

DETECTING DIABETIC RETINOPATHY USING IMAGE PROCESSING AND DEEP LEARNING TECHNIQUES

Mylavarabhatla Pratyush^{*1}, Bhavika Arigala^{*2}, Kshitij Dhar^{*3}, MS Vishvesh^{*4},
Mayuri K P^{*5}

^{*1,2,3,4} Student, Dept. of CSE, Sir MVIT, Bengaluru, Karnataka, India.

^{*5} Asst. Professor, Dept. of CSE, Sir MVIT, Bengaluru, Karnataka, India.

ABSTRACT

Diabetic Retinopathy is an eye disease caused due to the damage of blood vessels present in the tissue at the back of the eye. Because of this, the vision could be lost if the level of diabetes reaches very high. In our paper, fundus images containing diabetic retinopathy have been taken into deliberation. The idea behind this paper is to propose and summarize an automated knowledge model of Diabetic Retinopathy Detection technique. The proposed model has been trained by using Deep Convolutional Neural Network (CNN). These Deep Learning models can evaluate the characteristics such as blood vessels, fluid drip, exudates, hemorrhages and microaneurysms into various classifications. This model will calculate the severity level of the patient's eye and will be helpful in recognizing the proper class of severity of diabetic retinopathy images. Image Processing techniques are assessed based prior to the model execution.

Keywords: Diabetic Retinopathy, fundus, Classification, CNN, Deep Learning, Image Processing.

I. INTRODUCTION

Diabetes mellitus, popularly known as diabetes, is a chronic disease which occurs when the pancreas does not secrete enough insulin, or the body is unable to process it properly. This disease affects the circulatory system slowly, including blood vessels of the retina. As diabetes progresses, the vision of a patient may start to deteriorate and eventually lead to diabetic retinopathy. One of the most common retinal complications related to diabetes is Diabetic Retinopathy (DR). It is asymptotic and hence, most of the patients remain unaware of the condition unless it slowly begins to impair their vision. Diabetic retinopathy progresses through the following stages - in the earliest stage, diabetic retinopathy may cause no symptoms or only mild vision problems due to the occurrence of micro aneurysms (MA). The second stage, moderate non-proliferative retinopathy (NPDR) is where the blood vessels nourishing the retina might distort and swell with the progress of disease, losing their ability of blood transportation. The third stage, severe non-proliferative retinopathy results in deprived blood supply to the retina due to the blockage of a greater number of blood vessels, hence signaling the retina for the growth of fresh blood vessels. Proliferative diabetic retinopathy (PDR), is the advanced stage where the growth features secreted by the retina activate the proliferation of the new blood vessels that are growing along the interior of the retina in some vitreous gel, filling the eye. Therefore, consistent screening of diabetic retinopathy in the earliest stage is essential in order to avoid further complications and to control spreading of the disease.

Diagnosis of DR is done by evaluation of fundus (retinal) images. The detection of blood vessels from the fundus images is usually a tedious process. The features are extracted from the raw images in the dataset using image processing techniques. In recent times, computer vision with deep neural networks is being used to train a model which improves the level of accuracy than any other neural network models. Many features have common intensity properties which are distinguished using geometric features and correlations. The detection of many of the features such as the blood vessels, hard and soft exudates, microaneurysms and hemorrhages can be done quite accurately, and the deep learning models can quantify the features into different classes. Micro aneurysms are small areas of balloon-like swellings in the retina's tiny blood vessels [7] and exudates/lesions are typically manifested as random whitish/yellowish patches of varying sizes, shapes and locations [10]. Retinal hemorrhage is bleeding from the blood vessels in the retina, inside the eye. Color fundus images are used to study eye diseases like diabetic retinopathy. The automatic processing and analysis of retinal images is time saving and can provide objective detection to the ophthalmologists in due time.

