

RETINAL BLOOD VESSEL SEGMENTATION:
METHODS AND IMPLEMENTATIONS

A Thesis Submitted to the
College of Graduate and Postdoctoral Studies
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

Zhexin Jiang

©Zhexin Jiang, August 2017. All rights reserved.

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 my 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
3B36 Engineering Building
57 Campus Drive
University of Saskatchewan
Saskatoon, Saskatchewan
Canada
S7N 5A9

ABSTRACT

Since the retinal blood vessel has been acknowledged as an indispensable element in both ophthalmological and cardiovascular disease diagnosis, the accurate segmentation of the retinal vessel tree has become the prerequisite step for automatic or computer-aided diagnosis systems. In the attempt to fulfill the need of the accurate, robust, and fast automated vessel segmentation method, this thesis, therefore, has investigated different works of image segmentation algorithms and techniques, including unsupervised and supervised methods. Further, the thesis has developed and implemented two systems of the accurate retinal vessel segmentation.

The methodologies explained and analyzed in this thesis, have been selected as the most efficient approaches to achieve higher precision, better robustness, and faster execution speed, to meet the strict standard of the modern medical imaging. Based on the intensive investigation and experiments, this thesis has proposed two outstanding implementations of the retinal blood vessel segmentation.

The first implementation focuses on the fast, accurate and robust extraction of the retinal vessels using unsupervised techniques, by applying morphology-based global thresholding to draw the retinal venule structure and centerline detection to extract the capillaries. Besides, this system has been designed to minimize the computing complexity and to process multiple independent procedures in parallel.

The second proposed system has especially focused on robustness and accuracy in regardless of execution time. This method has utilized the full convolutional neural network trained from a pre-trained semantic segmentation model, which is also called the transfer deep learning. This proposed method has simplified the typical retinal vessel segmentation problem from full-size image segmentation to regional vessel element recognition.

Both of the implementations have outperformed their related works and have presented a remarkable scientific value for future computer-aided diagnosis applications. What's more, this thesis is also a research guide which provide readers with the comprehensive knowledge on how to research on the task of retinal vessel segmentation.

ACKNOWLEDGEMENTS

The work presented in this thesis would not have been possible without the assistance and cooperation of many people. I would first like to thank my supervisor Dr. Seokbum Ko of the Department of Electrical and Computer Engineering at the University of Saskatchewan. The door to Prof. Ko office was always open whenever I ran into a trouble spot or had a question about my research or writing. He consistently allowed this paper to be my own work but steered me in the right the direction whenever he thought I needed it.

I would also like to thank my lab members who were involved in the experiments for this research: Hao Zhang, Yi Wang, Juan Yopez and Suganthi Venkatachalam. Without their passionate participation and input, the experiments could not have been successfully conducted. The days in our lab were hard work, sometimes frustrating, but I always had fun.

Last but not least, I must express my very profound gratitude to my family, to my parents, and to my wife for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

Who can number the clouds in wisdom
or who can stay the bottles of heaven
When the dust groweth into hardness
and the clods cleave fast together
(Job 38:37-38)

CONTENTS

Permission to Use	i
Abstract	ii
Acknowledgements	iii
Contents	v
List of Tables	vii
List of Figures	viii
List of Abbreviations	x
1 Introduction	1
2 Motivation: why retinal blood vessel segmentation matters	3
2.1 The retinal blood vessels	3
2.1.1 Indicator of retinal and systemic diseases	4
2.1.2 Automated retinal blood vessel segmentation	5
2.2 Applications for the retinal blood vessel segmentation	6
2.2.1 Computer aided diagnostic system of diabetic retinopathy	6
2.2.2 Biometric Security identification system	7
3 Existing research: standing on the shoulders of giants	8
3.1 Retinal blood vessel databases	8
3.1.1 DRIVE	9
3.1.2 STARE	10
3.1.3 CHASE_DB1	10
3.1.4 HRF	11
3.2 Performance measures	11
3.2.1 Accuracy	12
3.2.2 Sensitivity and Specificity	12
3.2.3 Other factors to consider	13
3.3 Related works and methodology	15
3.3.1 Related works	15
3.3.2 Methodology	18
4 Implementation: proposed methods of retinal blood vessel segmentation	26
4.1 Fast, accurate and robust retinal blood vessel segmentation system	26
4.1.1 Preprocessing	27
4.1.2 Venules structure extraction	27

4.1.3	Capillaries detection	29
4.1.4	Post-processing	30
4.2	Retinal blood vessel segmentation using fully convolutional network with transfer learning	31
4.2.1	Preprocessing	32
4.2.2	Data augmentation	33
4.2.3	Network architecture for training and testing	34
4.2.4	Merging and overlapping	35
4.2.5	Post-processing	35
5	Performance evaluation: which one is the best	46
5.1	Experiment setup	46
5.1.1	Data preparation	46
5.1.2	Execution environment setup	47
5.2	Performance measurement	49
5.3	Performance comparison and analysis	53
6	Conclusion and apocalypse: why deep learning is the ultimate tool for medical image applications	63
6.1	Final review	63
6.2	Medical imaging, deep learning, and transfer learning	64
6.3	Future work	65
	References	66

LIST OF TABLES

5.1	The arrangement of databases for the cross-database training and testing. . .	47
5.2	Performance of the proposed unsupervised implementation on four databases.	50
5.3	Performance of the proposed unsupervised implementation on STARE and HRF databases (only healthy retinal images included).	50
5.4	Linear and parallel execution time per image of the proposed unsupervised implementation on four databases.	51
5.5	Performance of the proposed supervised implementation on four databases (single-database test).	51
5.6	Performance of the supervised implementation on four databases (cross-database test).	52
5.7	Performance comparison with the related state-of-the-art works on DRIVE database.	54
5.8	Performance comparison with the related state-of-the-art works on STARE database.	55
5.9	Performance comparison with the related state-of-the-art works on CHASE_DB1 database.	56
5.10	Execution time comparison with a single image from DRIVE database. . . .	57

LIST OF FIGURES

2.1	The retinal color image from the HRF database [25].	3
2.2	Example of the automated process with input on the left and output on the right (images are from the DRIVE database [43]).	5
3.1	Different greyscale distribution before (left) and after (right) the top-hat transform.	19
3.2	By transforming the fully connected layers into convolutional convolution layers, the fully convolutional network can make dense predictions for per-pixel tasks like semantic segmentation	23
3.3	Result comparison of the two PASCAL VOC [4] images segmentation where the left column shows the segmentation result of the fully convolutional network; the column in the middle is the ground truth image; the right column is the original color image. (Results is generated by the work of [40].)	24
4.1	Functional diagram of the proposed unsupervised retinal vessel segmentation system.	27
4.2	Preprocessing phase: (a) original color image, (b) color image after black ring removal, (c) greyscale image after transform through green channel.	28
4.3	Venules structure extraction phase: (a) image after preprocessing, (b) image after top-hat transform, (c) image after intensity thresholding.	29
4.4	Comparison with different thresholds: (a) intensity thresholding with very a strict threshold, (b) intensity thresholding using proposed threshold, (c) intensity thresholding with a loose threshold.	30
4.5	Capillaries detection phase: (a) image after preprocessing, (b) image after black-and-white inversion, (c) image after centerline highlighting, (d) the final result of capillaries detection after connectivity test.	36
4.6	Post-processing phase: (a) image after venules structure extraction, (b) image after centerline detection, (c) image after overlapping, (d) the final result of retinal vessel segmentation after de-noising.	37
4.7	Functional diagram of the proposed supervised retinal vessel segmentation system.	38
4.8	The retinal images taken with unbalanced illumination (left) and unfocused camera (right).	39
4.9	Preprocessing phase: (a) original color image, (b) image after “black ring” replacement, (c) image after Gaussian smooth (only applied to the outer area), (d) final look of the preprocessed image.	40
4.10	A full-size image will be divided into multiple 50×50 image grids and then be resized into 500×500	41
4.11	The 500×500 retinal color image grids (first row) and their coresponding ground truth images (second row).	42
4.12	The input (first row) and output (second row) image slices of the testing process	43

4.13	The architecture of the fully convolutional AlexNet for retinal vessel segmentation	44
4.14	The comparison of the full-size results: (a) original color image, (b) the ground truth image, (c) the image after merging and overlapping, (d) the final result after post-processing.	45
5.1	The trending chart of accuracy, loss, and learning rate during the training process (generated by DIGITS).	58
5.2	Single-database test result comparison of the image from DRIVE (first row) and STARE (second row) databases: (a), (d) are original color images; (b), (e) are ground truth images; (c), (f) are single-database testing results; . . .	59
5.3	Single-database test result comparison of the image from DRIVE (first row) and STARE (second row) databases where (a), (d) are original color images; (b), (e) are ground truth images; (c), (f) are single-database testing results. .	60
5.4	Cross-database test result comparison of the image from DRIVE (first row) and STARE (second row) databases where (a), (d) are original color images; (b), (e) are ground truth images; (c), (f) are cross-database testing results. .	61
5.5	Comparison between single-database results and cross-database results of the image from DRIVE (first row) and STARE (second row) databases where (a), (d) are ground truth images; (b), (e) are single-database testing results; (c), (f) are cross-database testing results.	62

LIST OF ABBREVIATIONS

DRIVE	Digital Retinal Images for Vessel Extraction
STARE	Structured Analysis of the Retina
CHASE_DB1	Child Heart and Health Study in England, Database 1
HRF	High-Resolution Fundus
PASCAL VOC	Pattern Analysis, Statistical Modelling and Computational Learning, Visual Object Classes

CHAPTER 1

INTRODUCTION

The retinal blood vessel has been acknowledged as an indispensable element in both ophthalmological and cardiovascular disease diagnosis such as glaucoma and diabetic retinopathy. The attributes of retinal vasculature including length, width, tortuosity, branching pattern and angles will contribute to the diagnostic result. However, manual segmentation of retinal blood vessels, although possible, is a time consuming and repetitive work, and it also requires professional skills for even the thinnest vessel could contribute to the differential diagnosis list. In order to assist ophthalmologists with this complex and tedious work, the demand for the fast automated analysis of the retinal vessel images arises. Chapter 2 will further discuss the clinic value of retinal blood vessel and how it will assist the automated diagnosis system. Based on the demand, this thesis presents the author's persistent efforts on fulfilling the need of the accurate, robust, and fast automated vessel segmentation, from existing methods to proposed implementations.

The main body of the thesis about retinal blood vessel segmentation is divided into two parts: methods and implementation. Since the study of retinal blood vessel segmentation was firstly published in 1989 [11], almost 30 years have passed, magnificent improvement has been made to this study with the development of public retinal image databases, advanced image processing algorithms, and supervised machine learning techniques. To understand how those development has effectively contributed to the performance of retinal vessel segmentation, in Chapter 3 the thesis has investigated their abilities to recognize the principle features of the retinal blood vessel. In addition, this Chapter has also introduced the most cited public retinal databases and performance evaluation metrics. All of the investigated resources and methods have been applied to the proposed implementations, which are included in Chapter 4.

More specifically, Chapter 4 has proposed two implementations for different medical scenes. The first implementation has proposed a fast, accurate and robust retinal vessel segmentation system, and this work has presented a novel approach to extract blood vessels from the retinal fundus, by using morphology-based global thresholding to draw the retinal venule structure and centerline detection method for capillaries. Meanwhile, the system has been designed to minimize the computing complexity and has processed multiple independent procedures in parallel, managing to shorten the execution time. This implementation is designed for portable diagnosis systems with healthy retinal images as inputs. The second implementation has utilized the fully convolutional neural networks and transfer learning, training a vessel segmentation network based on a pre-trained semantic segmentation neural network, which is trained by numerous divided retinal image grids to identify vessel elements regionally. Therefore unlike other related works, this proposed method has simplified the typical retinal vessel segmentation problem from full-size image segmentation to regional vessel element recognition. The results will be merged and restored back into the complete full-sized binary images. Eventually, additional unsupervised image post-processing techniques are applied to this proposed method so as to refine the final result. The second implementation, although requiring a long training process and execution time, is highly precise and robust among both healthy and abnormal retinal images.

The performance evaluation has been analyzed in Chapter 5. There the proposed implementations have been compared with all the other outstanding related works through the single-database test and cross-database test. In this way, the thesis has proven the proposed works to be state of the art.

Finally, in Chapter 6, the thesis has discussed why the proposed way of utilizing deep learning and transfer learning techniques in medical imaging should be encouraged, and the research plans to be carried on in the next stage.

The implementation in Section 4.1, “A fast, accurate and robust retinal blood vessel segmentation system”, has been published by Elsevier B.V. on Journal of Biocybernetics and Biomedical Engineering in the year of 2017 [23].

CHAPTER 2

MOTIVATION: WHY RETINAL BLOOD VESSEL SEGMENTATION MATTERS

2.1 The retinal blood vessels

The retinal blood vessels within the human eyeball include arteries and capillaries. The central retinal artery enters the eye through the optic nerve and splits into the superior and inferior branches. These branches then keep branching out more, like the branches of a tree, until they form a very fine network of very thin blood vessels called capillaries (see Figure 2.1).



Figure 2.1: The retinal color image from the HRF database [25].

It is mainly at the capillaries that oxygen and nutrients leave the blood, entering the retina, and then carbon dioxide with waste products leave the retina and pass into the blood to be taken away. Most of the problems caused by conditions affecting retinal blood vessels

do so by either blocking these capillaries or causing them to become leaky. The capillaries join up to form branch veins and these then join at the optic nerve to form the central retinal vein that dives into the optic nerve on its way towards the heart.

Because of the very close relationship between the retina and the artery, the eyeball becomes a unique window to observe the health of the human in a direct and noninvasively way. In other words, the retina is able to present the symptom of disorders right on the retinal blood vessels. This section, therefore, will discuss the clinic value of the retinal blood vessels, and the motivation for the development of the automated retinal blood vessel segmentation algorithms. Besides, the section here will briefly introduce some popular applications of the retinal blood vessel segmentation.

2.1.1 Indicator of retinal and systemic diseases

In the beginning, retinal imaging is primarily used in ophthalmic clinics for the detection of diabetic retinopathy, age-related macular degeneration, glaucoma, retinal neoplasms, and etc. Gradually, researchers have found that the appearance of the retinal blood vessels is an important diagnostic indicator for much systemic pathology, including diabetes mellitus, hypertension, cardiovascular and cerebrovascular disease, and atherosclerosis [15]. More and more research is supporting the fact that the retinal blood vessel may provide a lifetime summary measure of genetic and environmental exposure, and may, therefore act as a valuable risk marker for future systemic diseases [12]. Furthermore, The use of the retinal vessel characteristics provides early identification of people at risk due to diverse disease processes [45].

The artery and capillary of the retinal blood vessel have many observable features, including diameter, color, tortuosity, and transparency. To measure these features, the boundaries of the blood vessels have to be precisely depicted, and then the measurements are to be provided to clinical diagnosis, treatment evaluation, and medical research. In order to be fully engaged in measuring and studying the critical features, some researchers start to develop the method of retinal blood vessel segmentation.

2.1.2 Automated retinal blood vessel segmentation

Manual segmentation of retinal blood vessels, at first, although possible, is so time consuming and repetitive, and it also requires professional skills. Especially for many diseases such as diabetes and hypertension, patients are required to take regular ocular screening in order to detect retinopathy in early stages. Manually segmentation of each patient's retinal vessel for each time is impossible for ophthalmologists. There has to be an automated way of segmentation before retinal vessel segmentation could be practiced into real life. In order to assist ophthalmologists with this complex and tedious work, the demand for the fast automated analysis of the retinal vessel images arises.

In this thesis, the entire automated process of retinal vessel segmentation is unmanned, with no initial point, no pre-setting, no adjustable threshold or any other human interaction. The input is a retinal color image and the output will be a binary image of the vessel and non-vessel pixels. Figure 2.2 gives an example of the automated process.

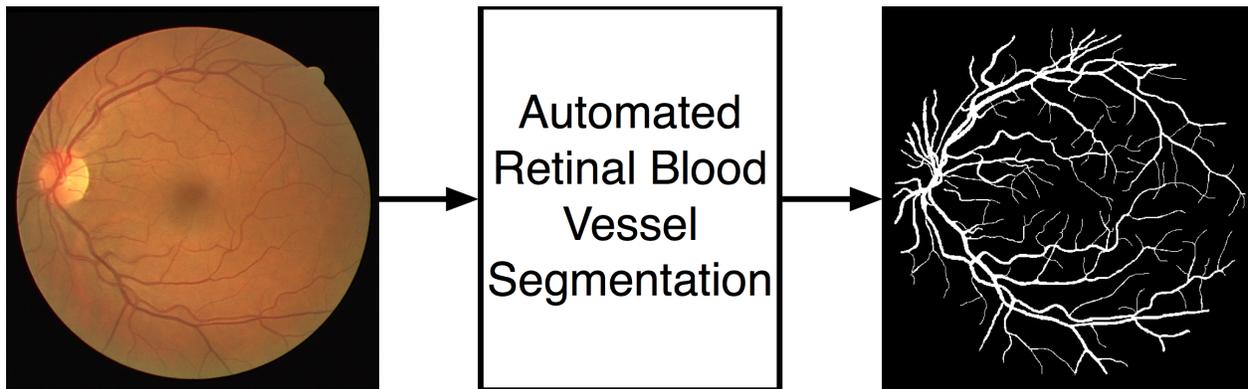


Figure 2.2: Example of the automated process with input on the left and output on the right (images are from the DRIVE database [43]).

Unlike typical foreground and background segmentation in normal image processing, there are three problems making the retinal vessel segmentation task even harder. First, the retinal color image tends to be red everywhere, thus having a lower contrast than normal image segmentation. Second, most retinal color images suffer from unbalanced illumination and make it harder to recognize background. Third, the symptom for retinopathy has unexpected color and shape, thus making it more difficult to separate vessel from noises. Under such circumstance, the research of the retinal blood vessel segmentation has brought much attention and

been developed.

2.2 Applications for the retinal blood vessel segmentation

Most of the applications for the retinal blood vessel segmentation are the clinical application, owing to the extraordinary diagnostic function. The computer aided diagnostic system is one of the most popular application, and the diagnostic system of diabetic retinopathy is the hottest topic because of the great number of people suffering from diabetics. This section will talk about how the automated diagnosis techniques bring benefits. Besides, there are still other applications for retinal vessel segmentation that are also popular, such as biometric identification, and the thesis will briefly introduce this application as well. These applications reflect demand of people, and the automated segmentation of the retinal blood vessels is a prerequisite for these applications.

2.2.1 Computer aided diagnostic system of diabetic retinopathy

Diabetes is a disease that affects about 5.5% of the population worldwide. About 10% of all diabetic patients have diabetic retinopathy [43]. Diabetic retinopathy is nowadays the most frequent cause of vision loss among adults [16]. It is a progressive disease, beginning with mild or nonexistent symptoms. Microaneurysms are often the first clinical sign of diabetic retinopathy, appearing like small deep red dots. As the disease develops, increased tortuosity of the retinal blood vessels may result in the weakening of the outer walls and precede the formation of microaneurysms, which can leak fluid into the retina and cause swelling of the macula [15]. This is why sometimes it will be too late once the symptom of diabetic retinopathy is noticed. Therefore, comprehensive regular eye examinations are essential for people with diabetes, and The World Health Organization has also advised yearly ocular screening to patients [43] because if diabetic retinopathy is detected at an early stage, it can be effectively treated with laser therapy.

However, the limited number of ophthalmologists can not meet the high demand for

the regular eye examination. Take Canada for example, according to the profile from the Canadian Medical Association [1], there are only 1221 ophthalmologists available within the country, and 3.4 ophthalmologists per 100,000 people on average. Hence, doctors and computer scientists are working on an automated solution in order to cover the large diabetic population while reducing the clinical burden on retinal specialists. The automated segmentation of the retinal blood vessels is a prerequisite for the computer aided diagnostic system of diabetic retinopathy. With the segmentation of the vessels, the system is able to interpret the retinal image by study the diameter, tortuosity, and the history pattern of the vessel.

2.2.2 Biometric Security identification system

The retinal blood vessel can also serve as a biometric identifier. Since the biometric identifiers, such as fingerprint, human face, DNA, iris, and retina are the unique and measurable characteristics used to identify and describe individuals, they have been attempted or already applied to automated personal recognition and identification systems. Especially when there requires high-security assurance, biometric identifiers offer safer and more convenient verification than conventional ways of recognition. Since the retinal vessel tree was found to be unique for each individual and can be used for biometric identification, retinal scanning recognition has become one of the most well-known identification technologies.

The foundation of retinal recognition is the recognition of the pattern of the retinal blood vessels. Usually, the retinal image is captured through the fundus camera. Once captured, it will be analyzed by computer, and the pattern of the retinal blood vessel structure will be extracted and stored as the identification. The automated segmentation of the retinal blood vessels is a prerequisite before the computer could study the pattern of the vessel structure. Some researchers have studied how to extract features from the retinal vessel for individual authentication based on vessel ground truth images provided by public databases [13], while others have utilized the existing automated retinal vessel segmentation algorithms to develop their biometric authentication systems [37, 10].

