Medical imaging algorithm research for diagnosis of ocular diseases

Tan, Ngan Meng

2015

Tan, N. M. (2015). Medical imaging algorithm research for diagnosis of ocular diseases.
Doctoral thesis, Nanyang Technological University, Singapore.

<https://hdl.handle.net/10356/62191>

<https://doi.org/10.32657/10356/62191>

NANYANG TECHNOLOGICAL UNIVERSITY

Medical Imaging Algorithm Research For Diagnosis Of Ocular Diseases

Tan Ngan Meng

A thesis submitted to the Nanyang Technological University
in partial fulfillment of the requirement for the Degree of
Doctor of Philosophy

2015

Abstract

Color retinal fundus images provide visual documentation of the health of a person's retina. With the widespread adoption of higher quality medical imaging techniques and data, there are increasing demands for medical image-based computer-aided diagnosis (CAD) systems to manage large volumes of data, provide objective assessments for decision support and help in labour-intensive observer-driven tasks. This thesis focuses on the development of 2-dimensional color retinal image analysis algorithms for automated optic cup localization in glaucoma, the leading cause of irreversible blindness worldwide.

Traditionally, the optic cup is automatically segmented using image processing-based methods, often with many hand-crafted heuristics. With the incorporation of learning-based techniques, the accuracy of medical image-based CAD systems has improved significantly and are now widely accepted and adopted by medical practitioners. In this dissertation, three novel approaches for automatic localization of the optic cup in retinal fundus images are presented.

In the first work, a boundary-based cup detection approach using vessel kinks is presented. The key contribution in this work is its close modeling relationship with the clinical grading protocol to identify the optic cup, providing explicit visual evidence. Experimental results demonstrated that the novel use of vessel kinks as cup boundary key points guidance provides improved accuracy performance over existing retinal image-processing based strategies.

Although the use of vessel kinks is highly desirable and provides additional visual evidence, accurate detection and interpretation of these small vessel bends can, at times, be challenging. Instead, in the second work, a novel region-based unsupervised learning approach for automatic optic cup localization is proposed. This approach requires no training procedure, and utilizes domain knowledge and region-based features

in a similarity-based label propagation and refinement scheme to obtain an estimated cup region. The promising result suggests that learning-based techniques are capable of accurate automatic optic cup localization.

Recently, supervised superpixel-based cup localization has demonstrated superior performance. In the third work, a study on the limitations of this state-of-the-art classification framework is presented and an alternative generalized multi-scale approach is proposed, with improved stability and performance. This approach offers a stable and robust solution to reduce classification performance variations due to repeated random sampling of training samples. Furthermore, it integrates and unifies multiple superpixel resolutions for better boundary adherence. Extensive experimental results demonstrates the improved robustness and accuracy in optic cup localization against existing methods.

In summary, three approaches for optic cup localization are proposed. This thesis demonstrate that using vessel kinks as cup margin key points is highly desirable and provides additional visual evidence. The challenges in the detection of these key points are also discussed. Alternatively, a region-based unsupervised learning approach is presented. Experimentally, it was shown that in the absence of ground-truth labels, this approach is able to achieve higher or comparable accuracy to the boundary-based and existing retinal image processing-based approaches. Lastly, the limitations of the state-of-the-art supervised superpixel-based cup localization approach are studied and improved with a novel multi-scale multi-model framework, which offers stability and improved accuracy.

Acknowledgments

As many people may know, the road in PhD can be frustrating and exhausting at times. Having the company of family and friends with positive ideas and energies can often help in realigning and refreshing the fatigued mind. I am incredibly grateful and lucky to have many such wonderful companions around me. No amount of acknowledgement is enough for the advice and sacrifices of these colleagues and friends.

First, I must extend my deepest gratitude to my advisors, Professor Goh Wooi Boon and Dr. Jimmy Liu Jiang for their help, encouragement and discussions. I am immensely grateful for their time, advice and opportunities given to undertake this meaningful research. To Professor Goh, thank you for your humility, respect and trust throughout this mentorship. To Dr. Jimmy, thank you for your patience, support and encouragement to pursue this journey. I deeply treasure the experiences, opportunities and rapport. Heartfelt thanks also go to my fellow colleagues Dr. Damon Wong, Dr Cheng Jun and Ms Zhang Zhuo from the Institute for Infocomm Research for their generosity and helpfulness.

I am also particularly grateful to Dr. Xu Yanwu, Mr Jayachandran Maniyeri, Mr Yin Fengshou and close friends for their time, constant encouragements and inspirations in this journey. I sincerely appreciate your kindness and patience. Additionally, I would like to thank the PhD examiners for finding time to read and examine this report. Your reviews have been insightful.

My family has been my pillar of strength and happiness. I am indebted to my wife, who has been ever supportive throughout this PhD experience. Thank you for reminding me to always learn, be happy and to know and understand myself, so that I can love and understand others. To my parents, I thank them for showing me how to love, support and treasure my loved ones unconditionally. To the little ones, thank you for coming

into our lives. Your laughs, innocence and inquisitiveness never fail to brighten up our days. Without all your support and faith, it would not be possible for me to complete this.

Contents

Abstract	i
Acknowledgments	iii
List of Figures	ix
List of Tables	xiii
1 Introduction	1
1.1 Motivation	2
1.2 Scope and Objectives	3
1.3 Contributions of Thesis	4
1.4 Thesis Organization	5
2 Medical Background and Related Works	6
2.1 The Human Eye	6
2.1.1 Structure of the Human Eye	6
2.1.2 Image Acquisition of the Retina	9
2.2 A Medical Overview of Ocular Diseases	19
2.2.1 Glaucoma	19
2.3 Related Works	23
2.3.1 Optic Disc Localization and Segmentation	23
2.3.2 Optic Cup Segmentation	27
2.4 Evaluation Dataset	35
2.4.1 ORIGA ^{-light}	35
2.5 Evaluation Criteria	37
2.5.1 ORIGA ^{-light} Evaluation Criteria	37
2.5.2 Superpixel Classification Evaluation Criteria	38

3	The Detection and Use of Vessel Kinking to Improve Optic Cup Detection in Digital Retinal Fundus Images	40
3.1	Introduction	40
3.1.1	What are Vessel Kinks?	41
3.1.2	Prior Works	41
3.2	Methodology	42
3.2.1	Vessel Trajectory Detection	43
3.2.2	Kinks Detection	49
3.2.3	Optic Cup Feature Point Integration	50
3.2.4	Optic Cup Boundary Detection	52
3.3	Experiments	56
3.3.1	Comparison of methods	56
3.4	Model Training and Computation Speed	58
3.4.1	SVM Model Training	58
3.4.2	Computation Speed	58
3.5	Discussion	59
3.6	Summary	61
4	Domain Prior Based Superpixel Propagation for Unsupervised Optic Cup Localization	62
4.1	Introduction	62
4.2	Methodology and Contributions	63
4.2.1	Superpixel Segmentation	63
4.2.2	Multi-Scale Blood Vessel Extraction	65
4.2.3	Optic Cup Origin Prior - Central Retinal Vessel	65
4.2.4	Superpixel Feature Representation	67
4.2.5	Label Initialization and Propagation	68
4.2.6	Label Refinement	68
4.2.7	Region-based Ellipse Fitting	70
4.3	Experiments	70
4.3.1	Comparison to Prior Arts	70

4.3.2	Comparison of methods within proposed framework	71
4.4	Parameter Selection and Other Considerations	73
4.4.1	K-means Parameter Selection	73
4.4.2	Computation Speed	73
4.4.3	Alternative Features	74
4.5	Summary	74
5	Multi-Scale Superpixel Classification for Supervised Optic Cup Local-ization	76
5.1	Introduction	76
5.2	Methodology and Contributions	77
5.2.1	Pre-processing	79
5.2.2	Multi-Scale Superpixel Feature Representation	79
5.2.3	Multi-Scale Superpixel Classification	80
5.2.4	Model Integration Strategy	81
5.2.5	Testing Phase	82
5.3	Experiments	83
5.3.1	Experimental Setup	83
5.3.2	Evaluation Criteria	84
5.3.3	Experiments to validate motivations	85
5.3.4	Superpixel Classification Performance	88
5.3.5	Comparison with existing methods.	90
5.4	Discussion on Parameter Settings	99
5.4.1	Sparsity Regularizer λ	99
5.4.2	Computation Speed	102
5.5	Summary	102
6	Conclusion and Future Works	103
6.1	Principal Results	103
6.2	Future Research Directions	106
	References	109

Appendix A List of Publications	127
Appendix B List of Abbreviations	128
Appendix C Pseudo Codes	129

List of Figures

2.1	The Human Eye	7
2.2	Anatomy and schematic diagram of the human eye.	8
2.3	The first indirect ophthalmoscope. Image from [5].	10
2.4	Retinal image capture using the fundus camera.	11
2.5	Illustrations of the angle of coverage, and field of view capture in a retinal fundus camera.	12
2.6	Retinal nerve fibre layer (RNFL) analysis using the GDxVCC.	14
2.7	Optic nerve imaging using HRT.	16
2.8	Optic nerve imaging using OCT.	17
2.9	Anatomy of the Optic Nerve Head	20
2.10	Progression of Optic Cup Enlargement Changes in Glaucoma	20
2.11	A Contrast of Vision	21
2.12	Cup-to-Disc Ratio	22
2.13	General Appearance of the Optic Disc in a Retinal Fundus Image	24
2.14	Scaling disparity of optic cup.	27
2.15	Extent of optic cup excavation and its effect in pallor appearance variation.	28
2.16	Variation of Pallor Intensity.	29
2.17	Viewpoint perspective.	29
2.18	Challenges in Occlusion.	30
2.19	Close-up Appearance of the Optic Cup	31
2.20	Sample Images and Ground-Truth in ORIGA ^{-light} dataset	36
3.1	(<i>Left</i>) Optic disc showing kinking of vessels (black arrows). (<i>Right</i>) Illustration of vessel kinking.	41

3.2	Schematic diagram of the proposed system. The main contributions are indicated in the colored boxes.	43
3.3	Different settings for a and ϵ for the enhancement of thin vessels. (<i>Top-Left</i>) RGB Image, (<i>Top-Centre</i>) $a = 20$ $\epsilon = 4$, (<i>Top-Right</i>) $a = 1$ $\epsilon = 4$, (<i>Bottom-Left</i>) $a = 4$ $\epsilon = 100$, (<i>Bottom-Centre</i>) $a = 4$ $\epsilon = 0.01$, (<i>Bottom-Right</i>) $a = 5$ $\epsilon = 1$, The set of parameters leading to the transformed image in the lower right corner was used. For illustration purposes the intensity ranges for each image was stretched to $[0, 255]$	46
3.4	Three stage patch generation for vessel edge detection	49
3.5	Simulated vessel edge(red), approximated straight line segment(blue), and kink point(yellow) are detected at (<i>Left</i>) $\gamma = 60^\circ$, (<i>Middle</i>) $\gamma = 100^\circ$ and (<i>Right</i>) $\gamma = 159^\circ$	50
3.6	(<i>Top-Left</i>) Original Image. (<i>Top-Right</i>) Small vessel enhancement using wavelet transform. (<i>Bottom-Right</i>) Detected vessels (in blue), and vessel kinks points (in green) along vessels; (<i>Bottom-Left</i>) Optic cup feature points (in blue) comprising kinks and optic cup color extrema	51
3.7	Comparison of MVEE and B2AC on set of generated typical sparse key points (<i>Left</i>) Minimal difference when points are well-distributed and representative. (<i>Centre</i>) and (<i>Right</i>) show an over-fit by B2AC when the point distribution is not representative and a more robust recovery for MVEE.	55
3.8	Sample segmentation results (in blue) of different approaches and ground-truth(in red). <i>From Top</i> : IET [51], JET [54], LS [131], KinkB2AC, Proposed Method - KinkMVEE. [Best viewed in color]	57
3.9	Images with low contrast, arising from media opacity. [Best viewed in color]	60
3.10	Absence of physical blood vessel landmarks in the temporal region. [Best viewed in color]	60
4.1	Flowchart of the proposed methodology.	64
4.2	Region segments obtained using superpixel segmentation. Each region segment is compact and uniform, while adhering to the image boundaries.	64

4.3	<i>Left:</i> Green Channel of the input image. <i>Middle:</i> Extracted Blood Vessel. <i>Right:</i> Generated blood vessel mask.	66
4.4	Illustration of optic cup origin with neighborhood prior	67
4.5	k -Means clustered region segments.	69
4.6	Final segmentation (blue) against manual ground-truth segmentation (red).	72
4.7	Cup Detection with different k values for clustering.	73
5.1	Flowchart of the proposed learning framework.	78
5.2	Cup detection errors with different number of superpixels. Figure and caption is reproduced from with permission from [138].	80
5.3	Overview of the major steps in this approach. <i>From Left :</i> Contrast Enhancement; Multi-scale Superpixel segmentation, vessel removal and feature classification; Multiple base model training and selection; Model integration and Ellipse-fitting. [Best viewed in color]	83
5.4	Effects of proposed contrast normalization. (<i>Top</i>) Original image, (<i>Bot- tom</i>) Contrast normalized images. Finer details of the tissues becomes clearer and distinct.	87
5.5	Boxplot illustrating variance in localization accuracy.	88
5.6	Rank-sorted performance plots for trained models. <i>Top :</i> Models sorted by prediction accuracy. <i>Bottom :</i> Models sorted by training m_1 error. Trained models of Scale 1024 are shown here. [Best viewed in color]	89
5.7	Superpixel classification performance comparison in terms of ROC curves, at different scales ($S_1 = 512$ in blue, $S_2 = 1024$ in red, $S_3 = 2048$ in green. Mean performance of single model.	91
5.8	Superpixel classification performance comparison in terms of ROC curves, at different scales ($S_1 = 512$ in blue, $S_2 = 1024$ in red, $S_3 = 2048$ in green. Performance of Majority Voting.	92
5.9	Superpixel classification performance comparison in terms of ROC curves, at different scales ($S_1 = 512$ in blue, $S_2 = 1024$ in red, $S_3 = 2048$ in green. Performance of the proposed model integration.	93

5.10	Illustration of localized cup boundary at superpixel scale $S_1 = 512$ (<i>Top-Left</i>), $S_2 = 1024$ (<i>Top-Right</i>), $S_3 = 2048$ (<i>Bottom-Left</i>), Proposed integration approach (<i>Bottom-Right</i>). The ground-truth boundary is shown in blue. [Best viewed in color]	94
5.11	Sample Results. <i>From Top</i> : Original Image, Contrast Normalized, $S_1 = 512$, $S_2 = 1024$, $S_3 = 2048$, Majority Voting, Proposed Method. [Best viewed in color]	97
5.12	Selected SVM base Models using different Sparsity constraints. The 3 most frequently selected models are shown for illustration. [Best viewed in color]	101

List of Tables

2.1	Overview comparison of retinal acquisition devices.	18
2.2	Categorization of Existing Optic Disc Methods	26
2.3	Taxonomy of Existing Optic Cup Segmentation Methods	34
2.4	Clinical Grading Information in ORIGA ^{-light} dataset	37
3.1	Selected Parameters for Gabor Wavelet Implementation	45
3.2	List Of Features For Vessel Edge Classification	47
3.3	Cup localization performance comparison	56
4.1	Features used	68
4.2	Cup localization performance comparison	71
4.3	Features comparison Cup localization performance	74
5.1	Performance comparison between proposed contrast normalization and intra-image learning	86
5.2	Performance of Superpixel Classification	90
5.3	Cup localization performance comparison	98
5.4	Influence of Sparsity Measure	99

Chapter 1

Introduction

Computer-aided Diagnosis (CAD) utilizes computer technology and analysis to effectively handle large volumes of data used in diagnostic process [115]. As high quality medical imaging techniques and modalities gain widespread adoption, this has facilitated a greater role for CAD in medical imaging for their processing and analysis capabilities.

Medical image-based CAD is an interdisciplinary research area used in radiology and medical imaging for the interpretation of medical images. It has shown to be effective as radiologists use the CAD output as a "second opinion" and this has helped improve the diagnostic accuracy, consistency and productivity of the final decision made by clinicians [28, 41]. Successful examples includes CAD for early detection of breast cancer using digital mammogram [28, 41], detection of colonic polyps using computed tomographic (CT) colonography [28, 94], and detection of lung nodules for chest radiography [28, 127]. Examples of commercial companies using medical image-based CAD includes Hologic¹, and iCAD².

While medical image-based CAD has shown to be promising, the approaches and methods used vary widely. Often these would depend on factors such as the required specific application and imaging modality. The presence of imaging artifacts such as noise and motion, which may be introduced during data acquisition, can also have a significant penalty on the performance of the algorithms. In comparison to generic methods customized to suit the domain application, methods that integrate explicit application-specific domain knowledge, such as anatomical or image modality prior knowledge, usual-

¹ <http://www.hologic.com/>

² <http://www.icadmed.com/>

ly have an edge in terms of accuracy, precision and computational efficiency. Nonetheless, the selection of such an appropriate approach to the task is often a challenging issue [95].

In addition, the task of objective evaluation and comparison of algorithms in medical image analysis can be difficult. Some of the reasons include a lack of definitive gold standard (caused by inter-and intra-observer variability) and tedious data collection [21].

Color retinal fundus³ images provides visual documentation of the health of a person's retina, and is often the most fundamental image used by eye specialists to assess retinal abnormalities. A reliable ocular-image CAD system, which can provide objective assessments, measurements and reduce labor-intensive observer-driven tasks for ophthalmologists, has always been one of the ultimate aims in this area of research. Although there are a number of imaging technologies, such as Heidelberg Retinal Tomograph (HRT) and Optical Coherence Tomography (OCT), which can be used to view the retina, digital fundus images still remains highly relevant for establishing a baseline. Furthermore, retinal fundus cameras are widely available and prevalent in primary healthcare, thus, are highly suitable for disease screening.

This thesis focuses on the design, implementation and evaluation of image analysis algorithms using color retinal fundus images taken from patients suffering varying degrees of glaucoma, a major ocular disease. In particular, it focuses on methods to identify the damages caused by glaucoma on the nerve fibres of the optic nerve. For a fair performance evaluation, a large set of retinal fundus images from the online glaucoma analysis database, ORIGA^{-light} [147] will be used. This database consists of images from a population-based cross-sectional study with graded glaucoma pathological signs.

1.1 Motivation

Globally, glaucoma is reported to be the second leading cause of blindness, and estimated to visually impair 60 million people by 2010 and 80 million by 2020 [99]. It is the most prevalent of eye diseases in the United States and around the world. There are approximately 2.5 million diagnosed cases of glaucoma in the United States and 60 million worldwide; however, this figure is significantly underreported due to the slow,

³ Fundus: refers to the bottom on an organ, in the context of this thesis, the back of the eye, where the retina, optic nerve and macula resides

asymptomatic progression of the disease, resulting in low disease detection until vision lost sets in. In Singapore, the Tanjong Pagar Study[37] and the Singapore Malay Eye Study[35, 117] found that a 3% prevalence of glaucoma amongst Singaporean adults of 40 years and above. The study also reports that more than 90% of the patients are unaware of their conditions[35, 117]. Similarly, under-diagnosis rate ranges from 50% in developed countries like North America to as high as 90% in developing countries like India.[17]

Vision loss is often a late symptom of glaucoma. When glaucoma patients are referred to the ophthalmologists, it was found that severe permanent visual loss has occurred, and often, the disease is then managed using expensive surgery. At the same time, glaucoma is a social burden, with heavy direct and in-direct costs, since visual loss reduces an individual's capacity to work and live independently.

Although glaucomatous damage is irreversible, early diagnosis and treatment of glaucoma has shown to reduce or halt the risk of visual field loss [46]. Hence, early detection is critical to prevent blindness and aid in glaucoma management. A medical image-based CAD system for glaucoma is a technological-means of assisting this process, especially in data analysis, management and objective assessments for decision support.

The motivation of this thesis is to explore methods for automatic retinal image analysis for glaucoma in 2-dimensional fundus images. As glaucoma causes optic nerve damage and increases the size of the optic cup, it is important to localize and measure the optic cup. Although it is essential and may seem simple, localization of the cup is a highly challenging task, especially for inexperienced clinicians. Furthermore, automatic methods to localize the optic cup are much less reported, perhaps due to the entry barrier of basic fundus image interpretation training. This meaningful task of automatic optic cup localization is selected as the focus of this research.

1.2 Scope and Objectives

In this thesis, a review on the current approaches for automatic optic nerve segmentation is presented, and new approaches to determine the optic nerve damage from glaucoma in retinal images are proposed. Specifically, a study on the background of this deadly disease, the recommended practices of diagnosis, challenges, and the existing automated

methods will be surveyed. Inspired by recent successful machine learning techniques adapted on object recognition, new learning-based approaches for optic cup localization approaches are investigated and proposed. The ORIGA^{-light} database [147] will be used to evaluate the accuracy of the proposed methods.

1.3 Contributions of Thesis

This thesis proposes several novel approaches for optic cup localization. More specifically, it exploits and adopts popular and proven machine learning techniques to localize the optic cup. The main contributions of this thesis are as follows:

First, an automated vessel kink detection method, which is modeled after a frequently-used clinical clue was developed to improve optic cup boundary detection. Using vessel kinks and color information as boundary key points, a minimum enclosing ellipsoid method is proposed to represent the detected optic cup boundary. Initial experimental evaluation demonstrates encouraging improved results over existing methods.

Second, a novel unsupervised learning region-based approach for optic cup localization was proposed. Here, image pixels are clustered into boundary-adhering superpixel regions, and a set of features is extracted from each region for feature representation. Next, region labels are initialized using domain priors and propagated using a similarity measure. Finally, a label refinement scheme is performed to return the final localized optic cup. Unlike previous reported approaches, this method requires no training samples, and is not limited by the absence of vessel kinks. Furthermore, it was experimentally demonstrated that such learning-based approach is also able to achieve competitive performance.

Third, a study of the limitations of the current state-of-the-art supervised superpixel-based learning for optic cup localization, was done and a novel optimal model selection approach was proposed to overcome some of these limitations. A major shortcoming in the previous approaches, is the variations in classification performances due to random sampling of training samples. Experimental results demonstrated that the proposed approach, which selects and integrates multiple models and multiple scales, yielded improved performance stability and increased accuracy. Experimental results on the ORIGA database demonstrates a 7.12% improvement. In addition, against the intra-image [138] method, the proposed approach does not assume a restricted CDR range of 0.2-0.9.

1.4 Thesis Organization

This report has 6 chapters and is organized in the following manner:

Chapter 2 introduces an overview of the medical background of glaucoma, imaging aspects, followed by a review of the current state-of-the-art automatic localization and segmentation of the optic disc and cup. The image dataset and criteria used for the evaluations of the proposed approaches are also presented.

In Chapter 3, a system which automatically detect vessel kink points and integrates with optic cup color information for boundary-based optic cup key point detection is presented. Clinically, these visual cues are often used, when available, to estimate the cup margin. Tested on a large dataset, the effectiveness and the use of vessel kinks, and the performance of such boundary-based approach is discussed.

In Chapter 4, the first unsupervised learning approach of optic cup localization is proposed. Superpixels are used to cluster and partition an image into boundary-adhering segments, for features extraction. Domain priors are then used to initialize the unlabeled segments for similarity-based label propagation. Finally a label refinement is performed. Detailed experiments are carried out to validate each step in the framework, demonstrating its effectiveness.

Chapter 5 introduces the multi-scale supervised approach for optic cup detection. A study on the limitations of the current state-of-the-art supervised superpixel-based learning for optic cup localization is discussed, and a generalized multi-scale multi-model approach, which overcomes most of the limitations, is proposed. Experimental results demonstrate the improved stability and performance from this framework.

Chapter 6 summarizes the contributions in this research and the areas for future research works to be explored.

Appendix A contains a list of publication works resulting from this thesis.

Appendix B contains a list of abbreviations used in this thesis.

Appendix C contains the pseudo codes for the algorithms proposed in this thesis.

Chapter 2

Medical Background and Related Works

This chapter will cover two main sections, the medical background of the ocular disease, and the existing computerized approaches in literature. An introduction on the structure of the human eye, the image acquisition of the retina and medical overview of the ocular disease will be covered. The clinical retinal image assessment protocols will also be described. In the second half of this chapter, a review of the influential and exciting works that are related to this research along with the available dataset will be detailed.

2.1 The Human Eye

Aristotle (384 BC - 322 BC) classified the five sensory organs of the body: sight, smell, taste, touch, and hearing. The human eye, the organ of vision shown in Fig. 2.1, provides the capability to see color, detect motion, identify shapes, gauge distance and speed, judge the size of faraway objects, see in three dimensions even though images fall on the retina in two[43]. In the brain itself, neurons devoted to visual processing number in the hundreds of millions and take up about 30 percent of the cortex, as compared with 8 percent for touch and just 3 percent for hearing[43].

