# PATTERN RECOGNITION AND MACHINE LEARNING FRAMEWORK FOR AUTOMATED ANALYSIS OF RETINAL IMAGES

Thesis

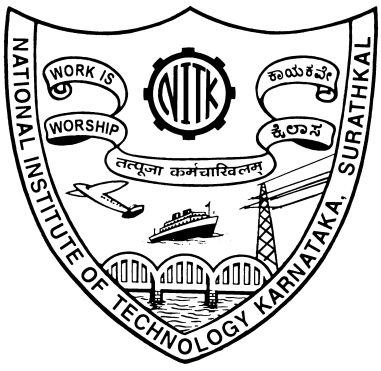
Submitted in partial fulfilment of the requirements for the degree of

## DOCTOR OF PHILOSOPHY

by

**CHETAN L SRINIDHI**

(135032EC13F01)




**Department of Electronics and Communication Engineering**

**National Institute of Technology Karnataka**

**Surathkal, Mangalore - 575025, India**

**June, 2019**

# DECLARATION

I hereby ***declare*** that the Research Thesis entitled **Pattern Recognition and Machine Learning Framework for Automated Analysis of Retinal Images** which is being submitted to the ***National Institute of Technology Karnataka, Surathkal*** in partial fulfillment of the requirements for the award of the Degree of ***Doctor of Philosophy*** is a ***bona fide report of the research work carried out by me***. The material contained in this thesis has not been submitted to any University or Institution for the award of any degree.

**Chetan L Srinidhi**

Register No.: 135032EC13F01

Department of Electronics and Communication Engineering

Place: NITK, Surathkal.

Date:

# CERTIFICATE

This is to ***certify*** that the Research Thesis entitled **Pattern Recognition and Machine Learning Framework for Automated Analysis of Retinal Images**, submitted by **Chetan L Srinidhi** (Register Number: 135032EC-13F01) as the record of the research work carried out by him, is *accepted as the Research Thesis submission* in partial fulfillment of the requirements for the award of degree of ***Doctor of Philosophy***.


**Dr. Aparna P**

Research Guide

Assistant Professor

Department of Electronics and Communication Engineering

NITK Surathkal - 575025

**Dr. Jeny Rajan**

Research Co-Guide

Assistant Professor

Department of Computer Science and Engineering

NITK Surathkal - 575025


**Chairman - DRPC**

(Signature with Date and Seal)

*To My Beloved Parents, Teachers and Friends*

# ACKNOWLEDGEMENTS

We are at the end of the line! Even though it seemed far away at the beginning, now I realize how fast the time travels. Fortunately, I enjoyed this journey very much, and I had the opportunity to find lots of friends, travel around, collaborate with brilliant international colleagues and learn many valuable life lessons. This achievement was not possible without the support of many friends, colleagues and my family. I would like to express my sincere appreciation to all these people.

First and foremost, I would like to express my sincere gratitude to Dr. Aparna P, my promoter and advisor. I really appreciate her in the first place, that she trusted and invited me to join NITK. I really could not achieve this without her endless support, trust, patience and honest feedback. I am extremely grateful for her, for the freedom which gave me always, to explore all different dimensions of research and personal life. I thank you, ma'am. Also, I express my sincere gratitude to co-research supervisor Dr. Jeny Rajan for his guidance and support throughout this research work.

Besides, my advisors, I wholeheartedly convey my sincere appreciation to my research collaborator and mentor Prof. Jayanthi Sivaswamy, Dean Academics, International Institute of Information Technology (IIIT)-Hyderabad. Her sheer joy and enthusiasm toward research motivated me throughout my PhD. Under her guidance, not only I got rigorous training research design and execution but also enjoyed the freedom to pursue my dreams and I will always to indebted to her for that. She has been an exceptional mentor and a very compassionate human being, which was apparent in her unconditional support during my difficult times of my life. No word can justify my sincere gratitude and respect for her. Thank you so much madam.

A special acknowledgement goes to Prof. T. Laxminidhi, Head, Department of Electronics and Communication Engineering, NITK, for the constant support and encouragement. My sincere gratitude to Prof. Muralidhar Kulkarni, Prof. U. Sripathi Acharya and Prof. M. S. Bhat for their valuable support and guidance in academic activities. My heartfelt thanks to Dr. Jidesh P, and Dr. Ramesh Kini (RPAC members) for their valuable and insightful comments at every phase of my Ph.D duration. I am extremely grateful to Prof. Uma Maheshwar Rao, Director, NITK, for providing the necessary infrastructure and resources to accomplish my research work. Also, I would like to extend my sincere thanks to the entire staff of both NITK and IIIT Hyderabad, specially Subrahmanya Karanth and Rateesh, who have helped me with many logistics matters and without them this wouldn't have been possible. The financial support provided by Ministry of Human Resource and Development, Govt. of India for my research work was indispensable and I am grateful.

During my Ph.D. I had great opportunities to visit and learn with many good researchers in several institutes. I sincerely appreciate my friends Karthik Gopinath

(Ph.D candidate @ ETS Montreal), Samrudhdhi Rangrej, Syed Tabish (Ph.D candidates @ McGill University), Priyadarshi Rath (MS student @ UMASS), Govinda, Sukanya, Jhanavi, Sukesh, Raghav (MS students @ IIITH) and finally Joyneel (B.Tech student @ IIITH). Indeed I have discussed and learnt a lot of research ideas from them, which has helped in many different ways. I also enjoyed all the laughter, funny chats and appreciate all your help.

I am lucky to have many nice and kind friends around me. I would like to thank them all, especially Praveen GB (Ph.D candidate @ BITS, Goa), Christian Chavarro-Nieto (Post-Doctoral Researcher, Harvard Medical School) and his wife Catherine for being supportive and helpful in many difficult situations of both research and personal life. I really thank all my close friends Shareef Babu, Raghavendra MANS, Jhnanesh Somayaji, Sarwesh, Yajunath, Jayaram, Abhishek, Sreedhar, Shankarnarayan, Punneth TR, Mahesh Kumar, Abdul, Bheem, Girish for being supportive in all phase of my research life. A special heartfelt appreciation to my wife Keerthi Shetty for all the unconditional support, care and love which you have provided throughout my research career. Thank you, Keerthi. Finally, I will never forget the many wonderful lunches, dinners and fun trips we've done together. Loving thanks to my friends.

Always the most important part, my deep gratitude goes to my father, mother, sister and my beloved four leg brothers - Leo as well as Puppy, for their unconditional love, support and encouragement in my life. I would dedicate this work to my father and mother, who have spent their entire life for the betterment of my well being. I have nothing to give them till now, but I am sure this work would make them proud.

Finally, I would like to extend my gratitude and sincere appreciation to my Ph.D examiners Prof. Rajbabu Velmurugan (Department of Electrical Engineering, IIT-Bombay) and Prof. Palaniappan Ramaswamy (School of Computing, University of Kent, UK) for their careful and insightful comments, which prompted us to look more carefully at several aspects of the thesis. The detailed feedback has been very helpful in improving the quality of the thesis. Further, I would also like to thank the DTAC members Prof. Muralidhar Kulkarni (DRPC, Chairman), Prof. Annappa and Dr Rekha Bhat for their kind support and valuable suggestions.

# ABSTRACT

The retina is one of the few locations in the human body that allows direct non-invasive visualization of its anatomical components. A comprehensive analysis of retinal microvasculature structures provides potential clinical biomarkers towards early diagnosis and prognosis of systemic and neurodegenerative diseases. The research focus of this thesis is to develop a series of novel pattern recognition and machine algorithms for automated analysis of retinal vasculature which includes - segmentation of vascular tree, classification of vessels into artery/vein and identification of vessel bifurcation and crossover points. Besides, several geometrical properties at crossover points are analysed to study and quantify the influence of various systemic diseases.

Accurate segmentation of retinal vessels is challenging due to the varying nature of vessel calibre, the proximal presence of pathological lesions, strong central vessel reflex and relatively low contrast images. Most existing methods mainly rely on carefully designed hand-crafted features to model the local geometrical appearance of vasculature structures, which often lacks the discriminative capability in segmenting vessels from a noisy and cluttered background. To address this issue, a novel visual attention guided unsupervised feature learning (VA-UFL) approach is proposed to automatically learn the most discriminative features, without complex domain expertise. The VA-UFL approach inherits the combined knowledge of visual attention mechanism and multi-scale contextual information to selectively visualize the most relevant part of the structure in a given local patch. The experiment results show that the proposed approach is shown to be robust to segmentation of thin vessels, strong central vessel reflex, complex crossover structures and fares well on abnormal cases. Further, the discriminative features learned via visual attention mechanism is superior to handcrafted features, and it is easily adaptable to various kind of datasets, where generous training images are often scarce.

Detection and classification of vessel junctions are extremely challenging due to spatially varying nature of vessel calibre, which often results in a very close appearance of false bifurcation or crossover points. Existing approaches model the orientation of vessels in a local neighbourhood, without explicitly considering the vessel shape information, which might aid in resolving ambiguities. To address this problem, a novel vessel keypoint descriptor (VKD) is proposed, which is derived from

the projection of log-polar transformed binary patches. The VKD along with shape based features aids in accurate localization of junctions and classifying them into bifurcations/crossovers. Evaluation results on five challenging datasets show that the designed system is robust to changes in resolution and other variations across datasets.

Several geometrical properties at crossover points are analysed to detect and quantify the morphological changes linked to hypertension, stroke and other systemic diseases. To this end, a complete system for detecting arteriovenous (AV) nicking is presented. The entire system is solely based on analysis of vessel morphology, which requires no knowledge of artery/vein class label for vessel segments. Both local orientation and width of vessels are estimated with the aid of VKD, to detect and quantify the vascular changes at crossover points for identifying AV nicking. The proposed solution indicate that the crossover geometrical properties can be considered as essential biomarkers in assessing the progression of various systemic diseases.

Separation of arteries and veins is a fundamental prerequisite in the automatic detection of vessel-specific biomarkers associated with systemic and neurodegenerative diseases. In this thesis, a novel graph search metaheuristic approach is proposed for the automatic separation of arteries/veins (A/V) from color fundus images. The proposed method exploits local information to disentangle the complex vascular tree into multiple subtrees, and global information to label these vessel subtrees into arteries and veins. Based on the anatomical uniqueness at vessel crossing and branching points, the vascular tree is split into multiple subtrees containing arteries and veins. Further, the identified vessel subtrees are labelled with A/V based on a set of handcrafted features trained with random forest classifier. The experimental results demonstrate the superiority of the proposed approach in outperforming state-of-the-art methods for A/V separation.

**Keywords :** Retinal Image, Vessel Segmentation, Visual Attention, Unsupervised Feature Learning, Arteriovenous Nicking, Vessel Keypoints, Keypoint Descriptor, Vessel Width, Artery/Vein Classification, Graph Traversal, Depth-First Search.

# Contents

# List of Figures

# List of Tables

# ABBREVIATIONS

| | |
|---|---|
| AVN | Arteriovenous Nicking |
| A/V | Artery/Vein |
| Acc | Accuracy |
| AUC | Area Under the Curve |
| AVR | Arteriolar-to-Venular Diameter Ratio |
| CVR | Central Vessel Reflex |
| CAD | Computer-Aided Diagnosis |
| DFS | Depth-First Search |
| DR | Diabetic Retinopathy |
| GT | Ground Truth |
| MCC | Mathews Correlation Coefficient |
| RF | Random Forest |
| ROI | Region of Interest |
| RGB | Red Green Blue |
| SLO | Scanning Laser Ophthalmoscope |
| Se | Sensitivity |
| Sp | Specificity |
| t-SNE | t-distributed Stochastic Neighbor Embedding |
| UFL | Unsupervised Feature Learning |
| VA-UFL | Visual Attention Guided Unsupervised Feature Learning |
| VKD | Vessel Keypoint Descriptor |

# Chapter 1

# Introduction

Retinal diseases are of the most significant public health concern in the working and aged population worldwide. For instance, Diabetic Retinopathy (DR), Glaucoma and Age-related Macular Degeneration (AMD) are the leading causes of blindness in the ageing population (Tham et al. (2014); Wong et al. (2014); Yau et al. (2012)). In a recent study (Yau et al. (2012)), it is estimated that there are 93 million people with DR, followed by AMD which accounts for 7-8 % of total blindness worldwide (Wong et al. (2014)). Another major leading cause of blindness is the Glaucoma, characterized by progressive damage to the optic nerve. It has been projected that there are about 64.3 million people with Glaucoma (Tham et al. (2014)). All these retinal diseases are likely to increase by three folds as a consequence of exponential ageing population, diabetes, lifestyle changes, and other risk factors. Such statistics naturally drives a considerable amount of research dedicated to developing computer-aided diagnostic (CAD) tool for the automated diagnosis of retinal pathologies, mainly for low and middle-income countries.

Retinal microcirculation offers a unique non-invasive way to study the early manifestation of several diseases affecting the human circulatory system. Changes in retinal vascular geometrical patterns such as width, tortuosity, branching angle, junction exponents and fractal dimension have been investigated as candidate biomarkers in various ocular, systemic and neurodegenerative diseases (Abramoff et al. (2010); Cheung et al. (2017); Heringa et al. (2013); McGrory et al. (2017)). Data from long-term population-based studies have demonstrated a consistent link between the retinal microvascular changes with incident clinical stroke (Wong et al. (2001b)), hypertension (Cheung et al. (2012)), arteriosclerosis (Hubbard et al. (1999)), dementia (Frost et al. (2013)), and other cerebral small vessel diseases (Wong et al. (2002)).

For instance, the narrowing of arteries and widening of veins is a significant indicator of the progression of diabetic retinopathy (Abramoff et al. (2010)), hypertension (Hubbard et al. (1999)), and various other cardiovascular abnormalities (Wong et al. (2006)). Specifically, the arteriolar-to-venular diameter ratio (AVR) is a prognostic indicator of stroke, cerebral atrophy, cognitive decline and myocardial infarction (Niemeijer et al. (2011)). Therefore, an accurate analysis and quantification of vessel specific morphological changes may provide an early insight into better understanding the pathophysiology of the disease conditions.

The retinal fundus photography is an excellent non-invasive technique most commonly used to analyse and quantify the vascular abnormalities in large-scale clinical settings, due to its speed and affordability. Manual analysis of retinal color fundus image is extremely time-consuming and requires an enormous amount of painstaking manual process. Hence, developing a computer aided diagnostic tool is of paramount importance in large-scale retinal disease screening programs.

## 1.1 Retinal Imaging

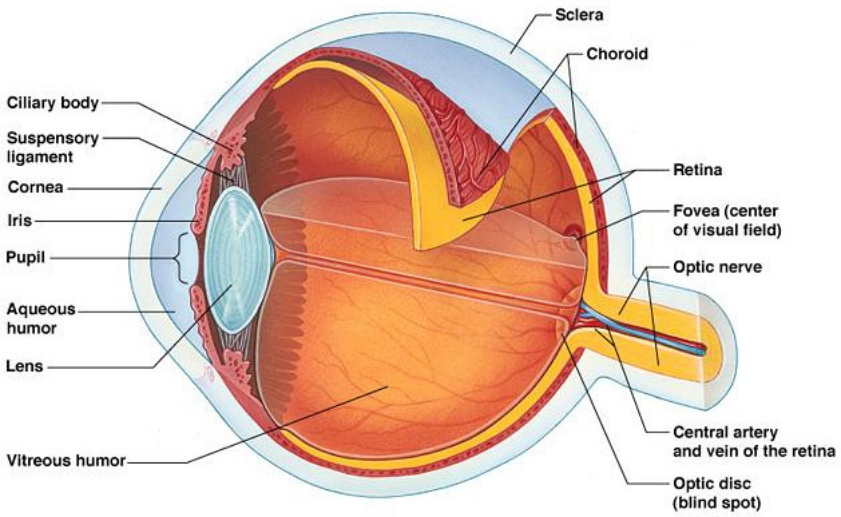



Figure 1.1: Anatomy of the human eye (Hogan (1971)).

Imaging the retina is of prime importance for early diagnosis, monitoring disease progression, treatment and surgery planning of various ocular diseases. The retina is the most important part of the eye, which is formed by a thin layer of photosensitive neural cells, which lies at the back (i.e. the fundus) of the eye, as illustrated in Figure 1.1. The light waves from the external world pass through the vitreous and projects directly on the retina, which are then converted to electric signals and

transmitted to the brain via the optic nerves. The image of the retina is acquired using a specialized microscope attached with a camera called the fundus camera. The fundus is the interior surface of the eye including the retina, optic disc, macula, and fovea. A typical human fundus photograph captured from the fundus camera is shown in Figure 1.2. The microstructures present in the retina are the optic disc, macula and blood vessels such as arteries and veins.

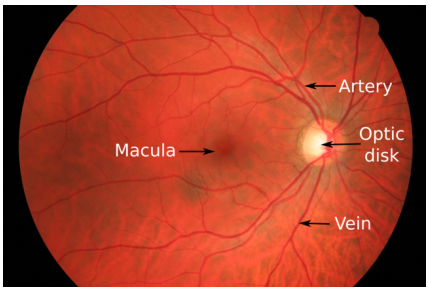
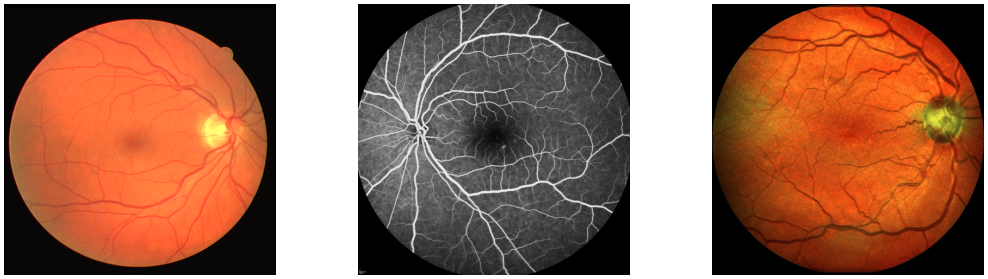


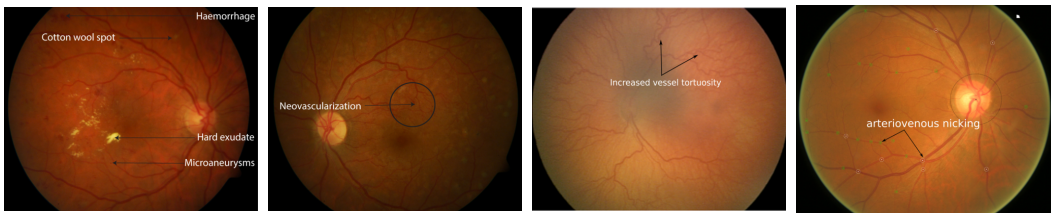

Figure 1.2: A color fundus photograph of the retina (Odstrcilik et al. (2013)).

The fundus photography is obtained by a projection of 3-D retinal tissue onto the 2-D imaging plane using reflected light. Among the different imaging modalities shown in Figure 1.3, the color fundus photography, fluorescein angiography (FA) and scanning laser ophthalmoscope (SLO) images are the most widely used imaging techniques for analysis of retinal images. Fluorescein angiography is a process of imaging vascular flow within the retina and surrounding tissue by injecting a fluorescent dye into the blood stream. This dye fluoresces a different color when light from a particular wavelength reaches it. Compared to fundus imaging, SLO imaging provides very high contrast and finely detailed low-resolution images of size 1K-1.5K pixels.



(a) (b) (c)

Figure 1.3: Retinal imaging modalities. (a) The color fundus photography from DRIVE dataset (Staal et al. (2004)); (b) Fluorescein Angiography (FA) (Alipour et al. (2014)); (c) Scanning Laser Ophthalmoscope (SLO) from IOSTAR dataset (Zhang et al. (2016)).




(a) (b) (c) (d)

Figure 1.4: (a) Common signs of DR (Diaretdb1 (2007)); (b) Common sign of Proliferative DR (PDR) (Diaretdb1 (2007)); (c) Common sign of ROP (Oloumi et al. (2015)), with kind permission from Elsevier; (d) Arteriovenous nicking (Nguyen et al. (2013a)).

## 1.2 Clinical Significance of Retinal Images and Its Challenges

The retina is the only part of the human body that allows direct non-invasive visualization of its anatomical components. Many large-scale population-based studies (Doubal et al. (2009); Smith et al. (2004); Wong and Mitchell (2007); Wong et al. (2001b, 2003)) have been conducted to find statistical correlations between a disease and structural changes in the vascular system of the retina. Such longitudinal clinical studies have shown strong and consistent relationship between changes in retinal microvasculature with number of conditions like hypertension (Sharrett et al. (1999); Tso and Jampol (1982)), stroke (Doubal et al. (2009); Wong et al. (2001a)) and other cardiovascular diseases (Wong et al. (2001b)). Hence, a comprehensive analysis of retinal vasculature is a key task in the automated early detection of these pathologies.

The key motivations which necessitate the analysis of retinal images and its vasculature are as follows:

- One of the significant landmark study conducted in Yau et al. (2012) has shown a strong link between DR and structural changes in retinal vasculature. The earliest sign of DR is the presence of tiny capillary dilations called microaneurysms, usually appearing near thin vessels as shown in Figure 1.4 (a). Although the segmentation of vasculature has no direct role in the assessment of these pathologies, it is necessary, as pathological structures (such as, microaneurysm and exudates) have similar visual features as high-curvature and junction points of thin vessels (Youssef and Solouma (2012)). The knowledge

of shape and characteristic of the retinal vasculature can also aid in accurate localization of other degenerative anatomical structures like the optic disc or macula (Hoover and Goldbaum (2003)).

- The Proliferative Diabetic Retinopathy (PDR) (Welikala et al. (2014)) is an advanced stage of DR which is primarily characterized by abnormal growth of new blood vessels, which is termed as neovascularization. The abnormal growth of new vessels is triggered mainly to compensate for the damaged blood vessels caused by DR. The new blood vessels appear in loopy structures mostly near the optic disc region or the veins, as shown in Figure 1.4 (b). Since conventional vessel segmentation algorithms consider a vessel as a thin elongated smooth linearly varying structure, it often fails to detect complex structures like neovascularization (see Figure 1.4 (b)).

- The recent study in Cheung et al. (2011a) has shown a link between Retinopathy of Prematurity (ROP) and the temporal changes of retinal vessel width and tortuosity. One of the earliest changes visible in vessel morphology is the increased vessel tortuosity (Grisan et al. (2008); Sodi et al. (2012)), as depicted in Figure 1.4 (c). The increase in vessel tortuosity has shown a link with the progression of ROP, hypertensive retinopathy (Sutter and Helbig (2003)), and some rarer hereditary retinopathies.

- The ratio of arteries to veins width (A/V ratio), arteriovenous (AV) nicking as depicted in Figure 1.4 (d) have shown to be associated with hypertension, cardiovascular and other systemic diseases (Nguyen et al. (2013a); Wong and Mitchell (2007); Wong et al. (2001a,b)). The Blue Mountains eye population-based study (Smith et al. (2004)) proved the hypothesis that small vessel structural changes may precede the development of severe hypertension. The analysis of such retinal microvasculature requires precise tools to extract the vessel tree, quantify the morphological changes and evaluate the condition of the patients.

Many automated retinal image analysis approaches have been proposed in the literature. Although most of the existing methods have been successful in achieving performance close to the trained human observers, there still exists many significant challenges in clinical scenarios. The following are some of the current challenges related to automated retinal image analysis.

- During acquisition of retinal fundus images, different 3D retinal structures with varying depth are projected onto a 2D image with maximum intensity. This leads to the overlapping of non-vascular structures and decrease in the visibility of thin vessels in low-contrast imaging. Due to the use of low quality fundus camera, various imaging artefacts such as blur, noise, uneven illumination and inter-camera variability may also be introduced.

- The majority of the existing approaches have mainly focused on segmenting large vessel structures. Accurate delineation of thin vessel structures is still very challenging and an open problem. Since large and small vessels differ in their size, shape, and contrast, they should be tackled differently, because applying the same technique might tend to over segment the other.

- The presence of abnormalities such as exudates, haemorrhages, cotton wool spots and microaneurysms, structures with strong central vessel reflex, bifurcations/crossover regions, close parallel and highly curved vessels pose a significant challenge for accurate analysis and quantification of vascular structures.

## 1.3 Objectives of the Thesis

The main objectives of this thesis is to develop novel image analysis solutions for comprehensive assessment of retinal vascular network, which are listed as follows:

1. To propose a novel fully automated retinal vessel segmentation method that can effectively deal with uneven illumination, segmentation of thin vessels, strong central vessel reflex, complex crossover structures and segmentation in the presence of abnormalities.

2. To propose a novel end-to-end junction detection and junction classification system, followed by analysis and quantification of crossover biomarkers in retinal images.

3. To propose a novel automated method for separation of arteries and veins from retinal color fundus images.

## 1.4 Contributions

The main focus of this thesis is to develop novel automated image analysis solutions for the comprehensive assessment of vascular structures in retinal images. The proposed approaches are mainly inspired by the human visual perception, followed by analysis of various unique geometrical properties of retinal vascular structures such as orientation, scale, curvature, width and color information. The proposed framework play a vital role in developing a multifactorial decision support system by integrating the findings of various analysis components into automated retinal pre-screening solutions. The main contributions of this thesis can be summarized as follows:

- A comparative study of the recent retinal vessel segmentation methods is presented with the following objectives. First, the most crucial preprocessing steps that are involved in the accurate segmentation of vessels are discussed in detail. Second, the state-of-the-art segmentation techniques are reviewed, which are classified into different categories based on their main principle. Third, the quantitative analysis of these methods is carried out in terms of its sensitivity, specificity, accuracy and area under the curve, along with the newly introduced performance metrics in the current literature. Fourth, the advantages and limitations of the existing approaches are discussed thoroughly. Finally, an insight into existing problems and possible future directions are provided for developing robust segmentation techniques.

- A novel visual attention guided unsupervised feature learning (VA-UFL) approach is proposed to automatically learn the most discriminative features for segmenting vessels in retinal images. The VA-UFL approach inherits the knowledge of both biologically inspired visual attention mechanism and multi-scale contextual information under a single framework. The proposed method selectively pay attention to the most relevant part of the structure in a given local patch, by encoding rich hierarchical information into the unsupervised filtering learning process to generate a set of discriminative features. This learned feature set underscores the interclass differences (between the vessel and background pixels) to be large, while keeping intra-class differences (between vessel pixels) to be small, resulting in the accurate segmentation of

vessels even in the presence of cluttered background. The proposed scheme is validated on five publicly available retinal datasets, including both the RGB and SLO images. The extensive experimental analysis demonstrates the effectiveness of the proposed approach in handling all the challenging cases, compared with the state-of-the-art methods.

- A novel vessel keypoint descriptor (VKD) is introduced, which is derived from the projection of log-polar transformed binary patches around vessel points. VKD is used to design a two-stage solution for junction detection and classification. In the first stage, the keypoints detected using VKD are refined using curvature orientation information to extract candidate junctions. True junctions from these candidates are identified in a supervised manner using a random forest classifier. In the next stage, a novel combination of local orientation and shape based features is extracted from the junction points and classified using a second random forest classifier. Evaluation results on five datasets indicate that the designed system is robust to changes in image resolution and other variations across datasets. Besides, several geometrical properties at crossover points are analysed further to detect and quantify the morphological changes linked to hypertension, stroke and other systemic diseases. The proposed solution indicate that the crossover properties can be considered as crucial biomarkers in assessing the progression of various microvascular diseases.

- A novel graph search metaheuristic approach is proposed for the automatic separation of arteries/veins (A/V) from color fundus images. The method exploits local information to disentangle the complex vascular tree into multiple subtrees, and global information to label these vessel subtrees into arteries and veins. From a given binary vessel map, a graph representation of the vascular network is constructed representing the topological and spatial connectivity of the vascular network. Based on the anatomical uniqueness at vessel crossing and branching points, the vascular tree is split into multiple subtrees containing arteries and veins. Finally, the identified vessel subtrees are labeled with A/V based on a set of hand-crafted features trained with random forest classifier. The proposed method has been tested on four different publicly available retinal datasets, while outperforming the state-of-the-art approaches.

## 1.5 Outline

This thesis is organized into six chapters as follows.

Chapter 2 presents a comparative study of the recent segmentation techniques proposed for the automatic segmentation of vessels in retinal color fundus images. This chapter is divided into 4 main parts: (1) image pre-processing techniques; (2) different categories of segmentation methods; (3) summary of the recent state-of-the-art segmentation methods, with a focus on existing segmentation challenges; (4) a brief overview of the publicly available retinal image datasets, followed by the validation measures employed and thorough assessment of segmentation methods; and (5) a general discussion followed by the directions for future research in developing robust segmentation methods.

Chapter 3 presents an unsupervised feature learning approach for segmenting retinal vessels by incorporating the idea of biologically inspired visual attention mechanism into unsupervised filter learning technique. An extensive experimental validation is carried out on five publicly available retinal datasets, demonstrating the robustness of the proposed approach in handling challenging segmentation cases compared with the state-of-the-art methods.

Chapter 4 introduces a novel vessel keypoint descriptor, which is used to design a two-stage solution for junction detection and classification. This descriptor is subsequently used in Chapter 5, for identifying and analysing vessel keypoints in retinal images. The next part of the chapter presents an analysis of several geometrical properties at vessel crossover points to detect and quantify the morphological changes linked to several systemic diseases.

Chapter 5 presents a novel formulation of graph search metaheuristic method for the automatic identification of artery/vein (A/V) from retinal images. Various anatomical uniqueness at vessel keypoints is exploited to divide the vascular tree into multiple subtrees containing arteries and veins. The identified vessel subtrees are labeled with A/V based on a set of hand-crafted features trained with random forest classifier.

Chapter 6 concludes this thesis by providing a general summary of the presented research work and discuss ideas for future work to realize automated retinal prescreening solutions.

# Chapter 2

# Comparative Study of Recent Retinal Vessel Segmentation Methods

Retinal vessel segmentation is a crucial step towards the accurate visualization, diagnosis, early treatment and surgery planning of ocular diseases. For the last two decades, a tremendous amount of research has been dedicated to developing automated methods for segmenting vessels from retinal color fundus images. In this chapter, a systematic review of the most recent retinal vessel segmentation methods has been carried out. The objectives of this study are as follows: First, the most crucial preprocessing steps that are involved in the accurate segmentation of vessels are discussed thoroughly. Second, the recent state-of-the-art segmentation techniques are reviewed, which are classified into different categories based on their main principle. Third, the quantitative analysis of these methods is carried out in terms of its sensitivity, specificity, accuracy and area under the curve, along with the newly introduced performance metrics in the current literature. Fourth, the advantages and limitations of the existing methods are discussed in detail. Finally, an insight into existing problems and possible future directions are provided for building a successful automated retinal image analysis system.

## 2.1 Background

Retinal vessel segmentation is a fundamental step in the accurate visualization and quantification of various retinal pathologies. Changes in vascular morphology such as shape, tortuosity, branching pattern and width provide accurate early detection of many retinal diseases (Abramoff et al. (2010)). Clinically, retinal fundus images are often routinely acquired for mass screening of various abnormalities. Manual segmentation of vessel structures from these fundus images is often tedious, time-consuming and error-prone, especially for large population screening. Therefore, there is a need for the CAD tool which can reduce the number of manual operators with an increase in speed, accuracy and reproducibility mainly for large population-based screening programs.

Over the past two decades, a tremendous amount of research has been devoted to segmenting the vessel structures from retinal color fundus images. Despite the fact, segmentation of retinal vessels remains a challenging task, due to the presence of abnormalities, varying size and shape of the vessels, non-uniform illumination and anatomical variability between subjects. Numerous fully automated, semi-automated methods have been reported in the literature which was quite successful in achieving segmentation accuracy on par with trained human annotators. Despite this, there is considerable scope for further improvements due to the various challenges posed by the complex nature of vascular structures. Some of the existing problems include segmentation in the presence of abnormalities, segmentation of thin vessel structures and segmentation near the bifurcation and crossover regions.

In this chapter, retinal vessel segmentation methods published in the recent years (2012-2017) are considered for the review. This study provides a comprehensive insight on latest developments that has progressed over recent years, highlighting the key design challenges in the development of new methods. The advancements in technological innovations in imaging modalities (like, Fluorescence Angiography (FA), Scanning Laser Ophthalmoscope (SLO), Optical Coherence Tomography (OCT)) along with the combination of adaptive optics to FA, SLO and OCT has provided both spectacular spatial resolution as well as the depth information for clear visualization of anatomical components. Besides, standard acquisition protocol across different imaging devices have helped the research community to explore and validate their methods on more general and diverse datasets.

This chapter is structured as follows. In Section 2.2, an overview of the articles published in the recent years, along with their selection criteria are discussed in detail. The image pre-processing steps employed before vessel segmentation methods are detailed in Section 2.3. The different categories of segmentation methods are described in detail in Section 2.4, along with their advantages and limitations. Subsequently, the summary of the recent state-of-the-art segmentation methods with a focus on existing segmentation challenges is presented in Section 2.5. A brief overview of the publicly available retinal image datasets, followed by the validation measures employed and thorough assessment of segmentation methods is presented in Section 2.6. A general discussion on recapitulating the main points and future trends in retinal vessel segmentation are discussed in Section 2.7, followed by chapter summary in Section 2.8.

## 2.2 Methodology of the Review

In this study, a comprehensive literature search has been performed on the recent retinal vessel segmentation techniques published in the recent years (2012-2017). A total of 56 peer-reviewed articles are listed in Table 2.1, which were selected from the Google Scholar, PubMed and the Web of Science databases. The selection criteria include articles from peer-reviewed journals and conferences related to the retinal vessel segmentation methods.

The *acronyms* for the algorithms in Table 2.1 are as follows: Gradient orientation analysis (GOA), Morphological transformation (MT), Line strength measure (LSM), Gabor filter response (GFR), Decision trees (DT), Hysteresis binary classifier (HBC), Modified line operator (MLO), Support vector machines (SVM), Gray level co-occurence matrix (GLCM), Neural network (NN), Local binary pattern (LBP), Dual Gaussian (DG), Second derivative of Gaussian (SDG), Multi-scale matched filtering (MSMF), Shape and region features (SRF), Conditional random field (CRF), Structured output support vector machine (SOSVM), Gray voting (GV), Gabor wavelet (GW), Gaussian mixture model (GMM), First and second order features (FSOF), Multi-scale texton dictionary (MTD), Convolutional neural networks (CNN), Random forest (RF), Lattice neural network with dendritic processing (LNNDP), Deep autoencoders (DAE), Deep convolutional neural network (DCNN), Particle filtering (PF), Nearest neighbor search (NNS), Bar - Combination of Shifted Filter Responses (B-COSFIRE), Matched filtering with multi-wavelet

kernels (MFMK), Adaptive thresholding (AT), Local radon transform (LRT), Generalized Gabor function (GGF), Contour reconstruction (CR), Curvelet transform (CT), Matched filtering (MF), Laplacian of Gaussian filter (LoG), Locally adaptive derivative on orientation score (LAD-OS), Hessian matrix (HM), Local entropy thresholding (LET), Medialness function (MEF), Non-local means filtering (NMF), Multi-scale line detection (MSLD), Complex continuous wavelet transform (CCWT), Length filtering (LF), Hysteresis thresholding (HT), Percentile thresholding (PT), Neighbourhood estimator before filling (NEBF), Anisotropic diffusion filtering (ADF), Orientation aware detector (OAD), Path opening filter (POF), Tensor voting framework (TVF), Maximum a posteriori estimation (MAP), Transductive inference (TI), Orientation scores (OS), First order derivative of Gaussian (FDOG), Morphological top-hat transform (MTHO), Bit plane slicing (BPS), Region growing (RG), Fuzzy classification (FC), Morphological component analysis (MCA), Morlet wavelet transform (MWT), Adaptive global thresholding (AGT), Phase congruency (PC), Contrast limited adaptive histogram equalization (CLAHE), Bayesian model with spatial constraint (BMSC), Graph cut (GC), Level sets (LS), Retinex-based inhomogeneity correction (RIC), Local phase-based enhancement (LP), Infinite perimeter active contour (IPAC), Directional response vector map (DRVM), Bottom-hat transform (BHT), Bee colony swarm optimization (BSO), Fuzzy c-means (FCM), Pattern search optimization (PSO), Direction map (DM), Brain inspired multi-scales and multi-orientations (BIMSO).

## 2.3 Pre-Preprocessing

Image pre-processing is an essential prerequisite for accurate segmentation of retinal vessels. Vessel segmentation is quite difficult due to various imaging conditions such as noise, intensity inhomogeneity, poor visibility of thin vessel structures, anatomical variations, and other imaging artefacts. These artefacts are often inherited mainly from the image acquisition process (because of low-quality image acquisition devices). Thus, before applying any vessel segmentation techniques, few preprocessing steps are employed to improve the segmentation accuracy. In general, there are three pre-processing steps involved before the vessel segmentation.

1. Removal of central vessel reflex.

2. Intensity inhomogeneity correction.

Table 2.1: Summary of the vessel segmentation approaches presented in 56 papers considered in this study.

| | Authors | Methods | Datasets |
|---|---|---|---|
| | | **Supervised** | |
| | Fraz et al. (2012b) | GOA + MT + LSM + GFR + DT | DRIVE, STARE, CHASE_DB1 |
| | Condurache and Mertins (2012) | HBC | DRIVE, STARE |
| | Fraz et al. (2014) | DG + SDG + GFR + LSM + MT + DT | CHASE_DB1 |
| | Welikala et al. (2014) | MLO + SVM | MESSIDOR |
| | Rahebi and Hardalaç (2014) | GLCM + NN | DRIVE, STARE |
| | Fathi and Naghsh-Nilchi (2014) | LBP + NN | DRIVE, STARE |
| | Ganjee et al. (2014) | MSMF + SRF | STARE |
| | Orlando et al. (2017) | CRF + SOSVM | DRIVE, STARE, CHASE_DB1, HRF |
| | Wang et al. (2015) | CNN + RF | DRIVE, STARE |
| | Vega et al. (2015) | LNNDP | DRIVE, STARE |
| | Dai et al. (2015) | GV+ GW + GMM | DRIVE, STARE |
| | Roychowdhury et al. (2015a) | FSOF + GMM | DRIVE, STARE, CHASE_DB1 |
| | Zhang et al. (2015) | MTD + NN | DRIVE |
| | Li et al. (2016) | DAE | DRIVE, STARE, CHASE_DB1 |
| | Liskowski and Krawiec (2016) | DCNN | DRIVE, STARE, CHASE_DB1 |
| | Maninis et al. (2016) | DCNN | DRIVE, STARE |
| | Wu et al. (2016) | DCNN + PF + NNS | DRIVE |
| | Fu et al. (2016b) | CNN + CRF | DRIVE, STARE |
| | Strisciuglio et al. (2016) | B-COSFIRE + SVM | DRIVE, STARE |
| | Abbasi-Sureshjani et al. (2015) | BIMSO | DRIVE, IOSTAR |
| | | **Unsupervised** | |
| **Matched filtering** | Wang et al. (2013) | MFMK + AT | DRIVE, STARE |
| | Krause et al. (2016) | LRT | DRIVE |
| | Azzopardi et al. (2015) | B-COSFIRE | DRIVE, STARE, CHASE_DB1 |
| | Kovács and Hajdu (2016) | GGF + CR | DRIVE, STARE, HRF |
| | Kar and Maity (2016) | CT + MF + LoG | DRIVE, STARE, DIARETDB1 |
| | Zhang et al. (2016) | LAD-OS | DRIVE, STARE, CHASE_DB1, HRF, IOSTAR, RC-SLO |
| **Multi-scale approach** | Yu et al. (2012) | HM + LET | DRIVE, STARE, HRF |
| | Moghimirad et al. (2012) | MEF + HM | DRIVE, STARE |
| | Budai et al. (2013) | HM + HT | DRIVE, STARE |
| | Nguyen et al. (2013b) | MSLD | DRIVE, STARE, REVIEW |
| | Fathi and Naghsh-Nilchi (2013) | CCWT + AT + LF | DRIVE, STARE |
| | Zheng et al. (2013) | HM + NMF | DRIVE |
| | Annunziata et al. (2016) | NEBF + HM + PT | STARE, HRF |
| | Abdallah et al. (2015) | ADF + HM | DRIVE, STARE |
| | Yin et al. (2015) | OAD + GW + POF | DRIVE. STARE |
| | Christodoulidis et al. (2016) | MSLD + TVF | HRF |
| **Vessel tracking** | Yin et al. (2012) | MAP | REVIEW |
| | Yin et al. (2013) | MAP | DRIVE, STARE, REVIEW |
| | Zhang et al. (2014) | MAP + MSLD | REVIEW |
| | De et al. (2014) | TI | DRIVE |
| | Bekkers et al. (2014) | OS | HRF, REVIEW |
| **Mathematical morphology** | Fraz et al. (2012b) | FDOG + MTHT + BPS | DRIVE, STARE, MESSIDOR |
| | Fraz et al. (2013) | FDOG + MTHT + RG | DRIVE, STARE |
| | Sigurosson et al. (2014) | POF + FC | DRIVE |
| | Imani et al. (2015) | MCA | DRIVE, STARE |
| **Thresholding based approach** | Roychowdhury et al. (2015b) | AGT | DRIVE, STARE, CHASE_DB1 |
| | Mapayi et al. (2015a) | AT + GLCM | DRIVE, STARE |
| | Mapayi et al. (2015b) | AGT + CLAHE + PC | DRIVE, STARE |
| **Model-based approach** | Xiao et al. (2013) | BMSC | DRIVE, STARE |
| | Salazar-Gonzalez et al. (2014) | GC | DRIVE, STARE, DIARETDB1 |
| | Zhao et al. (2014) | ADF + LS + RG | DRIVE, STARE |
| | Zhao et al. (2015a) | RIC + LP + GC | DRIVE, STARE, ARIA, VAMPIRE |
| | Zhao et al. (2015b) | LP + IPAC | DRIVE, STARE, VAMPIRE |
| **Other general approach** | Hassanien et al. (2015) | BSO + FCM + PSO | DRIVE, STARE |
| | Frucci et al. (2016) | DM | DRIVE |
| | Lázár and Hajdu (2015) | MSMF + BHT + DRVM + RG | DRIVE, STARE, HRF |

3. Vessel enhancement.

These techniques don't necessarily guarantee to achieve higher segmentation accuracy. It depends on the developed method and application at hand. Therefore, the reader should be mindful about when and what preprocessing method to be employed for a given problem. All the preprocessing techniques have been experimented on the green channel of the RGB retinal image. This is because the green channel exhibits better contrast between the vessels and the background (Marin et al. (2011), Fraz et al. (2012b), Azzopardi et al. (2015), Nguyen et al. (2013b), Zhao et al. (2014)). In Mendonca and Campilho (2006), different color components were investigated and found that the green channel exhibits the highest contrast between vessels and background. Once the green channel is extracted from RGB retinal image, the following preprocessing steps are performed.

### 2.3.1 Removal of Central Vessel Reflex

Retinal vessels usually appear darker than the background surface because of lower reflectance. It includes a light streak running longitudinally along the vessel centre known as Central Vessel Reflex (CVR). It is more prominent in arteries than in veins, because this phenomenon occurs at longer wavelengths that are more responsive to the blood oxygen content (Narasimha-Iyer et al. (2008)). It is more visible in younger individuals than in adults (Fraz et al. (2014)). This vessel reflex must be removed because the pixels intensities in the middle of the vessel are much lower than its surroundings leading to false detection of two close vessels instead of a single vessel.

The classical technique for removal of CVR is morphological opening. In this method, the morphological opening is performed using a three-pixel diameter disc, defined on a square grid by using eight-connectivity as a structuring element. Ricci and Perfetti (2007) initially proposed a basic line detector for vessel segmentation which efficiently deals with vessel reflex. This method is based on the average grey level response of a fixed line length passing through the target pixel at different orientations. The limitation of the basic line detector is that it tends to merge adjacent parallel vessels and produces false detections near bifurcation/crossover regions (see Figure 2.1). To overcome this limitation, Nguyen et al. (2013b) proposed a method based on the response of basic line detector on various scales and varying the length of the line. This multi-scale line detector can recognize vessel reflex as a part of a vessel since the average line response is not affected much because the

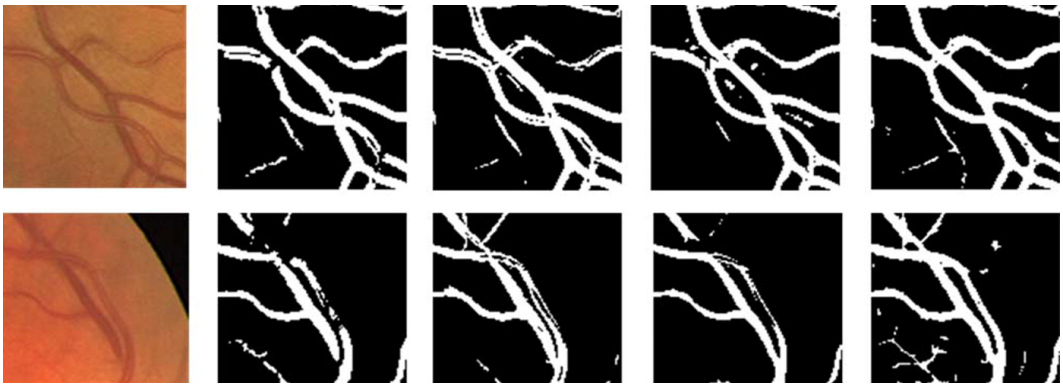


Figure 2.1: Segmentation results on ROI showing the performance of various existing methods in the presence of CVR, close parallel vessels and bifurcation/crossover regions. First column: original image; Second column: segmentations of Staal et al. (2004) method; Third column: Soares et al. (2006) method; Fourth column: Ricci and Perfetti (2007) method; and Fifth column: Nguyen et al. (2013b) method. (Illustrations based on material from Nguyen et al. (2013b), with kind permission from Elsevier).

central reflex constitutes an only small number of pixels compared to its surrounding vessel pixels (see Figure 2.1).

### 2.3.2 Intensity Inhomogeneity Correction

During image acquisition, retinal fundus images often contain background intensity variation caused by non-uniform illumination. This influences the performance of the vessel segmentation algorithms. Among the several methods proposed in the literature, histogram equalization is commonly employed for intensity inhomogeneity correction. The drawback of this approach is that it cannot handle color images and pixels intensities spanning the whole range of display devices. Another popular enhancement technique is gamma correction, which enhances images that are either too dark or too bright. But this method is image dependent. Foracchia et al. (2005) proposed a well-known technique for luminosity and contrast normalization in retinal images. In this method, the Luminosity and contrast variability in the background part of the image is estimated and then used for the normalization of the whole image. Normalization is performed on both intra and inter images.

One of the most promising approach used in the literature is the Contrast-Limited Adaptive Histogram Equalization (CLAHE) technique proposed in Pizer et al. (1987). In this technique, the image is divided into contextual regions that are then individually enhanced using histogram equalization. The clipping level of the histogram is chosen by computing the local histogram mapping function. This can

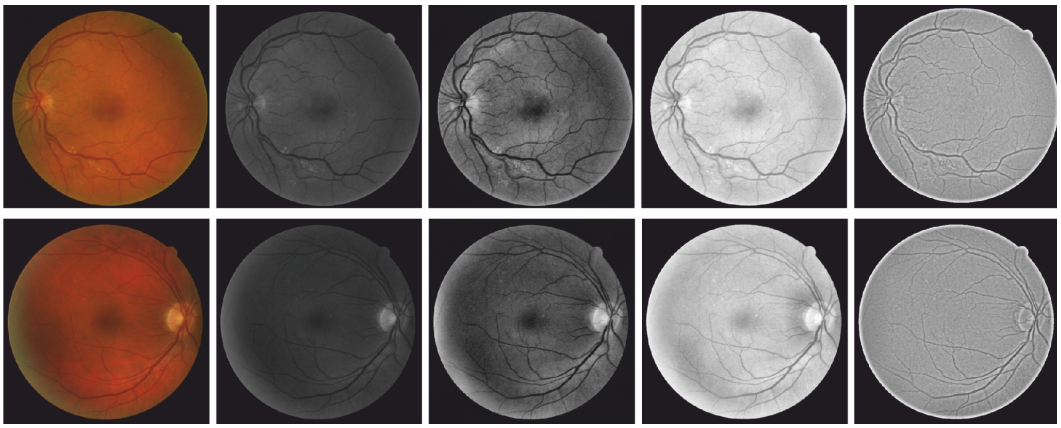


Figure 2.2: A comparative study on image-wise enhancement techniques. First column: two example images from the DRIVE dataset; Second column: the green channel image; Third and Fourth column: results after applying Histogram Equalization and Gamma correction image enhancement methods; Fifth column: results after applying Retinex. (Illustrations based on material from Zhao et al. (2015a), with kind permission from PloS one).

reduce the amplification of noise in similar regions. Azzopardi and Petkov (2013), Welikala et al. (2014), Sigurosson et al. (2014), Zhao et al. (2014) have adopted this technique. But the problem with this method is that if there is a bright or dark lesion next to the vessel, the lesion location is further enhanced leading to difficulty in distinguishing between the vessel and lesion structure.

Fraz et al. (2013) proposed a method, based on subtracting an estimate of the image background (which is obtained by applying an arithmetic mean kernel by using decimation) from the original image. The most recent and effective technique for intensity inhomogeneity correction was proposed by Zhao et al. (2015a) and is based on Retinex theory (see Figure 2.2). In Zhao et al. (2015a), an image is modelled as a multiplication of two components, the reflectance and the illumination. The reflectance component of the image reveals the object of interest more clearly, and it can be considered as the enhanced version of an image.

### 2.3.3 Vessel Enhancement

The retinal image constitutes varying anatomical structures in their size, complex shapes and orientation such as blood vessels, optical disc, and background tissues. The problem with vessel segmentation is the complexity of vascular structures near the thin vessels and bifurcation/crossover regions. So, it is often necessary to enhance vessel structures. Chaudhuri et al. (1989) were the first to investigate the vessel enhancement technique by applying matched filters. It is based on Gaussian shape

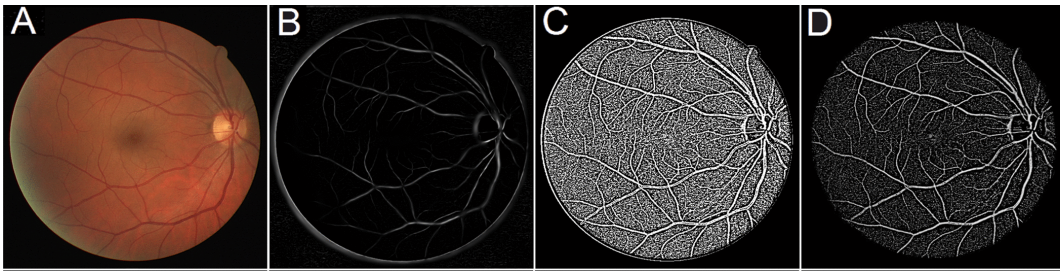



Figure 2.3: Vessel enhancement results using different methods. (a) A randomly chosen image from the DRIVE dataset; (b) Eigenvalue-based (Frangi et al. (1998)); (c) Wavelet-based (Bankhead et al. (2012)); and (d) Local Phase-based (Zhao et al. (2015a)) enhancements on (a). (Illustrations based on material from Zhao et al. (2015a), with kind permission from PloS one).

modelling of vessel cross-sectional profile of the retina. But the major drawback of this approach is, it cannot effectively enhance vessels of varying width at a single scale. On the other hand, even by adopting multiple filters with various scales, some fine vessels cannot be detected due to the low-density contrast and relatively heavy background noise.

Most of the retinal vessel enhancement filters found in the literature are based on image intensity profile. Some of the prominent intensity based filters are matched filters (Chaudhuri et al. (1989)), amplitude-modified second order Gaussian filter (Gang et al. (2002)), Eigenvalue based filter (Frangi et al. (1998)), multi-scale linear operators (Palomera-Pérez et al. (2010)), wavelet (Bankhead et al. (2012), Fathi and Naghsh-Nilchi (2013)), Gabor filter (Soares et al. (2006)) and COSFIRE filter (Azzopardi et al. (2015), Azzopardi and Petkov (2013)). But the drawback with the above-mentioned intensity based filters is, they are more prone to non-uniform illumination present in the image. In contrast, filters based on local phase information (Zhao et al. (2015a), Zhao et al. (2015b)) of an image is an emerging technique which can avoid the problems faced by the intensity based filters. Local Phase (LP) filters can be estimated by quadrature filters under the concept of monogenic signals for two or higher dimensions. These LP filters can show consistent results at the bifurcation and crossover regions when compared to the other parts of the vessels (see Figure 2.3). Thus, indeed it can be applied to solve the complex segmentation problem at vessel crossings and bifurcations regions.

Another well-known vessel enhancement technique reported in the literature is the use of morphological filters, such as top hat transform (Vega et al. (2015), Roychowdhury et al. (2015b), Sigurosson et al. (2014), Fraz et al. (2012b), Fraz et al.

(2012b), Fraz et al. (2014)). These morphological filters can enhance the image regions by estimating local background by a morphological opening operation, which is then subtracted from the original image. One of the recent technique reported in the literature is matched filters with multi-wavelet kernels (MFMK) proposed by Wang et al. (2013). These MFMK methods can separate vessels from clutter edges and bright localized areas such as lesions. Finally, the effects of each preprocessing step might affect each segmentation technique differently making any implicit generalization difficult. The comparison of the above preprocessing steps and their interaction with segmentation performance is quite complex and beyond the scope of this study.

## 2.4 Segmentation Methods

In this section, the most recent techniques are reviewed that are proposed for the retinal vessel segmentation task in the recent years. The Table 2.1 summarizes the vessel segmentation methods presented in 56 papers published between 2012 and 2017, along with the datasets used for performance evaluation. In this review, the main focus is on methods applied to color fundus images, FA and SLO images, since most of the earlier methods have been evaluated on these three modalities.

Broadly, all the previously published segmentation methods can be divided into two main categories, supervised and unsupervised. Supervised methods require a manually annotated set of training images for classifying a pixel either as a vessel or a non-vessel in previously unseen images. In these methods, the features extracted for each vessel pixel, along with their true label are used to learn a model of a classifier. Most of the techniques in this category are k-nearest neighbors, Support Vector Machine (SVM), Neural Networks (NN), Gaussian Mixture Models (GMM), AdaBoost, Conditional Random Fields (CRF) and the recent Convolutional Neural Networks (CNN) etc. On the contrary, unsupervised methods can segment the vessels without requiring any labelled annotations. In general, most of these techniques are based on the response of matched filters, morphological processing, vessel tracing, thresholding, region growing, multi-scale approaches, etc. In the following section, the essence of these techniques are briefly explained, along with the performance of vessel segmentation methods in various challenging conditions, followed by strength and weakness of these methods.

### 2.4.1 Supervised Methods

Supervised methods are based on pixel classification where the primary objective is to segment retinal vessels by training a set of manually annotated gold standard images. Often these gold standard images are annotated by experienced ophthalmologists. The most discriminative set of feature vectors must be selected from the set of training images for better classification of the vessel and non-vessel structures. In general, the performance of supervised methods is better than that of unsupervised ones.

Fraz et al. (2012b, 2014) proposed a novel method in this category based on a combination of multiple feature extraction techniques such as orientation analysis of gradient vector field, morphological transformation, line strength measure, and Gabor filter response. Classification of pixels was performed using an ensemble of weak learners such as decision trees. Most of the earlier techniques were evaluated on DRIVE and STARE dataset which consists of a fewer number of images. In contrast, Fraz et al. (2012b) created a new database known as CHASE_DB1 which was primarily used to study the cardiovascular risk factors in younger individuals. The addition of CHASE_DB1 provides an opportunity for the researchers to evaluate their method on this challenging dataset due to the presence of CVR and illumination artefacts.

Condurache and Mertins (2012) adopted a binary classification technique based on hysteresis classification paradigm. Their method is well suited for binary classification problem afflicted by significant class skew and overlap between classes such as vessels and background structures. Welikala et al. (2014) proposed a dual classification approach for the segmentation of new blood vessels. A method based on gray level co-occurrence matrix (GLCM) and neural network (NN) classifier was proposed in Rahebi and Hardalaç (2014). A novel segmentation technique based on non-linear invertible orientation scores is presented in Abbasi-Sureshjani et al. (2015). According to the authors, the method is robust against noise, non-uniform illumination and contrast variability. In addition to preserving the vessel connectivity, it has higher sensitivity and detects the small vessels better than the state-of-the-art methods for both RGB and SLO images.

An effective technique to deal with pathological images was proposed in Ganjee et al. (2014). Their method adopted a combination of multi-scale matched filtering

along with shape and region-based features to distinguish between the vessel and lesion structures. A combination of first and second order gradient features along with Gaussian Mixture Model (GMM) classifier was used in Roychowdhury et al. (2015b). Strisciuglio et al. (2016) proposed an efficient method for segmenting vessels in pathological images using a selective set of B-COSFIRE filters utilizing different feature selection methods. For the first time in the literature, conditional random fields for retinal vessel segmentation have been proposed by Orlando et al. (2017). In their method, a fully connected CRF is trained using structured output SVM in a supervised manner. Their approach was able to solve the poor segmentation problem due to weak priors assigned to an elongated structure as in the case of standard pairwise potential.

