

Faculdade de Engenharia da Universidade do Porto



Universidade do Porto
Faculdade de Engenharia
FEUP

Classification Approach for Diagnosis of Arteriosclerosis using B-Mode Ultrasound Carotid Images

CATARINA DE BRITO CARVALHO

Mestrado em Engenharia Biomédica

AURÉLIO CAMPILHO
Professor Catedrático da Universidade do Porto,
Faculdade de Engenharia

RUI ROCHA
Professor Adjunto do Instituto Superior de Engenharia do Porto,
Instituto Politécnico do Porto

October 22, 2012

I would live to study, and not study to live.
Francis Bacon

Acknowledgments

It would not have been possible to write this master thesis without the help, all the willingness, the sharing of knowledge and advices given by Professor Aurélio C. Campilho, who was a real tutor and supervisor, on both an academic and a personal level. The support of my second supervisor, Professor Rui António Rocha and his availability in providing me his unpublished work was also invaluable.

My sincere thanks goes also to Dr. Elsa Azevedo from the department of neurology of the Hospital de S. João for the given images and to the INEB-Instituto Nacional de Engenharia Biomédica.

I thank my fellow labmates of Biomedical Bioimaging group, to Cátia Silva for the stimulating discussions, help and fun and to David. In particular, I am grateful to Dr. José R. Maseda for the several brainstorming sessions and the invaluable suggestions.

Last but not the least, I would like to thank my family for the unwearied support.

Abstract

According to the World Health Organization (WHO), arteriosclerosis is the third leading cause of worldwide deaths. Anatomically, this pathology is characterized by arteries' hardening or the increasing of the thickness between the intima and media arterial layers.

Arteriosclerosis is commonly diagnosed using carotid ultrasonography. The ultrasound images are very challenging because they have a significant amount of speckle and generally poor quality. The above-mentioned intima-media thickness (IMT) is often measured manually by medical experts or using some commercial software. However, the manual segmentation is very time consuming and subjective and the commercial software lack the detection of large IMT.

The main purpose of this work was the development of an automatic classification approach to improve a previous work developed within the Bioimaging Group of INEB. This method uses the instantaneous coefficient of variation as primary edge detector, a fuzzy classification stage, that returns a score map for each interface, and a dynamic programming function to construct the final contours. It is suitable to real-time processing and does not required any user interaction. The herein presented method intents to improve the fuzzy classification method (of the previous work) by replacing the heuristic approach by a classification based scheme.

We followed a pattern recognition approach that can be described by the following sequence. Firstly, the important features are computed. Then, a feature selection method is applied to reduce the dataset dimensionality, the computational cost and feature redundancy. In this method, the classification error is used as evaluation measure. The next step is the classifiers training and evaluation. Four classifiers were trained and evaluated: a quadratic, a k-nearest neighbor, a parzen and a support vector classifier. To evaluate their performance, a cross validation method was used. Two image databases were used to train and test the proposed method.

The classification result is a fuzzy score map where each pixel is labeled with its likelihood of belonging to the searched interface. These maps (one for each searched interface) are then applied to the already developed dynamic programming function and the final interfaces are obtained.

Resumo

Segundo a Organização Mundial de Saúde, a arteriosclerose é a terceira causa de morte nos países desenvolvidos. Anatomicamente, esta doença é caracterizada pelo endurecimento das paredes ou pelo aumento do espessamento entre a camada íntima e média das artérias.

Esta doença é usualmente diagnosticada recorrendo a exames de ultra-sonografia da carótida. As imagens provenientes do ultra-som são muito complicadas de classificar devido à presença de grandes quantidades de ruído associado, normalmente, com baixa qualidade da imagem. O parâmetro de avaliação desta doença é a espessura entre a camada íntima e a camada média (IMT). Esta IMT pode ser medida manualmente, por médicos, ou usando algumas aplicações comerciais. Apesar disto, a segmentação manual é muito morosa e subjectiva e as aplicações comerciais ainda não conseguem segmentar grandes espessamentos da IMT.

O principal objectivo deste trabalho é o desenvolvimento de uma abordagem de classificação completamente automática para ser usada como alternativa a uma outra desenvolvida no grupo de bio imagem do INEB. Este método usa o coeficiente de variação instantâneo como detector principal de bordos, uma etapa de classificação difusa, que retorna mapas de pontuação para cada interface, e uma função de programação dinâmica que constrói o contorno final. Esta metodologia é executada em tempo-real e não necessita de nenhuma interacção com o utilizador. O método aqui proposto pretende desenvolver uma metodologia alternativa à classificação difusa (o trabalho anterior) substituindo a abordagem heurística por uma baseada em classificação.

Seguimos uma abordagem baseada em reconhecimento de padrões que pode ser descrita pela seguinte sequência. Inicialmente, as características mais importantes são calculadas. Depois é aplicado um método de selecção de características para reduzir o tamanho dos conjuntos de dados, os tempos de computação e a redundância entre características. O erro de classificação é usado nesta altura como medida de avaliação. O passo seguinte é o treino e avaliação do classificador. Foram treinados e avaliados quatro classificadores. Estes são um classificador quadrático, um classificador de k-vizinhos mais próximos, um classificador parzen e um classificador de vectores de suporte. Para avaliar as performances destes classificadores foi utilizado um método de validação cruzada. Este procedimento foi aplicado em duas bases de imagens.

O resultado da classificação é um mapa de probabilidades onde cada pixel é etique-

tado com a sua probabilidade de pertencer à interface procurada. Estes mapas são então aplicados a uma função de programação dinâmica e as interfaces finais são obtidas.

Contents

Acknowledgments	v
Abstract	vii
Resumo	ix
Content	xiii
List of figures	xvii
List of tables	xix
Abbreviations	xxi
1. Introduction	1
1.1 Motivation	1
1.2 What is Arteriosclerosis?	2
1.3 Arteriosclerosis Diagnosis	3
1.4 System Overview	7
1.5 Objectives and Principal Contributions	8
1.6 Outline	9
2. Carotid Image Segmentation: An Overview	11
2.1 State-of-the-art	11
2.1.1 Edge Detection using Intensity and Gradient Profiles	12
2.1.2 Dynamic Programming	15
2.1.3 Active Contours	17
2.1.4 Nakagami Modeling	21
2.1.5 Hough Transform	21
2.1.6 Watershed Transform	23
2.1.7 Cubic Spline and RANSAC Methods	24

2.1.8	Classification Methods	25
2.1.9	Integrated Approach	27
2.2	Performance Evaluation	29
2.3	Summary	33
3.	From Features to Classes in US Carotid Images	35
3.1	Feature Measurements	36
3.2	Feature Analysis	43
3.2.1	Feature Redundancy	43
3.2.2	Pixels Selection of LI and MA Classes	43
3.2.3	Removing Misclassified Pixels	45
3.2.4	Dataset Resampling	46
3.3	Dimensionality Reduction	46
3.4	Classification Methods	48
3.4.1	Quadratic Bayes Classifier	48
3.4.2	Parzen Classifier	49
3.4.3	K-Nearest Neighbor Classifier	50
3.4.4	Support Vector Classifier	51
3.5	Classification Output and Dynamic Programming	53
3.6	Methodology Summary	54
4.	Experimental Results and Discussion	57
4.1	The Image Dataset	58
4.2	Dataset Resampling	59
4.3	Feature Selection	61
4.3.1	Feature Selection using Rocha's Database	62
4.3.2	Feature Selection using Molinari's Database	63
4.3.3	Feature Selection Summary	64
4.4	Classification Performance	64
4.4.1	Classifiers Parameters	66
4.4.2	Receive Operator Curve	67
4.4.3	Classifier Validation	67
4.5	Classification Output	69
4.5.1	Score Maps	69
4.5.2	Dynamic Programming	70
4.6	Statistical Analysis	74
4.6.1	Performance Measurements for Interfaces	75
4.6.2	Performance Measurements for the IMT	76
4.6.3	Bland-Altman Plots	79
4.6.4	Good and Bad Results	82

5. Conclusion	85
Appendix A. Commercial software	87
Appendix B. Differences of Three Tested Methods	89
Appendix C. Classification Results For Both Databases	91
Bibliography	93

List of Figures

1.1	Portuguese population mortality of NCDs	1
1.2	Artery structure. An anatomical and ultrasound view.	2
1.3	Johann Czermak	2
1.4	Carotid artery	3
1.5	Arteriosclerosis diagnosis exams	4
1.7	Example of ultrasound principle and acquisition of carotid ultrasound image	5
1.6	K. T. Dussik	5
1.8	Example of A-mode and B-mode scan.	6
1.9	Intravascular ultrasound	7
1.10	Basic flowchart of previously developed work	8
1.11	Basic flowchart of the work previously developed with the new approach adjustments	9
2.1	Carotid structure on a B-mode ultrasound image	12
2.2	Intensity profile of a carotid B-mode ultrasound image	13
2.3	Typical intensity profile before and after using FOAM operator	14
2.4	Molinari <i>et al.</i> ACT method	15
2.5	Loizou <i>et al.</i> method	19
2.6	Delsanto <i>et al.</i> CULEXsa method	20
2.7	Geometry of Nakagami modeling	21
2.8	Golemati <i>et al.</i> method	22
2.9	Molinari <i>et al.</i> method based in watershed transform	23
2.10	Rocha <i>et al.</i> method	25
2.11	Molinar <i>et al.</i> CALEXia method	26
2.12	Vertical distance	29
2.13	Hausdorff and Polyline distance metric	30
2.14	Example of Bland-Altman method.	32
3.1	Flowchart of the proposed method	36
3.2	Pixels' representation for the three types of classes	36

3.3	Intensity based features - intensity magnitude and maximum intensity between pixel and lumen axis.	38
3.4	Intensity based features - Gaussian windows and its displacements	39
3.5	Gradient based features - Gradient in horizontal and vertical directions and gradient orientation	40
3.6	Gradient and column profile based features	41
3.7	Specific MA interface features	43
3.8	Ground-truth images: incoherent and coherent	45
3.9	Reference pixels	45
3.10	Pixels exclusion - problem and solution	46
3.11	Sequential forward selection without and with wrapper based method	47
3.12	Two experimental datasets	48
3.13	Quadratic classifier applied to the experimental datasets	49
3.14	Parzen classifier applied to the experimental datasets	50
3.15	Voronoi tessellation	51
3.16	K-nearest neighbor classifier applied to the experimental datasets	52
3.17	SVM principle	52
3.18	SVM application to the experimental datasets	53
3.19	Classification results and final contour	54
3.20	Flowchart of the complete methodology	55
4.1	Methodology flowchart	57
4.2	Rocha's and Molinari's image examples	58
4.3	Reference pixels and LI and MA datasets structure	60
4.4	Scatter diagram for two pairs of features using the Rocha's LINW dataset	61
4.5	Classifiers ROC for Rocha's and Molinari's database	68
4.6	Score maps of Rocha's databases using different feature selection methods	70
4.7	Score maps of Molinari's databases using different feature selection methods	71
4.8	Error and geometric coefficient, λ , for the dynamic programming method using Rocha's and Molinari's database	73
4.9	Final interfaces obtained with different feature selection methods for an image of Rocha's database	74
4.10	Final interfaces obtained with different feature selection methods for an image of Molinari's database	75
4.11	Failed detections for Molinari's and Rocha's databases	76
4.12	Bland-Altman plot for the comparison of mean IMT measurements for Molinari's database using In-In method	79
4.13	Bland-Altman plots for the comparison between GT IMT measurements for Rocha's database	80

4.14	Bland-Altman plots for the comparison between automatic and GT IMT measurements for Rocha's database	81
4.15	Rocha's database good results	82
4.16	Justification for bad detection in Rocha's database	83
4.17	Molinari's database results	84
4.18	Justification of bad results for Molinari's database	84
B.1	Differences between methods for Molinari's database	89
B.2	Differences between methods for Rocha's database	90
C.1	Results for Rocha's database	91
C.2	Results for Molinari's database	92

List of Tables

2.1	Methodologies summary	34
3.1	Measured Features	44
4.1	An example of the resampling method for the LI dataset: original and re-sampled class sizes of Fig. 4.3 (a)	60
4.2	An example of the resampling method for the MA dataset: original and resampled class sizes of Fig. 4.3 (a)	61
4.3	Selected features and its ranking order using the In-In method and the wrapper approach	63
4.4	Cross-validation mean error for the trained classifiers using the different feature selection methods	65
4.5	Chosen classifiers specifications	66
4.6	Classifiers validation error	67
4.7	Geometric coefficients used by the dynamic programming function for each database	72
4.8	Performance measurements (mm) for interfaces for pixel pairs using different features selection methods	77
4.9	Performance measurements for minimum, mean, maximum and IMT measures (mm) and the coefficient of variation for the mean IMT (%) using different features selection methods	78

Abbreviations

ACT	Automatic Computer-Based Tracings
CAILRS	Carotid Artery Intima Layer Regional Segmentation
CALEXia	Carotid Layer Extraction using an Integrate Approach
CAMES	Completely Automated Multi-resolution Edge Snapper
CARES	Completely Automated Robust Edge Snapper
CAUDLES-EF	Carotid Automated Ultrasound Double Line Extraction System Using Edge Flow
CCA	Common Carotid Artery
CMUDS	Carotid Measurement Using Dual Snakes
CT	Computed Tomography
CULEXsa	Completely User-Independent Layer Extraction based on signal analysis
CV	Coefficient of Variation
CVS	Cardiovascular Disease
ECA	External Carotid Artery
FN	False Negative
FOAM	First-Order Absolute Edge Operator
FOM	Figure-of-Merit
FP	False Positive
GT	Ground-Truth
HD	Hausdorff Distance
HT	Hough Transform
ICA	Internal Carotid Artery
ICOV	Instantaneous Coefficient of Variation
IM	Intima-Media
IMT	Intima-Media Thickness
IVUS	Intravascular Ultrasound
KNNC	K-Nearest Neighbor Classifier
LI	Lumen-Intima
LIFW	Lumen-Intima Far Wall
LINW	Lumen-Intima Near Wall
MA	Media-Adventitia
MAD	Mean Absolute Distance
MAFW	Media-Adventitia Far wall
MANW	Media-Adventitia Near wall
MR	Magnetic Resonance
MSC	Mean Shift Classifier
MSE	Mean Square Error
NA	Not-Applicable
ParzenC	Parzen Classifier
PDM	Polyline Distance Metric
PRED	Pattern Recognition Edge Detection algorithm
PT	Proposed Traces
QDC	Quadratic Discriminant Classifier

RANSAC	Random Sample Consensus algorithm
RF	Radio-Frequency
ROC	Receive Operator Curve
ROI	Region of Interest
SDL	Signed Distance to Lumen
SVC	Support Vector Classifier
TN	True Negative
TP	True Positive
VD	Vertical Distance
WHO	World Health Organization

Introduction

1.1 Motivation

According to the World Health Organization [WHO, 2012], the cardiovascular diseases (CVDs) were the leading cause of deaths in the so called western countries, in 2010. In the case of the Portuguese population, Fig. 1.1 shows that the CVDs are responsible for 37% of all deaths caused by noncommunicable diseases (NCD). A NCD is a medical condition or disease which is non-infectious and non-transmissible between persons. WHO estimates that 86% of all deaths in Portuguese population are caused by NCDs.

CVDs are disorders of the heart and blood vessels and include the cerebrovascular diseases. Arteriosclerosis is the third leading cause of death, among the CVDs.

This large incidence increased the interest of the scientific community in the develop-

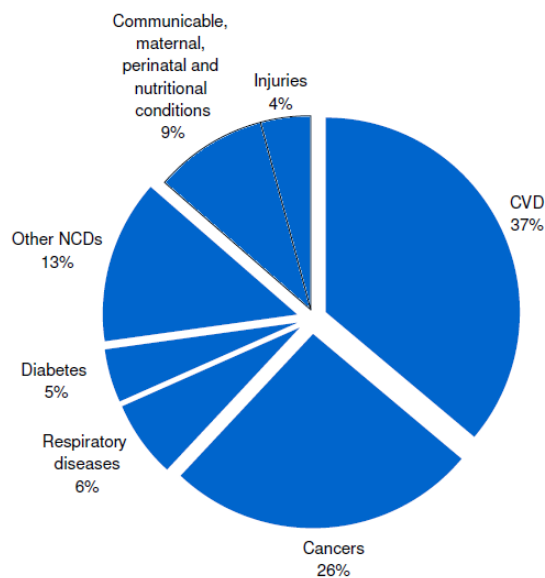


Fig. 1.1: Portuguese population mortality of NCDs in 2010, according to WHO [2012].

ment of better diagnosis techniques, for improving this disease detection at early stages, reducing the diagnosis subjectivity and making the analysis of large images databases easier.

1.2 What is Arteriosclerosis?

Arteries are muscular and elastic tubes that transport blood and its components. These are composed by three layers (tunicas): the intima, the media and the adventitia. The lumen is the artery interior through which the blood flows. Figure 1.2 (a) presents an anatomical view and Fig. 1.2 (b) presents an ultrasound view of the artery structures.

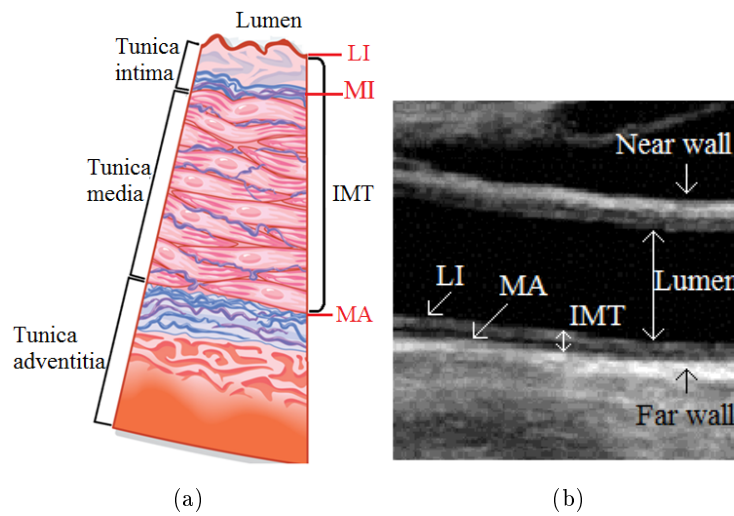


Fig. 1.2: Artery structure. An anatomical and ultrasound view: (a) artery anatomy [Art. Encyclopædia Britannica Online, 2012]; the LI is the lumen-intima interface, MA is the media-adventitia interface, MI is the media-intima interface, IMT is the intima-media thickness; (b) an ultrasound view of artery structures; the near wall is the wall nearest to the ultrasound probe and the far wall is the farthest one.

The most relevant arteries interfaces mentioned in this report are the lumen-intima (LI), the intima-media (IM) and the media-adventitia (MA). The LI is the interface between the lumen and the intima, the IM is the interface between the intima and the media and the MA is the interface between the media and the adventitia (Fig. 1.2). The distance between the LI and the MA is called the intima-media thickness (IMT).

Arteriosclerosis was not diagnosed until 1852, when Johann Czermak (Fig. 1.3) autopsied an ancient Egyptian. Later, Sir Marc Armand Ruffer also identified evidences of this disease in 3000-year-old Egyptian mummies [Allam et al., 2011].



Fig. 1.3: Johann Czermak. [Library, 2012].

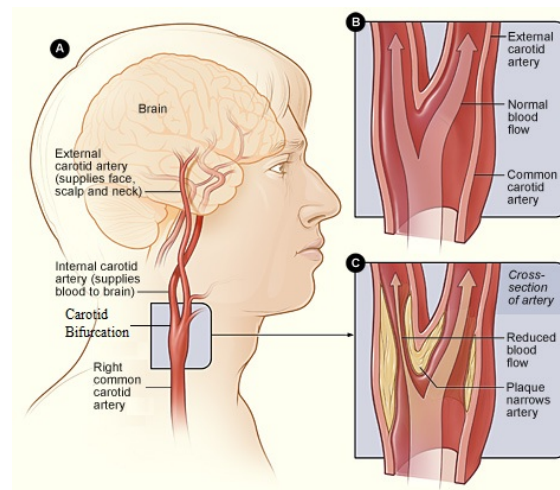


Fig. 1.4: Carotid artery: A- location of the right carotid artery in the head and neck; B- detailed image of a normal vessel; C- vessel with atheroma. Adapted from Vascular and Endovascular Surgery [July].

Marchand was the first to introduce the term “atherosclerosis” and described it as an association of fatty degeneration and vessel stiffening [Anitschkow et al., 1933].

The degenerative changes responsible for the loss of arteries elasticity are called arteriosclerosis. Its incidence is greater with increasing age. This should not be confused with atherosclerosis, which is a specific form of arteriosclerosis and is caused by the deposition of fats, cholesterol, calcium and other substances in the intima layer. It is characterized by a thickening of the intima layer and by chemical changes of the media layer. The fat is progressively accumulated and produces a protrusion towards the lumen [Seeley et al., 2012]. Depending on the size of this protrusion, different stages of the disease can be achieved. When the IMT is lower than 1mm, the artery is at a normal state, when this thickness is between 1 and 1.3mm, it is considered as artery thickening and when the IMT is higher than 1.3mm, it means that there is an atheroma. The latest state is the artery asymptomatic severe stenosis ($>50\%$ of the lumen diameter) or occlusion. This last stage normally requires a carotid endarterectomy to prevent stroke or aneurism formation [Kumar et al., 2010].

1.3 Arteriosclerosis Diagnosis

There are several medical imaging modalities commonly used to diagnose arteriosclerosis: angiography (using X-ray or magnetic resonance (MR)), computed tomography (CT), echocardiography and ultrasonography [Schmidt, 2007].

Angiography is an X-ray imaging technique used to visualize the interior of blood vessels or organs by injecting a dye. It has some disadvantages like the use of radiation, the probability of allergic reaction to the dye and the possibility of damaging the blood

vessels with the catheter insertion [Schmidt, 2007]. The MR can also be used to perform angiography exams. This technique is expensive and it can cause allergic reaction to dye or vessel rupture. Figure 1.5 (a) shows an X-ray angiography and Fig. 1.5 (b) shows MR angiography (the arrows point to the carotid arteries).

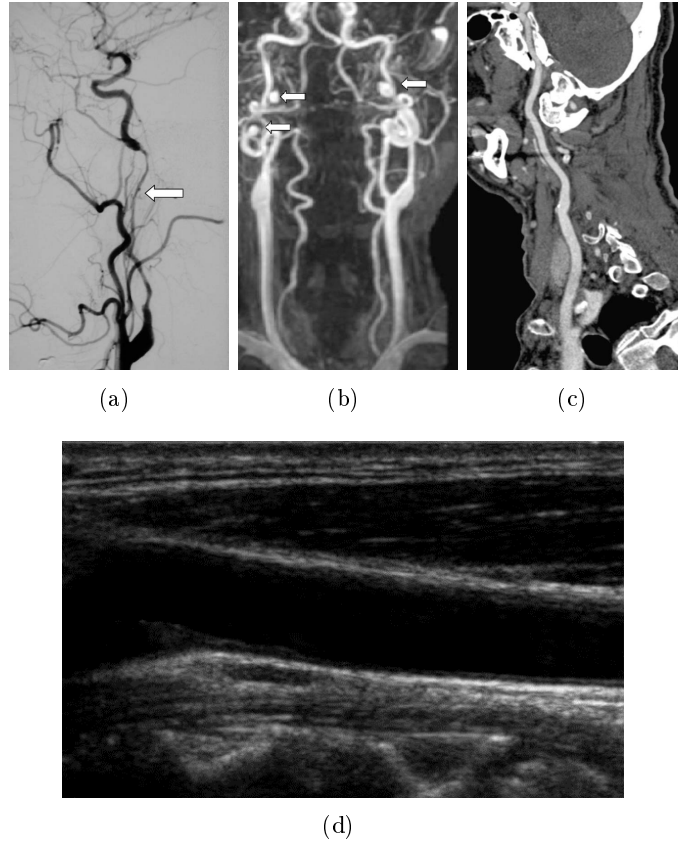


Fig. 1.5: Arteriosclerosis diagnosis exams: (a) conventional angiography; (b) MR angiography; (c) CT; (d) ultrasonography [Schaller, 2007].

In addition, arteriosclerosis diagnosis can also be done using CT imaging (presented in Fig. 1.5 (c)). This technique uses X-Ray energy and the final result is a set of slices of body sections or a 3D reconstruction of these slices [Schmidt, 2007]. Although this is not associated with the risk of vessel rupture or allergic reaction, the use of radiation can still be harmful.

At last, ultrasonography is one of the most common techniques used for arteriosclerosis detection because it is less expensive and faster than the other techniques. It is precise, comfortable and does not bring any risks related to radiation, the catheter insertion or allergic reaction. Historically, the first practical realization of ultrasound imaging appeared during World War I for submarine detection. These attempts were then applied to industrial applications for metals testing. The Austrian K. T. Dussik (Fig. 1.6), the so-called “father of diagnostic ultrasound”, was the first to apply ultrasound for medical diagnosis, in 1941. He tried to detect the brain ventricles using echotransmission [Roelandt, 2008].

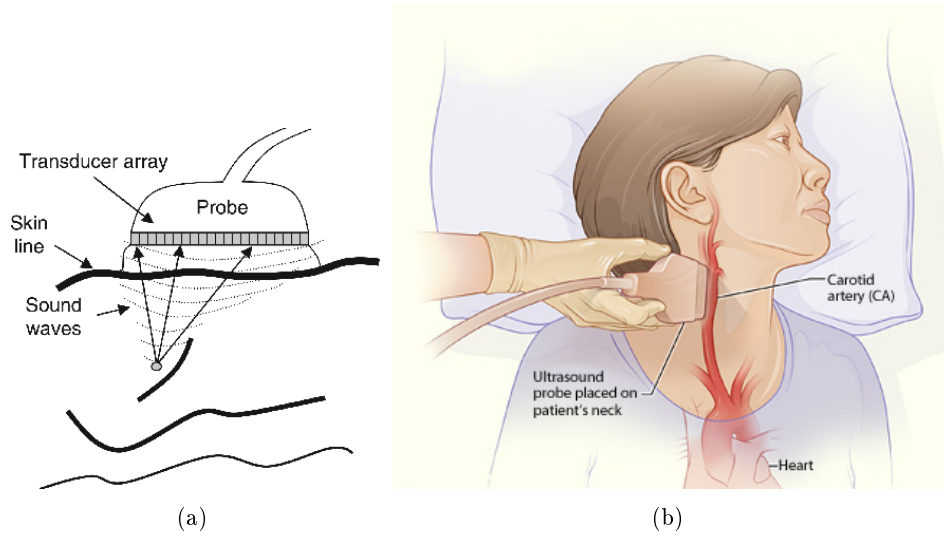


Fig. 1.7: Example of ultrasound principle and acquisition of carotid ultrasound image: (a) ultrasound imaging technique; (b) acquisition of carotid ultrasound image. Adapted from Vascular and Endovascular Surgery [2012].

A major step forward was the introduction in 1968 of electronic beam steering using phased-array technology. Since the mid-1970s, electronic scanners have been available from many companies and many ultrasound features have been improved [Suetens, 2002].

This technique represents an image of the biological tissue by transmitting sound waves focused beams into the body and then receiving the reflected echo (Fig. 1.7). In theory, a wave motion results from the periodic medium-dependent propagation of the particles vibration around their resting positions. Sound waves are pressure waves that spread by alternately compressing and decompressing the medium they are traveling in. The speed of the sound wave propagation is a function of compression and density of the medium. As a result, the frequencies used in vascular ultrasonography usually have a range from 2–10 MHz. With this technique, it is possible to capture tissue structures in images like the one shown in Fig. 1.5 (d) [Schäberle, 2004].

An ultrasound image system is typically composed by a workstation and a probe (Fig. 1.7 (a)). The probes have an array of transducers coupled to the probe body. The transducer elements are responsible for the emission and reception of the acoustic waves, that propagate into the body, where it encounters reflecting surfaces and small scatterers (Fig. 1.7 (a)). The way that these objects reflect or scatter the wave depends on the layer acoustic impedance and a proportional sound level is returned to the transducer. The time lapse between the sound wave transmission and its reception is defined as the pulse-echo cycle.



Fig. 1.6: K. T. Dussik. [Woo, 2012].

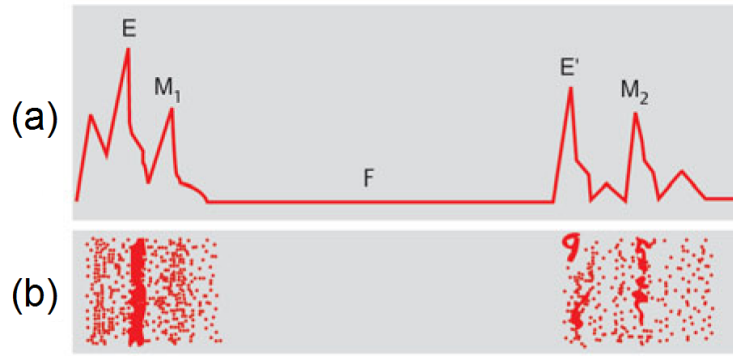


Fig. 1.8: Example of A-mode and B-mode scans: (a) A-mode; (b) B-mode display: echo amplitudes are converted to spots of varying brightness. E and E' are hyperechogenic structures, M_1 and M_2 are structures hypoechoic and F is a structure with no echo [Schmidt, 2007].

This enables the measurement of the distance from the echo source to the receptor. The transducer then converts the acoustic wave to an electrical signal that can be amplified, processed, stored and displayed [Iniewski, 2009, Prince and Links, 2006].

The carotid ultrasound image acquisition is done with the patient in the supine position with the head turned slightly to the opposite side, as shown in Fig. 1.7 (b).

There are three different ultrasound imaging modes: the A-mode, the B-mode and the M-mode. The A-mode returns a signal of the amplitude of tissue echo (Fig. 1.8 (a)). Structures with high density and that are hyperechogenic are characterized by high echo peaks. The principle of B-mode imaging is the conversion of the reflected ultrasound peaks on spots of brightness proportional to their intensity, as shown in Fig. 1.8 (b). A large echo peak of A-mode represents a large quantity of bright spots in B-mode. This means that structures with high density and echogenicity are brighter than hypoechoic structures. At last, the M-mode scanning generates a time-motion trace that records the motion of acoustic reflectors such as heart valves and myocardial walls over time.

Ultrasonography can also be performed using Doppler techniques. These can be divided in continuous-wave (CW) Doppler, pulsed Doppler, duplex sonography, power Doppler and spectral Doppler. The essential difference between basic ultrasonography and sonography using Doppler is that the last technique analyzes frequency changes that result from the relative distance between the transmitter and the receiver. Besides these, ultrasound images can also result from color-coded duplex ultrasound, which combines the presentation of two-dimensional morphologic information with superimposed flow data of a defined area displayed in color. An important modality of the color-coded duplex ultrasound technique is the intravascular ultrasound (IVUS), represented in Fig. 1.9. This is commonly used in arteriosclerosis diagnosis and it is a catheter-based imaging modality. The 360° IVUS view shows the three layers of the arterial wall architecture [Schäberle, 2004].

From all the techniques mentioned above, B-mode ultrasonography is one of the most

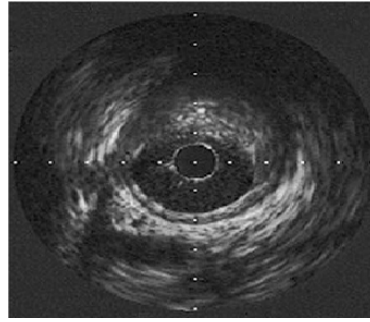


Fig. 1.9: Intravascular ultrasound [Schaller, 2007].

commonly used to diagnose arteriosclerosis because it is one of the most basic ultrasound method, is easy to use, it does not require a high computational effort and it is not invasive. However, it entails a poor image quality, caused by the speckle noise, the low contrast level and the presence of “shadows”. A shadow occurs when either a strong reflector such as a gas/tissue boundary or a highly attenuating structure hides a deeper lying organ. The shadow effect is very common in the presence of calcified atheromas.

Currently, the arteriosclerosis diagnosis can be done by manual measurement of the IMT in B-mode ultrasound images which produces subjective results and is time consuming. Due to this, several methods for semi-automatic and automatic measurement of the IMT have been published. The main objective of these studies is to perform an automatic segmentation of the artery walls in order to determine the IMT. Eliminating the user dependence, these diagnosis would be less subjective and the results would be more accurate and reliable. There are already commercial systems that perform IMT measurements automatically or semi-automatically but they fail on the large plaques detection and are very sensitive to image noise. More information about these can be found in appendix A.

1.4 System Overview

This herein proposed work is based in a previous work developed by the INEB Bioimaging group [Rocha et al., 2011]. The flowchart presented in Fig. 1.10 presents an overview of this method.

The first step of the presented work is the image reading, followed by the ROI and lumen axis detection. The ROI detection is a three classes approach and the *Otsu* algorithm [Otsu, 1979] is used to distinguish each class. The lumen axis estimation is obtained by applying a triangle thresholding algorithm, developed by Zack et al. [1977], measuring the distance map between the black pixels and the white ones, computing the local maxima and constructing the final lumen axis using a dynamic programming function. The next block is the fuzzy edge map estimation. This is computed using the ICOV (instantaneous coefficient of variation). A variation of this map, the fuzzy intima score map, is then computed for the LI interface. To obtain the MA interface, the fuzzy adventitia score map

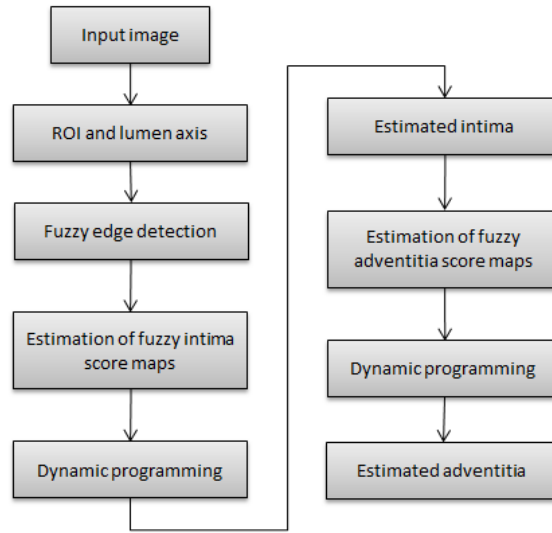


Fig. 1.10: Basic flowchart of previously work, developed by Rocha et al. [2012].

is computed using the fuzzy valley edge map and the fuzzy step edge map. A score map is obtained for each interface and the final contours are constructed by passing these maps to the dynamic programming function. This work is explained in detail in section 2.1.8.

Figure 1.11 presents the proposed approach with the replaced blocks highlighted.

The proposed method introduces an alternative way to produce the score maps which substitutes the heuristic approach of Rocha et al. [2012] by an automatic classification based scheme. This classification method intends to automatically define the discriminant features for each one of the interfaces and give to each one a classification weight. This information is embedded in a classification function that returns the score map for each interface.

The main difference in comparison with the previous method is the new pattern recognition approach, the automatic definition of the classifier and the reduction of the previous method subjectivity (related with the features).

1.5 Objectives and Principal Contributions

The main advantages of the proposed method is the capability to evaluate images with and without atheroma in horizontal and diagonal lumen vessel images achieving good results even in low quality images. This is important because the mentioned commercial software packages fail in the analysis of images with large atheroma and low quality. It can also be applied to different datasets only requiring an initial training using a representative population of the total dataset. Another advantage is the use of less heuristic variables, like in Rocha et al. [2012], the selection of the best features and the automatic definition of the classifiers parameters.

