightweight computationally efficient model, compact such as depth-wise separable convolutions. There are many versions of Mobile Net available like MobileNetV1, MobileNetV2, and MobileNetV4 (**Figure 7**).

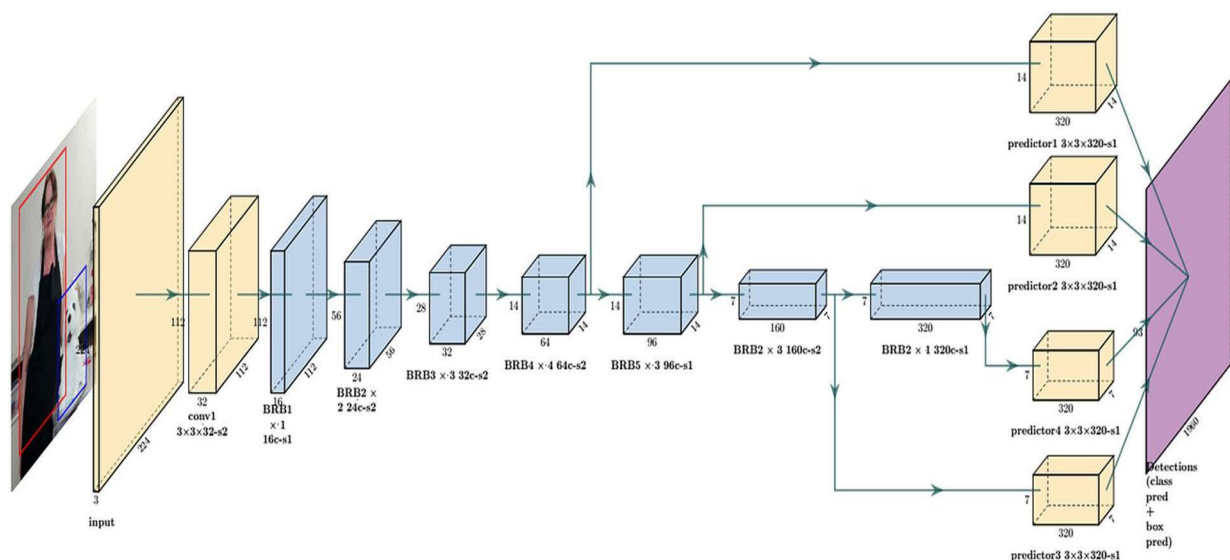


Figure 7. Architecture of MobileNet.

Efficient Net: It is a CNN architecture also a scaling method called Efficient Net which consistently scale each and every of the given dimensions of width, depth and

resolution using a combination coefficient, not like conventional work that arbitrary scales all these factors (Figure 8).

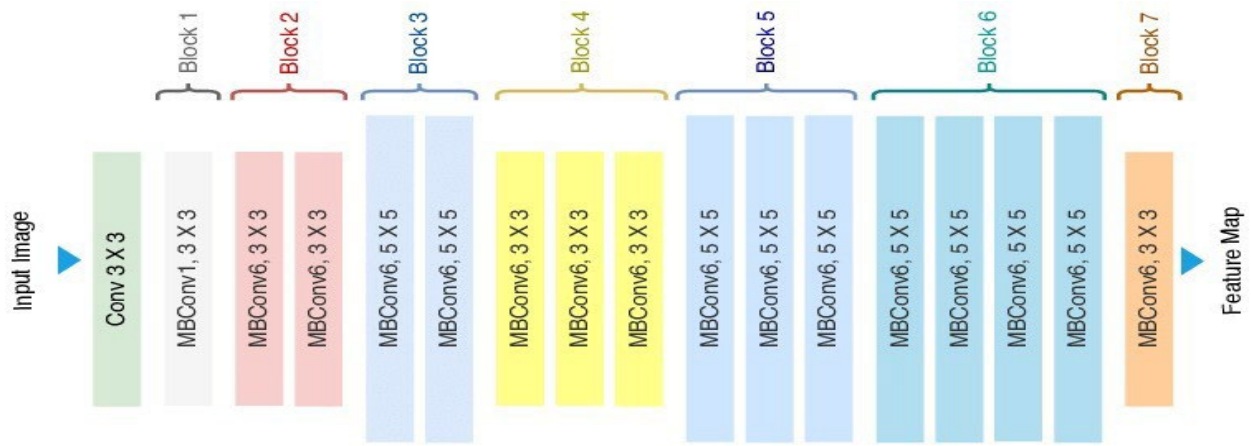


Figure 8. Architecture of EfficientNet.

Retina Net: To overcome class imbalance during training time retinaNet is one stage object detection architecture that uses the focused loss function to centralise learning on concrete negative instance. Loss in Focal gives a modifying phrase to cross entropy loss. Retinanet contains two task specific sub and backbone network to

form a single cohesive network. The backbone provides for computation of feature map over whole i/p image and off-the-self CNN. Convolution bounding box regression is obtained by second subnet where convolutional object classification is obtained by first on the output of backbone (Figure 9).