Recently, Cheung et al. (2014) observed that the micro-vascular structures (thin vessels) in the retina play a vital role in the early diagnosis of diseases such as stroke, alzheimers and other micro-vascular diseases. In contrast, segmentation of thin vessel structures in the retina is extremely difficult compared to the large vessels. To overcome this limitation, Dai et al. (2015) proposed a gray voting technique based on local gray level statistics. Their method enhances small vessel structures and has shown promising results. Further, a general extension of Local Binary Pattern (LBP) operator was proposed in Fathi and Naghsh-Nilchi (2014), with NN classifier to segment the thin vessels. Zhang et al. (2015) adopted a multiscale texton dictionary and NN classifier to classify vessel/non-vessel pixels. A Gabor filter bank is used to extract features which are then used as initial key points to initialize k -means clustering which builds a texton dictionary ultimately.

Inspired by the success of deep learning techniques in computer vision applications, Convolutional Neural Network (CNN) has emerged as a promising approach to solve segmentation problems in medical imaging applications. In this context, Wang et al. (2015) proposed a segmentation method by employing CNN and ensemble of random forests. Later, Vega et al. (2015) adopted a new generation of neural network known as Lattice Neural Network (LNN). Some of the interesting properties of LNN are, there are no convergence problems with a single layer feed forward neural network and does not involve any hidden layers. Li et al. (2016) proposed a cross-modality approach using deep autoencoders. A deep neural network is used to model the relationship between the retinal image and the vessel map. The proposed method works well in the presence of CVR, pathological images, and thin

vessel structures.

The most recent approach by Liskowski and Krawiec (2016) proposed a deep neural network model. Their method was able to achieve an area under the curve (AUC) of 0.99, which is significantly better than all previously published methods in the literature. The technique is also robust to CVR and performs reasonably well on pathological images. A novel CNN architecture was proposed in Maninis et al. (2016) to solve both the retinal vessel and optic disc segmentation problem. Wu et al. (2016) proposed a DCNN architecture under a probabilistic tracking framework, designed with generalized particle filtering technique to extract retinal vessel tree. Later, Fu et al. (2016b) formulated the vessel segmentation as a boundary detection problem using fully connected CNN model. A fully connected CRF's is utilized further to take into account the long-range interaction between the pixels.

### 2.4.2 Unsupervised Methods

The unsupervised segmentation methods work without prior knowledge and labelled ground truths. In general, most of these methods are rule-based techniques which include conventional matched filtering, morphological processing, vessel tracing, thresholding, region growing, multi-scale approaches, etc. The unsupervised techniques generally have a higher speed and lesser computational complexity compared to supervised ones.

#### 2.4.2.1 Matched Filtering

Matched filtering (MF) involves convolving an image with a 2-D Gaussian template. These methods exploit the fact that the vessel cross-sectional intensity profile can be modelled as a Gaussian shaped curve. But these assumptions often fail in the presence of CVR. Hence several variants of the MF have been proposed in the literature to overcome this limitation. The earliest approach of using 2-D matched filters for vessel segmentation was proposed in Chaudhuri et al. (1989). Later, Wang et al. (2013) extended the idea of matched filtering in combination with multi-wavelet kernels to separate the blood vessels from lesion structures. To enable the faster and parallel implementation, Krause et al. (2016) employed local Radon transforms for retinal vessel segmentation. They have segmented 20,000 images of size 2,048 $\times$ 1,536 in about 3 hours on an NVIDIA Ge-Force GTX680. This is one of the most recent methods in implementing the segmentation algorithm on the GPU-based platform with

very low time complexity. Another well-known technique in this category is based on B-COSFIRE (Bar - Combination of Shifted Filter Responses) filter proposed by Azzopardi et al. (2015). The idea behind their approach is that the COSFIRE filter can selectively respond to bar-shaped (linear) structures such as vessels.

Kovács and Hajdu (2016) proposed a self-calibration technique, which can be used to transform a trained model to retinal images of different resolution, field-of-view (FOV), noise level, etc. The segmentation method was based on template matching using Gabor filters and contour reconstruction strategies. Kar and Maity (2016) proposed a combined approach using curvelet transform, matched filtering and Laplacian of Gaussian filter. Their experimental results show that the method performed well on both pathological as well as noisy retinal images. Zhang et al. (2016) proposed a segmentation technique based on maximizing the multi-scale second-order Gaussian derivatives filter response in the orientation score domain. The method has been evaluated on six publicly available databases including the SLO images such as IOSTAR and RC-SLO datasets. Their approach has shown better segmentation performance at vessel crossings, CVR, closely parallel and tiny vessels.

#### 2.4.2.2 Multi-scale Approach

Retinal vessel structures appear at multiple scales and orientation in an image. This property has been exploited in the literature to capture vessel structures of varying width and direction. The advantage of these methods is both major and minor vessels can be segmented effectively. Among the techniques proposed in this category, the earliest approach by Yu et al. (2012) is based on Eigen analysis of the Hessian matrix, followed by second-order local entropy thresholding. Their method had a few drawbacks: first, it tends to over segment the major vessels including pathological lesions; second, it under-segments the thin vessels. To overcome these limitations, Moghimirad et al. (2012) proposed a method based on multi-scale medialness function initially intended for tubular structure extraction. Most of the tubular structure extraction schemes are based on Eigen analysis of the Hessian matrix. But the problem with these methods is that the response of the Eigenvalues of the Hessian matrix is weaker in the area of bifurcation/crossover points. To address these issues, authors in Moghimirad et al. (2012) proposed a combined approach of Hessian matrix and 2-D medialness function. The advantage of their method is that it provides better

segmentation at bifurcation and crossover regions. Zheng et al. (2013) proposed multi-scale Hessian matrix based non-local filtering approach to enhance the vessel structures and suppress the background noise.

Nguyen et al. (2013b) introduced the concept of multiscale line detection (MSLD) for vessel segmentation. This idea was the initial extension of the basic line detection method proposed by Ricci and Perfetti (2007). In comparison to Ricci and Perfetti (2007), their method overcomes the limitation of poor segmentation near two closely parallel vessels and at crossover points. Later, Yin et al. (2015) modified the idea proposed in Nguyen et al. (2013b) to design a novel orientation aware detector to segment both major and minor vessels. Their method works well in the presence of CVR, close vessels, and crossover points. The main limitation of the line detector methods is that it is unable to segment the smallest vessels. This drawback has been addressed by Christodoulidis et al. (2016) by adopting a hybrid approach based by combining MSLD and multi-scale tensor voting procedure. This approach has shown a better segmentation performance on thin vessel structures compared to other previous methods.

Fathi and Naghsh-Nilchi (2013) proposed a complex continuous wavelet transform (CCWT) with an adaptive histogram based thresholding method for segmentation of vessels. Vessel diameter estimation has also been addressed by employing circular structure descriptor on the centerline of the vessels. Budai et al. (2013) proposed an improvement over the vessel enhancement method initially proposed in Frangi et al. (1998). A novel approach based on inpainting technique was suggested by Annunziata et al. (2016) to inpaint the false vessel structures (such as exudates) in pathological images. Their method significantly reduces the number of false positives in pathological images. A multiscale line tracking procedure was proposed in Abdallah et al. (2015) based on the fact that the vessel structures appear at multiple scales and orientation in an image.

#### 2.4.2.3 Vessel Tracing

In vessel tracing methods, initial seed points are chosen either manually or automatically both on the edges and centerline of the vessels. Given these initial seed points, the entire vessel tree is traced by following vessel centerline based on local information. Since the vessels are connected in the retina, tracing methods can follow a whole tree structure without explicitly monitoring the background. These methods

provide a precise vessel connectivity information at branching and crossovers points for early detection of many systemic diseases.

The method in this category was proposed in Yin et al. (2012, 2013) based on maximum a posteriori (MAP) formulation. The initial vessel edge points are detected iteratively using local grey level statistics and vessel continuity properties. A Gaussian-shaped curve is fitted to the intensity profile of local vessel cross-section to estimate the vessel appearance. Similar to the previous work, Zhang et al. (2014) proposed a combined approach using MAP criterion and multiscale line detection method. In their work, the authors succeed in differentiating between normal, branching and crossover vessels. A principled way of addressing the crossover issue has been solved in De et al. (2014) based on transductive learning approach. Their method performs better in resolving many complex bifurcation/crossover points. Inspired by the modelling of cortical columns in the primary visual cortex, Bekkers et al. (2014) proposed a multi-orientation vessel connectivity analysis near bifurcation and crossing regions. Among all the tracing methods, their approach was able to track and measure width even at multiple locations successfully.

#### 2.4.2.4 Mathematical Morphology

Mathematical morphology is a powerful tool based on set theory concept, mainly used for extracting complex image structures that provide useful representation and description of region shapes such as features, boundaries, skeletons and convex hulls. These methods exploit the fact that the vessels are linear and connected in the retina. These methods are also known for its speed and noise resistance. The main drawback of these methods is, they fail to model the highly curved vessels which are more prominent in younger individuals.

Among the methods in this category, Fraz et al. (2013, 2012b) proposed a novel approach for identifying both vessel centerline and segmentation of vascular tree. The vessel centerlines are identified using the first order derivative of a Gaussian filter followed by a multi-directional morphological top-hat transform to segment the vessels. The main limitations in Fraz et al. (2013, 2012b) is difficulty in modelling the highly curved and tortuous vessels. Later, Sigurosson et al. (2014) addressed this challenge by proposing a novel approach based on path opening filter followed by a fuzzy set theory based data fusion technique. Their method demonstrated an ability to distinguish between major and minor vessels with better accuracy than

the other methods. The traditional problem of vessel segmentation in the presence of abnormalities was tackled in Imani et al. (2015). In their approach, morphological component analysis (MCA) combined with Morlet Wavelet Transform (MWT) was used to separate vessels from other lesion structures which are crucial in clinical settings for the assessment of abnormal cases.

#### 2.4.2.5 Thresholding Based Approach

Among the most recent thresholding based approaches, Roychowdhury et al. (2015b) proposed a method which can efficiently handle vessel segmentation in the presence of abnormalities as well as thin vessel structures. Their method adopted an iterative vessel segmentation approach by employing global adaptive thresholding followed by a novel stopping criteria. Further, Mapayi et al. (2015a) proposed a local adaptive thresholding technique based on GLCM energy information for segmenting retinal vessels. Their method provides robust segmentation for both grayscale and green channel intensity of RGB retinal images. Later, Mapayi et al. (2015b) proposed an approach based on global thresholding with phase congruency and CLAHE for segmentation of vessels.

#### 2.4.2.6 Model Based Approach

Model-based approaches consist of vessel profile models and deformable models. In vessel profile models, the vessel cross-sectional intensity profile is modelled as a Gaussian-shaped curve or mixture of Gaussians in the case of CVR. In deformable models, both active contour and level set based approaches are employed. Among the methods in this category, Xiao et al. (2013) presented a Bayesian-based segmentation approach which takes into account the spatial information. Their method results in better performance in the detection of both narrow and low contrast vessels. Further, Salazar-Gonzalez et al. (2014) presented a combined framework for both vessel and optic disc segmentation problem. Their method adopted a graph cut technique to segment the vessel tree followed by MRF image reconstruction and compensation factor method to segment the optic disk. Zhao et al. (2014), adopted two separate techniques for extracting both major and minor vessels from retinal images. For major vessels, a level set method based on region-scalable fitting energy function is applied and for minor vessels, a region growing approach is adopted.

Recently, a novel inhomogeneity correction method based on Retinex theory and LP analysis was proposed in Zhao et al. (2015a) for vessel enhancement. In

their work, the authors have shown the comparative analysis of various existing preprocessing methods in the literature along with their proposed one (see Figure 2.2, Figure 2.3). Adopting the similar preprocessing technique proposed in Zhao et al. (2015a), Zhao et al. (2015b) proposed a segmentation approach based on infinite perimeter active contour model utilizing hybrid region information of the image. For the first time in the literature, the authors in Zhao et al. (2015a) have evaluated their segmentation technique on FA images which could open further research directions in retinal image analysis.

#### 2.4.2.7 Other General Approach

The methods published in this category include those that belong to general image processing based techniques that are adapted for segmenting retinal vessels. Hassanien et al. (2015) proposed the idea of two-level optimization for segmenting thin vessel structures. The first level includes finding the vessel clusters with artificial bee colony swarm optimization using fuzzy c-means fitness function. In the second level, a pattern search algorithm is enhanced by adding shape descriptors as an additional feature in the fitness function. Later, Frucci et al. (2016) used the concept of the directional map for segmenting vessel structures. The advantage of their method is, it is computationally less intensive and requires no preprocessing steps. Finally, Lázár and Hajdu (2015) proposed a hybrid region growing approach based on directional response vector similarity of pixels along with the nearest neighbor classifier.

## 2.5 Summary of Vessel Segmentation Methods

Among the recent methods published in the literature, supervised methods based on deep learning architecture (Li et al. (2016); Liskowski and Krawiec (2016)) has surpassed all other techniques in obtaining the performance very close to (or even outperforming) trained human observers. These methods are able to successfully address most of the challenges posed by the complex nature of vessel structures. The success of these methods is mainly due to larger computational power and the addition of large datasets. The main advantage of CNN based methods is that they don't require any carefully hand-crafted features and complex domain expertise. In contrast, these methods are able to successfully learn the complex nature of vessel structures even in the most challenging cases (like the presence of thin vessel structures and abnormalities). A new wave of CNN methods have been recently proposed

in the literature such as Fu et al. (2016b); Wu et al. (2016) manifesting a special interest in the community towards the development of newer deep learning based architectures for retinal vessel segmentation.

Table 2.2 depicts the segmentation challenges addressed by recent state-of-the-art methods. The techniques proposed in Christodoulidis et al. (2016); Dai et al. (2015); Fathi and Naghsh-Nilchi (2014); Kovács and Hajdu (2016); Roychowdhury et al. (2015b); Sigurosson et al. (2014); Xiao et al. (2013); Zhao et al. (2014) has explicitly addressed the challenges of segmenting thin vascular structures. Christodoulidis et al. (2016) proposed a combined approach of MSLD along with multiscale tensor voting for detection of major and minor vessel structures. This is one of the first kind in segmentation methods that have specifically considered the segmentation of thin vessels and has shown remarkable performance. The problem of CVR, poor segmentation near bifurcation/crossover points and close parallel vessels have been addressed in Nguyen et al. (2013b). Other methods like Zhao et al. (2015a), Zhao et al. (2015b) has considered the problem of intensity inhomogeneity and vessel enhancement techniques on low contrast images. Although the most recent methods have shown interest in addressing the segmentation challenges, there is still room for future improvements towards the building of successful retinal CAD systems.

## 2.6 Review of Validation

In this section, the quantitative assessment of segmentation methods proposed in the current state-of-the-art is analysed, including the datasets employed and measures performed. Some significant findings of the existing techniques and few open problems have also been discussed.

### 2.6.1 Retinal Image Databases

Retinal image segmentation methods in the literature are evaluated using different quantitative measures and on different publicly available datasets. Most of the existing state-of-the-art methods have been evaluated using publicly available: DRIVE (Staal et al. (2004)), STARE (Hoover et al. (2000)), CHASE_DB1 (Fraz et al. (2012b)) and HRF (Odstrcilik et al. (2013)) datasets. Few methods have also made an attempt to validate their technique on other databases like ARIA (Aria (2006)), DIARETDB1 (Diaretdb1 (2007)), MESSIDOR (Messidor (2017)), REVIEW (Al-Diri et al. (2008)), VAMPIRE (Perez-Rovira et al. (2011)), IOSTAR (Zhang et al.

Table 2.2: Segmentation challenges addressed by recent state-of-the-art methods. The acronyms stand for: robust to pathological lesions (PL), able to segment thin vessel structures (TV), robust to central vessel reflex (CVR), better segmentation near bifurcation/crossover regions (BC), robust to intensity inhomogeneity, blur and noise present in the image (IN).

| **Authors** | **PL** | **TV** | **CVR** | **BC** | **IN** |
|---|---|---|---|---|---|
| Fraz et al. (2012b) | Yes | — | Yes | — | — |
| Fathi and Naghsh-Nilchi (2014) | — | Yes | — | — | — |
| Fraz et al. (2014) | — | — | Yes | — | — |
| Ganjee et al. (2014) | Yes | — | — | — | — |
| Orlando et al. (2017) | — | — | — | — | Yes |
| Dai et al. (2015) | — | Yes | — | — | — |
| Roychowdhury et al. (2015a) | Yes | — | — | — | — |
| Li et al. (2016) | Yes | Yes | Yes | Yes | Yes |
| Liskowski and Krawiec (2016) | Yes | Yes | Yes | Yes | Yes |
| Fu et al. (2016b) | Yes | — | — | — | — |
| Zhang et al. (2016) | Yes | Yes | Yes | Yes | Yes |
| Strisciuglio et al. (2016) | Yes | – | – | – | – |
| Wang et al. (2013) | Yes | — | — | — | — |
| Azzopardi et al. (2015) | — | — | — | Yes | — |
| Kovács and Hajdu (2016) | — | Yes | — | — | — |
| Kar and Maity (2016) | Yes | — | — | — | Yes |
| Moghimirad et al. (2012) | — | — | — | Yes | — |
| Zheng et al. (2013) | — | — | — | — | Yes |
| Nguyen et al. (2013b) | — | — | Yes | Yes | — |
| Annunziata et al. (2016) | Yes | — | — | — | — |
| Yin et al. (2015) | Yes | — | Yes | Yes | — |
| Christodoulidis et al. (2016) | Yes | Yes | — | — | — |
| Yin et al. (2012) | — | — | — | Yes | — |
| Zhang et al. (2014) | — | — | Yes | Yes | — |
| De et al. (2014) | — | — | — | Yes | — |
| Bekkers et al. (2014) | — | — | — | Yes | — |
| Sigurosson et al. (2014) | — | Yes | — | — | — |
| Imani et al. (2015) | Yes | — | — | — | — |
| Roychowdhury et al. (2015b) | Yes | Yes | — | — | — |
| Xiao et al. (2013) | — | Yes | — | — | — |
| Zhao et al. (2014) | — | Yes | — | — | — |
| Zhao et al. (2015a) | — | — | — | Yes | Yes |
| Zhao et al. (2015b) | — | — | — | — | Yes |

(2016)) and RC-SLO (Zhang et al. (2016)). An overview of the aforementioned publicly available datasets are provided in Table 2.3.

The earliest datasets like DRIVE and STARE includes a fewer number of images ranging from 20 to 40. All of these images mainly consists of healthy images and fewer pathological cases. With the recent advancements in newer high-resolution fundus camera, it is now possible to obtain much higher resolution ($3504 \times 2336$) images as in HRF dataset, when compared to lower resolution ($768 \times 584$) images in case of DRIVE. A thorough evaluation of segmentation methods particularly on abnormal cases is required for clinical scenarios. Such images can be found in STARE, ARIA and HRF datasets, which contain instances of DR, AMD, and Glaucoma. Recently, Fraz et al. (2012b) introduced a new database CHASE_DB1, which includes images of multi-ethnic school children used primarily for quantifying vascular changes in the retinal, in relation to cardiovascular disease. Recently, Zhang et al. (2016) published two newer datasets: IOSTAR and RC-SLO based on the SLO technique. These two datasets cover a wide range of challenging cases, such as high curvature changes, CVR, micro-vessels and crossings/bifurcations.

### 2.6.2 Validation Measures

Four primary validation measures have been commonly employed to evaluate the performance of vessel segmentation methods. They are $Se$ (Sensitivity), $Sp$ (Specificity), $Acc$ (Accuracy) and $AUC$ (Area under the curve). Several newer validation measures for vessel segmentation have also been reported in the literature which is provided in Table 2.4.

Since retinal vessel segmentation is a binary classification problem, the commonly accepted measures include: i) True positive ($TP$) - Number of correctly classified vessel pixels, ii) False negative ($FN$) - number of incorrectly classified vessel pixels, iii) True negative ($TN$) - Number of correctly classified background pixels and iv) False positive ($FP$) - Number of wrongly classified background pixels. Based on these critical measures, different performance parameters can be estimated such as $Se$, $Sp$, $Acc$ and $AUC$. Sensitivity indicates the capability of the algorithm to correctly detect retinal vessels, while specificity indicates the ability to distinguish all other non-vessel structures. Accuracy measures the ratio of correctly classified pixels (both vessel and non-vessel) to the total number of pixels in the image field-of-view. All these measures are obtained through the pixel to pixel comparison between

Table 2.3: Overview of publicly available retinal image datasets.

| Dataset | Year | Description | Image size and FOV | Vessel groundtruth (Reference standard) |
|---|---|---|---|---|
| STARE | 2000 | Total 20 color fundus images. out of which 10 are healthy and 10 are pathological. | $700 \times 605$ $35°$ | Groundtruths are annotated by two human observers. |
| DRIVE | 2004 | Contains 40 color fundus images. It is divided into 20 for training and 20 for testing. | $584 \times 565$ $45°$ | Groundtruths are annotated by two human observers. |
| MESSIDOR | 2004 | Contains 1200 color fundus images. Contains pathological signs such as microaneurysms, hemorrhages, neovascularization and hard exudates. | $1440 \times 960$, $2240 \times 1488$, $2304 \times 1536$ $45°$ | No vessels groundtruth are available. Reference standard is available for grading of diabetic retinopathy and the risk of macular edema. |
| ARIA | 2006 | Total of 212 color fundus images: first group: 92 AMD images second group: 59 DR images third group: 61 normal images | $768 \times 576$ $50°$ | The reference standard for optic disc, blood vessel tracking and fovea location is marked by two clinical experts. |
| DIARETDB1 | 2007 | Consists of 89 retinal images. Out of which 84 contain signs of DR and 5 are normal. | $1500 \times 1152$ $50°$ | Manually segmented retinal vasculature is not available for this database. |
| REVIEW | 2008 | Contains 16 images with 193 annotated vessel segments consisting of 5066 profile points | – | Vessel widths are manually marked by three independent experts. |
| CHASE_DB1 | 2011 | Contains 28 color fundus images. Out of which 20 corresponds to testing and 8 for training. | $1280 \times 960$ $30°$ | Vessel segments and width are annotated by two clinical experts. |
| HRF | 2011 | Contains 45 color fundus images: first set: 15 healthy images second set: 15 DR images third set: 15 glaucoma images | $3304 \times 2336$ $60°$ | Reference standard was provided by three clinical experts. |
| VAMPIRE | 2011 | 8 retinal FA images. Out of which 4 contains AMD. | $3900 \times 3072$ $200°$ | Groundtruth was provided by three clinical experts. |
| IOSTAR | 2015 | Includes 30 SLO images | $1024 \times 1024$ $45°$ | Vessels are annotated by a group of experts |
| RC-SLO | 2015 | Contains 40 SLO image patches | $360 \times 320$ – | Vessels are annotated by a group of experts |

Table 2.4: Overview of validation measures used for quantitative assessment of retinal vessel segmentation methods.

| **Validation measure** | **Description** |
|---|---|
| Sensitivity ($Se$)/Recall ($Re$) | $\frac{TP}{TP+FN}$ |
| Specificity ($Sp$) | $\frac{TN}{TN+FP}$ |
| Accuracy ($Acc$) | $\frac{TP+TN}{TP+TN+FP+FN}$ |
| Area under curve ($AUC$) | $\frac{Se+Sp}{2}$ |
| Precision ($Pr$) | $\frac{TP}{TP+FP}$ |
| Matthews Correlation Coefficient ($MCC$) | $\frac{\frac{TP}{N}-S\times P}{\sqrt{P\times S\times 1-S\times 1-P}}$ $N = TP + TN + FP + FN,$ $S = TP + FN \times N,$ $P = TP + FP \times N$ |
| F1 score ($F$) | $\frac{2\times Pr\times Re}{Pr+Re}$ |
| G - mean ($G$) | $\sqrt{Se \times Sp}$ |
| Dice Coefficient ($DC$) | $\frac{2(A\cap B)}{A+B}$ |
| Quality Evaluation Function ($QEF$) | $f(C, A, L) = C \times A \times L = CAL$ |

automated segmentation and reference ground truth. Most of these measures are suitable if the class data is balanced (equal number of positive and negative classes). But in the case of retinal images, the negative class samples (background pixels) outnumber the positive class samples (vessel pixels). By contrast, the presence of a class imbalance in the data has a profound impact on these performance measures.

The aforementioned validation measures provide global information on segmentation quality without taking into account that the detected pixels are part of a vessel structure with specific features. Hence, more suitable validation measures for class imbalance data are included in Orlando et al. (2017) and Azzopardi et al. (2015), such as Matthews Correlation Coefficient ($MCC$), $F1 - score$ and $G - mean$. The $MCC$ is a correlation coefficient between the ground truth and the predicted binary segmentation which returns a value between -1 and +1 with +1 indicating a perfect prediction, 0 no better than random, and -1 a total disagreement between prediction and ground truth. The $F1 - score$ is the harmonic mean of precision and recall which achieves maximum value of 1 when the segmentation of the positive class is perfect, and lowest value of 0 when the segmentation is entirely wrong. Similarly, the $G - mean$ is a metric that measures the balance between $Se$ and $Sp$ by taking their geometric mean, returning a value between 0 and 1. The most common overlap metric, the Dice coefficient (DC) is also used for comparing the agreement between the manual annotations and the result of the segmentation method. The DC ranges from 0 (no agreement) to 1 (perfect agreement).

Recently, Gegundez-Arias et al. (2012), proposed a new measure called Quality Evaluation Function ($QEF$) based on the characterization of vascular structures as connected segments with the measurable area and length. This measure is sensitive to anatomical vascularity features. Hence, it is highly desirable in evaluating segmentation methods on pathological images. The $QEF$ function has also shown a high degree of matching with human quality perception when compared to other measures reported in the literature.

### 2.6.3 Assessment of Vessel Segmentation Methods

The performance evaluation of segmentation methods significantly varies across the datasets. The main reasons for this are

- Input images of different resolution across the data-sets, as shown in Table 2.3.
- Different morphological attributes of the images like pathological lesions, varying tissue structures, intensity homogeneity, and noise inherited due to various scanning protocols.
- Intra and inter-observer variability among the ground-truth annotations, both within and across datasets.

In this regard, the assessment of segmentation methods is considered on an individual dataset for fair evaluation of previously published methods. The four performance metrics are considered for evaluation: $Se$, $Sp$, $Acc$ and $AUC$. The segmentation techniques are compared on the four main datasets: DRIVE, STARE, CHASE_DB1 and HRF, as depicted in Table 2.5. The values presented in the table are taken as it is reported in the corresponding articles.

Table 2.5 depicts the performance evaluation of different methods on DRIVE dataset. Among these methods, the highest reported $Se$ of 0.9094 (Condurache and Mertins (2012)), $Sp$ of 0.9870 (Budai et al. (2013)), $Acc$ of 0.9767 (Wang et al. (2015)) and $AUC$ of 0.9790 (Liskowski and Krawiec (2016)). No single method was able to achieve the best results concerning all the metrics. When compared to all other datasets, vessel segmentation on DRIVE is relatively easy as most of the images are healthy with no pathological signs, vessels are visible, and images are not much affected by illumination. One of the key observation is that, most of the methods

Table 2.5: Comparison of segmentation methods on the DRIVE, STARE, CHASE_DB1 and HRF datasets.

| | Methods | DRIVE | | | | STARE | | | | CHASE_DB1 | | | | HRF | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | *Se* | *Sp* | *Acc* | *AUC* | *Se* | *Sp* | *Acc* | *AUC* | *Se* | *Sp* | *Acc* | *AUC* | *Se* | *Sp* | *Acc* | *AUC* |
| Supervised methods | 2nd human observer | 0.7760 | 0.9724 | 0.9472 | — | 0.8952 | 0.9384 | 0.9349 | — | 0.8105 | 0.9711 | 0.9545 | — | | | | — |
| | Fraz et al. (2012b) | 0.7406 | 0.9807 | 0.9480 | 0.9747 | 0.7548 | 0.9763 | 0.9534 | 0.9768 | 0.7224 | 0.9711 | 0.9469 | 0.9712 | — | — | — | — |
| | Condurache and Mertins (2012) | **0.9094** | 0.9591 | 0.9516 | 0.9726 | 0.8902 | 0.9673 | 0.9595 | 0.9791 | — | — | — | — | — | — | — | — |
| | Fraz et al. (2014) | — | — | — | — | — | — | — | — | 0.7259 | 0.9770 | 0.9524 | 0.9760 | — | — | — | — |
| | Rahebi and Hardalaç (2014) | 0.7365 | 0.9707 | 0.9461 | 0.9564 | 0.6902 | 0.9804 | 0.9527 | 0.9462 | — | — | — | — | — | — | — | — |
| | Fathi and Naghsh-Nilchi (2014) | 0.7649 | — | 0.9449 | — | 0.7505 | — | 0.9460 | — | — | — | — | — | — | — | — | — |
| | Ganjee et al. (2014) | — | — | — | — | 0.7437 | — | 0.9542 | — | — | — | — | — | — | — | — | — |
| | Orlando et al. (2017) | 0.7760 | 0.9730 | — | 0.9507 | 0.8951 | 0.9387 | — | — | 0.7277 | 0.9712 | — | 0.9524 | 0.7874 | 0.9584 | — | 0.9359 |
| | Wang et al. (2015) | 0.8173 | 0.9733 | **0.9767** | 0.9475 | 0.8104 | 0.9791 | **0.9813** | 0.9751 | — | — | — | — | — | — | — | — |
| | Vega et al. (2015) | 0.7444 | 0.9600 | 0.9412 | — | 0.7019 | 0.9671 | 0.9483 | — | — | — | — | — | — | — | — | — |
| | Dai et al. (2015) | 0.7359 | 0.9720 | 0.9418 | — | 0.7769 | 0.9550 | 0.9364 | — | — | — | — | — | — | — | — | — |
| | Roychowdhury et al. (2015a) | 0.7250 | 0.9830 | 0.9520 | 0.9620 | 0.7720 | 0.9730 | 0.9510 | 0.9690 | — | — | — | — | — | — | — | — |
| | Zhang et al. (2015) | 0.7812 | 0.9668 | 0.9504 | — | — | — | — | — | — | — | — | — | — | — | — | — |
| | Li et al. (2016) | 0.7569 | 0.9816 | 0.9527 | 0.9738 | 0.7726 | **0.9844** | 0.9628 | 0.9879 | 0.7507 | **0.9793** | **0.9581** | 0.9716 | — | — | — | — |
| | Liskowski and Krawiec (2016) | 0.7811 | 0.9807 | 0.9535 | **0.9790** | **0.9289** | 0.9710 | 0.9667 | **0.9930** | **0.8793** | 0.9668 | 0.9577 | **0.9845** | — | — | — | — |
| | Fu et al. (2016b) | 0.7294 | — | 0.9470 | — | 0.7140 | — | 0.9545 | — | — | — | — | — | — | — | — | — |
| | Strisciuglio et al. (2016) | 0.7777 | 0.9702 | 0.9454 | 0.9597 | 0.8046 | 0.9710 | 0.9534 | 0.9638 | — | — | — | — | — | — | — | — |
| | Abbasi-Sureshjani et al. (2015) | 0.7695 | 0.9742 | 0.9477 | 0.9525 | — | — | — | — | — | — | — | — | — | — | — | — |
| | Wang et al. (2013) | — | — | 0.9461 | 0.9543 | — | — | 0.9521 | 0.9682 | — | — | — | — | — | — | — | — |
| | Azzopardi et al. (2015) | 0.7655 | 0.9704 | 0.9442 | 0.9614 | 0.7716 | 0.9701 | 0.9497 | 0.9563 | 0.7585 | 0.9587 | 0.9387 | 0.9487 | — | — | — | — |
| | Kovács and Hajdu (2016) | 0.7450 | 0.9793 | 0.9494 | 0.9722 | 0.8034 | 0.9786 | 0.9610 | 0.9836 | — | — | — | — | 0.7502 | **0.9868** | **0.9674** | — |
| | Kar and Maity (2016) | 0.7548 | 0.9792 | 0.9616 | — | 0.7577 | 0.9788 | 0.9730 | — | — | — | — | — | — | — | — | — |
| | Zhang et al. (2016) | 0.7743 | 0.9725 | 0.9476 | 0.9636 | 0.7791 | 0.9758 | 0.9554 | 0.9748 | 0.7626 | 0.9661 | 0.9452 | 0.9606 | 0.7978 | 0.9717 | 0.9556 | **0.9608** |
| | Yu et al. (2012) | 0.7233 | — | 0.9426 | — | 0.7112 | — | 0.9463 | — | — | — | — | — | 0.7938 | — | 0.9566 | — |
| | Moghimirad et al. (2012) | 0.7852 | — | 0.9659 | 0.9580 | 0.8133 | — | 0.9756 | 0.9678 | — | — | — | — | — | — | — | — |
| | Budai et al. (2013) | 0.6440 | **0.9870** | 0.9572 | — | 0.5800 | 0.9820 | 0.9386 | — | — | — | — | — | 0.6690 | 0.9850 | 0.9610 | — |
| | Nguyen et al. (2013b) | — | — | 0.9407 | — | — | — | 0.9324 | — | — | — | — | — | — | — | — | — |
| | Fathi and Naghsh-Nilchi (2013) | 0.7768 | 0.9759 | 0.9581 | 0.9516 | 0.8061 | 0.9717 | 0.9591 | 0.9680 | — | — | — | — | — | — | — | — |
| | Annunziata et al. (2016) | — | — | — | — | 0.7128 | 0.9836 | 0.9562 | 0.9655 | — | — | — | — | 0.7128 | 0.9836 | 0.9581 | — |
| | Abdallah et al. (2015) | 0.5879 | — | 0.9155 | — | 0.6145 | — | 0.9402 | — | — | — | — | — | — | — | — | — |
| | Yin et al. (2015) | 0.8957 | — | 0.9506 | — | 0.8886 | — | 0.9315 | — | — | — | — | — | — | — | — | — |
| | Christodoulidis et al. (2016) | — | — | — | — | — | — | — | — | — | — | — | — | **0.8506** | 0.9582 | 0.9479 | — |
| | Yin et al. (2013) | 0.6522 | 0.9710 | 0.9267 | — | 0.7248 | 0.9666 | 0.9412 | — | — | — | — | — | — | — | — | — |
| | Fraz et al. (2012b) | 0.7152 | 0.9759 | 0.9430 | — | 0.7311 | 0.9680 | 0.9442 | — | — | — | — | — | — | — | — | — |
| | Fraz et al. (2013) | 0.7302 | 0.9742 | 0.9422 | — | 0.7318 | 0.9660 | 0.9423 | — | — | — | — | — | — | — | — | — |
| | Imani et al. (2015) | 0.7524 | 0.9753 | 0.9523 | — | 0.7502 | 0.9745 | 0.9590 | — | — | — | — | — | — | — | — | — |
| | Roychowdhury et al. (2015b) | 0.7395 | 0.9782 | 0.9494 | 0.9672 | 0.7317 | 0.9842 | 0.9560 | 0.9673 | 0.7615 | 0.9575 | 0.9467 | 0.9623 | — | — | — | — |
| | Mapayi et al. (2015a) | 0.7632 | 0.9634 | 0.9461 | 0.9658 | 0.7626 | 0.9657 | 0.9510 | 0.9781 | — | — | — | — | — | — | — | — |
| | Mapayi et al. (2015b) | 0.7391 | 0.9569 | 0.9377 | — | 0.5031 | 0.9567 | 0.9221 | — | — | — | — | — | — | — | — | — |
| | Xiao et al. (2013) | 0.7513 | 0.9792 | 0.9529 | — | 0.7147 | 0.9735 | 0.9476 | — | — | — | — | — | — | — | — | — |
| | Salazar-Gonzalez et al. (2014) | 0.7512 | — | 0.9412 | — | 0.7887 | — | 0.9441 | — | — | — | — | — | — | — | — | — |
| | Zhao et al. (2014) | 0.7354 | 0.9789 | 0.9477 | — | 0.7187 | 0.9767 | 0.9509 | — | — | — | — | — | — | — | — | — |
| | Zhao et al. (2015a) | 0.7440 | 0.9780 | 0.9530 | 0.8610 | 0.7860 | 0.9750 | 0.9510 | 0.8810 | — | — | — | — | — | — | — | — |
| | Zhao et al. (2015b) | 0.7420 | 0.9820 | 0.9540 | 0.8620 | 0.7800 | 0.9780 | 0.9560 | 0.8740 | — | — | — | — | — | — | — | — |
| | Hassanien et al. (2015) | 0.7210 | 0.9710 | 0.9388 | — | 0.6490 | 0.9820 | 0.9467 | — | — | — | — | — | — | — | — | — |
| | Frucci et al. (2016) | 0.6700 | 0.9860 | 0.9590 | — | — | — | — | — | — | — | — | — | — | — | — | — |
| | Lázár and Hajdu (2015) | 0.7646 | — | 0.9458 | — | 0.7248 | — | 0.9492 | — | 0.7102 | — | 0.9533 | — | — | — | — | — |

that reported the highest performance measures on DRIVE belong to the class of supervised methods.

Among the methods on STARE dataset, the highest reported $Se$ of 0.9289 (Liskowski and Krawiec (2016)), $Sp$ of 0.9844 (Li et al. (2016)), $Acc$ of 0.9813 (Wang et al. (2015)) and $AUC$ of 0.9930 (Liskowski and Krawiec (2016)). As per the observation, mostly supervised and multiscale methods have outperformed other methods in both DRIVE and STARE datasets. As discussed earlier in Section 2.4.1 and Section 2.4.2.2, learning based methods (supervised) and multiscale methods (based on Hessian matrix), have shown remarkable performance. The main reasons for the success of these methods are: (i) the ability to inherit the multiscale nature of vascular structures, which appears as both thin and thick structures at various scales and orientations. (ii) learning the complex vessel structures directly from the training data, as opposed to handcrafting those features based on domain expertise. The main drawback of the supervised methods is that it requires pixel-level annotations for learning the correspondence during the training phase and much more computationally expensive than the other category of methods.

Among the published methods on CHASE_DB1 data-sets, the highest reported $Se$ of 0.8793 (Liskowski and Krawiec (2016)), $Sp$ of 0.9793 (Li et al. (2016)), $Acc$ of 0.9581 (Li et al. (2016)) and $AUC$ of 0.9845 (Liskowski and Krawiec (2016)). This dataset is newly released compared to earlier DRIVE and STARE databases. The resolution of the images in CHASE_DB1 is of $1280 \times 960$, as opposed to $584 \times 565$ in DRIVE and $700 \times 605$ in STARE datasets. Another dataset which contains high-resolution images ($3304 \times 2336$) is the HRF, having both healthy and diseased ones. The highest reported metrics in this dataset are $Se$ of 0.8506 (Christodoulidis et al. (2016)), $Sp$ of 0.9868 (Kovács and Hajdu (2016)), $Acc$ of 0.9674 (Kovács and Hajdu (2016)) and $AUC$ of 0.9608 (Zhang et al. (2016)). Most of the segmentation methods on HRF dataset have evaluated their performance by resizing the image to a much lower resolution thus reducing the computational overhead. But the drawback with resizing the images is, the thin vessel structures are masked out leaving behind only the larger vessels, which are comparatively easier to segment. In contrast, Kovács and Hajdu (2016) proposed a blind calibration technique that can be used to transform a trained model to different resolution datasets without compromising on accuracy. Their method is well suited for clinical scenarios where the images are acquired from various scanners and protocols.

## 2.7 Discussion

The need for computer-aided detection of retinal vessels has motivated the medical imaging community to develop better segmentation methods in the last two decades. Almost all the techniques from conventional image processing to sophisticated machine learning (including deep learning techniques) have been extensively explored. Most existing methods in the literature are quite successful in addressing issues like intensity inhomogeneity, CVR, segmentation of complex vessel structures (near bifurcation and crossover regions) and extraction of thin vessels. But still, there are some open and unsolved challenges which need to be explored. Some of the limitations which needs to be addressed are:

- Segmenting vessels in the presence of abnormalities.
- Segmentation of both major and minor vessel structures in retinal image.
- Developing a robust segmentation method that works well on different images with various resolutions, acquired across multiple imaging devices.

These significant challenges have recently driven the retinal imaging community to tackle these issues. Most of the methods are designed to handle healthy images without accounting for abnormalities. The techniques in Sigurosson et al. (2014), Palomera-Pérez et al. (2010), Bekkers et al. (2014), Fraz et al. (2013) addressed, in particular, the problem of vessel segmentation in the presence of abnormalities. Most of the researchers evaluated their methods on DRIVE and STARE datasets although, a recent trend has moved towards CHASE_ DB1 and HRF databases. A large number of images from these datasets is normal excluding ARIA and MESSIDOR. The number of images in these datasets is very limited ranging from 20 to 45 and are relatively easier to segment. From Table 2.1, it can be observed that only fewer methods evaluated their performance on ARIA and MESSIDOR although these datasets publicly released a long time ago. Further, a large number of existing methods are evaluated on the retinal images of adults. The morphological attributes of retinal images of premature infants and children's are entirely different than that of adults. The choroidal vessels, CVR, and other illumination artefacts are more prominently visible in young individuals than the older population (Fraz et al. (2014)). Hence, there is a further scope for developing newer segmentation methods that performs well for both adult as well as paediatric retinal images.

The existing methods in the literature are mainly focused on segmenting large vessel structures leaving behind the thin and low contrast vessels. To the best of our knowledge, only a fewer number of existing studies have addressed the segmentation of thin vessels in fundus images. In this regard, methods in Zhao et al. (2014), Roychowdhury et al. (2015b), Dai et al. (2015), Li et al. (2016), Zhang et al. (2016), Christodoulidis et al. (2016) have made an attempt to specifically address the problem of thin vessel segmentation. Among these methods, the method in Christodoulidis et al. (2016) has achieved a significant state-of-the-art results compared to previous methods for segmentation of thin vessels. But still there exists difficulty in segmenting vessels which lies at extremely low contrast and junction locations.

Current segmentation methods are mainly evaluated on color fundus images. Most of these images are often corrupted by noise, blur, and non-uniform illumination. Hence, it is often difficult for accurate segmentation of thin and low contrast vessel structures. The advancements in newer retinal imaging modalities (like FA, SLO, OCT) and with the integration of adaptive optics (Felberer et al. (2015)) to these modalities has provided a better spatial as well as transverse resolution. These newer imaging modalities provide better visualization of vessels and other microstructures. Further, this opens new research direction in better understanding the structural and physiological changes that affect the retinal vasculature. Some of the newer methods like Perez-Rovira et al. (2011), Abbasi-Sureshjani et al. (2015), Zhang et al. (2016) have made an attempt to evaluate the segmentation performance on FA and SLO images. Further, this provides an opportunity for the researchers to validate their technique on high resolution, high contrast, and finely detailed images.

## 2.8 Summary

In this chapter, a comprehensive review of the recent state-of-the-art retinal vessel segmentation methods is presented. The various complexities and challenges involved in developing robust segmentation techniques have been discussed, including the most crucial image preprocessing steps that have not been addressed earlier in the literature. In addition, the strengths and weakness of each category of segmentation methods with focus to current challenges are discussed in detail. Further, an in-depth quantitative evaluation of state-of-the-art approaches on the individual dataset is assessed. Although many articles have been published on the automated

segmentation of retinal vessels, still there exists room for further improvements. Most of the previous methods have dealt with the fewer images mostly healthier ones, typically segmenting the larger vessel structures and much lower resolution input images. Thus, some of the challenges that remain open to the research community are segmentation in the presence of abnormalities, accurate segmentation of thin vessel structures and segmentation in the presence of non-uniform illumination and various other imaging artefacts.

# Chapter 3

# Automatic Segmentation of Retinal Vascular Structures from Color Fundus Images

Accurate segmentation of retinal vessels from color fundus images play a significant role in the early diagnosis of various ocular, systemic and neuro-degenerative diseases. This chapter presents a novel visual attention guided unsupervised feature learning (VA-UFL) approach, to automatically learn the most discriminative features for segmenting vessels in retinal images. The proposed VA-UFL approach captures both the knowledge of visual attention mechanism and multi-scale contextual information, to selectively visualize the most relevant part of the structure in a given local patch. This allows to encode a rich hierarchical information into unsupervised filtering learning process to generate a set of discriminative features, that aid in the accurate segmentation of vessels, even in the presence of cluttered background. The proposed method is validated on the five publicly available retinal datasets: DRIVE, STARE, CHASE_DB1, IOSTAR and RC-SLO. The experimental results show that the proposed approach significantly outperforms the state-of-the-art methods in terms of sensitivity, accuracy and area under the receiver operating characteristic curve across all five datasets.

## 3.1 Background

Segmentation and quantification of the retinal vascular tree provide important clinical biomarkers through the analysis of its geometrical properties, that aid in early diagnosis of various diseases such as diabetes (Abramoff et al. (2010)), stroke (Baker et al. (2010)), hypertension (Rosenbaum et al. (2016)), arteriosclerosis (Wong et al. (2001a)), and cerebral small vessel diseases (London et al. (2013)). Changes in vessel geometrical patterns such as width, tortuosity, fractal dimension and branching angle provide an early insight into the progression of the aforementioned diseases. Therefore, an accurate delineation of retinal vascular structures from color fundus image is of interest. Manual segmentation of vessel tree is often tedious, time-consuming and prone to large intra and inter-observer variability. Besides, manual delineation often requires careful interpretation of images, which is very cumbersome and painful especially for large population-based screening programs. Hence, there is a need for the automated segmentation of retinal vessels for accurate quantification of vascular changes, along with the entire disease course.

Segmenting retinal vessel tree is extremely challenging due to multi-scale nature of varying vessel calibre, the presence of strong central vessel reflex, the proximal presence of pathological lesions, close parallel and highly curved vessels and complex crossover regions. In the past decades, many solutions have been proposed with techniques ranging from conventional matched filtering (Chaudhuri et al. (1989)) to more recent convolutional neural network (Fu et al. (2016a); Liskowski and Krawiec (2016)) based approaches. The existing techniques can be broadly divided into two main categories: unsupervised and supervised methods (Srinidhi et al. (2017a)). A thorough discussion of these methods is presented in Chapter 2. In general, most of the existing techniques mainly rely on carefully designed hand-crafted filters to inherently model the local geometrical appearance of vascular structures. For instance, the hand-crafted filters such as Gabor filters (Soares et al. (2006)), multi-scale derivative of Gaussian (Niemeijer et al. (2004)), matched filters (Chaudhuri et al. (1989)), ridge detector (Staal et al. (2004)), line detector (Ricci and Perfetti (2007)), wavelet transform (Zhang et al. (2017, 2016)), moment invariant features (Marin et al. (2011)), first and second order derivatives of Gaussian (Roychowdhury et al. (2015a)), and response of COSFIRE filters (Azzopardi et al. (2015); Strisciuglio et al. (2016)), to name a few. These hand-crafted filters were designed based on complex domain

knowledge and require careful parameter tuning to achieve optimal segmentation performance, across a wide variety of data. Besides, the response of these filters often poorly represents the appearance of thin vessel structures, crossover and bifurcation regions, highly curved tortuous vessels and susceptible to non-illumination present in an image. The other drawback of these methods is that the extracted features often lack the discriminative capability to predict the actual class label, even in the presence of similar looking cluttered objects.

To address the aforementioned limitations, many automatic feature learning algorithms (Fu et al. (2016a); Liskowski and Krawiec (2016); Maninis et al. (2016); Wu et al. (2016)) have been proposed to learn the feature representations directly from the training data. These approaches are primarily motivated by the success of deep learning (DL) methods, which is applied in various computer vision applications such as object recognition, scene classification, semantic segmentation etc. The success of these methods is critically dependent on an enormous amount of labelled training data, which is typically expensive in medical imaging applications. To address these shortcomings, many unsupervised feature learning (UFL) algorithms (Coates et al. (2011); Hinton et al. (2006); Lee et al. (2007); Ranzato (2014); Vincent et al. (2008)) have been proposed to automatically learn the feature representations, only from a set of unlabelled data. Automatic feature learning enables to encode rich hierarchical information that learns to map complex functions from input to output, directly from the data, without depending on hand-crafted features.

The main bottleneck for the accurate segmentation of retinal vessels comes from the multiscale nature of varying vessel calibre, poor visibility of low contrast thin vessels, the proximal presence of pathological structures and poor vessel connectivity at complex junction locations. The automatic features learned from these challenging locations often lack discriminative capability in accurately identifying vessel pixels from similar-looking cluttered objects. This is mainly because, the traditional UFL approaches encodes feature representation from a limited input patch size, which is often referred to as "receptive field". The selection of this receptive field mainly depends on the object of interest, which is to be encoded. In the case of retinal vessels, the structure of interest varies significantly, resulting in difficulty in choosing an appropriate receptive field size, that fits for a wide range of input data. Further, various size of receptive fields encodes different *contextual* information (inter-pixel dependencies), thereby resulting in different feature representation for the same pixel

centred on a patch.

To alleviate this problem, a novel UFL approach is proposed, which is primarily inspired by the visual attention mechanism in the human visual system. This is based on the idea that, humans can pay attention *selectively* to the part of the image, instead of processing the whole scene in its entirety (Borji and Itti (2013); Cheung et al. (2016)). Such a selection mechanism is often referred to as "visual attention prediction". In this work, the idea of a visual attention mechanism is leveraged with the unsupervised feature learning approach, to automatically learn the most relevant hierarchical features from unlabeled data. The proposed visual attention guided unsupervised feature learning (VA-UFL) approach automatically learns to selectively pay attention to the most pertinent part of the structure in a given input patch, and use this information to selectively encode features for subsequent classification. The feature learned via this approach offers an edge over tradition UFL methods, by exploring both the notion of *selection mechanism* and the *multi-scale contextual information*, under a single framework. This allows to learn the most relevant hierarchical features at multiple scales, which encodes sufficient information about the most salient region, as well as the multi-scale context for a given input patch.

The key contributions of this chapter can be summarized as follows.

i. The idea of visual attention mechanism is explored, which learns to selectively pay attention to the most relevant structure and capture multi-scale contextual information, from a given local patch. This, in turn, drives the subsequent UFL approach to encode the rich hierarchical information for the automatic feature discovery in retinal images.

ii. The proposed visual attention model, when leveraged with the UFL approach, aims to automatically learn the most discriminative set of features that underscores the interclass differences (between the vessel and background pixels) to be large, while keeping intra-class differences (between vessel pixels) to be small, resulting in the accurate segmentation of retinal vessels.

iii. The proposed approach is validated on five publicly available retinal datasets, including both the RGB and SLO images. The extensive experimental analysis demonstrates the effectiveness of the proposed approach in handling all the challenging cases, compared with the state-of-the-art methods.

The rest of this chapter is organized as follows. Section 3.2, explains in detail

the proposed framework. In Section 3.3, the detailed information about the dataset used, quantitative measures employed, parameter settings, followed by experimental results are provided. The key findings, followed by future directions are discussed in Section 3.4. Finally, the chapter summary is presented in Section 3.5.

## 3.2 Proposed Method

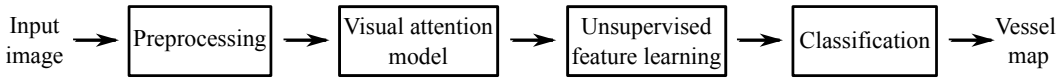
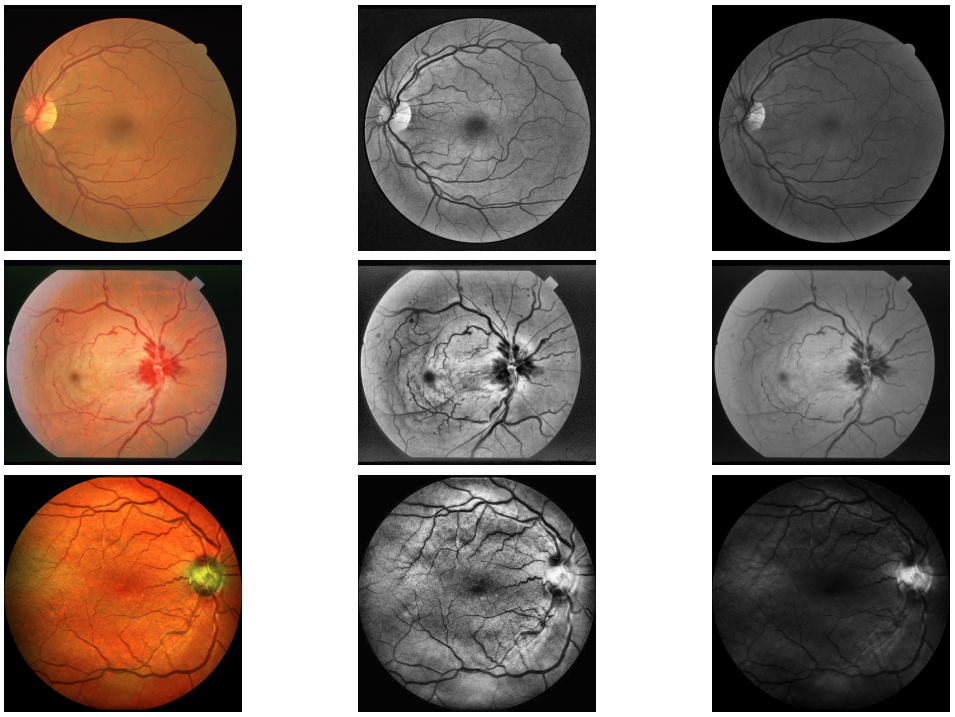



Figure 3.1: An overview of the proposed retinal vessel segmentation approach.

The overall framework of the proposed retinal vessel segmentation approach is illustrated in Figure 3.1. Given an image, the image preprocessing steps are first applied to remove the non-uniform illumination and contrast variability. Next, the input patches are extracted at random locations from the preprocessed image. Each input image patch is modelled with a visual attention mechanism by applying a retinal transformation. A filter bank/dictionary is then trained using these patches based on K-means clustering. Given the learned filter bank and a set of labelled training images, the features are obtained corresponding to each input patches. These features are then trained using random forest (RF) (Breiman (2001)) classifier to predict the label of an unknown patch from the test image. The details are presented next.

### 3.2.1 Preprocessing

Retinal fundus image often exhibits non-uniform illumination acquired due to the geometrical properties of the retinal surface and various other complex imaging conditions such as the presence of pathologies, pupil dilation and involuntary eye movement etc (Joshi and Sivaswamy (2008)). This greatly influences the performance of the segmentation methods especially for the region close to the periphery of the retina, which often exhibits many false detections. Further, luminosity and contrast variability pose a significant problem for the accurate delineation of thin vessel structures. Therefore, in this work, the method proposed in Joshi and Sivaswamy (2008) is adopted for the retinal image enhancement, which is based on the geometrical property of the retinal surface. Compared to the most widely used method in the literature (Foracchia et al. (2005)), the approach in Joshi and Sivaswamy (2008)



(a) (b) (c)

Figure 3.2: Visualization of image pre-processing steps. (a) An example image; (b) Enhanced image after applying the contrast limited adaptive histogram equalization (CLAHE) algorithm; (c) Enhanced image after applying technique proposed in Joshi and Sivaswamy (2008). The images shown belongs to the DRIVE (first row), STARE (second row) and IOSTAR (third row) datasets.

utilizes the non-uniform sampling scheme on the polar grid to better estimate the degradation component of the acquired image. A sample visualization of image pre-processing steps is shown in Figure 3.2. The approach is shown to be robust in enhancing the thin vessel structures, which subsequently improves the vessel segmentation performance. In the experiments, only the green channel of the retinal fundus image is considered, since it exhibits a better contrast between vessel and background when compared to the other channels (Azzopardi et al. (2015); Mendonca and Campilho (2006); Niemeijer et al. (2004); Soares et al. (2006); Zhang et al. (2016)).

### 3.2.2 Visual Attention Modelling

Given an image, the goal is to classify a patch centred on a pixel of interest, as belonging to a vessel or background. The patches $X = \{x_1, x_2, ..., x_m\}$ are sampled at random locations from an image, as shown in Figure 3.3. For each patch $x \in \mathbb{R}^p$, a binary label $y \in \{0, 1\}$ is assigned, where 0 – denotes a background pixel; and 1

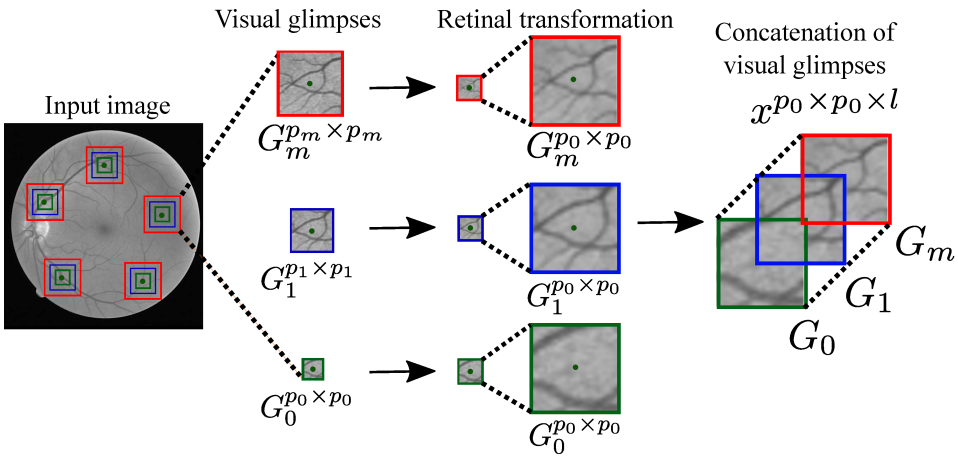



Figure 3.3: An illustration of visual attention modelling. *Note:* the center green dot marks the pixel of interest $(i, j)$.

