Design and Implementation System to Measure the Impact of Diabetic Retinopathy Using Data Mining Techniques

Dr. Karim Hashim Al-Saedi [1], Dr. Razi Jabur Al-Azawi[2], Rasha Asaad Kamil[3]

AL-Mustansiriyah University, Baghdad, Iraq[1]
University of Technology, Baghdad, Iraq[2]
Informatics Institute for Postgraduate Studies, UITC, Baghdad, Iraq[3]

Abstract: *In this research, an accurate measurement system of diabetic retinopathy is developed and investigated using data mining techniques. Diabetic retinopathy is an injury produced in the blood vessels of retina because of diabetes. Such a disease leads to a loss of in a patient's vision. Hence, an early analysis of diabetic retinopathy using an accurate and fast technique provides the patient with enough protection treatment time. The color fund us image can be used to automatically detect and realize the various lesions of diabetic retinopathy and its normal features, respectively. The specifications of the normal color fund us images are analyzed and classified by the extraction method into normal or abnormal. Therefore, the abnormal image will then be categorized into three levels: Mild, moderate, and Severe. To predict the unknown class, an association rule and SVM classifier are used. The results are promising to support the patients and the accelerating process. They further hope to overcome many problems in this field and in any future research.*

Keywords: *Accurate Measurement, Diabetes, Retina, Data Mining Techniques, Diabetic Retinopathy.*

1. INTRODUCTION

Nowadays worldwide, the major problems diabetic patients suffer from are the diabetic retinopathy and blindness. The greatest challenge to current health care is the severe progress of diabetes. Hence, the early detection via regular screen and timely intervention highly helps control the disease [1].The distorted vision of diabetes is represented by bleeding or the accumulation of fluid in the retina. However, the symptoms can only be noticed by the diabetic patients when they start suffering from some sight disorders. Diabetic retinopathy signs are abnormal vessels; exudates are yellow-colored lipid leaks from blood vessels. These exudates have two forms; hard and soft exudates. Microaneurysm is a small dark red spot swell in capillaries. The hemorrhages, on the other hand, look similar to microaneurysm, but they are located in the deeper layers of the retina with unorganized shapes and sizes [2].The various ways of medical images can be classified based on the following criteria: texture, neural networks, and data mining task. The types obtained from the last criterion are considered the best among all other types because they help improve the accuracy of the classification[3].A large amount of data extracted from the retinal images is provided to a database while the proposed algorithm helps discover how new images are classified according to the useful information stored in the database. The increase number of diabetic patients in health care systems indicates an increase in the data as well. Hence, data mining techniques are becoming necessary in information organization and in useful novel and hidden knowledge extraction. They further enhance the computing research[4], and play a vital role when researching diabetes. This is because such research will provide valuable assets to the community. Therefore, these techniques are able to detect the required and needed knowledge in this regard[5].Many of eye specialists take extra time and effort to diagnose diabetic symptoms; therefore, the automatic diagnostic systems are very helpful. The common reason of blindness is the diabetic retinopathy; it causes damage in the retina due to a fluid leak in the blood vessels. In this work, the proposed system is believed to be able to analyze the retinal image, measure the impact of DR, and display a report, as shown in Figure 1.
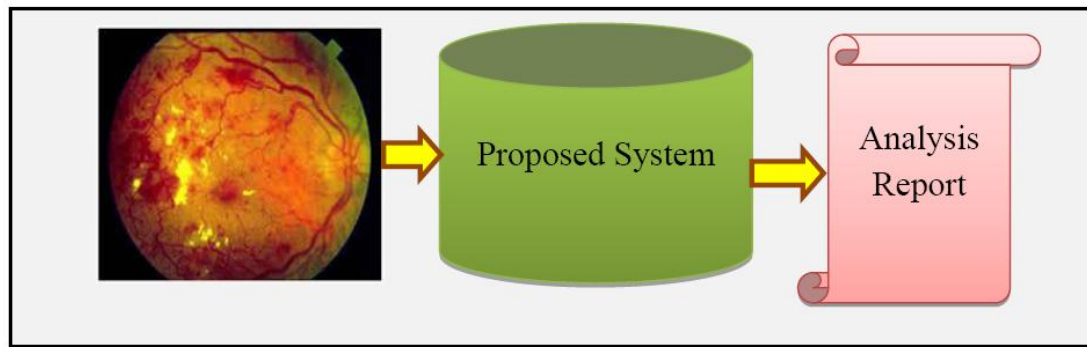
Figure1._General View of the Proposed System_