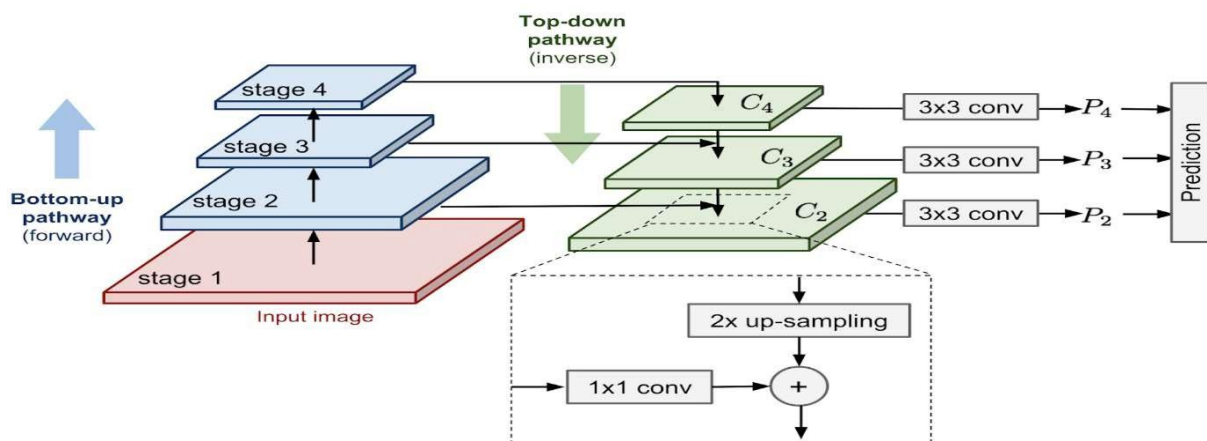


Figure 9. Architecture of RetinaNet.

AlexNet: Alex net has 8 layers with itself learnable parameters. This Alexnet model has eight layer where three layer are Max pooling layer fully connected layer and output layer where remaining 5 layers are connected with Relu activation function except output layer. By

using Alexnet the training process could be complete nearly six times faster with Relu activation function. By employing dropout layer to stop their model from overfitting. ImageNet dataset used here to train the model with up to 14 million images ('alexnet') (Figure 10).

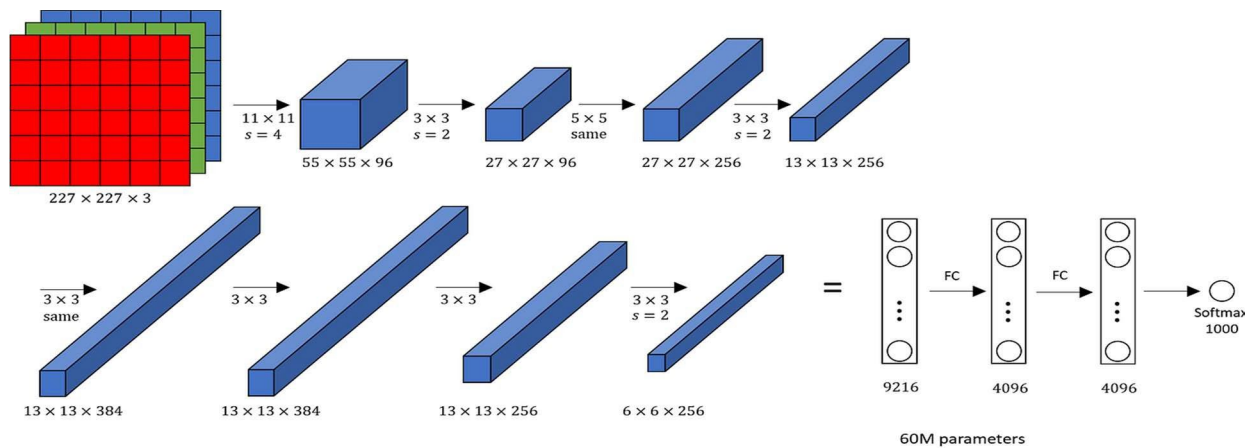


Figure 10. Architecture of AlexNet.

Vision transformer: A transformer-like design is used across picture patches in the vision transformer, image categorization design. The series of vectors is put into a typical transformer encoder, after division of an image into fixed-size sub images and linearly implanting each

one, along with adding position implanting. The conventional method of classification performs incorporating an extra learnable parameter "classification token" into the sequence (Figure 11 and Table 1).

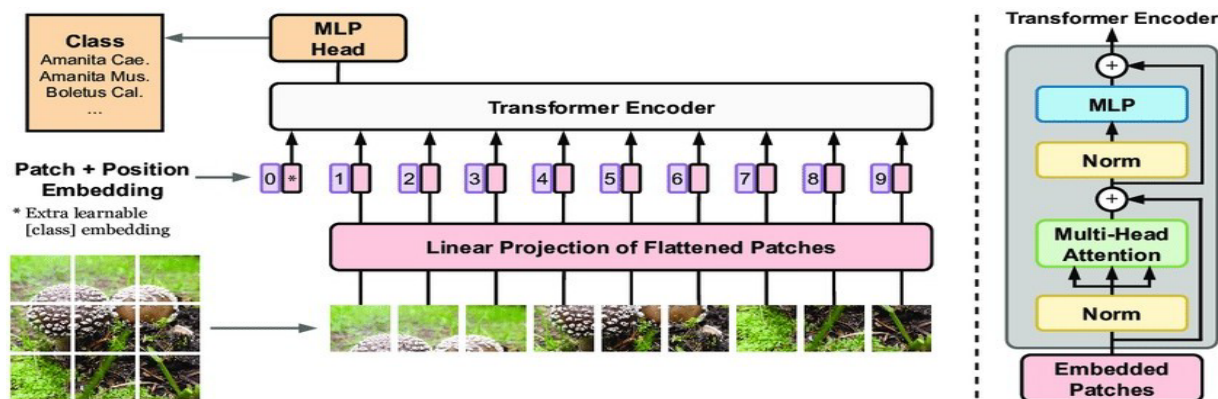


Figure 11. Architecture of Vision Transformer.