– denotes a vessel pixel. Let $Y = \{y_1, y_2, ..., y_m\}$ be the corresponding labels of the input data $X$.

The input patch $x \in \mathbb{R}^p$ centred on a pixel of interest $(i, j)$ is modelled by constructing visual glimpses $G(i, j) = \{G_0, G_1, ..., G_m\}$. The visual glimpses consists of multi-scale patches $G_0, G_1, ..., G_m$ having dimension $p_0 \times p_0, p_1 \times p_1, ..., p_m \times p_m$ pixels, respectively, such that the dimension of the patch ($G_0 << G_1 << ... << G_m$), as shown in Figure 3.3. Then a retinal transformation $T$ is applied for each visual glimpse $G$, to model the visual attention mechanism.

A Retinal transformation $T$ is similar to fovea of the human retina, where it encodes a higher spatial resolution at a region close to the centre (pixel of interest) and progressively lower resolution as one moves away from the centre, in a non-linear fashion. This transformation $T$ extracts a high resolution information around the pixel of interest $(i, j)$, but encodes a low resolution information for pixels further from $(i, j)$, as shown in Figure 3.3. This lower resolution representation is often referred to as "visual glimpse" (Larochelle and Hinton (2010)).

The retinal transformation $T$ is applied to visual glimpses $G$ as: $x = T(G(i, j)) = T(G_0^{p_0 \times p_0}, G_1^{p_1 \times p_1}, ..., G_m^{p_m \times p_m})$ (where, $x \in \mathbb{R}^p$ is the input patch centred on a pixel of interest $(i, j)$), which is computed by down-sampling each multi-scale patches to the dimension equal to $(p_0 \times p_0)$ pixels using bi-cubic interpolation. Where, the glimpse $G_0^{p_0 \times p_0}$ – represents the patch with highest spatial resolution, and $G_1^{p_0 \times p_0}, G_2^{p_0 \times p_0}, ..., G_m^{p_0 \times p_0}$ – corresponds to the successive lower resolution patches for location $(i, j)$, as shown in Figure 3.3. The output of retinal transformation $T$ results in a single input patch $x^{p_0 \times p_0 \times l}$, which is obtained by concatenating the vi-

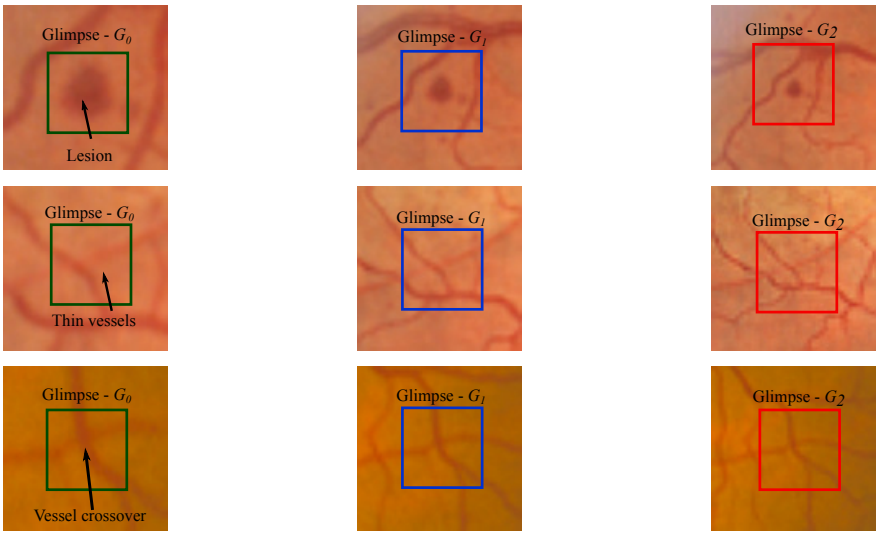



Figure 3.4: Visualization of sample visual glimpse patches containing different retinal structures of interest (from top to bottom: retinal lesion (first row), area surrounded by thin vessels (second row) and a crossover location (third row). The object of interest is positioned at the centre pixel of a patch). (a) The higher resolution patch ($20 \times 20$ pixels); (b) and (c) Corresponds to the progressive lower resolution patches, obtained by first extracting a higher resolution patch of size $40 \times 40$ and $60 \times 60$ pixels respectively, followed by down-scaling to $20 \times 20$ pixels.

sual glimpses $G = \{G_0^{p_0 \times p_0}, G_1^{p_0 \times p_0}, ..., G_m^{p_0 \times p_0}\}$; where, $l = 0, 1, ..., m$ – denotes the number of levels of visual glimpse $G$ for a pixel location $(i, j)$.

The intuition behind the proposed approach is that, when humans tend to focus selectively on a particular object/scene, the region centred to the eye fixation point will exhibit higher spatial resolution than the area surrounding the object of interest (Borji and Itti (2013)). The region centred to the eye fixation is regarded as the most critical part of a scene (salient region) while, the region surrounding the object of interest often provides contextual information (inter-object relationship), which aid in accurate localization/classification of an object in the presence of clutter. This allows to draw conclusion that at what level of spatial context, the extracted features contribute to the true label prediction of a patch centred on a pixel of interest.

The visualization of sample visual glimpse patches for different retinal regions of interest is shown in Figure 3.4. The goal is to classify the object of interest which is the centre pixel of a patch, as either belonging to a vessel or background. The contextual information present in the region of interest (ROI) (see Figure 3.4 (a)) often lack the discriminative capability in predicting the true class label which is often surrounded by similar looking cluttered background. For instance, a dark lesion

such as microaneurysm (see the first row of Figure 3.4) which usually appears near the region of thin vessels, exhibits the similar visual appearance of high-curvature and junction points of small vessels. Thus, posing a significant challenge in predicting the actual class label, given the limited neighborhood information. While, the patches (see the first row of Figure 3.4 (b) and Figure 3.4 (c)) provides much larger contextual information in which a lesion appears to be at the isolated location, often not connected to any vessel fragments.

Similarly, the region consisting of low contrast thin vessels, large and small vessel crossings, poor vessel connectivity at complex junction locations (see the second and third row of Figure 3.4), often hinder the segmentation performance mainly for the detection of thin vessels. In contrast, by taking into account the most salient region - encoded by higher resolution glimpse ($G_0$) and the contextual information - encoded by lower resolution glimpses ($G_1, G_2$), provides the most discriminative features by encoding the knowledge of both inter-structure relationship as well as the most salient region, in a given local patch. Thus, facilitating the subsequent classifier in accurately predicting the vessel pixels, even in the presence of similar looking background structures.

### 3.2.3 Visual Attention Guided Unsupervised Feature Learning (VA-UFL)

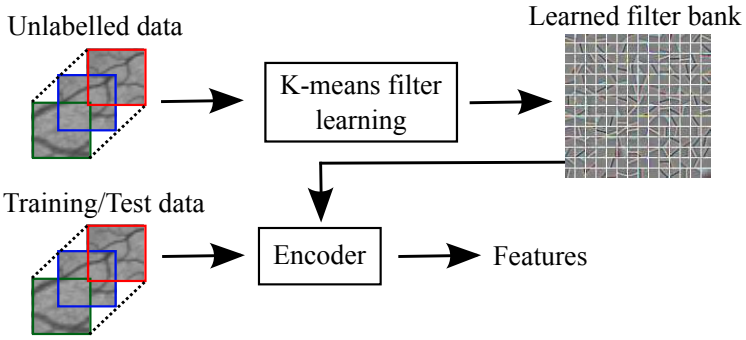



Figure 3.5: An overview of the visual attention guided unsupervised feature learning (VA-UFL) framework.

The goal of any UFL algorithm is to learn a good latent representation from only unlabelled data $U$. Given an input patch $x$, the aim is to train a dictionary $D$, which encodes the knowledge of the distribution of $x$. This trained dictionary $D$ is often referred to as learned filter bank/weights. Next, an encoder $E(x; D) = s$ is defined, that maps an input patch $x \in \mathbb{R}^p$ to a new representation $s \in \mathbb{R}^K$ where,

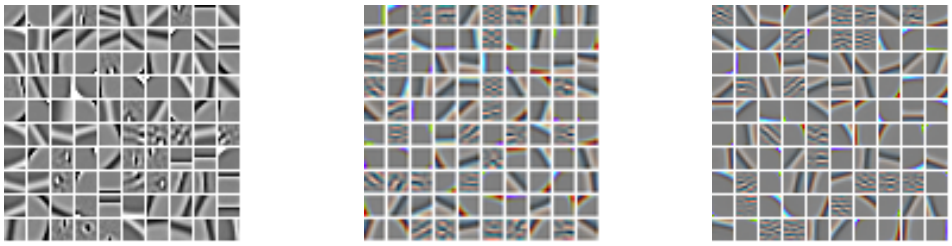
$s$ is known as latent representation or feature vector. The feature vector ($s$) along with their true label can be passed to any machine learning classifier to predict the label of an unknown patch from $s$. The overall framework of the UFL approach is illustrated in Figure 3.5.

Off-the-shelf, several unsupervised feature learning algorithms have been proposed in the past, such as sparse coding (Lee et al. (2007)), sparse auto-encoders (Ranzato et al. (2007)), restricted Boltzmann machines (RBMs) (Hinton et al. (2006)), denoising auto-encoders (Vincent et al. (2008)), K-means clustering (Coates et al. (2011)), deep belief networks (DBN) (Ranzato et al. (2008)), and many others as well. Among the approaches, K-means clustering (Coates et al. (2011)) has been successfully applied for unsupervised filter learning, often competing with the state-of-the-art techniques, due to its speed and scalability. In this work, the unsupervised learning module is adopted based on K-means clustering (Coates et al. (2011)) for extracting a discriminative feature set, for the task of retinal vessel segmentation. The main reason behind choosing the K-means clustering as unsupervised filter learning is of two folds: K-means tends to learn the sparse projections of the input data - (i) given a sufficiently large amount of training images corresponding to the input dimensionality of the image; (ii) apply whitening to remove correlation between the data points. The above two assumptions hold in the case where, there is an abundant amount of unlabelled data which is a typical scenario in the medical imaging domain, due to its expensive manual annotations. The major advantages of K-means over other UFL approaches are its speed, scalability and no hyper-parameter tuning involved.

Given a set of randomly sampled visual glimpse patches $x$, a dataset $X = \{x_1, x_2, ..., x_m\}$ is constructed, where $x \in \mathbb{R}^p$. Having obtained input data $X$, the data preprocessing is performed followed by unsupervised filter learning to learn a Dictionary $D$.

#### 3.2.3.1 Data Preprocessing

It is standard practice in conventional deep learning frameworks to carry-out local brightness and contrast normalization, before generating features from the input data. Two types of preprocessing steps are generally performed: global contrast normalization and zero-phase component analysis (ZCA) whitening (Bell and Sejnowski (1997)).



(a) (b) (c)

Figure 3.6: Example bases learned using K-means for varying level ($l$) of input visual glimpses. (a) Visual glimpse, with $l = 0$; (b) Visual glimpse, with $l = 1$; (c) Visual glimpse, with $l = 2$.

In global contrast normalization, the local brightness and contrast normalization are performed by normalizing every patch $x$ by subtracting the mean and dividing by the standard deviation of its elements. After normalization, the patch level whitening is performed using ZCA transform to remove the linear correlations among neighbouring pixels. The ZCA transformation is summarized as follows. Given a normalized data $X$:

1. find $\mu$ and $\Sigma$, such that

$$\mu = \text{mean}(X) \text{ and } \Sigma = \text{cov}(X) = \frac{1}{n}\sum_{i=1}^{n} XX^T; \tag{3.1}$$

2. find eigenvalue and eigenvector, such that

$$\sum = V\Lambda V^T; \tag{3.2}$$

3. finally, the patch level whitening is computed as

$$x_{ZCA} = V(\Lambda + \epsilon_{ZCA}\mathbb{I})^{-1/2}V^T x, \tag{3.3}$$

where, $\epsilon_{ZCA}$ is a small constant controlling the trade-off between whitening and noise amplification. For simplicity, $x_{ZCA}$ is denoted as $x$, which is used as an input to the rest of the pipeline.

#### 3.2.3.2 K-means Filter Learning

Given the whitened data $x$, a reconstructive dictionary $D \in \mathbb{R}^{p\times K}$ is learned with $K$ elements or atoms, which is accomplished by solving the following optimization

problem

$$< D, s >= \arg\min_{D,s} \|x - Ds\|_2^2 \quad s.t. \|s\|_0 \leq 1, \tag{3.4}$$

where, $\|x - Ds\|_2^2$ denotes the reconstruction error, $s \in \mathbb{R}^K$ is the code vector corresponding to the input $x$, and $D = [d_1, d_2, ..., d_K] \in \mathbb{R}^{p \times K}$ ($K > p$, makes the dictionary overcomplete). Here each column of the dictionary are the centroids learned by K-means. A sample visualization of the learned dictionary (filter bank) for varying level ($l$) of input visual glimpses ($G$) is shown in Figure 3.6. It is observed that the higher levels of input visual glimpses produce more discriminative visual features such as, sharp vessel edges (see Figure 3.6 (b) and Figure 3.6 (c)) compared to the noisy filter bases learned without attention modelling, as shown in Figure 3.6 (a).

#### 3.2.3.3 Encoder

An encoder is a function $E(x; D) = s$, which is defined as a non-linear mapping function that transforms a new input patch $x_i \in \mathbb{R}^p$ to a latent representation $s_i \in \mathbb{R}^K$ with $D$ fixed. Where, $s_i$ is the feature vector corresponding to a new input patch $x_i$.

More concisely, given $D$ the learned filter bank using K-means, the latent representation $s_i$ is found for a new input patch $x_i$ by solving the following optimization problem

$$s_i = \arg\min_{s} \|x_i - Ds\|_2^2 + \lambda \|s\|_1, \tag{3.5}$$

where, $\lambda$ is a parameter that controls the reconstruction error and sparsity. The following $L_1$ optimization can be solved efficiently using approaches (Gregor and LeCun (2010); Lee et al. (2007)). Both the positive and negative components of sparse code $s$ are considered into separate features as: $s_K = max\{0, s_K\}$ and $s_{K+n} = max\{0, -s_K\}$. Note: the resulting feature vector dimension is $s \in \mathbb{R}^{2K}$. The overall steps of the proposed VA-UFL algorithm is summarized in Algorithm 1.

### 3.2.4 Classification

To distinguish between the vessel and non-vessel pixels, a RF classifier is used, due to its capability of performing both classification and feature selection implicitly. It is robust against overfitting, outliers and high dimensional imbalanced data. This is true in this case, where only $9 - 14\%$ of the total pixels belongs to vessels while, the

**Algorithm 1:** Visual attention guided unsupervised feature learning (VA-UFL) approach.

**Input :** $U$ – set of unlabelled images;
$x$ – the input patch centred on a pixel of interest;
$l$ – number of levels of visual glimpse;
$K$ – number of atoms in the dictionary.

**Output:** $D$ – learned filter bank/dictionary;
$s_i$ – feature vector for a new input patch $x_i$.

1. Extract input patches $x$ at random locations from a set of unlabelled images $U$;
2. Apply retinal transformation $T$ to each input patch $x$ to construct visual glimpses $G$ of $l$ levels;
3. Pre-process the patch by normalizing brightness and contrast, as well as ZCA whitening defined in Equation (3.3);
4. Train a dictionary $D$ using K-means clustering defined in Equation (3.4);
5. Compute the feature vector $s_i$ for a new input train/test visual glimpse patch $x_i$ via sparse encoding defined in Equation 3.5;

**return** $\boldsymbol{D}$,$\boldsymbol{s_i}$.

rest belongs to non-vessels, leading to a highly skewed dataset.

A RF is a combination of $N_T$ decision trees which are trained independently using bootstrap samples drawn with replacement from the training set. Each node in a tree is split using a randomly selected subset of $m$ features ($m = \sqrt{d}$) (where $d$ is the dimensionality of the feature vector), which is chosen according to the decrease in the Gini index as recommended in Breiman (2001). The RF returns the probability of being a vessel or non-vessel for each selected feature input, based on the majority of the trees returning a positive response. The default setting for the classifier is the number of trees $N_T = 100$ trees; and the feature vector dimension $d = (2 \times K) = (2 \times 100) = 200$ (where $K$ is the number of dictionary atoms).

During training, given a set of labelled training images $X_{(i)}$, $Y_{(i)}$, $i = 1, 2, ..., m$; where, $X_{(i)}$ are the training images; and $Y_{(i)} \in \{0, 1\}$ are the corresponding class labels (where, 0 - indicates a background pixel, and 1 - indicates a vessel pixel); the goal is to classify a patch $x_{(i)}$ centred on a pixel of interest as either vessel or background. In order to do this, the visual glimpse patches $x_{(i)}$ are first extracted for every pixels, from each of the training images $X_{(i)}$. Given the learned dictionary $D$ using unlabelled images $U$, the features $s_{(i)}$ are extracted corresponding to each input visual glimpse patch $x_{(i)}$, using the encoder defined in Section 3.2.3.3. The feature set $s_{(i)}$ corresponding to each input patch $x_{(i)}$ along with their true label $y_{(i)}$ is used to train a RF classifier of $N_T$ trees. In the testing stage, the probability of a

patch $P(vessel = 1|s_{(i)})$, $\forall(x_{(i)}, y_{(i)})$ that belongs to a set of test image is estimated by feeding into the trained RF classifier.

## 3.3 Experiments and Results

### 3.3.1 Datasets

To validate the proposed approach, five publicly available retinal datasets were used: DRIVE (Staal et al. (2004)), STARE (Hoover et al. (2000)), CHASE_DB1 (Fraz et al. (2012c)), and the two new Scanning Laser Ophthalmoscopy (SLO) image datasets namely: IOSTAR (Zhang et al. (2016)), and RC-SLO (Zhang et al. (2016)).

The DRIVE dataset consists of 40 color fundus images with a resolution of $565 \times 584$ pixels ($px$). The dataset is divided into training and test sets, each of which contains 20 images. The STARE dataset includes 20 fundus images of resolution $700 \times 605$ $px$, out of which 10 contains pathological signs. A leave-one-out cross-validation was performed on STARE. The CHASE_DB1 contains 28 fundus images of resolution $1280 \times 960$ $px$, out of which first 20 images are used for testing and the last 8 images for training (Fraz et al. (2012c)). Manually segmented binary vessel maps are provided by two human annotators for DRIVE, STARE and CHASE_DB1 datasets. Besides, the proposed method is also validated on the two new publicly available: IOSTAR and RC-SLO datasets. The IOSTAR consists of 30 - SLO images of resolution $1024 \times 1024$ $px$ and the RC-SLO contains 40 image patches of resolution $360 \times 320$ $px$. All the vessel pixels in both the datasets are annotated by a group of experts. Half random split was employed for both IOSTAR and RC-SLO datasets (Abbasi-Sureshjani et al. (2015); Zhang et al. (2016)). The performance metrics of the proposed segmentation approach are computed by considering all the pixels within the field of view (FOV).

### 3.3.2 Evaluation Metrics

In-order to quantitatively compare the performance of the binary segmentation results with the corresponding manual ground truths, the five different performance measurements are obtained based on: the number of true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN). The following are the evaluation metrics that were used to compare the proposed approach with the state-of-the-art vessel segmentation methods: *Sensitivity* ($Se$) (also know as *Recall* ($Re$)),

*Specificity* ($Sp$), *Accuracy* ($Acc$), *Matthews correlation coefficient* ($MCC$), *F1-score* ($F1$) and *Area under the ROC curve* ($AUC$).

$$Se = \frac{TP}{TP + FN},\ Sp = \frac{TN}{TN + FP},\ Acc = \frac{TP + TN}{N},$$
$$MCC = \frac{TP/N - S \times P}{\sqrt{P \times S \times (1 - S) \times (1 - P)}},\ F1 = \frac{2 \times Pr \times Re}{Pr + Re}, \tag{3.6}$$

where, N = TN + TP + FN + FP, S = (TP + FN) / N, P = (TP + FP) / N and Pr = TP / (TP + FP).

The $MCC$ and *F1-score* are often used in the performance analysis of segmentation result on class imbalance datasets. This is true in case of retinal images, where only a small portion of pixels belongs to vessels (only around 9% - 14% vessel pixels), while the others are counted as background pixels. The $MCC$ returns a value between -1 and +1, with +1 indicating a perfect prediction, and -1 a completely incorrect prediction. Finally, the receiver operating characteristic (ROC) curves are computed with the true positive fractions ($Se$) versus false positive fractions ($1 - Sp$) by varying the threshold on the probability map. The $AUC$ is calculated to quantify the performance of the segmentation algorithm, where the $AUC$ value of 1 indicates a perfect segmentation.

### 3.3.3 Parameters Setting

The proposed retinal vessel segmentation approach mainly depends on the three main parameters: $p_0$ – (the size of input visual glimpse patch $x$), $l$ – (the number of levels of visual glimpse) and $K$ – (the number of dictionary atoms (learned filter size)). These parameters play a critical role in obtaining the optimal performance of retinal vessel segmentation approach. The details are presented next.

In the experiments, $K = 100$ filters are chosen by randomly sampling 100000 patches from the training sets of three datasets namely - DRIVE, STARE and CHASE_DB1. Figure 3.7 (a) shows the effect of varying the input patch size ($p_0$), without visual attention mechanism ($l = 0$), on the accuracy of segmentation performance. It is observed that the input patch size of $(9 \times 9)$ $px$ gave the best results, across all three datasets of varying image resolutions. This is because the selected patch size can capture the varying nature of vessel calibre (such as thin and thick vessels) across multiple scales and multiple orientations. Further, the chosen patch size is found to be strongly dependent on average vessel calibre across each dataset.

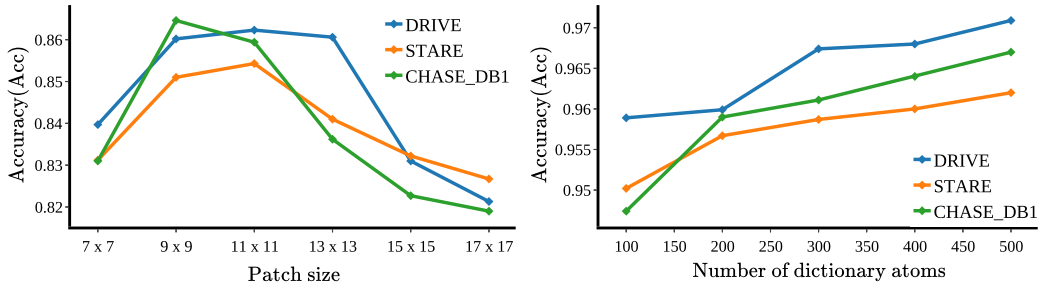



Figure 3.7: Influence of parameter selection. (a) Classification accuracy vs. varying input patch size ($p_0$); (b) Classification accuracy vs. varying number of learned filters $K$.

Table 3.1: Performance analysis for varying levels ($l$) of visual glimpse. Note: the input patch with $l = 0$ is the one without having any visual attention mechanism.

| Datasets | $l$ | $Se$ | $Sp$ | $Acc$ | $AUC$ | $MCC$ | $F1$ |
|---|---|---|---|---|---|---|---|
| DRIVE | 0 | 0.7665 | 0.9088 | 0.8602 | 0.8990 | 0.6444 | 0.6737 |
| | 1 | 0.8387 | 0.9296 | 0.9011 | 0.9405 | 0.6886 | 0.7077 |
| | 2 | **0.8644** | **0.9667** | **0.9589** | **0.9701** | **0.7421** | **0.7607** |
| STARE | 0 | 0.7894 | 0.8832 | 0.8510 | 0.9010 | 0.6391 | 0.6442 |
| | 1 | 0.8134 | 0.9178 | 0.8945 | 0.9372 | 0.6732 | 0.6818 |
| | 2 | **0.8325** | **0.9746** | **0.9502** | **0.9670** | **0.7398** | **0.7698** |
| CHASE_DB1 | 0 | 0.7576 | 0.8815 | 0.8646 | 0.9108 | 0.6485 | 0.6695 |
| | 1 | 0.7823 | 0.9297 | 0.9173 | 0.9303 | 0.6653 | 0.7012 |
| | 2 | **0.8297** | **0.9663** | **0.9474** | **0.9591** | **0.6927** | **0.7189** |
| IOSTAR | 0 | 0.7640 | 0.9074 | 0.8543 | 0.9094 | 0.6536 | 0.6831 |
| | 1 | 0.7853 | 0.9393 | 0.9236 | 0.9453 | 0.6739 | 0.7032 |
| | 2 | **0.8269** | **0.9669** | **0.9564** | **0.9663** | **0.7057** | **0.7354** |
| RC - SLO | 0 | 0.7756 | 0.9176 | 0.9056 | 0.9165 | 0.6598 | 0.6854 |
| | 1 | 0.8149 | 0.9432 | 0.9304 | 0.9488 | 0.6812 | 0.7019 |
| | 2 | **0.8488** | **0.9666** | **0.9581** | **0.9678** | **0.7029** | **0.7200** |

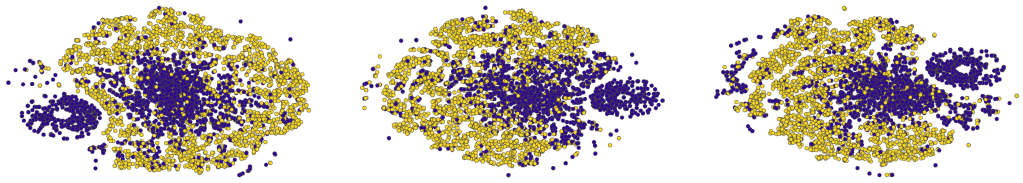
It is experimentally found that, the mean vessel calibre varies from $3.4 \pm 1.6$ $px$ for the DRIVE, $4.4 \pm 2.6$ $px$ for the STARE, $5.4 \pm 3.6$ $px$ for the CHASE_DB1, $6.3 \pm 2.5$ $px$ for the IOSTAR and $5.2 \pm 2.5$ $px$ for the RC-SLO datasets (Zhang et al. (2016)). Hence, the chosen patch size is well within the range of average vessel calibre across all five datasets and thus exhibits an improved performance for a fixed patch size ($p_0$), even with varying image resolutions across datasets.

Next, to assess the influence of visual attention mechanism on filter learning, two levels ($l$) of visual glimpse are adopted: $G_1, G_2$ where, $l = 2$; on a fixed input patch ($G_0$), which is of size $p_0 = (9 \times 9)$ $px$. Table 3.1 depicts the vessel segmentation performance for varying levels of the visual glimpse, across all five datasets.

It is observed that there is a significant boost in the segmentation performance, with an average $Se > 10/5/7/6/7\%$, $Sp > 6/9/8/6/5\%$, $Acc > 10/10/8/10/5\%$, $AUC > 7/7/5/6/5\%$, $MCC > 10/10/4/5/4\%$, and $F1 > 9/13/5/5/4\%$ across DRIVE/STARE/CHASE_DB1/IOSTAR/RC-SLO datasets, respectively. This is mainly because the higher level of visual glimpses learns to selectively pay attention to the most relevant part of a scene (such as vessel structures) in a given local patch. Thus providing a discriminative capability in predicting the true class label ($TP's$), even in the presence of noisy and cluttered background. The ability of the proposed approach to correctly identify the true positives is clearly reflected in the values of $Se, Acc, MCC, F1$, with an average increase in 4% across all five datasets, when compared with the performance obtained without any visual attention mechanism ($l = 0$ in Table 3.1). Further, there is a boost in the values of all performance metrics of $\approx 4\%$ at every level ($l$) of the visual glimpse, clearly indicating an influence of visual attention mechanism on vessel segmentation performance.

Thus in all the experiments, the input patch size of $(9 \times 9)$ $px$ with $l = 2$ levels of visual glimpse is empirically chosen for evaluating the segmentation performance against the state-of-the-art methods. This is because, the higher levels of visual glimpses will generally result in improved vessel segmentation performance (as shown in Table 3.1). But this improved performance often comes at the price of a slower prediction time with increasing levels of visual glimpse, due to increase in the dimensionality of input patch $x$, which is centred on each pixel of interest $(i, j)$ to be classified as vessel or background. Hence the value $l$ is restricted to 2, to balance between the segmentation performance and the computational complexity.

The segmentation performance is also evaluated by varying the number of dictionary atoms ranging from $K = \{100, 200, 300, 400, 500\}$, as shown in Figure 3.7 (b). For this experiment, 100000 patches of size $(9 \times 9)$ $px$ are randomly sampled with $l = 2$ glimpses for learning the filter size $K$. It is shown that the proposed segmentation approach achieves an average $Acc > 0.95$, when chosen filter size $K = 500$. But this improved segmentation accuracy comes at the price of a slower prediction time, which is inappropriate for a standard clinical setting. In the experiments, it is observed that there is a very marginal improvement in the $Acc$ of the segmentation performance between the filter size of $K = 100$ to 500. Hence, the size of the filter is empirically set as $K = 100$ throughout the experiments, which offer a good compromise between training/testing speed and segmentation performance.



(a) (b) (c)

Figure 3.8: Feature embedding visualization via (t-SNE) for subset of observations obtained from test sets of all five datasets, with varying levels ($l$) of visual glimpse. (a) Visual glimpse with $l = 0$; (b) Visual glimpse with $l = 1$; (c) Visual glimpse with $l = 2$. The violet and yellow colors correspond to the vessel and background, respectively. Best viewed in color.

Besides, there are few other parameters such as $\epsilon_{ZCA}$ which is set at 0.1, that provides a good balance between filter sharpness and noise amplification. The sparsity parameter $\lambda$ is chosen to be 1, which yield best results on all datasets at fairly faster convergence rate. A RF classifier of $N_T = 100$ trees is trained using a $2 \times K = 200$ dimension feature vector. Further, increasing the number of trees have shown to result in better segmentation performance at a cost of increased time complexity (Breiman (2001)). Hence, the RF of 100 trees is used for classification across all five datasets. The performance analysis of vessel segmentation vs. different classifiers is also shown in Appendix A.

### 3.3.4 Visualization of Discovered Features

In this section, in-depth experimental validation of the central hypothesis of the proposed approach is presented, by visualizing the effect of visual attention mechanism on unsupervised filter learning. The learned projections are visualized using a fast implementation of t-distributed Stochastic Neighbor Embedding (t-SNE) (Van Der Maaten (2014)) using the default set of parameters. This technique is adopted due to its extensive ability to preserve neighborhoods and clusters in lower dimensional subspace, compared to other dimensionality reduction techniques such as PCA, Isomap etc.

For visualizing the learned features, a random subset of 100000 observations is selected from test sets of all five datasets, due to the extensive computational complexity of t-SNE. The features are extracted from an input visual glimpse patch of size $(9 \times 9)$ $px$ with $l = 2$ glimpse, to mimic the visual attention mechanism, as described in Section 3.2.2. Figure 3.8 depicts the projection of a subset of observations for varying levels of visual glimpses. It is observed that, there is a clear

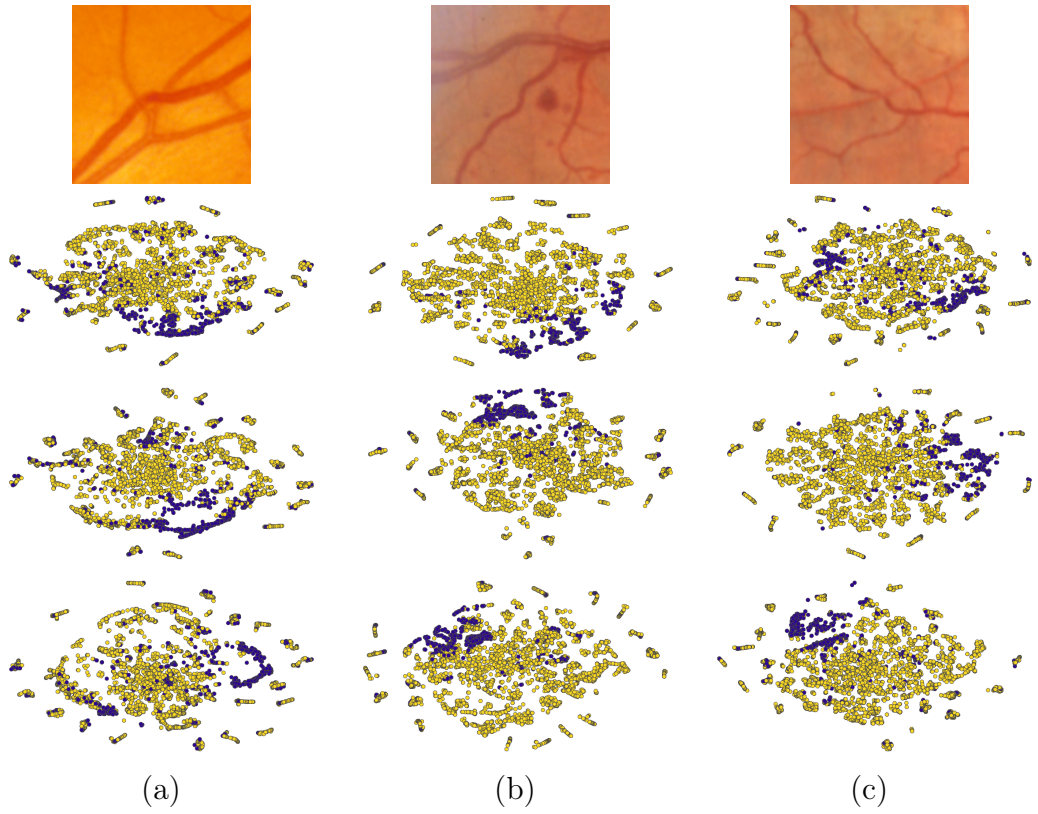



Figure 3.9: Feature embedding visualization via (t-SNE) of sample ROI's (containing (a) complex crossover structures; (b) pathology; (c) thin vessels) for varying levels ($l$) of visual glimpse. First row: original image; Second row: visual glimpse with $l = 0$; Third row: visual glimpse with $l = 1$; Fourth row: visual glimpse with $l = 2$. The violet and yellow colors correspond to the vessel and background, respectively. Best viewed in color.

visual separation between the vessel and background classes obtained with $l = 2$ visual glimpse, as shown in Figure 3.8 (c). This infers that the higher levels of visual glimpse generally captures larger contextual information and thereby providing more discriminative features for subsequent class-assignment. Whereas, for the lower levels of visual glimpse (see Figure 3.8 (a) and Figure 3.8 (b)), class separation is clearly inferior, indicating that there is a strong overlap between the vessel and background classes, due to minimal neighborhood information encountered during filter learning process.

To further substantiate the hypothesis, the learned features of sample ROI's are visualized for various challenging cases such as segmentation in the presence of pathology, segmenting thin vessels and complex crossover structures, as shown in Figure 3.9. The learned features via visual attention mechanism have shown to be superior in exhibiting a discriminative class separation between vessel and background

(see second and third row), compared to the one without attention mechanism (see the first row). The proposed approach is also shown to be robust to pathology and thin vessels structures, where higher levels of visual glimpse often aid in predicting the true class label, even in the presence of similar looking cluttered objects (see second and third column). This is consistent with the hypothesis proposed in Section 3.2.2, that the region containing thin vessels and pathology often exhibits similar visual appearance, thereby posing a significant challenge for accurate segmentation of retinal vessels. Hence, visual attention mechanism is shown to capture the most relevant structure (salient region), as well as the neighborhood information (context), across multiple scales in a given local patch. Thus, the visual attention mechanism aims to drive the unsupervised filter learning process to generate a more compact and distinctive feature set for the subsequent classification.

### 3.3.5 Vessel Segmentation Results

The sample qualitative results of the proposed approach are shown in Figure 3.10, on five datasets namely: DRIVE, STARE, CHASE_DB1, IOSTAR and RC-SLO. It is observed that the proposed approach segments most of the thin vessel structures, preserves vessel connectivity at junction locations and also shown to be robust even in the presence of illumination artefacts (see third and sixth row). The improved performance is mainly due to the discriminative capability of unsupervised filter learning, which is combined with the visual attention mechanism to leverage both the salient and contextual information from multiple scales. This hybrid information facilitates to accurately identify vessel pixels from the clutter background (such as lesions), differentiating thin and thick vessels and different crossover vessel structures.

Table 3.2 and Table 3.3, depicts the performance of the proposed approach with the existing state-of-the-art supervised and unsupervised approaches. The proposed method achieves an average $Se$ of greater than 7%, compared with the existing approaches validated on DRIVE, CHASE_DB1, IOSTAR and RC-SLO datasets, and a higher $Se$ value than all the supervised approaches on STARE dataset. This increased $Se$ is mainly because of the discriminative feature set, which takes into account of large contextual information in predicting the true class label from similar-looking background structures. For the DRIVE dataset, a $Acc/AUC$ of 0.9589/0.9701 is obtained, which is significantly higher than all the unsupervised approaches. For the CHASE_DB1 dataset, a $Acc$ of 0.9474 is obtained, which is higher

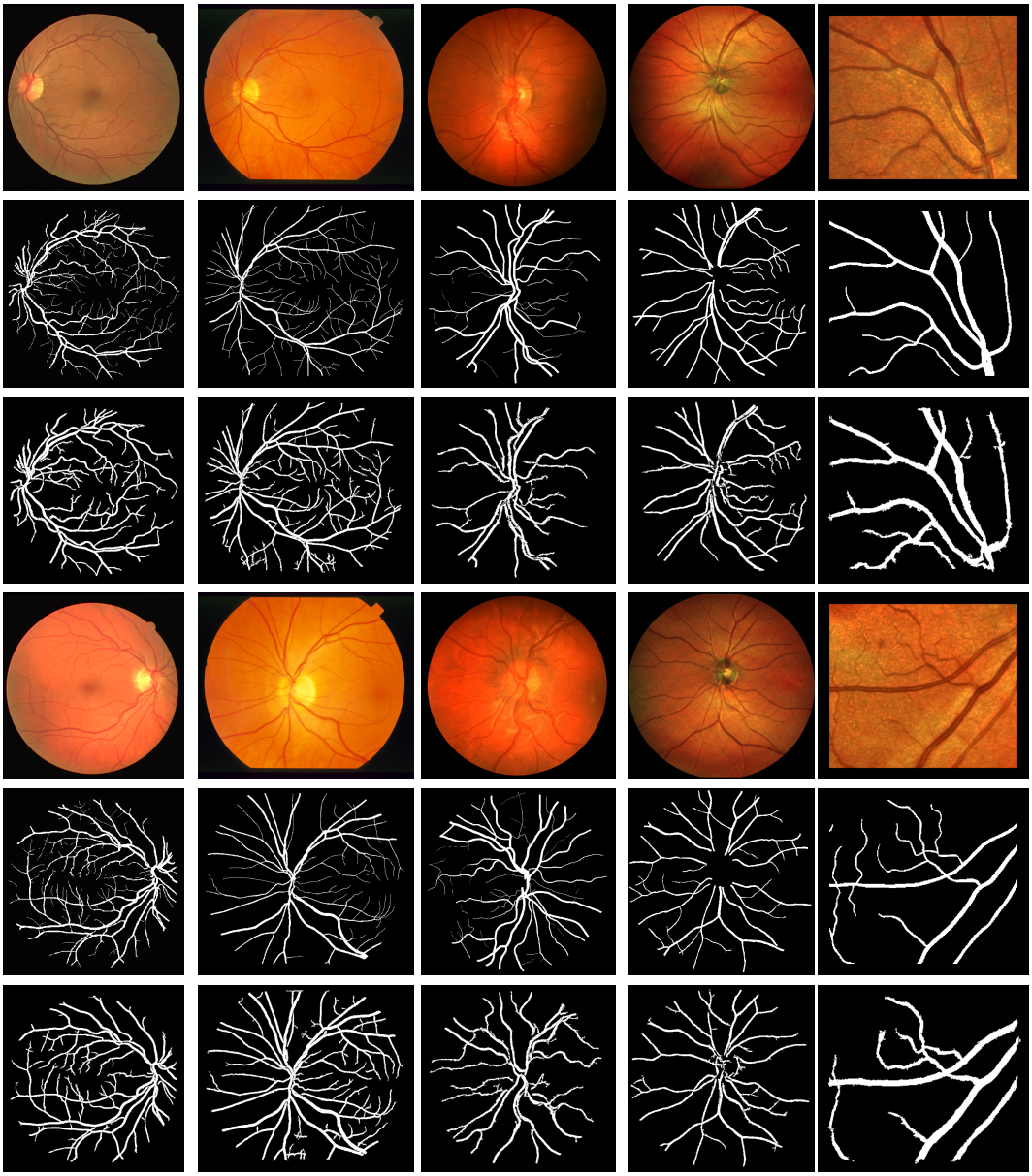



Figure 3.10: Segmentation results obtained on DRIVE (a), STARE (b), CHASE_DB1 (c), IOSTAR (d) and RC-SLO (e) datasets. First and fourth row: original images; Second and fifth row: manual segmentation results; Third and sixth row: segmentation results of the proposed approach.

Table 3.2: Performance comparison on the DRIVE, STARE and CHASE_DB1 datasets.

| | Methods | Year | DRIVE | | | | STARE | | | | CHASE_DB1 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | *Se* | *Sp* | *Acc* | *AUC* | *Se* | *Sp* | *Acc* | *AUC* | *Se* | *Sp* | *Acc* | *AUC* |
| | 2nd human observer | - | 0.7760 | 0.9724 | 0.9472 | - | 0.8952 | 0.9384 | 0.9349 | - | 0.8105 | 0.9711 | 0.9545 | - |
| Unsupervised methods | Mendonca and Campilho (2006) | 2006 | 0.7344 | 0.9764 | 0.9452 | - | 0.6996 | 0.9730 | 0.9440 | - | - | - | - | - |
| | Martinez-Perez et al. (2007) | 2007 | 0.7246 | 0.9655 | 0.9344 | - | 0.7506 | 0.9569 | 0.9410 | - | - | - | - | - |
| | Al-Diri et al. (2009) | 2009 | 0.7282 | 0.9551 | - | - | 0.7521 | 0.9681 | - | - | - | - | - | - |
| | Lam et al. (2010) | 2010 | - | - | 0.9472 | 0.9614 | - | - | 0.9567 | 0.9739 | - | - | - | - |
| | Zhang et al. (2010) | 2010 | 0.7120 | 0.9724 | 0.9382 | - | 0.7177 | 0.9753 | 0.9484 | - | - | - | - | - |
| | You et al. (2011) | 2011 | 0.7410 | 0.9751 | 0.9434 | - | 0.7260 | 0.9756 | 0.9497 | - | - | - | - | - |
| | Fraz et al. (2012a) | 2012 | 0.7152 | 0.9759 | 0.9430 | - | 0.7311 | 0.9680 | 0.9442 | - | - | - | - | - |
| | Roychowdhury et al. (2015b) | 2015 | 0.7395 | 0.9782 | 0.9494 | 0.9672 | 0.7317 | 0.9842* | 0.9560 | 0.9673 | 0.7615 | 0.9575 | 0.9467* | 0.9623* |
| | Azzopardi et al. (2015) | 2015 | 0.7655 | 0.9704 | 0.9442 | 0.9614 | 0.7716 | 0.9701 | 0.9497 | 0.9563 | 0.7585 | 0.9587 | 0.9387 | 0.9487 |
| | Yin et al. (2015) | 2015 | 0.7246 | 0.9790 | 0.9403 | - | 0.8541* | 0.9419 | 0.9325 | - | - | - | - | - |
| | Zhao et al. (2015b) | 2015 | 0.7420 | 0.9820 | 0.9540 | 0.8620 | 0.7800 | 0.9780 | 0.9560 | 0.8740 | | | | |
| | Zhang et al. (2016) | 2016 | 0.7743 | 0.9725 | 0.9476 | 0.9636 | 0.7791 | 0.9758 | 0.9554 | 0.9748 | 0.7626* | 0.9661* | 0.9452 | 0.9606 |
| | Kovács and Hajdu (2016) | 2016 | 0.7450 | 0.9793* | 0.9494 | 0.9722* | 0.8034 | 0.9786 | 0.9610* | 0.9836* | - | - | - | - |
| | Zhao et al. (2017) | 2017 | 0.7820* | 0.9790 | 0.9570* | 0.8860 | 0.7890 | 0.9780 | 0.9560 | 0.8850 | - | - | - | - |
| Supervised methods | Niemeijer et al. (2004) | 2004 | - | - | 0.9416 | 0.9294 | - | - | - | - | - | - | - | - |
| | Staal et al. (2004) | 2004 | - | - | 0.9441 | 0.9520 | - | - | 0.9516 | 0.9614 | - | - | - | - |
| | Soares et al. (2006) | 2006 | 0.7332 | 0.9782 | 0.9466 | 0.9614 | 0.7207 | 0.9747 | 0.9480 | 0.9671 | - | - | - | - |
| | Ricci and Perfetti (2007) | 2007 | - | - | 0.9595 | 0.9558 | - | - | 0.9584 | 0.9602 | - | - | - | - |
| | Lupascu et al. (2010) | 2010 | 0.7200 | - | 0.9597† | 0.9561 | - | - | - | - | - | - | - | - |
| | Marin et al. (2011) | 2011 | 0.7067 | 0.9801 | 0.9452 | 0.9588 | 0.6944 | 0.9819 | 0.9526 | 0.9769 | - | - | - | - |
| | Fraz et al. (2012c) | 2012 | 0.7406 | 0.9807 | 0.9480 | 0.9747† | 0.7548 | 0.9763 | 0.9534 | 0.9768 | 0.7224 | 0.9711 | 0.9469 | 0.9712 |
| | Orlando et al. (2017) | 2016 | 0.7897 | 0.9684 | - | - | 0.7680 | 0.9738 | - | - | 0.7277 | 0.9712 | - | - |
| | Li et al. (2016) | 2016 | 0.7569 | 0.9816† | 0.9527 | 0.9738 | 0.7726 | 0.9844 | 0.9628 | 0.9879† | 0.7507 | 0.9793 | 0.9581 | 0.9716 |
| | Liskowski and Krawiec (2016) | 2016 | 0.7750 | 0.9795 | 0.9518 | 0.9747† | 0.7766 | 0.9854† | 0.9638† | 0.9868 | 0.7544 | 0.9846† | 0.9610† | 0.9801† |
| | Zhang et al. (2017) | 2016 | 0.7861 | 0.9712 | 0.9466 | 0.9703 | 0.7882 | 0.9729 | 0.9547 | 0.9740 | 0.7644 | 0.9716 | 0.9502 | 0.9706 |
| | **Proposed method** | 2017 | **0.8644**† | **0.9667** | **0.9589** | **0.9701** | **0.8325**† | **0.9746** | **0.9502** | **0.9670** | **0.8297**† | **0.9663** | **0.9474** | **0.9591** |

* Best values among unsupervised methods.
† Best values among supervised methods.

Table 3.3: Performance comparison on the IOSTAR and RC-SLO datasets.

| Methods | Datasets | Year | *Se* | *Sp* | *Acc* | *AUC* | *MCC* |
|---|---|---|---|---|---|---|---|
| Zhang et al. (2016) | IOSTAR | 2016 | 0.7545 | **0.9740** | 0.9514 | 0.9615 | **0.7318** |
| **Proposed method** | IOSTAR | 2017 | **0.8269** | 0.9669 | **0.9564** | **0.9663** | 0.7057 |
| Zhang et al. (2016) | RC-SLO | 2016 | 0.7787 | **0.9710** | 0.9512 | 0.9626 | **0.7327** |
| **Proposed method** | RC-SLO | 2017 | **0.8488** | 0.9666 | **0.9581** | **0.9678** | 0.7029 |

Table 3.4: Comparison of the proposed approach in terms of average $MCC$ and $F1$-score.

| Methods | Year | DRIVE | | STARE | | CHASE_DB1 | |
|---|---|---|---|---|---|---|---|
| | | $MCC$ | $F1$ | $MCC$ | $F1$ | $MCC$ | $F1$ |
| 2nd human observer | – | 0.7601 | 0.7881 | 0.7225 | 0.7401 | 0.7475 | 0.7686 |
| Azzopardi et al. (2015) | 2015 | 0.7475 | – | 0.7335 | – | 0.6802 | – |
| Orlando et al. (2017) | 2016 | 0.7556 | 0.7857 | 0.7417 | 0.7644 | 0.7046 | 0.7332 |
| Zhang et al. (2017) | 2016 | **0.7673** | **0.7953** | **0.7608** | **0.7815** | **0.7324** | **0.7581** |
| **Proposed method** | 2017 | 0.7421 | 0.7607 | 0.7398 | 0.7698 | 0.6927 | 0.7189 |

than all the unsupervised approaches. The performance of the proposed method is also validated on SLO images (see Table 3.3), which includes challenging cases like thin vessels, complex crossovers and closely spaced parallel vessels. The proposed method obtains slightly noticeable improvements in $Acc/AUC$ of 0.9564/0.9663 and 0.9581/0.9678 on IOSTAR and RC-SLO dataset respectively, compared with Zhang et al. (2016). In Table 3.4, the proposed method is also compared with the existing approaches in terms of average $MCC$ and $F1$-score across DRIVE, STARE and CHASE_DB1 datasets. The proposed technique has shown to perform better in terms of $MCC$ and the $F1$-score, indicating that the proposed approach is also robust to class imbalance (fewer vessel pixels when compared to pixels belonging to background) in the dataset.

For the sake of fair comparison with the existing state-of-the-art methods, the sensitivity values are extracted from the ROC curve for a fixed specificity, reported in two best methods among supervised and unsupervised approaches, as shown in Table 3.5. Compared with the unsupervised approaches, the proposed method achieved best $Se$ value for a fixed $Sp$, reported in Kovács and Hajdu (2016); Zhao et al. (2017) on DRIVE; Kovács and Hajdu (2016); Yin et al. (2015) on STARE; and Roychowdhury et al. (2015b); Zhang et al. (2016) on CHASE_DB1, respectively. Whereas, among the supervised ones, a higher $Se$ is achieved in comparison with Li et al. (2016); Liskowski and Krawiec (2016) on DRIVE and STARE, inferring that the proposed approach is also competent with other deep learning based approaches. The methods in Kovács and Hajdu (2016); Zhang et al. (2017, 2016) strongly rely on carefully designed hand-crafted features (such as Gabor and Wavelet-based filtering techniques) compared with the proposed approach that learns the feature representation directly from the raw data, without the explicit need for complex domain knowledge.

Table 3.5: Comparison of sensitivity and specificity values.

| | DRIVE | | | STARE | | | CHASE_DB1 | | |
|---|---|---|---|---|---|---|---|---|---|
| | Methods | $Se$ | $Sp$* | Methods | $Se$ | $Sp$* | Methods | $Se$ | $Sp$* |
| **Unsupervised** | Zhao et al. (2017) | 0.7820 | 0.9790 | Yin et al. (2015) | 0.8541 | 0.9419 | Roychowdhury et al. (2015b) | 0.7615 | 0.9575 |
| | Kovács and Hajdu (2016) | 0.7450 | 0.9793 | Kovács and Hajdu (2016) | 0.8034 | 0.9786 | Zhang et al. (2016) | 0.7626 | 0.9661 |
| | **Proposed method** | **0.7939** | **0.9790** | **Proposed method** | **0.8660** | **0.9419** | **Proposed method** | **0.7862** | **0.9575** |
| | | **0.7935** | **0.9793** | | **0.8060** | **0.9786** | | **0.7670** | **0.9661** |
| **Supervised** | Liskowski and Krawiec (2016) | 0.7750 | 0.9795 | Liskowski and Krawiec (2016) | **0.7766** | **0.9854** | Liskowski and Krawiec (2016) | **0.7544** | **0.9846** |
| | Li et al. (2016) | 0.7569 | 0.9816 | Li et al. (2016) | 0.7726 | 0.9844 | Zhang et al. (2017) | **0.7644** | **0.9716** |
| | **Proposed method** | **0.7930** | **0.9795** | **Proposed method** | 0.7736 | 0.9854 | **Proposed method** | 0.7459 | 0.9846 |
| | | **0.7647** | **0.9816** | | **0.7773** | **0.9844** | | 0.7512 | 0.9716 |

* specificity cut-off: specific point on the ROC curve for extracting sensitivity values.

### 3.3.6 Cross-Validation

The performance of the proposed method is evaluated by training and testing on a set of images from different databases, to check the robustness and practicability of the approach. Table 3.6 shows the cross-validation performance between datasets with the existing state-of-the-art methods. A slightly decrease in the *Acc* value of 0.9489 is obtained by training on STARE and testing on DRIVE dataset, when compared to training and testing on STARE alone. This is because the STARE contains pathological signs in almost 10 of its images, while DRIVE contains mostly healthier ones. Further, the $Acc/AUC$ values of 0.9640/0.9702 is obtained by training on DRIVE and tested on STARE, which is marginally greater than $Acc/AUC$ of 0.9589/0.9701 obtained by training and testing on DRIVE alone. A decreasing trend is also observed in $Acc/AUC$ values when trained on STARE and tested on CHASE_DB1, compared to training and testing on STARE alone. Overall, there is a clear indication that the learned feature set doesn't strongly rely on the specific training set, and are highly adaptable to varying image resolutions, robust to pathological signs and various other imaging artefacts present in the data.

### 3.3.7 Performance Analysis on Challenging Cases

Segmenting the retinal vessels from color fundus image is challenging due to the presence of strong CVR, low contrast thin vessel structures, close parallel and highly curved vessels, close bifurcation and crossover regions and pathological lesions such as exudates, haemorrhages and microaneurysms.

