# A Convolutional Neural Network for Classification of Diabetic Retinopathy

***Md. Shahadat Mollah***
***ID: 2014-3-60-075***
***Jamal Mostafa Eahsan***
***ID: 2014-3-60-056***
***Nafia Nowrin***
***ID: 2014-3-60-052***

A thesis submitted in partial fulfilment of the requirements for the degree of Bachelor of Science in Computer Science and Engineering


**Department of Computer Science and Engineering**
**East West University**
**Dhaka-1212, Bangladesh**

**September, 2019**

# Declaration

We, hereby, declare that the work presented in this thesis titled “Diabetic Retinopathy Classification using Convolutional Neural Network” is the outcome of the investigation performed by us under the supervision of Dr. Taskeed Jabid, Associate Professor, Department of Computer Science and Engineering, East West University. We also declare that no part of this thesis has been or is being submitted elsewhere for the award of any degree or diploma.

.....
Dr Taskeed Jabid
Supervisor

.....
Md.Shahadat Mollah
ID: 2014-3-60-075

.....
Jamal Mostafa Eahsan
ID: 2014-3-60-056

.....
Nafia Nowrin
ID: 2014-3-60-052

# Letter of Acceptance

We, hereby declare that this thesis is the students own work and best effort of ours. All other sources of information used have been acknowledged. This thesis has been submitted with our approval.

. . . . . .

Dr. Taskeed Jabid

Associate Professor

Department of Computer Science and Engineering

East West University

Aftabnagar, Dhaka-1212, Bangladesh

**Supervisor & Chairperson**

# Abstract

As the years pass by, more and more diseases are being diagnosed and there has been a gradual increase in the number of patients suffering from diabetic. With the advancing technology, medical sciences have been fast-forwarding towards machines that help in the diagnosis of the disease in every field. In our thesis, we worked with diabetic retinopathy images to classify between retinopathy eye and normal eye using a convolutional neural network. The convolutional neural network is very efficient for image classification problems as it extracts the features from the images that help us differentiate between the normal eye and retinopathy eye. The implementation of such technology has enabled the doctors to provide faster and better treatments to patients suffering from diabetic. Many diseases can be diagnosed through such systems in hospitals. Some of the hospitals have already started using these systems and the results are quite impressive. This will benefit both the people and the doctors as the treatments will be provided within a short period and the patients don't have to suffer through diabetic retinopathy. We developed a convolutional neural network model to classify diabetic retinopathy images in [1] dataset. The dataset is an open-source diabetic retinopathy dataset that has about 35126 images of various patients appropriate for our research.

# Acknowledgments

We would like to express our deepest gratitude to the almighty Allah for His blessings on us. Our special thanks go to our supervisor, Dr. Taskeed Jabid, who gave us this opportunity, initiated us into the field of “Computer Vision and Deep Learning", and without whom this work would not have been possible. His encouragements, visionaries and thoughtful comments and suggestions, unforgettable support at every stage of our Bachelor of Science study were simply appreciating and essential. His ability to make us capable enough to finally answer our own questions correctly is something valuable and what we have learned and we would try to emulate, if ever we get the opportunity.

There are numerous other people too who have shown me their constant support and friendship in various ways, directly or indirectly related to our academic life. We will remember them in our hearts and hope to find a more appropriate place to acknowledge them in the future.

**Md. Shahadat Mollah**
**September,2019**

**Jamal Mostafa Eahsan**
**September,2019**

**Nafia Nowrin**
**September,2019**

# Table of Contents

# List of Figures

# List of Tables

# Chapter 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 the 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 vessels may lead to anomalies like microaneurysms, hemorrhages, hard exudates, and cotton wool spots develop during the different phases of diabetic retinopathy [2,3].

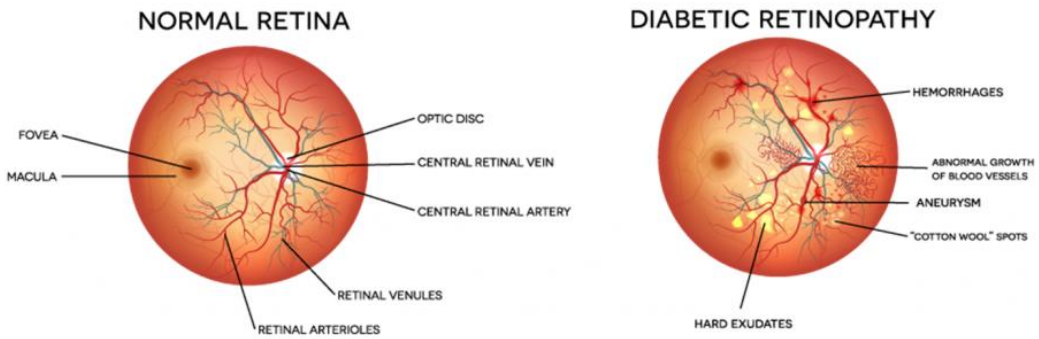



Figure 1.1: Features in normal retina and DR image

Figure 1.2: Difference between normal vision and DR vision

As of 2017, an estimated 425 million people had diabetes worldwide [4]. This represents 8.8% of the adult population [4]. Diabetic retinopathy affects up to 80 percent of those who have had diabetes for 20 years or more [5]. The longer a person has diabetes, the higher his or her chances of developing diabetic retinopathy [6]. In developed countries, one of the alarming causes of blindness is diabetic retinopathy [7] Each year in the United States, diabetic retinopathy accounts for 12% of all new cases of blindness. It is also the leading cause of

blindness in people aged 20 to 64[8]. For 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 [9]. At least 90% of new cases could be reduced with proper treatment and monitoring of the eyes [10]. But the symptoms for diabetic retinopathy do not show up clearly in the early stages, which makes it even harder for the ophthalmologist to prevent the patient from being blind. The number of patients 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.

In our thesis project, we have developed a model using CNN that can actively detect any kind of irregularities in the diabetic retinopathy images of the patients. With this system, we intend to make the detection of diabetic retinopathy stages and make the process easier for further treatment. This detection model can help the ophthalmologists suggest their patients' treatment cycle and save enough time and also improve their eye conditions.

Detection of diabetic retinopathy is quite sensitive, hence more challenging. The results obtained must be accurate as it can be a life-changing decision for the patients. Hence, along with this system, doctors' presence is crucial. This system cannot replace doctors but it will simply speed up the treatment process and save patients' lives.

# Chapter 2

# Background Study

In our research we used convolutional neural network to classify diabetic retinopathy images. We also used transfer learning method in our work. This chapter describes diabetic retinopathy, the convolutional neural networks and transfer learning method in details.

## 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 the retina, which may cause blindness and this process is what we know as diabetic retinopathy. The two types of diabetic retinopathy are NPDR (non-proliferative diabetic retinopathy) and PDR (proliferative diabetic retinopathy) where NPDR (non-proliferative diabetic retinopathy) can be subdivided into mild non-proliferative diabetic retinopathy, moderate non-proliferative diabetic retinopathy, severe non-proliferative diabetic retinopathy. Proliferative however refers to whether there is any neovascularization (abnormal blood vessel growth) present or not. The stages of diabetic retinopathy are described below:

- **NPDR (non-proliferative 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 (non-proliferative diabetic retinopathy): with one or more microaneurysms present in retina. There are approximately 40% of the people with diabetes have signs of mild NPDR [11].

- Moderate NPDR (non-proliferative diabetic retinopathy): multiple microaneurysms can be found in retina along with retinal hemorrhages, venous beading and spots 10 of cotton wool is also formed. 16% of the people who has moderate NPDR will show a tendency to develop PDR in about a time span of one year [12].

