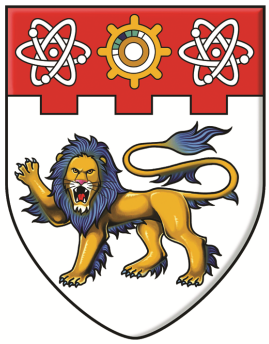
MULTIORGAN DETECTION AND SEGMENTATION STRATEGIES FOR 2D ULTRASOUND IMAGES OF THE THYROID GLAND

SUBBARAO NIKHIL NARAYAN

2016



NANYANG TECHNOLOGICAL UNIVERSITY

MULTIORGAN DETECTION AND SEGMENTATION STRATEGIES FOR 2D ULTRASOUND IMAGES OF THE THYROID GLAND

SUBBARAO NIKHIL NARAYAN
SCHOOL OF ELECTRICAL & ELECTRONIC ENGINEERING

2016

MULTIORGAN DETECTION AND SEGMENTATION STRATEGIES FOR 2D ULTRASOUND IMAGES OF THE THYROID GLAND

SUBBARAO NIKHIL NARAYAN

SUBBARAO NIKHIL NARAYAN

School of Electrical & Electronic Engineering

A thesis submitted to the Nanyang Technological University
in partial fulfilment of the requirements for the degree of
Doctor of Philosophy

2016

Thesis dedicated to my loving wife, Rohini.

Acknowledgments

I would like to take this opportunity to express my sincere gratitude to my supervisor, Assoc. Prof. Pina Marziliano, for providing me with an opportunity to undertake my PhD studies under her supervision and helping me establish a clinical collaboration with Tan Tock Seng Hospital, Singapore. Thank you for allowing me to undertake overseas research attachments and present papers at all the conferences.

I would like to thank Nanyang Technological University, Singapore for providing me with the financial support during my PhD studies.

I am grateful to my co-supervisor Prof. Nadia Magnenat Thalmann for providing me with the lab support and the facilities necessary to conduct research without glitches.

I would like to thank my clinical collaborators Dr. Christopher G.L. Hobbs and Dr. Jeevendra Kanagalingam for providing me with the clinical need and the necessary datasets along with the ground truth to validate my algorithms.

Thanks to Tian Jing, Talal Bin Amin, Samuel Deslauriers-Gauthier, Peter Arnoud Bensen and Pina Marziliano for volunteering to get your thyroid glands scanned and in effect helping me to obtain the datasets necessary to carry out my research work.

Thanks to my team at Siemens Corporate Research and Technologies Centre, Bangalore for providing me with an amazing opportunity to work on state-of-the-art medical imaging technologies during my internship.

Special thanks to Assoc. Prof. Chandra Sekhar Seelamantula for providing me with an opportunity to visit Spectrum Labs at Indian Institute of Science, Bangalore as a researcher for 4 months.

Thanks to all my Singapore friends, Vittal Premchandran, Kaushik Sunder, Govindrajan Sridharan, Cecil Wilson, Tejas Canchi, Vinay Adiga, Anirudh Kashyap, Akhila Anirudh, Nitin Shivaraman, Yasir Tahir, Vikram Shenoy Handiru, Supriya S., Yogeshwaran Jayaraman, Andrew Ee, Amrish Nair, Lavanya Bhaskaran, Srinivasan Yadunathan, Vinod Veera Reddy, Divya Venkatraman, Jayachandra Dakhala and Sathiya Narayanan for making my stay a memorable one. A very special thanks to Dr. Anil Kumar for all the technical discussions that we have had over a cup of coffee over the past 4 years.

I would not have been able to finish my PhD studies had it not been for the constant support and encouragement by my wife, Rohini. Thank you, with all my heart, for being so understanding and motivating during my PhD studies. I would not wish the difficulties of a long distance relationship on anyone, but if they all went as well as ours, they just might be worth it.

Last but not the least, I would like to thank my family and relatives for constantly

providing me with all the encouragement and care. In particular I would like to thank my father Dr. S.D. Subba Rao for all the wonderful discussions we have had with regard to research in medicine. You have been a constant source of inspiration to me and will always remain so. Had I not made a pact with my mother at the start of my PhD studies, I would never have been able to embark on this wonderful journey that is a PhD. It is this pact that got me my Rohini and my PhD. Thank you for all the wonderful things that you have provided me with in this life. I would also like to thank my brother Dr. Anil Kumar for all the wonderful discussions that we have had in the last four years on my research work.

Abstract

The thyroid gland is an endocrine gland that is responsible for iodine metabolism and hormone regulation in the human body. The thyroid gland is susceptible to disorders such as Goitre, Hashimotos Thyroiditis, cancer, etc., which often present with lumps in the neck. Any physician who examines the gland, first palpates the region where the lump is present and orders for a follow up investigation to make a diagnosis. This is done with the help of imaging modalities such as ultrasound, computed tomography (CT) and magnetic resonance imaging (MRI) . The choice of imaging modality is dependent on the severity of the condition prevalent with the disorder. Imaging with CT or MRI modalities are recommended when the treatment of the disorder involves surgically operating the gland. But in a clinical setting, which is where a majority of patients are treated, ultrasound is the gold standard imaging modality used to screen the thyroid gland for disorders.

The aim of this research is to devise automatic landmark based multiorgan detection and segmentation algorithms for freehand 2D ultrasound images of the thyroid gland. Three new methods are proposed to automatically detect and segment multiple organs in the US images of the thyroid gland. The organs that are considered for this research are: (a) the thyroid gland; (b) the carotid artery; (c) the trachea; and (d) the muscles. All of our methods make use of speckles and imaging artefacts in the image, which are otherwise considered to be noise, as sources of information to detect and segment the organs in ultrasound images. The speckle related pixels are determined by the application of Hessian based blob detectors on the ultrasound image. We demonstrate the application of our methods to perform guided interventions, volumetric analysis and computer aided diagnosis.

In the first of the three methods, the speckle related pixels are clustered into three echogenic classes based on its relative brightness in a kernel. An energy based model is then used to detect the carotid artery in the labelled image (image whose pixels are labelled with the cluster number). The carotid artery is used as a landmark to detect and segment the remaining organs in the image. In the second method, an agglomerative clustering constrained with a similarity metric is proposed to cluster the speckle related pixels into an unknown number of echogenic classes. The total number of class labels generated at the end of the clustering determines the number of echogenic classes in the image. This process is viewed as a quantization of the speckle related pixels into levels that are determined by the tissue echogenicity. The enhancement artefact is then detected in the quantized image which is in turn used as a landmark to detect the carotid

artery. The carotid artery is used as a landmark to detect the remaining organs. All organs are segmented by the application of region based local phase methods. In the last algorithm, the speckle related pixels are classified into their respective echogenic classes by the use of local phase based methods. The carotid artery and the trachea are detected from the binary image containing the hypoechoic pixels as foreground pixels. The pixels of the carotid artery and trachea along with the pixels that lie in the region between them are used to train a random forests classifier to classify the remaining pixels that belong to the hyperechoic tissue into the thyroid gland and background pixels.

The proposed methods are compared with each other and with state-of-the-art methods to prove the efficacy of the algorithms in performing medical image analysis. Manual segmentations obtained from two trained sonographers are used as ground truth to validate the algorithms. In all, 1091 images from five databases of annotated images are used in our experiments. The images are pre-processed by a novel annotation removal and image restoration algorithm that restores images to a high quality (average PSNR $> 38dB$) . All of our methods yield good quality segmentation results with Probabilistic Rand Index (PRI) > 0.83 and Boundary Error (BE) $<$ 1mm . Analysis of the results using the Dice co-efficient as the metric shows that the proposed methods performs better than the state-of-the-art methods.

Table of Contents

List of Figures

List of Tables

List of Acronyms and Notations

Acronyms

^{123}I	Iodine-123: Radioactive isotope of Iodine with molar mass of 122.9 grams/mol, page 2
^{131}I	Iodine-131: Radioactive isotope of Iodine with molar mass of 130.91 grams/mol, page 2
2D	Two Dimensional, page iii
2D+T	Time series 2D US image, page 30
3D	Three Dimensional, page 1
AWMF	Adaptive Weighted Median Filter, page 28
B	Breast, page 25
BDCT	Block Discrete Cosine Transform, page 42
BE	Boundary Error, page iv
C	Cardiac, page 25
CAD	Computer Aided Diagnosis, page 8
CRT	Cathode Ray Tube, page 13
CT	Computed Tomography, page iii
dB	Decibel, page iv
DSC	Dice Co-efficient, page 47
ENT	Ear, Nose and Throat, page 35

F	Foetus, page 25
FN	False Negative, page 47
FNAB	Fine Needle Aspiration Biopsy, page 7
FOV	Field of View, page 13
FP	False Positive, page 47
GAC	Geodesic Active Contour, page 25
GBM	Graph Based Model, page 25
GC	Graph Cut, page 25
GCE	Global Consistency Error, page 45
GDM	Geometric Deformable Model, page 25
GE	General Electric, page 53
GHz	Giga Hertz = 10^9 Hertz, page 47
GT	Ground Truth, page 35
GUI	Graphics User Interface, page 35
IJV	Internal Jugular Vein, page 8
IQ	In phase Quadrature, page 19
JCR	Joint Classification Regression, page 58
K	Kidney, page 25
k-NN	k Nearest Neighbour, page 102
k-NN	k Nearest Neighbour, page 92
L	Liver, page 25
LCD	Liquid Crystal Display, page 13
LS	Level Set, page 25

LS	Longitudinal Scan, page 35
MAP	Maximum A-Posteriori, page 58
MedTech	Medical Technology, page 1
MHz	Frequency in Megahertz = 10^6 Hertz, page 12
ml	millilitre, page 4
mm	Millimetre, page iv
MODS	Multi-Organ Detection and Segmentation, page 9
MRI	Magnetic Resonance Imaging, page iii
MSE	Mean Squared Error, page 102
MSE	Mean Squared Error, page 92
MSER	Maximally Stable Extremal Resions, page 58
MSL	Marginal Space Learning, page 57
NC	Normalized Cut, page 25
NN	Neural Network, page 102
NN	Neural Network, page 92
NN	Neural Networks, page 28
NPR	Normalized Probabilistic Rand, page 45
PBT	Probabilistic Boosting Tree, page 58
PET	Positron Emission Tomography, page 2
Pl	Pelvis, page 25
POCS	Projection Onto Convex Sets, page 37
PPV	Positive Predictive Value, page 47
Pr	Prostate, page 25

PRI	Probabilistic Rand Index, page iii
PSNR	Peak Signal to Noise Ratio, page iv
RBF	Radial Basis Function, page 51
RE	Refinement Error, page 46
RF	Radio frequency, page 15
RF	Random Forest, page 102
RF	Random Forest, page 92
SAS	Statistical analysis system, page 6
SE	Sensitivity, page 47
SE	Structuring Element, page 97
SLIC	Simple Linear Iterative Clustering, page 58
SP	Specificity, page 47
SPS	Speckle Patch Similarity, page 69
SRAD	Speckle Reducing Anisotropic Diffusion, page 17
SRR	Similarly Reflective Region, page 69
T	Thyroid, page 25
TGC	Time Gain Compensation, page 13
TN	True Negative, page 47
TP	True Positive, page 47
TS	Transverse Scan, page 35
TV	Thyroid Volume, page 102
TV	Thyroid Volume, page 92
US	Ultrasound, page iii

VOI	Variation of Information, page 45
cm	Centimetre, page 1
gms	Weight in Grams, page 1

Notations

$\bar{E}_i$	Energy inside the closed curve representing the candidate for carotid artery, page 64
$\bar{E}_o$	Energy outside the closed curve representing the candidate for carotid artery, page 64
$\bar{f}_r(x, y)$	$1 - f_r(x, y)$, page 40
$\bigtriangledown$	The gradient operator, page 65
$\bigtriangleup$	Set of all pixels that were removed during the annotation removal process, page 42
$\boldsymbol{\psi}$	Number of images in a dataset, page 45
$\cdot$	Element wise multiplication , page 41
$\delta(t)$	Dirac delta function, page 21
$\hat{V}(x, y)$	Hilbert transform of $V(x, y)$, page 15
∞	Infinity, page 46
λ	Class label of a pixel after clustering by the SPS method, page 73
log	Logarithm to the base 10, page 44
$\mathbb{I}$(condition)	Indicator function that results true if the condition is satisfied, page 45
$\mathbf{r}_m(k)$	Intensity of the k^{th} pixel in an $M_{R1} \times M_{R1}$ ROI defined around the m^{th} speckle pixel in $\mathcal{F}_s$., page 69

$\mathbf{x}$	Row vector of the co-ordinates $(x, y) \in \mathbb{Z}$, page 23
ν	Speed of sound in a tissue, page 15
$\otimes$	The convolution operator, page 15
$\phi(t)$	Local phase of a signal $f(t)$, page 21
$\phi(x, y)$ or $\phi(\mathbf{x})$	Local phase of a 2D signal, page 24
σ	Shape parameter for distributions (Rayleigh, Gaussian etc.), page 15
σ_x	Pulse width of the transmitting ultrasound wave, page 15
σ_y	Beam width of the transmitting ultrasound wave, page 15
τ	User defined threshold, page 62
Area_T	Cross-sectional area in the transverse scan of 2D US image, page 6
Length_L	Length of Thyroid gland in the longitudinal scan of 2D US image, page 6
$\tilde{f}$(x,y)	Restored image after the application of POCS algorithm, page 42
A_L	Surface area of the left lobe of the thyroid gland, page 28
A_l	Cross-sectional area in the longitudinal plane of 3D US volume, page 5
A_R	Surface area of the right lobe of the thyroid gland, page 28
A_t	Cross-sectional area in the transverse plane of 3D US volume, page 5
B	Block size, page 42
$b_0(x, y)$	Binary image obtained by setting all non zero pixels in $f(\mathbf{x})$ as foreground, page 40
$b_a(x, y)$	Binary annotation mask, page 40
C	Constant used to compute the dosage of radio-iodine, page 4
D	Dosage of ^{131}I, page 4
D	Maximum anterio-posterior depth of the gland under a transverse scan, page 4

D_L	Maximum depth of the left lobe of the thyroid gland under the transverse scan, page 28
D_l	Maximum depth in the longitudinal scan, page 5
D_R	Maximum depth of the right lobe of the thyroid gland under the transverse scan, page 28
D_t	Maximum depth in the transverse scan, page 5
$E(\mathbf{x})$	Binary image that contains the carotid candidates that have a shape of an ellipse, page 63
Ecc	Eccentricity of the Ellipse, page 73
F	Energy functional employed to localize the carotid artery in the three category MODS algorithm, page 64
$f(\mathbf{x})$	2D US image, page 40
$f(t)$	1D continuous time signal, page 15
$f(t)_{\mathrm{A}}$	Analytic signal at instant t, page 21
$f(t)_{\mathcal{H}}$	Hilbert transform of signal $f(t)$, page 21
$f(x, y)$	Amplitude of B-Mode ultrasound image at (x, y), page 15
$f(x, y)$	Intensity of 2D US image at (x, y), page 40
f_0	Centre frequency of the transmitting ultrasound wave, page 15
$f_{\mathrm{car}}(\mathbf{x})$	Binary image containing the pixels belonging to the segmented carotid artery as its foreground, page 93
f_{MH}	Binary mask employed to determine the intensity of pixels belonging to the hyper-echoic regions in $f(\mathbf{x})$, page 89
$f_{\mathrm{musc}}(\mathbf{x})$	Binary image containing the pixels belonging to the segmented muscles as its foreground, page 93
$f_{\mathrm{M}}(\mathbf{x})$	Monogenic signal, page 23

$f_{\mathrm{R}}(\mathbf{x})$ 2D signal after applying the Riesz transform to the band-limited signal $f_b(\mathbf{x})$, page 23

$f_{\mathrm{thy}}(\mathbf{x})$ Binary image containing the pixels belonging to the segmented thyroid gland as its foreground, page 95

$f_{\mathrm{trc\_cand}}(\mathbf{x})$ Binary image containing the pixels belonging to the segmented trachea as its foreground, page 91

$f_b(\mathbf{x})$ 2D bandlimited signal, page 23

$f_e(\mathbf{x})$ Even component of the monogenic signal $f_M(\mathbf{x})$, page 73

F_GT Set of pixels belonging to the bounding box containing the carotid artery segmented manually by the expert, page 97

F_GT Set of pixels belonging to the bounding box containing the carotid artery segmented manually by the expert, page 92

$f_o(\mathbf{x})$ Odd component of the monogenic signal $f_M(\mathbf{x})$, page 73

$f_r(x, y)$ 2D image obtained after multiplying $r(x, y)$ with $f(x, y)$, page 40

F_S Set of pixels belonging to the bounding box containing the carotid artery segmented by the Self-Learning based MODS method, page 97

F_S Set of pixels belonging to the bounding box containing the carotid artery segmented by the Self-Learning based MODS method, page 92

$f_{\mathrm{an}}(\mathbf{x})$ Binary image containing the pixels belonging to an-echoic regions as its foreground, page 62

$f_{\mathrm{avg}}(\mathbf{x})$ Average intensity inside a window $w(x, y)$, page 61

$f_{\mathrm{ho}}(\mathbf{x})$ Binary image containing the pixels belonging to hypo-echoic regions as its foreground, page 62

$f_{\mathrm{hr}}(\mathbf{x})$ Binary image containing the pixels belonging to hyper-echoic regions as its foreground, page 62

$f_{\mathrm{hypo}}(\mathbf{x})$ Binary image containing pixels belonging to the hypo-echoic regions as foreground that is obtained bu directly thresholding the local phase $\phi(\mathbf{x})$ of an input ultrasound image $f(\mathbf{x})$, page 89

$f_{\text{l\_ho}}(\mathbf{x})$	Result of applying connected component labelling to $f_{\text{ho}}(\mathbf{x})$, page 63
$f_Q(\mathbf{x})$	Quantized image, page 73
$G(u, v)$	BDCT of $g(x, y)$, page 42
$G(x, y)$	Gaussian distributed white noise, page 15
$g(x, y)$	Annotation free image obtained after multiplying $f_r(x, y)$ with $\bar{b}_a(x, y)$, page 41
$g_y(x_i)$	Intensity of pixels in $\mathcal{N}_x$, page 65
g_{yl}	Result of low pass filtering the signal $g_y(x_i)$, page 65
$H(f(\mathbf{x}))$	Hessian of function $f(\mathbf{x})$, page 60
$H(f_a(\mathbf{x}))$	Entropy of automatically segmented image $f_a(\mathbf{x})$, page 46
$H(f_m^k(\mathbf{x}))$	Entropy an image $f_m(\mathbf{x})$ segmented manually by the k^{th} segmentor, page 46
$H(f_s(\mathbf{x}))$	Entropy of a generic segmented (manual or automatic) image $f_s(\mathbf{x})$, page 46
$h_1(\mathbf{x})$ $h_2(\mathbf{x})$	anti-symmetric filters required to obtain $f_{\text{R}}(\mathbf{x})$ from $f_b(\mathbf{x})$), page 24
$H_{\text{E}}(u)$	Edge filter in frequency domain, page 21
$h_{\text{L}}(t)$	Ideal line filter in time domain, page 21
$H_{\text{L}}(u)$	Ideal line filter in frequency domain, page 21
$I(f_a(\mathbf{x}), f_m^k(\mathbf{x}))$	Mutual information, page 46
$I(x, y)$	2D Echogenicity model for tissue intensity at the spatial location, page 15
$I_{\text{avg\_m}}$	Average intensity of the m^{th} SRR, page 73
K	Total number of labels in an image after the application of connected component labelling , page 40
L	Maximum length of the Thyroid gland in mm under a longitudinal scan, page 4
L	Total number of clusters, page 73

$l_{\max}$	The class label that contains the maximum number of pixels associated with it in an image, page 40
M	Height of a 2D US image, page 33
N	Width of a 2D US image, page 33
N_x	Width of the window used to describe a Neighbourhood in the three category clustering based segmentation algorithm, page 61
N_y	Height of the window used to describe a Neighbourhood in the three category clustering based segmentation algorithm, page 61
n_{a}	number of pixels that have an average value above $f_{\mathrm{avg}}(\mathbf{x}) + \tau$, page 62
n_{b}	number of pixels that have an average value below $f_{\mathrm{avg}}(\mathbf{x}) - \tau$, page 62
n_{e}	number of pixels that have an average value $f_{\mathrm{avg}}(\mathbf{x}) : f_{\mathrm{avg}}(\mathbf{x}) - \tau \leq f_{\mathrm{avg}}(\mathbf{x}) \leq f_{\mathrm{avg}}(\mathbf{x}) + \tau$, page 62
$P1f$ and $P2f$	Projection operators applied on $f(\mathbf{x})$ in the POCS algorithm, page 42
$p_s(f(t), \sigma)$	PDF of Rayleigh distribution for speckles in the signal $f(t)$ for a given σ, page 15
r	Slope of a digital line, page 73
$r(x, y)$	Binary image containing only the pixels with class label $l_{\max}$ as foreground pixels, page 40
$S(\mathbf{r}_m, \mathbf{r}_n)$	Sum of Squared Differences in pixel intensities contained in image patches $\mathbf{r}_m(k)$ and $\mathbf{r}_n(k)$, page 69
$s(x, y)$	Class label of pixel (x, y) after applying connected component labelling to $f(\mathbf{x})$, page 40
s_{mn}	Similarity between two patches of images centred at m^{th} and n^{th} speckle pixels in $\mathcal{F}_s$., page 69
t	Index of continuous time signal, page 15
$T(x, y)$	Acoustic Impedance at (x, y), page 15

U	24 hour radio-iodine uptake, page 4
u	Angular frequency in radians, page 21
v	Angular frequency in radians, page 21
$V(x, y)$	Amplitude of envelope detected echo at (x, y), page 15
V_{lobe}	Volume of individual thyroid lobe, page 4
V_p	Volume of thyroid gland estimated using Planimetry method, page 5
V_{CT}	Volume of thyroid gland manually estimated from a CT scan, page 29
W	Maximum width of the Thyroid gland in mm under the transverse scan, page 4
$w(x, y)$	Windowing function used to describe a Neighborhood around the pixel at (x, y), page 61
x	Index of a discrete time signal, page 15
$x_{\text{c\_car}}$ and $y_{\text{c\_car}}$	The row and column indices of the centroid of the segmented carotid artery $f_{\text{car}}(\mathbf{x})$, page 95
x_{y_m}	The location of maximum gradient value in g_{yl}, page 65
y	Index of a discrete time signal, page 15
$\mathcal{F}_{\text{Neg\_Hyper}}$	Set of hyper-echoic pixels belonging to the negative training set, page 95
$\mathcal{F}_{\text{Neg\_Hypo}}$	Set of hypo-echoic pixels belonging to the negative training set, page 95
$\mathcal{F}_{\text{Neg}}$	Set of pixels belonging to the negative training set, page 95
$\mathcal{F}_{\text{Pos}}$	Set of pixels belonging to the positive training set, page 95
$\mathcal{F}_{\text{Train}}$	Set of all training samples, page 95
$\mathcal{F}_s$	set of speckle related pixels, page 61
$\mathcal{H}(\S)$	Heaviside function, page 73
$\mathcal{N}$	Set of all pixels belonging to the nodes of the skeleton that is obtained by applying morphological skeletonization algorithm on $f_{\text{ho}}(\mathbf{x})$, page 63

$\mathcal{N}_x$	Set of all non-zero pixels in a single column of $f_{\text{hr}}(x, y)$, page 65
$\mathcal{T}$	Set of intensity peaks used to determine the pixels associated with annotation, page 40
$\mathbb{E}$	Expected index value used to determine NPR for qualitative image analysis, page 45
DoG	Difference of Gaussian, page 22
E12	Manual segmentation by expert 1 used as GT to quantify the segmentation by expert 2, page 53
E21	Manual segmentation by expert 2 used as GT to quantify the segmentation by expert 1, page 53
exp	Exponential function, page 15
FC	Phase congruency or feature assymmetry, page 73
h(t)	Analytic filter in 1D obtained from line and edge filters, page 21
PD	Publicly available dataset, page 33
PDF	Probability Density Function, page 15
PPD	Dataset used to test the pre-processing algorithms, page 34
S1	Dataset acquired at Tan Tock Seng Hospital, Singapore, page 33
S2	Dataset acquired at Mt. Elizabeth Hospital, Singapore, page 33
sin	sine function, page 15
SS	Dataset used to train and test supervised segmentation algorithms, page 33
V	Volume of Thyroid gland in millilitres, page 4

Chapter 1

Introduction

In an era dominated by changing lifestyles, the rate at which the general world population is contracting diseases of various sorts is increasing by the day. The rate is so alarming that governments around the world have reserved a very large proportion of their financial budget to improve healthcare. Such is the importance of innovation in medical technology (MedTech) in recent times. Innovation in MedTech can be anything from a novel way to dispose surgical needles to complex automated robots for surgery that incorporate sophisticated algorithms as a part of its workflow.

There has been a significant effort in the research community over the last couple of years to streamline the process of MedTech innovation under one banner called the "Biodesign" process [1]. The Biodesign process starts with the identification of a clinical need followed by the innovation stage that deals with ideation and brainstorming and ends with the implementation of the idea. Adhering to the stages outlined in the biodesign process, theme of the research work is in the Otorhinolaryngology field of medicine and the clinical need is to devise automatic methods to segment multiple organs in ultrasound images of thyroid gland.

1.1 The thyroid gland

The thyroid gland is an endocrine gland that is responsible for iodine metabolism and hormone regulation in the human body. It is a H-shaped organ located at the base of the neck and is surrounded by the trachea posteriorly, the parathyroid glands, the strap muscles, the longus colli muscle, the carotid artery, internal jugular vein and the neuro vascular bundle. The gland has two lobes that are connected together by a bridge of soft tissue called isthmus. The size and weight of the thyroid gland are influenced greatly by the lifestyle and gender. Under normal circumstances the thyroid gland weighs about 30gms and is about 4-6cm long and about 3cm wide [2]. Figure 1.1 shows the location of the thyroid gland at the base of the neck along with the surrounding anatomical structures.

Screening the thyroid gland for disorders can be done using any of the following imaging modalities: (1) Ultrasound (2D/3D) (2) Scintigraphy (3) Positron Emission Tomog-

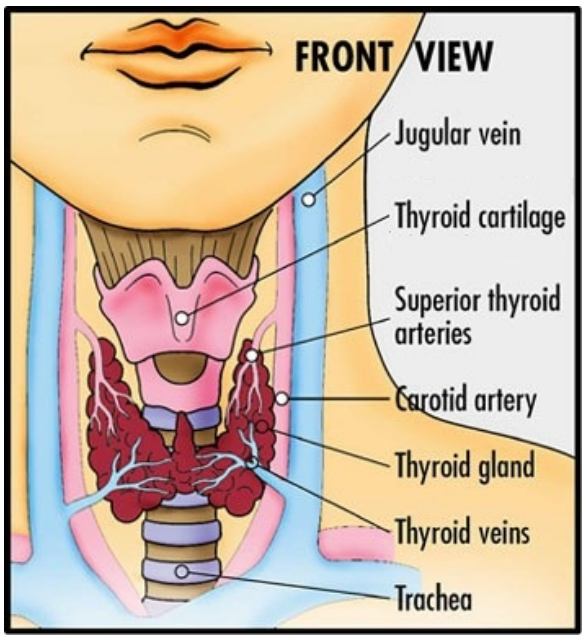


Figure 1.1: Thyroid and its surrounding anatomical structures

raphy (PET) (4) Computed Tomography (CT) and (5) Magnetic Resonance Imaging (MRI). Of all the modalities, ultrasound is the most preferred choice amongst clinicians and radiologists as it is portable, does not require any ionising contrast agents, radiation free and easy to learn. Scintigraphy measures the activity of the thyroid gland after the gland has been infused with a radionuclide, usually radio-iodine I-123 (^{123}I) or radio-iodine I-131 (^{131}I). A PET scan is also used to observe the activity of the thyroid gland. CT and MRI scans are also popular choices to scan the thyroid gland but are expensive and not as portable as ultrasound imaging devices. In principle, ultrasound imaging is used as a first level screening imaging modality and CT/MRI are used later to confirm the diagnosis. Ultrasound imaging is still a popular tool to perform guided interventions such as Fine Needle Aspiration Biopsy or Cannulation of the Internal Jugular Vein [3]. Figure 1.2 shows the thyroid gland seen under each of the imaging modalities[1] just discussed. Figure 1.3 shows the ultrasound image of the thyroid gland with the boundaries of the different organs visible marked by an experienced sonographer.

1.2 Motivations

1.2.1 Volumetric analysis

Disorders of the thyroid gland such as Goitre, Graves Disease, Hashimoto's Thyroiditis and Thyroid Carcinoma [4] are caused due to imbalance in Iodine uptake. The disorders result in the enlargement of the gland and the differential diagnosis of the disorders is

[1]Source: www.thyroidimaging.com

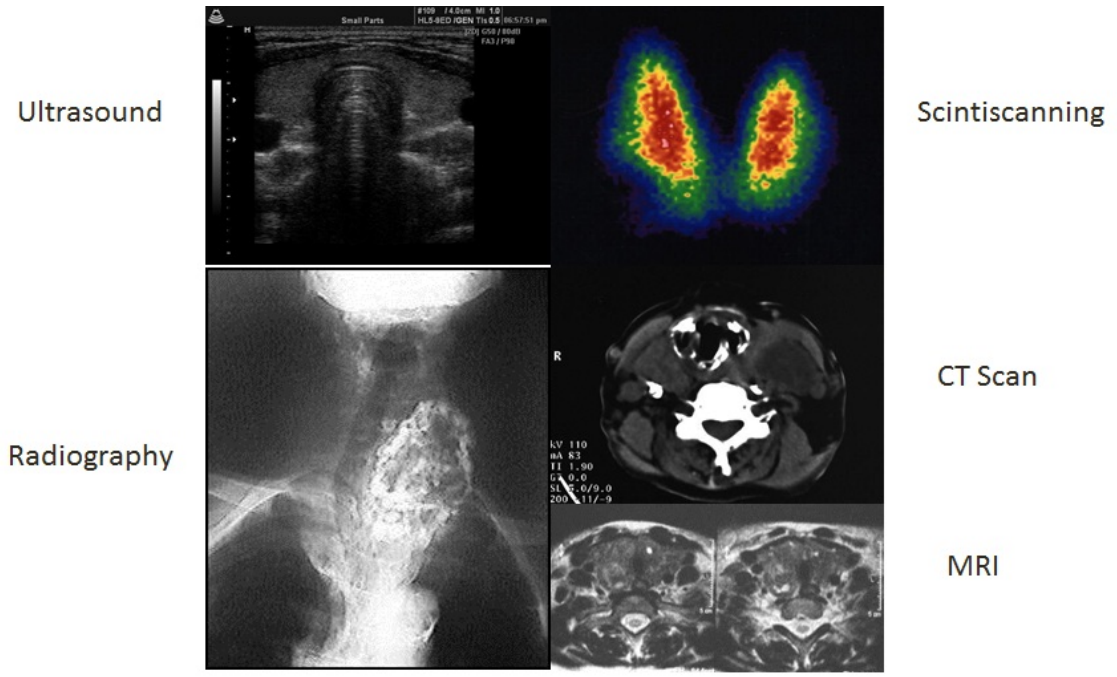
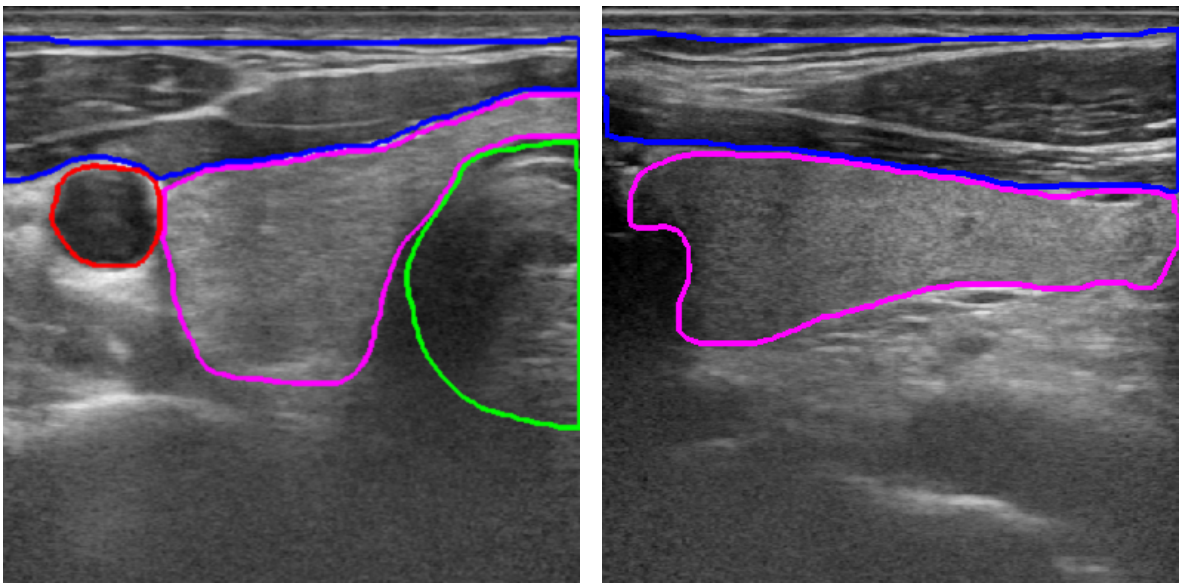


Figure 1.2: Imaging modalities used to screen the thyroid gland for disorders



(a) (b)

Figure 1.3: Ultrasound image of the right lobe of the thyroid gland in (a) transverse and (b) longitudinal scans, depicting the boundaries of the thyroid (in purple), the carotid artery (in red), the muscles (in blue) and the trachea (in green) that have been manually delineated by a trained sonographer.

done based on the volume of the diseased gland when compared to that of the normal gland. The volume of the normal thyroid gland is dependent on the body weight, height, age and gender of an individual and can vary between 5 to 20ml [3] . Not only is the estimate on the thyroid volume used to diagnose the type of disease affecting the thyroid gland [5] it is also used to calculate the dosage requirements of ^{131}I, that is used to treat graves disease [6]. The dosage of ^{131}I is calculated using the formula of Equation (1.1):

$$D = V \times [100\%/U] \times C \tag{1.1}$$

where D is the ^{131}I dosage , V is the volume in millilitres of the thyroid gland which is manually estimated using any modalities described in Section 1.1, U is the 24 hour radio-iodine uptake and C is a constant [6]. If the volume is over-estimated, then the dosage of ^{131}I given is higher which may subsequently lead to carcinoma in the thyroid gland. In the event of underestimation, the probability of recurrence of the disorder is higher. So it is important that the volume of the thyroid gland be estimated as accurately as possible. There are three different methods to estimate the volume of the thyroid gland depending on the modality of imaging used to screen the thyroid gland. They are:

(1) The ellipsoid method [7]

(2) The planimetry method [8]

(3) The cavaliari method [9]

1.2.1.1 The ellipsoid method

The ellipsoid method is the preferred method to estimate the volume of the thyroid gland when ultrasound is used to scan the gland. This method was first proposed by Brown et al., [7] where the shape of the thyroid gland is approximated to be an ellipse and the volume of each lobe is calculated by Eq. (1.2):

$$V_{\text{lobe}} = \frac{\pi}{6} \times L \times W \times D, \tag{1.2}$$

where L is the maximum length of the gland when observed under a longitudinal scan, W is the maximum width of the gland under the transverse scan and D is the maximum anterio-posterior height of the gland under a transverse scan. The parameters of Eq. (1.2) are measured manually by a technician or a doctor with the help of the digital calipers present in the ultrasound workstation. Figure 1.4 shows the measurements made on an ultrasound image of the thyroid gland. The final thyroid volume is obtained by adding the volumes of the two lobes.

1.2.1.2 The Planimetry method

This method is employed when MRI or CT modalities are used to scan the thyroid gland [8]. In this method, the surface area of the gland in each slice of the gland is

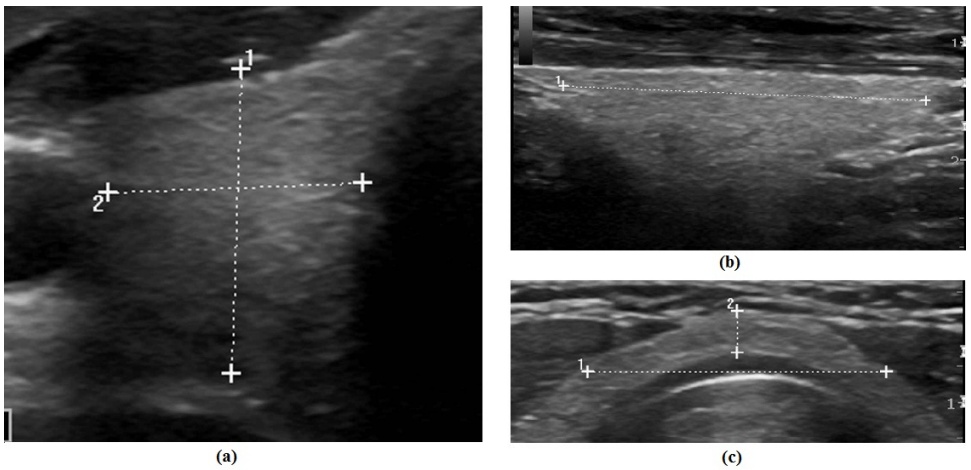


Figure 1.4: (a). Measuring the maximum width (W) and depth (D) of a thyroid gland on a transverse Ultrasound scan of the gland; (b). Measuring the maximum length(L) of a thyroid gland on a longitudinal Ultrasound scan of the gland; (c). Measuring the dimensions of the isthmus on a transverse scan of the gland. Source: [10]

calculated manually and is then multiplied by parameters such as the slice thickness and slice gap to obtain an estimate on the volume of the thyroid gland. This method is applicable to all modalities that facilitate 3D imaging of the gland, including the 3D Ultrasound imaging modality. In the case of 3D ultrasound imaging modality, the volume is calculated according to Equation (1.3) [11]:

$$V_p = \frac{\pi}{6} \times D_l \times 4\frac{A_l}{\pi \times D_l} \times 4\frac{A_t}{\pi \times D_t} \tag{1.3}$$

where, D_l and D_t are the maximum depth in the longitudinal and transverse sections of the scan respectively, A_l and A_t are cross-sectional areas in the longitudinal and transverse planes, respectively.

1.2.1.3 The Cavaliari method

Similar to the planimetry method of Section 1.2.1.2, the Cavaliari method is employed to estimate the volume of the thyroid gland when it is imaged using 3D modalities such as the CT imaging or MRI. In this method, a grid of equidistant test points are overlaid on the image and the points that are close to the boundary points of the thyroid gland (manually marked) used to determine the surface area of the thyroid gland in each slice of the 3D volume. The volume of the thyroid gland is estimated by adding the surface areas of every slice [9]. Figure 1.5 illustrates the thyroid volume estimation using the Cavaliari method on MRI volume of the thyroid gland. The test grid is used as a reference to compute the surface area in square centimetres\millimetres.

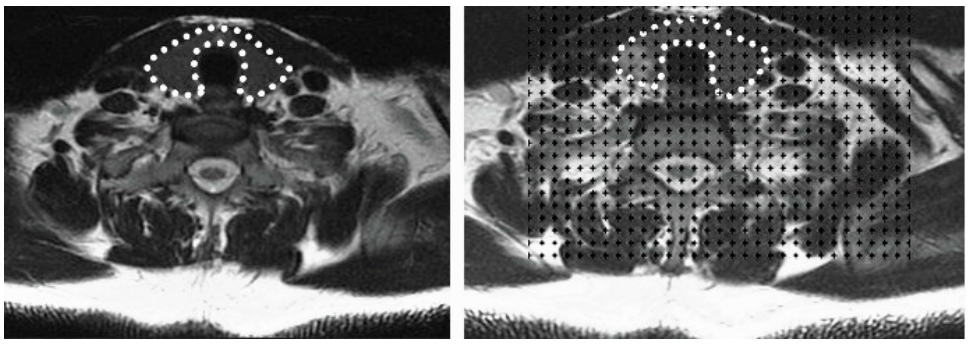


Figure 1.5: Illustration of application of the Cavaliari method on two consecutive slices of MRI images of the thyroid gland. The white dots denotes the boundary of the gland manually outlined by an expert radiologist. Source: [9]

1.2.1.4 Alternates to the ellipsoid method

Owing to the popularity of ultrasound imaging to screen the thyroid gland, the ellipsoid model of Section 1.2.1.1 is often used to estimate the volume of the thyroid gland. This method is however known to underestimate the thyroid volume [6, 9, 12–14]. There has been a constant effort in the scientific community to determine the cause for this and come up with better ways to estimate the volume of the thyroid. While some argue that the inter-observer variability plays a major role in the estimation [11], there are others [15] who argue that the $\frac{\pi}{6} = 0.524$ factor of Eq.(1.3) is responsible for the error in estimation. Based on the work by Brunn et al., [15] the World Health Organization has recommended the use of 0.479 instead of 0.524 as the correction factor in Eq. (1.3). Shabana et al., [16] suggest that any value in the range $[0.494, 0.554]$ can be used as the correction factor for the bias in Eq.(1.3) to get a good estimate of the thyroid volume. Ruggeri et al. [17] and Trimboli et al., [18] have proposed the linear model of Equation (1.4)

$$\text{Predicted Volume} = 1.24101 \times V_{\text{lobe}} + 3.66287 \tag{1.4}$$

where V_{lobe} is the volume of the gland estimated using Eq.(1.2) or by:

$$V_{\text{lobe}} = \text{Area}_{\text{T}} \times \text{Length}_{\text{L}}. \tag{1.5}$$

to predict the volume of the thyroid gland. The constants in Eq. (1.4) are generated by using the SAS system software [2]. The authors in [17] have reported an accuracy of 43% using this method by comparing the predicted volume estimates to the post surgical estimate obtained using the Archimedes Principle.

Even with such improvements over the ellipsoid method, the volumes estimated are still reported to be underestimates of the actual thyroid volume (obtained by applying Archimedes Principle only in cases where the thyroid gland was surgically removed) [9]. This is because, US imaging, in general, requires several parameters that needs to be set before acquiring the images and if the settings are off by a small margin, the boundaries

[2] `http://www.sas.com/en_us/software/analytics/stat.html`

of the tissues may appear diffuse in the display. Under such circumstances, the lower boundary of the thyroid gland may not be seen and the operator will have to rely on his/her experience to approximate the boundary to estimate the volume. Thus, whatever be the model used, the volumes will be underestimated as long as the boundary of the thyroid gland is manually detected. One possible way to overcome this will be to devise methods that can automatically detect the boundary of the thyroid gland and highlight the boundaries on the display.

1.2.2 Guided interventions

The use of ultrasound to perform guided interventions is discussed in this Section. The terms guided intervention and procedure are used interchangeably in the rest of the chapter and refers to the insertion of a sterile needle into the skin under the ultrasound probe. Depending on the type of intervention and the organ of interest (nodule\vessel), there are four methods of ultrasound guidance [19, 20]:

(1) Indirect method: In this method, the ultrasound probe is first positioned on the neck and a mark is placed on the skin at the location that corresponds to the organ of interest. The probe is then removed and the procedure is performed without the visual aid provided by the ultrasound display.

(2) Direct method: Contrary to the indirect method, the entire procedure is monitored continuously on the ultrasound display in this method.

(3) Mechanical guides: In this method, an external fixture is attached to the ultrasound probe. The fixture has a slot through which the needle can be inserted. This method is known to be cumbersome and is seldom used to perform guided interventions although the accuracies reported with this method are close to 100%.

(4) Free-hand method: This method is similar to the direct method and differs only in the way the needle is inserted into the skin.

Of the four methods, the free-hand method is preferred by the medical community to perform guided interventions [21]. Ultrasound is used as a guide to: (a) perform biopsy; and (b) catheterize the internal jugular vein in the head and neck region of the human body.

1.2.2.1 Fine Needle Aspiration Biopsy (FNAB)

This is a test that is performed to assess the malignancy of thyroid nodules. Thyroid nodules are abnormal lumps that develop in the thyroid gland. The majority of thyroid nodules are benign cysts or swellings caused by thyroid inflammation. In FNAB, a needle is used to extract some cells from the thyroid nodule. The location on the skin where the needle needs to be inserted is determined by placing the ultrasound probe over the region where the lumps are present and visualizing the nodule on the display in real-time. The

sampled cells are then sent to a pathologist to determine whether the cells are malignant. If the cell yield is sufficient, FNAB is an accurate method to differentiate benign from cancerous nodules. However, in 20% of patients, either the cell yield is insufficient to obtain a diagnosis or the pathology report does not allow the diagnosis of cancer to be completely excluded. Incorrect parameter settings at the time of acquisition which leads to diffuse tissue boundaries is most likely the cause for this. A better cell yield can be obtained if the boundary between the thyroid gland and the nodule is clearly demarcated automatically and displayed to the operator performing the biopsy in real-time. This requires the thyroid gland to be detected and segmented first followed by the application of methods to detect and segment the nodules within the segmented thyroid gland.

1.2.2.2 Cannulation of the internal jugular vein (IJV)

Ultrasound guided cannulation is defined as "ultrasound scanning being performed to verify the presence and position of a suitable target vessel before skin puncture followed by real-time ultrasound imaging to guide the needle tip throughout the vessel puncture process" [22]. Cannulation of vessels is necessary to obtain central venous access for haemodialysis, administration of antibiotics\fluids and for haemodynamic monitoring. Owing to the anatomical location of the IJV and the ease with which it can accommodate catheters of various sizes, the IJV is the preferred vessel to obtain central venous access [23]. However, its proximity to the common carotid artery results in complications such as laceration of pleura, subclavian artery puncture, hematoma, catheter malposition, fragment embolization, air embolism, venous thrombosis, and infection are included, some of which may cause death if not treated immediately [24–29]. Such complications can be reduced if the carotid artery is automatically detected, segmented and tracked in real-time during the procedure.

1.3 Problem statement and Objectives

When examining a thyroid gland with ultrasound, a physician (an Otorhinolaryngologist) makes a note of the following features visible in the image: (a) tissue echogenicity; (b) vascularity of blood vessels; and (c) size of the anatomical structures of interest. The detection of such features in a software setting that help a doctor to make a diagnosis is termed as Computer Aided Diagnosis (CAD) . While the tissue echogenicity and the size of anatomical structures provide information on the malignancy risk of the disorder, the vascularity of blood vessels tells us if there are any clots in the vessel. The size of the anatomical structures also help in determining the kind of treatment that needs to be administered to cure the disease once the diagnosis is made. For example, from the discussion in Section 1.2.1, the length, width and depth of the thyroid gland are needed to estimate the volume of the thyroid gland which is then used to determine the amount of radioiodine that needs to be administered to treat thyroid disorders such as Goitre etc. Automatically estimating the volume of the thyroid gland will eliminate the subjectivity

associated with manual volume estimation and helps in determining the right quantity of medication that is needed for the therapy. This application of ultrasound imaging is called diagnostic imaging. Apart from diagnostic imaging, ultrasound is used to perform guided interventions such as cannulation of interal jugular vein to obtain central venous access. It was seen in Section 1.2.2.2 that accidental puncture of the common carotid artery is one of the most common complications of this procedure. Automatically detecting, segmenting and tracking the common carotid artery will help in preventing the complications that arise during guided interventions using ultrasound.

While there are automatic methods proposed to: (a) segment the thyroid gland for volumetric analysis; or (b) detect, segment and track the carotid artery for guided interventions; there is no single algorithm that performs both at the same time. Moreover, these methods are semi-automatic methods that require manual initializations to perform the operations. All algorithms make use of pre-processing algorithms to remove speckle noise and imaging artefacts. But a review into the literature on ultrasound image formation suggests that the echogenicity of a tissue is due to the speckles in the image and the imaging artefacts provide information regarding the characteristics of the tissue surrounding it. Further, the methods are not invariant to image acquisition settings and rely heavily on the brightness information present in the images to segment the organs. Automatic organ detection is absent in a majority of the algorithms.

The aim of this research work is to perform automatic landmark based multi organ detection and segmentation (MODS) in free hand 2D ultrasound images of the thyroid gland. The following objectives are identified to accomplish the goal:

(1) Identify suitable landmarks associated with organs that need to be segmented and eliminate the need for manual initializations to perform MODS.

(2) Devise methods to perform tissue characterization that are invariant to image acquisition settings.

(3) Detect and segment the thyroid gland, the common carotid artery, the muscles and trachea in 2D ultrasound images of the thyroid gland.

1.4 Original contributions

The contributions of this research work are:

- An in depth review of the state-of-the-art methods to: (a) perform binary and multi-organ segmentation; and (b) perform qualitative, quantitative and statistical validation of the segmentation results.
- A preprocessing algorithm to automatically remove annotations from ultrasound images and accurately restore the image. The restored images are of high quality with average PSNR $> 38dB$.

- Methods to perform echogenicity based tissue characterization using speckles and localization of organs using imaging artefacts as landmarks. The speckle related pixels are determined by the application of Hessian based blob detectors on the ultrasound image.

- Three new methods to perform MODS in 2D ultrasound images of the thyroid gland. Of the three algorithms, two are unsupervised and one is a semi-supervised algorithm. All three algorithms perform better than the state-of-the-art methods.

  Method 1: The speckle related pixels are clustered into three echogenic classes based on its relative brightness in a kernel. An energy based model is then used to detect the carotid artery in the labelled image (image whose pixels are labelled with the cluster number). The carotid artery is used as a landmark to detect and segment the remaining organs in the image.

  Method 2: An agglomerative clustering constrained with a similarity metric is proposed to cluster the speckle related pixels into an unknown number of echogenic classes. The total number of class labels generated at the end of the clustering determines the number of echogenic classes in the image. This process is viewed as a quantization of the speckle related pixels into levels that are determined by the tissue echogenicity. The enhancement artefact is then detected in the quantized image which is in turn used as a landmark to detect the carotid artery. The carotid artery is used as a landmark to detect the remaining organs. All organs are segmented by the application of region based local phase methods.

  Method 3: The speckle related pixels are classified into the respective echogenic classes by the use of local phase based methods. The carotid artery and the trachea are detected from the binary image containing the hypoechoic pixels as foreground pixels. The pixels of the carotid artery and trachea along with the pixels that lie in the region between them are used to train a random forests classifier to classify the remaining pixels that belong to the hyperechoic tissue into the thyroid gland and background pixels.

- The application of the proposed methods to perform guided interventions, volumetric analysis and computer aided diagnosis.

1.5 Organization of the thesis

The rest of the thesis is organized as follows:

Chapter 2 contains the background information necessary to understand the methods proposed in Chapters 4 and 5. This chapter also contains an in-depth review of state-of-the-art methods proposed to segment the thyroid gland and the carotid artery. The importance of speckle related pixels for tissue characterization and imaging artefacts for anatomical landmarks is discussed in this chapter.

Chapter 3 contains the details regarding the datasets and validation protocol used in our experiments. The novel preprocessing algorithm to automatically remove annotations

from ultrasound images is described in this chapter. State-of-the-art methods proposed to segment the thyroid gland and the carotid artery described in Chapter 2 are compared with each other using a multitude of metrics that determine the quality of segmentation results.

Chapter 4 will describe the unsupervised methods proposed to perform MODS using speckle related pixels and imaging artefacts for tissue characterization and organ detection, respectively.

Chapter 5 will describe the semi-supervised method for MODS in 2D ultrasound images.

Chapter 6 will describe the applications of the proposed MODS methods to perform volumetric analysis and guided interventions using 2D ultrasound images of the thyroid gland. Possible future works of this research are also discussed in this chapter.

Chapter 2

Background

In this chapter we introduce the concepts related to ultrasound image formation. The motivations to use speckles and imaging artefacts for tissue characterization and anatomical landmark detection are discussed in this chapter. Further, a review of the state-of-the-art detection and segmentation algorithms proposed for ultrasound images of the thyroid gland is also undertaken in this chapter. The importance of treating ultrasound images as monogenic signals and its application to perform tissue characterization is discussed at the end of the chapter.

2.1 Basics of ultrasound imaging

Ultrasound imaging deals with forming a brightness image of the echo that is scattered back from a tissue when an ultrasound wave is incident on it. A typical ultrasound imaging device has the following five functional blocks [30]: (a) Pulser; (b) Transducer; (c) Beamformer; (d) Signal Processor; (e) Display Processor and (f) Display Unit; as shown in the block diagram of Figure 2.1.

