DIABETIC RETINOPATHY DETECTION USING IMAGE-PROCESSING

Supervisor:Professor Dr.Md. HaiderAli
Co-supervisor: Mohammad Hammad Ali

Asif UzZaman (Asif) **12301018**

Shadaab Kawnain Bashir **13301092**

DepartmentofComputerScienceandEngineering

SchoolofEngineeringandComputerScience

BRACUniversity

Submittedon: 20th April2016

DECLARATION

We,herebydeclarethatthisthesisisbasedontheresultsfoundbyourselves.Materialsof work foundbyotherresearcheraremen tionedby reference.ThisThesis,neither inwholeor in part,hasbeen previouslysubmitted foranydegree.

Signatureof Supervisor

SignatureofAuthor

ProfessorDr.Md.HaiderAli

Asif UzZaman (Asif)

Shadaab Kawnain Bashir

ABSTRACT

Diabetic retinopathy is a leading problem throughout the world and many people are losing their vision because of this disease. The disease can get severe if it is not treated properly at its early stages. The damage in the retinal blood vessel eventually blocks the light that passes through the optical nerves which makes the patient with Diabetic Retinopathy blind. Therefore, in our research we wanted to find out a way to overcome this problem and thus using the help of convolutional neural network (ConvNet), we wereable to detect multiple stages of severity for Diabetic Retinopathy.There are other processes present to detect Diabetic Retinopathy and one such process is manual screening, but this requires a skilled ophthalmologist and takes up a huge amount of time. Thus our automatic diabetic retinopathy detection technique can be used to replace such manual processes and theophthalmologist can spend more time taking proper care of the patient or at least decrease the severity of this disease.

ACKNOWLEDGEMET

This thesis was suggested by Lecturer Mohammed Hammad Ali, and was accepted by Professor Dr. Md. Haider Ali, Chairperson, Department of Computer Science and Engineering, BRAC University. This is the work of Asif UzZaman (Asif) and Shadaab Kawnain Bashir students of BRAC University, studying Computer Science and Engineering from the year 2012. The document has been prepared as an effort to compile the knowledge obtained by us during these four years of education and produce a final thesis which innovatively addresses one of the issues of research in computer vision.

We would like to express our gratitude to Almighty Allah (SWT) who gave us the opportunity, determination, strength and intelligence to complete our thesis.

We would like to thank our supervisor, Professor Dr. Md. Haider Ali sincerely and co-supervisorMohammed Hammad Ali for their consistent supervision, guidance and unflinching encouragement in accomplishing our work.

We also want to acknowledge Ms. Dilruba Showkat, Ex-Lecturer, Department of Computer Science and Engineering, BRAC University for convincing us to work with computer vision.

Last but not least, our gratitude towards BRAC University and our supervisor for allowing us to use one of the PC's from graphics lab.

Contents

Table of Figures

1. INTRODUCTION

Diabetes is a chronic disease and the numbers of diabetes patients are increasing at a very rapid pace, which may eventually lead to vital organ failure. In most cases, it may affect heart, kidney and there can be complications in the eye. As it a metabolic disease, therefore the body is unable to produce insulin which eventually increases the glucose level in the blood. When the glucose level of the blood vessel in retina is increased the vision becomes blurred and without proper treatment it can lead to complete blindness, this process of damage within the retina is called diabetic retinopathy. Excess amount of glucose in the blood vessel may lead to anomalies like microaneurysms, hemorrhages, hard exudates and cotton wool spots develop during the different phases of diabetic retinopathy [1, 2].

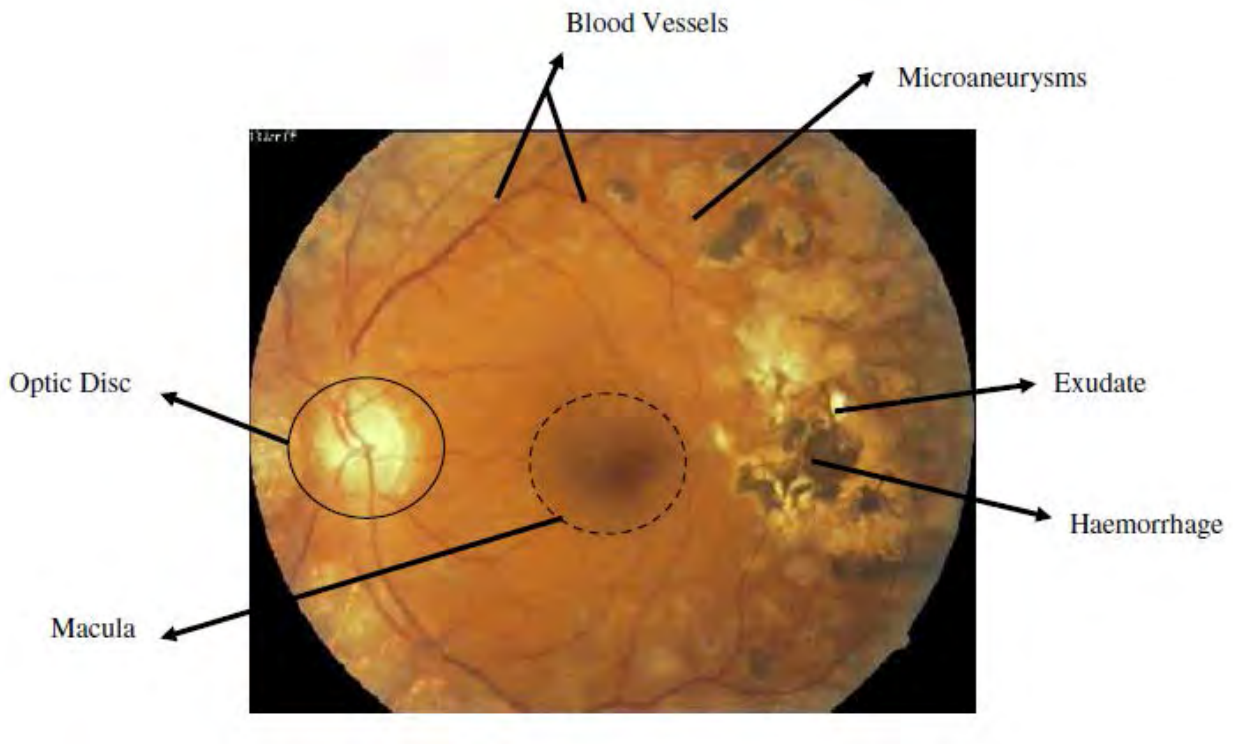

Figure 1: Features in DR image [3]

