DETECTION OF DIABETIC RETINOPATHY USING CNN

A MINI PROJECT REPORT

Submitted by

NICY F

REG NO : LTKM17MCA042

RESHMA ABRAHAM

REG NO : LTKM17MCA044

to

The APJ Abdul Kalam Technological University

In partial fulfillment for the award of the degree of

MASTER OF COMPUTER APPLICATIONS

Thangal Kunju Musaliar College of Engineering
Kerala

NOVEMBER 2019

DECLARATION

We undersigned hereby declare that the project report ("Detection of diabetic retinopathy using convolutional neural network(CNN)"),submitted for partial fulfillment of the requirements for the award of degree of Master of Computer Applications of the APJ Abdul Kalam Technological University, Kerala is a bonafide work done by us under supervision of Prof.Fousia M Shamsudeen. This submission represents our ideas in our own words and where ideas or words of others have been included, we have adequately and accurately cited and referenced the original sources. We also declare that we have adhered to ethics of academic honesty and integrity and have not misrepresented or fabricated any data or idea or fact or source in my submission. We understand that any violation of the above will be a cause for disciplinary action by the institute and/or the University and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been obtained. This report has not been previously formed the basis for the award of any degree, diploma or similar title of any other University.

Place
Date

Nicy F
Reshma Abraham

Thangal Kunju Musaliar College of Engineering
Department of Computer Applications

C E R T I F I C A T E

This is to certify that, this report entitled **DETECTION OF DIABETIC RETINOPATHY USING CNN** is a bonafide record of the work submitted by **NICY F (LTKM17MCA042)**, **RESHMA ABRAHAM(LTKM17MCA044)** to the **APJ Abdul Kalam Technological University** in partial fulfillment of the requirements for the award of the Degree of **Master of Computer Applications** is a bonafide record of the project work carried out by them under our guidance and supervision. This report in any form has not been submitted to any other University or Institute for any purpose.

Internal Supervisor   Mini Project Coordinator   Head of the Department

ACKNOWLEDGEMENT

First and for most we thank GOD almighty and our parents for the success of this project. We owe sincere gratitude and heart full thanks to everyone who shared their precious time and knowledge for the successful completion of our project.

We are extremely grateful to **Dr. Nadera Beevi S**, Head of the Department, Department of Computer Applications, for providing us with best facilities.

We would like to thank our mini project coordinator, **Prof.Sheera Shamsu** and our guide **Prof.Fousia M Shamsudeen**, Department of Computer Applications, who motivated us throughout the project.

We profusely thank all other faculty members in the department and all other members of TKM College of Engineering, for their guidance and inspirations throughout our course of study.

We owe our thanks to all friends and all others who have directly or indirectly helped us in the successful completion of this project.

Nicy F
Reshma Abraham

Abstract

Diabetic retinopathy (DR) is a grievous and most common retinal disease, are a common complication of diabetes and one of the major causes of blindness in humans.it occurs mainly due to the damage of retinal blood vessels by the high glucose blood level, the different extent of microstructure, such as micro-aneurysms, hard exudates could occupy the retinal area. Usually, there are no early visible symptoms of Diabetic Retinopathy. In the current clinical diagnosis, the detection mainly relies on the ophthalmologist examining the color fundus image and then scrutinize the patient's condition. This detection is strenuous and prolonged, which results in more errors.

The paramount objective of this project is to present a method to detect diabetic retinopathy on the fundus images by using deep learning-based Convolutional Neural Network.CNN has recently turn out to be a reassuring approach in image analysis which includes 3 major difficult challenges: pre-processing,classification, and detection. In this research, we used the publicly available Kaggle dataset of retina images to train an ensemble of deep Convolution Neural Network(CNN) to encode the rich features and improve the classification for different stages of DR. The experimental results show that the proposed model detects all the stages of DR unlike the current methods and performs better and provide an accuracy of 90%.

Contents

List of Figures

Chapter 1

Introduction

Diabetic Mellitus, commonly known as diabetes, is a metabolic disorder that root high blood glucose level over an elongated period.Individuals who have diabetic Mellitus have a higher risk of developing Diabetic Retinopathy due to damages of the retinal blood vessel to provide relevant treatment and prevent visual loss ,it's extremely important to categorize and automate the diagnosis of diabetic retinopathy based upon the severity,so we use Convolutional neural network(CNN) for automatically diagnosing. On the Wiscons in epidemiologic study of diabetic retinopathy in 2003 classify Diabetic Retinopathy into 5 Sages including Stage 1)No apparent Retinopathy,2)Mild Non-proliferative DR, 3)Moderate non-proliferative DR,4)Severe non-proliferative DR,5)Proliferative DR

The Convolutional neural network (CNN) is a category of neural network that is proven to be effective in image recognition and classification.CNN extends the regular neural network by adding the operations of convolution,non-linearity,and sub sampling. The purpose of convolution is to extract features from the input images by convoluting the input images with some specially chosen small square matrices, certain image processing effects such as edge detection, sharpening could be realized. Another operation called Rectified Linear Unit (ReLU) that could be used after every convolution operation is a nonlinear operation by replacing all negative values in the feature operation by replacing all negative pixel values in the feature map by zero. The purpose of Re LU is to introduce non-linearity. The third operation is called pooling or sub-sampling reduces the dimensionality of each feature map while retaining the most important information.

1.1 Objective

The main objective of the project is to automatic detection of diabetic retinopathy using CNN with high accuracy and to identifying the stages of diabetic retinopathy by image processing technique.

Chapter 2

Literature Survey

2.1 Theoretical Investigations

Diabetic retinopathy is a common complication of diabetes and one of the major causes of blindness in the active population. Many of the complications of DR can be prevented by blood glucose control and timely treatment.Since the varieties and the complexities of DR, it is difficult for detection and manual diagnosis is time consuming.

2.2 Optimisation studies

Fousia.M.Shamsudeen[8] stated that the Contrast Limited Adaptive Histogram Equalization (CLAHE) is one of the widely establihed contrast enhancement scheme for fundus images.But the disadvantage in CLAHE is the contrast and quality of the enhanced image heavily depend on the colour model used to represent the contextual image.In her article that investigate the colour model, adequate channel to be equalized and the histogram specification which offer maximum grey level contrast and image quality in fundus images, when enhanced with CLAHE.. She has been observed that RGB model out performs HSV model, the adequate channel to be equalized is the green and exponential histogram is the apt choice for fundus imagery.

T.A.Samoora[24] stated that the computerized image analysis has been successfully applied in various medical applications since 1982.Image segmentation has been seen as a wide application at that stage before the1990s. Recently, more artificial intelligence-based approaches such as supervised methods have become popular in this area. Pattern recognition and machine learning methods are very productive due to their success in the medical image analysis system.

Fousia.M.Shamsudeen[7] stated that the an objective method was demonstrated to identify the optimum value of gain in the modified sigmoid transform. The best trade-off between dynamic range compression and saturation is obtained at the objectively defined gain.