Figure 3.11, shows the qualitative performance of the proposed approach on sample ROI's containing various challenging cases. The proposed method produces far less false positives, more clean segmentation compared to the most recent state-of-the-art methods (Liskowski and Krawiec (2016); Orlando et al. (2017)). The approach is also shown to preserve various complex vascular structures (first and second row) and segments thin vessels accurately even at very low contrast regions (third row). This robustness is mainly due to the learnt filters that leverage the visual attention mechanism by taking into account of multi-scale contextual information, to produce a set of more discriminative features. Compared to the conventional methods such as Christodoulidis et al. (2016); Roychowdhury et al. (2015a), the proposed approach is capable of segmenting both large and small vessel structures within a sin-

Table 3.6: Performance comparison with cross-validation between databases.

| | Methods | *Acc* | *AUC* |
|---|---|---|---|
| Test images from: **DRIVE** | | | |
| Model trained on **STARE** | Soares et al. (2006) | 0.9397 | - |
| | Ricci and Perfetti (2007) | 0.9266 | - |
| | Marin et al. (2011) | 0.9448 | - |
| | Fraz et al. (2012c) | 0.9456 | **0.9697** |
| | Li et al. (2016) | 0.9486 | 0.9677 |
| | Liskowski and Krawiec (2016) | 0.9416 | 0.9605 |
| | Zhang et al. (2017) | 0.9447 | 0.9593 |
| | **Proposed method** | **0.9489** | 0.9676 |
| Model trained on **CHASE_DB1** | Li et al. (2016) | **0.9484** | **0.9605** |
| | **Proposed method** | 0.9301 | 0.9476 |
| Test images from: **STARE** | | | |
| Model trained on **DRIVE** | Soares et al. (2006) | 0.9327 | - |
| | Ricci and Perfetti (2007) | 0.9464 | - |
| | Marin et al. (2011) | 0.9528 | - |
| | Fraz et al. (2012c) | 0.9493 | 0.9660 |
| | Li et al. (2016) | 0.9545 | 0.9671 |
| | Liskowski and Krawiec (2016) | 0.9505 | 0.9595 |
| | Zhang et al. (2017) | 0.9488 | 0.9676 |
| | **Proposed method** | **0.9640** | **0.9702** |
| Model trained on **CHASE_DB1** | Li et al. (2016) | **0.9536** | **0.9620** |
| | **Proposed method** | 0.9334 | 0.9464 |
| Test images from **CHASE_DB1** | | | |
| Model trained on **DRIVE** | Li et al. (2016) | 0.9429 | **0.9628** |
| | **Proposed method** | **0.9476** | 0.9621 |
| Model trained on **STARE** | Fraz et al. (2012c) | 0.9415 | **0.9565** |
| | Li et al. (2016) | 0.9417 | 0.9553 |
| | Zhang et al. (2017) | **0.9458** | 0.9538 |
| | **Proposed method** | 0.9403 | 0.9521 |

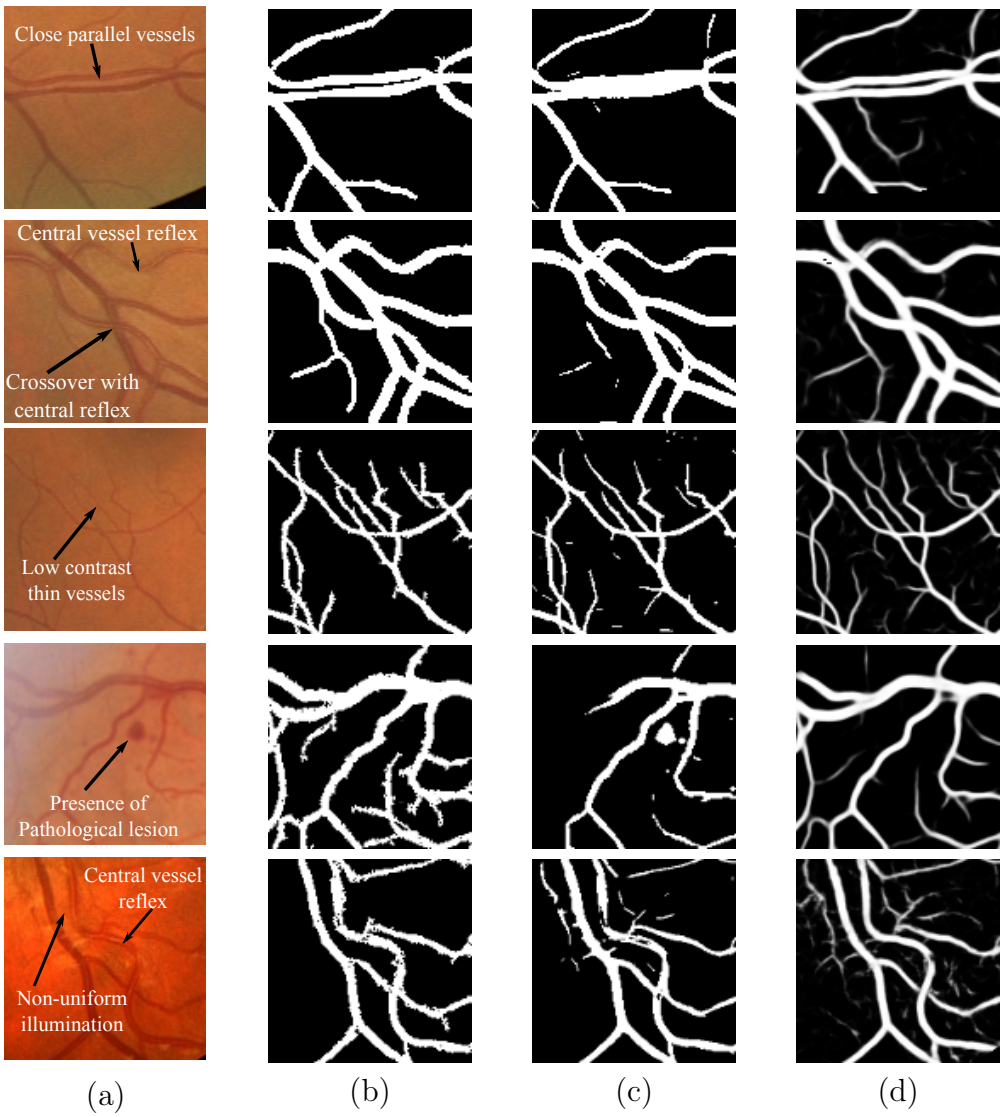



Figure 3.11: Qualitative assessment of vessel segmentation results on challenging cases. (a) Original image; (b) Proposed approach; (c) Orlando et al. (2017); (d) Liskowski and Krawiec (2016).

gle framework, without the explicit need for sophisticated techniques to tackle them separately. Further, the method is also shown to be robust to pathological lesions such as microaneurysms (see the fourth row) and even the CVR phenomenon (see the fifth row) compared to Orlando et al. (2017). The significant advantage of the proposed method is its ability to handle most of the challenging cases, without the explicit need for any complex pre/post-processing operations. Hence, the proposed approach is more reliable and suitable for developing a CAD tool for the automated analysis of retinal vasculature.

Table 3.7: Computational complexity vs. accuracy (*Acc*) of the state-of-the-art methods on DRIVE/STARE datasets, respectively. Note: the segmentation time is recorded for processing one DRIVE/STARE image.

| **Methods** | Year | Running time | *Acc* |
|---|---|---|---|
| Soares et al. (2006) | 2006 | 3 mins | 0.9466/0/0.9480 |
| Marin et al. (2011) | 2011 | 1.5 mins | 0.9452/0.9526 |
| Fraz et al. (2012c) | 2012 | 2 mins | 0.9480/0.9534 |
| Roychowdhury et al. (2015a) | 2015 | 3.11/6.7 s | 0.9490/0.9560 |
| Roychowdhury et al. (2015b) | 2015 | 2.45/3.95 s | 0.9520/0.9510 |
| Li et al. (2016) | 2016 | 1.2 mins | 0.9527/**0.9628** |
| Zhang et al. (2017) | 2016 | 23.4 s | 0.9466/0.9547 |
| **Proposed approach** | 2018 | **45 s** | **0.9589**/0.9502 |

### 3.3.8 Computational Cost

The running time for the vessel segmentation approach is reported in Table 3.7, starting from visual attention modelling to the prediction of class labels using RF classifier, which takes approximately 45 seconds per image using an unoptimized MATLAB code, on a machine equipped with 2.2GHz, Intel i3-2330 processor and 8GB of RAM. The main computational bottleneck is the encoder part of the unsupervised filter learning approach, which takes roughly 20 seconds per image for solving $L_1$ optimization in the sparse coding step. This can be solved more efficiently using a very recently proposed soft thresholding scheme (Fawzi et al. (2015)), which jointly learns both the dictionary and a linear classifier in a single optimization step. This could be interesting future work in this direction. It is observed that based on the current parameter setting (see Section 3.3.3), the proposed approach offers a considerable prediction time suitable for the standard clinical setting, without compromising on segmentation accuracy across all five datasets.

## 3.4 Discussion

This chapter presented visual attention guided unsupervised feature learning framework for the task of segmenting retinal vessels from color fundus images. Segmenting retinal vessels is challenging due to the multi-scale nature of varying vessel calibre, complex crossover patterns, the proximal presence of pathological lesions and non-uniform illumination inherited during the image acquisition process. These challenges often pose great difficulty in obtaining a discriminative feature set for the

accurate classification of vessel pixels. Most previous methods (Azzopardi et al. (2015); Kovács and Hajdu (2016); Roychowdhury et al. (2015a); Zhang et al. (2017, 2016)) have focused extensively on carefully designing hand-crafted features, which are generally based on complex domain knowledge, and requires an enormous amount of manual tuning of parameters to achieve optimal segmentation performance. These limitations are addressed by leveraging the idea of visual attention mechanism into conventional unsupervised feature learning framework. The proposed scheme aims to produce a more versatile segmentation approach through the design of an automatic learned discriminative feature set, using a set of only unlabelled image samples. Therefore, the proposed method is capable of capturing the geometrical richness of retinal vasculature - occurring at multiple-scales and multiple-orientations, without the need for any expensive manual annotations and complex domain expertise.

The proposed approach has several appealing properties. First, the visual attention mechanism is capable of capturing the rich contextual information, which can ignore the clutter present in a local neighborhood, by focusing only on the pixel of interest. Second, the integration of visual attention mechanism into unsupervised filter learning encourages the intra-class similarities to be small (between the vessel pixels), and emphasizes the inter-class differences to be large (between the vessel and background pixels). Third, the VA-UFL approach requires only three main parameters: $p_0$, $l$ and $K$ during its training, which is generally chosen by cross-validation. Further, the chosen parameters should be fixed only once, which is independent of varying image resolutions and requires practically no manual tuning for various kinds of datasets.

Extensive experiments are carried out on five publicly available retinal datasets containing various challenging cases to illustrate the broad applicability of the VA-UFL method. It is shown that the proposed approach is highly competent and outperforms state-of-the-art methods by a value of $Se > 0.82$ across all five datasets. In addition, the proposed approach consistently performed well in terms of global performance metrics such as $MCC$ and $F1$-score, demonstrating that the method is also robust to class imbalance present in the data (fewer vessel pixels compared to background). The significant boost in $Se$ value is mainly due to the discriminative power of learned filter bank, which leverage the strength of visual attention mechanism, to accurately identify true positives from a similar-looking noisy background.

The proposed approach is also shown to perform well on various challenging

cases such as: segmenting thin vessels from low contrast regions, segmentation in the presence of pathological lesions and strong CVR. The improved segmentation performance is mainly because of the visual attention mechanism, that takes into account complex neighborhood information, to accurately segment vessels from its complex noisy background. This is also true in the case of projection of learned features, which is visualized via t-SNE in low-dimensional subspace. It is observed that there is a significant class separation between the vessel and background pixels with higher levels of the visual glimpse, indicating that visual glimpse is inherently capturing the most relevant structures in addition to the contextual information, to learn a set of most discriminative basis.

Although the proposed approach successfully segments complex retinal vascular structures, some misclassification has been observed in the region of bright and large exudates. These false positives can further be removed by employing a simple pre/post-processing steps (such as image inpainting) or by integrating vessel shape prior information into the filter learning approach. Besides that, the segmentation of pathological images can be even further improved by incorporating more hybrid contextual models such as auto-context (Tu and Bai (2010)). The auto-context iteratively utilizes the posterior distribution of labels along with the image features, to obtain a more compact and discriminative features suitable for highly overlapping classes. Furthermore, the discriminative capability of the UFL approach can be further improved by associating each class label information with the dictionary atom during the filtering process. This hybrid information effectively reduces both the reconstruction and classification error using a unified objective function.

## 3.5 Summary

In summary, the proposed approach inherits the advantages of visual attention mechanism and fully utilizes the potential of unsupervised feature learning for representing most discriminative features for classification. This allows to explore the space of both selection mechanism and multi-scale contextual information under a single framework, without the need for any complex feature learning modules. Extensive experimental analysis on five publicly available retinal datasets demonstrates the superior performance of the proposed approach with respect to other state-of-the-art vessel segmentation methods. The effectiveness of the proposed approach is evident by the significant improvement in the value of sensitivity compared to all

previously published methods and even outperforms second human observer. Further, the method is also shown to be robust to several challenging image structures such as central vessel reflex, complex crossover patterns, closely parallel and highly curved vessels, thin vessels and performs reasonably well on pathological images. In conclusion, the excellent performance of the approach demonstrates the applicability for real-time computer-aided diagnosis and large-scale retinal disease screening programs.

# Chapter 4

# Automatic Junction Classification and Analysis of Crossover Biomarkers in Retinal Images

Retinal vessel keypoint detection and classification is a fundamental step in tracking the physiological changes that occur in the retina, which is linked to various ocular and systemic diseases. In this chapter, a novel vessel keypoint descriptor (VKD) is introduced, which is derived from the projection of log-polar transformed binary patches around vessel points. VKD is used to design a two-stage solution for junction detection and classification. First, the keypoints are detected using VKD and are refined using curvature orientation information to extract candidate junctions. True junctions from these candidates are identified via supervised manner. Second, a novel combination of local orientation and shape based features are extracted from the junction points and classified into bifurcations and crossovers, outperforming the state-of-the-art approaches. Subsequently, several geometrical properties at crossover points are analysed to detect and quantify the morphological changes linked to hypertension, stroke and other systemic diseases. The proposed solution indicate that the crossover properties can be considered as crucial biomarkers in assessing the progression of various microvascular diseases.

## 4.1 Background

One of the basic tasks in vessel tree analysis is identifying keypoints such as junctions, as they serve as landmarks useful in the registration of retinal images and in biometrics. Clinically, vessel keypoints such as bifurcation points are of interest in assessing diseases such as retinopathy of prematurity - via the branch angle at these points. Retinal vessels vary significantly in their size, shape, and orientation. Further, in detecting keypoints, local as well contextual information is useful in resolving ambiguities and overcoming problems due to the variability in morphology and orientation.

Automatic detection of junctions which are a specific type of keypoint has attracted much interest. Junction detection and classification have been attempted in the past using vessel maps or models. Skeletonization of vessel maps in Abbasi-Sureshjani et al. (2016); Aibinu et al. (2010); Fathi et al. (2013); Favali et al. (2016) is followed by analysis in a window around a vessel pixel, to assess if the pixel is a keypoint, such as bifurcation or crossover. Both, circular (Fathi et al. (2013)) and square-shaped windows (Aibinu et al. (2010)) have been considered. Evidence gathering by pixel voting followed by a refinement based on the distance between pixels is typically used to locate bifurcations and crossovers. In contrast, the model-based techniques employ shifted Gabor filters (Azzopardi and Petkov (2013)) or eigen analysis of Hessian at the vessel point (Su et al. (2012)). More recently, orientation scores derived by correlating image patches with anisotropic wavelets (Abbasi-Sureshjani et al. (2016); Favali et al. (2016)) have been proposed as a basis to detect junctions and subsequently classify them as bifurcations or crossovers. This approach was quite successful in modelling different kinds of complex junction locations. Existing approaches for detecting keypoints, model the orientation of the vessel in a local neighbourhood without explicitly considering the larger context or shape information which might aid in handling ambiguities. In this chapter, a learning-based approach is proposed that combines orientation and shape information for detecting and classifying keypoints. The proposed strategy is shown to be robust, accurate and reliable.

The most prominent clinical biomarker associated with hypertension and incident stroke is the arteriovenous (AV) nicking - characterized by the steep decrease in vein width at both sides of the vessel crossovers (Wong and Mitchell (2007, 2004)). This

phenomenon occurs due to the increased blood pressure in the arteries, which in turn exerts a pressure on the veins, leading to compressed veins at artery-vein crossing locations. In clinical practice, assessment of AV nicking is done manually - which is extremely time-consuming and subject to large intra and inter observer variability (Hubbard et al. (1999)).

Automated detection of AV nicking from color fundus images has gained a very recent interest with the seminal work of Nguyen et al. (2013a). Thereafter, a few methods have been proposed in the literature based on unsupervised and supervised techniques. Majority of the existing methods starts by first extracting a vascular tree, followed by vessel crossover point identification, A/V segment classification, vein width estimation and finally AV nicking quantification in a sequential order.

Among the unsupervised ones, Nguyen et al. (2013a) proposed a fully automated method for detection of AV nicking based on the aforementioned sequence of techniques. In their approach, AV nicking quantification is represented as a continuous valued output, which is often difficult to clinically interpret the actual severity level of an AV nicking crossover point. In contrast among supervised ones, Roy et al. (2014) proposed a combination of vessel cross-sectional intensity profile and second order derivative features to estimate the vein widths at a crossover point. These vein widths are further trained using a random forest classifier to compute the severity level of an AV nicking. Pereira et al. (2014) adopted a similar approach to Nguyen et al. (2013a) for estimation of abnormal vein widths, which are used as a set of features to train a support vector machine classifier. The major drawback of these existing methods is that they are strongly dependent on A/V classification of crossover segments, which is extremely challenging for low contrast and noisy images. Further, the features extracted from low-contrast crossover regions will be generally noisier, leading to erroneous computation of AV nicking severity levels.

To address the above mentioned challenges, a machine learning-based approach is proposed to precisely quantify the AV nicking severity level of a crossover point. The proposed approach is based on the estimated vessel widths of all crossing segments, without the explicit need for artery/vein classification. This is based on a vessel keypoint descriptor (VKD), which was proposed and illustrated for the detection and classification of vessel junctions (Srinidhi et al. (2017b)). Here, the space of both local orientation and width of vessels is explored with the aid of VKD, to detect and quantify the vascular changes at interest points for identifying AV nicking. The

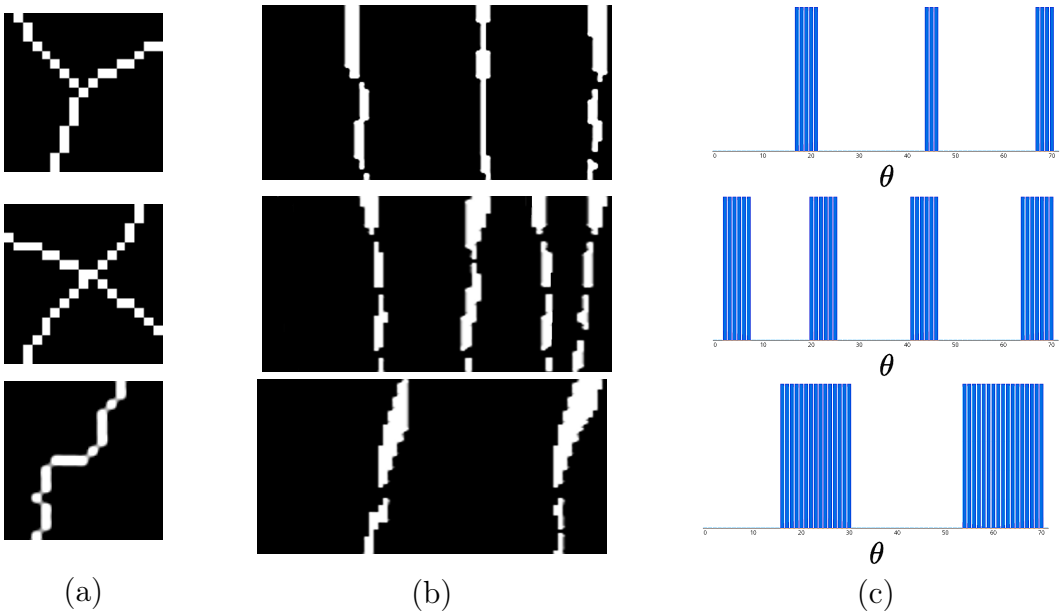



Figure 4.1: Illustration of VKD. (a) (from top to bottom): Sample bifurcation, crossover and a non-junction patch; (b) The corresponding Log-Polar maps (with x-axis being $\theta$); (c) The final descriptor $V_p(n)$.

main contributions of this chapter are summarised as follows:

i. An end-to-end system for retinal vasculature analysis: junction detection, junction classification and vessel width measurement.

ii. A formulation of AV nicking severity level quantification in terms of estimated vessel widths of all crossing vessel segments. This removes the need for knowledge of the identity of the vessel segments.

iii. A thorough validation of AV nicking quantification approach using an unsupervised and a supervised technique on a publicly available retinal dataset.

The remainder of this chapter is organized as follows. Section 4.2, explains in detail the proposed junction classification approach, followed by analysis of crossover biomarkers in Section 4.3. Section 4.4, defines the experimental setup, dataset used and results of the Junction classification and AV nicking quantitative approach. Finally, the key findings as well as possible future directions are discussed in Section 4.5.

## 4.2 Automatic Junction Classification Approach

In this section, the proposed "Vessel Keypoint Descriptor" (VKD) is explained in detail, which is designed to operate on a binary vessel map and extract points of interest in the vessel tree. All vessel maps are skeletonized first to handle varying vessel calibre within and across images. Next, around every vessel point $p$, a region of interest (ROI) $R_p(x, y)$ is considered followed by, applying a log-polar transform to obtain $R_p(r, \theta)$. The log mapping preserves information close to a vessel point, while increasingly compressing the information as one moves away from the vessel point in a non-linear fashion. Figure 4.1 shows sample ROI's with different vessel patterns of interest and the corresponding log-polar mapped results. The number of vertical lines depends on the pattern (straight vessel segment, crossing vessels, branching vessels etc.), and their position depends on the orientation of the vessels in $R_p(x, y)$.

Next, a vertical projection of $R_p(r, \theta)$ is computed to obtain a vector $R_p(\theta)$. In order to build robustness to spurious vessels and varying vessel calibre, the range of $r$ is limited to an interval $[R_1, R_2]$, where as $\theta \in [0°, 360°]$. Since $r$ and $\theta$ have to be discrete variables (say $m, n$), it is hence, sampled with sampling rates of $\Delta_r$ and $\Delta_\theta$. The obtained projection $R_p(\theta)$ provides a count of the number of pixels in a vessel fragment at a specific $\theta$. This is binarized with a threshold set at 1 to obtain a vessel keypoint descriptor $V_p(n)$, which is shown as a 1-D function (for a typical junction and non-junction) in Figure 4.1. Since a cluster of responses in $V_p(n)$ correspond to only one vessel fragment, the vessel edges are located by computing a first order difference of $V_p(n)$ as follows:

$$V'_p(n) = abs(V_p(n+1) - V_p(n)). \tag{4.1}$$

The number of vessel branches $N$ at a point $p$ is found as:

$$N(p) = \frac{1}{2} \sum_n V'_p(n). \tag{4.2}$$

Finally, the obtained set of key points: $C = \{p_i | N(p_i) > 2\}$.

Figure 4.2 (a) and Figure 4.2 (c) show two challenging segmented vessel patches, where complex vessel patterns can be seen. The detected keypoints are marked in

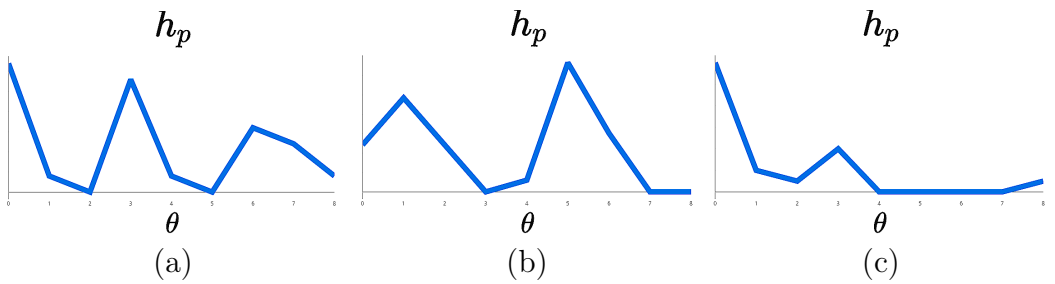
(a) (b) (c) (d)

Figure 4.2: Challenges in vessel keypoint detection. Red ellipses point to artefacts due to segmentation and skeletonization. (a), (c) Skeletonized vessel maps; (b), (d) Keypoints detected with VKD.

blue in Figure 4.2 (b) and Figure 4.2 (d). These are seen mostly in the vicinity of junctions. A common type of artefact produced by segmentation are spurs - appearing as a part of a vessel segment which leads to noisy keypoints. These are handled by setting $R_1 > k$ pixels; with $k = 2$ or higher value. The effect of skeletonization also produces an artefact at vessel crossovers, namely a structure that resembles two closely placed bifurcations as depicted in Figure 4.2 (d). These are also tackled with a judicious choice of $R_1$ and $R_2$. It should be noted that since the vessel map is about one pixel thick, the parameters need to be fixed only once. Empirically, it is found that $R_1 = 3px$; $R_2 = 8px$; $\Delta_r = 0.1$; and $\Delta_\theta = 5°$ produce the best results. A sample image, the corresponding thinned vessel map and detected keypoints are shown in Figure 4.5 (a) through Figure 4.5 (c).

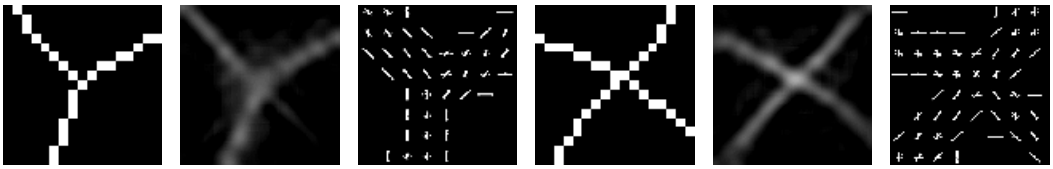
The proposed keypoint descriptor is used to design a system for vessel analysis. The input to this system is a skeletonized vessel map and output is a set of bifurcation and crossover points. The system design consists of the following steps: candidate junction selection, junction detection followed by classification. Details are presented next.

### 4.2.1 Junction Candidates Selection




(a) (b) (c)

Figure 4.3: Curvature orientation histograms for sample: (a) Bifurcation; (b) Crossover; and (c) Non-junction points.



(a) (b) (c) (d) (e) (f)

Figure 4.4: Visualization of LD and HOG feature descriptors for sample bifurcation (a) and crossover (d) patches. (b), (e) and (c), (f) are the corresponding line response and the HOG map, respectively.

Given a skeletonized vessel map, a set of keypoints are extracted using VKD as described in the Section 4.2. The obtained keypoints $C$ appear as clusters of points close to junctions (as shown in Figure 4.5 (c)). These are refined to obtain the desired candidate junction points by noting that: a junction point is where vessels of two or three different orientations meet. This can be characterised by the entropy of the histogram of vessel orientations; points close to a junction will exhibit a higher entropy than those away (Ram et al. (2009)). For each point in $C$, a $3 \times 3$ neighborhood $\omega$ is considered and a Hessian matrix is computed for every pixel in $\omega$, and its second eigenvector is used to construct a curvature orientation histogram (COH) $h_p$. Finally, the entropy of COH $E(h_p)$ is obtained and a non-maximum suppression with a radius of 12 pixels is done to obtain the set of final candidate junctions $C_j$, which are shown in Figure 4.5 (d). Figure 4.3, illustrates the distinct signatures in COH for bifurcation, crossover as well as a non-junction point, respectively.

### 4.2.2 Junction Detection

A supervised learning technique is proposed for detecting true junctions. The VKD ($V$) and COH ($h$) features are extracted at every candidate points. These two features capture the vessel orientation information in a local neighborhood. Hence, they exhibit an excellent discriminative ability in distinguishing between a true junction $J_n$ and a non-junction $NJ_n$. For each $C_j$, features are extracted from a $17 \times 17$ neighborhood and concatenated into a 1-D long vector $X_J = (V, h)$. A Random Forest (RF) (Breiman (2001)) classifier of 200 trees was trained. A sample result of the detected junctions are shown in Figure 4.5 (e).

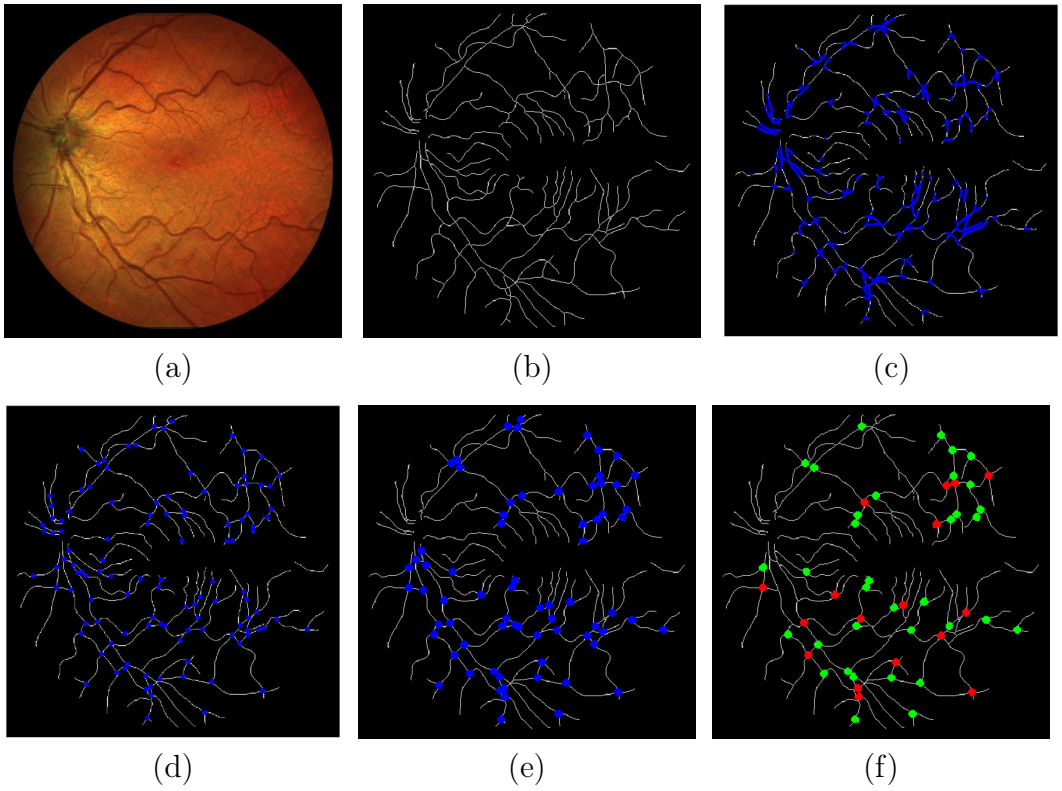



Figure 4.5: Junction detection and classification results. (a) A sample image; (b) Skeletonized vessel map; (c) Result of VKD; (d) Candidate junctions; (e) Detected true junctions; (f) Identified bifurcations (green) and crossovers (red).

### 4.2.3 Junction Classification

Bifurcation versus crossover classification is a non-trivial task, since it poses several challenges which include: ambiguity due to proximal presence of bifurcation and crossover points, multiple bifurcations/crossovers in a small region and parallel vessels. These are addressed by using a combination of features: orientation based $(V, h)$ and shape based (response of basic Line Detector (LD) (Ricci and Perfetti (2007)) and Histogram of Oriented Gradients (HOG) (Dalal and Triggs (2005)) features. The LD $(L)$ and HOG $(O)$ primarily capture the local vessel shape information and can be used to accurately identify bifurcations and crossovers, as shown in Figure 4.4. For each detected true junction $J_n$, the features $(V, h, L, O)$ are computed in a $17 \times 17$ neighborhood. The line response at each pixel is obtained as explained in Ricci and Perfetti (2007), which is vectorized to form a 1-D vector $L$. These features are concatenated to derive the final feature vector. A RF classifier with 500 trees is trained with these features to classify junctions $J_n$ into bifurcations $J_b$ and crossovers $J_c$. A sample result of the junction classification is shown in Figure 4.5 (f).

## 4.3 Analysis of Crossover Biomarkers

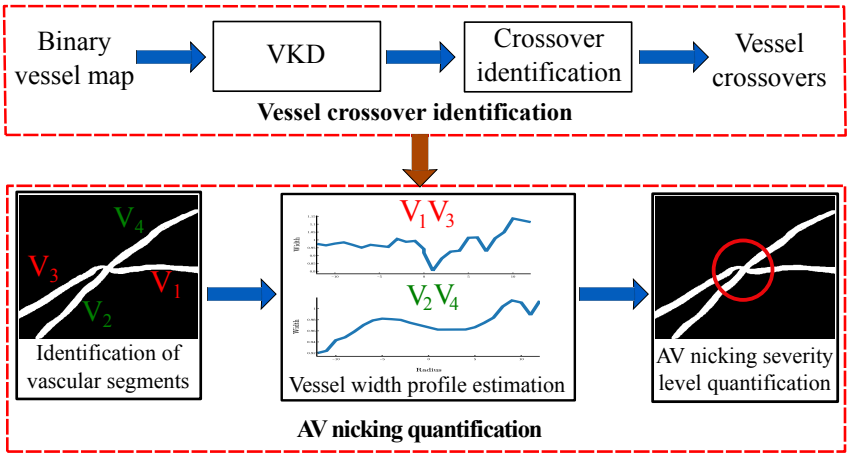



Figure 4.6: An overview of the proposed end-to-end solution for arteriovenous (AV) nicking quantification.

Retinal anatomy is such that arteries and veins crossover each other, at some places in the retinal surface. The pressure exerted by an artery on an underlying vein in some disease conditions causes a compression of the vein, and this, in turn, appears as a thinning of the vein at the crossover point. Such a phenomenon is called "AV nicking". Generally, this phenomenon occurs in the case of small veins and arteries. Thus for detection of AV nicking, morphological changes at crossover points are sufficient to identify AV nicking. Hence, the proposed method aim to address AV nicking detection and quantification *without* the knowledge of a vessel segment being an artery or a vein. An overview of the proposed framework for automatic quantification of AV nicking is illustrated in Figure 4.6. The framework consists of two main stages: (i) vessel crossover identification; (ii) AV nicking detection and severity level quantification.

In the first stage (see Section 4.2), the system takes a binary vessel map as input and detects a set of keypoints using the VKD, which are processed to identify junctions. The detected junctions are then classified into bifurcations and crossovers. In the second stage, the widths of the crossing vessel segments are estimated at multiple points on the segments and are used to predict the severity level of an AV nicking crossover location. Next, the details of the proposed pipeline are presented.

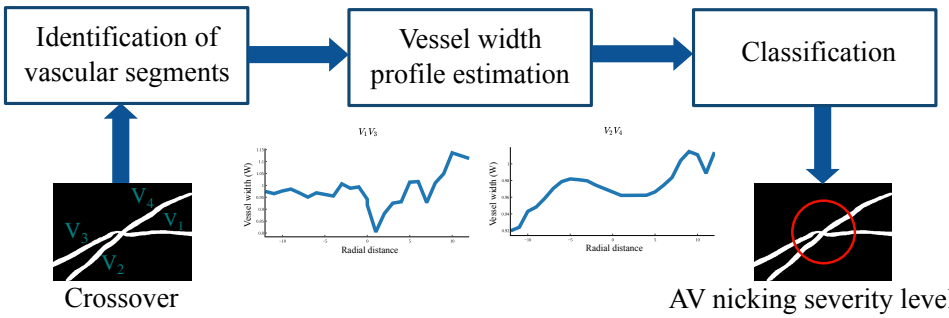



Figure 4.7: An overview of the proposed AV nicking quantification approach.

### 4.3.1 AV Nicking Detection and Quantification

AV nicking detection and quantification is hampered by poor vessel connectivity at crossover locations, ambiguity between artery and vein segments at low contrast regions, as well as subtle changes in vein width. However, the VKD is a rich descriptor, which in addition to identifying keypoints like crossover points in the vessel network, can also aid in identifying the number of vessel segments and their thickness. Such information can be leveraged to identify the vascular abnormalities, if present, including nicking. Hence, a three-step solution is proposed:

- Identify the individual pair of vascular segments at crossover locations;
- Accurately measure the individual vessel widths;
- Quantify AV nicking severity into one of 3 levels.

The overall AV nicking quantification approach is illustrated in Figure 4.7.

#### 4.3.1.1 Identification of Vascular Segments

A vessel crossover arises when an artery crosses a vein, leading to four vascular segments. Of these four, diagonally opposite pair of vessel fragments will belong to the same class - either an artery or a vein. In practice, there can be more than four branches associated with a crossover point. This is due to a branching and crossover point occurring very close to each other.

Figure 4.8, shows two sample AV nicking (top 2 rows) and two normal crossover patches (bottom 2 rows), and their corresponding VKDs. The last row is an example of five vessel fragments appearing to be converging at a crossover point. Since, the VKD represents the presence of at least one vessel point at a specific angle within a window of $\geq M_{min}$, each cluster of responses in VKD indicates a vessel fragment.

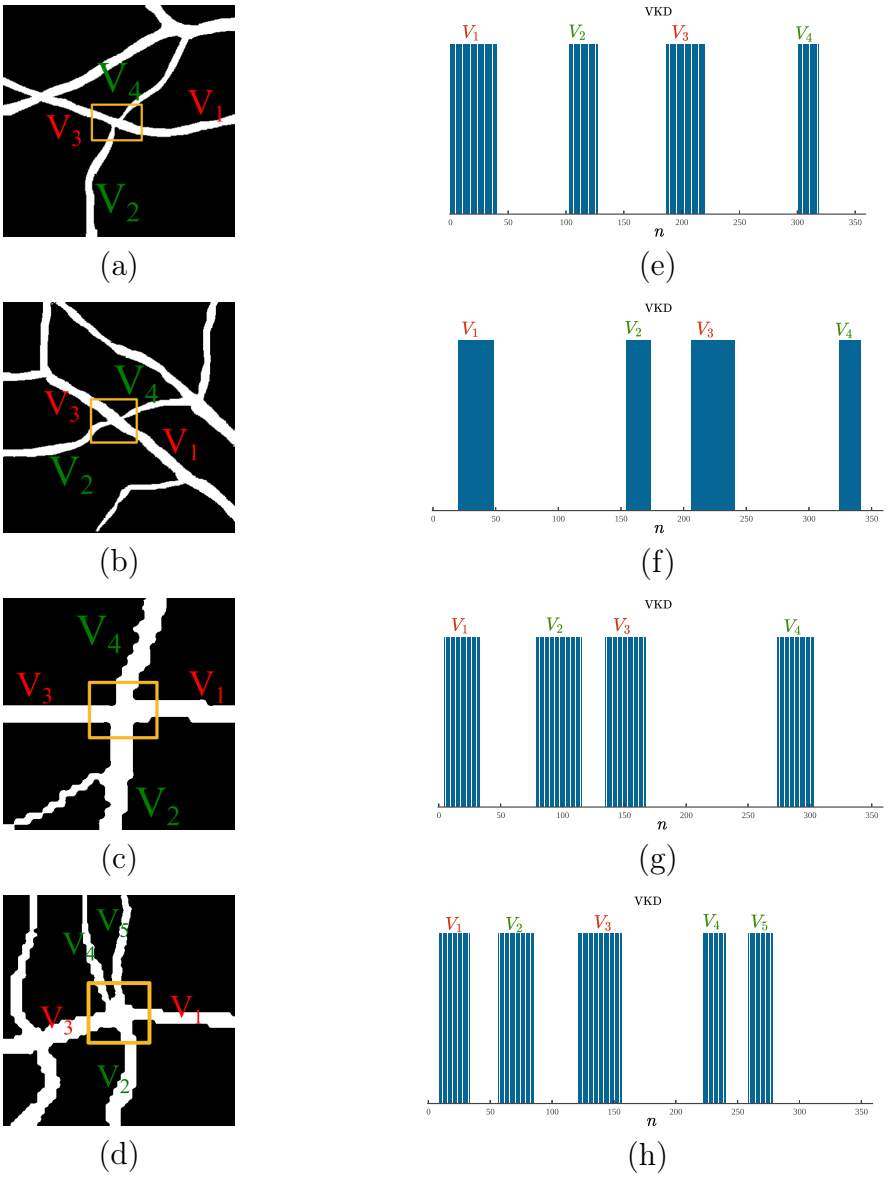



Figure 4.8: Illustration of VKD for AV nicking and normal crossover patches. (a) and (b) AV nicking crossovers with corresponding VKD are in (e) and (f) respectively; (c) and (d) Normal crossovers with corresponding VKD are in (g) and (h), respectively.

From this fact, the two observations at a crossover point $c$ are made.

i. A pair of vessel fragments which belong to a single segment are *collinear*. This implies that a pair of vessel fragments to be an artery/vein must be separated by $n = \pm 180^\circ$ in the VKD. This is seen from the examples in Figure 4.8, where labelled fragments which are collinear are shown in red or green (first column). Their corresponding VKD (second column) exhibits the required $\pm 180^\circ$ angular separation between the pairs of fragments. This also shows how the unpaired fragment can be easily identified in the five fragments case (last row).

ii. The angular span of a vessel (measured with $c$ as origin) is *correlated* to its thickness. This implies a positive correlation between the width of a vessel and the corresponding cluster in VKD. This can be observed to be true from the difference in the vessel width patterns in nicking versus normal cases as shown in Figure 4.8.

The above two observations are used to identify the vessel segments at a crossover point, as explained next.

Given a detected junction point $J$, the VKD($J$) is given as

$$\begin{aligned}\text{VKD}(J) =& \Big[u\Big(n - n_1\Big) - u\Big(n - (n_1 + \epsilon_1)\Big)\Big] \\ &+ \Big[u\Big(n - n_2\Big) - u\Big(n - (n_2 + \epsilon_2)\Big)\Big] \\ &+ \cdots + \Big[u\Big(n - n_B\Big) - u\Big(n - (n_B + \epsilon_B)\Big)\Big] \\ =& \sum_{i=1}^{B} \Big[u\Big(n - n_i\Big) - u\Big(n - (n_i + \epsilon_i)\Big)\Big], \end{aligned} \tag{4.3}$$

where, $\Big[u\Big(n - n_i\Big) - u\Big(n - (n_i + \epsilon_i)\Big)\Big] \triangleq V_i$ – refers to a specific vessel fragment $V_i$ belonging to either an artery/vein; $\epsilon$ is a small constant representing the angular span of the vessel; $n \in [0^\circ, 360^\circ]$ - assuming an angular sampling rate of $1^\circ$; and $B$ is the number of vessel fragments converging at a junction point, which is typically $\geq 3$.

Next, if a junction is a crossover point $c$, then there exists a fragment $V_{ip}$ for

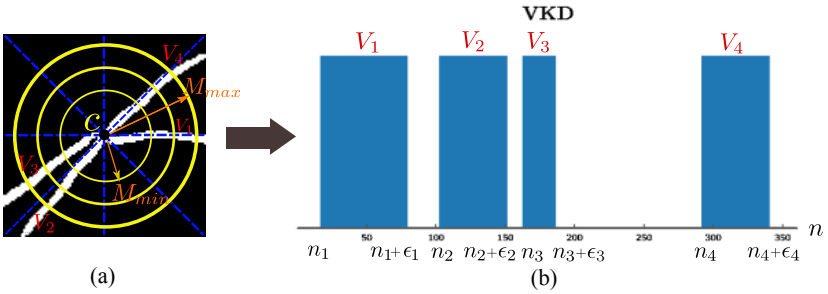



Figure 4.9: Identification of corresponding pair of vascular fragments ($V_i$ and $V_{ip}$). (a) Sample patch with a crossover point; (b) The corresponding VKD.

every vessel fragment $V_i$ belonging to the same class (artery/vein) such that:

$$V_{ip} = u\Big(n - (n_i \pm 180^\circ \pm \gamma)\Big) - u\Big(n - \big((n_i \pm 180^\circ \pm \gamma) + \epsilon_i\big)\Big), \tag{4.4}$$

where, $V_i$ and $V_{ip}$ are vessel fragment pairs belonging to same vessel segment (artery/vein), which are separated by an angle of $180^\circ \pm \gamma$, as illustrated in Figure 4.9 (a). $\gamma$ is a small factor which accounts for highly *curved* crossover vessel segments. In Figure 4.9, ($V_1$ *and* $V_3 = V_{1p}$) or ($V_2$ *and* $V_4 = V_{2p}$) are examples of ($V_i$ *and* $V_{ip}$).

Further, if $c$ is a crossover point, then there exists two scenarios:

i. A *simple* vessel crossover: a pair of vessels crossing each other at a point (see Figure 4.8 (a-c)). In this case, VKD is of the form

$$\text{VKD}(c) = \sum_{i=1}^{2} \Big[V_i + V_{ip}\Big]. \tag{4.5}$$

ii. A *complex* vessel crossover: a pair of vessels crossing at a point and there exists a *proximal* bifurcation (see Figure 4.8 (d)). In this case, VKD is of the form

$$\text{VKD}(c) = \sum_{i=1}^{2} \Big[V_i + V_{ip}\Big] + V_k, \tag{4.6}$$

where, $V_k \triangleq \Big[u\Big(n - n_k\Big) - u\Big(n - (n_k + \epsilon_k)\Big)\Big]$ corresponds to a lone vessel fragment resulting from a nearby bifurcation as shown in Figure 4.8 (d) ($V_k$ - refers to vessel fragment $V_5$). *Note*: a complex crossover point generally consists of $V_i > 4$ vessel fragments. Although this is a very rare condition, it is still take into account in order to build robust approach, which works well irrespective of various complex crossover patterns.

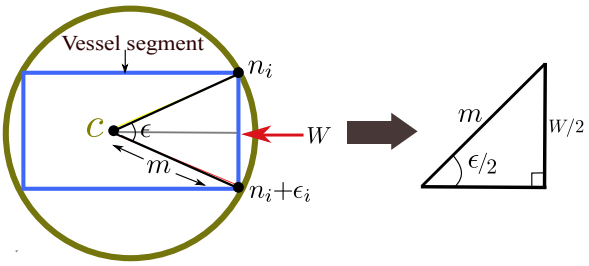



Figure 4.10: Vessel width estimation for a sample vessel fragment. $c$ is the crossing point; $\epsilon$ is the angular span between vessel end points; $W$ is the cross-sectional vessel width; and $m$ is the radial distance.

Thus, given a crossover point $c$, VKD(c) is analysed to first identify if $c$ represents a *simple* or *complex* case, and next identify the vessel segments of interest using the appropriate form of VKD.

#### 4.3.1.2 Vessel Width Profile Estimation

The AV nicking quantification requires an estimation of vessel widths of all crossing vessel segments. The vessel widths at a crossover point is estimated using VKD.

Each vessel segment at a crossover point $c$ can be characterised by the cross-sectional width $W$ and an angular span $\epsilon$ measured with respect to $c$. Since the vessel characteristics can change along its length, both $W$ and $\epsilon$ can also vary along its length. Consider a cross-section of the vessel at a radial distance of $m$, with the end points of vessel subtending an angle $\epsilon$ at $c$ (see Figure 4.10). It can be noted that $W$ is the chord length in a circle of radius $m$. Hence, $W$ can be computed from its relation to subtended angle and radius as

$$W = 2 \times m \times \sin(\epsilon/2)\,. \tag{4.7}$$

In an ideal scenario, the width of a normal vessel segment is expected to be constant over a limited range of radial distance $|m| \in [M_{min}, M_{max}]$ whereas, the angular span will vary with the distance. However, in practice, due to imperfect vessel segmentation, the width of normal vessels is also variable along the vessel length. Thus, both $W$ and $\epsilon$ modelled as functions of radial distance $m$ can exhibit variations.

A sample AV nicking patch is shown in Figure 4.11 (a). The width profile in Figure 4.11 (c) and Figure 4.11 (d) are for nicking ($V_1, V_3$) and normal ($V_2$,$V_4$) vessel segments respectively. In this plot, the crossover point corresponds to the

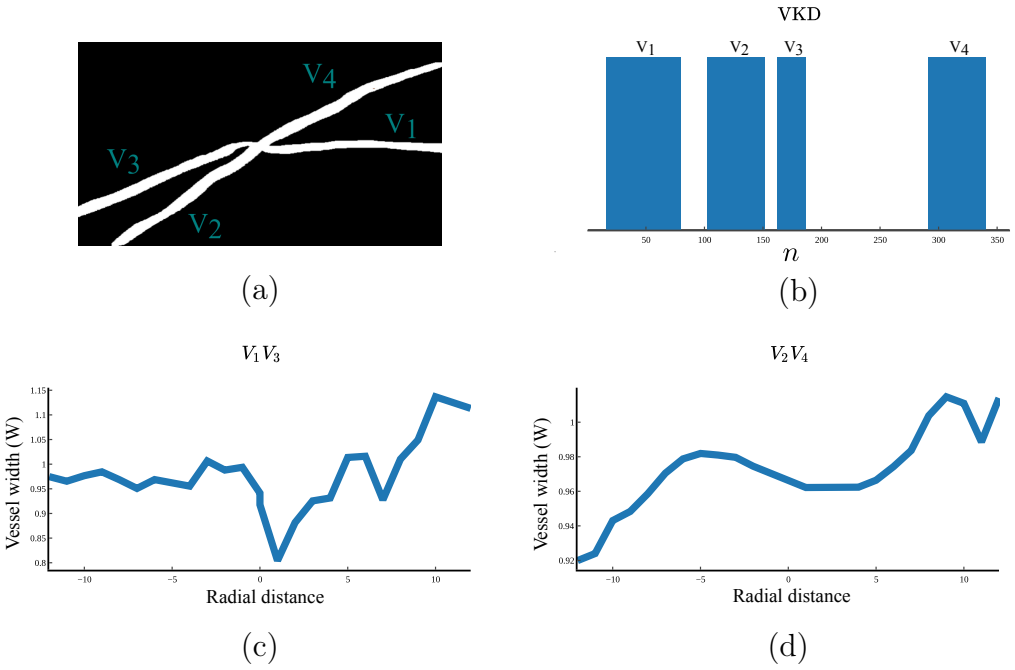



Figure 4.11: Vessel width profile estimation. (a) Sample AV nicking patch; (b) Corresponding VKD; (c) and (d) The vessel width profiles ($W$) for the segments ($V_1, V_3$) and ($V_2, V_4$), respectively. *Note*: the vessel widths are shown at different scales for (c) and (d).

zero radial distance point. The nicking segment exhibits a gradual downward trend (Figure 4.11 (c)) as one approaches towards the origin and a sharp minimum close to origin. Further, as one moves past the origin, the vessel width jumps to a higher value and remains more or less constant. The plot for normal vessel segment in Figure 4.11 (d) is at a magnified scale, which enhances the variability in width due to imperfect vessel segmentation. These observations are consistent with trends in VKD (see Figure 4.11 (b)), which shows a significant decrease in angular span of responses (*the third set*) for the nicking vessel fragment $V_3$ and similar responses for the normal fragments ($V_2$, $V_4$).

In summary, the proposed approach solely rely on VKD to estimate the angular span $\epsilon$ of a vessel segment as well as its width $W$ to quantify AV nicking. While, the former may be sufficient for the task, $W$ is also considered to build robustness to width variations – often introduced by binarization in the vessel segmentation process. In contrast, earlier methods rely on A/V classification and use only the vein width to assess the AV nicking severity (Nguyen et al. (2013a); Pereira et al. (2014); Roy et al. (2014)).

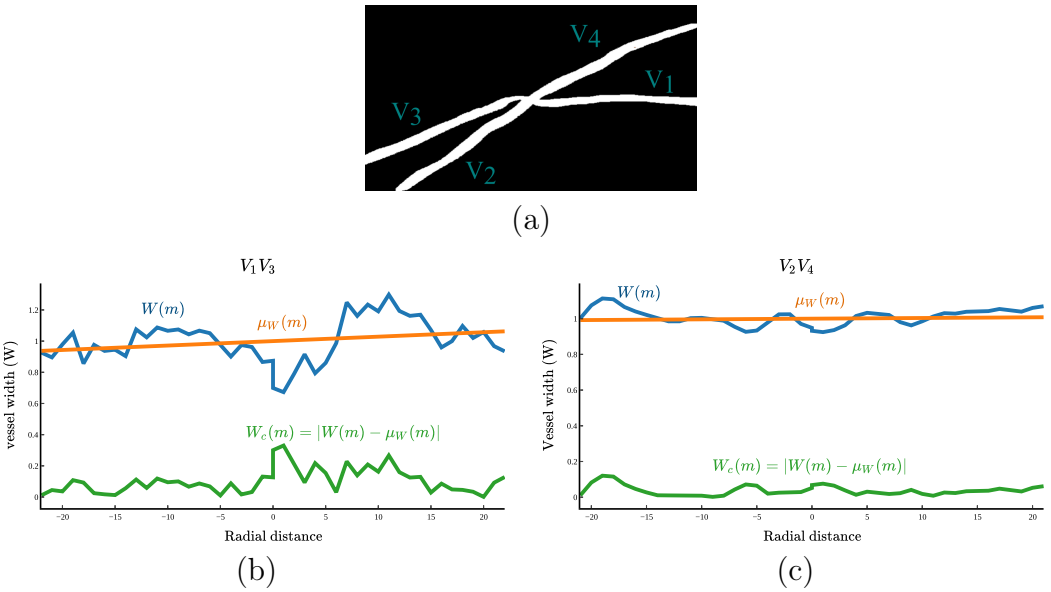



Figure 4.12: Vessel width as a function of radial distance. (a) A sample AV nicking patch; (b) and (c) Vessel width profiles $W(m)$ (blue), and mean-corrected result $W_c(m)$ (green) for nicking ($V_1$ $V_3$) and normal ($V_2$ $V_4$) segments, respectively. $\mu_W$ is the best fit line for $W$.

#### 4.3.1.3 AV Nicking Classification

The two different solutions for automatic classification of AV nicking has been proposed. Nicking is quantified at one of four severity levels: $0, 1, 2, 3$; with 0 denoting normal; 1 denoting mild; 2 denoting moderate; and 3 denoting severe nicking case. Of these two solutions, one is based on an ***unsupervised*** approach to underscore the richness of VKD while, the other is based on a ***supervised*** approach.

***i) Unsupervised Approach:*** In this approach, the mean vessel width is used to characterize a normal vessel, and the deviation is analysed and used to quantify the AV nicking severity.

Given a crossover location $c$, for each pair of vessel segments, the vessel width ($W$) profile is available from the previous stage (see Section 4.3.1.2). This profile varies significantly both within and across images. This needs to be compensated first. A straight line is fit to the width profile $W$ of every segment to estimate the mean width ($\mu_W$) of the segment. The mean-corrected width profile has information about nicking behaviour if any. This profile is found using an absolute difference (since sign is not of interest) as follows

$$W_c(m) = \Big(|W(m) - \mu_W(m)|\Big), \tag{4.8}$$

where, $W_c(m)$ is the mean-corrected width profile and $m$ is the radial distance index. A sample crossover patch with nicking is shown in Figure 4.12, along with the corresponding $W$ and $W_c$ profiles. A sharp rise in the peak is observed near the crossover point in $W_c$ for $(V_1V_3)$ due to nicking. Comparatively, $W_c$ remains almost constant for the normal vessel segment $(V_2V_4)$. Since the change in vessel width is captured well by $W_c$, it can be used to quantify AV nicking. The change is captured by computing the variation in $W_c$ via the mean absolute difference, as it is generally less sensitive to extreme outliers that can occur due to poor vessel segmentation process. This is defined as

$$\sigma_W = \frac{1}{n} \sum_{m=-M_{max}}^{M_{max}} \Big| W_c(m) - \overline{W_c} \Big|, \tag{4.9}$$

where, $|M_{max}|$ – is the maximum radial distance considered in $W_c$ computation; $\overline{W_c}$ represents the mean value of $W_c$; and $\sigma_W \in [0, 1]$ represents the mean change in vessel width. Since, AV nicking occurs only in the region close to a crossover point (see Figure 4.9 (a)), the computation of $\sigma_W$ is limited to $|m| \in [M_{min} : M_{max} = \alpha M_{min}]$ for both positive and negative values of $m$. Defining $M_{max}$ in terms of $M_{min}$ serves to adapt the $\sigma_W$ computation to a given crossover patch and minimise tuning.

Since every crossover point has a pair of vessel segments (artery-vein), $\sigma_W$ is computed for both vessel segments and combined to obtain the final AV nicking severity level, which is computed as

$$\text{AVN}(c) = max\Big(\sigma_W^1, \sigma_W^2\Big), \tag{4.10}$$

where, $\sigma_W^1$ and $\sigma_W^2$ are the mean change in widths of two crossing vessel segments. $\text{AVN}(c) \in [0, 1]$ represents the AV nicking severity of a crossover point $c$ ; $\text{AVN}(c) = 0$ indicates a normal crossover location; while, $\text{AVN}(c) > 0$ indicates mild to severe AV nicking. Finally, the severity class label for a crossover patch is obtained by thresholding $\text{AVN}(c)$.

***ii) Supervised Approach:*** The motivation for exploring a supervised method is to improve the classification performance. Both angular span $(\epsilon)$ and cross-sectional width $(W)$ quantify the vessel width. These are used as features to train a RF classifier. This combined feature set should be robust to improper vessel binarization acquired during the segmentation process and works well irrespective of various

image resolutions.

The feature vector ($F$) corresponding to each individual crossing vessel segments (artery-vein) is defined as a tuple based on the estimated $\epsilon$ and $W$ as

$$F = [m, \epsilon(m), W(m)], \quad \forall m \in [-M_{max}, +M_{max}] \tag{4.11}$$

where, $F$ is the feature vector corresponding to a crossing vessel segment $V_s$ (which may be A/V). At a crossover, the two such feature vectors are obtained – one corresponding to an artery and the other to a vein.

As done earlier for computation of $\sigma_W$, the range of $m$ cannot be chosen to be $[M_{min} : M_{max} = \alpha M_{min}]$, since this will result in a feature vector with variable length depending on $M_{min}$ for a crossover patch. Hence, $|m| \in [M_{min} : M_{max} = M_{min} + l]$ pixels are chosen. The length of the final feature vector for each vessel segment $V_s$ is $[2(l+1) \times 3]$.

Thus, given a patch with crossover point $c$, the features $F$ corresponding to each crossover vessel segments $V_s$ are derived using Equation (4.11). The label ($L_c$) for a crossover patch (AV nicking severity) is estimated based on the label $L_s$ of each crossing vessel segment $V_s$ as follows

$$\begin{aligned} &\text{if } (L_s == \text{normal}), \text{ then } L_c = \text{normal} \\ &\text{else if } (L_s == \text{nicking}), \text{ then } L_c = \text{nicking}. \end{aligned} \tag{4.12}$$

A RF classifier (Breiman (2001)) with 300 trees was trained with the obtained features $F$ along with vessel labels $L_s$. Finally, the label $L_c$ of a crossover point is determined based on the predicted labels $L_s$ of each crossover vessel segment $V_s$, as defined in Equation (4.12).