2.1.1 Structure of the Human Eye

The eye is a complex structure, with an approximate length of 24 to 25mm, where light passes through the cornea and a transparent crystalline lens, which assists in focusing

Figure 2.1: The Human Eye - The figure shows the anatomy of the human eye, taken from University of Maryland Medical Center (UMMC).

light, onto the retina through the pupil. The amount of light which enters the eye is controlled by the iris, which has the ability to enlarge or contract, while the lens converges the incoming light rays to a sharp focusing point onto the retina by lengthening and shortening its width. The retina, situated at the back of the eye, is approximately 0.5 mm thick [44] (Fig. 2.2a). In the center of the retina is the optic nerve, an oval white area measuring about 2mm (height) x 1.5mm (width) across. At approximately 17° (4.5-5 mm), or 2.5 disc diameters away from the optic nerve, lies a blood vessel-free reddish region, known as the fovea, Fig. 2.2b. The centre of the fovea is also called the macula. The entire retina is a circular area of between 30 and 40 mm in diameter [98]. It is a multi-layered tissue, and is covered with two types of light-sensitive photoreceptor cells, rods and cones, to convert light rays into electrical impulses. As shown in Fig. 2.2c, a radial cross-section of the retina shows that the ganglion cells (the output neurons of the retina) occupy the innermost of the retina nearest to the lens and front of the eye, and the rods and cones photosensors lie at the outermost in the retina against the pigment epithelium and choroid [44].

(a) Sagittal view of the human eye with an schematic enlargement of the retina. Image from [44]

(b) Example of a retina image.

(c) Schema of the layers of the developing retina around 5 months gestation (Modified from Odgen,1989). The rods and cones are denoted as R and C respectively. Image from [44].

Figure 2.2: Anatomy and schematic diagram of the human eye.

The rods are most sensitive to light and are responsible for vision in low light and peripheral vision. The cones which are concentrated in the fovea, handle color vision and details. The retina collects the signals and sends the impulses, through a million nerve fibers that are carried by the optic nerve, to the visual cortex of the brain. These impulses are then interpreted and processed by the brain. The optic nerve on the retina is also called the blind spot as it is insensitive to light. Often, the fovea, optic nerve, retina and retinal vasculature are referred to as the ocular fundus structures.

2.1.2 Image Acquisition of the Retina

2.1.2.1 Fundus image capture

A color retinal fundus camera is widely used to photograph the retina. The fundus camera is a complex optical imaging device with a low power microscope and an attached camera, Fig. 2.4a. Its optical operation principles is designed based on the indirect ophthalmoscope, Fig. 2.3. The general working principle described below is based on the overview provided in [90].

Light is first generated from either viewing lamp or the electronic flash, and is passed through a set of filters, mirrors and a series of lenses for focusing. A mask on the uppermost lens is then used to shape the light into a doughnut. Based on the Gullstrand principle, the ring of light is projected on the cornea, through the pupil. The resulting retinal image then leaves the cornea through the un-illuminated portion of the doughnut, Fig. 2.4b. This space within the ring allows a separation of both the incoming and outgoing illumination. The outgoing light continues through the central aperture of the mirror, through the astigmatic correction device and the diopter compensation lenses, and then back to the single lens reflex camera system.

Clinically, this camera is used by ophthalmologists and trained medical professionals to monitor and discover evidence of ocular abnormalities for immediate feedback, diagnosis and treatment of retina diseases. The fundus photographs are then kept as visual records to document the ophthalmoscopic appearance of a patient's retina.

In the market, fundus cameras are described by their angle of coverage. This is derived from the optical angle of acceptance of the fundus camera lens and can range between 20° , to 140° . An angle of 30° , considered the normal angle of view, creates a film image

Figure 2.3: The first indirect ophthalmoscope. Image from [5].

with a magnification of 2.5x [104]. Wide-angle fundus cameras capture images between 45° and 140° and provide proportionately less retinal magnification. In this research, the images used are mostly taken using a 45° digital retinal camera, Fig. 2.4c, (model CR-DGi with a 10D SLR digital camera backing; Canon, Tokyo, Japan).

Capturing the retinal fundus photographs can be performed either through a dilated or non-dilated pupil. The main advantage of dilation is to allow better view and image capture of the retina. However, even with dilation, the quality of the fundus image can still be affected by additional difficulties such as the media opacity due to cataracts. Non-mydratic¹ retinal fundus cameras allows digital photographs of the eye to be captured through a small pupil size (in between 2.0- to 4.0mm) without the need and discomfort of pupil dilation. However, there may be a trade-off in the image acquisition quality.

Another consideration in fundus photography acquisition, is the field of capture within the retina. The Early Treatment Diabetic Retinopathy Study (ETDRS) has established a standard seven-field photographic protocol used as the gold standard in fundus photography, Fig. 2.5b (where Field 1 = optic disc; Field 2 = macula; Field 3 = temporal to macula; Field 4 = superior temporal; Field 5 = inferotemporal; Field 6 = superior nasal; Field 7 = inferior nasal). For the purposes of assessing the optic disc, a Field 1 capture is usually used.

¹Without causing dilation to the pupil.

(a) A simple schematic of the fundus camera. Image from [75].

(b) "Donut" of light. Image from [90].

(c) Canon CR-DGi. A Non-Mydriatic Retinal Camera allows digital images of the eye to be captured through a small pupil size without pupil dilation. Image from *Retinal Physician*[96].

Figure 2.4: Retinal image capture using the fundus camera.

(a) Depiction of different angle of coverage on the retina. Image from [90].

(b) Seven standard fields of the ETDRS protocol on a right eye. The optic nerve is highlighted in yellow.

Figure 2.5: Illustrations of the angle of coverage, and field of view capture in a retinal fundus camera.

2.1.2.2 Other Imaging Devices

Over the past 15 years, other imaging technologies have emerged to detect early glaucomatous optic nerve damage and monitor for changes. Examples of such devices includes nerve fiber layer polarimetry (GDx nerve fiber analyzer), Heidelberg Retinal Tomography (HRT) and Optical Coherence Tomography (OCT). A brief introduction on these technologies will be described below.

GDx nerve fiber analyzer, Carl Zeiss Meditec. The GDx nerve fiber analyzer is a scanning-laser polarimeter developed by Laser Diagnostic Technologies, which uses the reflected polarised light to measure the thickness of the retinal nerve fibre layer. The retinal nerve fibre layer (RNFL) is innermost retinal layer closest to the vitreous, containing axons and ganglion cells which converge to the fibres of the optic nerve [1]. As studies have shown that the structural loss of ganglion cells, which makes up the RNFL, has an association with functional visual field defect [114], the use of the GDx nerve fiber analyzer offers a means to measure and track RNFL thickness.

The scanning-laser polarimetry technique relies upon the birefringent qualities of the RNFL, where the phase shift in the polarization of the light (retardation) passing through the nerve fibers tissues are measured (see Fig. 2.6b for depiction). The degree of change of polarization is proportional to the thickness of the RNFL tissues, where higher retardation denotes thicker tissues[25]. A high-resolution image of 256 by 256 pixels of the RNFL topographical map is created, where each pixel represents a measure of the retardation of the laser scan at its location [6]. For each test, three serial scans are obtained, and these measurements are compared against a normative database of sex, race, and age matched normals [2]. Fig. 2.6c shows an example of the GDx report printout of the nerve fibre analysis thickness.

Confocal Scanning Laser Ophthalmoscopy (HRT), Heidelberg Retinal Tomograph. The principle of confocal microscopy was first introduced and patented by Marvin Minsky in 1957 [76] to perform precision microscopy. The key element in the confocal approach is the use of spatial filtering techniques to eliminate out-of-focus light or in specimens whose thickness exceeds the immediate plane of focus [4]. The Heidelberg Retinal Tomograph (HRT) is a confocal laser scanning system which uses a rapid scanning laser diode to obtain multiple cross-section images to produce a three-dimensional high-resolution topographic image of the optic nerve 2.7a.

(a) The GDxVCC. Image from Carl Zeiss Meditec GDxVCC Brochure. (b) Working principles of the GDx scanning-laser polarimeter. Image from Carl Zeiss Meditec GDxVCC Brochure.

(c) GDx Printout Patient with Early POAG. Image from [6].

Figure 2.6: Retinal nerve fibre layer (RNFL) analysis using the GDxVCC.

A 670-nm diode laser is used to emit a beam, which is redirected by two oscillating mirrors to focus on the x-axis and y-axis along a plane of focus perpendicular to the optic nerve. A $15^\circ \times 15^\circ$, two-dimensional reflected image (384x384 image resolution) from this plane is then captured. Subsequent reflected images are obtained from different image planes of an equidistance longitudinal resolution depth. For example, using a 4mm depth, a maximum of 64 image planes will be captured. These cross-sections are then combined and reconstructed to form a three-dimension contour map of the optic nerve surface. For analysis, the software requires the operator to manually define the outline of the optic disc. This allows a reference plane to be established for the computer to perform an automatic analysis computing the stereometric parameters, classification of the eye and depth profile comparison against the Moorfields Regression Analysis (MRA) data. The OU report from the HRT machine depicting the various measurement results is shown in Fig. 2.7b.

Optical Coherence Tomography (OCT). Using a technique known as low-coherence interferometry, to establish the echo time delay and magnitude of backscattered light reflected off an object-of-interest, the Optical Coherence Tomography (OCT) is a non-invasive imaging technique used to scan the structured layers of the retina [105].

The working principles of OCT is similar to ultrasound, where light reflections, instead of sound, is being used to provide the cross-sectional images, Fig. 2.8a. Using a beam splitter, the optical beam (from the source) is redirected into two arms: a reference arm (usually a mirror), and a sample arm where light is diffused and reflected off the intended tissue structures. Light from both the reference and sample arms travels back to the beam splitter and recombines to form an interference pattern, which is sensed by a photodetector, if the light from both arms have similar optical distance [105]. In Time-Domain OCT (TD-OCT), a reflectivity profile of the tissue sample, called an A-scan, is obtained by scanning the position of the oscillating mirror in the reference arm. Using interferometry to record the optical path length of received photons, the OCT is able to reject background signal while accepting light directly reflected from surfaces of interest. Each pixel in the A-scan represents the reflection intensity at that position, and a cross-sectional tomograph (B-scan) can be produced by laterally combining a series of these axial depth scans (A-scan).

(a) Schematic diagram of a confocal scanning laser system in the HRT II. Image from Heidelberg Engineering Product Literature [3].

(b) A sample of the HRT Glaucoma Module v.3 Report. Image from Heidelberg Engineering Product Literature [3].

Figure 2.7: Optic nerve imaging using HRT.

(a) Schematic diagram of the time-domain OCT(TD-OCT), spectral domain OCT(SD-OCT), swept-source OCT(SS-OCT) and the adaptive optic system that can be added to the SD-OCT system (AO-OCT). Image from [105].

(b) A high resolution SD-OCT scan revealing the underlying tissue structures of the retina. Image from [39].

Figure 2.8: Optic nerve imaging using OCT.

Recent advances in OCT have led to the higher speed, higher resolution Spectral Domain OCT (SD-OCT), which encodes the reflected light in the Fourier domain, to the latest Swept-Source OCT(SS-OCT), which encodes the spectral components in time, providing greater details. An example of a SD-OCT scan is shown in Fig. 2.8b.

Role of Retina Fundus Imaging. Although, the above mentioned newer technologies provides alternative measurement of optic disc and retinal nerve fiber layer structure, they cannot quantify other irregularities of the optic nerve head such as disc hemorrhage, optic disc pallor and peripapillary atrophy [56]. The GDx provides good retinal nerve fiber layer thickness measurement, however, its measurements have found to be affected by media opacity and other retina conditions. In HRT, although it provides a three-dimensional topographic view of the optic nerve, the perspective is obtained from a reconstruction of the cross-sectional image(i.e. the image is not a true view of the optic nerve, but an approximate representation). Furthermore, manual outlining of the optic disc boundary is required before the software analysis is performed. This limitation also results in difficulties to compare quantitative measurements between subsequent follow-up. Lately, the versatility of OCT, for being able to provide both RNFL and optic nerve data, has been gaining popularity. However, it have been reported that its results changes with each new generation device and are not backwards compatible [61]. In

Table 2.1: Overview comparison of retinal acquisition devices.

Devices	Advantages	Limitations
Fundus Camera	Stable technology. True color of retina. Low Cost and widespread use.	Does not capture layer thickness or information.
GDx nerve fiber analyzer	Provides thickness of nerve fibre layer.	Measurements can be affected by existing eye conditions. Pseudo Color image.
Heidelberg Retinal Tomography (HRT)	Provides topographic (layer thickness) image of the optic nerve.	Topographic image is an approximate representation. Pseudo Color image.
Optical Coherence Tomography (OCT)	Multiple cross-sectional (A-scans) of retina. Cross-sectional tomograph (B-scan). Sub-layer retinal tissue thickness.	Results change with each new generation device and are not backwards compatible. Pseudo Color image. High Cost.

addition, OCT machines are unable to provide progression analysis as there is no means to measure the same tissue during follow-up exams.

Retinal fundus photographs, on the other hand, provides the advantages of full color which helps to distinguish between cupping and pallor, has a stable technology and widespread usage for community-level screening [6]. Furthermore, the record of a retinal photograph will not go out of date, and is often used as a baseline for clinical evaluation and comparison. This makes it the ideal media to use for glaucoma disease screening purposes. A summarized comparison table of these devices is shown in Table. 2.1.

2.2 A Medical Overview of Ocular Diseases

Ocular diseases are diseases and disorders of the eye. According to the World Health Organization(WHO), there are 285 million visually impaired people worldwide and 80% of all visual impairment can be avoided or cured [88]. In 2010, a study by the WHO found that the top causes of blindness are cataracts, glaucoma and age-related macular degeneration [91] and, the leading cause of irreversible blindness is glaucoma. This section will cover an overview of the medical aspects of glaucoma and the structural changes it brings to the retina as it progresses.

2.2.1 Glaucoma

Glaucoma is a serious and lifelong eye-threatening disease, and the leading cause of irreversible blindness which affects the optic nerve head [92, 99]. The optic nerve head is made up of bundles of ganglion cell axons which leaves the retina and enter the front of the optic nerve (also known as optic disc). The optic cup is a pale cavity-like depression in the optic nerve head, and the area between the optic nerve and cup is known as the rim, Figure 2.9.

Glaucoma is a progressive optic neuropathy that structurally damages retinal ganglion cells and other mechanisms in the visual pathway. As this disease progresses, the retinal ganglion cells in the optic nerve head gets damaged and eventually results in ganglion cell death. This leads to a collapse and enlarged excavation of the optic cup within the optic nerve head, Fig. 2.10.

Figure 2.9: Anatomy of the Optic Nerve Head - Clinical features of the optic nerve.

Figure 2.10: Progression of Optic Cup Enlargement Changes in Glaucoma - (Left) A normal optic nerve head. (Centre) Early glaucoma optic nerve head. (Right) Advanced glaucoma optic nerve head. Image from Bay EyesCataract&LaserCenter [30].

Unfortunately, glaucoma does not have any early warning signs [38]. Patients with early glaucoma do not have visual symptoms. They usually first develop a “tunneling vision”, the lost of their peripheral vision, and eventually central vision, (see Fig. 2.11). Studies have also reported that 90% of those affected are unaware of their condition [37, 108] and caring for glaucoma can be costly. Without treatment, a substantial number of patients will experience significant progression of the disease leading to total permanent blindness within a few years. Several studies demonstrated that the estimated progression of optic nerve fiber loss in glaucoma can be in the order of 9-63% over a five-year period [20, 60, 57].

Although glaucoma cannot be cured, early diagnosis and treatment of glaucoma is reported to able to reduce or halt the risk of visual field loss [46], as well as being cost efficient[65]. Intraocular pressure (IOP) lowering (e.g. by eye drops, laser, surgery) is the therapeutic approach that can reduce the risk of glaucoma development and slow down the rate of disease progression. Due to the irreversible nature of the disease, early detection of glaucomatous changes is crucial to allow for early treatment prior to the onset of functional visual loss and aid in glaucoma management. A medical image-based computer-aided diagnosis for glaucoma is a means of taking such an active role.

Figure 2.11: A Contrast of Vision - (Left) Normal’s patient view of a scene. (Right) Glaucoma’s patient view of a scene. Note the loss of vision around the peripheral. Image from NIH [80]

2.2.1.1 Glaucoma Assessment

Eye doctors usually use several tests to detect glaucoma. These are usually categorized as functional or image-based assessment. Functional assessment includes measuremen-

t of IOP (tonometry) and visual field examination. Image-based evaluation relies on optic nerve head imaging devices to determine optic nerve head structural damage or thinning of the retinal nerve fiber layer (RNFL). Examples of such imaging devices are Retinal Fundus Camera, Heidelberg Retinal Tomography (HRT) and Optical Coherence Tomography (OCT). Nonetheless, most of these methods have their own limitations.

Measurement of IOP was reported to have poor sensitivity of around 50% [114]. This is partially due to cases of normal tension glaucoma, where patients have a condition in which optic nerve damage and vision loss have developed even with a normal pressure inside the eye. Visual field examination is often time-consuming and found to be unreliable due to factors such as patient's fatigue and learning effects. Image-based evaluation using HRT and OCT are highly costly and are usually only available at tertiary hospitals, thus limiting its outreach.

Figure 2.12: Cup-to-Disc Ratio - The vertical cup-to-disc ratio is used as clinical measure in assessing glaucoma.

On the other hand, retinal fundus cameras are commonly found and used at primary care for assessment of the retinal and optic nerve. Unlike IOP, the appearance of the optic nerve does not fluctuate from day to day, and unlike visual field tests, it is not dependent on patient co-operation [87]. In addition, compared to functional assessment, direct inspection of the optic disc seems to have the highest accuracy [109]. The study

in [109] also showed in particular that Cup-to-Disc Ratio (CDR), Fig. 2.12, is the most important single disc parameter for glaucoma. It is an important clinical indicator of glaucoma[26], and measures the ratio between the vertical height of the optic cup against the vertical optic disc height. Currently, CDR assessment is performed manually by the ophthalmologists to gauge and monitor optic cupping size and degeneration of the optic nerve head in glaucoma. In this thesis, the objective is to design algorithm approaches to directly assess the optic nerve changes in retinal fundus images.

2.3 Related Works

As mentioned in the previous section, to assess the optic nerve head for glaucoma in retinal fundus photography, the optic disc and optic cup have to be inspected and delineated. In a clinical environment, this is usually manually outlined by an ophthalmologist. However, the task of manual annotation is often labor intensive, leading to researchers developing automated processes for disc and cup segmentation.

2.3.1 Optic Disc Localization and Segmentation

In general, the optic disc appears as a bright elliptical region in the red channel of the RGB space, or brightness channel in other color spaces, Fig. 2.13. Based on this observation, works to determine the optic disc have been extensively reported. Methods of optic disc boundary detection usually involves two steps: localization and boundary segmentation. The former aims to find the location of the optic disc center, while the latter delineates the boundary of the optic nerve head from the retina automatically. Broadly, these approaches can be categorized as intensity-based approach, shape-based approach and learning-based approach, or a combination of these approaches.

In the first category, [112, 51, 45] proposed locating the optic disc by identifying the area with the highest variation in intensity (1-10%) of adjacent pixels. Similar works also search for the brightest region in the fundus image as the optic disc using PCA [66, 67], morphological operators[128, 11] and template-matching[89, 63] techniques and fine-tune the localized optic disc contour using watershed[128], Hough transform [107, 150, 23] and ellipse-fitting [70, 133] methodologies.

Figure 2.13: General Appearance of the Optic Disc in a Retinal Fundus Image - (Left) Original Color Fundus Image. (Right) Red Channel of the Fundus Image. The optic disc appears as the brightest circular-elliptic region.

In the second category, shape-based approaches in the form of deformable models, level sets and geometric models have been proposed with reported improved accuracy compared to intensity-based methods. In the local deformable model approaches, Osareh [89] and Mendels [74] proposed using GVF-snake-based gradient energy minimization approaches. Parametric deformable models using modified ASM [67, 66] were reported to delineate the optic disc boundary with high experimental accuracy. Yin [143] proposed using the Circular Hough transform combined with active shape models for better optic disk boundary detection. Lowell and colleagues [72] used a variant of Hu’s circular deformable model with an elliptical model and variable edge-strength dependent stiffness to the same task. Similarly, [122] used the Chan-Vese formulation for optic disc localization. Approaches using level sets includes [70, 133], which uses a variational level set formulation without re-initialization to segment the optic disc. Besides using local shape to detect the optic disc, Hoover and colleagues [49] presented a voting-based fuzzy convergence method to detect the optic disc by finding the origination of the blood vessel network using global geometric shape relationship. This framework uses the geometric relationship between the optic disc and main blood vessels arcades and was reported to be more robust in finding the optic disc in the presences of drusen², exudates³ and imaging artifacts such as uneven illumination. Ravishankar et. al. [102] as well as Foracchia et. al.[36] also reported a more robust optic disc localization using a similar framework.

²Yellowish deposits under the retina.

³Lipid residues that are associated with vascular leakage from damaged capillaries

In the third category, learning-based methods were used to discriminate features to find the optic disc. Xu *et al.*[137] extended the GVF-snake-based approach with an unsupervised k-means clustering of the contour points to reduce the influence of blood vessels occlusion and detect the optic disc boundary. Abramoff *et al.*[8] used pixel features and a set of Gaussian Gabor wavelet steerable filter bank of responses as features in a k-nearest neighbor (kNN) classifier. A hard classification is then performed using a majority vote of the kNNs on each pixel. Tobin *et al.* [58] used the retina vasculature along with the retina luminance, vessel density, average vessel thickness and orientation in a two-class Bayesian classifier to determine the optic disc location. No segmentation of the optic disc was done. Niemeijer *et al.* [85] introduced a similar pixel-based classification learning method by using kNN regressor and a circular template to estimate the distance of each pixel to a image feature in the fundus image.

A summarized comparison table of these methods are compiled in Table 2.2.

Table 2.2: Categorization of Existing Optic Disc Methods

Method	Approach	Year	Techniques Used	Localization	Segmentation
Sinthanayothin [112]	Intensity	1999	Thresholding	✓	✗
Inoue [51]	Intensity	2005	Thresholding	✓	✓
Hatanaka [45]	Intensity	2010	Thresholding, Canny Edge Detection, Spline Fitting	✓	✓
Walter [128]	Intensity	2001	Thresholding, Morphology Operator, Watershed Transformation	✓	✗
Aquino [11]	Intensity, Local Shape	2010	Intensity Variation, Morphology Operator, Low-Pass Filter, Prewitt Edge Detection, Otsu Thresholding, Circular Hough Transform	✓	✓
Osareh [89]	Intensity, Local Shape	2002	Template Matching, Morphology Operator, Gradient Vector Flow (GVF) Snake	✓	✓
Lalonde [63]	Intensity, Local Shape	2001	Hausdorff-based template matching, Canny Edge Detection, Pyramidal Decomposition using Haar-based discrete wavelet transform	✓	✓
Sekhar [107]	Intensity, Local Shape	2008	Morphology Operator, Threshold, Circular Hough Transform	✓	✓
Cheng [23]	Intensity, Local Shape	2011	Thresholding, Morphology Operator, Mean Filter, Constrained Elliptical Hough Transform	✓	✓
Rangayyan [150]	Intensity, Local Shape	2008	Median Filter, Canny Edge Detection, Circular Hough Transform	✓	✓
Liu [70]	Intensity, Local Shape	2009	Thresholding, Variational Level-Set, Ellipse Fitting	✓	✓
Wong [133]	Intensity, Local Shape	2008	Thresholding, Variational Level-Set, Ellipse Fitting	✓	✓
Mendels [74]	Intensity, Local Shape	1999	Local Minima Detection, Morphology Operator, Gradient Vector Flow (GVF) Snake	✓	✓
Yin [143]	Intensity, Local Shape	2011	Hough transform, Active Shape Model, Principal Component Analysis (PCA), Ellipse Fitting	✓	✓
Lowell [72]	Intensity, Local Shape	2004	Template Matching, Circular Deformable Model	✓	✓
Tang [122]	Intensity, Local Shape	2006	Thresholding, Chan-Vese (CV) Active Contour Model	✓	✓
Foracchia [36]	Geometric Shape	2004	Weighted Residual Sum of Squares (RSS), Simulated Annealing (SA)	✓	✗
Ravishankar [102]	Intensity, Geometric Shape	2009	Morphology Operator, Hough Transform, Circular Hough Transform	✓	✓
Hoover [49]	Intensity, Geometric Shape	2003	Fuzzy Convergence, Illumination Equalization, Fisher's Linear Discriminant Statistic	✓	✗
Li [66, 67]	Intensity, Local Shape, Learning	2003	Principal Component Analysis (PCA), modified Active Shape Model (ASM)	✓	✓
Xu [137]	Intensity, Local Shape, Learning	2007	Circular Hough Transform, Modified GVF Snake, K-Means Clustering	✓	✓
Abramoff [8]	Intensity, Learning	2007	Color Opponency Gaussian Filter Bank, Gabor wavelet, Principal Component Analysis (PCA), k-Nearest Neighbor (kNN) Classifier	✓	✓
Tobin [58]	Intensity, Geometric Shape, Learning	2006	Bayesian Classifier	✓	✗
Niemeijer [85]	Intensity, Local Shape, Learning	2009	Gaussian Filter Bank, k-Nearest Neighbor (kNN) Classifier, k-Nearest Neighbor (kNN) Regression, Circular Template Matching, Sequential Floating Feature Selection (SFFS)	✓	✗

2.3.2 Optic Cup Segmentation

Many methods have been reported for optic disc detection, but significantly fewer methods have been presented for optic cup detection, which is much more challenging. Here, some of the primary challenges faced in optic cup segmentation from a computer vision standpoint and the existing works found in literature will be presented.

2.3.2.1 Challenges

Figure 2.14: Scaling disparity of optic cup. - (*From Left to Right*) CDR of 0.16, 0.56, 0.8, 0.91 respectively.

