ΠΑΝΕΠΙΣΤΗΜΙΟ ΙΩΑΝΝΙΝΩΝ

ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΕΠΙΣΤΗΜΗΣ ΥΛΙΚΩΝ

Προγνωστική Μοντελοποίηση της Εξέλιξης της Αθηρωματικής Πλάκας

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

Βασιλική Ι. Κίγκα

Ιωάννινα, 2023

«Η υλοποίηση της διδακτορικής διατριβής συγχρηματοδοτήθηκε από την Ελλάδα και την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) μέσω του Επιχειρησιακού Προγράμματος «Ανάπτυξη Ανθρώπινου Δυναμικού, Εκπαίδευση και Δια Βίου Μάθηση», 2014-2020, στο πλαίσιο της Πράξης «Ενίσχυση του ανθρώπινου δυναμικού μέσω της υλοποίησης διδακτορικής έρευνας Υποδράση 2: Πρόγραμμα χορήγησης υποτροφιών ΙΚΥ σε υποψηφίους διδάκτορες των ΑΕΙ της Ελλάδας».

Επιχειρησιακό Πρόγραμμα
Ανάπτυξη Ανθρώπινου Δυναμικού,
Εκπαίδευση και Διά Βίου Μάθηση

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

UNIVERSITY OF IOANNINA

SCHOOL OF SCIENCES AND TECHNOLOGIES

DEPARTMENT OF MATERIALS SCIENCE AND ENGINEEIRING

Predictive Modeling of the Atheromatic Plaque Growth

PhD Thesis

Vassiliki I. Kigka

Ioannina, 2023

«The implementation of the doctoral thesis was co-financed by Greece and the European Union (European Social Fund-ESF) through the Operational Programme «Human Resources Development, Education and Lifelong Learning» in the context of the Act "Enhancing Human Resources Research Potential by undertaking a Doctoral Research" Sub-action 2: IKY Scholarship Programme for PhD candidates in the Greek Universities».

«Η έγκριση της διδακτορικής διατριβής από το Τμήμα Μηχανικών Επιστήμης Υλικών της Σχολής Θετικών Επιστημών του Πανεπιστημίου Ιωαννίνων δεν υποδηλώνει αποδοχή των γνωμών του συγγραφέα Ν. 5343/32, άρθρο 202, παράγραφος 2».

Ημερομηνία αίτησης της κ. Βασιλικής Κίγκα: 18.1.20I9

Ημερομηνία ορισμού Τριμελούς Συμβουλευτικής Επιτροπής: 23.1.2019

Μέλη Τριμελούς Συμβουλευτικής Επιτροπής:

Επιβλέπων

Δημήτριος Ι. Φωτιάδης - Καθηγητής, Τμήμα Μηχανικών Επιστήμης Υλικών, Πολυτεχνική Σχολή, Πανεπιστήμιο Ιωαννίνων

Μέλη

Λεωνίδας Γεργίδης - Αναπλ. Καθηγητής Τμήμα Μηχανικών Επιστήμης Υλικών, Πολυτεχνική Σχολή, Πανεπιστήμιο Ιωαννίνων,

Λάμπρος Μιχάλης - Καθηγητής, Τμήμα Ιατρικής, Πανεπιστήμιο Ιωαννίνων

Ημερομηνία ορισμού θέματος: 23. 1.2019

Θέμα: «Προγνωστική Μοντελοποίηση της Εξέλιξης της Αθηρωματικής Πλάκας»

Ημερομηνία ορισμού Επταμελούς Εξετατισκής Επιτροπής: 10- 10-2023

ΔΙΟΡΙΣΜΟΣ ΕΠΤΑΛΜΕΛΟΥΣ ΕΞΕΤΑΣΤΙΚΗΣ ΕΠΙΤΡΟΠΗΣ

1. **Δημήτριος Φωτιάδης**, Καθηγητής, Τμήμα Μηχανικών Επιστήμης Υλικών, Πολυτεχνική Σχολή, Πανεπιστήμιο Ιωαννίνων
2 **Λεωνίδας Γεργίδης**, Αναπληρωτής. Καθηγητής, Τμήμα Μηχανικών Επιστήμης Υλικών, Πολυτεχνική Σχολή, Πανεπιστήμιο Ιωαννίνων
3 **Λάμπρος Μιχάλης**, Καθηγητής, Τμήμα Ιατρικής, Σχολή Επιστημών Υγείας, Πανεπιστήμιο Ιωαννίνων
4 **Φραγκίσκα Σιγάλα**, Αναπληρώτρια Καθηγήτρια, Τμήμα Ιατρικής, Σχολή Επιστημών Υγείας, Πανεπιστήμιο Αθηνών
5 **Γεώργιος Ματσόπουλος**, Καθηγηής, Σχολή Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών, Εθνικό Μετσόβιο Πολυτεχνείο
6 **Αικατερίνη Νάκα**, Καθηγήτρια, Τμήμα Ιατρικής, Σχολή Επιστημών Υγείας, Πανεπιστήμιο Ιωαννίνων
7 **Αντώνιος Σακελλάριος**, Επίκουρος. Καθηγητής, Τμήμα Μηχανολόγων και Αεροναυπηγών Μηχανικών, Πανεπιστήμιο Πατρών

Έγκριση Διδακτορικής Διατριβής με βαθμό «Άριστα» στις 30.11.2023

Ο Πρόεδρος του Τμήματος

Απόστολος Αυγερόπουλος
Καθηγητής

Η Γραμματέας του Τμήματος

Μαρία Κόντου

Date of application of Mrs. Vassiliki Kigka: 18.1.2019

Date of Appointment of PhD Advisory Committee: 23.1.2019

Members of the Three-member Advisory Committee:

Supervisor

Dimitrios I. Fotiadis - Professor, Department of Materials Science and Engineering, School of Engineering, University of Ioannina

Members

Leonidas Gergidis - Associate Professor, Department of Materials Science and Engineering, School of Engineering, University of Ioannina,

Lampros Michalis - Professor, Department of Medicine, University of Ioannina

Date of dissertation topic assignment: 23.1.2019

Topic: Prognostic Modeling of Atherosclerotic Plaque Progression

Date of appointment of the seven-member examination committee: 10.10.2023

SEVEN-MEMBER EXAMINATION COMMITTEE

1. **Dimitrios Fotiadis**, Professor, Department of Materials Science and Engineering, School of Engineering, University of Ioannina
2. **Leonidas Gergidis**, Associate Professor, Department of Materials Science and Engineering, School of Engineering, University of Ioannina
3. **Lambros Michalis**, Professor, Department of Medicine, Faculty of Health Sciences, University of Ioannina
4. **Fragkiska Sigala**, Associate Professor, Department of Medicine, School of Health Sciences, University of Athens
5. **Georgios Matsopoulos**, Professor, School of Electrical and Computer Engineering, National Technical University of Athens
6. **Aikaterini Naka**, Professor, Department of Medicine, School of Health Sciences, University of Ioannina
7. **Antonios Sakellarios**, Assistant. Professor, Department of Mechanical and Aeronautical Engineering, University of Patras

Approval of PhD Dissertation with grade "Excellent" on 30.11.2023.

The Chairman of the Department

Apostolos Avgeropoulos
Professor

The Secretary of the Department

Maria Kontou

Dedication

To my Daughter, my Husband and my Parents.

Acknowledgements

First of all, I would like to thank my family and my husband, who all these years support my dreams and my effort. I sincerely need to cordially thank my supervisor Dr. Dimitrios I. Fotiadis, Professor, Dept. of Materials Science and Engineering, University of Ioannina for his support and great assistance in order to complete this thesis. I would also like to thank the members of my advisory committee, Dr. Lampros Michalis, Professor of Cardiology, Medical school, University of Ioannina and Dr. Leonidas Gergidis, Associate Professor, Dept. of Materials Science and Engineering, University of Ioannina, for their support and cooperation.

I should also thank all my colleagues I have worked with, especially Dr. Eleni Georga, Dr. Vassiliki Potsika, Dr. Vassilis Tsakanikas and Dr. Panagiotis Siogkas, with whom we spent many years together, sharing good and bad times.

Furthermore, I would like to thank Dr. Igor Koncar with whom we had an excellent collaboration through the last two years of my PhD thesis, Dr. Laura Ludovica Gramegna, who provides the brain MRI annotations and Dr. Dimitra Loggitsi and Dr. Stathis Detorakis, who contributed significanty to the CTA images annotation.

This work was part funded by: the European Commission [Project SMARTOOL, “Simulation Modeling of coronary ARTery disease: a tool for clinical decision support — SMARTool” under grant agreement number: 689068, Project TAXINOMISIS, under grant agreement number 755320]. The last 16 months of this PhD thesis, the work was funded by National Scholarships’ Foundation (IKY, Greece), through the Operational Program "Human Resource Development, Education and Lifelong Learning", 2014-2020, within the framework of the Action "Strengthening human resources through the for the implementation of doctoral research Sub-action 2: IKY grant program for doctoral candidates of Greek universities".

Table of Contents

Chapter 1 Introduction 1

1.1 Physiology of the cardiovascular system 1

1.2 Central nervous system vasculature 6

1.3 Atherosclerosis 7

1.4 Imaging of atherosclerosis 12

Chapter 2 Literature Overview 17

2.1 Introduction 17

2.2 3D reconstruction based on computed tomography angiography 18

2.3 Prediction of coronary artery disease 33

2.4 Prediction of carotid artery disease 42

2.5 Serum Markers in Carotid Artery Disease 45

2.6 Contribution of this thesis 62

Chapter 3 3D arterial Reconstruction using CTA 66

3.1 Introduction 66

3.2 Methodology for 3D inner and outer wall reconstruction 67

3.3 Methodology for atherosclerotic plaque characterization 78

3.4 Dataset 85

3.5 Results 86

3.6 Discussion 104

Chapter 4 Coronary Artery Disease Prediction 111

4.1 Introduction 111

4.2 Coronary artery disease prediction using machine learning 112

4.2.1 Methodology 112

4.2.2 Dataset 118
4.2.3 Results 119
4.3 Prediction of the progression of coronary artery disease using machine learning 124
4.3.1 Methodology 124
4.3.2 Dataset 125
4.3.3 Results 125
4.4 Site specific prediction of atherosclerotic plaque progression using machine learning 127
4.4.1 Methodology 127
4.4.2 Dataset 130
4.4.3 Results 130
4.5 Site specific prediction of PCI stenting using machine learning 131
4.5.1 Methodology 131
4.5.2 Dataset 133
4.5.3 Results 133
4.6 Discussion 134
Chapter 5 Carotid artery disease prediction 139
5.1 Introduction 139
5.2 Methodology 140
5.3 Diagnostic prediction of carotid artery disease 147
5.3.1 Problem definition 147
5.3.2 Dataset 147
5.3.3 Results 149
5.4 Diagnostic prediction of the vulnerability of carotid artery plaques 153
5.4.1 Problem definition 153
5.4.2 Dataset 154
5.4.3 Results 158
5.5 Discussion 178

Chapter 6 Silent brain lesions detection 181
6.1 Introduction 181
6.2 Association of carotid artery disease with silent brain lesions 182
6.2.1 Brain MRI analysis 182
6.2.2 Methodology 185
6.2.3 Dataset 186
6.2.4 Results 188
6.3 Discussion 193
Chapter 7 Conclusions and future work 197
7.1 Conclusions 197
7.2 Future work 201
References 203
Appendix 231
Author's Publications 235
Journal Publications 235
Conference Publications 236
Book Chapters 240
Short CV 241

List of Abbreviations

Acronym	*Meaning*
2D (p 13)	Two Dimensional
3D (p 16)	Three Dimensional
%AS (p 24)	% Area of Stenosis
%DS (p 24)	% Degree of Stenosis
ACEi	Angiotensin Converting Enzyme Inhibitor
AHA	American Heart Association
ANN	Artificial Neural Network
Apo	Apolipoprotein
ARB	Angiotensin Receptor Blocker
ASCVD	Atherosclerosis and Atherosclerotic Cardiovascular Disease
ASCVD	Atherosclerosis and atherosclerotic cardiovascular disease
AUC	Area Under the Curve
BMI	Body Mass Index
CA	Coronary Angiography
CABG	Coronary Artery Bypass Grafting
CAD	Coronary Artery Disease
CAD-RADS	Coronary Artery Disease Reporting and Data System
CAM	Cell Adhesion Molecule
CART	Classification and Regression Tree
CAS	Carotid Artery Stenosis
CCA	Common Carotid Artery
CCL	C-C Motif Chemokine Ligand
CEA	Carotid Endarterectomy
CFS	Correlation-based Feature Selection
CI	Confidence Interval
COPD	Chronic Obstructive Pulmonary Disease
CP	Calcified Plaques

Acronym	*Meaning*
CRP	C Reactive Protein
CT	Computed Tomography
CTA	Computed Tomographic Angiography
CV	Cardiovascular
CVD	Cardiovascular Disease
DBP	Diastolic Blood Pressure
DICE	Dice Coefficient
DM	Diabetes Mellitus
ECA	External Carotid Artery
ESC	European Society of Cardiology
ESS	Endothelial Shear Stress
FABP4	Fatty Acid Binding Protein 4
FFNN	feed-forward neural network
FFR	Fractional Flow Reserve
FOV	Field of View
GDF15	Growth Differentiation Factor 15
GFR	Glomerular filtration rate
GMM	Gaussian Mixture Model
Hb	Haemoglobin
Hct	Hematocrit
hcy	homocysteine
HD	Hausdorff Distance
HDL	High Density Lipoprotein
hs-CRP	high-sensitivity C Reactive Protein
HU	Hounsfield Unit
ICA	Internal Carotid Artery
ICAM	Intercellular Adhesion Molecule
IGF-1	Insulin-like Growth Factor 1

Acronym	*Meaning*
IL	Interleukin
IMT	Intima Media Thickness
IVUS	Intravascular Ultrasound
IVUS-RF	Intravascular Ultrasound Radiofrequency
LAD	Left Anterior Descending Artery
LCX	Left Circumflex artery
LDL	Low Density Lipoprotein
LMCA	Left Main Coronary Artery
LMT	Logistic Model Tree
Lp-PLA(2)	Lipoprotein-associated Phospholipase A(2)
LR	Logistic Regression
MACE	Major Adverse Cardiovascular Events
MI	Myocardial Infarction
MIP	Maximum Intensity Projection
ML	Machine Learning
MLA	Minimal Lumen Area
MLD	Minimal Lumen Diameter
MMP	Matrix Metalloproteinase
MPO	Myeloperoxidase
MRI	Magnetic Resonance Imaging
NB	Naïve Bayes
NC	Necrotic Core
NCP	Non Calcified Plaques
NGAL	Neutrophil Gelatinase associated Lipocalin
NIRS	Near Infrared Spectroscopy
NPV	Negative Predictive Value
NRS	Napkin Ring Sign
Nt-pro	N-terminal pro-brain natriuretic peptide
OCT	Optical Coherence Tomography

Acronym	*Meaning*
OPG	Osteoprotegerin
OPN	Osteopontin
OR	Odds Ratio
ox-LDL	oxidized Low Density Lipoprotein
PAD	Peripheral Arterial Disease
PAOD	Peripheral Arterial Occlusive Disease
PB	Plaque Burden
PCI	Percutaneous Coronary Intervention
PCSK-9	Proprotein Convertase Subtilisin Kexin type 9
PDGF	Platelet Derived Growth Factor
PET-CT	Positron Emission Tomograph- Computed Tomography
PPV	Positive Predictive Value
PSF	Point Spread Function
PSV	Peak systolic velocity
PTP	Pre Test Probability
PTX	Pentraxin-3
RANTES	Regulated on Activation Normal T-cell Expressed and Secreted
RBF	Radial Basis Function
RCA	Right Coronary Artery
RF	Random Forest
RFE	Recursive Feature Elimination
ROC	Receiver Operating Characteristic
ROI	Region Of Interest
RT	Random Tree
SBI	Silent Brain Infarct
SBP	Systolic Blood Pressure
SD	Standard Deviation
SHAP	SHAPley Additive exPlanations

Acronym	*Meaning*
sIL-6R	Soluble Interleukin-6 receptor
sLOX-1	Soluble Lectin-like Oxidized LDL receptor-1
SMC	Smooth Muscle Cell
SMOTE	Synthetic Minority Oversampling Technique
SSIM	Structural Similarity Index Measure
STP	Standard Temperature and Pressure
SVM	Support Vector Machine
TAT	Thrombin Antithrombin complex
TC	Total Cholesterol
TIA	Transient Ischemic Attack
TIMP	Tissue Inhibitor of Metalloproteinase
TNF-α	Tumor Necrosis Factor alpha
TOF	Time Of Flight
TRL	Triglyceride-Rich Lipoprotein
UMC	Universitair Medisch Centrum Utrecht
US	Ultrasound
USMI	IRCCS Azienda Ospedaliera Universitaria San Martino Ospedaliera Universitaria San Martino-Isr-Istituto Nazionale per La Ricerca Sul Cancro
VCAM	Vascular Cell Adhesion Molecule
VEGF-A	Vascular Endothelial Growth Factor A
VH-IVUS	Virtual Histology-Intravascular Ultrasound
VWF	Von Willebrand Factor
WSS	Wall Shear Stress
XGB	Extreme Gradient Boosting

List of Symbols

Symbol	*Meaning*
$\bar{I}_{lumen}$	Mean lumen intensity
E_{CV}	Chan-Vese energy function
E_{shape}	Shape comparison term
E_{ψ}	Labelling term
cp_{thres}	Threshold of the calcified plaques
d_1	Distance transform for the pixels around the extracted centerline
$f_{1,lumen}$	Updated membership function for the lumen
$f_{2,outer}$	Updated membership function for the outer wall
f_{cp}	Membership function for the calcified plaques
f_{lumen}	Membership function for the lumen
f_{outer}	Membership function for the outer wall
$g^{bell}(x; a, b, c)$	Bell-shaped membership function
$g^{sigm}(x; a, b)$	Sigmoidal membership function
l_{thres}	Threshold of the lumen
w_i	Estimated weight to multiply the probability in the level set method
w_{vessel}	Image weight based on the vesselness measure
r	Pearson's correlation
R^2	Degree of correlation
ml	Mean luminal intensity around the extracted centerline
φ	Initial image shape
3D (p14)	Three dimensional
2D (p12)	two-dimensional
SMARTFFR	Smart Fractional Flow Reserve

List of Figures

Figure 1.1 The circulation system [2]. 2

Figure 1.2 Heart anatomy [3]. 3

Figure 1.3 The coronary arteries [2]. 4

Figure 1.4 Arterial morphology [4]. 5

Figure 1.5: Basic Anatomy of carotid arterial tree [14]. 7

Figure 1. 6 Stages in the development of atherosclerotic lesions [20]. 9

Figure 1. 7 Current guidelines for carotid artery disease management [30]. 12

Figure 1.8 Imaging Techniques of Coronary Artery disease [41]. 14

Figure 3.1 The proposed methodology outline [204]. 67

Figure 3.2 An example of the implementation of the Vesselness filter in a CTA image [204]. 68

Figure 3.3 An example of the implementation of blooming effect removal technique [204]. 69

Figure 3.4 Example of a successfully extracted coronary artery centerline using the vesselness/intensity cost function [204]. 70

Figure 3.5 Membership functions distributions for lumen and outer wall over HU [204]. 72

Figure 3.6 Lumen Segmentation example, a) acquired image, b) a factor 0.1, c) a factor 0.6 [204]. 76

Figure 3.7 Example of 3D reconstruction of coronary artery in baseline and follow-up [204]. 77

Figure 3.8 (a) 3D reconstruction of a carotid artery bifurcation, (b) the 2D CTA slice without segmentation and (c) the 2D segmentation of the inner wall and outer wall [206]. 77

Figure 3.9 Region of Interest (ROI) [204]. 78

Figure 3.10 Membership functions distribution for CP over HU [204]. 79

Figure 3.11 An example of the segmentation procedure: a) the acquired image, b) inner wall, outer wall and CP [205]. 80

Figure 3.12 a) Lumen (red) segmentation and CP (green) segmentation across the vessel centerline, b) Lumen (red) and CP (green) segmentation in a 2D CTA slice, c) Lumen (green) and outer wall (grey) suface per 0.5 mm. 81

Figure 3.13 An example of the segmentation procedure: a) the acquired image, b) inner wall, outer wall and NCP plaques [205]. 82

Figure 3.14 3D reconstructed models for the lumen, the outer wall and the CP [204]. 83

Figure 3.15 3D reconstructed models of a coronary arterial bifurcation for the lumen, the outer wall, the CP and the NCP. 83

Figure 3.16 3D reconstructed models of a coronary arterial tree for the lumen, the outer wall, the CP and the NCP. 84

Figure 3.17 Diagram of a coronary stenosis. 88

Figure 3.18 Lumen and CP objects detected by a) the proposed methodology, and b) the medical expert annotation [204]. 89

Figure 3.19 Bland Altman and Correlation plots for the DS1. 91

Figure 3.20 Bland Altman and Correlation plots for the DS2. 91

Figure 3.21 Bland Altman and Correlation plots for the MLA. 91

Figure 3.22 Bland Altman and Correlation plots for the MLD. 92

Figure 3.23 Bland Altman and Correlation plots for the PB. 92

Figure 3.24 Comparison between the inner wall area per 0.5mm derived by the presented methodology and the inner wall area per 0.5mm derived by an IVUS based approach. 93

Figure 3.25 Comparison of the proposed methodology (with and without blooming incorporation) with IVUS based methodology. 95

Figure 3.26 Bland-Altman and correlation plots for CTA and VH-IVUS for the CP volume and the CP length of lesion. 96

Figure 3.27 Bland-Altman and correlation plots for CTA and VH-IVUS for the NCP volume and NCP length of lesion. 97

Figure 3.28 Bland-Altman and correlation plots for CTA and VH-IVUS for the CP area. 97

Figure 3.29 Bland-Altman and correlation plots for CTA and VH-IVUS for the NCP area. 98

Figure 3.30 Manual and automated segmentation for the lumen (a, b), for the CP (c, d) and the NCP (e,f) [205]. 102

Figure 3.31 First Validation Strategy..... 103

Figure 3.32 Second Validation Strategy. 103

Figure 3.33 Correlation Analysis for the inner wall area..... 104

Figure 3.34 Correlation Analysis for the outer wall area..... 104

Figure 4. 1 Overall procedure for the SMARTFFR index calculation. 113

Figure 4.2 Flow chart depicting the distribution of the cohort in CAD-severity groups based on the CTA imaging at the follow-up step. in total, 287 patient imagings (125 in Class 1 and 62 in Class 2) were analyzed [232]. 114

Figure 4.3 Overall pipeline of the proposed methodology [232]..... 118

Figure 4. 4 Confusion Matrix..... 120

Figure 4.5 Feature importance based on mean SHAP values. The number of the existing CP and the highest coronary DS are indicated as the most significant features..... 122

Figure 4.6 Input Features Contribution Table (blue, features with negative effect; yellow, features with positive effect). The most significant features that contribute positively to the output target are thyroid stimulating hormone, medication therapy of beta blockers, aspartate aminotransferase, diabetes, and minimum lumen area, whereas the most significant feature with negative effect on the output is the number of the CP at the baseline analysis of patient imaging. 123

Figure 4.7 a) 3D reconstruction, b) ESS distribution at the luminal surface, c) normalized subendothelial LDL concentration [234]. 129

Figure 4.8 Prediction performance (sensitivity versus 1-specificity) of the XGB tree ensembles across each fold. 133

Figure 5. 1 Pipeline of Predictive Models Overall Training and Evaluation Procedure based on Machine Learning Techniques. 142

Figure 5.2 Curation Procedure Output. 142

Figure 5.3 Diagnostic Prediction of CAS>50% clinical problem. 147

Figure 5.4 Normalized confusion matrix regarding the diagnosis of carotid artery disease presence. .. 151

Figure 5.5 ROC curve analysis regarding the diagnosis of carotid artery disease presence. 151

Figure 5.6 Tuning of threshold value based on ROC analysis regarding the diagnosis of carotid artery disease presence. 152

Figure 5.7 Tuning of threshold value based on balanced accuracy regarding the diagnosis of carotid artery disease presence. 152

Figure 5.8 Diagnostic Prediction of vulnerable plaque components. 154

Figure 5. 9 ROC curve analysis regarding Subproblem 1. 169

Figure 5.10 ROC curve analysis regarding Subproblem 2. 169

Figure 5.11 ROC curve analysis regarding Subproblem 3. 170

Figure 5.12 ROC curve analysis regarding Subproblem 4. 170

Figure 5. 13 ROC curve analysis regarding Subproblem 5. 171

Figure 5.14 ROC curve analysis regarding Subproblem 6. 171

Figure 5. 15 ROC curve analysis regarding Subproblem 7. 172

Figure 5.16 Tuning of threshold value based on ROC analysis regarding Subproblem 1..... 173

Figure 5.17 Tuning of threshold value based on balanced accuracy regarding Subproblem 1. 173

Figure 5. 18 Tuning of threshold value based on ROC analysis regarding Subproblem 2..... 174

Figure 5. 19 Tuning of threshold value based on balanced accuracy regarding Subproblem 2. 174

Figure 5. 20 Tuning of threshold value based on ROC analysis regarding Subproblem 3..... 175

Figure 5. 21 Tuning of threshold value based on balanced accuracy regarding Subproblem 3. 175

Figure 5. 22 Tuning of threshold value based on ROC analysis regarding Subproblem 4..... 176

Figure 5. 23 Tuning of threshold value based on balanced accuracy regarding Subproblem 4. 176

Figure 5. 24 Tuning of threshold value based on ROC analysis regarding Subproblem 5..... 177

Figure 5. 25 Tuning of threshold value based on balanced accuracy regarding Subproblem 5. 177

Figure 6.1: Example of 3D TOF image, where the red point depicts the lumen of the CCA..... 182

Figure 6. 2: 2D TOF images of the right and left carotid arteries are shown at the sides of the image with the axial slice locations for the T1-weighted sequence illustrated in yellow. The corresponding T1-weighted slices, from CCA to ICA, are shown in the centre of the image. 183

Figure 6.3 Chronic white matter ischemia. 184

Figure 6. 4 Examples of a) cortical, b) lacunar and c) small subcortical infarcts. 185

Figure 6.5 Posterior circulation stroke. 185

Figure 6.6 Plaque stenosis (arrow) on the left side that corresponds to the brain lesion (encircled) on the right side..... 186

List of Tables

Table 2.1 Coronary Lumen Detection Methodologies based on CTA modality. 26

Table 2.2 Plaque Characterization Methodologies based on CTA modality. 27

Table 2.3 Methodologies for 3D reconstruction of carotid arteries using CTA imaging modality. 32

Table 2.4 Non-imaging CAD Severity Diagnosis Methods based on Machine Learning [100]. 36

Table 2.5 Pre-test CAD Severity Scores based on Genomics. 38

Table 2.6 Imaging based CAD prediction studies. 41

Table 2.7: Summary of studies for the carotid artery disease stratification. 44

Table 2.8 Serum biomarkers related to clinical outputs of carotid artery disease. 53

Table 2.9 A summary of studies investigating biomarkers related to carotid artery disease diagnosis. .. 55

Table 2.10 A summary of studies investigating biomarkers related to carotid atherosclerotic plaque vulnerability. 56

Table 2.11 A summary of studies investigating biomarkers related to symptomatic carotid artery disease. 58

Table 2.12 A summary of studies investigating biomarkers related to future stroke event. 59

Table 2.13 A summary of studies investigating biomarkers related to CV mortality. 60

Table 2.14 Biomarkers related with different clinical outputs of carotid artery disease in i) Symptomatic, Asymptomatic & Controls, ii) Asymptomatic, iii) Symptomatic and in iv) General population. 61

Table 3.1 A summary of the parameters of the membership functions for the lumen and the outer wall.72

Table 3.2 Segmentation Validation metrics for CTA images. 89

Table 3.3 Comparison of geometrical metrics derived by the presented methodology with those derived by an IVUS based approach. 90

Table 3.4 Comparison of geometrical metrics derived by the presented methodology with the incorporation of blooming effect removal, without the blooming effect removal and with those derived by an IVUS based approach. 94

Table 3.5 Validation metrics for the CP. 99

Table 3.6 Validation metrics for the NCP. 99

Table 3.7 CTA and VH-IVUS correlation metrics. 99

Table 3.8 CTA and VH-IVUS validation metrics. 100

Table 3.9 Validation metrics against manual annotations for the lumen segmentation. 100

Table 3.10 Validation metrics against manual annotations for the CP segmentation. 101

Table 3.11 Validation metrics against manual annotations for the NCP segmentation. 101

Table 3.12 Comparison of the proposed methodology with the other methodologies for the NCP detection. 109

Table 4. 1 Definition of the utilized CAD risk classes. 113

Table 4.2 SVM RFE Algorithm [242]. 116

Table 4.3 Imaging and non-imaging data utilized. 119

Table 4.4 Evaluation of the CAD risk-prediction problem over 10-fold using imaging and non-imaging data. 120

Table 4.5 Evaluation of the CAD risk-prediction problem over 10-fold using only non-imaging data. 121

Table 4.6 Variables used in the current proposed model. 125

Table 4.7 Classification Performance. 126

Table 4.8 Confusion matrix of Case 3-J48. 126

Table 4.9 Features ranking selection by Case 3-J48. 127

Table 4.10 Performance of the utilized prediction models. 131

Table 4.11 Evaluation of the CAD risk prediction problem over the 10-folds 134

Table 4.12 Comparison of the proposed methodology with existing in the literature studies. 135

Table 5. 1 Univariate Null hypothesis Significance Tests. 141

Table 5.2 CAD Risk Stratification Model Training and Evaluation Procedure. 144

Table 5.3 The Easy Ensemble Algorithm [241]. 144

Table 5.4: Development of Predictive Models. 146

Table 5.5 Baseline Characteristics of training and external validation cohort. 148

Table 5. 6 Subjects characteristics in CAS group and non-CAS group..... 149

Table 5.7 Evaluation of the carotid artery disease risk prediction problem over the 10-folds. 150

Table 5.8 Validation of the carotid artery disease risk problem using an external validation dataset.... 150

Table 5.9 Roc curve analysis identifying symptomatic patients according to intraplaque parameters. . 155

Table 5.10 Baseline Characteristics of training datasets regarding Subproblems 1-5. 156

Table 5.11 Baseline Characteristics of training datasets regarding Subproblem 6. 157

Table 5.12 Baseline Characteristics of training datasets regarding Subproblem 7. 158

Table 5.13 Subjects characteristics in high risk plaques group and low risk plaques group regarding Subproblem 1. 159

Table 5.14 Subjects characteristics in high risk plaques group and low risk plaques group regarding Subproblem 2. 159

Table 5.15 Subjects characteristics in high risk plaques group and low risk plaques group regarding Subproblem 3. 160

Table 5.16 Subjects characteristics in high risk plaques group and low risk plaques group regarding Subproblem 4. 161

Table 5.17 Subjects characteristics in high risk plaques group and low risk plaques group regarding Subproblem 5. 161

Table 5.18 Subjects characteristics in high risk plaques group and low risk plaques group regarding Subproblem 6. 162

Table 5.19 Subjects characteristics in high risk plaques group and low risk plaques group regarding Subproblem 7. 164

Table 5.20 Evaluation of the carotid artery disease risk prediction problem based on the vulnerability of plaques over the 10-folds, regarding Subproblem 1. 165

Table 5. 21 Evaluation of the carotid artery disease risk prediction problem based on the vulnerability of plaques over the 10-folds, regarding Subproblem 2. 166

Table 5.22 Evaluation of the carotid artery disease risk prediction problem based on the vulnerability of plaques over the 10-folds, regarding Subproblem 3. 166

Table 5.23 Evaluation of the carotid artery disease risk prediction problem based on the vulnerability of plaques over the 10-folds, regarding Subproblem 4. 167

Table 5.24 Evaluation of the carotid artery disease risk prediction problem based on the vulnerability of plaques over the 10-folds, regarding Subproblem 5. 167

Table 5.25 Evaluation of the carotid artery disease risk prediction problem based on the vulnerability of plaques over the 10-folds, regarding Subproblem 6. 168

Table 5.26 Evaluation of the carotid artery disease risk prediction problem based on the vulnerability of plaques over the 10-folds, regarding Subproblem 7. 168

Table 5. 27 Summary of the existing in the literature machine learning based studies in comparison with our proposed approach. 180

Table 6.1 Baseline characteristics of Dataset 1. 187

Table 6.2 Baseline characteristics of Dataset 2. 187

Table 6. 3 Baseline characteristics of Dataset 3. 187

Table 6.4 Subjects characteristics in the Class 0 (Absence of Brain lesion) and Class 1 (Presence of Brain lesion) for cortical, lacunar, small subcortical and active infarct, regarding Dataset 1. 189

Table 6.5 Odds Ratio analysis for the association of DS and the presence of Brain lesion, regarding Dataset 2. 189

Table 6.6 Subjects characteristics for continuous input features in the Class 0 (Absence of Brain lesion) and Class 1 (Presence of Brain lesion) for cortical, lacunar, small subcortical and active infarct, regarding Dataset 2. 190

Table 6. 7 Odds Ratio analysis for the association of CVD related risk factors and the presence of Brain lesions, regarding Dataset 3. 191

Table 6. 8 Comparison of our approach with existing in the literature studies. 194

Abstract

This PhD thesis aims to develop models for the predictive modeling of atherosclerotic plaque progression, both in coronary and carotid arteries. In this thesis, active contour models and dynamic threshold segmentation techniques have been implemented for the segmentation of the inner wall, outer wall, CP and NCP in coronary and carotid arteries, using computed tomography angiography images. Additionally, through this thesis machine learning models that utilize both imaging and non-imaging data for the prediction of coronary artery disease were developed, whereas models using only non-imaging data were developed for the carotid artery disease prediction.

The first chapter presents the physiology of the cardiovascular system. More specifically, the function of the circulatory system, the anatomy and function of the heart, the coronary and carotid arteries' anatomy, are presented. Then, the pathophysiology of atherosclerosis and atherosclerosis risk factors are presented. Finally, this chapter reports the imaging modalities of atherosclerosis, both invasive and non-invasive, and the advantages and the disadvantages of each technique in clinical practice.

In the second chapter of this thesis, an extensive presentation of the existing in the literature methods for the three-dimensional reconstruction of the coronary and carotid arteries and the localization of atherosclerotic plaques, both at an automated level and at a non-automated level, is performed. Then, existing studies for the prediction of coronary and carotid artery disease, utilizing either standard statistical analysis techniques or machine learning techniques, are presented. Finally, in this chapter, all the existing biomarkers for the diagnosis and prediction of carotid disease and the mechanism by which they participate in the pathogenesis of the disease, as well as the existing studies in the literature that demonstrate their importance, are presented.