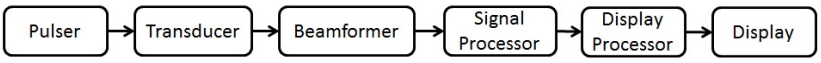


Figure 2.1: Block diagram of an Ultrasound Imaging System

The pulser block is responsible for activating a short burst of electric pulses at a centre frequency of about 2-10 MHz . The transducer then converts the input electric energy into a compression wave. Transducers can be one of the following types: (a) single element piston type; (b) multi-element annular type; (c) phased array; (d) linear array; and (e) curvature array type. State-of-the-art ultrasound scanners use linear or phased array transducers in 1D or 2D configuration to obtain a 2D image or 3D volume. The aperture of the transducer is additionally provided with a focussing unit to direct the ultrasound wave emitted by the transducer. The transducer is usually made of a piezoelectric material with a plastic covering. The depth to which an ultrasound wave

can penetrate is dependent on the frequency of operation. The higher the frequency of the transmitted ultrasound wave, the lower is the penetration depth. Ultrasound waves also undergo an exponential attenuation with increasing depth [31]. So the frequency of the transmit pulse has to be adjusted depending on the depth of the anatomic structure of interest under the probe. Often the transducers have a limit on the transmit frequency and if the depth of the organ is beyond the range of the transducer, another transducer with the desired range is used for scanning purposes.

While the pulser and transducer blocks of Figure 2.1 form the transmitter part of an ultrasound imaging system, the rest, except for the beamformer, are a part of the receiver system. The beamformer is responsible for focussing and steering the ultrasound wave during the transmit part of the imaging cycle and receive the backscattered echo after certain time delay. The beamformer uses several Time Gain Compensation(TGC) amplifiers and analogue to digital converters to receive the echo and reconstruct the signal. The received echo from the beamformer is then processed in the Signal Processing unit where it is filtered, envelope detected and log compressed. The reconstructed echo is rescaled according to the resolution of the display unit and suitable geometric corrections are applied in the display processing unit before being displayed on the CRT/LCD monitor . An operator can enhance the quality of the images by varying the TGC gains, Field of Views(FOV) , image compression parameters and envelope detection parameters.

2.1.1 Properties of Ultrasound Waves

Ultrasound waves undergo reflection, refraction, diffraction, scattering and absorption as it passes through different media. When an ultrasound wave strikes a boundary that is larger than its wavelength, it undergoes reflection. Depending on the velocity of the wave at the boundary between the different media and based on the angle of incidence, it undergoes refraction according to Snell's law. On the other hand, when an ultrasound wave encounters physical structures that are of a size comparable to its wavelength, it undergoes scattering. When the cells in a soft tissue are hit by an ultrasound wave, they gain energy and act as independent sources of ultrasound waves and scatter it in all directions. Such cells in the soft tissue are called scatterers. The ultrasound waves scattered by the scatterers undergo either constructive or destructive interference depending on its phase. This results in an overall increase/decrease of ultrasound energy that reaches back the receiver of the ultrasound machine. Fluid filled regions such as cysts or blood vessels appear dark in an ultrasound image as they have almost no scatterers present in it. The reflection and scattering of ultrasound waves are responsible for the general appearance of tissues. The refraction and diffraction of ultrasound waves results in the formation of artefacts in the image.

2.1.2 Tissue echogenicity

Echogenicity is defined as the ability of a tissue to reflect ultrasound waves. Echogenicity of a tissue directly translates to its brightness on the image. Referring to the ultrasound nomenclature in [32], a tissue being imaged by US is said to be: (1) *Anechoic*

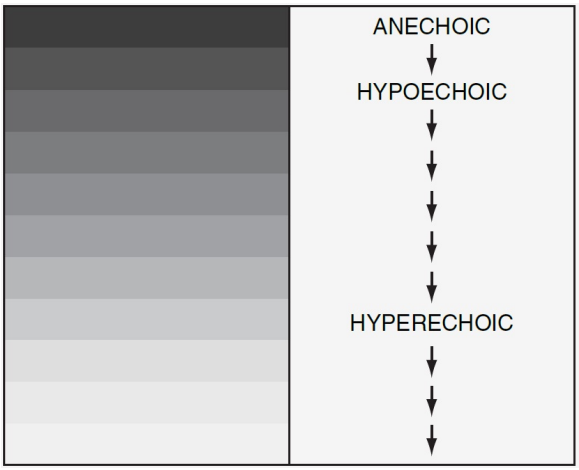


Figure 2.2: Echogenicity grading in an ultrasound image. *Source:* [32]

Tissue	Speed of sound (m/s)
Lung	600
Fat	1460
Aqueous humor	1510
Liver	1555
Blood	1560
Kidney	1565
Muscle	1600
Lens of eye	1620
Skull bone	4080

Table 2.1: Speed of sound in different tissues. Source: [34]

if it does not reflect any US waves; (2) *Hypoechoic* if it is less reflective compared to the neighbouring tissue; (3) *Hyperechoic* if it is more reflective compared to the neighbouring tissue; or (4) *Iso-Echoic* if it is as reflective as another tissue. Echogenicity is always defined with respect to the global brightness of a tissue, rather than the local brightness [2,32,33]. A hyperechoic tissue appears brighter than a hypoechoic tissue and iso-echoic tissues have the same brightness on a US image. Echogenicity as a grading of gray levels from Anechoic to Hyperechoic is shown Figure 2.2.

2.2 Speckle in ultrasound images

Ultrasound imaging is a coherent imaging modality, like that of lasers and the presence of speckles gives rise to the unique texture in US images. Speckle is defined as a deterministic, random pattern that appears in an image due to the interference between the ultrasound waves scattered by cells in a tissue that are of a size smaller than its wavelength. Modelling the statistics of speckles in ultrasound images has received a great deal of attention in the last couple of decades [35–37]. A detailed account on the statistics of speckles in ultrasound images can be found in [34] and [37]. While some argue that speckles follow a Rician [38] or the K-Distributions [39], the Rayleigh distribution is the widely accepted statistical model for fully developed speckles in ultrasound images [40]. The PDF of Rayleigh distribution for speckles is given by:

$$p_s(f(t), \sigma) = \frac{f(t)}{\sigma^2} \exp(\frac{-f(t)^2}{2\sigma^2}) \tag{2.1}$$

where, $f(t)$ is the signal intensity at t and σ is the shape parameter. The Rayleigh distribution can be employed to simulate speckles in an image by applying the convolution model [40–42]. According to this model, given $I(x, y)$ (echogenicity model for tissue intensities) and Gaussian distributed white noise $G(x, y)$, the acoustic impedance at (x, y) is given by:

$$T(x, y) = I(x, y) \times G(x, y). \tag{2.2}$$

The RF echo detected by the beamformer is then derived as:

$$V(x, y) = h(x, y) \otimes T(x, y), \tag{2.3}$$

where, $\otimes$ is the convolution operator and

$$h(x, y) = \sin(k_0 x) \exp(-x^2/2\sigma_x^2) \exp(y^2/2\sigma_y^2). \tag{2.4}$$

Here, σ_x and σ_y respectively represent the pulse-width and beam-width of the transmitting ultrasound wave. For a centre frequency of f_0, $k_0 = 2\pi f_0/\nu$, where ν is the speed of sound in a tissue. Table 2.1 lists the speed of sound in various tissues in the human body [34]. The amplitude of the B-mode envelope detected echo is given by:

$$f(x, y) = |V(x, y) + j\hat{V}(x, y)| \tag{2.5}$$

where, $\hat{V}(x, y)$ is the Hilbert transform of $V(x, y)$.

Figure 2.3 shows $f(x, y)$ generated with the following parameters: (a)$\sigma_x = 10$; (b) $\sigma_y = 15$; (c) $\nu = 1560$; and (d) $f_0 = 10$MHz. The echogenicity model used to simulate the speckle field has two regions R_0 and R_1 with pixel intensities equal to 0.2 and 1, respectively. The intensity histograms of the corresponding pixels in $f(x, y)$ are shown in Figures 2.3(c) and 2.3(d). It can be seen from the figures that the pixel intensities follow a Rayleigh distribution with its shape determined by σ_x and σ_y.

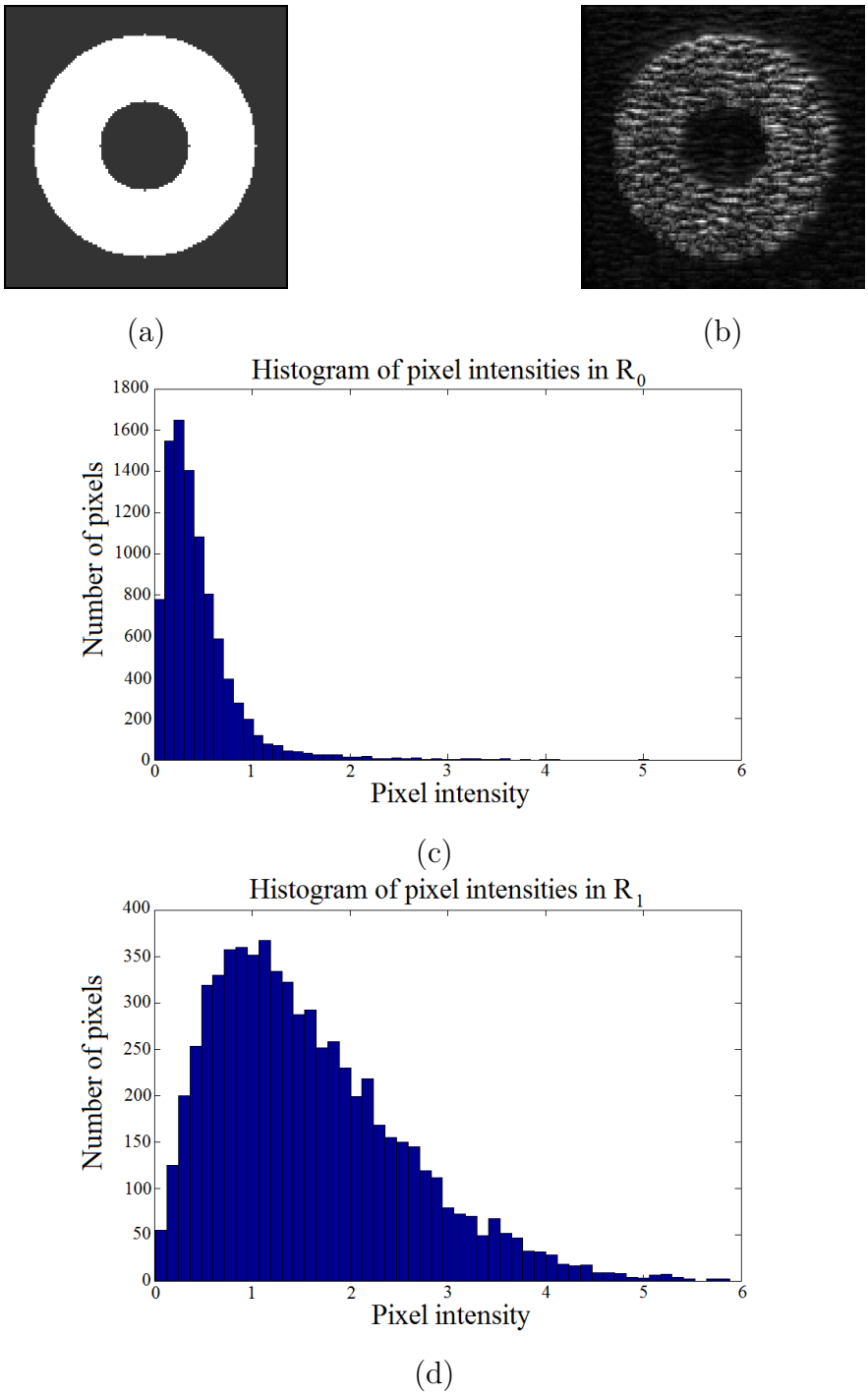


Figure 2.3: Simulating speckles using the convolution model. (a) Echogenicity model $I(x, y)$ with two regions R_0 and R_1 having intensities 0.2 (in gray) and 1 (in white), respectively; (b) B-mode amplitude image $f(x, y)$; (c) Histogram of pixel intensities in R_0 of $f(x, y)$; and (d) Histogram of pixel intensities in R_1 of $f(x, y)$

2.2.1 Role of speckles in ultrasound image analysis

Speckles in ultrasound image is always treated as noise and in a majority of segmentation algorithms, a first step is to filter the image to remove speckle noise. A majority of the filters are variants of the anisotropic diffusion filter [43]. One popular variant is the Speckle Reducing Anisotropic Diffusion(SRAD) filter introduced by Yu and Acton [42]. The wavelet diffusion filter [44] is another variant of the anisotropic diffusion filter which clubs wavelet decomposition with anisotropic diffusion to remove speckle noise. The Rayleigh-trimmed anisotropic filter [45] makes use of the statistical distribution of the speckles to despeckle the images. Eva et al., [46] propose yet another variant called gradient anisotropic diffusion filter to remove speckle noise from the images. Chinrungrueng and Suvichakorn [47] propose a modification of the Savitzky-Golay filter to eliminate speckle noise from ultrasound images. Apart from variants of anisotropic diffusion filters, wavelets [48] and adaptive weighted median filters [49] have also been used to filter ultrasound images to remove speckle noise. The gradient anisotropic diffusion filters and adaptive weighted median have been used to remove speckle noise in ultrasound images of the thyroid gland [46, 50].

2.2.2 Speckles for tissue characterization

Speckles appear as a direct consequence of the interaction between the scatterers in a tissue and the ultrasound waves incident on them. Therefore speckles in an ultrasound image are an embodiment of the tissue echogenicity. Thus, it would be inappropriate to treat speckles as noise and filter the ultrasound images to remove speckles.

Looking at the histograms in Figure 2.3(c) and 2.3(d), it can be seen that the peaks of the distributions in the two echogenic regions are well separated. This suggests that the pixels associated with speckles can be easily classified into their respective echogenic regions when the parameters of the distributions are known. This is the idea behind the Rayleigh Mixture Model proposed by Seabra et al. [51], to categorize plaques in ultrasound images. However, determining the parameters of the mixture model is a cumbersome task. In this research work, we explore unsupervised methods to classify the speckle related pixels into the respective echogenic regions for tissue categorization.

2.3 Artefacts in ultrasound images

In order to reconstruct the image at the beamformer it is assumed that: (a) the ultrasound wave travels in a straight line; (b) the tissue through which the wave travels is homogeneous; and(c) the velocity of the wave remains constant in the tissue. Artefacts are formed in an ultrasound image when these assumptions are not met [52]. The physics involved with the formation of imaging artefacts and its significance as a landmark to detect anatomical structures are discussed in this section.

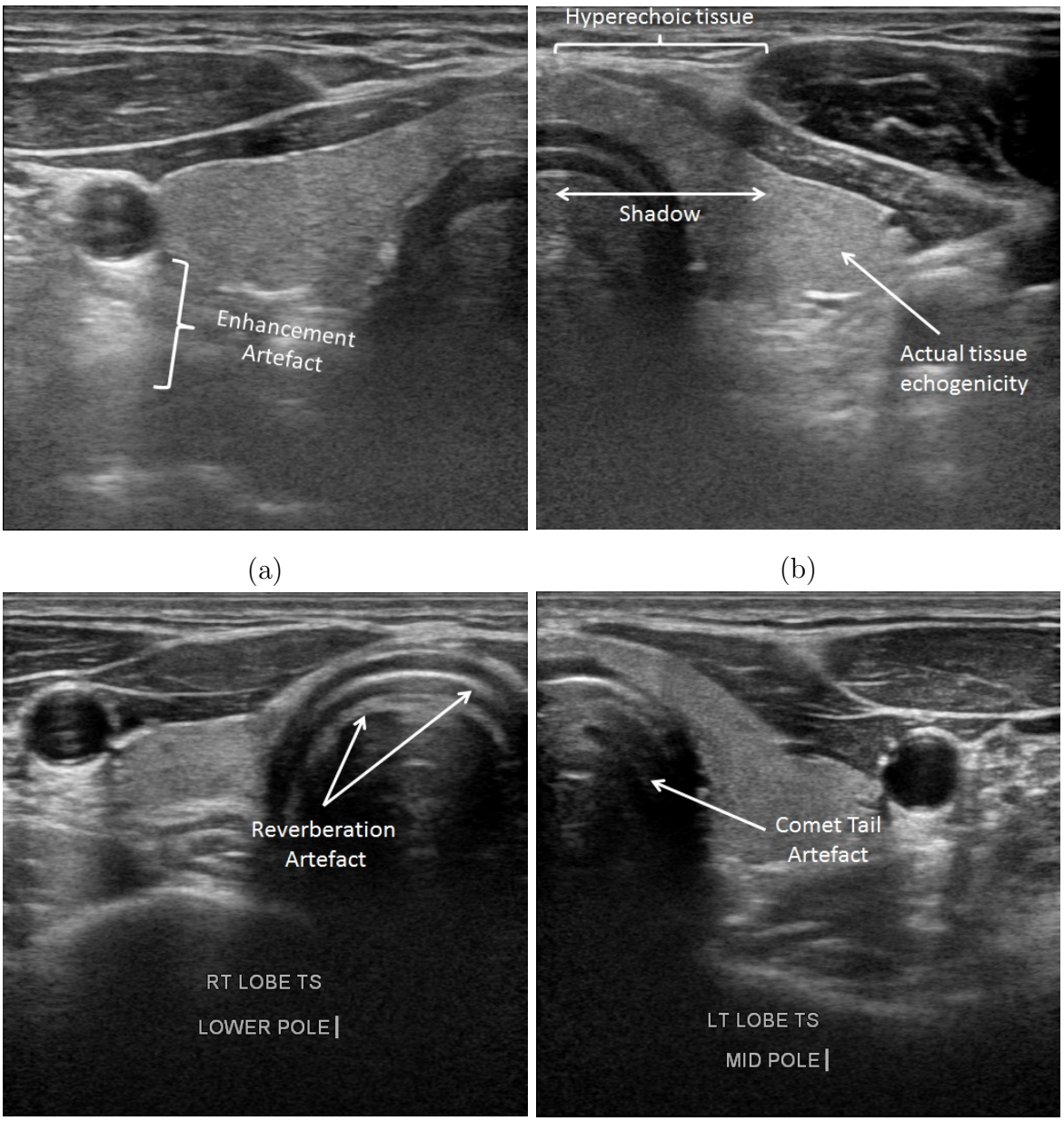


Figure 2.4: Artefacts in an ultrasound image. (a) Enhancement artefact; (b) Shadow artefact; (c) Reverberation artefact; and (d) Comet-tail artefact.

2.3.1 Formation

As an ultrasound wave travels through a tissue, a part of its energy is absorbed and dissipated as heat. Absorption of waves increases exponentially with depth. Thus the intensity of the received echo decreases with increasing depth. In order to compensate for the absorption loss, the TGC amplifiers are used at the beamformer to amplify the

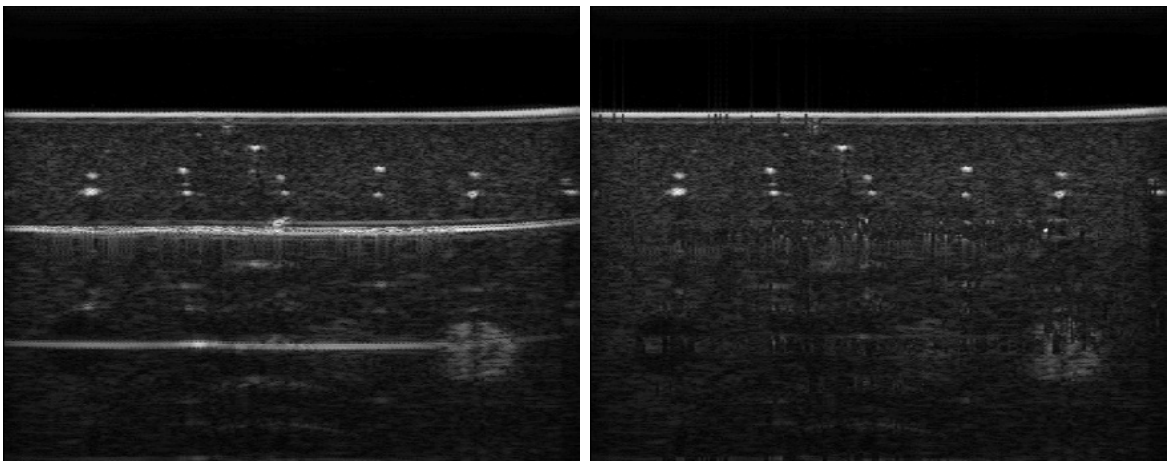
echos emanating from deeper levels. The absorption loss at a given depth is estimated based on the assumption that the tissue is homogeneous i.e., the echogenicity of the tissue remains constant between the transducer and the maximum depth permitted by the frequency of operation. Based on this assumption, the gains of the amplifier are fixed for a given depth. But at times when there is an anechoic structure present close to the transducer, the waves do not get absorbed at that depth and the energy gets transferred to the tissue underneath it. Under such circumstances, the intensity of the back scattered echo reaching the receiver is higher than anticipated. Subsequent amplification of the high intensity echo will result in the region directly below the anechoic structure to appear brighter than expected. Such regions of increased brightness under an anechoic or a hypoechoic structure is called an *enhancement artefact*. Enhancement artefacts are typically present under blood vessels in an ultrasound image. Figure 2.4(a) shows an enhancement artefact underneath the carotid artery in an ultrasound image of the thyroid gland.

Another artefact commonly found in ultrasound images is the *shadow artefact*. This type of artefact occurs when there is a hyperechoic tissue close to the transducer. In this case, the waves incident on the tissue get reflected back to the transducer. Reconstruction of the envelope detected echo will result in a shadow underneath the tissue which obscures other tissues below it. Shadow artefacts are also known to occur due to refraction of waves by a hypoechoic tissue. Figure 2.4(b) shows a shadow artefact under the hyperechoic fatty tissue in the ultrasound image.

Artefacts are also known to occur in an ultrasound image when there is a high impedance mismatch between the media through which the wave travels. The occurrence of *reverberation artefacts* and *comet tail artefacts* are associated with this phenomenon. When there is a high impedance mismatch between two media, the ultrasound wave reflected from the posterior boundary of the media gets reflected back into the same media upon return to the anterior boundary. This process repeats until there are no more reflections. The energy of the wave decreases with each reflection and there is a delay between the two reflections. This delayed reflection is interpreted as an echo appearing from a deeper region and is reconstructed as a ghost image under the reconstruction of the previous echo. This type of artefact appears as equidistant hyperechoic bands in an ultrasound image. These are termed as reverberation artefacts. Reverberation artefacts, when they take on a tapering shape are termed as comet-tail artefacts. Figures 2.4(c) and 2.4(d) shows reverberation and comet-tail artefact, respectively, inside the trachea in an ultrasound image of the thyroid gland

2.3.2 Role of artefacts in ultrasound image analysis

Similar to speckles, artefacts are considered as noise when performing ultrasound image analysis. One of the early attempts to automatically identify and remove artefacts was by Bylund et al., [54] where the reverberation artefact is assumed to be additive noise that is uncorrelated to the uncorrupted signal. Removal of the reverberation artefact is done on the *In phase Quadrature* (IQ) demodulated RF data at the beamformer. The



(a) (b)

Figure 2.5: Automatic reverberation removal in an ultrasound image [53]. (a). Input image with reverberation artefact; and (b). Image after the removal of reverberation artefact

authors implement an optimized Wiener filter to remove reverberation artefacts. Tay et al., [55] use wavelet decomposition on IQ data to remove reverberation artefacts. Here the IQ data is modelled as a sum of the artefact signal, the actual reflected wave and an additive zero-mean white Gaussian noise component. Win et al., [53] follow a different approach to identify and suppress the reverberation artefacts. The authors model the time of flight between successive echo to detect reverberation artefact. The regions corresponding to reverberation artefacts are replaced by the data that correlates well with the envelope detected signal of the adjacent echos. Figure 2.5 shows the results of reverberation artefact removal in a synthetic image by this method.

2.3.3 Artefacts as anatomical landmarks

From the discussion in Section 2.3.1, it can be inferred that the enhancement artefact and the shadow artefact are embodiments of the tissue directly above it. With respect to the ultrasound images of the thyroid gland, an enhancement artefact is always present under the carotid artery [2, 33]. The goal of this research work is to use the enhancement artefact as a landmark to detect the carotid artery in freehand 2D ultrasound image acquisitions of the thyroid gland.

2.4 Local phase in ultrasound images

In this section we demonstrate the relationship between the local phase of an ultrasound image and its echogenicity. The local phase of a signal and its spatial position are equivariant and it is invariant to the signal energy. This property of local phase is called

the *split of identity* and is a desirable feature for low level computer vision tasks such as detecting lines or edges in images [56]. An in depth analysis of local phase and its importance can be found in [57] and [58]. The concepts of local phase and its derivation from the analytic signal described in [57] are summarized in the next Section.

2.4.1 The analytic signal in 1D

The analytic function of a signal $f(t),\ t \in \mathbb{R}$, is given by:

$$f(t)_{\mathrm{A}} = f(t) - if(t)_{\mathcal{H}}, \tag{2.6}$$

where $f(t)_{\mathcal{H}}$ is the Hilbert transform of $f(t)$. The argument, $\arg f(t)_{\mathrm{A}}$, of the analytic signal represents the local phase, $\phi(t)$ of the signal. The phase of a signal provides information regarding the state of the signal in a local neighbourhood. In general, at any given instant t_0, the signal $f(t)$ is (i) maximum at t_0 if $\phi(t_0) = 2\pi k$; (ii) close to zero at t_0 and monotonically increasing from negative to positive value if $\phi(t_0) = \frac{\pi}{2} + 2\pi k$; (iii) close to zero at t_0 and monotonically decreasing from positive to negative value if $\phi(t_0) = -\frac{\pi}{2} + 2\pi k$; and (iv) minimum at t_0 if $\phi(t_0) = \pi + 2\pi k$. This property of local phase is used to determine the presence of lines and edges in a signal by convolving it with the appropriate line or edge filters.

An ideal line filter in time domain is given by:

$$h_{\mathrm{L}}(t) = -\delta(t+1) + 2\delta(t) - \delta(t-1). \tag{2.7}$$

Its Fourier Transform is given by:

$$H_{\mathrm{L}}(u) = 1 - 2\cos(u), \tag{2.8}$$

where, u is the angular frequency in radians. The line filter is related to the edge filter by the relation:

$$H_{\mathrm{E}}(u) = \begin{cases} -iH_{\mathrm{L}}(u) & u < 0 \\ iH_{\mathrm{L}}(u) & u \geq 0 \end{cases}. \tag{2.9}$$

The analytic filter consisting of both edge and line filters is expressed in complex form as:

$$h(t) = h_{\mathrm{L}}(t) - ih_{\mathrm{E}}(t) \tag{2.10}$$

Analytic filters that satisfy Eq.(2.9) and Eq.(2.10) are called *quadrature filters*. Convolving the input signal $f(t)$ with the analytic filter of Eq.(2.10) is the same as convolving the analytic signal $f_{\mathrm{A}}(t)$ of Eq.(2.6) with the real part of the analytic filter, i.e.,

$$f(t) \otimes h(t) = f_{\mathrm{A}}(t) \otimes \mathrm{Re}\{h(t)\}, \tag{2.11}$$

where, $\otimes$ refers to the convolution operator.

The line filter of Eq.(2.7) has a frequency higher than the spectrum of natural signals. Such filters are convolved with a Gaussian of the required bandwidth to fit its frequency in the spectrum of the signal. A natural signal may have lines and edges at different frequencies which leads to the need for the quadrature filter to be tuned at different bandwidths depending upon the spectrum of the signal. An alternate approach in computer vision is to rescale the signal (image) to a higher or lower resolution using the concept of image pyramids and apply a quadrature filter of a given bandwidth to the rescaled signal.

2.4.2 Quadrature band-pass filters and phase stability

Quadrature filters are essential to estimate the local phase in a signal as discussed in the previous sub-section. There are a number of quadrature filters proposed in literature for phase estimation. A comparative analysis of some of these can be found in [59]. For the sake of completeness, Table 3 in [59] that contains the equations of different quadrature filters along with its properties has been reproduced in Table 2.2. Quadrature filters designed to detect lines and edges should be insensitive to the DC component of the signal. Looking at the equations for the filters in Table 2.2, it can be observed that the Gabor filter and the Difference of Gaussians (DoG) filter has a non-zero DC value and are thus poor choices for filters to detect lines and edges in signals.

Other than the requirement for the quadrature filter to have zero DC value, the stability of phase in scale-space also affects the choice of the filter for phase estimation. The robustness of phase information can be studied by drawing phase diagrams that contain the response of the filters at high frequency (low scale) at the bottom of the image and low frequency (high scale) at the top of the image. A quadrature filter can be considered to be completely stable if it contains vertical lines only in the phase diagram [57]. The filters in [59] are compared based on their phase stability. It was shown that the family of filters similar to the Sarkar and Boyer's filter are optimal to detect edges and lines in images while the zero DC corrected Gabor filters are a good choice to measure image velocity. A similar study was undertaken in [60] to choose the filters for measuring image velocity and binocular disparity.

It was shown in [57] that the scaling of a signal to estimate the local phase results un an unpredictable behaviour at singular points i.e., the points where the analytic signal goes through the origin of the complex plane. This unwanted behaviour of phase can be avoided by selecting filters that are stable across all scales i.e., by using scale invariant filters for phase estimation. One such filter is proposed in [61] for local phase estimation in 2D signals.

2.4.3 The monogenic signal and local phase in 2D signals

We have so far discussed the local phase estimation in 1D signals by making use of quadrature filters and the 1D analytic signal. Extension of the same to 2D signals is not simple as the phase estimation methods for 1D signals do not take into consideration the orientation of the signal with respect to the axes. In the case of 2D signals, the

Filter	Fourier Domain	Bandwidth	Tuning Frequency
Gabor	$n_c\ \exp(-\frac{\sigma^2}{2}(u-u_0)^2)$	$\frac{\ln(\frac{\sigma u_0+\sqrt{2\ln 2}}{\sigma u_0-\sqrt{2\ln 2}})}{\ln 2}$	u_0
log-Normal	$n_c\ \exp(\frac{\ln^2(\frac{u}{u_0})}{2\ln^2(\kappa_\beta)})$	$-\frac{2\sqrt{2}}{\sqrt{\ln 2}}\ln(\kappa_\beta)$	u_0
Gaussian derivative	$n_c\ u^a\ \exp(-\sigma^2u^2)$	$\frac{\ln(\frac{W(-1,c)}{W(0,c)})}{2\ln 2}$ $c=-\frac{4^{-(1/a)}}{e}$	$\frac{\sqrt{a}}{\sigma}$
DoG	$n_c\ (e^{(-\frac{\sigma'^2u^2}{2})}-e^{(-\frac{\sigma^2u^2}{2})})$	$\star$	$\frac{2}{\sigma}\sqrt{\frac{\ln(\gamma)}{(\gamma^2-1)}}$ $\gamma=\frac{\sigma'}{\sigma}$
Cauchy	$n_c\ u^a\exp(-\sigma u)$	$\frac{\ln(\frac{W(-1,c)}{W(0,c)})}{\ln 2}$ $c=-\frac{2^{-(1/a)}}{e}$	$\frac{a}{\sigma}$
Deriche	$n_c\ u^a\frac{4\alpha^3}{(u^2+\alpha^2)^2}$	$\star$	$\alpha\sqrt{\frac{a}{4-a}}$
Sarkar and Boyer	$*$	$\star$	$\star$

* $2n_cu^a\alpha^2\tilde{m}\frac{(2\sin(\psi)-\tilde{m}\cos(\psi))u^2+\alpha^2(\tilde{m}\cos(\psi)+3\tilde{m}\cos(\psi)+2\sin(\psi))}{(u^2+\alpha^2)((u+\alpha\tilde{m})^2+\alpha^2)(u-\alpha\tilde{m})^2+\alpha^2)}$.

⋆ Analytic function does not exist. It can however be numerically approximated. Refer to [59] for more details.

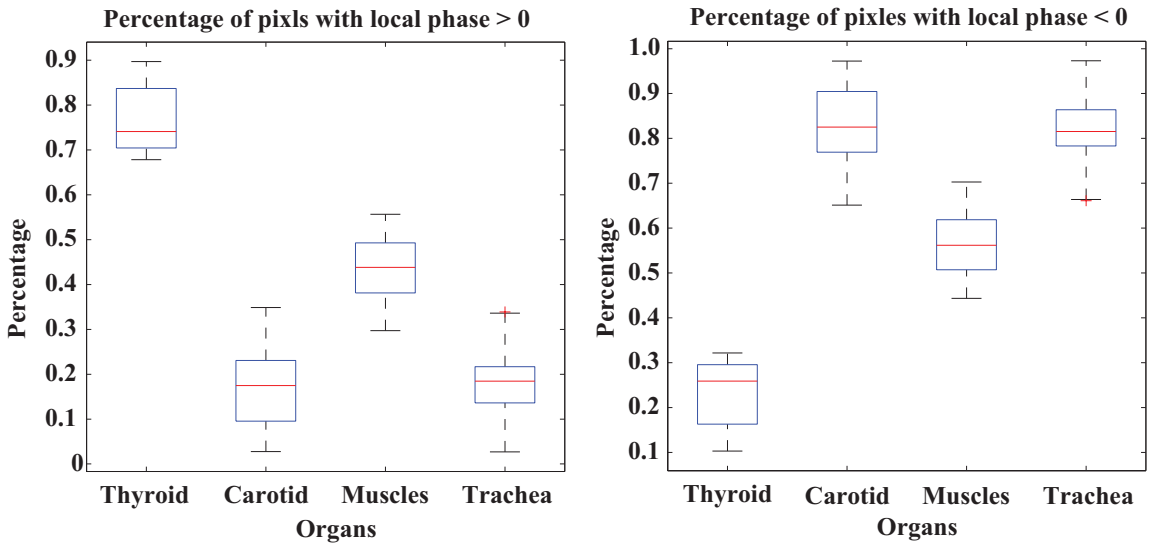
Table 2.2: List of quadrature filters along with its equations and properties.

local phase can be estimated by using steerable filters [62]. This approach is however not isotropic i.e., the energy of the signal will not remain constant when the orientation of the signal is changed and hence the split of identity property of the analytic signal is not satisfied. Felsberg and Sommer [56] proposed an analytic signal that retains the split of identity property in 2D by making use of Riesz transform which is a 2D generalization of the Hilbert transform. This analytic signal in 2D is called the *monogenic signal*. The algebra of *quaternions* is used to define a monogenic signal. The spatial location or the co-ordinates, $(x, y) \in \mathbb{R}$ ($\in \mathbb{Z}$ for images), of a 2D signal is denoted by the row vector $\mathbf{x}$ for the sake of simplicity. A monogenic signal is given by:

$$f_{\mathrm{M}}(\mathbf{x}) = f_b(\mathbf{x}) - (i, j)f_{\mathrm{R}}(\mathbf{x}) \tag{2.12}$$

where, $f_b(\mathbf{x})$ is the signal band-limited by using any of the quadrature filters mentioned in Section 2.4.2 and $f_{\mathrm{R}}(\mathbf{x})$ is signal $f_b(\mathbf{x})$ after applying Riesz transform. The Riesz transformed signal is obtained by convolving $f_b(\mathbf{x})$ with the following two anti-symmetric filters:

$$h_1(\mathbf{x}) = \frac{-x}{2\pi(x^2+y^2)^{\frac{3}{2}}} \text{ and } h_2(\mathbf{x}) = \frac{-y}{2\pi(x^2+y^2)^{\frac{3}{2}}}. \tag{2.13}$$



(a) (b)

Figure 2.6: Local phase in organs present in the ultrasound image of Figure 1.3. (a) Percentage of pixels with positive local phase; and (b) Percentage of pixels with negative local phase.

The transformed signal is given by:

$$f_{\mathrm{R}}(x,y) = \sqrt{(h_1(\mathbf{x}) \otimes f_b(\mathbf{x}))^2 + (h_2(\mathbf{x}) \otimes f_b(\mathbf{x}))^2}. \tag{2.14}$$

The local phase of the signal is then estimated as:

$$\phi(x,y) = \tan^{-1}\left(\frac{f_{\mathrm{R}}(\mathbf{x}))}{f_b(\mathbf{x}))}\right) \tag{2.15}$$

2.4.4 Echogenicity and local phase

While local phase is a popular feature to detect lines and edges in an image, it is used to describe regions in this research work. The box plots of local phase in each organ of Figure 1.3 is shown in Figure 2.6. From the plots it can be observed that the pixels corresponding to the carotid artery, muscles and trachea have negative local phase and those corresponding to the thyroid gland have positive local phase. The regions corresponding to hyperechoic tissues can be separated from the regions corresponding to the hypoechoic tissues by thresholding the local phase. This concept is used in Chapters 4 and 5 to perform MODS in ultrasound images.

2.5 Ultrasound image segmentation

Ultrasound image segmentation has been a hot topic of research for over two decades. Attempts at segmenting ultrasound images can be dated back to the early 80's when US imaging had just started to gain importance as a diagnostic device in a clinical setting. Since then, there have been tens of thousands of papers published on ultrasound image segmentation. A good insight into the sheer magnitude of research that has gone into this field can be found in a survey on US image segmentation algorithms by Noble and Boukerroui [63], and that was just until the year 2005. Eleven years have elapsed and the interest in the field has not lost its momentum yet. This can be attributed to the ever changing ultrasound imaging technology. The advances in transducer technology, beam forming technology etc., have led to the introduction of 2D+T, 3D and 3D+T US image acquisition systems along with improvements made to the baseline 2D US imaging system. It is this advance in imaging technology that necessitates the need to upgrade the segmentation algorithms to keep up with the technological growth.

US image segmentation algorithms can be categorized according to the clinical need that is being addressed like segmenting the ventricular boundaries in an Echocardiogram [64], segmenting lesions in the Breast [65], segmenting tumour in the Liver [66], segmenting the Prostate gland [67] or segmenting the Thyroid gland [49] and many more. Another way of categorization can be based on the underlying principle used to segment the images. Sonka et al., [68] outline the following five basic principles for generic image segmentation: (a). Mean Shift; (b). Geometric Deformable Models (GDM); (c). Graph based model (GBM); (d). Fuzzy connectivity; and (e) Optimal Single and Multiple Surface Segmentation. A detailed description of each of the basic principles can be found in [68].

A non-exhaustive list of the research work carried out from 2006 to 2016 in the domain of medical imaging is provided in Table 2.3. The list categorized according to the basic principle employed and the organ of interest. The focus here is on the organs that can be simultaneously be scanned by the three most popular imaging modalities: CT, MRI and Ultrasound. Accordingly the following organs are considered: Breast, Heart (Cardiac), Thyroid, Liver, Kidneys, Foetus, Pelvis and Prostate. In addition to US imaging, CT and MRI are also included in this study to highlight the invariance of the basic principles to the modality of imaging. From the table it is evident that Level Sets (of GDM) and Graph Cuts (of GBM) are the two most preferred algorithms to segment any organ of any modality. Both are semi-automatic methods that require manual initializations in the form of initial contours (Level Sets) or seed points for the foreground and background (Graph Cut) pixels to carry out the segmentation.

The shortcomings of the Level Set based methods such as that employed in [78] and graph cut based algorithms such as that used in [105] is that they are only effective in a single organ segmentation scenario. Unless the algorithms are re-initialized with a different set of initial contours (level sets) or seed points (graph cuts) for different organs, these methods cannot be employed in a multi-organ segmentation framework. The GDM and GBM based algorithms are inherently iterative algorithms that employ some form

Basic principle	Modality	Organs of interest								References
		B	C	T	L	K	F	Pl	Pr	
Mean shift	US/MRI/CT	✓	✓		✓				✓	[69] [70] [71] [72]
GDM (LS)	US/MRI/CT	✓	✓	✓	✓	✓	✓	✓	✓	[73] [74] [75] [76] [77] [78] [79] [80] [81] [82] [83] [84] [85] [86] [87] [88]
GDM (GAC)	US/MRI/CT	✓	✓	✓	✓	✓	✓		✓	[89] [90] [91] [92] [93] [94] [95] [96] [97] [98] [82] [99] [100] [101] [102]
GBM (NC)	CT/MRI		✓							[103] [104]
GBM (GC)	US/MRI/CT	✓	✓	✓	✓	✓	✓	✓	✓	[105] [106] [107] [108] [109] [110] [111] [112] [113] [114] [115] [116] [117] [118] [119] [120] [121] [122]
Fuzzy connectivity	US/CT/MRI	✓	✓		✓	✓	✓	✓		[123] [124] [125] [126] [127] [128] [129]
Optimal surface	CT		✓		✓	✓			✓	[130] [131] [132] [132]

GDM = Geometric Deformable Model; LS = Level Set; GAC = Geodesic Active Contours; GBM = Graph Based Model; NC = Normalized Cut; GC = Graph Cut; CT = Computed Tomography; MRI = Magnetic Resonance Imaging; and US = Ultrasound.
B = Breast; C = Cardiac; T = Thyroid; L = Liver; K = Kidney; F = Foetus; Pl = Pelvis; and Pr = Prostate.
✓indicates that there is at least one algorithm based on a particular basic principle for the corresponding organ of interest.

Table 2.3: Ultrasound image segmentation algorithms categorized according to the basic principle used.

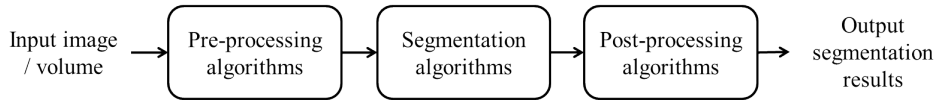


Figure 2.7: Workflow of segmentation algorithms proposed for ultrasound images of the thyroid gland.

of an energy/error minimization constraint to converge to the optimal segmentation at every iteration. Thus, re-initializing these algorithms for multi-organ segmentation will be computationally expensive.

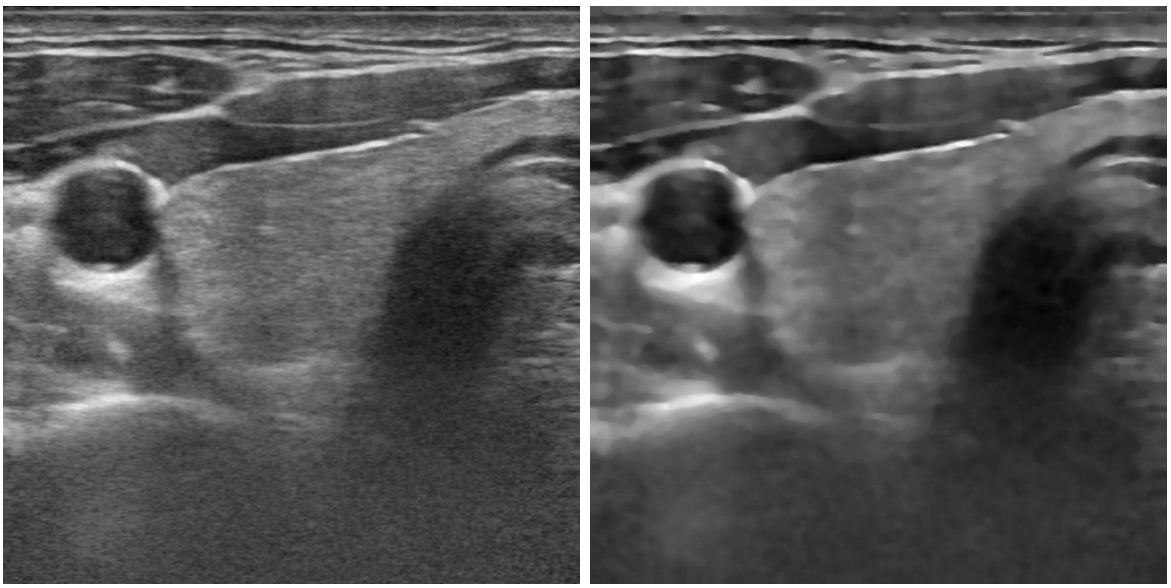
Returning to the list in Table 2.3, it can be seen that [46,78,81,94–96,98] are the only works related to segmenting the ultrasound images of the thyroid gland and conforming to the basic principles outlined in [68]. The rest are either for CT/MRI or for organs other than the thyroid gland. It can also be observed that except for GDM based methods, other basic principles are not employed for US images of the thyroid gland. The graph based methods in [114] and [115] are for CT volumes of the head and neck. However, there are algorithms proposed in literature that employ methods other than the basic principles previously listed to segment US images of the thyroid gland. These are discussed in detail in Sec.2.5.1.

2.5.1 Segmentation in ultrasound images of the thyroid gland

The workflow employed by a majority of segmentation algorithms devised for ultrasound images of the thyroid gland is illustrated in Figure 2.7. Given an input image, it is first preprocessed to remove speckles and imaging artefacts by methods listed in Sections 2.2.1 and 2.3.2, respectively. Segmentation algorithms are then applied on the preprocessed images to segment the thyroid gland or the carotid artery for volumetric analysis or guided interventions. The segmentation results are post-processed to remove segmentation leaks and the results are displayed for further analysis by the doctors. The various methods proposed to segment the ultrasound images of the thyroid gland and the workflow followed by each of the algorithms is discussed in this section. These are categorized according to their application being: (a) for volumetric analysis; or (b) for guided interventions.

2.5.1.1 For volumetric analysis

To date there have been three attempts to segment the thyroid gland for volumetric analysis. The first of the three attempts was by Kolloez et al. [46] where the segmentation is performed on 3D ultrasound volumes of the thyroid gland. The input volumes are first preprocessed by curvature flow filters to remove speckle noise. Geodesic active contours are then employed to segment the thyroid gland. The dimensions of the thyroid gland are then estimated from the segmentation results and the ellipsoid method (refer to Section 1.2.1.1 for details) is used to determine the volume of the thyroid gland. The volume estimates are compared to the manual volume determined from CT scans. The authors



(a) (b)

Figure 2.8: Results of filtering an ultrasound image by the adaptive weighted median filter (AWMF) with a neighbourhood of size 9×9 pixels. (a) Input image; and (b) Results of filtering by AWMF.

note that the thyroid volumes estimated from the ultrasound volume are an underestimate of the actual thyroid volume. The algorithm is semi-automatic as it requires manual initializations to segment the thyroid gland.

The remaining two methods proposed by Chang et al. [50] and Garg et al. [133], are supervised segmentation algorithms to segment the thyroid gland in 2D ultrasound images. Both methods differ only in the classifier used to segment the gland. In both methods, the ultrasound images are first pre-processed by the application of the adaptive median filter (AWMF) to remove speckle noise. Figure 2.8 shows the results of applying AWMF on an ultrasound image of the thyroid gland. A sliding window is then used to extract texture based features such as: (a) Co-efficient of Local Variation; (b) Block Difference of Inverse Probabilities; (c) Normalized Multi-Scale Intensity Difference;(d) Haar wavelet based features; and (e) histogram features from the preprocessed image. The features are used to train a radial basis function neural networks classifier (feed forward neural networks in [133]) to classify the pixels in the image into the thyroid and the non-thyroid classes. The pixels labelled as the thyroid gland are set as foreground pixels in the segmented image. The segmented image is post-processed by the application of morphological operations to remove false positive results. As before, the dimensions of the thyroid gland are estimated from the segmentation results. These are used to calculate the surface area of the thyroid gland which is then used to estimate the volume of the gland by the following equation:

$$V = a \times (A_L \times D_L + A_R \times D_R) + b \tag{2.16}$$

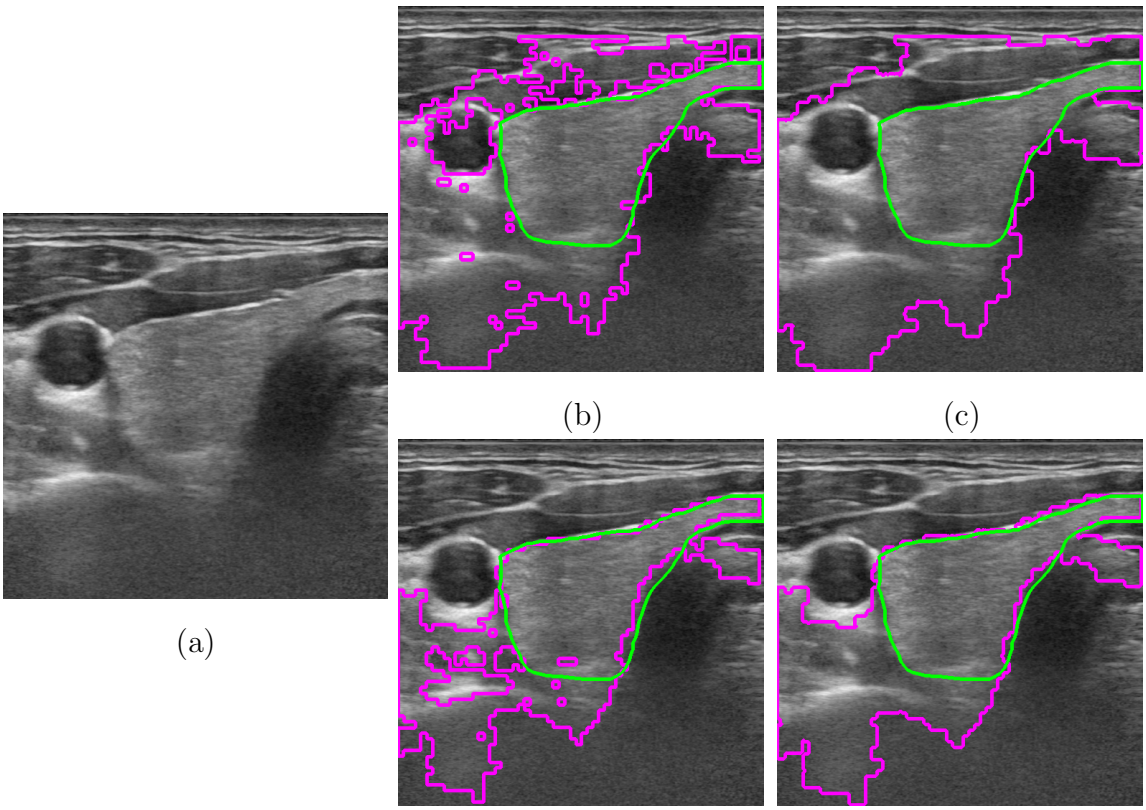


Figure 2.9: Results of segmenting the transverse scan of the thyroid gland by the method of [49] when: (b-c) the training class for the negative class includes pixels of the carotid only; and (d-e) the training class for the negative class includes pixels of both the carotid and muscles. (b,d) Results before post-processing; and (c,e) Results after post-processing.

where, A_L and A_R correspond to the surface areas of the left and right thyroid lobes, respectively, under the longitudinal scan and D_L and D_R correspond to the maximum depth of the gland in the right and left lobes, respectively, under the transverse scan. The parameters a and b are estimated by the Particle Swarm Optimization (PSO) using the fitness function:

$$Fitness = [a \times (A_L \times D_L + A_R \times D_R) + b] - V_{CT}. \tag{2.17}$$

The volume of the thyroid gland V_{CT} is estimated manually on a CT scan by the planimetry method 1.2.1.2.

Supervised segmentation algorithms require training samples of the positive class and negative class to train the classifiers. Often, the training samples for the positive class is derived from the ground truth labellings provided by experts and that of negative samples are randomly chosen from regions in the image that do not correspond to the positive

training samples in the training images. Since the negative class is not specific, the sensitivity of the algorithm to detect the negative samples in the test set will decrease. This is illustrated in Figure 2.9, where the segmentation algorithm of [50] results in two different segmentations when two different sets of training samples are used for the negative class. It can also be seen that this method tends to over segment the thyroid gland even after applying post-processing methods to prevent segmentation leaks. The performance of the supervised algorithms is subject to the quality of the training datasets and is in turn dependent on the inter-observer variability of the ground truth labellings. The nature of the features also dictate the performance of the segmentation algorithms. Since the features used in [50] and [133] are based on the intensity of the pixels, the algorithms are sensitive to the changes in the pixel intensity when the TGC is changed.