A poster related with this work was already presented during the I3S - Third Scientific

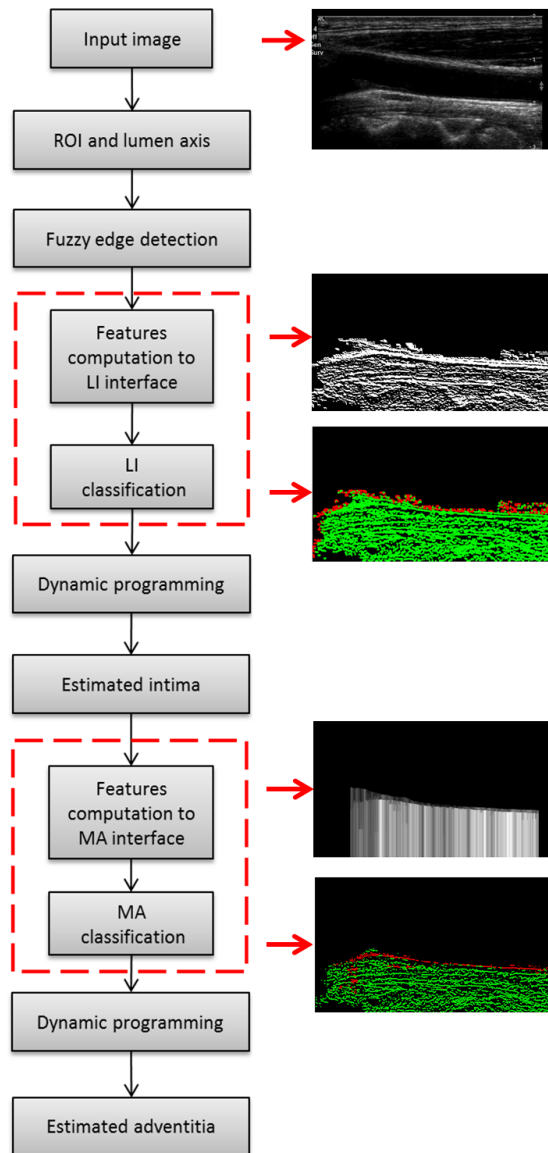


Fig. 1.11: Basic flowchart of the work previously developed by Rocha et al. [2012] with the new approach adjustments. The main contribution of the herein proposed work is the replacement of the interfaces fuzzy score maps by a classification approach. The new approach is composed of two main stages, the features computation and the interfaces classification.

Retreat, at Póvoa de Varzim in 2012 [Carvalho et al., 2012].

1.6 Outline

This dissertation is structured by chapters and according to the following sequence. The presented chapter introduces the anatomical and medical contexts and the diagnosis techniques used nowadays to diagnose the arteriosclerosis disease. Chapter 2 reports the

state-of-the-art for the IMT measurement using ultrasound carotid image segmentation. It also presents the performance methods commonly used in the evaluation of these approaches. The chapter 3 describes the computed set of features, the dataset dimensionality reduction problem and the classification stage. A brief description of the dynamic programming algorithm is given [Rocha et al., 2012]. Chapter 4 introduces the obtained results and chapter 5 presents the conclusion, the problems associated with the new methodology and the future work.

Carotid Image Segmentation: An Overview

The vessel diameter and the intima-media thickness are the two most used indicators for the diagnosis of arteriosclerosis. The segmentation and classification techniques herein presented in this chapter are focused on estimating the artery wall interfaces in order to measure the IMT.

The reference values for the IMT and for carotid diameter (in men and women) are the following [Ligouri et al., 2001]:

Carotid diameter in men:	6.52 ± 0.98 mm;
Carotid diameter in women:	5.11 ± 0.87 mm;
IMT - Normal:	$IMT \leq 1.0$ mm;
IMT - Thickening:	$1.0 \text{ mm} < IMT < 1.3 \text{ mm}$;
IMT - Plaque:	$IMT \geq 1.3$ mm.

Figure 2.1 (a) presents a typical ultrasound image of a non pathologic carotid with a normal IMT. Figure 2.1 (b) shows an ultrasound image of a pathologic carotid with plaque.

This section is organized according to the methodologies, by describing the basic principles, performance, advantages and limitations of the different methods together with a brief discussion of the performance evaluation approaches. At the end, a summary table is presented.

2.1 State-of-the-art

This section addresses the most important studies developed for the computerized IMT assessment. Several methods are described and some conclusions are drawn. The presented methods are based in: intensity and gradient profiles, dynamic programming, active contours, Nakagami modeling, Hough transform, watershed transform, cubic spline and RANSAC, classification and integrated approaches. Each of these is described below.

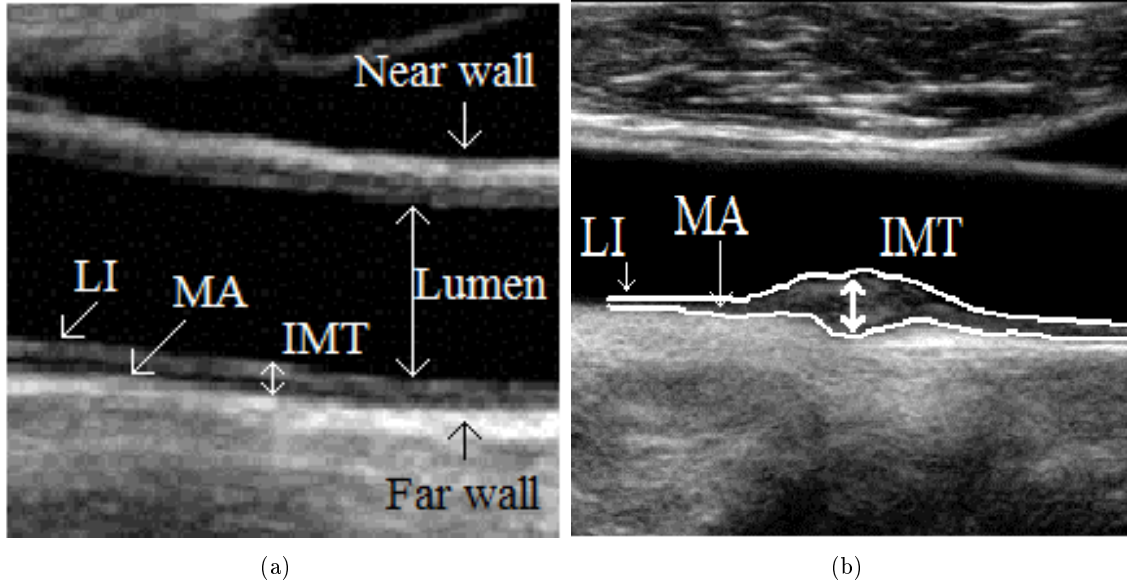


Fig. 2.1: Artery structure on a B-mode ultrasound image: (a) normal carotid; (b) pathological carotid. LI - lumen-intima interface, MA - media-adventitia interface and IMT - Intima-Media Thickness.

2.1.1 Edge Detection using Intensity and Gradient Profiles

The first approaches to perform carotid wall segmentation are based in edge detection [Pignoli and Longo, 1988, Touboul et al., 1992]. In this technique, the authors proposed the measurement of the echo peaks and they related them with the respective carotid layer (Fig. 2.2).

Pignoli and Longo [1988] were the first researchers to introduce computer methods in clinics for IMT measurement. Their approach is based on the image appearance and intensity profile. As Fig. 2.2 reports, the intensity profile has higher responses in MA interfaces (of near and far wall), medium response in the transition between the LI and the MA interfaces and a very low response in the lumen. Using these different responses, they estimated the IMT [Molinari et al., 2010d].

Touboul et al. [1992] adopted the same research structure and principles but their technique was implemented at a multi-centric clinical and on epidemiological studies [Molinari et al., 2010d].

A few years later, in 2001, Ligouri et al. [2001] proposed a segmentation technique also based in edge detection, but using image gradients instead of the intensity profile. This procedure was composed by two phases. The first phase corresponded to the selection of the Region of Interest (ROI), made by a human operator, and the second phase consisted in a pattern recognition edge detection (PRED) algorithm used to find the pixels belonging to the two required interfaces (LI and MA) of each wall. The final results were compared with a commercial software, the "PROSOUND" [Prosound, 2012], and the IMT

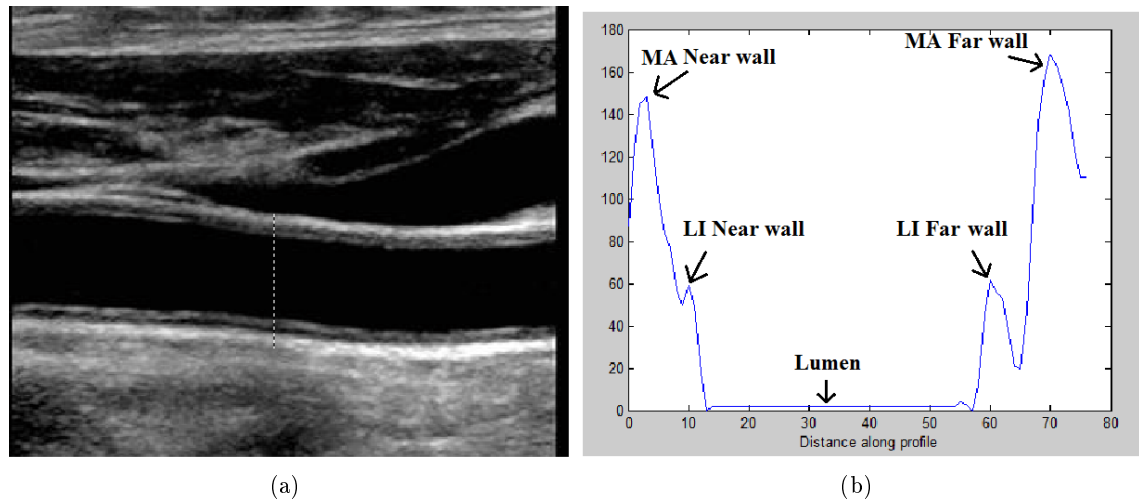


Fig. 2.2: Intensity profile of a carotid B-mode ultrasound image: (a) column selection (dashed line); (b) intensity profile of the dashed line. Each peak of the intensity profile corresponds to an interface of the artery wall. The two peaks at the left correspond to the MA and LI interfaces of the near wall (NW) and the right ones correspond to the far wall (FW) [Pignoli and Longo, 1988].

maximum difference obtained between these two systems was 0.2 mm. The strengths of this approach are the algorithm capability to work with images acquired from different hardware equipments and a low standard deviation for IMT (0.05 mm) in straight carotid images. Two disadvantages are the user-dependency and the failure on vessel curves.

In 2005, Stein et al. [2005] proposed a semi-automatic IMT border detection program and evaluated the results by analyzing the results' reproducibility and accuracy and the decreasing of the detection time. This approach segments the CCA, the carotid bulb and the ICA. A semi-automatic and a manual measurement were performed. The semi-automatic procedure determines the ROI based on the image intensity and vessel morphology. After that, the LI and MA interfaces are automatically detected based on image intensity and gradient information. The mean IMT measurement of the CCA was about $12.0 \pm 6.0 \mu\text{m}$ and the coefficient of variation (CV) between the medical expert and the semi-automatic method was of 3.2%. This approach has a time reduction of 46% in comparison with manual segmentation but has a low accuracy for the maximum IMT measurement of carotid bulb and is user-dependent.

Faita et al. [2008] described a real-time automatic technique based on a pattern recognition approach and a "First-Order Absolute Moment Edge Operator" (FOAM), that is particularly robust against the speckle. This filter enhances the image variations as it can be seen in Fig. 2.3. The image processing scheme is divided into 4 steps: filtering process using FOAM; heuristic search, which corresponds to the detection of the interfaces LI, IM and MA; outlier removal; and IMT evaluation. The IMT measurement error was $10.0 \pm 38.0 \mu\text{m}$ [Molinari et al., 2010d]. The strong points of this approach are the better

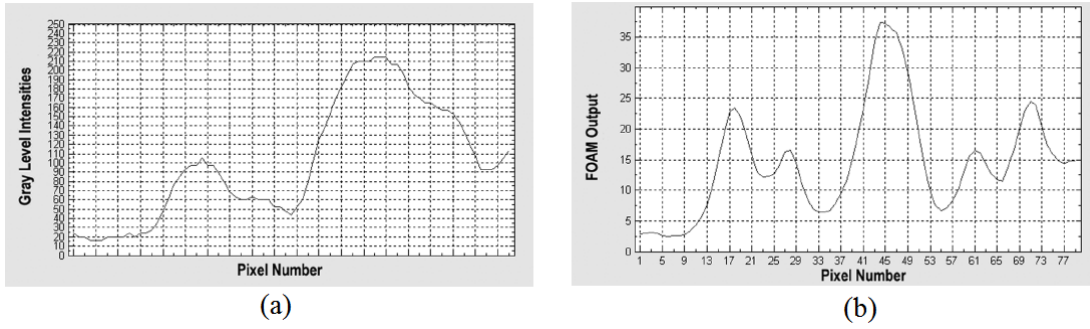


Fig. 2.3: Typical intensity profile after and before using FOAM operator: (a) typical intensity profile of a ROI column; (b) intensity profile of the same column after the filtering process with the FOAM operator [Faita et al., 2008].

signal-to-noise ratio obtained with FOAM, the less inter-observer and intra-observer variability and the real-time implementation (despite of the absence of time quantification). However, this approach has also weaknesses like the user-dependency and the impossibility to process images with plaques.

Molinari et al. [2009] presented an algorithm for the automatic computer-based tracings (ACT) of the CCA, characterized by four main features: user-independence, suitability to normal and pathological images, robustness to noise and independence of ultrasound scanners. This methodology was implemented in three steps. The first is the image cropping and ROI selection followed by the lumen estimation, using local statistics, after filtering the original image with a first order local statistic filter used to attenuate the speckle (Fig. 2.4 (a) and (b)). The last step is the adventitia layers estimation. This process is based on pixel intensity distribution along each column and the location of maxima and minima along the corresponding intensity profile (Fig. 2.4 (c)). The final detected boundaries are presented in Fig. 2.4 (d).

In order to demonstrate the independence of ultrasound equipments, this methodology was tested on 300 images, from three different ultrasound equipments. The adventitia boundaries were erroneously located in 2,7% of the cases. This confirms the robustness of the signal processing approach. The algorithm also has a good behavior on the detection of adventitia layers, even in the presence of atheroma. The total processing time was 35 seconds. The IMT measurements were not presented.

Some more recent works, proposed by Molinari et al. [2011b, 2012d], are applying new methods for edge detection. In 2011, Molinari et al. [2011b] developed a method for double line extraction using edge flow (CAUDLES-EF). This method is divided in two stages. The first one is the far adventitia estimation and the second stage is the double line border extraction. The edge flow magnitude and edge flow direction are obtained computing the intensity edge flow and the texture edge flow. At last, these two features were integrated with the flow propagation, in eight directions, and the boundaries were detected. The final stage is the removal of outliers. The IMT measurement bias error is

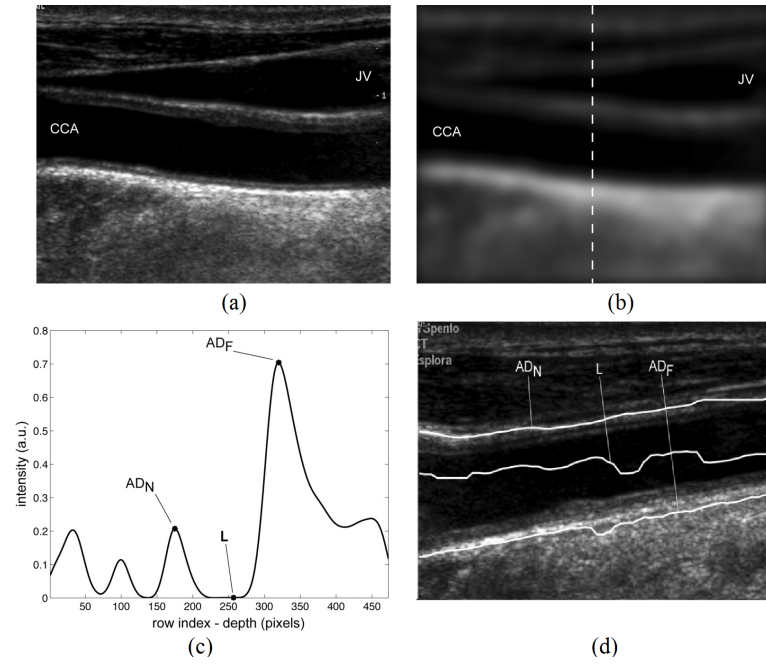


Fig. 2.4: ACT method of Molinari et al. [2009]: (a) original image; (b) image after speckle noise reduction with Gaussian filter (the white dashed line marks a column of the image); (c) intensity profile of image (b); (d) automatically detected boundaries of the ROI and lumen points plotted on the original image (AD_N = near adventitia; AD_F = far adventitia; L = lumen of the CCA).

$-43 \pm 9.3 \mu\text{m}$. This algorithm has some functional errors like the overlap of LI and MA interfaces and the inaccurate MA detection. It is also sensitive to the defined parameters like the ROI height (which discards the possibility of detection of large plaques).

Finally, in 2012, the same authors developed a completely automated multi-resolution edge snapper (CAMES) [Molinari et al., 2012d] that uses a multi-institutional database and performs a clinical validation. This method is divided in two stages: the first is the automated detection of the carotid artery based on a combination of Gaussian kernels with several known scales; the second stage consists of the automated segmentation of LI and MA interfaces using the FOAM edge detector, a gradient-of-gaussian filter followed by heuristic based peak detection and location. The IMT measurement error is $78 \pm 112 \mu\text{m}$. This method is accurate, has a low computational demand and it is very fast. It is robust to noise and suitable for several databases (with images acquired with different US probes and equipments).

2.1.2 Dynamic Programming

Dynamic programming is a method for solving optimization problems without performing an exhaustive evaluation of all the combinations between the system variables [Ballard and Brown, 1982].

It was firstly applied to the segmentation of carotid ultrasound images in 1994 by

Gustavsson et al. [1994]. Gustavsson et al. [1994] proposed an automatic technique that extracts the vessel echo intensity, edge strengths and boundary continuity from each pixel of an image previously normalized and smoothed. These features are embedded, as weight terms, into a cost function. The result of the cost function indicates the probability of the pixel being located at an artery interface. The first step is to detect MA boundaries, for the near and far walls. These boundaries will be used as starting point for the dynamic programming to detect the LI interfaces. The result of automatic IMT measurement obtained with this procedure was an inter-method mean IMT of $87.0 \pm 23.0 \mu\text{m}$ with a CV of 3.2%. One of the best aspects of this procedure is the low computational complexity and an accuracy equal to manual tracings. The main disadvantage is the impossibility of manual correction in case of wrong segmentation.

Three years later, the same research group headed by Wendelhag et al. [1997] presented another algorithm based on dynamic programming but, in this case, it includes optional interactive modifications by the human operator. This study used 500 images from CCA, carotid artery bulb and common femoral artery. The image features measured were the same used in the previous study [Gustavsson et al., 1994] and were embedded into a cost function. The weight of each feature is adjusted by training procedures. The estimated values of the three boundary features (echo intensity, intensity gradient and boundary continuity) are included as weighed terms in the cost function so that each image point is associated with a specific cost that in turn correlates with the likelihood of that point being located at the echo interface [Wendelhag et al., 1997]. The maximum IMT measured was $40.0 \pm 36.0 \mu\text{m}$. Manual correction of LI and MA were performed in 17% of all CCA images. The whole manual procedure of IMT measurement takes 45 minutes while the presented technique takes 15 to 18 minutes. This approach has some advantages like the reduction of computational time, although it is still too long, and the lesser dependency on the readers experience. One disadvantage is the frequent detection errors of the carotid artery bulb caused by the presence of plaque.

Another study based on dynamic programming was reported by Liang et al. [2000] in 2000. This is a sequel of the two previous studies [Gustavsson et al., 1994, Wendelhag et al., 1997]. The first step is to build a cost function for each boundary and define the cost terms. The definition of these cost terms are made by a training procedure which uses medical expert data. The next step is a multiscale dynamic programming, which is the implementation of the dynamic programming in different scale spaces, produced by Gaussian filters with different standard deviations. This operation allows the analysis of different boundaries because they have different responses depending on the scale. The final step is the measurement of the far wall IMT. This procedure allows human intervention to correct the detected interfaces. This correction is embedded into the cost function and monitored by a fuzzy expression to decrease the subjectiveness. The IMT measurement error was $42.0 \pm 20.0 \mu\text{m}$. As strong points the authors pointed out the reduction of subjectiveness, the robustness, the reliability due to the training process and the reduction of the overall

analysis time. The manual reading time was about 3.5 minutes while the automatic detection was 0.7 minutes. The main weakness of this technique is the inability to segment and measure images with plaques.

One of the last performed studies based in dynamic programming was presented by Holdfeldt et al. [2008] in 2008 and is an adaptation of the method proposed by Liang et al. [2000] to perform IMT assessment in images sequences. The image sequences are divided by frame and for each frame a candidate boundary pixel is defined. After this a movement model is considered using the candidates movement between frames. The method was tested using real and synthetic data. The results for the synthetic data were better than for the real data. Besides, this method has a lower error for single frames than for multiple frames which indicates that the model is not optimized to follow the boundary's movement.

Lee et al. [2010] developed a dynamic programming method using a directional Haar-like filter. This method used dynamic programming to achieve a global minimum search of image features extracted using the referred filter. The dynamic programming uses a cost function defined by an image feature term, obtained by the horizontal Haar-filter rotated between $[-45^\circ, 45^\circ]$ and a geometrical force term (boundary smoothness). The obtained boundary is the polyline that minimizes the defined cost function. The IMT detection error is between $[43-44\mu\text{m}]$. This method's advantages is the capability to perform IMT detection in sloppy carotids, the low sensibility to noise and the suitability to images with weak edges. Some disadvantages are the lack on plaques segmentation and the user-dependency.

2.1.3 Active Contours

An active contour or snake is a controlled contour that elastically adapts to the closest target [Dougherty, 2009]. The snake is controlled by an energy function and it continuously evolves in order to reduce its energy. The snake's energy function (E_{snake}) depends on the internal and external energy and is defined as:

$$E_{snake} = E_{internal} + E_{external} \quad (2.1)$$

The internal energy ($E_{internal}$) depends on the intrinsic properties of the snake, such as its length or curvature. The external energy ($E_{external}$) depends on factors such as image structure and constraints imposed by the user [Dougherty, 2009].

Gutierrez et al. [2002] was one of the first to use active contours for IMT measurement. Their approach was based on active contour and multi-resolution analysis. The operator selects a ROI that is convolved with the corresponding partial derivatives of a Gaussian filter. Based on the extracted features, a scaled artery image is used to identify the approximated positions of the near and far wall. Two complementary images are obtained based on gradient value in y-direction. One enhances pixel values transitions from high to low echoes and the other enhances pixel values transitions from low to high echoes. The

next phase is the contour modeling by the geometric deformed model. In this model, a set of vertices connected by straight line segments or edges forms the basic contour structure. The contour deformation is caused by a combination of forces which acts on the vertexes. This model is defined by three forces. The internal force, the external force and the damping force. It has a good accuracy and low coefficient of variation but it may need manual corrections. The mean IMT measurement error was $30.0 \pm 60.0 \mu\text{m}$.

Cheng et al. [1999, 2002] started the work on arteriosclerosis diagnosis back on 1999. These two methods were based on an adaptation of *Cohen's* snake. The first step is the image normalization to prevent the snake to diverge and oscillate. It requires a user selection of two lumen points used to initialize the active contour that will estimate the intima boundary. Then, the *Macleod* operator is used to enhance the interfaces. The edge intensity, the gradient, the thickness and the continuity are considered in order to avoid the noise effect. The processing time is reduced by adding a vertical external force which attracts the snake downward. This method is robust to noise and the results, according to the authors, are acceptable. The Mean-square-error (MSE) was used to evaluate this method. The results for the far wall segmentation were $62.3 \pm 60.5 \mu\text{m}$ and $38.4 \pm 68.3 \mu\text{m}$ for the LI interface and the MA interface, respectively [Molinari et al., 2010d]. The processing time is 30 seconds to 1 minute per image and the hardware used to run this algorithm was not specified.

Loizou et al. [2007] used snake to performance in carotid segmentation. The image is first normalized using the method introduced by Elatrozy et al. [1998]. In order to reduce the speckle, a linear scalar filter (*Ismv*) based on the mean and variance of the pixel neighborhood was used. After this, the initial contour location is defined. This step is very important because a bad initialization can make the snake converge to a wrong location. So, it is essential to place the initial snake contour as close as possible to the boundary of interest. This approach is organized into the following steps (Fig. 2.5): image loading, ROI selection and cropping (Fig. 2.5 (a) and (b)); despeckling of the selected area applying the *Ismv* filter (Fig. 2.5 (c)); binarization of the cropped area and dilation (Fig. 2.5 (d) and (e)); elimination of erroneous small edges that may trap the snake and labeling of connected components (Fig. 2.5 (f)); extraction of the segmented contour matrix by locating points and their coordinates on the adventitia and constructing an interpolation B-spline (Fig. 2.5 (g)); mapping the detected contour from the contour matrix in order to form the initial snake contour for the adventitia layer (Fig. 2.5 (h)); Application of snake segmentation based on the energy function proposed by Williams and Shah (Fig. 2.5 (i)).

The strengths of this methodology are the image quality and snake initialization improvement, the insignificant difference between manual and automatic IMT measurements and the small rate (5%) of bad initialization of the snake. The evaluation of the method was based on the IMT maximum, mean and minimum and their CV. The evaluation was made using two sets of manual contours traced in different periods of time. This set of measurements was compared with the one obtained by the automatic method.

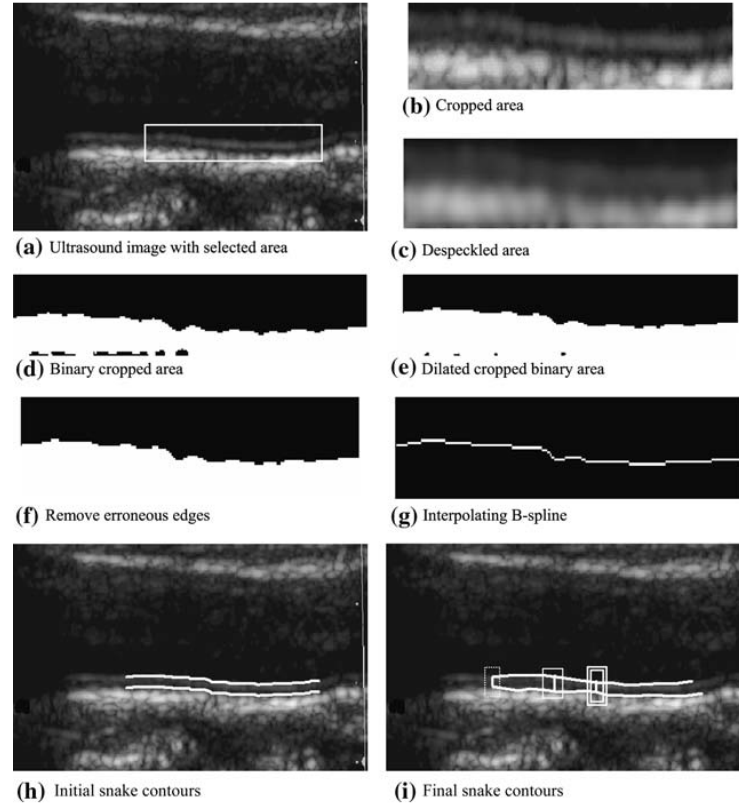


Fig. 2.5: Method developed by Loizou et al. [2007]: (a) original ultrasound image with selected area; (b) cropped area; (c) despeckled area; (d) binary cropped area; (e) dilated area after removal of small edges; (f) interpolation with B-spline; (g) detected initial contours; (h) final contour after the snake deformation.

In 2009, Loizou et al. [2009] presented another method that uses snakes for the IMT assessment. The significant difference between this method and the one presented in [Loizou et al., 2007] is that the latter performs the assessment of the IMT, the media layer thickness and the intima-media-complex in several age groups. The Williams and Shah snake was used and the first detected interfaces was the LI. Then, this interface is displaced 6 pixels downward and the upper side of the media layer interface is used to initialize the snake in order to detect the MA interface. The mean IMT error is $14 \pm 37 \mu\text{m}$.

A few years later, the same group led by Petroudi et al. [2011] developed another snake based method. This methodology uses the Chan-Vese snake after the image normalization and the speckle removal. The snake model tries to separate the image into regions based on pixel intensity and introduces a specific energy function that takes under consideration the curve length and the area inside the curve. The obtained regions are then used to initialize the snake that will detect the LI and MA interfaces. The IMT error is $0.01 \pm 0.01 \text{mm}$.

Other work that exploited the snakes potential in the carotid artery wall detection was presented by Delsanto *et al.* Delsanto et al. [2007]. This work is based on a concept

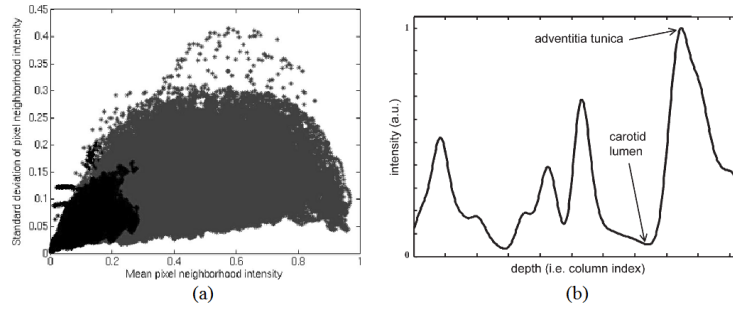


Fig. 2.6: CULEXsa method developed by Delsanto et al. [2007] : (a) distribution of normalized intensity mean and standard deviation of a 10×10 square neighborhood for each pixel belonging to a B-Mode ultrasound image of a normal CCA; (b) Intensity profile relative to a column of a B-Mode CCA image.

of completely user-independent layer extraction signal analysis (CULEXsa). CULEXsa is characterized by three phases. The first phase is the ROI identification, which is composed by image smoothing, determination of the adventitia wall using the intensity profile, and an individualization of the carotid lumen. The lumen pixels are the minimum intensity points of the neighborhood profile whose mean intensity and variance belongs to the lowest intensity pixels class (Fig. 2.6 (a)). The second phase corresponds to a gradient-based segmentation composed by image filtering with a Gaussian filter and edges enhancement using the image intensity gradients (Fig. 2.6 (b)). The last phase is the segmentation refinement through an active contour.

This algorithm does not require the ROI selection. The results are statistically comparable to the ones obtained by the human trained operator, presenting an average error of the LI and the MA interfaces segmentation lower than 1 pixel, which corresponds to an average error equal to 7%. The processing time is about 20 seconds, using Matlab and a dual core processor of 2.5 GHz. This method is affected by noise sensibility, blood scattering and image quality, which can preclude the proper ROI selection and the correct CCA segmentation.

Bastida-Jumilla et al. [2010] developed another method for the IMT assessment using snakes. This method uses anisotropic diffusion to create homogeneous and clearly separated regions inside the image followed by a geodesic active contour to detect the interfaces. This algorithm is based on thresholding of a geodesic curve at each iteration and mathematical morphological operations. The detected curves are smoothed by fitting a polynomial of low order. This method was tested only in 5 non-asymptomatic images and requires a ROI selection by the user. The mean IMT measurement error is 12.86 pixels.

Already in 2012, Molinari et al. [2012b] presented another study describing the performance of CMUDS (Carotid Measurement Using Dual Snakes). CMUDS is divided in three stages: the first is the automated carotid recognition and tracing of far adventitia wall, the second is the ROI determination followed by initialization of two snakes model (one for LI and another for MA interfaces) and the third stage is the convergence of the

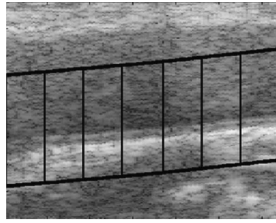


Fig. 2.7: Geometry of Nakagami modeling Molinari et al. [2010d].

final LI and MA tracings. The novelty of this method is that the snake's initialization are performed by shifting the detected adventitia far wall 3 mm upwards for LI interface and 0.1 mm for the MA interface. After the snakes' initialization, three parameters (for each interface) are computed to define the snakes energy. These are the traces continuity, the stopping force strength and the attraction term. This snake's term uses as main feature the FOAM map. Besides these, a mutual constraint force between the two snakes is defined. The CMUDS IMT error is $199 \pm 205 \mu\text{m}$. This method is very dependent on the first stage and the distance between the detected adventitia wall and the snakes' initialization point is very subjective.

2.1.4 Nakagami Modeling

The Nakagami method models the intensity of a small ROI to perform the segmentation (Fig. 2.7). This technique states that if a ROI contains the carotid wall, its intensity is characterized by a specific pattern and the presence of speckle noise [Molinari et al., 2010d]. Destrempes et al. [2009] proposed a segmentation strategy based on Nakagami mixture modeling and stochastic optimization. They considered small vertical ROIs containing the IMT complex and analyzed the radio frequency (RF) signal. The authors modeled the vertical RF signals as a mixture of three Nakagami distributions. These distributions follow the next assumptions: the lumen corresponds to the distribution with lower mean; the IMT corresponds to the mixture; the adventitia corresponds to the distribution with higher mean. The authors reported a MSE of $21 \pm 13 \mu\text{m}$ for the LI FW interface and of $0.12 \pm 7 \mu\text{m}$ for the MA FW interface [Molinari et al., 2010d].

2.1.5 Hough Transform

The Hough transform (HT) is a popular tool in applications on image segmentation because of its robustness. The original HT was designed to detect straight and curved lines and this method can be applied to images with edges already detected. An advantage of this approach is the robustness of the segmentation results because segmentation is not too sensitive to imperfect data or noise [Sonka et al., 2008].

Golemati et al. [2007] applied this image segmentation technique to the problem of carotid artery wall detection. The HT was used to identify the arterial wall in longitudinal

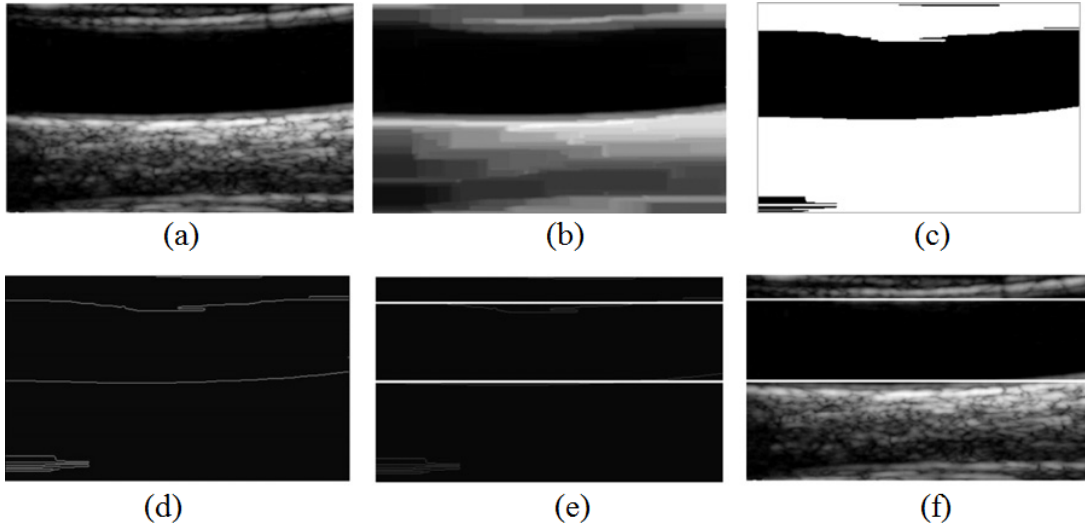


Fig. 2.8: Method developed by Golemati et al. [2007]: (a) original (cropped) image; (b) image after filtering and morphological closing; (c) image (binary) after thresholding; (d) result of edge detection in binary image; (e) result of HT on edge image; (f) result of HT on original image.

sections and lumen boundaries in transverse sections of B-mode ultrasound images. For each subject, two images were recorded, one corresponding to a longitudinal section and the other to a transverse section. The main steps of this method are: the reduction of the image area (Fig. 2.8 (a)) made by a morphological closing, followed by a thresholding to remove low intensity areas and the definition of the first and last non zero line and columns (Fig. 2.8 (b) and (c)); preprocessing, where the image was smoothed using a Gaussian filter with kernel size of 7×7 and subjected to a morphological closing; edge detection, using the *Sobel* gradient operator (Fig. 2.8 (d)); HT application, which is different for transverse or longitudinal images; and selection of dominant circle and lines (Fig. 2.8 (e)). To define IMT boundaries, two lines should be detected. The final result is presented in Fig. 2.8 (f).

The validation of this methodology was made by comparison of automatic traces against human traces. In healthy subjects the systolic IMT values were $55.0 \pm 6.0 \mu\text{m}$ and the diastolic ones were $61.0 \pm 5.0 \mu\text{m}$. This means that the IMT values change between cardiac phases. The main advantages of this method include the simplicity, the low computational cost and, according to the authors, the validity of the results. It also allows the analysis of different image projections and video records. The time required to identify two lines in both projections did not exceed 5 seconds in a sequence of 30 images. This algorithm was executed using a computer with 2.2 GHz. In longitudinal images, the accuracy was particularly high (>0.96) indicating a considerable reliability in lumen detection. As weak points there is the high dependency on image preprocessing, the low system sensitivity (<0.83) and the incapability of processing vessels with atheromas.