- Severe NPDR (non-proliferative diabetic retinopathy): in this case severe forms of intraretinal 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 [12].

- **PDR (proliferative diabetic retinopathy):** At this stage, there is circulation problem depriving the retina of oxygen and hence in PDR small 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 growths, 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 [12].

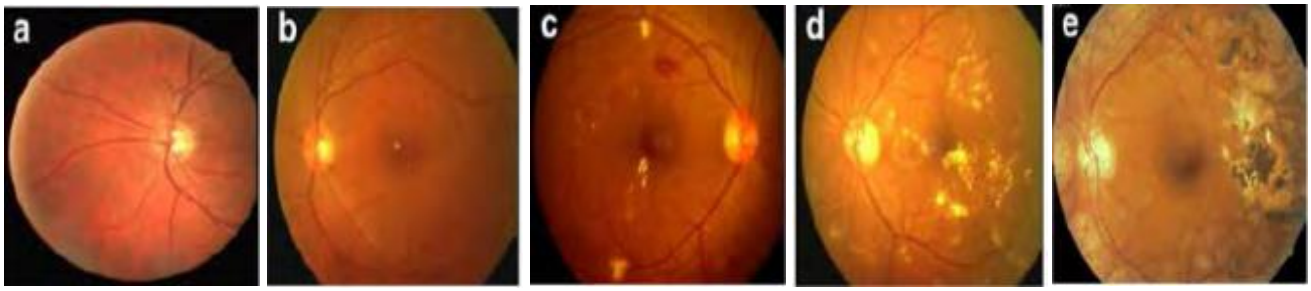



Figure 2.1: (a) Normal (b) Mild DR (c) Moderate DR (d) Severe DR (e) PDR [3]

## 2.2 Convolutional Neural Network

In machine learning, a convolutional neural network (CNN, or ConvNet) is a class of deep, feed-forward artificial neural networks, most commonly applied to analyzing visual imagery. CNNs use a variation of multilayer perceptron designed to require minimal preprocessing [13]. They are also known as shift invariant or space invariant artificial neural networks (SIANN), based on their shared-weights architecture and translation invariance characteristics [14, 15].

Convolutional networks were inspired by biological processes [16]. in that the connectivity pattern between neurons resembles the organization of the animal visual cortex. Individual cortical neurons respond to stimuli only in a restricted region of the visual field known as the receptive field. The receptive fields of different neurons partially overlap such that they cover the entire visual field.

CNNs use relatively little pre-processing compared to other image classification algorithms. This means that the network learns the filters that in traditional algorithms were hand-engineered. This independence from prior knowledge and human effort in feature design is a major advantage. They have applications in image and video recognition, recommender systems and natural language processing.

Convolutional Neural Network architectures make the explicit assumption that the inputs are images, which allows us to encode certain properties into the architecture. These then make the forward function more efficient to implement and vastly reduce the amount of parameters in the network. The description of layers in convolutional neural networks described below.

- Convolution Layer: The Conv layer is the core building block of a Convolutional Network that does most of the computational heavy lifting. Three hyperparameters control the size of the output volume: the depth, stride and padding. The Output Size of a convolutional layer is:

$$\left[\frac{n+2p-f}{s}+1 * \frac{n+2p-f}{s}+1\right]$$

where p = padding, s = stride, f = number of filters, n = image width = image height.

- Max Pooling Layer: Max pooling is a sample-based discretization process. Max pooling is done by applying a max filter to (usually) non-overlapping sub regions of the initial representation. The Output Size of a Max-Pooling layer is:

$$\left[\frac{n-f}{s}+1 * \frac{n-f}{s}+1\right]$$

where, s = stride, f = number of filters, n = image width = image height.

- Average Pooling Layer: Average pooling layer reduces the variance and complexity in the data. It also performs down-sampling by dividing the input into rectangular pooling regions and computing the average values of each region. The Output Size of a Average-Pooling layer is:

$$\left[\frac{n-f}{s}+1 * \frac{m-f}{s}+1\right]$$

where, s = stride, f = number of filters, n = image width = image height.

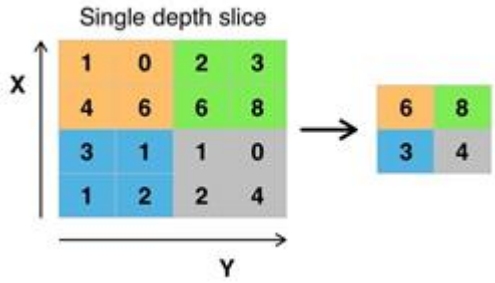



Figure 2.2: Max pooling with a 2x2 filter and stride = 2.

- Concat Layer: The Concat layer concatenates its multiple input blobs to one single output blob.
- Dropout Layer: A dropout layer randomly sets input elements to zero with a given probability. Dropout is a technique used to improve over-fit on neural networks.
- Fully Connected Layer: The fully connected (FC) layer in the CNN represents the feature vector for the input. This feature vector holds information that is vital to the input.
- Softmax Layer: Softmax assigns decimal probabilities to each class in a multiclass problem. Those decimal probabilities must add up to 1.0. This additional constraint helps training converge more quickly than it otherwise would.

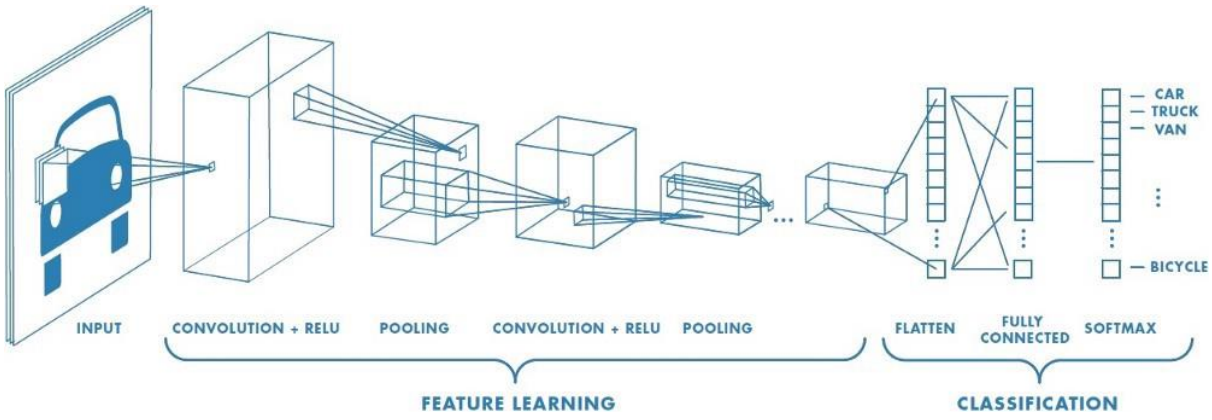



Figure 2.3: Typical CNN architecture

## 2.3 Transfer Learning

Transfer learning is a research problem in machine learning that focuses on storing knowledge gained while solving one problem and applying it to a different but related problem. In transfer learning a model developed for a task is reused as the starting point for a model on a second task.

