بِسْمِ ٱللَّهِ ٱلرَّحْمَٰنِ ٱلرَّحِيمِ

SUDAN UNIVERSITY OF SCIENCE AND TECHNOLOGY

COLLEGE OF GRADUATE STUDIES

Automatic Diagnosis of Non-proliferative Diabetic Retinopathy Using Artificial Neural Network

التشخيص التلقائي لاعتلال شبكية العين الناتج عن السكري باستخدام الشبكة العصبية الاصطناعية

A THESIS SUBMITTED IN PARTIAL FULFILMENT FOR DEGREE OF M.sc IN BIOMEDICAL ENGINEERING

By:

Mawada Abdalbasit Saeed

Supervisor:

Dr. Zienab Adam Mustafa

2022

الآيــة

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

قال تعالى: قَالُوا سُبْحَانَكَ لَا عِلْمَ لَنَا إِلَّا مَا عَلَّمْتَنَا ۖ إِنَّكَ أَنْتَ الْعَلِيمُ الْحَكِيمُ ﴿32﴾

سورة البقرة الآية 32

Dedication

As well as everything that I do, I would be honored to dedicate this dissertation to ***my parents****. The two person that gave the tools and values necessary to be where I am standing today. My parents support me on every step I make, and decision I take*

Also, I dedicate it to my supporters who support me in every steps in my life, and help me to follow my dream, who look forward and always had high hope in my success to my ***brothers and sisters'*** *life is nothing without you.*

To my ***husband*** *who encourages and stays by my side all the time*

To those who I chose to spend my days with, share joy and grief, laughs, and gained experience together to ***my friends***

Acknowledgement

First, I thank ALLAH for his blessing us with the strength to complete this project

I cannot express enough thanks to my supervisor for her continued support and encouragement: ***Dr.* Zeinab Adam Mustafa** *who has answered my questions, assisted me with writing and was positive support for this project.*

*I further wish to thank several individuals (****Eng. Abdelazim Ibrahim Ahmed, Eng. Hossam Hatim, Dr. Rayyan Ibrahim)***

Table of contents

List of figures

List of tables

List of Abbreviations

DR: Diabetic Retinopathy……..
OD: Optical Disk……
PDR: Proliferative Diabetic Retinopathy……
PDR: Proliferative Diabetic Retinopathy……
NPDR: Non-proliferative Diabetic Retinopathy……
MAs: Microaneurysms……
CWS: Cotton Wall Spot……..
RMS: Root Mean Square……
SD: Standard Deviation……..
AM: Arithmetic Mean……
DIP: Digital Image Processing……
ANN: Artificial Neural Network……
CAD: Computer Aided Diagnosis……..
SVM: Support Vector Machine……..

ABSTRACT

Diabetic retinopathy is an eye disease caused by diabetes mellitus which affects the retina. It leads the retina blood vessels to swell, these damages the retina and may lead to blindness if it does not early and accurately detect. The most effective treatment is early detection through regular screenings, but fundus images are poor quality to be diagnostic**.**

In this research, we present an automatic diagnosis system to classify in which stage the non-proliferative diabetic retinopathy using artificial neural network. The grading of the severity level of DR is based on detecting and analyzing the early clinical signs associated with the disease, such as microaneurysms' proposed method consists of five stages: pre-processing, segmentation of optic disk, detection of MAs, feature extraction and classification. Mathematical morphology operation is used for pre-processing. K-mean technique used for segmentation of MAs, Artificial neural network is used for classification of the disease stage. A database of 261 color images are used in order to evaluate the performance of the developed system. The system achieves 99.0% of sensitivity ,95.833% of specificity and 98.18% of accuracy.

المستخلص

اعتلال الشبكية السكري هو أحد داء السكري التي تصيب العين. ويحدث هذا المرض بسبب تلف الاوعية الدموية بالأنسجة الحساسة للضوء الموجودة في الجزء الخلفي للعين(الشبكية)، وفي البداية قد لا ينتج عن اعتلال الشبكية السكري اية اعراض وإذا لم قد يتسبب في مشاكل بسيطة في الرؤية إذا لم يتم تشخيصه وعلاجه مبكرا يمكن ان يسبب ذلك العمى الدائم. لضمان فاعلية العلاج يجب الكشف المبكر من خلال الفحوصات المنتظمة، ولكن معظم الصور لشبكية العين رديئة الجودة، مما يجعل عملية اكتشافه صعبة وزيادة احتمالية حدوث خطأ في التشخيص.

يقدم هذا البحث نظام ألي لتصنيف درجة الاصابة بالمرض باستخدام الشبكة العصبية الصناعية ، تم استخدام خمسة مراحل وهي قبل المعالجة للصورة باجراء عمليات حسابية ، تجزئة القرص البصري ، الكشف عن الاوعية المتفجرة ، استخراج السمات من الصورة ، واخيرا مرحلة التصنيف بأستخدام الشبكة العصبية الصناعية.

قاعدة بيانات تتكون من مئتان وواحد وستون صورة ملونة تم استخدامهم في النظام. حقق النظام 98.18% من الدقة وحساسية 99.00% وخصاصية 95.833%.

CHAPTER ONE
INTRODUCTION

1.1. General Overview:

Diabetic retinopathy (DR) is one of the complications of diabetes and a leading cause of blindness in the world [1]. The tiny blood vessels inside the retina are damaged due to diabetes and result in various vision-related problems and it may lead to complete vision loss without early detection and treatment [2]. The retina is the light-sensitive layer of cells at the back of the eye that convert light into electrical signals. The signals are sent to the brain which turns them into the images. The retina needs a constant supply of blood, which it receives through a network of tiny blood vessels [3].

Over time, a persistently high blood sugar level can damage these blood vessels and most diabetes patients will never go for a fundamental evaluation of DR since it is a time-consuming procedure It affects 80% of the people who have diabetes. It does not cause any symptom during its initial stage and some people may notice minute changes in their vision [2].

Diabetic retinopathy falls into two main classes **Non-proliferative diabetic retinopathy (NPDR)** is characterized by structural damage to small retinal blood vessels, cause them to dilate, leak, or rupture. Visible retinal lesions include microaneurysms, at first alone but later accompanied by one or more of the following: (dot or blot hemorrhages, hard exudates (HE), soft exudates or cotton-wool spots (CWS), intraretinal microvascular abnormalities (IRMA), and venous beading).The word “proliferative” refers to whether there is neovascularization (abnormal blood vessel growth) in the retina. As the disease progresses, it may evolve into **proliferative diabetic retinopathy (PDR)**, which is defined by the presence of neovascularization and has a greater potential for serious visual consequences [4].

NPDR is Further Subdivided Based on Retinal Findings:

- Early NPDR – At least one microaneurysm present on retinal exam.
- Moderate NPDR – Characterized by multiple microaneurysms, dot-and-blot hemorrhages, venous beading, and/or cotton wool spots.
- Severe NPDR – In the most severe stage of NPDR, you will find cotton wool spots, venous beading, and severe intraretinal microvascular abnormalities (IRMA) [4].