## 4.4 Experiments and Results

### 4.4.1 Junction Classification

Experiments were done on five publicly available datasets. The details are provided in Table 4.1. Ground truth (GT) labels for bifurcations and crossovers were derived for ARIA and CHASE_DB1 from two independent markers, while for DRIVE, IOSTAR and STARE were shared by Fathi et al. (2013), Abbasi-Sureshjani et al. (2016). As three independent markings exist for DRIVE (Fathi et al. (2013), Abbasi-

Table 4.1: Dataset split used in experiments. $J_n$ denotes the number of junctions. Subscripts denote the type: bifurcations (b) and crossovers (c).

| Dataset | Images | Training | Testing |
|---|---|---|---|
| | | $J_n(J_b/J_c)$ | $J_n(J_b/J_c)$ |
| DRIVE (Staal et al., 2004) | 40 | 2684 (1923/761) | 2064 (1513/551) |
| IOSTAR (Abbasi-Sureshjani et al., 2016) | 24 | 757 (550/207) | 939 (663/276) |
| STARE (Hoover et al., 2000) | 20 | 2561 (1266/1295) | 2489 (1291/1198) |
| CHASE_DB1 (Fraz et al., 2012b) | 28 | 1575 (698/877) | 1335 (580/755) |
| ARIA (Aria, 2006) | 40 | 972 (682/290) | 1168 (815/353) |
| **Total** | **152** | **8549** | **7995** |

Sureshjani et al. (2016), Azzopardi and Petkov (2013)), a majority consensus was taken to derive the GT. A half random split was done to generate the training and test sets from each dataset. The number of vessel junctions (bifurcations and crossovers) considered for evaluation is listed in Table 4.1. The performance of the proposed approach was evaluated with the following metrics: sensitivity/recall ($Se$), specificity ($Sp$), accuracy ($Acc$), precision ($Pr$), $F_1$-score ($F$) and area under the ROC curve ($AUC$). Detected junctions within a radius of 4 pixels of GT points were taken to be true positives ($TPs$).

#### 4.4.1.1 Junction Detection Results

A class imbalance exists between junctions and non-junctions classes in the set of candidates. This was handled with oversampling by rotating the patches in arbitrary directions. An RF classifier of 200 trees was used for classification. The obtained values for different performance metrics are listed for the proposed method and the state-of-the-art method (Abbasi-Sureshjani et al. (2016)) in Table 4.2. The proposed approach can be seen to outperform BICROS (Abbasi-Sureshjani et al. (2016)) on both DRIVE and IOSTAR, while it achieved the best $Se$, $Sp$ and $F$ on all datasets. The higher $Se$ is due to the discriminative power of VKD in differentiating vessel junctions from the complex background. The COHD is able to eliminate the false positives among the clusters of $C$ leading to a higher $Sp$. Results reported by Aibinu et al. (2010) on DRIVE and STARE are only on a subset of images precluding a thorough comparison. The GRAID method in Núñez et al. (2015) is evaluated using two different GT's on DRIVE, whereas the data provided in Azzopardi and Petkov (2013) is the only one publicly available. Hence, the comparison is restricted to

Table 4.2: Comparative evaluation of junction detection performance.

| Method | Dataset | *AUC* | *Acc* | *Se* | *Sp* | *Pr* | *F* |
|---|---|---|---|---|---|---|---|
| BICROS | DRIVE | - | - | 0.74 | - | 0.81 | 0.67 |
| | IOSTAR | - | - | 0.65 | - | 0.77 | 0.61 |
| **Proposed method** | DRIVE | 0.79 | 0.75 | 0.79 | 0.66 | 0.81 | 0.80 |
| | IOSTAR | **0.90** | **0.85** | **0.87** | 0.81 | **0.94** | **0.90** |
| | STARE | 0.83 | 0.78 | 0.80 | 0.73 | 0.89 | 0.85 |
| | CHASE_DB1 | **0.90** | 0.81 | 0.80 | **0.84** | 0.93 | 0.86 |
| | ARIA | 0.79 | 0.71 | 0.70 | 0.74 | 0.87 | 0.78 |

Table 4.3: Comparative evaluation of junction classification performance.

| Method | Dataset | *AUC* | *Acc* | *Se* | *Sp* | *Pr* | *F* |
|---|---|---|---|---|---|---|---|
| BICROS | DRIVE | - | 0.83 | 0.59 | 0.91 | - | - |
| | IOSTAR | - | 0.83 | 0.67 | **0.93** | - | - |
| **Proposed method** | DRIVE | 0.93 | 0.87 | 0.86 | 0.87 | 0.70 | 0.77 |
| | IOSTAR | **0.94** | 0.89 | **0.88** | 0.89 | 0.76 | 0.82 |
| | STARE | 0.91 | 0.85 | **0.88** | 0.84 | 0.62 | 0.73 |
| | CHASE_DB1 | 0.91 | 0.87 | 0.83 | 0.88 | 0.67 | 0.74 |
| | ARIA | 0.93 | **0.90** | 0.84 | 0.92 | **0.81** | **0.83** |

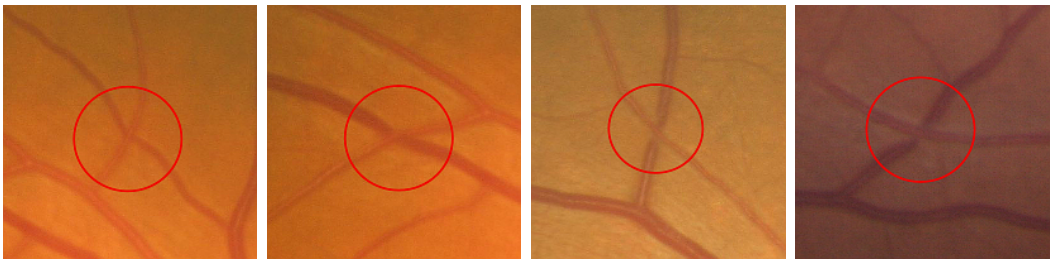
Azzopardi and Petkov (2013). The results on CHASE_DB1 and ARIA data sets are also consistently high. Overall, the obtained results indicate that the proposed method is robust to changes in image resolution.

#### 4.4.1.2 Junction Classification Results

Once again, a class imbalance is observed between the bifurcation and crossover classes which was handled by oversampling - by rotating the patches in arbitrary directions. An RF classifier with 500 trees was trained. Table 4.3 depicts the bifurcation versus crossover classification performance. There is a boost in *Se* of $\approx$ 20% relative to BICROS method. Azzopardi and Petkov (2013) reports only vessel bifurcation performance, without taking into account the complex bifurcation/crossover structures. Similarly, Aibinu et al. (2010); Fathi et al. (2013) have reported the bifurcation/crossover classification performance independently against the background. Hence, the proposed approach cannot be directly compared with those methods. In general, most of the previous techniques are not able to model the complex bifurcation/crossover points whereas the proposed approach is able to better discriminate even a complex bifurcation/crossover points which have not been addressed previously in the literature. Hence, the proposed method is more robust and reliable for junction classification task.

Table 4.4: Distribution of 90 AV nicking patches in the dataset.

| **AV nicking severity level** | **0** (normal) | **1** (mild) | **2** (moderate) | **3** (severe) |
|---|---|---|---|---|
| Number of crossover patches | 48 | 16 | 14 | 12 |



(a) (b) (c) (d)

Figure 4.13: Sample patches in different severity classes. a) Normal (0); b) Mild (1); c) Moderate (2); d) Severe (3).

### 4.4.2 AV Nicking Quantification

The binary vessel map required as input to the system was obtained from color retinal images using a multi-scale line detection method (Nguyen et al. (2013b)).

#### 4.4.2.1 Dataset

The proposed solution was assessed using a public dataset of fundus images acquired in two population-based studies: *Blue Mountains Eye Study* (BMES) (Cugati et al. (2007); Mitchell et al. (1995); Wang et al. (1999)); and the *Singapore Malay Eye Study* (SiMES) (Cheung et al. (2011b); Foong et al. (2007)). The dataset has 90 crossover patches extracted from the 47 high-resolution ($401 \times 401$ pixels) images (Nguyen et al. (2013a)). Each of these 90 crossover locations was manually graded by two independent experts at the Centre for Eye Research Australia using a 4-scale grading system (from 0 for normal, 1 for mild, 2 for moderate and 3 for severe nicking case). Any disagreement between two experts was reassessed in a joint session to derive a single grading for each crossover point. The distribution of AV nicking severity levels for the 90 crossover points is shown in Table 4.4, and sample patches with different AV nicking severity levels are shown in Figure 4.13.

Three-levels of discriminations which are of clinical relevance were considered for assessing AV nicking severity. First is the [0] vs. [1,2,3], which is the discrimination of normal from other abnormal cases. Second is a variant of the first [0,1] vs. [2,3], which clubs mild and normal cases, and aims to discriminate it from moderate to

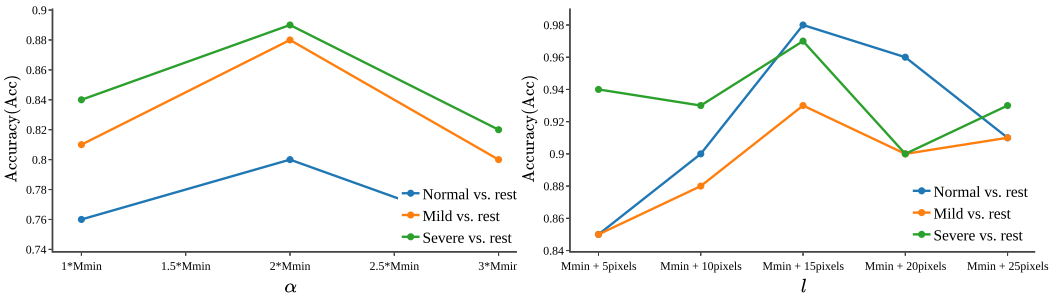



Figure 4.14: Effect of ROI parameter choice on classification accuracy for proposed (a) Unsupervised; (b) Supervised approaches.

severe AV nicking cases. This is referred to as 'mild vs. rest' for convenience. The third is [0,1,2] vs. [3], which is the discrimination of severe nicking cases from all others. All assessment was done with four performance metrics: sensitivity ($Se$), specificity ($Sp$), accuracy ($Acc$) and area under the ROC curve ($AUC$).

#### 4.4.2.2 Parameter Selection

The proposed AV nicking quantification solution is governed by two main parameters: i) identification of appropriate ROI around a crossover point; ii) effective feature-set that precisely quantifies decreasing vessel widths. These two parameters play a critical role for accurate identification and quantification of AV nicking at a crossover location. The details are presented next.

*i)* ***Identification of ROI:*** The parameters of ROI, namely $M_{min}$ and $M_{max}$, greatly influences the identification of AV nicking at a crossover point. This is because of the variable nature of vessel thickness and changes in vessel width around a crossover point. Ideally, the choice of ROI should be independent of image resolution both within and across datasets.

Figure 4.14 shows the effect of varying size of ROI on AV nicking detection performance. The immediate vicinity of the crossover point $c$ belongs to multiple (at least 4) vessel fragments of different thickness, and hence the $M_{min} = 0$ leads to ambiguity in vessel width measurement. A solution to find the correct $M_{min} > 0$ is as follows. In VKD computation, as the radial distance $m$ is increased, the number of distinct vessel signatures at $c$ starts at 1 and becomes $\geq 4$. The latter point is taken as $M_{min}$, and it corresponds to the diameter of the largest crossing vessel segment. The value of $M_{max}$ is fixed based on this $M_{min}$. In the case of unsupervised approach, $M_{max} = \alpha M_{min}$. The effect of varying $\alpha$ on the accuracy of AV nicking classification

Table 4.5: Effect of choice of vessel width characterization on classification performance with the *unsupervised* approach. $W$: cross-sectional width and $\epsilon$: angular span.

| Classification | Feature set | AUC | Acc | Se | Sp |
|---|---|---|---|---|---|
| Normal vs. rest | $W$ | **1.00** | **1.00** | **1.00** | **0.92** |
| | $\epsilon$ | 0.79 | 0.72 | 0.24 | 0.92 |
| Mild vs. rest | $W$ | **0.94** | **0.90** | **0.88** | **0.92** |
| | $\epsilon$ | 0.76 | 0.73 | 0.24 | 0.92 |
| Severe vs. rest | $W$ | **0.92** | **0.88** | **0.83** | **0.94** |
| | $\epsilon$ | 0.81 | 0.76 | 0.36 | 0.94 |

Table 4.6: Effect of choice of vessel width characterization on classification performance with the *supervised* approach. $W$: cross-sectional width and $\epsilon$: angular span. All results are with 3-fold validation.

| Classification | Feature set | AUC | Acc | Se | Sp |
|---|---|---|---|---|---|
| Normal vs. rest | $W$ | 0.96 | 0.92 | 0.90 | 0.92 |
| | $\epsilon$ | 0.84 | 0.85 | 0.78 | 0.92 |
| | $W + \epsilon$ | **1.00** | **0.98** | **0.99** | **0.92** |
| Mild vs. rest | $W$ | 0.96 | 0.94 | 0.90 | 0.92 |
| | $\epsilon$ | 0.85 | 0.80 | 0.57 | 0.92 |
| | $W + \epsilon$ | **0.96** | **0.93** | **0.90** | **0.92** |
| Severe vs. rest | $W$ | 0.93 | 0.90 | 0.70 | 0.94 |
| | $\epsilon$ | 0.88 | 0.85 | 0.57 | 0.94 |
| | $W + \epsilon$ | **0.98** | **0.97** | **0.91** | **0.94** |

can be seen in Figure 4.14 (a). The accuracy peaks for $\alpha = 2$, and this value is chosen in the experiments. In the case of supervised approach, $M_{max} = [M_{min} + l]$ pixels and the effect of varying $l$ on accuracy is shown in Figure 4.14 (b). The first peak is at $l = 15$, which is chosen for the experiments. Both these choices of ROI are independent of varying image resolutions both within and across datasets as they are adapted to a given crossover patch.

The accuracy of AV nicking quantification peaks only at certain limited ROI (see Figure 4.14) is because, the AV nicking which is characterized by decrease in vein widths generally occurs only at a region very close to a crossover point. Further, as one moves away from this limited ROI, the vein width gradually increases and remains more or less constant (see Figure 4.11 (c)). Hence, the choice of ROI is very crucial in determining the optimal performance of an AV nicking quantification system.

*ii)* ***Effectiveness of Features***: Next, the results for assessing the relative importance of the features ($\epsilon$ and $W$) for AV nicking classification are presented. Table 4.5 and Table 4.6, depicts how each individual vessel width features affect the performance of AV nicking classification with unsupervised and supervised approach, respectively. Although, $\epsilon$ is a quantifier for vessel width, it is highly susceptible to irregularities due to vessel segmentation process. This can be seen in the low $Se$ values for all levels of discriminations (see Table 4.5 and Table 4.6), when relying only on $\epsilon$. This clearly indicates the failure in accurately identifying the true positives. Even a slight change in the vessel structure can lead to erroneous result in AV nicking quantified values. In contrast, the cross-sectional width $W$ has a better discriminative capability in estimating vessel width changes, as indicated by an improvement in $Se$ by at least $> 10\%$ in almost all levels of discriminations for both unsupervised and supervised cases. Finally, in the case of supervised approach, a combination of $W$ and $\epsilon$ (see Table 4.6) is seen to be most effective for classification at all severity levels. This combination is equally successful in the two extreme tasks: discriminating between normal and all nicking cases ([0] vs. [1,2,3]), as well as between severe and all other cases ([3] vs. [0,1,2]), with $AUC \geq 0.98$.

#### 4.4.2.3 Evaluation of Proposed AV Nicking Classification

AV nicking is essentially a multi-class classification problem with normal (0) and 3 severity (1, 2, 3) levels. The proposed method for classification is assessed in a one *vs.* all manner for two reasons: the first being the insufficient number of cases available for some classes; and the second is to facilitate a direct comparison with results in Nguyen et al. (2013a); Roy et al. (2014).

Comparative evaluation was done against two methods. One being Nguyen et al. (2013a); and the other with COSFIRE filter (Azzopardi and Petkov (2013)), which was initially designed to detect only vascular bifurcations from retinal images. The COSFIRE filter was extended in this work for AV nicking quantification such that, no artery-vein classification was required to enable a fair comparison against the proposed VKD-based approach. The crossover points were identified first with VKD and the responses of a bank of Gabor filters (which forms the core of COSFIRE) were used for AV nicking quantification. The responses were mapped to log-polar coordinates $(\rho, \phi)$ where, $\rho \in [M_{min}, M_{max}]$ and $\phi$ sampled at every $1^\circ$ interval. The values of $M_{min}$ and $M_{max}$ were chosen as described in Section 4.4.2.2. The other

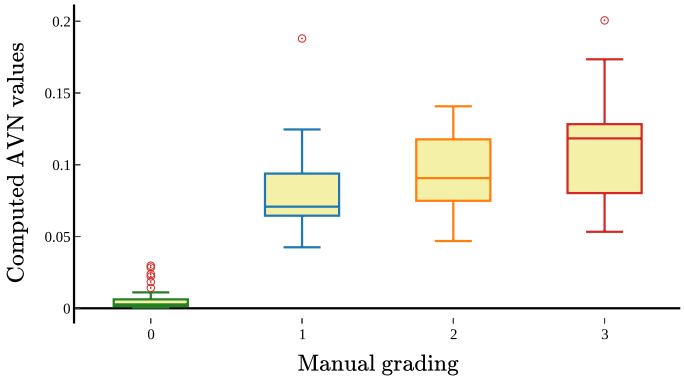



Figure 4.15: Computed AV nicking values (AVN($c$)) for different nicking severity classes for the unsupervised approach.

Table 4.7: Comparative analysis of the proposed *unsupervised* methods. '*' - denotes the proposed extension of the method reported in Azzopardi and Petkov (2013).

| Classification | Method | AUC | Acc | Se | Sp |
|---|---|---|---|---|---|
| Normal vs. rest | **Proposed method** | **1.00** | **1.00** | **1.00** | **0.92** |
| | Nguyen et al. (2013a) | 0.86 | 0.80 | 0.67 | 0.92 |
| | Azzopardi and Petkov (2013)* | 0.86 | 0.84 | 0.36 | 0.92 |
| Mild vs. rest | **Proposed method** | **0.94** | **0.90** | **0.88** | **0.92** |
| | Nguyen et al. (2013a) | 0.92 | 0.88 | 0.77 | 0.92 |
| | Azzopardi and Petkov (2013)* | 0.87 | 0.81 | 0.32 | 0.92 |
| Severe vs. rest | **Proposed method** | **0.92** | 0.88 | **0.83** | **0.94** |
| | Nguyen et al. (2013a) | 0.91 | **0.89** | 0.58 | 0.94 |
| | Azzopardi and Petkov (2013)* | 0.88 | 0.81 | 0.45 | 0.94 |

two parameters $\lambda$ (wavelength) and $\theta$ (orientation) of the Gabor filters were chosen as the default set of parameters as described in Azzopardi and Petkov (2013).

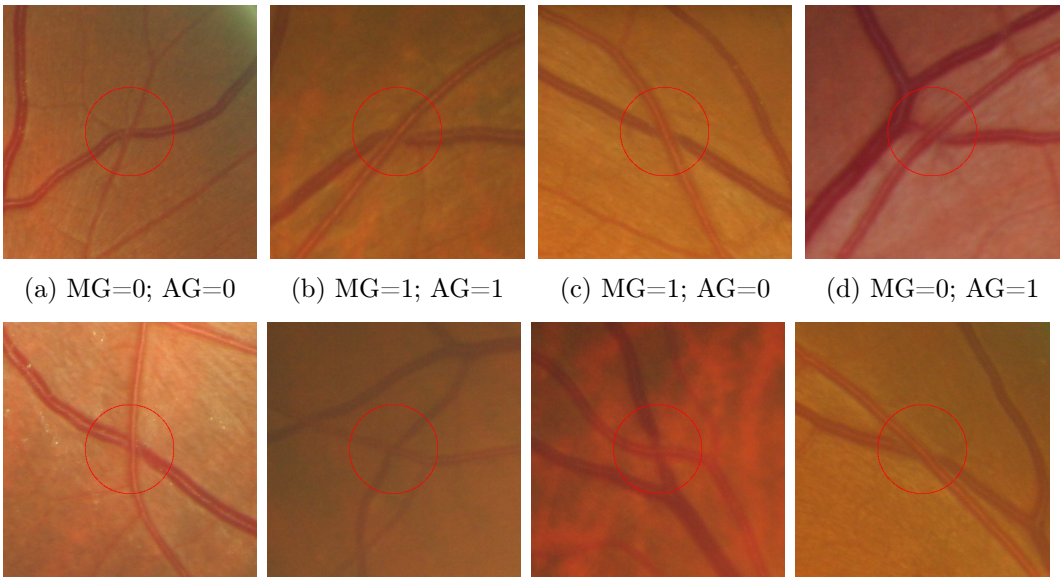
*i)* ***Unsupervised Approach:*** In this approach, either $\epsilon$ or $W$ can be used to quantify the AV nicking. However, based on the results in Table 4.5, the $W$ is chosen for the detailed classification assessment. The box-whisker plot in Figure 4.15 shows the range of computed AV nicking values AVN($c$) corresponding to the 90 crossover patches against manual grading (groundtruth). Overall, a clear association appears to exist between the objective measurement (AVN($c$)) and subjective grading. It can be noted that the nicking classes (1, 2, 3) are clearly distinguishable from the normal (0) class whereas, among the nicking classes, this is not the case especially between classes 1 and 2. The plot also indicates that the class 3 is quite discriminable against all other classes. These observations are also consistent with the classification results in Table 4.5.

Table 4.7, also shows that the proposed method outperforms Nguyen et al. (2013a) and Azzopardi and Petkov (2013), for all 3 types classification. All $Se$ values are reported for a fixed $Sp = 0.92$ or 0.94 to facilitate fair comparison. The improvement in $Se$ is $> 30\%$ for normal versus other classes while, $Se$ is at least $> 11\%$ for the other two cases. The proposed method has less false positives compared to Azzopardi and Petkov (2013)*, mainly because the COSFIRE filter selectively responds only to the local vessel patterns (such as shapes) rather than the width of the local vessel structures. The proposed method also outperforms Nguyen et al. (2013a) and Azzopardi and Petkov (2013)*, in terms of $AUC$ and $Acc$. The time for complete AV nicking quantification approach, including junction detection, was found to be about 46 seconds per image, using an unoptimized MATLAB code, on a 2.2GHz Intel i3-2330 processor with 2GB of RAM.

*ii)* ***Supervised Approach:*** A 3-fold validation was adopted for evaluation. Samples of each of the 4 classes were divided into 3 sets of equal size. A 300-tree RF classifier was trained using two sets selected from each class. Testing was done on the remaining set of patches from each class. Since the dataset (of 90 patches) has significant class imbalance, class weighting was done, which renders classification cost-sensitive by penalizing misclassification of the minority class (Breiman (2001)). Thus, the minority class samples were given a larger weight (i.e., higher misclassification cost).

Sample results of the supervised classification are shown in Figure 4.16. These include successfully predicted cases (first two columns) and failure cases (last two columns). It is observed that the proposed method handles even very low contrast crossover patches well (see Figure 4.16 (a, b, f)). This is primarily because, the proposed solution is not dependent on A/V classification of crossover segments and depends solely on vessel morphology. The performance in Figure 4.16 (a, e), underscores this benefit as in these cases, both arteries and veins appear very similar due to central vessel reflex, and this would be a challenging case for any automatic artery/vein classification approach. The successful prediction of the case in Figure 4.16 (a, f) is also remarkable given that the vessel width change is markedly subtle. Most failures (see Figure 4.16 (c, d, g, h)) are due to discontinuity in the vessel structures at crossover location which leads to an erroneous binary vessel map.

The confusion matrix for the proposed supervised nicking classification is presented in Table 4.8, which shows that the proposed method correctly classifies 79



(a) MG=0; AG=0 (b) MG=1; AG=1 (c) MG=1; AG=0 (d) MG=0; AG=1

(e) MG=0; AG=0 (f) MG=1; AG=1 (g) MG=1; AG=0 (h) MG=0; AG=1

Figure 4.16: Sample classification results obtained with the supervised approach. First two columns: successfully predicted cases; Last two columns: failure cases. *Note*: 'M(A)G' denote manual (automatic) grading; '0' refers to negative class; and '1' for positive class. (a-d) Results for mild vs. rest separation; (e-h) Results for severe vs. rest separation.

Table 4.8: Confusion matrix of the proposed *supervised* method.

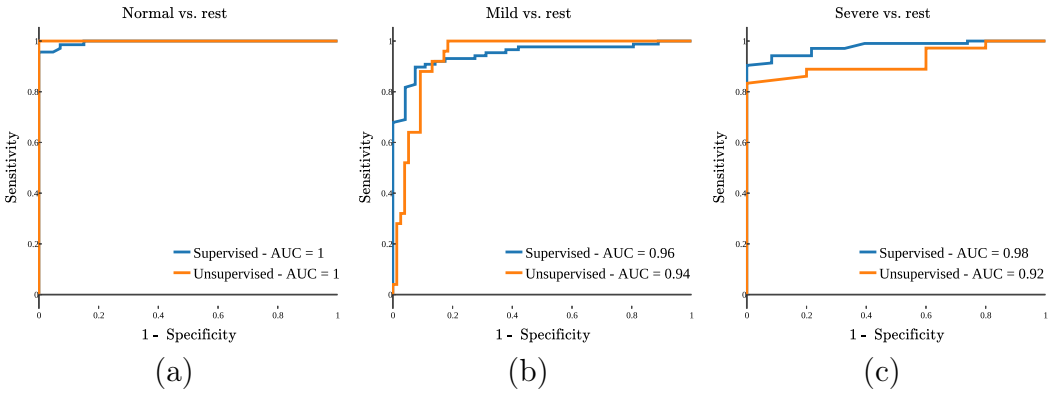
| | | **Predicted** | | | | **Total** |
|---|---|---|---|---|---|---|
| | **Severity** | 0 | 1 | 2 | 3 | |
| **Actual** | 0 | **45** | 2 | 1 | 0 | 48 |
| | 1 | 1 | **13** | 1 | 1 | 16 |
| | 2 | 0 | 2 | **11** | 1 | 14 |
| | 3 | 0 | 1 | 1 | **10** | 12 |
| **Correctly predicted** | | 45 | 13 | 11 | 10 | **79/90** |
| **Accuracy** | | **93.75%** | **81.25%** | **78.57%** | **83.33%** | **87.77%** |

Table 4.9: Confusion matrix of the method in Roy et al. (2014).

| | | **Predicted** | | | | **Total** |
|---|---|---|---|---|---|---|
| | **Severity** | 0 | 1 | 2 | 3 | |
| **Actual** | 0 | 40 | 10 | 1 | 0 | 51 |
| | 1 | 8 | 5 | 2 | 2 | 17 |
| | 2 | 4 | 1 | 6 | 3 | 14 |
| | 3 | 1 | 0 | 1 | 9 | 11 |
| **Correctly predicted** | | **40** | **5** | **6** | **9** | **60/93** |
| **Accuracy** | | **78.43%** | **29.41%** | **42.85%** | **81.81%** | **64.51%** |

Table 4.10: Confusion matrix of the method in Azzopardi and Petkov (2013)*. *Note*: '*' - denotes the proposed extension of the method reported in Azzopardi and Petkov (2013).

| | | Predicted | | | | Total |
|---|---|---|---|---|---|---|
| | **Severity** | 0 | 1 | 2 | 3 | |
| **Actual** | 0 | 41 | 5 | 2 | 0 | 48 |
| | 1 | 5 | 7 | 2 | 2 | 16 |
| | 2 | 1 | 2 | 9 | 2 | 14 |
| | 3 | 0 | 1 | 3 | 8 | 12 |
| **Correctly predicted** | | **41** | **7** | **9** | **8** | **65/90** |
| **Accuracy** | | **85.41%** | **43.75%** | **64.28%** | **66.66%** | **72.22%** |




(a) (b) (c)

Figure 4.17: ROC curves for the unsupervised and supervised approaches. (a) Normal vs. rest; (b) Mild vs. rest; (c) Severe vs. rest. All results are for 3-fold validation for the supervised approach.

crossover points out of total 90 crossovers, with an accuracy rate of 87.77% across all severity levels. Two supervised versions namely: Roy et al. (2014) and Pereira et al. (2014) exists for AV nicking classification. The method in Pereira et al. (2014) reports an average *AUC* of 85% for normal vs. rest and severe vs. rest on the same dataset while, Roy et al. (2014) reports a classification *Acc* of 64.51% on a slightly larger (93) set of patches. Table 4.9 and Table 4.10 presents the confusion matrix for Roy et al. (2014) and Azzopardi and Petkov (2013)*, respectively. Comparing the figures tabulated in the two confusion matrices, it is concluded that there is an improved classification performance in almost all levels of AV nicking classification for the proposed method, indicating that the system is more accurate and reliable for developing an automated AV nicking quantification solution. The time for complete AV nicking quantification approach, including junction detection, takes about 60 seconds per image, using an unoptimized MATLAB code, on a machine equipped with 2.2GHz Intel i3-2330 processor and 2GB of RAM.

The receiver operating characteristic (ROC) plots for AV nicking classification are presented in Figure 4.17 for both the proposed unsupervised and supervised approaches. An $AUC$ of $\geq 0.92$ is obtained across all levels of discriminations for both approaches. Both normal vs. rest and severe vs. rest classifications are seen to be accurately handled by the proposed supervised solution with a $AUC \geq 0.98$, demonstrating that the proposed system is effective in detecting both normal and severe AV nicking cases equally well. This is mainly due to the discriminative power of VKD in detecting a very subtle change in vessel widths at a crossover location.

## 4.5 Discussion and Summary

In this chapter, a complete solution for the assessment of retinal vasculature mainly at vessel keypoints has been presented. This includes junction detection, crossover identification, vessel width estimation and AV nicking quantification. The entire pipeline is based on a vessel keypoint descriptor that was introduced in Srinidhi et al. (2017b).

The main advantage of the proposed solution is that it does not depend on A/V classification or intensity-based features for quantifying AV nicking, in contrast to the earlier methods (Nguyen et al. (2013a); Pereira et al. (2014); Roy et al. (2014)). The proposed approach starts with a binary vessel map and looks at AV nicking quantification as a generic vessel width estimation problem, irrespective of the type of vessel. The proposed approach was also compared against the extension of COSFIRE filter (Azzopardi and Petkov (2013))* to AV nicking quantification, requiring no knowledge of artery-vein labels to further evaluate the robustness of the proposed VKD approach. This was demonstrated to lead better quantification of vascular changes at all crossover locations, irrespective of the type of crossing vessel segment.

The vessel width derived from VKD in the case of unsupervised approach, serves as an effective objective measure of AV nicking severity while, the features based on variations in the vessel width and angular span in case of supervised approach, serves to effectively represent the vessel segments (normal/abnormal). In the latter case, an AV nicking is probabilistic (and continuous-valued), which is attractive as such possibility is unavailable in the human grading system. Further, the proposed objective level measure can be useful in strengthening the correlation between AV nicking and microvascular diseases. Any similar measurements when manually done

are subjected to large intra and inter-observer variability. The proposed methods are robust to variations in image contrast, vessel definition and require practically no tuning. Based on the superior classification performance ($Se = ACC = AUC = 1.0$ and $Sp = 0.92$) with the unsupervised method for normal (0) vs. nicking (at any severity level) cases, it is concluded that such a simple approach is adequate if only nicking detection is of interest. Whereas, the supervised method is preferable if it is also of interest to segregate cases by severity levels.

Most failures in AV nicking quantification (refer, Figure 4.16 (c, d, g, h)) are mainly due to missing vessel connectivity information at crossover locations, inherited during vessel segmentation process. Many state-of-the-art vessel segmentation methods (Azzopardi et al. (2015); Liskowski and Krawiec (2016)) are highly susceptible to varying contrast or non-uniform illumination, the presence of low-intensity thin vessel structures and complete occlusion of vessel crossovers at peripheral areas. These challenges often lead to an incomplete representation of vascular connectivity between tiny vessel crossovers, proximal multiple junction locations and crossing of vessels with a very low angle. The above limitations can be further addressed using a more sophisticated method for closing the broken gaps in the binarized vessel maps (Zhang et al. (2018)); or by using a very recently proposed robust segmentation methods (Srinidhi et al. (2018)).

The proposed VKD can also be easily extended to A/V classification post nicking identification, as the segment which exhibits nicking will be a vein, and any vessel tracing method can be used to propagate this label along the vessel tree. In addition, the VKD can also be further extended to find the branching angle and tortuosity of vessels, which can further aid in extending quantitative vessel analysis for early diagnosis of diseases such as hypertensive retinopathy, diabetic retinopathy, diabetic neuropathy and dementia.

# Chapter 5

# Automated Method for Retinal Artery/Vein Separation via Graph Search Metaheuristic Approach

Separation of the vascular tree into arteries and veins is a fundamental prerequisite in the automatic diagnosis of retinal biomarkers associated with systemic and neurodegenerative diseases. In this chapter, a novel graph search metaheuristic approach is presented for the automatic separation of arteries/veins (A/V) from color fundus images. The method exploits local information to disentangle the complex vascular tree into multiple subtrees, and global information to label these vessel subtrees into arteries and veins. Given a binary vessel map, a graph representation of the vascular network is constructed representing the topological and spatial connectivity of the vascular structures. Based on the anatomical uniqueness at vessel crossing and branching points, the vascular tree is split into multiple subtrees containing arteries and veins. Finally, the identified vessel subtrees are labeled with A/V based on a set of hand-crafted features trained with random forest classifier. The proposed method has been tested on four different publicly available retinal datasets with an average accuracy of 94.7%, 93.2%, 96.8% and 90.2% across AV-DRIVE, CT-DRIVE. INSPIRE-AVR and WIDE datasets, respectively. These results demonstrate the superiority of the proposed approach in outperforming state-of-the-art methods for A/V separation.

## 5.1 Background

Retinal microcirculation offers a unique non-invasive way to study the early manifestation of several diseases affecting the human circulatory system. Changes in retinal vascular geometrical patterns (such as width, tortuosity, branching angle, junction exponents and fractal dimension) specific to artery/vein (A/V) are investigated as candidate biomarkers with incident clinical stroke (Wong et al. (2001b)), hypertension (Cheung et al. (2012)), arteriosclerosis (Hubbard et al. (1999)), dementia (Frost et al. (2013)), and other cerebral small vessel diseases (Wong et al. (2002)). For instance, the narrowing of arteries and widening of veins is a significant indicator of the progression of diabetic retinopathy (DR) (Abramoff et al. (2010)), hypertension (Hubbard et al. (1999)), and various other cardiovascular abnormalities (Wong et al. (2006)). Specifically, the arteriolar-to-venular diameter ratio (AVR) is a prognostic indicator of stroke, cerebral atrophy, cognitive decline and myocardial infarction (Niemeijer et al. (2011)). Therefore, accurate analysis and quantification of vessel-specific morphological changes may provide an early insight into better understanding the pathophysiology of the disease conditions.

The retinal fundus photography is an excellent non-invasive technique most commonly used to analyse and quantify the vascular abnormalities in large-scale clinical settings, due to its speed and affordability. Manual separation of A/V from color fundus image is extremely time-consuming and requires an enormous amount of painstaking manual process. Hence, developing an automated tool for separation of A/V is of paramount importance in large-scale retinal disease screening programs.

Many methods have been introduced in the past for retinal A/V separation, with methods focused on either graph-based (Dashtbozorg et al. (2014); Estrada et al. (2015); Hu et al. (2015); Joshi et al. (2014); Pellegrini et al. (2018); Rothaus et al. (2009)) or feature based techniques (Claudia Kondermann (2007); Grisan and Ruggeri (2003); Huang et al. (2018a,b); Mirsharif et al. (2013); Niemeijer et al. (2009); Vazquez et al. (2010); Welikala et al. (2017); Xu et al. (2017); Zamperini et al. (2012)). Among these methods, graph-based techniques rely on establishing a graph structure by uniquely representing an entire vessel tree into multiple subtrees based on utilizing the vessel connectivity information at crossing and bifurcation points. These identified vessel subtrees are further separated into A/V segments based on the pixel-wise classification of vessel centerlines. On the other hand, the

feature-based techniques solely rely on pixel-level intensity information to classify vessels into A/V.

Among the feature based methods, the earliest approach for A/V separation was proposed by Grisan and Ruggeri (2003). They presented a classification technique only in a well-defined concentric zone around the optic disc region. A similar approach based on the combination of clustering and vessel tracing method was also proposed in Vazquez et al. (2010). Claudia Kondermann (2007), Niemeijer et al. (2009) and Mirsharif et al. (2013) explored a wide set of pixel-wise features, along with different set of classifiers to obtain an optimal A/V labeling. Zamperini et al. (2012) proposed effective feature set based on color, spatial location and vessel width, across different region-of-interest (ROI) measured from optic disc region. More recently, Welikala et al. (2017) presented a method based on a deep neural network approach. Their method was validated on a large population-based cohort study dataset known as UK Biobank. Xu et al. (2017) proposed first and second order textural features, along with the intensity level features to obtain a discriminative feature set for A/V separation. A novel feature extraction technique based on luminosity and reflection properties of vascular structures has been explored in Huang et al. (2018b). A genetic search based feature selection technique for high dimensional data has been recently proposed in Huang et al. (2018a).

The significant limitations of the feature based approaches are two folds. First, due to the input image acquisition process, retinal images exhibit varying contrast and luminosity, often resulting in difficulty in distinguishing A/V segments of thin and peripheral vessels. Second, the absence of vessel connectivity information leads to difficulty in precisely tracking A/V segments of branching and crossover points. To address these issues, graph-based approaches have gained increasing interest by incorporating the structural characteristics of the retinal vascular tree. These methods exploit the distinct nature of the underlying retinal vascular connectivity pattern, that the arteries and veins will cross each other, but never with themselves (Dashtbozorg et al. (2014); Estrada et al. (2015); Joshi et al. (2014)).

Based on this assumption, several graph-theoretic approaches have been explored in the past to improve the A/V classification performance. The earliest method was proposed by Rothaus et al. (2009) based on a semi-automatic technique, by solving a constrained graph search optimization problem. Joshi et al. (2014) presented a technique by dividing the entire vessel tree into individual subtrees, by finding

an optimal path using the Dijkstra algorithm. These individual subtrees are further labeled as A/V based on a set of orientation, width, and intensity features. Dashtbozorg et al. (2014) proposed a similar strategy by first subdividing the entire vascular tree into multiple sub-graphs based on the type of intersection points, followed by the assignment of A/V label to each vessel sub-graphs based on a set of intensity features. Estrada et al. (2015) presented an A/V classification method for both fundus as well as scanning laser ophthalmoscope (SLO) images. They constructed a global likelihood model based on carefully designed domain-specific features to estimate the underlying vascular topology. A novel graph-based metaheuristic approach exploiting the vascular connectivity was proposed by Hu et al. (2015). Finally, Pellegrini et al. (2018) proposed a novel graph cut based global optimization technique for optimal A/V separation in an ultra-wide field of view (UWFoV) SLO images, requiring no manual intervention.

Despite the considerable improvements in graph-based techniques, retinal A/V separation still suffers some difficulties. Most existing graph-based approaches mainly rely on geometrical analysis of vessel keypoints to efficiently exploit the underlying structural characteristics of a vascular network. In particular, the traditional approaches (Dashtbozorg et al. (2014); Estrada et al. (2015); Hu et al. (2015); Joshi et al. (2014); Pellegrini et al. (2018)) utilizes only the orientation, width and intensity level information to address the challenging "crossover issue" encountered during a graph search process. However, one important bottleneck is that these techniques do not take into account the curvature characteristics of vascular segments, which is often critical for successful disentanglement of a highly curved crossover segment. For example, at a highly curved crossover point, the angular separation between A/V pairs will be highly skewed depending on the nature and complexity of crossing vessel segments. The vessel geometrical properties such as width, tortuosity, bifurcation/crossover angle are also prone to variations in the presence of the diseased conditions. Thus, any geometrical modelling without the inclusion of curvature information will usually produce a highly suboptimal representation of the vascular network. Further, an imperfect separation of crossing vessel pairs leads to an erroneous propagation of labels throughout the entire graph search path, resulting in inaccurate A/V separation.

In this chapter, a novel graph search metaheuristic approach is presented for the automatic A/V separation from retinal color fundus images. Here, the previously

proposed vessel keypoint descriptor (VKD) (Srinidhi et al. (2017b) (Refer, Chapter 4)) is extended to incorporate the curvature characteristics of crossing vessel segments, along with the orientation and width information. This curvature property acts as a unique feature to VKD to aid in resolving the possible conflict, in assigning an A/V label of the highly curved crossover point. Besides, a novel graph search metaheuristic algorithm has also been presented to generate anatomically meaningful vessel subtrees, by searching the space of possible connectivity of vascular networks.

The key contributions of this chapter can be summarized as follows:

i. An extended vessel keypoint descriptor is proposed which integrates curvature with the orientation and width information to precisely disentangle all crossing vessel pairs into corresponding A/V segments.

ii. A novel depth-first search based graph search metaheuristic algorithm is proposed to accurately identify all A/V vessel subtrees from a given vascular topology.

iii. The results of the proposed method has been extensively validated on four challenging publicly available retinal datasets, including images from two different imaging modalities - fundus as well as UWFov-SLO images.

The chapter is organized as follows: The details of the proposed methodology is presented in Section 5.2. Datasets and experimental results are described in Section 5.3. Finally, the key findings as well as the limitations of the work are discussed in Section 5.4, followed by summary in Section 5.5.

## 5.2 Methods

An overview of the proposed solution for automatic separation of A/V trees from retinal images is shown in Figure 5.1. The pipeline consists of four main stages: identification of vessel keypoints, a graph representation of the vascular network, vessel subtree extraction and subtree A/V labeling. In the first stage, given a binary vessel map, the vessel keypoints are identified using a "vessel keypoint descriptor" (VKD). In the second stage, the identified vessel keypoints are considered as nodes and vessel segments as edges that represents an undirected graph structure of the vascular network. In the third stage, the anatomically meaningful vessel subtrees are identified using a novel graph search metaheuristic approach by exploiting the

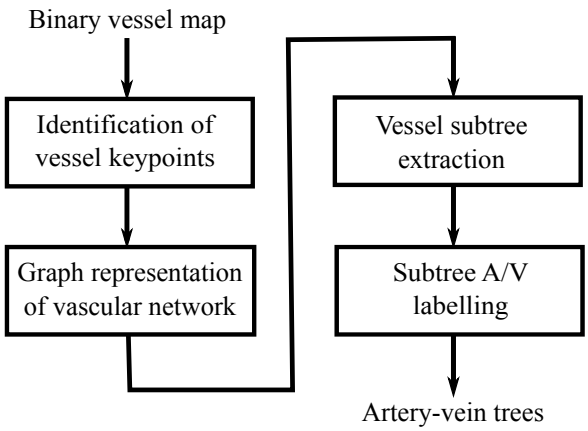



Figure 5.1: An overview of the proposed A/V separation approach.

structural connectivity of the vascular network. Finally, in the fourth stage, each vessel subtrees are given a unique A/V label based on carefully designed hand-crafted features, which are then trained using a random forest (RF) classifier.

### 5.2.1 Identification of Vessel Keypoints

The proposed approach starts by identifying vessel keypoints such as bifurcations, crossovers and vessel endpoints from a given binary vessel tree. A ROI $R_p(x, y)$ is extracted for every vessel point $p$ in a vessel map. The vessel centerline extraction process is eliminated to preserve the vessel connectivity information and to avoid errors introduced by thinning operation. A log-polar transform (LPT) is applied to the ROI to obtain $R_p(m, n)$ where, $m$ and $n$ are the radial and angular indices respectively. The LPT preserves information close to a vessel point while, increasingly compressing the information as one moves away from the vessel point, in a non-linear fashion. Figure 5.2 (first and second row) shows the ROI for different vessel patterns of interest (branching vessels, crossing vessels, straight vessel segment and a vessel endpoint) and the corresponding log-polar mapped results. It can be observed that the number of vertical lines depend on the pattern of the vessel segment and their position depends on the orientation of vessels in $R_p(m, n)$. Where, $m \in [M_{min}, M_{max}]$; and $n$ is the angle index, which is obtained by sampling the angular variable at $1^\circ$ interval, and hence $n \in [0^\circ, 360^\circ]$. A vertical projection of $R_p(m, n)$ results in a vector $R_p(n)$, as illustrated in Figure 5.2 (third row). In order to build robustness to spurious vessels and varying vessel calibre, projections at a limited set of radii ($m$) are considered. The obtained projection $R_p(n)$ provides a count of the number of pixels in a vessel fragment at a specific angle $n$. Since, only the presence or absence

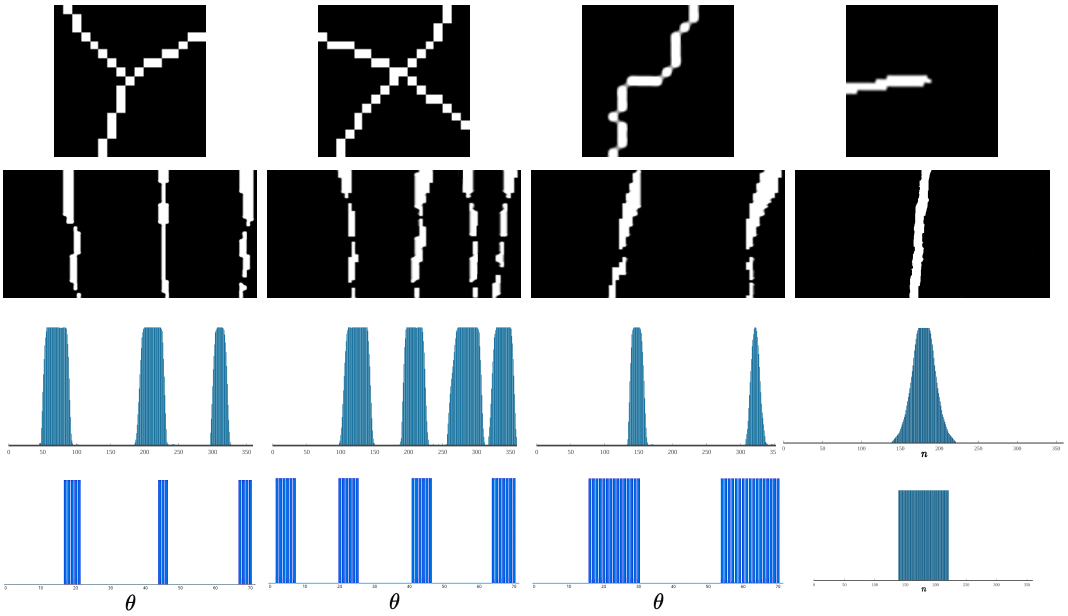



Figure 5.2: Illustration of VKD. Row 1: (a) Sample bifurcation; (b) Crossover; (c) Non-junction; and (d) A vessel endpoint patch. Row 2: The corresponding log-polar maps (with x-axis being $n$). Row 3: The projections of log-polar maps ($R_p(n)$). Row 4: Thresholding applied to $R_p(n)$; this is the VKD descriptor.

of vessel at a particular angle is of interest, $R_p(n)$ is thresholded, with $threshold = 1$ to obtain the VKD $V_p(n)$.

The VKD for four sample ROI's are shown in Figure 5.2 (fourth row). Each vessel fragment gives rise to a cluster of responses in $V_p(n)$. Hence, the vessel edges are found by computing the first order difference of $V_p(n)$ as

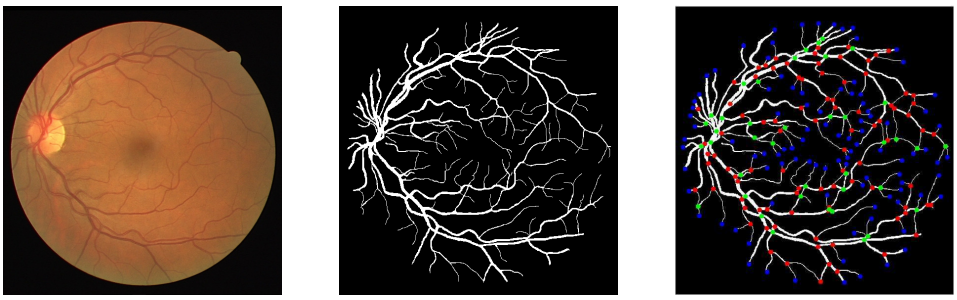
$$V'_p(n) = \Big| V_p(n+1) - V_p(n) \Big|. \tag{5.1}$$

The number of vessel branches $S$ at a point $p$ is given as

$$S(p) = 0.5 \times \sum_n V'_p(n). \tag{5.2}$$

Finally, a set of keypoints is obtained as: $C = \{K | 1 \leq S(p) \geq 3\}$. The detected keypoints appear in the form of clusters close to junctions, which are then refined to localize the vessel keypoints. The desired candidate keypoints $w$ are identified from cluster $C$ by computing the entropy for every keypoint $K$ in a $3 \times 3$ local neighbourhood, followed by non-maximum suppression within a radius of 12 pixels.

Identifying and classifying true keypoints from given $w$ is challenging, due to



(a) (b) (c)

Figure 5.3: (a) Original image; (b) Binary vessel map; (c) Identified vessel keypoints. (Note: the color red represents bifurcation point, green represents crossover point and blue represents vessel endpoints).


(a) (b)

Figure 5.4: (a) Graph representation of retinal vascular tree; (b) A sample ROI enlarged for visualization.

the proximal presence of bifurcation and crossover points, nearby multiple junction locations, close parallel and highly curved vessels. To this end, a combination of four features is employed such as $\Big(V_w(n), R_w(n)\Big)$; basic line detector response (LD) (Ricci and Perfetti (2007)); and histogram of oriented gradients (HOG) (Dalal and Triggs (2005)), computed at every candidate keypoints $w$ in a $(17 \times 17)$ neighborhood. RF classifier (Breiman (2001)) with 500 trees is trained with these features to identify the bifurcation $(b)$, crossover $(c)$ and vessel end $(e)$ points. The detected vessel keypoints for a sample image is shown in Figure 5.3. These keypoints form the input to the A/V separation module, which is described next.

### 5.2.2 Graph Representation of Vascular Network

A graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ is constructed with $\mathcal{V}$ being the node and $\mathcal{E}$ being the edge of a graph. The edge $\mathcal{E}_{ij}$ represents a vessel segment that connect two nodes $(\mathcal{V}_i, \mathcal{V}_j)$ in graph $\mathcal{G}$. There exists three different types of node $\mathcal{V}$ in $\mathcal{G}$ such as:

1. Bifurcation nodes $(\mathcal{V}_b)$ – correspond to vessel bifurcation points $(b)$ (vessels of same type (A/V) bifurcates into two branches).

2. Crossover nodes ($\mathcal{V}_c$) – correspond to vessel crossover points ($c$) (vessels of two different types (artery-vein) crossover each other).

3. Vessel end nodes ($\mathcal{V}_e$) – correspond to vessel end points ($e$).

The graphical illustration of different types of node are shown in Figure 5.4.

### 5.2.3 Vessel Subtree Extraction

In order to identify anatomically meaningful vessel trees (A/V), the graph $\mathcal{G}$ is divided into multiple subtrees $\mathcal{S}_\mathcal{G}$ using a novel graph search metaheuristic approach. The proposed graph search method is based on the two *anatomical* uniqueness of retinal vasculature:

- at *bifurcation* points, only vessels of same type bifurcate into different branches (i.e., arteries will bifurcate into arteries and same with veins).

- at *crossover* points, arteries and veins will cross each other, but never with themselves.

Based on these assumptions, a two-step solution is proposed for extracting the vessel subtrees from a graph $\mathcal{G}$ as follows: (i) identification of A/V segments at a crossover location; (ii) depth-first search (DFS) based graph search approach.

#### 5.2.3.1 Identification of A/V Segments at Crossover Location

A crossover point is a location where an artery crosses a vein, often leading to four vascular fragments. Out of these four, the two diagonally opposite vessel pairs belong to the same class (A/V) as shown in Figure 5.5 (first and second column). In practice, there can be more than four fragments, due to branching and crossover points occurring very close to each other as shown in Figure 5.5 (third and fourth column). The main bottleneck for the accurate extraction of vessel subtrees exist in identifying the appropriate A/V vessel pairs at crossover location. To this end, the curvature characteristics of all crossing vessel fragments is modelled with the aid of VKD, to tackle the challenging crossover issue encountered during graph search.

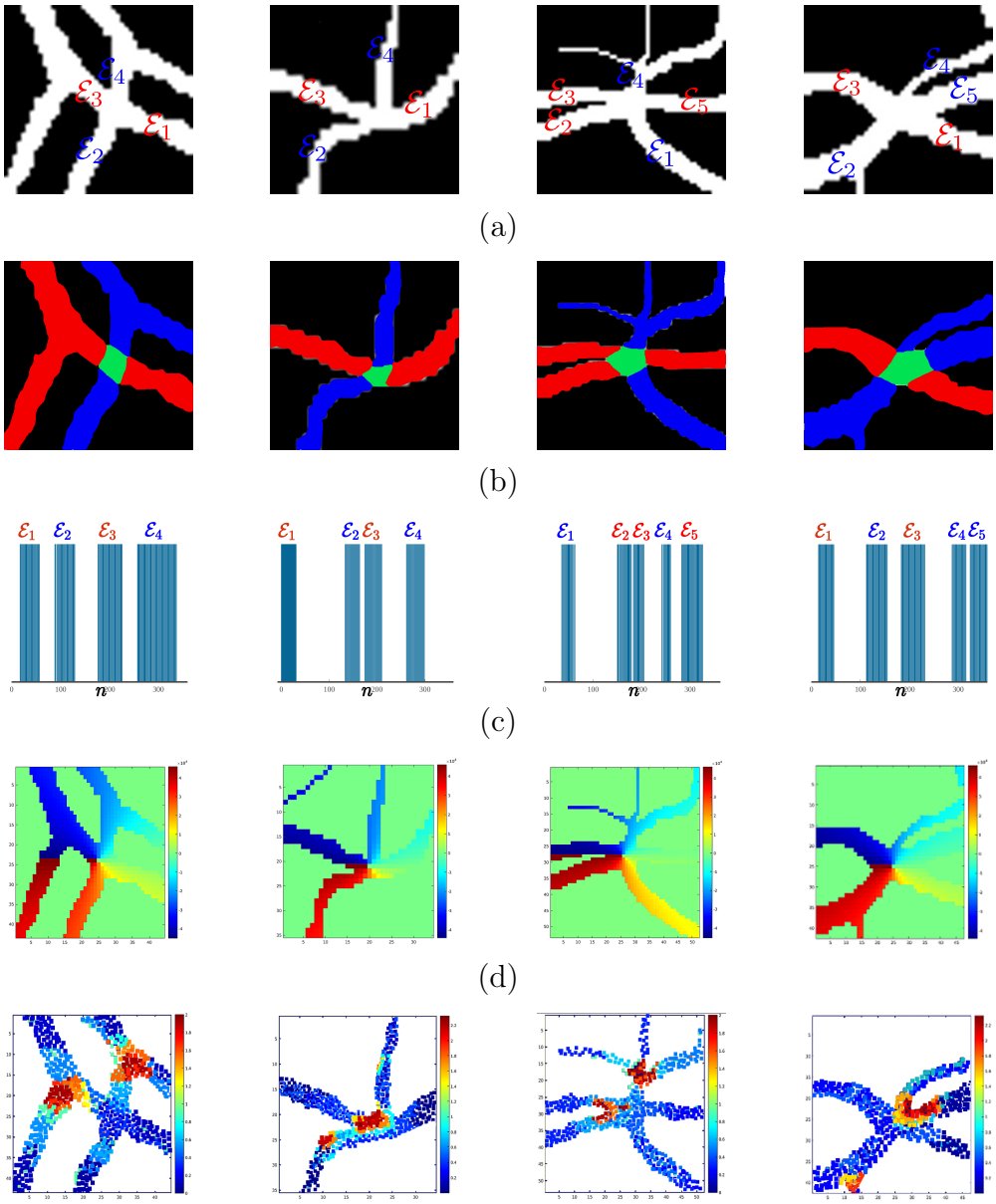
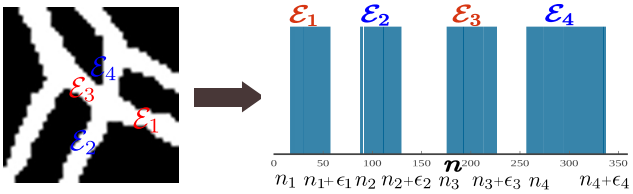



Figure 5.5: (a) Binary vessel map; (b) The corresponding A/V labels (Note: the color red represents an artery, blue represents a vein and green represents a crossover point); (c) The corresponding VKD $V_p(n)$ (with x-axis being $n$); (d) The corresponding orientation heat-maps; (e) The corresponding curvature heat-maps.




(a) (b)

Figure 5.6: Identification of corresponding pair of vascular fragments ($\mathcal{E}_i$ and $\mathcal{E}_{ip}$). (a) A sample patch with a crossover point; (b) Corresponding VKD.