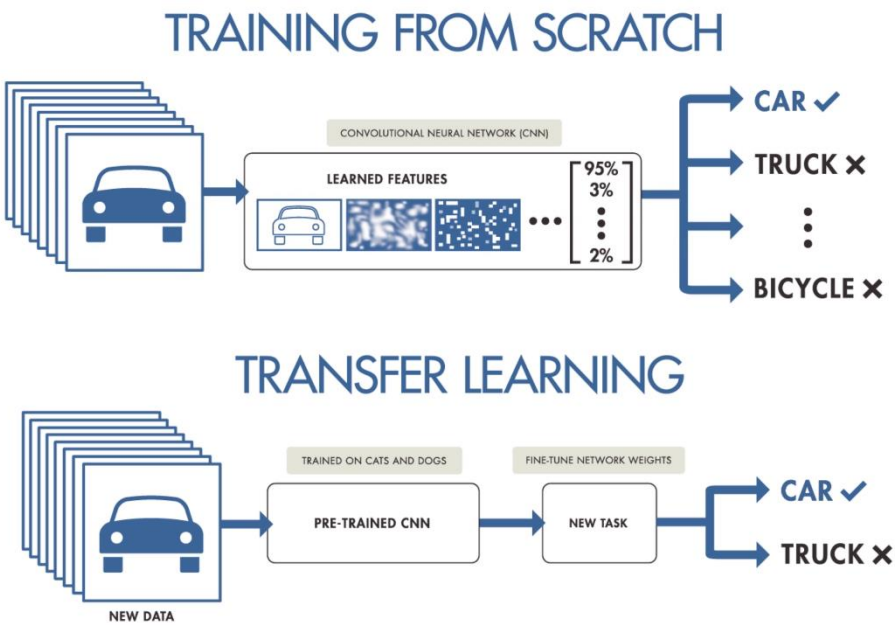



Figure 2.4: Typical transfer learning architecture with CNN

It is a popular approach in deep learning where pre-trained models are used as the starting point on computer vision and natural language processing tasks given the vast compute and time resources required to develop neural network models on these problems and from the huge jumps in skill that they provide on related problems.

Transfer learning is related to problems such as multi-task learning and concept drift and is not exclusively an area of study for deep learning.

Nevertheless, transfer learning is popular in deep learning given the enormous resources required to train deep learning models or the large and challenging datasets on which deep learning models are trained.

Transfer learning only works in deep learning if the model features learned from the first task are general.

This form of transfer learning used in deep learning is called inductive transfer. This is where the scope of possible models (model bias) is narrowed in a beneficial way by using a model fit on a different but related task.

### 2.3.1 Inception-v3

We have also used convolutional neural network with transfer learning using Inception-v3. There are several architectures that are available like AlexNet, Inception-v3, VGGNet, GoogLeNet, ResNet etc. All these architectures however use convolutional 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 Inception-v3 [17], this architecture was published on IEEE Xplore. Benchmark of inception-v3 on the ILSVRC 2012 classification

challenge validation set demonstrate substantial gains over the state of the art: 21.2% top-1 and 5.6% top-5 error for single frame evaluation using a network with a computational cost of 5 billion multiply-adds per inference and with using less than 25 million parameters [18]. With 42 layers deep, the computation cost is only about 2.5 higher than that of GoogLeNet [19], and much more efficient than that of VGGNet [20]. However, the downside of the Inception-v3 is that it is more expensive to evaluate and uses a lot of GPU memory and parameters.

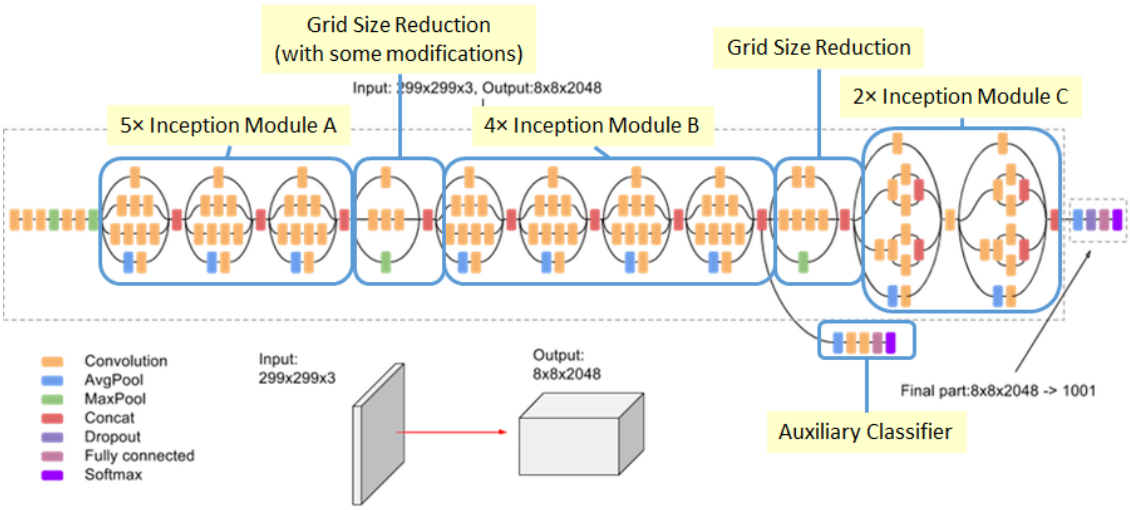



Figure 2.5: inception v3 model architecture

# Chapter 3

# Related Work

Several approaches have been done to classify diabetic retinopathy from images. In previous years many feature based model is being used to classify diabetic retinopathy images. Blood vessels, area of exudates, microaneurysms, optic disc and texture features are the common features that has been used to classify diabetic retinopathy images. In modern literature there are neural networks especially convolutional neural networks have been used to classify diabetic retinopathy images.

## 3.1 Feature Based Model

FBR is quite different from a variety of object recognition algorithms [21,22,23]. These algorithms rely on a single type of feature or a set of features. In [21] 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 vessels. They have used Top-Hat transform to enhance the blood vessels. Feature selection is carried out to extract only the significant features Texture features for classification into normal, NPDR and PDR stages. They have used two datasets (DIARETDB0, DIARETDB1) which contains 219 images where 194 images has retinopathic lesions. They have achieved the Specificity, sensitivity and accuracy 97.3%, 85% and 89.57% respectively. In [22] their main task was to detect the retinal changes of diabetes patient's eye such as microaneurysms, hard exudates, soft exudates, hemorrhage etc., 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 [23] 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 [24] exudates in color fundus were detected as well as classify the severity of the lesions using SVM classifier. In

[25] they also did the same type of work as that of [24] but in their work they have used ANN classifier. In [26] 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.
In [27] they have used the background subtraction technique and morphological technique to obtain the exudate candidate regions and the basic properties of exudates are extracted as feature for classification by the support vector machine (SVM) classifier. Then the efficient fovea localization method based on Image Relative Subtraction technique is applied. Finally, the system assesses the severity of DME based on the distance between exudates and the fovea.

## 3.2 Deep Learning-Based Model

In present day deep learning is very popular in computer vision. Deep learning (also known as deep structured learning or hierarchical learning) is part of a broader family of machine learning methods based on learning data representations, as opposed to task-specific algorithms. Learning can be supervised, semi-supervised or unsupervised. Deep learning architectures such as deep neural networks, deep belief networks and recurrent neural networks have been applied to fields including computer vision, speech recognition, natural language processing, audio recognition, social network filtering, machine translation, bioinformatics, drug design and board game programs, where they have produced results comparable to and in some cases superior to human experts.

In recent literatures there are also several approaches that used deep convolution neural network to classify diabetic retinopathy images. Since convolutional neural networks are scalable for large datasets it is more suitable to use convolutional neural network for f diabetic retinopathy image classification. Deep learning was used in [28] to classify E-Ophtha and Messidor-2 image dataset for a data-driven deep learning algorithm as a novel diagnostic tool for automated DR detection. The algorithm processed color fundus images and classified them as healthy (no retinopathy) or having DR, identifying relevant cases for medical referral.