CHAPTER 3

EXISTING RESEARCH: STANDING ON THE SHOULDERS OF GIANTS

3.1 Retinal blood vessel databases

Rome wasn't built in a day. The development of the science and technology today is the accumulation of wisdom since thousands of years ago. When people stand on the shoulders of giants, they will surely be able to see further. It happens the same when scientists discover truth by building on previous discoveries. Thanks to those who have contributed to the area of retinal blood vessel segmentation, especially researchers who have established the publicly available retinal blood vessel databases, scientists will start easier and will be able to compare the performance with each other.

Normally the study of retinal blood vessel segmentation starts from importing public retinal blood vessel database, where there offer researchers with retinal color images and the corresponding information. Some of the databases provide vessel ground truth images, which shows exactly where each vessel pixel is located. With those databases, researchers are able to design their algorithms and compare their performances in the same criterion.

Currently there exist 9 publicly available retinal blood vessel databases, among which CHASE_DB1 [5], DRIVE [43], HRF [25], STARE [7] databases contain both retinal color images and retinal blood vessel ground truth images, while DiaRetDB1 V2.1 [24], Messidor [14], REVIEW [9], ROC [36], and VICAVR[6] databases just provide retinal color images but without labeled images. Although the above databases are all decent in quality and contain both normal and abnormal retinal images, however, as the study of vessel segmentation requires the vessel ground truth as a golden standard, this section, therefore, will just introduce

the following 4 databases that contain vessel ground truth images in section 3.1.1 ~ 3.1.4. Most of the retinal blood vessel segmentation methodologies are evaluated on DRIVE and STARE databases.

3.1.1 DRIVE

The name of the DRIVE (Digital Retinal Images for Vessel Extraction) database has well expressed its purpose – to enable comparative studies on segmentation of blood vessels in retinal images. The DRIVE database consists of 40 color retinal images, obtained from a diabetic retinopathy screening program in The Netherlands. The 40 images are randomly selected from 400 diabetic subjects between 25-90 years of age, and 33 do not show any sign of diabetic retinopathy while 7 show signs of mild early diabetic retinopathy. Each image is JPEG compressed. The images have been acquired using a Canon CR5 non-mydratic camera with a 45-degree field of view. Each image has been captured using 8 bits per color plane at 768 by 584 pixels. The field of the view of each image is circular with a diameter of approximately 540 pixels, and each image has been cropped around the field of the view. For each image, a mask image is provided that delineates the field of the view.

The set of 40 images has been divided into a training and a test set, both containing 20 images. For the training images, a single manual segmentation of the vasculature is available. For the test cases, two manual segmentations are available; one is used as gold standard, the other one can be used to compare computer-generated segmentations with those of an independent human observer. All human observers that manually segmented the vasculature were instructed and trained by an experienced ophthalmologist, and they were asked to mark all pixels for which they were for at least 70% certain that they were the vessel.

In general, the image quality in DRIVE database is desirable and contains just 7 abnormal retinal images with mild disease. It can represent the retinal conditions of the majority of people.

3.1.2 STARE

The STARE database belongs to the STARE (Structured Analysis of the Retina) Project, which has been conceived and initiated at the University of California, San Diego, and it has been funded by the U.S. National Institutes of Health. The STARE database contains 400 retinal color images. The images have been acquired using a Topcon TRV-50 fundus (bottom of the eyeball) camera with a 35-degree field of view. Each image has been captured using 8 bits per color plane at 605 by 700 pixels, and the approximate diameter of the field of view is 650 by 500. 20 of the images can be used for blood vessel segmentation because they are with the vessel ground truth images. The 20 images have been manually segmented by two different experts. The segmented results of the second expert have shown many more of the thinner vessels than the results of the first expert. Usually, the performance is computed with the segmentation of the first expert as the ground truth.

Among those 20 images with ground truth, only 9 images are healthy retinal images, while the other 11 images show signs of 8 kinds of retinal diseases, mild or severe. 3 of the images even suffer from decreased sharpness. Therefore, the STARE database is the most complicated database among all the others, and it always tests the noise-resistance of an algorithm.

3.1.3 CHASE_DB1

The CHASE_DB1 database is a subset of the CHASE (Child Heart and Health Study in England) dataset, which contains retinal images of multiethnic children. CHASE_DB1 is the only subset in CHASE that has vessel ground truth images. The retinal images of both of the eyes of each child were recorded with a handheld NM-200-D fundus camera made by Nidek Co. Ltd., Gamagori, Japan. The images have been captured with a 35-degree field of view, and each image has a resolution of 1280×960 pixels with tagged image file (TIF) format. The images have been captured in subdued lighting using flash and illumination settings of 3, and levels have been adjusted by the operator in the event of perceived over or under exposure. A fixation target has been used, and focused images have been captured centered on the optic disc with full field illumination. To avoid poor quality, the images have been

displayed immediately on a color screen, allowing image capture to be repeated.

The CHASE_DB1 database includes 28 images in total, which are collected from left and right eyes of 14 children. The vessel ground truth images were manually segmented by two human observers. There has no record of the symptom of the 28 retinal images, but they are all in good quality and contrast.

3.1.4 HRF

The HRF (High-Resolution Fundus) image database has been established by a collaborative research group to support comparative studies on automatic segmentation algorithms on retinal fundus images. The images captured have used a Canon CR-1 fundus camera with a field of view of 45 degree and different acquisition setting. The segmentation dataset of HRF has 45 fundus images in total, including 15 images of healthy patients, 15 images of patients with diabetic retinopathy, and another 15 images of glaucomatous patients. The ground truth images for vessel segmentation are available for each image and have been generated by a group of experts working in the field of retinal image analysis and clinicians from the cooperated ophthalmology clinics. The images in HRF database has the resolution of 3504 by 2336, which is the highest among the other databases.

3.2 Performance measures

To describe the performance of the medical research, accuracy alone is not adequate enough to present the features and drawbacks. Besides, the term “accuracy” has an ambiguous meaning owing to the different experimental settings. This section will introduce how to judge the performance of a method is good or not, by applying a few mathematical tools which are commonly used to measure the performance of the retinal blood vessel segmentation algorithms, including accuracy, sensitivity, specificity, and also will discuss some other factors that should be considered in the performance measure. In order to calculate accuracy, sensitivity, and specificity, each result – the segmented binary vessel image, will be compared pixel to pixel with its corresponding ground truth image from the database. If the pixel of the result is labeled the same as the pixel of the ground truth image, it is labeled correctly.

Otherwise, this pixel is labeled wrong.

3.2.1 Accuracy

Accuracy is a common standard which most related works have applied. It reflects the proportion of pixels that are correctly classified as vessel or non-vessel (or background). There used to be two existing methods of calculating accuracy; the first method only takes pixels inside the field of view into account, or in another word, the outside “black ring” area will not be counted, while the second method takes all the pixels into account. However, because not all of the databases provide the field of view ground truth images, for instance, the field of view ground truth images are not available in STARE and CHASE_DB1 databases, the first method of calculating accuracy cannot guarantee fair comparison on these databases. In order to be generic for performance comparison, this work uses the second way to measure accuracy (see Equation 3.1).

$$Accuracy = \frac{\textit{Number of pixels that are correctly classified}}{\textit{Total number of pixels}} \quad (3.1)$$

3.2.2 Sensitivity and Specificity

Sensitivity (also called the true positive rate) and specificity (also called the true negative rate) are statistical measures of the performance of a binary classification test. In this thesis, the retinal blood vessel segmentation result is a pixel-based classification, and all pixels are classified either as vessel or non-vessel. In this case, there will be four cases after the segmentation/classification: i) the true positive where a pixel is identified as vessel in both the ground truth and segmented image; ii) the true negative where a pixel is classified as a non-vessel in the ground truth and the segmented image; iii) the false negative where a pixel is classified as non-vessel in the segmented image but as a vessel pixel in the ground truth image, iv) the false positive where a pixel is marked as vessel in the segmented image but non-vessel in the ground truth image. In this theory, sensitivity and specificity can be calculated through Equation (3.2) and Equation (3.3).

$$Sensitivity = \frac{\textit{True positive}}{\textit{True positive} + \textit{False negative}} \quad (3.2)$$

$$\textit{Specificity} = \frac{\textit{True negative}}{\textit{True negative} + \textit{False positive}} \quad (3.3)$$

Note that for databases with the field of view ground truth images, such as DRIVE and HRF databases, the Sensitivity and Specificity are computed considering only pixels inside the field of view. However, for databases do not include the feature of the field of view, like DRIVE and HRF databases, the Sensitivity and Specificity are calculated based on the total number of pixels. To sum up, sensitivity reflects the ability of the algorithm to detect the vessel pixels, and specificity shows the ability to avoid noise pixels.

3.2.3 Other factors to consider

The last section has presented performance metrics that could be not measured or compared, however, there are still some important factors that are somewhat mentioned in the related works. These factors cannot be directly measured, such as robustness and execution complexity, but are meanings to completely assess the capability of the algorithm. These metrics are common in scientific research and very important for medical applications, especially robustness. Hence, this section will discuss the importance of these metrics, the difficulty in measuring these metrics, and how to effective compare these metrics with other published work.

Robustness

In computer science, robustness is the ability of a computer system to cope with errors during execution and cope with erroneous input. However, robustness in medical imaging is the ability of an algorithm to deal with inputs of poor quality and new medical cases. More specifically, robustness in the retinal blood vessel segmentation refers to the ability to maintain the good performance when the input retinal image is of poor quality (e.g. defocused, bad illumination), or the input image is taken by another camera, or even the image contains symptoms that the algorithm has never dealt before. If an algorithm remains about the same performance under such adverse conditions, the algorithm is robust. On the contrary, if the algorithm performs worse after any of the above conditions takes place, the

algorithm has a bad robustness.

The way of testing the robustness of a method of retinal blood vessel segmentation is limited. Because retinal color images with vessel ground truth are scarce and expensive, for most researchers, taking new retinal images is not possible. The only feasible solution is conducting cross-database test among the given 4 public databases in Section 3.1. There is a dilemma for researchers to prepare the cross-database test because the best way to make the algorithm robust is to adapt the algorithm to as many databases as possible, especially the resources are so limited. However, to prove the cross-database test objective, the tested algorithm shall not know anything about the testing database while being developed. Researchers must not be tempted to use the testing database when developing the algorithm.

No matter unsupervised or supervised method, the cross-database test shall provide at least two datasets, one for training/developing, the other testing. Those two separate datasets shall come from different databases because the different screening environment and groups of patients can ensure the objectivity.

Execution complexity

It is common sense that to be precise and accurate is the most important thing for doctors making the diagnosis. In some cases, diabetes and hypertension patients are required to take regular ocular screening in order to detect retinopathy in early stages. However, patients who are inconvenient to move or live distantly from the city will be less approachable for the location-specific treatment. Hence, there are needs for a handy and portable automated solution. A portable platform, in the current trend, means energy efficient and less computation, thus requesting the execution complexity of the automated retinal vessel segmentation algorithm to be smaller.

It is not likely to analysis the complexity of the published algorithms, but the execution complexity can be reflected if the execution time and the computation platform are provided in the published works. And the fact is, owing to the extraction of vessels from retinal images is truly a harsh task, most published works which perform well have appeared to have a long execution time (presented in Section 3.3). As precision and computing complexity is always a trade-off, some hardware implementations have managed to accelerate the execution speed

in the price of accuracy lost. It is not to say the decreasing in accuracy has diminished the value of the works, and in fact, they are great progress. This thesis wants to encourage more researchers who devote in this subject to provide as much execution information as possible, such as programming language, execution environment, execution platform, execution time, and etc., in order to provide an intuitive way of comparing execution complexity and make more progress.

3.3 Related works and methodology

3.3.1 Related works

The proposed systems from this thesis are enlightened by numerous previous works and based on both classic image processing algorithms and cutting edge machine learning techniques. In the past decades since 1989, hundreds of papers about retinal blood vessel segmentation have been published, and plenty of methods have been proposed. According to Fraz's survey [18] in the year of 2012, the existing retinal segmentation techniques on 2-D retinal images can be summarized into six categories, which are (i) supervised pattern recognition, (ii) mathematical morphology, (iii) matched filtering, (iv) vessel tracking, (v) model based approaches and (vi) parallel/hardware approach. Nowadays, since the hybrid methods, which combine several different categories of methods together, have been more and more popular because of its great performance, it is very hard to be categorized into specific groups. Therefore, in another way of categorization which is more accepted these years, the existing works can be divided into two big categories: supervised and unsupervised methods.

Supervised method

Supervised method use extracted feature vectors, in other words, the labeled training data, to train a classifier, in order to automatically classify the retinal blood vessel and non-vessel pixels from a retinal color image. Such algorithm will learn a set of rules of vessel extraction on the basis of the training dataset. The labeled training dataset is very important in supervised methods because the knowledge of vessel segmentation are directly gained

from the manually segmented images by ophthalmologists. Hence, the performance of the supervised methods in the single-database test usually appears to be better than that of unsupervised ones. The most cited related works with supervised methods are the works of [43, 41, 38, 33, 34]. The work of [43] has used kNN-classifier and sequential forward feature selection to classify the feature vectors. The work of [41] has composed the feature vectors with both the grayscale intensity and two-dimensional Gabor wavelet transform. The work of [33] has even constructed 41-dimensional feature vector, based on the local intensity structure, spatial properties, and geometry at multiple scales. The work of [34] has applied neural network for pixel classification and has computed a 7-dimensional feature vector. The work of [38] has utilized line operators and support vector machines and has achieved the highest accuracy among the above-mentioned works in the single-database test. However, its accuracy has dropped significantly in the cross-database test.

In recent 5 years, with the development of hardware support and the maturation of the neural network techniques, deep learning has become a growing trend in general data analysis and has been termed one of the 10 breakthrough technologies in MIT Technology Review in 2013 [3]. Some works have utilized the deep learning techniques in the retinal blood vessel segmentation and have achieved surprisingly good results [44, 32, 20, 31]. The work of [44] has combined two superior classifiers – using the convolutional neural network to perform as a trainable hierarchical feature extractor and the random forest to work as a trainable classifier. The work of [32] has also applied the convolutional neural network, trained by a large amount of augmented and preprocessed images with global contrast normalization, zero-phase whitening, geometric transformations and gamma corrections. The work of [20] has remolded the task of segmentation as a problem of cross-modality data transformation from retinal image to vessel map, using a wide and deep neural network with strong induction ability.

Generally speaking, the supervised methods appeared to have a better result in comparison with the unsupervised methods, but the supervised methods shall not be compared with the unsupervised methods directly until the cross-database tests have been conducted. More specifically, the classifier shall be trained in one or more databases but tested in another different database, which is also called the robustness test (introduced in Section 3.2.3).

Some supervised methods have achieved a decent accuracy in the single-database test, but perform badly during the cross-database test, such as the work of [38], while some works do not perform the cross-database test, such as the works of [41, 32, 20].

Unsupervised method

The unsupervised classification methods intend to find inherent patterns of retinal vessels directly from the retinal color images and decide whether the pixels are part of the vessel or non-vessel. Unlike the supervised methods, unsupervised methods do not require training data and the training process, and usually with, not necessarily, a higher robustness and faster execution speed. As we mentioned earlier, the unsupervised methods can be further categorized into four groups – mathematical morphology, matched filtering, vessel tracking, and model-based approaches.

Mathematical morphology containing a set of image processing techniques is one of the most famous approaches for image segmentation. It extracts image components that are useful while smoothing the rest area. The morphological operation has the advantage of speed and noise resistance in identifying specific shapes such as features, boundaries, skeletons and convex hulls, by applying structuring elements to grayscale or binary images [35, 19, 17, 8].

Matched filtering techniques usually convolve a 2-D kernel (or a structural element) for blood vessel cross-profile identification (typically a Gaussian or Gaussian-derivative profile). The kernel is rotated into many different directions to model a feature in the image at some unknown position and orientation, and the matched filter response indicates the presence of this feature. Such techniques are very effective to detect vessel centerlines [35, 42, 17].

In most cases, vessel tracking algorithms are more effectively used in conjunction with matched filters of the morphological operators, such as the works of [42] and [46]. Tracking a vessel means to follow the vessel centerline guided by local information and try to find the path which best matches a vessel profile model, through which not only the centerline but also the widths of each individual vessel will be accurately extracted.

Model-based approaches such as vessel profile models in the work of [29], extracting retinal vessels by using explicit vessel models, are designed to handle both normal and pathological retinas with bright and dark lesions simultaneously. Some other methods using deformable

models such as parametric models and geometric models are not as effective as the former one.

3.3.2 Methodology

This section will introduce the methodologies that applied to the proposed two implementations. As the thesis in the next chapter will propose two retinal blood vessel segmentation systems, in order to make it smoother to read, the principles of each methodology applied will be introduced here as the preparation.

The first proposed system, belonging to the unsupervised method, has applied mathematical morphology and matched filtering. During the processes, the retinal images shall firstly be transformed into greyscale and then go through the top-hat transform, intensity thresholding and centerline highlighting.

The second proposed system is a supervised method, which has applied the deep learning techniques and the training of the fully convolutional neural network based on the pre-trained neural network, which is also the application of transfer learning.

The above-mentioned methodologies will be introduced here in Section 3.3.2. The detailed system design of the proposed two methods will be presented in Section 4.1 and Section 4.2.

Morphology processing

Two operations belonging to the mathematical morphology theory will be applied to the proposed system, which is top-hat transform and morphology erosion. The top-hat method is utilized to redistribute the greyscale intensity from a preprocessed greyscale retinal image, in order to generate a characteristic feature for vessel/non-vessel classification, while the erosion operation helps to de-noise the image and smooth the vessel edges in the final post-processing stage.

The principle of morphology processing is to simplify the image data through retaining their essential characteristics from a shape and removing other extraneous elements. The basic operations in mathematical morphology processing are erosion, dilation, opening, and closing, which deduce the top-hat transform – one of the most efficient way to do feature extraction, background equalization, and image enhancement. The mathematical definitions

of a white top-hat transform can be formulated as Equation (3.4),

$$T_W(f) = f - f \circ b \quad (3.4)$$

where f is a greyscale image; b is a greyscale structuring element; the symbol \circ denotes the opening operation; T_W denotes the white top-hat transform, where the object elements are highlighted brighter than their surroundings.

Since the top-hat transform is sensitive to ridges and peaks (the sharp changes in greyscale intensity values), when provided with a preprocessed greyscale retinal vessel image, it will be capable of highlighting the vessel structure while smoothing the rest non-vessel background in different illumination conditions, by considering the edges of vessels as peaks and the flat fundus as the background. However, not just vessels but also noises (such as non-vessel tissues) are likely to be introduced into the results during this process. To solve this, the proposed solution will be discussed in Section 3.3.2.

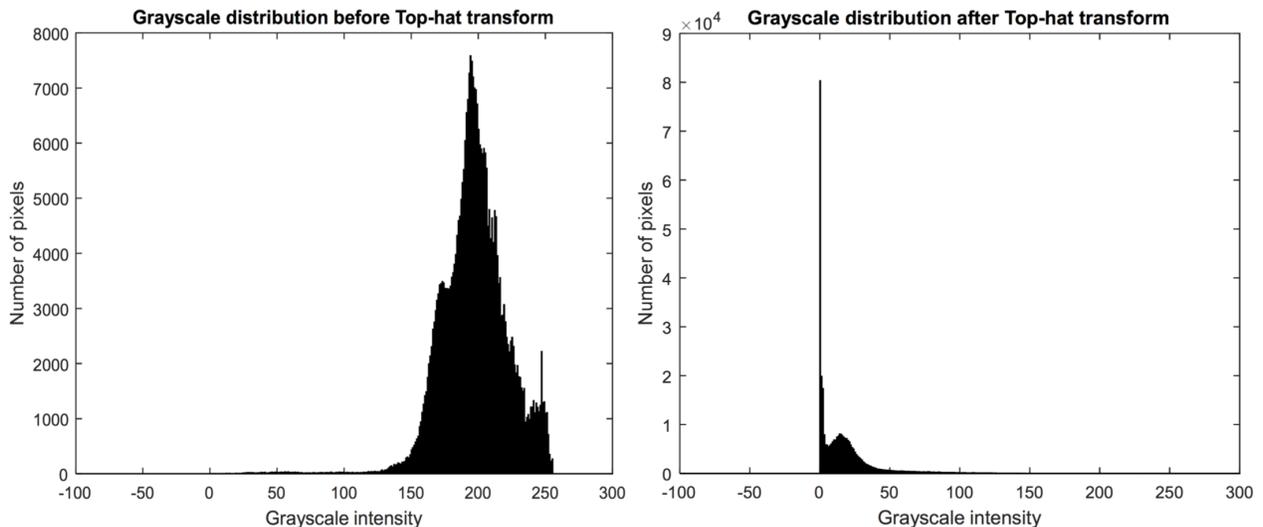


Figure 3.1: Different grayscale distribution before (left) and after (right) the top-hat transform.