One year later, Stoitsis *et al.* Stoitsis et al. [2008], one of the authors of the previous

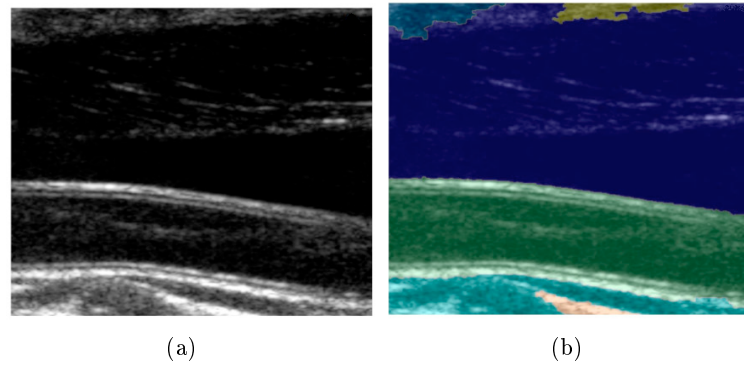


Fig. 2.9: Molinari et al. [2010c] method based in watershed transform: (a) original image; (b) image segmentation based in watershed transform.

referred work, presented an automatic methodology that allows the extraction of circles from transverse sections. In this methodology, the HT was used to initialize an active contour. The procedure is composed by a preprocessing stage, computation of a gradient vector flow field, HT application and contour estimation. This methodology has the following advantages: it is very simple and has a low computational cost; the results obtained for healthy patients were better using the HT's initialization of the active contour. The weak points were the low statistical sampling (14 images), worse results using combined methodology for plaque images in comparison with the results obtained only with the active contour. This methodology is very dependent on external variables as the parameters of the threshold, the kernel size of the smoothing filter and the morphological closing.

2.1.6 Watershed Transform

In 2010, Molinari et al. [2010c] proposed an approach for IMT assessment based on morphological operations using watershed transform to segment the image. The image was eroded, reconstructed and binarized using the Otsu's criterion in order to obtain the watershed's initialization markers.

After the region segmentation, a Canny edge operator was applied to each of the segmented regions and the lumen is detected, allowing the ROI definition. Figure 2.9(a) represents the input image and Fig. 2.9(b) represents the segmented image. The following step is the classification of the ROI pixels using a K-means clustering method, where K is the number of used clusters. This classifies the pixels in three classes: artery lumen, intima and media layers and adventitia layer. The point of transition between the lumen's cluster and intima and media cluster corresponds to the LI interface whereas the transition between the intima and media cluster and the adventitia cluster corresponds to the MA interface. The IMT error is $15 \pm 170 \mu\text{m}$. The main advantage of this method is the good lumen detection (for images with low noise) even in images with the jugular vein present. The average computational time is 18 ± 3 seconds on a dual 2.5 GHz PowerPc with 8 GB of

RAM in a Matlab framework. As it was said, the artery detection is not robust in images with high noise ratio. Besides, the method is very sensitive to the size of the erosion structural element.

2.1.7 Cubic Spline and RANSAC Methods

In 2010 and 2011, Rocha et al. [2010, 2011] reported a method for semi-automatic segmentation of carotid wall that looks for the best smooth curves according to a new gain function. This approach starts with the selection of two or three points of the lumen region, which allows the estimation of the lumen axis and the ROI selection (Fig. 2.10 (a) and (b)). Then, the edge estimation is performed by: estimating the dominant gradient direction at the edges, selecting the edges of interest that define the final edge map and, at last, determining the valley edge map. The edge detection was performed by using a new smoothing filter, developed by the author, inspired in *Tauber's* anisotropic diffusion model, which takes advantage of the instantaneous coefficient of variation (ICOV) edge detector. The edge selection only keeps the edges that are compatible with the adventitia boundaries and this selection is based on their gradient orientation, their distance to the lumen axis and their signed distance to the lumen boundaries. The valley edge map (Fig. 2.10 (e)) is viewed as a subset of the edge map (Fig. 2.10 (d)) and consist on the edges associated to profiles with two intensity peaks, being one of these usually lower. This procedure does not consider profiles with only one peak or a weak lower peak. The MA interface is detected using the random sample consensus (RANSAC) algorithm. It uses *a priori* knowledge of the distribution of inliers and sets of 5 points are used to define each spline. The best spline is selected by the algorithm. The consensus of the fitted spline is measured by a gain function that integrates the response of several discriminant features of the carotid boundaries. This discriminant features are the distance of the spline to any edge point, the distance to valley edge points, the angle between the orientation of the normal to the spline and to the intensity gradient and the signed distance to the lumen boundary (SDL) (Fig. 2.10 (c)). The estimated adventitia boundaries are presented in Fig. 2.10 (f). The LI is detected by combining dynamic programming, smooth intensity thresholding surfaces and geometric snakes.

The validation method was done against three manual tracings and the dataset used consisted of 50 longitudinal images. The detection error was similar to those observed in manual segmentation for 95% of far walls (FW) and 73% of near walls (NW). The strong points of this method are the robustness to highly degrading factors, like heavy noise, graphical marks, missing data and occlusion of the lumen region by atheroma. Also the capability of adjusting to flexible tubular shapes, similarity between manual tracings and automatic tracings, a good performance in the segmentation of far wall boundaries, which are the most used in clinical practice, and the ability to segment large plaques. This algorithm median processing time is 28.5 seconds. The segmentation of the NW is complex. This method was implemented in Matlab on a computer with an Intel Core Duo

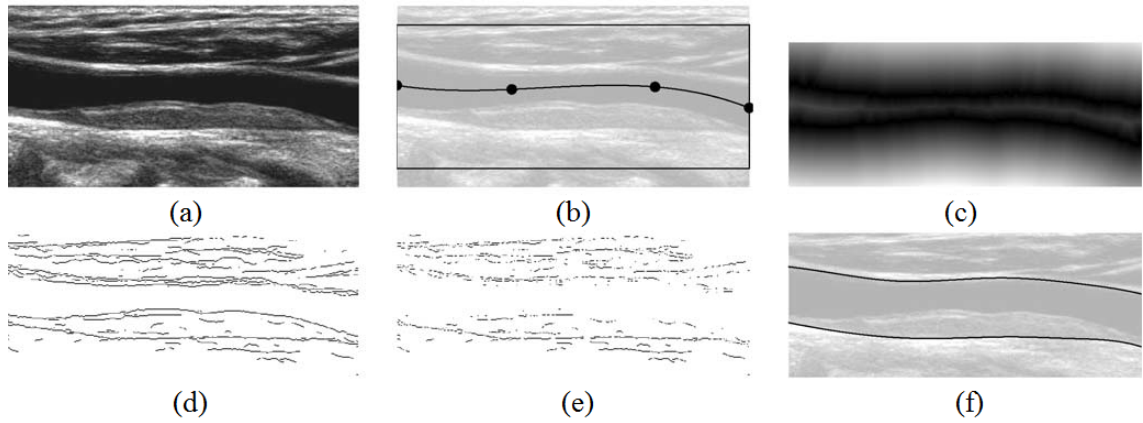


Fig. 2.10: Rocha et al. [2011] method : (a) input image; (b) points entered by the user, the interpolated lumen medial axis and the ROI; (c) absolute values of SDL; (d) edge map; (e) valley edge map; (f) Estimated adventitia boundaries.

processor at 2.13 GHz and a 2 GB RAM.

2.1.8 Classification Methods

This section presents the classification approaches that integrate some of the previously explained theoretical principles.

The first work was reported by Molinari et al. [2010a]. This work presents a methodology called carotid artery layer extraction using an integrated approach (CALEXia). This approach consists on the extraction of geometric features, line fitting and classification. The output of the algorithm is the tracings of the near and far adventitia layers. The first step is the image smoothing with a Gaussian filter and the computation of the vertical gradient column-wise. For each column, the algorithm searches for seed points, which are the high intensity local maxima surrounded by low intensity local minima. These seed points have a high probability of belonging to high intensity structures (Fig. 2.11 (a) and (b)). The next step is to connect those seed points (Fig. 2.11 (c)). To overcome the problem of false positives and over-segmentation, line segments are first combined, fitted and connected (Fig. 2.11 (c), (d) and (e)) and then a linear perceptron that combines geometric features discards all the line segments that have a high probability of including a vessel lumen, obtaining the final contour (Fig. 2.11 (f)).

The validation of this method is done using the mean of the three human tracings. The mean absolute error for the NW is 0.07 ± 0.07 mm and the far wall error is 0.17 ± 0.24 mm. The wrong detections can be originated when the CCA is parallel to the jugular vein and by deeper plaques protruding into the artery lumen. This method is very noise sensitive, it can confuse the jugular vein with the carotid artery, the smoothing process can change the location of the edges, it is unable to process images with large atheromas and it does not perform IMT measurements. The advantages of this method are the good results for

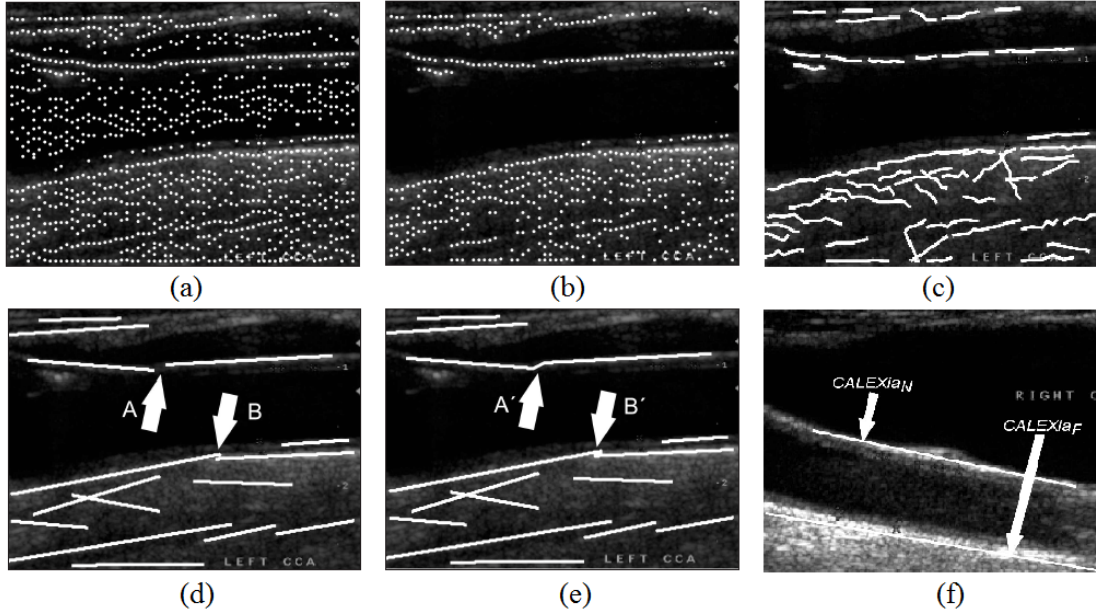


Fig. 2.11: Molinari et al. [2010a] CALEXia method: (a) detected local maxima seed points; (b) selected seed points based on the linear discriminator model; (c) line segments after fitting; (d) valid line segments; (e) connected line segments. White arrows A and B indicate a couple of valid but disconnected line segments, white arrows A' and B' show the connection of the line segments; (f) final result for healthy carotid.

normal and curved vessels and vessels with plaque. The dataset is very large (100 normal CCAs, 50 CCAs with an increased IMT, 30 with fibrous plaque and 20 anechoic plaques).

Rocha et al. [2012] presented a new IMT assessment method which uses fuzzy classification to construct the final interfaces. This method uses the ICOV edge detector as Rocha et al. [2010, 2011], an edge fuzzy classification method and a dynamic programming procedure. It is suitable for the detection of both near and far wall LI and MA interfaces. The first step is the ROI and lumen axis automatic detection. The ROI detection is performed using two Otsu's threshold [Otsu, 1979]. That separate the three pixels classes (dark pixels, corresponding to lumen and hypoechogenic tissues, bright pixels, corresponding to letters or graphical marks and medium intensity pixels, corresponding to echogenic tissue). The lumen axis detection is performed inside the already detected ROI and follows four steps. In the first one, a triangle threshold algorithm is used to set the lumen pixels to black; in the second step, the distance between dark pixels and the closest white pixel in the same image column is measured; the next step consists in the detection of the local maxima of the distances in the gradient direction; in the last step, a dynamic programming algorithm is used to delineate the lumen axis, which is the contour with the highest score. At this point, a second ROI is defined using the lumen information combined with the maximum expected carotid vertical diameter.

Then the fuzzy edge map is computed, from the ICOV image, using robust statistics

based on the *Tukey's* function. Then, the fuzzy valley edge map is also computed. It represents the edges associated with a valley shaped intensity profile. To detect the LI interface, the fuzzy edge map is passed to a score function to construct a LI score map. The final LI interface is obtained applying this score map to the dynamic programming algorithm. This algorithm is based on a gain function that combines the edge information, from the fuzzy score maps, with a geometric term. The MA interface detection uses both the fuzzy edge map and the fuzzy valley map to compute a MA score map and the search is conditioned by the already detected LI interface position. The MA final contour is also obtained using the referred dynamic programming algorithm.

The IMT error was evaluated by comparing the automatic detections with the three manual tracings performed by two experts. The mean error is 0.207 mm for NW and 0.16 mm for FW. This method was computed using an Intel Core 2 Duo processor at 2.13 GHz and the mean time required to segment an image was 2.1s. The main advantages of this methodology are its suitability to segment large plaques, its low sensibility to noise, its complete automation and the user independence.

Meiburger et al. [2011] developed a method called CAILRS (carotid artery intima layer regional segmentation). This is based on scale-space multi-resolution analysis and on the mean shift classifier (MSC), which is a non-parametric technique for the estimation of the density gradient. This method's first stage is the carotid recognition (equal to Molinari et al. [2012d, 2011b, 2010a]) and the second is the LI and MA border estimation. This is divided in the determination of the ROI, regional wall segmentation using the MSC, reconstruction of the interfaces and the final refinement checks. The MSC proposed method was firstly proposed by D and P [1997] and uses a search window of a certain radius r initialized in a chosen location. Inside this search window, the differences between the local mean and the center of the window is calculated and the resultant vector is the so called, mean shift vector. A fundamental property of this classifier is its proportionality to the gradient of the probability density at a considered point.

This method's IMT error is $-35 \pm 186 \mu\text{m}$ and its disadvantages are the high dependency on the first detected adventitia wall, the MSC only distinguishes one class (of the three image classes) in the image and the impossibility to segment large plaques.

2.1.9 Integrated Approach

The methods presented in this section integrate different techniques and methodologies, some of them already described in the previous sections.

In 2010, Molinari et al. [2010b] presented an approach that integrates two algorithms already presented, the CULEXsa [Delsanto et al., 2007] and the CALEXia [Molinari et al., 2010a]. These two automated techniques are complementary and use different foundations for segmentation. They have completely different architectures and they offer different performances on LI and MA interfaces. Therefore, the goal of this research was to calibrate IMT measurements combining CULEXsa and CALEXia segmentations of LI and

MA interfaces, by using a greedy approach.

The general concept of the greedy algorithm is an iterative technique that searches the global optimum of a problem by a series of local optimizations. This kind of algorithms are commonly used for error correction and performance optimization.

CULEXsa and CALEXia lack on noise robustness, but the noise effect is different in each one. This method starts with the representation of the automatic tracing of CULEXsa and CALEXia. Firstly it is assumed that the greedy algorithm boundary is CULEXsa tracing because this presents a lower global error, for LI boundaries, in comparison with CALEXia. After this, the first vertex of CALEXia is swapped with the closest one of CULEXsa and the system error is computed. This process is iterated for all CULEXsa points and the error is computed to all the swaps. If the error to a certain vertex for CALEXia swap is lower than the error with CULEXsa, then CULEXsa's point is substituted by the CALEXia's point. This is done to all the computed interfaces. At the end, the resultant boundaries are interpolated by a bicubic spline, to smooth out any minor oscillation in the boundary profile.

The validation of this method is done using the polyline distance metric (PDM). This metric was chosen because the authors stated that it is more robust and a reliable indicator of the distance between two given boundaries. They evaluated three features, the mean system error, the error per vertex and the IMT metrics in the greedy framework. The advantages of this methodology are the good results in the presence of noise, vessel pathologies and different carotid appearances, high accuracy and robustness and a total processing time of 4.5 ± 0.2 seconds. This method was implemented in Matlab and executed on a dual 2.5GHz PowerPC with 8MB of RAM. The mean IMT error obtained with the greedy approach was lower than the error obtained with the performance of CULEXsa and CALEXia and has a value of $83.1 \pm 61.8 \mu\text{m}$. The main disadvantages are the need to compute the errors in every system interaction and the high IMT error.

In 2011, Molinari et al. [2011c] developed an edge detection methodology, called CARES, that combines the CALEX and the FOAM methods. CALEX was used to automatically trace the profile of far adventitia wall. The CARES algorithm modified the FOAM edge operator to perform a robust heuristic search of the maxima points. The final error is $32 \pm 141 \mu\text{m}$ and the computational time is 2.3 seconds (average time on the 300 images database). The disadvantages of CARES are the low reproducibility and the impossibility of segmenting large plaques.

Molinari et al. [2011a] also presented a work where different segmentation techniques were fused in order to obtain a high-performance carotid IMT measurement. This method fuses the CALEXia, CULEXsa and the watershed transform. The best combination of the three referenced methods first combines the CULEXsa with the CALEXia and then combines the result with the watershed transform. The mean IMT error is $46 \pm 47 \mu\text{m}$. This combination decreases the sensibility to noise.

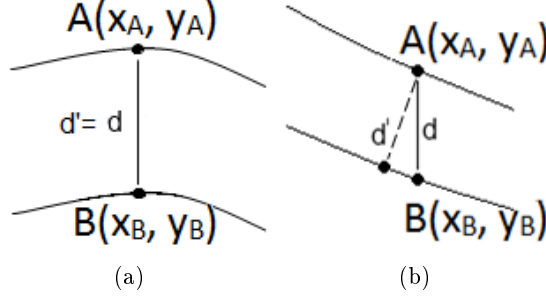


Fig. 2.12: Vertical distance: (a) between two horizontal contours; (b) between two non-horizontal contours. A and B represent the two points where the true diameter d' , is overestimated by the vertical distance, d .

2.2 Performance Evaluation

This section describes the performance measures used to evaluate the final interfaces of the presented methods. This evaluation is done using reference contours, which are arteries interfaces manually traced by medical experts. The evaluation result gives the affinity between the two compared contours.

Vertical Distance

The vertical distance (VD) is one of the easiest and commonly used distance to measure the differences between two contours. It is defined by the modulus of the difference between the ordinates of two pixels with the same abscissa. This measure is defined in eq. 2.2 and illustrated in Fig. 2.12 (a).

$$d(A, B) = |y_A - y_B| \quad (2.2)$$

Despite its common use, VD is a rough estimation of the true distance, measured by the *Euclidean* distance [Deza and Deza, 2009]. This measure can be associated with some error, like the one presented in Fig 2.12 (b). The measured vertical distance is presented by d but the real distance between the two contours is d' .

Hausdorff Distance

The Hausdorff distance (HD) measures the distance between two sets of points. It measures the farthest distance between them.

Given two finite point sets, $A = \{a_1, \dots, a_p\}$ and $B = \{b_1, \dots, b_q\}$, the Hausdorff distance is defined as:

$$HD = \max(d_{AB}, d_{BA}) \quad (2.3)$$

where $d_{AB} = \max_{a \in A} \min_{b \in B} \|a - b\|$ is the directed Hausdorff distance from A to B and d_{BA} is

the inverse. Figure 2.13 (a) illustrates a reference for the computation of d_{AB} and d_{BA} .

The directed Hausdorff distance identifies the point $a \in A$ that is farthest from any point of B and measures the distance from a to its nearest neighbor in B (using the given norm $\|\cdot\|$).

The major disadvantage of this distance measure is that only distances between vertices are measured. Therefore, this measure only have significant results when the two compared set have the same number of vertices [Molinari et al., 2010d].

Polyline Distance Metric

In 2000, Suri et al. [2000] proposed a new measurement method to evaluate automatic tracings called polyline distance metric (PDM). This method is based on the measurement of the distance of each boundary vertex to the other boundary segments. The measured distance becomes very robust because of its high degree of independence from the number of points of each boundary [Molinari et al., 2010d].

The polyline distance from the vertex v to the boundary B can be defined by equation 2.4.

$$d(v, B) = \min_{s \in B} d(v, s) \quad (2.4)$$

Considering two boundaries A and B , presented in Fig. 2.13 (b), it is possible to define the distance $d(v, s)$ between the vertex v and a segment s .

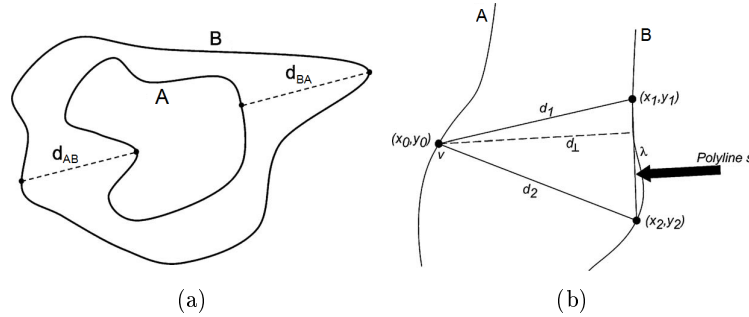


Fig. 2.13: (a) Hausdorff distance [Molinari et al., 2010d]. The distance d_{AB} is the shortest distance between the most distant point of B and the closest point of A . The distance d_{BA} is the shortest distance between the most distant point of B and the closest point of A (b) Polyline distance metric between two boundaries A and B [Molinari et al., 2010d].

The polyline distance between two boundaries (A and B) is defined as:

$$D(A, B) = \frac{d(A, B) + d(B, A)}{(\text{number of vertex of } A + \text{number of vertex of } B)} \quad (2.5)$$

Mean Absolute Distance

The mean absolute distance (MAD) is one of the most commonly used evaluation measures. It gives the mean absolute differences between the proposed traces (PT) and a ground truth (GT). This is defined as:

$$MAD = \frac{1}{N} \sum_{y=1}^N |PT(y) - GT(y)| \quad (2.6)$$

where N is the number of points of each boundary and y is the ordinates of two contours with the same abscissa. It is very simple to implement and it returns acceptable results [Molinari et al., 2010d].

This measure, like the VD, can be misused if the artery is not horizontal because the distance measured will not correspond to the true artery diameter. In this case, the measure overestimates the diameter.

Mean Square Error

The mean square error (MSE) goal is to compare two contours by providing a quantitative score that describes the degree of similarity/fidelity or, conversely, the level of error/distortion between them [Wang and Bovik, 2009].

Suppose that PT_i with $i = 1, 2, \dots, N$; and GT_i with $i = 1, 2, \dots, N$; are two finite-length contours, where N is the number of contour pixels. The MSE between the contours is:

$$MSE(PT, GT) = \frac{1}{N} \sum_{i=1}^N (PT_i - GT_i)^2 \quad (2.7)$$

Coefficient of Variation

The coefficient of variation (CV) measures the variability between two estimates of the same measure, for instance, IMT mean. It can be used to assess the inter and intra variability between two different methods/operators [Gustavsson et al., 1994]. CV is defined as:

$$CV = \frac{\sigma}{(\sqrt{2}\mu)} \times 100\% \quad (2.8)$$

where μ is the pooled mean and σ is the standard deviation for the differences between the manual and automatic tracings.

Figure-of-Merit

A figure-of-merit (FOM) is a performance measure that establishes the ratio between the correct responses and the total ones. The FOM is used by some of the presented methods. It measures the ratio of the difference between the ground-truth and the automatic method. Equation 2.2 represents the used FOM evaluation criterion, where *Method* represents the method under evaluation.

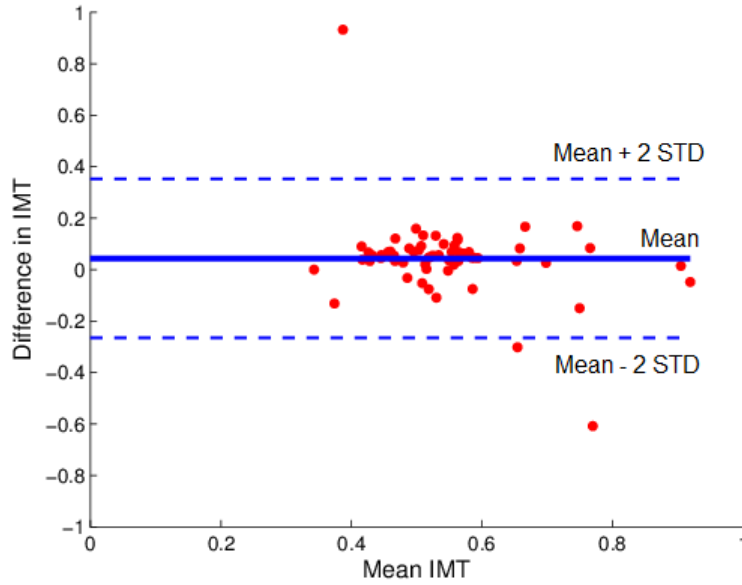


Fig. 2.14: Example of Bland-Altman method. The filed line represents the mean difference of IMT values and the dashed lines represent the $\pm 2 \times$ standard deviation.

$$FOM = \left(1 - \frac{|GT_{IMT} - Method_{IMT}|}{GT_{IMT}} \right) \times 100$$

Bland-Altman Method

The Bland-Altman method consists in representing the difference between two estimates of the same measure as a function of the corresponding means. The limits of agreement can be defined as $\mu \pm 2\sigma$, where σ is standard deviation of the estimates and μ is the mean [Bland and Altman, 1986]. Figure 2.14 illustrates a typical Bland-Altman plot. The filed line represents the mean value (μ) of the difference between the estimated IMT and IMT of reference and the dashed lines represent the $\mu \pm 2\sigma$ for the differences between the estimated IMT and IMT of reference.

Classification Evaluation

There are multiple performance measures that can be used to evaluate the performance of a classification algorithm.

Detection and classification performance metrics use different parameters, namely: true positive (TP), which corresponds to the true lesion areas that are detected by the algorithm; false positive (FP), that corresponds to non-lesion areas that are detected by the algorithms as being real lesions; true negative (TN), which corresponds to non-lesion areas that were correctly classified; and false negative (FN), that represents lesions that were classified incorrectly [Mendonça, 2011].

Based on these parameters, the following evaluation measures are used:

- Accuracy = $(TP+TN)/(TP+TN+FP+FN)$. Which is a measure of the global performance of the algorithm in what concerns correct decisions.
- Sensitivity = $TP/(TP+FN)$. Refers to how often the algorithm reports that an abnormality exists in the instances where it actually exists.
- Specificity = $TN/(TN+FP)$. Refers to how frequently the algorithm correctly reports normal when no abnormality exists.
- Precision = $TP/(TP+FP)$. Precision is the fraction of detections that are relevant, which means that quantifies the correct classification from all the lesions detected.

A Receive Operator Curve (ROC) summarizes the space of possible trade offs between sensitivity and specificity [Mendonça, 2011]. It is used to make comparisons of observer performance between two different observation conditions or parameters [Costardou, 2005].

2.3 Summary

Table 2.1 summarizes the presented techniques and evaluation methods. This table presents the authors and methodologies used, the year of publication, the evaluation method, some results, the computational cost and the number of images used. It also indicates the images with plaques and if the method is automatic.

The best method is the Meiburger et al. [2011]. The IMT error of this method is $-16 \pm 258 \mu m$. This method uses 300 images, and some of them with plaques. This method was evaluated using the polyline distance metric and the FOM. It takes 1.87 s to classify an image.

This table also shows that the most recent methods have been focusing in classification and integrated approaches. However, images with plaques are not yet commonly processed.

Tab. 2.1: Methodologies summary

Author	Method	Evaluation criteria	IMT error (μm)	Process. time	Num. of images	Plaque presence?	Auto?
Pignoli and Longo [1988]	ED	-	-	-	-	-	no
Touboul et al. [1992]	ED	MAD	-	-	-	yes	no
Ligouri et al. [2001]	ED	MAD	-	-	133	yes	no
Stein et al. [2005]	ED	MAD+CV	-	5.5 \pm 0.5 min	300	yes	no
Faita et al. [2008]	ED	MAD	-	-	150	no	no
Molinari et al. [2011b]	ED	PDM+FOM	-43 \pm 9.0	-	300	no	yes
Molinari et al. [2012d]	ED	PDM+FOM	78 \pm 112	15s	300	no	yes
Gustavsson et al. [1994]	DP	MAD+CV	-	-	121	no	yes
Wendelhag et al. [1997]	DP	MAD	-	-	500	no	yes
Liang et al. [2000]	DP	MAD+CV	-	42s	458	yes	yes
Lee et al. [2010]	DP	MAD	[43 44]	-	75	no	yes
Gutierrez et al. [2002]	AC	MAD+CV	30.0 \pm 60	-	180	no	yes
Cheng et al. [1999, 2002]	AC	MSE	-	30-60s	32	yes	no
Loizou et al. [2007]	AC	MAD+HD+CV	-	-	100	no	no
Delsanto et al. [2007]	AC	MAD	-	-	120	yes	yes
Loizou et al. [2009]	AC	MAD+CV	-	-	-	no	yes
Molinari et al. [2009]	AC	PDM	-	-	300	yes	yes
Bastida-Jumilla et al. [2010]	AC	MAD	-	-	5	no	yes
Petroudi et al. [2011]	AC	MAD	90 \pm 100	-	30	no	yes
Molinari et al. [2012b]	AC	PDM+FOM	30 \pm 284	-	665	no	yes
Molinari et al. [2010c]	WT	PDM	150 \pm 170	18 \pm 3s	200	no	yes
Destrepes et al. [2009]	NM	MAD+HD	-	-	30	-	no
Golemati et al. [2007]	HT	MAD	-	5s**	10 sequences	yes	yes
Stoitsis et al. [2008]	HT	MAD	-	-	10 sequences	yes	yes
Rocha et al. [2011]	CS	MAD+CV	-	28.5s	50	yes	no
Molinari et al. [2010a]	C	MAD	-	-	200	no	yes
Meiburger et al. [2011]	C	PDM+FOM	-16 \pm 258	1.87s	300	no	yes
Rocha et al. [2012]	C	MAD+CV	NW:190 FW: 150	2.1s -	50 -	yes -	yes -
Molinari et al. [2010b]	IA	PDM	-	4.5 \pm 0.2s	200	no	yes
Molinari et al. [2011c]	IA	PDM+FOM	32 \pm 141	2.3s	300	no	yes
Molinari et al. [2011a]	IA	PDM	46.3 \pm 46.9 μm	117.2 \pm 1.6s	200	no	yes

ED - Edge Detection; DP - Dynamic Programming; AC - Active Contour; WT - Watershed; NM - Nakagami Modeling;

HT- Hough Transform; CS - Cubic Spline; C- Classification; IA - Integrated Approach.

MAD - Mean Absolute Distance; PDM - Polyline Distance Metric; HD - Hausdorff Distance; MSE - Mean Square Error.

FOM - Figure of Merit.

From Features to Classes in US Carotid Images

Pattern Recognition is the background strategy used to develop the method for layer labeling in US carotid images. A basic explanation of the involved concepts from a theoretical and a practical points of view are given.

To understand Pattern Recognition it is necessary to understand the concept "pattern". According to Bow [2002], a "pattern" can be defined as a quantitative or structural description of an object or some other entity of interest.

Pattern recognition is a process of categorizing any sample of measured or observed data as a member of one of several classes. Machine vision intends to discover and reproduce the human mechanism of object recognition [Duda et al., 2000]. The combination of pattern recognition with machine vision, when accurate and reliable, is an useful and challenging instrument in several applications like the IMT measurement.

Figure 3.1 gives an overview of the methodology for labeling the carotid interfaces. As this figure shows, this method is composed of three main stages: the feature computation stage, the interface classification and the interface detection.

The feature measurements and selection stage (for LI and MA interfaces) is composed of the feature measurements, feature analysis and the dimensionality reduction. Firstly, the features for each interface are computed, allowing the discrimination of the classes. Then, the feature analysis is performed and some consideration of the used ground-truths is done. At last, the dimensionality reduction is performed to reduce the computational cost and the feature redundancy. The next stage is the interface classification (for LI and MA interfaces) which uses the previously computed features to create a score map for each interface. Each pixel of this map is labeled with its likelihood of belonging to the searched interface. In the third stage, the interfaces are detected by applying the interfaces score maps to a dynamic programming function [Rocha et al., 2012].

This chapter is organized in six sections: feature measurements, feature analysis, dimen-

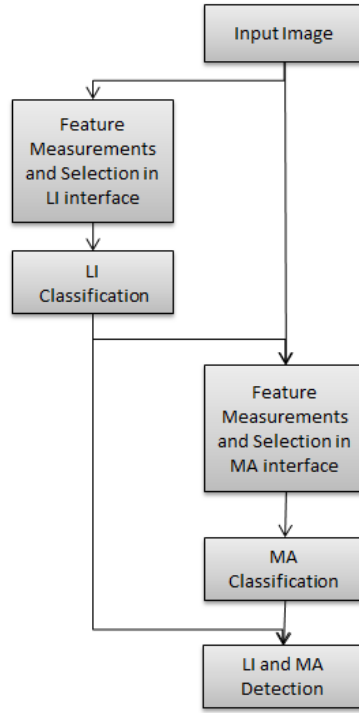


Fig. 3.1: Flowchart of the proposed method. The main stages are the features computations, the interfaces classification and the interfaces detection.

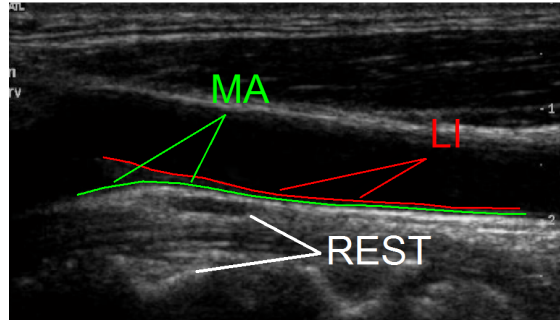


Fig. 3.2: Pixels' representation for the three types of classes.

sionality reduction, classification methods, dynamic programming function and methodology summary.

3.1 Feature Measurements

The first stage of the proposed method is the measurement, at each pixel, of a set of features. Each sample (pixel) will be characterized by a set of features and a class. Each image has three classes of pixels, the lumen-intima (LI) pixels, the media-adventitia (MA) pixels and all the other pixels (REST class). These three classes are represented in Fig. 3.2.

For each pixel, several features are computed. These features allow the differentiation between classes.

The features computed for the LI and the MA interfaces are almost the same but, the responses of a given feature for the LI is different from the response of the same feature for the MA. Six more features were computed for the MA interface, in comparison with the LI interface. They are pixel-based or region-based. This combination intends to relate the pixel information with the pixel neighboring information.

In order to simplify the features description, they are grouped by intensity, neighborhood, gradient orientation and column profile features. Each one is explained and justified. Some of the used features were firstly computed by Rocha et al. [2011] and are marked with *. The most important one is the instantaneous coefficient of variation (ICOV) edge detector, developed by Yu and Acton [2004].

This edge detector measures the edge strength in speckled images, like the ultrasound images. The value of the ICOV edge detector is given by equation 3.1:

$$ICOV(x, y) = \sqrt{\frac{\left| (1/2) \|\nabla I(x, y)\|^2 - (1/16)(\nabla^2 I(x, y))^2 \right|}{(I(x, y) + (1/4)\nabla^2 I(x, y))^2}} \quad (3.1)$$

where I represents the image intensity, ∇ , ∇^2 , $\|\cdot\|$ and $|\cdot|$ are the gradient, Laplacian, gradient magnitude and absolute value, respectively. As shown in equation 3.1, the ICOV combines information of the image intensity with the first and second derivative. This measure intends to allow for balanced and well localized edge strength measurements in bright regions as well as in dark regions. Its good performance was already demonstrated for edge-preserving speckle-reducing anisotropic diffusion on ultrasound images [Yu and Acton, 2004].