2. RELATED WORK

A lot of research has been contributed to this field, such as: the automatic detection of diabetic retinopathy in non-dilated RGB retinal fundus images. This study uses an automated method for classifying Microaneurysm, the first clinical sign of DR. The classification is further based on the number of Microaneurysm to identify the severity of DR. The pre-processing model contains many stages, such as: the green channel that is used for RGB retinal fundus image and as a gray scale image. Then, it is also used in image normalization, contrast enhancement, and in threshold as well to obtain a binary image. Later, blood vessels are removed, after the removal of noises. Microaneurysm is finally detected to be classified into four stages as: no DR, Mild DR, Moderate DR, and Severe DR [6]. The identification of diabetic retinopathy stages in human retinal images produces a novel method in the identification of DR levels. Such an identification is done in the first early stage (Microaneurysm) through categorizing the levels into some grades[7]. The detection of red lesions and hard exudates in color fundus images has also been suggested and investigated [1]. The investigation is focused on the automatic method suggested for the detection of the normal features, such as that of blood vessels, optic disc, and fovea. Then, the red lesions are detected and the classifier finally divides the hard exudates into moderate and severe ones based on whether their location is inside or outside the fovea of a macula. An automated detection of diabetic retinopathy using retinal optical images is introduced by [8]. The automated system of DR uses Krisch's edge detection algorithm in the extraction of blood vessels from the retinal image. Then, the input image is pre-processed to extract the vessels, and the features are later extracted and classified into: normal, Mild NPDR, Moderate NPDR, and Severe NPDR. The pre-processing model contains a gray scale conversion, histogram, histogram equalization and image adjustment. The diabetic retinopathy stages of identification using the retinal images are suggested by [2]. In these stages, an association rule with C4.5 classifier is proposed; it involves an image acquisition, enhancement, and feature extraction. Hence, the severity of the disease can be predicted and classified. In early 2016, the diabetic retinopathy analysis using image mining to detect type2 diabetes is suggested by [9]. To detect a diabetic retinopathy, a classifying system has been proposed to classify the severity of the disease into: a standard database DIARETDB1 and HRF into Mild and Severe groups by KNN classifier. The system takes an input image, resizes it and uses a green channel band. The pre-processing stage contains Laplace edge detection and histogram threshold. DWT and GLCM features are classified into normal, Severe DR and Mild DR. In this paper, a data mining technique is used to provide high accuracy in DR measurement system. Consequently, this paper will focus on the stage of diabetic identification depending on the presence of those features in the following table:-

TABLE I	
Diabetic Retinopathy Grading	
Grade	Description
Grade 0: Normal	(MA = 0) AND (H = 0)
Grade 1: Mild NPDR	(0 < MA <= 5) AND (H = 0)
Grade 2: Moderate NPDR	((5 < MA < 15) OR (0 < H < 5)) AND (NV = 0)
Grade3: Severe NPDR	(MA >= 15) OR (H >=5) OR (NV = 1)

3. PROBLEM STATEMENT

Diabetic retinopathy is a vascular disease that affects blood vessels in the retina, causing partial or total blindness to diabetic patients. In the first five years of the disease, the patient who suffers from