Table 1. Dataset.

S.No	Year	Dataset	Country	Size	Resolution	Annotation	Device	Task	No of studies
1	2004	MESSIDOR	France	1200	1440 x 960 2240 x 1488 2304 x 1536	Image level	Topcon TRC NW6	DR grading	3
2	2007	DIARETDB1	Finland	89	1500 x 1152	Pixel level	Zeiss FF450+	DR grading lesion detection	2
3	2009	eyePACS	United state	887 02	Varying	Image level	Topcon TRC NW6	DR grading	3
4	2010	MESSIDOR-2	France	1748	1440 x 960 2240 x 1488 2304 x 1536	Image level	Topcon TRC NW6	DR grading	3
5	2012	HEI-MED	United state	169	Varying	Pixel level	Zeiss Visucam Pro, fundus camera	DR grading lesion detection	2
6	2013	E-Optha	France	463	2544 x 1996 1440 x 960	Pixel level	CanonCR DGI, TopCon TRC NW6	DR grading lesion detection	1
7	2019	Drive	Netherland	80	564 x 584	Pixel level	Canon CR5	DR grading	2
8	2013	DRiDB	Croatia	50	768 x 584	Pixel level	Zeiss VISUCAM 200, DFC	DR grading lesion detection	
9	2015	EyePACS 2015	United state	887 02	433 x 289 5148 x 3456	Image level	4 camera	DR grading	

10	2017	IDRID	India	516	4288 x 2848	Image and pixel level	Kowa VS-10a, Fundus camera	DR grading lesion detection	1
11	2019	APTOS 2019	India	5590	Varying	Image level	Variety of common conventional cameras	DR grading	3
12	2019	DDR	China	136 73	Varying	Image and pixel level	Variety of common conventional cameras	DR grading lesion detection	2
13	2022	DIARETDB0	Finland	516	1500 x 1152	Pixel level	Zeiss FF450+	DR grading lesion detection	1
14	2019	ODIR	India	5590	Varying	Image level	Topcon 3 camera		1
15	2007	ROC	China	136 73	768 x 576 1058 x 1061 1386 x 1391	Image level	-	Microaneurysms detection	2
16	2021	FGADR	-	-	Varying	Image level	-	DR grading Lesion detection	1
17	2021	Shankara Netralaya	India	-	Varying	Image level	-	DR grading	1
18	2021	RFMID	-	3200	Varying	Image level	-	DR grading	1
19	2004	STARE	United State	400	700 x 605	Image level	-	DR grading	1
20	2010	ORIGA	Singapore	650	1500 x 2000 3500 x 5000	Image level	Variety of common conventional cameras	Glaucoma grading	1
21	2014	Chase	-	28	999 x 960	Image level	-	DR grading	2
22	2022	RIDB	United State	-	-	Images	-	DR grading	1

Data preprocessing

Preprocessing is a typical first step in deep learning research to prepared raw input that the network can accept, which can improve DL model's training performance. Varied cameras and cameras setting lead to varied quality problems which is size, noise artifacts, contrast, lighting, and sharpened areas in the fundus

photos. During the DL models training phase, these heterogeneities could obscure certain specific characteristics of the DR feature. As a result, we give a quick overview of the preparation processes for the DR classification using fundus images in the section. The fundus images undergo many major procedures, including image denoising, augmentation and normalization (**Table 2**).

Table 2. Data preprocessing approaches of methods.

Preprocessing approaches	Methods	Role of methods
Enhancement	Contrast enhancement: Histogram equalization retinex based methods. Adaptive histogram equalization CLAHE Illumination correction: Top-Hat transform homomorphic filtering Color space transformation: RGB to grayscale conversion RGB to HSL/HSC RGB to CIE	Here these given image enhancement methods are used to enhance fundus image appearance and other important information.
Denoising and normalization	Denoising: Gaussian filtering Wiener filtering Median filtering Mean filtering	To remove unwanted potential noise in captured pictures and keep away from feature biasness given methods are used.
Image augmentation	Geometric transformation: Rotating, affine, shifting, rescaling, cropping, projective Color transformation: Brightness transformation Contrast transformation. Color space transformation. GANs-based image generation	To increase the size of training data and avoid overfitting image augmentation is required.