The list below presents the features computed for the LI interface:

1. Intensity based - features that use intensity values directly or indirectly:
 - Intensity magnitude - it has high intensity values at the searched interfaces. This information is obtained from the normalized input image (Fig. 3.3 (a));
 - Mean ($Intensity_LI_{Mean}$ of equation 3.2), maximum ($Intensity_LI_{Max}$ of equation 3.3) and standard deviation (std) ($Intensity_LI_{Std}$ of equation 3.4) of intensity value between a pixel and the lumen axis - measures the maximum intensity between the studied pixel and the lumen-axis. It is known that the LI interface is placed near the first high intensity peak between the lumen axis and the artery tissue (Fig. 3.3 (b));

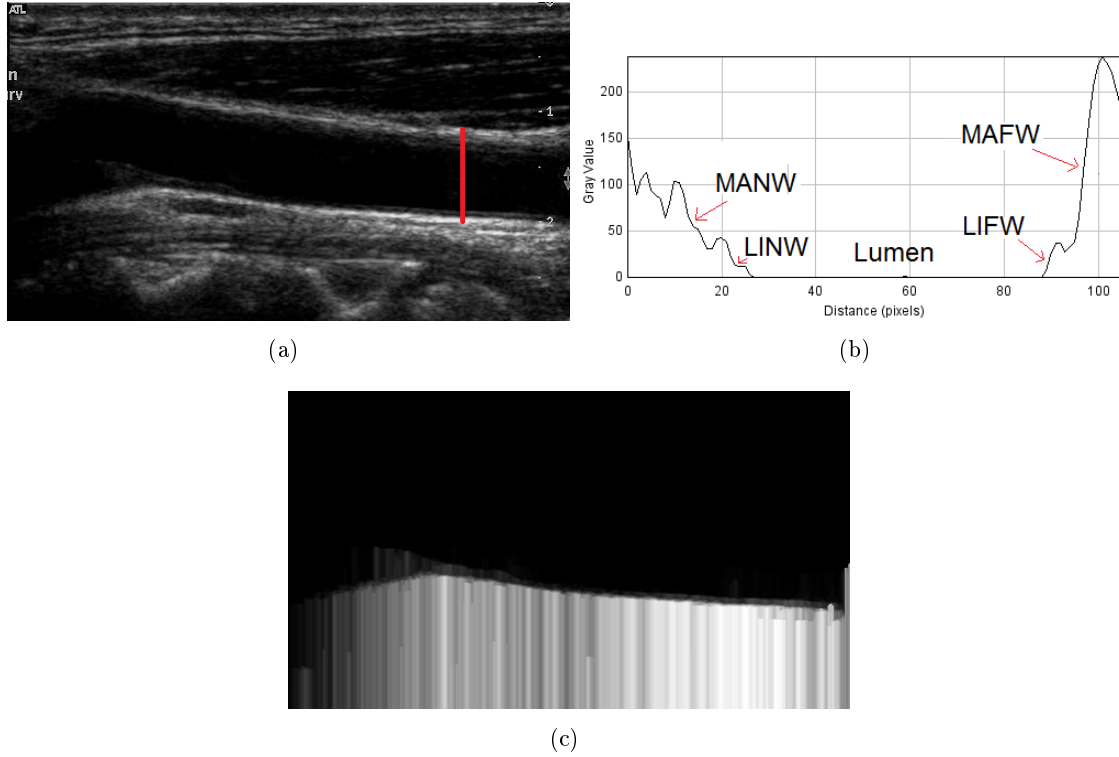


Fig. 3.3: Intensity based features: (a) intensity magnitude image; (b) intensity profile of intensity magnitude image. The profile was measured in the red line presented in (a); (c) maximum of intensity between pixel and lumen axis.

$$Intensity_LI_{Mean} = mean(I(c, r_LA : r_px)) \quad (3.2)$$

$$Intensity_LI_{Max} = max(I(c, r_LA : r_px)) \quad (3.3)$$

$$Intensity_LI_{Std} = std(I(c, r_LA : r_px)) \quad (3.4)$$

where I is the input image, c is the studied column of I , r_LA is the lumen axis row and r_px is the studied pixel row.

2. Neighborhood based - features that give information about the neighborhood properties. All of these neighborhood features (I_{Gauss}) were obtained by the convolution of a Gaussian low pass filter ($Gauss_{filter}$) of size h with the input normalized image. Equation 3.5 presents this convolution operation:

$$I_{Gauss} = Gauss_{filter}(h) * I \quad (3.5)$$

- Gaussian image displaced below the pixel under analysis. The displacements are of 3, 5, 7 and 20 pixels. The used Gaussian filter has a h equal to the performed displacement. For instance, the Gaussian filter used for a displacement of 7

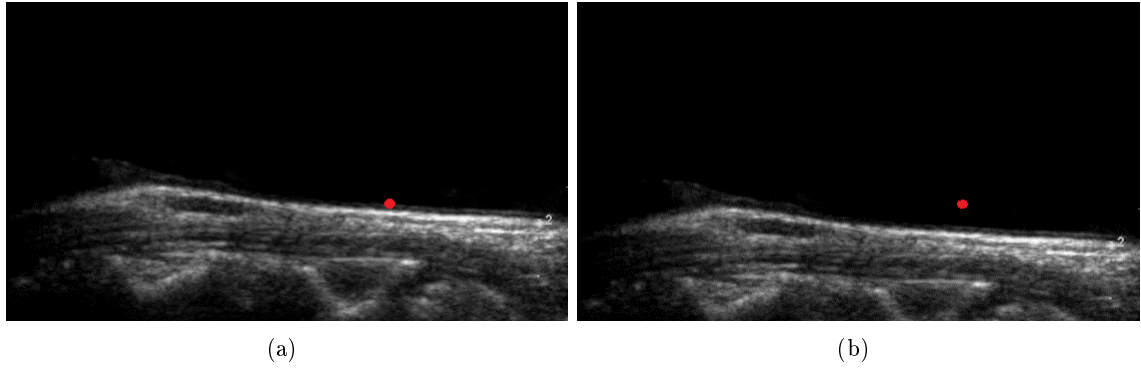


Fig. 3.4: Gaussian windows and its displacements: (a) Gaussian image displaced 3 pixels below; (b) Gaussian image displaced 20 pixels below. The red dot represents the pixel with coordinate (338,221) on both images.

pixels has also a h size of 7 pixels. Figure 3.4 (a) illustrates the image resultant from a window displacement of 3 pixels below the studied pixel and Fig. 3.4 (b) illustrates the image resultant from a window displacement of 20 pixels below the studied pixel. The red dot represents the pixel with coordinate (338,221) in both images. As these figures show, the red dot in the image obtained with a displacement of 3 pixels is placed near the LI interface and the red dot in the image obtained with a displacement of 20 pixels is at the lumen. The displacement allows us to relate the analyzed pixel with the respective neighbors. The feature that uses a displacement of 20 pixels is very useful to relate the studied pixel position with the lumen position and is only used for the below displacement.

- Gaussian image displaced above the pixel under analysis with Gaussian filter h of 3, 5 and 7 pixels. These features are similar to the ones described above but instead of studying the neighbors below, is the neighbors above that pixel that are studied.
- Gaussian image displaced to the left of the analyzed pixel with a Gaussian filter h of 3, 5 and 7 pixels. These features are similar to the one described above but instead of studying the neighbors above or below, it is the left ones that are studied.
- Gaussian image displaced to the right of the analyzed pixel with a Gaussian filter h of 3, 5 and 7 pixels. These features are similar to the ones described above but instead of studying the neighbors above, below or left, it is the right neighbors that are studied.

3. Gradient based - features that use gradient responses in an $n \times n$ region:

- Gradient in horizontal direction (G_x) - high responses for high horizontal vari-

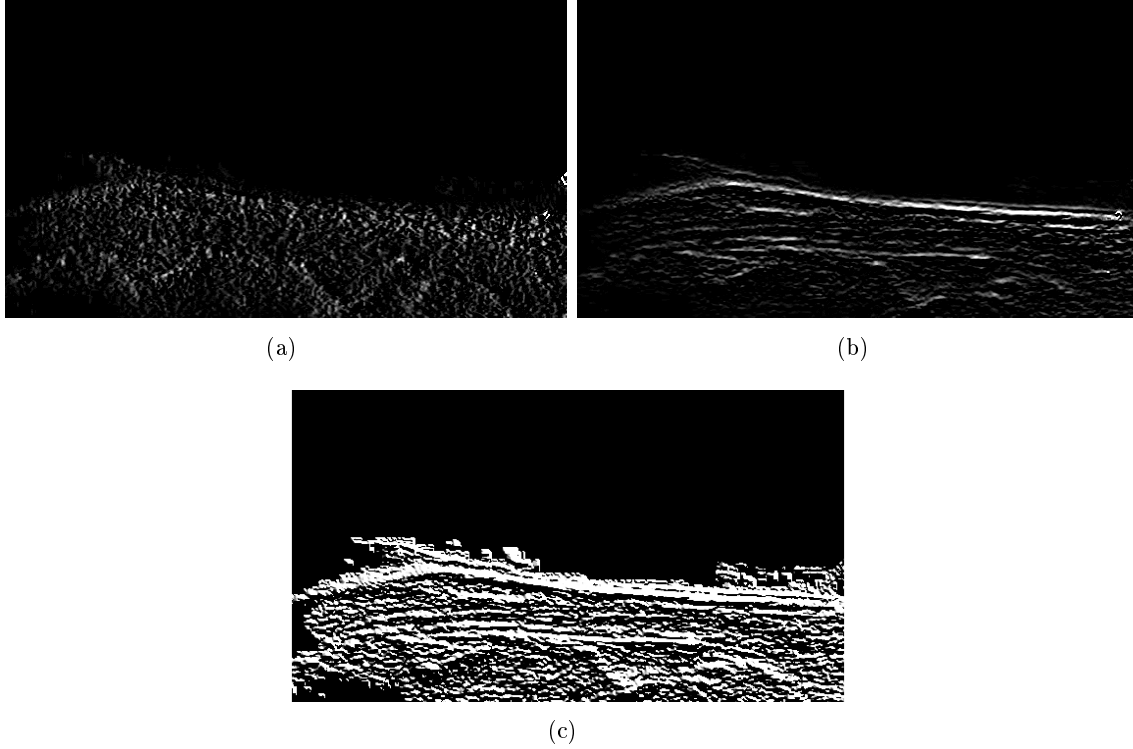


Fig. 3.5: Gradient based features: (a) gradient in horizontal direction; (b) gradient in vertical direction; (c) gradient magnitude.

ations (Fig. 3.5 (a)). The horizontal gradient is defined by: $G_x = \frac{\delta I}{\delta x}$, where I is the input image and x is the horizontal direction;

- Gradient in vertical direction (G_y) - high responses for high vertical variations (Fig. 3.5 (b)). The vertical gradient is defined by: $G_y = \frac{\delta I}{\delta y}$, where I is the input image and y is the vertical direction
- Gradient orientation* - it combines the gradient on horizontal and vertical direction returning the arctangent of the real parts of G_x and G_y (Fig. 3.5 (c)). The gradient orientation ($atan2$) is computed by equation 3.6 [Gdeisat and Lilley]:

$$atan2(G_x, G_y) = \begin{cases} \arctan(\frac{G_x}{G_y}), & G_x > 0 \\ \arctan(\frac{G_x}{G_y}) + \pi, & G_y \geq 0, G_x < 0 \\ \arctan(\frac{G_x}{G_y}) - \pi, & G_y < 0, G_x < 0 \end{cases} \quad (3.6)$$

- Instantaneous coefficient of variation (ICOV)* (Fig. 3.6 (a)) [Rocha et al., 2011, Yu and Acton, 2004];
- Mean ($ICOV_LI_{Mean}$ of equation 3.7), maximum ($ICOV_LI_{Max}$ of equation 3.8) and standard deviation ($ICOV_LI_{Std}$ of equation 3.9) of the ICOV

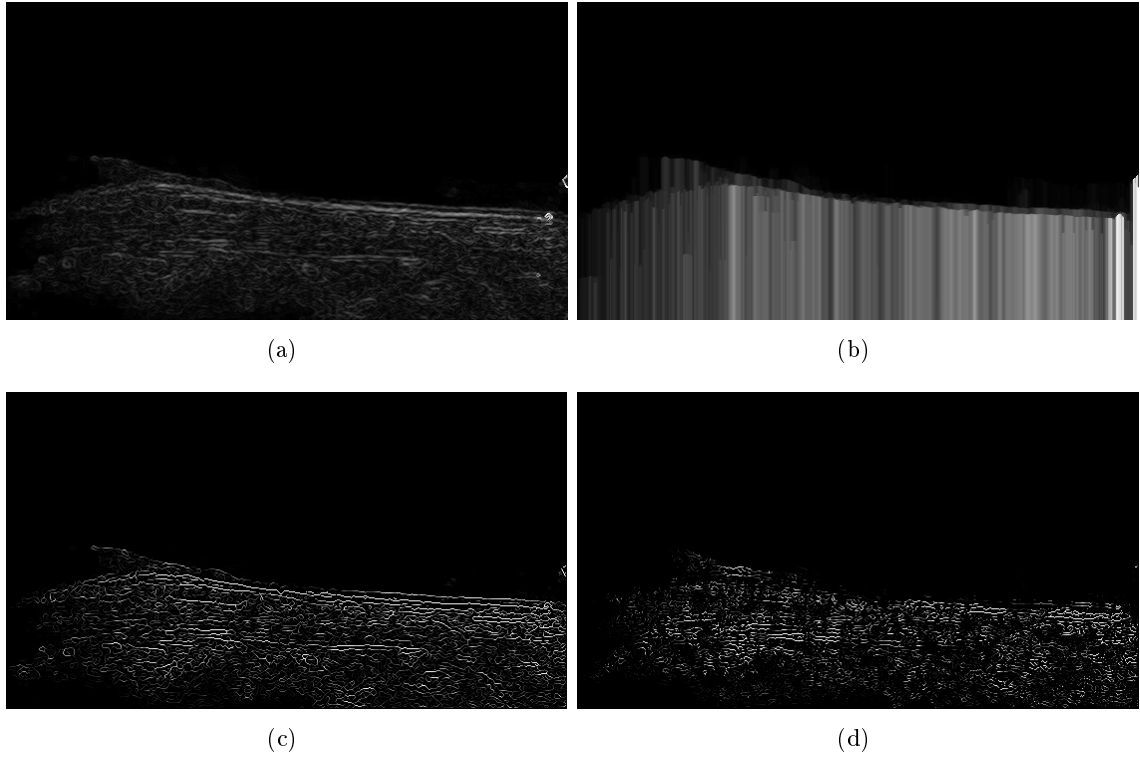


Fig. 3.6: Gradient and column profile based features: (a) ICOV; (b) maximum of ICOV between pixel and lumen axis; (c) fuzzy step edges; (d) fuzzy valley edges.

between the studied pixel and the lumen axis - analyses the column ICOV variations (Fig. 3.6 (b)).

$$ICOV_LI_{Mean} = mean(ICOV(c, r_LA : r_px)) \quad (3.7)$$

$$ICOV_LI_{Max} = max(ICOV(c, r_LA : r_px)) \quad (3.8)$$

$$ICOV_LI_{Std} = std(ICOV(c, r_LA : r_px)) \quad (3.9)$$

where $ICOV$ is the ICOV image, c is the studied column of the $ICOV$ image, r_LA is the lumen axis row and r_px is the studied pixel row.

4. Column profile - features based in the intensity/ICOV of profile of a given column:

- Fuzzy step edge* - it has high responses in regions with a step intensity profile like the LI interface. This happens because the lumen usually has low ICOV pixels and the LI interface has high ICOV values. This feature combines information of the gradient orientation with the ICOV intensity (Fig. 3.6 (c));
- Fuzzy valley edge* - it has high responses in edges associated to a valley shaped intensity profile like the MA interface edges. This happens because the adventitia and the intima layers have higher intensity than the media layer. This

feature combines the intensity information, given by the edges associated to a valley shaped intensity profile, with the gradient orientation information given by the fuzz step edge (Fig. 3.6 (d)).

The features computed for the MA interface were the same as the ones computed for the LI, plus the following six features:

1. Intensity based:

- Mean ($Intensity_MA_{Mean}$ of equation 3.10), maximum ($Intensity_MA_{Max}$ of equation 3.11) and standard deviation ($Intensity_MA_{Std}$ of equation 3.12) of the intensity value between the studied pixel and the LI interface - measures the column intensity variations (Fig. 3.7 (a)). It's known that the MA is associated with the first high intensity peak below the LI interface, as Fig. 3.3 shows.

$$Intensity_MA_{Mean} = mean(I(c, r_LI : r_px)) \quad (3.10)$$

$$Intensity_MA_{Max} = max(I(c, r_LI : r_px)) \quad (3.11)$$

$$Intensity_MA_{Std} = std(I(c, r_LI : r_px)) \quad (3.12)$$

where I is the input image, c is the studied column of I , r_LI is the LI interface row and r_px is the studied pixel row.

2. Gradient based:

- Mean ($ICOV_MA_{Mean}$ of equation 3.13), maximum ($ICOV_MA_{Max}$ of equation 3.14) and standard deviation ($ICOV_MA_{Std}$ of equation 3.15) of the ICOV between the LI interface and studied pixel - analyses the column ICOV variations (Fig. 3.7 (b)).

$$ICOV_MA_{Mean} = mean(ICOV(c, r_LI : r2)) \quad (3.13)$$

$$ICOV_MA_{Max} = max(ICOV(c, r_LI : r2)) \quad (3.14)$$

$$ICOV_MA_{Std} = std(ICOV(c, r_LI : r2)) \quad (3.15)$$

where $ICOV$ is the ICOV image, c is the studied column of the $ICOV$ image, r_LI is the LI interface row and r_px is the studied pixel row.

Table 3.1 summarizes the measured features. A brief description and identification (ID) are given to each feature.

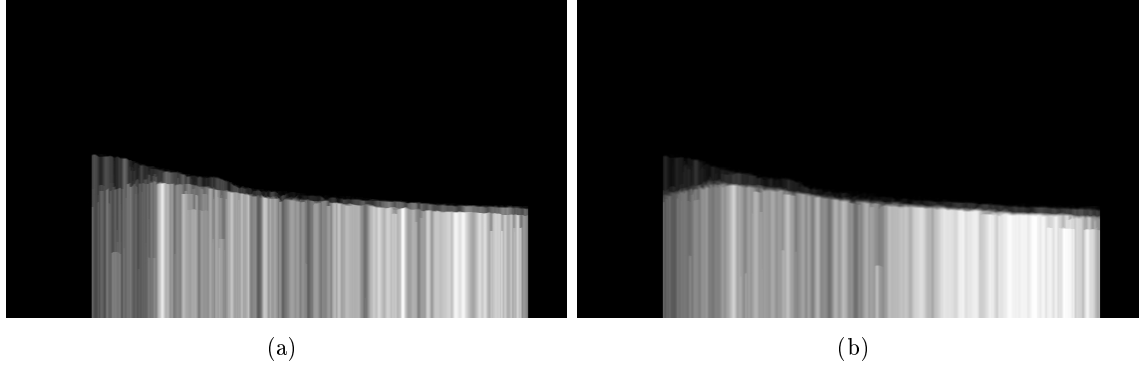


Fig. 3.7: Specific MA interface features: (a) maximum ICOV between the studied pixel and the LI interface previously detected; (b) maximum intensity between the studied pixel and the LI interface previously detected.

3.2 Feature Analysis

The definition of the feature set needs to be grounded by the type of image properties we want to measure quantitatively. On the other hand, we need to be able to detect and reduce feature redundancy and even a misclassified ground-truth. The ground-truths used by the developed method are contours plotted manually by medical experts.

This section addresses four major considerations: the feature redundancy; the selection of the pixels used as training references; the removing of misclassified pixels; and the dataset resampling.

3.2.1 Feature Redundancy

The features to be measured should be based in the assumption that they could lead to a good interfaces detection, which means that it is expected that the computed features have good information about the interfaces. The computational cost also needs to be considered. The larger the dataset, the higher the cost.

To reduce the redundancy of features and the computational cost, a feature selection method was used. It reduces the number of features and the computational cost without compromising significantly the classification performance. This issue will be addressed in more detail in section 3.3.

3.2.2 Pixels Selection of LI and MA Classes

The given ground-truth is a manual trace of the LI and MA interfaces of the far and the near walls. This kind of ground-truths are always subject to human error. It is also known that the expert's fatigue, the repetition of similar images or the lack of concentration contributes to a higher human error rate.

Figure 3.8 shows some examples of ground-truths. In Fig. 3.8 (a) and (b) two

Tab. 3.1: Measured Features

ID	Feature name	Description
F1	Gradient in x	gradient in horizontal direction
F2	Gradient in y	gradient in vertical direction
F3	Fuzzy step edge	edge intensity profiles
F4	Fuzzy valley edge	valley intensity profiles
F5	Gradient orientation	arctangent of gradient in x and y components
F6	Intensity magnitude	normalized image intensity
F7	ICOV magnitude	ICOV magnitude
F8	Maximum intensity (px-LA)	maximum intensity between pixel and lumen axis
F9	Maximum ICOV (px-LA)	maximum ICOV between pixel and lumen axis
F10	Mean intensity (px-LA)	mean intensity between pixel and lumen axis
F11	Mean ICOV (px-LA)	mean ICOV between pixel and lumen axis
F12	Std intensity (px-LA)	Standard deviation of intensity between pixel and lumen axis
F13	Std ICOV (px-LA)	Standard deviation of ICOV between pixel and lumen axis
F14	Maximum intensity (px-LI)	maximum intensity between pixel and LI
F15	Maximum ICOV (px-LI)	maximum ICOV between pixel and LI
F16	Mean intensity (px-LI)	maximum intensity between pixel and LI
F17	Mean ICOV (px-LI)	maximum ICOV between pixel and LI
F18	Std intensity (px-LI)	Standard deviation of intensity between pixel and LI
F19	Std ICOV (px-LI)	Standard deviation of ICOV between pixel and LI
F20	Gaussian up 3	Gaussian image with window size of 3 pixels displaced 3 pixels above
F21	Gaussian down 3	Gaussian image with window size of 3 pixels displaced 3 pixels below
F22	Gaussian up 5	Gaussian image with window size of 5 pixels displaced 5 pixels above
F23	Gaussian down 5	Gaussian image with window size of 5 pixels displaced 5 pixels below
F24	Gaussian up 7	Gaussian image with window size of 7 pixels displaced 7 pixels above
F25	Gaussian down 7	Gaussian image with window size of 7 pixels displaced 7 pixels below
F26	Gaussian left 3	Gaussian image with window size of 3 pixels displaced 3 pixels left
F27	Gaussian right 3	Gaussian image with window size of 3 pixels displaced 3 pixels right
F28	Gaussian left 5	Gaussian image with window size of 5 pixels displaced 5 pixels left
F29	Gaussian right 5	Gaussian image with window size of 5 pixels displaced 5 pixels right
F30	Gaussian left 7	Gaussian image with window size of 7 pixels displaced 7 pixels left
F31	Gaussian right 7	Gaussian image with window size of 7 pixels displaced 7 pixels right
F32	Gaussian below 20	Gaussian image with window size of 20 pixels displaced 20 pixels below

ID- features identification; px - studied pixel; LA - lumen-axis; LI - Lumen-Intima; Std - standard deviation

questionable chosen interfaces are visible . Figure 3.8 (c) and (d) present two logical and coherent traces.

It is clear from these examples that there are situations where the medical experts have difficulty in detecting the correct interface.

The proposed solution to this problem is a specific ground-truth selective method. We considered that only the pixels that are local maxima of the ICOV at the gradient direction, specific for each wall, are reference candidates. This is justified because the ICOV has high responses at the interfaces and because we know which gradient direction is expected for each wall. For instance, at the near wall, the gradient direction points towards the up, because the intensity decreases toward the lumen. The opposite occurs at the far wall.

After obtaining the reference candidates, only those that are coincident with the ground-truth are kept and used as reference. Figure 3.9 (a) shows the used references (represented in red) and the REST pixels (presented in white).

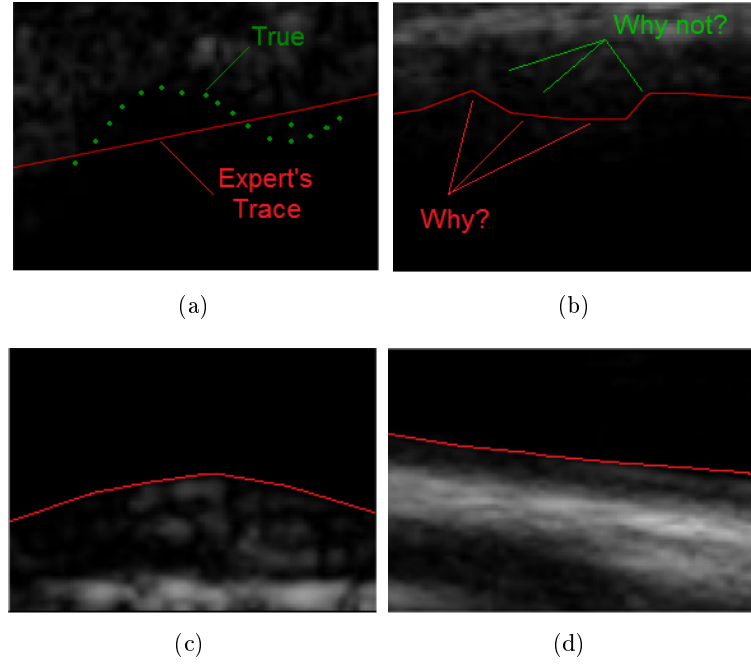


Fig. 3.8: Ground-truth images: (a) and (b) represent incoherent ground-truths; (c) and (d) represent coherent ground-truths.

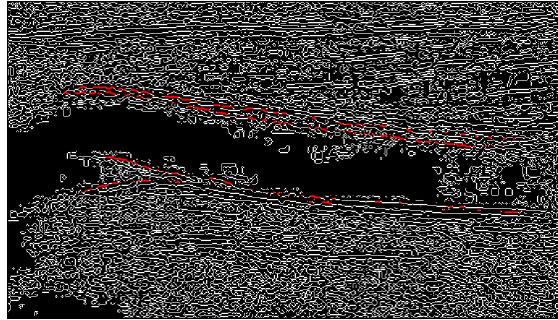


Fig. 3.9: Reference selected pixels; The white pixels represent the maximum ICOV at the gradient direction pixels and the red pixels represent the reference contours for MA and LI classes.

3.2.3 Removing Misclassified Pixels

The REST class candidates are the pixels that are maxima of the ICOV in the gradient direction and are not LI or MA pixels. Figure 3.10 (a) shows three situations of possible pixel incorrect labeling. The numbers presented in the figure are related with the removal criteria described below. Situation 1 describes a MA pixel that can be classified as REST. Situation 2 describes a LI pixel that can also be classified as REST. Situation 3 describes a case where a muscle bundle can be classified as MA class (sometimes the features of this structure are similar to the ones of the MA).

To solve these problems, three removal criteria were used. These are presented below

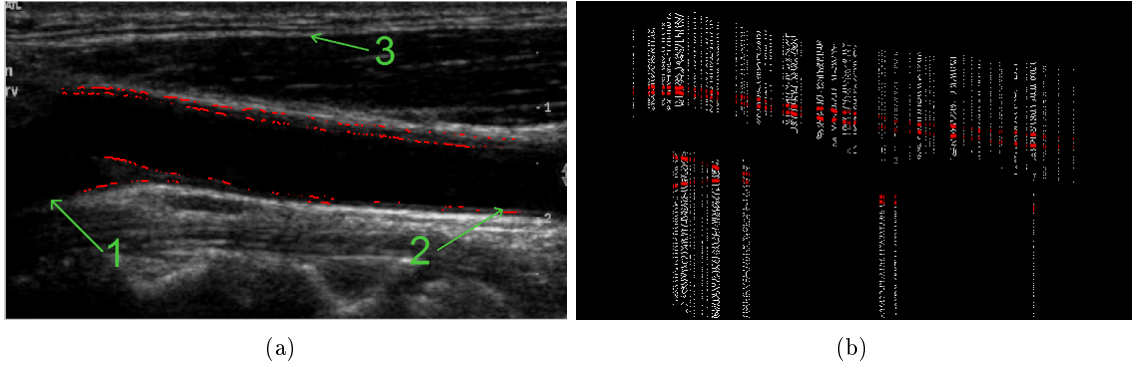


Fig. 3.10: Pixels exclusion: (a) problematic pixels; (b) final result of pixels selection. The red pixels represent the LI and MA pixels and the others represent the REST pixels.

and enumerated according to the numbers in Fig. 3.10 (a).

1. Only pixels belonging to the same image longitudinal range of the ground-truth are used;
2. Only pixels that have LI and MA reference in the same column are considered;
3. Only pixels that are located not more than 2 lumen diameters above the lumen axis are used. This criterion was used only for the near wall interface, in order to decrease the possibility of the muscle bundle detection.

The final result of this selection of pixels is shown in Fig. 3.10 (b).

3.2.4 Dataset Resampling

As it is easy to understand by the analysis of Fig. 3.10 (b), the number of the reference pixels from the three classes is very different. This can cause a biased classification towards the class with more samples, which is the REST class.

The solution to this problem is the implementation of a resampling method in order to balance the classes size. This solution will be broadly discussed in section 4.2.

3.3 Dimensionality Reduction

The main goals of feature selection is to select a subset of d features from the given set D , $d < D$, without significantly degrading the performance of the recognition system [Pudil et al., 1994] and to reduce the computational cost.

One of the most commonly used criterion to find the optimal number of features is the inter-intra class distance measure. This measure combines the inter class (S_B) with the intra class (S_W), defined by equation 3.17.

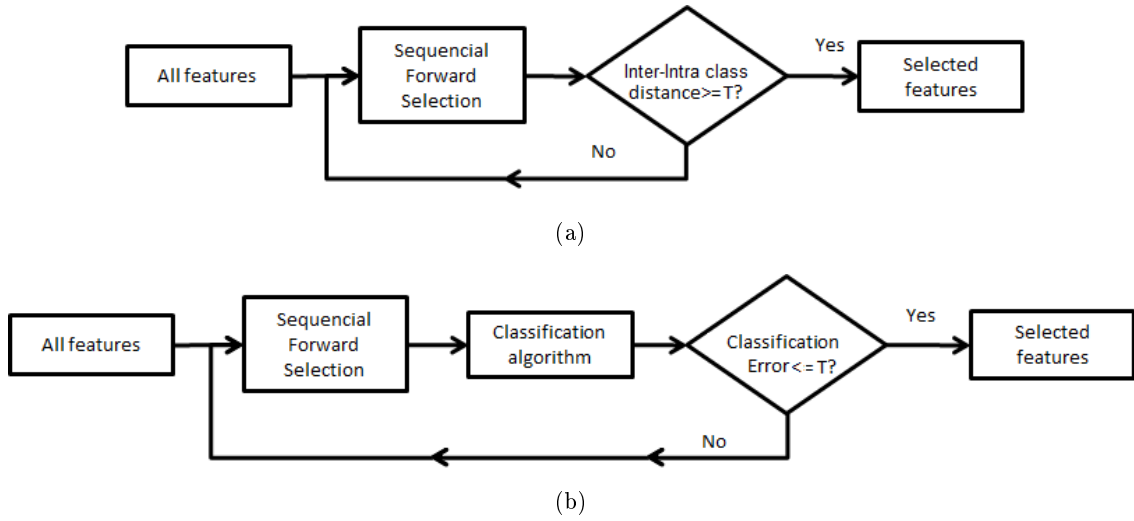


Fig. 3.11: Features selection: (a) sequential forward selection without and (b) with wrapper based method.

$$S_B = \sum_{i=1}^C \frac{n_i}{n} (m_i - m)(m_i - m)^T \quad (3.16)$$

$$S_W = \sum_{i=1}^C \frac{n_i}{n} \hat{\Sigma}_i \quad (3.17)$$

where m_i and $\hat{\Sigma}_i$, $i=1, \dots, C$ are the sample means and covariance matrices of each class (with n_i samples) and m is the sample mean.

The inter-intra class distance ($In - In$) if obtained by equation 3.18.

$$InIn = trace \left(\frac{S_B}{S_W} \right) \quad (3.18)$$

where, *trace* is the sum of the elements on the matrix main diagonal.

A wrapper based methodology can also be used (Fig. 3.11 (b)). It uses the classifier algorithm itself as part of the features subsets evaluation [Kohavi and H. John, 1997]. This method allows a better selection of features because these are chosen according to the selected subset adaptability to the classification method. A disadvantage of this method is that it can induce the classifier overfitting. Figure 3.11 (a) represents the sequential forward selection using the Inter-Intra class distance as evaluation criterion and Fig. 3.11 (b) shows the sequential forward selection using a wrapper based evaluation method.

The feature selection can be done using the sequential forward selection (SFS) algorithm (Fig. 3.11 (a)). This selects iteratively the best feature from the total dataset until a defined performance threshold, T , is achieved. This T should reduce the dataset dimensionality without significantly compromising the classifier performance.

Besides the two presented feature selection methods, and in order to verify the results

with and without feature selection, a dataset without dimensionality reduction was also used.

3.4 Classification Methods

This section gives a brief description of the four classification models used to obtain the final classifier. These were the quadratic Bayes classifier (QDC), the Parzen classifier (ParzenC), the k-nearest neighbor classifier (KNNC) and the support vector classifier (SVC). To illustrate the classifiers properties, we use a subset of samples from 2 classes of pixels, the LI and the OTHERS classes. One of the used subset is extracted from the near wall (NW) and the other is from the far wall (FW). The dimensionality of the used datasets were reduced to 2 features in order to allow a 2D representation. For the dataset of NW, the feature 1 corresponds to F12 and the feature 2 corresponds to F23. For the FW dataset, the feature 1 used was the F11 and the feature 2 was the F23. Figure 3.12 represents the feature space created by the two datasets.

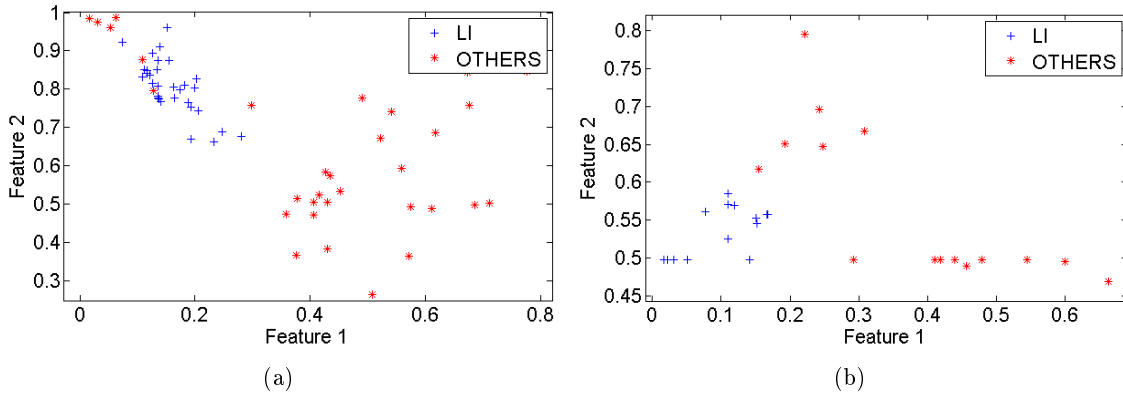


Fig. 3.12: Two experimental datasets. (a) dataset set of NW; (b) dataset of FW. The red markers correspond to the LI class and the blue ones correspond to the OTHERS class.

3.4.1 Quadratic Bayes Classifier

One fundamental statistical approach to the problem of pattern classification is the Bayesian decision theory. It is based on the quantification of the trade-offs between various classification decisions using probability and the costs associated with such decisions. It makes the assumption that the decision problem is posed in probabilistic terms, and that all of the relevant probability values are known [Duda et al., 2000].

