Coventry University

كلية الشرق الأوسط
Middle East College

Project Submission
in Partial Fulfillment of the requirements for the Degree of Master of Science in Information Technology (MSc- IT)

Automated Detection and Classification of Diabetic Retinopathy using Deep Learning for Smart Health in Oman

Author: **(Wafa Rashid Mohammed Aladawi-PF17F1780**

Supervisor: **Dr. C. Jayakumari**

Academic Year: 2019/2020

كلية الشرق الأوسط
Middle East College

MSc Project

Declaration of Originality

This project is all my own work and has not been copied in part or in whole from any other source except where duly acknowledged. As such, all use of previously published work (from books, journals, magazines, internet etc.) has been acknowledged within the main report to an item in the References or Bibliography lists.

I also agree that an electronic copy of this project may be stored and used for the purposes of plagiarism prevention and detection.

Copyright Acknowledgement

I acknowledge that the copyright of this project report, and any product developed as part of the project, belong to Coventry University.

Signed: Dated:

Office Stamp

MSc (IT) Project
Declaration by Examiners

I / We have examined this report titled……

……

……

submitted by……ID No…….

in partial fulfillment of the requirements of MSc (IT) Course during……semester.

Signature of Supervisor

Signature of 2nd Marker

Name of Supervisor
Date:

Name of 2nd Marker
Date:

Acknowledgement

First of all, I would like to thank my supervisor Dr. C. Jayakumari for her guidance and support for me during my project. She has been a wonderful mentor for me, I learnt a lot from her vast knowledge and experience while her supervision to my project. I would like to express my gratitude to her for encouraging my research and for allowing me to grow as a research scientist. Without her assistance, I wouldn't be able to complete my project successfully.

The deepest gratitude and thanks go to the Middle East College (MEC) for providing countless support and valuable guidance. Their support has enabled me to go against all odds and complete this project properly. Also, I would like to thank them for giving me the opportunity to use their resources and complete my project in the college.

I would also like to thank the consultant ophthalmologist Dr, Nassra Al-Habsi from Al-Nahdha Hospital who collaborated in the development of this project and for providing clinical guidance.

Finally, my gratitude also goes to my wonderful parents, my husband, and other members of my family for their countless support.

Abstract

Diabetic retinopathy is one of diabetes complication that causes damage to retina of eye and may leads to a gradual decrease in vision. Early detection of diabetic retinopathy is critical to avoid blindness. In Oman, retinal screening still in manual process which can be costly and time-consuming because of large diabetic population. This study aims to automatically detect and classify the different DR grade using deep learning algorithms. Predicting the diabetic retinopathy presence is performed using convolutional neural network (CNN) which provide a great advantage of screening for diabetic retinopathy. This study makes use of Kaggle dataset to train the proposed model. The experiment results show that deep learning has an excellent performance when detecting and classifying the diabetic retinopathy grades for Kaggle dataset.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

DM	Diabetes Mellitus
DR	Diabetic Retinopathy
NPDR	Non-proliferative Diabetic Retinopathy
PDR	Proliferative Diabetic Retinopathy
DME	Diabetic Macula Edema
MOH	Ministry of Health
WHO	World Health Organization
T1DM	Type 1 Diabetes Mellitus
T2DM	Type 2 Diabetes Mellitus
GDM	Gestational diabetes mellitus
AI	Artificial intelligent
ML	Machine Learning
DL	Deep learning
CNN	Convolutional Neural Networks
Adam	Adaptive Moment Estimation
SVM	Support Vector Machine
RNN	Recurrent neural network
RBM	Restricted Boltzmann machine
DRIVE	Digital Retinal Images for Vessel Extraction

1 Chapter 1: Introduction

1.1 Background

Diabetes Mellitus (DM) is a serious chronic disease that occurs when the pancreas doesn't produce enough insulin (a hormone that regulates the level of sugar or glucose in the blood) or when the body cannot use the insulin effectively. Symptoms of diabetes begin to appear and develop gradually in diabetes patients such as frequent urination, weight loss, nausea and vomiting, excessive thirst and appetite, and slow healing wounds.

DM can be categorized into three types which are: Type 1 diabetes (T1D), Type 2 diabetes (T2D), and Gestational diabetes mellitus (GDM). T1D occurs when body immune system mistakenly attacks and destroys beta-cells in the pancreas that produce insulin and cause permanent damage. The cause of this attack is not known specifically. There may be genetic and environmental factors. Therefore, the patient with the first type of diabetes should receive insulin daily, throughout his life. Whereas the second type of diabetes is the most common type of diabetes among people, and it is between 90 and 95 percent of the total number of patients with diabetes (Who.int, 2019). T2D starts as an insulin resistance, that means that the body cannot use insulin efficiently and this stimulates the pancreas to produce more insulin until it becomes unable to keep pace with demand. So, decreasing insulin production in the body leads to high blood sugar. The exact cause of T2D is unknown, but can include heredity factors, lack of exercise, weight gain, and there may also be other health factors and environmental causes. The third type of diabetes is gestational diabetes which occurs only during pregnancy and women more likely to develop to type 2 diabetes within 5 to 10 years (Cold et al., 2019).

Diabetes is a significant public health problem and is one of most targeted diseases for action by world leaders. Over the past few decades, there has been significant rise in the prevalence of diabetes among people around the world. An estimated '422 million' adults worldwide were diagnosed with diabetes in 2014, compared to 108 million in 1980. The prevalence of diabetes has doubled globally since 1980, rising from 4.7% to 8.5% among adults (Apps.who.int, 2019). Diabetes in all its types can lead to multiple complications in the body and increased the risk of premature death in general. Potential complications include heart attack, stroke, kidney failure, amputation, blindness and nerve damage.

Diabetic retinopathy (DR) is one of diabetes complication that affects the eye and occurs as a result of damage to the blood vessels of light-sensitive tissues in the back of retina. The hyperglycemia continuous causes to a blockage in the tiny blood vessels that nourish the retina. Thus, the eye tries to create new blood vessels, but these vessels don't grow properly and can bleeding easily. DR in its early stages doesn't cause any changes in vision, but with the development of the disease, the level of vision worsens significantly may lead to blindness.

The disease risks associate with the length of diabetes and therefore more than 75% diabetes patients have DR after 20 years of the diabetes (Who.int, 2019).

The exact distinction between the different grades of diabetic retinopathy to define each grade's features make the recognition task to classify DR very challenging. It has been shown that periodic screening of DR and timely treatment reduces the risk of blindness (Raman et al., 2018). Hence, DR screening and automated detection of diabetic retinopathy is critical to address these challenges as well as to detect the disease at early stage to prevent blindness with appropriate treatment using advances technical medical developments (Singh Sisodia, Nair and Khobragade, 2017).

1.2 Problem Statement

Diabetes is a global epidemic and about 15% of Omani adults are infected with this disease, and this percentage rises to 25% for those over the age of sixty years. The proportion of people with diabetes from 2008 to 2018 increased more than 3%, and there are more than 7500 diabetes patients in the Sultanate of Oman annually (Alwatan.com, 2019). Therefore, the competent authorities must be concerted efforts to reduce the prevalence of this disease. Currently in Oman, diabetic retinopathy is detected by manual process that needs optometrist to examine retina photographs and consequently this process consumes a lot of a time. As aforementioned, the number of diabetics is increased and that means ophthalmologist workload will increase because they will spend much time to detect diabetic retinopathy. Therefore, ophthalmologists will not be able to take care of the patients with their full potential. So, an automated system is required to detect DR, reduce time and cost while diagnose diabetes patients.

1.3 Aim & Objectives

Based on the aforementioned problem, the main aim of this study is to develop an effective system for diabetic retinopathy screening that can contribute towards improving DR management. The system will make use of deep learning algorithm in order to detect important features on the retina photographs effectively and classify patients into the correct DR grades. This automated system will facilitate a reduction in the burden of manual grading for the ophthalmologists and help ease the pressure on the eye screening centers in Oman.

The main objectives of this research are as follows:

1. To study about DR disease and its impact in Oman.
2. Study about the Deep Learning algorithm using python for image processing.
3. Early detection and classification of DR using Deep Learning algorithm to prevent blindness.
4. Reduce time and cost taken for diagnosis of DR.

5. To evaluate the accuracy of deep learning algorithms for detection and classification of DR.

1.4 Thesis Workflow

The following chart describe the workflow for this study:

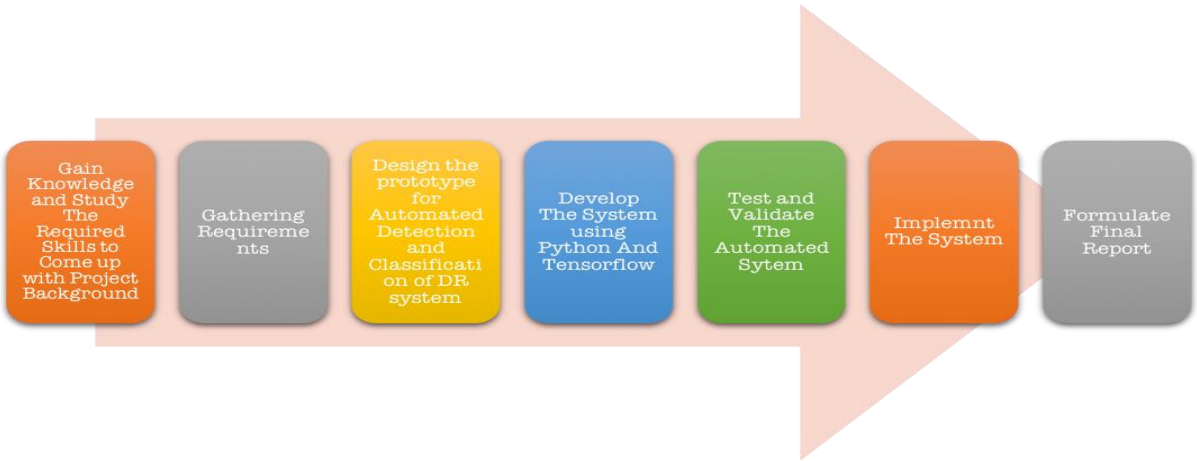


Figure 1: Research Workflow

1.5 Motivation

The main purpose behind selecting this topic for this research that Oman is one of the developing countries and there is shortage of resources to overcome some problems. One of these problems is the number of diabetics is increased day by day and there is no effective system that assist ophthalmologists to detect DR and classify it. The significance of the proposed research is to overcome the existing problems that face the ophthalmologists in the DR screening, such as:

1. Manual diagnosis of patients with diabetes by the ophthalmologist
2. Time taken to diagnosis DR

During searching about these problems, the researcher got motivated to work with this topic that would assist a huge amount of people in their healthy lives. As there are inadequate ways to detect about diabetic retinopathy in Oman, an automated system will be developed that makes use of deep learning algorithms with the help of image preprocessing. The system will facilities detection process and give prediction about diabetic retinopathy.

1.6 Thesis Orientation

The thesis is organized into seven chapters, each chapter concentrating on different part of the research work. The following is a short description of contents for each chapter:

Chapter 1 is the **Introduction** of the thesis. This chapter provides the background of the basic information on diabetes mellitus and diabetic retinopathy, as well as the prevalence of diabetes

mellitus worldwide and, particularly, in Oman. Also, provides an overview of current problems about screening diabetic retinopathy in Oman and presents the aims and objectives of this thesis. Moreover, motivations that have led to this thesis are described here.

Chapter 2 is **Literature Review**. This chapter consists of "Literature Review" of different topics that related to this thesis. In addition, this chapter described the "Related Works" of existing researches done by others, and that are related to our research work.

Chapter 3 is **Methodology** description that used for this study. Research methodology and project methodology are described in this chapter that were used to complete this study. As well as this chapter provide justification for each methodology selected.

Chapter 4 is **Project Management**. This chapter illustrate project management methodology that has been selected to complete this project. Moreover, this chapter describes different components of project management such as tasks schedule, work breakdown structure (WBS), milestone, deliverables, project management risks, etc.

Chapter 5 is **Requirement Gathering & Analysis**. This chapter contains description of data gathering methods that used for this research as well as the analysis of data that were gathered.

Chapter 6 is **Design and Implementation**. The proposed methodology framework is explained in this chapter. Also, this chapter describe all software and hardware tools that used for this study. Finally, explained the selected algorithm for this study and describe the different stages of project implementation.

Chapter 7 is **Experimental Result Analysis**. This chapter provide analysis of experimental results of deep learning algorithm that applied for this study. This chapter makes use of plots and different kind of comparison graphs to show the DL results.

Chapter 8 is **Critical Appraisal**. This chapter provides systematically assess the scientific research to identify the strengths and weaknesses as well as to assess the research finding. Also, this chapter present the critical reflection done by student.

Chapter 9 is **Social, Legal, Ethical Considerations & Project Sustainability**. This chapter briefly discussed the social, legal, ethical considerations and sustainability of this project.

Chapter 10 is **Conclusion and Future Works**. This chapter concludes this thesis and summaries the main results and achievements. As well as, the future work of this study is mentioned in this chapter.

2 Chapter 2: Literature Review

2.1 Diabetes Mellitus

Diabetes is a chronic disease that occurs when the pancreas fails to produce enough insulin or when the body cannot use the insulin effectively that it produces. Insulin is a hormone that regulates the level of sugar in the blood. Hyperglycemia is a common effect of uncontrolled diabetes and, over time, causes serious damage to many organs, especially nerves and blood vessels. Diabetes mellitus can be categorized into three types which are: Type 1 Diabetes (T1D), Type 2 Diabetes (T2D) and Gestational diabetes. T1D is known as insulin-dependent diabetes, which begins in adulthood or infancy, and is characterized by the lack of insulin production. This type of diabetes requires daily insulin use, and there are no means of prevention so far. The symptoms of T1D are excessive urination, thirst, persistent hunger, weight loss, visual disturbances, and feeling tired. T2D is known as non-insulin-dependent due to the inability of the body to use insulin effectively. The symptoms of T2D are often the same symptoms of T1D because of obesity and physical laziness, but it is often less clear. Gestational diabetes occurs during pregnancy because of a hyperglycemia which the glucose rate increases above the normal rate without reaching the rate necessary for the diagnosis of diabetes. Gestational diabetes is diagnosed by prenatal screening, not by the above symptoms. women who develop gestational diabetes are likely to develop type 2 diabetes in the future (Mayo Clinic, 2019).

Although diabetes is a chronic disease and cannot be cured, but there is a great need for treatment and control it for several reasons. The most important reason is to prevent the occurrence of complications associated with diabetes disease, or to alleviate these complications if occurred through conservation of the level of sugar in the blood (Cold et al., 2019). Diabetes causes many complications which divided into two types which are: Microvascular Complications including diabetic retinopathy, Diabetic Nephropathy and Neuropathy (Nerve Disease) and Macrovascular Complications include Cardiovascular Disease, Diabetic Foot Disease, Dermopathy.

2.2 Diabetic Retinopathy

Diabetics may experience complications at the retinal level caused by high blood sugar. This condition is called diabetic retinopathy (DR). DR is one of diabetes complication that caused by damage to blood vessels in light-sensitive tissues in the back of the eye (retina). Diabetic retinopathy in its early stages does not cause any changes in vision, but with the development of the disease, the level of vision worsens significantly may leading to blindness. Diabetes Retinopathy is specified by indications of retinal ischemia (retinal microvascular anomalies, microaneurysms, hemorrhages, intra-venous gauge variations from the norm, cotton-fleece spots, and neovascularization) as well as indications of expanded retinal vascular porousness. Vision deficiency can result from a few systems, including neovascularization causes vitreous discharge and additionally retinal separation, macular edema, and retinal narrow nonperfusion.

Based on these indications, retinopathy is divided into non-proliferative diabetic retinopathy (NPDR) and proliferative retinopathy (PDR). The NPDR is additionally partitioned into mild-NPDR, moderate-NPDR, and severe-NPDR. People with diabetes fear loss of vision because diabetic retinopathy leads to several diseases (Wu, 2013).

Most researches state that diabetic retinopathy should be treated in the early stage of its prevalence which minimizes the permanent blindness caused by the longer duration of diabetes. Besides, the studies have shown that the retina of a diabetic person can tell the impact which affects the hyperglycemia in the period. Based on this research, it is better to treat the diabetic retinopathy that have long-term diabetes to them improve the vision and for better health purpose (Nei.nih.gov, 2019).

