A HIGHLY ACCURATE AND ROBUST

RETINAL VESSEL SEGMENTATION ALGORITHM

A Thesis Submitted to the College of

Graduate Studies and Research

in Partial Fulfillment of the Requirements

for the Degree of Master of Science

in the Department of Electrical and Computer Engineering

University of Saskatchewan

Saskatoon

By

SEN AN

PERMISSION TO USE

In presenting this thesis in partial fulfilment of the requirements for a postgraduate degree from the University of Saskatchewan, I agree that the libraries of this university may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which this thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Head of the Department of Electrical and Computer Engineering

Department of Electrical and Computer Engineering, University of Saskatchewan

57 Campus Drive, Saskatoon, Saskatchewan, S7N 5A9

ABSTRACT

Diabetes is one of the most common diseases that affects human beings' health. Its complications can cause death, and it sometimes causes vision loss. Diabetic retinopathy (DR) is the complication of diabetes that leads to vision loss. Many people who lack treatment for diabetes experience DR, especially in remote areas. Researchers are looking for ways to detect DR at an early stage with lower costs to decrease the occurrence of vision loss. Segmentation is a method that could help in DR detection. Segmentation is a method used to convert a continuous color or grayscale image into a binary ground truth image. Continuous images have low contrast for vessels, but the binary ground truth would provide higher contrast for observing vessel structure. In addition, patients need periodic examination of the retinal blood vessels. Binary ground truth images are easier to store and compare every three or six months. Vessel segmentation relies on specific algorithms. Two metrics are used here for comparison; one is vessel tracking ability, and the other is robustness. These two metrics are essentially accuracy, but the training sources are different. With respect to vessel tracking ability, the training and testing parts of the algorithm are drawn from the same database without any interference from other databases. To assess robustness, training and testing are done in different databases to test whether the algorithm can be extended to other databases. Similarly, single-database testing is used here to reflect vessel tracking ability, and cross-database testing is chosen to reflect robustness. Before 2012, papers on this topic focused only on vessel tracking ability, whose training and testing are based in the same database. After 2012, Fraz et al. proposed cross-database testing to measure the robustness performance. In a clinic, robustness is as crucial as vessel tracking ability. In this thesis, robustness is the main target, along with lower degradation for vessel tracking ability. This thesis achieved 0.5% higher cross-database testing accuracy than Fraz's work. The performance improvement is mainly

from the optimized classifier input. The binary classifier input is processed by pre-processing and Mathematical Morphology (MM) components, so the data distribution is better for the following classification.

ACKNOWLEDGMENTS

CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

The retina is an important tissue for human beings' vision because of its light sensitivity. It plays an important role in vision formation. Light passes through the cornea and lens and finally reaches the retina [1]. The retina is light-sensitive tissue that helps to form vision signals and passes them to the brain.

If this light-sensitive tissue is blocked in the light route, part of the vision will be missing. This blocking could occur in the middle or on the periphery, and the type of blocking depends on the disorder. Central vision loss is caused by age-related macular degeneration, while peripheral vision loss is caused by glaucoma [2]. Though age-related macular degeneration and glaucoma are not the only reasons for vision loss, they are quite common.

The retina is a reflector for other potential disorders, not just for vision disorder. The retina is associated not only with eye diseases but also with diabetes. Diabetic retinopathy is a general cause of vision loss. It is a complication of diabetes, and it happens with newly grown capillary vessels [3]. The detailed pathology and symptoms will be introduced later.

For elder people, regular checking should be done to look for potential diseases such as macular degeneration, glaucoma, or diabetes. For retina-related diseases, periodic examination would prevent the occurrence of vision loss in an early stage. Physicians could prescribe medication while the disease is still curable.

Eye examinations are usually non-invasive and inexpensive compared to blood examinations. Because of the connection between the retina and other diseases, people could even be alerted to potential eye diseases or diabetes when they get their regular eye examinations.

Current eye examination devices are confined to a place, an environment, and operators. These devices are usually large and unmovable, and they rely on the operators' clinical experience. Because of these restrictions and lack of devices and doctors, people in remote areas cannot receive services. In India, for example, a large number of people lose their vision because of diabetes and other eye diseases. It is necessary to develop easily-carried automatic diagnostic devices. An algorithm for disease analysis is needed in the market. If regular retinal examinations could be provided inexpensively, people could deal with diabetes in an early stage, and the number of people experiencing vision loss would decrease drastically.

In the future, examinations will not have strict place or environment requirements. For example, people could be given a normal retinal examination when get new glasses. Moreover, every village has its own portable retinal tester for their villagers. These are beneficial to decrease the occurrence of vision loss, and portable eye examination devices are extremely helpful for remote poor areas [4].

1.1 RETINA-RELATED DISORDERS

Because retina is a core tissue in vision formation, the damage to the retina can directly affect peoples' vision. With affected vision, lots of people may encounter various difficulties in their daily activities [2].

Diabetic retinopathy (DR) and macular degeneration are two common retinal disorders that can cause vision loss directly, and these two retinal disorders can be detected using image processing techniques. Chapter 1.1.1 and Chapter 1.1.2 introduce these two disease generally, and Chapter 1.2 introduces the relationship between these two disorders and this thesis.

1.1.1 DIABETIC RETINOPATHY

Diabetes is a leading cause of health issues worldwide. It can cause death not only because of itself but also mostly because of its complications, especially for elder people. Diabetes gradually changes the blood sugar. In some cases, diabetes also causes vision loss for patients. People usually become aware of diabetes at a late stage because it is hard to discover based only on symptoms. Usually people will be subjected to a blood examination after their health issues become serious. Diabetes can cause vision loss. Because diabetes develops with changes in blood sugar, block capillary vessels more easily become clocked. When a block occurs, body must grow new nourishing vessels to support downstream tissues [3]. The retina fundus has rich vessel tissues, so it can be observed easily whether vessels have grown. First, micro aneurysms appear with the change in blood gradients. In a captured retinal image, some dark spots or dark areas will appear. This is the mild non-proliferative retinopathy stage. After micro aneurysms increase, the body grows more capillary vessels to nourish the downstream tissues. This is the moderate non-proliferative retinopathy stage. These micro aneurysms might swell and leak fluid into the retinal fundus, and this leaking fluid could condense and block the light-sensitive retina. This is a cause of vision loss. When the situation worsens, the proliferative retinopathy stage is too serious to recover from. According to previous research, retinal analysis is quite helpful for the early detection of diabetic retinopathy [5, 6]. People easily ignore blood sugar increasing. A retinal examination can help to alert people of the need for an accurate blood examination before it is too late. Especially in remote and poor areas, people do not seek health examinations until their issues become serious. If the retinal examination can be implemented through a portable, easily-operated machine, people would have a chance to go to a local clinic for a short retinal examination and an estimate of their health status.

Retinal image analysis is used to analyze and estimate potential patient problems. Before the actual analysis step, segmentation is an important pre-step for retinal analysis. The original retinal fundus images captured by devices are RGB (Red, Green, Blue) color images. The contrast for vessels, especially for capillary vessels, is weak. The method of analysis might ignore some of these thin capillary vessels, decreasing the detection accuracy. Segmentation is a step to convert three-channel RGB images into binary ground truth images. Binary images are discrete images, and it is more efficient for width and curvature measurement. In addition, binary images are easier to store in patients' medical records.

1.1.2 MACULAR DEGENERATION

The macula is the central part of the retina; it plays an extremely important role in vision formation, and its color is yellow. In addition, the macula is the most sensitive tissue in the retina. The macula provides central and sharp vision, which is the field straight ahead of the human eye. This field affects daily life the most, such as reading, writing, and driving. If macular disorders occur, the central vision will be blocked, blurred, or distorted. People might have trouble in their daily activities.

Macular degeneration in different stages is categorized by the size of drusen, which are yellow deposits under the retina. Drusen are formed by aging as a result of some undisposed waste. Ophthalmologists look for drusen when they are trying to assess for the occurrence of age-related macular degeneration (AMD).

A retinal examination can alert a doctor to macular degeneration, especially among elder people. This examination is helpful for the early detection of age-related macular degeneration to prevent the occurrence of vision loss. Moreover, a retinal examination is usually non-invasive, and people can be examined periodically.

Macular degeneration happens for various reasons; according to clinic surveys, an unhealthy lifestyle, such as smoking, and family history are the most common reasons for this retinal disorder. To prevent this disorder, maintaining a healthy lifestyle, a stable cholesterol level, and normal blood pressure are efficient ways.

1.2 MOTIVATION

Diabetes threatens people not only by itself, but it can also affect people's health mostly by its complications and these complications could cause death. Diabetic retinopathy is a complication of diabetes and lots of people have lost their vision because of it. Diabetic retinopathy is the most common diabetic eye disease and it is caused by changes in the blood vessels of the retina.

Diabetic retinopathy has different stages. At the earliest stage, microaneurysms occur. They are swelling areas in the retinal blood vessels. As the disease progresses, some blood vessels which nourish the retina are blocked. Later, many more blood vessels are blocked. The signals sent by the retina for nourishment trigger the growth of new blood vessels.

Vessel segmentation focuses on highlighting vessel structures and highlighting newly grown vessels. Segmentation is significantly important in retinal inspection, especially for diabetic retinopathy to prevent vision loss.

In eye diseases detection system, segmentation is a key component before other specific algorithms. Every disease detection needs its own algorithm. Segmentation is a common step in order to increase the detection accuracy. The general diagram is shown in Figure 1.1. The processing procedure starts after image capture, and ends in a displaying screen. The device uses segmentation algorithm and disease analyzing algorithms. Also, user interface is needed to give feedbacks to the user. Automatic algorithms need to be developed in order to broaden

the usage of this kind of device. This thesis proposes a high-accuracy and robust retinal vessel segmentation algorithm.
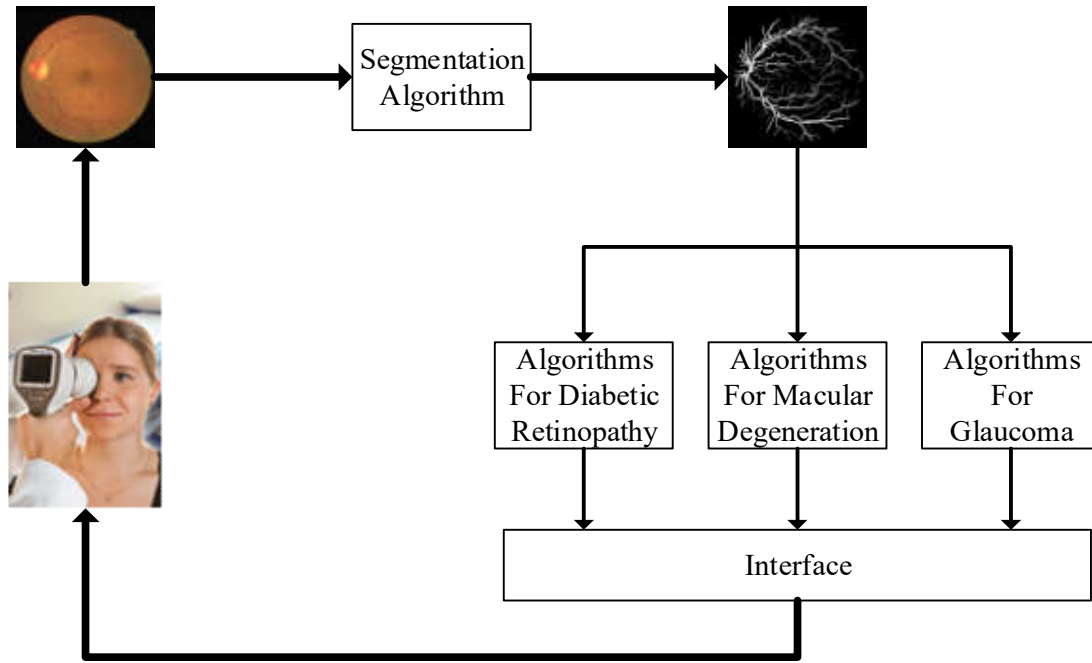
Figure 1.1: General block diagram of eye disease detection system

1.3 SEGMENTATION

In Figure 1.1, segmentation plays as a shared component for several disease-detecting algorithms. Because segmentation is to convert continuous images to ground truth images, this is helpful to improve analyzing performance for subsequent algorithms, such as algorithm for Diabetic Retinopathy (DR), algorithm for macular degeneration, and algorithm for glaucoma. Vessel segmentation gets rid of the distracting part in the original image.

1.3.1 THE IMPORTANCE OF SEGMENTATION

Segmentation is used to convert continuous color or grayscale images to discrete binary images. The segmentation component can be considered as a binary classifier, with one class for vessels and the other class for non-vessels.

Retinal images are usually captured by medical devices or obtained from public databases. These color or grayscale images usually have low contrast between vessels and the background, so retinal vessels are easily mixed into the background, especially thin vessels. It becomes harder for subsequent manual or automatic analysis to identify vessels. Even experienced ophthalmologists cannot distinguish vessels in low-contrast conditions. This might cause ignorance of newly grown vessels in diagnosis. In the worst case, early treatment might be delayed. Binary images are better in highlighting vessel structure, and binary images could decrease the occurrence of errors in diagnosis. Segmentation is a procedure used to highlight vessel structures and increase vessel contrast to enable subsequent algorithms to mark the vessels. It serves as a pre-processing step before other algorithms are employed, and it can increase the analysis accuracy of other algorithms. Low vessel contrast makes it more difficult to identify retinal disorders. The algorithms used target different retinal disorders. If vessels can be highlighted first, it is easier for algorithms to detect newly grown capillary vessels, abnormal vessel bifurcation, and even great vessel curvature. Newly grown vessel detection is much more meaningful in a clinic setting. It could reflect the presence of diabetes and the stage of Diabetic Retinopathy (DR). With the growth of capillary vessels, ophthalmologists could conclude that more regular vessels are blocked because of increased blood sugar.

In this thesis, the target is to develop a better segmentation algorithm. The two metrics used to measure segmentation performance are vessel tracking ability and robustness. To achieve better segmentation accuracy, a supervised classification method is used. To prevent

potential drawbacks of the supervised methods, pre-processing and mathematical morphology components are implemented to extend the usage of the algorithm and increase the robustness performance. The method is described in detail in the Chapter 3.

1.3.2 SEGMENTATION DATA SOURCE

In current research, two public databases are mainly used; one is DRIVE (Digital Retina Images for Vessel Extraction), which is provided by a Dutch institution (Image Science Institute), and the other is STARE (STructured Analysis of the REtina), which is run by the University of California. As shown in Table 1.1, other image sources are related to this topic. In this table, REVIEW and CHASE_DB1 are listed. REVIEW (REtinal Vessel Image set for Estimation of Widths) is provided by the University of Lincoln, and vessel widths are documented in this database. CHASE_DB1 (Child Heart And Health Study in England) is another database for retinal vessel segmentation, created by Kingston University London. Some of the properties of these databases are listed in Table 1.1. The most common ones for retinal vessel segmentation are DRIVE and STARE. These two databases were made available to the public in the early 2000s, and they are widely accepted by researchers worldwide. Nearly every retinal vessel segmentation algorithm has been tested on these two databases.

Image resolutions are shown in Table 1.1. The highest resolution appears in the REVIEW dataset. However, REVIEW categorizes its images into four types, HRIS (the High Resolution Image Set), VDIS (the Vascular Disease Image Set), CLRIS (the Central Light Reflex Image Set), and KPIS (the Kick Point Image Set). Every type has a different image resolution. Even though its average resolution is higher than those of the other three databases, the lack of consistency brought inconvenience in programming and comparison.

DRIVE, STARE, and CHASE_DB1 all contain manually labelled ground truth images. These manually labelled images can be used for supervised classification. These manually labelled images are the standards for segmentation. In another words, manually labelled images are 100% correctly classified images.

Table 1.1: Comparison of different databases

	Height (pixel)	Width (pixel)	Bit Width	Number of Retinal Images for Segmentation	Institution
DRIVE	584	565	8	20+20	Image Science Institute
STARE	605	700	8	20	University of California National Institutes of Health
REVIEW	2,438 (4) 1,024 (8) 1,440 (2) 119 (2)	3,584 (4) 1,360 (8) 2,160 (2) 288 (2)	8	4+8+2+2	University of Lincoln
CHASE_DB 1	960	999	8	28	Kingston University London

The DRIVE database has 40 images for segmentation, 20 for training and 20 for testing. STARE has 20 images in total, and CHASE_DB1 has 28 images in total. CHASE_DB1 was created in the 2010s, so it is not widely used in retinal vessel segmentation papers. In contrast, DRIVE and STARE have been in every paper on this topic.

1.4 THESIS OUTLINE

In this thesis, the segmentation algorithm is introduced and every single component is explained in detail. Chapter 1 gives brief introduction of the related research field. In Chapter 2, previous works have been discussed, including some of their drawbacks. Chapter 3 explains the proposed algorithm and its components in detail, while Chapter 4 demonstrates the algorithm performance. In Chapter 5, conclusion and future work will be given.

2 BACKGROUND

To perform retinal vessel segmentation, color fundus images are imported from the DRIVE and STARE databases. In the DRIVE database, the images are 584 pixels by 565 pixels size. In the STARE database, the images are 605 pixels by 700 pixels. Color images need to be converted into binary ground truth images. The inputs are continuous color images, and the outputs are binary ground truth images. The algorithm is basically a binary classifier.

The measurement of algorithm performance is the same as the standard to measure binary classifiers. The vessel tracking ability and robustness will be used to measure the performance.

2.1 SUPERVISED VS. UNSUPERVISED CLASSIFICATION

Classifiers are categorized as supervised classifiers or unsupervised classifiers. According to previous papers, supervised classifiers achieve higher accuracy than unsupervised classifiers in general. Supervised classification involves a training set to train the model, and the classification is based on this model. For example, in this thesis, the training component uses samples to obtain parameters for a Gaussian Mixture Model (GMM). After the parameters are obtained, the classification component can calculate the vessel class probability and non-vessel class probability. According to these two probabilities, an unknown sample can be classified into vessel class or non-vessel class. Therefore, our thesis uses a supervised method. Unsupervised methods do not have a training component, so the parameters are not obtained by a training component but instead result from an empirical calculation or prior knowledge.