Once top-hat transform is completed, the intensity is reconstructed into another distribution. A demonstration of intensity changes before and after the transform is given in Figure 3.1, where about 87% pixels have the intensity of 0 after redistribution. Meanwhile, a small gap could be visually speculated in the right histogram between intensity value 0 and 25, implying there exists a possible threshold to distinguish the vessel/non-vessel pixels. According

to this, the greyscale intensity is selected as the characteristic feature for vessel classification, assuming that most non-vessel pixels have intensity value close to 0.

Intensity thresholding

Image thresholding enjoys a central position in applications of image segmentation, because of its intuitive properties, the simplicity of implementation and computational speed. The proposed system will utilize global intensity thresholding after the top-hat reconstruction of the greyscale image, through which a binary image will be generated according to the Equations (3.5) and (3.6),

$$g(x, y) = \begin{cases} 1, & \text{if } f(x, y) > T \\ 0, & \text{if } f(x, y) \leq T \end{cases} \quad (3.5)$$

where $f(x, y)$ corresponds to the input image, $g(x, y)$ represents the binary image after thresholding, and T is the threshold. Any point (x, y) in the image at which $f(x, y) > T$ is recognized as an object point, or vessel element; otherwise, the point is regarded as a background point, or non-vessel element. The threshold value T in this system is an empirical threshold, rather than using approaches like Otsu’s method to generate one. The reason behind this mainly owing to small vessels are very thin structures, and they usually present low local contrast and hard to be distinguished with their surrounding noises (or non-vessel tissues). In this case, the vessel segmentation process is divided into two phases: the venules structure extraction and the capillary detection. For the venules segmentation, there will not require an optimal threshold to classify the distribution, but a more “strict” threshold (larger than optimal) to eliminate the noises while sacrificing some certain thin vessels, in order to achieve a higher positive predictive value ¹ As a complement, another different method will be specifically utilized for capillary detection.

Centerline highlighting

Because both venules and capillaries are crucial for ophthalmological diagnosis, and the venules segmentation process intends to have ignored most of the thin vessels, in this case,

¹Positive predictive value: the probability that subjects with a positive screening test truly have the disease. In this thesis, it reflects the ability of the system to identify pixels as part of the vessels which truly belong to the vessel class.

capillaries need to be retrieved through centerline detection. Because of the fineness of capillaries, the centerline of a thin vessel can be used to approximate itself, thus completing the function of a vessel extraction system.

Centerline highlighting is the preliminary but the most important step for capillaries detection. It intends to highlight the candidate pixels which could possibly be the ridges of the vessels, by applying a first-order derivative filter orthogonally to the main orientation of the vessel. In this case, derivative values with opposite signs will be obtained on both hillsides of the vessel. In other words, there will be positive values on one side of the vessel cross section while negative values on the other, by which a sign matrix will be obtained. Since the opposite signs indicate two hillsides of a vessel, the presence of centerline thus can be reflected by just scanning for the sign formats. After the pixels of a centerline were located, the intensity value of them will be replaced by the summation of the absolute values of the two horizontal neighbor pixels, in order to highlight the centerline. A particular kernel is proposed here which is the difference of offset Gaussians filters. The difference of offset Gaussians filter is famous for its peak sensitivity and noise resistance, shown in Equation (3.7).

$$\begin{bmatrix} -1 & -2 & 0 & 2 & 1 \\ -2 & -4 & 0 & 4 & 2 \\ -1 & -2 & 0 & 2 & 1 \end{bmatrix} \quad (3.7)$$

However, because the above process only works for the vertical direction, the image itself or the kernel must be rotated in order to be performed in all directions, and this will be discussed in the next section. The idea of centerline highlighting is from the work of [35] and [17].

Fully convolutional neural network

The fully convolutional neural network is one of the most famous application of deep learning, which is an improvement of artificial neural networks, consisting of more layers that permit higher levels of feature abstraction and improved predictions from data [30]. Especially, the convolutional neural network has proven to be a powerful tool for various of computer vision tasks such as image classification and segmentation. Recently, medical image analysis groups

across the world are quickly entering this field and applying convolutional neural networks and other deep learning methodologies to a wide variety of applications, and extraordinary results are emerging continuously [21].

The convolutional neural network, inspired by the organization of the animal visual cortex, is a feed-forward artificial neural network where the connectivity pattern is in between its neurons. Individual cortical neurons respond to stimuli in a restricted region of space which also known as the receptive field. The receptive fields of different neurons partially overlap such that they tile the visual field. The response of an individual neuron to stimuli within its receptive field can be approximated mathematically by a convolution operation [2].

The fully convolutional network, proposed by the computer vision group of the University of California, Berkeley [40], is derived from the convolutional neural network, which, in theory, is normally comprised of one or more convolutional layers (often with a subsampling step) and then followed by one or more fully connected layers as in a standard multilayer neural network. The major feature that makes a fully convolutional network different from the convolutional neural network is the transformation of all fully connected layers into convolution layers (see Figure 3.2), through which, a fully convolutional network is able to operate on an input of any size, and produces an output of corresponding spatial dimensions. In this case, some classification networks, such as the AlexNet [28], can be used for the end-to-end, pixels-to-pixels for semantic segmentation, instead of outputting the classification prediction scores.

Normally each layer of data in a typical convolutional neural network is a three-dimensional array of size $h \times w \times d$, where h and w are spatial dimensions, and d is the feature or channel dimension. A normal convolutional neural network is built on translation invariance. A fully convolutional neural network naturally operates on an input of any size, and produces an output of corresponding (possibly resampled) spatial dimensions.

A real-valued loss function composed with an fully convolutional neural network defines a task. If the loss function is a sum over the spatial dimensions of the final layer (see Equation 3.8, where x_{ij} is the data vector at location (i, j) in a particular layer),

$$l(x; \theta) = \sum_{ij} l'(x_{ij}; \theta) \tag{3.8}$$

its gradient will be a sum over the gradients of each of its spatial components. Thus stochastic

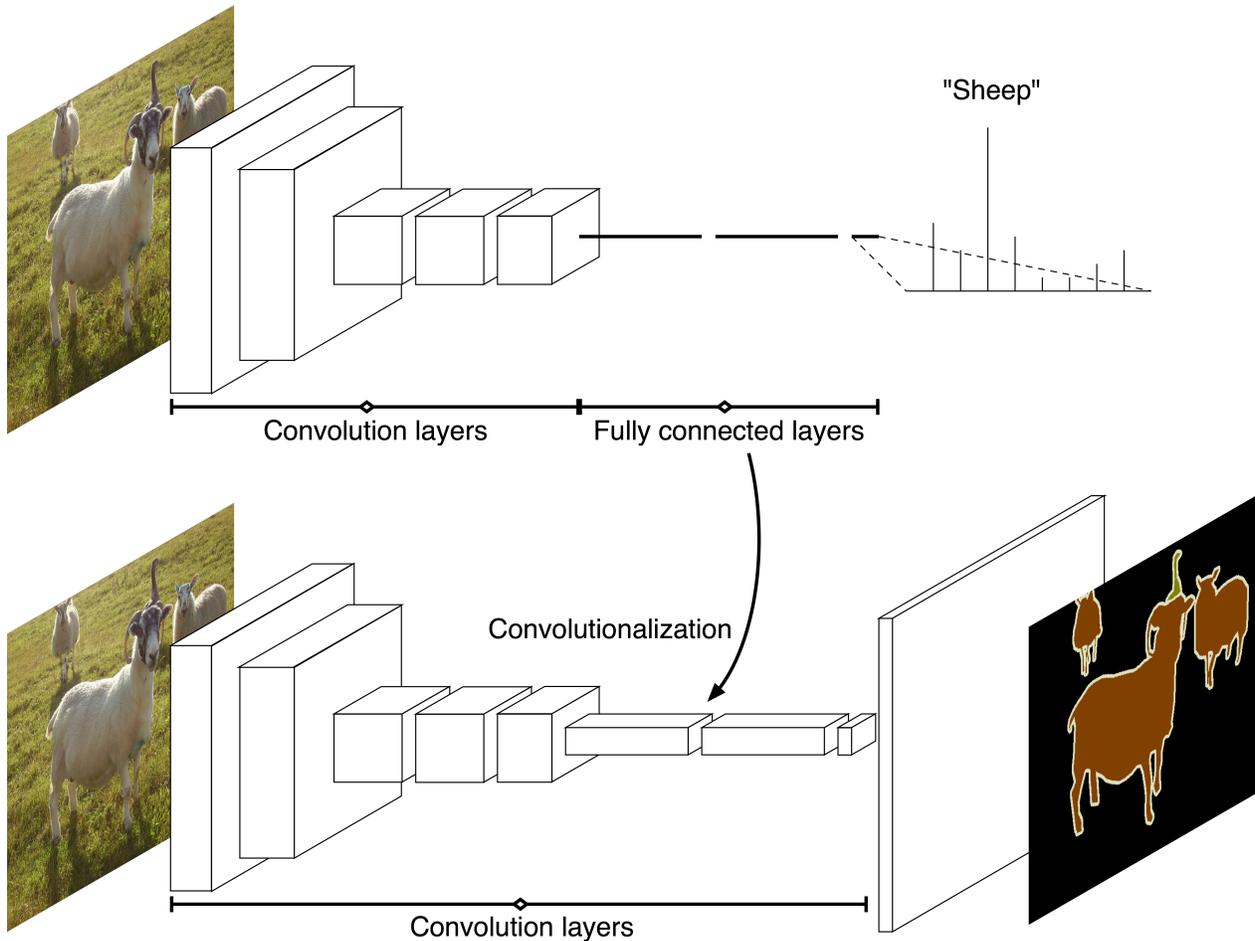


Figure 3.2: By transforming the fully connected layers into convolutional convolution layers, the fully convolutional network can make dense predictions for per-pixel tasks like semantic segmentation

gradient descent on l computed on whole images will be the same as the stochastic gradient descent on l' , taking all of the final layer receptive fields as a mini batch. When these receptive fields overlap significantly, both feedforward computation and back-propagation are much more efficient when computed layer-by-layer over an entire image instead of independently patch-by-patch.

The work of [40] has defined and detailed the space of fully convolutional networks, and have adapted contemporary classification networks, such as AlexNet, the VGG net, and GoogLeNet, into fully convolutional networks and transfer their learned representations by fine-tuning to the segmentation task. The documentation of [2] has defined a skip architecture that combines semantic information from a deep, coarse layer with appearance information

from a shallow, fine layer to produce accurate and detailed segmentations. With this fully convolutional network architecture, the fully convolutional version of AlexNet proposed by the work of [40] has proven to be state-of-the-art in semantic segmentation on PASCAL VOC dataset (Pattern Analysis, Statistical Modelling and Computational Learning, Visual Object Classes) [4]. Figure 3.3 shows two sample segmentation results comparison, where presenting the ability of the fully convolutional network to separate the closely interacting objects and to refine the edges of the structures.

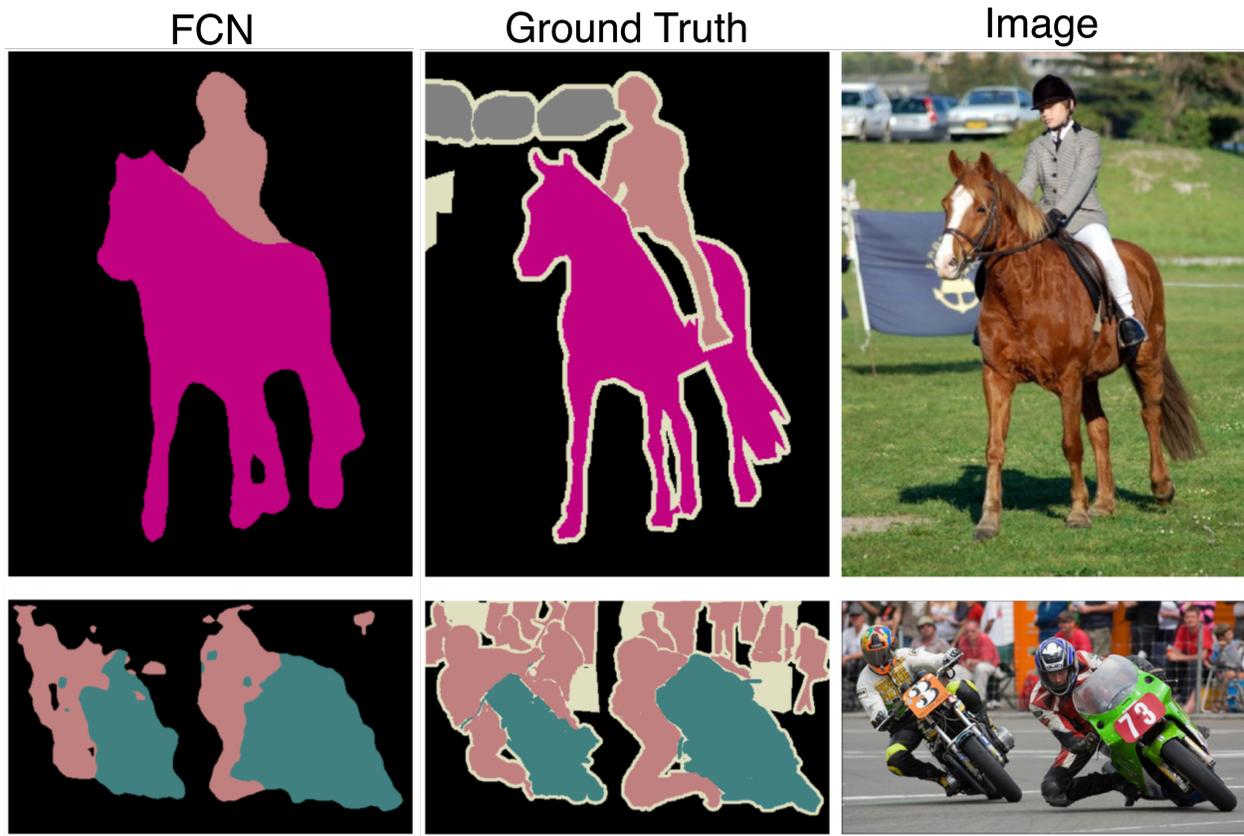


Figure 3.3: Result comparison of the two PASCAL VOC [4] images segmentation where the left column shows the segmentation result of the fully convolutional network; the column in the middle is the ground truth image; the right column is the original color image. (Results is generated by the work of [40].)

The fully convolutional version of AlexNet is the one to be applied in this thesis for vessel segmentation, and the model having generated the results in Figure 3.3 will also be utilized as a pre-trained model so as to continue to train the new network that works for the retinal blood vessel segmentation. This innovative use of the neural network is also called transfer

learning and is especially beneficial for medical imaging.

Transfer learning

The convolutional neural network is remarkably helpful in medical imaging. However, training a full convolutional neural network from scratch is a challenge. First, convolutional neural networks require large numbers of labeled training data, which is somehow very difficult to realize in the medical domain where the expert annotation is expensive and the diseases are rare. Second, training a deep convolutional neural network requires large computational and memory resources, without which the training process would be extremely time-consuming. For a task such like the retinal blood vessel segmentation, the total number of the retinal color images from all the four databases is 133, which is far away from the training requirement of the fully convolutional neural network.

Transfer learning is the perfect solution when there have insufficient publicly available datasets for training the network from scratch. With transfer learning, the convolutional neural network models pre-trained from natural image dataset, such as ImageNet [39], can be used for the new medical task at hand. The convolutional neural network models pre-trained from natural image dataset or from a different medical domain are used for a new medical task at hand. In one scheme, a pre-trained convolutional neural network is applied to an input image and then the outputs are extracted from layers of the network. The extracted outputs are considered features and are used to train a separate pattern classifier, thus solving the training data insufficiency [21].

CHAPTER 4

IMPLEMENTATION: PROPOSED METHODS OF RETINAL BLOOD VESSEL SEGMENTATION

4.1 Fast, accurate and robust retinal blood vessel segmentation system

The first proposed implementation ¹ of the retinal vessel segmentation system, entitled by fast, accurate and robust, is a novel approach that combines the strength of matched filter and mathematical morphological while avoiding their weaknesses. The proposed segmentation system is schematically described by the functional block diagram as shown in Figure 4.1, where we identify the four main processing phases. In the first preprocessing phase, the goal is to transform the retinal color image into the greyscale image. The greyscale image then becomes the input for both venules structure segmentation and capillaries detection phases. These two phases are designed in parallel, which focuses on the different scale to segment venules and capillaries separately. The venules structure segmentation phase has applied morphology-based global thresholding to roughly draw the retinal venule structure, and at the same time, the capillaries detection phase utilizes matched filter method to detect the centerline of the thin vessels, which is just right the capillaries. Since these two processes are independent, therefore once supported with multithread technology, the execution time will be shortened. The last phase includes overlapping and de-noising the results from the previous phases in order to create an image of a clear vascular structure with abundant thin vessels.

¹“A fast, accurate and robust retinal blood vessel segmentation system”, 2017 published by Elsevier B.V. on Journal of Biocybernetics and Biomedical Engineering

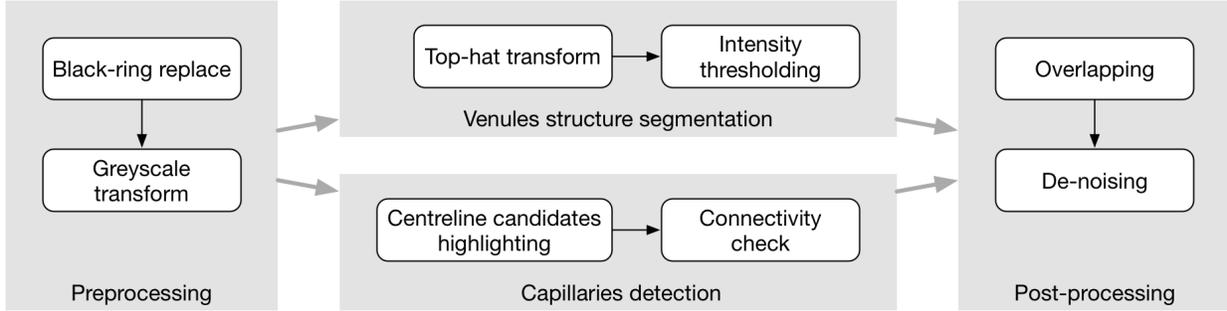


Figure 4.1: Functional diagram of the proposed unsupervised retinal vessel segmentation system.

4.1.1 Preprocessing

To provide a greyscale output with greater contrast, the proposed preprocessing will firstly remove the “black ring” and replace it with the color of the average value of the fundus background (see Figure 4.2(b)). In order to get the average level of the background, three random areas (each area is 50×50) inside the field of the view will be selected. As the color image is indeed a 3-dimension matrix, each selected area is actually a $50 \times 50 \times 3$ matrix. These matrixes are used to estimate the average value of the color within the field of the view. This procedure creates the uniformity of the background and it is quite important for the later top-hat transform because a more balanced background will improve the quality of segmenting the object elements.

After replacing the black ring area with the color of the average value, the optimized color image will be transformed into greyscale through green channel, which has been applied in several works [43, 22], as it naturally presents a higher contrast between vessels and fundus background. Finally, the greyscale image will be black-white inverted (see Figure 4.2(c)), just to be in line with the output standard – vessel in white and background in black.

4.1.2 Venules structure extraction

The object for venules structure extraction is to extract large vessels as much as possible, but leave the thin ones to capillaries detection. Based on the work of preprocessing, the first step in this stage is to directly perform top-hat transform (see Figure 4.3(b)). However, many

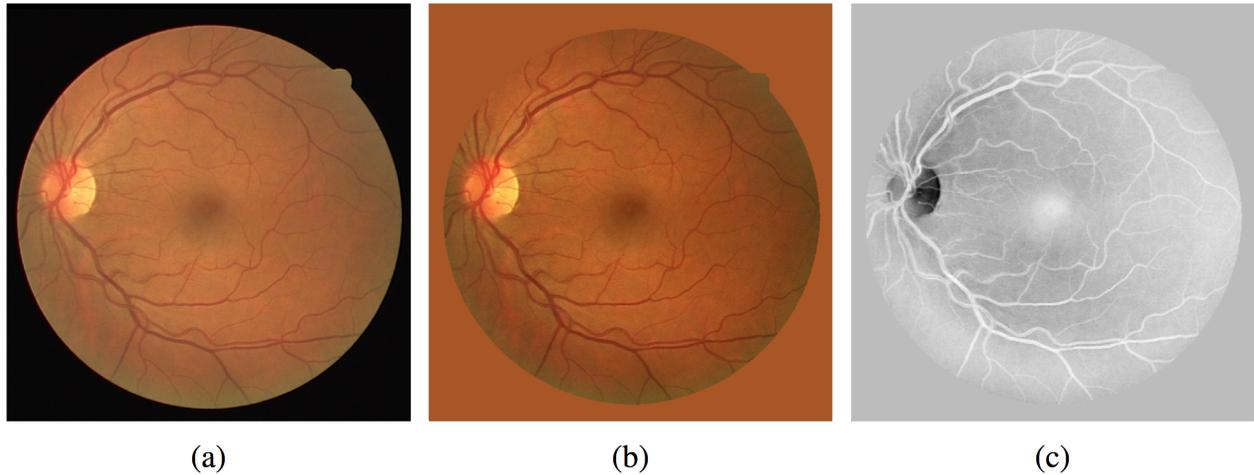


Figure 4.2: Preprocessing phase: (a) original color image, (b) color image after black ring removal, (c) greyscale image after transform through green channel.

other retinal tissues such as the optic disc, macula, fovea, and posterior pole will become the noises as they have a relatively close contrast to vessels. To eliminate these noises, the global intensity thresholding is used here. It classifies vessel and non-vessel elements with minimum computing complexity. Figure 4.3(c) shows a clear vessel skeleton after thresholding.

