DETECTION OF COTTON WOOL FOR DIABETIC RETINOPATHY BY USING NEURAL NETWORK

BUI HUY TOAN

A THESIS REPORT SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF ENGINEERING IN COMPUTING IN ENGINEERING SYSTEMS
INTERNATIONAL COLLEGE
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG
ACADEMIC YEAR 2018
KMITL-2018-IC-M-011-01

DETECTION OF COTTON WOOL FOR DIABETIC RETINOPATHY BY USING NEURAL NETWORK

BUI HUY TOAN

A THESIS REPORT SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF ENGINEERING IN COMPUTING IN ENGINEERING SYSTEMS
INTERNATIONAL COLLEGE
KING MONGKUT'S INSTITUTE OF TECHNOLOGY LADKRABANG
ACADEMIC YEAR 2018
KMITL-2018-IC-M-011-01

THESIS TITLE	Detection of Cotton Wool for Diabetic Retinopathy by using Neural Network.
STUDENT NAME	Bui Huy Toan
STUDENT ID	59610024
DEGREE	Master of Engineering
PROGRAMME	Computing in Engineering Systems (International Program)
ADVISOR	Dr. Ukrit Watchareeruetai
CO-ADVISOR	Asst. Prof. Dr. Noppadol Maneerat

Abstract

Diabetes is a group of common diseases which lead to diabetic retinopathy in human eyes. For more information, diabetic retinopathy is a cause of blindness and vision loss. For this reason, several researches based on diabetic retinopathy have been conducted throughout the years. In diabetic retinopathy, there are four common symptoms which are exudates, cotton wool, hemorrhages, and microaneurysms. Many studies on exudates, hemorrhage, and microaneurysms have been done and showed effective performances. However, there is very less effective method to detect cotton wool. Therefore, it is necessary to find a method to detect cotton wool for diabetic retinopathy detection.

This study proposes an image processing method which can separate between diabetic retinopathy patients and normal patients based on cotton wool and segments the location of cotton wool in retinal fundus images. Firstly, this study employs median filter to enhance the quality of images, and all images are also rescaled into the same resolution. Next, feature extraction was employed to take important information from pixels of the image, and a feedforward neural network is built to classify the cotton wool. And then, all the selected pixels were converted back to rebuild the classification image, and a simple shape descriptor was operated to improve the classification accuracy. Finally, the input image could be marked by using the classification region, and performance evaluation is computed. The experiment was conducted on 60 images selected from two well-known public datasets which are DIARETDB01 and

MESSIDOR. In diabetic retinopathy patient classification, the result shows a promising performance that is 100%, 93.33%, and 96.66% for sensitivity, specificity, and accuracy, respectively. Also, for patients who have cotton, the method helps to segment regions of cotton wool in retinal fundus image and shows the result that is 91.57%, 99.39%, and 99.36% for sensitivity, specificity, and accuracy, respectively

Acknowledgment

The author would like to express the gratitude to people who have been supported and helped physically or mentally to complete this thesis.

Firstly, I would like to sincerely thank my advisor, Dr. Ukrit Watchareeruetai, and my co-advisor, Asst. Prof. Dr. Noppadol Maneerat for their advices in study accomplishment as well as writing corrections of this work. Their expertise in research experience and technology knowledge helped to orientate my study and inspired me during my entire study. It is impossible for me to complete this study without their kind help and support.

Next, that is my gratitude to Assoc. Prof. Dr. Chaiwat Nuthong, lecturers and all staffs in International College. I highly appreciate the knowledge that I have learned from all lecturers in the past time. The support from the staffs helps me to have a good condition to do my study as well as do administrative issues.

Finally, I would like to thank my family and all my friends for their spiritual support. In particular, my parents are always behind me and unconditionally supported me throughout my challenging time. Their encouragements are always an uncountable assistance for me.

Bangkok, August 2018

Bui Huy Toan

Table of contents

List of figures

List of tables

Chapter 1

Introduction

In this chapter, the author would like to give a description of the problem and how it motivates us to conduct this study.

1.1 Motivation and problem description

Diabetes [1] is a group of disease of human's blood glucose (blood sugar) level, which is involved with insulin. Insulin is a hormone that helps the human body to absorb sugar from the blood into cells to produce energy. Diabetes is diagnosed when the body does not create and uses insulin normally and can be divided into 3 main types of diabetes: type 1, type 2, and gestational diabetes.

- In type 1, the body cannot secrete insulin. Body immune system attacks and destroys cells in pancreas that secretes insulin. The number of patients falls into this case is approximately 10% of all the diabetes cases.
- In type 2, which is more common, the body either cannot produce enough insulin or use insulin well to keep whole body activities work normally. Because your body system is too insensitive to insulin or too little insulin is produced. This takes approximately 90% of the cases for diabetic patients.
- Gestational diabetes is a form of diabetes that occurs during pregnancy. This case only appears in 24^{th} and 28^{th} week of pregnancy. Most women no longer have diabetes after the baby is born. Therefore, this case is unnecessary to be considered.

Diabetes may cause a few serious problems in the body. Diabetes can damage eyes, kidneys and even lead to heart attack or cerebral thrombosis which are blood clots

in the blood vessels. To prevent those effects, early detection of diabetes is very important to help all patients meet the suitable treatments as soon as possible. One of common solution for diabetes detection is using a blood test. However, it usually takes a few days to have result. Besides, the number of diabetic patients was estimated rising to over 382 million people throughout the world in 2013 [17]. Moreover, lack of specialists, who are needed to conduct the blood test, extends the process of blood test for weeks or even a month. For that reason, this study would like to explore a new approach of diabetes analysis by using diabetic retinopathy, which is a consequence disease of diabetes.

In the very early stage, diabetic retinopathy causes some lesions which are known as exudates, cotton wool, hemorrhages, and microaneurysms in human eyes. Exudates are liquids of proteins, lipids or cellular elements that leaked from blood veils or damaged cells due to inflammation. Cotton wools are blood-lacking events of a very small amount of tissues. Hemorrhages are events that blood vessels bleed into the center of the eye. Microaneurysms are dilations of capillary (also known as tiny blood vessels which supply blood to cells). Those lesions damage human vision and cause blindness or partial vision loss. Fortunately, those lesions can be revealed under an eye analysis by a doctor. However, symptoms may be only familiar with specialists, but not a normal person or nurse. More importantly, the number of specialists compared with patients is a problem that deserves consideration, since the number of patients is very huge and they all need medical diagnoses as soon as possible. Besides, after a few stressful working hours, doctors tend to be exhausted and will make many consecutive mistakes. Therefore, the diagnoses from doctors may be not always accurate all the time. For that season, the problem motivated us to propose an automatic system to detect diabetic retinopathy symptoms in retinal fundus image by using computer vision.

Computer vision is a field in artificial intelligence and computer science. The goal of computer vision is to make electronic devices be able to see, process relevant information and provide an appropriate result from digital images or videos in the way that human vision does [25]. With technology development, machines have become more convenient and powerful. Therefore, the machines could replace humans in many tasks. A machine with specific abilities in computer vision can solve various problems such as object classification, object recognition, or object tracking. Besides, machines can work continuously much longer than human without any interruption by weariness and the stableness of computer can prevent human mistakes. Therefore, a computer

system can support human to make a better efficiency. For example, a vehicle tracking system can observe hundreds of vehicles at the same time to find out the violations of traffic laws or follow a suspicious vehicle for weeks or even months to help in criminal investigation [15].

The abilities of machines in computer vision are different one-by-one, since the main factor is about the algorithm that computers employ. A high complexity algorithm may require more resources such as processing time, memory space or even both of them, while a low complexity algorithm may not be efficient enough to solve the problem and give a poor performance. Studies in computer vision are to find an appropriate algorithm to deal with the problems that improve performance, reduce the resource requirement, or advantage both of them. Nowadays, many computer vision studies prefer to employ image processing techniques and machine learning algorithms to solve different problems, since the combinations between them can partially reduce the complexity by learning without being programmed to perform a specific task but still provide an effective result based on requirements. In machine learning, neural networks are well-known techniques which are widely used to solve pattern recognition problems. Neural networks are simple to construct and often appear to be good ability to generalize and respond to unexpected patterns. For that season, this study applies the image processing and neural network as the main methods to deal with diabetic retinopathy.

Many studies of diabetic retinopathy using image processing and machine learning have been done for exudates, hemorrhages, microaneurysms and give various efficiency achievements [9, 12, 14, 26]. However, researchers on cotton wool have not been good enough since the performance is still low efficient. Some of studies combine exudates and cotton wool as the same object for detection and may lead to a confusing result and unsuitable treatment [10]. Cotton wool spots are similar to exudates in color but different in shape and size. More importantly, despite the similarities, cotton wool spots and exudates are not the same things. Therefore, they may tell different problems of eyes in diabetic retinopathy. The appearance of exudates and cotton wools can be revealed by differences between Fig. 1.1 and Fig. 1.2.

As the result, the whole problems as stated above lead us to propose a new method to detect cotton wool in retinal fundus image. The full objective is described in the following section.

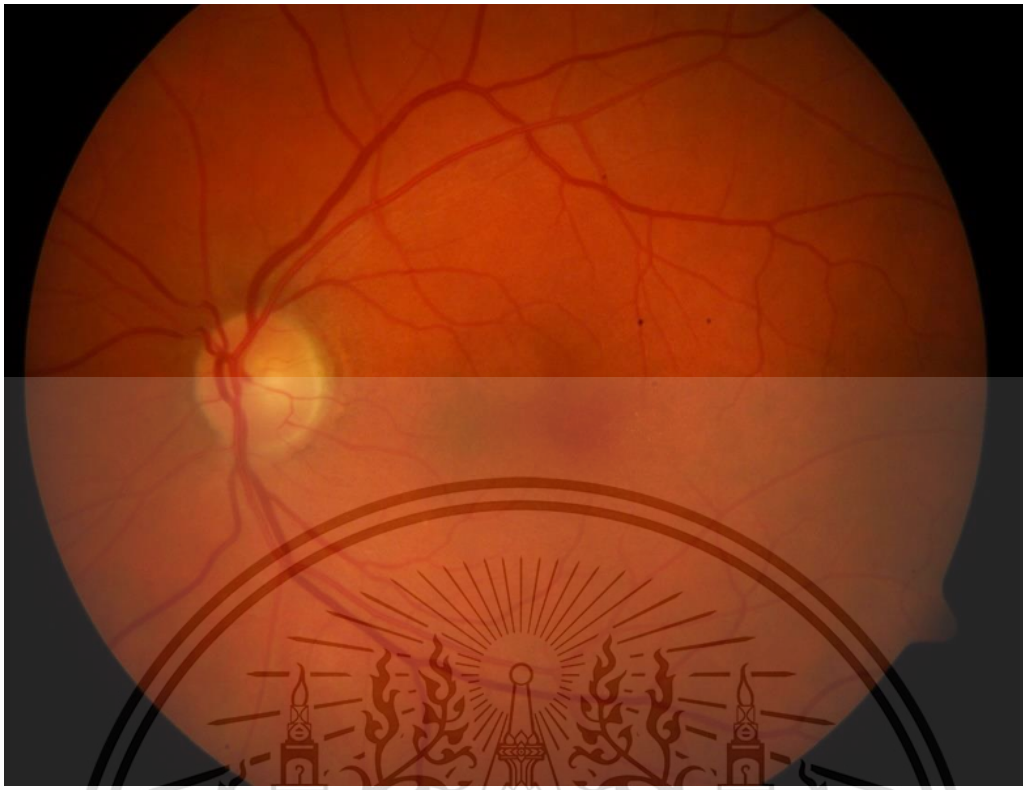


Figure 1.1 Example of a healthy eye image.

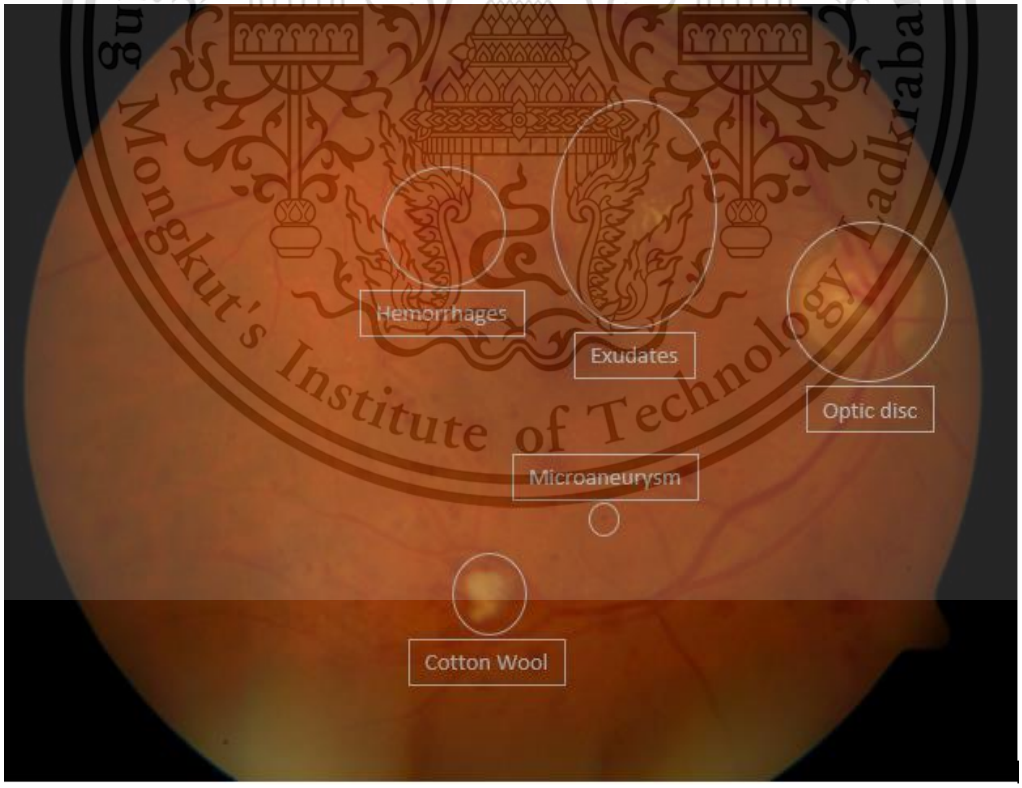


Figure 1.2 Example of cotton wool image

Source: T. Kauppi *et al.*, "DIARETDB1 diabetic retinopathy database and evaluation protocol," Proceedings of Medical Image Understanding and Analysis, Aberystwyth, Wales, 2007.

1.2 Objective of the research

The aim of this study is to satisfy the following objectives:

- To develop an automatic system that takes a raw image as input, extracts optic disc, and detects cotton wool in the image as output.
- To separate the cotton wool image and non-cotton-wool image in order to distinguish between diabetic retinopathy patients and normal patients to save time and effort for doctors.
- To segment regions of cotton wool and non-cotton-wool in the eye images.

Scope of the study:

- The proposed method is tested using global datasets which make our method be comparable with other methods.

1.3 Thesis organization

This thesis consists of five chapters including an introduction, background knowledge and literature review, methodology, experimental results, and discussions. The first chapter is the introduction and the rest are expressed as follows:

Chapter 2 gives the background knowledge of the techniques that was used in this study.

Chapter 3 describes full details of the method. This chapter also explains how to extract the features as well as the way of training and testing processes in a neural network.

Chapter 4 shows results from experiments in the binary image. The results are presented by graphs as well as summarized in tables.