diabetes does not complain from any symptoms. Accordingly, if he/she neglects audits and periodic examination of the retina, the disease will control a large part of the retina and cause large blood spots leak. This may in return lead to a retinal detachment from the surrounding wall. These advanced stages are difficult to be treated; therefore, the focus is to discover the disease in its early stages and determine its level. This further requires an immediate treatment to keep the remaining vision. Here comes the need for the automatic system to detect diabetic retinopathy in its early stages before it is fully blindness. Ophthalmologists divide DR into four stages, where the first stage is a normal retina that is free from any abnormalities. The second stage is Mild (NPDR); that is, small red colored swelling spots (dots) shown on the walls of the retina called (Microaneurysm). Moderate (NPDR) is the third stage of the disease where Microaneurysm begins to rupture and appears with small or medium-sized blood on the surface of the retina. Exudates are the fourth stage which is called the severe (NPDR) level. The hard exudates are displayed due to the fragile and weak walls of the blood vessels of the retina. Such weak walls lead to liquid protein leakage from the blood vessels to the surface of the retina. This type of exudates is characterized by yellowish groups of a clear boundary. Soft exudates are the second type of exudates which is characterized by blurred boundaries of a whitish color. These exudates cause blockages in the blood vessels of the retina, and prevent the arrival of food and oxygen for its tissue. They further lead to the emergence of the so-called (Neovascularization), winding new blood vessels; consider Figure2. These vessels are very weak and fragile; they are formed due to blockages in the blood vessels of the retina when delivering food and oxygen to its tissues. Its weak construction causes a large blood area on the surface of the diseased retina. This will consequently lead to the loss of parts of the vision and so the patient will see some of the features of an image or notices dark parts that obscure the natural vision.
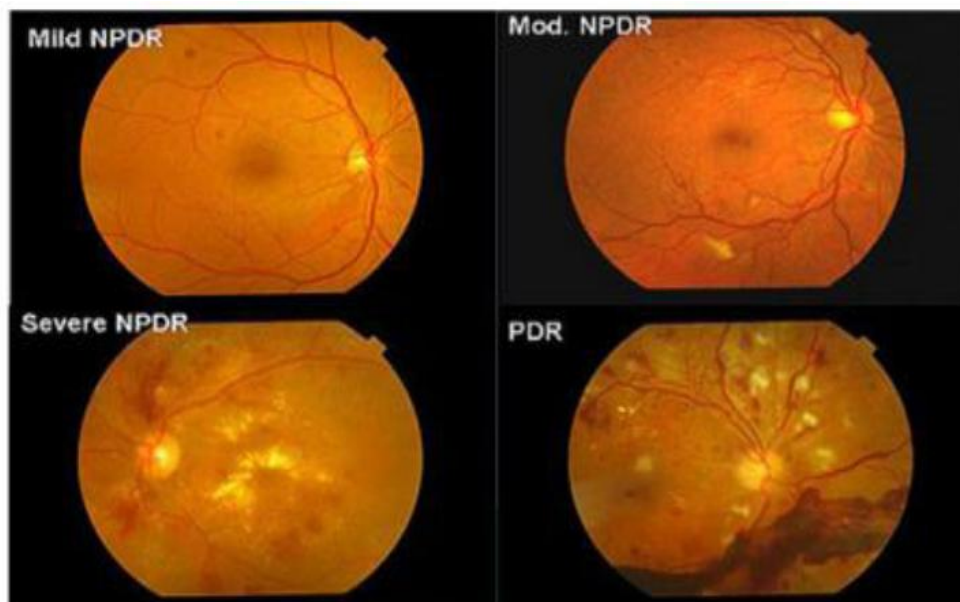

Figure2. *Diabetic Retinopathy Stages*

4. RESEARCH OBJECTIVES

This proposal aims to present an accurate system to measure the impact of diabetes on retina using data mining techniques. Such techniques help increase the classification accuracy of the data. The following procedures will be taken into account when examining the DR:

- extracting the best features of the retinal images using image processing techniques;

- building a new system that works on the classification of retinal images with accurate measurement based on the extracted features;

- adopting an accurate measurement system based on the standard levels as divided by Ophthalmologists; and

- Introducing a report on the patient's case that shows the impact of the disease on the retina. The proposed system reviews the report to the specialist to be able to identify the required treatment.

5. THE PROPOSED SYSTEM

The proposed system is built by combining two modules: the Image processing techniques module, and the data mining techniques module. Each of these modules contains several sub-modules, components, and sub-components. The modules work of the system proposed can be summarized as follows: the Image Processing Techniques Module contains an image pre-processing stage that helps reduce the undesired information of an image. It also increases image contrast, and helps in the extraction and segmentation of the desired features. To reduce false detection, the vessel extraction step is performed and the optic disc is removed. Then, the desired features analysis is used in DR detection. The Data Mining Techniques Module is used to categorize the output features from the image processing module. The classification algorithms will be applied to measure the impact of diabetes disease on the retina. The architecture of this proposed system is shown in Figure 3.