To find a proper threshold, there are various approaches such as adaptive thresholding and Otsu’s method, but here the system proposes a fixed empirical threshold. Generally speaking, the adaptive thresholding method usually generates a different threshold for each individual image to best fit its own intensity distribution, but such methods simply increase computing complexity because the system has to learn the strategy to analyze the image and to make decisions. For approaches like Otsu’s method, aiming to find the best barrier to differentiate the intersected intensity classes, is not ideal for segmenting retinal blood vessels. Because the non-vessel tissues become very similar after being filtered by green channel during the preprocessing phase. In this case, the Otsu’s method in retinal vessel segmentation will be more likely to preserve those unwanted non-vessel tissues and preform similarly to a loose threshold (see Figure 4.4(c)). Therefore, a more strict threshold than the Otsu’s optimal one is proposed for this system, to reduce most noises while sacrificing thin vessels. Figure 4.3 shows the entire process of the current phase and Figure 4.4 provides the comparison of results using three different thresholds.

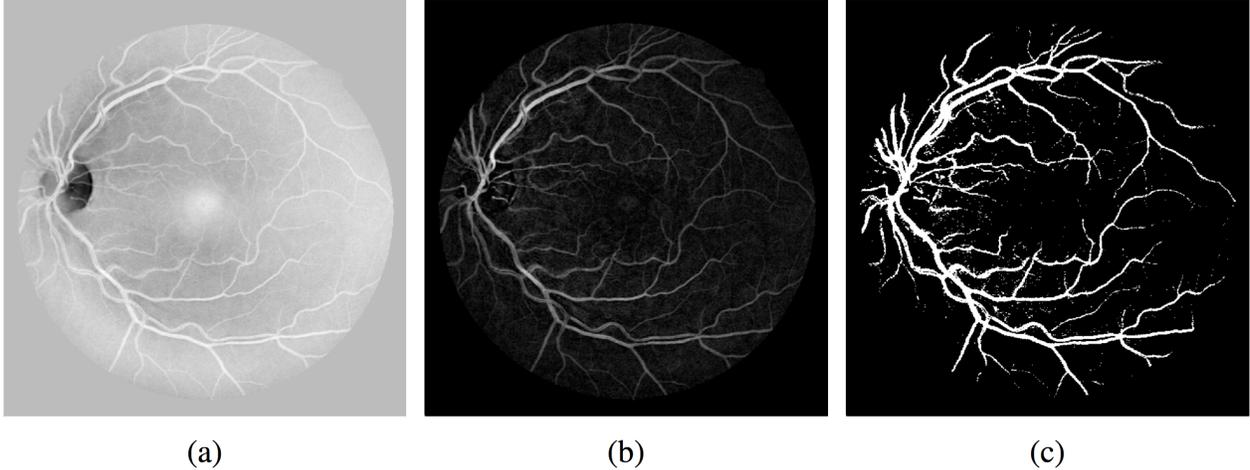


Figure 4.3: Venules structure extraction phase: (a) image after preprocessing, (b) image after top-hat transform, (c) image after intensity thresholding.

4.1.3 Capillaries detection

While extracting venules structure, capillaries detection will also be running simultaneously with the same input image. But the input image needs to be black-and-white reversed (see Figure 4.5(b)) in order to maintain the consistency, for the result after this phase shall be the vessel in white and background in black.

As mentioned previously, the most important step of detecting centerline pixels is applying the first-order derivative filter (given in Equation (3.7)) orthogonally to the main orientation of the vessel centerline in all directions. To enable the detection in all directions, the image itself will be rotated instead of rotating the kernel in avoid of losing image information. The results after each rotation will be combined into one to realize the complete coverage of every direction. During the experiment, we found the best rotation step is 10 degrees (therefore 18 directions in total) for preserving most capillaries while maintaining a good running time. Thus, this entire process can be simply regarded as processing 18 different images and overlapping their results, while the process of each iteration in detail has already been explained in Section 3.3.2. Figure 4.5(c) shows the result of centerline highlighting.

The second part of capillaries detection is to eliminate the centerline candidates that have low connectivities. The connectivity between the current pixel and every of its surrounding neighbours will be examined. If connected, they will be categorized into the same group.

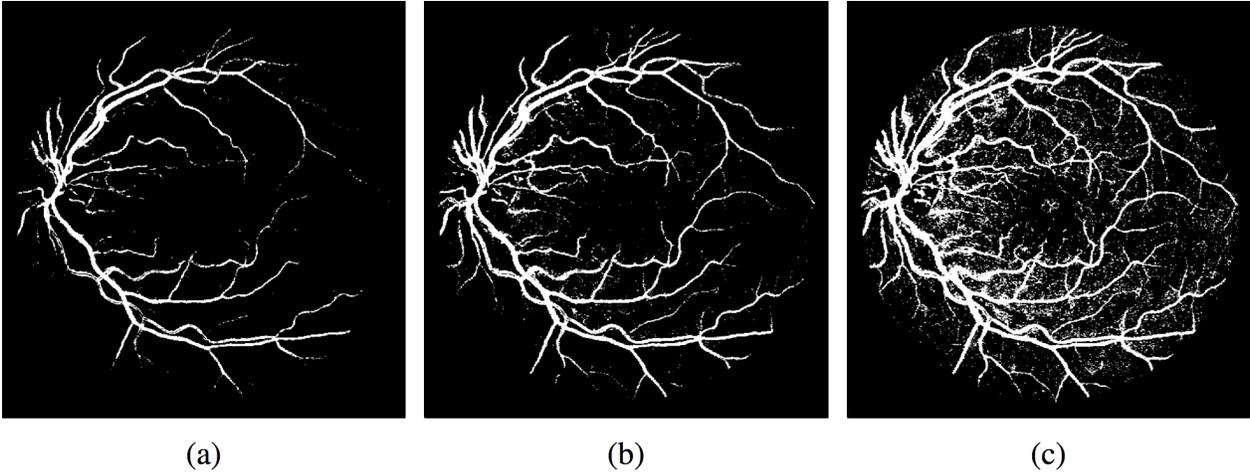


Figure 4.4: Comparison with different thresholds: (a) intensity thresholding with very a strict threshold, (b) intensity thresholding using proposed threshold, (c) intensity thresholding with a loose threshold.

When the examination for every pixel is done, the groups with its member size smaller than a fixed empirical threshold value will be removed. The smaller the threshold is, the more details it preserves, but a smaller threshold also increases the risk of keeping noises, thus pushing more pressure to the de-noising part. On the other hand, although larger threshold may bring in the risk of removing necessary capillary fragments, but it will eliminate more noises. Besides, capillary fragments can hardly have any scientific value in real situation and that is why this system proposes a fairly larger threshold for connectivity check. Figure 4.5(d) shows the final result after connectivity test and Figure 4.5 presents the whole process of capillaries detection.

4.1.4 Post-processing

The final stage of retinal vessel segmentation is post-processing, which contains image overlapping and morphology de-noising. The word “overlapping” here means to combine the venules structure image (see Figure 4.3(c)) and capillaries image (see Figure 4.5(d)) into one, by simply applying the logic OR operation to both of them. The result is shown in Figure 4.6(c).

After overlapping, some noises which have been existing since venules structure extraction was accomplished need to be removed. The reason to keep these noises until the overlapping

phase is because the noises may be overlapped or reconnected with vessel centerlines to enhance the vessels.

The overlapped image will be de-noised by applying erosion operation from the mathematical morphology theory. Erosion operation has the ability to shrink the edge of the larger object and remove small spots, and in this way, not just the coarse edges of the vessels will be polished but also the scattering noises will be removed. Figure 4.6(d) shows the final result after de-noising and Figure 4.6 presents the entire post-processing process.

4.2 Retinal blood vessel segmentation using fully convolutional network with transfer learning

The second proposed implementation of the retinal vessel segmentation system is enlightened by transfer learning, which provides theoretical support to apply the fully convolutional neural network decently in retinal blood vessel segmentation. As described in Section 3.3.2, the original use of the fully convolutional version of AlexNet is semantic segmentation², which is a very efficient way for the machine to visualize and understand the image. The most difficult part in retinal blood vessel segmentation is to distinguish vessels and all the other surrounding tissues / lesions, and the use of efficient semantic segmentation helps to eliminate the noises. There are three innovative points which have eventually made this proposed work successful. First, the proposed method has shifted, or in another word, simplified the typical retinal vessel segmentation problem from full-size image segmentation to regional vessel element recognition. This is to say, vessel pixels are to be recognized from region to region and merged together in the end. Second, because of this problem shifting, the training data, therefore, can be augmented from a hundred to a hundred thousand, which guarantees the effectiveness of deep network training. Third, the proper method of fine tuning the pre-trained semantic segmentation model has made the regional segmentation task much easier. This pre-trained semantic segmentation model is the fully convolutional version of

²Semantic segmentation: the pixel-to-pixel and end-to-end segmentation which labels each pixel within the image with a class of objects (e.g. car, person, cat, and etc.) and another class of non-objects (e.g. ground, sky, water, etc.).

AlexNet, which well performs the pixel-to-pixel and end-to-end segmentation.

Figure 4.7 presents the functional block diagram of the proposed segmentation system and there are three stages of the system. Before processing, the data gathered from the databases will be firstly assigned into the training group and testing group, which will be discussed in Section 5.1.1. In both training and the testing phases, both of the original training and testing datasets will be preprocessed to enhance the image contrast. Next, every image in each dataset will be split into 50×50 image slice, which is the process of data augmentation. The procedure of data augmentation is the essence of this proposed system, not only because it will greatly increase the amount of the dataset, but also because this simplifies the vessel segmentation task from global vascular tree segmentation into regional vessel portion segmentation. After the preparation of image slices, the training data will be utilized to fine tune the pre-trained fully convolutional AlexNet. Then the testing data will be fed to the tuned network in order to test the performance of the tuned network. The results generated by the network will be collected. In the last stage, the collected result slices will be merged into full-size and post-processed with de-noising techniques. The following sections will comprehensively discuss the three phases and each processing module.

4.2.1 Preprocessing

The retinal blood vessel color images, owing to the difficulties in taking photos through the pupil, more or less have unbalanced illumination. Usually, the vessels in the dark are very hard to tell because of the low contract. In another case, some images were taken under the defocused condition, thus making the vessel appearing blurred. Figure 4.8 presents a retinal image with unbalanced illumination and another image with the defocused camera.

To improve the image quality, all the images will be firstly removed from the “black ring” area outside the field of the view, and then conducted the contrast enhancement. When removing the “black ring”, the removed area will be replaced with the color of the mean intensity from the red-green-blue channels respectively. In this way, the background of the retinal image is balanced (see Figure 4.9(b)). Next, the Gaussian blur will be applied particularly to the replaced area, in order to avoid the edge effect when conducting the contrast enhancement. With the smoothed edge of the field of the view and the the balanced

background, contrast enhancement serves a better purpose of highlighting vessels from the fundus (see Figure 4.9(c)). At last, a new black background will be placed at where the “black ring” used to be. This will prevent the edge effect during the training phase (see Figure 4.9(d)). Figure 4.9 presents the entire process of preprocessing a retinal image.

4.2.2 Data augmentation

The typical data augmentation process in image deep learning including flipping, rotation, scaling, and cropping, are used to boost the training performance, especially when there are not sufficient dataset. Image cropping strategy is used here in this proposed implementation, where one full-size image is going to be cropped into multiple image slices. However, cropping the retinal color image and enlarge the size of the training dataset is not the original intention. It is the fact that the fully convolutional AlexNet performs the best in regional vessel portion segmentation that requires image cropping in order to prepare the training data. As a result, this process of image cropping happens to benefit in increasing the size of the dataset.

During the process of data augmentation, each full-size image after preprocessing will be cropped into 50×50 image slices, and each image slice is half overlapped with its neighbors. The 50×50 image slices will then be enlarged into 500×500 , in order to magnify the details. Figure 4.10 presents the general idea of this data augmentation process. Note that the full-size ground truth images will also be divided into the 50×50 images slices in the same rule.

In this way, the total number of the retinal blood vessel image has been increased from 133 full-size images to 84843 image slices. The purpose of training, in this case, becomes generating a fully convolutional network that works for labeling vessel pixels from a small region of the retinal color image. Technically speaking, this makes the segmentation task easier because the features appearing within a small area of the retinal color are as simply as the reddish and tubular object. Figure 4.11 shows a pair of image slice and its ground truth image.

4.2.3 Network architecture for training and testing

This section will mainly present the architecture of the fully convolutional network applied in the proposed work. The entire procedure of the training process does not require human interference after the network architecture has been designed and training strategy has been set up. The training setup and strategy information including platform, solver type, learning rate and etc. will be included in 5.1. The computer will do the rest of the work to output the trained network, which will be utilized for testing and generating the binary vessel segmentation results. Figure 4.12 shows some samples of the input and output images of the testing process. These results will be recovered into full-size and post-processed in the last step.

Figure 4.13 presents the architecture of the fully convolutional AlexNet applied in the retinal vessel segmentation task, the framework is defined in the work of [40], where there has combined layers of the feature hierarchy and refines the spatial precision of the output. The input is a slice of 500×500 RGB retinal color image, and the output is a 500×500 binary image.

There are in total 8 convolution layers and 1 deconvolution layer in the network. Each convolution layer outputs a BLOBs (binary large object), depicted as white blocks in Figure 4.13, is the input for the following next layer, and is labeled with their *height* \times *width* \times *depth*. The kernel size of the first Convolution layer is 11, with the stride and padding value to be 4 and 100, while the kernel size of the second layer is 5, and its stride and padding value are 1 and 2. These two layers are each followed by a max pooling layer and a local response normalization layer. The kernel size of the pooling layer is 3 and the stride value is 2, and the local size of the normalization layer is 5. The third, the fourth, and the fifth convolution layer have the same kernel sizes of 3, as well as the same stride and padding values of 1. Another max pooling layer, same as the previous one, follows the fifth convolution layer. The sixth, the seventh, and the eighth convolution layers do not do padding and their stride value are 1. The kernel sizes of them are 6, 1, 1, respectively. For the last deconvolution layer, its kernel size is 63 and stride is 32.

4.2.4 Merging and overlapping

As previously mentioned, the training data will be used to fine tune this pre-trained network, and the trained network will be utilized for generating the segmentation results. Shown in Figure 4.12, these result pieces have to be recovered into the full-size binary segmentation results. Since each segmentation slice has a reference number to identify its belonging and the exact location of its belonged image, the results will be merged easily. When the overlapping areas containing conflict pixels between the neighbor slices, the system performs OR operation to determine whether the pixel belongs to the vessel class or not. Figure 4.14 (c) shows the full-size result after the merge of the small image slices. The full-size results will be further processed to be smoother and more accurate.

4.2.5 Post-processing

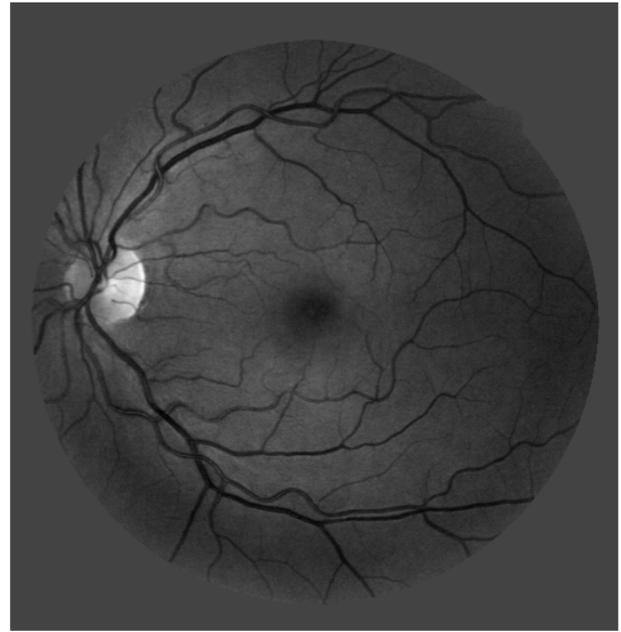
As the semantic segmentation of the retinal vessels appears to be wider than the vessel in ground truth image, and there are noises after merging and overlapping. Therefore, the post-processing step is expected in order to slightly increase the accuracy and look nicer.

To restore the actual width of the vessel, a small threshold window of 9×9 will be applied to the image after merging and overlapping as shown in Figure 4.14 (c). This window works with the pixels along the vessel edges. When an edge pixel is detected, this pixel becomes the center of the 9×9 window, and the window area will be transformed into grayscale. Next, based on the greyscale distribution of the window, a threshold will be generated using the Otsu's method. If the intensity of the centered pixel is lower than this threshold, this centered pixel will be considered as the vessel, otherwise, this pixel will be classified as background.

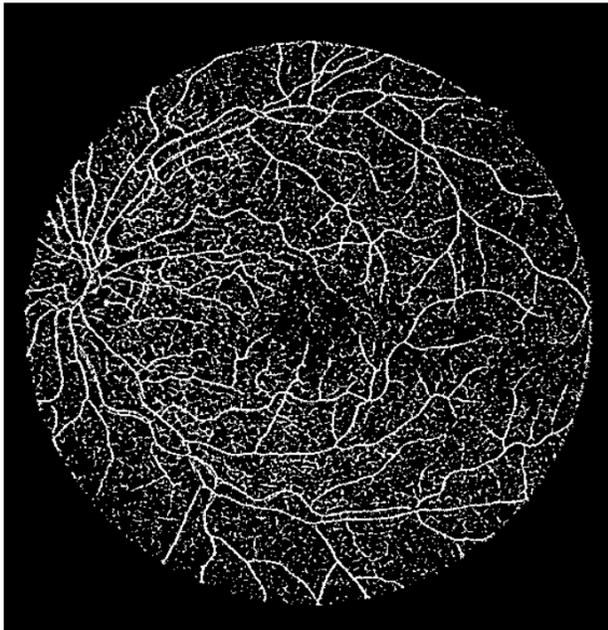
Figure 4.14 (d) presents the final look of the image after post-processing. Compared with the image before post-processing in Figure 4.14 (c), the noises have been eliminated and the width of the vessels is more close to the ground truth in Figure 4.14 (b).