Chapter 5 concludes the study results and shows the author opinions. This chapter also discusses advantages and disadvantages of this research.

Chapter 2

Background knowledge and literature reviews

This chapter provides a short summary of background knowledge to help readers be able to well understand the main content of the proposed system. This also gives a review of existing methods as well.

2.1 Background knowledge

2.1.1 Image processing and feature extraction

In computer science, image processing is a part that makes an analysis or transformation on the digital image [25]. For more information, image processing is a type of signal processing that takes an image as input and extracts the image in different formats or characteristics as output. The goal of image processing focuses on processing a raw image (smoothing, sharpening, contrasting, and so on) to enhance the quality of the image or to highlight the appropriate characteristics.

In image processing, feature extraction is to solve the problem that describes an image by transformation of raw image into a set of most relevant information [7]. Feature set (also known as feature vector) is a set of values that represent the specific characteristics of image. The features extracted from image may completely change the way that we see from the raw image. In fact, most of information of images are usually redundant, only some features are relevant that should be extracted to make decisions based on demands. Features can be divided into low and high level of complexity. Low-

level features which are well-known such as intensity level, shape, size of object. The high-level features are semantic features which present concept of objects based on human knowledge. The use of features is significantly different based on purpose and none of features can guarantee the best recognition performance. For example, color is useful to distinguish between night time and day time landscapes, while circularity is better choice to distinguish between a ball and a tree image. Therefore, feature extraction, which is a process of constructing a set of features, plays an important role in classification and recognition tasks.

2.1.2 Color space and color transfer

Color space is a color organization which visualizes a color by predetermined values and rules for physical devices such as a computer. It reproduces a color in both analog and digital representations. Some of well-known color spaces used in computer vision are RGB, $CIELAB$, HSV and so on.

Color transfer, which is proposed by Reinhard et al. in 2001 [20, 21], is a technique to transfer color casts from an image to other images. Color transfer can remove an undesired color cast as well as shuffle desired color casts in the image. As an explanation, firstly, color transfer analyzes image, extracts desired characteristics from the color channel in both source and destination images. Then, it applies source characteristics to correct the destination image characteristics. The characteristics extraction and correction are described in the formula as follows.

$$L_0 = \frac{\sigma_t}{\sigma_s}(L_s - \mu_S) + \mu_t \tag{2.1}$$

where L represents a two-dimensional plane in the image, s, t and o are the source, target and output images, respectively. μ and σ are mean and standard deviation of source and target images, respectively. For a color image which has three dimensions, the process of color transfer will be repeated for each component plane of the image.

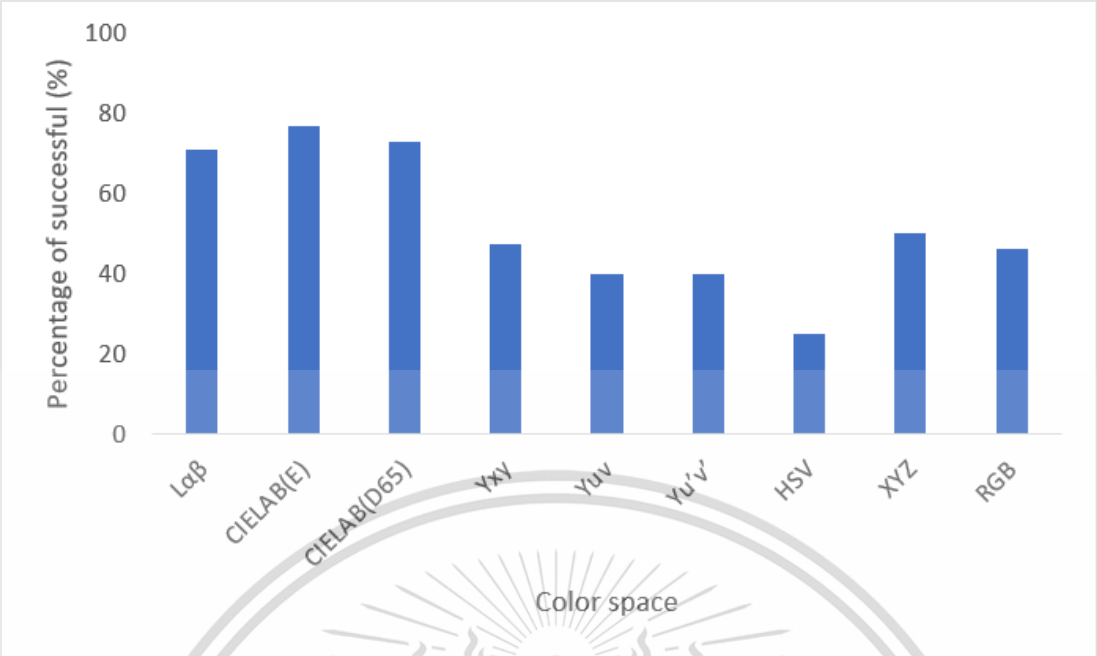


Figure 2.1 Average percentage of successful color transfer of each color space.
Source: E. Reinhard and T. Pouli, "Colour spaces for colour transfer," Proceedings of Computational Color Imaging, Heidelberg, Berlin, pp. 1-15, 2011.

In color transfer, the quality of result is depended on the way that divides the image into three component planes. Each color plane in color space has to be independent to each other. Therefore, in the first publication, $L\alpha\beta$ color space, which proposed by Ruderman et al. in 1998 [23], was the chosen color space to perform color transfer. In 2011, Reinhard and Pouli published that color transfer can be performed in many other color spaces [22]. However, the overall success rates of color transfer are significantly different between color spaces accordingly as presented in Fig. 2.1. The performance graph shows that color transfer between image technique gives the highest performance on CIELAB(E) color space (also known as CIE L*a*b* or sometimes abbreviated as simply “Lab” color space with “E” is the standard illuminant) and the lowest performance on HSV color space.

2.1.3 Maximum entropy

Theory of entropy by Shannon [24], which is concerned with the information theory, states that the probability distribution which best represents the current state of knowledge is the one with largest entropy. Therefore, by maximizing the entropy, we

can maximize the information gain at the solution which fits out prior knowledge but no assumptions beyond what is known. Azarbad et al. [2] apply to image thresholding based on maximum entropy. For a grayscale image, each pixel value is represented by i so the probability of a pixel to be i can be denoted as:

$$P_i = \frac{n_i}{n}; i = 0,1,2 \dots ,255. \tag{2.2}$$

where n is the total of pixels in the image, n_i is the number of pixels that have value i. To calculate a threshold t which maximize entropy value, the image is divided into two regions: background region which is represented by b, and object region which is presented by o. Therefore P_o and P_b are probabilities of object and background, respectively.

$$P_o = \sum_{i=0}^{t-1} P_i \tag{2.3}$$

$$P_b = \sum_{i=t}^{255} P_i \tag{2.4}$$

Entropy of object and background probability distributions are therefore calculated as follows:

$$H_o(t) = -\sum_{i=0}^{t-1} \frac{P_i}{P_o} \log_2 \frac{P_i}{P_o} \tag{2.5}$$

$$H_b(t) = -\sum_{i=t}^{255} \frac{P_i}{P_b} \log_2 \frac{P_i}{P_b} \tag{2.6}$$

thus, the entropy function of the image is:

$$H(t) = H_o(t) + H_b(t) \tag{2.7}$$

The maximum entropy thresholding attempts to maximize $H(t)$ by changing the value t. Therefore, we can find the optimal threshold value t which have the maximum entropy of $H(t)$ as follows:

$$t = Arg\big(max\ H(t)\big);\ t = 1,2,\dots,255. \tag{2.8}$$

2.1.4 Feedforward neural network

Artificial neural networks are models whose functionality is loosely based on the animal brain [6]. To be more specific, they are computing systems which are constructed to learn the activities or characteristics of the data and classify new unknowns. Among neural network models, feedforward neural network, which is common to use for the classification purposes, is the first and simplest type of neural networks.

In feedforward neural network, the input values will be fed in one direction from the input layer, through hidden layer and to output layer of the model. Layers of nodes in feedforward neural network are interconnected in a feed-forward way and there is no cycle or loop (Fig. 2.2). The output is calculated from the input which is fed forwardly through a series of weights. The sum of weights and input is calculated in each node throughout the network.

In the training process, the difference between predicted and actual result responses an error value which can be used to adjust the weight values in connections in the whole model over iterations. The model adjusts the weight values in whole connections to improve accuracy by using an error value. The process will be repeatedly conducted as stated above for whole training data. Finally, the process will be terminated when it meets a stop-condition such as the maximum limit of iterations, maximum running time, or minimum gradient throughout a number of iterations.

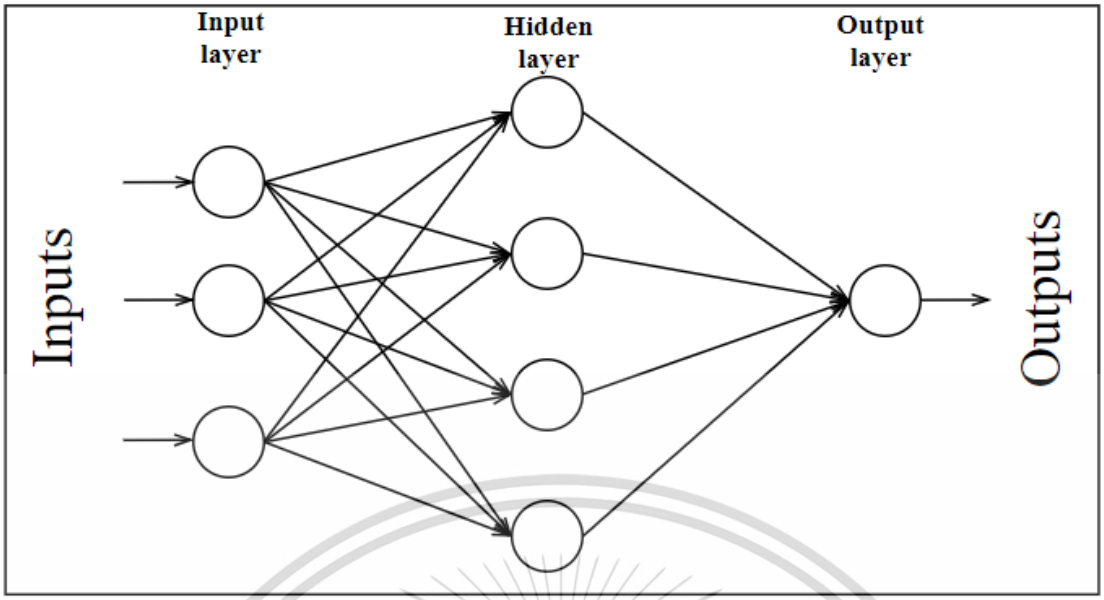


Figure 2.2 A model of neural network

In the testing process, images are fed exactly in the same way of training process. However, weights of the network are no longer changed in this step but the performance values of training, testing (and sometimes validation) process are extracted instead. The testing process, which presents the flexibility of model, since the model is tested by the unknown data, plays the most important role to make the final decision in model evaluation.

2.1.5 Shape descriptor

Shape analysis [4] is a technique that uses geometry of shape in object recognition in computer to interpret and process the object. The goal of shape analysis is to detect the similar shape objects that fit together in the database. Shape analysis is related to statistical analysis of geometric shape. The objects have to be represented in digital form that means to be stored in an electronic device such as computer in order to be detected or recognized. In shape analysis, before any comparison, the given objects have to be simplified which is called a shape descriptor. The purpose of shape descriptor is to ignore the useless information and account all key information to make objects be easier to be handled, to be stored and to be compared. A complete shape descriptor is a representation which can be used to reestablish the original object. In

this study, we employ two shape descriptor parameters which are eccentricity and solidity.

Eccentricity

In an object, major axis is the longest line between two points in the object [18]. Minor axis is the second longest line between two points and perpendicular with major axis. Major and minor axes are represented in Fig. 2.3.

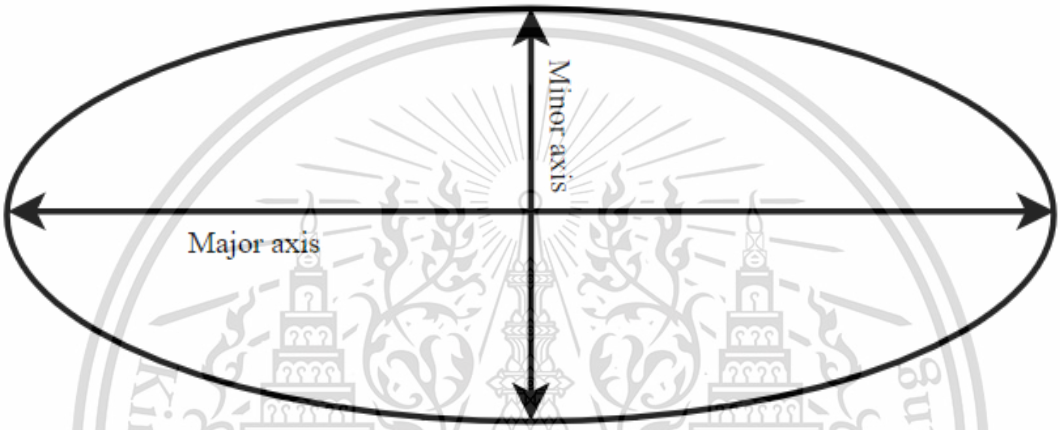


Figure 2.3 Major and minor axes

Eccentricity (also known as a measure of aspect ratio), which is the proportion of the major axis and minor axis, can be computed in the formula as follows:

$$\text{Eccentricity} = \frac{Major\ axis}{Minor\ axis} \tag{2.9}$$

Solidity

For an object, convex hull, which is used to cover and represent an object (convex or non-convex), is the smallest polygon that can contain the whole input object [25]. The convex hull consists of two parts: 1) object regions which belong to the object and convex hull and 2) concave regions which belong to the convex hull but not the object (Fig. 2.4).

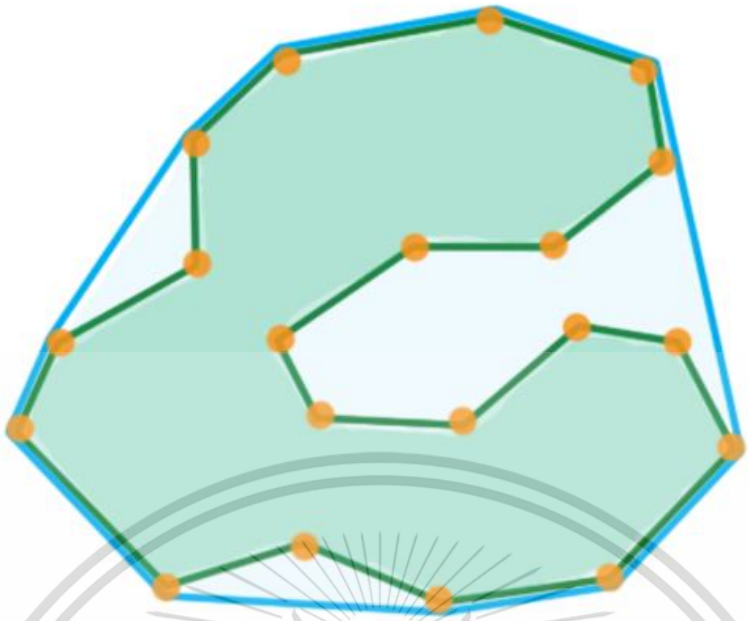


Figure 2.4 Convex hull of an object

Source: Video game physics tutorial -part II: Collision detection for solid object. Retrieved from http://www.toptal.com/game/video-game-physics-part-ii-collision-detection-for-solid-objects.

