Diabetic Retinopathy Detection Using Deep Learning

(Mrs. Neeta Chavan, Assistant Professor in Department of **Electronics and Telecommunication Engineering, Vivekanand Education Society's Institute of Technology**)

Authors Name: Himanshu Agrawal
Department of **Electronics and Telecommunication,**
Vivekanand Education Society's Institute of Technology,
Mumbai 400071.
Email:2016.himanshu.agarwal@ves.ac.in

Authors Name: Atharva Musale
Department of **Electronics and Telecommunication,**
Vivekanand Education Society's Institute of Technology,
Mumbai 400071.
Email:2016.atharva.musale@ves.ac.in

Authors Name: Garima Dwivedi
Department of **Electronics and Telecommunication,**
Vivekanand Education Society's Institute of Technology,
Mumbai 400071.
Email:2016.garima.dwivedi@ves.ac.in

Authors Name: Mitali Kalayandasani
Department of **Electronics and Telecommunication,**
Vivekanand Education Society's Institute of Technology,
Mumbai 400071.
Email:2016.mitali.kalyandasani@ves.ac.in

***Abstract*— Diabetic Retinopathy is one of the leading cause of blindness for people who are suffering from diabetes. Early stage detection is vital to forestall permanent blindness. The number of people suffering from the disease is rising every year. For detection of diabetic retinopathy, the current method requires inspection of colored fundus images by a retina specialist. This is a tedious process and in the regions typically in rural areas where proper facilities and retina specialist are not present it becomes challenging. We have developed a system that can automate the task of detecting the disease using deep learning which uses convolutional neural networks. We have trained and tested our network and have achieved accuracy of 0.96, sensitivity of 0.94 and specificity of 0.97. Our proposed model is simple and efficient with respect to computational time and space is concerned.**

Keywords— Diabetic Retinopathy, Deep Learning, Convolutional Neural Networks, DenseNet121.

I. Introduction

Diabetic Retinopathy is one of the most leading cause of blindness among diabetic patients. People with Diabetes can have an eye condition known as Diabetic Retinopathy (DR), which can cause vision loss and blindness in people. This is when high blood sugar levels cause damage to the light-sensitive layer of tissue in the back of an eye (i.e. blood vessels in the retina). The number of people with diabetes was 425 million in 2017 and is projected to increase 642 million by 2040 [1] because of aging global population, urbanization and sedentary lifestyles. In the early stage of the disease, the person can show mild or even no signs of disease. This type is referred as Non-proliferative diabetic retinopathy (NPDR). In NPDR blood vessels are weakened and can lead to Proliferative diabetic retinopathy (PDR) [2]. PDR can cause oxygen deprivation in the retina. Over time this can cause floaters and swelling in the retina [3].

The current method of DR screening is based upon colored fundus image of eye inspected by a retina specialist. This method is tedious and a person with deep knowledge is required. In rural areas there are underdeveloped hospitals and often such specialist are absent. Also as the number of DR is increasing it is becoming difficult to detect the disease. With the help of deep learning we have come up with a solution that can automate the screening and can help reduce medical cost and also providing help in rural areas [4].

Deep Learning is a subfield of Machine Learning. The design of the network is based on the idea of neurons connected in the human brain [5]. Deep learning is a machine learning technique which is able to identify structure and patterns in the data by training on lot of real life examples. Deep Learning can be used in voice recognition, text classification, image recognition, etc. For example, self-driving cars uses to deep learning to recognize pedestrians and stop signs. Over the last few years, this field has gained a lot of momentum and is even outperforming humans in some tasks.

In this paper we have developed a grading system capable of identifying severity levels (0: No DR, 1: Mild DR, 2: Moderate DR, 3: Severe DR, 4: Proliferate DR). It is based upon Convolutional Neural Networks (CNNs), a class of deep neural networks. CNNs are capable of extracting different patterns in the images which then the system is capable of classifying the severity levels. We first made the architecture from scratch which gave sub-par performance. We then used transfer learning approach which gave better results.

II. Related Work

In the past various scientists and doctors worked on the field of early detection of Diabetic Retinopathy. There were several research going on to solve the problem exclusively by image processing. For example, in the year 2013, Adarsh et al [6], used image processing techniques for detection of DR through the detection of retinal blood vessels, exudate, micro-aneurysms and texture features. This achieved accuracies of 96% and 94.6% on the public 89 and 130 image databases DIARETDB0 and DIARETDB1 respectively.

