Detection and Classification of Diabetic Retinopathy Using Convolutional Neural Networks and Fundus Images

Udayan Birajdar, Sanket Gadhave, Shreyas Chikodikar, Shubham Dadhich
Departmnet of Computer Engineering
NBN Sinhgad School of Engineering

Abstract

There are countless number of people with diabetes around the world. Diabetic Retinopathy (DR), a major complication of diabetes, is a retinal disease that results in blindness. Early detection can prevent or delay the loss of vision due to DR. For early detection, it is essential for diabetic patients to undergo frequent retinal tests. Diagnosing DR using fundus images is a complex process and thus difficult and time consuming task requiring professional expertise to identify if the DR traits are present in eye. We propose a system that can provide immediate feedback thus effectively simplifying the process. Convolutional Neural Network (CNN) architecture based on fundus image database to accurately diagnose DR with minimum efforts on the part of patient and doctor is developed. The database contains multiple cases of retinal hemorrhage, micro-aneurism, cotton-woolspots, etc. The retinal image of the patient is fed to the system where image is processed to detect signs of DR. The network is also trained to classify various stages of DR depending on severity such as nonproliferative DR or proliferative DR. The output contains a report detailing if DR is present or not in the retinal image in question and if present, which stage it is.

Key words: Diabetic Retinopathy , Convolution Neural Networks, Artificial Intelligence

I. INTRODUCTION

It is estimated that 415 million people are living with diabetes in the world, which is estimated to be 1 in 11 of the world's adult population. Diabetes is a disease in which the body's ability to produce or respond to the hormone insulin is impaired, resulting in abnormal metabolism of carbohydrates and elevated levels of glucose in the blood. One of the major areas where we can observe the effect of diabetes is the human eye. This phenomenon in medical terms is called as diabetic retinopathy. If left untreated diabetic retinopathy can lead to permanent blindness. Thus early detection plays a crucial role in treating diabetic retinopathy. Traditional methods to detect diabetic retinopathy is complex and includes lot of efforts on the part of patient and doctors hence the proposed system helps in detecting diabetic retinopathy without the patient requiring to be physically being present. The proposed system works on Convolutional Neural Networks in which a database of human eyes affected with diabetic retinopathy are used for training. The database considered is MESIDOR database which consists of over two thousand images. The trained model will detect the presence of diabetic retinopathy and classify it based on severity. The classification is based on four categories which include proliferate diabetic retinopathy which is further divided into three types :low, medium and high and non-proliferate diabetic retinopathy which indicates that there is no problems with the eye. The programming language used is python and libraries used are tensor flow and keras. The tensorflow model is trained on graphics processing unit to achieve faster training. The final model can successfully classify the fundus image of human eye based on DR levels.

II. LITERATURE REVIEW

Sr No.	Title	Publication	Technique Used	Problem Identified
1.	Detection of Diabetic Retinopathy and Catract by vessel extraction	IEEE, 2017	Retinal Image Segmentation	• Only gives binary output(whether yes or no)
2.	Fundus Image Texture Features Analysis in Diabetic Retinopathy Diagnosis	IEEE, 2017	Naive Bayes , SVM and KNN	• Unable to distinguish AMD and DR
3.	Diabetic Retinopathy Screening Based on CNN	IEEE, 2017	Convolutional Neural Networks and Machine Learning	• Accuracy upto 77%
4.	Red Lesion Detection using Dynamic Shape Features for DR	IEEE, 2015	Computer Aided Diagnostic and ROI extraction	• Automatic DR Grading unavailable • Less Precision

In 'Detection of diabetic Retinopathy and Catract by vessel extraction'[1] retinal segmentation is used to extract various features like haemorrhages that help to detect disease and do treatment on them. The segmentation used in this particular paper consists of green filter and then using of clahe algorithm to enhance the features. Furthermore Support Vector Machines are used to detect the Diabetic Retinopathy and also Feature Vector Machine is used. The drawback of this research paper is that , it only gives binary output that is whether diabetic retinopathy is present or absent. No information or detailed report is presented indicating the amount of DR or the severity of diabetic retinopathy.

In 'Fundus Image Texture Features Analysis in Diabetic Retinopathy Diagnosis' [2] the paper investigates texture feature capabilities from fundus images to differentiate between DR , age related macular degeneration (AMD) and normal. In the four experiments are were designed for two types of databases namely DIARETDB0 and STARE. The classifiers used are Naïve Bayes , SVM and KNN.In case of multiclass classification for images , distinguishing AMD and DR has been a challenge and therefore gives lower accuracy.