Solidity, which is the proportion of the total pixels of the object regions and the total pixels in convex hull region, can be computed as follows:

$$Solidity = \frac{Object\ area}{Convex\ hull\ area} \tag{2.10}$$

2.2 Literature reviews

Cotton wool detection is a part of diabetic retinopathy detection, which is used to detect bright lesions of diabetic retinopathy in order to prevent blindness. For detection of bright lesions which are exudates and cotton wool, optic disc is a trouble, since optic disc is very bright and has nearly the same color as cotton wool and exudates. Therefore, Sreng et al. [26] proposed to use the maximum entropy technique to remove the optic disc and detect exudates in the image. The result of the technique is tested to work with 100 images from Bhumibol Adulyadej hospital and shows 93%

accuracy of optic disc detection. Therefore, the technique is sufficient to detect the location of optic disc.

Harini and Sheela [9] applied Fuzzy C-Mean and morphological operation to detect exudates and microaneurysm. Firstly, blood vessel is extracted using Fuzzy C-Mean and morphological operation. Then, optic disc is removed by finding the brightest point in image and calculating the circle. After that, microaneurysm is extracted using AND operation and contrast-limited adaptive histogram equation (CLAHE). Exudates are detected using a set of features such as contrast, correlation, energy, homogeneity. Finally, a model of support vector machine is applied for classification. Experiments in this study are done using 75 images selected from DIARETDB0 and DIARETDB1, 45 images for testing and 30 images for training. The final result shows that sensitivity, specificity, and accuracy are 100%, 95.83%, and 96.67%, respectively. The study shows a good performance result and the size of data is kind of large.

Kaur et al. [14] proposed a supervised approach using decision tree to detect hemorrhage in retinal fundus images. Firstly, all images were resized to the same size. Then, blood vessels and fovea, which is the red center region in retinal fundus image, is removed since they have the same color with hemorrhage. After that, hemorrhage is extracted using Otsu's thresholding and feature extraction such as local range, standard deviation, mean from green component. Finally, a model of decision tree is employed for classification. Experiment in this study is done using 50 images selected from DIARETDB0. Performance shows that sensitivity and specificity values are 90.42% and 93.53%, respectively. the study can be considered as efficient since a large number of images from a global data is used and the final result shows effective performance.

Irshad et al. [11] implemented Gabor filter bank and global thresholding to detect cotton wool spots. At the start, the system eliminates blood vessel using morphological closing operation. Then, the system employs Gabor filter [5] to enhance the quality of candidate region, analyze the direction of texture, and determine the frequency of object. Cotton wool and other lesions contain more variations which make them have high frequency. After that, global thresholding technique was applied to binarize the images. Evaluation of the study is based on the detected blobs. If a detected blob has cotton wool, it will be true positive. If a detected blob does not have cotton wool, it will be false positive. If a blob of lesion cannot be detected, it will be false negative. True negative is not considered in this research. As the result, the overall

performance (also known as accuracy) cannot be computed. The study employs 9 cotton wool images taken from AFIO hospital and shows that sensitivity and positive predicted values are 82.21% and 82.23%, respectively. The dataset can be considered as small since it consists only 9 cotton wool images and does not have normal images. Besides, the private data is also a weakness since it makes a difficulty for comparison.

Hashim and Hashim [10] proposed a region-based approach to classify abnormal and normal regions. Firstly, the proposed method employs Haralick's texture features [8] for feature extraction. Then, support vector machine and multi-layer perceptron were independently experimented to classify the regions. Because the method takes regions as input, researchers had to manually crop image using Photoshop. Consequently, the system is not fully automatic, since the optic disc was removed manually from regions. The performance of system was tested on 930 selected regions (487 normal regions and 443 abnormal regions) taken from DIARETDB01, DRIVE and STARE datasets. Support vector machine shows that performances of sensitivity, specificity, and accuracy are 82.39%, 62.42%, and 71.94%, respectively. Multi-layer perceptron shows that performance of sensitivity, specificity, and accuracy are 71.78%, 76.49, 74.53%, respectively. Multi-layer perceptron seems to be better in accuracy while support vector machine has a better sensitivity value. Performance of study is not good enough since most of performance values are lower than 80%.

Chapter 3

Proposed method

This study proposes an automatic method to recognize cotton wool in retinal fundus images. This method consists of five main processes (Fig. 3.1): 1) Pre-processing which removes redundant border, scaling the image resolution, smoothing image quality, 2) optic disc extraction which uses an adjustment of maximum entropy to remove optic disc, 3) feature processing consists of two sub-processes: color transfer which applies color cast between images and feature extraction which build a feature set from image, 4) cotton wool classification which uses a pre-built feature set as characteristic extraction and a feed-forward neural network as a classifier, 5) post-processing which employs shape description to reduce misclassification. Each process is fully explained in details in the following sections.

3.1 Pre-processing

An original raw image (Fig. 3.2) sometime may contain noises and unnecessary small details. Therefore, a pre-processing step is quite useful and necessary to improve quality and smooth images.

Firstly, we detect the location of eye. RGB input images were decomposed into red plane, green plane and blue plane. In the green plane, we build a mask image (Fig. 3.3) by manually thresholding at value 10 in range from 0 to 255. Then coordinates of leftmost, rightmost, topmost, and bottommost white pixel is detected.

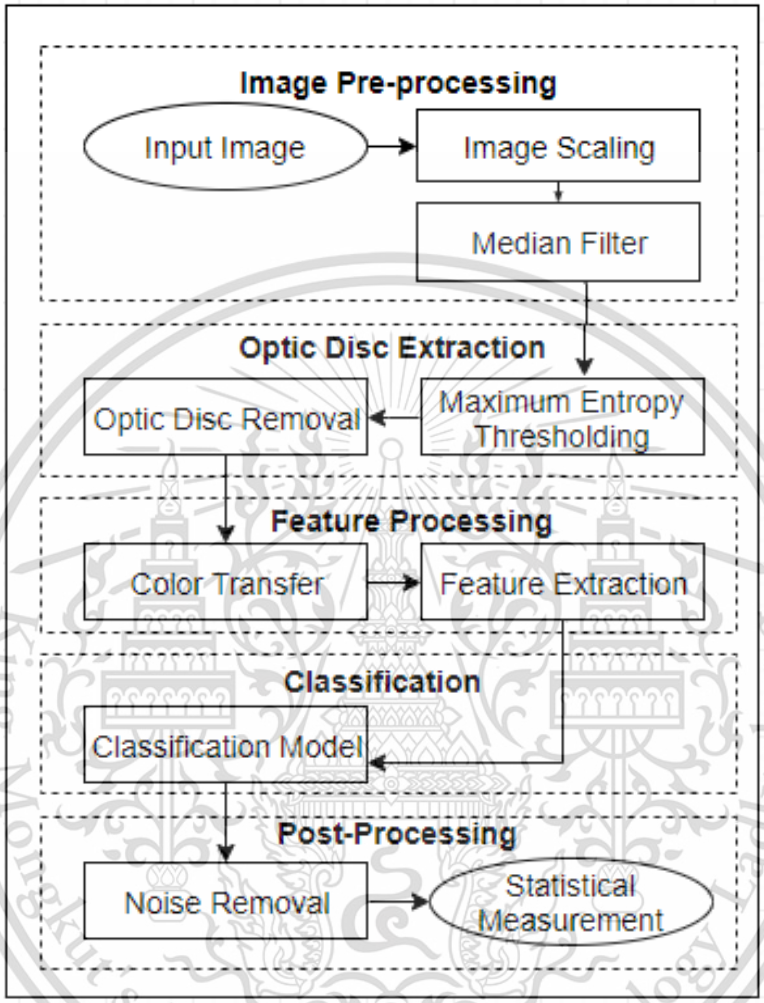


Figure 3.1 Overview of the system.

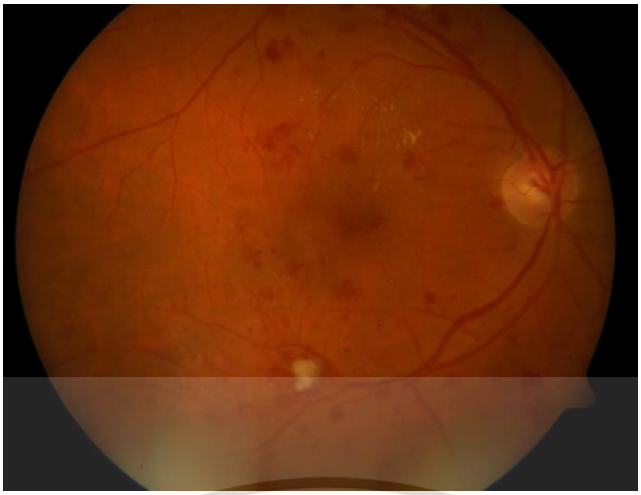


Figure 3.2 Raw image.

Figure 3.3 Example of mask image.

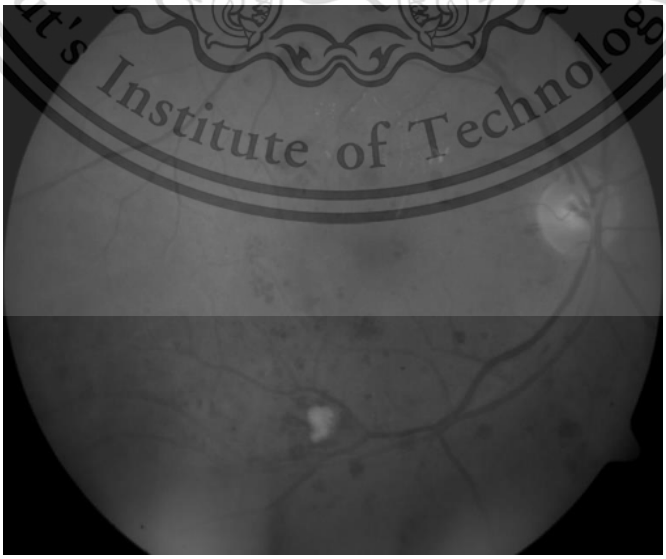


Figure 3.4 Image after preprocessing.

The black empty spaces between eye and border of the image were removed based on those pixel coordinates (Fig. 3.4), since there is no necessity in that region.

Secondly, salt-and-pepper is the most common noise which usually appears in low quality devices while recording images. Therefore, we apply a smoothing technique, which is called 2D median filter [16], in each plane of image to guarantee that the noises will be completely removed. The median filter can be expressed as follows:

$$f(x,y) = \underset{(m,n)\in W_{xy}}{median} \{g(m,n)\} \quad (3.1)$$

where $f(x,y)$ is the output for each given coordinate x,y from the median filter, m, n are coordinates of pixel g in the window W_{xy} which has size 3×3.

In addition, since the source images may come from different data with different resolutions. To make images be easier to be handled and the method be adaptive to different situations, we rescale all images into the same resolution which is manually chosen at 1200×1000 in this study.

3.2 Optic disc extraction

In human eyes, optic disc is a component that is responsible to receive visual information from object and transmits to brain. In the retinal fundus image, optic disc is the largest yellowish component. Optic disc is very similar to cotton wool in color which make optic disc be misclassified with cotton. Therefore, we need an additional step to remove optic disc before cotton wool classification to prevent misclassifying. Fortunately, optic disc is usually the biggest white-yellow object in eye images. For that reason, we take optic disc size as an advantage in order to detect and remove the optic disc.

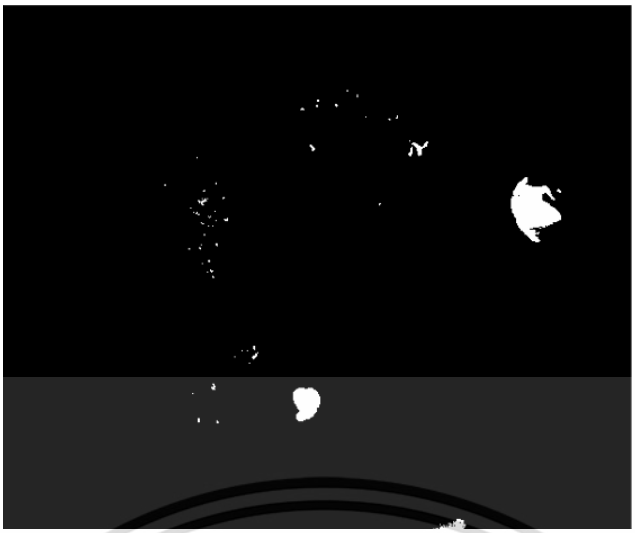


Figure 3.5 Maximum entropy thresholding.

Figure 3.6 Optic disc location.

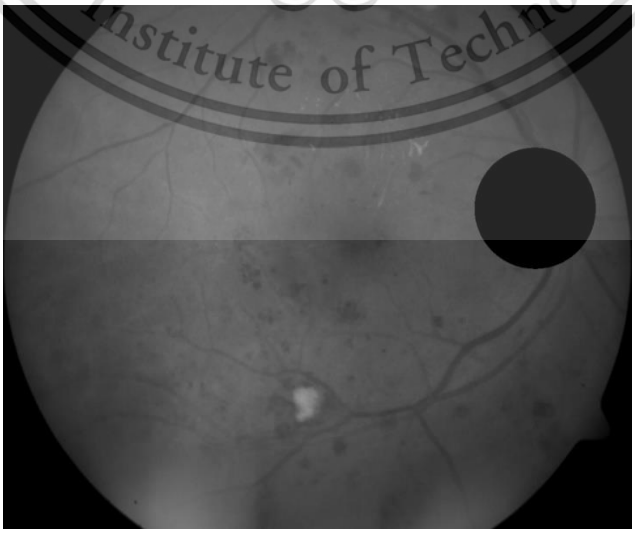


Figure 3.7 Image after optic disc removal.

This study employs a statistic technique called image thresholding based on the maximum entropy [2, 26] in order to detect the location of the optic disc. The calculation of maximum entropy technique was explained earlier in section 2.1.3. After employing the technique, optic disc is the largest white region after thresholding (Fig 3.5). The position of optic disc is detected and marked by a white circle (Fig 3.6) which has the same centroid with the optic disc region and radius is manually chosen at 37 pixels (fixed with the image resolution at 1200 ×1000). Then, the marked white region is used to remove optic disc in retinal fundus image (Fig. 3.7).

3.3 Feature processing

In fact, a classification model does not learn in the same way as human brain does. Therefore, the model cannot understand image directly and may need some help to do that. As a result, a feature processing, that can extract useful features before feeding images into the classification model, plays an important role. This process consists of two main steps: 1) color transfer between images which is used to remove undesired color casts and 2) feature extraction which is used to extract image characteristics.