2.3 Diabetes and diabetic retinopathy in Oman

Haddad and Saad, 1998 in their research studied the prevalence of DR in Oman. 500 diabetic patients were selected to detect the presence of diabetic retinopathy in Al Buraimi hospital in Oman. The presence of diabetic retinopathy was graded into different grades as mild NPDR, moderate-severe NPDR and proliferative retinopathy (PDR). 212 out of 500 patients was detected with diabetic retinopathy. 128 patients were detected with mild-NPDR (25.6% of the total patients), 20 patients with moderate-NPDR (4%), and 64 patients with proliferative diabetic retinopathy (12.8%). They stated in their study that diabetic retinopathy can be overcome by controlling the glycaemia, blood lipids should be low and controlled blood pressure (hypertensive patients) can lower the effect of retinopathy in the patients.

Alrawahi et al., 2012 in their studies found that the prevalence of diabetic retinopathy in Oman was more among males (51.6%) compared to females (36.5%). They stated that the prevalence of diabetic retinopathy was associated with significant factors such as gender, long duration of diabetes, family history of diabetes and poor glycemic control. People with diabetes in Oman have high rates of complications associated with diabetes. it was found that more than 14% of Oman's diabetic population suffer from diabetic retinopathy as shown in Figure 2 (Al-Shookri et al., 2019).

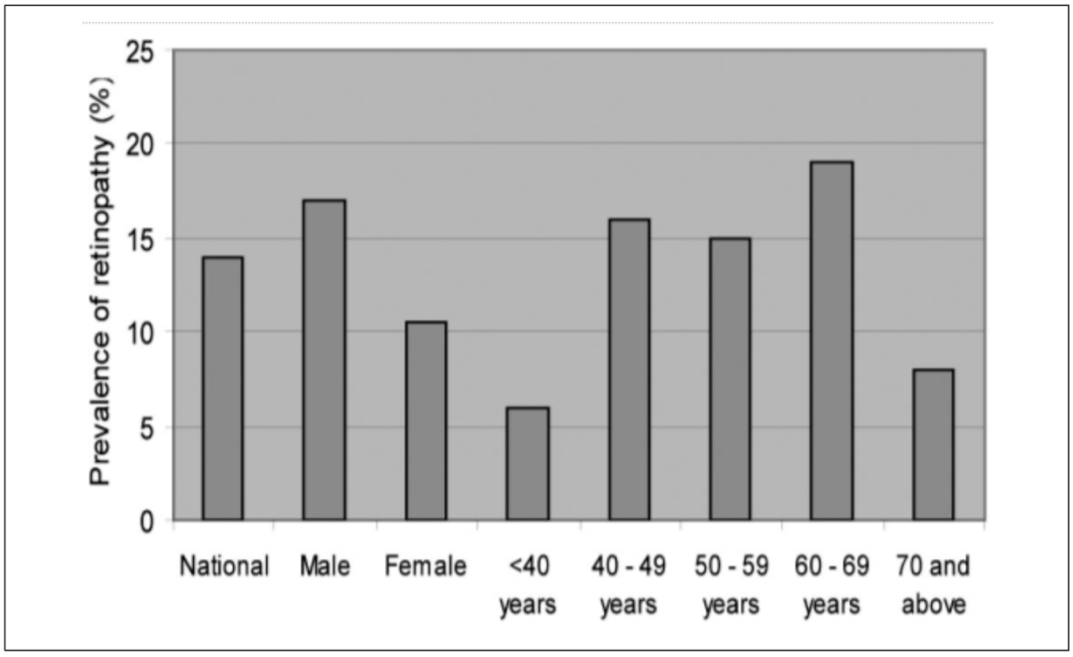


Figure 2: Prevalence of diabetic retinopathy in Oman (Al-Shookri et al., 2019)

2.4 Artificial Intelligence

Artificial intelligence is the ability of the machine to simulate the human mind and the way it works, such as its ability to think, discover, and benefit from past experience. Since the development of the computer in the mid-twentieth century, it has been discovered that the computer can perform more complex tasks (Rada, 1986).

Artificial intelligence is going beyond human levels on a number of different complicated tasks. Nowadays, artificial intelligence has received a great interest in the field of medicine. Therefore, doctors and radiologists rely on artificial intelligence for accurate and correct results. Artificial intelligence has become widely used in the medical field, using software to improve and develop human knowledge in the analysis of complex medical data. There are many uses of artificial intelligence in the field of health, for example used in the modernization and development of medicines and treatments, used in the process of screening and knowledge and detection of various diseases, used in maintaining the patients safety by monitoring their health and care, and also used in the development Good systems and protocols for the treatment of patients (Ramesh et al., 2004).

In this regard, artificial intelligence focused on the diagnosis of diabetes complications through eye examination, due to the critical importance of the treatment of this disease. Recently, AI techniques have been used to detect retinal damage after algorithms have been trained in thousands of images, which help doctors diagnose and treat patients more effectively and quick (Lu et al., 2018).

2.5 Deep Learning

Deep learning (DL) is a new technique and new branch of learning a machine technology under the term which now the most popular "Artificial Intelligence". Many other companies like google, Facebook, etc. have been using this method for data analysis to know about the surfing of the individual user. DL is a subset of machine learning (ML) introduced to bring ML closer to its original objectives. Deep learning models learns to accomplish classification tasks directly from text, images, or audio. DL models can attain sophisticated accuracy, occasionally exceeding human performance (Singh et al., 2007).

Nowadays, deep learning has become a new trend in incorporating in the field of medicines. A lot of researches have also been done on the detection of diabetic retinopathy using deep learning techniques as well as in machine learning. Deep learning is a very important tool for the health care purpose which allows to predict the disease in a patient at an early stage. Deep learning technique helps to specify the disease and specific treatment for a patient. Software based on deep learning technique for diabetic retinopathy showed high sensitivity and specificity for detecting the diabetic retinopathy (Gulshan et al., 2016).

There are various studies on algorithms for diabetic retinopathy. Ting et al., 2018 in their studies suggested that more research should be done for detecting diabetic retinopathy using deep learning technique in order to specifically treat the patients according to the type of disease stages more easily. Further develop the deep learning techniques with algorithms in diabetic retinopathy screening programs lead to better care of the patients. Gargeya and Leng, 2017 in their study shows a great deal of effective use of deep learning technique in detecting the type of diabetic retinopathy in a patient because the world is leading to an era of technology so is the medicine. This technology is efficient in saving time and treating quickly if further development is done successfully and cost-effective as well.

2.6 Image Preprocessing

The most important step in predicting using deep learning models is image preprocessing to improve the quality of images and therefore improve predicting results. Building an effective deep learning model needs carefully consider the input data format. There are several techniques which enhances the pixels of the image, filtering is also done and detection of the blood vessels through these images and comparing it with the recorded data to predict the results according to the statistical data to get a clear vision about the severity of the disease caused by the diabetic retinopathy via different preprocessing of the images and then integrating them with encoders to convert the data and provide the information in a readable form. This new technology has great advantages in analyzing the diabetic retinopathy. They are combined to improve the accuracy of the results automatically. They also have different transformation scales for colored images. There are different methods for preprocessing the retinal photographs of poor quality using deep learning method such as Grayscale conversion,

Adaptive Histogram Equalization, Discrete Wavelet Transform, Matched filter and Fuzzy C-means segmentation for preprocessing (Sayed et al., 2017).

2.7 Deep Learning Methods

The deep learning methods have deep learning algorithms which have subsets in it. The most used deep learning algorithms currently available are supervised, unsupervised and reinforcement learning (Vogt, 2018). These subsets have further categories which are the following:

2.7.1 Convolutional neural network (CNN)

A convolutional neural network is a specific type of model which accepts the 2- dimensional images as input data and time series as well. This method is based on linear operation, the layer is always present in one of the networks. The output of the operated convolution is in non-linear activation function by means of a pooling function (Vogt, 2018). The pooling function changes the location of the output and represents the invariants to small data and does subsampling of the input. The most common pooling function is max pooling which gives the maximum output in rectangular form. The pooling and convolution are layered together to obtain a classified way. The main advantage of CNN is their accuracy in image recognition. The disadvantages of CNN are that it has a high cost, a lot of training required and also a good GPU to run to properly (COŞKUN et al., 2017).

2.7.2 Recurrent neural network (RNN)

A recurrent neural network is a type of model which not only processes the current data but also the previous outputs as well. These previous outcomes are encoded in hidden form. The recurrent neural network can access the previous data automatically which is very useful in the sequencing data. The memories are encoded in the internal form in a recurrent neural network. The recurrent neural networks are made up of in such a way that they use backpropagation, which computes the gradients. The problem in this method is that the data becomes immeasurable or extremely small because a number of data is accumulated. The pros of RNN is that they are friendly to GPU and have a simple lock step computation. The cons of RNN is that they have vanishing gradient and gradient exploding problem. They cannot be used in deep models (Du and Swamy.,2013).

2.7.3 Restricted Boltzmann machine (RBM)

Restricted Boltzmann machine (RBM) is a type of model which works on 2 variables which are connected to each other. The hidden processing helps the restricted Boltzmann machine to produce the stimulation on the forward pass while the visible layers help them to work on the rebuilding of the backward pass as well. This algorithm is helpful in reducing, classifying and many more things. The main advantage of RBM is that the samples created looks like generated from your data and pattern completion can also be done by RBM. The disadvantage of RBM

is that they are tricky to use with different algorithms and missing values. The loss cannot be tracked in RBM (Harrison, 2018).

2.7.4 Autoencoders

Autoencoders neural networks have 3 layers in which the output is connected directly with the input layers. There are a greater number of visible units than the hidden units. They find more useful input data to represent the output data. The major advantage of using Autoencoders is that they are able to learn non-linear feature and reduce the dimensional stability. The drawback of Autoencoders is that they are expensive computationally to train, cannot be uninterpreted and the underlying maths is more difficult (Vogt, 2018).

2.8 Libraries

Libraries are sort of machine learning algorithms, but they differ in their implementation method. There are different types of software developed in contrast to machine learning algorithms to simply process the machine tools automatically which execute the tasks set by the user. They also execute different neural networks and functions (Deeplearning.net, 2019). However, they are theoretically same for both the machine learning algorithms. These are some of the most commonly used libraries listed below:

2.8.1 Keras

It is the same as Lasagne, Keras is a high-level state wrapper which keeps running over Theano. Moreover, it is additionally ready to keep running over TensorFlow. In contrast to Lasagne, which continuously utilizes some Theano code, Keras won't demonstrate any of the fundamental work. It is intended to limit overhead, to consider quick and simple prototyping of machine learning calculations (Keras.io, 2019).

2.8.2 TensorFlow

TensorFlow is a Python library and stands separated by being the just one of the significant libraries created from the beginning by a real enterprise, while the others have their source in the exploration network. TensorFlow has some incorporated personal satisfaction instruments, for example, Tensor Board, which enables the client to effectively create charts envisioning things, for example, learning rate, model loads, misfortune capacities and that's just the beginning. TensorFlow is likewise the main library that can disseminate the remaining task at hand crosswise over GPUs on a similar gadget as well as on a few associated gadgets, which can be a noteworthy computational pro (TensorFlow, 2019).

2.8.3 Theano

Theano is a Python library for scientific articulations in Python, created with the objective of encouraging examination in profound learning. With Theano you can characterize, upgrade, and assess numerical articulations proficiently, utilizing multi-dimensional exhibits. Its linguistic structure is like NumPy and this joined with advanced local machine code makes

Theano an incredible asset, particularly when actualizing machine learning calculations. Theano enhances the selection of articulations before calculation and would then be able to make an interpretation of it either to C++ that the program will be kept running on the CPU (Deeplearning.net, 2019).

2.8.4 Lasagne

Lasagne is a python encoder which is simple in building and training the difficult machine learning algorithms. This library used to develop and train the neural networks in "Theano". Lasagne used with Theano but cannot be used as a replacement. They are very helpful tools (Lasagne.readthedocs.io, 2019).

2.9 Publicly Available Datasets of Retinal Image

Datasets for machine learning have a number of lists and has a vast range to specifically answers the question arouses in the mind of an individual. Datasets are highly important and should be of a good quality specifically for the detection of the diabetic retinopathy using deep learning methods. There are different public datasets can be used for deep learning models as following:

2.9.1 Kaggle datasets

Kaggle.com is an online community that created by google LLC. Kaggle.com is one of the most popular sites midst artificial intelligence engineers and data scientists. It provides a lot public dataset that can be used for deep learning project. It can be used in academic papers but highly depends upon the dataset. Kaggle includes a good number of datasets for analyzing diabetic retinopathy in order to compare the results (Kaggle.com, 2019).

2.9.2 Messidor-2 dataset

Messidor-2 is a type of dataset which is an original extended from Messidor database particularly for diabetic retinopathy. It has approximately 1748 images. These images were taken by the Topcon TRC NW6 non-mydriatic fundus image of retina at an angle of 45 degrees (Guillaume PATRY, 2019).

2.9.3 DRIVE dataset

This DRIVE dataset (Digital Retinal Images for Vessel Extraction) is developed to compare the blood vessels segmentation in the retinal images which are caused by the long-term diabetes mellitus. Different algorithms compare their results with this database and share it with the researchers. There are different ways to download these datasets and inspect and evaluate the result according to them. The images were taken from a diabetic retinopathy screening program held in the Netherland. It has a good collection of the dataset for the evaluation of the diabetic retinopathy that ages between 25-90. It shows 7 indications for mild and early diabetic retinopathy. The format of the image is JPEG compressed. The images were

specifically taken by Canon CR5 non-mydriatic 3CCD camera at an angle of 45 degrees. They are both divided into training and test sets (Isi.uu.nl, 2019).

2.10 Challenges and Limitations

Notwithstanding the promising results achieved using deep learning, there are still many unresolved challenges facing the clinical applications of deep learning for healthcare. One of the main issues faced deep learning applications is data quality. Medical data is very noisy, heterogeneous, incomplete and ambiguous, unlike other domains where the data are clean and well prepared. Training DL models using large datasets and variegate is difficult and requires considering numerous issues, such as data discrepancy, repetition, and missing-values (Miotto et al., 2017). Paranjpe and Kakatkar, 2014 pointed out in their paper that varying illumination across the retina images is one the major challenges facing the automated detection of diabetic retinopathy. Also, they stated that the similarity of retinal features such as optic disc, blood vessels, and macula is another challenge. All these challenges provide numerous opportunities and possibilities of future research to enhance this field. Therefore, there would be a promising future for deep learning in healthcare.

2.11 Related Work

There are numerous methodologies of automated detection and classification of diabetic retinopathy. Several methods are available from the literature that based on neural networks, fuzzy C-means clustering, mathematical morphology, pattern-recognition, etc. Table shows a comparison of previous studies that carried out for diagnosing diabetic retinopathy

In 2016, Maher and Dhopeshwarkar presented algorithms for an automated detection of diabetic retinopathy using Support vector machine (SVM). Their proposed methodology made use of DIARETDB1 of total 89 images and achieved sensitivity of 94 %, specificity of 94.7% and accuracy of 95.38%.

In a work carried by Gargeya and Leng, 2017, automated DR detection has been developed and evaluated. They have used local datasets for training, whereas they used the public MESSIDOR-2 and E-Ophtha datasets for testing. Their model achieved AUC of 0.97, sensitivity of 94% and sensitivity of 98%.

Raju et al., 2017 in their work built a system for automatic diagnosis of diabetic retinopathy using convolutional neural network (CNN). Publicly available Kaggle dataset were used to train the network their model. Around 35,000-images were used for training network and achieved 80.28% sensitivity, 92.29% specificity on training and achieved accuracy of 93.28% on the validation datasets.

Dutta et al., 2018 proposed a model of automated knowledge to identify the major antecedents of diabetic retinopathy. In their study, trained proposed model with three types of deep learning methods which were Back Propagation-NN, Deep Neural Network (DNN) and Convolutional

Neural Network (CNN). After testing models with CPU, they concluded that DNN outperformed CNN on training accuracy as well as validation accuracy.

Hemanth, Deperlioglu and Kose, 2019 in their study proposed solution for diagnosing diabetic retinopathy from retinal images (MESSIDOR) using image processing and deep learning. They used HE, and CLAHE approaches for retinal images enhancement, whereas they used convolutional neural network to perform DR classification. Different performance parameters were achieved such as 97% of accuracy, 94% of sensitivity, 98% of specificity, 94% of precision, 94% of FScore, and 95% of GMean.