Scaling. The size of the optic cup appears differently from one eye to another. In particular, the scaling disparity can range from 0.1 - 1.0 cup-to-disc ratio (CDR), Fig. 2.14. As such, features and details that can be extracted from the neuro-retinal rim (the area between the optic disc and cup) or from the optic cup may fluctuate (less discriminative) and become biased. The difficulty in estimating the size of the optic cup in the absence of a prior can greatly affect accurate delineation of its boundary. Henceforth, for robust optic cup localization, it is important to extract reliable and stable descriptors in the retinal image across multiple scales.

Variation in Pallor Appearance The variation in pallor appearance can be traced to the following factors. First, the optic cup is a cavity-like depression within the optic nerve and the extend of an optic cup excavation varies from one individual to another, (see Fig. 2.15). A punch-out optic cup appears as a well-defined boundary within the optic disc, while a shallow, saucer-like optic cup produces little or gradual boundary in

the retinal image, which may sometimes be an early, subtle sign of damage . Examples of shallow cup region and punch-out optic cup is illustrated in Fig. 2.16. Second, image acquisition from different lighting environments and mixed retina image cameras can affect the photometric quality, transform and aspect ratio of the retina image captured. Although some of these changes can be compensated by normalization and filtering, they may also introduce new distortions and affect the image contents.

Figure 2.15: Extent of optic cup excavation and its effect in pallor appearance variation. -

Viewpoint Perspective The optic nerve is a three dimensional structure and projects different appearances when observed at different viewing angles or when seen in different field of capture (ETDRS Field 1 - Optic Disc Centered, Field 2 - Macula Centered). In general, when the optic nerve is being assessed, the Field 1 image is used, while a Field 2 image is used to appraise the state of the macula. In addition, viewpoint perspective variation may result from a common population condition known as tilted optic nerve [130], where the optic nerve exits the eye at an oblique angle, compared to the normal optic nerve which exits the sclera at a 90° angle. Fig. 2.17 shows instances of these images. The challenge of optic cup viewpoint perspective invariance thus, requires additional known information (eg. inferring degree of tilt) and this complex task can be towering if good generalization is required as a model across all other unseen image instances.

Figure 2.16: Variation of Pallor Intensity. - (*Top Row*) Examples of a shallow and diffused cup pallor region. (*Bottom Row*) Examples of a 'punched-out' optic cup.

Figure 2.17: Viewpoint perspective. - (*Top-Left*) Field 2 image captured image. (*Top-Right*) Field 1 image of the same subject. (*Bottom*) Example of tilted optic nerve.

Occlusion Similar to object and human-based recognition, distortion and misrepresentation from occlusion can lead to a degradation of visual appearance and affect the task of optic nerve assessment in retinal fundus imaging. In particular, if the key salient features of the optic cup boundary are occluded, the performance of the localization algorithm may also be affected. Examples of possible occlusion in this domain are media opacity from cataracts, artifacts from image acquisition or from pre-existing medical conditions. Fig. 2.18 shows the effects of media opacity on the appearance of the optic nerve.

Figure 2.18: Challenges in Occlusion. - (*Top*) Media opacity artifacts from image acquisition. (*Bottom*) Effects of cataracts on the visual appearance of the optic nerve.

2.3.2.2 Related Works

In this section, some of the noteworthy works and trends in this field will be discussed and highlighted. Optic cup detection in digital retinal fundus photography can be broadly classified along two imaging strategies based on multiview images [137, 8, 53] or single

Figure 2.19: Close-up Appearance of the Optic Cup - (Left) In RGB. Note that the blood vessels are more prevalent in the nasal sector. (Right) Green Channel of the Optic Cup. This channel is most commonly used as the pallor has the best contrast.

images [51, 70, 133, 120, 45, 54, 139, 27]. For each image strategy, the methods can be further classified based on their approaches. This is summarized in Table 2.3.

Multiview image strategies utilize depth information obtained from stereo geometry setups, which can be fixed (stereo images)[137, 8] or estimated [53]. In stereo images, pixel depth information is calculated by modeling the fixed geometric configuration. This is then used to detect the cup, such as by pixel classification [8]. Xu and colleagues [137] proposed using depth information derived from stereo images, local gradient magnitude, shape as global information and a smoothing term as the energy function for their deformable model. Stereo information has also been estimated by calculating the motion displacement field from sequential images [53], and cup boundary points were estimated by detecting depth discontinuity points. Although multiview methods can provide a measure of depth, there are additional acquisition requirements, such as the need for customized stereo retinal fundus camera hardware setups, or the need to ensure consistent image quality in sequentially acquired images.

Single retinal images are more straightforward to acquire and analyze for screening, although the lack of depth information necessitates the use of other visual indicators [51, 70, 133, 120, 45, 54, 139, 27] for optic cup detection.

One of the most heavily used visual indicator is the optic cup pallor (also known as optic cup color), (see Fig. 2.19). In [51], Inoue *et al.* described a method to segment

the optic cup in the green channel by a discriminatory analysis of the optic disk pixels. A threshold level that separates the histogram of the ROI image into two classes is determined by the maximum value of the decentralization between the two classes. In another pallor-based only approach, Hatanaka *et al.* directly obtained the cup-to-disk ratio in [45] from a detected optic disk, by averaging vertical pixel intensity profiles within the optic disk on the blue channel to determine the cup area edge. Beyond using intensity profiles and histograms, Liu *et al.* [70] and Wong *et al.* [133] have also demonstrated methods using level sets for optic cup segmentation.

In an earlier work to introduce domain knowledge into optic cup segmentation, Tan *et al.* have proposed using a probabilistic Gaussian mixture model to regulate the temporal optic cup boundary and fuse the result with variational level set for optic cup segmentation [120]. Although level set-based methods have the advantages of being independent from contour parametrization, and are adaptable to topological changes, they are highly susceptible and influenced by noise, poor image gradient contrast and muddled background. As such, faced with a shallow temporal cup boundary, the level set segmentation can be vulnerable to leaks and shocks in this zone. Using domain knowledge of the optic cup, the proposed approach generated a closer-fitting temporal cup boundary compared with [70]. Instead of regulating the temporal region of the optic cup, Joshi *et al.* [54] propose a symmetric optic cup shape domain prior. The authors proposed to first perform blood vessel removal, followed by cup detection in the LAB space using an empirically determined dynamic threshold. This returns a thresholded cup pallor. To avoid the blood vessels at the nasal cup boundary, the authors assume that the optic cup will be symmetric. The nasal optic cup boundary was then formed by mirroring the temporal optic cup boundary around the vertical axis of symmetry of the optic disc.

Another significant work incorporating domain knowledge for cup segmentation is by the use of vessel kinking. Wong *et al.* first put forward the idea of automatic detection of vessel kinks [132], which are the bendings of small vessels as the vessels traverse from the optic disc to cup. Vessels around an estimated cup boundary are extracted and identified using a Bayesian rule, comprising of edge and wavelet prior probability coefficients. The kinks are then identified if the directional angular changes of the vessel edges exceed 20° . An ellipse fitting approach is then used on these identified points to obtain the optic cup

boundary. However, the approach was only tested on a small dataset of 27 images. In another recent work, Yin *et al.* [145] proposed the domain prior of finding the Nasal Side Vessel Wall (NSVW), to reduce the effects of blood vessel interweavement, and integrate with an Active Shape Model (ASM) approach for optic cup segmentation.

Recently, learning-based methods have gain popularity in the tasks of medical image segmentation. A machine learning framework was proposed in [139] by Xu *et al.* to detect the optic cup. Features extracting color difference between cup and disc regions, as well as the elliptical shape of the optic cup are derived from each sliding window. These features are then selected based on group sparsity constraint and fitted in a non-linear kernelized Support Vector Regression (ϵ -SVR). The final cup region is then obtained by applying non-maximal suppression. The authors note that their method is computationally heavy (7.2 minutes per image) and has a limitation in detecting large optic cup as the non-maximal suppression suppresses the rim of the cup.

Table 2.3 is a summary of the current noteworthy optic cup segmentation methods in this literature review.

Table 2.3: Taxonomy of Existing Optic Cup Segmentation Methods

Method	Year	Strategy	Pallor-based	Domain Prior	Machine Learning	Techniques Used
Inoue [51]	2005	Single	✓	✗	✗	Discriminatory Threshold
Xu [137]	2007	Multi-View	✓	✗	Unsupervised	Circular Hough Transform, Modified GVF Snake, K-Means Clustering
Abramoff [8]	2007	Multi-View	✓	✓	Supervised	Color Opponency Gaussian Filter Bank, Gabor wavelet, Principle Component Analysis (PCA), k-Nearest Neighbor (kNN) Classifier
Wong [133]	2008	Single	✓	✗	✗	Threshold, Variational Level-Set, Ellipse Fitting
Liu [70]	2009	Single	✓	✗	✗	Threshold, Variational Level-Set, Ellipse Fitting
Wong [132]	2009	Single	✓	✓	✗	Threshold, Variational Level-Set, Canny Edge Detector, Gabor wavelet filter, Bayesian Rule
Hatanaka [45]	2010	Single	✓	✗	✗	Vessel Erase, Intensity Profile, Averaging, Smoothing, Threshold
Tan [120]	2010	Single	✓	✓	✗	Variational Level Set, Expectation Maximization, Convex Hull
Joshi [54]	2010	Single	✓	✓	✗	Threshold A Channel of LAB, Morphology operator, Intensity Mapping, Dynamic Threshold, Symmetric Cup
Xu [139]	2011	Single	✓	✓	Supervised	Sliding Window Feature Extraction, Group Sparsity Constraint, Support Vector Regression, Non-maximal Suppression
Yin [145]	2012	Single	✓	✓	✗	Active Shape Model, Frangi Filter, Nasal Side Vessel Wall (NSVW), Convex Hull
Joshi [53]	2012	Multi-View	✓	✓	✗	Circular Hough Transform, Chan-Vese (CV) Active Contour Model, Image registration by Discontinuity Displacement Field, Threshold A Channel of LAB, Morphology operator, Intensity Mapping, Dynamic Threshold, Confidence Value Mapping, Radial Smoothing.

2.4 Evaluation Dataset

An introduction of the dataset and metrics used for evaluation in this work is presented. The optic cup segmentation methodologies are evaluated on the ORIGA^{-light} database, as it is the only public dataset, to our knowledge, that consists of the manual demarcation of the optic disc and cup boundaries by clinical experts.

2.4.1 ORIGA^{-light}

The ORIGA^{-light} (An Online Retinal Fundus Image Database for Glaucoma Analysis and Research) [147] database is an online depository which contains of 650 annotated retina images (482 normal and 168 glaucoma cases) from the Singapore Malay Eye Study (SiMES) [35]. This is a population-based cross-sectional study of the Malay ethnic group, aged 40-79, on the prevalence of and risk factors for age-related eye diseases in Singapore. This study was conducted from August 2004 to June 2006. The glaucoma images are distributed amongst the dataset, (*i.e.* They are not all grouped together in the front or at the end of the dataset). The first 325 images from the dataset, forms Set I, denoted as Set_I and the rest forms Set II, denoted as Set_{II} .

The color fundus images were acquired using a 45° FOC Canon CR-DGi retinal fundus camera with a Canon 10D single-lens reflex camera (SLR). The color images were captured at an image resolution of 3072 x 2048 pixels and saved in JPEG format. The pathological signs in each image are graded following a clinical glaucoma grading protocol defined by the Centre of Eye Research, Australia. The grading information for each image is shown in Table 2.4, with manual segmentation of the optic cup and disc boundaries by trained grading professionals from the Singapore Eye Research Institute using ORIGA^{-GT}, an image grading tool provided for manual segmentation and ground-truth building. To perform the grading, the user manually defines points along the optic disc and cup boundaries, and an ellipse, based on Direct Least Square Fitting Algorithm (B2AC) [34], is used to fit the input points. Each image was manually assessed by the grading team after discussions and mutual agreement among the team. To avoid bias in the grading processing, no clinical information was provided at the time of grading. Samples of the retinal images in the dataset, as well as the ground truth optic disc and cup, are shown in Fig. 2.20.

Figure 2.20: Sample Images and Ground-Truth in ORIGA-light dataset - (*From Left to Right*) Original ROI image (800x800) in the dataset, Ground-truth Optic Disc Mask, Ground-truth Optic Cup Mask, Ground-truth overlaid on original image.

Table 2.4: Clinical Grading Information in ORIGA^{-light} dataset. Table adapted from [147]

Annotation Term	Details
Eye	Left eye / Right eye
Image Quality	Disc presented / No Disc presented
Gradability	Gradable / Not Gradable
Field	Retina image taken in Field 1 - 6
Disc Size	Large Disc / Middle Disc / Small Disc
CDR	Cup-to-Disc Ratio, automatically calculated
ISNT rule	Follow ISNT rule / Not follow ISNT rule
RNFL	Retinal Nerve Fiber Layer thinning / Not thinning
Notch	Yes / No
Disc Haemorrhage	Presented / not presented
Alpha PPA	Presented / not presented
Beta PPA	Presented / not presented
Comments	Additional information

2.5 Evaluation Criteria

2.5.1 ORIGA^{-light} Evaluation Criteria

The following are the criteria recommended for evaluation of the ORIGA^{-light} dataset in [147]. Here, E_{dt} refers to the detected cup region, E_{gt} refers to the ground-truth labelled ellipse, D_{dt} is the vertical height of the detected cup, D_{gt} is the vertical height of the ground-truth cup, DD is the vertical height of optic disc.

2.5.1.1 Non-overlap Ratio (m_1)

The non-overlap ratio, m_1 , measures how well the segmented area matches with the clinical segmentation. It is calculated by measuring the error between the number of pixels in the intersection of segmentation and reference, divided by the number of pixels in the union of segmentation and reference. A perfect segmentation returns a value of 0 and a value of 1 means there is no overlap at all between the segmentation and reference. The highest possible value for m_1 is 1.

$$m_1 = 1 - \frac{\text{area}(E_{dt} \cap E_{gt})}{\text{area}(E_{dt} \cup E_{gt})}, \quad (\text{Eq. 2.1})$$

2.5.1.2 Absolute Area Difference (m_2)

The absolute area difference, m_2 , gauges the extent of the segmentation away from the clinical ground-truth. This metric is measured by the absolute difference in area between the segmentation and reference, divided by the area of the reference. This value is 0 when there is a perfect segmentation, and larger than zero otherwise. It is important to note that the perfect value of 0 can also be obtained for a non-perfect segmentation, as long as the area of that segmentation is equal to the area of the reference. Note that there is not maximum value for this criteria.

$$m_2 = \frac{|area(E_{dt}) - area(E_{gt})|}{area(E_{gt})}, \quad (\text{Eq. 2.2})$$

2.5.1.3 Absolute CDR Error (δ)

The absolute CDR error (δ) measures the CDR variation between the vertical height of the cup segmentation and cup reference, divided by the vertical height of the reference disc. This value is 0 when a perfect segmentation is met, and larger than zero otherwise. It is important to note that the perfect value of 0 can also be obtained for a non-perfect segmentation, when both the area of that segmentation and area of the reference are equal.

$$\delta = \frac{|D_{dt} - D_{gt}|}{DD} \quad (\text{Eq. 2.3})$$

2.5.2 Superpixel Classification Evaluation Criteria

Here, TP and TN refers to the true positives and true negatives, and FP and FN denotes the false positives and false negatives, respectively.

2.5.2.1 Sensitivity (P_+) and Specificity (P_-)

Sensitivity, defined here as (P_+), measures the ratio of actual positives which are correctly identified, while specificity, defined here as (P_-), is the proportion of actual negatives which are correctly identified.

$$P_+ = \frac{TP}{TP + FN}, P_- = \frac{TN}{TN + FP} \quad (\text{Eq. 2.4})$$

2.5.2.2 Balanced Accuracy ($\bar{P}$)

Balanced Accuracy is a performance measure commonly used to evaluate a classifier's performance on an imbalanced dataset. It is defined as

$$\bar{P} = \frac{P_+ + P_-}{2} \quad (\text{Eq. 2.5})$$

2.5.2.3 Receiver Operating Characteristic Curve (ROC)

The receiver operating characteristic curve is a plot of the true positive rate (Sensitivity) against the false positive rate (1 - Specificity) at different classifier threshold settings. The top left corner of the plot is considered the most desirable point, where the false positive rate is 0, and the true positive rate is 1. This plot is often used in binary classifications to study the prediction output of a classifier. The area under the ROC curve (AUC) is a measure of the overall performance of the classifier. The maximum value of the AUC is 1, indicating a perfectly accurate classification, while an AUC of 0.5, depicted as a diagonal line from the left bottom to the top right corner of the ROC curve, indicates pure chance or random guess.

Chapter 3

The Detection and Use of Vessel Kinking to Improve Optic Cup Detection in Digital Retinal Fundus Images

3.1 Introduction

The detection of the optic cup is important for glaucoma assessment. However, it is difficult to visually discern the optic cup using only color information from digital retinal fundus images. A clinical phenomenon known as the pallor-cupping discrepancy describes an under-estimation of the cup boundary if only color information is used. In the early stages of glaucoma, enlargement of the cup may progress ahead of that of the area of pallor. A potential pitfall in interpreting optic nerve head cupping is to look only at the area of pallor and miss the larger area of cupping [101]. A better estimation of the optic cup boundary can be obtained from an assessment of kinking of vessels at the temporal cup margin, which is an accepted clinical practice [52].

This chapter first describes the motivation of using visual features of vessel kinks, which are small vessel bends along the cup margin. Next, a framework, which automatically detects and integrates the kinking of vessels with color information, is introduced to improve optic cup detection.

Figure 3.1: (Left) Optic disc showing kinking of vessels (black arrows). (Right) Illustration of vessel kinking.

3.1.1 What are Vessel Kinks?

Kinking refers to the mechanical bending of *small* vessels as they traverse across the surface of the neuroretinal rim and fall into the depression of the optic cup. When projected onto a 2D plane, such as on a single retinal image, the vessel trajectory can be observed to bend or 'kink' across the optic cup boundary, indicating a change in topology due to the optic cup excavation. An illustration is provided in Figure 3.1.

3.1.2 Prior Works

The detection and use of kinks for optic cup detection is a recent approach and there are few published methods in literature. It is also a challenging task to detect small vessel kinks due to its size and visibility. Wong *et al.* [133] first reported an approach where kinks are detected through localized patch generated from a preliminary cup boundary obtained by level-set. A morphological approach is then used to detect vessel kinks around the cup estimate boundary. However, the approach is highly reliant on the initial

cup estimate from pde-based level-set segmentation, which is susceptible to ‘leaks’ or ‘shocks’ during curve evolution.

A similar implementation by [55], described the use of r-bends, or kinks, by first identifying vessels via trench detection, and subsequently, r-bends were detected along the detected vessels in using a prescribed region-of-support approach. Multiple steps, which involved selecting only r-bends within a specified circular fit were used and a polynomial fitting was employed to determine the optic cup boundary. However, vessel bendings were used as the only indicators of the optic cup. Although the exclusive use of kinks may be ideal in some cases, in practice, often, there are some limitations, such as when insufficient kinks are detected, or when poor vessel visibility precludes the detection of vessels for kinks analysis. Under these circumstances, it may be useful to integrate kinks with other sources of cup information.

3.2 Methodology

A framework is proposed to detect and use the kinking of vessels, which are topological indicators of the optic cup depression, in order to improve optic cup detection. Unlike previous approaches, in this work, vessel kink detection is performed 1) independent of cup estimate, and, 2) solely focused on the temporal region of the cup, where the cup boundary is usually weaker. Furthermore, the detected vessel kinks are used as key points to integrate with color (or pallor) information.

The motivation of this approach is based on the observation that 1) the automatic detection of the optic cup boundary using prior approaches often fail, and ‘leak’, in the temporal region (Fig. 3.1(*Left*)), where the boundary is usually the weakest; 2) exclusive use of vessel kinks to segment the cup boundary without integration of color information is not reliable, especially in cases where small vessels are absent or when their visibility is poor.

A disc image is first used to detect vessel edges using a support vector machine classifier with hybrid edge and wavelet features. Piece-wise probing is used to detect locations of kinks along the vessels. Optic cup feature points are integrated using these detected kinks and color-based indicators of the optic cup boundary. The set of feature points are then combined using a convex optimization-based minimum volume enclosing ellipse to obtain the optic cup boundary. The overall system is presented in Fig. 3.2.

Figure 3.2: Schematic diagram of the proposed system. The main contributions are indicated in the colored boxes.

3.2.1 Vessel Trajectory Detection

The main objective of this step is to detect the trajectory of the blood vessels. Instead of performing accurate vessel segmentation, the goal is to detect kinking along vessel edges. This reduces the complexity of the problem, since it reduces the dependence on granularity of the bank of pre-defined matched filters on the vessel caliber.

A three-step approach is proposed to detect plausible vessel edges with trajectories. First, vessels are enhanced using Gabor wavelets; Second, vessel edges are extracted using the Canny edge detector, which helps to link short vessel segments; Third, features are extracted from local patches for vessel edge classification.

3.2.1.1 Vessel Enhancement using Gabor Wavelets

Often, a main challenge in the detection of vessel kinks is the poor visibility of small vessels. Existing edge detection methods do not perform well, in general, as they require extensive tuning to balance the sensitivity of under-detection (*i.e.* little or no vessel edge responses) and over-detection (*i.e.* noise). Instead, the use of Gabor wavelets is investigated to perform two functions, a) enhance the visibility of vessel edges, and b) as an analyzing wavelet for vessel edge classification. The two-dimensional Gabor wavelet is suitable as it offers the advantages of tuning specific frequencies for vessel enhancement and noise filtering, and directional selectivity. Previous application of Gabor wavelets was used to perform supervised multi-scale retinal vessel segmentation [113]. An introduction of the two-dimensional Gabor wavelets used in this work is described here for completeness and better understanding. Letting coordinates in the real plane $\mathbb{R}^2$ be represented by $\mathbf{x} = [x \ y]^T$, the Gabor wavelet is defined as

$$\psi(\mathbf{x}) = e^{jk_0(\mathbf{x})} e^{-\frac{1}{2}\|A\mathbf{x}\|^2} \quad (\text{Eq. 3.1})$$

where $k_0 = [k_x \ k_y]$ represents the frequency of the complex exponential, $A = \text{diag}[\epsilon^{-\frac{1}{2}} \ 1]$, and $\epsilon \geq 1$ represents the wavelet elongation factor.

The wavelet transform is defined as

$$W(\mathbf{x}) = \max |T(\mathbf{b}, \theta, a)| \quad (\text{Eq. 3.2})$$

where

$$T(\mathbf{b}, \theta, a) = C^{-\frac{1}{2}} a^{-1} \int \psi^*(a^{-1} r_{-\theta}(x - b)) f(\mathbf{x}) d^2x \quad (\text{Eq. 3.3})$$

and $\mathbf{b}$ is the displacement, $r_{-\theta}$ the rotation operation with angle θ and a is the dilation parameter which controls the scale of the transform, and C is a normalizing constant. The

Table 3.1: Selected Parameters for Gabor Wavelet Implementation

Parameter	Symbol	Value
<i>Elongation</i>	ϵ	1
<i>Dilation</i>	a	5
<i>Rotation Angle</i>	θ	$0^\circ - 170^\circ$ (default)
<i>Rotation Angle Stepsize</i>	$\Delta\theta$	10° (default)
<i>Complex Exponential Frequency</i>	k_0	$[0, 3]$ (default)

wavelet transform T is performed across the retinal image, and the maximum response of the transform is taken as θ varies from 0° to 170° , where $\Delta\theta = 10^\circ$.

In tuning the wavelet transform, the elongation parameter, ϵ , and dilation (or scale) parameter, a , are adjusted to enhance vessel edges. The scale of the transform controls how well fine vessel edges and contours are enhanced; having too small a value results in a noisy grainy image whereas a large value of a misses the small vessels. Similarly, a larger value for the elongation term, ϵ , emphasizes the linearity of vessels, at the expense of neglecting vessel curvature. In order to enhance vessel edges for kinks detection, a number of experiments were conducted and it was found empirically that values of $a = 4, \dots, 10$ and $\epsilon = 1, \dots, 10$ were suitable. In this work, values of $a = 5$ and $\epsilon = 1$ were found to offer optimal enhancement, in terms of preserving the curvature trajectory and increasing contrast against non-vessel elements in the transformed image. A value $k_0 = [0, 3]$ is kept as the low frequency component for complex exponential. This setting is recommended and consistent with the experimental findings in [113], as it was found to enable the wavelet transform to present strong responses for pixels associated with blood vessels. Various settings of a and ϵ are shown in Fig. 3.3. In the enhanced image, it can be seen that the thick vessels are enhanced comparably to thinner vessels, due to the scaling parameter, a . The final parameters used in this implementation is listed in Table. 3.1.

3.2.1.2 Vessel Edge Detection

Next, the enhanced vessel edges are detected from the wavelet transform image by applying a Canny edge detector. The characteristics of Canny edge detection are particularly suitable for the detection of vessels traversing the optic cup boundary. Differences in the appearance of the background as the vessel crosses from the disc to the cup result in a variation of the visibility of vessels. Using a single threshold for vessel edge detection,

Figure 3.3: Different settings for a and ϵ for the enhancement of thin vessels. (*Top-Left*) RGB Image, (*Top-Centre*) $a = 20$ $\epsilon = 4$, (*Top-Right*) $a = 1$ $\epsilon = 4$, (*Bottom-Left*) $a = 4$ $\epsilon = 100$, (*Bottom-Centre*) $a = 4$ $\epsilon = 0.01$, (*Bottom-Right*) $a = 5$ $\epsilon = 1$, The set of parameters leading to the transformed image in the lower right corner was used. For illustration purposes the intensity ranges for each image was stretched to $[0, 255]$.

