AUTOMATIC SEGMENTATION OF INTRA-RETINAL CYSTS FROM OPTICAL COHERENCE TOMOGRAPHY SCANS

A THESIS

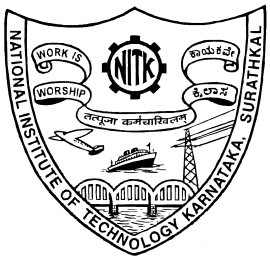
Submitted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

by

GIRISH G N

(148006CS14F07)



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

NATIONAL INSTITUTE OF TECHNOLOGY KARNATAKA
SURATHKAL, MANGALORE - 575025, INDIA

OCTOBER, 2018

ಮಾತೃ ದೇವೋ ಭವ, ಪಿತೃ ದೇವೋ ಭವ, ಆಚಾರ್ಯ ದೇವೋ ಭವ

मातृ देवो भव, पितृ देवो भव, आचार्य देवो भव

MATRU DEVO BHAVA, PITRU DEVO BHAVA, ACHARYA DEVO BHAVA

This thesis is
dedicated to
My Parents and Teachers.

DECLARATION

I hereby ***declare*** that the Research Thesis entitled **AUTOMATIC SEGMENTATION OF INTRA-RETINAL CYSTS FROM OPTICAL COHERENCE TOMOGRAPHY SCANS** which is being submitted to the ***National Institute of Technology Karnataka, Surathkal*** in partial fulfillment of the requirements for the award of the Degree of ***Doctor of Philosophy*** is a ***bona fide report of the research work carried out by me***. The material contained in this thesis has not been submitted to any University or Institution for the award of any degree.

GIRISH GN
Register No.: 148006CS14F07
Department of Computer Science and Engineering
National Institute of Technology Karnataka
Surathkal-575025

Place: NITK - Surathkal

Date: 25 OCTOBER 2018

CERTIFICATE

This is to *certify* that the Research Thesis entitled **AUTOMATIC SEGMENTATION OF INTRA-RETINAL CYSTS FROM OPTICAL COHERENCE TOMOGRAPHY SCANS**, submitted by **GIRISH G N** (Register Number: 148006CS14F07) as the record of the research work carried out by him, is *accepted* as the *Research Thesis submission* in partial fulfillment of the requirements for the award of degree of ***Doctor of Philosophy***.

Dr. Jeny Rajan
Research Guide
Assistant Professor
Department of Computer Science and Engineering
National Institute of Technology Karnataka
Surathkal-575025

Chairman - DRPC
Department of Computer Science and Engineering
National Institute of Technology Karnataka
Surathkal-575025
(Signature with Date and Seal)

ACKNOWLEDGEMENTS

It has been a great privilege to be a part of the student fraternity of this esteemed organization, National Institute of Technology Karnataka (NITK), Surathkal, India, during my doctoral research. I would like to acknowledge a deep sense of gratitude to Department of Computer Science and Engineering (CSE), NITK, for bestowing me with an opportunity to undertake and successfully complete my life-time dream of receiving a doctorate.

I submit my whole-hearted reverence to my supervisor, Dr. Jeny Rajan, Assistant Professor, Department of CSE, NITK Surathkal, India, for his continuous guidance and support. I would also take this opportunity to convey my heartfelt reverence to my research collaborators and mentors, Dr. Abhishek R. Kothari, Pink City eye and retina center, Jaipur, India, and Dr. Sohini Roychowdhury, Assistant Professor, Department of Electrical and Computer Engineering, University of Washington, Bothell, USA; for their invaluable guidance and encouragement throughout my research. My research supervisor, and mentors' fondness for ethics, and their immense expertise in research, provided strong foundation for my further ventures in my area of interest. I am indebted to these three scholars in guiding me through my doctoral studies.

I am greatly indebted to highly insightful experts in my Doctoral Research Progress Assessment Committee- Dr. Basavaraj Talwar , Assistant Professor, Department of CSE, NITK Surathkal, India and Dr. A.V. Narasimhadhan, Assistant Professor, Department of ECE, NITK Surathkal, India. Their timely suggestions provided clarity and confidence to strive towards successful completion of my research. Their constructive feedback resulted in continual flow of ideas for completion of my research thesis. I am thankful to Dr. Manu Basavaraju, Assistant Professor, Department of CSE, NITK Surathkal, India and Dr. Shashidhar G. Koolagudi, Associate Professor, Department of CSE, NITK Surathkal, India for their support and assistance during my studies at NITK Surathkal. I would also like to thank Dr. Alwyn Roshan Pais, Associate Professor and Chairman, Doctoral Research Program Committee, Department of CSE, NITK Surathkal for his support and valuable suggestions during my doctoral studies.

I whole heartedly express my gratitude to Prof. P. Santhi Thilagam, Head of the Department, Department of Computer Science, NITK Surathkal. It is my honour to have received her invaluable support all along my time here at this institution. I would be failing in my duty if I do not acknowledge the services and support by the faculty and staff of Department of CSE, NITK Surathkal, for providing access to valuable resources for successful completion of my studies. I feel immensely proud to acknowledge the

facilities and kindhearted support by Prof. K. Uma Maheshwar Rao, respected Director of NITK Surathkal.

I would like to thank Dr. Jing Wu, Christian Doppler Laboratory for Ophthalmic Image Analysis, Department of Ophthalmology, Medical University of Vienna, for providing OPTIMA cyst segmentation challenge dataset, and Prof. Sina Farsiu, Vision and Image Processing Lab, Department of Biomedical Engineering, Duke University, for making available the Duke DME dataset crucial for the present study.

I am ever indebted to Dr. Shrinivasa Naika C.L., Assistant Professor, Department of CSE, University B.D.T. College of Engineering, Davanagere. He is the primary reason for my inclination towards research and has always been a pillar of support throughout my doctoral studies as well. He introduced and encouraged me to venture out in the field of research, and has been helping me all along, in academic as well as non-academic lines. He started off as a teacher in my life, and ended up being my Guru. I should also thank his family, for providing warmth of home, consoling and pacifying me during crests and troughs of my life.

I should thank my prior organization, Karnataka State Council for Science and Technology (KSCST), IISc Campus, Bangalore, for teaching me the way to move forward in research, during my stay there as a Project Associate in the pilot project. Special thanks to my Principal Investigator and Executive Secretary Dr. S.G.S. Swamy, my colleagues Priyanka, Venkatesh and other staff for their support during my employment period there.

I heartily thank the Department of Biotechnology, Government of India, for awarding me with the Travel Grant to be a part of "38th International Conference of IEEE Engineering in Medicine Biology Society" at Orlando, USA on 2016. I also thank my friends Ashwini H.S., Avin, Bharath, Pradeep, Arvind Gautham, Radhika, and Mr. & Mrs. Corne for making this visit memorable by assisting in travel, visa and accommodation.

It is needless to say that this sort of research venture would have never been possible without the blessings and cooperation of all my family members. My grandfather, Late. N.V. Thipperudrappa, Retd. teacher and his childhood best friend Dr. Sanna Honnappa, MD, are those who monitored me closely and inspired me to study, which eventually led me towards doctoral research. I lost my grandparents (Late. N.V. Thipperudrappa and Late. Sharadamma) when I was about to begin my doctoral research and eventually, his best friend Dr. Sanna Honnappa filled my grandfather's place.

My eternal love to my parents, G. Nagarajappa and T.N. Annapoornamma, my beloved sisters G.N. Ranjitha, Bhoomika A.P., Rashmi P., Sowmya Virupaksha, Chandrakala, Sanidhya S.C., Anusha C.S., maternal grandparents S.P Rudramuni and Prabhavathi G.S. for raising me during my diploma course. Uncle Dr. S.R. Channaveerappa, aunt T.N. Nagarathna, and cousins T.N. Veeresh, R. Siddesh, R. Annapurna, Sulok S.C., Dileep G.S., Naveen G.S., Vishwanth G.R., Gurumurthi G.P., Mahadeva Swamy G.R.

and others have always stood by me with a smile, to lend a supporting hand in fulfilling all my responsibilities.

A special thanks to my closest of friends, Rahul K., Prabhu Prasad B.M., Pavan Kumar B.N, Virupaksha H.S., Gadhigesh Kumar G., Ullas K., Anand C., Muralidhara S.R., Soumya Vaidya, Arvind Belure, Kiran G.V., Darshan M.K., and Yuvraj H.P. who were there with me in all ups and downs of my life. When I was planning to do my PhD, and decided to continue my career in R&D, they have unconditionally supported me morally, financially and physically. This thesis would never have been so effortless without the continuous support of these people.

I would like to express my gratitude from the bottom of my heart to admirable friends Ms. Chaitanya S. Nayak, research scholar, Department of MACS, NITK Surathkal and Ms. Swathi S., research assistant, IIT Madras for their exemplary patience in reading and providing critical reviews for my articles and reports, including corrections of my grammatical errors, which has contributed to the quality of this research work. Both of them were with me during my good and bad days at NITK.

My special thanks to my research lab partners and colleagues- Anil Kumar, Bheemappa, Yamanappa, Chetan Srinidhi, Sharmas Vali, Keerthi Shetty, Khyamling, Dr. Raghu J. Mandya, Mahadevi, Manjunath Mulimani, Manjunath K.V., Manjunath S.H., Sooraj Bhat, Vishnu Swaroop, Shashidhar, Sanjay Bankapur, Praveen, Sachin, Vishal, Shiva, Ambikesh, Prakash, Uday, Dr. Nagaraj P.Y., Sunil Meti, Chandra Shaker Balure, Anoop B.N., YVS Murthy, Saraswathi, Ravi, Bijay, Dr. Krishna Kumar P., Dr. Venkatesh Lamani, Dr. Goutham Simha G.D., and Dr. Sudeep P.V. for their support during my stay at NITK. I would also like to thank my friends from G.V. Ploytechnic, J.N.N. College of Engineering and U.B.D.T. College of Engineering.

I have learned a lot as a mentor to my beloved M.Tech students- Naren Rao, Anima V.A., Remya K. Sudheesh, Bibhash Thakur, Priya Suryavanshi, Banoth Saikumar; and my B.Tech students- Gaurav Verma, Abhishek, Aadheeshwar, Suchand, Ankit Tripathi, Guru Pradeep Reddy, Ashritha Kandiraju, Prajwala T.M.. I am thankful for their enthusiasm and brain storming, during my stay at image processing lab.

Life at NITK gave me seamless avenues for innovation, exposure, learning, development along with entertainment through extra-curricular activities. I am thankful to Kannada Vedike cultural club, Incident committee and Student council for providing opportunity to utilize my leadership and management skills in conducting annual college festivals and events such as, “Incident”, “Parva”, “Ek Bharath, Shresht Bharath” and others. I am grateful to Prof. Suresh M. Hegde, Dean Students’ Welfare, Dr. Ramesh M.R., Faculty Advisor of Kannada Vedike Club, Dr. Manoj and Dr. Hem Prasad Nath, SAS officers, for always motivating me during the organization of these events.

I am very happy to have a big family from Kannada Vedike including Meti, Manoj, Vikas, Thejas, Bansode, Pavan Walikar, Kishor, Ullas, Yogeesh, Yatish, Tsunami Pra-

jwal, Keerthana, Anusha Betta, Anikethan, Sachin Karigar, Akash Borale, DG, Ashay, Raju, Sourabh, Gagan Sappani, Harshith, Madhu, Bharath, Abhishek Nayak, Raghu, Koushik, Santhosh, Teena, Shreyasvi, Arpitha, Venki, Krishnamurthy, Thilak, Sudarshan, Niranjan, Suma, Sanjana, Likitha, Pratyush Giri, and 4B, 5A, 5B, D wings' hostel mates, and others for working with me in organizing all the events successfully.

Everyone always believed that, I am smart guy, hardworking and a topper. On the contrary, I have an untold story about my academic career. There was once a time when I was a school dropout. It took several years to fight all the hurdles and brace myself up. What I have learnt from this and would like to convey to the new and upcoming research aspirants, is that, self help is the best help and patience is the key.

Finally, I would love to thank everyone who is directly and indirectly responsible for the successful completion of my doctoral research work. . . :-)

GIRISH G N

Place: NITK - Surathkal

Date: 25 OCTOBER 2018

ABSTRACT

Retinal cysts are formed by accumulation of fluid in the retina caused by leakage due to blood retinal barrier breakdown from inflammation or vascular disorders. Analysis of retinal cystic spaces holds significance in detection, treatment and prognostication of several ocular diseases like age-related macular degeneration, diabetic macular edema, etc. Segmentation of intra-retinal cysts (IRCs) and their quantification is important for retinal pathology and severity characterization. In recent years, automated segmentation of intra-retinal cysts from optical coherence tomography (OCT) B-scans has gained significance in the field of retinal image analysis.

In this thesis, a benchmark study is conducted to compare existing methods to identify the factors affecting IRC segmentation from OCT scans. A modular approach is employed to standardize the different IRC segmentation algorithms followed by analysis of variations in automated cyst segmentation performances and method scalability across image acquisition systems are done by using publicly available cyst segmentation challenge dataset (OPTIMA cyst segmentation challenge). Such exhaustive analysis on the scalability of OCT cyst segmentation methods in terms of methodological and input data variations has not been done before.

An efficient cyst segmentation technique must be capable of performing cyst identification and delineation with minimum errors. Several methods proposed in the literature fail to delineate cysts up to their true boundary. To address this problem, an unsupervised vendor dependent method using marker controlled watershed transformation is proposed in this thesis. The method is based on two stages- k-means clustering technique is used to identify cysts in the form of marker, followed by topographical based watershed transform for final segmentation. Qualitative and quantitative evaluation of

the proposed method is carried out against ground truth obtained from two graders on OPTIMA cyst segmentation challenge dataset (Spectralis Vendor OCT scans). Obtained results show that the proposed method outperformed other considered unsupervised methods.

Several segmentation methods have been proposed in the literature for IRC segmentation on vendor-specific OCT images, but these lack generalizability across imaging systems. To address this issue, a fully convolutional network (FCN) model for vendor-independent IRC segmentation is proposed in this thesis. The proposed FCN was trained using the OPTIMA cyst segmentation challenge dataset (with four different vendor-specific images, namely, Cirrus, Nidek, Spectralis and Topcon). This method counteracts image noise variabilities and model over-fitting by data augmentation and hyper-parameter optimization. Additionally, sensitivity analysis of the model hyper-parameters (depth and receptive field size) is performed to optimize the proposed FCN. The Dice Correlation Coefficient of the proposed method outperforms the algorithms published in the OPTIMA cyst segmentation challenge.

Deeper FCNs exhibit better feature learning capabilities than shallower networks but those are computationally intensive due to large number of computation parameters and may be prone to vanishing gradient problem. To address this issue, a depthwise separable convolutional filter based end-end convolutional neural network architecture with swish activation functions is proposed in this thesis. OPTIMA cyst segmentation challenge dataset with four different vendor scans were used to evaluate the proposed architecture for vendor independent IRC segmentation task. Obtained experimental results show that the proposed method significantly reduced the number of computation parameters compared to regular convolution based FCN.

Keywords: Optical Coherence Tomography; Segmentation; Retinal Image Anaysis; Retinal Cyst; Cystoid Macular Edema; Convolutional Neural Networks; Deep Learning; k-Means Clustering; Watershed Transformation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBRIVATIONS AND NOMENCLATURE

S_σ	Smoothing parameter of NLM filter
$\hat{\beta}_{\text{ML}}$	Estimated scale parameter of Gamma distribution.
$\hat{\rho}_{\text{ML}}$	Estimated shape parameter of Gamma distribution.
$\hat{\sigma^2}$	Estimated noise variance.
$\hat{\sigma_d}$	Geometrical spread parameter.
$\hat{\sigma_r}$	Photometric spread parameter.
$\hat{B}_e$	Estimated bias.
$\sigma()$	Sigmoid function.
BCL_{net}	Binary cross entropy loss funcion of the nueral network.
$C_{\min}$	Cluster with minimum centroid.
CB	Catchment basin.
CB_a	Catchment basin area threshold.
CB_h	Width to height ratio of catchment basin.
$CB_{\max}$	Maximum intensity threshold of catchment basin.
CL_i	Cross entropy loss for input i.
G_1	Grader 1.
$G_1 \cap G_2$	Intersetion of both graders.
G_2	Grader 2.
k	Number of clusters.
k-NN	k-nearest neighbor classifier.
k_h	Height of the convolution filter.
k_w	Width of the convolution filter.

M_k	Marker image.
N	Number of filters in a convolutional layer.
R_h	Region width to height ratio.
R_p	Region area.
$R_e()$	Rectified linear unit activation function.
$S()$	Segmentation function.
$s()$	Swish activation function.
S_E	Minimum squared error criterion.
S_t	Structuring element.
T_{depth}	Total number of parameters in a depthwise separable convolutional layer.
T_{Loss}	Total loss value of nueral network.
T_{param}	Total number of parameters in a regular fully convolutional layer.
T_{point}	Total number of parameters in point-wise convolution stage.
T_{sep}	Total number of parameters in separable convolution stage.
2D	2-Dimensional.
3D	3-Dimensional.
AG	Amsler Grid.
ANN	Artificial neural network.
AS	Automaic segmented.
AUB	Adaptive and unbiased bilateral.
BM	Bruchs membrane.
BMEIS	Boundary of myoid and ellipsoid of inner segments.
BNLM	Bayesian non local means.
CFP	Color fundus photography.
CLAHE	Contrast-limited adaptive histogram equalization
CME	Cystoid macular edema.

CNN	Convolution neural network.
DC	Dice coefficient.
DME	Diabetic macular edema.
DSCN	Depthwise separable convolutional neural network.
ELM	External limiting membrane.
FA	Fundus fluorescein angiography.
FCN	Fully convolutional neural network.
FN	False negative.
FNLM	Fast non local means.
FP	False positive.
GCL	Ganglion cell layer.
GLCM	Gray level co-occurrence matrix.
GT	Ground Truth
GTDP	Graph theory and dynamic programming.
ILM	Internal limiting membrane.
INL	Inner nuclear layer.
IPL	Inner plexiform layer.
IRC	Intra-retinal cyst.
KR	Kernel regression.
ML	Maximum likelihood.
MME	Micro-cystic macular edema.
MWS	Marker-controlled watershed.
NB	Naive Bayes.
NFL	Nerve fiber layer.
NLM	Non local means.
OCSC	OPTIMA cyst segmentation challenge.

OCT	Optical coherence tomography.
OCTSEG	Optical coherence tomography segmentation and evaluation GUI tool.
ONL	Outer nuclear layer.
OPL	Outer plexiform layer.
PED	Pigment epithelial detachment.
ReLU	Rectified Linear Unit
RF	Random forest.
RNFL	Retinal nerve fiber layer.
ROI	Region of interest.
SD-OCT	Spectral domain optical coherence tomography.
SNR	Signal-to-noise ratio.
SRF	Sub-retinal fluid.
SVM	Support vector machine.
TN	True negative.
TP	True positive.
UFNLM	Unbiased fast non local means.
VA	Visual Acuity.
WHO	World Health Organization.

CHAPTER 1

INTRODUCTION

1.1 Human Eye and Retina

The human visual system mainly comprises of two functional parts- eye and brain. Eye acts as camera to capture light information that is converted to electrical impulses and transmitted to the brain, which functions as image processing module for perception. Eye is the second most complex organ in human body after the brain. Anatomy of the human eye is shown in Figure. 1.1. Eye has various components performing unique functions that coordinate with each other and with other parts of the human body facilitating three dimensional, colored moving visualization in a variety of light conditions.

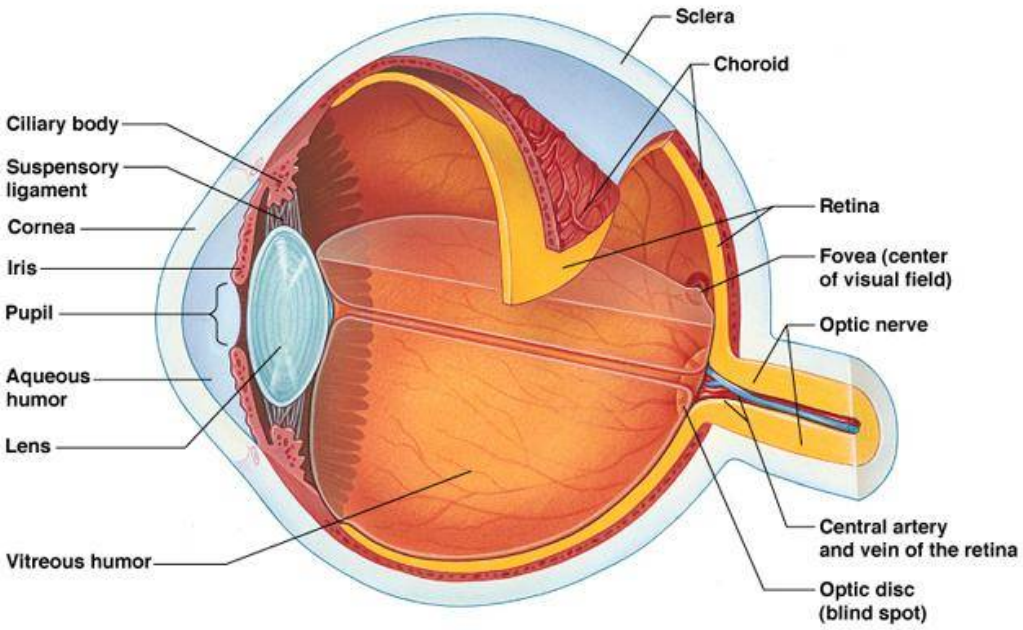


Figure 1.1: Anatomy of the human eye. (Hogan and JA Weddell, 1971)

The eye's anatomical structures are arranged in three layers, namely outer fibrous tunic, intermediate vascular tunic and the innermost retina. The outer protective fibrous tunic is composed of cornea and sclera. Cornea transmits and focuses lights into the eye. Sclera is continuous with the cornea and forms the supporting wall of the eyeball. The middle layer, or the vascular tunic, consists of choroid, ciliary body, pigmented epithelium and iris. Iris is visible as the coloured part of the eye that functions like a diaphragm and helps in regulating the amount of light that enters the eye. The anterior chamber between the cornea and iris as well as the posterior chamber between the lens and the iris is occupied by the aqueous humour, a fluid that nourishes the anterior structures.

The lens is a transparent structure that focuses light rays onto the retina. Lens is attached to the ciliary body by suspensory ligament, which transmit muscular forces for changing the lens shape to focus on objects at various distances. The vitreous body is a clear, jelly-like substance that fills the eye. The innermost layer or retina, gets oxygenation from choroidal blood vessels and retinal vessels. The retina is a multilayered neural tissue that lines the back of the eye, senses light and creates electrical impulses that travel through the optic nerve to the visual cortex, which is the part of the brain that controls our senses of sight.

Retina is a thin optically transparent tissue consisting of several layers. The main function of the retina is to convert received light to neural signals, which are forwarded to brain for visual perception. Figure. 1.2 shows cross sectional view of the retina. Retinal layers are organized into two distinct functional components namely neurosensory retina and retinal pigment epithelium. Neurosensory layers of the retina contains total 9 layers. Starting from the vitreous surface, they are the internal limiting membrane (ILM), retinal nerve fiber layer (RNFL), ganglion cell layer (GCL), inner plexiform layer (IPL), inner nuclear layer (INL), outer plexiform layer (OPL), outer nuclear layer (ONL), external limiting membrane (ELM) and the layer of photoreceptors. Retinal pigment epithelium (RPE) is the outermost layer of the retina, this layer provides metabolic

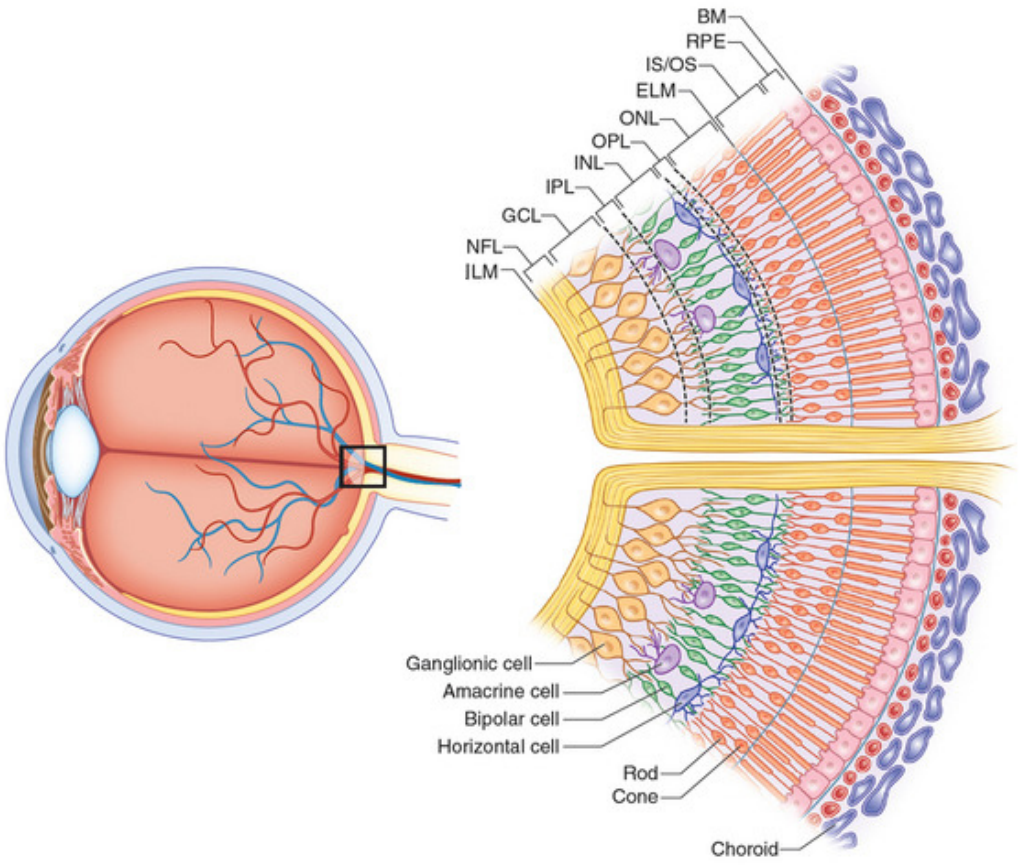


Figure 1.2: Cross sectional view of the retina. (Image courtesy: (Retina, 2015))

support to the nuerosensory layers of the retina and it absorbs incident scattered light.

1.2 Retinal Cyst and Cystoid Macular Edema

A retinal cyst is a fluid-filled space in the retina, and the presence of edema and cysts increases the thickness of the involved retina. Retinal cysts have pathological significance and are present in several eye disorders like age related macular degeneration, diabetic retinopathy, retinal vein occlusion, ocular inflammation, diabetic macular edema (DME) etc. Such *intra-retinal cysts* (IRCs) in the macular region (Cystoid Macular Edema) is an important impediment to retinal function and can cause central vision loss

(WHO, 2007).

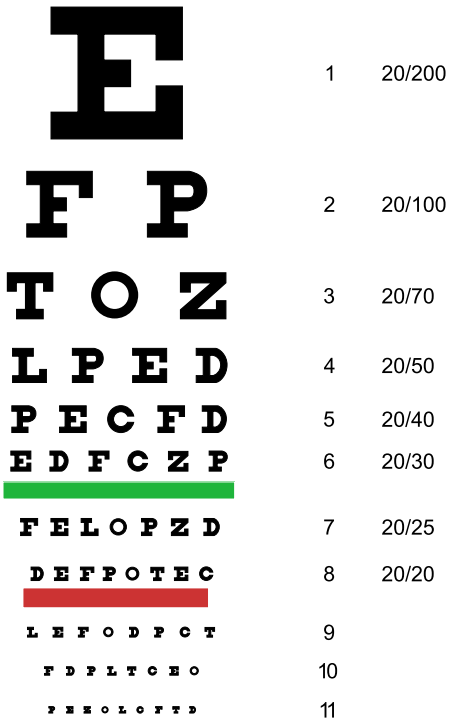
Cystoid macular edema (CME) following cataract surgery was initially reported by Irvine in 1953 and is known as the Irvine-Gass syndrome (Irvine, 1953). CME develops when excess fluid accumulates within the macular retina. This is thought to occur following disruption of the blood retinal barrier because of the aforementioned eye disorders. This fluid accumulation in retina reduces macular retinal function. Medical studies shows that visual acuity can be reasonably correlated from the volume of retinal cystic fluids and their location in the retina (Wilkins *et al.*, 2012).

1.3 Diagnosis of CME

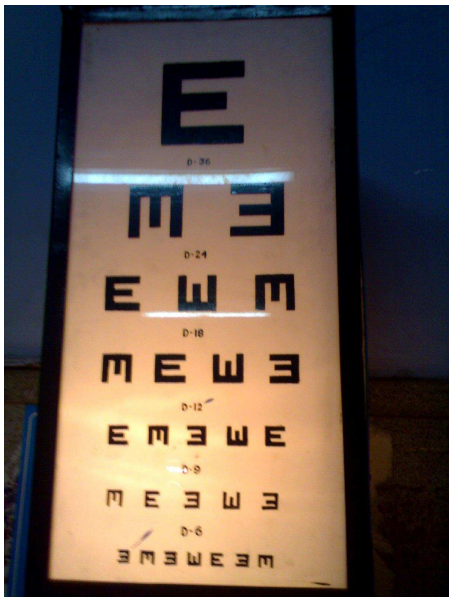
Primary symptoms of CME are decreased, blurry or wavy vision, faded colors, and distortion near or at the center of vision. Ocular diseases can be characterized using several techniques. Initial Visual Acuity (VA) test is used to measure central vision according to the size of letters or symbols viewed on Snellen chart or E chart respectively. Figure. 1.3 shows an example of Snellen and E chart.

Amsler Grid test (AG) is used to detect visual distortion caused in macular disorders. Amlser Grid contains horizontal and vertical lines with a small dot at the center of the grid and is shown in Figure. 1.4(a). Patient is asked to view the AG from each eye separately. If the patient has macular pathology, he/she may notice distortion or they may be unable to see some lines in the AG. Both VA and AG tests are psychophysical tests and needs patient cooperation. For clinical anatomical evaluation of the macula, ophthalmoscopy or slit lamp biomicroscopy test (with a 78 or 90 diopter aspheric lens) is used. Figure. 1.4(b) shows slit lamp system.

Imaging techniques can also be used for the detection of cystic fluids in the retina. Color fundus photography (CFP) is used for imaging of the macula. With stereoscopic view of the retina using CFP, ophthalmologist can identify thickening of retina that is

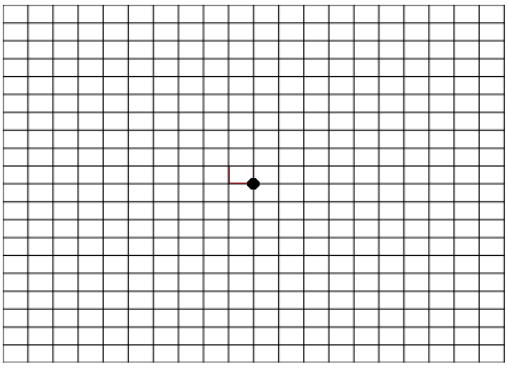


(a) Snellen chart



(b) E chart

Figure 1.3: Different type of charts used in visual acuity test.



(a)

(b)

Figure 1.4: Amsler grid and slit lamp machine.

(a) (b)

Figure 1.5: Color fundus photography and fundus fluorescein angiography scan of a retina with CME.

caused due to intra-retinal fluid accumulation. Fundus fluorescein angiography (FA) is widely available and used for dynamic evaluation of the retinal vasculature. In FA, the amount of fluorescein leakage depends on the dysfunction of the retinal vascular endothelium. Apart from being a significant diagnostic modality, Fundus Fluorescein Angiography also improves the accuracy for planning the treatment of CME. Figure. 1.5 shows CFP and FA images of a retina with CME. Figure. 1.5(b) shows fluorescein leakage into the cystoid spaces is distributed radially in retina forming the classic honeycomb like structure appearance near fovea. Slit-lamp bio-microscopy, fundus photography and FA techniques gives 2D *en face* view of the retina and are prone to differences in subjective assessment to quantify and diagnose retinal pathology. An advanced imaging technique called optical coherence tomography (OCT) was introduced in clinical practice for the purpose of high resolution cross-sectional imaging (as opposed to *en face*) and for objective measurement of retinal structures.

1.4 Optical Coherence Tomography

OCT is a non-invasive imaging technique that uses low coherence light for resolving internal structures of biological tissues and obtains cross-sectional, high-resolution images of the retina and can detect retinal thickness (Huang *et al.*, 1991). It is a widely used diagnostic technique for several ocular diseases.

Figure. 1.6 shows a schematic diagram of OCT Scanner. A low coherence light source (LCS) emits light towards beam splitter (BS). Part of this beam is projected on to reference mirror (RM) and the rest to the sample (SMP) to be scanned. Reflected light from both RM and SMP are combined at BS, creating interference patterns based on the interferometry concept. When the reflected light beams from RM and SMP are *in phase*, this generates constructive inference. At other times, it results in destructive inference. Photo detector (PD) receives combined reflected light with interference pattern and forward the signal to a digital signal processor (DSP) unit. DSP unit uses Fourier transform to create a linear A-scan of the point scanned. Several adjacent A scans along a linear path of scanning are combined to give a B-scan image. An OCT scan typically consists of several B scan images taken over a fixed area, with a fixed protocol.

Figure. 1.7 shows a commercially available OCT Scanner. OCT provides excellent visualization of several retinal structures like cysts, exudates and retinal layers. Figure. 1.8 shows OCT B-scan of the retina with *en face* macular projection. The dark green line with arrow head shown on the *en face* macular image (left side of the figure) corresponds to the projection line of the OCT B-scan shown on right side of the figure. Figure. 1.9 shows OCT scans of the retina with normal and cystic fluids obtained using Cirrus and Spectralis vendor OCT machines.