Table 1: Comparison of previous studies that carried out for diagnosing diabetic retinopathy

No.	Authors	Year	Datasets	Method	accuracy	sensitivity	specificity
4	(Hemanth, Deperlioglu and Kose, 2019)	2019	MESSIDOR	CNN	97%	94%	98%
6	(Cao, 2019)	2019	Local datasets	Bayesian model	93.5%	0.949	0.928
7	(Sahoo and Sekhar, 2019)	2019	STARE, DRIVE, DIARETDB1 and DIARETDB0	Wavelet based Image Segmentation	95.72%	95.87%	96.2%
9	(Mahiba and Jayachandran, 2019)	2019		Hybrid Structure Descriptor and Modified CNNs	98.41%		
10	(Hassan et al., 2019)	2019	Zhang dataset and Rabbani dataset	CNN	94.33%		
11	(Shanthi and Sabeenian, 2019)	2019	Messidor	CNN	96.6% and 96.2%, 95.6% and 96.6%		
	(Wan, Liang and Zhang, 2018)	2018	Kaggle	CNN	95.68%		
5	(Li et al., 2018)	2018	Internal datasets	CNN		92.5%	98.5%
1	(Raju et al., 2017)	2017	Kaggle	CNN	93.28%	80.28%	92.29%
3	(Gargeya and Leng, 2017)	2017	local dataset, MESSIDOR 2 and E-Ophtha			94%	98%
2	(Maher and Dhopeshwarkar, 2016)	2016	DIARETDB1	SVM	95.38%	94 %	94.7%
8	(Jadhav and Patil, 2016)	2016	DRIVE and STARE	Wavelet Transform	95%		

3 Chapter 3: Methodology

As it is indicated in the title, this chapter explain in detail what does research methodology means. The chapter will explain first the research approach that has been chosen to apply in this research. Furthermore, the author explains the research process and the different methods that used for data collection that have applied as well as state the sources of data. Then, the chapter goes on to discusses the software development methodology which used to complete the related product.

3.1 Research Methodology

Research methodology is a set of systematic technique that used to solve the problem of research, as to how research is done scientifically (Godwill, n.d.). Number of activities was performed by the researcher to complete the research work. Under this section, the researcher identified the actions that have been taken to investigate the problem of research and described the techniques used to identify information to understand the problem. Outlining the research methodology is important to explain the purpose of the study and define the problem statement clearly to that allow reader to critically evaluate a research's overall reliability and validity. In this section, the researcher answered two main questions in this report which is: How the data was collected and, how it was analyzed. The figure below states the flowchart of research process.

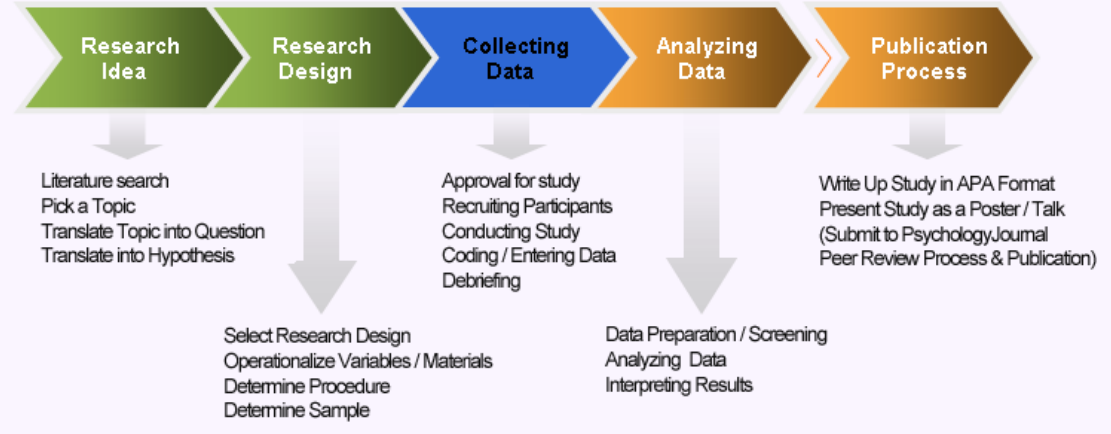


Figure 3:Research Process Flowchart

3.1.1 Research approach (Qualitative vs Quantitative)

The method that should be used in a research is an issue that must be considered carefully by each researcher who wants to acquire correct and scientifically reliable results, because the method has direct impact on the entire research process. From this aspect, selecting an appropriate method is crucial for scholars to achieve reliable and accurate results (Rahman, 2016). The most significant thing the researchers make during research method selection is differentiate between quantitative and qualitative data. The table (1) below shows the features of qualitative and quantitative research methods.

Table 2: Features of Qualitative & Quantitative Research Methods

Qualitative research	Quantitative Research
The aim is a complete, detailed description.	The aim is to classify features, count them, and construct statistical models in an attempt to explain what is observed.
Researcher may only know roughly in advance what he/she is looking for.	Researcher knows clearly in advance what he/she is looking for.
Recommended during earlier phases of research projects.	Recommended during latter phases of research projects.
The design emerges as the study unfolds.	All aspects of the study are carefully designed before data is collected.
Researcher is the data gathering instrument.	Researcher uses tools, such as questionnaires or equipment to collect numerical data.
Data is in the form of words, pictures or objects.	Data is in the form of numbers and statistics.
Subjective – individuals' interpretation of events is important, e.g., uses participant observation, in-depth interviews etc.	Objective: seeks precise measurement & analysis of target concepts, e.g., uses surveys,
Qualitative data is 'richer', time consuming, and less able to be generalized.	Quantitative data is more efficient, able to test hypotheses, but may miss contextual detail.
Researcher tends to become subjectively immersed in the subject matter.	Researcher tends to remain objectively separated from the subject matter.

Based on the above table that describe the features of qualitative and quantitative research methods, the researcher was used both methods simultaneously in order to satisfy the objectives of this research. The researcher used a combination of quantitative and qualitative to ensure that the understanding of research problem is enhanced by combining different methods of research than either qualitative or quantitative methods alone. In this research, the researcher collected both qualitative data (text, images) and quantitative data (numbers). Furthermore, one advantage of mixing both approach in the research is that the strengths of one approach could make up for the weaknesses of other approach. Also, the researcher was able to use more available tools to collect more comprehensive data. This provides outcomes with a comprehensive perspective of the research problem.

3.1.2 Sources of Data

There are two different sources of data that the researchers could use to gather data for particular research which are: primary data collection and secondary data collection. In this particular research, the researcher was used both sources, primary and secondary data in order

to achieve an accurate and objective results. The primary data was collected from ophthalmologists and patients that have diabetes especially those who have diabetic retinopathy. The secondary data was collected from online datasets (Kaggle), journal papers, books, websites and newspaper.

3.1.2.1 Primary data

Primary data refers to an original data that collected by a researcher for a specific purpose for the first time. Primary data is collected directly from primary sources for the main research purpose. Gathering Primary data is consumed time and expensive compared to the secondary data gathering. The following methods were used to collect the primary data for this particular research:

- Questionnaire: The researcher used questionnaire method because it is faster and cheaper than other methods of collection data. The researcher conducted a survey with 50 people who have diabetes mellites from different age by using SurveyMonkey.
- Interview: The researcher conducted face to face interview with ophthalmologist from An-Nhada Hospital.

Justification for using primary data

Primary data collection is critical because it draw the data directly from the field. The researcher got specific information about the research. Primary data is more reliability than secondary data because it is collected by the concerned researcher. Furthermore, primary data could give a realistic view to the researcher about the research topic.

3.1.2.2 Secondary data

Secondary data is the data was gathered by other parties or collected from other researches. Secondary data methods are analyzing the data that have been already gathered by others for some purpose. For this research, literature review is the main method that used to collect secondary data.

- Literature Review: The researcher is pre-reviewed journal papers, books, conference papers and dissertations that are relevant to the research topic. The main topics that are reviewed are:

  - → Diabetes Mellites
  - → Diabetic Retinopathy
  - → Artificial Intelligence
  - → Image Preprocessing
  - → Deep Learning

- Acquiring online datasets: For the purpose of the research, the researcher downloaded datasets (retina images) from Kaggle.com that are publicly available with Csv file that consist the label of the images (normal, NPDR, PDR)

Justification for using Secondary data

Secondary data helped the researcher to enhance the understanding of the problem statement. The researcher used secondary data due to its cost effectiveness. Others have already collected the data; thus, the researcher did not need to spend any time, effort or cost into the data collection phases. Through using secondary data, the researcher avoided the difficulties related to the data collection process. Furthermore, published researches give the reporter more information about the research problem. Secondary data allows the researcher to benefit from the results of previous studies and access the data that may was impossible for the researcher to collect easily such as number of diabetics in Oman, complications of diabetes, symptoms of diabetic retinopathy, etc. With the help of secondary data, the researcher can make the primary data more specific as well as the researcher was also able to make out what are the deficiencies and gaps and the further data that need to be gathered.

3.1.3 Instruments for Data Collection

Data collection is an important component for conducting a research and which refers to process of collecting data from all the related resources to find the answers of the research problem. Data Collection is a critical aspect of any type of research study. Inaccurate data collection could influence the findings of the research and lead to invalid findings (Unicef-irc.org, 2019). As mentioned above, the researcher used mixed methods of data collection to collect more comprehensive data. The following methods are used in this research:

- **Questionnaire:** Questionnaire is a research instrument for gathering data which consist a series of questions that related to a research problem for the purpose of collecting data from respondents. This an instrument of data collection allows the researcher to collect data from a large audience. For this research, the researcher used quantitative questionnaire that contain closed-ended questions with multiple choice about diabetes mellites and diabetic retinopathy. The researcher searched for diabetes people from his society and mailed the questionnaire through WhatsApp to them because it is faster and consumed low cost. The results of questionnaire of this research analyzed using quantitative instruments such as bar-charts, pie-charts, and percentages.

- **Interview**: Interview is qualitative method of data collection and is the most common data collection method used in qualitative healthcare research. This method is the most appropriate where detailed insights are needed from individual participants. For this research, the researcher conducted a face to face interview with ophthalmologist by using a semi-structured interview that contain of several key questions that assist the

researcher to define the fields to be explored. A complete description of a research problem is offered to the researcher through interview.

3.2 Software Development Methodology

Software development methodology is a framework of software engineering that follows a development life cycle approach (SDLC) and involves different processes for developing an information system. Developers should choose an appropriate software methodology to build the system based on the aims and objectives of the project (S, 2017). There are various software development methodologies as shown below in table (2):

	Pros	Cons
Waterfall	▲ Easy to understand and functional Simple enough to handle as model is rigid ▲ Saves significant amount of time ▲ Allows for easy testing and analysis	▲ Only matches precise needs ▲ Not applicable for maintenance projects ▲ No option to know possible outcome of a project ▲ Not excellent for long and ongoing projects
Agile	▲ Adaptive approach that responds to changes favorably ▲ Allows for direct communication to maintain transparency ▲ Improved quality by finding and fixing defects quickly and identifying expectation mismatches early.	▲ Focuses on working with software and lacks documentation efficiency ▲ Chances of getting off-track as outcome are not clear
Prototype	▲ Gives clear idea about the functional process of the software ▲ Reduces the risk of failure in a software functionality ▲ Assists well in requirement gathering and the overall analysis	▲ Chances of extension in management cost ▲ Excessive involvement of client can affect processing ▲ Too many changes affect the workflow of the software
Spiral model	▲ Risk factors are considerably reduced ▲ Excellent for large and complex projects ▲ Allows for additional functionality later ▲ Suitable for highly risky projects with varied business needs	▲ Costly model in software development ▲ Failure in risk analysis phase may damage the whole project ▲ Not appropriate for low-risk projects ▲ Might get continued and never finish

Table 3: Software Development Methodologies

3.2.1 Selected Software Development Methodology

Based on the above table that describe the various software methodologies, the suitable methodology for "Automated Detection and Classification of Diabetic Retinopathy using Deep Learning" project was Agile methodology. The researcher used agile methodology because it uses the concept of constant feedback and iteration in order to improve the system

that is under development. Agile methodology is considered as unstructured model and its process is divided into individual models.

Justification for using Agile Development Methodology

The project is developed using the concept of data science "deep learning technique", agile software development methodology made the project easier and organized as the researcher could use cycles. With each cycle, the researcher could learn something new and were able to iterate back to improve the results. Moreover, one of the core characteristics of agile development is iterating continuously. Deep learning project required multiple iterations before reaching final results and for developing prediction model, the researcher found that the agile methodology is best methodology for his project. Also, deep learning project has multiple outputs before reach any conclusion and the researcher should share these intermediate outputs such as sensitivity, specificity and accuracy of training stage, so agile methodology allows the researcher to share these outputs. Data science project need to define objectives and follow critical path to achieve them and using agile methodologies would be the appropriate for that.

3.2.2 Agile Development Methodology Process

"Automated Detection & Classification of Diabetic Retinopathy using Deep Learning" project is developed in a cyclic iterative loop that leading to continuous data modeling, preparation, and evaluation. This approach is started by an iterative loop between business-understanding and data-understanding, and then move to an iterative loop between data-preparation and data-modeling, which therefore shift to an evaluation stage, which divides its results to deployment stage and return to the business-understanding. Phases of agile methodology framework are illustrated in figure 2 and described in detail as below:

1. **Business-understanding phase** includes understanding rules and business objectives of the project and creating a project plan.
2. **Data-understanding phase** involves collecting data with high quality as well as exploring it.
3. **Data-preparation phase** involves data preprocessing to import it into training and testing models.
4. **Modeling phase** involves selecting an appropriate DL technique to be used for the project.
5. **Evaluation phase** reviews and evaluate the process based on the business objectives. Evaluate if the model achieved project objectives or need further data.
6. **Deployment phase** presents the results in a convenient format to the users.

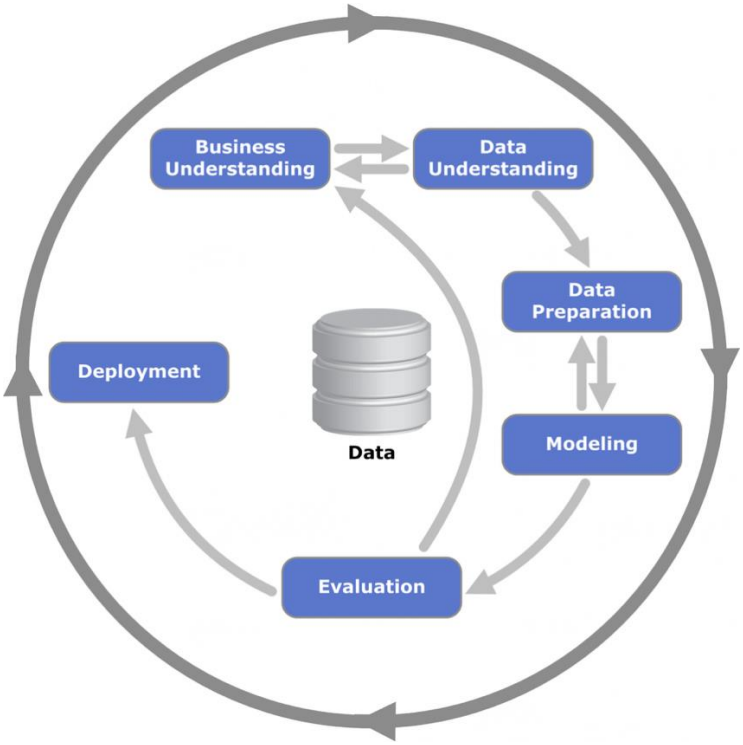


Figure 4: Agile Development Methodology Process

4 Chapter 4: Project Management

Project management (PM) is the main objective to guarantee that the final project is success. PM refers generally to the practice of starting, planning, executing, monitoring, and closing the work to achieve project objectives at the specified timeline. The project becomes successful when all project objectives have been followed and achieved within time frame. For this project, an agile project management has been selected and which is an iterative process that emphasis on the incessant monitoring and improvement of deliverables. This chapter describes different components of project management such as tasks schedule, work breakdown structure (WBS), milestone, deliverables, project management risks, etc.

4.1 Project Phases

During this project, different phases is determined by the project developer to complete the project. In the concept of project management, the project is organized into five phases as the following:

- Initiation Phase: This is the first phase of the project management lifecycle. At this stage, the project is defined at a broad level and initiated. During this phase, project objective is defined, and the project charter is completed.

- Planning Phase: After the project is initiated, it needs a plan for scope, duration, resources, risk, etc. to guide the project team, as well as assists them understand the scope and timeline of the project

- Execution Phase: This phase takes the longest amount of time of overall project time line as well as require more efforts than other phases. At this stage, acquiring all essential resources that required to accomplish the project. During this phase, deliverables of the project is developed and completed.

- Monitoring and Control Phase: This phase is occurred at the same time of Execution Phase. During this phase, the project team should monitor tasks for scope creep and calculate the key performance indicators (KPI) to prevent project failure.