Hubel et.al [6] stated that The history of CNNs started with the experiments conducted by Hubel and Wiesel in 1959 [25]. They found that the cells in animals visual cortex are responsible for detecting the light in the receptive fields.

In 1990, LeCun et al. [31] published a paper where they described the modern framework of CNN. They introduced a neural network called LeNet-5 which is used to classify the handwriting digits. They used back propagation algorithm to train the neural network. However, due to the lack of data and the computation power at that time, the proposed networks couldn't perform well in large-scale problems. After that, many researchers have developed methods to overcome some problems encountered in training deep CNNs Different machine learning based approaches have shown better performance in this area. Support Vector Machine(SVM),decision tree,Naive Bayes and Random Forest[10],[11] are examples of these and have been investigated by researchers.

Rubini and Kunthavai [21] proposed to apply hessian-based candidate selection before the feature extraction and classification using a support vector machine (SVM) classifier In 2006, a deep convolutions neural network (CNN) structure was designed for different imaging tasks of natural images [12] that intended to classify different features of color images.

Mookiah et al. [15] proposed a system that used hybrid features, including exudate vessel area, texture, and entropy, for DR classification.

AlexNet[1],VGGNet[1] and the inception of the architecture of GoogLeNet[14] have been applied successfully to a large variety of object identification and semantic segmentation.

Bhatkar and Kharat [2] explored the use of the multilayer perception neural network as the classifier to process extracted features, such as a 64-point discrete cosine transform,and other statistical features,including entropy and Euler's number.

Krizhevsky et al. [1] contributed to the watershed transformation of the image by using the ImageNet challenge in 2012, and they proposed a CNN model called AlexNet,winning the challenge with a significant margin. Due to the good performance of the CNN model, the medical image analysis community has considered the implementation of CNN.

Xiaoling wang et.al [33] stated that they also use the above CNN architecture for DR stage classification. In these they go through the three models and find out that the important feature of AlexNet is the introduction of the ReLU nonlinearity into the training of neural network and also support "dropout" and advantage of VGG16 is its simple and standardized approach of constructing the hidden layers.the average of 5 fold cross-validation accuracy obtained by Inception NETv3 with the highest accuracy 63.3%

Wang et al. [19] proposed a novel architecture that classifies the images as normal abnormal, referable non-referable DR and gets the high AUC on a normal and referable DR task 0.978 and 0.960 respectively and specificity is 0.5. Their proposed method uses three networks: main, attention and crop. The main network uses the Inception model that is trained on ImageNet where the attention network highlights different types of lesions in the images and crop networks crop the high attention image.

Quelle [7] proposed three CNN models for binary classification and detected DR lesions. They also used the Kaggle and DiaretDb1 data set for training and testing respectively. Diabetic retinopathy has five (5) stages to classify the occurrence of diseases.

The stage-wise classification is discussed by Chanrakumar and Kathirvel [25] introduced the CNN model with a dropout regularization technique trained on the Kaggle dataset and tested on DRIVE and STARE dataset. The accuracy achieved by their model is 94%. They manually performed an augmentation and pre-processing steps by using an image editing tool accuracy achieved by their model is 94%. Memon et al. [27] stated that the pre-processing is done on the dataset and they used nonlocal means de-noising and added a delta value to get the equal level of brightnessof theimages. For evaluation, the over all kappa score accuracy is0.74, for the validation purpose, 10% of the images were used.

Pratt et al. [9] proposed a CNN architecture used for classifying five stages but could not classify the mild stage accurately, due to the nature of architecture. Another limitation is that they used the skewed dataset of Kaggle that led to the high specificity with the trade-off in low sensitivity

Yang et al. [30] proposed DCNN (Deep Convolution Neural Network) for two stages of DR (normal and NPDR). The pre-processed data is given as input to the two networks (local and global). Lesions are highlighted and sent to the global network for grading. They used class weight and kappa scores for evaluation of the model. However, the PDR stage was not considered in their work

Zhangetal. [12]was one of the first researchers who worked on detection of retinal blood vessels based on neural network techniques. They used some self organizing map (SOM) as pre-processing techniques to train the network to get retinal blood vessels,and their proposed unsupervised method gave a good retinal vessels segmentation. Their method consists of three steps. First, a multi dimensional feature vector from the green channel of the RGB retinal image intensity is constructed, and then vessels enhanced intensities feature is achieved by using morphological operation.
Zillya et al. [13] proposed a method to identify of retinal blood vessels by using ensemble learning late fusion on convolutional neural network (CNN). Their method is based on the following three steps,
1) An entropy sampling method is used to select the informative points to reduce the computational complexity for performing the uniform sampling.
2) A CNN model based on convolutional filters was used.
3) A softmax logistic classifier is used to fuse the output of all learned filters, and to test the trained model on DRIVE database. Their observation was based on the graph cut algorithm because the output of the classifier is subjected to an unsupervised graph followed by a convex hull transformation to achieve the final segmentation of retinal vessel
In [16], [23], the authors checked the performance of the Kaggle dataset over different CNN models
Garcia et al. [8] proposed a method of using the right and left eye images separately and applied CNN (Alexnet, VGGnet16, etc.). The pre-processing and augmentation phases were performed on the dataset to improve the contrast of images. They achieve the best results on VGG16 with no fully connected layer and achieved 93.65% specificity, 54.47% sensitivity, and 83.68% accuracy. However, DR stages were not explicitly classified in their work.

Tanetal. [11]proposed a seven-layer convolutional neural network (CNN) method. Their method not only detects the retinal blood vessels but also the optic disc and fovea of the retinal image. Due to detection of more than one feature at the same time, the accuracy of segmenting the retinal

blood vessels was negatively affected.
They used the normalization technique before segmentation to remove non-uniform background and noise in retinal images in such way it makes consistency in background lighting and contrast. The contribution of their work contains the selection of every effective vessels pixel in fundus image of three color channels before feeding into following CNN.

Liskowski and Krawiec[16] proposed an unsupervised deep neuralnetwork(DNN) for analyzing retinal blood vessels. They used pre-processing steps such as global contrast normalization, zero-phase whitening, and augmentation using geometric transformations and gamma corrections for enhancing the intensities of blood vessels against their background.

Yao et al. [32] proposed a method based on the CNN and post-processing steps to get the retinal vessels segmentation. The CNN checks each pixel with its neighbor soft he retinal fundus image for making better contrast of retinal vessels against their background. Segmentation of retinal fundus image is achieved by using binarization as post-processing. The binarization contains two steps: First, they used local multi-scale and global normalization imaging technique to achieve the initial binarization results. They named this step as local binarization.

Mahapatra et al. [4] implemented a method for retinal vessels detection which is based on the local saliency map and CNN model to get the vessels image. They used the unsupervised information from local saliency maps and the supervised information from the trained convolutional neural network. The final result is achieved by a combination of both saliency maps from unsupervised information and they train CNN from supervised information.