2.5.1.2 For guided interventions

Segmentation methods in the field of guided interventions mainly focus on segmenting and tracking the common carotid artery in 2D and 2D+T US images. A majority of the methods are variants of the star algorithm proposed by Friedland et al. [134] to detect and track ventricular cavity boundaries in 2D+t US images. In this method the ventricular cavity is modelled as an ellipse whose parameters are estimated by applying Hough 1D transform on edge pixels along pre-defined radial lines emanating from a seed point. A simple gradient based edge detector is used in this method to obtain the edge pixels. Abolmaesumi et al. [135] model the carotid as an ellipse and estimate the radius of the ellipse along each radial line using the Kalman filter. Guerrero et al. [136] use the parametric form of an ellipse to model the carotid artery and estimate the ellipse parameters using the extended Kalman filter. The methods in [135] and [136] are called the star Kalman and the star extended Kalman algorithms, respectively. The spokes ellipse algorithm proposed by Wang et al. [137] segments the carotid artery in 2D US images by fitting an ellipse using least squares to the edge pixels obtained using an edge function similar to [134]. The techniques described so far to segment the carotid artery are semi-automatic algorithms that require manual initializations in the form of seed points provided by expert sonographers. Figure 2.10 shows the results of applying the aforementioned algorithms on an ultrasound image of the thyroid gland. It can be observed that except for the Star Extended Kalman algorithm, the rest either under-segment or over-segment the carotid artery.

2.6 Summary

- The properties of ultrasound waves to undergo reflection and scattering are responsible for the tissue echogenicity in an ultrasound image. Whereas the properties of interference, diffraction and refraction result in the formation of imaging artefacts in ultrasound images.

- Speckles and imaging artefacts are embodiments of tissue echogenicity and can re-

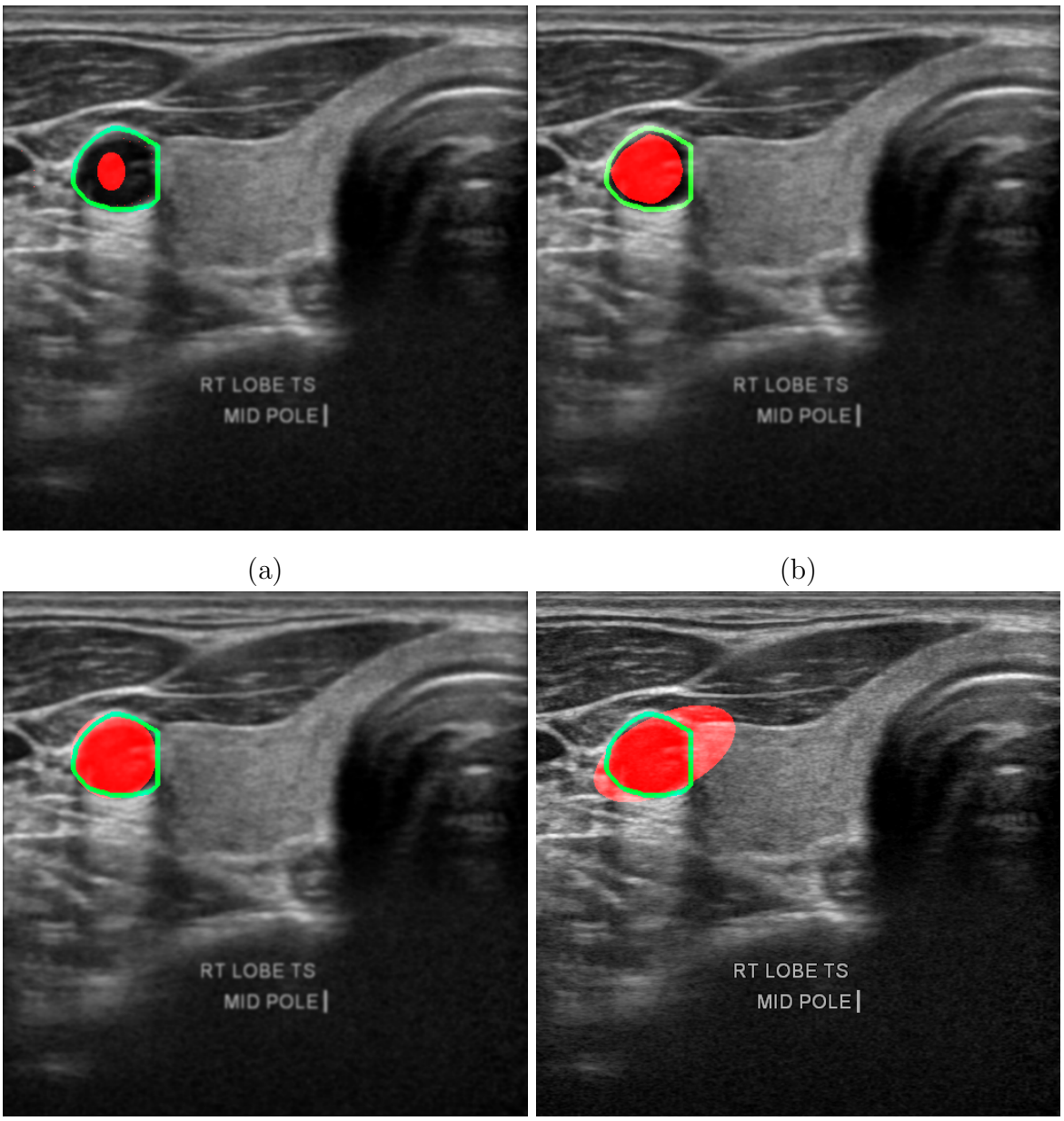


Figure 2.10: Results of carotid segmentation by existing methods where the pixels in red belong to the segmented carotid and pixels in green belong to the boundary marked by a sonographer. (a) Star algorithm. (b) Star Kalman algorithm. (c) Star Extended Kalman algorithm. and (d) Spokes ellipse algorithm.

spectively be used to perform tissue characterization and detect anatomical structures of interest in an ultrasound image.

- Local phase can be employed to differentiate between the hyperechoic and the hypoechoic regions in an ultrasound image.
- Existing methods to segment the thyroid gland in 2D ultrasound images are not generic in terms of its applicability to images acquired from different ultrasound imaging devices. Thus there is a definite need for unsupervised segmentation algorithms for the ultrasound images of the thyroid gland that perform well on multi-vendor datasets.
- Methods proposed to segment the carotid artery are semi-supervised algorithms. The shape of the carotid artery is approximated to be an ellipse whose parameters are estimated by the application of Kalman filters or 1D Hough transform. Except for the Star extended Kalman algorithm, the rest under-segment or over-segment the carotid artery.

Chapter 3

Datasets, Pre-Processing and Validation Protocol

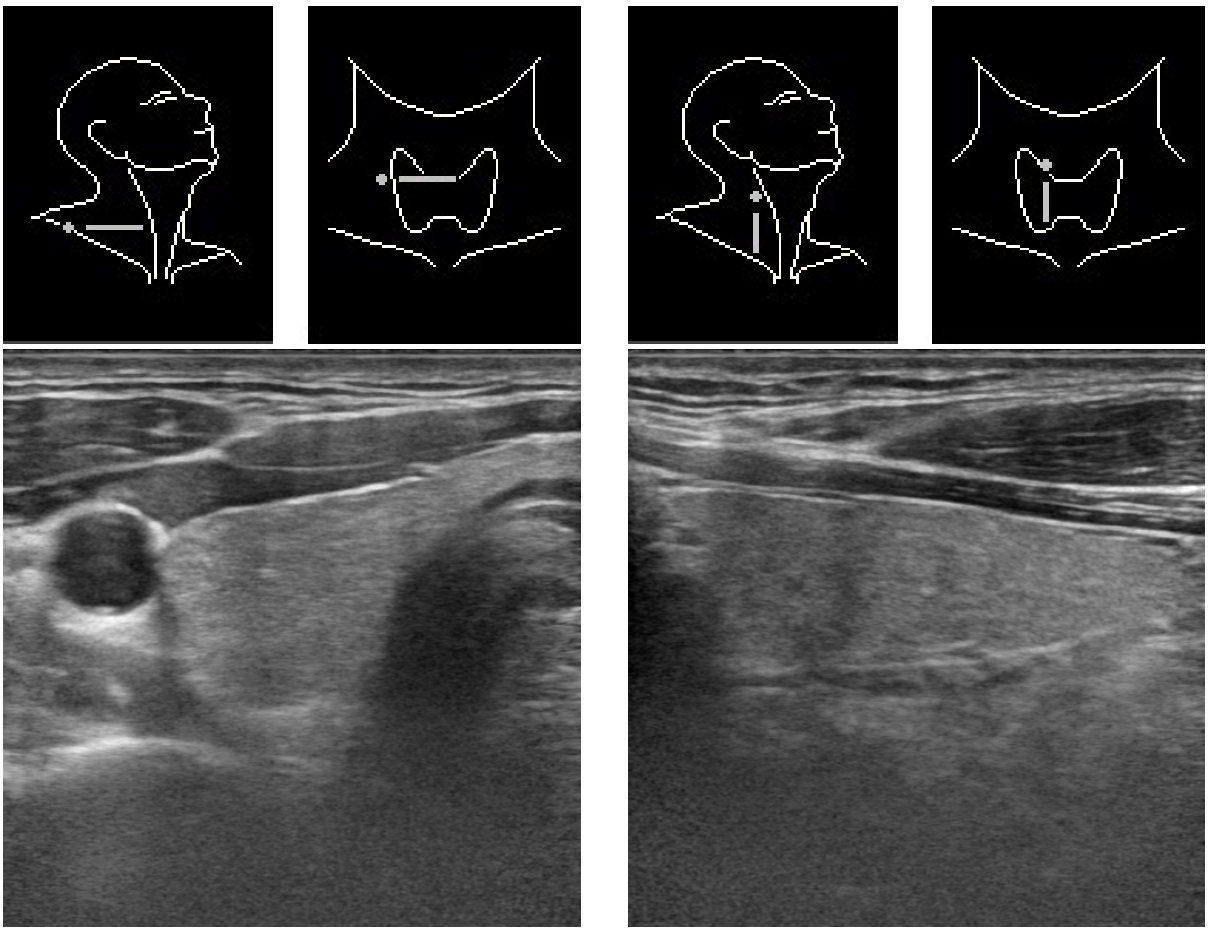
3.1 Datasets

In clinical practice, the thyroid gland is scanned in two orthogonal planes. The images in the two planes are called the transverse and longitudinal scans. In transverse scans, the ultrasound probe is held horizontal and the gland is scanned from top to bottom. In the longitudinal scan, the ultrasound probe is held vertical and the gland is scanned side ways. Figure 3.1 shows the transverse and longitudinal ultrasound scans of the thyroid gland along with the schematics that depict the probe positions.

To perform experiments, anonymized US images in B-mode were retrospectively collected from two hospitals:

- *Set 1 (S1)*: Images in this set were acquired at the Tan Tock Seng Hospital, Singapore, using Hitachi HI Vision Avius Ultrasound equipped with an L75 5-18MHz probe. All images have a resolution $M \times N$ equal to 768×1024.
- *Set 2 (S2)*: Images in this set were acquired at Mt. Elizabeth Hospital, Singapore, using GE Logiq Ultrasound equipped with L6-12-RC probe. The images have a resolution $M \times N$ equal to 614×820 .
- *Supervised Set (SS)*: This dataset consists of 24 images (15 transverse and 9 longitudinal) to train and 16 images (11 transverse and 5 longitudinal) to test supervised learning-based segmentation algorithms. Of the 15 training images in transverse scan 12 are from Set 1 and 3 are from Set 2. All images (train and test) in the longitudinal scan are from Set 1.
- *Public Dataset (PD)*: This is a publicly available dataset[1] that is used to test the performance of the proposed methods to detect, segment and track the common carotid artery. The dataset consists of images acquired using Ultrasnix and Toshiba ultrasound machines [138].

[1] `http://splab.cz/en/research/zpracovani-medicinskych-signalu/databaze/artery`



(a) (b)

Figure 3.1: Transverse (a) and longitudinal (b) ultrasound scan of the thyroid gland.

- *Pre-processing Dataset (PPD)*: This dataset is used to test the annotation removal algorithm that shall be described in the sections to follow. Images in this dataset are of thyroid gland with abnormalities. The position and type of abnormality is annotated in the image and therefore act as ideal candidates to test the annotation removal algorithm. The images were acquired at the Tan Tock Seng Hospital, Singapore, using Hitachi HI Vision Avius Ultrasound equipped with an L75 5-18MHz probe.

The details of the datasets are provided in Table 3.1. The datasets consist of a mix of images that are of the right or left lobes of the thyroid gland. The images were acquired at the mid or lower pole with the patient in supine position. Images that were affected by abnormalities leading to organ occlusion are excluded from the studies involving multi-organ segmentation. All images were acquired at a centre frequency of 10 MHz. Since the datasets were retrospectively collected, the Time Gain Compensation (TGC) setting used during imaging is unknown.

Dataset	Patients\ Volunteers	# Images	TS	LS	GT	FOV Resolution	Pixel size (mm)
S1	6\6	34	20	14	34	511×510	0.0741
S2	10\-	18	18	-	6	450×432	0.0893
SS	-\-	40	26	14	40	S1 & S2	S1 & S2
PD	17\-	971	971	-	971	-	-
PPD	28\-	28	28	-	NA	511×510	0.0741

S1 = Set 1. S2 = Set 2. SS = Supervised Set. PD = Public Database. PPD = Pre-Processing Database. TS = Transverse Scan. LS = Longitudinal Scan. GT = Ground Truth. FOV = Field of View. mm = Millimetres. NA = Not Applicable.

Table 3.1: Details of datasets used in experiments

Figure 3.2 shows one image from each database that is listed above. The region in the middle of the images in Figure 3.2 (a)-(c), where the tissue/organ is visible, is termed as the Field-of-View (FOV) of an ultrasound image. The region around the FOV contains the image acquisition settings and details regarding any observation that is made within the FOV of the image. All processing happens within the FOV of the image. The method to extract the FOV for ultrasound image analysis is described in Section 3.3.1.1.

3.1.1 Ground Truth (GT)

Two expert Otorhinolaryngologists who are also trained sonographers at Tan Tock Seng Hospital, Singapore, were asked to manually delineate the organs (Muscles, trachea, thyroid and carotid) using the ITK-SNAP software [139]. The polygon tool provided in the software is used to segment the organs of interest (thyroid, carotid artery, muscles and the trachea). Figure 3.3 shows the process of manually segmenting the organs using the ITK-Snap software.

3.2 Annotations in ultrasound images

Often, the thyroid gland is scanned by a trained sonographer who then sends the images to a doctor (Otorhinolaryngologist \ Ear, Nose and Throat (ENT) specialist) to make a diagnosis. The sonographers annotate the images to provide information regarding the scan to the doctors who are physically not present during the procedure. Figure 3.4 shows a US image with annotations. The sonographer has marked "LT LOBE TS" and "CYST" to indicate that the image is that of the left lobe of the thyroid gland in the transverse scan which has a cyst of width 11.0 mm and depth of 6.8 mm. The extremities of the cyst are highlighted with the help of markers that resembles a "+" symbol.

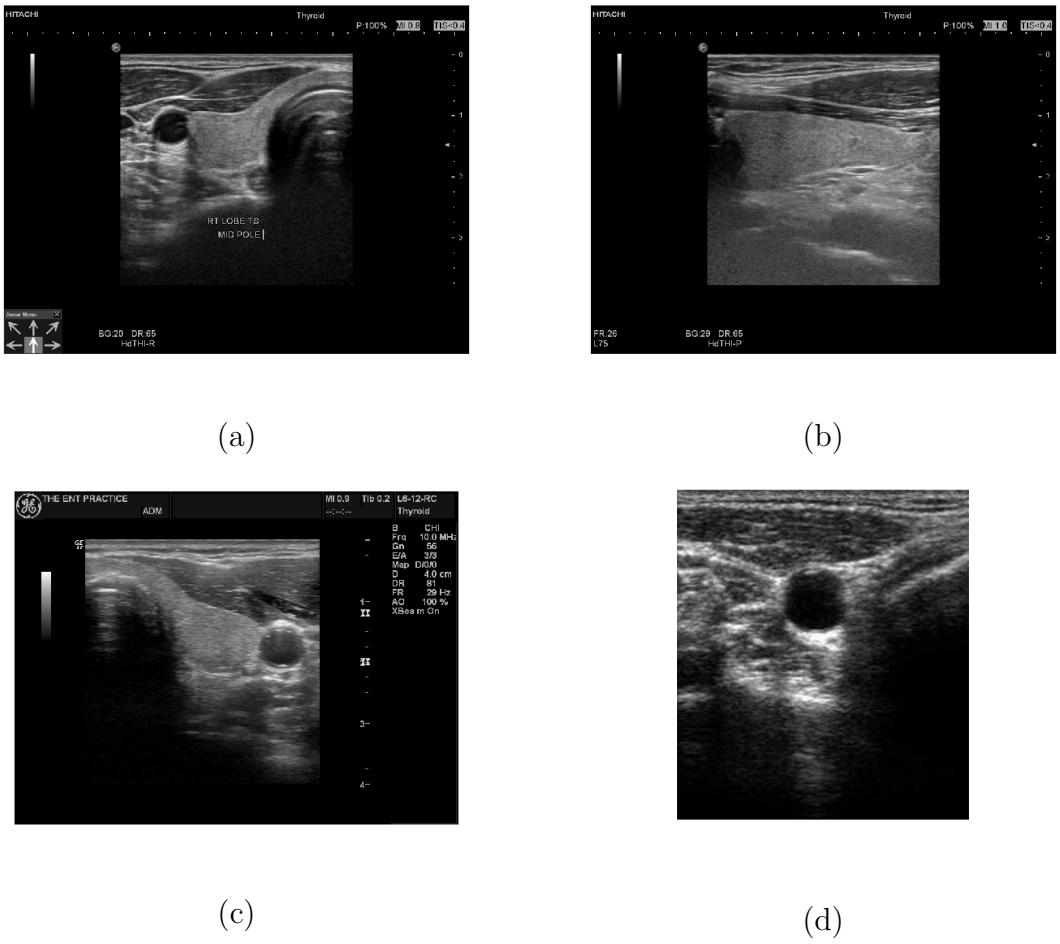
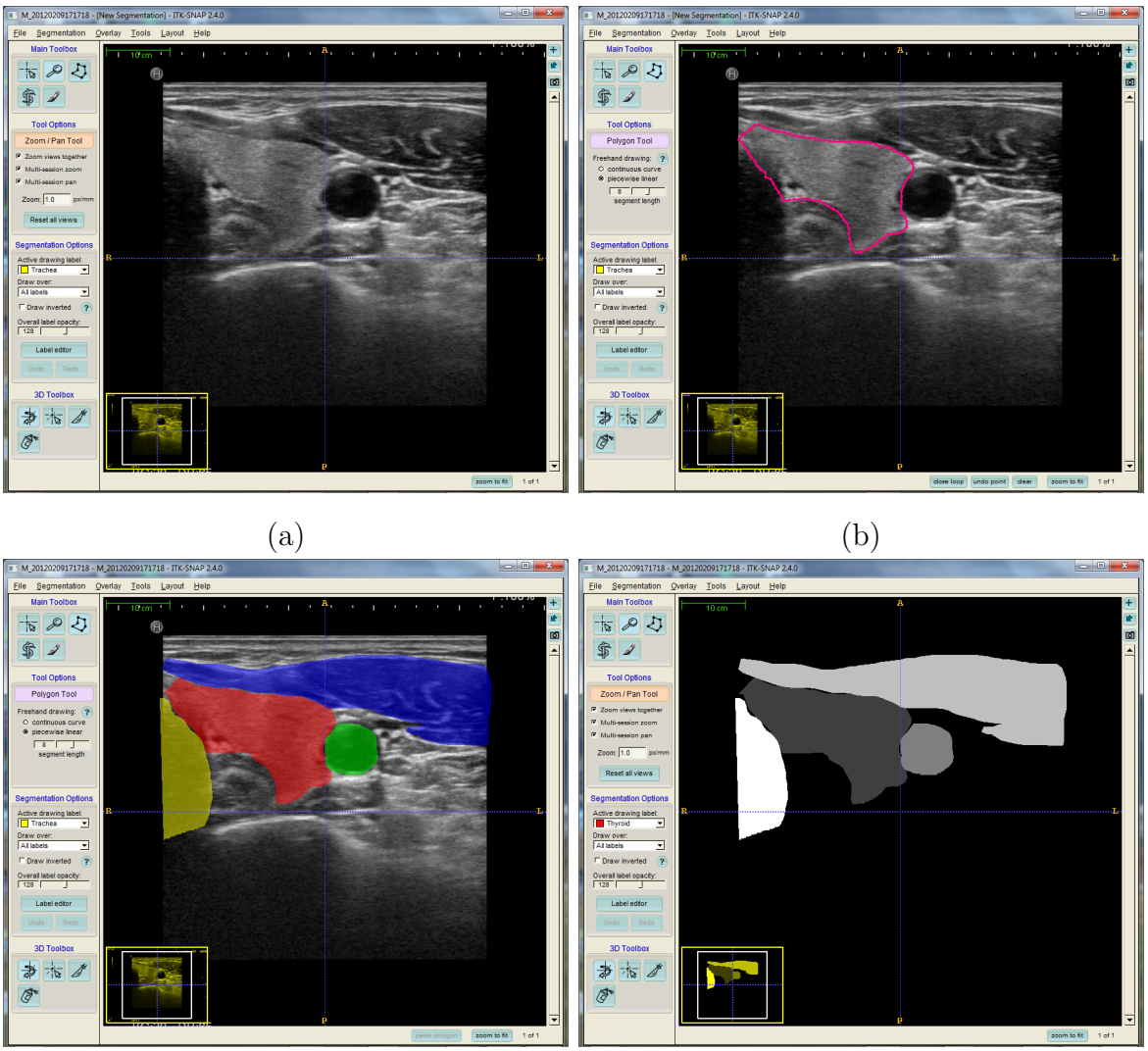


Figure 3.2: Sample images from the datasets used in the research work. (a) S1- Transverse;(b) S1- Longitudinal;(c) S2- Transverse; and (d) PD

3.3 Pre-processing for annotation removal

While annotations in US images help a doctor in making a diagnosis, it adversely affects the automatic methods that are designed to analyse the images. Looking at Fig. 3.4, it can be seen that the annotations obscure the boundaries of tissues where it is present. Application of gradient based boundary detectors [140, 141] to the images with annotation will result in its edges being detected as a part of the boundary that belongs to a tissue. As an example, the application of the Canny edge detector to Fig. 3.5(a) results in Fig. 3.5 (b) that shows edge pixels in soft tissue region where there should have been none if it weren't for the annotations. It can also be observed from Fig. 3.5(b) that the boundary of the cyst is deformed by the annotations. Such an effect is undesirable for computer aided diagnosis algorithms such as [142] that rely on the accurate representation of the boundary of the abnormality to assess the malignancy risk associated with it.



(a) (b)

(c) (d)

Figure 3.3: Example of using ITK-Snap to manually segment organs in ultrasound images. (a) ITK Snap GUI with input US image in the transverse scan;(b) Manually marked boundary of the thyroid gland (in pink);(c) Manually segmented organs (red - thyroid; green - carotid; blue - muscles; and yellow - trachea); and (d) Manual segmentation mask in gray scale.

The purpose of an ultrasound image analysis algorithm is for it to be used by a doctor as a second opinion to affirm the diagnosis s/he makes. It goes without saying that the images that are to be analysed will be annotated by a trained sonographer unless a doctor is physically present at the time of scanning. In this section, an annotation removal algorithm is proposed to detect and remove the annotations and restore the annotation removed image the image without annotations.

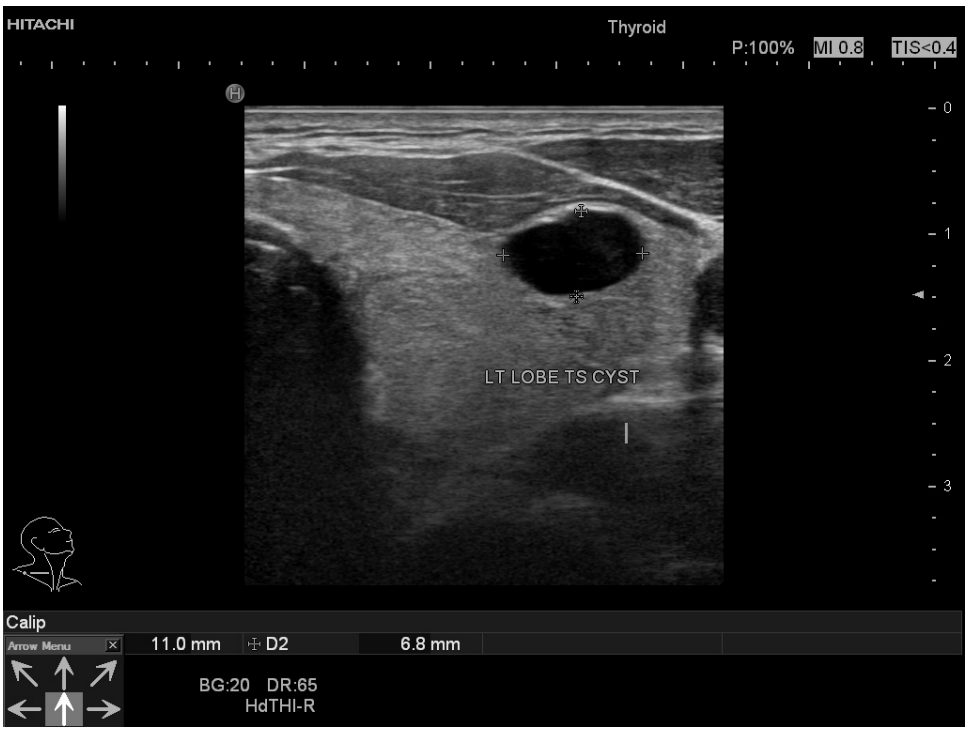
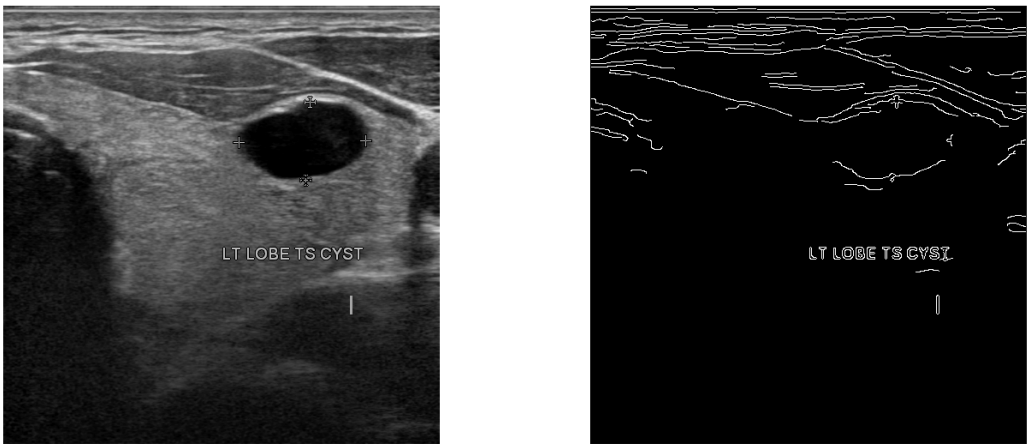


Figure 3.4: 2D ultrasound image of the thyroid gland in transverse scan with annotations.



(a) (b)

Figure 3.5: Results of applying Canny edge detector to the Field of View (FOV) in the image of Fig. 3.4. (a) The FOV. (b) Edge map.

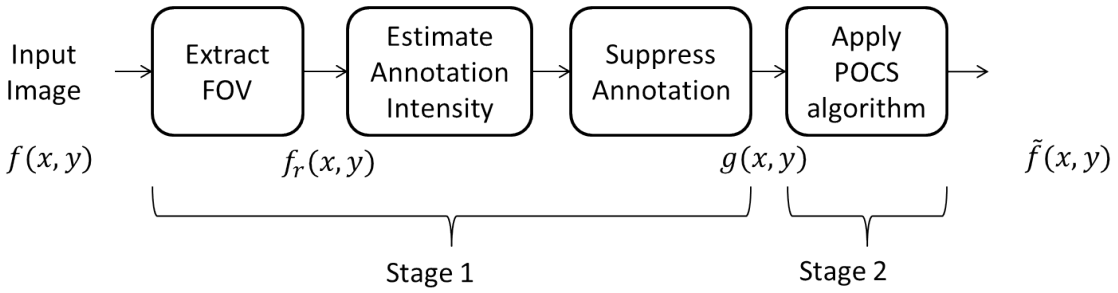


Figure 3.6: Flow diagram of the annotation removal algorithm

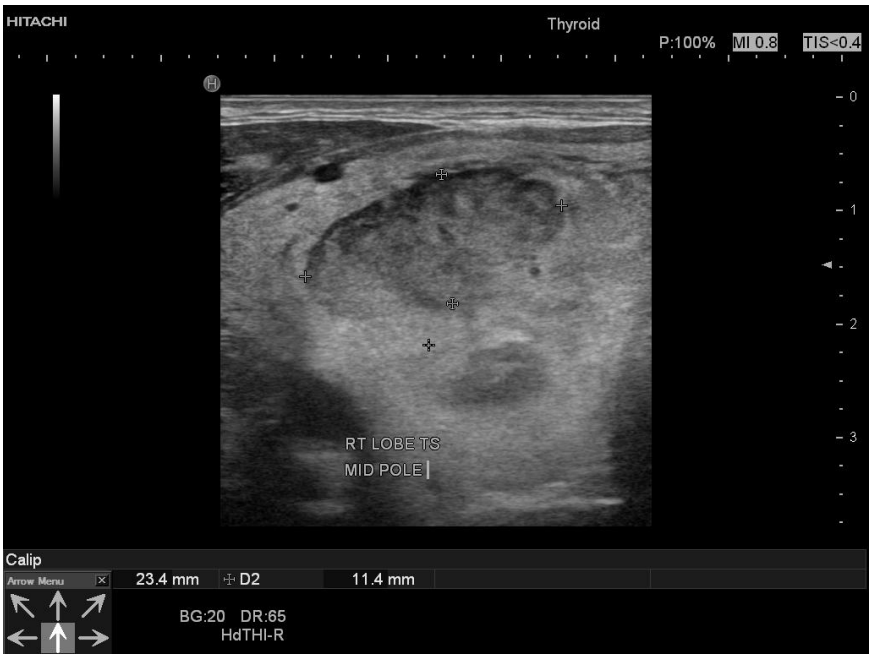


Figure 3.7: Input 2D Ultrasound Image $f(x, y)$ of the thyroid gland with annotations.

3.3.1 Annotation Removal

Figure 3.6 shows the flow diagram of the annotation removal algorithm. The first stage in the algorithm is to determine the intensities of the annotations in the FOV and to remove the artefacts.

3.3.1.1 Extracting the FOV

The FOV is the is the largest component in the input US image. Denoting the input US image by $f(x, y)$, a binary image $b_0(x, y)$ is obtained by Eq.(3.1):

$$b_0(x,y) = \begin{cases} 1 & \text{if } f(x,y) > 0 \\ 0 & \text{otherwise} \end{cases} . \tag{3.1}$$

Applying 2D connected components algorithm on the binary image $b_0(x, y)$ results in a labelled image $s(x, y)$ with K components, given by

$$s(x, y) = l; \quad l \in [1, K]. \tag{3.2}$$

Letting $l_{\max}$ to be the label of the component in $s(x, y)$ with the maximum number of pixels in it, we define a binary mask $r(x, y)$ such that

$$r(x, y) = \begin{cases} 1 & \text{if } s(x, y) = l_{\max} \\ 0 & \text{otherwise} \end{cases}. \tag{3.3}$$

The image containing the FOV, $f_r(x, y)$ and its complement $\bar{f}_r(x, y)$ are given by

$$f_r(x, y) = f(x, y) \times r(x, y) \tag{3.4}$$

$$\bar{f}_r(x, y) = f(x, y) - f_r(x, y). \tag{3.5}$$

3.3.1.2 Estimating annotation intensity

Annotations in the FOV have the same intensity as the pixels in the regions surrounding the FOV. To determine the intensities of the annotation, the histogram of $\bar{f}_r(x, y)$ is obtained. The non-zero bins of the histogram represent the intensities of the annotation. However, it should be noted that pixels in FOV that do not belong to the annotation may also have intensities overlapping with it. In order to minimize the false categorization of non-annotation pixels as belonging to annotations, the gray level range is empirically divided into three sub-ranges. Assuming the gray level range of the input image to be [0,255], the sub-ranges are empirically defined as [0,100], [101,200] and [200-255]. The peaks of histogram in each of the three sub-ranges are determined to be the intensities of the annotation.

3.3.1.3 Suppressing annotations

Let the intensity peaks in each of the three sub-ranges determined in Section 3.3.1.2 be represented by the set $\mathcal{T}$ given by

$$\mathcal{T} = \{T_1, T_2, T_3\}. \tag{3.6}$$

The binary annotation mask $b_a(x, y)$ is defined as

$$b_a(x, y) = \begin{cases} 1 & \text{if } f_r(x, y) \in \mathcal{T} \\ 0 & \text{otherwise} \end{cases}. \tag{3.7}$$

The binary annotation mask is further subjected to dilation with a 3×3 structuring element. Dilation is done so as to avoid discontinuities in the restored image.

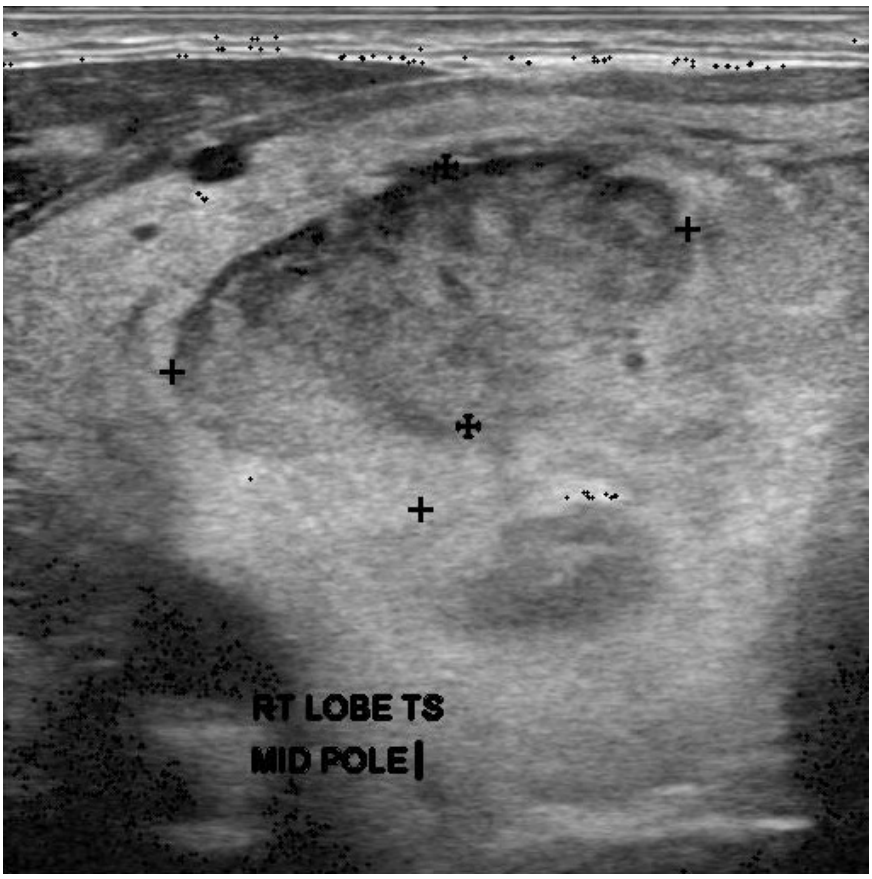


Figure 3.8: The Field of View (FOV) with annotations removed.

The image, $g(x, y)$ with the artefacts removed is obtained by multiplying the FOV $f_r(x, y)$ with $\bar{b}_a(x, y)$, the complement of the binary artefact mask as given by Eq.(3.8):

$$g(x, y) = f_r(x, y) \cdot \bar{b}_a(x, y). \tag{3.8}$$

Figure 3.8 shows the FOV in Fig. 3.7, with the annotations removed using the method described above. It can be seen that the annotations are completely removed from the FOV. It can also be seen that a few pixels associated with soft tissue are also removed in the process, but these are restored to a fair degree of accuracy by the restoration scheme described in Section 3.3.2.

3.3.2 Image Restoration

Perhaps the most studied topic in image processing after enhancement of images is that of image restoration. One way to restore the image is by the method of Convex Projections or more commonly known as Projection onto Convex Sets (POCS) method [143]. Hirani and Totsuka [144] apply the POCS algorithm to interactively remove noise

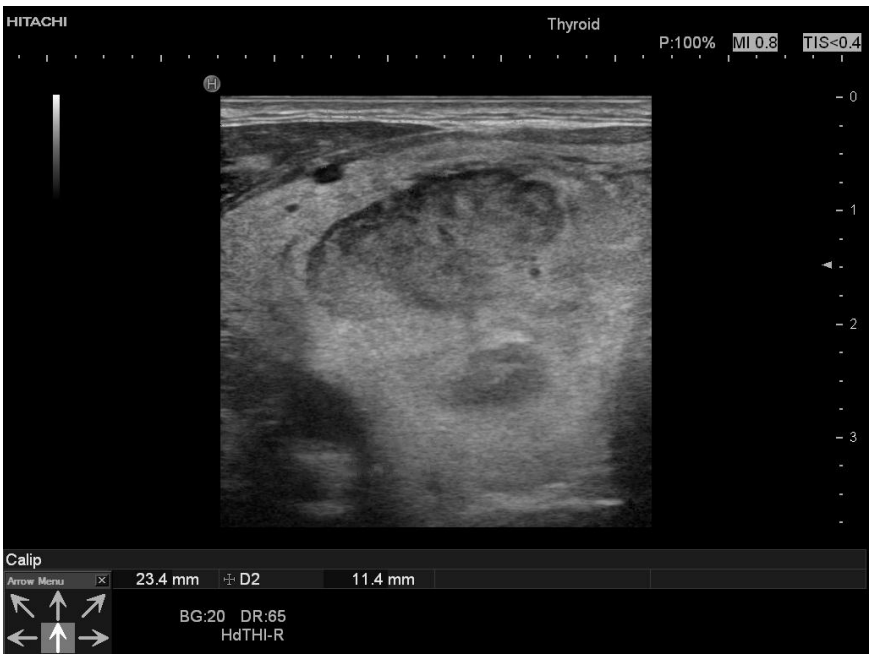


Figure 3.9: Restored image $\tilde{f}$(x,y).

and restore the image. Ogawa and Haseyama [145] have used it to restore texture in images with text matter embedded in it. Zhu et al. [146] treat the restoration of the tampered images with watermarks as an irregular sampling problem and use POCS to restore the tampered parts of the image. Santago and Link [147] make use of POCS to restore Cardiac Magnetic Resonance Images that are affected by motion artefacts. In order to restore the annotation removed pixels in our algorithm, an approach similar to the one observed in [146] is used with a slight modification. A modified POCS based approach using Block Discrete Cosine Transforms (BDCT) is used to restore the image.

In our method, the input image $g(x, y)$ is divided into blocks of size $B \times B$ pixels with a block overlap of $B/2$ pixels, $B > 0$. Block Discrete Cosine Transform (BDCT) is applied to each block to obtain $G(u, v)$. The first projection operator P_1 is given by:

$$P_1 f = \begin{cases} G(u,v) & \text{if } u - v < a; 1 < a < B \\ 0 & \text{otherwise} \end{cases} \tag{3.9}$$

where $G(u, v)$ is the BDCT of $g(x, y)$. Application of this projection operator, which is essentially a low pass filter, band-limits the image block.

Let $\triangle$ denote the set of all pixels that were removed during the annotation removal process. The second projecton operator P_2 is given by:

$$P_2 f = \begin{cases} f(x,y) & (x,y) \in \triangle \\ g(x,y) & (x,y) \notin \triangle \end{cases} . \tag{3.10}$$

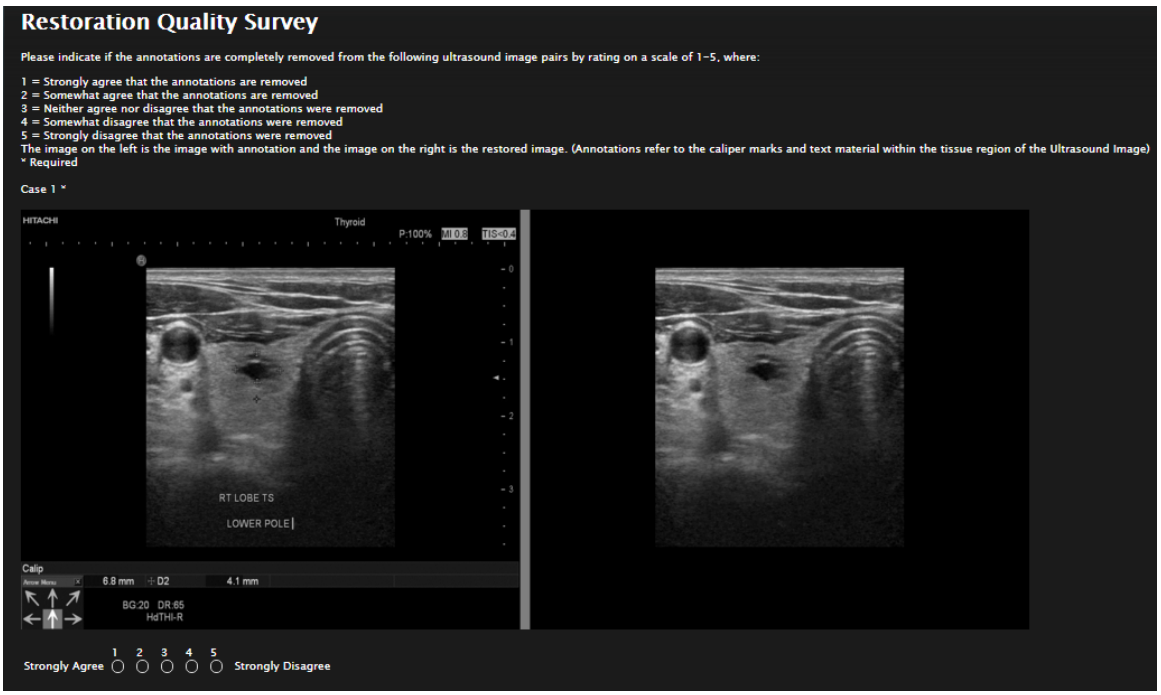


Figure 3.10: Screenshot of the online survey conducted to perform subjective analysis of restoration quality.

The image is iteratively restored using the relation:

$$f^{(n+1)} = P_2 P_1 f^{(n)} \tag{3.11}$$

with

$$f^{(1)} = P_2 P_1 g. \tag{3.12}$$

Figure 3.9 shows the restored image $\tilde{f}(x, y)$.

3.4 Validation Protocol

The metrics used to: (a) validate the annotation removal algorithm and the MODS algorithms of Chapter 4; and (b) assess the performance of existing segmentation algorithms are discussed in this section. The following notations are used to describe the metrics:

- Input image: $f(\mathbf{x})$
- Image resolution: $M \times N$
- Pixel co-ordinates: $\mathbf{x} = (x, y)$
- Pixel indices: $i, j \in [1, MN]$

- Pixel set: $X = \mathbf{x}_i$
- Restored image: $\tilde{f}(\mathbf{x})$
- Number of experts: K
- Automatically segmented image: $f_a(\mathbf{x})$
- Image segmented by the k^{th} expert: $f_m^k(\mathbf{x})$
- Segmentation label at pixel i in $f_a(\mathbf{x})$: $l_i^a \in [0, L]$, $L \to \mathbb{Z}$
- Segmentation label at pixel i in $f_m^k(\mathbf{x})$: $l_i^{mk} \in [0, L]$, $L \to \mathbb{Z}$
- Segmentation label at pixel i in ψ^{th} image of the dataset containing manual segmentations: $l_i^{mk\psi} \in [0, L]$, $L \to \mathbb{Z}$

3.4.1 Preprocessing algorithm

3.4.1.1 Qualitative analysis

A subjective analysis is performed to assess the quality of the restored image. An online survey[2] is prepared where the participants (clinicians and medical imaging researchers) are asked to give a rating on a scale of 1-5, indicating their opinion on the quality of restoration. A score of 1 suggests that the participant strongly agrees that all annotations are removed and a score of 5 suggests that a participant strongly disagrees that the annotations are removed. Figure 3.10 shows a screen shot of the online survey page.

3.4.1.2 Quantitative analysis

The Peak-Signal to Noise Ratio (PSNR) metric is used to quantify the performance of the annotation removal algorithm. For an 8-bit gray scale US image, the PSNR of the restored image is given by:

$$\text{PSNR} = 10 \cdot \log \left(\frac{255^2}{\frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [f(\mathbf{x}) - \tilde{f}(\mathbf{x})]^2} \right) \tag{3.13}$$

[2] https://sites.google.com/site/snikhilnarayan/survey-2

3.4.2 Segmentation algorithms

The segmentation algorithms are qualitatively and quantitatively evaluated for performance using numerous metrics that shall be discussed in the next Section. While the Qualitative metrics allow for the comparison of different algorithms, Quantitative metrics measure the accuracy of the segmentation results with respect to the GT segmentations.

3.4.2.1 Qualitative analysis

There are four metrics to assess the quality of the segmented image. These are: (a) the Probabilistic Rand Index (PRI) [148]; (b) the Global Consistency Error (GCE) [149]; (c) Variation of Information (VOI) [150]; and (d) Boundary Distance Error (BE) [151].

(i) *PRI*: Given K manual segmentations of image $f(\mathbf{x})$ the PRI for a segmentation algorithm is given by:

$$\begin{aligned}\mathrm{PRI}(f_a(\mathbf{x}), f_m^k(\mathbf{x})) = \frac{1}{\binom{MN}{2}} \sum_{\substack{i,j \\ i \neq j}} [\mathbb{I}(l_i^a = l_j^{mk}) p_{ij} \\ + \mathbb{I}(l_i^a \neq l_j^{mk})(1 - p_{ij})]\end{aligned} \tag{3.14}$$

where, $\mathbb{I}$ is an indicator function and

$$p_{ij} = \frac{1}{K} \sum_{k=1}^{K} [\mathbb{I}(l_i^{mk} = l_j^{mk})]. \tag{3.15}$$

with $f_m^k(\mathbf{x})$ and $f_a(\mathbf{x})$, $k \in [1, K]$, representing the manually and automatically segmented images, respectively. The results of a segmentation algorithm are considered to be of good quality if the PRI $\in [0, 1]$ has a value that is close to one.

An extension to the PRI is the Normalized Probabilistic Rand (NPR) Index [152, 153]. The NPR Index is given by:

$$\mathrm{NPR} = \frac{\mathrm{PRI} - \text{expected index}}{1 - \text{expected index}}. \tag{3.16}$$

If there are $\boldsymbol{\psi}$ images in the dataset, then the expected index $\mathbb{E}$ is given by:

$$\mathbb{E} = \frac{1}{\binom{MN}{2}} \sum_{\substack{i,j \\ i \neq j}} \left[p'_{ij} p_{ij} + (1 - p'_{ij})(1 - p_{ij}) \right] \tag{3.17}$$

where,

$$p'_{ij} = \frac{1}{\boldsymbol{\psi}} \sum_{\psi} \frac{1}{K} \sum_{k} \mathbb{I}(l_i^{mk\psi} = l_j^{mk\psi}) \tag{3.18}$$

With a little bit of mathematical jugglery, the expected index $\mathbb{E}$ can be written as:

$$\mathbb{E} = \frac{1}{\boldsymbol{\psi}} \sum_{\psi} \frac{1}{K} \sum_{k} \text{PRI}(f_a^{\psi}(\mathbf{x}), f_m^k(\mathbf{x})) \tag{3.19}$$

The NPR index is used here to justify the use of constants in one of the MODS method of Chapter 4.

(ii) *GCE*: This metric is defined based on the assumption that the results of the automatic segmentation are a refinement of the GT manual segmentation. Using the notations described at the beginning of the Section, the refinement error *RE* is given by:

$$RE(f_a(\mathbf{x}), f_m^k(\mathbf{x}), \mathbf{x}_i) = \frac{|l_i^a \backslash l_i^{mk}|}{l_i^a} \tag{3.20}$$

where, $|.|$ denotes the cardinality of the set and $\backslash$ denoted set difference. The GCE metric is then defined as:

$$GCE = \frac{1}{MN} \min\{\sum_i RE(f_a(\mathbf{x}), f_m^k(\mathbf{x}), \mathbf{x}_i), \sum_i RE(f_m^k(\mathbf{x}), f_a(\mathbf{x}), \mathbf{x}_i)\} \tag{3.21}$$

A good segmentation algorithm has GCE $\in [0, 1]$ values close to 0.

(iii) *VOI*: The VOI metric measures the quality of segmentation in terms of the information difference between the segmentation results of the automatic algorithm and the manual segmentations. The VOI metric is defined as:

$$VOI = H(f_a(\mathbf{x})) + H(f_m^k(\mathbf{x})) - 2I(f_a(\mathbf{x}), f_m^k(\mathbf{x})) \tag{3.22}$$

where, for a generic segmented image $f_s(\mathbf{x})$ of size $M \times N$ that has n_l pixels with the label l, the entropy $H(f_s(\mathbf{x}))$ is given by:

$$H(f_s(\mathbf{x})) = -\sum_{l=1}^{L} \frac{n_l}{MN} \log \frac{n_l}{MN} \tag{3.23}$$

and the mutual information $I(f_a(\mathbf{x}), f_m^k(\mathbf{x}))$ is given by:

$$I(f_a(\mathbf{x}), f_m^k(\mathbf{x})) = \sum_{l'=1}^{L} \sum_{l=1}^{L} \frac{n_{l,l'}}{MN} \log \frac{n_{l,l'}}{MN} \frac{n_l}{MN} \frac{n'_l}{MN}. \tag{3.24}$$

The results of a segmentation algorithm are of good quality if VOI $\in [0, \infty]$ is close to zero.

(iv) *BE*: Boundary error is the maximum of the pairwise Euclidean distance between the boundary pixels in $f_a(\mathbf{x})$ and $f_m^k(\mathbf{x})$. It should be noted that the boundary error here is measured in terms of the number of pixels. In order to obtain the error in millimetres, this should be multiplied by the pixel size in mm. A good segmentation algorithm produces the lowest boundary errors i.e., BE close to 0.

3.4.2.2 Quantitative analysis

The segmentation algorithms are quantitatively evaluated by using metrics that measure the amount of overlap between the GT and the segmentation results. The overlap measures used for validation are: (1) Sensitivity (SE); (2) Specificity (SP); (3) Dice Coefficient (DSC) and (4) True Positive Rate (TPR) or Positive Predictive Value (PPV) given by:

$$\mathrm{SE} = \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FP}}; \mathrm{SP} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}}; \tag{3.25}$$

$$\mathrm{PPV} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}}; \quad \mathrm{DSC} = \frac{2 \times \mathrm{TP}}{2 \times \mathrm{TP} + \mathrm{FP} + \mathrm{FN}} \tag{3.26}$$

where TP, TN, FP and FN stand for True Positives, True Negatives, False Positives and False Negatives, respectively.

3.5 Experimental results

The algorithms were implemented using Matlab 2012 and executed on an Intel(R) Core(TM) i5-2400, 3.10GHz processor equipped with 4GB RAM.

3.5.1 Annotation removal algorithm

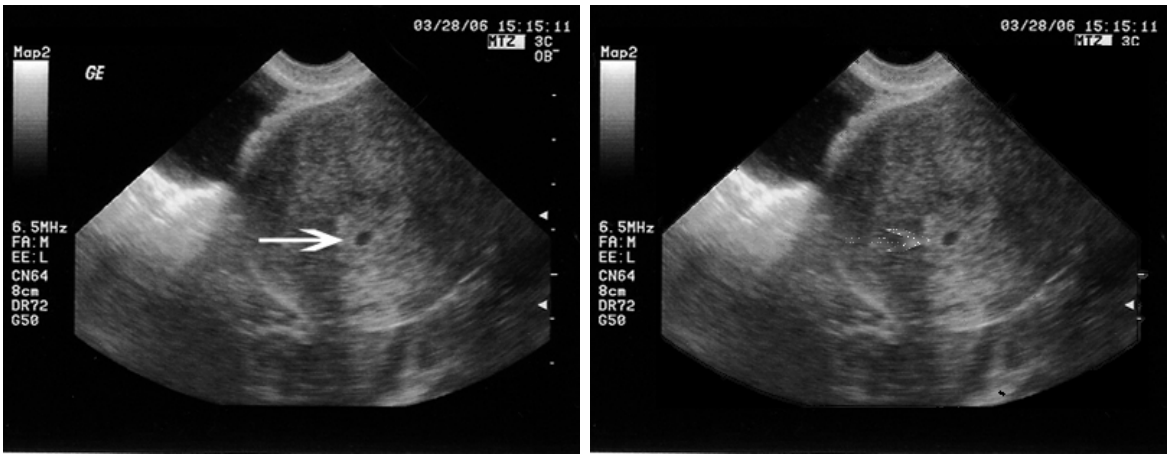
The PPD dataset is used to validate the performance of the annotation removal algorithm. This dataset consists of 18 images with annotation and 10 images without annotation. Each of the 18 images has at least one nodule with caliper marks and text embedded in its FOV by the examining sonographer. The BDCT is applied on image blocks of size 64×64 pixels with an overlap of 32 pixels and $B = 16$.

