INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact factor: 4.295
(Volume3, Issue4)
Available online at www.ijariit.com

Automatic Detection of Diabetic Retinopathy using deep Convolutional Neural Network

Vishakha Chandore
IIT, Delhi
vishakhachandore@gmail.com

Shivam Asati
IIT, Delhi
Shivamasati786@gmail.com

Abstract: ***The purpose of this project is to design an automated and efficient solution that could detect the symptoms of DR from a retinal image within seconds and simplify the process of reviewing and examination of images. Diabetic Retinopathy (DR) is a complication of diabetes that is caused by changes in the blood vessel of the retina and it is one of the leading causes of blindness in the developed world. Currently, detecting DR symptoms is a manual and time-consuming process. Recently, fully-connected and convolutional neural networks have been trained to achieve state-of-the-art performance on a wide variety of tasks such as speech recognition, image classification, natural language processing, and bioinformatics. In our approach, we trained a deep Convolutional Neural Network model on a large dataset consisting around 35,000 images and used dropout layer techniques to achieve higher accuracy.***

Keywords: ***Diabetic Retinopathy, Fundus Images, Neural Network, Macula, Optic Disc.***

1. INTRODUCTION

Diabetes is a group of metabolic diseases in which a person has high blood sugar, either because the body does not produce enough insulin, or because the cells do not respond to the insulin that is produced. Chronically high blood sugar from diabetes is associated with damage to the tiny blood vessels in the retina, leading to diabetic retinopathy. The retina detects light and converts it to signals sent through the optic nerve to the brain. Diabetic retinopathy can cause blood vessels in the retina to leak fluid or hemorrhage (bleed), distorting vision. In its most advanced stage, new abnormal blood vessels proliferate (increase in number) on the surface of the retina, which can lead to scarring and cell loss in the retina.

Diabetic retinopathy is one of the common complications of diabetes. It is a severe and widely spread eye disease. It damages the small blood vessels in the retina resulting in loss of vision. The risk of the disease increases with age and therefore, middle aged and older diabetics are prone to Diabetic Retinopathy. The progression from no retinopathy to PDR can take 2 decades or more, and this slow rate enables DR to be identified and treated at an early stage. Development and progression of DR are related to duration and control of diabetes. DR in its early form is often asymptomatic but amenable to treatment.
Diabetic retinopathy may progress through four stages:

1. Mild nonproliferative retinopathy - Small areas of balloon-like swelling in the retina's tiny blood vessels, called microaneurysms, occur at this earliest stage of the disease. These microaneurysms may leak fluid into the retina.
2. Moderate non-proliferative retinopathy - As the disease progresses, blood vessels that nourish the retina may swell and distort. They may also lose their ability to transport blood. Both conditions cause characteristic changes to the appearance of the retina and may contribute to DME.
3. Severe non-proliferative retinopathy - Many more blood vessels are blocked, depriving blood supply to areas of the retina. These areas secrete growth factors that signal the retina to grow new blood vessels.
4. Proliferative diabetic retinopathy (PDR) - At this advanced stage, growth factors secreted by the retina trigger the proliferation of new blood vessels, which grow along the inside surface of the retina and into the vitreous gel, the fluid that fills the eye. The new blood vessels are fragile, which makes them more likely to leak and bleed. Accompanying scar tissue can contract and cause retinal detachment—the pulling away of the retina from underlying tissue, like wallpaper peeling away from a wall. Retinal detachment can lead to permanent vision loss.

Problems in Diabetic Retinopathy Detection

Diabetic Retinopathy is the leading cause of blindness in the working-age population of the developed world. It is estimated to affect over 93 million people. The US Center for Disease Control and Prevention estimates that 29.1 million people in the US have diabetes and the World Health Organization estimates that 347 million people have the disease worldwide. Diabetic Retinopathy (DR) is an eye disease associated with long-standing diabetes. Progression to vision impairment can be slowed or averted if DR is detected in time, however, this can be difficult as the disease often shows few symptoms until it is too late to provide effective treatment.

Currently, detecting DR is a time-consuming and manual process that requires a trained clinician to examine and evaluate digital color fundus photographs of the retina. By the time human readers submit their reviews, often a day or two later, the delayed results lead to lost follow up, miscommunication, and delayed treatment.