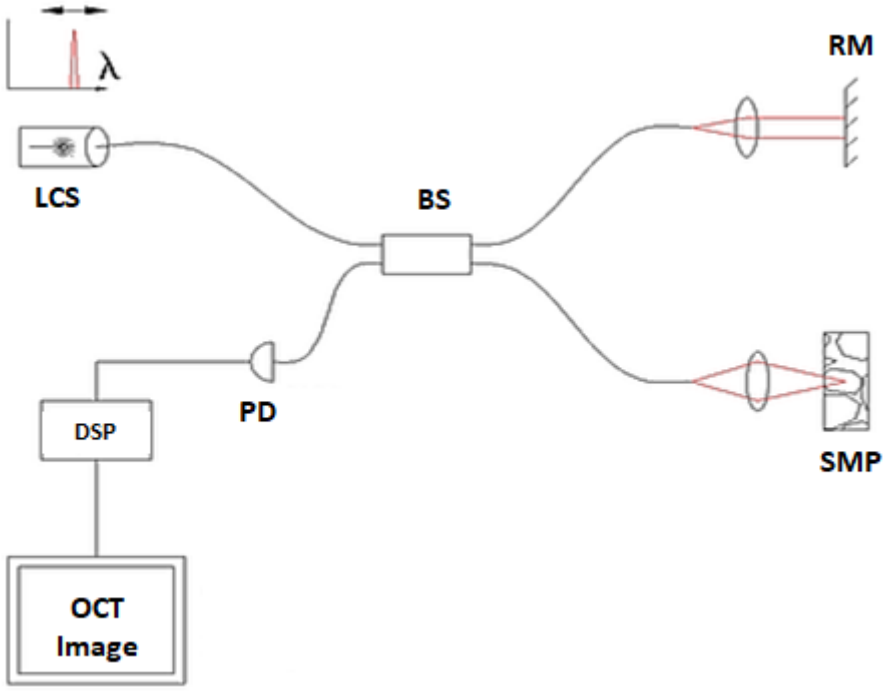


Figure 1.6: Schematic diagram of OCT Scanner. Components include: low coherence light source (LCS), beamsplitter (BS), reference mirror (RM), sample (SMP), photo detector (PD), and digital signal processor (DSP). (Image courtesy: Wikipedia)

Figure 1.7: Cirrus optical coherence tomography scanner.

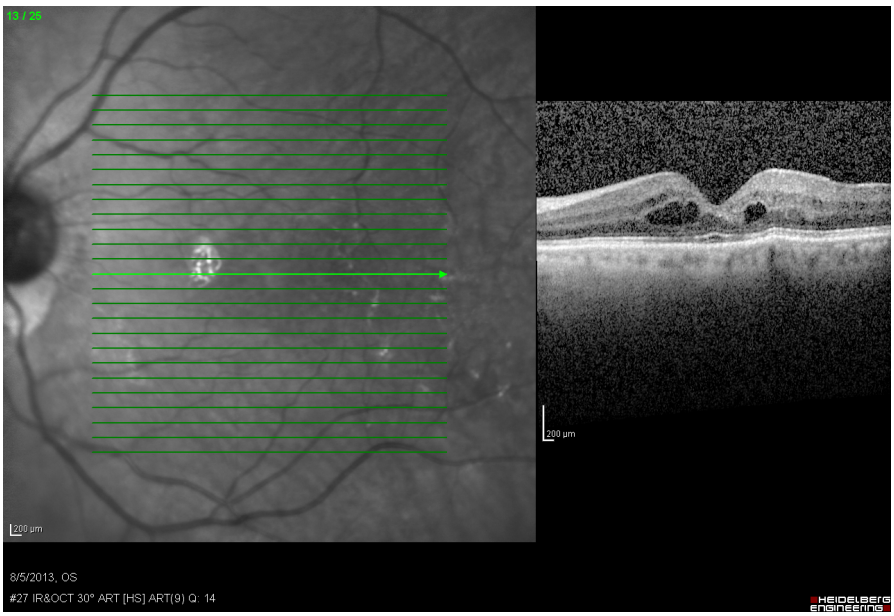
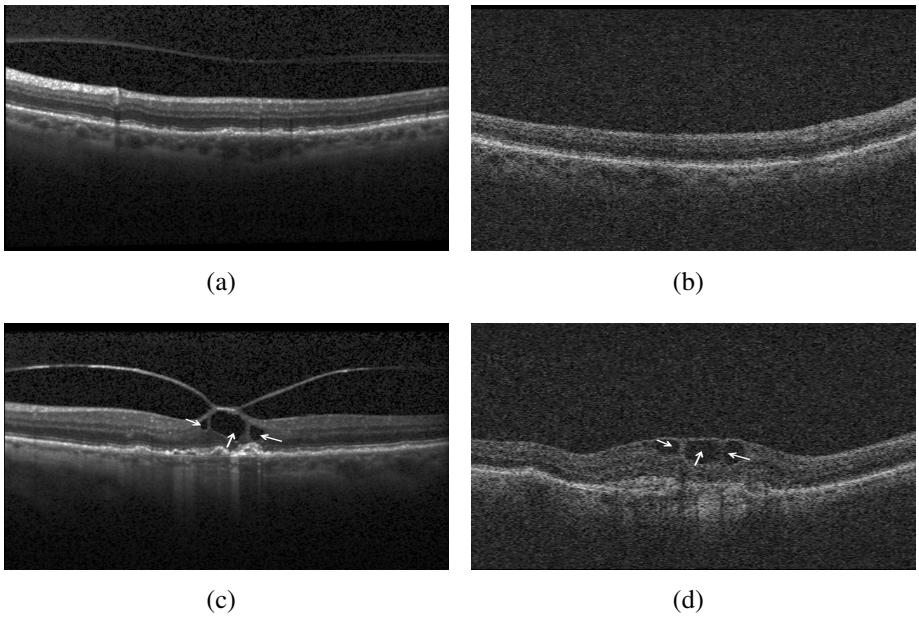


Figure 1.8: *En face* macular projection and its corresponding OCT B-Scan of a retina. (Image Courtesy: http://eyewiki.aao.org/ (Garg, 2014))



(a) (b)

(c) (d)

Figure 1.9: Optical coherence tomography scans of retina from two different vendors: (a), (b) normal retina; (c), (d) retina with cystoid macular edema. (White colored arrow shows cystic fluids). (a) and (c) obtained from Spectralis imaging system, (b) and (d) obtained from Cirrus imaging system. (Image courtesy: OPTIMA cyst challenge dataset (OPTIMA, 2015))

.

1.5 Motivation and Problem Statement

World Health Organization (WHO) estimated that by the year 2020, there will be 75 million blind people and 314 million partially blind people in the world (WHO, 2007). Visual morbidity from CME is a major public health problem globally. If detected early and treated effectively, most of this visual impairment from CME can be prevented or cured.

Medical studies show that visual acuity can be fairly correlated to the location and volume of retinal thickening, intra-retinal fluid and retinal cystic fluid spaces in retinal tissue (Wilkins *et al.*, 2012). Manually analyzing OCT data for the detection of cysts and measuring their volume is time consuming and needs expertise, and currently is not performed routinely, leading to qualitative assessment of the scan alone. Automated cyst segmentation and quantification methods enable accurate quantification for faster diagnosis and treatment.

Development of automated cyst segmentation techniques for OCT scans pose several challenges such as, speckle noise in the scans, poor signal-to-noise ratio (SNR), retinal vessel shadows, and scan intensity difference across the vendors.

1.5.1 Problem Statement

Spatial location and volume of accumulated intra-retinal cystic fluids in the retina can be used for quantification of anatomical disturbance due to disease in the retina and can be correlated to visual morbidity due to cystoid macular edema. This thesis aims to identify factors affecting intra-retinal cyst segmentation from optical coherence tomography scans. Furthermore, this thesis reports on the development of unsupervised and supervised automatic intra-retinal cyst segmentation techniques using optical coherence tomography scans for aiding diagnosis process in cases with cystoid macular edema.

Research Objectives:

1. Identification of factors affecting intra-retinal cyst segmentation from OCT B-scans.
2. Automatic segmentation of intra-retinal cystic fluid spaces from OCT B-scans.

1.6 Major Contributions

A comparative assessment of existing automated intra-retinal cyst segmentation methods on OCT B-scans is conducted to identify the factors affecting IRC segmentation. For benchmarking purposes, a modular approach is employed for standardizing the different segmentation algorithms. Further, we analyze the variations in automated cyst segmentation performances and method scalability across two image acquisition systems (Cirrus and Spectralis) using the publicly available OPTIMA cyst segmentation challenge (OCSC) dataset.

An automatic unsupervised vendor-dependent IRC segmentation method using marker-controlled watershed transform on OCT B-scans is proposed. This method is based on two stages – k-means clustering technique to identify cysts in the form of markers and topographical based watershed transform for final segmentation. Qualitative and quantitative evaluation of proposed method is carried out against ground truth obtained from two graders on OCSC dataset.

Methods proposed for IRC segmentation on vendor-specific OCT images lacks generalizability across imaging systems. To address this limitation, a fully convolutional network (FCN) model for vendor-independent IRC segmentation is proposed. This method counteracts image noise variability and trains FCN models on OCT sub-images from the OCSC dataset (with four different vendor-specific images, namely, Cirrus, Nidek, Spectralis, and Topcon). Further, data augmentation and model hyper-parametrization is performed to prevent model over-fitting.

Deeper FCNs are computationally intensive due to regular convolutional filters which

results in large number of computation parameters. These Deeper FCNs are prone to vanishing gradient problem. Hence, to reduce the number of computation parameters a depth-wise separable convolutional neural network architecture is proposed. To prevent vanishing gradient problem Swish activation function is employed in the proposed architecture. OCSC dataset with four different vendor scans are used to evaluate the proposed architecture.

1.7 Organization of this Thesis

This thesis is organized as follows:

Chapter 2 summarizes the survey of automated methods developed for segmentation of IRCs and presents a comparative study on different IRC segmentation techniques to provide insights on factors affecting IRC segmentation from OCT images.

Chapter 3 presents a vendor dependent unsupervised IRC segmentation technique based on marker-controlled watershed transformation with both qualitative and quantitative evaluation.

Chapter 4 presents a fully convolutional neural network architecture for vendor independent IRC segmentation. Model based denoising technique is also proposed to enhance the quality of the OCT scans prior to segmentation. Hyper-parametrization and domain specific data augmentation techniques are employed to build customized fully conventional neural network architecture for IRC segmentation task.

Chapter 5 presents a depthwise separable convolutional neural network architecture for vendor independent IRC segmentation. Various analysis are conducted in this chapter to overcome the limitations of regular fully convolutional neural network architectures. Finally, **Chapter 6** concludes the thesis and provides future work directions.

CHAPTER 2

COMPARATIVE STUDY OF AUTOMATED INTRA-RETINAL CYST SEGMENTATION TECHNIQUES USING OCT B-SCANS

In the recent years, automated segmentation of intra-retinal cysts using optical coherence tomography B-scans has gained significant importance in the field of retinal image analysis. Automated segmentation of cystic spaces in retinal scans is a challenging task due to noise, scan intensity variations, varied cyst morphology, confounding retinal structures and pathologies. The objective of this Chapter[1] is to compare different intra-retinal cyst segmentation algorithms for comparative analysis and benchmarking purposes.

In this Chapter, a modular approach is employed for standardizing the different IRC segmentation algorithms. Further, analysis of the variations in automated cyst segmentation performances and method scalability across image acquisition systems is conducted using the publicly available cyst segmentation challenge dataset (OPTIMA cyst segmentation challenge (OPTIMA, 2015)). Several key automated methods are comparatively analyzed using quantitative and qualitative experiments. Our analysis demonstrates the significance of variations in SNR, retinal layer morphology and post-processing steps on the automated cyst segmentation processes. Finally, this Chapter provides insights towards the scalability of automated methods across vendor-specific imaging modalities to provide guidance for retinal pathology diagnostics and treatment processes.

[1]The work described in this Chapter has been published as: **G. N. Girish**, V. A. Anima , Abhishek R. Kothari, P. V. Sudeep, Sohini Roychowdhury and Jeny Rajan, "A Benchmark Study of Automated Intra-retinal Cyst Segmentation Algorithms using Optical Coherence Tomography B-Scans". *Computer Methods and Programs in Biomedicine*, Volume 153, pages 105-114, ISSN 0169-2607, Elsevier, 2018.

2.1 Brief Review of Automated IRC Segmentation Methods

The existing OCT intra-retinal cyst segmentation methods can be classified into two categories: semi-automated and fully automated. The semi-automated methods require manual intervention to define initial markers for each cyst. These methods are time and manual labor intensive owing to the large numbers of frames that need to be manually examined to define the markers (Fernandez, 2005; Zheng *et al.*, 2013). Thus, fully automated segmentation methods were developed to overcome these limitations. Figure. 2.1 highlights the categorization of existing automated methods for intra-retinal cyst segmentation. The automated methods are initially categorized as 2D or 3D segmentation techniques followed by further classification into unsupervised (Wilkins *et al.*, 2012; Wieclawek, 2015; Rashno *et al.*, 2017; Chen *et al.*, 2012) and supervised approaches (Quellec *et al.*, 2010; Gonzalez *et al.*, 2013; Swingle *et al.*, 2014, 2015; Chiu *et al.*, 2015; Lang *et al.*, 2015; Zhang *et al.*, 2015; Schlegl *et al.*, 2015; Lee *et al.*, 2017; de Moura *et al.*, 2017).

Wilkins *et al.* (2012) automatically segmented cysts from OCT B-scans using empirically obtained intensity threshold. Segmenting cysts with empirical threshold makes it difficult to generalize this method for clinical applications. Post-processing steps employed in their algorithm were not sufficient, which in turn affected sensitivity and specificity adversely.

Gonzalez *et al.* (2013) used watershed transform followed by region merging and texture analysis for detecting retinal cysts. The result of watershed segmentation contained over-flooded regions due to improper regional minima. Hence, these regions were combined into larger catchment basins by region merging approach using intensity profiles of each region. However, after region merging, there were large number of sets of regions that did not correspond to cystic fluids. Gray level co-occurrence matrix (GLCM) and Gabor filter based texture features were computed from the remaining

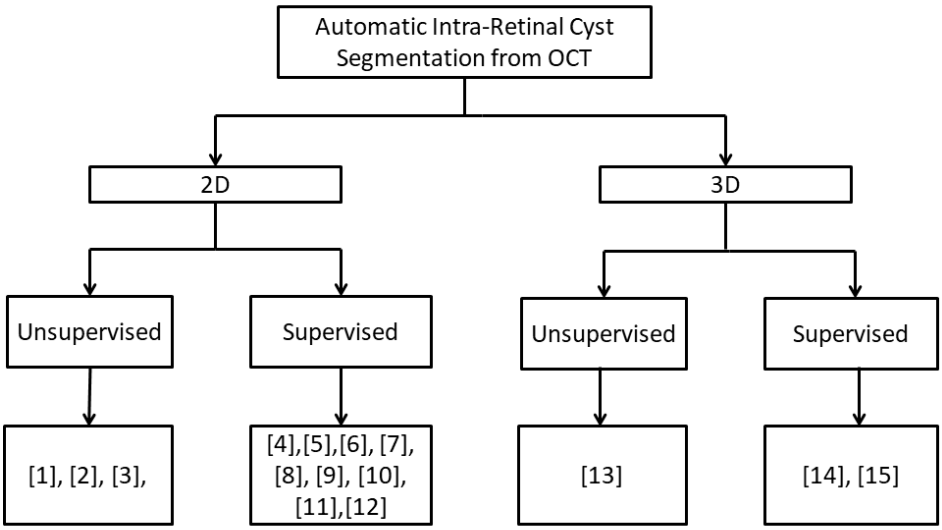


Figure 2.1: Taxonomy of segmentation approaches used for automated intra-retinal cyst segmentation.
[Ref: unsupervised [1, 2, 3, 13]-(Wilkins *et al.*, 2012; Wieclawek, 2015; Rashno *et al.*, 2017; Chen *et al.*, 2012) and supervised [4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15]-(Pilch *et al.*, 2013; Gonzalez *et al.*, 2013; Swingle *et al.*, 2014, 2015; Chiu *et al.*, 2015; Lang *et al.*, 2015; Schlegl *et al.*, 2015; Lee *et al.*, 2017; de Moura *et al.*, 2017; Zhang *et al.*, 2015; Quellec *et al.*, 2010)]

regions. Finally, these regions were classified using the random forest (RF) classifier (Breiman, 2001), support vector machine (SVM) (Hearst *et al.*, 1998) and naive Bayes (NB) classifier (Ng and Jordan, 2002) into cystic and non-cystic spaces based on extracted features.

k-means clustering and k-nearest neighbor (k-NN) techniques were employed by Pilch *et al.* (Pilch *et al.*, 2013). k-means clustering (Hartigan and Wong, 1979) was used to segment cystic structures, and subsequently Hu-moment (Hu, 1962) features were extracted from each segmented region. Finally, segmented regions were classified using k-NN classifier (Cover and Hart, 1967).

Swingle *et al.* (2014, 2015) proposed automated pseudo- and micro-cysts detection

algorithm and also identified their spatial distribution in primate retina. This algorithm was based on pixel-wise classification approach. Lang *et al.* (Lang *et al.*, 2015) used both intensity and spatial features with RF classifier for classifying pseudo-cyst pixels from OCT B-scans. Pixel-wise classification approaches suffer from presence of speckle noise and shadow artifacts which are always present in OCT scans.

Chiu *et al.* (2015) proposed an automatic retinal layer and fluid segmentation approach based on kernel regression (KR), graph theory and dynamic programming (GTDP) framework. Initially, a KR-based classification was used to estimate fluid and retinal layer positions. Then these classification results were used for finer segmentation using GTDP. This algorithm was evaluated on 110 B-scans from 10 patients with severe diabetic macular edema pathology. This method produced better segmentation correlation coefficient but was computationally expensive compared to normal GTDP algorithm.

Wieclawek (2015) proposed an approach utilizing different image processing techniques such as non-linear complex diffusion and mathematical morphology operations. In particular, h-minima transform and image reconstruction by erosion and geodesic operations were used to segment the cysts from OCT images. The obtained results were qualitatively reported.

Schlegl *et al.* (2015) used convolution neural network (CNN) to segment the intraretinal fluid (IRF) and sub-retinal fluid (SRF) from retinal OCT scans. The CNN was trained with approximately 300,000 2D image patches which were extracted from 157 OCT image volumes available at the Vienna Reading Center. The trained CNN model achieved an overall accuracy of 96% across three classes (Normal retinal tissue, IRF and SRF) using pixel-wise classification approach.

Similar patch based CNN approach was proposed by (Lee *et al.*, 2017) to segment IRF from 1289 OCT images. However the model was trained on OCT images obtained from single vendor OCT machine and could not be generalized.

de Moura *et al.* (2017) proposed intensity and texture based supervised method for

the identification of IRCs. A heterogeneous set of 326 intensity and texture based features was extracted from corresponding IRC and normal patches in OCT images. Then, Relief-F and L0 feature measures were used to identify optimal feature subset. Finally, linear Bayes normal classifier, quadratic Bayes normal classifier and k-NN classifier were used for the discrimination of IRC and normal patches. The method was evaluated on 363 and 360 patches of IRC and normal patches extracted from 51 OCT images. The LDC classifier with L0 feature selector provided better accuracy (0.91) than Relief-F and other classifiers.

However, in patch-based segmentation methods, the selection of an appropriate patch size has a considerable impact on segmentation results. Selection of small patches leads to inaccurate segmentation of bigger cysts whereas larger patches result in segmentation of non-retinal vascular regions as cystic structures.

Very recently, a neutrosophic sets and graph-cut algorithms based approach was proposed by Rashno (Rashno *et al.*, 2017) to segment IRCs from OCT images. Initially the images were transformed into neutrosophic domain and candidate cyst regions were segmented using fuzzy based clustering technique. Finally, fine segmentation was performed on the clustered candidate regions using graph-cut technique. The method was evaluated on three different datasets (UMN, OPTIMA and Duke) obtained using Spectralis vendor OCT scans. Even though the method was unsupervised, lot of parameters had to be manually identified for fuzzy clustering objective function. Choosing appropriate number of clusters and initial coarse segmentation affects the performance of the method.

Automated 3D volumetric segmentation of retinal cysts has been attempted (Quellec *et al.*, 2010; Chen *et al.*, 2012; Zhang *et al.*, 2015). 3D volumetric analysis of retinal cysts from OCT requires densely sampled data with a higher number of C-scans, otherwise small cystic structures can go undetected.

The key challenges posed by the OCT images to most existing methods for automated segmentation of cystic regions include pixel-level variabilities due to noise,

image intensity variations, varied cyst morphology, confounding retinal structures and complex pathologies.

Based on the existing methods for automated cyst segmentation from OCT images, a generic methodological framework is proposed in Figure. 2.2. This framework consists of four main steps: (1) pre-processing; (2) retinal layer segmentation; (3) cyst segmentation; and (4) post-processing. Since OCT images contain varying degrees of additive speckle noise, a pre-processing module is required for quality enhancement and equalization of the OCT images. Finally, the post-processing step is implemented to reduce the incorrectly segmented non-cystic regions (i.e., false positive regions). In this Chapter, existing automated retinal cyst segmentation methods are standardized based on the work-flow shown in Figure. 2.2 and comparatively analyzed to evaluate the significance of the automated methods with respect to input data and output metrics.

This Chapter makes three key contributions. First, a modular approach to standardize existing OCT cyst segmentation methods is presented for methodological benchmarking purposes. The methodological contributions from significant automated OCT cyst segmentation methods are reviewed and comparatively discussed. Secondly, quantitative and qualitative analysis through experiments are presented for evaluation of the existing automated OCT cyst segmentation methods. It is observed from the results that supervised OCT segmentation methods achieve higher cyst segmentation recall when compared to unsupervised approaches with degradation in segmentation precision across data sets with variable scan qualities. Lastly, OCT images from two different image acquisition systems (Cirrus and Spectralis) are comparatively analyzed for scalability limitations owing to the image-level variabilities introduced by imaging systems. This Chapter also provide novel insights into the limitations of automated cyst

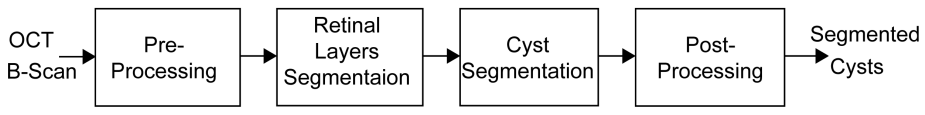


Figure 2.2: Generic framework of automated intra-retinal cyst segmentation system.

Table 2.1: OPTIMA cyst segmentation challenge dataset description

Vendor	Train set	# B-Scans	Test set	# B-Scans
Cirrus	4	584	4	512
Nidek	3	384	3	15
Spectralis	4	196	4	112
Topcon	4	512	4	270
Total	15	1676	15	909

segmentation tasks for retinal diagnostic and screening purposes.

2.2 Materials

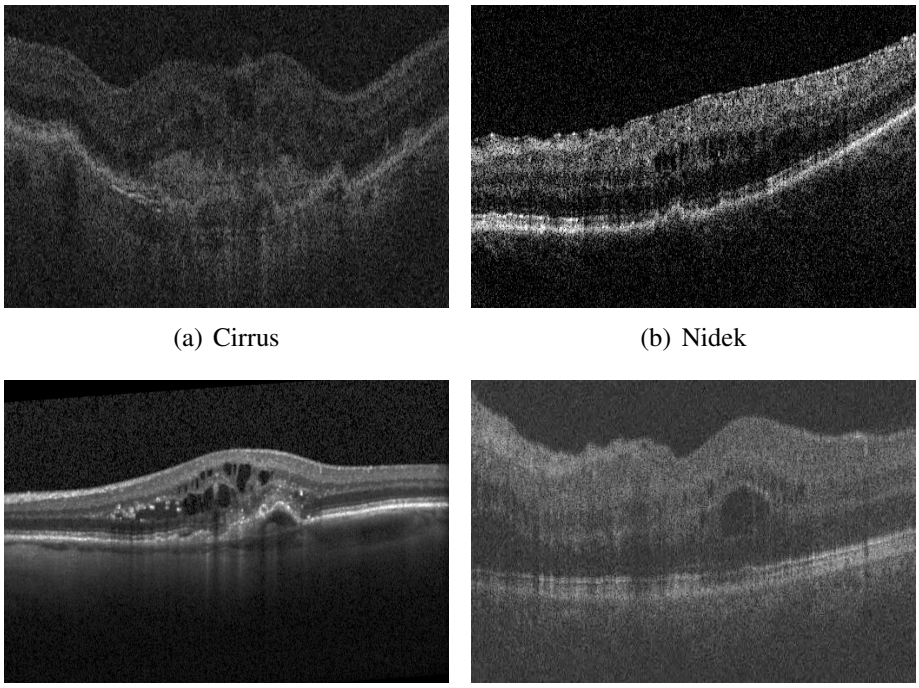
2.2.1 Dataset

Comparative analysis of the existing automated intra-retinal OCT cyst segmentation methods is performed using publicly available OPTIMA cyst segmentation challenge dataset (OPTIMA, 2015). OCSC dataset contains 30 SD-OCT scans with 2585 B-scans with random presence of IRC obtained from CME subjects using four different imaging systems, namely Zeiss Cirrus, Nidek, Spectralis Heidelberg and Topcon. Ground truth (GT) for all scans are annotated by two expert ophthalmologists.

The selected OCT scans are acquired over 6×6 mm of the macula and foveal center from subjects with CME. The OCT frames (B-scans) represent gray-scale images with resolution of $[496 \times 512]$ and $[496 \times 1026]$ pixels, respectively. The description of data set is presented in Table 2.1 and a sample OCT B-scan frame from each vendor is shown in Figure. 2.3. This dataset can be obtained from Christian Doppler Laboratory for Ophthalmic Image Analysis, Department of Ophthalmology, Medical University of Vienna[2].

In this comparative study, OCT scans from Cirrus and Spectralis image acquisition

[2]optima@meduniwien.ac.at



(a) Cirrus (b) Nidek

(c) Spectralis (d) Topcon

Figure 2.3: Sample retinal OCT B-scans from four different vendors of OCSC dataset. (Image courtesy: OPTIMA cyst challenge dataset (OPTIMA, 2015)).

systems are used to evaluate the methods, since the data sets from these systems demonstrate moderate to severe pathological features when compared to the other imaging systems.

2.2.2 Performance Metrics

For quantitative analysis of the automated OCT cyst segmentation methods, the prediction of cystic regions from the automated methods are compared with the manually annotated ground truth region images, which are provided for the input data sets from two independent trained ophthalmologists. In this work, we perform segmentation of cystic regions followed by the evaluation of the number of true Positive (TP), false positive (FP) and false negative (FN) regions. Here, TPs are defined as actual cystic regions that are automatically detected, whereas the non-cystic regions detected as cysts

by an automated algorithm are FPs, and the actual cystic regions that are undetected by an automated algorithm are FNs. The output metrics for evaluation of the automated segmentation algorithms are defined as precision (2.1) and recall (2.2).

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \tag{2.1}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \tag{2.2}$$

As an additional measure for methodological evaluation, the Dice coefficient (DC) (Dice, 1945) (2.3) analyzes the correlation of automated segmentation (AS) vs. GT per patient image stack.

$$\text{DC} = 2 \cdot \frac{|\text{AS} \cap \text{GT}|}{|\text{AS}| + |\text{GT}|} \tag{2.3}$$

2.3 Methods

2D intra-retinal cyst segmentation methods published till March-2015 are selected for performing the comparative study. We have chosen 2D methods instead of 3D segmentation methods because of the spatial resolution between scans is too high and which can results miss out small cystic structures. The methodological contributions of automated OCT cyst segmentation techniques considered for this comparative benchmark study are summarized below.

Wilkins *et al.* (2012) proposed an automated cyst segmentation method using empirically obtained global intensity threshold. For reducing the speckle noise, the authors applied a fast bilateral filter on the SNR balanced and normalized OCT B-scans. In addition, the top-bottom search mechanism with the intensity threshold was employed to identify the RPE-nerve fiber layer (NFL) boundaries. Next, regional thresholding was used for the rejection of FPs. In (Wilkins *et al.*, 2012), the method was evaluated using 19 OCT-volumes (16 captured from eyes of patients with vitreo-retinal disease and 3

control eyes) and the mean sensitivity and specificity were reported as 0.91 and 0.96, respectively. However, cyst segmentation using empirical thresholding is not feasible in a clinical setup and thus, the post-processing steps mentioned in (Wilkins *et al.*, 2012) are insufficient for generalizable removal of FPs.

Gonzalez *et al.* (2013) used Watershed (WS) transformation for detecting cysts followed by region merging and texture analysis. In (Gonzalez *et al.*, 2013), the retinal layers were initially segmented using a graph cut segmentation approach. After delimiting the region of interest (ROI) between ILM and RPE layers, WS was applied to segment the cyst regions. After WS segmentation, the authors combined the over-flooded regions to form larger catchment basins by region merging approach and performed the post-processing steps using 14 different parameters for each region.

The detected cysts were then classified with texture descriptors based on GLCM and Gabor filters and by using classifiers such as naive-Bayes, support vector machine or random forest with 10-fold cross validation. The comparative evaluation in (Gonzalez *et al.*, 2013) indicates that GLCM+SVM and Gabor+SVM have segmentation accuracies of 0.8293 and 0.8244, respectively. The limitations of this method include over-flooding of WS due to improper regional minima-estimation and the large number of parameters while post-processing. Hence, this method suffers from generalizability in a clinical setup.

In (Pilch *et al.*, 2013), Pilch *et al.* segmented the cysts with the k-means clustering technique and classified them using k-nearest neighbor. Beforehand, the speckle noise in B-scans was removed with Bayesian estimation method, followed by implementation of an automated active contour model for retinal layer segmentation. The experiments discussed in (Pilch *et al.*, 2013), using eight OCT B-scans with 130 cysts, indicate that the retinal structures are over-segmented with k-means clustering method if k is set too low and improper delineation of segmented retinal micro-structures occurs when k is set too high. Segmented regions were classified using Hu-moment features and k-NN classifier.

Swingle *et al.* (2014, 2015) proposed an automated pseudo-cyst detection algorithm, based on pixel-wise classification approach, and identified the spatial distribution of cysts in the retinal fundus images. Prior to classification, B-scans were normalized and ILM and Bruchs membrane (BM) retinal boundaries were segmented. The authors trained an RF classifier with 14 features extracted from manually segmented data. Next, intensity profiles and image-based features were also estimated for the test data set. Finally, RF-based classification is performed on the test images, where each test-image pixel was probed for being cyst-like. Due to the presence of spurious pixels in the classification output, post-processing was performed to remove all connected components below the threshold of 30 pixels. In (Swingle *et al.*, 2014), the proposed method was evaluated on 49 OCT B-scans obtained from five subjects with Micro-cystic macular edema (MME) and the results demonstrated automated classification accuracy of 0.846 and DC of 0.75 obtained against two manual graders. However, this pixel-wise classification approach for cysts is limited in its performance by the presence of speckle noise and shadow artifacts in OCT images.

Lang *et al.* (2015) used both intensity and spatial features with an RF classifier for classifying pseudo-cyst pixels from OCT B-scans. Initially B-scans were normalized, followed by extraction of eighteen features, of which sixteen were intensity based and two were spatial position based features. The system was trained on the features extracted from manually segmented data. After extracting features for test data, the RF classifier was used with 60 trees, each with minimum terminal leaf size of 10. Experimental analysis was performed using leave-one-out approach. Further, this method was assessed on images obtained from 10 MME patients and 10 control candidates. For MME data, precision of 0.85, recall of 0.79 and DC of 0.98 was reported with this algorithm. In case of control data, this method correctly labeled all the pixels.

Wieclawek's segmentation method involved different image processing techniques including non-linear complex diffusion and mathematical morphology operations (Wieclawek, 2015). The authors segmented the cysts from OCT B-scans using h-minima transform

and image reconstruction by erosion and geodesic operations. The segmented regions above and below the NFL and RPE layers and the regions connected to image boundaries were removed for reducing FPs. However, in this work, the performance of the cyst segmentation method were qualitatively reported.

2.4 Experimental Setup

Experiments analysis for automated segmentation of cysts using the methods described above is performed on OCT images from Spectralis and Cirrus imaging systems, and the segmentation outcomes are comparatively evaluated. For standardization purposes, the pre-processing module for the removal of additive speckle noise from the OCT B-scans precedes the retinal layer segmentation step. For each method described in section 2.3, the automated segmentation algorithm module is followed by post-processing modules. The segmentation methods are implemented in MATLAB on a standalone computer with following specifications: 64-bit operating system, Core i7 3700 CPU and RAM size 8 GB. The functional modules for the automated OCT segmentation methods are described below.

2.4.1 Pre-processing and Retinal Layers Segmentation

Since the segmentation methods are sensitive to noise, all the input OCT images are denoised first. Recently proposed adaptive and unbiased bilateral (AUB) filter (Sudeep *et al.*, 2016*a*) is employed for denoising OCT scans. In (Sudeep *et al.*, 2016*a*), a three parameter Gamma distribution function is used to fit the noisy (observed) OCT B-scans and a maximum likelihood (ML) approach is used to estimate the parameters of the Gamma distribution, based on which, the *bias* $\hat{B}_e$ parameter is computed. Next, $\hat{B}_e$ is subtracted from the output of an adaptive version of the conventional bilateral filter (Tomasi and Manduchi, 1998). The various parameters associated with the AUB filter

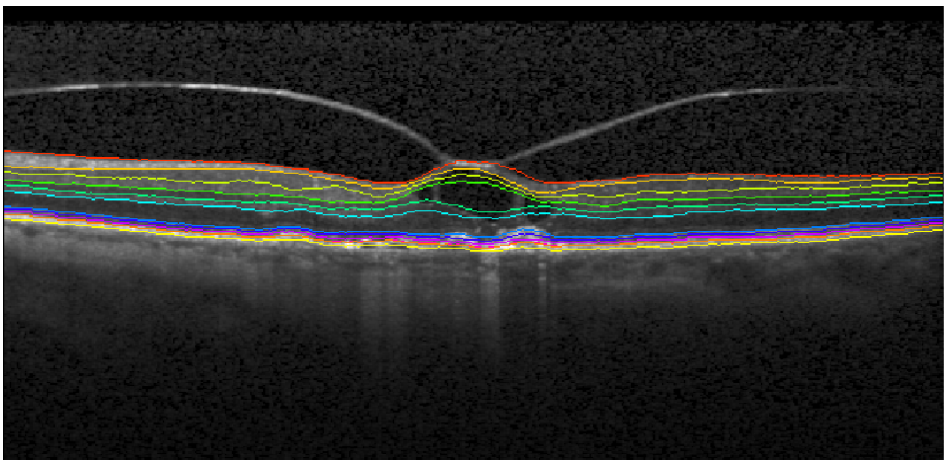


Figure 2.4: OCT image of retina to visualize the order and position of the boundaries of layers. (From top: 1- internal limiting membrane (ILM); 2- retinal nerve fiber layer boundary (RNFL); 3- ganglion cell layer (GCL); 4- inner plexiform layer (IPL); 5 - outer plexiform layer (OPL); 6- outer nuclear layer (ONL); 7- Boundary of myoid and ellipsoid of inner segments (BMEIS); 8- Inner Segment/Outer Segment (IS/OS) junction; 9- Outer Photoreceptor (OPR); 10- Retinal Pigment Epithelium (RPE); 11- Bruch's Membrane Complex (BMC). (The Iowa Reference Algorithms (Retinal Image Analysis Lab, Iowa Institute for Biomedical Imaging, Iowa City, IA) (Li *et al.*, 2006; Garvin *et al.*, 2009; Abramoff *et al.*, 2010))

are as follows: Half width of the filter = 5, geometrical parameter $\hat{\sigma}_d$=1.5, photometric spread $\hat{\sigma}_r$=2.5$\hat{\sigma}$ and 0.6$\hat{\sigma}$ for Cirrus and Spectralis respectively. Where $\hat{\sigma^2}$ denotes the estimated noise variance from the given input image. The estimated bias $\hat{B}_e$ is empirically estimated as 32 and 15 for the Spectralis and Cirrus scans, respectively.