1.2. Problem Statement:

- Diabetic retinopathy is globally the primary cause of visual impairment and blindness in diabetic patients
- Retinal image is essential and crucial for ophthalmologists to diagnose diseases and precise the severity of it for early detection. Unfortunately, in a normal situation the retinal images are low-quality images.

1.3. Objectives

1.3.1. General objectives

The aim is to design automatic diagnosis system for non-proliferative diabetic retinopathy.

1.3.2. Specific objectives:

- Improve Non-proliferative diabetic retinopathy detection and classification using artificial neural network and identify diabetic at which stage.
- Reduce the percentage of error by choosing the appropriate method in each stage to provide good diagnostic accuracy.

1.4. Methodology:

The proposed method is a four stage CADe system for Classification fundus image as non-proliferative diabetic retinopathy moderate stage or severe stage. The first stage comprises pre-processing and the next stage is segmentation of anatomical structures and pathological parts, the third stage is featuring extraction and the final stage is classification.

1.5 Thesis Layout:

This Thesis consist of six chapters, chapter one is an introduction, chapter two introduces the literature reviews, chapter three provides the theoretical background. Research methodology described in chapter four, the result obtained, and dis of the thesis result given in chapter five, and finally in chapter six conclusion and future work should be considered.

CHAPTER TWO

LITERATURE REVIEWS

2.1 Introduction:

Early detection of DR prevent patient from blindness, in the past years many papers have developed algorithms to detect in which stage non-proliferative diabetic retinopathy is and below some of them.

S. Kumar et al, (2020) presented an automated early DR detection scheme through improved segmentation strategies of blood vessels and optic discs. This paper mainly focused on red lesion features such as microaneurysm and hemorrhages to detect the early stage of DR using the DIARETDB1 database. The sensitivity and specificity obtained with this improved model were 87% and 93% respectively [5].

S.S.Chowhan et al, *(*2018*)* They presented the segmentation of blood vessels of non-proliferative diabetic retinopathy; the assessment of segmentation was done on DRIVE dataset. They applied a morphological operation for detection and segmentation of blood vessels from retinal images. The experimental results express the retinal vessels can be effectively detected and segmented [6].

H.Y.Taso et al, (2018) built a prediction model for the DR in type 2 diabetes mellitus, so they could detect the related risk factors more accurately and exercised early prevention strategies for diabetic retinopathy in the most high-risk population. They used data mining techniques including Support Vector Machine (SVM), Logistic Regression, Artificial Neural Networks (ANN) and Decision Trees. The maximum AUC, Accuracy, and sensitivity of 0.839, 0,795 and 0.933 are achieved with Support Vector Machine classifier and reached maximum specificity of 0.750 with Decision Trees classifier. This method identified using of insulin and duration of diabetes as novel interpretable

features to assist with clinical decisions in identifying the high-risk populations for diabetic retinopathy [7].

M.D. Saleh et al, *(*2017) they provided an automated diagnosis system for DR integrated with a user-friendly interface. The system extracted some retinal features, such as optic disc, fovea, and retinal tissue for easier segmentation of dark spot lesions in the fundus images. Then classified of the correctly segmented spots into MAs and HAs. The system quantified the severity level of DR based on the number and location of MAs and HAs. A database of 98 color images were used to evaluate the performance of the developed system. the proposed system achieved 84.31% and 87.53% values in terms of sensitivity for the detection of MAs and HAs respectively [8].

Al-Jarrah et al, (2017) they proposed a novel morphology-based algorithm for detecting retinal lesions and classifying each case. First, the proposed algorithm detects the three DR lesions, namely hemorrhages, microaneurysms and exudates. Second, they defined and extracted a set of features from detected lesions. The set of selected feature emulates what physicians looked for in classifying NPDR case. Finally, they designed an artificial neural network (ANN) classifier with three layers to classify NPDR to normal, mild, moderate, and severe. Bayesian regularization and resilient backpropagation algorithms are used to train ANN. The accuracy for the proposed classifiers based on Bayesian regularization and resilient backpropagation algorithms are 96.6 and 89.9, respectively. The obtained results were compared with results of the recent published classifier. the proposed classifier outperforms the best in terms of sensitivity and specificity [9].

H.Pratt et al*,* (2016) proposed a Convolutional Neural Networks (CNN) approach to diagnose DR from retinal fundus images. They developed a network with CNN architecture and data augmentation which can identify the features such as micro-aneurysms, exudate and hemorrhages on the retina

images using the KAGGLE database. They reached accuracy and sensitivity of 75% and 95% respectively [10].

M.Shahin et al, (2012) proposed a system for automated classification of normal, and abnormal retinal images by automatically detecting the blood vessels, hard exudates microaneurysms, entropy and homogeneity. This paper mainly focused on computing blood vessels area, exudates area, microaneurysms area, entropy, and homogeneity from the processed retinal images. finally fed to the artificial neural network (ANN) classifier for the automatic classification. accuracy is 92 %, a sensitivity of more than 88 %, and a specificity of 100% obtained with this automated system [11].

Table (2.1): comparing b/w pervious study

Name of The Research	Author Year of publishing	Method use	Result
An Automated Early Diabetic Retinopathy Detection Through Improved Blood Vessel and Optic Disc Segmentation	Shailesh Kumara, Abhinav Adarsha, Basant Kumara, Amit Kumar Singh 2020	Mathematical Morphology Operation	Sensitivity = 87% Specificity =93%
Retinal Vessel Segmentation Of Nonproliferative Diabetic Retinopathy	Santosh S. Chowhan, Rakesh S. Deore, Sachin A. Naik. 2018	Morphological Operations	

Predicting Diabetic Retinopathy And Identifying Interpretable Biomedical Features Using Machine Learning Algorithms	Hsin-Yi Tsao1,2, Pei-Ying Chan3,4 And Emily Chia-Yu Su 2018	Support Vector Machines	AUC= 79.5%
Non-Proliferative Diabetic Retinopathy Symptoms Detection and Classification Using Neural Network	Mohammad A Al-Jarrah [1] , Hadeel Shatnawi [1] 2017	Bayesian Regularization and Resilient Backpropagation Algorithms	AUC=96.9 %
Convolutional Neural Networks for Diabetic Retinopathy	Harry Pratta,∗, Frans Coenenb, Deborah M Broadbentc, Simon P Hardinga,C, Yalin Zhenga,C 2016	Convolutional Neural Network	AUC=75% Sensitivity = 95%

CHAPTER THREE
THEORETICAL BACKGROUND

3.1 The human eye

The human eye is a paired sense organ that reacts to light and allows vision. Rod and cone cells in the retina are photoreceptive cells which can detect visible light and convey this information to the brain. It signals information, which is used by the brain to elicit the perception of color, shape, depth, movement, and other features [12].