The third chapter describes the proposed methodology for the three-dimensional reconstruction of the inner and outer wall of the coronary and carotid arteries and for the identification and characterization of atherosclerotic plaques (calcified and non-calcified plaques). In addition to this, different processes for validating the proposed methodology are presented, as well as the innovative aspect of the present methodology compared to the existing literature.

The fourth chapter of the thesis aims to present machine learning models for the prediction of coronary artery disease, predicting the obstructive coronary artery disease, the progression of the disease and the

placement of an endovascular stent. The proposed models were trained with non-imaging and imaging data, geometry and blood flow based data.

In the fifth chapter of this paper, machine learning models were proposed to diagnose and identify subjects with asymptomatic carotid disease and participants with the presence of high-risk atherosclerotic plaques, using typical medical records as input.

Finally, in the sixth chapter, the association of carotid artery disease with the presence of clinically asymptomatic brain lesions, was presented. More specifically, the aim of this chapter is to correlate ultrasound markers of the carotid artery, as well as characteristics of each patient (demographic, clinical, hematological, biochemical data and risk factors) with the presence of clinically asymptomatic brain lesions in the ipsilateral hemisphere.

The seventh and last chapter of this paper constitutes a discussion section, related to the contribution of the proposed PhD thesis, as well as to possible future research steps.

Περίληψη

Η παρούσα διδακτορική διατριβή στοχεύει στην ανάπτυξη μοντέλων για την προγνωστική μοντελοποίηση της αθηρωματικής πλάκας, τόσο σε στεφανιαία αγγεία όσο και στις καρωτίδες αρτηρίες. Στα πλαίσια της διατριβής αυτής αναπτύχθηκαν τεχνικές επεξεργασίας αξονικής αγγειογραφίας για την τριστάστατη επεικόνιση των στεφανιαίων αγγείων και καρωτίδων, καθώς επίσης και μοντέλα μηχανικής μάθησης που αξιοποιούν τόσο απεικονιστικά όσο και μη απεικονιστικά δεδομένα για την πρόβλεψη της στεφανιαίας και καρωτιδικής νόσου. Στόχος της διατριβής είναι η επικύρωση των προτεινόμενων μεθοδολογιών και η ανάπτυξη ολοκληρωμένων μοντέλων για την διαστρωμάτωση του κινδύνου στην στεφανιαία και καρωτιδική νόσο.

Η έρευνα της διατριβής ξεκίνησε έπειτα από την έλλειψη αυτοματοποιημένων μεθοδολογιών για τον αυτόματο εντοπισμό του εσωτερικού και εξωτερικού τοιχώματος των στεφανιαίων και καρωτίδων αρτηριών και την τρισδιάστατη ανακατασκευή του πλήρους αρτηριακού δέντρου, καθώς και έπειτα από την μελέτη διαφόρων περιορισμών στον αυτόματο εντοπισμό και χαρακτηρισμό της αθηρωματικής πλάκας. Στη συνέχεια, η διατριβή βασίστηκε στην αξιοποίηση απεικονιστικών δεδομένων που προκύπτουν από μη επεμβατικές απεικονιστικές τεχνικές για την καλύτερη διαχείρηση της στεφανιαίας και καρωτιδικής νόσου.

Στο πρώτο κεφάλαιο αρχικά παρουσιάζεται η φυσιολογία του καρδιαγγειακού συστήματος. Πιο συγκεκριμένα παρουσιάζεται η λειτουργία του κυκλοφορικού συστήματος, η ανατομία και η λειτουργία της καρδιάς, οι στεφανιαίες και καρωτίδες αρτηρίες και η ανατομία τους. Στη συνέχεια παρουσιάζεται η παθοφυσιολογία της αθηροσκλήρωσης, καθώς και παράγοντες κινδύνου για εμφάνιση της αθηροσκλήρωσης και κάποια επιδημιολογικά στοιχεία για την στεφανιαία και καρωτιδική νόσο Τέλος στο κεφάλαιο αυτό παρουσιάζονται οι τρόποι απεικόνισης της αθηροσκλήρωσης, τόσο οι επεμβατικοί, όσο και οι μη επεμβατικοί, τα πλεονεκτήματα και τα μειονεκτήματα αυτών στην κλινική πράξη.

Στο δεύτερο κεφάλαιο της παρούσας διατριβής γίνεται μια εκτενής παρουσίαση της υπάρχουσας βιβλιογραφίας σχετικά με τις μεθόδους τρισδιάστατης ανακατασκευής των στεφανιαίων και καρωτίδων αρτηριών και εντοπισμού των αθηρωματικών πλακών, τόσο σε αυτοματοποιημένο επίπεδο, όσο και σε μη αυτοματοποιημένο επίπεδο. Στη συνέχεια, γίνεται μία λεπτομερής αναφορά σε ήδη υπάρχουσες μελέτες για τη διαστρωμάτωση του κινδύνου για στεφανιαία και καρωτιδική νόσο, αξιοποιώντας είτε τεχνικές τυπικής στατιστικής ανάλυσης είτε τεχνικές μηχανικής μάθησης. Τέλος, στο κεφάλαιο αυτό,

παρουσιάζονται συνολικά όλοι οι υπάρχοντες βιοδείκτες για τη διάγνωση και πρόβλεψη της καρωτιδικής νόσου και ο μηχανισμός με τον οποίο συμμετέχουν στη παθογένεση της νόσου, καθώς και οι υπάρχουσες στη βιβλιογραφία μελέτες που καταδεικνύουν τη σημαντικότητα τους.

Στο τρίτο κεφάλαιο παρουσιάζεται η προτεινόμενη μεθοδολογία για την τρισδιάστατη ανακατασκευή του εσωτερικού και του εξωτερικού τοιχώματος των στεφανιαίων και καρωτίδων αρτηριών και για τον εντοπισμό και χαρακτηρισμό των αθηρωματικών πλακών. Πιο συγκεκριμένα, οι αθηρωματικές πλάκες που εντοπίζονται διαχωρίζονται σε ασβεστοποιημένες και μη- ασβεστοποιημένες πλάκες, που αποτελούνται κυρίως από λιπίδια και ινώδη ιστό. Επιπλέον παρουσιάζονται διαφορετικοί τρόποι επικύρωσης της προτεινόμενης μεθοδολογίας, καθώς επίσης και η καινοτόμος πτυχή της παρούσας μεθοδολογίας συγκριτικά με την υπάρχουσα βιβλιογραφία.

Το τέταρτο κεφάλαιο της διατριβής στοχεύει στην παρουσίαση μοντέλων μηχανικής μάθησης για τη διαστρωμάτωση του κινδύνου για στεφανιαία νόσο, προβέποντας την παρουσία στεφαναίας παθολογικής νόσου, είτε την εξέλιξη της νόσου, είτε την τοποθέτηση ενδαγγειακού στεντ. Τα προτεινόμενα μοντέλα εκπαιδεύτηκαν με μη απεικονιστικά και απεικονιστικά δεδομένα, που προκύπτουν από την γεωμετρία των αγγείων, καθώς και την μοντελοποίηση της αιματικής ροής των αγγείων.

Στο πέμπτο κεφάλαιο της παρούσας εργασίας, προτάθηκαν μοντέλα μηχανικής μάθησης για τη διάγνωση και την αναγνώριση ατόμων με ασυμπτωματική καρωτιδική νόσο και για την ανίχνευση ατόμων με παρουσία αθηρωματικών ασταθών πλακών υψηλού κινδύνου, χρησιμοποιώντας ως είσοδο τυπικά ιατρικά δεδομένα.

Τέλος, στο έκτο κεφάλαιο παρουσιάστηκε η συσχέτιση της καρωτιδικής νόσου με την παρουσία κλινικά ασυμπτωματικών εγκεφαλικών βλαβών. Πιο συγκεκριμένα, στόχος του κεφάλαιου αυτού ήταν η συσχέτιση υπερηχογραφικών δεικτών της καρωτιδικής αρτηρίας, καθώς και χαρακτηριστικά του κάθε ασθενούς (δημογραφικά, κλινικά, αιματολογικά, βιοχημικά δεδομένα και παράγοντες κινδύνου) με την παρουσία κλινικά ασυμπτωματικών εγκεφαλικών βλαβών στο ομόπλευρο ημισφαίριο.

Το έβδομο και τελευταίο κεφάλαιο της παρούσας εργασίας αποτελεί μια ενότητα συζήτησης, που σχετίζεται με τη συνεισφορά της προτεινόμενης διδακτορικής διατριβής, καθώς και με τα πιθανά μελλοντικά ερευνητικά βήματα.

Chapter 1 Introduction

1.1 Physiology of the cardiovascular system

1.2 Central nervous system vasculature

1.3 Atherosclerosis

1.4 Imaging of atherosclerosis

1.1 Physiology of the cardiovascular system

1.1.1 Introduction

The cardiovascular (CV) system is responsible for providing blood supply throughout the body and its purpose is the transportation of nutrients, oxygen, carbon dioxide, hormones, and blood cells to and from the cells in the body. The main components of the CV system are the heart, the arteries, the veins and the capillaries. The heart and the vessels are responsible for providing adequate blood flow to all parts of the body. The CV system is regulated by a myriad of stimuli including changing blood volume, hormones, electrolytes, osmolarity, medications, adrenal glands, kidneys and much more [1].

The heart is the organ that pumps the blood through the vessels. It pumps blood directly into the arteries, more specifically the aorta or the pulmonary artery. Blood vessels, including arteries, capillaries, and veins, which transport blood to and from the heart, consist of the circulation system, as it is shown in Figure 1.1. Blood vessels play a significant role in the function of circulation system, since they control the amount of the blood flow to specific parts of the body. Arteries are distinguished between large and small arteries and are responsible to transport the blood away from the heart. Large arteries receive the highest pressure of blood flow and are more thick and elastic to accommodate the high pressures, whereas small arteries, such as arterioles, have more smooth muscle which contracts or relax to regulate blood flow to specific portions of the body. On the other hand, capillaries branch off of arterioles and

are made from a single cell layer. This thin layer allows the exchange of nutrients, gases, and waste with tissues and organs. Also, the veins transport blood back to the heart.

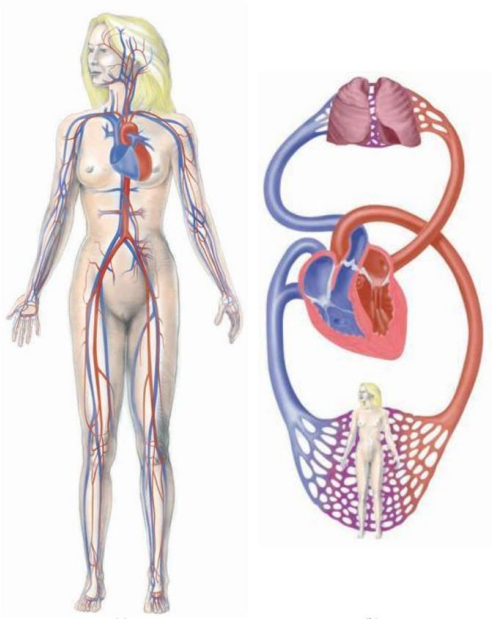

Figure 1.1 The circulation system [2].

1.1.2 The Cardiovascular system

The heart, a muscular organ, which is located between the lungs, in the middle compartment of the human chest, supplies blood through the blood vessels of the circulatory system. It is surrounded by a membrane called pericardium, which is a fibroserous sac that encloses the heart and the roots of great vessels. It is a cone-shaped organ and its size in an adult is about 12 cm in length, 8 to 9 cm in breadth at the broadest part, and 6 cm in thickness. The heart is subdivided by septa into right and left halves, and a constriction subdivides each half of the organ into two cavities, the atrium (upper cavity) and the ventricle (lower cavity). Therefore, the heart consists of four chambers, the right and left atria and the right and left ventricles [3].

The heart includes also four valves, two atrioventricular and two semilunar. The right atrioventricular, known also as tricuspid, separates the right atrium from the right ventricle preventing backflow into the atrium. The left one, known as bicuspid or mitral valve, separates the left atrium from the left ventricle. The semilunar valves are the pulmonary and the aortic. The first one separates the right ventricle from the pulmonary arteries and prevents backflow after ventricular contraction, while the second one

separates the left ventricle from the aorta and prevents backflow after ventricular contraction. Anatomical features of the heart are shown in Figure 1.2 [3].

The arterial supply of the heart is provided by the right and left coronary arteries which arise from the aorta immediately above the aortic valve. Most of the blood from the heart walls drains into the right atrium through the coronary sinus, which lies in the posterior part of the atrioventricular groove. The heart is innervated by sympathetic and parasympathetic fibers of the autonomic nervous system via the cardiac plexuses situated below the arch of the aorta.

Coronary arteries are the only branches of the ascending aorta. Traditionally, a coronary artery has been described as any artery or arterial branch that carries blood to the cardiac parenchyma. The cardiac parenchyma is defined as any structure located in the pericardial cavity and includes not only the myocardium, but also structures such as the pulmonary truck, the superior vena cava, and the semilunar valves. Coronary arteries are located on the epicardial surface of the heart [4].

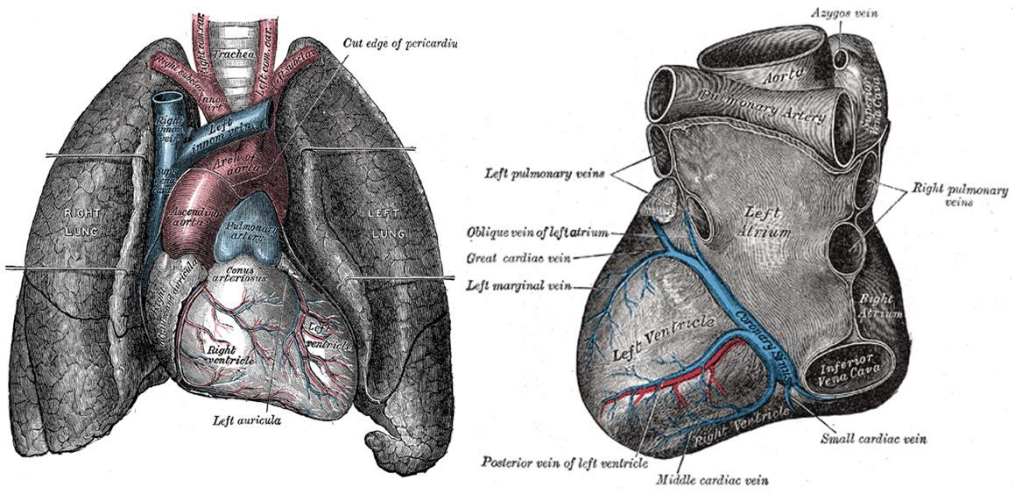

Figure 1.2 Heart anatomy [3].

There are two main coronary arteries, the left main coronary artery (LMCA) and the right coronary artery (RCA). The LMCA artery originates from the left sinus of valsalva and travels anteriorly and leftward. It is positioned between the left atrial appendage and the pulmonary trunk and is divided into two major branches: the left anterior descending (LAD) and left circumflex (LCX) arteries. The LAD extends from the left main and curves around the pulmonary trunk prior to entering the anterior

interventricular groove and extending to the apex and it then extends distally to the apex within the inferior interventricular sulcus towards the crux of the heart. It then provides branches to the inferior walls of both ventricles. The LCX artery provides two major branches: septal perforator arteries and diagonal branches. The septal perforator arteries branch at right angles from the anterior descending artery and supply the anterior two thirds of the intraventricular septum. On the other hand, the RCA artery originates from above the right cusp of the aortic valve and supplies blood to the right ventricle, the right atrium, and the sinoatrial and atrioventricular nodes. The coronary arteries, as it is shown in Figure 1.3, run along the outer surface of the heart [5].

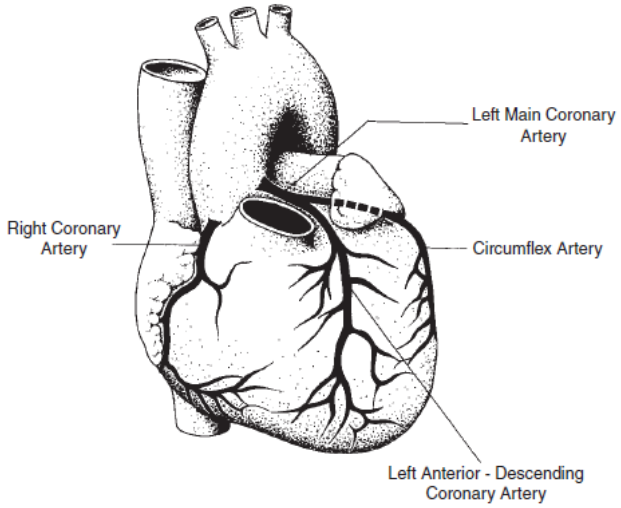

Figure 1.3 The coronary arteries [2].

1.1.3 Arterial morphology

The arteries are the blood vessel that deliver oxygen-rich blood from the heart to the tissues of the body. In general, the arteries are distincted into two main categories: the elastic and the muscular. The elastic arteries are characterized by large diameters and are located close to the heart. On the other hand, the location of the muscular arteries is at the periphery, such as the femoral and the iliac arteries. Each artery is considered as a muscular tube lined by smooth tissue, which consists of three histologically distinct regions, the intima, the media and the adventitia layer, as it is shown in Figure 1.4 [4, 6, 7].

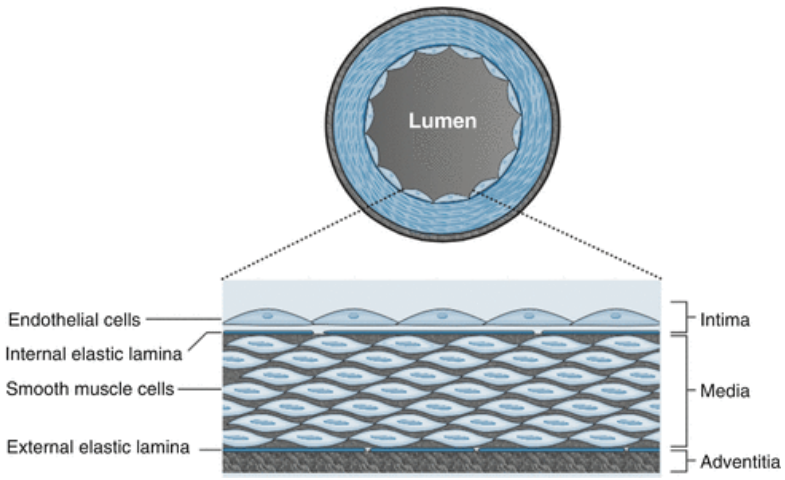

Figure 1.4 Arterial morphology [4].

The intima is the inner layer of arteries and veins. This layer consists of an elastic membrane lining and smooth endothelium that is covered by elastic tissues. An endothelial membrane lies between the lumen and the intima wall layer. Often, the internal elastic lamina separates the tunica intima from the tunica media. This layer is structured by a smooth lining of endothelial cells, which provides the opportunity to the arteries to cope with the high-pressure surges of blood from the heart because they have very thick muscular walls [8]. The pathological alterations of intimal components is directly associated with atherosclerosis, a common disease of arterial wall, since the intima layer and its endothelial lining contribute to the inner smooth surface of blood vessels, the anticoagulant characteristics of the vessel wall and the expression of adhesion molecules walls [8]. Atherosclerosis is a common disease of arterial wall, involving the deposition of calcium, collagen fibers, cellular waste products and fatty substances, resulting to the build-up of atherosclerotic plaque. These alterations contribute to significant changes in the mechanical behavior of atherosclerotic arteries, which significantly differ from the healthy ones.

The media is the middle layer of the artery, composed mainly of smooth muscle cells (SMCs), a network of elastic and collagen fibrils and elastic laminae which separate the media into a number of fiber-reinforced layers. The media consists of a highly organized three-dimensional network of elastin, vascular SMC and collagen with extracellular matrix proteoglycans. As it was found recently [9] the media behaves mechanically as if its material properties were homogeneous.

The adventitia is the outer layer composed primarily of thick bundles of collagen fibrils arranged in helical structures and fibroblast cells. The tunica externa, also known as the tunica adventitia (or adventitia for short), is the outermost tunica (layer) of a blood vessel, surrounding the tunica media. It is

mainly composed of collagen and, in arteries, is supported by an external elastic lamina. The collagen serves to anchor the blood vessel to nearby organs, giving it stability [10].

Undoubtedly, the artery wall constituents have a significant growth interest in the mechanical properties of biological soft tissue, since there is a strong belief that the mechanical factors play a significant role on the atherosclerotic plaque development and progression [11].

1.2 Central nervous system vasculature

The common carotid artery (CCA) is a large elastic artery, which provides the main blood supply to the head and neck region. There is one CCA on either side of the body and these arteries differ in their origin. The left CCA arises from the aortic arch within the superior mediastinum, whilst the right CCA arises from the brachiocephalic trunk posterior to the right sternoclavicular joint. The CCA ascends lateral to the trachea and esophagus within the deep cervical fascia, the carotid sheath, with the internal jugular vein and the vagus nerve [12]. The CCA ascends in the neck medial to the internal jugular vein and normally has no branches. Occasionally, the superior thyroid artery arises proximal to the bifurcation into internal and external carotid arteries (ECAs). The bifurcation is usually located at the level of the superior border of the thyroid cartilage. Variations in the levels at which the carotid bifurcates are more often above this position than below. The ECA, supplying the extracranial structures of the head, gives off several branches before its terminal bifurcation into the internal maxillary and superficial temporal arteries. These are the superior thyroid, ascending pharyngeal, lingual, facial, occipital, and posterior auricular arteries. The internal carotid artery (ICA) proceeds posteromedially to enter the carotid canal at the base of the skull without giving off any branches. On the medial side of the bifurcation lie the small, oval carotid body, a chemoreceptor, and the carotid sinus, a pressure receptor intrinsic to the wall of the common and internal carotid arteries. It supplies the forehead, nose, eyes and the ipsilateral cerebral hemisphere [13]. The basic anatomy of carotid arterial tree is shown in Figure 1.5 , below.

Near its bifurcation, the CCA forms two specialized structures, the carotid sinus and the carotid body. The carotid sinus is a dilation of the base of the ICA, which is involved in relaying information about the arterial blood pressure to the hypothalamus. It is therefore referred to as a baroreceptor and is innervated by the carotid branch of the glossopharyngeal nerve. On the other hand, the carotid body is an oval structure, located posterior to the carotid bifurcation, involved in relaying information about the arterial chemical composition to respiratory centers in the brainstem. Like the carotid sinus, it is innervated by

the carotid branch of the glossopharyngeal nerve. The carotid body is surrounded by a fibrous capsule and consists of multiple lobules divided by septa. Within each lobule, there are two types of cells: glomus (type I) cells and sustenacular (type II) cells. The glomus cells are involved in storing peptides, such as neurotensin, and amines, such as adrenaline, noradrenaline and dopamine. The sustentacular cells separate the glomus cells from an extensive network of fenestrated sinusoids. The carotid body is a chemoreceptor stimulated by hypercapnia, hypoxia and increased hydrogen ion concentration (low pH). In response to these changes, the carotid body changes the rate and volume of respiration via a reflex involving the respiratory centres in the brainstem [12].

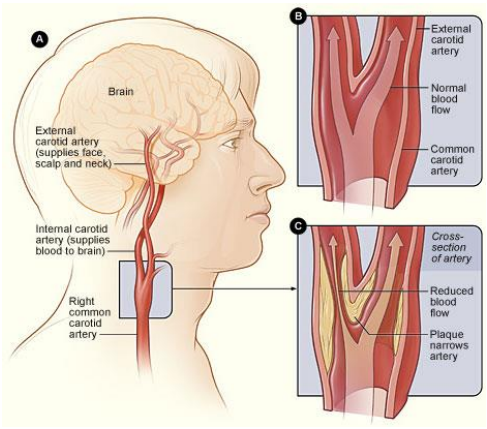

Figure 1.5: Basic Anatomy of carotid arterial tree [14].

1.3 Atherosclerosis

1.3.1 Introduction

Atherosclerosis, the leading cause of morbidity and mortality worldwide, derives its name from the Greek words “athere” meaning the soft lipid-rich material in the centre of atheroma and “sclerosis”, referring to connective tissue in the plaques [15]. The most commonly affected by atherosclerosis arteries include those of medium and large size, such as the coronary arteries, the aorta and cerebral arteries.

Atherosclerosis is a chronic inflammatory disease caused by high concentrations of lowdensity lipoprotein (LDL) cholesterol [16] and is characterized by the vascular obstruction from the deposition of atherosclerotic plaque, which results in reduced blood flow [17]. Atherosclerosis begins with the

accumulation of lipid laden foam cells in the intima layer of the artery. Lipid retention is the first step in the pathogenesis of atherosclerosis which is followed by the chronic inflammation at susceptible sites in the walls of the major arteries [18], causing plaque growth.

1.3.2 Mechanism of Atherosclerosis

During the past years the understanding of atherosclerosis mechanism has undergone a remarkable progression, since atherosclerosis was considered as a cholesterol storage disorder. Currently, atherosclerosis is considered as an inflammatory disorder and the appreciation of arterial remodeling has expanded attention beyond stenoses evident by angiographic imaging modality to include the biology of nonstenotic atherosclerotic plaques [19]. Atherosclerotic plaque development starts from the accumulation of lipids in the arterial wall to advanced atherosclerotic lesions, characterized by an inflammatory response, the proliferation of SMCs and the thickening of the arterial wall. The mechanism of atherosclerosis pathogenesis is a complex process and its major steps are described below.

Atherosclerotic disease refers to the development of the atherosclerotic plaque in the inner wall of the arteries. The inner layer of coronary arteries is lined by endothelial cells, which are in contact with the blood and normally resist the attachment of the white blood cells streaming past them. Arterial endothelial cells, as it is shown in Figure 1. 6b, express adhesion molecules able to capture leukocytes on their surfaces. The alterations of endothelium permeability in combination with changes in the composition of the extracellular matrix beneath the endothelium allow the retention of cholesterol-containing LDL particles in the artery wall. Biochemically modified LDL particles may induce the leykocyte adhesion, whereas modified particles undergo endocytosis by monocyte-derived macrophages, leading to intracellular cholesterol accumulation.

When monocytes which are the most numerous white blood cells in plaques, reside in the arterial wall, they differentiate into macrophages. Inside the growing atheroma, the mononuclear phagocytes surround the lipoprotein particles and are transformed to foam cells. The formation of atheroma involves the recruitment of SMCs from the middle layer of the arterial wall into the tunica intima (Figure 1. 6c). The proliferation of resident intimal SMCs and the synthesis of extracellular medium macromolecules. These macromolecules include elastin, proteoglycans and collagen. SMCs and macrophages die in advanced lesions by apoptosis. The extracellular lipid which is derived from the dead cells can build up in the plaque composing the lipid pool or necrotic core (NC) of the plaque. Plaque lesions generally cause stenoses and produce limited blood flow which can lead to tissue

ischaemia (Figure 1. 6d). Thrombi are generated when the fibrous cap of a plaque ruptures and enables blood coagulation components to come into contact with the thrombogenic plaque [20].

Among prominent risk factors that promote the atherosclerosis are hypertension, diabetes mellitus (DM), dyslipidemia, obesity, sedentary lifestyle, family history, smoking and genetic predisposition. Atherosclerosis diagnosis is clinical and the definitive diagnosis is made through imaging tests. Disease management plan includes some behavior modifications, such as the physical activity with low caloric diet, rich in fiber component, whereas the main class of drugs used in treatment are the antiplatelet drugs and the antiatherogenic drugs [18].

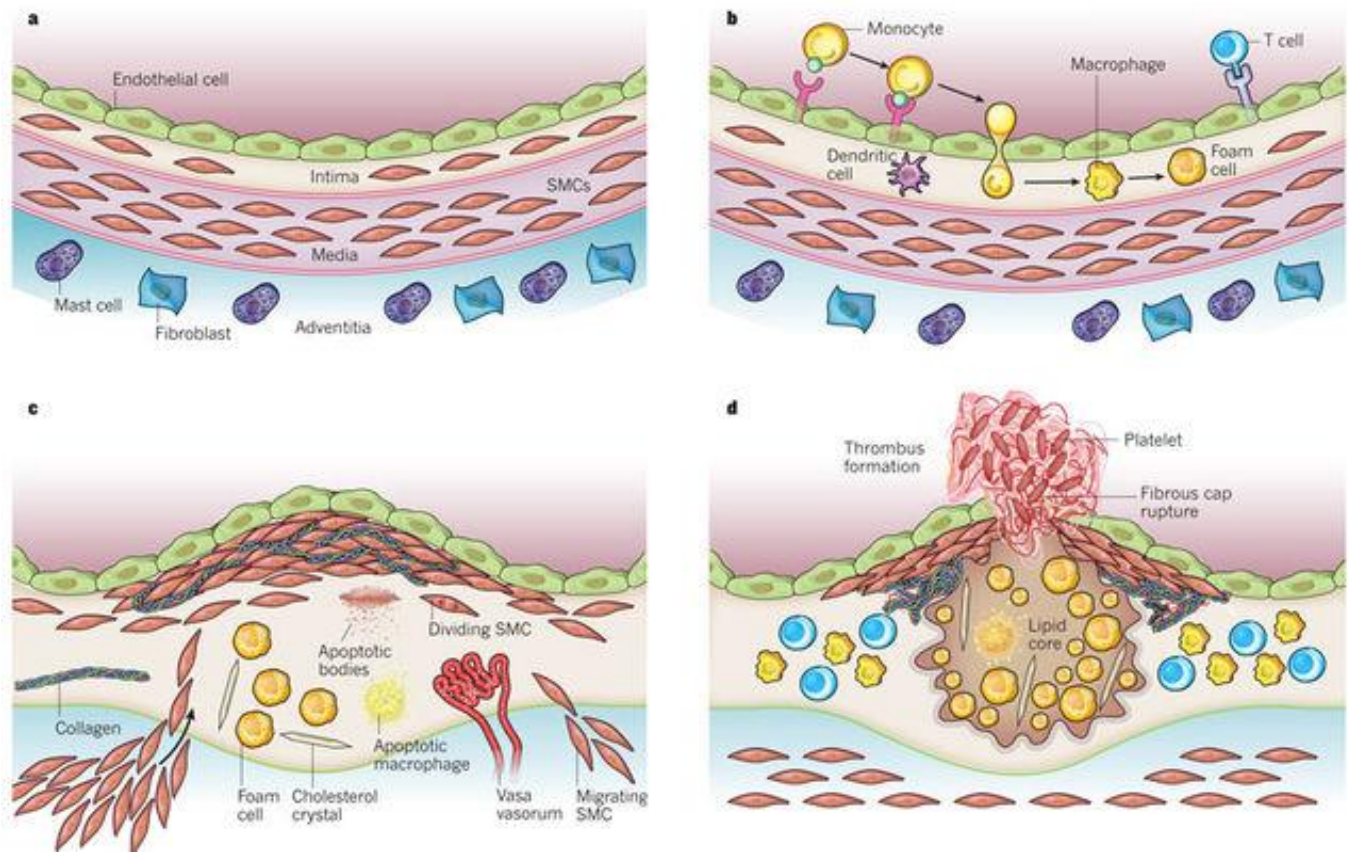

Figure 1. 6 Stages in the development of atherosclerotic lesions [20].

1.3.3 Coronary Artery Disease

Coronary artery disease (CAD) is a common public health disease, due to the high mortality rate of its acute clinical manifestations, such as major adverse coronary events (MACE) and the related costs for the health care system, as high as 27% of total cardiovascular disease (CVD) costs. According to the

latest epidemiological available data (WHO Mortality Database) more than 4 million deaths each year across Europe are caused by CVD, reaching 45% of all deaths. Specifically, the number of deaths from CVD is higher in women reaching the 49% of all deaths and 40% in male gender. CAD accounts for the 20% of total deaths for the females and 19% respectively for the men [21].

The gold standard of CAD diagnosis is the invasive coronary angiography (CA), a widely used procedure in the clinical practice, according to the guidelines for assessing CVD. CA reduces drastically the misdiagnosis of the high burden CAD patients but concurrently it may be used excessively. It has been reported that only 41% of patients undergoing CA have obstructive CAD, meaning that more than 50% of the procedures may be unnecessary [22]. Nowadays both the European and the American guidelines put great emphasis on the pre-test patient stratification, according to their risk of having an obstructive CAD, potentially leading to MACE, by the use of clinical PTP (pre-test probability) models. The main economic and process-of-care advantage of PTP models relies in balancing the diagnostic accuracy in identifying patients likely to have obstructive CAD with accuracy in identifying those who are unlikely to benefit from potentially expensive testing and who may be managed conservatively. Reduction in unnecessary testing means saving time, anxiety, and cost for patients, limiting radiation exposure and in lowering the rate of false-positive test results that could lead to more invasive, unnecessary procedures. The PTP models will thus serve as a "System-Firewall" identifying the high risk ones who need to be appointed to more thoroughly diagnostic assessment [23, 24]. At the moment there are various such validated diagnostic models of the PTP of stable obstructive CAD, which most of them are reported in Chapter 2 of the current thesis. Briefly, in 1979 Diamond and Forrester (the Diamond-Forrester score) [25] proposed an easy and simple PTP model of obstructive CAD, still recommended by the American Heart Association (AHA) guidelines. The European Society of Cardiology (ESC) guidelines for stable CAD have replaced the DF score with 2 new revised scores (CAD Consortium Basic and CAD Consortium Clinical) [26, 27].

1.3.4 Carotid Artery Disease

Carotid artery disease, the build-up of atherosclerotic plaques in carotid bifurcations, is a highly prevalent and devastating disease of our times, with enormous socioeconomic burden. It constitutes the primary cause of cerebrovascular events and ischaemic stroke, and accounts for up to 30% of all strokes. In the European Union, this translates to more than 150.000 deaths annually and over €12 billion per year in direct and indirect costs.

The disease is characterized by the development of an atherosclerotic plaque inside the artery wall that reduces blood flow to the brain and increases the risk for transient ischemic attack (TIA) or stroke. Some carotid plaques are soft and tend to corrupt or to create irregular projections in the carotid lumen. This process consists an “inner trauma” of the human organism and in order to heal it, thrombus is created through platelets. This thrombus can grow and interrupt the blood flow into the carotid artery. This situation provokes the stroke, which consists a medical emergency that occurs when the brain loses all or much of its blood supply.

Carotid artery disease often develops slowly and is either asymptomatic or symptomatic, with the first clinical signs of the condition might be a stroke or TIA. A TIA is a temporary shortage of blood flow to the brain which may cause burning sensation or weakness in one side of the body, inability to control the movements of the limbs (upper and lower), loss of vision from the one eye and inability to talk clearly.