According to a study which was conducted by World Health Organization, it shows that the number of diabetes patients will increase from 130 million to 350 million over the next 20 years [4].In developed countries, one of the alarming cause of blindness is diabetic retinopathy [5] and for the developing countries this problem is even more dangerous as they do not have the proper screening technologies to overcome the prevention from this disease, however 75% of the people with diabetic retinopathy lives in the developing countries[6]. The symptoms for diabetic retinopathy do not show up in the early stages, which makes it even harder for the ophthalmologist to prevent the patient from being blind. The number of patient with diabetic retinopathy is increasing, which will increase the workload for the ophthalmologist because most of their time will be spent to detect diabetic retinopathy. As a result, they will not be able to take care of the patient with their full potential.

The main goal of our thesis is to make the diabetic retinopathy detection system automated so that the specialist can take proper care of their patients and do not have to worry about the detection process.

1.1 MOTIVATION

The main reason behind choosing this topic for our thesis is that we live in a developing country where there is always shortage of resources to overcome any problem. Our country has a population of 16 million and around 3.2 million people have diabetes, from where approximately 50% of them have diabetic retinopathy [7]. In developing countries like Bangladesh people usually treat themselves with a whole body checkup for diabetes but they do not treat their eyes because they have a little knowledge about the fact that diabetes can affect their eyes. Therefore we want to raise an awareness that diabetic retinopathy can be very dangerous if it is not treated properly in its early stages.

With the help of image processing, we want to help our medical facilities so that the detection process becomes easier and none of the patient goes blind because of diabetic retinopathy. Therefore, using machine learning we are going to train our dataset to give the best possible outcome to the ophthalmologist so that they can worry less about the detection of diabetic retinopathy and focus more on the proper treatment of the patients. Since everything around us is getting digitalized, we wanted to incorporate the medical sector with computer science and thus bring better changes in the people's life.

1.2 THESIS OUTLINE

Chapter 1 is the formal introduction of the thesis. We have discussed our motivation and approach towards our proposed topic in this chapter.

Chapter 2 is the background study that covers all the important basics needed for this research along with their formal definitions and representations. In this chapter we have described the details about Diabetic Retinopathy, concepts of Convolutional Neural Network and the VGGNet architecture.

Chapter 3 focuses on the proposed work. Firstly, we have described our whole workflow to achieve the results following our proposal. Then we have discussed about our dataset and data processing. Later, we have elaborated our Convolutional Neural Network architecture that we have used. At the end, we discussed about the hardware and toolkits that were used.

Chapter 4 is the result analysis part where we have explained our acquired results following our approach towards proposed work. We have attached our accuracy, sensitivity and specificity graph in this chapter.

Chapter 5 we have talked about the limitations of our project as well as, we mentioned about some approaches to overcome those limitations.

Chapter 6 we have discussed our future goals regarding this research.

2. BACKGROUND STUDY

2.1 DIABETIC RETINOPATHY

Diabetes occurs when our body is not being able to produce sufficient insulin therefore it leads towards high glucose level which in many cases causes damage in the blood vessels of retina, which may cause blindness and this process is what we know as diabetic retinopathy.The two types of diabetic retinopathy are NPDR (nonproliferative diabetic retinopathy) and PDR (proliferative diabetic retinopathy) where NPDR(nonproliferative diabetic retinopathy) can be subdivided into mild nonproliferative diabetic retinopathy, moderate nonproliferative diabetic retinopathy, severe nonproliferative diabetic retinopathy. Proliferative however refers whether there is any neovascularization (abnormal blood vessel growth) present or not. The stages of diabetic retinopathy are described below:

- NPDR(nonproliferative diabetic retinopathy): it occurs when retinal capillaries are damaged due to hyperglycemia and as the capillary walls are weakens there is a small outpouching of the vessel lumens which is known as microaneurysms. These microaneurysms eventually cause the rupture to form hemorrhage, the vessels leak and cause the fluid to flow all over the retina. It can be further divided into the following categories:
    - Mild NPDR(nonproliferative diabetic retinopathy): with one or more microaneurysms present in retina. There are approximately 40% of the people with diabetes have signs of mild NPDR [8].
    - Moderate NPDR(nonproliferative diabetic retinopathy): multiple microaneurysms can be found in retina along with retinal hemorrhages, venous beading and spots

of cotton wool is also formed. 16% of the people who has moderate NPDRwill show tendency to develop PDR in about a time span of one year [9].

- Severe NPDR(nonproliferative diabetic retinopathy): in this case severe forms of intra retinal microvascular abnormalities are found and along with this cotton wool spots and venous beading are also present in this stage. The "4-2-1 rule" is usually used to diagnose in this stage. The patient can be diagnosed if he/she has the following complicacies: diffuse intra retinal hemorrhages and microaneurysms in 4 quadrants, venous beading in ≥ 2 quadrants, or IRMA in ≥1 quadrant. Approximately 50% of the patient with severe NPDR can eventually have PDR with a year [9].

- PDR(proliferative diabetic retinopathy): At this stage, there is circulation problem depriving the retina of oxygen and hence in PDRsmall abnormal blood vessel starts to grow along the retinal wall. Like a film of a camera, the retina sits at the back of the eye and because of these abnormal blood vessel growth, a gel like fluid is filled at the back of the eye which makes the vision blurry and in extreme cases complete blindness is also possible as the light rays cannot be received by the optical nerve [9].

