DIABETIC RETINOPATHY DETECTION using DEEP LEARNING

Ashmit Pareekh  
Department of Electronics and Telecommunication Engineering,  
Pimpri Chinchwad College of Engineering and Research,  
Pune - 412101, India.

Hrishikesh Patil  
Department of Electronics and Telecommunication Engineering,  
Pimpri Chinchwad College of Engineering and Research,  
Pune - 412101, India.

Mohit Rade  
Department of Electronics and Telecommunication Engineering,  
Pimpri Chinchwad College of Engineering and Research,  
Pune - 412101, India.

Prof M. Andhare  
Department of Electronics and Telecommunication Engineering,  
Pimpri Chinchwad College of Engineering and Research,  
Pune - 412101, India.

Abstract

Diabetic retinopathy or commonly abbreviated as DR is a condition in diabetes brought about by change in the vessels of the retina and one of the significant reasons for visual impairment in the dynamic populace because of diabetes. Numerous issues brought about by DR can be forestalled by ideal treatment. Since the diabetic retinopathy progresses through several stages, it is really difficult for diabetic retinopathy detection and grading of such images is a time-consuming manual process.

In this work we are trying to find an automatic way to classify the given image into different stages of diabetic retinopathy. This project uses convolutional neural network (CNNs) power for DR detection and uses a publicly available dataset of Aptos competition available on Kaggle for training these models. The dataset had around 3663 high definition fundus images, each of resolution 4000 * 2000. At the end the algorithm is able to classify the given image into different stages of DR.

Keywords : Deep Learning, Diabetic retinopathy, Convolutional neural network

1. Introduction

Diabetic retinopathy is a medical condition in which the retina of an individual having had diabetes mellitus for several years gets damaged. The abnormalities that characterise diabetic retinopathy appear in a predictable order [1]. In its early days diabetic retinopathy is usually symptomless but as it progresses, the formation of new abnormal blood vessels takes place inside the retina. Based on the progression of symptoms, diabetic retinopathy is of two types,

a) Non proliferative diabetic retinopathy - The basic stage, usually symptomless.

b) Proliferative diabetic retinopathy - Advanced stage, abnormal growth of blood vessels takes place.

Diabetic retinopathy is best diagnosed with a dilated eye exam[2]. For this exam, a doctor uses drops placed in your eyes which widens(dilate) the pupils of an eye which allows the doctor to have a better view inside your eyes. These drops may cause your close vision to blur until they wear off, several hours later. During the exam, your eye doctor will look for the following abnormalities :-

- Abnormal blood vessels
- Swelling, blood or fatty deposits in the retina
- Bleeding in the eye (vitreous)
- Retinal detachment
- Abnormalities in your optic nerve
Based on the observation, the classification of DR is shown in table 1.

Early detection of diabetic retinopathy condition is critical for good prognosis relies on skilled clinicians and is a time-intensive process. This is a challenge in areas that lack access to a high-tech clinical facility. Also, as the nature of DR the grading of DR affected eyes depends on the expertise of the examiner. At last as there is an increase in numbers of both, diabetic cases and diabetic retinopathy all around the world, the manual methods for diagnosis of diabetic retinopathy may not be able to keep pace with demand for testing services.

Deep learning or hierarchical learning, is a type of machine learning which uses a layered algorithmic architecture to analyze data. In deep learning models, data is passed through a stack of layers which form a neural network, each layer uses the output obtained from the previous layer to produce its results. These layers use some mathematical functions that turn raw input into meaningful output. Deep learning models become more accurate as they process more data, the network actually learns from previous results to extend their ability to form correlations and

Table 1.- Classification of Diabetic Retinopathy

<table>
<thead>
<tr>
<th>Proposed Disease Severity Level</th>
<th>Observed abnormalities</th>
</tr>
</thead>
<tbody>
<tr>
<td>No diabetic retinopathy</td>
<td>No abnormality</td>
</tr>
<tr>
<td>Mild DR</td>
<td>Only 1 microaneurysm is present</td>
</tr>
<tr>
<td>Moderate DR</td>
<td>More than just microaneurysms but less severe NPDR</td>
</tr>
<tr>
<td>More than just microaneurysms but less severe Nonproliferative Diabetic Retinopathy</td>
<td>International Classification</td>
</tr>
<tr>
<td></td>
<td>Any of the following and no signs of PDR</td>
</tr>
<tr>
<td></td>
<td>● More than 20 intraretinal hemorrhages in each of four quadrants</td>
</tr>
<tr>
<td></td>
<td>● Definite venous beading in two or more quadrants</td>
</tr>
<tr>
<td></td>
<td>● Prominent IRMA in one or more quadrants</td>
</tr>
<tr>
<td>Proliferative Diabetic Retinopathy</td>
<td>One of both of the following:</td>
</tr>
<tr>
<td></td>
<td>● Neovascularization</td>
</tr>
<tr>
<td></td>
<td>● Vitreous/preretinal hemorrhage</td>
</tr>
</tbody>
</table>

connections. These layers of neural networks can perform several tasks, which include prediction, classification etc. For processing image data, Convolutional neural networks also called CNN are used. In CNN models data is filtered through a cascade of multiple convolution layers, with each successive layer using the output from the previous one to inform its results. It is capable of performing segmentation, object (i.e. lesion or region of interest) detection and classification.
2. Related work