Unfortunately, there is no effective known cure for diabetic retinopathy and the present treatments available are just management strategies at best. So its very important to detect the disease in its early stages.

Clinicians can identify DR by the presence of lesions associated with the vascular abnormalities caused by the disease. While this approach is effective, its resource demands are high. The expertise and equipment required are often lacking in areas where the rate of diabetes in local populations is high and DR detection is most needed. As the number of individuals with diabetes continues to grow, the infrastructure needed to prevent blindness due to DR will become even more insufficient. The need for a comprehensive and automated method of DR screening has long been recognized, and previous efforts have made good progress using image classification, pattern recognition, and machine learning. With color fundus photography as input, the goal of this project is to push an automated detection system to the limit of what is possible – ideally resulting in models with realistic clinical potential. Given a image of left and right eye of the patient, the main aim of the project is to classify the eye status among one of the following classes 0 - No DR, 1 - Mild, 2 - Moderate, 3 - Severe, 4 - Proliferative DR. So our task is to create an automated analysis system capable of assigning a score based on the above scale.

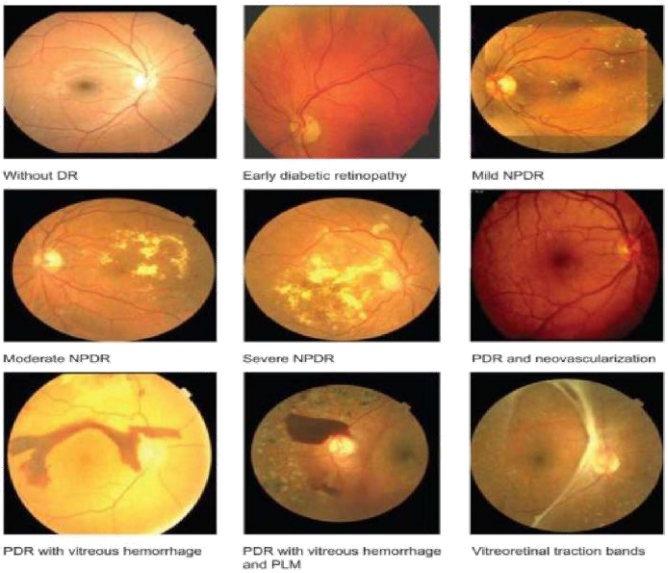


Fig. 1 Various Diabetic Retinopathy Stages

Features to Distinguish between a healthy and a non-healthy eye

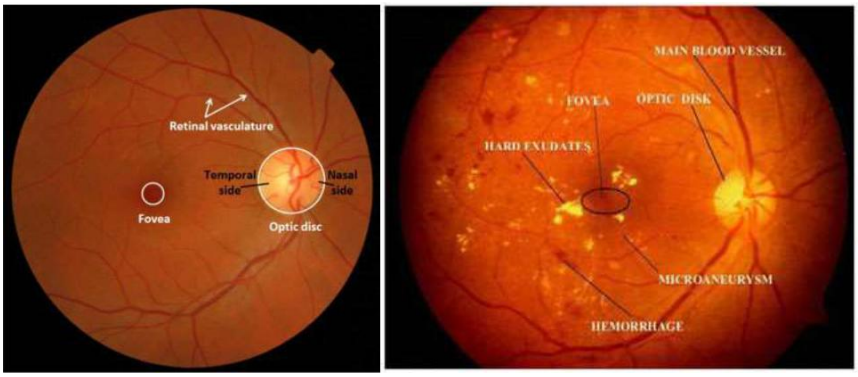


Fig. 2 FUNDUS images features

Optic Disc: OD is the region of the posterior pole where the vasculature and retinal nerve axons enter and leave the eye. The OD in a healthy retinal image usually appears as a bright yellowish and elliptical object marked by surface vessels. The presence of pathologic changes occurring at the site.

Of the OD, such as neo-vascularization from DR or changes to the physiologic c p due to glaucoma, can also affect its appearance dramatically.