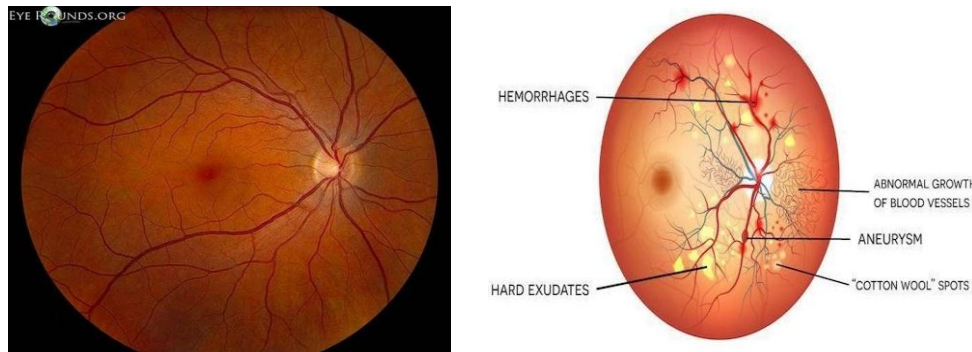


Figure 1: Normal eye (left) and Defected eye (right)

II. RELATED WORK

A lot of work has been done in the field of Diabetic Retinopathy depending on the research area and field of interest. Tons of papers have been published with the vision of detecting the disease, using various technologies and procedures, few of them being Image processing methods, Machine learning models, Data mining techniques, etc. The comparative study shows there is still a lack of Deep Learning methods to be implemented as far as Diabetic Retinopathy is concerned.

Thomas et al. [2] in their paper present to users, the contribution of image processing in the diagnosis of Diabetic Retinopathy. They reveal various efficient algorithms for the detection of optic disc and retinal exudates. They show the role of image processing using three key ways - Image Enhancement, using contrast stretching techniques, Mass Screening, including detection of pathologies and retinal features, and Monitoring, including feature detection and registration of retinal images.

Diego et al. [6] present a new supervised method for blood vessel detection in digital retinal images. This method uses a neural network (NN) scheme for pixel classification and computes a 7-D vector composed of grey-level and moment invariants-based features for pixel representation. The process stages are identified as - Original fundus image pre-processing for grey-level homogenization and blood vessel enhancement, Feature extraction for pixel numerical representation, Classification to label the pixel as vessel or non-vessel and finally, Postprocessing for filling pixel gaps in detected blood vessels and removing falsely detected and isolated vessel pixels.

Shradha et al. [7] produced edge maps based on Kirsch edge detection methods. The blood vessels in retinal images are classified using KNN classifier. They make use of 2D Gabor wavelet which proves to be efficient in enhancing vessel contrast while filtering out any noise. Information from wavelet responses at different scales are combined through the supervised classification framework, allowing proper segmentation of vessels of various widths.

Doshi et al. [13] presented the design, architecture and implementation of deep convolutional neural networks for automatic detection and classification of diabetic retinopathy from color fundus retinal images. They discussed the quadratic kappa metric, which is used to evaluate the prediction results.

Huang et al. [17] proposed a new convolutional network architecture, referred to as Dense Convolutional Network (DenseNet). It establishes direct connections between any two layers with the same feature-map size. They showed that DenseNets scale naturally to hundreds of layers, while exhibiting no optimization difficulties.

Suvajit et al. [19] in their paper proposed an optimal method for Diabetic Retinopathy using Deep Learning. They begin with applying a median filter to remove noise and enhance the images for better extraction of features, followed by Canny edge detection methods for feature extraction. They use Fuzzy C Means (FCM) to find out cluster levels of the training data which leads to better training accuracy. They use both images and statistical data to train using three neural network models - Feed forward NN, DNN and CNN to justify the differences in using statistical method and Image processing method. Results prove both CNN and DNN models to be effective in terms of image processing. Due to CPU training time of CNN getting affected in their study, DNN outperforms CNN for training accuracy as well as validation accuracy.

Meshram et al. [20] proposed the prominent CLAHE (Contrast Limited Adaptive Histogram Equalization) based image enhancement algorithm in view of the difficulties that arrive while applying standard enhancement techniques in Diabetic Retinopathy. The step enhances the contrast of image and distinguishes the details of the vessel's appearances. Finally, image segmentation is performed. The image is then converted into a group of white and black pixels to extract the blood vessels.

Kirange et al. [24] in the state-of-art methods use Kirsch's template technique for vessel extraction. They then use morphological operations for optic disc removal, followed by calculating the area of microaneurysms and hemorrhages which along with Gabor features ensemble together for accurate classification. For classification purposes, they use various algorithms like multi-class SVM, KNN, Naïve Bayes, Neural networks and Decision trees for comparison. The results prove the average accuracy of Naïve Bayes Classifier with Gabor features is comparatively better.