Fu et al. [10] developed a novel method for detecting retinal blood vessels based on probability map using CNN. They formulated the vessel extraction with consideration of boundary detection problem. The CNN is used to generate retinal. vessels probability map. The vessel probability map differentiates vessels and background pixels in low contrast region. Their method performs better in the pathological images due to robustness of differentiating vessels and background pixels in low contrast region. Afterwards, they utilized fully-connected Conditional Random Fields (CRFs) with combination of the discriminative vessel probability

Maji et al. [5] proposed a method by using Conv Net ensemble based CNN architecture for processing the fundus color image in order to get retinal blood vessel. They used ensemble learning based on multiple models which seek to promote diversity among models with a combination, and it helps to reduce the problem of overfitting of training data. The main contribution of their method was an integration of ensemble learning with ConvNet to improve the generalization and this approach was a heuristics independent. In general, it provides a useful solution to solve complex medical data like retinal images.

Dutta et al. [18] used Kaggle dataset with three deep learning models (Feed Forward Neural Network (FNN), Deep Neural Network (DNN), and Convolutional Neural Network (CNN). They used 2000 images out of 35128 images with a 7:3 validation split. They applied many pre-processing steps (median mean, STD deviation,etc.) and then trained their model on the training dataset. The Best training accuracy of 89.6% was obtained on DNN.

Shaohua Wan et.al [23] stated that in addition to the following architecture they adopt GoogleNet, ResNet and analyze how well this model does with the DR image classification. GoogleNet used a 22layers deep CNN,which is small and more speedy than VggNet and smaller and more precise than AlexNet and the network structure is more complex than VGGnet adding 'Inception' layers to the network structure each an 'Inception' layer contain six convolutional and one pooling operation, which decreases the thickness of fusion feature image and got accuracy.

Sergio Bortolin junior et al. [20] has proposed an automated detection of microaneurysm and hemorrhages in color eye fundus images. This methods consists offive methods: pre-processing,enhancement of low intensity structure,detection of blood vessels, elimination of blood vessels, and elimination of fovea. Green channel and CLAHE are used for pre-processing. Enhancement of low intensity has been achieved with the helo of applying alternating sequential filtering(ASF). Detection of blood vessels and elimination of blood vessels was performed by applying ASF and morphology opening with multiscale structuring element

Sarni Suhaila Rahim et al. [2] proposed several techniques for detection of microaneurysm. In system I,they have used adaptive histogram equalization,discrete wavelet transform, and filtering and morphology process for pre-processing. Area of pixels, mean and standard deviation are the extracted features of DR. Decision tree,K-nearest neighbor,polynomial kernel S V Mand Radial basis function(RBF) kernel SVM have been used for classification. Result of system I has been shown in [3]. They used histogram equalization, shade correction, vessel segmentation, and morphological operation for pre-processing. Area of pixels, mean, standard deviation are extracted features from the pre-processed fundus images. Decision tree, KNN, and SVM has been used for detection of diabetic retinopathy.

Sopharak A et al. [3] proposed hybrid method for fine MAs detection from nondilated DR retinal images,using mathematical morphology,naive Bayes classifier.

Adal K M et al. [14 used scale-adopted blob analysis and semi-supervised learning for automated detection of microaneurysms and evaluate the performance on ROC competition database R.A.Welikalaetal. [17]used two vessel segmentation methods,such as standard line operator and modified line operator and latter apply SVM for dual classification.

Gondal et al. [29] proposed a CNN model for the referable Diabetic Retinopathy (RDR). They used two publicly available datasets Kaggle and DiaretDB1,where the Kaggle dataset is used for training and DiaretDB1is used for testing. They are doing binary classificationas normal and mild stages are considered as non-referable DR where the rest of the three stage are used as referable DR.The performance of the CNN model is evaluated based on binary classification resulting in sensitivity 93.6% and specificity 97.6% on DiaretDB1.

ResNet is put forward by Kaiming.He stated that it trains the dataset with less error rate and its quite prominent though the number of parameters is less than VGGNet.The ResNet architecture is decent that greatly accelerate the training of ultra-deep neural network and improve the accuracy of the model and they got a 90% accuracy by using this model.

Toufique Ahmed Soomro et.al [26] presents a comprehensive review of the principle and application of deep learning in retinal image analysis.it contains an overview of deep learning methods and explains CNN, an explanation of the importance of segmentation by using CNN, details of measuring parameters and the user database in the existing methods.The state of art of retinal vessel segmentation methods based on deep learning, and analyze the methodologies of each method, analyze the main issues in retinal vessel methods based deep learning then concluding remarks and future research directions.

We can summarize all these works of DR Classification problems into two groups. The first one is the binary classification of DR i.e. either the patient has DR or not. The problem in this method is we cannot identify the severeness of the disease even after knowing that a patient has DR. The solution to this problem is multi-class classification. In multi-class classification, we classify DR into five different classes or stages as discussed in the introduction section. But most of the related work is unable to properly classify all the stages of DR, especially the early stages. It is impor tant to predict the DR at early stages for the cure as in later stages it is difficult to cure the disease and can lead to blindness. To the best of our knowledge, no other work has detected the mild stages of DR, by using the Kaggle dataset which we have used for our research. Our model can detect the mild stage and performs better than the current state of the art. Moreover, intherelated work, no one has shown the effect of a balanced dataset. The imbalanced dataset can lead to bias in the classification's accuracy. If the samples in the classes are equally distributed as in the case of a balanced dataset then the network can learn the features properly, but in case of unequal distribution network outperforms for highly sampled class. Moreover, the current CNN architectures for DR detection lack to analyze the effect of different hyperparameters tuning (meta-learning) and its implications.

Chapter 3

Methodology

IMAGE PREPROCESSING

3.1 Pre-processing Techniques

Pre-processing is a common name for operations with images at. the lowest level of abstraction - both input and output are intensity images.
The aim of pre-processing is an improvem-ent of the image data that suppresses unwilling distortions or enhances some image features important for further process-ing, although geometric transformations of images (e.g. rotation, scaling, translation) are classified among pre-processing. The acquired data are usually messy and come from different sources. To feed them to the ML model (or neural network), they need to be standardized and cleaned up. More often than not, preprocessing is used to conduct steps that reduce the complexity and increase the accuracy of the applied algorithm.
Image processing could be simple tasks like image resizing. In order to feed a dataset of images to a convolutional network, they must all be the same size.
Data preprocessing techniques might include:

3.1.1 Convert color images to grayscale

In certain problems it is useful to lose unnecessary information from the images to reduce space or computational complexity. For example, converting the colored images to gray scale images. This is because in many objects, color is not necessary to recognize and interpret an image. Gray scale can be good enough for recognizing certain objects. Because color images contain more information than black and white images, they can add unnecessary complexity and take up more space in memory. Converting it to gray scale reduces the number of pixels that need to be processed.
Humans perceive color through wavelength-sensitive sensory cells called cones. There are three different types of cones, each with a different sensitivity to electromagnetic radiation (light) of different wavelength. One type of cone is mainly sensitive to red light, one to green light, and one to blue light. By emitting a controlled combination of these three basic colors (red, green and blue), and hence stimulate the three types of cones at will, we are able to generate almost any perceivable color. This is the reasoning behind why color images are often stored as three separate image matrices; one storing the amount of red (R) in each pixel, one the amount of green (G) and

one the amount of blue (B). We call such color images as stored in an RGB format.
In gray scale images, however, we do not differentiate how much we emit of the different colors; we emit the same amount in each channel. What we can differentiate is the total amount of emitted light for each pixel; little light gives dark pixels and much light is perceived as bright pixels. When converting an RGB image to gray scale, we have to take the RGB values for each pixel and make as output a single value reflecting the brightness of that pixel. One such approach is to take the average of the contribution from each channel: (R+B+C)/3. However, since the perceived brightness is often dominated by the green component, a different, more ”human-oriented”, method is to take a weighted average, e.g.: 0.3R + 0.59G + 0.11B.