Macula: The macula or macula lut ea is an oval-shaped pigmented area near the center of the retina of the human eye. It has a diameter of around 5.5 mm (0.22 in).The fovea is located near the center of the macula. It is a small pit that cont ains the largest concentration of cone cells. The macula is thus responsible for the central, high-resolution, color vision that is possible in good li ght; and this kind of vision is impaired if the macula is damaged, for example in macular degeneration.

Exudates: These yellow flecks are called exudates. They are the lipid residues of serious leakage from damaged capillaries. The commonest cause is diabetes.

Microaneurysms: A microaneurysm is a tiny aneurysm, or swelling, in the side of a blood vessel. In people with diabetes, microaneurysms are sometimes found in the retina of the eye. These miniature aneurysms can rupture and leak blood.

Hemorrhages: Retinal hemorrhage is bleeding from the blood vessels in the retina, inside your eye. Your retina is the thin layer that lines the back of your eye. Medical conditions, such as diabetes, high blood pressure, anemia, or leukemia. Eye problems, such as macular degeneration, or a bulging of the blood vessels in the retina. You may have no symptoms. You may have a sudden or gradual loss of vision, ranging from mild to severe. You may have blind spots.

Types of eye diseases (With their corresponding FUNDUS image)

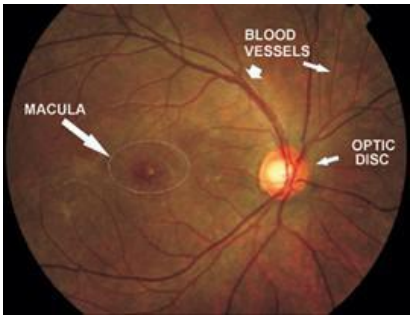
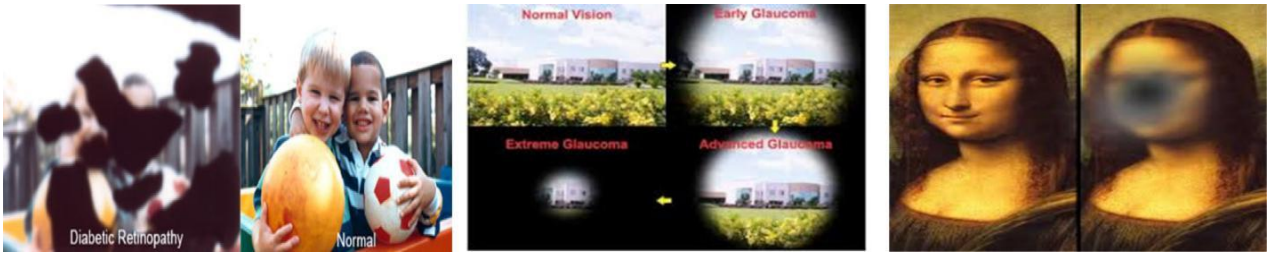


Fig. 3 - A Healthy Fundus

(a) (b) (c)



(d) (e) (f)

Fig. 4 Retinal image of an eye having (a) Diabetic Retinopathy (b) Glaucoma (c) AMD (age-related macular degeneration); Image saw by people having disease (d) Diabetic Retinopathy (e) Glaucoma
(f) AMD

2. RELATED WORK

Generally speaking, one may split all methods used to detect diabetic retinopathy into morphological or rule based methods and machine learning methods which have become increasingly more widespread.

Morphological Methods -A common theme in the literature of morphological approaches has been the need to split the problem up into first identifying the normal features or parts of the retina, such as the blood vessels, fovea, and optic disc, and then attempting to identify and possibly localize exudates and hemorrhages.

Some early approaches used various image processing techniques, for example, Pinz et al, (Pinz et al. 1998)[1] used the gradient based method and hough transform to map and localize blood vessels, the optic disc, and the fovea. Chaudhuri et al (Chaudhuri et al. 1989)[2] used two dimensional matched filters to map the network of blood vessels in the retina, a technique adopted by many later works.

Sinthanayothin et al (Sinthanayothin et al. 1999)[3] find blood vessels by performing PCA on image gradients and inputting the results to a neural network. Additionally, they localize the optic disc through simple intensity variations in image patches and find the fovea through matched filters. Filtering (Saiprasad Ravishankar 2009)[4] and Segmentation (Thomas Walter and Erginay 2002) [5] are also morphological methods.