- Closure Phase: At the final stage, the project is formally closed, and the final deliverables are released as well as hand over the project documentation.

Figure 5:Project Management Lifecycle

4.2 Project Task list

Project task list is the main base for any project, and which also known as a checklist. Project task list describe the steps that the project team should follow to complete full project. It shows the main tasks as well as the sub-tasks of a project with the timeline. The Table below shows the task list for "Automated Detection and Classification of Diabetic Retinopathy using Deep Learning for Smart Health in Oman" project with time frame.

Nr.	PSP	Task	Duration	Start	End	Predecessor
1	1	Automated Detection and Classification of Diabetic Rtinopathy using Deep Learning for Smart Health in Oman	169	18/03/19	03/09/19	
2	1.1	Initiating	29	18/03/19	16/04/19	
3	1.1.1	Develop Project Charter	11	19/03/19	30/03/19	
4	1.1.1.1	Identify Goals and Objectives	5	20/03/19	25/03/19	
5	1.1.1.2	Develop Strategies and Plans	13	26/03/19	08/04/19	
6	1.1.1.3	Research Previous Experience	10	19/03/19	29/03/19	
7	1.1.1.4	Develop Project Charter	10	20/03/19	30/03/19	
8	1.1.2	Develop Preliminary Project Scope Statement	16	31/03/19	16/04/19	
9	1.1.2.1	Develop High Level Work Breakdown Structure	7	31/03/19	07/04/19	4
10	1.1.2.2	Conduct Peer Review	15	01/04/19	16/04/19	4
11	1.1.2.3	Prepare Preliminary Project Scope Statement	0	16/04/19	16/04/19	7
12	1.2	Planning	36	16/04/19	22/05/19	11
13	1.2.1	Define Scope	5	16/04/19	21/04/19	
14	1.2.1.1	Document Scope Management Plan	4	16/04/19	20/04/19	
15	1.2.1.2	Specify Deliverables and Acceptance Criteria	2	17/04/19	19/04/19	
16	1.2.1.3	Define Scope	0	21/04/19	21/04/19	
17	1.2.2	Develop Project Schedule	21	19/04/19	10/05/19	
18	1.2.2.1	Build Work Breakdown Structure	3	19/04/19	22/04/19	16
19	1.2.2.2	Develop Resource Plans	10	20/04/19	30/04/19	15
20	1.2.2.3	Prepare Project Estimates	11	28/04/19	09/05/19	
21	1.2.2.4	Define Dependencies and Develop Project Schedule	9	01/05/19	10/05/19	14
24	1.2.3	Develop Risk Plans	12	10/05/19	22/05/19	21
25	1.2.3.1	Document Risk Management Plan	4	10/05/19	14/05/19	14
26	1.2.3.2	Identify Risks	9	11/05/19	20/05/19	19
27	1.2.3.3	Analyze Risks	12	09/05/19	21/05/19	
28	1.2.3.4	Document Risk Management Plans	0	22/05/19	22/05/19	27
29	1.3	Design & Coding	49	22/05/19	10/07/19	
30	1.3.1	Design Framework	16	22/05/19	07/06/19	18
31	1.3.1.1	Define framework stages and activities	9	22/05/19	31/05/19	
32	1.3.1.2	Design framework content formats	13	25/05/19	07/06/19	
33	1.3.2	Build the Framework	42	27/05/19	08/07/19	32
34	1.3.2.1	Train DL model	14	27/05/19	10/06/19	32
35	1.3.2.2	Test DL model	25	10/06/19	05/07/19	34
36	1.3.2.3	Validate DL model	37	03/06/19	10/07/19	35
37	1.3.3	Assure Quality	45	26/05/19	10/07/19	
38	1.3.3.3	Conduct Project Reviews	45	26/05/19	10/07/19	
39	1.3.3.4	Facilitate Continuous Improvement	46	25/05/19	10/07/19	
40	1.4	Monitoring and Controlling	54	23/06/19	16/08/19	
41	1.4.1	Monitor and Control Project Work	43	23/06/19	05/08/19	
42	1.4.1.1	Manage Action Items	28	23/06/19	21/07/19	27
43	1.4.1.2	Manage Project Records	41	25/06/19	05/08/19	26
44	1.4.2	Schedule Control	60	24/06/19	23/08/19	
45	1.4.2.1	Track Status	37	24/06/19	31/07/19	
46	1.4.2.2	Maintain Project Schedule	30	01/07/19	31/07/19	36
47	1.4.2.3	Maintain Work Plans	47	30/06/19	16/08/19	
48	1.5	Closing	17	17/08/19	03/09/19	47
49	1.5.1	Close Project	17	17/08/19	03/09/19	
50	1.5.1.1	Conduct Post-Project Review	11	17/08/19	28/08/19	47
51	1.5.1.2	Close Out the Project Records	5	29/08/19	03/09/19	50

Table 4: Project Schedule

4.3 Work Breakdown Structure (WBS)

A work breakdown structure (WBS) is a high-level method that shows the critical work tasks of a project in a graphical chart called hierarchical tree structure. WBS divide individual tasks into subtasks and organizes the members' work into manageable sections. The figure below shows the hierarchical tree structure (WBS) of the project.

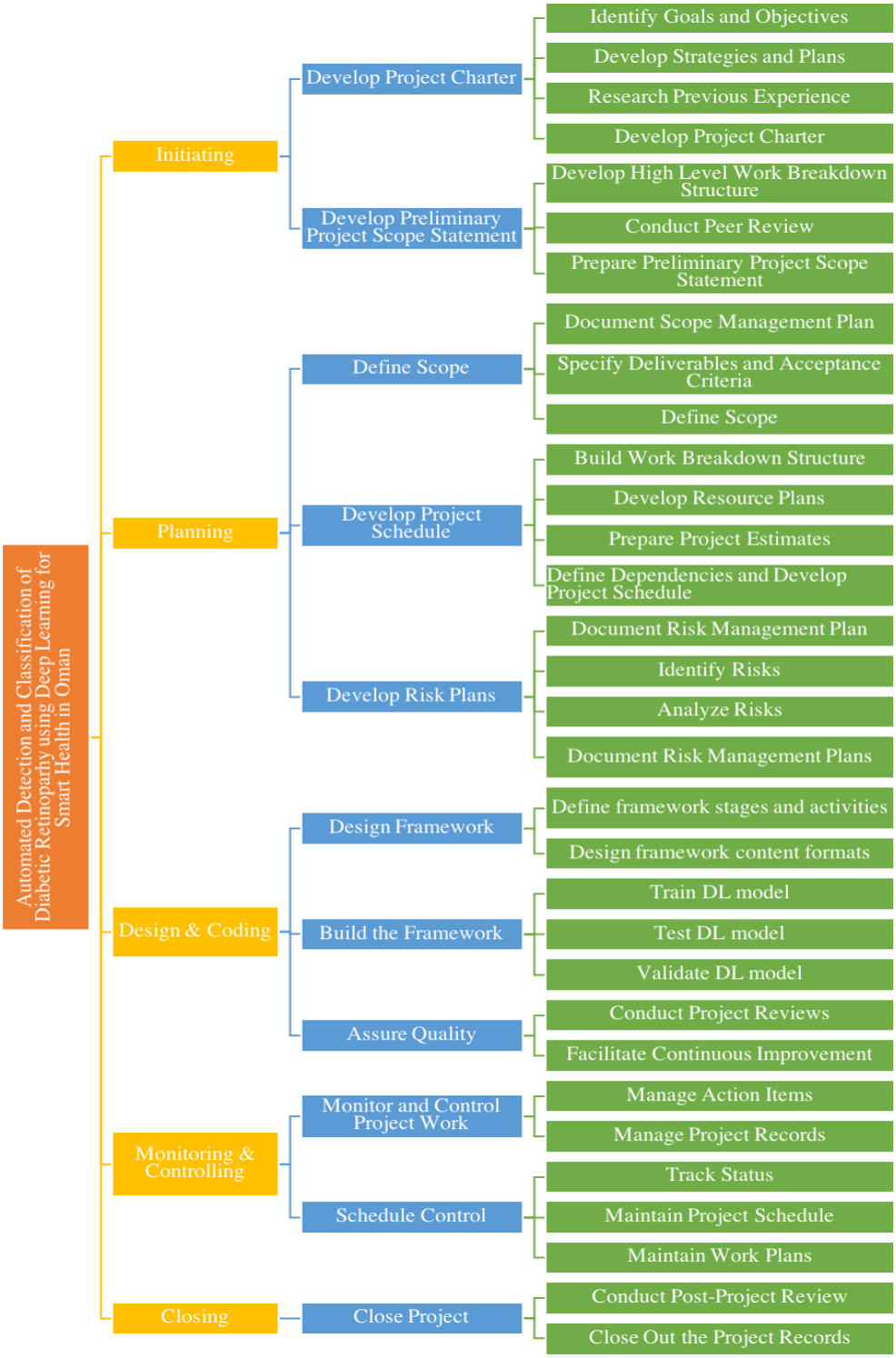


Figure 6: A work breakdown structure (WBS)

4.4 Project Management Risks

Throughout the life of the project, the project team should ensure that risks are actively identified as early as possible in the project so as to reduce their influence.

4.4.1 Project Risks

The following table describe the risks for the project:

Risks	Description	Occurrence of Risk	Impact
Project Schedule	The project tasks may take longer time than expected. Delayed in the completion of the project tasks typically effect the project benefits and might loss of its competitive advantage.	low	High
Project Performance	The final project may fail to produce the expected results such as sensitivity, specificity, & accuracy.	Medium	Medium
Data availability	The collected data may not enough to train and test deep learning model to achieve high accuracy.	low	High
High cost	The cost of resources that are required for this project is high such as laptop.	Medium	low
Acquiring real data images with labels	Collecting real fundus images from a leading hospital in Oman for this project is too difficult.	High	Low
Version Compatibility	While running the DL model, many errors occurred because sometimes the version of anaconda and TensorFlow do not compatible with python packages version.	High	High
Device Storage	To get high accuracy, the project task requires to train more images, but the model cannot train huge number of images because the device storage is not enough.	High	High

Table 5: Project Risks

4.4.2 Mitigation Plan

The table below shows the strategy of mitigation for the project risks:

Risk	Strategy of Mitigation
Project Schedule	Make a great effort to complete the project on time.

Project Performance	Increase the number of retina images that used to train the model to get high accuracy as well as do not used a complicated algorithm.
Data availability	Collect online dataset from various websites such as Kaggle and Medissor-2.
High cost	Use the available recourses or buy equipment with low cost.
Acquiring real data images with labels	Get No-Objection Certificate from the Hospital to collect real data.
Version Compatibility	Download an appropriate python packages version that suitable for anaconda and TensorFlow version.
Device Storage	Increased the device storage to be able to handle huge number of images.

Table 6: Strategy of Mitigation

4.5 Project Budget

The following project budget plan is used for this project to estimate the total cost of a project includes a detailed estimate of all costs:

Expenditure Description	Budget Requested (OMR)	Justification for Expenditures
1. Equipment		
Laptop (MacBook Pro)	700 OMR	Laptops are required for the project developer to complete the project such build DL algorithms, write the final reports, etc.
Hard Disk	30 OMR	Hard Disk with more than 500 GB is required for project backup
Printer	50 OMR	Printer is needed to print out the final report.
2. Online Courses		
DataCamp Courses	45 OMR	Project developer need an online course to increase his knowledge in artificial intelligent, Machine learning, and Deep learning to complete the project.
3. Software		
MS project	10 OMR	MS project is required to prepare project plan.
Microsoft Office (Word)	40 OMR	To write the final report, the researcher needed for Microsoft Office.
4. Consultant Fee		
Atoms Lab (privet company)	10 OMR	Some consultations are required for project developer to clarify about some topic related to the project.
4. ITA course		

Artificial intelligence (Python & TensorFlow) course	25 OMR	Attend this course to increase awareness about Python & TensorFlow

Table 7: project budget plan

4.6 Project Assumptions

The following assumptions are related to project and which were made during project plan preparation:

- Once new information and issues are exposed, the project plan might change.
- The project developer will make huge effort to complete project goals and objectives.
- The project developer will stand by the guidelines identified within the project plan.
- Project will be completed within the time specified in the project timeline.
- The project developer will always available to complete project tasks and objectives.
- Collect enough fundus images to train and test model and consequently get high accuracy.

4.7 Gantt Chart for The Research Project

The figure below shows the timeline for overall research project:

Task	Mar		Apr				May				Jun				Jul				Aug				Sep	
No. Weeks	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4	1	2
Select Project Title																								
Prepare Project Proposal																								
Chapter 1 Introduction																								
Chapter 2 Literature Review																								
Chapter 3 Methodology																								

Chapter 4 Requirement Gathering & Analysis																						
Chapter 5 Project Management																						
Chapter 6 Design & Implementation																						
Chapter 7 Discussion & Results																						
Chapter 8 Conclusion																						
Review Final Report & Submit																						
Final Presentation																						

Figure 7: Gantt Chart for The Research Project

5 Chapter 5: Requirement Gathering & Analysis

This chapter contains description of data gathering methods that used for this research as well as the analysis of data that were gathered.

5.1 Study Population

Due the increasing prevalence of diabetes and diabetic retinopathy in Sultanate of Oman, this research emphasis on the diabetes patients in Oman with different age group of people.

5.2 Data Collection

To achieve the research objectives and response to the research questions, different data collection methods used which deals with two types of data source: primary and secondary. The task of collecting data took place after identifying the research problem evidently.

5.2.1 Collecting Data Procedure

Questionnaires and interview were used to collect primary data, while the secondary data was collected from literature review and online datasets.

5.2.1.1 Primary Data collection through questionnaire

The questionnaire is one of the most prominent tools used in scientific research. The questionnaire consists of a variety of questions, which are related to each other in a way that achieves the goal sought by the researcher through the problem posed by his research (Research-Methodology, 2019). For this research purpose, set of questions has been prepared related to the research topic to collect primary data from respondents using SurveyMonkey (Surveymonkey.com, 2019). Because to free account, SurveyMonkey allowed the researcher to prepare ten questions only. The questionnaire prepared for diabetes patients to obtain statistically information about diabetes mellitus and diabetic retinopathy. The researcher prepared structured questionnaire which used only close ended questions that does not require much effort to show the result. The questionnaire was sent to 50 respondents (diabetes patients) through WhatsApp. The results of the questionnaire were analyzed as below:

Table 8 shows the personal profile of the participants that were involved in the questionnaire. Out of 50 participants, 26 were males and 24 were females. Most of the participants were between the age group of 25 -64 years.

Table 8: Personal profile of the participants

Participants profile	Categories	Percentage/number
Gender?		
	Male	26%

	Female	24%
Age?		
	Under 18	2%
	18-24	2%
	25-34	18%
	35-44	10%
	45-54	18%
	55-64	26%
	65+	24%

Table 9 represents participants information about diabetes and diabetic retinopathy. 26 % of total participants have type 1 diabetes and 32 % have type 2 diabetes. Whereas 42 % of total participants don't know which type of diabetes they have. The duration of diabetes mellitus in these participants ranged from 11 to 15 years. 93 % of total participants have a family history of diabetes mellitus. The analysis showed that 38 % of participants have diabetic retinopathy. 22% of total participants undergo their eye checkup every three months and 16% participants every six months. Whereas 62% of participants undergo their eye checkup yearly and that means they don't have diabetic retinopathy.

Table 9: Diabetic Profile & DR Related Information of Participants

profile	Categories	Percentage/number
Type of diabetes mellitus		
	Type 1 Diabetes (T1D)	26%
	Type 2 Diabetes (T2D)	32%
	Don't know	42%
Duration of diabetes mellitus in years		
	0-5	24%
	6-10	24%
	11-15	30%
	16-19	12%
	20-25	6%
	26+	4%
Family history of diabetes mellitus		
	Yes	62%
	No	20%
	I don't know	18%
Complication of diabetes mellitus		
	Diabetic Retinopathy	38%
	Diabetic Nephropathy	8%
	Neuropathy (Nerve Disease)	16%
	Dermopathy	14%
	Diabetic Foot Disease	24%
	Cardiovascular Disease	14%
	Hearing Impairment	10%
	Alzheimer's Disease	0%

	No complication	36%
How often you undergo an eye checkup?		
	Every three months	22%
	Every six months	16%
	Yearly	62%

Table 10 represents knowledge of participants about diabetes mellitus & eye health. 91% of total participants think that good diabetes control may prevent diabetic retinopathy, whereas 4% don't think so and 6% they don't know. A high percentage (80 %) of participants think diabetes can lead to blindness. Only 12 % participants didn't know if diabetes can lead to blindness or not. The analysis showed that 52 % of participants think that diabetic retinopathy is a treatable condition, whereas 14 % participants don't think so.