For cyst segmentation, retinal layer segmentation is an important pre-processing step because it provides information regarding variations in morphology and thickness of retinal layers, which leads to refine the cyst segmentation results. Among different automated retinal layer segmentation methods (Li *et al.*, 2006; Mayer *et al.*, 2010; Lang *et al.*, 2013), the Iowa reference algorithm (Retinal Image Analysis Lab, Iowa Institute for Biomedical Imaging, Iowa City, IA) (Li *et al.*, 2006; Garvin *et al.*, 2009; Abramoff *et al.*, 2010) is used in this Chapter to obtain the eleven segmented retinal layers as shown in Figure. 2.4.

2.4.2 Cyst Segmentation and Post-processing

The cyst segmentation methods described in section 2.3 are implemented with method specific post-processing steps as follows. For (Wilkins *et al.*, 2012), cyst segmentation method is implemented with empirically computed gray level threshold of 49 and 37 for Spectralis and Cirrus scans, respectively. Post-processing is achieved with the gray level intensity standard deviation of 41, by removing regions less than 10 pixels and regions outside ROI (NFL-RPE). This step reduces FPs.

For Gonzalez method (Gonzalez *et al.*, 2013), post processing is performed after WS segmentation as follows: elongated cysts $(t_1 = 1.5; t_2 = 2; p_{occ} = 0.6)$, small cysts (regions lesser than $n_{min} = 10$ pixels), cysts next to image slides, cysts above ILM and below OPL are removed to reduce FPs followed by classification (Gonzalez *et al.*, 2013).

In the implementation of k-means clustering and k-NN method (Pilch *et al.*, 2013), pre-processed B-scans are automatically segmented using k-means algorithm. Optimal value of k is empirically determined as 3. Next, the gray-level pixel intensities of B-scans are clustered as hyper reflective layer, hypo-reflective layer and cystic or low intensity regions, respectively. Finally, the clusters are post-processed to remove outliers using boundary layers of retina (ILM and RPE), pseudo-cyst ratio (lesser than 10 pixels) and width to height ratio (width is 4 times greater than height). After this post-processing step, texture features are extracted using Hu-moments and k-NN classification ($k = 6$) is performed with manually segmented training data from 4 OCT volumes.

Swingle *et al.* (2014) is implemented to extract 14 intensity features on manually segmented data from the 4 OCT volumes as follows: 2 features are obtained from the the voxel intensity before and after gray-scale morphological closing operation and 9 features are retrieved after Gaussian filtering at various scales. Next, 2 features are extracted by Laplacian and Laplacian of Gaussian (LoG) of the image and a final feature

is the relative distance of intra-retinal cyst to the retinal boundary (ILM-Boundary of myoid and ellipsoid of inner segments (BMEIS)). Finally, an RF classifier is trained from the features with 60 trees and the test data is probed. Here, all connected components lesser than 30 pixels are ignored to remove spurious pixels in the post-processing step.

The method by Lang *et al.* (Lang *et al.*, 2015) is implemented with 15 features that are extracted from the OCT images, followed by the aforementioned steps for training RF classifier with parameters mentioned in (Lang *et al.*, 2015). Next, the test data is probed to classify cyst and non-cyst pixels. To generate final segmentation results, two-stage thresholding approach is employed with probability threshold of 0.5 and 0.85, respectively. Finally, all connected components lesser than 5 pixels are ignored to remove spurious pixels.

For the method by Wieclawek (Wieclawek, 2015), value of $h = 3$ and $h = 4$ are used to compute h-minima transform on Spectralis and Cirrus scans, respectively. Next, h-minima transformed B-scans are thresholded to obtain the segmented results. The segmented regions above and below NFL and RPE layers, and regions connected to image boundaries are removed to reject FPs .

2.5 Results and Analysis

The quantitative and qualitative comparative assessment of the automated cyst segmentation methods is presented below. Since low precision is indicative of FPs (over-detection) and low recall is indicative of missing patients with abnormalities (under-detection), high precision and high recall values are desired for an ideal automated cyst segmentation method.

2.5.1 Quantitative Assessment

Table 2.2 and Table 2.3 demonstrate precision and recall for the methods under analysis against grader 1 (G_1), grader 2 (G_2) and intersection of both the graders ($G_1 \cap G_2$) on the Spectralis and Cirrus scans, respectively. Here, observe that cyst segmentation results obtained with mathematical morphology based approach proposed by Wieclawek, achieve higher recall across input dataset variabilities (i.e., mean recall of 0.38 and 0.59 on Spectralis and Cirrus scans, respectively, with G_1) as compared to other unsupervised method (Wilkins *et al.*, 2012). However, segmentation precision varies largely with the image quality (i.e., mean precision of 0.35 and 0.09 on Spectralis and Cirrus scans, respectively, with G_1).

Among the supervised cyst segmentation methods, Swingle *et al.* method (Swingle *et al.*, 2014) outperforms other supervised methods in terms of recall while preserving marginal precision rates across dataset variabilities. Also, trends in degradation of precision with respect to image quality, similar to that of the unsupervised methods, is observed. Thus, from Table 2.2 and Table 2.3, it can be inferred that supervised methods outperform unsupervised methods in terms of overall precision and recall. Also, texture-based supervised methods such as GLCM, Gabor and Hu-moments are found to be efficient discriminators for cystic regions. Additionally, from Table 2.2 and Table 2.3, also observe that the variations in performances with respect to manual graders due to inter-grader variability (OPTIMA, 2015).

Further, comparative assessment of cyst segmentation methods on a sample Spectralis OCT B-scan are presented in Figure. 2.5. The original B-scan and GT provided by G_1 (see the marked region in red colour) are shown in Figures. 2.5(a) and 2.5(b), respectively. It can be observed from Figure. 2.5(c) that Wilkins method has failed to identify some cysts due to intensity variations, and it has also failed to delineate the cysts up to their actual boundary. In Figure. 2.5(d), the Wieclawek method shows over-segmentation by considering consecutive cysts as a single entity in an image. Here, the boundaries between cysts are missed and FPs in the ONL layer are observed.

It is observed from Figure. 2.5(e) that the Gonzalez method introduces an additional cyst (on the left side of the retina) and fails to segment consecutive cysts due to improper regional minima-estimate and over-flooding. Although the Pilch method is able to detect both micro and macro-cysts, it fails to delineate the cysts to their actual boundary as shown in Figure. 2.5(f)). In case of Swingle method and Lang method, all the identified consecutive cysts are merged into large cysts (see Figures. 2.5(g) and 2.5(h), respectively) due to the lack of boundary discrimination around each individual cyst.

Besides precision and recall, DC is also evaluated to validate the experimental results and plotted in Figures. 2.6 and 2.7 for Spectralis and Cirrus scans, respectively. In these box plots, the line inside the box indicates the median value and as a rule of thumb, any segmentation method performs well if the corresponding median in the box plot is high. Here, the range of the box gives the spread of the DC-values that are computed from different patient image volumes. It can be noticed from Figure. 2.6 and Figure. 2.7 that the Lang method has the highest median values for both Spectralis and Cirrus scans, respectively.

It is observed that automated sub-retinal layer segmentation techniques are prone to erroneous segmentation on pathological B-scans, when compared to normal subjects. For this analysis, an evaluation on IRC segmentation is conducted using automated retinal segmentation method against manual retinal layer ground-truth provided by the ophthalmologist on a Spectralis vendor test volume with 7 pathological OCT B-scans. Ophthalmologist manually segmented 6 retinal layer boundaries (ILM, RNFL, OPL, ONL, BMEIS and RPE) using GIMP (GNU Image Manipulation Program) software. A tablet computer with stylus is used for freeform (drawing) marking the boundaries as if writing on a flat surface. Grader also has the option to erase/add/correct markings. The grader made the best estimate for marking the boundaries of layers, which were completely obscured or lost due to CME.

Obtained mean Dice correlation results of the cyst segmentation methods using both automated and manual layer segmentation is reported in Table 2.4. Reported

Table 2.2: Mean (standard deviation) of precision and recall of compared methods on Spectralis OCT scans.

Method	G_1		G_2		$G_1 \cap G_2$	
	Precision	Recall	Precision	Recall	Precision	Recall
Wilkins *et. al* (Wilkins *et al.*, 2012)*	0.2415 (0.2573)	0.4584 (0.3073)	0.2429 (0.2607)	0.4878 (0.3477)	0.2423 (0.2580)	0.4620 (0.3259)
Wieclawek (Wieclawek, 2015)*	0.3578 (0.2210)	0.3863 (0.3341)	0.3700 (0.2337)	0.4282 (0.3533)	0.3807 (0.2483)	0.4204 (0.3383)
GLCM+NB (Gonzalez *et al.*, 2013)†	0.5354 (0.4403)	0.6563 (0.4619)	0.5373 (0.4417)	0.6667 (0.4714)	0.5463 (0.4333)	0.6458 (0.4533)
GLCM+SVM (Gonzalez *et al.*, 2013) †	0.5625 (0.4270)	0.4688 (0.3840)	0.5875 (0.4250)	0.5521 (0.4682)	0.5625 (0.4270)	0.4688 (0.3840)
GLCM+RF (Gonzalez *et al.*, 2013) †	0.5800 (0.4214)	0.5208 (0.4429)	0.3239 (0.3761)	0.4167 (0.5000)	0.2656 (0.3078)	0.2604 (0.3234)
Gabor+NB (Gonzalez *et al.*, 2013) †	0.4000 (0.3031)	0.3750 (0.2846)	0.3997 (0.2986)	0.3646 (0.2792)	0.3889 (0.2940)	0.3542 (0.2753)
Gabor+SVM (Gonzalez *et al.*, 2013) †	0.4167 (0.2887)	0.3542 (0.2753)	0.3505 (0.2865)	0.3229 (0.2554)	0.3824 (0.2893)	0.3437 (0.2730)
Gabor+RF (Gonzalez *et al.*, 2013) †	0.2667 (0.3485)	0.1979 (0.2342)	0.2281 (0.2818)	0.1979 (0.2342)	0.0833 (0.1667)	0.0104 (0.0208)
Pilch *et. al* (Pilch *et al.*, 2013) †	0.4067 (0.4611)	0.5833 (0.4410)	0.4125 (0.4686)	0.6250 (0.4787)	0.4067 (0.4611)	0.5833 (0.4410)
Swingle *et. al* (Swingle *et al.*, 2014) †	0.3500 (0.2428)	0.8270 (0.2617)	0.3353 (0.2537)	0.8140 (0.2852)	0.3493 (0.2677)	0.6994 (0.3012)
Lang *et. al* (Lang *et al.*, 2015)†	0.8000 (0.1897)	0.5388 (0.3403)	0.7355 (0.2742)	0.5053 (0.3519)	0.7880 (0.2114)	0.4816 (0.3134)

† Supervised Methods.
* Unsupervised Methods

Table 2.3: Mean (standard deviation) of precision and recall of compared methods on Cirrus OCT scans.

Method	G_1		G_2		$G_1 \cap G_2$	
	Precision	Recall	Precision	Recall	Precision	Recall
Wilkins *et. al* (Wilkins *et al.*, 2012)*	0.1352 (0.1706)	0.1959 (0.3381)	0.1344 (0.1704)	0.2383 (0.4133)	0.1352 (0.1706)	0.2047 (0.3548)
Wieclawek (Wieclawek, 2015)*	0.0968 (0.0889)	0.5989 (0.4192)	0.0837 (0.0712)	0.6234 (0.4449)	0.0944 (0.0848)	0.6048 (0.4262)
GLCM+NB (Gonzalez *et al.*, 2013) †	0.3783 (0.2766)	0.3238 (0.2695)	0.3529 (0.2716)	0.2779 (0.2192)	0.3917 (0.2833)	0.2660 (0.2282)
GLCM+SVM (Gonzalez *et al.*, 2013) †	0.5618 (0.4484)	0.2772 (0.2296)	0.6610 (0.4552)	0.2209 (0.1928)	0.6610 (0.4552)	0.2209 (0.1928)
GLCM+RF (Gonzalez *et al.*, 2013) †	0.5101 (0.3821)	0.1475 (0.1556)	0.6643 (0.4508)	0.0838 (0.0899)	0.6875 (0.4732)	0.0619 (0.0633)
Gabor+NB (Gonzalez *et al.*, 2013) †	0.2202 (0.1381)	0.6150 (0.2656)	0.2143 (0.1390)	0.6466 (0.2107)	0.2122 (0.1378)	0.6693 (0.1920)
Gabor+SVM (Gonzalez *et al.*, 2013) †	0.1934 (0.1166)	0.5628 (0.3139)	0.1852 (0.1127)	0.6155 (0.2677)	0.1714 (0.1002)	0.6064 (0.2369)
Gabor+RF (Gonzalez *et al.*, 2013) †	0.1446 (0.1371)	0.6734 (0.0752)	0.1236 (0.2071)	0.6344 (0.0975)	0.1478 (0.1400)	0.6265 (0.1975)
Pilch *et. al* (Pilch *et al.*, 2013) †	0.0936 (0.0939)	0.7674 (0.2539)	0.0927 (0.0899)	0.7735 (0.2327)	0.0918 (0.0939)	0.7565 (0.2504)
Swingle *et. al* (Swingle *et al.*, 2014) †	0.1784 (0.1821)	0.9452 (0.0997)	0.1642 (0.1622)	0.9995 (0.1183)	0.1758 (0.1966)	0.9576 (0.0736)
Lang *et. al* (Lang *et al.*, 2015)†	0.6428 (0.4738)	0.0822 (0.0888)	0.5000 (0.5773)	0.0710 (0.0910)	0.6250 (0.4787)	0.0991 (0.1040)

† Supervised Methods.
* Unsupervised Methods

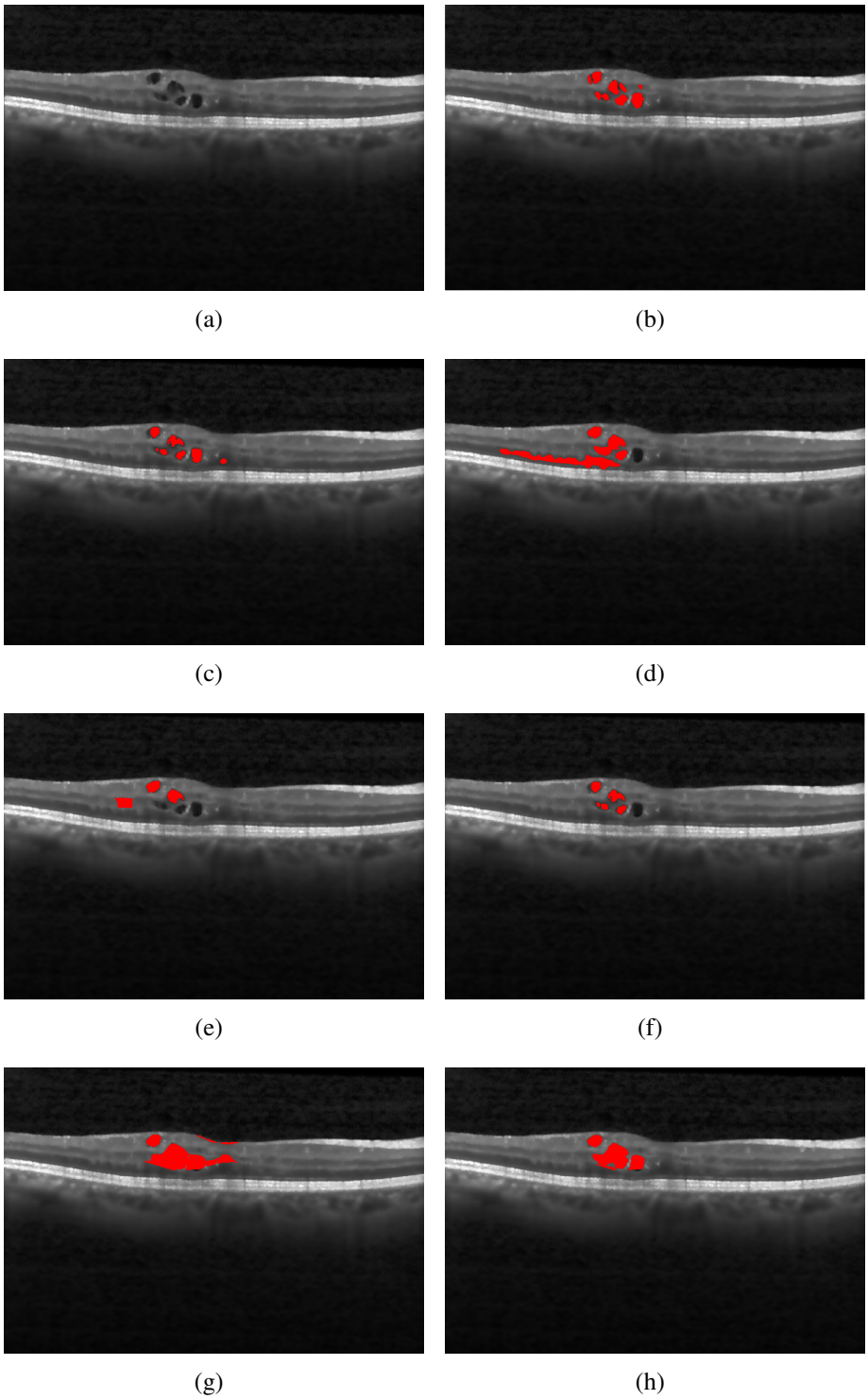


Figure 2.5: Results of different automated intra-retinal cyst segmentation methods on Spectralis scans against the GT from G_1 (a) Original B-scan, (b) GT (c) Wilkins Method (Wilkins *et al.*, 2012), (d) Wieclawek Method (Wieclawek, 2015), (e) Gonzalez Method (Gonzalez *et al.*, 2013), (f) Pilch Method (Pilch *et al.*, 2013), (g) Swingle Method (Swingle *et al.*, 2014), (h) Lang Method (Lang *et al.*, 2015).

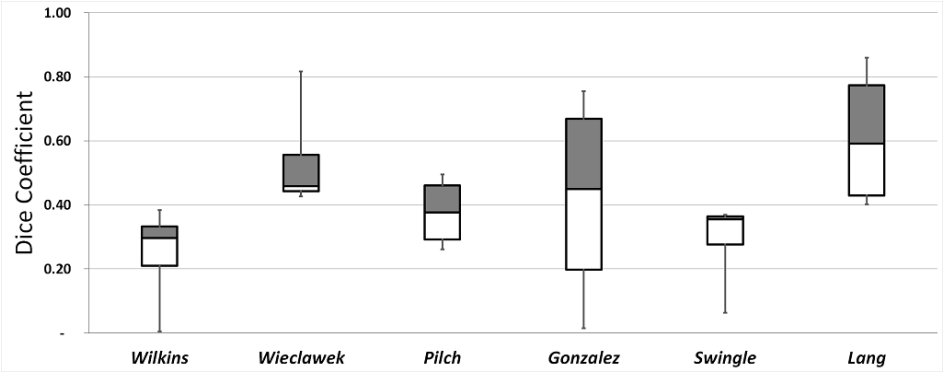


Figure 2.6: Whisker Box plot for the Dice correlation coefficient against intersection of both the graders on Spectralis dataset for Wilkins Method (Wilkins *et al.*, 2012) (0.30), Wieclawek Method (Wieclawek, 2015) (0.46), Pilch Method (Pilch *et al.*, 2013) (0.38), Gonzalez Method (Gonzalez *et al.*, 2013) (0.45), Swingle Method (Swingle *et al.*, 2014) (0.35), Lang Method (Lang *et al.*, 2015) (0.59),) .

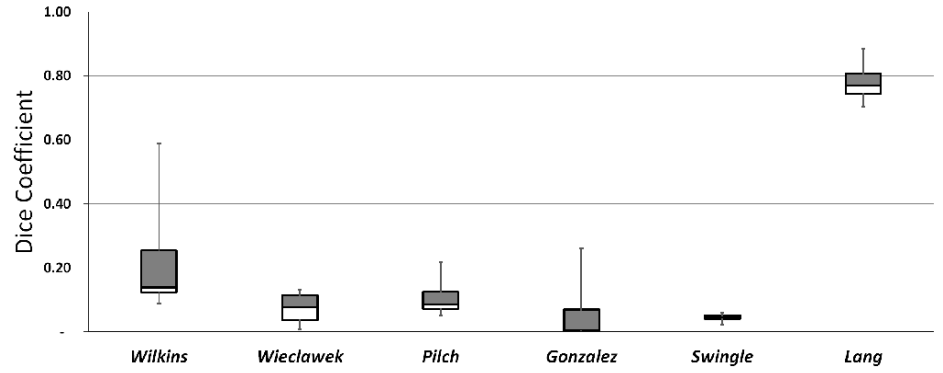


Figure 2.7: Whisker Box plot for the Dice correlation coefficient against intersection of both the graders on Cirrus dataset for Wilkins Method (Wilkins *et al.*, 2012) (0.14), Wieclawek Method (Wieclawek, 2015) (0.08),Pilch Method (Pilch *et al.*, 2013) (0.12), Gonzalez Method (Gonzalez *et al.*, 2013) (0.00), Swingle Method (Swingle *et al.*, 2014) (0.05), Lang Method (Lang *et al.*, 2015) (0.77),) .

results in Table 2.4 shows that methods proposed by Wilkins (Wilkins *et al.*, 2012), Weiclawek (Wieclawek, 2015) and Pilch (Pilch *et al.*, 2013) are robust to initial layer segmentation errors (minimal difference in the Dice rate between manual and automated layer segmentation), whereas, Gonzalez (Gonzalez *et al.*, 2013), Swingle (Swingle *et al.*, 2014) and Lang (Lang *et al.*, 2015) methods are more sensitive to layer segmentation errors (observe the significant variation in Dice rate obtained between manual and automated layer segmentation). Methods (Wilkins, Weiclawek and Pilch) utiliz-

Table 2.4: Mean Dice correlation of the different cyst segmentation methods using automated and manual layer segmentation on Spectralis vendor OCT scans using ground truth provided by G_1

Method	Manual Grader	Iowa reference algorithm
Wikkins (Wilkins *et al.*, 2012)	0.3702	0.3628
Weiclawek (Wieclawek, 2015)	0.4130	0.4058
Pilch (Pilch *et al.*, 2013)	0.3094	0.3060
Gonzalez (Gonzalez *et al.*, 2013)	0.2397	0.0150
Swingle (Swingle *et al.*, 2014)	0.4225	0.3437
Lang (Lang *et al.*, 2015)	0.5467	0.4487

ing RPE as the lower boundary of the ROI are robust to initial layer segmentation errors compared to methods (Gonzalez, Swingle and Lang) utilizing inner layer as ROI boundaries. In case of pathological cases it is difficult to accurately segment inner layers of the retina for both manual and automated algorithms.

In addition, an analysis on the accuracy of different cyst segmentation algorithms on segmenting cysts located in the inner sub-retinal layers is performed. Table 2.5 shows the mean recall rate of segmentation methods on different inner layers of the retina with Spectralis vendor OCT scans. It is observed that layer thresholding limiting in retina to select ROI affects the performance of the methods. The methods in Wilkins *et al.* (Wilkins *et al.*, 2012) and Weiclawek (Wieclawek, 2015) fails to segment the cysts between ILM-RNFL layers due to ROI of these methods restricted between (RNFL-RPE). Method proposed by Gonzalez (Gonzalez *et al.*, 2013) efficiently segments the cysts in all the inner layers of the retina, when compared to other methods. All the segmentation methods perform well in hypo-reflective layers compared to hyper-reflective layers (observe the recall rate in Table 2.5). Supervised methods proposed in Swingle *et al.* (Swingle *et al.*, 2014) and Lang *et al.* (Lang *et al.*, 2015) utilizes the intensity and texture based features, and therefore both these methods outperformed other methods in mean recall rate at respective inner layers.

Additionally, to evaluate the performance of segmentation methods on normal B-scans, experiments are conducted using first 20 normal B-scans of volume 2 in the

Table 2.5: Accuracy of cyst segmentation methods in different intra-retinal layers on Spectralis vendor OCT scans using ground truth provided by G_1.

Method	ILM-RNFL	RNFL-GCL	GCL-IPL	IPL-INL	INL-OPL	OPL-ONL
Wikkins (Wilkins *et al.*, 2012)	0	0.3333	0.2857	0.2839	0.1739	0.6097
Weiclawek (Wieclawek, 2015)	0	1	0.2857	0.5061	0.3478	0.5121
Pilch (Pilch *et al.*, 2013)	0	0.3333	0.2857	0.1604	0.2173	0.1951
Gonzalez (Gonzalez *et al.*, 2013)	0.5	1	0.2857	0.3086	0.3043	0.1707
Swingle (Swingle *et al.*, 2014)	1	0.6666	0.5714	0.8888	1	0
Lang (Lang *et al.*, 2015)	0.5	0	0.4285	0.5061	0.5217	0

Table 2.6: Mean Dice correlation of the different cyst segmentation methods on normal B-scans obtained using Cirrus and Spectralis vendor OCT system

Method	Cirrus Scan	Spectralis Scan
Wikkins (Wilkins *et al.*, 2012)	0.3032	8.51E-05
Weiclawek (Wieclawek, 2015)	2.53E-05	1
Pilch (Pilch *et al.*, 2013)	4.16E-06	0.4507
Gonzalez (Gonzalez *et al.*, 2013)	1.17E-10	0.8500
Swingle (Swingle *et al.*, 2014)	1.39E-08	0.0501
Lang (Lang *et al.*, 2015)	1	0.9500

Cirrus and Spectralis vendor test set. Table 2.6 shows the mean Dice correlation rate on both vendor normal B-scans. It can be inferred from the results that all the methods perform well in the Spectralis scans compared to Cirrus scans. Methods proposed by Wieclawek in (Wieclawek, 2015) and Lang *et al.* in (Lang *et al.*, 2015) perform better compared to other methods in both vendor scans (higher Dice-coefficient) rate. Method proposed by Lang (Lang *et al.*, 2015) *et al.* preserves higher precision and efficiently reduces the false positives when compared to other methods as shown in results of Tables 2.2, 2.3 and 2.6.

2.5.2 Qualitative Assessment

This sub-section discuss about the qualitative limitations of the aforementioned methods to segment cysts from OCT images. Future works that address these limitations can significantly enhance the accuracy of automated cyst segmentation and retinal pathol-

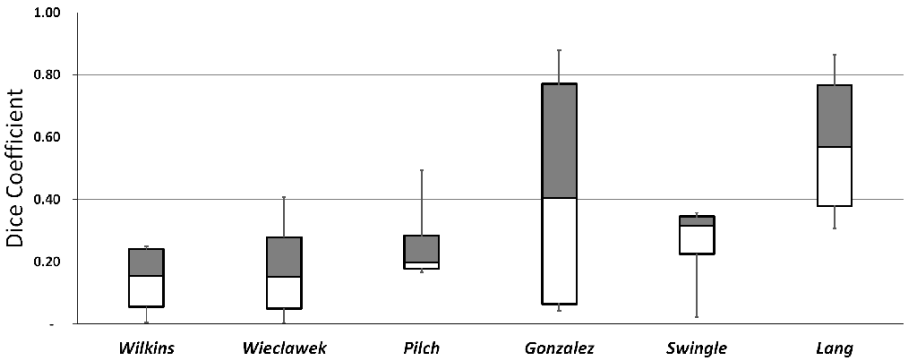
ogy diagnostics.

First, for the Wilkins method, the number of FPs are high and precision is low due to the lack of proper post-processing. For e.g., retinal blood vessel shadows and elongated cysts are not removed here. This method also suffers from manual thresholding, thus leading to low generalizability. It is observed that for a high threshold value, the number of FPs tremendously increases while for a low threshold value the number of FNs increase. Similar observations are found for the Wieclawek method.

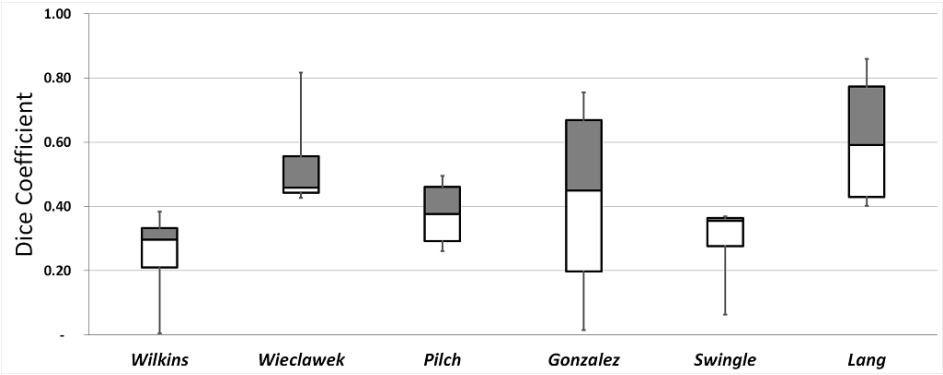
Furthermore, the Cirrus scans are more noisy and they have a low mean precision value when compared to the Spectralis scans. With the Gonzalez method, the Bayesian classifier outperforms SVM and RF classifiers. This indicates that the selection of classifier plays an important role for cyst segmentation. Also, GLCM features are found to produce better accuracy than Gabor filters. From this analysis, we can infer that texture-based features have a significant role in discriminating cystic structures from other regions in OCT B-scans.

Further, it can be observed from the recall values with Pilch method, the method performs well in the low intensity regions even though it is more prone to generating FPs in the presence of pixel intensity variations and blood vessel shadows. Experiments with the Pilch method also demonstrates that Hu-moment features extracted from the segmented regions are sufficient for separating cysts from non-cysts. Besides, is observed that several factors including blood vessel shadows, epi-retinal membrane folds and distortions in layers can affect the method by Swingle *et al.* and Lang *et al.*. Also, notice that the empirically selected threshold values for classifiers yields better segmentation accuracies when compared to the default threshold value of 0.5. Finally, it can be observed that the three stage post-processing of classifier output with the method proposed in (Lang *et al.*, 2015) leads to higher precision by reducing FPs when compared to that in the Swingle method.

Apart from cyst segmentation, an analysis on the impact of the denoising module on the overall cyst segmentation process is conducted. The DC box plots charts



(a)

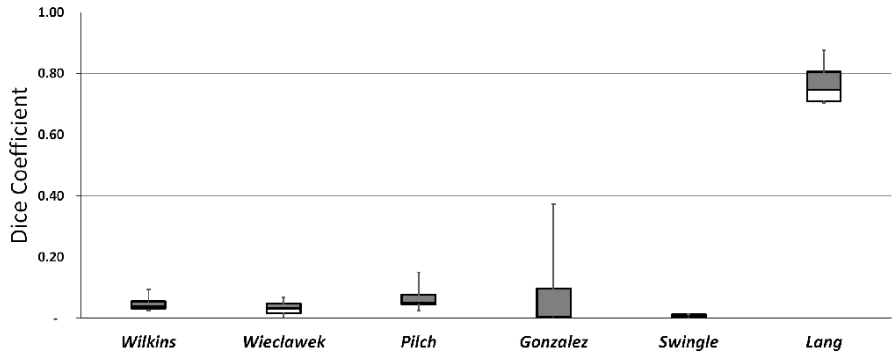


(b)

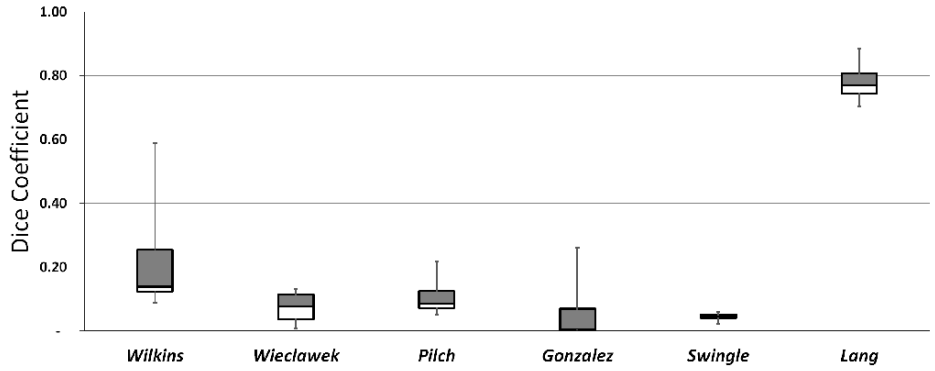
Figure 2.8: Results of DC in effect of denoising process on different automated intra-retinal cyst segmentation methods on Spectralis scans against the ground truth obtained from grader 1 ; (a) Before Denoising, (b) After Denoising.

in Figure. 2.8 and Figure. 2.9 show the variations in precision and recall for different segmentation methods before and after denoising. Here, observe that the DC values gradually increase when images are segmented after denoising. Thus, it can be inferred that the presence of image noise significantly distorts the segmentation of intra-retinal cystic regions.

From the experiments, it is observed that some of the automated segmentation methods are effective only in detecting the cyst locations but not at delineation of the actual lesion boundary, while other methods are sensitive to method parameters, thus lacking generalizability. It is acclaimed that a very important goal of any automated cyst segmentation technique should be able to segment the cyst boundary accurately for



(a)



(b)

Figure 2.9: Results of DC in effect of denoising process on different automated intra-retinal cyst segmentation methods on Cirrus scans against the ground truth obtained from grader 1 ; (a) Before Denoising, (b) After Denoising.

volumetric assessment and follow-up treatment protocols. This Chapter is aimed at providing guidance for the selection of an optimal automated intra-retinal layer and cyst segmentation method, given a specific clinical/research set-up. Various analysis conducted in this Chapter demonstrates that for clinical applications, the most significant metrics for selection of an automated segmentation method are: high recall rate and DC, followed by high precision and low-variability across vendor-specific imaging systems.