Table 3.2: List Of Features For Vessel Edge Classification

Feature	Description
$R_\mu(E_j), G_\mu(E_j), B_\mu(E_j)$	Mean color along edge E_j
$R_\sigma(E_j), G_\sigma(E_j), B_\sigma(E_j)$	Std deviation of color along edge E_j
$W_\mu(E_j)$	Mean wavelet response along edge E_j
$W_\sigma(E_j)$	Std deviation of wavelet response along edge E_j
$G_\nabla(E_j)$	Mean gradient response along edge E_j

such as in the Prewitt detector, may fail to detect the vessel edge, especially within the optic cup. Furthermore, the emphasis on linking of edge segments in the Canny algorithm is advantageous to obtain a longer and more complete run of the vessel course for kinking analysis.

In this work, the Canny edge detection algorithm is applied on $W(x)$ to generate a set of edges. A mask based on the detected optic disc is applied to remove edges outside the optic nerve head. To address any vessel bifurcations, branching pixels along each edge are selected if they have more than two 8-connected neighbours. These pixels are used to separate the edge into the component branches. Any branching segment less than 50 pixels, equivalent to 250 μm , is removed. If more than one possible path exists for the segment, only the path with the least deviation from the prior points, considered inwards from the optic disc rim, is kept while other paths are discarded. This reduces the likelihood of detecting vessel bends due to minor vessels splitting off from the vessel under consideration. A consolidated set $\mathbf{E}$ of p edges $\mathbf{E} = \{E_j\}_{j=1,\dots,p}$ is formed from the prior edge detection after applying these constraints, where $E_j = \{\mathbf{e}_k\}_{k=1,\dots,q}$ denotes the set of connected pixels for each edge.

3.2.1.3 Feature Extraction

In order to differentiate vessels from non-vessel edges in $\mathbf{E}$, the following features (see Table 3.2) are extracted to build a discriminative learning model for vessel edge classification.

- (i) *Mean Color:* Vessel color is a distinctive exploitable feature of vessels within the optic nerve head. Compared to other optic nerve structures, vessels have higher red content, lower green and blue content.

- (ii) *Std Dev of Color*: This is based on the observation that the deviations of color along a vessel edge should be minimal compared to non-vessel edges.
- (iii) *Wavelet response*: The magnitude of the wavelet transform represents the response of the attuned vessel enhancement kernel. The response should be high and relatively consistent for vessels.
- (iv) *Gradient*: Mean modulus of the derivative of the edge pixels in the green color space domain. The green domain is chosen as it offers the highest contrast among the other RGB color spaces for vessels. This has also been observed in previous works. The gradient, $G_{\nabla}(E_j)$, is defined as,

$$G_{\nabla}(E_j) = \text{mean}(\|\nabla \mathbf{e}_k\|)_{k=1,\dots,q} \quad (\text{Eq. 3.4})$$

where $\nabla \mathbf{e}_k$ represents the partial derivatives in the vertical and horizontal directions for each pixel in the j -th edge. To avoid bias and over-representation, the features were normalized to the range $[0, 1]$, with zero mean and unit variance.

These features are chosen empirically based on visual observations to describe the blood vessels, through different feature representations such as color, wavelet responses, and gradient to differentiate them from spurious noise and/or image artifacts. The image gradient from the red, blue and grey channel was investigated, however, the vessels were found to have the largest contrast in the green channel. This finding is also largely consistent with many literature works such as in [112, 49, 36, 113] performing retinal vessel segmentation. The mean color, mean wavelet response and mean gradient response each offers highly distinctive characteristics to represent the vessel edges. However, as the image quality from the population-based ORIGA^{-light} dataset is not consistent, the combination of these set of features from Table 3.2 offer a better discriminative model to be learnt in order to classify vessel edges across different image conditions.

3.2.1.4 Local Patch Generation and Vessel Edge Classification

In order to improve the detection of vessel edges belonging to smaller vessels, the extent of edges under consideration are limited to patches of size 128x128 pixels, which is about

Figure 3.4: Three stage patch generation for vessel edge detection

$\frac{1}{3}$ the size of the optic disc, to consider local vessel information. The use of such a patch-based approach avoids biasing the features towards the larger vessels, which could lead to under-detection of the smaller temporal vessels. These patches are formed in three stages as illustrated in Fig. 3.4. In the first stage, non-overlapping patches are defined within the optic disc boundary. These patches are limited to the temporal side of the optic disc, where the smaller kink-containing vessels have a higher occurrence. In the second stage, additional patches are defined with their horizontal centres at the intersection of the patches in the first step to enhance vessel continuity. In the final stage, patches are defined on the remaining regions not encompassed in the prior steps. The edges within each patch are considered independently of other patches, and a 9-dimensional feature vector for each within-patch edge candidate is generated using only local information within that patch. Within-patch edges are classified, and detected vessel edges from a patch are merged with other detected vessel edges and morphologically thinned to obtain the final vessel edge trajectories.

3.2.2 Kinks Detection

A two-step approach is proposed to identify kinking along vessel edges using piecewise probing. The high variability of the shape of the vessel, in particular considering the

Figure 3.5: Simulated vessel edge(red), approximated straight line segment(blue), and kink point(yellow) are detected at (Left) $\gamma = 60^\circ$, (Middle) $\gamma = 100^\circ$ and (Right) $\gamma = 159^\circ$.

presence of sharp kinking points, leads to difficulties in parametrically modeling the vessel trajectory.

Instead, a non-parametric approach is adopted to analyze the vessel course. First, each vessel segment course is subsampled as a series of connected straight lines segments, with a 5 pixel deviation tolerance. This allows isolated edges and spurs to be removed, while the general vessel course still remains sufficiently represented. Next, for each vessel segment, the angle between each two connected lines are then calculated using $\gamma = \cos^{-1}(\frac{\vec{U} \cdot \vec{V}}{\|\vec{U}\| \|\vec{V}\|})$ where $\vec{U}$ and $\vec{V}$ represents two connected vessel line segment. Kink points are identified when the $0^\circ < \gamma < 160^\circ$. A illustration of detected vessel kinks using simulated data is shown in Fig. 3.5.

3.2.3 Optic Cup Feature Point Integration

In the detection of the optic cup in practice, a combination of color and kinking information is used to accurately determine the extent of the optic cup. To form a baseline estimate from previous color-based methods, a system of reference is used, based on clinical nomenclature: Inferior(I), Superior(S), Nasal (N) and Temporal(T). Each of these extrema locations represent an extent of the optic cup boundary defined at $N, S, T, I = \{0, \frac{\pi}{2}, \pi, -\frac{\pi}{2}\}$ with respect to the optic disc centre. Using a previous method [134], a close estimate of the I, S and N regions can be obtained due to a more obvious cup boundary in those areas. However, due to the lack of a distinctly, visible cup edge

Figure 3.6: (Top-Left) Original Image. (Top-Right) Small vessel enhancement using wavelet transform. (Bottom-Right) Detected vessels (in blue), and vessel kinks points (in green) along vessels; (Bottom-Left) Optic cup feature points (in blue) comprising kinks and optic cup color extrema .

in the temporal side, optic cup segmentation tends to undersegment (segmentation leak towards the temporal edge of the optic disc) temporally and is unreliable in this region. In order to circumvent this problem, only the inferior, superior and nasal extents are extracted to form a baseline estimate of the extents of the optic cup in the corresponding region and discard the temporal extrema point.

The detected kinks which are located in the temporal region are added to the remaining I,S,N extrema points to generate a set of optic cup feature points $P = \{x_m\}_{m=1,\dots,M}, M \geq 4$, which incorporate color boundary and contour information. Fig. 3.6 illustrates the process flow used to obtain the integrated optic cup key feature points.

3.2.4 Optic Cup Boundary Detection

To integrate these key points and obtain the optic cup boundary, Wong *et al.* [131] have previously applied the direct least-squares ellipse-fitting method from Fitzgibbon *et al.* (B2AC) [34]. This was motivated by clinical observation that the shape of the cup tends to be an ellipse. However, direct ellipse fitting can lead to erroneous skewed results, particularly if the points in P are not well distributed. Instead, a more suitable approach based on convex optimization is proposed.

Since these key points can be sparse, incorrectly extracted and not adequately representative of an ellipse, using direct least-squares fitting [34] to provide the best ellipsoidal boundary fit for every point in a sparse and noisy set of point distribution can often lead to problems of bias and overfitting [97]. Instead, a minimum volume enclosing ellipse (MVEE), that encapsulates the optic cup feature points, is proposed to obtain the optic cup boundary. The advantage of using MVEE is that the method focuses on determining the ellipse which best encompasses all points in the point distribution with the least volume or, in this case, area. This is, in contrast, to least squares method, which attempts to find the ellipse which minimizes the distance for all points. In this section the minimum volume ellipsoid based on [59, 16, 79, 125] will be introduced.

Generally, for $x \in \mathbb{R}^n$, a full-dimensional ellipsoid $E_{Q,c}$ in $\mathbb{R}^n$, with center $c \in \mathbb{R}^n$, and shape matrix $Q \in \mathbb{S}_{++}^n$ is defined as:

$$E_{Q,c} = \{x \in \mathbb{R}^n | (x - c)^T Q (x - c) \leq 1\} \quad (\text{Eq. 3.5})$$

with the volume of $E_{Q,c}$, denoted as $\text{Vol}(E_{Q,c})$, is given as

$$\text{Vol}(E_{Q,c}) = v_0 \det Q^{-\frac{1}{2}} \quad (\text{Eq. 3.6})$$

where v_0 is the volume of the unit ball in $\mathbb{R}^n$. Suppose the set $S \subseteq \mathbb{R}^n$ is bounded and has a non-empty interior. The minimum volume ellipsoid, also known as the Löwner-John ellipsoid [59], that contains the set S is parameterized as

$$\varepsilon = \{x | \|Ax + b\|_2 \leq 1\} \quad (\text{Eq. 3.7})$$

where $A \in \mathbb{S}_{++}^n$. To compute the minimum volume ellipsoid, where $\text{Vol}(E_{Q,c})$ is proportional to $\det A^{-1}$, the problem can be expressed as:

$$\begin{aligned} & \underset{A,b}{\text{minimize}} && \log \det A^{-1} \\ & \text{subject to} && \sup_{x \in S} \|Ax + b\|_2 \leq 1 \end{aligned} \tag{Eq. 3.8}$$

where $A \in \mathbb{S}_{++}^n$ and $b \in \mathbb{R}^n$. Though the objective and constraint functions are both convex, the constraint function is only tractable in certain special cases [16]. In this work, the problem of finding the minimum volume ellipsoid that encloses $S = \{x_1, x_2, \dots, x_k\} \subseteq \mathbb{R}^n$ will be considered. More specifically, the key interest is determining the minimum volume ellipsoid encloses the convex hull of S . Geometrically, this is equivalent to finding the minimum volume ellipsoid that circumscribes the polyhedron $\text{conv}\{x_1, \dots, x_k\}$.

The convex optimization problem of computing the minimum volume ellipsoid can now be expressed as:

$$\begin{aligned} & \underset{A,b}{\text{minimize}} && \log \det A^{-1} \\ & \text{subject to} && \|Ax_i + b\|_2 \leq 1, \quad i = 1, \dots, k \end{aligned} \tag{Eq. 3.9}$$

where $A \in \mathbb{S}_{++}^n$.

The ellipsoid method was solved by Khachiyan [59] to demonstrate the polynomial time algorithm for linear programming. During the past decades, extensive research efforts have been developed on computing the MVEE. Kumar and Yildirim [62] provide more applications and solutions to the MVEE problem for point sets. In [62], they categorized these solutions by first-order algorithms [59], second order algorithms using interior point methods [118] and a fusion of both [59]. In this paper, the framework and implementation of [79], to solve for minimum volume ellipsoid based on [62, 118] to improve on the work of Khachiyan [59], will be used. The assumption for the formulation guarantees that the minimum volume ellipsoid is full-dimensional [118].

Since S is not centrally symmetric, a "lifting" to S to $\mathbb{R}^{n+1}$ is defined in [118] as

$$S' := \{\pm q_1, \dots, \pm q_n\}, \text{ where } q_i^T = \begin{bmatrix} x_i \\ 1 \end{bmatrix}; \quad i = 1, \dots, k. \tag{Eq. 3.10}$$

The minimum volume ellipsoid (MVEE) of the original set S can be recovered from the intersection of H with MVEE of the lifted points, q_i by,

$$\text{MVEE}(S) = \text{MVEE}(S') \cap H \text{ where } H = \{(x \in \mathbb{R}^{n+1} : x_{n+1} = 1\} \quad (\text{Eq. 3.11})$$

Since S' is centrally symmetric, $\text{MVEE}(S')$ is centred at the origin. This observation provide the "lifted" primal optimization problem, whose solution is used to compute $\text{MVEE}(S)$:

$$\begin{aligned} (\mathbf{P}(S)) \quad & \underset{M}{\text{minimize}} \quad -\log \det M \\ \text{subject to} \quad & q_i^T M q_i \leq 1, \quad i = 1, \dots, k \end{aligned} \quad (\text{Eq. 3.12})$$

$M \in \mathbb{R}^{(n+1) \times (x+1)}$ is symmetric and postive definite

where $M \in \mathbb{R}^{(n+1) \times (x+1)}$ is the decision variable.

The Lagrangian dual of $(\mathbf{P}(S))$ is equivalent to

$$\begin{aligned} (\mathbf{D}(S)) \quad & \underset{u}{\text{maximize}} \quad \log \det V(u) \\ \text{subject to} \quad & e^T u = 1, \\ & u \geq 0, \end{aligned} \quad (\text{Eq. 3.13})$$

where the linear operator $V(u) = \sum_{i=1}^n u_i q_i q_i^T$, $u \in \mathbb{R}^n$ is the decision variable and e is the vector of ones.

Let u^* denote the optimal value of u , and U^* denote the matrix with diagonal entries of u^* . Let P denote the $n \times k$ matrix with points $p_1, \dots, p_k$ being the columns of the matrix. Finally, the dual optimal solution obtained is

$$\begin{aligned} \text{MVEE}(S) &= E_{Q^*, c^*} \\ E_{Q^*, c^*} &:= \{x \in \mathbb{R}^n | (x - c^*)^T Q^* (x - c^*) \leq 1\} \end{aligned} \quad (\text{Eq. 3.14})$$

where

$$Q^* := \frac{1}{d} (P U^* P^T - P u^* (P u^*)^T)^{-1}, \quad C^* := P u^* \quad (\text{Eq. 3.15})$$

Figure 3.7: Comparison of MVEE and B2AC on set of generated typical sparse key points (*Left*) Minimal difference when points are well-distributed and representative. (*Centre*) and (*Right*) show an over-fit by B2AC when the point distribution is not representative and a more robust recovery for MVEE.

3.2.4.1 ‘Toy’ Example

To illustrate the utility of MVEE over B2AC for optic cup feature points, distributions of synthetic data based on some typically observed locations of extracted key points are generated. Both MVEE and B2AC algorithms were applied to these distributions and the ellipsoidal plots are shown in Fig. 3.7. It can be seen that when the points are well distributed and representative, both B2AC and MVEE techniques produce the same ellipse. However, when there are variations in the distribution of the optic cup feature points, the B2AC approach may result in an ellipse which extends beyond the feature points. This is because the B2AC uses a least-squares error measure as the algebraic criterion for ellipse-fitting, as opposed to MVEE, which approximates the ellipse based on the area defined the optic cup feature key points. As such, in contrast, the MVEE approach is able to extract a more acceptable and stable optic cup contour, and is more suitable for this implementation.

Computationally, from [59, 125], it is shown that for small dimensions, in this implementation $d = 2$, the computation complexity of MVEE can be solved in $O(d^{O(d)}m)$, where d denotes the dimensional space, and m denotes the number of points. For B2AC, its computation complexity of is $O(6m)$. As both MVEE and B2AC are used to fit s-

Table 3.3: Cup localization performance comparison

Data	<i>Set_I</i>			<i>Set_{II}</i>			<i>Overall</i>		
Evaluation criteria	m_1	m_2	δ	m_1	m_2	δ	m_1	m_2	δ
<i>KinkMVEE</i>	0.397	0.525	0.138	0.394	0.500	0.135	0.396	0.512	0.136
<i>KinkB2AC</i>	0.447	0.680	0.205	0.432	0.600	0.185	0.440	0.640	0.195
<i>IET</i> [51]	0.566	2.291	0.427	0.555	2.46	0.462	0.560	2.376	0.445
<i>JET</i> [54]	0.492	0.97	0.242	0.554	1.128	0.300	0.523	1.049	0.271
<i>LS</i> [131]	0.476	0.702	0.140	0.471	0.663	0.157	0.474	0.683	0.149
<i>Error reduction to</i> [51]	29.86%	77.08%	67.68%	29.00%	79.67%	70.78%	29.29%	78.45%	69.44%
<i>Error reduction to</i> [54]	19.31%	45.88%	42.98%	28.88%	55.67%	55.00%	24.28%	51.19%	49.82%
<i>Error reduction to</i> [131]	16.60%	25.21%	1.43%	16.35%	24.59%	14.01%	16.46 %	25.04%	8.72%

parse optic cup feature points, where $m \leq 20$ in practice, the differences in computational efficiency is not immediately obvious.

3.3 Experiments

3.3.1 Comparison of methods

The performance of this approach is evaluated on the ORIGA^{-light} dataset, as mentioned in Section 2.4.1. The evaluation metrics tested using non-overlap ratio, m_1 , absolute area difference (m_2), absolute CDR error (δ) are detailed in Section 2.5.1.

KinkMVEE denotes the full implemented approach, while KinkB2AC refers to the use of the direct least squares ellipse fitting on the detected boundary key points. LS denotes the baseline level-set approach by [131]. In addition, existing methods for optic cup segmentation in non-stereoscopic fundus images developed by Inoue *et al.* (IET) [51] and Joshi *et al.* (JET) [54] were implemented and tested on the same test dataset. To ensure that differences in the optic cup segmentation were due to methodology and not to errors in optic disc segmentation, the ground-truth optic disc was used as the common disc segmentation and starting point for all the methods evaluated.

The results of the experiments are shown in Table 3.3. In general, the overall performance metrics suggest that the proposed system of optic cup detection by integrating vessel kinks with color information improves upon the detection results when only color-based methods are used. Comparing IET, which uses a discriminative threshold-based approach, with JET, which uses regional domain prior, it can be observed that, in general,

Figure 3.8: Sample segmentation results (in blue) of different approaches and ground-truth (in red). *From Top :* IET [51], JET [54], LS [131], KinkB2AC, Proposed Method - KinkMVEE. [Best viewed in color]

approaches using domain priors usually offers higher accuracy against similar approaches without prior knowledge. Comparing JET with LS, it is observed that the level-set based approach offers an improved accuracy, due to its use of local gradient information. Furthermore, from Fig. 3.8, it is clear that due to the assumption that the cup is symmetric, JET tends to overestimate the cup boundary on the nasal and temporal regions. Comparing LS to KinkB2AC, it can be observed that the use of vessel kinks as key point regulators reduces both the m_1 and m_2 errors. Lastly, comparing KinkB2AC with KinkMVEE, there is a significant accuracy improvement in all three metrics. In particular, it is observed that using the minimum volume enclosing ellipse (MVEE) to represent the same sparsely distributed key points for optic cup boundary, results in a more robust optic cup estimation.

3.4 Model Training and Computation Speed

3.4.1 SVM Model Training

To train the SVM model parameters used in vessel edge segment classification, images from 5 sample eyes of varying degrees of optic cup visibility were selected. Using the proposed approach, a set of 146 training edges were selected based on patches generated from these eyes. To avoid a bias of the support vectors, the training set is composed of 75 vessel edges and 71 non-vessel edges. The radial basis function (RBF) is used as the kernel function. The edge set was used to train the SVM model using a grid search in a leave-one-out cross-validation approach, and the optimal parameters for C and γ produced an overall cross-validation accuracy of 93.8%. This model was used for the rest of the evaluation for vessel edge classification.

3.4.2 Computation Speed

The computation speed of the proposed method was also evaluated using MATLAB R2013b, using on a quad-core 3.4GHz CPU and 12GB RAM. The Gabor wavelet blood vessel enhancement requires 0.38s, and on average, vessel edge detection and feature extraction took 6.01s. The vessel edge classification requires 0.53s, and kinks detection took 1.9s. The proposed minimum volume enclosing ellipse (MVEE) procedure required

only 0.5s. On average, each image requires 9.3-10.8 seconds to be processed, and is largely dependent on the number of vessel edges detected for each image.

3.5 Discussion

Although experimentally, it was demonstrate that the use of vessel kinks can help improve the optic cup detection, however, it should be noted that there are some limitations and considerations for the effective use of kinking for cup boundary detection. Like many existing optic cup detection methods, there is a need for an accurate detection of the optic disc as a form of initialization for the optic cup detection. Furthermore, in this large database testing, there were no kinks detected in 98 out of 650 images (15.08%). In such a scenario, the approach defaults to using only color information from a single method. This can be combined with other color-based methods.

Even for images with vessel kinks, there were several instances of detection failure. The failure to detect vessel kinks can be attributed to two main aspects, 1) the inability in automatically detect blood vessels reliably, and 2) the absence of physical vessel kinks in the temporal region. In the case of the former, it was found that the proposed blood vessel extraction approach results in inferior performances in poor quality images, such as the examples shown in Fig. 3.9. As highlighted in Sec. 2.3.2.1, these images are examples of occlusion, whereby the media opacity can occur during image acquisition or as a pre-existing medical condition. In the latter case, as shown in Fig. 3.10, there were not physical vessel landmarks present.

A recent approach for optic cup detection using kinking was reported by Joshi *et al.* [55]. In that approach, kinks were the only feature used to determine the cup boundary. In their discussion section[55], it was reported that the largest obstacle faced in the cup detection was mainly attributed to difficulties in finding reliable visual feature extraction. In such cases, their approach will fail to return a cup region. This is a common limitation in such boundary-based approaches, where the lack of reliable boundary features keypoints contributes to a less consistent result. In contrast, in the same scenario, our proposed approach has the added advantage of additional color information as a supplement. However, the limitations of reliable boundary features keypoints still exists.

Figure 3.9: Images with low contrast, arising from media opacity. [Best viewed in color]

Figure 3.10: Absence of physical blood vessel landmarks in the temporal region. [Best viewed in color]

Another key consideration to improve the detection of kink feature landmarks, is to obtain a sufficient size dataset of manually labeled vessel kinks ground-truth. This would also largely help in the accurate modeling of automatic kinks detection with fewer parameter tuning, as well as, provide further possible insights on how outlier vessel kinks may be removed. The lack of interpretation of useful vessel kinks and their inter-relationship, is still a limitation in the existing key point boundary-based approaches.

3.6 Summary

A system of using vessel kinks as regulators of the optic cup boundary from non-stereo fundus images has been proposed. This is based on the clinical protocol that kink in vessels within the optic disc are useful as markers of a change in topology due to the depression of the optic cup. In this approach, kinking detection is focused in the temporal side of the optic disc, where these kinks are most informative. This is in contrast to other color-based indicators which are unreliable due to the reduced visibility of the optic cup in this region. Detected kinks and canonical points were used as feature points to regulate the cup boundary using a convex optimization-based minimum volume enclosing ellipsoid technique. This technique approximates the cup ellipse based on the area bounded by these key points and produces a more stable result. Tested on the ORIGA^{-light} clinical dataset, it is experimentally demonstrated that the use of vessel kinks can help improve cup detection accuracy. However, it is important to note that, while such key points boundary-based approaches follows the clinical protocols and provide explicit visual evidence, these naturally occurring vessel kinks are missed in more than 15% of the images. Although there are some limitations in such boundary-based key point approaches, this is an important step towards the implementation of a more robust automatic system for the screening and early detection of glaucoma using retinal fundus photographs.

Chapter 4

Domain Prior Based Superpixel Propagation for Unsupervised Optic Cup Localization

4.1 Introduction

In the previous chapter, a boundary-based approach using a vessel kink detection model was presented and a review of its strengths and limitations were discussed. The main design consideration of the boundary-based approach was to simulate and model the clinical practice of identifying clinical visual cues for the optic cup boundary. Although it was shown experimentally to improve the cup localization results, several challenges remain. The main challenges faced were the occasional absence of vessel kinks and the inability to automatically detect and interpret these key points reliably when they were present. This limited the performance of the general approach. Furthermore, the lack of manually labeled vessel kinks ground-truth and meticulous parameter tuning and post-processing for accurate modeling increases the complexity of designing a robust key point boundary-based approach. These limitations are common across most boundary-based approaches.

In this chapter, an alternative novel region-based approach is investigated. The proposed method uses unsupervised learning, with the help of domain priors to find hidden structure in unlabeled data for optic cup localization. In contrast to previous supervised approaches [139, 8], this framework requires no training procedure. To the best of the

author’s knowledge there are no other earlier works that performs unsupervised learning to localize the optic cup. Therefore, this chapter presents a significant contribution towards learning-based methods for optic cup localization.

4.2 Methodology and Contributions

Unsupervised learning is a type of machine learning, where, learning is performed on the raw input data without labels, optimization criterion(s), or feedback. Its main goal is to find patterns within the input data and build a representation. Examples of unsupervised learning are dimensionality reduction and clustering. Unsupervised learning is a powerful machine learning technique, and some successful applications includes face recognition [126], object recognition [33], human actions activity [84] and natural language processing [42].