3.3.1 Color transfer:

A technique called color transfer [20-22] is applied to make color characteristics between image to be consistent. As mentioned earlier in section 2.1.2, since the overall performance of color transfer is significantly different depending on the chosen color space, this study applies the color transfer technique in CIELAB(E) color space that gives the best performance. CIELAB(E) (also known as CIE L*a*b* or sometimes abbreviated as simply "Lab" with "E" is the standard illuminant) is a color space which consists of three component planes: L^*, b^*, a^*. An image is chosen to be the target image (I_t) to lend color casts. Firstly, the target image (I_t) and source image (I_s) are converted into CIELAB(E) color space. Then mean and standard deviation of each color channel in the source image and target image are computed. After that, the color transfer uses the computed mean and standard deviation to calculate the new pixel values of source image in each color channel. The process can be expressed in the formula as follows and in Fig 3.8:

$$l_o = \frac{\sigma_t^l}{\sigma_s^l}(l_s - \mu_s^l) + \mu_t^l \tag{3.2}$$

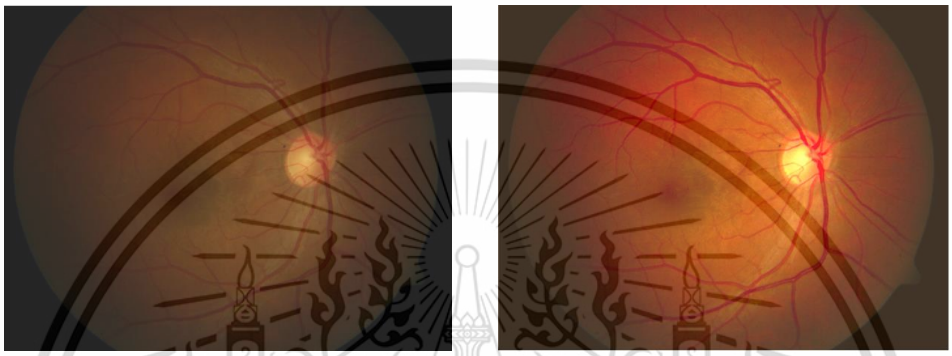
$$b_o = \frac{\sigma_t^b}{\sigma_s^b}(b_s - \mu_s^b) + \mu_t^b \tag{3.3}$$

$$a_o = \frac{\sigma_t^a}{\sigma_s^a}(a_s - \mu_s^a) + \mu_t^a \tag{3.4}$$

where l, b, a are individual pixels in L^*, b^*,a^* planes, respectively. o, s and t are output, source and target images, respectively and σ and μ are their respective standard deviation and mean.

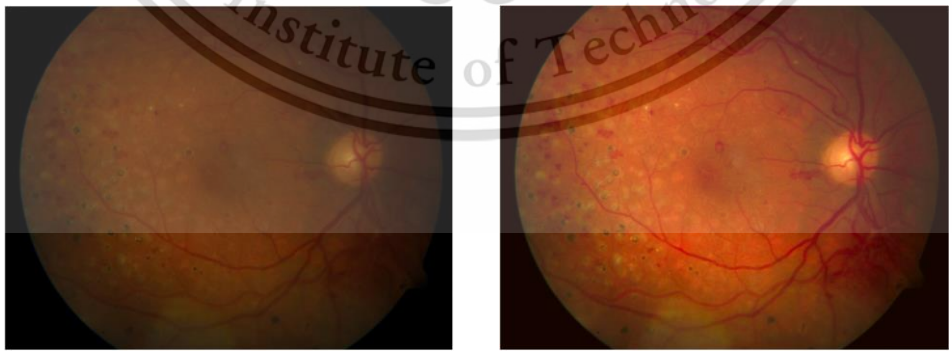
Finally, the source image is converted back to RGB color space to be prepared for the feature processing.

a) Target image



b) Source image c) Output image

d) Source image e) Output image



f) Source image g) Output image

Figure 3.8 Color transfer between images.

3.3.2 Feature extraction:

We calculate the following nine features on pixel-based in order to extract information from each pixel in the image.

- Red component: In RGB color space, it is an elemental plane which represents intensity of red color in image (Fig. 3.9).

Figure 3.9 Image on red plane.

- Green component: In RGB color space, it is also known as an element plane which represents intensity of green color in image (Fig. 3.10).

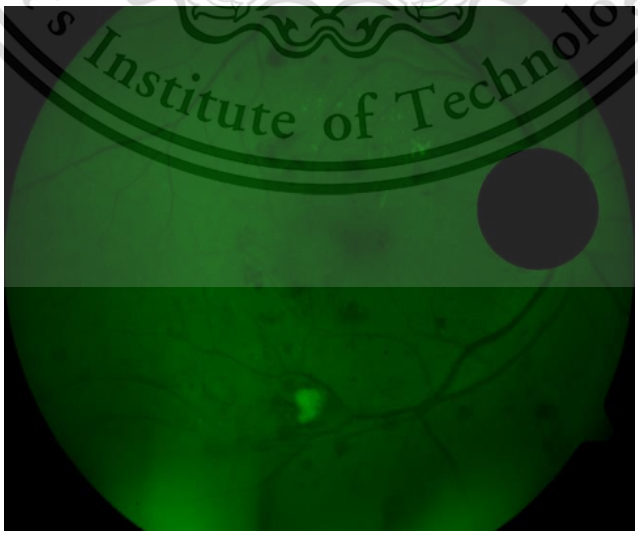


Figure 3.10 Image on green plane.

- Intensity: It is a transformation from RGB color space into 2-D plane which describes the average of intensity of the three-color planes and can be expressed as follows:

$$I(x,y) = 0.299 \times R(x,y) + 0.587 \times G(x,y) + 0.114 \times B(x,y) \quad (3.5)$$

where I is the intensity output value, x, y are coordinate pixel, R, G and B are red, green and blue planes, respectively.

- Brightness: It is a value which can illustrate the human visual perception. It reflects the source of light in the image and is expressed as follows:

$$Y(x,y) = 0.213 \times R(x,y) + 0.715 \times G(x,y) + 0.072 \times B(x,y) \quad (3.6)$$

where Y is the brightness output value, x, y are coordinates, R, G and B are red, green and blue planes, respectively.

- Saturation plane: In HSI color space, it is a component plane which can be transformed from RGB color space as follows:

$$S(x,y) = 1 - \frac{3}{R(x,y) + G(x,y) + B(x,y)} [\min(R(x,y), G(x,y), B(x,y))] \quad (3.7)$$

where S is saturation output value, x, y are coordinates, R, G and B are red, green and blue plane, respectively.

- Local range from intensity: in a grayscale image, each pixel and its neighbors in window size 3×3 are employed in order to calculate the difference between maximum and minimum pixel value as follows:

$$L(x,y) = \max_{(m,n) \in W_{xy}} \big(I(m,n)\big) - \min_{(m,n) \in W_{xy}} \big(I(m,n)\big) \quad (3.8)$$

where L is the local range output value of pixel coordinate x, y, m, n are coordinates pixel in intensity in window W_{xy} which has size 3 × 3, max, min are the maximum and minimum, respectively.

- Local mean from intensity: for every individual pixel, we compute the mean between the pixel and its neighbor pixels in the window size 3 × 3. The computing follows the equation as below:

$$\bar{I}(x,y) = \frac{1}{N} \sum_{(m,n)\in W_{xy}} I(m,n) \tag{3.9}$$

where $I(m,n)$ is the value of each pixel in intensity in window W_{xy} which has size 3 × 3 and N is the total number of pixels in window W_{xy}.

- Local variance from intensity: after mean calculation, we compute the local variance by the pixel values and the mean in window size 3 × 3 as follows:

$$V(x,y) = \frac{1}{N-1} \sum_{(m,n)\in W_{xy}} \left(I(m,n) - \bar{I}(x,y)\right) \tag{3.10}$$

where I and $\bar{I}$ is pixel value and local mean value in intensity in window W_{xy} which has size 3 × 3, N is the number of pixels in window W_{xy}.

- Local entropy from intensity: each pixel in grayscale is used to calculate the entropy value based on window size 9 × 9 as follows:

$$H(x,y) = -\sum_{k=0}^{255} P_k(x,y) log_2 P_k(x,y) \tag{3.11}$$

$$P_k(x,y) = \frac{C_k(x,y)}{N}; k = 0,1,2 \ldots,255. \tag{3.12}$$

where C_k is the total number of pixels which have value k in window W_{xy} corresponding to coordinate x, y and N is the total number of pixels in window W_{xy}, P_k is the fraction which describes the probability of pixels has value k ($0 \leq P_k \leq 1$), H is the entropy value.

3.4 Classification

This study employs a feedforward neural network model as a classifier in order to learn from information of feature and compute a threshold to classify unknown data. The model consists of one input layer, one hidden layer, and one output layer. The layers in model are interconnected and each output axon, which is the total net input, can be computed as follows:

$$X_{total} = \sum_{i=1}^{N} (w_i x_i) + w_b . b \tag{3.13}$$

where X_{total} is the output in the axon as well as the input for activation function, w are the weights axon, x are input values from data, b is bias and w_b is the weight corresponding to bias.

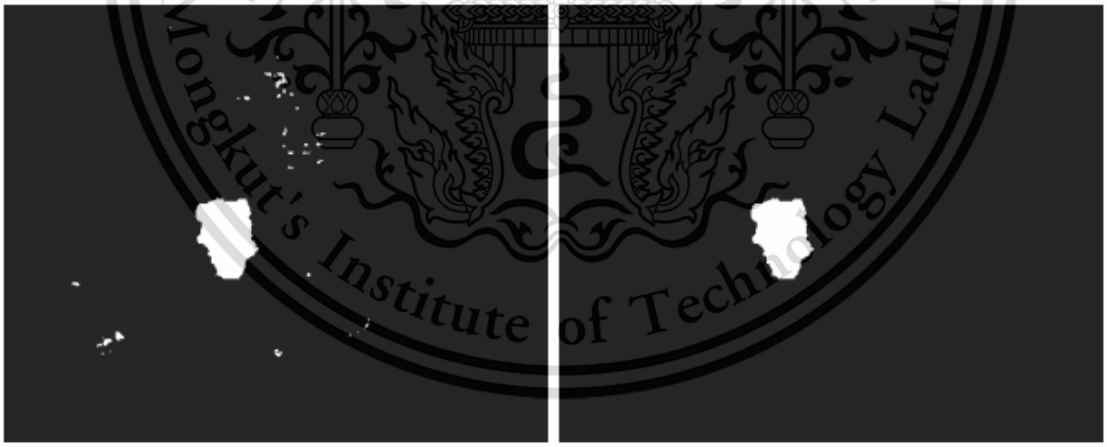
In the task of classification, the output values are typically discrete. Therefore, there are two activation functions which are recommended such as log-sigmoid function (also known as Fermi function) and tan-sigmoid function (also known as hyperbolic tangent sigmoid function). Fermi function generates the output value in the range of 0 and 1, whereas the tan-sigmoid function generates the value from -1 to 1. Because of that reason, the essential difference between range value makes the log-sigmoid function more difficult than tan-sigmoid function, where its results are close to 0, to learn something far from the threshold value [19]. Hence, the tan-sigmoid function is chosen as activation function for our model. For activation function at each node, the total axon, which is the output in Eqs. (3.12), is the input of following equation:

$$F(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{3.14}$$

where $F(x)$ is the output of each node, x are the output from the axon. The predicted outputs from the model are compared with the actual output taken from ground truth to calculate performance.

3.5 Post-processing

After the classification step, a binary image, which consists of cotton wool and noises, is formed. A post-processing step is therefore necessary to reduce misclassified points and improve performance of the system. Based on the characteristics of cotton wool, which are fat and short in shape and usually bigger than exudates and noises in size, we employ a simple noise reduction based on area which is the number of pixel in each object. In this study, blobs, which have area less than 200 pixels, are removed. Finally, two shape descriptor parameters, which are eccentricity and solidity, are conducted.

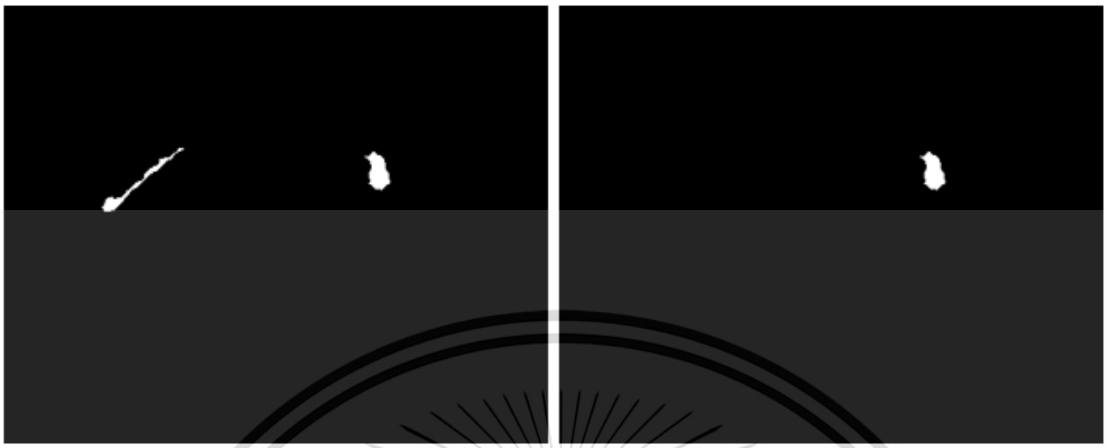


a) Before elimination by size b) After elimination by size

Figure 3.11 Application of size for noises elimination.

By observation, it is clear to notice that some points of noise are long and thin while cotton wool spots are commonly thick and short as stated above. Therefore, eccentricity, which is quite powerful to separate thin and short objects, is applied to distinguish cotton wool spots and noises. As the result, the value of aspects ratio, which

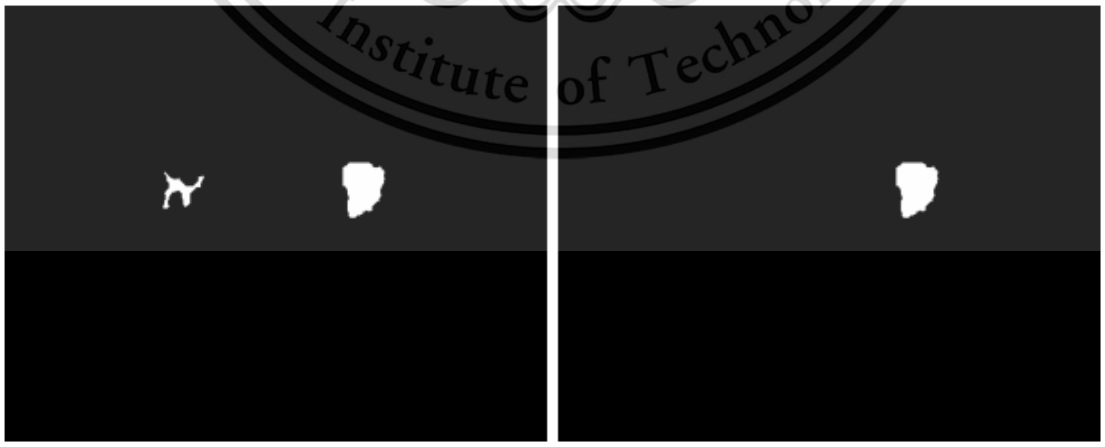
is manually chosen as 2.2 in this case, can partially separate the difference between noises and cotton wool and help to improve the performance (Fig. 3.6).



a) Before eccentricity process b) After eccentricity process

Figure 3.12 Application of eccentricity for noises elimination.

Equally important, some points of noise may have particularly unpredictable free shape, which makes it pass the test of eccentricity. Therefore, another process called solidity analysis, which can calculate the convexity of the object, was employed. Cotton wool is a chubby object with less concave, while noises have many more concaves inside. For that reason, the convexity can classify between cotton wool and noises in order to improve performance. In this study, the value of solidity was chosen manually as 0.6 (Fig. 3.7).



a) Before solidity process b) After solidity process

Figure 3.13 Application of solidity for noises elimination.