Given a crossover point $c$ (see Figure 5.6), the VKD($c$) is given as

$$\begin{aligned}\text{VKD}(c) = &\Big[u\Big(n-n_1\Big) - u\Big(n-(n_1+\epsilon_1)\Big)\Big] \\ &+\Big[u\Big(n-n_2\Big) - u\Big(n-(n_2+\epsilon_2)\Big)\Big] \\ &+\cdots+\Big[u\Big(n-n_S\Big) - u\Big(n-(n_S+\epsilon_S)\Big)\Big] \\ &=\sum_{i=1}^{S}\Big[u\Big(n-n_i\Big) - u\Big(n-(n_i+\epsilon_i)\Big)\Big], \end{aligned} \tag{5.3}$$

where, $\Big[u\Big(n-n_i\Big) - u\Big(n-(n_i+\epsilon_i)\Big)\Big] \triangleq \mathcal{E}_i$ – refers to a specific vessel fragment ($\mathcal{E}_i$) belonging to either an artery/vein; $\epsilon$ represents the angular span of a vessel fragment; $n \in [0°, 360°]$ - assuming an angular sampling rate of 1°; and $S$ is the number of vessel fragments converging at a crossover point, which is typically $\geq 4$. Figure 5.5 shows four sample artery-vein crossings (first row) and their corresponding VKDs in the third row. The first two columns represent a case of *simple* crossovers ($S = 4$ vessel fragments) and the last two columns represent a *complex* crossover case ($S > 4$ vessel fragments). The cluster of responses in VKD indicates the presence of at least one vessel fragment at a specific angle, within a window of $\geq M_{min}$. The value of $M_{min}$ corresponds to the diameter of the largest crossing vessel segment.

From a given *crossover* point $c$, a *three* important observations is made:

1. For every vessel fragment $\mathcal{E}_i$, there exists a vessel pair $\mathcal{E}_{ip}$ belonging to the same class (A/V), which is often separated by $n = \pm 180°$ in the VKD such that

$$\mathcal{E}_{ip} = u\Big(n-(n_i \pm 180° \pm \gamma)\Big) - u\Big(n-\big((n_i \pm 180° \pm \gamma)+\epsilon_i\big)\Big), \tag{5.4}$$

where, $\mathcal{E}_i$ and $\mathcal{E}_{ip}$ are vessel fragment pairs belonging to the same vessel seg-

ment (artery/vein), which are separated by an angle of $180° \pm \gamma$, as illustrated in Figure 5.5 (c). $\gamma$ is a small factor which accounts for highly *curved* crossover vessel segments. In Figure 5.5 (first and second column), ($\mathcal{E}_1$ and $\mathcal{E}_3 = \mathcal{E}_{1p}$) or ($\mathcal{E}_2$ and $\mathcal{E}_4 = \mathcal{E}_{2p}$) are examples of ($\mathcal{E}_i$ and $\mathcal{E}_{ip}$). This is seen from the examples in Figure 5.5, where the labeled fragments that are collinear are shown in red or blue (third row). This implies that a pair of vessel fragments $(\mathcal{E}_i, \mathcal{E}_{ip})$ which belongs to a single segment (A/V) are *collinear*.

2. The angular span ($\epsilon$) of a vessel fragment which is measured with respect to $c$ as the origin is *correlated* to its thickness. This implies a positive correlation between the width of a vessel and the corresponding cluster in VKD as shown in Figure 5.5 (c). Hence for convenience, the angular span ($\epsilon$) of a vessel fragment is denoted as the vessel *width*. From Figure 5.5 (a and c), it is observed that the width of the crossing vessel pairs $(\mathcal{E}_i, \mathcal{E}_{ip})$ remains almost constant, irrespective of A/V, which can be written as

$$\epsilon\{\mathcal{E}_i\} \cong \epsilon\{\mathcal{E}_{ip}\}, \tag{5.5}$$

where, $\epsilon$ denotes the vessel width. This is also observed to be true even for complex crossover locations as shown in Figure 5.5 (third and fourth column).

3. Ideally, the vessel fragment pairs ($\mathcal{E}_i$, $\mathcal{E}_{ip}$) are given the same class label (A/V), if the angular separation of $\mathcal{E}_i \cong \mathcal{E}_{ip} \pm 180°$. But this may not always be true for high curvature crossing vessel segments as well as complex crossovers as shown in Figure 5.5 (second and fourth column). The *curvature* $\kappa$ is defined as "the rate of change of orientation" and determines the shape of the vessel segment. From Figure 5.5 (a and e), it is observed at vessel crossings, the orientation of each vessel segment changes slightly in most cases. This implies that the rate of change of orientation for a pair of crossing vessel fragments ($\mathcal{E}_i$ and $\mathcal{E}_{ip}$) remain constant. It is also observed to be true for complex crossover locations as shown in Figure 5.5 (second and fourth column). In addition, a set of experiments has been performed to detail the significance of curvature information for separation of A/V segments of different simple and challenging crossover cases, which are provided in detail in the Appendix B.

   Hence, this observation is made to estimate the curvature ($\kappa$) of crossing vessel

fragments ($\mathcal{E}_i$ and $\mathcal{E}_{ip}$) based on the following theorem.

*Theorem 1:* Suppose $m = R(n)$ represents a polar parametrization of a plane curve, then the curvature at any point $(m, n)$ is given by

$$\kappa(n) = \frac{\left|R(n)^2 + 2[R'(n)]^2 - R(n)R''(n)\right|}{\{R(n)^2 + [R'(n)]^2\}^{3/2}}, \tag{5.6}$$

where, $m$ and $n$ are radial and angular indices respectively.

*Proof:* Let $\vec{r}(t) = (x(t), y(t))$ be a vector-valued function that traces a smooth curve $\alpha$. Then the curvature ($\kappa$) of $\alpha$ at any point $\vec{r}(t)$ is given by (refer, Section 1.4 of Gray (1996))

$$\kappa(t) = \frac{||\vec{r'}(t) \times \vec{r''}(t)||}{||\vec{r'}(t)||^3}. \tag{5.7}$$

Suppose that the curve $\alpha$ is given in the polar form $m = R(n)$, then curve can be parametrized as $\vec{r}(n) = (m\ \cos(n),\ m\ \sin(n),\ 0) = (R(n)\ \cos(n),\ R(n)\ \sin(n),\ 0)$. Next, $\vec{r'}(n)$ is computed by differentiating $\vec{r}(n)$ as

$$\vec{r'}(n) = -R(n)\sin(n) + R'(n)\cos(n),\ R(n).\cos(n) + R'(n)\sin(n),\ 0 \tag{5.8}$$

Similarly, compute $\vec{r''}(n)$ by differentiating $\vec{r'}(n)$ as

$$\begin{aligned} \vec{r''}(n) = R''(n)\cos(n) - 2R'(n)\sin(n) - R(n)\cos(n), \\ R''(n)\sin(n) + 2R'(n)\cos(n) - R(n)\sin(n),\ 0. \end{aligned} \tag{5.9}$$

Now, compute the cross product $(\vec{r'}(n) \times \vec{r''}(n))$ as

$$\begin{vmatrix} \vec{i} & \vec{j} & \vec{k} \\ -R(n)\sin(n) + R'(n)\cos(n) & R(n)\cos(n) + R'(n)\sin(n) & 0 \\ R''(n)\cos(n) - 2R'(n)\sin(n) - R(n)\cos(n) & R''(n)\sin(n) + 2R'(n)\cos(n) - R(n)\sin(n) & 0 \end{vmatrix}$$

$$= 0,\ 0,\ 2(R'(n))^2 + (R(n))^2 - R(n). \tag{5.10}$$

Next, $||\vec{r'}(n) \times \vec{r''}(n)||$ is obtained as

$$\begin{aligned} &\sqrt{[2(R'(n))^2 + (R(n))^2 - R(n)R''(n)]^2} \\ &= |2(R'(n))^2 + (R(n))^2 - R(n)R''(n)|. \end{aligned} \tag{5.11}$$

Similarly, $||\vec{r'}(n)||$ is computed as

$$\sqrt{(R'(n))^2 + (R(n))^2} = [(R'(n))^2 + (R(n))^2]^{1/2} \tag{5.12}$$

Finally, the curvature ($\kappa$) of a plane polar curve is obtained at any point $(m, n)$ by substituting Equation 5.11 and Equation 5.12 in Equation 5.7 as

$$\kappa(n) = \frac{\left|R(n)^2 + 2[R'(n)]^2 - R(n)R''(n)\right|}{\{R(n)^2 + [R'(n)]^2\}^{3/2}}. \tag{5.13}$$

The above three observations at a crossover point are considered to identify the corresponding pair of vascular fragments $(\mathcal{E}_i, \mathcal{E}_{ip})$ that belongs to A/V segment, as explained next.

If $c$ is a *crossover* point, then there exists two scenarios:

1. A *simple* vessel crossover: a pair of vessels crossing each other at a point (see Figure 5.5 (first and second column)). In this case, VKD is of the form

$$\text{VKD}(c) = \sum_{i=1}^{2} \Big[\mathcal{E}_i + \mathcal{E}_{ip}\Big]. \tag{5.14}$$

   In this case, $\mathcal{E}_i$ and $\mathcal{E}_{ip}$ are given the same class label (A/V), if any of the two following *conditions* are satisfied:

   (a) the angular separation between two vessel pairs $(\mathcal{E}_i, \mathcal{E}_{ip})$ is often separated by $\pm 180°$:

$$\mathcal{E}_i \cong \mathcal{E}_{ip} \pm 180°; \tag{5.15}$$

   (b) the width of the crossing vessel pairs remains almost constant:

$$\epsilon\{\mathcal{E}_i\} \cong \epsilon\{\mathcal{E}_{ip}\}; \tag{5.16}$$

   (c) the curvature $\kappa$ of crossing vessel pairs is relatively constant:

$$\kappa\{\mathcal{E}_i\} \approx \kappa\{\mathcal{E}_{ip}\}; \tag{5.17}$$

2. A *complex* vessel crossover: a pair of vessels crossing at a point and there exists a *proximal* bifurcation (see Figure 5.5 (third and fourth column)). In

this case, VKD is of the form

$$\text{VKD}(c) = \sum_{i=1}^{2} \Big[\mathcal{E}_i + \mathcal{E}_{ip}\Big] + \mathcal{E}_l, \tag{5.18}$$

where, $\mathcal{E}_l \triangleq \Big[u\Big(n - n_l\Big) - u\Big(n - (n_l + \epsilon_l)\Big)\Big]$ corresponds to a lone vessel fragment resulting from a nearby bifurcation as shown in Figure 5.5 (third and fourth column). $\mathcal{E}_l$ - refers to vessel fragments $\mathcal{E}_3$ and $\mathcal{E}_4$ in third and fourth column, respectively. (*Note*: a complex crossover point generally consists of $\mathcal{E}_i > 4$ vessel fragments). Although, this is a very rare condition, it is still taken into account in order to build robust approach that works well irrespective of various complex crossover patterns. In this case, $\mathcal{E}_i$ and $\mathcal{E}_{ip}$ are given the same class label (A/V), if any of the two *conditions* similar to that defined for simplex crossover case are satisfied. In addition, the lone vessel fragment $\mathcal{E}_l$ is assigned a class label (A/V), if the following condition is satisfied

$$\epsilon\{\mathcal{E}_l\} \cong \epsilon\{\mathcal{E}_i\} \cong \epsilon\{\mathcal{E}_{ip}\}. \tag{5.19}$$

From Figure 5.5 (third and forth column), $\mathcal{E}_4$ and $\mathcal{E}_3$ are the lone vessel fragments $\mathcal{E}_l$, that are assigned the same class label as $\mathcal{E}_2$ and $\mathcal{E}_5$ in third and fourth column, respectively.

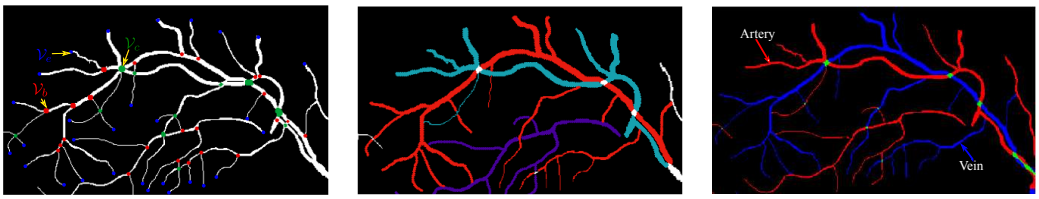
Thus, given a crossover point $c$, VKD($c$) is analysed first to identify if $c$ represents a *simple* or *complex* case and next identify the corresponding vessel pairs of the same class (A/V) using the appropriate form of VKD.

#### 5.2.3.2 Depth-First Search (DFS) Based Graph Search

To generate anatomically meaningful vessel subtrees $\mathcal{S}_\mathcal{G}$, a DFS based graph search algorithm is employed to efficiently search the space of possible vascular networks. In addition, the analysis using VKD near the vessel crossover points are used to aid in precise labeling of all crossing segments into corresponding A/V pairs. The method of extraction of individual subtrees $\mathcal{S}_\mathcal{G}$ involves the following steps.

1. Given an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ as shown in Figure 5.7 (a), initialize the graph search using DFS at arbitrary bifurcation node $\mathcal{V}_b$, with a label $l_i$ where, $i = 1, 2, ..., n$.

2. Find the neighbours associated with the node $\mathcal{V}_b$ denoted as: $\mathcal{N}_{\mathcal{V}_b}$.




(a) (b) (c)

Figure 5.7: (a) Binary vessel map (*Note:* the color red represents bifurcation nodes ($\mathcal{V}_b$), green represents crossover nodes ($\mathcal{V}_c$) and blue represents vessel end nodes ($\mathcal{V}_e$)); (b) The corresponding extracted vessel subtrees $\mathcal{S}_\mathcal{G}$, with different labels shown with different colors; (c) The corresponding subtree A/V labeling (arteries are shown in red and veins in blue).

3. If the associated neighbour $\mathcal{N}_{\mathcal{V}_b}$ is:

   - vessel end node ($\mathcal{V}_e$), then backtrack the DFS to $\mathcal{V}_b$.
   - vessel bifurcation node ($\mathcal{V}_b$), then continue the graph search with DFS.
   - vessel crossover node ($\mathcal{V}_c$), then set flag $\mathcal{F} = 1$ at $\mathcal{V}_c$.

4. For each crossover location $c$ (corresponding to $\mathcal{V}_c$ at which $\mathcal{F} = 1$) proceed with the following steps:

   - first identify if $c$ represents a *simple* or *complex* crossover point by computing VKD($c$) as given in Equation (5.14) and Equation (5.18).
   - find the vessel fragment pair $\mathcal{E}_{ip}$ corresponding to a vessel fragment $\mathcal{E}_i$ at which $\mathcal{F} = 1$, using the conditions defined in Equation (5.15), (5.16), (5.17), (5.19).
   - restart the graph search with DFS from the vessel fragment $\mathcal{E}_{ip}$, with the same label $l_i$ as assigned for vessel fragment $\mathcal{E}_i$.

5. The graph search is continued until there are no more nodes $\mathcal{V}$ to visit in the current search path.

6. Search for new paths that are left unlabeled (unvisited) in the graph $\mathcal{G}$. If found, start the DFS with a new label $l_i$ at an arbitrary bifurcation node $\mathcal{V}_b$, belonging to the new search path and continue from Step 2.

7. Continue the graph search until the entire graph $\mathcal{G}$ is labeled (as shown in Figure 5.7 (b)) and no more paths to visit in the vessel tree.

**Algorithm 2:** Subtree generation $\mathcal{S}_\mathcal{G}$

**Data:** Binary vessel tree $\mathcal{G} = (\mathcal{V}, \mathcal{E})$

**Parameters:** $\mathcal{V}_b$ = vessel bifurcation nodes
$\mathcal{V}_c$ = vessel crossover nodes
$\mathcal{V}_e$ = vessel end nodes
$\mathcal{T}_\mathcal{V}$ = type of node $\mathcal{V}$
$\mathcal{F}$ = flag raised at crossover nodes $\mathcal{V}_c$
$\mathcal{N}_\mathcal{V}$ = neighbours of node $\mathcal{V}$
$\mathcal{E}_i, \mathcal{E}_{ip}$ = corresponding vessel pairs

**Result:** Vessel subtrees $\mathcal{S}_\mathcal{G}$ with labels $l_i$ where, $i = 1, 2, .., n$

```
for i ← 1 to V do
    if T_V = V_b then
        for i ← 1 to N_{V_b} do
            start DFS at node V_b and assign a label l_i
            find N_V for node V_b
            if N_V = V_e then
                backtrack the DFS to V_b
            else if N_V = V_b then
                continue graph search with DFS
            else
                N_V = V_c; set flag F = 1 at V_c
            end
            while F = 1 do
                if c = simple crossover then
                    find the corresponding vessel pairs {E_i, E_ip} as defined in
                      Equation (5.15), (5.16), (5.17)
                else
                    c = complex crossover then
                    find the corresponding vessel pairs {E_i, E_ip} as defined in
                      Equation (5.15), (5.16), (5.17), (5.19)
                end
                if E_ip has already a label l_i then
                    terminate the DFS
                else
                    restart the DFS at E_ip and
                    continue the search with label l_i
                end

            end
        end
    end
end
```

The pseudocode for generating vessel subtrees $\mathcal{S}_\mathcal{G}$ from a graph $\mathcal{G}$ is illustrated in Algorithm 2.

There exists some misinterpretation of node labels as a result of vessel segmentation process (see Appendix C, Appendix D), which affects the optimal traversal of DFS based graph search algorithm. The following are the typical errors arises during node classification: (i) bifurcation node is wrongly classified as a crossover node and vice-versa; (ii) vessel end node is wrongly classified as bifurcation node. The detailed steps to correct these node errors are provided in Appendix D.

Each vessel subtree $\mathcal{S}_\mathcal{G}$ corresponds to different labels $l_i$, where $i = 1, 2, .., n$ as shown in Figure 5.7 (b) (*Note:* the subtrees generated doesn't imply A/V segments at this stage, and each vessel subtree is independent of the other having different labels $l_i$). Hence, to obtain the optimal A/V labels for the entire graph $\mathcal{G}$, each vessel subtree is given a unique label (A/V) based on a set of hand-crafted features extracted from vessel pixels corresponding to each vessel subtree, as described next.

### 5.2.4 Subtree A/V Labeling

For the identified vessel subtrees, the final goal is to assign an A/V label based on a set of hand-crafted features extracted from the vessel pixels, corresponding to each vessel subtree $\mathcal{S}_\mathcal{G}$. Since retinal images often exhibit varying contrast and luminosity, each input image is processed using the method proposed in Foracchia et al. (2005). A set of 66-D hand-crafted feature vector (as shown in Table 5.1) is extracted for every vessel pixel, and further normalized to zero mean and unit standard deviation. Some of these features were adopted in Dashtbozorg et al. (2014); Grisan and Ruggeri (2003); Niemeijer et al. (2011) and have shown to be robust for A/V separation. The vessel width and cross-sectional intensity features are detailed in Appendix E and Appendix F, respectively. These features were trained using RF classifier of 200 trees to predict the A/V label of each vessel subtree $\mathcal{S}_\mathcal{G}$. A thorough discussion on the performance comparison of different classifiers and feature selection techniques are also provided in Section 5.4.2.

A vessel subtree $\mathcal{S}_\mathcal{G}$ is assigned with a label artery (A), if the majority of the vessel pixels associated with a subtree are classified as arteries; else if the majority of the pixels are classified as veins, it is labeled with a vein (V). Further, to prevent the erroneous A/V separation as result of graph search analysis, the probability of individual vessel segment (belonging to a vessel subtree) being an A/V is also

Table 5.1: List of features extracted for A/V classification.

| Index | Feature description |
|---|---|
| 1-9 | Vessel pixel intensity in Red, Green, Blue Hue, Saturation, Brightness and Lab color space. |
| 10-45 | Mean, standard deviation, minimum and maximum intensities of Red, Green, Blue, Hue, Saturation, Brightness and Lab vessel maps. |
| 46-55 | Vessel centerline pixel intensity in a Gaussian blurred ($\sigma = 2, 4, 8, 12, 16$) of Red and Green Channel. |
| 56 | Vessel width* ($\epsilon$) calculated at the centerline pixel location of the binary vessel map. |
| 57-66 | Vessel cross-sectional intensity profile† in Red and Green channel. |

* The vessel width ($\epsilon$ as defined in Equation 5.5) of a segment is measured with respect to centerline pixel as the origin.
† The vessel profile was determined using VKD with respect to every centerline pixels, within a ROI of [$M_{min} = 3$ pixels, $M_{max} = 7$ pixels].

calculated, without considering the result of graph search approach. A vessel segment is assigned a label artery (A), if the probability of being an artery $P_A \geq 0.9$ (90% of the pixels are being classified as arteries); else if the probability of being a vein $P_V \geq 0.9$, it is assigned with a label vein (V). A thorough performance analysis of A/V separation without and with segment-wise analysis is provided in the Appendix G. A subtree A/V labeling for a sample ROI is shown in Figure 5.7 (c) with color red representing arteries and blue representing veins.

## 5.3 Experiments and Results

### 5.3.1 Materials

The effectiveness of the proposed approach is validated by testing on four different publicly available datasets: AV-DRIVE (Qureshi et al. (2013)), CT-DRIVE (Dashtbozorg et al. (2014)), INSPIRE-AVR (Niemeijer et al. (2011)) and WIDE (Estrada et al. (2015)). The AV-DRIVE consists of 40 images ($565 \times 584$ pixels) derived from earlier DRIVE dataset (Staal et al. (2004)), with ground truth (GT) A/V labels marked for all vessel pixels. Since three different human graders manually classified all the vessel pixels, a majority consensus was taken to arrive at the final A/V labeling. The CT-DRIVE consists of 20 images corresponding to the DRIVE test set (Staal et al. (2004)) with A/V labels graded only for vessel centerline pixels. The INSPIRE-AVR consists of 40 images ($2392 \times 2048$ pixels) with associated A/V labels

obtained from Dashtbozorg et al. (2014) for vessel centerline pixels, only. Finally, the WIDE dataset provides 30 SLO-images ($3900 \times 3072$ pixels) along with their A/V GT labels obtained from Estrada et al. (2015).

The input binary vessel map for all four datasets was obtained automatically from the raw color image using the method proposed in Azzopardi et al. (2015). Since, CT-DRIVE, INSPIRE-AVR, and WIDE dataset consists of manual A/V markings for centerline pixels, the binary vessel map is subsequently thinned to obtain vessel centerlines (Zhang and Suen (1984)). The obtained centerline maps usually consist of various misinterpretation of vascular structures as a result of thinning operation, which was further corrected using the technique proposed in Dashtbozorg et al. (2014). This refinement aimed at resolving critical issues such as missing vessel segment, splitting of crossover into two nearby bifurcations and the deletion of a false vessel segment.

### 5.3.2 Evaluation Metrics

The performance validation was carried out using three metrics: *Sensitivity* ($Se$), *Specificity* ($Sp$) and *Accuracy* ($Acc$).

$$Se = \frac{TP}{TP + FN},\ Sp = \frac{TN}{TN + FP},\ Acc = \frac{TP + TN}{N},$$

where, $TP/FP$ denotes true/false positives; $TN/FN$ denotes true/false negatives; and the total $N = TP + TN + FN + FP$. In the proposed method, the arteries are considered as positives and veins as negatives. Accordingly, sensitivity is defined as how well the method can detect arteries, while specificity indicates how well it can detect veins.

### 5.3.3 A/V Separation Evaluation

In the following subsections, the results of the proposed A/V separation method is presented on each of the four retinal datasets. The proposed approach is validated at each stage of the pipeline starting from vessel subtree extraction to the final A/V labeling of the entire vascular network.

The method "*DFS-search*" refers to the output at vessel subtree extraction stage, where the A/V labels are manually assigned for individual vessel subtrees. The "*RF-only*" refers to the classification of vessel pixels into A/V by using hand-crafted features (shown in Table 5.1) trained with RF classifier (without considering any

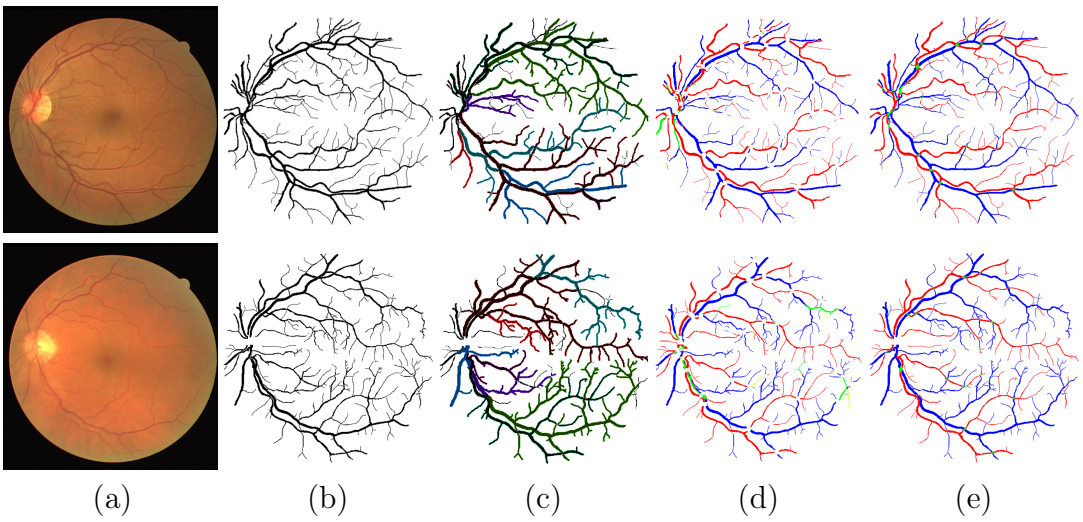



Figure 5.8: AV-DRIVE separation results. Top row: the proposed method's best result ($Acc$ = 0.985). Bottom row: the proposed method's worst result ($Acc$ = 0.889). (a) Original image; (b) Binary vessel map; (c) Identified vessel subtrees; (d) A/V separation results; (e) Corresponding A/V ground truth. (*Note*: the correctly labeled arteries/veins are shown in red/blue, respectively. While, incorrectly labeled arteries/veins are shown in yellow/green, respectively). Best viewed in color.

graph-search analysis). Finally, "*DFS-search with RF*" refers to the method's final output after subtree A/V labeling stage, which uses combined knowledge of graph-search analysis and RF classifier for predicting the final A/V label of the entire vessel tree.

#### 5.3.3.1 AV-DRIVE Dataset

The AV-DRIVE consists of pre-partitioned training and test set, each of which contains 20 images. The sample A/V separation results on DRIVE images are shown in Figure G.1. It is observed that the proposed method accurately classifies most of the A/V segments in the DRIVE set, including the region of thin and low contrast vessels. Most of the false classifications are due to the presence of nearby junction points and missing vessel connectivity, especially around the optic disc area. A similar ambiguity has also been observed with expert human graders in assigning A/V labels in and around the optic disc, which further showcase the difficulty of this A/V separation problem.

Table 5.2, shows the performance of the proposed approach with state-of-the-art methods on the AV-DRIVE dataset. The proposed method has shown a significant improvement in $Se$ of 0.966, with a relatively high $Sp$ of 0.929, clearly indicating that the system is capable of identifying arteries than the veins. This is also shown

Table 5.2: Comparative analysis of the proposed methods on AV-DRIVE dataset.

| **Method** | Accuracy | Sensitivity | Specificity | Time (sec) |
|---|---|---|---|---|
| Estrada et al. (2015) | 0.935 | 0.930 | **0.941** | 131.32 |
| Xu et al. (2017) | 0.923 | 0.915 | 0.929 | **5.00** |
| Welikala et al. (2017) | 0.912 | – | – | – |
| Hu et al. (2015) | 0.880 | – | – | – |
| Huang et al. (2018a) | 0.720 | 0.709 | 0.738 | |
| DFS-search | 0.896 | 0.918 | 0.875 | 124.70 |
| RF-only | 0.746 | 0.779 | 0.713 | 36.40 |
| **DFS-search + RF** | **0.947** | **0.966** | 0.929 | 161.10 |

Table 5.3: Comparative analysis of the proposed methods on CT-DRIVE dataset.

| **Method** | Accuracy | Sensitivity | Specificity | Time (sec) |
|---|---|---|---|---|
| Estrada et al. (2015) | 0.917 | 0.917 | **0.917** | **131.32** |
| Dashtbozorg et al. (2014) | 0.874 | 0.900 | 0.840 | – |
| Niemeijer et al. (2011) | N/A | 0.800 | 0.800 | – |
| DFS-search | 0.909 | 0.924 | 0.895 | 109.00 |
| RF-only | 0.753 | 0.781 | 0.726 | 34.00 |
| **DFS-search + RF** | **0.932** | **0.950** | 0.915 | 143.00 |

to be consistent with Figure G.1, where a majority of the wrongly classified vessel pixels belongs to the veins (shown in green).

#### 5.3.3.2 CT-DRIVE Dataset

A leave-one-out cross-validation was adopted for evaluation of 20 images of the CT-DRIVE dataset. Classification of vessel pixels which are higher than 3 pixels wide is considered for performance assessment (Dashtbozorg et al. (2014)). The sample A/V separation results on the CT-DRIVE dataset is shown in Figure G.2. The proposed method achieved a very high $Se$ of 0.950, which is $\approx$ 3% greater than previous approaches on the same dataset, as shown in Table 5.3. Similar to the AV-DRIVE dataset, much of the false positives are in the optic disc region. The thinning of closely spaced junction structures especially in the optic disc region, often resemble spurious and isolated fragments, leading to difficulty in obtaining any meaningful information about the vessel morphology. Henceforth, a marginal decrease in A/V performance is observed on CT-DRIVE dataset compared to AV-DRIVE, as shown in Table 5.3. A similar trend has also been observed with earlier methods (Dashtbozorg et al. (2014); Niemeijer et al. (2011)), that depends on centreline map as an input to their approach.

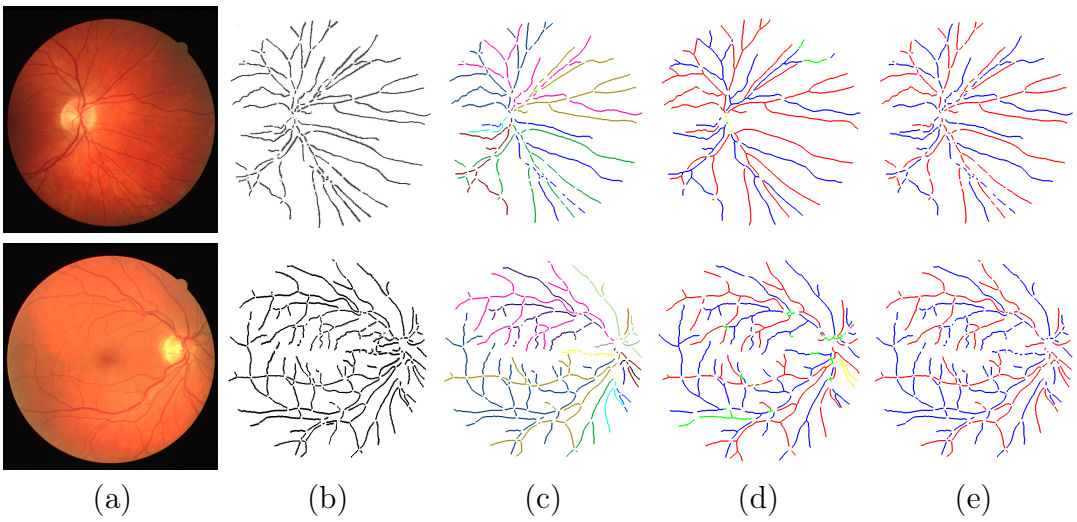



Figure 5.9: CT-DRIVE separation results. Top row: the proposed method's best result ($Acc = 0.981$). Bottom row: the proposed method's worst result ($Acc = 0.864$). (a) Original image; (b) Binary vessel map; (c) Identified vessel subtrees; (d) A/V separation results; (e) Corresponding A/V ground truth. (*Note*: the correctly labeled arteries/veins are shown in red/blue, respectively. While, incorrectly labeled arteries/veins are shown in yellow/green, respectively). Best viewed in color.

#### 5.3.3.3 INSPIRE-AVR Dataset

For the evaluation of A/V separation on INSPIRE-AVR dataset, a 2-fold cross-validation has been adopted. In this approach, the dataset is divided into two random equal sized partitions, each containing 20 images. One partition is used as train set while, the other partition as test set and vice-versa. Examples of A/V separation on INSPIRE-AVR dataset is shown in Figure G.3. Table 5.4 shows the comparative analysis of the proposed approach with state-of-the-art methods. The proposed method obtain an $Acc$ value of 0.968 and $Se$ of 0.969, which is 5% higher than the most recent results reported in Estrada et al. (2015). The INSPIRE-AVR dataset contains only fewer vessel structures and complex junction locations, which subsequently resulted in better A/V separation compared to AV/CT-DRIVE datasets.

The proposed method has been further evaluated by considering only six major A/V vessels within a region of 1-DD to 1.5-DD, as this ROI was normally used for AVR calculation (Niemeijer et al. (2011)). The proposed method obtain a $Acc$ value of 0.991, which is 1%($Acc = 0.971$)/15%($Acc = 0.84$) greater than previously reported methods (Dashtbozorg et al. (2014); Niemeijer et al. (2011)), respectively. For the sake of fair comparison, the proposed method obtain a $Se$ value of 0.981

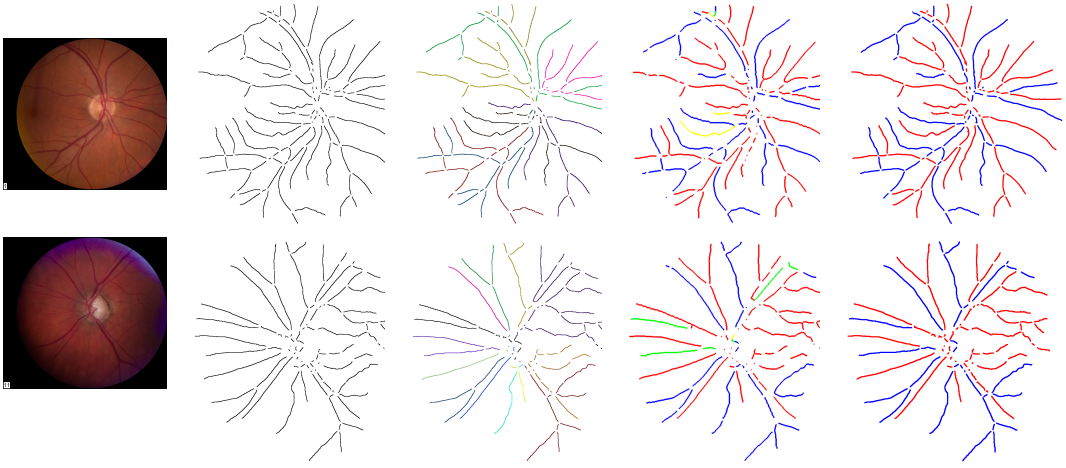



Figure 5.10: INSPIRE-AVR separation results. Top row: the proposed method's best result ($Acc = 0.989$). Bottom row: the proposed method's worst result ($Acc = 0.926$). (a) Original image; (b) Binary vessel map; (c) Identified vessel subtrees; (d) A/V separation results; (e) Corresponding A/V ground truth. (*Note*: the correctly labeled arteries/veins are shown in red/blue, respectively. While, incorrectly labeled arteries/veins are shown in yellow/green, respectively). Best viewed in color.

Table 5.4: Comparative analysis of the proposed methods on INSPIRE-AVR dataset.

| **Method** | Accuracy | Sensitivity | Specificity | Time (sec) |
|---|---|---|---|---|
| Estrada et al. (2015) | 0.909 | 0.915 | 0.902 | 117.68 |
| Dashtbozorg et al. (2014) | 0.849 | 0.910 | 0.860 | – |
| Niemeijer et al. (2011) | N/A | 0.780 | 0.780 | – |
| Huang et al. (2018b) | 0.851 | N/A | N/A | – |
| DFS-search | 0.921 | 0.945 | 0.897 | 62.13 |
| RF-only | 0.682 | 0.691 | 0.674 | 14.95 |
| **DFS-search + RF** | **0.968** | **0.969** | **0.966** | **77.08** |

from the ROC curve, for a fixed $Sp$ of 0.860 reported in Dashtbozorg et al. (2014); Niemeijer et al. (2011). Thus, the proposed approach is also shown to be reliable in developing an automated solution for quantifying AVR in retinal images.

#### 5.3.3.4 WIDE Dataset

In the WIDE dataset, a 2-fold cross-validation approach is adopted by randomly assigning images into two sets, each containing 15 images. The sample A/V results of the WIDE dataset is shown in Figure 5.11. When compared with other three datasets, A/V separation on the WIDE dataset is most challenging because of very low-contrast noisy images. These artefacts are more prominent in the peripheral regions containing thin and low contrast vessel structures as shown in Figure 5.11.

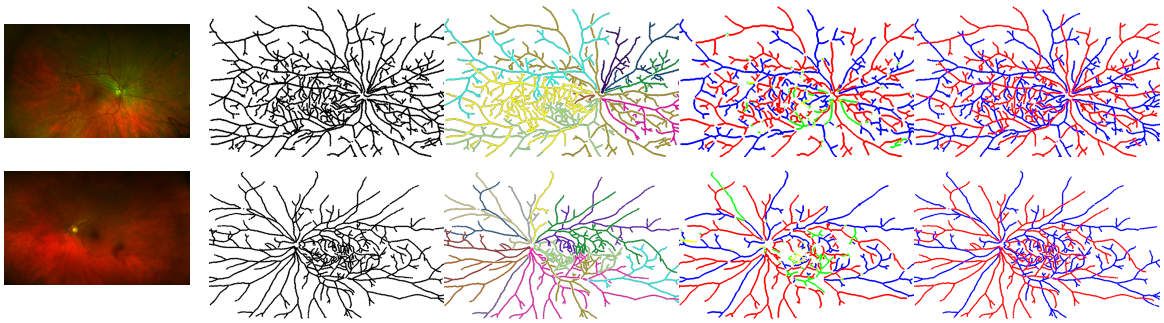



Figure 5.11: WIDE separation results. Top row: the proposed method's best result ($Acc$ = 0.939). Bottom row: the proposed method's worst result ($Acc$ = 0.827). (a) Original image; (b) Binary vessel map; (c) Identified vessel subtrees; (d) A/V separation results; (e) Corresponding A/V ground truth. (*Note*: the correctly labeled arteries/veins are shown in red/blue, respectively. While, incorrectly labeled arteries/veins are shown in yellow/green, respectively). Best viewed in color.

Table 5.5: Comparative analysis of the proposed methods on WIDE dataset.

| **Method** | Accuracy | Sensitivity | Specificity | Time (sec) |
|---|---|---|---|---|
| Estrada et al. (2015) | **0.910** | 0.910 | **0.909** | 777.35 |
| Pellegrini et al. (2018) | 0.862 | N/A | N/A | – |
| DFS-search | 0.861 | 0.891 | 0.831 | 209.40 |
| RF-only | 0.741 | 0.763 | 0.719 | 54.58 |
| **DFS-search + RF** | 0.902 | **0.923** | 0.882 | **263.98** |

Besides, the number of extracted vessel subtrees are significantly large due to the wider FOV followed by relatively lower spatial resolution, which is a typical case in SLO modality. Nevertheless, the proposed method was able to achieve a better $Acc/Se$ value of 0.902/0.923, which is comparatively higher than the most recent method proposed in Pellegrini et al. (2018), as shown in Table 5.5. Many of the false detections are mainly observed in the proximal region of multiple junction locations, thin and very low-contrast vessel areas (see Figure 5.11). The extracted vessel subtrees are also found to be suboptimal in these locations mainly because of the presence of erroneous vessel structures formed during vessel thinning operation. Furthermore, the pixel-wise features extracted from these regions were also found to be less discriminative, often leading to many false classifications.

## 5.4 Discussion

The proposed method consists of series of interlinked stages, where the performance of each stage depends on its previous stage output. Hence, to validate the robust-

ness of the approach, the performance at each stage is evaluated starting from: (i) identification of vessel subtrees - which is referred to as "DFS-search"; (ii) A/V classification using only hand-crafted features - which is referred to as "RF-only"; and (iii) finally, subtree A/V labeling stage (which combines both the knowledge of graph search and hand-crafted features) - which is referred to as "DFS-search with RF". Further, the relative importance of hand-crafted features are further investigated using different feature selection techniques, as well as various classifiers to examine the impact on final A/V labeling.

### 5.4.1 Performance Analysis of Each Stage Output

The proposed three-stage refinement steps accurately predicts the A/V labeling from four different datasets, including images from fundus, as well as SLO image modalities. In Section 5.2, it was shown how each of these steps contributes to yield a more accurate solution progressively. The experimental analysis also confirms this view, where it has been empirically shown, how each stage output improves upon the previous stage.

The A/V separation at the output of DFS-search is evaluated by manually assigning A/V labels for individual vessel subtrees. The proposed method obtain an average $Acc > 86\%$ across all four datasets (see Table 5.2 - Table 5.5), while depending solely on the knowledge of graph search. This underscores the richness of metaheuristic approach - which efficiently exploits local as well as global vessel connectivity information to precisely track all the A/V segments from a given vascular network. A highest *Acc* of 0.921 is observed on INSPIRE-AVR dataset, while the lowest *Acc* of 0.861 is on the WIDE dataset. This is because the INSPIRE-AVR dataset consists of a fewer number of graph linking structures - including the number of complex crossovers when compared to the WIDE dataset.

A substantial improvement is observed with a mean *Acc* of 15% (AV-DRIVE), 15.6% (CT-DRIVE), 23.9% (INSPIRE-AVR) and 12% (WIDE), when compared with DFS-search to the RF-only stage. The pixel level intensity-based features have shown to be vulnerable to varying image conditions such as resolution, contrast and illumination artefacts both within and across datasets. Finally, the combination of DFS-search with RF have shown a modest improvement in the *Acc* value of 5.1% (AV-DRIVE), 2.3% (CT-DRIVE), 4.7% (INSPIRE-AVR) and 4.1% (WIDE) from DFS-search to DFS-search with RF stage. This consistent improvement strongly

Table 5.6: Performance comparison (*Acc*) of different combination of classifiers and feature selection techniques on AV-DRIVE / INSPIRE-AVR datasets.

| Classifier/ Feature selection | NB | LDA | $k$-NN | SVM | RF |
|---|---|---|---|---|---|
| None | 0.909/ 0.945 | 0.921/ 0.923 | 0.875/ 0.938 | 0.886/ 0.902 | **0.947**/ **0.968** |
| SFS | 0.914/ 0.938 | 0.908/ 0.924 | 0.893/ 0.918 | 0.898/ 0.902 | 0.923/ 0.954 |
| Lasso | 0.926/ 0.906 | 0.897/ 0.892 | 0.862/ 0.873 | 0.887/ 0.884 | 0.909/ 0.927 |
| ENet | 0.897/ 0.910 | 0.885/ 0.903 | 0.871/ 0.891 | 0.869/ 0.876 | 0.916/ 0.935 |

* The acronym stands for 'SFS' - Sequential Forward Selection; 'LASSO' - Least Absolute Shrinkage and Selection Operator; 'ENet' - Elastic Net; 'NB' - Naive Bayes; 'LDA' - Linear Discriminant Analysis; '$k$-NN - $k$-Nearest Neighbor; 'SVM' - Linear Support Vector Machine; and 'RF' - Random Forest classifiers.

indicates that the system is more accurate while relying on more complex knowledge of vessel connectivity as well as pixel-level feature information for classifying A/V.

### 5.4.2 Influence of Feature Selection vs. Classifiers

In this section, the relative significance of hand-crafted features is analysed using a combination of various feature selection techniques with different classifiers. The experimental results on AV-DRIVE and INSPIRE-AVR datasets are reported in Table 5.6 with same dataset split as mentioned in Section 5.3.3.1 and Section 5.3.3.3, respectively.

To evaluate the importance of features on classification accuracy, a combination of feature selection techniques such as sequential forward selection, LASSO (Tibshirani (1996)), and ENet (Zou and Hastie (2005)) was adopted, with different classifiers such as Naive Bayes, LDA, $k$-NN, SVM, and RF. The RF classifier with no feature selection proved to be the most reliable combination with a *Acc* of 0.947/0.968 across AV-DRIVE and INSPIRE-AVR, respectively. This indicates that all the selected features given in Table 5.1 have a strong influence on final A/V labeling and have shown discriminative ability, even with varying imaging conditions such as contrast and luminosity. Further, the selected feature set captures both intensity (raw pixel intensities in RGB, HSI and Lab color space) as well as structural level information (such as vessel width) that aid in accurate separation of arteries from veins. Compared to different classifiers, RF has shown better performance due to its ability to perform both classification and feature selection implicitly.

### 5.4.3 Computation Time

The proposed automated A/V separation method was developed in MATLAB R2017 (MathWorks, Inc.), with an average computation time of 4.3 minutes per image (across all four datasets) using an Intel Core i7-5960 CPU at 3.00 GHz. The running time at each stage of the A/V separation pipeline is also reported as shown in Table 5.2 - Table 5.5. The DFS-based graph search is the most computationally intensive stage, which can be further improved by utilizing parallel processing capabilities and through more efficient implementation.

## 5.5 Summary

In this chapter, a comprehensive graph search metaheuristic approach is presented for separating arteries from veins in retinal images. The proposed method formulates the A/V separation as a graph search problem by incorporating sophisticated graph-theoretic knowledge with the domain-specific priors, to accurately identify A/V segments across the entire vascular tree. In this work, the vessel curvature criteria is explored in addition to orientation and width information, to precisely disseminate the A/V label information of highly curved crossovers, encountered during graph search process. Such limitation was not addressed previously in many state-of-the-art methods (Dashtbozorg et al. (2014); Estrada et al. (2015); Hu et al. (2015); Joshi et al. (2014); Pellegrini et al. (2018)) that utilizes graph theoretic knowledge, unlike the proposed approach. The inclusion of hybrid knowledge of both local and global vessel connectivity during graph traversal often helps in identifying anatomically meaningful vessel subtrees by searching the space of possible vascular networks. The proposed method was validated on four different challenging datasets including images from the narrow field (AV-DRIVE, CT-DRIVE, and INSPIRE-AVR datasets) and wide field-of-view (WIDE dataset) fundus photographs, with remarkable differences in resolution, quality, and acquisition protocol. Overall, there is a considerable improvement in A/V separation performance across the entire range of diversity and might enable real-time analysis in the future.

# Chapter 6

# Conclusions and General Discussion

Retinal fundus images provide vital information about the early manifestation of various diseases related to the eye, cardiovascular and neurodegenerative diseases. A comprehensive assessment of retinal vascular network provides important clinical biomarkers through the analysis of its geometrical properties, that aid in early diagnosis of various diseases such as diabetes, stroke, hypertension, arteriosclerosis and cerebral small vessel diseases. The detection and quantification of change in microvascular geometrical patterns lead to early insights into the progression of the aforementioned diseases. The major steps towards the analysis of retinal vasculature include - segmentation of vessel tree, identification and classification of vessel branching/crossover points, separation of vessel tree into arteries and veins and quantification of various geometrical properties with diseases. The most accurate and reliable quantitative analysis of retinal vascular tree requires an automated framework that can be used for various clinical and large scale pre-screening solutions. This dissertation mainly focus on pattern recognition and machine learning approaches for automatic analysis of retinal images and a multifactorial analysis framework, with a goal of automated retinal vasculature analysis and quantification in clinical settings.

In Chapter 2, a comprehensive review of the recent state-of-the-art retinal vessel segmentation methods is presented. The various complexities and challenges involved in developing robust segmentation techniques have been discussed thoroughly, including the most crucial image preprocessing steps that have not been addressed earlier in the literature. Besides, the strengths and weakness of each category of segmentation methods, with a focus to current challenges are also presented.

Further, an in-depth quantitative evaluation of state-of-the-art approaches on the individual dataset is assessed. Finally, some of the existing challenges such as segmentation in the presence of abnormalities, accurate segmentation of thin vessels and segmentation in the presence of non-uniform illumination are discussed in detail with the effect on segmentation performance.

The existing challenges in retinal vessel segmentation are addressed in Chapter 3, by proposing a novel visual attention guided unsupervised feature learning approach. The proposed method inherits the advantages of visual attention mechanism and fully utilizes the potential of unsupervised feature learning for representing most discriminative features for pixel-wise classification. This allows to explore the space of both selection mechanism and multiscale contextual information under a single framework, without the need for any handcrafted feature learning modules. The proposed approach has several appealing properties. First, the visual attention mechanism is capable of capturing the rich contextual information, which can ignore the clutter present in a local neighbourhood, by focusing only on the pixel of interest. Second, the integration of visual attention mechanism into unsupervised filter learning encourages the intra-class similarities to be small (between the vessel pixels), and emphasizes the inter-class differences to be large (between the vessel and background pixels). The effectiveness of the proposed approach is evident by the significant improvement in the segmentation of several challenging image structures such as central vessel reflex, complex crossover patterns, closely parallel and highly curved vessels, thin vessels and performs reasonably well on pathological images.

Chapter 4, presents a complete solution for the assessment of retinal vasculature mainly at vessel keypoints. This includes junction detection, bifurcation/crossover identification, vessel width estimation and AV nicking quantification. The entire pipeline is based on a proposed novel vessel keypoint descriptor. The main advantage of the proposed solution is that it does not depend on A/V classification or intensity-based features for quantifying AV nicking, in contrast to the earlier methods (Nguyen et al. (2013a); Pereira et al. (2014); Roy et al. (2014)). The proposed approach starts with a binary vessel map and looks at AV nicking quantification as a generic vessel width estimation problem, irrespective of the type of vessel. The results demonstrated to lead better quantification of vascular changes, in assessing the severity of mild to severe AV nicking cases.

Finally, in Chapter 5, a comprehensive graph search metaheuristic approach is

presented for separating arteries from veins in retinal images. The proposed method formulates the A/V separation as a graph search problem by incorporating sophisticated graph theoretical knowledge with the domain-specific priors, to accurately identify A/V segments across the entire vascular tree. The vessel keypoint descriptor proposed in Chapter 4 is extended to include the vessel curvature criteria, in addition to orientation and width information. This helps to precisely disseminate the A/V label information of not only simple vessel crossovers but also highly curved crossovers, typically encountered during vascular graph traversal. Such limitation was not addressed previously in many state-of-the-art methods (Dashtbozorg et al. (2014); Estrada et al. (2015); Hu et al. (2015); Joshi et al. (2014); Pellegrini et al. (2018)) that utilizes graph theoretic knowledge, unlike the proposed approach. The inclusion of hybrid knowledge of both local and global vessel connectivity during graph traversal often helps in identifying anatomically meaningful vessel subtrees by searching the space of possible vascular networks.

Although, the proposed vessel segmentation approach (in Chapter 3) works reasonably well, while, there are some misclassifications in the region of bright and large exudates. These false positives can further be removed by employing a simple pre/post-processing steps (such as image inpainting) or by integrating vessel shape prior information into the filter learning approach (Annunziata et al. (2016); Zhang et al. (2017)). The segmentation of pathological images can be even further improved by incorporating more hybrid contextual models such as auto-context technique (Tu and Bai (2010)). The auto-context iteratively utilizes the posterior distribution of labels along with the image features, to obtain a more compact and discriminative features suitable for highly overlapping classes (such as thin vessels and lesions). The discriminative capability of the unsupervised feature learning approach can be further improved by associating each class label information with the dictionary atom during the filtering process. This hybrid information effectively reduces both the reconstruction and classification error using a unified objective function.

As a future work, the proposed framework can be used an initial step in identifying various clinical biomarkers (as discussed in Chapter 1) that are vital in predicting retinal diseases at an early stage. Accurately detecting and quantifying these biomarkers requires a large datasets, which is free from noise and doesn't consist of large intra and inter-observer variability for the ground truth annotations. Further, acquiring a rich set of diverse clinical metadata (such as age, gender, life style fac-

tors, blood pressure and medical information such as diabetes, stroke, heart diseases) will further facilitate more accurate analysis and helps in early detection of various systemic and neuro-degenerative diseases.

With the advancements in deep learning techniques, several breakthrough methods have been proposed for segmenting retinal vessels (Fu et al. (2016a); Liskowski and Krawiec (2016)) reaching performance close to human observers. The success of these methods is strongly dependent on a large number of the labeled training set. Acquiring such a set of clinical annotations is often a tedious and expensive task, especially in the field of medical image analysis. To overcome this limitation, methods (Costa et al. (2018); Zhao et al. (2018)) based on Generative Adversarial Nets (GAN's) have been proposed for synthesizing realistic-looking retinal images, requiring a minimal set of training samples. The method proposed in Chapter 5 can be further extended using a graph convolutional network (GCN) (Defferrard et al. (2016)) by modelling the similar vessel connectivity information (discussed in Chapter 5, Section 5.2.2) under CNN framework, to obtain a more accurate A/V labelling.

In summary, this thesis paves the way towards the development of automated retinal image analysis tool based on pattern recognition and machine learning framework. The extensive validations illustrate that the proposed scheme is beneficial for the comprehensive analysis and quantification of retinal vasculature and can be reliably used to assist clinicians in early pre-screening of retinal microvascular abnormalities.

# Appendices

## A Performance Analysis of Proposed Retinal Vessel Segmentation vs. Different Classifiers

Table A.1 shows the performance analysis of different classifiers on retinal vessel segmentation approach proposed in Chapter 3. It is observed that the RF classifier performs well in terms of $Sp, Acc, AUC$, while still maintaining a very high $Se$ value. But this is not the case with other classifiers. RF classifier is better than other classifiers, due to its capability of performing both classification and feature selection implicitly. It is robust against overfitting, outliers and high dimensional imbalanced data. This is true in the proposed case, where only 9 to 14% of the total pixels belongs to vessels while, the rest belongs to non-vessels, leading to a highly skewed dataset.

## B Role of Integrating Curvature in Separating A/V at Vessel Crossover Locations

A set of experiments has been performed to detail the significance of curvature information for separation of A/V segments of different simple and challenging crossover cases, as shown in Figure B.1. In this figure, the first four rows (1 - 4) correspond to simple crossovers, and the last four rows (5 - 8) correspond to challenging crossover cases.

Among the simple vessel crossover case (*row*: 1 - 4), the identification of artery-vein segments is relatively easier, since the orientation and width condition criteria (defined in Equation 5.15 and Equation 5.16 of the Chapter 5) satisfies in most of these cases. Whereas in challenging crossover case (*row*: 5 - 8), the A/V segments are still being precisely separated, despite their deviations from the normal angular separation between corresponding vessel pairs ($\pm 180°$), as well as the vessel width.