In this work, the state-of-art SLIC (Simple Linear Iterative Clustering) superpixels are used to obtain region segments from the retinal image. Features are extracted from each superpixel unit, and the feature representations are generalized using clustering. As the clusters are not labeled, domain priors are used to provide initialization for further label propagation and refinement.

The contributions here are threefold. First, a new domain prior, optic cup origin is identified. The optic cup origin is derived from the physiological understanding that the central retinal vessels traces its origin from the optic cup before extending to the rest of the retinal. Second, the features of the optic nerve head are extracted from superpixels, which are obtained from low-level grouping and have more natural and descriptive features than pixel based techniques. Third, the domain knowledge comprising optic cup origin and cup pallor, and the extracted features from superpixels are then used to drive a similarity-based label propagation and refinement scheme for the optic cup localization. The framework of the overall approach is illustrated in Fig. 4.1.

4.2.1 Superpixel Segmentation

Instead of pixel-level classification, labels are assigned for each superpixel segment for efficiency. This bottom-up segmentation of superpixels not only reduces the complexity

Figure 4.1: Flowchart of the proposed methodology.

Figure 4.2: Region segments obtained using superpixel segmentation. Each region segment is compact and uniform, while adhering to the image boundaries.

of our problem, but also summarizes image redundancy and provides highly desirable spatial and local boundaries adhering properties.

Superpixels are acquired using the SLIC (Simple Linear Iterative Clustering) algorithm [9] which groups pixels by adapting a k -means clustering approach. Each SLIC superpixel corresponds to a cluster in a five-dimensional color (CIELAB) and image location plane space. The advantages of using the SLIC algorithm to perform superpixel segmentation are 1) its efficiency to generate compact, nearly uniform superpixel segments with $O(N)$ complexity, and, 2) only requiring a single parameter, which is to specify the number of superpixels. In this work, the number of superpixels is specified

to be 1152, which is about 550 pixels per superpixel segment. A example of the typical segmentation results is shown in Fig. 4.2.

4.2.2 Multi-Scale Blood Vessel Extraction

To reduce the effects on rim/cup classification, a multi-scale difference-of-closing vessel extraction algorithm is used. This is first introduced by [102]. A morphological closing operation is defined as dilation followed by erosion. The formulations for the dilation, erosion and closing operations are as follows:

$$\text{Dilation: } A \oplus B = A_1(x, y) = \sup_{i,j \in b} (A(x - i, y - i) + B(i, j)) \quad (\text{Eq. 4.1})$$

$$\text{Erosion: } A \ominus B = A_2(x, y) = \inf_{i,j \in b_1} (A(x - i, y - i) + B_1(i, j)) \quad (\text{Eq. 4.2})$$

$$\text{Closing: } C' = (A \oplus B_2) \ominus B_2 - (A \oplus B_1) \ominus B_1 \quad (\text{Eq. 4.3})$$

where A is the input image, B and B_1 are the structuring elements used for dilation and erosion respectively.

The advantage of this method is its flexibility in quick parameter tuning to extract only the thicker blood vessels, and the vessel extraction is simple and fast, requiring only morphological operations. In the implementation, a disk structuring element of radius of 16 pixels and 3 pixels for the dilation and erosion operations are used respectively. As the images are processed at superpixel scale, a blood vessel mask is generated based on superpixels that overlap at least 70% of the extraction vessels. Figure. 4.3 demonstrates some of the results using the multi-scale blood vessel extraction technique.

4.2.3 Optic Cup Origin Prior - Central Retinal Vessel

Based on discussions with clinical graders, it was understood that a portion of the central retinal vessel trunk that bends inwards, towards the centre of the optic disc, is considered as part of the optic cup. This is derived from the physiological understanding that the central retinal vessels traces its origin from the optic cup before extending to the rest of the retinal.

66
Figure 4.3: *Left:* Green Channel of the input image. *Middle:* Extracted Blood Vessel. *Right:* Generated blood vessel mask.

Figure 4.4: Illustration of optic cup origin with neighborhood prior

In order to estimate this optic cup origin automatically, the vessel junctions are first localized from extracted skeletonised vessels in the optic nerve. The optic cup origin point is then determined based on a weighed mean of these vessel junctions. The weight assigned to the c -th vessel junction is denoted as $w_c = (1 - \sqrt{(\frac{x_c}{a})^2 + (\frac{y_c}{b})^2})^2$, with a and b representing half of the optic disc width and height respectively. A strong optic cup origin region prior is defined by using the superpixels that lies within the vicinity of the optic cup origin point, excluding the blood vessels. Examples of the proposed automatic optic cup origin detection are shown in Figure. 4.4. Once the strong optic cup origin region prior is established, the rest of the superpixels are defined as unlabeled regions.

4.2.4 Superpixel Feature Representation

In this proposed approach, positional, color, texture and appearance information, which are relevant to characterise the cup and the rim, are extracted from each superpixel. In each i -th superpixel, the feature vector $\mathbf{f}_i$ consists of the euclidean distance to the optic cup origin center (d_i), mode RGB colors (r_i, g_i, b_i), mode LAB colors (l_i, a_i, b_i), color moments of the HSV colors (1st to 4th moments) (C_i^H, C_i^S, C_i^V)[116], image moment invariants RGB colors M_i^I and statistical texture descriptors from RGB channels

Table 4.1: Features used

Features	Description	Dimension
Positional	d_i	1
Color	$(r_i, g_i, b_i), (l_i, a_i, b_i), (C_i^H, C_i^S, C_i^V)$	18
Texture	T_i^R, T_i^G, T_i^B	18
Appearance	M_i^I	21

(T_i^R, T_i^G, T_i^B) (average gray level, average contrast, measure of smoothness, third moments, uniformity and entropy). In total, for each superpixel, 58 dimensional features are extracted. A summary of these features extracted is found in Table. 4.1. Features are then standardized to reduce the influence of scale differences among features, which makes the values of each feature in the data have zero-mean and unit-variance.

4.2.5 Label Initialization and Propagation

A similarity-based label propagation is introduced to propagate superpixel labels using measured feature similarity between strong prior superpixels and unlabeled ones. This returns the initial label l of a unlabeled superpixel, defined as

$$l_i = \frac{1}{N} \sum_{j=1}^N s_{i,j}, \quad (\text{Eq. 4.4})$$

where $s_{i,j} = e^{-\frac{d_{i,j}}{\sigma_f}}$ is the similarity between i -th unlabeled superpixel and the j -th strong prior superpixel and σ_f regulates the feature noise sensitivity. The correlation distance between the i -th unlabeled superpixel and the j -th strong prior superpixel $d_{i,j}$ is defined as

$$d_{i,j} = 1 - \frac{(f_i - \bar{f}_i)(f_j - \bar{f}_j)'}{\sqrt{(f_i - \bar{f}_i)(f_i - \bar{f}_i)'} \sqrt{(f_j - \bar{f}_j)(f_j - \bar{f}_j)'}}, \quad (\text{Eq. 4.5})$$

where $\bar{f}_i = \frac{1}{n} \sum_t f_i^t$ and $\bar{f}_j = \frac{1}{n} \sum_t f_j^t$.

After each superpixel get its initial label, a threshold $\theta_p = 0.5$ is used to determine the binary cup labels, *i.e.* $l_i \geq \theta_p$ is cup (1), otherwise is rim (0).

4.2.6 Label Refinement

The initial labels of superpixels are obtained by using optic cup origin prior and feature similarity measurement. Inspired by the bag-of-words method [32], a k-means based label

Figure 4.5: k -Means clustered region segments.

refinement is applied to the feature similarity measure before using pallor prior to refine the labels of each superpixel.

4.2.6.1 K-Means Clustering based Refinement

For all the superpixels within a given disc, excluding the ones corresponding to blood vessels, a k -means ($k = 10$) clustering is used to quantify each superpixel, Fig. 4.5. This quantization allows the cup segments to have variety of representation and is used as a local codebook. Using the k -means cluster index of each superpixel, the refined label l' of the i -th superpixel is defined as

$$l'_i = \alpha l_i + (1 - \alpha)[k_i \in \Omega], \quad (\text{Eq. 4.6})$$

where k_i is k -means index of the i -th superpixel, Ω is the complete set of indexes of strong prior superpixels, $[\cdot]$ is a logical expression, its value is 1 when the condition is true, otherwise it is 0, *i.e.*, $[k_i \in \Omega] = 1$ when $k_i \in \Omega$, otherwise $[k_i \in \Omega] = 0$, and α is a weighted parameter.

4.2.6.2 Pallor Prior based Final Refinement

In this work, pallor prior is utilized as a conservative initial estimate of the cup area. As the optic cup pallor appears as the brightest region within the optic nerve head, empirically taking each superpixel as the unit, the top $p_G = 30\%$ of the highest intensity (denoted as $G_{\max}$) in the green channel (in RGB color space) for the cup pallor, as well as the top $p_L = 25\%$ of the highest lightness (denoted as $L_{\max}$) in the lightness domain (in LAB color space) are used. The green channel is used as for two main reasons, it has the largest dynamic range among the RGB channels, and the optic cup appears most distinct. The Lightness domain of the LAB is used as it is perceptually uniform and closely agrees with the human perception of lightness. The cup pallor prior label, l_i^P of the i -th superpixel, is formed by fusing both thresholding results in the two color channels, calculated as $l_i^P = (l_i^G + l_i^L)/2$. The final label of the i -th superpixel is $l_i'' = [(l_i' + l_i^P) > 0]$.

4.2.7 Region-based Ellipse Fitting

In order to obtain a unique cup region to integrate and represent the superpixel segments, a simple ellipse-fitting method is used. In this implementation, the ellipse-fitted optic cup is represented by its centre and elongation parameters, (u, v, v, ν) .

4.3 Experiments

This approach was validated on the ORIGA^{-light} [147] dataset, of 650 images. Image set S_I and image set S_{II} both consists of 325 images each. The experiments are evaluated using the same criteria as [139] - non-overlap ratio (m_1), relative absolute area difference (m_2), absolute CDR error (δ). Details regarding the dataset and evaluation metrics can be found in Section 2.4.1 and Section 2.5.1 respectively.

4.3.1 Comparison to Prior Arts

The pixel based level-set segmentation method of [131] iteratively seeks the local gradient minimum within the optic disc and assigns pixels within the ellipse-fitted level-set contour as the cup region. The state-of-the-art sliding window based method [139] identifies a

Table 4.2: Cup localization performance comparison

Data Evaluation criteria	S_I			S_{II}			Overall		
	m_1	m_2	δ	m_1	m_2	δ	m_1	m_2	δ
<i>Full using Corr</i>	0.317	0.360	0.100	0.327	0.316	0.112	0.322	0.338	0.106
<i>Full using Eucli</i>	0.349	0.397	0.120	0.359	0.335	0.131	0.354	0.366	0.126
<i>Full w/o Prop</i>	0.433	0.405	0.180	0.463	0.415	0.215	0.448	0.41	0.198
<i>Full w/o Pallor</i>	0.397	0.355	0.120	0.397	0.347	0.131	0.397	0.351	0.126
<i>Full w/o kRef</i>	0.384	0.425	0.122	0.392	0.349	0.139	0.388	0.387	0.131
<i>Full w/o LRef</i>	0.462	0.399	0.140	0.485	0.395	0.163	0.474	0.397	0.152
<i>Pallor Only</i>	0.677	0.644	0.221	0.715	0.675	0.266	0.696	0.660	0.244
<i>pixel based [131]</i>	0.476	0.702	0.140	0.471	0.663	0.157	0.474	0.683	0.149
<i>window based [139]</i>	0.268	0.315	0.091	-	-	-	-	-	-
<i>Relative error reduction to [131]</i>	33.4%	48.72%	28.57%	30.57%	52.33%	28.66%	32.07%	50.51%	28.86%
<i>Relative error reduction to [139]</i>	-18.28%	-14.29%	-9.89%	-	-	-	-	-	-

cup as a whole, by ranking all the cup candidate regions obtained with sliding windows and produces a single detection result. In contrast, the proposed region-based approach utilizes domain prior from the optic cup origin as initial cup labels, which are then propagated, using a similarity-measure, and refined to identify the cup region.

As shown in Table 4.2, compared with the pixel based method [131], the proposed approach reduces the cup localization error significantly in all three evaluation criteria. Unlike the pixel-based method [131] which considers only local gradient information, the proposed method is more superior as it exploits additional regional features and domain priors, making it less susceptible to local optimal conditions. Compared with the window based method [139], the proposed method had a minor trade-off in cup localization accuracy. However, it is worth mentioning that training samples are needed in [139] and the listed results are the average of both training and testing samples.

4.3.2 Comparison of methods within proposed framework

(*Full using Corr*) and (*Full using Eucl*) denotes the fully proposed method using correlation distance and Euclidean distance as their similarity metric respectively; (*Full w/o Prop*) refers to the proposed framework excluding the similarity-based propagation stage; (*Full w/o Pallor*) excludes the pallor prior based refinement stage; (*Full w/o kRef*) excludes K-Means clustering; (*Full w/o LRef*) excludes the label refinement stage; and (*Pallor Only*) uses the pallor prior solely to localize the optic cup.

From the results listed in Table 4.2, it is observed that the correlation distance measure achieves better accuracy than the Euclidean distance for label propagation in this specific

Figure 4.6: Final segmentation (blue) against manual ground-truth segmentation (red).

application. Comparing *(Full using Corr)* and *(Full w/o Prop)* shows the effectiveness of propagation stage, and the results of *(Full w/o LRef)* demonstrate the advantages of Label refinement, using the combination of both K-Means and Pallor prior, to reduce localization errors. It can also be concluded that the K-Means based refinement aids more in reducing m_2 and δ errors, while Pallor prior is more beneficial in reducing m_1 error. Due to the complementary effect of both refinements, their combination has an add-on effect to further improve the overall optic cup localization. A sample of the experimental results from the proposed unsupervised method is shown in Fig. 4.6.

Figure 4.7: Cup Detection with different k values for clustering.

4.4 Parameter Selection and Other Considerations

4.4.1 K-means Parameter Selection

A further examination is done to explore the performance and stability of the proposed approach with different settings. For clustering, the number of clusters, k , is a primary parameter which influences both clustering and the refinement process. As shown in Fig. 4.7, one can observe that the errors become stable and has the best trade-off at $k = 10$.

When the number of clusters becomes too little (*i.e.* $k < 10$), the size of each cluster region becomes larger. This results in an improvement in the overlap error, m_1 , but a higher absolute CDR error (δ), due to ambiguously labeled superpixel segments being clustered in the same region. Similarly, when (*i.e.* $k > 10$), the features in the oversegmented regions becomes less distinctive, leading to a higher overlap error, m_1 .

4.4.2 Computation Speed

The computation speed of the proposed method was also evaluated using MATLAB R2013b, using on a quad-core 3.4GHz CPU and 12GB RAM. The morphological blood vessel extraction process takes 0.85s, and superpixel segmentation requires 0.67 sec. The

Table 4.3: Features comparison Cup localization performance

Data	S_I			S_{II}			Overall		
Evaluation criteria	m_1	m_2	δ	m_1	m_2	δ	m_1	m_2	δ
<i>Full Approach*</i>	0.307	0.378	0.092	0.308	0.297	0.104	0.307	0.337	0.098
<i>Full Approach</i>	0.317	0.360	0.100	0.327	0.316	0.112	0.322	0.338	0.106
<i>Relative error reduction</i>	3.15%	-5.00%	8.00%	5.81%	6.01%	7.14%	4.66%	0.30%	7.55%

most computationally heavy module in this approach, is the feature extraction for each superpixel, which requires 14.96 secs. In addition, k -means requires 1.4 secs, and the label initialization and propagation took 0.46 secs. On average, each image requires 19.41 seconds to be processed.

4.4.3 Alternative Features

Recently, a supervised superpixel-based intra-image strategy is proposed in [138] with state-of-the-art accuracy in cup localization error. In this section, an investigation of the performance of this proposed unsupervised learning framework using the histogram-based features exemplified in the intra-image[138] approach is experimentally validated. The details of these features are documented in Sec. 5.2.2.

(*Full Approach**) describes the proposed approach using the alternative features described in [138], while (*Full Approach*) denotes the proposed approach using the features described in Sec. 4.2.4. As can be observed from Table. 4.3, in general, comparing the use of different features in the same unsupervised framework, the alternative features proposed in [138] provides an overall relative error reduction in all three evaluation metrics. In particular, the new features offers a performance improvement with an overall relative error reduction of 4.66%, and 7.55% in non-overlap ratio, m_1 , and absolute CDR error, δ . A possible reason for this could be attributed to the higher dimension of the new features (*i.e.* 768 dimension vs 58 dimensions), which provides better separation capabilities. This experiment also highlights the flexibility of the proposed framework in incorporating new and more discriminative features.

4.5 Summary

Unlike the boundary-based approach in Chapter 3, in this chapter, a novel unsupervised learning region-based framework is presented. The motivation of this work is to explore

the use of learning-based methods for optic cup localization. In this approach, region segments are obtained using superpixel segmentation, which reduces image redundancy and allows regions to be learned and classified. Using domain priors extracted from the primary structures of the optic disc, a new optic cup origin domain prior is proposed to provide initialization for the cup labels. A similarity-based propagation and refinement scheme is then used to provide additional contexture information to refine the region labels. An optic cup region is obtained after a simple region-based ellipse fitting operation. Experimentally, it was demonstrated that the proposed learning-based method is capable of providing performance accuracy that is comparable to against existing state-of-the-art image-processing method. In addition, each contribution in this framework is separately verified and discussed. A further examination on the parameter selections and alternative features were also reviewed. The latter was provided to highlight the flexibility of this framework in adopting different types of features.

Chapter 5

Multi-Scale Superpixel Classification for Supervised Optic Cup Localization

5.1 Introduction

Recently, the use of superpixels for feature classification has been widely adopted in retinal imaging for diseases like pathological myopia [140], age-related macula degeneration (AMD) [135] and glaucoma [121, 138, 24]. In particular, for optic cup detection, the supervised superpixel-based classification [138] [24] has shown to achieve state-of-the-art performance against other existing approaches [131][55][144]. In [24], a centre surround statistics(CSS) feature descriptor based on biological inspired features were used as features for a non-linear support vector machine (SVM) to perform the task of optic cup segmentation. In [138], besides using the general supervised superpixel classification method, domain prior based on intra-image learning, where each model is learnt from samples from each test image without pre-labeling, is introduced to overcome the illumination differences between training and testing images. A refinement scheme is then used to include structural priors and local context for the final cup detection to further boost its performance. However, the approach assumes a CDR ranging between 0.2 to 0.9, which leads to low accuracy in small and large cup localization. Furthermore, at larger superpixel scales, this will lead to fewer training samples, thus, affecting the overall performance. In Chapter 4, an unsupervised approach[121] is proposed using domain priors, derived from cup pallor and optic cup origin, together with features extracted

from superpixels. A label refinement using k -means clustering is then used to achieve the task of optic cup localization. Nonetheless, the accuracy of the supervised approaches [138] [24] outperforms the unsupervised method.

In this chapter, two limitations in the existing superpixel-based framework are studied and identified as the motivation for this work. First, an alternative strategy to reduce the effects of varying illumination between images is proposed and experimentally validated. Next, a general multi-scale superpixel classification strategy is proposed to improve the accuracy and robustness for optic cup localization. This work is build upon the existing single scale superpixel classification framework in [138] and [121](in Chapter 4), and incorporated two novel contributions. First, a multiple model selection and integration scheme using sparse learning is proposed to provide stability in performance variations, arising from repeated random samples selection of training data. The intuition of sparsity is that many natural signals are sparse and can be concisely represented when expressed in the correct basis [18]. Applications employing sparse learning were demonstrated in data compression[18], object detection [10] and feature selection [83]. Similarly, in this work, sparse learning is used to concisely select, represent and integrate the multiple base modes. Second, multiple superpixel resolutions are integrated for better optic cup boundary adherence and localization. This general framework is also adaptable to other similar classification-based ocular disease CAD applications.

5.2 Methodology and Contributions

Similar to the setup of the work in [138], a disc image is used as the input image. First, blood vessels are extracted and the input disc image is enhanced by a contrast normalization scheme. Next, the contrast enhanced image is then segmented into superpixels and blood vessels which overlap with the superpixels are removed. Features are then extracted across multiple superpixel scales, and multiple classification models are trained for each respective scale. To obtain an unique label for each pixel with higher accuracy, optimal superpixel classification models are selected and integrated using a sparse learning approach. The final optic cup area is then identified by using an ellipse to enclose all the pixels predicted as ‘cup’. The framework of the proposed approach is shown in Fig. 5.1.

Figure 5.1: Flowchart of the proposed learning framework.

5.2.1 Pre-processing

5.2.1.1 Multi-Scale Blood Vessel Extraction

Similar to the previous unsupervised learning approach (in Chapter 4), blood vessels are first extracted using the multi-scale approach proposed by [102]. Details of the approach and parameters used can be found in Section. 4.2.2. As the images are processed at superpixels scale, a blood vessel mask is acquired using superpixels which overlap the extracted vessels by at least 70%.

5.2.1.2 Contrast Normalization

The influence of illumination variances between training and learning images can affect the performance of the trained machine learning models. To reduce these effects, a pre-processing using histogram stretching is performed. For each *RGB* channel in a disc image, the image intensities are represented as a histogram and stretched to $[0, 255]$ to expand the dynamic range and achieve consistency across all the images. This pre-processing histogram normalization accomplishes two goals, 1) it provides better illumination alignment for all images, 2) it helps to enhance the contrast between the optic cup and rim.

5.2.2 Multi-Scale Superpixel Feature Representation

As presented in Section. 4.2.1, SLIC superpixels offers the advantage of reducing the image redundancy while providing desirable spatial and local boundary adherence properties. Using a single parameter, the number of superpixel segments can be generated for an input image.

Experimental studies in [138] demonstrated the effects of different superpixel scales and their resulting performance accuracy (see Fig. 5.2). In particular, it was found that larger superpixel segments provides richer and more discriminative features compared to smaller superpixels, but are prone to over-segmentation in ambiguously-labeled boundary superpixels. In contrast, smaller superpixel regions offers less distinctive features but are able to adhere to the cup boundary with greater precision. This trade-off is unavoidable in a single scale classification framework. Instead of using a “one-size-fits-all” strategy, this

Figure 5.2: Cup detection errors with different number of superpixels. Figure and caption is reproduced from with permission from [138].

approach proposes to merge multiple superpixel scales segments to unify their advantages for better boundary fit.

Using the superpixels as building blocks, features can be extracted to describe each segment. Examples of features that had been used to represent superpixels in object recognition and human body segmentation, includes size, shapes, texture, colors, boundary contours, and thumbnail appearances [124, 78]. Following the features extracted in [138], for each j -th superpixel at superpixel scale S_c (*i.e.* $S_c \in \{S_1, \dots, S_\beta\}$), a feature vector $\mathbf{f}_j^c$ which consists of position information (coordinates and distance of each superpixel centroid with respect to the disc centre) denoted by (x_j, y_j, ρ_j) , mean RGB colors (r_j, g_j, b_j) and a 256-bin histogram (h_j^r, h_j^g, h_j^b) for each R, G, B channel is extracted. The features are then each normalized to the range of $[0, 1]$ using a L_1 -normalization of the histogram, to avoid the effects of magnitude differences between features. Finally, for each scale S_c , the feature matrix of all superpixels $\mathbf{F}^c$ can be obtained.

5.2.3 Multi-Scale Superpixel Classification

At each scale S_c , α linear SVM models are trained to predict each test superpixel to be either disc(-1) or cup (+1). For each test superpixel, with feature $\mathbf{f}_j^c$, using a pre-learned

linear SVM model $\mathbf{w}_j^c$ at scale S_c , the predicted label can be obtained by $(\mathbf{w}_j^c)^\top \mathbf{f}_j^c$ (*i.e.*, the superpixel is classified as cup(+1) when $(\mathbf{w}_j^c)^\top \mathbf{f}_j^c \geq 0$ with this pre-learned model, otherwise it is classified as disc(-1)).

Using a one-time random sampled superpixels with manual labels (determined by an overlapping ratio with ground truth cup), a base model can be learned using a linear SVM. However, since one time random sampling only learns a model which represents a small area in the feature space, its performance may vary significantly. To overcome this bias, multiple base models can be obtained by repeated random sampling. Furthermore, in order to avoid unbalanced training data distribution, equal positive and negative samples are selected for each training round. In total, $\alpha\beta$ base models are trained for β scales. At each scale, the α base models can be used to get α predictions simultaneously without much additional cost. This is performed simply by concatenating each single model projection vector $\mathbf{w}_j^c$, as columns, to form a projection matrix $\mathbf{W}^c$. The prediction labels for a single scale can then be found by $(\mathbf{W}^c)^\top \mathbf{F}_j^c$.

5.2.4 Model Integration Strategy

5.2.4.1 Implementation I : Majority Voting

To integrate the multiple base models, a majority voting strategy is included as a baseline comparison. Although this scheme is known for its simple implementation, it has been found to be just as effective as other complicated methods. Majority voting does not assume any prior knowledge on the individual classifiers' accuracy and requires no training. In this framework, this strategy is also suitable as the individual classifiers can be assumed to be independent.