3.1.2 Standardize images

One important constraint that exists in some machine learning algorithms, such as CNN, is the need to resize the images in your dataset to a unified dimension. This implies that our images must be preprocessed and scaled to have identical widths and heights before fed to the learning algorithm.

3.1.3 Data augmentation

Another common pre-processing technique involves augmenting the existing dataset with perturbed versions of the existing images. Scaling, rotations and other affine transformations are typical. This is done to enlarge your dataset and expose the neural network to a wide variety of variations of your images.When applying convolutional neural networks for image classification, it can be challenging to know exactly how to prepare images for modeling, e.g. scaling or normalizing pixel values.
Further, image data augmentation can be used to improve model performance and reduce generalization error and test-time augmentation can be used to improve the predictive performance of a fit model.
Image data augmentation is a technique that can be used to artificially expand the size of a training dataset by creating modified versions of images in the dataset. Training deep learning neural network models on more data can result in more skillful models, and the augmentation techniques can create variations of the images that can improve the ability of the fit models to generalize what they have learned to new images.
Convolutional neural network, or CNN, can learn features that are invariant to their location in the image. Nevertheless, augmentation can further aid in this transform invariant approach to learning and can aid the model in learning features that are also invariant to transforms such as left-to-right to top-to-bottom ordering, light levels in photographs, and more.
Image data augmentation is typically only applied to the training dataset, and not to the validation or test dataset. This is different from data preparation such as image resizing and pixel scaling; they must be performed consistently across all datasets that interact with the model.

Image Resizing
Images gathered from Internet will be of varying sizes. Due to presence of fully connected layers in most of the neural networks, the images being fed to network will be required of a fixed size (unless you are using Spatial Pyramid Pooling before passing to dense layers). Because of this, before the image augmentation happens, let us preprocess the images to the size which our network needs. With the fixed sized image, we get the benefits of processing them in batches.

Scaling
Having differently scaled object of interest in the images is the most important aspect of image diversity. When your network is in hands of real users, the object in the image can be tiny or large. Also, sometimes, object can cover the entire image and yet will not be present totally in image (i.e cropped at edges of object)

Translation
We would like our network to recognize the object present in any part of the image. Also, the object can be present partially in the corner or edges of the image. For this reason, we shift the object to various parts of the image. This may also result in addition of a background noise.

Rotation (at 90 degrees)
The network has to recognize the object present in any orientation. Assuming the image is square, rotating the image at 90 degrees will not add any background noise in the image.

Flipping
This scenario is more important for network to remove biasness of assuming certain features of the object is available in only a particular side.

Lighting Conditions
This is a very important type of diversity needed in the image dataset not only for the network to learn properly the object of interest but also to simulate the practical scenario of images being taken by the user. The lighting condition of the images are varied by adding Gaussian noise in the image.

Adding Salt and pepper noise
Salt and Pepper noise refers to addition of white and black dots in the image. Though this may seem unnecessary, it is important to remember that a general user who is taking image to feed into your network may not be a professional photographer. His camera can produce blurry images with lots of white and black dots. This augmentation aides the above mentioned users.

Perspective Transform
In perspective transform, we try to project image from a different point of view. For this, the position of object should be known in advance. Merely calculating perspective transform without knowing the position of the object can lead to degradation of the dataset. Hence, this type of augmentation has to be performed selectively. The greatest advantage with this augmentation is that it can emphasize on parts of object in image which the network needs to learn.

CONVOLUTIONAL LAYER

Diabetic Retinopathy canbe automatically detected by using a convolution neural network.It can provide more accurate detection results.it uses various architecture for finding a better accurate result.

3.2 Convolutional neural network

CNN stands for Convolutional Neural Network which is a specialized neural network for processing data that has an input shape like a 2D matrix like images. CNN's are typically used for image detection and classification. Images are 2D matrix of pixels on which we run CNN to either recognize the image or to classify the image. Identify if an image is of a human being, or car or just digits on an address. Like Neural Networks, CNN also draws motivation from brain . We use object recognition model proposed by Hubel and Wiesel.

3.3 Convolutional layer

Convolutional layers are the main building blocks of the CNNs. They learn the feature representation of the input images by performing convolutions over the inputs. The convolutional layer consists of several kernels that are used to compute different features from the input images. They ensure that the local connectivity neurons are connected to a small region of the input which is known as the receptive field. The extracted feature maps are calculated by convolving the input with the kernels and then add the bias parameters to the feature. The convolutional layer has many kernels, and they are applied to the input image to calculate the output feature map. Each kernel is shared by all special locations of the input. The advantage of weight sharing is to reduce the complexity of the model and the training process of the network becomes easier. Mathematically, consider x as the input image, W is the kernel, and b is the bias for the convolutional layer. The feature map z generated from this layer is calculated as:
z=Wx+b

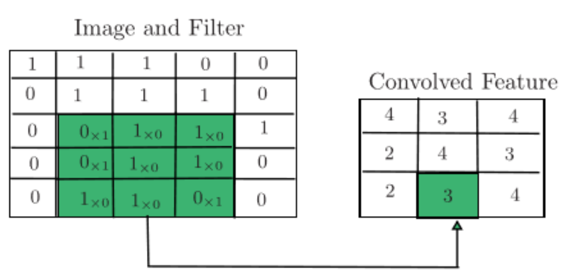


Figure 3.1: convolution operation

3.4 Activation layer

CNNs have some linear components and nonlinear components. The activation functions are the nonlinear components that follow the convolutional layers to introduce the non line arities to CNN to detect the non linear features and toi mprove the CNN performance. Rectified Linear Unit (ReLU) is one of the most popular activation functions used in CNNs. It has been shown that CNNs can be trained efficiently using ReLU [33]. ReLU is defined as: a=max (z, 0) Where z is the input to the activation function and a is the output. ReLU keeps the positive part of the input and prunes the negative part to zero. Another version of ReLU is leaky ReLU (LReLU) [34] that defines a parameter in range (0, 1) range (0, 1) to compress the negative part rather than mapping it to zero. This makes a small and non-zero gradient when the unit is not active (negative value). Exponential Linear Unit (ELU) is another activation function that enables faster learning of CNN and improves the accuracy of the classification task. Like ReLU and LReLU. Mathematically, ELU is defined as: . Where is a controlling parameter to saturate ELU for negative inputs. The last activation function we want total about is the sigmoidfunction[36]. The sigmoid function is used often in artificial neural networks to introduce the nonlinearity in the model. It takes real numbers and squashes them into range (0, 1). Mathematically, the sigmoid function is defined as:

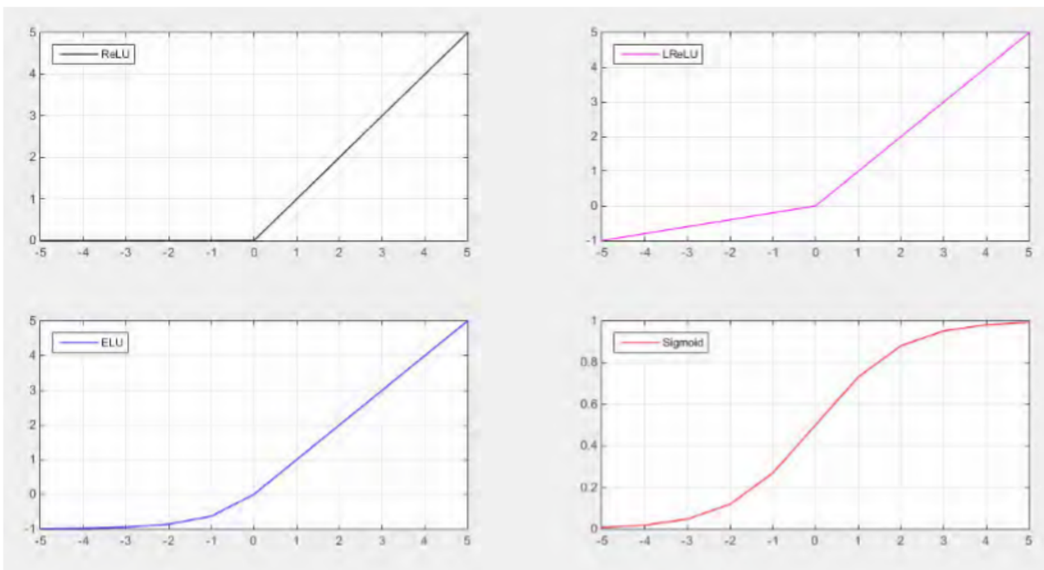


Figure 3.2: Different Activation Function

3.5 Pooling layer

The purpose of the pooling layer is to ensure the shift-invariance and lowers the computational burden by reducing the resolution of the feature maps. It is usually placed after the convolutional layer. It takes the feature map which is generated from the convolutional layer and outputs a single value for each receptive field (poolingwindow) according to the pooling operation. The pooling layer performs max-pooling [37], sum, and mean pooling [38]. Figure 3.3 shows different pooling operations. Also, there are other versions of pooling layers proposed for some tasks, such as Spatial Pyramid Pooling (SPP) [38] that can generate a fixed length of features regardless of the input size

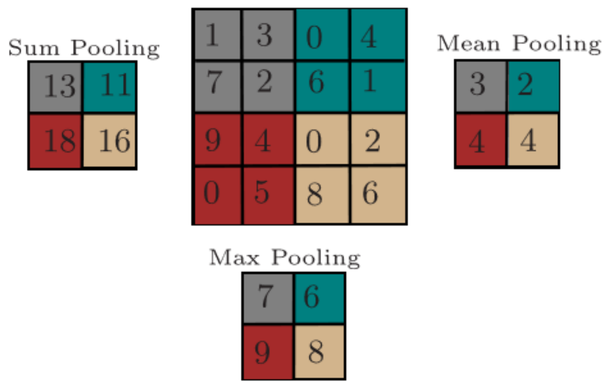


Figure 3.3: Different pooling layer

3.6 Fully conneted layer(FC)

(FC) in classification tasks, fully connected layers (FC) [12] are used at the end of the CNN after the convolutional layers and the pooling layers. Fully connected layers aim to generate specific semantic information. Then neurons in the fully connected layers have full connection to all neurons in the previous layer. It can be considered as a special case of a convolutional layer with the receptive field size is equal to one. Usually, dropout is used after the fully connected layers to avoid CNN from overfitting.

Difficulties: REGULARIZATION one of the most problematic issues regarding CNN training is overfitting. Overfitting happens when the model fits too well to the training dataset, and it cannot generalize to new examples that were not in the training dataset. So, overfitting is an unneglectable problem in deep CNNs.There are many proposed solutions to reduce overfitting effectively, such as Dropout [40], L1 regularization, and L2 regularization [39]. In deep learning, Dropout is widely used as regularization after the fully connected layers. It deletes or deactivates some neurons so that not all connections between the layers are activated at that time during training. It can also be applied after the convolutional layers. However,it is not preferable to add them in the first layers because dropout causes the information to get lost and if the information is lost in the first layers,it will be also lost for the whole network and this will affect the performance of the network. During

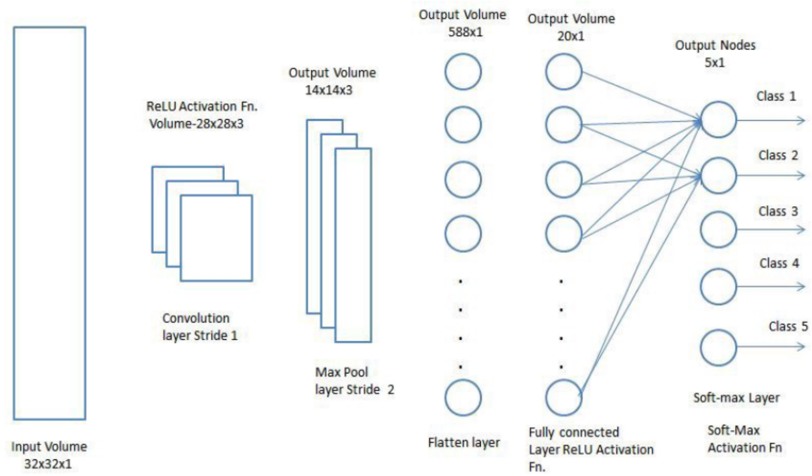


Figure 3.4: Fully connected layer

testing time, dropout layers are bypassed and they are not active.