Limitations: Several different computational strategies have been used in efforts to solve the problem of automated diabetic retinopathy. These attempts have been limited both by the amount of data available to researchers in this area and in the variety of methods used to solve the problem. The need to do so stems from the difficulty in formalizing the difference between similarly colored components such as hemorrhages versus blood vessels and exudates versus optic discs. In certain images, doing so is challenging even for a trained professional.

While various methods have had varying success identifying and localizing components of the retina, they typically operate on images with a good amount of contrast, few occlusions of retinal objects and few, if any, manifestations of the retinal disease. The last fifteen years have seen, however, a steady increase in the literature that attempts to tackle the problem of detecting not only retinal components in diseased eyes but also retinopathic components.

Machine learning methods - On the other hand, a number of attempts have been made to use machine learning to automatically locate manifestations of retinopathy. Examples include M. Melinscak et al[6], an automatic segmentation of blood vessels in fundus images. It contains deep max-pooling convolutional neural networks to segment blood vessels. It is deployed 10-layer architecture for achieving a maximum accuracy but worked with small image patches. It contains a preprocessing for resizing and reshaping the fundus images. It carried around 4-convolutional and 4-max pooling layer with 2 additional fully connected layers for vessel segmentation. Srivastava [7], a key idea of randomly drop units along with their connections during the training. His work significantly reduces the over fitting and gives improvements over other regularization techniques. Also, improves the performance of neural networks in vision, document classification, speech recognition etc.

Mrinal Haloi [8], a new deep learning based computer-aided system for microaneurysm detection. Comparing another deep neural network, it required less preprocessing, vessel extraction and more deep layers for training and testing the fundus image dataset. It consists of five layers which include convolutional, max pooling and Softmax layer with additional dropout training for improving an accuracy. It achieved low false positive rate.

Limitations: most of these attempts have come from the medical imaging community. There remain, however, many techniques from recent computer vision literature that have not been tried. Conversely, many of the computer-vision based attempts seem insufficient in practical terms because they do not generalize well when presented with the variety of real data observed by practitioners.

3. DIABETIC RETINOPATHY DETECTION SYSTEM

Our detection approach entails training a Convolutional Neural Network (CNN) to classify the level of DR in images. So, for training data, we were provided with approximately 35,126 labeled high-resolution images from Kaggle (online platform for data science competitions) taken under a variety of imaging conditions.

A clinician has rated the presence of diabetic retinopathy in each image on a scale of 0 to 4, according to the following scale:

0 - No DR
1 - Mild
2 - Moderate
3 - Severe
4 - Proliferative DR

Images were labeled with a subject id as well as either left or right (e.g. 1_left.jpeg is the left eye of patient id 1).

We then further augmented our training data to twice its original size. Then, we divided our training data itself into 20% validation data to fine tune our model. In all, there are roughly 61,000 training images and 14,000 validation images with 5 classification labels.

Preprocessing: Since the original images are fairly large (say, 3000x2000 pixels on average) and most of them contained a fairly large significant black border. We started removing most of these black borders but before that, as we required square matrix images as the input of our network, the images were first resized to say 3000 x 3000 (in the case of 3000 x 2000) by adding extra black borders and then resizing these images to 448 x 448. As if this does not take care correctly, we may have distorted FUNDUS images which may lose its original circular shape.

A primary step involved in the preprocessing is resizing the images into dimensions of 448x448.

Before proceeding into the architecture for classification.