III. METHODOLOGY

The dataset we have used is MESSIDOR - 2. It contains fundus images examined for Diabetic Retinopathy where each examination contains two macula-centered eye fundus images (one per eye). The dataset contains 874 examinations (1748 images) which are populated by the Ophthalmology department of Brest University Hospital (France). The MESSIDOR - 2 dataset is publicly available for download free of cost. After performing data augmentation on the images by flip and rotation techniques, we succeeded in generating 12,605 images in total.

The first step in detecting Diabetic Retinopathy is to process the fundus images before passing them on to the Deep Learning model. For image processing, we first convert the RGB fundus images to grayscale. This is done to filter out the noise in the image. Each color component is then extracted into red, blue and green components. Green component is the best for detecting blood vessels. [20] Since each blood vessel has different contrasts, we use CLAHE (Contrast Limited Adaptive Histogram Equalization). CLAHE uses several histograms, each corresponding to a distinct section of the image, and utilizes them to redistribute the lightness values of the image which in turn, improves the contrast of the image. After this, a gaussian filter is applied to the image to remove any noise or distortion from the image. A bit further pre-processing is done at a later stage right before fitting the model directly using Keras functions. The table below shows the class distribution of our images in the dataset.

Table 1: Class Distribution of fundus images in Dataset

CLASS	NAME	NUMBER OF FUNDUS IMAGES	PERCENTAGE
0	No DR	1805	14.3%
1	Mild DR	2700	21.4%
2	Moderate DR	2700	21.4%
3	Severe DR	2700	21.4%
4	Proliferative DR	2700	21.4%

Pre-Trained Deep Convolutional Neural Network (CNN)

As a preliminary evaluation of accuracy and other metrics, we begin by supplying our input images to a Convolution Neural Network for the purpose of classification into the desired five classes, indicating the severity of the disease. The general model architecture of a typical ConvNet is as shown in the figure below.

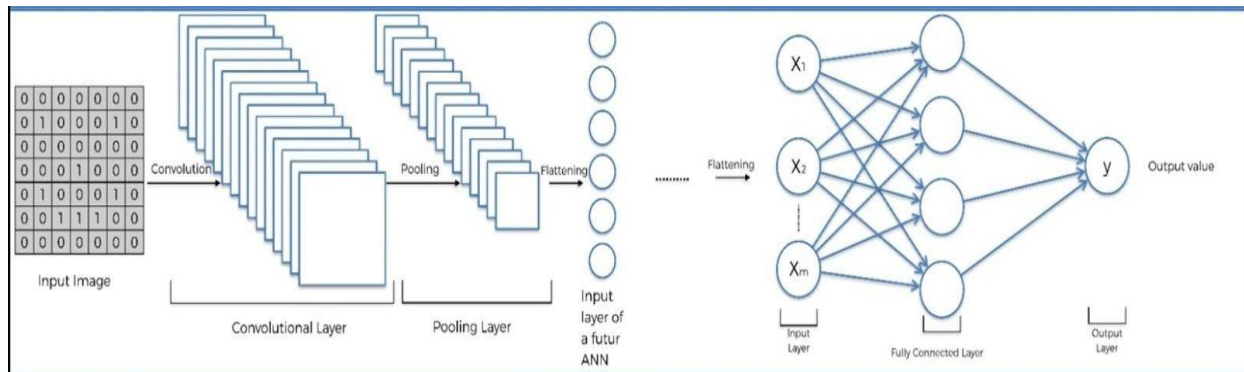


Figure 2: General CNN architecture

CNN Architecture

Our proposed CNN architecture consists of a network which takes images with resolution 512 * 512, as input. The architecture consists of 3 sets of combination of convolution, pooling and dropout layers stacked on top of each other. It is then followed by 2 sets of fully connected hidden layers, and finally an output layer.

1) Convolutional Layer

This is the first layer which is laid out to structure the CNN. In mathematical terms, it is a combined integration of two functions, and it shows how one function modifies the shape of another.

$$(f * g)(t) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} f(\tau) g(t - \tau) d\tau$$

The input image is mapped against a Feature Detector which gives us a Feature Map, which forms the backbone of this layer by reducing the dimensionality of input images and preserving only the important features of the image. The layer consists of a kernel or set of filters. Each filter is convolved against the input image to extract its features and form a new layer.

The first convolutional layer consists of 32 filters, followed by second with 64 filters and finally a third with 128 filters, each of size 3*3. Each filter convolves over its input with a stride of 1.