5.1. Image Acquisition

The retinal images are taken from the general dataset. Images are re-sized to 700x500 pixels. This dataset consists of 100 images, corresponding to four types: normal, Mild NPDR, Moderate NPDR and Severe NPDR.
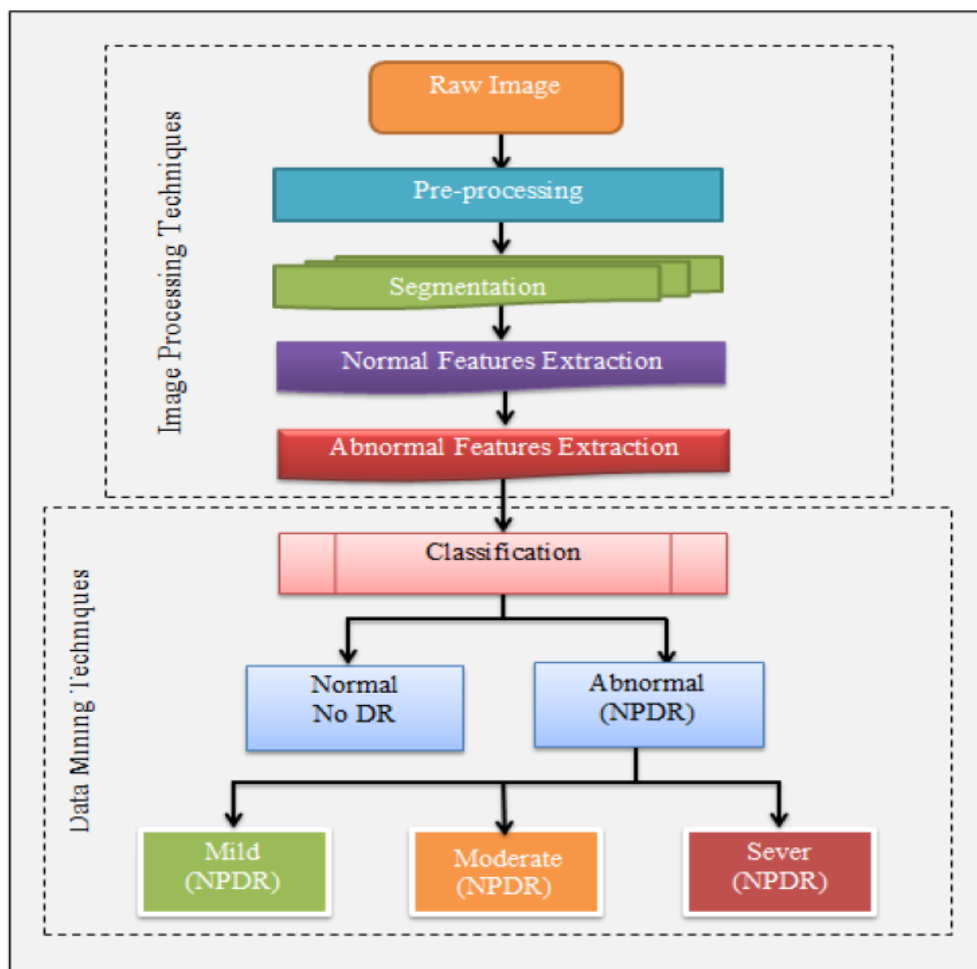

Figure3. *The Architecture of the Proposed System*

5.2. Image Enhancement

To improve the quality of the images, a limited contrastive, and adaptive histogram equalization (CLAHE) is applied. The input image should be smoothed and its edges are sharpened before segmentation, using median filter, mean filter, and edge detection filters. In an enhancement step, mathematical morphological operations are used. This step is a main phase that can lead to have successful or failure results. An enhanced image is segmented using a threshold method to extract the blood vessels and optic disc. Then, a fuzzy algorithm is used to exudate segmentation, and to obtain a region of interest (ROI) by ignoring the undesired region that confuses the result. Segmentation phase is performed to simplify the analysis of the colored fundus retinal images.

5.3. Feature Extraction

Vessels, exudates, microaneurysms and haemorrhagethey are clinically significant retinal features extracted from the retinal images; consider Figure 4below:
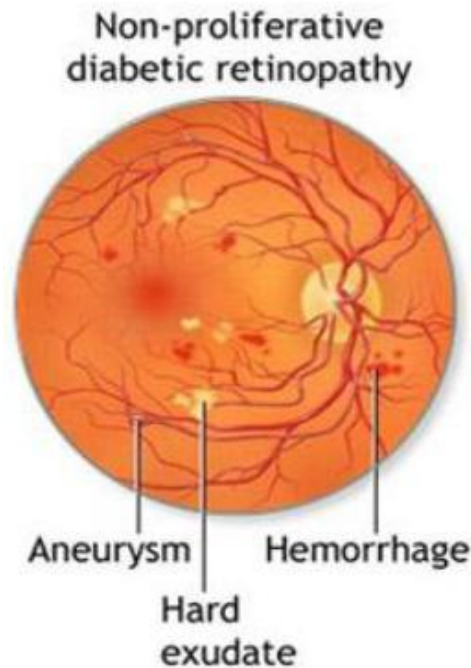

Figure4. *Presence of Exudates, Microaneurysm, and Haemorrhage.*

Features extracted are divided into two types: normal and abnormal features: The normal features in the retinal image are the blood vessels and the optic disc extracted using mathematical morphological operations. These extracted features are sometimes used, removed or neglected to avoid misclassification. The remained features should be analyzed to detect the abnormalities in the retinal image, such as Microaneurysm, Hemorrhages, and Exudates.