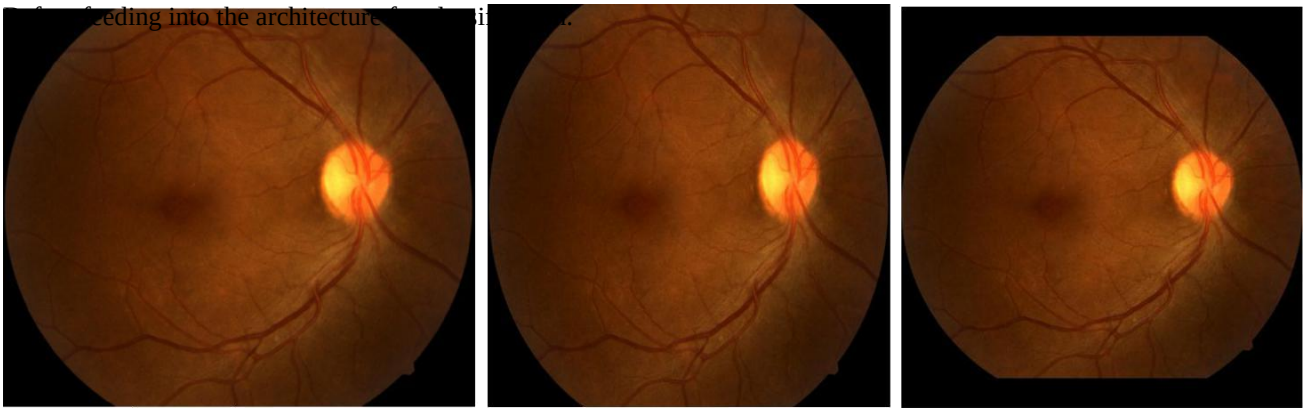


Fig. 5 (a) Image without preprocessing (say, 2000x3000) , (b) Image preprocessed without adding black borders(448x448), (c) Image preprocessed with adding black borders(448x448)

3.1 Architecture

3.1.1 Overall Architecture

In Image recognition, a Convolutional Neural Network (CNN) is a type of feed-forward artificial neural network in which the connectivity pattern between its neurons is inspired by the organization of animal visual cortex, whose individual neurons are arranged in such a way that responds to overlapping regions tiling the visual field. In deep learning, the convolutional neural network uses a complex architecture composed of stacked layers in which it is particularly well-adapted to classify the images. For multi-class classification, this architecture is robust and sensitive to each feature present in the images.

Now we are ready to describe the overall architecture of our CNN. As depicted in the table, the net contains fifteen layers with weights; out of which thirteen are convolutional and the remaining two are fully connected.

The output of the last fully-connected layer is fed to a 5-way softmax which produces a distribution over the 5 class labels. Our network maximizes the multinomial logistic regression objective, which is equivalent to maximizing the average across training cases of the log-probability of the correct label under the prediction distribution.

We also used three consecutive convolution layers followed by a maxpool layer to make our model deeper and increase the accuracy. Every convolutional and fully connected (dense) layer is followed by a ReLU layer to make the training faster.

The input image was 448x448 (RGB) and was reduced to 2x2 through our convolutional and dropout layers, and then fed into a fully connected layer with 1024 neurons to learn from. Then these 1024 neurons make predictions using softmax.

Table 1. – Deep Convolutional Neural Network Architecture

S. No.	Layer	Units	Filter	Stride	Size	Padding
1	Input				448	
2	Conv	32	5	2	224	2
3	Conv	32	3		224	1
4	MaxPool		3	2	111	0
5	Conv	64	5	2	56	2
6	Conv	64	3		56	1
7	Conv	64	3		56	1
8	MaxPool		3	2	27	0
9	Conv	128	3		27	1
10	Conv	128	3		27	1
11	Conv	128	3	2	27	1
12	MaxPool		3		13	0
13	Conv	256	3		13	1
14	Conv	256	3		13	1
15	Conv	256	3	2	13	1
16	MaxPool		3		6	0
17	Conv	512	3		6	1
18	Conv	512	3		6	1
19	Dropout		3			
20	Dense	1024	3			
21	Maxout	512	3			
22	Dropout		3			
23	Dense	1024	3			
24	Maxout	512	3			

3.1.2 ReLU Units

The standard way to model a neuron's output f as a function of its input x is with $f(x) = \tanh(x)$ or $f(x) = (1 + e^{-x})^{-1}$. In terms of training time with gradient descent, these saturating nonlinearities are much slower than the non-saturating nonlinearity $f(x) = \max(0; x)$. This activation function induces the sparsity in the hidden units. Also, It has been shown that the deep neural networks can be trained efficiently compared than sigmoid and logistic regression activation function. Deep convolutional neural networks with ReLUs train several times faster than their equivalents with tanh units. Faster learning has a great influence on the performance of large models trained on large datasets.