3.7 Work flow of CNN

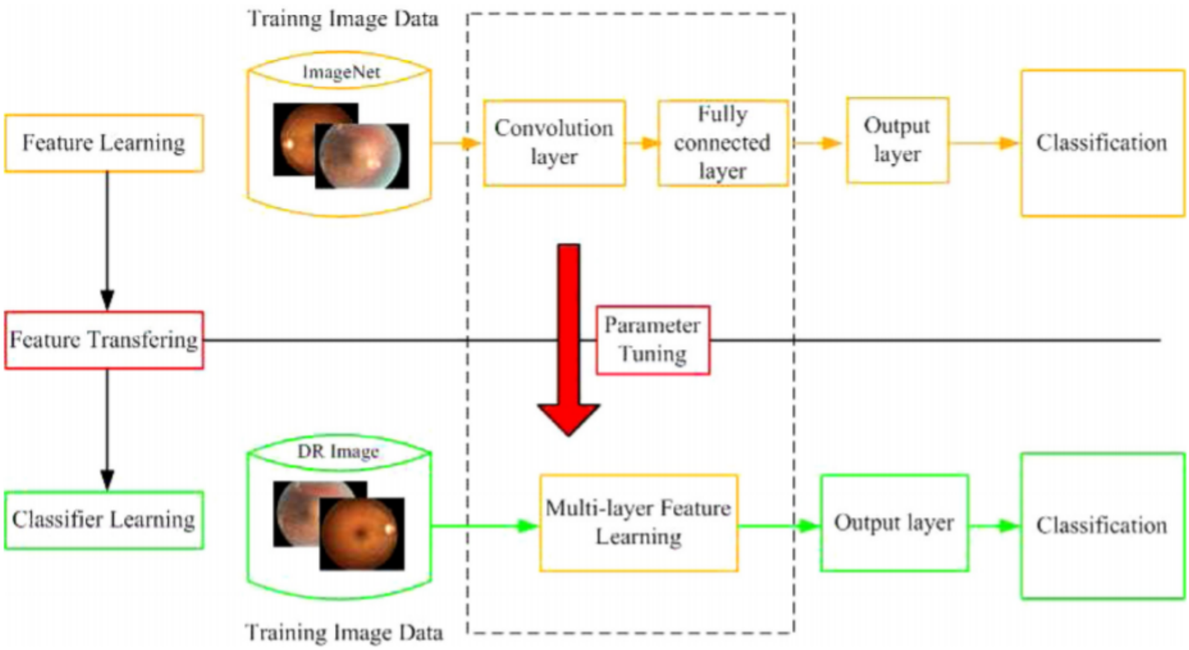


Figure 3.5: CNN flow diagram

Chapter 4

Software and Hardware specifications

Spyder

Spyder is a powerful scientific environment written in Python, for Python, and designed by and for scientists, engineers and data analysts. It offers a unique combination of the advanced editing, analysis, debugging, and profiling functionality of a comprehensive development tool with the data exploration, interactive execution, deep inspection, and beautiful visualization capabilities of a scientific package.
Beyond its many built-in features, its abilities can be extended even further via its plugin system and API. Furthermore, Spyder can also be used as a PyQt5 extension library, allowing developers to build upon its functionality and embed its components, such as the interactive console, in their own PyQt software Hardware

RAM : 8 GB
Processor : 64 bit Intel i5

Language-Python

Python is an interpreter, object-oriented, high-level programming language with dynamic semantics. It is high-level built-in data structures, combined with dynamic typing and dynamic binding, make it very attractive for Rapid Application Development, as well as for use as a scripting or glue language to connect existing components. Python's simple, easy to learn syntax emphasizes readability and therefore reduces the cost of program maintenance. Python supports modules and packages,which encourages program modularity and code reuse. The Python interpreter and the extensive standard library are available in source or binary form without charge for all major platforms and can be freely distributed. Python is often compared too their interpreted languages such as Java,JavaScript, Perl, Tcl, or Smalltalk. Comparisons to C++, Common Lisp and Scheme can also be enlightening. In this section, I will briefly compare Python to each of these languages. These comparisons concentrate on language issues only. In practice,the choice of a programming language is often dictated by other real world constraints such as cost, availability, training, and prior investment, or even emotional attachment. Since these aspects are highly variable, it seems a waste of time to consider them much for this comparison.

Important libraries installed are:

Tensorflow

Machine learning is a complex discipline. But implementing machine learning models is far less daunting and difficult than it used to be, thanks to machine learning frame works—such as Google's TensorFlow—that ease the process of acquiring data, training models, serving predictions, and refining future results. Created by the Google Brain team, TensorFlow is an open-source library for numerical computation and large-scale machine learning. TensorFlow bundles together a slew of machine learning and deep learning (aka neural networking)models and algorithms and makes them useful by way of a common metaphor. It uses Python to provide a convenient front-end API for building applications with the framework while executing those applications in high performance C++.

TensorFlow can train and run deep neural networks for hand written digit classification, image recognition, word embedding's, recurrent neural networks, sequence-to-sequence models for machine translation, natural language processing, and PDE (partial differential equation) based simulations. Best of all, TensorFlow supports production prediction at scale, with the same models used for training.

TensorFlow benefits

The single biggest benefit TensorFlow provides for machine learning development is abstraction. Instead of dealing with the nitty-gritty details of implementing algorithms, or figuring out proper ways to hitch the output of one function to the input of another, the developer can focus on the overall logic of the application. TensorFlow takes care of the details behind the scenes. TensorFlow offers additional conveniences for developers who need to debug and gain introspection into TensorFlow apps. The eager execution mode lets you evaluate and modify each graph operation separately and transparently, instead of constructing the entire graph as a single opaque object and evaluating it all at once. The Tensor Board visualization suite lets you inspect and profile the way graphs run by way of an interactive, web-based dashboard. TensorFlow also gains many advantages from the backing of an A-list commercial outfit in Google. Google has not only fuelled the rapid pace of development behind the project, but created many significant offerings around TensorFlow that make it easier to deploy and easier to use: the above-mentioned T P U silicon for accelerated performance in Google's cloud; an online hub for sharing models created with the framework;in-browserand mobile-friendly in carnations of the framework; and much more.

keras

Keras contains numerous implementations of commonly used neural-network building blocks such as layers, objectives, activation functions, optimizers, and a host of tools to make working with image and text data easier to simplify the coding necessary for writing Deep Neural Network code. In addition to standard neural networks, Keras has support for convolutional and recurrent neural networks. It supports other common utility layers like dropout, batch normalization, and pooling.

Chapter 5

Results

The extracted different features of retinal fundus image using CNN are analyzed and categorized that are presented with brief description along with its performance measures. To discriminate the retinal images as normal and abnormal, it has to be standardized using preprocessing techniques. The work comprises of detection of different features which results in high computational and time complexity.We found that there existed over-fitting phenomenon, in order to work out the over-fitting problem, we use transfer learning and hyperparameter-tuning methods to more accurately classify the fundus images.The better accuracy that we obtained using DDR is 90%

5.1 Performance Metrics

Accuracy
It is the ratio of number of correct predictions to the total number of input samples.