In [29] paper they train 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 [30] they have compared between three transfer learning method on our dataset. Where they have Achieve 37.43% ,50.03% and 63.23% on AlexNet, VGG16 and InceptionNet V3 repetitively .

In our thesis we used convolutional neural network to classify diabetic retinopathy dataset. We tried with a convolutional neural network built from scratch as well as with transfer learning that uses Inception V3 [31] model.

# Chapter 4

# Methodology

In our research we worked on diabetic retinopathy detection Image Dataset [1]. We developed a convolutional neural network from scratch to classify diabetic retinopathy images. We also used transfer learning from Inception v3 model which was pre-trained with Image-net [31] method in our work. This chapter describes the model of our convolutional neural networks and transfer learning method in details. The methodology is depicted in figure 4.1.

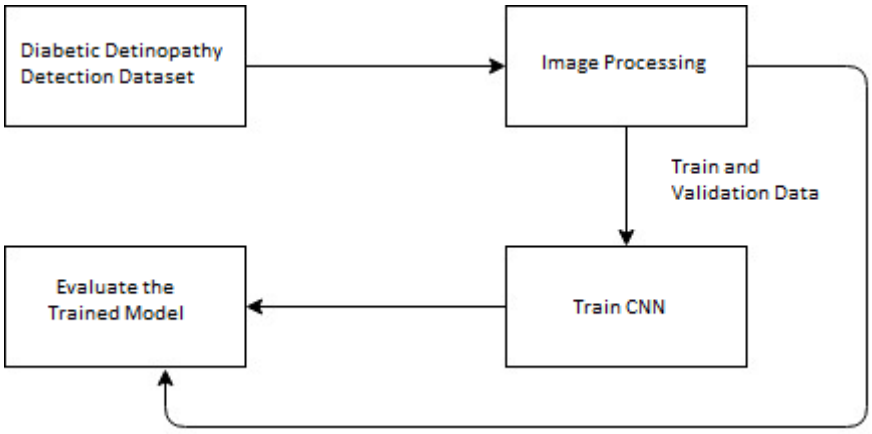



Figure 4.1: Methodology of Optical Coherence Tomography Scans Classification

## 4.1 Dataset

We have used diabetic retinopathy detection for classification dataset for our research. Retinal images were provided by EyePACS, a free platform for retinopathy screening. The dataset consists of 35,126 images grouped into 5 subclasses. The 5 subclasses are No DR, Mild, Moderate, Severe, Proliferative DR. where No DR contains 25810, Mild contains 2443, Moderate contains 5292, Severe contains 873 and Proliferative DR contains

708 images. The retinopathy images are divided into two parts: a training set with 28103 images, testing set with 7023 images.

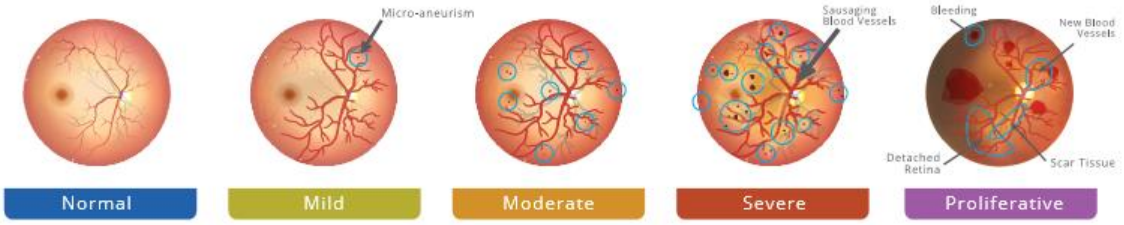



Figure 4.2: Stages of Diabetic Retinopathy

The images in the dataset come from different models and types of cameras, which can affect the visual appearance of left vs. right. Some images are shown as one would see the retina anatomically (macula on the left, optic nerve on the right for the right eye). Others are shown as one would see through a microscope condensing lens (i.e. inverted, as one sees in a typical live eye exam). There are generally two ways to tell if an image is inverted:

- It is inverted if the macula (the small dark central area) is slightly higher than the midline through the optic nerve. If the macula is lower than the midline of the optic nerve, it's not inverted.
- If there is a notch on the side of the image (square, triangle, or circle) then it's not inverted. If there is no notch, it's inverted.

Like any real-world data set, you will encounter noise in both the images and labels. Images may contain artifacts, be out of focus, underexposed, or overexposed. Our major aim of this study is to develop robust algorithms that can function in the presence of noise and variation.

## 4.2 Pre-Processing

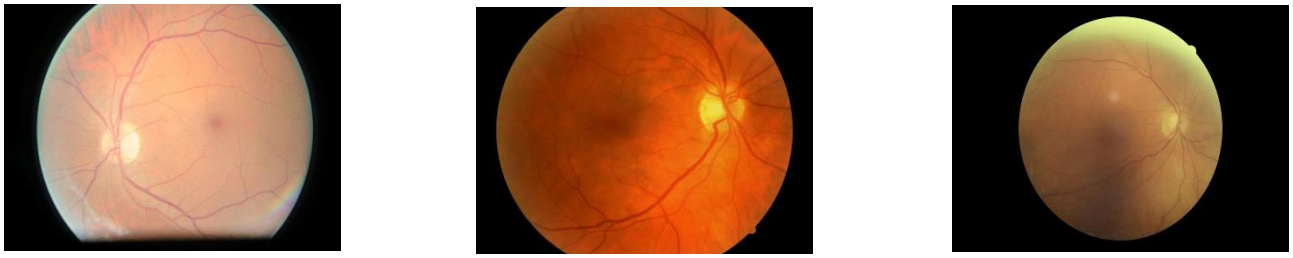
We have applied some image pre-processing technique to increase efficiency to our system. First, we re-sized all our images to 256 x 256 x 3 to increase processing time, get

more features and also to fit in our convolutional neural network model. Next, we have applied the following pre-processing techniques:

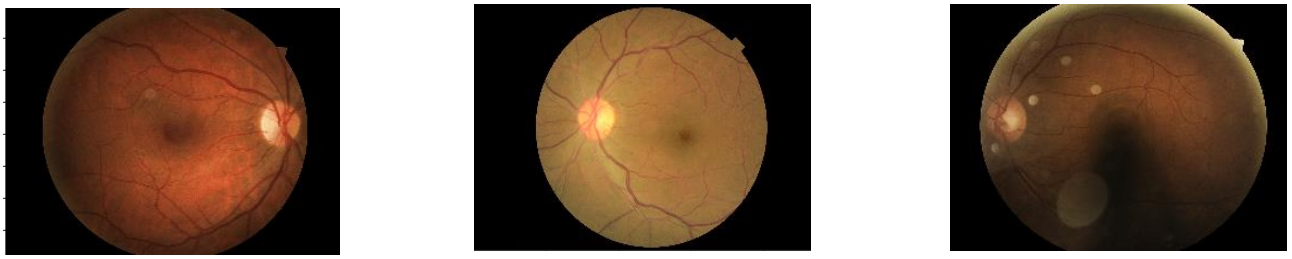
- **feature wise center**: False. Set input mean to 0 over the dataset, feature-wise.
- **sample wise center**: False. Set each sample mean to 0.
- **feature wise standard normalization**: False. Divide inputs by standard of the dataset, feature-wise
- **sample wise standard normalization**: False. Divide each input by its std.
- **ZCA whitening**: False. Apply ZCA whitening.
- **zca_epsilon:** 1e-06
- **rotation range**: 0. Degree range for random rotations.
- **width shift range**: 0.0, Randomly shift images horizontally.
- **height shift range**: 0.0, Randomly shift images vertically.
- **preprocessing_function**: None
- **horizontal_flip**: False
- **Fill mode**: “nearest”. Point outside the boundaries of the input are filled.