2) Pooling Layer

Pooling partitions the activation maps into a set of rectangles and collects the maximum value in the sub region. [12]

We perform Max Pooling in between our convolution layers for further dimensionality reduction. It helps in reducing the number of nodes in the future fully connected nodes without damaging the performance. We use a kernel of size 2*2 to perform this operation. Max pooling takes the maximum over the 2*2 regions in the depth slice of input.

3) Activation Functions

We have used ReLU (Rectifier Linear Unit) as the activation function for our hidden layers. It is a piecewise linear function that will output the input directly if it is positive, otherwise, it will output zero. It overcomes the vanishing gradient problem, allowing models to learn faster and perform better. Mathematically, it is defined as $y = \max(0, x)$.

Whereas for our output layer, we have used a softmax activation function to predict multinomial probability distribution. Softmax is used to normalize the outputs, converting them from weighted sum values into probabilities that sum to one. Mathematically, softmax is formulated as –

$$\sigma(\vec{z})_i = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}}$$

4) Dropout Layer

It is very likely for a deep neural network to overfit a training dataset with few examples. A single model can be used to simulate having many different network architectures by randomly dropping out nodes during training. This is called Dropout and offers a very computationally cheap and remarkably effective regularization method to reduce overfitting and improve generalization error in deep neural networks. Usage of dropout near fully connected layer avoids excessive generation of parameters. We set a hyperparameter indicating the probability at which outputs of the layer are dropped out, or inversely, the probability at which outputs of the layer are retained, to be 0.5 in our model.

5) Flattening

Flattening is converting the data into a 1-dimensional array for inputting it to the next layer. We “flatten” the output of the convolutional layers to create a single long feature vector which is then connected to the final classification model, which is called a fully connected layer.

6) Full Connection

The last layer of our CNN model, the Fully Connected Layer received the flattened nodes as an input layer which it uses to build an Artificial Neural Network in order to bind the features into attributes to predict the classes desired. Fully connected layers connect every neuron in one layer to every neuron in another layer, in order to classify the images, respectively.

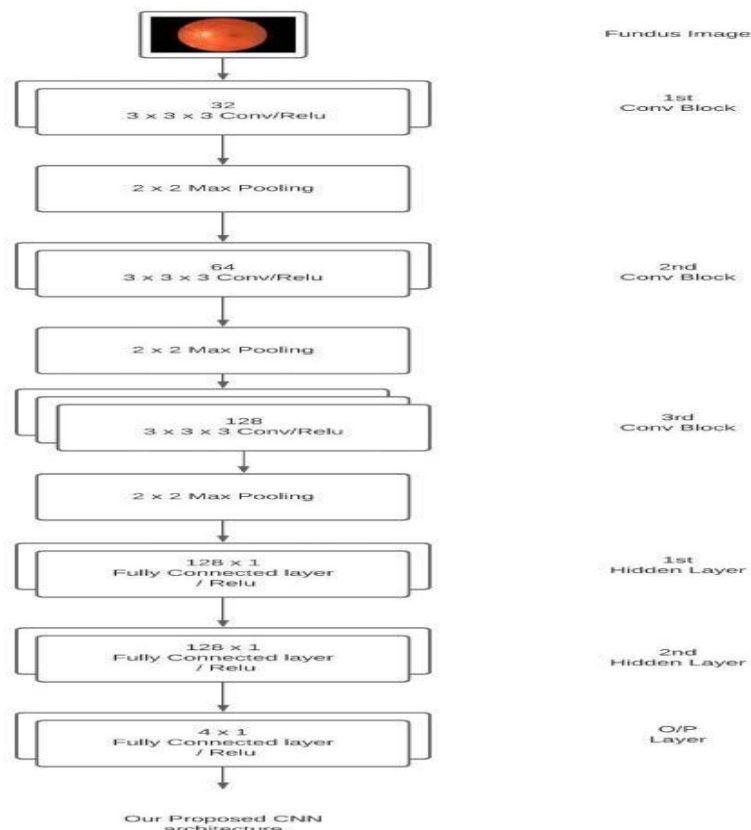


Figure 3: Proposed CNN Architecture using VGG16

Fine Tuning Using VGG16 and DenseNet121

On training our model through CNN and measuring the results, we found out the resulting accuracy was not up to the mark. Instantaneously, the validation accuracy had badly suffered. We realized our model was getting overfitted, possibly due to training of such a large deep neural network on a small sample set of images. To

overcome this issue, we proceeded to perform Transfer Learning using the VGG16 CNN model. Transfer learning is often performed when you have a small number of samples in your dataset to train a large model on.

