Learning based classification of OCT Images with Diabetic Macular Edema and Dry-Age Related Macular Degeneration with Recurrent Neural Networks

Author

Albab Ahmad Khan

0000117865

Fall 2015-MS-85(CE)

Supervisor

Dr. Muhammad Usman Akram

DEPARTMENT OF COMPUTER ENGINEERING

COLLEGE OF ELECTRICAL & MECHANICAL ENGINEERING

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY

ISLAMABAD, PAKISTAN

August 20, 2019

Learning based classification of OCT Images with Diabetic Macular Edema and Dry-Age Related Macular Degeneration with Recurrent Neural Networks

Author

Albab Ahmad Khan

0000117865

A thesis submitted in partial fulfilment of the requirements for the degree of

MS Computer Engineering

Thesis Supervisor:

Dr. Muhammad Usman Akram

Thesis Supervisor's Signature: ____

DEPARTMENT OF COMPUTER ENGINEERING

COLLEGE OF ELECTRICAL & MECHANICAL ENGINEERING

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY,

ISLAMABAD, PAKISTAN

August 20, 2019

Copyright Statement

Declaration

I hereby declare that I have written this thesis titled as "*Transfer learning based classification of Age Related Macular Degeneration(AMD) and Diabetic Macular Edema(DME) from OCT Scans with Recurrent Convolutional Neural Networks(RCNN's)*" completely on the basis of my personal efforts under the sincere guidance of my supervisor Dr. Usman Akram. All citations with references to all sources used in this thesis have been mentioned clearly and contents of this thesis have not been plagiarized. I certify that this work contains no material which has been accepted for the award of any degree, or in any university or any previously published material except where due references have been made in the text.

Signature of Student

Albab Ahmad Khan

0000117865

Language Correctness Certificate

This thesis has been read by an English expert and is free of typing, syntax, semantic, grammatical and spelling mistakes. Thesis is also according to the format given by the university.

Signature of Student

Albab Ahmad Khan

0000117865

Signature of Supervisor

Dr. M Usman Akram

Acknowledgements

I am eternally grateful to Allah Almighty; The Creator of All Universes and everything that has ever existed. I would never have been able to complete this insurmountable task without His guidance. I could really feel the divine help all along my thesis. Indeed, it sounded an uphill task, but Allah was always there whenever I found myself in any blind alley.

I want to extend my deepest gratitude to my parents. They provided me the moral support and always encouraged me whenever I got stuck at any point.

I want to thank my supervisor Dr. Muhammad Usman Akram. Their immense guidance played an instrumental role in shaping up my dissertation work. He taught me how to carry out research actively and effectively. Veritably, he is the architect of this work and I want to give him maximum credit for that. I used to bother him over and over again, usually numerous times in a matter of hour or so, but he always obliged. You have my utmost respect, sir.

I am indebted to BIOMISA LAB for providing me equipment for my experimental work. I'm forever in the debt of my university for covering the expenses incurred throughout.

Lastly, I want to pay my earnest gratitude to my friends who always boosted my morale whenever I was depressed. They are the true heroes and equally deserve all the recognition and praise. There are certain unsung heroes as well and I can't name them all but I do like to take this opportunity to send my warmest regards to whosoever played his part in making me what I'm today. I owe it to all of them.

ABSTRACT

The recent emergence of machine learning and deep learning methods for medical image analysis has enabled the development automated computer aided systems that can assist ophthalmologists in making better decisions about a patient's maculopathic diseases. In particular, Maculopathy is the field where a person can have several diseases such as Age related Macular Degeneration (AMD) and Diabetic Macular Edema (DME). Both of the diseases can lead to permanent loss of vision. Early detection of these diseases can reduce the level of disease severity and this early detection of disease can facilitate the patient with better and right diagnosis as well as proper medication. Several researches have already been conducted in order to identify and predict AMD and DME from the images of the patients by using dictionary learning based classification, SMO, Support Vector Machines (SVM), Principal Component Analysis (PCA), Convolutional Neural Networks (CNN) and Wavelet Convolutional Neural Networks. All of the research conducted by far was on scan level where the individual scan was classified as AMD, DME or Normal.

Therefore it was a hassle for the ophthalmologist in order to observe all scans and then deciding the category of disease, a patient is suffering. In this research we have proposed a novel approach where CNN and RNN (LSTM) are combined to detect the retinal disease on patient level. CNNs are used for feature extraction where LSTM is used for the classification of the image into 3 classes mainly AMD, DME and Normal Patients. After training R-CNN, we tested the model on scan as well as on patient level. On scans level, the model is giving 99.69% accuracy while it was tested on 647 images of Duke's dataset. The model is tested on patient level where every scan of the patient is categorized as AMD, DME Normal and then on the basis of most occurrences, the model decides to which class a particular patient belongs. Model has been tested for 6 patients. Out of 6 patients, 5 patients have been correctly classified and 1 patient has been wrongly classified. Thus we achieve 83% accuracy of the model on patient level.

Table of Contents

LIST OF FIGURES

LIST OF ABBREVIATIONS

AMD	Age related macular degeneration
DME	Diabetic Macular Edema
RNN	Recurrent Neural Networks
SP	Specificity
SN	Sensitivity
SVM	Support Vector Machines
CNN	Convolutional Neural Networks
SMO	Sequential Minimal Optimization
RCNN	Recurrent Convolutional Neural Networks
LSTM	Long Short-Term Memory
OCT	Optical Coherence Tomography

Chapter 1: Introduction

Diabetes becomes one of the major causes of inability to see among people and is termed as blindness. As per the statistics of World Health Organization (WHO), the numbers of diabetic patients across the globe in 1980 were approximately 108 million whereas in the latest survey of 2014, the number of diabetic patients has reached up to 422 million [1]. As per the statistics of WHO, the global percentage increase of diabetic patients among the adults over the age of 18 has reached from 4.7% (1980) to 8.5% (2014). Diabetes is drastically increasing in low income and middle-income countries as compared to the rest. Diabetes became a major reason of 1.6 million deaths in 2016 and another 2.2 million people died because of high blood glucose. Moreover, as per WHO estimations, it is estimated that by 2030, diabetes will become the seventh largest cause of human death which is quite early alarming and requires serious attention of health community [1]. Diabetes is one of the prime reasons for the complications of retina which results in the loss of human vision. The inner most layer of the eye ball is called retina and human vision is formed because of this layer. There are two main regions of retina i.e. peripheral region and macular region. The later region i.e. macular region is also known as Macula and is of prime importance because of the formation of central vision whereas the peripheral region of the retina is responsible for the formation of side vision. Macular region or Macula sometimes, is also commonly referred as retina. The formation of central vision and side vision is shown in the Figure 1.1. The image shows a vision of a normal person not suffering from any kind of irregularity or disease.