With suspicion of carotid artery disease presence, the patient should undergo some imaging examinations, including as first line examination the US of carotid arteries. Supplementary examinations are the CTA of the brain, the angiography of the cervical and intracranial vessels and the brain MRI, basically for evaluating the extent and severity of extracranial carotid stenoses [28].

Regarding the management of carotid artery disease, it involves lifestyle alterations, medication therapy and in some cases surgery. More specifically, in compliance with the current guidelines (shown in Figure 1.7), asymptomatic patients with 70% stenosis in their carotid artery, are considered to be at high risk of cerebrovascular events and are therefore directed to surgical intervention, carotid endarterectomy (CEA) or carotid artery stenting. In contrast, asymptomatic patients with <60% stenosis are considered to be at low-intermediate risk, and unless other confounding factors exist, they are subjected to medical treatment alone. For recently symptomatic patients, i.e. patients who have had symptoms, such as TIA or stroke (minor non disabling strokes) within the previous 6 months, the cut-off point of 50% stenosis for surgical intervention, is used. However, it is not uncommon these days to treat symptomatic patients with 50-69% stenosis but not vulnerable plaque features, as determined by US, MRI or CT angiography with medical treatment alone. For previously symptomatic patients, i.e. patients with symptoms before the last 6 months, treatment is usually conservative and involves the administration of medical therapy [29, 30].

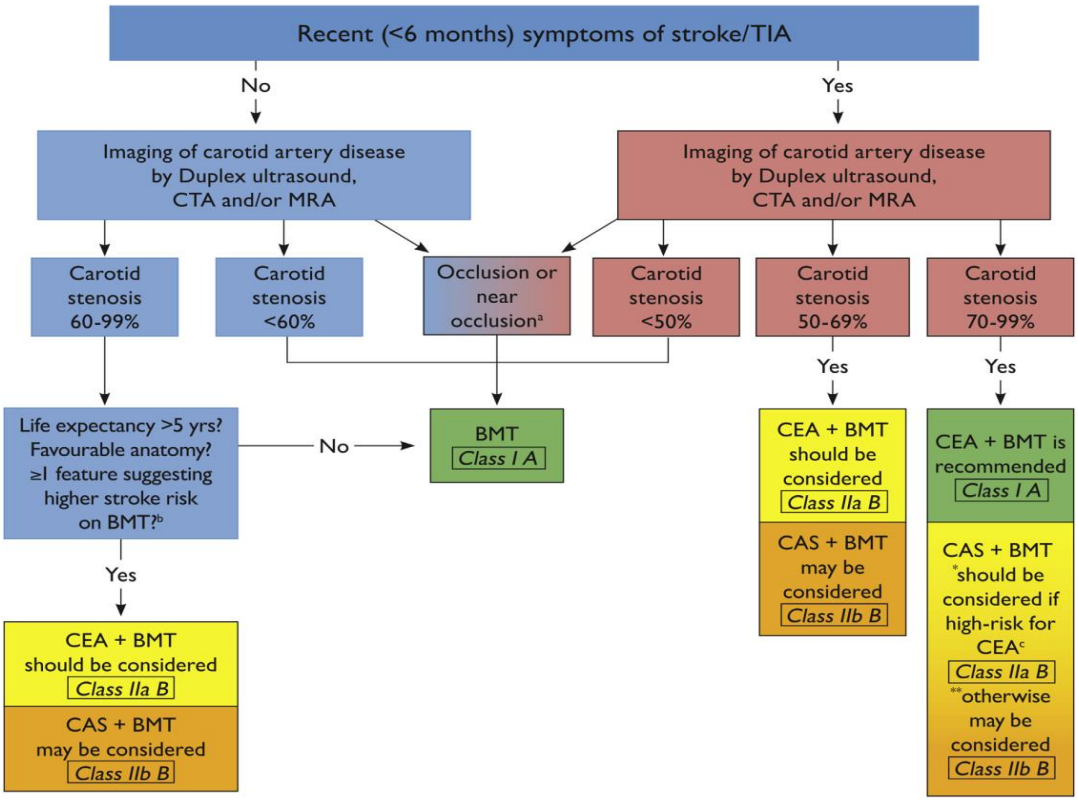

Figure 1. 7 Current guidelines for carotid artery disease management [30].

1.4 Imaging of atherosclerosis

1.4.1 Imaging of Coronary Artery Disease

CAD, the most common type of heart disease, is according to the `AHA one of the main causes of death worldwide [31]. Due to the awareness of the mortality rate of CAD, different cardiac imaging modalities, mostly imlemented shown in Figure 1.8, have been developed over the last years to provide accurate and early diagnosis of CAD and to support patient's management. CA is a predominant method for the detection of CVD. However, despite the broad implementation of this technique, angiographic imaging portrays only the two-dimensional (2D) silhouette of the lumen and fails to depict the entire circumference of the vessel wall. Contrary to this, different modalities overcome this limitation, investigating the changes of the composition of atherosclerotic plaque and evaluating accurately the arterial pathology [32].

Thus, alternative invasive and non-invasive imaging modalities, which have their own strengths and weaknesses are nowadays available and able to detect atherosclerosis. More specifically, invasive

[intravascular ultrasound (IVUS), optical coherence tomography (OCT), Near-infrared spectroscopy (NIRS)] and non-invasive techniques [computed tomography angiography (CTA), magnetic resonance imaging (MRI), positron emission tomography computed tomography (PET-CT)] provide reliable assessment of the lumen pathology, quantification of the plaque burden (PB) and characterization of atherosclerotic plaque composition [33, 34].

Invasive Cardiovascular imaging modalities

IVUS is a widely used technique, not only in clinical applications, but also in research, especially in the study of plaque evolution, due to its ability to visualize the vessel wall with increased accuracy. The required equipment to perform IVUS modality consists of a catheter, a transducer, a pullback device and a scanning console. IVUS catheter incorporates an US transducer combined usually with an inflatable balloon either with a stent, or without, in order to expand narrowed areas. Recent developments in IVUS image processing combined with the analysis of intravascular ultrasound radiofrequency (IVUS-RF) backscatter signal allow the evaluation of the atherosclerotic plaque composition [35]. There are several advantages for the use of IVUS imaging in the evaluation of CAD. First of all, the full circumference of the vessel wall and not just two surfaces can be visualized, due to its tomographic orientation. Thus, a reliable assessment of vessels is provided, including cases, such as ostial or bifurcation stenoses, diffusely diseased segments, eccentric plaques, which are difficult to be assessed by angiography [36]. Moreover, IVUS has the potential to quantify the percentage of narrowing. Its penetrating nature provides remarkable images of the atherosclerotic plaque and provides insight into the nature of the plaque. Finally, another important benefit by the use of IVUS is that it may reveal what in the past has been referred to as “re-stenosis”, which may be a recurrence of the plaque buildup that may have previously been removed. Additionally, IVUS provides the physician the ability to identify buildup that may have been missed during angiogram or angioplasty. On the other hand, IVUS still has indigenous limitations, such as several artifacts and speckle noise, which affect the clinical interpretation of the IVUS image and the processing of medical images for computer-aided diagnosis, regardless of the type of the catheter [34, 37].

OCT imaging modality utilizes near infrared light and relies on measuring the time delay and the magnitude of the backscattered light to generate the final images. OCT provides a detailed visualization of the vessel wall, better plaque characterization than IVUS, visualization of plaque micro-features, such as the presence of microcalcifications, neovascularization, and plaque hemorrhage [38].. In addition to

this, OCT could also detect stent fracture and early thrombus adhesion at the site of stent fracture. Although OCT provides a high image axial resolution, its poor signal penetration prevents the visualization of the outer vessel wall. In addition to this, OCT modality fails to provide the any details about the vessel geometry and as result information about the distribution of the plaque into the three dimensional (3D) space [34, 39]

NIRS uses the absorption, emission or scattering of light in the near-infrared portion of the electromagnetic spectrum (700-3000 nm) by atoms or molecules to determine sample composition or characteristics. This modality relies on the principle that different organic molecules absorb and scatter NIRS light with different degrees and wavelengths. This CV imaging has the potential to determine the chemical composition of atherosclerotic plaques *in vivo*, such as the lipid-rich plaques, improving possibly the patient risk management and the guiding therapy selection [40]. NIRS takes advantages of the non-ionizing radiation dose, which does not damage the tissue and of the good depth of penetration, which is about 2 to 3 mm. On the other hand, NIRS modality does not provide any details about the vessel geometry and the distribution of plaque in the 3D space.

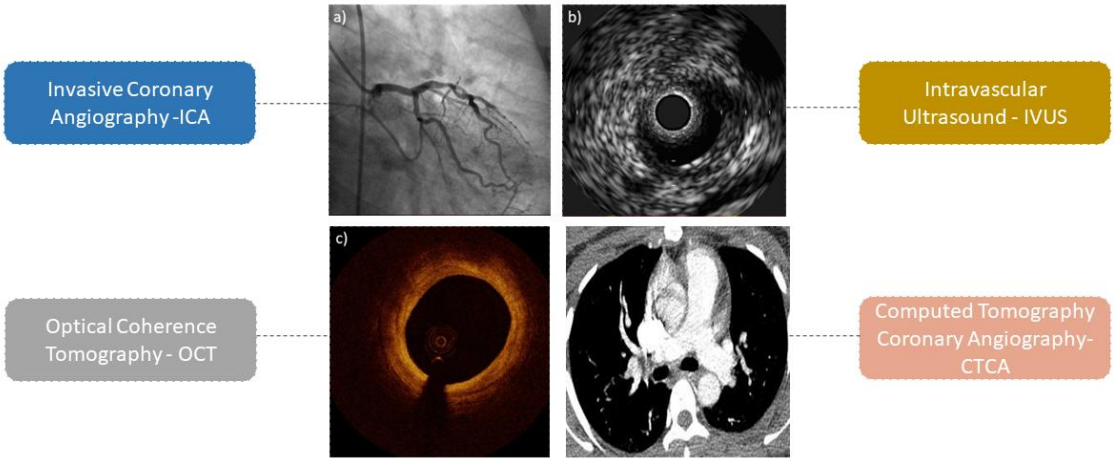

Figure 1.8 Imaging Techniques of Coronary Artery disease [41].

Non-invasive Cardiovascular imaging modalities

Cardiac MRI, an imaging technique with a high potential to visualize vessel anatomy, is based on the differences in biophysical response of the tissue to an electromagnetic radiofrequency pulse application within a strong, static magnetic field. MRI is a diagnostic imaging modality with excellent soft tissue contrast and therefore able to analyze the components of atherosclerosis plaque components [51,52].

Computed tomography coronary angiography

Computed Tomography (CT) modality is a medical imaging technique that utilizes computer-processed X-rays to produce tomographic images of specific parts of the body. CT scanners take advantages of the digital geometry processing to generate a 3D image of the internals of an object from a large series of 2D X-ray images taken around a single axis of rotation. CT scan, contrary to the conventional X-ray imaging, forms a full 3D computer model of patient's body organs. Since its invention in the 1970s, CT scan modality is considered as the ultimate choice for the clinical diagnosis of body organs.

CT images are acquired on the basis of the ability of body organs to block x-rays, which is called radio density. Attenuation values are recorded by CT scanner systems in a plane of a finite slice thickness for the whole cross-section, where every component of the cross section image represents a pixel. More specifically, attenuation value is represented mathematically as

$$I_t = I_0 e^{-\mu \Delta x}, \tag{1.1}$$

where I_t represents the attenuation value of the examined object, I_0 corresponds to the intensity in the beam path without any obstruction and μ is the linear transformation coefficient of a specific material.

However, in clinical CT dataset, the attenuation value is represented in Hounsfield Units (HU) scale. HU scale transforms the original linear attenuation into one in which the radiodensity of distilled water at standard temperature and pressure (STP) is defined as 0 HU, whereas the radiodensity of air at STP is defined as -1000 HU. More specifically, in a voxel with average linear attenuation coefficient given as μ, the corresponding HU value is given by:

$$HU = 1000 * \frac{\mu - \mu_{water}}{\mu_{water} - \mu_{air}}, \tag{1.2}$$

where μ_{water} and μ_{air} are the liner attenuation coefficients for water and air, respectively.

However, due to the "motion" of the heart, CTA has not been widely utilized for imaging the cardiac system. Imaging of coronary arteries requires high spatial acquisition imaging systems and additionally a high temporal resolution is required to emulate the dynamic heart as a "static organ". Recently, advancements in medical imaging introduced the multi-slice CT system that guarantees both higher spatial resolution and lower acquisition speed time. Dual source technique reduced the acquisition time by imaging in half rotation, whereas the multi slice system improved the spatial resolution. Thus, high diagnostic accuracy has been achieved with multislice CT scanners (64 slice and higher), and in selected patients coronary CTA is regarded as a reliable alternative to invasive coronary angiography [42].

CTA is a cardiac imaging modality for the detection of CAD and the exclusion of significant CAD. The basic clinical value of CTA is its high NPV and the portion of patients who could undergo this modality consists of symptomatic patients with intermediate CV risk [43]. Different studies have indicated that the introduction of 64-slice and 320- slice systems allows not only the identification of coronary stenosis, but also provides information about the composition and the distribution of atherosclerotic plaque in the 3D space [34, 44, 45]. While CTA modality can non-invasively detect atherosclerotic plaques and coronary artery stenosis, it is directly associated with the exposure in high radiation dose. This is even eminent in the hybrid cardiac imaging field, where the patients are additionally exposed to even higher radiation dose from nuclear perfusion scanning [46].

1.4.2 Imaging of Carotid Artery Disease

The technical improvement of imaging of carotid artery disease has been shifted from 2D to 3D and various imaging modalities have been implemented for the CAS diagnosis, including US, CT and magnetic resonance angiography (MRA).

US is considered as the first line examination for CAS diagnosis, due to its accuracy and non-invasive nature. This modality fuses two modes: the B-mode in which images are produced by the reflected sound waves and the colour-Doppler US mode that visualizes the motion of the moving parts or fluid (blood) to measure the speed and other flow parameters [47]. CTA is also considered as an accurate and non-invasive technique for imaging intracranial and extracranial blood vessels and was found to be an excellent examination for the detection of carotid occlusion and categorization of stenosis in either the 0%–29% or >50% ranges. MRA is also a non-invasive modality that uses the combination of MRI technology and intravenous contrast to visualize blood vessels. Compared to US, MRA was hound to be more accurate in identifying severe cases of CAS (>70%) and cases of total occlusion [48].

Chapter 2 Literature Overview

2.1 Introduction

2.2 Three-dimensional reconstruction based on Computed Tomography

2.3 Prediction of Coronary Artery Disease

2.4 Prediction of Carotid Artery Disease

2.5 Serum Markers in Carotid Artery Disease

2.6 Contribution of this thesis

2.1 Introduction

The recent years, technical breakthroughs, in the field of coronary and carotid imaging and identification of novel biomarkers, have enabled the progress of CVDs risk stratification. A remarkable growth in scientific publications on personalized medicine in the field of CVDs has been notified, including the use of new medical image processing techniques and the implementation of ML techniques for the early disease diagnosis and prognosis. The aim of this chapter is to present the existing in the literature studies regarding the techniques for CTA image processing, techniques for CAD and carotid disease prediction and to present the carotid artery disease related biomarkers.

More specifically, at first techniques for the CTA image analysis for the detection of coronary and carotid arteries and the distribution of atherosclerotic plaques, are presented. Most of these studies are based on well-known image processing techniques, whereas others have been developed based on existing software tools with the need of manual user interaction. The technical pipeline of these studies, their evaluation and their accuracy are presented. Subsequently, a presentation of techniques for CAD risk stratification is performed. The presented studies distinct into non-imaging based studies and imaging based studies and their outcomes vary according to the proposed problem definition of each

study. More specifically, most of the studies aim to detect either the future presence of CAD, as it is defined by DS and number of occluded vessels or the progression of CAD.

Subsequently, we present studies regarding the carotid artery disease prediction. Studies concerning the prediction of carotid artery disease, taking advantage of ML techniques, are presented and their outcomes are basically based on the presence of either CAS or the presence of high risk carotid plaques. Additionally, a detailed review for serum biomarkers related to carotid artery disease presence, is reported.

2.2 3D reconstruction based on computed tomography angiography

Technological developments of CV imaging have led to the development of different techniques for the 3D reconstruction of coronary and carotid arteries. CTA modality is a noninvasive imaging modality, which utilizes computer-processed X-rays to produce tomographic images of arteries and has experienced a remarkable progress in the last years in the field of 3D reconstruction of coronary [49, 50] and carotid arteries [48]. The progressive improvements of CT scanners, as far as the spatial resolution and the acquisition time are concerned, allow the visualization of the entire coronary arterial tree and carotid bifurcation. More specifically, the 2D CTA slices visualize accurately the inner wall, the outer wall and the atherosclerotic plaques of arteries and provide information about their composition. Thus, the detailed visualization of CTA slices in combination with the tomographic orientation of CTA, allow the reliable reconstruction of artery inner wall, outer wall, atherosclerotic plaques and their distribution in the 3D space. In addition to this, CTA imaging allows the reconstruction either of a bifurcated artery segment or the full reconstruction of the arterial tree.

2.2.1 3D reconstruction of coronary arteries based on CTA

Regarding the coronary artery reconstruction, different methodological solutions have been implemented for the accurate coronary vessel detection, including approaches based on the topological thinning [51], the particle filtering [52], the graph-based analysis [53], vessel tracking, active contours [54], minimal cost path computations [55] and fuzzy connectedness [56].

The basic aim of medical image processing is the segmentation process that assigns labels to the voxels. As far as the CTA images analysis is concerned, existing studies are dedicated to provide the coronary vessel geometry, whereas other studies aimed to detect and characterize the atherosclerotic plaques. More specifically, most of the studies aim to detect the coronary lumen and to evaluate the 3D coronary

artery geometry derived by the CTA slices, while other studies aim to detect the PB region and characterize its composition, by distinguishing the atherosclerotic plaque into calcified plaques (CP) and non-calcified (NCP) plaque.

Wang *et al*. [56] proposed a methodology for the segmentation and quantification of stenosed coronary arteries based on CTA images, using a localized region-based level sets framework. Primarily, as a preprocessing step, the entire heart is detected employing a mathematical morphology algorithm in order to accurate extract the coronary arteries. The coronary artery detection was achieved by implementing a Hessian-matrix based filter to enhance the visualization of vessel structures and subsequently, the initial coronary artery surface is estimated by a localized thresholding algorithm. Finally, the accurate lumen border detection was achieved by implementing a region-based level set algorithm.

On the other hand, Bouraoui *et al.* [57] proposed an automated algorithm for the segmentation of coronary arteries based on mathematical morphology techniques and discrete geometric tools. More specifically, the main steps of the algorithm is the CTA image preprocessing, the heart region localization, employing grey-level hit-or-miss opening and the implementation of a region-growing algorithm. The accuracy of this proposed methodology was finally validated using 60 CTA images by an expert cardiologist, who confirmed that the 90% of the images voxels were correctly segmented, whereas the remaining 10% of the CTA images voxels concerned two pathological data with low image quality, presenting important calcification, in the presence of stent.

Shahzad *et al.* [58] in a different attempt, presented a methodology for the automatic detection and quantification of coronary artery stenoses in CTA images. In this approach, the coronary centerlines were firstly extracted a two-point minimum cost path approach for the initialization of lumen segmentation, which was then achieved using graph cuts. Subsequently, a robust kernel regression model was applied to quantify the expected healthy coronary artery lumen diameter and then the lumen stenoses were computed by estimating the difference between estimated and expected diameter profiles. The validation dataset consisted of 30 testing datasets and the methodology was compared both by quantitative CA and manual assessment based on 3 experts' observers annotation. In addition to this, as far as the lumen segmentation is concerned, the dice coefficient (DICE) value for the health and diseased vessels was 0.68 and 0.65, respectively.

In a different direction, Chen *et al.* in 2014 [59] proposed a methodology for the segmentation of coronary arteries in a semi-automatic manner. The first step of their methodology is to isolate the heart

and coronary arteries region and to search for the probable location of coronary arteries by 3D region growing algorithm. Subsequently, the complete and accurate detection of coronary arteries is achieved applying discrete wavelet transformation and $\lambda-$mean operation. The proposed methodology was evaluated using 20 datasets and compared using the annotations edited by an expert radiologist using a commercial software workstation. The maximum calculated distance error was 2.2 mm, whereas the overlap ratio was 86.8%. The proposed methodology was directly compared to Yang *et al.*[60] study, where the maximum distance error and the overlap ratio were 2.3 mm and 68 %, respectively.

Kitamura *et al.* [61] presented a novel segmentation approach for the lumen and atherosclerotic plaque segmentation, considering the vessel as a tubular structure, whereas CP are more likely globular. Their segmentation idea was based on multi-label graph cuts utilizing higher-order potentials to impose shape priors. A standardized evaluation framework, presented in the medical image computing and computer assisted intervention (MICCAI) segmentation challenge 2012, was utilized and a sensitivity of 51.1% and a PPV of 33.3% compared to CTA reference, whereas the DICE as far as the lumen segmentation is concerned was 0.74 for the diseased coronary vessels and 0.73 for healthy ones.

In the same rationale Chen *et al.* [62] in 2015 proposed a novel segmentation for the detection of coronary artery pathologies. The proposed methodology consists of two basic steps: the implementation of a 3D region growing algorithm to initially segment the coronary arteries and subsequently, a vessel-texture discrimination algorithm was applied to detect the location of the vessel. The presented methodology was evaluated using a commercial software and an existing study based on a level set algorithm [60]. The efficiency of the proposed method was analyzed, computing the overlapping metric and the Hausdorff distance (HD). The average overlapping metric of the proposed approach was 0.96 and the mean HD was 1.08.

Yuanzhi *et al*. [63] presented an active contour framework with accurate shape and size constraints on the vessel cross-sectional planes to accurately segment the vessel. The proposed approach includes a multiscale vessel axis tracing in a 3D CTA images, a vessel boundary delineation on the cross-sectional planes, based on the extracted axis and finally the deformation of the vessel boundary surface, which is voxelized to produce the vessel segmentation. The novel aspect of the proposed methodology is its ability to achieve an accurate point detection in problematic CTA image region, such as vascular pathological regions and to avoid the disconnected and incomplete segmentation of the vessels. The

presented study is evaluated using two publicly available databases, the Rotterdam Coronary Artery evaluation framework datasets and the mean true positive ratio is 96.37±1.05.

On the other hand, in a recently published study, proposed by Athanasiou *et al.* [64], a semi-automated methodology was presented for the identification of the coronary lumen, the outer wall and the CP, based on a Gaussian mixture model. More specifically, the CTA image voxels were classified into lumen, outer wall and CP, utilizing a 3-component Gaussian Mixture Model (GMM). The setting parameters of GMM was adapted to each CTA, based on a set of region, manually annotated by an expert. The validation procedure included the comparison of the volume, area and length of the lumen and outer wall, as well as the CP volume against IVUS findings and promising results were extracted.

Zhu *et al.* [65] in 2022, in another attempt, proposed a U-shaped network based on spatio-temporal feature fusion structure to segment coronary arteries from 2D CTA slices, by combining features of multiple levels and different receptive fields separately to get more precise boundaries. The proposed methodology achieves the mean DICE of 0.87. In a similar approach, Mirunalini *et al.* [66] proposed a deep learning based model to segment and reconstruct the coronary artery, using the U-Net model for the segmentation and the Maximum Intensity Projection (MIP) reconstruction algorithm to analyze the presence of stenosis.

The studies related to coronary lumen detection are presented in Table 2.1.

Although extensive methodologies have been proposed for the automated segmentation of the coronary lumen based on 3D CTA images, several methodologies have been proposed for the atherosclerotic plaque region detection and the characterization of its composition. It has been indicated that the characterization and quantification of atherosclerotic plaque is of high significance for the patient risk stratification and contributes to the improvement of patient's management. Atherosclerotic plaque composition is classified into CP, NCP and mixed plaque. As far as the CP is concerned, its identification is well established even at the early stages of CTA usage in terms of calcium score and is considered as a strong predictor of coronary events [67, 68], while the identification of NCP remains still a challenging problem.

Detection of NCP based on non-invasive imaging modalities was performed in a study proposed by Schroeder *et al.* [69], who attempted to characterize the atherosclerotic plaque, using CTA modality and analyzing thirty-four plaques. In this approach, the atherosclerotic plaque was classified into soft,

intermediate and CP using an HU-based analysis. More specifically, the mean density values of soft, intermediate plaques and CP was 14 ± 26 HU, 91 ± 21 HU and CP 419 ± 194 HU, respectively. This approach was compared with the IVUS findings and a reliable performance of CTA, in terms of characterization of coronary lesion configuration, was indicated. In the same manner, Hur *et al*. [70] attempted to determine the diagnostic accuracy of CTA imaging modality to quantify and characterize the atherosclerotic plaque region. The atherosclerotic plaque composition was analyzed based on HU criteria, and more specifically the mean CT values for soft, fibrous, mixed and CP were 54±13 HU, 82±17 HU, 162±57 HU, 392±155 HU, respectively. The proposed methodology was evaluated using thirty-nine patients who underwent both CTA and IVUS and the utilized stenotic sites were totally sixty one. The correlation coefficients for the measurements of the lumen, vessel, plaque area, and percentage of luminal obstruction were 0.712, 0.654, 0.753 and 0.799, respectively and thus, it was observed that the luminal obstruction correlated well with this derived by the IVUS modality. However, a significant overlap was observed among the soft and the fibrous plaques and the classification of NCP as vulnerable or stable plaques, based on HU threshold analysis, remains still limited.

In 2011, Papadopoulou *et al*. [71] evaluated the ability of CTA modality to detect and quantify the atherosclerotic plaques, using as reference standard the IVUS modality. The CTA and IVUS co-registration step was developed, based on a dedicated software, MeVisLab software, whereas the co-registered region of interest (ROI) was considered for plaque analysis. Coronary inner and outer wall detection was performed manually implementing a stepwise approach. The atherosclerotic plaques were identified manually and classified into two categories, the NCP and the mixed/classified plaques. The evaluation procedure showed promising results, and the derived correlation coefficient for any type of atherosclerotic plaque volume was 0.91. However, the basic limitation of the presented methodology was the selection of the utilized dataset, where all of the patients presented high prevalence of the disease and the diagnostic accuracy might be lower in populations with lower prevalence.

Leber *et al.* [72], in another study, investigated the accuracy of CTA to detect different types of coronary plaques (CP, NCP and mixed plaques), in comparison with IVUS modality. The lumen and the outer wall of coronary arteries was empirically determined based on a display setting that was empirically determined, so that the CTA image equaled the IVUS image in size and pattern and permitted the exact separation between the inner wall, the outer wall, the atherosclerotic plaque and the surrounding tissue. In this study, the optimal setting to detect plaque and outer vessel boundaries ranged

from 395 to 809 HU, whereas the intensity of the lumen ranges from 165 to 339 HU. In the validation procedure, 38 vessels were totally utilized and only 23 of 145 coronary lesions were misclassified, whereas the correlation coefficient for the plaque volume was 0.69. The results of this study demonstrated the CTA ability to detect and quantify atherosclerotic plaque, while a further improvement of CTA is required for the assessment of very small and distal coronary plaques. In the same rationale, Voros *et al*. [73] attempted to investigate the accuracy of CTA to detect the coronary inner and outer wall and evaluated their findings against IVUS with radiofrequency backscatter analysis. The CTA image analysis, as far as the inner and outer wall borders detection is concerned, was applied by an experienced user [74, 75], whereas the atherosclerotic plaque detection was applied based on an HU based analysis. More specifically, the CTA voxels were classified as CP with attenuation values higher than 150 HU, as high density NCP with 30-149 HU and as low density NCP with −100-30 HU. The accuracy of their study was investigated, by extracting geometrical and compositional features, such as the minimal lumen diameter (MLD), the degree of stenosis (DS), the minimal lumen area (MLA), the percentage of area stenosis (%AS) and the volume and percentage of 3 plaque components: CP, high density NCP, and low density NCP. Correlation coefficients of CTA derived features against the Virtual Histology-Intravascular Ultrasound (VH-IVUS) derived features indicate an important correlation. More specifically, the correlation coefficient for the MLD, the %DS, the MLA, the %AS were 0.59, 0.43, 0.65, 0.44, respectively, whereas the correlation coefficient for the NCP volume and CP volume were 0.84, 0.65, respectively. This work was considered as the first prospective study to implement a 3D, quantitative analysis of co-registered CTA and VH-IVUS datasets, providing promising results for the CTA accuracy. On the other hand, since manual measurements of coronary plaques are time-consuming, computational attempts have been undertaken to automate the extraction of coronary plaques using the CTA imaging modality. More specifically, Dey *et al.* [76] attempted to evaluate the accuracy of developed automated algorithm [77] for the automated detection and quantification of NCP and CP. The proposed methodology combines HU density scan specific thresholds for the inner wall and the atherosclerotic plaques, and more specifically the mean HU value for the NCP was 242.6±25.9 HU, whereas for the CP 535±67.4 HU. The proposed developed algorithm was validated against IVUS manual image analysis and totally 22 NCP were utilized. A strong correlation of NCP volumes derived by the proposed algorithm against IVUS image analysis was indicated and the correlation coefficient was 0.92. In a similar manner, Brodoefel *et al.* [78] evaluated the CTA accuracy to detect and characterize the atherosclerotic plaque region in comparison with VH-IVUS. The plaque

characterization analysis was performed using a dedicated software, the SUREPlaque™ software, that automatically provides the vessel boundaries, while the 3 different components of atherosclerotic plaques, the fatty, fibrous and CP were extracted based on HU density values. More specifically, lipid plaque components range from -150 to 60 HU, lipid plaque from 61 to 149 HU and CP from 150 to 1300 HU. The methodology was validated in 14 patients, using 22 coronary lesions and good correlation was achieved for the entire plaque volume (r^2=0.76) and NCP volume (r^2= 0.84), whereas the correlation for the plaque composition proved very poor and insignificant.

On the other hand, Boogers *et al.* [79] aimed to prove the feasibility of CTA imaging modality to automatically quantify the coronary plaque, using a co-registration algorithm of CTA and IVUS data. The CTA images analysis was performed using a dedicated and extended version of the QAngio CT software (QAngio CT 1.1, Medis medical imaging systems, Leiden, the Netherlands), which provides the lumen and outer vessel wall borders. The proposed methodology was evaluated in comparison with IVUS images analysis using patients with 146 coronary lesions and the correlation coefficients for the MLA, the Lumen AS, the PB, the Remodeling index were 0.75, 0.79, 0.64 and 0.56, respectively. Based on Boogers *et al.* study, another approach was introduced by Graaf *et al*. [80], who investigated the automatic quantitative assessment of coronary stenosis and plaque constitution on CTA, compared to VH-IVUS. The lumen and vessel wall detection was applied based on the previously described methodology [79], while the plaque classification was based on two different approaches. More specifically, in the first approach different fixed HU cut-off values were used for classifying the plaque tissue, −30 to 75 HU, for NC, 76–130 HU for fibro-fatty, 131–350 HU for fibrotic, and higher than 351 for dense calcium. The selection of these HU value ranges was based on Brodoefel *et al.* [81] study and optimized using different datasets. As far as the second approach for threshold definition is concerned, its selection is user independent and is based on the principle that the plaque attenuation values is directly associated with the lumen intensity value. The study, proposed by Graaf *et al.* [80], was validated using 109 vessels, of which 69 revealed atherosclerosis. As far as the fixed thresholds is concerned, good correlations were achieved for volumes of fibrotic tissue (r = 0.695), fibro-fatty tissue (r = 0.714), NC (r = 0.523) and dense calcium (r = 0.736), whereas for the dynamic threshold tissue classifier, the correlation plaque volume for NC was 0.479 and the volumes of dense calcium, fibrotic and fibro-fatty volumes were all significantly overestimated by CTA image analysis.

However, the previously described studies are implemented either using manual estimations [70-72, 82], or utilizing available software [78, 80, 83]. A different approach was proposed by Jawaid *et al*. [84], who implemented a support vector machine (SVM) classifier, after computing the radial profiles by averaging the CTA image intensity in rings around the vessel centerline to identify the abnormal coronary segments. A derivative-based method was also applied to localize the position and the length of the NCP.

The above described methodologies for atherosclerotic plaque detection are summarized in Table 2.2.

Table 2.1 Coronary Lumen Detection Methodologies based on CTA modality.

Studies	**Methodology**	**Dataset**	**Results**
Wang *et al*. [56]	localized region-based level sets	clinical data provided by St Thomas'hospital, London	Accurate lumen segmentation
Bouraoui *et al.* [57]	mathematical morphology techniques, discrete geometrical tools	60 CTA images	90% correct segmentation
Shahzad *et al.*[58]	robust kernel regression	30 testing datasets compared to quantitative CA	DICE (healthy vessels): 68% DICE (diseased vessels): 65%
Chen *et al.* [59]	3D region growing Discret wavelet transformation	manually edited by a radiologist using commercial software	Overlap ratio 86.8 %
Kitamura *et al.* [61]	multi-label graph cuts	MICCAI 2012, Quantitative CT	DICE (healthy vessels): 73% DICE (diseased vessels): 74%
Chen *et al.* [62]	3D region growing vessel-texture discrimination algorithm	commercial software from GE Healthcare and the level-set method proposed by Yang et al.	Overlapping metric 0.96 Mean HD: 1.08
Yuanzhi *et al*. [63]	Active contour models	Rotterdam Coronary Artery evaluation framework datasets	True positive ratio: 96.37±1.05
Athanasiou *et al.* [64]	Gaussian Mixture Model	IVUS	Lumen volume correlation: 0.76
Zhu *et al.* [65]	U-shaped network	30 patients	DICE of 0.87
Mirunalini *et al*. [66]	a deep learning based model	50 patients	SSIM values of 0.84829, 0.85739, 0.83857, 0.8459, .8531 for different rotation angles

Table 2.2 Plaque Characterization Methodologies based on CTA modality.