VGG16

VGG16 [1] is a CNN architecture proposed by K. Simonyan and A. Zisserman in 2014 which became widely famous for its high accuracy results. It has 16 layers in total, with 13 of them being convolutional layers themselves. The weights in re-used layers may be used as the starting point for the training process and be adapted in response to the new problem. We load the pre-trained weights of this model so that we can utilize the useful features the model has learned for our task. The architecture of this model is indicated by the figure below.

1. Architecture

The input to the network is an image of dimensions (512, 512, 3). The first two layers have 64 channels of 3*3 filter size and same padding. Then after a max pool layer of stride (2, 2), two layers which have convolution layers of 256 filter size and filter size (3, 3). This is followed by a max pooling layer of stride (2, 2) which is the same as the previous layer. Then there are 2 convolution layers of filter size (3, 3) and 256 filters. After that there are 2 sets of 3 convolution layers and a max pool layer. Each has 512 filters of (3, 3) size with the same padding. This image is then passed to the stack of two convolution layers. In these convolution and max pooling layers, the filters we use is of the size 3*3.

We perform transfer learning by removing the last layer of our CNN model, and applying the model as input to this VGG16 to classify using its pre-existing network. The results are noted and compared with our proposed CNN model.

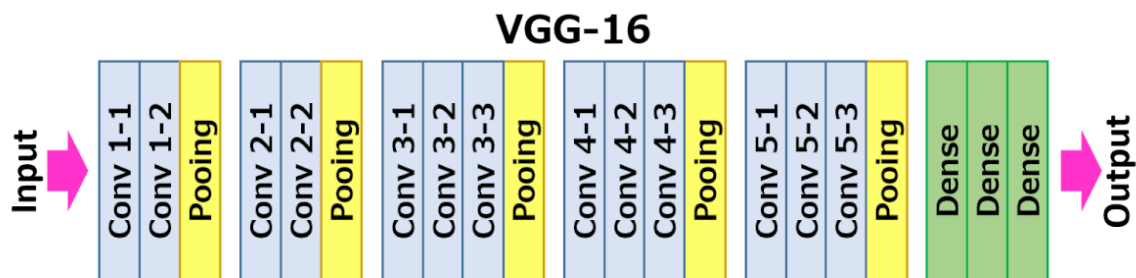


Figure 4: VGG16 Model and Architecture

The result on performing transfer learning using VGG16 was a lot better than classical model but still quite unsatisfactory considering the precision required when dealing with medical examination procedures. We went ahead to use another popular CNN architecture, called DenseNet121 to perform and compare the obtained results with our existing architecture.

DenseNet121

DenseNet is a type of convolutional neural network that utilizes dense connections between layers, through DenseBlocks, where we connect all layers (with matching feature-map sizes) directly with each other. To preserve the feed-forward nature, each layer obtains additional inputs from all preceding layers and passes on its own feature-maps to all subsequent layers. A DenseNet architecture has several compelling advantages including reduction in number of parameters, alleviating the vanishing gradient problem, and strengthening feature propagation [13]. A DenseNet network consists of $L(L+1)/2$ direct connections, where L is the no. of layers. To summarize, in a DenseNet architecture, each layer is connected to every other layer, hence the name Densely Connected Convolutional Network.

1. Architecture

Each architecture of DenseNet consists of four DenseBlocks with a variable number of Dense layers. In a dense block, each layer adds some features on top of the existing feature maps. The DenseNet-121 has [6,12,24,16] layers in the four dense blocks, respectively. As we can see from the figure below, the first part of the architecture consists of 64 filters of size 7*7 with a stride 2 convolution layer followed by a 3*3 stride 2 max pooling layer. Every dense block has two convolutions, with 1x1 and 3x3 sized kernels. In dense block 1, this is

repeated 6 times, in dense block 2 it is repeated 12 times, in dense block 3, 24 times and finally in dense block 4, 16 times. Each dense block is followed by a transition layer. In the transition layer, the number of channels is reduced to half. Transition layers applies batch normalization using down sampling, an essential step in CNN. The final dense block is followed by a Classification layer that accepts the feature maps of all the layers of the network in order to perform the classification. It would be essential to highlight that a DenseNet can have very narrow layers, say $k=12$, where k is the number of feature maps.