Figure 1.1: Vision of a normal person. (A): Central vision of normal person, (B): Side vision of a normal person

There are several visual impairments that occur within the macular region termed as maculopathy. Some of its commonly known forms are Diabetic Macular Edema (DME), macular degeneration also known as Age-related Macular Edema and Central Serous Retinopathy. We however will be considering AMD and DME in detail. Diabetic Macular Edema can be found in huge numbers among diabetic people. Its early discovery is vital in order to prevent loss of vision. DME is already affecting 6% of diabetics which has resulted in more than 20 million case around the world [4]. AMD is one of the causes of central vision loss and going by its name i.e. age-related macular edema, it can be found in people with age over 50 years. According to a research, it is found out that United Kingdom has been a victim of AMD with more than 250000 adults suffering from it [4].

Maculopathy symptoms do not appear early therefore, it is very crucial to discover the disease in early stages. Identifying the disease earlier is vital for in time treatment. Some techniques are being utilized for treatment identification and the common among them are Fundus Fluorescein (FFA) Fundus Photography and Optical Coherence Tomography (OCT). FFA is sort of a severe technique as it involves injection of sodium fluorescein dye injected into human body. This results in the highlighting of blood vessels and specifically the retinal abnormalities in the body. This technique is sometimes termed as invasive because injecting a dye might be harmful and can trigger allergies in various people. Fundus photography however is quite non-invasive. It shows the fundus of the retina which is also known as the top of the retina that displays all the macular abnormalities that occur from the choroidal vascular structures. Third and the most commonly used technique is called OCT. OCT helps in the discovering early symptoms of maculopathy by capturing cross-sectional retina using Michelsons' interferometer [2]. OCT can be used for visualizing oculopathy in detail.

1.1 Diagnosis of the Disease

One of the major reason's maculopathy is not diagnosed in its early stages is because it is often painless. Even though it is difficult to diagnose in its early stages, it still can be figured out if regular clinical examination is performed to check the changes occurring within the macula. This can help in prevention of maculopathy. Ophthalmologist diagnose maculopathy by dilating the pupil to clearly examine the fundus of the retina. There is often certain macular swelling within macula which is the indication of Macular Edema. Similarly, if drusen and scars are observed in any aged subject then it is a clear indication of AMD. Clinicians around the world normally rely on OCT but some of them have also a strong recommendation for AMSLER Grid examination which helps in measuring the central vision loss. The examination works for both the eyes where the subjects are asked to mark abnormalities on the charts they see. If the visions appear to be distorted and wavy then it's a clear indication of maculopathy, severity of which is dependent upon the distortion of their vision. Looking at Figure 1.2 we can see the Amsler grid view of a healthy person.

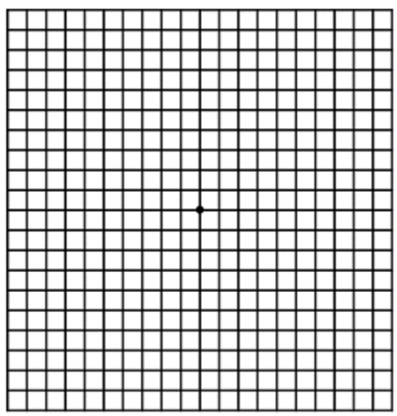


Figure 1.2: Amsler grid view of a healthy person

1.2 Motivation

Eye sight is one of the most important sense in human body but it is being majorly comprised and lost due to retinal disorders associated with diabetes. According to statistics published by World Health Organization, 80% of the blindness is avoidable if it is treated in time. But there are certain ophthalmologists who are unaware and they lack knowledge of the treatment and diagnosis which has caused a raise in the disease. Moving towards the rural areas, treatment costs in great amount because it involves travelling and treatment also requires a lot of money. Since the disease is always diagnosed at the advanced stages so its treatment also costs accordingly. Treatment is not very beneficial at that stage causing vision loss which doesn't seems to bring back the sight of the subject. This raised need of a computer aided diagnostic solution which can diagnose AMD and DME at an early stage thus bringing a relief among the people, especially in the rural areas. It is always crucial to diagnose maculopathy since it is one of the major causes of blindness among subjects. 80% of blindness cases are due to maculopathy and the chances are greater when patients are suffering from diabetes [3]. OCT imaging provides us with a lot of details related to macula and maculopathy but doctors and diagnosticians still face problems identifying the disease in rural areas. This could be due to lack of experts and enough efficient systems which might help in diagnosis. A retinal computer aided system would be of great help in this regard. R-CAD not only helps doctor in screening the retina but also provides rapid solutions. These systems can be installed in rural areas to enhance basic health facilities among the inhabitants. Following the clinical health care standards, these systems can help in diagnosing maculopathy in early stages and also helps to tract retina to reconstruct the effected portion. All of the novel systems and procedures up till now are providing solutions at the scanned level. We propose to provide the diagnoses at patient level which will be more effective for relative subjects.

1.3 Problem Statement

The conventional work in the detection and classification of Age-Related Macular Degeneration (AMD), AMD, Diabetic Macular Edema (DME) and Normal patients from OCT Scans is done with high precision at the scan level till now. But the classification is required at patient level. Therefore, a novel system was required where the classification of the disease should have been done at the patient level by considering all the scans of a patient to make the ease of process and rapid result regarding the conclusion of disease various scans.

1.4 Scope and Objectives

Eye diseases are very critical in diagnosis. Vision loss is majorly caused by DME i.e. Diabetic Macular Edema which is most certainly irreversible in individuals having diabetes. United States health and care has spent $500 million in prevention of eye diseases so far because the cases of vision loss were expected to increase affecting 300 million people till the year 2025 [5]. Similarly, AMD i.e. Age-Related Macular Edema if not treated in time can cause vision loss that cannot be cured. Both of these fore mentioned diseases were somewhat diagnosed through OCT images but the systems could only scan them at the very basic level. The scope of thesis revolves around classification of the OCT images of each patient to find rapid solutions and aid the doctors in diagnosis of diseases at early levels. The classification will be done through Recurrent Convolution Neural Networks. The thesis proposes to meet the following objectives through the research conducted.

- To develop a classification system that could identify retinal information through OCT images.
- To identify the disease by scanning images of AMD, DME affected and normal patients.
- The system will be able to classify images at patient level using RCNNs.
- To help doctors and experts in identification of maculopathy at early stage in a much speedy way than previously used CADs such that it is cost effective and helps in giving rapid results.

1.5 Thesis Organization

The rest of the thesis is organnized according to the following manner.

Chapter 2 gives a detailed review of already presented methods and algorithms for classification of maculopathy. It summarizes how OCT images were classified and which systems were used to extract information in diagnosis of AMD and DME.