Table A.1: Comparison of different classifiers on retinal vessel segmentation performance. Note: the acronym 'NB' - Naive Bayes (Friedman et al. (2001)); $k$-'nn' - $k$ - Nearest Neighbor (Friedman et al. (2001)); 'LDA' - Linear Discriminant Analysis (Friedman et al. (2001)); 'AB' - AdaBoost with 100 trees (Breiman (2001)); 'SVM' - Support Vector Machine (Vapnik (2013)); and 'RF' - Random Forest classifiers (Breiman (2001)).

| Datasets/Classifiers | | NB | LDA | $k$-nn | SVM | AB | RF |
|---|---|---|---|---|---|---|---|
| DRIVE | *Se* | **0.9070** | 0.8565 | 0.7075 | 0.8324 | 0.6666 | 0.8644 |
| | *Sp* | 0.4614 | 0.8826 | 0.8101 | 0.8987 | 0.8810 | **0.9667** |
| | *Acc* | 0.4990 | 0.8804 | 0.7865 | 0.8603 | 0.8629 | **0.9589** |
| | *AUC* | 0.6842 | 0.8696 | 0.7588 | 0.8655 | 0.7738 | **0.9701** |
| STARE | *Se* | **0.8727** | 0.7522 | 0.6291 | 0.7504 | 0.6094 | 0.8325 |
| | *Sp* | 0.4152 | 0.8805 | 0.7101 | 0.8710 | 0.8734 | **0.9746** |
| | *Acc* | 0.4650 | 0.8666 | 0.6978 | 0.8547 | 0.8447 | **0.9502** |
| | *AUC* | 0.6439 | 0.8164 | 0.6696 | 0.8107 | 0.7414 | **0.9670** |
| CHASE_DB1 | *Se* | **0.8860** | 0.8438 | 0.7299 | 0.8447 | 0.6345 | 0.8297 |
| | *Sp* | 0.7600 | 0.8596 | 0.8156 | 0.9014 | 0.8273 | **0.9663** |
| | *Acc* | 0.7688 | 0.8585 | 0.7985 | 0.8813 | 0.8138 | **0.9474** |
| | *AUC* | 0.8230 | 0.8517 | 0.7727 | 0.8730 | 0.7309 | **0.9591** |
| IOSTAR | *Se* | **0.8767** | 0.8483 | 0.7587 | 0.8147 | 0.7012 | 0.8269 |
| | *Sp* | 0.5401 | 0.8591 | 0.8315 | 0.8902 | 0.7928 | **0.9669** |
| | *Acc* | 0.5653 | 0.8583 | 0.8042 | 0.8619 | 0.7859 | **0.9564** |
| | *AUC* | 0.7084 | 0.8537 | 0.7814 | 0.8524 | 0.7470 | **0.9663** |
| RC-SLO | *Se* | 0.8677 | 0.8237 | 0.7325 | **0.8878** | 0.7570 | 0.8488 |
| | *Sp* | 0.5968 | 0.8872 | 0.8547 | 0.9045 | 0.8235 | **0.9666** |
| | *Acc* | 0.6162 | 0.8826 | 0.8635 | 0.9101 | 0.8187 | **0.9581** |
| | *AUC* | 0.7322 | 0.8554 | 0.7936 | 0.8961 | 0.7902 | **0.9678** |

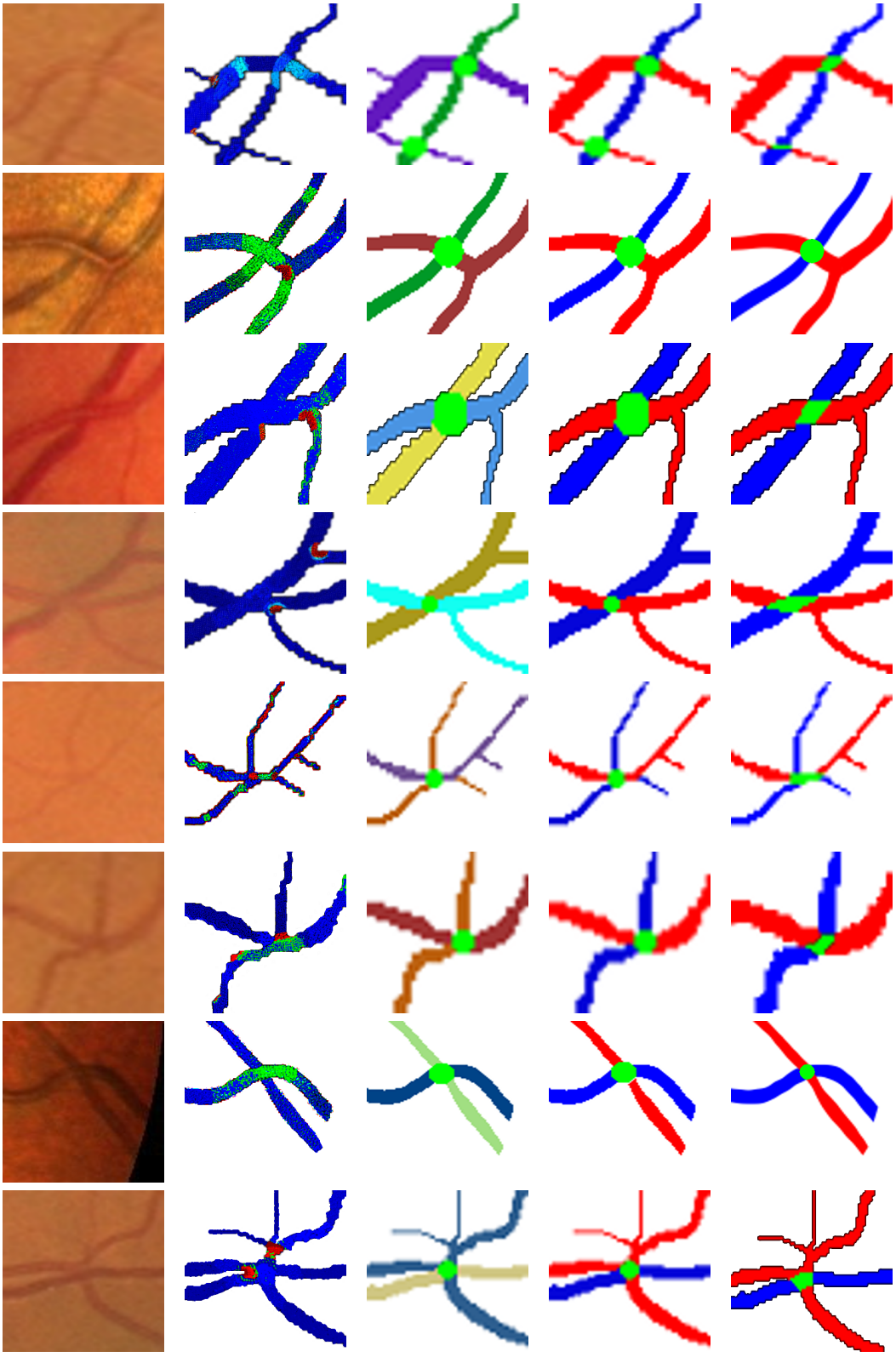



Figure B.1: Sample retinal image patches in different categories. (a) Original image; (b) curvature map; (c) the corresponding vessel subtrees; (d) A/V separation result; (e) A/V vessel groundtruth. The color of the curvature maps are scaled between the maximum and minimum values of the curvature ($\kappa$) in each image patch.

This is because, at bifurcation points, the vessel curvature is high due to the sudden change of orientation (branching vessels) compared to crossover points; where the orientation of each crossing vessel segment changes only slightly in most cases (as observed in *row*: 5 - 8). Thus, the proposed approach leverage the vessel curvature information, along with the orientation and width criteria, to precisely disseminate the A/V labels of not only simple crossovers, but also highly curved complex vessel crossovers.

# C Performance Analysis of Node Classification

The performance of node classification is quantified across DRIVE, INSPIRE-AVR and WIDE datasets in Table C.2, Table C.3 and Table C.4, respectively. Table C.1 shows the presence of a number of vessel keypoints in training and test set of each dataset, respectively. Since, the goal is to classify the identified vessel keypoints into one of the three categories: bifurcation node ($\mathcal{V}_b$) / crossover node ($\mathcal{V}_c$) / vessel end node ($\mathcal{V}_e$). Henceforth, it is essentially a multi-class classification problem, wherein, the confusion matrix is reported along with their performance measures for all the three datasets, individually.

Table C.1: Dataset split used in the experiments. $\mathcal{V}$ denotes the number of vessel nodes. Subscripts denote the type: bifurcations ($b$), crossovers ($c$) and end points ($e$).

| **Dataset** | **Images** | **Training** | | | **Testing** | | |
|---|---|---|---|---|---|---|---|
| | | $\mathcal{V}_b$ | $\mathcal{V}_c$ | $\mathcal{V}_e$ | $\mathcal{V}_b$ | $\mathcal{V}_c$ | $\mathcal{V}_e$ |
| **DRIVE** | 40 | 1923 | 761 | 3933 | 1513 | 551 | 2966 |
| **INSPIRE-AVR** | 40 | 430 | 140 | 1041 | 359 | 157 | 897 |
| **WIDE** | 30 | 1734 | 677 | 3612 | 2217 | 904 | 4618 |
| **Total** | **110** | **4087** | **1578** | **8586** | **4089** | **1612** | **8481** |

For an individual class $C_i$ (bifurcation/crossover/vessel end node), the assessment is defined by $TP_i$; $FP_i$; $TN_i$; $FN_i : Acc_i$; $Pr_i$; $Re_i$ *and* $F1 - score_i$ are calculated from the counts for $C_i$. Quality of the overall classification is usually assessed in two ways: macro-averaging and micro-averaging. The macro-averaging is considered, since it treats all classes equally (which is average of the same measures calculated for $C_1, ..., C_l$). The following are the performance measures adopted for the evaluation of multi-class node classification: *Precision* ($Pr_M$), *Recall* ($Re_M$), Accuracy ($Acc$) and *F1-score* ($F1 - score_M$).

Table C.2: Confusion matrix and vessel keypoints detection performance on **DRIVE** dataset. *Note:* $\mathcal{V}_b$ - bifurcation nodes, $\mathcal{V}_c$ - crossover nodes, $\mathcal{V}_e$ - vessel end nodes, $Pr$ - precision, $Re$ - recall, $Acc$ - accuracy and subscript $(M)$ denotes macro averaging.

| | | **Predicted** | | | |
|---|---|---|---|---|---|
| | **Nodes** | $\mathcal{V}_b$ | $\mathcal{V}_c$ | $\mathcal{V}_e$ | **Total** |
| **Actual** | $\mathcal{V}_b$ | 1347 | 166 | 0 | 1513 |
| | $\mathcal{V}_c$ | 85 | 466 | 0 | 551 |
| | $\mathcal{V}_e$ | 59 | 0 | 2907 | 2966 |
| | **Correctly predicted** | **1347** | **466** | **2907** | **4720/5030** |
| **Performance metric** | | $\mathcal{V}_b$ | $\mathcal{V}_c$ | $\mathcal{V}_e$ | **Average** |
| $Pr_{(M)}$ | | 0.9034 | 0.7370 | 1.00 | **0.8801** |
| $Re_{(M)}$ | | 0.8902 | 0.8457 | 0.9800 | **0.9053** |
| $Acc$ | | 0.9383 | 0.9500 | 0.9882 | **0.9588** |
| $F1-score_{(M)}$ | | 0.8967 | 0.7876 | 0.9890 | **0.8911** |

Table C.3: Confusion matrix and vessel keypoints detection performance on **INSPIRE-AVR** dataset. *Note:* $\mathcal{V}_b$ - bifurcation nodes, $\mathcal{V}_c$ - crossover nodes, $\mathcal{V}_e$ - vessel end nodes, $Pr$ - precision, $Re$ - recall, $Acc$ - accuracy and subscript $(M)$ denotes macro averaging.

| | | **Predicted** | | | |
|---|---|---|---|---|---|
| | **Nodes** | $\mathcal{V}_b$ | $\mathcal{V}_c$ | $\mathcal{V}_e$ | **Total** |
| **Actual** | $\mathcal{V}_b$ | 342 | 17 | 0 | 359 |
| | $\mathcal{V}_c$ | 4 | 153 | 0 | 157 |
| | $\mathcal{V}_e$ | 8 | 0 | 889 | 897 |
| | **Correctly predicted** | **342** | **153** | **889** | **1384/1413** |
| **Performance metric** | | $\mathcal{V}_b$ | $\mathcal{V}_c$ | $\mathcal{V}_e$ | **Average** |
| $Pr_{(M)}$ | | 0.9661 | 0.9000 | 1.00 | **0.9553** |
| $Re_{(M)}$ | | 0.9526 | 0.9745 | 0.9910 | **0.9727** |
| $Acc$ | | 0.9526 | 0.9745 | 0.9910 | **0.9727** |
| $F1-score_{(M)}$ | | 0.9593 | 0.9357 | 0.9954 | **0.9634** |

Table C.4: Confusion matrix and vessel keypoints detection performance on **WIDE** dataset. *Note:* $\mathcal{V}_b$ - bifurcation nodes, $\mathcal{V}_c$ - crossover nodes, $\mathcal{V}_e$ - vessel end nodes, $Pr$ - precision, $Re$ - recall, $Acc$ - accuracy and subscript $(M)$ denotes macro averaging.

| | | **Predicted** | | | |
|---|---|---|---|---|---|
| | **Nodes** | $\mathcal{V}_b$ | $\mathcal{V}_c$ | $\mathcal{V}_e$ | **Total** |
| **Actual** | $\mathcal{V}_b$ | 1936 | 281 | 0 | 2217 |
| | $\mathcal{V}_c$ | 23 | 881 | 0 | 904 |
| | $\mathcal{V}_e$ | 312 | 0 | 4306 | 4618 |
| **Correctly predicted** | | **1936** | **881** | **4306** | **7123/7739** |

| **Performance metric** | $\mathcal{V}_b$ | $\mathcal{V}_c$ | $\mathcal{V}_e$ | **Average** |
|---|---|---|---|---|
| $Pr_{(M)}$ | 0.8524 | 0.7581 | 1.00 | **0.8701** |
| $Re_{(M)}$ | 0.8732 | 0.9745 | 0.9324 | **0.9267** |
| $Acc$ | 0.9204 | 0.9600 | 0.9596 | **0.9466** |
| $F1-score_{(M)}$ | 0.8626 | 0.8527 | 0.9650 | **0.8934** |

$$Pr_{(M)} = \frac{\sum_{i=1}^{l} \frac{TP_i}{TP_i+FP_i}}{l},\ Re_{(M)} = \frac{\sum_{i=1}^{l} \frac{TP_i}{TP_i+FN_i}}{l},\ Acc = \frac{\sum_{i=1}^{l} \frac{TP_i+TN_i}{TP_i+FN_i+FP_i+TN_i}}{l},$$

$$F1-score_{(M)} = \frac{\sum_{i=1}^{l} \frac{2\times Pr_{(M)}\times Re_{(M)}}{Pr_{(M)}+Re_{(M)}}}{l}.$$

Where, $TP_i/FP_i$ denotes true/false positives for individual class $C_i$; $TN_i/FN_i$ denotes true/false negatives for individual class $C_i$; $l$ denotes the total number of class labels (in proposed case $l = 3$); and $M$ denotes macro-averaging.

# D Graph Modifications

The following are the typical errors arises during node classification that affects the optimal traversal of DFS based graph search algorithm:

**1) Bifurcation ($\mathcal{V}_b$) node is wrongly classified as crossover ($\mathcal{V}_c$) node.**

This occurs when multiple bifurcation nodes exist close to each other, or numerous bifurcations and crossover nodes coexist within a small region. Figure D.2 illustrates this phenomenon, where a sample region-of-interest (ROI) in top row consists of an undetected bifurcation point; and the bottom row containing a misclassified bifurcation point - due to the presence of multiple nearby bifurcations and crossover points.

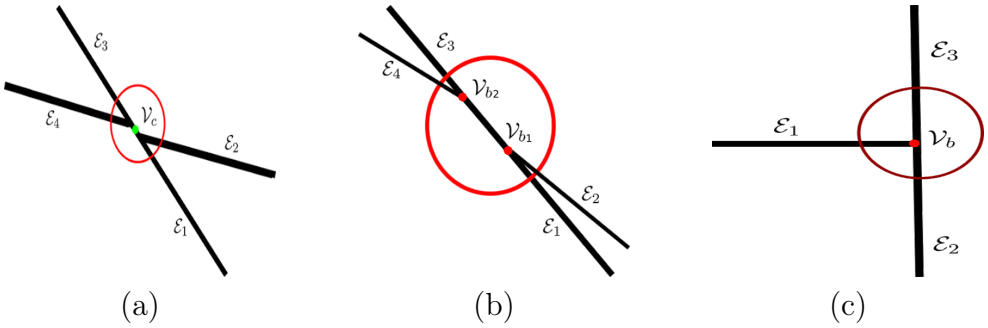



Figure D.1: Typical errors during node classification. (a) Two nearby bifurcations wrongly classified as a crossover ($\mathcal{V}_c$); (b) crossover wrongly classified as two close bifurcations ($\mathcal{V}_b$); (c) vessel end node wrongly identified as bifurcation ($\mathcal{V}_b$).

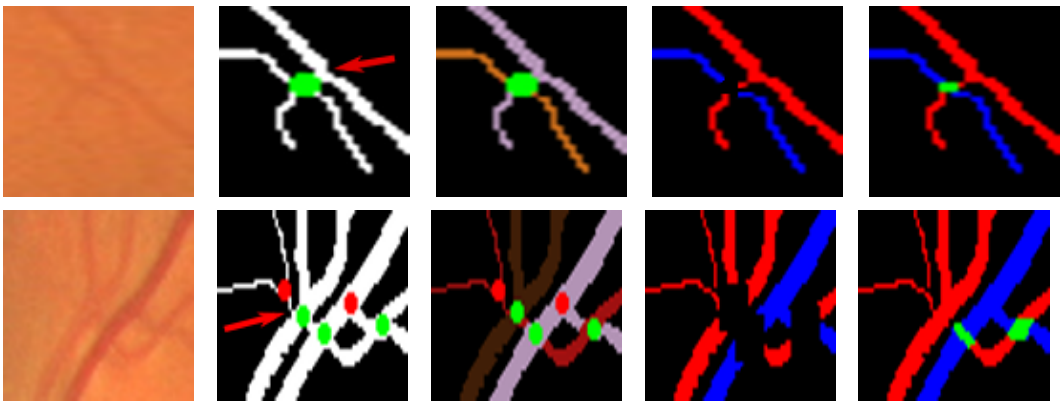



Figure D.2: A sample region-of-interest illustrating bifurcation and crossover nodes close to each other. (a) Original image patch; (b) vessel map with detected nodes (the red circles indicates bifurcation and green circles indicates crossover point); (c) identified vessel subtrees; (d) A/V separation results; (e) corresponding A/V groundtruth. **Top row:** the red arrow (see column (b)) indicates undetected bifurcation point; **Bottom row:** the red arrow (see column (b)) indicates misclassified bifurcation point.

This situation is illustrated and explained in more detail in Figure D.1 (a). Accordingly, the segments $(\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \mathcal{E}_4)$ must belong to the same vessel type whereas, due to misclassification $(\mathcal{E}_1, \mathcal{E}_3)$ appears to be crossing $(\mathcal{E}_2, \mathcal{E}_4)$ vessels. During DFS based graph traversal, due to crossover condition the segments $(\mathcal{E}_1, \mathcal{E}_3)$ will be given label $l_1$ and the segments $(\mathcal{E}_2, \mathcal{E}_4)$ will be given different label $l_2$. Although, DFS stage projects it as two different labels $(l_1, l_2)$, the final A/V labelling happens only at subtree A/V labelling stage depending on features extracted from the pixels across two crossing segments.

Figure D.2 (c) (bottom row) depicts a similar trend, where the three identified vessel subtrees have been predicted with correct labels after final A/V labelling of vessel subtrees (see Figure D.2 (d) (bottom row)). In the experimental analysis, it is found that these kind of misclassifications often doesn't induce any false classifications in final A/V results, due to the combined power of both graph search and pixel-wise processing capabilities.

**2) Bifurcation ($\mathcal{V}_b$) node is wrongly classified as vessel end node ($\mathcal{V}_e$).**

Experimentally, such misclassifications are not found during the vascular graph traversal.

**3) Crossover ($\mathcal{V}_c$) node is wrongly classified as bifurcation node ($\mathcal{V}_b$).**

Figure D.3 illustrates this situation, where a single crossover node is being classified as multiple close bifurcation nodes. This occurs when two close vessels crossover each other with a very low angle. This problem has been addressed by evaluating the two following conditions (see Figure D.1 (b)):

- for every bifurcation node $\mathcal{V}_b$, find its associated neighbours. If the associated neighbour is also a bifurcation node, then check whether the *distance* between two nodes is less than certain *threshold* $(th)$ as

$$dist(\mathcal{V}_{b1}, \mathcal{V}_{b2}) \leq th, \tag{D.1}$$

  where, $th$ is chosen empirically as 15 pixels.

- if the associated neighbour is a bifurcation node, then also check the vessel width $(\epsilon)$ of two bifurcating segments $(\mathcal{E}_1, \mathcal{E}_2/\mathcal{E}_3, \mathcal{E}_4)$ at $\mathcal{V}_{b1}, \mathcal{V}_{b2}$ matches with each other

$$\epsilon\{\mathcal{E}_1\} \cong \epsilon\{\mathcal{E}_3\} \vee \epsilon\{\mathcal{E}_2\} \cong \epsilon\{\mathcal{E}_4\}. \tag{D.2}$$

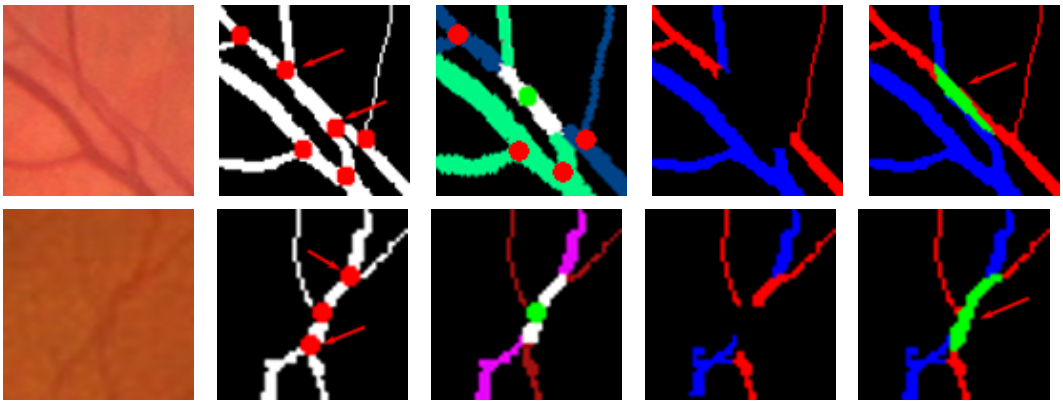



Figure D.3: A sample region-of-interest illustrating misclassified crossover as false bifurcation nodes. (a) Original image patch; (b) vessel map with detected nodes (the red circles indicates bifurcation and green circles indicates crossover point); (c) identified vessel subtrees (with corrected crossover node shown in green); (d) A/V separation results; (e) corresponding A/V groundtruth. **Top row:** the red arrow (see column (b)) indicates crossover being detected as two close bifurcation nodes; **Bottom row:** the red arrow (see column (b)) crossover being detected as three close bifurcation nodes.

If all of the above two conditions are satisfied, then the two bifurcation nodes are merged as a single crossover node by finding the midpoint of $\mathcal{V}_{b1}, \mathcal{V}_{b2}$.

**4) Crossover ($\mathcal{V}_c$) node is wrongly classified as vessel end node ($\mathcal{V}_e$).**

Experimentally, such misclassifications are not found during the vascular graph traversal.

**5) Vessel end node ($\mathcal{V}_e$) node is wrongly classified as bifurcation node ($\mathcal{V}_b$).**

Figure D.4 shows the challenging situation corresponding to the false detection of bifurcation node. The false bifurcation node is often created when two vessels are very close to each other, but they don't crossover. This problem has been solved by simply finding the angle between two vessel segments as illustrated in Figure D.1 (c). If orientation of $\angle\mathcal{E}_1\mathcal{E}_3 = 90° \pm 10° \ \wedge \ \angle\mathcal{E}_1\mathcal{E}_2 = 90° \pm 10°$; then the detected bifurcation are declared as false node and hence, this error is corrected by replacing it with a vessel end node during graph search stage.

**6) Vessel end node ($\mathcal{V}_e$) node is wrongly classified as crossover node ($\mathcal{V}_c$).**

Experimentally, such misclassifications are not found during the vascular graph traversal.

**7) Failure cases.**

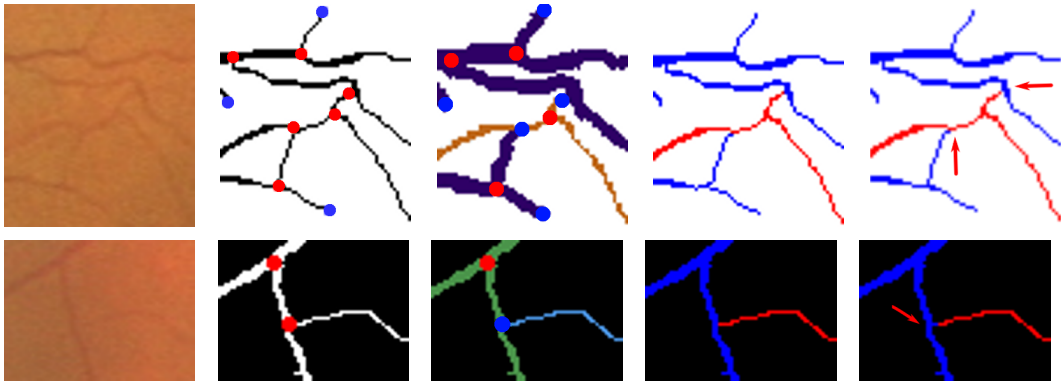



Figure D.4: A sample region-of-interest illustrating false bifurcation nodes. (a) Original image patch; (b) vessel map with detected nodes (the red circles indicates bifurcation and blue circles indicates vessel end point); (c) identified vessel subtrees (with corrected vessel end nodes (blue circle)); (d) A/V separation results; (e) corresponding A/V groundtruth. **Column (e)** the red arrow indicates creation of false bifurcation links.

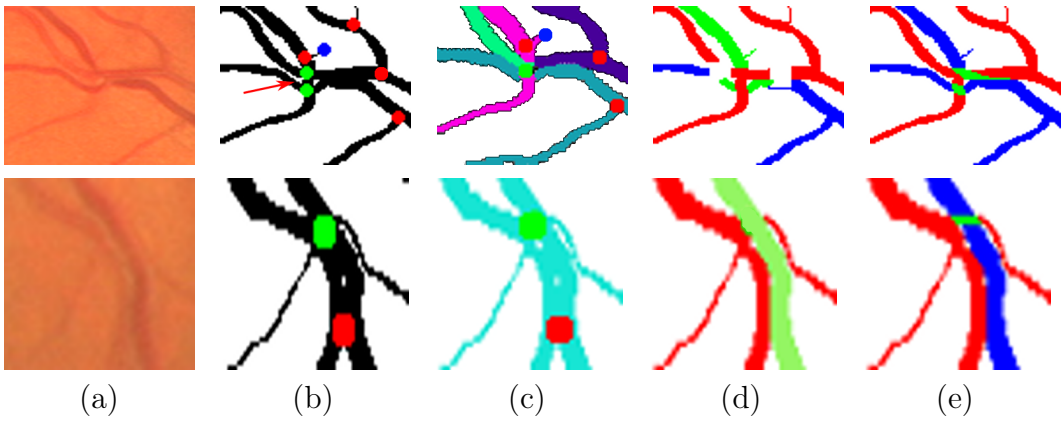



Figure D.5: A sample region-of-interest illustrating complex failure cases. (a) Original image patch; (b) vessel map with detected nodes (the red circles indicates bifurcation, green circles indicates crossover and blue circles indicates vessel end point); (c) identified vessel subtrees; (d) A/V separation results (green segment corresponds to misclassified veins); (e) corresponding A/V groundtruth.

Although, the proposed DFS based graph search method works reasonably well for most of the above mentioned challenging cases, but there exist some misclassifications which showcase the limitations of the method.

The limitations of the DFS based graph traversal are as follows:

- Presence of nearby multiple crossover points often leads to ambiguity in finding the appropriate vessel pairs $(\mathcal{E}_i, \mathcal{E}_{ip})$ for continuation of graph traversal, as shown in Figure D.5 (top row).

  The misidentified vessel subtrees (as shown in Figure D.5 (c) (top row)) are as follows:

  (i) The segments shown in green and purple (top row of column (c)) appears to be two different vessel subtrees, although belongs to the same segment.

  – This occurs because of failure in appropriately identifying corresponding A/V segments, which are based on orientation, width and curvature properties (defined in Equation 5.15, Equation 5.16 and Equation 5.17 of the Chapter 5) of crossing vessel segments. If none of these properties is satisfied, the DFS search is terminated and looks for other unvisited nodes for graph traversal. These two color (green and purple) indicates that it has been traced by two different search paths.

  (ii) the two segments (shown in pink) appears to be single vessel subtree, although belongs to two different vessels.

  – This happens when two segments satisfy the crossover conditions based on orientation, width and curvature characteristics of crossing segments.

- Figure D.5 (d) (bottom row) indicates the misclassified segment due to the merging of two different vessel types (A/V). Such misclassifications are difficult to solve with the current graph search method. This limitation has been partially addressed during subtree A/V labeling stage (as explained in Section 5.2.4) where, the A/V of each individual vessel segments (belonging to a vessel subtree) are further labelled without considering the result of graph search analysis.

# E Vessel Width Calculation

The vessel width ($\epsilon$) of a segment is measured using VKD (Srinidhi et al. (2017b)) for every centerline pixels of the binary vessel map, as illustrated in Figure E.1.

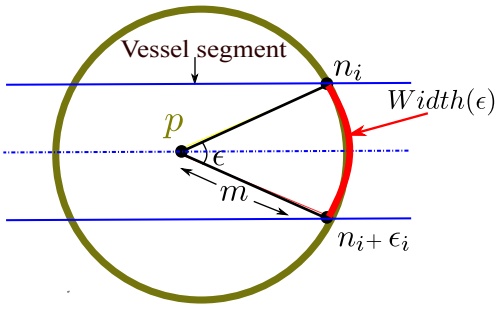



Figure E.1: Vessel width estimation for a sample vessel segment. $p$ is a point on the centerline pixel location $(x, y)$ of the binary vessel map. $\epsilon$ is the angular span between vessel intersection points $(n_i, n_i + \epsilon_i)$, which is referred to as *vessel width*; and m is the radial distance.

Consider a point $p(x, y)$ on a centerline map to which a log-polar transform is applied to obtain $p(m, n)$. Where, $m$ and $n$ are the radial and angular indices, respectively. The value $'m'$ is empirically chosen as 8 pixels, so that it covers the width of the widest vessels in all four datasets (AV-DRIVE, CT-DRIVE, INSPIRE-AVR, and WIDE), and must be fixed only once during training. Further, the angle index $n$ is obtained by sampling the angular variable at every $1^\circ$ interval, and hence $n \in [0^\circ, 360^\circ]$.

The angular span of the vessel segment at a radial distance $m$ is denoted as the *vessel width* $(\epsilon)$, with the end points of vessel $(n_i, n_i + \epsilon_i)$ subtending an angle $\epsilon$ at point $p$, as shown in Figure E.1. The sample vessel width measurements for artery-vein segments are shown in Figure E.2 (d). Note that, the width of the vein segment is relatively larger than an artery, which is an important feature in distinguishing artery-vein segments. Also note that in Figure E.2 (d), the y-axis corresponds to the normalised width values, obtained by dividing with its mean width corresponding to the entire vessel length.

# F Vessel Profile Estimation

The vessel profile was determined using VKD (Srinidhi et al. (2017b)) for every centerline pixels in red and green channel image. This is briefly described as follows: first an ROI $R_p(x, y)$ is extracted for every centerline pixels $p(x, y)$ in the red and green channel map. A log-polar transform is applied to the ROI to obtain $R_p(m, n)$ as illustrated in Figure F.1. Where, $m \in [M_{min}, M_{max}]$ is the radial index; and $n$ is the angle index, which is obtained by sampling the angular variable at $1^\circ$ interval, and hence $n \in [0^\circ, 360^\circ]$.

Next, a vertical projection of $R_p(m, n)$ is computed which results in a vector

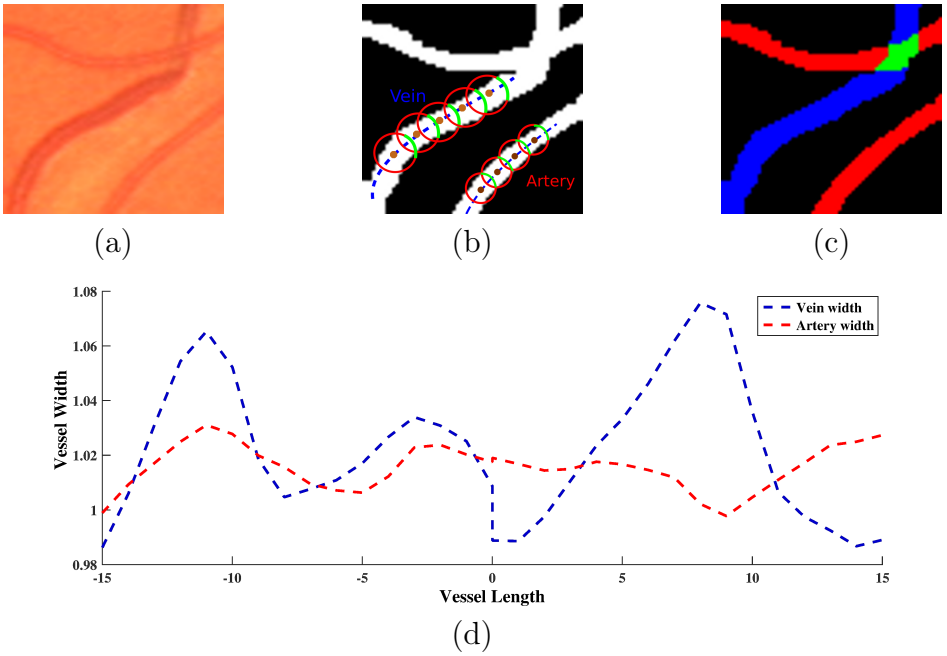



Figure E.2: A sample vessel width estimation for artery-vein segments. (a) Original image patch; (b) the corresponding binary vessel map illustrating the width estimation for a sample artery and vein segments; (c) the corresponding A/V ground truth; (d) artery and vein vessel width (shown in red and blue, respectively) calculated for a sample of 30 centerline pixels, as shown in (b); *Note:* the y-axis corresponds to the normalised width values, obtained by dividing with its mean width corresponding to the entire vessel length.

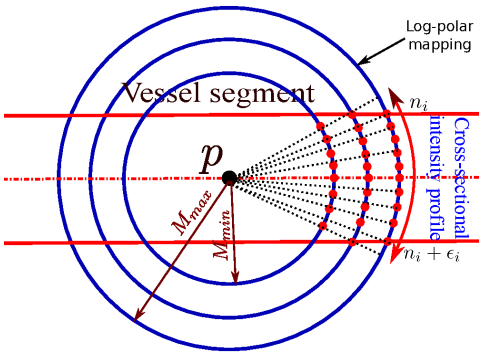



Figure F.1: Vessel cross-sectional intensity profile for a sample vessel segment. $p$ is a point on the centerline pixel location $(x, y)$ of the red/green channel image. $M_{min}$ and $M_{max}$ are the inner and outer most radii, respectively. The angular orientations: $n \in [n_i,\ n_i + \epsilon_i]$ (*red dots*), along with which the vessel intensity profile is estimated for artery-vein segments.

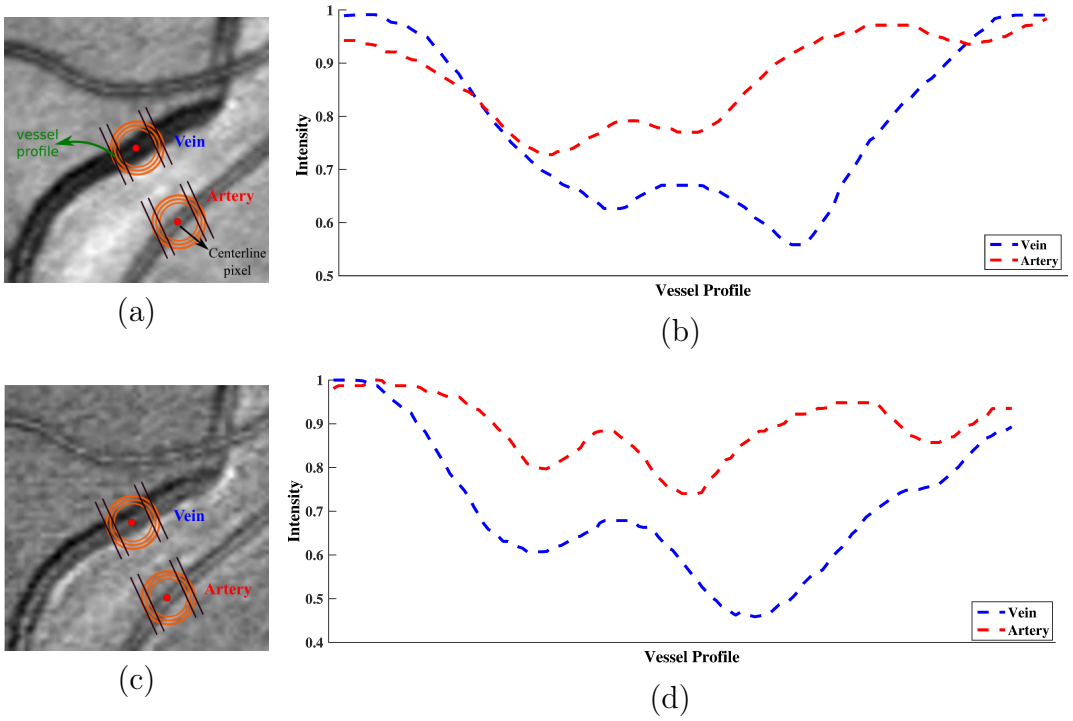



Figure F.2: (a) A sample retinal image patch (shown in *green* channel), with its corresponding vessel intensity profiles on the artery and vein are shown in (b); (c) a sample retinal image patch (shown in *red* channel), with its corresponding vessel intensity profiles on the artery and vein are shown in (d). *Note:* the intensity values in y-axis are normalized between 0 and 1; x-axis corresponds to vessel profile in the angular range $(n_i, n_i + \epsilon_i)$ sub-sampled at 10 equally spaced angular locations.

$R_p(n)$. The $R_p(n)$ provides a count of the number of pixels in a vessel fragment at specific angle $n$. Further, the ROI $R_p(n)$ is limited to a range $n \in [n_i, n_i + \epsilon_i]$ to estimate the cross-sectional intensity profile as shown in Figure F.1 (the red dots). A mean vessel intensity value is computed at every specific angle $n$ (along the projection), which is then sub-sampled at 10 equally spaced angular locations between $(n_i, n_i + \epsilon_i)$ to obtain the final cross-sectional intensity feature $(I_p(n))$. The value $M_{min} = 3$ pixels and $M_{max} = 7$ pixels are chosen experimentally, according to the vessel widths (both wide and small vessels) of all four retinal datasets (AV-DRIVE, CT-DRIVE, INSPIRE-AVR, and WIDE). Figure F.2 (b) and Figure F.2 (d) shows the sample cross-sectional vessel intensity profile in green and red channel image, respectively.

# G Performance Analysis of A/V Separation Without and With Segment-Wise Analysis During Subtree A/V Labeling Stage

Table G.1: Performance analysis of proposed A/V separation without and with segment-wise analysis during subtree A/V labeling stage.

| Dataset | Method | Accuracy | Sensitivity | Specificity |
|---|---|---|---|---|
| AV-DRIVE | Without segment-wise analysis | 0.922 | 0.947 | 0.897 |
| | Segment-wise analysis | **0.947** | **0.966** | **0.929** |
| CT-DRIVE | Without segment-wise analysis | 0.918 | 0.931 | 0.906 |
| | Segment-wise analysis | **0.932** | **0.950** | **0.915** |
| INSPIRE-AVR | Without segment-wise analysis | 0.951 | 0.959 | 0.943 |
| | Segment-wise analysis | **0.968** | **0.969** | **0.966** |
| WIDE | Without segment-wise analysis | 0.860 | 0.871 | 0.849 |
| | Segment-wise analysis | **0.902** | **0.923** | **0.882** |

The performance of proposed A/V separation are evaluated on the two following cases:

1. Without considering the individual vessel segment-wise analysis (belonging to a vessel subtree) to label A/V again.

2. Considering individual vessel segment-wise analysis to obtain final A/V labels.

Figure G.1 (d, e), Figure G.2 (d, e), Figure G.3 (d, e) and Figure G.4 (d, e) shows the sample A/V separation results, without and with segment-wise analysis on AV-DRIVE, CT-DRIVE, INSPIRE-AVR and WIDE datasets, respectively. Table

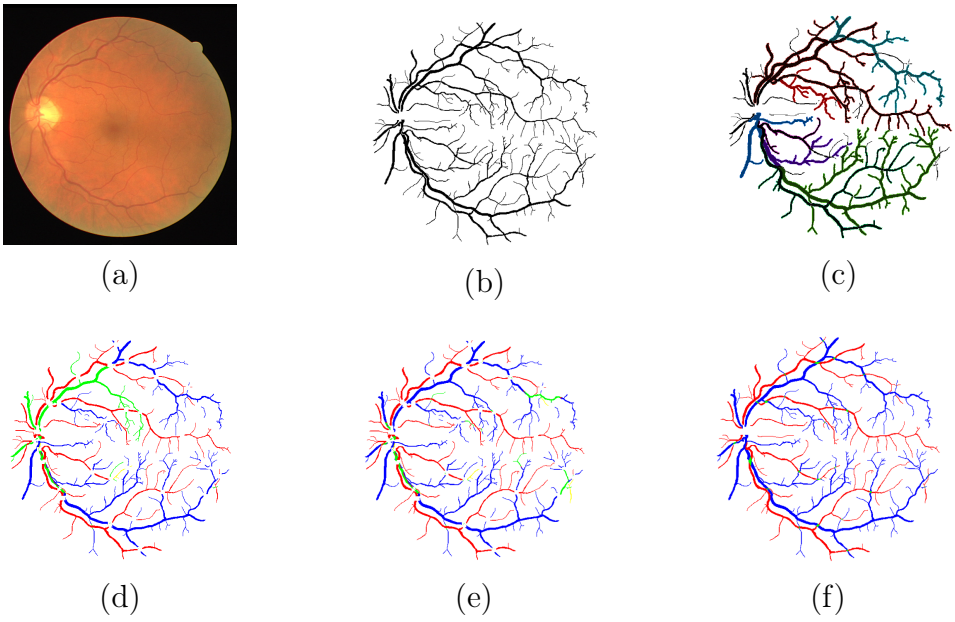



Figure G.1: AV-DRIVE A/V separation results. (a) Original image; (b) binary vessel map; (c) identified vessel subtrees; (d) A/V separation results without segment-wise analysis; (e) final A/V separation results (with segment-wise analysis); (f) corresponding A/V ground truth. (*Note*: the correctly labeled arteries/veins are shown in red/blue, respectively. While, incorrectly labeled arteries/veins are shown in yellow/green, respectively).

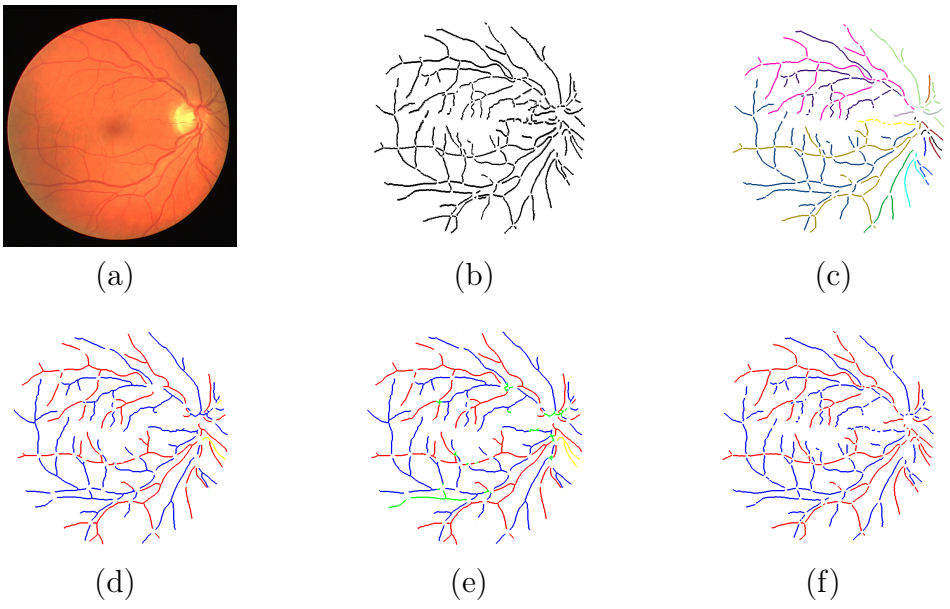



Figure G.2: CT-DRIVE A/V separation results. (a) Original image; (b) binary vessel map; (c) identified vessel subtrees; (d) A/V separation results without segment-wise analysis; (e) final A/V separation results (with segment-wise analysis); (f) corresponding A/V ground truth. (*Note*: the correctly labeled arteries/veins are shown in red/blue, respectively. While, incorrectly labeled arteries/veins are shown in yellow/green, respectively).

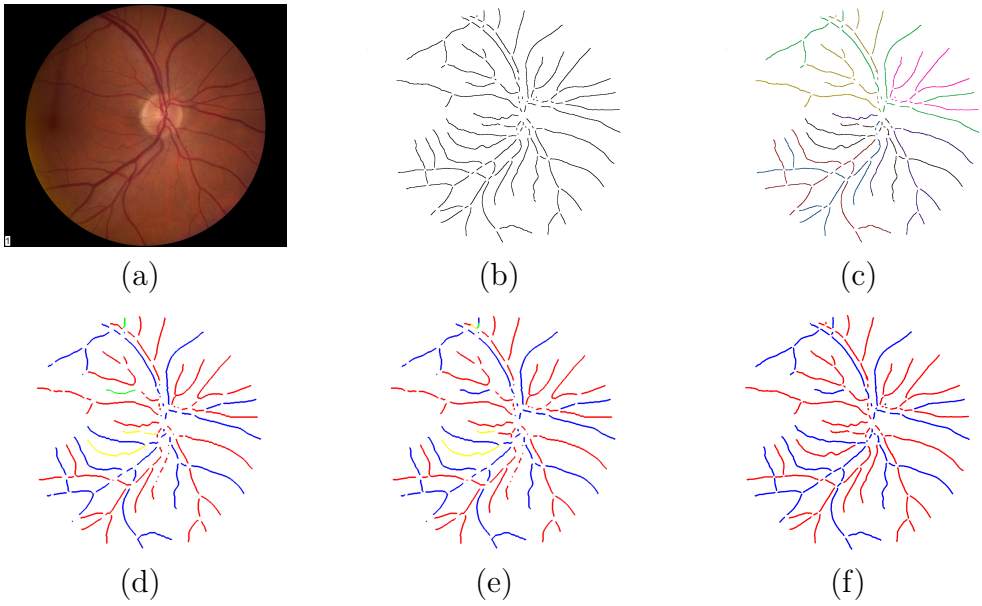



Figure G.3: INSPIRE-AVR A/V separation results. (a) Original image; (b) binary vessel map; (c) identified vessel subtrees; (d) A/V separation results without segment-wise analysis; (e) final A/V separation results (with segment-wise analysis); (f) corresponding A/V ground truth. (*Note*: the correctly labeled arteries/veins are shown in red/blue, respectively. While, incorrectly labeled arteries/veins are shown in yellow/green, respectively).

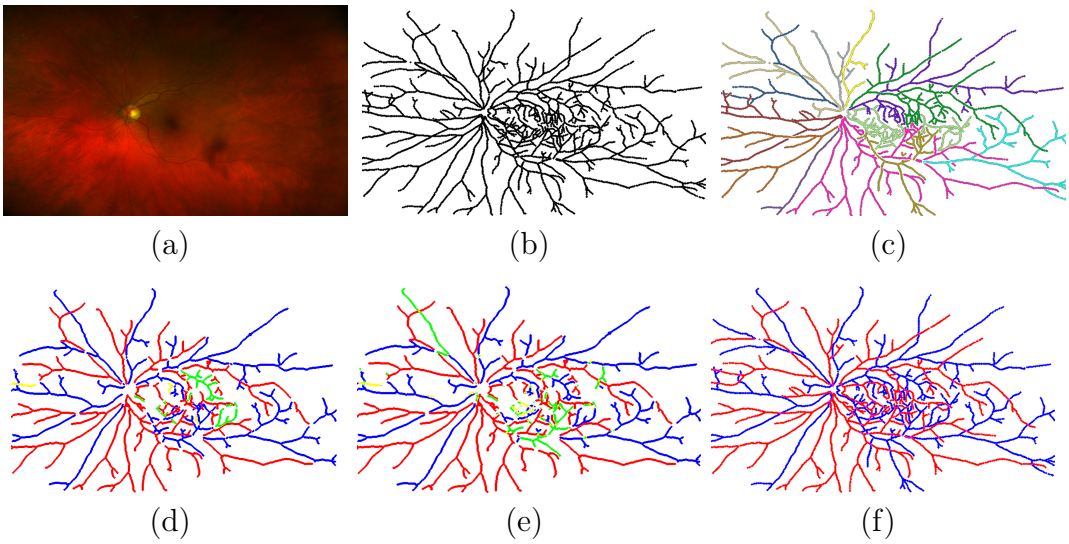



Figure G.4: Wide A/V separation results. (a) Original image; (b) binary vessel map; (c) identified vessel subtrees; (d) A/V separation results without segment-wise analysis; (e) final A/V separation results (with segment-wise analysis); (f) corresponding A/V ground truth. (*Note*: the correctly labeled arteries/veins are shown in red/blue, respectively. While, incorrectly labeled arteries/veins are shown in yellow/green, respectively).

G.1 depicts the performance analysis of proposed A/V separation without and with segment-wise analysis during subtree A/V labeling stage. It has been observed from Table G.1, that the performance of proposed approach (*Acc*) improves by a factor of 2.5% on AV-DRIVE, 1.4% on CT-DRIVE, 1.7% on INSPIRE-AVR and 4.2% on the WIDE dataset. A significant improvement in *Acc* value was observed after segment-wise analysis on the WIDE dataset, because the extracted vessel subtrees are found to be suboptimal, especially in the region of multiple close bifurcation and crossover points. Further, the identified vessel subtrees are significantly large in WIDE, due to wider field-of-view (FOV) when compared to other datasets.

# H Evaluation of A/V Separation as a Function of Segmentation Quality

An in-depth experimental analysis has been carried out to evaluate the proposed method's performance as a function of segmentation quality. In this regard, the three different vessel segmentation scenarios are considered:

- **Ground truth** – manual segmentation,
- **B-COSFIRE** – our current segmentation approach (Azzopardi et al. (2015)),
- **Multi-scale line detection** (**MSLD**) – a less accurate segmentation approach (Nguyen et al. (2013b)).

For generating binary vessel maps for the above two methods, the publicly available implementations are adopted:

- **B-COSFIRE** (*https://in.mathworks.com/matlabcentral/fileexchange/49172–trainable-cosfire-filters-for-curvilinear-structure-delineation-in-images*).
- **MSLD** (*https://people.eng.unimelb.edu.au/thivun/projects/retinal_ segmentation/*).

For the aforementioned experiments, the images from DRIVE (Qureshi et al. (2013)) and IOSTAR (Zhang et al. (2016)) datasets are considered, as it contains manually segmented ground truth vessel maps. Whereas, the images from INSPIRE-AVR (Dashtbozorg et al. (2014)) and WIDE (Estrada et al. (2015)) datasets doesn't contain any manually segmented vessel maps and hence, these datasets are not considered for further analysis.

Table H.1: Performance analysis of different vessel segmentation methods.

| Image No | Se | | Sp | | Acc | | F1-score | | MCC | |
|---|---|---|---|---|---|---|---|---|---|---|
| | B-COSFIRE | MSLD | B-COSFIRE | MSLD | B-COSFIRE | MSLD | B-COSFIRE | MSLD | B-COSFIRE | MSLD |
| **Drive-04** | **0.9840** | 0.9731 | **0.8297** | 0.6875 | **0.9706** | 0.9484 | **0.9839** | 0.9628 | **0.8147** | 0.7855 |
| **Drive-07** | 0.9528 | **0.9854** | **0.9280** | 0.7847 | 0.9511 | **0.9710** | 0.9731 | **0.9844** | 0.7254 | **0.7797** |
| **IOSTAR-02** | **0.9723** | 0.9659 | **0.8251** | 0.6501 | **0.9613** | 0.9422 | **0.9789** | 0.9597 | **0.7437** | 0.7096 |
| **IOSTAR-31** | 0.9892 | **0.9984** | **0.6828** | 0.3849 | **0.9674** | 0.9615 | **0.9828** | 0.9799 | **0.6875** | 0.5886 |

Table H.2: Performance analysis of proposed A/V Classification method on different vessel segmentation approaches. *Note:* GT - groundtruth segmentation.

| Image No | Accuracy | | | Sensitivity | | | Specificity | | |
|---|---|---|---|---|---|---|---|---|---|
| | GT | B-COSFIRE | MSLD | GT | B-COSFIRE | MSLD | GT | B-COSFIRE | MSLD |
| **Drive-04** | **0.9867** | 0.9618 | 0.9531 | **0.9914** | 0.9812 | 0.9676 | **0.9820** | 0.9424 | 0.9387 |
| **Drive-07** | **0.9927** | 0.9472 | 0.9583 | **0.9982** | 0.9705 | 0.9695 | **0.9873** | 0.9239 | 0.9471 |
| **IOSTAR-02** | **0.9937** | 0.9851 | 0.9704 | **0.9874** | 0.9789 | 0.9522 | **1.00** | 0.9913 | 0.9887 |
| **IOSTAR-31** | **0.9907** | 0.9777 | 0.9363 | **0.9937** | 0.9628 | 0.8727 | 0.9878 | 0.9926 | **1.00** |

It has been shown in previous studies (Azzopardi et al. (2015); Orlando et al. (2017); Zhang et al. (2016)) that the *Matthews Correlation Coefficient* (MCC) provides a more reliable estimate of the quality of binary segmentation with highly imbalanced class ratios. In case of retinal images, only a small portion of pixels ($\approx 9-14\%$) belong to vessels, while the rest belong to the background. This has also been observed from our experiments (see Table H.1) that the images having a smaller difference in $F1-score$ exhibits a more significant difference in their $MCC$ value, indicating a better estimate of the segmentation quality. Henceforth, both $F1-score$ and $MCC$ measures are reported in the evaluation of vessel segmentation methods.

A sample A/V separation results for different vessel segmentation approaches are shown in Figure H.1, Figure H.2, Figure H.3 and Figure H.4. It is observed that most of the false classifications are in the region of merging of close parallel vessels, discontinuity in vessel structures - especially at crossover points, splitting of vessel structure into too close vessels - due to central vessel reflex and missing of thin and peripheral vessels.

Table H.2 depicts the performance of the proposed A/V separation approach on different vessel segmentation methods. It is true that the A/V separation performance improves progressively with increase in quality of segmented binary vessel map. This has been observed from Table H.2 that, for every 3% increase in $MCC$ value results in $\approx 1\%$ improvement in accuracy of A/V classification performance. Besides, when compared with A/V separation on ground truth vessel map, the cur-

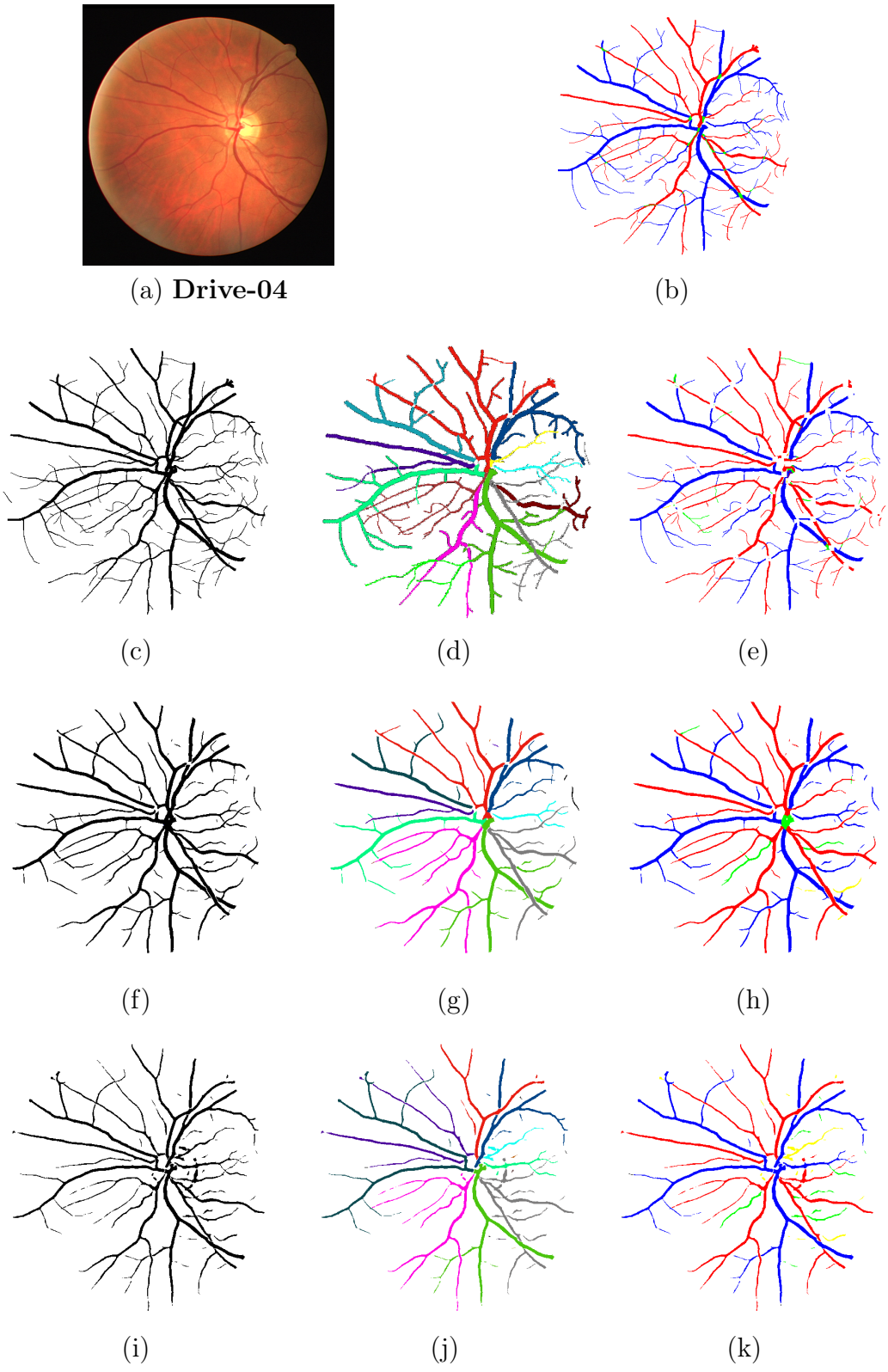


Figure H.1: (a) Original image (**Drive-04**); (b) corresponding A/V ground truth; (c) manual ground truth segmentation; (d) and (e) corresponding vessel subtrees and A/V separation result, respectively; (f) segmentation using B-COSFIRE method (Azzopardi et al. (2015)); (g) and (h) corresponding vessel subtrees and A/V separation result, respectively; (i) segmentation using MSLD method (Nguyen et al. (2013b)); (j) and (k) corresponding vessel subtrees and A/V separation result, respectively. (*Note*: the correctly labeled arteries/veins are shown in red/blue, respectively. While, incorrectly labeled arteries/veins are shown in yellow/green, respectively).