### 4.2.1 ZCA Whitening

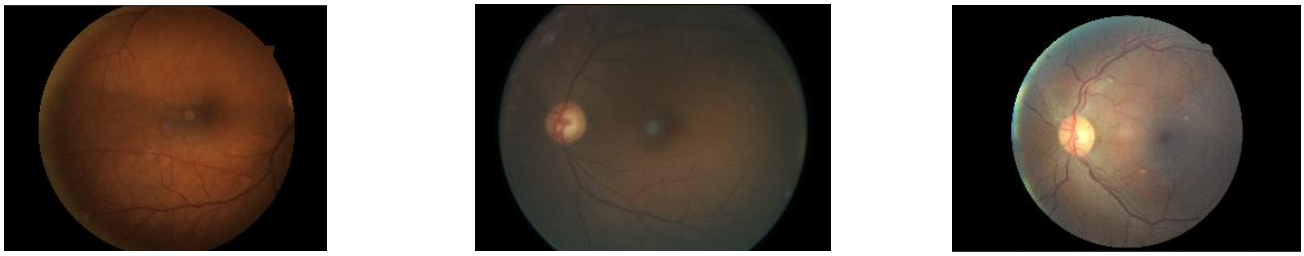
A whitening transform of an image is a linear algebra operation that reduces the redundancy in the matrix of pixel images. Less redundancy in the image is intended to better highlight the structures and features in the image to the learning algorithm. Typically, image whitening is performed using the Principal Component Analysis (PCA) technique. More recently, an alternative called ZCA (learn more in Appendix A of this tech report) shows better results and results in transformed images that keeps all of the original dimensions and unlike PCA, resulting transformed images still look like their originals.



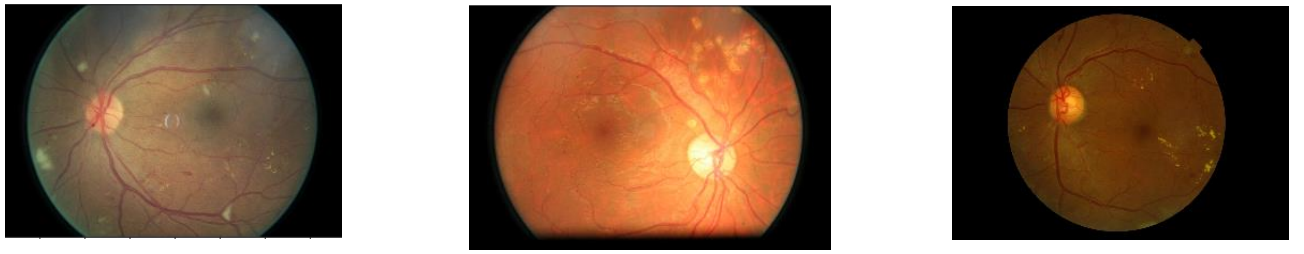
1. No DR



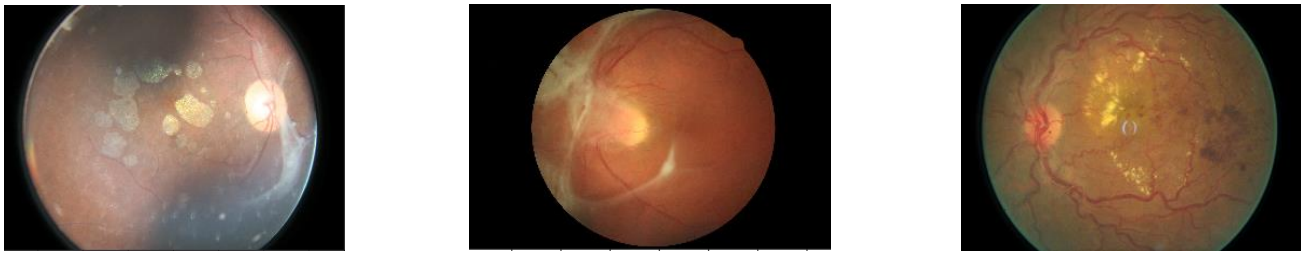
2. Mild



3. Moderate



4. Severe



5. Proliferative DR

Figure 4.3: Images from each class of the dataset

## 4.3 Training Convolutional Neural Network Model

We have developed a convolutional neural network model from scratch to classify diabetic retinopathy images.

We trained our dataset for 10 epochs with SGD optimizer with learning rate 0.01, decay 0.0, momentum 0 and nesterov False;

### 4.3.1 Stochastic Gradient Descent (SGD)

For normal Gradient Descent we need to process whole dataset that is very inefficient and expensive. Solution for solving this problem is random choice of next example that will help to update trainable parameters. In network training we take random batch of samples for each iteration and then do update for θ. This is a Stochastic Gradient Descent that widely used in training networks.

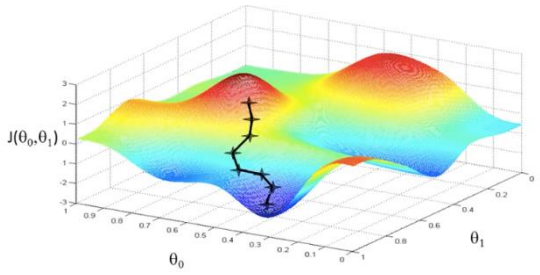



Figure 4.4: Visualization of SGD

Stochastic gradient descent is an optimization algorithm that estimates the error gradient for the current state of the model using examples from the training dataset, then updates the weights of the model using the back-propagation of errors algorithm, referred to as simply backpropagation [32].

### 4.3.2 Learning Rate

A neural network learns or approximates a function to best map inputs to outputs from examples in the training dataset.
The learning rate hyperparameter controls the rate or speed at which the model learns. Specifically, it controls the amount of apportioned error that the weights of the model are updated with each time they are updated, such as at the end of each batch of training

examples. A large learning rate allows the model to learn faster, at the cost of arriving on a sub-optimal final set of weights. A smaller learning rate may allow the model to learn a more optimal or even globally optimal set of weights but may take significantly longer to train.

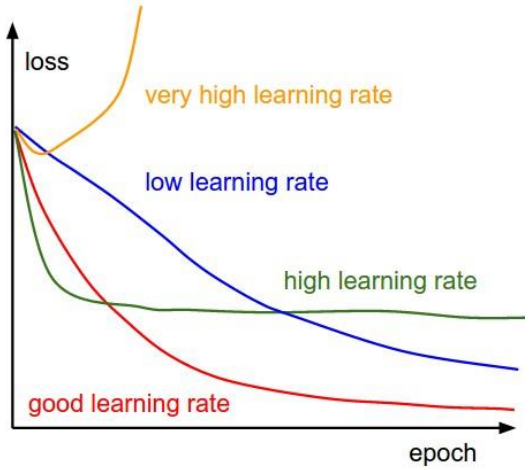



Figure 4.5: Different between types of learning rate

### 4.3.3 Decay

When training neural networks, it is common to use "weight decay," where after each update, the weights are multiplied by a factor slightly less than 1. This prevents the weights from growing too large, and can be seen as gradient descent on a quadratic regularization term