In **Chapter 3**, the proposed system is discussed in detail. (To be added)

Chapter 4 includes all the experimental results that are discussed in detail with all desired figures and tables. The comparisons of our methods with other state-of-the-art techniques are also given in this chapter.

Chapter 5 concludes the thesis and reveals future work of proposed system.

Chapter 2 : Literature Review

This chapter gives overview of the papers that provided aid in conducting this research. A summarized table is also shown in the end describing how each method is used and what results they produce.

2.1. Overview

Maculopathy is a very critical disease which is being worked upon. The criticality of its diagnosis lies in the early discovery of the diesease because its treatment if done late can subsequently result in vision loss. OCT images are very helpful in classification of AMD and DME, also some recent researches show the work done in classification of OCT images for diagnosis. Following are some of the researches conducted by certain researchers and also considered for this thesis.

E. Mousavi, R. Kafieh and H. Rabbani proposed a model [4] which performs classification of AMD and DME through OCT using a dictionary learning method. The writer's basic objective was early diagnosis of AMD and DME for which this classification model was used. The dataset on the model was working consisted of 45 subjects among which there were 15 patients with AMD, 15 with DME and the rest 15 were normal patients. For working on the model, authors do not consider the retina layer segmentation because it could have raise the risk factor for classification of abnormal images. The purpose of writer was to extract HOG features of AMD, DME and Normal OCT images. The study consists of learning discriminative atoms for classification of these images using 3 different kind of dictionary methods. The 3 methods were FDDL, LRSD and COPAR. Among these methods, FDDL was found to give results with highest accuracy i.e. 98.3% where the other methods could give accuracy upto 97.3% which isn't a bad result either and easily comparable withother models. As mentioned before, early diagnosis was also an important point of consideration for the auhors therefore, the reaserch paper illustrated how abnormal images of patients were diagnosed through abnormal detection. This detection was done on only 4% of all B-scans of a volume thus fulfilling the requirement of early diagnosis.

In this study of Yu Wang and his fellow authors, the focus is on AMD and DME along with its diagnosis. The SD-OCT images were used to get a scanned image of retina. The proposed system has outperformed certain models which were already working on this disease. This [5] study was based on creating a CAD model havig Local Configuration Patterns(LCP) which worked in favor of the classification problem under consideration. Correlation -based Feature Subjet (CFS) was used for screening features of OCT images based on LCP. Two algorithms were used by the author in this paper, SMO and LR. SMO was the one algorithm which could perform the best with an accuracy of up to 99.3 % up till now. In the end they propose and automated macular diease detection system which provides aid to different eye experts and diagnosticians of this field for ealry detection which is the most crucial. The model is also helpful in conducting different medical activities and applications.

Another paper was written by Ruaa Adeeb Abdulmunem Al-falluji [6] which focused on the system that detects only Diabetic Macular Edema(DME) using OCT images. The system works in different phases and the first of which is preprocessing phase that consists of firstly noise removal in the retinal images. The second part of this phase is flatenning of images which can be done by Local binary feature extraction. In the end , final features are transferred to SVM for classification. This system gave sensitivity of 100% and specifity reached up to 86.67%. The accuracy however was reached up to 93.33%.

A research article was found written by Joan Massich [7] which focuses on classification of SD-OCT images through binary local patterns. An automatic classification of DME images is done and following tasks were covered in this study.

Effect on each pre-processing step was studied

How different feature extracting strategies influence the results

The impact of codebook size

In the end there was a comparison among different classification strategies

The main focus of this study was on the fore-mentioned points. Also adding to the list were the results for the potential of 3D features leading to a high-level representation of 2D features which were obtained from the local patches. The study showed how SVM was worked upon in combination with BoW approach and similarly how RF classifier was used with global mapping. Joan Massich also emphasized on pre-processing steps, the performance of which was fluctuating in any case later but if flattening of images and alignment is used it might increase the performance every time. The research article shows that flattening is not as efficient in case of RPE. After flattening however, the author has mentioned the use of LBP invariant for reduction of patterns that were encoded. Once the task is done the non-rotational LBP is further studied further. The framework which was developed in this article was later tested on 32 patients. The acquired results for the tested framework were: Sensitivity was captured at 81.2% and specificity was observed at 93.7%. The study was concluded by stating that 3D features and 2D higher representation achieved the best results.

Another study [8] showed a fully automated algorithm for detection of DME and Dry AMD. The research proposes and algorithm that work on OCT images to give the desired results. The algorithm used multiscale histograms which worked as feature vectors in an SVM based classifier. The OCT data consisted of 45 subjects the division of which was: 15 normal subjects, 15 patients with AMD and 15 patients with DME. This data set was actually used for cross-validation further. This classifier proposed by Leo. A. Kim and his fellow writers was working correctly with identifying dry age-related macular edema patients. 100 % patients were correctly identified for AMD in this case. Now looking at Diabetic Macular Edema, the author got 100% identification results. There was a slight variation in case of normal subjects. The result identification rate for normal subjects was 86.67%. According to writers' claim the algorithm can be an effective and helpful tool for diagnosis of diseases related to ophthalmology.

The author S. P. K. Karri and his fellow authors worked on an algorithm in this study [9] that could identify retinal pathologies using OCT images. The study mentions improvement of CNN based GoogLeNet which helps in making the predictions better as compared to random initialization training. Filtered characteristics become easy to understand using by identification of certain features. The study has worked upon the sets involving data of AMD and DME. Just like other researches, this one also considered data for normal subjects as well. Since the main goal of this study was to identify pathology so a fine-tuned CNN was worked upon to give enhanced results than previously used classical learning. So, the algorithm demonstrates that there are different models if trained on non- medical images can result in fine-tuned algorithms working in a better for classification OCT images. The study involves dataset working on 45 patients which are divided into sets of 15 for each type of subjects i.e. 15 for AMD, 15 for DME and 15 normal subjects. After a number of experiments performed upon data, the study could only consider the model with accuracy up to 94%. The rest of the evaluations were performed on that model.

Zhongyang Sun and Yankuj Sun worked on convolutional networks and proposed a scheme [12] which used FCN i.e. fully convolutional networks for diagnosis of maculopathy among patients. Various OCT images were gathered from the sampled data set. Dataset was gathered from some clinical records and from Duke's university. So, the authors presented an FCN model which was trained on 900 labelled OCT images. These images consisted of AMD samples, DME samples and samples from normal subjects. Two of the classifiers were used for validation and classification of images and the effectiveness of disease recognition. ScSPM and V3 were the classifiers which worked on Clinical Data, selecting randomly 300 AMD, 300 DME and 300 normal subjects' images for training. The remaining data set was utilized for testing purposes. ScSPM provided them with accuracy of 98.69% on clinical data and V3 achieved an accuracy of 99.69 in case of clinical data. Dukes' data set comprising of 45 samples in total, 15 of them of AMD patients, 15 of DME and the remaining 15 were of the normal subjects. The classifiers provided 100% accuracy for this dataset. The basic advantage of FCN proposed by the authors was that it didn't require any sort of flattening or cropping to enhance the classification results.