Study	**Plaque**			**Methodology**	**Comparison**	**Results**
	CP	**NCP**	**intermediate**			
Schroeder *et al.* [69]	✓	✓	✓	HU-based analysis	IVUS	Nonparametric Kruskal-Wallis test revealed a significant difference of plaque density among the three groups
Hur *et al*. [70]	✓	✓	Fibrous, mixed	HU-based analysis, soft:54±13 HU, fibrous: 82±17 HU, mixed:162±57 HU, CP:392±155 HU	IVUS	Lumen Area: r=0.712, Plaque Area: r=0.753 Vessel Area: r=0.654, DS: r=0.799
Papadopoulou *et al*. [71]	✓	Mixed, CP	-	Inner, outer wall & plaques manually edited	IVUS	
Leber *et al.* [72]	✓	✓	mixed	manually annotated	IVUS	Missclassified 23/145 plaques, plaque volume: r=0.69
Voros *et al*. [73]	✓	✓	High density NCP	Inner & outer wall manually edited, high density NCP: 30-149 HU, low density NCP:-100-30 HU, CP: >150 HU	IVUS	vessel plaque volumes: r = 0.51
Dey *et al.* [76]	✓	✓	-	scan-specific attenuation threshold	IVUS	Mean plaque volume correlation (r = 0.92)
Brodoefel *et al.* [78]	✓	✓	fibrous	SUPERplaque software inner & outer wall, NCP: -150-60 HU, CP: 150-1300 HU, fibrous:61-149	VH-IVUS	NCP (%)28.2 ± 6.1 (CT), 29.9 ± 5.2 (IVUS) Fibrous (%)53.2 ± 8.7 (CT), 55.3 ± 12.2 (IVUS) CP (%) 18.7 ± 12.5 (CT), 14.4 ± 9.1 (IVUS)

2.2.2 3D reconstruction of carotid arteries based on CTA

Regarding the imaging of carotid arteries, CTA is considered as a promising non-invasive imaging modality utilized for the assessment of the carotid stenosis, supplementary to the use of US imaging. The main advantage of the CTA in carotid imaging is the ability to image the artery bifurcations, from the aortic arch to the brain parenchyma [85]. On the other hand, the high radiation dosage, for the need of intravenous contrast and the high intensity artifacts, consists of the acknowledged drawbacks of CTA.

Different studies have been conducted for the lumen carotid detection and the identification and characterization of carotid atherosclerotic plaques. More specifically, Sanderse *et al.* [86] presented an automatic initialization algorithm to detect the carotid arteries providing a fully automated approach for vessel centerline detection and segmentation. The carotid arteries are detected in axial slices of the volume of interest by applying a circular Hough transform. A hierarchical clustering approach was used to select carotid related signals in the Hough space, whereas a feedback architecture was introduced to successfully detect the range of the vessel diameters. The presented methodology was trained using 20 patient datasets and tested using 31 patient datasets. An overall detection accuracy of 0.88 was achieved. The main limitation of the proposed methodology was the fact that 21% of the overall cases cannot be successfully detected. In another attempt, Van Velsen *et al.* [87] presented a study based on the detection of lumen-like structures by analyzing the density, edge and ridge knowledge. Upon initialization with a start and endpoint, the lumen path is automatically defined in 3D image data instead of 2D MIP images, and the extracted lumen path is subsequently used as initialization for automatic lumen segmentation with a level set method. The shape of the vessel lumen is defined by evolving a surface toward the lumen boundaries using a speed term, which is defined based on image information and a smoothness constraint, while a predefined lumen path was used to assure a good initialization even if a stenosis or obstruction is present. In this work, the required image information for the speed term is based on image intensity and edge information. Manniesing and Niessen [88], in another study, presented an automatic method that segments the internal carotid arteries (ICAs) across the difficult locations of the skull base in CTA. The proposed methodology consists of the following three steps: (i) the entropy profiling to select the lower region that includes the skull base, (ii) the selection of a rough vessel segmentation which is used as input for a Hough transform to detect the 2D circular shapes in this plane and finally, (iii) the center points which are used to initialize a level set which evolves with a prior shape constraint on its topology. This proposed methodology was validated using 20 carotid arteries of 10 patient datasets and finally 18 out of 20 were successfully detected automatically.

On the other hand, Cuisenaire [89] developed a fully automated vessel extraction and segmentation tool, specifically used for head and neck region. This work focuses particularly on the initial centerline extraction technique. It uses a locally adaptive front propagation algorithm that attempts to find the optimal path connecting the ends of the vessel, typically from the lowest image of the scan to the circle of willis in the brain. It uses a patient adapted anatomical model of the different vessels both to initialize and constrain this fast marching, eliminating the need for manual selection of seed points. The initial centerline of each vessel was extracted and the vessels were segmented using the initialized 3D active surfaces using the spline-snake to identify true centerline from the segmented mask edges.

Another semi-automated technique for lumen segmentation of the carotid bifurcation in CTA using the level set approach was proposed by Manniesing *et al.* [90]. Firstly, the central vessel axis is obtained using path tracking between three user-defined seed points, using Dijkstra algorithm and voxel based backtracking method. Secondly, starting from this path, the segmentation is automatically obtained using a level set coupled with fast marching method. The cost and speed functions for path tracking and segmentation make use of intensity and homogeneity slice-based image features. The method was validated on a large dataset of 234 carotid bifurcations of 129 ischemic stroke patients with atherosclerotic disease. The results were compared to manually obtained lumen segmentations. Parameter optimization was carried out on a subset of 30 representative carotid bifurcations. With the optimized parameter settings, the method successfully tracked the central vessel paths in 201 of the remaining 204 bifurcations (99%) which were not part of the training set. The comparison with manually drawn segmentations shows that the average overlap between the method and observers is similar (for the inter-observer set the results were 92% vs. 87% and for the intra-observer set 94% vs. 94%). Similarly, Vukadinovic *et al.* [91] presented a level-set based and GentleBoost classification based approach for the segmentation of the carotid artery outer vessel wall and the atherosclerotic plaque components. This methodology consists of three steps: (i) the lumen segmentation based on the level set methodology [92], (ii) implementation of GentleBoost classifier to detect calcium pixels, (iii) implementation of GentleBoost classifier for voxels inside and outside the carotid artery lumen. The proposed methodology was trained and optimized using 20 datasets, whereas it was evaluated on 80 datasets. The reference standard of the implemented validation procedure was expert's manual annotation and the similarity index for the outer vessel wall and the atherosclerotic plaque was 0.92±0.03 and 0.81±0.06, respectively. In addition to this, the estimated correlation coefficients for the

vessel volume, the plaque volume, the fibrous volume, the lipid volume and the calcium volume were 0.96, 0.81, 0.89, 0.46 and 0.95, respectively.

Freiman *et al.* [93] proposed a practical methodology for the modeling of the total aortic arch and of the entire carotid vasculature, starting with an automated ROI identification which is followed by the implementation of an automatic watershed-based segmentation of the aorta. Aorta segmentation primarily relies in the prior knowledge of its anatomy, structure, location and its brightness. Subsequently, a Gaussian distribution was estimated for the carotid vasculature intensity, and finally carotid artery segmentation was applied, using graph min-cut segmentation approach, followed by a semi-automated graph-based refinement. The proposed methodology was validated using four different datasets and totally 66 patients, achieving HD and DICE of 1.55 and 0.91, respectively. Tang ***et al.*** [94] proposed a semi-automatic carotid lumen segmentation method and CAS (carotid artery stenosis) quantification using a level-set based approach, which was initiated with a centerline obtained from user defined seed points. This study consists of three steps: (i) the implementation of an iterative minimum cost path approach to extract the carotid vessel centerline, (ii) the implementation of a level set approach to extract initial carotid lumen boundary and (iii) the removal of the side branches in the segmented lumen. The obtained DICE of this proposed methodology was 90.2%, whereas the mean absolute surface distance was 0.34 mm. As far as the CAS quantification is concerned, the average error was 15.7% and 19.2% for cross-sectional diameter-based stenosis and cross-sectional area-based stenosis, respectively.

Dos Santos *et al*. [95] proposed a semi-automatic method in which they segmented vessel walls and surrounding tissues with an adaptive segmentation and region growing algorithms. The slice interval is determined for the segmentation and analysis process. The initial seed point in the lumen is set manually in the first slice to specify the artery side to be segmented and the exact position of the vessel. A typical diameter of 6 mm is used to create a ROI around the lumen and to generate a histogram. Gaussian fitting is executed, and 90% of the area under the curve (AUC) is chosen as the initial blood attenuation HU interval. The next step is the segmentation of the lumen using the initial seed and the attenuation interval, with a region-growing algorithm. The use of an adaptive attenuation interval in each slice prevents the common problem of dilution during region growing. By fitting the attenuation value histogram onto a Gaussian curve, it is possible to decrease the influence of the less frequent attenuation values representing surrounding tissues. After the initial slice is obtained, segmentation proceeds

automatically along the vessel and throughout the entire CT image stack. If the stenosis is so severe that no pixel identifies the lumen, the same initial seed position is used for the next slice. In this study, they concluded that there was no significant difference between the maximum percent stenosis value obtained using the semi-automated tool and those obtained using manual measurements.

Hemmati *et al.* [96] proposed a methodology for the carotid lumen segmentation based basically on a 3D level set model. Firstly, the grey levels of the CTA images were uniformly enhanced by the mean shift smoothing and then the vessel centerlines of the carotid arteries were extracted using a 3D Hessian based fast marching shortest path algorithm, whereas the final carotid artery segmentation was achieved utilizing a 3D level set function. This presented methodology was evaluated using 14 different CTA volumes and the achieved DICE was 0.85.

Bozkurt *et al*. [97] provided a novel method for carotid artery lumen segmentation on CTA images by using automatic vessel segmentation with inverse approach, in which vessel segmentation is performed after bone region is segmented and eliminated. The region growing and random walk segmentation methods are utilized in the elimination of bone region and the vessel segmentation. The seed points in the mentioned methods are not manually determined by any starting point. In automatic segmentation, seeds are automatically selected from the experimentally determined intervals according to the local histogram. The stages of preprocessing and post-processing are utilized for better segmentation. The tracking of vessel centers based on continuity is employed for 3D reconstruction and 3D imaging of the vessels. Experiments were conducted with different data sets including various CTA images by using the mentioned methods. As a result, DICE above 92% was achieved together 99% accuracy.

.

Table 2.3 Methodologies for 3D reconstruction of carotid arteries using CTA imaging modality.

Study	**Methodology**	**Detection**	**Results**
Sanderse *et al.* [86]	circular Hough transform	Vessel detection and centerline	20 datasets (training set), 31 datasets (test set), 88% accuracy
Manniesing & Niessen [88]	Hough transform and a level set which evolves with a prior shape constraint	Vessel detection	20 carotid arteries (10 patients), 90% accuracy (18/20 correctly identified)
Cuisenaire *et al*. [89]	locally adaptive front propagation algorithm (centerline extraction)	Vessel detection and centerline	
Vukadinovic *et al.* [91]	Level set methodology, GentleBoost classification framework ($lumen_{threshold}$ = 320HU, fibrous $tissue_{threshold}$=130 HU, $lipid_{threshold}$=60 HU)	Lumen, outer wall, plaque components	DICE: 0.92±0.03 (outer wall), 0.81±0.06 (plaque) in 80 datasets, vessel volume (r^2=0.96), plaque volume (r^2=0.81), fibrous volume (r^2=0.89), lipid volume (r^2=0.46), calcium volume (r^2=0.95)
Freiman *et al*. [93]	Automatic watershed-based segmentation of the aorta, graph-cut optimization framework for carotid segmentation	Carotid lumen	71 multicenter clinical CTA datasets, DICE of 84.5% (SD = 3.3%) and MSSD 0.48 mm (SD = 0.12 mm.)
Hemmati *et al.* [96]	3D Hessian based fast marching shortest path (centerline extraction), 3D level set (lumen segmentation)	Carotid lumen	14 CTA datasets, 0.85 similarity index accuracy
Guha *et al.* [98]	geodesic path propagation algorithm based on fuzzy distance transform.	Carotid vessel	eight patients' CTA image, the accuracy of all the segmented data was verified by overlaying the phantoms with the original images
dos Santos et al. [95]	adaptive segmentation algorithm and region growing	CAS	no significant difference between the maximum percent stenosis value obtained using the semi-automated tool and those obtained using manual measurements
Bozkurt *et al*. [97]	region growing and random walk segmentation methods	Carotid vessel	DICE 92%
Saba *et al.* [99]	HU attenuation analysis	Lipids, fibrous tissue, calcium plaque	Relation between the monochromatic kiloelectron volt values and the HU range of the plaque

2.3 Prediction of coronary artery disease

2.3.1 Non-imaging CAD prediction

In this section, we are going to provide a cohesive state of the art on ML methods for obstructive CAD prediction, focusing on studies employing non-imaging data and imaging based data. We primarily drew on our recently published literature review, where we analysed and contrasted the relevant literature with respect to the acquired dataset, the examined feature space, the employed predictive modelling schemes and their discriminative or predictive capacity [100].

As it is shown in Table 2.4, the detection of obstructive CAD, defined as a ≥50% diameter stenosis in at least one main coronary artery vessel assessed by CA or CTA, is formulated as a binary classification problem on the basis of a confined set of demographic, clinical, and biohumoral characteristics. We should clarify that the negative class is defined: either as (i) the complement of the positive class, or as (ii) individuals with no presence of CAD. On this basis, parametric (e.g. neural networks) or non-parametric (e.g. kernel methods) classification algorithms have been combined with feature evaluation techniques, aiming at specifying an accurate decision hyperplane and, in parallel, identifying the most informative features with respect to CAD severity.

Kurt *et al.*[101] demonstrated that, irrespectively of the classification algorithm, typical heart disease risk factors [i.e. age, sex, body mass index (BMI), smoking status, DM, hypertension, hypercholesterolemia, and family history of CAD] yield a high number of false positive predictions. In particular, a LR based model, a classification and regression tree (CART), a feed-forward neural network (FFNN) and a radial basis function network (RBF) have all resulted in a relatively high sensitivity (ranging from 89.5-92.3), however, their specificity was <50%. Correlation-based feature selection (CFS), a filter-based feature selection approach, identified treadmill stress testing-related characteristics (i.e. Duke Treadmill Score, and post exercise recovery period with persistent electrocardiographic ST-segment changes) amongst the most informative features with respect to CAD severity diagnosis [102]; a FFNN fed additionally with information on smoking, DM, and high density lipoprotein (HDL) attained 88.4% accuracy. Alizadehsani *et al.* by exploring the predictive potential of kernel-based methods, and particularly of SVM, both in feature selection and classification tasks [103]: (i) evaluated the discriminative capability of 54 features, concerning demographic, clinical, electrocardiographic, echocardiographic, and laboratory data (via the SVM weight vector), and (ii) demonstrated that kernel-based methods (i.e. SVM and bagging SVM) outperform both FFNN and naïve Bayes (NB), exhibiting 93.4% and 92.7% accuracy, and high sensitivity and specificity rates, as well. In a subsequent study, Alizadehsani *et al*. assessed the diagnostic accuracy of the same feature set with respect to the level of stenosis at LAD artery, LCX

artery, and right coronary artery (RCA)] separately, formulating a 2-class problem where a ≥50% diameter stenosis characterizes a stenotic artery [104]. The DS of LAD, LCX and RCA was diagnosed with 86.14%, 83.17% and 83.5% accuracy, respectively by (i) adopting a common feature set for all arteries, encompassing the 24 top ranked features according to a combined info-gain index, and (ii) employing a multiple kernel learning algorithm. Nahar *et al.* using the UCI Cleveland heart disease dataset, showed that knowledge-based feature selection may increase considerably the sensitivity of predictions [105]. Ensemble learning of the specified dataset, using a rotation forest with FFNN as the base classifier, reached 91.2% accuracy, 95.6% sensitivity, and 86.7% specificity [106].

Unlike black-box ML techniques, fuzzy rule-based classifiers provide interpretable decisions. Tsipouras *et al*. proposed an optimized fuzzy model for the diagnosis of CAD severity considering traditional CV risk factors as well as two non-invasive features, namely carotid–femoral and augmentation index. Their four-stage methodology encompassed: (i) induction of a decision tree, (ii) extraction of the rule base from the decision tree, in disjunctive normal form and formulation of a crisp model, (iii) transformation of the crisp set of rules into a fuzzy model, and (iv) optimization of the parameters of the fuzzy model [107]. The optimized fuzzy model resulted in 73.4% accuracy, 80.0% sensitivity and 65.2% specificity, presenting comparable performance with a FFNN (73.9% accuracy) and significantly better results than an adaptive neuro-fuzzy inference system (56.8% accuracy), both applied to the same task.

ML naturally arises as favourable solutions to CAD severity diagnosis when omics big data need to be exploited (

Table 2.5). The Corus CAD algorithm, which constitutes an exemplar case, was developed via a combination of microarray and RT-PCR gene expression data analysis, collected from age and sex-matched patients with symptoms suggestive of CAD [108]. The Corus CAD method is decomposed into the following phases: (i) feature selection based on an unsupervised cluster analysis, which yielded to the identification of meta-genes, (ii) a ridge linear regression model of the age, sex, and the selected genes. In [108], a 0.75 (0.70 – 0.80) AUC is reported with respect to the identification of the obstructive CAD patients. Dogan *et al*. exploited a comprehensive dataset of genome-wide DNA methylation data, SNP data and phenotype data [i.e. age, gender, systolic blood pressure (SBP), HDL, total cholesterol (TC) level, haemoglobin (Hb) A1C, smoking status, and the use of statins] derived from the Framingham Heart Study [109]. Random forest (RF) classification was employed for the identification of CAD vs control individuals, while addressing the class-imbalance issue

through a heuristic repeated under sampling approach. Sequential feature selection, by the ROC AUC of each feature, retained four CpG sites (cg26910465, cg11355601, cg16410464 and cg12091641), two SNPs (rs6418712 and rs10275666), age and gender. This classifier was capable of classifying symptomatic CAD with an accuracy, sensitivity and specificity of 0.78, 0.75, and 0.80, respectively. In contrast, a model using only conventional CAD risk factors as predictors had an accuracy, sensitivity and specificity of 0.65, 0.42, and 0.89, respectively.

Table 2.4 Non-imaging CAD Severity Diagnosis Methods based on Machine Learning [100].

Study	**Dataset-** **Feature set**	**Outcome**	**Methods**	**Performance**	
Kurt *et al*., **[101]**	n=1245 subjects, typical heart disease risk factors	Class 1 n=865:≥50% stenosis in at least one coronary artery vessel Class 0 ($n=380$): Otherwise	Classification: LR, CART, FFNN, RBF Evaluation: Training, Test, Validation sets (60%-20%-20%)	LR Acc.:79.5 Sens.:92.3 Spec.: 45.6	CART Acc.:79.9 Sens.:92.3 Spec.: 47.1
				FFNN Acc.:79.1 Sens.: 91.7 Spec.: 45.6	RBF Acc.:76.7 Sens.: 89.5 Spec.: 42.6
Verma *et al*., **[102]**	n=335 subjects who were suspected for CAD, Smoking, DM, clinical, electrocardiographic data	Class 1 (48.9%): CAD Class 0 (51.1%): No CAD	Feature Selection: CFS Clustering: k-means Classification: FFNN, LR, Fuzzy unordered rule induction algorithm (FURIA), Decision tree (C4.5). Evaluation: 10-fold cross-validation	FFNN Acc.: 88.40	FURIA Acc.: 82.80
Alizadehsani *et al*. **[103]**	n=303 subjects, demographic, clinical, electrocardiographic, echocardiographic, and laboratory data	Class 1 (n=865) ≥50% stenosis in at least one coronary artery vessel in CA Class 0 (n=380) Otherwise	Feature selection: SVM weights Classification: SVM, NB, Bagging of SVMs, FFNN Evaluation: 10-fold cross-validation	SVM Acc.: 93.39 Sens.: 95.37 Spec.: 88.51	Bagging SVM Acc.: 92.74 Sens.: 95.37 Spec.: 86.21
				FFNN Acc.: 87.13 Sens.: 90.28 Spec.: 79.31	NB Acc.: 55.37 Sens.: 38.89 Spec. :96.55
Alizadehsani *et al*. **[104]**	n=303 subjects, demographic, clinical, electrocardiographic,	Problem I Class 1 – LAD stenotic: ≥50% Class 0 Otherwise Problem II	Feature selection Approach I Different feature set for each artery: SVM weights Approach II Common feature set for all arteries: Information Gain	Approach I	Approach II
				LAD Acc.:85.81 Sens.:92.66 Spec.:76.19	LAD Acc.:86.14 Sens.: 90.96 Spec.:79.37

Study	Dataset- Feature set	Outcome	Methods	Performance	
	echocardiographic, and laboratory data	Class 1 – LCX stenotic: ≥50% Class 0: Otherwise Problem III Class 1 – RCA stenotic: ≥50% Class 0: Otherwise	Classification: SVM with kernel fusion Evaluation**:** 10-fold cross-validation	LCX Acc.: 77.23 Sens.: 69.75 Spec.: 82.07	LCX Acc.: 83.17 Sens.: 90.9 Spec.:72.22
				RCA Acc.:81.85 Sens. 68.4 Spec. 89.95	RCA Acc.: 83.50 Sens.: 87.01 Spec.: 78.57
Nahar *et al.*, **[105]**	The UCI Heart Disease Dataset n=303	Class 1 n=165:≥50% diameter stenosis. Class 0 n=138: Otherwise	Feature Selection: CFS, Knowledge-based feature selection Classification: Sequential minimal optimization Evaluation: Approach I 10-fold cross-validation Approach II Training – Test sets	Approach II – Knowledge-based feature selection: Acc.: 77.95, Sens.: 0.811 Approach II – Knowledge-based feature selection combined with CFS: Acc.: 83.83, Sens.: 0.919	
Karabulut & Ibrikci, **[106]**	The UCI Heart Disease Dataset n=303, clinical, electrocardiographic, and laboratory data	Class 1 – Existence of CAD n=165 :≥50% diameter stenosis. Class 0 – Absence of CAD n=138 Otherwise	Classification: Rotation forest with FFNNs as the base classifier Evaluation**:** 10-fold cross-validation	Acc.: 91.2 Sens.: 95.6 Spec.: 86.7 AUC: 0.915	
Tsipouras *et al.*, **[107]**	$n = 199$ subjects demographic, clinical, echocardiographic, and laboratory data	Class 1 $n = 110$: ≥50% diameter stenosis in at least one coronary artery vessel Class 0 $n = 89$: without any narrowing visible in CA	Optimized fuzzy model Evaluation: 10-fold stratified cross-validation	Acc.: 73.4, Sens.: 80.0, Spec.: 65.2	

(Continued

Table 2.5 Pre-test CAD Severity Scores based on Genomics.

Study	Dataset	Outcome	Methods	Feature Set	Performance
CORUS CAD **[108, 110]**	COMPASS cohort *n*=610	Obstructive CAD: ≥50% stenosis in at least 1 vessel (13% obstructive CAD, ≥50% diameter stenosis)	Feature evaluation: Hierarchical clustering, identification of meta-genes Classification: Ridge regression Evaluation: Leave one patient out cross-validation	Age, sex, microarray and RT-PCR gene expression data	ROC AUC: 0.75 (0.70 – 0.80)
Dogan *et al* **[109]**	Training Set *n*=1545 Symptomatic CHD: *n=173* Normal: *n*=1372	Presence of CAD	Sequential feature selection, ROC AUC , Classification: RF, Evaluation: Training – Test sets Sets	Genome-wide DNA methylation and SNP data, Phenotype	Acc. 0.77 Sens. 0.75| Spec. 0.80

2.3.2 Imaging CAD prediction

Several studies in the literature have investigated the role of typical risk factors for the CAD prediction. Nevertheless, existing studies have proved that the atherosclerotic plaque is non-uniformly distributed in the same vessel sites and is identified in different patient's artery segments. This phenomenon cannot explain the existence of the established risk factors which are assumed to affect the entire vasculature of patients [111]. Thus, a new atherosclerosis risk factors category has been introduced, the artery geometry based risk factors, which have led to the concept that the initial vessel geometry and its local mechanical forces may play a crucial role in the initiation and progression of the atherosclerotic plaque. Identification of high-risk atheromatic plaques and the prediction of atheromatic plaque progression consists a significant task in biomedical research area [112-115].

Thus, in this section, studies related to the CAD prediction based on imaging features are going to be presented and are summarized in Table 2.6, below. More specifically, the advancement of computational modeling has provided insights into the atheromatic plaque progression and the potential for in vivo wall shear stress (WSS) calculations was first demonstrated using a fusion of IVUS and X-ray angiography to acquire the 3D artery geometry. Existing studies confirmed the direct association of low WSS artery regions with the atheromatic plaque formation [113, 114], while PREDICTION study [115] concluded that low baseline WSS regions cause a decrease of luminal area and increase of PB. Thus, further longitudinal studies area now justified to evaluate whether WSS calculations combined with artery anatomical imaging can more accurately predict clinical events, which may be possibly implemented using non-invasive coronary imaging [116]. CTA is a promising non-invasive coronary imaging, providing useful prognostic information of atherosclerosis progression, since it is able to accurately detect the inner and outer wall of coronary arteries and identify the CP and NCP, as well [117]. Based on this direction, Liu *et al* [118] investigated the effect of biomechanical factors for the progression of atherosclerosis, utilizing 365 coronary segments of 3 mm and concluded that the decrease in MLA was independently predicted by low baseline Von Mises stress (-0.73 ± 0.13 mm^2). Kolossváry *et al*. [119] aimed to detect atherosclerotic plaques with a napkin-ring sign (NRS), implementing a supervised approach based on predefined measurements, so-called radiomics features [120]. NRS is considered as a significant prognostic biomarker of MACE. Different texture features derived by a radiomic dataset are able to identify the atherosclerotic plaques with and without NRS [119]. Another existing in the literature study proposed a ML based model to discriminate between patients with acute or chronic Myocardial Infarction (MI) and control subjects, using as input the CTA based calcium score.

This study resulted in an AUC of 0.78 [121]. Except for the examination of automated analysis and diagnostics performance of CTA imaging modality, prognostic evaluation has also been applied in CTA imaging. More specifically, survival analysis was performed in different patient groups with a CV risk and it was indicated that ML based models exhibited a larger AUC (0.79) than the individual clinical and CTA metrics (Framingham risk score: 0.61, segment stenosis score: 0.64) [122]. Similarly to this approach, another predictive model was developed using CTA imaging features, based on the DS. This prognostic model provided a risk score for all-cause death and non-fatal MI and resulted in an AUC of 0.771 [123].

Table 2.6 Imaging based CAD prediction studies.

Study	**Dataset-** **Feature set**	**Outcome**	**Methods**	**Performance**
Stone *et al.* [115]-PREDICTION study	endothelial shear stress and arterial plaque characteristics	Progression of CAD	Statistical analysis	low baseline WSS regions cause a decrease of luminal area and increase of PB
Liu *et al* [118]	365 coronary segments of 3 mm	Progression of CAD	Statistical analysis	the decrease in MLA was independently predicted by low baseline Von Mises stress (-0.73 ± 0.13 mm^2)
Mannil *et al.* [121]	30 patients with chronic MI, 30 subjects without cardiac abnormality	Detection of MI	decision tree C4.5 (J48), KNN locally weighted learning, Random Forest	AUC: 0.78
Motwani *et al.* [122]	10 030 patients clinical & CTA parameters	predict 5-year all-cause mortality	gain ranking, model building with a boosted ensemble algorithm, and 10-fold stratified cross-validation	AUC: 0.79
van Rosendael *et al.* [123]	8844 patients, CTA risk scores	≥3 year follow-up for MI and death	Statistical analysis, XGBoost model	AUC: 0.771

2.4 Prediction of carotid artery disease

Carotid artery disease is considered as the main risk factor of cerebrovascular events and ischemic stroke. Prediction of carotid artery disease enables the early diagnosis of the disease, as it is defined by the presence of CAS or the presence of vulnerable carotid plaques. Different studies, summarized in Table 2.7, have been presented in the literature for the prognosis of carotid disease, taking advantage of statistical analysis techniques and ML models.

More specifically, Jamthikar *et al.* [124] proposed a ML based algorithm for the development of a CV-stroke risk stratification tool, consisting of an online and offline system and concluded that ML-based integrated model with the event-equivalent gold standard as artery DS, is powerful and offers low cost and high performance for the CV-stroke risk assessment. Firstly, a principal component analysis was implemented for the selection of the most significant input features and then, a RF classifier was implemented to predict the risk of each test subject. The algorithm was validated using a 10-fold cross validation scheme and the final system named "AtheroRisk-Integrated" was compared against "AtheroRisk-Conventional". Left and right CCAs of 202 Japanese patients were retrospectively examined to obtain 395 US scans. AtheroRisk-Integrated system [AUC =0.80, $P<0.0001$, 95% confidence interval (CI): 0.77 to 0.84] showed an improvement of ~18% against AtheroRisk-Conventional ML (AUC =0.68, $P<0.0001$, 95% CI: 0.64 to 0.72).

In a same attempt Verde *et al.* [125] investigated and compared the performance of conventional ML classifiers, capable of identifying the presence of carotid disease, exploring as input heart rate variability features and evaluating the classification scheme in terms of accuracy, precision, recall and F-measure. Totally, 126 patients (89 individuals suffering from carotid diseases and 37 healthy subjects) participated in this study and their pathological state was identified by analyzing the intima media thickness (IMT) value, evaluated with the B-mode US. In another study, Greco *et al.* [126] taking advantage of statistical based models, developed a model for the prediction of individuals of high risk for carotid DS. A multivariable regression analysis was implemented to identify the major risk factors for the pathological condition for the DS (>50%) and the C-statistic measure was computed for the evaluation of the proposed model. The independent input features were demographics, comorbidities, and lifestyle characteristics. Finally, the proposed predicting scoring system with a more efficient C-statistic value was created for the identification of those individuals with a higher probability of a pathological DS.

De Weerd *et al*. [127] proposed a carotid artery disease prediction model for the identification of individuals with a CAS >50% and >70% and for the detection of the most efficient predictors of CAS >50 % and 70%. Age, sex, history of vascular disease, SPB and diastolic blood pressure (DBP), TC/HDL ratio, DM and current smoking were identified as the most significant predictors of stenosis (>50% and >70%). In addition to this, the proposed model discriminated well between participants with and without stenosis, with an AUC corrected for over optimism of 0.82 for moderate stenosis and of 0.87 for severe stenosis.

Khanna *et al.* [128], in another attempt, proposed a nonlinear model for the 10-year prediction carotid image phenotypes, which were automatically measured utilizing AtheroEdge™ system. More specifically, nine conventional CV risk factors, in combination with five types of carotid image phenotypes were adapted in a non-linear mathematical model for 10 year prediction of carotid phenotypes. In this study, totally 206 Japanese patients participated, and totally 407 US scans left/right CCA were selected. This fused prediction model concluded that the age and the Hb presented the highest influence on the 10 year carotid phenotypes, whereas the AUC for the five types of carotid phenotypes were 0.96, 0.94, 0.90, 1.0, and 1.0 for the average IMT, the maximum IMT, the minimum IMT, the wall variability and the total plaque area, respectively.

Araki *et al.* [129] proposed an innovative risk stratification model for the classification of high and risk carotid plaques or the classification between symptomatic and asymptomatic plaques. The overall model was based on the grayscale morphology of the US carotid wall, while Hao-wen Li *et al.* [130] conducted an observational study to distinguish the high risk plaques (vulnerable plaques) and the low risk (stable) carotid artery plaques, which were classified using high-resolution MRI modality. Three hundred twenty-six patients with ischemic stroke (189 patients with vulnerable plaque and 81 patients with stable plaque) and 432 normal controls were included in this study. Finally, ADAMTS7 polymorphisms of both rs7173743 and rs3825807 were associated with carotid plaque vulnerability but not the prevalence of ischemic stroke.

Table 2.7: Summary of studies for the carotid artery disease stratification.

Study	**Methodology**	**Input**	**Output**	**Results**
Jamthikar *et al.* [124]	ML based algorithm (RF classifier)	47 risk factors 34 were image-based phenotypes, 13 risk factors (patients' demographics, blood biomarkers)	Low risk- High risk based on the artery DS	AUC =0.80, P<0.0001, 95% CI: 0.77 to 0.84
Verde *et al.* [125]	SVM, Bayesian Classification, Decision Tree, Multilayer Perceptron, LMT, Instance-based Learning algorithm	Heart rate variable features	Healthy -Pathological based on IMT evaluated with the B-mode US	Accuracy: 0.72
Greco *et al.* [126]	multivariable regression analysis	Sex, age, race, marital status, smokers, high blood pressure, comorbidities (High cholesterol, CAD, PAD, DM), Family history, BMI, Dietary habits, Exercise	Identifications of individuals whose CAS >50%	C statistic = 0.753
de Weerd *et al.* [127]	Multivariable LR	age, sex, DM, history of coronary and/or cerebrovascular disease, medication information, blood pressure, lipid levels, current smoking, waist circumference and BMI	Identifications of individuals whose CAS >50% (moderate stenosis) and >70% (severe stenosis)	AUC: 0.82 (moderate stenosis), 0.87 (severe stenosis)
Khanna *et al.* [128]	Non-linear parametric model	Typical risk factors (ethnicity, gender, age, artery type, BMI, Hb, hypertension, LDL, smoking), Types of carotid automated image phenotypes	Five carotid artery plaques	AUC: 0.96 (average IMT),
Hao-wen Li *et al.* [130]	multivariate LR analyses	Male, Age, BMI, Hypertension, DM, Hyperlipidemia, Previous history of TIA, Previous history of stroke, smoking, TG, TC, LDL, HDL, FBG, Hb	distinguish vulnerable and stable carotid plaques	ADAMTS7 variants rs3825807 and rs7173743 associated with the risk for carotid plaque vulnerability

2.5 Serum Markers in Carotid Artery Disease

Regarding the contribution of serum biomarkers to the pathogenesis of atherosclerosis, biomarkers distinct into inflammatory, endothelial and cell adhesion, matrix-degrading or proteolysis biomarkers, lipid and metabolic ones and their contribution to the carotid atherosclerosis pathophysiology, is described below.

Inflammatory biomarkers

Inflammation process has a significant role in mediating all stages of atherosclerotic disease. Immune cell types (monocytes, macrophages, T-cells and neutrophils) and specialized lipid mediators, have a significantly contribution to the vascular inflammation and are activated by risk factors present in the vascular wall, such as shear stress, oxidized lipoproteins and oxidative stress [131]. C-reactive protein (CRP) is considered as one of the most significant biomarkers of inflammation and the measurement of both CRP and high-sensitivity CRP (hs-CRP) is widely used in clinical practice for the vascular disease stratification. Several studies have indicated that hs-CRP is associated with the presence of unstable CAS [132], the presence of ICA stenosis [133] and the detection of vascular risk patients [134]. Elevated levels of hs-CRP are associated with lower echogenicity of carotid plaques, suggesting a relation between the hs-CRP and the potential vulnerability of the plaques [135], whereas vulnerable atherosclerotic plaques indicated upregulation of hs-CRP [136]. Nevertheless, increased levels of hs-CRP are independently associated with high risk of ischemic stroke [137] and among the risk factors for acute Anterior Circulation Stroke [138].