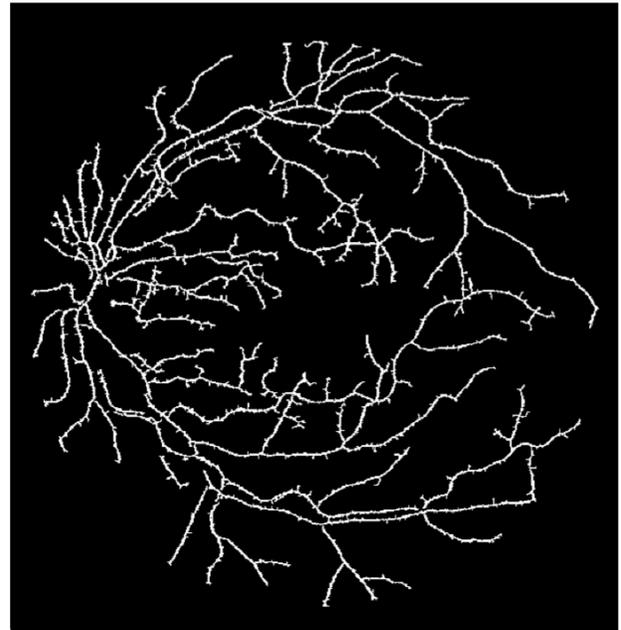
(a)



(b)

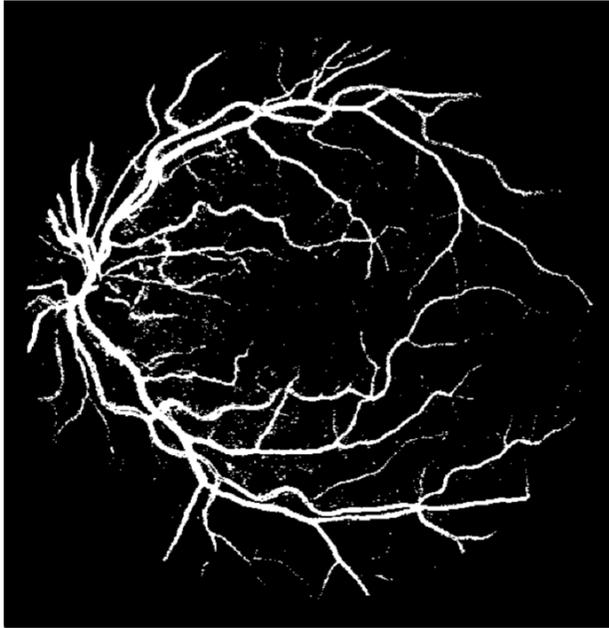


(c)

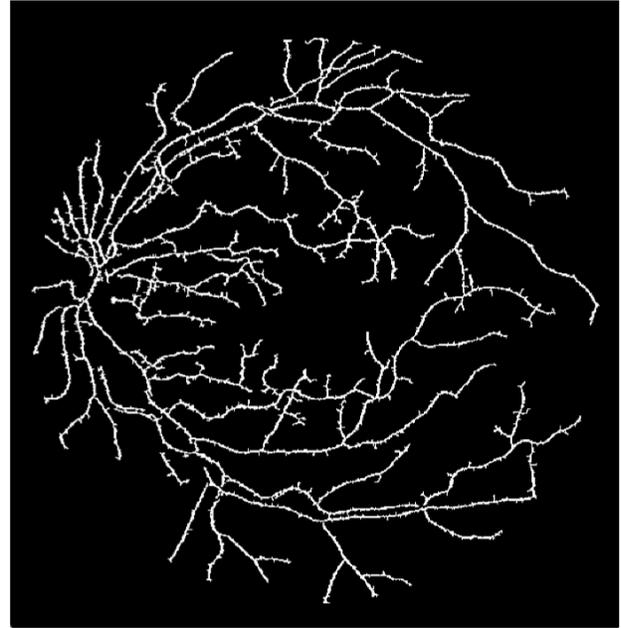


(d)

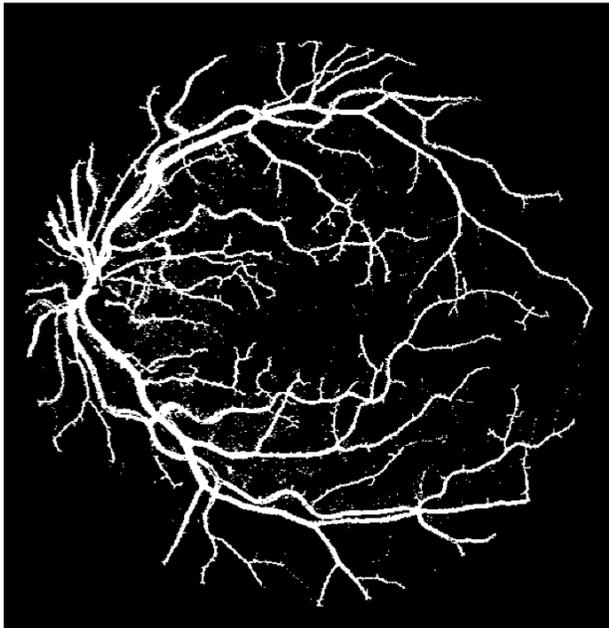
Figure 4.5: Capillaries detection phase: (a) image after preprocessing, (b) image after black-and-white inversion, (c) image after centerline highlighting, (d) the final result of capillaries detection after connectivity test.



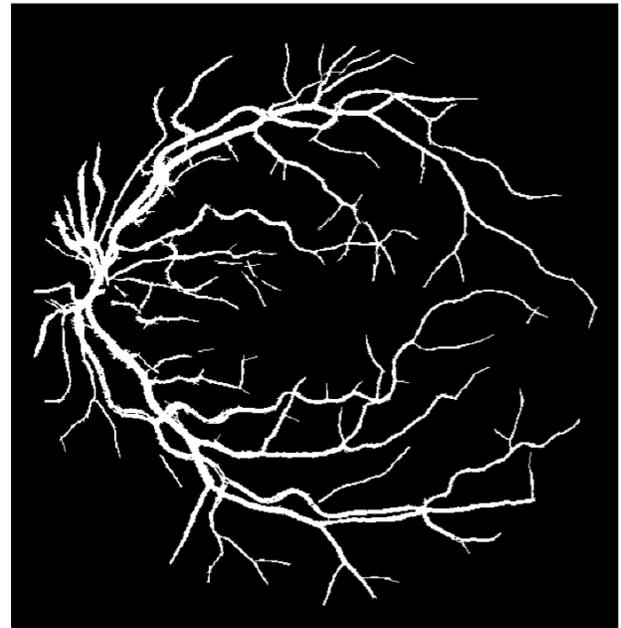
(a)



(b)



(c)



(d)

Figure 4.6: Post-processing phase: (a) image after venules structure extraction, (b) image after centerline detection, (c) image after overlapping, (d) the final result of retinal vessel segmentation after de-noising.

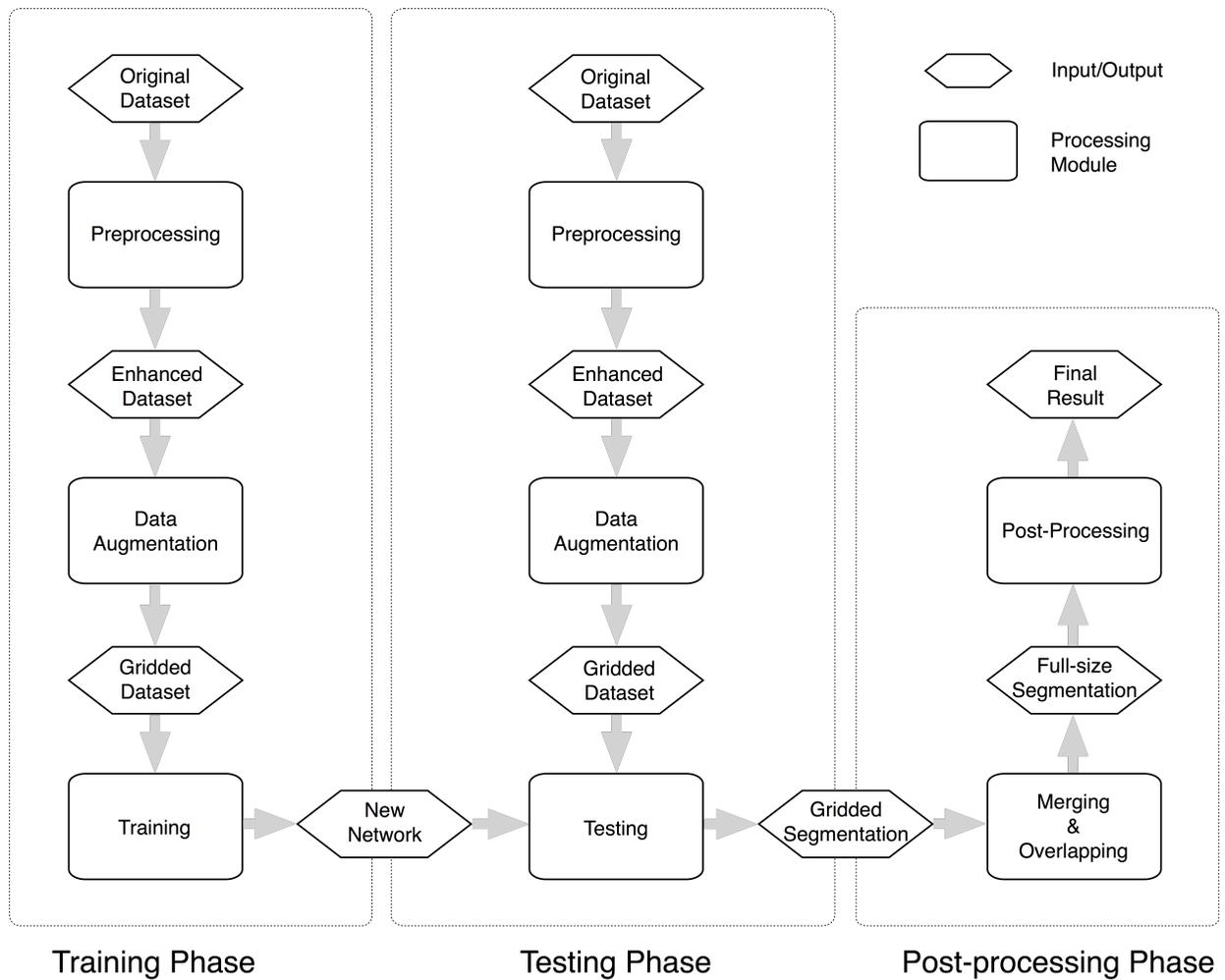


Figure 4.7: Functional diagram of the proposed supervised retinal vessel segmentation system.

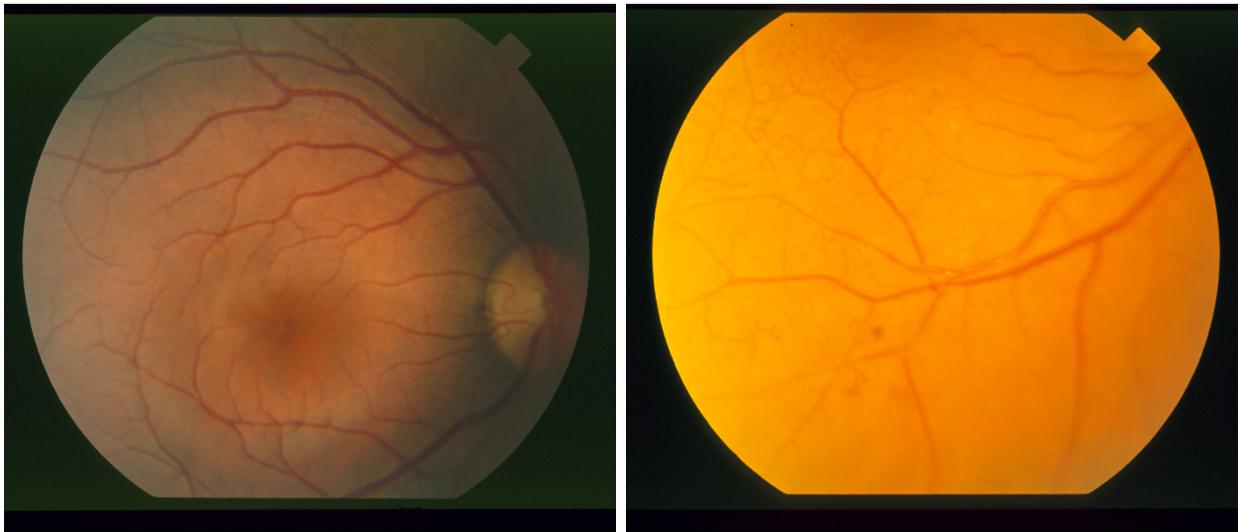
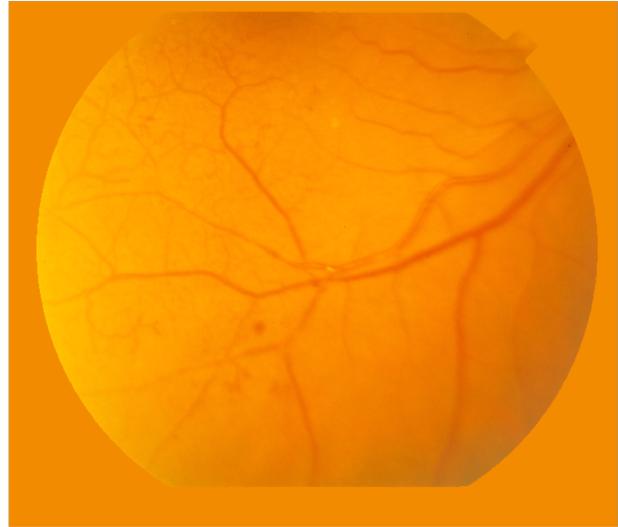


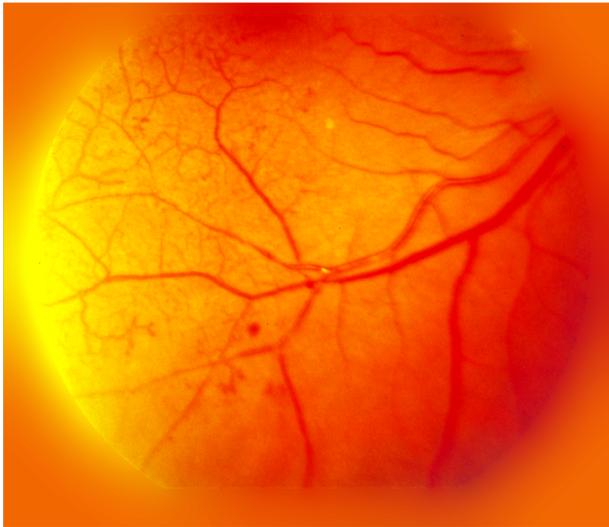
Figure 4.8: The retinal images taken with unbalanced illumination (left) and unfocused camera (right).



(a)



(b)



(c)



(d)

Figure 4.9: Preprocessing phase: (a) original color image, (b) image after “black ring” replacement, (c) image after Gaussian smooth (only applied to the outer area), (d) final look of the preprocessed image.

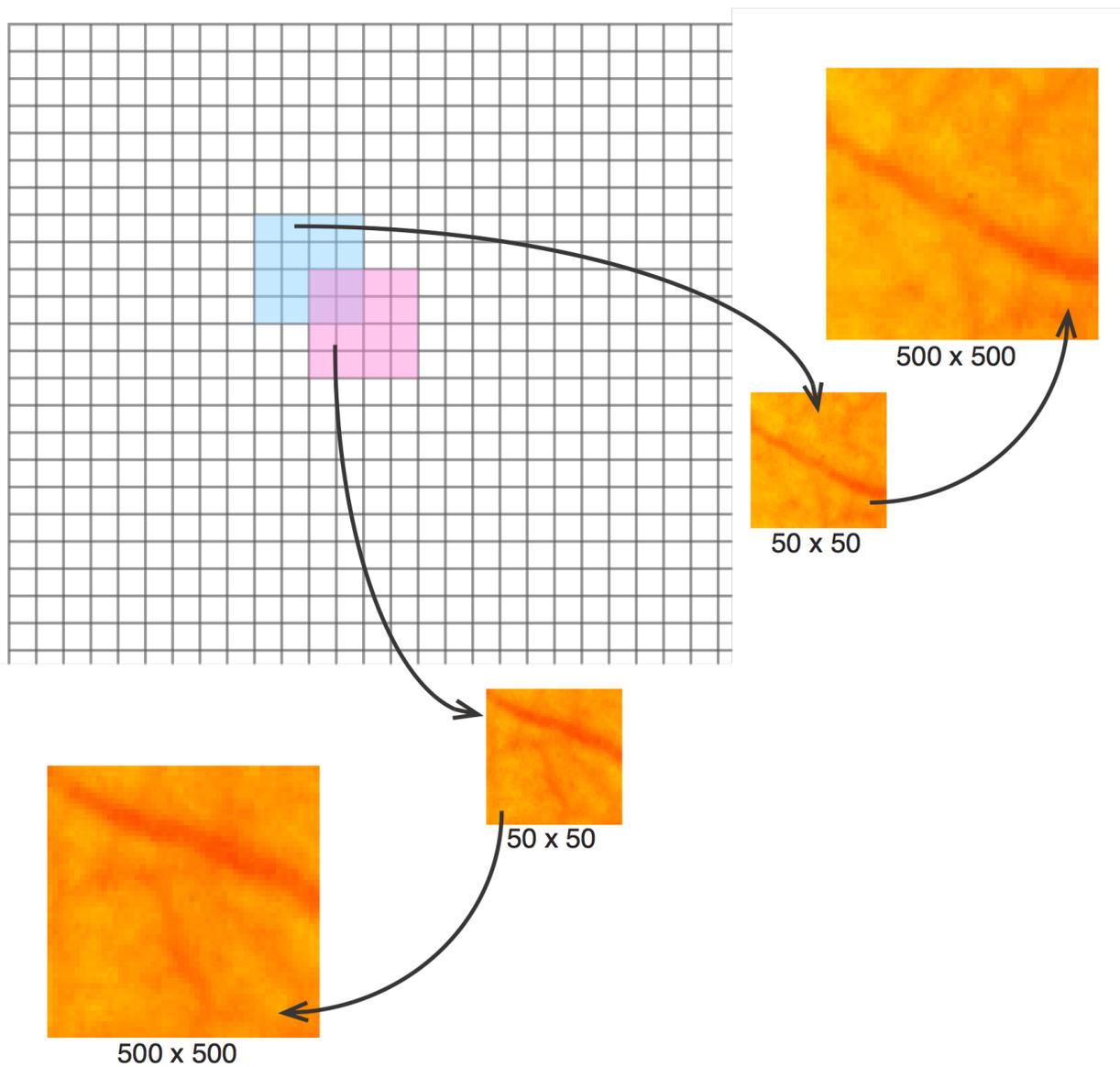


Figure 4.10: A full-size image will be divided into multiple 50×50 image grids and then be resized into 500×500 .

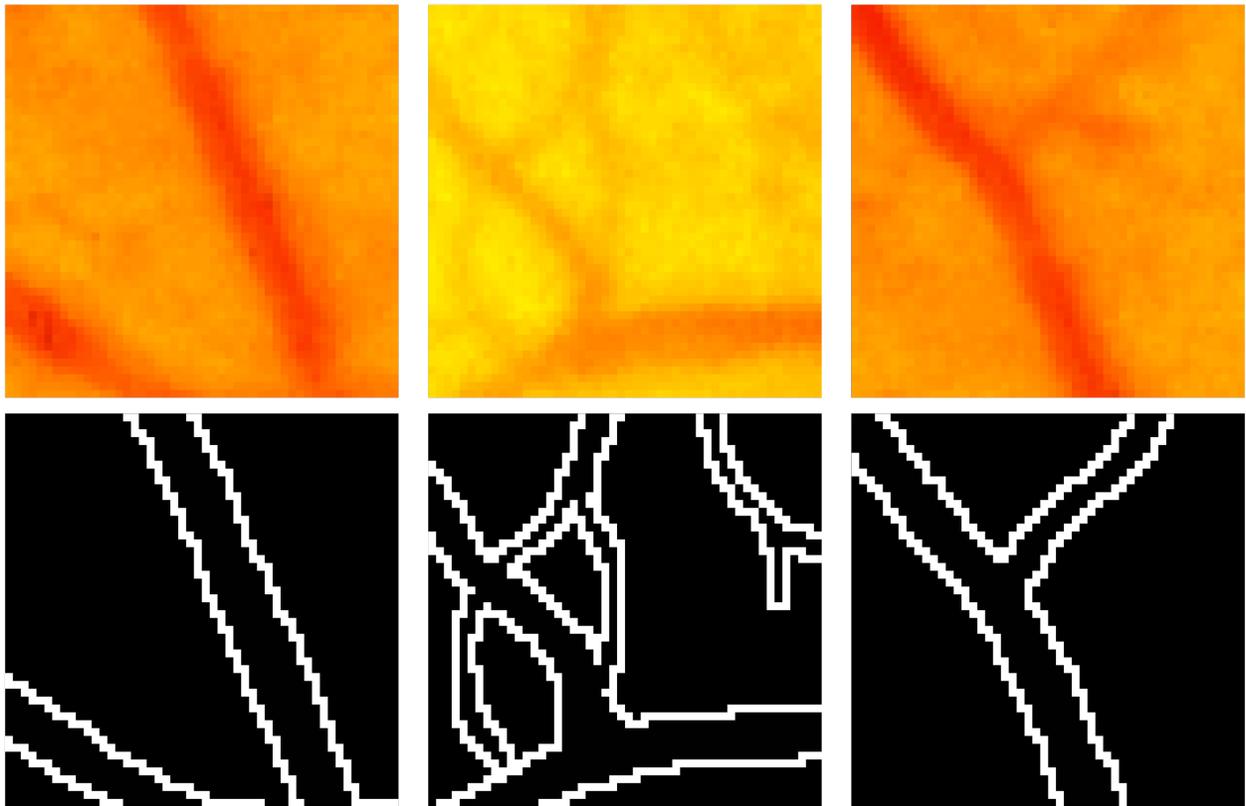


Figure 4.11: The 500×500 retinal color image grids (first row) and their corresponding ground truth images (second row).