In 'Diabetic Retinopathy Screening Based on CNN; [3] A Convolutional Neural Network is used as a classifier. Initially the database used is MESSIDOR which contains images of diabetic retinopathy affected eyes. These images are preprocessed. In preprocessing different techniques are used and finally the results of each techniques are calculated. A few image processing techniques used include : AHE , Gauss Noise , Grayscale Images , Green Channel , Mixed Transformation . The best result obtained is from mixed transformation. However the accuracy obtained is near to 80% which is quite less. Also multiclass classification is not given in this paper which is also a drawback as only binary output is present.

In 'Red Lesion detection using dynamic shape features for diabetic retinopathy ' [4] ,a novel method for automatic detection of both microaneurysms and hemorrhages in color fundus images is described and validated. The main contribution is a new set of shape features, called Dynamic Shape Features, that do not require precise segmentation of the regions to be classified. These features represent the evolution of the shape during image flooding and allow to discriminate between lesions and vessel segments. The method is validated per-lesion and perimage using six databases, four of which are publicly available.On the Messidor database, when detecting images with diabetic retinopathy, the proposed method achieves an area under the ROC curve of 0.899, comparable to the score of human experts, and it outperforms state-of-the-art approaches. However less precision is available in this system and automatic DR grading is also not present causing it to fail multiclass classification as well.

III. CONVOLUTION NEURAL NETWORKS

Convolution Neural Networks or CNN is a part of detection of different images and recognition of other images which the computer has never seen. CNN plays an important role in image recognition as it is more accurate than any other system yet developed . The working of CNN takes place through 4 layers. Similar to neural networks there are deep CNN learning models. The three layers of CNN include Convolution , Rectified Linear Unit or ReLu , Pooling and finally Fully Connected Layer artificial neural network layer.

An image is nothing but an array of pixels. In a normal computer system what happens is when an image is given for comparison and recognition then a normal computer compares the pixel densities and checks for any similarities. If there are similarities then it is given as a match or not a match. However it may happen that 2 images are similar but the actual images have different colours.This means that the pixel densities do not match but they are similar. This gives an inaccurate answer that images do not match but in reality they do.

The functioning of convolution layer is totally different than normal working of computer. In convolution layer what happens is if a particular data set of images is given to train then the CNN first shrinks the images so that the training process of neural network can happen efficiently. The shrinking of images is done by all the three layers stacked together and finally this data is then forwarded to fully connected layer which has output as the classification.

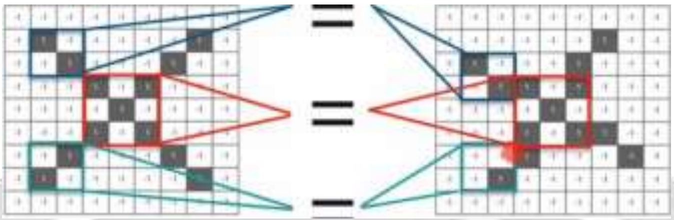

Figure 1: Feature Extraction using CNN

a. Convolution Layer

This is the first Layer of CNN , the convolution layer. In convolution layer once the features are extracted the system lines up the features and the actual image and then the multiplication of corresponding pixel happens and then there is addition of pixels and division by number of pixels. This step happens across the whole image and finally a new table consisting of all numeric values is obtained. This new table obtained is containing values of pixels with the features extracted which helps for further functioning. Consider the Figure number 2. The necessary features are extracted then the values of pixels are compared with each pixel, the pixels are added and then divided by total number of pixels. This gives the output of convolution layer. This process is done with the help of a 3x3 filter. Such a process helps to reduce the image without loss of any features.

Figure 2: Convolved Image

b. Rectified Linear Unit Layer

The ReLu layer is Rectified Linear Unit Layer. The working of this layer is simple. After the feature extraction the ReLu Layer deactivates the pixels which are not necessary and only keeps the pixels which are important. From the output of Convolution Layer we get positive as well as negative values. The positive values include successful finding of features and negative values include that those areas are not much important. The ReLu Layer takes the data of the pixels and if the value is positive then the value is kept as it is and if it is negative then it is converted to a zero. The figure 3 shows the exact working. If the value of x is greater than 0 then it retains the value and if it is below 0 then it returns 0.