This study [13] is based on the working of an automatic algorithm which served to scan OCT images of the patients suffering macular edema problem. This methodology is not supported by any other image denoising or cropping techniques for assessment of various retinal layers and to abnormalities of them. The classification of these images was based on a two staged schema which provided sub-systems comprising of adaptive feature learning and also worked upon diagnosis and its score. The proposed model worked in 2 stages. The first of the stages was based on WCNN i.e. Wavelet-based Convolutional Neural Networks. WCNN helped in generation of CNN codes which helped in feature extraction of 3-D volumes. Moving on to the second stage, the scoring was gained for abnormalities that were found in the existing OCT images. Authors in this research used 2 different types of datasets for validation and training purposes. The first dataset was from Topcon device consisting of 30 images of AMD, 30 images of DME and 30 images of normal people.

Second dataset was from Dukes' university again, consisting of 45 images which were equally divided among AMD, DME and normal subject categories. So, after applying cross validation of 10 iterations the two-class classification problem applied on the first dataset provided with the precision of 99.50%. Whereas for three-class classification problem, dataset of Dukes' university was used which provided precision of 98.83% after cross validation of 10 iterations.

2.2. Summary

The following table shows the summary of all the results for methods and processes used in the papers described earlier.

Table 2.1: Summary of previous methods implemented and accuracies achieved on same dataset

Sr. No	Authors	Method	Type of Maculopathy Disease examined	Dataset	Year	Results
1	E. Mousavi,	Dictionary Learning based Classifiers • COPAL • FDDL • LRSD Final Method: FDDL	AMD DME	Duke University:45 • AMD: 15 • DME:15 • Normal:15	2019	Accuracy: 98.3%
2	Yu Wang,	SMO	AMD DME	Duke University:45 • AMD: 15 • DME:15 • Normal:15	2016	Accuracy: 98 %

3	Ruaa Adeeb Abdulmunem Al-falluji	SVM	DME	Duke University:30 • DME: 15 • Normal:15	2016	Accuracy: 93.33%
4	Guillaume Lemaître	SVM	DME	SERI :32 • 16 DME • 16 Normal	2016	Sp: 93.7%
5	Pratul P. Srinivasan,	SVM	AMD DME	Duke University:45	2014	Cross Val: AMD: 100% DME: 100% Normal: 86.67%
6	S. P. K. KARRI	CNN(GoogLENet)	AMD DME	Duke University:45	2017	Accuracy: 94%
7	Khaled Alsaih	PCA+SVM	DME	SERI :32 • 16 DME • 16 Normal	2017	SP&SE: 87.5%
8	Zhongyang Sun	FCN(ScSPM,V3)	AMD, DME	Clinical Data Dukes Data	2019	Accuracy: Dukes: 100% Clinical:99.69
9	Reza Rasti	WCNN	AMD, DME	Dukes Data Topcon device	2017	Precision: 2-class classification:99.50% 3-class classification:98.83%

Chapter 3. Methodology

This chapter includes the details of the method used in this research and the modeling applied to get the desired results upon the OCT images. A detailed analysis of the methodology and modeling will be explained further.

3.1. Maculopathy Background

Maculopathy is basically a disease which is caused in the central part of retina and causes vision loss of not treated on time. Therefore, the diagnosis of this disease is very crucial at the earlier stages as it can be treated at initial stages or the steps taken initially can lead to precautionary methods. A variety of methods and systems have been proposed for early diagnosis of this disease, the summary of whom have been provided in Chapter 2. This study however proposes a model comprised of CNN (Convolutional Neural Networks) and LSTM classifier for image classification leading to early diagnosis.

The disease can be found in different types. Two of the most common types we will be considering for this study are AMD and DME. AMD or Age-related Macular Edema is caused after you reach a certain age but this doesn't cause permanent blindness. A serious and another form of maculopathy is occurred when a person is suffering from diabetes. Diabetes causes abnormalities in the retina of the person leading him/her towards maculopathy which can ultimately result in blindness and it is termed as DME or Diabetic Macular Edema. The main cause of this blindness is the late discovery of the disease which leads to reduced chances of prevention. This study proposes to solve the problem through a combined work of CNN and LSTM classifier which provides with the solution of the above-mentioned problem.

[Adding Images]

3.2. Methodology Overview

Up till now studies and researches have been working on a variety of classification problems. The authors worked more on CNN being a separate algorithm. None of them have tried to combine it with another deep learning algorithm to get enhanced results. We have proposed a methodology in which we have brought together CNN and LSTM for image classification. CNN serves to be feature extractor which fill in LSTM to get the images classified as desired. CNN is best suited for feature extracting and LSTM is well suited for producing efficient results on large sequential problems. Since our dataset consists of images obtained from videos which is again a product of sequential images so we have tried to get the best of the combination of the above mentioned. The methodology involves working of different convolutional layers on the data set and then application of LSTM for classification.

Following diagram shows a more comprehensive overview of methodology proposed for this research.

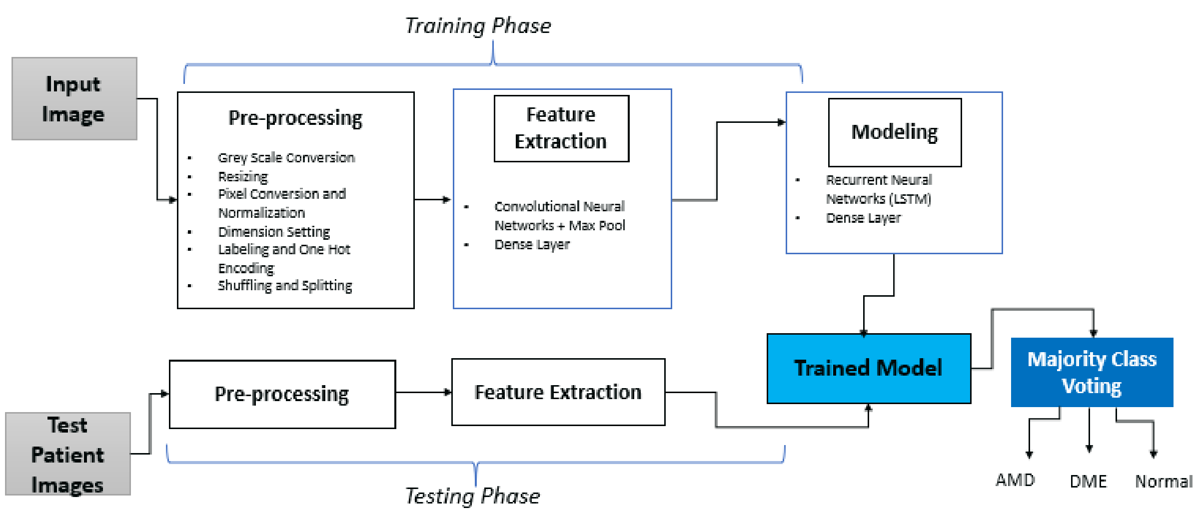


Figure 3.4: System overview

3.3. Data Set

The data set for this research was obtained from Duke's University which has already been referred by a lot of other authors in various other researches. It comprises of total 45 sampled OCT images the details of which are as following:

i. 15 AMD images
ii. 15 DME images
iii. 15 Normal images

3.4. Preprocessing

Preprocessing in our methodology was required and consisted of the following steps.

