# Diabetic Retinopathy Detection Using Deep Learning

**Anand Kumar[1], Sumit Saikia[2], Srijan Saumya[3] , Moumita Dutta[4]**
[1, 2, 3, 4] Dept of Computer Science & Engineering
[1, 2, 3, 4] SRM Institute of Science and Technology, Chennai.

***Abstract***- *Diabetic Retinopathy is an eye disorder which causes vision blurriness and blindness in diabetic patients. Currently, detection of Diabetic Retinopathy involves manual methods in which physical examination is done by a trained eye physician. This consumes a lot of time of the physician which could have been devoted to other patients. This paper tries to tackle this issue by using computer vision to not only detect this disease, but also automating this procedure using neural network to give results of many patients within a short time frame.*

***Keywords***- Diabetic Retinopathy, Neural Network, Convolutional Neural Network, Supported Vector Machine.

## I. INTRODUCTION

Diabetic retinopathy is an eye disease caused by diabetes that can lead to loss of vision or even complete blindness. Diabetic retinopathy accounts for 12% of all new cases of blindness in the United States, and is the leading cause of blindness for people aged 20 to 64 years. If caught early enough, progression to vision impairment can be slowed if not altogether stopped, however, this is often difficult because symptoms may appear too late to provide effective treatment. Diabetic retinopathy (DR) has been estimated to affect about 93 million people globally, though only half are aware of it. There are four main stages of Diabetic Retinopathy; in its most advanced stage, abnormal blood vessels propagate on the surface of the retina, which can lead to scarring and cell loss in the retina.

Currently, diagnosing DR is a slow and arduous process that requires trained doctors to analyze color photographs of retinas. They then classify the level of deterioration the patient’s eye has experienced into one of four categories. While this process is effective, it is very slow. It takes about 2 days to get back results and after that time it may be harder to reach the patient. Furthermore, in areas where access to trained clinicians or suitable equipment is limited, individuals are left without any support. As the number of people with diabetes increases this system will become even more insufficient.

We propose a model for classifying retina images as having DR using convolutional neural networks trained with transfer learning. The input to the model is a pre-processed 256px x 256px retina image, and the output is a class label indicating whether or not the retina has DR.

## II. BACKGROUND

Historically, image analysis and classification has mostly focused on “low level image analysis tasks” like feature extraction and basic color normalization coupled with classical machine learning classification models like regression, SVMs and random forests. These algorithms would usually manage a small set of manually identified features (“in the order of tens”) which tend to limit their classification ability. Progress was made by the introduction of automated extraction of high dimensional sets of image features (on order of thousands). Dimensionality reduction techniques (e.g sparse regression) were then used to supply the construction of simple linear classifiers for the data [1].

More recently, leveraging Convolutional Neural Networks to perform image classification has become a very popular technique, particularly in the biomedical field. Their invariance to noise, orientation, image quality, lighting and intra-class variation offers a critical robustness that makes them especially suitable for biomedicine. Early work from the research group of JurgenSchmidhuber won the ICPR best paper awards in 2012 and 2013 by focusing on algorithmic work for mitotic figure detection using neural networks [2]. Furthermore, neural networks have been used in identifying metastic breast cancer, [1] brain lesion segmentation[4], cancer diagnosis [5] and other areas where humans have previously been forced to manually classify a test result.
A joint study [1] by Harvard Medical School and MIT worked on furthering techniques in Histopathological image analysis for Metastic breast cancer. Their approach obtained a near human-level classification performance using a 27-layer neural network architecture. Similarly, researchers from Cambridge University, Imperial College London, and others, [6] introduced an 11-layer 3D.

## III. DATASET AND PREPROCESSING

We use a dataset of retina images from a recent Kaggle competition 1.These are a set of high resolution retina images taken in a variety of conditions, including different cameras, colors, lighting and orientations. For each person we have an image of their left and right eye, along with a DR classification diagnosed by a clinician. There is considerable noise and variation in the data set due to these differing conditions. We describe our data preprocessing in the following paragraphs, and the number of samples we use for each data partition in the experiments section.

A key part of setting up our pipeline was preprocessing our data, the color retina images. Despite coming from Kaggle, which has a reputation for having clean data, these images required a hefty amount of preprocessing before we could use them in our neural network. The provided retina image were of different dimensions and resolutions, were taken by different cameras, were in different orientations, and were sometimes not even aligned or cropped similarly. The size of the dataset was also more than 38 gigabytes (35,126 images), which was intractable to handle with our computational resources.