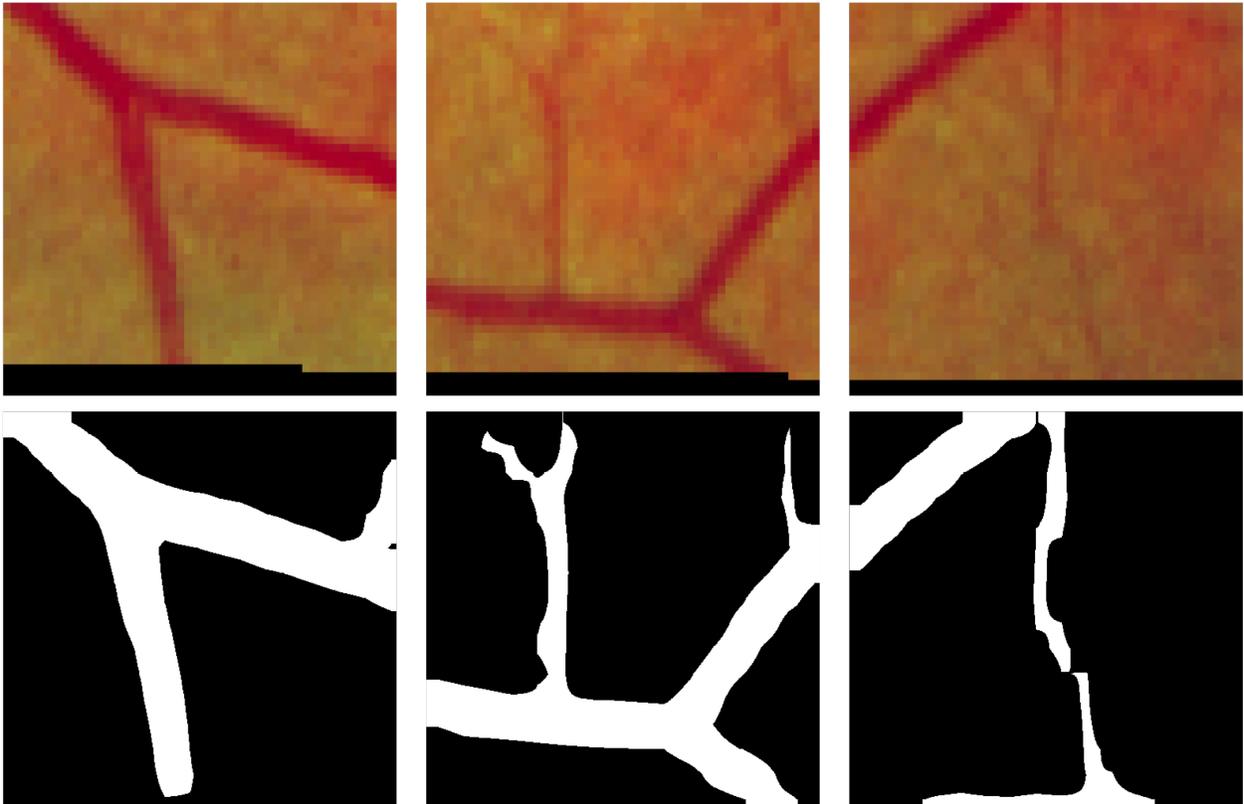


Figure 4.12: The input (first row) and output (second row) image slices of the testing process

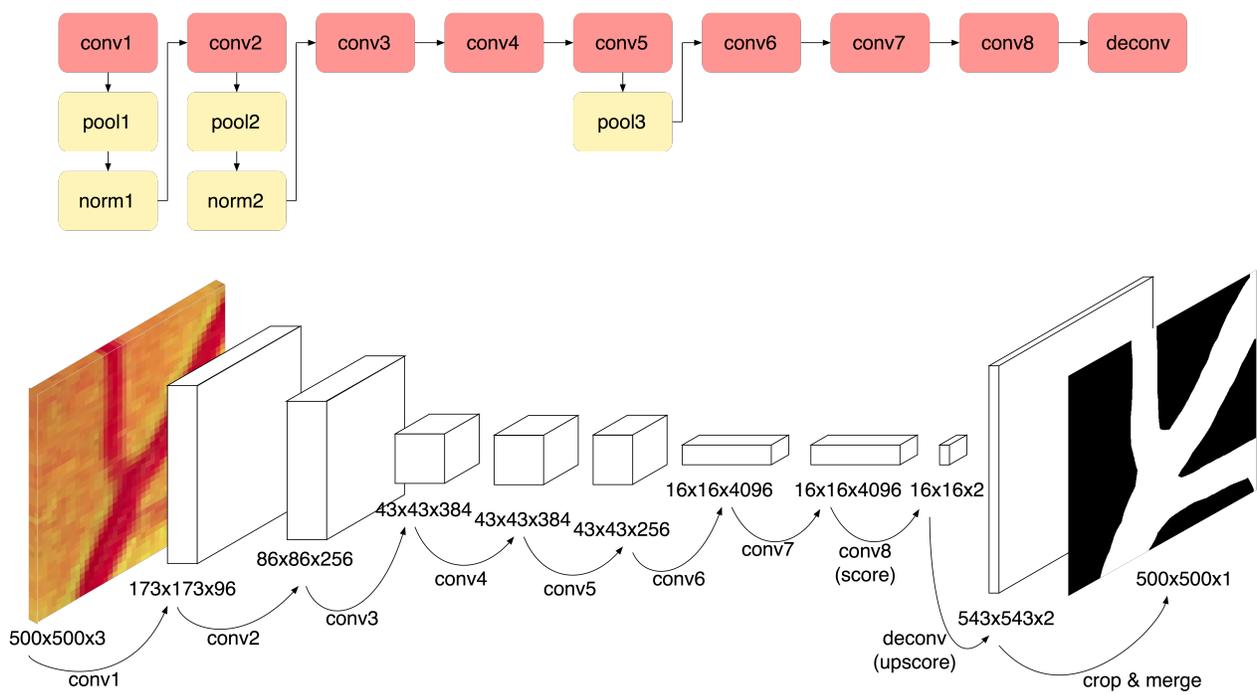
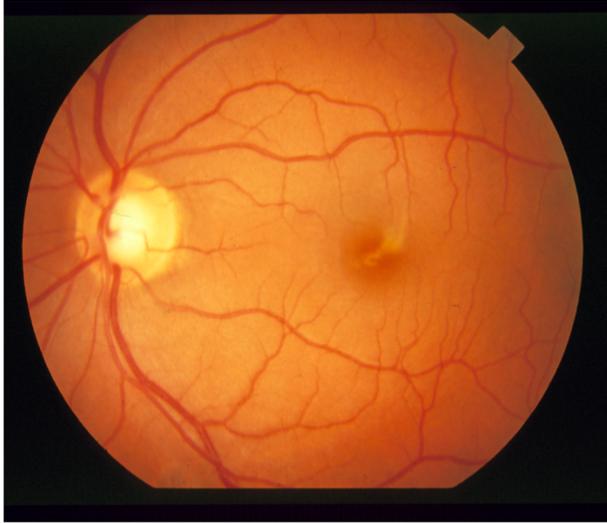
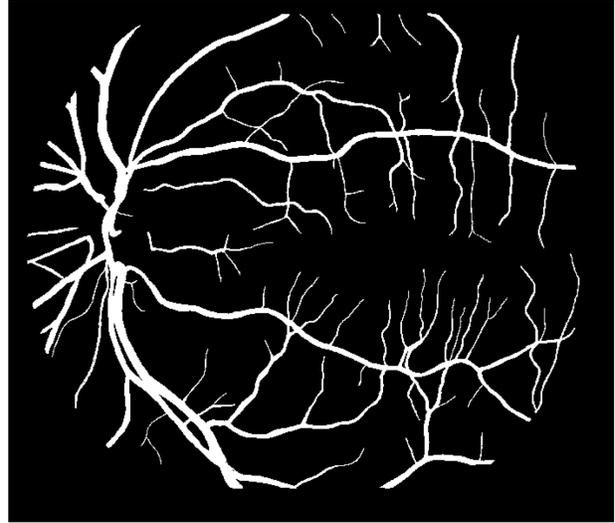


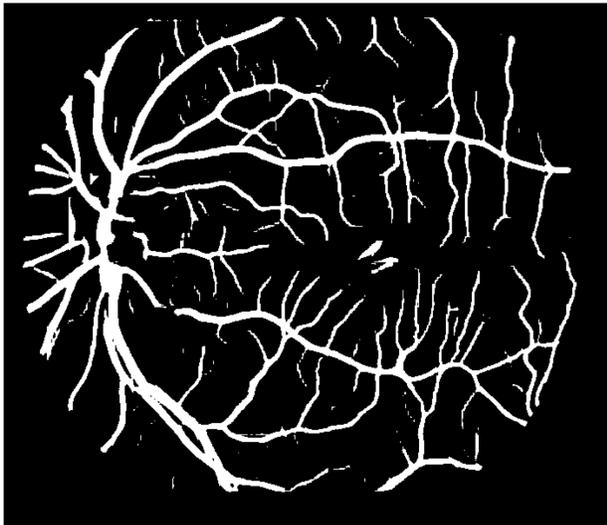
Figure 4.13: The architecture of the fully convolutional AlexNet for retinal vessel segmentation



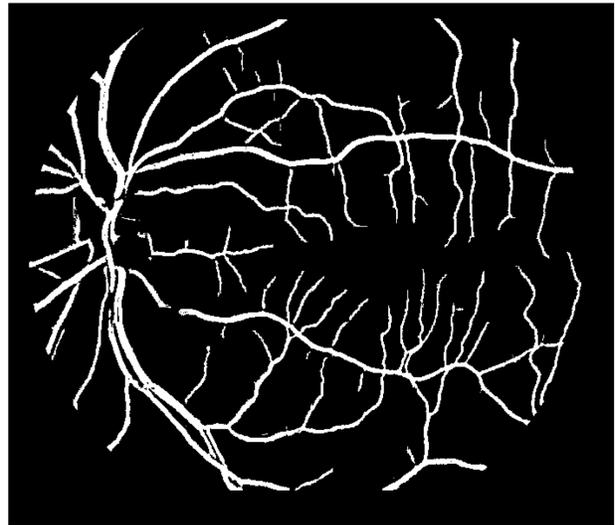
(a)



(b)



(c)



(d)

Figure 4.14: The comparison of the full-size results: (a) original color image, (b) the ground truth image, (c) the image after merging and overlapping, (d) the final result after post-processing.

CHAPTER 5

PERFORMANCE EVALUATION: WHICH ONE IS THE BEST

5.1 Experiment setup

Experimental settings for the retinal vessel segmentation are subtle but crucial, especially for the data preparation of the supervised learning. This section will briefly introduce the preparation of categorizing training and testing data, configuration of training parameters, and execution environment. With those information, this proposed research may be reproduced by other researchers and compared fairly with other related works.

5.1.1 Data preparation

Unsupervised

The first implementation proposed in Section 4.1, is an unsupervised method, has been designed based on the study of the databases from DRIVE and STARE. There will not be the cross-database test, but we will investigate its robustness based on its performance on CHASE_DB1 and HRF databases. In the first stage, all the 133 retinal images from the four databases will be used to evaluate the performance of this system.

Next, only healthy retinal images will be grouped and used to evaluate the performance of the system again, and this will assess its practical usability for normal people. As STARE and HRF databases provide the classification of healthy retinal images and their ground truth images, we will use these two databases to test the proposed system one more time. STARE database has in total 10 healthy retinal images, while HRF database has 15 healthy retinal

images.

Supervised

. The second implementation proposed in Section 4.2, is a supervised method using deep learning techniques. In this case, Single-database test and cross-database test are both performed to fully and objectively evaluate its performance. The single-database test just involves the DRIVE and the STARE databases, and the 60 full-size images from those two databases will firstly be augmented into 33680 image slices, 50% of the images slices (16840 images) will be randomly selected as the training data. Then, all of the 33680 image slices will be used for testing and processed into full-size for performance evaluation. The cross-database test involves all the four databases, and their full-size images will also have to be augmented into image slices in the beginning. Next, there will be four groups, and each group is made up of three databases for training and another database for testing. Table 5.1 shows the training and testing strategy for the cross-database test.

Table 5.1: The arrangement of databases for the cross-database training and testing.

Group	Training Dataset	Testing Dataset
1	CHASE_DB1 & DRIVE & HRF	STARE
2	CHASE_DB1 & DRIVE & STARE	HRF
3	CHASE_DB1 & HRF & STARE	DRIVE
4	DRIVE & HRF & STARE	CHASE_DB1

5.1.2 Execution environment setup

Unsupervised

The first proposed system has been implemented on two platforms and two languages. One is developed in MATLAB and executed on a personal computer with an Intel Core i7 CPU 4850HQ at 2.3 GHz and 16 GB 1600 MHz DDR3 RAM. We have performed both linear and parallel execution model in MATLAB. The other implementation is written in Python and

executed on a Raspberry Pi 3 Model B with a quad-core 64-bit ARM Cortex A53 at 1.2 GHz and 1GB of LPDDR2-900 SDRAM. We only perform the linear execution on Raspberry Pi because of the limitation of the hardware.

The second test of this unsupervised method is to investigate its performance when dealing with just healthy retinal images. Since only STARE and HRF databases provide the classification of healthy and abnormal retinal images, this test will just include these two databases.

Supervised

The training process of the proposed method has been conducted on a Linux server with a NVIDIA Tesla K40c GPU, an Intel Xeon CPU E5-2630 v2 at 2.60GHz and 32 GB 1333 MHz DDR3 RAM. The training of the fully convolutional network is conducted on an interactive deep learning GPU training system called DIGITS ¹.

There are four groups of the training process in this work. Each training process has 30 training epochs. Since transfer learning is based on a pre-trained model, the initial learning rate we set is 0.0001 and will be divided by 10 each time stepping down. The solver type we use is Stochastic gradient descent. Generated by DIGITS, Figure 5.1 visualizes the training process of one of the training groups. The upper subgraph in 5.1 presents the loss and accuracy trend throughout the epochs from 1 to 30. The lower subgraph in 5.1 shows the change of learning rate in three stages.

The testing process has been executed on the personal computer with a software framework called Caffe, which is developed by Yangqing Jia as part of his Ph.D. at UC Berkeley and supports many different types of deep learning architectures geared towards image classification and image segmentation, such as the fully convolutional network. Caffe supports GPU based acceleration using CuDNN of NVIDIA. The new model generated by DIGITS will be sent to this personal computer for image segmentation. The test program is written in Python.

¹DIGITS is a deep learning GPU training system developed by NVIDIA. It is a new system for developing, training and visualizing deep neural networks, which brings deep learning closer and more friendly into the hands of engineers and data scientists. DIGITS can be used to rapidly train the highly accurate deep neural network for image classification, segmentation and object detection tasks

The testing process has been executed on the personal computer with a software framework called Caffe ², which supports many different types of deep learning architectures geared towards image classification and image segmentation, such as the fully convolutional neural network. Caffe supports GPU based acceleration using CuDNN of Nvidia. The new model generated by DIGITS will be sent to this personal computer for image segmentation. The test program is written in Python.

5.2 Performance measurement

As previous mentioned, the proposed two implementations will be measured and compared to the related ones with three metrics; accuracy, sensitivity, and specificity if provided. Accuracy reflects the proportion of pixels that are correctly classified as vessel or non-vessel. Sensitivity (true positive rate) and specificity (true negative rate) are statistical measures of the performance of a binary classification test, where sensitivity reflects the ability of the algorithm to detect the vessel pixels while specificity is the ability to detect non-vessel pixels (or the ability to avoid noise pixels).

Unsupervised

Table 5.2 presents the performance of the proposed unsupervised implementation on DRIVE, STARE, CHASE_DB1, and HRF databases in terms of accuracy, sensitivity, and specificity. These metrics have already been discussed in Section 3.1.

From Table 5.2, the system performs obviously better on DRIVE and STARE databases than on CHASE_DB1 and HRF databases. As previously introduced, this system has been designed based on the data from STARE and DRIVE databases, therefore, even this is an unsupervised implementation, the segmentation strategy still has a bias for the familiar databases. Especially for the result of CHASE_DB1 database, its sensitivity is the lowest among the others, implying its ability to correctly identify vessels is not as good as the others. Yet, the accuracy of CHASE_DB1 and HRF databases is still high.

²Caffe is a deep learning framework, originally developed by Yangqing Jia as part of his Ph.D. at UC Berkeley. It is open source, under a BSD license. It is written in C++, with a Python interface, and supports GPU based acceleration using CuDNN of Nvidia.

Table 5.2: Performance of the proposed unsupervised implementation on four databases.

Database	Accuracy	Sensitivity	Specificity
DRIVE	0.9597	0.8375	0.9694
STARE	0.9579	0.7767	0.9705
CHASE_DB1	0.9443	0.6310	0.9650
HRF	0.9499	0.7953	0.9594

Next, this unsupervised method has been tested by only healthy retinal images. From Table 5.3, the accuracy of STARE has increased to 0.9594 instead of the accuracy of 0.9579 in Table 5.2. Although The accuracy of HRF has dropped to 0.9465, yet there is a huge increase in sensitivity from 0.7953 to 0.9463, which implies the high precision in vessel pixel classification.

Table 5.3: Performance of the proposed unsupervised implementation on STARE and HRF databases (only healthy retinal images included).

Database	Accuracy	Sensitivity	Specificity
STARE	0.9594	0.7398	0.9833
HRF	0.9465	0.9463	0.9467

We have measured the average execution time for processing one image. Because this proposed system has a parallel execution structure, therefore, we have provided both the linear execution time and the parallel execution time. Table 5.4 shows both of the linear and parallel execution time per image of the proposed unsupervised implementation on all four databases. Because the image resolution from different databases differs from each other, the execution time will increase as the image resolution gets larger.

Supervised

For the single-database test, Table 5.5 presents the performance of the proposed supervised implementation on both DRIVE and STARE databases in terms of accuracy, sensitivity, and specificity. It is clear that the accuracy of all the four databases is very high, especially that

Table 5.4: Linear and parallel execution time per image of the proposed unsupervised implementation on four databases.

Database (Image Resolution)	Linear Execution Time	Parallel Execution Time
DRIVE (565 × 584)	3.79 s	1.68 s
STARE (700 × 605)	4.67 s	2.17 s
CHASE_DB1 (999 × 960)	10.25 s	4.34 s
HRF (3504 × 2336)	85.46 s	34.16 s

of the STARE database. Meanwhile, the CHASE_DB1 has the highest sensitivity and the STARE database also has the highest specificity.

Table 5.5: Performance of the proposed supervised implementation on four databases (single-database test).

Database	Accuracy	Sensitivity	Specificity
DRIVE	0.9624	0.7540	0.9825
STARE	0.9734	0.8352	0.9846
CHASE_DB1	0.9668	0.8640	0.9745
HRF	0.9650	0.8010	0.8010

Figure 5.3 shows two set of result comparison from DRIVE and STARE database respectively. The testing results in Figure 5.3 (c) and (f) are extremely close but just slightly thicker than the ground truth presented in Figure 5.3 (b) and (e).

In order to further assess the robustness of the proposed system, we have also conducted the cross-database test on those four databases. Table 5.6 shows the four groups of results respectively on DRIVE, STARE, CHASE_DB1, and HRF databases. Each result of a database is generated from the network trained by the other three databases (this training strategy has already been introduced in the previous Table5.1).

From Table 5.5, it is clear that the accuracy of DRIVE and STARE databases are incredibly high, which implies the success of the application of transfer learning and the design of the proposed system. Figure 5.3 shows two set of result comparison from DRIVE and STARE database respectively. The testing results in Figure 5.3 (c), (f) look very close to

the ground truth in Figure 5.3 (b), (e). This great achievement proves the proper use of the pre-trained semantic segmentation model – the fully convolutional AlexNet, and also the transform/split of the segmentation task – from global to regional. To further prove the correctness of the utilization of the fully convolutional neural network and the advantage of the proposed system, we have conducted the cross-database test to assess its robustness.

In order to further assess the robustness of the proposed system, we have also conducted the cross-database test on those four databases. Table 5.6 shows the four groups of results respectively on DRIVE, STARE, CHASE_DB1, and HRF databases. Each result of a database is generated from the network trained by the other three databases (this training strategy has already been introduced in the previous Table5.1).

Table 5.6: Performance of the supervised implementation on four databases (cross-database test).

Database	Accuracy	Sensitivity	Specificity
DRIVE	0.9593	0.7121	0.9832
STARE	0.9653	0.7820	0.9798
CHASE_DB1	0.9591	0.7217	0.9770
HRF	0.9662	0.7686	0.9826

From Table 5.6, the accuracy of all the four databases except HRF has dropped a bit in contrast to their single-database test results in Table 5.5. However, the sensitivity of HRF has decreased. Yet, this cross database result remains to be excellent and stable in comparison with other works, which will be discussed later in Section 5.3. Figure 5.4 shows two set of result comparison from DRIVE and STARE databases respectively, while Figure 5.5 presents the comparison between the single and cross-database results from DRIVE and STARE database. The cross-database result in Figure 5.4 (c) from DRIVE database presents a minor difference in comparison with the ground truth image in Figure 5.4 (b). Also by observing Figure 5.5 (b) and 5.5 (c), the difference between the single-database and the cross-database results are small, even though the result in 5.5 (c) is generated by the fully convolutional network trained by the other three irreverent databases. However, the cross-database result from STARE database in Figure 5.5 (e) has a block of noise in the middle.

This is caused by insufficient types of training data. From the original color image in Figure 5.4 (d), there exists a block of a yellow lesion in the middle. Such medical case is rare and never exists elsewhere in databases other than STARE, therefore, the fully convolutional network makes mistake because it has not been trained for such case. Aside from this, the rest of vessel structure appears close to both the ground truth image in Figure 5.5 (d) and the single-database result in Figure 5.5 (e).