i. Grey Scaling
ii. Resizing

3.4.1. Grey Scale

The OCT images initially used for this study were in RGB form. A conversion into grey-scale was required for further processing and improved classification results. Open CV is the fundamental library which provides us with the variety of functions to support machine learning and computer vision. In this study it helped to convert the images from RGB format into grey scale having channel equal to1.

3.4.2. Resizing

Data set from Duke's University comprised of images was mostly in the found with the dimensions 512*496. It was required to get a size of 128x128 for an individual image which was the minimum size of the dataset in order to get the results of more specific retinal area to be observed for maculopathy disease.

3.4.3. Pixel Conversion and Normalization

The data extracted from images in normally in pixels. Further processing required data to be in decimal form and thus pixels were converted into floating values. These values then required scaling as the scale ranged from 0 to 255. Therefore, each value was normalized to get the values with in the range of 0 to 1 to get a standard data. Normalization in data is done to bring the data on the same scale so the results can be as accurate as possible. As there was data with different scales so normalization was done through: $\frac{Xi}{255}$.

3.4.4. Dimension Setting

After scaling the dataset, every image was tested to have the channel equivalent to 1. The channel gives the information of the image whether it is coloured or grey scale. In this step, there's a check whether the channel of the input image is 1 or not. 1 states if the channel is grey scale. If the channel is 1, a further one more dimension is added in order to feed it into the model.

3.4.5. Labeling and One Hot Encoding

One hot encoding is done to get the categorical data into a numerical form such that calculations and processing can be applied on it. Our study also required one hot encoding because images were labelled according to the type of patients they belonged to. After labelling they are encoded using one-hot encoding technique.

3.4.6. Shuffling and Splitting

After one-hot encoding data was shuffled. This shuffling involved images and labels which were one hot encoded. Shuffling was done to remove biasness from data. Once the images and labels were shuffled, the data was later split. Splitting was done to train the model on 80% data and rest 20% of data was used for validation purpose.

3.5. Modelling

Technology is improving every day and its evolution is helping mankind in problem solving. Not only it aids in finding solutions to specific problems but it also provides a variety of ways. Artificial intelligence and machine learning are one the ways which serve with data manipulation and analysis.

As this study involves working with image processing therefore, one of the most reliable algorithms is CNN. We have used a mixture of CNN and LSTM as Convolutional Neural Networks will be helping in feature extraction and LSTM will be used as a classifier for classification of images of maculopathy. Details of each of them are discussed below.

3.5.1. CNN in Detail

CNN (Convolutional Neural Networks) were introduced in 1990 but they were not a center of attention back then. Researchers didn't pay much heed because they had high requirements in terms of computational machines and required data set which were comparatively supposed to be huge. Arrival of GPUs in tech world made people focus towards CNN as it is state of the art feature extractor and classifiers. CNN is along with layers is described below and also illustrated in the figure.

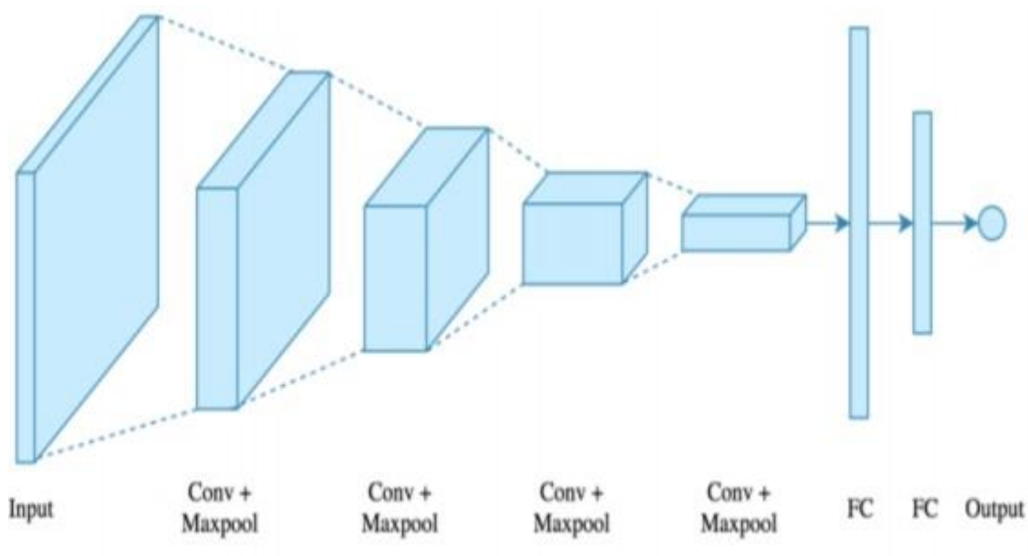


Figure 3.5: Architecture of CNN layers

i. Convolutional Layer:

This layer plays a vital role in convolutional neural networks because it works as a feature extractor for an image. These features represent various orientations or curves of certain edges at the initial layers. Going into the deeper layers, high level domain specific features are extracted by applying convocation to the output generated by the previous layers. Each layer may have varying number of sizes and filters. This size is defined by three hyper parameters that are

- Stride: After every convolution, there occurs a jump horizontally and vertically which is referred to as stride.
- Depth: Depth is the total number of filters in layer.
- Padding: Padding is the borders in the form of rows and columns that are put around the image given as input to complete convolution.

ii. Pooling Layer:

This layer is a down sampling layer which helps reduction of parameters and overfitting. Some common techniques in layer are max pooling, average pooling and L2-norm, where we will be considering max pool for this study. It revolves around the idea of considering a small neighborhood around for example 2x2 and then it replaces the neighborhood with a single value that could be probably a maximum value.

iii. Fully Connected Layer:

Once the features are gathered after convolutional layer and pooling layer, they are sent forward to the fully connected layer which serves the purpose of classification. This layer consists of neurons which are fired when they are presented with the matching class.