In the year 2016, Varun Gulshan and Lily Peng [7] demonstrated the use of neural networks to detect the diabetic retinopathy. They proposed a CNN approach for diagnosing DR from digital fundus images and accurately classifying its severity. They also developed a network with CNN architecture and data augmentation to identify the features for the classification. They trained this network using a high-end graphics processor unit (GPU) on the publicly available Kaggle dataset and demonstrate impressive results, particularly for a high-level classification task. On the data set of 80,000 images used our proposed CNN achieves a sensitivity of 95% and an accuracy of 75% on 5,000 validation images.

In the year 2017, Kele Xu et al [8] and Manoj Raju, Venkatesh Pagidimarri et al [9], also used deep convolutional neural networks to challenge the problem. Thus, in this paper, they explored the use of deep convolutional neural network methodology for the automatic classification of diabetic retinopathy using color fundus image. Kele Xu et al obtained an accuracy of 94.5% on our dataset, outperforming the results obtained by using classical approaches. Also Manoj Raju, Venkatesh Pagidimarri et al used approximately 35,000 images to train the network. They achieved sensitivity of 80.28%, specificity of 92.29% an accuracy of 93.28% on the validation dataset of ~53,000 images.

III. Dataset

The image dataset consisted of fundus images of eye. The images were acquired from Kaggle. The images on Kaggle were provided by Arvind Eye Clinic. The public dataset consisted of total 3662 images. Out of the total images 2662 images were used for training, 500 for validation, and 500 for testing.

The images were classified as five stages of severity. The five stages were No DR, Mild DR, Moderate DR, Severe DR and Proliferative DR. The images had varied levels of lightening, height and width of images ranged from couple hundreds to thousands. Some images were also out of focus, underexposed and overexposed.

IV. Methods

A. Convolutional Neural Network (CNN)

Convolution Neural Networks (CNN) also known as CovNets are a category of Neural Networks that have proven effectively in recognition and classification specifically in the areas that deals with images. They are currently been used in the areas of self-driving cars, robotics, facial recognition, healthcare, etc. The architecture of CNN consist of four layers (Convolution layer, ReLu layer, Pooling layer, Fully Connected layer) [10]

- The convolution layer is used to detect different features in the images like edges, corners, shapes, objects, etc. Initial convolutional layer detects features like lines, edges but as we go deeper into the network it can identify full objects like face, dog, cars, etc.
- The ReLu layer is used to increase the non-linearity of the image. It is used to strip away the excessive fat to provide a better feature extraction. This increases the speed of the training and helps to avoid vanishing gradient problem.
- Pooling layer is used to reduce the number of parameters in the network. This is done by dimensionality reduction of the image. It reduces overfitting, computational load and training time.
- The last layer in the network is the fully connected layer. The features extracted from the previous layers are combined and aggregated in this layer to perform classification.

B. Preprocessing

The images were of very high resolution and of different sizes. Hence the images were resized to 224x224 pixels. The images contained lot of black region which contributed to noise. So we cropped it to reduce the noise which increased the accuracy and reduced the computational time. No other heavy preprocessing was done as we found out that it didn't improve the results and also increased unnecessary computational time.

C. Data Augmentation

During each epoch random augmentation of images was performed to reduce overfitting. The images were randomly flipped horizontally, flipped vertically, zoomed, and rotated. To mitigate the less number of images and class imbalance as shown in Figure 1, data augmentation helped give variance in the dataset.

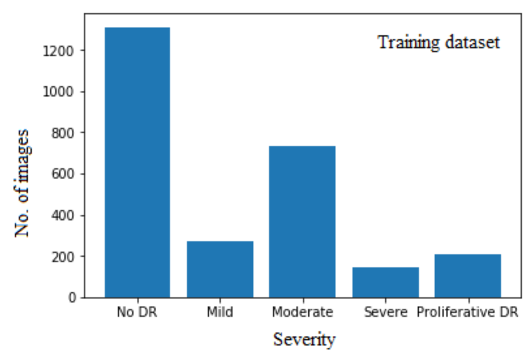

Figure 1: No. of images in training dataset to demonstrate class imbalance problem

D. Transfer Learning

We built model from scratch and also used transfer learning approach. For transfer learning we have used DesnseNet121 network. The last fully connected layer of the network was removed, then a transfer learning scenario was followed [11] by treating the remaining network components as a fixed feature extractor for the new dataset. The weights were pre-initialized from the network which was used to train ImageNet dataset.