Pentraxin-3 (PTX-3), another acute phase protein, constitutes a potential inflammatory biomarker and is shown to be independently associated with the severity of carotid atherosclerosis [139] [140]. Additionally, Shindo *et al*. investigated the prognostic significance of PTX-3 of the vulnerability of carotid plaques through immunohistochemical analysis. [136]

Interleukin (IL)-6, a pleiotropic proinflammatory cytokine, has been shown to be localized into inflammatory cells in the vulnerable plaques and its elevated levels are associated with high risk of atherosclerotic plaques [135] [136]. Additionally, IL-6 was shown to be associated with the presence of ICA stenosis [9], and also through a genetic association study, it was indicated that IL-6 is associated with ICA stenosis [17]. On the other hand, tumor Necrosis Factor alpha (TNF-α), an inflammatory cytokine involved in early inflammatory events, has been correlated with the prevalence and severity of CAS [140] and with high risk carotid plaques. Vulnerable atherosclerotic plaque showed upregulation of

the TNF-α and increased levels of TNF-α are directly linked with high size carotid plaques [136] [141]. Immunohistochemistry analysis indicated that TNF-α combined with hypoxia and oxidized LDL markedly increased MMP-7 expression, which is directly associated with the symptomatic carotid artery disease [142], whereas immunohistochemistry analysis of plaques after CEA indicated that TNF-α was significantly increased in symptomatic patients [143].

Endothelial and cell adhesion biomarkers

Cell adhesion molecules (CAMs) are responsible for the regulation of the inflammatory response and the endothelial function. Selectins, a family of cell-surface glycoproteins, are involved in the rolling and anchoring of leukocytes on the vascular wall, whereas intercellular adhesion molecules (ICAMs) and vascular cell adhesion molecules (VCAMs), induce firm adhesion of inflammatory cells at the vascular surface. Expression of VCAM-1, ICAM-1, and L-selectin has been consistently observed in atherosclerotic plaques and their soluble forms have been identified in the circulation [144].

More specifically, VCAM-1 showed a significant association with the ICA stenosis [133] and the CV mortality [145]. As for the high risk plaques, vulnerable atherosclerotic plaques indicated upregulation of VCAM-1 [136]. ICAM-1, an endothelial- and leukocyte-associated transmembrane protein, showed also a significant association with CV mortality [145] and with the presence of ICA stenosis [146].

Selectins, a family of CAMs, have also been indicated to be associated with carotid artery disease, with E-selectin gene variants to be independently and significantly associated with ICA stenosis [146], whereas high risk atherosclerotic plaques showed its upregulation [136]. On the other hand, lower values of L-selectin were associated with atherosclerotic plaque size [141].

Additionally, NGAL is found in granules of activated human neutrophils and has been proposed as a valuable biomarker for the detection of unstable carotid plaques in asymptomatic patients [147].

Matrix-degrading or proteolysis biomarkers

MMPs contribute to the degradation of both matrix and non-matrix proteins, involved in the process of plaque destabilization and cap erosion and this function takes place in the extracellular environment. Different studies have indicated that MMPs play a significant role in the detection of vulnerable high risk atherosclerotic plaques in patients with advanced CAS. It was indicated that the combination of MMP-1, MMP-7 and Tissue Inhibitor of Metalloproteinase (TIMP)-1 demonstrated the highest positive predictive value (PPV) 89.4% and negative predictive value (NPV) 60.1% for patients correctly

classified as individuals with unstable and stable carotid lesions by means of blood sample analysis [148]. Additionally, levels of MMP-9 were significantly elevated in individuals with unstable atherosclerotic plaques in comparison with those with stable ones [149], whereas immunohistochemistry analysis indicated that the mRNA levels of MMP-2, MMP -7, MMP -9 and MMP -14 were elevated in vulnerable plaques, among which expression of MMP-2 and MMP -14 were the highest [150]. In a similar study, Sigala *et al*. [151] concluded that MMP-9 is directly related to plaque instability.

In another direction, other studies attempted to correlate levels of MMPs with the presence of symptomatic carotid artery disease. Elevated levels of MMP-2 and MMP-9 were observed in patients with symptomatic carotid artery disease in comparison with those without symptoms [149]. On the other hand, Abbas *et al*. [142] concluded that carotid artery disease patients had significantly high plasma levels of MMP-7, compared with healthy individuals, with the highest levels of MMP-7 in patients with symptoms within the last 2 months [142]. ICA stenosis was shown through a genetic association study to be associated with MMP-3 and MMP-9 gene variants [146]. In addition to this, total mortality was also independently associated with elevated plasma levels of MMP-7 [142], while higher serum MMP-9 levels in the acute phase of ischemic stroke were associated with increased risk of mortality and major disability [152].

Lipid biomarkers.

Lipid factors are, together with inflammatory factors, the main actors in the onset, evolution and destabilization of the atherosclerotic plaque. Firstly, LDL is independently related with the presence and extent of subclinical early systematic atherosclerosis [153]. Low levels of LDL are likely to prevent large artery atherosclerosis [154], could reduce the ischemic complications and affects the plaque stability and antithrombus formation [155]. Other studies have shown that increased levels of LDL were independent risk factor for the occurrence of CAS [140] and showed positive associations with ischemic stroke [156].

The oxidation of LDL is considered as a significant atherogenic modification of LDL within the vascular wall, where oxidized low-density lipoprotein (ox-LDL) can trigger the expression of adhesion molecules on the cell surface, and thus stimulate the activation of endothelial cells These adhesion molecules mediate the rolling and adhesion of blood leukocytes, that adhere to the endothelium and migrate into the intima, causing the activation of macrophages, the release of proinflammatory cytokines and the production of proteolytic enzymes, contributing to the matrix degradation and plaque destabilization

[157]. The role of circulating ox-LDL has gained considerable attention and low levels of ox-LDL are a promising therapeutic target against atherosclerosis [158]. In a study conducted by Sigala *et al.* [151], it was indicated that ox-LDL levels were associated with the presence of clinical symptoms of carotid artery disease. Additionally, there are efforts to clarify the correlation between the morphology of human atherosclerotic plaques and the ox-LDL levels in plasma and plaques. It was shown that elevated ox-LDL levels are related with the vulnerability to rupture [159]. Ox-LDL levels are also considered as significant biomarker for the prognosis of carotid artery disease. In a study by Markstad *et al*. [160], the authors indicated that ox-LDL leads to the release of Soluble lectin-like oxidized LDL receptor-1 (sLOX-1) from endothelial cells and that circulating levels of sLOX-1 are associated with the risk of ischemic stroke, whereas Wang *et al.* [161] showed that elevated levels of ox-LDL were associated with the high risk mortality and poor functional outcome within one year after stroke onset.

On the other hand, high HDL is characterized by its antioxidant, antithrombotic, anti-inflammatory and antiapoptotic characteristics and may play a significant protective role in acute stroke, protecting and limiting the ischaemia on the blood–brain barrier and on the parenchymal cerebral compartment [162]. Lower levels of HDL were independently associated with an increased risk of having echolucent, rupture-prone atherosclerotic plaques [163] [164] [165]. Moreover, HDL contributes as a prognostic marker for the severity of stroke, since low baseline HDL (≤35 mg/dL) at admission was associated with higher stroke severity and poor clinical outcome during follow-up in patients with atherosclerotic ischemic stroke [166].

Triglyceride-rich lipoproteins (TRLs), a pool of lipoproteins that includes chylomicrons, very LDLs, intermediate-density lipoproteins and other remnant lipid metabolism particles, appear to promote atherogenesis independently of LDL [167]. Elevated levels of TRLs seem to be associated with plaque echolucency, which is characterized by increased lipid content and macrophage density plaques. Echolucent carotid plaques were proven to be related with higher risk for future ischemic stroke, particularly in previously symptomatic individuals, for restenosis after CEA, as well as for MI [165] [168]. On the other hand, in a study presented by Kofoed *et al*. [168], it was shown that TRLs are elevated in patients with CAS higher than 50%, compared with controls.

Although circulating lipoprotein-associated phospholipase A(2) (Lp-PLA(2)) has been considered as a novel biomarker for the CVDs, its correlation between the atherosclerotic plaque expression of Lp-PLA(2), inflammation, stability and the presence of clinical symptoms, especially for cerebrovascular

disease remains poorly defined. Nevertheless, in a study conducted by Mannheim *et al.* [169], it was shown that symptomatic carotid artery plaques are characterized by increased levels of Lp-PLA(2), strongly supporting the role of Lp-PLA(2) in the pathophysiology and clinical representation of cerebrovascular disease. In addition to this, Lp-PLA2 consists a significant biomarker for the management of asymptomatic patients with carotid artery disease, since its elevated levels is directly associated with high grade of CAS [170], and with the presence of unstable atherosclerotic plaques [170] [138]. Regarding the prognostic significance of Lp-PLA2, it has been shown that its activity is an independent predictor for coronary heart disease and ischemic stroke in the general population [171] and Lp-PLA2 in its highest levels had an increased risk of recurrence after the first ischemic stroke [172].

ApolipoproteinS (Apos) are the protein components of plasma lipoproteins, which consist of a core of triglyceride and cholesterol esters and a peripheral region of phospholipid, sphingolipid and protein. The most relevant subtypes are considered the Apo A-I (the main protein on HDL), Apo B-100 (the main protein on LDL), Apo C-II (important in chylomicrons and very LDL, activates lipoprotein lipase), and Apo E (present in chylomicrons, very LDL, and intermediate density lipoprotein, allowing the binding of these lipoproteins to the hepatocytes).

ApoA-I (or ApoA1) levels may be clinically useful for the diagnosis of stroke and for the differentiation between ischemic and hemorrhagic strokes [173], whereas reduced ApoA-I levels are risk factors for a first ischemic stroke and elevated ApoA-I is considered as a risk factor for a first hemorrhagic stroke [174]. A meta-analysis by Paternoster *et al.* [175] indicated a clear association of APOe with carotid IMT, suggesting the possibility of a specific association with large artery ischemic stroke.

Proprotein convertase subtilisin kexin type 9 (PCSK-9) is a protease produced at the liver and is detectable in human plasma. It consists a key regulator of the metabolism of LDL, and recently has been suggested to participate in the development of atherosclerosis [176]. Chan *et al*. [177] showed that serum PCSK-9 levels is considered as an independent predictor of carotid IMT and may contribute to increased risk of subclinical carotid atherosclerosis, independently of conventional risk factors. On the other hand, *Xie et al.* [178] concluded that plasma PCSK-9 levels are associated with 10-year progression of atherosclerosis.

Metabolic biomarkers

Proinflammatory chemerin, leptin, and resistin are considered as adipokines which influence the vascular wall function. Association of circulating adipokines with the cerebrovascular symptomatology and the carotid plaque vulnerability was investigated and it was shown that low levels of chemerin and elevated levels of restinin were related to the plaque instability, the risk of stroke and the severity of carotid artery disease [179].

On the other hand, adiponectin, an anti-inflammatory and vasculoprotective adipokine, may act as a novel prognostic biomarker for atherosclerosis in stroke, since it was shown to be related with the risk of ischemic stroke [180] [181]. In a study by Gustafsson *et al.* [182], the role of circulating adiponectin to the vascular function and morphology was investigated and the authors concluded that elevated levels are associated with less arterial pathology, while Saarikoski *et al.* [183] supported the role of adiponectin in the pathophysiology of early atherosclerosis.

Fatty acid binding protein 4 (FABP4) has been also shown to play an important role in macrophage cholesterol trafficking and has been considered as a key factor connecting the vascular and lipid accumulation with the inflammation process. Increased levels of FABP4 are associated with the presence of carotid artery disease, plaque instability and adverse outcome in patients with carotid atherosclerosis, since the highest mRNA levels of FABP4 have been observed in patients with the most recent symptoms [184] [185].

Elevated levels of homocysteine (hcy) have been associated with carotid plaque development and hyperhomocysteinemia has been described as an independent CV risk factor. More specifically, in a study presented by Alsulaimani et al. [186], it was shown that elevated hcy was independently associated with plaque morphology and increased plaque area, subclinical markers of stroke risk, whereas high hyperhomocysteinemia prevalence in patients with extracranial cerebrovascular disease was confirmed by Alvarez et al. [187]. In addition to this, higher total hcy levels were associated with asymptomatic carotid artery disease [188].

Osteoprotegerin (OPG) is a secretory glycoprotein which belongs to the TNF receptor family and its high concentration has been related with the CV and vascular disease and contributes to the atherosclerotic plaque stability. Studies have shown that higher levels of OPG have been observed in asymptomatic plaques related to symptomatic ones [189] and all plaques exhibited calcification were

significantly higher in asymptomatic patients. The effect of OPG in atherosclerotic plaques was confirmed by a study by Schiro et al., in which OPG was significantly elevated in symptomatic individuals related to asymptomatic group [190].

In Table 2.8, all the types of biomarkers related to carotid artery disease and their associations with the disease, are presented.

In addition to the contribution of each type of biomarker to atherosclerosis pathophysiology, different studies have confirmed the association of serum biomarkers with the presence of CAS, the diagnosis of vulnerability of carotid atherosclerotic plaques, the presence of symptomatic carotid artery disease and their prognostic value for the future stroke and CV mortality.

Regarding the detection of carotid artery disease presence, there different studies that aim to identify the most significant serum biomarkers associated either with the presence of CAS greater than 50% [132] [133] [140] [146] [168], or with the early detection of subclinical carotid atherosclerosis [134] [183]. In these studies, the presence of carotid artery disease was mainly assessed by carotid US and the analysis of data was implemented using conventional statistical techniques, as it is reported in Table 2.9.

Nevertheless, the focus of CAS diagnosis was also shifted from pure stenosis quantification to plaque characterization and more specifically to the detection of high risk vulnerable plaques. This has led to further advancements in the existing imaging tools and identification of high risk plaque related imaging characteristics and to the identification of plaque histology characteristics related to high risk plaques [191]. Table 2.10 indicates all the existing methodologies for the detection of vulnerable plaques. In most of these studies, the instability of atherosclerotic plaques was considered after the immunohistochemistry analysis of carotid plaques [136], whereas, in other studies vulnerable plaques was considered after the implementation of carotid US [141]. All types of serum biomarkers seem to be associated with the vulnerability of plaques and some of them have been also detected as useful biomarkers for the diagnosis of CAS, such as the hs-CRP, the PTX-3, the IL-6 and the TNF-α, concluding that the measurement of inflammatory biomarkers contribute both to the diagnosis and the monitoring of the disease progression.

Additionally inflammatory TNF-α, matrix degrading biomarkers (MMP-2, MMP-7, MMP-9), ox-LDL and metabolic biomarkers (restinin, OPN, OPG) have been shown to be associated with symptomatic

carotid artery disease, as presented in Table 2.11. Early detection of symptomatic carotid artery disease patients may provide a cost-effective disease detection and management strategy.

Apart from the diagnostic based circulating biomarkers, in this section, biomarkers with high prognostic value for the progression of carotid artery disease are also reported. More specifically, biomarkers associated with high risk of future stroke or CV event are described in Table 2.12. Circulating biomarkers which constitute the typical lipid patient profile (LDL, TC, triglyceride, ox-LDL, Lp-PLA2), inflammatory biomarkers such as hs-CRP, TNF-α, IL-6, adipokines (Adiponectin, Leptin, FABP4) and hcy are related to high risk of future stroke. Meta-analyses, conducted by Gorgui *et al.* [62] and Gairolla *et al*. [63], concluded that levels of adiponectin and leptin are significantly associated with ischemic stroke, showing that adipokines may have a cause-effect relation with carotid artery disease. As for the CV mortality, typical hs-CRP biomarker, endothelial and cell adhesion biomarker (VCAM-1, ICAM-1), MMPs (MMP7. MMP-9), ox-LDL and FABP4 seem to be associated with the CV mortality, as it is shown in Table 2.13. An indicative study by Zhong *et al.* [152], concluded that the high value of MMP-9 was associated with high CV mortality and major disability. In this study, data collection from 3,186 participants were analyzed and 767 of them have experienced major disability or died. On the other hand, Wang *et al.* [161] in another study collected biochemical markers from 3688 patients and concluded that the patients in the highest ox-LDL quartile had a higher risk of one year stroke mortality.

In Table 2.14, we summarize all the presented studies, based on their participants and their clinical characteristics, providing specific biomarkers for the risk stratification of asymptomatic participants, of symptomatic participants (either with diagnosis of carotid artery disease or undergoing CEA), of symptomatic, asymptomatic and controls participants and of the general population.

Table 2.8 Serum biomarkers related to clinical outputs of carotid artery disease.

Type of Biomarkers	Biomarkers	Diagnosis of Carotid Artery Disease	Plaque Vulnerability	Symptomatic Carotid Artery Disease	Future Stroke Event	Stroke Severity	CV Mortality
Inflammatory	hs-CRP	[133] [132, 134]	[135] [136]		[137] [192] [138]		[193]
	PTX-3	[140] [139]	[136]				
	IL-6	[133] [146]	[135] [136]				
	TNF-α	[140]	[136] [141] [148]	[142] [143]			
Endothelial and Cell Adhesion Biomarkers	NGAL		[147]				
	VCAM-1		[136]				[145]
	ICAM-1	[146]					[145]
	E-selectin	[146]	[136]				
	L-selectin		[141]				
Matrix Degrading or Proteolysis	MMP-1		[148]				
	TIMP-1		[148]				
	MMP-2		[150]	[149]			
	MMP-3	[146]					
	MMP-7		[148] [150]	[142]			[142]
	MMP-9	[146]	[149] [150] [151]	[149]			[152]

Type of Biomarkers	Biomarkers	Diagnosis of Carotid Artery Disease	Plaque Vulnerability	Symptomatic Carotid Artery Disease	Future Stroke Event	Stroke Severity	CV Mortality
	MMP-14		[150]				
Lipid Related	LDL	[140]			[156]		
	TC				[156]		
	triglyceride				[156]		
	ox-LDL		[159]	[151]	[160]		[161]
	HDL		[164] [163]			[166] [156]	
	TRL	[168]	[165] [168]				
	Lp-PLA2	[170]	[170] [138]	[169]	[171] [172]		
	apoA-I					[173] [174]	
	apoE						
	PCSK9	[177]					
Metabolic Biomarkers	restinin		[179]	[179]			
	adiponectin	[183]		[180] [181]			
	leptin			[181]			
	FABP4	[184]	[184]	[184]			
	homocysteine	[188] [186]			[186]		
	OPG		[189]	[189] [190]			

Table 2.9 A summary of studies investigating biomarkers related to carotid artery disease diagnosis.

Study	**Biomarker**	**Methodology**	**Dataset**	**Output**
Puz *et al.* [132]	hs-CRP	laboratory tests, US examination, statistical analysis	65 patients with ICA stenosis > 50% (39 symptomatic) and 30 healthy	Patients with ICA stenosis had significantly higher serum concentrations and CRP values than the individuals from the control group ($p = 0.009$)
Horn *et al.* [134]	hs-CRP	laboratory tests, US examination, statistical analysis	subclinical and advanced (INVADE study-n = 3,092, > 55 years)	rate of both subclinical and advanced stages of atherosclerosis was higher in patients with pathological hs-CRP
Debing *et al.* [133]	hs-CRP, IL-6, sVCAM-1	statistical analysis, high-resolution B-mode US	180 patients with ICA stenosis, 180 age-matched and sex-matched controls.	levels of hs-CRP, VCAM-1, and IL-6 in the CEA group were significantly higher than in the control group
Yi *et al.* [140]	PTX-3, TNF-α, LDL	CTA, statistical analysis	206 patients with ischemic stroke	plasma levels of PTX-3, TNF-α, and LDL were increased significantly in the CAS group patients vs. the CAS-free
Knoflach *et al*. [139]	PTX-3	statistical analysis, high-resolution B-mode US	132 young men, 205 young women, 562 individuals 55 to 94 years old	PTX-3 level was independently associated with prevalent
Biscetti *et al.* [146]	IL-6, ICAM-1, E-selectin, MMP-3, MMP-9	genetic association study	933 individuals (344 patients with ICA stenosis-CEA and 589 controls)	IL-6, ICAM-1, MMP-3, and MMP-9 gene polymorphisms were independently associated with ICA stenosis
Kofoed *et al.* [168]	TRL	high-resolution B-mode US and computerized image analysis	66 controls and 323 patients with CAS ≥ 50%	fasting and postprandial triglyceride-rich lipoproteins are elevated in patients with CAS of ≥ 50%
Chan *et al*. [177]	PCSK-9	biochemical analysis, carotid US	95 asymptomatic subjects	serum PCSK-9 remained an independent predictor of mean carotid IMT
Saarikoski *et al.* [183]	adiponectin	US data on carotid IMT	2,147 young adults	low serum adiponectin concentration is independently related with increased carotid IMT early atherosclerosis
Holm *et al*. [184]	FABP4	enzyme immunoassay, statistical analyses	28 asymptomatic, 31 symptomatic, 202 patients with acute ischemic stroke	FABP4 levels were higher in patients with carotid atherosclerosis
Jia *et al.* [188]	hcy	carotid duplex US examination	5393 Chinese participants	hcy> 19.3μmol/L was considered as an independent indicator of asymptomatic CAS
Alsulaimani *et al.* [186]	hcy	ultrasonographic assessment of plaque morphology measured by gray-scale median	1327 stroke-free subjects	increasing hcy was associated with an increasing risk

Table 2.10 A summary of studies investigating biomarkers related to carotid atherosclerotic plaque vulnerability.

Study	**Biomarker**	**Methodology**	**Dataset**	**Output**
Yamagami *et al.* [135]	hs-CRP, IL-6	statistical analyses	246 patients, including 80 patients with a history of stroke/TIA	IL-6 and hs-CRP were negatively correlated with carotid plaque echogenicity
Shindo *et al.* [136]	hs-CRP, PTX-3, IL-6, TNF-α, VCAM-1	histological analysis, statistical analysis	58 patients with CAS	vulnerable group showed upregulation of proinflammatory cytokines, endothelial activation markers and inflammation markers. and downregulation of anti-inflammatory markers
Andersson *et al.* [141]	TNF-a	US procedure, measurement serum markers, statistical analysis	1,016 subjects	Plaque size was also related to increased levels of TNF-α
Pelisek *et al*. [148]	TNF-α, MMP-1, TIMP-1, MMP-7	Measurement of serum markers by ELISA assays, Multiscore analysis	patients (n = 101) were classified as histologically stable (n = 37) or unstable (n = 64).	Circulating levels of MMP-1, MMP-7, TIMP-1, and TNF-α were significantly enhanced in patients with unstable plaques
Guo *et al*. [150]	MMP-2, MMP-7, MMP-14	carotid plaque specimens, histology and immunohistochemistry analysis	64 patients	The mRNA levels of MMP-2, MMP-7, MMP-9 and MMP-14 were elevated in vulnerable plaques
Alvarez *et al.* [149]	MMP-9	histopathologic analysis, immunohistochemistry (macrophage count, T lymphocytes, activated T lymphocytes)	40 patients with CAS	MMP-9 was also significantly higher in the symptomatic group and in patients with unstable plaques
Eilenberg *et al.* [147]	NGAL, MMP-9	histological investigation, statistical analysis	83 patients with asymptomatic CAS	Circulating NGAL and MMP-9/NGAL are significantly increased in asymptomatic patients with vulnerable carotid atherosclerotic plaques
Sarlon-Bartoli *et al.* [170]	Lp-PLA2	laboratory measurements, histological assessment and immunohistochemistry of carotid plaques, statistical analyses	42 patients (neurological symptoms were present in 16, unstable plaques in 23)	Plasma Lp-PLA2 level was independently associated with unstable carotid plaques

Study	Biomarker	Methodology	Dataset	Output
Yang *et al.* [138]	Lp-PLA2	laboratory examination, carotid ultrasonography and grouping, statistics analysis	100 patients with acute anterior circulation stroke and 50 noninfarction subjects (control group)	Hs-CRP and Lp-PLA2 levels were significantly higher in vulnerable plaque group than in mixed plaque group and stable plaque group
K. Nishi *et al*. [159]	Ox-LDL	histopathological characteristics of plaques, immunohistochemical analysis, statistical analysis	44 patients	The ox-LDL level was significantly higher in vulnerable than stable plaques
Mathiesen *et al.* [163]	HDL	ultrasonography, statistical Analysis	216 with CAS, 223 control subjects	Low levels of HDL are associated with an increased risk of having echolucent, rupture-prone atherosclerotic plaques
Peters *et al*. [164]	HDL	Measurement of IMT, statistical analysis	984 individuals	Low levels of HDL are related to echolucency of the carotid intima-media
Nordestgaard *et al*. [165]	TRL, HDL	US imaging methods, histological characterization	111 asymptomatic, 135 symptomatic patients with CAS, 44 ipsilateral ischaemic strokes	Vulnerable plaques are associated with elevated levels of triglyceride-rich lipoproteins and with reduced levels of HDL
Kofoed *et al.* [168]	TRL	high-resolution B-mode US and computerized image analysis	66 controls and 323 patients with CAS ≥ 50%	Fasting and postprandial triglyceride-rich lipoproteins are elevated in patients with CAS of ≥ 50% compared with controls
Gasbarrino *et al.* [179]	Restinin, Chemerin	association of circulating adipokines and carotid plaque instability	n=165 symptomatic and asymptomatic patients	Low chemerin and high resistin levels were associated with plaque instability
Holm *et al.* [184]	FABP4	enzyme immunoassay, statistical analyses	28 asymptomatic, 31 symptomatic, 202 acute ischemic stroke	FABP4 is linked to plaque instability in patients with carotid atherosclerosis
Davaine *et al.* [189]	OPG	OPG measurement, histological and immunological analyses, statistical analysis	73 carotid plaques (49 asymptomatic and 24 symptomatic)	Circulating OPG levels were higher in the plasma of asymptomatic patients

Table 2.11 A summary of studies investigating biomarkers related to symptomatic carotid artery disease.

Study	**Biomarkers**	**Methodology**	**Dataset**	**Output**
Abbas *et al.* [142]	MMP-7	immunohistochemistry, statistical analyses	182 consecutive patients with moderate (50–69%), 23 healthy controls	MMP-7 could contribute to plaque instability
Schneiderman *et al.* [143]	TNF-α	lesion analysis, statistical analysis	40 symptomatic, 38 asymptomatic patients with progressive stenosis	TNF-α was significantly increased in symptomatic patients
B. Alvarez *et al.* [149]	MMP-2, MMP-9	MMP-2 and MMP-9 measurement, statistical analysis	40 patients with CAS	Elevated MMP-9 concentration is associated with carotid plaque instability
Sigala *et al.* [151]	Ox-LDL	Immunohistochemistry, data analysis and statistics	36 patients undergoing CEA, 20 controls	Ox-LDL was increased in symptomatic patients
Gasbarrino [61]	Restinin	Measurement of circulation adipokines, data analysis and statistics	165 neurologically symptomatic and asymptomatic patients	restinine levels were significantly elevated in symptomatic
A. Schiro *et al.* [190]	OPN, OPG	Measurement of circulating EMPs, platelet MPs (PMPs) and inflammatory markers, statistical analysis	-	OPN and OPG were significantly elevated in the symptomatic

Table 2.12 A summary of studies investigating biomarkers related to future stroke event.

Study	Biomarkers	Methodology	Dataset	Output
Zhou *et al.* [137]	hs-CRP	Meta-analysis	2436 ischemic strokes, 655 hemorrhagic strokes from 66,560 participant	When comparing the highest with the lowest hs-CRP category, the pooled RR of ischemic strokes was 1.46
Yang *et al.* [138]	hs-CRP	Laboratory examination, carotid ultrasonography and grouping, statistics analysis	100 patients with acute anterior circulation stroke and 50 controls	Hs-CRP and Lp-PLA2 are among the risk factors for anterior circulation stroke
Ma *et al.* [192]	hs-CRP, TNF-α, IL-6	Data collection, statistical analysis	288 ischemic stroke patients and 300 controls	Hs-CRP, TNF-α, and IL-6 are considered to be important markers of the body's inflammatory state in ischemic stroke
Gu *et al.* [156]	LDL, TC, triglyceride	Baseline information collection, statistical analysis	Six cohort studies in China with 267, 500 participants	TC, LDL, and triglyceride showed positive associations with ischemic stroke
Markstad *et al.* [160]	Ox-LDL	Analyses of the plaque tissue: cytokines and chemokines, measurement of sLOX-1 in blood samples, statistical analysis	202 patients undergoing CEA	Ox-LDL induces the release of sLOX-1 from endothelial cells and that
Oei *et al.* [171]	Lp-PLA2	Measurement of Lp-PLA2 activity, statistical analysis	308 coronary heart disease cases, 110 ischemic stroke cases, and a random sample of 1820 subjects	Lp-PLA2 activity is an independent predictor of ischemic stroke
Elkind *et al.* [172]	Lp-PLA2	Measurement of Lp-PLA2 activity, statistical analysis	467 patients	Stroke patients with Lp-PLA2 activity levels in the highest quartile had an increased risk of recurrence after first ischemic stroke
Alsulaimani *et al.* [186]	hcy	Assessment of hcy, assessment of carotid atherosclerosis, statistical analysis	1327 stroke-free subjects	elevated hcy was independently associated with plaque morphology
Gorgui *et al.* [62]	Adiponectin	systematic review and meta-analysis	-	increased adiponectin levels were associated with an increase in risk for ischemic stroke
Gairolla *et al.* [63]	Adiponectin, leptin	systematic review	-	adiponectin and leptin are significantly associated with stroke

Table 2.13 A summary of studies investigating biomarkers related to CV mortality.

Study	**Biomarkers**	**Methodology**	**Dataset**	**Output**
Mayer *et al.* [193]	hs-CRP	Clinical and laboratory data collection, statistical analysis	1065 patients with neurological asymptomatic carotid atherosclerosis	The risk of all-cause and CV mortality significantly increased in patients with elevated serum levels of hs-CRP
Hoke *et al.* [145]	VCAM-1, ICAM-1	Measurement of CAMs, statistical analysis	855 patients	significant association between CV mortality and ICAM-1
Abbas *et al.* [142]	MMP-7	Immunohistochemistry, statistical analyses	182 consecutive patients with moderate (50–69%) levels, 23 healthy controls	high plasma levels of MMP-7 were independently associated with total mortality
Zhong *et al.* [152]	MMP-9	Data collection, statistical analysis	3186 participants, 767 participants (24.6%) experienced major disability or died	Higher log MMP-9 was associated with death and major disability.
Wang *et al.* [161]	Ox-LDL	Biochemical indicators, diagnosis of stroke and stroke subtype classification, statistical analysis	3688 patients	Patients in the highest ox-LDL quartile had a higher risk of 1-year stroke mortality
Holm *et al.* [184]	FABP4	Enzyme immunoassay, statistical analyses	asymptomatic (n = 28), symptomatic (n = 31), patients with acute ischemic stroke (n = 202)	levels of FABP4 were significantly associated with total and CV mortality

Table 2.14 Biomarkers related with different clinical outputs of carotid artery disease in i) Symptomatic, Asymptomatic & Controls, ii) Asymptomatic, iii) Symptomatic and in iv) General population.

i) Symptomatic, Asymptomatic & Controls	
Carotid Artery Disease Diagnosis	hs-CRP [132] [134] [133], IL-6 [133] [146], sVCAM-1 [133], ICAM-1 [146], E-selectin [146], MMP-3 [146], MMP-9 [146], TRL [168], FABP4 [184]
Carotid Atherosclerotic Plaque Vulnerability	hs-CRP [135] , IL-6 [135], TRL [165] [168], HDL [165] [163], Restinin [179], Chemerin [179], FABP4 [184], Osteoprotegerin [189], Lp-PLA2 [170] ,TNF-α [148], MMP-1 [148], TIMP-1 [148], MMP-7 [148]
Symptomatic Carotid Artery Disease	MMP-7 [142], TNF-α [143] [192], hs-CRP [192], IL-6 [192], Ox-LDL [151], Restinin [61]
CV Mortality	MMP-7 [142], FABP4 [184]
ii) Asymptomatic	
Carotid Artery Disease Diagnosis	PCSK-9 [177], hcy [186]
Carotid Atherosclerotic Plaque Vulnerability	NGAL [147], MMP-9 [147]
Future Stroke Event	Homocysteine [186]
CV mortality	hs-CRP [193]
iii) **Symptomatic patients (either with diagnosis of carotid artery disease or undergoing CEA)**	
Carotid Artery Disease Diagnosis	PTX -3 [140], TNF-α [140], LDL [140]
Carotid Atherosclerotic Plaque Vulnerability	MMP-9 [149], Lp-PLA2 [138]
Future Stroke Event.	hs-CRP [137] [138], MMP-2 [149], MMP-9 [149], Ox-LDL [160], Lp-PLA2 [171]
iv) **General population**	
Carotid Artery Disease Diagnosis	Adiponectin [183], hcy [188], PTX-3 [139]
Carotid Atherosclerotic Plaque Vulnerability	MMP-2 [150], MMP-7 [150], MMP-14 [150], Ox-LDL [159], HDL [164], TNF-α [141]
Future Stroke Event.	LDL [156], TC [156], triglycerides [156], Lp-PLA2 [172]
CV Mortality	MMP-9 [152], Ox-LDL [161], VCAM-1 [145], ICAM-1 [145]

2.6 Contribution of this thesis

It is clear from the abovementioned literature overview regarding the coronary and carotid lumen and atherosclerotic plaques detection (Table 2.1, Table 2.2, Table 2.3) that the majority of the studies are based either on the vessel or the atherosclerotic plaque detection. Most of the studies, concerning lumen detection, are performed using typical image processing techniques, whereas atherosclerotic plaque detection is performed based on HU analysis, with specific ranges for each plaque component. Additionally, the evaluation process of the existing studies includes either the comparison with other imaging techniques or the qualitatively evaluation.

In this thesis, we focus on developing an overall pipeline for the reconstruction of the entire coronary arterial tree and the carotid artery bifurcation non-invasively using CTA, which provides 3D models of the lumen, the outer wall, the CP and the NCP in order to overcome the literature limitations, regarding the implemented methodology and the validation strategy. Our methodology for reliable segmentation of the CTA images renders it a useful tool in the clinical and research arena.

In particular, our approach uses active contours to detect the inner, outer wall and CP, whereas the NCP segmentation is performed using a fully adaptive HU- based threshold technique in order to be independent by the different CTA scanners. Additionally, our proposed pipeline integrates an algorithm for the blooming effect removal, in cases of high HU intensities values such as in CP or stented arteries. The proposed segmentation method is validated using experts' manual annotations, but also to enhance our validation results it is compared with an invasive imaging modality i.e. IVUS and VH-IVUS modality. The DICE coefficient for the lumen segmentation is 0.75, whereas for the CP and NCP was 0.72 and 0.7, respectively, which is similar or better than most of the other studies presented in the current literature.