Image enhancement: Contrast and light level in homogeneities in retinal images can be resolved using techniques like contrast enhancement, lighting and color space enhancement. The retinal contrast enhancement is used to highlight frontage features, while lighting

rectifying lowers the uniformity error in retinal images. CLAHE is one of the most efficient approach, controlling noise amplification and enhancing local contrast. Lighting correction is another technique, with gamma and logarithmic corrections being popular. Retinal image enhancement techniques based on deep learning, such as

Generative Adversarial Networks (GAN), have gained attention. CycleGAN, a model for retinal fundus picture enhancement, can produce output pictures for any input image. CBAM also called convolutional block attention modules can maintain color and texture consistency with input images. Cycle-CBAM outperforms CycleGAN in DR classification tasks. Dynamic retinal image feature limitation based on unpaired training parameters can help reduce artifacts and improve outcomes.

Denoising and normalization: To remove noises in the images such as Gaussian noise, salt and pepper noise, affects fundus pictures, which presents another major processing problem. This issue is commonly solved with filtering-based techniques. when contrasted with the mean, Gaussian, and Wiener linear filtering techniques. As a more reliable non-linear filtering method, the median filter effectively eliminates edge blurring and salt and pepper noise. Normalizing picture intensity is another common practice to lessen feature bias during training.

Image augmentation: It is commonly recognized that in order to prevent overfitting, DL models usually require a large quantity of training data. However, data augmentation is used to increase the amount and variety of the data, which can improve the resilience and accuracy of the models, because of absence of higher labeled data and class imbalance. Geometric and color alterations are examples of traditional data augmentation techniques for fundus imaging. Rotation, flipping, and cropping are examples of geometric

transformations that modify an image's geometry while maintaining the CNN's invariance to changes in location and orientation. Based on fundamental picture modifications, these approaches are simple and straightforward to use, although they sometimes miss some fascinating parts and suffer from the padding effect. In order to maintain the CNN's invariance to variations in illumination and color, augmentation in the color channel space is another useful technique. The elimination of lighting bias from pictures is the main benefit of color transformation; yet there is a risk of color information loss with this technique, and additional memory requirements, transformation expenses, and training time must all be considered. GANs, which create fake test with properties like the actual dataset, have gained popularity recently for data augmentation. Fundus image synthesis has been out by Zhu et al., using a combination of GAN and adversarial auto encoder training. The vascular tree was created using the former, and non-vascular characteristics were created using the latter. To produce a high-resolution (SR) picture, an Image Super Resolution (ISR) approach based on GANs in which every pixel's relevance was specified as a loss. The outcomes of the experiment demonstrated that the SR photos performed better than the pixel-weighting techniques. Reducing the extreme disparity between fundus pictures in various classes in DR classification is a significant task. Lastly, fundus pictures were synthesized using cGAN, a modified U-net, RF-GAN 1, RF-GAN2 and the masks and DR grading labels (**Figure 12**).

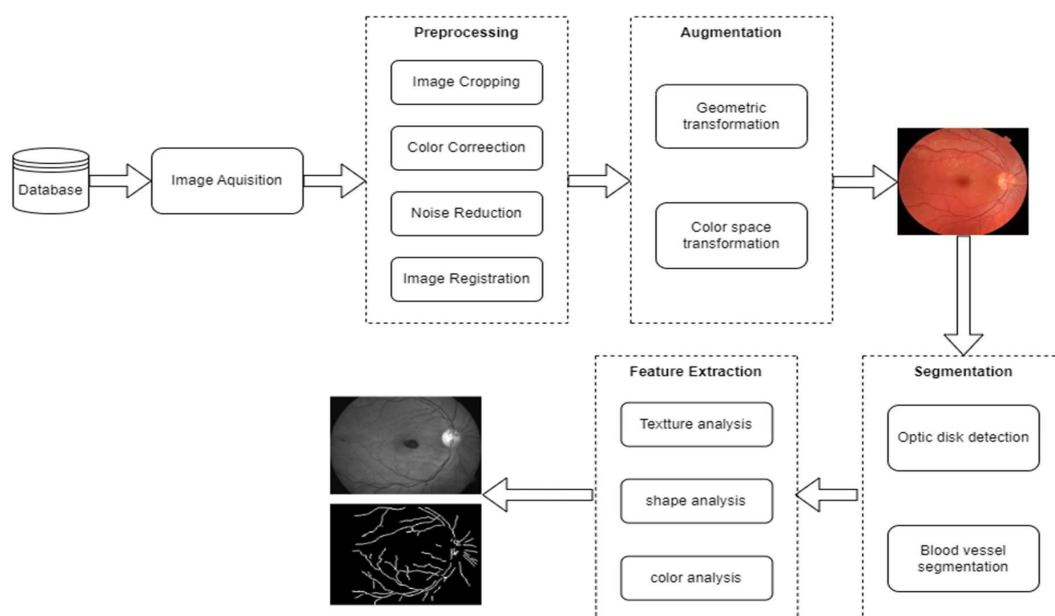


Figure 12. Preprocessing, augmentation, segmentation, and feature extraction of fundus images.

Segmentation

During classification of diabetic retinopathy, segmentation of fundus is a very necessary step to segment different feature and area of interest from

preprocessed fundus images. Preprocessed images have been divided into different small parts or regions where each represent different features which are fovea, micro blood vessels, OD and different class DR lesions which includes Micro aneurysms, Hemorrhage, and exudates.