3.1.1 Structure of the human eye

Humans have two eyes, situated on the left and the right of the face. The eyes sit in bony cavities called the orbits, in the skull. There are six extraocular muscles that control eye movements. The front visible part of the eye is made up of the whitish sclera, a colored iris, and the pupil. A thin layer called the conjunctiva sits on top of this. The front part is also called the anterior segment of the eye.

The eye is not shaped like a perfect sphere, rather it is a fused two-piece unit, composed of an anterior (front) segment and the posterior (back) segment. The anterior segment is made up of the cornea, iris, and lens. The cornea is transparent and more curved, and is linked to the larger posterior segment, composed of the vitreous, retina, choroid and the outer white shell called the sclera. The cornea is typically about 11.5 mm (0.3 in) in diameter, and 0.5 mm (500 μm) in thickness near its Centre. The posterior chamber constitutes the remaining five-sixths; its diameter is typically about 24 mm. The cornea and sclera are connected by an area termed the limbus. The iris is the pigmented circular structure concentrically surrounding the centre of the eye, the pupil, which appears to be black. The size of the pupil, which controls the amount of light entering the eye, is adjusted by the iris' dilator and sphincter muscles [13].

Light energy enters the eye through the cornea, through the pupil and then through the lens. The lens shape is changed for near focus (accommodation) and is controlled by the ciliary muscle. Photons of light falling on the light-sensitive cells of the retina (photoreceptor cones and rods) are converted into electrical signals that are transmitted to the brain by the optic nerve and interpreted as sight and vision. An overall view is given in figure (3.1) [13].

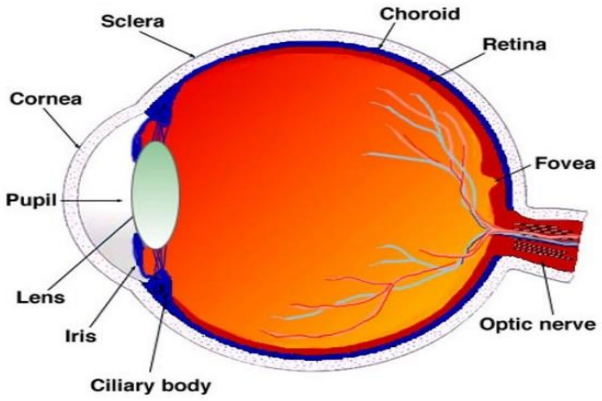


Fig (3.1): Structure of human eye [13].

3.1.2 Optical disc:

The optic disc or optic nerve head is the point of exit for ganglion cell axons leaving the eye. Because there are no rods or cones overlying the optic disc, it corresponds to a small blind spot in each eye [14].

3.1.3 Retina:

The retina is a thin layer of tissue that lines the back of the eye on the inside. It is located near the optic nerve. The purpose of the retina is to receive light that the lens has focused, convert the light into neural signals, and send these signals on to the brain for visual recognition. The retina processes light through a layer of photoreceptor cells. These are essentially light-sensitive cells, responsible for detecting qualities such as color and light-intensity. The retina processes the information gathered by the photoreceptor cells and sends this information to the brain via the optic nerve. Basically, the retina processes a picture from the focused light, and the brain is left to decide what the picture is Due to the

retina’s vital role in vision, damage to it can cause permanent blindness. Conditions such as retinal detachment, where the retina is abnormally detached from its usual position, can prevent the retina from receiving or processing light. This prevents the brain from receiving this information, thus leading to blindness [14].

3.2 Abnormalities associated with the eye:

Abnormalities associated with the eye can be divided into two main classes, the first being disease of the eye, such as cataract, conjunctivitis, blepharitis, and glaucoma. The second group is categorized as systemic diseases such as hypertension, arteriosclerosis and diabetes.

3.2.1 Diabetes:

Diabetes mellitus is a condition of chronically elevated blood glucose concentrations which give rise to its main symptom of passing large quantities of sweet–tasting urine. The fundamental underlying abnormality is a net (relative or absolute) deficiency of the hormone insulin. Insulin is essentially the only hormone that can lower blood glucose. There are two main types of diabetes: type 1 is caused by an autoimmune destruction of the insulin-producing in the pancreas (absolute deficiency); and type 2 is a result of both impaired insulin secretion and resistance to its action often secondary to obesity (relative deficiency. Diabetes is common and is becoming more common. In absolute numbers, globally there are 463 million people aged 20–79 with known diabetes in 2019, projected to rise to 700 million in 2045 [14].

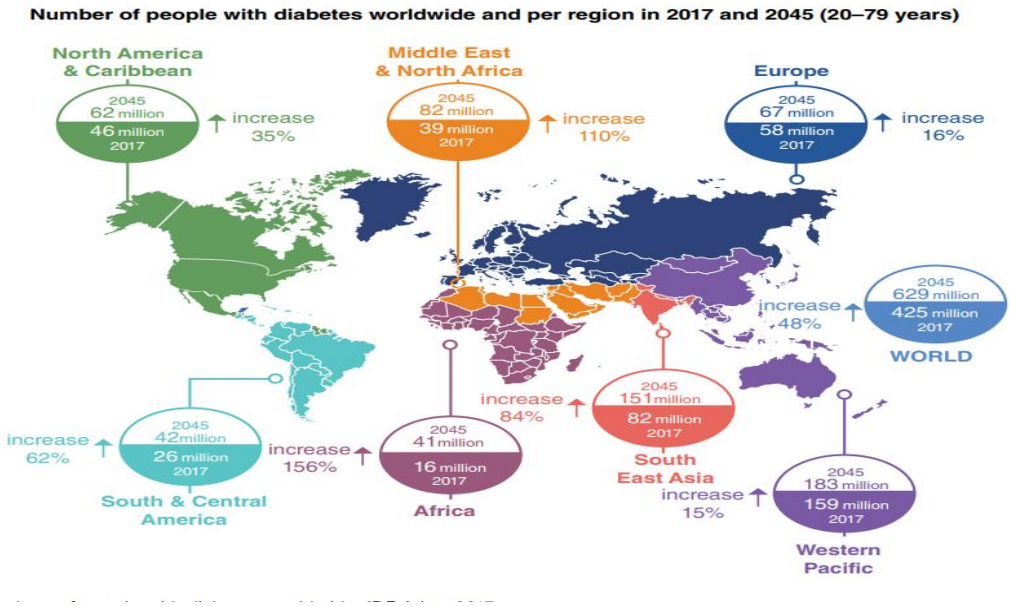


Fig (3.2): Rising numbers of people with diabetes worldwide. IDF Atlas. [14]

- **Type one Diabetes** is an autoimmune disease that attacks the pancreas, this means the pancreas is unable to produce insulin leading to an insulin deficiency. Most commonly appears in childhood or adolescence, but people of any age can develop type one diabetes.
- **Type two Diabetes** means that the body does not respond to the insulin that the pancreas creates. Typically, the pancreas increases the supply of insulin initially, but over time, it reduces the amount of insulin it produces. When the pancreas is unable to produce enough insulin, it leads to a build-up of glucose in the blood stream. Type 2 diabetes most often develops in people over age 45 years, but more and more children, teens, and young adults are also developing it. due to a variety of lifestyle habits, such as energy-dense diets and a lack of physical activity. Insulin resistance causes type 2 diabetes.