Images in the dataset that do not have any annotation embedded in its FOV are used to quantify the restoration results. Annotations in the form of text are introduced at random locations within the FOV of the image which are then input to the annotation removal algorithm. The PSNR is calculated for the restored image to measure the quality of restoration. The simulation is repeated for 50 times for every image in the dataset. Table 3.2 shows the PSNR values averaged over 50 iterations for all the images in the dataset. It can be seen that the images have PSNR values greater than 38dB suggesting the high quality of restoration.

Case Number	**1**	**2**	**3**	**4**	**5**
PSNR	39.78	40.48	38.94	39.62	39.67
Case Number	**6**	**7**	**8**	**9**	**10**
PSNR	39.02	40.33	40.11	41.06	40.91

* PSNR = Peak Signal to Noise Ratio

Table 3.2: Restoration Quality Analysis. The proposed pre-processing algorithm produces high quality restoration of the annotation removed image as evidenced by high PSNR values ($> 39dB$)



(a) (b)

Figure 3.11: Restoration of a Ultrasound image randomly downloaded from the internet. (a) Input. (b) Output.

The algorithm is also executed on a random Ultrasound image downloaded from the internet to check the generality of the algorithm. Figure 3.11(a) shows an Ultrasound image of a foetus with an arrow mark, given as input to the algorithm. Figure 3.11(b) shows the output after annotation removal and restoration. Figure 3.12 shows the results of the online survey conducted to assess the quality of restoration subjectively. It can be seen than more than 70% of the participant pool, which included clinicians and medical imaging researchers, agreed that the annotations had indeed been fully removed from the images and that the restored image resembled the original image without any distortion.

The algorithm, in general, is able to suppress the annotations completely at regions where the texture is homogenous and restore the image without any distortion. It is observed that the regions in the restored image that originally have annotations at gray level discontinuities such as edges of cystic nodule walls etc., tend to appear blurry. This is due to the fact that the projection operator used in the POCS algorithm acts as a low pass filter. It can be inferred from the experiments that the algorithm produces high

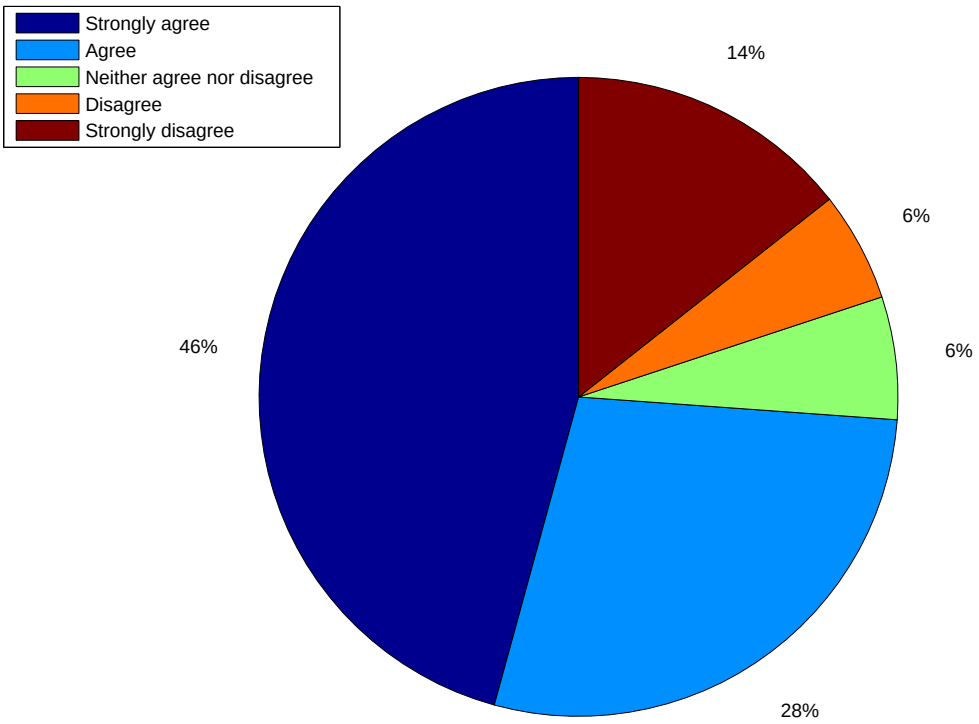


Figure 3.12: Survey results. Roughly 75% of the participant pool agreed that the annotation had been completely removed and that the restored image did not have any distortion in it.

quality restorations of the annotation removed images. Hence, the annotation removal algorithm can be used as a pre-processing tool to perform ultrasound image analysis in images with annotations.

3.5.2 Segmentation algorithms

In this section we analyse the performance of the methods introduced in Chapter 2 to segment the thyroid gland and the carotid artery. The metrics described in Section 3.4 are used to validate the algorithms.

3.5.2.1 Segmenting the carotid artery

The semi-automatic methods of [135], [134], [136] and [137] are implemented for this study. The centroid of the GT carotid segmentation (pixels of carotid that lie in the

Method	Initial Parameters
Spoke Ellipse	# Radial Lines: 30; Ellipse Fit: Least Squares (trace constraint)
Star	# Radial Lines: 30; $\text{Edge}_{\text{thr}} = 50$ Ellipse Fit: Hough 1D Transform
Star Kalman	# Radial Lines: 50; $\text{r}_{\max} = 75$ $\mathbf{Q} = 3; R = 20; M = 10$ Initial State Vector = $\text{r}_{\max}/2$
Star Extended Kalman	# Radial Lines: 50; $\text{r}_{\max} = 85$ $\mathbf{Q} = \text{diag}(2, 2, 0.1), R = 20, M = 5$ Initial State Vector = $[\text{r}_{\max}/2, \text{r}_{\max}, 0]^{\text{T}}$

* Descriptions for the notations used in each algorithm can be found in the respective papers.

Table 3.3: Parameters used to initialize the carotid segmentation methods.

intersection of the two manually segmented regions) is chosen as the seed point to initialize the algorithms. The recommended parameter settings needed to initiate the algorithms are listed in Table 3.3. It was noted in Section 2.5.1.2 that the carotid artery for the input ultrasound image in Figure 2.10 is under segmented or over segmented by all the methods. This is in fact true for all the images in our database which is evident from the DSC values in Table 3.4. The average overlap measured by the DSC metric for every method, except for the spokes ellipse algorithm on images in the dataset S2 and the star extended Kalman algorithm for images in the dataset S1, is less than 76%. High values of sensitivity and specificity can be attributed to the manual initializations of the algorithm.

From the plots in Figure 3.13, it can be observed that the median values of PRI, GCE, VOI and BE respectively stand at 0.97, 0.015, 0.1 and 7 for the Spokes Ellipse algorithm; 0.94, 0.0275, 0.27 and 22 for the Star algorithm; 0.97, 0.010, 0.1 and 9 for the Star Kalman algorithm; and 0.98, 0.009, 0.07 and 5 for the Star Extended Kalman algorithm. Combining this information with the DSC values of Table 3.4, it can be inferred that the existing methods often tend to under-segment the carotid artery. The Star Extended Kalman algorithm proposed by Gurrero et al. [136] performs better than the remaining algorithms both quantitatively and qualitatively.

It is also interesting to note that, except for the Star Kalman algorithm, the DSC values of set S1 differ from that of S2 by more than 7%. This is due to the fact that the boundary pixels of the carotid artery are determined by the use of intensity-based gradients. This suggests that the performance of the existing methods to segment the carotid artery is specific to the dataset used and therefore cannot be generalized to all datasets.

Method	Dataset	**SE**	**SP**	**DSC**	**PPV**
Spoke Ellipse	S1	0.982 ± 0.030	0.935 ± 0.090	0.742 ± 0.206	0.667 ± 0.244
[137]	S2	0.990 ± 0.008	0.962 ± 0.036	$\mathbf{0.839 \pm 0.073}$	0.755 ± 0.120
Star [134]	S1	0.935 ± 0.089	0.693 ± 0.338	0.404 ± 0.239	0.531 ± 0.361
	S2	0.875 ± 0.104	0.813 ± 0.350	0.303 ± 0.236	0.371 ± 0.334
Star Kalman	S1	0.999 ± 0.001	0.576 ± 0.212	0.700 ± 0.154	0.982 ± 0.044
[135]	S2	0.999 ± 0.001	0.566 ± 0.197	0.702 ± 0.166	0.998 ± 0.003
Star Extended	S1	0.987 ± 0.032	0.932 ± 0.099	$\mathbf{0.827 \pm 0.167}$	0.793 ± 0.177
Kalman [136]	S2	0.982 ± 0.018	0.930 ± 0.143	0.757 ± 0.125	0.692 ± 0.212

⋆ Results are averaged over both experts.
S1 = Set 1. S2 = Set 2. SE = Sensitivity. SP = Specificity. DSC = Dice Co-efficient. PPV = Positive Predictive Value.
Best results are highlighted in **Bold** font.

Table 3.4: Quantitative analysis of existing methods to segment the carotid artery. It is observed that while Spoke Ellipse algorithm has a good overlap with the GT for the images in Set 2, the Star Extended Kalman algorithm has good overlap with GT for images in Set 1.

3.5.2.2 Segmenting the thyroid gland

The thyroid segmentation algorithm of [49] is implemented by first determining the probable thyroid region followed by applying adaptive weighted median filtering and morphological closing of the image. Then, features such as: (a) Coefficient of Local Variation; (b) Block Difference of Inverse Probabilities; (c) Normalized Multi-Scale Intensity Difference;(d) Haar wavelet based features; and (e) histogram features are extracted in a 16×16 rectangular window. A radial basis function (RBF) neural network is trained using the features extracted from the images in the training set of dataset SS. The parameters used at every stage in our implementation are the same as those reported in [49]. Each test image in the dataset SS is divided into blocks of size 16×16 pixels, with 50% overlap between the blocks. The blocks are then classified into the thyroid or non-thyroid region by the trained classifier. The resulting image is post-processed by the recommended morphological operations. A similar approach is followed in implementing the method proposed by Garg et al. [133] with the following changes: (a) feature set is changed to histogram based features and multi-scale Haar wavelet based features; and (b) the classifier changed to feed-forward neural networks. It should be noted here that the regions belonging to the thyroid gland form the positive class and regions not belonging to the thyroid gland (carotid, muscles and the trachea) form the negative class. Overall 66 samples (33 positive and 33 negative) are generated to train the neural network classifier.

From the plots in Figures 3.14 and 3.15, it can be seen that both methods result in low quality segmentations of the thyroid gland. This is also confirmed by the low DSC

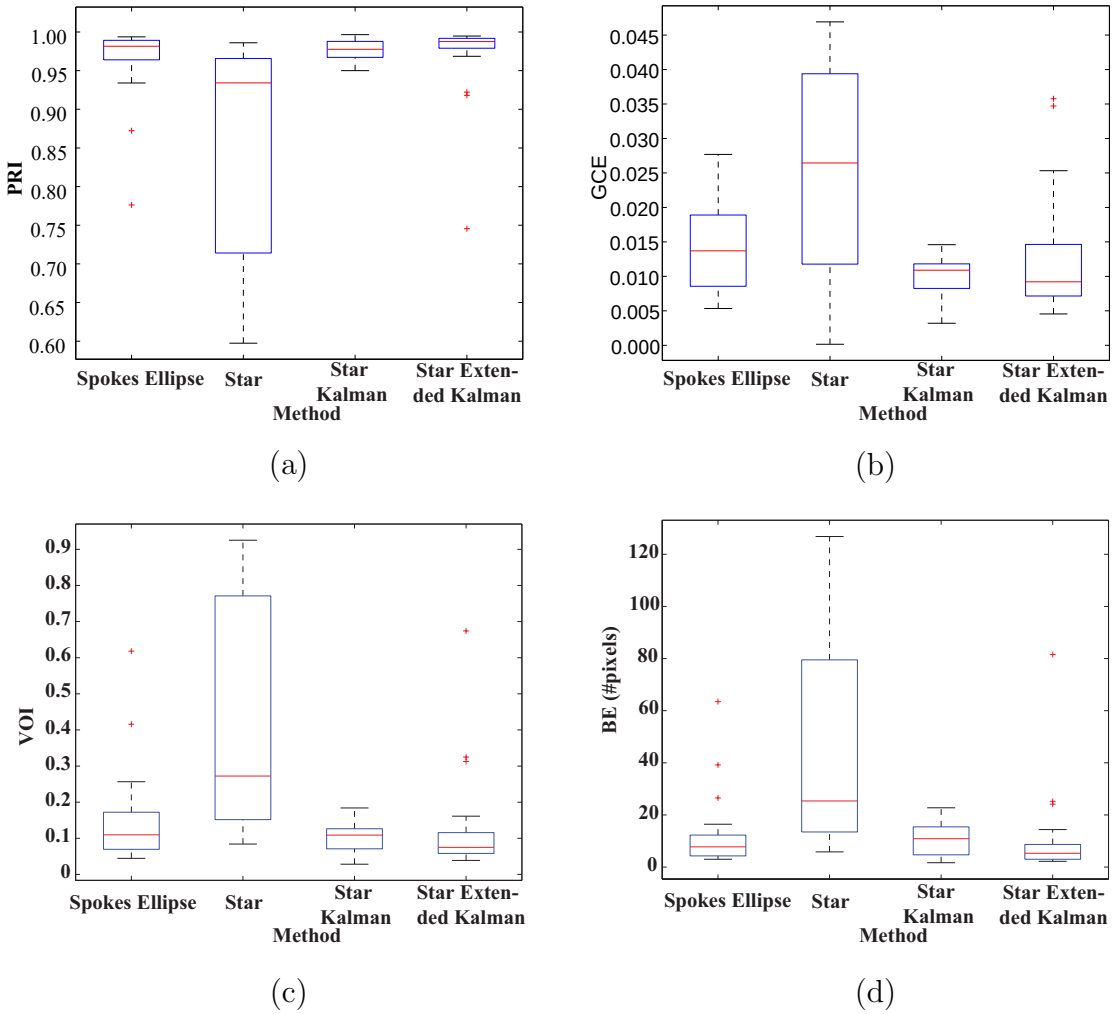


Figure 3.13: Qualitative analysis of existing methods to segment the carotid artery. (a) Probabilistic Rand Index (PRI). (b) Global Consistency Error (GCE). (c) Variation of Invormation (VOI). and (d) Boundary Error (BE) in pixels. It can be seen that for the Star Extended Kalman method, the PRI value is close to 1 with GCE, VOI and BE lower than the rest. Thus, the results of Star Extended Kalman method are better than that obtained by other algorithms.

values in Table 3.5.

3.5.3 Inter-observer variation

Overlap statistics of Section 3.4.2.2 are used to determine the inter-observer variation and the results are summarized in Table 3.6. The low SP values of E12 when compared to E21 for all the organs suggest the possibility of expert 1 under-segmenting the organs

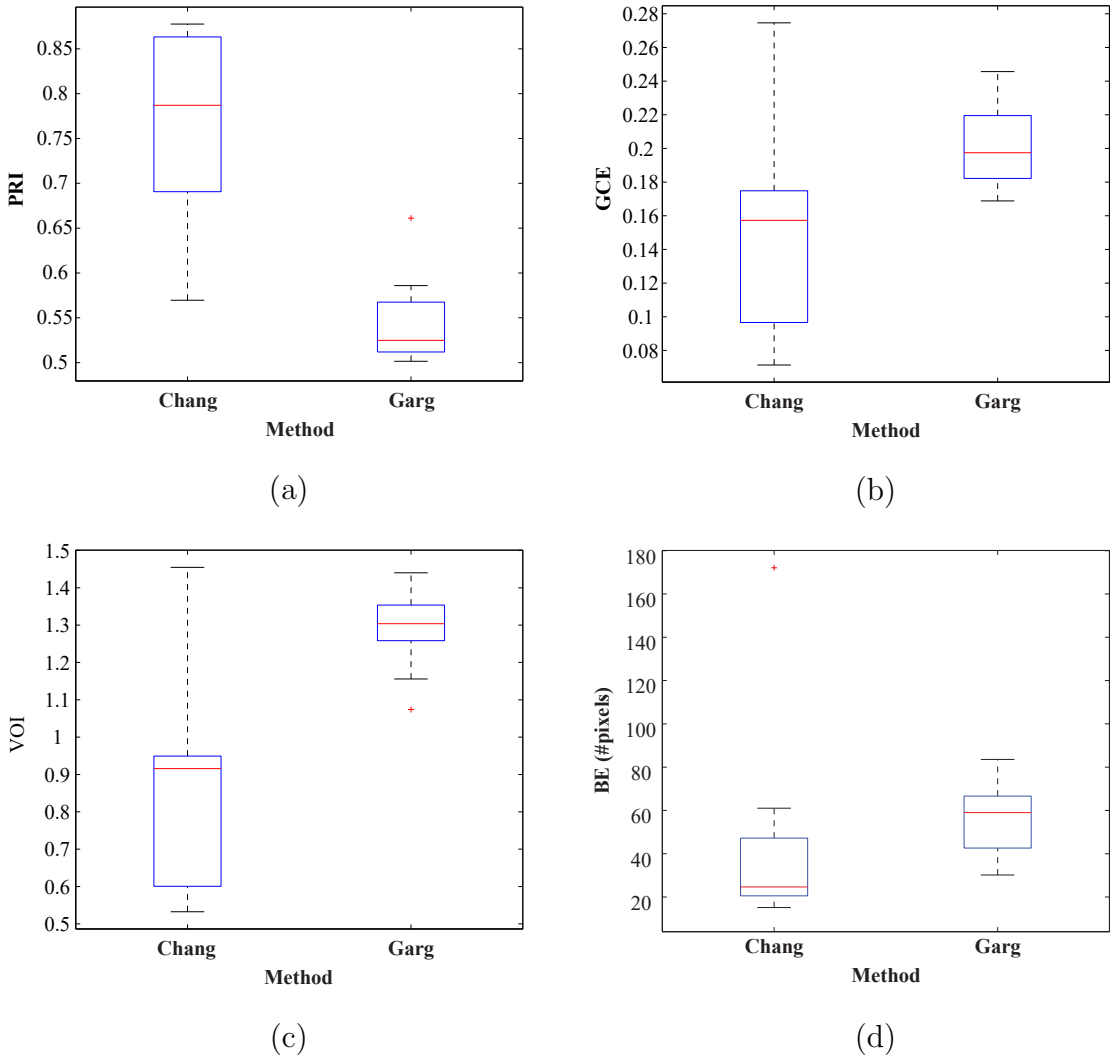


Figure 3.14: Qualitative analysis of existing methods to segment the thyroid gland in transverse scan. (a) Probabilistic Rand Index (PRI). (b) Global Consistency Error (GCE). (c) Variation of Invormation (VOI). and (d) Boundary Error (BE) in pixels. It can be seen that for the method by Chang, the PRI value is close to 1 with GCE, VOI and BE lower than that of the method by Garg. Thus, the results of Chang are better than that of Garg.

when compared to expert 2.

3.6 Summary

- The datasets used to conduct experiments were introduced in this chapter. Five different sets of images are used in our experiments. Images of dataset S1 and

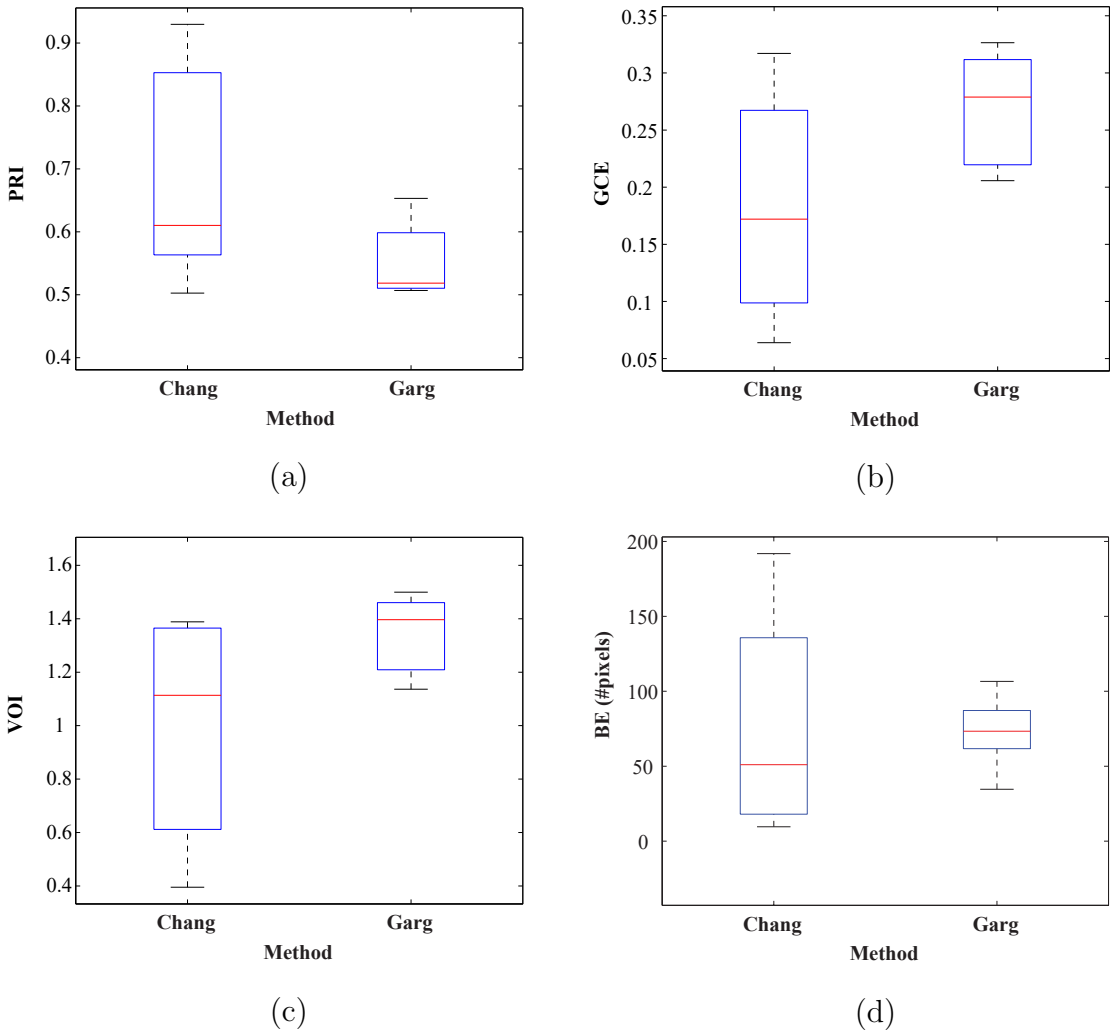


Figure 3.15: Qualitative analysis of existing methods to segment the thyroid gland in longitudinal scan. (a) Probabilistic Rand Index (PRI). (b) Global Consistency Error (GCE). (c) Variation of Invormation (VOI). and (d) Boundary Error (BE) in pixels. From the plots it can be inferred that the method of Chang et al., [49] produces better quality results than that by the method of Garg et al., [133] for longitudinal scans.

S2 were acquired using Hitachi and GE ultrasound imaging devices, respectively. Images in the public dataset were acquired using Toshiba and Ultrasonix imaging devices. The purpose of employing a multi-vendor dataset is to test the generality of the proposed methods.

- Annotations are embedded in ultrasound images by sonographers to provide more information about the findings in an ultrasound scan to the consulting physicians. Such annotations are a deterrent to the successful execution of automatic seg-

Method	Dataset	Scan	**SE**	**SP**	**DSC**	**PPV**
Chang [49]	SS	T	0.891 ± 0.106	0.662 ± 0.258	**0.589± 0.199**	0.598 ± 0.263
	SS	L	0.856 ± 0.127	0.457 ± 0.389	**0.435 ± 0.376**	0.464 ± 0.431
Garg [133]	SS	T	0.524 ± 0.157	0.955 ± 0.077	0.411 ± 0.074	0.264 ± 0.060
	SS	L	0.421 ± 0.206	0.772 ± 0.370	0.388 ± 0.212	0.266 ± 0.161

⋆ Results are averaged over both experts.
SS = Supervised Set. T = Transverse. L = Longitudinal. SE = Sensitivity. SP = Specificity. DSC = Dice Co-efficient. PPV = Positive Predictive Value.
Best results are highlighted in **Bold** font.

Table 3.5: Quantitative analysis of existing methods to segment the thyroid gland. It can be inferred from the DSC overlap values that the method proposed by Chang et al., [49] performs better than that of Garg et al., [133] to segment the thyroid gland.

Organ	GT	**SE**	**SP**	**DSC**	**PPV**
Thyroid	E12	0.993 ±0.006	0.832 ±0.088	0.890 ±0.050	0.961 ±0.029
	E21	0.965 ±0.024	0.961 ±0.029	0.889 ±0.051	0.833 ±0.088
Carotid	E12	0.999 ±0.001	0.855 ±0.072	0.901 ±0.030	0.959 ±0.042
	E21	0.996 ±0.002	0.959 ±0.042	0.901 ±0.049	0.856 ±0.072
Muscles	E12	0.997 ±0.002	0.620 ±0.146	0.749 ±0.117	0.978 ±0.021
	E21	0.926 ±0.026	0.978 ±0.021	0.749 ±0.117	0.620 ±0.146
Trachea	E12	0.990 ±0.009	0.786 ±0.181	0.833 ±0.173	0.895 ±0.186
	E21	0.974 ±0.019	0.895 ±0.186	0.833 ±0.173	0.786 ±0.181

E12: Manual segmentation by expert 1 used as GT to quantify the segmentation by expert 2; E21: Manual segmentation by expert 2 used as GT to quantify the segmentation by expert 1.
GT = Ground Truth. SE = Sensitivity. SP = Specificity. DSC = Dice Co-efficient PPV = Positive Predictive Value.

Table 3.6: Interobserver variation.

mentation algorithms. A novel annotation removal algorithm was proposed in this chapter that can automatically detect and remove annotations from ultrasound images. High quality restoration (PSNR > 38dB) of the annotation removed images was achieved by the use of POCS algorithm.

- Four qualitative metrics and Four quantitative metrics were introduced to validate the segmentation algorithms. It is found that the existing methods to segment the thyroid gland and the carotid artery provide good quality segmentation results but fall short of segmenting the actual boundaries of the organs under consideration. This is evidenced by the low DSC values of the segmentation results suggesting

the under/over segmentation of the organs. Hence better methods are needed to accurately segment the organs.

- All the segmentors and all organs are considered together at the time of computing the qualitative metrics. Thus, these can be used to compare different algorithms with each other. However, since all the organs are together considered for qualitative analysis, these metrics cannot be used to compare multi-organ segmentation algorithms with binary segmentation algorithms. The quantitative metrics on the other hand are specific to each segmentor and organ. The quantitative metrics with respect to individual segmentors have to be: (a) considered separately; or (b) averaged; to compare different algorithms with each other. Unlike the qualitative metrics, the quantitative metrics can used to compare the results of a multi-organ segmentation algorithm with that of a binary segmentation algorithm.

- The qualitative metrics are used to select the best of the two methods proposed in Chapter 4 while the quantitative methods are used to compare the performance of the proposed methods with the existing methods.

- Manual segmentation by two expert sonographers are used as ground truth to validate the segmentation algorithms. It was observed that one of the experts consistently under-segmented the organs when compared to the other expert.

Chapter 4

Unsupervised MODS methods

In this chapter two new methods called the: (a) Three Category MODS system; and (b) Multi-Category MODS System; are introduced to perform MODS in US images of the thyroid gland. A cascade of unsupervised classifiers are employed to perform the segmentation in the proposed methods.

4.1 State-of-the-art

A majority of US image segmentation algorithms are based on level-sets [154–156] and active contours [157, 158]. These methods are semi-automatic and are restricted to perform binary segmentations in images. The concept of multi-organ segmentation is relatively new and only a handful of algorithms have been proposed for CT and US imaging modalities.

One of the early attempts to perform MODS in US images was made by Georgescu et al. [159]. This was followed by the introduction of methods to segment the foetal anatomic structures in US images of the abdomen by Zhou et al. [69] and Carniero et al [160, 161]. These methods employ a technique called the marginal space learning (MSL) to detect and segment objects of interest. A bounding box given by $B = (x, y, \alpha, \sigma)$ is used to describe the appearance of the object in the image. Hence, segmentation is formulated as placing a rectangular bounding box around the detected anatomical object. Here, x and y refer to the co-ordinates of the top left corner of the bounding box, α refers to the orientation of the box with respect to the image\volume and σ refers to the scale of the bounding box. Figure 4.1 shows a generic MSL framework where a cascade of classifiers are employed to perform MODS in US images. In the MSL framework, each classifier is trained on a subset of the bounding box parameter space $\theta = \{x, y, \alpha, \sigma\}$. Referring to Figure 4.1, the position classifier acts only on the x, y parameters of the input samples. The samples classified as belonging to the positive class form input to the orientation classifier. As the name suggests, the orientation classifier filters input candidates based on its orientation with respect to the base of the image. The candidates filtered by the orientation classifier are fed into the scale classifier which outputs if the region of the image under consideration belongs to an anatomical or not. It has been reported

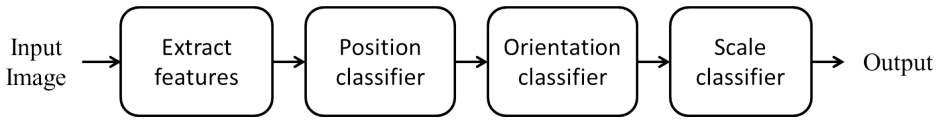


Figure 4.1: MSL Framework for MODS in US images.

that the computational complexity of the algorithm is reduced three fold with the use of MSL [160]. Image features based on Haar wavelets are used in the classification process. Figure 4.2 shows the commonly used kernels to generate Haar wavelet based features. The Probabilistic Boosting Tree (PBT) is used as the classifier in all three stages of the framework.

Feng et al., [162] have proposed another cascade of random forest classifiers to segment the US images of the liver. In this method, the first classifier of the cascade is trained on features extracted from regions belonging to the super pixels as determined by the Simple Linear Iterative Clustering (SLIC) algorithm [163] to segment the liver parenchyma. The second classifier is trained on features extracted from maximally stable extremal regions (MSER) [164] to segment various organs within the segmented liver parenchyma. Rahmatullah et al. [165] use a two stage cascade of AdaBoost classifiers to segment US images of the foetus. The first stage classifier makes use of the global features obtained from Haar wavelets to detect possible regions containing anatomical structures of interest. The second stage classifier which is trained on local features such as local phase and local energy, is used to detect the presence of organs in the candidate regions. In all of the methods, intensity based Haar wavelet features are used in the classification process. This would mean that the performance of the existing methods is closely linked to the brightness of the image. Hence, existing methods perform sub-optimally on datasets that have been acquired at different time gain compensation settings and hence cannot be generalized. Although authors in [165] use local phase which is invariant to signal energy to categorize the organs, the primary detector is still based on Haar wavelets and hence is expected to perform poorly on multi-vendor datasets. It is interesting to note that none of the algorithms encode the relative spatial positions of the organs as a feature in the classification process.

Similar to the MODS in US images, the methods proposed for CT imaging modality also employ a cascade of classifiers to detect and segment multiple organs. Liu et al. [166] have proposed a maximum a posteriori (MAP) framework that makes use of probabilistic atlases and Gaussian Mixture Models to perform MODS. Kohlberger et al., [167] employ MSL with PBT to detect regions containing anatomical structures of interest in CT volumes of liver, lungs and kidney. Level sets are then used to segment organs within the detected regions. Wels et al., [168] follow an approach similar to [167] to segment MRI volumes of the brain. Criminisi et al., [169] make use of decision forests trained with Haar wavelets to perform MODS in CT volumes. Glocker et al. [170] have proposed a Joint Classification and Regression (JCR) model to segment multiple organs in CT volumes of the abdomen. Of the methods discussed thus far, both for CT and US imaging, this is perhaps the only algorithm that makes use of the relative spatial locations of the

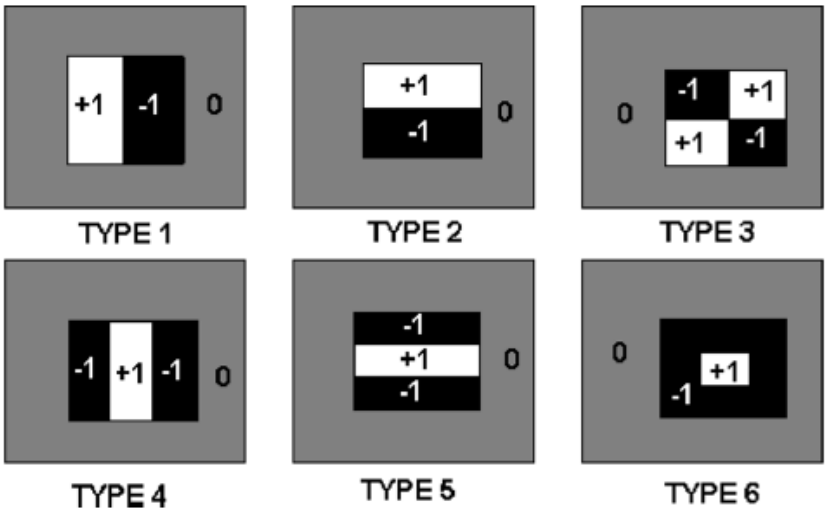


Figure 4.2: Image features used in [160]

organs to perform MODS. Intensity based Haar features in conjunction with relative spatial locations of organs determined from signed distance maps are used to build the JCR model. Since the methods introduced in this Chapter employ spatial constraints to detect and segment organs in US images, the JCR model is modified for use in US images and compared with our methods in Section 4.5.

There have been attempts to employ shape contexts to perform MODS in CT imaging modality [171, 172]. Use of shape contexts for MODS in freehand US imaging is challenging for two reasons: (a) additional instrumentation is necessary to capture the probe position in world coordinates at the time of image acquisition; and (b) organs in a US scan undergo deformation under probe pressure. The imaging plane determines the shape of the organ visible under the US scan. This is evident from the images in Figure 3.1 where the shapes of the thyroid gland are different in two different imaging planes. In order to use shape contexts, the imaging plane has to be determined by the use of markers attached to the probe which are tracked with the help of an external tracking device such as a video camera. This information can be used to categorize the images according to the imaging plane and a cascade of classifiers can be used to detect and segment organs in each plane.

Shape information for supervised segmentation can be used if the deformation can be successfully modelled and corrected. Deformation of tissues in US images vary with the applied pressure. There are no standards set for the optimal pressure that has to be applied by the sonographers to scan the gland. Therefore, it is probable for two different sonographers to apply different pressures while scanning the gland. Under such circumstances, a shape-based supervised segmentation algorithm trained using the dataset of one sonographer may not perform well on the dataset of another sonographer, which renders the algorithm to be of little use in a clinical setting.

The methods that have been proposed thus far employ supervised learning to perform MODS. Pixels within the regions of the image marked as belonging to an anatomical structure by an expert are treated as samples for the positive training class. Negative training samples are generated automatically from regions of the image that are located at an empirically determined distance from the anatomical structure of interest. It was seen in Section 3.5.2.2 that classifiers trained with automatically generated samples for the negative training class resulted in sub-optimal performance of the algorithm. Moreover, supervised segmentation algorithms require a large database of annotated datasets to train the classifiers. Acquiring such a large database is expensive and time consuming, thus prompting the need for unsupervised methods for MODS.

Not only do the unsupervised methods not need large databases of annotated datasets, they can also be used for exploratory data analysis. Unsupervised algorithms have previously been employed to perform tissue characterization and binary segmentation [173–177]. Mean shift based segmentation algorithms have been used in [178] and [179] to segment CT and US images respectively. Linguraru et al., [180] make use of graph cuts to perform multi-organ segmentation in CT images of the abdomen. Unsupervised clustering based segmentation algorithms such as Mean-Shift [181] and Normalized Cuts [182] can effectively be extended to perform multi-organ segmentation if the clusters are formed based on some criteria related to image formation and tissue characterization. Motivated by the discussion on speckles and artefacts in Chapter 2, three methods are proposed in this chapter that utilize pixels associated with speckles and imaging artefacts to perform MODS in US images.

4.2 Speckle related pixels

In all of our methods, we cluster speckle related pixels to determine the tissue echogenicity. It was seen in Section 2.2 that speckles in US images directly correspond to the number of scatterers in the tissue and hence determine the echogenicity of the tissue. Given the nature of the scattered US waves to undergo constructive or destructive interferences, it can be assumed that the speckles appear blob like in B-Mode US images. Hence, a Hessian based blob detector is used to determine the set of pixels associated with speckles. Given an input image $f(\mathbf{x})$, the set of pixels belonging to a speckle is given by:

$$\mathcal{F}_s = \{\mathbf{x} : |H(f)| > 0 \text{ and } \frac{\partial^2 f(\mathbf{x})}{\partial^2 x} > 0\} \tag{4.1}$$

where,

$$H(f) = \begin{pmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial^2 x} & \frac{\partial^2 f(\mathbf{x})}{\partial x \partial y} \\ \frac{\partial^2 f(\mathbf{x})}{\partial y \partial x} & \frac{\partial^2 f(\mathbf{x})}{\partial^2 y} \end{pmatrix} \tag{4.2}$$

is the Hessian matrix and $|H(f)|$ is its determinant.

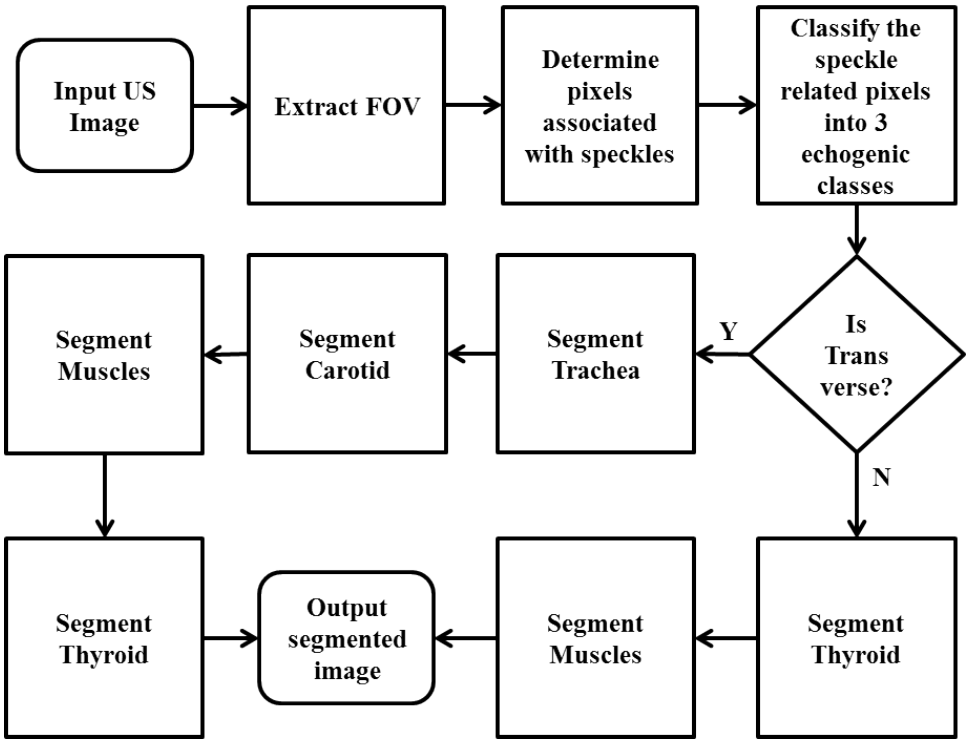


Figure 4.3: Three category clustering based MODS framework

4.3 Three Category MODS system

In the first of the three methods, the set of speckle related pixels, $\mathcal{F}_s$, are clustered into one of the three echogenic levels: (a) Anechoic; (b) Hypoechoic; and (c) Hyperechoic. The framework proposed to perform MODS by this method is depicted in Figure 4.3. Given the fact that the echogenicity of a tissue is expressed in terms of its brightness with respect to the tissue surrounding it, the goal is to cluster speckle related pixels based on its intensity profile in a neighbourhood of pixels associated with speckles. The average value of speckle pixels in the neighbourhood is used to determine its echogenicity.

Denoting the neighbourhood of a speckle related pixel by a window $w(m,n), m \in [1, N_x]$ and $n \in [1, N_y]$, centred at (x, y), the average intensity in the window is given by:

$$f_{\text{avg}}(\mathbf{x}) = \frac{1}{N_x N_y} \sum_{x-\frac{N_x}{2}}^{x+\frac{N_x}{2}} \sum_{y-\frac{N_y}{2}}^{y+\frac{N_y}{2}} f(\mathbf{x}). \tag{4.3}$$

The coordinates of a pixel (x, y) is denoted by a row vector $\mathbf{x}$ for the sake of simplicity. A speckle related pixel is said to be:

(1) hyperechoic if $n_{\text{b}} > n_{\text{a}} + n_{\text{e}}$

(2) anechoic if $n_{\text{a}} > n_{\text{b}} + n_{\text{e}}$ and

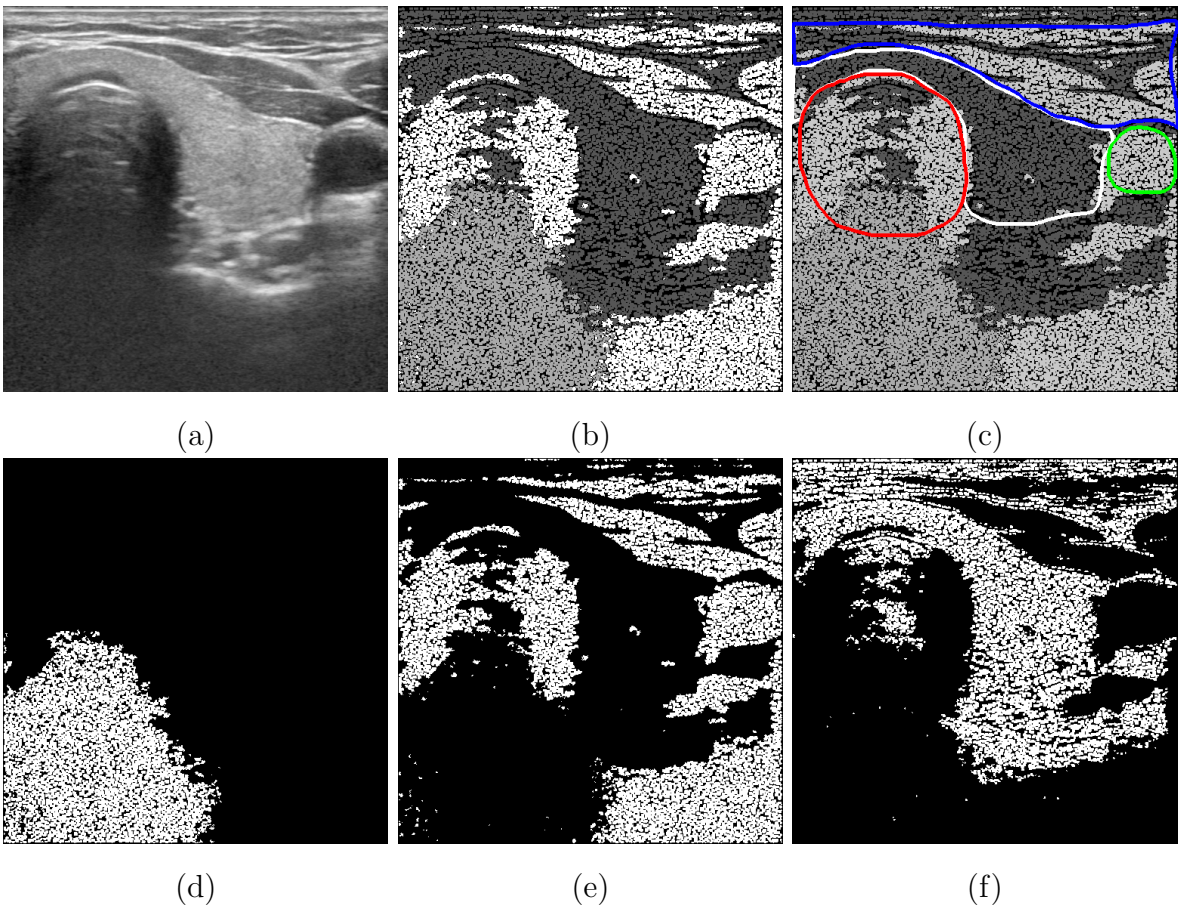


Figure 4.4: Three category clustering of speckle related pixels. (a) Input US image in the transverse scan; (b) Results of clustering where pixels with: (i) lowest brightness corresponds to anechoic regions, (ii) intermediate corresponds to brightness to hypoechoic regions and (iii) highest brightness corresponds to hyperechoic regions; (c) Clustering results with GT overlay (green = carotid; white = thyroid; blue = muscles; and red = trachea); (d) Binary image $f_{\text{an}}(\mathbf{x})$ with pixels belonging to anechoic regions as foreground pixels; (e) Binary image $f_{\text{ho}}(\mathbf{x})$ with pixels belonging to hypo-echoic regions as foreground pixels; and (f) Binary image $f_{\text{hr}}(\mathbf{x})$ with pixels belonging to hyper-echoic regions as foreground pixels.

(3) hypooechoic if $n_{\text{e}} > n_{\text{a}} + n_{\text{b}}$,

where, n_{a} is the number of pixels that have an average value above $f_{\text{avg}}(\mathbf{x}) + \tau$; n_{b} is the number of pixels that have an average value below $f_{\text{avg}}(\mathbf{x}) - \tau$; and n_{e} is the number of pixels that have an average value $f_{\text{avg}}(\mathbf{x}) : f_{\text{avg}}(\mathbf{x}) - \tau \leq f_{\text{avg}}(\mathbf{x}) \leq f_{\text{avg}}(\mathbf{x}) + \tau$. The binary images $f_{\text{hr}}(\mathbf{x})$, $f_{\text{ho}}(\mathbf{x})$ and $f_{\text{an}}(\mathbf{x})$ that respectively have hyper, hypo and anechoic pixels as foreground pixels are shown in Figure 4.4.

Looking at Figure 4.4(c) it can be seen that the carotid artery, trachea and muscles are entirely contained in the hypoechoic regions while the thyroid is entirely contained in

(a) (b) (c)

Figure 4.5: Detecting and segmenting the trachea. (a) Image depicting the relationship between the trachea in the Hypo-echoic regions and the Anechoic regions. Here, the region in Gray denotes the anechoic regions, the region in white denoted the Hypo-echoic regions and the GT overlay for the trachea is in red. (b) Detected trachea. (c) Segmented trachea (after setting the pixels in the convex hull of the detected trachea as foreground pixels).

the hyperechoic region. The methods to detect the carotid artery, trachea and muscles in $f_{\text{ho}}(\mathbf{x})$ and to estimate the boundary of the thyroid gland in $f_{\text{hr}}(\mathbf{x})$ are discussed in the Sections to follow.

4.3.1 Segmenting the Trachea

The trachea is the first organ to be segmented in this method. From Figures 4.4(b-e), it can be seen that the trachea in $f_{\text{ho}}(\mathbf{x})$ (in red in Figure 4.4(c)) sits on top of the anechoic pixels in $f_{\text{an}}(\mathbf{x})$. In order to detect and segment the trachea, connected component labelling is applied on $f_{\text{ho}}(\mathbf{x})$ to get a labelled image $f_{\text{l\_ho}}(\mathbf{x})$. The component of $f_{\text{l\_ho}}(\mathbf{x})$ that shares its lower boundary with $f_{\text{an}}(\mathbf{x})$ is the trachea. Pixels belonging to this component and its convex hull in $f_{\text{ho}}(\mathbf{x})$ are segmented the as trachea. Figure 4.5 illustrates the steps leading to the segmentation of the Trachea.

4.3.2 Segmenting the Carotid

The carotid artery is entirely contained in $f_{\text{ho}}(\mathbf{x})$ and casts an enhancement artefact underneath it. The pixels of the artefact belong to $f_{\text{hr}}(\mathbf{x})$. The carotid artery, although circular in shape, appears as an ellipse due to deformation under the applied probe pressure. Hence, a two stage approach is followed to detect the carotid in $f_{\text{ho}}(\mathbf{x})$. In the first stage, the components of $f_{\text{l\_ho}}(\mathbf{x})$ that resemble an ellipse with hyperechoic regions below it are chosen as probable candidates for the carotid artery. In the second stage, an energy based deformable model is applied to detect the carotid artery from the short listed candidates.

Looking at Figure 4.4, it can be seen that none of the regions that belong to the foreground resemble an ellipse. So to detect the elliptical carotid candidates in $f_{\text{ho}}(\mathbf{x})$, the image is first morphologically filled and skeletonized. The nodes of the skeleton are represented by the set $\mathcal{N}$. Every node in $\mathcal{N}$, an ellipse encompassing the node and its adjacent node is fit by treating the two nodes as its focii. The major axis of the ellipse is along the line joining the adjacent nodes. Its length, a, is determined to be twice the distance between the midpoint of the line segment joining the focii and the boundary pixel in $f_{\text{ho}}(x, y)$ that is closest to the two focii. The length of the minor axis, b, of the ellipse is found in a similar way on the line perpendicular to the major axis and passing through the mid point of the focii. The ellipses that have a hyperechoic region below it are chosen as candidates for the carotid artery.

Denoting the binary image containing elliptical carotid candidates by:

$$E(\mathbf{x}) = \begin{cases} 1 & \text{inside ellipse} \\ 0 & \text{outside ellipse} \end{cases} \tag{4.4}$$

the energy functional F for the deformable model is defined as:

$$F = \int_{\text{inside ellipse}} |f(\mathbf{x}) - \bar{E}_i|^2 dxdy + \int_{\text{outside ellipse}} |f(\mathbf{x}) - \bar{E}_o|^2 dxdy \tag{4.5}$$

where

$$\bar{E}_i = \frac{\int f(\mathbf{x}) f_{\text{ho}}(\mathbf{x}) E(\mathbf{x}) dxdy}{\int E(\mathbf{x}) dxdy} \tag{4.6}$$

and

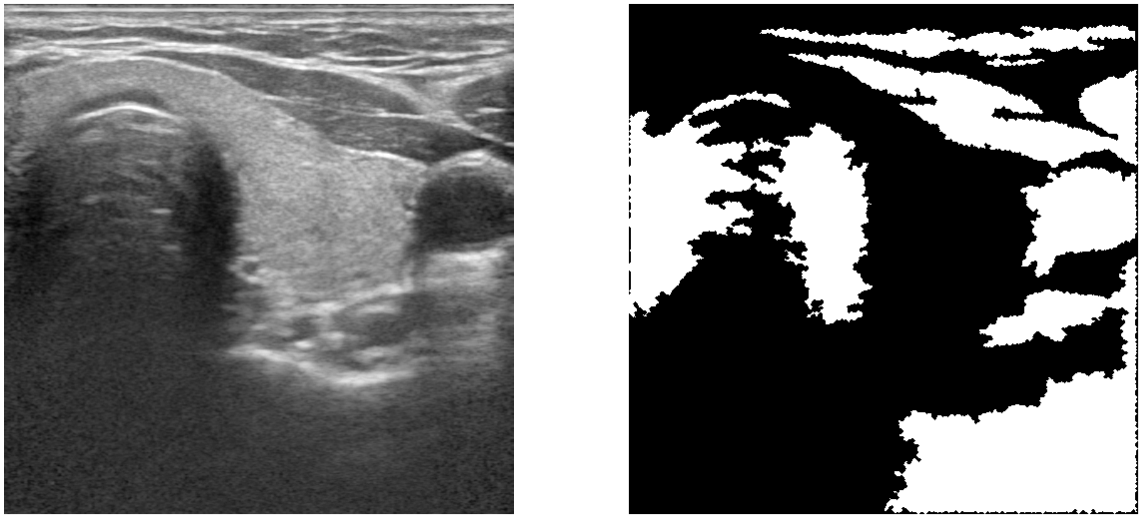
$$\bar{E}_o = \frac{\int f(\mathbf{x}) f_{\text{ho}}(\mathbf{x}) (1 - E(\mathbf{x})) dxdy}{\int (1 - E(\mathbf{x})) dxdy}. \tag{4.7}$$

The elliptical carotid candidate that has the lowest value of F and that has the lowest internal energy (first term in the sum of Eq. (4.5)) is segmented as the carotid. Figure 4.6 shows the process of segmenting the carotid artery.