The obtained results indicate that the proposed methodology allows reliable and automated detection of the luminal and vessel wall borders and fast and accurate characterization of plaque type in CTA images. The previous presented methodologies have limitations such as that they are based on high quality images or on thresholding techniques. Our method can handle any kind of CTA data, while it combines arterial reconstruction with plaque characterization. In addition to this, all the potential validation strategies have been performed for the evaluation of our methodology. Our methodology has been integrated into useful tools clinically evaluated for both the coronary and carotid artery disease [194].

Accurate CTA image analysis has led to the development of CAD risk prediction models using geometry-based risk factors. The overall 3D reconstruction pipeline has been utilized for the calculation of geometry based features, such as the DS, the MLA, the MLD, the presence and volume of CP. Additionally, biomechanical features, such as the SmartFFR and WSS, have been utilized for CAD risk prediction models. As it is shown in Table 2.6, statistical analysis and ML models have been implemented to predict the progression of the CAD, the prediction of MI and the CV-related mortality, using as input imaging and non-imaging data.

This thesis focuses on presenting a pipeline for the imaging based CAD risk prediction, which includes the following steps: (i) CTA image analysis, (ii) calculation of imaging-based parameters, (iii) dataset collection and curation, (iv) problem definition, (v) class imbalance handling, (vi) feature selection and classification and finally (vii) the model evaluation. Four different approaches have been proposed for CAD prediction, which include as outcome the prediction of obstructive CAD (2-class or 3-class problems) and the site-specific CAD progression and PCI placement. The achieved accuracies were 0.81, 0.67, 0.74 and 0.78, indicating promising results for the CAD risk prediction.

Comparing with other existing studies in the literature, our proposed methodology takes into account imaging features, deriving from the non-invasive CTA imaging modality, whereas other existing studies calculate imaging features from IVUS modality. Regarding the first approach, the implementation of ML models for CAD prediction, using as input the combination of both imaging (geometrical and biomechanical) constitutes a novelty of my thesis. It is clear from the literature that assessing and predicting CAD severity was achieved traditionally using statistical modelling and not ML. In spite of the reported good discrimination ability of such parametric regression models, a recent systematic review demonstrated the paucity of external validation and head-to-head comparisons, the poor reporting of their technical characteristics as well as the variability in outcome variables, predictors and prediction horizons, which limits their applicability in evidence-based decision making in healthcare [195]. Thus, the limitations of statistical analysis techniques are overcome through this thesis with the implementation of classical ML models for CAD prediction. In conclusion, the work of this thesis related to CAD is in line with the latest CAD guidelines regarding the necessity of using non-invasive CTA imaging for diagnosis of CAD. More specifically, using our methodology, besides the CAD diagnosis which is provided by CTA, prediction of coronary stenosis or disease progression will be delivered to the clinician to further support the treatment selection.

Regarding the carotid artery disease prediction, through this thesis, attempts have been undertaken to detect the presence of carotid disease, as it is defined by CAS and the presence of carotid plaques vulnerability, as they are defined either by histology based features or by the presence of symptomatic disease Most of the existing studies, concerning the CAS presence and high risk plaques presence, are basically based on inflammatory biomarkers (hs-CRP, PTX-3, IL-6, TNF-α), ICAM-1, VCAM-1, E-selectin and MMPs and are basically based on statistical analysis. In this thesis, we present a ML based model for the detection of CAS presence, taking into account demographics, clinical data, typical risk factors and medication therapy, whereas for the histology based high risk plaques detection, lipid related biomarkers are used, as well. Comparing to other similar studies in the literature, our approach achieves a higher accuracy (0.78 for CAS detection) and 0.67 (average for high risk plaque detection), based only on typical health records, easily recorded by a general practitioner. External validation dataset have been also used for the validation of ML model focused on CAS presence detection. More specifically, 512 individuals were utilized for the evaluation of our model and the overall accuracy was 0.89. Comparing to the similar existing in the literature studies, as it is shown in Table 5.27, the achieved accuracies either with the training dataset or with the external validation datasets, are higher. In addition to this, in this thesis, our aim is to contribute to the carotid artery disease early detection only for the asymptomatic individuals. Thus, retrospective datasets only for asymptomatic participants were utilized for the development of the diagnosis of CAS. As it is shown in Table 2.14, there are only two studies, which aim to detect CAS presence in asymptomatic participants.

Moreover, diagnostic prediction of high risk plaques has been defined based on histology related components, such as the total collagen, the smooth muscle cells, the neutrophils, the lipid and the macrophages. Comparing with other existing in the literature studies (shown in Table 2.10), which basically define high risk plaques based on US-derived features, the histology based clinical outcome definition constitutes a novelty of the proposed models. However, regarding the existing in the literature studies for high risk plaques detection, there are not similar studies in the literature, which use typical medical records for the plaque vulnerability detection to be directly compared with our proposed models.

Additionally, through this thesis, ML based models have been developed for the detection of high risk plaques, as they are defined by the presence of symptomatic carotid artery disease. Two ML models have been developed and their accuracies were 0.79 and 0.65. Based on the existing in the literature

studies, it is found that TNF-α, MMPs, ox-LDL and metabolic biomarkers are associated with the symptomatic disease (Table 2.8) and the association of TNF-α and MMPs (MMP-8, TIMP-2, MMP-9/TIMP-1) has been confirmed through this thesis. In addition to this, in this thesis, it was established the association of biomarkers, such as OPG, OPN, leptin, RANTES, FABP4 and GDF15 with symptomatic artery disease.

The last contribution of the proposed thesis is the investigation of the relation between carotid artery disease and silent brain lesions. A dataset of 211 participants was utilized for this analysis and US related features, clinical features, risk factors, hematological and biochemical features were used for the analysis. CAS presence, previous MI, alcohol abuse and previous CABG are positive factors for SBIs presence and high values of SBP, triglycerides, creatine and HbA1c are also associated with the presence of SBIs.

Chapter 3 3D arterial Reconstruction using CTA

3.1 Introduction

3.2 Methodology for 3D inner and outer wall reconstruction

3.3 Methodology for atherosclerotic plaque characterization

3.4 Dataset

3.5 Results

3.6 Discussion- Beyond the state of the art

3.1 Introduction

Due to its less invasive nature, a great interest has been focused on the imaging and diagnosis of CAD and carotid artery disease using CTA modality. Different studies (described in detail in Chapter 2) have indicated that CTA modality is able to analyze accurately the coronary and the carotid artery wall and provides not only the detection and quantification of the atherosclerotic plaque [196, 197], but also the classification of its composition [198, 199]. In addition to its high accuracy, CTA provides robust prognostic information in patients with suspected CAD and allows the risk stratification as well, when CAD is present [200], while it can be used for prediction of plaque growth based on computational modelling [201, 202]. Moreover, CTA provides a promising potential in detecting certain vulnerable lesions in risk of embolic events [203].

In this chapter, we present a new semi-automated methodology for 3D coronary and carotid artery reconstruction using CTA modality, implementing the active contour models for the segmentation of 2D CTA images. In order to investigate the accuracy of our methodology, we implemented a validation procedure for the inner wall and outer wall detection, utilizing both manual annotations and the IVUS modality for the coronary arteries and only manual annotations for the carotid arteries. Additionally, we examine the accuracy of the CTA modality to identify the PB region and specify the atherosclerotic plaque composition. More specifically, a methodology for the identification and volumetric quantification of CP and NCP is proposed, using density measurements, quantified by

HU. The detection of CP is achieved by an active contour based approach incorporating prior shapes, whereas the NCP is detected by an adaptive threshold based technique. The proposed methodology for coronary atherosclerotic plaque detection is evaluated using the corresponding VH-IVUS images and manual annotations, whereas the carotid atherosclerotic plaque detection is evaluated using only manual annotations.

3.2 Methodology for 3D inner and outer wall reconstruction

The proposed methodology includes 7 stages, as it is shown in Figure 3.1, below [204-206]. In the first stage, the CTA images are acquired and then they are pre-processed to detect the vessel silhouette. In the third stage, a blooming effect removal technique is implemented and in the fourth stage, a centerline extraction approach of the vessel is applied. In the fifth stage, two weight functions for the lumen and the outer wall of arteries are estimated. In the sixth stage, an extension of active contour models for the lumen and the outer wall segmentation is implemented. Finally, in the last seventh stage the 3D surfaces for the lumen and the outer wall are constructed, employing the Fast Marching cubes algorithm.

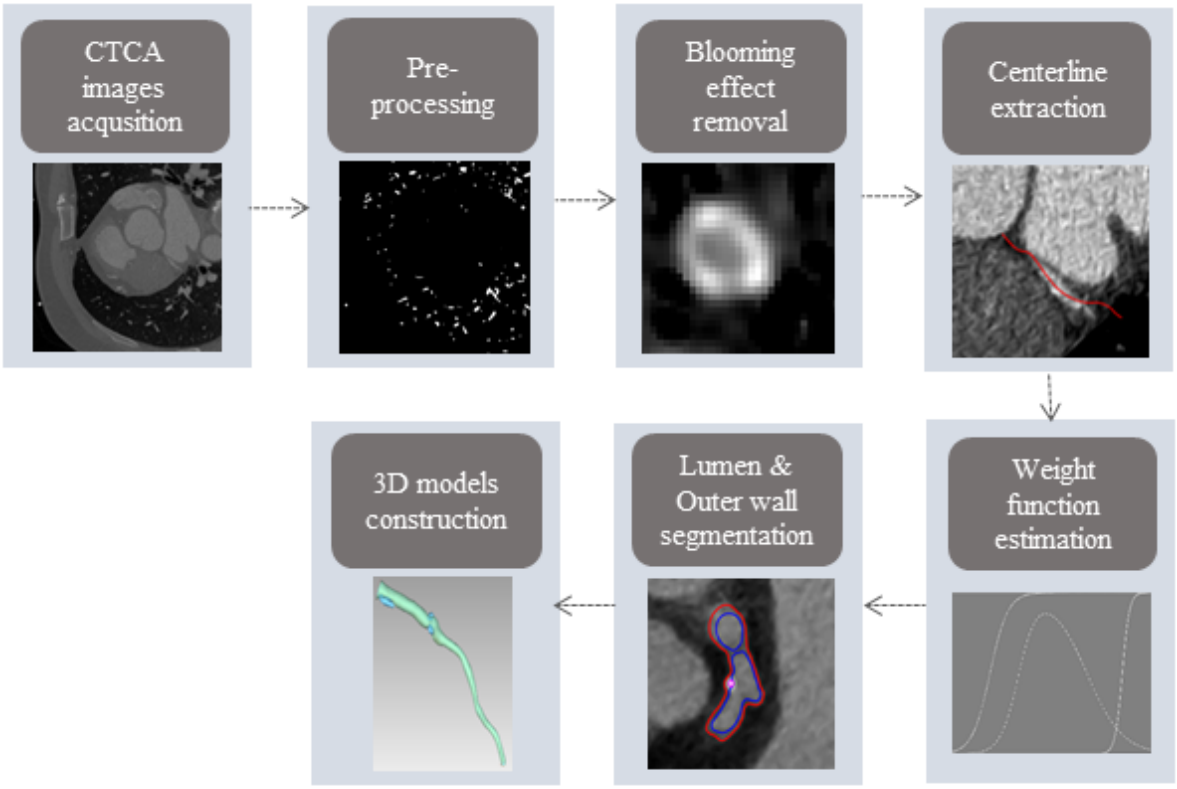

Figure 3.1 The proposed methodology outline [204].

3.2.1 Preprocessing

The image preprocessing step is applied in the axial DICOM acquired slices to remove irrelevant details of the CTA images. A vessel enhancement filter, the Frangi Vesselness filter [207] is implemented to identify tubular structures and limit the ROI to vessel candidate regions. In Figure 3.2, an example of the implementation of the Vesselness filter is shown.

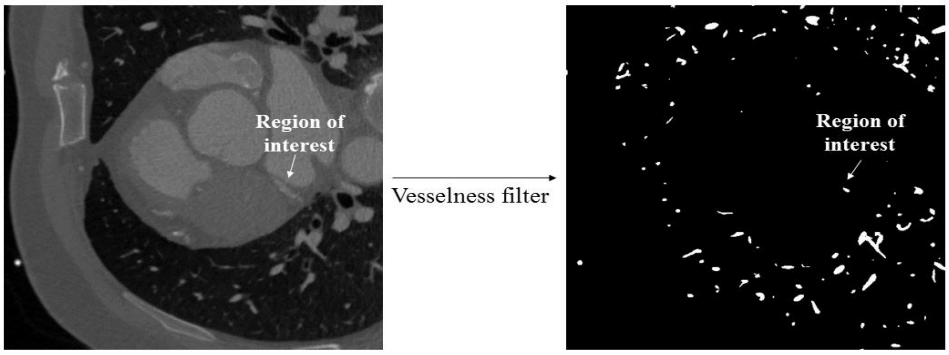

Figure 3.2 An example of the implementation of the Vesselness filter in a CTA image [204].

3.2.2 Blooming effect removal

The blooming effect is a typical CTA image artifact, where small high density objects are illustrated thicker with smeared edges. This artifact affects the visualization and the quantification of small structures, such as the calcifications, as well as, the visualization of metal stents. Thus, in this step we aim to remove the blooming effect, by applying the Blind deconvolution approach [208]. In general, an output CTA image can be modeled by the convolution of the deblurred input image with the point spread function (PSF) of the system, as it is shown in Eq.(3.1):

$$f(x,y) = h(x,y) * g(x,y), \tag{3.1}$$

where $h(x,y)$ corresponds to the PSF of the CT system, $g(x,y)$ is the real input structure and $f(x,y)$ represents the output image. In this approach, since the PSF of each CT system is not defined, we estimate the CT system's PSF, using a Gaussian kernel and subsequently we implement on the high intensities CTA image's regions a deconvolution technique to acquire the deblurred CTA image [209]. The estimated Gaussian kernel is considered as a symmetric Gaussian lowpass filter of size 5 with a standard deviation 0.8.

An example of the blooming effect removal step is illustrated below, in Figure 3.3.

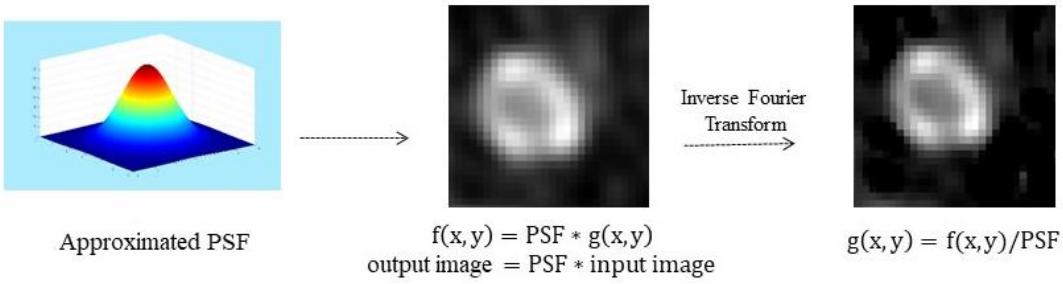

Figure 3.3 An example of the implementation of blooming effect removal technique [204].

3.2.3 Centerline extraction

The centerline is mainly required for creating an initial vessel mask for the vessel segmentation algorithm. However, the centerline extraction stage still remains a challenging task, since the size of the vessels is small and several reconstruction artifacts are observed. In the proposed methodology, a minimum cost path approach is implemented for the centerline extraction, based on Metzt *et al*. [210] approach.

The proposed centerline extraction methodology is quite simple and, therefore easy to implement, since the main requirement is the starting point and the ending point of the vessel to extract the corresponding centerline. The cost function, which is considered for the minimum cost path approach is a combination of the lumen and vessel weight.

Firstly, we extract the image weight based on the vesselness measure (w_{vessel}) [207]. Subsequently, we compute the value of the top 50% of the image intensities, which are higher than 100 HU, considering only the parts of the image, where the w_{vessel} measure is higher than 0. This computed value *ml* is very significant, since it is used for the extraction of the lumen weight. More specifically, the lumen weight is extracted by using a generalized bell-shaped membership function and it is defined as:

$$w_{lumen} = 0.9 \cdot \frac{1}{1+\left|\frac{x-c}{a}\right|^{2b}} + 0.1, \tag{3.2}$$

where $a = 0.02$, b is the minimum value between $ml - l_{thres}$ and 500, and c is the value of $ml + cp_{thres}$. Heuristically, the threshold of the lumen (l_{thres}) and the CP (cp_{thres}) were defined with the value of 80 HU and 400 HU, respectively. More details can be found in the Appendix (Threshold selection).

The considered cost function V for the minimum path approach is a combination of the vessel and the lumen weight and is defined by:

$$V = w_{vessel} \cdot w_{lumen} \tag{3.3}$$

In order to calculate the shortest distance from a list of points to all other pixels in an image volume, a Multistencil Fast Marching Method is implemented based on the approach described in [211]. An example of the above procedure is depicted in Figure 3.4.

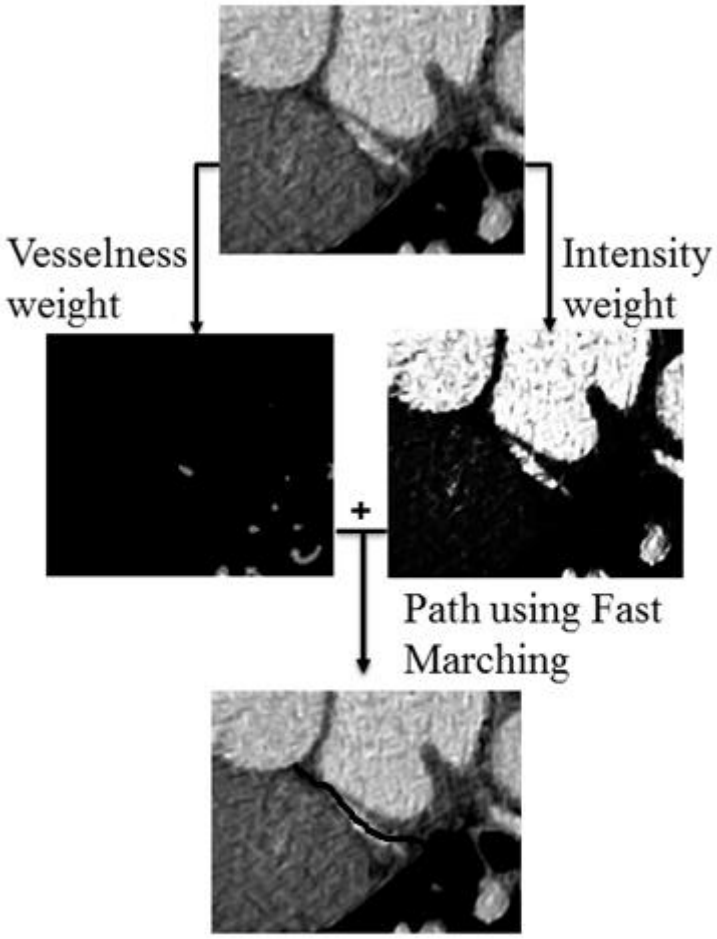

Figure 3.4 Example of a successfully extracted coronary artery centerline using the vesselness/intensity cost function [204].

3.2.4 Estimation of weight function for the inner and outer wall

Similarly to the previous step, in this stage two different membership functions for the lumen and the outer wall are computed, aiming to compensate different protocols for discriminating the lumen and the outer wall. These membership functions are all adapted to the mean vessel intensity across the

centerline, assuming that this corresponds to the mean lumen intensity. More specifically, the mean lumen intensity $\bar{I}_{lumen}$ is calculated, taking into consideration only the pixels of the image, whose intensities are higher than 100 HU and their Euclidean distance from the extracted centerline is less than 5.

For the lumen a generalized bell-shaped membership function is used, whereas for the outer wall a sigmoidal membership functions is implemented, as it is shown in Figure 3.5. The generalized bell-shaped membership function is defined as:

$$g^{bell}(x; a, b, c) = \frac{1}{1+\left|\frac{x-c}{a}\right|^{2b}}, \tag{3.4}$$

whereas the sigmoidal membership function is defined as:

$$g^{sigm}(x; a, b) = \frac{1}{1+e^{-a(x-b)}}, \tag{3.5}$$

where x is the image and a, b, c are the defined parameters.

For the lumen, the membership function is given by:

$$f_{lumen} = (1-\varepsilon) \cdot g^{bell}(x; a_{lumen}, b_{lumen}, c_{lumen}) + \varepsilon, \tag{3.6}$$

where $a_{lumen} = 0.02$, $b_{lumen} = min([max([\bar{I}_{lumen} - l_{thres} \quad 150]) \quad 500]) - 0.01$ and $c_{lumen} = \bar{I}_{lumen} + cp_{thres}$.

The membership function for the outer wall and the plaques is in both cases a sigmoidal function and is given by:

$$f_{outer} = (1-\varepsilon) \cdot g^{sigm}(x; a_{outer}, b_{outer}) + \varepsilon, \tag{3.7}$$

where $a_{outer} = 0.02$ and $b_{outer} = min\big(200, max([\bar{I}_{lumen} - l_{thres} - ncp_{thres} \quad 100])\big)$. The threshold value for the lumen (l_{thres}) as previously stated is 80 HU. The intensity threshold for NCP (ncp_{thres}) is defined by the value of 50 HU. The value of the (ncp_{thres}) is defined heuristically and based on the current literature [212, 213]. More details can be found in the Appendix (Threshold selection).

The parameter ε is a weight of the membership functions and in all cases, it is defined by the value of 0.05. In Table 3.1, we summarize the different values of the parameters of the membership functions for each component.

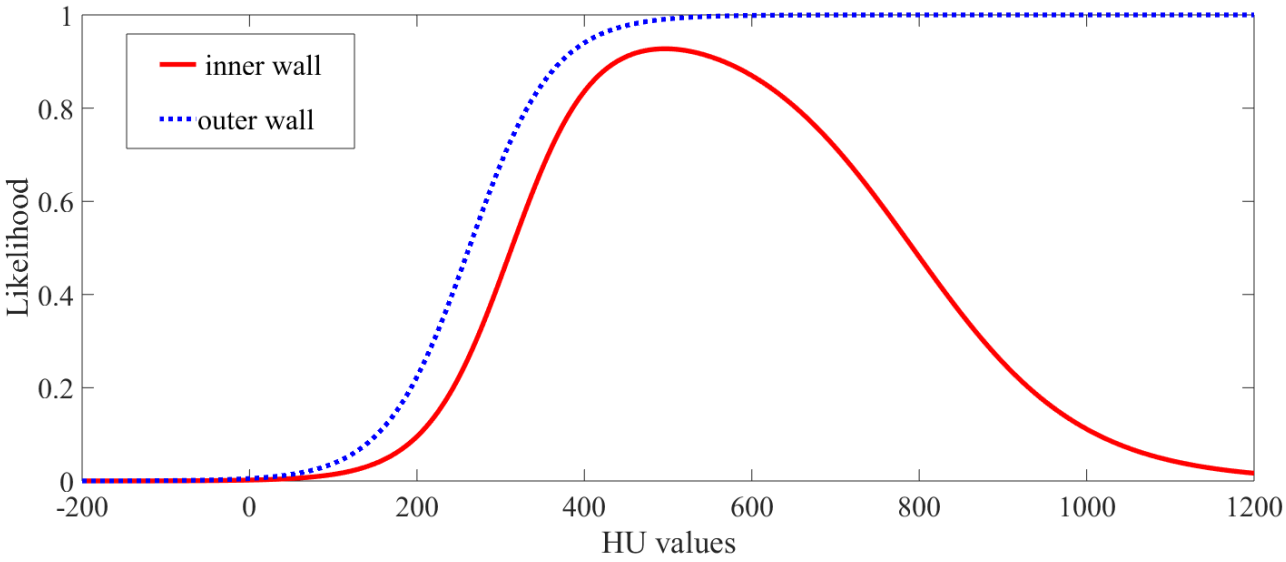

Figure 3.5 Membership functions distributions for lumen and outer wall over HU [204].

Table 3.1 A summary of the parameters of the membership functions for the lumen and the outer wall.

Parameters	a	b	c
Lumen	0.02	$min([max([\bar{I}_{lumen} - l_{thres} \quad 150]) \quad 500]) - 0.01$	$\bar{I}_{lumen} + cp_{thres}$.
Outer wall	0.02	$min(200, max([\bar{I}_{lumen} - l_{thres} - ncp_{thres} \quad 100]))$	-

3.2.5 Lumen segmentation

In this step, an extension of the active contour models [214] is implemented for the lumen segmentation. This approach is based on regional measures and does not depend on any edge definition. In other words, the boundaries of the detected objects are not necessarily defined by the gradient. The main improvement of our lumen level set segmentation approach is that it incorporates a prior shape [215], aiming to segment an object whose shape is similar to the given prior shape which is independent of translation, scaling and rotation, from a background where there are several objects. For the lumen segmentation, the prior shape is a tubular mask across centerline with a small radius.

Update of lumen intensities

Our purpose is to implement the membership function in the part of the CTA image, which is near to the extracted centerline. Thus, first, in order to update the membership function for the lumen, we calculate the Euclidean distance transform only of the pixels around the extracted centerline. This estimated distance d_1 limits the ROI, since based on its value, we considered only the pixels whose distance transform is lower than $^4/_{p_x}$, where p_x is the pixel spacing of the DICOM image and 4 is an approximate value for the radius of the lumen. In other words, we assume that: $f_{lumen} = 0 \rightarrow pixels: d_1 > {}^4/_{p_x}$ and in the same manner $f_{outer} = 0 \rightarrow pixels: d_1 > {}^4/_{p_x}$. For the lumen pixels, the updated membership function is given by:

$$f_1 = f_{lumen} \cdot g^{sigm}(d_1; a_{2,lumen}, b_{2,lumen}), \tag{3.8}$$

where $a_{2,lumen} = -0.5$ and $b_{2,lumen} = {}^2/_{p_x}$, whereas for the outer wall plaques pixels, the updated membership function is given by:

$$f_{2,outer} = f_{outer} \cdot g^{sigm}(d_1; a_{2,outer}, b_{2,outer}), \tag{3.9}$$

where $a_{2,outer} = -0.5$ and $b_{2,outer} = {}^{2.5}/_{p_x}$.

Approximation of an initial binary image

For the implementation of the Level Set method approach, an approximation of an initial image-shape φ is required. This image is a binary image, which includes 0's as background pixels and 1's as foreground pixels. The intensity threshold value for the estimation of the initial image is $^{w_i}/_2$. Thus, the pixel value of initial image is 1, when $f_1 \cdot w_i$ is higher than $^{w_i}/_2$, whereas it is 0, when $f_1 \cdot w_i$ is lower than $^{w_i}/_2$. The parameter w_i is an estimated weight to multiply the probability in the level set method and it is defined 1000.

Calculation of the speed function

The Level set methods have been widely used in the field of machine vision for segmentation problems, since they are used for the modelling of evolving curves or surfaces. The basic idea behind the level set approaches is the presentation of the interface of a surface, using a higher dimensional function, which is called the level set function. This means that the 2D curve could be

described by the 3D level set function, where the additional dimension t represents the time. In the proposed methodology, the detected curve $C \in \Omega$ is represented by the zero level set of a Lipschitz function , such that:

$$\begin{aligned} C &= \partial\omega = \{(x,y,t) \in \Omega : \varphi(x,y,t) = 0\}, \\ inside(C) &= \omega = \{(x,y,t) \in \Omega : \varphi(x,y,t) > 0\}, \\ inside(C) &= \omega = \{(x,y,t) \in \Omega : \varphi(x,y,t) < 0\}, \end{aligned} \tag{3.10}$$

where x, y are the spatial coordinates of the 2D image, t is the dimension of time and $\omega \subset \Omega$.

In this stage, a level set based variational method using prior shapes is implemented for the lumen segmentation and our aim is to incorporate shape priors into the Chan-Vese's model for segmentation. This approach is based on Cremers *et al*. [216] and Chan *et al.* [215] studies, in which besides the basic level set segmentation function φ, a shape function ψ and a labelling function L are introduced. The key idea of this methodology is that the defined prior shape is compared with the region where both the level set function φ and the labelling function L are positive.

In the presented methodology, the speed function that is impemented to evolve the level set curve is based on Chan *et al.* approach [215]. The defined speed function is a Chan-Vese energy function, combined with prior shapes and with a labelling function and is given by:

$$E(\varphi, \psi, L, c_1, c_2) = E_{CV} + E_{shape} + E_{\psi}, \tag{3.11}$$

where E_{CV} is the Chan-Vese energy funtion, E_{shape} is a shape comparison term and E_{ψ} is a labelling term.

The Chan-Vese energy funtion [214] is widely used in medical image segmentation approaches and it is defined as:

$$E_{CV}(c_1, c_2, C) = \int_{inside(C)} (u(x,y) - c_1)^2 dxdy + \int_{outside(C)} (u(x,y) - c_2)^2 dxdy, \tag{3.12}$$

Disretizing the above speed function equation and writing it as a pixelwise function, it gives

$$E_{CV}(x,y) = (u(x,y) - c_1)^2 - (u(x,y) - c_2)^2, \tag{3.13}$$

where u is the image, x, y are the spatial coordinates of the 2D image, C is the segmentation curve and c_1, c_2 are the average greyscale intensity values inside and outside of C, respectively.

The shape comparison term is defined as:

$$E_{shape}(\varphi, L, \psi) = \int_{\Omega} \left(H(\varphi)H(L) - H(\psi)\right)^2 dxdy, \tag{3.14}$$

where H is the Heaviside function and $H(\varphi)H(L)$ represents the intersection of $\varphi > 0$ and $L > 0$. The labelling function E_ψ is a term that indicates if the lumen is segmented successfully, since then this term will be small and it is given by:

$$E_\psi(\psi, c_1, c_2) = \int_\Omega \left[(u - c_1)^2 H(\psi) + (u - c_2)^2 \big(1 - H(\psi)\big)\right] dxdy, \quad (3.15)$$

Sparse Field Algorithm implementation

As previously mentioned, the key idea of Level Set approaches is that only the area, where $\varphi(x, y) \approx 0$ is important to accurately represent the curve. In this approach, a sparse field algorithm approach, proposed by Whitaker *et al*. [217] is implemented to maintain an accurate and minimal representation of φ. Once the initial φ is defined, the algorithm returns fully initialized arrays for the label map and for an updated φ. Both arrays are of the same size and the label map records the status of each point. Once the energy function has been computed for the part of the image which is around the centerline, the level sets may be deformed in order to minimize some of the energy function. Thus, a sparse field approach is implemented twice to update φ near the zero level set. First, based on the initial φ and on a positive factor α, which controls the speed and curvature of the level set, the algorithm results in a new φ_{lumen}. Subsequently, based on the resulted φ_{lumen}, the algorithm is implemented again, using a higher value of factor α , to achieve a smoother lumen shape. In the first step, a lower value of factor α was applied, in order to provide fast segmentation in the whole image. In the second implementation of sparse field algorithm, a higher value of factor α is selected and applied only at the ROI in order to provide smooth and accurate segmented objects which depict only the coronary arteries.In the first step, the factor α is defined 0.1, whereas in the second step factor α is defined 0.6, as it is shown in Figure 3.6. The value of α factor affects only the speed of the segmentation and it does not affect the quality of the segmentation. The number of iterations for the sparse field algorithm impementation is set to 200.

3.2.6 Outer wall segmentation

Similarly to the previous step, a Level Set model is implemented for the outer wall segmentation. However, in this stage the initial φ is based on the lumen shape, as segmented in the previous stage and on the updated values of outer wall intensities. More specifically, the required φ of this stage is also a binary image, which includes 1's for the pixels, where either the $w_i \cdot f_{2,outer} + w_i \cdot f_{2,plaque}$ is

higher than $^{w_i}/_2$ or the segmented lumen (φ_{lumen}) has a positive value. Furthermore, the energy function for this stage, is calculated only for the pixels, where φ_{lumen} is higher than -0.1. The sparse field algorithm implementation process for the outer wall is exactly the same as in the previous stage described.

Figure 3.6 Lumen Segmentation example, a) acquired image, b) a factor 0.1, c) a factor 0.6 [204].

3.2.7 3D surface construction

In this stage, an isosurface of data from each different extracted φ array is computed to construct the mesh surfaces. In this step the algorithm marching cubes, proposed by Lorensen and Cline [218], is implemented to construct 3D surfaces for the lumen and the outer wall. Marching cubes extracts a polygonal mesh of an isosurface from a 3D discrete scalar field, by proceeding through it. A triangulation approach is implemented, and connecting the detected border points of each CTA image, 3D models are constructed. An example of 3D reconstruction of coronary arteries in the baseline and follow-up for one case is shown in Figure 3.7, below.

In Figure 3.8 below, we demonstrate (a) the 3D reconstruction of a carotid artery bifurcation, (b) the 2D CTA slice without segmentation and (c) the 2D segmentation of the inner wall and outer wall.

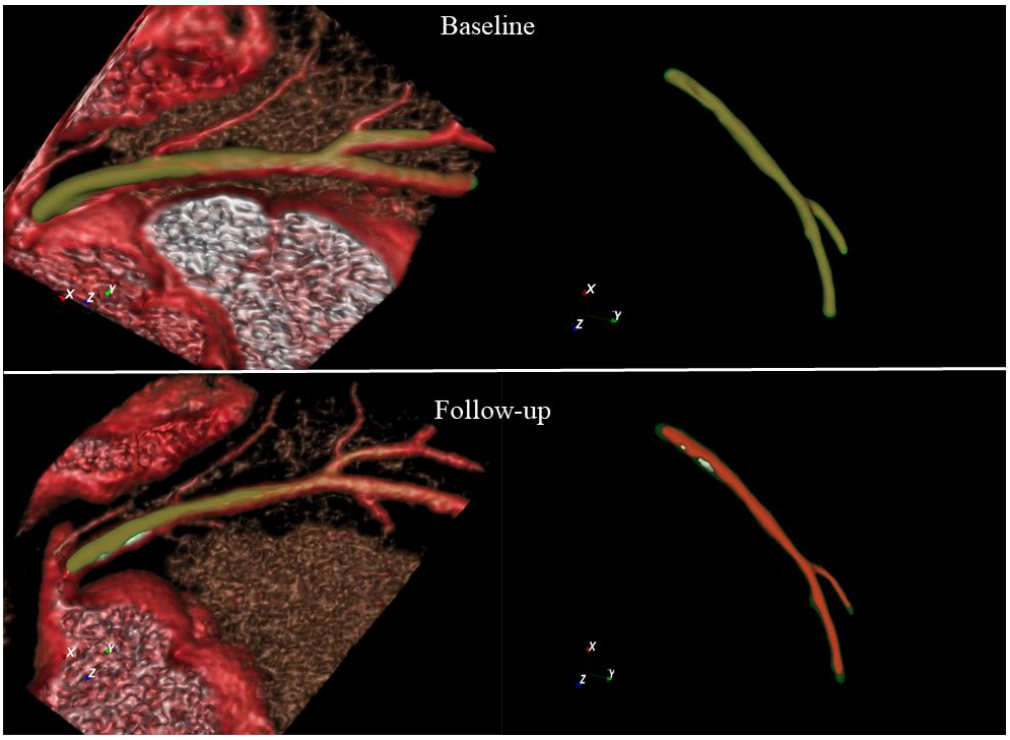

Figure 3.7 Example of 3D reconstruction of coronary artery in baseline and follow-up [204].