3.5.2. LSTM Networks

RNN were normally used for working on problems related to sequences but it didn't give the expected results to the researchers. Although RNNs work fine for short sequences but they do not give up to the mark results for longer sequences used for longer period of time. Therefore, LSTM comes in place.

LSTM is an invariant of RNN which is used for remembering long term sequences with the help of addition and multiplication. The information received by LSTM is passed through various cell states and decided further if is to be used or to be forgotten. The following diagram shows the overview of LSTM followed by the details of the layers used in it.

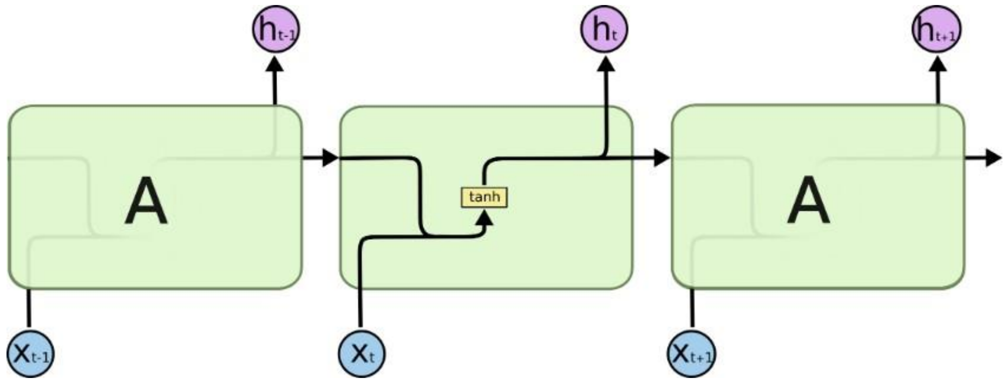


Figure 3.5.2: Architecture of LSTM

i. Forget Gate:

As the by the name, this gate forgets or neglects the information that does not play a vital role for the model. Hence, this information is removed using multiplication filter which ultimately results in enhancement of LSTM's efficiency.

ii. Input Gate:

This gate is used to provide input in three simple steps. The first state, similar to Forget Gate, helps in deciding exactly what to be added into the cell. The second step involves a vector creation of the probable values to be added into the cell. The third and final step involves multiplication of the regulatory filter obtained from step one with the vector created in the later step. In the end the result is added in to the cell state using addition.

iii. Output Gate:

Output gates helps in determining the relevant information and its display as output.

3.6. Working with CNN and LSTM

As mentioned before, for this research we will be using CNN and LSTM where CNN will be used to extract features from the data set of Dukes' University. These extracted features will further be used for classification by feeding them into LSTM.

Going into the details of it, we have used a convolutional neural network having three convolutional layers along with Max Pooling. To avoid any chances of over fitting, Dropout layers were used in order to remove some of the already learned features. The figure below gives a comprehensive view of the architecture of CNN used for this study.

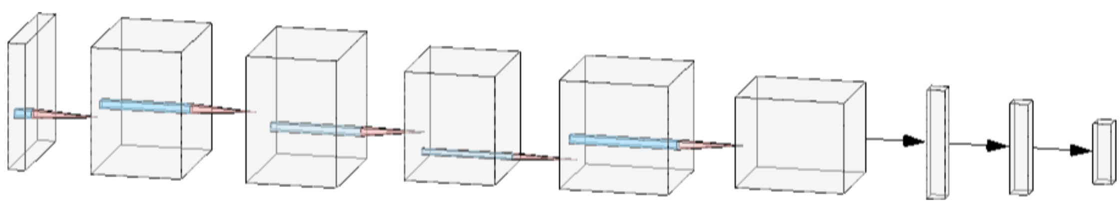


Figure 3.6: Architectural Diagram of CNN and LSTM combined

3.6.1. Layers used for Modelling

The working on various layers and their details are mentioned below along with a diagram providing a complete information in all layers.

i. Input Layer:

The images in the data set were actually in RGB format and greater sizes which required a huge amount of computational resource and greater time consumption. Therefore, making the model resource effective, the images were resized to a size of 128x128 and also converted from RGB to grey scale image having a single channel.

ii. Convolutional Layer 1:

Next layer in our model is the first convolutional layer which is implemented using Convolution2D. It is basically used for extraction of features that are related to curves and edges in the images. Convolution2D function in this layer uses 32 filters with a size od 3x3 for learning the features. It also uses an activation function ‘RELU’ which always helpful and rapid in convergence of the model.

iii. Convolutional Layer 2:

Second convolutional layer is again used with same parameters as the previous one. This layer is used to give a chance to our model to learn every basic feature existing in the data set. It was implemented using Convolution2D with parameters including filters 32 and size of 3x3.

iv. Pooling Layer 1:

Pooling layer is helpful in minimizing the spatial field and works on feature map. It helps in minimal resource utilization and least computational cost by reducing parameters. For this layer we have used MaxPooling because it is a common function used for feature extraction in images.

v. Dropout Layer 1:

This layer as mentioned below aids in avoiding overfitting of the model. The model basically learns the basic features and using this layer, some of the features are removed. For this model the output received as input for this layer gives us certain features. 50% of these neurons are dropped randomly which learned features from previous layers.

vi. Convolutional Layer 3:

After dropout layer 3, we have used the third convolutional layer for learning the deterministic features of the images. These are the high-level features obtained after regularization of the model and these will be used for differentiation of images. This layer is using 64 filters with size of 3x3and the activation method of RELU.

vii. Pool Layer 2:

After extraction of high-level features, we again use a pooling layer to refine the features. Similarly, we will reduce the parameters, keep the relevant and features important for further classification.

viii. Dropout Layer 2:

Second dropout layer will regularize the model and randomly choose the neurons at the rate of 50%.

ix. Flattening Layer 1:

Flatten layer is used to make output used for CNN into a single long vector that is actually a feature and further used for classification. In this model we have used this layer to convert the data into a single dimensional array.

x. Dense Layer 1:

This layer is also known as Fully Connected Layer which is used for classification purpose. This layer is given the out parameter so that this layer can have specified number of outputs. Each of the layers will be representing a single class.

xi. Reshape Layer:

LSTM requires a specific type of input for classification. Reshape layer is used for this purpose. The output received from Dense Layer is reshaped by this layer into 1x64 and later fed into LSTM layer.

xii. LSTM Layer:

This layer will receive and input of 1x64 after reshaping of data and the return sequences for its output are set to true.

xiii. Flattening Layer 2:

After feeding data into LSTM, we again send the output received into second flattening layer for conversion into a form to be fed into dense layer.

xiv. Dense Layer 2:

This layer is again used to classify the data into specific classes after setting the output units.

xv. Dense Layer 3:

This is the last layer which will be ultimately classifying the data into desired classes. Data should be classified into 3 classes namely AMD, DME and Normal. So, 3 is set as output along Softmax as the activation function.