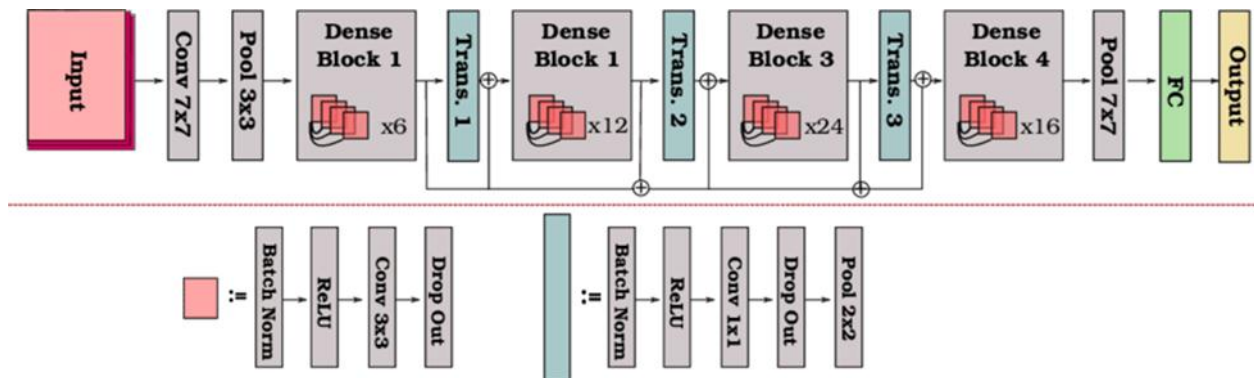


Figure 5: DenseNet121 Model and Architecture

IV. RESULTS AND EVALUATION

Image Processing

Input images are scaled down to 224×224 . Figure below shows an input image from our dataset along with its pre-processed output using python library OpenCV. VGG16 is a 16-layer deep convolutional network which is designed for image detection of 224×224 input images and uses small, 3×3 convolutional layers and 2 max-pooling layers. The depth of VGG allows it to extract more complex features from various images. DenseNet121 is a 121-layer deep convolutional network which is also known as the Densely Connected Convolutional Network. Its layers relate to feature maps of each foregoing layer which is used as the input for each subsequent layer. The reason for choosing DenseNet121 is that majorly reduces the number of parameters needed at a relatively low computational costs for a better performance.

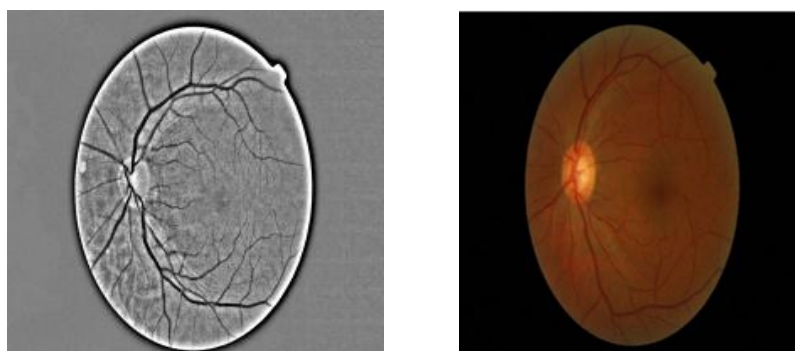


Figure 6: Input Fundus Image(left) and Fundus Image after Image Processing (right)

Model Analysis

In medical performance measurements, specificity, sensitivity and accuracy are crucial evaluation metrics. They are evaluated as-

$$SE = TP / (TP + FN) \quad SP = TN / (TN + FP)^*$$

$$Accuracy = (TN + TP) / (TN + FP + FN + FP)^*$$

*TP – True Positive, FP – False Positive, TN – False Negative, FN – False Negative

Sensitivity (true positive rate or recall) measures how likely the test is positive for someone who has diabetic retinopathy. Specificity (true negative rate) measures how likely the test is positive for someone who does not have diabetic retinopathy. Positive predictive value is also called as Precision. Accuracy measures the diabetic and non-diabetic patients from the database.