V. Experimental Results

We divided the image dataset into training, validation and testing dataset. The training dataset had 2662 images, validation dataset had 500 images and testing dataset had 500 images. The images were resized to 224x224 which DenseNet121 network used the same size for training the ImageNet images. As the images had huge black region around the eye, it had to be cropped to reduce the noise. We also didn't use any special preprocessing as the image quality was already excellent and even if we did there was no significant improvement in the results. The network was trained on GPU of Kaggle Notebooks. It uses Nvidia Tesla P100 GPU. The network was trained for 50 epochs and used batch size of 32.

A. Single Label vs Multi Label

Initially we used single label for predicting the labels. The results were sub-par and we thought there was room for improvement. We stumbled upon Lex Toumbourou [12], as he demonstrated the use of creating Ordinal targets. That means if the target is a certain class, then it encompasses all the classes before it.

B. Model Made from Scratch

Convolution layer, MaxPooling layer, BatchNormalization, and Fully Connected layer was used to build the architecture. After every two Convolution layer, MaxPolling and BatchNormalization was used. This was repeated for five times. Upper Convolution layer had 64 filters, and deeper into the network the Convolution layer contained 512 filters. The filter size used was (3, 3) and activation function as ReLu. In the end a fully connected layer was used with Dropout of 50% and activation function as sigmoid. The loss function used was binary crossentropy and optimizer used was adam. The learning rate used was initialized with 0.01, and during training if the validation loss does not decreases after every 5 epochs then the learning rate was reduced by factor of 0.2.

As seen in the figure 2 after around 35-40 epochs the loss and accuracy of training and validation almost seemed constant and there was no improvement in the results after that. The results of the model are shown in the table 1. The confusion matrix is shown the figure 3.

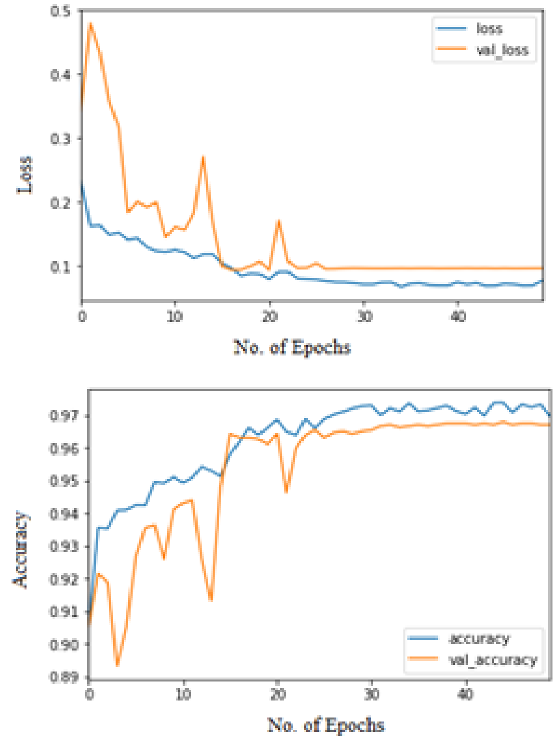

Figure 2: Loss and Accuracy Graphs for model made from scratch

Sr. No.	Parameters	Values
1	Training Accuracy	0.93
2	Validation Accuracy	0.92
3	Testing Accuracy	0.93
4	Sensitivity	0.87
5	Specificity	0.97