Table 10: Questions Related to Diabetes Mellitus & Eye Health

Questions Related to Diabetes Mellitus & Eye Health	Categories	Percentage/number
Do you think that good diabetes control may prevent diabetic retinopathy?		
	Yes	91%
	No	4%
	I don't know	6%
Do you think diabetes can lead to blindness?		
	Yes	80%
	No	8%
	I don't know	12%
Do you think diabetic retinopathy is a treatable condition?		
	Yes	52%
	No	14%
	I don't know	34%

5.2.1.2 Primary Data collection through Interview

An interview is an important tool for obtaining information from human resources and is used in various fields such as medicine, journalism, education and staff selection. Interview is defined as a verbal interaction between two people in a confrontational position. One of them, the interviewer, tries to provoke some information or expressions about the other person, which is about his or her views and beliefs (Research-Methodology, 2019). For the purposes of this research, face to face interview was conducted with ophthalmologist. The interview was conducted to collect valuable information about the ophthalmology field, specifically on diabetic retinopathy. The researcher asked ophthalmologist different questions about diabetic retinopathy, and the screening process for diabetic retinopathy.

5.2.1.3 Secondary Data collection through literature review

A literature review is an objective and critical summary that evaluates the existing researches that are related on a topic under consideration for study (Research-Methodology, 2019). For this research, the researcher read the published researches and got a good grasp of diabetes mellitus and diabetic retinopathy. A literature review described what others have published in the form of a set of summaries and critically discussed the results that others found in their researches about sensitivity, specificity, and accuracy of automated detection of DR. Moreover, literature review provides a comparative table of the performance of deep learning techniques that were used by the other researches to detect DR. The researcher emphasis on the journal papers that are recently published. The main keywords used to search for journal papers are diabetes mellitus, diabetic retinopathy, deep learning, machine learning, CNN, fundus image, etc.

5.2.1.4 Secondary Data collection through online datasets (fundus image)

The most significant data for this research is fundus images which collected from online sources. Kaggle.com is one of the most popular online community of machine learners and data scientists which allow users to find image datasets, CSVs, financial time-series, movie reviews, etc. (Kaggle.com, 2019). For this research study, the researcher downloads dataset of high-resolution color retina images and CSVs file. The dataset was released in the 2015 diabetic retinopathy detection competition (https://www.kaggle.com/c/diabetic-retinopathy-detection). The image dataset divided into two folders: one for training (8,409 images) and other for testing (11,381 images). The dataset contains left and right retina images for every patient (e.g. right.jpeg means the right eye of patient). Training dataset contains normal images and images of different grades of diabetic retinopathy including NPDR (grade-1, grade-2, grade-3) and PDR (grade-4). While the CSVs file is train labels that contain the labels for those images in training dataset as shown in figure 8.

Table 11: DR Grades Details

DR Grade	Description
Normal	Diabetic retinopathy absent
Mild-NPDR	Microaneurysms only
Moderate-NPDR	1. Microaneurysms 2. One or more of exudates
Severe-NPDR	Any of the following: 1. Microaneurysms and exudates 2. Intra retinal micro vascular abnormalities 3. Venous beading
PDR	Either or both of the following:

	1. New vessels elsewhere 2. New vessels on the optic 3. Neovascularization

A1 | image

	A	B	C
1	image	level	
2	10_left	0	
3	10_right	0	
4	13_left	0	
5	13_right	0	
6	15_left	1	
7	15_right	2	
8	16_left	4	
9	16_right	4	
10	17_left	0	
11	17_right	1	
12	19_left	0	
13	19_right	0	
14	20_left	0	
15	20_right	0	
16	21_left	0	
17	21_right	0	
18	22_left	0	
19	22_right	0	

Figure 8: CSVs file

6 Chapter 6: Design and Implementation

In this chapter, first an introduction is given to diabetic retinopathy as well as its impact. Then the proposed methodology framework is explained that implemented in neural network, by building a CNN model to train the system to detect and classify diabetic retinopathy.

6.1 Diabetic Retinopathy (DR)

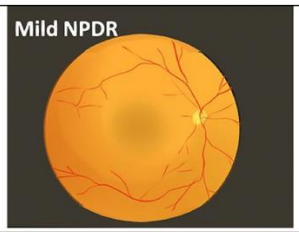
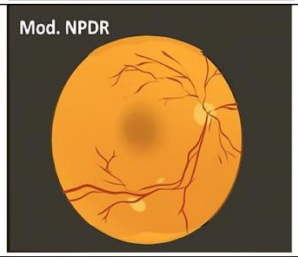
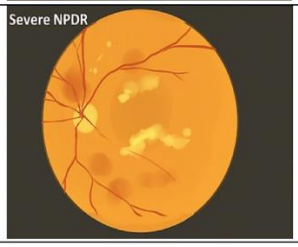
Diabetic retinopathy is one of the eye complications caused by diabetes which occurs when diabetes affects the small blood veins and therefore affect vision. This disease occurs due to damage of blood vessels in the light-sensitive tissues in the back of the eye (retina). This can happen to anyone with type 1 diabetes or type 2 diabetes. The greater duration of diabetes, the likelihood of complications in the eye. At early stages, diabetic retinopathy symptoms may not appear clearly and only causes minor vision problems, but it can eventually lead to blindness. As the condition worsens, symptoms of diabetic retinopathy may include pockets or dark lines floating in the vision, blurred vision, vision is volatile, vision is weak for colors, dark or empty areas in vision, loss of sight. Diabetic retinopathy often affects both eyes (Nei.nih.gov, 2019).

High blood sugar can lead to a blockage in the tiny blood vessels that nourish the retina. In response, the eye tries to create new blood vessels, but these vessels do not grow properly and can leak easily. There are two types of diabetic retinopathy:

I. Early diabetic retinopathy

This type is more common and is known as non-proliferative diabetic retinopathy (NPDR). When infected with this type, do not grow new blood vessels but the walls of blood vessels of the retina weaken and cause leakage of blood and fluid inside the retina (Wu, 2013). NPDR may progress through three stages as shown in table 12:

Table 12: Non-Proliferative Diabetic Retinopathy Stages

NPDR stage	Features	Image
Mild NPDR	Small areas of balloon-like swelling in the retina's tiny blood vessels, called microaneurysms, occur at this earliest stage of the disease. These microaneurysms may leak fluid into the retina.	
Moderate NPDR	As the disease progresses, blood vessels that nourish the retina may swell and distort. They may also lose their ability to transport blood. Both conditions cause characteristic changes to the appearance of the retina and may contribute to DME.	
Severe NPDR	Many more blood vessels are blocked, depriving blood supply to areas of the retina. These areas secrete growth factors that signal the retina to grow new blood vessels.	

II. Advanced diabetic retinopathy

Diabetic retinopathy can develop into this advanced and severe type of disease known as proliferative retinopathy (PDR). When this type occurs, the damaged blood vessels close which then causing the growth of new and abnormal blood vessels within the retina and can leak into the transparent gel that fills the middle of the eye (the glass body).

The tissue damage caused by the growth of the new blood vessels can cause detach the retina from the back of the eye. If the blood vessels overlap with the normal flow of fluids outside the eye, the pressure can accumulate inside the eyeball (Wu, 2013). This can cause damage the nerves that transmits the images from the eye to the brain (optic nerve). The figure 9 is an image that shows PDR symptoms.

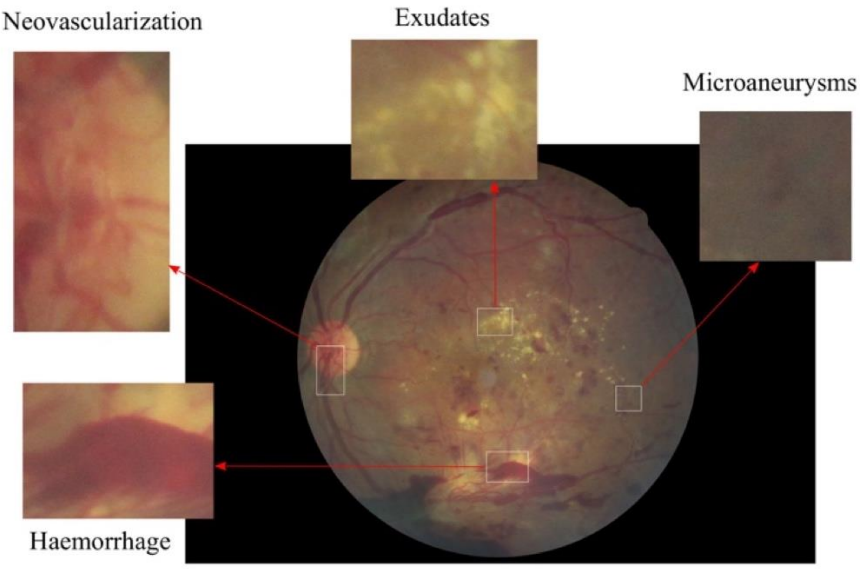


Figure 9: Example image of Proliferative Retinopathy (PDR)

III. Diabetic macular edema (DME)

Diabetic macular edema (DME) occurs when fluid accumulation in an area of macula because of leaking blood vessels. Macular is an important part of the retina that controls the most abilities of detailed vision. Diabetic retinopathy is a disease that destroys small blood vessels in the retina, leading to impaired vision. If these blood vessels are left untreated, they begin to increase the pressure in the eye and leak the fluids, causing the DME to occur (Wu, 2013).

6.2 The Importance of Deep Learning

As it has already mentioned above, diabetic retinopathy is a disease that affects the eyes of a patient. Mainly, it caused by damage of blood vessels of the light-sensitive tissues at the back of the eye. In many countries, deep learning is one of the crucial methods used for the detection of diabetic retinopathy. It is an advanced learning method that accurately detects diabetic retinopathy. This type of learning method plays an important part to develop a new solution of critical disease like cancer, diabetes, diabetic retinopathy, etc. A study explained that a deep learning advance method used to detect the symptoms of diabetic retinopathy (Abràmoff, et al., 2016). Due to the increasing prevalence of diabetes disease, the demand for diabetic retinopathy detection methods is also increasing.

6.3 The Significant of Early Detection of Diabetic Retinopathy

The cause of diabetic retinopathy is blindness and is one of the most common complications of diabetes mellitus. Therefore, doctors insist on the importance of prevention and early detection of DR. Symptoms of diabetic retinopathy do not appear at early stage until the disease

progresses, and sometimes the condition cannot be treated. Therefore, early detection of diabetic retinopathy helps to obtain effective treatment for minimizing damage to vision as well as early detection of diabetic retinopathy makes avoid blindness for 90% of patients. In Oman, diabetes patients diagnose first manually in the primary health care establishments of the Ministry of Health (MOH). Then, if the patients are affected with diabetic retinopathy, they will transfer them to secondary health care establishment. Sometimes, some DR patients are transferred late to the secondary health care establishment and this leads to patient's condition worsens and likelihood to lose their vision. Moreover, the workload of the ophthalmologist is too big because they follow manual way to diagnose diabetes patients and might take 30 min for each patient. Thus, an urgent need for automated system to assist ophthalmologist in the primary health care to detect and classify diabetic retinopathy.

6.4 Proposed Method

The proposed method makes use of convolutional neural network (CNN) to detect and classify diabetic retinopathy. Two main types of diabetic retinopathy which are non- proliferative diabetic retinopathy (NPDR) and proliferative diabetic retinopathy. The initial stage is data collection which are collected from online public dataset 'Kaggle'. The second step is image preprocessing that performed during training itself. The Feature selection stage is carried out by convolutional layers to extract only the significant features for DR grades. After the model is trained, the saved model will used to predict diabetic retinopathy and classify into different stage. The block diagram for automated detection and the classification of diabetic retinopathy is as shown in the Figure 10.

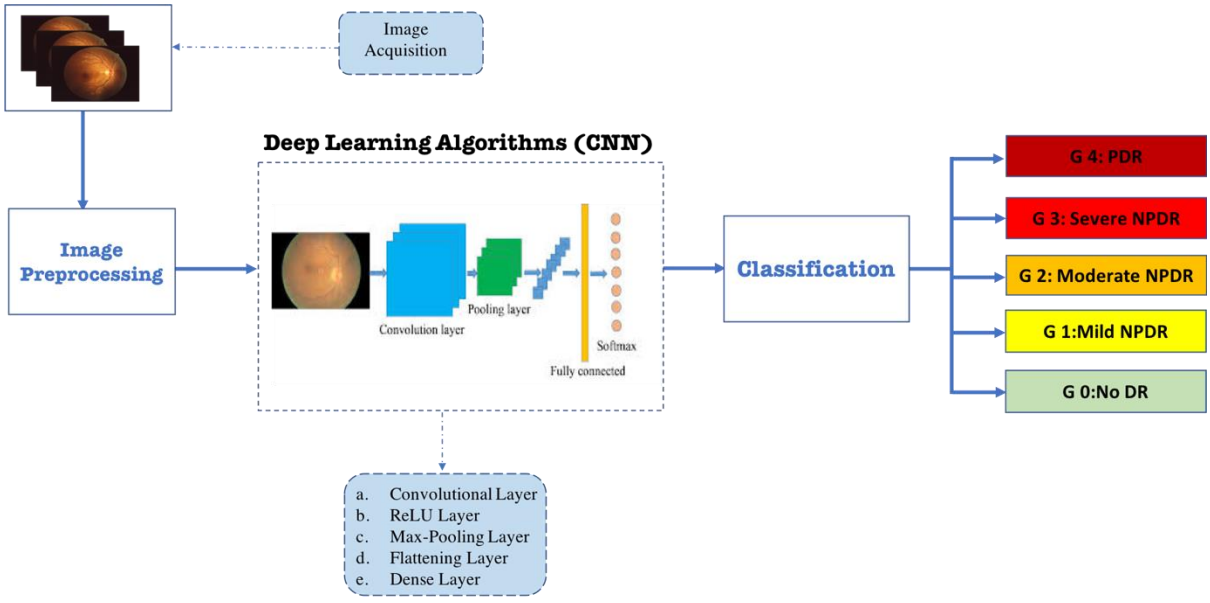


Figure 10: Proposed Method for DR Detection and Classification

6.5 Implementation

6.5.1 6.5.1: Tools and Technologies

The following table shows hardware and software parameters that used for this work:

Table 13: Software & Hardware Parameters

Software & Hardware	Parameters
Operating system	MacBook Pro (13-inch, 2016, Two Thunderbolt 3 ports)
Processor	2 GHz Intel Core i5
Memory	8 GB 1867 MHz LPDDR3
Graphics	Intel Iris Graphics 540 1536 MB
Environment	Keras, TensorFlow
Python	Version 3.7
TensorFlow	Version 1.13.1
Anaconda Navigator	Version 1.9.7
Spyder editor	Version 3.3.5

- **Anaconda Navigator**: An open source innovation designed for Python data-science and machine learning that allows to launch applications to streamline packages management and deployment.
- **Spyder:** an open source and powerful scientific environment designed for and by engineers, scientists, and data-analysts. It provides functionality of a comprehensive development tool and combination of the advanced editing, debugging and analysis. Spyder comes installed with anaconda navigator.
- **TensorFlow:** An open source library for dataflow programming that makes machine learning easier and faster.
- **Keras:** Python library which is on top of TensorFlow machine learning framework that developed by Google in collaboration with Brain Team. Keras allows users for fast and easy prototyping and supports convolutional-networks and recurrent-networks as well as combinations of both.
- **Convolutional Neural Network**: A Convolutional Neural Network (CNN/ ConvNet) is a branch of deep learning algorithm and is a form of Feedforward-Neural-Networks. The figure 11 shows a typical schema of CNN. CNN used for image analysis and interpretation, including medical imaging. The first part is feature extractor which involves of convolutional and max-pooling layers, while the second part of CNN is the classifier (fully connected layer) that performs non-linear transformations of the extracted features (Pratt et al., 2016).