To start, we had to transform the images in such a way that it would be feasible for a neural network or any learning algorithm to converge in a reasonable time. This consisted of resizing each image 256px by 256px. While this helped in making the necessary computation less intensive, it did not help with the fact that the lighting, orientation, and alignment were not similar across images.

To reduce this variation in the images, as has been the standard in other high-performing papers according to Andrej Karpathy, each image was rescaled to have the same radius (the eyeball) and each pixel had its color subtracted by the local average, mapping the average to 50% gray 2. A local average was used to account for the varying lighting conditions of the images, given that these images are taken by illuminating the retina, and an ill-aligned lighting source creates a gradient of illumination across the image that a local average can largely remove. The edges of the images were also clipped since there is a great variation on the boundaries or edges of the images. We then trained the convolutional neural networks with these preprocessed images.

## IV. IMPLEMENTATION

The diagnosis of diabetic retinopathy (DR) through colour fundus images requires experienced clinicians to identify the presence and significance of many small features which, along with a complex grading system, makes this a difficult and time consuming task.The expertise and equipment required are often lacking in areas where diabetic retinopathy detection is most needed. Most of the work in the field of diabetic retinopathy has been based on disease detection or manual extraction of features, but this paper aims at automatic diagnosis of the disease into its different stages using deep learning.

The use of using Deep CNN for detecting DR can increase the accuracy of its treatment, by reducing the scope of errors in detecting it. The treatment span can be reduced in this process because the detection of the problem and its types will be in a short period of time using this model.One of the challenges of the project was the limitations of the computational resources we had available. Deep CNN models take a long time to train, and the resources necessary to train them become expensive over time.To reduce this variation in the images, as has been the standard in other high-performing papers according to Andrej Karpathy, each image was rescaled to have the same radius (the eyeball) and each pixel had its color subtracted by the local average, mapping the average to 50% gray 2.

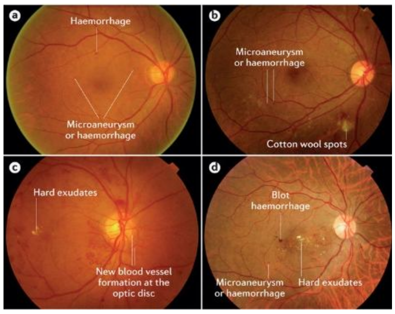

*Fig. 4.1 Retinal Image*

Develop a DR grading system capable of classifying fundus images based on location, number and type of retinal lesion.

Binary models were trained and tested on the Kaggle and MMDR dataset. Our model was the highest performing CNN.