Table 1: Scores for model made from scratch

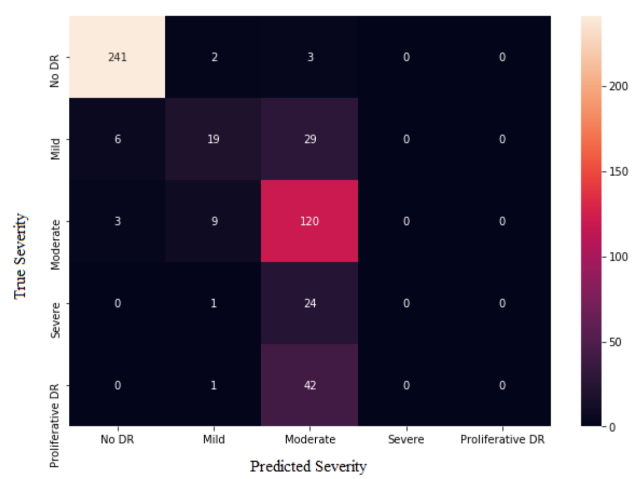

Figure 3: Confusion Matrix for model made from scratch

C. Using Transfer Learning

Even though the pre trained network was not trained on fundus images of the eye it can help improve the results and also reduce the training time [13] as weights are pre-initialized rather than randomly initialized. We used DenseNet121 architecture which consists of 121 layers. Using pre trained network, it was able to learn more features and help avoid vanishing gradient problem [14]. In the end GlobalAveragePooling was used along with 50% Dropout and sigmoid as activation fuction. All the hyper parameters were same as described above. The graph of loss and accuracy of training and validation is seen in figure 4. Accuracy, Sensitivity and Specificity are have been mentioned in table 2. The confusion matrix is shown in figure 5.

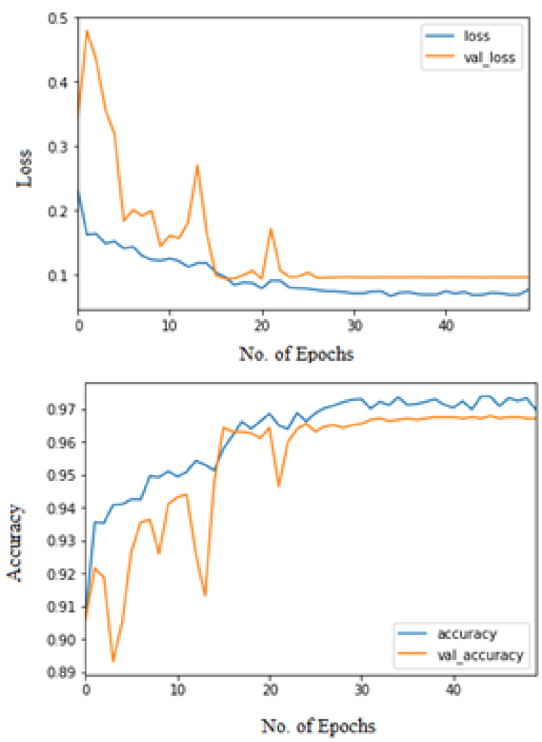

Figure 4: Loss and Accuracy Graphs for model made using pre trained network

Sr. No.	Parameters	Values
1	Training Accuracy	0.96
2	Validation Accuracy	0.96
3	Testing Accuracy	0.96
4	Sensitivity	0.94
5	Specificity	0.97

Table 2: Scores for model made using pre trained network

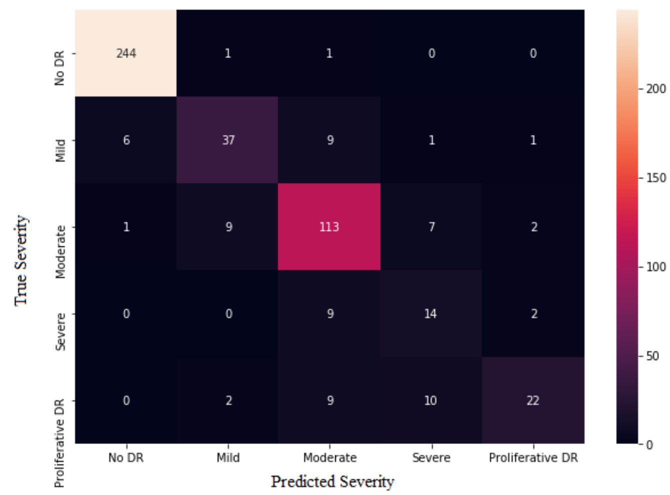

Figure 5: Confusion Matrix for model made using pre trained network

D. Graphical User Interface (GUI)

Using python module tkinter, a GUI was built for ease and simplicity so it can be used by the people who are not experts in the field. An image can be uploaded in the program and it will display the message of the severity of DR. The screenshot of the program is shown in figure 6.