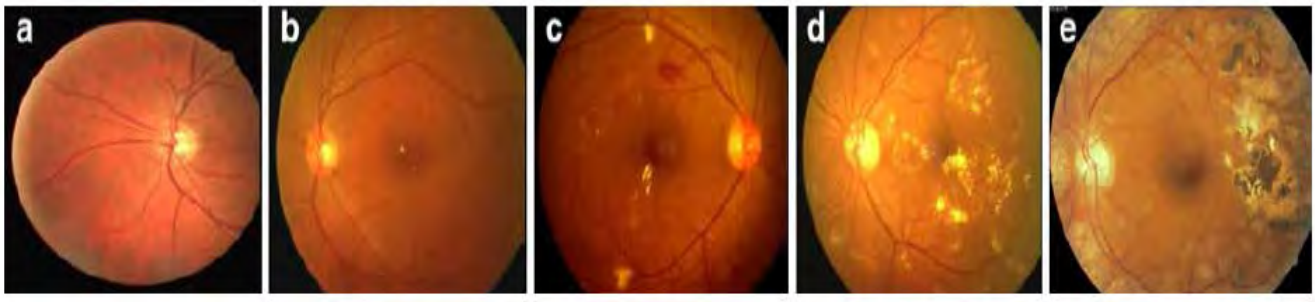

Figure 2: (a) Normal (b) Mild DR (c) Moderate DR (d) Severe DR (e) PDR [3]

2.2 CONVOLUTION NEURAL NETWORK

Convolutional Neural Network and Neural Networks are quite similar as they are made up of neurons which have weight and bias functions that are learnable. After a neuron receives an input it performs dot-product and optionally follows it with a non-linearity [10]. In general, it receives a raw image as an input and at the end it generates a class of scores. In convolutional neural network there is a loss function at the end layer (fully-connected layer).

The main advantage of using convolutional neural network is that it assumes that the input it receives is always an image which indeed helpsto pass certain parameters into the architecture. However, because of this assumption we are able to implement forward function more efficiently and also this will help to reduce the parameters in the network.

2.2.1 ARCHITECTURE OVERVIEW

This architecture has a 3D volume of neurons and the biggest advantage of this convolutional neural network is that the inputs are only considered as images. One of the main differences between Neural Network and Convolutional Neural Network is that Convolutional Neural Network has one more dimension than the Neural Network. Therefore, Convolutional Neural Networkhas 3 dimensions in total: **width**, **height** and **depth**. This depth means that it is the third dimension of the activation volume not the full depth of the Network. This whole Convolutional Neural Network is built on top of LAYERS [10].

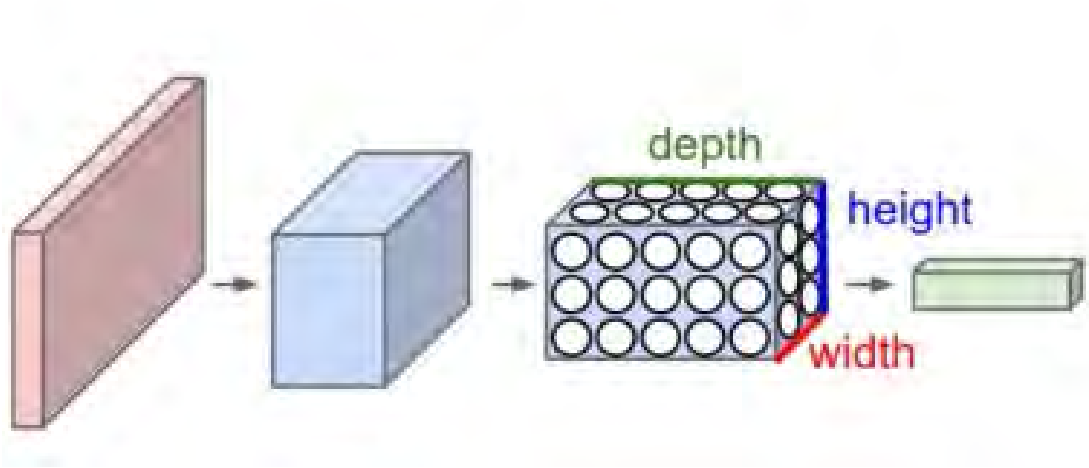

Figure 3: 3D output volume of neuron activation [10]

2.2.2 LAYERS USED TO BUILD CONVNET

There are three main layers on which this architecture can be built: **Convolutional Layer**, **Pooling Layer** and **Fully-Connected Layer**. When these layers are put on top of one another they form the desired full convolutional neural network architecture [10].

Architecture of convolutional neural network:

- INPUT these holds the raw pixel values of the input images.
- CONVOLUTIONAL layer will calculate the desired output after testing on a database by computing a dot product between the weight and regions that are connected to the input volumes.
- RELU layer is used to do an element wise activation function where the thresholding is done at level zero. No change in the overall volume takes place because of RELU layer.
- POOL layer is used to reduce the size/volume of both the width and height.

- FULLY CONNECTED layer will calculate the class scores at the end of the convolutional neural network.

2.3 VGGNet ARCHITECTURE

We have used convolutional neural network in our thesis and for that there are several architectures that are available like AlexNet, VGGNet, GoogLeNet and the latest one used by Microsoft is the ResNet. All these architectures however useconvolutional neural network as their basic mode of operation but there are changes in the architecture because of the filter sizes and because different architecture has used different sizes of the depth weight. There are also improvements in the accuracy of the results due to these slight modifications and over the years these architectures helped to improve in the field of image classifications. In our thesis, we have implemented VGGNet [11], this architecture was proposed from the network of Karen Simonyan and Andrew Zisserman and with their proposed network they became the runner up in ILSVRC 2014. Their main aim was to show that the depth of the network is very important for good performance. Their best network has 16 layers of convolutional layers as well as fully connected layers and they used a filter size of 3x3 throughout the architecture. It was also seen that even with a slightly weak classification performance, the VGG convolutional neural network features can outperform GoogLeNet in multiple transfer learning tasks. As a result of which VGGNet is one of the most preferred choice of architecture within the community where features extraction is concerned. However the downside of the VGGNet is that it is more expensive to evaluate and uses a lot of GPU memory and parameters.