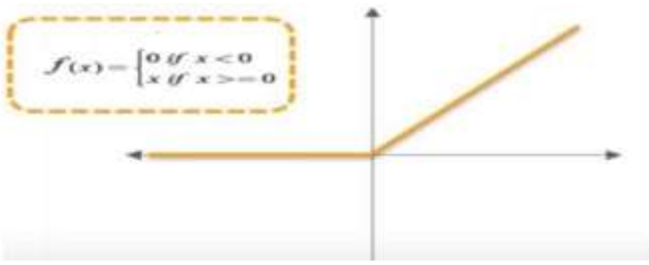

Figure 3 : ReLu Layer

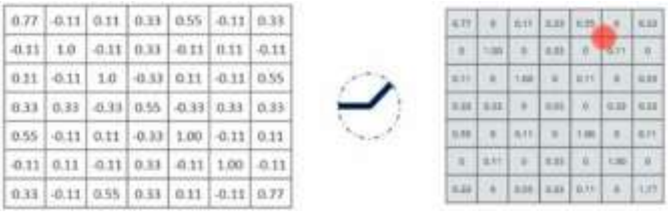

Figure 4 : ReLu Example

c. Pooling Layer

The Pooling layer does a simple job of reducing the input size. For the pooling layer generally a stride is selected. The size of stride is 2x2 or 3x3. That stride is walked over the image and the maximum value from the image is extracted and stored in a different table. This helps to reduce the size of the input data a lot. The pooling can be max pooling or min pooling. The max pooling takes the maximum value of the stride whereas min pooling takes minimum value. Figure 5 shows max pooling layer.

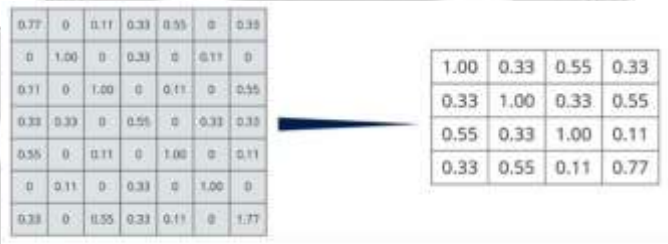

Figure 5 : Max Pooling

d. Layer Stacking

Finally this process of Convolution ReLu and Pooling is carried out once more or any number of times as long as the input becomes small enough for processing. This is carried out by stacking the layers one over the other shown in the image below. After Stacking the layers the output obtained is in numerical form. This output is given as input to the neural network and predictions are made with the help of it. This is how Convolution Neural Networks work.

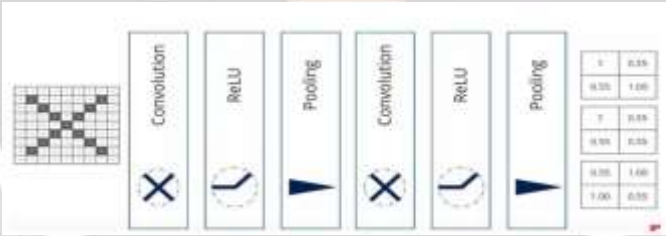

Figure 6 : Layer Stacking

e. Fully Connected Layer

This is the final layer in convolutional neural network . It takes the input as numbers from previous layers and feeds the data to the neural network. The fully connected layer consists of multiple neurons and a connection of those neurons to each other. The output of fully connected layer consists of number of classifications. It consists of feed forward neural network and error back propagation to reduce the errors and get better accuracy. It helps to fine tune the weights and biases.