Chapter 4

Results and discussion

This chapter shows the results and explains the way of performance analysis. The information about dataset is also supplied and an explanation of the way how dataset is designed in experiment. The performance calculation is stated up in two different ways which can supply a comprehensive view for this study and the effects of the proposed method as well.

4.1 Experimental setup

The experiments in this research were conducted on 60 images which are selected from two public datasets: DIARETDB01 [13] and MESSIDOR [3]. These two datasets, which are well-known for researches in diabetic, will contribute more reliability for the result of this study.

The ground truth for the dataset was made manually using Photoshop CS6, since provided ground truth from DIARETDB01 [13] dataset does not completely fix to the cotton wool regions and MESSIDOR [3] dataset does not provide any ground truth.

In this study, we test the method using k-fold cross-validation. The convention of k-fold cross validation usually uses 10-folds (k equal 10). However, each fold will become too small which is unnecessary and inconvenient. Therefore, the 60 images were divided into two-fold for cross-validation (1-fold for training and 1-fold for testing). Each fold consists of 30 images: 15 images of cotton and 15 images of healthy

eye. Training fold is used to train the system and testing fold is to calculate the performance. The training and testing fold are later inverted to each other and the whole process is performed again to make the experiment result become more realizable.

4.2 Measure

The results of this study were separately evaluated by two ways which decide unit of value:

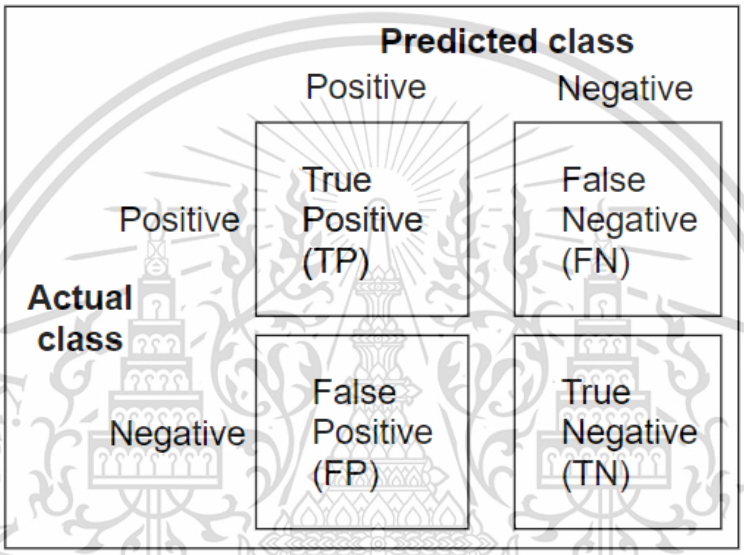


Figure 4.1 Confusion Matrix.

Image-based: used to measure the success of classification between cotton wool images and non-cotton-wool images. Therefore, cotton wool images will be positive and non-cotton-wool images will be negative values.

Pixel-based: used to measure the success of segmentation between cotton wool regions and non-cotton-wool regions in an image.

Therefore, true positive, true negative, false positive, and false negative values will be defined in Fig 4.1 and as follows:

- True Positive (TP): It is cotton wool image/pixel and is detected as cotton wool images/pixel.
- True Negative (TN): It is non-cotton-wool image/pixel and is detected as non-cotton-wool image/pixel.
- False Positive (FP): It is non-cotton-wool image/pixel and is detected as cotton wool image/pixel.

- False Negative (FN): It is cotton wool image/pixel and is detected as non-cotton-wool image/pixel.

The classification performance of this study is indicated by three factors:

- Accuracy: It is used to indicate the overall performance, which can provide the overall look of performance based on all related values.
- Sensitivity: It is used to indicate the performance of true positive classification which represents the proportion of correctly identified true positive value over predicted true positive value.
- Specificity: It is used to indicate the performance of true negative classification which represents the proportion of correctly identified true negative value over predicted true negative value.

The three factors above are calculated as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4.1}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{4.2}$$

$$Specificity = \frac{TN}{TN + FP} \tag{4.3}$$

where TP, TN, FP, FN are true positive, true negative, false positive and false negative values, respectively.

4.3 Result evaluation

Experimental result in this study is evaluated in two different methods: 1) image-based evaluation which classifies diabetic retinopathy images and non-diabetic retinopathy images based on cotton wool and 2) pixel-based evaluation which segments the position of cotton wool in diabetic retinopathy fundus images.

4.3.1 Image-based evaluation

Based on the first objective of the research, we compute the result performance on image-based which separates the cotton wool images and non-cotton-wool images. The result of from 60 images (Table 4.1), which were divided into training and testing cases and inversely, shows that 96.66% for accuracy of images can be classified correctly. With 100% for sensitivity, which means all cotton wool images have a correct classification, we are quite confident that all the potential diabetic patients will meet the appropriate treatment. And 93.33% for specificity, which shows the performance of recognition of normal patients, proves that the proposed method truly helps to reduce a huge effort, time and money for doctors, patients, and relevant medical staffs from unnecessary diagnoses later.

Table 4.1 Performance on image-based

	TP	FP	TN	FN	Sensitivity	Specificity	Accuracy
Fold 1	15	2	13	0	100.00	86.66	93.33
Fold 2	15	0	15	0	100.00	100.00	100.00
Average					100.00	93.33	96.66

4.3.2 Pixel-based evaluation

Among the cotton wool images, we compute the performance of segmentation between cotton wool and non-cotton-wool regions. For a cotton wool image, we compute the performance on pixel-based, which means the system detects the region of cotton wool based on each marked pixel in ground truth images. Ground truth images are binary images whose values are either 1 or 0 corresponding to cotton wool regions in the raw images, 1 means cotton wool pixel and 0 means normal pixel. The samples of classification results and their ground truth are shown in binary images and in RGB images (Fig. 4.2 and Fig. 4.3).

By observation, we can see that the performances are changed image-by-image, since the complexity of cotton wool and noises in size and shape are very diverse. The average performance for fold 1 shows that sensitivity, specificity, and accuracy are 88.97%, 99.70%, and 99.68% (Fig 4.4 and Table 4.3), respectively. The average performance for fold 2 shows that sensitivity, specificity, and accuracy are 94.18%, 99.62%, and 99.60% (Fig 4.5 and Table 4.4), respectively. The final average performance shows that, 91.57% of cotton wool region can be segmented correctly,

99.39% of non-cotton-wool regions were truly classified, and 99.36% is the average accuracy between cotton wool and non-cotton-wool (Table 4.2). The proposed method obviously outperforms the previous works and proves that it is a reliable and efficient method since all average performance value is better than 90%.

Table 4.2 Performance comparison.

Method	Sensitivity	Specificity	Accuracy
SVM by Hashim et al. [10]	82.39%	62.42%	71.94%
MLP by Hashim et al. [10]	71.78%	76.49%	74.53%
Proposed method	91.57%	99.39%	99.36%

(a) Classification region (b) Ground truth c) Image result.

Figure 4.2 Example of image result and ground truth.

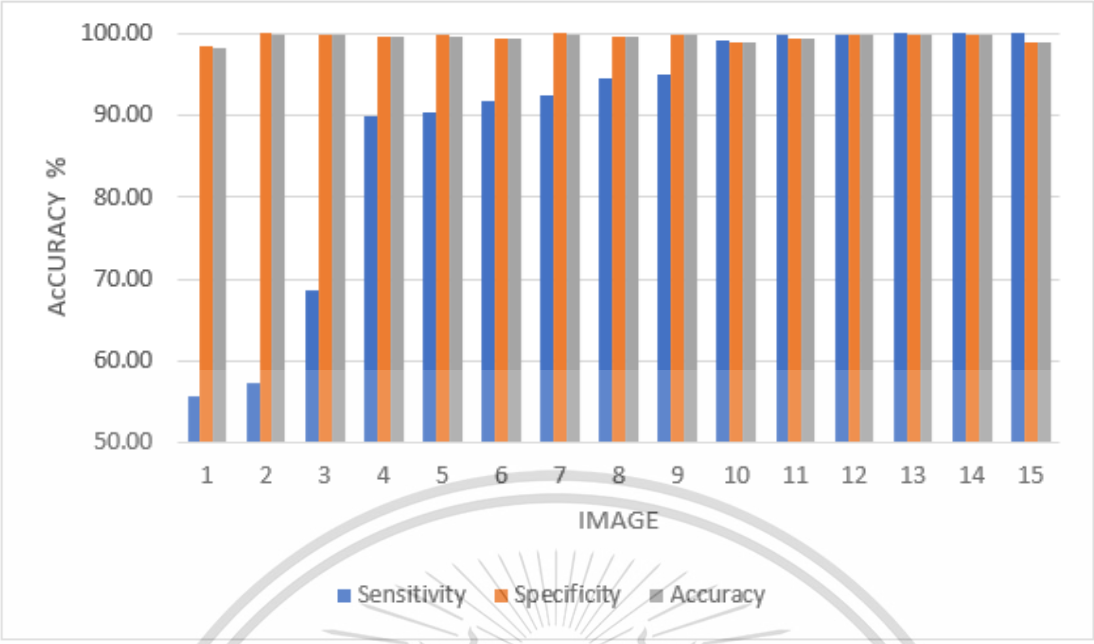


Figure 4.3 Performance on pixel-based for 15 cotton wool images of fold 1.

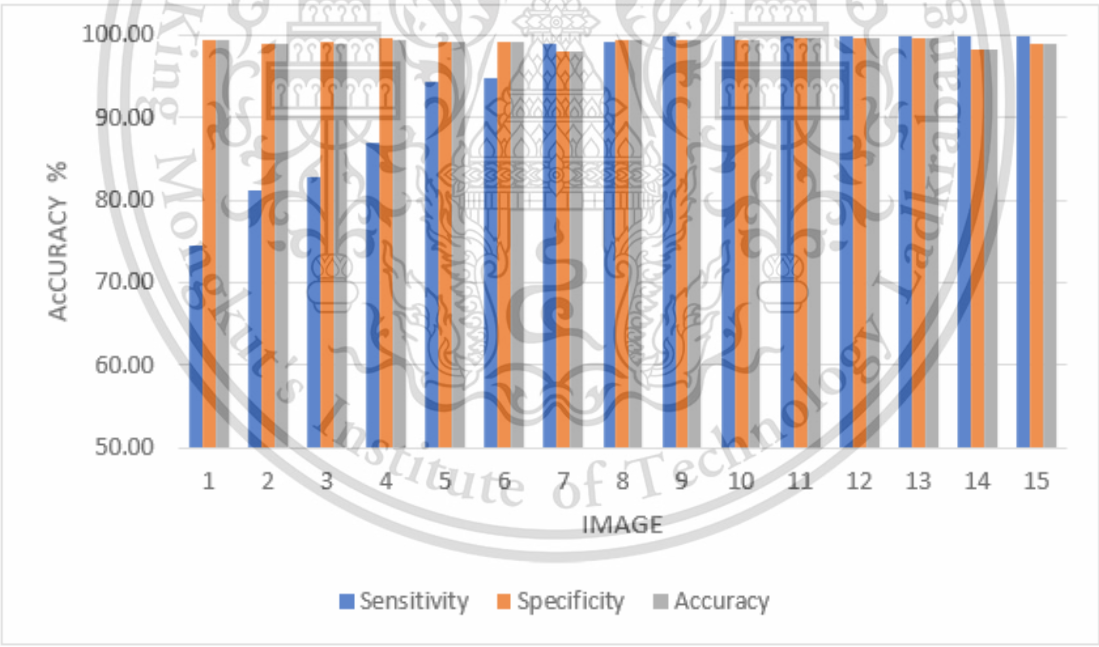


Figure 4.4 Performance on pixel-based for 15 cotton wool images of fold 2.

Table 4.3 Performance on pixel-based for 30 images of fold 1.

Image No.	Sensitivity	Specificity	Accuracy
1	55.69	98.45	98.09
2	57.26	100.00	99.92
3	68.71	99.78	99.76
4	89.95	99.56	99.53
5	90.41	99.72	99.71
6	91.78	99.32	99.31
7	92.42	99.96	99.94
8	94.52	99.55	99.54
9	94.97	99.78	99.75
10	99.20	98.79	98.79
11	99.73	99.40	99.40
12	99.85	99.72	99.72
13	100.00	99.89	99.89
14	100.00	99.88	99.88
15	100.00	99.01	99.01
16	N/A[1]	99.32	99.32
17	N/A	99.81	99.81
18	N/A	100.00	100.00
19	N/A	100.00	100.00
20	N/A	100.00	100.00
21	N/A	100.00	100.00
22	N/A	100.00	100.00
23	N/A	100.00	100.00
24	N/A	100.00	100.00
25	N/A	100.00	100.00
26	N/A	100.00	100.00
27	N/A	100.00	100.00
28	N/A	100.00	100.00
29	N/A	100.00	100.00
30	N/A	100.00	100.00
Average	88.97[2]	99.70	99.68

[1] In non-cotton-wool images (image 16 to 30), there is not cotton wool pixel. Therefore, sensitivity cannot be calculated.

[2] The average sensitivity in this table is calculated from image 1 to 15.

Table 4.4 Performance on pixel-based for 30 images of fold 2.

Image No.	Sensitivity	Specificity	Accuracy
1	74.40	99.51	99.45
2	81.14	99.10	99.05
3	82.79	99.25	99.05
4	87.09	99.59	99.55
5	94.41	99.16	99.14
6	94.75	99.22	99.21
7	99.07	98.13	98.13
8	99.18	99.35	99.35
9	99.86	99.41	99.42
10	100.00	99.58	99.58
11	100.00	99.68	99.68
12	100.00	99.78	99.78
13	100.00	99.68	99.68
14	100.00	98.40	98.40
15	100.00	99.03	99.04
16	N/A[1]	100.00	100.00
17	N/A	100.00	100.00
18	N/A	100.00	100.00
19	N/A	100.00	100.00
20	N/A	100.00	100.00
21	N/A	100.00	100.00
22	N/A	100.00	100.00
23	N/A	100.00	100.00
24	N/A	100.00	100.00
25	N/A	100.00	100.00
26	N/A	100.00	100.00
27	N/A	100.00	100.00
28	N/A	100.00	100.00
29	N/A	100.00	100.00
30	N/A	100.00	100.00
Average	94.18[2]	99.62	99.60

[1] In non-cotton-wool images (image 16 to 30), there is not cotton wool pixel. Therefore, sensitivity cannot be calculated.

[2] The average sensitivity in this table is calculated from image 1 to 15.

Chapter 5

Conclusions and future works

This thesis presents a cotton wool detection method, which can classify the cotton wool and non-cotton-wool images and segment the regions of cotton wool in cotton wool images. While the classification on images-based helps to know whether a patient is positive with diabetic retinopathy, the cotton wool segmentation on pixel-based is useful to help doctor partially detect the spot of cotton wool and improves the accuracy of detection. In order to reach the stated purposes, image was firstly pre-processed to remove noises. Then, color transfer technique was employed onto each image before extracting useful feature such as green, red and gray planes. Next, a classification model, which is a feed-forward neural network, was used to feed the whole features in order to train and test. Last but not least, a post-processing step, which applies shape description technique, was used to remove noises and improve the performance accuracy.