Table 2.1: Classification methods (DRIVE database)

No	Type	Methods	Year	SN	SP	Acc	AUC
1.	Unsupervised Methods	Hoover [7]	2000	---	---	---	---
2.		Zana [8]	2001	0.6971	---	0.9377	0.8984
3.		Jiang [9]	2003	---	---	0.9212	0.9114
4.		Mendonca [10]	2006	0.7344	0.9764	0.9452	---
5.		Lam [11]	2008	---	---	---	---
6.		Al-Diri [12]	2009	0.7282	0.9551	---	---
7.		Lam [13]	2010	---	---	0.9472	0.9614
8.		Miri [14]	2011	0.7352	0.9795	0.9458	---
9.		Fraz [15]	2011	0.7152	0.9759	0.9430	---
10.		You [16]	2011	0.7410	0.9751	0.9434	---
11.	Supervised Methods	Niemeijer [17]	2004	---	---	0.9416	0.9294
12.		Soares [18]	2006	0.7332	0.9782	0.9461	0.9614
13.		Staal [5]	2004	---	---	0.9441	0.9520
14.		Ricci [19]	2007	---	---	0.9595	0.9558
15.		Lupascu [20]	2010	0.7200	---	0.9597	0.9561
16.		Marin [21]	2011	0.7067	0.9801	0.9452	0.9588
17.		Fraz [6]	2011	0.7406	0.9807	0.9480	0.9747

According to the principles of supervised and unsupervised classification methods, supervised classification relies on the properties of the training samples. If the training samples have a general offset, the parameters obtained after the training session will have an offset as well. The classification accuracy will drop with normal samples.

Table 2.1 gives a comparison of previous methods. This comparison includes SN (sensitivity), SP (specificity), Acc (accuracy), and AUC (area under the curve). These metrics are explained in Chapter 2.2 . Previous methods are categorized into unsupervised methods and supervised methods. The methods listed in Table 2.1 are tested on the DRIVE database, and those in Table 2.2 are tested on the STARE database.

Table 2.2: Classification methods (STARE database)

No	Type	Methods	Year	SN	SP	Acc	AUC
1.	Unsupervised Methods	Hoover [7]	2000	0.6747	0.9565	0.9264	---
2.		Zana [8]	2001	---	---	---	---
3.		Jiang [9]	2003	---	---	0.9009	---
4.		Mendonca [10]	2006	0.6996	0.9730	0.9440	---
5.		Lam [11]	2008	---	---	0.9474	0.9392
6.		Al-Diri [12]	2009	0.7521	0.9681	---	---
7.		Lam [13]	2010	---	---	0.9567	0.9739
8.		Miri [14]	2011	---	---	---	---
9.		Fraz [15]	2011	0.7311	0.9680	0.9442	---
10.		You [16]	2011	0.7260	0.9756	0.9497	---
11.	Supervised Methods	Niemeijer [17]	2004	---	---	---	---
12.		Soares [18]	2006	0.7207	0.9747	0.9479	0.9671
13.		Staal [5]	2004	---	---	0.9516	0.9614
14.		Ricci [19]	2007	---	---	0.9584	0.9602
15.		Lupascu [20]	2010	---	---	---	---
16.		Marin [21]	2011	0.6944	0.9819	0.9526	0.9769
17.		Fraz [6]	2011	0.7548	0.9763	0.9534	0.9768

To compare the performance of supervised and unsupervised classification methods, the average accuracy was calculated on the DRIVE and STARE databases. Figure 2.1 shows the average accuracy of supervised and unsupervised methods on DRIVE. As indicated, only SN exhibits a closer value between supervised and unsupervised methods. For the other three metrics, supervised ones all exhibit higher values than unsupervised ones.

The situation is the same for STARE. As shown in Figure 2.2, supervised methods all have higher values than unsupervised ones.
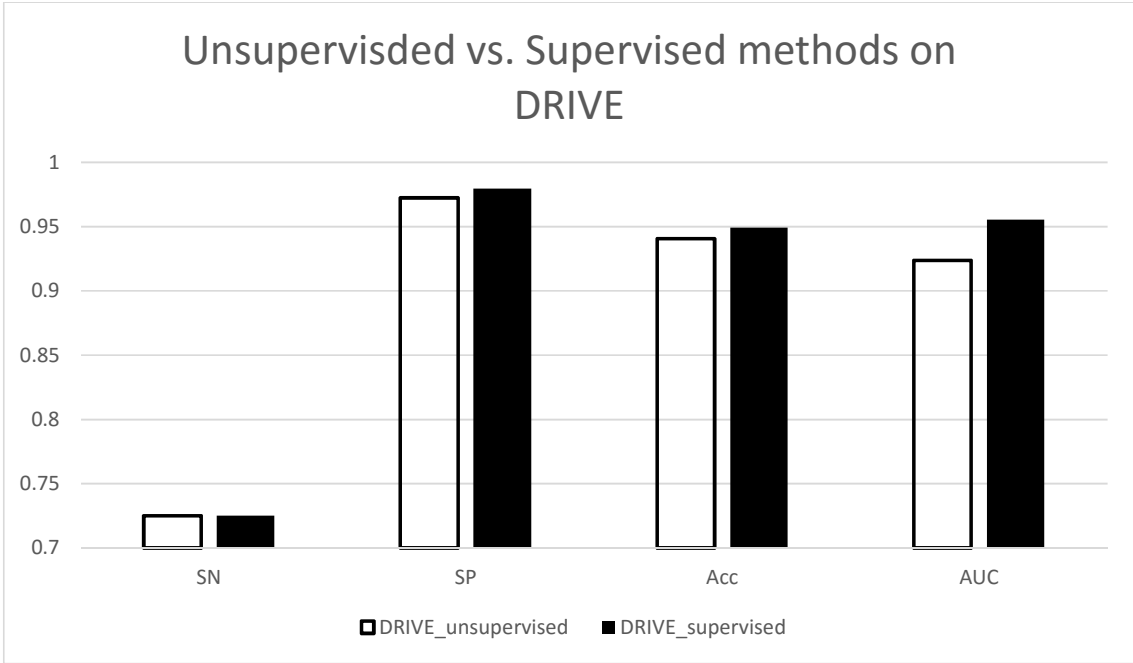
Figure 2.1: Comparison of classification methods on DRIVE
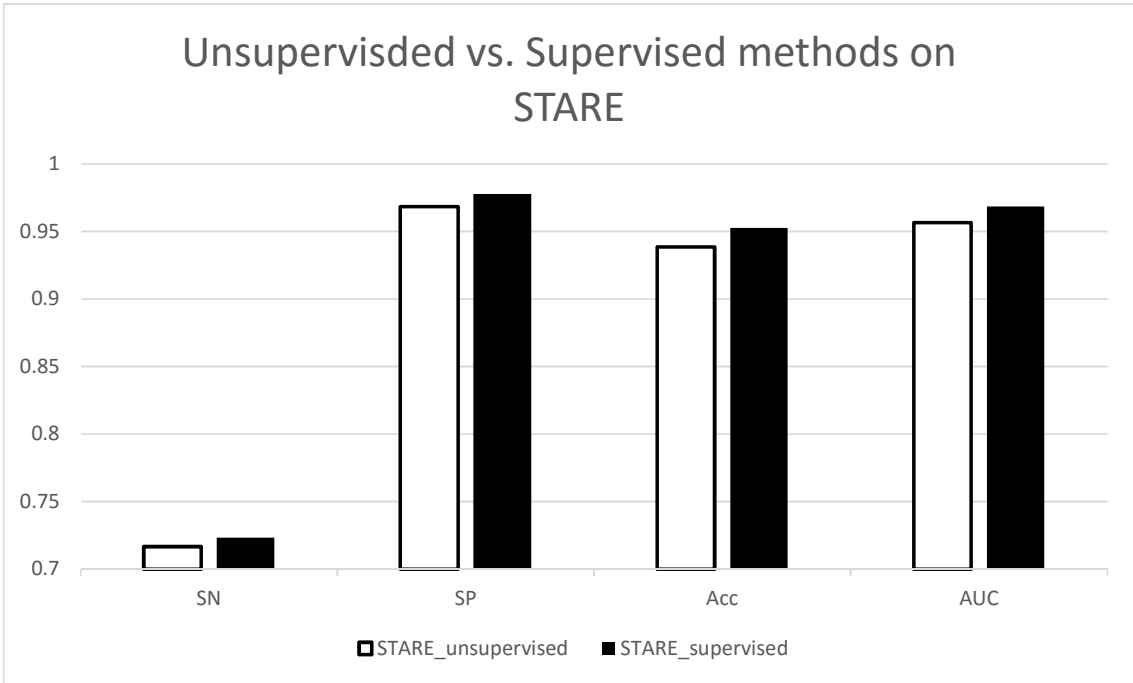
Figure 2.2: Comparison of classification methods on STARE

In general, supervised methods achieve better performance than unsupervised ones. However, one potential drawback in this comparison is that the comparison is done within the same database. In other words, the comparisons of supervised and unsupervised SN, SP, Acc, and AUC are all done in the same database. For the supervised methods, the training and classification sessions are done in the same database. If an unknown database were used, the classification on this new database might drop.

To differentiate our comparative metrics from others, two measurements are used here. One is vessel tracking ability, and the other is robustness. Vessel tracking ability measures how well the algorithm can highlight the vessel structure without interference from other databases. Robustness indicates how much the algorithm relies on training databases.

In general, supervised methods have better vessel tracking ability with worse robustness because they rely on the properties of the training samples. Unsupervised methods have better robustness with lower vessel tracking ability because they do not employ training samples to help build the model.

In this thesis, a supervised method is selected because the first priority is the vessel tracking ability. The target of this thesis is to achieve higher robustness with less degradation with respect to vessel tracking ability.

2.2 BINARY CLASSIFIER PERFORMANCE

As mentioned in a previous section, segmentation is helpful for the early detection of diabetic retinopathy. Segmentation is a common step for other algorithms; it is quite useful in practical medical devices.

In retinal vessel segmentation, grayscale images are inputted to the classifier, and the classifier will classify every pixel into the vessel class or non-vessel class. The input for the

classifier is grayscale images, and the output is ground truth images. Thus, the classification procedure is a binary classification. Binary classification has several metrics to measure its performance, such as sensitivity and specificity.

Table 2.3 shows four cases of classification results; each case can be represented by TP, TN, FN, or FP. Another important metric of accuracy is the summation of TP and TN, and accuracy is the most important metric in retinal vessel segmentation.

Table 2.3: Binary classifier performance

	Original	Classified result	Performance
Correct classification	Vessel	Vessel	True positive (TP)
	Non-vessel	Non-vessel	True negative (TN)
Incorrect classification	Vessel	Non-vessel	False negative (FN)
	Non-vessel	Vessel	False positive (FP)

The method used to measure the segmentation method is the same as those to evaluate binary classifiers. The metrics are true positive (TP), true negative (TN), false positive (FP), false negative (FN), accuracy (ACC), and receiver operating characteristic (ROC) curve. Among these, ACC is the first priority for performance evaluation.

Though ACC is close to 100% in the algorithms, it is impossible to reach 100% for two reasons. The first is that the checking procedure is done pixel by pixel, and no algorithm can hit every pixel. The second reason is that the standard images are labelled by human beings, and different ophthalmologists will have different opinions on the vessel marking in the same image. New methods are not mature enough to be accepted by all researchers, so the pixel-by-pixel comparison method is still widely used to evaluate algorithm performance. At the same time, researchers are still improving algorithms to improve accuracy, even accuracy that is

measured pixel by pixel. Accuracy is a straightforward way to evaluate algorithm performance. Moreover, accuracy can be sorted into two types, vessel tracking ability and robustness. Vessel tracking ability is assessed within one database, which means that the training and classification are done in the same database. Vessel tracking ability is directly related to single-database testing, while robustness is directly related to cross-database testing. In cross-database testing, the training and testing steps are done in two different databases. In general, the DRIVE and STARE databases are widely used. Cross-database testing yields two sets of data, one from the training on DRIVE and testing on STARE and the other from the training on STARE and testing on DRIVE. These two metrics use the accuracy value differently, but they reflect different algorithm performance.

2.3 ACCURACY MEASUREMENT AT THE IMAGE LEVEL

The two public databases, DRIVE and STARE, both provide manually labelled images for comparison purposes. Manually labelled images make it easier to compare different algorithms because classified images can be compared using these manually labelled images, which are labelled by experts with rich clinic experience, so they can be considered standard ones.

Figure 2.3 is a manually labelled ground truth image, and Figure 2.4 is a classified ground truth image. Because they are binary images, scanning and comparing every pixel is the easiest way to compare them. This method has a significant disadvantage. Vessel connectivity is not counted in this metric, and some unreliable method could contribute to high accuracy. Researchers are attempting to develop a more practical method to measure algorithm performance, but the current proposals are not mature enough to be widely used.
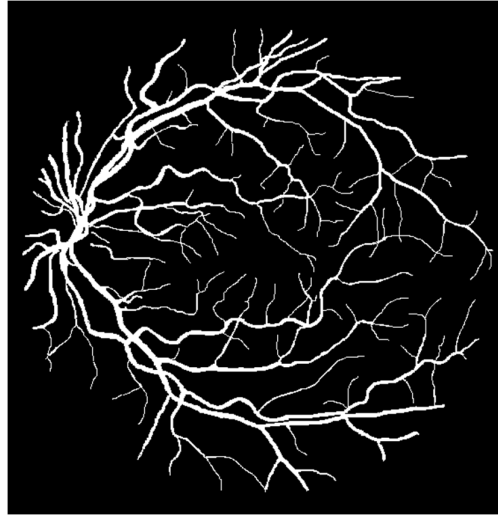
Figure 2.3: Manually-labelled image
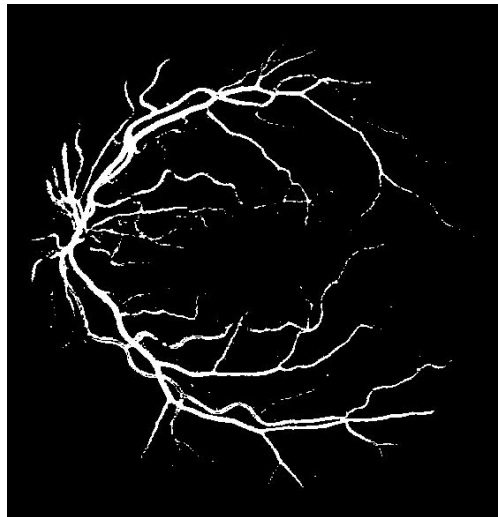
Figure 2.4: Classified image

2.4 ACCURACY MEASUREMENT AT THE PIXEL LEVEL

Accuracy is measured pixel by pixel, so the measuring program must scan pixels column by column. If a classified pixel is the same as a manually labelled one, this pixel can be counted as a correctly classified pixel, and the accuracy will increase. Figure 2.5 shows a manually

labelled matrix, or a manually labelled image, and it can be considered a standard matrix. Figure 2.6 shows the image or matrix after classification; green pixels are correctly classified pixels, while red ones are incorrectly classified pixels. Green ones contribute to accuracy, and red ones detract from it.

1	1	1	1	1	1
1	1	1	1	1	1
1	1	0	1	1	1
1	1	0	1	1	1
1	0	1	0	1	1
1	1	1	1	1	1

Figure 2.5: Manually-labelled matrix

1	1	1	1	1	1
1	1	1	1	1	1
1	1	0	1	1	1
1	1	0	1	1	1
1	0	0	0	1	1
1	1	0	1	0	1

Figure 2.6: Classified matrix

3 METHODOLOGY AND ANALYSIS

The inputs to the algorithm are three-channel RGB images obtained from the DRIVE and STARE databases. The outputs of the algorithm are binary ground truth images whose vessel structures have been highlighted. The highlighted vessel structures are in white, and the background is in black. According to the function of the algorithm, the algorithm can be considered as a binary classifier. The two classes are the vessel class and the non-vessel class. The job of the algorithm is to classify every pixel of an unknown image into the vessel class or the non-vessel class. The decision rule is based on Bayesian classification theory. Every pixel in the vessel class is assigned to '1' (1 bit), while pixels in the non-vessel class are assigned to '0' (1 bit). To improve the algorithm performance, an independent thin vessel highlighting step has been performed separately to highlight thin vessels.