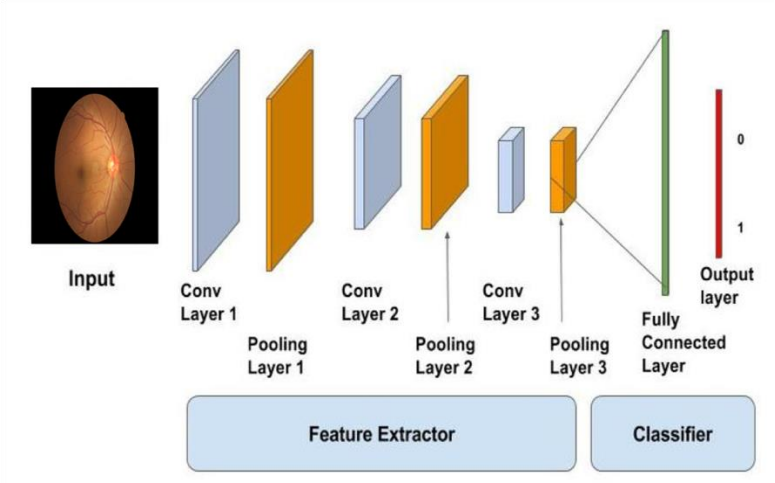


Figure 11: Convolutional Neural Network schema

The above figure illustrates the basic CNN architecture which designed as a series of stages. The training formula in CNN is sequential process that involves various iterations to enhance the parameters and extract distinguishing characteristics from input images. The first three training phases consist of convolutional layers and pooling layers. The output layer consists of three fully-connected layers, while and the last layer is the softmax function that showing whether there is diabetic retinopathy or not.

6.5.2 The Action Plan

- **Phase 1 - Dataset Collection:** The initial stage of the proposed method is data collection. The retina dataset that used for this research is provided by Kaggle diabetic retinopathy detection competition that was released in the 2015 for diabetic retinopathy detection competition (https://www.kaggle.com/c/diabetic-retinopathy-detection). Kaggle dataset holds 88,702 high-resolution retina images that captured under a variety of imaging conditions. The dataset also contains some poor-quality retinal images which make it difficult for algorithm to classify the DR accurately as illustrated in Figure 12.

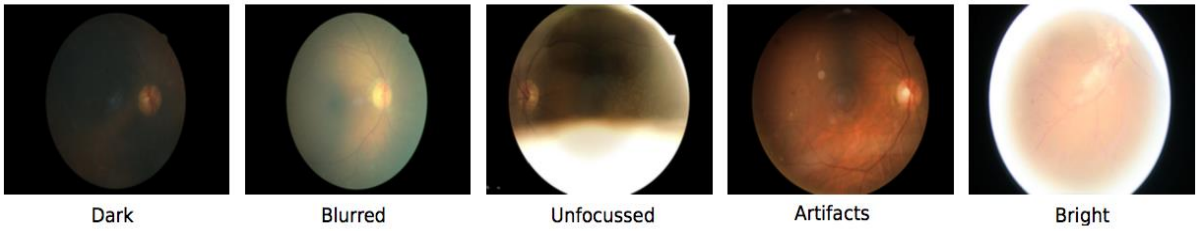


Figure 12: Some Poor-Quality Retinal Images of Kaggle Dataset

Dataset consists high-resolution color retina images which labelled into five classes (No DR, Mild, Moderate, Severe, Proliferative DR). The images were graded by ophthalmologists based on standard severity scale. The left and right images were reported for each patient. The dataset is split into training set with 1000 images and testing set with

300 images. To improve the training accuracy, the images resized into 512*512*3. Table 14 illustrates the distribution between the different DR grades in the training set of the dataset.

DR grade	Grade name	No. image	Percentage
G-0	No DR	722	72.2%
G-1	Mild	69	6.9%
G-2	Moderate	164	16.4%
G-3	severe	28	2.8%
G-4	PDR	17	1.7%

Table 14: Distribution Between the Different DR Grades in the Training Set

- **Phase 2 - Import Required Libraries and Split the Dataset:** To build CNN models for this work, different python packages are downloaded using anaconda and then imported such as: NumPy for expressing images, matplotlib for graph plotting, Pandas, OpenCV, etc. Kaggle dataset is used for this work and already split into training images and testing images.

```
#import package
import numpy as np # linear algebra
import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)
import os
import random
import sys
import cv2
import matplotlib
from subprocess import check_output

from keras.models import Sequential
from keras.layers import Dense, Conv2D, MaxPooling2D, Dropout, Flatten
from keras.preprocessing.image import ImageDataGenerator, array_to_img, img_to_array, load_img
from keras.optimizers import Adam
from sklearn.model_selection import train_test_split
from keras.utils import to_categorical
```

Figure 13: The imported packages

- **Phase 3 - Image Preprocessing and Augmentation:** Gathering datasets usually follows a number of preprocessing steps before they are feed into the neural network model. This study doesn't perform any manual preprocessing before training the models. The CNN model makes use of Keras preprocessing which is "Keras.preprocessing. image.ImageDataGenerator" that performed during training itself. This includes flipping the image vertically or horizontally, image rotation, shifting the image left / right / top / bottom by some amount, zoom or shear the image etc.

  The dataset provided by Kaggle is extremely unequal in number for five grades of diabetic retinopathy. Most of images in the dataset belongs to the grade 0 (no DR) and only a very few images belong to grade 3 (NPDR- severe) and grade 4 (PDR). Thus, to uphold the equilibrium among images number, the dataset was augmented by performing images rotations.

```
aug = ImageDataGenerator(rotation_range=30, # randomly rotate images in the range (degrees, 0 to 180)
                         width_shift_range=0.1, # randomly shift images horizontally (fraction of total width)
                         height_shift_range=0.1, # randomly shift images vertically (fraction of total height)
                         shear_range=0.2,
                         zoom_range=0.2, # randomly zoom into images
                         horizontal_flip=True, # randomly flip images
                         fill_mode="nearest")
```

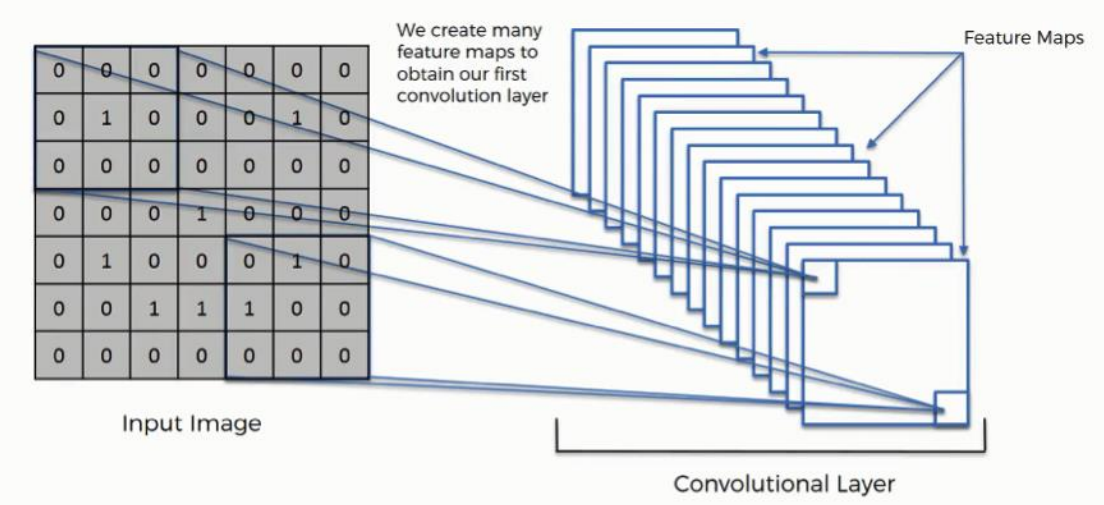
Figure 14: Code for Data Augmentation

- **Phase 4 - Building the CNN Model:** The Sequential model is selected for this study because it's the easiest way to build a model in "Keras" and allows to build the model layer by layer. To implement CNN, convolutional layers were stacked up first and then followed by "max-pooling" layers. The "Dropout" layers were included to avoid overfitting. Then flatten layer added that convert the matrix into a liner array.

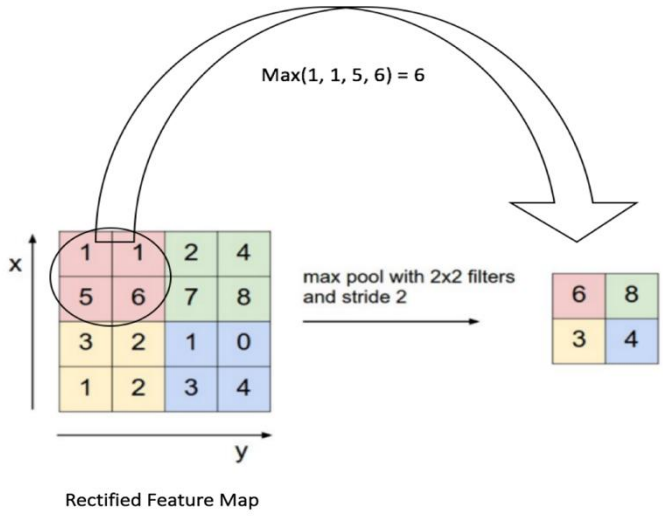
  I. ***Convolution:*** The main aim of convolutional layers is feature extraction from input images. Convolutional layers maintain the spatial relationship between pixels by learning the features of images using small squares of images. Each image considered as a matrix of pixel values. This model consists of three convolution layers (Conv2D) that will deal with input images, which are seen as two-dimensional matrices and gives the responses of that filter at each spatial position. The Convolutional Layer comprises of a set of filters.

  Kernel size is the size of convolutional filters matrix. This work makes use of the popular choice of kernel sizes which is (3*3). As well as, the input shape is specifying by the first layer. The input shape of input images rescaled into (512*512*3).

  The acquired matrix is known also the feature-map. There is also additional operation called Rectified Linear Activation 'ReLU' and used after each convolutional operation to make the neural network easier to train and achieves better performance.

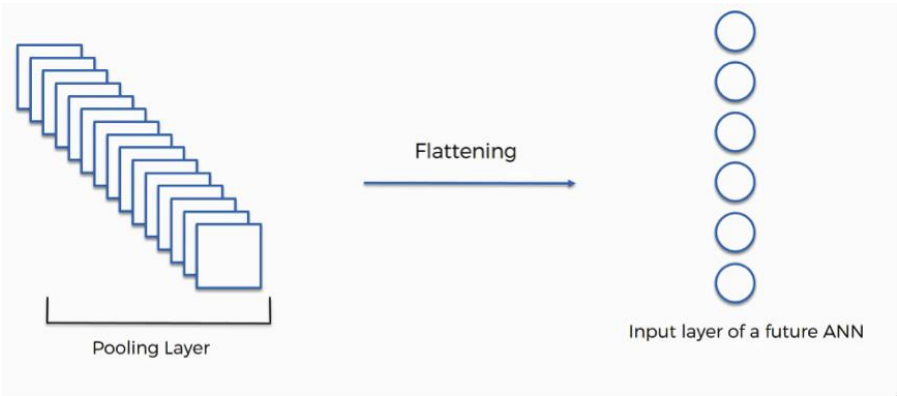


II. ***Polling:*** Pooling layers are also called subsampling and used on activation maps to avoid overfitting. Pooling Layers decreases the dimensionality of every feature map but preserves the most significant data. These layers would often follow Convolutional Layers in architecture of CNN.



III. **Dropouts:** Dropouts is a regularization technique that used in the model to reduce the overfitting in the neural networks and improve generalization of neural network. Dropouts drop randomly the connections between neurons and force the network to discover new paths.

IV. ***Flatting:*** Flatten layer is the last stage of CNN model and it is connected to a fully-connected layer. Flatten layer convert the data into one-dimensional array and create a single long feature vector to input it to the next layer.



- **Phase 5 – Full Connection:** The fully-connected layer 'Dense' is added and followed by 'softmax' function. Dense layer is used for output layer that perform classification on the features extracted by the Conv2D layers and down-sampled by the max-pooling layers.

Every node in the Dense layer is linked to every node in the previous layer. This step use 'Softmax' activation function that turns numbers into probabilities that sum to one. This function outputs a vector and the CNN model makes its "prediction" based on which choice has the highest probability. The following figure illustrates the final network structure that used for this study:

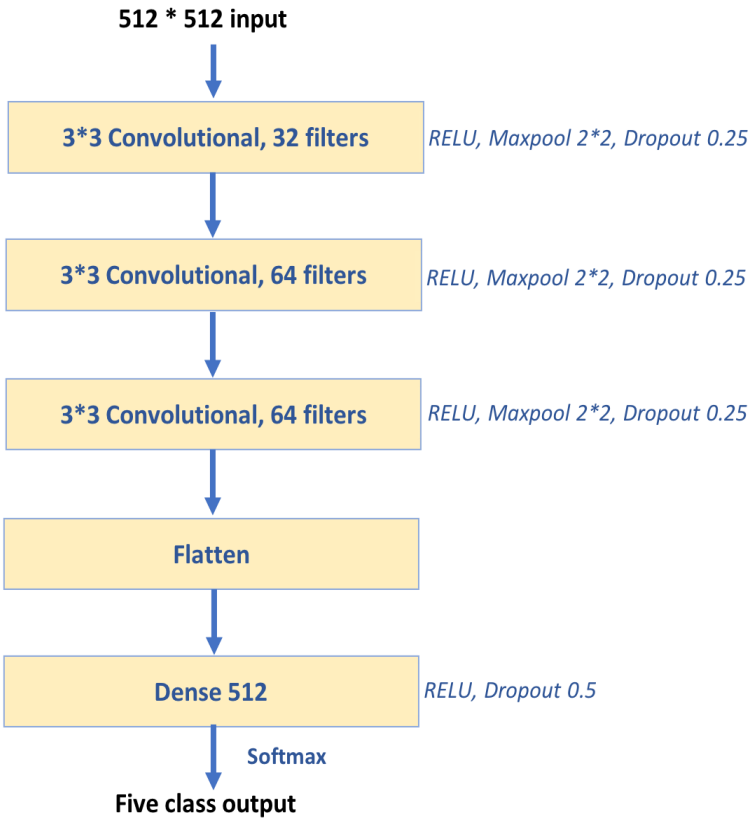


Figure 15: The Final Model Structure

- **Phase 6 - Compiling the Model:** Compiling the model and configure the learning process should take place before training a model via the compile method which expects three parameters: optimizer, loss function, and metrics.

  - Optimizer: The optimizer is one of the two arguments needed to compile Keras model. 'Adam' optimizer which stand for Adaptive Moment Estimation and is used to control the learning rate throughout training. Learning rate defines how quickly calculate the optimal weights for the model. The smaller learning rate might result to more accurate weights, but the time will be longer that taken to calculate the weights. In the compiling model, the optimizer is instantiated before passing it to "model.compile()".
  - Loss function: Compiling model makes use of "binary_crossentropy" loss function. The lower score of loss function indicates that the model is performing better.

- Metrics: The metrics function used to judge the model performance. The 'accuracy' metric is used for this model to show the accuracy score on the validation set when training the model.

```
# returns our fully constructed deep learning + Keras image classifier
opt = Adam(lr=INIT_LR, decay=INIT_LR / EPOCHS)      #lr=INIT_LR, decay=INIT_LR / EPOCHS

model.compile(loss="binary_crossentropy", optimizer=opt, metrics=["accuracy"])
return model
```

Figure 16: Compiling Model

- **Phase 7 - Training the Model:** To train the model, "fit () function" is used with following parameters: training data (Xtrain), target data (trainY), validation data (Xval, valY), and the number of epochs. For validation data, the dataset split into Xval and valY. Epochs refers to the number of times that the model will be cycle through the data. The more epochs are set, the more the model will be enhanced up to reach a certain point. For this model, 15 epochs are set with 64 of batch size.

```
# train the network
print("training network...")
sys.stdout.flush()
#class_mode ='categorical', # 2D one-hot encoded labels
H = model.fit_generator(aug.flow(Xtrain, trainY, batch_size=BS), \
    validation_data=(Xval, valY), \
    steps_per_epoch=len(trainX) // BS, \
    epochs=EPOCHS, verbose=1)
```

Figure 17: Training Stage

- **Phase 8 - Saving Model:** After trained the model, the model is saved into the disk to load it later in order to make predictions as show in figure 18.

```
# save the model to disk
print("Saving model to disk")
sys.stdout.flush()
model.save("/tmp/mymodel")
```

Figure 18: How to save the model

- **Phase 9 -Testing:** Once the network has been trained with enough learning examples, it reaches a point where you present it with an entirely new set of inputs it's never seen before and see how it responds.

  The retinal fundus images are tested on convolutional neural network model that already trained. The test images were resized into the same size as that of the model in which it was

trained with, so that the test can be performed accurately. Then, the fully connected layers are converted into the original convolutional layer filter size and then a class score is assigned to all the input images. With the assist of these class scores will then be able to differentiate between the five different grades of DR namely, Normal (0), Mild-DR (1), Moderate-DR (2), Severe-DR (3), PDR (4).