Experiments were done using 60 images which are selected from 2 public datasets, DIARETDB01 and MESSIDOR. The result indicates that the method can work effectively for all of objectives. The first objective is fully reached since the proposed method can automatically remove optic disc from the raw image. The second objective, which is important to reduce the time of doctors and relevant medical staffs from normal patients, is also successfully achieved. Finally, the experimental results have shown a reliable performance for the third objective, which is helpful to reduce the effort to locate the position of cotton wool from the doctors and to improve

diagnostic accuracy. In addition to that, we can also apply this method for other similar diseases in eye images with a minor modification.

Cotton wool detection is a part of diabetic retinopathy detection system, which is designed to work with whole symptom of diabetic retinopathy. Notably, the cotton wool detection method gives an efficient result itself and maybe even more accurate in the combination system.

There are still plenty of techniques which can be used to improve cotton wool segmentation performance in future work. With 91.57 % sensitivity, there is a gap that we need to fill. Besides, after combining the whole diabetic retinopathy detection method, few methods may be conflicted to each other or the whole method become more complexity which may take more time for processing system. Additionally, more data is needed to test comprehensively the whole method.

Bibliography

[1] "Global report on diabetes," World Health Organization, 2016, available at: http://www.who.int/diabetes/global-report/en/, last accessed on: 24.05.2018.

[2] M. Azarbad, A. Ebrahimzade, and V. Izadian, "Segmentation of infrared images and objectives detection using maximum entropy method based on the bee algorithm," International Journal of Computer Information Systems and Industrial Management Applications (IJCISIM), vol. 3, pp. 26-33, 2011.

[3] E. Decencière, X. Zhang, G. Cazuguel, B. Lay, B. Cochener, C. Trone, et al., "Feedback on a publicly distributed image database: The MESSIDOR database," Image Analysis and Stereology, vol. 33, pp. 231-234, 2014.

[4] L. D. Floriani and M. Spagnuolo, Shape Analysis and Structuring, Springer, 2008.

[5] D. Gabor, "Theory of communication," Journal of the Institution of Electrical Engineers, vol. 93, no. 26, pp. 429-441, 1946.

[6] K. Gurney, An Introduction to Neural Networks, UCL Press, 1997.

[7] I. Guyon, S. Gunn, M. Nikravesh, and L. A. Zadeh, Feature Extraction: Foundations and Applications, Springer, 2006.

[8] R. M. Haralick, K. Shanmugam, and I. h. Dinstein, "Textural Features for Image Classification," Proceedings of IEEE Transactions on Systems, Man, and Cybernetics, pp. 610-621, 1973.

[9] R. Harini and N. Sheela, "Feature extraction and classification of retinal images for automated detection of diabetic retinopathy," Proceedings of International Conference on Cognitive Computing and Information Processing, pp. 1-4, 2017.

[10] M. F. Hashim and S. Z. M. Hashim, "Diabetic retinopathy lesion detection using region-based approach," Proceedings of 8th Malaysian Software Engineering Conference (MySEC), pp. 306-310, 2014.

[11] S. Irshad, M. Salman, M. U. Akram, and U. Yasin, "Automated detection of cotton wool spots for the diagnosis of hypertensive retinopathy," Proceedings of Cairo International Biomedical Engineering Conference (CIBEC), pp. 121-124, 2014.

[12] A. S. Jadhav and P. B. Patil, "Detection of exudates for diabetic retinopathy using wavelet transform," Proceedings of International Conference on Power, Control, Signals and Instrumentation Engineering, pp. 568-571, 2017.

[13] T. Kauppi, V. Kalesnykiene, J.-K. Kamarainen, L. Lensu, I. Sorri, A. Raninen, et al., "DIARETDB1 diabetic retinopathy database and evaluation protocol," Proceedings of Medical Image Understanding and Analysis, Aberystwyth, Wales, 2007.

[14] N. Kaur, J. Kaur, M. Acharyya, N. Kapoor, S. Chatterjee, and S. Gupta, "A supervised approach for automated detection of hemorrhages in retinal fundus images," Proceedings of International Conference on Wireless Networks and Embedded Systems, pp. 1-5, 2016.

[15] K. Klubsuwan, W. Koodtalang, and S. Mungsing, "Traffic violation detection using multiple trajectories evaluation of vehicles," Proceedings of International Conference on Intelligent Systems, Modelling and Simulation, pp. 220-224, 2013.

[16] J. S. Lim and S. Jae, Two-Dimensional Signal and Image Processing, Prentice Hall PTR, pp. 469-476, 1990.

[17] S. Melmed, K. S. Polonsky, P. R. Larsen, and H. M. Kronenberg, William's Textbook of Endocrinology (13th edition), Elsevier, 2016.

[18] R. Mente, B. V. Dhandra, and G. Mukarambi, "Image recognition using shape descriptor: Eccentricity and color," IBMRD's Journal of Management and Research, vol. 3, no. 1, pp. 210-216, 2014.

[19] C. Özkan and F. S. Erbek, "The comparison of activation functions for multispectral Landsat TM image classification," Photogrammetric Engineering and Remote Sensing, vol. 69, pp. 1225-1234, 2003.

[20] E. Reinhard, M. Adhikhmin, B. Gooch, and P. Shirley, "Color transfer between images," IEEE Computer Graphics and Applications, vol. 21, no. 5, pp. 34-41, 2001.

[21] E. Reinhard, A. O. Akyuz, M. Colbert, and C. E. Hughes, "Real-time color blending of rendered and captured video," Proceedings of Interservice/Industry Training, Simulation, and Education Conference, 2004.

[22] E. Reinhard and T. Pouli, "Colour spaces for colour transfer," Proceedings of Computational Color Imaging, Heidelberg, Berlin, pp. 1-15, 2011.

[23] D. L. Ruderman, T. W. Cronin, and C.-C. Chiao, "Statistics of cone responses to natural images: Implications for visual coding," The Bell System Technical Journal, vol. 15, pp. 2036–2045, 1998.

[24] C. E. Shannon and W. Weaver, The Mathematical Theory of Communication, University of Illinois Press, pp. 379–423, 623–656, 1949.

[25] M. Sonka, V. Hlavac, and R. Boyle, Image Processing, Analysis, and Machine Vision (4th Edition), Cengage, 2015.

[26] S. Sreng, N. Maneerat, D. Isarakorn, R. Varakulsiripunth, B. Pasaya, J. Takada, et al., "Feature extraction from retinal fundus image for early detection of diabetic retinopathy," Proceedings of IEEE Region 10 Humanitarian Technology Conference, pp. 63 - 66, 2013.

Appendices

Appendix A: Retinal fundus image dataset

The examples of raw, optic disc removal, and pixel-based result on binary image, and ground truth are shown in Table A.1. This gives an overlook on process for all images which are used in this study.

Table A.1 Process on cotton wool images and ground truth

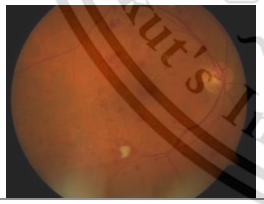
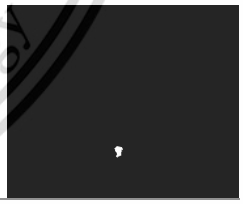
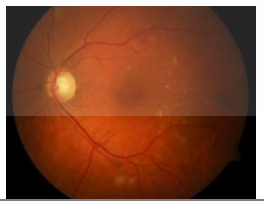
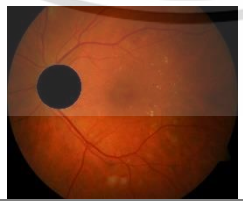
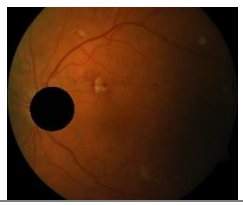
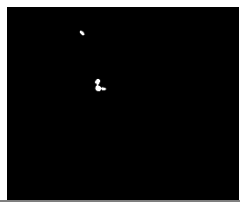
No.	Raw image	OD removal	Binary result	Ground truth
1				
2				
3				
4				

Table A.1 Process on cotton wool images and ground truth (continue).

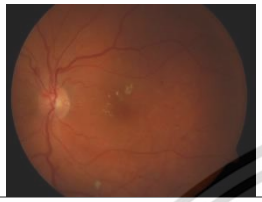
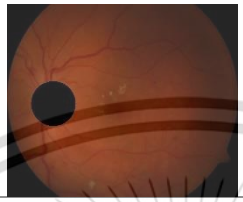
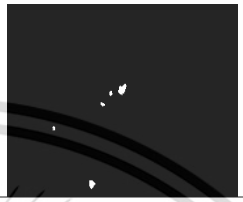
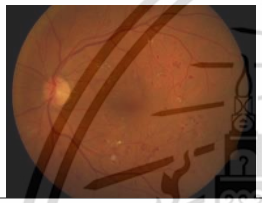
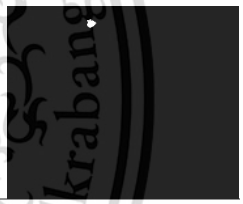
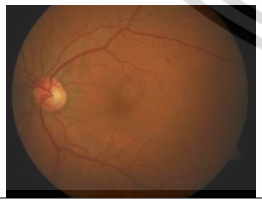
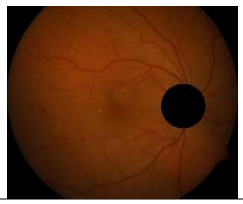
No.	Raw image	OD removal	Binary result	Ground truth
5				
6				
7				
8				
9				
10				
11				

Table A.1 Process on cotton wool images and ground truth (continue).

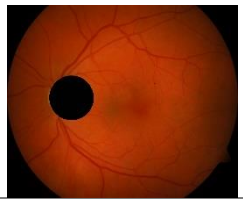
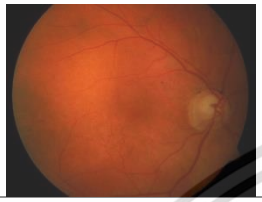
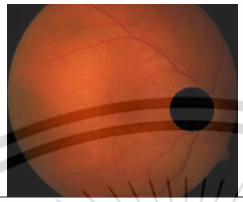
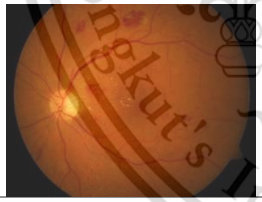
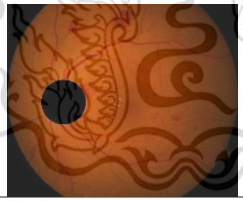
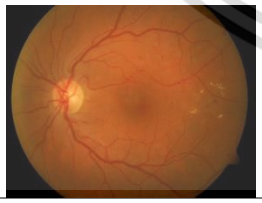
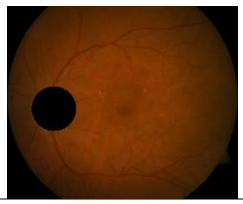
No.	Raw image	OD removal	Binary result	Ground truth
12				
13				
14				
15				
16				
17				
18				

Table A.1 Process on cotton wool images and ground truth (continue).

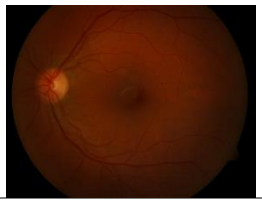
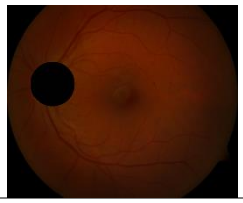
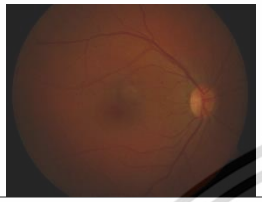
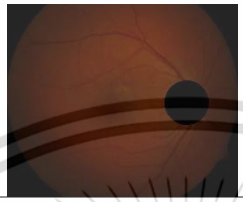
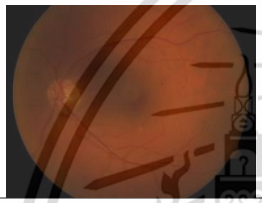
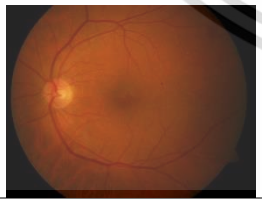
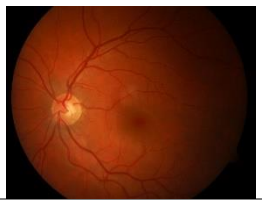
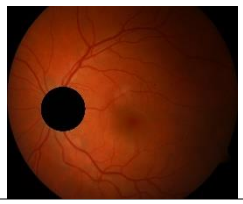
No.	Raw image	OD removal	Binary result	Ground truth
19				
20				
21				
22				
23				
24				
25				

Table A.1 Process on cotton wool images and ground truth (continue).

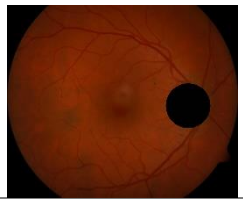
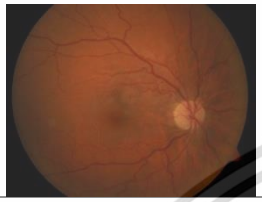
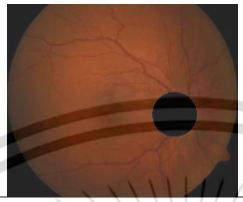
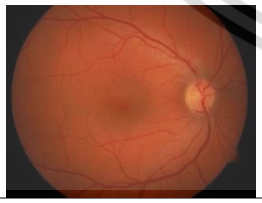
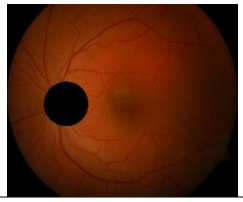
No.	Raw image	OD removal	Binary result	Ground truth
26				
27				
28				
29				
30				
31				
32				

Table A.1 Process on cotton wool images and ground truth (continue).

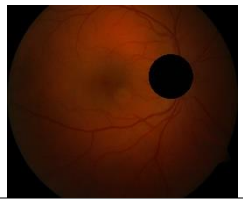
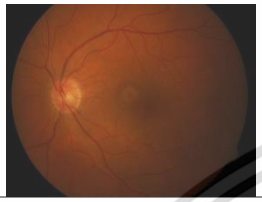
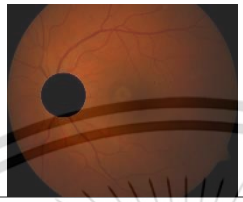
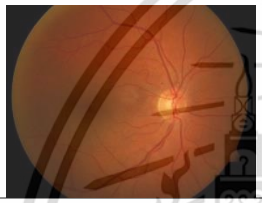
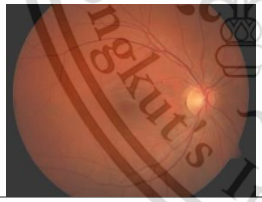
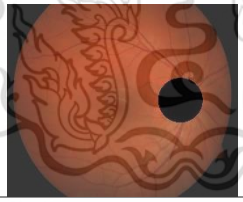
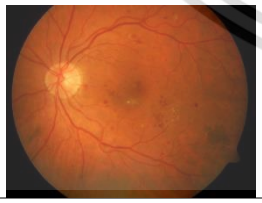
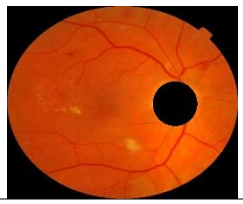
No.	Raw image	OD removal	Binary result	Ground truth
33				
34				
35				
36				
37				
38				
39				

Table A.1 Process on cotton wool images and ground truth (continue).