3.1 PRE-PROCESSING

The images obtained from public databases are three-channel RGB images, and every pixel is 8 bits wide. In these RGB images, the green channel gives the highest contrast for blood vessels, so, in the following steps, mainly green channel images are used. The red channel has the highest average pixel value and the lowest contrast between vessels and non-vessels, so it can be used to generate a mask. Because this mask needs the central area to be 255 (8 bits) and the outer region to be 0 pixels (8 bits), red channel images are easy to dichotomize because their central values are high.
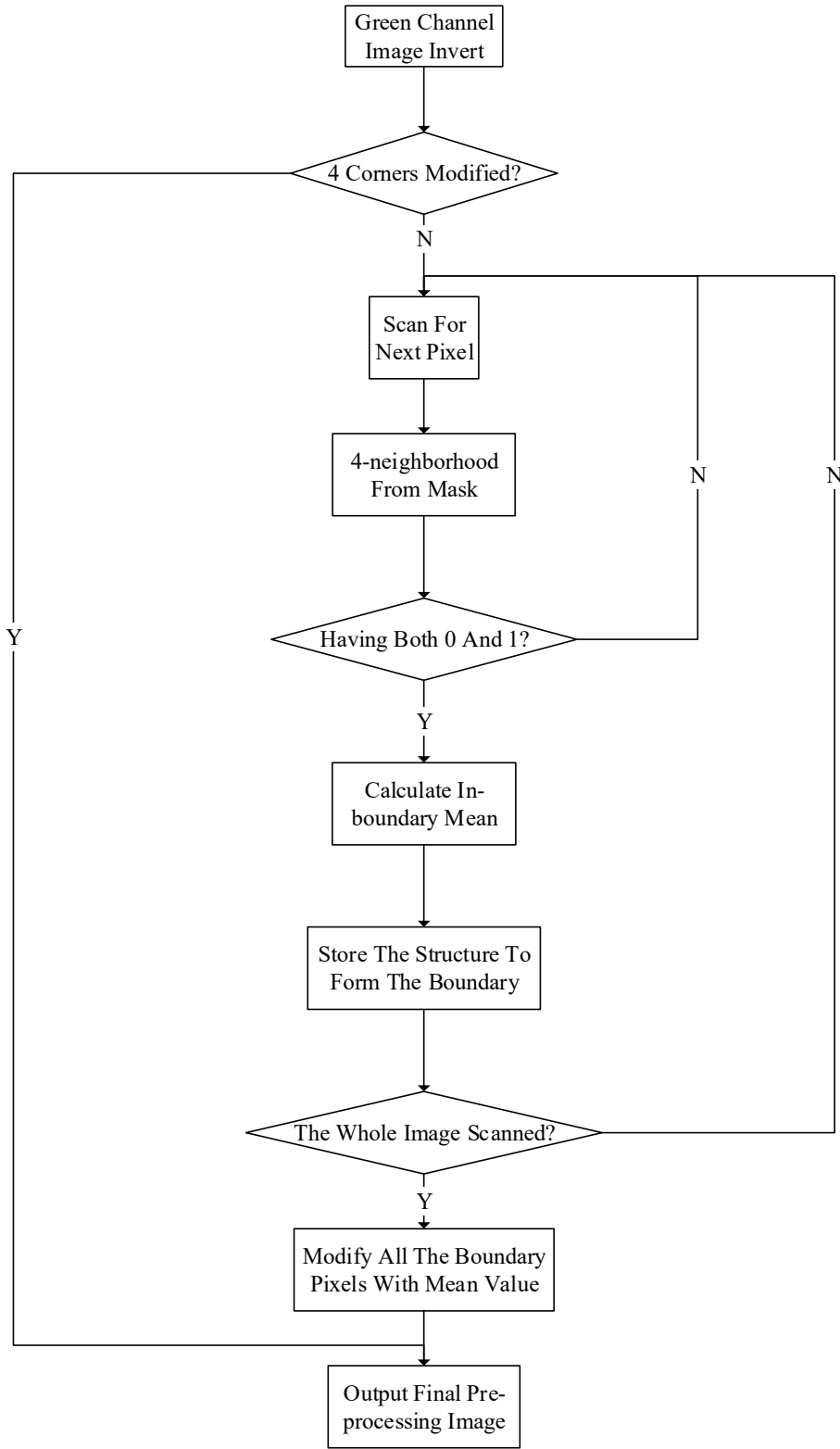
Figure 3.1: Flowchart for pre-processing

Pre-processing is a common component in image processing and its main target is to help subsequent steps to improve in accuracy or another performance measure. In retinal vessel segmentation, the central color area is the area that the program is concerned with, while the outer black area is meaningless to the segmentation. In this pre-processing step, the outer black area is eliminated and filled with a color similar to that of the central area. Pre-processing is an iterative procedure, and the mask is modified in every loop to mark the target pixel and tell the program when to stop. In this pre-processing step, four neighborhoods are selected, and only pixels inside the border are counted to calculate the mean value, and the central pixel is replaced by this mean value. However, the four neighborhoods are not necessarily close to the central pixel; it can be further pixels with another distance, which is 10 pixels in this thesis because 10 pixels was found to be the best distance after some experiments. This method is mainly based on Soares' paper [18]. Here, optimization has been done, with the distance upgraded to 10 pixels.

Figure 3.2: Original image

Figure 3.3: Pre-processed image

Figure 3.1 shows the flow chart for the pre-processing component in this thesis. The main function of this component is to eliminate the outer black area and to adjust its values so that they are as close as possible to that of the inner fundus background. As the flow chart indicates, this pre-processing component is an iterative procedure. In this procedure, four neighborhoods are used instead of eight to save computational time, and the distance was set at 10 pixels after some experiments. Scanning is done column by column. As stated previously, original RGB images have the best contrast for vessels in the green channel. The red channel has a higher average value throughout the fundus region, so it is more convenient to generate a mask from a red channel image. Every channel is 8 bits wide, and the value range is from 0 to 255. The value of 50 is set as the only threshold to generate a mask. The value over 50 is set to 255 (8 bits), and the pixels whose values are not over 50 are set to 0 (8 bits). This gives a perfect indicator to mark where the outer black area is, and it can be used to eliminate the outer black area in the following sections. The neighborhood from the mask is used to make a judgement, and the neighborhood from the original image is used to perform a calculation. If the four neighborhoods have both 0s and 1s, this central pixel is the potential boundary point; if the four

neighborhoods have all 0s, it means that they are in the black area, outside the region of interest. If they have all 1s, it means that they are in the fundus region, and the program does not need to modify these values. Because the neighborhood values influence the central pixel, all potential pixels can be modified only after the scanning of all pixels in the image. Every potential pixel must be stored in a structure to be modified later. The structure mainly needs to store the location and mean value of the central pixel. After every potential boundary pixel is obtained, this round can be finished, and all potential pixels can be modified.
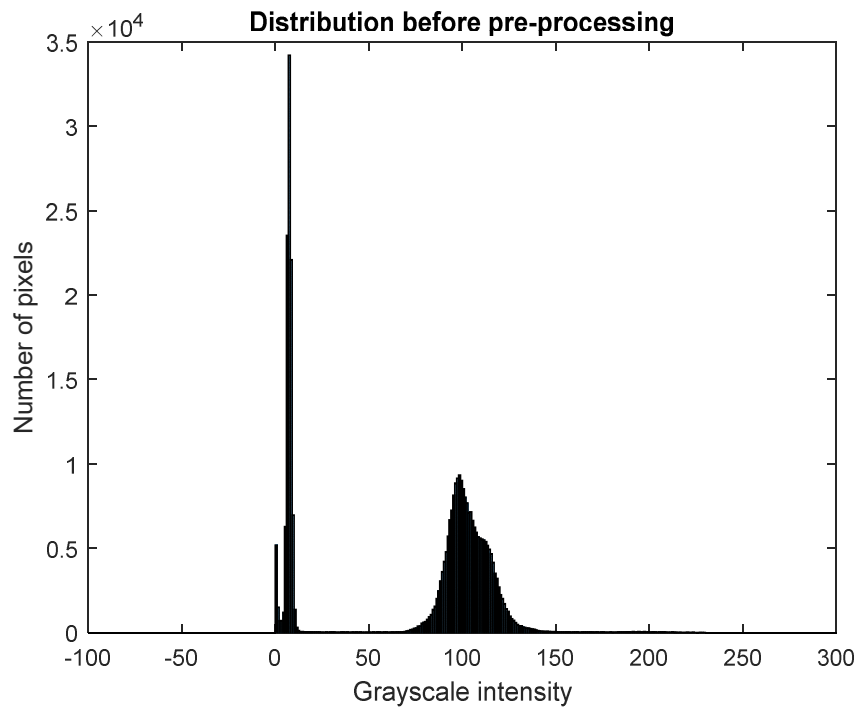
Figure 3.4: Grayscale distribution before pre-processing

The two images are shown in Figures 3.2 and 3.3 to demonstrate the performance of the pre-processing component. Figure 3.2 shows the original image from public database, and Figure 3.3 shows the result after pre-processing, where the outer black area has been eliminated and the color has been replaced by the central fundus color. This operation is helpful for subsequent steps.
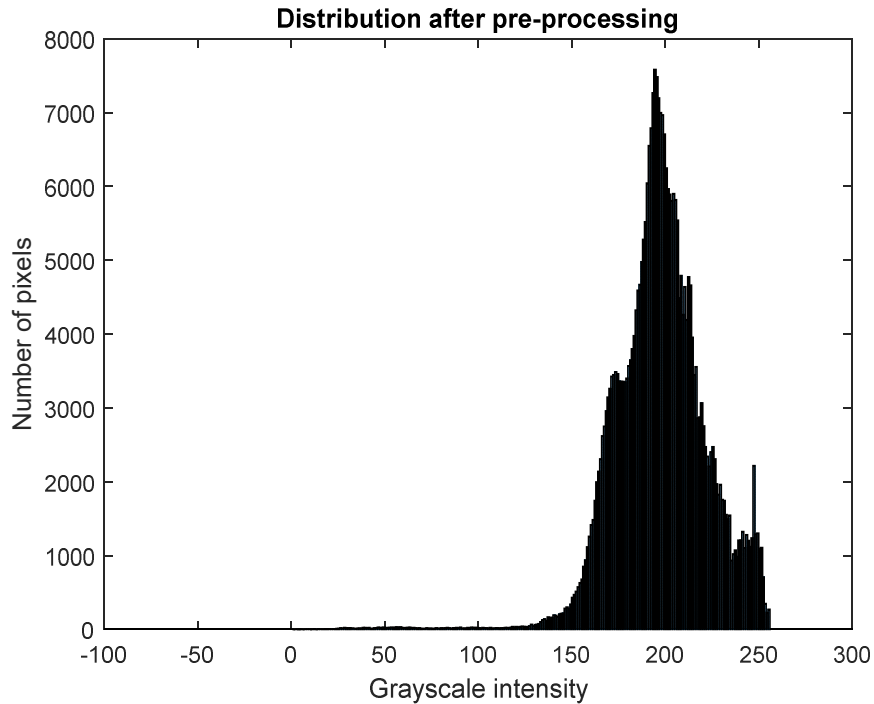
Figure 3.5: Grayscale distribution after pre-processing

To obtain a better understanding of the usage of pre-processing component, grayscale distribution histograms are shown in Figure 3.4 and Figure 3.5. Figure 3.4 shows the histogram before pre-processing, and it matches Figure 3.2 in the green channel. Figure 3.5 matches Figure 3.3. Before pre-processing, the grayscale intensities are the highest in two distributed ranges, [0 20] and [50 150]. The [0 20] range reflects the outer black area, and this area is meaningless for further processing. Moreover, it is harmful for the training procedure. Vessel structures and fundus are in the [50 150] range, and their values are too close to be classified. In the subsequent reconstruction step, vessel structures are assigned different values from the background.

Figure 3.5 has a continuous range, [150 255]. This range has already been expanded compared to Figure 3.4. Figure 3.4 has most of its samples in [75 140], while Figure 3.5 has

25

most of its samples in [125 250]. The range in Figure 3.5 is large, which means that the contrast is higher than in Figure 3.2. Higher contrast can help the classification achieve better accuracy because overlapping will have fewer samples.

3.2 MATHEMATICAL MORPHOLOGY (BINARY IMAGE)

The mathematical morphology theory was first proposed in 1987 by Robert et al. in [22]. It was a milestone in image processing; the technique provided a new view of image analysis, as other techniques focus mainly on frequency-space analysis. In frequency analysis, various types of transforms are usually needed to change the original values, which might cause problems when the values increase or decrease, especially for classification. If the original values increase or decrease, the classification might lose accuracy. Mathematical morphology does not require a change in the original values, as it usually uses other existing values to replace the original value, which means that the pixel value is replaced by another pixel value.

The theory is based on set theory. An original image can be considered a set (set A), and Structural Element (SE) is the description of an operation, such as shifting. The result might be a subset of set A or the extension of set A. For an erosion operation (Chapter 3.2.1), the result will be a subset of the original set. However, for a dilation operation (Chapter 3.2.2), the result will be an extension of the original set.

Mathematical morphology is constructed by erosion and dilation operations. These two operations are basic operations, and other advanced operations are constructed by these two. Opening, closing, and top-hat operations are all based on these two basic operations, erosion and dilation. However, mathematical morphology can be implemented with binary images, grayscale images, and color images, and the theory itself is the same.

Mathematical morphology has been widely used in many fields, such as image sharpening, image segmentation, image blurring, and other tasks. Here, mainly the top-hat transform has been used because the top-hat transform is sensitive to ridge-like and peak-like structures. In retinal images, vessels are like ridges on the flat fundus. As a result, the top-hat transform is sensitive to vessels, and it can highlight vessel-like structures.

3.2.1 EROSION (BINARY)

The basic operations in mathematical morphology are erosion and dilation. (3.1) contains the definition of the erosion operation.

$$A \ominus S_E = \{(i,j) | f(i,j)_{S_E} \in A \text{ and } (i,j) \in A\} \tag{3.1}$$

In (3.1), A is the original set (original image). The symbol (i,j) refers to the points in set A. This definition is used to select a subset of A, and every element in this subset needs to meet the requirement. The requirement is that the original element from A and the shifting element by S_E is in set A. The symbol S_E represents a shifting operation, and this is the core component in mathematical morphology, a structural element. $f(i,j)_{S_E}$ is a shifting operation of point (i,j) by structural element S_E. The definition of shifting is given in (3.2).

$$f(i,j)_{S_E} = \{(i',j') | i' = i + s_i, j' = j + s_j, (s_i, s_j) \in S_E\} \tag{3.2}$$

The erosion operation can obtain the subset of set A; the number of elements in $A \ominus S_E$ should be less than the number of elements in A. In fact, the erosion operation can eliminate some of the elements in A; usually, the program will eliminate scattering noise. The type of

noise that can be eliminated depends on the shape of Structural Element (SE). The erosion operation can only eliminate noise smaller than SE.

The erosion operation can obtain different results only by changing the structural element S_E, and different SEs should be selected in accordance with different usages. For instance, a linear structural element should be used to highlight linear shapes, so a line-shaped SE should be chosen. Because retina vessels are isotropic, a disk-shaped structural element has been chosen.

0	0	0	0	0	0	0
0	1	1	1	0	0	0
1	1	1	1	1	1	1
0	1	1	1	1	1	1
0	0	1	0	0	0	0

Figure 3.6: Original matrix for demonstration

Figure 3.6 shows the original matrix in an explanatory demonstration for Mathematical Morphology (MM) erosion, and this matrix's size is 5 by 7 pixels in rows and columns. It is a binary matrix, and the principle for binary MM is addressed in this section. 0s can be considered as the background, while 1 can be treated as an object or shape. Figure 3.7 shows the SE in this demonstration. It is a disk-shaped SE whose radius is one pixel. Because the distance from the diagonal pixels to the central pixel is greater than 1, the diagonal pixels are not considered in the disk-shaped Structural Element (SE) whose radius is one pixel. After the erosion operation, the shape shown in Figure 3.6 will shrink because of the MM erosion. In other words, the number of 1s will decrease after the erosion.

0	1	0
1	1	1
0	1	0

Figure 3.7: Structural Element (SE)

In Figure 3.8, the principle of MM erosion is shown, especially for binary images. The green and red crosses are constructed by the Structural Element (SE) from Figure 3.7, and these two crosses have different center pixels. The pixels covered by the green cross are all 1s, while the pixels covered by the red cross are three 1s and one 0. The central pixel and shifting pixels (peripheral pixels) are all in original shape. According to the definition of Mathematical Morphology (MM) erosion, the green central pixel should be kept after erosion, while the red central pixel is discarded because it does not meet requirements to be kept.

Figure 3.8: Explanation of erosion at the pixel level

Figure 3.9 is the resulting matrix after Mathematical Morphology (MM) erosion. It demonstrates Figure 3.8, and the green center is kept, while the red center is discarded. After comparison, the shape in Figure 3.9 is smaller than the one in Figure 3.6, which means that the original shape has been diminished.

Figure 3.9: Result matrix after erosion of Figure 3.6

3.2.2 DILATION (BINARY)

Here, the definition of the dilation operation is given.

$$A \oplus S_D = \bigcup_{(i,j) \in A} f(i,j)_{S_D} \tag{3.3}$$

In (3.3), A is the original set, or the original image. The symbol S_D represents a structural element, and it can be considered as the shifting vector. It restricts the direction and distance of shifting. $f(i,j)_{S_D}$ is the shifting operation of element (i,j) by structural element S_D. The union operation gathers all shifting pixels from element (i,j). The dilation set is usually bigger than the original set, and the dilation set can be considered as an extension of the original set A.

From the definition, dilation is a combination of at least two sets, so the number of elements in $A \oplus S_D$ is larger than that of A. Shapes in the original image will be enlarged, especially on the edges.

Figure 3.10: Explanation of dilation at the pixel level

A demonstration is shown here to help explain Mathematical Morphology (MM) dilation. The original matrix is the same as the one in Figure 3.6, and the Structural Element (SE) is from Figure 3.7. This SE disk-shaped with a radius equal to 1 pixel, but its shape is similar to

a cross. In Figure 3.10, two green crosses show the shifting areas after shifting of the central

pixels, and the shifting is based on elements in SE. These two green crosses can cover the area

whose elements are 0s and 1s. Because dilation result is union of sets, the area covered by a

green cross will contain 1s after dilation. Every 1 in the original matrix needs shifting by SE,

and the green cross for every 1 in the original image needs to be united. Figure 3.11

demonstrates the performance of dilation. The original shape has been enlarged.

0	1	1	1	0	0	0
1	1	1	1	1	1	1
1	1	1	1	1	1	1
0	1	1	1	1	1	1
0	0	1	1	1	1	1

Figure 3.11: Result matrix after dilation of Figure 3.6

The erosion operation will usually shrink shapes in the original image, and some of the

scattering noise smaller than the Structural Element (SE) will be gone. In contrast, the dilation

operation will enlarge the shapes in the original image. The level of shrinkage or enlargement

depends on the SE. As a result, erosion and dilation could change the original image

significantly, which is not very helpful in image processing. Image recovery is also needed to

keep the main information from the original image. Erosion results need to be recovered by

dilation, and dilation results need to be recovered by erosion. This leads to the definition of

mathematical morphology opening and closing.

3.2.3 OPENING AND CLOSING (BINARY)

(3.4) is the definition of opening, while (3.5) is the definition of closing. These two operations are formed by the basic operations, erosion and dilation. The orders are different, and so are the results. Opening can be used for de-noising, and closing might be used to connect objects.

$$\psi(A) = (A \ominus S_E) \oplus S_D \qquad (3.4)$$

$$\phi(A) = (A \oplus S_D) \ominus S_E \qquad (3.5)$$

As indicated by (3.4) and (3.5), the right side of the equations are formed by $\ominus$ and $\oplus$, which are erosion and dilation. The orders are different. Opening is done with erosion first and then dilation, while closing is done inversely. Erosion can eliminate the scattering noise, and dilation can recover the shapes. Closing is done first by dilation and then erosion. Dilation will expand the shapes, and some of the shapes might connect together if the gaps were smaller than the Structural Element (SE). Erosion will recover the shapes' sizes.