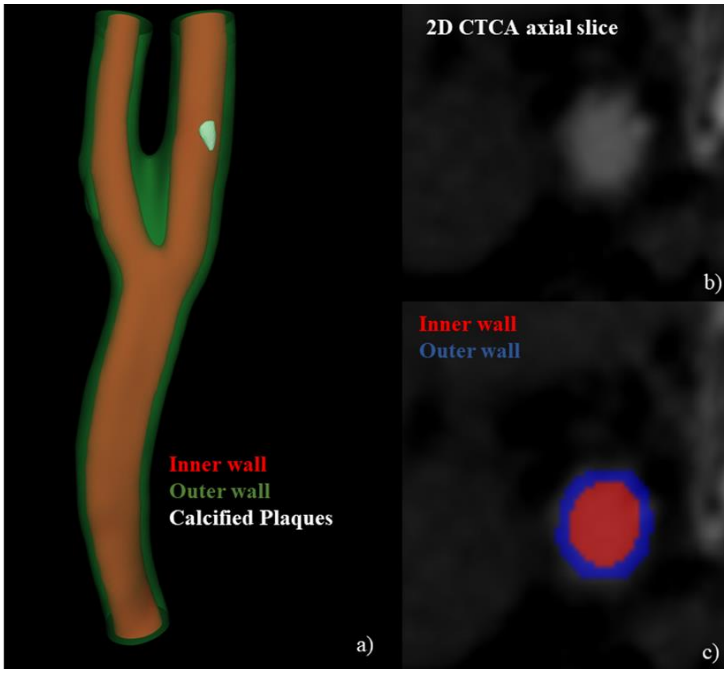

Figure 3.8 (a) 3D reconstruction of a carotid artery bifurcation, (b) the 2D CTA slice without segmentation and (c) the 2D segmentation of the inner wall and outer wall [206].

3.3 Methodology for atherosclerotic plaque characterization

The proposed methodology is based on the accurate detection of the lumen and the outer wall of coronary and carotid vessels, as described previously. The detection and characterization of atherosclerotic plaques is implemented in the ROI, which is inside the outer wall and outside the lumen, as it is shown in Figure 3.9, below.

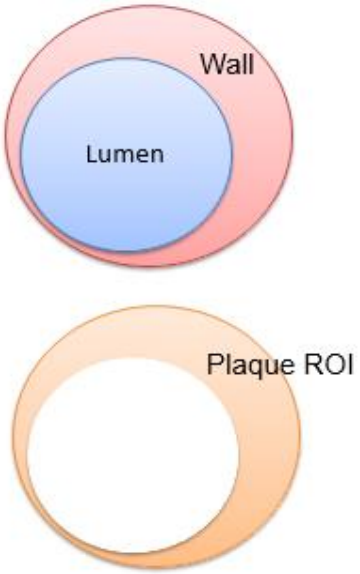

Figure 3.9 Region of Interest (ROI) [204].

The characterization of atherosclerotic plaque into CP and NCP requires the accurate detection of the lumen and the outer wall. The proposed methodology includes four steps: In the first step a weight function for CP is estimated to approximate the potential HU values ranges for the CP. In the second step, an extended approach of active contour models without edges is implemented to detect the CP. In the third step, an adaptive intensity range is extracted based on the mean lumen intensity for the detection of NCP. Finally, in the fourth step, the 3D geometrical models for the inner and the outer wall, the CP and the NCP are constructed, employing the fast marching approach.

3.3.1 Estimation of weight function for the CP

In this step, similarly to the lumen and the outer wall, a membership function for the the CP is estimated, aiming to compensate a different protocol for discriminating the lumen CP into the ROI.

Contrary to the CP intensity ranges, the NCP intensity ranges are close to the outer wall HU values and as a result, a membership function cannot be extracted for the NCP detection. The utilized membership function for the CP, illustrated in Figure 3.10, is a sigmoidal function, adapted to each CTA image, since it is based on $\bar{I}_{lumen}$. The $\bar{I}_{lumen}$ value as stated previously, corresponds to the mean intensity values of the top of half image intensities, which are higher than 100 HU, taking into consideration only the image's pixels which correspond to the initial vessel mask, extracted by the Frangi vesselness filter [207].

The membership function for the CP (f_{cp}) is a sigmoidal function which is given by:

$$g^{sigm}(x;a,b)=\frac{1}{1+e^{-a(x-b)}}, \tag{3.16}$$

where x is the image and a, b are the defined parameters.

In our case the f_{cp} is given by:

$$f_{cp} = (1-\varepsilon)\cdot g^{sigm}(x;a_{cp},b_{cp})+\varepsilon, \tag{3.17}$$

where $a_{cp}=0.05,\ b_{cp}=\bar{I}_{lumen}+cp_{thres}$. The cp_{thres} is defined by the value of 400 HU. As stated previously, more details can be found in Appendix (Threshold selection).

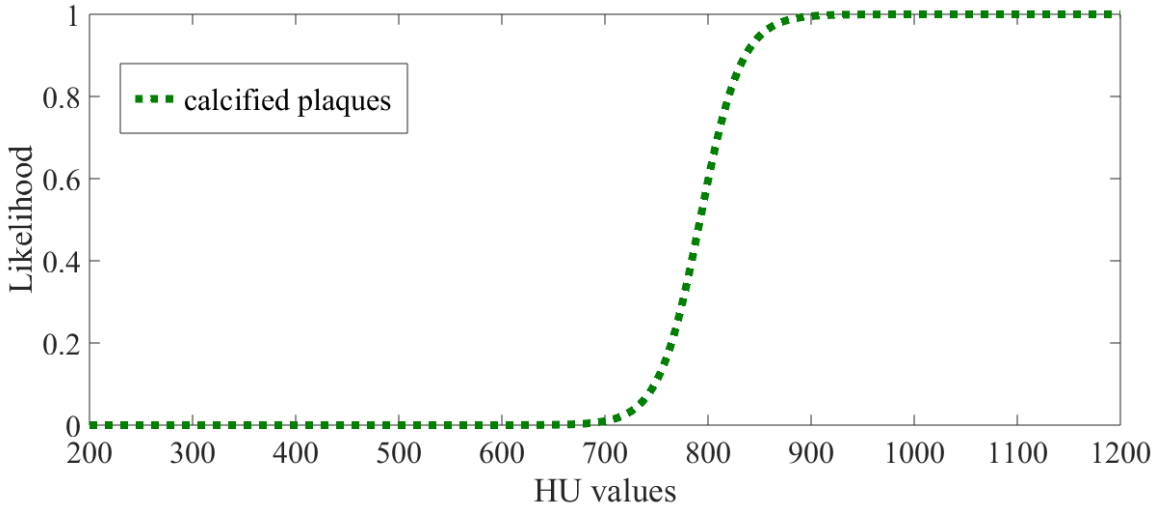

Figure 3.10 Membership functions distribution for CP over HU [204].

3.3.2 CP segmentation

In this stage, the Level Set method [214, 217, 219, 220] is applied in the ROI of the outer wall. The main idea of CP segmentation is exactly the same with this proposed in the previous chapter for the

segmentation of the inner and the outer wall. The initial φ for the CP segmentation is based only on the updated plaques intensity function. In other words, the initial φ is also a binary image, which includes 1's for the pixels, where $w_i \cdot f_{cp}$ is larger than $w_i/2$. For this phase, only a sparse field algorithm implementation is required, as the segmented objects are relatively smaller and the a factor is 0.5. An example of the segmentation procedure is demonstrated in Figure 3.11.

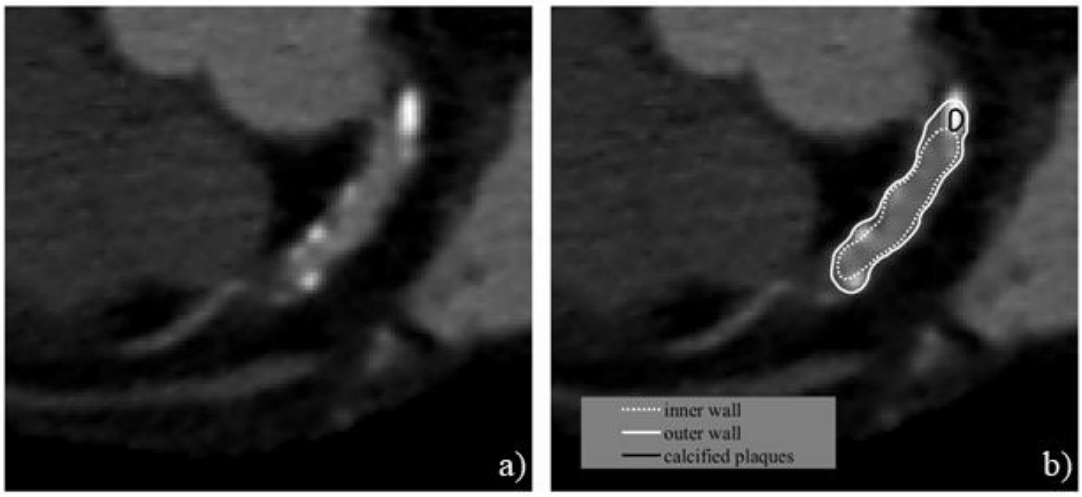

Figure 3.11 An example of the segmentation procedure: a) the acquired image, b) inner wall, outer wall and CP [205].

In addition to this, we demonstrate in Figure 3.12 below, an example of the centreline extraction, in combination with the lumen and CP segmentation, across the length of the centreline. The segmentation procedure is indicated not only across all the length of the vessel centreline, but also in a 2D CTA slice. Moreover, in Figure 3.12, we illustrate the DS, by providing the exact lumen and outer wall surfaces in mm^2 per 0.5mm.

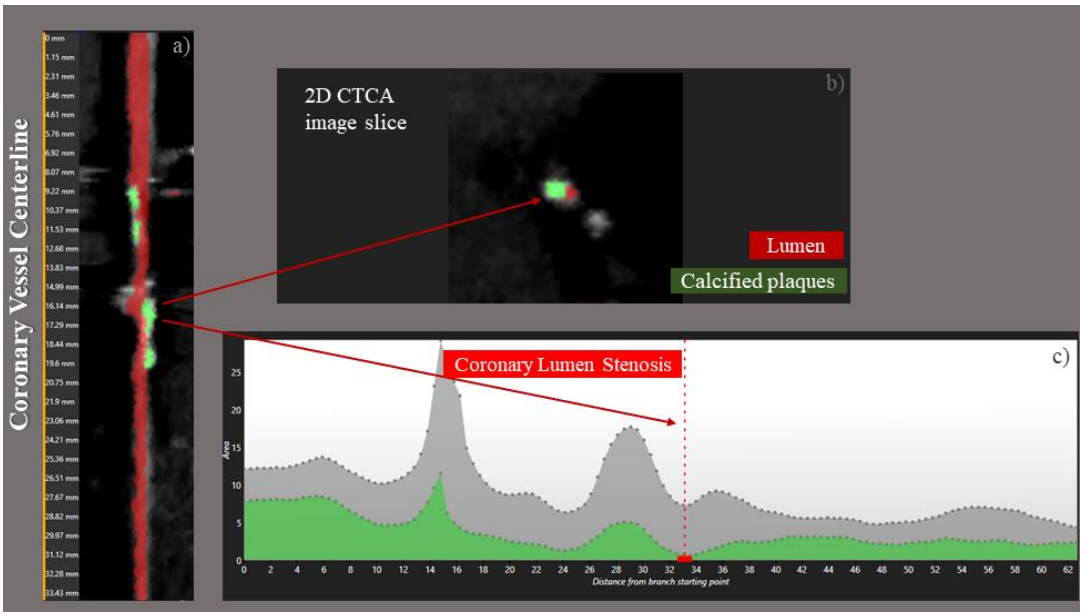

Figure 3.12 a) Lumen (red) segmentation and CP (green) segmentation across the vessel centerline, b) Lumen (red) and CP (green) segmentation in a 2D CTA slice, c) Lumen (green) and outer wall (grey) suface per 0.5 mm.

3.3.3 NCP segmentation

Contrary to the CP, the NCP usually have a lower intensity, which makes the detection of the NCP a challenging problem. Different studies have indicated that the luminal intensity is clearly related to the intensity inside the atherosclerotic plaque [221]. Thus, in this study due to the lower intensities of NCP, the NCP cannot be successfully identified by a level set approach, since the intensity ranges are close to the intensity ranges of the outer wall. The identification of NCP is achieved by a dynamic thresholding technique. Thresholding is a conventional segmentation technique, in which the image pixels are partitioned depending on their intensity value, using an appropriate threshold value. However, in the case of dynamic thresholding segmentation, the threshold value is not a constant value and relies on the mean intensity values of pixels, which correspond to the lumen region.

The main idea of this approach is to adapt the threshold values of NCP into the luminal intensity, by defining a critical value threshold. More specifically the range of NCP HU values is extracted based on the mean luminal intensity (*ml)*. The *ml* is computed after the implementation of the Frangi

vesselness filter. The *ml* value is the mean value of the highest of half image intensities, which are higher than 100 HU, considering only the parts of the CTA image, which are potential coronary vessels. After the definition of the *ml*, the range of NCP intensities is defined from 100 HU to the *ml* value. In addition to this, the aforementioned segmentation approach is implemented in the ROI, which is located inside the segmented outer wall and outside the inner wall. An example of the segmentation procedure is demonstrated in Figure 3.13.

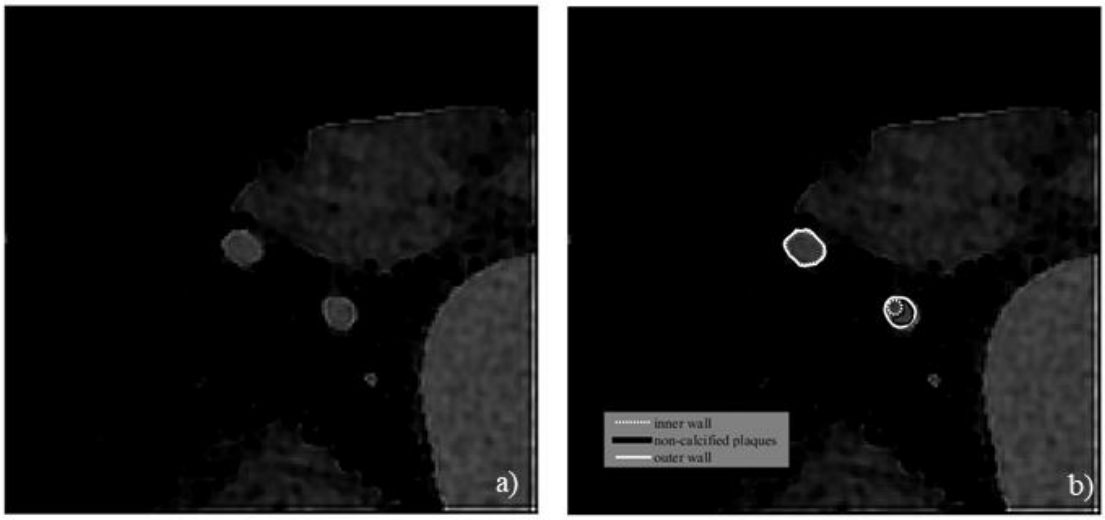

Figure 3.13 An example of the segmentation procedure: a) the acquired image, b) inner wall, outer wall and NCP plaques [205].

3.3.4 3D surface reconstruction

The Marching cubes algorithm, which has been proposed by Lorensen and Cline [222], is applied to construct the 3D surfaces for the the CP and the NCP. More specifically, a triangle topology is defined by constant density surfaces, applying the divide- and- conquer approach to create inter-slice connectivity. In Figure 3.14, a 3D reconstructed geometry of the lumen, the outer wall and the CP back-projected in volume rendering is illustrated. Moreover, in Figure 3.15 and Figure 3.16, lumen, outer wall and both types of atherosclerotic plaques (CP and NCP) are shown for coronary arterial tree and coronary artery bifurcation, respectively.

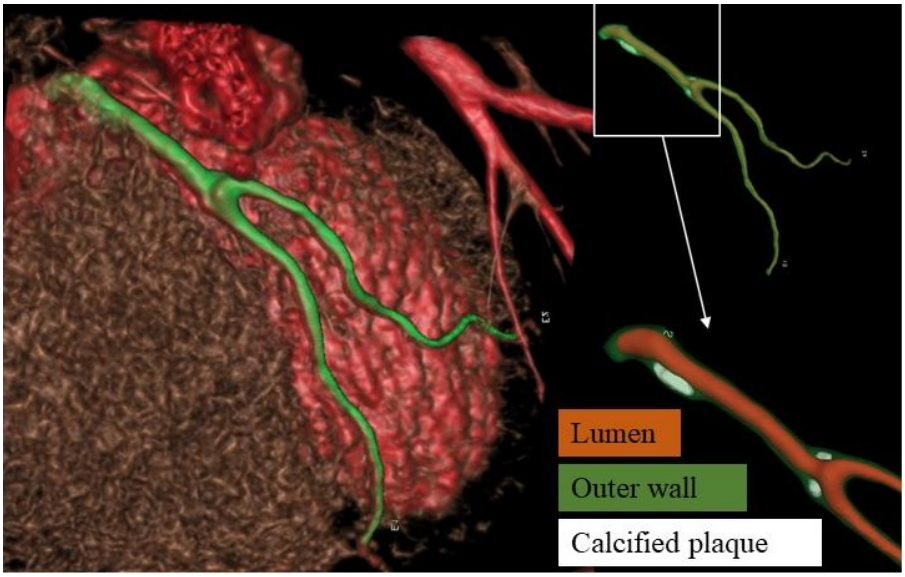

Figure 3.14 3D reconstructed models for the lumen, the outer wall and the CP [204].

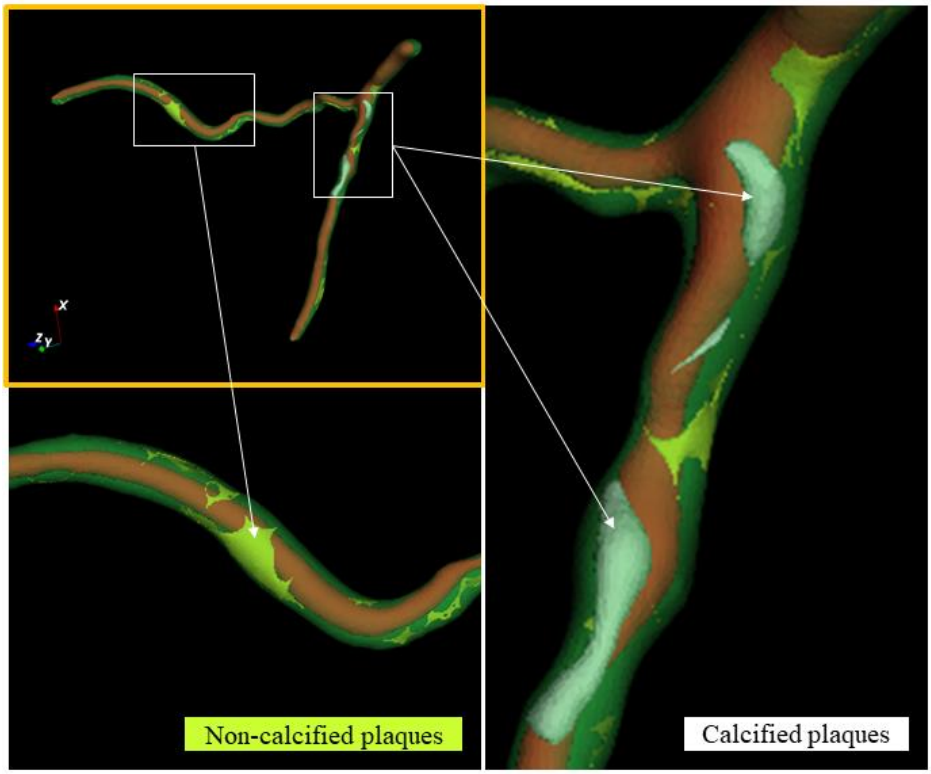

Figure 3.15 3D reconstructed models of a coronary arterial bifurcation for the lumen, the outer wall, the CP and the NCP.

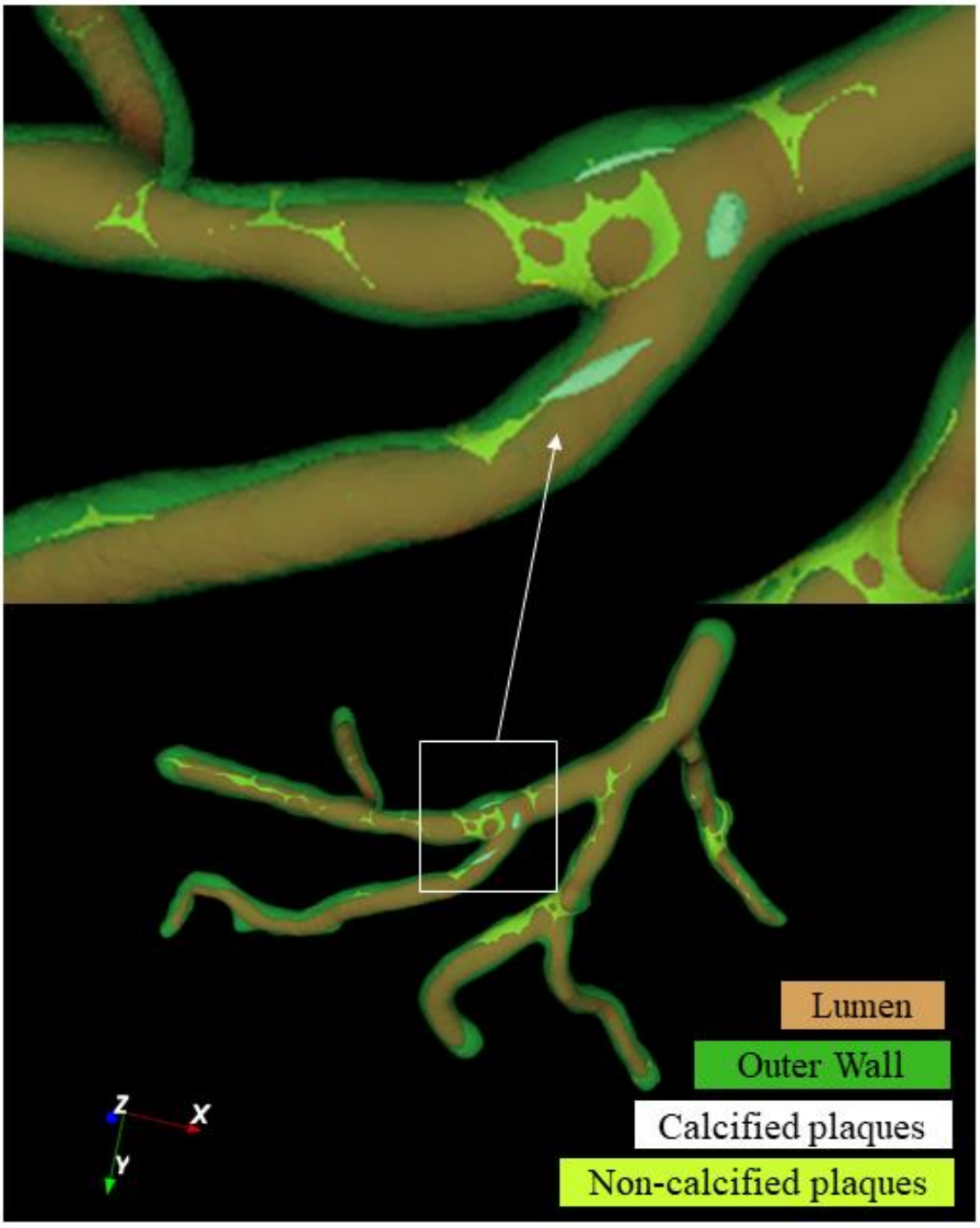

Figure 3.16 3D reconstructed models of a coronary arterial tree for the lumen, the outer wall, the CP and the NCP.

3.4 Dataset

3.4.1 Dataset for 3D coronary artery reconstruction

3.4.1.1 Dataset for inner wall and outer wall segmentation

The data used for the validation of the inner and outer wall segmentaion, were acquired from 12 patients, who underwent CTA imaging for clinical purposes. Our validation dataset consists of twelve coronary arteries, deriving from six different medical centers. Two arteries were completely healthy (no stenosis present), nine arteries had an intermediate stenosis (seven had a 30%-50% DS and two had a 50%-70% DS) and one artery was fully occluded (>90% DS). Five arteries were scanned with a 64-slice Dual Source Siemens SOMATOM Definition Flash® CT scanner, two arteries were scanned with a Philips Brilliance 64 CT Scanner® and the remaining five arteries were scanned with a 64-slice General Electric Medical Systems Discovery PET-CT 690® scanner.

IVUS and biplane X-ray angiography were acquired from eight patients who underwent coronary catheterization. CTA imaging was also acquired for the same patients. Registration between the imaging modalities was performed using major landmarks common for all modalities, such as the bifurcations or large CP.

3.4.1.2 Dataset for CP and NCP segmentation

Dataset to compare with VH-IVUS modality

The proposed methodology was validated using imaging data from 18 patients, who underwent CA, IVUS and CTA imaging for clinical purposes. All of the patients enrolled in the study had a significant luminal stenosis (>50%), which was initially confirmed by the CTA imaging and further investigated by VH-IVUS. This imaging dataset was provided by the Heart Institute, University of Sao Paulo, São Paulo, Brazil and the CTA images acquisition was performed using a 64slice MDCT scanner (Aquillion 64TM, Toshiba Medical Systems, Japan), while the IVUS examination was performed using a 20 MHz electronic multi-array 2.9 F catheter (Eagle Eye®, Volcano Corporation Inc) connected to a dedicated console (InVision Gold®, Volcano Corporation Inc., San Diego, CA, USA).

Dataset to compare with manual expert's annotations

The imaging data were acquired from 27 patients, who underwent CTA imaging for clinical purposes and one of their coronary arteries was selected for manual expert's analysis. 8 coronary arteries had no significant stenosis (<30%), 8 had an intermediate stenosis (30%-50%), 8 had a significant stenosis (50%70%) and 3 coronary arteries were fully occluded with a stenosis higher than 70%. In addition to this, the utilized imaging dataset was derived by six different medical centers and the coronary arteries were scanned with different CTA scanners (64-slice Dual Source Siemens SOMATOM Definition Flash® CT scanner, Philips Brilliance 64 CT Scanner®, 64-slice General Electric Medical Systems Discovery PET-CT 690®).

3.4.2 Dataset for 3D carotid artery reconstruction

A total of 16 patients (4 females and 12 males) with a mean age of 76 years (range: 55–86 years) were used for the validation of the carotid artery reconstruction. 14 patients have no previous event, whereas 2 of them have a following stroke event. CTA was performed on either a Philips ICT (256 rows) or an Iqon (64rows) with the same protocol. Bolus tracking was used with the ROI placed in the aortic arch. 80 ml of Imeron 350 was administered with a flow rate of 4ml/s followed by a 50ml saline chaser. Collimation and reconstruction matrix was 0.9mm

3.5 Results

3.5.1 Results for coronary artery reconstruction

3.5.1.1 Results for inner wall and outer wall segmentation

Artery reconstruction using manual annotations

Reconstruction of arteries was successfully obtained in 12 patients, 2 RCA, 8 LAD and 2 LCX, by the proposed methodology in a semi-automatic manner, whereas an expert radiologist manually annotated in the corresponding segments the lumen and the outer wall.

The aim of the validation process is to assess the accuracy and the quality of the proposed methodology. In this study, our purpose is to validate our methodology, using as gold standard a medical expert's annotations. It is known that the lack of expert-annotated datasets remains one of the main challenges in medical image processing [223]. Generating high-quality expert-derived

annotations in CTA images is time-consuming and requires a specialized in CTA imaging field medical expert.

Artery and CP reconstruction using IVUS

Reconstructions of arteries based on IVUS modality is performed using the study proposed by Bourantas *et al.* [224]. This approach combines the IVUS and X-ray angiography and based on the 3D luminal centerline, derived from two angiographic projections, it places the lumen and the media-adventitial borders detected by IVUS frames onto the centerline.

Metrics for evaluating 3D image segmentation

The lumen and the outer wall reconstructed by the two different modalities were compared using as metrics for the 3D image segmentation the DICE and the HD [225]. Additional metrics were used for the validation of the proposed methodology against the IVUS based reconstructed segments. For this purpose we used for comparison two types of DS (DS1, DS2), the PB, the MLA and the MLD. In more detail, the DICE, also called overlap index, is an overlap based metric, which is widely used to compare directly automatic and expert's annotations segmentations. It is considered as a statistical validation metric to evaluate the performance of both the reproducibility of manual segmentations and the spatial overlap accuracy of automated segmentation [226]. DICE is defined as:

$$DICE = \frac{2|S_g^1 \cap S_t^1|}{|S_g^1| + |S_t^1|}, \tag{3.18}$$

where S_t^1 presents the automatic segmentation and S_g^1 presents the expert's annotations.

HD is a spatial distance based metric and is commonly used in the evaluation of image segmentation as dissimilarity measure. HD between two finite point sets A and B is defined as:

$$HD(A,B) = max\big(h(A,B), h(B,A)\big), \tag{3.19}$$

where $h(A,B) = \max_{a \in A} \min_{b \in B} \|a - b\|$.

The 3D models reconstructed by CTA and IVUS were compared using as validation metrics the DS1, the DS2, the PB, the MLA and the MLD. Each 3D model was sliced per 0.5 mm, to estimate the validation metrics. Based on the literature, two different ways of DS estimation were used [227]. More specifically, DS1 and DS2 are given by:

$$DS1 = \frac{B-A}{B} \cdot 100\%, \tag{3.20}$$

$$DS2 = \frac{C-A}{C} \cdot 100\%, \tag{3.21}$$

where A is the luminal diameter at the site of maximal narrowing, B is the diameter of the normal distal coronary artery beyond the bulb where the artery walls are intersected and C is the diameter of estimated original width of the coronary artery at the site of maximal narrowing, as it is shown in Figure 3.17 below.

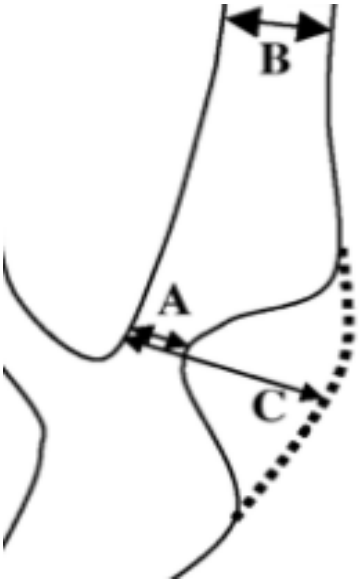

Figure 3.17 Diagram of a coronary stenosis.

The PB is extracted based both on the inner and the outer wall contours areas per 0.5 mm and is given by:

$$PB = \frac{outer\_wall\_area - inner\_wall\_area}{outer\_wall\_area}, \tag{3.22}$$

Comparison using manual annotations

For both quantitative evaluation and comparison purposes, we present in Table 3.2 the DICE and HD distributions obtained by the comparison of the proposed segmentation methodology and the expert's manual segmentation. The HD demonstrates the degree of resemblance between the two models which are superimposed on one another. Thus, the lower HD implies better segmentation, while a higher DICE implies higher accuracy of the segmentation.

The validation process indicates a good agreement, since the mean value of DICE is 0.749, while the mean value of HD is 1.746. In Figure 3.18, an example of the lumen 3D models reconstructed by the proposed methodology and by the expert's annotation, is shown. The two reconstructed arteries indicate a similar geometry. The semi-automated reconstructed artery is smoother than the manually segmented, since a pixel by pixel segmentation may not result in a smooth shape.

Table 3.2 Segmentation Validation metrics for CTA images.

Cases	Arteries	DICE	HD
#1	LAD	0.847	0.837
#2	LAD	0.675	1.910
#3	LAD	0.777	0.860
#4	LAD	0.674	2.423
#5	LAD	0.759	1.589
#6	LAD	0.810	1.510
#7	LAD	0.574	2.671
#8	LAD	0.751	1.800
#9	LCX	0.847	1.296
#10	LCX	0.722	1.830
#11	RCA	0.781	2.341
#12	RCA	0.778	1.880

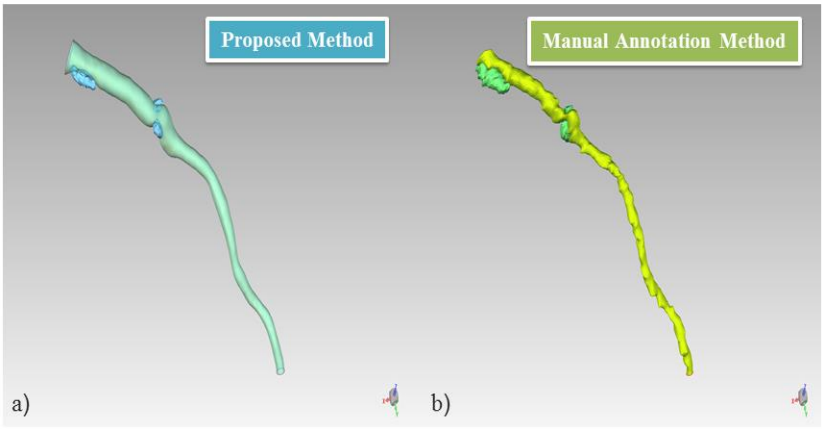

Figure 3.18 Lumen and CP objects detected by a) the proposed methodology, and b) the medical expert annotation [204].