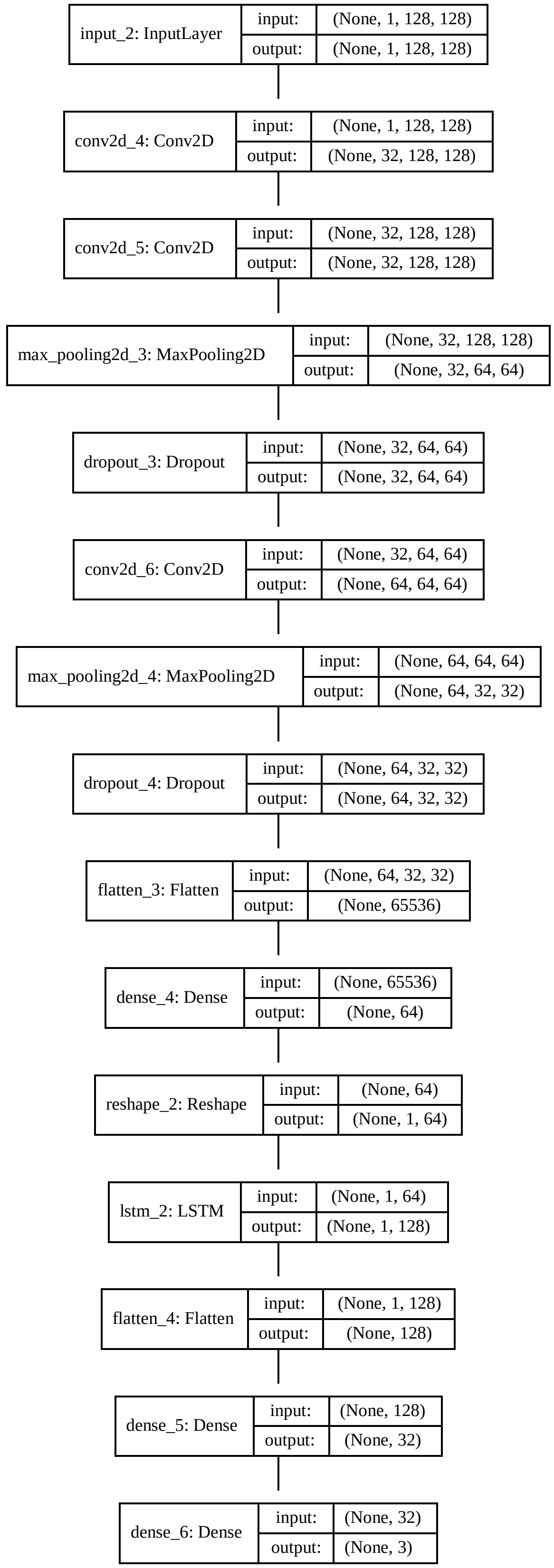

input_2: InputLayer
input: (None, 1, 128, 128)
output: (None, 1, 128, 128)
conv2d_4: Conv2D
input: (None, 1, 128, 128)
output: (None, 32, 128, 128)
conv2d_5: Conv2D
input: (None, 32, 128, 128)
output: (None, 32, 128, 128)
max_pooling2d_3: MaxPooling2D
input: (None, 32, 128, 128)
output: (None, 32, 64, 64)
dropout_3: Dropout
input: (None, 32, 64, 64)
output: (None, 32, 64, 64)
conv2d_6: Conv2D
input: (None, 32, 64, 64)
output: (None, 64, 64, 64)
max_pooling2d_4: MaxPooling2D
input: (None, 64, 64, 64)
output: (None, 64, 32, 32)
dropout_4: Dropout
input: (None, 64, 32, 32)
output: (None, 64, 32, 32)
flatten_3: Flatten
input: (None, 64, 32, 32)
output: (None, 65536)
dense_4: Dense
input: (None, 65536)
output: (None, 64)
reshape_2: Reshape
input: (None, 64)
output: (None, 1, 64)
lstm_2: LSTM
input: (None, 1, 64)
output: (None, 1, 128)
flatten_4: Flatten
input: (None, 1, 128)
output: (None, 128)
dense_5: Dense
input: (None, 128)
output: (None, 32)
dense_6: Dense
input: (None, 32)
output: (None, 3)

Figure 3.6.1: Architecture of all layers of proposed R-CNN

In the end model is compiled and the metric used for getting the performance score is ‘accuracy’. Detailed results are discussed in Chapter 4.

Chapter 4. Results

This chapter gives details about the results gained from the application of the methodology explained in the chapter before. The results are obtained from two types of data sets, the details of which are given below.

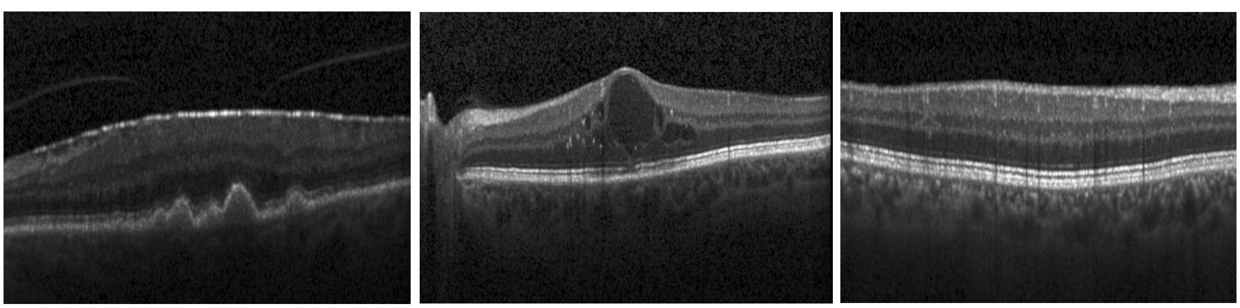
4.1. Data Set

The data set was obtained from two different sources on which the methodology was tested. One dataset was obtained from Dukes' University and the other was a local data set obtained from AFIO. The details of each are described as following.

4.1.1. Dataset: Dukes' University

Dukes' University dataset had images of 45 patients or subjects. For proceeding with the testing, 647 images were selected on scan level and further on patient level. These images belonged to different classes namely AMD, DME and Normal subjects.

Following are some of the scans which show the difference between AMD, DME and Normal Scans which are acquired from Duke's Dataset:



(A) (B) (C)

Figure 4.1: (A) Scan of a patient having AMD. (B) Scan of a patient having DME. (C) Scan of a Normal patient.

4.2. Levels of Classification for Testing

Data was classified on two levels which are described as below.

4.2.1. Scan Level

On scan level, images were classified by separately classifying each image into its separate class. Different images were classified as whether they belonged to AMD, DME or Normal.