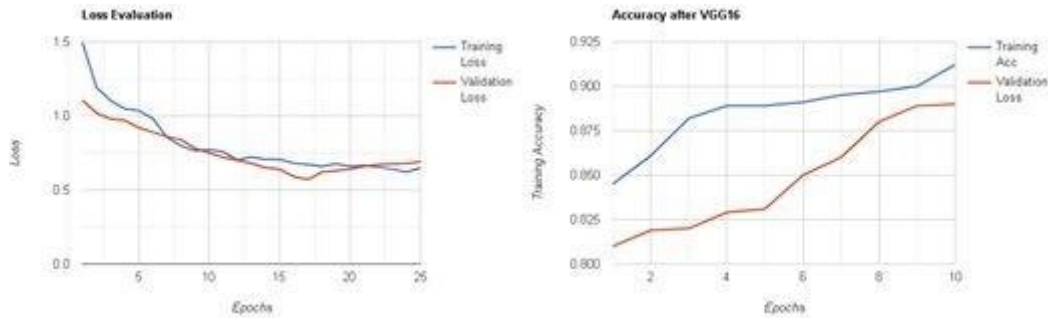


Figure 7: Training v/s Validation loss(left) and Training v/s Validation accuracy (right) for VGG16

Our model was initially trained on a VGG16 model and achieved an incompetent accuracy of 72.48% over 320 epochs. In order to improve this, we trained it using DenseNet121 model post which it quickly converges to provide a solid 89.91% testing accuracy. The graphs have been provided below, to compare the accuracies and to record any signs of overfitting or underfitting by noting down the training and validation losses.

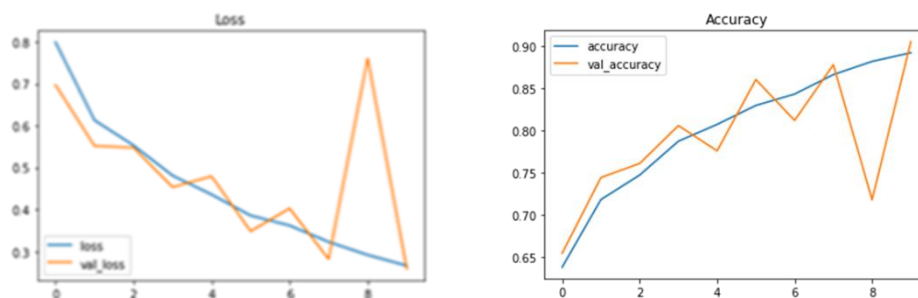


Figure 8: Training v/s Validation loss(left) and Training v/s Validation accuracy (right) for DenseNet-121

V. CONCLUSION

Our research paper proposes a system for DR detection and classification at different stages. Accurate processing of fundus images is very essential to obtain proper features. In Diabetic retinopathy, retinal blood vessels are damaged due to fluid leakage from these vessels. Different lesions, i.e., exudes, haemorrhages, microaneurysms, and textures are used to detect the stage of DR. Use of several image processing techniques for DR lesion detection have been discussed and evaluated. We trained our dataset on a VGG16 model and were able to achieve 72.48% accuracy in terms of identifying the classes of retinal images. To improve our accuracy, we performed Transfer Learning using the DenseNet121 CNN model. The proposed method can be improved, and the specificity of implementation can be increased by using some additional steps. A reliable fundus image analysis system is developed to give an ophthalmologist the most extensive view of the retinal state to be able to diagnose diabetic retinopathy.

VI. REFERENCES

- [1] Huan Wang, Wynne Hsu, Kheng Guan Goh, Mong Li Lee. "An Effective Approach to Detect Lesions in Color Retinal Images", IEEE (2000)
- [2] Thomas Walter, Jean-Claude Klein, Pascale Massin, and Ali Erginay. "A Contribution of Image Processing to the Diagnosis of Diabetic Retinopathy—Detection of Exudates in Color Fundus Images of the Human Retina", IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 21, NO. 10, (2002).
- [3] Chanjira Sinthanayothin, Viravud Kongbunkiat, Suthee Phoojaruenchanachai, Apichart Singalavanija. "Automated Screening System for Diabetic Retinopathy", IEEE (2003)
- [4] Jesse Vislisl and Thomas Oetting, "Diabetic Retinopathy", EyeRounds.org (2010)
- [5] Estevão S. Gedraite, Murielle Hadad. "Investigation on the Effect of a Gaussian Blur in Image Filtering