In the Bayes theory, an *a priori* probability is converted in a *a posteriori* probability by the equation:

$$p(c_j|x) = \frac{p(x|c_j) \times p(c_j)}{p(x)} \quad (3.19)$$

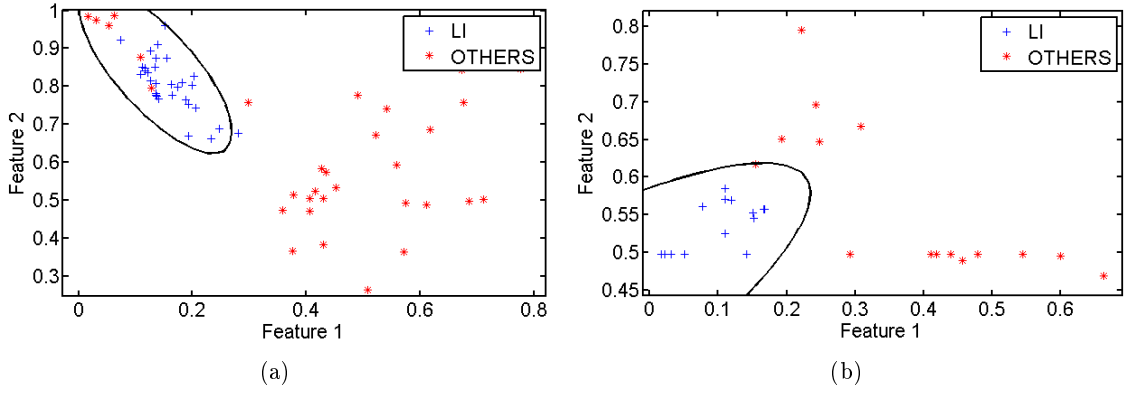


Fig. 3.13: QDC applied to the NW (a) and FW (b) experimental datasets. The red markers correspond to the LI class and the blue ones correspond to the OTHERS class. The black line corresponds to the decision boundary.

where x is the observation and c_j is the j^{th} class. This equation shows that, by observing the value of x , the $p(c_j)$ (*prior* probability) can be converted to the $p(c_j|x)$ (*posterior* probability). The $P(x|c_j)$ is the class conditional probability. The $p(x)$ is a scale factor that ensures that the sum of $p(x)$ is equal to one. More information about the Bayesian theory can be found in Duda et al. [2000].

The QDC is based in the Bayes theorem and has a multivariate density. The general multivariate normal density was transformed in a density discriminant function and the result is:

$$g_i(x) = -\frac{1}{2}(x - \mu_i)^t \Sigma_i^{-1} (x - \mu_i) - \frac{d}{2} \ln 2\pi - \frac{1}{2} \ln |\Sigma_i| + \ln P(c_i) \quad (3.20)$$

where μ is the observations mean, Σ is the covariance matrix, $|\Sigma|$ and Σ^{-1} are its determinant and inverse, respectively, and $P(c_i)$ is the *prior* probability of class i .

In the two-category case, the decision surfaces are hyperquadratics. Figure 3.13 shows two examples of a QDC applied to the experimental datasets. The classification error for the NW dataset was 9.68% and to the FW dataset was 3.44%.

3.4.2 Parzen Classifier

The ParzenC is a non-parametric method that is based in an unknown probability density function (pdf) estimation associated with a given set of points. According to the Parzen windows pdf estimation method, if N data points, $x_i \in R^l, i = 1, 2, \dots, N$, are given and they follow an unknown distribution, their pdf can be estimated using the function [Theodoridis and Koutroumbas, 2010]:

$$p(x) = \frac{1}{Nh^l} \sum_{i=1}^N \phi\left(\frac{x - x_i}{h}\right) \quad (3.21)$$

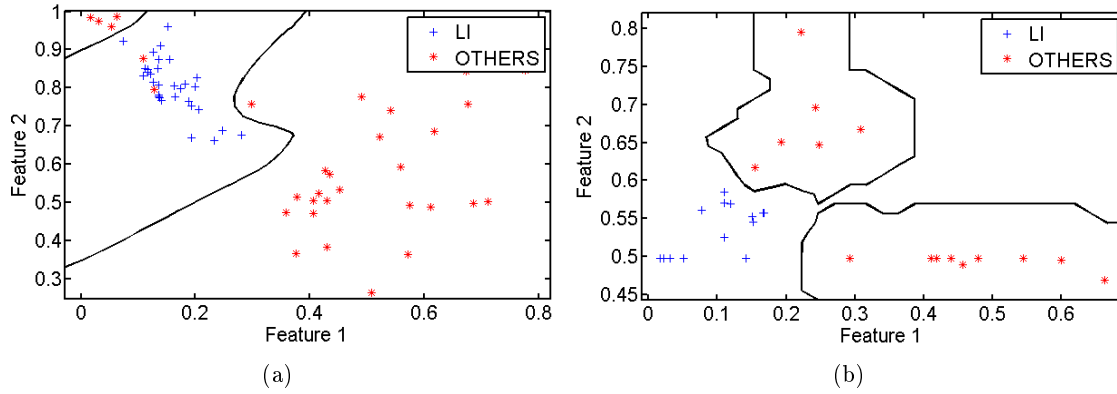


Fig. 3.14: ParzenC applied to the NW (a) and FW (b) experimental datasets. The red markers correspond to the LI class and the blue ones correspond to the OTHERS class. The black line corresponds to the decision boundary.

for sufficiently large N and sufficiently small values of h , which is the window length defined by the user. ϕ is an appropriately defined kernel function and l is related with the dimensional features space. A commonly used kernel function is the Gaussian, and in this case the expression becomes:

$$p(x) = \frac{1}{Nh^l} \sum_{i=1}^N \frac{1}{(2\pi)^{l/2} h^l} \exp \left(- \frac{(x - x_i)^T (x - x_i)}{2h^2} \right) \quad (3.22)$$

Figure 3.14 shows two examples of a ParzenC applied to the the experimental datasets. The classification error for the NW dataset was 3.22% and to the FW dataset was 0%.

3.4.3 K-Nearest Neighbor Classifier

The KNNC is a particular case of the Bayes theory. The goal is to estimate the value of the unknown pdf at a given point x . According to the k-nearest neighbor estimation technique, the following steps are performed:

1. Choose a value for k .
2. Find the distance between x and all training points x_i , $i = 1, 2, \dots, N$. Any distance measure can be used (e.g., Euclidean, Mahalanobis).
3. Find the k -nearest points to x .
4. Compute the volume $V(x)$ in which the k -nearest neighbors lie.
5. Compute the estimate by:

$$p(x) = \frac{k}{NV(x)} \quad (3.23)$$

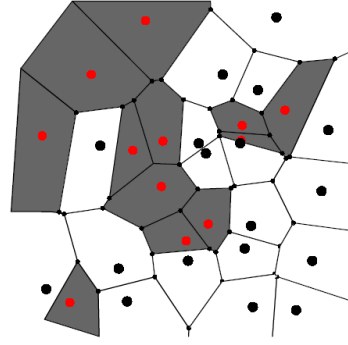


Fig. 3.15: Voronoi tessellation - This example shows two classes (red and black dots) and the regions decision boundaries [Duda et al., 2000].

If the Euclidean distance [Deza and Deza, 2009] is employed and the distance between the k -furthest neighbor and x is ρ , the volume $V(x)$ is equal to [Theodoridis and Koutroumbas, 2010]:

- $V(x) = 2\rho$ in the 1-dimensional space;
- $V(x) = \pi\rho^2$ in the 2-dimensional space;
- $V(x) = \frac{4}{3}\pi\rho^3$ in the 3-dimensional space;

The region volume ($V(x)$) is defined by the Voronoi cells. These partition the features space. Each cell is defined by a training pattern from the belonging class. The decision regions are the cell boundaries. Figure 3.15 shows a set of training patterns, generated by two classes, the Voronoi cells of each pattern and the decision regions of the nearest neighbor classifier [Duda et al., 2000].

Figures 3.16(a) and (b) show two examples of a k -nn classifier applied to the experimental datasets. The classification error for the NW dataset was 3.22% and to the FW dataset was 0%.

3.4.4 Support Vector Classifier

The Support Vector Machines are learning systems that use a hypothesis space of linear functions in a high dimensional feature space. In SVC the main goal is to choose the decision boundary for which the margin is maximized. The maximum margin solution can be defined using computational learning theory [Cristianini and Shawe-Taylor, 2000].

The support vectors are the training samples that define the optimal separating hyperplane and are the most difficult patterns to classify. These are considered as the most informative patterns for the classification task. Figure 3.17 represents these patterns. The three support vectors are shown as solid dots and a b margin is used to separate the support vectors from the optimal hyperplane.

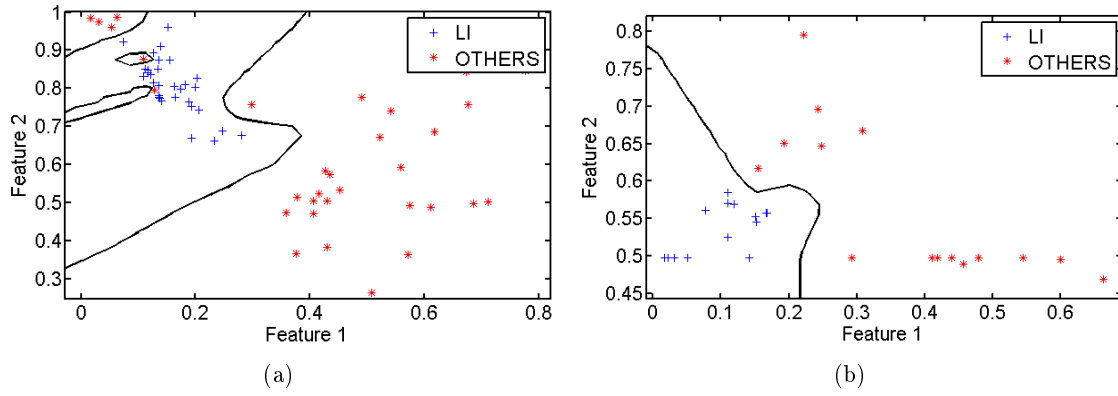


Fig. 3.16: KNNC applied to the NW (a) and FW (b) experimental datasets. The red markers correspond to the LI class and the blue ones correspond to the OTHERS class. The black line corresponds to the decision boundary.

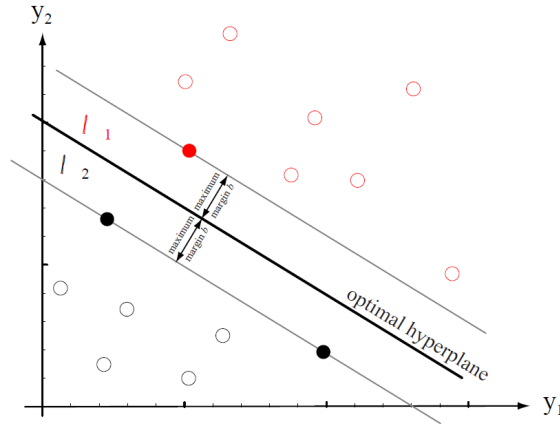


Fig. 3.17: SVM classification with margin b . The support vectors are represented by the filled dots [Cristianini and Shawe-Taylor, 2000].

The maximal margin classifier forms the strategy of the first SVC. The function that represents a geometric margin of 1 is defined by:

$$\langle w \cdot x^+ \rangle + b = +1 \quad (3.24)$$

$$\langle w \cdot x^- \rangle + b = -1 \quad (3.25)$$

where w is the weight vector of the functional margin. The functional margin presented in Fig. 3.17 has a margin of b .

Training a SVC consists of finding the optimal hyperplane that maximizes the distance from the nearest training patterns. The SVC trained at the proposed approach is based in a radial kernel and searches for the distance maximization between the support vectors [Cristianini and Shawe-Taylor, 2000].

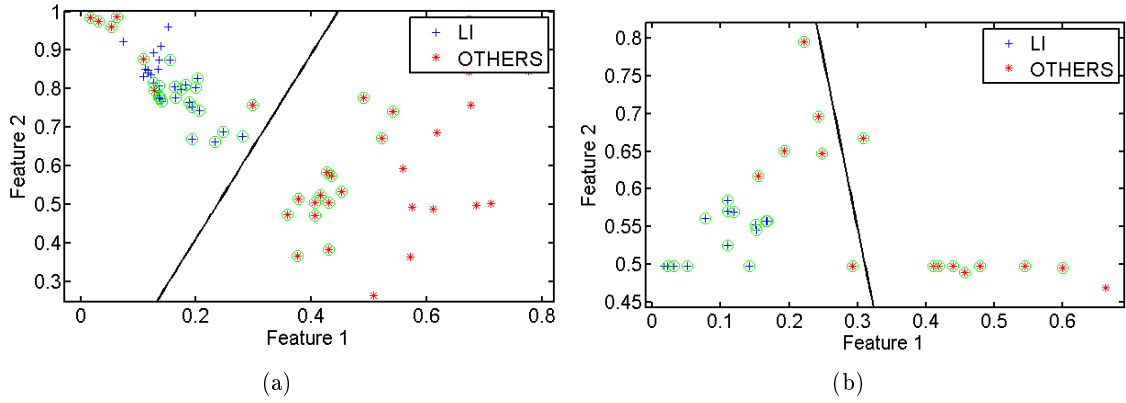


Fig. 3.18: SVC applied to the NW (b) and FW (c) experimental datasets. The red markers correspond to the LI class, the blue ones correspond to the OTHERS class and the green circles correspond to the used support vectors. The black line corresponds to the decision boundary.

Figure 3.18(a) and (b) show two examples of a SVC applied to the experimental datasets. The circles show the support vectors. The classification error for the NW dataset was 11.29% and to the FW dataset was 20.69%.

3.5 Classification Output and Dynamic Programming

The result of the interfaces classification is a score map. In this map, the probability of a pixel belonging to the searched interface is represented by an intensity. The final contour is obtained by passing the score map to a dynamic programming algorithm used in Rocha et al. [2012].

The dynamic programming approach used is based on a gain function with two terms: one is the embedded edge information from the fuzzy score map (of each interface) and the other is a geometric smoothness information. This algorithm scans the score map, from the first to the last column, and searches for the contour that maximizes the gain function *Gain* described below (equation 3.26):

$$Gain = \psi(x_1, y_1) + \sum_{j=2}^N [\psi(x_j, y_j) + \lambda \rho(x_j, y_j)] \quad (3.26)$$

where N is the number of columns of the ROI, (x_j, y_j) are the coordinates of the contour pixel at column j , $\psi(x_j, y_j)$ is the score map, $\rho(x_j, y_j)$ is a geometric term and λ is a weight parameter.

Figure 3.19 (a) presents the result of the LI classification and Fig. 3.19 (b) presents the MA classification result. These score maps will then be used by the dynamic programming function and the final interface is presented in Fig. 3.19 (c).

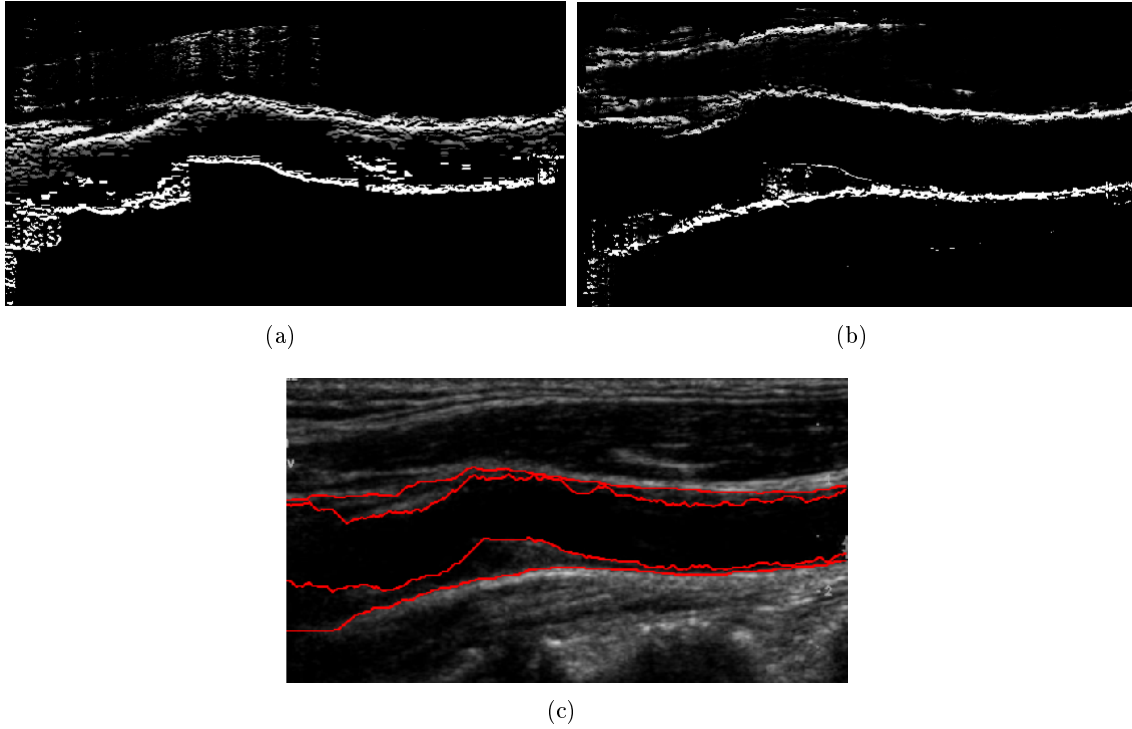


Fig. 3.19: Classification results and final contour: (a) presents the score map of the LI interface for both near and far wall; (b) presents the score map of the MA interface for both walls; (c) represent the final result obtained by the dynamic programming function.

3.6 Methodology Summary

The global methodology flowchart is presented in Fig. 3.20, where all the steps described before are included.

This figure shows that the methodology for the LI (red rectangles) and the MA (blue rectangles) interfaces is almost the same. The differences between these two are that the MA interface requires a previous classification of the LI interface and six more features are measured to this interface.

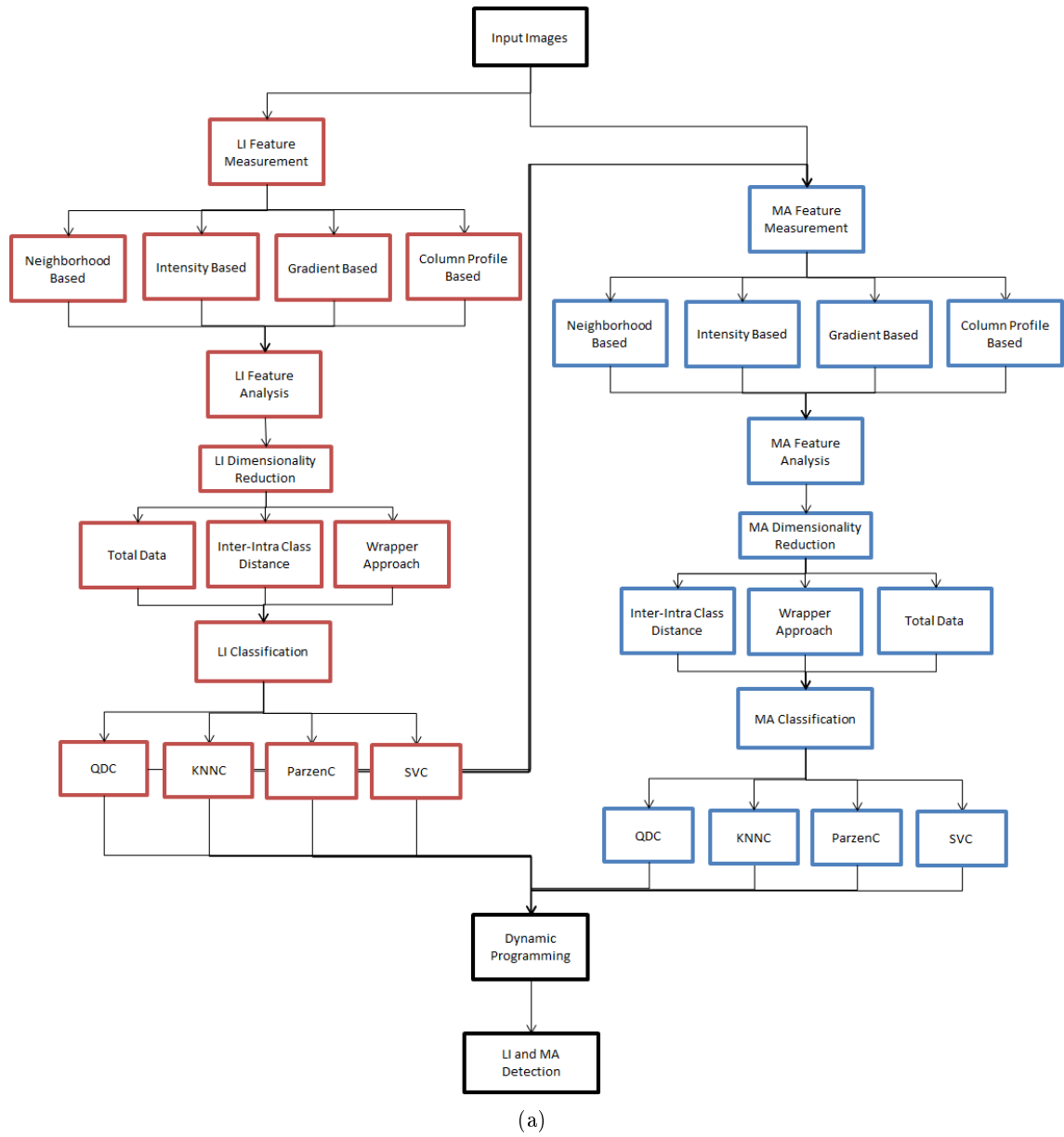


Fig. 3.20: Flowchart of the complete methodology. The most important stages are presented in this flowchart. The red rectangles represent the LI stages and the blue rectangles represent the MA stages.

Experimental Results and Discussion

This chapter presents the results of the method proposed in this thesis. The first section describes the input image dataset. The resampling problem for balancing the number of samples per class is addressed in section 4.2. The feature selection is described in section 4.3. Section 4.4 addresses the training of classifiers and the final results are shown in section 4.5. At last, section 4.6 shows the statistical analysis of the obtained results. Here some performance measurements and the Bland-Altman plots are shown as well as some examples of the final segmentation are given.

The flowchart presented in Fig. 4.1 presents all the methodology stages addressed through this chapter. Each of these blocks are related with a section itemized above.

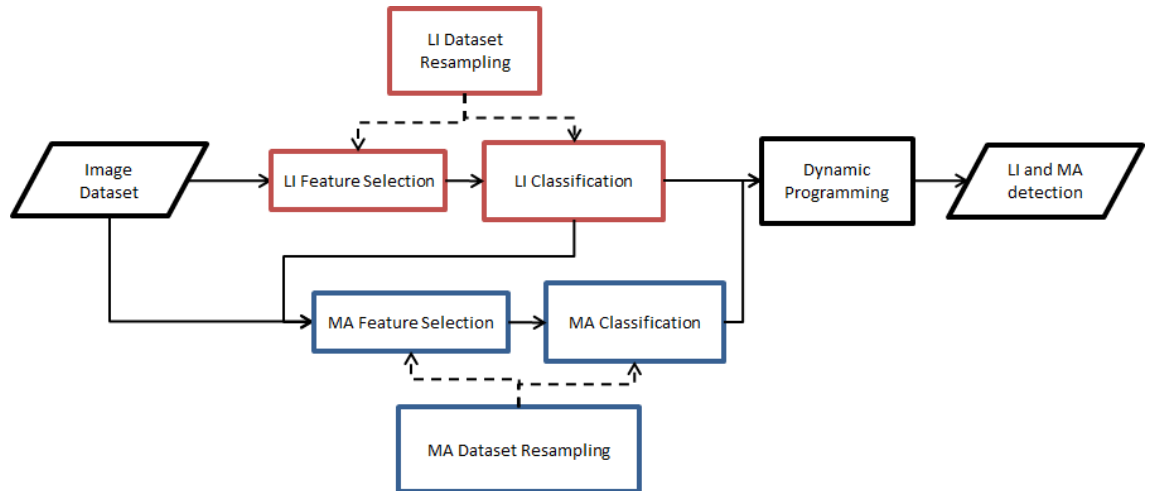


Fig. 4.1: Complete methodology flowchart

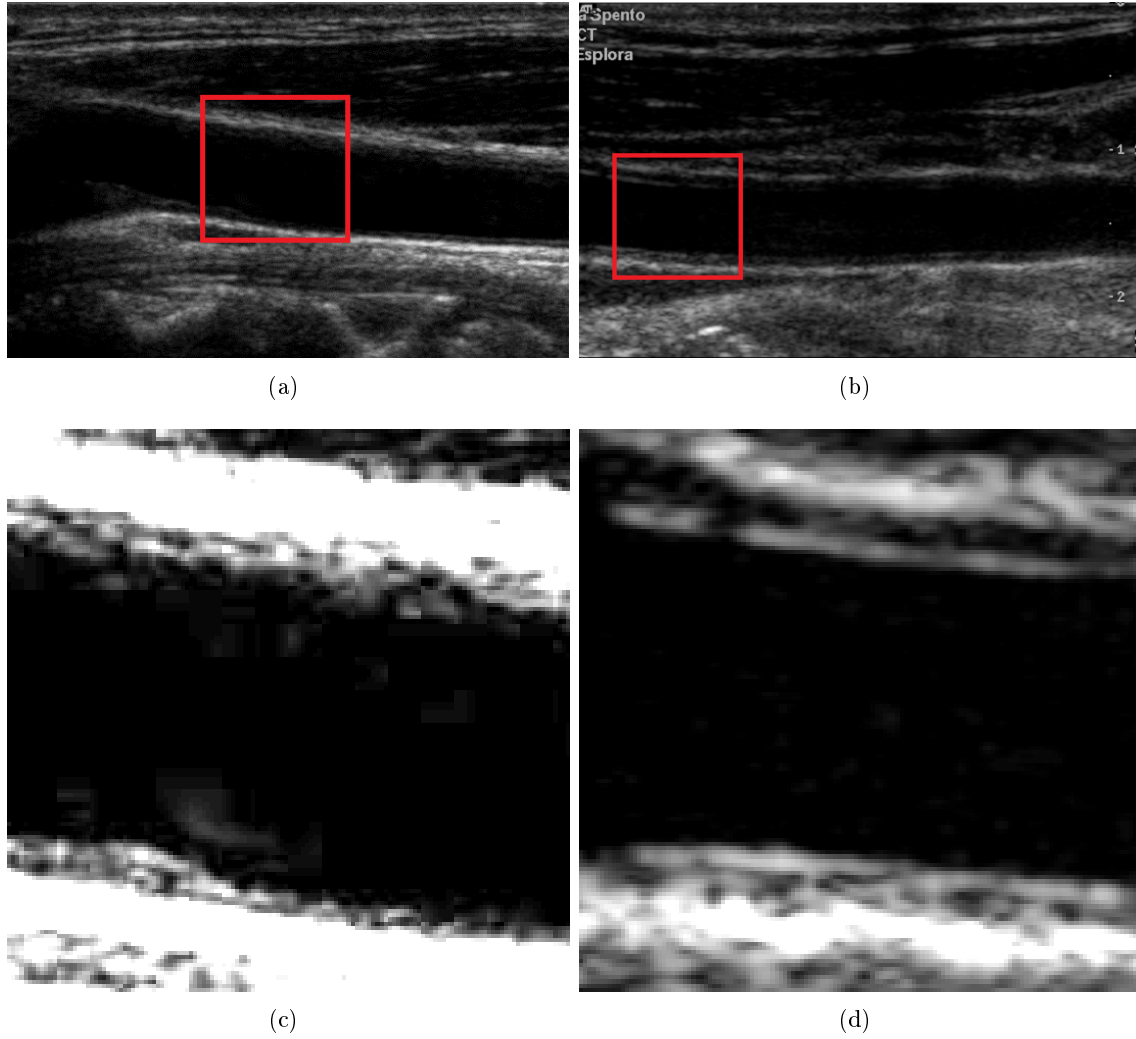


Fig. 4.2: Rocha's and Molinari's image examples: (a) Rocha's image and (b) Molinari's image. Image (c) represents the effect of compression to JPEG format for the red rectangle of (a); and (d) represents the Molinari's PNG with image compression for the red rectangle of (b).

4.1 The Image Dataset

This methodology was applied to two image databases: the Rocha's [Rocha et al., 2011] (Fig. 4.2 (a)) and the Molinari's [Molinari et al., 2011a, 2012c] (Fig. 4.2 (b)) databases. The Rocha's database is composed of 50 longitudinal images of the common carotid artery, in JPEG format. This is a lossy compression type ([Wiggins et al., 2001]) and an example of information loss is presented in Fig. 4.2 (c).

The Molinari's database is composed of 90 longitudinal images of the common carotid artery. The original image format is DICOM and they were converted to PNG format. This is a lossless compression technique [Wiggins et al., 2001] and as Fig. 4.2 (d) shows the data preservation.

The image pixel size in the two databases was normalized to 0.09 mm. All images of both databases were acquired using a Philips ATL HDI 5000 [Rocha et al., 2011, Molinari et al., 2012a].

The methodology used is the same to both databases but the classifiers were trained separately for each database.

Twenty-three images of Rocha’s database were used for training. This training set has one image from each patient. Twenty-five images from the Molinari’s database were used for training. Once again, only one image per patient was chosen from this database. This selection is done to have a representative and unbiased training sample.

The ethical issues were taken under consideration and the patients’ personal information is not revealed anywhere.

Rocha’s database has three ground-truths manually traced by 2 medical experts for each arterial interface. These are referred to as MA1 and MA2 ground-truths, both from the same expert but traced in different occasions, and MB1 ground-truth which is from another expert. The traced interfaces are the LI of the near wall (LINW), the LI of the far wall (LIFW), the MA of the near wall (MANW) and the MA of the far wall (MAFW). Molinari’s database has only one ground-truth (MM1) and it is just for the far wall interfaces (LIFW and MAFW).

4.2 Dataset Resampling

The dataset resampling is an auxiliary procedure used at the training stage. This is performed after defining the input image dataset. Figure 4.1 shows this procedure enclosed to the total methodology.

As it was already mentioned in section 3.2.4, the different classes have very different number of pixels. To overcome this problem we used a resampling method based in a subsample architecture. As it can be seen in Fig. 4.3 (a), the LI and MA pixels (red pixels) are much less than the REST pixels (white pixels).

Figure 4.3 (b) shows the organization of each dataset, by grouping different classes of pixels. The LI dataset is composed by LI, MA and REST pixels. These three classes dataset is used to train a 2 classes classifier. One of the two classes is the LI class (LI pixels) and the other is the OTHERS class (MA and REST pixels). For the MA dataset there are only two kinds of pixels, the MA and the REST pixels. So, the MA dataset is composed by the MA and REST classes.

Before developing the resampling method, some considerations are taken:

- LI dataset: The LI dataset should have 50% of pixels from the LI class and a 50% of pixels from the OTHERS class (25% of MA pixels and 25% of REST pixels). These percentages were chosen to the test situation where the main goal is to discriminate the LI pixels from the MA and REST pixels.

Tab. 4.1: An example of the resampling method for the LI dataset: original and resampled classes sizes. For the LI dataset of Fig. 4.3 (a).

Class	Number of pixels	Class percentage in LI dataset(%)
Class size without resampling		
LI	125	2.63
MA	180	3.79
REST	4447	93.58
Class size with resampling		
LI	125	50.2%
MA	62	24.9%
REST	62	24.9%

- MA dataset: The MA dataset should have 50% of pixels from the MA class and a 50% of pixels from the REST class. These percentages create a balanced dataset of both classes.

Figure 4.3 (a) is used to illustrate the LI resampling and table 4.1 presents the original and resampled sizes of each class of this figure for the LI dataset. The 125 pixels from the LI class are selected, fulfilling the 50% of pixels from the LI class, and 62 pixels from the MA and REST class are also selected, fulfilling the others 50% ($\approx 25\%$ for each class).

The same figure (4.3 (a)) is also used to illustrate the MA resampling and table 4.2 presents the original and resampled sizes of each class of this figure for the MA dataset. This method is slightly different from the LI resampling method. This is a two classes problem only composed by MA and REST pixels. So, all the pixels from the MA class are selected and the same number of pixels from the other class is randomly selected.

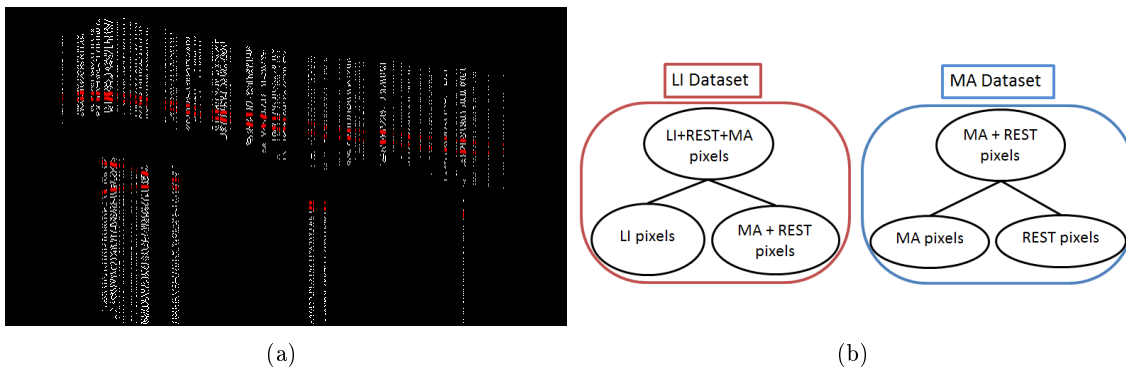


Fig. 4.3: (a) Reference pixels. The white pixels represent the pixels from the REST class and the red ones represent the pixels from the LI and MA classes; (b) LI and MA datasets structure. The LI dataset is composed by LI, MA and REST pixels and the MA dataset is composed by MA and REST pixels.

Tab. 4.2: An example of the resampling method for the MA dataset: original and resampled classes sizes. Example for MA dataset of Fig. 4.3 (a).

Class	Number of pixels	Class percentage in MA dataset(%)
Class sizes without resampling		
MA	180	3.89
REST	4447	96.10
Class sizes with resampling		
MA	180	50%
REST	180	50%

4.3 Feature Selection

Figure 4.1 shows, the feature selection blocks and its relations with the rest of the methodology.

Four datasets, one for each interface, were created from Rocha's database training images. The samples of Molinari's database training images were also grouped in LIFW dataset and MAFW dataset. These datasets combine the information of the pixel feature with the class.

The feature selection method was applied to each dataset and only the selected features were used afterwards.

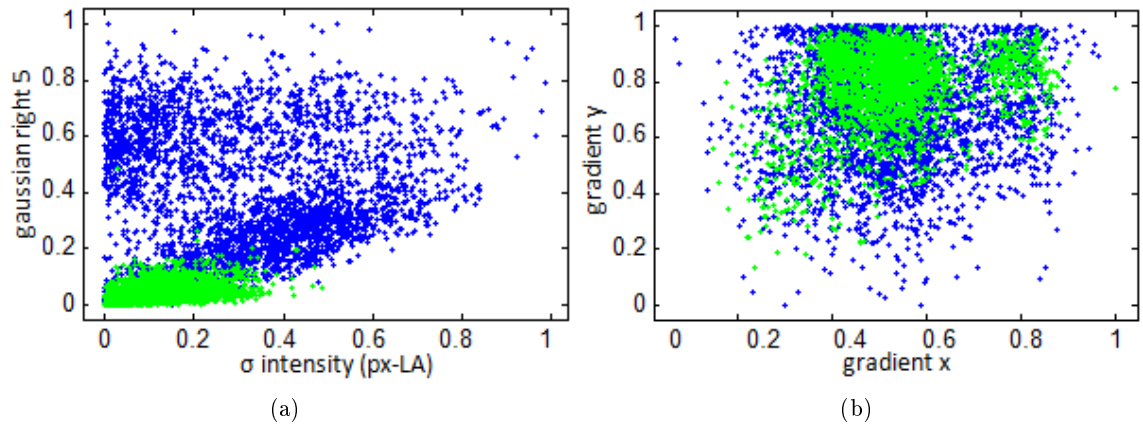


Fig. 4.4: Scatter diagram for a discriminant pair of features (a) and a non discriminant one (b), using the Rocha's LINW dataset. The green points represent the LI class and the blues represent the REST class.

Figure 4.4 shows the relation between two pairs of features. Fig. 4.4 (a) presents two features (standard deviation of intensity between lumen axis and pixel and a gaussian filtered image, with window size of 5, displaced to the right) that discriminate well the LI and REST classes. Figure 4.4 (b) presents two features (gradient in x and gradient in y) that do not discriminate well the same two classes and can be considered disadvantageous

to the classifier.

To evaluate the best feature selection method, each interface dataset was divided into 2 sub-sets, a training subset (80% of the total dataset) and a validation subset (20% of the total dataset).