CNNs are most commonly used in object segmentation, object recognition and classification. Hence a variety of models have been developed till now. One such model named Alexnet which was proposed by Alex Krizhevsky, Ilya Sutskever and Geoffrey Hinton had 60 million parameters won the 2012 ImageNet competition with a top-5 error rate of 15.3%, compared to the second place top-5 error rate of 26.2%. Another architecture, VGG16 had around 138 million parameters and it became a runner up in the ImageNet contest in 2014. The winner of ILSVRC, GoogLeNet architecture is also known as Inception Module goes deeper in parallel paths with different receptive field sizes and achieved a top-5 error rate of 6.67%. In 2015, Resnet achieved a top 5 error rate of 3.6%.

Efforts have been made to implant deep learning models for medical diagnosis. Research conducted[4] used two fundus image datasets, Messidor-1 and kaggle dataset in which images were graded from 0 to 4 based on the severity of DR. Several CNN architectures were tested with particular focus on GoogLeNet. The dataset was split into 9:1 ratio into train set and validation set. The model achieved a validation accuracy of 66%.

3. Proposed work

The aim of our work is the detection of DR and classification of a given fundus image into 5 classes indicating the severity of diabetic retinopathy as shown in table 1 before. Generally, the procedure involved in solving such problems is as shown in figure 1.

First, the image is acquired from source. Then preprocessing techniques are applied to the images. The preprocessing is done to remove noise from images which makes the images more suitable for the training process. Pre-processing techniques involve resizing, reduction in noise and image contrast.

Later the image set will be split into 2 sets: train set and validation set. The train set will be used for training the model. During the training phase, the model will learn the parameter and will try to classify the images into the five different classes. Once the training is complete, the hyperparameters are tuned to make the model more accurate. Once the model of optimum accuracy is obtained, the made model was used to predict some sample images from validation set and confusion matrix was used to assess the performance.

3.1 Data acquisition

The data for training of a model was obtained from a kaggle competition[3]. The dataset had around 3662 images. The data obtained was highly unbalanced and had around 1800 images out of 3662 of grade 0. The figure 2 shows the image distribution of the dataset.
3.2 Data processing

The images obtained from the source were somewhat noisy. And as shown in figure 3, the red plane of image was noisy while the blue plane did not have a great amount of information, most of the information was present in the green plane of image so our first step was to discard the red and blue plane of the image.

The next step in preprocessing was the use of CLAHE (contrast limited adaptive histogram equalisation), this was done because the images of the dataset were dark, CLAHE improves the contrast in image. After using CLAHE, the resultant images had noise, so median filtering was applied to the images. Figure 5 shows image with CLAHE while figure 6 shows image after applying median filtering to CLAHE’d image.
3.3 Splitting of dataset and training of model :-

The dataset was splitted into a 90% training set and 10% validation set. After splitting the dataset, the model was trained on the dataset. The model used was VGG16, which had 13 convolutional layers with filters of size 3x3, 5 Max Pooling layers, followed by 4 fully connected layers.

4. Results

A total of 362 images were used for testing the developed model. The images had to be preprocessed just like the train set by which noise and uniform illumination was obtained. The prepared model showed the test accuracy of 80% with a loss of 0.4994, precision, recall, f1-score and support score in prediction and classification of given images into 5 grades is shown in figure 6.

![Figure 6 : Performance of model](image)

5. Conclusion

If a clinician manually does the detection of DR, then he has to go for detection of individual features. Detection of features is a time intensive process and a skilled clinician is required to grade the image, as a human is involved, it can be prone to errors. Our system gives a way to reduce time required for analysing the fundus image and grade the image instantly. Our model was trained on a smaller dataset of 3662 images which had to be processed for better visibility of features and we obtained performance as shown in figure 6.

The performance of our cnn model is satisfactory and can still be improved. For training our model we used a dataset of 3662 images, however datasets having even more images are available and can be tested. Further to denoise the images some more steps are required. The prepared model can be tuned, so the size of the end model is large and it is difficult to deploy so some different architecture of CNN model with the same results would be appreciable.
6. References