5.2.4.2 Implementation II : Model Selection based on Sparsity Constraint

In the training phases, for each pixel, $\alpha\beta$ base models are used to classify it as disc or cup. To obtain a unique and accurate classification with minimal computation cost, the following problem is solved:

$$\min_{\boldsymbol{\omega}} \sum_{a=1}^{\tau} \|\mathbf{l}_a - \boldsymbol{\omega}^\top \mathbf{d}_a\|_2^2 + \lambda \|\boldsymbol{\omega}\|_1 \quad (\text{Eq. 5.1})$$

where $\mathbf{d}_a$ is the $\alpha\beta$ predictions from the linear SVM base models and, l_a is the training label. The first term is to minimize the prediction error, and the second term is regularized by λ to enforce the sparsity of $\boldsymbol{\omega}$, which means, only a few linear base models are selected for the unique final prediction for each pixel. In this implementation, the SLEP [69] toolbox is used to train the integration model. The solution of $\boldsymbol{\omega}$, denoted as $\boldsymbol{\Omega}$, indicates which base models are chosen to generate the final feature D_s of each pixel, *i.e.*, the i -th base model is selected when $|\omega_i| > 10^{-3}$. A unique integration model Ψ is then obtained by consolidating the weight parameters of the selected base models in $\boldsymbol{\Omega}$. Likewise, it is worthwhile noting that the majority voting scheme assumes $\boldsymbol{\omega} = \frac{1}{\alpha\beta} \times \mathbf{1}$.

In real applications, only a small number of selected models are used to compute the intermediate results for making final classification. This allows a boost to the performance without much additional cost.

5.2.5 Testing Phase

To summarize, in the testing phase, four steps are sequentially performed to obtain the final pixel-level label from the original image. 1) Preprocessing for blood vessel extraction and contrast enhancement, 2) Superpixel segmentation, blood vessel removal and feature extraction at multiple scales, $S_c \in \{512, 1024, 2048\}$, 3) Using the selected base models, a feature vector for the s -th pixel, $\mathbf{D}_s$, is formed by concatenating the respective classification decision values of selected models indicated by $\boldsymbol{\omega}$. 4) The integration model $\boldsymbol{\Omega}$ is then used to assign each pixel with a unique label $\boldsymbol{\Omega}^\top \mathbf{D}_s$.

As the given ground-truth for optic cup segmentation is an elliptical reference [147], after obtaining the final labels for each pixel, a minimum ellipse that encloses all positive labeled pixels is computed to produce the final optic cup boundary estimate. Another major reason to introduce ellipse fitting is to include superpixels lying on the blood vessels which had been previously removed. Moreover, the use of ellipse-fitting to produce an accurate single optic cup region has also been widely used and reported in related works [131, 139, 138, 24, 121]. In this implementation, the ellipse-fitted optic cup is represented by its centre and elongation parameters, (u, v, ν) . This is also consistent with the previous approach introduced in Chapter. 4. The major steps of the proposed approach is illustrated in Fig. 5.3. Pseudo codes for both the learning and testing phases can be found in Appendix. C.

Figure 5.3: Overview of the major steps in this approach. *From Left :* Contrast Enhancement; Multi-scale Superpixel segmentation, vessel removal and feature classification; Multiple base model training and selection; Model integration and Ellipse-fitting. [Best viewed in color]

5.3 Experiments

5.3.1 Experimental Setup

This approach is implemented in MATLAB R2013b, using on a quad-core 3.4GHz CPU and 12GB RAM. A total of 650 2-D fundus photographs from the online clinical dataset

ORIGA^{-light} [147] were used for the experiments. They were divided in two sets Set_I and Set_{II} , which consisted of 168 images with glaucoma and 482 normal images. Ground-truth segmentation of the optic cup, provided as an ellipse, were drawn manually by a team of grading experts. Details on this dataset can be found in Sec. 2.4.1.

In the learning phase, 150 images from image set Set_I were used as the training set. This setting is consistent with [139][138]. A 2-fold cross-validation was performed to determine the optimal parameters for the linear SVM classifiers by setting the parameters as $C \in \{10^{-3}, 10^{-2}, \dots, 10^2, 10^3\}$. Similar to [138], the optimal SVM parameter was determined to be $C = 100$, and the remaining models were trained by repeated random sampling, using the same optimal regularization parameter. A total of $\alpha = 100$ models were trained for each superpixel scale, $S_i \in \{512, 1024, 2048\}$. For sparse model selection, the optimal parameter was chosen from $\lambda \in \{1024, 2048, \dots, 32768\}$ by using cross-validation.

5.3.2 Evaluation Criteria

As the fundamental basis of this approach is based on superpixel classification, an assessment of the multi-models' superpixel classification performance, at superpixel scale precision is performed. Similarly, to evaluate the localization of the unique optic cup for a given retinal image, the pixel-scale evaluation metrics introduced in Sec. 2.5.1 are measured.

5.3.2.1 Superpixel Classification - for superpixel scale precision evaluation

ROC curves presents a comprehensive representation to summarize the accuracy of the superpixel classification predictions of our learning models, where each point on the curve represents the true-positive rate and false-positive rate.

As the linear models are trained using partial images from image set Set_I , Area Under ROC (AUC) and a balanced accuracy with a fixed 85.0% specificity are assessed on image set Set_{II} . The details of the evaluation criteria for superpixel classification are introduced in Sec. 2.5.2

5.3.2.2 Cup Localization - for pixel scale precision evaluation

Cup localization error is evaluated using the non-overlap ratio (m_1), absolute area difference (m_2), and absolute CDR error (δ), as introduced in Sec. 2.5.1

It is worth mentioning that m_1 is the only evaluation metric related to cup localization accuracy. This is also widely used for accuracy evaluation in general object detection challenges [29]. Absolute CDR error (δ), however, is associated to glaucoma diagnosis and used in many previous works [139][138][121], thus is also considered in this work, but m_1 is the dominant criteria in the experiments. The absolute area difference (m_2) is tabulated in the final performance comparison table for a fair evaluation across all previous works.

The experiments using these criteria are carried out on Set_I and Set_{II} , and compared with existing state-of-the-art methods. This evaluation is comparable to many real applications, where the main objective is to localize the unique cup as accurately as possible to the ground truth. Additionally, the influence of parameters and computation speed of the overall performance are also reported and discussed.

5.3.3 Experiments to validate motivations

Two sets of experiments were performed to identify two current limitations in the existing superpixel-based cup localization framework. In the first experiment, the effects of illumination variations between images are examined to verify if the contrast normalization scheme mitigates these influences. In the second experiment, the consequences of random sampling of training data, and its influence on the accuracy in optic cup localization are studied.

5.3.3.1 Motivation 1 : Alternative strategy to reduce illumination variations in replacement of intra-image learning

In the existing superpixel-based approach of [138], an intra-image learning is proposed to avoid illumination inconsistencies between training and testing images. Though this is highly desirable, the approach assumes that each image has a CDR of between 0.2 to 0.9. In most cases, this assumption may be valid, however, in extreme cases of small ($CDR \leq 0.2$) and large ($CDR \geq 0.9$) optic cups, the training samples becomes insufficient.

Table 5.1: Performance comparison between proposed contrast normalization and intra-image learning

	<i>Overall (Set_I & Set_{II})</i>	
	m_1	δ
<i>Proposed Contrast Normalization</i>	0.263	0.094
<i>No Contrast Normalization</i>	0.287	0.108
<i>Pre-Learned [138]</i>	0.289	0.089
<i>Intra-Image [138]</i>	0.267	0.081

Furthermore, the intra-image learning approach is unable to localize CDR sizes beyond its assumed range. Instead, a histogram-based contrast normalization approach is proposed, as an alternative to intra-image strategy, to reduce this inter-image illumination variations.

From Fig. 5.2, it was observed that the cup localization results performs most favorably at superpixel scale 2048. Table 5.1 compares the results when the proposed contrast normalization approach was adopted and was not adopted within the same experimental framework. The mean errors of the $\alpha = 100$ models at superpixel scale 2048 are used in this experiment for a fair comparison. The results show that the proposed normalization pre-processing method improves feature discriminative power, and is able to boost the overall cup localization performance in m_1 and δ . Comparing the implementation without contrast normalization against the pre-learned approach in [138], a similar m_1 performance is achieved with, however, a slightly higher δ . This could be due to the performance variations from the models trained with randomly selected samples. Compared to the intra-image learning, this approach achieves a relative error reduction by 1.50% in overlap error (m_1) but has a trade-off in 16.0% δ increment, without using the proposed multiple models and multiple scale fusion strategy. Visually, as shown in Fig. 5.4, the consistency in the dynamic range from this approach also helps to enhance the contrast between the optic cup and rim.

5.3.3.2 Motivation 2 : Performance variations of SVM models trained with random sampled data

As observed in Fig. 5.2, the cup detection error varies at different superpixel scales. Similarly, it is observed in this experiment that, the cup localization performance varies

Figure 5.4: Effects of proposed contrast normalization. (*Top*) Original image, (*Bottom*) Contrast normalized images. Finer details of the tissues becomes clearer and distinct.

with different random training data samples. To illustrate these variabilities, a box plot is computed for m_1 and δ errors for all the $\alpha\beta$ models in this framework (Fig. 5.5).

To further explore and understand the performance variabilities, 3 measures (classification prediction accuracy, overlap error m_1 on training set and overlap error m_1 on testing set) for all the α training models in the same scale are tabulated. Each model is ranked according to their respective measures, and sorted according to their prediction accuracy, Fig. 5.6(*Top*), and training m_1 error, Fig. 5.6(*Bottom*).

From the two plots, it can be observed that a trained model with high prediction accuracy does not necessarily result in a high pixel-level cup localization performance, m_1 . Likewise, models that perform well in m_1 during training, does not produce the best overall m_1 results. However, a trained model with poor prediction accuracy or high m_1 training error, often results in weak overall m_1 localization performance. This further highlights the challenges in finding optimal models for accurate cup localization, where models selected based on high accuracy or low training overlap error does not guarantee the performance for cup localization in testing.

Figure 5.5: Boxplot illustrating variance in localization accuracy.

5.3.4 Superpixel Classification Performance

In this section, an evaluation of the trained multi-models' superpixel classification performance for each superpixel scale (S_1, S_2, S_3) is performed. To avoid bias from the training samples used, the assessment is focused on the 325 images from Set_{II} of the ORIGA^{-light} dataset.

For clarity, base models from each superpixel scale will be modeled as a single ROC curve to represent the 100 ROC curves. For each scale, the single ROC curve is found by calculating the mean of all true positive points at each false positive interval from the 100 ROC curves. The AUC derived from the single ROC curve is also found to be about the same as the mean AUCs calculated from each base model, as the variances between each base model's performance is minimal. Next, similar to our proposed framework which integrates multiple models, a majority voting scheme, which is simple and has been shown to be effective in many applications, is used as a baseline comparison. As the integration model assigns a label for each pixel, in order to obtain a distinct decision value for a given superpixel segment, an averaging operation on the predicted labels of all pixels within each superpixel unit is performed. Lastly, the performance of our model integration approach is also compared.

Figure 5.6: Rank-sorted performance plots for trained models. *Top* : Models sorted by prediction accuracy. *Bottom* : Models sorted by training m_1 error. Trained models of Scale 1024 are shown here. [Best viewed in color]

Results of the superpixel classification performances are shown in Fig. 5.7, 5.8, 5.9 and Table 5.2. It is observed that the classification performances for single model is the

Table 5.2: Performance of Superpixel Classification

Evaluation Criteria	Superpixel Scale		
	$S_1 = 512$	$S_2 = 1024$	$S_3 = 2048$
$AUC_{SingleModel}$	0.950 ± 0.00088	0.955 ± 0.00078	0.957 ± 0.00083
$AUC_{MajorVote}$	0.948	0.946	0.945
$AUC_{ModelSelect}$	0.957	0.957	0.957
$\bar{P}_{SingleModel}(\%)$	90.163 ± 0.00022	90.726 ± 0.00020	90.935 ± 0.00021
$\bar{P}_{MajorVote}(\%)$	90.992	90.979	91.068
$\bar{P}_{ModelSelect}(\%)$	91.041	91.070	91.157

highest at superpixel scale $S_3 = 2048$, Fig. 5.7, for both AUC and balanced accuracy. Using multiple models, the *Majority Voting* scheme, which assigns *equal weights* to all trained models, is shown to have improved performances at 85% sensitivity, where the test samples are easier to differentiate and both the optimal and sub-optimal models have consistent decisions. This also demonstrates that multiple model integration can improve the accuracy, as compared to single model approaches. However, in test samples that are more challenging and ambiguous, the optimal and sub-optimal models have different decisions. This is evident in Fig. 5.8, where the tail end of the ROC curves starts to taper at a false positive rate of 0.3, leading to a decrease in overall AUC performance when compared against single model approaches. Compared to *Majority Voting*, the proposed model selection scheme, which *selects and integrates optimal models with different weights*, produces consistent predictions for both easy and challenging test cases, thus improving both AUC and balanced accuracy.

5.3.5 Comparison with existing methods.

The cup localization performance within the proposed framework and against several other published existing techniques is compared. *Majority Voting* denotes using the majority voting integration strategy, as detailed in Section. 5.2.4.1. The mean performance of all the trained $\alpha\beta$ models, and the mean performance of the models in each superpixel scale, S_1, S_2, S_3 are also evaluated. Other established approaches such as *window-based* [139], *pixel-based* level set [131] and superpixel-based methods (*Intra-image* [138], *Pre-learned* [138], *Unsupervised* Chap. 4) were also included in this experiment. The evaluation performance of these methods are listed in Table 5.3.

Figure 5.7: Superpixel classification performance comparison in terms of ROC curves, at different scales ($S_1 = 512$ in blue, $S_2 = 1024$ in red, $S_3 = 2048$ in green. Mean performance of single model.

Figure 5.8: Superpixel classification performance comparison in terms of ROC curves, at different scales ($S_1 = 512$ in blue, $S_2 = 1024$ in red, $S_3 = 2048$ in green). Performance of Majority Voting.

Figure 5.9: Superpixel classification performance comparison in terms of ROC curves, at different scales ($S_1 = 512$ in blue, $S_2 = 1024$ in red, $S_3 = 2048$ in green). Performance of the proposed model integration.

Figure 5.10: Illustration of localized cup boundary at superpixel scale $S_1 = 512$ (*Top-Left*), $S_2 = 1024$ (*Top-Right*), $S_3 = 2048$ (*Bottom-Left*), Proposed integration approach (*Bottom-Right*). The ground-truth boundary is shown in blue. [Best viewed in color]

In general, it is observed that the superpixel-based approaches offers a significant performance improvement compared to pixel-based [131] and window-based [139] methods. Comparing the mean performance of the base models (mean of $\alpha\beta$ models, means of S_1 , S_2 and S_3 models) with existing superpixel-based cup localization methods, it can be observed that the proposed alternative contrast normalization method improves feature discrimination power and provides an improvement in cup overlap error, m_1 , but has a slight trade-off in the absolute CDR error δ .

In addition, the mean performances at each individual scale (S_1, S_2, S_3) was almost similar. A similar conclusion was made in [138], where it was observed that the cup localization accuracy is comparable when the size of each superpixel segment ranges between 80-350 pixels. The misclassification increases when the size falls out of that range.

Although from Table 5.3, the overall results of the three superpixel scales are similar; but for a specific test image with comparable m_1 results at different scales, the localized cup boundaries may not appear as identical. The proposed multiple scale fusion, however, can further reduce the localization error with closer boundary fit. An example of this can be seen in Fig. 5.10.

Compared with the mean models and existing single superpixel-scale approaches ([139],[138],[121]), the combination of multiple models from multiple scales, as used in *Majority Voting* and in the proposed model selection approach, offers an improvement in cup localization accuracy. This demonstrates that the combination of multiple scales and multiple models has a complementary effect in reducing localization errors and improving classification accuracy. However, from Fig. 5.5, it can be observe that individual trained models have large variances. These performance variations from individual models is mainly caused by the bias of randomly selected training samples. Multi-model integration is able to improve localization performance, mainly because it has an equivalent effect of increasing the training sample density and reducing the sampling bias.

Compared to the *Majority Voting* scheme, the proposed model selection approach yielded improvements in both performance criteria. The error reduction is due to the learnt model selection strategy which is able to select and integrate optimal models across different scales. This result also suggests that the use of sparse learning does

reduce redundancy and improve the classification performance. Furthermore, measured against the current state-of-the-art intra-image superpixel-based method [138], the proposed model selection approach offers robust cup localization with lower or similar cup localization performance at minimal additional computation cost. An illustration of the sample results from the experiments is shown in Fig. 5.11.

Figure 5.11: Sample Results. *From Top* : Original Image, Contrast Normalized, $S_1 = 512$, $S_2 = 1024$, $S_3 = 2048$, Majority Voting, Proposed Method. [Best viewed in color]

Table 5.3: Cup localization performance comparison

Data	Set_I			Set_{II}			Overall (Set_I & Set_{II})		
Evaluation criteria	m_1	m_2	δ	m_1	m_2	δ	m_1	m_2	δ
Proposed	0.238	0.292	0.079	0.259	0.258	0.082	0.248	0.275	0.081
Majority Voting	0.257	0.357	0.088	0.260	0.284	0.084	0.259	0.320	0.086
Mean performance of all $\alpha\beta$ models	0.265	0.405	0.097	0.261	0.318	0.089	0.263	0.362	0.093
Mean performance of $S_1 = 512$ models	0.266	0.403	0.097	0.262	0.318	0.089	0.264	0.361	0.093
Mean performance of $S_2 = 1024$ models	0.264	0.400	0.096	0.261	0.316	0.089	0.263	0.358	0.093
Mean performance of $S_3 = 2048$ models	0.266	0.412	0.098	0.260	0.320	0.090	0.263	0.366	0.094
pixel based [131]	0.476	0.702	0.140	0.471	0.663	0.157	0.474	0.683	0.149
window based [139]	0.268	0.315	0.091	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
Unsupervised Superpixel based [121]	0.317	0.630	0.100	0.327	0.316	0.112	0.322	0.338	0.106
Pre-learned Superpixel based [138]	0.277	0.314	0.087	0.301	0.285	0.091	0.289	0.300	0.089
Intra-Image Superpixel based [138]	0.265	0.313	0.079	0.269	0.267	0.082	0.267	0.290	0.081
Error reduction to [131]	50.00%	58.40%	43.57%	45.10%	61.09%	47.77%	47.68%	59.74%	45.64%
Error reduction to [139]	11.19%	7.30%	13.19%	N.A.	N.A.	N.A.	N.A.	N.A.	N.A.
Error reduction to [121]	24.92%	53.65%	21.00%	20.80%	18.35%	26.79%	22.98%	18.64%	23.58%
Error reduction to Pre-learned [138]	14.08%	7.01%	9.20%	13.95%	9.47%	9.89%	14.19%	8.33%	8.99%
Error reduction to Intra-Image[138]	10.19%	6.71%	–	3.72%	3.37%	–	7.12%	5.17%	–
Error reduction to Majority Voting	7.39%	18.21%	10.23%	0.38%	9.15%	2.38%	4.25%	14.06%	5.81%

5.4 Discussion on Parameter Settings

In this section, the influence of the regularization parameter for sparse model selection and the computation speed of this proposed framework are reviewed.

5.4.1 Sparsity Regularizer λ

Identifying and using only the effective models can provide higher localization precision with minimal computational cost increase. The role of the regularizer λ , is to penalize and enforce sparsity on ω for optimal model selection. Fig. 5.12 shows the selected SVM base models across different superpixel scales, and its associated weights as λ increases. The influence of different λ settings on the overall performance is shown in Table 5.4.

From Table 5.4, it is observed that $\lambda = 4096$ provides the best overall results, selecting 6% of the models, roughly 2% at each scale. As λ increases or decreases, the number of selected models changes. When too few models are selected, the generalization condition becomes overly strict; conversely, selecting too many models introduces redundancy, thereby, also significantly affecting the overall localization accuracy.

On closer inspection of Fig. 5.12, it can be observed that the weight of the selected model(s) in $S_3 = 2048$, is the highest. A similar observation is reported in [138], where a superpixel of scale 2048 provides the lowest localization errors. However, as demonstrated experimentally, the linear combination of selected optimal models from multiple scales is able to enhance the localization performance with higher classification accuracy, provide better adherence to the optic cup boundary and improved performance stability, with minimal additional computational costs. Experimentally, it was also noted that these

Table 5.4: Influence of Sparsity Measure

λ	Selected Models				m_1	δ
	S_1	S_2	S_3	$\{S_1, S_2, S_3\}$		
1024	8	6	7	21	0.251	0.083
2048	7	7	6	20	0.250	0.082
4096	6	7	5	18	0.248	0.081
8192	7	3	2	12	0.249	0.083
16384	5	2	1	8	0.251	0.083
32768	3	2	1	6	0.253	0.085

selected models may not perform well in training and testing (in terms of m_1 and δ) individually.

Figure 5.12: Selected SVM base Models using different Sparsity constraints. The 3 most frequently selected models are shown for illustration. [Best viewed in color]

5.4.2 Computation Speed

Computation speed was also evaluated in the experiments. The experimental setup used is mentioned in Sec. 5.3.1. The contrast normalization process takes 0.12s per image, while the morphological-based vessel extraction operation required 0.85s. Superpixel segmentation, using SLIC, is performed in 0.64s for a 400×400 image. Training and testing of linear SVM classifiers is also quick. The most computationally heavy process is feature extraction for each superpixel and at multiple scales, and the overall computation time increases in a near linear relation with the number of superpixels. On average, the overall computation for a given test image takes 27.97s. In comparison, the intra-image learning in [26], operating using a similar computation hardware, takes 20.2s per image. In short, this proposed approach provides a gain of 6.74% in cup localization accuracy at an additional 38.46% increase in computational cost. The improved robustness and accuracy in cup localization makes a worthwhile trade-off as increases in computational cost can be often be addressed by using faster computational hardware and more optimized algorithm implementations.

5.5 Summary

Based on the state-of-the-art superpixel classification approach, this chapter has presented an optimal model integration approach to robustly localize the optic cup in retinal images. The framework provides two major contributions. Firstly, motivated by the performance variations due to random selection of training samples, a sparsity-based optimal models selection and integration solution for robust and stable optic cup localization is presented. Secondly, the multiple model selection is extended to include multiple superpixel scales for better optic cup boundary adherence and richer feature descriptions.

Experimentally, it was demonstrated that this generalized multi-model framework is able to produce optic cup localization accuracy that is 7.12% more accurate than the existing state-of-the-art intra-image learning method [138], without a restricted CDR range of 0.2 to 0.9. It is believed that the proposed framework is general enough to be adapted to other similar classification-based ocular disease CAD applications.

Chapter 6

Conclusion and Future Works

This chapter summarizes the research work of the thesis, which is to design and develop medical image analysis algorithms for optic nerve damage in glaucoma in 2-dimensional retinal fundus images. This thesis proposes three approaches to segment and localize the optic cup. To this end, several principal results and conclusions were reached.

6.1 Principal Results

This thesis first introduced a boundary-based technique using vessel kinks as key point landmarks. Biologically, kinks are vessel directional topological indicators of the optic cup boundary when vessels traverse across the cup depression. This is a highly recommended and widely used clinical cue in the clinical assessment protocol for medical doctors and clinical graders.

Inspired by this, a framework was proposed in Chapter 3 to model the detection of vessel kinking, together with cup pallor information, in order to improve existing optic cup detection. Unlike previous approaches[133, 55], in this work, vessel kinks detection is performed without the influence of an initial cup estimate, and solely focused on the temporal region of the cup, where the cup boundary is usually weaker. This is also the recommended grading practice for human graders, so as to reduce cup estimate bias. The detected vessel kinks and canonical points(from color information) are then used as boundary feature points to regulate the cup boundary using a convex optimization-based minimum volume enclosing ellipsoid technique. This technique provides a more stable result for such boundary-based approach as it approximates the cup ellipse based on the

area bounded by these key points. A key advantage of the boundary-based framework is its close modeling relationship with the clinical grading protocol, providing explicit visual evidence. However, a discussion on the limitations and challenges of such boundary-based key points approach was also articulated.

This led to the proposal of a novel region-based approach using unsupervised learning. Drawing from the experience from the previous work, which requires hand-crafting of visual features (eg. vessel kink), a hypothesis was set up to investigate the use of learning-based methods for optic cup localization. In particular, an unsupervised approach was chosen to perform this task due to its advantages of not requiring training examples, and its data-driven approach to uncover the underlying structure of the data. In this approach, region segments are obtained using superpixel segmentation, which reduces image redundancy and allows regions to be learned and classified. Using domain priors extracted from the primary structures of the optic disc, a new optic cup origin domain prior is proposed to provide initialization for the cup labels. A similarity-based propagation and refinement scheme is then used to provide additional contexture information to refine the region labels. An optic cup region is obtained after a simple region-based ellipse fitting operation. Tested on the population-based ORIGA^{-light} dataset, it was demonstrated that learning-based methods is capable of providing comparable and reliable performance against existing state-of-the-art image-processing method. This is likely due to the increased amount of information drawn from feature extraction to better represent each region, and cluster similar segments together. Each component in the unsupervised framework was individually verified experimentally and discussed. It was observed that the best performance was obtained using the correlation distance as the similarity metric during the label propagation stage. Additional experimental results were presented to demonstrate that the performance of the proposed unsupervised learning framework can be improved if more discriminative image features is employed. This highlighted the flexibility of the proposed framework and the ease in which it can incorporate and adapt future research work.