The first approach to select the most important features used the inter-intra (In-In) method (see 3.3) as evaluation criterion. The stop criterion was defined at the point where the evaluation criterion value is higher than 99.9% of the maximum of the evaluation criterion. Since this is a very basic feature selection evaluation, a wrapper based approach was also used.

Table 4.3 presents the selected features for each interface/database and for each feature selection method (using In-In method or wrapper method). The numbers presented in the table are the ranking order in the selected subset. So, the most important features are the ones that will be firstly chosen, and so on. If a feature has no indication of the ranking order, that feature and its samples were not used by that dataset. A *Not Applicable* feature is a feature that was not used by a specific dataset. This situation happens in the case of features that were only computed for the MA dataset and, therefore, are not used by the LI datasets.

4.3.1 Feature Selection using Rocha's Database

There are some comments for Rocha's database that can be derived from table 4.3:

- **LINW**: The three most important features selected by the In-In method were the F5, F12 and F23. These three were also selected by the wrapper approach. In fact, these two methods share the same most important feature, which is the F12. The wrapper approach selects all ones that use intensity and ICOV variation between the studied pixel and the lumen axis, the gradient features and some neighborhood features.
- **LIFW**: The In-In method selects 7 features (F2, F6, F11, F12, F20, F22 and F24), being F11 the most important of these. Besides this, it also selects some neighborhood, gradient and intensity features. The wrapper approach selects 9 features (F1, F2, F3, F8, F23, F24, F25 and F28). The most important selected feature using the wrapper approach is F8.
- **MANW**: To discriminate this interface, 10 features were selected using the In-In method. These were F3, F4, F7, F10, F16, F17, F18, F19, F23 and F32. The wrapper approach selected 7 features, F3, F12, F14, F15, F16, F23 and F30. Only 3 of them are coincident. The most important feature selected by the In-In method (F23) is the second most important one using the wrapper approach. The features selected by the In-In method are based on the ICOV, intensity variations between pixel and LI interface, column profile features and neighborhood information. The wrapper

Tab. 4.3: Selected features and its ranking order using the In-In method and the wrapper approach.

ID	Feature Name	Rocha's								Molinari's			
		LINW		LIFW		MANW		MAFW		LIFW		MAFW	
		In	Wr	In	Wr	In	Wr	In	Wr	In	Wr	In	Wr
F1	Gradient in		11		3							10	
F2	Gradient in Y		2	3	7								
F3	Fuzzy step edge				4	4	5			2		1	
F4	Fuzzy valley edge	4	6			8						5	
F5	Gradient orientation	3	15		9			4	5	3			
F6	Intensity magnitude			5							2	15	
F7	ICOV magnitude					10				5			
F8	Maximum intensity (px-LA)		13		1			3	6	1	1		
F9	Maximum ICOV (px-LA)		9								3		
F10	Mean intensity (px-LA)		7			2			8			2	
F11	Mean ICOV (px-LA)		8	1									
F12	σ intensity (px-LA)	1	1	4			7					7	
F13	σ ICOV (px-LA)		5							4			
F14	Maximum intensity (px-LI)	NA		NA		1				NA			1
F15	Maximum ICOV (px-LI)	NA		NA		4				NA			
F16	Mean intensity (px-LI)	NA		NA		5	6			NA		6	
F17	Mean ICOV (px-LI)	NA		NA		3			7	NA		13	
F18	σ intensity (px-LI)	NA		NA		6			1	NA			
F19	σ ICOV (px-LI)	NA		NA		9				NA			3
F20	Gaussian up 3	5	14	6									
F21	Gaussian down 3							2				16	
F22	Gaussian up 5		12	7						6		3	
F23	Gaussian down 5	2	10		2	1	2					14	
F24	Gaussian up 7		3	2	6			1	2			9	2
F25	Gaussian down 7				5							8	
F26	Gaussian left 3											11	
F27	Gaussian right 3												
F28	Gaussian left 5				8								
F29	Gaussian right 5											4	
F30	Gaussian left 7		4				3		3				
F31	Gaussian right 7								4				
F32	Gaussian below 20					7						12	

ID - Feature identification; In - In-In method; Wr - Wrapper approach; px - studied pixel; LA - lumen-axis; LI - Lumen-Intima; NA - Not Applicable.

approach also selects some variation features (between pixel and LI interface), a column profile feature and two features of neighborhood information.

- **MAFW**: The result of the feature selection methods for this dataset was the most similar. Three of the 8 features selected by the wrapper approach were common to the In-In method. The In-In method selected 4 features (F5, F8, F21 and F24). This interface requires less features to be discriminated in comparison with the MANW.

4.3.2 Feature Selection using Molinari's Database

Some comments for the Molinari's database can also be made from table 4.3:

- **LIFW**: Six features were selected for this interface using the In-In method. These

were F3, F5, F7, F8, F13 and F22. The most important selected feature is F8. The second most important feature is F3 and the third is F5. The features selected using the wrapper approach were F6, F8 and F9. These are features based in intensity and ICOV.

- **MAFW**: To discriminate this interface, the In-In method selected 16 features (F1, F3, F4, F6, F10, F12, F16, F17, F21, F22, F23, F24, F25, F26, F29 and F32). The most important ones are F3 and F10. Half of these features (F21, F22, F23, F24, F25, F26, F29 and F32) are based in the neighborhood information and the other half are based in intensity, ICOV, gradient and column profile. The wrapper approach selected only three features from this dataset. These were F14, F19 and F24. Only one of these three is common to the In-In selected features.

4.3.3 Feature Selection Summary

Table 4.3 can also be used to compare the feature selection differences between the two databases. For Molinari’s database, the wrapper approach used less features than the In-In method. For Rocha’s database, the opposite happened and the number of features selected using the wrapper approach was always higher than the number of features selected by the In-In method, except for the MANW dataset.

The few number of features selected by the wrapper approach for the Molinari’s database, could indicate an over-fitting and lack of generalization of the features to the specific training conditions. A reason to test the two feature selection methods (In-In and wrapper) is to assess if this over-fitting occurs or not. Besides these two methods, the NoFeatSel method (without feature selection) was also used to obtain the final results.

Another difference between the two studied databases is that, for Rocha’s database, the samples used at the training stage and the test stage are different. This difference is related to the lossy compression of these images.

The training stages for both databases only consider as interface candidates the pixels that are local maxima of the ICOV in the gradient direction. The test stage should consider the same candidates as the training stage, but, for Rocha’s database, this would result in a very noise sensitive method. In order to avoid this problem, the test stage for Rocha’s database consider all the pixels as interface candidates.

4.4 Classification Performance

To avoid an optimistically biased evaluation, a cross-validation method was used to evaluate the classifiers performance. In this validation method, a fraction of objects is used in the training subset and the remaining samples are used for testing. This fraction is rotated over the available set of objects and the performance results are averaged [Duin et al., 2007]. These results are used to estimate the classifiers error.

Tab. 4.4: Cross-validation mean error for the trained classifiers using the different feature selection methods. The best result for each classifier is written in bold and the best result for each interface is underlined in red.

Dataset	Classifier	Mean error		
		In-In (%)	wrapper (%)	total data (%)
Rocha's	LINW	KNNC	8.64	7.73
		SVC	10.23	9.13
		ParzenC	9.09	7.70
		QDC	10.45	<u>9.39</u>
	LIFW	KNNC	7.06	6.77
		SVC	9.12	7.43
		ParzenC	7.58	6.84
		QDC	10.01	5.89
	MANW	KNNC	9.33	9.77
		SVC	9.50	9.93
		Parzen	9.12	8.90
		QDC	9.28	9.33
	MAFW	KNNC	8.37	8.37
		SVC	10.34	10.02
		Parzen	7.55	10.34
		QDC	10.84	6.55
Molinari's	LIFW	KNNC	5.01	4.53
		SVC	6.26	4.94
		ParzenC	4.46	4.85
		QDC	6.26	<u>7.14</u>
	MAFW	KNNC	4.71	5.93
		SVC	4.18	9.93
		ParzenC	4.07	5.35
		QDC	5.73	4.20

The cross-validation method uses 5-folds, where one of the 5 partitions is used as test set and the other 4 are used as training set. The final classifier is the one with the lowest mean error obtained by the cross-validation. This evaluation performance is done for all the classifiers and feature selection methods. Since the cross-validation method has a high processing time, this method was performed using only half size of the datasets.

Table 4.4 presents the mean classification error (obtained from the 5 cross-validation training sets rotations) of cross-validation for each database/dataset, classifier and feature selection method. The best classifier of each feature selection method is highlighted in table 4.4. This table also presents the best classifier for each dataset underlined in red.

For the Rocha's LINW and MANW datasets, the ParzenC was the best classifier using the NoFeatSel method. For the LIFW and MAFW datasets, the best cross-validation mean error was obtained using the QDC. The best feature selection method for the LIFW was the wrapper and for the MAFW was the NoFeatSel method. For the Molinari's LIFW and MAFW datasets, the best cross-validation mean classification error was obtained using the combination of ParzenC with the NoFeatSel method.

For Rocha's database, the best cross-validation mean error for each feature selection method is as follows:

- **In-In method:** using the KNNC for the LINW and LIFW datasets and the ParzenC for the MANW and MAFW datasets.
- **Wrapper approach:** using the ParzenC for the LINW dataset, the QDC for the LIFW and MANW datasets and the KNNC for the MAFW dataset.
- **NoFeatSel method:** using the ParzenC for the LINW and MANW datasets and using the QDC for the LIFW and MAFW datasets.

For the Molinari's database, the best cross-validation mean error for each feature selection method is as follows:

- **In-In method:** using the ParzenC for the LIFW and MAFW datasets.
- **Wrapper approach:** using the ParzenC for the LIFW dataset and the QDC for the MAFW.
- **NoFeatSel method:** using the ParzenC for the LIFW and MAFW datasets.

From all the tested classifiers, the SVC was the only one that was not selected for any interface or method.

4.4.1 Classifiers Parameters

Each trained classifier is characterized by its type and some intrinsic parameters. This is presented in table 4.5.

Tab. 4.5: Chosen classifiers specifications

Database	Method	Classifier Parameters			
		LINW	LIFW	MANW	MAFW
Rocha's	In-In	KNNC k=1	KNNC k=1	ParzenC pw=0.106	ParzenC pw=0.082
	Wrapper	ParzenC pw=0.059	QDC m=9	QDC m=7	KNNC k=7
	NoFeatSel method	ParzenC pw=0.088	QDC m=25	ParzenC pw=0.157	QDC m=32
Molinari's	In-In	NA NA	ParzenC pw=0.087	NA NA	ParzenC pw=0.222
	Wrapper	NA NA	ParzenC pw=0.147	NA NA	ParzenC pw=0.247
	NoFeatSel method	NA NA	KNNC k=9	NA NA	QDC m=3

NA - Not Applicable

All the classifiers parameters were automatically optimized by the own classification functions. More information can be found in Duin et al. [2007].

4.4.2 Receive Operator Curve

The classifiers receive operator curve (ROC) for Rocha's database are presented in Fig. 4.5 (a)-(d). The classifiers ROC for the Molinari's database are presented in Fig. 4.5 (e) and (f).

For Rocha's database, the classifiers for the LINW interface detection (Fig. 4.5 (a)) that present a better ratio between the false negatives and false positives are the ones using the NoFeatSel method and the wrapper approach. For the LIFW (Fig. 4.5 (b)), this relation is not so significant and no accurate conclusion follows from these curves, except that both methods have similar classification errors. The best classifier for the MANW (Fig. 4.5 (c)) used the NoFeatSel method and for the MAFW (Fig. 4.5 (d)) used a wrapper approach.

For Molinari's database, the classifiers for the LIFW interface detection (Fig. 4.5 (e)) and for the MAFW interface detection (Fig. 4.5 (f)) presented very similar ROCs. For the LIFW dataset, the best classifier used the wrapper approach. The worst classifier for this interface is the NoFeatSel method. The best classifier for the MAFW is the one using the wrapper approach.

4.4.3 Classifier Validation

The last step of the trained classifier evaluation is its validation. This is performed using the 20% of samples selected from the total dataset. This sub-set was not involved in any of the previous stages. By applying this sub-set, we obtain the classifier behaviour with new data.

Table 4.4.3 presents these results. Each of them are studied according to the initial feature selection method.

Tab. 4.6: Classifiers validation error. The bold values represent the best validation error for each dataset and database.

Database	Feature Selection	Validation Error (%)			
		LINW	LIFW	MANW	MAFW
Rocha's	In-In	9.26	8.85	8.68	10.53
	Wrapper	8.10	5.90	8.19	11.68
	NoFeatSel method	8.42	7.37	8.19	12.17
Molinari's	In-In	NA	8.3	NA	8.82
	Wrapper	NA	5.93	NA	11.76
	NoFeatSel method	NA	6.32	NA	8.82

For Rocha's database, the best validation error was obtained for the LIFW using the wrapper approach. This feature selection method resulted in 3 of the best validation errors.

The best validation error for Molinari's database was obtained using the wrapper approach for the LIFW dataset. The best training error was obtained using the NoFeatSel

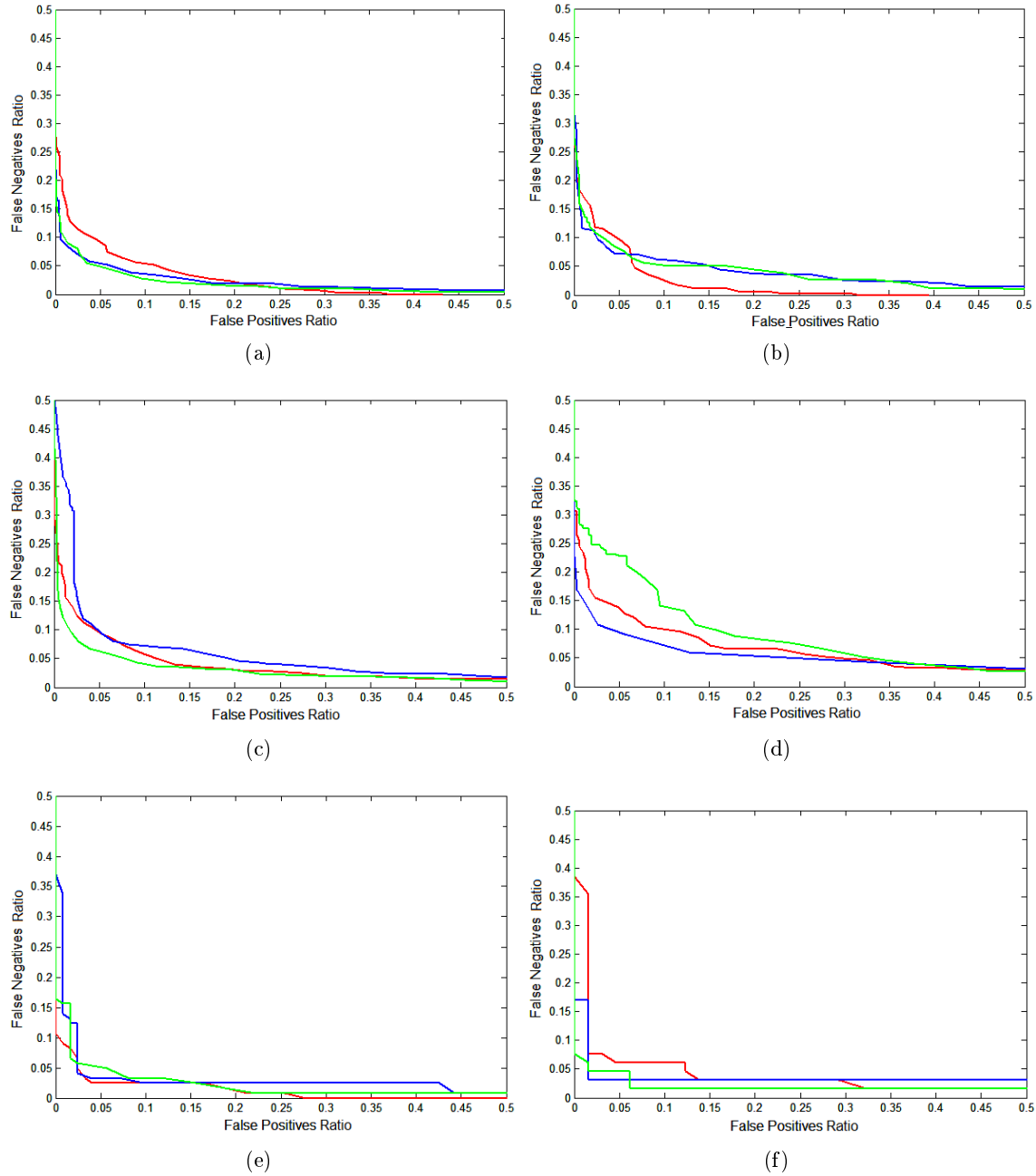


Fig. 4.5: Classifier's ROC for Rocha's and Molinari's database: (a) classifier of LINW, (b) LIFW, (c) MANW and (d) MAFW of Rocha's database; (e) and (f) are the classifiers ROC of LIFW and MAFW, respectively, for Molinari's database. The ROCs obtained with the In-In method are presented in red, the ROCs obtained with the wrapper approach are presented in blue and the ROCs obtained without the feature selection method are presented in green.

method. For the MAFW dataset, the validation using the In-In approach and the NoFeat-Sel method were equal and better than the one obtained with the wrapper approach.

4.5 Classification Output

After the feature measurement and the classification performance evaluation, the classification output is obtained. Figure 4.1 shows the classification block enclosed to the complete methodology.

This section is divided into the score maps presentation and the dynamic contours definition, which includes the assessment of the geometric coefficient and the final contour construction. An example of these steps is given for Rocha's and Molinari's databases using the three evaluated feature selection methods.

4.5.1 Score Maps

The classification output is a probabilistic score map which translates the pixel likelihood of belonging to the searched interface. Four score maps are obtained for Rocha's database and two for Molinari's database.

Figure 4.6 shows the score maps for one image of Rocha's database, using the In-In method (Fig. 4.6 (a) and (b)), the wrapper approach (Fig. 4.6 (c) and (d)) and the NoFeatSel method (Fig. 4.6 (e) and (f)). The differences between these three score maps are significant. The LINW score map obtained using the wrapper approach is similar to the one obtained by the NoFeatSel method. The LINW and LIFW score maps obtained using the In-In method presented much noise in the surrounding tissue and the three MANW and MAFW score maps are quite different.

Figure 4.7 shows the score maps for one image of Molinari's database, using the In-In method (Fig. 4.7 (a) and (b)), the wrapper approach (Fig. 4.7 (c) and (d)) and the NoFeatSel method (Fig. 4.7 (e) and (f)) to perform the feature selection. As is shown in Fig. 4.7, the score map for the LIFW interface obtained with the In-In method has low detection of tissue below the interface and its interface detection is similar to the one obtained with the NoFeatSel method. The wrapper approach results in a medium response to the tissue below the interface and it is the score map with the worst result. For the MAFW interface, the In-In method gives the best MA detection, in comparison with the others two. The wrapper approach has a response similar to the one obtained using the NoFeatSel method. The last method is the one that detects more tissue below the searched interface. These figures also show that the MAFW interface and the muscle bundle place below the MAFW have equivalent responses because they have similar anatomical structures.

These differences occurs in the minority of the images. This happens for 9 images of the 27 tested from Rocha's database and 4 of the 90 tested from Molinari's database. Appendix B shows some of these images.

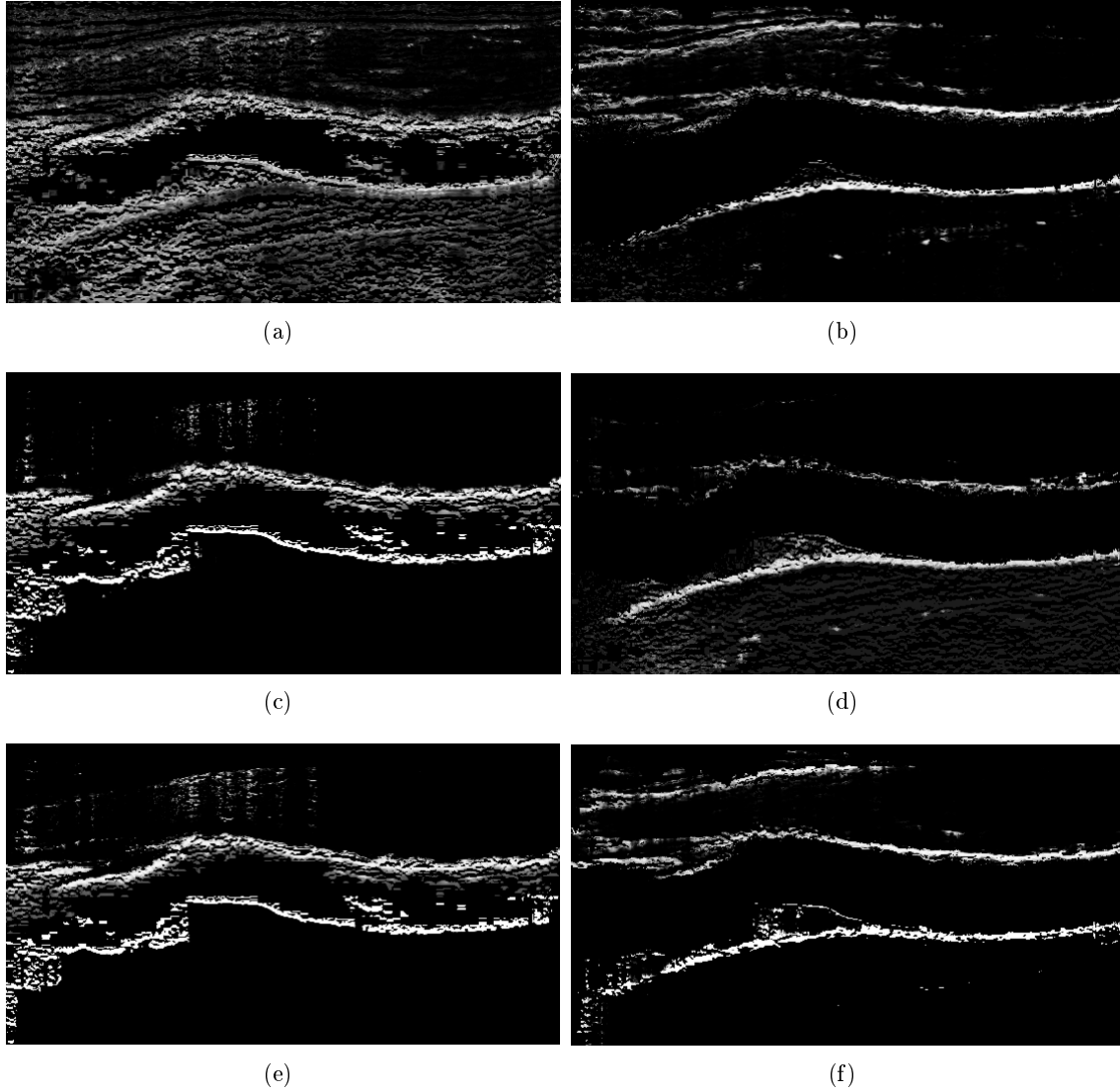


Fig. 4.6: Score maps of LIFW (left column) and MAFW (right column) for one image of Rocha's database using the In-In method ((a) and (b)), the wrapper approach ((c) and (d)) and the NoFeatSel method ((e) and (f)). The pixel intensities represent their probability of belonging to the searched interface.

4.5.2 Dynamic Programming

The fourth stage of the developed method is the application of the dynamic programming function to the obtained score maps (Fig. 4.1)

The dynamic programming used in Rocha et al. [2011] combines the score map with a geometric smoothness information. The definition of the geometric coefficients (3.26) for each interface of each database is presented below. The final contour obtained using the three different feature selection methods is also presented.

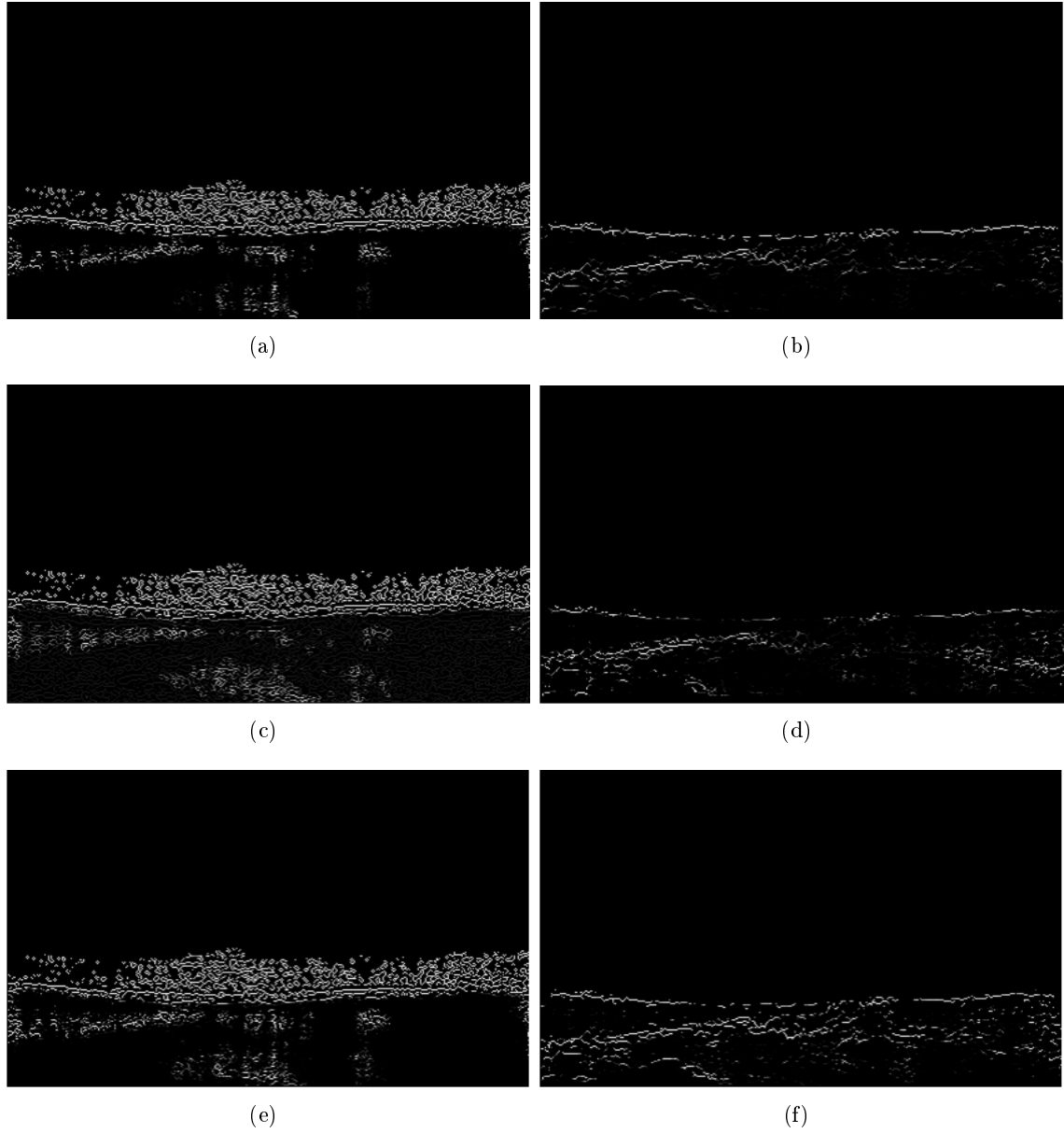


Fig. 4.7: Score maps of LIFW (left column) and MAFW (right column) for one image of Molinari's database using the In-In method ((a) and (b)), the wrapper approach ((c) and (d)) and the NoFeatSel method ((e) and (f)). The pixel intensities represent their probability of belonging to the searched interface.

Geometric Coefficients Definition

The dynamic programming geometric coefficients were chosen by evaluating the mean distance error obtained for each coefficient using the test set. Table 4.5.2 shows the chosen coefficients.

For Rocha's database, the geometric coefficients were evaluated between 0 and 55 and for Molinari's database, these were evaluated between 0 and 100. The difference between

evaluated ranges was caused by computational limitations related with the test stage for Rocha's database. As it was already mentioned, the test stage for Rocha's database use all the image pixels as interfaces candidates, while the test stage of Molinari's database only uses the local maxima of the ICOV in the gradient direction pixels as interfaces candidates, making the Rocha's database test stage much more time-consuming than the test stage of Molinari's database.

The plots presented in Fig. 4.8 (a) - (d) represent the error evolution as the geometric coefficient increases for Rocha's database and Fig. 4.8 (e) and (f) represent the error evolution as the geometric coefficient increases for Molinari's database.

Table 4.5.2 presents the selected geometric coefficients.

Tab. 4.7: Geometric coefficients used by the dynamic programming function for each database

Database	Geometric coefficient			
	LINW	LIFW	MANW	MAFW
Rocha's	0	5	50	55
Molinari's	NA	90	NA	95

As table 4.5.2 shows, the geometric coefficients selected for Molinari's database are higher than the ones selected for Rocha's database. This is caused by the high sensibility of Rocha's method to noise in comparison with the Molinari's method and also because Molinari's database has few images with plaques or large IMTs, unlike Rocha's database, leading to a higher smoothness of the contour.

After defining the geometric coefficient used by the dynamic programming function for each interface, the final contours are drawn.

Final Interfaces

The detection of the LI and MA interfaces is the last step of the proposed method (Fig. 4.1).

The final interfaces for an image of Rocha's database, using the different feature selection methods, are presented in Fig. 4.9. Figure 4.9 (a) presents the result obtained using the In-In method, Fig. 4.9 (b) reports the result using the wrapper approach and Fig. 4.9 (c) presents the result obtained using the NoFeatSel method.

For the image presented in Fig. 4.9, the results obtained with the NoFeatSel method and with the wrapper approach are similar and are better than the result of the In-In method. This method is worse in the detection of the plaque. It is important to say that the results presented in Fig. 4.9 are illustrative. In other images the differences between methods are less significant or even negligible. Appendix B shows the images where the difference between methods were more significant.

The final interfaces for an image of Molinari's database, using the different methods, are

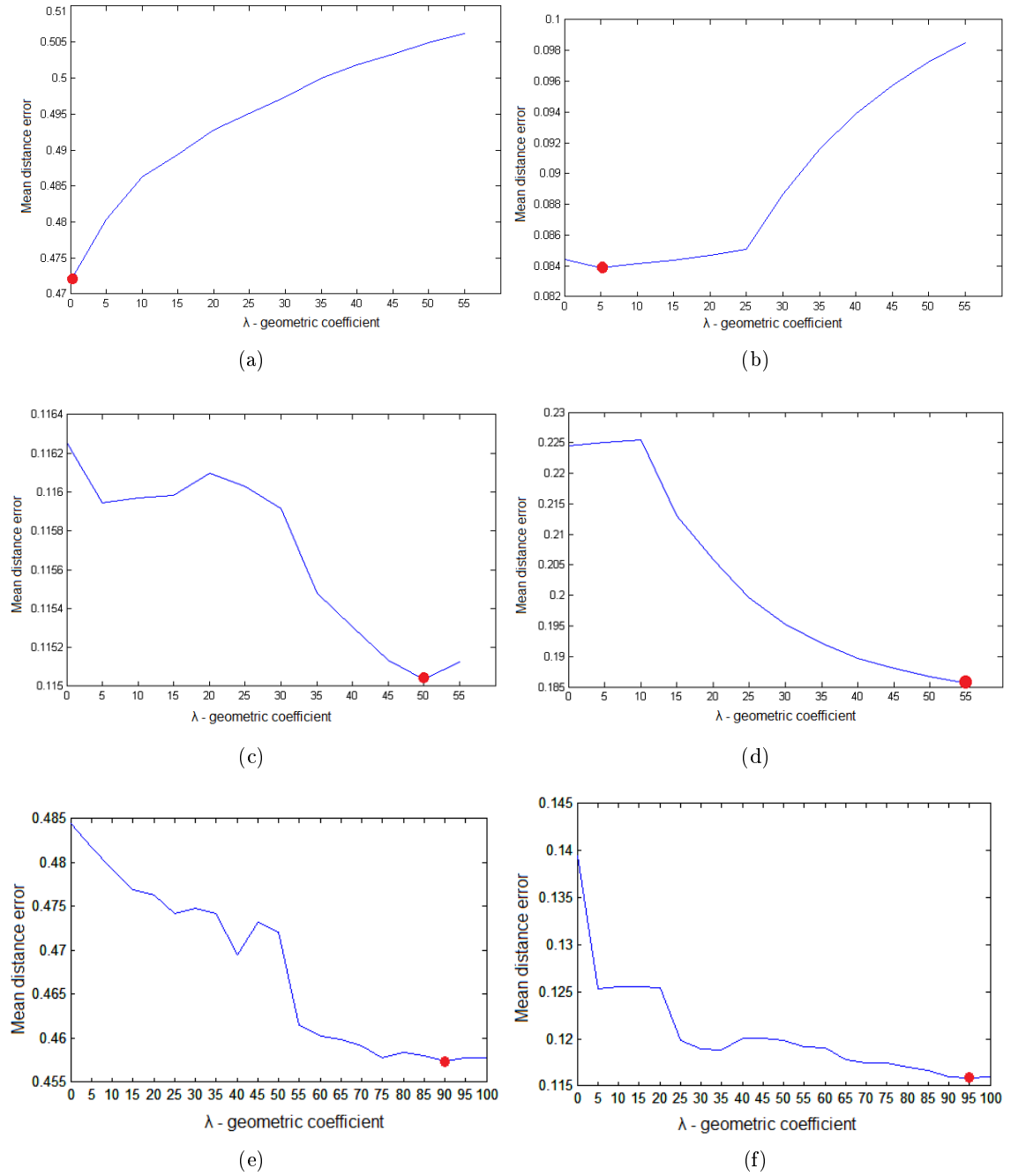


Fig. 4.8: Error and geometric coefficient, λ , for the dynamic programming using Rocha's database ((a)-(d)) and Molinari's database ((e) and (f)): (a) $\lambda = 0$ for LINW, (b) $\lambda = 5$ for LIFW, (c) $\lambda = 50$ for MANW and (d) $\lambda = 55$ for MAFW; (e) $\lambda = 90$ for LIFW and (f) $\lambda = 95$ for MAFW. The red dot corresponds these points.

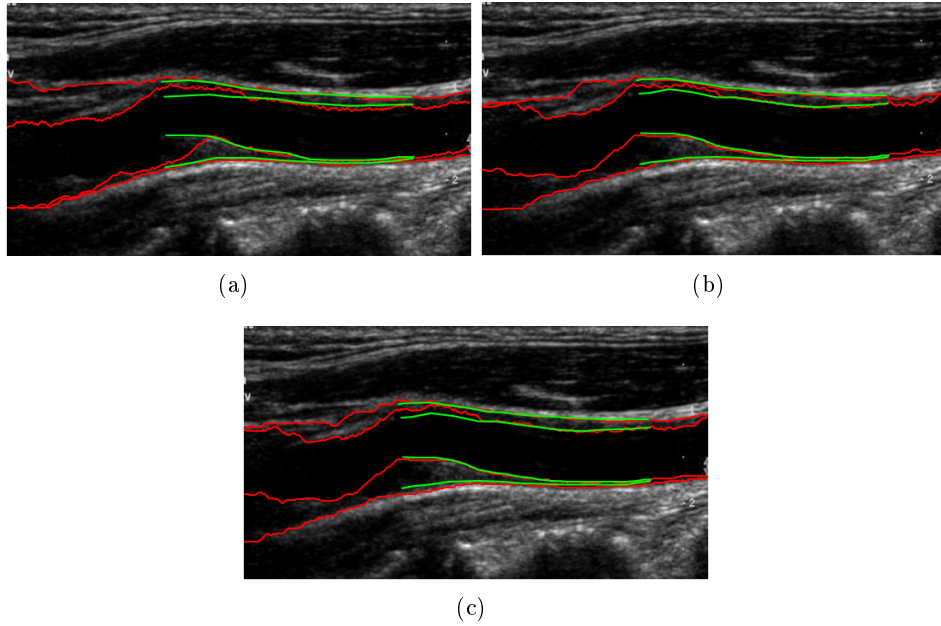


Fig. 4.9: Final interfaces obtained with different feature selection method for Rocha’s database: (a) In-In method, (b) wrapper approach and (c) NoFeatSel method. The green contours are the MA1 traces and the red contours are the ones obtained by the proposed method.

presented in Fig. 4.10. Figure 4.10(a) presents the result obtained using the In-In method, Fig. 4.10 (b) shows the result using the wrapper approach and Fig. 4.10 (c) presents the result obtained using the NoFeatSel method.