Erosion and dilation can change the size of shapes individually. However, in opening and closing, erosion and dilation occur in pairs, so the sizes of shapes will be recovered. In this way, the main information of the original image will be kept.

Opening and closing play different roles in image processing. Opening is always used to eliminate scattering noise, and closing can be used to connect different shapes in the original image.

Opening and closing consist of the basic operations, erosion and dilation. A more advanced operation, the top-hat operation, consists of opening and closing as well. Top-hat is

the most beneficial mathematical morphology method for retinal vessel segmentation in this thesis.

3.3 MATHEMATICAL MORPHOLOGY (GRAYSCALE IMAGE)

Mathematical morphology is a bit different for binary and grayscale images, but the core concepts are the same. In a binary image, the erosion operation depends on the shifting operation and whether the shifted point is still in the image. If it is still in the original image, the point will be replaced by 1; otherwise, the point will be replaced by 0. In a grayscale image, the erosion operation relies mainly on finding the minimum value over the Structural Element (SE). The SE is a binary mask whose target point is in the middle of the mask, and the 1s in the mask mark the detecting region. Finally, grayscale erosion is used to find the minimum value in the neighboring pixels marked by 1 in the SE.

For the dilation operation, a grayscale operation is used to find the maximum value in the neighboring region marked in the SE. In the binary dilation operation, shifted pixels will be included in the result set. As in grayscale erosion, the region marked by SE needs to be detected, and the dilation operation is used to find the maximum in this detecting region.

Binary mathematical morphology can be included in grayscale mathematical morphology, as grayscale is an extension of binary.

Because all mathematical morphology operations consist of erosion and dilation, all other operations can be derived from these basic concepts. Opening and closing consists of erosion and dilation. It can be understood further with the grayscale concepts, and demonstrations will be given in the following sections.

33

3.3.1 EROSION (GRAYSCALE)

Erosion is the basic technique in mathematical morphology; it can be used to eliminate the noise. Figure 3.12 is an original grayscale image, and the red circle represents Structural Element (SE). The reason for the red circle here is to demonstrate the performance of the mathematical morphology erosion. The red circle's diameter is 30 pixels, the vertical distance from the upper rectangle and the middle rectangle is 33 pixels, and the distance from the middle rectangle and the lower one is 15 pixels. The erosion performance relies on the SE. Here, the diameter of the SE is 30 pixels, and it can decrease the area of the rectangles. On the right upper side, one point is left because only this point and its neighbor are both in the original area.

Erosion can be used to remove noise in the original image. In the middle of the original image, three small spots are located from top to bottom. After the erosion operation, they are all removed from the original image.
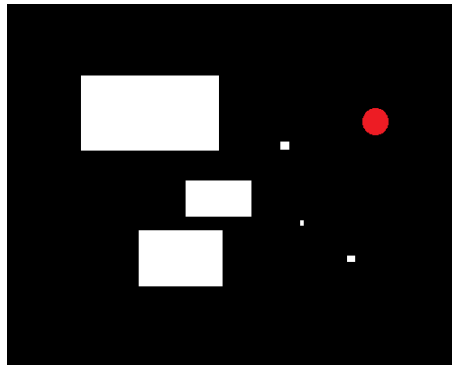

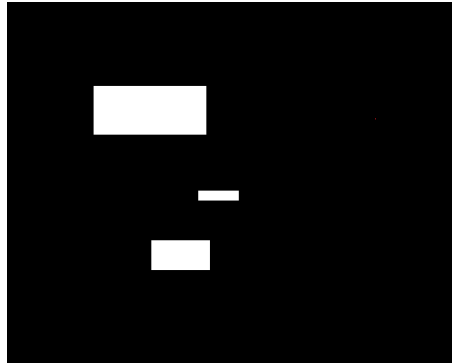Figure 3.12: Original grayscale image for demonstration

Figure 3.13: Mathematical Morphology (MM) erosion of Figure 3.12

3.3.2 DILATION (GRAYSCALE)

As stated before, dilation is the operation that can enlarge the area of the original shapes. In Figure 3.14, three rectangles in an original image have been enlarged after a dilation operation. The distance between the upper two rectangles is 30 pixels, so the Structural Element (SE) cannot cover the entire gap between the two. However, the distance between the lower two rectangles is 15 pixels, and it can be covered by the SE. If the SE were bigger than the gap, shapes could be connected together, as shown in Figure 3.14, where two lower rectangles have been connected together. Other separate shapes have been enlarged. After dilation, the sizes of shapes in the image have been enlarged significantly. The level of enlargement depends on the SE. In addition, the performance of linear SE and disk SE differs.
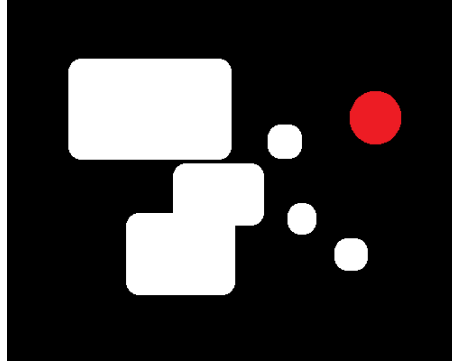
Figure 3.14: Mathematical Morphology (MM) dilation of Figure 3.12

3.3.3 OPENING (GRAYSCALE)

The opening operation is a combination of erosion and dilation: first erosion and then dilation. Because the first step is erosion, the noise in the original image has already been removed before the dilation. In addition, after erosion, the area of the rectangles decreases, and then the next dilation operation can restore the area to the original area. Because the Structural Element (SE) is round, the edges of the three rectangles are smoothed. In Figure 3.15, the scattering noise smaller than the SE has been removed. The edges of other separate shapes become smoother. The opening operation is an efficient way to eliminate noise and to keep the main information from the original image.
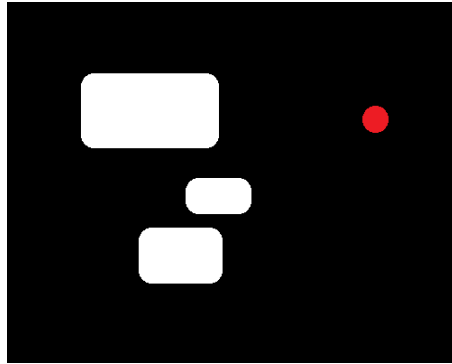
Figure 3.15: Mathematical Morphology (MM) opening of Figure 3.12

3.3.4 CLOSING (GRAYSCALE)

Closing is another combination of erosion and dilation: first dilation and then erosion. Because dilation is done first, the noise will be increased, and the erosion step cannot eliminate the noise spots. In addition, because the dilation is done first, the lower two rectangles can be connected, while the upper two rectangles cannot. Because the vertical distance between the upper two rectangles is 33 pixels, while the diameter of the Structural Element (SE) is 30 pixels. In Figure 3.16, parts of the original image have been connected together. Closing can be considered as an important method of image blurring.

Figure 3.16: Mathematical Morphology (MM) closing of Figure 3.12

3.4 TOP-HAT RECONSTRUCTION

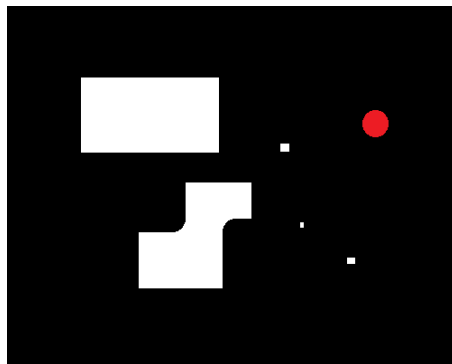The top-hat transform can be derived from previous definitions. The definition of top-hat transform is listed in (3.6). The bottom-hat transform is another form of the top-hat transform. The bottom-hat transform is defined in (3.7).

$$T_{top}(f) = f - \psi(f) \tag{3.6}$$

$$T_{bottom} = \phi(f) - f \tag{3.7}$$

The top-hat transform is sensitive to peaks, while the bottom-hat transform is sensitive to valleys. The top-hat operation is suitable for this application because vessels can be considered as ridges in the flat fundus. Moreover, the top-hat transform is not only sensitive to ridges and peaks but it can balance the background as well. In original images, the background is continuous in grayscale. Only the vessels have some significant changes in value, while other parts have smooth changes in grayscale values. After the top-hat transform, only sharp changes will be left, and smoothing changes can be removed. These smoothing changes have no meaning for vessel segmentation.

The top-hat transform is not only used for extraction of ridges and peaks but can also balance the background in different illumination conditions. Ridges and peaks have sharp changes in grayscale intensity values, and the other background area has only a smoothing change in gray level intensity.
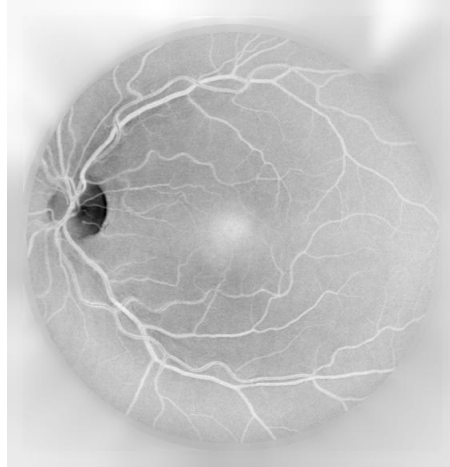
Figure 3.17: Pre-processed image

With different Structural Element (SE)s, the top-hat transform is sensitive to different shapes. If the SE is round, the top-hat transform is sensitive to isotropic structures, such as vessels. The radius of this disk shape is sensitive to vessels of different widths. After processing by various radius disk-shaped SEs, several new images will be generated with vessels retained while the background is eliminated. These images can be subjected to an "OR" operation to retain as many as vessel-like structures as possible. The "OR" operation is essentially an overlapping, and images of the same size will be overlapped. Every pixel will be replaced with the maximum value in the same position.
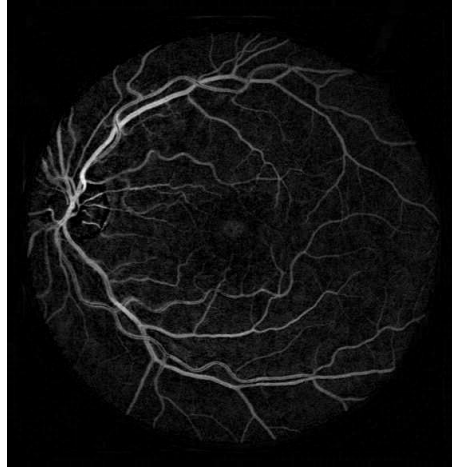
Figure 3.18: Reconstructed image

Figure 3.17 is an image after pre-processing. It has been transformed into a grayscale image. Figure 3.18 is the image after reconstruction. Vessels have been retained in the reconstructed image.

3.5 FEATURE GENERATION

Feature generation is the preliminary step for the training and classification procedure because the program needs feature or features to measure the clustering performance for input samples. If the data can be clustered into several groups, this is helpful for training and classification, and it also improves the classification accuracy. In previous papers, several or even dozens of features have been used to do the training and classification. More features do not necessarily mean higher accuracy.

In this thesis, only one feature has been used to do the training and classification. The grayscale intensity value is the only feature inputted to the trainer and classifier. The reason that only the grayscale value used here is that fewer features mean less computational time,

and other components, such as the mathematical morphology and pre-processing components, work to optimize this feature to compensate for its potential drawbacks.

The mathematical morphology operation is a helpful way to optimize feature generation because Mathematical Morphology (MM) is a transform that can narrow down the range of original grayscale values. Background values can be narrowed down close to 0 (8 bits), and vessels can be closer to 1 (8 bits).

Using grayscale as the only feature, this segmentation algorithm also achieves high accuracy, around 96%. This accuracy is very comparable with those of previous methods. Because only the grayscale value has been used, the MM operations are very time-efficient.

Table 3.1: Comparison of the number of features

Author	Year	Number of features
Ricci [6]	2007	3
Fraz [7]	2012	9
Marin [10]	2011	7
Proposed	2015	1

3.6 TRAINING

Through pre-processing and Mathematical Morphology (MM) reconstruction, images are transformed into optimized grayscale images. The demonstration is shown in Figure 3.18. The pre-processed image is lighter in general, while the reconstructed one is much darker in the background. The image after pre-processing, Figure 3.17, has an unbalanced background, and the vessel contrast is low because the whole image is lighter. After reconstruction, the background is balanced, and vessel-like structures are highlighted. The vessel contrast is better in Figure 3.18, where the background is much darker and the vessels are much lighter. The

reconstruction component plays a role in re-distributing grayscale values. The reconstruction step makes the background pixels closer to 0 (8 bits) and the vessel pixels closer to 255 (8 bits). This is helpful for further binary classification.
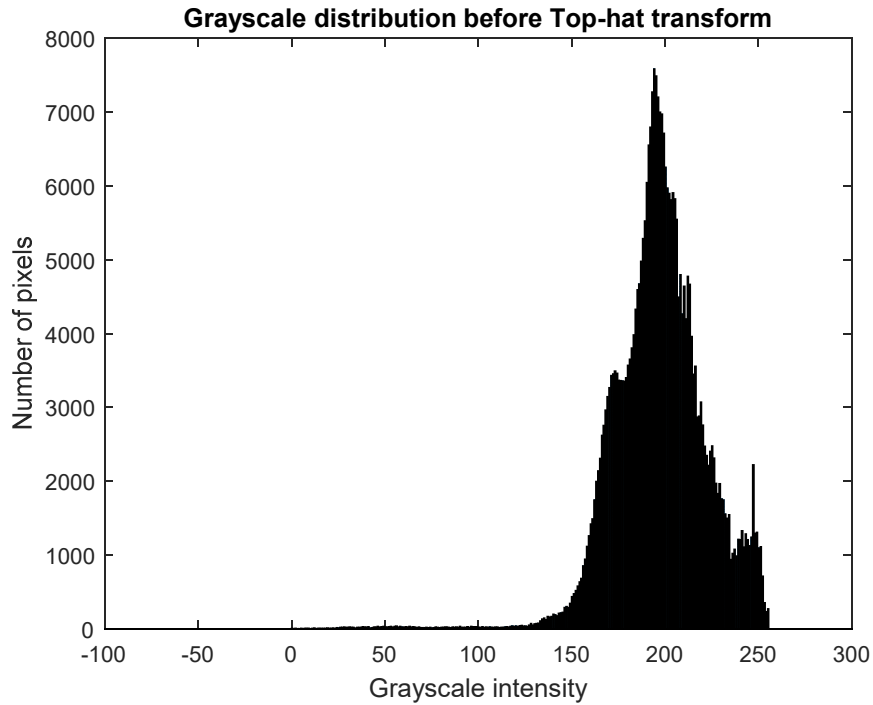
Figure 3.19: Grayscale distribution before top-hat transform

Figure 3.19 is the histogram of the grayscale distribution of Figure 3.17, which is the image after pre-processing. The distribution is quite continuous from 0 to 255 because the pre-processed image is still in 8 bits. From this grayscale distribution, it is very difficult to distinguish between vessel areas and background areas. This demonstrates that top-hat reconstruction is necessary to build a classification model.
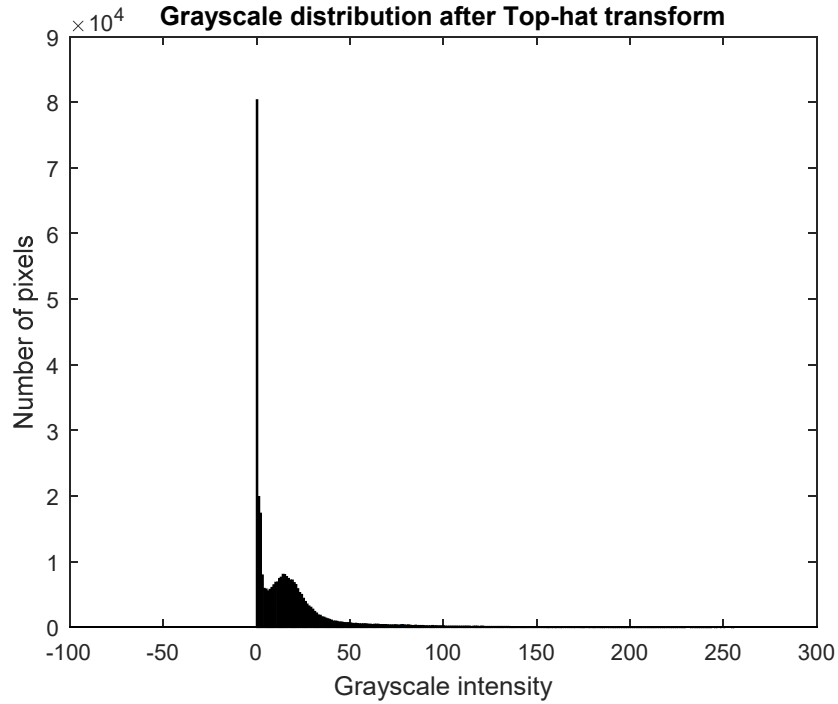
Figure 3.20: Grayscale distribution after top-hat transform

Figure 3.20 is the grayscale distribution after top-hat reconstruction. An obvious gap is shown in this histogram, and this gap might be the division value to differentiate between vessels and background. Because of the shape of this distribution, Gaussian Mixture Model (GMM) is a quite straightforward way to model the distribution, and it can provide a standard to classify unknown samples. In the training procedure, GMM will be trained and parameters will be obtained for the classification component.