- and Segmentation" January (2011).
- [6] Diego Marín, Arturo Aquino, Manuel Emilio Gegúndez-Arias, and José Manuel Bravo. "A New Supervised Method for Blood Vessel Segmentation in Retinal Images by Using Gray-Level and Moment Invariants-Based Features", IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 30, NO. 1 (2011)
 - [7] Shradha Mirajkar, M. M. Patil "Feature Extraction of Diabetic Retinopathy Images". International Journal of Computer Applications (2013)
 - [8] Muhammad Waseem Khan. "Diabetic Retinopathy Detection using Image Processing: A Survey. ", International Journal of Emerging Technology & Research, Volume 1, Issue 1 (2013)
 - [9] Gupta S, Karandikar AM." Diagnosis of Diabetic Retinopathy using Machine Learning". Journal of Research and Development, Volume 3, Issue 2 (2015)
 - [10] Karen Simonyan & Andrew Zisserman "Very Deep Convolutional Networks for Large Scale Image Recognition" ICLR (2015)
 - [11] Ketki S. Argade, Kshitija A. Deshmukh, Madhura M. Narkhede, Nayan N. Sonawane, Sandeep Jore Automatic Detection of Diabetic Retinopathy using Image Processing and Data Mining Techniques.", International Conference on Green Computing and Internet of Things (2015)
 - [12] T Chandrakumar & R Kathirvel "Classifying Diabetic Retinopathy using Deep Learning Architecture", International Journal of Engineering Research & Technology, Vol. 5 Issue 06 (2016)
 - [13] Darshit Doshi, Aniket Shenoy, Deep Sidhpura and Dr. Prachi Gharpure. "Diabetic Retinopathy Detection using Deep Convolutional Neural Networks", International Conference on Computing, Analytics and Security Trends (CAST) (2016)
 - [14] Vishakha Chandore and Shivam Asati. "Automatic Detection of Diabetic Retinopathy using deep Convolutional Neural Network" International Journal Of Advance Research, Ideas and Innovations in Technology, Volume 3, Issue 4 (2017)
 - [15] Yashal Shakti Kanungo, Bhargav Srinivasan and Dr. Savita Choudhary. "Detecting Diabetic Retinopathy using Deep Learning", IEEE (2017)
 - [16] Elia J. Duh, Jennifer K. Sun, Alan W. Stitt JCI . "Diabetic retinopathy: current understanding, mechanisms, and treatment strategies by Insight". (2017)
 - [17] Gao Huang, Zhuang Liu, Laurens van der Maaten, Kilian Q. Weinberger . "Densely Connected Convolutional Networks" 28th January, (2018).
 - [18] Junjun Pan, Zhifan Yong, Dong Sui and Hong Qin. "Diabetic Retinopathy Detection Based on Deep Convolutional Neural Networks for Localization of Discriminative Regions", International Conference on Virtual Reality and Visualization (2018)
 - [19] Suvajit Dutta, Bonthala CS Manideep, Syed Muzamil Basha, Ronnie D. Caytiles and N. Ch. S.N. Iyengar. "Classification of Diabetic Retinopathy Images by Using Deep Learning Models", International Journal of Grid and Distributed Computing (2018)
 - [20] S P Meshram & M S Pawar "Extraction of Retinal Blood Vessel from Diabetic Retinopathy Imagery using Contrast Limited Adaptive Histogram Equalization" International Journal of Recent Advances in Engineering & Technology (IJRAET) Volume - 6, Issue - 8 (2018)
 - [21] Miss Sheetal N. Pore, Dr. C M Jadhao, Prof. S. S. Mhaske. "Detection of Diabetic Retinopathy in Retinal Images" IJCRT, Volume 6, Issue 2 April (2018)
 - [22] Sehrish Qummar , Fiaz Gul Khan, Sajid Shah , Ahmad Khan. "A Deep Learning Ensemble Approach for Diabetic Retinopathy Detection" October (2019)
 - [23] Nour Eldeen, Mohammed Loey, Mohammed Hamed and Nasr Eldin. "Deep Transfer Learning Models for Medical Diabetic Retinopathy Detection" December (2019)
 - [24] Dr. D. K. Kirange, Dr. J.P. Chaudhari, Dr. K. P. Rane, Dr. K. S. Bhagat and Dr. Nandini Chaudhri. "Diabetic Retinopathy Detection and Grading Using Machine Learning", International Journal of Advanced Trends in Computer Science and Engineering, Volume 8, No.6 (2019)
 - [25] Borys Tymchenko, Philip Marchenko and Dmitry Spodarets. "Deep Learning Approach to Diabetic Retinopathy Detection" March (2020)