For the presented image (Fig. 4.10), the results obtained with the NoFeatSel method is the worst of the three. The interface obtained with the In-In method (Fig. 4.10) (b) is also miss classified. The motive for this difference relies on the different score maps. The differences between methods for the majority of Molinari’s database is small. Appendix B shows the images where the difference between methods were more significant.

4.6 Statistical Analysis

In this section the statistical analysis of the final test results is presented. These results are presented in millimeters and in percentages. The MAD, MSE, CV, FOM and classification performance are measured. The results are also presented using Bland-Altman plots.

The developed method was tested in 65 images from Molinari’s database and 27 images from Rocha’s database. These images are different from the ones used for the training procedure. The results obtained for the Molinari’s database are compared with the respective ground-truth (MM1). The results obtained for Rocha’s database are compared with the three given ground-truths (MA1, MA2 and MB1).

The automatic detection of the IMT failed completely (IMT error higher than 1 mm)

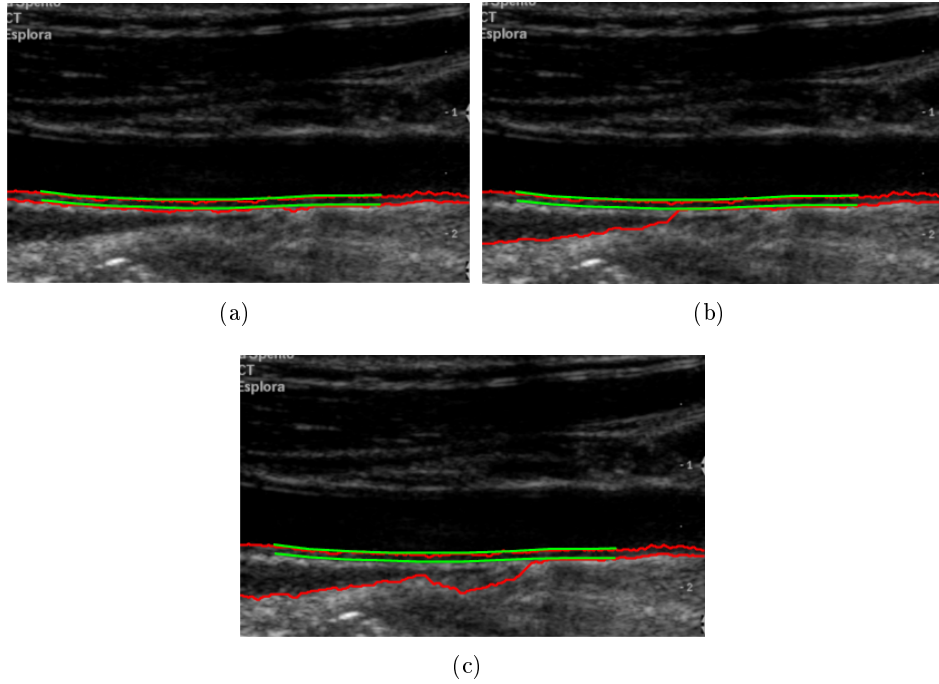


Fig. 4.10: Final interfaces obtained with different feature selection method for Molinari's database: (a) In-In method, (b) wrapper approach and (c) NoFeatSel method. The green contours are the MA1 traces and the red contours are the ones obtained by the proposed method.

in one of Rocha's images, using the NoFeatSel method approach (Fig. 4.11 (a)). For Molinari's database, only one IMT was miss-detected and it occurred for the wrapper approach (Fig. 4.11 (b)). The presented statistical results do not consider these two failed detections.

This method was compiled using Matlab 2011b, on a computer equipped with an Intel Core i7 quad core processor at 2.67 GHz. The mean time, t , required to segment an image was 1.73 min for Rocha's database and 20.49s for Molinari's database. The significant difference between the computational times was caused because the Molinari's test stage only considers the pixels that are local maxima of the ICOV in the gradient direction as interface candidates and the Rocha's test stage considers all the pixels as interface candidates.

4.6.1 Performance Measurements for Interfaces

This section presents the MAD, MSE, CV and FOM performance measures used to evaluate the proposed method. These measures were used to evaluate the inter-method (automatic versus manual operators), inter-observer (manual versus manual between two different operators) and intra-observer (manual versus manual, same operator on different periods) variability. The measurement of the inter-observer and intra-observer variability is only possible for Rocha's database, since it has three ground-truths (GT), being MA1 and MB1

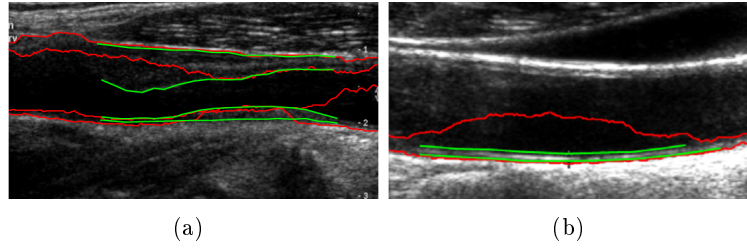


Fig. 4.11: Failed detections for Rocha's and Molinari's databases: (a) represent the failed detection for Rocha's database using the NoFeatSel method and (b) represent the failed detection obtained for Molinari's database using the wrapper approach.

from the same operator.

Table 4.8 presents the results for the performance measurement between interfaces. Each pixel of the automatic contour is compared with the corresponding one of the GT. This type of statistics allows an accurate comparison between pixel pairs and of its agreement with the different GTs. The best FOM results for each interface are presented in green and the worst in red.

For Rocha's database, the obtained FOMs were very good. The worst FOM for the LINW was 95.82% using the NoFeatSel method approach and the best was 96.89% using the In-In method. For the LIFW, the worst FOM was 98.48% using the In-In method and the best was 98.98% using the NoFeatSel method approach. This result was expected since the LIFW has higher contrast and well defined edges than the LINW. The performance presented by the NoFeatSel method was also expected because, since it uses all the features, the classification is more adaptable to different images (like the ones with and without plaque). For the MANW, the worst FOM was 96.95% using the In-In method and the best was 97.80% using the NoFeatSel method approach. For the MAFW, the worst obtained FOM was 97.72% using the NoFeatSel method approach and the best was 98.72% obtained using the wrapper approach. The maximum difference, for Rocha's database, between the best and worst FOMs does not exceed 1.07% which indicates that the developed methods have similar behaviours in the interface detection.

The worst FOM obtained for the LIFW was 98.52% using the In-In method and the best was 98.60%, using the wrapper approach. For the MAFW, the worst FOM was 98.73%, using the NoFeatSel method and the best was 98.97% using the In-In and wrapper methods. The performance values for the MAFW are similar to the ones obtained for the LIFW. The referred best results are presented in table 4.8 in green and the worst are presented in red.

4.6.2 Performance Measurements for the IMT

Table 4.9 presents the results obtained for the IMT mean, maximum, minimum and the coefficient of variation for the mean IMT (CVIMT) using the three different feature

Tab. 4.8: Performance measurements (mm) for interfaces for pixel pairs using different features selection methods. The best FOM results for each interface are presented in green and the worst in red.

Database	Comparison		LI				MA			
			MAD	MSE	CV(%)	FOM(%)	MAD	MSE	CV(%)	FOM(%)
Rocha's	(MA1,MA2)	NW	0.178	0.251	1.608	97.70	0.134	0.255	2.25	98.03
		FW	0.098	0.133	0.504	99.22	0.111	0.144	0.467	99.18
	(MA1,MB1)	NW	0.138	0.197	1.282	98.21	0.133	0.210	1.681	98.05
		FW	0.092	0.132	0.534	99.27	0.0871	0.132	0.513	99.35
	(MA2,MB1)	NW	0.120	0.190	1.33	98.45	0.131	0.192	1.457	98.09
		FW	0.079	0.122	0.522	99.37	0.1414	0.179	0.710	99.16
	In-In method									
	Rocha's	(MA1,Auto)	NW	0.263	0.403	2.75	96.62	0.192	0.309	2.53
FW			0.159	0.300	1.42	98.75	0.219	0.361	1.481	98.41
(MA2,Auto)		NW	0.242	0.400	2.88	96.89	0.204	0.328	2.687	96.95
		FW	0.193	0.304	1.30	98.48	0.202	0.364	1.562	98.53
Molinari's	(MB1,Auto)	NW	0.252	0.408	2.91	96.76	0.200	0.313	2.507	97.02
		FW	0.18	0.341	1.43	98.59	0.243	0.398	1.632	98.23
	(MM1,Auto)	FW	0.172	0.226	0.888	98.52	0.127	0.151	0.475	98.97
		wrapper approach								
Rocha's	(MA1,Auto)	NW	0.292	0.618	5.021	96.17	0.166	0.268	2.178	97.57
		FW	0.158	0.245	1.087	98.83	0.176	0.243	0.86	98.71
	(MA2,Auto)	NW	0.296	0.624	5.020	96.12	0.188	0.268	1.977	97.26
		FW	0.169	0.244	0.975	98.66	0.174	0.254	0.958	98.72
Molinari's	(MB1,Auto)	NW	0.300	0.627	5.031	96.05	0.0175	0.250	1.852	97.44
		FW	0.152	0.247	1.081	98.79	0.191	0.260	0.909	98.60
	(MM1,Auto)	FW	0.163	0.200	0.704	98.60	0.126	0.247	1.229	98.970
		NoFeatSel method								
Rocha's	(MA1,Auto)	NW	0.300	0.649	5.29	96.05	0.148	0.278	2.44	97.80
		FW	0.136	0.234	1.06	98.92	0.292	0.356	1.04	97.87
	(MA2,Auto)	NW	0.311	0.659	5.30	95.91	0.184	0.340	2.96	97.26
		FW	0.144	0.223	0.95	98.86	0.259	0.336	1.093	98.11
Molinari's	(MB1,Auto)	NW	0.318	0.662	5.31	95.82	0.175	0.306	2.59	97.40
		FW	0.129	0.223	1.02	98.98	0.313	0.378	1.075	97.72
Molinari's	(MM1,Auto)	FW	0.171	0.225	0.885	98.53	0.157	0.257	1.167	98.73

selection methods and the two databases. In table 4.9, the best results are presented in green and the worst are presented in red.

Using Rocha's database, the best CVIMT for the NW was 1.93% (MB1,Auto - using the wrapper approach) and the worse was 11.22% (MA1,Auto - using the NoFeatSel method). For the FW, the best CVIMT result was 0.96% (MB1, Auto - using the wrapper approach) and the worse was 4.66% (MA1,Auto - using the In-In method). The method that presented the best performance for Rocha's database was the wrapper approach.

Tab. 4.9: Performance measurements for mean, max and minimum IMT measures (mm) and coefficient of variation for the mean IMT (%) using different features selection methods. The best mean CVIMT results for each wall are presented in green and the worst in red.

Database			IMT min	IMT mean	IMT max	CVIMT(%)
Rocha's	(MA1,MA2)	NW	0.182	0.12	0.043	3.76
		FW	0.080	0.116	0.141	4.66
	(MA1,MB1)	NW	0.090	0.062	0.274	1.90
		FW	0.079	0.019	0.035	3.69
	(MA2,MB1)	NW	0.092	0.061	0.264	2.01
		FW	0.036	0.116	0.162	4.30
In-In method						
Rocha's	(MA1,Auto)	NW	0.040	0.193	0.132	8.32
		FW	0.028	0.099	0.1147	4.66
	(MA2,Auto)	NW	0.197	0.073	0.155	6.43
		FW	0.107	0.058	0.237	3.69
	(MB1,Auto)	NW	0.108	0.134	0.402	7.39
		FW	0.107	0.096	0.087	4.30
Molinari's	(GT,Auto)	FW	0,186	0,054	0,203	9.37
wrapper approach						
Rocha's	(MA1,Auto)	NW	0.069	0.070	0.427	2.90
		FW	0.028	0.022	0.054	1.87
	(MA2,Auto)	NW	0.240	0.047	0.400	2.47
		FW	0.110	0.128	0.167	2.26
	(MB1,Auto)	NW	0.149	0.028	0.200	1.93
		FW	0.107	0.016	0.090	0.96
Molinari's	(GT,Auto)	FW	0,236	0,093	0,201	17.58
NoFeatSel method						
Rocha's	(MA1,Auto)	NW	0.076	0.262	2.553	11.22
		FW	0.029	0.031	0.054	2.04
	(MA2,Auto)	NW	0.255	0.369	2.526	8.86
		FW	0.109	0.121	0.167	1.70
	(MB1,Auto)	NW	0.163	0.313	2.326	9.52
		FW	0.108	0.023	0.090	1.16
Molinari's	(GT,Auto)	FW	0,188	0,080	0,254	18.21

For Molinari's database, the best CVIMT was 9.37%, using the In-In method, and the worst was 18.21%, using the NoFeatSel method.

Table 4.9 shows that the best CVIMT was obtained using the wrapper approach in Rocha's database and the In-In method in Molinari's database. Therefore, we can conclude that the best method to classify the US carotid images of Rocha's database is based on a wrapper approach, using a ParzenC for the LINW dataset, a QDC for the LIFW and MANW datasets and a KNNC for the MAFW dataset (see table 4.5 and 4.4). The best

method to classify US carotid images of Molinari’s database is the In-In method using a ParzenC for the LIFW and MAFW datasets (see table 4.5 and 4.4).

4.6.3 Bland-Altman Plots

Besides the above mentioned measurements, the Bland-Altman plots can also be used to evaluate the proposed method. The Bland-Altman plots are presented for the selected methods.

Bland-Altman plots, like the one presented in Fig. 4.12, represent the IMT mean differences between the two different segmentations methods against the average of those values.

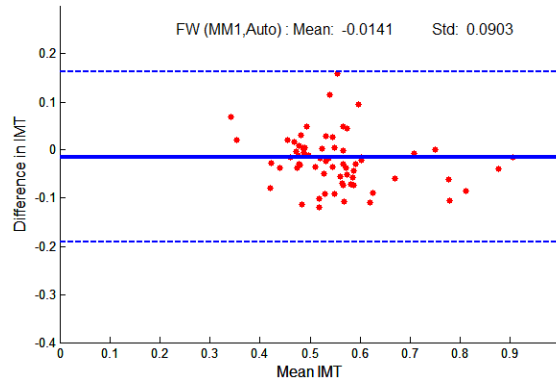


Fig. 4.12: Bland-Altman plot for the comparison of mean IMT measurements for Molinari’s database using In-In method. *Mean* stands for the IMT mean value of the two segmentation methods and *STD* is their standard deviation.

Figure 4.12 represents the Bland-Altman plot for Molinari’s database. Since this database only has one GT, that makes the inter-operator variability evaluation impossible. This graphic shows that the developed method does not generate outliers, its mean is around 0 mm and the standard deviation is 0.09 mm.

The same evaluation is performed for Rocha’s database, but, in this case, both the inter-observer and inter-method evaluation are possible. Figure 4.13 presents the inter operator Bland-Altman plots. These plots compare the IMT measurements from two traces for each wall.

The two GT, for the NW and the FW, that are more related are the (MA1,MB1) and the less related are the (MA1,MA2). The maximum absolute mean for the NW was 0.12 mm with a standard deviation of 0.060 mm and the minimum was 0.0059 mm with a standard deviation of 0.035 mm. The maximum absolute mean for the FW was 0.116 mm with a std of 0.021 mm and the minimum was 0.0025 mm with a std of 0.025 mm.

Figure 4.14 presents the Bland-Altman plots obtained for the inter-method variability analysis.

The two mean IMT measures, between two methods, for the NW and the FW that are

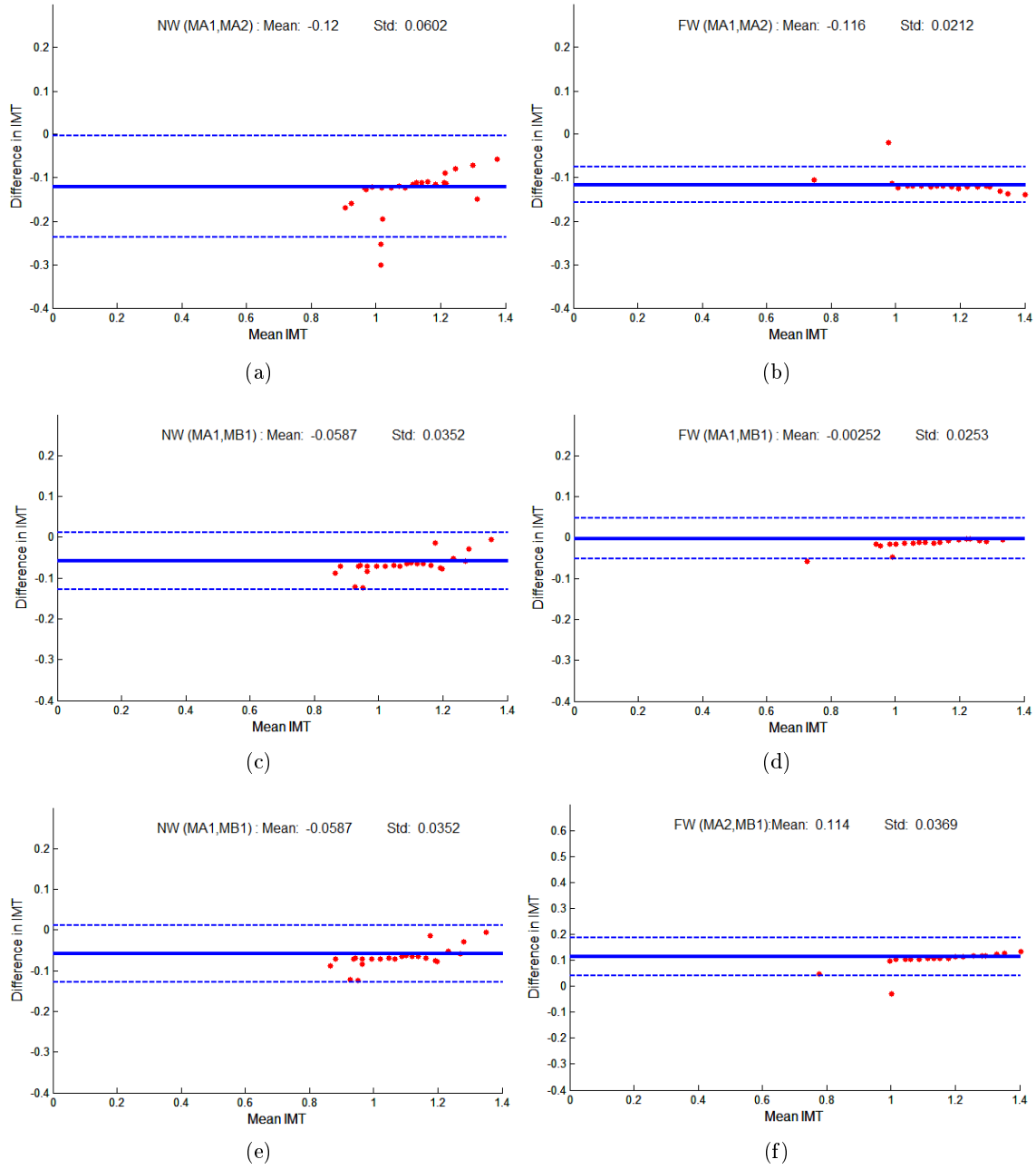


Fig. 4.13: Bland-Altman plots for the comparison between GT IMT measurements at the NW (left column) and FW (right column) in Rocha's database. (MA1,MA2) (a) and (b); (MA1, MB1) (c) and (d); (MA2,MB1) (e) and (f).

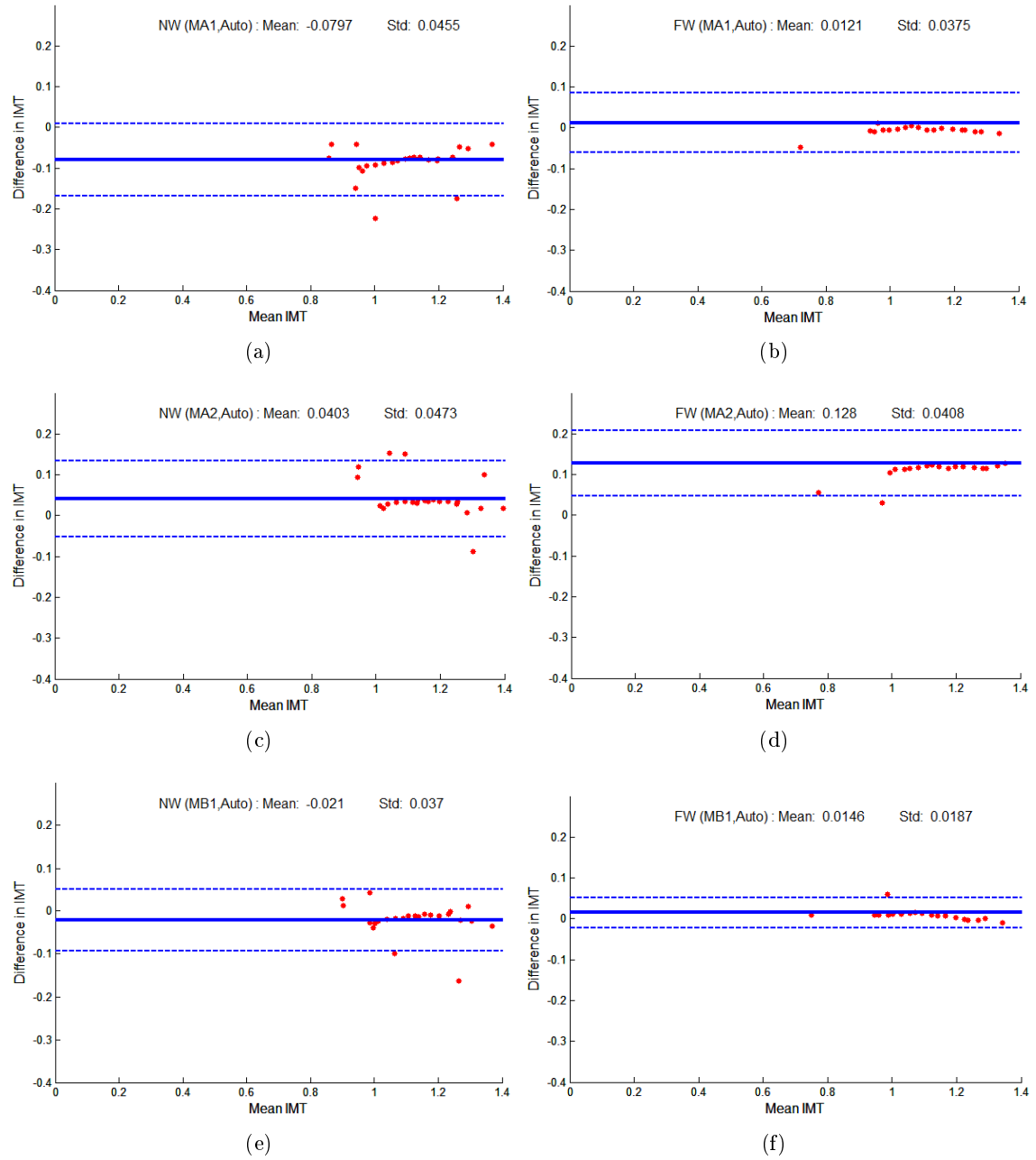


Fig. 4.14: Bland-Altman plots for the comparison between automatic and manual operators for Rocha's database of near wall (left column) and far wall (right column). (MA1,Auto) (a) and (b); (MA2,Auto) (c) and (d); (MB1, Auto) (e) and (f).

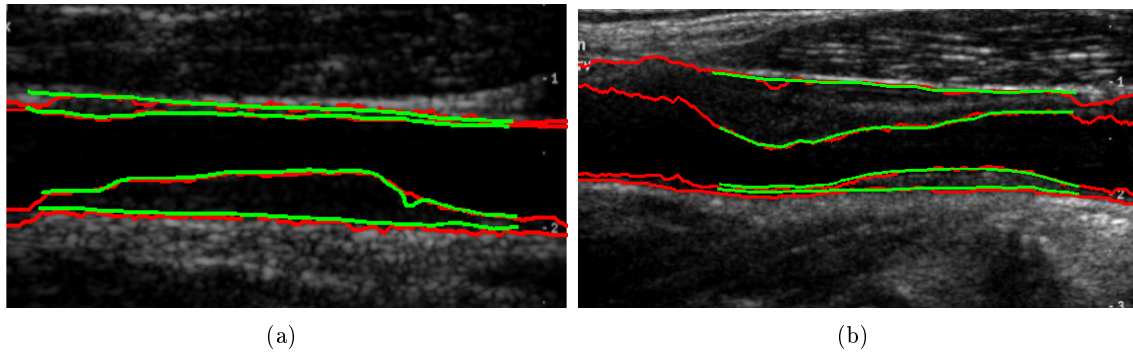


Fig. 4.15: Rocha’s database. (a) and (b) represent two good results. The green contour is the MA1 manual contour and the red contour is the one obtained by the proposed method.

more related are the (MB1,Auto) ones and the two less related are the (MA1,Auto) ones for the NW and the (MA2, Auto) for the FW. The maximum absolute mean of the NW was 0.080 mm with an standard deviation of 0.050 mm and the minimum was 0.021 mm with a standard deviation of 0.037 mm. The maximum absolute mean for the FW was 0.128 mm with a standard deviation of 0.041 mm. It is important to refer that the mean of the inter-observer variability between the MA1 and the MA2 was also the higher and had a mean of 0.116 mm with a standard deviation of 0.0212 mm. The minimum absolute mean for the FW was 0.0121 mm with a standard deviation of 0.038 mm.

The inter-method variability and the inter-operator variability are in the same range. The manual contour more related with the developed method is the MB1. The mean of the inter-observer variability between the MA1 and the MB1 is -0.0587 mm with a standard deviation of 0.0352 mm whereas the mean of the inter variability between the MB1 and the developed method is 0.021 mm, with a standard deviation of 0.037 mm. The higher inter method variability occurred between the MA2 and the automatic method, showing a mean of 0.128 mm and a standard deviation of 0.0408 mm.

These results indicate that the developed method has a variability similar to the one that occurs between two medical experts.

4.6.4 Good and Bad Results

After the statistical analysis of the results, the best method is selected and some good and bad results are shown. For Rocha’s database, the best method is the one using the wrapper approach while for Molinari’s database the best one is the In-In method.

As an example, Fig. 4.15 shows two good detections. These two good examples present two images with plaques, which are them selves difficult images.

Figure 4.16 presents two bad detections of interfaces with a different cause of miss-detection. The MA interface was wrongly detected in Fig. 4.16 (a) due to a small calcification in the intima media complex. The resultant score map (Fig. 4.16 (b)) shows

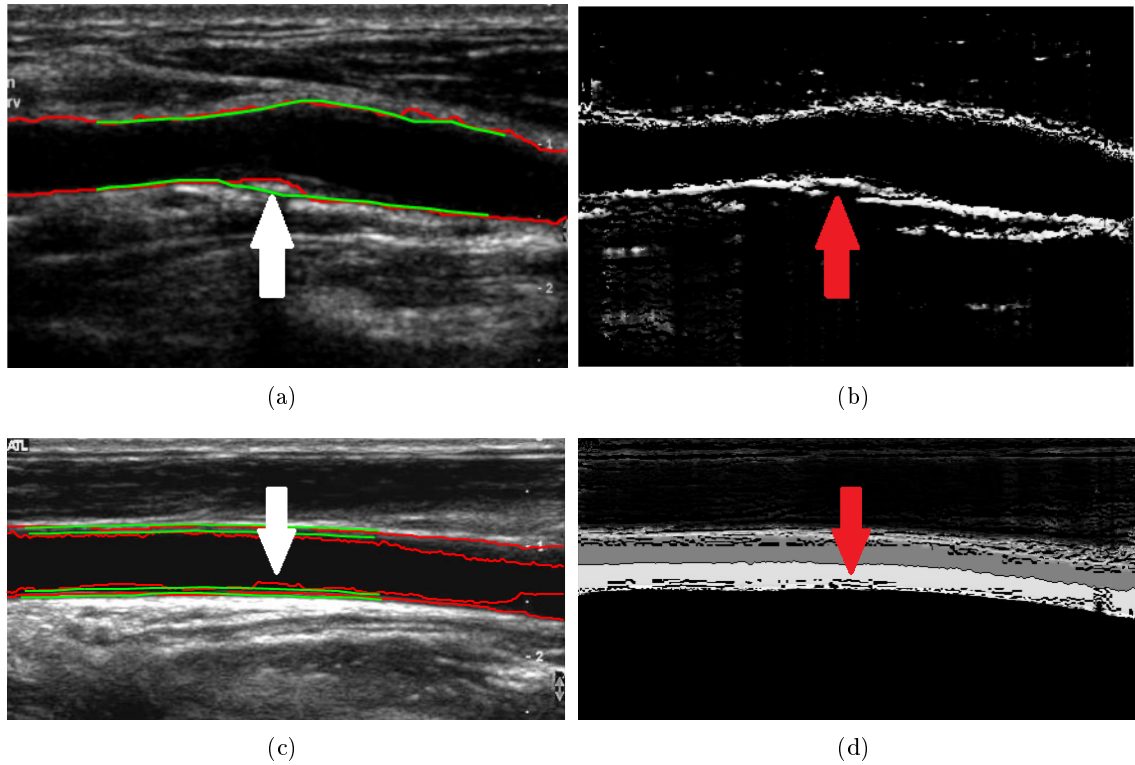


Fig. 4.16: Justification for bad detection in Rocha’s database: (a) and (c) represent the original images that caused the segmentation problems and (b) and (d) show the justification for these problems. These problems are caused by high echogenic structures and low adaptability of the classifier for different images.

the resultant shadow marked with an arrow. Figure 4.16 (c) presents an image that is very different from the others in Rocha’s database. The wrong classification in this image happens because the lumen has higher intensity values than the lumen of the majority of the other images. The lumen pixels are understood by the classifier as belonging. As a result, the score map presents pixels at the lumen with high probability of belonging to the LI class. Figure 4.16 (d) reveals this classifiers mistake (marked with an arrow). This problem could have been corrected if the training set had more images like this one.

Figure 4.17 (a) and (b) show two good results obtained with the In-In method for Molinari’s database. These figures show that the interfaces were correctly detected, even in diagonal arteries and in cases with a small IMT.

Figure 4.18 presents two bad interface detections. In this case, the IMT was underestimated in Fig. 4.18 (a). This was caused by the muscle highlighted in the NW. The resultant score map (Fig. 4.18 (b)) shows high noise below the searched interface. Figure 4.18 (c) presents an image with some noise in the lumen region. This results in a poor score map, presented in Fig. 4.18 (d). This shows that the developed method is sensitive to the lumen noise.

Appendix C shows more examples interfaces detection for both databases.

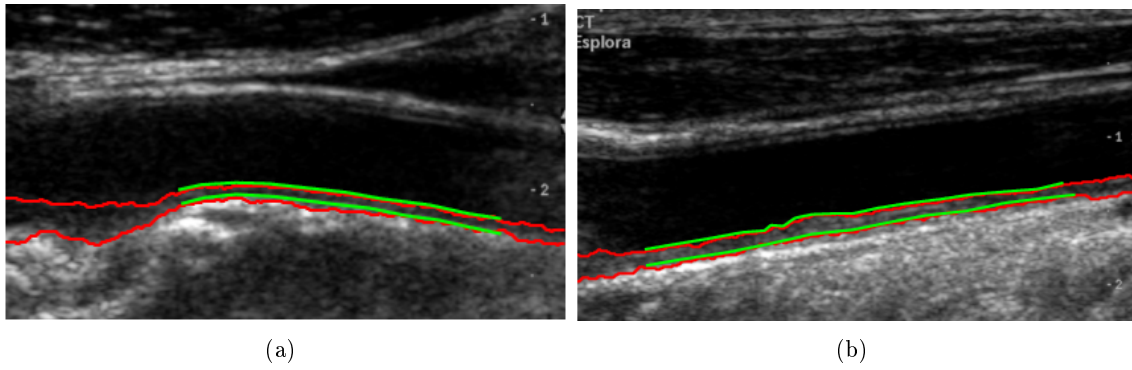


Fig. 4.17: Molinari's database. (a) and (b) represent two good results and (c) and (d) represent two bad results obtained using Molinari's database. The green contour is the MA1 manual contour and the red contour is the one obtained by the proposed method.

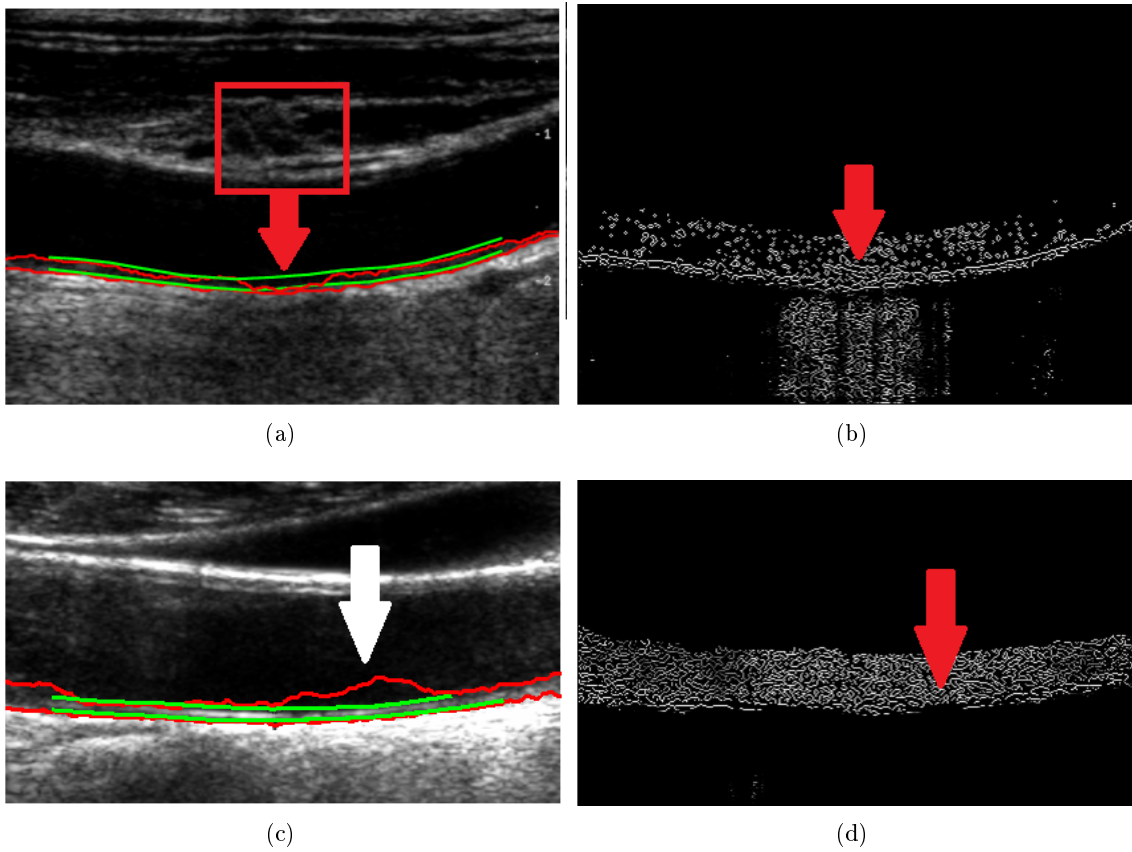


Fig. 4.18: Justification of bad results for Molinari's database: (a) and (c) represent the original images that caused the segmentation problems and (b) and (d) represent the justification to these problems. These problems are caused by an agglomerate of structures above the searched interface and sensitivity of the detection method to lumen noise.

Conclusion

The main goal of this work was the development of a classification approach to classify artery interfaces, in order to diagnose the arteriosclerosis. This classification approach was included, as an alternative interfaces classification method, in a previous work developed by Rocha et al. [2012] at the INEB - Bioimaging group. We tried to develop an approach as little empirical as possible to avoid subjectivity problems, like incompatibility to other databases or hardware.

The developed method is a new technique for the arteriosclerosis diagnosis since it uses a pattern recognition approach, which is rarely used by the state-of-the-art articles.