2.6 Summary

In this Chapter, the performances of a variety of supervised and unsupervised methods are compared using OCT images from Spectralis and Cirrus acquisition systems, and the cyst segmentation results are quantitatively and qualitatively analyzed. Such exhaustive analysis of existing OCT cyst segmentation methods has not been presented so far.

Analysis shows that the performance of the unsupervised methods highly depends on the quality of OCT B-scans and the supervised methods outperform unsupervised methods, even if the quality of the OCT B-scan is low. However, the limitations posed by supervised methods include: over-segmentation in the presence of several consequent cysts, and the need for training data. To overcome challenges mentioned in this Chapter, next Chapters of this thesis propose both unsupervised and supervised intraretinal cyst segmentation methods which can efficiently discriminate cyst and non-cyst pixels.

CHAPTER 3

MARKER CONTROLLED WATERSHED TRANSFORM FOR INTRA-RETINAL CYST SEGMENTATION

An automatic intra-retinal cyst segmentation method using marker-controlled watershed transform on OCT B-scans is proposed in this Chapter[3]. The proposed method is based on two stages – k-means clustering technique is used to identify cysts in the form of markers, followed by topographical based watershed transform for final segmentation. Qualitative and quantitative evaluation of proposed method is carried out against ground truth obtained from two graders on OCSC dataset Spectralis vendor OCT scans.

3.1 Introduction

In the literature, marker-controlled watershed (MWS) segmentation approach has given promising segmentation results on other medical image modalities such as mammograms (Xu *et al.*, 2011), computerized tomography (Yan *et al.*, 2006; Cristoforetti *et al.*, 2008), ultrasound (Gomez *et al.*, 2010) and magnetic resonance imaging (Cui *et al.*, 2009). This motivated us to explore MWS in segmenting retinal cysts from OCT images.

[3]The work described in this Chapter has been published in: **G. N. Girish**, Abhishek R. Kothari and Jeny Rajan, "Automated Segmentation of Intra-Retinal Cysts from Optical Coherence Tomography Scans Using Marker Controlled Watershed Transform", in *38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC'16)*. Orlando, Florida, USA: IEEE, Aug 2016.
G. N. Girish, Abhishek R. Kothari and Jeny Rajan, "Marker controlled watershed transform for intra-retinal cysts segmentation from optical coherence tomography B-scans", *Pattern Recognition Letters*, Elsevier, (In Press, 2017).

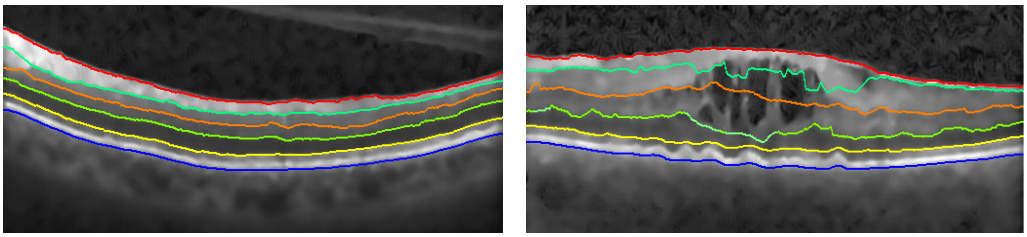
In OCT images, cysts are consecutive and congruent in nature and boundaries among them are very difficult to differentiate. The proposed approach overcomes this by locating cysts as markers using k-means clustering (coarse segmentation step) and segmenting them up to their true boundary using the gradient magnitude image as segmentation function in MWS (fine segmentation step). Method proposed in this Chapter makes three key contributions. First, an efficient marker generation technique for OCT images based on k-means clustering is proposed and the optimal value of k for distinguishing the OCT layers and abnormalities has been suggested. Second, post-processing steps for reducing the false positives is used. Third, delineates the cysts up to their true boundary, which is crucial in case of consecutive cysts and varied cyst morphology.

3.2 Methodology

OCT acquired images of the retina contain vitreous fluid region followed by retina and choroid regions. Detection of different layers of the retina is important for disease characterization (Al-Mujaini *et al.*, 2013). In the proposed method, pre-processing and retinal layer segmentation is performed and subsequently cysts are segmented using topographical based marker controlled watershed transform. Finally, segmented results are post-processed to reduce false positives.

3.2.1 Pre-processing

Performance of any segmentation algorithm depends on the quality of the input image. OCT images are generally corrupted with speckle noise. Hence, these images need to be despeckled prior to segmentation. Several methods have been proposed in the literature for de-speckling OCT images (Wilkins *et al.*, 2012; Gonzalez *et al.*, 2013; Sudeep *et al.*, 2016*a*). In the method proposed in this Chapter, Bayesian non local means (BNLM) filter (Coupe *et al.*, 2009) is applied to denoise OCT scans due to its



(a) (b)

Figure 3.1: Optical coherence tomography scan of the retina to visualize the order and position of the boundaries of layers (a) normal retina (b) retina with CME pathology. (From top: red- internal limiting membrane(ILM); light green- retinal nerve fiber layer boundary (RNFL) orange- inner plexiform layer (IPL); dark green- outer plexiform layer (OPL); yellow- outer nuclear layer (ONL); blue- retinal pigment epithelium (RPE)). (Segmented using OCTSEG (Optical Coherence Tomography Segmentation and Evaluation GUI) tool (Mayer *et al.*, 2010))

superior performance over other methods.

3.2.2 Retinal Layer Segmentation

Retina is a complex layered structure between vitreous fluid and choroid region in a OCT scan. Understanding the morphological characteristics and distortion in retinal layers is a primary task in any pathological analysis. Intra-retinal cysts are prone to occur in region between internal limiting membrane and outer plexiform layer layers of the retina. Segmentation of retinal layers play vital role in region of interest boundary determination for cyst segmentation. Methods have already been developed for retinal layer segmentation (Mayer *et al.*, 2010; Lang *et al.*, 2013).

In this Chapter, OCTSEG Tool (optical coherence tomography segmentation and evaluation GUI tool) (Mayer *et al.*, 2010) is used to automatically segment six retinal layers namely, ILM, RNFL, IPL, OPL, ONL, and RPE layer. Results of layer segmentation on normal retina and on a CME pathological case are shown in Figure. 3.1. It can be seen that OCTSEG tool is robust even in the presence of pathological changes in the retina and accurately estimates the boundary of intra-retinal layers. This is crucial for segmenting retinal images.

3.2.3 Intra-Retinal Cyst Segmentation using Marker Controlled Watershed Transform

Watershed transform algorithm segments the image into several regions using geodesic transform approaches (Beucher and Meyer, 1992). Improper regional minima imposition leads to over-segmentation problems in traditional watershed transform. Marker controlled watershed transform approach addresses this problem by initially marking the patterns to be segmented, followed by watershed computation on segmentation function (Beucher and Meyer, 1992; Lotufo and Falcao, 2000). Intra-retinal cysts are low intensity regions in the retinal OCT images and marker controlled watershed transform can be used to segment these regions effectively. To accomplish this, initially markers are generated for cystic regions followed by segmentation function creation and watershed computation.

Marker Generation and Segmentation Function Computation

Marker generation for watershed segmentation depends on scope of the problem. For intra-retinal cyst segmentation, the markers are cystic region locations. In the proposed approach, k-means clustering technique (Hartigan and Wong, 1979) is employed to identify cystic locations. k-means clustering algorithm groups the unlabeled data into k different clusters based on minimum squared error criterion S_E . It is an iterative clustering technique which aims to reduce the intra-cluster variance. The minimum squared error criterion S_E is given as (Hartigan and Wong, 1979),

$$S_E = \sum_{i=1}^{k} \sum_{p \in C_i} ||p - \mu_i||^2 \tag{3.1}$$

where $||p - \mu_i||$ is the Euclidean distance between mean vector (cluster center) μ_i of cluster C_i and pixel p of the image.

After performing k-means clustering, cluster with minimum centroid $C_{\min}$ is cho-

Algorithm 1 Marker generation using k-means clustering

Input: Image I, ILM and OPL layer boundaries, and number of clusters k.
Output: Marker Image M_k.
Initialize *k* random points as centroids.
Initialize marker image M_k with zeros of size equal to original image *I*.
For each pixel p of image I, calculate the squared error S_E between the cluster centroids μ_i and p using Eq. (3.1).
Assign each pixel p in I to the nearest center based on S_E.
After labeling of all the pixels in I, recalculate new centre μ_i for each cluster C_i using

$$\mu_i = \frac{1}{n} \sum_{k=1, p \in C_i}^{n} p_k \tag{3.2}$$

Repeat the process until cluster centroids are unchanged.
Find cluster $C_{\min}$ with minimum centroid $\mu_{\min}$ among all centroids μ_k.
Assign cluster pixels $p \in C_{\min}$ to 1 in image $I_{\min}$.
for each region $r \in C_{\min}$ **do**
    if r is connected to image extremes (boundaries) or is outside ILM and OPL layers **then**
        goto 12
    else
        Compute center of r: $[r_x, r_y] \leftarrow \text{centroid}(r)$.
        Create marker in M_k: $M_k(r_x, r_y) \leftarrow 1$.
    end if
end for
Return marker image M_k.

sen for marker generation as cysts are very low intensity regions in OCT scans. Subsequently, segmented outlier regions are removed in $C_{\min}$ and center pixel location (r_x, r_y) are computed for each remaining region r. After this, the pixel $M_k(r_x, r_y)$ in marker image M_k is labeled as 1. Algorithm 1 explains the marker generation process.

Next step in the proposed method is to compute the segmentation function $(S(x))$, which is another input to the MWS. $S(x)$ is computed by obtaining the gradient magnitude of the input image f (Eq. (3.3)). $S(x)$ is given as,

$$S(x) = \text{mag}(\nabla(f)) = \sqrt{g_x^2 + g_y^2} \tag{3.3}$$

where g_x and g_y are gradient along x and y dimensions of the image f.

Watershed Computation and Post-processing Operations

Marker controlled watershed approach proposed in this Chapter utilizes graph-cut and image foresting transform for creating catchment basins (Lotufo and Falcao, 2000). This method models segmentation function $S(x)$ as a connected digraph by representing each pixel as node and connectivity between pixels as edges. Computation of watershed catchment basin is done using marker image M_k and connectivity between nodes is determined by structuring element S_t. Initially, marker pixels of M_k are imposed on $S(x)$ by marking them as source nodes in the digraph. Then, catchment basin CB_i associated with marker pixel region R_i of $S(x)$ is given as (Lotufo and Falcao, 2000),

$$CB_i = \{x : S^*(R_i, x) \leq S^*(R_j, x), i \neq j\}, \tag{3.4}$$

where $S^*(R_i, x)$ is the shortest path from source region R_i to given pixel x (Lotufo and Falcao, 2000). The shortest path $S^*(R_i, x)$ is calculated as,

$$S^*(R_i, x) = \min\{C^*(s, x) : s \in R_i\}, \tag{3.5}$$

where $C^*(s, x)$ is the minimum cost of all the paths between source node s and destination node x. This can be computed as (Lotufo and Falcao, 2000),

$$C^*(x, y) = \min_i\{C(\pi_i(s \rightarrow x))\}, \tag{3.6}$$

where $\pi_i(s \rightarrow x)$ is a path from source s to destination x with connectivity defined by S_t.

Resultant catchment basins segmented using MWS contain true cystic regions and additional false positive regions. Watershed regions are post-processed to reduce FP. In OCT scans, intra-retinal cysts appear as low intensity and oval shaped regions. Hence, regions having intensity greater than catchment basin maximum intensity threshold ($CB_{\max}$) are removed. For removal of rectangular cystic structures, catchment basin

width to height ratio (CB_h) is used. Due to noise and shadow artifacts in OCT scans, some small structures are segmented as pseudo-cysts, which are removed using catchment basin area threshold (CB_a) pixels.

3.3 Experimental Results and Discussions

To evaluate the performance of the method proposed in this Chapter, several experiments were conducted on OPTIMA cyst segmentation challenge dataset (Spectralis vendor scans) (OPTIMA, 2015). Performance of the proposed segmentation algorithm was compared against k-means clustering, traditional watershed transform and other unsupervised state-of-the-art IRC segmentation methods.

3.3.1 Dataset Description

The OPTIMA cyst challenge dataset (Spectralis vendor scans) contains 8 volumes in total (6 with 49 frames and 2 with 7 frames). All OCT volumes were acquired over a 6×6 mm area of the macula centered on the fovea, from subjects with CME pathology. The provided OCT scans contain grayscale frames with a resolution 496×512 and 496×1026 pixels. Sample OCT B-scans from OPTIMA challenge dataset (obtained using Spectralis Heidelberg OCT) is shown in Figure. 2.3(c). This dataset was used for validation of the proposed segmentation algorithm because of the presence of variety of features and pathologies (including foveal scans, hard exudates, blood vessel shadows, epiretinal membrane, and ILM folds) compared to other vendor scans.

3.3.2 Experiments

Since OCT scans have significant speckle noise, prior to segmentation tasks, scans are denoised using BNLM filter (Coupe *et al.*, 2009). The BNLM filter is a variant of non-

local means filter which denoises the images using self similar non-local patches. The BNLM filter is executed with the following parameters: search window size 11×11, similarity window size 5×5 and $S_\sigma = 0.7$. After denoising, six retinal layers are automatically segmented using OCTSEG tool (Mayer *et al.*, 2010).

Intra-retinal cyst segmentation is done using marker controlled watershed transform. Initial marker image is generated using k-means clustering. An experimental analysis is conducted to verify the best choice for selecting value of k. Marker images are generated with different values of k (varying from 2–9) and the results are as shown in Figure. 3.2. It can be inferred from the images that retinal structures are over-segmented for smaller values of k. However, higher value of k leads to poor delineation of structures due to the presence of high number of clusters (see Figure. 3.2). Hence, it is evident that k-means clustering produces optimal results when $k = 3$.

Clusters formed with the obtained optimal value were grouped together according to hyper-reflective, hypo-reflective and very low intensity pixels. Clusters with hypo- and hyper-reflective pixels represent different retinal layers while the very low intensity clusters contain candidate cyst regions.

After clustering the image, cluster with minimal centroid is used to remove outliers. These are the regions connected to left and right extremes of the image (20 pixels) and regions outside the region of interest (ILM and OPL layers). Finally, center pixel of each remaining region is used to construct the marker image M_k (See Section 3.2.3 and Algorithm 1).

After marker generation, segmentation function is computed as explained in Section 3.2.3. Topographical watershed segmentation is performed as described in Section 3.2.3. In order to reduce false positives, catchment basin maximum intensity threshold $CB_{\max} \geq 125$, catchment basin width to height ratio $CB_h \geq 4$ and pseudo cysts with catchment basin area threshold $CB_a = 10$ are used. After removal of FP regions, resultant image contains the segmented output by the proposed method.

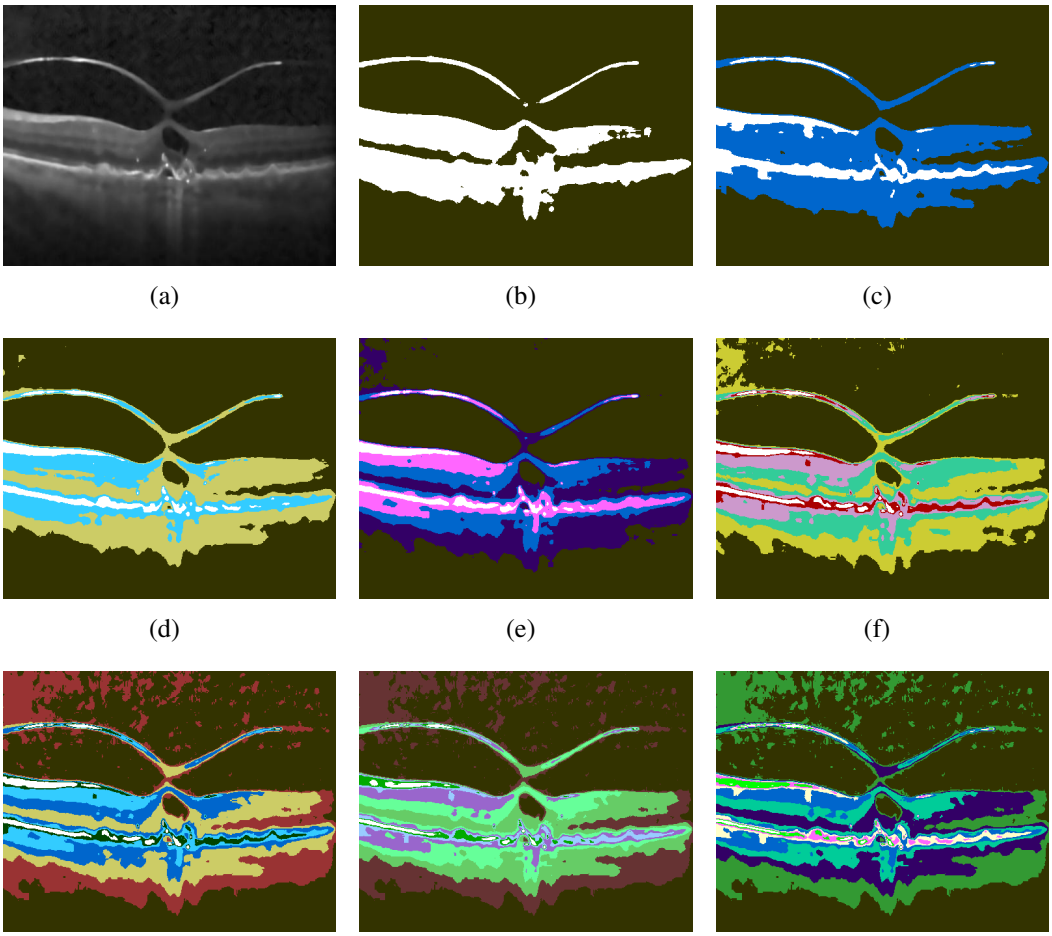


Figure 3.2: Segmentation result of a OCT image using k-means clustering with different number of clusters k. (Different colors represents different clusters) (a) Original image; (b) $k = 2$; (c) $k = 3$; (d) $k = 4$; (e) $k = 5$; (f) $k = 6$; (g) $k = 7$; (h) $k = 8$; (i) $k = 9$.

3.3.3 Results and Discussions

Evaluation of the proposed method is done qualitatively and quantitatively. To analyze the generalizability, segmentation is performed using 8 OCT volumes of Spectralis vendor scans. Obtained segmentation results are ascertained using ground truth obtained from two expert ophthalmologists (OPTIMA, 2015).

Table 3.1: Mean (standard deviation) of precision and recall for proposed method, k-means clustering, traditional watershed transform and state-of-the-art unsupervised segmentation methods against G_1, G_2, and $G_1 \cap G_2$ (computed using region-wise analysis).

Method	G_1		G_2		$G_1 \cap G_2$	
	Precision	Recall	Precision	Recall	Precision	Recall
k-means clustering	0.81 (0.26)	0.56 (0.35)	0.81 (0.25)	0.59 (0.35)	0.81 (0.25)	0.57 (0.34)
Traditional watershed	0.55 (0.25)	0.65 (0.36)	0.54 (0.25)	0.65 (0.35)	0.53 (0.24)	0.63 (0.37)
Wilkins (Wilkins *et al.*, 2012)	0.34 (0.28)	0.67 (0.39)	0.34 (0.27)	0.68 (0.40)	0.34 (0.27)	0.66 (0.39)
Wieclawek (Wieclawek, 2015)	0.57 (0.26)	0.39 (0.23)	0.57 (0.24)	0.41 (0.23)	0.57 (0.26)	0.41 (0.23)
Proposed	0.78 (0.14)	0.65 (0.27)	0.80 (0.15)	0.68 (0.28)	0.78 (0.15)	0.67 (0.28)

Quantitative Analysis

Region-wise precision (2.1) and recall (2.2) metrics are computed per OCT volume against ground truth of G_1, G_2, and $G_1 \cap G_2$.

Table 3.1 shows the result of mean precision and recall computed using region-wise analysis of the proposed segmentation algorithm, k-means clustering, traditional watershed transform and other unsupervised intra-retinal cyst segmentation techniques against G_1, G_2, and $G_1 \cap G_2$. For proposed method, recall rates are 0.65, 0.68, and 0.67 for G_1, G_2, and $G_1 \cap G_2$, respectively, whereas precision rates are 0.78, 0.80, and 0.78 for G_1, G_2, and $G_1 \cap G_2$, respectively. It can be seen that proposed method gives better recall rate while maintaining high precision rate.

Results are also compared with other unsupervised intra-retinal cyst segmentation methods. The same pre-processing and retinal layer segmentation steps are used across all the methods for unbiased comparison.

For k-means clustering technique, segmented regions are post-processed to remove FPs using, retinal boundary (ILM-OPL) threshold, region width to height ratio $R_h \geq 4$, and, region area $R_p \leq 10$. In traditional watershed segmentation technique, post-processing is carried out as mentioned in Gonzalez method (Gonzalez *et al.*, 2013), to

reduce FP with following parameters: elongated cysts ($t_1 = 1.5$, $t_2 = 2$ and $p_{\text{occ}} = 0.8$), small cysts ($n_{\min} < 10$ pixels), cysts next to image margins, and cysts outside the retinal boundary (ILM-OPL). Precision and recall rates of our method surpasses the results of k-means clustering and traditional watershed segmentation methods.

For the methods reported by Wilkins *et al.* and Wieclawek, core algorithm and post-processing steps are implemented as mentioned in the published papers (Wilkins *et al.*, 2012; Wieclawek, 2015). For Wilkins *et al.* (Wilkins *et al.*, 2012), cysts are segmented using empirically computed intensity threshold value $t = 40$. The standard deviation value $S_d = 41$ and pseudo cyst area $P_a = 10$ pixels are used to reject FP regions. In Wieclawek (Wieclawek, 2015), h-minima transform is obtained using empirically computed value $h = 6\%$ of the maximum normalized intensity threshold. To construct binary image, h-minima transformed image is thresholded. To improve the results, ILM-OPL layers are used instead of RNFL-RPE layers to limit the region of interest. In both the algorithms, segmented regions outside region of interest as well as regions connected to image margins are removed to reduce FP regions.

Precision obtained from the proposed method is improved by 44% and 21% compared to Wilkins *et al.* (Wilkins *et al.*, 2012) and Wieclawek (Wieclawek, 2015), respectively. In addition to this, recall of the proposed method also improved by around 1% and 26%, respectively, on GT of $G_1 \cap G_2$.

In the literature, cyst segmentation techniques have been evaluated based on region-wise analysis without considering individual pixels that are segmented into the cystic regions. In this study, we have additionally computed precision and recall metrics based on pixel-wise analysis to better characterize the performance of individual segmentation methods. For pixel-wise analysis, true positives are true cystic pixels detected by the algorithm, false positives are not actual cystic pixels identified as cystic pixels by the algorithm and false negatives are true cystic pixels that remain undetected by the algorithm. Precision and recall metrics are computed against grader gold standard using Eqs. (2.1) and (2.2), respectively. Results are tabulated in Table 3.2.

Table 3.2: Mean (standard deviation) of precision and recall for proposed method, k-means clustering, traditional watershed transform and state-of-the-art unsupervised segmentation methods against G_1, G_2, and $G_1 \cap G_2$ (computed using pixel-wise analysis).

Method	G_1		G_2		$G_1 \cap G_2$	
	Precision	Recall	Precision	Recall	Precision	Recall
k-means clustering	0.89 (0.09)	0.33 (0.25)	0.90 (0.09)	0.32 (0.23)	0.87 (0.10)	0.38 (0.26)
Traditional watershed	0.43 (0.24)	0.43 (0.25)	0.43 (0.24)	0.41 (0.23)	0.40 (0.22)	0.48 (0.26)
Wilkins (Wilkins *et al.*, 2012)	0.40 (0.34)	0.35 (0.27)	0.40 (0.34)	0.33 (0.26)	0.38 (0.34)	0.39 (0.29)
Wieclawek (Wieclawek, 2015)	0.72 (0.19)	0.33 (0.19)	0.73 (0.17)	0.31 (0.17)	0.68 (0.20)	0.36 (0.21)
Proposed	0.75 (0.21)	0.56 (0.19)	0.76 (0.22)	0.54 (0.19)	0.70 (0.23)	0.61 (0.20)

Precision and recall values with pixel-based analysis demonstrate some differences as compared to region-wise analysis. Even though a segmentation algorithm might segment cyst regions, the accuracy of correctly delineating the regions up to the boundary is crucial. Pixel-wise analysis provides a measure of accuracy of pixel classification and accurate delineation of identified cyst region. The importance of the same in quantitative analysis of the OCT image can not be overstressed. Therefore, this rigorous method is employed to test and represent the performance of segmentation algorithms along with region-wise analysis.

It is noted that recall values of the state-of-the-art methods reduced by around 10–20% in pixel-wise analysis compared to region-wise analysis (Table 3.2). It can be inferred from this result that these methods cannot reproducibly delineate the segmented region up to the boundary. In pixel-wise analysis for accuracy (recall rate), proposed method outperforms other compared methods while preserving the precision rate, and improved upon the performance of k-means method by around 23%, with delineation of cyst regions up to the true boundary. It also improved upon traditional watershed method by around 30% (precision) and 13% (recall) (see Table 3.2) by limiting over-flooding using a marker function to define proper regional minima. This method also performs favorably compared to method reported by Wilkins *et al.* (Wilkins *et al.*, 2012) and Wieclawek (Wieclawek, 2015) by improving precision and recall as well. Precision

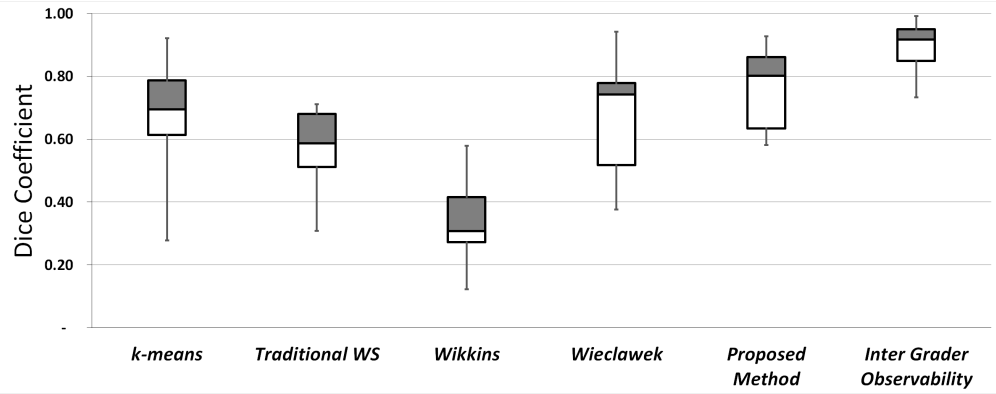


Figure 3.3: Whisker Box plot for the Dice correlation coefficient against intersection of both the graders for k-means clustering (median $= 0.69$), traditional watershed transform (median $= 0.59$), Wilkins (Wilkins *et al.*, 2012) (median $= 0.31$), Wieclawek (Wieclawek, 2015) (median $= 0.74$), Proposed segmentation method (median $= 0.80$) and inter grader observability between G_1 and G_2 (median $= 0.92$).

of the proposed method is lower compared to k-means clustering. This is due to false positives generated from over-flooding.

To measure the segmentation correlation between algorithm result and ground truth, Dice coefficient (2.3) (Dice, 1945) is used.

DC is computed between segmentation result of the methods and ground truth of $G_1 \cap G_2$ for each OCT volume, and it is plotted in Figure. 3.3. It is observed that the proposed method has higher overlap rate compared to other methods with a median of 0.80 and maximum correlation up to 0.95. k-means clustering method given DC coefficient (with a median of 0.69, maximum correlation up to 0.90 and minimum correlation of 0.35) due to failure of proper delineation of cystic regions.

Traditional watershed method has a median correlation of 0.59 and the maximum correlation of 0.73. The method of Wilkins *et al.* (Wilkins *et al.*, 2012) has a median DC of 0.31 and maximum of 0.57 as manual thresholding led to imperfectly delineated segmented regions. The method proposed by Wieclawek (Wieclawek, 2015) has maximum DC of 0.92 and a median DC of 0.75. Inter-grader segmentation correlation between two graders on the dataset has a median DC of 0.92, a maximum of 0.98 and a

minimum of 0.70. Due to intensity variations and confounding retinal pathologies, it is difficult to accurately delineate cystic fluids even for expert ophthalmologists in some OCT scans.

Qualitative Analysis

Segmentation results of the proposed and compared methods are shown in Figures. 3.4 and 3.5. In Figure. 3.4, the OCT B-scan does not show undulations in retinal layers and retinal fluid is accumulated in multiple individual consecutive cysts (see Figure. 3.4(a)). Figure. 3.4(b) shows the ground truth of $G_1 \cap G_2$, in which accumulated fluid is marked as different independent consecutive cystic locations in the retina. k-means clustering method segmented cysts in the retina as several broken cysts and fails to delineate them up to the boundary possibly due to intensity variations inside the cystic structures (see Figure. 3.4(c)).

Traditional watershed transform method segments independent consecutive cysts as one large cyst because of over-flooding in the catchment basin (see Figure. 3.4(d)). In Figure. 3.4(e), the method proposed by Wilkins *et al.* (Wilkins *et al.*, 2012) segments large cysts as several broken cysts. This method is sensitive to intensity variations, and it may fail to delineate the cystic fluids up to their true boundary (observe the marginal difference between the ground truth and segmented result). The method proposed by Wieclawek (Wieclawek, 2015) over-segments individual consecutive cystic regions as large cyst regions as seen in the Figure. 3.4(f). There are several false positives due to lack of post-processing steps (Figure. 3.4(f)). The method proposed by us delineates the cysts accurately up to the boundary by the two stage approach (see. Figure. 3.4(g)). Our method also segments two false positives due to scan intensity variations of the OPL (see Figure. 3.4(g)). Figure. 3.4(h) shows inter-grader observability between G_1 and G_2, in which G_1 segmented consecutive large cysts as several independent cysts (pink color marked boundaries) and G_2 segmented consecutive cysts as large cysts (yellow color marked boundaries).