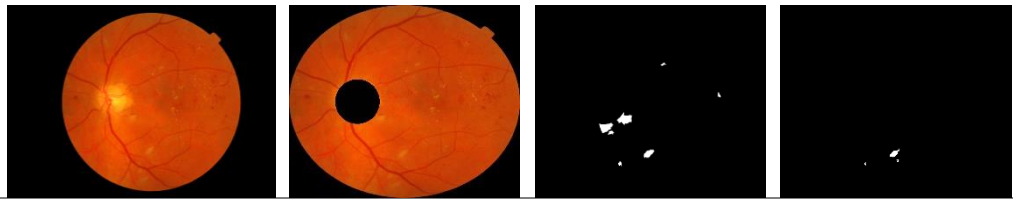
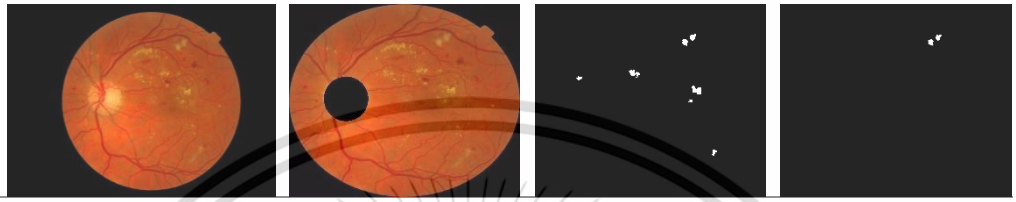
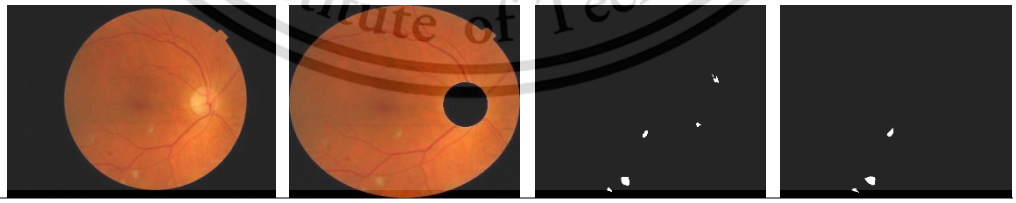
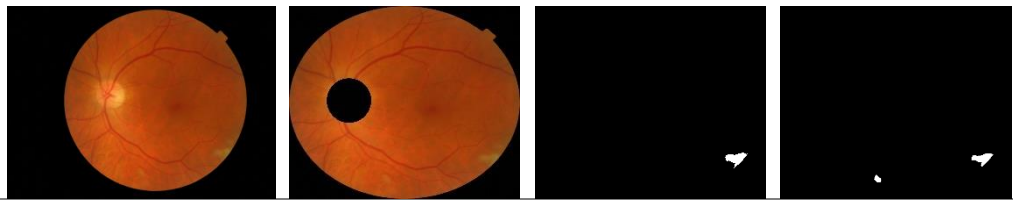
No.	Raw image	OD removal	Binary result	Ground truth
40				
41				
42				
43				
44				
45				
46				

Table A.1 Process on cotton wool images and ground truth (continue).

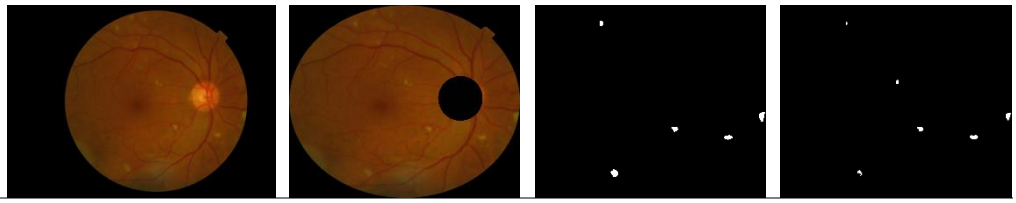
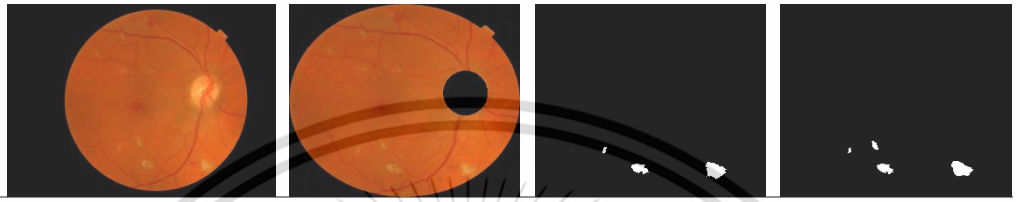
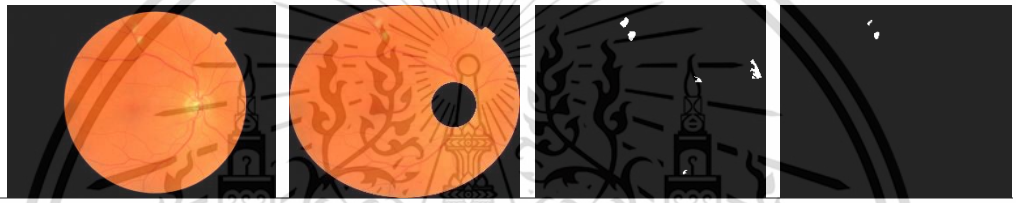
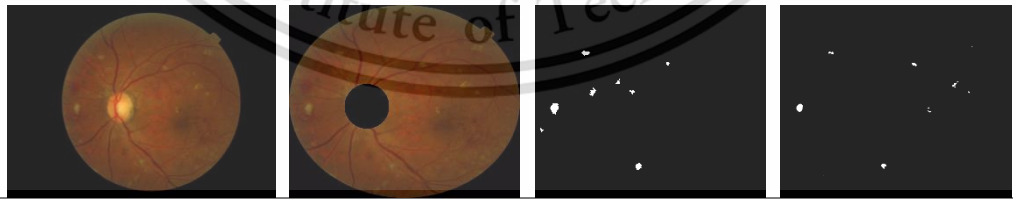
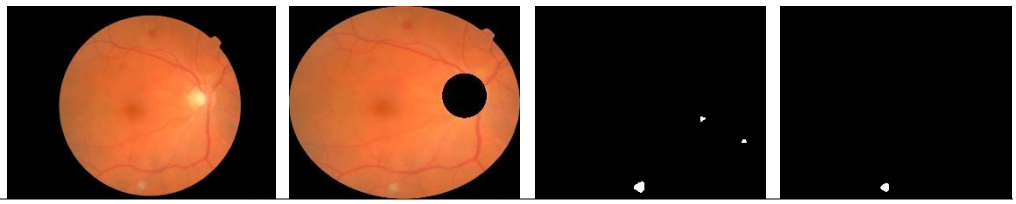
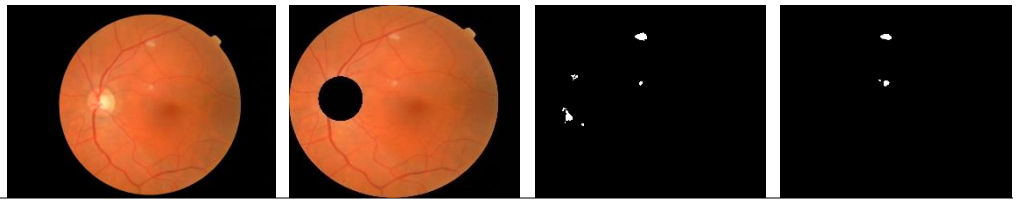
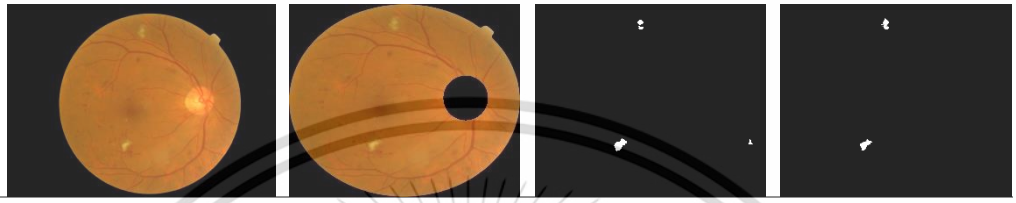
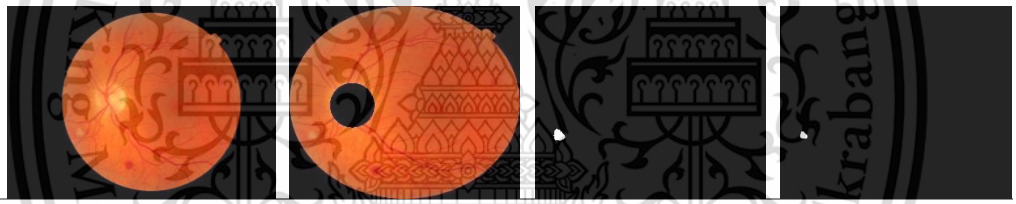
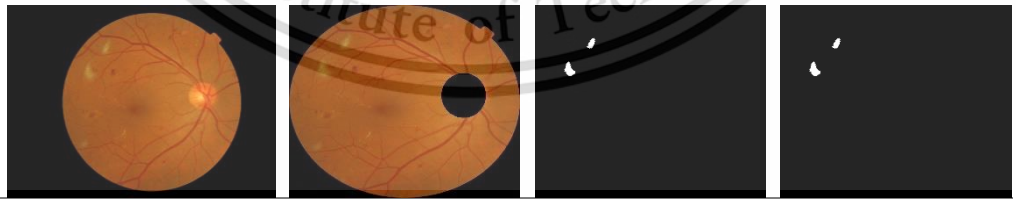
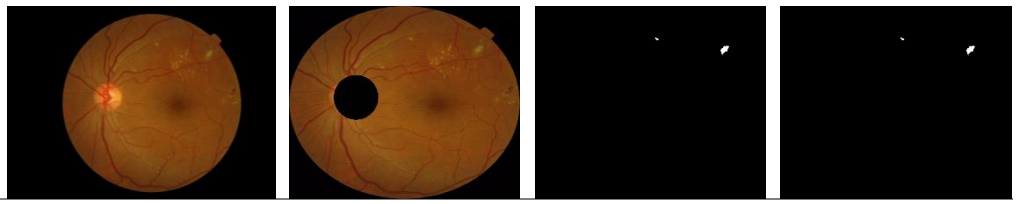
No.	Raw image	OD removal	Binary result	Ground truth
47				
48				
49				
50				
51				
52				
53				

Table A.1 Process on cotton wool images and ground truth (continue).

No.	Raw image	OD removal	Binary result	Ground truth
54				
55				
56				
57				
58				
59				
60				

Appendix B: Publication

Systems, Man, & Cybernetics Society

IEEE

IEEE Catalog Number: CFP1761U-USB
ISBN: 978-1-5386-0468-7

2017 IEEE 10th International Workshop on Computational Intelligence and Applications (IWCIA) Proceedings

IWCIA 2017

November 11-12, 2017

Hiroshima, JAPAN

2017 IEEE 10th International Workshop on Computational Intelligence and Applications
November 11-12, 2017, Hiroshima, Japan

Detection of Cotton Wool for Diabetic Retinopathy Analysis using Neural Network

Toan Bui
International College
King Mongkuts Institute of Technology Ladkrabang
Chalongkrung Rd., Ladkrabang, Bangkok, 10520 Thailand
toanbh.hust@gmail.com

Noppadol Maneerat
Faculty of Engineering
King Mongkuts Institute of Technology Ladkrabang
Chalongkrung Rd., Ladkrabang, Bangkok, 10520 Thailand
kmnoppad@yahoo.com

Ukrit Watchareeruetai
International College
King Mongkuts Institute of Technology Ladkrabang
Chalongkrung Rd., Ladkrabang, Bangkok, 10520 Thailand
ukrit.wa@kmitl.ac.th

***Abstract*—This paper presents an automatic segmentation method used to detect cotton wool spots in the retinal images for diabetic retinopathy disease. An early detection of cotton wool is important to prevent the dangerous damage which may cause blindness and vision loss. A preprocessing is applied to enhance image quality followed by optic disc removal. A feature extraction method is used to take useful elements from the image for increasing accuracy in classification step. A neural network model is employed for learning task and tested by k-fold cross validation. Our approach is evaluated by ground truth on DIARETDB1 public data. The result shows that cotton wool can be segmented by this method with 85.9% in sensitivity, 84.4% in specificity, and 85.54% in accuracy.**

***Index Terms*—Image processing, neural network, medical image processing, diabetic retinopathy, classification.**

I. Introduction

Diabetic retinopathy (DR) is known as one of the common diseases caused by high blood sugar. It may create some dangerous side-effects such as blindness and vision loss unless having an early detection for good treatment. The very first sign effect on eyes is known as exudates, cotton wool, hemorrhage, and micro aneurysms. Most of them can be easily detected by specialists. However, the number of specialists is very less (around one thousand) compared with patients raising year by year (around six million patients) [1]. Therefore, an automatic detection method should be useful for helping in early detection and timely medication. Cotton wools (as shown in Fig. 1) are nerve fiber layer infarcts or pre-capillary arterial occlusion. In other words, they are an ischemic event of a very small amount of tissue. Cotton wool is common in diabetic retinopathy and hypertension. It has a similar feature with exudates and optic disc in color but different in size and morphology.

Cotton wool can be detected using image processing and computer vision. In addition, machine learning is one of useful methods for classification task. This issue has inspired a lot of analysts in biomedical and ophthalmology. However, researches based on cotton wool are limited in number and accuracy as well. Most of them used private databases that cause a lot of difficulties for later analysis and comparison [3]. Some research analysts use global data but they can reach limited performance 74.53% accuracy [4]. Some of them do not make clear efficient accuracy between exudates and cotton wool, which should be separately detected for different treatment. Therefore, they will be considered as a different investigation and incomparable result [3].

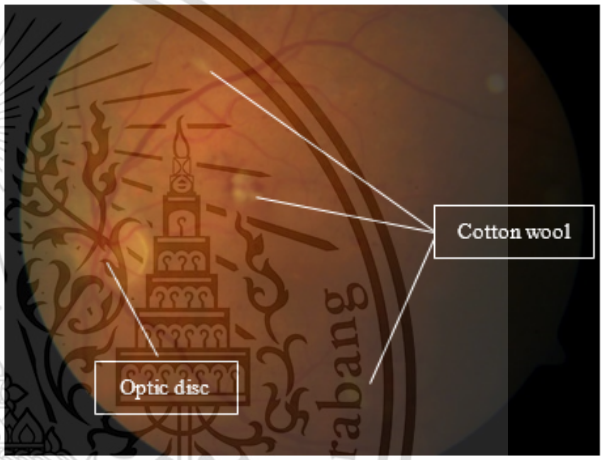


Fig. 1: Example of cotton wool spot in a retina image [2]

In this article, a method is used to detect cotton wool only and consider exudates and optic disc as noise element in analysis process using the neural network with feature extraction. The validation of the proposed method is examined on DIARETDB1 [2], a public database for scientific research purpose. The database consists of 89 images, along with ground truth made independently by four medical experts, using a software tool provided for the image. The previous research based on similar data should be considered for comparison [4].