7 Chapter 7: Experimental Result Analysis

In the previous chapter, the proposed system of this study and implementation of algorithms have been discussed. The previous chapter demonstrated how the dataset collected, description of dataset, and algorithms that used. This chapter discusses about the results that achieved from the experiments upon the implementation of these algorithms. The dataset divided two parts, training and testing dataset. The result of the training and testing dataset will be show in this chapter. As mentioned in the previous chapter that this study makes use of supervised machine learning algorithms. Convolutional neural network method is the selected deep learning approach to achieve research objectives. In this study, the deep learning algorithms was proposed for images classification of diabetic retinopathy. From this study, the following points are observed:

7.1 Performance on Dataset

As mentioned in the previous chapter, this study used retinal fundus from Kaggle dataset. Kaggle dataset comprises more than 80,000 images (Pratt et al., 2016) and this project used 1000 images only for training and 300 images for testing because this project performed in CPU (MacBook pro) and this is a big limitation faced this study. With just 1000 images (750 for training and 250 for validation) the model only achieved 89% of training accuracy. From the plot of accuracy in the following figure 19, the trend for accuracy still rising for the last few epochs and that means if the model trained for more images it will obtain higher accuracy.

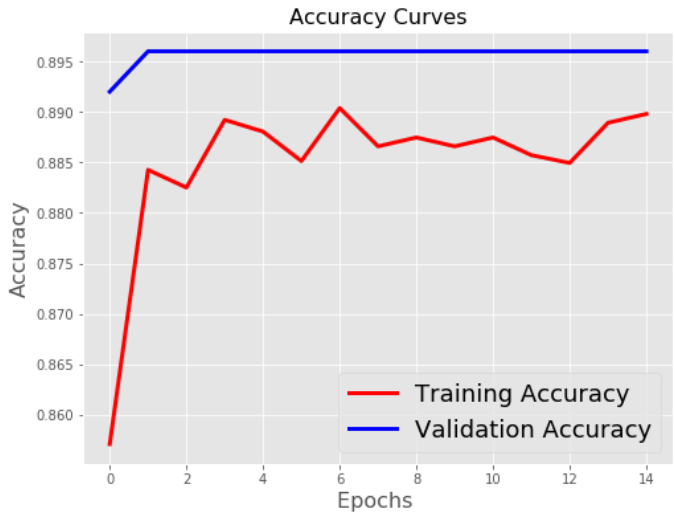


Figure 19: Training Accuracy with 1000 images

7.2 Performance on Image Preprocessing

Image pre-processing is used commonly in the medical image and is essential for feature extraction in classification of diabetic retinopathy. Retinal fundus images are often deteriorated due to the high noise and different sources of interference in the imaging processes. Improving

the appearance and visual quality of images might help in the medical images interpretation and also might improve the diagnostic decision. Preprocessing of the retina image can help to improve detect the features that required and makes feature extraction easier as well as obtained better performance.

As mentioned in the previous chapter, the preprocessing is done during training itself in this work. The experiment results on the Kaggle dataset, CNN model achieved 89% of accuracy with preprocessing of retina images. From the experiment results, can observed that the image preprocessing techniques has a vast influence on DR detection and boost the performance.

7.3 Performance on Tune Parameters

The tune parameters like learning rate, epochs, etc., can improve the performance of CNN model. Huge number of epochs not necessarily improve the model performance. There is no magic rule for how specify the number of epochs. According to many experiments, the project specified only 15 epochs. Figure 20.a illustrates that the model works best with 15 epochs. Whereas in the figure 20.b, the model was trained on 30 epochs, but the last epoch achieved 80% of accuracy. That means, increasing the number of epochs will not always enhance the model performance.

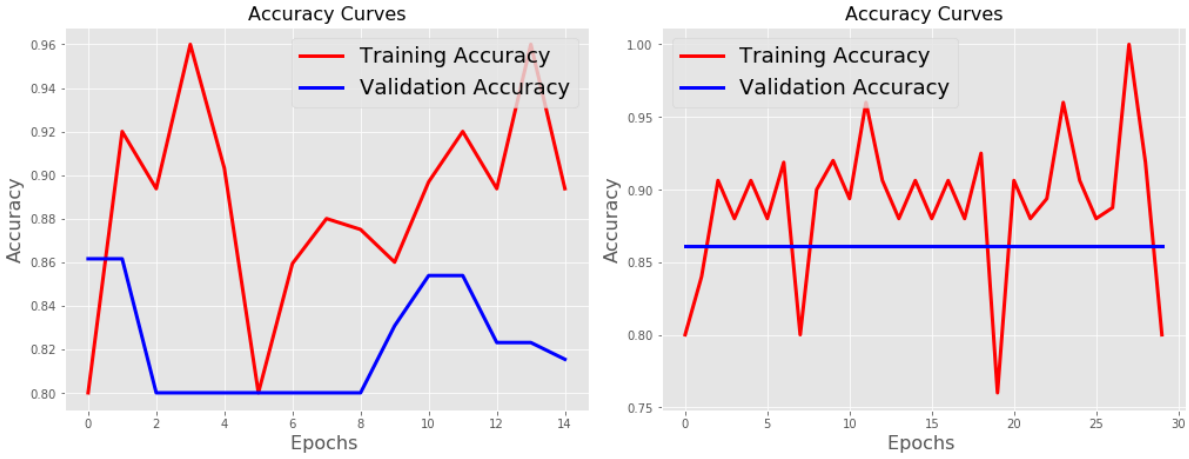


Figure 20: a) Training with 15 epochs, b) training with 30 epochs

Also, the dropout layer in the CNN model improved the performance because it a way to prevent the networks from Overfitting. Moreover, the neural network for this project used Adam optimizer because it has very little memory requirement as well as due it's computationally efficient.

7.4 Performance on Loss Function

The model in this work is compiled using 'binary_crossentropy loss'. This may seem unreasonable because for multiclassification should use 'categorical_crossentropy loss'. Binary_crossentropy loss penalize each output node independently.

One of the experiments carried out by this study using 100 images achieved 90+ of accuracy using 15 epochs and 'binary_crossentropy' loss function. Meanwhile, the same model achieved 85% only by using categorical_crossentropy even using 50 epochs. Moreover, the training loss and validation loss are too high while using categorical_crossentropy. The following figures shows that 'binary_crossentropy' loss function obtained better performance than other loss function.

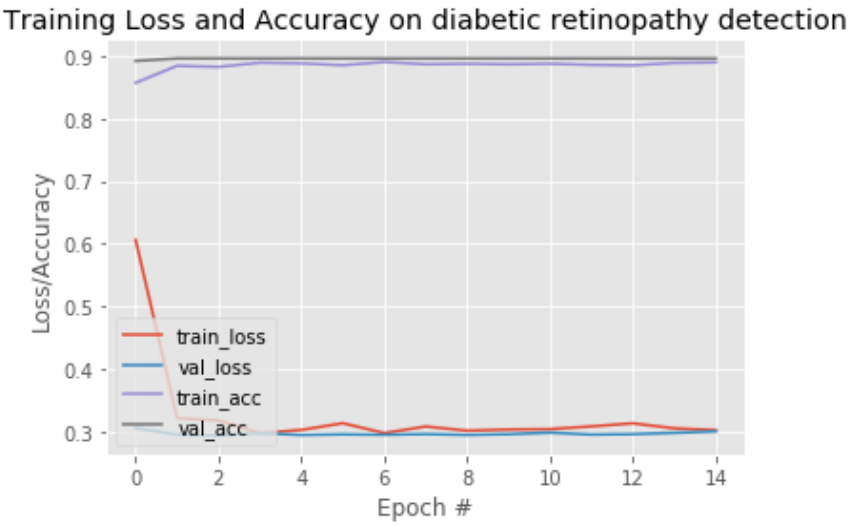


Figure 21: Training with 'binary_crossentropy' loss function

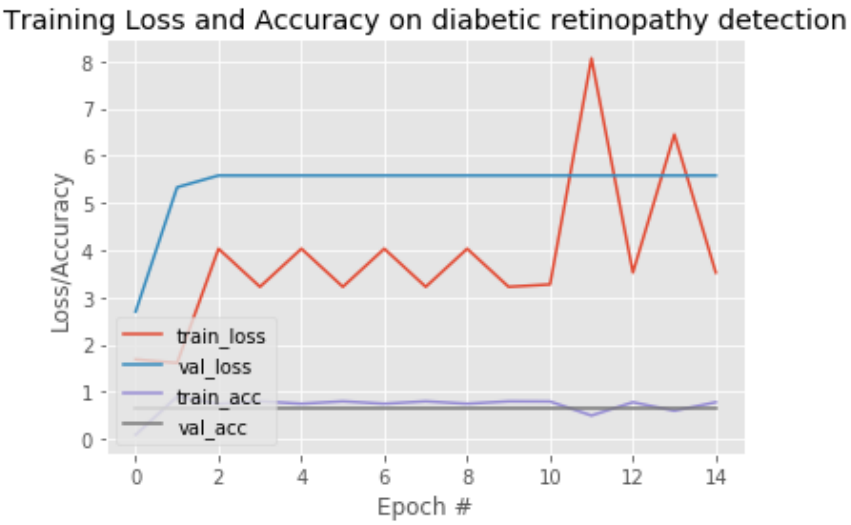


Figure 22: Training with 'gategorical_crossentropy' loss function

7.5 Performance on Data Augmentation

Training network with not enough data is the major reason for Overfitting. However, there is one very effective way to cope Overfitting which is Data Augmentation. Data Augmentation is the process of artificially generating more images from the images that already available by orientation, rotation, changing the size, etc. of the original images.

7.6 Summary

For this work, the supervised deep learning models were built to perform classification of input images. The supervised technique that used for this work allows a system to fetch automatically the representations that required for feature detection from raw data. Convolutional Neural Network (CNN) is used because it is an efficient recognition algorithm for image processing and pattern recognition. Furthermore, CNN is designed for collecting images of two-dimensional and easily manipulate them in search operations. The models were trained to learn the attributes for the retinal fundus images.

The CNN models were initially trained with labelled datasets on 1000 images to reach a significant level. 512*512*3 was the input shape of images. A training dataset was divided into 64 batches. The 'binary_crossentropy' loss function used with 15 epochs to compile the model and the low learning rate of 0.001 was used. This model achieved 89% of training accuracy and 90% of validation accuracy. The validation-accuracy was greater than training-accuracy. This means that the model has been very well circulated. The following figure shows the training loss and accuracy on DR detection using CNN with 1000 images.

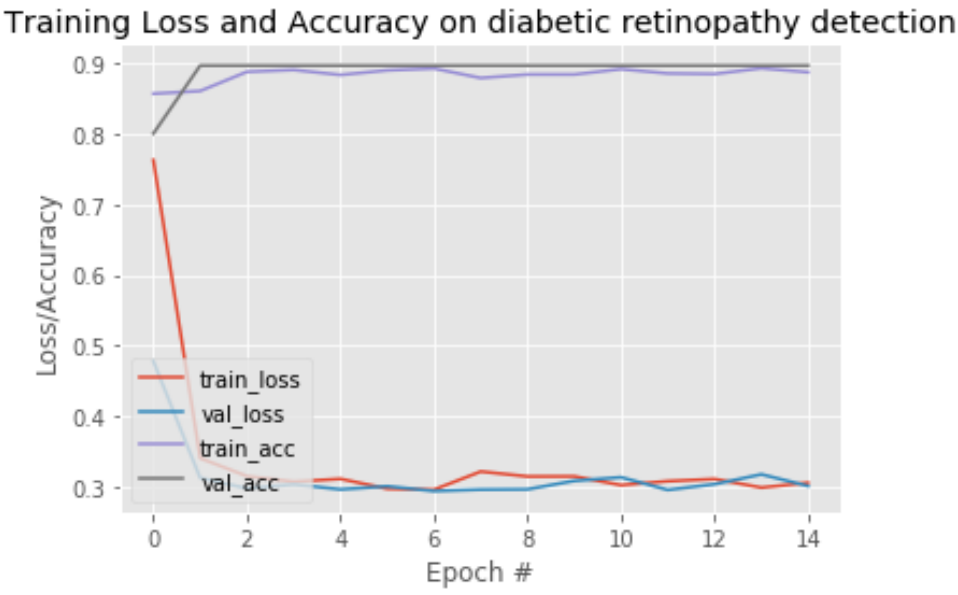


Figure 23: Training Loss and Accuracy with 1000 images

8 Chapter 8: Critical Appraisal

8.1 Introduction

Automatic classification of diabetic retinopathy is an emerging research area aimed at reducing the workload inherent in the traditional grading process. The main objective of this study is to develop a deep learning algorithm to automatically detect and grade diabetic retinopathy using the dataset from the 'Kaggle diabetic retinopathy competition'.

This study makes use of DL method to build the model by using Convolutional Neural Network (CNN). Chapter 2 reviews the findings of previous studies that conducted on automated DR screening. Whereas chapter 6 presented the proposed method that used to achieve research objectives. The CNN model for this work achieved good performance on DR detection. This trained CNN is found to be able to classify thousands of images within a short period of time making them effective in real time.

The output results generated from the research work have achieved the major study objective, where the CNN model developed through this research has proven to be effective in detecting and classifying diabetic retinopathy stages. This project will be able to assist ophthalmologists in clinical diagnosis to detect diabetic retinopathy and therefore effectively avoiding or delaying DR. Furthermore, this study makes a significant contribution to knowledge in the field of medical image processing. This is achieved by proposing a new use of image processing techniques in screening for diabetic retinopathy, as described in chapter 6.

8.2 Research Achievements

The introductory chapter discussed the issues of screening of diabetic retinopathy and highlighted the main objectives of this study. Based on the that objectives, the following represent the main achievements to this research:

8.2.1 Project Achievement

Table 15: Project Achievement

S.NO	Achievement	Status
1	Project proposal (Appendix A)	Completed
2	Mid-term review – project presentation	Completed
3	Literature review (reading many papers related to the research topic and comparing between their results) (chapter 2)	Completed
4	Face to face interview with one ophthalmologist from Al-Nhada Hospital	Completed
5	Analyzing the interview and confirm the outcomes (chapter 5)	Completed
6	Prepare electronic questionnaire and distributed among diabetes patients through WhatsApp (Appendix E)	Completed

7	Analyzing the results of the questionnaire (chapter 5, Appendix E)	Completed
8	Python & TensorFlow Certificate (Appendix D)	Completed

8.2.2 Publication Achievement

Table 16: Publication Achievement

S.NO	Achievement	Status
1	• Paper: Recent innovations in Automated Detection and Classification of Diabetic Retinopathy (Appendix C) • Certificate (Appendix C) Available on : https://www.ijitee.org/download/volume-8-issue-10/	Completed

8.2.3 Development Achievement

Table 17: Development Achievement

S.NO	Achievement	Status
1	Preprocessing retina images	Completed
2	Build CNN model	Completed
3	Train CNN model with Kaggle dataset	Completed
4	Test model	Completed

8.3 Limitations

Just like any automated system, even the powerful deep learning models has some limitations. The basic limitations of this project were hardware specification. A higher GPU-specification is thought to build more accurate CNN models if trained with a larger dataset, but this work used CPU and effect the final results. The model was trained on 512*512*3-pixel size images because the memory size of the graphics card not able to work with a larger pixel size. The accuracy of DR detection might improve if the model trained with larger pixel size images. Likewise, the training dataset is relatively small compared to other studies that have recently conducted on DR screening because as previously mentioned that the memory size of system could not able to train for more images. Furthermore, images were obtained using different hardware, or not present on training phase, might sometimes result in reduced the accuracy. Luckily, this limitation can be circumvented by training on larger datasets or fine-tuning the model on the new data.

8.4 Challenges

During the course of the work, this project faced a number of challenges. Lack of knowledge in deep learning was the biggest challenge that the project faced. Furthermore, the low proficiency in English which negatively affected the performance in writing the research project. On other hand, having the deep understanding of the terminologies of research topic

helped in writing the research project well. Also, project duration that given for project is comparatively not enough because this type of project needs more time.

8.5 Student Reflection

This study which is about "Automated Detection and Classification of Diabetic Retinopathy using Deep Learning" might be the first project conducted in Oman. This study aims to provide automated system that assist the clinicians in the primary healthcare to improve management and clinical diagnostic of diabetic retinopathy. Quantitative and qualitative methodologies were used for this study to gather requirements, whereas Agile methodology was used for project completion. The process of conducting interview and questionnaire and collecting responses provided a good chance to practice interpersonal communication skills. Furthermore, this study provided a good opportunity to improve writing and speaking skills.