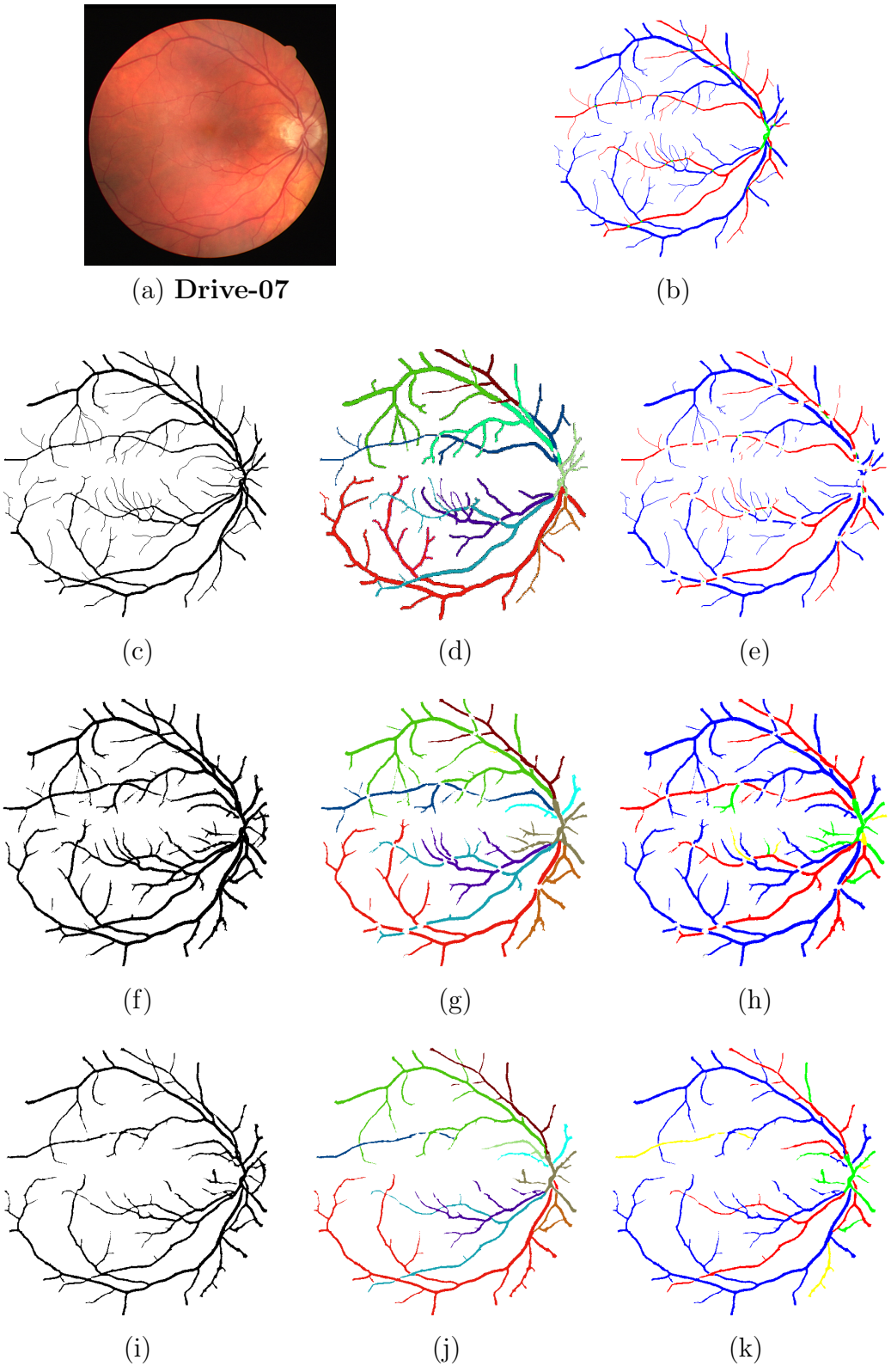


Figure H.2: (a) Original image (**Drive-07**); (b) corresponding A/V ground truth; (c) manual ground truth segmentation; (d) and (e) corresponding vessel subtrees and A/V separation result, respectively; (f) segmentation using B-COSFIRE method (Azzopardi et al. (2015)); (g) and (h) corresponding vessel subtrees and A/V separation result, respectively; (i) segmentation using MSLD method (Nguyen et al. (2013b)); (j) and (k) corresponding vessel subtrees and A/V separation result, respectively. (*Note*: the correctly labeled arteries/veins are shown in red/blue, respectively. While, incorrectly labeled arteries/veins are shown in yellow/green, respectively).

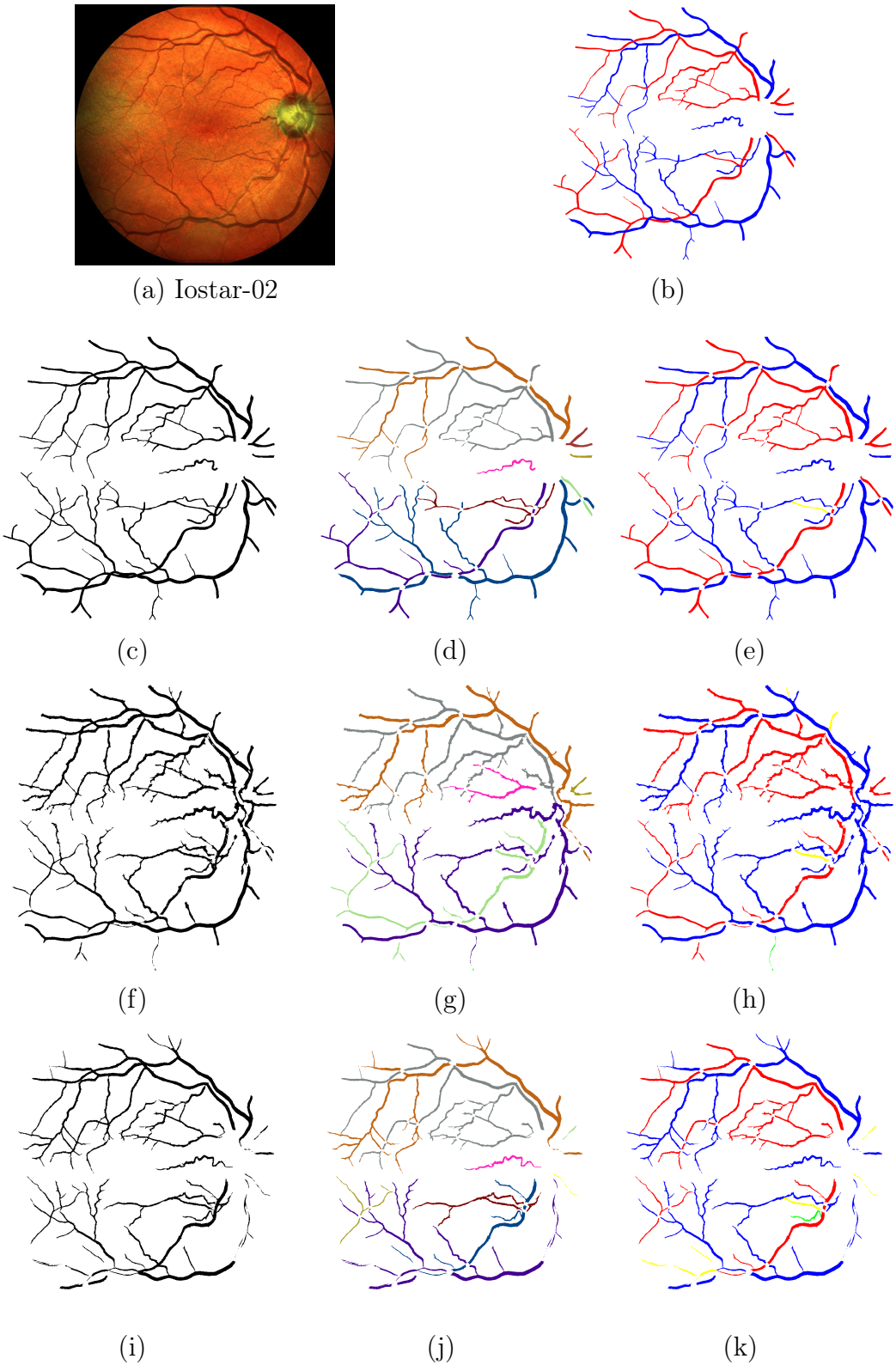
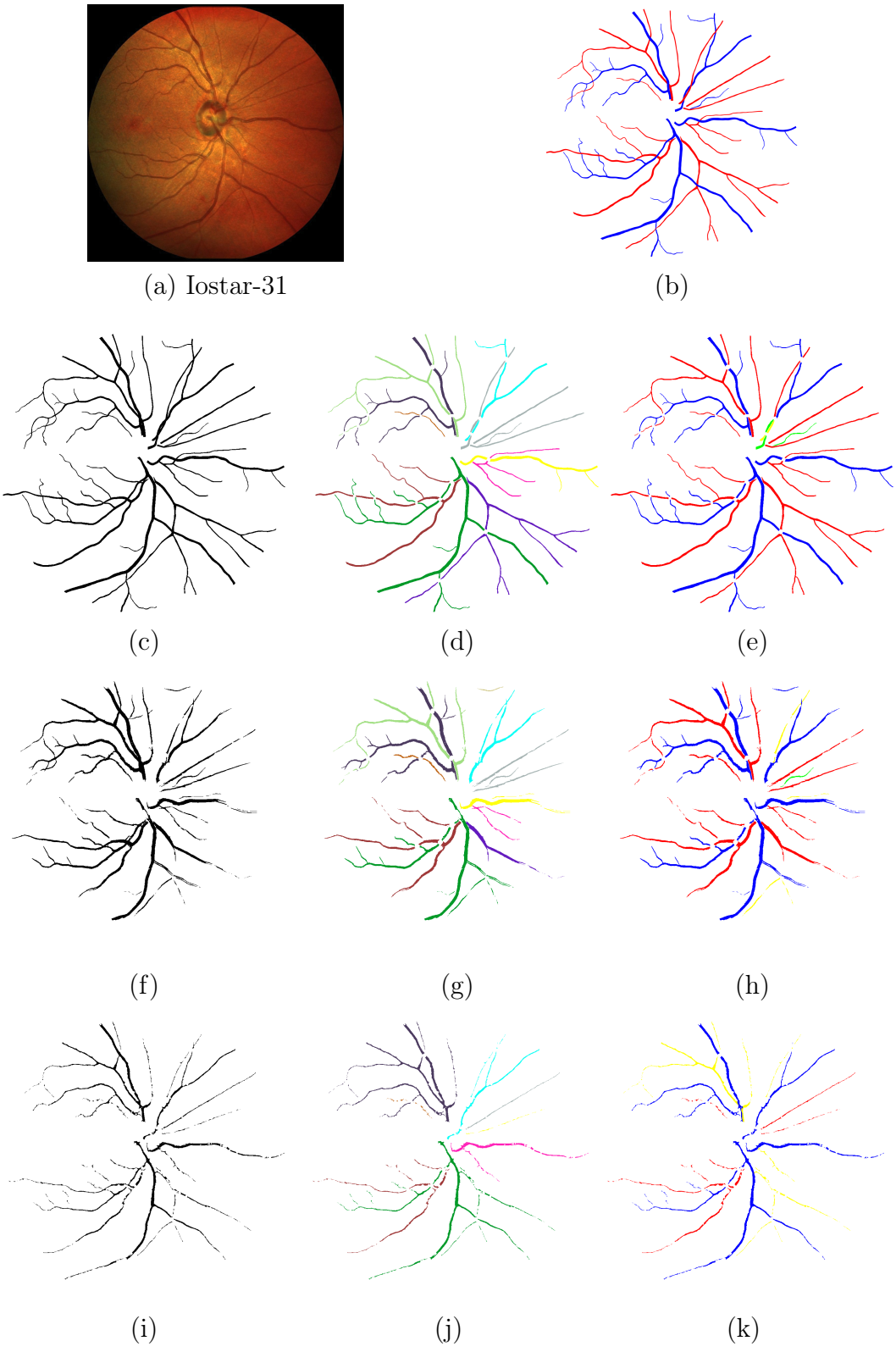


Figure H.3: (a) Original image (**Iostar-02**); (b) corresponding A/V ground truth; (c) manual ground truth segmentation; (d) and (e) corresponding vessel subtrees and A/V separation result, respectively; (f) segmentation using B-COSFIRE method (Azzopardi et al. (2015)); (g) and (h) corresponding vessel subtrees and A/V separation result, respectively; (i) segmentation using MSLD method (Nguyen et al. (2013b)); (j) and (k) corresponding vessel subtrees and A/V separation result, respectively. (*Note*: the correctly labeled arteries/veins are shown in red/blue, respectively. While, incorrectly labeled arteries/veins are shown in yellow/green, respectively).



(a) Iostar-31 (b) (c) (d) (e) (f) (g) (h) (i) (j) (k)

Figure H.4: (a) Original image (**Iostar-31**); (b) corresponding A/V ground truth; (c) manual ground truth segmentation; (d) and (e) corresponding vessel subtrees and A/V separation result, respectively; (f) segmentation using B-COSFIRE method (Azzopardi et al. (2015)); (g) and (h) corresponding vessel subtrees and A/V separation result, respectively; (i) segmentation using MSLD method (Nguyen et al. (2013b)); (j) and (k) corresponding vessel subtrees and A/V separation result, respectively. (*Note*: the correctly labeled arteries/veins are shown in red/blue, respectively. While, incorrectly labeled arteries/veins are shown in yellow/green, respectively).

rent segmentation approach (B-COSFIRE) yields a slight decrease in accuracy, while the MSLD approach results in a moderate decline in overall classification performance. Although, the proposed method performance is strongly dependent on the quality of input binary vessel map, the above experimental results signify that the method still obtains a more satisfactory A/V separation even on a less accurate noisy vessel map.

# References


Abbasi-Sureshjani, S., Smit-Ockeloen, I., Bekkers, E., Dashtbozorg, B., and t. H. Romeny, B. M. (2016). Automatic detection of vascular bifurcations and crossings in retinal images using orientation scores. In *IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, pages 189–192.

Abbasi-Sureshjani, S., Smit-Ockeloen, I., Zhang, J., and Ter Haar Romeny, B. (2015). Biologically-inspired supervised vasculature segmentation in slo retinal fundus images. In *International Conference on Image Analysis and Recognition*, pages 325–334.

Abdallah, M. B., Malek, J., Azar, A. T., Montesinos, P., Belmabrouk, H., Monreal, J. E., and Krissian, K. (2015). Automatic extraction of blood vessels in the retinal vascular tree using multiscale medialness. *Journal of Biomedical Imaging*, 2015:1.

Abramoff, M. D., Garvin, M. K., and Sonka, M. (2010). Retinal imaging and image analysis. *IEEE Reviews in Biomedical Engineering*, 3:169–208.

Aibinu, A., Iqbal, M., Shafie, A., Salami, M., and Nilsson, M. (2010). Vascular intersection detection in retina fundus images using a new hybrid approach. *Computers in Biology and Medicine*, 40(1):81–89.

Al-Diri, B., Hunter, A., and Steel, D. (2009). An active contour model for segmenting and measuring retinal vessels. *IEEE Transactions on Medical Imaging*, 28(9):1488–1497.

Al-Diri, B., Hunter, A., Steel, D., Habib, M., Hudaib, T., and Berry, S. (2008). Review - a reference data set for retinal vessel profiles. In *30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 2262–2265.

Alipour, S. H. M., Rabbani, H., and Akhlaghi, M. (2014). A new combined method based on curvelet transform and morphological operators for automatic detection of foveal avascular zone. *Signal, Image and Video Processing*, 8(2):205–222.

Annunziata, R., Garzelli, A., Ballerini, L., Mecocci, A., and Trucco, E. (2016). Leveraging multiscale hessian-based enhancement with a novel exudate inpainting technique for retinal vessel segmentation. *IEEE Journal of Biomedical and Health Informatics*, 20(4):1129–1138.

Aria (2006). *Online Retinal Image Archive*. `http://www.eyecharity.com/aria_online/` (accessed June 3, 2016).

Azzopardi, G. and Petkov, N. (2013). Automatic detection of vascular bifurcations in segmented retinal images using trainable cosfire filters. *Pattern Recognition Letters*, 34(8):922–933.

Azzopardi, G., Strisciuglio, N., Vento, M., and Petkov, N. (2015). Trainable cosfire filters for vessel delineation with application to retinal images. *Medical Image Analysis*, 19(1):46–57.

Baker, M. L., Wang, J. J., Liew, G., Hand, P. J., De Silva, D. A., Lindley, R. I., Mitchell, P., Wong, M.-C., Rochtchina, E., Wong, T. Y., Wardlaw, J. M., and Hankey, G. J. a. (2010). Differential associations of cortical and subcortical cerebral atrophy with retinal vascular signs in patients with acute stroke. *Stroke*, 41(10):2143–2150.

Bankhead, P., Scholfield, C. N., McGeown, J. G., and Curtis, T. M. (2012). Fast retinal vessel detection and measurement using wavelets and edge location refinement. *PloS one*, 7(3):e32435.

Bekkers, E., Duits, R., Berendschot, T., and ter Haar Romeny, B. (2014). A multi-orientation analysis approach to retinal vessel tracking. *Journal of Mathematical Imaging and Vision*, 49(3):583–610.

Bell, A. J. and Sejnowski, T. J. (1997). The "independent components" of natural scenes are edge filters. *Vision Research*, 37(23):3327–3338.

Borji, A. and Itti, L. (2013). State-of-the-art in visual attention modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(1):185–207.

Breiman, L. (2001). Random forests. *Machine Learning*, 45(1):5–32.

Budai, A., Bock, R., Maier, A., Hornegger, J., and Michelson, G. (2013). Robust vessel segmentation in fundus images. *International Journal of Biomedical Imaging*, 2013.

Chaudhuri, S., Chatterjee, S., Katz, N., Nelson, M., and Goldbaum, M. (1989). Detection of blood vessels in retinal images using two-dimensional matched filters. *IEEE Transactions on Medical Imaging*, 8(3):263–269.

Cheung, B., Weiss, E., and Olshausen, B. A. (2016). Emergence of foveal image sampling from learning to attend in visual scenes. *CoRR*, abs/1611.09430.

Cheung, C. S., Butty, Z., Tehrani, N. N., and Lam, W. C. (2011a). Computer-assisted image analysis of temporal retinal vessel width and tortuosity in retinopathy of prematurity for the assessment of disease severity and treatment outcome. *Journal of American Association for Pediatric Ophthalmology and Strabismus*, 15(4):374–380.

Cheung, C. Y., Ikram, M. K., Chen, C., and Wong, T. Y. (2017). Imaging retina to study dementia and stroke. *Progress in Retinal and Eye Research*, 57:89–107.

Cheung, C. Y., Ikram, M. K., Sabanayagam, C., and Wong, T. Y. (2012). Retinal microvasculature as a model to study the manifestations of hypertension. *Hypertension*, 60(5):1094–1103.

Cheung, C. Y., Ong, Y. T., Ikram, M. K., Ong, S. Y., Li, X., Hilal, S., Catindig, J.-A. S., Venketasubramanian, N., Yap, P., Seow, D., Chen, C. P., and Wong, T. Y. (2014). Microvascular network alterations in the retina of patients with alzheimer's disease. *Alzheimer's & Dementia*, 10(2):135–142.

Cheung, C. Y., Tay, W. T., Mitchell, P., Wang, J. J., Hsu, W., Lee, M. L., Lau, Q. P., Zhu, A. L., Klein, R., Saw, S. M., et al. (2011b). Quantitative and qualitative retinal microvascular characteristics and blood pressure. *Journal of Hypertension*, 29(7):1380–1391.

Christodoulidis, A., Hurtut, T., Tahar, H. B., and Cheriet, F. (2016). A multi-scale tensor voting approach for small retinal vessel segmentation in high resolution fundus images. *Computerized Medical Imaging and Graphics*, 52:28–43.

Claudia Kondermann, Daniel Kondermann, M. Y. (2007). Blood vessel classification into arteries and veins in retinal images. In *Proc. SPIE*, volume 6512, pages 6512–6512–9.

Coates, A., Ng, A., and Lee, H. (2011). An analysis of single-layer networks in unsupervised feature learning. In *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, volume 15, pages 215–223.

Condurache, A. P. and Mertins, A. (2012). Segmentation of retinal vessels with a hysteresis binary-classification paradigm. *Computerized Medical Imaging and Graphics*, 36(4):325–335.

Costa, P., Galdran, A., Meyer, M. I., Niemeijer, M., Abràmoff, M., Mendonça, A. M., and Campilho, A. (2018). End-to-end adversarial retinal image synthesis. *IEEE Transactions on Medical Imaging*, 37(3):781–791.

Cugati, S., Cumming, R. G., Smith, W., Burlutsky, G., Mitchell, P., and Wang, J. J. (2007). Visual impairment, age-related macular degeneration, cataract, and long-term mortality: The Blue Mountains Eye Study. *Archives of Ophthalmology*, 125(7):917–924.

Dai, P., Luo, H., Sheng, H., Zhao, Y., Li, L., Wu, J., Zhao, Y., and Suzuki, K. (2015). A new approach to segment both main and peripheral retinal vessels based on gray-voting and gaussian mixture model. *PloS one*, 10(6):e0127748.

Dalal, N. and Triggs, B. (2005). Histograms of oriented gradients for human detection. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 886–893.

Dashtbozorg, B., Mendonça, A. M., and Campilho, A. (2014). An automatic graph-based approach for artery/vein classification in retinal images. *IEEE Transactions on Image Processing*, 23(3):1073–1083.

De, J., Li, H., and Cheng, L. (2014). Tracing retinal vessel trees by transductive inference. *BMC Bioinformatics*, 15.

Defferrard, M., Bresson, X., and Vandergheynst, P. (2016). Convolutional neural networks on graphs with fast localized spectral filtering. In *Advances in Neural Information Processing Systems (NIPS)*, pages 3844–3852.

Diaretdb1 (2007). *Diabetic retinopathy database and evaluation protocol*. `http://http://www.it.lut.fi/project/imageret/diaretdb1/` (accessed May 13, 2016).

Doubal, F. N., Hokke, P. E., and Wardlaw, J. M. (2009). Retinal microvascular abnormalities and stroke: a systematic review. *Journal of Neurology, Neurosurgery & Psychiatry*, 80(2):158–165.

Estrada, R., Allingham, M. J., Mettu, P. S., Cousins, S. W., Tomasi, C., and Farsiu, S. (2015). Retinal artery-vein classification via topology estimation. *IEEE Transactions on Medical Imaging*, 34(12):2518–2534.

Fathi, A. and Naghsh-Nilchi, A. R. (2013). Automatic wavelet-based retinal blood vessels segmentation and vessel diameter estimation. *Biomedical Signal Processing and Control*, 8(1):71–80.

Fathi, A. and Naghsh-Nilchi, A. R. (2014). General rotation-invariant local binary patterns operator with application to blood vessel detection in retinal images. *Pattern Analysis and Applications*, 17(1):69–81.

Fathi, A., Naghsh-Nilchi, A. R., and Mohammadi, F. A. (2013). Automatic vessel network features quantification using local vessel pattern operator. *Computers in Biology and Medicine*, 43(5):587–593.

Favali, M., Abbasi-Sureshjani, S., ter HaarRomeny, B., and Sarti, A. (2016). Analysis of vessel connectivities in retinal images by cortically inspired spectral clustering. *Journal of Mathematical Imaging and Vision*, 56(1):158–172.

Fawzi, A., Davies, M., and Frossard, P. (2015). Dictionary learning for fast classification based on soft-thresholding. *International Journal of Computer Vision*, 114(2):306–321.

Felberer, F., Rechenmacher, M., Haindl, R., Baumann, B., Hitzenberger, C. K., and Pircher, M. (2015). Imaging of retinal vasculature using adaptive optics SLO/OCT. *Biomedical Optic Express*, 6(4):1407–1418.

Foong, A. W., Saw, S.-M., Loo, J.-L., Shen, S., Loon, S.-C., Rosman, M., Aung, T., Tan, D. T., Tai, E. S., and Wong, T. Y. (2007). Rationale and methodology for a population-based study of eye diseases in malay people: The Singapore Malay Eye Study (SiMES). *Ophthalmic Epidemiology*, 14(1):25–35.

Foracchia, M., Grisan, E., and Ruggeri, A. (2005). Luminosity and contrast normalization in retinal images. *Medical Image Analysis*, 9(3):179–190.

Frangi, A. F., Niessen, W. J., Vincken, K. L., and Viergever, M. A. (1998). Multiscale vessel enhancement filtering. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 130–137.

Fraz, M., Barman, S., Remagnino, P., Hoppe, A., Basit, A., Uyyanonvara, B., Rudnicka, A., and Owen, C. (2012a). An approach to localize the retinal blood vessels using bit planes and centerline detection. *Computer Methods and Programs in Biomedicine*, 108(2):600–616.

Fraz, M., Remagnino, P., Hoppe, A., Rudnicka, A., Owen, C., Whincup, P., and Barman, S. (2013). Quantification of blood vessel calibre in retinal images of multi-ethnic school children using a model based approach. *Computerized Medical Imaging and Graphics*, 37(1):48–60.

Fraz, M., Remagnino, P., Hoppe, A., Uyyanonvara, B., Rudnicka, A., Owen, C., and Barman, S. (2012b). Blood vessel segmentation methodologies in retinal images – a survey. *Computer Methods and Programs in Biomedicine*, 108(1):407–433.

Fraz, M. M., Remagnino, P., Hoppe, A., Uyyanonvara, B., Rudnicka, A. R., Owen, C. G., and Barman, S. A. (2012c). An ensemble classification-based approach applied to retinal blood vessel segmentation. *IEEE Transactions on Biomedical Engineering*, 59(9):2538–2548.

Fraz, M. M., Rudnicka, A. R., Owen, C. G., and Barman, S. A. (2014). Delineation of blood vessels in pediatric retinal images using decision trees-based ensemble classification. *International Journal of Computer Assisted Radiology and Surgery*, 9(5):795–811.

Friedman, J., Hastie, T., and Tibshirani, R. (2001). *The Elements of Statistical Learning*, volume 1. Springer series in statistics New York, NY, USA.

Frost, S., Kanagasingam, Y., Sohrabi, H., Vignarajan, J., Bourgeat, P., Salvado, O., Villemagne, V., Rowe, C., Macaulay, S. L., Szoeke, C., et al. (2013). Retinal vascular biomarkers for early detection and monitoring of alzheimer's disease. *Translational Psychiatry*, 3(2):e233.

Frucci, M., Riccio, D., di Baja, G. S., and Serino, L. (2016). Severe: Segmenting vessels in retina images. *Pattern Recognition Letters*, 82:162–169.

Fu, H., Xu, Y., Lin, S., Wong, D. W. K., and Liu, J. (2016a). Deepvessel: Retinal vessel segmentation via deep learning and conditional random field. In *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 132–139.

Fu, H., Xu, Y., Wong, D. W. K., and Liu, J. (2016b). Retinal vessel segmentation via deep learning network and fully-connected conditional random fields. In *IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, pages 698–701.

Gang, L., Chutatape, O., and Krishnan, S. M. (2002). Detection and measurement of retinal vessels in fundus images using amplitude modified second-order gaussian filter. *IEEE Transactions on Biomedical Engineering*, 49(2):168–172.

Ganjee, R., Azmi, R., and Gholizadeh, B. (2014). An improved retinal vessel segmentation method based on high level features for pathological images. *Journal of Medical Systems*, 38(9):108.

Gegundez-Arias, M. E., Aquino, A., Bravo, J. M., and Marin, D. (2012). A function for quality evaluation of retinal vessel segmentations. *IEEE Transactions on Medical Imaging*, 31(2):231–239.

Gray, A. (1996). *Modern Differential Geometry of Curves and Surfaces with Mathematica*. CRC Press, Inc., 1st edition.

Gregor, K. and LeCun, Y. (2010). Learning fast approximations of sparse coding. In *Proceedings of the 27th International Conference on Machine Learning (ICML)*, pages 399–406.

Grisan, E., Foracchia, M., and Ruggeri, A. (2008). A novel method for the automatic grading of retinal vessel tortuosity. *IEEE Transactions on Medical Imaging*, 27(3):310–319.

Grisan, E. and Ruggeri, A. (2003). A divide et impera strategy for automatic classification of retinal vessels into arteries and veins. In *Proceedings of the 25th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, volume 1, pages 890–893.

Hassanien, A. E., Emary, E., and Zawbaa, H. M. (2015). Retinal blood vessel localization approach based on bee colony swarm optimization, fuzzy c-means and pattern search. *Journal of Visual Communication and Image Representation*, 31:186–196.

Heringa, S. M., Bouvy, W. H., van den Berg, E., Moll, A. C., Kappelle, L. J., and Biessels, G. J. (2013). Associations between retinal microvascular changes and dementia, cognitive functioning, and brain imaging abnormalities: A systematic review. *Journal of Cerebral Blood Flow & Metabolism*, 33(7):983–995.

Hinton, G. E., Osindero, S., and Teh, Y.-W. (2006). A fast learning algorithm for deep belief nets. *Neural Computation*, 18(7):1527–1554.

Hogan, M. J. (1971). Histology of the human eye. *An Atlas and Textbook*.

Hoover, A. and Goldbaum, M. (2003). Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels. *IEEE Transactions on Medical Imaging*, 22(8):951–958.

Hoover, A. D., Kouznetsova, V., and Goldbaum, M. (2000). Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response. *IEEE Transactions on Medical Imaging*, 19(3):203–210.

Hu, Q., Abràmoff, M. D., and Garvin, M. K. (2015). Automated construction of arterial and venous trees in retinal images. *Journal of Medical Imaging*, 2:2–2–16.

Huang, F., Dashtbozorg, B., Tan, T., and ter Haar Romeny, B. M. (2018a). Retinal artery/vein classification using genetic-search feature selection. *Computer Methods and Programs in Biomedicine*, 161:197–207.

Huang, F., Dashtbozorg, B., and ter Haar Romeny, B. M. (2018b). Artery/vein classification using reflection features in retina fundus images. *Machine Vision and Applications*, 29(1):23–34.

Hubbard, L. D., Brothers, R. J., King, W. N., Clegg, L. X., Klein, R., Cooper, L. S., Sharrett, A., Davis, M. D., and Cai, J. (1999). Methods for evaluation of retinal microvascular abnormalities associated with hypertension/sclerosis in the atherosclerosis risk in communities study. *Ophthalmology*, 106(12):2269–2280.

Imani, E., Javidi, M., and Pourreza, H.-R. (2015). Improvement of retinal blood vessel detection using morphological component analysis. *Computer Methods and Programs in Biomedicine*, 118(3):263–279.

Joshi, G. D. and Sivaswamy, J. (2008). Colour retinal image enhancement based on domain knowledge. In *Sixth Indian Conference on Computer Vision, Graphics Image Processing*, pages 591–598.

Joshi, V. S., Reinhardt, J. M., Garvin, M. K., and Abramoff, M. D. (2014). Automated method for identification and artery-venous classification of vessel trees in retinal vessel networks. *PloS one*, 9(2):e88061.

Kar, S. S. and Maity, S. P. (2016). Blood vessel extraction and optic disc removal using curvelet transform and kernel fuzzy c-means. *Computers in Biology and Medicine*, 70:174–189.

Kovács, G. and Hajdu, A. (2016). A self-calibrating approach for the segmentation of retinal vessels by template matching and contour reconstruction. *Medical Image Analysis*, 29:24–46.

Krause, M., Alles, R. M., Burgeth, B., and Weickert, J. (2016). Fast retinal vessel analysis. *Journal of Real-Time Image Processing*, 11(2):413–422.

Lam, B. S. Y., Gao, Y., and Liew, A. W. C. (2010). General retinal vessel segmentation using regularization-based multiconcavity modeling. *IEEE Transactions on Medical Imaging*, 29(7):1369–1381.

Larochelle, H. and Hinton, G. E. (2010). Learning to combine foveal glimpses with a third-order boltzmann machine. In *Advances in Neural Information Processing Systems (NIPS)*, pages 1243–1251.

Lee, H., Battle, A., Raina, R., and Ng, A. Y. (2007). Efficient sparse coding algorithms. In *Advances in Neural Information Processing Systems (NIPS)*, pages 801–808.

Li, Q., Feng, B., Xie, L., Liang, P., Zhang, H., and Wang, T. (2016). A cross-modality learning approach for vessel segmentation in retinal images. *IEEE Transactions on Medical Imaging*, 35(1):109–118.

Liskowski, P. and Krawiec, K. (2016). Segmenting retinal blood vessels with deep neural networks. *IEEE Transactions on Medical Imaging*, 35(11):2369–2380.

London, A., Benhar, I., and Schwartz, M. (2013). The retina as a window to the brain from eye research to cns disorders. *Nature Reviews Neurology*, 9:44–53.

Lupascu, C. A., Tegolo, D., and Trucco, E. (2010). Fabc: Retinal vessel segmentation using adaboost. *IEEE Transactions on Information Technology in Biomedicine*, 14(5):1267–1274.

Lázár, I. and Hajdu, A. (2015). Segmentation of retinal vessels by means of directional response vector similarity and region growing. *Computers in Biology and Medicine*, 66:209–221.

Maninis, K.-K., Pont-Tuset, J., Arbeláez, P., and Van Gool, L. (2016). Deep retinal image understanding. In *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pages 140–148.

Mapayi, T., Viriri, S., and Tapamo, J.-R. (2015a). Adaptive thresholding technique for retinal vessel segmentation based on glcm-energy information. *Computational and Mathematical Methods in Medicine*, 2015.

Mapayi, T., Viriri, S., and Tapamo, J.-R. (2015b). Comparative study of retinal vessel segmentation based on global thresholding techniques. *Computational and Mathematical Methods in Medicine*, 2015.

Marin, D., Aquino, A., Gegundez-Arias, M. E., and Bravo, J. M. (2011). 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.

Martinez-Perez, M. E., Hughes, A. D., Thom, S. A., Bharath, A. A., and Parker, K. H. (2007). Segmentation of blood vessels from red-free and fluorescein retinal images. *Medical Image Analysis*, 11(1):47–61.

McGrory, S., Cameron, J. R., Pellegrini, E., Warren, C., Doubal, F. N., Deary, I. J., Dhillon, B., Wardlaw, J. M., Trucco, E., and MacGillivray, T. J. (2017). The application of retinal fundus camera imaging in dementia: A systematic review. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 6:91–107.

Mendonca, A. M. and Campilho, A. (2006). Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction. *IEEE Transactions on Medical Imaging*, 25(9):1200–1213.

Messidor (2017). *Methods for Evaluating Segmentation and Indexing Techniques Dedicated to Retinal Ophthalmology.* `http://messidor.crihan.fr/index-en.php`.

Mirsharif, Q., Tajeripour, F., and Pourreza, H. (2013). Automated characterization of blood vessels as arteries and veins in retinal images. *Computerized Medical Imaging and Graphics*, 37(7):607–617.

Mitchell, P., Smith, W., Attebo, K., and Wang, J. J. (1995). Prevalence of age-related maculopathy in australia: The Blue Mountains Eye Study. *Ophthalmology*, 102(10):1450–1460.

Moghimirad, E., Rezatofighi, S. H., and Soltanian-Zadeh, H. (2012). Retinal vessel segmentation using a multi-scale medialness function. *Computers in Biology and Medicine*, 42(1):50–60.

Narasimha-Iyer, H., Mahadevan, V., Beach, J. M., and Roysam, B. (2008). Improved detection of the central reflex in retinal vessels using a generalized dual-gaussian model and robust hypothesis testing. *IEEE Transactions on Information Technology in Biomedicine*, 12(3):406–410.

Nguyen, U. T. V., Bhuiyan, A., Park, L. A. F., Kawasaki, R., Wong, T. Y., Wang, J. J., Mitchell, P., and Ramamohanarao, K. (2013a). An automated method for retinal arteriovenous nicking quantification from color fundus images. *IEEE Transactions on Biomedical Engineering*, 60(11):3194–3203.

Nguyen, U. T. V., Bhuiyan, A., Park, L. A. F., and Ramamohanarao, K. (2013b). An effective retinal blood vessel segmentation method using multi-scale line detection. *Pattern Recognition*, 46(3):703–715.

Niemeijer, M., Staal, J., van Ginneken, B., Loog, M., Abramoff, M. D., et al. (2004). Comparative study of retinal vessel segmentation methods on a new publicly available database. In *SPIE Medical Imaging*, volume 5370, pages 648–656.

Niemeijer, M., van Ginneken, B., and Abramoff, M. D. (2009). Automatic classification of retinal vessels into arteries and veins. In *SPIE Medical Imaging*, pages 72601F–72601F.

Niemeijer, M., Xu, X., Dumitrescu, A. V., Gupta, P., van Ginneken, B., Folk, J. C., and Abramoff, M. D. (2011). Automated measurement of the arteriolar-to-venular width ratio in digital color fundus photographs. *IEEE Transactions on Medical Imaging*, 30(11):1941–1950.

Núñez, J. M., Bernal, J., Sánchez, F. J., and Vilariño, F. (2015). Growing algorithm for intersection detection (graid) in branching patterns. *Machine Vision and Applications*, 26(2):387–400.

Odstrcilik, J., Kolar, R., Budai, A., Hornegger, J., Jan, J., Gazarek, J., Kubena, T., Cernosek, P., Svoboda, O., and Angelopoulou, E. (2013). Retinal vessel segmentation by improved matched filtering: evaluation on a new high-resolution fundus image database. *IET Image Processing*, 7(4):373–383.

Oloumi, F., Rangayyan, R. M., Casti, P., and Ells, A. L. (2015). Computer-aided diagnosis of plus disease via measurement of vessel thickness in retinal fundus images of preterm infants. *Computers in Biology and Medicine*, 66:316–329.

Orlando, J. I., Prokofyeva, E., and Blaschko, M. B. (2017). A discriminatively trained fully connected conditional random field model for blood vessel segmentation in fundus images. *IEEE Transactions on Biomedical Engineering*, 64(1):16–27.

Palomera-Pérez, M. A., Martinez-Perez, M. E., Benítez-Pérez, H., and Ortega-Arjona, J. L. (2010). Parallel multiscale feature extraction and region growing: Application in retinal blood vessel detection. *IEEE Transactions on Information Technology in Biomedicine*, 14(2):500–506.

Pellegrini, E., Robertson, G., MacGillivray, T., van Hemert, J., Houston, G., and Trucco, E. (2018). A graph cut approach to artery/vein classification in ultra-widefield scanning laser ophthalmoscopy. *IEEE Transactions on Medical Imaging*, 37(2):516–526.

Pereira, C., Veiga, D., Gonçalves, L., and Ferreira, M. (2014). Automatic arteriovenous nicking identification by color fundus images analysis. In *International Conference on Image Analysis and Recognition*, pages 321–328.

Perez-Rovira, A., Zutis, K., Hubschman, J. P., and Trucco, E. (2011). Improving vessel segmentation in ultra-wide field-of-view retinal fluorescein angiograms. In *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 2614–2617.

Pizer, S. M., Amburn, E. P., Austin, J. D., Cromartie, R., Geselowitz, A., Greer, T., ter Haar Romeny, B., Zimmerman, J. B., and Zuiderveld, K. (1987). Adaptive histogram equalization and its variations. *Computer Vision, Graphics, and Image Processing*, 39(3):355–368.

Qureshi, T. A., Habib, M., Hunter, A., and Al-Diri, B. (2013). A manually-labeled, artery/vein classified benchmark for the drive dataset. In *Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems*, pages 485–488.

Rahebi, J. and Hardalaç, F. (2014). Retinal blood vessel segmentation with neural network by using gray-level co-occurrence matrix-based features. *Journal of Medical Systems*, 38(8):85.

Ram, K., Babu, Y., and Sivaswamy, J. (2009). Curvature orientation histograms for detection and matching of vascular landmarks in retinal images. In *SPIE Medical Imaging, Image Processing*, volume 7259, pages 7259–7259–8.

Ranzato, M. (2014). On learning where to look. *arXiv preprint:1405.5488*.

Ranzato, M. A., lan Boureau, Y., and Cun, Y. L. (2008). Sparse feature learning for deep belief networks. In *Advances in Neural Information Processing Systems (NIPS)*, pages 1185–1192.

Ranzato, M. A., Poultney, C., Chopra, S., and Cun, Y. L. (2007). Efficient learning of sparse representations with an energy-based model. In *Advances in Neural Information Processing Systems (NIPS)*, pages 1137–1144.

Ricci, E. and Perfetti, R. (2007). Retinal blood vessel segmentation using line operators and support vector classification. *IEEE Transactions on Medical Imaging*, 26(10):1357–1365.

Rosenbaum, D., Kachenoura, N., Koch, E., Paques, M., Cluzel, P., Redheuil, A., and Girerd, X. (2016). Relationships between retinal arteriole anatomy and aortic geometry and function and peripheral resistance in hypertensives. *Hypertension Research*, 39(7):536–542.

Rothaus, K., Jiang, X., and Rhiem, P. (2009). Separation of the retinal vascular graph in arteries and veins based upon structural knowledge. *Image and Vision Computing*, 27(7):864–875.

Roy, P. K., Nguyen, U. T. V., Bhuiyan, A., and Ramamohanarao, K. (2014). An effective automated system for grading severity of retinal arteriovenous nicking in color retinal images. In *36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 6324–6327.

Roychowdhury, S., Koozekanani, D. D., and Parhi, K. K. (2015a). Blood vessel segmentation of fundus images by major vessel extraction and subimage classification. *IEEE Journal of Biomedical and Health Informatics*, 19(3):1118–1128.

Roychowdhury, S., Koozekanani, D. D., and Parhi, K. K. (2015b). Iterative vessel segmentation of fundus images. *IEEE Transactions on Biomedical Engineering*, 62(7):1738–1749.

Salazar-Gonzalez, A., Kaba, D., Li, Y., and Liu, X. (2014). Segmentation of the blood vessels and optic disk in retinal images. *IEEE Journal of Biomedical and Health Informatics*, 18(6):1874–1886.

Sharrett, A. R., Hubbard, L. D., Cooper, L. S., Sorlie, P. D., Brothers, R. J., Nieto, F. J., Pinsky, J. L., and Klein, R. (1999). Retinal arteriolar diameters and elevated blood pressure: the atherosclerosis risk in communities study. *American Journal of Epidemiology*, 150(3):263–270.

Sigurosson, E. M., Valero, S., Benediktsson, J. A., Chanussot, J., Talbot, H., and Stefansson, E. (2014). Automatic retinal vessel extraction based on directional mathematical morphology and fuzzy classification. *Pattern Recognition Letters*, 47:164–171.

Smith, W., Wang, J. J., Wong, T. Y., Rochtchina, E., Klein, R., Leeder, S. R., and Mitchell, P. (2004). Retinal arteriolar narrowing is associated with 5-year incident severe hypertension: the blue mountains eye study. *Hypertension*, 44(4):442–447.

Soares, J. V. B., Leandro, J. J. G., Cesar, R. M., Jelinek, H. F., and Cree, M. J. (2006). Retinal vessel segmentation using the 2-d gabor wavelet and supervised classification. *IEEE Transactions on Medical Imaging*, 25(9):1214–1222.

Sodi, A., Guarducci, M., Vauthier, L., Ioannidis, A. S., Pitz, S., Abbruzzese, G., Sofi, F., Mecocci, A., Miele, A., and Menchini, U. (2012). Computer assisted evaluation of retinal vessels tortuosity in fabry disease. *Acta Ophthalmologica*, 91(2):e113–e119.

Srinidhi, C. L., Aparna, P., and Rajan, J. (2017a). Recent advancements in retinal vessel segmentation. *Journal of Medical Systems*, 41(4):70.

Srinidhi, C. L., Aparna, P., and Rajan, J. (2018). A visual attention guided unsupervised feature learning for robust vessel delineation in retinal images. *Biomedical Signal Processing and Control*, 44:110–126.

Srinidhi, C. L., Rath, P., and Sivaswamy, J. (2017b). A vessel keypoint detector for junction classification. In *IEEE 14th International Symposium on Biomedical Imaging (ISBI)*, pages 882–885.

Staal, J., Abramoff, M. D., Niemeijer, M., Viergever, M. A., and van Ginneken, B. (2004). Ridge-based vessel segmentation in color images of the retina. *IEEE Transactions on Medical Imaging*, 23(4):501–509.

Strisciuglio, N., Azzopardi, G., Vento, M., and Petkov, N. (2016). Supervised vessel delineation in retinal fundus images with the automatic selection of b-cosfire filters. *Machine Vision and Applications*, 27(8):1137–1149.

Su, R., Sun, C., and Pham, T. D. (2012). Junction detection for linear structures based on hessian, correlation and shape information. *Pattern Recognition*, 45(10):3695–3706.

Sutter, F. K. and Helbig, H. (2003). Familial retinal arteriolar tortuosity: A review. *Survey of Ophthalmology*, 48(3):245–255.

Tham, Y.-C., Li, X., Wong, T. Y., Quigley, H. A., Aung, T., and Cheng, C.-Y. (2014). Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis. *Ophthalmology*, 121(11):2081–2090.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B*, 58:267–288.

Tso, M. O. and Jampol, L. M. (1982). Pathophysiology of hypertensive retinopathy. *Ophthalmology*, 89(10):1132–1145.

Tu, Z. and Bai, X. (2010). Auto-context and its application to high-level vision tasks and 3d brain image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(10):1744–1757.

Van Der Maaten, L. (2014). Accelerating t-sne using tree-based algorithms. *Journal of Machine Learning Research*, 15:3221–3245.

Vapnik, V. (2013). *The Nature of Statistical Learning Theory*. Springer science & business media.

Vazquez, S., Cancela, B., Barreira, N., Penedo, M. G., and Saez, M. (2010). On the automatic computation of the arterio-venous ratio in retinal images: Using minimal paths for the artery/vein classification. In *Digital Image Computing: Techniques and Applications (DICTA), International Conference on*, pages 599–604.

Vega, R., Sanchez-Ante, G., Falcon-Morales, L. E., Sossa, H., and Guevara, E. (2015). Retinal vessel extraction using lattice neural networks with dendritic processing. *Computers in Biology and Medicine*, 58:20–30.

Vincent, P., Larochelle, H., Bengio, Y., and Manzagol, P.-A. (2008). Extracting and composing robust features with denoising autoencoders. In *Proceedings of the 25th International Conference on Machine Learning (ICML)*, pages 1096–1103.

Wang, J. J., Mitchell, P., Cumming, R. G., and Lim, R. (1999). Cataract and age-related maculopathy: The Blue Mountains Eye Study. *Ophthalmic Epidemiology*, 6(4):317–326.

Wang, S., Yin, Y., Cao, G., Wei, B., Zheng, Y., and Yang, G. (2015). Hierarchical retinal blood vessel segmentation based on feature and ensemble learning. *Neurocomputing*, 149:708–717.

Wang, Y., Ji, G., Lin, P., and Trucco, E. (2013). Retinal vessel segmentation using multiwavelet kernels and multiscale hierarchical decomposition. *Pattern Recognition*, 46(8):2117–2133.

Welikala, R., Dehmeshki, J., Hoppe, A., Tah, V., Mann, S., Williamson, T., and Barman, S. (2014). Automated detection of proliferative diabetic retinopathy using a modified line operator and dual classification. *Computer Methods and Programs in Biomedicine*, 114(3):247–261.

Welikala, R., Foster, P., Whincup, P., Rudnicka, A., Owen, C., Strachan, D., and Barman, S. (2017). Automated arteriole and venule classification using deep learning for retinal images from the uk biobank cohort. *Computers in Biology and Medicine*, 90:23–32.

Wong, T. and Mitchell, P. (2007). The eye in hypertension. *The Lancet*, 369(9559):425–435.

Wong, T. Y., Kamineni, A., Klein, R., Sharrett, A. R., Klein, B. E., Siscovick, D. S., Cushman, M., and Duncan, B. B. (2006). Quantitative retinal venular caliber and risk of cardiovascular disease in older persons: the cardiovascular health study. *Archives of Internal Medicine*, 166(21):2388–2394.

Wong, T. Y., Klein, R., Couper, D. J., Cooper, L. S., Shahar, E., Hubbard, L. D., Wofford, M. R., and Sharrett, A. R. (2001a). Retinal microvascular abnormalities and incident stroke: the Atherosclerosis Risk in Communities Study. *The Lancet*, 358(9288):1134–1140.

Wong, T. Y., Klein, R., Klein, B. E., Tielsch, J. M., Hubbard, L., and Nieto, F. (2001b). Retinal microvascular abnormalities and their relationship with hypertension, cardiovascular disease, and mortality. *Survey of Ophthalmology*, 46(1):59–80.

Wong, T. Y., Klein, R., Sharrett, A. R., Couper, D. J., Klein, B. E., Liao, D.-P., Hubbard, L. D., Mosley, T. H., investigators, A., et al. (2002). Cerebral white matter lesions, retinopathy, and incident clinical stroke. *JAMA*, 288(1):67–74.

Wong, T. Y. and Mitchell, P. (2004). Hypertensive retinopathy. *New England Journal of Medicine*, 351(22):2310–2317.

Wong, T. Y., Mosley, T. H., Klein, R., Klein, B. E., Sharrett, A. R., Couper, D. J., and Hubbard, L. D. (2003). Retinal microvascular changes and mri signs of cerebral atrophy in healthy, middle-aged people. *Neurology*, 61(6):806–811.

Wong, W. L., Su, X., Li, X., Cheung, C. M. G., Klein, R., Cheng, C.-Y., and Wong, T. Y. (2014). Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis. *The Lancet Global Health*, 2(2):e106–e116.

Wu, A., Xu, Z., Gao, M., Buty, M., and Mollura, D. J. (2016). Deep vessel tracking: A generalized probabilistic approach via deep learning. In *IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, pages 1363–1367.

Xiao, Z., Adel, M., and Bourennane, S. (2013). Bayesian method with spatial constraint for retinal vessel segmentation. *Computational and Mathematical Methods in Medicine*, 2013.

Xu, X., Ding, W., Abràmoff, M. D., and Cao, R. (2017). An improved arteriovenous classification method for the early diagnostics of various diseases in retinal image. *Computer Methods and Programs in Biomedicine*, 141:3–9.

Yau, J. W., Rogers, S. L., and Wong, T. Y. (2012). Global prevalence and major risk factors of diabetic retinopathy. *Diabetes Care*, 35(3):556–564.

Yin, B., Li, H., Sheng, B., Hou, X., Chen, Y., Wu, W., Li, P., Shen, R., Bao, Y., and Jia, W. (2015). Vessel extraction from non-fluorescein fundus images using orientation-aware detector. *Medical Image Analysis*, 26(1):232–242.

Yin, Y., Adel, M., and Bourennane, S. (2012). Retinal vessel segmentation using a probabilistic tracking method. *Pattern Recognition*, 45(4):1235–1244.

Yin, Y., Adel, M., and Bourennane, S. (2013). Automatic segmentation and measurement of vasculature in retinal fundus images using probabilistic formulation. *Computational and Mathematical Methods in Medicine*, 2013.

You, X., Peng, Q., Yuan, Y., ming Cheung, Y., and Lei, J. (2011). Segmentation of retinal blood vessels using the radial projection and semi-supervised approach. *Pattern Recognition*, 44(10):2314–2324.

Youssef, D. and Solouma, N. H. (2012). Accurate detection of blood vessels improves the detection of exudates in color fundus images. *Computer Methods and Programs in Biomedicine*, 108(3):1052–1061.

Yu, H., Barriga, S., Agurto, C., Zamora, G., Bauman, W., and Soliz, P. (2012). Fast vessel segmentation in retinal images using multi-scale enhancement and second-order local entropy. In *SPIE Medical Imaging, Computer-Aided Diagnosis*, volume 8315, pages 8315–8315–12.

Zamperini, A., Giachetti, A., Trucco, E., and Chin, K. S. (2012). Effective features for artery-vein classification in digital fundus images. In *Computer-Based Medical Systems (CBMS), 25th International Symposium on*, pages 1–6.

Zhang, B., Zhang, L., Zhang, L., and Karray, F. (2010). Retinal vessel extraction by matched filter with first-order derivative of gaussian. *Computers in Biology and Medicine*, 40(4):438–445.

Zhang, J., Bekkers, E., Chen, D., Berendschot, T. T. J. M., Schouten, J., Pluim, J. P. W., Shi, Y., Dashtbozorg, B., and t. H. Romeny, B. M. (2018). Reconnection of interrupted curvilinear structures via cortically inspired completion for ophthalmologic images. *IEEE Transactions on Biomedical Engineering*, 65(5):1151–1165.

Zhang, J., Chen, Y., Bekkers, E., Wang, M., Dashtbozorg, B., and ter Haar Romeny, B. M. (2017). Retinal vessel delineation using a brain-inspired wavelet transform and random forest. *Pattern Recognition*, 69:107–123.

Zhang, J., Dashtbozorg, B., Bekkers, E., Pluim, J. P. W., Duits, R., and ter Haar Romeny, B. M. (2016). Robust retinal vessel segmentation via locally adaptive derivative frames in orientation scores. *IEEE Transactions on Medical Imaging*, 35(12):2631–2644.

Zhang, J., Li, H., Nie, Q., and Cheng, L. (2014). A retinal vessel boundary tracking method based on bayesian theory and multi-scale line detection. *Computerized Medical Imaging and Graphics*, 38(6):517–525.

Zhang, L., Fisher, M., and Wang, W. (2015). Retinal vessel segmentation using multi-scale textons derived from keypoints. *Computerized Medical Imaging and Graphics*, 45:47–56.

Zhang, T. Y. and Suen, C. Y. (1984). A fast parallel algorithm for thinning digital patterns. *Commun. ACM*, 27(3).

Zhao, H., Li, H., Maurer-Stroh, S., and Cheng, L. (2018). Synthesizing retinal and neuronal images with generative adversarial nets. *Medical Image Analysis*, 49:14–26.

Zhao, Y., Liu, Y., Wu, X., Harding, S. P., and Zheng, Y. (2015a). Retinal vessel segmentation: An efficient graph cut approach with retinex and local phase. *PloS one*, 10(4):e0122332.

Zhao, Y., Rada, L., Chen, K., Harding, S. P., and Zheng, Y. (2015b). Automated vessel segmentation using infinite perimeter active contour model with hybrid region information with application to retinal images. *IEEE Transactions on Medical Imaging*, 34(9):1797–1807.

Zhao, Y., Zhao, J., Yang, J., Liu, Y., Zhao, Y., Zheng, Y., Xia, L., and Wang, Y. (2017). Saliency driven vasculature segmentation with infinite perimeter active contour model. *Neurocomputing*, 259:201–209.

Zhao, Y. Q., Wang, X. H., Wang, X. F., and Shih, F. Y. (2014). Retinal vessels segmentation based on level set and region growing. *Pattern Recognition*, 47(7):2437–2446.

Zheng, J., Lu, P.-R., Xiang, D., Dai, Y.-K., Liu, Z.-B., Kuai, D.-J., Xue, H., and Yang, Y.-T. (2013). Retinal image graph-cut segmentation algorithm using multiscale hessian-enhancement-based nonlocal mean filter. *Computational and Mathematical Methods in Medicine*, 2013.

Zou, H. and Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2):301–320.

# List of Publications

## Journal Publications

1. Srinidhi, C. L., Aparna, P., & Rajan, J. (2017). **Recent advancements in retinal vessel segmentation.** *Journal of medical systems*, 41(4), 70.

2. Srinidhi, C. L., Aparna, P., & Rajan, J. (2018). **A visual attention guided unsupervised feature learning for robust vessel delineation in retinal images.** *Biomedical Signal Processing and Control*, 44, 110-126.

3. Srinidhi, C. L., Aparna, P., & Rajan, J. (2019). **Automated Method for Retinal Artery/Vein Separation via Graph Search Metaheuristic Approach.** in *IEEE Transactions on Image Processing*, vol. 28, no. 6, pp. 2705-2718, June 2019.

## Conference Proceedings

1 Srinidhi, C. L., Rath, P., & Sivaswamy, J. (2017). **A vessel keypoint detector for junction classification.** In Biomedical Imaging (ISBI), IEEE 14th International Symposium on (pp. 882-885). (**Oral Presentation - Best Paper Award Runner Up**)

# Brief Bio-Data

**Name: Chetan L Srinidhi**

**Address**

Chetan L Srinidhi

Research Scholar, Department of Electronics and Communication Engineering National Institute of Technology Karnataka, Surathkal, India-575025 Email: srinidhipy@gmail.com

**Qualification**

M.Tech - Communication Systems, RV College of Engineering, Bangalore.

B.E - Electronics and Communication Engineering, Visvesvaraya Technological University (VTU), Belgaum.

**Research Interests**

Medical Image Processing, Pattern Recognition, Machine Learning and Deep learning.