The remainder of this article is organized as follows: Section

978-1-5386-0469-4/17/$31.00

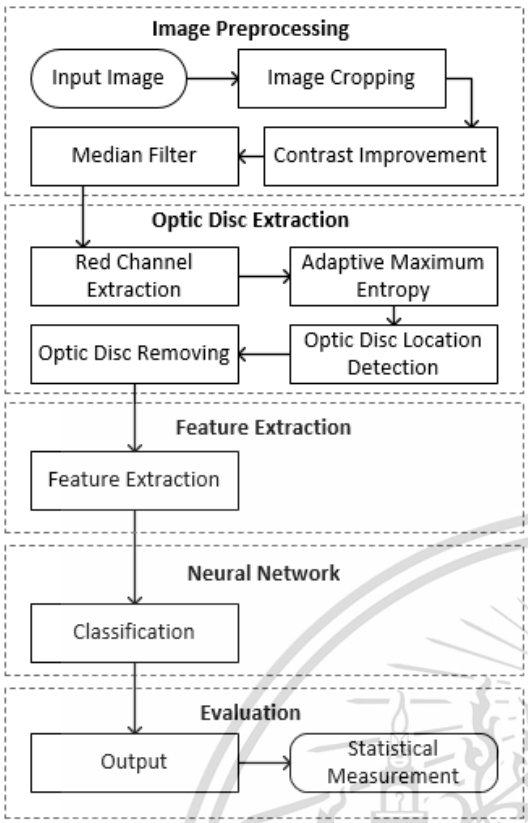


Fig. 2: Flow chart of the proposed method

II describes the proposed method for enhancing input image , feature extraction and classification step. Section III is the experimental setup, performance result and comparison. Finally, the last section concludes the whole paper and gives some idea about future work.

II. Proposed method

The overview of the proposed method is shown in Fig. 2. The flowchart describes the whole process for input image data and how it was evaluated by pixel samples for getting final performance. Each process is explained in the following section.

A. Preprocessing

Firstly, an input RGB image (Fig. 3a) is analyzed to detect the eye location in the image. A cropped image is extracted for reducing useless instances in next analysis. Secondly, the image is divided into the red, green and blue planes. A median filter is applied in each plane a for noise removal [5] that is salt or pepper appear in the image during collecting data step and it is calculated as:

$$\hat{f}(x,y) = \underset{(s,t)\in S_{xy}}{\text{median}}\{g(s,t)\} \tag{1}$$

where $F(x,y)$ is the response of the median filter at a given coordinate, S_{xy} is a $m \times n$ sub-image of input image g, x and y are coordinate in the input image. The result is done and showed in Fig. 3b.

B. Optic Disc Extraction

The enhanced image obtained from the previous step is used for detecting optics disc. Because the optic disc is considered as noise in detecting cotton wool, an early optic disc removing step can improve the final result. In this research, the image is converted into grayscale image and applied maximum entropy thresholding method [6], [7], [8] to extract optic disc. The equation is explained as follow:

$$H_o(t) = -\sum_{i=0}^{t-1}(\frac{P_i}{P_o}\log_2\frac{P_i}{P_o}) \tag{2}$$

$$H_b(t) = -\sum_{i=0}^{t-1}(\frac{P_i}{P_b}\log_2\frac{P_i}{P_b}) \tag{3}$$

$$H_t = H_o(t) + H_b(t) \tag{4}$$

where i is the value in the gray image, o and b represent values of object and background, respectively. P_i is the probability of the pixels being i. P_o and P_b represent the probabilities of object and background, respectively. $H_o(t)$ and $H_b(t)$ are the probabilities distribution for object and background, respectively. $H(t)$ is the entropy of grayscale image implied by threshold t.

With $H(t)$ is the maximum value, the optimal threshold value can be selected by maximizing the entropy of $H(t)$ which can be defined as:

$$t = ArgMax(H(t)) \tag{5}$$

where t is in a range between 0 and 255.

After the location of cotton wool is detected, a MATLAB function is used for detecting centroid of optic disc region. Based on the centroid, a black circle mark is drawn on optic disc (Fig. 3c).

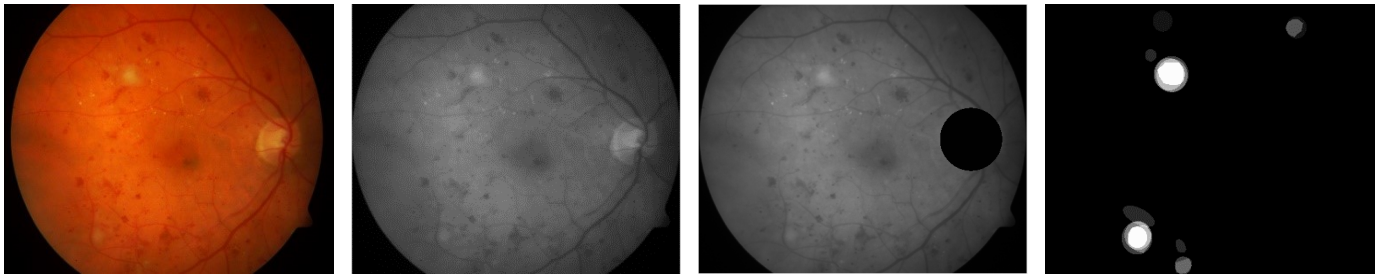
C. Feature Extraction

Because the regions of cotton wool are quite small, compared to the backgrounds, it makes our data unbalanced and causes the difficulties in the learning process. We use an approach that each image can be divided into pixel and each pixel of cotton wool and non-cotton wool is a learning instance. In that process, the number of non-cotton wool samples is reduced comparing with cotton wool samples for making data balancing approximately 60% for non-cotton wool and 40% for cotton wool.

Before putting the samples and their ground truth (Fig. 3d) into the neural network in the learning process, each sample will be analyzed to extract useful features as follows:

1) Red plane: a component plane color space represents the red element in the RGB image.

2) Green plane: a component plane color space represents the green element in the RGB image.



(a) Original image [2] **(b)** Preprocessing image **(c)** Optic disc extracted **(d)** Ground truth image [2]

Fig. 3: Images in processing and their ground truth

3) Intensity plane: a 2-D plane is used in range from 0 to 255 to represent pixel intensity as follows:

$$I = 0.2989 \times R + 0.5870 \times G + 0.1140 \times B \quad (6)$$

where R is the pixel value in the red plane, G is the pixel value in the green plane and B is the pixel value in the blue plane.

4) Local mean: each pixel in the intensity plane is assigned to produce mean value using its 8 neighbor pixels in a window size 3×3.

$$\mu = \frac{1}{N}\sum_{i=1}^{N}(x_i) \quad (7)$$

where x_i is the value of each pixel in window size 3×3 and N is the number of pixels (in this case, it is 9).

5) Local variance: after mean calculation, each pixel is considered again in window size 3×3 using its 8 neighbor values to produce the variance.

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2 \quad (8)$$

where x_i is the value of each pixel in window size 3×3, μ is the mean of considered pixel and N is the number of pixel.

6) Local brightness: the value represents visual perception in which is a source of light or light reflection source. It can be calculated by RGB color space as follows:

$$Y = 0.2126 \times R + 0.7152 \times G + 0.0722 \times B \quad (9)$$

7) Saturation plane: a component plane in HSI color space is transformed from RGB color space as follows:

$$S = 1 - \frac{3}{R+G+B}[min(R,G,B)] \quad (10)$$

8) Local range: in intensity plane: each pixel is considered with its 8 neighbor pixels in the window size 3×3. The pixel value is replaced by the range between the biggest value and smallest value. In window size I, the value response as:

$$L = Max(x_i) - Min(x_i), x_i \in I \quad (11)$$

9) Local standard deviation: each pixel in intensity plane is used to produce standard deviation value by 8 neighbor values in the window size 3×3.

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2} \quad (12)$$

where x_i is the value of each pixel in window size 3×3, μ is the mean of considered pixel and N is the number of pixels.

10) Local entropy: each pixel in the intensity plane is used to produce its entropy as shown in Eqs. 2, 3 using a window size 9×9.

D. Neural Network

A neural network model is built for classification step. The model consists of a hidden layer inside to produce the best result. The hidden layer is tested from 10, 15, 20, ..., 40 nodes to get the best result. In each node, a transfer function is evaluated to use for learning each sample.

In the task of classification, the output values are typically discrete, most of the time they are binary. There are two strong functions recommended: log-sigmoid function (Fermi function) and tan-sigmoid function (Hyperbolic tangent sigmoid function). Fermi function generates an output value in the range of 0 and 1. Whereas, the tan-sigmoid function generates values in the range from -1 to 1. The essential difference between range value makes the log-sigmoid function more difficult than tan-sigmoid to learning something far from threshold value where its results are close to 0 [9]. Therefore, the tan-sigmoid function is the better choice for our model.

$$F(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (13)$$

where x is the sum of whole features for each input sample which is analyzed by the previous step.

III. Experimental Results

For the evaluation, previous work focused on a number of the cotton wool blob in each image or in total. However, that approach makes some uncountable error in the border of each blob. In this study, we use an evaluation method based on a number of pixel samples. The method can detect cotton wool spot not only about the number of blobs but also calculate the percentage of the cotton wool blobs is correctly detected.

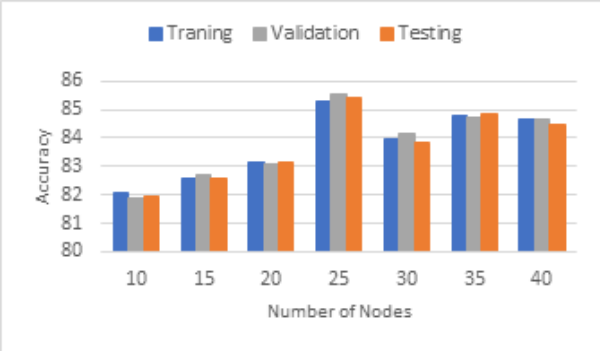


Fig. 4: The accuracy of the proposed methods with different number of hidden nodes

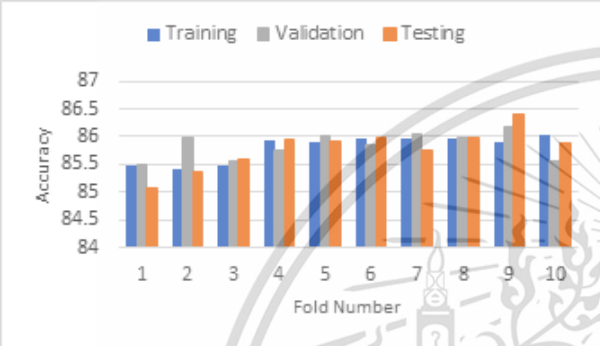


Fig. 5: The accuracy of the proposed method (10-fold cross-validation)

A process for testing the best result based on number of nodes is done and shown in Fig. 4. The accuracy is rising while the number of nodes was being increased (minimum around at 82% in case of 10 nodes) but it stops increasing to get the maximum accuracy value at 25 nodes (with 85.54%) and get decreased after that. So the best result obtained can be produced using 25 nodes in the layer.

$$Sensitivity = \frac{TP}{TP + FN} \tag{14}$$

$$Specificity = \frac{TN}{TN + FP} \tag{15}$$

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{16}$$

The proposed method is applied to DIARETDB1, a popular public database for any research based on eye diseases. 14 cotton wool images with resolution 1500×1152 pixels in 24-bit depth PNG format is used for extracting 248,772 samples. They consist of 138,376 non-cotton wool samples and 110,396 cotton wool samples. The experiment is done using k-fold cross validation with k equal 10. k-fold cross validation is used to ensure that the proposed method is not overfitting for this data. We divide our data into the neural network by 8 folds for training, 1 fold for validation, 1 fold for testing. The result demonstrates that our model is not only overfitting for this data only but also good in general since the result is changing time by time but not a large difference (Fig. 5).

TABLE I: Performance of classification (%)

Method	Sensitivity	Specificity	Accuracy
[4]	71.78	76.49	74.53
Proposed	85.90	84.40	85.54

For result comparison, we compute three common evaluation methods known as sensitivity (recall, true-positive-rate or TPR), specificity (true-negative-rate or TNR), and accuracy (ACC) using elements as true positive (TP), true negative (TN), false positive (FP) and false negative (FN) in Eqs. 14, 15, 16. According to the result in Table. I, our result gets 85.90%, 84.40%, and 85.54% for sensitivity, specificity, and accuracy, respectively. The proposed method is obviously better than previous work based on the similar database.

IV. Conclusion

Obviously, the proposed method has been successful in cotton wool detection using a fully automatic method. Many features are already used in extraction step and contributed an important role. In this case, they are color features as RGB space, intensity and the texture features as contrast, range, mean, and entropy. In order to reach more accuracy in future work, we can consider between extracting for more complex features and using more different machine learning algorithm as well.

Acknowledgment

This research has been financially supported by AUN/SEED-Net, JICA.

References

[1] T. Chetthakul, C. Deerochanawong, S. Suwanwalaikorn, N. Kosachunhanun, C. Ngarmukos, P. Rawdaree, et al., *Thailand diabetes registry project: Prevalence of diabetic retinopathy and associated factor in type 2 diabetes mellitus,* Journal of the Medical Association of Thailand, vol.89, pp.S27-36, 2006.

[2] T. Kauppi, V. Kalesnykiene, J.K. Kamarainen, L. Lensu, I. Sorri, A. Raninen, et al., *DIARETDB1 diabetic retinopathy database and evaluation protocol, Proceedings of 11th Conference on Medical Image Understanding and Analysis,* 2007.

[3] S. Irshad, M.U. Akram, M. Salman, and U. Yasin, *Automated detection of cotton wool spots for the diagnosis of hypertensive retinopathy,* Proceedings of 7th Cairo International Biomedical Engineering Conference, pp.121-124, 2014.

[4] M.F. Hashism and S.Z.M. Hashim, *Diabetic retinopathy lesion detection region-based approach,* Proceedings of 2014 8th Malaysian Software Engineering Conference (MySEC), pp.306-310, 2014.

[5] J.S. Lim, *Two-Dimensional Signal and Image Processing* Prentice Hall, pp. 469-476, 1990.

[6] M. Azarbad, A. Ebrahimzade, and V. Izadian, *Segmentation of infrared images and objectives detection using maximum entropy method based on the bee algorithm,* International Journal of Computer Information Systems and Industrial Management Applications (IJCISIM), vol.3, pp.026-033, 2011.

[7] C.E. Shannon and W. Weaver, *The Mathematical Theory of Communication,* University of Illinois Press, 1971.

[8] S. Sreng, N. Maneerat, D. Isarakorn, R. Varakulsiripunth, B. Pasaya, J. Takada, et al., *Feature Extraction from Retinal Fundus Image for Early Detection of Diabetic Retinopathy* Proceedings of IEEE Region 10 Humanitarian Technology Conference, 2013.

[9] D. Kriesel, *A Brief Introduction to Neural Network* 2007.

Author biography

Author: Mr. Bui Huy Toan

Degree: Master of Engineering

Date of Graduation: July 31st, 2018

Date of Birth: June 25th, 1993

Place of Birth: Phutho, Vietnam

Undergraduate and Graduate Education:

Master of Engineering in Computing in Engineering Systems,

King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand 10520, 2018.

The Degree of Bachelor in Information Technology,

Hanoi University of Science and Technology, Hanoi, Vietnam, 2016.

Major: Computing in Engineering Systems.

Presentations and Publication:

- Paper's name: "**Detection of Cotton Wool for Diabetic Retinopathy Analysis using Neural Network**", 11th – 12th, November 2017, IWCIA2017, Hiroshima, Japan.