For this project, open source software is used that enable of technology development and solving real world problems. Wide range of knowledge was developed from this study regarding diabetic retinopathy, artificial intelligence and deep learning through literature review and self-study of related topics. During the course of work, some challenges were faced such as lack of experience in the field of deep learning, but this challenge has been overcome through joining courses and reading more journal papers about the related topics. Another challenge was faced during the work which is the limited of hardware specification. Deep learning projects always need GPU to train DL models to get accurate results. Last but not least, this project will be the beginning of the journey to develop more projects related to artificial intelligence and deep learning.

9 Chapter 9: Social, Legal, Ethical Considerations & Project Sustainability

This chapter briefly discussed social, legal, ethical considerations and sustainability of this project.

9.1 Social, Legal, Ethical Considerations

As part of DL algorithms development process, they required consideration in a variety of social and ethical issues. Algorithm of the automated detection and classification of diabetic retinopathy developed based on ethical and social, legal considerations. This project doesn't need the research ethics and biosafety approval because it makes use of online datasets "Kaggle", so the project doesn't involve any human participants. As well as, this project will not put anyone at risk of physical harm, or injury. Furthermore, the generated results of this project be seen as not controversial and will not attract the attention of the security services. According to the interview that was conducted with ophthalmologist, this project will be clinically acceptable to help ophthalmologists for detecting diabetic retinopathy in Oman.

9.2 Project Sustainability

Nowadays, artificial intelligence technologies play a critical role in bringing social development to communities by improving management and clinical diagnosis. Though the impact of deep learning methods is usually not directly observed due to its virtual characteristics, there is no doubt that the automated system positively impacts the society in general. For example:

- → Economic impact: This project will help reduce the number of people with the advanced stage of diabetic retinopathy and thus reduce the amount of financial expenses for treatment.
- → Social impact: This project will be able to assist ophthalmologists in clinical diagnosis to detect diabetic retinopathy and therefore effectively avoiding or delaying DR and therefore reduce number of DR patients. Early and accurate detection of diabetic retinopathy is important because of its potential for reducing the number of cases of blindness around Oman society.
- → Environmental impact: This project doesn't have any environmental influence because all the resources that used for this project are safe.
- → Quality and Time impact: Human physicians will not be substituted by machines, but this project can absolutely help physicians to make better clinical decisions. An automated system can also help physicians by providing up-to-date medical information and providing patients the care and treatment in proper time. In addition, this project can assist to reduce diagnostic and therapeutic errors in the clinical practice.

10 Chapter 10: Conclusion and Future Work

This chapter concludes this thesis and summaries the main results and achievements. Moreover, the future plan with thesis project will be provided in this chapter.

10.1 Conclusion

Diabetic retinopathy is one of diabetes complication that damages retinal blood vessels and might leads to blindness. People get affected by diabetic retinopathy after suffering from diabetes mellitus for a long term and eventually lose their vision. This disease is now become common in most of the areas in Oman due to increased prevalence of obesity. The early detection of DR is essential and can avoid vision loss in more than 50% of all cases. AI-based technology has opened the ways to develop more accurate approaches for detecting and classifying the different DR complications.

Nowadays in Oman, diabetes retinopathy different stages are detected by retinal examination using manual process by ophthalmologists. This paper proposed a solution for detection and classification of Diabetic Retinopathy based on deep Learning. Most of existing deep learning algorithms required manual feature extraction stage to identify the different grades of diabetic retinopathy from fundus images. In this study, the proposed solution used Convolutional Neural Network (CNN) that is automatically extracts the features of all grades of diabetic retinopathy. The neural network architecture with dropout function yielded high DR classification accuracy. This study has shown that CNN model is an effective in term for image and has the ability to be trained to detect and classify the different stages of diabetic retinopathy from fundus images. This work has achieved around 89% of accuracy for classifying different grade of diabetic retinopathy with the Kaggle dataset.

Automated detection and classification using deep learning is a promising solution for screening huge number of patients with diabetes. Deep learning has shown impressive results of high accuracy in automated image analysis for DR detection. Artificial intelligence techniques cannot substitute the ophthalmologists' role in management and clinical diagnosis, although would act as a useful assistant in diabetic retinopathy screening and provide diagnostic supports. AI-based technology would help increase referral of patients with retinal diseases to the ophthalmologist for further treatment and therefore would assistance in reducing visual-impairment.

Generally, the results of DL algorithm demonstrate the ability of automated learning systems in simplifying existing DR screening programs in a time-efficient and cost-effective manner. Furthermore, the findings indicate that DL approaches provide powerful tools for refining patients' quality of life and it is suitable method for use in clinical daily practice. Implementing such deep learning algorithm can significantly reduce the rate of sight loss attributed to diabetic retinopathy, create a new diagnostic workflow for disease detection and enhance clinical management.

10.2 Future work

For any work, there is always constantly evolving. This project looks to implement novel techniques for further increasing the accuracy and to enlarge it to other diseases related to eye. Here is shortened list of points that project will take in the future to achieve its objectives:

- → **Hardware Implementation**: This study looking forward to implement CNN model in a hardware system because the hardware product will be the best solution for diabetes patient.
- → **More Datasets:** The CNN model will be trained in more datasets such as DRIVE, Messidor-2 and so on. Also, this project looking forward to use real images data to train the model.
- → **Different Algorithms**: Develop more DL algorithms to detect other eye diseases.
- → **Software Implementation**: Develop mobile app for this purpose. In this way people will be able to upload their data into the machine learning software and will let them know about their disease through mobile app whether it's in a good or bad condition.

11 Reference

Who.int. (2019). [online] Available at: https://www.who.int/blindness/Prevention%20of%20Blindness%20from%20Diabetes%20Mellitus-with-cover-small.pdf [Accessed 9 Jun. 2019].

Cold, F., Health, E., Disease, H., Disease, L., Management, P., Conditions, S., Problems, S., Disorders, S., Checker, S., Interviews, E., Boards, M., Answers, Q., Guide, I., Doctor, F., A-Z, C., A-Z, S., Medications, M., Identifier, P., Interactions, C., Drugs, C., Pregnant, T., Management, D., Obesity, W., Recipes, F., Exercise, F., Beauty, H., Balance, H., Relationships, S., Care, O., Health, W., Health, M., Well, A., Sleep, H., Teens, H., Kids, F., Pregnant, G., Trimester, F., Trimester, S., Trimester, T., Baby, N., Health, C., Vaccines, C., Kids, R., Cats, H., Dogs, H., Risks, H., Do You Really Need 10, 0., Elderly, D., Foods, F., We Eat, 0., Boards, M., Blogs, E. and Center, N. (2019). Types of Diabetes Mellitus. [online] WebMD. Available at: https://www.webmd.com/diabetes/guide/types-of-diabetes-mellitus#1 [Accessed 9 Jun. 2019].

Rada, R. (1986). Artificial intelligence. Artificial Intelligence, 28(1), pp.119-121.

Lu, W., Tong, Y., Yu, Y., Xing, Y., Chen, C. and Shen, Y. (2018). Applications of Artificial Intelligence in Ophthalmology: General Overview. Journal of Ophthalmology, 2018, pp.1-15.

Gulshan, V., Peng, L., Coram, M., Stumpe, M., Wu, D., Narayanaswamy, A., Venugopalan, S., Widner, K., Madams, T., Cuadros, J., Kim, R., Raman, R., Nelson, P., Mega, J. and Webster, D. (2016). Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. JAMA, 316(22), p.2402.

Singh, Y., Kumar Bhatia, P. and Sangwan, O. (2007). A Review of Studies On Machine Learning Techniques. International Journal of Computer Science and Security (IJCSS), 1(1), pp.70 - 84.

Sayed, S., Inamdar, D. and Kapre, S. (2017). Detection of Diabetic Retinopathy using Image Processing and Machine Learning. International Journal of Innovative Research in Science, Engineering and Technology, [online] 6(1), pp.99-107. Available at: http://www.ijirset.com [Accessed 11 Jun. 2019].

Ting, D., Pasquale, L., Peng, L., Campbell, J., Lee, A., Raman, R., Tan, G., Schmetterer, L., Keane, P. and Wong, T. (2018). Artificial intelligence and deep learning in ophthalmology. British Journal of Ophthalmology, 103(2), pp.167-175.

Gargeya, R. and Leng, T. (2017). Automated Identification of Diabetic Retinopathy Using Deep Learning. Ophthalmology, 124(7), pp.962-969.

Vogt, M. (2018). An overview of deep learning techniques. at - Automatisierungstechnik, 66(9), pp.690-703.

COŞKUN, M., YILDIRIM, Ö., UÇAR, A. and DEMIR, Y. (2017). AN OVERVIEW OF POPULAR DEEP LEARNING METHODS. European Journal of Technic, 7(2), pp.165-176.

Du, K. and Swamy, M. (2013). Recurrent Neural Networks. Neural Networks and Statistical Learning, pp.337-353.

Harrison, R. (2018). Continuous restricted Boltzmann machines. Wireless Networks.

Kaggle.com. (2019). Kaggle: Your Home for Data Science. [online] Available at: https://www.kaggle.com/ [Accessed 18 Jun. 2019].

Guillaume PATRY, H. (2019). Messidor - ADCIS. [online] ADCIS. Available at: http://www.adcis.net/en/third-party/messidor/ [Accessed 18 Jun. 2019].

Isi.uu.nl. (2019). DRIVE: Digital Retinal Images for Vessel Extraction. [online] Available at: http://www.isi.uu.nl/Research/Databases/DRIVE/ [Accessed 18 Jun. 2019].

TensorFlow. (2019). TensorFlow. [online] Available at: https://www.tensorflow.org [Accessed 19 Jun. 2019].

Keras.io. (2019). Home - Keras Documentation. [online] Available at: https://keras.io [Accessed 19 Jun. 2019].

Deeplearning.net. (2019). Welcome — Theano 1.0.0 documentation. [online] Available at: http://deeplearning.net/software/theano/ [Accessed 19 Jun. 2019].

Lasagne.readthedocs.io. (2019). Welcome to Lasagne — Lasagne 0.2.dev1 documentation. [online] Available at: https://lasagne.readthedocs.io/en/latest/ [Accessed 23 Aug. 2019].

Rahman, M. (2016). The Advantages and Disadvantages of Using Qualitative and Quantitative Approaches and Methods in Language “Testing and Assessment” Research: A Literature Review. Journal of Education and Learning, 6(1), p.102.

S, S. (2017). A Study of Software Development Life Cycle Process Models. SSRN Electronic Journal.

Godwill, E. (n.d.). Fundamentals of research methodology.

Research-Methodology. (2019). Data Collection Methods - Research-Methodology. [online] Available at: https://research-methodology.net/research-methods/data-collection/ [Accessed 3 Jul. 2019].

Dutta, S., Manideep, B., Basha, S., Caytiles, R. and Iyengar, N. (2018). Classification of Diabetic Retinopathy Images by Using Deep Learning Models. International Journal of Grid and Distributed Computing, 11(1), pp.99-106.

Miotto, R., Wang, F., Wang, S., Jiang, X. and Dudley, J. (2017). Deep learning for healthcare: review, opportunities and challenges. Briefings in Bioinformatics, 19(6), pp.1236-1246.

Mahiba, C. and Jayachandran, A. (2019). Severity analysis of diabetic retinopathy in retinal images using hybrid structure descriptor and modified CNNs. Measurement, 135, pp.762-767.

Wan, S., Liang, Y. and Zhang, Y. (2018). Deep convolutional neural networks for diabetic retinopathy detection by image classification. Computers & Electrical Engineering, 72, pp.274-282.

Hassan, B., Hassan, T., Li, B., Ahmed, R. and Hassan, O. (2019). Deep Ensemble Learning Based Objective Grading of Macular Edema by Extracting Clinically Significant Findings from Fused Retinal Imaging Modalities. Sensors, 19(13), p.2970.

Jadhav, A. and Patil, P. (2016). Detection of optic disc from retinal images using wavelet transform. 2016 International Conference on Signal Processing, Communication, Power and Embedded System (SCOPES).

Shanthi, T. and Sabeenian, R. (2019). Modified Alexnet architecture for classification of diabetic retinopathy images. Computers & Electrical Engineering, 76, pp.56-64.

Cao, K. (2019). Artificial intelligence on diabetic retinopathy diagnosis: an automatic classification method based on grey level co-occurrence matrix and naive Bayesian model. International Journal of Ophthalmology, 12(7), pp.1158-1162.

Sahoo, R. and Sekhar, C. (2019). Detection of Diabetic Retinopathy from Retinal Fundus Image using Wavelet based Image Segmentation. International Journal of Computer Applications, 182(47), pp.46-50.

Raju, M., Aswath, A., Pagidimarri, V., Barreto, R., Kadam, A. and Kasivajjala, V. (2017). Development of a Deep Learning Algorithm for Automatic Diagnosis of Diabetic

Retinopathy. International Medical Informatics Association (IMIA) and IOS Press. [online] Available at: https://www.researchgate.net/publication/322252818 [Accessed 24 Aug. 2019].

Gargeya, R. and Leng, T. (2017). Automated Identification of Diabetic Retinopathy Using Deep Learning. Ophthalmology, 124(7), pp.962-969.

Li, Z., Keel, S., Liu, C., He, Y., Meng, W., Scheetz, J., Lee, P., Shaw, J., Ting, D., Wong, T., Taylor, H., Chang, R. and He, M. (2018). An Automated Grading System for Detection of Vision-Threatening Referable Diabetic Retinopathy on the Basis of Color Fundus Photographs. Diabetes Care, 41(12), pp.2509-2516.

Hemanth, D., Deperlioglu, O. and Kose, U. (2019). An enhanced diabetic retinopathy detection and classification approach using deep convolutional neural network. Neural Computing and Applications.

Paranjpe, M. and Kakatkar, M. (2014). REVIEW OF METHODS FOR DIABETIC RETINOPATHY DETECTION AND SEVERITY CLASSIFICATION. International Journal of Research in Engineering and Technology, 03(03), pp.619-624.

Maher, R. and Dhopeshwarkar, M. (2016). Automatic Detection of Microaneurysms in digital fundus images using LRE. IOSR Journal of Computer Engineering (IOSR-JCE), 18(1), pp.11-17.

Surveymonkey.com. (2019). SurveyMonkey: The World's Most Popular Free Online Survey Tool. [online] Available at: https://www.surveymonkey.com/ [Accessed 5 Jul. 2019].

Ramesh, A., Kambhampati, C., Monson, J. and Drew, P. (2004). Artificial intelligence in medicine. Annals of The Royal College of Surgeons of England, 86(5), pp.334-338.

Mayo Clinic. (2019). Diabetes - Symptoms and causes. [online] Available at: https://www.mayoclinic.org/diseases-conditions/diabetes/symptoms-causes/syc-20371444 [Accessed 10 Jul. 2019].

Nei.nih.gov. (2019). [online] Available at: https://nei.nih.gov/sites/default/files/Diabetic-Retinopathy-What-You-Should-Know-508.pdf [Accessed 10 Jul. 2019].

Wu, L. (2013). Classification of diabetic retinopathy and diabetic macular edema. World Journal of Diabetes, 4(6), p.290.

Haddad, O. and Saad, M. (1998). Prevalence and risk factors for diabetic retinopathy among Omani diabetics. British Journal of Ophthalmology, 82(8), pp.901-906.

Alrawahi, A., Rizvi, S., Al-Riyami, D. and Al-Anqoodi, Z. (2012). Prevalence and Risk Factors of Diabetic Nephropathy in Omani Type 2 Diabetics in Al-Dakhiliyah Region. Oman Medical Journal, 27(3), pp.212-216.

Al-Shookri, A., Khor, G., Isa, H., Loke, S. and Al-Maskari, M. (2019). Type 2 diabetes in the sultanate of Oman. [online] Semanticscholar.org. Available at: https://www.semanticscholar.org/paper/Type-2-diabetes-in-the-sultanate-of-Oman.-Al-Shookri-Khor/1486238f4e53a5073439b02b8b12b80d55547f52 [Accessed 22 Aug. 2019].

Abràmoff, M., Lou, Y., Erginay, A., Clarida, W., Amelon, R., Folk, J. and Niemeijer, M. (2016). Improved Automated Detection of Diabetic Retinopathy on a Publicly Available Dataset Through Integration of Deep Learning. Investigative Opthalmology & Visual Science, 57(13), p.5200.

Pratt, H., Coenen, F., Broadbent, D., Harding, S. and Zheng, Y. (2016). Convolutional Neural Networks for Diabetic Retinopathy. Procedia Computer Science, 90, pp.200-205.

12 Appendix