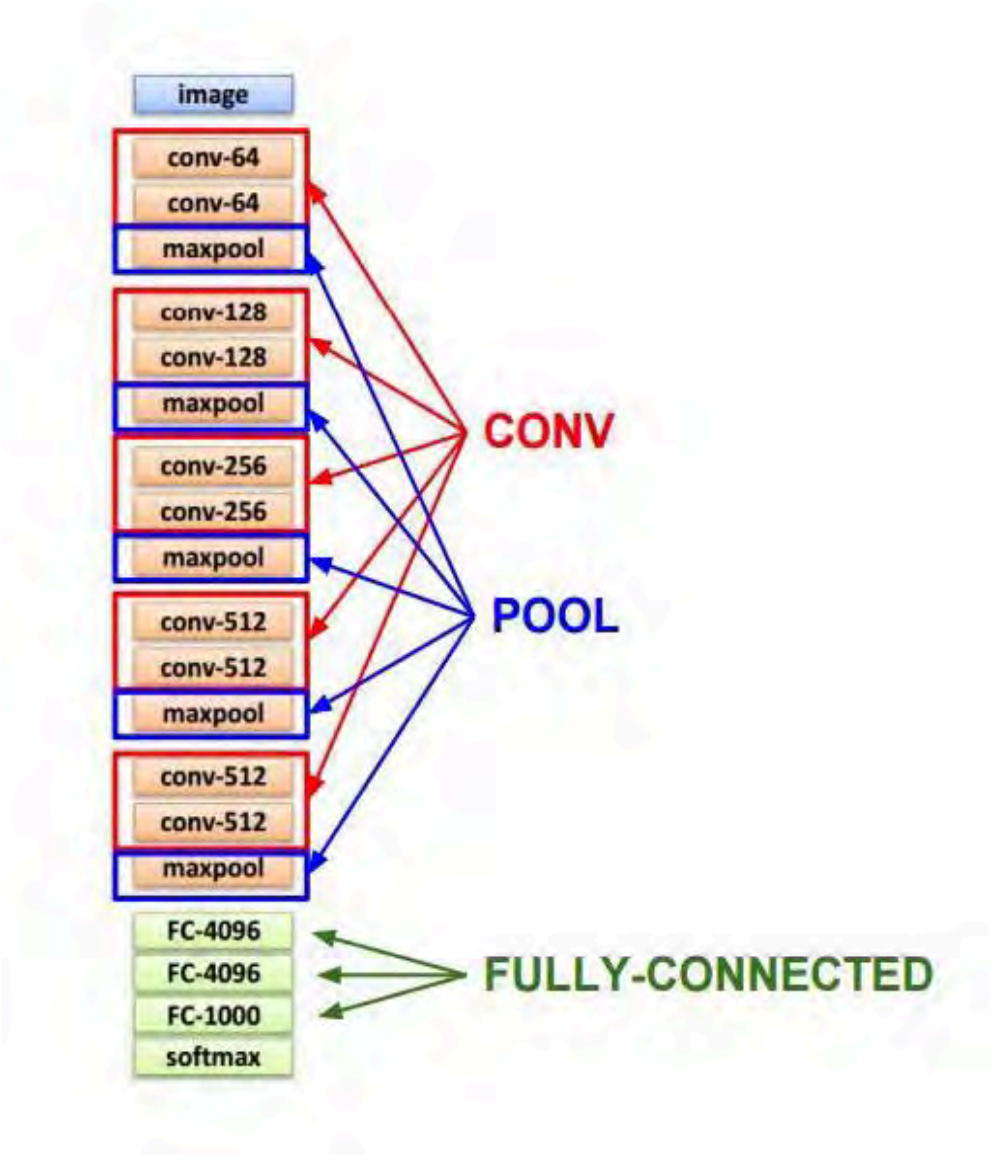

Figure 4: VGGNet Architecture [12]

For our understanding of the overall convolutional neural network we went through the lecture of Stanford University in YouTube. This gives us the basic knowledge of the whole process and gave us the idea for our thesis work.

2.4 RELATED WORK

While doing our research we have found out what other people has done in this topic. In [13] they have detected Microaneurysms using Naive Bayes to classify the disease stages and their main task was to extract the features of areas like blood vessel. In [14] their main task was to detect the retinal changes of diabetes patient's eye such as microaneurysms, hard exudates, soft exudates, hemorrhageetc., they wanted to monitor the changes in retinal images and from that they want to conclude whether the patient have diabetic retinopathy or not. In [15] they have used a new algorithm to detect the blood vessels efficiently, which is a key step to detect Diabetic Retinopathy. First they enhance the image and then curvelet transformation is applied to equalize the image, these pre-processed image helps in better extraction of the blood vessels. In [16] exudates in color fundus were detected as well as classify the severity of the lesions using SVM classifier. In [17] they also did the same type of work as that of [16] but in their work they have used ANN classifier.In [18] they first, localize and segment optic disc, they also did segmentation of retinal vasculature and then they localize macula and fovea and at last they were able to localize and segment diabetic retinopathy.

3. PROPOSED WORK

Our workflow of the automated system to detect diabetic retinopathy is given below:

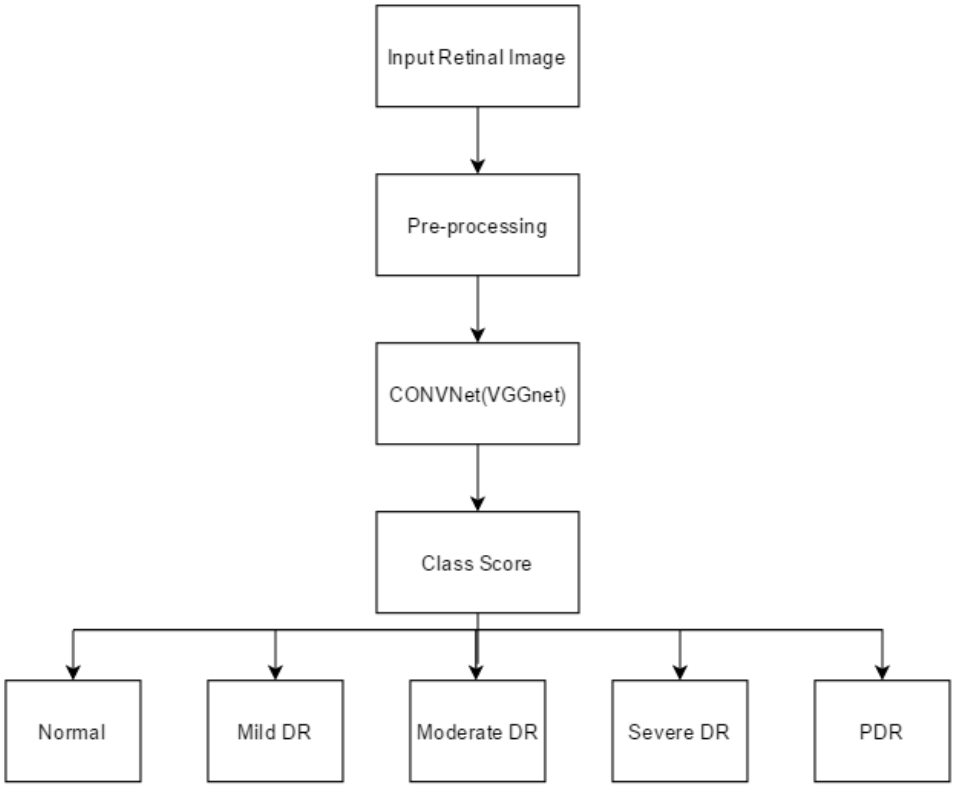

Figure 5: Block Diagram of DR detection and Severity Classification

3.1 RETINAL IMAGES

Our dataset contains 35,126 images for training the network and 53,576 images for testing the network.The five fairly unbalanced classes of images for diabetic retinopathy, is divided into 5 subclasses. The following table and graph will show the number of images and their percentages. Our dataset contains images which are taken at various lighting condition. There are five training labels [0,1,2,3,4] where the labels are named as normal, mild DR, moderate DR, severe DR and PDR respectively and for each patient there are two images of both right and left eye naming "right" and "left" and for each patient, which has a unique patient id. These dataset that we are using is taken from a competition which took place in kaggle.com.

Training dataset of 5 unbalance classes are given below:

Class	Name	Number of Image
0	Normal	25,810
1	Mild DR	2,443
2	Moderate DR	5,292
3	Severe DR	873
4	PDR	708

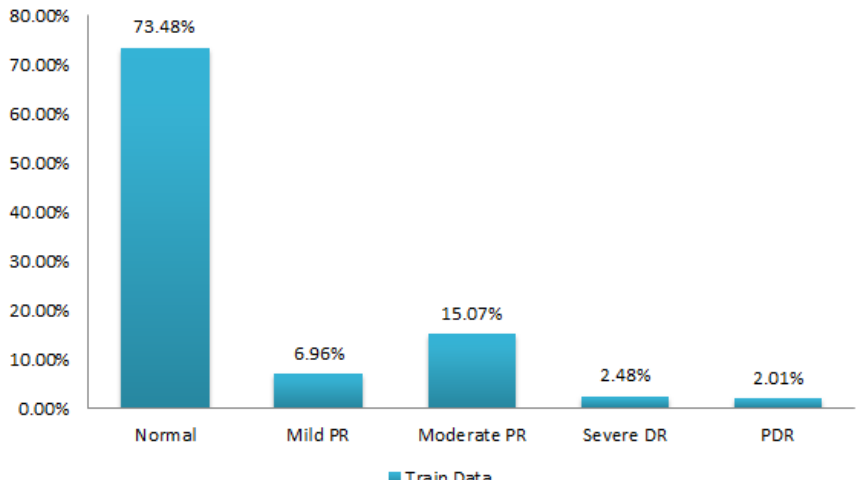

Figure 6: Training Dataset Labels in Percentage

Test dataset of 5 unbalance classes are given below:

Class	Name	Number of Image
0	Normal	39,533
1	Mild DR	3,762
2	Moderate DR	7,861
3	Severe DR	1,214
4	PDR	1,206

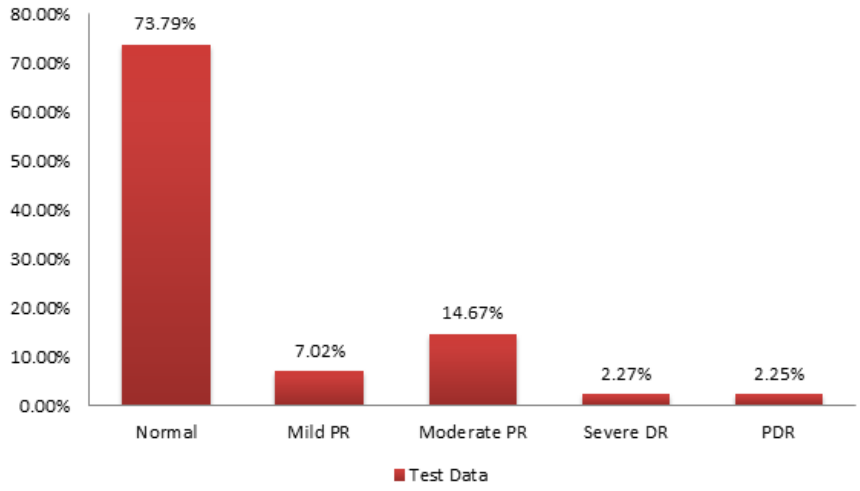

Figure 7: Test Dataset Labels in Percentage

3.2 PRE-PROCESSING

For preprocessing of the data we used the toolkit graphicsmagick. We changed the format of the image to png, and then we resized the images as 128x128pixels and the pseudo code is given below:

Pseudo Code

Preprocess Image(X, A)

X= image

A= desired size

Input: Raw jpeg image

Output: preprocessed png image

1. For each Image X
2. **Do**
3. Convert X to png format
4. Add black border of 1x1 pixel
5. Fuzz X to 10%
6. Trim X
7. Repage X
8. Gravity Centre X
9. Resize A
10. Black background X
11. **End loop**
12. **Return** X

3.3 NETWORK ARCHITECTURE OF VGGNet

The input RGB images for our training data was of size 128x128. We centered cropped our input images to get the desired pixel sizes for our experiments. The images are processed through a pile of convolutional layers where we have used a small filter size of 3x3. We have used a fixed convolutional stride of 1 pixel with a padding of 1 pixel for the 3x3 convolutionallayer. For max-pooling we have used a pixel window of 2x2 with stride 2 and 3x3 with stride 2.

We used a stack of convolutional layers (where with different depths we can achieve different architectures) which is carry forward by three fully connected layers. For the first two fully connected layers we have used 2048 channels and the third fully connected layer will have 5 channels which will perform classification for our five classes. The final layer is the sigmoid or softmax layer and however the fully connected layer remains similar in the entire network that we have used. Our hidden layers are equipped with the rectification layer, ReLU nonlinearity. Our next section will cover four of our best model of the VGGNet that we have used to train our dataset.