3.2.2 Diabetic retinopathy:

Diabetic retinopathy (DR) is a vascular disease of the retina which affects patients with diabetes mellitus. It is the number one cause of blindness in people between the ages of 20-70 worldwide. It falls into two main classes: non-proliferative and proliferative. The word “proliferative” refers to whether there

is neovascularization (abnormal blood vessel growth) in the retina. Early disease without neovascularization is called non-proliferative diabetic retinopathy (NPDR). As the disease progresses, it may evolve into proliferative diabetic retinopathy (PDR), which is defined by the presence of neovascularization and has a greater potential for serious visual consequences. diabetic retinopathy symptoms may include: Spots or dark strings floating in your vision (floaters) Blurred vision. Fluctuating vision. Impaired colour vision. Dark or empty areas in vision. Vision loss. Diabetic retinopathy usually affects both eyes.

Fig (3.3): (a)Normal vision(b)A simulation of what someone seen with advanced diabetic retinopathy [17].

Causes:

Over time, too much sugar in blood can lead to blockage in tiny blood vessels that nourish the retina, leading to cut off its blood supply. As a result, the eye attempts to grow new blood vessels. But these new blood vessels don't develop properly and can leak easily and rupture.

Types of diabetic retinopathy: In this more common form — called non-proliferative diabetic retinopathy (NPDR) — new blood vessels don’t grow(proliferating). Hyperglycemia results in damage to retinal capillaries. This weakens the capillary walls and results in small outpouchings of the vessel lumens, known as microaneurysms. Microaneurysms eventually rupture to form hemorrhages deep within the retina. The weakened vessels also become leaky,

causing fluid to seep into the retina. Fluid deposition under the macula, or macular oedema, interferes with the macula's normal function and is a common cause of vision loss in those with DR. Resolution of fluid lakes can leave behind sediment, like a receding river after a flood. This sediment is composed of lipid by-products and appears as waxy, yellow deposits called hard exudates. As NPDR progresses, the affected vessels eventually become obstructed. This obstruction may cause infarction of the nerve fibre layer, resulting in fluffy, white patches called cotton wool spots (CWS) [4].

NPDR IS FURTHER SUBDIVIDED BASED ON RETINAL FINDINGS:

- **Early NPDR** – At least one microaneurysm present on retinal exam.
- **Moderate NPDR** – Characterized by multiple microaneurysms, dot-and-blot hemorrhages, venous beading, and/or cotton wool spots.
- **Severe NPDR** – In the most severe stage of NPDR, you will find cotton wool spots, venous beading, and severe intraretinal microvascular abnormalities (IRMA).

proliferative diabetic retinopathy: This is the most advanced stage of diabetic retinopathy. Signals sent by the retina for nourishment will essentially cause the retina to produce new blood vessels to enhance nourishment. When these new blood vessels are produced, they are typically fragile and tend to break and produce bleeding causing floaters in the retina and vitreous cavity which, in severe cases, lead to vision loss and in some cases permanent blindness [4].

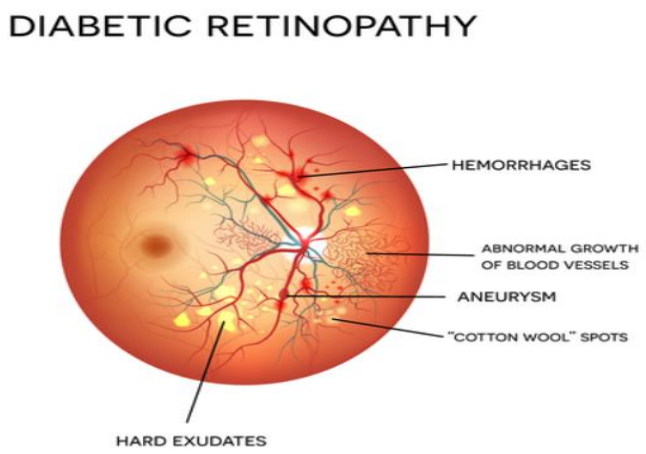


Fig (3.4) Example of retinal image with PDR [17].

3.3 Fundus Image:

Fundus photography involves photographing the rear of an eye; also known as the fundus. Specialized fundus cameras consisting of an intricate microscope attached to a flash enabled camera are used in fundus photography. The main structures that can be visualized on a fundus photo are the central and peripheral retina, optic disc, and macula. Fundus photography can be performed with colored filters, or with specialized dyes including fluorescein and indocyanine green [15].

Fig (3.5): Fundus Camera [15].

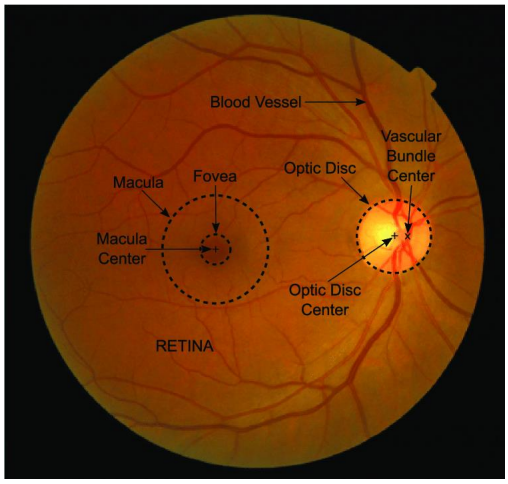


Fig (3.6): Normal fundus photographs of anatomical structure of the right eye, fundus has no sign of disease or pathology [17].

In each picture the macula is in the Centre of the image, and the optic disk is located towards the nose. Both optic disks have some pigmentation at the perimeter of the lateral side, which is considered normal (non-pathological). The orange appearance of the normal fundus is due to complexes of vitamin A as 11-cisretinaldehyde with opsin proteins in the retina (i.e., rhodopsin). The left image (right eye) shows lighter areas close to larger vessels, which is regarded as a normal finding in younger people.

3.4 Digital image processing:

Digital image processing is the use of a digital computer to process digital images through an algorithm. As a subcategory or field of digital signal processing, digital image processing has many advantages over analogue image processing. It allows a much wider range of algorithms to be applied to the input data and can avoid problems such as the build-up of noise and distortion during processing. Since images are defined over two dimensions (perhaps more) digital image processing may be modelled in the form of multidimensional systems. The generation and development of digital image processing are mainly affected by three factors: first, the development of computers; second, the development of mathematics (especially the creation and

improvement of discrete mathematics theory); third, the demand for a wide range of applications in environment, agriculture, military, industry and medical science has increased [3].