Two of the most important issues are the image quality of Rocha's database and the image compression using JPEG. This type of images does not return a good output to the interface detection because the developed method is sensitive to noise and more precisely to the lumen noise. The compression of these images result in low image resolution and in the pixelization effect. As in Molinari's database the images were not compressed with JPEG, the classification results were not affected by these drawbacks.

A significant disadvantage of the proposed method is the binary classification performed by the trained classifiers. This is not the better way to perform the interfaces classification. If a pixel is classified as belonging to the LI interface but it is not coincident with the reference contour, the classification would be wrong from a computational point of view but, from a medical point of view, that difference would not be significant. An alternative way to train the classifiers would be using the distance between the classified pixel and the reference one.

The developed method, as well as the ones presented in the state-of-the-art, reported better results for the far wall. This interface has usually more definition and contrast, and this property is related with the probe position during the ultrasonography.

The results obtained for Rocha's database and Molinari's database were different in many important aspects. The feature selection method that presented better results for Rocha's database was the one using the wrapper approach. The one selected for Molinari's

database was the In-In method. The selected features for each interface were also different. Only 3 features were coincident for the LIFW and 4 for the MAFW. For Rocha's LINW, the selected classifier was the KNNC. The LIFW, MANW and MAFW were classified using a ParzenC. For Molinari's database, the classifier that presented the best performance for the LIFW and the MAFW was the ParzenC.

All of the 27 tested images selected from Rocha's database were correctly detected and the same happened in the 90 images used as testing from the Molinari's database.

The presented results show that for Rocha's database, the developed method inter variability is similar to the intra observer variability. The performance measurements used to evaluate the final interfaces compare each interface pixel with the corresponding one of the ground-truth. The FOMs obtained for these interfaces were very high, between 96% and 98%. The CV obtained for the IMT mean measurement was between 0.96% and 8.32%. For Molinari's database, the FOM obtained between the pixels pairs was around 98.89% and the CV between the mean IMT measurements ranged between 9.37% and 18.21%. This last CV value is considered to be have high variability.

The processing time obtained for Rocha's (1.73 min) database is significantly higher in comparison with the one obtained for Molinari's database (20.49 seconds). This difference happened because all the pixels from Rocha's database were considered as interface candidates and only the maximum of the ICOV in the gradient direction were considered for Molinari's database.

One of the most important aspects of this work is the dependency on the dynamic programming function. This method should be improved in order to reduce the interfaces detection error.

As future work, the classifier training should be done according to the distance between the GT and the automatic contour. The ICOV detector could be replaced by the FOAM detector, which is becoming more frequent in the literature due to its reliable results. The classification strategy could also be redefined by detecting the MA interface and the LI interface afterwards. This could be done because the MA interfaces have a low classification error and high response to the edge detectors. At last, some more textural features could be used in order to obtain better information about the pixels neighborhood.

Commercial software

There are already commercial systems that perform IMT measurements. They can be automatic or semi-automatic, perform IMT assessment measures like maximum, mean and minimum, allow manual and automatic tracings and it the storage, among other features. This subject was deeply explored in Carvalho [2011].

These commercial systems are:

- “Vivid 7 Dimension / Vivid i” of General Electrics Vivid General Electrics [2012]. This software does a semi-automatic analysis of the near and far wall IMT. To perform this analysis, it is only necessary to follow four steps. The first one is the selection of the specific button from the ultrasound equipment, followed by the selection of 2 points at the arterial lumen. The third step is the recording of the results. The last step is the conversion of that data into medical report format. This software returns the maximum and mean IMT values Vivid General Electrics [2012].
- “SonoCalc IMT” of SonoSite Sonosite [2012]. This software is fully automatic when the processed images have good quality and of semi-automatic when this quality is low. With this software it is possible to measure IMT manually, do an analysis of multiple images and measure IMT maximum, mean and average values. It allows the storage of the patient’s information as well as the patient’s pathology. At the end, it prints a report with the patient’s results and the pathology reference values.
- “M’Ath” of imt-Intelligence in Medical Technologies imt-Intelligence in Medical Technologies [2012]. With this software it is possible to do a semi-automatic near and far wall IMT measurement and measure the stenosis degree. It is possible to store the patient’s data and measure the image quality using the results of the resultant contour. If this contour was too fragmented, the image quality is low. It returns the mean, maximum and minimum IMT. However, this software can not measure the IMT of images with plaque.

- “SDU-2200 Pro” of Shimadzu [Shimadzu 2012]. It allows the integration of a IMT measurement module that performs the automatic tracings and the IMT analysis.
- Prosound α 7 of ALOKA Prosound [2012]. The ProSound α 7 is a portable diagnostic ultrasound system that automatically extracts the maximum IMT and the mean IMT only by setting a ROI on a vertical view of the vessel.

Differences of Three Tested Methods

This appendix shows the images that presented a significant difference between the result of the three feature selection methods. Figure B.1 shows the most significant differences for the Rocha's database.

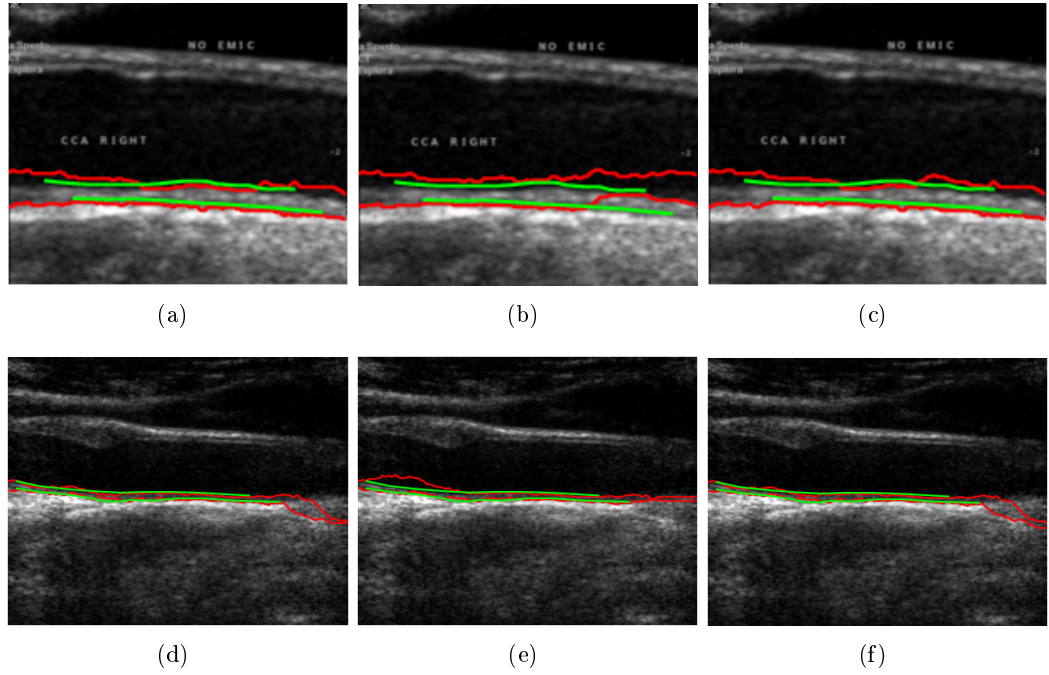


Fig. B.1: Differences between methods for Molinari's database. The left column represents the result using the In-In method, the middle column represent the wrapper approach and the right represents the result using the total dataset

Figure B.2 shows the most significant differences for the Molinari's database.

As this figures show, depending on the chosen feature selection method, the classification result can be very different. This is one of the reasons for any of these have been discharged from the potential methods until the analysis of the final results.

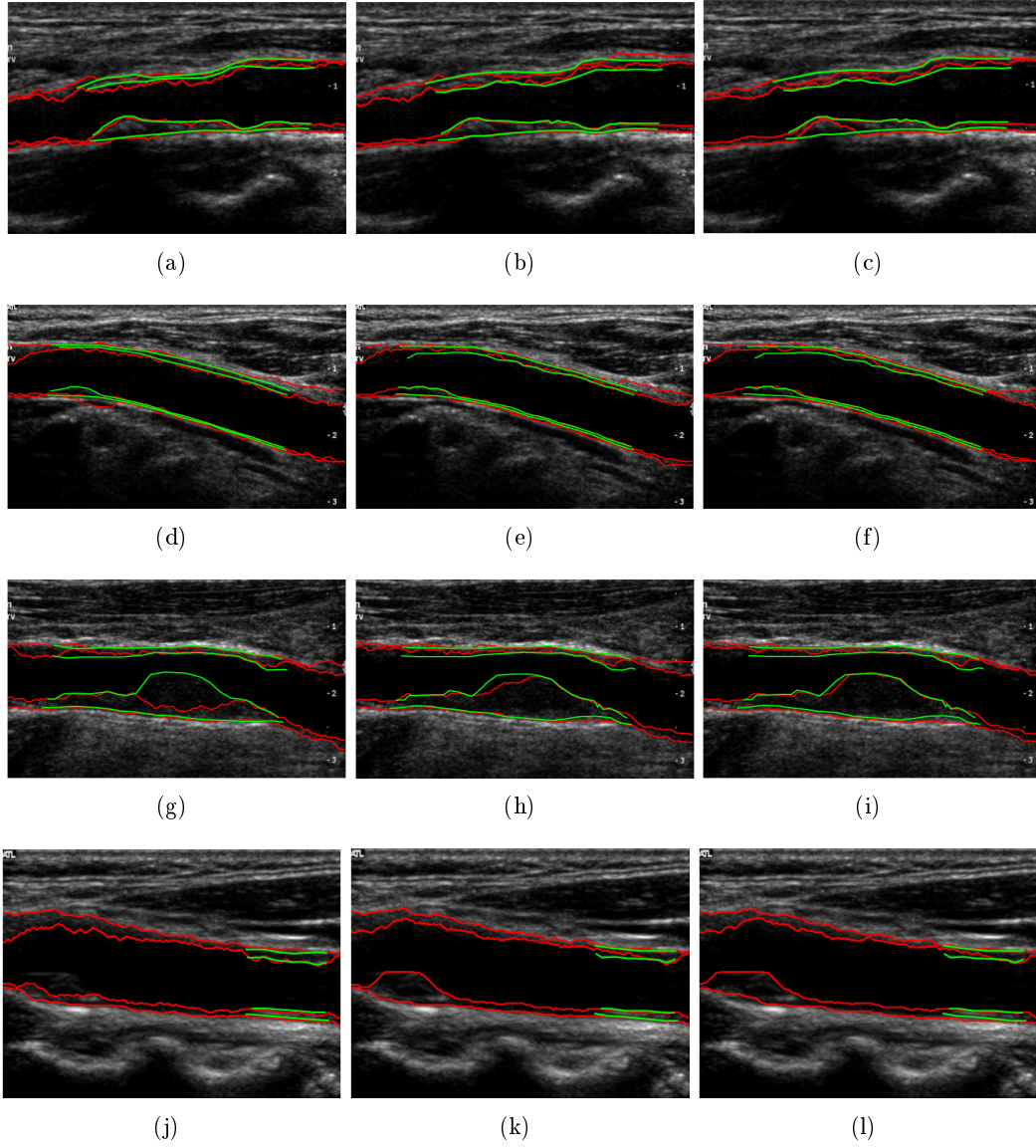


Fig. B.2: Differences between methods for Rocha's database. The left column represents the result using the In-In method, the middle column represent the wrapper approach and the right represent the result using the total dataset

Classification Results For Both Databases

This appendix shows more examples of detected interfaces.

Figure C.1 shows the classification result for 9 images of Rocha's database.

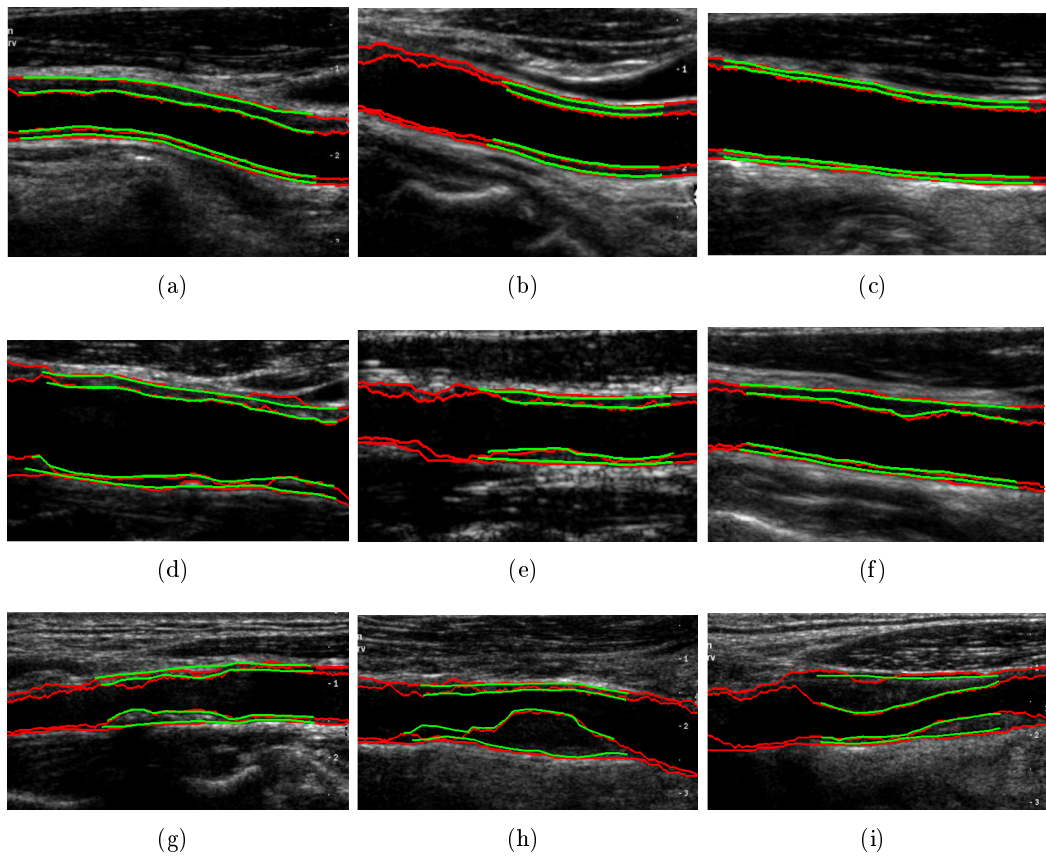


Fig. C.1: Results for Rocha's database: (a), (b) and (c) present arteries with low IMT; (d), (e) and (f) present arteries with some thickness; (g), (h) and (i) present arteries with plaques.

Figure C.1 (a) - (c) present arteries with low IMT. The interfaces detection of these

images is well performed. Figure C.1 (d) - (f) present arteries with some thickness. These images are more harder to classify but the final result is also good. Figure C.1 (g) - (i) present arteries with plaques. The interfaces detection for these images is also good, but MANW of Fig. C.1 (i) is little displaced.

Figure C.1 show the classification result for 9 images of Molinari's database.

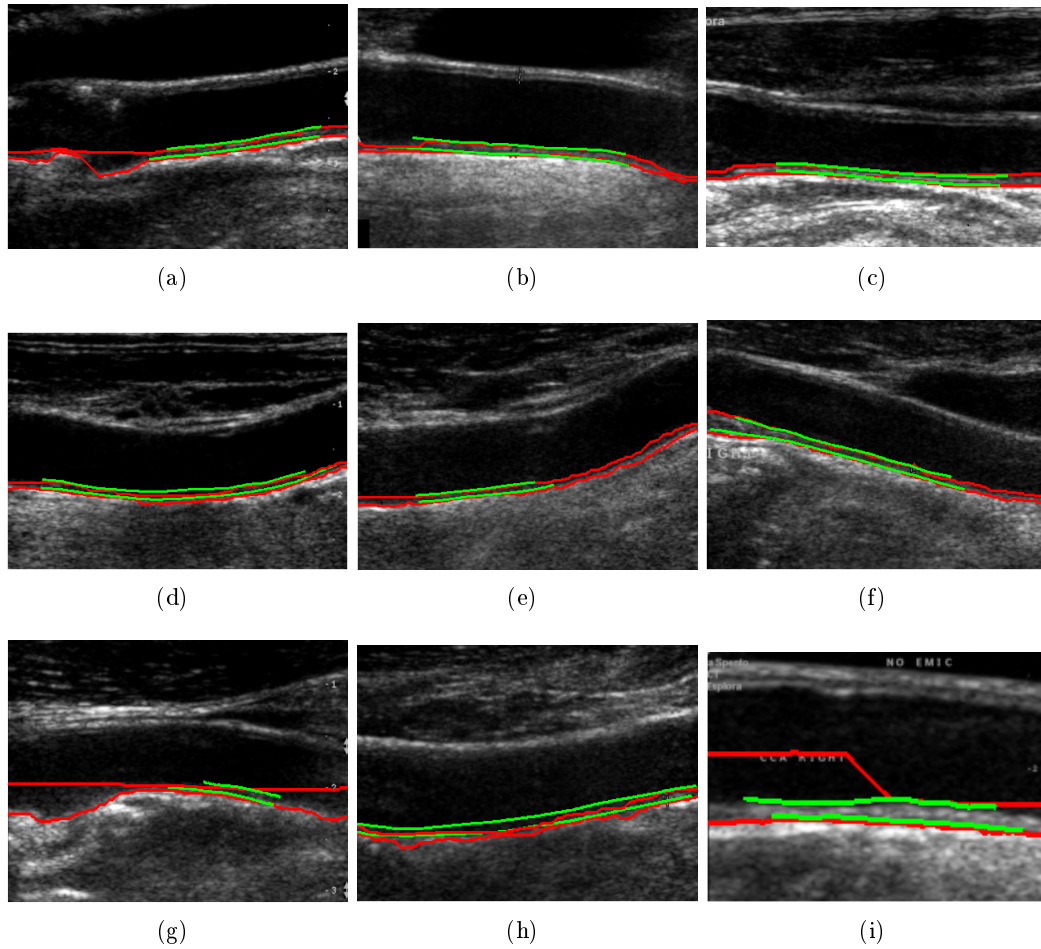


Fig. C.2: Results for Molinari's database: (a), (b) and (c) present horizontal arteries with low IMT; (d), (e) and (f) present sloping arteries; (g), (h) and (i) present some bad detections.

Figure C.2 (a) - (c) present horizontal arteries with low IMT. The interfaces detection of these images is well performed. Figure C.2 (d) - (f) present sloping arteries. Despite the slope, the interfaces detection is well performed. Figure C.2 (g) - (i) present bad detection cases. Figure C.2 (g) was wrongly detected because the high imposed geometric strength. Figure C.2 (h) present a case of subestimation of the IMT. This was caused by the lumen noise. Figure C.2 (i) present a wrong classification caused by external marks.

Bibliography

- Vivid General Electrics. Vivid general electrics. Website, July 2012. http://www.vivid7ultrasound.com/pdf/vivid7ultrasound.com/Vivid_7_IMT_Brochure.pdf.
- Adel H. Allam, Randall C. Thompson, Samuel Wann, Michael I. Miyamoto, Abd el-Halim Nur el Din, Gomaa Abd el Maksoud, Muhammad Al-Tohamy Soliman, Ibrahem Badr, Hany Abd el-Rahman Amer, M. Linda Sutherland, James D. Sutherland, and Gregory S. Thomas. Atherosclerosis in ancient egyptian mummies. *Jacc: Cardiovascular imaging*, 4:315–327, April 2011.
- N Anitschkow, Ludwig Aschoff, E T Bell, and Edmund V Cowdry. *Arteriosclerosis: A survey of the problem*, volume 1. 1933.
- Art. Encyclopædia Britannica Online. Art. encyclopædia britannica online. Website, July 2012. <http://www.britannica.com/EBchecked/topic/36874/artery?overlay=true&assemblyId=121565>.
- Dana Ballard and Christopher Brown. *Computer Vision*, pages 137–143. 1982.
- Bastida-Jumilla, Morales-Sanchez, Verdu-Monedero, Larrey-Ruiz, and Sancho-Gomez. Detection of the intima and media walls of the carotid artery with geodesic active contours. Proceedings of 2010 IEEE 17th International Conference on Image Processing, 2010.
- J. Martin Bland and Douglas G. Altman. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, 1:307–310, 1986.
- Sing-Tze Bow. *Pattern Recognition and Image Processing*. Marcel Dekker Inc, second edition, 2002.
- Catarina Carvalho. Interface gráfica para segmentação de imagens de ultrassonografia da carótida. Pratical report, Faculdade de Engenharia do Porto, FEUP, June 2011. Master Degree on Biomedical Engineering.
- Catarina Carvalho, Rui Rocha, Aurélio Campilho, and Elsa Azevedo. Classification approach for atherosclerosis diagnosis in B-mode ultrasound images. Technical report, Póvoa de Varzim, Portugal., 2012.
- Da-Chuan Cheng, Arno Schmidt-Trucksäss, Kuo-sheng Cheng, Markus Sandrock, Qin Pu, and Hans Burkhardt. Automatic detection of the intimal and the adventitial layers of the common carotid artery wall in ultrasound B-mode images using snakes. In *Proceedings*

- of the 10th International Conference on Image Analysis and Processing, pages 452–457, 1999.
- Da-Chuan Cheng, Arno Schmidt-Trucksas, Kuo-sheng Cheng, and Hans Burkhardt. Using snakes to detect the intimal and adventitial layers of the common carotid artery wall in sonographic images. *Computer Methods and Programs in Biomedicine*, 67:27–37, 2002.
- Lena Costaridou. *Medical Image Analysis Methods*. Taylor and Francis, 2005.
- Nello Cristianini and John Shawe-Taylor. *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press, 2000.
- Comaniciu D and Meer P. Robust analysis of feature spaces: color image segmentation. In *IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, pages 750–755, 1997.
- Silvia Delsanto, Filippo Molinari, Pierangela Giustetto, William Liboni, Sergio Badalamenti, and Jasjit S. Suri. Characterization of a completely user-independent algorithm for carotid artery segmentation in 2D ultrasound images. *IEEE Transactions on instrumentation and measurement*, 56(4):1265–1274, 2007.
- F. Destrempes, J. Meunier, M.F. Giroux, G. Soulez, and G. Cloutier. Segmentation in ultrasonic B-mode images of healthy carotid arteries using mixtures of nakagami distributions and stochastic optimization. *IEEE Trans. Med. Imaging*, 28(2):215–229., 2009.
- Michel Marie Deza and Elena Deza. *Encyclopedia of Distances*. Springer, 2009.
- Geoff Dougherty. *Digital Image Processing for Medical Applications*. Cambridge University Press, 2009.
- Richard O Duda, Peter E. Hart, and David G. Stork. *Pattern Classification*. Wiley-Interscience, second edition, 2000.
- R.P.W. Duin, P. Juszczak, P. Paclik, E. Pekalska, D. de Ridder, D.M.J. Tax, and S. Verzakov. Prtools4 - a matlab toolbox for pattern recognition, 2007. Version 4.1.
- T Elatrozy, A Nicolaides, T Tegos, and M Griffin. The objective characterization of ultrasonic carotid plaque features. *European Journal of Vascular and Endovascular Surgery*, 16:223–230, 1998.
- Francesco Faita, Vincenzo Gemignani, Elisabetta Bianchini, Chiara Giannarelli, Lorenzo Ghiadoni, and Marcelo Demi. Real-time measurement system for evaluating of the carotid intima-media thickness with a robust edge detector. *Ultrasound Med*, 27:1353–1361, 2008.
- Munther Gdeisat and Francis Lilley. One-dimensional phase unwrapping problem. Website, July 2007 . http://www.ljmu.ac.uk/GERI/CEORG_Docs/OneDimensionalPhaseUnwrapping_Final.pdf.
- Spyretta Golemati, John Stoitsis, Emmanouil G Sifakis, Thomas Balkizas, and Konstantina S. Nikita. Using the hough transform to segment ultrasound images of longitudinal and transverse sections of the carotid artery. *Ultrasound in Medicine and Biology*, 33(12): 1981–1992, 2007.

- Tomas Gustavsson, Quan Liang, Inger Wendelhag, and John Wikstrand. A dynamic programming procedure for automated ultrasonic measurement of the carotid artery. *Computers in Cardiology*, pages 297–300, 1994.
- MA Gutierrez, PE Pilon, SG Lage, L Kopel, RT Carvalho, and SS Furuie. Automatic measurement of carotid diameter and wall thickness in ultrasound images. *Computers in Cardiology*, 29:359–362, 2002.
- Peter Holdfeldt, Mats Viberg, and Tomas Gustavsson. A new method on dynamic programming for boundary detection in ultrasound image sequences. pages 3072–3074, August 2008.
- imt-Intelligence in Medical Technologies. imt- intelligence in medical technologies, July 2012. http://www.iimt.fr/lo_produitsUK.htm.
- Krzysztof Iniewski. *Medical Imaging -Principles, Detectors, and Electronics*. John Wiley Sons, Inc., Hoboken, New Jersey, 2009.
- Ron Kohavi and George H. John. Wrappers for featur subset selection. *Artificial Intelligence*, pages 273–324, 1997.
- Vinay Kumar, Abul K. Abbas, Nelson Fausto, and Jon Aster. *Robbins and Cotran pathologic basis of disease*. Saunders, 1600 John F. Kennedy Blvd. Ste 1800 Philadelphia, PA 19103-2899 Robbins and Cotran Pathologic Basis of Disease, 8/E, 8 edition, 2010.
- Yu-Bu Lee, Yoo-Joo Choi, and Myoung-Hee Kim. Boundary detection in carotid ultrasound images using dynamic programming and a directional haar-filter. *Computers in Biology and Medicine*, (40):687–697, 2010.
- Quan Liang, Inger Wendelhag, John Wikstrand, and Tomas Gustavsson. A multiscale dynamic programming procedure for boundary detection in ultrasonic artery images. *IEEE Transactions on medical imaging*, 19(2):127–142, February 2000.
- Science Photo Library. Johann czermak. Web photo, July 2012. <http://www.sciencephoto.com/media/224037/enlarge>.
- Consolatina Ligouri, Alfredo Paolillo, and Antonio Pietrosanto. An automatic measurement system for the evaluation of carotid intima-media thickness. *IEEE Transactions on instrumentation and measurement*, 50(6):1684–1691, December 2001.
- C. P. Loizou, C. S. Pattichis, M. Pantziaris, T. Tyllis, and A. Nicolaides. Snakes based segmentation of the common carotid artery intima media. *Med Bio Eng Comput*, 45: 35–49, 2007.
- Christos P. Loizou, Constantinos S. Pattichis, Andrew N. Nicolaides, and Marios Pantziaris. Manual and automated media and intima thickness measurements of the common carotid artery. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 56(5):983–994, 2009.
- Kristen M. Meiburger, Filippo Molinari, U. Rajendra Acharya, Luca Saba, Paulo Rodrigues, William Liboni, Andrew Nicolaides, and Jasjit S. Suri. Automated carotid artery intima layer regional segmentation. *Physics in Medicine and Biology*, 56:4073–4090, 2011.

- Ana Maria Mendonça. Evaluation and interpretation, 2011. Course Lecture on the Computer Aided Dignosis course - FEUP.
- Filippo Molinari, Guang Zeng, and Jasjit S. Suri. An integrated approach to computer-based automated tracing and its validation for 200 common carotid arterial wall ultrasound images - a new technique. *J Ultrasound Med*, 29:399–418, 2010a.
- Filippo Molinari, Guang Zeng, and Jasjit S. Suri. Greedy technique and its validation for fusion of two segmentation paradigms leads to an accurate intima-media thickness measure in plaque carotid arterial ultrasound. *Vascular Ultrasound*, 34:63–73, 2010b.
- Filippo Molinari, Guang Zeng, and Jasjit S. Suri. Carotid wall segmentation and imt measurement in longitudinal ultrasound images using morphological approach. *Computer Methods and Programs in Biomedicine*, 100(3):201–221, 2010c.
- Filippo Molinari, R.U. Acharya, G Zeng, K. M Meiburger, and J. S Suri. Completely automated robust edge snapper for carotid ultrasound imt measurement on a multi-institutional database of 300 images. *Springer Medical and Biological engineering and computing*, 49:11, 2011a.
- Filippo Molinari, Kristen M. Meiburger, Guang Zeng, Andrew Nicolaides, and Jasjit S. Suri. Caudles-ef: Carotid automated ultrasound double line extraction system using edge flow. *Journal of Digital Imaging*, 24:1059–1077, 2011b.
- Filippo Molinari, Kristen M. Meiburger, U. Rajendra Acharyac Luca Saba, Giuseppe Ledda, Guang Zeng, Sin Yee Stella Ho, Anil T. Ahuja, Suzanne C. Ho, Andrew Nicolaides, and Jasjit S. Suri. Ultrasound imt measurement on a multi-ethnic and multi-institutional database: Our review and experience using four fully automated and one semi-automated methods. *computer methods and programming in biomedicine*, page 15, May 2012a.
- Filippo Molinari, Kristen M. Meiburger, Luca Saba, U. Rajendra Acharya Mario Ledda, and Andrew Nicolaides Jasjit S. Suri. Constrained snake vs. conventional snake for carotid ultrasound automated imt measurement on multi-center data sets. *Ultrasonics*, 52(7):949–961, 2012b.
- Filippo Molinari, C Pattichis, G Zeng, L Saba, U Acharya, R Sanfilippo, A Nicolaides, and J. S. Suri. Completely automated multi-resolution edge snapper (CAMES) - a new technique for an accurate carotid ultrasound imt measurement: Clinical validation and benchmarking on a multi-institutional database. *IEEE Transactions on image processing*, 24:12, 2012c.
- Filippo Molinari, Constantinos S. Pattichis, Guang Zeng, Luca Saba, U. Rajendra Acharya, Roberto Sanfilippo, Andrew Nicolaides, and Jasjit S. Suri. Completely automated multi-resolution edge snapper-a new technique for an accurate carotid ultrasound imt measurement: Clinical validation and benchmarking on a multi-institutional database. *IEEE Transactions On Image Processing*, 21(3):1211–1222, 2012d.
- Fillipo Molinari, William Liboni, Pierangele Giustetto, Sergio Badalamenti, and Jasjit S. Suri. Automatic computer-based tracings (act) in longitudinal 2-d ultrasound images using different scanners. *Journal of Mechanics in Medicine and Biology*, 9(4):481–505, 2009.

- Fillippo Molinari, Guang Zeng, and Jasjit S. Suri. Review: A state of the art review on intima-media thickness (imt) measurement and wall segmentation techniques for carotid ultrasound. *Computer Methods and Programs in Biomedicine* 100, 100(3):201–221, 2010d.
- Fillippo Molinari, U. Rajendra Acharya, Guang Zeng, Kristen Meiburger, and Jasjit SSuri. Completely automated robust edge snapper for carotid ultrasound imt measurement on a multi-institutional database of 300 images. *Medical and Biological Engineering and Computing*, 49:965–945, 2011c.
- N Otsu. A threshold selection method from gray-level histograms. *IEEE Transaction System Man Cybern*, 9(1):62:66, 1979.
- Styliani Petroudi, Christos Loizou, Marios Pantziaris, and Constantinos Pattichis. A fully automated method using active contours for the evaluation of the intima-media thickness in carotid us images. 33rd Annual International Conference on IEEE EMBS, 2011.
- P. Pignoli and T. Longo. Evaluation of atherosclerosis with B-mode ultrasound imaging. *J. Nucl. Med*, 32:166–173, 1988.
- Jerry L. Prince and Jonathan M. Links. *Medical Imaging Signals and Systems*. Pearson Prentice Hall Bioengineering, 2006.
- Prosound. Prosound. Website, July 2012. <http://www.hitachi-aloka.co.jp>.
- P. Pudil, Novoviáčová, and J. Kittler. Floating search methods in feature selection. *Pattern Recognition Letters*, 15:1119–1125, 1994.
- Rui Rocha, Aurélio Campilho, Jorge Silva, Elsa Azevedo, and Rosa Santos. Segmentation of the carotid intima-media region in B-mode ultrasound images. *Image and vision computing*, 28:614–625, September 2010.
- Rui Rocha, Aurélio Campilho, Jorge Silva, Elsa Azevedod, and Rosa Santos. Segmentation of ultrasound images of the carotid using RANSAC and cubic splines. *Computer methods and programs in biomedicine*, 101:94–106, 2011.
- Rui Rocha, Jorge Silva, and Aurélio Campilho. Automatic segmentation of carotid B-mode images using fuzzy classification. *Medical and Biological Engineering and Computing*, 50(5):533–545, 2012.
- J. R. T. C. Roelandt. Seeing the invisible: A short history of cardiac ultrasound. *Eur J Echocardiography*, 1:8–11, 2008.
- W. Schäberle. *Ultrasonography in Vascular Diagnosis - A Therapy-Oriented Textbook and Atlas*. Springer, 2004.
- Bernhard Schaller. *Imaging of Carotid Artery Stenosis*. SpringerWienNewYork, 2007.
- Günter Schmidt. *Thieme Clinical Companions - Ultrasound*. Thieme, Rüdigerstrasse 14, 70469 Stuttgart, Germany, 2007.
- Rod Seeley, Trent Stephens, and Philip Tate. *Anatomia and Fisiologia*. Lusodidacta, 8 edition, 2012.

- Shimadzu. Shimadzu. Website, July 2012. <http://www.shimadzu.eu/medical/products/ultra/sdu2200/features.\aspx?page=IMT>.
- Milan Sonka, Vaclav Hlavac, and Roger Boyle. *Image Processing, Analysis and Machine Vision*. Thompson, 2008.
- Sonosite. Sonosite. Website, July 2012. <http://www.sonosite.com/products/sonocalc-IMT/demo/>.
- James Stein, Claudia Korcarz, Maureen E. Mays, Pamela Douglas, Mari Palta, Hongling Zhang, Tamara LeClaire, Diane Paine, David Gustafson, and Liexiang Fan. A semiautomated ultrasound border detection program that facilitates clinical measurement of ultrasound carotid intima-media thickness. *Journal of the American Society of Echocardiography*, 18(3):244–251, 2005.
- J. Stoitsis, S. Golemati, S. Kendrosa, and K. S. Nikita. Automated detection of the carotid artery wall in B-mode ultrasound images using active contours initialized by the hough transform. In *30th Annual International IEEE EMBS Conference*, 2008.
- Paul Suetens. *Fundamentals of Medical Imaging*. Cambridge University Press, 2 edition, 2002.
- J.S. Suri, R.M. Haralick, and F.H. Sheehan. Greedy algorithm for error correction in automatically produced boundaries from low contrast. *Pattern Anal. Appl*, 3(1), 2000.
- Sergios Theodoridis and Konstantinos Koutroumbas. *An introduction to Pattern Recognition: A MATLAB approach*. Elsevier Inc., 2010.
- P.J. Touboul, P. Prati, P.Y. Scarabin, V. Adrai, E. Thibout, and P. Ducimetiere. Use of monitoring software to improve the measurement of carotid wall thickness by B-mode imaging. *J. Hypertens*, 10:S37–41, 1992.
- Vascular and Endovascular Surgery. Vascular and endovascular surgery. Website, July 2012. <http://vascular.surgery.ucsf.edu/conditions--procedures/carotid-ultrasound.aspx>.
- Vascular and Endovascular Surgery. Vascular and endovascular surgery. Website, 2012 July. <http://vascular.surgery.ucsf.edu/conditions--procedures/carotid-artery-disease.aspx>.
- Zhou Wang and Alan C. Bovik. Mean squared error: Love it or leave it? - a new look at signal fidelity measures. *IEEE SIGNAL PROCESSING MAGAZINE*, 98, 2009.
- Inger Wendelhag, Quan Liang, Tomas Gustavsson, and John Wikstrand. A new automated computerized analyzing system simplifies readings and reduces the variability in ultrasound measurement of intima-media thickness. *Stroke*, 28:2195–2200, 1997.
- WHO. World health organization. Website, July 2012. <http://www.who.int/en/>.
- Richard H. Wiggins, H. Christian Davidson, H. Ric Harnsberger, Jason R. Lauman, and Patricia A. Goede. Image file formats: Past, present and future. *RadioGraphics*, 21(3): 789–798, 2001.

- Joseph Woo. Karl theodore dussik. Web photo, July 2012. <http://www.ob-ultrasound.net/dussikbio.html>.
- Yongjian Yu and Scott T. Acton. Edge detection in ultrasound imagery using the instantaneous coefficient of variation. *IEEE Transactions On Image Processing*, 13(12):1640–1655, 2004.
- G. W. Zack, W. E. Rogers, and S. A. Latt. Automatic measurement of sister chromatid exchange frequency. *Histochem Cytochem*, 25(7):741–753, 1977.