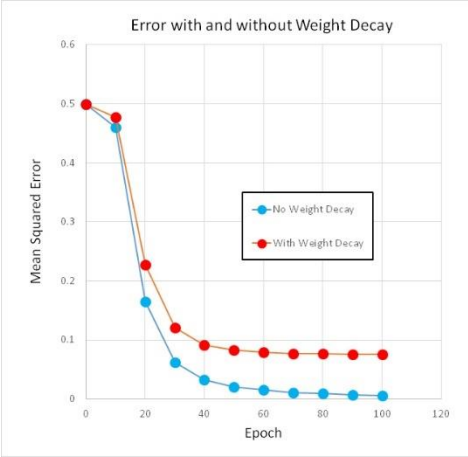



Figure 4.6: Visualization of Decay

### 4.3.8 Proposed CNN Model:

We took the preprocessed data from the input images, we first allocate random weight to the first layer of convolutional neural networks. The training is carried out by multinomial logistic regression using mini-batch gradient descent with momentum. We set out batch size to 10 and we considered our momentum as 0. 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. We have used 3 convolution layers along with a flattened layer. Throughout out model we have used ReLU activation function. For all positive values, ReLU is linear (identity) and zero for all negative values. ReLU used mainly for, it's cheap to compute as there is no complicated math. The model can therefore take less time to train or run. It converges faster. Linearity means that the slope doesn't plateau, or "saturate," when x gets large. It doesn't have the vanishing gradient problem suffered by other activation functions like sigmoid or tanh. It's sparsely activated. Since ReLU is zero for all negative inputs, it's likely for any given unit to not activate at all. This is often desirable.

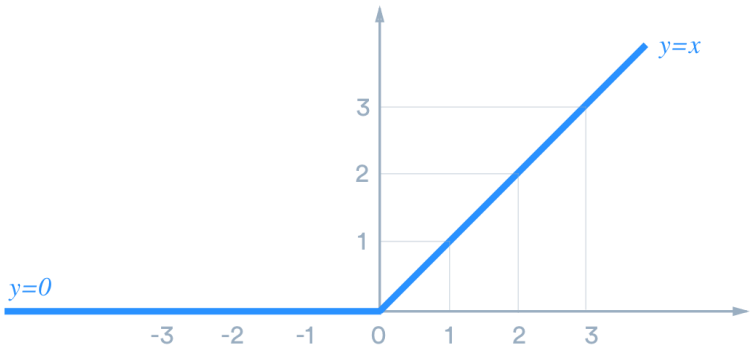



Figure 4.7: ReLu activation Function

We have used Maxpooling which is sample-based discretization process. The objective is to down-sample an input representation, reducing its dimensionality and allowing for

assumptions to be made about features contained in the sub-regions binned. This is done in part to help over-fitting by providing an abstracted form of the representation. As well, it reduces the computational cost by reducing the number of parameters to learn and provides basic translation invariance to the internal representation.

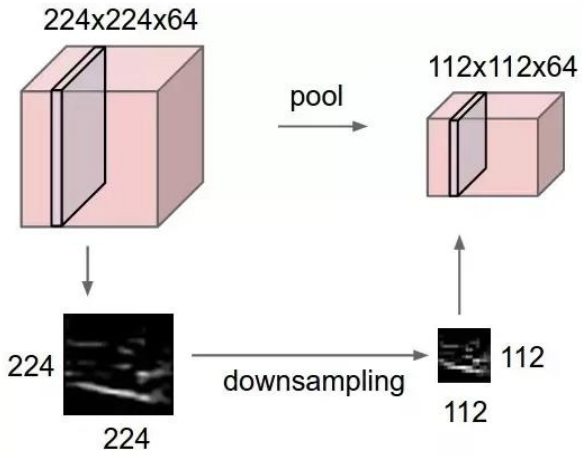



Figure 4.8: Representation of Max Pooling

We also used normalization in our convolution layers. Batch normalization is to limit covariate shift by normalizing the activations of each layer. This, supposedly, allows each layer to learn on a more stable distribution of inputs, and would thus accelerate the training of the network. Also changes appropriate learning rate wildly when there are high-order interactions between variables.

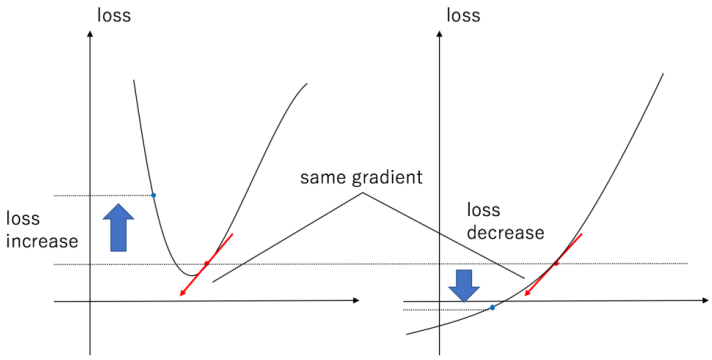



Figure 4.9: Representation of Batch Normalization

At end we flatten out convolution layer into Dense. Flattening involves transforming the entire pooled feature map matrix into a single column which is then fed to the neural network for processing.

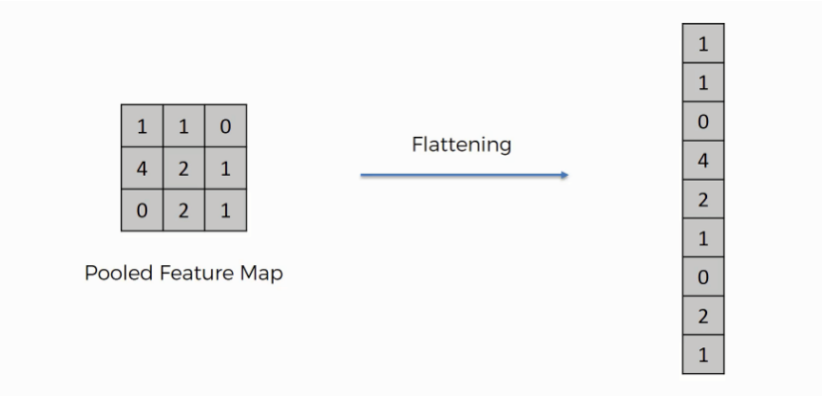



Figure 4.10: Representation of flatten layer

Finally, we build our model and train and test images with our model. below our model architecture.

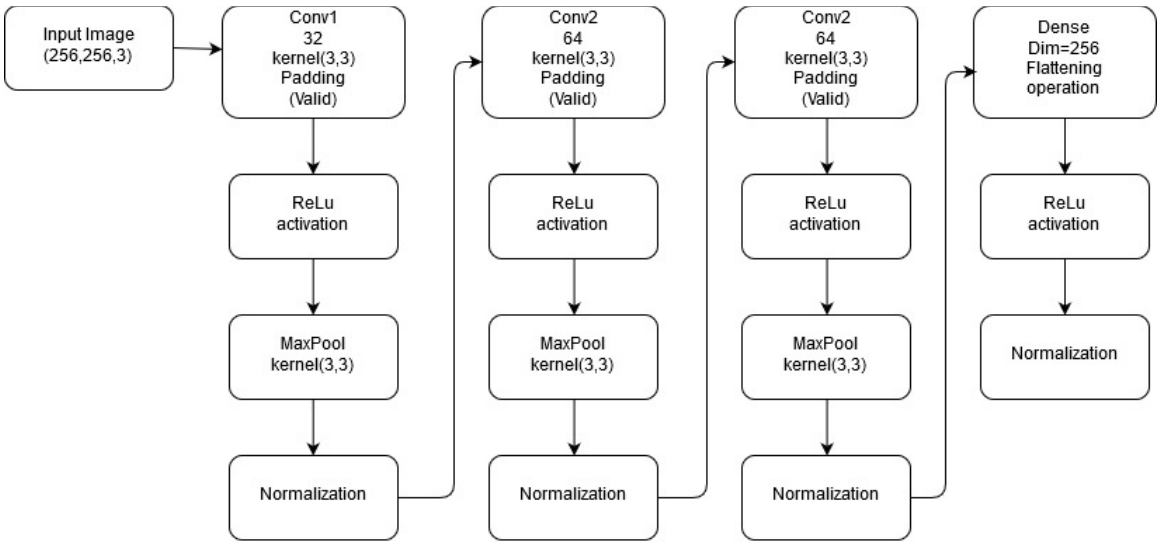



Figure 4.11: Architecture of our Convolutional Neural Network.