Different image segmentation techniques have been taken into consideration such as U-Net for blood vessels and od segmentation, morphological operation for blood vessel segmentation, 2D wavelet discrete transformation. Saranya et al., used a non-U-Net based FCN-8S architecture, multi class U-Net and deepLabV3+ model segmentation of blood vessels, dark area and other needy information. Y Zhou et al., used random angle rotation horizontal and vertical reversal, random translation and image distortion in fundus images. Li, Fan and Chen used multilayer thresholding, adaptive thresholding green channel extraction to detect blood vessels, dark region

shown in fundus image. Akram and Khan, used some canny edge detection methods, multi-scale filtering for segment different ROI. Some graph cuts of exudates, alpha expansion and creation of various classes during the time of segmentation of images while classification of DR. Aras et al., used Clustering based Automated Region Growing Segmentation (CARGS) techniques to detect tiny blood vessels in retina so that classification task completes easily. Gabor wavelet, multilayer thresholding, computer aided detection system adaptive histogram equalization also used to summarize and identified by segmentation of fundus images (**Figure 13**).

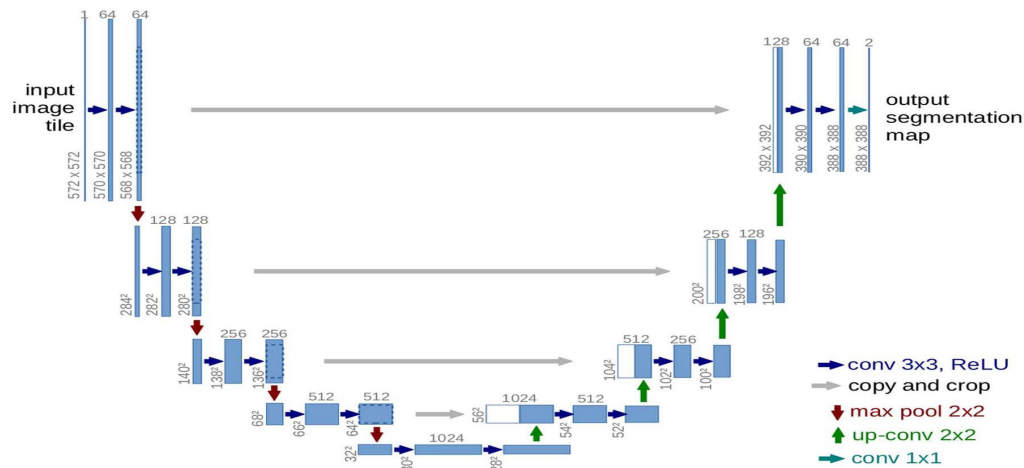


Figure 13. Segmentation of fundus image using U Net Architecture.

Feature extraction

Feature extraction methods are used to reduce large amounts of raw data into smaller, relevant data. These methods are divided into deep feature extraction and handcrafted features extraction.

Handcrafted features include LBP, LTP, SIFT, SURF, and HOG, and are used for DR classification. GLCM features are used for characterization of DR lesions, with accuracies of 97.31% and 93.86%, respectively. GLCM, GLRLM, and CRT are used for high-level texture feature extraction, with a 99.39% accuracy on the ODIR dataset. HOG descriptive features are used for DR lesion detection, with four types of features extracted for each

region of interest. SURF and spatial LBP are used for automated grading of DR lesions. GLCM with CNN performs best for glaucoma detection, with an accuracy of 93.16%. Handcrafted features like SURF and LOG are used for DR lesions classification, with SURF used for interest point detection and localization. Ali et al., used First order histogram equalization, second order co-occurrence matrix, wavelet and run-length matrix feature for extraction and classification. Aras et al., used an optic disk detection technique by Hough transform. Morphological operation, neural network techniques and thresholding techniques also used for feature extraction such as vessel tracking (**Figure 14**).

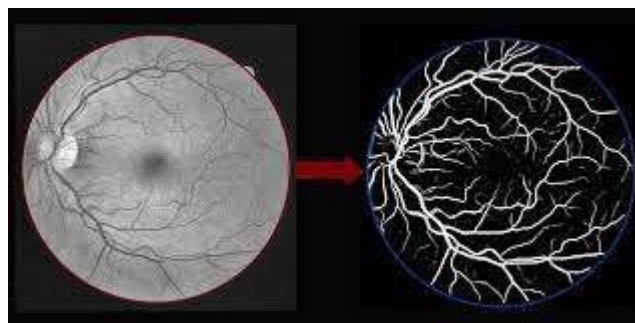


Figure 14. Segmentation and feature extraction.

Lesion detection

To measure the accuracy and performance of the system used in diabetic retinopathy classification, selection of best algorithms and lesion are very important steps to train any model. Based on selected lesion we can easily

identify the classification report of given system. There is various lesion present in our eyes which are microaneurysms, Hemorrhages and exudates (**Figure 15**). Each lesion has their capability and quality to detect severity of diabetic retinopathy. Some of well-known lesion are given below:

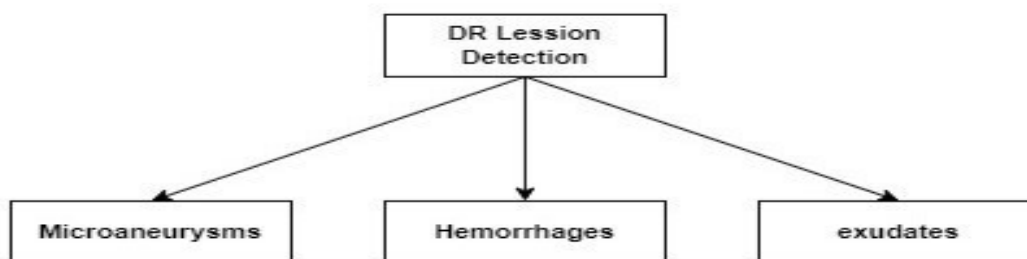


Figure 15. DR Lesion detection.