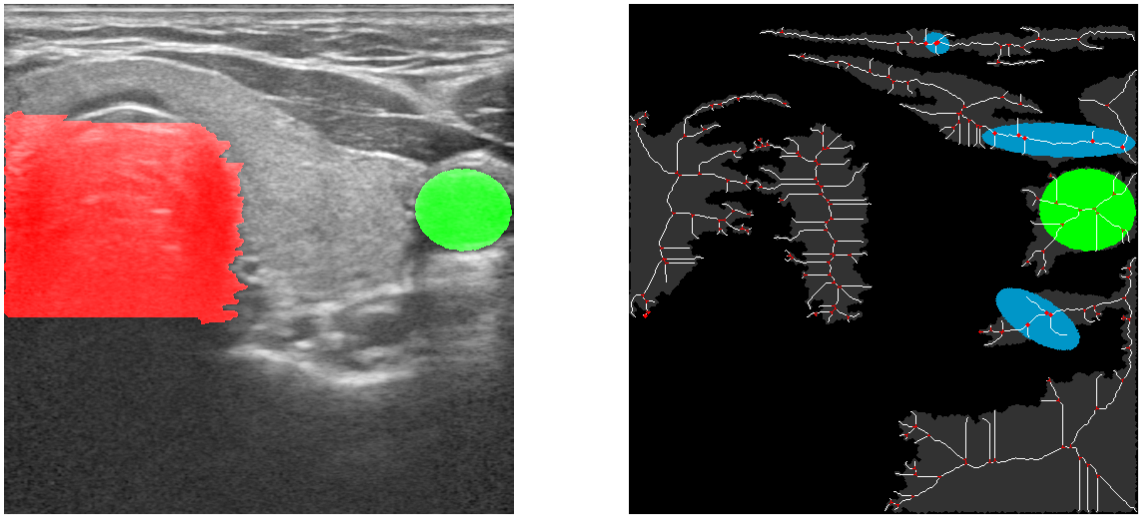
4.3.3 Segmenting the Muscles

In Transverse Scans: The carotid artery and trachea detected previously are used as landmarks to detect and segment the muscles. Muscles in transverse US scan lie above the thyroid, trachea and carotid artery. So the search region for the muscles is established above the highest point of the trachea or carotid artery in $f_{\text{l\_ho}}(\mathbf{x})$. The longest edge in $f_{\text{ho}}(\mathbf{x})$ above the carotid and trachea obtained by an edge detector belongs to the muscles. The component that contains this edge and the components that lie within the bounding box containing it in $f_{\text{l\_ho}}(\mathbf{x})$ are the muscles. The canny edge detector is used to detect the longest edge in our experiments.

In Longitudinal Scans: Since the carotid and trachea are not visible in the longitudinal



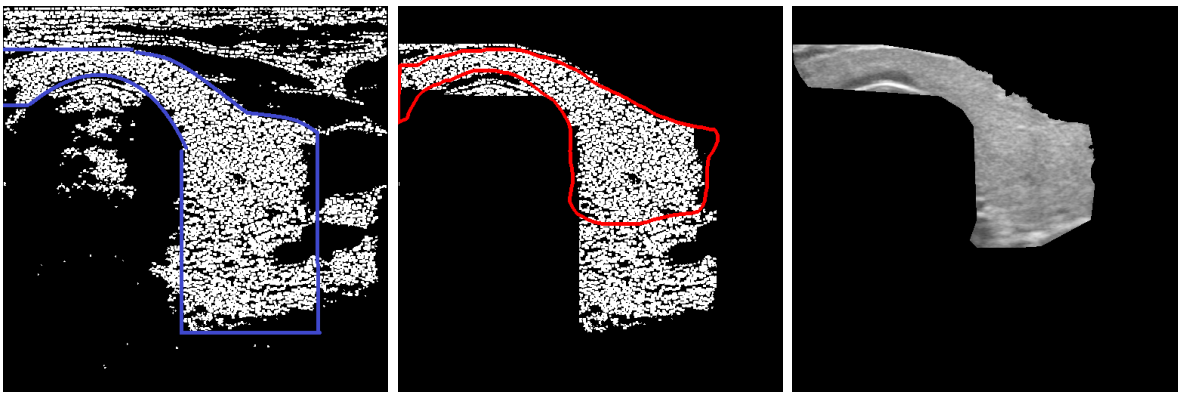
(a) (b)



(d) (c)

Figure 4.6: (a). Input ultrasound image in transverse scan; (b). Binary image $f_{\text{ho}}(\mathbf{x})$ representing hypo-echoic regions after morphologically filling the holes; (c). Skeleton of $f_{\text{ho}}(\mathbf{x})$ with node pixels (red), carotid candidates (blue ellipses) and carotid candidate having the lowest internal energy (green ellipse); (d). Segmented trachea and carotid

scan, the thyroid gland needs to be segmented first before segmenting the muscles. Once the thyroid is segmented, all the hypoechoic components of $f_{\text{l\_ho}}(x, y)$ having a height to



(a) (b) (c)

Figure 4.7: Establishing search region to detect and segment the thyroid gland. (a) Search region in the binary hyper-echoic image $f_{hr}(\mathbf{x})$ (region inside blue boundary); (b) GT over lay in the search region (in red); and (c) Segmented thyroid gland.

width ratio greater than half are chosen as that belonging to the muscles and segmented from the image.

4.3.4 Segmenting the Thyroid

In Transverse Scans: The thyroid gland shares two of its four boundaries with the muscles and the trachea. Hence, muscles and trachea are used as landmarks to segment the thyroid gland. Further, the thyroid gland is fully contained in the space enclosed by the trachea, muscles and the carotid artery. So the space to be segmented is restricted to the region in between these organs in $f_{hr}(\mathbf{x})$. Figure 4.7 shows the search space for the thyroid gland in $f_{hr}(\mathbf{x})$. From Figure 4.7(b), it can be seen that three boundaries of the thyroid gland are fully contained in the search space leaving only the lower boundary of the thyroid gland to be estimated.

The lower boundary of the thyroid gland is estimated by considering a signal g_y:

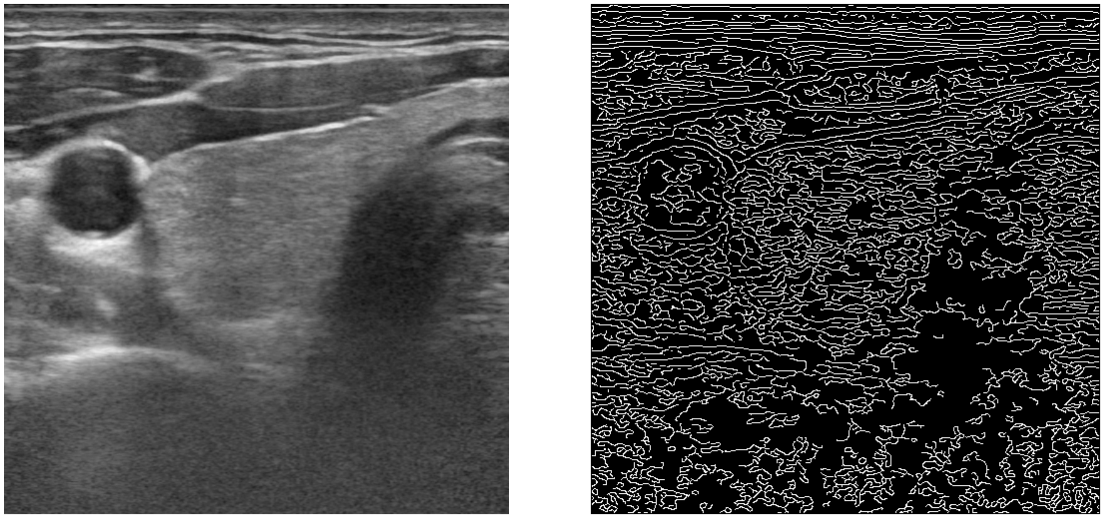
$$g_y(x_i) = f(x_i, y) \quad , 1 \leq y \leq N, x_i \in \mathcal{N}_x \text{ and } i \in [1, |\mathcal{N}_x|], \tag{4.8}$$

where,

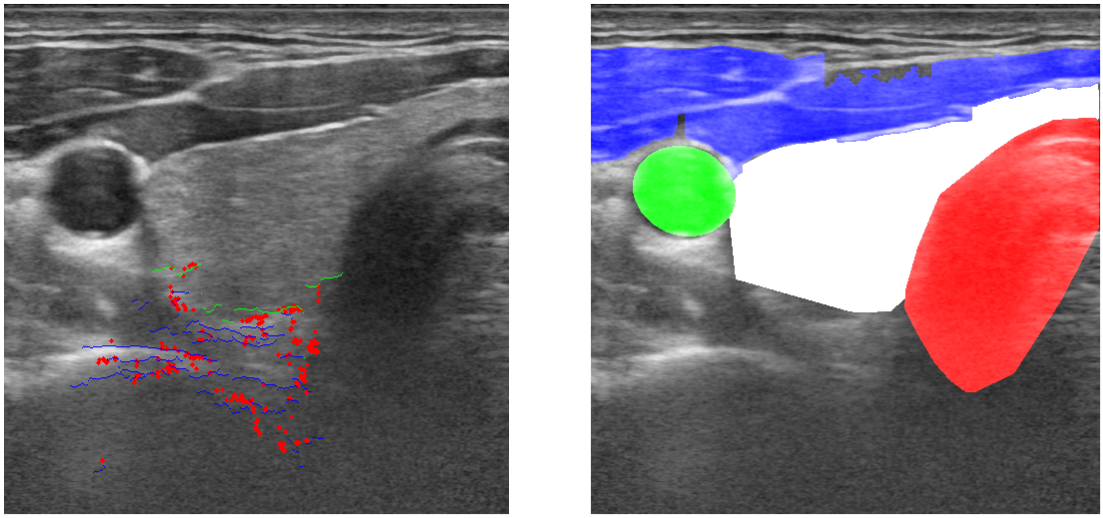
$$\mathcal{N}_x = \{x : f_{\text{hr}}(x, y) > 0 | y\} \tag{4.9}$$

denotes the set of all non-zero pixels in a single column of $f_{\text{hr}}(x, y)$. Thus, $g_y(x_i)$ is the intensity of a pixel at (x_i, y) in the input image $f(\mathbf{x})$. A low pass filter is then applied on the signal g_y with an empirically chosen cut off frequency equal to $\frac{|n_x|}{15}$. Let us denote the low pass filtered signal by g_{yl}. All the edge pixels in $f(\mathbf{x})$ (determined by the Canny edge detector) at x_{y_m}:

$$x_{y\_m} = \text{argmax}_{n_x} | \bigtriangledown g_{yl}|, \tag{4.10}$$



(a) (b)



(c) (d)

Figure 4.8: (a). Input ultrasound image in transverse scan; (b). Edge map after applying Canny edge detector to the input in (a); (c). Speckle pixels associated with maximum intensity gradient in each column (red), edges associated with these pixels and selected anterior boundary of thyroid (green); (d). Fully segmented image of thyroid gland

are possible candidates for the lower boundary of the thyroid. The edge pixels that are closer to the trachea and carotid artery are retained and a smooth curve is fit using splines

to estimate the lower boundary of the thyroid gland. All the pixels of $f_{\text{hr}}(x, y)$ that lie underneath the estimated anterior boundary are excluded from further analysis. The segmented thyroid is obtained by including all the pixels of $f_{\text{hr}}(x, y)$ that lie within the convex hull of the region bounded by the lower boundary of the muscles, upper boundary of the trachea, the carotid and the estimated lower boundary of the thyroid. Figure 4.8 shows the process of estimating the lower boundary of the thyroid gland along with the fully segmented image in the transverse scan.

In Longitudinal Scans: The thyroid gland is segmented in the longitudinal scans by considering both the hyperechoic and anechoic regions. A binary image $f_{\text{b}}(\mathbf{x})$ is formed with the speckle pixels of $f_{\text{hr}}(\mathbf{x})$ and $f_{\text{an}}(\mathbf{x})$ as foreground pixels. This image is morphologically filled and labelled by connected component analysis to get a labelled image $f_{\text{bl}}(\mathbf{x})$. The largest component corresponds to the thyroid and is retained as foreground pixels in $f_{\text{b}}(\mathbf{x})$. The rest are reset to background pixels. Similar to the segmentation of the thyroid gland in the transverse scan, a signal $g(x)$, given by:

$$g(x) = \sum_{y=1}^{N} f_{\text{b}}(x, y), \tag{4.11}$$

is low pass filtered to get a signal $g_l(x)$.

The spatial locations in $g(x)$ at which the signal has a local minima corresponds to the rows in $f(\mathbf{x})$ that may contain the lower boundary of the thyroid gland. The row in the image at which the difference between the successive minima values is the largest is the deepest point of the thyroid gland in the image. The foreground pixels in $f_{\text{b}}(\mathbf{x})$ below this point are reset to background pixels. The pixels contained in the convex hull of the remaining foreground pixels in $f_{\text{b}}(\mathbf{x})$ form the segmented thyroid gland. Figure 4.9 shows the process of segmenting the thyroid in longitudinal scans.

4.4 Multi-Category MODS system

In the previous section, the pixels in $\mathcal{F}_s$ were clustered into one of three echogenic classes based on its intensity. The results of this clustering were used to sequentially detect and segment the trachea, carotid artery, muscles and the trachea. The success of the algorithm depends on the successful detection of the trachea and the carotid artery that act as landmarks to detect the remaining two organs. While the detection of the trachea is straightforward, the detection of the enhancement artefact is crucial to successfully detect the carotid artery. Although, the enhancement artefact is hyperechoic to the thyroid gland (as a matter of fact, the enhancement artefact is hyperechoic to all the tissues in the scan), it is assigned the same echogenic class as the thyroid in the process of clustering the pixels. This resulted in the detection of a large number of false positive carotid candidates, some of which were falsely detected and segmented as the carotid artery by the subsequent energy based model. Thus, three echogenic classes are insufficient to capture the entire echogenic range of the tissues present in the scan. But the number of echogenic levels in a US image rarely remains constant and varies with

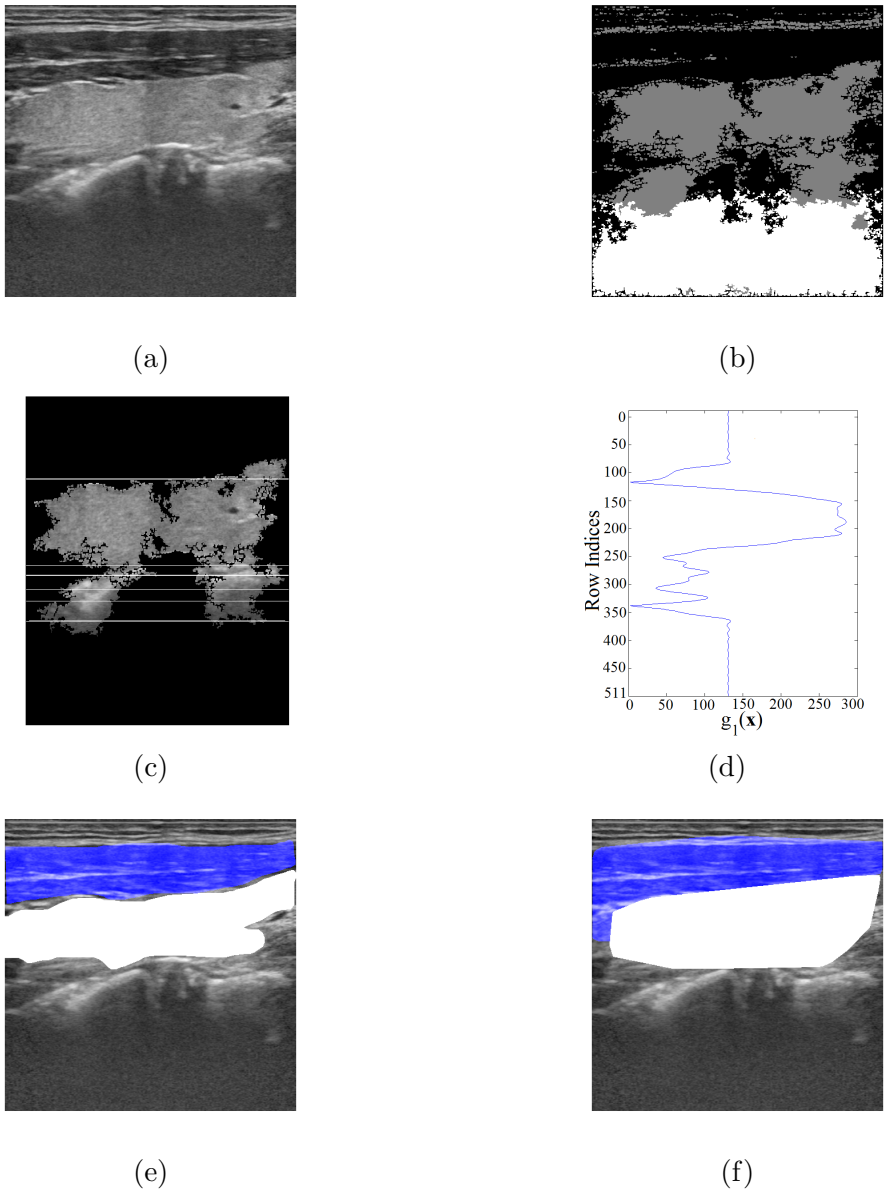


Figure 4.9: (a). Input ultrasound image in longitudinal scan; (b). Image with regions representing hyper-echoic and anechoic pixels as foreground; (c). Thyroid with local minima rows (in white) ; (d). $g_l(x)$ (On x-axis is the number of non-zero pixels and y axis is the row number ranging from 1 to the total number of rows in the image); (e) Manual segmentation (Ground truth) of thyroid (in white) and muscles (in blue) by an expert; (f). Thyroid (in white) and muscles (in blue) segmented by the three category clustering based segmentation algorithm.

probe motion and the image acquisition settings. This makes echogenicity based tissue characterization all the more challenging.

In order to deal with the dynamic nature of echogenic levels, a multi-category MODS method is proposed in this section that makes use of an agglomerative clustering scheme to determine the number of echogenic levels in the image. As before, the results of the clustering algorithm are used to detect and segment the four organs in the image. Tissues in the image that are iso-echoic regions are called Similarly Reflective Regions (SRR) and the number of clusters formed at the end of the agglomerative clustering process determines the number of SRR's present in the image. This can in a way be considered as a quantization step in the MODS process. The multi-category MODS framework is illustrated in the flow diagram of Figure 4.10.

4.4.1 Speckle patch similarity

In this algorithm, unlike clustering pixels based on its intensities, a similarity constraint is used to cluster speckle related pixels. This constraint is termed as the Speckle Patch Similarity (SPS) in the context of this research work and is given by:

$$s_{mn} = S(\mathbf{r}_m, \mathbf{r}_n) \leq \tau, \tag{4.12}$$

where,

$$S(\mathbf{r}_m, \mathbf{r}_n) = \sum_{k=1}^{M_{R1}^2} (\mathbf{r}_m(k) - \mathbf{r}_n(k))^2 \quad , m \neq n, \tag{4.13}$$

and τ is a pre-defined threshold. Here, $\mathbf{r}_m(k)$ refers to the intensity of the k^{th} pixel in an $M_{R1} \times M_{R1}$ ROI defined around the m^{th} speckle pixel in $\mathcal{F}_s$. The SPS constraint is applied to cluster speckle pixels into their respective echogenic levels based on the assumption that a small patch of image surrounding a speckle in an US image is similar to another patch of the same tissue or of any other iso-echoic tissue. This assumption is justified by referring to the image in Figure 4.11 which shows a simple pattern having four brightness levels corrupted by speckle noise. It can be observed that the patches of the image R1 and R2 around two speckle related pixels of the region with the same brightness are similar to each other. Similarly, the patches R3 and R4 in another region are similar to each other in terms of the brightness levels contained in the patch. By grouping all the patches of image around every pixel associated with a speckle that satisfy a the constraint of Eq. (4.12) under one label, the number of SRR's in the image is determined to be the total number of labels generated at the end of the grouping. The process of determining the number of SRR's is outlined in Algorithm 1.

Patch based segmentation algorithms for ultrasound imaging was first introduced by Ciurte et al. [183] to segment opthalmic 2D ultrasound images in an active contour framework. This method was further modified by the same authors in [105] by replacing the active contour formulation with a graph based algorithm. In an earlier discussion on medical image segmentation algorithms based on basic principles (Chapter 2, Sec. 2.5) it was seen that active contours (belonging to GDM) and min-cuts (belonging to GBM) are only effective in a single organ segmentation framework. Here in this research work, we extend the patch based algorithms to perform multi-organ segmentation. Unlike [183]

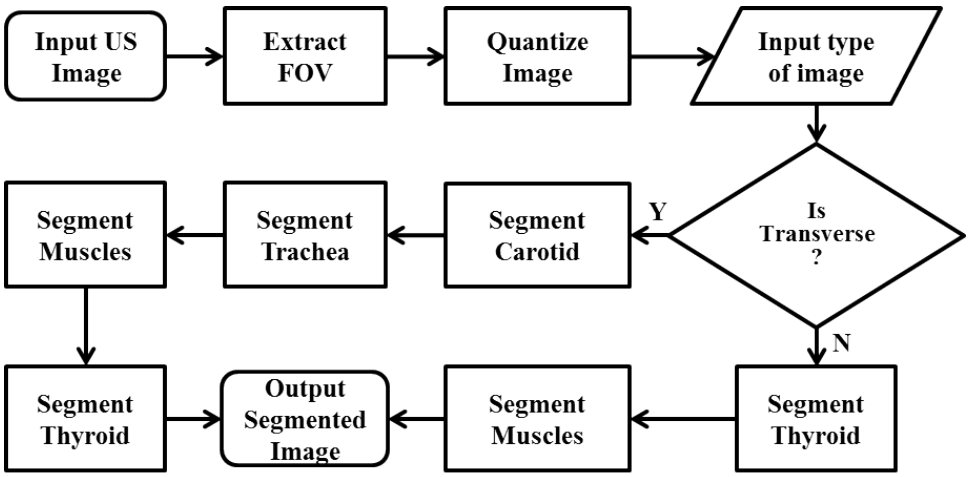


Figure 4.10: SPS Segmentation framework

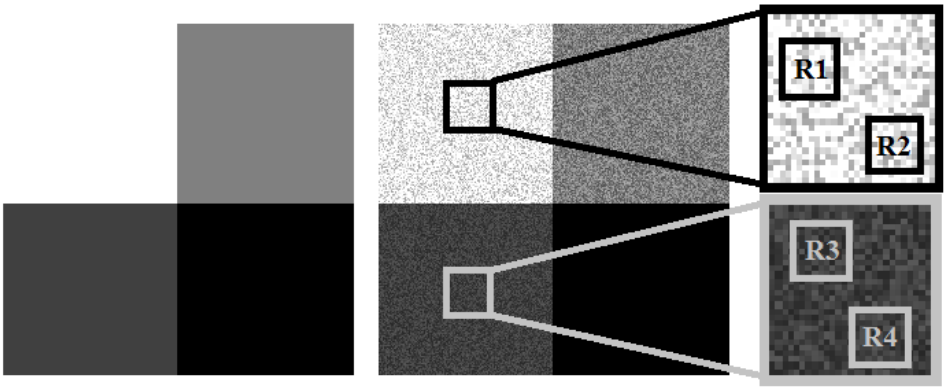


Figure 4.11: Illustration of the Speckle Patch Similarity on a simple image comprising of four brightness levels corrupted with speckle noise with $\sigma = 0.04$.

and [105] where Pearson's distance is used, we use SSD as a similarity metric in our algorithms. Experiments prove that the detection accuracy is higher with SSD metric than with the Pearson's distance as the metric.

The clustering of pixels according to the similarity constraint of Eq. (4.12) results in only an estimate of the number of SRR's in the image. One way to know the correctness of the estimate is through simulations. Figure 4.12 shows a synthetic image with 4 segments corrupted by speckle noise with $\sigma = 0.04$. Speckles within each segment belong to one SRR, hence the expectation is that the clustering algorithm should return 4 clusters. Applying Algorithm 1 to this image results in 29 SRR's instead of 4 SRR's. Thus, the SRR estimate of Algorithm 1 is an overestimate. This overestimation is overcome by merging clusters in a way that clusters belonging to the same SRR are assigned the same label. The class connectivity constraint (Theorem 1) is used to merge the clusters belonging to the same SRR. The pseudo-code in Algorithm 2 outlines the steps to be

Algorithm 1 Obtaining initial SRR estimate

1: **initialize** Choose a pixel $\vec{\mathbf{x}}_m$ from $\mathcal{F}_s$ and let its class label be $l = 1$
2: **repeat**
3: Calculate the pairwise similarity s_{mn}, with $m \neq n$
4: Assign to all unlabelled $\vec{\mathbf{x}}_n$ whose s_{mn} satisfy the SPS constraint, the class label l
5: Increment Class label; $l = l + 1$
6: Choose another unlabelled pixel $\vec{\mathbf{x}}_m$ from the set $\mathcal{F}_s$, assign the new class label to the pixel
7: **until** all the pixels in $\mathcal{F}_s$ have been labelled

taken to merge the clusters.

Theorem 1 (***Class Connectivity***)**.** *Given a set $\mathcal{L}$ of l class labels, if for every pixel associated with a class label λ_m in the image there exists at least one pixel with class label λ_n connected to it, where $m \neq n$ and $(m, n) \in [1, l]$ and $l \in \mathbb{N}$, then the pixels having the labels λ_m and λ_n are said to be class connected and the pixels belonging to the two classes can be merged into a single class λ.*

Proof. Let us denote the SRR's (quantization levels) by a class label set $\mathcal{L}$ and the image with pixels having the values from $\mathcal{L}$ as $f_{\mathrm{L}}(\mathbf{x})$:

$$\mathcal{L} = \{\lambda_1, \lambda_2, ..., \lambda_l\} \tag{4.14}$$

Let the set of pixels associated with each $\lambda_m,\ m \in [1, l]$ be denoted by $\mathcal{X}_m$:

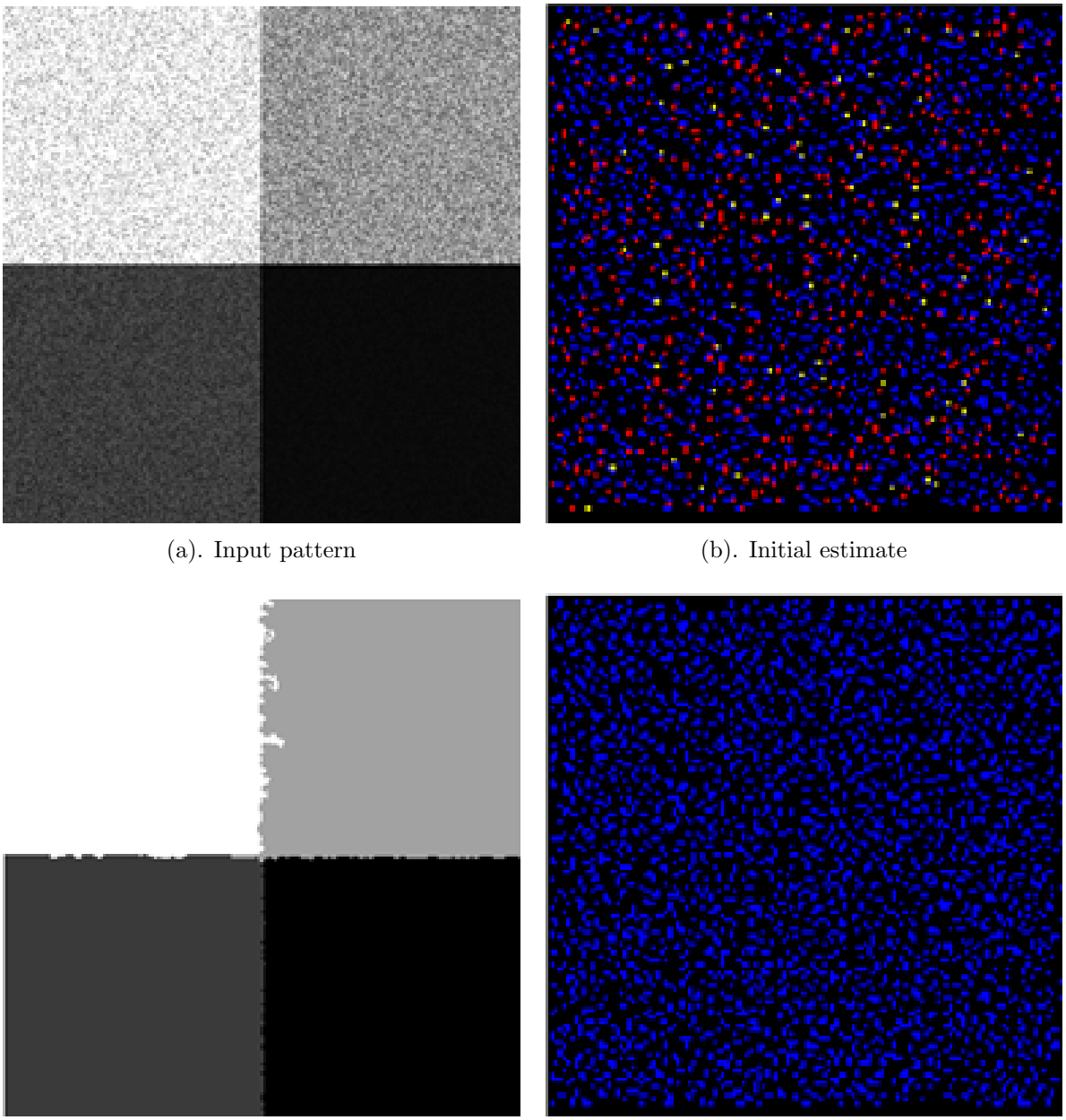
$$\mathcal{X}_m = \{\mathbf{x} : f_{\mathrm{L}}(\mathbf{x}) = \lambda_m\} \tag{4.15}$$
$$\mathcal{F}_s = \{\mathcal{X}_1, \mathcal{X}_2, ..., \mathcal{X}_l\} \tag{4.16}$$

Consider any two pixel sets $\mathcal{X}_m$ and $\mathcal{X}_n$ from $\mathcal{F}_s$ with the labels λ_m and λ_n respectively, where $m \neq n$ and $m, n \in [1, l]$. Let $\mathcal{X}_{mn}$ represent the union of the two sets and $\mathcal{L}_1$ the set of the two class labels:

$$\mathcal{L}_1 = \{\lambda_m, \lambda_n\} \tag{4.17}$$
$$\mathcal{X}_{mn} = \mathcal{X}_m \bigcup \mathcal{X}_n. \tag{4.18}$$

From Eq.(4.16), $\mathcal{X}_{mn} \subset \mathcal{F}_s$.

According to the definition of connected components [140], two pixels $\mathbf{x}_m$ and $\mathbf{x}_n$ are said to be connected in $\mathcal{X}_{mn}$ if there exists a path between them. The connected component can be assigned a new label λ_{mn}. Assuming $|\mathcal{X}_m| > |\mathcal{X}_n|$, when every pixel in $\mathcal{X}_n$ is connected to at least one pixel in $\mathcal{X}_m$, there can be K components in $\mathcal{X}_{mn}$ with



(a). Input pattern (b). Initial estimate

(c). Quantized pattern (d). Final estimate

Figure 4.12: Quantization results on an input pattern with four segments corrupted with speckle noise ($\sigma = 0.04$) showing the intermediate results after obtaining the initial and final SRR estimates

$|\mathcal{X}_{mn}| > K \geq 1$. Let the connected component label set be denoted by $\mathcal{L}_C$ given by:

$$\mathcal{L}_C = \{\lambda_{C1}, \lambda_{C2}, ..., \lambda_{CK}\} \tag{4.19}$$

Denoting $f_{\mathrm{LC}}(\mathbf{x})$ as the image with pixels at locations in $\mathcal{X}_m$ and $\mathcal{X}_n$ having labels in $\mathcal{L}_C$,

we can write,

$$\mathcal{X}_{mn} = \{\mathbf{x} : f_{\mathrm{LC}}(\mathbf{x}) \in \mathcal{L}_C\}. \tag{4.20}$$

From Eq. (4.17) and Eq. (4.18) we can write

$$\mathcal{X}_{mn} = \{\mathbf{x} : f_{\mathrm{L}}(\mathbf{x}) \in \mathcal{L}_1\}. \tag{4.21}$$

Comparing Eq. (4.20) and Eq. (4.21), it can be seen that all the pixels in $\mathcal{X}_{mn}$ that have the labels from the set $\mathcal{L}_C$ in $f_{\mathrm{LC}}(\mathbf{x})$ originally had values from a single set $\mathcal{L}_1$ in $f_{\mathrm{L}}(\mathbf{x})$. This means that for every pixel in $\mathcal{X}_n$ there was at least one pixel in $\mathcal{X}_m$ linked to it by its connected component. So all the pixels in the cluster(SRR) with label λ_n are linked to pixels with label λ_m and hence can be merged into a single super class λ by virtue of its pixel connectivity. □

Algorithm 2 Merging pixels to obtain final SRR estimate

1: **initialize** Construct an empty $l \times l$ accumulator matrix $\mathbf{A}$ with a_{mn} representing each element of $\mathbf{A}$ and $m, n \in [1, l]$
2: **repeat**
3: **for** $m = 1$ *to* l **do**
4: For every pixel having a label λ_m , define a $M_{R2} \times M_{R2}$ ROI around the pixel.
5: Determine the class label of the all pixels in the ROI excluding the pixel under consideration.
6: Increment the value of a_{mn}, if there exists at least one pixel with the class label λ_n, where $n \in [1, l]$, in the ROI.
7: If $a_{mm} < a_{mn}$ then re-assign all the pixels having class label λ_m with λ_n.
8: **end for**
9: **until** there are no more class re-assignments.

The number of cluster labels that remain after merging the clusters gives the final estimate of the number of SRR's in the image. Once the speckle related pixels are clustered into their respective SRR's, morphological operations are used to assign the SRR labels to the remaining pixels. Given an SRR, a binary image is formed by retaining the pixels at that level as foreground pixels and resetting the remaining pixels to background. The holes in the binary image whose spatial locations do not overlap with the pixels belonging to other SRR's are morphologically filled and the pixels are assigned the label of the SRR that is under consideration. The process is repeated until all the SRR's are covered. Let L denote the final estimate of the number of clusters and $I_{\mathrm{avg\_m}}$ denote the average intensity of the m^{th} SRR, then intensity of the pixels in the quantized image $f_{\mathrm{Q}}(\mathbf{x})$ takes on values in the range $[I_{\mathrm{avg\_1}}, I_{\mathrm{avg\_L}}]$ with $I_{\mathrm{avg\_1}} < I_{\mathrm{avg\_2}} ... < I_{\mathrm{avg\_L}}$.

4.4.2 Detecting and Segmenting the Carotid in Transverse Scan

The enhancement artefact is used as the landmark to detect and segment the carotid artery. From the discussion in the previous Section, the SRR with the highest average intensity is assigned the label L. So pixels in $f_{\mathrm{Q}}(\mathbf{x})$ that have an intensity $I_{\mathrm{avg\_L}}$ belong to the enhancement artefact. Since the carotid is theoretically anechoic, the pixels belonging to the carotid have the lowest average intensity $I_{\mathrm{avg\_1}}$ in $f_{\mathrm{Q}}(\mathbf{x})$. A binary image $f_{\mathrm{bC}}(\mathbf{x})$ is formed with all the pixels having intensity $I_{\mathrm{avg\_1}}$, that lie above the pixels with intensity $I_{\mathrm{avg\_L}}$, as foreground pixels. Connected component labelling is then performed on $f_{\mathrm{bC}}(\mathbf{x})$ to determine the carotid candidates. Since the carotid appears as an ellipse, for every component which is a carotid candidate an ellipse is constructed using linear least squares method [184]. The carotid artery is detected and segmented from the list of candidates by the use of local phase based methods.

In Chapter 2 several methods were discussed to determine the local phase in an image. In this method, the scale invariant method of [61] is used to obtain the local phase image $\phi(\mathbf{x})$. Once the local phase is calculated, the phase congruency or feature asymmetry is obtained as:

$$\mathrm{FC}(\mathbf{x}) = \frac{\lfloor [|f_e(\mathbf{x})| - |f_o(\mathbf{x})|] - T_s \rfloor}{\sqrt{f_e(\mathbf{x})^2 + f_o(\mathbf{x})^2} + \epsilon}, \tag{4.22}$$

where $\lfloor . \rfloor$ denotes the flooring function to zero,

$$T_s = e^{\frac{1}{MN}\sum_{\mathbf{x}} \log(\sqrt{f_e(\mathbf{x})^2 + f_o(\mathbf{x})^2})}. \tag{4.23}$$

Here, $f_e(\mathbf{x})$ and $f_o(\mathbf{x})$ refer to the even and odd band-pass filters, the details of which are given in Section 2.4. The carotid candidate with the highest number of pixels having $\mathrm{FC}(\mathbf{x}) > 0.45$ within it is chosen as the most likely candidate for the carotid. The pixels of this candidate are retained as foreground pixels in $f_{\mathrm{bC}}(\mathbf{x})$, with the rest reset to background pixels. The selected candidate is indeed the carotid and segmented in the image $f(\mathbf{x})$ if:

$$\sum_{\mathbf{x}} (f_{\mathrm{bC}}(\mathbf{x}))(1 - \mathbb{H}(\phi(\mathbf{x}))) > 0 \tag{4.24}$$

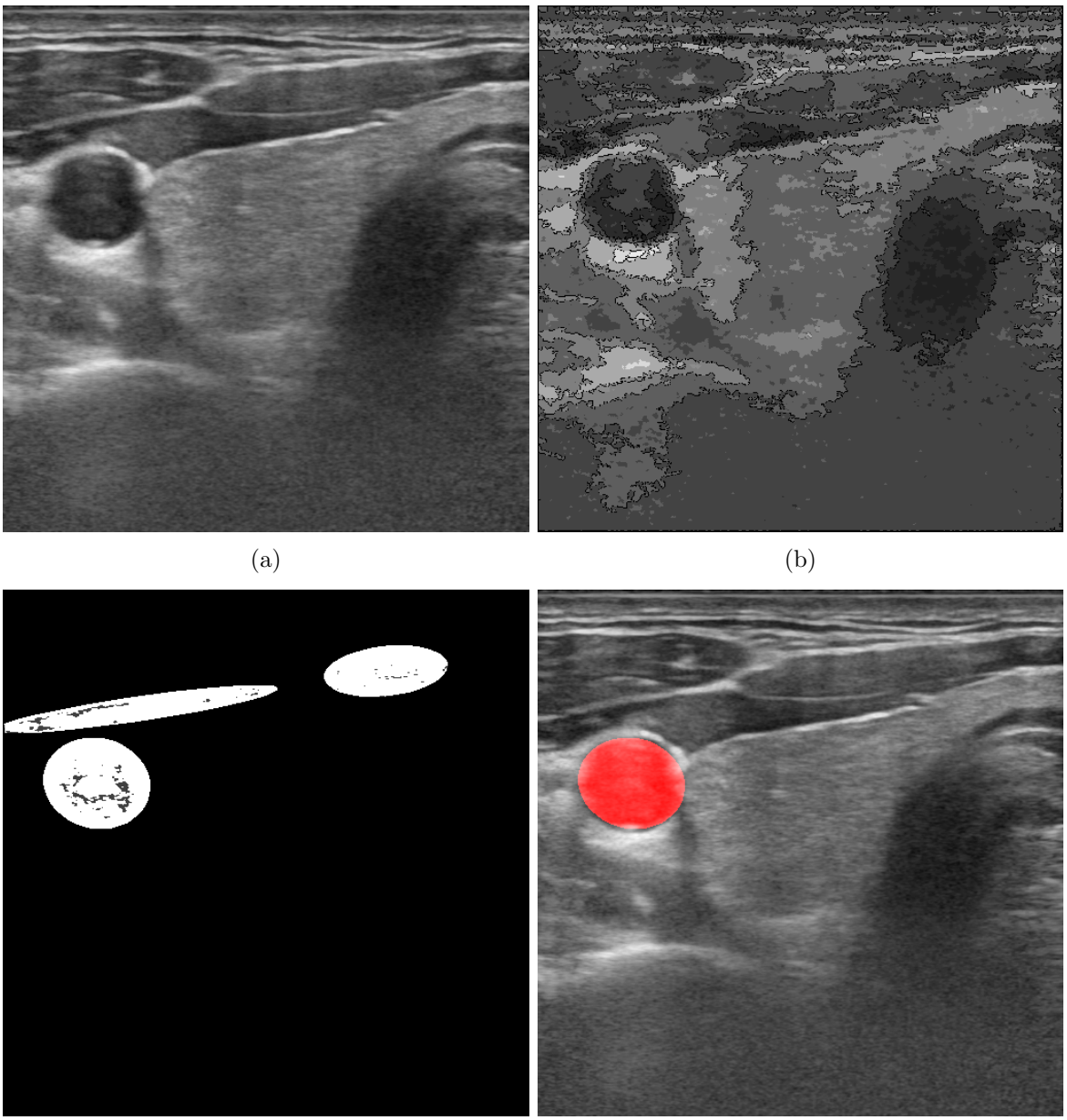
and

$$\mathrm{Ecc} = \sqrt{1 - \frac{b^2}{a^2}} > 0.75, \tag{4.25}$$

where, $\mathbb{H}(.)$ represents the Heaviside function and Ecc is the eccentricity of the candidate ellipse with major axis a and minor axis b. The constants of Eq. (4.22) and Eq. (4.23) are empirically chosen. Figure 4.13 shows the process of segmenting the carotid in transverse US scans of the thyroid gland.

4.4.3 Segmenting the remaining organs in Transverse Scans

Local phase based methods are employed once again to segment the rest of the organs in transverse US scans. Once the carotid artery is detected and segmented, the next



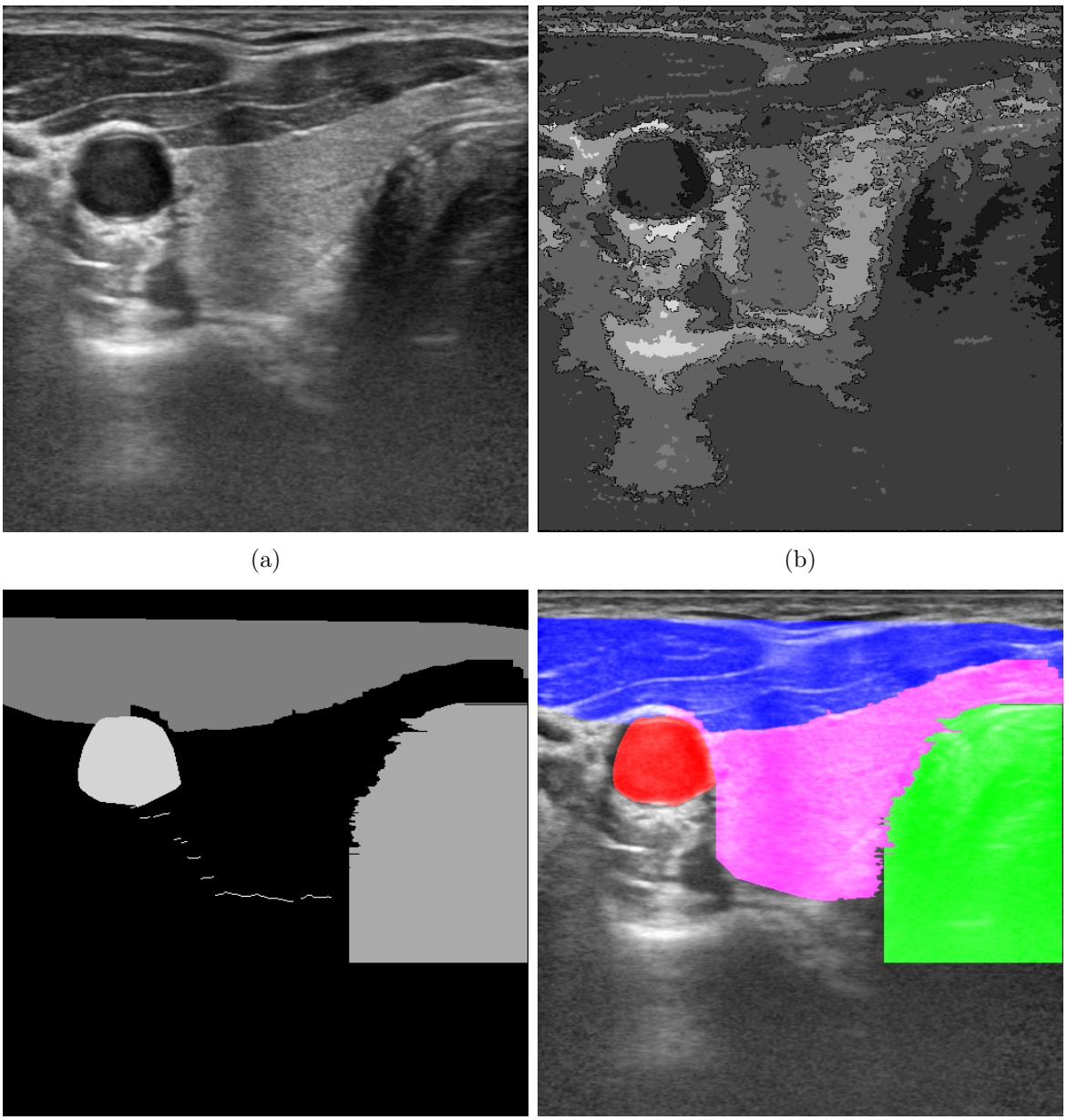
(a) (b)

(c) (d)

Figure 4.13: Sequence of images showing the segmentation of carotid in transverse US scans where: (a) is the input US image;(b) is the quantized image $f_Q(\mathbf{x})$; (c) is the image of the carotid candidates with pixels that have FC > 0.45 highlighted (in gray) and (d) is the segmented carotid (in red).

organ to be segmented is the trachea. The trachea is segmented by applying connected component analysis to the complement of the binary image $\phi(\mathbf{x})$ given by:

$$\phi_b(\mathbf{x}) = (1 - \mathbb{H}(\phi(\mathbf{x})). \tag{4.26}$$



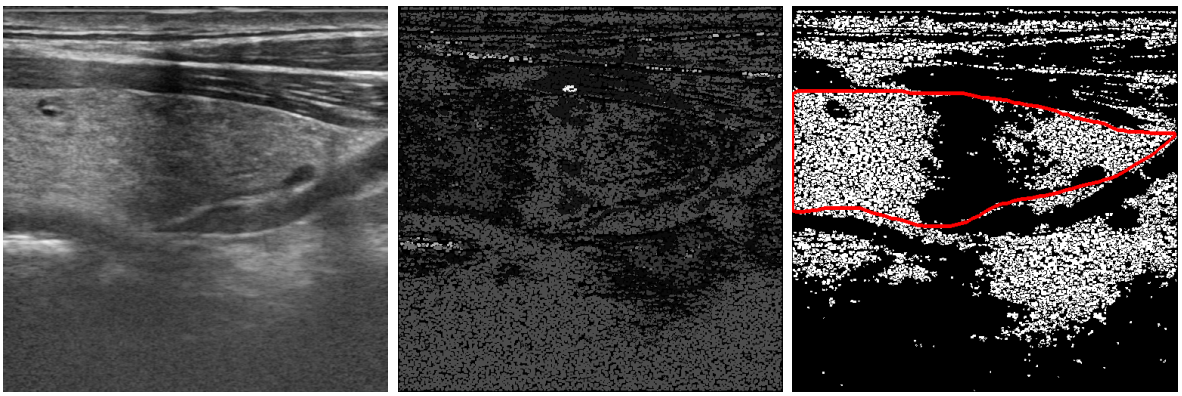
(a) (b)

(c) (d)

Figure 4.14: Estimating the lower boundary of the thyroid gland. (a) Input image in transverse scan; (b) Quantized image $f_Q(\mathbf{x})$; (c) Estimated lower boundary of thyroid gland and (d) Segmentation results with the thyroid in Magenta, the carotid in Red, muscles in Blue and the trachea in Green.

The largest and the deepest (towards the bottom of the image) component in $\phi_b(\mathbf{x})$ is segmented as the Trachea.

The carotid artery is used as the landmark to segment the muscles. The muscles are segmented by applying connected component analysis to the binary image $\phi_b(\mathbf{x})$ minus



(a) (b) (c)

Figure 4.15: Establishing search region to segment thyroid gland in longitudinal scans. (a) Input US image in longitudinal scan; (b) Quantized image $f_Q(\mathbf{x})$; and (c) Binary search region $f_{\text{bMask}}(\mathbf{x})$ with GT overlay (in red).

the pixels belonging to the segmented trachea. The largest component above the carotid is identified as the Omohyoid muscles. The components that lie within the bounding box around it belong to the strap muscles. The pixels of the components that belong to the Omohyoid and Strap muscles are together labelled as muscles in the segmented image.

Similar to the approach followed in Section 4.3.4, a search region is established in between the carotid artery, trachea and the muscles in a binary image $\bar{\phi}_b(\mathbf{x})$ given by:

$$\bar{\phi}_b(\mathbf{x}) = \lceil \mathbb{H}(\phi(\mathbf{x})) \rceil, \tag{4.27}$$

where, $\lceil . \rceil$ stands for the ceiling operator. Pixels that are outside the search region are set as background pixels in $\bar{\phi}_b(\mathbf{x})$. The edge pixels in $f(\mathbf{x})$ (determined by the Canny edge detector at a threshold of 0.4) that lie under a digital line given by [185]:

$$n = \lceil r \times m \rceil + c, \forall m \in [1, M] \text{and} |r| \leq 1, \quad \text{or} \tag{4.28}$$

$$m = \lceil \frac{n}{r} \rceil + c, \forall n \in [1, N] \text{and} |r| > 1, \tag{4.29}$$

where,

$$r = \frac{y_{\text{Cr}} - y_{\text{Tr}}}{x_{\text{Cr}} - x_{\text{Tr}}} \tag{4.30}$$

and

$$c = y_{\text{Cr}} - r x_{\text{Cr}} \text{ or } c = y_{\text{Tr}} - r x_{\text{Tr}}, \tag{4.31}$$

belong to the lower boundary of the thyroid gland. The digital line runs from the highest point in the trachea (close to the top of the image) to the centroid of the carotid artery. A spline is fit to the edge pixels that lie below the digital line of Eq. (4.29) to remove discontinuities in the lower boundary of the thyroid gland. All the foreground pixels in $\bar{\phi}_b(\mathbf{x})$ that lie below the lower boundary of the thyroid gland are reset to background

pixels. The foreground pixels that remain are segmented as the thyroid gland. Figure 4.14 shows the segmentation of the transverse US scans of thyroid gland along with the estimated lower boundary of the thyroid gland.