3.2 Reducing Over-fitting

3.2.1 Data Augmentation

The easiest and most common method to reduce over-fitting on image data is to artificially enlarge the dataset using label-preserving transformations. We employed transformed images to be produced from the original images with very little computation.In our implementation, the transformed images are generated in Python code on the CPU. The data augmentation consists of generating vertical and horizontal reflections with 50% probability. This increases the size of our training set by a factor of

2. Without this scheme, our network suffers from substantial over-fitting, which would have forced us to use much smaller networks.

(a) FUNDUS Image **(b) Image flipped horizontally** **(c) Image flipped vertically**

Fig. 7 - Data augmentation

3.2.2 Dropout

Combining the predictions of many different models is a very successful way to reduce test errors, but it appears to be too expensive for big neural networks that already take several days to train. There is, however, a very efficient version of model combination that only costs about a factor of two during training. The recently-introduced technique, called "dropout" consists of setting to zero the output of each hidden neuron with probability 0.5. The neurons which are "dropped out" in this way do not contribute to the forward pass and do not participate in back-propagation. So every time an input is presented, the neural network samples a different architecture, but all these architectures share weights. This technique reduces complex co-adaptations of neurons since a neuron cannot rely on the presence of particular other neurons. It is, therefore, forced to learn more robust features that are useful in conjunction with many different random subsets of the other neurons. At test time, we use all the neurons but multiply their outputs by 0.5, which is a reasonable approximation to taking the geometric mean of the predictive distributions produced by the exponentially-many dropout networks. Without dropout, our network exhibits substantial over-fitting. Dropout roughly doubles the number of iterations required to converge.

3.3 Details of learning

We trained our models using stochastic gradient descent with a batch size of 100 examples, the momentum of 0.9, and weight decay of 0.0005. We found that this small amount of weight decay was important for
the model to learn. In other words, weight decay here is not merely a regularizer: it reduces the model's training error. The update rule for w eight w was

$$v_{i+1} := 0.9 \cdot v_i - 0.0005 \cdot \epsilon \cdot w_i - \epsilon \cdot \left\langle \frac{\partial L}{\partial w}\Big|_{w_i} \right\rangle_{D_i}$$

$$w_{i+1} := w_i + v_{i+1}$$

$$\left\langle \frac{\partial L}{\partial w}\Big|_{w_i} \right\rangle_{D_i}$$

Where i is the iteration index, v is th e momentum variable, € is the learning rate and
is the average over the i_{th} batch D_i of the derivative of the objective with respect to w, evaluated at w_i We initialized the weights in ea ch layer from a zero-mean Gaussian distrib ution with standard deviation 0.01. We initialized the neuron biases in all the convolutional layers, as well as in the fully-connected hidden layers, with the constant 0. This initialization accelerates the early stages of learning by providing the ReLUs with posit ive inputs. We initialized the neuron biases in the remaining layers with the constant 0. We used an equal learning rate for all layers, which we adjusted manually throughout training.

4. ANALYSIS OF RESULT

The performance evaluation of any neural network is done on the basis of some specific parameters which decides whether the current model is justified for the dataset or not.

Evaluation Metrics: The measure ment of an accuracy for the network architecture is estimated by correctly classified DR suffered ima ges from the pool of images in the different da taset. Also, evaluate the algorithm which will be suffered by over-fitting or under-fitting could be visualized by plotting the training and validation loss. A whole objective is to minimizing the cost function of the deep convolutional neural network result s significantly reflected in the testing datasets.

In terms of diabetic retinopathy performance measurements, Specificity (SP), Sensitivity (SE) and Accuracy (Acc) are the crucial para meters for deciding the algorithms.
Four parameters which take part in measuring those performances are:

True Positive(TP) - Correctly dete cted DR images

True Negative (TN) - Correctly detected Non-DR images

False Positive(FP) - Number of Non-DR images are detected wrongly as DR imag s False Negative(FN) - Number of D R images are detected wrongly as Non-DR ima ges

At last, the Sensitivity, Specificity, and Accuracy are measured for each fundus ima ges available in the database.