3.5 Digital Fundus image Processing:

The mathematical background for digital fundus image processing will explained in three main steps:

Step (1) - Pre-processing: Image Enhancement.

Step (2) - Processing: image segmentation.

Step (3) - Post processing: Feature Extraction and selection.

3.5.1 Pre-processing: Image Enhancement:

Image enhancement is the process of adjusting digital images so that the results are more suitable for display or further image analysis. For example, you can remove noise, sharpen, or brighten an image, making it easier to identify key features.

Methods of image enhancement:

1-Filtering with morphological operators.

2-Histogram equalization.

3-Noise removal using a Wiener filter.

4-Linear contrast adjustment.

5-Median filtering.

6-Unsharp mask filtering.

7-Contrast-limited adaptive histogram equalization (CLAHE) [16].

Wiener Filter:

Weiner filter was used to remove the Gaussian noise from damaged image depending the estimating statistics for each neighbor per pixel. This filter relies on the strength of the noise (noise of contrast in damaged image). When there is large contrast, the filter with little smoothing. If the contrast is big, the filter will work on more smoothing. This filter produces better results than other filters used to enhancement the image [16].

3.5.2 Processing: image segmentation:

Image Segmentation is the process by which a digital image is partitioned into various subgroups (of pixels) called Image Objects, which can reduce the complexity of the image, and thus analysing the image becomes simpler.

We use various image segmentation algorithms to split and group a certain set of pixels together from the image. By doing so, we are actually assigning labels to pixels and the pixels with the same label fall under a category where they have some or the other thing common in them Using these labels, we can specify boundaries, draw lines, and separate the most required objects in an image from the rest of the not-so important ones[16].

Image segmentation Techniques:

Based on the image segmentation approaches and the type of processing that is needed to be incorporated to attain a goal, we have the following techniques for image segmentation.

1-Threshold Method.

2-Edge Based Segmentation.

3-Region Based Segmentation.

4-Clustering Based Segmentation.

5-Watershed Based Method.

6-Artificial Neural Network Based Segmentation[16].

K-means clustering:

K-means is one of the simplest unsupervised learning algorithms which can address the clustering problems, in general. The process follows a simple and easy way to classify a given image through a certain number of clusters which are fixed apriority.

The algorithm starts at this point where the image space is divided into k pixels, representing k group centroids. Now, each of the objects is then assigned to the group based on its distance from the cluster. When all the pixels are assigned to all the clusters, the centroids now move and are reassigned. These steps repeat until the centroids can no longer shift. At the convergence of this algorithm, we have areas within the image, segmented into ―K‖ groups where the constituent pixels show some levels of similarity [18].

3.5.3 Post processing: Feature Extraction and selection:

Feature extraction refers to the process of transforming raw data into numerical features that can be processed while preserving the information in the original data set. It yields better results than applying machine learning directly to the raw data Feature extraction can be accomplished manually or automatically.

Manual feature extraction requires identifying and describing the features that are relevant for a given problem and implementing a way to extract those features. In many situations, having a good understanding of the background or domain can help make informed decisions as to which features could be useful. Over decades of research, engineers and scientists have developed feature extraction methods for images, signals, and text. An example of a simple feature is the mean of a window in a signal. Automated feature extraction uses specialized algorithms or deep networks to extract features automatically from signals or images without the need for human intervention. This technique can

be very useful when you want to move quickly from raw data to developing machine learning algorithms. Wavelet scattering is an example of automated feature extraction. The features used in this research combined the shape and statistical feature. Basically, shape-based image retrieval consists of the measuring of similarity between shapes represented by their features. Some simple geometric features can be used to describe shapes. Usually, the simple geometric features can only discriminate shapes with large differences; therefore, they are usually used as filters to eliminate false hits or combined with other shape descriptors to discriminate shapes. They are not suitable to be standalone shape descriptors. A shape can be described by different aspects. These shape parameters are Centre of gravity, Axis of least inertia, Digital bending energy, Eccentricity, Circularity ratio, Elliptic variance, Rectangularity, Convexity, Solidity, Euler number, Profiles, Hole area ratio. They will be introduced in this section. The list of statistic feature which use in this research and its mathematical representation will discuss in the next chapter [18].

3.6 Classification using ANN:

Neural nets take inspiration from the learning process occurring in human brains. They consist of an artificial network of functions, called parameters, which allows the computer to learn, and to fine tune itself, by analysing new data. Each parameter, sometimes also referred to as neurons, is a function which produces an output, after receiving one or multiple inputs. Those outputs are then passed to the next layer of neurons, which use them as inputs of their own function and produce further outputs. Those outputs are then passed on to the next layer of neurons, and so it continues until every layer of neurons have been considered, and the terminal neurons have received their input. Those terminal neurons then output the result for the model. Every function, including the initial neuron receives a numeric input, and produces a numeric output, based on an internalized function, which includes the addition of a bias term, which is

unique for every neuron. That output is then converted to the numeric input for the function in the next layer, by being multiplied with an appropriate weight. This continues until one final output for the network is produced [10].

CHAPTER FOUR

METHODOLOGY

4.1. Introduction

The main objective of this research was to design an automatic system which diagnosis the stage of non-proliferative diabetic retinopathy from Fundus image using artificial neural network and classify the image into moderate stage or severe, using matlab to programmer the code.

4.2. Image data set

The data set was collected from (Kaggle Database – detection of non-proliferative diabetic retinopathy dataset) using fundus image, used 80% of the image for training and 20% of the image for the testing, the fundus image of the NPDR contain some features which make it differ from normal DR image, it contained 261 image . 168 images diagnosed as “moderate stage”, 93 images diagnosed as” severe stage”.

Table (4.1) Retinal fundus dataset

Category	**Number**
Moderate stage	168
Severe stage	93

The features that distinguish moderate image from severe are out pushing of microaneurysm, exudates and cotton wall spots.

4.3. Methodology:

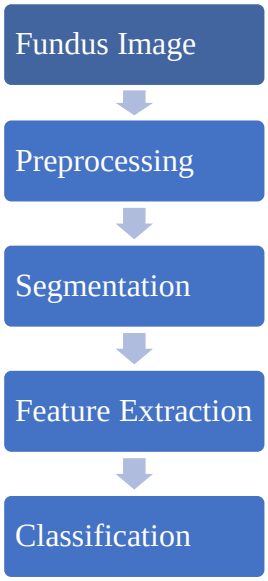


Fig (4.1) The proposed methodology used in this research as general.