4.2.2. Patient Level

Once the images were classified at scan level, they were later classified at patient level through majority class voting. 5/6 patients were classified accurately.

4.3. Performance Metrics

There are different performance metrics which can be used for checking the performance of your classifier. The metric used for this study is Accuracy. Accuracy was calculated as following equation (1):

$$Accuracy=\frac{tp+tn}{tp+tn+fp+fn} \qquad (1)$$

4.4. Final Results

The final results are shown in the form of a confusion matrix and their accuracy is obtained in the end showing the performance of the model.

As we know that we had two types of data which was classified at different levels. The details of these results are mentioned below.

4.4.1. Results from Duke's University

The results were obtained on scan level and patient level and the details of which are as following:

- **Scan level:** On scan level, following confusion matrix was achieved followed by the accuracy obtained.

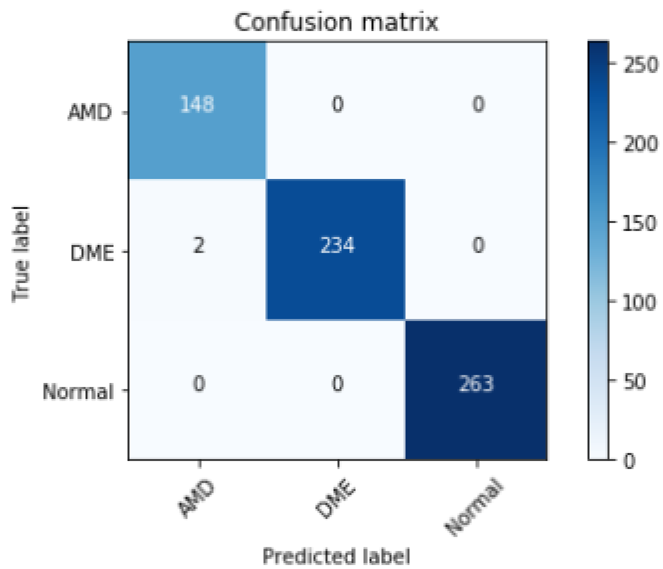


Figure 4.4: Confusion Matrix between Actual and Predicted labels of AMD, DME and Normal Scans

$$\text{Accuracy}=\frac{148+234+263}{148+234+263+2} = 99.69\%$$

So, the accuracy obtained on scan level on data set of Dukes' University is **99.69%.**

- **Patient Level:** On patient level, data of 6 patients was tested from which 5 were classified correctly. The accuracy obtained was **83.33%.**

Chapter 5: Conclusion

This chapter consists of the conclusion of the complete thesis and summarizes the contributions made to

5.1. Contribution and Conclusion

This research gives a solution to a problem related to maculopathy. The disease is occurred in the retina of human eye and causes certain visual impairments among humans. Maculopathy is one of the major causes of sight loss among diabetic and aged patients. This is a dangerous vision disease which if not diagnosed at the earlier stage then can lead to permanent vision loss among its victims. Therefore, as a solution to this problem, we proposed an approach of identifying the OCT images of the patients having this disease. There are two different types associated with this disease: i) Age-Related Macular Degeneration ii) Diabetic Macular Edema. Therefore, our work problem revolves around classification of the OCT images of AMD and DME and Normal images at the early stages.

The problem for this study was solved through CNN and LSTM. The data was obtained from Dukes' University which consisted of OCT images of 45 subjects which were divided into three different types. There were 15 patients of AMD, 15 patients of DME and 15 Normal subjects. This data was then fed into our model further. Before feeding the images into the model, pre-processing was done for the images to get efficient information. After preprocessing, data was then fed into CNN for feature extraction. Once the features were extracted then they had to be classified according to our classes using LSTM. The results obtained from this method were obtained by testing the model on some of the samples of Duke's Dataset which weren't the part of data while feeding into the model for training. While testing the scans from Dukes' University (647 scans) had an accuracy of 99.69% and 6 patients having the sequential scans were tested with the model where we got the accuracy of 83.33% as the model was able to correctly classify 5 patients.

References

[1] Mathers, Colin D., and Dejan Loncar. "Projections of global mortality and burden of disease from 2002 to 2030." *PLoS medicine* 3.11 (2006): e442.

[2] Fujimoto, J., and W. Drexler. "Introduction to optical coherence tomography." *Optical coherence tomography*. Springer, Berlin, Heidelberg, 2008. 1-45.

[3] Maculopathy, Oxford medicine online, Retrieved: August 6th, 2018.

[4] Mousavi, Elahe, Rahele Kafieh, and Hossein Rabbani. "Classification of dry age-related macular degeneration and diabetic macular edema from optical coherence tomography images using dictionary learning." *arXiv preprint arXiv:1903.06909* (2019).

[5] Wang, Yu, et al. "Machine learning based detection of age-related macular degeneration (AMD) and diabetic macular edema (DME) from optical coherence tomography (OCT) images." *Biomedical optics express* 7.12 (2016): 4928-4940.

[6] Al-falluji, Ruaa Adeeb Abdulmunem. "DME Detection using LBP Features." *International Journal of Computer Applications* 148.8 (2016).

[7] Lemaître, Guillaume, et al. "Classification of SD-OCT volumes using local binary patterns: experimental validation for DME detection." *Journal of ophthalmology* 2016 (2016).

[8] Srinivasan, Pratul P., et al. "Fully automated detection of diabetic macular edema and dry age-related macular degeneration from optical coherence tomography images." *Biomedical optics express* 5.10 (2014): 3568-3577.

[9] Karri, Sri Phani Krishna, Debjani Chakraborty, and Jyotirmoy Chatterjee. "Transfer learning based classification of optical coherence tomography images with diabetic macular edema and dry age-related macular degeneration." *Biomedical optics express* 8.2 (2017): 579-592.

[10] Alsaih, Khaled, et al. "Machine learning techniques for diabetic macular edema (DME) classification on SD-OCT images." *Biomedical engineering online*16.1 (2017): 68.

[11] People.duke.edu. (2019). *Srinivasan_BOE_2014*. [online] Available at: http://people.duke.edu/~sf59/Srinivasan_BOE_2014_dataset.htm [Accessed 3 Jul. 2019].

[12] Sun, Zhongyang, and Yankui Sun. "Automatic detection of retinal regions using fully convolutional networks for diagnosis of abnormal maculae in optical coherence tomography images." *Journal of biomedical optics* 24.5 (2019): 056003.

[13] Rasti, Reza, et al. "Wavelet-based convolutional mixture of experts model: An application to automatic diagnosis of abnormal macula in retinal optical coherence tomography images." *2017 10th Iranian Conference on Machine Vision and Image Processing (MVIP)*. IEEE, 2017.