A distribution like Figure 3.20 can be modelled by GMM. GMM is a combination of several individual Gaussian models. Every Gaussian model has its own configuration λ, which is shown in (3.8). Assuming that M Gaussian models have been used, M groups of $w_i, \boldsymbol{\mu}_i,$ and Σ_i will be generated to construct the total probability of an unknown sample.

$$\lambda = \{w_i, \mu_i, \Sigma_i\}, i = 1, \dots, M. \tag{3.8}$$

Unknown samples can obtain their own possibilities in GMM, according to (3.9). Assuming all parameters, the configuration of GMM has been achieved. Unknown sample x can obtain its own possibility $p(x|\lambda)$ by (3.9). In (3.9), the number of Gaussian models has been set as M. The summation of probabilities of sample x in every Gaussian model multiplied by their own weight can construct the possibility of sample x in GMM. In the classification module, the probability of the sample will be calculated, and it will use parameters obtained in the training session.

$$p(x|\lambda) = \sum_{i=1}^{M} w_i \, g(x|\mu_i, \Sigma_i) \qquad (3.9)$$

To obtain better understanding, a demonstration for GMM is shown in Figure 3.21. Only one dimension is used here for simplicity. The three independent Gaussian models in Figure 3.21 have the configuration of $\lambda_1 = \{0.1, -3, 1\}$, $\lambda_2 = \{0.5, 1, 1\}$, and $\lambda_3 = \{0.4, 5, 1\}$, and these three parameters are weight, mean value and standard deviation separately. The central Gaussian model has the highest weight, so it contributes the most to the probability. Figure 3.20, the distribution of the image grayscale, can be modelled by GMM with the same principle. Because grayscale is selected as the only feature here, this problem will become easier to understand.
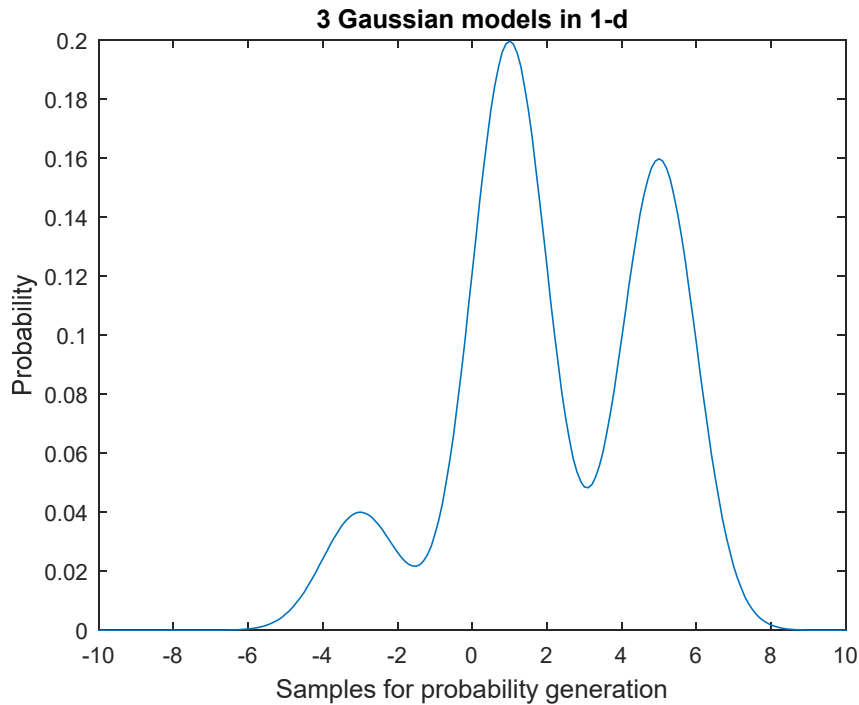
Figure 3.21: One-dimensional Gaussian Mixture Model (GMM)

The input data for the training session are grayscale images and corresponding manually labelled images. The grayscale images were obtained from the Mathematical Morphology (MM) reconstruction session, and the manually labelled images are from the DRIVE and STARE public databases. Every pixel in the grayscale image can be matched to a corresponding manually labelled pixel to be classified into one class, the vessel class or the non-vessel class. Before the iterative procedure for Gaussian Mixture Model (GMM) parameters is undertaken, two large sets of grayscale samples will be obtained according to the manually labelled images. Only one feature is used here, the grayscale value. The grayscale value varies from 0 to 255 (8 bits) because the image is 8 bits in width. After the previous reconstruction step, the range of background values was narrowed down closer to 0 (8 bits), and the range of vessel pixels was narrowed down closer to 255 (8 bits). This top-hat reconstruction step is helpful for the training procedure. The DRIVE and STARE databases

45

both provided manually labelled ground truth images to help compare the segmentation results. In addition, these manually labelled images can be used to train the classification model because they can help to classify every pixel into the vessel class or the non-vessel class.

The DRIVE database has 20 images to be trained, and STARE has 20 images to be trained. Let us assume that 1,000 pixels of one image are selected to form the vessel class, and the other 1,000 pixels are selected for the non-vessel class. Every training step requires 20,000 pixels to form the vessel class and the other 20,000 pixels to form the non-vessel class. Training Gaussian models is an iterative procedure, and the target is to get every individual Gaussian model's configuration, weight, mean value, and covariance. Assuming that 5 has been selected as the number of Gaussian models, two groups of Gaussian models need to be built, one group for the vessel class and the other group for the non-vessel class. Each group has five Gaussian models. In the following, iterative equations are listed to obtain the configuration of every Gaussian model.

$\Pr(i|\boldsymbol{x}_t, \lambda)$ is the intermediate variable for further calculation. (3.10) calculates the intermediate variable for every training sample for each Gaussian model. This equation is based on the existing configuration λ, so λ needs to have an initial value obtained initiated at the beginning.

$$\Pr(i|\boldsymbol{x}_t, \lambda) = \frac{\omega_i g(\boldsymbol{x}_t|\boldsymbol{\mu}_i, \Sigma_i)}{\sum_{k=1}^{M} \omega_k \, g(\boldsymbol{x}_t|\boldsymbol{\mu}_k, \Sigma_k)} \tag{3.10}$$

In (3.10), $g(\boldsymbol{x}_t|\boldsymbol{\mu}_i, \Sigma_i)$ is the Gaussian probability equation to calculate the probability of sample $\boldsymbol{x}_t$ probability in each Gaussian model with the configuration $\boldsymbol{\mu}_i, \Sigma_i$, and the definition of $g(\boldsymbol{x})_{\sigma,\mu}$ is in (3.10). (3.11) is the one-dimensional Gaussian distribution probability

equation. Because only one dimension, the grayscale value, is selected, a one dimensional equation is given here.

$$g(x)_{\sigma,\mu} = \sigma^{-1}(2\pi)^{-\frac{1}{2}}e^{-\frac{1}{2}(x-\mu)\sigma^{-2}(x-\mu)} \qquad (3.11)$$

In the training procedure, the Gaussian Mixture Model (GMM) model needs to update its configuration in every loop because GMM has a suboptimal initial configuration. $\overline{\omega_i}$ and $\overline{\mu_i}$ are calculation results to be assigned in the next loop. Because ω_i and μ_i change in every loop, $\Pr(i|x_t, \lambda)$ needs to be recalculated in every loop. (3.12) and (3.13) give the definitions of the calculation procedure. Once the program meets the specific range, the iterative procedure can stop and the final result will be obtained.

$$\overline{\omega_i} = \frac{1}{T}\sum_{t=1}^{T}\Pr(i|x_t, \lambda) \qquad (3.12)$$

$$\overline{\mu_i} = \frac{\sum_{t=1}^{T}\Pr(i|x_t, \lambda)x_t}{\sum_{t=1}^{T}\Pr(i|x_t, \lambda)} \qquad (3.13)$$

Gaussian Mixture Model (GMM) is a relatively straightforward model as a binary classifier. GMM can be considered as a combination of several Gaussian models, and the probability of the sample data is the summation of the probabilities of each Gaussian model. GMM is suitable for this application because the sample distribution is like two Gaussian distributions, one for vessel pixels and the other for non-vessel pixels. Configuration λ will be obtained for the vessel class, configuration λ' will be obtained for the non-vessel class, and

λ and λ' become classification parameters to calculate the vessel possibility and non-vessel possibility in the classification step.

3.7 CLASSIFICATION

The inputs for the classification session are continuous grayscale values after reconstruction. Its output should be binary values, or a ground truth image. This task of the classifier means that the classifier should be a binary classifier.

Here, one class is the vessel class, and the other is the non-vessel class. The standard to decide the class to which the sample belongs is the probabilities of the two classes. If the probability of the vessel class is bigger, the input sample should be put into the vessel class. If the probability of the non-vessel class is bigger, the input sample should be put into the non-vessel class.

After the training session, two groups of Gaussian models can be obtained, and each group has five independent Gaussian models. One group matches the vessel class, and the other matches the non-vessel class. Each of the 10 Gaussian models has its own configuration, mean value, covariance, and weight. The vessel class probability for the input sample can be accumulated by the five Gaussian models for the vessel class, and the non-vessel class probability of the input sample can be obtained by the five Gaussian models for the non-vessel class.

According to the Bayesian classification decision, the probability for the vessel class, C_v, is multiplied by the class-conditional probability $p(\boldsymbol{x}|C_v)$ and the class probability $P(C_v)$. The probability for the non-vessel class, $P(C_{n-v})$, is multiplied by $p(\boldsymbol{x}|C_{n-v})$ and $P(C_{n-v})$. $P(C_v)$ and $P(C_{n-v})$ are two prior probabilities, which are related to the two classes. The classification

decision is based on (3.14). If the probability for the vessel class is bigger, the sample x will be classified into the vessel class. Otherwise, the sample x will be classified into the non-vessel class.

$$p(x|C_v)P(C_v) > p(x|C_{n-v})P(C_{n-v}) \qquad (3.14)$$

The prior probabilities $P(C_v)$ and $P(C_{n-v})$ are calculated from the manually labelled images. Because these are binary ground truth images, it is easy to count the total number of vessel pixels and non-vessel pixels. (3.15) and (3.16) show how to calculate prior probabilities. Here, N_v represents the total number of vessel pixels, while N_{n-v} represents the total number of non-vessel pixels

$$P(C_v) = \frac{N_v}{N_v + N_{n-v}} \qquad (3.15)$$

$$P(C_{n-v}) = \frac{N_{n-v}}{N_v + N_{n-v}} \qquad (3.16)$$

(3.17) shows how to calculate class-conditional probability with Gaussian Mixture Model (GMM). In general, it is the summation of M Gaussian models. The unknown sample x has its own probability in every Gaussian model, and these M probabilities will be multiplied by P, the weight of every Gaussian model. Finally, these probabilities are added together to obtain the final probability of sample x in the vessel class, C_v, or the non-vessel class, C_{n-v}. The sample x can be n-dimensional data, but in this thesis, only a one-dimensional x is used in the classification module.

$$p(x \mid C_v) = \sum_{j=1}^{M} p(x|j, C_v) P_{v,j} \tag{3.17}$$

$$p(x \mid C_{n-v}) = \sum_{j=1}^{M} p(x|j, C_{n-v}) P_{n-v,j} \tag{3.18}$$

The classification probability is generated by summation of several Gaussian models, and the probability of every Gaussian model is defined in (3.11). In the classification step, every pixel in an image needs to be classified into the vessel class or the non-vessel class.

The images before classification have already been subjected to the top-hat transform, so the images that need to be classified are grayscale images whose vessels are closer to 0 (8 bits) and non-vessels are closer to 255 (8 bits), which is similar to Figure 3.18. The grayscale distribution is similar to Figure 3.20. The task of the classification step is to generate the vessel probability and non-vessel probability of every pixel in the image that needs to be classified.
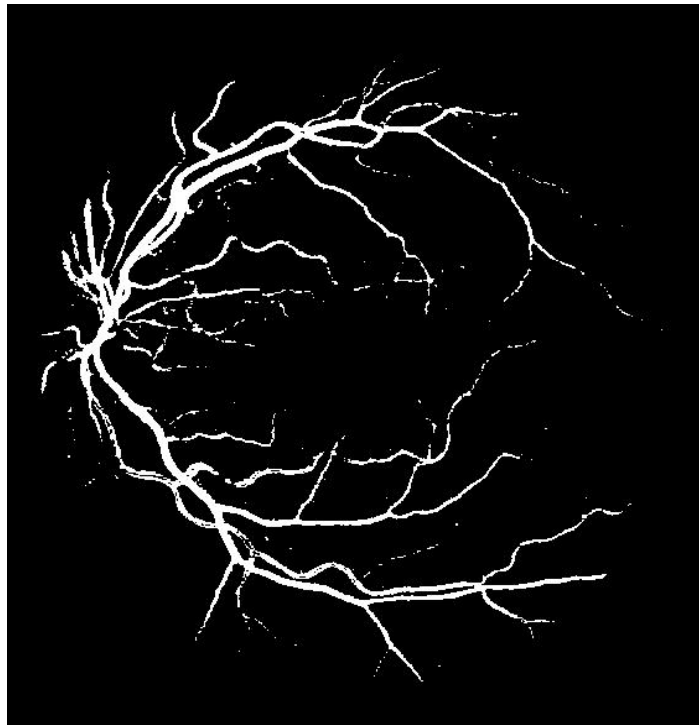

Figure 3.22: Classified image

Figure 3.22 is the binary ground truth image after classification. The vessel structure is already obvious in this image, but some thin vessel parts are missing because thin vessels' grayscale intensity is in the middle of the 0 to 255 (8 bits) range. In Figure 3.20, thin vessels' grayscale values are in the middle of the two clusters, so the vessel and non-vessel probabilities of thin vessels are closer. This causes thin vessels to be missing because some thin vessel pixels' non-vessel probability is bigger. The classified image in Figure 3.22 will contribute high accuracy. To improve the accuracy, an independent thin vessel tracking step is introduced.

3.8 THIN VESSEL DEFINITION

Thin vessels usually mean more in a clinic setting because the growth of thin vessels usually occurs with disease, such as diabetes. As stated previously, because of the blocking of nourishing vessels, the body needs to grow more thin vessels to nourish the retina. This is a symptom of mild non-proliferative retinopathy. As a result, thin vessel detection often indicates the presence of diabetes.

A potential drawback of the previous classification method is that thin vessels may be missing. The grayscale intensity values of thin vessels are in the middle of 0 to 255 (8 bits) range, and the vessel class and non-vessel class probabilities are close. As a result, it is hard to distinguish thin vessels from the background because thin vessels might very likely be classified into the non-vessel class.

A binary classifier is a bottleneck for thin vessel highlighting because the process is restricted by the performance of the binary classifier. For this reason, other biometrics need to be calculated to measure thin vessels. A derivative is a potential option because derivatives

could reflect sharp changes in grayscale values in some directions. The grayscale intensity will change in the horizontal direction near both for wide vessels and thin vertical vessels. This might reflect the presence of thin vessels.

In this thesis, a method for thin vessel detection is proposed. This method detects thin vessels separately. Wide vessel detection relies on classification theory, while thin vessel detection relies mainly on derivatives' properties.

In this newly proposed thin vessel highlighting method, two steps will be done sequentially:

1) Highlighting vessel centerlines:

The first step is to draw the skeleton of the vessels, or define the centerline of the vessels. In this step, wider vessels' centerlines and thin vessels will be retained.

2) Path-definition (Pixel connectivity elimination by threshold):

The second step is to find the connectivity of every pixel. The pixels that have a connectivity that is higher than a specific threshold will be retained; this step is helpful in eliminating the scattering noise.

3.8.1 CENTERLINE HIGHLIGHTING

Because the classifier exhibits poor performance in the detection of thinner vessels, the centerline highlighting component plays a complementary role for thin vessel detection. In the classification theory, thin vessels' grayscale intensity values are closer to that of the background. This makes it very difficult to separate thin vessels from the background.

The first step in thin vessel detection is centerline highlighting. This component's main task is to draw the skeleton of the vessel tree. Wide vessels' centerlines will be retained. For

thinner vessels, vessels can be left. A convolution kernel is used to highlight vessel centerlines. A separate component is based on deviation. Highlighting missing thin vessels could contribute to total accuracy.

To obtain every pixel's deviation, a convolution kernel is selected. This kernel is called the difference of Gaussians kernel, and it is derived from the Gaussian kernel. The deviation process is shown in Figure 3.25. The derivation is mainly composed of shifting and subtraction. Shifting is done for left shifting and right shifting. As shown in Figure 3.25, the difference of Gaussians kernel is obtained by subtracting the left shifting Gaussian kernel from the right shifting Gaussian kernel. This subtraction of two direction shifts could highlight edge changes, especially for a sharp change in one pixel range.

$$\begin{bmatrix} -1 & -2 & 0 & 2 & 1 \\ -2 & -4 & 0 & 4 & 2 \\ -1 & -2 & 0 & 2 & 1 \end{bmatrix}$$

Figure 3.23: Convolution kernel

$$[+ + --]$$

Figure 3.24: Sign combination

$$\begin{bmatrix} 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 2 & 4 & 2 \\ 0 & 0 & 1 & 2 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 2 & 1 & 0 & 0 \\ 2 & 4 & 2 & 0 & 0 \\ 1 & 2 & 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} -1 & -2 & 0 & 2 & 1 \\ -2 & -4 & 0 & 4 & 2 \\ -1 & -2 & 0 & 2 & 1 \end{bmatrix}$$

Figure 3.25: Derivation of the convolution kernel

Figure 3.26 is a convolved mesh figure by the difference of Gaussians kernel. The two edges of the same vessel yield different values. One is positive and the other is negative because one edge is a rising edge and the other is a dropping edge in the same scanning direction.