4.3.1. Image Pre-processing and enhancement:

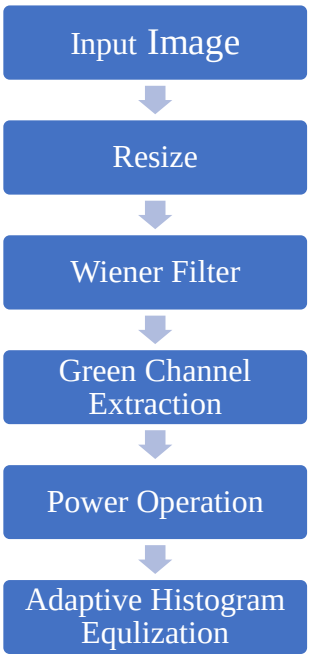


Fig (4.2) Pre-Processing stage

Pre-processing is a crucial stage for preparing the fundus image for segmentation since image quality varies according to the conditions of acquisition. The image could be acquired under some undesired conditions, such as unevenly illuminated, noisy or low-contrasted images, which affects the

performance of segmentation algorithm. Hence the acquired RGB image has to undergo a sequence of preprocessing steps, which are resizing, wiener filter , green-channel extraction, power operation , and adaptive histogram normalization.

The Input to the developed system is a color image of human retina, which is acquired by using a fundus camera, and its outputs are binary images depicting the presence of spot lesions (MAs) and also displaying the severity level of DR. Fig. (4.1) shows the flow chart of the developed system. A database of 261 low-resolution color images compressed by JPEG format was used in this work .

The input fundus images resized to (512*512) pixels decreases the computational time, as well as the storage space.

Wiener filter was used to remove a speckle noise that appears, after trying several filter sizes, we found that (3x3) is the best size which preserves the image details.

img_filt = wiener3d (img_org, [3 3]);

green-channel provides maximum local contrast among the image pixel values. As MAs are distinct from the other retinal features, the green-channel *IG* is first extracted from the RGB image. power operation to increase image contrast.

we used adaptive histogram equalization because the image needs to locally enhanced and divide to blocks so blood vessels, Mas and CWS clearly visible.

img_adapt = adapthisteq(img_filt(:,:,2).^2);

4.3.2. Segmentation of microaneurysms:

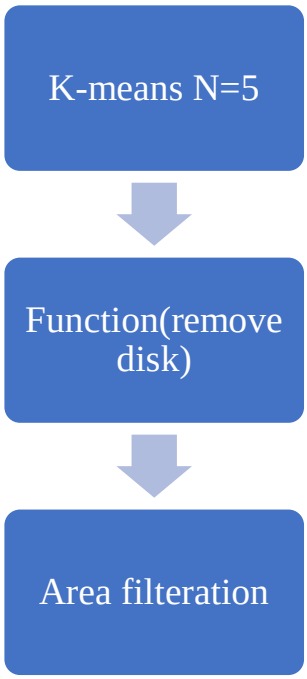


Fig (4.3) segmentation stage

Segmentation is a process that partitions an image into regions with the same properties or characteristics. It plays a major role in image analysis system by facilitating the description of anatomical structures and other regions of interest.

After applying adaptive histogram equalization to the image to increase contrast; the spot lesions microaneurysms will be segmented.

Segmentation done by K-means algorithm with two nonoverlapping classes. Background and noisy pixels were segmented into one class and the fundus image feature which consist of the spots, edam, blood vessels and optic disk of the fundus images were segmented into another class without any pixel belonging into two classes. The non-overlapping of this method made it suitable for this research work where it is only of interest to distinguish between the background and the main fundus image features, Number of clusters=5.

(Km_seg =1) means to choose the class that most dark (MAs).We found that the background, blood vessel will be districted also (undesired features), using and to get rid of undesired features.

Fundus_mask= ismember(km_seg,1) & im2bw (img_filt,0.1) / ismember(km_seg,5);

At the end of these steps, only one object may remain in most of the cases, which represents the optic disc. the image sometimes contains bright spot lesion(s) (i.e. cotton–wool), which may be similar to the optic disc in terms of size and intensity level. In such cases, using remove disk function helps and more effective.

Fundus_mask=removedisk(img_filt,fundus_mask);

4.4. Feature extraction:

4.4.1. Arithmetic mean

Arithmetic mean (AM) the arithmetic mean μ is the average of the Values {X1, X2, X3, …,Xm} located within a time window. It was calculated by equation:

$$\sigma = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(xi-\mu)^2} \qquad \text{……..(4.1)}$$

4.4.2. Standard Deviation

The standard deviation σ was calculated by equation (4.2) to measure how the values {X1, X2…,Xm} are spread out

$$\sigma = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(xi-\mu)^2} \qquad \text{…….……(4.2)}$$

4.4.3. D. Kurtosis

Kurtosis was calculated by equation (4.3) to measure the peak of the probability distribution of the data

$$ku = \frac{\mu_4}{\sigma^4} \qquad \text{……..(4.3)}$$

Where μ4 is the 4th moment of the mean, and given by:

$$\mu 4 = \frac{1}{m}\sum_{i=1}^{m}(xi-\mu)^4$$ ……..(4.4)

4.4.4. Skewness

Skewness was used to measure the asymmetry of the data. It was calculated by equation (4.5):

$$sk = \frac{\mu_3}{\sigma^3}$$ ……(4.5)

Where is the 3th moment about the mean, and given by:

$$\mu_3 = \frac{1}{m}\sum_{i=1}^{m}(xi-\mu)^3$$ ……..…..(4.6)

4.4.5. Entropy:

Entropy was used to measure the impurity associated with a random variable. The entropy H of a discrete variable X with possible values {x1,x2,…..,Xi} and probability mass function P(X) is given by:

$$H(X) = \sum_{i=1}^{M} p(Xi).\log_2 P(Xi)$$ ……..(4.7)

4.4.6. Root Mean Square (RMS):

Root mean square is a measure of the magnitude of a set of values. It is the square root of the arithmetic mean of the squares of the original values. It was calculated using the following equation:

$$RMS = \sqrt{\frac{\sum_{i=1}^{M} xi^2}{m}}$$ ……(4.8)

4.4.7. Correlation:

A correlation coefficient is a numerical measure of some type of correlation, meaning a statistical relationship between two variables. [a] The variables may be two columns of a given data set of observations, often called a sample, or two components of a multivariate random variable with a known distribution. Several types of correlation coefficient exist, each with their own definition and own range of usability and characteristics. They all assume values in the range from -1 to +1, where ±1 indicates the strongest possible agreement and 0 the strongest possible disagreement. tools of analysis, correlation coefficients present certain problems, including the propensity of some types to be distorted by outliers and the possibility of incorrectly being used to infer a causal relationship between the variables (for more, see Correlation does not imply causation).

4.4.8 Variance:

In probability theory and statistics, the coefficient of variation (CV), also known as relative standard deviation (RSD), is a standardized measure of dispersion of a probability distribution or frequency distribution. It is often expressed as a percentage and is defined as the ratio of the standard deviation σ to the mean μ (or its absolute value) **Last two features is Energy and Stander Deviation Error**.

4.5. Classification:

The selected features from the image were used to detect in which stage a non-proliferative diabetic retinopathy using feed forward Artificial Neural Network, Number of neurons =200.