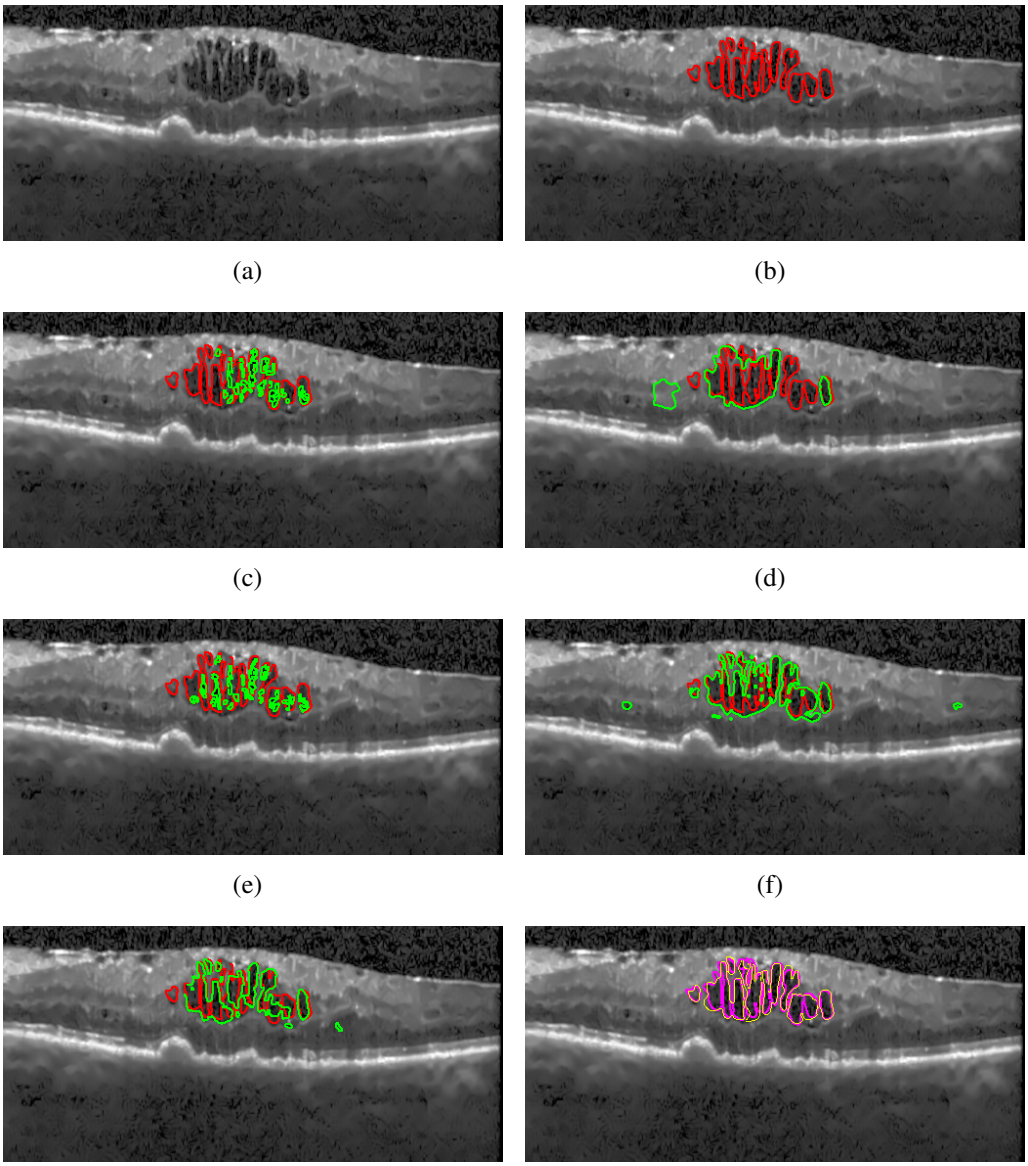


Figure 3.4: Results of proposed automatic cyst segmentation method and state-of-the-art methods in presence of limited retinal layer undulations: (a) Original B-scan; (b) Ground truth of $G_1 \cap G_2$; (c) k-means clustering; (d) Traditional watershed; (e) Wilkins method (Wilkins *et al.*, 2012); (f) Wieclawek method (Wieclawek, 2015); (g) Proposed method; (h) Inter grader observability (Pink: G_1, Yellow: G_2). (Red: Ground truth of $G_1 \cap G_2$, Green: Algorithm result)

Figure. 3.5 shows the segmentation of intra-retinal cyst in the presence of significant retinal layer undulation due to vitreomacular traction. The B-scan shown in Figure. 3.5 contains a cyst with vitreomacular traction and a pigment epithelial detachment (PED)

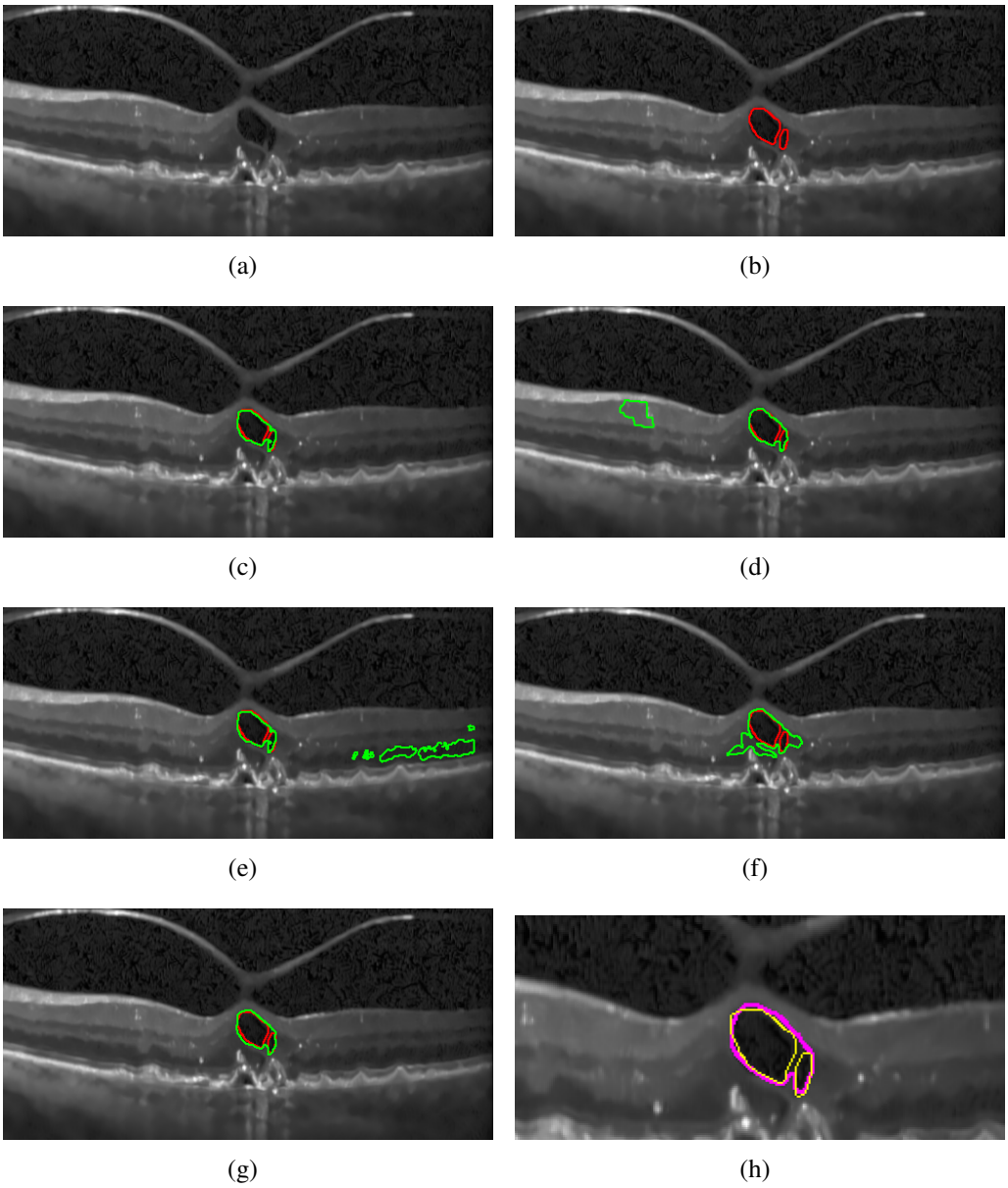


Figure 3.5: Results of proposed automatic cyst segmentation method and state-of-the-art methods in presence of significant retinal layer undulations due to vitreomacular traction with pigment epithelial detachment with cysoid macular edema: (a) Original B-scan; (b) Ground truth of $G_1 \cap G_2$; (c) k-means clustering; (d) Traditional watershed; (e) Wilkins method (Wilkins *et al.*, 2012); (f) Wieclawek method (Wieclawek, 2015); (g) Proposed method; (h) Inter grader observability (Pink: G_1, Yellow: G_2). (Red: Ground truth of $G_1 \cap G_2$, Green: Algorithm result)

in the foveal region of the retina (see Figure. 3.5(a)). Figure. 3.5(b) shows the ground truth of $G_1 \cap G_2$, in which accumulated fluid was marked as different independent

consecutive cystic locations in the retina. k-means clustering method segments the cystic fluid as one region with minimal difference in the correlation of the true boundary of the cyst (see Figure. 3.5(c), best viewed in zoom). Traditional watershed transform method effectively segments the cyst up to its true boundary; it also segment a false positive on left beside the foveal region (see Figure. 3.5(d)).

Wilkins *et al.* (2012) segments true cyst with minimal correlation difference in the true boundary delineation, along with false positives due to scan intensity variation and limited post-processing steps (see Figure. 3.5(e)). Wieclawek method (Wieclawek, 2015) over-segments the cyst due to the undulation in the retinal layers and intensity variations (see Figure. 3.5(f)).

Proposed method segments the cyst up to its true boundary with higher correlation to the ground truth and without segmenting any false positives inspite of distortions in retinal layers, vitreomacular traction as well as PED beneath the foveal region of the retina (see Figure. 3.5(g)). In Figure. 3.5(h), G_1 marked one cyst (red color marked boundaries), while G_2 marked two cysts (yellow color marked boundaries), displaying inter-grader variability. Inter-grader variability increases when acquired OCT scans are affected with significant retinal irregularities, scan intensity variations and low signal-to-noise ratio.

Discussions

The obtained results with the proposed method are better than those obtained with k-means clustering, traditional watershed, and state-of-the-art unsupervised cyst segmentation techniques by improving recall and segmentation correlation while preserving precision. Preserving higher precision plays a vital role in adapting the method to clinical applications.

Delineation of the cysts by utilizing our strategy of k-means clustering for marker generation followed by watershed transform has higher accuracy and reproducibility

than other tested methods. Proper delineation is an important factor where cyst area and volume quantification is required as in clinical treatment process. Average computational cost of the proposed method per B-scan is 1.31 seconds on a workstation with Intel (R) Xeon (R) CPU E5-2670 v3 2.30GHZ processor, 128 GB DDR3 RAM, 64 GB Windows 10 operating system and MATLAB computational environment.

The proposed method has following advantages:

- It is a fully automatic unsupervised segmentation technique.
- It follows a two stage segmentation approach, which initially locates the cysts and delineates them up to their true boundary.
- It is robust against retinal layer distortions, epiretinal membrane folds and blood vessel shadows.
- It is validated against benchmarked intra-retinal cyst segmentation dataset and gives better results compared to other reported unsupervised segmentation techniques.

There are certain limitations for the proposed method. The performance of our method depends significantly on the quality of input image. Watershed transform is very sensitive to noise and requires robust speckle reduction in pre-processing stage to improve the SNR, otherwise it affects the performance of the segmentation algorithm. Incorrect retinal layer segmentation also leads to the removal of candidate cysts during post-processing steps. Accurate segmentation of layers is in turn dependent on quality the of input OCT image.

3.4 Summary

This Chapter presented a novel unsupervised two-stage automatic intra-retinal cyst segmentation algorithm based on marker-controlled watershed transform. Experimental results show that the proposed method performs favorably compared to other reported

state-of-the-art methods for segmentation of intra-retinal cysts from OCT images. However, the proposed method is vendor dependent and sensitive to noise and image intensity variations. To address this issue an automatic supervised method using convolutional neural network is proposed in the next Chapter for vendor independent IRC segmentation.

CHAPTER 4

A FULLY CONVOLUTIONAL NEURAL NETWORK MODEL FOR AUTOMATED SEGMENTATION OF INTRA-RETINAL CYSTS FROM CROSS VENDOR OCT SCANS

All the methods described in the previous two Chapters are vendor dependent studies and this limits the complexity associated with accurate IRC segmentation across different manufacturing vendors. To address this problem, this Chapter[4] proposes a fully convolutional network model for vendor-independent IRC segmentation. Proposed FCN extracts more abstract and complex features from the cross vendor OCT scans that accurately defines cystic structures across different scan intensity variations and noise levels.

4.1 Introduction

Over the past few years, several automated IRC segmentation methods have been proposed in the literature. To the best of our knowledge, the existing methods are dependent on vendor-specific acquisition systems and require human intervention to determine acquisition and visualization-specific parameters. Additionally, segmentation of fluid-filled spaces in retinal volumes across different OCT vendors is a challenging

[4]The work described in this Chapter has been published as: **G. N. Girish**, Bibhash Thakur, Sohini Roychowdhury, Abhishek R. Kothari, and Jeny Rajan, "Segmentation of Intra-Retinal Cysts from Optical Coherence Tomography Images using a Fully Convolutional Neural Network Model". *IEEE Journal of Biomedical and Health Informatics*, IEEE, (In Press, 2018).

task due to the variations in the noise levels and pixel intensities across scans (see Figure. 2.3). The proposed method addresses these challenges posed to the development of automated vendor independent cyst segmentation methods.

Recent forays into the usage of convolutional neural networks for biomedical image segmentation have shown significant improvements over other approaches (Kamnitsas *et al.*, 2015; Lequan *et al.*, 2016). CNNs are a specific variant of standard neural network models, which take an image as an input and perform repetitive convolution operations to abstract features specific to the segmentation task. The learnable weights and biases of the neurons serve as convolution filters. These filters may be arranged in different ways to design architectural variants for various segmentation problems.

One such significant CNN model used in a variety of different medical image segmentation tasks, is the U-Net architecture proposed by Ronneberger *et al*. (Ronneberger *et al.*, 2015). U-Net is a fully convolutional network, such as the one popularized by Long and Shelhammer (Long *et al.*, 2015), where, if an image is provided as the input, the output obtained is its corresponding segmentation mask. Milletari *et al.* (Milletari *et al.*, 2016) further expanded this concept and converted their model to accept 3D image stacks as input. They also introduced the concept of residual connections described in the widely cited ResNet architecture (He *et al.*, 2016). The self-learning and abstract feature learning capabilities of CNNs for distinguishing subtle spatial variations motivate our choice of FCN in vendor-independent cyst segmentation tasks. So far, FCNs have not been utilized for IRC segmentation from vendor-independent OCT scans.

This Chapter makes two key contributions. First, a customizable state-of-the-art FCN model is presented that is capable of automating IRC area segmentation from OCT images across vendor-specific imaging systems. The sensitivity of model parameters, such as number of layers and kernel dimensions, towards the IRC area segmentation goals are analyzed. It is observed from the results that an optimally parametrized model can achieve higher recall rate of 0.66 while preserving the precision rate of 0.79 across multiple vendor data, when compared to state-of-the-art methods. Second, the

importance of OCT image pre-processing by image noise suppression using Gamma noise models, sub-retinal region of interest segmentation and optimal data augmentation methods are presented. It is observed that image pre-processing and domain-specific data augmentation methods significantly prevent model over-fitting while ensuring generalizability across vendor-specific imaging systems.

4.2 Brief Introduction to Neural Networks

Before diving into the methodology and framework of the proposed solution, this section provides brief introduction into the working of neural networks in general and convolutional neural networks in particular. Artificial neural networks (ANNs) are a type of machine learning algorithm which essentially are a stack of linear models connected in succession. A linear model is basically a function $f(x)$ which is defined as (Martin and Maes, 1979):

$$f(x) = W.x + b \tag{4.1}$$

Here, x is a vector of inputs to the function, W is a matrix of weights mapping the input vector to the output and b is a vector of bias parameters added to the output of the vector-matrix multiplication. The output of this linear model will also be a vector of the size defined by the weight matrix and bias vector. The simplest classification strategy can be implemented with just one linear model. In that case, x will be the input to be classified, let‘s say a vector of size $D \times 1$. The task is to classify the input into one of C classes and hence, the output of the linear model will be a vector of size $C \times 1$. So logically, the weight matrix W has to be of a size $C \times D$ and the bias vector will be of size $C \times 1$. The vector output from this model is called a score vector, as it typically contains the confidence value of each output class associated with the current input. Generally, the vector index with the greatest value associated with it is considered the class predicted by the model.

In machine learning, such types of models are used to make a computer learn to accurately solve complex problems. This is done with the process of optimization. Optimization in machine learning is an iterative process where a model learns to correctly classify the input by looking at a large number of similar examples given to the model, gradually improving its predictions. This is done with the aid of an error or loss function.

4.2.1 Optimization and Loss Functions

In standard machine learning workflow, the first task performed is to divide the data into a separate training, validation and test sets. The training and validation sets are used to optimize the model to make correct predictions and the test set is used to model the real world and judge how well the model performs in the specified task. No optimization is performed on the test set. After the division, the model is initialized with some random weights and the data in the training set is fed into the model in some order. The input passes through the random weights and outputs a score vector. This score vector is compared with the actual class output of the provided input using a loss or error function. Based on the magnitude of this loss function, the weights are updated such that the model gives a lower loss the next time this input is seen.

Cross-Entropy Loss

There are different loss functions proposed in literature to optimize such models. One of the popular ones is the cross-entropy loss and it is defined as (Goodfellow *et al.*, 2016):

$$CL_i = -\log(\frac{e^{f_{y_i}}}{\sum_j e^{f_j}}) \tag{4.2}$$

where, CL_i is the loss and y_i is the actual correct label for input i. f_j denotes the j-th element of the vector of class scores f. The cross-entropy loss is composed of the softmax function $\frac{e^{f_{y_i}}}{\sum_j e^{f_j}}$. The softmax function can be intuitively understood from a

probabilistic perspective. It can be can be interpreted as the (normalized) probability assigned to the correct label y_i given the input x and parameterized by W and b. The softmax function interprets the scores inside the output vector f as unnormalized log probabilities. Exponentiating these quantities therefore gives the unnormalized probabilities, and the division performs the normalization so that the probabilities sum to one (Li *et al.*, 2016). Therefore, the cross entropy loss minimizes the negative log likelihood of the correct class.

CL_i is the loss associated with a single input value. The total loss value T_{Loss} over all N_{all} inputs is defined as the mean of the individual loss function values for each input (Goodfellow *et al.*, 2016):

$$T_{Loss} = \frac{1}{N_{all}} \sum_i CL_i \tag{4.3}$$

Optimization: Stochastic Gradient Descent

Once the loss function value is computed, the weights in the model need to be updated such that the loss is minimized, i.e., the weights and bias of the model has to be such that the loss function reaches a global minima. This is done with the process of gradient descent. The gradient of a function with respect to its parameters is a vector of the magnitude of change of the function output along each direction, i.e., it is a vector of the partial derivatives of the function with respect to its parameters (Nielsen, 2015).

$$\nabla f(a, b, c, x) = \begin{bmatrix} \frac{\partial}{\partial a} f \\ \frac{\partial}{\partial b} f \\ \frac{\partial}{\partial c} f \end{bmatrix} \tag{4.4}$$

Here, f is a function taking an input x and having parameters a, b, and c. ∇f is then defined as the gradient of f. In the context of linear models, the gradient of T_{Loss} with respect to W and b of the model is calculated, and the corresponding weights and bias

are updated yielding W' and b' respectively (Nielsen, 2015).

$$W' = W - \alpha \frac{\partial}{\partial W} T_{Loss} \tag{4.5}$$

$$b' = b - \alpha \frac{\partial}{\partial b} T_{Loss} \tag{4.6}$$

In Eqs. (4.5) and (4.6), the parameter α refers to the step size or learning rate of the gradient descent. It defines the magnitude of change on the original parameters in each iteration. The learning rate is an important hyperparameter in machine learning algorithms, which defines how fast the model will train. A hyperparameter is a parameter that must be explicitly specified by the user and cannot be learned by a machine. A higher learning rate, which makes the loss converge faster, may not always be ideal as the updated parameters may overshoot the minima. The gradients are subtracted from the original values because a positive slope value of a function is always in the direction of the minima. W' and b' are updated such that T_{Loss} moves toward its minima.

In practice, instead of computing the loss over the entire training set, and then computing the gradient and then updating the parameters, a method called stochastic or mini-batch gradient descent is performed. In this approach, the gradient is calculated over a set batch of inputs and the weights are updated. These batches are fed into the model one after the other until the entire training set is exhausted. This is called an epoch. The size of the batches is generally constrained by the amount of physical memory of the machine, as the entire batch has to be fed into the memory at once.

4.2.2 Artificial Neural Networks

A single linear model is generally not powerful enough to emulate complex mappings from the input to the output classes. There is a need to stack these linear models one after another to better describe complex functions that occur in the real world. A stack

of linear models arranged in a manner such that the output of one linear model becomes the input of another linear model forms the basis for artificial neural networks. An ANN is just a stack of linear models with activation functions between them that map some complex relation between the input to some output.

Activation Functions

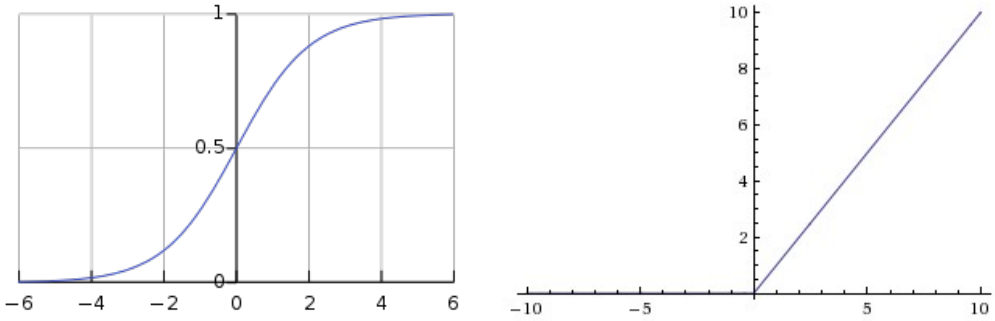
Activation functions or non-linear functions are functions applied on the output of a linear model to make it non-linear. This non-linear output may be passed as the input to another linear model. Activation functions are required because of the property of composition of linear models. This property states that, if f and g are two linear models and we have a composite function $y = f(g(x))$, then there always exists another linear model h that accurately describes y, i.e. (Nielsen, 2015),

$$y = f(g(x)) = h(x) \tag{4.7}$$

Practically, this means that multiple linear models without a non-linearity in between can be broken down into a single linear model. To better map complex relationships, a non-linear function is applied on the output of a linear model. The two most commonly used non-linear activation functions are the sigmoid function and the Rectified Linear Unit (ReLU) function. Figure. 4.1 shows sigmoid and ReLU activation functions. The sigmoid function takes a value and squashes it into the range 0 to 1. The sigmoid function is of the form (Goodfellow *et al.*, 2016):

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{4.8}$$

where, x is the input to the sigmoid function σ. The sigmoid function can be interpreted as a binary softmax function, where its value can be thought of as the confidence level between 2 distinct classes. If the value is leaning more toward 0 or 1, the model is more confident about the input belonging to that particular class. Median values near 0.5 can



(a) Sigmoid Activation (b) ReLU Activation

Figure 4.1: Non-linear activation functions. (Li *et al.*, 2016)

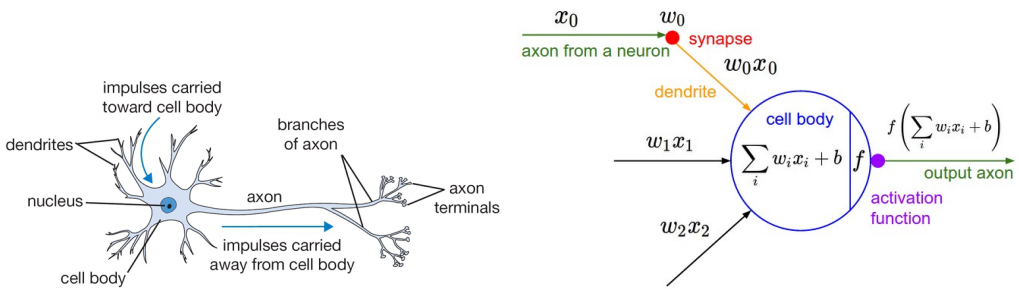
be treated as unsure observations. The sigmoid function is generally not used in any intermediate layers of an ANN because, if a sigmoid model output is 0 or 1, its gradient becomes 0, after which the model cannot update the weights. Sigmoid can be used in the last layer if it is a binary classification problem.

The ReLU is most commonly used activation function for intermediate layers. Its simple definition of thresholding the output at zero (4.9) greatly accelerates the training time of neural networks applying gradient descent (Goodfellow *et al.*, 2016).

$$R_e(x) = max(0, x) \tag{4.9}$$

It has been proved that a stack of one or more linear models followed by a non-linearity, work as universal approximators, i.e, they can approximate the output of any mathematical function with a tolerance value of ϵ (Nielsen, 2015). This concept is analogous to how the human brain works. As shown in Figure. 4.2(a), the working component of the human brain is a nerve cell, called a neuron.

A neuron accepts inputs from other neurons with variable synaptic strengths through its dendrites. Its output is propagated through the axon towards its terminals and is accepted by the dendrites of some other neurons with their respective strengths. In the computational model of the neuron, Figure. 4.2(b), multiple inputs coming in may



(a) Model of Nerve Cell (b) Neuron as a Linear Model

Figure 4.2: Models of biological and computational neurons. (Li *et al.*, 2016)

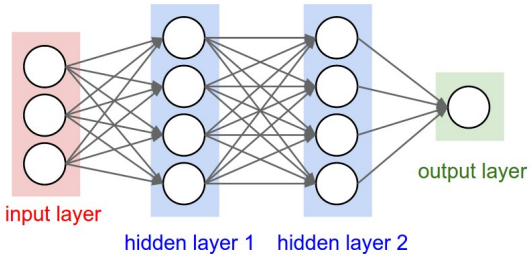


Figure 4.3: An artificial neural network with 2 hidden layers. (Li *et al.*, 2016)

be multiplied with some pre-defined weights and then added. An activation function is applied to the sum of the weighted inputs and the bias parameter of the neuron, thus functioning as a linear model passed through a non-linearity. The output is then passed to other neurons. A combination of these neurons arranged in an acyclic graph architecture forms a ANN.

Neurons are arranged in distinct layers. The neurons in a single layer accept inputs from the previous layer and passes outputs to the next layer, but individual neurons in a single layer are not connected. A simple ANN with fully connected layers is described in Figure. 4.3. A fully connected layer is a layer in which every neuron is connected to every neuron in the previous layer. The input and output layer out of a neural network do not have any weights assigned. Weights are present in the hidden layers which perform the linear computations.

Backpropagation

All the weights of an ANN are randomly assigned when it is first created. The weights are automatically updated based on the training data and its associated labels by computing the loss function at the output layer. The entire network then acts as a single model which accepts an input and outputs a class as the output. Based on the loss obtained, the weights of the hidden layers are updated such that the loss is minimized. Propagating the gradient of the loss function through the layers of the network is performed by a method called backpropagation.

Backpropagation or reverse-mode differentiation is a method by which we can check the influence of every input and intermediate node of a computational graph in the output. In the context of neural networks, it can be used to calculate the effect of every neuron in the loss function. Before computing the reverse gradient, for all the layers, the gradient of that layer with respect to the previous layer must be known. Then, starting from the output layer, the reverse is calculated using the chain rule of differentiation. Let the gradient of the loss function T_{Loss} with respect to the weights of the output layer p be δ_p^L. Then the gradient of the loss function with respect to the weights of the previous layer will be

$$\delta_{p-1}^{T_{Loss}} = \frac{\partial T_{Loss}}{\partial f_{p-1}} = \frac{\partial T_{Loss}}{\partial f_p} \frac{\partial f_p}{\partial f_{p-1}} \tag{4.10}$$

$$i.e., \delta_{p-1}^{T_{Loss}} = \delta_p^{T_{Loss}} \frac{\partial f_p}{\partial f_{p-1}} \tag{4.11}$$

where f_p and f_{p-1} signify the activated outputs of the p^{th} and $p-1^{th}$ layer respectively. Eq. (4.11) can be iteratively applied on every layer in sequence till the input layer is reached.

4.2.3 Convolutional Neural Networks

In Convolutional Neural Networks, the standard linear model of an ANN is replaced by a convolution operation. All other concepts of ANNs like Loss Functions, Optimizers, Activation Functions and Backpropagation carry forward to CNNs. The successful training and implementation of CNN architectures have been a landmark in the field of Computer Vision and Machine Intelligence. On the ImageNet classification challenge, widely regarded as the Olympics of Computer Vision, different CNN architectures have brought down the error rate from 28% in 2011 to 2.9% in 2016. In comparison, humans have an error rate of 5.1% (Russakovsky *et al.*, 2015). An understanding of the convolution operation and its use as a feature extractor is necessary to understand the significance of convolutional neural networks.

The Convolution Operation

Mathematically, a convolution is an integral that expresses the amount of overlap of one function g as it is shifted over another function f. Convolution of two functions f and g is given by (Weisstein, 2017):

$$f * g \equiv \int f(\tau)g(t-\tau)d\tau \tag{4.12}$$

where the symbol $f * g$ denotes convolution of f and g. In the context of image processing, a convolution is the dot product of a convolution matrix or kernel G being shifted over an input image matrix F. The kernel or filter is aligned with a portion of the image having the same size as the kernel itself, all the values in the corresponding positions are pairwise multiplied and finally added to get a single value as the output. This process can be clearly understood from Figure. 4.4.

It can be intuitively understood that the convolution output will have a higher value when the overlapped region of the image is very similar to the convolution kernel. As

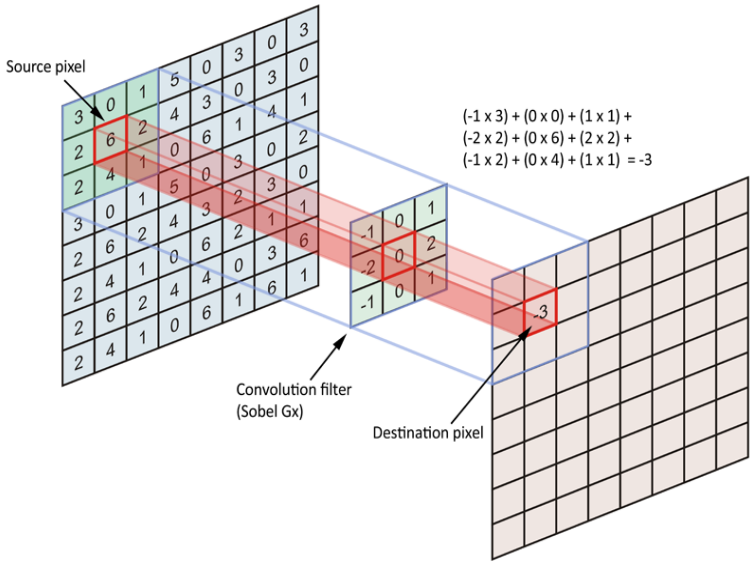


Figure 4.4: Convolution with a Sobel filter (Hartley, 2014)

a result, filters can be said to identify particular visual "elements" of an image (Zeiler and Fergus, 2014). The advantage of using convolutions for detecting these features of the image is positional invariance. Since one filter is slided over the entire image, it will detect the feature wherever it is located in the image. A collection of filters which work on the same input is grouped together and named a convolutional layer. The reason why CNNs work so well is because they function as automated hierarchical feature extractors. The first convolution layer accepts the image itself as the input. Later layers, on the other hand, take the output activation of the previous layer as its input. This forms a hierarchical structure.

CNN Components

Apart from a convolutional layer, CNNs stack pooling layers and fully connected layers to construct a complete architecture. A pooling layer is a downsampling operation performed on the input. This is done for two reasons: Reducing the number of parameters in the network and also to counter overfitting. It accepts two hyperparameters as input: the filter size and the stride. These parameters decide the extent of downsampling to

be performed. For example, a [2×2] pooling size with stride 2 will reduce the spatial resolution of the input by exactly half. The most common pooling operation performed is max-pooling, where the maximum value in a filter-size space in the input is given as the output.

Fully connected layers in CNNs function in the same way as in ANNs. They are generally used as the final layers of a CNN architecture to map the features extracted from convolutions to the corresponding class score outputs. Stacking convolutional layers with pooling, fully-connected layers, etc., in different arrangements form a CNN architecture. Coming up with the best architecture and hyperparameters for the specific problem at hand is the core research area when working with neural networks.

4.3 Proposed Methodology

The proposed method comprises of two stages, namely: pre-processing stage that includes image noise removal and retinal layer segmentation, followed by IRC segmentation stage. OCT scans are inherently affected by speckle noise, which influences the performance of automated IRC segmentation methods. The initial pre-processing module suppresses speckle noise and improves the IRC segmentation performance. The two methodological stages are described below.

4.3.1 Pre-processing: Denoising and Sub-retinal Layer Segmentation

OCT scans contain varying degrees of speckle noise as in other medical imaging modalities like ultrasound. The work in (Sudeep *et al.*, 2016*a*) experimentally showed that OCT images with presence of noise can be approximated with a Gamma distribution. Motivated by this work, approach given in (Sudeep *et al.*, 2016*b*) is implemented as Unbiased fast non local means (UFNLM) to denoise the OCT image data set. First,

each input image is denoised with a Fast NLM (FNLM) method proposed in (Froment, 2014) followed by the subtraction of bias due to speckle. Here, the bias term is estimated from the noisy image using the maximum likelihood method described in (Sudeep *et al.*, 2016*a*,*b*),

$$\hat{B}_e = \hat{\rho}_{\text{ML}} \cdot \hat{\beta}_{\text{ML}} \tag{4.13}$$

where $\hat{B}_e$ is the estimated bias and $\hat{\rho}_{\text{ML}}$, $\hat{\beta}_{\text{ML}}$ are the estimated shape and scale parameters (of the Gamma distribution) from the noisy image. Since the shape and scale parameters can be assumed to be consistent throughout all regions in an input OCT image, they can be estimated from any homogeneous image region (where the underlying intensity is constant). If we assume that many such piece-wise homogeneous regions may exist, then ρ and β can also be estimated as (Sudeep *et al.*, 2016*a*,*b*),

$$\hat{\rho}_{\text{ML}} = \text{mode}\left\{\rho_{\text{ML}(i,j)}\right\}, \qquad \hat{\beta}_{\text{ML}} = \text{mode}\left\{\beta_{\text{ML}(i,j)}\right\} \tag{4.14}$$

where, $\hat{\rho}_{\text{ML}}$ and $\hat{\beta}_{\text{ML}}$ represent the estimated values of ρ and β, respectively, for each pixel at location (i, j) using a neighborhood of size $[m \times n]$. In this work, the neighborhood window size is empirically determined as $[5 \times 5]$. Finally, the FNLM algorithm is executed with the following parameters: search window and similarity window sizes of the filter are $[11 \times 11]$ and $[7 \times 7]$, respectively. The smoothing parameter S_σ is empirically estimated across vendors. The estimated bias parameter $\hat{B}_e$ and smoothing parameter S_σ for images acquired with different vendors are reported in Table 4.1.

Table 4.1: Automatically estimated bias value '$\hat{B}_e$' and smoothing parameter 'S_σ' on different vendor OCT scans considered in this study for speckle noise reduction.

Vendor	Smoothing Parameter (S_σ)	Bias Value ($\hat{B}_e$)
Cirrus	10	15
Nidek	15	26
Spectralis	10	32
Topcon	10	9

Following the image denoising process, sub-retinal layer segmentation provides in-

formation regarding morphology variations and thickness of the layers. This information is utilized to refine the IRC segmentation results. In this Chapter, the Iowa Reference algorithm (Li *et al.*, 2006; Garvin *et al.*, 2009; Abramoff *et al.*, 2010) is used to segment 11 different retinal layers. As intra-retinal cysts are prone to occur in the retinal area enclosed by ILM to BMEIS layers, the OCT B-scans are cropped by delimiting this region as the ROI.

After delimiting the ROI of the OCT scans, contrast enhancement is performed using Contrast-limited adaptive histogram equalization (CLAHE) (Zuiderveld, 1994) to increase the intensity difference between the cystic and non-cystic regions. Kernel size of $(\frac{1}{8})^{\text{th}}$ of the image width and height and clip limit of 0.01 is used for CLAHE.

4.3.2 Cyst Segmentation: FCN Model

The FCN model (Long *et al.*, 2015) accepts denoised sub-retinal layer regions of OCT images as input and produces a prediction score matrix that is then used to construct a binary output mask corresponding to segmented cyst regions. Proposed FCN model is inspired by Ronneberger's U-Net architecture (Ronneberger *et al.*, 2015) that captures both local and global features from an input image to construct an accurate segmentation map. While global features indicate the exact location and relative size of the cystic region, the local features determine the exact cyst boundaries.

The proposed architecture follows a multi-stage two-phase approach. The first phase performs convolutions in five stages. Every stage has two convolutional layers followed by a down-sampling layer with max-pooling operation to garner a larger receptive field. The second phase aims to revert the activations of the first phase to the original resolution. This up-sampling is performed with trainable deconvolution layers in 4 stages, thus enabling end-to-end predictions regarding the location of IRCs.

For accurate segmentation both local and global features are taken into consideration. Thus, the features extracted from the previous stages, in phase one, are forwarded

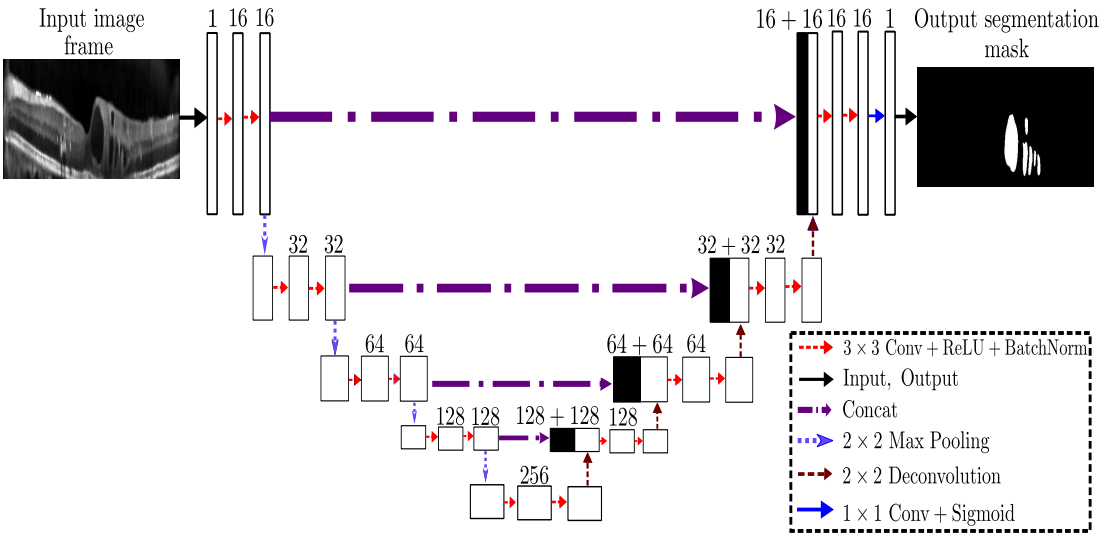


Figure 4.5: The proposed network architecture. The FCN model comprises of 18 convolutional layers with the number of kernels varying from 16 to 256.

and concatenated with the features in the corresponding output of up-sampled stages in phase two. The advantage of this architecture is that every pixel is considered as an individual training sample with per-pixel back-propagation error. This process significantly increases the size of the training data set, thereby avoiding model over-fitting.