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

Specificity
Specificity is defined as the proportion of actual negatives, which got predicted as the negative (or true negative).

$$Specificity = \frac{TN}{FP+TN}$$

Sensitivity

Sensitivity is a measure of the proportion of actual positive cases that got predicted as positive (or true positive).

$$Sensitivity = \frac{TP}{TP + FN}$$

Chapter 6

Conclusion

Diabetic Retinopathy is one of the complications of diabetes and is an important blinding disease. Effective and automated diagnosis to the degree of Diabetic Retinopathy lesions have an important clinical significance. Early diagnosis allows for early treatment, which is crucial because early detection can effectively prevent visual impairment. Diabetic Retinopathy automatic classification of fundus images can effectively assist doctors in Diabetic Retinopathy diagnosis, which can improve the diagnostic efficiency. Convolutional Neural Networks for detecting Diabetic Retinopathy and transfer learning are presented to classify Diabetic Retinopathy fundus images and automatic feature learning reduces the process of extracting the feature of fundus images. The best experimental classification accuracy is 90% and our results yield better accuracy on Diabetic Retinopathy image classification.

References

[1] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Image Net classification with deep convolutional neural networks," in Proc. Conf. Adv. Neural Inf. Process. Syst., 2012, pp. 1097–1105.

[2] A. P. Bhatkar and G. U. Kharat, "Detection of diabetic retinopathy in retinal images using MLP classifier," in Proc. IEEE Int. Symp. Nano electron. Inf. Syst., Dec. 2015, pp. 331–335.

[3] A. Sopharak, B. Uyyanonvara, and S. Barman, "Fine Microaneurysms Detection from Non-dilated Diabetic Retinopathy Retinal Images Using a Hybrid Approach," vol. II, pp. 6–9, 2012.

[4] D.Mahapatra,P.K.Roy,S.Sedai,andR.Garnavi,"Retinal image quality classification using saliency maps and CNNs," in Proc. Int. Workshop Mach. Learn. Med. Imag., 2016, pp. 172–179

[5] D. Maji, A. Santara, P. Mitra, and D. Sheet, "Ensemble of deep convolutional neural networks for learning to detect retinal vessels in fundus images," 2016, pp. 1–4, arXiv:1603.04833. [Online]. Available: https://arxiv.org/abs/1603.04833 [67]

[6] D.H.HubelandT.N.Wiesel,"Receptive fields and functional architecture of monkey striate cortex," J. Physiol., vol. 195, no. 1, pp. 215–243, 1968

[7] Fousiya Shamsudeen"An objective function based technique for devignetting fundus imagery using MST" Elseiver(2019)82-91

[8] Fousiya Shamsudeen"Enhancement of Fundus Imagery" 2016 International Conference on Next Generation Intelligent Systems (ICNGIS).

[9] G. Quellec, K. Charrière, Y. Boudi, B. Cochener, and M. Lamard, "Deep image mining for diabetic retinopathy screening," Med. Image Anal., vol. 39, pp. 178–193, Jul. 2017.

[10] G. García, J. Gallardo, A. Mauricio, J. López, and C. Del Carpio, "Detection of diabetic retinopathy based on a convolutional neural network using retinal fundus images," in Proc. Int. Conf. Artif. Neural Netw. New York, NY, USA: Springer, 2017, pp. 635–642.

[11] H.Pratt, F.Coenen, D.M.Broadbent, S.P.Harding, andY.Zheng, "Convolutional neural networks for diabetic retinopathy," Procedia Comput. Sci., vol. 90, pp. 200–205, Jan. 2016. [23]

[12] H.Fu,Y.Xu,D.W.K.Wong,andJ.Liu,"Retinal vessel segmentation via deep learning network and fully-connected conditional random fields," in Proc.IEEE13thInt.Symp.Biomed.Imag.(ISBI),Sep.2016,pp.698–701.

[13] J. H. Tan, U. R. Acharya, S. V. Bhandary, K. C. Chua, and S. Sivaprasad, “Segmentation of optic disc, fovea and retinal vasculature using a single convolutional neural network,” J. Comput. Sci., vol. 20, pp. 70–79, May 2017

[14] J. Zhang, Y. Cui, W. Jiang, and L. Wang, “Blood vessel segmentation of retinal images based on neural network,” in Image and Graphics. ICIG (Lecture Notes in Computer Science), vol. 9218, Y. J. Zhang, Eds. Cham, Switzerland: Springer, 2015

[15] J. Zilly, J. M. Buhmann, and D. Mahapatra, “Glaucoma detection using entropy sampling and ensemble learning for automatic optic cup and disc segmentation,”Comput.Med.Imag.Graph.,vol.55,pp.28–41,Jan.2017.

[16] K. M. Adal, D. Sidibe, S. Ali, E. Chaum, T. P. Karnowski, and F. Meriaudeau, “Automated detection of microaneurysms using scaleadapted blob analysis and semi-supervised learning,” Comput. Methods Programs Biomed., vol. 114, no. 1, pp. 1–10, 2016

[17] M. R. K. Mookiah et al., “Evolutionary algorithm based classifier parameter tuning for automatic diabetic retinopathy grading: A hybrid feature extraction approach,” Knowl.-Based Syst., vol. 39, pp. 9–22, Feb. 2013. [Online]. Available: http://www.sciencedirect.com/ science/article/pii/S0950705112002651

[18] P.LiskowskiandK.Krawiec, “Segmenting retinal blood vessels with deep neuralnetworks,”IEEETrans.Med.Imag.,vol.35,no.11,pp.2369–2380, Nov. 2016.

[19] R. A. Welikala et al., “Automated detection of proliferative diabetic retinopathy using a modified line operator and dual classification,” Comput. Methods Programs Biomed., vol. 114, no. 3, pp. 247–261, 2014.

[20] S. Dutta, B. C. Manideep, S. M. Basha, R. D. Caytiles, and N. C. S. N. Iyengar, “Classification of diabetic retinopathy images by using deep learning models,” Int. J. Grid Distrib. Comput., vol. 11, no. 1, pp. 89–106, Jan. 2018.

[21] S. Wang, Y. Yin, G. Cao, B. Wei, Y. Zheng, and G. Yang, “Hierarchical retinalbloodvesselsegmentationbasedonfeatureandensemblelearning,” Neurocomputing, vol. 149, pp. 708–717, Feb. 2015.

[22] Segio B. Junior and D. Welfer, “Automatic Detection of Microaneurysms Ang Haemorrhages in Color Eye Fundus Images,” Southeast Asian J Trop Med Public Heal., vol. 34, no. 4, pp. 751–757, 2003.

[23] S. S. Rubini and D. A. Kunthavai, “Diabetic retinopathy detection based on eigenvalues of the hessian matrix,” Procedia Comput. Sci., vol. 47, pp. 311–318, May 2015

[24] S. S. Rahim, C. Jayne, V. Palade, and J. Shuttleworth, “Automatic detection of microaneurysms in colour fundus images for diabetic retinopathy screening,” Neural Comput. Appl., vol. 27, no. 5, pp. 1149– 1164, 2016.