CHAPTER FIVE

RESULT AND DISCUSSION

5.1 RESULT

NPDR has caused serious public health and safety problems, and hence has become a major cause of blindness, so in this research we have collect 261 retinal fundus images and process it to facility the way of diagnosis we train 80% of the data and validate and test 20% by using feed forward neural network and the aim was to classify NPDR in which stage.

5.1.1 Classification performance

In medical image analysis the platform of classifier was evaluated typically by the measure sensitivity and specificity and accuracy and error rate instead of measuring direct accuracy of the classification.

Table (5.1): performance classification

Net type	Specificity	Accuracy	Sensitivity	Error rate
feedforward	95.833	98.18	99.00	1.817

The experimental results for our system and the five stage of processing done to the fundus images as follow:

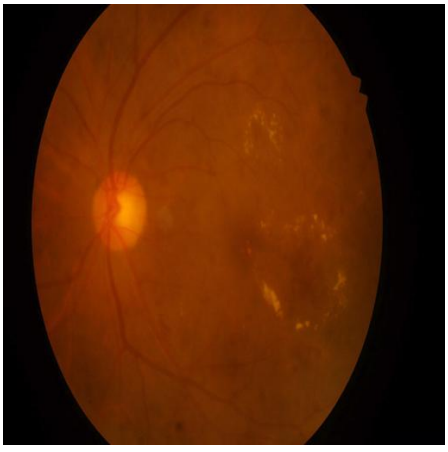


Fig (5.1): an original image from KAGGLE database (NPDR moderate stage).

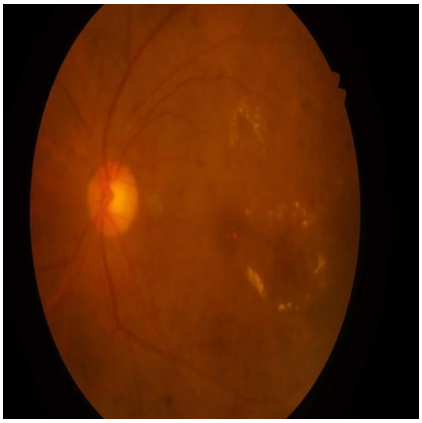


fig (5.2) A filtered image after resizing and applying wiener filter.

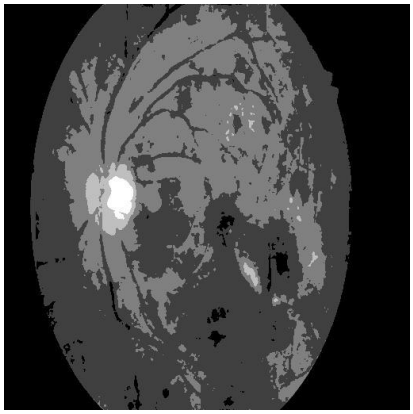
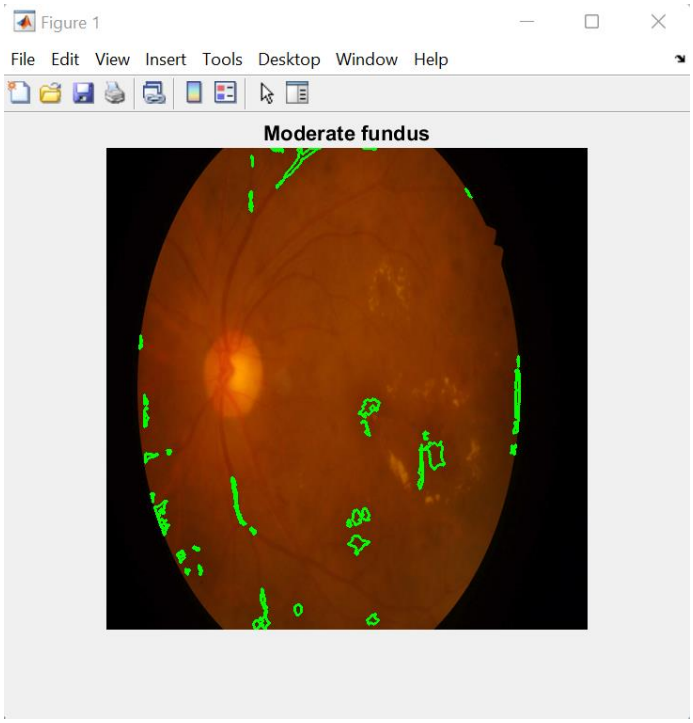


Fig (5.3) K-means segmentation technique to five cluster from dark to bright.

Fig (5.4) The optic disk segmentation.

Fig (5.4) Microaneurysms lesions segmentation.



Fig(5.5) the application of the MATLAB on pre diagnosed retinal image with (NPDR) moderate stage, the algorithm detected microaneurysms and colored it by green.

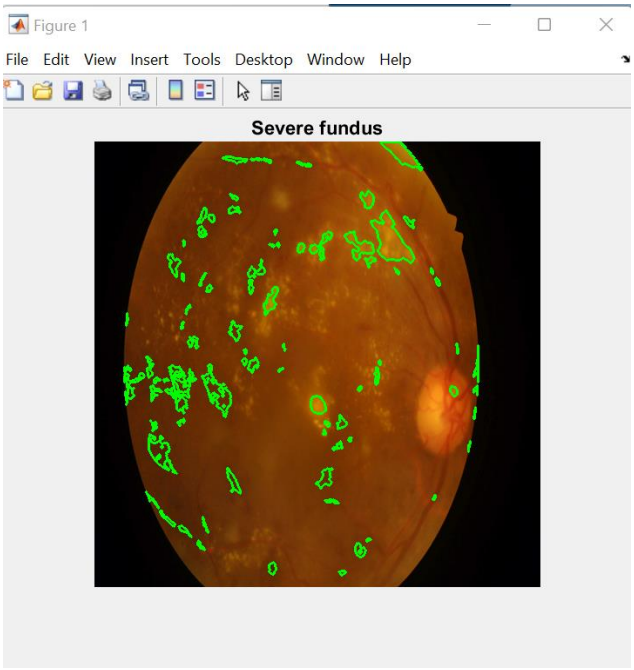


Fig (5.6) shows the application of the MATLAB on pre-diagnosed retinal image with (NPDR) severe stage, the algorithm detected microaneurysms and colored it by green.