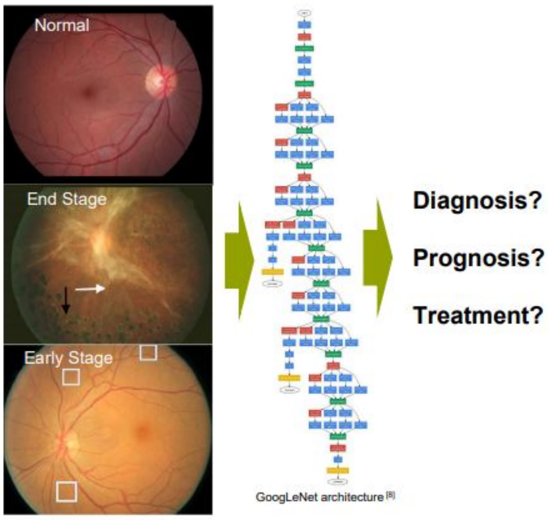


*Fig. 4.2 Training Neural Net*

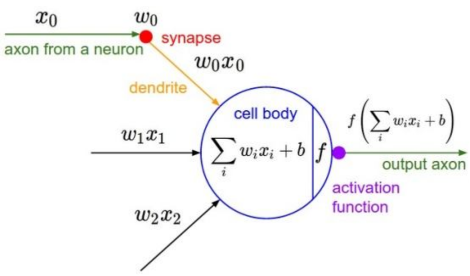


*Fig. 4.3 Activation Function*

## V. CONCLUSION

We present accurate deep convolutional network models for detecting both stage 4 diabetic retinopathy and referable diabetic retinopathy. Our model for detecting referable diabetic retinopathy yields better sensitivity and specificity than family doctors. Across all models, using data augmentation improved the resulting accuracy of the model. This is likely due to its ability to reduce overfitting by artificially increasing the size of the training set, given that smaller test sets are easier to overfit. Additionally, training the top two blocks of the Inception V3 network in addition to the final two fully-connected layers improved performance relative to just training the final two layers. This is likely because the top blocks in the base model represent higher-order features that are more specialized to the ImageNet task the model was originally trained for. So, re-training these layers likely allows us develop higher-level features specific to this task, increasing performance.

We obtained our best results using the Inception V3 model and ImageNet pre-trained weights with two additional fully-connected layers, training the two additional layers and the top two blocks of the base model. Given the degree to which data augmentation improved our results, it might be promising to explore other means of augmentation. One option to consider would be adding small random rotations to our training images, in addition to the flipping, noise, zoom, and shear transformations we applied.

One of the challenges of the project was the limitations of the computational resources we had available. Deep CNN models take a long time to train, and the resources necessary to train them become expensive over time. There are multiple network architectures and preprocessing/data augmentation techniques that could have been explored had there been more computational resources available. Because of these time and computational constraints, there was a tradeoff between training time and accuracy. Training time could have been sped up by shrinking the image size, but cursory experiments indicated that this would result in markedly lower accuracy. Furthermore, higher accuracy could have been achieved through randomized search on the hyper parameter space had these resources been available.

## VI. ACKNOWLEDGEMENT

We are greatly indebted to our college, SRM IST, Ramapuram, for providing us a healthy environment that is conducive to learning that made us reason out, question and think out of the box. It has also strengthened our confidence and has enabled us to conceptualize, realize our dream project. We are also grateful for the support provided by Dr.J.Jagadeesan, Head of the Department, Computer Science & Engineering, SRM IST, Ramapuram for providing us with such an opportunity.

## REFERENCE

[1] Dayong Wang Aditya Khosla, Rishab Gargeya Humayun Irshad Andrew H Beck Beth Israel Deaconess Medical Center, Harvard Medical School CSAIL, *Massachusetts Institute of Technology, Deep Learning for Identifying Metastatic Breast Cancer*.

[2] D. C. Cires¸an, A. Giusti, L. M. Gambardella, and J. Schmidhuber, Mitosis detection in breast cancer histology images with deep neural networks. In Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013.

[3] Gill, James M., et al. "*Accuracy of screening for diabetic retinopathy by family physicians*." The Annals of Family Medicine 2.3 (2004).

[4] Havaei, Mohammad, et al. "*Brain tumor segmentation with deep neural networks.*" Medical Image Analysis (2016).

[5] Fakoor, Rasool, et al. "*Using deep learning to enhance cancer diagnosis and classification."* Proceedings of the International Conference on Machine Learning. 2013.

[6] KonstantinosKamnitsasa , Christian Lediga , Virginia F.J. Newcombeb,c , Joanna P. Simpsonb ,Andrew D. Kaneb , David K. Menonb,c, Daniel Rueckerta , Ben Glockera, *Efficient Multi-Scale 3D CNN with fully connected CRF for Accurate Brain Lesion Segmentation*.

[7] Oquab, Maxime, et al "*Learning and transferring mid-level image representations using convolutional neural networks."* Proceedings of the IEEE conference on computer vision and pattern recognition. 2014.

[8] Abràmoff, Michael D., et al. "Automated analysis of retinal images for detection of referable diabetic retinopathy." JAMA ophthalmology 131.3 (2013).

[9] Oquab, Maxime, et al. *"Learning and transferring mid-level image representations using convolutionalneural networks."* Proceedings of the IEEE conference on computer vision and pattern recognition. 2014.

[10] Szegedy, Christian, et al. *"Rethinking the inception architecture for computer vision."*arXiv preprint arXiv:1512.00567 (2015)

[11] Srivastava, Nitish, et al. *"Dropout: a simple way to prevent neural networks from overfitting."* Journal of Machine Learning Research 15.1 (2014): 1929-1958.

[12] Kingma, Diederik, and Jimmy Ba. "*Adam: A method for stochastic optimization.*" arXiv preprint arXiv:1412.6980 (2014).