5.3 Performance comparison and analysis

This section will present the performance comparison with the related state-of-the-art works on DRIVE, STARE, and CHASE_DB1 databases. Because no relevant result of HRF database has been found from other’s work, the performance comparison of HRF database will be omitted. Three tables are presented in this section. Within each table, the results are separated into cross-database test and single-database test. As unsupervised methods do not have training process at all, they do not necessarily have single or cross-database tests. In order to fairly compare the performance between supervised and unsupervised methods, the results from unsupervised methods can be compared with the results from the cross-database test of the supervised methods. What’s more, the different works are categorized into three groups, which are unsupervised, supervised, and supervised with deep learning. The best values of the results from each group are marked with \diamond and \dagger .

Table 5.7 presents the performance comparison on DRIVE database. For the cross-database test, the first proposed method has the highest accuracy in both the unsupervised and supervised group, which surpasses the second proposed method just 0.0004. However, with this tiny difference, we can say both the proposed implementations stay on the top in terms of accuracy. Meanwhile, the second proposed method performs the best in the supervised group in terms of accuracy and specificity. The shortage in sensitivity of the second method, in contrast to the first one, is owing to the nature of the pre-trained semantic segmentation not being sensitive to small objects, while the manually labeled ground truth images of the DRIVE database contains rich capillaries. However, the specificity of our work beats all the other works in both groups, which implies the highest noise resistance. Along

with the outstanding performance on accuracy, the second proposed method is one of the bests in the DRIVE cross-database test. For the single-database test, the proposed method will be compared with only supervised methods. The accuracy and specificity of the second proposed method stay on the top. Although the sensitivity of the work of [44] and [31] is higher, their accuracy is much lower. Overall, both of the proposed methods outperforms the others in the DRIVE single-database test.

Table 5.7: Performance comparison with the related state-of-the-art works on **DRIVE** database.

Cross-database test				
	Algorithm	Accuracy	Sensitivity	Specificity
Unsupervised	Abdurrazaq et al.[8]	-	0.8214	0.9059
	Fraz et al.[17]	0.9430	0.7152	0.9768
	Lam et al.[29]	0.9472	-	-
	Mendonca et al.[35]	0.9463	0.7315	0.9781 [◊]
	Zhang et al.[47]	0.9476	0.7743	0.9725
	Proposed Method 1	0.9597 [◊]	0.8375 [◊]	0.9694
Supervised	Fraz et al.[19]	0.9456	0.7242	0.9792
	Marin et al.[34]	0.9448	-	-
	Ricci et al.[38]	0.9266	-	-
Supervised (Deep Learning)	Wang et al.[44]	0.9428	0.8431 [†]	0.9574
	Li et al.[31]	0.9486	0.7273	0.9810
	Proposed Method 2	0.9593 [†]	0.7121	0.9832 [†]
Single-database test				
	Algorithm	Accuracy	Sensitivity	Specificity
Supervised	Fraz et al.[19]	0.9480	0.7406	0.9807
	Marin et al.[34]	0.9452	0.7067	0.9801
	Ricci et al.[38]	0.9595	-	-
Supervised (Deep Learning)	Wang et al.[44]	0.9533	0.8173 [†]	0.9733
	Li et al.[31]	0.9527	0.7569	0.9816
	Proposed Method 2	0.9624 [†]	0.7540	0.9825 [†]

[◊] represents the best values among all the unsupervised methods.

[†] represents the best values among all the supervised methods.

Table 5.8 presents the performance comparison on STARE database, which shows that the performance of the second proposed method excels from all the other works. First, the first proposed method still excels in terms of accuracy of the unsupervised group. In the cross-database test, the second method performs the best on both accuracy and sensitivity, with

its specificity slightly lower than the work of [31]. Indeed, because the STARE database is the most complicated database which contains 8 kinds of retinal diseases and 3 of the images are even defocused, the fully convolutional network would not have enough knowledge to deal with those images perfectly if not being trained so. This is also the nature of deep learning. However, in the single-database test, our second implementation outperforms all the others in every three aspects. Especially, the advantage in accuracy of our second implementation is much more superior than other works. Hence, our work is no doubt state-of-the-art on STARE database.

Table 5.8: Performance comparison with the related state-of-the-art works on **STARE** database.

Cross-database test				
	Algorithm	Accuracy	Sensitivity	Specificity
Unsupervised	Fraz et al.[17]	0.9442	0.7311	0.9681
	Lam et al.[29]	0.9567	-	-
	Mendonca et al.[35]	0.9479	0.7123	0.9758
	Zhang et al.[47]	0.9554	0.7791 [◊]	0.9758 [◊]
	Proposed Method 1	0.9579 [◊]	0.7767	0.9705
Supervised	Fraz et al.[19]	0.9493	0.7010	0.9770
	Marin et al.[34]	0.9528	-	-
	Ricci et al.[38]	0.9452	-	-
Supervised (Deep Learning)	Wang et al.[44]	0.9413	0.7116	0.9675
	Li et al.[31]	0.9545	0.7027	0.9828 [†]
	Proposed Method 2	0.9653 [†]	0.7820 [†]	0.9798
Single-database test				
	Algorithm	Accuracy	Sensitivity	Specificity
Supervised	Fraz et al.[19]	0.9534	0.7548	0.9763
	Marin et al.[34]	0.9526	0.6944	0.9819
	Ricci et al.[38]	0.9646	-	-
Supervised (Deep Learning)	Wang et al.[44]	0.9621	0.8104	0.9791
	Li et al.[31]	0.9628	0.7726	0.9844
	Proposed Method 2	0.9734 [†]	0.8352 [†]	0.9846 [†]

[◊] represents the best values among all the unsupervised methods.

[†] represents the best values among all the supervised methods.

Table 5.9 presents the performance comparison on CHASE_DB1 database. As previously mentioned, the CHASE_DB1 database is not as popular as DRIVE and STARE databases, therefore there are fewer related works that have performed cross-database tested on this

database. In the unsupervised group, the result of [47] has finally surpassed our work in every aspect, but just a tiny bit. For the cross-database test, the accuracy and specificity of the second proposed method have outperformed all the other works. Although the sensitivity of the work of [47] is higher than the proposed second method, its much lower accuracy makes his work much less competitive. For the single-database test, the work of [31] seems to have threatened the predominant ranking of the proposed method, because its sensitivity and specificity is higher than the second method, and its accuracy in the single-database test is just a bit lower. However, when comparing its accuracy in the cross-database test, the work of [31] shows its insufficiency in robustness, which also reflects the excellent overall performance of the proposed work.

Table 5.9: Performance comparison with the related state-of-the-art works on CHASE_DB1 database.

Cross-database test				
	Algorithm	Accuracy	Sensitivity	Specificity
Unsupervised	Zhang et al.[47]	0.9457 [◊]	0.7562 [◊]	0.9675 [◊]
	Proposed Method 1	0.9443	0.6310	0.9650
Supervised	Fraz et al.[19]	0.9415	0.7103	0.9665
Supervised (Deep Learning)	Li et al.[31]	0.9429	0.7118	0.9791 [†]
	Proposed Method 2	0.9591 [†]	0.7217 [†]	0.9770
Single-database test				
	Algorithm	Accuracy	Sensitivity	Specificity
Supervised	Fraz et al.[19]	0.9469	0.7224	0.9711
Supervised (Deep Learning)	Li et al.[31]	0.9581	0.7507	0.9793 [†]
	Proposed Method 2	0.9668 [†]	0.8640 [†]	0.9745

[◊] represents the best values among all the unsupervised methods.

[†] represents the best values among all the supervised methods.

Table 5.10 shows the comparison of the execution time of the related works. The execution time is collected and calculated using the image from DRIVE database. The execution platforms are also provided along with each execution time. Take the factor of execution platform into consideration, the execution time of the first proposed implementation is in the third place in Table 5.10, and it has the shortest execution time on CPU platform. Although the GPU and FPGA implementations of [27] and [26] have a very short execution, the average accuracy of [27] and [26] is 0.9468 and 0.9234, respectively.

Table 5.10: Execution time comparison with a single image from **DRIVE** database.

Algorithm	Execution Time (per image)	Platform
Koukounis et al.[26]	0.049 s	FPGA Spartan 6
Krause et al.[27]	1.2 s	NVIDIA Geforce GPU GTX680
Zhang et al.[47]	12 s	Intel Core i7 CPU @ 2.7 GHz 4 GB RAM
Marin et al.[34]	90 s	Intel Core 2 Duo CPU @ 2.13 GHz 4 GB RAM
Fraz et al.[19]	100 s	Intel Core 2 Duo CPU @ 2.27 GHz 4 GB RAM
Lupascu et al.[33]	125 s	Intel Core 2 Duo CPU @ 3.16 GHz 3.25 GB RAM
Lam et al.[29]	780 s	Intel Core 2 Duo CPU @ 1.83 GHz 2 GB RAM
Proposed Method 1	1.68 s	Intel Core i7 CPU @ 2.3 GHz 16 GB RAM
Proposed Method 2	324.8 s	Intel Core i7 CPU @ 2.3 GHz 16 GB RAM

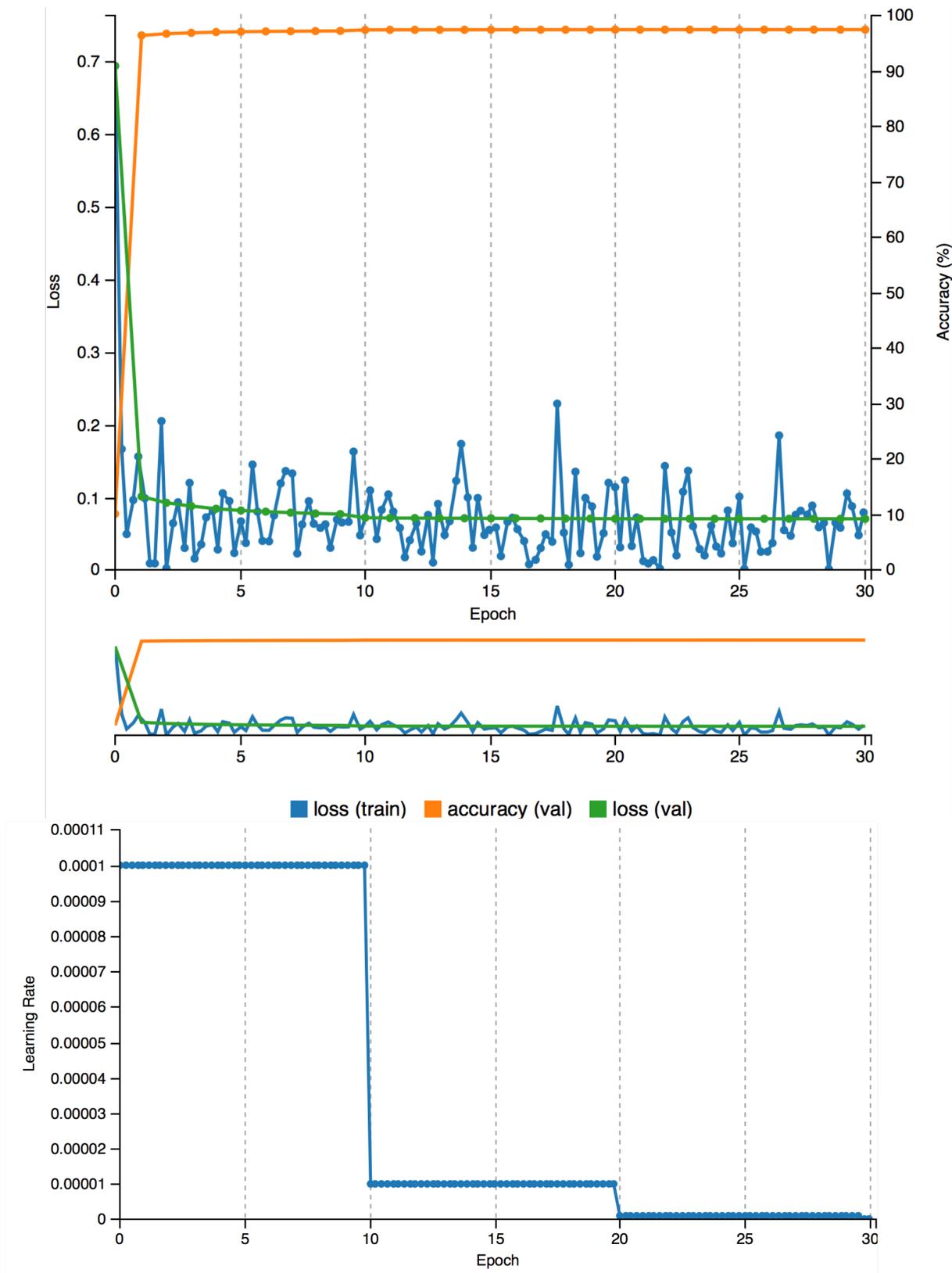


Figure 5.1: The trending chart of accuracy, loss, and learning rate during the training process (generated by DIGITS).

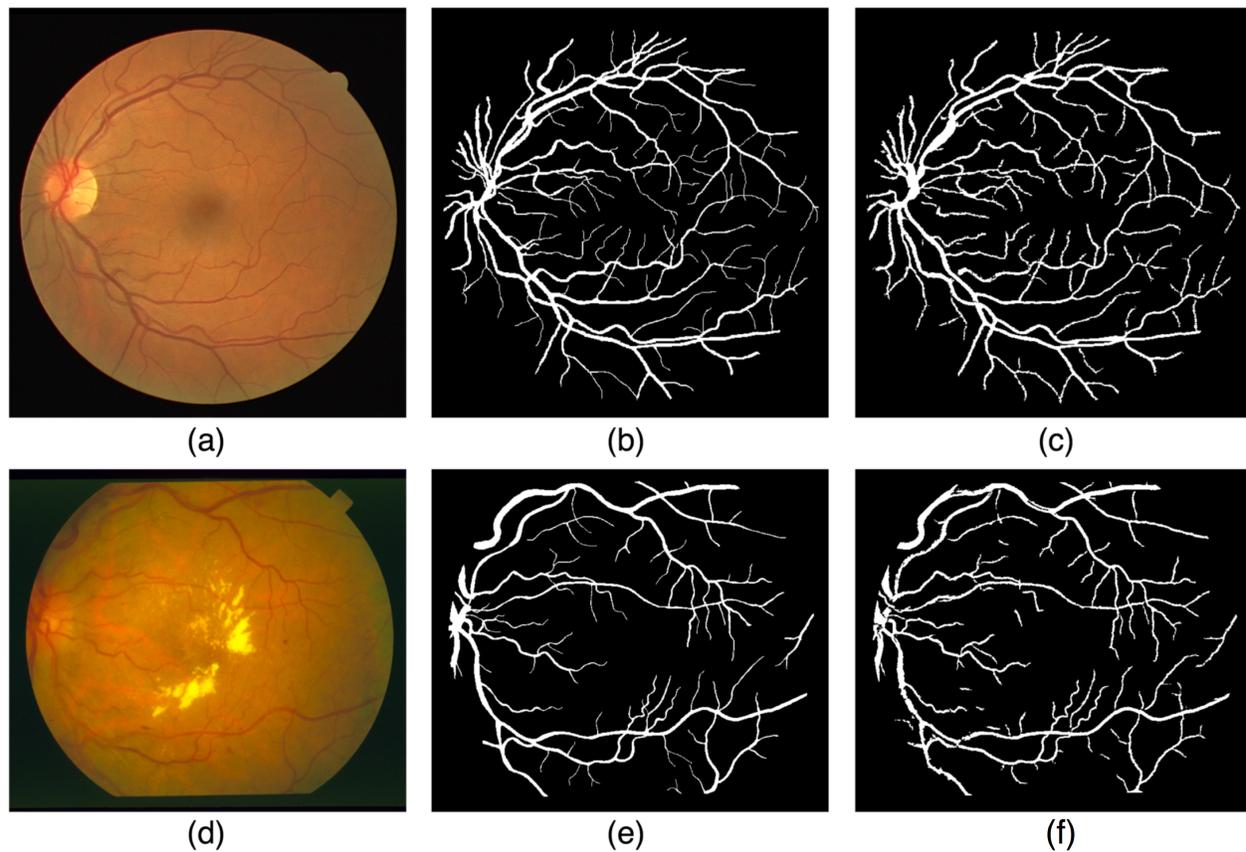


Figure 5.2: Single-database test result comparison of the image from DRIVE (first row) and STARE (second row) databases: (a), (d) are original color images; (b), (e) are ground truth images; (c), (f) are single-database testing results;

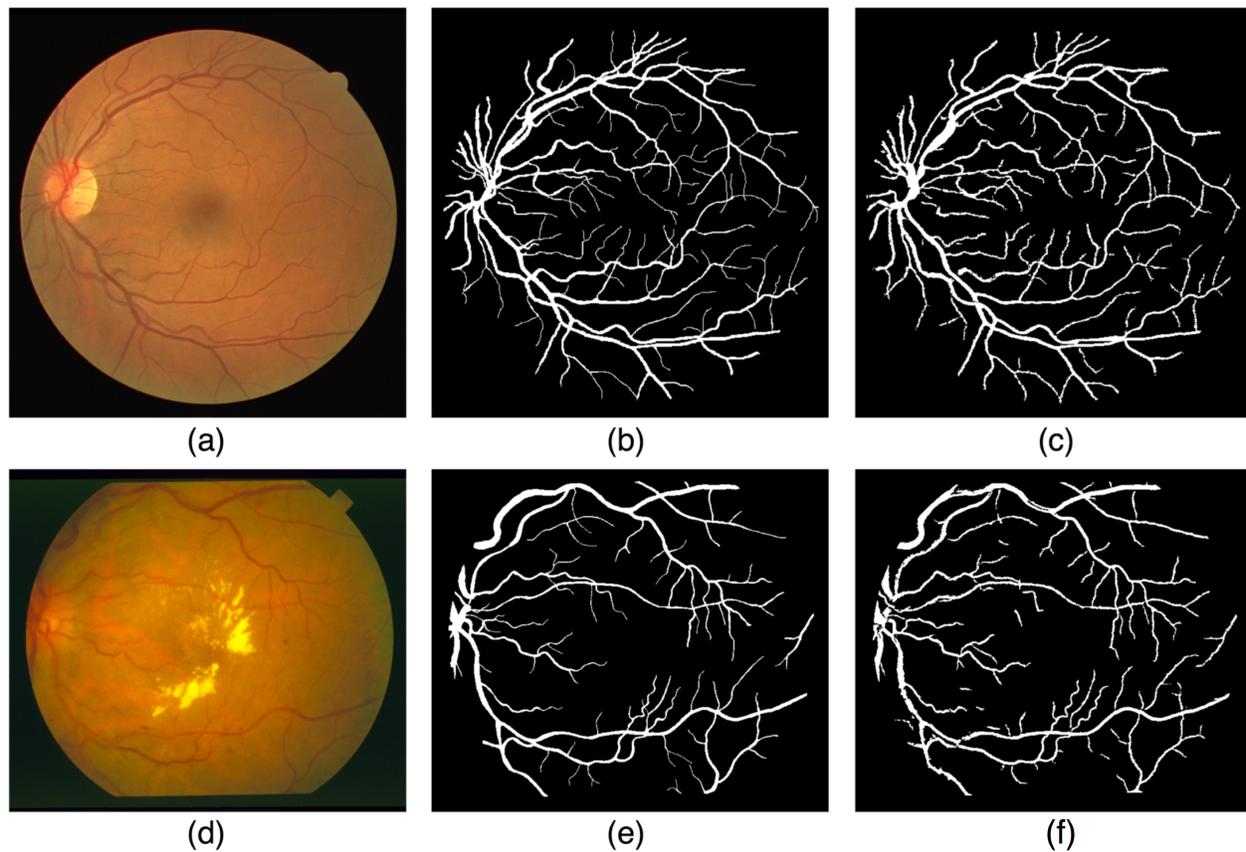


Figure 5.3: Single-database test result comparison of the image from DRIVE (first row) and STARE (second row) databases where (a), (d) are original color images; (b), (e) are ground truth images; (c), (f) are single-database testing results.