The proposed network architecture is shown in Figure. 4.5. The model is designed to process images of size $[256 \times 512]$. The convolution filter width and height ($[k_w \times k_h]$) is chosen as $[3 \times 3]$ across the entire network. This filter size is motivated by prior work (Simonyan and Zisserman, 2014) to ensure discriminative feature learning from pixel neighborhood while ensuring low parametrization when compared to larger filter sizes.

In the analysis/down-sampling phase, there are two $[3 \times 3]$ convolutions before a $[2 \times 2]$ max-pooling layer which reduces the resolution of the image exactly by half. All convolutions are followed by a ReLU activation function and batch normalization operation. The addition of batch normalization (Ioffe and Szegedy, 2015) speeds up the network training and convergence process. The number of filters (N) in each convolutional layer also doubles after every stage.

The second phase of the network up-samples the activations using deconvolution

or fractionally-strided convolution. It follows a path symmetric to the down-sampling path but the pooling operations are replaced with deconvolutions. The final layer in the network is a [1×1] convolution layer. The loss function of the network is computed by pixel-wise sigmoid activation function over the output of the final [1×1] layer with the binary cross entropy loss, which maps the input to a probability map having the same dimensions as the input image. The loss function of the network for a single input image with n_{out} pixels is given in (4.15),

$$BCL_{net} = -\sum_{i=1}^{n_{\text{out}}} \left(t(i) \log\left(\sigma(i)\right) + \left(1 - t(i)\right) \log\left(1 - \sigma(i)\right)\right) \tag{4.15}$$

where $t(i)$ is the actual binary output (target) and $\sigma(i)$ is the predicted binary output for the pixel i. The sigmoid activation function $\sigma()$ applied on weighted sum of the inputs and the final output y are given in (4.16).

$$\sigma(i) = \frac{1}{1 + e^{-y_i}}, \qquad y_i = \sum_{j=1}^{n} x_j w_{ji} \tag{4.16}$$

Here, *n* signifies the number of neurons present in the layer just before the sigmoid activation. Hence, the model output represents a pixel-wise likelihood of being cystic or non-cystic. The resulting binary image mask is then compared with the actual target t cyst mask, and with the aid of the loss function BCL_{net}, model weights are updated such that the loss is minimized for subsequent epochs. Thus, the FCN model is trained.

The proposed model (see Figure. 4.5) is designed with 18 total convolutional layers (denoted by red arrows in Figure. 4.5) across the two phases. The number of filters (N) for the 2 starting layers are assigned as 16. After each pooling layer, the number of filters are doubled for the next 2 convolutional layers. There are 4 pooling layers, making the maximum number of filters in a convolutional layer to be 256.

In the up-sampling phase, the number of convolutional filters is reduced by half after every up-sampling operation. Similar to the first phase, this is also performed 4 times, leaving the last 2 convolutional layers with $N = 16$ filters each. This is followed by a [1×1] convolutional layer with a filter to produce a [256×512] sized output image

mask, to match the input dimensions.

4.4 Experimental Setup

To evaluate the performance of the proposed method, experiments are conducted on the benchmarked OPTIMA cyst segmentation challenge dataset (OPTIMA, 2015). The proposed FCN architecture is implemented in Keras 1.0 (Chollet, 2015) with Tensorflow back-end on a workstation with a 64-bit Ubuntu operating system, Intel Xeon Processor E5-2600 (Intel, Mountain View, CA), solid state hard drive, 128 GB of RAM and the NVIDIA Quadro K2200 GPU with 4GB dedicated memory.

4.4.1 Data

The OPTIMA cyst challenge dataset contains OCT scans with cystoid macular edema obtained using four different vendors, namely Zeiss Cirrus, Nidek, Spectralis Heidelberg and Topcon (OPTIMA, 2015). This dataset consists of separate training and test subsets. Each OCT volume is acquired over 6×6 mm of the macula and foveal center from subjects with CME. Table 2.1 shows the dataset description. A sample OCT B-scan frame from each vendor is shown in Figure. 2.3.

4.4.2 FCN Model Training and Hyper-parametrization

Model hyper-parameters need to be optimally tuned for the final objectives. In the context of FCNs, the most important parameters/hyper-parameters that need to be optimized include: the number of weights and biases in the network, the number of layers, the number of filters/kernels in each layer and the learning rate of the model. The best combination of these parameters is identified by performing grid-search using the hold-out method (Kohavi *et al.*, 1995). Finally, the test data set is used for trained

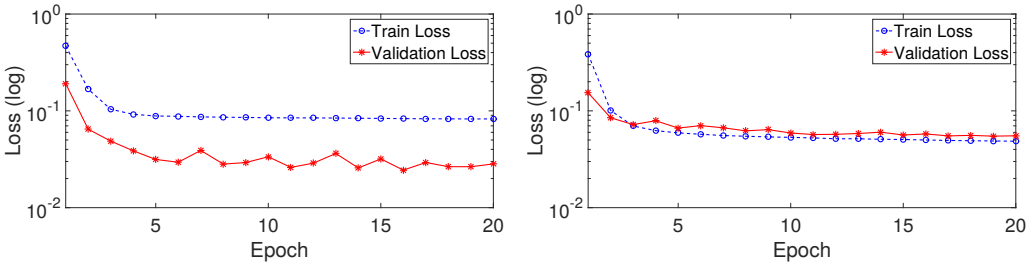
Table 4.2: Different FCN architectures formed by varying number of layers.

# Layers	Starting Filter Size ($k_w \times k_h \times N$)	*Depth*	# Parameters
4	$3 \times 3 \times 64$	0	112,513
6	$3 \times 3 \times 64$	1	367,809
10	$3 \times 3 \times 64$	2	1,866,817
14	$3 \times 3 \times 64$	3	7,707,457
18	$3 \times 3 \times 64$	4	31,054,145

model performance analysis. Several experiments have been designed to validate the FCN model architecture as described in section 4.3.2. As U-net model, in the proposed architecture two convolutional layers are always followed by a max-pooling layer or a deconvolution layer. Keeping this feature constant and the number of filters as described in (Ronneberger *et al.*, 2015), the number of layers can be varied for different architectures. An overview of these architectures is presented in Table 4.2. Here, *Depth* signifies the number of max-pooling layers, *parameters* signify the number of weights and biases in the network, and *layers* specify the convolutional layers.

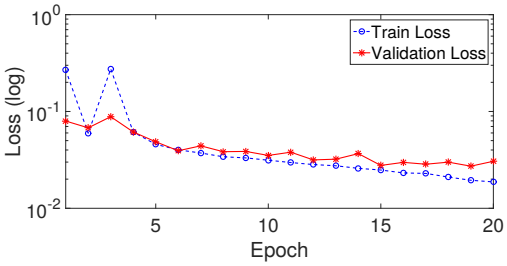
For experimental evaluation of the FCN parameters on the various architectures, one volume from each vendor is extracted from the training set to act as the validation set, i.e., the training and validation sets contain 11 and 4 volumes, respectively. First, these 3D OCT volumes are separated into 2D B-scans, which constitute training data of 1243 B-scans and validation data of 433 B-scans across all the vendors. As discussed in the section 4.3.1, only the ROI between the ILM and BMEIS sub-retinal layers are considered for intra-retinal cyst segmentation. All the images are then resized using bilinear interpolation to a standard resolution of $[256 \times 512]$. After resizing, the OCT scans are normalized to zero mean and unit variance.

The FCN network is trained with the binary cross-entropy loss function (refer (4.15)). The Adam optimizer (Kingma and Ba, 2014), which is a variant of stochastic gradient descent, is used to update the weights. The default hyper-parameters used for this optimizer are: $\beta_1 = 0.9$, $\beta_2 = 0.999$, $\epsilon = 1 \times 10^{-8}$. The learning rate is empirically chosen as 3×10^{-4} based on a short random search.

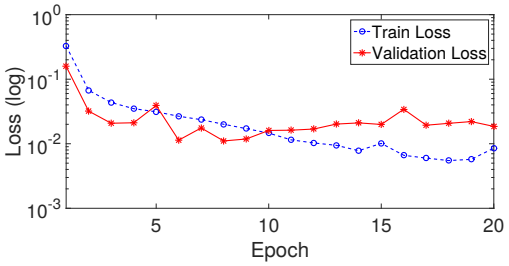


(a) Layers $= 4$, *Depth* $= 0$

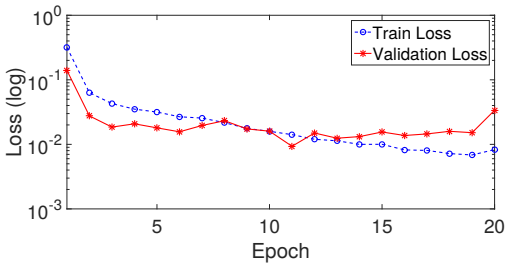
(b) Layers $= 6$, *Depth* $= 1$



(c) Layers $= 10$, *Depth* $= 2$



(d) Layers $= 14$, *Depth* $= 3$



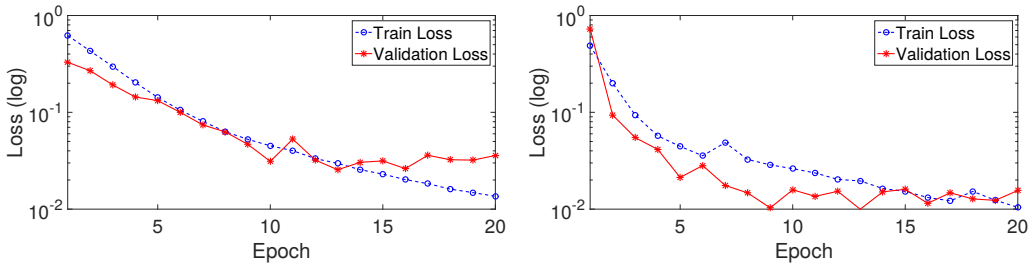
(e) Layers $= 18$, *Depth* $= 4$

Figure 4.6: Loss Value vs. Epoch graphs for FCN model *Depth*: [0, 1, 2, 3, 4] architectures after model training. Blue: Training Loss. Red: Validation Loss. (Plotted in Log Scale)

Initial network weights are assigned according to the *He initializer* (He *et al.*, 2015). No other pre-trained weights are used in the process. The FCN network is trained from scratch completely on the train data set. Nevertheless, the train data comprises of relatively less number of samples to train a full-scale FCN. Hence, data augmentation is applied to generalize the model and reduce over-fitting. Several data augmentation schemes are evaluated for their impact on training loss and model convergence. These include image gradient enhancement, edge enhancement, brightness and contrast variations. However, it is observed that augmented data generated by horizontal flipping, random shear, height, width and zoom shifts are best for IRC segmentation tasks. This observation is intuitive since domain knowledge suggests cysts to have variable shapes, structures and orientations but similar appearance with respective to neighboring tissues. Thus, domain-specific data augmentation process is key to a generalizable FCN model. The data transformations for augmentation are performed on-the-fly during training, thereby alleviating storage concerns. The proposed model is trained up to 120 epochs, beyond which, no further change in loss function is observed. Total train samples presented to the network after data augmentation for final training is $120 \times 1243 = 149,160$.

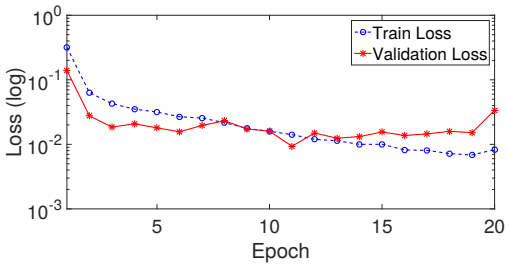
Figure. 4.6 show the training process for various FCN model architectures. *Depth* 0, 1, 2 models (see Figures. 4.6(a), 4.6(b) and 4.6(c)) exhibit limited learning capabilities and the training loss does not reduce significantly. *Depth* 3, 4 models (see Figures. 4.6(d) and 4.6(e)) can be further investigated for their receptive fields. The receptive field is defined as the effective area of the original input image covered by a convolution filter (Luo *et al.*, 2016). If the convolution filter size is kept the same throughout the network, the receptive field of a filter increases after every pooling operation. This is because the resolution of the image is reduced, but the size of the filter remains the same.

Even though both *Depth* 3 and *Depth* 4 models converge to a lower training loss, the latter one is chosen for its higher receptive field, which can account for more fea-



(a) Layers $= 18$, *Depth* $= 4$, Starting Filter Size $N = 16$

(b) Layers $= 18$, *Depth* $= 4$, Starting Filter Size $N = 32$



(c) Layers $= 18$, *Depth* $= 4$, Starting Filter Size $N = 64$

Figure 4.7: Loss Value vs. Epochs graphs for *Depth 4* architecture with different starting filter sizes. Blue: Training Loss. Red: Validation Loss. (Plotted in Log Scale)

tures. However, in *Depth* 4 model the rate of change in loss function demonstrates some over-fitting trends. Keeping this *Depth* 4 as constant, the number of parameters can be varied by changing the number of filters in each layer, thereby reducing over-fitting. Table 4.3 shows the number of parameters in the different architectures formed at *Depth* 4 by varying the starting filter size, and Figure. 4.6(e) shows their training processes, respectively.

It can be observed that both $N = 32$ and $N = 16$ starting filter sizes converge to low training losses with significantly less over-fitting trends. Thus, the model with a fewer number of parameters is chosen as the final architecture, because of higher generalizability and low over-fitting trends. The proposed *Depth* 4 FCN model requires 338 seconds to train per epoch (total 1234 samples in the training data), and 0.06 seconds per test image (total 909 samples in the test data) for prediction.

Table 4.3: Architectures designed by varying number of filters in *Depth* 4

Starting Filter Size ($k_w \times k_h \times N$)	# Parameters
$3 \times 3 \times 64$	31,054,145
$3 \times 3 \times 32$	7,771,297
$3 \times 3 \times 16$	1,946,705

Table 4.4: Mean (standard deviation) of precision and recall of the proposed method on different vendor OCT scans.

Vendor	G_1			G_2			$G_1 \cap G_2$		
	Precision	Recall	Dice	Precision	Recall	Dice	Precision	Recall	Dice
Cirrus	0.71 (0.21)	0.61 (0.38)	0.62 (0.35)	0.71 (0.21)	0.61 (0.38)	0.62 (0.35)	0.67 (0.18)	0.66 (0.40)	0.63 (0.35)
Nidek	0.83 (0.02)	0.75 (0.23)	0.77 (0.15)	0.81 (0.01)	0.75 (0.24)	0.76 (0.14)	0.74 (0.03)	0.80 (0.25)	0.75 (0.11)
Spectralis	0.82 (0.09)	0.62 (0.09)	0.71 (0.08)	0.84 (0.09)	0.64 (0.11)	0.72(0.10)	0.79 (0.09)	0.69 (0.10)	0.74 (0.09)
Topcon	0.82 (0.09)	0.70 (0.17)	0.75 (0.13)	0.82 (0.08)	0.73 (0.14)	0.77 (0.11)	0.76 (0.11)	0.78 (0.13)	0.77 (0.12)
Overall	**0.79 (0.12)**	**0.66 (0.22)**	**0.71 (0.20)**	**0.80 (0.12)**	**0.67 (0.22)**	**0.72 (0.19)**	**0.74 (0.12)**	**0.73 (0.23)**	**0.72 (0.19)**

4.5 Results and Analysis

Segmentation results of the proposed method are compared with manually graded ground-truth provided with dataset images (ascertained by two trained ophthalmologists). Pixel-wise analysis is conducted and precision and recall metrics are computed for each of the OCT volumes in the test dataset against GT from two independent graders (G_1, G_2) and their intersection ($G_1 \cap G_2$). Precision and recall metrics are given in (2.1) and recall (2.2).

To assess the correlation accuracy of the segmentation, Dice coefficient (ref (2.3)) is computed between segmented results of the algorithm and ground truth for each test OCT volume.

Table 4.4 shows the mean precision, recall and Dice coefficient results of the proposed method against G_1, G_2, and $G_1 \cap G_2$. It can be noticed that the proposed method gave highest mean recall rate on scans obtained from Nidek vendor compared to other

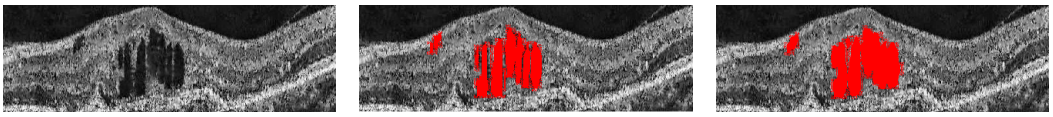
vendors (0.75 on G_1, 0.75 on G_2 and 0.80 on $G_1 \cap G_2$). The precision of the proposed method is stabilized across the three vendors (Nidek, Spectralis and Topcon). In case of Dice coefficient, scans obtained from Nidek and Topcon result in higher correlation on G_1 (0.77 and 0.75, respectively) and G_2 (0.76 and 0.77, respectively) compared to other two vendors (Cirrus and Spectralis). On Cirrus scans, precision, recall and Dice correlation rates are lower compared to other vendors due to high-intensity variations and noise.

Also, observe the variations in the results obtained against G_1, G_2, and $G_1 \cap G_2$, due to inter-observer variability between manually annotated GTs. The scan intensity variations and retinal vessel shadows often increase the difficulty in cyst identification and delineation for human graders, thus accounting for such inter-observer variability. Observe a variable degree of inter-observer variability across vendors for the IRC segmentation task.

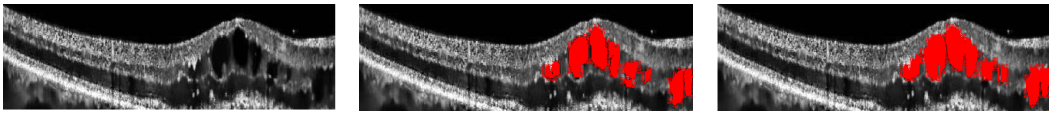
Qualitative results using the proposed segmentation method are shown in Figure. 4.8 and Figure. 4.9. Images in the first column of these figures show original scans from different vendors, the second column show GTs provided by $G_1 \cap G_2$, and the third column show the result of the proposed FCN model, respectively. It is observed that due to intensity variations on the sample Cirrus scan, the boundaries between individual cysts are ill-defined (see Figure. 4.8(a)). Hence, the proposed segmentation method merges individual cysts into larger cystic region due to small intensity variations in the boundaries of cysts (see Figure. 4.8(c)).

In the sample Nidek scan, the proposed method over-segments the cystic structures by merging two cysts into a single large cyst, thereby resulting in false positive (see Figures. 4.8(d) and 4.8(f)). In case of sample Topcon and Spectralis scans (see Figures. 4.8(g) and 4.8(j)), the proposed method finely segments the cystic structures when compared to the other two vendors owing to the clear boundary between the cystic structures (see Figures. 4.8(i) and 4.8(l)).

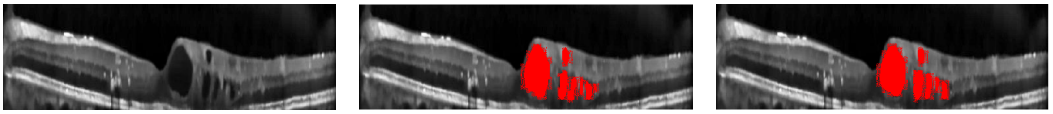
Figure. 4.9 shows the limited segmentation performance results of the proposed



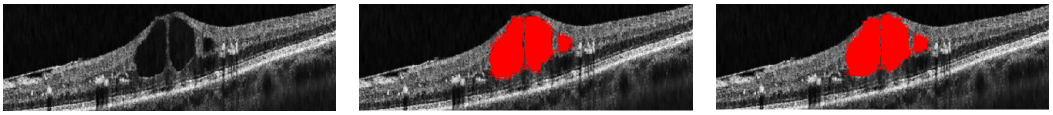
(a) Sample B-scan of Cirrus Vendor (b) Ground truth (c) Segmented output



(d) Original B-scan of Nidek Vendor (e) Ground truth (f) Segmented output



(g) Sample B-scan of Spectralis Vendor (h) Ground truth (i) Segmented output



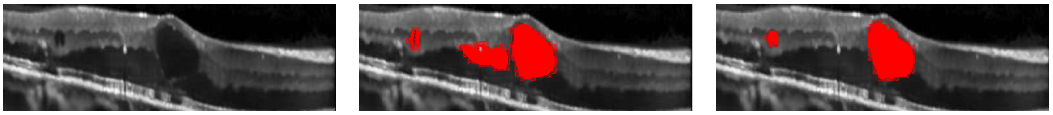
(j) Sample B-scan of Topcon Vendor (k) Ground truth (l) Segmented output

Figure 4.8: Results of proposed method on different vendor scans against the GT from $G_1 \cap G_2$.

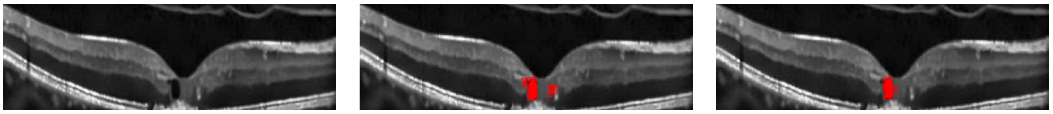
method on different vendor scans. For the sample Cirrus scan (see Figures. 4.9(a), 4.9(b) and 4.9(c)), the proposed method is unable to detect cysts due to low intensity variations between the cysts and neighboring non-cystic regions. In the sample Nidek scan (see Figure. 4.9(d)) the boundary between cysts beside the larger cysts is relatively unclear and the cyst merges with the ONL. Here, the model ignores the cyst by misclassifying it as ONL layer (see Figures. 4.9(e) and 4.9(f)).

In the sample Spectralis scan, the grader had marked a non-cystic region as cystic region on right to the foveal center. Here, the proposed method shows its robustness to such instances by not segmenting this region (see Figures. 4.9(g), 4.9(h) and 4.9(i)). The proposed method segments 6 FPs in the sample Topcon scan due to intensity variations (see Figures. 4.9(j), 4.9(k) and 4.9(l)). These limiting results demonstrate the vital role of scan intensity variations towards OCT cyst segmentation tasks.

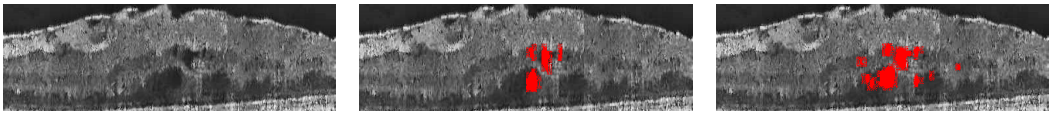
(a) Sample B-scan of Cirrus Vendor (b) Ground truth (c) Segmented output



(d) Sample B-scan of Nidek Vendor (e) Ground truth (f) Segmented output



(g) Sample B-scan of Spectralis Vendor (h) Ground truth (i) Segmented output



(j) Sample B-scan of Topcon Vendor (k) Ground truth (l) Segmented output

Figure 4.9: Limiting segmentation results by the proposed method for each vendor, against the GT from $G_1 \cap G_2$

The Dice coefficient using the proposed method when compared with the published OPTIMA Cyst Segmentation Challenge results (OPTIMA, 2015) are shown in Table 4.5. Here, the method proposed by de Sisternes *et al.* (de Sisternes *et al.*, 2015), uses 34 handcrafted features to train a machine learning model. While designing appropriate features manually across imaging vendors may be manually tedious and insufficient task, it is observed that the designed features unable to encompass all the OCT scan variations.

In (Venhuizen *et al.*, 2016), a cascaded CNN working on multiple resolutions is used to segment the IRC. This method requires training of 3 different networks independently. Moreover, the selection of patch sizes is found to impact the final segmentation results. Another method involving two-stage graph-cut segmentation approach is proposed by Oguz *et al.* (Oguz *et al.*, 2016), where the cost function for the graph-

Table 4.5: Comparison mean (standard deviation) of Dice coefficient of the proposed method against the results of OPTIMA cyst segmentation challenge on all vendor OCT volumes (Test set).

Method	G_1	G_2	$G_1 \cap G_2$
de Sisternes *et al.* (de Sisternes *et al.*, 2015)	0.68 (0.15)	0.67 (0.14)	0.69 (0.15)
Venhuizen *et al.* (Venhuizen *et al.*, 2016)	0.61 (0.19)	0.60 (0.19)	0.59 (0.19)
Oguz *et al.* (Oguz *et al.*, 2016)	0.60 (0.15)	0.59 (0.15)	0.60 (0.14)
Esmaeili *et al.* (Esmaeili *et al.*, 2016)	0.55 (0.27)	0.55 (0.27)	0.55 (0.28)
Haritz *et al.* (Gopinath and Sivaswamy, 2016)	0.23 (0.15)	0.23 (0.15)	0.23 (0.15)
Proposed Method	**0.71 (0.20)**	**0.72 (0.19)**	**0.72 (0.19)**

cut algorithm is computed using initial intra-retinal layer segmentation results. Since segmentation of inner retinal layers is more prone to error on pathological images than on normal ones, the cost function is found to be impacted significantly. Also, Haritz *et al.* (Gopinath and Sivaswamy, 2016) initially identified candidate cyst regions using a center-surround difference technique. In this method, local descriptors are extracted from cystic regions and classified using random forest classifiers. Nonetheless, this rule-based method is found to fail while segmenting small cyst regions.

Esmaeili *et al.* (2016) proposed an automated IRC segmentation method using 3D curvelet transform and K-SVD dictionary learning approaches. However, this method is proposed and validated only on Spectralis vendor OCT scans of the challenge dataset. Generalizing this method for vendor-independent IRC segmentation might affect the dictionary learning process and performance because of the higher level of noise present in other OCT vendor (Cirrus, Nidek, and Topcon) scans.

The method proposed in this paper uses a pre-processing module to equalize the OCT images across vendors. Next, the FCN model automatically captures both micro and macro-level features for better characterization of cystic structures. Thus, the proposed method outperforms reported challenge results with highest Dice rate (0.71 on G_1, 0.72 on G_2 and 0.72 on $G_1 \cap G_2$) with an improvement of 3% on ground truth provided by G_1 and 5% on G_2 compared to results reported in de Sisternes *et al.* (de Sisternes *et al.*, 2015). Additionally, slight variations are observed in the segmentation

Table 4.6: Mean Dice coefficient of cyst segmentation results on Duke DME dataset.

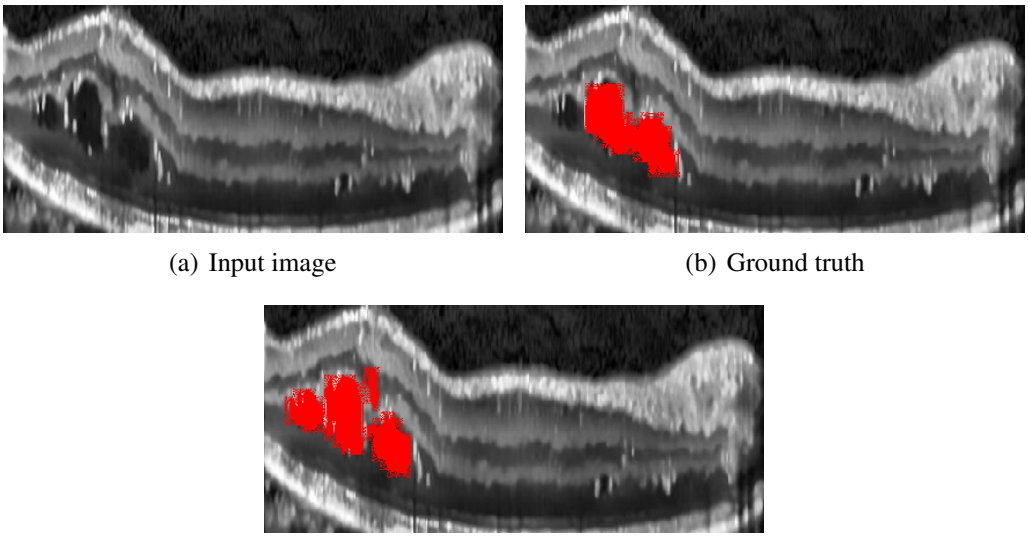
Method	Dice coefficient on $G_1 \cup G_2$
Proposed Method	0.53
KR + GTDP (Chiu *et al.*, 2015)	0.53

results for the comparative methods on the ground-truth provided by both graders due to inter-observer variabilities.

4.6 Validation on the Duke DME Dataset

Additional experiments are conducted on the publicly available Duke DME dataset (Chiu *et al.*, 2015) to validate the robustness of the proposed method for cross-vendor OCT scans. The Duke DME dataset contains a total of 110 B-scans with ground truth for both retinal layer boundaries and fluids. These images are pre-processed using UFNLM+CLAHE and the parameters specific to Spectralis vendor scans as mentioned in section 4.3.1 of this thesis. Retinal ROI (ILM-BMEIS) is obtained from the ground truth provided with the dataset, and the B-scans are cropped and resized to the resolution of $[256 \times 512]$ pixels. The pre-trained *Depth* 4 FCN model is tested on these B-scans. The results are given in Table 4.6. As compared with KR + GTDP method (Chiu *et al.*, 2015), the proposed FCN method gives mean DC of 0.53 on ground truth obtained from union of both graders, noteworthy that our model is not explicitly fine-tuned or trained on Duke DME dataset.

Figure. 4.10 shows the proposed segmentation method result on a B-scan from Duke DME dataset. In Figure. 4.10(b) ground truth provided by both graders missed a cyst on the left extreme of the B-scan and the marked cystic structures are merged even though the boundaries between cysts are clearly separated by the hard exudates (see Figure. 4.10(b)). In Figure. 4.10(c) the proposed segmentation method robustly segments all cysts without significant overlap between the consecutive cysts boundaries.



(a) Input image (b) Ground truth

(c) Segmented output

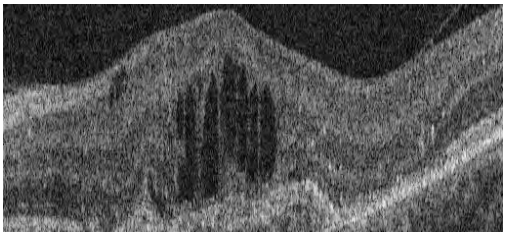
Figure 4.10: Result of the proposed segmentation method on Duke DME dataset B-scan: (a) Input image, (b) Ground truth, (c) Segmented output.

4.7 Effect of Pre-processing on Intra-Retinal Cyst Segmentation

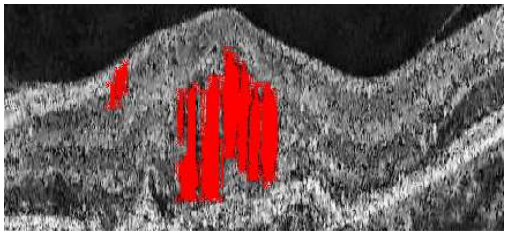
To evaluate the effect of pre-processing on intra-retinal cyst segmentation, automated cysts segmentation using the proposed FCN model analyzed using noisy, raw OCT scans and UFNLM denoised scans, respectively. As described in section 4.3.1 of this Chapter, we have performed denoising followed by contrast enhancement to improve the quality of OCT scans. For raw OCT scans analysis, the optimal model (*Depth* 4) is trained using raw OCT scans with the procedure as described in section 4.4.2 of this Chapter. Table 4.7 shows the mean Dice coefficient of segmentation results on the overall dataset. It can be inferred from the Table 4.7 that denoising (with UFLNM) improves the segmentation results by 50%. UFNLM method removes bias introduced after denoising image with FNLM filter, which in turn provides better segmentation results. Experiments are also conducted to analyze the effect of contrast enhancement technique on segmentation result. We observe 4% improvement on UFNLM + CLAHE compared to UFNLM, respectively (refer Table 4.7). Figure. 4.11 shows the segmentation results

Table 4.7: Mean Dice coefficient of cyst segmentation using different pre-processing techniques compared against ground truth provided by $G_1 \cap G_2$.

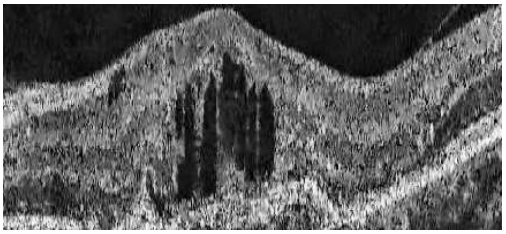
Pre-processing	Dice coefficient
Raw data	0.18
UFNLM	0.68
UFNLM + CLAHE	0.72



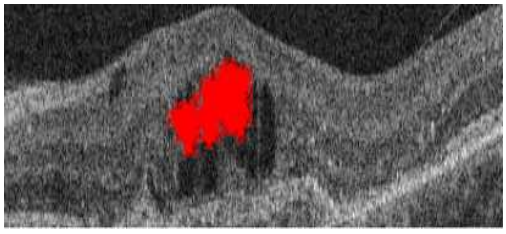
(a) Noisy raw image



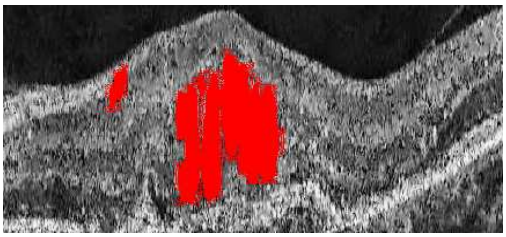
(b) Ground truth



(c) UFNLM + CLAHE pre-processed image



(d) Segmentation result on raw image



(e) Segmentation result on UFNLM + CLAHE pre-processed image

Figure 4.11: Result of the segmentation with and without pre-processing: (a) Noisy raw image, (b) Ground truth, (c) UFNLM+CLAHE pre-processed image, (d) Segmentation result on raw image, (e) Segmentation result on UFNLM + CLAHE pre-processed image.

on raw image and denoised images, respectively.