The convolution kernel in Figure 3.23 can only process vertical vessel centerlines. A rotation step will be performed later.
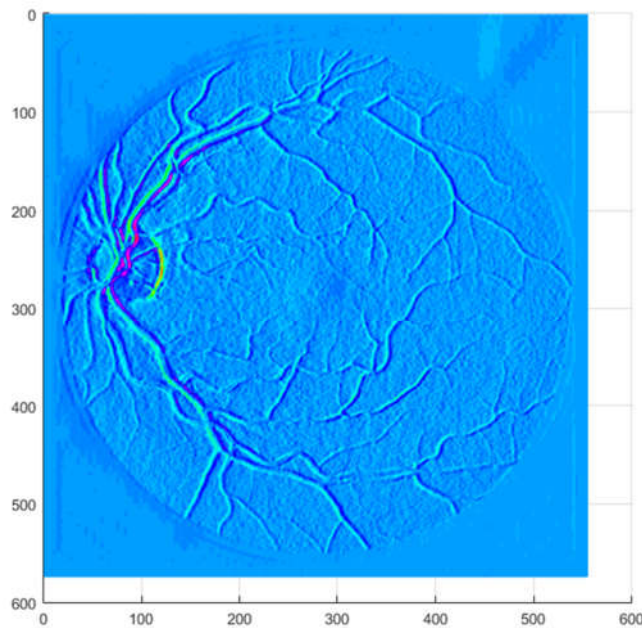

Figure 3.26: Convolved image by kernel

In Figure 3.26, vessels can be detected by scanning the sign combination. The sign combination is shown in Figure 3.24. Two positive signs indicate a continuous rising edge, and two negative signs indicate continuous dropping edge. This is a reflection of ridges, and it matches the vessel structures in grayscale images.
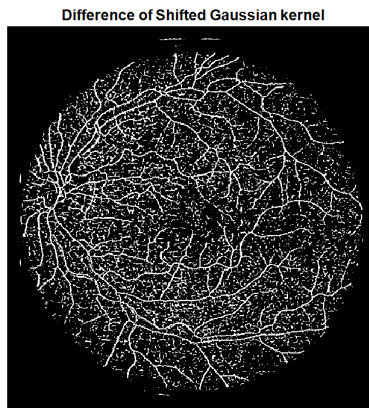
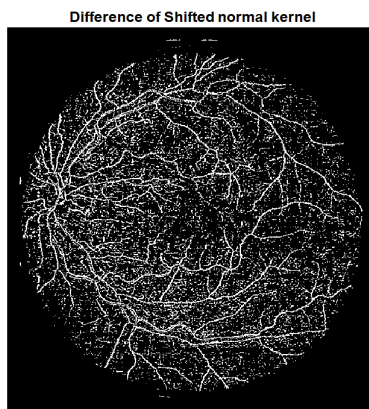Figure 3.27: Convolution kernel made by a shifted Gaussian kernel

Figure 3.28: Convolution kernel made by a difference kernel

Figure 3.27 and Figure 3.28 are comparisons of convolution results by different kernels. Figure 3.27 is the convolution result using the kernel in Figure 3.23; this kernel is the difference between the left-shifted Gaussian kernel and the right-shifted Gaussian kernel. The Gaussian kernel itself could blur the image. The blurred image has already had some scattering noise removed because the blurring procedure is related to every pixel's neighboring pixels. Here,

only the 3 by 3 Gaussian kernel affects the blurring result, so the blurring result could retain local properties while eliminating some scattering noise.

$$[-1\ 0\ 1]$$
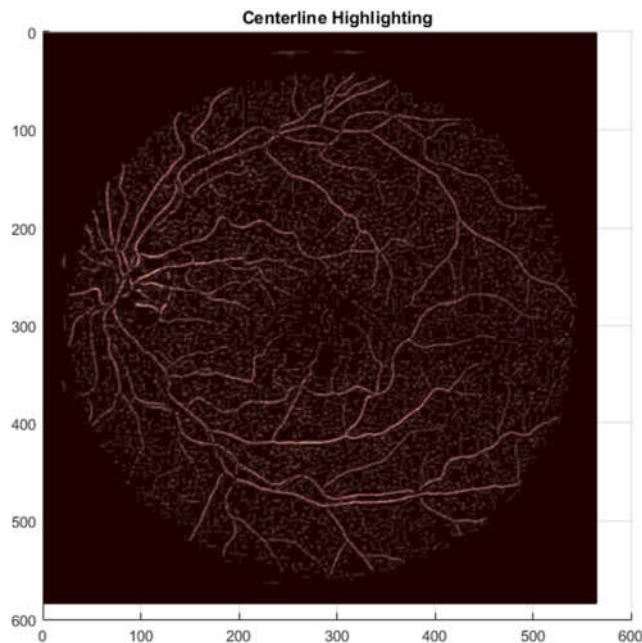
Figure 3.29: Difference kernel

Figure 3.30: Centerline highlighted image

In Figure 3.28, we can see that some small local areas are very noisy compared to the same location in Figure 3.27. The kernel is shown in Figure 3.29. This is the normal difference kernel without any blurring. This kernel is extremely sensitive to a change in grayscale values in the horizontal direction. The target in our thesis is sharp changes, but this kernel could also be sensitive to smoothing changes, which might occur in background areas. Background areas have similar grayscale values, and they have smoothing value changes. These changes are not

useful for segmentation. To eliminate this risk, Gaussian blurring is necessary to decrease the sensitivity of the kernel. As a result, the kernel shown in Figure 3.23 is much better because of its peak sensitivity and noise resistance. This is the reason that we chose the difference of Gaussians kernel.

After scanning the sign combination as shown in Figure 3.24, pixels meet the requirements to be modified. The absolute summation of the two neighboring pixels will replace the pixel's original value. This operation can only deal with the vertical vessels, and all directions would be processed. The final result of this operation would be shown in Figure 3.30.

In Figure 3.31, the flowchart of the centerline detection component is introduced. After a pre-processed image is obtained, the image needs to be inverted to adjust to the convolution kernel. To detect both edges of thin vessels, the difference of Gaussians kernel is used here. This kernel is sensitive to the edges of vessels. In addition, because of the size of the kernel, it is only sensitive to thinner vessels of one to three pixels in width. The convolution kernel is only sensitive to vertical vessels. After convolution, a sign matrix will be obtained and used for sign combination scanning. Only the specific sign combination is needed for the next operation. The combination of signs is shown in Figure 3.24. This combination of signs could reflect the presence of a thin vessel. After finding the combination of signs, the detected pixels' value will be replaced by the summation of the absolute values of the two horizontal neighboring pixels; the purpose of this operation is used to highlight the centerline of vessels. The algorithm until now has focused on vertical vessels. For vessels in other directions to be detected, the images need to be rotated; 10 degrees is selected as the rotation step here.
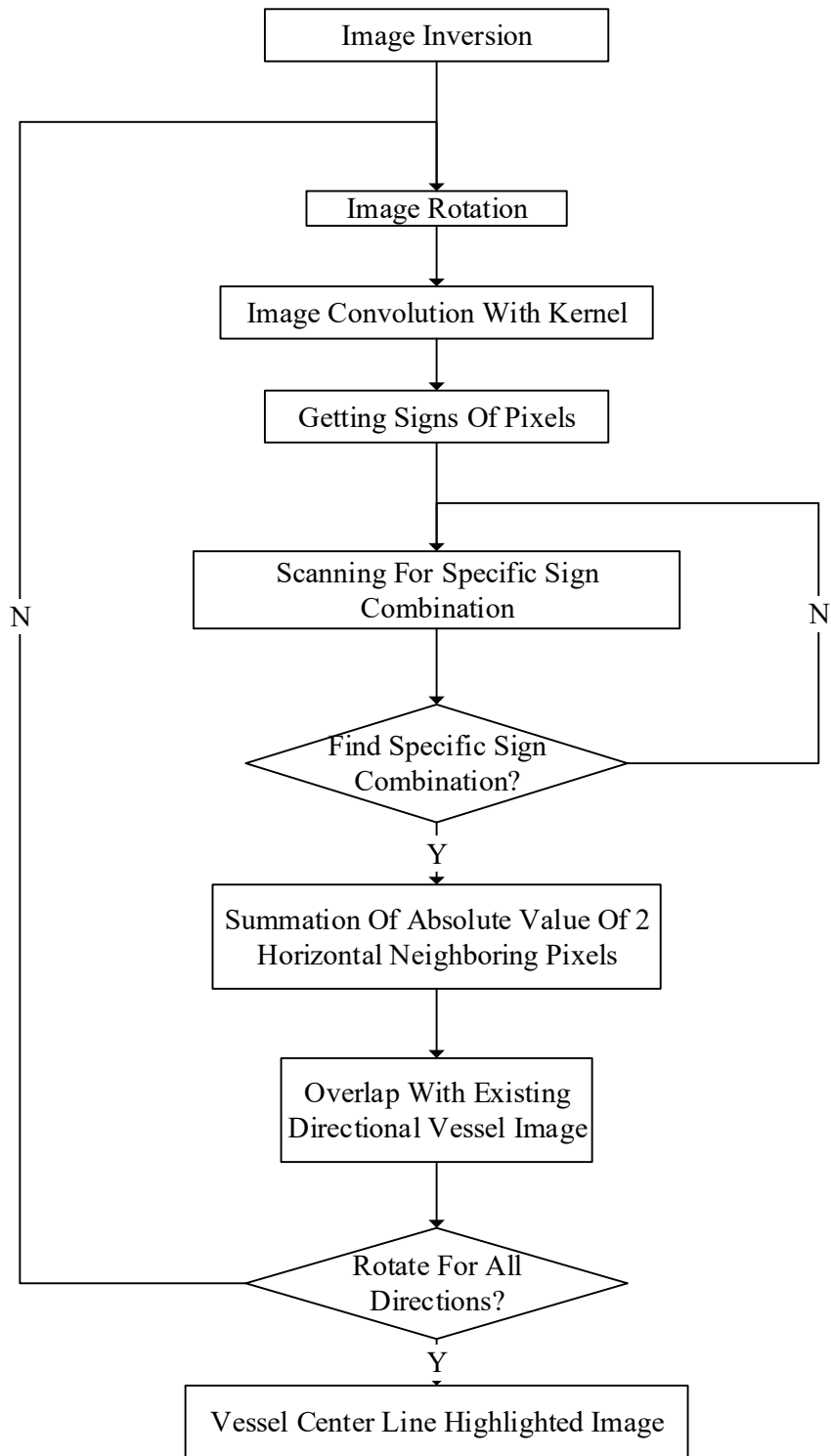
Figure 3.31: Flowchart for centerline highlighting

3.8.2 PATH-DEFINITION

After obtaining centerline highlighted images, such as Figure 3.27, an empirical threshold will be selected to dichotomize the images. In the manually labelled images, around 87% of pixels are background or non-vessel pixels, and the remaining 13% of pixels are vessel pixels. The size of DRIVE images are 584×565 pixels, so the 287,065th pixel's value is set as the threshold to dichotomize the DRIVE images. In contrast, STARE images are 605×700 pixels in height and width, so the threshold for STARE is set as the value of the 368,445th pixel. The first pixel is 0 (8 bits), the minimum value, and the last pixel is 255 (8 bits), the maximum value. In accordance with these two values, centerline highlighted images can be dichotomized into binary images. The following is the result for a binary image after using the empirical threshold. As shown in Figure 3.32, most thin vessels can be retained. Further processing is based on this image, and this part of thin vessel processing plays a complementary role in vessel segmentation.
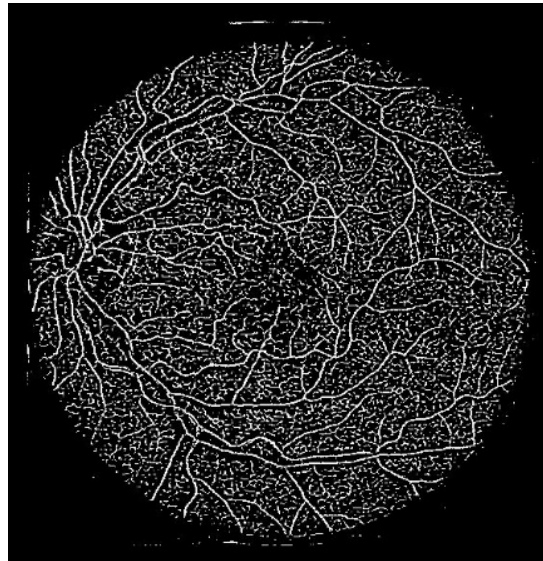

Figure 3.32: Image after centerline highlighting and empirical dichotomizing

Figure 3.32 shows a mix of vessel structures with scattering noise. To increase the accuracy, scattering noise needs to be eliminated as much as possible. In this path-definition component, the area, or connectivity serves as a measurement standard. Vessel pixels are connected with each other, and the area is significantly higher than the scattering noise.
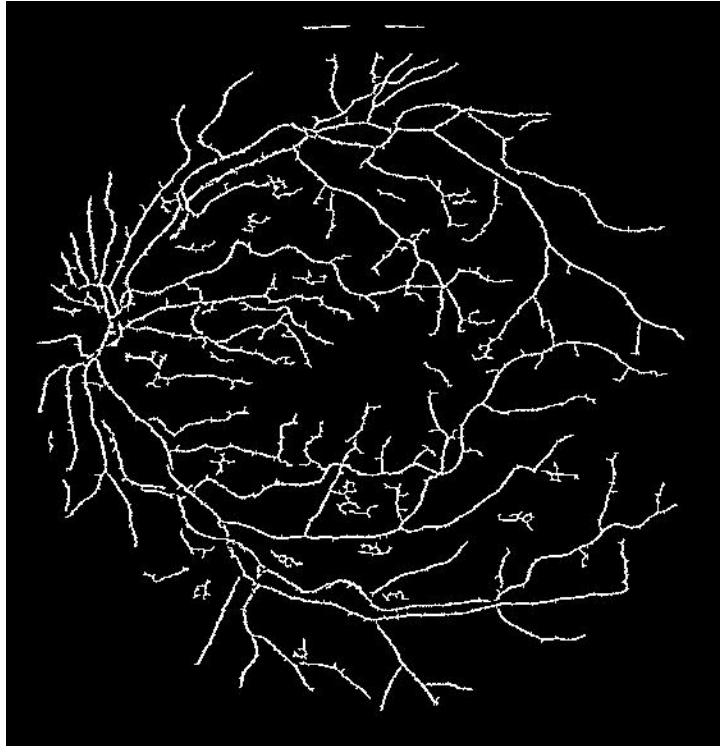
Figure 3.33: Image after path-definition

As can be seen in Figure 3.33, paths have been defined according to pixel connectivity. Pixels with greater connectivity can be kept, because vessel pixels always have other vessel pixels as neighbor. On the contrary, noise pixels have smaller connectivity, because they have local connectivity with a relatively smaller value. A threshold can be set to eliminate smaller area. 15 is selected as the threshold according to some experiments.

Figure 3.34: Flowchart for path-definition

Figure 3.34 shows the flowchart of the scattering noise removal algorithm. The algorithm's target is to retain potential vessel structures and remov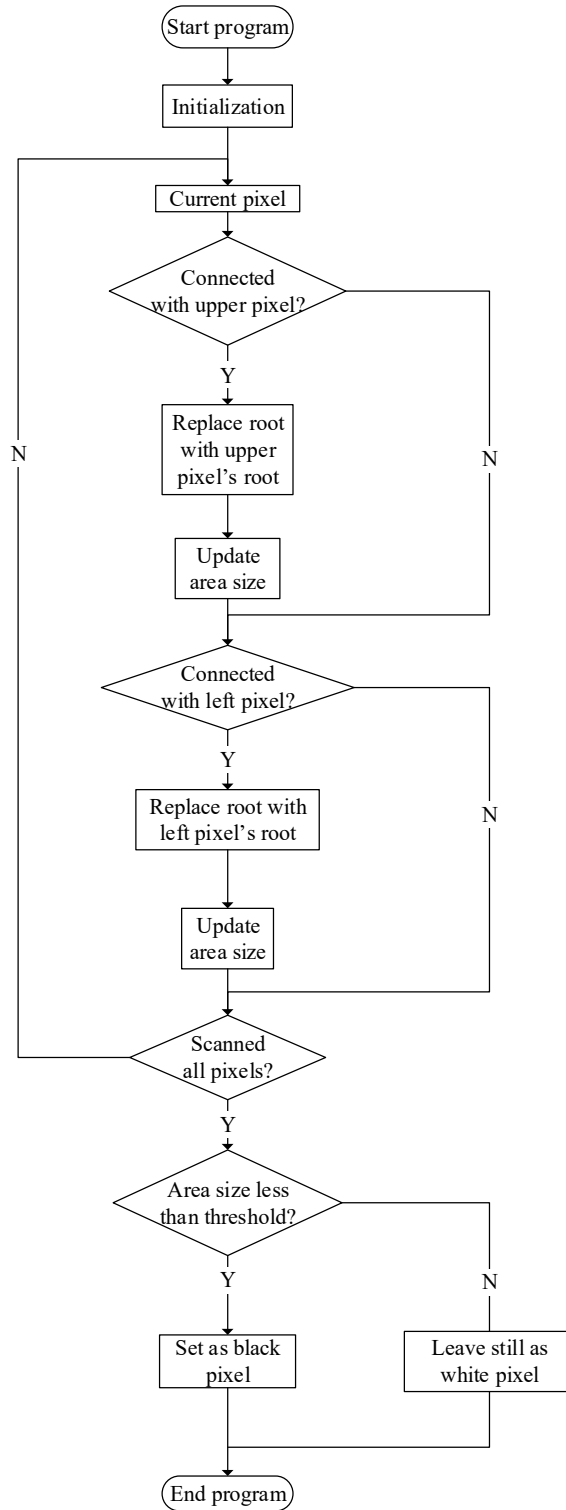e noise. The scanning direction is column by column. In every scanning pixel, two directions need to be checked for connectivity. The upper connection is checked first, followed by the left connection. In the scanning of each pixel, the pixel's root needs to be updated because each connecting area has its own root. Assuming that one pixel's has already been stored and the next scanning pixel is connected to it, the next pixel's root will be replaced by this pixel's root. In this iterative procedure, each pixel's root will be updated. The concept of this algorithm is based on a disjoint-set data structure.

The purpose of path-definition is to eliminate the area smaller than the threshold. The vessel structure has a huge area compared to those of other scattering structures.

This component was coded by my teammate Xinhong Liu, from the City University of Hong Kong.

3.9 MATHEMATICAL MORPHOLOGY DE-NOISING

Mathematical Morphology (MM) is not only used to do reconstruction before training and classification but can also be used to do post-processing, MM de-noising, which is used to eliminate scattering noise in the image. MM de-noising relies mainly on the MM erosion operation, which can remove noise smaller than its Structural Element (SE) (structural element). Because the erosion operation will shrink shapes and areas in the image, opening is selected to recover shapes and areas after the erosion operation. The noise eliminated is smaller than the SE, so the performance of MM de-noising depends on the shape of the SE. In retinal vessel segmentation, linear SEs can be used because vessels can be considered lines, and vessels can be retained by linear SEs.