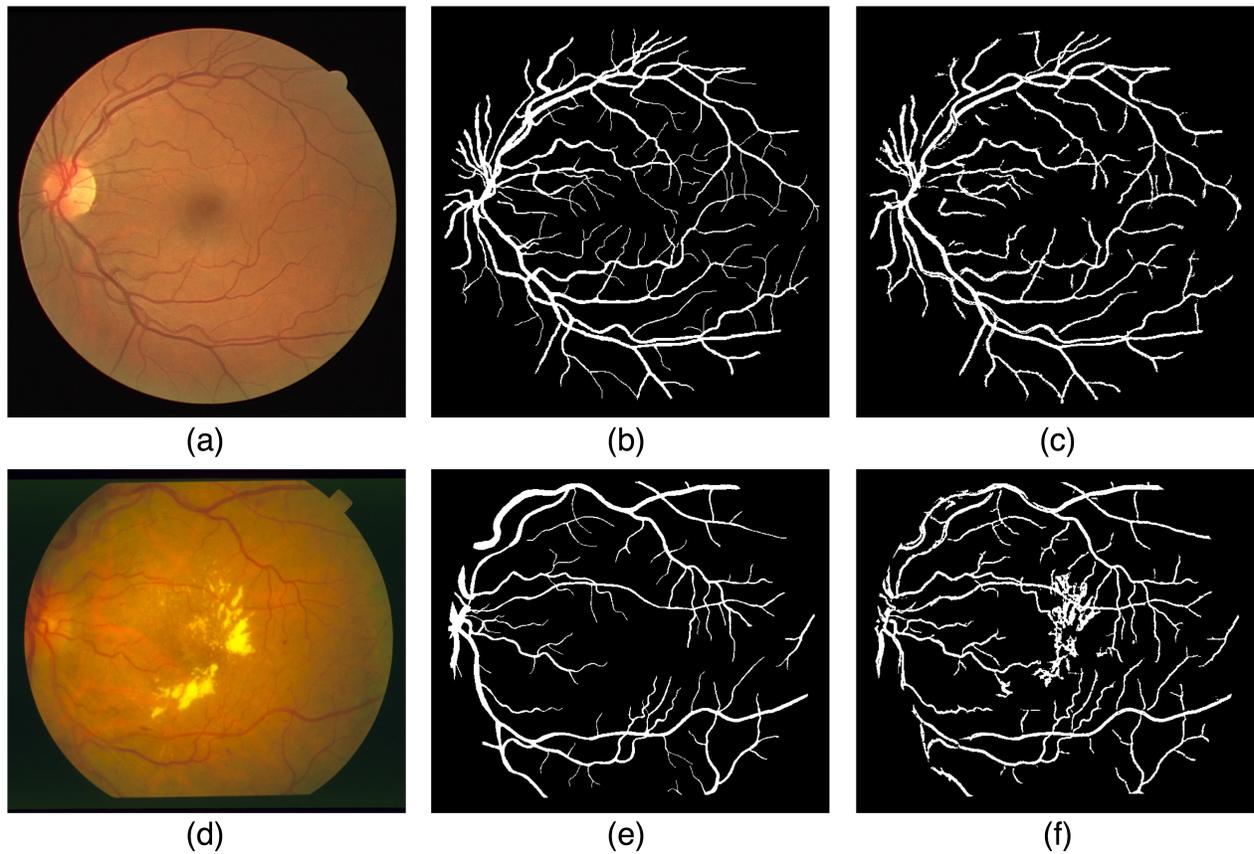


Figure 5.4: Cross-database test result comparison of the image from DRIVE (first row) and STARE (second row) databases where (a), (d) are original color images; (b), (e) are ground truth images; (c), (f) are cross-database testing results.

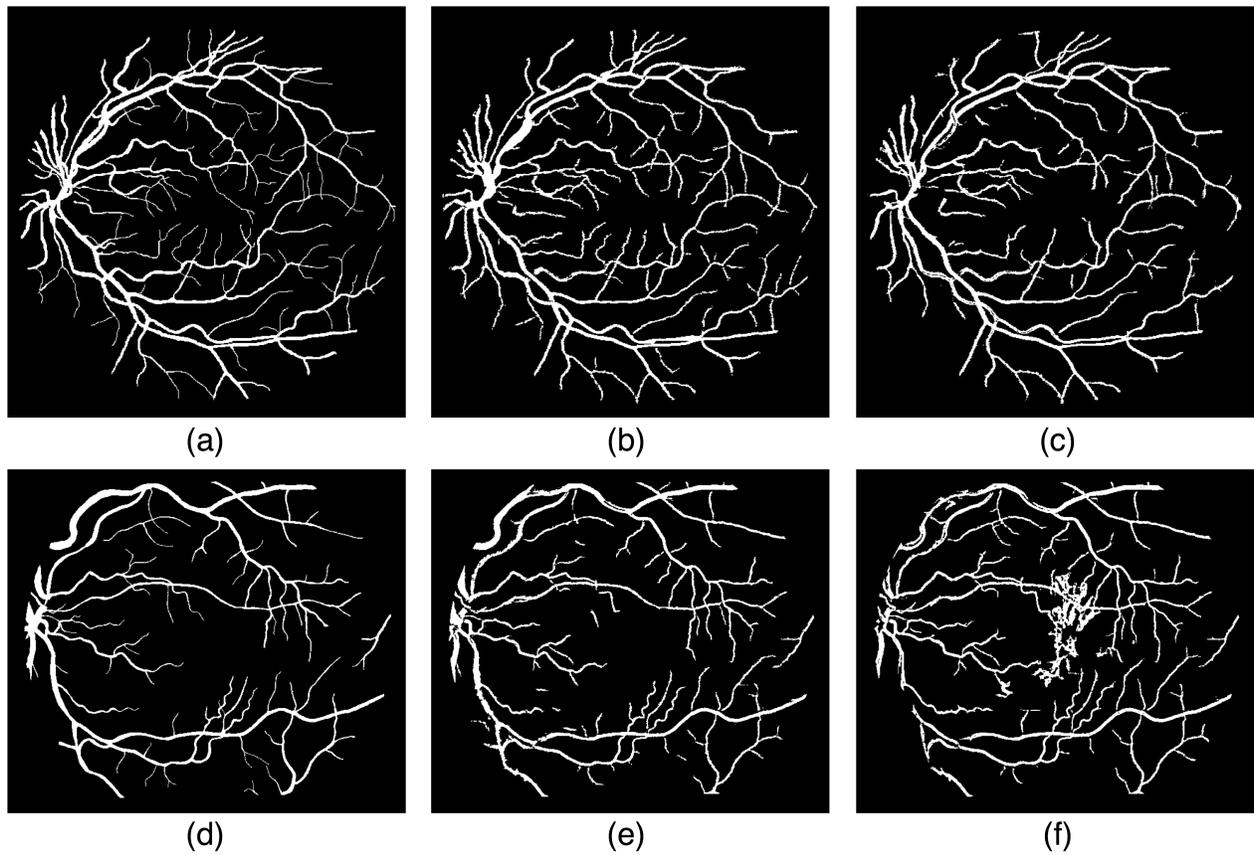


Figure 5.5: Comparison between single-database results and cross-database results of the image from DRIVE (first row) and STARE (second row) databases where (a), (d) are ground truth images; (b), (e) are single-database testing results; (c), (f) are cross-database testing results.

CHAPTER 6

CONCLUSION AND APOCALYPSE: WHY DEEP LEARNING IS THE ULTIMATE TOOL FOR MEDICAL IMAGE APPLICATIONS

6.1 Final review

This thesis has fully investigated and introduced the research background of the retinal blood vessel segmentation, including the motivation, the existing application, the publicly available databases, several image segmentation algorithms, and many existing implementations of retinal color image segmentation. Based on the investigation and understanding of the task, the thesis has developed and implemented two outstanding methods of retinal vessel segmentation.

The first implementation is a fast, accurate and robust retinal blood vessels segmentation system, which uses morphological processing technique to extract venules and applies matched filter algorithm to detect capillaries. These procedures are executed separately but simultaneously, in order to shorten the execution time while achieving a high accuracy.

The second implementation has proposed a supervised method to segment retinal blood vessel from the retinal color images with the help of the fully convolutional network and transfer learning. The proposed method has innovatively simplified and shifted a typical retinal vessel segmentation problem into regional semantic vessel element segmentation tasks, in this way the training data has been ideally augmented. Eventually, the cross-database results of the proposed method have outperformed almost all the other works in every aspect and prove to be state-of-the-art.

In general, by studying this thesis, the reader will be able to understand the importance of retinal blood vessels and how the proposed implementations will assist doctors and scientists to achieve higher goals.

6.2 Medical imaging, deep learning, and transfer learning

In medical imaging, the accurate diagnosis and assessment of a disease depend on both image acquisition and image interpretation. Image acquisition, over recent years, has improved magnificently. However, the interpretation of medical images has only recently begun to benefit from the development of computer technology. Most medical image interpretations are performed by doctors. But the expert interpretation is always expensive and rare, large variations across different interpreters. Many automated diagnostic tasks require an initial search process to detect abnormalities, and to quantify measurements and changes over time. Computerized tools, specifically image analysis and machine learning, are the key enablers to improve diagnosis, by facilitating identification of the findings that require treatment and to support the experts workflow. Among these tools, deep learning has developed rapidly and proved to be the state-of-the-art foundation, leading to accuracy and robustness [21].

Deep learning is an extremely effective tool for object recognition and localization in natural images. Medical image analysis researchers all over the world are quickly entering this field and applying deep convolutional neural networks and other deep learning methodologies to a wide variety of applications. Promising results are emerging, such as the one in this thesis. The outstanding performance of the second implementation has confirmed a popular question in the context of medical image analysis: Can the use of pre-trained deep convolutional neural networks, eliminate the need for training a deep convolutional neural network from scratch? Our extensive training and testing experiments, having been conducted for half a year, have demonstrated that through a proper transformation of the task, transfer learning and the pre-trained deep convolutional neural network are very useful and functional for medical image analysis.

Here is a funny analogy. Let's say the process of deep learning is like the way of training

a surgeon, who will be required to do large amounts of practice before holding a scalpel. The more opportunity a surgeon practice, the more accurate he will be. Transfer learning is like another mindset. Given the chances of practice are limited, another way to accelerating the surgeon education is to train a butcher.

Transfer learning is such an effective solution when there have insufficient publicly available ground truth and expert interpretations. The convolutional neural network models pre-trained from a different medical domain or even a natural image dataset can be used for a new medical task, which, in the proposed work, has been verified. Since there are more and more outstanding deep convolutional neural networks coming up, it can be predicted that the researchers will benefit from the gradually sophisticated use of transfer learning.

6.3 Future work

In the future, our first plan is the hardware implementation of the unsupervised method. As we mentioned previously in the introduction, in order to make eye examination more accessible to patients, the automated retinal vessel segmentation system is expected to be implemented on hardware, so as to be realized handy and portable in the future. Based on our experiments, Raspberry Pi is more suitable and realistic to be the implementation platform candidate. We have already implemented the proposed system on Raspberry Pi 3 Model B in python, and the current execution time takes 67.69 s per image (image from the DRIVE database) without losing accuracy. Next, we will continue to optimize the program and shorten the execution time on Raspberry Pi.

The second plan is to conduct more experiments on transfer learning and try different pre-trained models for retinal blood vessel applications, such as diabetic retinopathy. Since we have presented that the pre-trained model for semantic segmentation truly works for vessel segmentation, it will also worth trying to extract other retinal features such as lesion and degeneration.

REFERENCES

- [1] Canadian Ophthalmology Profile, Available at: <https://www.cma.ca/Assets/assets-library/document/en/advocacy/Ophthalmology-e.pdf>.
- [2] Convolutional Neural Networks (LeNet) DeepLearning 0.1 documentation. Available at: <http://deeplearning.net/tutorial/lenet.html>.
- [3] Deep learning - mit technology review. Available at: <https://www.technologyreview.com/s/513696/deep-learning/>.
- [4] The pascal visual object classes homepage. Available at: <http://host.robots.ox.ac.uk/pascal/VOC/>.
- [5] Retinal image analysis – the chase_db1 retinal image database. Available at: <https://blogs.kingston.ac.uk/retinal/chasedb1/>.
- [6] Vicavr database – varpa group. Available at: <http://www.varpa.es/research/ophtalmology.html>.
- [7] The stare project – structured analysis of the retina, January 2013. Available at: <http://cecas.clemson.edu/~ahoover/stare/>.
- [8] Ibrahim Abdurrazaq, Subhas Hati, and C. Eswaran. Morphology approach for features extraction in retinal images for diabetic retionopathy diagnosis. In *Proceedings of the International Conference on Computer and Communication Engineering 2008, ICCCE08: Global Links for Human Development*, pages 1373–1377, 2008.
- [9] Bashir Al-Diri, Andrew Hunter, David Steel, Maged Habib, Taghread Hudaib, and Simon Berry. Review – a reference data set for retinal vessel profiles. In *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 2262–2265. IEEE, August 2008.
- [10] Alauddin Bhuiyan, Akter Hussain, Ajmal Mian, Tien Y. Wong, Kotagiri Ramamohanarao, and Yogesan Kanagasingham. Biometric authentication system using retinal vessel pattern and geometric hashing. *IET Biometrics*, 6(2):79–88, March 2017.
- [11] S Chaudhuri, S Chatterjee, N Katz, M Nelson, and M Goldbaum. Detection of blood vessels in retinal images using two-dimensional matched filters. *IEEE transactions on medical imaging*, 8(3):263–269, 1989.

- [12] Ning Cheung, TienY Wong, and Lauren Hodgson. Retinal vascular changes as biomarkers of systemic cardiovascular diseases. In *Automated Image Detection of Retinal Pathology*. CRC Press, October 2009.
- [13] Alexandru Paul Condurache, Johannes Kotzerke, and Alfred Mertins. Robust retina-based person authentication using the sparse classifier. *European Signal Processing Conference*, (Eusipco):1514–1518, 2012.
- [14] Etienne Decencire, Xiwei Zhang, Guy Cazuguel, Bruno Lay, Batrice Cochener, Caroline Trone, Philippe Gain, Richard Ordonez, Pascale Massin, Ali Erginay, Batrice Charton, and Jean-Claude Klein. Feedback on a publicly distributed database: the messidor database. *Image Analysis & Stereology*, 33(3):231–234, August 2014.
- [15] Geoff Dougherty. *Medical Image Processing*. Springer Science+Business Media, 2011.
- [16] Donald S Fong, Lloyd P Aiello, Frederick L Ferris, and Ronald Klein. Diabetic retinopathy, October 2004.
- [17] M. M. Fraz, S. A. Barman, P. Remagnino, A. Hoppe, A. Basit, B. Uyyanonvara, A. R. Rudnicka, and C. G. Owen. An approach to localize the retinal blood vessels using bit planes and centerline detection. *Computer Methods and Programs in Biomedicine*, 108(2):600–616, 2012.
- [18] M.M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A.R. Rudnicka, C.G. Owen, and S.A. Barman. Blood vessel segmentation methodologies in retinal images a survey. *Computer Methods and Programs in Biomedicine*, 108(1):407–433, 2012.
- [19] Muhammad Moazam Fraz, Paolo Remagnino, Andreas Hoppe, Bunyarit Uyyanonvara, Alicja R. Rudnicka, Christopher G. Owen, and Sarah A. Barman. An ensemble classification-based approach applied to retinal blood vessel segmentation. *IEEE Transactions on Biomedical Engineering*, 59(9):2538–2548, 2012.
- [20] Huazhu Fu, Yanwu Xu, Damon Wing Kee Wong, and Jiang Liu. Retinal vessel segmentation via deep learning network and fully-connected conditional random fields. *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, pages 698–701, 2016.
- [21] H. Greenspan, B. van Ginneken, and R. M. Summers. Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique. *IEEE Transactions on Medical Imaging*, 35(5):1153–1159, 2016.
- [22] Adam Hoover, Valentina Kouznetsova, and Michael Goldbaum. Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. *IEEE Transactions on Medical Imaging*, 19(3):203–210, 2000.
- [23] Zhixin Jiang, Juan Yopez, Sen An, and Seokbum Ko. Fast, accurate and robust retinal vessel segmentation system. *Biocybernetics and Biomedical Engineering*, 37(3):412–421, 2017.

- [24] Sorri Iris Raninen Asta Voutilainen Raija Kamarainen Joni Lensu Lasse Kauppi Tomi, Kalesnykiene Valentina and Uusitalo Hannu. Diaretddb1 v2.1 – diabetic retinopathy database and evaluation protocol, May 2009. Available at: http://www.it.lut.fi/project/imageret/diaretddb1_v2_1/.
- [25] Thomas Kohler, Attila Budai, Martin F. Kraus, Jan Odstrcilik, Georg Michelson, and Joachim Hornegger. Automatic no-reference quality assessment for retinal fundus images using vessel segmentation. In *Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems*, pages 95–100, 2013.
- [26] Dimitris Koukounis, Christos Ttofis, Agathoklis Papadopoulos, and Theocharis Theocharides. A high performance hardware architecture for portable, low-power retinal vessel segmentation. *Integration, the VLSI Journal*, 47(3):377–386, 2014.
- [27] Michael Krause, Ralph Maria Alles, Bernhard Burgeth, and Joachim Weickert. Fast retinal vessel analysis. *Journal of Real-Time Image Processing*, 11(2):413–422, apr 2013.
- [28] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 25*, pages 1097–1105. Curran Associates, Inc., 2012.
- [29] Benson S. Y. Lam, Yongsheng Gao, and Alan Wee-Chung Liew. General retinal vessel segmentation using regularization-based multiconcavity modeling. *IEEE Transactions on Medical Imaging*, 29(7):1369–1381, 2010.
- [30] Yann LeCun, Yoshua Bengio, and Geoffrey Hinton. Deep learning. *Nature*, 521(7553):436–444, may 2015.
- [31] Qiaoliang Li, Bowei Feng, Linpei Xie, Ping Liang, Huisheng Zhang, and Tianfu Wang. A cross-modality learning approach for vessel segmentation in retinal images. *IEEE Transactions on Medical Imaging*, 35(1):109–118, 2016.
- [32] Pawel Liskowski and Krzysztof Krawiec. Segmenting retinal blood vessels with deep neural networks. *IEEE Transactions on Medical Imaging*, 35(11):2369–2380, 2016.
- [33] Carmen Alina Lupascu, Domenico Tegolo, and Emanuele Trucco. Fabc: Retinal vessel segmentation using adaboost. *IEEE Transactions on Information Technology in Biomedicine*, 14(5):1267–1274, 2010.
- [34] Diego Marín, Arturo Aquino, Manuel Emilio Gegúndez-Arias, and José Manuel 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*, 30(1):146–158, 2011.
- [35] Ana Maria Mendonca and Aurélio Campilho. Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction. *IEEE Transactions on Medical Imaging*, 25(9):1200–1213, 2006.

- [36] Meindert Niemeijer, Bram van Ginneken, Michael J Cree, Atsushi Mizutani, Gwenole Quellec, Clara I Sanchez, Bob Zhang, Roberto Hornero, Mathieu Lamard, Chisako Muramatsu, Xiangqian Wu, Guy Cazuguel, Jane You, AgustIn Mayo, Qin Li, Yuji Hatanaka, Beatrice Cochener, Christian Roux, Fakhri Karray, Marla Garcia, Hiroshi Fujita, and Michael D Abramoff. Retinopathy online challenge: Automatic detection of microaneurysms in digital color fundus photographs. *IEEE Transactions on Medical Imaging*, 29(1):185–195, January 2010.
- [37] Manjiri B. Patwari, Ramesh R. Manza, Yogesh M. Rajput, Manoj Saswade, and Neha Deshpande. Personal identification algorithm based on retinal blood vessels bifurcation. In *2014 International Conference on Intelligent Computing Applications*, pages 203–207. IEEE, March 2014.
- [38] Elisa Ricci and Renzo Perfetti. Retinal blood vessel segmentation using line operators and support vector classification. *IEEE Transactions on Medical Imaging*, 26(10):1357–1365, 2007.
- [39] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg, and Li Fei-Fei. Imagenet large scale visual recognition challenge. *International Journal of Computer Vision (IJCV)*, 115(3):211–252, 2015.
- [40] Evan Shelhamer, Jonathan Long, and Trevor Darrell. Fully convolutional networks for semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(4):640–651, 2017.
- [41] João V.B. Soares, Jorge J.G. Leandro, Roberto M. Cesar Jr., Herbert F. Jelinek, and Michael J. Cree. Retinal vessel segmentation using the 2-d gabor wavelet and supervised classification. *IEEE Transactions on Medical Imaging*, 25(9):1214–1222, sep 2006.
- [42] Michal Sofka and Charles V. Stewart. Retinal vessel centerline extraction using multiscale matched filters, confidence and edge measures. *IEEE Transactions on Medical Imaging*, 25(12):1531–1546, 2006.
- [43] Joes Staal, Michael D. Abramoff, Meindert Niemeijer, Max A. Viergever, and Bram van Ginneken. Ridge based vessel segmentation in color images of the retina. *IEEE Transactions on Medical Imaging*, 23(4):501–509, 2005.
- [44] Shuangling Wang, Yilong Yin, Guibao Cao, Benzhen Wei, Yuanjie Zheng, and Gongping Yang. Hierarchical retinal blood vessel segmentation based on feature and ensemble learning. *Neurocomputing*, 149(PB):708–717, 2015.
- [45] T Y Wong, Q Mohamed, R Klein, and D J Couper. Do retinopathy signs in non-diabetic individuals predict the subsequent risk of diabetes? *British Journal of Ophthalmology*, 90(3):301–303, March 2006.
- [46] Lili Xu and Shuqian Luo. A novel method for blood vessel detection from retinal images. *Biomedical engineering online*, 9:14, 2010.

- [47] Jiong Zhang, Behdad Dashtbozorg, Erik Bekkers, Josien P. W. Pluim, Remco Duits, and Bart M. ter Haar Romeny. Robust retinal vessel segmentation via locally adaptive derivative frames in orientation scores. *IEEE Transactions on Medical Imaging*, 35(12):2631–2644, dec 2016.