## 4.4 Transfer Learning

We applied also applied transfer learning in our research. We chose google inceptionv3 which was pre-trained Imagenet dataset. First Prepared the dataset, then in our previous CNN model, we are going to instantiate the InceptionV3 network from the keras.applications module, Them load the model and their weights but leaving out the last fully connected layer, since that is specific to the ImageNet competition. Finally, we compile the model selecting the optimizer, the loss function, and the metric. In this case we wear use SGD optimizer with the default learning rate of 0.01, and a “categorical_crossentropy” used in multiclass classification tasks — as loss function. We used a learning scheduler to set learning rate to 0.03 after 7 epochs.

# Chapter 5

# Experimental Results

In our thesis we used diabetic retinopathy detection dataset for our research. We developed a convolutional neural network from scratch to classify diabetic retinopathy images. We also used transfer learning from Inception v3 model which was pre-trained with Image-net [32] method in our work. This chapter describes the model of our convolutional neural networks and transfer learning method in details. The methodology is depicted in figure 5.1

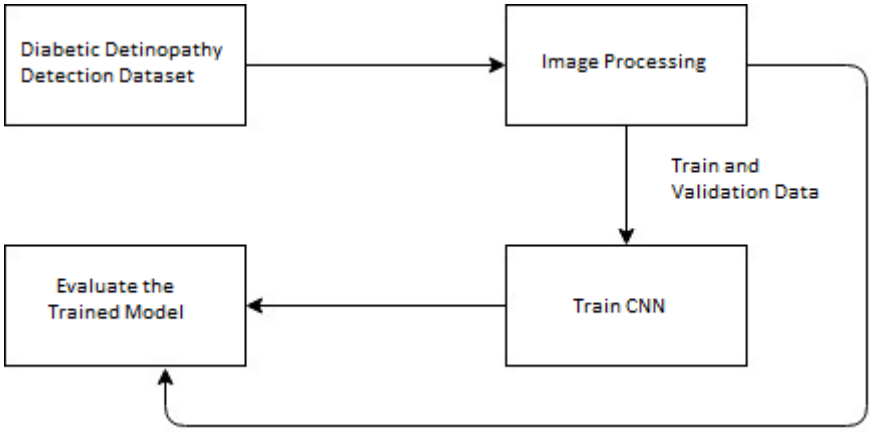



Figure 5.1: Methodology of diabetic retinopathy image classification

## 5.1 Evaluation of Model

Our dataset was divided into three parts: training, validation and evaluation. We used training and validation parts of the dataset while training the model and we used evaluation part of our dataset during the evaluation of our model. We resize the images of evaluation part in 224 x 224 x 3. We evaluated the accuracy of the model by true positive (TP), true negative (TN), false positive (FP) and false negative (FN) after classification.

$$Accuracy = \frac{TP}{TP + FP + TN + FN}$$

## 5.2 Obtained Results

The results obtained by running different models with diabetic retinopathy dataset are given in Table 5.1.

| Model | Accuracy |
|---|---|
| Harry Pratt's Proposed CNN in [28] | 75.12% |
| Proposed CNN (only disease) | 61.27% |
| Proposed CNN | 75.87% |
| Inception V3 (Transfer Learning) | 86.82% |

Table 5.1: Test results of models

We can see that our model gives more accuracy then the model proposed in the same dataset in [28]. We also see that Inception V3 has an accuracy of 86.82%. Since inception v3 model has been already pre-trained on image-net it transferred it learning in our dataset. So, inception v3 has better result as expected.

5.2. We see that the accuracy curve. almost every epoch gaining some accuracy.

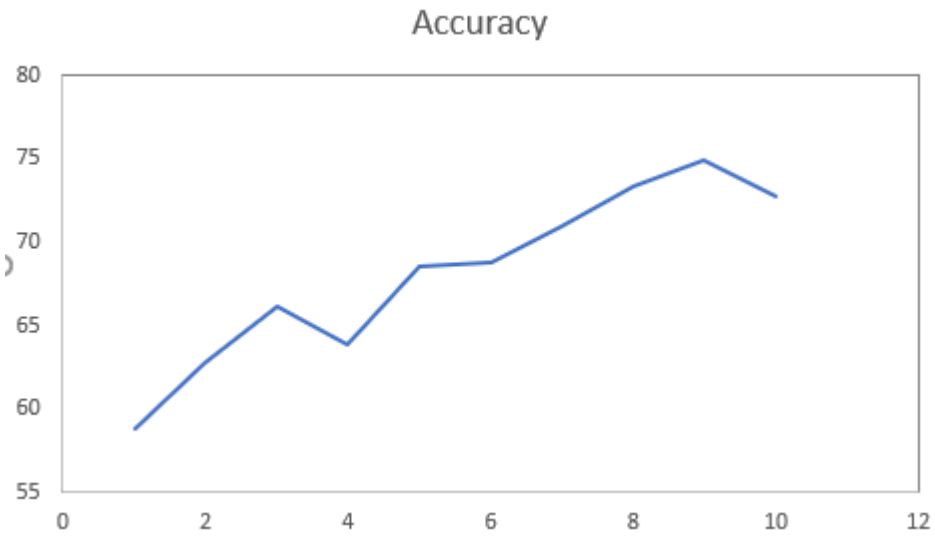



Figure 5.1: Accuracy curve of proposed model.

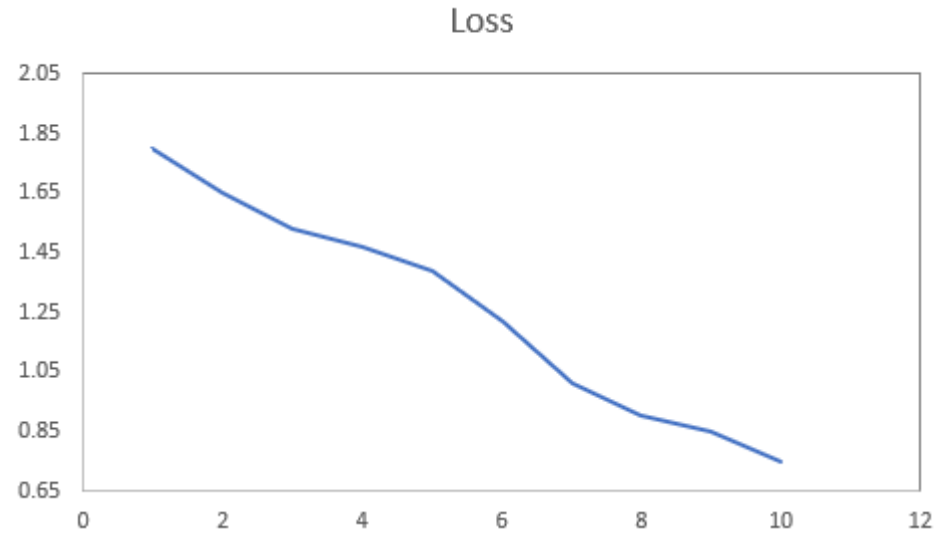



Figure 5.2: Loss curve of proposed model.