Proposed UFNLM + CLAHE pre-processing technique significantly improves the quality of the image (see Figure. 4.11(c)) by reducing the noise and enhancing the con-

trast between cyst and non-cystic pixels compared to the raw and noisy OCT image (see Figure. 4.11(a)). It can be seen from Figure. 4.11(d) that the proposed model partially segments few cystic structures on the raw image. Noise and poor contrast in the raw image (see Figure. 4.11(a)) significantly affects the distinguishing capability of the FCN model for IRC segmentation (see results in Figure. 4.11(d)) when compared to the UFNLM + CLAHE pre-processed image shown in Figure. 4.11(e).

4.8 Summary

This Chapter presents an FCN model-based vendor independent IRC area segmentation technique. The FCN model is customized for IRC area segmentation by utilizing denoised retinal OCT images to train it from scratch. Sensitivity analysis of the model hyper-parameters demonstrates that deeper networks exhibit better feature learning capabilities than shallower networks while higher receptive fields induce higher training losses that may lead to model over-fitting. However, deeper networks results in large number of computation parameters. Next Chapter address this issue by utilizing depthwise separable convolution filters instead of regular convolution filters.

CHAPTER 5

DEPTHWISE SEPARABLE CONVOLUTIONAL NEURAL NETWORK ARCHITECTURE FOR AUTOMATED SEGMENTATION OF INTRA-RETINAL CYSTS FROM CROSS VENDOR OCT SCANS

In previous Chapter, a pixel-level segmentation system of IRCs using FCN model was proposed, where a model based denoising technique and contrast enhancement improved the initial quality of OCT scans. Next, several varieties in network architectures were experimentally evaluated to obtain the optimal FCN hyperparameters. It was showed that although deeper networks with higher receptive fields improve IRC segmentation accuracy, they also increased the model computation complexity proportionally and lose generalizability with significantly deep networks.

In this Chapter[5], we address this limitation of increase in computational complexity for a deep neural network by introducing a novel depthwise separable encoder decoder architecture that is efficient for vendor independent IRC segmentation tasks. This Chapter has three key contributions. First, a depthwise separable convolutional neural network (DSCN) architecture comprising of an encoder-decoder pair is introduced that significantly reduces the computational parameters when compared to regular FCNs models. Second, the impact of an efficient activation function for reducing the effect of

[5]The work described in this Chapter has been submitted for possible publication as: **G. N. Girish**, Banoth Saikumar, Sohini Roychowdhury, Abhishek R. Kothari, and Jeny Rajan, "Depthwise Separable Convolutional Neural Network Model for Intra-Retinal Cyst Segmentation" in *Biomedical Signal Processing and Control*, Elsevier, 2018. **(Under Review)**

vanishing gradient in deeper networks is analyzed. Third, hyper-parameter analysis is presented to achieve optimal architecture for IRC segmentation.

5.1 Data and Method

The proposed method comprises of the following two stages: pre-processing and retinal layer segmentation performed for enhancement and region of interest limitation, followed by cyst segmentation using the proposed DSCN model. The publicly available OPTIMA cyst segmentation challenge dataset obtained from four different vendors namely Cirrus, Spectralis, Topcon, and Nidek is used for assessment of the proposed method. Detailed dataset description is provided in section 2.2.1 of Chapter 2.

5.1.1 Pre-processing: De-noising and Retinal Layer Segmentation

OCT scans are affected by speckle noise, which degrades the quality of the scans. Speckle noise in OCT scans can be approximated to Gamma Distribution (Sudeep *et al.*, 2016*c*). To reduce the effect of speckle noise, UFNLM filter is employed in this Chapter as described in section 4.3.1 of Chapter 4. Denoising is performed on OCT scans using FNLM filter then bias is deducted from the resultant scans. The estimated bias $\hat{B}_e$, shape $\hat{\rho}_{\text{ML}}$ and scale $\hat{\beta}_{\text{ML}}$ parameters of the Gamma distribution are calculated using Eqs. (4.13),(4.14).

Although denoising improves the quality of OCT scans, scan intensity variations among different vendors can lead to poor contrast. To further enhance the quality of the scans CLAHE algorithm is employed with threshold of $(\frac{1}{8})^{\text{th}}$ of maximum intensity as mentioned in section 4.3.1 of Chapter 4.

Typically, retinal OCT scans contain the retina along with vascular regions above and below the retina. Also, IRCs are found accumulated within the ILM and BMEIS layer. Thus, following OCT image contrast enhancement, the region of interest for IRC

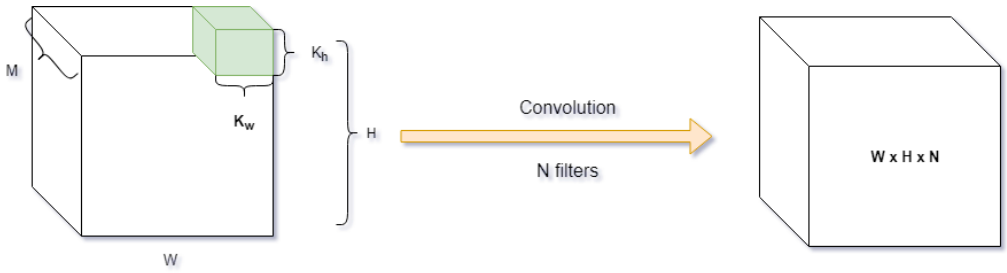


Figure 5.1: Regular convolution

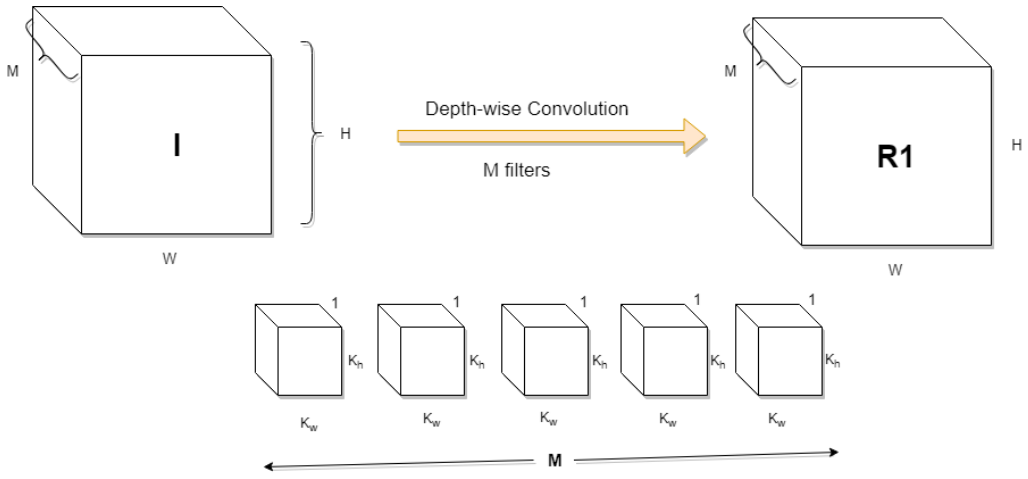
segmentation limited by utilizing ILM-BMEIS boundaries which are obtained using OCT-Explorer tool (Li *et al.*, 2006; Garvin *et al.*, 2009; Abramoff *et al.*, 2010).

5.1.2 Network Architecture for Intra-retinal Cyst Segmentation

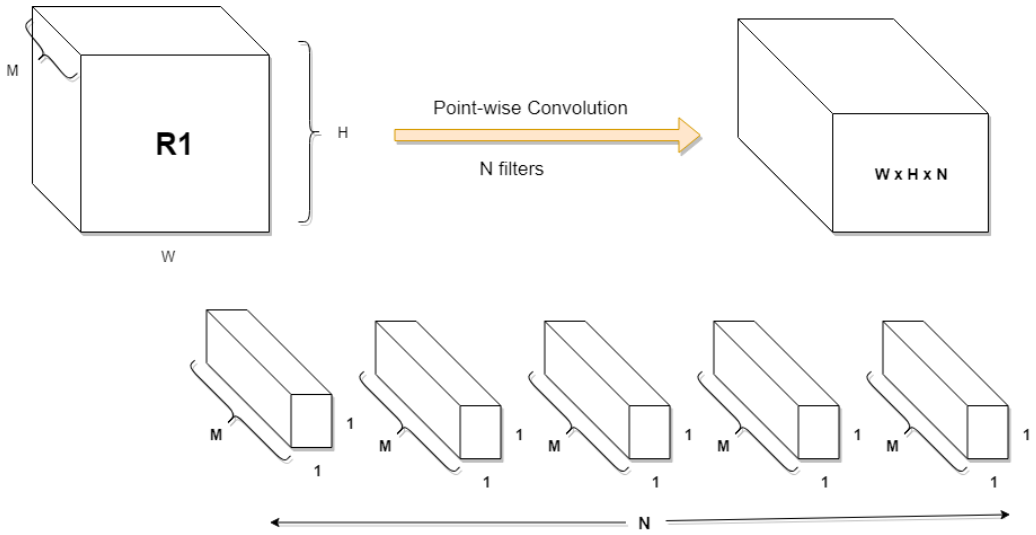
The proposed architecture is inspired by Google's Xception network. Xception network was proposed for classification of larger images in the ImageNet dataset (Chollet, 2016). Our proposed architecture is entirely based on depthwise separable convolution layers. A depthwise separable convolution is a superlative version of inception modules. These are performed in two steps– initially, separable convolutions in spatial domain over each channel of an input, followed by $[1 \times 1]$ point-wise convolution. On the other hand, a regular fully convolution layer takes an image with size $[W \times H \times M]$ as an input and produces output of feature maps with size $[W \times H \times N]$ by convolving filters with kernel size $[k_w \times k_h]$, as shown in Figure. 5.1. Total parameters (T_{param}) of a regular fully convolutional layer can be calculated as (Paul-Louis),

$$T_{param} = W \times H \times M \times k_w \times k_h \times N. \tag{5.1}$$

Figure. 5.2 shows depthwise separable convolution stages. Separable convolution stage takes an image of size $[W \times H \times M]$ as an input and performs spatial channel-wise convolutions with kernel size $[k_w \times k_h]$. Finally, it produces feature maps of size $[W \times H \times M]$, as shown in Figure. 5.2(a). The total number of parameters in this stage is



(a) Stage 1: Depthwise convolution



(b) Stage 2: Point-wise convolution

Figure 5.2: Detailed representation of depthwise separable convolutions

given as:

$$T_{sep} = W \times H \times M \times k_w \times k_h. \tag{5.2}$$

The point-wise convolution stage takes the output produced by separable convolution stage of size $[W \times H \times M]$ and performs $[1 \times 1]$ convolution to produce N number of output channels, as shown in Figure. 5.2(b). Total parameters in point-wise stage can be given as:

$$T_{point} = W \times H \times M \times 1 \times 1 \times N. \tag{5.3}$$

Finally, total number of parameters of a depthwise separable convolution layer is given as:

$$T_{depth} = T_{sep} + T_{point}. \tag{5.4}$$

For example, consider an input image with size $[5 \times 5 \times 3]$, filter with kernel size $[3 \times 3]$, and number of output channels as 6. The total number of parameters in the regular convolution layers will be calculated as $5 \times 5 \times 3 \times 3 \times 3 \times 6$ which results in 4050 number of parameters. In case of depth-wise separable convolutional layer– the number of parameters in first stage $T_{sep} = 5 \times 5 \times 3 \times 3 \times 3 = 675$, and in second stage $T_{point} = 5 \times 5 \times 3 \times 1 \times 1 \times 6 = 450$. Finally, total parameters in depth-wise convolutional layer $T_{depth} = 675 + 450 = 1125$. Hence, the depth-wise separable convolution layers are computationally efficient compared to regular convolutional layers.

The proposed architecture is two phased, comprising of an encoder and decoder, with multiple stages in each phase. In the first encoder phase, 4 stages exist with each group of depthwise separable convolutions, batch normalization, activation and down-sampling layer followed by residuals. The down-sampling layer with max-pooling operation is used to reduce the size of the receptive field for extracting features at micro and macro levels.

The second decoder phase reverts the operations of first phase with up-sampling operation to obtain original resolution of input image as output. Features extracted in encoder phase are concatenated with layers of the corresponding stage in decoder phase after up-sampling with same receptive field size, which adds the advantage in recalling past learned features and improves overall segmentation accuracy.

The proposed architecture is shown in Figure. 5.3. An input image of size $[256 \times 512]$ is processed through the network resulting in a same size output with the segmented mask. For convolution and depthwise separable convolutions, filters kernel size is fixed as $[3 \times 3]$ based on the prior work (Simonyan and Zisserman, 2014) to extract

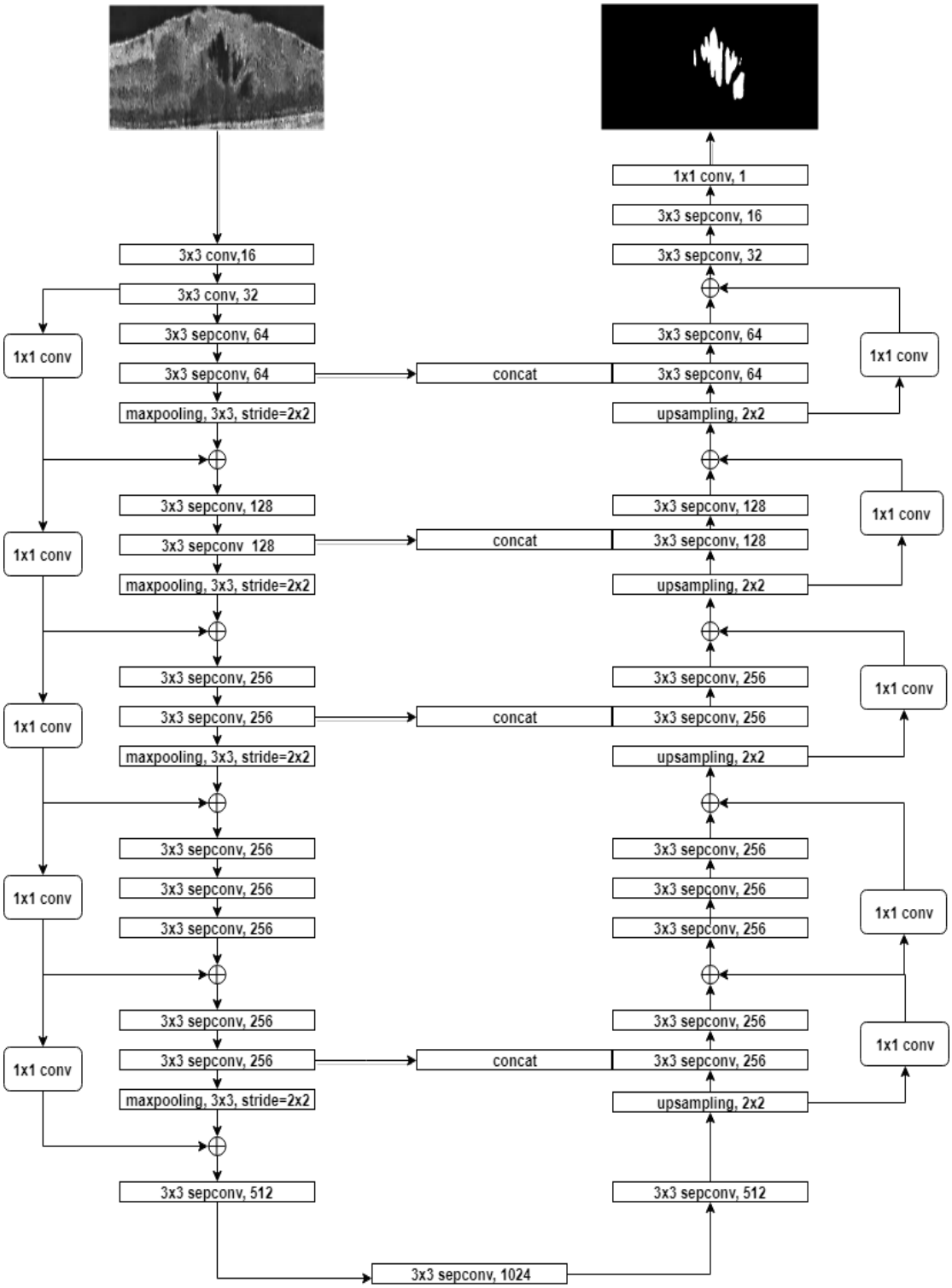


Figure 5.3: Proposed DSCN architecture.

discriminative features while preserving low parametrization. Batch normalization is performed after convolution layer to avoid internal covariate shift. After the batch normalization layer non-linear activation functions are employed.

In deeper networks, gradients accumulation in each update results in very large updates in the weights, and in turn, unstable network. The accumulated gradients at each update may results in overflow and NaN values in weights of the layers. This situation is called as gradient explosion or vanishing gradient problem (Hochreiter, 1998). Most activation functions such as sigmoid, tanh and ReLU are monotonic and their output value increases proportional to input, which results in vanishing gradient problem in deeper networks. Swish activation function is non-monotonic function, and output activations of Swish relies between linear and ReLU (Ramachandran *et al.*, 2017). Swish is unbounded above and bounded below, due to its non-monotonic attribute the output of Swish might decreases even when the input is increased. In case of explosion of gradient occurs then clipping gradient is the best method to prevent such a scenario (Pascanu *et al.*, 2012; Goodfellow *et al.*, 2016).

To prevent vanishing gradient problem in the proposed network, the Swish activation function is employed in all intermediate layers of except output layer. Swish activation function performed better than ReLU activation function in terms of accuracy on many datasets by allowing flow of small gradients (Ramachandran *et al.*, 2017; Arora *et al.*, 2016). In addition we have added skip connections and residuals with $[1 \times 1]$ convolutions to counteract vanishing gradient problem. The Swish activation function is defined as,

$$s(x) = x \times \sigma(x) \tag{5.5}$$

where x is an input tensor to function and $\sigma(x)$ is sigmoid activation function which is given in (4.16):

The final layer in the network is a $[1 \times 1]$ convolution layer with sigmoid activation function. The loss function of the network is binary cross entropy loss (Shore and Johnson, 1980) and it is computed using (4.15).

The proposed optimal model is designed with *Depth* = 4 and contains total 30 convolutional layers (3 regular convolutions and 27 depthwise separable convolutions). Initial 2 layers and final output layer are regular convolutional layers, and remaining layers are depthwise separable convolutional layers as in (Chollet, 2016).

5.2 Experimental Setup

The proposed architecture is implemented in Keras with Tensorflow as backend and experiments are conducted using Google's Colaboratory platform (GoogleColab, 2017). It's a Jupyter notebook environment that requires no setup to use and runs entirely in the cloud.

5.2.1 DSCN Model Training and Hyper-parametrization

The experiments are conducted using OCSC dataset. The training set is divided into train and validation set for hyper-parameter optimization. Several experiments are performed to validate different FCN architectures by making validation set of 433 B-scans (one volume from each vendor) from training set of 1676 B-scans (total 15 volumes).

While designing an optimal CNN architecture, the major factor is to optimize the parameters and hyper-parameters, such as depth of the network, total number of layers in network, number of filters in each layer, kernel size of the filter, and learning rate of the model. Combination of optimal parameters or hyper-parameters can be identified using grid search and short random search. In this work, several experiments are performed by varying depth of the network as shown in Table 5.1. Here, *Depth* refers to the number of max-pooling layers, *layers* signify number of convolutional layers. The learning rate is empirically chosen as 1×10^{-3} based on short random search. The Adam optimizer (Kingma and Ba, 2014) is used to minimize the loss of network.

Train set comprises less number of B-scans which results in over-fitting of the

Table 5.1: Different architectures formed by varying number of convolutional layers.

Total Layers	Starting Filter Size [$k_w \times k_h \times N$]	$Depth$	Total Parameters
7	3×3×16	0	16,097
11	3×3×16	1	68,705
15	3×3×16	2	268,129
22	3×3×16	3	850,529
30	3×3×16	4	2,665,313

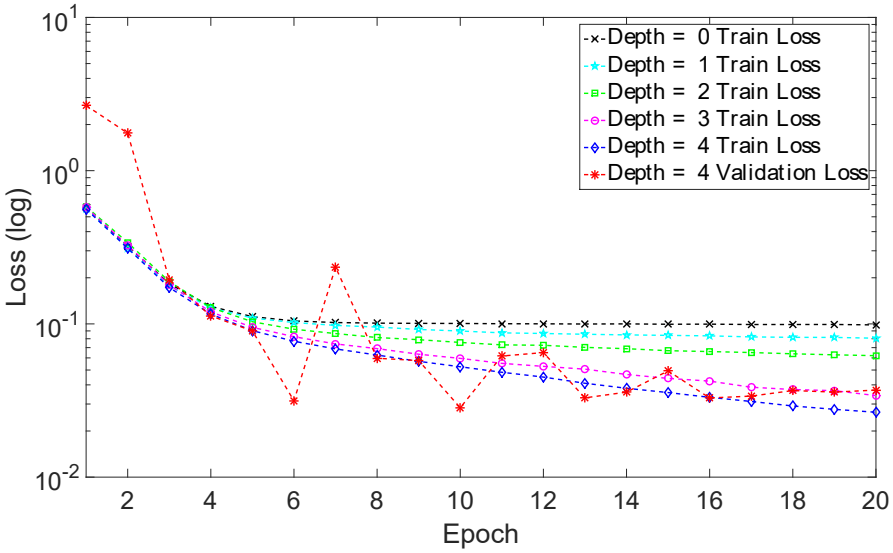


Figure 5.4: Loss vs. epoch graphs for proposed model $Depth$: [0, 1, 2, 3, 4] architectures after model training. (Plotted in log scale)

model. Hence, data augmentation technique is applied to generalize the model and prevent over-fitting of the model. Different augmentation techniques are applied such as horizontal flipping, image contrast and brightness variations, edge enhancement, and width, height and zoom shifts. On-the-fly data augmentation is performed during training the model by alleviating storage constraints.

Various architectures of the DSCN model is constructed by varying $Depth$ and starting filter size k for finding out optimal configuration. The loss graph across the different $Depth$ architectures is shown in Figure. 5.4. It can be inferred from the graph that the training loss is significantly reduced in $Depth = 4$ compared to $Depth = 0$, 1, 2 and

Table 5.2: Distinct architectures by varying number of filters in *Depth* = 4.

Starting Filter Size [$k_w \times k_h \times N$]	Total Parameters
3×3×16	2,665,313
3×3×32	10,481,345

3 which exhibits limited learning capabilities. Our earlier study in Chapter 4 reports that deeper architectures can account for more features thus yielding higher learning capabilities. Hence, $Depth$ =4 architecture is chosen over others.

To identify better receptive field of the $Depth = 4$ model, two architectures are formed by using $N = 16$ and $N = 32$ as starting filter sizes. Table 5.2 shows the number of parameters in these two architectures and Figure. 5.5 show their loss graph. We observe that starting filter size $N = 16$ converges to the lower training loss and exhibits lesser over-fitting trend compared to $N = 32$. Hence, final optimal architecture is formed by using $N = 16$ and $Depth = 4$. Due to lesser train samples, we have merged validation set with the training set thus yielding total 1676 B-scans as train samples and on-the-fly augmentation is performed for training the proposed optimal architecture. The proposed model is trained up to 100 epochs, beyond which, no further change in loss is observed. Total train samples presented to the network after data augmentation for final training is $100 \times 1676 = 167,600$.

5.3 Results and Analysis

The proposed architecture is validated using test datasets (909 B-scans) of the OCSC dataset. Segmentation results are compared with the ground truth of the OCSC dataset, ascertained by two expert ophthalmologists (G_1, G_2) and $G_1 \cap G_2$. Precision, recall and Dice coefficient metrics are used to compute the performance of the proposed method. Precision is the fraction of relevant instances among retrieved instances and recall is the fraction of relevant instances that have been retrieved over the total amount of relevant

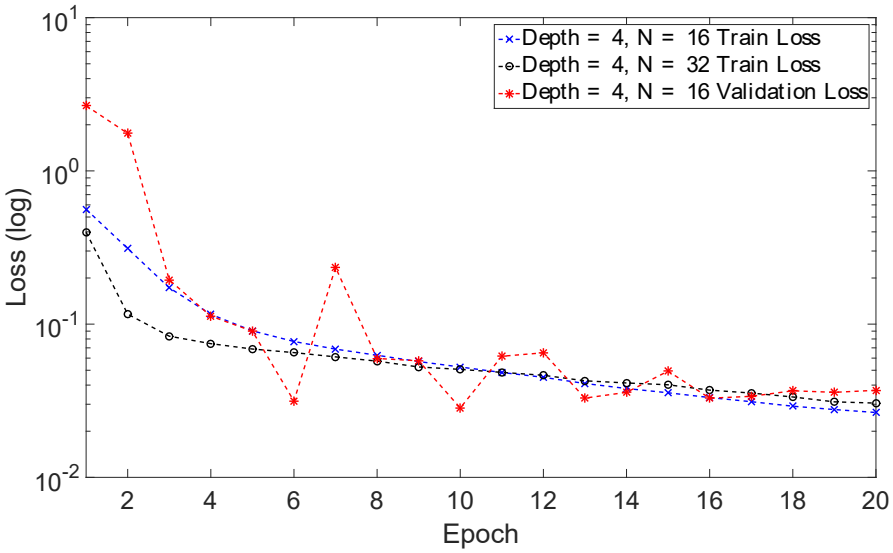


Figure 5.5: Loss vs. epoch graphs of distinct architectures with varying starting filter size in $Depth = 4$. (Plotted in log scale)

Table 5.3: Mean (standard deviation) of precision, recall and Dice of the proposed method.

Vendor	G_1			G_2			$G_1 \cap G_2$		
	Precision	Recall	Dice	Precision	Recall	Dice	Precision	Recall	Dice
Cirrus	0.83 (0.04)	0.62 (0.34)	0.66 (0.30)	0.86 (0.02)	0.61 (0.33)	0.66 (0.30)	0.79 (0.03)	0.67 (0.36)	0.67 (0.29)
Nidek	0.83 (0.05)	0.70 (0.28)	0.74 (0.20)	0.81 (0.02)	0.70 (0.30)	0.73 (0.20)	0.75 (0.01)	0.75 (0.31)	0.72 (0.18)
Spectralis	0.91 (0.06)	0.62 (0.06)	0.74 (0.05)	0.92 (0.06)	0.64 (0.07)	0.75 (0.06)	0.88 (0.06)	0.70 (0.07)	0.78 (0.06)
Topcon	0.88 (0.05)	0.68 (0.25)	0.75 (0.20)	0.88 (0.04)	0.70 (0.23)	0.77 (0.17)	0.83 (0.06)	0.75 (0.22)	0.78 (0.16)
Overall	**0.86 (0.03)**	**0.65 (0.03)**	**0.72 (0.04)**	**0.87 (0.04)**	**0.66 (0.04)**	**0.73 (0.04)**	**0.82 (0.05)**	**0.72 (0.03)**	**0.74 (0.05)**

instances. Dice coefficient reports the segmentation correlation index. Precision, recall and Dice coefficient are given in (2.1) , (2.2), and (2.3), respectively.

Table 5.3 present the results of proposed method compared with ground truth G_1, G_2, and $G_1 \cap G_2$ on precision, recall and dice co-efficient metrics. Our method gave a mean recall rate of 0.72 and precision rate 0.82 across the vendors on ground truth obtained from $G_1 \cap G_2$. The precision and recall rates reported across all the vendors shows that the proposed method provides better recall rate while preserving higher pre-

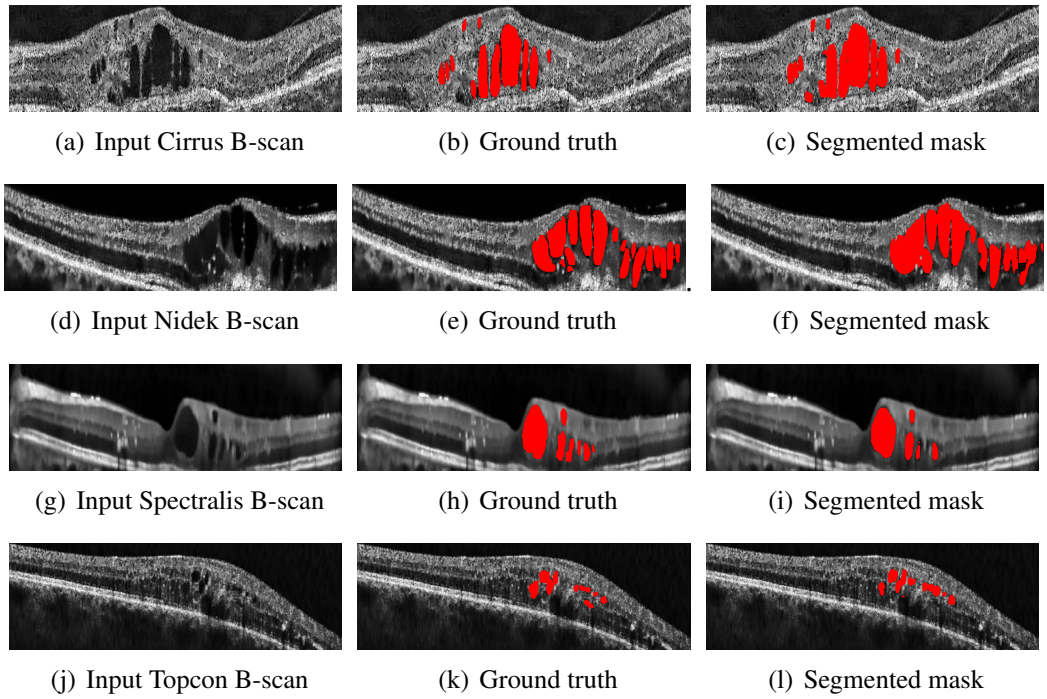
cision rates. Dice coefficient rate of our method shows higher segmentation correlation across the ground truths (0.72 on G_1, 0.73 on G_2, and 0.74 on $G_1 \cap G_2$).

It can be noticed that, precision and recall results on the Cirrus scans are lower compared to other vendor OCT scans due to the poor quality of the scans. Similar trend can be observed in the segmentation correlation results reported using Dice metric. Due to inter-grader observability, there is a marginal difference in the reported results of G_1, G_2, and $G_1 \cap G_2$. Lower standard deviation values of Dice metric on Spectralis vendor scans implies that proposed method performs accurate segmentation correlation on these scans compared to others due to the lower scan intensity variation and noise of the Spectralis scans.

For qualitative analysis, segmentation results of the proposed method are shown in Figure. 5.6. The images shown in first column are original B-scans of four different vendors (see Figures. 5.6(a), 5.6(d), 5.6(g), and 5.6(j)) and images of second column are the corresponding ground truth (see Figures. 5.6(b), 5.6(e), 5.6(h), and 5.6(k)). Resultant segmentation images of the proposed method are shown in third column (see Figures. 5.6(c), 5.6(f), 5.6(i), and 5.6(l)).

We can observe that the proposed method segmentation merges consecutive cysts into a larger cyst on Cirrus and Nidek scans (see Figures. 5.6(c) and 5.6(f)), which is due to the intra-scan intensity variations and ill-defined boundary between individual cysts (see Figures. 5.6(b) and 5.6(e)). In Spectralis and Topcon scans (see Figures. 5.6(i) and 5.6(l)), the proposed method is unable to detect the pseudo cystic regions due to low intensity variations between the cysts and neighboring non-cystic regions.

Table 5.4 compares the mean Dice coefficient of the proposed method with the OCSC challenge results and other state-of-the-art methods developed for OCSC dataset. de Sisterns (de Sisternes *et al.*, 2015) and Esmaeli (Esmaeili *et al.*, 2016) proposed handcrafted feature based supervised approaches, whereas Oguz (Oguz *et al.*, 2016) and Haritz (Gopinath and Sivaswamy, 2016) proposed unsupervised segmentation algorithms for vendor independent IRC segmentation on OCSC dataset. Generalizing these



(a) Input Cirrus B-scan (b) Ground truth (c) Segmented mask

(d) Input Nidek B-scan (e) Ground truth (f) Segmented mask

(g) Input Spectralis B-scan (h) Ground truth (i) Segmented mask

(j) Input Topcon B-scan (k) Ground truth (l) Segmented mask

Figure 5.6: Segmentation results of proposed model on four vendors OCT input scans from OCSC dataset.

methods across the vendors might affect the feature extraction and parameter selection process, which may lead to poor performance on cross vendors OCT IRC segmentation.

Venhuizen *et al.* (2016) proposed multi-scale patch based cascaded CNN, but the selection of appropriate patch sizes is tedious and it is found that patch-sizes affect final segmentation methods. Very recently, Gopinath (Gopinath *et al.*, 2017) proposed FCN architecture for IRC segmentation on OCSC dataset. Both these methods use regular convolutions, which may lead to larger parameters and require lot of computation. Our proposed method uses depthwise separable convolutions to reduce the computation complexity and outperforms other reported methods by providing mean Dice rate of 0.74 on OCSC dataset (see Table 5.4).

Experiments are also conducted to analyze the effect of vanishing gradient problem in deeper networks. Different activation functions, Swish and ReLU are used to analyze the performance and it has been found that results of Swish activation function outperform the results of ReLU activation function with 3% improvement in the Dice rate.

Table 5.4: Comparison based on Dice coefficient mean (standard deviation) of the proposed method against the results of OPTIMA cyst segmentation challenge on all vendor test OCT volumes.

Method	**DICE COEFFICIENT**		
	G_1	G_2	$G_1 \cap G_2$
de Sisternes *et al.* (de Sisternes *et al.*, 2015)	0.68 (0.15)	0.67 (0.14)	0.69 (0.15)
Venhuizen *et al.* (Venhuizen *et al.*, 2016)	0.61 (0.19)	0.60 (0.19)	0.59 (0.19)
Oguz *et al.* (Oguz *et al.*, 2016)	0.60 (0.5)	0.59 (0.15)	0.60 (0.14)
Esmaeili *et al.* (Esmaeili *et al.*, 2016)	0.55 (0.27)	0.55 (0.27)	0.55 (0.28)
Haritz *et al.* (Gopinath and Sivaswamy, 2016)	0.23 (0.15)	0.23 (0.15)	0.23 (0.15)
Gopinath *et al.* (Gopinath *et al.*, 2017)	0.71 (0.17)	0.72 (0.15)	0.72 (0.16)
Proposed FCN Method (Chapter 4)	**0.71 (0.20)**	**0.72 (0.19)**	**0.72 (0.19)**
Proposed method	**0.72 (0.04)**	**0.73 (0.04)**	**0.74 (0.05)**

Table 5.5: Comparison between depthwise separable convolution and regular convolution architectures.