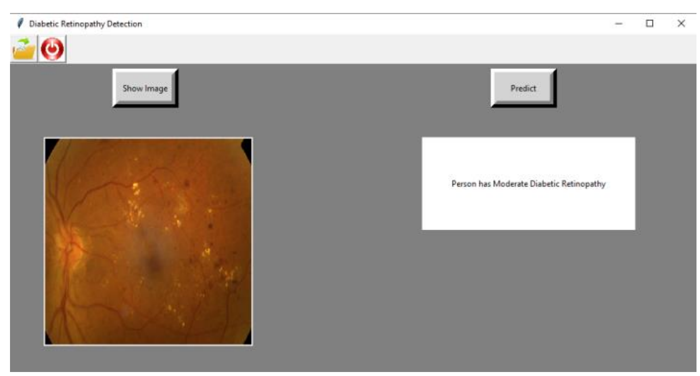

Figure 6: GUI of DR Detection

VI. Conclusion

In this paper, a brief review for the methods used for early detection of Diabetic Retinopathy was presented. Various processes involved in building the model were explained. The main use of this project is to make the model available in the rural areas such that doctors in rural parts can use this to detect the Diabetic Retinopathy in the early stages and help prevent permanent loss of vision.

VII. References

[1] "International Diabetes Federation", IDF diabetes atlas: 8th edn, https://www.diabetesatlas.org/(2017)

[2] "American Optometric Association", https://www.aoa.org/patients-and-public/eye-and-vision-problems/glossary-of-eye-and-vision-conditions/diabetic-retinopathy

[3] "Diabetes UK", https://www.diabetes.co.uk/diabetes-complications/diabetic-retinopathy.html

[4] Goh JK, Cheung CY, Sim SS, Tan PC, Tan GS, Wong TY. Retinal Imaging techniques for diabetic retinopathy screening. J Diabetes Sci Technol. 2016;10:282–94.

[5] "What Is Deep Learning AI? A Simple Guide With 8 Practical Examples", https://www.forbes.com/sites/bernardmarr/2018/10/01/what-is-deep-learning-ai-a-simple-guide-with-8-practical-examples/#5bf6762c8d4b

[6] Adarsh, P., Jeyakumari, D." Multiclass svm-based automated diagnosis of diabetic retinopathy. In: Communications and Signal Processing (ICCSP)", 2013 International Conference on. IEEE; 2013, p. 206–210.K. Elissa, "Title of paper if known," unpublished.

[7] Varun Gulshan, PhD; Lily Peng, MD, PhD; Marc Coram, PhD; Martin C. Stumpe, PhD; Derek Wu, BS; Arunachalam Narayanaswamy, PhD; Subhashini Venugopalan, MS; Kasumi Widner, MS; Tom Madams, MEng; Jorge Cuadros, OD, PhD; Ramasamy Kim, OD, DNB; Rajiv Raman, MS, DNB; Philip C. Nelson, BS; Jessica L. Mega, MD, MPH; Dale R. Webster, PhD, "Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs", 2016.

[8] Kele Xu, Dawei Feng, and Haibo Mi, "Deep Convolutional Neural Network-Based Early Automated Detection of Diabetic Retinopathy Using Fundus Image", School of Information and Communication, National University of Defense Technology, Wuhan 430019, China, 2017

[9] Manoj Raju, Venkatesh Pagidimarri, Ryan Barreto, Amrit Kadam, Vamsichandra Kasivajjala, Arun Aswath, "Development of a Deep Learning Algorithm for Automatic Diagnosis of Diabetic Retinopathy", International Medical Informatics Association (IMIA) and IOS Press, 2017.

[10] "Business Applications Of Convolutional Neural Networks", https://theappsolutions.com/blog/development/convolutional-neural-networks/

[11] Oquab M., Bottou L., Laptev I., Sivic J. Learning and transferring mid-level image representations using convolutional neural networks. In Proceedings of the IEEE conference on computer vision and pattern recognition, 2014:1717–1724.

[12] "Kaggle", https://www.kaggle.com/lextoumbourou/blindness-detection-resnet34-ordinal-targets

[13] Ravishankar, Hariharan & Sudhakar, Prasad & Venkataramani, Rahul & Thiruvenkadam, Sheshadri & Annangi, Pavan & Babu, Narayanan & Vaidya, Vivek. (2017). Understanding the Mechanisms of Deep Transfer Learning for Medical Images.

[14] Gao Huang, Zhuang Liu, Laurens van der Maaten, Kilian Q. Weinberger, "Densely Connected Convolutional Networks," arXiv:1608.06993v5 [cs.CV], Aug. 2016.