5.2 Discussion

Automated medical data processing is a current trend in bioinformatics analysis that is beneficial for disease detection and diagnosis. Diabetic retinopathy (DR) is one of the most serious complications of diabetes mellitus and is a major cause of blindness worldwide. No apparent clinical symptoms exist during the early stage of DR, but the disease causes fine lesions to form on the retina and eventually leads to blindness. Therefore, early diagnosis and treatment of DR are highly important. Thus, similar to other automated disease detection applications, small object detection in fundus images is vital for DR screening. In the early stages of DR, subtle changes occur in the retina, causing capillaries to expand and form microaneurysms (MAs). On a fundus image, these MAs usually appear as dark red dots with diameters ranging from 15 μm to 60 μm. These dots constitute less than 0.15% of the total pixels. Therefore, they are aptly termed “small objects”.

this paper proposes a small object detection method based on k-mean segmentation. firstly, enhances fundus image in many steps, divide the image to 5 clusters from (one to five) the darkest to the brightest object, choosing cluster one and five MAS and CWS for detection. after segmentation feed forward network used to train and testing, contain two hidden layer each layer with size 100 neurons, (trains 80% and testing 20%, repeat this process 200 times, for each 2 time from 200 if the accuracy consistent; swipe the data and repeat the training 20 times. After all steps

were done, we found the diagnosis of NPDR with accuracy 98.8% depending on the number of the image that we use the accuracy can increase and decrease.

Comparing this study with the previous study using the accuracy, sensitivity and specificity that we get it represent one of the acceptable performance research, in this table compare between this study performance and pervious study performance:

Table (5.2): comparing b/w this study performance and pervious study performance.

Name of the research	Accuracy	Sensitivity	Specificity
In the proposed system	98. 18%	99.00%	95.83%
An Automated Early Diabetic Retinopathy Detection Through Improved Blood Vessel and Optic Disc Segmentation	95%	87%	93%
Retinal Vessel Segmentation of Non-proliferative Diabetic Retinopathy	86.7%,	-	-
Predicting Diabetic Retinopathy And Identifying Interpretable Biomedical Features Using Machine Learning	79.5%	-	-

Algorithms			
Non-Proliferative Diabetic Retinopathy Symptoms Detection and Classification Using Neural Network	96.80%	-	-
Convolutional Neural Networks for Diabetic Retinopathy	75%	95%	-

CHAPTER SIX

CONCLUSION AND RECOMMENDATION

6.1. Conclusion

Millions of people worldwide suffer from severe visual impairments and blindness because of rapid progressions of silent ocular diseases. Studies have shown that early diagnosis combined with proper treatment can prevent or delay the severe visual problems resulting from these silent diseases However, early diagnosis requires periodic screening of a huge number of candidate patients which can be limited by the scarcity of ophthalmologists. Automatic retinal screening systems have the potential to facilitate the wide and periodic screening of candidate patients in order to enable early detection of retinal disease symptoms in an efficient and cost-effective manner. proposed technique provide system to diagnose NPDR from digital fundus images using MATLAB software evaluated from KAGGLE WEBSITE, where the proposed approach used to detect (NPDR)via 3 steps: fundus image preprocessing, MAs segmentation and Feature were extract from segmented part and then select the feature most relevant feature, combination of shape feature (3feature) and (statistical feature (10), thirdly and finally ANN classifier applied to find best classification accuracy. It is observed from the experimental result that the proposed system achieves better accuracy for Retinopathy with combined feature than each type of feature individually.

6.2 Recommendations

1-Instead of using digital fundus image Retinopathy progression may be analyzed by using various image modalities such as Optical Coherence Tomography and scanning laser polarimetry.

2-Design, complete, integrated, automated system to classify different type of Diabetic Retinopathy.

3- Using another classification and segmentation techniques.

References

[1] P. Saranya and S. Prabakaran, "Automatic detection of non-proliferative diabetic retinopathy in retinal fundus images using convolution neural network," *J Ambient Intell Humaniz Comput*, 2020, doi: 10.1007/s12652-020-02518-6.

[2] J. Vislisel and T. Oetting, "DIABETIC RETINOPATHY: FROM ONE MEDICAL STUDENT TO ANOTHER," 2010.

[3] C. Sinthanayothin, "Image Analysis for Automatic Diagnosis of Diabetic Retinopathy," 1999.

[4] Z.Lizong and F.Shuxin, "Detection of microaneurysms in fundus images based on an attention mechanism,"doi:10.1008/s12652-02518.

[5] S. Kumar, A. Adarsh, B. Kumar, and A. K. Singh, "An automated early diabetic retinopathy detection through improved blood vessel and optic disc segmentation," *Opt Laser Technol*, vol. 121, Jan. 2020, doi: 10.1016/j.optlastec.2019.105815.

[6] S. S. Chowhan, R. S. Deore, and S. A. Naik, "Retinal Vessel Segmentation of Non-Proliferative Diabetic Retinopathy," *International Journal of Applied Evolutionary Computation*, vol. 10, no. 1, pp. 30–38, Dec. 2018, doi: 10.4018/ijaec.2019010102.

[7] H. Y. Tsao, P. Y. Chan, and E. C. Y. Su, "Predicting diabetic retinopathy and identifying interpretable biomedical features using machine learning algorithms," *BMC Bioinformatics*, vol. 19, Aug. 2018, doi: 10.1186/s12859-018-2277-0.

[8] M. D. Saleh and C. Eswaran, "An automated decision-support system for non-proliferative diabetic retinopathy disease based on MAs and HAs

detection," *Comput Methods Programs Biomed*, vol. 108, no. 1, pp. 186–196, Oct. 2012, doi: 10.1016/j.cmpb.2012.03.004.

[9] Al-Jarrah MA, Shatnawi H. Non-proliferative diabetic retinopathy symptoms detection and classification using neural network. J Med Eng Technol. 2017 Aug;41(6):498-505. doi: 10.1080/03091902.2017.1358772. Epub 2017 Aug 8. PMID: 28786703.

[10] H. Pratt, F. Coenen, D. M. Broadbent, S. P. Harding, and Y. Zheng, "Convolutional Neural Networks for Diabetic Retinopathy," in *Procedia Computer Science*, 2016, vol. 90, pp. 200–205. doi: 10.1016/j.procs.2016.07.014.

[11] M. Shahin1, Taha E. Taha1, W. Al-Nuaimy2, S. El Rabaie1, OsamaF.Zahran1 and Fathi E. Abd El-Samie1Automated Detection ofDiabetic Retinopathy in Blurred Digital Fundus Images2012.

[12] "Handbook of Diabetes."

[13] Ratna Bhargavi V*1, Ranjan K. Senapati2, Bright Lesion Detection in Color Fundus Images Based on Texture Features,2016.

[14] J. Vislisel and T. Oetting, "DIABETIC RETINOPATHY: FROM ONE MEDICAL STUDENT TO ANOTHER," 2010.

[15] (https://en.wikipedia.org/wiki/Fundus_photography).

[16]H.Rajaa,H.Hanna and H.Gaida" Noise Removed by Processing the Lightness and Chromatic Components Basic on YcbCr Color Space," *Journal of Babylon University/Pure and Applied Sciences/ No.(9)/ Vol.(24): 2016.*

[17] www. Wikipedia-pdr.com

[18] L. Zhang, S. Feng, G. Duan, Y. Li, and G. Liu, "Detection of microaneurysms in fundus images based on an attention mechanism," *Genes (Basel)*, vol. 10, no. 10, Oct. 2019, doi: 10.3390/genes10100817.