Lately, the use of superpixels for feature classification has been widely adopted in retinal imaging. In particular, for optic cup detection, the supervised superpixel-based classification approach proposed in [138] has been shown to achieve state-of-the-art performance. A general multi-scale superpixel classification was proposed, which builds

upon the single scale superpixel optic cup localization framework in [138] and [121](in Chapter 4, and incorporated two novel contributions. First, a multiple model selection and integration scheme using sparse learning was proposed to provide stability in performance variations, arising from repeated random samples selection of training data. Second, multiple superpixel resolutions were integrated for better optic cup boundary adherence and localization.

On the whole, the proposed multi-scale multi-model superpixel cup localization framework has demonstrated its ability to produce optic cup localization accuracy that is 7.12% higher than the current state-of-the-art intra-image learning method [138]. Furthermore, it is not restricted within a CDR range of 0.2 to 0.9, and offers better boundary adherence, with robust performance.

It has been observed the supervised learning-based approaches for optic cup localization can produce highly competitive results, compared to the unsupervised approach. Unfortunately, in order to deliver state-of-the-art performance, ground-truth labels are required for the supervised learning algorithm to minimize the errors and generate the resulting learning model(s). In the absence of these labels, such an approach would not be feasible. This is a common challenge faced in medical imaging applications, where obtaining sufficient numbers of ground-truth labels for real patient data can be a difficult and costly affair. Under these circumstances, the proposed unsupervised learning approach, which is not based on the same assumptions, offers a more practical alternative that also gives reasonably accurate results. The additional benefit is that with the flexible framework, the accuracy of the optic cup localization can be further improved if more discriminative features are discovered and later integrated into the proposed unsupervised learning framework. Even though the proposed boundary-based approach that uses vessel kinks does not perform as well as the learning-based counterparts, its strength is that it provides a visual appreciation and interpretation of how and where these key points are extracted. This offers another form of valuable image-based visual evidence feedback to medical professionals and such clinically-supported evidence can help make automated CAD systems more acceptable to clinicians.

Although in this thesis, it is demonstrated that the proposed approaches are effective for the task of optic cup localization, it is believed that these proposed frameworks are

also general enough to be adapted to other similar classification-based ocular disease CAD applications such as, the localization of drusens in Age-Related Macular Degeneration (AMD) and the localization of peripapillary atrophy in degenerative/severe myopia. While there is still a long way from replicating clinical graders who can cognitively recognize optic nerve damage, and demarcate the optic cup boundary, under various challenging settings, the author hope that this thesis has taken a positive step in providing a solution. In the section below, future works are proposed to further investigate and improve competency and versatility in this area.

6.2 Future Research Directions

During the course of this research journey, it was found that it would be meaningful to further investigate and utilize the intrinsic relationship between the different visual features and clinical clues(domain priors) for glaucoma diagnosis. Techniques for feature selection, includes minimum-redundancy-maximum-relevance [93], unsupervised feature learning and deep learning [14, 64], sparse group lasso [81, 13, 73], have shown to be effective for applications in face recognition, digit recognition, life science and breast cancer data. However, there has not been any works to investigate, the relationship of extracted image features with domain priors, for glaucoma diagnosis on retinal images. The study of this relationship could bring more insights into 1) the overall performance and discriminative features of automatic glaucoma diagnosis, 2) better interpretation of clinical glaucoma diagnosis observations, 3) possible useful understanding of physiological color, size and shape morphology normal and/or glaucomatous optic nerves. Another possible area of work is to investigate the union of both boundary-based and region-based techniques proposed in this thesis.

Lately, it has recently been shown in some computer vision applications, e.g. gait recognition [136] and prostate segmentation [19], that higher accuracy can be achieved using reconstruction from samples. Motivated by this, a optic cup localization framework based on local patch reconstruction has been investigated, and preliminary work in this area is detailed in [141]. In particular, two types of local patches, i.e. grids and superpixels were used to show the variety, generalization ability and robustness of the proposed

framework. However, the performance, in terms of non-overlap ratio, m_1 , still lacks behind the multi-scale superpixel classification approach presented in Chap. 5. A possible future work will be to embed the approach onto the proposed multi-scale superpixel classification framework.

Another potential research extension will be to explore means to fuse both the proposed (supervised and unsupervised) approaches together into a unified framework. One such framework is the semi-supervised learning framework. Semi-supervised learning (SSL) is a technique that attempts to build better classifiers by using large amount of unlabeled data, together with small amount of labeled data [106, 148]. Often, in medical imaging, labels or "ground-truth" requires manual annotation efforts by medical experts, which can be difficult, expensive and time-consuming to obtain. Semi-supervised learning addresses this issue by utilizing small(limited) amount of marked(known/annotated) labels, together with large amount of unlabeled data, and learn a predictor that predicts future test data better than the predictor learned solely from the labeled training data alone. Researchers have reported positive performances using SSL-based methods in extracting diagnoses and investigation results from Electronic Health Records[129], prediction of cancer patient's survival rate [12], breast cancer diagnosis [68]. However, in general, it is widely accepted that unlabeled data is only likely to help if there exists a link (cluster assumption or manifold assumption) between the marginal data distribution and the target function to be learnt [110]. Some representative research works in SSL includes using generative likelihood model [86, 48, 40], low density separation [22, 142], co-training [15, 50] and multi-view learning [146], and graph-based approaches [149, 71].

A possible direction along the same line will be to investigate a robust graph-based semi-supervised learning (GSSL) for glaucoma diagnosis. GSSL has been shown to be effective in 1)scalability to large data, 2) modeling the manifold structures in high dimensional spaces, 3) good performance and ease in implementation. The extension from the proposed unsupervised methodology fits well with this framework. However, some of the major challenges in GSSL is 1) to reduce the influence of noisy labels, which can result in poor graph transduction, 2) imbalanced labelling in the ratio of positive and negative labels, which may cause a bias during learning [71]. Fergus and colleagues [100] proposed using probabilistic latent semantic analysis (pLSA) as a prior to filter these noisy labels.

Lastly, as mentioned earlier, it is often difficult to find sufficient or true exemplary labels, especially in a screening setup. Unlike supervised approaches where correct labels are provided, reinforcement learning emphasizes on learning from direct interaction with the environment using rewards as feedback, without exemplary supervision or complete models of the environment [119]. Successful applications of RL has been reported in backgammon [123], cellular telephone network system [111], helicopter control [82], marketing strategies[7] and robotics [47]. Other works in recommending diagnostic tests for solitary pulmonary nodule (SPN) [31] and segmentation of the prostate in ultrasound images [103] have also been reported. A possible future direction could be to develop an online reinforcement learning framework for glaucoma diagnosis. This provides several advantages over existing automated retinal image-based glaucoma diagnosis. First, an online RL framework offers real-time (life-long) learning, which allows the system to adapt their knowledge and behaviour over time, and reduces the sample complexity when learning new task [77]. Second, this framework has the ability to explore (learn new knowledge) and exploit (apply learnt knowledge) its states and actions. Third, it has the ability to perform sequential decision making using partially observable states to improve its observability of the environment. Intuitively, the online RL system is similar to reasoning in human decision making. Due to learning through interaction with its environment, the algorithm is forced to learn from its own mistakes from the reward/penalty feedback and can learn a better policy.

References

- [1] Glossary of Ophthalmology Terms, Heidelberg Engineering. <http://www.heidelbergengineering.com/international/education/glossary-of-terms/>. Accessed: Jul 2012.
- [2] HANDBOOK of Ocular Disease Management, Review of Optometry. <http://cms.revoptom.com/handbook/>. Accessed: Jul 2012.
- [3] Heidelberg engineering academy quick guides & tutorials. <http://www.heidelbergengineering.com/us/academy-education/quick-guides/>. Accessed: May 2014.
- [4] Laser scanning confocal microscopy. <http://www.micro.magnet.fsu.edu/primer/techniques/confocal/>. Accessed: May 2012.
- [5] The ophthalmoscope. <http://www.mrcophth.com/Historyofophthalmology/ophthalmoscope.html>. Accessed: Jun 2014.
- [6] Optic Nerve and Retinal Nerve Fiber Imaging. http://eyewiki.aao.org/Optic_Nerve_and_Retinal_Nerve_Fiber_Imaging. Accessed: Jul 2012.
- [7] Naoki Abe, Naval Verma, Chid Apte, and Robert Schroko. Cross channel optimized marketing by reinforcement learning. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '04, pages 767–772, New York, NY, USA, 2004. ACM.
- [8] M. D. Abramoff, W. L. Alward, E. C. Greenlee, L. Shuba, C. Y. Kim, J. H. Fingert, and Y. H. Kwon. Automated segmentation of the optic disc from stereo color photographs using physiologically plausible features. *Invest. Ophthalmol. Vis. Sci.*, 48:1665–1673, Apr 2007.

REFERENCES

- [9] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk. Slic superpixels compared to state-of-the-art superpixel methods. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, PP(99), 2012.
- [10] S. Agarwal, A. Awan, and D. Roth. Learning to detect objects in images via a sparse, part-based representation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 26(11):1475–1490, Nov 2004.
- [11] A. Aquino, M.E. Gegundez-Arias, and D. Marin. Detecting the optic disc boundary in digital fundus images using morphological, edge detection, and feature extraction techniques. *Medical Imaging, IEEE Transactions on*, 29(11):1860–1869, nov. 2010.
- [12] E. Bair and R. Tibshirani. Semi-supervised methods to predict patient survival from gene expression data. *PLoS Biol.*, 2(4):E108, Apr 2004.
- [13] Semy Bengio, Fernando Pereira, Yoram Singer, and Dennis Strelow. Group Sparse Coding. In *Advances in Neural Information Processing Systems*, 2009.
- [14] Yoshua Bengio, Pascal Lamblin, Dan Popovici, Hugo Larochelle, Universit De Montral, and Montral Qubec. Greedy layer-wise training of deep networks. In *NIPS*. MIT Press, 2007.
- [15] Avrim Blum and Tom Mitchell. Combining labeled and unlabeled data with co-training. In *Proceedings of the eleventh annual conference on Computational learning theory, COLT’ 98*, pages 92–100, New York, NY, USA, 1998. ACM.
- [16] Stephen. Boyd and Lieven. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.
- [17] J. M. Burr, G. Mowatt, R. Hernandez, M. A. Siddiqui, J. Cook, T. Lourenco, C. Ramsay, L. Vale, C. Fraser, A. Azuara-Blanco, J. Deeks, J. Cairns, R. Wormald, S. McPherson, K. Rabindranath, and A. Grant. The clinical effectiveness and cost-effectiveness of screening for open angle glaucoma: a systematic review and economic evaluation. *Health Technol Assess*, 11(41):1–190, Oct 2007.

REFERENCES

- [18] E.J. Candes and M.B. Wakin. An introduction to compressive sampling. *Signal Processing Magazine, IEEE*, 25(2):21–30, March 2008.
- [19] Yihui Cao, Yuan Yuan, Xuelong Li, Baris Turkbey, PeterL. Choyke, and Pingkun Yan. Segmenting images by combining selected atlases on manifold. In Gabor Fichtinger, Anne Martel, and Terry Peters, editors, *Medical Image Computing and Computer-Assisted Intervention MICCAI 2011*, volume 6893 of *Lecture Notes in Computer Science*, pages 272–279. Springer Berlin Heidelberg, 2011.
- [20] J. Caprioli. Clinical evaluation of the optic nerve in glaucoma. *Trans Am Ophthalmol Soc*, 92:589–641, 1994.
- [21] V. Chalana and Y. Kim. A methodology for evaluation of boundary detection algorithms on medical images. *IEEE Trans Med Imaging*, 16:642–652, Oct 1997.
- [22] Olivier Chapelle and Alexander Zien. Semi-supervised classification by low density separation.
- [23] J. Cheng, J. Liu, D. W. Wong, F. Yin, C. Cheung, M. Baskaran, T. Aung, and T. Y. Wong. Automatic optic disc segmentation with peripapillary atrophy elimination. *Conf Proc IEEE Eng Med Biol Soc*, 2011:6224–6227, 2011.
- [24] J. Cheng, J. Liu, Y. Xu, F. Yin, D.W.K. Wong, N.M. Tan, D. Tao, C. Cheng, T. Aung, and T.Y. Wong. Superpixel classification based optic disc and optic cup segmentation for glaucoma screening. *IEEE Trans Med Imaging*, 32(6):1019–1032, 2013.
- [25] S. Da Pozzo, R. Marchesan, and G. Ravalico. Scanning laser polarimetry - a review. *Clin. Experiment. Ophthalmol.*, 37(1):68–80, Jan 2009.
- [26] T. Damms and F. Dannheim. Sensitivity and specificity of optic disc parameters in chronic glaucoma. *Invest. Ophthalmol. Vis. Sci.*, 34:2246–2250, Jun 1993.
- [27] Wong Wing Kee Damon, Jimmy Liu, Tan Ngan Meng, Yin Fengshou, and Wong Tien Yin. Automatic detection of the optic cup using vessel kinking in digital retinal fundus images. In *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on*, pages 1647 –1650, may 2012.

REFERENCES

- [28] Kunio Doi. Computer-aided diagnosis in medical imaging: Historical review, current status and future potential. *Computerized Medical Imaging and Graphics*, 31:198–211, 2007.
- [29] Mark Everingham, Luc Gool, Christopher K. Williams, John Winn, and Andrew Zisserman. The pascal visual object classes (voc) challenge. *Int. J. Comput. Vision*, 88(2):303–338, June 2010.
- [30] Bay EyesCataract&LaserCenter. Glaucoma . <http://www.retinalphysician.com/articleviewer.aspx?articleID=100656>. Accessed: Jul 2012.
- [31] S.J. Fakhri and T.K. Das. Lead: a methodology for learning efficient approaches to medical diagnosis. *Information Technology in Biomedicine, IEEE Transactions on*, 10(2):220 –228, april 2006.
- [32] L. Fei-Fei and P. Perona. A bayesian hierarchical model for learning natural scene categories. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 2, pages 524 – 531 vol. 2, june 2005.
- [33] R. Fergus, P. Perona, and A. Zisserman. Object class recognition by unsupervised scale-invariant learning. In *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, volume 2, pages II–264–II–271 vol.2, June 2003.
- [34] A.W. Fitzgibbon, M. Pilu, and R.B. Fisher. Direct least squares fitting of ellipses. In *Pattern Recognition, 1996., Proceedings of the 13th International Conference on*, volume 1, pages 253 –257 vol.1, aug 1996.
- [35] A. W. Foong, S. M. Saw, J. L. Loo, S. Shen, S. C. Loon, M. Rosman, T. Aung, D. T. Tan, E. S. Tai, and T. Y. Wong. Rationale and methodology for a population-based study of eye diseases in Malay people: The Singapore Malay eye study (SiMES). *Ophthalmic Epidemiol*, 14:25–35, 2007.
- [36] M. Foracchia, E. Grisan, and A. Ruggeri. Detection of optic disc in retinal images by means of a geometrical model of vessel structure. *Medical Imaging, IEEE Transactions on*, 23(10):1189 –1195, oct. 2004.

REFERENCES

- [37] P. J. Foster, F. T. Oen, D. Machin, T. P. Ng, J. G. Devereux, G. J. Johnson, P. T. Khaw, and S. K. Seah. The prevalence of glaucoma in Chinese residents of Singapore: a cross-sectional population survey of the Tanjong Pagar district. *Arch. Ophthalmol.*, 118:1105–1111, Aug 2000.
- [38] Glaucoma Research Foundation. Glaucoma facts and stats. http://www.glaucoma.org/learn/glaucoma_facts.html. Accessed: Feb 2010.
- [39] E. M. Frohman, J. G. Fujimoto, T. C. Frohman, P. A. Calabresi, G. Cutter, and L. J. Balcer. Optical coherence tomography: a window into the mechanisms of multiple sclerosis. *Nat Clin Pract Neurol*, 4(12):664–675, Dec 2008.
- [40] Akinori Fujino, Naonori Ueda, and Kazumi Saito. A hybrid generative/discriminative approach to semi-supervised classifier design. In *Proceedings of the 20th national conference on Artificial intelligence - Volume 2*, AAAI’05, pages 764–769. AAAI Press, 2005.
- [41] Fiona J. Gilbert, Susan M. Astley, Maureen G.C. Gillan, Olorunsola F. Agbaje, Matthew G. Wallis, Jonathan James, Caroline R.M. Boggis, and Stephen W. Duffy. Single reading with computer-aided detection for screening mammography. *N. Engl. J. Med.*, 359(16):1675–1684, Oct 2008.
- [42] John Goldsmith. Unsupervised learning of the morphology of a natural language. *Computational linguistics*, 27(2):153–198, 2001.
- [43] Denise Grady. The vision thing: Mainly in the brain. *Discover*, 14:57–66, June 1993.
- [44] R. Nelson H. Kolb, E. Fernandez and B. W. Jones. Webvision: The organization of the retina and visual system 2005 [online]. <http://webvision.med.utah.edu/>. Accessed: Mar 2011.
- [45] Yuji Hatanaka, Atsushi Noudo, Chisako Muramatsu, Akira Sawada, Takeshi Hara, Tetsuya Yamamoto, and Hiroshi Fujita. Automatic measurement of vertical cup-to-disc ratio on retinal fundus images. In David Zhang and Milan Sonka, editors,

- Medical Biometrics*, volume 6165 of *Lecture Notes in Computer Science*, pages 64–72. Springer Berlin / Heidelberg, 2010.
- [46] A. Heijl, M. C. Leske, B. Bengtsson, L. Hyman, B. Bengtsson, and M. Hussein. Reduction of intraocular pressure and glaucoma progression: results from the Early Manifest Glaucoma Trial. *Arch. Ophthalmol.*, 120(10):1268–1279, 2002.
- [47] Todd Hester and Peter Stone. TEXPLORE: Real-time sample-efficient reinforcement learning for robots. *Machine Learning*, To Appear, 2012.
- [48] Alex D. Holub, Pietro Perona, and MaxWelling. Exploiting unlabelled data for hybrid object classification. In *Advances in Neural Information Processing Systems, Workshop in Inter-Class Transfer*, 2005.
- [49] A. Hoover and M. Goldbaum. Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels. *Medical Imaging, IEEE Transactions on*, 22(8):951–958, aug. 2003.
- [50] Zhi hua Zhou, De chuan Zhan, and Qiang Yang. Semi-supervised learning with very few labeled training examples. In *Twenty-Second AAAI Conference on Artificial Intelligence (AAAI-07)*, pages 675–680, 2007.
- [51] N. Inoue, K. Yanashima, K. Magatani, and T. Kurihara. Development of a simple diagnostic method for the glaucoma using ocular Fundus pictures. *Conf Proc IEEE Eng Med Biol Soc*, 4:3355–3358, 2005.
- [52] J. B. Jonas, W. M. Budde, and S. Panda-Jonas. Ophthalmoscopic evaluation of the optic nerve head. *Surv Ophthalmol*, 43(4):293–320, 1999.
- [53] G. Joshi, J. Sivaswamy, and S.R. Krishnadas. Depth discontinuity-based cup segmentation from multiview color retinal images. *Biomedical Engineering, IEEE Transactions on*, 59(6):1523–1531, june 2012.
- [54] G.D. Joshi, J. Sivaswamy, K. Karan, and S.R. Krishnadas. Optic disk and cup boundary detection using regional information. In *Biomedical Imaging: From Nano to Macro, 2010 IEEE International Symposium on*, pages 948–951, april 2010.

REFERENCES

- [55] G.D. Joshi, J. Sivaswamy, and S.R. Krishnadas. Optic disk and cup segmentation from monocular color retinal images for glaucoma assessment. *Medical Imaging, IEEE Transactions on*, 30(6):1192–1205, june 2011.
- [56] Miriam Karmel. Optic Disc Evaluation: Back to Basics. <http://www.aao.org/publications/eyenet/200504/glaucoma.cfm>. Accessed: Jul 2012.
- [57] M. A. Kass, D. K. Heuer, E. J. Higginbotham, C. A. Johnson, J. L. Keltner, J. P. Miller, R. K. Parrish, M. R. Wilson, and M. O. Gordon. The Ocular Hypertension Treatment Study: a randomized trial determines that topical ocular hypotensive medication delays or prevents the onset of primary open-angle glaucoma. *Arch. Ophthalmol.*, 120(6):701–713, Jun 2002.
- [58] Jr. Kenneth W. Tobin, Edward Chaum, V. P. Govindasamy, Thomas P. Karnowski, and Omer Sezer. Characterization of the optic disc in retinal imagery using a probabilistic approach. volume 6144, page 61443F. SPIE, 2006.
- [59] Leonid G. Khachiyan. Rounding of polytopes in the real number model of computation. *Math. Oper. Res.*, 21(2):307–320, May 1996.
- [60] R. Komulainen, A. Tuulonen, and P. J. Airaksinen. The follow-up of patients screened for glaucoma with non-mydratic fundus photography. *Int Ophthalmol*, 16(6):465–469, Nov 1992.
- [61] C. G. Krader. Fundus photography still a valuable tool in glaucoma care. <http://ophthalmologytimes.modernmedicine.com/ophthalmologytimes/news/tags/glaucoma-360/fundus-photography-still-valuable-tool-glaucoma-care?page=full#>. Accessed: May 2014.
- [62] P. Kumar and E. A. Yildirim. Minimum-volume enclosing ellipsoids and core sets. *Journal of Optimization Theory and Applications*, 126(1):1–21, 2005.
- [63] M. Lalonde, M. Beaulieu, and L. Gagnon. Fast and robust optic disc detection using pyramidal decomposition and hausdorff-based template matching. *Medical Imaging, IEEE Transactions on*, 20(11):1193–1200, nov. 2001.

- [64] Honglak Lee, Yan Largman, Peter Pham, and Andrew Y. Ng. Unsupervised feature learning for audio classification using convolutional deep belief networks. In *Advances in Neural Information Processing Systems 22*, pages 1096–1104. 2009.
- [65] P. P. Lee, J. G. Walt, J. J. Doyle, S. V. Kotak, S. J. Evans, D. L. Budenz, P. P. Chen, A. L. Coleman, R. M. Feldman, H. D. Jampel, L. J. Katz, R. P. Mills, J. S. Myers, R. J. Noecker, J. R. Piltz-Seymour, R. R. Ritch, P. N. Schacknow, J. B. Serle, and G. L. Trick. A multicenter, retrospective pilot study of resource use and costs associated with severity of disease in glaucoma. *Arch. Ophthalmol.*, 124(1):12–19, Jan 2006.
- [66] H. Li and O. Chutatape. A model-based approach for automated feature extraction in fundus images. In *Computer Vision, 2003. Proceedings. Ninth IEEE International Conference on*, pages 394 –399 vol.1, 2003.
- [67] Huiqi Li and O. Chutatape. Automated feature extraction in color retinal images by a model based approach. *Biomedical Engineering, IEEE Transactions on*, 51(2):246–254, feb. 2004.
- [68] Ming Li and Zhi-Hua Zhou. Improve computer-aided diagnosis with machine learning techniques using undiagnosed samples. *Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on*, 37(6):1088 –1098, nov. 2007.
- [69] J. Liu, S. Ji, and J. Ye. *SLEP: Sparse Learning with Efficient Projections*. Arizona State University, 2009.
- [70] J. Liu, D. W. K. Wong, J. H. Lim, H. Li, N. M. Tan, Z. Zhang, T. Y. Wong, and R. Lavanya. Argali: An automatic cup-to-disc ratio measurement system for glaucoma analysis using level-set image processing. In Ratko Magjarevic, Chwee Teck Lim, and James C. H. Goh, editors, *13th International Conference on Biomedical Engineering*, volume 23 of *IFMBE Proceedings*, pages 559–562. Springer Berlin Heidelberg, 2009.
- [71] W. Liu, J. Wang, and S.-F. Chang. Robust and scalable graph-based semisupervised learning. *Proceedings of the IEEE*, 100(9):2624 –2638, sept. 2012.

REFERENCES

- [72] J. Lowell, A. Hunter, D. Steel, A. Basu, R. Ryder, E. Fletcher, and L. Kennedy. Optic nerve head segmentation. *Medical Imaging, IEEE Transactions on*, 23(2):256–264, Feb. 2004.
- [73] Julien Mairal, Francis Bach, Jean Ponce, and Guillermo Sapiro. Online learning for matrix factorization and sparse coding. *J. Mach. Learn. Res.*, 11:19–60, March 2010.
- [74] F. Mendels, C. Heneghan, and J.P. Thiran. Identification of the optic disc boundary in retinal images using active contours. In *Proceedings of the IMVIP Conference.*, page 103115, 1999.
- [75] Quigley MG and Dub P. A new fundus camera technique to help calculate eye-camera magnification: A rapid means to measure disc size. *Archives of Ophthalmology*, 121(5):707–709, 2003.
- [76] M. Minsky. Microscopy apparatus, December 19 1961. US Patent 3,013,467.
- [77] Thomas M. Mitchell. *Machine Learning*. McGraw-Hill, Inc., New York, NY, USA, 1 edition, 1997.
- [78] G. Mori, Xiaofeng Ren, A.A. Efros, and J. Malik. Recovering human body configurations: combining segmentation and recognition. In *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, volume 2, pages II–326–II–333 Vol.2, June 2004.
- [79] N. Moshtagh. Minimum volume enclosing ellipsoid. <https://sites.google.com/site/nimamoshtagh/software>. Accessed: 10/07/2012.
- [80] USA National Institute of Health. Facts About Glaucoma . http://www.nei.nih.gov/health/glaucoma/glaucoma_facts.asp. Accessed: Jul 2012.
- [81] Sahand Negahban, Pradeep Ravikumar, Martin Wainwright, and Bin Yu. A unified framework for high-dimensional analysis of ℓ_1 -estimators with decomposable regularizers. In Y. Bengio, D. Schuurmans, J. Lafferty, C. K. I. Williams, and A. Culotta, editors, *Advances in Neural Information Processing Systems 22*, pages 1348–1356. 2009.