CNN architecture type	# Parameters	DC (in %)
FCN Method (Chapter 4, U-Net based $Depth = 4$)	1,946,705	72
Regular Convolution Network (Xception based $Depth = 4$)	20,404,257	72
DSCN (proposed network)	2,665,313	74

Further, we have analyzed the effect of regular convolution and depthwise separable convolutions on the proposed architecture. We trained optimal $Depth = 4$ architecture by replacing depthwise separable convolutions with regular convolutions from scratch for 100 epochs. The obtained results show that regular convolution model requires 86% more parameters than the proposed DSCN architecture (see Table 5.5). Table 5.5 also reports the comparison of number parameters in FCN architecture proposed in Chapter4 with proposed DSCN architecture. We also observed that proposed DSCN architecture outperforms regular convolution architecture and Chapter4 FCN architecture by providing 3% improvement in mean Dice rate. To analyze the generalizability of the proposed architecture for vendor independent IRC segmentation, publicly available Duke DME dataset is also used. The 110 B-scans of Duke DME datasets are pre-processed as mentioned in section 5.1.1. The architecture is not explicitly trained on Duke DME dataset. Our method provides Dice rate of 0.54 on ground truth obtained from $G_1 \cap G_2$

by outperforming the results reported in (Chiu *et al.*, 2015).

5.4 Summary

In this Chapter, an efficient end-end convolutional neural network architecture is proposed for vendor independent IRC cyst segmentation. The proposed model utilizes depthwise separable convolutional filters instead of regular convolutional filters, which significantly reduces the computational parameters compared to regular FCN. To prevent vanishing gradient problem, swish activation function is employed. Skip connections used in the proposed architecture effectively combines higher-order semantic information from down-sampling phase to up-sampling phase thus improving the segmentation accuracy.

CHAPTER 6

CONCLUSIONS AND FUTURE WORK

6.1 Conclusions

One of the objectives of this thesis was to identify factors affecting intra-retinal cyst segmentation from OCT scans. OCT images are generally affected by speckle noise, signal intensity variations across scans, etc., and retinal scans have inherent anatomical structures which produce artifacts, such as blood vessel shadows, layered architecture disruption in disease, etc. IRC segmentation poses several challenges due to the above. This thesis presents a comparative assessment of existing automated intra-retinal cyst segmentation methods on OCT B-scans. The standardized methodological modules and cyst segmentation experiments conducted in this thesis demonstrate that factors such as pixel intensity variations, noise, blood vessel shadows and retinal layer distortions can impact automated cyst segmentation accuracy. In addition, pre-processing and post-processing steps are found to play a vital role in automated cyst segmentation processes.

A novel unsupervised two-stage automatic intra-retinal cyst segmentation algorithm based on marker-controlled watershed transform has been proposed in this thesis to address the accurate delineation of cystic structures. In this method, cystic regions are initially detected and then they are delineated up to their true boundary. Spectralis vendor OCT scans of OPTIMA cyst segmentation challenge dataset were used to validate the proposed method. The proposed MWS method efficiently segmented the cystoid regions with a high recall rate of 0.68 while maintaining precision rate at 0.80. Results were analyzed by conducting pixel-based evaluation in addition to region-based evaluation. Obtained results show that there was a marginal difference between the region- and pixel-based analysis.

To address generalizability issues of IRC segmentation on cross vendor OCT scans, a fully convolutional neural network architecture is proposed in this thesis. The proposed architecture is customized for IRC segmentation and trained from scratch. Domain specific image enhancement, data augmentation and hyper-parameter optimizations were performed to design optimal architecture. The proposed optimal FCN model with lower receptive field (starting filter size 16) and *Depth* 4 results in a computationally efficient yet robust vendor-independent model for IRCs. Additionally, domain-specific data augmentation methods are found to improve model training and convergence rates. The proposed model is evaluated qualitatively and quantitatively on OPTIMA cyst challenge dataset (four vendor OCT scan sets). The results demonstrate that the proposed method efficiently segments the IRCs by providing mean Dice rate of 0.71 on scans obtained across four different vendors. The generalizability of the proposed method was also tested using additional publicly available Duke DME dataset and it performed well without additional training on this dataset.

Deeper FCNs leads to larger number of computation parameters and may be prone to vanishing gradient problem. Hence, this thesis proposes a depthwise separable convolutional neural network architecture to address these issues. Proposed DSCN reduces the computational parameters compared to regular FCNs. To reduce the effect of vanishing gradient in deeper networks, an efficient activation function has been employed in the proposed architecture. The proposed architecture is robust to scan intensity variations and image noise variability across the vendors by providing segmentation results with a mean dice rate of 0.74 on OPTIMA cyst segmentation challenge dataset. Generalizability analysis of the proposed method is done using publicly available Duke DME dataset. Without any further training and fine-tuning of the proposed method on this dataset, proposed method produces a mean dice rate of 0.54.

6.2 Future Work

The proposed IRC segmentation techniques in this thesis are able to segment intra-retinal cysts in OCT scans upto their true boundary. Proposed CNN architectures are able to segment the cysts efficiently across different vendor scans. However, proposed CNN methods fail to delineate the cysts accurately when the quality of the scans is poor. Hence, future work may be directed towards cystic boundary detection or cyst counting tasks by suitably modifying the FCN parameters, goals and loss functions. It is noteworthy that for area-based segmentation tasks, penalizing FP pixels more than FN pixels in the loss function leads to smaller cysts getting missed while over detection and merged cysts are observed for higher penalization of FNs over FPs. Future efforts can be directed towards the implementation of such modified proposed architectures for the detection of intra-retinal/sub-retinal fluid spaces, pigment epithelial detachment and other macular degenerative pathologies.

It has been observed from studies conducted in this thesis, in future work several factors must be considered while designing an automated cyst segmentation technique. The desirable features to be included in prospective algorithms would include (but not be limited to):

- Image denoising.
- Removal of blood vessel and hard exudate shadow artifacts prior to segmentation.
- Retinal layer segmentation.
- Detection of candidate regions and accurate delineation of lesion boundaries.
- Post-processing operations for removal of false positive candidate regions.

REFERENCES

Abramoff, M. D., **M. K. Garvin**, and **M. Sonka** (2010). Retinal imaging and image analysis. *IEEE Reviews in Biomedical Engineering*, **3**, 169–208. ISSN 1937-3333. xiii, 25, 75, 95

Al-Mujaini, A., **U. K. Wali**, and **S. Azeem** (2013). Optical coherence tomography: Clinical applications in medical practice. *Oman Medical J.*, **28**(2), 86 – 91. ISSN 1999768X. 42

Arora, R., **A. Basu**, **P. Mianjy**, and **A. Mukherjee** (2016). Understanding deep neural networks with rectified linear units. *CoRR*, **abs/1611.01491**. 99

Beucher, S. and **F. Meyer** (1992). The morphological approach to segmentation: the watershed transformation. *Optical Engineering-New York-Marcel Dekker Inc-*, **34**, 433–433. 44

Breiman, L. (2001). Random forests. *Machine Learning*, **45**(1), 5–32. ISSN 1573-0565. 15

Chen, X., **M. Niemeijer**, **L. Zhang**, **K. Lee**, **M. D. Abramoff**, and **M. Sonka** (2012). Three-dimensional segmentation of fluid-associated abnormalities in retinal oct: Probability constrained graph-search-graph-cut. *IEEE Trans. Med. Imaging,*, **31**(8), 1521–1531. ISSN 0278-0062. 14, 15, 17

Chiu, S. J., **M. J. Allingham**, **P. S. Mettu**, **S. W. Cousins**, **J. A. Izatt**, and **S. Farsiu** (2015). Kernel regression based segmentation of optical coherence tomography images with diabetic macular edema. *Biomed. Opt. Express*, **6**(4), 1172–1194. 14, 15, 16, 88, 107

Chollet, F. (2015). Keras. 78

Chollet, F. (2016). Xception: Deep Learning with Depthwise Separable Convolutions. *ArXiv e-prints*. 95, 100

Coupe, P., **P. Hellier**, **C. Kervrann**, and **C. Barillot** (2009). Nonlocal means-based speckle filtering for ultrasound images. *IEEE Trans. Image Process.*, **18**(10), 2221–2229. ISSN 1057-7149. 42, 47

Cover, T. and **P. Hart** (1967). Nearest neighbor pattern classification. *IEEE Trans. info. Theory*, **13**(1), 21–27. 15

Cristoforetti, A., **L. Faes**, **F. Ravelli**, **M. Centonze**, **M. D. Greco**, **R. Antolini**, and **G. Nollo** (2008). Isolation of the left atrial surface from cardiac multi-detector ct images based on marker controlled watershed segmentation. *Med. Eng. Phys.*, **30**(1), 48 – 58. ISSN 1350-4533. 41

Cui, Y., **Y. Tan**, **B. Zhao**, **L. Liberman**, **R. Parbhu**, **J. Kaplan**, **M. Theodoulou**, **C. Hudis**, and **L. H. Schwartz** (2009). Malignant lesion segmentation in contrast-enhanced breast mr images based on the marker-controlled watershed. *Med. Phy.*, **36**(10), 4359–4369. ISSN 2473-4209. 41

de Moura, J., **P. L. Vidal**, **J. Novo**, **J. Rouco**, and **M. Ortega** (2017). Feature definition, analysis and selection for cystoid region characterization in optical coherence tomography. *Procedia Computer Science*, **112**, 1369 – 1377. ISSN 1877-0509. Knowledge-Based and Intelligent Information & Engineering Systems: Proceedings of the 21st International Conference, KES-20176-8 September 2017, Marseille, France. 14, 15, 16

de Sisternes, L., **J. Hong**, **T. Leng**, and **D. L. Rubin** (2015). A machine learning aproach for device-independent automated segmentation of retinal cysts in spectral domain optical cohorence tomography images. URL https://optima.meduniwien.ac.at/fileadmin/Challenge2015/Sisternes-CystChallenge15.pdf. 86, 87, 104, 106

Dice, L. R. (1945). Measures of the amount of ecologic association between species. *Ecology*, **26**(3), 297–302. 21, 53

Esmaeili, M., **A. M. Dehnavi**, **H. Rabbani**, and **F. Hajizadeh** (2016). Three-dimensional segmentation of retinal cysts from spectral-domain optical coherence tomography images by the use of three-dimensional curvelet based k-svd. *Journal of Medical Signals and Sensors*, **6**(3), 166. 87, 104, 106

Fernandez, D. C. (2005). Delineating fluid-filled region boundaries in optical coherence tomography images of the retina. *IEEE Trans. Med. Imag.*, **24**(8), 929–945. 14

Froment, J. (2014). Parameter-free fast pixelwise non-local means denoising. *Image Processing On Line*, **4**, 300–326. 74

Garg, P. G. (2014). Cystoid macular edema. URL http://eyewiki.aao.org/Cystoid_Macular_Edema. xii, 9

Garvin, M. K., **M. D. Abramoff**, **X. Wu**, **S. R. Russell**, **T. L. Burns**, and **M. Sonka** (2009). Automated 3-d intraretinal layer segmentation of macular spectral-domain optical coherence tomography images. *IEEE Trans. Med. Imag*, **28**(9), 1436–1447. ISSN 0278-0062. xiii, 25, 75, 95

Gomez, W., **L. Leija**, **A. V. Alvarenga**, **A. F. C. Infantosi**, and **W. C. A. Pereira** (2010). Computerized lesion segmentation of breast ultrasound based on marker-controlled watershed transformation. *Med. Phy.*, **37**(1), 82–95. ISSN 2473-4209. 41

Gonzalez, A., **B. Remeseiro**, **M. Ortega**, **M. G. Penedo**, and **P. Charln**, Automatic cyst detection in oct retinal images combining region flooding and texture analysis. *In Proceedings of the 26th IEEE International Symposium on Computer-Based Medical Systems*. 2013. ISSN 1063-7125. xiii, 14, 15, 22, 26, 30, 31, 32, 33, 34, 35, 42, 50

Goodfellow, I., **Y. Bengio**, and **A. Courville**, *Deep learning*. MIT Press, 2016. 64, 65, 67, 68, 99

GoogleColab (2017). Google colaboration research centre. URL https://www.colab.research.google.com. 100

Gopinath, K. and **J. Sivaswamy** (2016). Domain knowledge assisted cyst segmentation in OCT retinal images. *CoRR*, **abs/1612.02675**. 87, 104, 106

Gopinath, K., **J. Sivaswamy**, **A. P. Reddy**, and **S. G. Agraharam** (2017). Segmentation of retinal cysts from optical coherence tomography volumes via selective enhancement. *arXiv preprint arXiv:1708.06197*. 105, 106

Hartigan, J. A. and **M. A. Wong** (1979). Algorithm as 136: A k-means clustering algorithm. *J. R. Stat. Soc. Ser. C (Appl. Stat.)*, **28**(1), 100–108. ISSN 00359254, 14679876. 15, 44

Hartley, T. (2014). When parallelism gets tricky: Accelerating floyd-steinberg on the mali gpu. URL https://bit.ly/2s55z53. xiv, 72

He, K., **X. Zhang**, **S. Ren**, and **J. Sun**, Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. *In Proceedings of the IEEE International Conference on Computer Vision*. 2015. 81

He, K., **X. Zhang**, **S. Ren**, and **J. Sun**, Deep residual learning for image recognition. *In Proceedings of the IEEE Conf. Comput. Vis. Pattern Recognit.*. 2016. 62

Hearst, M. A., **S. T. Dumais**, **E. Osuna**, **J. Platt**, and **B. Scholkopf** (1998). Support vector machines. *IEEE Intelligent Systems and their applications*, **13**(4), 18–28. 15

Hochreiter, S. (1998). The vanishing gradient problem during learning recurrent neural nets and problem solutions. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, **6**(02), 107–116. 99

Hogan, M. J. and **J. JA Weddell** (1971). Histology of the human eye: an atlas and textbook. xii, 1

Hu, M.-K. (1962). Visual pattern recognition by moment invariants. *IEEE Trans. Info. Theory*, **8**(2), 179–187. 15

Huang, D., **E. A. Swanson**, **C. P. Lin**, **Schuman**, *et al.* (1991). Optical coherence tomography. *Science*, **254**(5035), 1178–1181. 7

Ioffe, S. and **C. Szegedy** (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*. 76

Irvine, S. R. (1953). A newly defined vitreous syndrome following cataract surgery. *Am. J. Ophthalmol.*, **36**(5), 600b. ISSN 0002-9394. 4

Kamnitsas, K., **L. Chen**, **C. Ledig**, **D. Rueckert**, and **B. Glocker** (2015). Multi-scale 3d convolutional neural networks for lesion segmentation in brain mri. *Ischemic Stroke Lesion Segmentation*, **13**. 62

Kingma, D. P. and **J. Ba** (2014). Adam: A method for stochastic optimization. *CoRR*, **abs/1412.6980**. 79, 100

Kohavi, R. *et al.*, A study of cross-validation and bootstrap for accuracy estimation and model selection. *In IJCAI*, volume 14. Stanford, CA, 1995. 78

Lang, A., **A. Carass**, **M. Hauser**, **E. S. Sotirchos**, **P. A. Calabresi**, **H. S. Ying**, and **J. L. Prince** (2013). Retinal layer segmentation of macular oct images using boundary classification. *Biomed. Opt. Express*, **4**(7), 1133–1152. 25, 43

Lang, A., **A. Carass**, **E. K. Swingle**, **O. Al-Louzi**, **P. Bhargava**, **S. Saidha**, **H. S. Ying**, **P. A. Calabresi**, and **J. L. Prince** (2015). Automatic segmentation of microcystic macular edema in oct. *Biomed. Opt. Express*, **6**(1), 155–169. xiii, 14, 15, 16, 23, 27, 30, 31, 32, 33, 34, 35, 36

Lee, C. S., **A. J. Tyring**, **N. P. Deruyter**, **Y. Wu**, **A. Rokem**, and **A. Y. Lee** (2017). Deep-learning based, automated segmentation of macular edema in optical coherence tomography. *Biomed. Opt. Express*, **8**(7), 3440–3448. 14, 15, 16

Lequan, Y., **H. Chen**, **Q. Dou**, **J. Qin**, and **P. A. Heng** (2016). Automated melanoma recognition in dermoscopy images via very deep residual networks. *IEEE Trans. Med. Imag.*. 62

Li, F.-F., **A. Karpathy**, and **J. Johnson** (2016). Cs231n: Convolutional neural networks for visual recognition. *URL: http://cs231n.github.io*. xiv, 65, 68, 69

Li, K., **X. Wu**, **D. Z. Chen**, and **M. Sonka** (2006). Optimal surface segmentation in volumetric images-a graph-theoretic approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **28**(1), 119–134. ISSN 0162-8828. xiii, 25, 75, 95

Long, J., **E. Shelhamer**, and **T. Darrell**, Fully convolutional networks for semantic segmentation. *In Proceedings of the IEEE Conf. Comput. Vis. Pattern Recognit.*. 2015. 62, 75

Lotufo, R. and **A. Falcao**, The ordered queue and the optimality of the watershed approaches. *In Mathematical Morphology and its Applications to Image and Signal Processing*. Springer, 2000, 341–350. 44, 46

Luo, W., **Y. Li**, **R. Urtasun**, and **R. Zemel**, Understanding the effective receptive field in deep convolutional neural networks. *In Advances in Neural Information Processing Systems*. 2016. 81

Martin, N. and **H. Maes**, *Multivariate analysis*. Academic press, 1979. 63

Mayer, M. A., **J. Hornegger**, **C. Y. Mardin**, and **R. P. Tornow** (2010). Retinal nerve fiber layer segmentation on fd-oct scans of normal subjects and glaucoma patients. *Biomed. Opt. Express*, **1**(5), 1358–1383. xiv, 25, 43, 48

Milletari, F., **N. Navab**, and **S.-A. Ahmadi**, V-net: Fully convolutional neural networks for volumetric medical image segmentation. *In 2016 Fourth International Conference on 3D Vision (3DV)*. IEEE, 2016. 62

Ng, A. Y. and **M. I. Jordan**, On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes. *In Advances in neural information processing systems*. 2002. 15

Nielsen, M. A. (2015). Neural networks and deep learning. *URL: http://neuralnetworksanddeeplearning.com/*. 65, 66, 67, 68

Oguz, I., **L. Zhang**, **M. D. Abràmoff**, and **M. Sonka**, Optimal retinal cyst segmentation from oct images. *In SPIE Med. Imag.*. International Society for Optics and Photonics, 2016. 86, 87, 104, 106

OPTIMA (2015). Optima cyst segmentation challenge. URL http://optima.meduniwien.ac.at/optima-segmentation-challenge-1/. xii, 9, 13, 19, 20, 28, 47, 49, 78, 86

Pascanu, R., **T. Mikolov**, and **Y. Bengio** (2012). Understanding the exploding gradient problem. *CoRR*, **abs/1211.5063**. 99

Paul-Louis (). About Separable Convolution. https://bit.ly/2wLNOfR. Accessed : 7-5-2018. 95

Pilch, M., **K. Stieger**, **Y. Wenner**, **M. N. Preising**, **C. Friedburg**, **E. Meyer zu Bexten**, and **B. Lorenz** (2013). Automated segmentation of pathological cavities in optical coherence tomography scans. *Invest. Ophthalmol. Vis. Sci.*, **54**(6), 4385. xiii, 15, 22, 26, 30, 31, 32, 33, 34, 35

Quellec, G., **K. Lee**, **M. Dolejsi**, **M. K. Garvin**, **M. D. Abramoff**, and **M. Sonka** (2010). Three-dimensional analysis of retinal layer texture: Identification of fluid-filled regions in sd-oct of the macula. *IEEE Trans. Med. Imag.*, **29**(6), 1321–1330. ISSN 0278-0062. 14, 15, 17

Ramachandran, P., **B. Zoph**, and **Q. V. Le** (2017). Searching for activation functions. *CoRR*, **abs/1710.05941**. 99

Rashno, A., **B. Nazari**, **D. D. Koozekanani**, **P. M. Drayna**, **S. Sadri**, **H. Rabbani**, and **K. K. Parhi** (2017). Fully-automated segmentation of fluid regions in exudative age-related macular degeneration subjects: Kernel graph cut in neutrosophic domain. *PloS one*, **12**(10), e0186949. 14, 15, 17

Retina, R. (2015). http://www.retinareference.com/. URL http://www.retinareference.com/anatomy/. xii, 3

Ronneberger, O., **P. Fischer**, and **T. Brox**, U-net: Convolutional networks for biomedical image segmentation. *In International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, 2015. 62, 75, 79

Russakovsky, O., **J. Deng**, **H. Su**, **J. Krause**, **S. Satheesh**, **S. Ma**, **Z. Huang**, **A. Karpathy**, **A. Khosla**, **M. Bernstein**, *et al.* (2015). Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, **115**(3), 211–252. 71

Schlegl, T., **A.-M. Glodan**, **D. Podkowinski**, **S. M. Waldstein**, **B. S. Gerendas**, **U. Schmidt-Erfurth**, and **G. Langs** (2015). Automatic segmentation and classification of intraretinal cystoid fluid and subretinal fluid in 3d-oct using convolutional neural networks. *Invest. Ophthalmol. Vis. Sci.*, **56**(7), 5920–5920. 14, 15, 16

Shore, J. and **R. Johnson** (1980). Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy. *IEEE Trans. Info. Theory*, **26**(1), 26–37. 99

Simonyan, K. and **A. Zisserman** (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*. 76, 97

Sudeep, P., **S. I. Niwas**, **P. Palanisamy**, **J. Rajan**, **Y. Xiaojun**, **X. Wang**, **Y. Luo**, and **L. Liu** (2016*a*). Enhancement and bias removal of optical coherence tomography images: An iterative approach with adaptive bilateral filtering. *Comput. in Bio. and Med.*, **71**(Supplement C), 97 – 107. ISSN 0010-4825. 24, 42, 73, 74

Sudeep, P., **P. Palanisamy**, **J. Rajan**, **H. Baradaran**, **L. Saba**, **A. Gupta**, and **J. S. Suri** (2016*b*). Speckle reduction in medical ultrasound images using an unbiased non-local means method. *Biomedical Signal Processing and Control*, **28**, 1–8. 73, 74

Sudeep, P., **P. Palanisamy**, **J. Rajan**, **H. Baradaran**, **L. Saba**, **A. Gupta**, and **J. S. Suri** (2016*c*). Speckle reduction in medical ultrasound images using an unbiased non-local means method. *Biomedical Signal Processing and Control*, **28**, 1–8. 94

Swingle, E. K., **A. Lang**, **A. Carass**, **O. Al-Louzi**, **S. Saidha**, **J. L. Prince**, and **P. A. Calabresi**, Segmentation of microcystic macular edema in cirrus oct scans with an exploratory longitudinal study. *In Proc. SPIE*, volume 9417. 2015. 14, 15, 23

Swingle, E. K., **A. Lang**, **A. Carass**, **H. S. Ying**, **P. A. Calabresi**, and **J. L. Prince**, Microcystic macular edema detection in retina oct images. *In SPIE Med. Imag.*. International Society for Optics and Photonics, 2014. xiii, 14, 15, 22, 23, 26, 28, 30, 31, 32, 33, 34, 35

Tomasi, C. and **R. Manduchi**, Bilateral filtering for gray and color images. *In 1998 Sixth International Conference on Computer Vision*. 1998. 24

Venhuizen, F., **M. J. van Grinsven**, **B. van Ginneken**, **C. C. Hoyng**, **T. Theelen**, and **C. I. Sanchez** (2016). Fully automated segmentation of intraretinal cysts in 3d optical coherence tomography. *Invest. Ophthalmol. Vis. Sci.*, **57**(12), 5949–5949. 86, 87, 105, 106

Weisstein, E. W. (2017). Convolution. *MathWorld–A Wolfram Web Resource. URL: http://mathworld.wolfram.com/Convolution.html*. 71

WHO (2007). Global initiative for the elimination of avoidable blindness: action plan 2006-2011. 4, 10

Wieclawek, W., Automatic cysts detection in optical coherence tomography images. *In 2015 22nd International Conference Mixed Design of Integrated Circuits Systems (MIXDES)*. 2015. xiii, xiv, 14, 15, 16, 23, 27, 30, 31, 32, 33, 34, 35, 50, 51, 52, 53, 54, 55, 56, 57

Wilkins, G. R., **O. M. Houghton**, and **A. L. Oldenburg** (2012). Automated segmentation of intraretinal cystoid fluid in optical coherence tomography. *IEEE Trans. Bio. Eng.*, **59**(4), 1109–1114. ISSN 0018-9294. xiii, xiv, 4, 10, 14, 15, 21, 22, 26, 28, 30, 31, 32, 33, 34, 35, 42, 50, 51, 52, 53, 54, 55, 56, 57

Xu, S., **H. Liu**, and **E. Song** (2011). Marker-controlled watershed for lesion segmentation in mammograms. *Journal of Digital Imaging*, **24**(5), 754–763. ISSN 1618-727X. 41

Yan, J., **B. Zhao**, **L. Wang**, **A. Zelenetz**, and **L. H. Schwartz** (2006). Marker-controlled watershed for lymphoma segmentation in sequential ct images. *Med. Phy.*, **33**(7Part1), 2452–2460. ISSN 2473-4209. 41

Zeiler, M. D. and **R. Fergus**, Visualizing and understanding convolutional networks. *In European conference on computer vision*. Springer, 2014. 72

Zhang, L., **W. Zhu**, **F. Shi**, **H. Chen**, and **X. Chen**, Automated segmentation of intraretinal cystoid macular edema for retinal 3d oct images with macular hole. *In 2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*. 2015. ISSN 1945-7928. 14, 15, 17

Zheng, Y., **J. Sahni**, **C. Campa**, **A. N. Stangos**, **A. Raj**, and **S. P. Harding** (2013). Computerized assessment of intraretinal and subretinal fluid regions in spectral-domain

optical coherence tomography images of the retina. *American Journal of Ophthalmology*, **155**(2), 277 – 286.e1. ISSN 0002-9394. 14

Zuiderveld, K., Contrast limited adaptive histogram equalization. *In Graphics gems IV*. Academic Press Professional, Inc., 1994. 75

PUBLICATIONS BASED ON THIS THESIS

Journal Papers

1. **G. N. Girish**, Abhishek R. Kothari and Jeny Rajan, "Marker controlled watershed transform for intra-retinal cysts segmentation from optical coherence tomography B-scans", *Pattern Recognition Letters*, Elsevier, 2017 (**In Press**).

2. **G. N. Girish**, V. A. Anima , Abhishek R. Kothari, P. V. Sudeep, Sohini Roychowdhury and Jeny Rajan, "A Benchmark Study of Automated Intra-retinal Cyst Segmentation Algorithms using Optical Coherence Tomography B-Scans". *Computer Methods and Programs in Biomedicine*, Volume 153, pages 105-114, ISSN 0169-2607, Elsevier, 2018.

3. **G. N. Girish**, Bibhash Thakur, Sohini Roychowdhury, Abhishek R. Kothari, and Jeny Rajan, "Segmentation of Intra-Retinal Cysts from Optical Coherence Tomography Images using a Fully Convolutional Neural Network Model", *IEEE Journal of Biomedical and Health Informatics*, IEEE, 2018 (**In Press**).

4. **G. N. Girish**, Banoth Saikumar, Sohini Roychowdhury, Abhishek R. Kothari and Jeny Rajan, "Depth-wise Separable Convolutional Neural Network Model for Segmentation of Intra-Retinal Cysts from Optical Coherence Tomography Images", in *Biomedical Signal Processing and Control*, Elsevier, 2018. **(Under Review)**.

Conference Papers

1. **G. N. Girish**, Abhishek R. Kothari and Jeny Rajan, "Automated Segmentation of Intra-Retinal Cysts from Optical Coherence Tomography Scans Using Marker Controlled Watershed Transform", in *38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC'16)*. Orlando, Florida, USA: IEEE, Aug 2016.

Girish G N

CONTACT INFORMATION

Research Scholar
Department of Computer Science and Engineering
National Institute of Technology Karnataka
Surathkal, Mangalore
India-575025
Web Links: *Google Scholar*, *Research Gate*, *LinkedIn*

Cell: +91-9844289249
girishanit@gmail.com
girish.cs14f07@nitk.edu.in

RESEARCH INTERESTS

Image processing and computer vision, retinal image analysis, optical coherence tomography, object recognition and tracking, facial image analysis.

EDUCATION

Ph.D., Computer Science and Engineering, *Expected:* 2018
National Institute of Technology Karnataka, Surathkal, India

- Coursework CGPA: 9.25/10
- Area of Research: *Retinal Image Analysis*
- Thesis Title: *Automatic Segmentation of Intra-Retinal Cysts from Optical Coherence Tomography Scans*
- Advisor: Dr. Jeny Rajan

Master of Technology, Computer Science and Engineering, Sep 2014
Visvesvaraya Technological University (VTU), Belgaum, India

- Grade: First Class with Distinction (71.95%)
- University B.D.T. College of Engineering, Davanagere, India
- Area of Research: *Biometrics and Image Processing*
- Thesis Title:*A Comparative Study on Face Recognition Using MB-LBP and PCA*
- Advisor: Dr. Shrinivasa Naika C.L.

Bachelor of Engineering, Computer Science and Engineering, May 2012

- Grade: First Class with Distinction (70.26%)
- J.N.N. College of Engineering, Shimoga, India
- Major Project Title: *Fingerprint Recognition System*

RESEARCH EXPERIENCE

Research Scholar Dec 2014 to present
Department of Computer Science and Engineering,
National Institute of Technology Karnataka
Supervisor: Dr. Jeny Rajan

Project Associate Oct 2014 to Dec 2014
Karnataka State Council for Science and Technology (KSCST), Indian Institute of Science (IISc) Campus, Bangalore.
Title of Project: Implementation of Virtual Laboratory (VL) to improve the quality of education in the backward taluks of Karnataka.
Principal Investigator: Dr. S.G.S. Swamy

PUBLICATIONS

Journal Publications

1. **G. N. Girish**, Bibhash Thakur, Sohini Roychowdhury, Abhishek R. Kothari, and J. Rajan, "Segmentation of Intra-Retinal Cysts from Optical Coherence Tomography Images using a Fully Convolutional Neural Network Model", *IEEE Journal of Biomedical and Health Informatics*, IEEE, 2018.

2. **G. N. Girish**, V. A. Anima , Abhishek R. Kothari, P. V. Sudeep, Sohini Roychowdhury and J. Rajan, "A Benchmark Study of Automated Intra-retinal Cyst Segmentation Algorithms using Optical Coherence Tomography B-Scans". *Computer Methods and Programs in Biomedicine*, Volume 153, pages 105-114, ISSN 0169-2607, Elsevier, 2018.

3. **G. N. Girish**, Abhishek R. Kothari and J. Rajan, "Marker controlled watershed transform for intra-retinal cysts segmentation from optical coherence tomography B-scans", *Pattern Recognition Letters*, Elsevier, 2017, ISSN 0167-8655.

Conference Publications

1. **G. N. Girish**, Abhishek R. Kothari and J. Rajan, "Automated Segmentation of Intra-Retinal Cysts from Optical Coherence Tomography Scans Using Marker Controlled Watershed Transform", in *38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC'16)*. Orlando, Florida, USA: IEEE, Aug 2016.

2. Sooraj Kumar R., **G. N. Girish**, P. B. Ramteke and S. G. Koolagudi, "Text Independent Automatic Accent Identication System for Kannada Language", in *International Conference on Data Engineering and Communication Technology (ICDECT-2016)*. Pune, INDIA: Springer, Mar 2016.

3. Narendra Rao T.J., **G. N. Girish**, and J. Rajan, "An Improved Contextual Information Based Approach for Anomaly Detection via Adaptive Inference for Surveillance Application", in *International Conference on Computer Vision and Image Processing (CVIP-2016)*. IIT Roorkee, INDIA: Springer, Feb 2016. **(Awarded as International Association for Pattern Recognition (IAPR) Best Student Paper)**

4. **G. N. Girish**, S. Naika C L, and P. K. Das, "Effect of Modified Convolution on Local Descriptor Based Face Recognition", in *Eighth International Multi Conference on Information Processing (IMCIP-2014)*. Bangalore, INDIA: Elsevier, July 2014.

5. **G. N. Girish**, S. Naika C L, and P. K. Das, "Face recognition using MB-LBP and PCA: A comparative study", in *International Conference on Computer Communication and Informatics (ICCCI -2014)*. Coimbatore INDIA: IEEE, January 2014.

Book Chapter

1. Narendra Rao T.J., **G. N. Girish**, Mohit P. Tahiliani and Jeny Rajan, "Anomalous Event Detection Methodologies for Surveillance Application - An Insight", *Handbook of Research on Advanced Concepts in Real-Time Image and Video Processing*, pages=1–27, IGI Global Publishers, 2017.

Awards

- **Department of Biotechnology, Govt. of India - Travel Grant Award 2016** for attending 38th IEEE EMBC'16 at Orlando, USA — Aug 2016
- **International Association for Pattern Recognition (IAPR) Best Student Paper** at CVIP-2016, IIT Roorkee — Feb 2016
- **MHRD Fellowship for PhD** — Dec 2014
- **2nd Topper of the College in M.Tech University Examinations** for the academic year 2013-14 — Sep 2014

Software Skills

Computer Programming:

- Python, MATLAB, C, and C++

Deep Learning Library:

- Keras

References

Dr. Jeny Rajan
Asst. Professor in Dept. of CSE,
National Institute of Technology Karnataka,
Surathkal, India
E-mail: jenyrajan@gmail.com
Phone: +91-7829430838

Dr. Sohini RoyChowdhury
Asst. Professor in Dept. of Electrical Engineering
Division of STEM, 223-Beardslee Building,
University of Washington, Bothell, USA.
E-mail: roych@uw.edu
Phone: +1 (785)-320-1035

Dr. S.G.S. Swamy
Executive Secretary
Karnataka State Council for Science & Technology
IISc Campus, Bangalore, India
E-mail: swamy@kscst.iisc.ernet.in
Phone: +91-9448515976

Dr. Abhishek R. Kothari
Director & Chief Vitreoretinal Surgeon
Pink City Eye and Retina Center
Jaipur, India
E-mail: dr.a.kothari@gmail.com
Phone: +91-9928033335