3.3.1 MODEL CONFIGURATION

Model A:

Name	Batch	Channel	Height	Width	FilterSize/Stride
Input	128	3	128	128	
Conv	128	32	128	128	3/1
Max pool	128	32	64	64	3/2
Conv	128	64	64	64	3/1
Max pool	128	64	32	32	3/2
Conv	128	128	32	32	3/1
Conv	128	128	32	32	3/1
Max pool	128	128	16	16	3/2
Conv	128	256	16	16	3/1
Max pool	128	256	16	16	3/2
FC	128	2048			
FC	128	2048			
Sigmoid	128	5			

Below is the representation of total memory and parameter for Model A:

Name	Dimension	Memory	Weight
Input	[128x128x3]	128*128*3=49K	0
Conv3-32	[128x128x32]	128*128*32=524K	(3*3*3)*32=864
Pool2	[64x64x32]	64*64*32=131K	0
Conv3-64	[64x64x64]	64*64*64=262K	(3*3*32)*64=18432
Pool2	[32x32x64]	32*32*64=65K	0
Conv3-128	[32x32x128]	32*32*128=131K	(3*3*64)*128=73728
Conv3-128	[32x32x128]	32*32*128=131K	(3*3*128)*128=147456
Pool2	[16x16x128]	16*16*128=32K	0
Conv3-256	[16x16x256]	16*16*256=65K	(3*3*128)*256=294912
Pool2	[8x8x256]	8*8*256=16K	0
FC-2048	[1x1x2048]	1*1*2048=2048	(8*8*256)*2048=33554432
FC-2048	[1x1x2048]	1*1*2048=2048	2048*2048=4194304
Sigmoid	[1x1x5]	1*1*5=5	2048*5=10249

TOTAL MEMORY	5.64MB/image (only for forward propagation) * 2 for backward propagation
TOTAL PARAMETER	38 million parameters

Model B:

Name	Batch	Channel	Height	Width	FilterSize/Stride
Input	128	3	128	128	
Conv	128	32	128	128	3/1
Max pool	128	32	64	64	3/2
Conv	128	64	64	64	3/1
Max pool	128	64	32	32	3/2
Conv	128	64	32	32	3/1
Max pool	128	64	16	16	3/2
Conv	128	128	16	16	3/1
Max pool	128	128	8	8	3/2
Conv	128	256	8	8	3/1
Max pool	128	256	4	4	2/2
FC	128	2048			
FC	128	2048			
Sigmoid	128	5			

Below is the representation of total memory and parameter for Model B:

Name	Dimension	Memory	Weight
Input	[128x128x3]	128*128*3=49K	0
Conv3-32	[128x128x32]	128*128*32=524K	(3*3*3)*32=864
Pool2	[64x64x32]	64*64*32=131K	0
Conv3-64	[64x64x64]	64*64*64=262K	(3*3*32)*64=18432
Pool2	[32x32x64]	32*32*64=65K	0
Conv3-64	[32x32x64]	32*32*64=65K	(3*3*64)*64=36864
Pool2	[16x16x64]	16*16*64=16K	0
Conv3-128	[16x16x128]	16*16*128=32K	(3*3*64)*128=73728
Pool2	[8x8x128]	8*8*128=8K	0
Conv3-256	[8x8x256]	8*8*256=16K	(3*3*128)*256=294912
Pool2	[4x4x256]	4*4*256=4K	0
FC-2048	[1x1x2048]	1*1*2048=2048	(4*4*256)*2048=8388608
FC-2048	[1x1x2048]	1*1*2048=2048	2048*2048=4194304
Sigmoid	[1x1x5]	1*1*5=5	2048*5=10249

TOTAL MEMORY	4.71MB/image (only for forward propagation) * 2 for backward propagation
TOTAL PARAMETER	13 million parameters

Model C:

Name	Batch	Channel	Height	Width	FilterSize/Stride
Input	128	3	128	128	
Conv	128	32	128	128	3/1
Max pool	128	32	64	64	3/2
Conv	128	64	64	64	3/1
Max pool	128	64	32	32	3/2
Conv	128	128	32	32	3/1
Conv	128	128	32	32	3/1
Max pool	128	128	16	16	3/2
Conv	128	256	16	16	3/1
Max pool	128	256	8	8	3/2
FC	128	2048			
FC	128	2048			
Sigmoid	128	5			

Below is the representation of total memory and parameter for Model C:

Name	Dimension	Memory	Weight
Input	[128x128x3]	128*128*3=49K	0
Conv3-32	[128x128x32]	128*128*32=524K	(3*3*3)*32=864
Pool2	[64x64x32]	64*64*32=131K	0
Conv3-64	[64x64x64]	64*64*64=262K	(3*3*32)*64=18432
Pool2	[32x32x64]	32*32*64=65K	0
Conv3-128	[32x32x128]	32*32*128=131K	(3*3*64)*128=36864
Conv3-128	[32x32x128]	32*32*128=131K	(3*3*128)*128=73728
Pool2	[16x16x128]	16*16*128=32K	0
Conv3-128	[16x16x128]	16*16*128=32K	(3*3*128)*128=294912
Pool2	[8x8x128]	8*8*128=8K	0
Conv3-256	[8x8x256]	8*8*256=16K	(3*3*128)*256=294912
Pool2	[4x4x256]	4*4*256=4K	0
FC-2048	[1x1x2048]	1*1*2048=2048	(2*2*256)*2048=8388608
FC-2048	[1x1x2048]	1*1*2048=2048	2048*2048=4194304
Sigmoid	[1x1x5]	1*1*5=5	2048*5=10249

TOTAL MEMORY	5.5MB/image (only for forward propagation) * 2 for backward propagation
TOTAL PARAMETER	13.3 million parameters

Model D:

Name	Batch	Channel	Height	Width	FilterSize/Stride
Input	128	3	128	128	
Conv	128	32	128	128	3/1
Max pool	128	32	64	64	3/2
Conv	128	64	64	64	3/1
Max pool	128	64	32	32	3/2
Conv	128	128	32	32	3/1
Conv	128	128	32	32	3/1
Max pool	128	128	16	16	3/2
Conv	128	256	16	16	3/1
Max pool	128	256	16	16	3/2
FC	128	2048			
FC	128	2048			
Softmax	128	5			

Below is the representation of total memory and parameter for Model D:

Name	Dimension	Memory	Weight
Input	[128x128x3]	128*128*3=49K	0
Conv3-32	[128x128x32]	128*128*32=524K	(3*3*3)*32=864
Pool2	[64x64x32]	64*64*32=131K	0
Conv3-64	[64x64x64]	64*64*64=262K	(3*3*32)*64=18432
Pool2	[32x32x64]	32*32*64=65K	0
Conv3-128	[32x32x128]	32*32*128=131K	(3*3*64)*128=73728
Conv3-128	[32x32x128]	32*32*128=131K	(3*3*128)*128=147456
Pool2	[16x16x128]	16*16*128=32K	0
Conv3-256	[16x16x256]	16*16*256=65K	(3*3*128)*256=294912
Pool2	[8x8x256]	8*8*256=16K	0
FC-2048	[1x1x2048]	1*1*2048=2048	(8*8*256)*2048=33554432
FC-2048	[1x1x2048]	1*1*2048=2048	2048*2048=4194304
Softmax	[1x1x5]	1*1*5=5	2048*5=10249

TOTAL MEMORY	5.64MB/image (only for forward propagation) * 2 for backward propagation
TOTAL PARAMETER	38 million parameters

3.3.2 TRANING

We took the preprocessed data from the input images, we first allocate random weight to the first layer of convolutional neural network. Then with the help of forward and backward propagation it tries to adjust the weight so that the weight can be optimized along the way. With these samples of data, forward pass is done so that it can get an accurate level of prediction about the image that it is trying to train. However back propagation helps to reduce the error which takes place during forward propagation.Along the way, small updates are made to the parameters. The training is carried out by multinomial logistic regression using mini-batch gradient descent with momentum. We set out batch size to 128 and we considered our momentum as 0.9. The training was adjusted by the weight decay and the dropout regularization for the first two fully connected layers and our drop out ratio was set to 0.5. Our initial learning rate was set to 0.01 and it was decreased by a factor 10 when the accuracy level stops improving.

Throughout our convolutional neural network we have used filter size of 3x3 with stride 1. The main reason behind choosing the filter size of 3x3 is that having three 3x3 filter produces lesser parameters that using a single 7x7 filter produces. Considering three filters with a channel of C produces $3*(3^2*C^2) = 27C^2$ whereas, if for a single 7x7 filter with a channel of C produces $(7*7*C^2) = 49C^2$. Therefore, for our own convenience we have chosen a smaller filter size to keep the number of parameters less.

3.3.3 TESTING

The images are tested on the already trained convolutional neural network. The test input images are resized to the same size as that of the model in which it is trained with, so that the testing can be done accurately. The fully connected layers are then converted into the original convolutional layer filter size, then a class score is assigned to all the test images. With the help of these class scores we will then be able to distinguish between the 5 different classes of diabetic retinopathy namely, Normal (0), Mild DR (1),Moderate DR (2), Severe DR (3), PDR (4).

3.4IMPLEMENTATION HARDWARE AND TOOLKITS

We used a graphics card from NVidia, and the model is GTX 780 which has 3GB of memory andour PC has an onboard RAM of 16GB. The operating system that we first tried to used is Ubuntu 14.04 LTS but many of our repositories were locked therefore we had to eventually downgrade to Ubuntu 12.04 LTS. It approximately it took around 7-8 hours to train our dataset using this configuration. We used Cuda version 7.5 and along with it we used the cuDNN which helped us with a lot of libraries. We have also used libraries of Theano, Lasagne and Pylearn.

4. RESULT AND ANALYSIS

Our approach was to use different combinations of convolutional layers and maxpool layers for our work. We have used combinations of 9, 10 and 11convolutional layersand using those we came up with four of our best models namely model A, B, C and D. In model A and D, there was 9convolutional layersfollowed by model B which has 10 layers and C with 11 convolutional layers. After we are done with our experiment we can conclude that 9 layers of network worked better in our system. In figure 8, we have calculated our level of accuracy in contrast to the original pathological labels that was provided with our dataset. We tested our accuracy level with the four best models that we had. Model A and D follow the same architecture but both have different level of accuracy because their activation function is different. Model A uses sigmoid function for its activation whereas model D uses softmax. Several parameters such as True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) are calculated. These parameters are calculated by comparing the classifier outcome with the number of normal and abnormal images from the database. For an abnormal image, the result is True Positive (TP) if the outcome of classification is abnormal and the result is False Negative (FN) if the classifier output is normal. For normal image, the result is True Negative (TN), if the classifier output is normal and False Positive (FP) if the classification outcome is abnormal. In a given image dataset, these parameters, TP,TN, FP, FN are used in the calculation of the Accuracy, Sensitivity (SN) and specificity (SP). Performance of the classifier can be measured in terms of sensitivity, specificity and accuracy.

$$\text{Sensitivity} = TP / (TP+FN)$$

Sensitivity is measure of the percentage of abnormal images classified as abnormal.

$$\text{Specificity} = TN / (TN+FP)$$

Specificity gives the measure of normal images that are classified correctly as normal.

$$\text{Accuracy} = (TP+TN)/ (TP+FN+TN+FP)$$

It is the measure of total number of well classified normal and abnormal images.