5.4. Data Mining

The data mining techniques are used to extract useful, rich and hidden information from the retinal image database. This step helps obtain the required knowledge in decision-making about the charts, diagrams, or reports. The normal retinal image should have no DR abnormalities, such as that of Microaneurysm, Hemorrhages, Hard Exudates, and Soft Exudates.

6. CONCLUSION

This research clearly shows that the results are future aspirations for applying data mining techniques in the expectations of medical database problems. This system may be introduced as a training tool for medical students. It is also considered a helpful hand for Ophthalmologists. It is further hoped for the system to be expanded and improved by combining other medical features, as well as data mining techniques.

REFERENCES

[1] J. Jose and J. Kuruvilla, "Detection of Red Lesions and Hard Exudates in Color Fundus Images," vol. 3, no. 10, pp. 8583–8588, 2014.

[2] A. Thammastitkul and B. Uyyanonvara, "Diabetic Retinopathy Stages Identification Using Retinal Images," pp. 20–23, 2016.

[3] N. D. Panse, "Glaucoma and Diabetic Retinopathy Diagnosis using Image Mining," vol. 117, no. 5, pp. 14–16, 2015.

[4] S. Dua, V. Jain, and H. W. Thompson, "Patient classification using association mining of clinical images," *2008 5th IEEE Int. Symp. Biomed. Imaging From Nano to Macro, Proceedings, ISBI*, pp. 253–256, 2008.

[5] A. A. T. Shraddha Jalan, "Review paper on Diagnosis of Diabetic Retinopathy using KNN and SVM Algorithms," pp. 128–131, 2015.

[6] S. Sb and V. Singh, "Automatic Detection of Diabetic Retinopathy in Non-dilated RGB Retinal Fundus Images," *Int. J. Comput. Appl.*, vol. 47, no. 19, pp. 26–32, 2012.

[7] A. Alaimahal and S. Vasuki, "Identification of diabetic retinopathy stages in human retinal image," vol. 2, no. 2, pp. 808–811, 2013.

[8] S. Ganesh and A. M. Basha, "Automated Detection of Diabetic Retinopathy Using Retinal Optical Images," no. 4, pp. 136–144, 2015.

[9] K. A. Anant, "Diabetic Retinopathy Analysis using Image Mining to Detect Type 2 Diabetes," vol. 5, no. 1, pp. 37–42, 2016.

AUTHORS' BIOGRAPHY

Dr. Karim Hashim Al-Saedi, received the PhD degree in computer science (Data Mining and Network Security) from the Universiti Sains Malaysia (USM) in 2013, and M.Sc. in computer science at University of Technology, Baghdad, Iraq, in 2005 He is working as Lecture in Department of Computer Science, College of science, at AL-Mustansiriyah University, Baghdad, Iraq. Also work as Visitor Lecture in Informatics Institute for Postgraduate Studies, UITC, Baghdad, Iraq. His research interests include the areas of Data Mining, Advanced Internet Security and monitoring, Medical Image Assessment, and Machine learning. He is a Member of the Internet Society (ISOC), since 2012. Malaysia, Member of the UNCTAD Virtual Institute, since 2006. Geneva, Switzerland, and Member of the Iraq Computer Society since 1997. Baghdad, Iraq.

Ass.Prof. Dr. RaziJabur Al-Azawi, Teaching at the University of Technology, Designation is Dr. Eng. I got the scientific degree of assistant professor at 2009 and specialized modeling and computer simulation , specialty flour is Informatics also holds a Bachelor of Laser Engineering , I supervised on numerous theses to students in undergraduate and postgraduate and I have a lot of papers at International Journals its impact factor , I participated in numerous local and international conferences.

Rasha Asaad Kamil, Bachelor of computer science from University of Baghdad at 2005, Higher Diploma degree in Information Technology / Website Technique from Informatics Institute for Postgraduate Studies / Iraqi Commission for Computer and Informatics at 2013.Now I'm a Master degree student in Informatics Institute for Postgraduate Studies, in Image Processing and Data Mining Techniques on a Biomedical approach. I have two papers published in (IJERD) journal, March 2015.