Microaneurysms: Microaneurysms (MAs) are microscopic, localized bulges in the walls of capillaries, which are tiny blood vessels, that resemble balloons. They are frequently observed in the retina of the eye and serve as a distinguishing feature of a number of retinal disorders, most notably diabetic retinopathy. Diabetes complications such as diabetic retinopathy can damage tiny blood vessels in retina due to persistently higher blood sugar levels. Micro aneurysms as well as other abnormalities including hemorrhages, exudates (fluid leakage), and macular edema can result from this injury. A thorough eye examination can help identify microaneurysms; methods like Optical Coherence Tomography (OCT) and fluorescein angiography are particularly useful in this regard. When fundus photography is used, they show up as tiny, spherical red spots on the retina. When determining the course of diabetic retinopathy and directing treatment choices, microaneurysm existence and severity, in addition to other retinal abnormalities, are critical markers. Preventing vision loss and other consequences linked to diabetic retinopathy, including microaneurysms, requires early identification and therapy of the condition.

Hemorrhages: The term hemorrhages, which can also be written hemorrhages, describes how blood leaks out of broken blood arteries and onto adjacent tissues or areas. They can be caused by a variety of things, such as trauma, illness, or medical disorders, and they can happen in different regions of the body. Hemorrhages are frequently linked to retinal illnesses, this includes macular degeneration, DR, hypertensive retinopathy, and vein occlusion in retina, in the context of eye and retinal health. Hemorrhages, for instance, can happen in diabetic retinopathy when the retina's fragile blood vessels

sustain damage from persistently elevated blood sugar levels. These hemorrhages might appear as thick blood pools (vitreous hemorrhage) inside the eye, bigger blotchy patches (flame-shaped hemorrhages), or tiny dot-like spots inside the eye. Similar to this, high blood pressure in hypertensive retinopathy can cause damage to the retina's blood vessels and leaking, which can cause hemorrhages. A thorough eye examination can reveal the presence of retinal hemorrhages, frequently using OCT, fluorescein angiography, fundus photography. The degree and existence of retinal hemorrhages can be crucial markers for determining the severity and course of underlying retinal disorders. Depending on the underlying cause and severity, treatment approaches for retinal hemorrhages might vary, but they usually focus on treating the underlying issue to stop additional damage and maintain eyesight.

Exudates: One kind of fluid that seeps from blood arteries and collects in surrounding tissues is called exudate. Exudates are frequently observed in the retina and are linked to several retinal illnesses, most notably diabetic retinopathy, in the context of eye health. Prolonged elevated sugar level in blood can damage the small blood vessel of retina in diabetic retinopathy. Liquid, proteins, and lipids may seep into the surrounding retinal tissue when these blood vessels leak. Exudates, which are seen as yellowish-white deposits on the retina, can occur as a result of this fluid buildup. Hard exudates and soft exudates are the two categories into which exudates in diabetic retinopathy are often divided. Using methods like fluorescein angiography, fundus photography or OCT, it is possible to see exudates in the retina during a thorough eye exam. To detect and classify DR in different grades, exudates can be used (**Figure 16**).

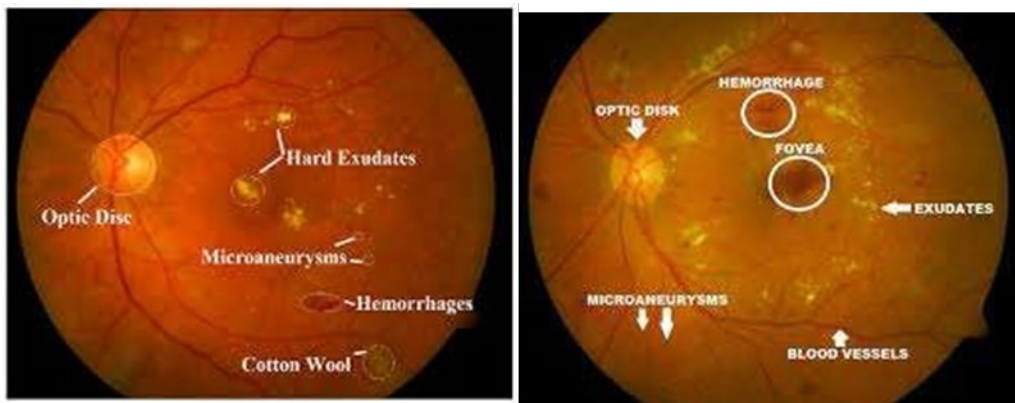


Figure 16. Microaneurysms, hemorrhages, optic disk and exudates in fundus images.