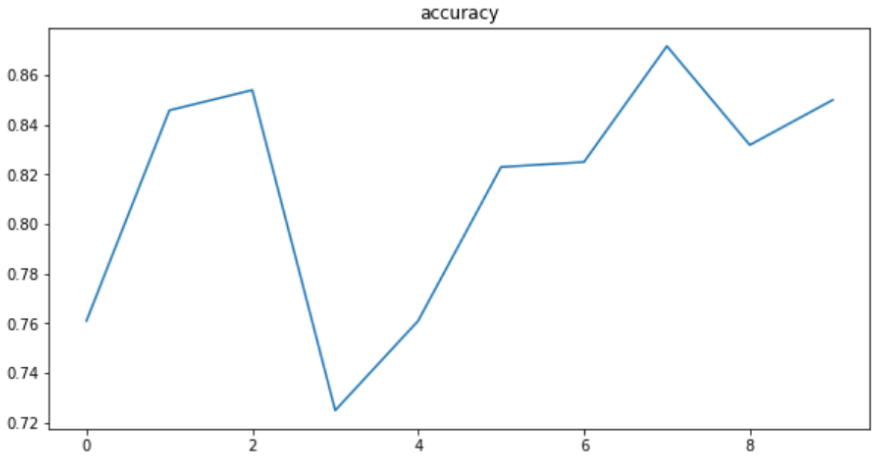



Figure 5.3: Accuracy curve of Inception V3 (Transfer Learning).

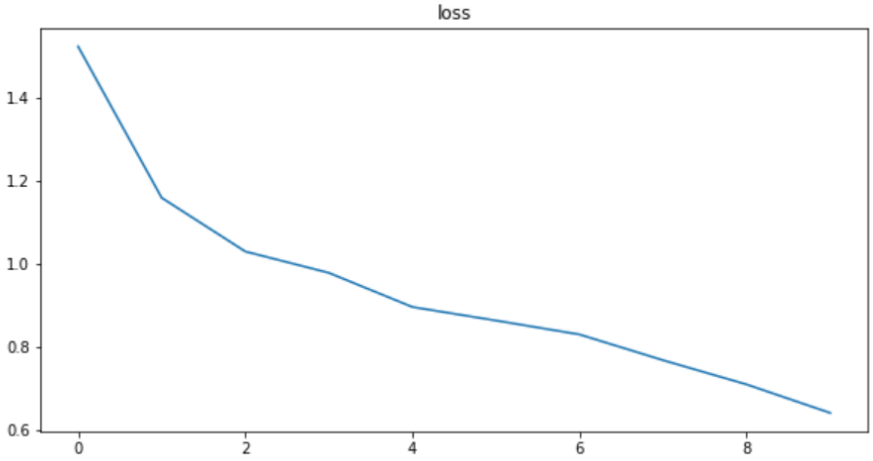



Figure 5.4: Loss curve of Inception V3 (Transfer Learning).

# Chapter 6

# Conclusion

In our paper, we proposed a system which will be able to detect diabetic retinopathy from the image of the eye of a patient. With our proposed system, the doctors can however spend 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.

# Bibliography


1. https://www.kaggle.com/c/diabetic-retinopathy-detection 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. International Diabetes Federation (2017). *IDF Diabetes Atlas, 8th edn*. Brussels, Belgium: International Diabetes Federation.

4. Kertes PJ, Johnson TM, eds. (2007). *Evidence Based Eye Care*. Philadelphia, PA: Lippincott Williams & Wilkins. ISBN 978-0-7817-6964-8.

5. Caroline MacEwen. "diabetic retinopathy". Retrieved August 2, 2011.

6. 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.

7. Engelgau, Michael, Linda Geiss, Jinan Saaddine, Jame Boyle, Stephanie Benjamin, Edward Gregg, Edward Tierney, Nilka Rios-Burrows, Ali Mokdad, Earl Ford, Giuseppina Imperatore, K. M. Venkat Narayan. "The Evolving Diabetes Burden in the United States." Annals of Internal Medicine, 1 June 2004. Web. 22 Apr. 2014.

8. "Orbis", Orbis.org, 2016. [Online]. Available: http://www.orbis.org. [Accessed: 01-Apr2016].

9. Tapp RJ, Shaw JE, Harper CA, de Courten MP, Balkau B, McCarty DJ, Taylor HR, Welborn TA, Zimmet PZ (June 2003). "The prevalence of and factors associated with diabetic retinopathy in the Australian population". *Diabetes Care*. **26** (6): 1731–7.

10. 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.

11. N. Cheung, P. Mitchell and T. Wong, "Diabetic retinopathy", The Lancet, vol. 376, no. 9735, pp. 124-136, 2010.

12. Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” in Proceedings of the IEEE, 1998, pp. 2278–2324.

13. M. Matsugu, K. Mori, Y. Mitari, and Y. Kaneda, “Subject independent facial expression recognition with robust face detection using a convolutional neural network,” vol. 16, no. 5, 2003, pp. 555 – 559, advances in Neural Networks Research: IJCNN ’03.

14. W. Zhang, K. Itoh, J. Tanida, and Y. Ichioka, “Parallel distributed processing model with local space-invariant interconnections and its optical architecture,” vol. 29, no. 32. OSA, Nov 1990, pp. 4790–4797.

15. W. E. Grimson, “On the recognition of curved objects,” vol. 11, no. 6. Washington,DC, USA: IEEE Computer Society, Jun. 1989, pp. 632–643.

16. https://cloud.google.com/tpu/docs/inception-v3-advanced

17. Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jonathon Shlens, Zbigniew Wojna ” Rethinking the Inception Architecture for Computer Vision” *11 Dec 2015 (this version, v3)*

18. Christian Szegedy , Wei Liu , Yangqing Jia , Pierre Sermanet , Scott Reed , Dragomir Anguelov , Dumitru Erhan , Vincent Vanhoucke , Andrew Rabinovich “Going Deeper with Convolutions”

19. Karen simonyan & andrew zisserman “very deep convolutional networks for large-scale image recognition” 10 apr 2015.

20. 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.

21. 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.

22. 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.

23. M. Gandhi and D. Dhanasekaran, "Diagnosis of Diabetic Retinopathy Using Morphological Process and SVM Classifier", International conference on Communication and Signal Processing, 2013.

24. 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.

25. 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.

26. Yu, Ting & Ma, Yide & Li, Wen. (2015). Automatic detection of exudates and fovea for grading of diabetic macular edema in fundus image.

27. Rishab Gargeya, Theodore Leng, Automated Identification of Diabetic Retinopathy Using Deep Learning, Ophthalmology, Volume 124, Issue 7,2017, Pages 962-969, ISSN 0161-6420,

28. Harry Pratt, Frans Coenen, Deborah M. Broadbent, Simon P. Harding, Yalin Zheng,Convolutional Neural Networks for Diabetic Retinopathy, Procedia Computer Science, Volume 90,2016, Pages 200-205, ISSN 1877-0509.

29. X. Wang, Y. Lu, Y. Wang and W. Chen, "Diabetic Retinopathy Stage Classification Using Convolutional Neural Networks," *2018 IEEE International Conference on Information Reuse and Integration (IRI)*, Salt Lake City, UT, 2018, pp. 465-471

30. C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the inception architecture for computer vision,” vol. abs/1512.00567, 2015.

31. J. Deng, W. Dong, R. Socher, L. Li, K. Li, and L. Fei-Fei, “Image-net: A large-scale hierarchical image database,” in 2009 IEEE Conference on Computer Vision and Pattern Recognition, June 2009, pp. 248–255.

32. https://machinelearningmastery.com/learning-rate-for-deep-learning-neural-networks/