REFERENCES

- [82] Andrew Y. Ng, H. Jin Kim, Michael I. Jordan, and Shankar Sastry. Inverted autonomous helicopter flight via reinforcement learning. In *In International Symposium on Experimental Robotics*. MIT Press, 2004.
- [83] Feiping Nie, Heng Huang, Xiao Cai, and Chris H. Ding. Efficient and robust feature selection via joint l_2 , l_1 -norms minimization. In J.D. Lafferty, C.K.I. Williams, J. Shawe-Taylor, R.S. Zemel, and A. Culotta, editors, *Advances in Neural Information Processing Systems 23*, pages 1813–1821. Curran Associates, Inc., 2010.
- [84] JuanCarlos Niebles, Hongcheng Wang, and Li Fei-Fei. Unsupervised learning of human action categories using spatial-temporal words. *International Journal of Computer Vision*, 79(3):299–318, 2008.
- [85] Meindert Niemeijer, Michael D. Abrmoff, and Bram van Ginneken. Fast detection of the optic disc and fovea in color fundus photographs. *Medical Image Analysis*, 13(6):859 – 870, 2009. Includes Special Section on Computational Biomechanics for Medicine.
- [86] Kamal Nigam, Andrew Kachites McCallum, Sebastian Thrun, and Tom Mitchell. Text classification from labeled and unlabeled documents using em. In *Machine Learning*, pages 103–134, 1999.
- [87] Optic-disc.org. Glaucoma Evaluation Tutorial . http://www.optic-disc.org/tutorials/glaucoma_evaluation_basics/page13.html. Accessed: Jul 2014.
- [88] World Health Organization. Visual Impairment and Blindness Fact Sheet N°282 . <http://www.who.int/mediacentre/factsheets/fs282/en/>. Accessed: 10/07/2012.
- [89] A. Osareh, M. Mirmehdi, B. Thomas, and R. Markham. Comparison of colour spaces for optic disc localisation in retinal images. In *Pattern Recognition, 2002. Proceedings. 16th International Conference on*, volume 1, pages 743 – 746 vol.1, 2002.

- [90] M. E. Tyler P. J. Saine. Fundus photography overview, excerpted from ophthalmic photography: Retinal photography, angiography, and electronic imaging, 2nd edition. <http://www.opsweb.org/?page=fundusphotography>. Accessed: Jun 2014.
- [91] D. Pascolini and S. P. Mariotti. Global estimates of visual impairment: 2010. *Br J Ophthalmol*, 96(5):614–618, May 2012.
- [92] D. Pascolini, S. P. Mariotti, G. P. Pokharel, R. Pararajasegaram, D. Etya’ale, A. D. Negrel, and S. Resnikoff. 2002 global update of available data on visual impairment: a compilation of population based prevalence studies. *Ophthalmic Epidemiol*, 11:7–115, 2004.
- [93] Hanchuan Peng, Fuhui Long, and Chris Ding. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27:1226–1238, 2005.
- [94] Nicholas Petrick, Maruf Haider, Ronald M. Summers, Srinath C. Yeshwant, Linda Brown, Edward M. Iuliano, Adeline Louie, J. Richard Choi, and Perry J. Pickhardt. Ct colonography with computer-aided detection as a second reader: Observer performance study. *Radiology*, 246(1):148–156, 2008. PMID: 18096536.
- [95] D. L. Pham, C. Xu, and J. L. Prince. A Survey of Current Methods in Medical Image Segmentation. In *Annual Review of Biomedical Engineering*, volume 2, pages 315–338. 2000.
- [96] Retinal Physician. Nonmydriatic Cameras in Retinal Practice . <http://www.1800clearly.com/glaucoma1.htm>. Accessed: Jul 2012.
- [97] M. Pilu, A.W. Fitzgibbon, and R.B. Fisher. Ellipse-specific direct least-square fitting. In *Image Processing, 1996. Proceedings., International Conference on*, volume 3, pages 599 –602 vol.3, sep 1996.
- [98] S.L. Polyak. *The retina*. University of Chicago Press, 1941.

REFERENCES

- [99] H. A. Quigley and A. T. Broman. The number of people with glaucoma worldwide in 2010 and 2020. *Br J Ophthalmol*, 90:262–267, Mar 2006.
- [100] P. Perona R. Fergus, L. Fei-Fei and A. Zisserman. Learning object categories from Internet image searches. *Proc. Of IEEE*, 8(98), 2010.
- [101] S. Freedman S. E. Moroi R. R. Allingham, K. Damji and G. Shafranov. *Shields’ Textbook of Glaucoma*. Lippincott Williams and Wilkins, Philadelphia PA, 2004 (7th edition).
- [102] S. Ravishankar, A. Jain, and A. Mittal. Automated feature extraction for early detection of diabetic retinopathy in fundus images. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 210 –217, june 2009.
- [103] F. Sahba, H.R. Tizhoosh, and M.M.A. Salama. A reinforcement learning framework for medical image segmentation. In *Neural Networks, 2006. IJCNN ’06. International Joint Conference on*, pages 511 –517, 0-0 2006.
- [104] Patrick J. Saine and Marshall E. Tyler. *Ophthalmic Photography: Retinal Photography, Angiography, and Electronic Imaging*. Butterworth-Heinemann Medical, 2nd edition, 2002.
- [105] J. S. Schuman. Spectral domain optical coherence tomography for glaucoma (an AOS thesis). *Trans Am Ophthalmol Soc*, 106:426–458, 2008.
- [106] Matthias Seeger. Learning with labeled and unlabeled data. In *Tech. Rep., University of Edinburgh*, 2001.
- [107] S. Sekhar, W. Al-Nuaimy, and A.K. Nandi. Automated localisation of retinal optic disk using hough transform. In *Biomedical Imaging: From Nano to Macro, 2008. ISBI 2008. 5th IEEE International Symposium on*, pages 1577 –1580, may 2008.
- [108] S. Y. Shen, T. Y. Wong, P. J. Foster, J. L. Loo, M. Rosman, S. C. Loon, W. L. Wong, S. M. Saw, and T. Aung. The prevalence and types of glaucoma in malay people: the Singapore Malay eye study. *Invest. Ophthalmol. Vis. Sci.*, 49:3846–3851, Sep 2008.

REFERENCES

- [109] D. H. Sim and L. G. Goh. Screening for glaucoma in the Chinese elderly population in Singapore. *Singapore Med J*, 40(10):644–647, Oct 1999.
- [110] Aarti Singh, Robert D. Nowak, and Xiaojin Zhu. Unlabeled data: Now it helps, now it doesnt. In *Advances in Neural Information Processing Systems*, 2008.
- [111] Satinder Singh and Dimitri Bertsekas. Reinforcement learning for dynamic channel allocation in cellular telephone systems. In *In Advances in Neural Information Processing Systems: Proceedings of the 1996 Conference*, pages 974–980. MIT Press, 1997.
- [112] C. Sinthanayothin, J. F. Boyce, H. L. Cook, and T. H. Williamson. Automated localisation of the optic disc, fovea, and retinal blood vessels from digital colour fundus images. *Br J Ophthalmol*, 83(8):902–910, Aug 1999.
- [113] J.V.B. Soares, J.J.G. Leandro, R.M. Cesar, H.F. Jelinek, and M.J. Cree. Retinal vessel segmentation using the 2-d gabor wavelet and supervised classification. *Medical Imaging, IEEE Transactions on*, 25(9):1214 –1222, sept. 2006.
- [114] A. Sommer, J. M. Tielsch, J. Katz, H. A. Quigley, J. D. Gottsch, J. Javitt, and K. Singh. Relationship between intraocular pressure and primary open angle glaucoma among white and black Americans. The Baltimore Eye Survey. *Arch. Ophthalmol.*, 109(8):1090–1095, Aug 1991.
- [115] D. Steward. What is computer-aided diagnosis? *Semin. Vet. Med. Surg. Small Anim.*, 11(2):74–84, May 1996.
- [116] Markus Stricker and Markus Orengo. Similarity of color images. In *SPIE Conference on Storage and Retrieval for Image and Video Databases III*, volume 2420, pages 381–392, 1995.
- [117] D. H. Su, T. Y. Wong, W. L. Wong, S. M. Saw, D. T. Tan, S. Y. Shen, S. C. Loon, P. J. Foster, and T. Aung. Diabetes, hyperglycemia, and central corneal thickness: the Singapore Malay Eye Study. *Ophthalmology*, 115:964–968, Jun 2008.

REFERENCES

- [118] Peng Sun and Robert M. Freund. Computation of minimum-volume covering ellipsoids. *Oper. Res.*, 52(5):690–706, October 2004.
- [119] Richard S. Sutton and Andrew G. Barto. Reinforcement learning i : Introduction, 1998.
- [120] N. M. Tan, J. Liu, D. K. Wong, F. Yin, J. H. Lim, and T. Y. Wong. Mixture model-based approach for optic cup segmentation. *Conf Proc IEEE Eng Med Biol Soc*, 2010:4817–4820, 2010.
- [121] N-M. Tan, Y. Xu, J. Liu, W.B. Goh, C. Cheung, T. Aung, and T.Y. Wong. Domain prior based superpixel propagation for optic cup localization. In *IEEE International Symposium on Biomedical Imaging (ISBI)*, pages 880–883, 2013.
- [122] Yandong Tang, Xiaomao Li, A. von Freyberg, and G. Goch. Automatic segmentation of the papilla in a fundus image based on the c-v model and a shape restraint. In *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, 0 2006.
- [123] Gerald Tesauro. Temporal difference learning and td-gammon. *Commun. ACM*, 38(3):58–68, March 1995.
- [124] Joseph Tighe and Svetlana Lazebnik. Superparsing: Scalable nonparametric image parsing with superpixels. In *Proceedings of the 11th European Conference on Computer Vision: Part V, ECCV’10*, pages 352–365, Berlin, Heidelberg, 2010. Springer-Verlag.
- [125] Michael J. Todd and E. Alper Yildirim. On khachiyan’s algorithm for the computation of minimum-volume enclosing ellipsoids. *Discrete Appl. Math.*, 155(13):1731–1744, August 2007.
- [126] M.A Turk and AP. Pentland. Face recognition using eigenfaces. In *Computer Vision and Pattern Recognition, 1991. Proceedings CVPR ’91., IEEE Computer Society Conference on*, pages 586–591, Jun 1991.

REFERENCES

- [127] B. van Ginneken, B.M. Ter Haar Romeny, and M.A. Viergever. Computer-aided diagnosis in chest radiography: a survey. *Medical Imaging, IEEE Transactions on*, 20(12):1228–1241, Dec 2001.
- [128] Thomas Walter and Jean-Claude Klein. Segmentation of color fundus images of the human retina: Detection of the optic disc and the vascular tree using morphological techniques. In *Proceedings of the Second International Symposium on Medical Data Analysis*, ISMDA '01, pages 282–287, London, UK, UK, 2001. Springer-Verlag.
- [129] Z. Wang, A. D. Shah, A. R. Tate, S. Denaxas, J. Shawe-Taylor, and H. Hemingway. Extracting diagnoses and investigation results from unstructured text in electronic health records by semi-supervised machine learning. *PLoS ONE*, 7(1):e30412, 2012.
- [130] Matthew T. Witmer, Curtis E. Margo, and Mitchell Drucker. Tilted optic disks. *Survey of Ophthalmology*, 55(5):403 – 428, 2010.
- [131] D. W. Wong, J. Liu, J. H. Lim, N. M. Tan, Z. Zhang, S. Lu, H. Li, M. H. Teo, K. L. Chan, and T. Y. Wong. Intelligent fusion of cup-to-disc ratio determination methods for glaucoma detection in ARGALI. *Conf Proc IEEE Eng Med Biol Soc*, 2009:5777–5780, 2009.
- [132] D. W. K. Wong, J. Liu, J. H. Lim, H. Li, and T. Y. Wong. Automated detection of kinks from blood vessels for optic cup segmentation in retinal images. volume 7260, page 72601J. SPIE, 2009.
- [133] D. W. K. Wong, J. Liu, J.H. Lim, X. Jia, F. Yin, H. Li, and T. Y. Wong. Level-set based automatic cup-to-disc ratio determination using retinal fundus images in argali. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pages 2266 –2269, 2008.
- [134] Damon W. K. Wong, Jiang Liu, Ngan Meng Tan, Zhuo Zhang, Shijian Lu, Joo Hwee Lim, Huiqi Li, and Tien Yin Wong. Enhancement of optic cup detection through an improved vessel kink detection framework. volume 7624, page 762439. SPIE, 2010.

REFERENCES

- [135] D.W.K. Wong, Jiang Liu, Ngan-Meng Tan, Fengshou Yin, Xiangang Cheng, G.C.M. Cheung, M. Bhargava, and Tien Yin Wong. Automatic localization of the macula in a supervised graph-based approach with contextual superpixel features. In *Pattern Recognition (ICPR), 2012 21st International Conference on*, pages 2063–2066, 2012.
- [136] Dong Xu, Yi Huang, Zinan Zeng, and Xinxing Xu. Human gait recognition using patch distribution feature and locality-constrained group sparse representation. *Image Processing, IEEE Transactions on*, 21(1):316–326, Jan 2012.
- [137] Juan Xu, Opas Chutatape, Eric Sung, Ce Zheng, and Paul Chew Tec Kuan. Optic disk feature extraction via modified deformable model technique for glaucoma analysis. *Pattern Recogn.*, 40:2063–2076, July 2007.
- [138] Y. Xu, J. Liu, S. Lin, D. Xu, C. Y. Cheung, T. Aung, and T. Y. Wong. Efficient optic cup detection from intra-image learning with retinal structure priors. *Med Image Comput Comput Assist Interv*, 15(Pt 1):58–65, 2012.
- [139] Y. Xu, D. Xu, S. Lin, J. Liu, J. Cheng, C. Y. Cheung, T. Aung, and T. Y. Wong. Sliding window and regression based cup detection in digital fundus images for glaucoma diagnosis. *Med Image Comput Comput Assist Interv*, 14(Pt 3):1–8, 2011.
- [140] Yanwu Xu, Jiang Liu, Zhuo Zhang, Ngan Meng Tan, D.W.K. Wong, Seang Mei Saw, and Tien Yin Wong. Learn to recognize pathological myopia in fundus images using bag-of-feature and sparse learning approach. In *Biomedical Imaging (ISBI), 2013 IEEE 10th International Symposium on*, pages 888–891, 2013.
- [141] Yanwu Xu, Ying Quan, Yi Huang, Ngan MengTan, Ruoying Li, Lixin Duan, Lin Chen, Huiying Liu, Xiangyu Chen, D.W.K. Wong, M. Baskaran, S. Perera, Tin Aung, Tien Yin Wong, and Jiang Liu. Local patch reconstruction framework for optic cup localization in glaucoma detection. In *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*, pages 5418–5421, Aug 2014.

REFERENCES

- [142] Zenglin Xu, Rong Jin, Jianke Zhu, Irwin King, and Michael R. Lyu. Efficient convex relaxation for transductive support vector machine. In *Advances in Neural Information Processing Systems*, 2008.
- [143] Fengshou Yin, Jiang Liu, Sim Heng Ong, Ying Sun, D. W K Wong, Ngan Meng Tan, C. Cheung, M. Baskaran, Tin Aung, and Tien Yin Wong. Model-based optic nerve head segmentation on retinal fundus images. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, pages 2626–2629, Aug 2011.
- [144] Fengshou Yin, Jiang Liu, D. W K Wong, Ngan-Meng Tan, C. Cheung, M. Baskaran, Tin Aung, and Tien Yin Wong. Automated segmentation of optic disc and optic cup in fundus images for glaucoma diagnosis. In *Computer-Based Medical Systems (CBMS), 2012 25th International Symposium on*, pages 1–6, 2012.
- [145] Fengshou Yin, Jiang Liu, D.W.K. Wong, Ngan Meng Tan, Jun Cheng, Ching-Yu Cheng, Yih Chung Tham, and Tien Yin Wong. Sector-based optic cup segmentation with intensity and blood vessel priors. In *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pages 1454–1457, Aug 2012.
- [146] Jun Yu, Meng Wang, and Dacheng Tao. Semisupervised multiview distance metric learning for cartoon synthesis. *Image Processing, IEEE Transactions on*, 21(11):4636 –4648, nov. 2012.
- [147] Z. Zhang, F. S. Yin, J. Liu, W. K. Wong, N. M. Tan, B. H. Lee, J. Cheng, and T. Y. Wong. ORIGA(-light): an online retinal fundus image database for glaucoma analysis and research. *Conf Proc IEEE Eng Med Biol Soc*, 2010:3065–3068, 2010.
- [148] Xiaojin Zhu. Semi-supervised learning literature survey. In *Comput. Sci., Univ. Wisconsin-Madison, Madison, WI, Tech. Rep. 1530*, 2006.
- [149] Xiaojin Zhu, Zoubin Ghahramani, and John Lafferty. Semi-supervised learning using gaussian fields and harmonic functions. In *ICML*, pages 912–919, 2003.

REFERENCES

- [150] Xiaolu Zhu and Rangaraj M. Rangayyan. Detection of the optic disc in images of the retina using the hough transform. In *Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE*, pages 3546 –3549, aug. 2008.

Appendix A

List of Publications

- i Ngan-Meng Tan; Yanwu Xu; Jiang Liu; Wooi Boon Goh; Cheung, C.; Aung, T.; Tien Yin Wong, "Domain prior based superpixel propagation for optic cup localization," In IEEE International Symposium on Biomedical Imaging (ISBI), pages 880-883, 2013
- ii Ngan-Meng Tan; Yanwu Xu; Jiang Liu; Wooi Boon Goh; Fengshou, Yin.; Tien Yin Wong, "Multi-Scale Superpixel Classification for Optic Cup Localization," In IEEE International Symposium on Biomedical Imaging (ISBI), pages 137-140, 2014
- iii Ngan-Meng Tan; Yanwu Xu; Wooi Boon Goh; Jiang Liu, "Robust Multi-Scale Superpixel Classification for Optic Cup Localization", Computerized Medical Imaging and Graphics, Available online 13 October 2014, ISSN 0895-6111, (<http://www.sciencedirect.com/science/article/pii/S0895611114001530>).

Appendix B

List of Abbreviations

Acronym	Definition
ASM	Active Shape Model
AUC	Area under the (ROC) curve
CAD	Computer-aided Diagnosis
CDR	Cup-to-Disc Ratio
CT	Computed Tomographic
ETDRS	The Early Treatment Diabetic Retinopathy Study
GVF	Gradient Vector Flow
HRT	Heidelberg Retinal Tomograph
IOP	Intraocular pressure
ISNT	Inferior, Superior, Nasal, Temporal
kNN	k-nearest neighbor
MRA	Moorfields Regression Analysis
MVEE	Minimum volume enclosing ellipse
OCT	Optical Coherence Tomography
PCA	Principal component analysis
PLSA	Probabilistic Latent Semantic Analysis
RBF	Radial Basis Function
RGB	Red, Green, Blue channels of a color image
ROC	Receiver Operating Characteristic Curve
RNFL	Retinal Nerve Fibre Layer
RSS	Residual Sum of Squares
SD-OCT	Spectral Domain Optical Coherence Tomography
SFFS	Sequential Floating Feature Selection
SLIC	Simple Linear Iterative Clustering
SS-OCT	Swept Source Optical Coherence Tomography
SSL	Semi-supervised Learning
SVM	Support Vector Machines
SVR	Support Vector Regression
TD-OCT	Time Domain Optical Coherence Tomography
WHO	World Health Organization

Appendix C

Pseudo Codes

Algorithm 1 Domain Prior Based Superpixel Propagation for Unsupervised Optic Cup Localization

```
1: procedure LEARNING AND TESTING PHASE
Input: ROI Image  $\mathcal{Q}$ 
2:    $S_c \leftarrow 1152$ 
3:    $k \leftarrow 10$ 
Output: Localized Cup Region  $pred\_cup$ 
4:    $\mathcal{BV} = \text{Blood\_Vessel\_Extraction}(\mathcal{Q})$ 
5:    $\mathcal{VJ} = \text{Central\_Retinal\_Vessel\_Junction}(\mathcal{BV})$ 
6:    $\mathcal{PR}_1 = \text{Optic\_Cup\_Origin\_Prior\_Point\_Region}(\mathcal{VJ})$ 
7:    $\mathcal{IT}_G = \text{Intensity\_Thresholding\_Green\_Channel}(\mathcal{Q})$ 
8:    $\mathcal{IT}_L = \text{Intensity\_Thresholding\_L\_Channel}(\mathcal{Q})$ 
9:    $\mathcal{PR}_2 = (\mathcal{IT}_G + \mathcal{IT}_L)/2$ 
10:   $SP = \text{SLIC\_Superpixel\_Segmentation}(\mathcal{Q}, S_c)$ 
11:   $\mathbf{F} \leftarrow \{\}$ 
12:  for  $j \leftarrow 1 : |SP|$  do      /* j-th superpixel  $SP_j$  */
13:     $f_j = \text{Feature\_Extraction}(SP_j)$ 
14:     $\mathbf{F} \leftarrow [\mathbf{F}; f_j]$ 
15:   $\mathbf{B} = \text{k\_Means}(k, \mathbf{F}, \mathcal{Q}, \mathcal{BV})$  /* k-means codebook  $\mathbf{B}$  */
16:  for  $j \leftarrow 1 : |SP|$  do
17:     $l_i = \text{Label\_Init\_Similarity}(\mathcal{PR}_1, f_j, \mathbf{F})$ 
18:     $l'_i = \text{k\_Means\_Refinement}(l_i, f_j, \mathbf{B})$ 
19:     $l''_i = \text{Pallor\_Prior\_Refinement}(l', \mathcal{PR}_2)$ 
20:   $pred\_cup = \text{Ellipse\_Fitting}(l''_i, u, v, \nu)$  /* Ellipse fitting */
```

Algorithm 2 Multi-Scale Superpixel Classification for Supervised Optic Cup Localization - Learning Phase

```
1: procedure LEARNINGPHASE
Input: Training Disc Images  $\mathcal{K}_i$  ( $i=1, \dots, 150$ ) with Ground-Truth Cup Masks  $C_i$ 
2:    $S_c \leftarrow \{512, 1024, 2048\}$ 
3:    $\beta \leftarrow |S_c| = 3$ 
Output: Selected base models  $\Omega$ , Final integration model  $\Psi$ 
4:   for  $i \leftarrow 1 : 150$  do
5:      $Blood\_Vessel\_Extraction(\mathcal{K}_i)$ 
6:      $\mathcal{K}'_i = Contrast\_Normalization(\mathcal{K}_i)$ 
7:     for  $S_c \leftarrow \{512, 1024, 2048\}$  do
8:        $SP^c = SLIC\_Superpixel\_Segmentation(\mathcal{K}'_i, S_c)$ 
9:        $L^c = Get\_Superpixel\_Label(SP^c, C_i)$ 
10:      for  $j \leftarrow 1 : |SP^c|$  do /* j-th superpixel  $SP^c_j$  */
11:         $f^c_j = Feature\_Extraction(SP^c_j)$ 
12:      for  $S_c \leftarrow \{512, 1024, 2048\}$  do
13:         $\mathbf{W}^c \leftarrow \{\}$ 
14:        for  $\alpha \leftarrow 1 : 100$  do /* Base Models */
15:           $w^c_\alpha = Linear\_SVM\_Training(f^c, L^c)$ 
16:          /* Training with equal positive and negative random samples */
17:           $\mathbf{W}^c \leftarrow [\mathbf{W}^c; w^c_\alpha]$ 
18:       $\omega = Sparse\_Model\_Selection(f^c, L^c, \mathbf{W}^c)$ 
19:       $\Psi \leftarrow \{\}, \Omega \leftarrow \{\}$ 
20:      for  $i \leftarrow 1 : \alpha\beta$  do
21:        if  $abs(\omega_i) > 10^{-3}$  then
22:           $\Omega \leftarrow [\Omega; w_i]$ 
23:           $\Psi \leftarrow [\Psi; \omega_i]$ 
```

Algorithm 3 Multi-Scale Superpixel Classification for Supervised Optic Cup Localization - Testing Phase

```

1: procedure TESTINGPHASE
Input: Test Disc Image  $\mathcal{Q}$ 
2:    $S_c \leftarrow \{512, 1024, 2048\}$ 
3:    $\beta \leftarrow |S_c| = 3$ 
4:    $\mathbf{F}^c \leftarrow \{\}$ 
Output: Localized Cup Region, ( $pred\_cup$ )
5:    $Blood\_Vessel\_Extraction(\mathcal{Q})$ 
6:    $\mathcal{Q}' = Contrast\_Normalization(\mathcal{Q})$ 
7:   for  $S_c \leftarrow \{512, 1024, 2048\}$  do
8:      $SP^c = SLIC\_Superpixel\_Segmentation(\mathcal{Q}', S_c)$ 
9:     for  $j \leftarrow 1 : |SP^c|$  do      /* j-th superpixel  $SP_j^c$  */
10:       $f_j^c = Feature\_Extraction(SP_j^c)$ 
11:       $\mathbf{F}^c \leftarrow [\mathbf{F}^c; f_j^c]$ 
12:    $pred = \Psi^\top \Omega^\top \mathbf{F}^c$ 
13:    $pred\_cup = Ellipse\_Fitting(pred, u, v, \nu)$       /* Ellipse fitting */

```