[25] Shaohua wan, yan ling , “Deep convolutional neural network for diabetic retinopathy detection by image classification,” . Appl., vol. 27, no. 5, pp. 1149– 1164, 2018

[26] T. A. Soomro , J. Gao, T. M. Khan, A. F. M. Hani, M. A. U. Khan, and M. Paul, “Computerised approaches for the detection of diabetic retinopathy using retinal fundus images: A survey,” Pattern Anal. Appl., vol. 20, no. 4, pp. 927–961, 2017.

[27] T.ChandrakumarandR.Kathirvel,“Classifying diabetic retinopathy using deep learning architecture,” Int. J. Eng. Res. Technol., vol. 5, no. 6, pp. 19–24, Jun. 2016.

[28] Toufique ahamad somroo,” Deep Learning Models for Retinal Blood Vessels Segmentation” Received May 7, 2019, accepted May 24, 2019, date of publication June 3, 2019, date of current version June 12, 2019.

[29] W. R. Memon, B. Lal, and A. A. Sahto, “Diabetic retinopathy,” The Prof. Med. J., vol. 24, no. 2, pp. 234–238, 2017

[30] W. M. Gondal, J. M. Köhler, R. Grzeszick, G. A. Fink, and M. Hirsch, “Weakly-supervised localization of diabetic retinopathy lesions in retinal fundus images,” in Proc. IEEE Int. Conf. Image Process. (ICIP), Sep. 2017, pp. 2069–2073.

[31] Y. Yang, T. Li, W. Li, H. Wu, W. Fan, and W. Zhang, “Lesion detection and grading of diabetic retinopathy via two-stages deep convolutional neural networks,”inProc.Int.Conf.Med.ImageComput.Comput.-Assist. Intervent. Cham, Switzerland: Springer, 2017, pp. 533–540.

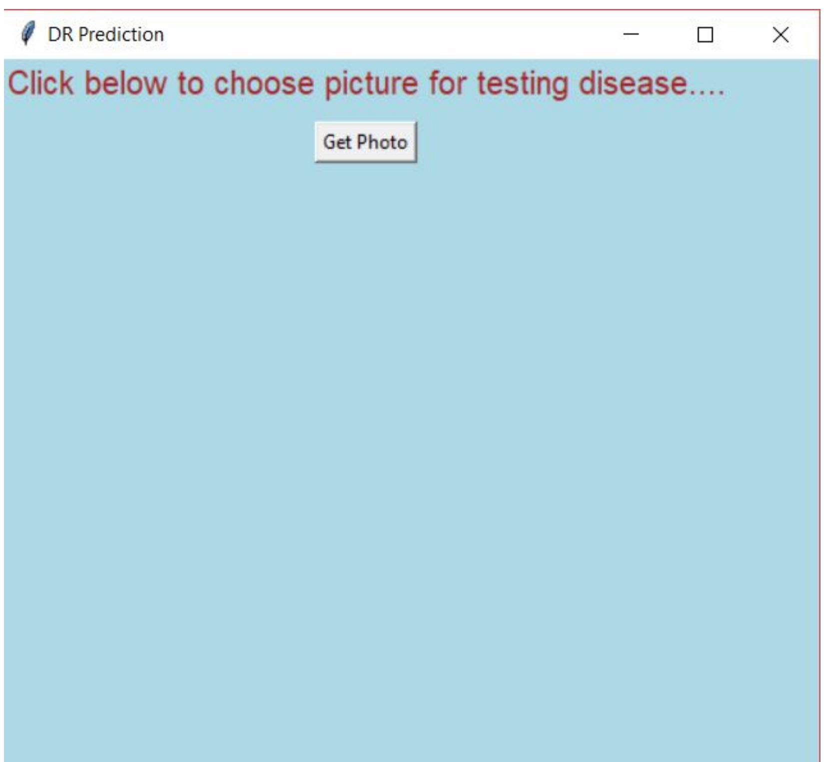
[32] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521, pp. 436–444, May 2015

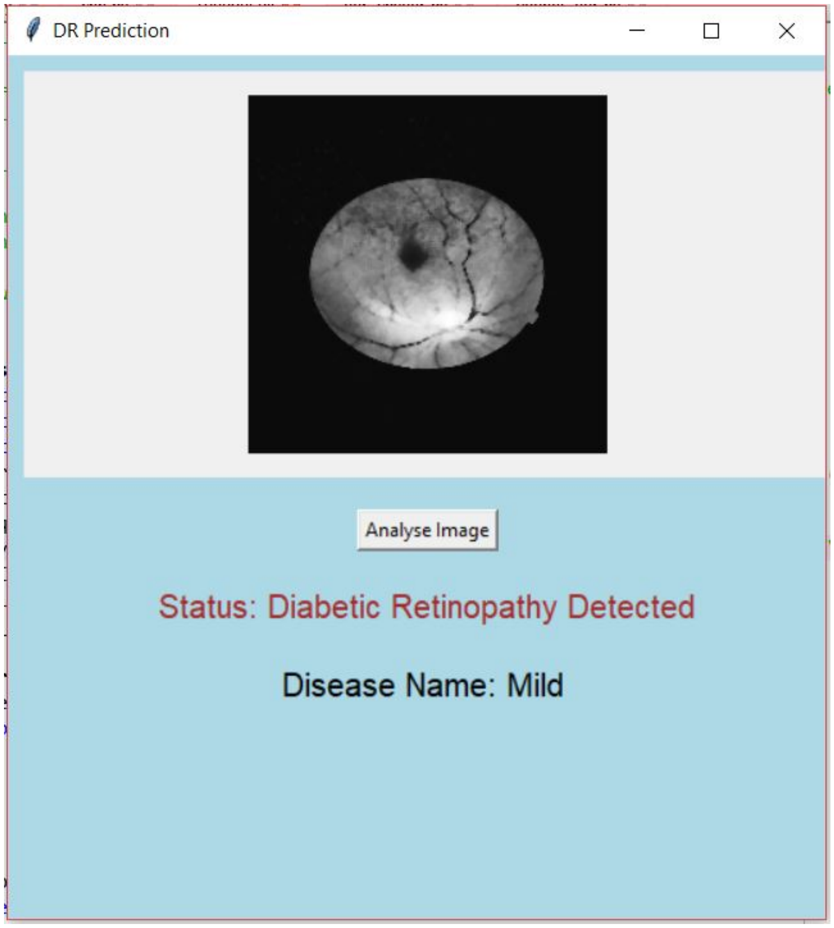
[33] Z. Yao, Z. Zhang, and L.-Q. Xu, “Convolutional neural network for retinal blood vessel segmentation,” in Proc. 9th Int. Symp. Comput. Intell. Design (ISCID), vol. 1, Dec. 2016, pp. 406–409.

[34] Xiaoling wang,Yongjin Lu “Diabetic retinopathy stage classification using cnn”, In ieee conference on information reuse and integration for data science 2018

Chapter 7

Appendices




DR Prediction
Analyse Image
Status: Diabetic Retinopathy Detected
Disease Name: Mild