Figure 3.35 is the overlapped result of wider vessel extraction and thin vessel extraction, and it still has some scattering noise. The wider vessels are derived from the classification component, while the thin vessels are derived from the thin vessel extraction component. Both of these components bring noise into this overlapped result. A linear SE is selected to retain line-shaped vessels.

Figure 3.36 is the result after MM de-noising. Scattering noise has been removed from Figure 3.35. In Figure 3.36, some curving vessels have been removed, which could not be avoided because of the property of the linear Structural Element (SE). Some steps of the de-noising procedure are detailed below.
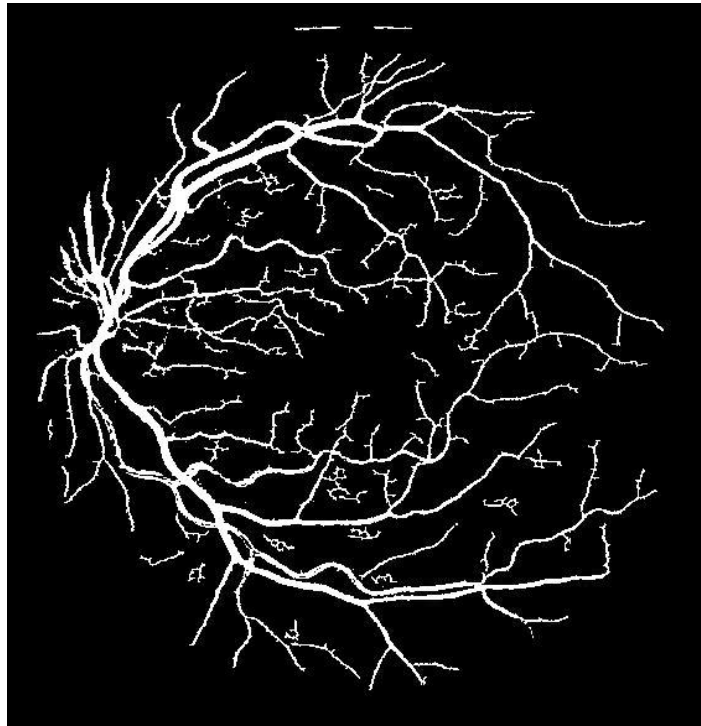

Figure 3.35: Image before Mathematical Morphology (MM) de-noising
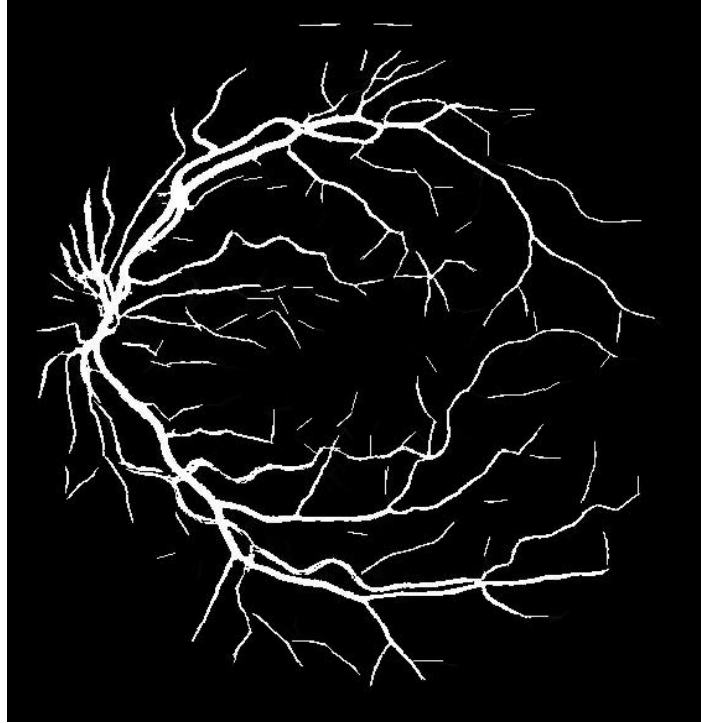
Figure 3.36: Image after Mathematical Morphology (MM) de-noising

Because of the linear Structural Element (SE), directional results need to be shown to explain the principle. Figure 3.37, Figure 3.38, Figure 3.39, and Figure 3.40 are four images with different linear SEs. These four SEs are only different in angles, which are 0, 45, 90, and 135 degrees, respectively. At 0 degrees, horizontal vessels can be retained, along with some nearly horizontal vessels. Vessels do not have a fixed angle in the retina fundus, and every SE at a different angle can retain some piece of the vessel. Finally, different pieces of the vessel can be used to construct the whole vessel structure.
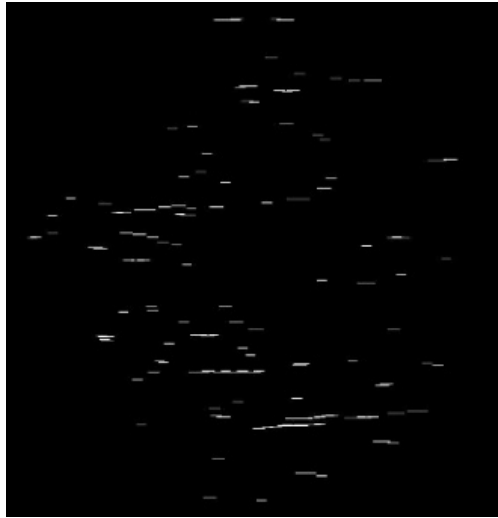
Figure 3.37: Directional de-noising at 0 degrees

Figure 3.38: Directional de-noising at 45 degrees
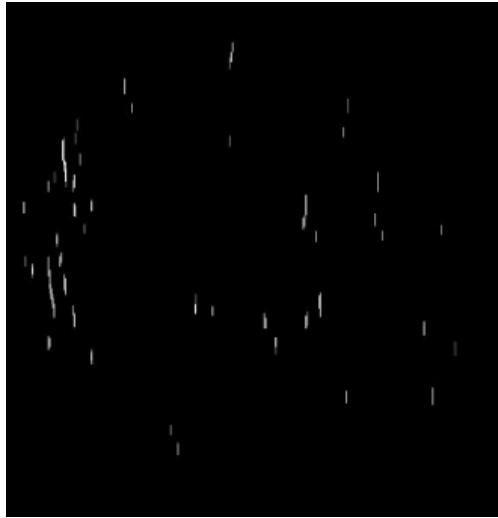
Figure 3.39: Directional de-noising at 90 degrees
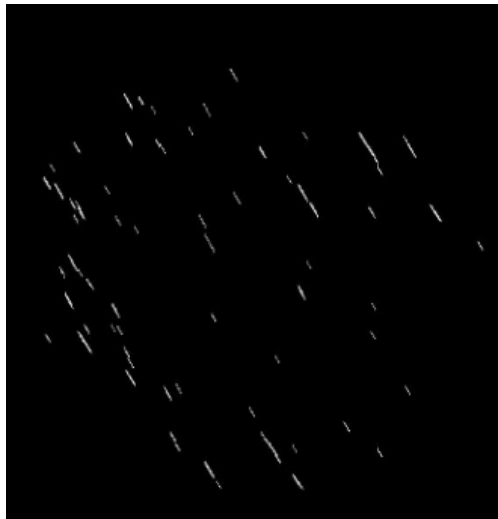
Figure 3.40: Directional de-noising at 135 degrees

The following is the explanation of how Mathematical Morphology (MM) de-noising works in the horizontal direction. Figure 3.41 is the original image matrix, and all pixels are 0 or 1. The linear Structural Element (SE) is three pixels in a line in the horizontal direction. If three or more pixels appear in one line, they will be retained; otherwise, one or two pixels in a

line will be removed. This means that longer lines are easier to retain and smaller lines will be removed. The scattering noise smaller than three pixels will be eliminated.
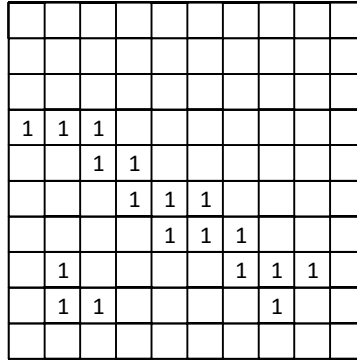
Figure 3.41: Image pixel matrix

Figure 3.42 is the image matrix after Mathematical Morphology (MM) de-noising. Green circles are retained pixels, and the normal 1s are removed pixels. During this processing, linear structures are retained.

Figure 3.42: 0-degree linear Structural Element (SE) de-noising at the pixel level

3.10 SYSTEM FLOWCHART

Figure 3.43 is the flowchart of the whole algorithm; as indicated, the algorithm is mainly divided into two branches. The right branch is for wider vessels. The wider vessel segmentation part can contribute more to the accuracy because the supervised classification can achieve satisfactory performance in wider vessel classification. In the righ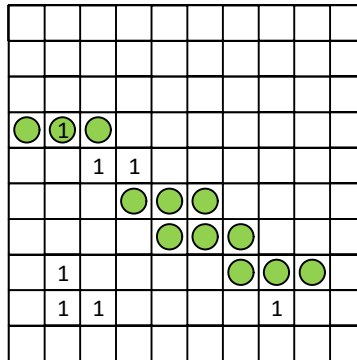t branch in the flowchart, classification theory is used to divide pixels into two classes, the vessel class and the non-vessel class. A Gaussian Mixture Model (GMM) is selected for training and classification. In the training process, the parameters of GMM were obtained. In GMM, two groups of parameters will be generated, one for the vessel class and the other for the non-vessel class. The parameters include the weight, mean value, and standard deviation for every Gaussian model in the two groups. Assuming that every group has five individual Gaussian models, 10 groups of weight, mean value, and standard deviation values will be obtained in the training session. Although the wider vessel classification can contribute more to accuracy, thinner vessels will be missing in huge amounts. The reason for using the two branches is that the classification theory is not suitable for thin vessels because thin vessels usually have a grayscale value mid-way between 0 (8 bits) and 255 (8 bits), making a determination hard for the classifier. The vessel class and non-vessel class probabilities are close. The left branch is for thin vessels. Thinner vessel highlighting is mainly based on deviation changes. As mentioned before, vessels appear with changes in grayscale value. One edge has a positive deviation, and the other edge has a negative edge. The convolution kernel in this step is sensitive to these two edges, and thin vessel values can be highlighted.
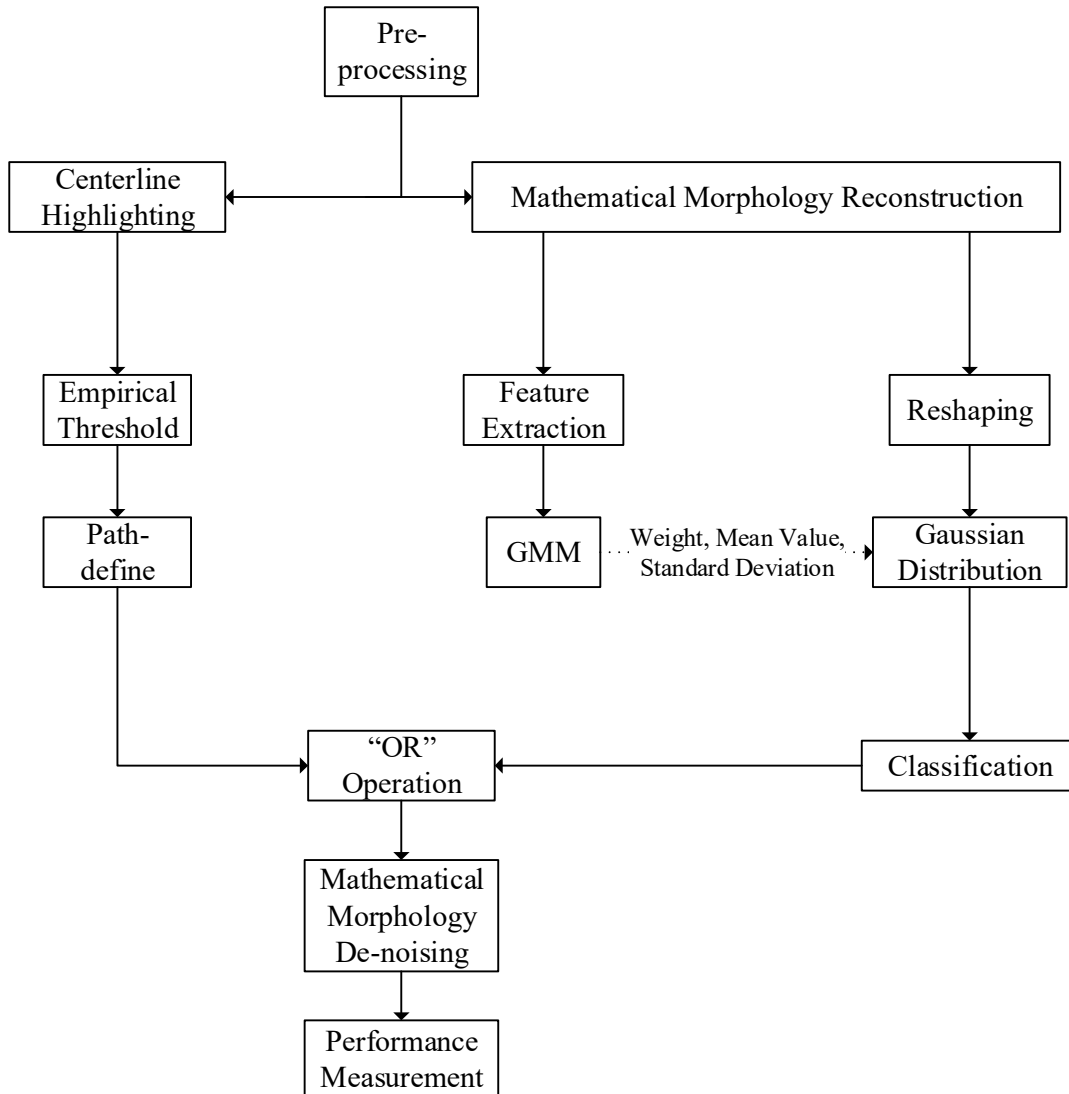
Figure 3.43: System flowchart

The two branches are independent. They share the pre-processing and de-noising components. After obtaining two result images, they need to be overlapped. In the overlapped images, wider vessels and thinner vessels will both be retained with scattered noise. After obtaining overlapped images with scattered noise, Mathematical Morphology (MM) can be implemented to eliminate smaller noise. In addition, in the whole algorithm flowchart, the final

step is the performance measurement, where accuracy and other metrics can be obtained to be compared.
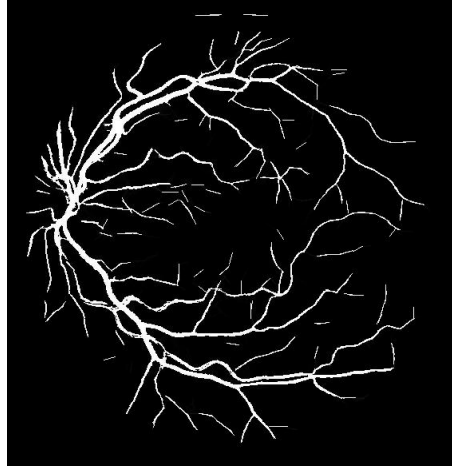
Figure 3.44: Ground truth image after processing

Figure 3.44 is an overlapped sample to show the result. Figure 3.44 is the one obtained after Mathematical Morphology (MM) de-noising. As shown in the image, smaller structures have been removed. Linear structures are retained because of the property of the Mathematical Morphology (MM) Structural Element (SE). Here, a linear structuring element has been used, so the line structures have been retained. The length of the line is around 15 pixels, while lines smaller than 15 pixels are gone. In Figure 3.45, a comparative image is shown. The yellow part is the correctly classified pixels, the red ones are missing pixels, and the green ones are wrongly classified pixels.
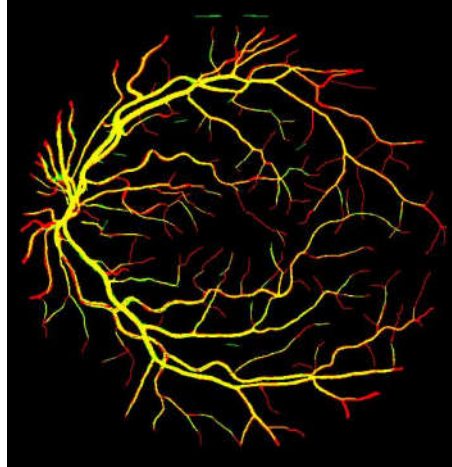
Figure 3.45: Comparison image with a manually labelled one

4 EXPERIMENTS

DRIVE and STARE are the two databases, which conduct the tests. In this thesis, two metrics are used in order to measure the algorithm performance. 'Vessel tracking ability' is directly related to 'single-database testing', while 'robustness' is directly related to 'cross-database testing'. Single-database testing is a classification method whose training and classification parts are in the same database. Cross-database testing's training and classification parts are in two different databases, and these two databases need to exchange.

4.1 VESSEL TRACKING AND ROBUSTNESS PERFORMANCE

According to previously published papers, supervised classification methods achieved higher accuracy than unsupervised methods in general. Because accuracy is the first priority for medical devices, supervised classification method is selected here. On one hand, supervised methods can contribute to higher accuracy in retinal vessel segmentation. On the other hand, supervised methods rely too much on the training set, so it is easily affected by the properties of the training set. It cannot easily extended to other databases. Researchers are optimizing supervised methods for wide extension. In the current thesis, a supervised method is used for better robustness performance with a lower degradation of vessel tracking ability.

Because a supervised method was chosen, two metrics are used here to measure its performance. One metric is vessel tracking ability and the other is robustness. Vessel tracking ability is the performance of the algorithm with respect to how many vessels it can detect, assessed by single-database testing. Single-database testing is done within one database without the interference of other databases, so it can reflect only the vessel tracking ability of

the algorithm. After testing the vessel tracking ability of the algorithm, the robustness of the supervised method needs to be determined to compare it with that of other algorithms. Cross-database testing can reflect robustness performance. As the names of these two types of testing, single-database testing is based on training on one database, and cross-database testing is based on training on another database.