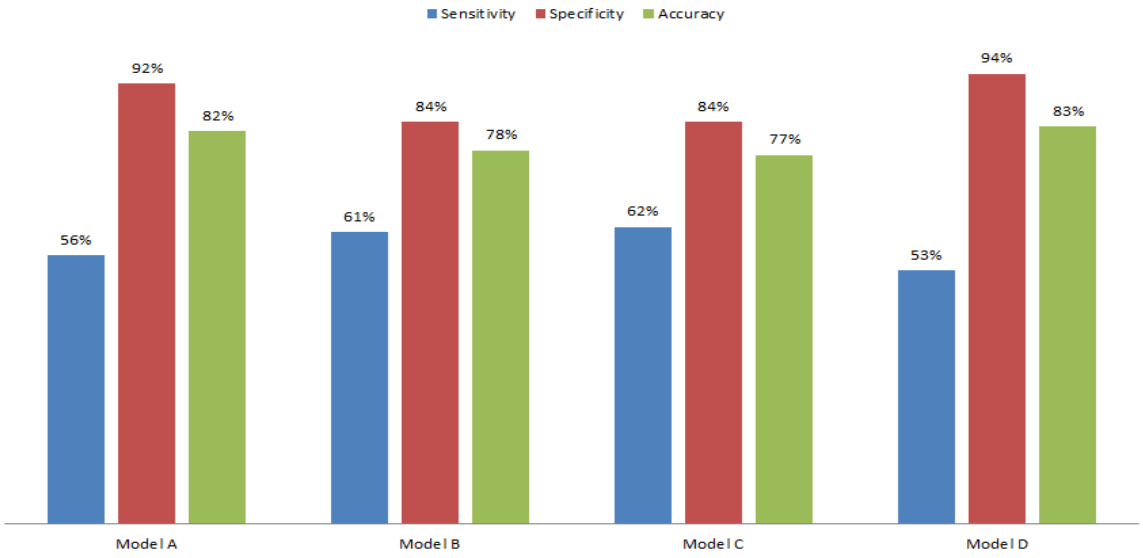

Figure 8: Classifier Performance Graph

5. LIMITATIONS

We were confined with 128x128 pixel size images and were not able to work with a larger pixel size. With our graphics card of NVidia GTX780 we could not train our dataset with a larger pixel which did actually affect the result of our thesis. With increase in the memory size of the graphics card the accuracy will increase because in that case our system will be able to work in a larger image size to detect the problem more accurately.

We had to work mostly at labs which also had a fixed opening and closing time therefore we were able to train and test less number of models.

6. FUTURE PROSPECTS

We want to develop an online based service where doctors can get yearly subscription and through which they can easily check for the diabetic retinopathy very easily without having to worry about program or its internal code. They will simply update the images into our online repository and our system will provide them with the details of every single image along with the level of severity. This will be a very convenient in which every single doctor can get access to our system through internet as in today's world internet is accessible by almost everyone.

We want to implement the latest and greatest algorithm or architecture available for computer vision for making our system more intelligent to find out the stages of the diabetic retinopathy.

7. CONCLUSION

In our paper, we proposed a system which will be able to detect diabetic retinopathy from the image of an eye of a patient. With our proposed system, the doctors can however spent less time on the overall detection process and can take more care of his/her patient. However, we were not being able to bring out the best accurate results because of our hardware limitations but accuracy level can be increased if someone uses a better GPU because in that case they will be able to have larger input image. Moreover, in this paper it is shown that the diseases are classified according to their severity level using convolutional neural network. This convolutional neural network can work easily with images as we do not have to manually mention the features within the image, the network does these automatically by itself while the images are ready to get trained. This overall work is very important because early detection is very important for a patient with diabetic retinopathy because without that the patient can go blind in extreme cases. So hopefully, if we can integrate this system with medical science then many doctors will be able to save the vision of their patients.

8. REFERENCES

1. R. Frank, "Diabetic Retinopathy", *New England Journal of Medicine*, vol. 350, no. 1, pp. 48-58, 2004.
2. Y. Ng, R. Acharya U, R. Rangayyan and J. Suri, *Ophthalmological imaging and applications*. .
3. O. Faust, R. Acharya U., E. Ng, K. Ng and J. Suri, "Algorithms for the Automated Detection of Diabetic Retinopathy Using Digital Fundus Images: A Review", *J Med Syst*, vol. 36, no. 1, pp. 145-157, 2010.
4. World Diabetes, A newsletter from the World Health Organization, 4, 1998.
5. G. Ong, L. Ripley, R. Newsom, M. Cooper and A. Casswell, "Screening for sight-threatening diabetic retinopathy: comparison of fundus photography with automated color contrast threshold test", *American Journal of Ophthalmology*, vol. 137, no. 3, pp. 445-452, 2004.
6. "Orbis", *Orbis.org*, 2016. [Online]. Available: http://www.orbis.org. [Accessed: 01- Apr- 2016].
7. A. KHAN, "AnisurRahman Khan: 3.2 million people in Bangladesh suffer from diabetes", *Arrkhan.blogspot.com*, 2013. [Online]. Available: http://arrkhan.blogspot.com/2013/10/32-million-people-in-bangladesh-suffer.html. [Accessed: 01- Apr- 2016].
8. J. Nayak, P. Bhat, R. Acharya U, C. Lim and M. Kagathi, "Automated Identification of Diabetic Retinopathy Stages Using Digital Fundus Images", *J Med Syst*, vol. 32, no. 2, pp. 107-115, 2007.
9. N. Cheung, P. Mitchell and T. Wong, "Diabetic retinopathy", *The Lancet*, vol. 376, no. 9735, pp. 124-136, 2010.

10. "CS231n Convolutional Neural Networks for Visual Recognition", *Cs231n.github.io*, 2016. [Online]. Available: http://cs231n.github.io/convolutional-networks/. [Accessed: 01- Apr- 2016].
11. "Visual Geometry Group Home Page", *Robots.ox.ac.uk*, 2016. [Online]. Available: http://www.robots.ox.ac.uk/~vgg/research/very_deep/. [Accessed: 01- Apr- 2016].
12. J. Clune, *Artificial Intelligence*, 1st ed. .
13. R. Maher, D. Panchal and J. Kayte, "Automatic Diagnosis Microaneurysm Using Fundus Images", *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 5, no. 10, 2015.
14. N. Thomas and T. Mahesh, "Detecting Clinical Features of Diabetic Retinopathy using Image Processing", *International Journal of Engineering Research & Technology (IJERT)*, vol. 3, no. 8, 2014.
15. B. SINGH and K. JAYASREE, "Implementation of Diabetic Retinopathy Detection System for Enhance Digital Fundus Images", *International Journal of advanced technology and innovation research*, vol. 7, no. 6, pp. 0874-0876, 2015.
16. M. Gandhi and D. Dhanasekaran, "Diagnosis of Diabetic Retinopathy Using Morphological Process and SVM Classifier", *International conference on Communication and Signal Processing*, 2013.
17. E. Shahin, T. Taha, W. Al-Nuaimy, S. Rabaie, O. Zahranand ..Abd El-Samie, "Automated Detection of Diabetic Retinopathy in Blurred Digital Fundus Images", 2012.
18. P. Prentasic. “Detection of Diabetic Retinopathy in Fundus Photographs”. University of Zagreb, Faculty of Electrical Engineering and Computing Unska 3, 10000 Zagreb, Croatia. 2013. pp 6-8.