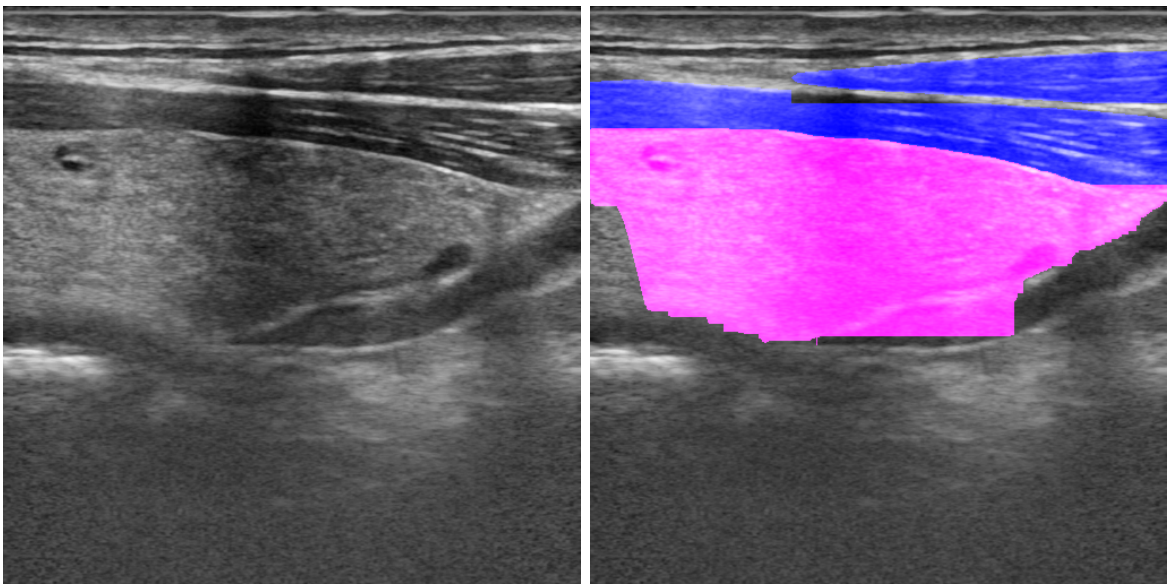
4.4.4 Segmenting the longitudinal scans

In longitudinal scans, the thyroid gland is the first organ to be detected and segmented. The segmented thyroid is then used as a landmark to segment the muscles. The thyroid gland is detected by applying intensity based thresholds to the quantized image $f_{\mathrm{Q}}(\mathbf{x})$. A search range is first established by forming a binary image $f_{\mathrm{bMask}}(\mathbf{x})$ whose foreground pixels are those that have an $I_{\mathrm{avg\_}m}$ in the range $[I1, I2]$, $I1 < I_{\mathrm{avg\_}m} < I2$, in $f_{\mathrm{Q}}(\mathbf{x})$. This range is empirically determined to be [50,200]. Figure 4.15 shows the search region of the thyroid gland in longitudinal scan. It can be seen from Figure 4.15(c) that some portions of the thyroid gland are missing due to the shadow artefact passing through the tissue. This is the drawback of using intensity based methods to segment US images. To suppress the effect of the shadow artefacts on the segmentation results, a local phase based method similar to the one in the previous section is employed to segment the thyroid gland in the search region.

To segment the thyroid gland in the search region, a binary image $\bar{\phi}_b(\mathbf{x})$ is obtained using Eq.(4.27). Connected component labelling is then applied on this image. The pixels that belong to the largest component in $\bar{\phi}_b(\mathbf{x})$ are retained as the foreground pixels. The component is that of the thyroid gland if:

$$\sum_{\mathbf{x}} f_{\mathrm{bMask}}(\mathbf{x})\bar{\phi}_b(\mathbf{x}) > 0. \tag{4.32}$$

The segmented thyroid gland is used as the landmark to segment the muscles. Accordingly, a search region is established above the segmented thyroid gland in the binary image $1 - f_{\mathrm{bMask}}(\mathbf{x})$. Connected component labelling is then performed on the binary image to get a labelled image. For each row in the labelled image, the histogram of class labels is obtained and the label at which the histogram peaks is stored in an accumulator array. The total number of pixels belonging to each of the labels in the accumulator array is then determined and the array is sorted in descending order. The top two labels in the sorted accumulator array belong to the Omohyoid and Strap muscles. The pixels that belong to these labels are labelled as muscles in the segmented image. Figure 4.16 shows the results of MODS in the longitudinal scan.



(a) (b)

Figure 4.16: Results of applying SPS based MODS algorithm on longitudinal scans. (a) Input; and (b) Segmented image (thyroid in purple and muscles in blue).

Similarity Metric	Carotid detection accuracy in transverse scans (Set 1 - Transverse images + Set 2)	Thyroid detection accuracy in longitudinal scans (Set 1 - Longitudinal images)
Pearson's distance	73.07%	100%
SSD metric	100%	100%

Table 4.1: Comparison between Pearson's distance and SSD similarity metrics in terms of the detection accuracy (in %) of (a) the carotid artery in transverse scans of Set 1 and Set 2; and (b) the thyroid gland in longitudinal scans of Set 1.

4.5 Experimental results

In this Section, the metrics for qualitative and quantitative analysis introduced in Chapter 3 are used to validate the performance of the MODS methods proposed in this Chapter.

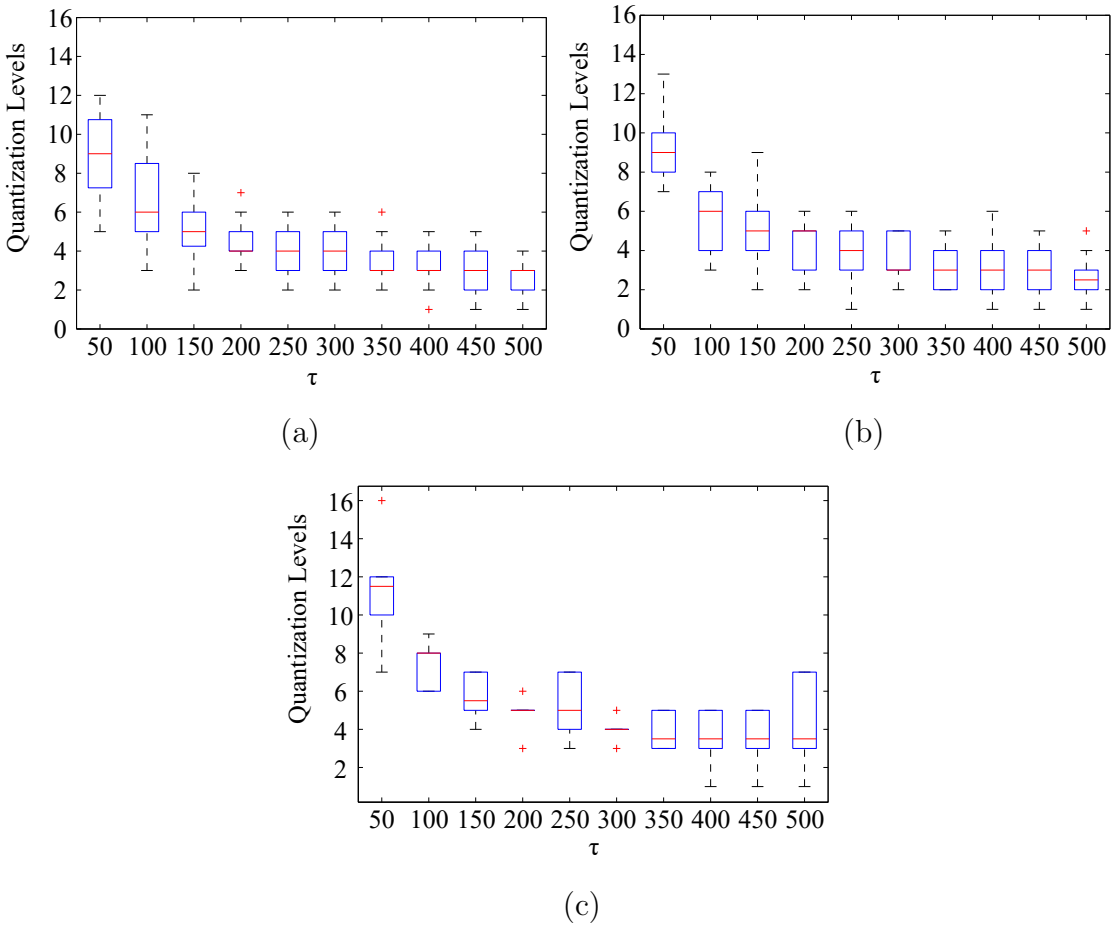


Figure 4.17: Variation of estimated quantization levels with changes in threshold τ. (a) Set 1: Transverse Scan; (b) Set 1: Longitudinal Scan; and (c) Set 2: Transverse Scan. From the plots it can be observed that the number of quantization levels (clusters formed) do not vary significantly For $\tau > 100$.

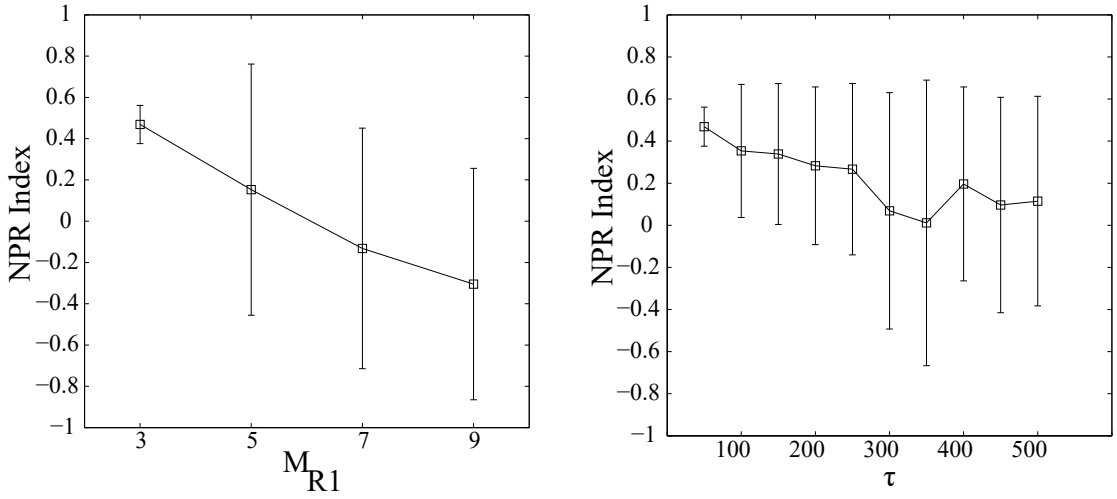
4.5.1 Parameter selection

4.5.1.1 Three Category MODS system

The three category MODS system was implemented by setting the window size $N_x \times N_y$ to be equal to 41×41. A threshold of $\tau_1 = 10$ was used to classify the speckle pixels into the three classes. The size of false positive candidate was set to $M_2 = 50$ while segmenting the carotid.

4.5.1.2 Multi-Category MODS system

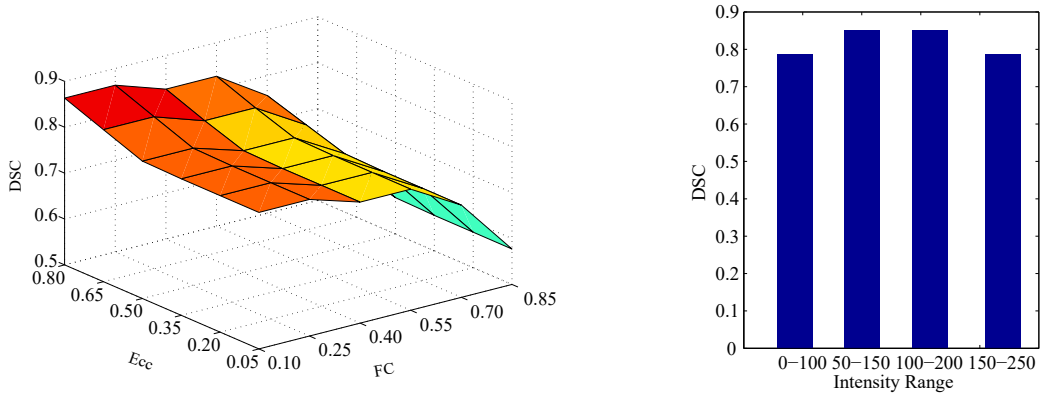
The SRR estimation method of Section 4.4.1 is an agglomerative clustering algorithm. The number of clusters formed converges to one (trivial solution) at large τ. In our experiments, it is found that this happens for $\tau > 10000$. Nevertheless, the number



(a) NPR vs M_{R1} ($\tau = 100$)

(b) NPR vs τ ($M_{R1} = 3$)

Figure 4.18: Performance analysis of the proposed quantization scheme with respect to: (a) changes in the ROI size M_{R1} when the threshold τ is held constant at 100 and (b) changes in the threshold τ when the ROI size M_{R1} is held constant at 3. The Normalized Probabilistic Rand (NPR) index has highest values at M_{R1}=3 and $\tau = 100$.



(a)

(b)

Figure 4.19: Sensitivity analysis for the constants used to segment: (a) the carotid artery. Here Ecc and FC refer to the eccentricity and phase congruency, respectively; and (b) the thyroid gland in transverse and longitudinal scans, respectively.

of SRR's formed remains stable at 5 ± 2 for τ between 100 and 400. Further, from the performance plots of Figure 4.18, it can be inferred the best performance of the SRR estimation algorithm is seen at $\tau = 100$ and $M_{R1} = 3$. Thus, all experiments

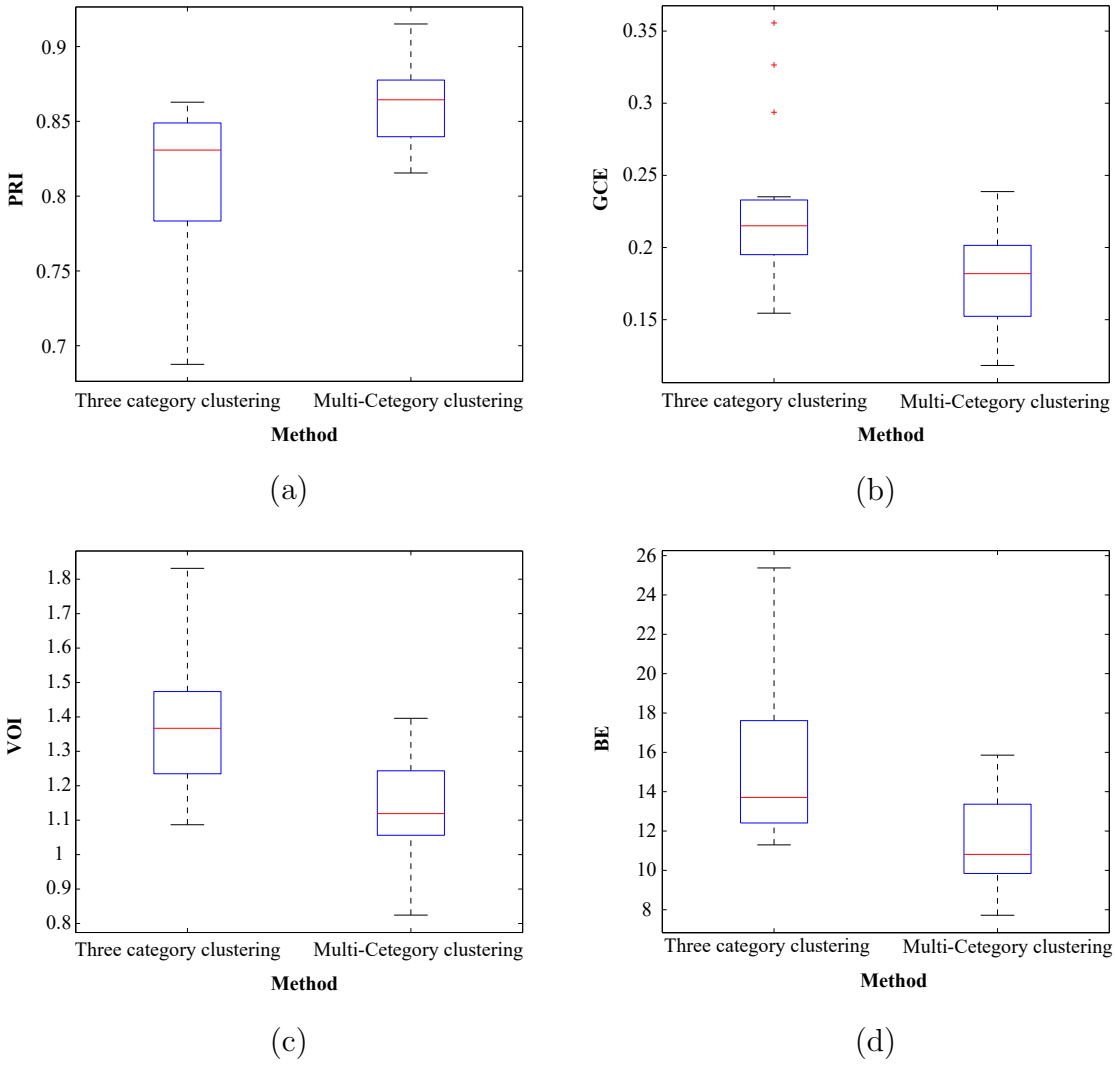


Figure 4.20: Qualitative analysis of clustering methods to segment all organs in transverse scan. Comparison of MODS methods for US images in the longitudinal scan using (a) Probabilistic Rand Index (PRI). (b) Global Consistency Error (GCE). (c) Variation of Information (VOI). and (d) Boundary Error (BE in pixels). It can be inferred from the plots that the multi-category clustering based segmentation produces better quality results with PRI close to 1, GCE close to 0, VOI and BE lower than that of the three category clustering based segmentation.

are conducted by setting τ and M_{R1} equal to 100 and 3, respectively. With regard to the choice of metric used for similarity, Table 4.1 shows a comparative account of the detection accuracies when Pearson's metric and SSD metric are used to estimate the number of SRR's. It can be seen that the Pearson's distance helps to positively identify the carotid artery only in 73% of the cases in the transverse scans while it can positively

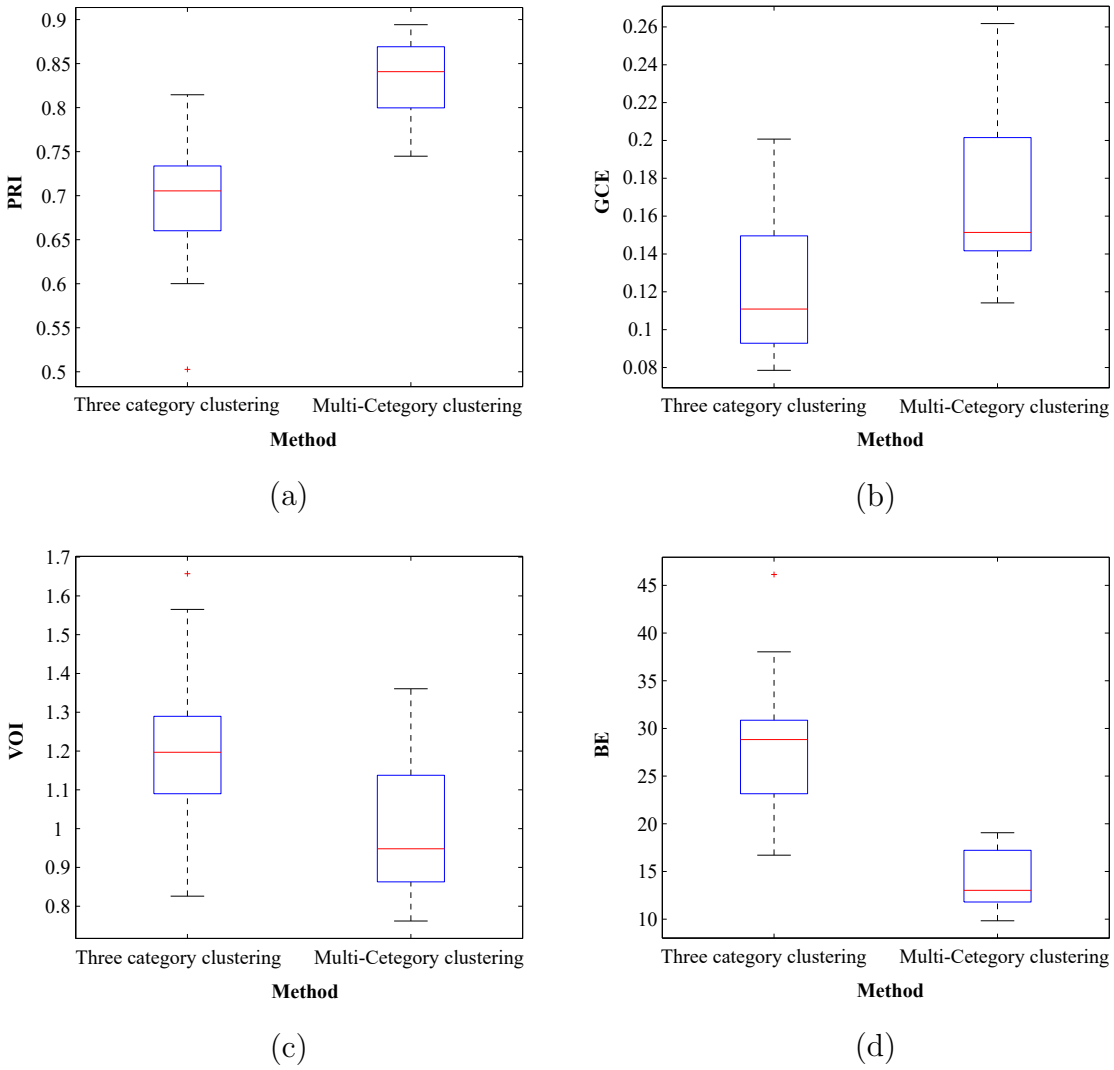


Figure 4.21: Qualitative analysis of clustering methods to segment all organs in longitudinal scan. (a) Probabilistic Rand Index (PRI). (b) Global Consistency Error (GCE). (c) Variation of Information (VOI). and (d) Boundary Error (BE in pixels). It can be inferred from the plots that the multi-category clustering based segmentation produces better quality results with PRI close to 1, GCE close to 0, VOI and BE lower than that of the three category clustering based segmentation.

help to detect the thyroid gland in 100% of the cases in the longitudinal scan. The SSD metric on the other can detect to 100% both the carotid artery and the thyroid gland in the transverse and longitudinal scans, respectively. Thus, SSD metric is the metric of choice in our multi-category MODS framework.

A sensitivity analysis is undertaken to determine the values of the three constants: (a) phase congruency (FC); (b) the eccentricity of the ellipse (Ecc); and (c) the intensity

Method	Dataset	Scan	**SE**	**SP**	**DSC**	**PPV**
Multi-category	S1	T	0.977 ± 0.010	0.855 ± 0.078	**0.854 ± 0.040**	0.862 ±0.063
	S1	L	0.930 ± 0.044	0.940 ± 0.034	**0.853 ± 0.069**	0.790 ± 0.114
	S2	T	0.958 ± 0.018	0.873 ± 0.083	**0.810 ± 0.031**	0.767 ± 0.081
	SS	T	0.970 ± 0.013	0.864 ± 0.083	**0.840 ± 0.040**	0.827 ± 0.069
	SS	L	0.941 ± 0.043	0.936 ± 0.035	**0.866 ± 0.072**	0.818 ± 0.116
Three category	S1	T	0.977 ± 0.017	0.819 ± 0.090	0.822 ±0.048	0.842 ±0.101
	S1	L	0.933 ± 0.075	0.857 ± 0.107	0.811 ± 0.096	0.805 ± 0.162
	S2	T	0.863 ± 0.122	0.636 ± 0.303	0.487 ± 0.270	0.460 ± 0.339
	SS	T	0.955 ± 0.056	0.740 ± 0.270	0.702 ± 0.246	0.705 ± 0.280
	SS	L	0.966 ± 0.015	0.844 ± 0.095	0.850 ± 0.051	0.867 ± 0.060
JCR [170]	SS	T	0.679 ± 0.102	0.922 ± 0.077	0.490 ±0.067	0.337 ±0.060
	SS	L	0.449 ± 0.038	0.930 ± 0.062	0.468 ± 0.064	0.315 ±0.053
Chang [49]	SS	T	0.891 ± 0.106	0.662 ± 0.258	0.589 ± 0.199	0.598 ± 0.263
	SS	L	0.856 ± 0.127	0.457 ± 0.389	0.435 ± 0.376	0.464 ± 0.431
Garg [133]	SS	T	0.524 ± 0.157	0.955 ± 0.077	0.411 ± 0.074	0.264 ± 0.060
	SS	L	0.421 ± 0.206	0.772 ± 0.370	0.388 ± 0.212	0.266 ± 0.161

⋆ Results are averaged over both experts.
† Methods whose names are italicized represent multi-organ segmentation algorithms. Unsupervised algorithms are listed first followed by supervised algorithms for every organ. The results of each algorithm are sorted according to the dataset used from S1 to SS.
T = Transverse. L = Longitudinal. S1 = Set 1. S2 = Set 2. SS = Supervised Set. SE = Sensitivity. SP = Specificity. DSC = Dice Co-efficient. PPV = Positive Predictive Value.
Best results are highlighted in **Bold** font.

Table 4.2: Performance analysis of the thyroid segmentation algorithms. It can be inferred from the Dice Co-efficient values that the Multi-Categoty clustering based segmentation algorithm outperforms the remaining algorithms ⋆ †.

range [I1, I2]. From the plots in Fig. 4.19, it can be inferred that the algorithm is robust to changes in Ecc and changes in FC up to 0.55. The algorithm performs at its best at low values of FC and high values of Ecc. Accordingly, the values of FC and ECC are set to 0.45 and 0.77, respectively, in our experiments. For the thyroid gland in longitudinal scans, the best performance is seen when the intensity is chosen in the range [50,200].

4.5.2 Validation

The datasets and the metrics introduced in Chapter 3 are used to validate the MODS methods proposed in this chapter. The Joint Classification Regression (JCR) algorithm proposed by Glocker et al. [170] to perform multi-organ segmentation in CT images of the abdomen method is modified for use on US images so that it can be compared with

Method	Dataset	Scan	**SE**	**SP**	**DSC**	**PPV**
Multi-Category	S1	T	0.998 ± 0.002	0.886 ± 0.077	**0.887 ± 0.037**	0.897 ± 0.065
	S2	T	0.999 ± 0.001	0.836 ± 0.072	**0.881 ± 0.031**	0.939 ±0.054
	SS	T	0.997 ± 0.002	0.908 ± 0.056	**0.892 ± 0.031**	0.882 ±0.063
Three Category	S1	T	0.995 ± 0.005	0.845 ±0.183	0.809 ±0.151	0.805 ±0.181
	S2	T	0.986 ± 0.011	0.783 ±0.355	0.680 ±0.311	0.620 ±0.307
	SS	T	0.992 ± 0.008	0.841 ±0.286	0.751 ±0.261	0.695 ±0.266
Star [134]	S1	T	0.935 ± 0.089	0.693 ±0.338	0.404 ±0.239	0.531 ±0.361
	S2	T	0.875 ± 0.104	0.813 ±0.350	0.303 ±0.236	0.371 ±0.334
Star Kalman [135]	S1	T	0.999 ± 0.001	0.576 ±0.212	0.700 ±0.154	0.982 ±0.044
	S2	T	0.999 ± 0.001	0.566 ±0.197	0.702 ±0.166	0.998 ±0.003
Star Extended Kalman [136]	S1	T	0.987 ± 0.032	0.932 ±0.099	0.827 ±0.167	0.793 ±0.177
	S2	T	0.982 ± 0.018	0.930 ±0.143	0.757 ±0.125	0.692 ±0.212
Spoke Ellipse [137]	S1	T	0.982 ± 0.030	0.935 ±0.090	0.742 ±0.206	0.667 ±0.244
	S2	T	0.990 ± 0.008	0.962 ±0.036	0.839 ±0.073	0.755 ±0.120
JCR [170]	SS	T	0.965 ± 0.040	0.748 ±0.204	0.503 ±0.191	0.417 ±0.219

⋆ Results are averaged over both experts.

† Methods whose names are italicized represent multi-organ segmentation algorithms. Unsupervised algorithms are listed first followed by supervised algorithms for every organ. The results of each algorithm are sorted according to the dataset used from S1 to SS.
T = Transverse. L = Longitudinal. S1 = Set 1. S2 = Set 2. SS = Supervised Set. SE = Sensitivity. SP = Specificity. DSC = Dice Co-efficient. PPV = Positive Predictive Value.
Best results are highlighted in **Bold** font.

Table 4.3: Performance analysis of the carotid segmentation algorithms. It can be inferred from the Dice Co-efficient values that the Multi-Categoty clustering based segmentation algorithm outperforms the remaining algorithms ⋆ †.

our methods. We use the adaptive weighted median filter of [49] is with a window size of 5×5 to filter the image for noise removal in our implementation of [170]. Five variants of the box features are extracted from the filtered image. The box sizes vary between 5mm and 10mm, and displacements of the boxes are drawn from a [0,10mm] interval. In all, 369 features are extracted per pixel in every organ from the training images in the dataset SS. A total of 50 trees are trained by the method of Bagging. Each tree is trained on a random subset of images that corresponds to 10% of the number of training samples. From the pool of 369 features, 40 features are evaluated at each split node and for each feature 10 different thresholds at equal intervals in the range of feature responses is employed to evaluate the information gain. Signed distance maps are used to measure the pairwise distance between every pixel in one organ to the nearest pixel of another organ. In addition to the comparison with the JCR algorithm, the segmentation results of the carotid artery and thyroid gland are compared with the existing methods for the

Method	Dataset	Scan	**SE**	**SP**	**DSC**	**PPV**
Multi-Category	S1	T	0.975 ± 0.023	0.856 ± 0.130	**0.838± 0.074**	0.850 ± 0.120
	S1	L	0.980 ± 0.026	0.689 ± 0.237	0.712 ± 0.203	0.823 ± 0.246
	S2	T	0.984 ± 0.010	0.729 ± 0.151	**0.786 ± 0.100**	0.878 ± 0.083
	SS	T	0.978 ± 0.021	0.806 ± 0.135	**0.817 ± 0.075**	0.859 ± 0.118
	SS	L	0.993 ± 0.008	0.819 ± 0.103	**0.878 ± 0.049**	0.960 ± 0.040
Three catrgory	S1	T	0.980 ± 0.017	0.737 ± 0.164	0.780 ± 0.116	0.862 ± 0.116
	S1	L	0.959 ± 0.038	0.859 ± 0.120	**0.788 ± 0.106**	0.780 ± 0.201
	S2	T	0.983 ± 0.016	0.601 ± 0.254	0.676 ± 0.195	0.855 ± 0.117
	SS	T	0.977 ± 0.017	0.676 ± 0.202	0.732 ± 0.148	0.841 ± 0.114
	SS	L	0.978 ± 0.025	0.839 ± 0.114	0.853 ± 0.051	0.894 ± 0.106
JCR [170]	SS	T	0.917 ± 0.035	0.624 ± 0.258	0.556 ± 0.160	0.551 ± 0.145
	SS	L	0.885 ± 0.043	0.940 ± 0.031	0.745 ± 0.108	0.632 ± 0.148

⋆ Results are averaged over both experts.
† Methods whose names are italicized represent multi-organ segmentation algorithms. Unsupervised algorithms are listed first followed by supervised algorithms for every organ. The results of each algorithm are sorted according to the dataset used from S1 to SS.
T = Transverse. L = Longitudinal. S1 = Set 1. S2 = Set 2. SS = Supervised Set. SE = Sensitivity. SP = Specificity. DSC = Dice Co-efficient. PPV = Positive Predictive Value.
Best results are highlighted in **Bold** font.

Table 4.4: Performance analysis of the algorithms to segment the muscles. It can be inferred from the DSC values that other than that of longitudinal scans of Set 1, the Multi-category clustering based segmentation algorithm performs better than other algorithms to segment the muscles ⋆ †.

respective organs discussed in Chapter 2. The results of the comparison are summarized in Tables 4.2, 4.3, 4.4 and 4.5.

It can be inferred from the DSC values in Tables 4.2-4.5, that the Multi-Category MODS system method performs better than the existing methods used to segment both individual and multiple organs. While the Multi-Category segmentation algorithm has high Sensitivity and Specificity values for all the datasets and for all the organs, the same cannot be said of the three category segmentation algorithm. The three category segmentation algorithm has high sensitivity and specificity values only for organs in images of dataset S1. Still, both perform better than the JCR method of [170]. Low sensitivity and high specificity values by the thyroid segmentation algorithms of Garg et al. [133] and JCR suggest over segmentation of the thyroid gland with the pixels not belonging to the thyroid gland falsely labelled as thyroid gland. On the other hand the method of Chang et al. [49] has high sensitivity and low specificity. This suggests that a lot of pixels belonging to the thyroid gland were incorrectly identified as background pixels in the segmented image. With regard to segmenting the carotid artery, except for the extended Kalman algorithm of [186], the remaining do not perform as well as our

Method	Dataset	Scan	**SE**	**SP**	**DSC**	**PPV**
Multi-Category	S1	T	0.977 ± 0.018	0.880 ± 0.092	$\mathbf{0.850 \pm 0.053}$	0.835 ± 0.096
	S2	T	0.956 ± 0.040	0.900 ± 0.078	$\mathbf{0.832 \pm 0.085}$	0.804 ± 0.156
	SS	T	0.972 ± 0.026	0.907 ± 0.080	$\mathbf{0.856 \pm 0.057}$	0.826 ± 0.110
Three category	S1	T	0.963 ± 0.031	0.876 ± 0.122	0.780 ± 0.110	0.744 ± 0.177
	S2	T	0.977 ± 0.031	0.291 ± 0.376	0.249 ± 0.298	0.555 ± 0.294
	SS	T	0.955 ± 0.034	0.767 ± 0.350	0.607 ± 0.296	0.607 ± 0.232
JCR [170]	SS	T	0.862 ± 0.036	0.804 ± 0.146	0.526 ± 0.131	0.407 ± 0.132

⋆ Results are averaged over both experts.
† Methods whose names are italicized represent multi-organ segmentation algorithms. Unsupervised algorithms are listed first followed by supervised algorithms for every organ. The results of each algorithm are sorted according to the dataset used from S1 to SS.
T = Transverse. L = Longitudinal. S1 = Set 1. S2 = Set 2. SS = Supervised Set. SE = Sensitivity. SP = Specificity. DSC = Dice Co-efficient. PPV = Positive Predictive Value.
Best results are highlighted in **Bold** font.

Table 4.5: Performance analysis of the algorithms to segment the trachea. From the DSC values, it can be inferred that the Multi-category clustering based segmentation algorithm performs better than other algorithms to segment the trachea ⋆ †.

methods.

From the plots of Figures 4.20 and 4.21, it can be seen that the Multi-category MODS system performs better than the Three Category MODS segmentation algorithm. A full comparison with the MODS systems is carried out in the next Chapter.

4.6 Summary

- An in depth review of the state-of-the-art multi-organ segmentation algorithms was undertaken in this chapter.
- Two new unsupervised methods were proposed for MODS in the ultrasound images of the thyroid gland.
- The first of the two methods is called the Three Category MODS system. In this method, the speckle related pixels are clustered into three echogenic classes based on its relative brightness in a kernel. An energy based model is then used to detect the carotid artery in the labelled image (image whose pixels are labelled with the cluster number). The carotid artery is used as a landmark to detect and segment the remaining organs in the image.
- The second method is called the Multi-Category MODS system. In this method, an agglomerative clustering constrained with a similarity metric is proposed to cluster

the speckle related pixels into an unknown number of echogenic classes. The total number of class labels generated at the end of the clustering determines the number of echogenic classes in the image. This process is viewed as a quantization of the speckle related pixels into levels that are determined by the tissue echogenicity. The enhancement artefact is then detected in the quantized image which is in turn used as a landmark to detect the carotid artery. The carotid artery is used as a landmark to detect the remaining organs. All organs are segmented by the application of region based local phase methods.

- From the results of the qualitative analysis, it can be inferred that the Multi-Category MODS system performs better than the three category MODS system.
- The results of the quantitative analysis proves that the Multi-Category MODS algorithm outperforms the existing methods proposed to perform both individual and multi-category segmentation in ultrasound images of the thyroid gland.

Chapter 5

Self-learning for MODS in US images

Self-learning-based approaches for computer vision problems are relatively new in the community and have recently found applications to image denoising [187] and image/video deblocking [188]. In a first of its kind attempt, a self-learning based approach is proposed in this Chapter to perform landmark based multi-organ segmentation in US images of the thyroid gland. In this method, an initial unsupervised algorithm drives the subsequent supervised algorithm that is trained on the results of the unsupervised algorithm. This is where our method differs from semi-supervised methods. Semi-supervised methods use unsupervised algorithms to restrict the search region in the image and then employ supervised algorithms trained on external annotated datasets to perform segmentation. We start with the premise that the input image is unlabelled and use local-phase-based methods to detect and segment hypo-echoic anatomical structures (carotid, trachea and muscles) in an unsupervised manner. The relative position of the segmented hypo-echoic structures with respect to the hyper-echoic anatomical structures (subcutaneous fat, thyroid, artefacts and soft tissue) are used as priors to automatically generate the training samples to train a Random Forests classifier and classify the unlabelled pixels. Figure 5.1 shows the self-learning MODS framework.

5.1 Detecting Hypo-Echoic regions

The first step of this method is to detect the hypoechoic regions in the image. This is done by using local phase based methods. Local phase, $\phi(\mathbf{x})$, and local energy, E, are estimated using the methods outlined in Section 2.4. The scale invariant method of [61] is used in this method to estimate the local phase and energy in the input US image. Once the local features are calculated, the binary image $f_{\text{hypo}}(\mathbf{x})$ containing the

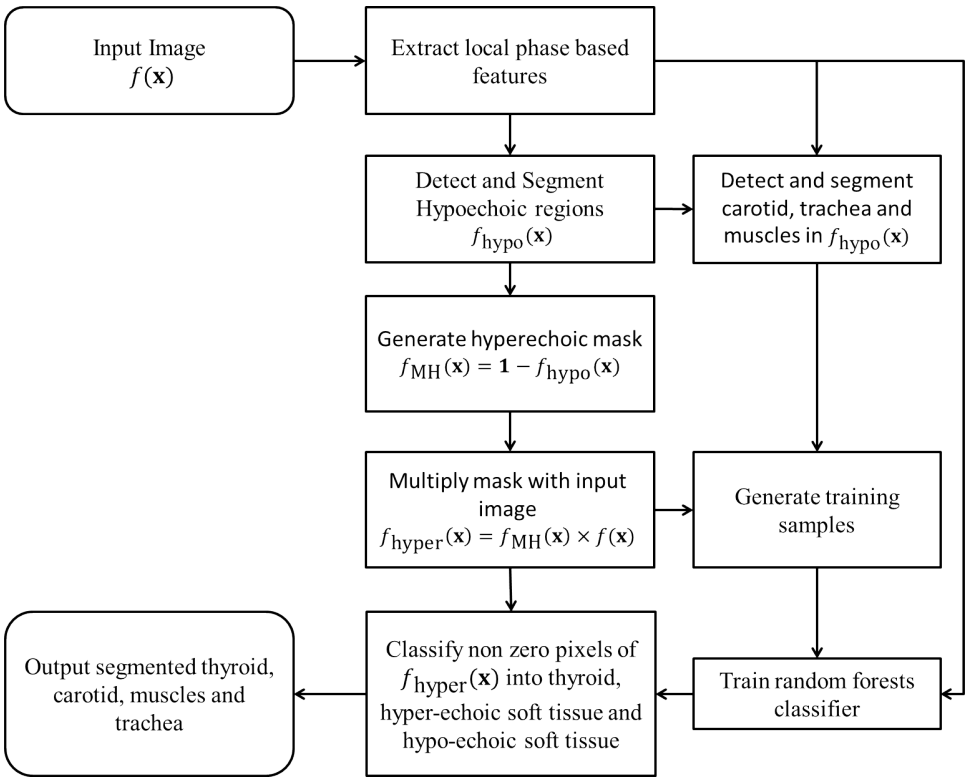


Figure 5.1: Flow diagram of the Self-learning MODS system.

hypoechoic regions is formed by Eq. (5.1):

$$f_{\text{hypo}}(\mathbf{x}) = \begin{cases} 1 & \text{if } \phi(\mathbf{x}) < 0 \\ \\ 0 & \text{Otherwise} \end{cases} . \tag{5.1}$$

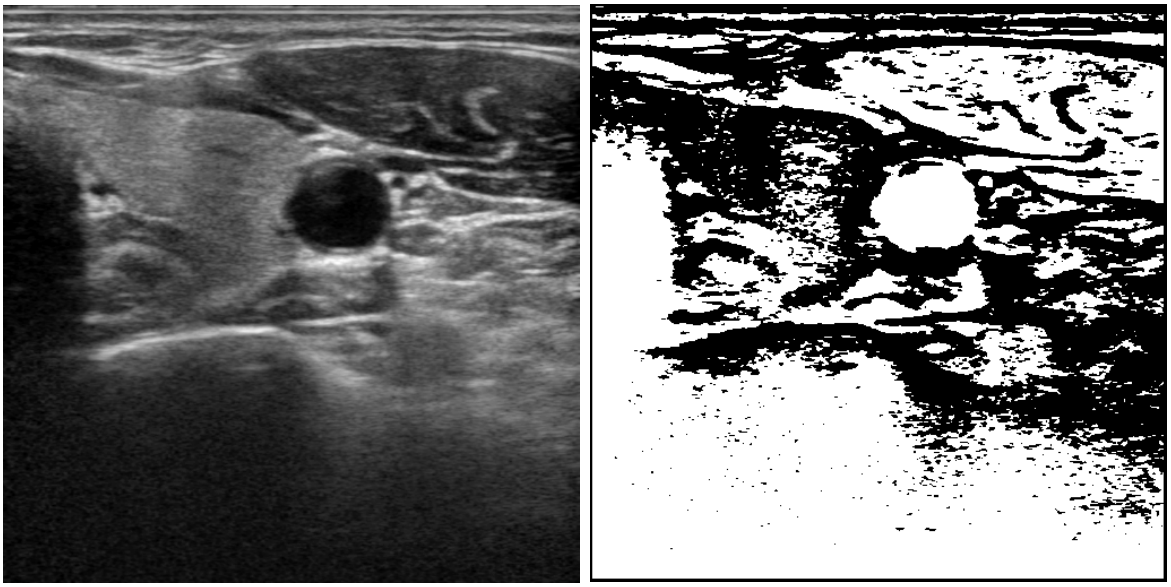
The complement of $f_{\text{hypo}}(\mathbf{x})$ represents the binary mask for the hyper-echoic regions $f_{\text{MH}}(\mathbf{x})$:

$$f_{\text{MH}}(\mathbf{x}) = 1 - f_{\text{hypo}}(\mathbf{x}) \tag{5.2}$$

Figure 5.2 shows an input image of a thyroid gland in the transverse scan and its corresponding hypo-echoic binary image $f_{\text{hypo}}(\mathbf{x})$.

5.2 Segmenting Hypo-Echoic regions

The hypoechoic anatomical landmarks of significance in the transverse scan are the trachea, the carotid and the strap muscles.



(a) Input (b) $f_{\text{hypo}}(\mathbf{x})$

Figure 5.2: (a)Ultrasound image of the left lobe of a thyroid gland in the transverse scan and (b) its corresponding binary hypo-echoic region.

5.2.1 Segmenting the trachea

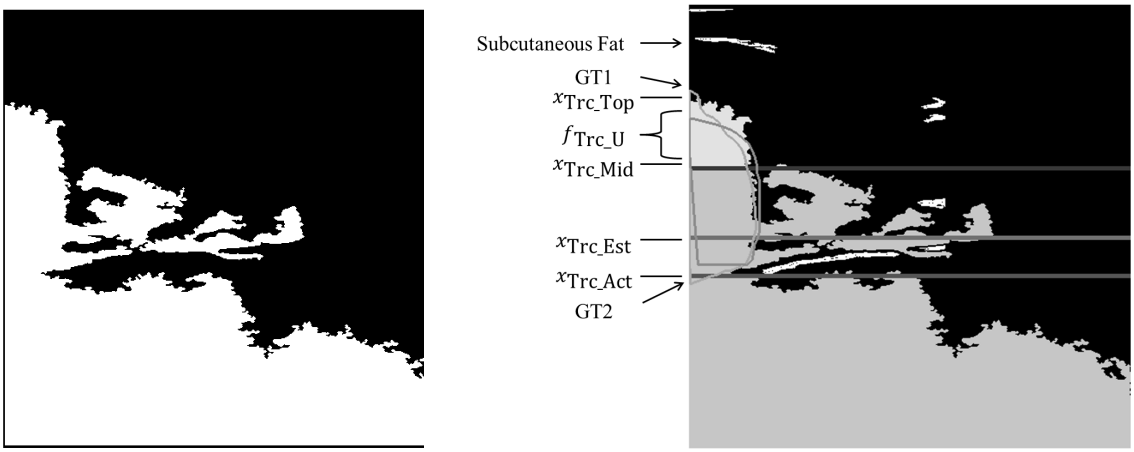
The trachea is found in between the two lobes of the thyroid gland under the isthmus. It is cylindrical in shape and runs through the length of the throat. In 2D US images of thyroid gland, only the top of the cross-section is visible as a semi-circle when the probe is placed directly over the isthmus. But, during routine scans of the thyroid gland, one lobe is scanned at a time, thus making it possible to visualize only a portion of the top tracheal surface. The methods to (a) detect the top tracheal boundary; and (b) estimate its anterior boundary, are discussed in this section on segmenting the trachea in 2D US images of the thyroid gland. The steps to be followed to segment the trachea are outlined in Algorithm 3 and are described in detail in the lines that follow.

Detection: Trachea is detected and segmented by subjecting $f_{\text{hypo}}(\mathbf{x})$ to morphological erosion and connected component analysis. The largest hypo-echoic component that is close to the bottom of the image is labelled as the tracheal candidate $f_{\text{trc\_cand}}(\mathbf{x})$. Best results for the tracheal candidate are observed when experiments are performed with the disk type structural element of radius 2 pixels for morphological erosion and 8 connectivity for connected component analysis. Figure 5.3a shows the detected tracheal candidate.

Segmentation: US imaging of the thyroid gland is usually performed at centre frequencies around 10Mhz. At this frequency of operation, the lower boundary of the trachea is not visible in the images and has to be approximated by making use of anatomical landmarks such as the soft tissue layer in between the thyroid gland and the Omohyoid muscles. The cross-section of trachea being circular and with only the top boundary of this cross-

Algorithm 3 Segmenting the trachea

1: **initialize** Choose the binary image $f_{\text{hypo}}(\mathbf{x})$ which contains the hypo-echoic regions as its foreground
2: Select the largest hypo-echoic component attached to the bottom of the image as the tracheal candidate $f_{\text{Trc\_Cand}}(\mathbf{x})$
3: Determine $x_{\text{Trc\_Est}}$ which denotes the approximate depth of the anterior boundary of trachea
4: Determine the axis of symmetry $x_{\text{Trc\_Mid}}$
5: Flip a copy of the upper half of the trachea about the axis of symmetry to segment the trachea
6: Refine the segmentation results by stretching the segmented trachea up to its true depth $x_{\text{Trc\_Act}}$



(a) (b)

Figure 5.3: Segmenting the trachea. (a). The Tracheal candidate: $f_{\text{trc\_cand}}$. (b) Estimating the lower boundary of trachea (GT1 = Ground Truth by Expert 1; GT2 = Ground Truth by Expert 2;$x_{\text{trc\_top}}$ = Top non-zero row in $f_{\text{trc\_cand}}$; $f_{\text{trc\_cand}}$ = Upper half of trachea; $x_{\text{trc\_mid}}$ = axis of symmetry; $x_{\text{trc\_est}}$ = Approximate row index of lower boundary of trachea; $x_{\text{trc\_tct}}$ = Actual row index of lower boundary of the trachea)

section visible as a semi-circle, it is necessary to determine the axis-of symmetry about which a copy of the top surface can be flipped so that the lower half of the trachea can be estimated.

The approximate depth of the lower boundary $x_{\text{trc\_est}}$ is determined to be the row $x \in [1, m]$ in $f_{\text{trc\_cand}}(\mathbf{x})$ at which the foreground pixels occupy more than half the number of columns n of the image in that row. The axis of symmetry is then determined to be the arithmetic mean of $x_{\text{trc\_est}}$ and the index of the first non-zero row $x_{\text{trc\_top}}$ in $f_{\text{trc\_cand}}(\mathbf{x})$.All the pixels that lie in between $x_{\text{trc\_mid}}$ and $x_{\text{trc\_est}}$ in $f_{\text{trc\_cand}}(\mathbf{x})$ belongs to the upper part

of the trachea $f_{\text{trc\_u}}(\mathbf{x})$. An initial segmentation of the trachea is obtained by flipping a copy of $f_{\text{trc\_u}}(\mathbf{x})$ about $x_{\text{trc\_mid}}$ so that the segmented trachea $f_{\text{trc}}(\mathbf{x})$ spans between $x_{\text{trc\_top}}$ and $x_{\text{trc\_est}}$.

This is not the final segmentation as a closer look at the image in figure 5.3b, shows that the actual depth of the trachea $x_{\text{trc\_act}}$ is much lower than $x_{\text{trc\_est}}$. From Fig. 5.3(b), it can be seen that the lowermost point of the trachea (where the contours of the Ground Truth segmentations provided by two expert doctors (GT1 and GT2) approximately meet) is at the same depth as that of the soft tissue that separates the thyroid gland from the Omohyoid muscles. The segmentation results are refined after the thyroid gland is segmented by the method described in Section 5.3.

5.2.2 Segmenting the carotid

The carotid artery appears hypo-echoic in the US image of the thyroid gland due to the presence of blood in it. The inherent characteristic of fluid filled tissues is that they contain very few scatterers. This means that the speckle signature of fluid filled structures such as the carotid artery or cysts is very less. Thus the carotid artery is the component in $f_{\text{hypo}}(\mathbf{x})$, that has the lowest number of speckle related pixels in $f(\mathbf{x})$ and the highest average local energy, E. The local energy is computed using the method described in Section 2.4. The steps involved in segmenting the carotid artery are outlined in Algorithm 4.

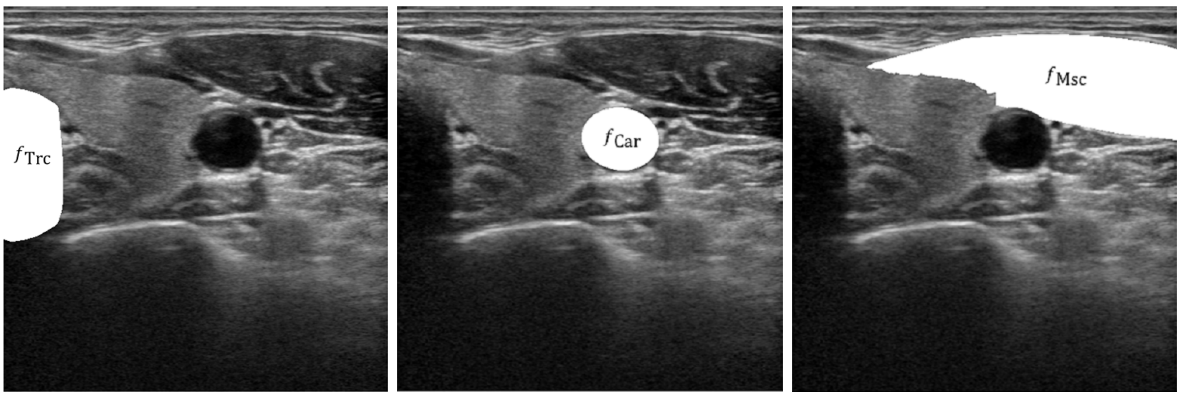
Assuming that the speckles in US images appear to be like blobs, the pixels related to speckles are obtained by applying the Hessian of Gaussian method [166]. The eigenvalues of the Hessian matrix obtained in an image patch that is filtered by a Gaussian kernel indicate the curvature of the blob-like structure in the direction denoted by its eigenvector.

Let λ_1 and λ_2 denote the eigenvalues of the Hessian Matrix:

$$H(f) = \begin{pmatrix} \frac{\partial^2 f(\mathbf{x})}{\partial^2 x} & \frac{\partial^2 f(\mathbf{x})}{\partial x \partial y} \\ \frac{\partial^2 f(\mathbf{x})}{\partial y \partial x} & \frac{\partial^2 f(\mathbf{x})}{\partial^2 y} \end{pmatrix}. \tag{5.3}$$

A pixel is said to belong to a speckle if both $\lambda_1 < 0$ and $\lambda_2 < 0$. The scale of Gaussian kernel plays an important role in the detection of speckle related pixels in the image. The higher the scale, the lower is the detection of significant speckle related pixels. Significant speckle related pixels refer to a group of pixels that actually resemble a blob. With high scales, the possibility of obtaining a large number of unconnected pixels as speckles is high. Scale values in the range 4-10 were found to provide optimal results in the proposed method. It is this step that helps in positively detecting the carotid in the presence of cysts or other disorders in the thyroid gland.

Let the binary image $f_{\text{hypo}}(\mathbf{x})$ minus the largest tracheal component be denoted by $f_{\text{car}}(\mathbf{x})$. All the pixels that do not belong to the carotid component are reset to background in $f_{\text{car}}(\mathbf{x})$. To segment the carotid artery, an edge detector (Canny's with 0.4 threshold) is first applied to the input image $f(\mathbf{x})$ to obtain an edge map $f_{\text{edge}}(\mathbf{x})$. For



(a) (b) (c)

Figure 5.4: Segmented (a) trachea $f_{\text{trc}}(\mathbf{x})$, (b) carotid $f_{\text{car}}(\mathbf{x})$ and (b) muscles $f_{\text{musc}}(\mathbf{x})$ in the input US image of Figure 5.2(a).