Ricci [19] achieved the best single-database testing accuracy in 2007, but it dropped drastically on cross-database testing, according to Fraz [6]. This means that Ricci's algorithm has an outstanding vessel tracking ability but very poor robustness. Because of this drawback, Fraz proposed another algorithm to solve the robustness issue. In Fraz's work [6], robustness is improved with the loss of some vessel tracking ability. In our thesis, robustness performance was improved further with less degradation of vessel tracking ability.

Fraz's work [6] has more practical meaning because it is much easier to extend it to medical devices. Cross-database testing is more important at the algorithm level. In a clinic, medical devices must always deal with new images and new databases. The images captured by medical devices are always new to the devices, and the properties might be much different from those of the training set. Algorithms in devices must be robust enough to produce an accurate analysis. If the embedded algorithm relies too much on the training set's properties, the analysis will be debated by human beings. This makes medical devices useless. In a broader sense, Fraz's work has great meaning for research and industry.

Our algorithm achieved better performance than Fraz's [6]. As mentioned above, Fraz's work has broad applications in research and industry, so our thesis has even more important applications. In both single-database testing and cross-database testing, our thesis exhibits better performance in accuracy. The reason that ours achieves better tracking ability is that the mathematical morphology reconstruction method is used instead of frequency-based methods because mathematical morphology is a replacement method, not a value-changing method like

frequency-based methods. Frequency-based methods change original values after some kind of transform, and this is detrimental to a binary classifier. The reason for the better robustness is the pre-processing step because it narrows the range of grayscale values, and this is beneficial for the normalization of images with different illumination conditions.

4.2 SINGLE-DATABASE TESTING

In 2007, Ricci et al. proposed a method of segmentation, and this method has thus far achieved the best vessel tracking ability [19]. However, the method has very poor robustness performance, as proved by Fraz [6]. Its robustness is very poor after cross-database testing. Fraz et al. proposed another method to deal with the robustness weakness. In [6], cross-database testing was proposed. This testing is meaningful in measuring algorithm robustness. Robustness is important in medical applications. Because Fraz's paper used a frequency-based method, it lost some performance. Fraz's paper [6] was the first reference paper in the current thesis; it gained much better robustness with a slightly worse vessel tracking ability.

With supervised algorithm, training and testing may be conducted on two different databases, such as testing in the DRIVE database and training in the STARE database or testing in the STARE database and training in the DRIVE database. In addition, training and testing can be done within the same database, DRIVE or STARE. If training and testing are done in the same database, single-database accuracy will be achieved. If training and testing are done in different databases, cross-database accuracy will be obtained. Single-database and cross-database accuracy are used to measure the different aspects of the algorithm.

The method used to calculate accuracy was introduced above. In general, accuracy is achieved pixel by pixel. Every correctly classified pixel contributes to accuracy, while wrongly classified ones are detrimental to accuracy. The DRIVE testing set has 20 images, as does the

STARE testing set. Every image has its own accuracy after calculation of the total number of correctly classified pixels. The database's accuracy is the average value from 20 images.

In this section, single-database accuracy is given. This accuracy is the average value of 20 images. In DRIVE, single-database testing is based on training on DRIVE. In STARE, single-database testing is based on training on STARE.

Table 4.1 gives the accuracy for DRIVE single-database testing. This testing is based on training on the DRIVE database; the highest values are those in Ricci's [19] work in 2007 and Lupascu's [20] work in 2010.

Table 4.1: Single-database testing accuracy on DRIVE

	Year	Accuracy
Ricci [19]	2007	0.9595
Lupascu [20]	2010	0.9597
Marin [21]	2011	0.9452
Fraz [6]	2012	0.948
Lam [13]	2010	0.9472
Staal [5]	2004	0.9441
Proposed method	2015	0.9548

Table 4.2 shows the single-database accuracy on the STARE database, and testing and training are both done on the STARE database. The best results were achieved in Ricci's [19] work in 2007. In general, Ricci's [19] work exhibits the best performance on single-database testing.

In general, this thesis did not achieve the best accuracy in single-database testing, but it achieved the best accuracy in cross-database testing. This will be discussed in the section on cross-database testing.

Table 4.2: Single-database testing accuracy on STARE

	Year	Accuracy
Ricci [19]	2007	0.9584
Marin [21]	2011	0.9526
Fraz [6]	2012	0.9534
Lam [13]	2010	0.9567
Staal [5]	2004	0.9516
Proposed method	2015	0.9552

4.3 CROSS-DATABASE TESTING

In contrast to single-database testing, cross-database testing is based on training on another database, such as training on DRIVE and testing on STARE or training on STARE and testing on DRIVE.

Cross-database testing can test algorithm performance with the interference of other databases. Because of differing illumination conditions, the image quality is different in each database. If the training is done on a database with lighter images and the testing is done on darker images, the false negative rate will be high. In contrast, the false positive rate will be high. Other factors will also decrease the accuracy significantly.

Ricci [19] achieved the highest accuracy in single-database testing, but it dropped drastically in cross-database testing. According to Fraz [6], the worst case of Ricci's algorithm dropped to 92% in cross-database testing. This is unacceptable because testing will be done with a new database containing images from patients. Though Ricci [19] achieved the best accuracy in single-database testing, its robustness performance is poor. For this reason, Fraz et al. proposed another metric to measure algorithm robustness in [6].

Table 4.3 gives a comparison of cross-database testing in three previous papers and the current thesis, which achieved the best cross-database testing accuracy.

Table 4.3: Cross-database testing accuracy

	DRIVE (trained on STARE)	STARE (trained on DRIVE)
Fraz [6]	0.9456	0.9493
Ricci [19]	0.9266	0.9464
Marin [21]	0.9448	0.9528
Proposed method	**0.9496**	**0.9532**

4.4 COMPARISON

In Figure 4.1, a comparison bar diagram has been shown in order to demonstrate the performance improvement. Three metrics are used in this bar diagram: the average of single-database testing, the average of cross-database testing and the average of the above two values.
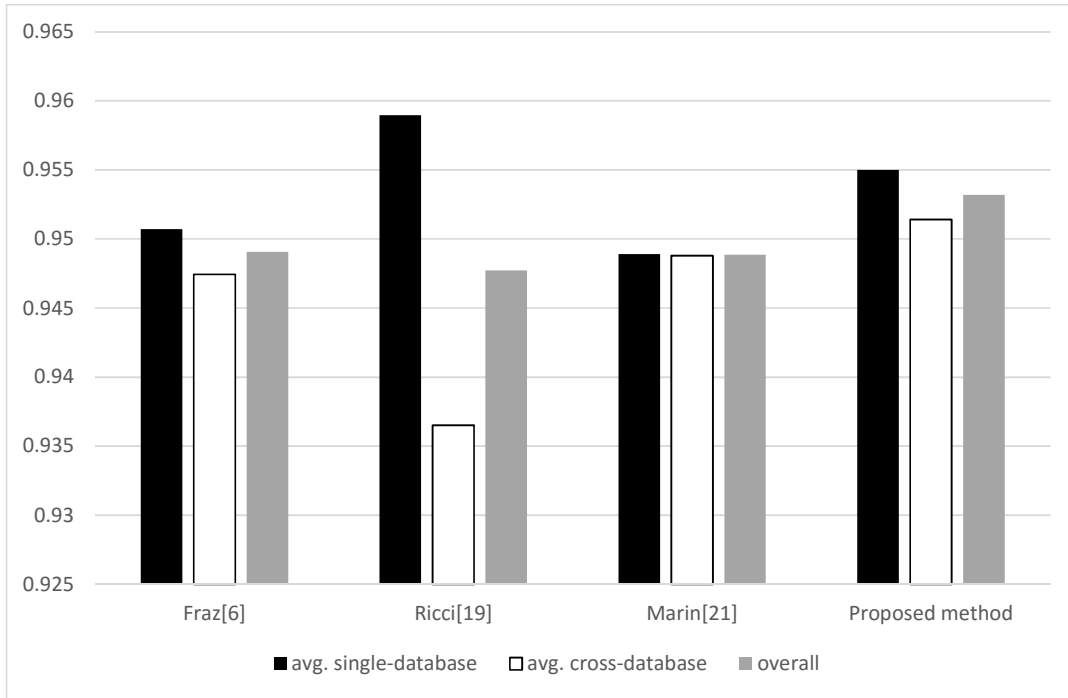
Figure 4.1: Comparison with previous works

DRIVE and STARE have their own singe-database testing values separately. The black bars in Figure 4.1 show the average of these two values of four different works. The first three are previous works, and the right one is from this thesis. The highest average of single-database testing comes from Ricci's work [19]. But this works achieve poor cross-database testing accuracies.

The white bars demonstrate average of cross-database testing accuracies, while the gray bars are the average values of the above two. The proposed work in this thesis has the highest values in the middle bars and the right bars, which means the proposed work has the best performance on cross-database testing and general performance.

4.5 TIME EFFICIENCY

The most time-consuming aspect of the algorithm is the Gaussian Mixture Model (GMM) model training component. In previous papers, researchers have argued that GMM has poor time efficiency. According to the experiments in the current thesis, GMM will decrease time efficiency when the sample size is large. Consequently, the training sample could not be large. In this thesis, 2,000 pixels are selected from each image. Though the number of training samples is small, the classification will not lose much accuracy because of optimization, which is discussed below.

In the training and classification module, only one feature, the grayscale value, is used. Pre-processing and mathematical morphology optimized the grayscale value for the training and classification module. Because it is a binary classifier, clustering and a large distance to the cluster center are two beneficial factors. Pre-processing can help to shrink the grayscale values to a narrower range, and the mathematical morphology reconstruction can help to increase the distance among cluster centers.

Because of these optimizations, Gaussian Mixture Model (GMM) is still good for classification even with a smaller training sample set. As a result, time efficiency is satisfactory even with the use of GMM. In the experiments, the classification time for each image is less than 0.9 seconds.

Table 4.4: Time efficiency comparison

		DRIVE		STARE	
		Training	Classification	Training	Classification
Mendonca [10]	Unsupervised methods	< 2.5 min		< 3 min	
Ricci [19]	Supervised methods	---		---	
Fraz [6]		100s	---	49s	---
Soares [18]		<8h	~100s	<8h	~100s
Marin [21]		---	<90s	---	<90s
Proposed	Supervised		~0.9s		~0.9s

In Table 4.4, time efficiency values have been listed. In this research field, previous works usually lack of time metric, because researchers always focus on the performance of classifier. It is difficult to conclude that the proposed method in this thesis has the best time efficiency due to the lack of time metric information of other works.

5 CONCLUSION AND FUTURE WORK

Retinal diseases, such as Diabetic Retinopathy and Macular Degeneration, affect peoples' life. In severe cases, patients would lose their vision. In order for eye disease analyzing, various algorithms have been proposed for specific diseases. However, these methods will get a lower accuracy without vessel segmentation, since segmentation usually plays as a pre-step for other disease detection algorithms. As a result, segmentation can increase the accuracy of disease detection.

Segmentation is essentially a binary classifier, so the methods can be categorized into either unsupervised or supervised. Because of higher accuracy of supervised methods, this thesis proposes a segmentation method based on supervised classifier. The advantages of the proposed algorithm are high-accuracy and better robustness.

In Chapter 4.4 , the proposed work in this thesis has been proved that it has achieved the best performance on cross-database testing, or robustness performance. The main important reason of the best performance is due to the optimized input for binary classifiers. The images pass pre-processing component and Mathematical Morphology (MM) component before entering the classifier. These two components play as optimizing components in this thesis. The images are originally colorful. Pre-processing component converts colorful images into continuous grayscale images, and MM reconstruction module redistributes grayscale values. This redistribution is beneficial for the following classifier. In training and classification part, mainly Gaussian Mixture Model (GMM) is used to model the grayscale distribution. If grayscale values can be redistributed within several clusters, GMM can easily model the distribution with higher accuracy. In other words, if grayscale values can be corrected either closer to 0 or to 255 (8 bits), the classification would be much easier. Top-hat reconstruction

is very suitable for this task. It can extend the contrast between vessels and background, because it can whiten the vessels and blacken the background.

In this thesis, another new thought has been implemented as well, which is a separate thin-vessel processing algorithm. In traditional methods, vessel segmentation always relies on the classifier. Thin vessels are hard to be extracted by classifiers, because they always in overlapping areas among different clusters. Therefore, it is necessary to develop an efficient way to handle thin vessels. Derivative is used in this thesis, because it is sensitive to rising or falling edges in some direction. In general, thin vessels are more important than thick vessels in pathology, because disease starts with thin vessel growth.

In the current algorithm, the most time-consuming part is pre-processing component. It is difficult to find an efficient data structure to improve the speed due to the nature of the iterative procedure. Hardware implementation is a good option for speed optimization, because some hardware architectures have parallel functions, such as GPU and FPGA. In the future, the proposed algorithm in this thesis could be implemented into hardware platform overall or partly according to the usage.

LIST OF REFERENCES

[1]   [Wikipedia]. Available: https://en.wikipedia.org/wiki/Retina.

[2]   [Retinal International]. Available: http://www.retina-international.org/eye-conditions/symptoms-understanding/loss-of-peripheral-vision/.

[3]   [National Eye Institution]. Available: https://nei.nih.gov/health/diabetic/retinopathy.

[4]   [Eye Quick]. Available: http://www.eyequick.com/.

[5]   J. Staal, M. D. Abramoff, M. Niemeijer and M. A. Viergever, "Ridge-based vessel segmentation in color images of the retina," IEEE Transaction on Medical Imaging, vol. 23, no. 4, pp. 501-509, 2004.

[6]   M. M. Fraz, P. Remagnino, A. Hoppe and B. Uyyanonvara, "An Ensemble Classification-Based Approach Applied to Retinal Blood Vessel Segmentation," IEEE Transactions on Biomedical Engineering, vol. 59, no. 9, pp. 2538-2548, 2012.

[7]   A. Hoover, V. Kouznetsova and M. Goldbaum, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," IEEE Transactions on Medical Imaging, vol. 19, no. 3, pp. 203-210, 2000.

[8]   F. Zana and J.-C. Klein, "Segmentation of vessel-like patterns using mathematical morphology and curvature evaluation," IEEE Transactions on Image Processing, vol. 10, no. 7, pp. 1010-1019, 2001.

[9]   X. Jiang and D. Mojon, "Adaptive local thresholding by verification-based multithreshold probing with application to vessel detection in retinal images," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 25, no. 1, pp. 131-137, 2003.

[10] A. M. Mendonca and A. Campilho, "Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction," IEEE Transaction on Medical Imaging, vol. 25, no. 9, pp. 1200-1213, 2006.

[11] B. S. Lam and H. Yan, "A Novel Vessel Segmentation Algorithm for Pathological Retina Images Based on the Divergence of Vector Fields," IEEE Transactions on Medical Imaging, vol. 27, no. 2, pp. 237-246, 2008.

[12] B. Al-Diri, A. Hunter and D. Steel, "An Active Contour Model for Segmenting and Measuring Retinal Vessels," IEEE Transactions on Medical Imaging, vol. 28, no. 9, pp. 1488-1497, 2009.

[13] B. Y. Lam, Y. Gao and A. W.-C. Liew, "General Retinal Vessel Segmentation Using Regularization-Based Multiconcavity Modeling," IEEE Transactions on Medical Imaging, vol. 29, no. 7, pp. 1369-1381, 2010.

[14] M. S. Miri and A. Mahloojifar, "Retinal Image Analysis Using Curvelet Transform and Multistructure Elements Morphology by Reconstruction," IEEE Transactions on Biomedical Engineering, vol. 58, no. 5, pp. 1183-1192, 2010.

[15] M. M. Fraz, S. A. Barman, P. Remagninoa, A. Hoppea, A. Basitb, B. Uyyanonvarac, A. R. Rudnickad and C. G. Owend, "An approach to localize the retinal blood vessels using bit planes and centerline detection," Computer Methods and Programs in Biomedicine, vol. 108, no. 2, p. 600–616, 2012.

[16] Q. Peng, X. You, L. Zhou and Y.-m. Cheung, "Retinal Blood Vessels Segmentation Using the Radial Projection and Supervised Classification," in International Conference on Pattern Recognition (ICPR), Istanbul, 2010.

[17] [Image Science Institute]. Available: http://www.isi.uu.nl/Research/Publications/publicationview.php?id=866.

[18] J. B. Soares, J. G. Leandro, R. M. Cesar and H. F. Jelinek, "Retinal vessel segmentation using the 2-D Gabor wavelet and supervised classification," IEEE Transactions on Medical Imaging, vol. 25, no. 9, pp. 1214-1222, 2006.

[19] E. Ricci and R. Perfetti, "Retinal Blood Vessel Segmentation Using Line Operators and Support Vector Classification," IEEE Transactions on Medical Imaging, vol. 26, no. 10, pp. 1357-1365, 2007.

[20] C. A. Lupascu, D. Tegolo and E. Trucco, "FABC: Retinal Vessel Segmentation Using AdaBoost," IEEE Transactions on Information Technology in Biomedicine, vol. 14, no. 5, pp. 1267-1274, 2010.

[21] D. Marin, A. Aquino, M. E. Gegundez-Arias and J. M. Bravo, "A New Supervised Method for Blood Vessel Segmentation in Retinal Images by Using Gray-Level and Moment Invariants-Based Features," IEEE Transactions on Medical Imaging, vol. 30, no. 1, pp. 146-158, 2010.

[22] R. M. Haralick, S. R. Sternberg and X. Zhuang, "Image Analysis Using Mathematical Morphology," IEEE Transaction on Pattern Analysis and Machine Intelligence, Vols. PAMI-9, no. 4, pp. 532-550, 1987.

VITA

The master of science degree candidate, Sen An, was awarded a bachelor's degree in engineering by the Harbin Institute of Technology, China. He enrolled at the University of Saskatchewan in September 2013 to pursue his master's degree. His major was computer science, but he recently changed it to computer engineering, and he is focusing on image processing and software-hardware interface design. In his master's degree studies, image processing is the main target of his thesis. At the same time, a hardware-friendly view has been considered in the development of the software algorithm.