$$SE = TP / TP + F N$$

$$SP = TN / TN + FP$$

$$Accuracy = TN + TP / TN + FP + FN + FP$$

$$Precision = TP / TN + TP$$

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

Since our model is very large (having fifteen convolutional layers and two dense layers), it is not easy to train such model on our personal laptops. Therefore, we tried two to simplify our model (using few layers only) and train it on a rather very small dataset with even less dimension of each image. We also simplified our labels to only 2 classes (not having DR - class 0; having DR - class 1). On this simplified model and dataset, we got the following results –

TABLE 2 - CONFUSION MATRIX RESULTS

actual results \ Predicted Results	Class 0	Class 1
Class 0	45 %	10 %
Class 1	5 %	40 %

Table 3 - Performance evaluation report

Label	Precision
Class -0	0.81
Class -1	0.88

These are only the initial phase results which were just an experiment. Before the presentation, we will further test our complete data to improve both performance and accuracy.

HARDWARE AND SOFTWARE REQUIREMENT

For augmentation and preprocessing, we used Python Imaging Library (PIL) for contrast adjustment, color balance adjustment, rotate or crop. At preprocessing stage, black border and resizing are done with the numpy package. Also, a special python library, JIT provided by Anaconda, continuum analytics was used to decrease the computation time. Convolutional Neural Network (CNN), multi-layer deep architecture are implemented using caffe[10] libraries open sourced by BVC. For handling our large dataset, Graphics Processing Unit was needed. To meet this requirement, HPC (High performance Computing), IIT Delhi[9] is used with 2x NVIDIA K40 (12GB, 2880 CUDA cores) GPU and 2x E5-2680 v3 2.5GHz/12-Core CPU having RAM of 64 GB.

CONCLUSION AND FUTURE WORK

Till now, we have not tested the complete test data, therefore, our first target would be to achieve the same. Our main focus would be to design such a detection system which is highly accurate and precise.

Among other existing supervising algorithms, most of them are requiring more pre-processing or post-processing stages for identifying the different stages of the diabetic retinopathy. Also, other algorithms mandatorily requiring manual feature extraction stages to classify the fundus images.

In our proposed solution, Deep convolutional Neural Network is a wholesome approach to all level of diabetic retinopathy stages. No manual feature extraction stages are needed. Our network architecture with dropout techniques yielded significant classification accuracy. This architecture has some setbacks such as an additional stage augmentation are needed for the images taken from a different camera with a different field of view. Also, our network architecture is complex and computation-intensive requiring high-level graphics processing unit to process the high-resolution images when the level of layers stacked more.

We can also implement our whole model as an application on mobile phones, so as to make diabetic retinopathy detection easier and time-saving.

REFERENCES

[1] Pinz, A.; Berngger, S.; Datlinger, P.; and Kruger, A. 1998. Mapping the human retina. IEEE Transactions on Medical Imaging 17:606–619.
[2] Chaudhuri, S.; Chatterjee, S.; Katz, N.; Nelson, M.; and Goldbaum, M. 1989. Detection of blood vessels in retinal images using two-dimensional matched filters. In IEEE Transactions on Medical Imaging, volume 8, 263–269.
[3] Sinthanayothin, C.; Boyce, J. F.; Cook, H.; and Williamson, T. H.1999. Automated localization of the optic dic, fovea, and retinal blood vessels from digital color fundus images. In Br J. Opthalmol., volume 83, 902–910
[4] SaiprasadRavishankar, Arpit Jain, A. M. 2009. Automatic feature extraction for early detection of diabetic retinopathy in fundus images.CVPR.
[5] Thomas Walter, Jean-Claude Klein, P. M., and Erginay, A. 2002.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 21(10):1236–1243.4
[6] M. Melinscak, P. Prentasic, S. Loncaric, “Retinal Vessel Segmentation using Deep Neural Networks”, VISAPP(1),(2015):577-582
[7] N.Srivastava, G.Hinton, A.Krizhevsky, I Sutskever, R Salakhutdinov, “Dropout: A simple way to prevent Neural networks from overfitting”, Journal of Machine learning research (2014) 1929-1958.
[8] MrinalHaloi, “Improved Microaneurysm detection using Deep Neural Networks”, Cornel UniversityLibrary(2015),arXiv:1505.04424.
[9] High-Performance Computing, IIT Delhi, http://supercomputing.iitd.ac.in/
[10] Deep-learning framework, http://caffe.berkeleyvision.org/.