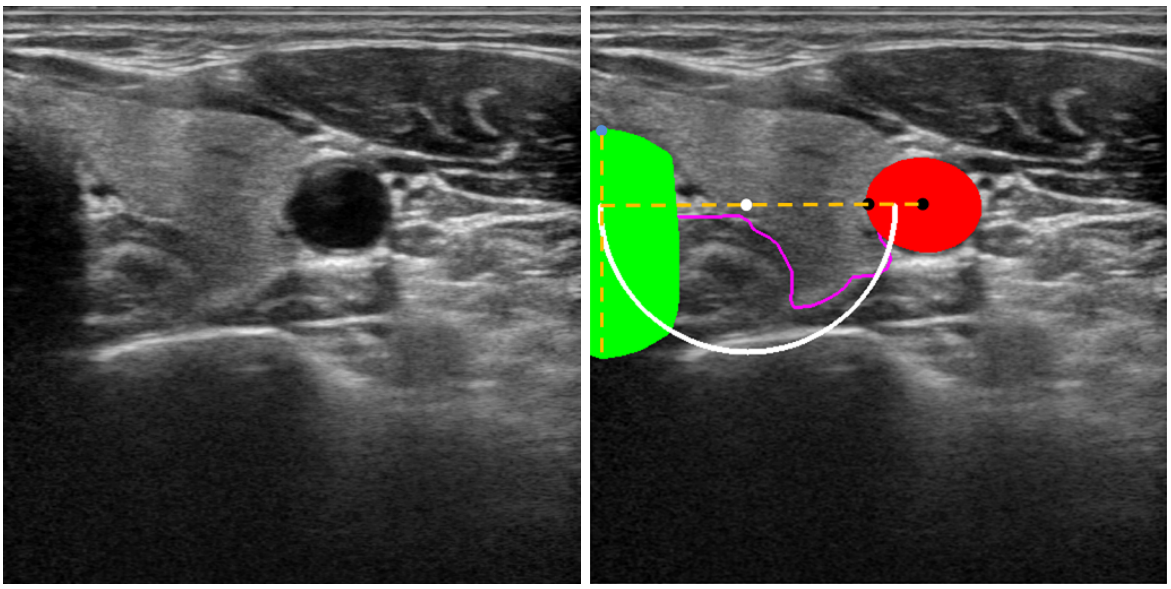
every pixel on the boundary of the detected carotid in $f_{\text{car}}(\mathbf{x})$ the pixel that is closest to it (in terms of Euclidean distance) is retained in $f_{\text{edge}}(\mathbf{x})$ with the rest set to background pixels. The carotid is then segmented by fitting an ellipse using the least squares method with the trace constraint to the points retained in $f_{\text{edge}}(\mathbf{x})$. Figure 5.4(b) shows the segmented carotid $f_{\text{car}}(\mathbf{x})$ for the input in Figure 5.2.

Algorithm 4 Segmenting the carotid artery

1: **initialize** Choose the binary image $f_{\text{hypo}}(\mathbf{x})$ which contains the hypo-echoic regions as its foreground
2: Reset to background in $f_{\text{hypo}}(\mathbf{x})$ all the pixels that correspond to the trachea in $f_{\text{Trc}}(\mathbf{x})$
3: Apply Gaussian filter with $\sigma \in [4, 10]$ to $f_(\mathbf{x})$
4: Determine the Hessian matrix for the region within the Gaussian Kernel
5: Obtain the eigenvalues of the Hessian matrix
6: Determine the pixels belonging to Speckles
7: Apply the binary mask $f_{\text{hypo}}(\mathbf{x})$ on $f_(\mathbf{x})$ and retain only the pixels belonging to speckles as carotid candidates $f_{\text{Car}}(\mathbf{x})$
8: Segment as carotid artery the component in $f_{\text{hypo}}(\mathbf{x})$ that has the lowest number of speckles in $f_{\text{Car}}(\mathbf{x})$

5.2.3 Segmenting the muscles

The largest component of $f_{\text{hypo}}(\mathbf{x})$ that is above the carotid and trachea belongs to the strap muscles. The muscles are segmented by including all the components that lie within the rows of $f_{\text{hypo}}(\mathbf{x})$ that contain the largest component above the carotid and trachea as foreground pixels in a binary image $f_{\text{musc}}(\mathbf{x})$. In longitudinal scans the largest component of muscle in $f_{\text{hypo}}(\mathbf{x})$ is identified to be the one that lies spatially above the



(a) (b)

Figure 5.5: Approximating the anterior boundary of the thyroid gland in transverse scans. (a) Input US image in the transverse scan; (b) Semicircle of interest (in white) that contains the lower boundary of the thyroid gland (in purple). The diameter of this semicircle is the distance between the mid-point of the points in black (the centroid of carotid and point on its boundary closest to the trachea in the same line) and the point in column of the image containing the highest point on trachea (point of intersection of line segments in orange).

largest component in $f_{\text{MH}}(\mathbf{x})$. Figure 5.4(c) shows the segmented muscles $f_{\text{musc}}(\mathbf{x})$ for the input in Figure 5.2.

5.3 Thyroid and Soft Tissue Segmentation

This part of the algorithm deals with classifying the foreground pixels of $f_{\text{MH}}(\mathbf{x})$, into thyroid gland and non-thyroid tissue, respectively by using the Random Forests [189] classifier. The training samples for the classifier come from within the image to be segmented and not from a pool of external training dataset. This is the self-learning part of the proposed method.

The positions of the carotid and trachea relative to the thyroid are used to generate the training pixels in the transverse scans. Figure 4.11 illustrates the procedure to generate the training samples. The non zero pixels in $f_{\text{MH}}(\mathbf{x})$ that lie in between the carotid boundary and trachea boundary in the row containing the centroid of the carotid are chosen as training pixels for the thyroid gland. The non-zero pixels in $f_{\text{trc}}(\mathbf{x})$ and $f_{\text{car}}(\mathbf{x})$ that lie on the same line are chosen as the training pixels for carotid and trachea, respectively.

Since the organ of interest here is the thyroid, the training pixels of the thyroid forms the positive class and that of the carotid and trachea form the negative class. The tissue that below the thyroid gland has the same echogenicity as that of the soft-tissue above the muscles. The location of this tissue in $f(\mathbf{x})$ will determine the bottom boundary of the thyroid and the trachea. Hence, pixels that have the highest intensity in every column of the image above the muscles are also chosen as training samples for the negative class in addition to the pixels of carotid and trachea. As this method: (a) generates training samples; (b) trains classifier; and (c) performs the classification; choosing all the pixels of the segmented carotid and trachea for training samples for negative class will lead to longer execution times. In longitudinal scans, the training pixels for the thyroid are obtained from the non-zero pixels in $f_{\text{MH}}(\mathbf{x})$ that lie mid-way between the muscles and the region of $f_{\text{hypo}}(\mathbf{x})$ located spatially underneath the largest component in $f_{\text{MH}}(\mathbf{x})$.

Denoting the row and column indices of the centroid of the carotid artery by $x_{\text{c\_car}}$ and $y_{\text{c\_car}}$, the set of all hypo-echoic pixels that form a part of the negative training set is given by:

$$\mathcal{F}_{\text{Neg\_Hypo}} = \{(x,y): \quad f_{\text{car}}(x,y) > 0,\ f_{\text{trc}}(x,y) > 0,\ x = x_{\text{c\_car}} \text{ and } y \in [1,N]\}, \tag{5.4}$$

where, N is the number of columns in the image. The hyper-echoic soft-tissue pixels above the muscles also form a part of the negative training set given by:

$$\mathcal{F}_{\text{Neg\_Hyper}} = \{(x,y): \ \forall_{x_i} f(x,y) > f(x_i,y);\ x, x_i \in [1, x_{\text{msc\_top}}], x \neq x_i \text{ and } y \in [1,N]\}. \tag{5.5}$$

The set of all the pixels that form the negative training set is given by:

$$\mathcal{F}_{\text{Neg}} = \mathcal{F}_{\text{Neg\_Hypo}} \cup \mathcal{F}_{\text{Neg\_Hyper}} \tag{5.6}$$

Similarly, the set of pixels that form a part of the positive training set is given by:

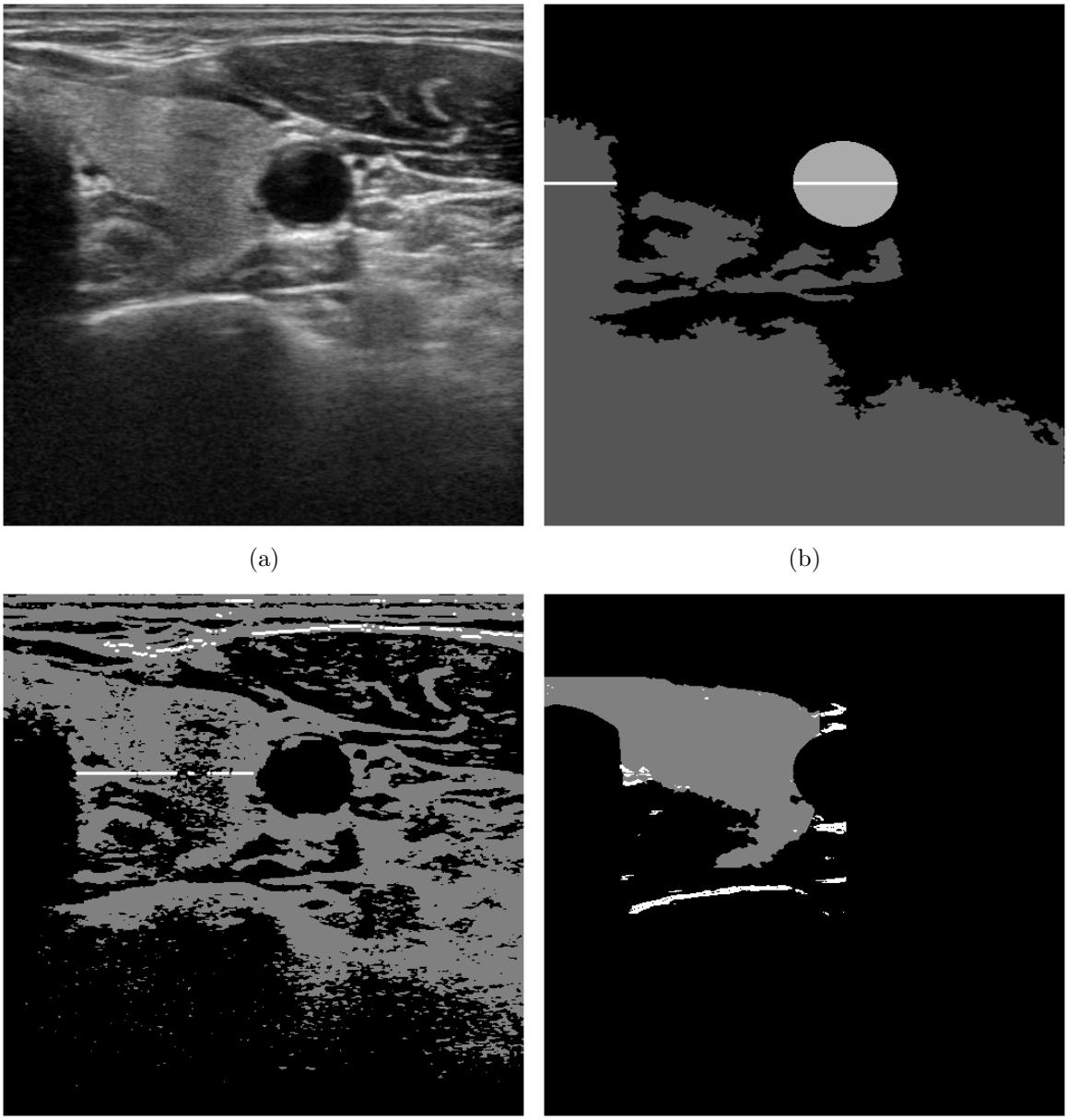
$$\mathcal{F}_{\text{Pos}} = \{(x,y): \ f_{\text{MH}}(x,y) > 0,\ x = x_{\text{c\_car}} \text{ and } y \in [1,N]\}. \tag{5.7}$$

The set of all the training samples is given by:

$$\mathcal{F}_{\text{Train}} = \mathcal{F}_{\text{Pos}} \cup \mathcal{F}_{\text{Neg}} \tag{5.8}$$

The remaining non-zero pixels of $f_{\text{MH}}(\mathbf{x})$ that are not a part of the training set are classified into the thyroid and non-thyroid tissue. Random forests with 500 trees that are trained with local phase, energy and orientation as features are used to perform the supervised classification of pixels in $f_{\text{MH}}(\mathbf{x})$. It is the use of noise robust local phase based features that makes it possible for the use of a pixel level classifier in the proposed method. The binary image $f_{\text{thy}}(\mathbf{x})$ with the pixels of the positive class as foreground pixels contains the segmented thyroid gland.

In certain images, the hyper-echoic tissue underneath the thyroid gland may be absent. This will lead the segmentation to leak into the region below the thyroid gland. Under such circumstances some sort of regularization needs to be applied to prevent



(a) (b)

(c) (d)

Figure 5.6: Segmenting thyroid and soft tissue in transverse scans using self supervised learning: (a) input image; (b). training pixels for carotid and trachea (pixels in white indicate training samples); (c) training samples for the thyroid and soft tissue (pixels in white at the top and bottom are for soft tissue and thyroid, respectively); and (d) segmented thyroid (gray) and soft tissue (white)

segmentation leaks. In transverse scans, the lower boundary of the thyroid gland is fully contained in an imaginary semicircle whose centre lies in the row of the image containing the centroid of the carotid as illustrated in Figure 5.5. All the pixels that lie outside and below this circle are identified as false positive thyroid pixels and are reset to background

in the segmented image. The assumption of the lower boundary of the thyroid gland to be fully contained in a circle is valid as the standard geometrical approximation of a thyroid is an ellipsoid in the literature on volumetric analysis of the thyroid gland [10,14].

Denoting the column index of the pixel that lies on the boundary of the carotid artery and facing the segmented trachea by $y_{\text{car\_trc}}$, the mid point on the line connecting this point and the centroid of carotid is given by:

$$(x_{ct}, y_{ct}) = (x_{\text{c\_car}}, \frac{y_{\text{c\_car}} + y_{\text{car\_trc}}}{2}) \tag{5.9}$$

In section 5.2.1, the row index of the pixel corresponding to the top of the trachea was denoted by $x_{\text{trc\_top}}$. Denoting the column index of the same pixel to be $y_{\text{trc\_top}}$, let the point of intersection of a vertical line passing through $y_{\text{trc\_top}}$ and $x_{\text{c\_car}}$ be denoted by $(x_{\text{c\_car}}, y_{\text{trc\_top}})$. The radius of the semicircle containing the lower boundary of the thyroid gland is then given by:

$$r_{\text{thy}} = |\frac{y_{\text{trc\_top}} - y_{ct}}{2}|. \tag{5.10}$$

All the pixels of $f_{\text{thy}}(\mathbf{x})$ that lie below this semi-circle are reset to 0. The binary image $f_{\text{thy}}(\mathbf{x})$ now contains the thyroid gland fully segmented. Figure 5.6 shows the result of segmenting the thyroid gland in transverse scan using the self-learning method.

It was observed in Fig.5.3, that the actual lower boundary of the trachea lies at the depth of the soft-tissue separating the thyroid gland and the Omohyoid muscles. This depth $x_{\text{trc\_act}}$ is estimated to be the depth of the largest hyper-echoic component resulting after the classification by the self-learning method just described. Looking at Fig. 5.6(d), this corresponds to the row index of the bright streak (in white) directly underneath the thyroid gland (in gray) that spans almost the width of the thyroid gland itself. The lower boundary of the trachea in $f_{\text{trc}}(\mathbf{x})$ is extended up to this point. This completes the MODS in US images of the thyroid gland using the self-learning method.

5.4 Experimental results and discussion

Figure 5.7 shows the results of the proposed MODS algorithm in transverse and longitudinal US scans of the thyroid gland. The success of the carotid detector in Section 5.2.2 depends on the size of the blob and the size of the structuring element (SE). The plot in Fig. 5.8 shows that the best detection accuracy is obtained for a disk type SE of size 5 pixels and blob size greater than 4 and less than 10. Here, the carotid is said to be detected if $F_S \cap F_{GT} \neq \varnothing$, where, F_S and F_{GT} are the set of pixels belonging to the bounding boxes containing the carotid segmented by the proposed method and GT carotid segmentation, respectively. The proposed algorithm is validated for performance by the qualitative and quantitative measures described in Section 3.4. The results of qualitative analysis are summarized in Table 5.1. The results of quantitative analysis on transverse and longitudinal scans are summarized in Tables 5.2 and 5.3, respectively.

Referring to Table 5.1, it can be seen that for both transverse and longitudinal scans, the PRI values are close to 1, GCE values are close to 0, and VOI values are also low. It

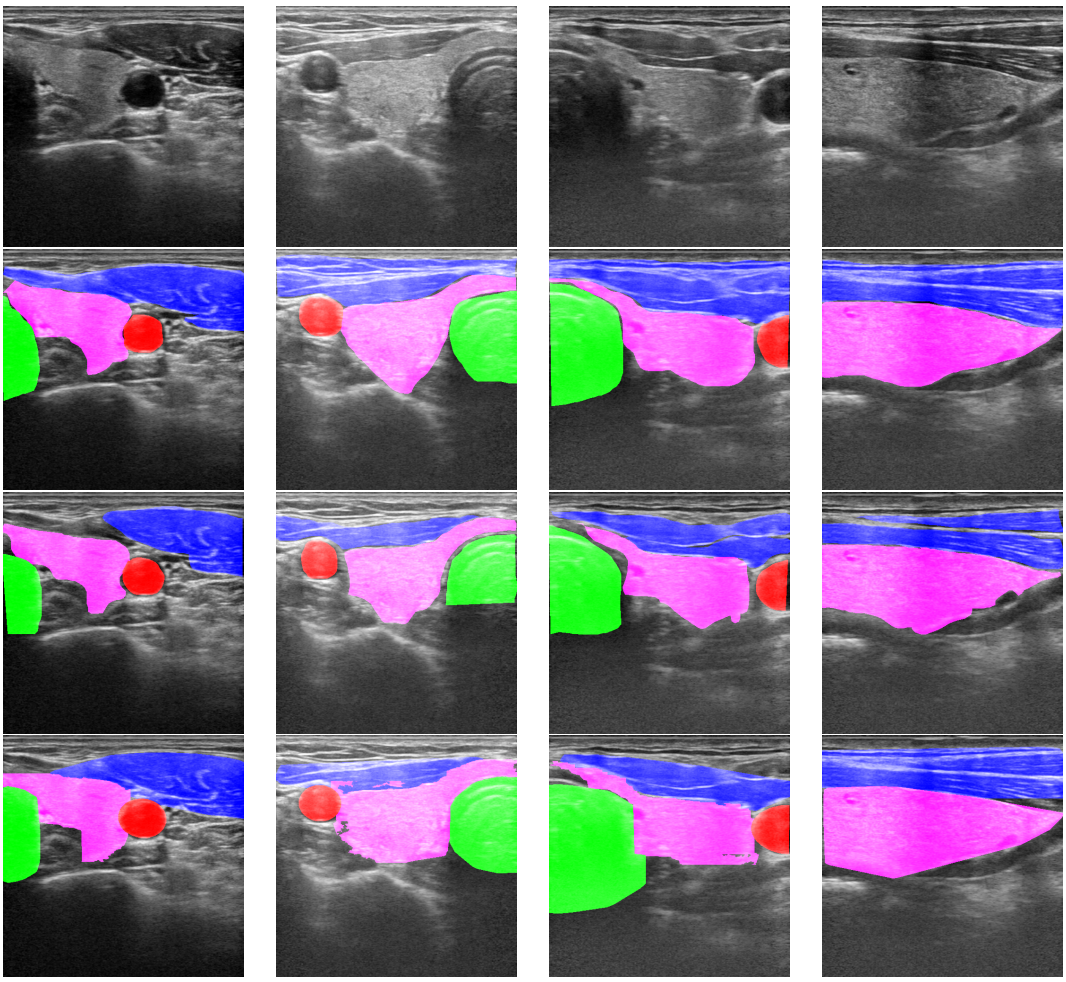


Figure 5.7: MODS results (bottom row) in input US images of the normal thyroid gland (top row) with the second and third rows consisting of manual segmentations by expert 1 and expert 2, respectively.

Dataset	**PRI**	**GCE**	**VOI**	**BE (pixels)**	**BE** (mm)
Transverse	0.860 ±0.037	0.179 ±0.047	1.133 ±0.220	11.965 ±2.900	0.89 ±0.22
Longitudinal	0.835 ±0.043	0.169 ±0.041	1.01 ±0.187	14.0 ±3.05	1.17 ±0.21

PRI = Probabilistic Rand Index. GCE = Global Consistency Error. VOI = Variation of Information BE = Boundary Error.

Table 5.1: Qualitative analysis of the proposed MOS algorithm. It can be seen that PRI is close to 1 and GCE is close to 0 which suggests that the algorithm is capable od producing good quality results.

is interesting to note that the BE values are approximately around 11 pixels or close to 1 mm, which means that the proposed segmentation algorithm can produce results at 1

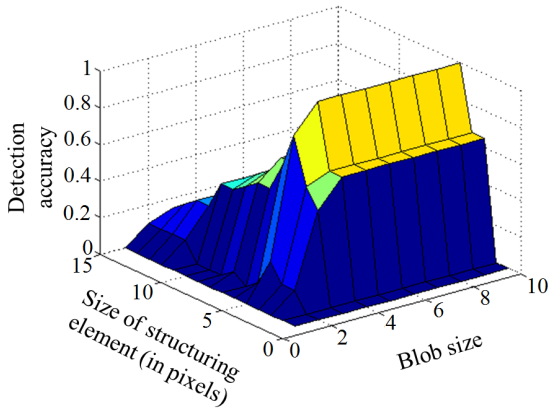


Figure 5.8: Plot of the carotid detector accuracy vs. blob size and size of the structuring element.

Organ	GT	**SE**	**SP**	**DSC**	**PPV**
Thyroid	E1	0.973 ±0.025	0.863 ±0.087	0.857 ±0.050	0.867 ±0.098
	E2	0.963 ±0.024	0.905 ±0.075	0.843 ±0.047	0.800 ±0.087
Carotid	E1	0.997 ±0.003	0.886 ±0.0767	0.874 ±0.043	0.878 ±0.103
	E2	0.996 ±0.002	0.941 ±0.057	0.877 ±0.043	0.829 ±0.085
Muscles	E1	0.992 ±0.006	0.738 ±0.123	0.826 ±0.084	0.952 ±0.030
	E2	0.967 ±0.017	0.920 ±0.080	0.843 ±0.061	0.785 ±0.085
Trachea	E1	0.980 ±0.022	0.923 ±0.050	0.882 ±0.071	0.854 ±0.117
	E2	0.971 ±0.019	0.933 ±0.059	0.844 ±0.081	0.785 ±0.129

⋆ E1 = Expert 1. E2 = Expert 2. SE = Sensitivity. SP = Specificity. DSC = Dice Co-efficient. PPV = Positive Predictive Value. GT = Ground Truth.

Table 5.2: Quantitative analysis on the transverse scans ⋆.

Organ	GT	**SE**	**SP**	**DSC**	**PPV**
Thyroid	E1	0.947 ±0.031	0.903 ±0.073	0.863 ±0.046	0.834 ±0.075
	E2	0.922 ±0.048	0.933 ±0.076	0.820 ±0.079	0.747 ±0.124
Muscles	E1	0.995 ±0.007	0.625 ±0.223	0.734 ±0.189	0.971 ±0.025
	E2	0.963 ±0.035	0.851 ±0.180	0.782 ±0.104	0.769 ±0.123

⋆ E1 = Expert 1. E2 = Expert 2. SE = Sensitivity. SP = Specificity. DSC = Dice Co-efficient. PPV = Positive Predictive Value. GT =Ground Truth.

Table 5.3: Quantitative analysis on the longitudinal scans ⋆

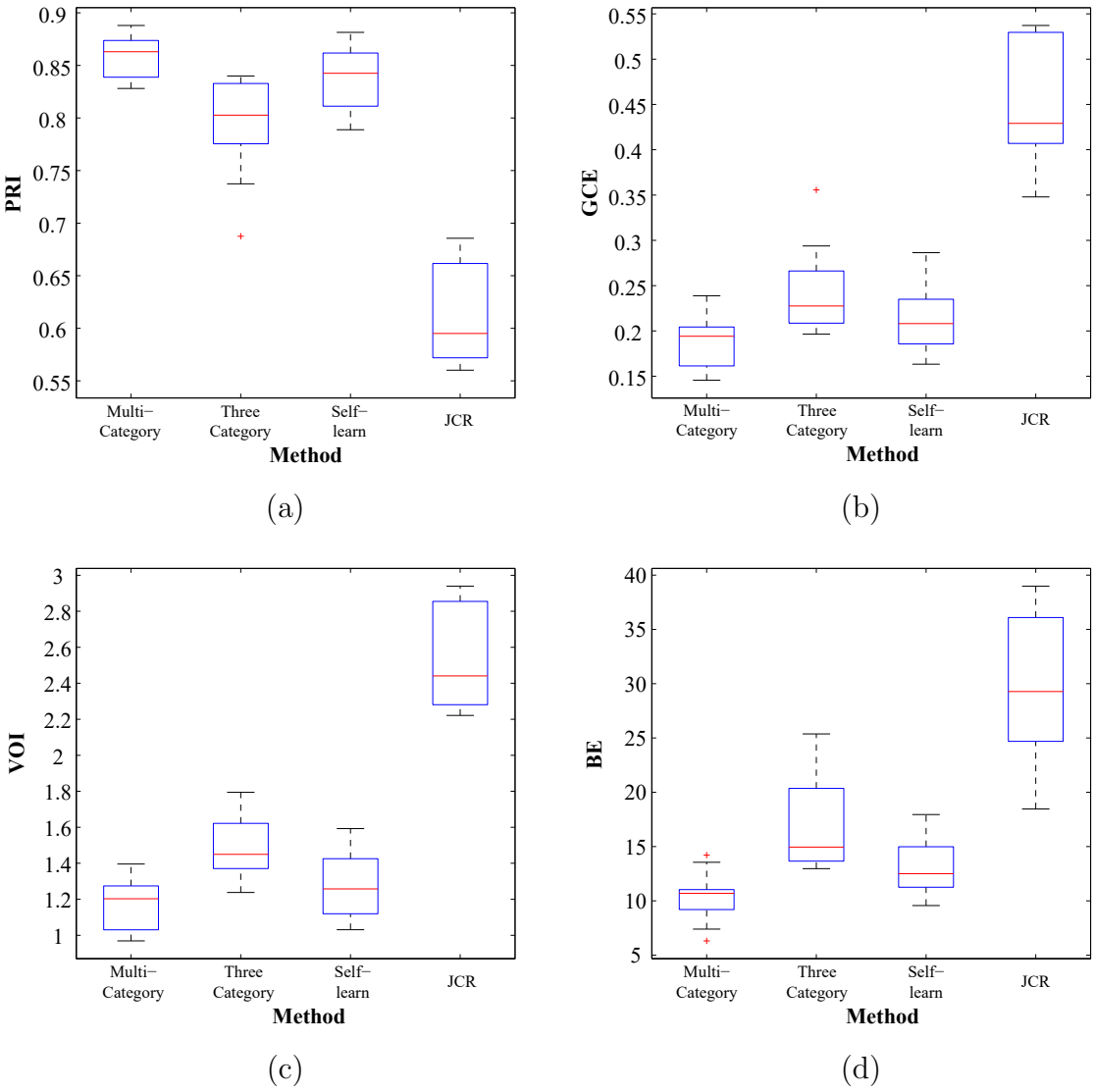


Figure 5.9: Comparison of MODS methods for US images in the transverse scan using (a) Probabilistic Rand Index (PRI). (b) Global Consistency Error (GCE). (c) Variation of Information (VOI). and (d) Boundary Error (BE). Of the four methods that are compared it can be seen that high quality results are obtained by the multi-category clustering algorithm which is followed by the Self-learning and Three category segmentation algorithms. The results of the JCR algorithm, however, are not of good quality.

mm accuracy. Such high accuracy is preferred when performing volume analysis of the thyroid gland where an offset of the segmentation result by more than a few millimetres can result in a large underestimation or overestimation of the gland volume. Looking at the plots in Figures 5.9 and 5.10, it can be observed that the quality of segmentation by the self-learning method is close to the Multi-Category segmentation algorithm of Section

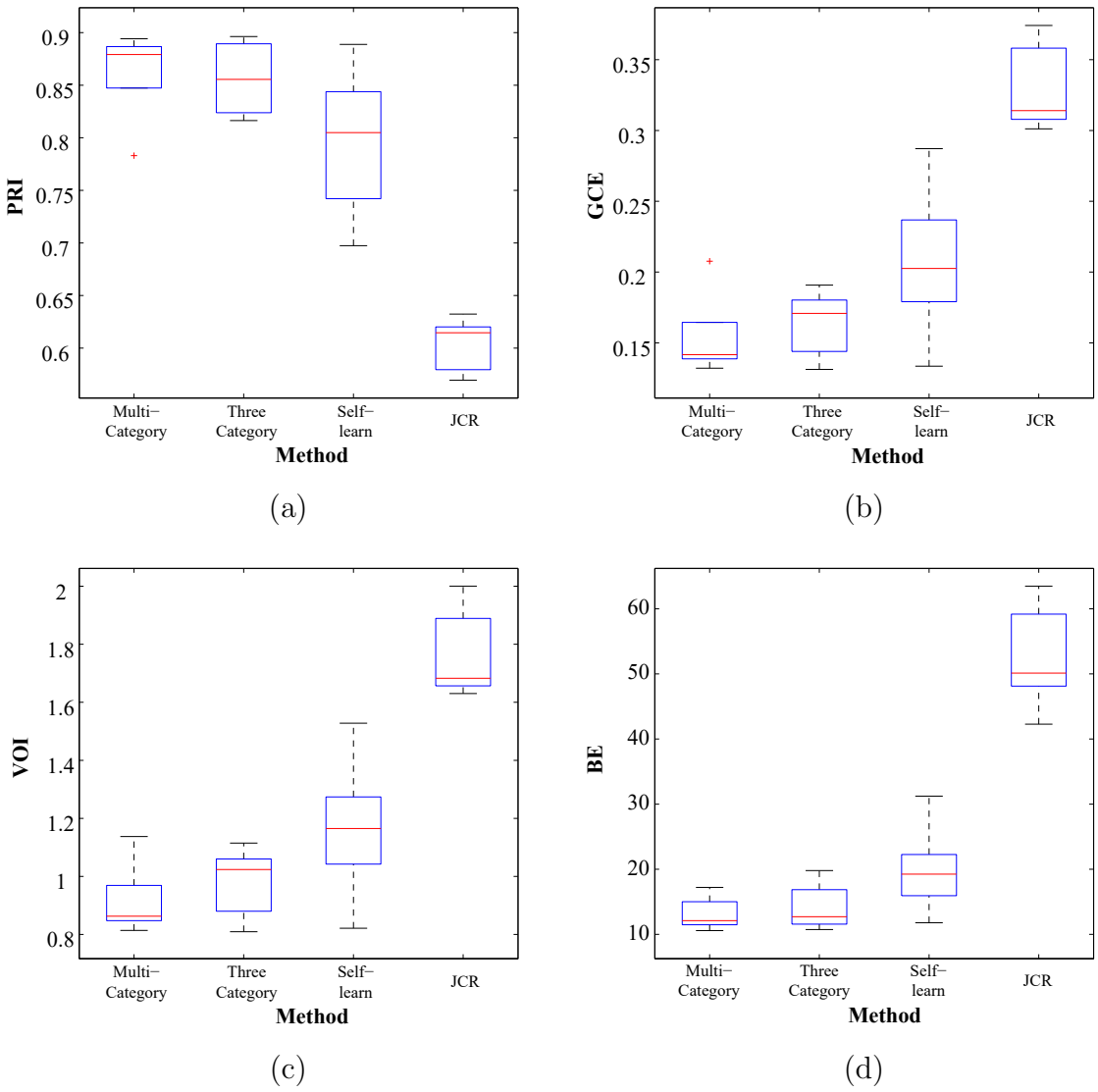


Figure 5.10: Comparison of MODS methods for US images in the longitudinal scan using (a) Probabilistic Rand Index (PRI). (b) Global Consistency Error (GCE). (c) Variation of Information (VOI). and (d) Boundary Error (BE). It can be inferred from the plots that, for longitudinal scans, high quality results are obtained by the multi-category clustering algorithm, followed by the Three category and Self-learning based segmentation algorithms. The results of the JCR algorithm, are again not of good quality.

4.4 and better than the JCR method of [170]. Images in the testing set of the dataset SS were used to generate the plots as the JCR is a supervised segmentation algorithm. But these plots do not provide any information with regard to the accuracy of segmenting individual organs. To measure this we employ the methods introduced in Section 3.4.2.2.

From Tables 5.2 and 5.3, it can be seen that the average Sensitivity is above 0.94 for

all the anatomical structures in both the scans. This suggests that the probability of not detecting any anatomical structure of interest is low. The average specificity values for all the organs except the muscles are above 0.9 when compared to the GT provided by Expert 2 and above 0.85 when the GT is that of Expert 1. High values of specificity for thyroid in both the scans, carotid and trachea in the transverse scans indicate that the probability of the algorithm to misclassify each of the anatomical structures is low. The muscles when seen in a US image has layers separated by fatty tissue layers. While Expert 1 had a tendency to include the fatty tissue as a part of the muscle region, Expert 2 excluded such layers from the region. This explains the discrepancy in the specificity and DSC values between the two observers. This also explains why the Specificity values are below 0.85 for the muscles. High values of average PPV values indicate the low occurrence of false positive segmentations and better performance of the proposed method. The average DSC value for both the observers is more than 0.84 for all the organs in transverse scans, while in the longitudinal scans it is around 0.84 and 0.76 for thyroid and muscles, respectively.

Table 5.4 shows the performance of the proposed method on multi-vendor datasets. It can be seen that the proposed method performs well on multi-vendor datasets demonstrating the robustness of the proposed method to changes in gain settings.

5.4.1 On the choice of classifier

In Section 5.3 a pixel level Random Rorests (RF) classifier was implemented to classify the pixels belonging the hyper-echoic regions into the thyroid and non-thyroid tissue. Since the classification is pixel based with known number of classes and low dimensionality of the feature vector (3 in this case), it is natural to think of using clustering based approaches to perform classification. Table 5.5 summarizes the results of the proposed method, in terms of DSC values averaged over all the images, when classification is performed by using nearest neighbour clustering (k-NN with k=10). In addition, a Neural Networks (NN) classifier with mean squared error (MSE) goal of 0.0001 is also implemented for comparative purposes. It can be seen that the performance of all three classification schemes are comparable to each other with the RF classifier performing slightly better than the others. Hence, a pixel level RF classifier is chosen in the proposed method. The times taken by the proposed method to detect and segment all the organs in transverse (longitudinal) scans are 22.44 ± 3.52 (42.61 ± 6.07)s, 1.15 ± 0.17 (4.87 ± 0.33)s and 20.26 ± 4.32 (57.09 ± 42.52)s respectively for RF based, k-NN based and NN based self-learning methods. RF based method is chosen in this research work for its segmentation accuracy over the other two methods. The k-NN based method can be used in applications that demand real time execution of the MODS algorithm.

5.4.2 Applications

The goal of this work is to provide with a unified framework to: (a) detect carotid for guided interventions; and (b) segment thyroid for volume estimation.

Dataset	**Thyroid**	**Carotid**	**Muscles**	**Trachea**
S1	0.85 ±0.05	0.88 ±0.04	0.84 ±0.07	0.86 ±0.08
S2	0.81 ±0.03	0.88 ±0.03	0.79 ±0.10	0.83 ±0.09

S1 = Dataset 1. S2 = Dataset 2. DSC = Dice Co-efficient.

Table 5.4: Performance on multi-vendor datasets in terms of DSC overlap values averaged over both the experts.

	GT	RF		k-NN		NN	
		T	**L**	**T**	**L**	**T**	**L**
Thy	E1	**0.86±0.05**	**0.86±0.05**	0.85 ±0.08	0.84	0.85 ±0.04	0.71 ±0.14
	E2	**0.85 ± 0.05**	**0.82± 0.08**	0.83 ±0.05	0.81 ±0.09	0.82 ±0.05	0.70 ±0.12
Time (s)		22.34 ±3.52	42.56 ±6.06	**1.04±0.17**	**4.82±0.33**	19.50 ±4.30	56.93 ±42.5

⋆ E1 and E2 stand for expert 1 and expert 2, respectively; RF with 500 trees, k-NN with k=10 and MSE goal of NN = 0.0001. RF = Random Forests. k-NN = K-Nearest Neighbors. NN = Neural Networks. MSE = Mean Squared Error. GT = Ground Truth. s = Time in Seconds.
Best results are highlighted in **Bold** font.

Table 5.5: DSC overlap values of the segmented thyroid gland (**Thy**) using self-learning with different classifiers. While segmentation using the k-NN classifier is the fastest amongst the three, The RF classifier is seen to produce the best results in terms of the DSC metric. ⋆

5.4.2.1 Carotid detection for guided interventions

The proposed method is compared with existing methods by executing it on a public database consisting of 971 thyroid US images in transverse scans. The acquisition details and the link to the database can be found in [138]. The results of comparison are provided in Table 5.6. From the table it can be observed that the proposed method performs better than the reported performance in [138] of other methods on the same dataset. It should also be noted that the database consists of images acquired using Ultrasonix and Toshiba US imaging machines, thereby proving again the robustness of the proposed method to detect carotid artery in multi-vendor datasets under unknown acquisition settings. The proposed algorithm is capable of detecting and segmenting the carotid artery in the presence of abnormalities in the thyroid gland. Figure 5.11 shows the results of the carotid detector on two images of the thyroid gland and one general US image of the neck (Fig. 5.11b) with abnormality.

5.4.2.2 Thyroid segmentation for volumetric analysis

The proposed algorithm can be extended to thyroid volume estimation by considering the segmented thyroid in Section 5.3. The depth and width of the thyroid are obtained at approximately the centre of the segmented thyroid in between the carotid and the

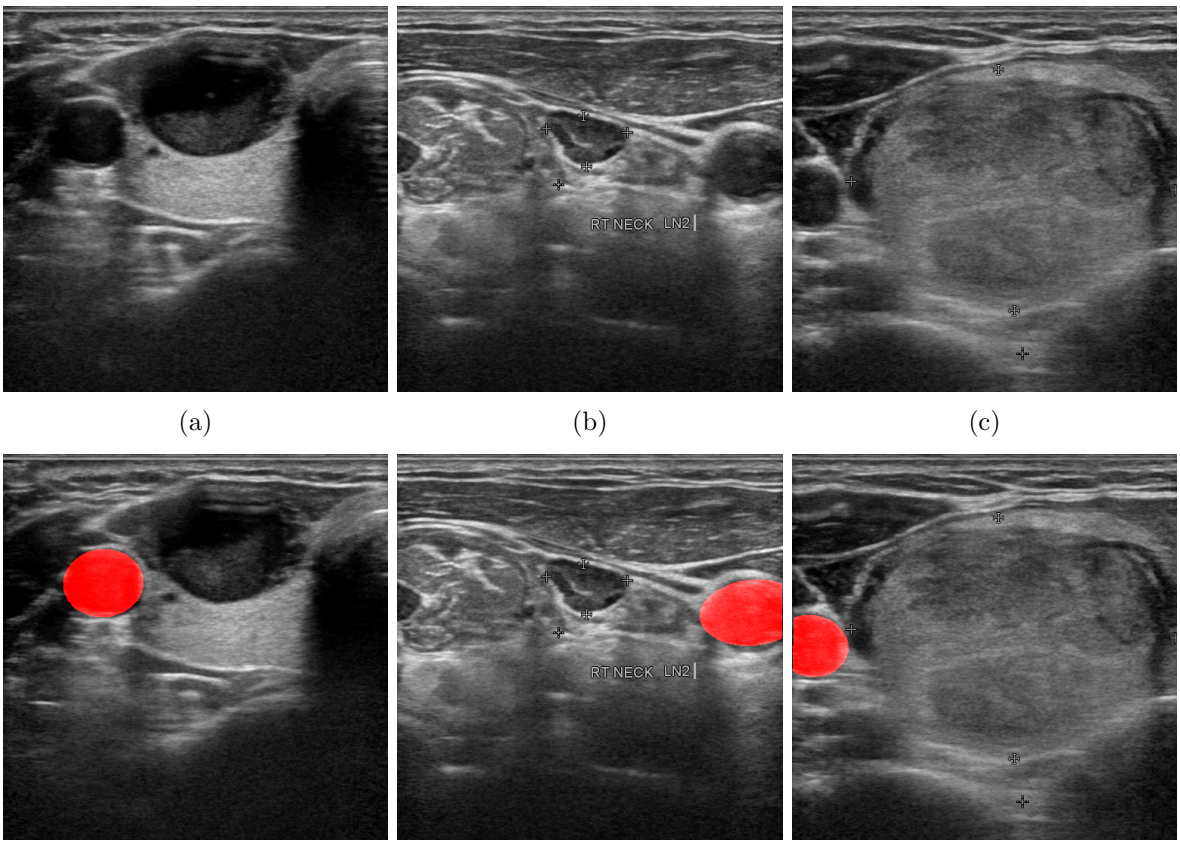


Figure 5.11: Carotid artery detection and segmentation results (bottom row in red) for input US images (top row) of: (a, d) thyroid gland with cyst;(b, e) thyroid gland with hypoechoic nodule; and (c, f) thyroid gland with isoechoic nodule.

	Proposed method [a]	Riha et al. [138] [b]	Golemati et al. [186] [b]
SE (%)	**99**	97	83
(# Images)	(971)	(971)	(750)

[a] Images used are from the dataset made publicly available by Riha et al. [138]
[b] Values and images are as reported in [138] for comparison.
SE = Sensitivity.
Best results are highlighted in **Bold** font.

Table 5.6: Comparison with existing carotid detection algorithms.

Methods	Diff [a]				MSE [b]
	Case 1 (ml)	Case 2 (ml)	Case 3 (ml)	Case 4 (ml)	
	Ellipsoid Model of Eq. (1.3)				
Brown et al. [7]	0.086	0.261	0.002	0.726	0.269
Brunn et al. [15]	**0.072**	**0.219**	**0.002**	**0.608**	**0.225**
	Linear Model of Eq. (1.4)				
Ruggeri et al. [17]					
V_{lobe} using Eq. (1.3)	0.133	0.403	0.004	1.119	0.415
V_{lobe} using Eq. (1.5)	2.717	0.450	3.019	1.914	5.100

[a] Absolute difference between automatically estimated TV and GT.
[b] Mean square error between automatically estimated TV and GT.
Diff = Difference in volume. MSE = Mean Squared Error. TV = Thyroid Volume. GT = Ground Truth.
Best results are highlighted in **Bold** font.

Table 5.7: Volume estimation of thyroid lobe by the proposed method. Of the different variants of Ellipsoid method, the method proposed by Brunn et al. [15] is seen to give the best results.

trachea in the transverse scans. The length of the thyroid gland is the maximum length of the segmented thyroid in the longitudinal scan. By plugging in the estimated length, width, area, and depth of the segmented thyroid gland into Eq. (1.3), its volume can be calculated. Table 5.7 shows the results of volume estimation of thyroid lobes on 4 cases that were acquired at Tan Tock Seng Hospital, Singapore, using Hitachi HI Vision Avius Ultrasound equipped with an L75 5-18MHz probe. Of the four cases, 3 cases were of the right thyroid lobe and one was of the left thyroid lobe. The ground truth volume is determined manually by an expert sonographer with the help of digital calipers provided in the US machine. It can be observed from the table that the volume estimation method proposed by Brunn et al. [15] gives the lowest MSE with respect to the GT lobe volume which proves that the correction for bias does indeed help in improving the estimation accuracy of the ellipsoid method. The full volume of the thyroid gland can be obtained by summing the volumes of each lobe together with the isthmus. This research presents preliminary results of TV estimation using the proposed method. For a complete picture on the usefulness of the proposed method for volume estimation, the results have to be compared with the volume estimated using MRI and post-operative volume estimated using water displacement techniques which is considered as one of the immediate future works.

5.5 Summary

- A self-learning multiorgan detection and segmentation algorithm is proposed in this chapter.
- In this method, the speckle related pixels are classified into the respective echogenic classes by the use of local phase based methods. The carotid artery and the trachea are detected from the binary image containing the hypoechoic pixels as foreground pixels. The pixels of the carotid artery and trachea along with the pixels that lie in the region between them are used to train a random forests classifier to classify the remaining pixels that belong to the hyperechoic tissue into the thyroid gland and background pixels.
- From the plots in Figures 5.9 and 5.10, it can be inferred that the Multi-Category segmentation introduced in Chapter 4 is the best for MODS in US images of the thyroid gland and is closely followed by the Self-learning method introduced in this chapter.
- Applications of this method to track the carotid artery and perform volumetric analysis are also discussed in this chapter.

Chapter 6

Conclusion and Future Works

6.1 Conclusion

Segmentation algorithms for the ultrasound images of the thyroid gland are developed with the aim of helping a doctor to perform: (a) volumetric analysis for radiotherapy; or (b) guided interventions. While binary segmentation algorithms to segment the thyroid gland or carotid artery do exist, a unified framework that can do both has not been proposed so far. The motivations to take up this research work on devising multi-organ detection and segmentation algorithms for the ultrasound images of the thyroid gland are described in detail in Chapter 1.

The current methods proposed to segment the thyroid gland and the carotid artery are semi-automatic. These algorithms require the organs to be manually detected by an expert (in the form of seed point initializations or curve initializations around the organ of interest) following which the segmentation is performed by automated methods such as geodesic active contours, classification by neural networks or shape fits by ellipse, etc. The challenge is really in automating the organ detection part of the segmentation workflow. Given that the image acquisitions in our research work are freehand 2D ultrasound image acquisitions, there needs to be a way for the computer to recognize the organs in question and then apply the appropriate segmentation algorithm for the organ. One approach that is commonly followed for ultrasound images is by the use of external markers attached to the ultrasound probe. A tracking device (video camera for example) is used to track the movement of the markers and then this information is used to co-register the ultrasound images with a 3D volume of the gland obtained by CT or MRI imaging modalities. Once the images are co-registered, atlas based methods or methods based on marginal space learning may be employed to detect the organs of interest and then segment the detected organs. This approach is quite tedious and computationally expensive.

Interestingly, the properties of ultrasound waves and its interaction with tissues provide a way to detect organs without the need for external markers and tracking devices. A detailed description on the physics behind ultrasound image formation and the properties of ultrasound waves are discussed in Chapter 2. The properties of ultrasound waves to undergo reflection and scattering results in the echogenicity of a tissue and its

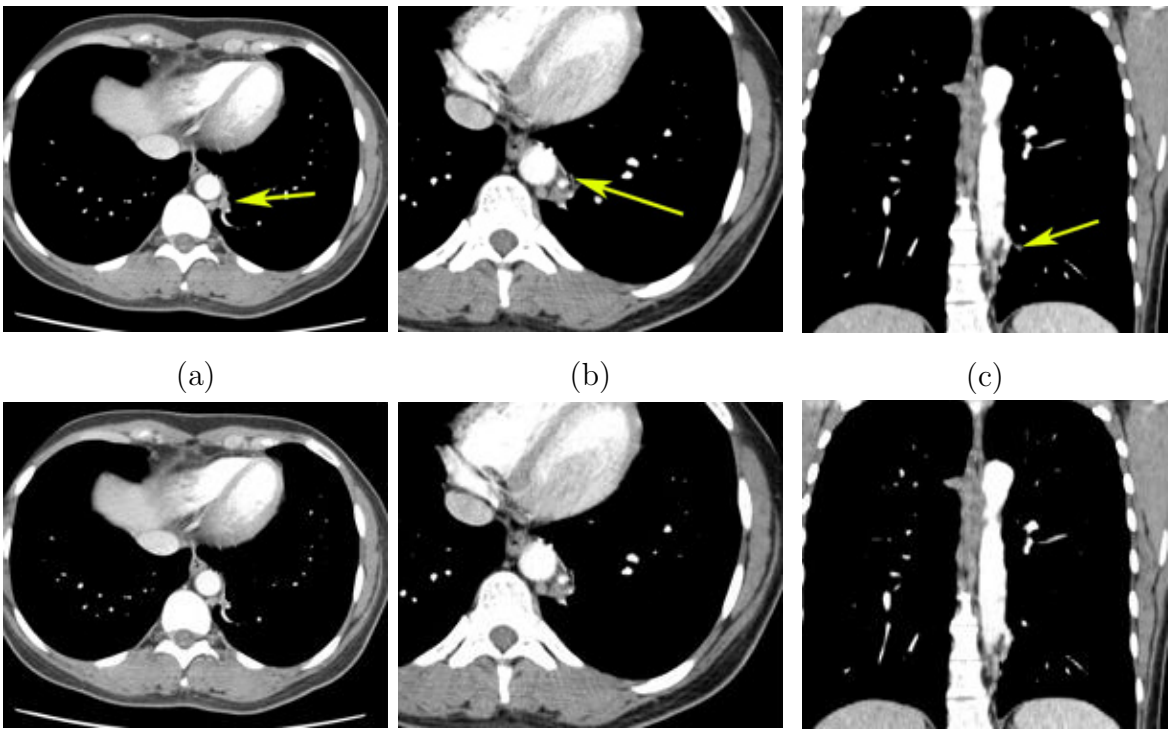
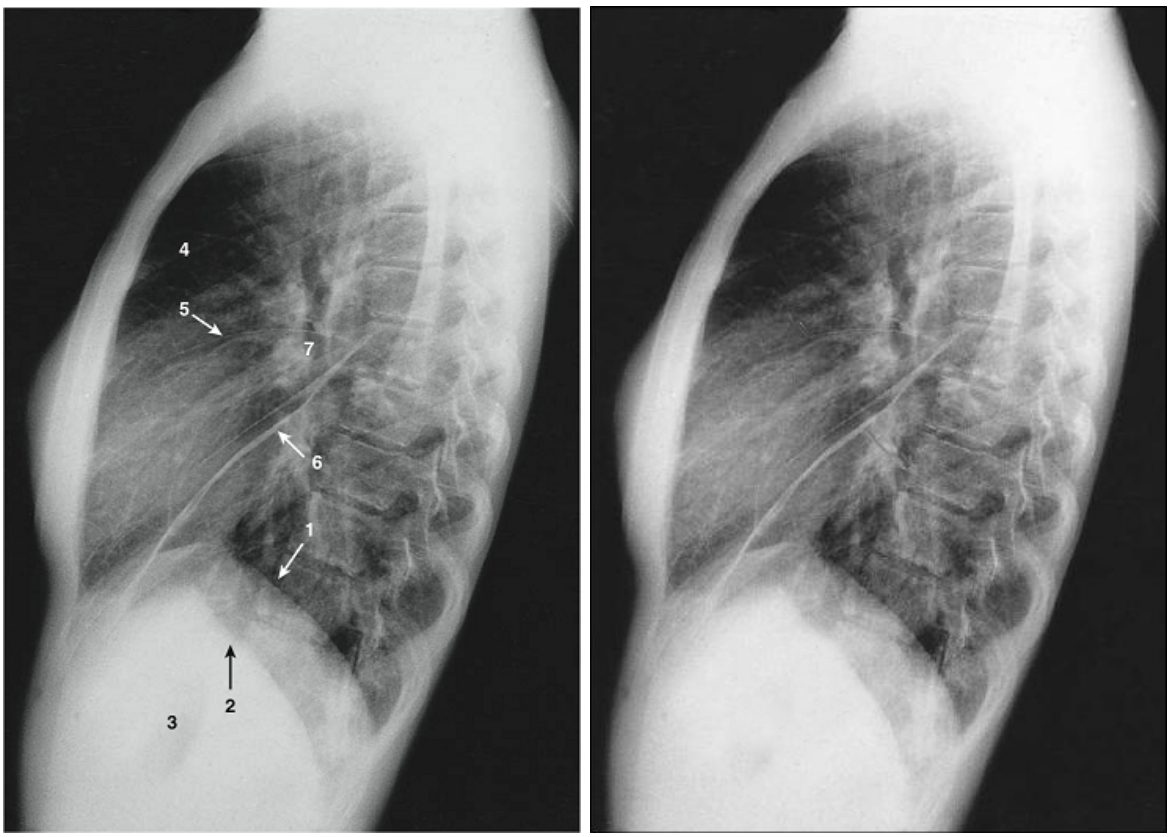


Figure 6.1: Annotation removal and image restoration in CT images. (a)-(c) Input; and (d)-(f) Output.

property to undergo refraction, interference and diffraction results in the formation of imaging artefacts. The enhancement artefacts are of particular interest in this research work as they predominantly occur under blood vessels and can be used as a landmark to detect the carotid artery. Enhancement artefacts appear hyperechoic in ultrasound images. Since speckles are embodiments of tissue echogenicity, speckle related pixels are used to perform echogenicity based tissue characterization to automatically detect the artefacts. The detection and segmentation algorithms need to be: (a) robust to changes in the acquisition parameters; and (b) applicable to multi-vendor datasets under varying conditions of tissue brightness. This is achieved by the use of local phase based methods which are insensitive to changes in signal energy and yet preserve the structural integrity of the organs in its phase. Thus, tissue characterization using speckles, anatomic structure localization using enhancement artefacts and region based segmentation using local phase forms the core of the MODS methods proposed in this research work.