V. SYSTEM ARCHITECTURE

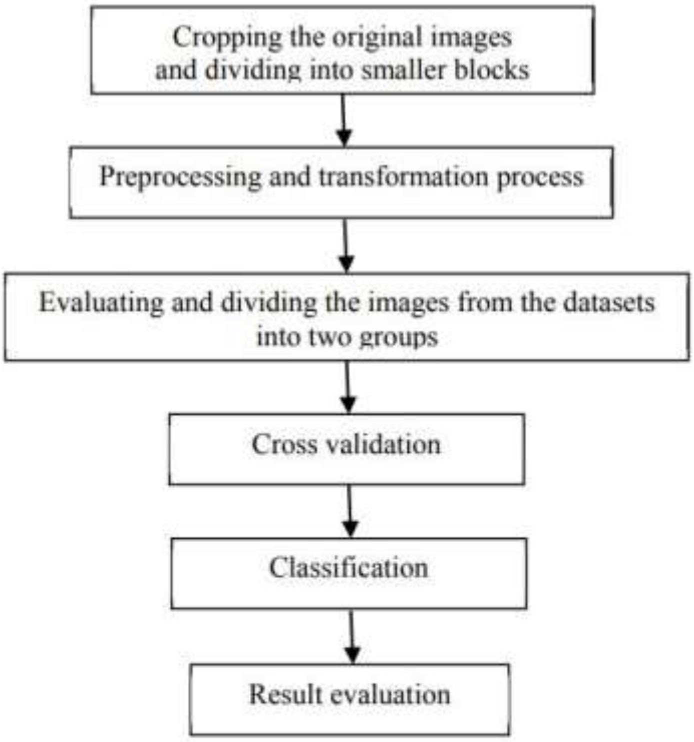

Figure 7 : Flowchart of the system

i. Getting the data sets is one of the tricky jobs. In the proposed paper the dataset is gathered from various online free databases like MESIDOR and also from some doctors of India. Some of the entities in the dataset have been changed or reversed so as to have a variety in the dataset.
ii. Image Processing forms a small part of CNN . As clear and fine images need to be provided to the Neural Network. The images are cleaned and they are resized to fit properly. Also the images are converted to gray scale instead of RGB to ease the job.
iii. According to the shape of various veins in the eyes the region of interest is calculated by the CNN. Also depending on the hemorrhages and cotton wool spots the disease intensity is detected.
iv. In order to extract these features from the images the images the proposed model uses CNN which automatically extracts the features by using convolution.
v. Multiple layers are stacked for the CNN the layers are shown in the Figure 8.
vi. The cross entropy function is used for calculating the loss of the output and the learning rate is set to 0.0001
vii. The Neural Network is trained for 20 epochs that is 20 cycles and then it is ready to make predictions.

Layers	Description
Input Layer	24x24x3 images
Convolution-1	Convolution and rectified linear activation (ReLU).
Pool-1	Max pooling.
Convoution-2	Convolution and rectified linear activation.
Pool-2	Max pooling.
Local-3	Fully connected layer with ReLU
Local-4	fully connected layer with ReLU
softmax	Classification result

Figure 8 : Layers used in proposed CNN

VI. MATHEMATICAL MODEL

Let S be Closed system defined as,
S = { Ip, Op, Ss, Su, Fi, A}

Training a Convolutional Neural Network by training it using images of eyes having DR and performing various actions from the set of actions A so that Su state can be attained.

S ={Ip, Op, Ss, Su, Fi, A}

Where,

Ip1= {Dtr | Dtr € D }
Where,
Dtr – Images of an eye
D – Dataset containing Images of eyes

Output set :
Op1= { F | F is Acceptance of relevant Data }
Op2= { F' | F' is a set of features selected for extraction }
Op3= { F'' | F'' is a set of features learned by the CNN through Training }
Op4= { c | c is a trained model used to make predictions}

Set of actions = A={F1,F2,F3,F4}
Where,
F1 = Accepting Relevant Data
F2 = Using Convolution Layer to extract features
F3 = Training the Deep Neural Network
F4 = Model Architecture Successful

Ss- Loading Images → Extracting Features → Training Model → Making Predictions

Su- Success state is when the CNN can accurately detect and classify an eye affected with Diabetic Retinopathy.

Fi- Failure state is when the trained model makes an incorrect prediction.

VI. CONCLUSION

This paper represents a new efficient way for detection and classification of diabetic retinopathy for the human eye. With the help of CNN the features are easily extracted. The proposed system may give accuracy for up to 85%. Also this system acts supplementary for the doctors and does not focus on replacing the doctors. Furthermore it can also perform multiclass classification and gives report about the intensity of Diabetic Retinopathy.

VII. FUTURE SCOPE

As this system if proposed only for laptops our prime future scope is to implement such a system on android devices . If the classification and detection is done on handheld smartphones it will be even more beneficial and effective. Another future scope is to detect other diseases related to DR other than just PDR or NPDR.

VIII. REFERENCES

[1] Prof.Dr.S.S.Chorage, Sayali S. Khot. (2017). Detection of Diabetic Retinopathy and Cataract by Vessel Extraction from Fundus Images. *International Conference On Electronics,Communication and Aerospace Technoogy.*
[2] Devvi Sarwinda , Alhadi Bustamam , Aniati M. Arymurthy. (2017). Fundus Image Texture features analysis in Diabetic Retinopathy Diagnosis. *International Conference On Sensing Technology ,IEEE*
[3] Lama Seoud, Thomas Hurtut, Jihed Chelbi,Farida Cheriet and J.M.Pierre Langlois. (2015). Red lesion detection using Diabetic Retinopathy Screening. *IEEE Transactions on medical Imaging* .
[4] Monika Andonova, Jarmila Pavloviþova, Slavomir Kajan, Milos Oravec, Veronika Kurilova. (2017). Diabetic Retinopathy scanning based on CNN. *5th International Symposium* .