Comparison of existing DR techniques

The comparison of output of accuracy and loss value during training, testing and validation is discussed for the testing of the data using various machine learning approaches such as Vgg19, DenseNet, ResNet50 and Vision Transformer. Different classifiers have been used by researchers to obtain the best accuracy and precision to detect DR at early stage. Many preprocessing techniques have been used to enhance contrast, brightness, removal of noises and other unnecessary information by different techniques such as contrast limited histogram equalization, Adaptive histogram equalization, green channel extraction, filtering to remove noises in image, gray scale conversion, color correction. Different types of segmentation processes

have been used to extract features from fundus images such as blood vessels detection, optic disk detection, dark region detection and other ROI detection with the help of different morphological operations, U Net, non-U Net and multi class U Net architecture, 2-D wavelet discrete transformation, angle rotation, translation, multilayer thresholding, matched filtering, canny edge detection and many more techniques. For feature extraction in fundus images such as Microaneurysms, hemorrhages and exudates, some filtering has been applied to detect tiny blood vessels, dark region, and optic disk (Table 3). In all these proposed models, different types of classifiers have been used with different accuracy, precision, specificity, and sensitivity. Below figure is the tabular format of the result provided for classification followed by the output graph:

Table 3. Comparison of existing DR techniques.

S No	Architecture	Accuracy	Sensitivity	Specificity	Precision	F1 score
1	ResNet	0.924	0.854	0.943	0.977	0.918
2	AlexNet	0.946	0.953	0.938	0.937	0.945
3	GoogleNet	0.937	0.895	0.989	0.99	0.939
4	VGG19	0.966	0.949	0.984	0.985	0.967

Proposed system

There are various deep learning models which are used to check for the severity of diabetic retinopathy at early stage such as Vgg, GoogleNet or Inception, ResNet, Xception, vision transformer, mobileNet, efficientNet, alexNet, denseNet, mobileNet, RetinaNet. In this proposed system, the author introduces the steps for methodology of diabetic retinopathy classification. To enhance the given image, the author used her some contrast and brightness enhancement techniques which are contrast limited histogram equalization, gray scale conversion, shade correction and median filtering to remove unwanted noise in the image. In the process of segmentation, the author extract meaning components such as edges, lines, blood vessels and dark region in the image. To extract these features the author proposes some techniques which have been applied over

preprocessed image are a CNN non-U Net model, multilayer and adaptive thresholding to extract tiny edges and vein in the retina. The process of feature extraction, the author suggested some processes to sharpen the output and extract such meaningful information after sharpening the output obtain from segmentation that is the image components extracted from segmentation is sharpened in this step to classify the image with different classifiers. The methods used for feature extraction are morphological operation such as dilation, erosion, multiclass thresholding, vessel tracing, optic disk detection by Hough transform. Here, the author used resNet50 classification model for execution for this dataset (Figure 17). In the proposed model resNet50 provided best result with ODIR dataset of up to 97.55% among other classification models. The proposed resNet50 model contains 50 layers which are distributed as following parameters:

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	7×7, 64, stride 2				
conv2_x	56×56	3×3 max pool, stride 2				
		$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
conv3_x	28×28	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 8$
conv4_x	14×14	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 23$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 36$
conv5_x	7×7	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	1×1	average pool, 1000-d fc, softmax				
FLOPs		1.8×10^9	3.6×10^9	3.8×10^9	7.6×10^9	11.3×10^9

Figure 17. Layer distribution of ResNet50 Architecture.

Evaluation parameter

To evaluate the overall execution of the proposed system, the author used different evaluation parameters which produce sufficient result to classify the severity of diabetic retinopathy by classification with the help of confusion matrix. A confusion matrix contains True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN) result which is very helpful to calculate the given evaluation parameter. In this proposed model five parameter have been used to evaluate the overall performance of proposed model which are following:

Precision: It is the ration of total true positive classified values with respect to total positively classified values which is the sum of total true positive and false positive (TP+FP).

$$Precision = \frac{TP}{TP + FP}$$

Accuracy: It is the ration of total number of values classified as TP and FN with respect to the summation of values classified as TP, FP, TN, FN for the measurement of accuracy. The accuracy of any model can be obtained by given formula:

$$Accuracy = \frac{TP}{TP + FP + TN + FN}$$

Sensitivity: Sensitivity (SN) is calculated as the number of correct positive predictions divided by the total number of positives. It is also called Recall (REC) or True Positive

Rate (TPR). The sensitivity of any model can be obtained by given formula:

$$Sensitivity = \frac{TP}{TP + FN}$$

Specificity: Specificity (SP) is calculated as the number of correct negative predictions divided by the total number of negatives. The specificity of any model can be obtained by given formula:

$$Specificity = \frac{TN}{TN + FP}$$

F1 score: It is the harmonic mean of precision and recall. The F1 score of any model can be obtained by given formula:

$$F1\ Score = \frac{2 * Precision * Recall}{Precision + Recall}$$

RESULT AND DISCUSSION

At the first stage of preprocessing, the proposed model converts RGB color space image into HSI color space then segmentation, feature extraction, and training of proposed model has been completed by using resNet50 classifier. Different types of machine learning models have been implemented to check the severity level of diabetic retinopathy. The proposed model produces higher accuracy of 96.56% as compared to existing deep learning models which is much improved as compared to other machine learning models. Yang et al., used a hybrid CNN model which produces the accuracy of 96.26% with

APTOS2019 and MESSIDOR2 datasets. Saeed, Hussain and Aboalsamh proposed a two state transfer learning model which produces the accuracy of 96% with EyePACS and MESSIDOR dataset. (IEEE Engineering in Medicine and Biology Society) proposed an ensemble deep learning model by combining ResNet152, Inception ResNet V2, Inception V3 and Integrated model which produced accuracy up to 92.56%. The author proposed a model by using a CNN architecture of ResNet50 which produces an improved accuracy of 96.565 as compared to other ML models.