The datasets and the validation protocol used in this research are discussed in Chapter 3. Five different datasets from four different vendors are used in our experiments. One of the five is a publicly available dataset that is used to test the detection accuracy of the carotid artery. One dataset is used to measure the performance of the pre-processing algorithm and the rest are used to validate the MODS methods. A novel algorithm is proposed in this chapter to remove annotations from ultrasound images. A histogram



(a) (b)

Figure 6.2: Annotation removal and image restoration in X-Ray images. (a) Input; and (b) Output.

based approach is followed to detect the annotations within the FOV of an image and POCS based methods are employed to restore the regions of images that were obscured by the annotation. The application of this algorithm is not restricted to ultrasound images alone. Figures 6.1 and 6.2 show the results of the annotation removal algorithm on CT and X-Ray images, respectively. The performance of annotation removal algorithm is qualitatively assessed by conducting online surveys where the participants agreed that the annotations were removed without any trace from the images. Experimental results also showed that the restored images are of high quality with average PSNR>38dB. On the other hand, four qualitative metrics and five quantitative metrics are employed to validate the segmentation algorithms. While the qualitative metrics allow for the comparison of different algorithms, the quantitative metrics are used to compare the segmentations with respect to the expert segmentor. Care must be taken when comparing different algorithms using the qualitative metrics as multi-organ segmentation algorithms cannot be compared to binary segmentation algorithms using the qualitative metrics. Existing methods are validated using the nine metrics and it is found that the current methods tend to under/over segment the organs under consideration.

Two new methods to perform unsupervised MODS in ultrasound images of the thyroid gland are proposed in Chapter 4. The first of the two methods is called the Three category MODS system. In this method, the speckle related pixels are clustered into three echogenic classes based on its relative brightness in a kernel. An energy based model is then used to detect the carotid artery in the labelled image (image whose pixels are labelled with the cluster number). The carotid artery is used as a landmark to detect and segment the remaining organs in the image. The second method is called the Multi-category MODS system. In this method, an agglomerative clustering constrained with a similarity metric is proposed to cluster the speckle related pixels into an unknown number of echogenic classes. The total number of class labels generated at the end of the clustering determines the number of echogenic classes in the image. This process is viewed as a quantization of the speckle related pixels into levels that are determined by the tissue echogenicity. The enhancement artefact is then detected in the quantized image which is in turn used as a landmark to detect the carotid artery. The carotid artery is used as a landmark to detect the remaining organs. All organs are segmented by the application of region based local phase methods. The proposed methods outperform the existing methods to segment both individual and multiple organs. Amongst the two algorithms proposed, the multi-category algorithm is found to perform better than the three category algorithm.

A self-learning MODS algorithm is proposed in Chapter 5. In this method, the speckle related pixels are classified into the respective echogenic classes by the use of local phase based methods. The carotid artery and the trachea are detected from the binary image containing the hypoechoic pixels as foreground pixels. The pixels of the carotid artery and trachea along with the pixels that lie in the region between them are used to train a random forests classifier to classify the remaining pixels that belong to the hyperechoic tissue into the thyroid gland and background pixels. The applications of this MODS method to perform guided interventions and volumetric analysis are also discussed in that chapter.

6.2 Future Works

The following directions are identified as potential future works:

- The speckle based clustering and quantization methods for tissue characterizations can be seen as despeckeling methods. One of the future works is to compare the clustering and quantization methods to existing despeckling methods such as that proposed in [190].
- To test the validity of the use of imaging artefacts as landmarks to detect and segment multiple organs in ultrasound images of the abdomen, prostate etc..
- The primary motivation for taking up this research work on developing MODS systems was for it to be used as a platform to perform guided interventions and

volumetric analysis. The next step will be to design a robotic guided intervention system that will employ the carotid detection and segmentation algorithms of Chapters 4 or 5 to automatically determine the plane of needle insertion for FNAB or cannulation of the IJV. This will help to prevent accidental punctures of the carotid artery during the procedure.

- The shape of the thyroid gland is assumed to be an ellipsoid when determining its volume by Eq. (1.2). It is widely accepted that this assumption leads to the underestimation of the thyroid volume. One of the next steps is to reconstruct the 3D volume of the thyroid gland by making use of the segmented thyroid gland in successive frames and registering the images in two orthogonal planes. This can be done without the use of markers by making use of the tracheal rings as the interest points for the registration algorithm. For a given subject, the distance between the tracheal rings is known to remain constant [191]. By detecting the trachea in the transverse scans and the tracheal rings in the longitudinal scans, the two images can be registered and the segmentation of the thyroid gland in the registered 3D volume can be used to derive a mathematical model to estimate the thyroid volume in cubic millimetres. Of course, this would require extensive validation with the water displacement method to establish the accuracy of the method.

- One of the assumptions in the formulation of the three MODS methods is that all organs are present in the images that are input to the system. Based on this assumption the algorithms work in a sequential manner by detecting one organ at a time. The first organ to be detected is used as a landmark to detect the next organ and so on. The algorithm fails when the first organ is not detected or incorrectly detected. This can happen when the image does not contain the first organ of interest and contains the rest or when the first organ to be detected is occluded. Thus, one of the future works will be to develop an unsupervised MODS system that detects all the organs simultaneously without depending on the successful detection of other organs.

- The goal of this research work was to devise MODS methods that can help in guided interventions and volumetric analysis. Now that we have proposed three methods to perform MODS, the next step will be to validate it for guided interventions and volumetric analysis by acquiring large datasets and conducting experiments.

Author's Publications

Journals

[1] Narayan N.S., Marziliano P., Kanagalingam J. and Hobbs C.G.L., Speckle Patch Similarity for Echogenicity based Multi-Organ Segmentation in Ultrasound Images of the Thyroid Gland , *accepted for publication in the IEEE Journal of Biomedical and Health Informatics*,2015.

[2] Narayan N.S., Marziliano P., Kumar H.J.R. and Seelamantula C.S., Landmark Driven Self Learning Based Approach to Multi-Organ Segmentation in Ultrasound Images of the Thyroid Gland, *submitted to IEEE Trans. Medical Imaging.*

Conferences

[1] Kumar H.J.R, Seelamantula C.S., N.S. Narayan and Marziliano P., "Automatic Segmentation of Common Carotid Artery in Transverse Mode Ultrasound Images", *Accepted for presentation at IEEE International Conference on Image Processing(ICIP 2016), Sept. 25-28, 2016*

[2] Narayan N.S. and Marziliano P., Echogenicity based approach to detect, segment and track the common carotid artery in 2D ultrasound images, *in Proc. 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE EMBC)*, pp. 2989 - 2992 2015.

[3] Narayan N.S., Marziliano P., Kanagalingam J. and Hobbs C.G.L., Speckle in Ultrasound Images: Friend or Foe? *IEEE International Conference on Image Processing(ICIP 2014)*, pp. 5816 - 5820, 2014.

[4] Narayan N.S., Marziliano P. and Hobbs C.G.L., Automatic removal of manually induced artefacts in 2D ultrasound Images of Thyroid Gland, *in Proc. 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE EMBC)*, pp. 33993402, 2013.

Bibliography

[1] S. A. Zenios, J. Makower, and P. G. Yock, *Biodesign : the process of innovating medical technologies*, ser. Cambridge texts in biomedical engineering. Cambridge University Press, 2010.

[2] B. B. Tempkin, *Ultrasound scanning : principles and protocols*. St. Louis, Mo. : Elsevier Saunders, 2009.

[3] J. G. Hubbard, W. B. Inabnet, and C. y. Lo, *Endocrine Surgery: Principles and Practice*. Springer, 2009.

[4] R. V. Lloyd, *Endocrine pathology: differential diagnosis and molecular advances*. Springer, 2010.

[5] A. J. Wood, A. R. Hermus, and D. A. Huysmans, "Treatment of benign nodular thyroid disease," *New England Journal of Medicine*, vol. 338, no. 20, pp. 1438–1447, 1998.

[6] J. W. van Isselt, J. M. de Klerk, P. P. van Rijk, A. P. van Gils, L. J. Polman, C. Kamphuis, R. Meijer, and F. J. Beekman, "Comparison of methods for thyroid volume estimation in patients with graves' disease," *European journal of nuclear medicine and molecular imaging*, vol. 30, no. 4, pp. 525–531, 2003.

[7] M. C. Brown and R. Spencer, "Thyroid gland volume estimated by use of ultrasound in addition to scintigraphy," *Acta Oncologica*, vol. 17, no. 4, pp. 337–341, 1978.

[8] S. Bontha, D. Dupuy, J. Monchik, and I. Jackson, "Use of volume derived from three-dimensional computed tomography to follow-up a retrosternal compressive toxic multinodular goiter treated with I-131," *Endocrinologist*, vol. 11, no. 3, pp. 247–249, Jun. 2001.

[9] U. E. Vurdem, N. Acer, T. Ertekin, A. Savranlar, O. Topuz, and M. Keceli, "Comparison of three volumetric techniques for estimating thyroid gland volume," *Turkish Journal of Medical Sciences*, vol. 42, no. 1, pp. 1299–1306, Dec. 2012.

[10] R. Malago, M. DOnofrio, M. Ferdeghini, W. Mantovani, C. Colato, P. Brazzarola, M. Motton, and R. P. Mucelli, "Thyroid volumetric quantification comparative

evaluation between conventional and volumetric ultrasonography," *Journal of Ultrasound in Medicine*, vol. 27, no. 12, pp. 1727–1733, 2008.

[11] V. Brauer, P. Eder, K. Miehle, T. Wiesner, H. Hasenclever, and R. Paschke, "Interobserver variation for ultrasound determination of thyroid nodule volumes," *Thyroid*, vol. 15, no. 10, pp. 1169–1175, 2005.

[12] R. Hermans, R. Bouillon, K. Laga, P. Delaere, B. D. Foer, G. Marchal, and A. Baert, "Estimation of thyroid gland volume by spiral computed tomography," *European radiology*, vol. 7, no. 2, pp. 214–216, 1997.

[13] C. Swinyard, "Methods for volumetric determination of fresh endocrine glands," *The anatomical record*, vol. 74, no. 1, pp. 71–78, 1939.

[14] B. Nygaard, T. Nygaard, L. I. Jensen, P. Søe-Jensen, K. Gerhard Nielsen, M. Fugl, L. Hegedüs *et al.*, "Thyroid volume measured by ultrasonography and ct," *Acta Radiologica*, vol. 43, no. 3, pp. 269–274, 2002.

[15] J. Brunn, U. Block, G. Ruf, I. Bos, W. Kunze, P. Scriba *et al.*, "Volumetric analysis of thyroid lobes by real-time ultrasound," *Deutsche medizinische Wochenschrift (1946)*, vol. 106, no. 41, pp. 1338–1340, 1981.

[16] W. Shabana, E. Peeters, and M. De Maeseneer, "Measuring thyroid gland volume: Should we change the correction factor?" *American Journal of Roentgenology*, vol. 186, no. 1, pp. 234–236, 2006.

[17] M. Ruggieri, A. Fumarola, A. Straniero, A. Maiuolo, I. Coletta, A. Veltri, A. Di Fiore, P. Trimboli, P. Gargiulo, M. Genderini *et al.*, "The estimation of the thyroid volume before surgery: an important prerequisite for minimally invasive thyroidectomy," *Langenbeck's Archives of Surgery*, vol. 393, no. 5, pp. 721–724, 2008.

[18] P. Trimboli, M. Ruggieri, A. Fumarola, M. D'Alò, A. Straniero, A. Maiuolo, S. Ulisse, and M. D'Armiento, "A mathematical formula to estimate in vivo thyroid volume from two-dimensional ultrasonography," *Thyroid*, vol. 18, no. 8, pp. 879–882, 2008.

[19] A. G. Randolph, D. J. Cook, C. A. Gonzales, and C. G. Pribble, "Ultrasound guidance for placement of central venous catheters: a meta-analysis of the literature," *Critical care medicine*, vol. 24, no. 12, pp. 2053–2058, 1996.

[20] P.-A. C. Abboud and J. L. Kendall, "Ultrasound guidance for vascular access," *Emergency medicine clinics of North America*, vol. 22, no. 3, pp. 749–773, 2004.

[21] G. Turker, F. N. Kaya, A. Gurbet, H. Aksu, C. Erdogan, and A. Atlas, "Internal jugular vein cannulation: an ultrasound-guided technique versus a landmark-guided technique," *Clinics*, vol. 64, no. 10, pp. 989–992, 2009.

[22] M. Lamperti, A. R. Bodenham, M. Pittiruti, M. Blaivas, J. G. Augoustides, M. Elbarbary, T. Pirotte, D. Karakitsos, J. LeDonne, S. Doniger *et al.*, "International evidence-based recommendations on ultrasound-guided vascular access," *Intensive care medicine*, vol. 38, no. 7, pp. 1105–1117, 2012.

[23] D. Karakitsos, N. Labropoulos, E. De Groot, A. P. Patrianakos, G. Kouraklis, J. Poularas, G. Samonis, D. A. Tsoutsos, M. M. Konstadoulakis, and A. Karabinis, "Real-time ultrasound-guided catheterisation of the internal jugular vein: a prospective comparison with the landmark technique in critical care patients," *Critical Care*, vol. 10, no. 6, pp. 162–170, 2006.

[24] D. P. Odasso, "Internal jugular vein and carotid artery anatomic relation as determined by ultrasonography," *Anesthesiology*, vol. 85, no. 1, pp. 35–8, 1996.

[25] P. M. Maietta, "Accidental carotid artery catheterization during attempted central venous catheter placement: A case report," *American Association of Nurse Anesthetists (AANA) journal*, vol. 80, no. 4, p. 251, 2012.

[26] M. Reuber, L. Dunkley, E. Turton, M. Bell, and J. Bamford, "Stroke after internal jugular venous cannulation," *Acta Neurologica Scandinavica*, vol. 105, no. 3, pp. 235–239, 2002.

[27] S. Mitchell and R. A. Clark, "Complications of central venous catheterization," *American Journal of Roentgenology*, vol. 133, no. 3, pp. 467–476, 1979.

[28] J. Rajinikanth, E. Stephen, and S. Agarwal, "Complication of central venous cannulation," *Canadian Journal of Surgery*, vol. 51, no. 5, p. E113, 2008.

[29] M. Blaivas and S. Adhikari, "An unseen danger: Frequency of posterior vessel wall penetration by needles during attempts to place internal jugular vein central catheters using ultrasound guidance*," *Critical care medicine*, vol. 37, no. 8, pp. 2345–2349, 2009.

[30] J. U. Quistgaard, "Signal acquisition and processing in medical diagnostic ultrasound," *IEEE Signal Processing Magazine*, vol. 14, no. 1, pp. 67–74, 1997.

[31] A. Herment, J. Guglielmi, P. Dumée, P. Peronneau, and P. Delouche, "Limitations of ultrasound imaging and image restoration," *Ultrasonics*, vol. 25, no. 5, pp. 267–273, 1987.

[32] A. Dunstan, C. Silkowski, and C. Odwin, *Emergency Medicine Sonography: Pocket Guide to Sonographic Anatomy and Pathology.* Jones & Bartlett Learning, 2010.

[33] J. P. McGahan and B. B. Goldberg, *Diagnostic ultrasound.* New York : Informa Healthcare, 2008.

[34] H. Stippel, "Speckle suppression, segmentation and registration of medical ultrasound images," Ph.D. dissertation, Ghent University, 2003.

[35] C. B. Burckhardt, "Speckle in ultrasound b-mode scans," *IEEE Transactions on Sonics and Ultrasonics*, vol. 25, no. 1, pp. 1–6, 1978.

[36] R. F. Wagner, S. W. Smith, J. M. Sandrik, and H. Lopez, "Statistics of speckle in ultrasound b-scans," *IEEE Transactions on Sonics and Ultrasonics*, vol. 30, no. 3, pp. 156–163, 1983.

[37] J. M. Thijssen, "Ultrasonic speckle formation, analysis and processing applied to tissue characterization," *Pattern Recognition Letters*, vol. 24, no. 4, pp. 659–675, 2003.

[38] T. Eltoft, "Modeling the amplitude statistics of ultrasonic images," *IEEE Transactions on Medical Imaging*, vol. 25, no. 2, pp. 229–240, 2006.

[39] V. Dutt and J. F. Greenleaf, "Adaptive speckle reduction filter for log-compressed b-scan images," *IEEE Transactions on Medical Imaging*, vol. 15, no. 6, pp. 802–813, 1996.

[40] S. Balocco, C. Gatta, J. M. Ferré, and P. Radeva, "Ultrasound despeckle methods," in *Ultrasound Imaging*. Springer, 2012, pp. 49–71.

[41] J. Bamber and R. Dickinson, "Ultrasonic b-scanning: a computer simulation," *Physics in medicine and biology*, vol. 25, no. 3, p. 463, 1980.

[42] Y. Yu and S. T. Acton, "Speckle reducing anisotropic diffusion," *IEEE Transactions on Image Processing*, vol. 11, no. 11, pp. 1260–1270, 2002.

[43] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629–639, 1990.

[44] Y. Yue, M. M. Croitoru, A. Bidani, J. B. Zwischenberger, and J. W. Clark Jr, "Nonlinear multiscale wavelet diffusion for speckle suppression and edge enhancement in ultrasound images," *IEEE Transactions on Medical Imaging*, vol. 25, no. 3, pp. 297–311, 2006.

[45] Y. Deng, Y. Wang, and Y. Shen, "Speckle reduction of ultrasound images based on rayleigh-trimmed anisotropic diffusion filter," *Pattern Recognition Letters*, vol. 32, no. 13, pp. 1516–1525, 2011.

[46] E. N. Kollorz, D. A. Hahn, R. Linke, T. W. Goecke, J. Hornegger, and T. Kuwert, "Quantification of thyroid volume using 3-D ultrasound imaging," *IEEE Transactions on Medical Imaging*, vol. 27, no. 4, pp. 457–466, 2008.

[47] C. Chinrungrueng and A. Suvichakorn, "Fast edge-preserving noise reduction for ultrasound images," *IEEE Nuclear Science Symposium Conference Record*, vol. 3, pp. 849–854, 2000.

[48] H. Rabbani, M. Vafadust, P. Abolmaesumi, and S. Gazor, "Speckle noise reduction of medical ultrasound images in complex wavelet domain using mixture priors," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 9, pp. 2152–2160, 2008.

[49] C.-Y. Chang, Y.-F. Lei, C.-H. Tseng, and S.-R. Shih, "Thyroid segmentation and volume estimation in ultrasound images," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 6, pp. 1348 –1357, 2010.

[50] ——, "Thyroid Segmentation and Volume Estimation in Ultrasound Images," *IEEE International Conference on Systems, Man and Cybernetics*, pp. 3442–3447, Oct. 2008.

[51] J. Seabra, F. Ciompi, O. Pujol, J. Mauri, P. Radeva, and J. Sanches, "Rayleigh mixture model for plaque characterization in intravascular ultrasound," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 5, pp. 1314–1324, 2011.

[52] A. Nilsson, "Artefacts in sonography and doppler," *European Radiology*, vol. 11, no. 8, pp. 1308–1315, 2001.

[53] K. K. Win, J. Wang, C. Zhang, and R. Yang, "Identification and removal of reverberation in ultrasound imaging," *5th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, pp. 1675–1680, 2010.

[54] N. E. Bylund, M. Andersson, and H. Knutsson, "Interactive 3D filter design for ultrasound artifact reduction," *IEEE International Conference on Image Processing*, vol. 3, pp. III–728–31, 2005.

[55] P. C. Tay, S. T. Acton, and J. Hossack, "A transform method to remove ultrasound artifacts," *IEEE Southwest Symposium on Image Analysis and Interpretation*, pp. 110–114, 2006.

[56] M. Felsberg and G. Sommer, "The monogenic signal," *IEEE Transactions on Signal Processing*, vol. 49, no. 12, pp. 3136–3144, 2001.

[57] G. H. Granlund and H. Knutsson, *Signal processing for computer vision*. Springer Science & Business Media, 2013.

[58] A. V. Oppenheim and J. S. Lim, "The importance of phase in signals," *Proceedings of the IEEE*, vol. 69, no. 5, pp. 529–541, 1981.

[59] D. Boukerroui, J. A. Noble, and M. Brady, "On the choice of band-pass quadrature filters," *Journal of Mathematical Imaging and Vision*, vol. 21, no. 1-2, pp. 53–80, 2004.

[60] D. J. Fleet and A. D. Jepson, "Stability of phase information," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 12, pp. 1253–1268, 1993.

[61] M. Mellor and M. Brady, "Phase mutual information as a similarity measure for registration," *Medical image analysis*, vol. 9, no. 4, pp. 330–343, 2005.

[62] W. T. Freeman and E. H. Adelson, "The design and use of steerable filters," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, no. 9, pp. 891–906, 1991.

[63] J. Noble and D. Boukerroui, "Ultrasound image segmentation: A survey," *IEEE Trans. on Medical Imaging*, vol. 25, no. 8, pp. 987–1010, 2006.

[64] M. Mignotte, J. Meunier, and J. Tardif, "Endocardial boundary estimation and tracking in echocardiographic images using deformable template and markov random fields," *Pattern Analysis & Applications*, vol. 4, no. 4, pp. 256–271, 2001.

[65] A. Madabhushi and D. Metaxas, "Combining low-, high-level and empirical domain knowledge for automated segmentation of ultrasonic breast lesions," *IEEE Transactions on Medical Imaging*, vol. 22, no. 2, pp. 155 –169, 2003.

[66] J. Xu, K. Chen, X. Yang, D. Wu, and S. Zhu, "Adaptive level set method for segmentation of liver tumors in minimally invasive surgery using ultrasound images," in *The 1st International Conference on Bioinformatics and Biomedical Engineering (ICBBE)*, 2007, pp. 1091 –1094.

[67] D. Shen, Y. Zhan, and C. Davatzikos, "Segmentation of prostate boundaries from ultrasound images using statistical shape model," *IEEE Transactions on Medical Imaging*, vol. 22, no. 4, pp. 539 –551, 2003.

[68] M. Sonka, V. Hlavac, and R. Boyle, *Image processing, analysis, and machine vision*. Cengage Learning, 2014.

[69] S. K. Zhou, F. Guo, J. Park, G. Carneiro, J. Jackson, M. Brendel, C. Simopoulos, J. Otsuki, and D. Comaniciu, "A probabilistic, hierarchical, and discriminant framework for rapid and accurate detection of deformable anatomic structure," *IEEE 11th International Conference on Computer Vision*, pp. 1–8, 2007.

[70] M. J. Stoutjesdijk, J. Veltman, H. Huisman, N. Karssemeijer, J. O. Barentsz, J. G. Blickman, and C. Boetes, "Automated analysis of contrast enhancement in breast mri lesions using mean shift clustering for roi selection," *Journal of Magnetic Resonance Imaging*, vol. 26, no. 3, pp. 606–614, 2007.

[71] Z. Qian, R. Huang, D. Metaxas, and L. Axel, "A novel tag removal technique for tagged cardiac mri and its applications," *2007 4th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 364–367, 2007.

[72] T. Liu, H. Zhou, F. Lin, Y. Pang, and J. Wu, "Improving image segmentation by gradient vector flow and mean shift," *Pattern Recognition Letters*, vol. 29, no. 1, pp. 90–95, 2008.

[73] L. Massoptier and S. Casciaro, "A new fully automatic and robust algorithm for fast segmentation of liver tissue and tumors from ct scans," *European radiology*, vol. 18, no. 8, pp. 1658–1665, 2008.

[74] W. Xie, Y. Ma, and Y. Li, "Breast mass segmentation in digital mammography based on pulse coupled neural network and level set method," *SPIE Sensing Technology+ Applications*, pp. 95 010J–95 010J, 2015.

[75] C.-L. Huang, "Breast mass segmentation on breast mri using the shape-based level set method," *Biomedical Engineering: Applications, Basis and Communications*, vol. 26, no. 04, p. 1440006, 2014.

[76] Y.-L. Huang, Y.-R. Jiang, D.-R. Chen, and W. K. Moon, "Level set contouring for breast tumor in sonography," *Journal of digital imaging*, vol. 20, no. 3, pp. 238–247, 2007.

[77] D. Cremers, S. J. Osher, and S. Soatto, "Kernel density estimation and intrinsic alignment for shape priors in level set segmentation," *International Journal of Computer Vision*, vol. 69, no. 3, pp. 335–351, 2006.

[78] M. M. Hossain, K. AlMuhanna, L. Zhao, B. K. Lal, and S. Sikdar, "Semiautomatic segmentation of atherosclerotic carotid artery wall volume using 3d ultrasound imaging," *Medical physics*, vol. 42, no. 4, pp. 2029–2043, 2015.

[79] D. Grosgeorge, C. Petitjean, J. Caudron, J. Fares, and J.-N. Dacher, "Automatic cardiac ventricle segmentation in mr images: a validation study," *International journal of computer assisted radiology and surgery*, vol. 6, no. 5, pp. 573–581, 2011.

[80] V. Mihalef, R. I. Ionasec, P. Sharma, B. Georgescu, I. Voigt, M. Suehling, and D. Comaniciu, "Patient-specific modelling of whole heart anatomy, dynamics and haemodynamics from four-dimensional cardiac ct images," *Interface Focus*, vol. 1, no. 3, pp. 286–296, 2011.

[81] D. K. Iakovidis, M. A. Savelonas, S. A. Karkanis, and D. E. Maroulis, "A genetically optimized level set approach to segmentation of thyroid ultrasound images," *Applied Intelligence*, vol. 27, no. 3, pp. 193–203, 2007.

[82] K. Suzuki, R. Kohlbrenner, M. L. Epstein, A. M. Obajuluwa, J. Xu, and M. Hori, "Computer-aided measurement of liver volumes in ct by means of geodesic active contour segmentation coupled with level-set algorithms," *Medical physics*, vol. 37, no. 5, pp. 2159–2166, 2010.

[83] A. Hoogi, C. F. Beaulieu, G. M. Cunha, E. Heba, C. B. Sirlin, S. Napel, and D. L. Rubin, "Adaptive local window for level set segmentation of ct and mri liver lesions," *arXiv preprint arXiv:1606.03765*, 2016.

[84] A. Abdalbari, X. Huang, and J. Ren, "Segmentation of the liver from abdominal mr images: a level-set approach," *SPIE Medical Imaging*, pp. 94 133L–94 133L, 2015.

[85] O. Gloger, K. D. Tonnies, V. Liebscher, B. Kugelmann, R. Laqua, and H. Volzke, "Prior shape level set segmentation on multistep generated probability maps of mr datasets for fully automatic kidney parenchyma volumetry," *IEEE transactions on medical imaging*, vol. 31, no. 2, pp. 312–325, 2012.

[86] Y. Song, H. Wang, Y. Liu, C. Li, G. E. Tasian, Z. Gong, and D. Zhao, "An improved level set method for segmentation of renal parenchymal area from ultrasound images," *Journal of Medical Imaging and Health Informatics*, vol. 5, no. 7, pp. 1533–1536, 2015.

[87] M. Choong, M. Seng, S. S. Yang, A. Kiring, and K. T. K. Teo, "Foetus ultrasound medical image segmentation via variational level set algorithm," *International Conference on Intelligent Systems Modelling and Simulation*, pp. 225–229, 2012.

[88] S. Chen and R. J. Radke, "Level set segmentation with both shape and intensity priors," *International Conference on Computer Vision*, pp. 763–770, 2009.

[89] M. Alemán-Flores, L. Álvarez, and V. Caselles, "Texture-oriented anisotropic filtering and geodesic active contours in breast tumor ultrasound segmentation," *Journal of Mathematical Imaging and Vision*, vol. 28, no. 1, pp. 81–97, 2007.

[90] W. Wang, L. Zhu, J. Qin, Y.-P. Chui, B. N. Li, and P.-A. Heng, "Multiscale geodesic active contours for ultrasound image segmentation using speckle reducing anisotropic diffusion," *Optics and Lasers in Engineering*, vol. 54, pp. 105–116, 2014.

[91] M. A. Duarte, A. V. Alvarenga, C. M. Azevedo, M. J. G. Calas, A. F. Infantosi, and W. C. Pereira, "Evaluating geodesic active contours in microcalcifications segmentation on mammograms," *Computer methods and programs in biomedicine*, vol. 122, no. 3, pp. 304–315, 2015.

[92] J. Folkesson, E. Samset, R. Y. Kwong, and C.-F. Westin, "Unifying statistical classification and geodesic active regions for segmentation of cardiac mri," *IEEE Transactions on Information Technology in Biomedicine*, vol. 12, no. 3, pp. 328–334, 2008.

[93] Y. Zhou, W.-R. Shi, W. Chen, Y.-l. Chen, Y. Li, L.-W. Tan, and D.-Q. Chen, "Active contours driven by localizing region and edge-based intensity fitting energy with application to segmentation of the left ventricle in cardiac ct images," *Neurocomputing*, vol. 156, pp. 199–210, 2015.

[94] D. E. Maroulis, M. A. Savelonas, D. K. Iakovidis, S. A. Karkanis, and N. Dimitropoulos, "Variable background active contour model for computer-aided delineation of nodules in thyroid ultrasound images," *IEEE Transactions on Information Technology in Biomedicine*, vol. 11, no. 5, pp. 537–543, 2007.

[95] M. Savelonas, D. Maroulis, D. Iakovidis, S. Karkanis, and N. Dimitropoulos, "A variable background active contour model for automatic detection of thyroid nodules in ultrasound images," *IEEE International Conference on Image Processing 2005*, vol. 1, pp. I–17, 2005.

[96] M. A. Savelonas, D. K. Iakovidis, I. Legakis, and D. Maroulis, "Active contours guided by echogenicity and texture for delineation of thyroid nodules in ultrasound images," *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, no. 4, pp. 519–527, 2009.

[97] E. N. Kollorz, D. A. Hahn, R. Linke, T. W. Goecke, J. Hornegger, and T. Kuwert, "Quantification of thyroid volume using 3-d ultrasound imaging," *IEEE transactions on medical imaging*, vol. 27, no. 4, pp. 457–466, 2008.

[98] E. A. Mylona, M. A. Savelonas, and D. Maroulis, "Automated adjustment of region-based active contour parameters using local image geometry," *IEEE transactions on cybernetics*, vol. 44, no. 12, pp. 2757–2770, 2014.

[99] H. T. Huynh, I. Karademir, A. Oto, and K. Suzuki, "Liver volumetry in mri by using fast marching algorithm coupled with 3d geodesic active contour segmentation," *Computational Intelligence in Biomedical Imaging*, pp. 141–157, 2014.

[100] ——, "Computerized liver volumetry on mri by using 3d geodesic active contour segmentation," *American journal of roentgenology*, vol. 202, no. 1, p. 152, 2014.

[101] J. Oh, D. R. Martin, and X. Hu, "Partitioned edge-function-scaled region-based active contour (p-esrac): Automated liver segmentation in multiphase contrast-enhanced mri," *Medical physics*, vol. 41, no. 4, p. 041914, 2014.

[102] J. Anquez, E. D. Angelini, and I. Bloch, "Segmentation of fetal 3d ultrasound based on statistical prior and deformable model," *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 17–20, 2008.

[103] M. Elattar, E. Wiegerinck, N. Planken, H. van Assen, J. Baan Jr, H. Marquering *et al.*, "Automated normalized cut segmentation of aortic root in ct angiography," *XIII Mediterranean Conference on Medical and Biological Engineering and Computing*, pp. 1821–1824, 2014.

[104] D. Wei, C. Li, and Y. Sun, "Medical image segmentation and its application in cardiac mri," *Biomedical Image Understanding, Methods and Applications*, pp. 47–89, 2015.

[105] A. Ciurte, X. Bresson, O. Cuisenaire, N. Houhou, S. Nedevschi, J.-P. Thiran, and M. B. Cuadra, "Semi-supervised segmentation of ultrasound images based on patch representation and continuous min cut," *PloS one*, vol. 9, no. 7, p. e100972, 2014.

[106] L. Liu, W. Qin, R. Yang, C. Yu, L. Li, T. Wen, and J. Gu, "Segmentation of breast ultrasound image using graph cuts and level set," *International Conference on Biomedical Image and Signal Processing*, pp. 1–4, 2015.

[107] Z. Zhou, W. Wu, S. Wu, P.-H. Tsui, C.-C. Lin, L. Zhang, and T. Wang, "Semi-automatic breast ultrasound image segmentation based on mean shift and graph cuts," *Ultrasonic imaging*, 2014.

[108] N. Saidin, U. K. Ngah, H. A. M. Sakim, D. N. Siong, M. K. Hoe, and I. L. Shuaib, "Density based breast segmentation for mammograms using graph cut and seed based region growing techniques," *International Conference on Computer Research and Development*, pp. 246–250, 2010.

[109] B. Uzkent, M. J. Hoffman, E. Cherry, and N. Cahill, "3-d mri cardiac segmentation using graph cuts," *Image and Signal Processing Workshop*, pp. 47–51, 2014.

[110] G. Funka-Lea, Y. Boykov, C. Florin, M.-P. Jolly, R. Moreau-Gobard, R. Ramaraj, and D. Rinck, "Automatic heart isolation for ct coronary visualization using graph-cuts," *International Symposium on Biomedical Imaging: Nano to Macro*, pp. 614–617, 2006.

[111] D. Grosgeorge, C. Petitjean, J.-N. Dacher, and S. Ruan, "Graph cut segmentation with a statistical shape model in cardiac mri," *Computer Vision and Image Understanding*, vol. 117, no. 9, pp. 1027–1035, 2013.

[112] X. Alba, F. i Ventura, M. Rosa, K. Lekadir, C. Tobon-Gomez, C. Hoogendoorn, and A. F. Frangi, "Automatic cardiac lv segmentation in mri using modified graph cuts with smoothness and interslice constraints," *Magnetic resonance in medicine*, vol. 72, no. 6, pp. 1775–1784, 2014.

[113] D. Grosgeorge, C. Petitjean, and S. Ruan, "Joint segmentation of right and left cardiac ventricles using multi-label graph cut," *International Symposium on Biomedical Imaging (ISBI)*, pp. 429–432, 2014.

[114] R. Kéchichian, S. Valette, M. Sdika, and M. Desvignes, "Automatic 3d multiorgan segmentation via clustering and graph cut using spatial relations and hierarchically-registered atlases," *International MICCAI Workshop on Medical Computer Vision*, pp. 201–209, 2014.

[115] V. Fortunati, R. F. Verhaart, F. van der Lijn, W. J. Niessen, J. F. Veenland, M. M. Paulides, and T. van Walsum, "Tissue segmentation of head and neck ct images for treatment planning: a multiatlas approach combined with intensity modeling," *Medical physics*, vol. 40, no. 7, p. 071905, 2013.

[116] L. Massoptier and S. Casciaro, "Fully automatic liver segmentation through graph-cut technique," *29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 5243–5246, 2007.

[117] S. Esneault, N. Hraiech, E. Delabrousse, and J.-L. Dillenseger, "Graph cut liver segmentation for interstitial ultrasound therapy," *29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 5247–5250, 2007.

[118] A. M. Ali, A. A. Farag, and A. S. El-Baz, "Graph cuts framework for kidney segmentation with prior shape constraints," *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 384–392, 2007.

[119] M. Oda, T. Nakaoka, T. Kitasaka, K. Furukawa, K. Misawa, M. Fujiwara, and K. Mori, "Organ segmentation from 3d abdominal ct images based on atlas selection and graph cut," *International MICCAI Workshop on Computational and Clinical Challenges in Abdominal Imaging*, pp. 181–188, 2011.

[120] Y. Pauchard, T. Fitze, D. Browarnik, A. Eskandari, I. Pauchard, W. Enns-Bray, H. Pálsson, S. Sigurdsson, S. J. Ferguson, T. B. Harris *et al.*, "Interactive graph-cut segmentation for fast creation of finite element models from clinical ct data for hip fracture prediction," *Computer methods in biomechanics and biomedical engineering*, pp. 1–11, 2016.

[121] J. Huang, J. F. Griffith, D. Wang, and L. Shi, "Graph-cut-based segmentation of proximal femur from computed tomography images with shape prior," *Journal of Medical and Biological Engineering*, vol. 35, no. 5, pp. 594–607, 2015.

[122] D. Mahapatra, "Graph cut based automatic prostate segmentation using learned semantic information," *IEEE 10th International Symposium on Biomedical Imaging*, pp. 1316–1319, 2013.

[123] C. G. Looney, "Fuzzy connectivity clustering with radial basis kernel functions," *Fuzzy Sets and Systems*, vol. 160, no. 13, pp. 1868–1885, 2009.

[124] P. K. Saha, J. K. Udupa, E. F. Conant, D. P. Chakraborty, and D. Sullivan, "Breast tissue density quantification via digitized mammograms," *IEEE Transactions on Medical Imaging*, vol. 20, no. 8, pp. 792–803, 2001.

[125] J. Fleureau, M. Garreau, A. Simon, R. Hachemani, and D. Boulmier, "Assessment of global cardiac function in msct imaging using fuzzy connectedness segmentation," *Computers in Cardiology*, pp. 725–728, 2008.

[126] U. Kurkure, A. Pednekar, R. Muthupillai, S. D. Flamm, I. A. Kakadiaris *et al.*, "Localization and segmentation of left ventricle in cardiac cine-mr images," *IEEE transactions on biomedical engineering*, vol. 56, no. 5, pp. 1360–1370, 2009.

[127] Y. Zhou and J. Bai, "Multiple abdominal organ segmentation: an atlas-based fuzzy connectedness approach," *IEEE Transactions on Information Technology in Biomedicine*, vol. 11, no. 3, pp. 348–352, 2007.

[128] S. Rueda, C. L. Knight, A. T. Papageorghiou, and J. A. Noble, "Feature-based fuzzy connectedness segmentation of ultrasound images with an object completion step," *Medical image analysis*, vol. 26, no. 1, pp. 30–46, 2015.

[129] W. Zhang, J. Liu, J. Yao, and R. M. Summers, "Segmenting the thoracic, abdominal and pelvic musculature on ct scans combining atlas-based model and active contour model," *SPIE Medical Imaging*, pp. 867 008–867 008, 2013.

[130] K. Li, X. Wu, D. Z. Chen, and M. Sonka, "Optimal surface segmentation in volumetric images-a graph-theoretic approach," *IEEE transactions on pattern analysis and machine intelligence*, vol. 28, no. 1, pp. 119–134, 2006.

[131] X. Zhang, J. Tian, K. Deng, Y. Wu, and X. Li, "Automatic liver segmentation using a statistical shape model with optimal surface detection," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 10, pp. 2622–2626, 2010.

[132] X. Li, X. Chen, J. Yao, X. Zhang, and J. Tian, "Renal cortex segmentation using optimal surface search with novel graph construction," *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 387–394, 2011.

[133] H. Garg and A. Jindal, "Segmentation of thyroid gland in ultrasound image using neural network," *Proceedings of 4th International Conference on Computing, Communications and Networking Technologies*, pp. 1–5, 2013.

[134] N. Friedland and D. Adam, "Automatic ventricular cavity boundary detection from sequential ultrasound images using simulated annealing," *IEEE Transactions on Medical Imaging*, vol. 8, no. 4, pp. 344–353, 1989.

[135] P. Abolmaesumi, S. E. Salcudean, W.-H. Zhu, M. R. Sirouspour, and S. P. DiMaio, "Image-guided control of a robot for medical ultrasound," *IEEE Transactions on Robotics and Automation*, vol. 18, no. 1, pp. 11–23, 2002.

[136] J. Guerrero, S. E. Salcudean, J. Mcewen, B. Masri, S. Nicolaou *et al.*, "Real-time vessel segmentation and tracking for ultrasound imaging applications," *IEEE Transactions on Medical Imaging*, vol. 26, no. 8, pp. 1079–1090, 2007.

[137] D. C. Wang, R. Klatzky, B. Wu, G. Weller, A. R. Sampson, and G. D. Stetten, "Fully automated common carotid artery and internal jugular vein identification and tracking using b-mode ultrasound," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 6, pp. 1691–1699, 2009.

[138] K. Říha, J. Mašek, R. Burget, R. Beneš, and E. Závodná, "Novel method for localization of common carotid artery transverse section in ultrasound images using modified viola-jones detector," *Ultrasound in medicine & biology*, vol. 39, no. 10, pp. 1887–1902, 2013. [Online]. Available: splab.cz/en/research/zpracovani-medicinskych-signalu/databaze/artery

[139] P. A. Yushkevich, J. Piven, H. Cody Hazlett, R. Gimpel Smith, S. Ho, J. C. Gee, and G. Gerig, "User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability," *Neuroimage*, vol. 31, no. 3, pp. 1116–1128, 2006.

[140] R. C. Gonzalez and E. W. Richard, "Digital image processing," *Prentice Hall Press*, 2002.

[141] A. K. Jain, *Fundamentals of digital image processing.* Englewood Cliffs, NJ : Prentice Hall, 1989.

[142] I. Legakis, M. A. Savelonas, D. Maroulis, and D. K. Iakovidis, "Computer-based nodule malignancy risk assessment in thyroid ultrasound images," *International Journal of Computers and Applications*, vol. 33, no. 1, p. 29, 2011.

[143] D. Youla and H. Webb, "Image restoration by the method of convex projections: Part 1 - theory," *IEEE Transactions on Medical Imaging*, vol. 1, no. 2, pp. 81 –94, 1982.

[144] A. Hirani and T. Totsuka, "Dual domain interactive image restoration: basic algorithm," *Proceedings of the International Conference on Image Processing*, vol. 1, pp. 797–800, 1996.

[145] T. Ogawa and M. Haseyama, "Pocs-based iterative reconstruction algorithm of missing textures," *IEEE International Conference on Image Processing*, vol. 3, pp. III–101–III–104, 2007.

[146] X. Zhu, A. Ho, and P. Marziliano, "Image authentication and restoration using irregular sampling for traffic enforcement applications," *First International Conference on Innovative Computing, Information and Control*, vol. 3, pp. 62–65, 2006.

[147] P. Santago, K. Link, W. Snyder, S. Rajala, J. Worley, and Y. Han, "Restoration of cardiac magnetic resonance images," *Proceedings of Third Annual IEEE Symposium on Computer-Based Medical Systems*, pp. 60–67, 1990.

[148] R. Unnikrishnan and M. Hebert, "Measures of similarity," in *Seventh IEEE Workshops on Application of Computer Vision (WACV)*, vol. 1. IEEE, 2005, pp. 394–394.

[149] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," *Proceedings of the 8th IEEE International Conference on Computer Vision*, vol. 2, pp. 416–423, 2001.

[150] M. Meil, "Comparing clusterings: an axiomatic view," *Proceedings of the 22nd international conference on Machine learning*, pp. 577–584, 2005.

[151] J. Freixenet, X. Muñoz, D. Raba, J. Martí, and X. Cufí, "Yet another survey on image segmentation: Region and boundary information integration," *Proceedings of the European Conference on Computer Vision*, pp. 408–422, 2002.

[152] R. Unnikrishnan, C. Pantofaru, and M. Hebert, "A measure for objective evaluation of image segmentation algorithms," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops*. IEEE, 2005, pp. 34–34.

[153] ——, "Toward objective evaluation of image segmentation algorithms," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 6, pp. 929–944, 2007.

[154] Q. Yang and D. Boukerroui, "Ultrasound image segmentation using local statistics with an adaptive scale selection," in *9th IEEE International Symposium on Biomedical Imaging (ISBI)*. IEEE, 2012, pp. 1096–1099.

[155] M. Lee, Y. Chen, S. Kim, and K. Kim, "Geometric active model for lesion segmentation on breast ultrasound images," *IEEE 11th International Conference on Computer and Information Technology (CIT)*, pp. 150–157, 2011.

[156] A. Belaid, D. Boukerroui, Y. Maingourd, and J.-F. Lerallut, "Implicit active contours for ultrasound images segmentation driven by phase information and local maximum likelihood," *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pp. 630–635, 2011.

[157] D. Barbosa, T. Dietenbeck, J. Schaerer, J. D'hooge, D. Friboulet, and O. Bernard, "B-spline explicit active surfaces: An efficient framework for real-time 3-d region-based segmentation," *IEEE Transactions on Image Processing*, vol. 21, no. 1, pp. 241–251, 2012.

[158] X. Qian and B.-J. Yoon, "Contour-based hidden markov model to segment 2d ultrasound images," *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 705–708, 2011.

[159] B. Georgescu, X. S. Zhou, D. Comaniciu, and A. Gupta, "Database-guided segmentation of anatomical structures with complex appearance," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 429–436, 2005.

[160] G. Carneiro, B. Georgescu, S. Good, and D. Comaniciu, "Detection and measurement of fetal anatomies from ultrasound images using a constrained probabilistic boosting tree," *IEEE Transactions on Medical Imaging*, vol. 27, no. 9, pp. 1342–1355, 2008.

[161] G. Carneiro, F. Amat, B. Georgescu, S. Good, and D. Comaniciu, "Semantic-based indexing of fetal anatomies from 3-d ultrasound data using global/semi-local context and sequential sampling." *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2008.

[162] X. Feng, X. Shen, Q. Wang, J.-B. Kim, Z. Hao, Y. Hwang, W.-C. Bang, J. D. Kim, and J. Kim, "Learning based ensemble segmentation of anatomical structures in liver ultrasound image," *SPIE Medical Imaging*, pp. 866–947, 2013.

[163] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk, "Slic superpixels compared to state-of-the-art superpixel methods," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, pp. 2274–2282, 2012.

[164] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide-baseline stereo from maximally stable extremal regions," *Image and vision computing*, vol. 22, no. 10, pp. 761–767, 2004.

[165] B. Rahmatullah, A. T. Papageorghiou, and J. A. Noble, "Integration of local and global features for anatomical object detection in ultrasound," *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2012*, pp. 402–409, 2012.

[166] X. Liu, M. G. Linguraru, J. Yao, and R. M. Summers, "Organ pose distribution model and an map framework for automated abdominal multi-organ localization," *Medical Imaging and Augmented Reality*, pp. 393–402, 2010.

[167] T. Kohlberger, M. Sofka, J. Zhang, N. Birkbeck, J. Wetzl, J. Kaftan, J. Declerck, and S. K. Zhou, "Automatic multi-organ segmentation using learning-based segmentation and level set optimization," *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2011*, pp. 338–345, 2011.

[168] M. Wels, Y. Zheng, G. Carneiro, M. Huber, J. Hornegger, and D. Comaniciu, "Fast and robust 3-d mri brain structure segmentation," *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2009*, pp. 575–583, 2009.

[169] A. Criminisi, J. Shotton, and S. Bucciarelli, "Decision forests with long-range spatial context for organ localization in ct volumes," in *MICCAI Workshop on Probabilistic Models for Medical Image Analysis*, 2009.

[170] B. Glocker, O. Pauly, E. Konukoglu, and A. Criminisi, "Joint classification-regression forests for spatially structured multi-object segmentation," *Computer Vision–ECCV 2012*, pp. 870–881, 2012.

[171] T. Okada, M. Linguraru, M. Hori, Y. Suzuki, R. Summers, N. Tomiyama, and Y. Sato, "Multi-organ segmentation in abdominal ct images," *Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 3986–3989, 2012.

[172] R. Wolz, C. Chu, K. Misawa, M. Fujiwara, K. Mori, and D. Rueckert, "Automated abdominal multi-organ segmentation with subject-specific atlas generation," *IEEE Transactions on Medical Imaging*, vol. 32, no. 9, pp. 1723–1730, 2013.

[173] G. E. Mailloux, M. Bertrand, R. Stampfler, and S. Ethier, "Texture Analysis Of Ultraosund B-Mode Images by Segmentation," *Ultrasonic Imaging*, vol. 6, no. 3, pp. 262–277, 1984.

[174] W. Lu and J. Tan, "Segmentation of ultrasound fetal images," *Proceedings of the Society of Photo-Optical Instrumentation Engineers (SPIE)*, vol. 4203, pp. 81–90, 2000.

[175] H. Liu, C. Xie, Z. Chen, and Y. Lei, "Segmentation of ultrasound image based on morphological operation and fuzzy clustering," *Third IEEE International Workshop on Electronic Design, Test and Applications, DELTA*, p. 4 pp., 2006.

[176] W.-L. Lee, Y.-C. Chen, Y.-C. Chen, and K.-S. Hsieh, "Unsupervised segmentation of ultrasonic liver images by multiresolution fractal feature vector," *Information Sciences*, vol. 175, no. 3, pp. 177 – 199, 2005.

[177] N. Archip, R. Rohling, P. Cooperberg, H. Tahmasebpour, and S. Warfield, "Spectral clustering algorithms for ultrasound image segmentation," *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, pp. 862–869, 2005.

[178] P. R. Bai, Q. Y. Liu, L. Li, S. H. Teng, J. Li, and M. Y. Cao, "A novel region-based level set method initialized with mean shift clustering for automated medical image segmentation," *Computers in biology and medicine*, vol. 43, no. 11, pp. 1827–1832, 2013.

[179] T. Grenier, C. Revol-Muller, F. Davignon, O. Basset, and G. Gimenez, "Multiparametric smoothing based on mean shift procedure for ultrasound data segmentation," *Proceedings of the EUSIPCO05*, 2005.

[180] M. G. Linguraru, J. A. Pura, V. Pamulapati, and R. M. Summers, "Statistical 4d graphs for multi-organ abdominal segmentation from multiphase ct," *Medical image analysis*, vol. 16, no. 4, pp. 904–914, 2012.

[181] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603–619, 2002.

[182] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 888–905, 2000.

[183] A. Ciurte, N. Houhou, S. Nedevschi, A. Pica, F. L. Munier, J.-P. Thiran, X. Bresson, and M. B. Cuadra, "An efficient segmentation method for ultrasound images based on a semi-supervised approach and patch-based features," in *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. IEEE, 2011, pp. 969–972.

[184] W. Gander, G. Golub, and R. Strebel, "Least-squares fitting of circles and ellipses," *BIT Numerical Mathematics*, vol. 34, no. 4, pp. 558–578, 1994.

[185] D. Jayachandra and A. Makur, "Directional variance: A measure to find the directionality in a given image segment," *Proceedings of 2010 IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 1551–1554, 2010.

[186] S. Golemati, J. Stoitsis, E. G. Sifakis, T. Balkizas, and K. S. Nikita, "Using the hough transform to segment ultrasound images of longitudinal and transverse sections of the carotid artery," *Ultrasound in medicine & biology*, vol. 33, no. 12, pp. 1918–1932, 2007.

[187] D.-A. Huang, L.-W. Kang, Y.-C. Wang, and C.-W. Lin, "Self-learning based image decomposition with applications to single image denoising," *IEEE Transactions on Multimedia*, vol. 16, no. 1, pp. 83–93, 2014.

[188] C.-H. Yeh, L.-W. Kang, Y.-W. Chiou, C.-W. Lin, and S.-J. Fan Jiang, "Self-learning-based post-processing for image/video deblocking via sparse representation," *Journal of Visual Communication and Image Representation*, vol. 25, no. 5, pp. 891–903, 2014.

[189] A. Criminisi and J. Shotton, *Decision forests for computer vision and medical image analysis*. Springer Science & Business Media, 2013.

[190] J. Seabra and J. M. Sanches, "Ultrasound speckle/despeckle image decomposition for tissue analysis," *Ultrasound Imaging*, pp. 73–95, 2012.

[191] Å. Randestad, C.-E. Lindholm, and P. Fabian, "Dimensions of the cricoid cartilage and the trachea," *The Laryngoscope*, vol. 110, no. 11, pp. 1957–1961, 2000.