CONCLUSION

This paper proposes a machine learning based on CNN architecture of ResNet50 which contains multiple steps to enhance image quality and other meaningful information by performing preprocessing, segmentation, feature extraction and classification respectively. This model produces better accuracy as compared to other machine learning models which are denseNet, AlexNet, VGG. To extract features from segmented images a U-Net architecture has been used here to improve accuracy of fundus image with ODIR dataset. Thus, to reduce diabetic retinopathy at an early stage the proposed deep learning can be used to check the severity of fundus image. This proposed model has become very helpful to detect grading of any diabetic retinopathy patient and classify whether the patient is affected by this illness or not. The proposed model uses the evaluation parameters such as accuracy, precision, sensitivity, specificity, and F1 score.

REFERENCES

1. Akram MU, Khan SA (2013). Multilayered thresholding-based blood vessel segmentation for screening of diabetic retinopathy. *Eng Comput.* 29:165-173.
2. Albahli S, Nazir T, Irtaza A, Javed A (2021). Recognition and Detection of Diabetic Retinopathy Using Densenet-65 Based Faster-RCNN. *Comput Mater Continua.* 67(2).
3. Ali A, Qadri S, Khan Mashwani W, Kumam W, Kumam P, et al (2020). Machine learning based automated segmentation and hybrid feature analysis for diabetic retinopathy classification using fundus image. *Entropy.* 22(5):567.
4. Aras RA, Lestari T, Nugroho HA, Ardiyantol (2016). Segmentation of retinal blood vessels for detection of diabetic retinopathy: A review. *Commun Sci Technol.* 1(1).
5. Bala R, Sharma A, Goel N (2014). CTNet: Convolutional transformer network for diabetic retinopathy classification. *Neural Comput Appl.* 36(9):4787-4809.
6. Dutta S, Manideep BC, Basha SM, Caytiles RD, Iyengar NC (2018). Classification of diabetic retinopathy images by using deep learning models. *Int J Grid Distrib.* 11(1):89-106.
7. Incir R, Bozkurt F (2024). A study on effective data preprocessing and augmentation method in diabetic retinopathy classification using pre-trained deep learning approaches. *Multimed Tools Appl.* 83(4):12185-12208.
8. Jabbar MK, Yan J, Xu H, Ur Rehman Z, Jabbar A (2022). Transfer learning-based model for diabetic retinopathy diagnosis using retinal images. *Brain Sci.* 12(5):535.
9. Kandel I, Castelli M (2020). Transfer learning with convolutional neural networks for diabetic retinopathy image classification. A review. *Appl Sci.* 10(6):2021.
10. Karki SS, Kulkarni P (2021). Diabetic retinopathy classification using a combination of efficientnets. *Int Conf Emerg Smart Comput Inform (ESCI).* 68-72.
11. Li Q, Fan S, Chen C (2019). An intelligent segmentation and diagnosis method for diabetic retinopathy based on improved U-NET network. *J Med Syst.* 43:1-9.
12. Lin CL, Wu KC (2023). Development of revised ResNet-50 for diabetic retinopathy detection. *BMC bioinformatics.* 24(1):157.
13. Aatila M, Lachgar M, Hrimech H, Kartit A (2021). Diabetic retinopathy classification using ResNet50 and VGG-16 pretrained networks. *Int J Comput Eng Data Sci.* 1(1):1-7.
14. Nahiduzzaman M, Islam MR, Islam SR, Goni MO, Anower MS, et al (2021). Hybrid CNN-SVD based prominent feature extraction and selection for grading diabetic retinopathy using extreme learning machine algorithm. *IEEE Access.* 9:152261-152274.
15. Saeed F, Hussain M, Aboalsamh HA (2021). Automatic diabetic retinopathy diagnosis using adaptive fine-tuned convolutional neural network. *IEEE Access.* 9:41344-41359.
16. Saranya P (2022). Blood vessel segmentation in retinal fundus images for proliferative diabetic retinopathy screening using deep learning. *Vis Comput.* 38(3):977-992.
17. Shaukat N, Amin J, Sharif MI, Sharif MI, Kadry S, et al (2023). Classification and segmentation of diabetic retinopathy: A systemic review. *Appl Sci.* 13(5):3108.
18. Akram MU, Khalid S, Tariq A, Khan SA, Azam F (2014). Detection and classification of retinal lesions for grading of diabetic retinopathy. *Comput Biol Med.* 45:161-171.
19. Zhou D, Kang B, Jin X, Yang L, Lian X et al (2021). Deepvit: Towards deeper vision transformer. *arXiv.* 2103:11886.
20. Zhou Y, Wang B, Huang L, Cui S, Shao L (2020). A benchmark for studying diabetic retinopathy: segmentation, grading, and transferability. *IEEE Trans Med Imaging.* 40(3):818-828.