Comparison using IVUS findings

As far as the comparison with IVUS process is concerned, we demonstrate in Table 3.3 the values of the comparison metrics. The metrics of the 3D models derived by our methodology are correlated

with those derived by the IVUS findings, and the Bland altman and correlation plots for DS1, DS2, MLA, MLD, and PB are demonstrated in Figure 3.19, Figure 3.20, Figure 3.21, Figure 3.22 and Figure 3.23, respectively. It is clear, that the correlation between the two methodologies is statistically significant for all the reconstructed cases.

Table 3.3 Comparison of geometrical metrics derived by the presented methodology with those derived by an IVUS based approach.

Cases	**Method**	**DS1(%)**	**DS2(%)**	**PB(/0.5mm)**	**MLA(mm^2)**	**MLD(mm)**
#1	CT	0.29	0.53	0.53±0.09	3.73	1.10
	IVUS	0.43	0.47	0.54±0.17	3.72	1.11
#2	CT	0.17	0.33	0.45±0.09	7.2	1.49
	IVUS	0.27	0.37	0.5±0.07	8.37	1.64
#3	CT	0.64	0.55	0.45±0.11	0.95	0.75
	IVUS	0.54	0.46	0.38±0.2	3.81	1.15
#4	CT	0.27	0.33	0.5±0.07	5.46	1.32
	IVUS	0.17	0.33	0.42±0.2	7.45	1.62
#5	CT	0.54	0.56	0.55±0.09	0.83	0.78
	IVUS	0.5	0.62	0.51±0.08	3.18	1.03
#6	CT	0.34	0.48	0.46±0.09	5.14	1.28
	IVUS	0.24	0.33	0.54±0.09	4.28	1.19
#7	CT	0.54	0.72	0.48±0.08	3.05	0.87
	IVUS	0.5	0.62	0.52±0.09	4.79	1.24
#8	CT	0.43	0.42	0.36±0.09	1.47	0.79
	IVUS	0.3	0.51	0.26±0.22	2.51	0.9

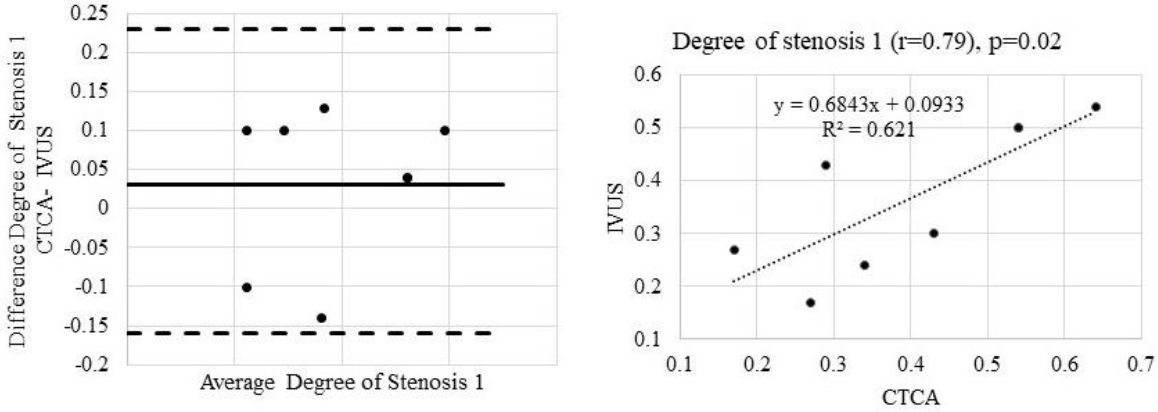

Figure 3.19 Bland Altman and Correlation plots for the DS1.

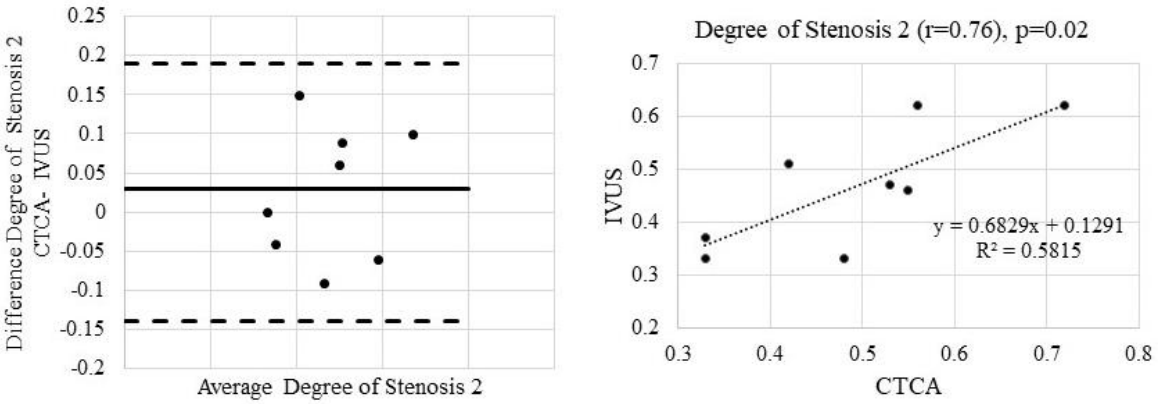

Figure 3.20 Bland Altman and Correlation plots for the DS2.

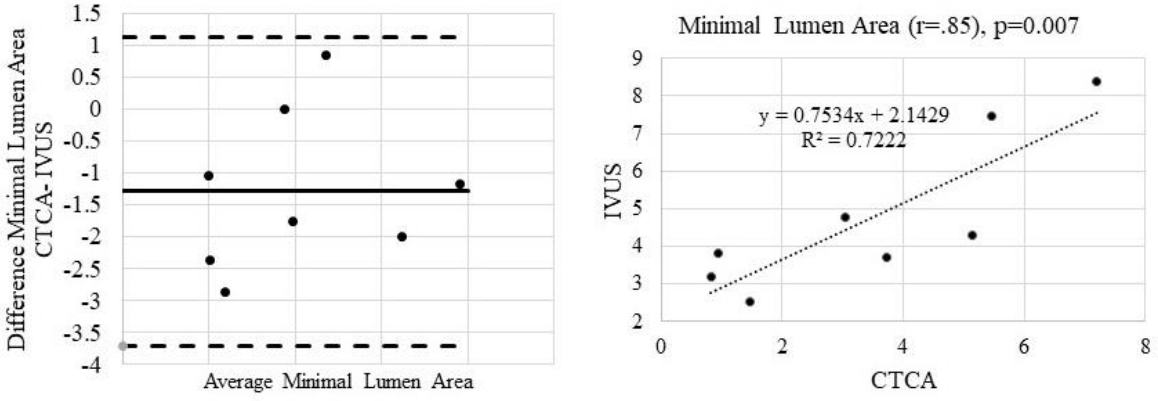

Figure 3.21 Bland Altman and Correlation plots for the MLA.

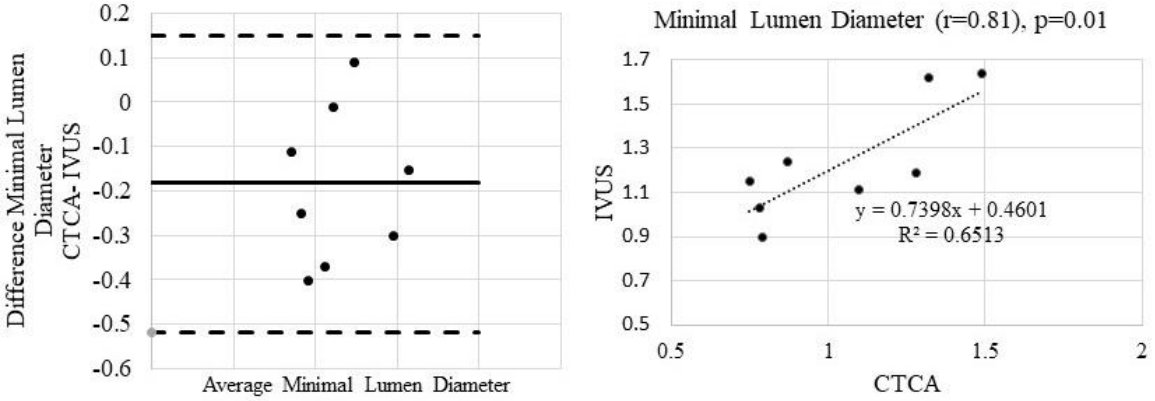

Figure 3.22 Bland Altman and Correlation plots for the MLD.

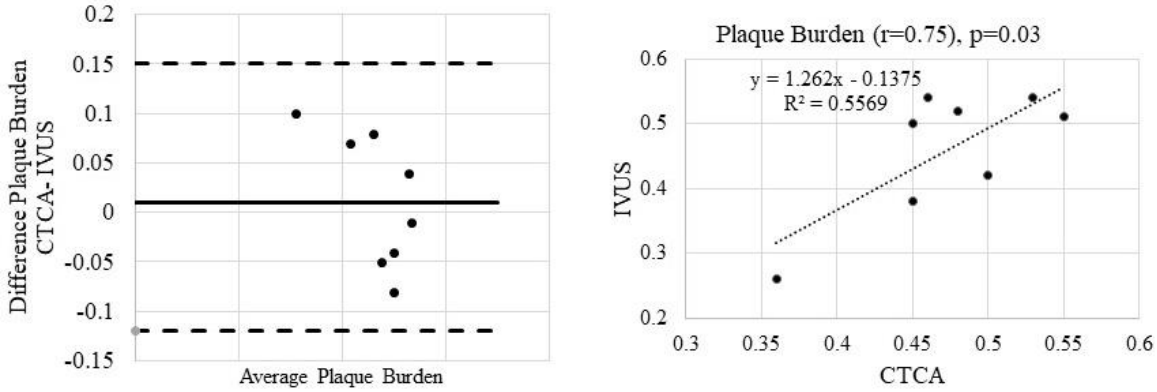

Figure 3.23 Bland Altman and Correlation plots for the PB.

The results above indicate a good correlation between the presented methodology and an already validated 3D reconstruction methodology [224], which utilizes the IVUS imaging modality. In some cases, the presented methodology indicates an excellent agreement, compared to the IVUS modality as it is shown in Figure 3.24, below. It is observed the calculated surfaces per 0.5 mm of the 3D coronary model derived by CTA based method are similar to those derived by IVUS based method.

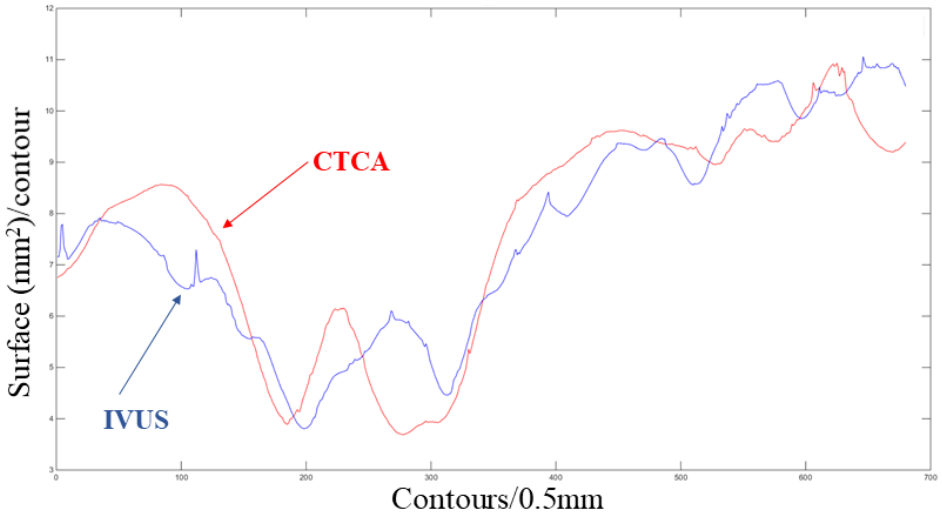

Figure 3.24 Comparison between the inner wall area per 0.5mm derived by the presented methodology and the inner wall area per 0.5mm derived by an IVUS based approach.

Evaluation of blooming effect removal using IVUS findings

As far as the validation of blooming effect removal is concerned, we demonstrate in Table 3.4 the geometrical features derived by the CTA based methodology with and without the incorporation of blooming effect removal, in comparison with the IVUS based reconstruction methodology. It is observed that the geometrical features derived by the methodology which incorporates the blooming effect removal indicates a better correlation with those derived by the CTA based methodology, which do not incorporate the blooming effect removal. In addition to this, in Figure 3.25, we demonstrate that the mean error of the lumen area per 0.5mm derived by the CTA based methodology when the blooming effect removal is incorporated, is lower than the respective error of the lumen area per 0.5 mm derived by the CTA based methodology without the incorporation of blooming effect removal, when it is compared to the IVUS based methodology.

Table 3.4 Comparison of geometrical metrics derived by the presented methodology with the incorporation of blooming effect removal, without the blooming effect removal and with those derived by an IVUS based approach.

Cases	**Metrics**	**CTA**	**CTA Blooming**	**IVUS**
#1	DS1	0.30	0.31	0.27
	DS2	0.54	0.53	0.46
	PB	0.56	0.52	0.52
	PB/0.5mm	0.53±0.09	0.48±0.17	0.54±0.17
#2	DS1	0.38	0.28	0.16
	DS2	0.36	0.34	0.33
	PB	0.54	0.50	0.37
	PB/0.5mm	0.50+0.07	0.46+0.06	0.36+0.2
#3	DS1	0.62	0.60	0.22
	DS2	0.72	0.70	0.45
	PB	0.47	0.42	0.64
	PB/0.5mm	0.48±0.08	0.47±0.09	0.52±0.09
#4	DS1	0.56	0.47	0.28
	DS2	0.59	0.51	0.68
	PB	0.53	0.57	0.73
	PB/0.5mm	0.55±0.09	0.58±0.08	0.51±0.08
#5	DS1	0.44	0.41	0.29
	DS2	0.41	0.41	0.60
	PB	0.36	0.39	0.50
	PB/0.5mm	0.36±0.09	0.37±0.09	-

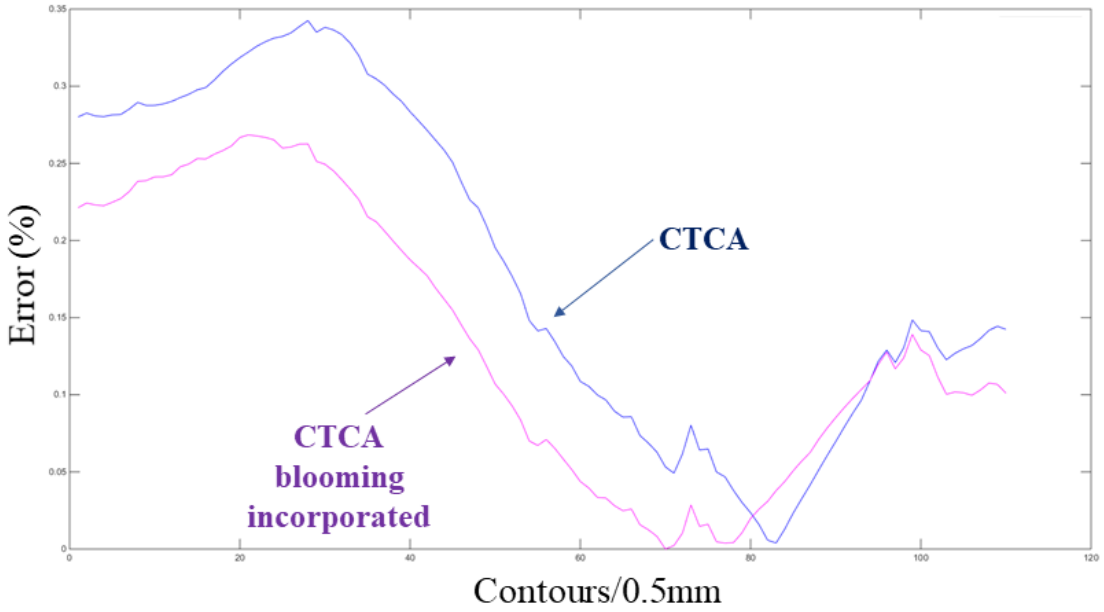

Figure 3.25 Comparison of the proposed methodology (with and without blooming incorporation) with IVUS based methodology.

3.5.1.2 Results for CP and NCP segmentation

In the validation procedure, 18 coronary arteries and 90 CTA and VH-IVUS slices were identified, registered and used for the validation procedure. After the implementation of the proposed methodology and the analysis of VH-IVUS slices, we identified 6 CP lesions and 12 NCP lesions. As far as the 3D comparison procedure is concerned, we present in Table 3.5 and Table 3.6 the volume and the length of the lesions obtained by the proposed methodology and by the VH-IVUS for CP and NCP, respectively.

The correlation between CTA and VH-IVUS is extracted using Bland Altman analysis and linear regression correlation analysis. The Bland Altman plots and the correlation plots are shown in Figure 3.26 for the CP volume and CP length of lesion, whereas in Figure 3.27, the corresponding plots for the NCP volume and length of the lesion are shown. As far as the 2D comparison procedure is concerned, in Figure 3.28 and Figure 3.29 the Bland Altman analysis and the correlation plot for the CP area and NCP area are shown, respectively. The comparison results

indicate that the Pearson's correlation (*r*) for the CP volume, the length of the lesion and the area was 0.93, 0.84 and 0.85, while the Degree of correlation (R^2) was 0.85, 0.71 and 0.72, for the CP volume, the CP length of the lesion and the CP area, respectively.

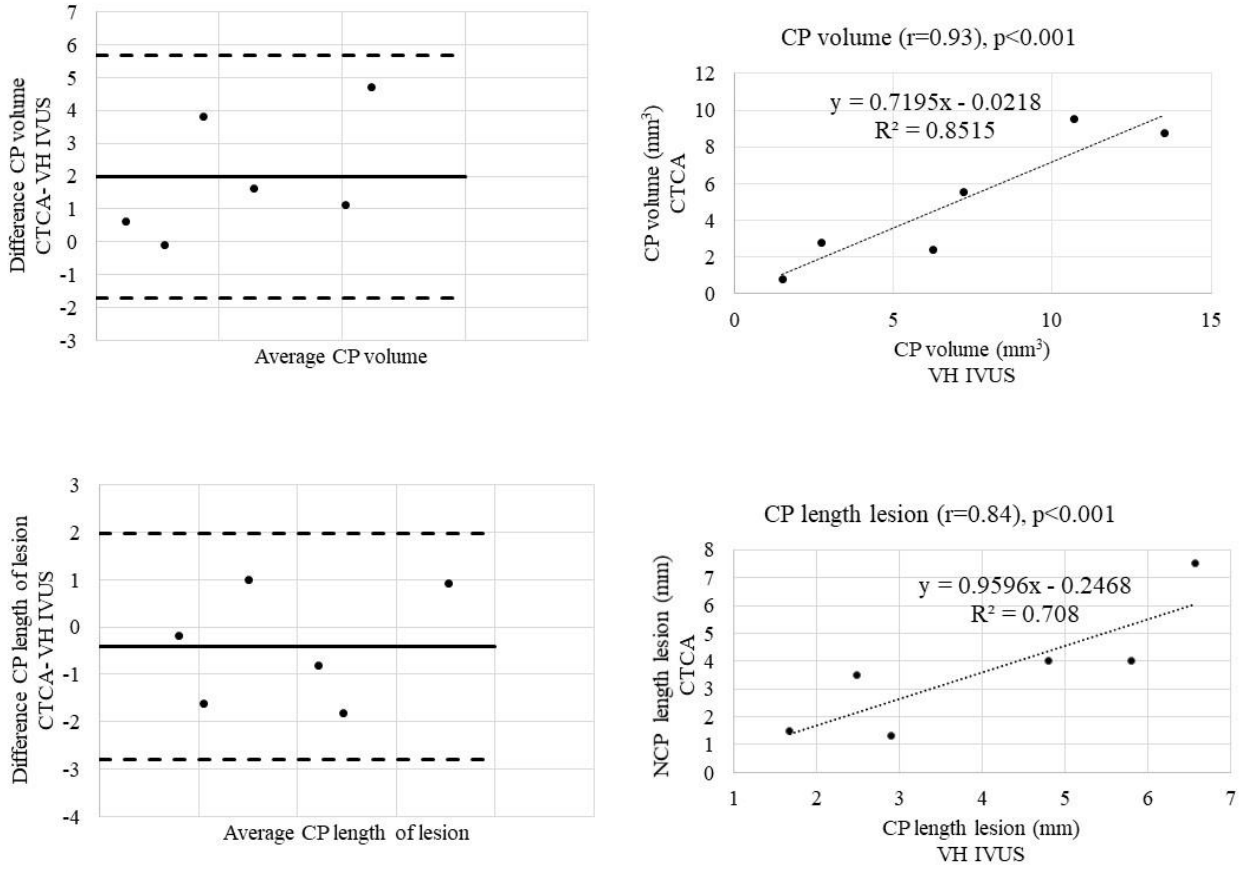

Figure 3.26 Bland-Altman and correlation plots for CTA and VH-IVUS for the CP volume and the CP length of lesion.

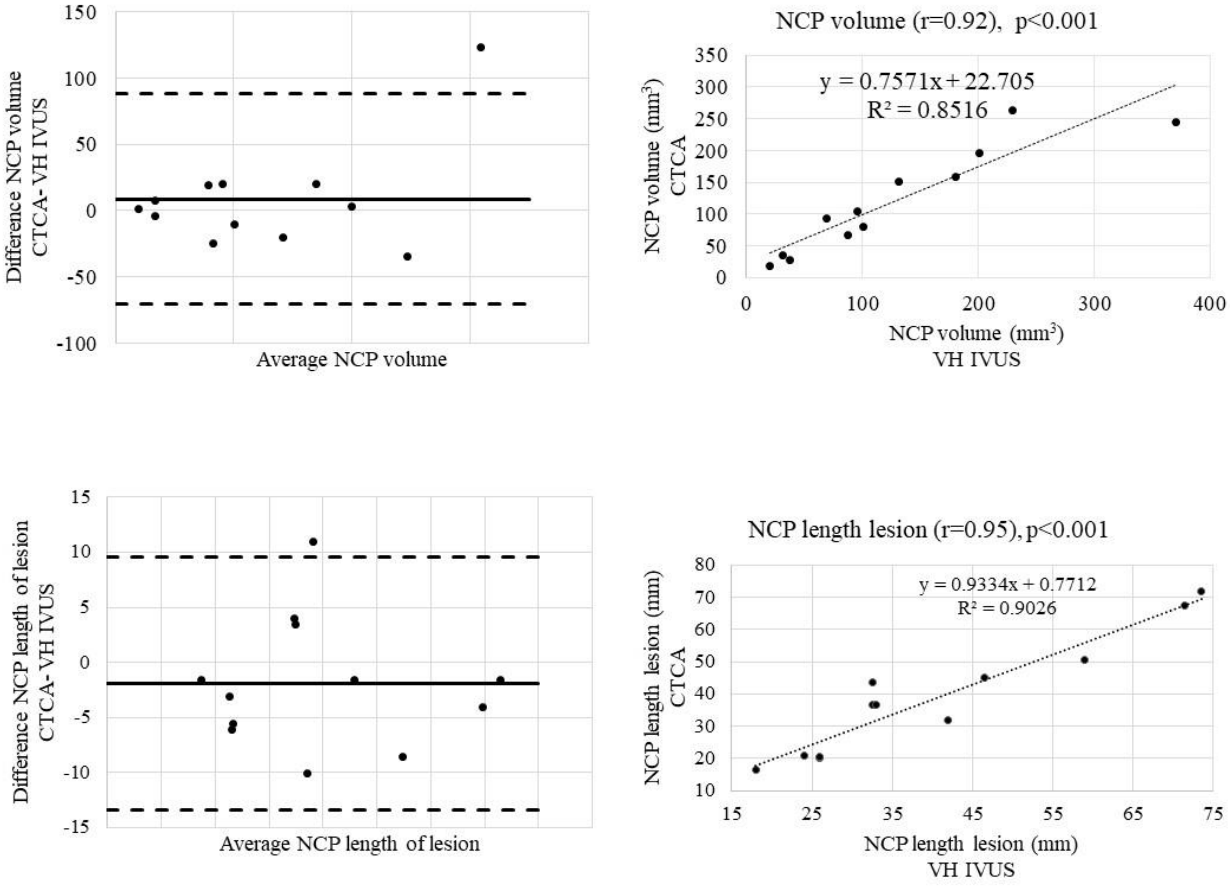

Figure 3.27 Bland-Altman and correlation plots for CTA and VH-IVUS for the NCP volume and NCP length of lesion.

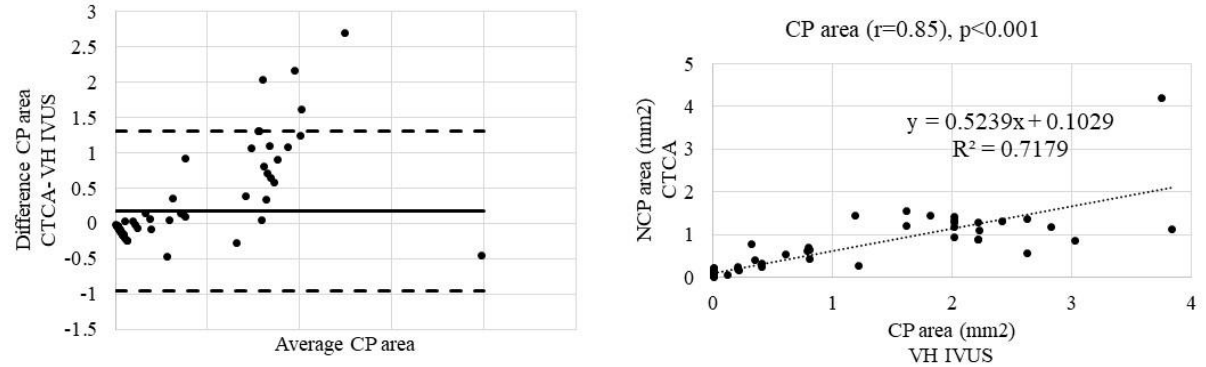

Figure 3.28 Bland-Altman and correlation plots for CTA and VH-IVUS for the CP area.

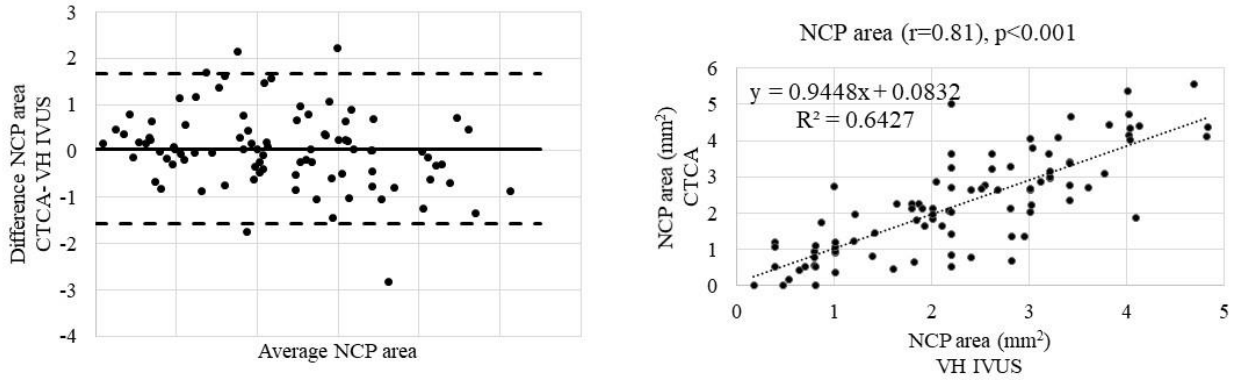

Figure 3.29 Bland-Altman and correlation plots for CTA and VH-IVUS for the NCP area.

The corresponding correlation values of NCP are 0.92, 0.95 and 0.81 for Pearson's correlation (*r*), and 0.85, 0.9, 0.64 for the Degree of correlation (R^2), for the plaque volume, the length of lesion and the area, respectively. The mean values of the plaque volume extracted by the proposed methodology are 6.98±4.59 mm^3 and 120.59±83.11 mm^3 for CP and NCP, respectively, whereas the corresponding mean values of the plaque volume by VH-IVUS are 5±3.58 mm^3 for CP and 129.28±101.3 mm^3 for NCP. The mean value of the CP length of the lesion is 3.63±2.25 mm and for the NCP lesion is 40.375±18.62 mm, while the mean value of the CP length of lesion based on VH-IVUS is 4.04±1.97 mm and of NCP lesion 46±18.29 mm. In addition to this, the mean value of the plaque area based on CTA and VH-IVUS is 0.59±0.99 mm^2 and 0.42±0.61 mm^2 for CP, respectively, and 2.31±1.17 mm^2 and 2.26±1.38 mm^2 for NCP, respectively. The previously described correlation metrics and the mean values of the validation metrics are shown in Table 3.7 and Table 3.8, respectively.

Table 3.5 Validation metrics for the CP.

Case	CP volume VH-IVUS (mm^3)	CP volume CTA (mm^3)	Length of the CP lesion VH-IVUS (mm)	Length of the CP lesion CTA (mm)
#1	2.81	2.72	2.9	1.3
#2	0.86	1.49	1.68	1.5
#3	9.54	10.69	4.81	4
#4	5.56	7.2	6.57	7.5
#5	8.8	13.51	5.81	4
#6	2.41	6.24	2.49	3.5

Table 3.6 Validation metrics for the NCP.

Case	**NCP volume VH-IVUS (mm^3)**	**NCP volume CTA (mm^3)**	**Length of the NCP lesion VH-IVUS (mm)**	**Length of the NCP lesion CTA (mm)**
#1	246.39	369.84	59	50.5
#2	158.73	179.63	71.5	67.5
#3	80.42	100.69	26	20
#4	18.39	20.16	24	21
#5	29.11	37.09	26	20.5
#6	151.12	131.24	42	32
#7	93.89	69.25	32.5	36.5
#8	197.48	200.69	73.5	72
#9	35.12	31.14	18	16.5
#10	105.29	95.05	33	36.5
#11	67.74	87.39	32.5	43.5
#12	263.38	229.2	46.5	45

Table 3.7 CTA and VH-IVUS correlation metrics.

	Correlation metrics	**Pearson's correlation (r)**	**Degree of correlation (R^2)**
CP	CP volume vs VH-IVUS	0.93	0.85
	CP length of lesion vs VH-IVUS	0.84	0.71
	CP area vs VH-IVUS	0.85	0.72
NCP	NCP volume vs VH-IVUS	0.92	0.85
	NCP length of lesion vs VH-IVUS	0.95	0.90
	NCP area vs VH-IVUS	0.81	0.64

Table 3.8 CTA and VH-IVUS validation metrics.

	Metrics	CTA	VH-IVUS
CP	volume	6.98±4.59	5±3.58
	length of lesion	3.63±2.25	4.04±1.97
	area	0.59±0.99	0.42±0.61
NCP	volume	120.59±83.11	129.28±101.3
	length of lesion	40.375±18.62	38.46±18.29
	area	2.31±1.17	2.26±1.38

Comparison with manual annotations

The utilized data were acquired from 27 patients, who underwent CTA imaging for clinical purposes. 1350 CTA slices were utilized and manually annotated for the lumen segmentation validation, whereas 78 CTA slices and 47 CTA slices were utilized for the CP and NCP segmentation validation procedure, respectively. In Table 3.9, we present the DICE and HD for the lumen segmentation, whereas in Table 3.10 and Table 3.11, we present the DICE and HD distributions for the CP and NCP, respectively. The mean value of DICE was 0.72±0.08, 0.7±0.09 and 0.62±0.07, for the lumen, CP and NCP, respectively, whereas the mean HD value was 1.95 ±0.45, 1.74 ±0.34 and 1.95 ±0.36 for lumen, CP and NCP, respectively. In Figure 3.30, we show the manual and the automated segmentation for the lumen (Figure 3.30a and Figure 3.30b), the CP (Figure 3.30c and Figure 3.30d) and the NCP (Figure 3.30e and Figure 3.30f).

Table 3.9 Validation metrics against manual annotations for the lumen segmentation.

Cases	**Lumen**		**Cases**	**Lumen**	
	DICE	**HD**		**DICE**	**HD**
#1	0.79±0.1	2.39±0.86	**#15**	0.77±0.05	2.43±0.18
#2	0.83±0.05	2.07±0.35	**#16**	0.66±0.14	1.6±0.43
#3	0.67±0.09	1.69±0.57	**#17**	0.52±0.12	1.78±0.58
#4	0.8±0.06	1.77±0.55	**#18**	0.68±0.08	1.588±0.73
#5	0.79±0.04	1.63±0.3	**#19**	0.81±0.1	1.86±0.66
#6	0.72±0.1	2.09±0.37	**#20**	0.77±0.08	1.68±0.43
#7	0.64±0.12	1.54±0.16	**#21**	0.58±0.13	2.03±0.34
#8	0.75±0.07	2.22±0.28	**#22**	0.68±0.12	1.9±0.57
#9	0.65±0.09	2±0.31	**#23**	0.59±0.1	2.04±0.73
#10	0.66±0.07	2.3±0.17	**#24**	0.78±0.07	1.64±0.34

Cases	Lumen		Cases	Lumen	
	DICE	HD		DICE	HD
#11	0.64±0.03	2.31±0.38	**#25**	0.73±0.09	1.84±0.44
#12	0.73±0.06	2.9±0.97	**#26**	0.82±0.05	2.05±0.32
#13	0.65±0.03	2.24±0.4	**#27**	0.8±0.08	1.67±0.62
#14	0.81±0.07	1.43±0.23			

Table 3.10 Validation metrics against manual annotations for the CP segmentation.

Cases	CP	
	DICE	HD
#1	0.79±0.12	1.33±0.16
#2	0.66±0.11	2.38±0.27
#3	0.67±0.09	2.56±0.23
#4	0.62±0.13	1.27±0.24
#5	0.61±0.15	1.89±0.38
#6	0.62±0.06	1.82±0.58
#7	0.72±0.07	1.58±0.36
#8	0.71±0.05	1.31±0.21
#9	0.78±0.04	1.94±0.51
#10	0.68±0.07	1.61±0.63
#11	0.72±0.08	1.87±0.32
#12	0.73±0.17	1.35±0.43
#13	0.76±0.09	1.82±0.15

Table 3.11 Validation metrics against manual annotations for the NCP segmentation.

Cases	NCP	
	DICE	HD
#1	0.69±0.08	2.39±0.56
#2	0.65	1.41
#3	0.78±0.06	1.42±0.2
#4	0.51±0.09	2.98±0.34
#5	0.54±0.08	1.28±0.38
#6	0.65±0.08	2.25±0.46
#7	0.49±0.08	1.89±0.56

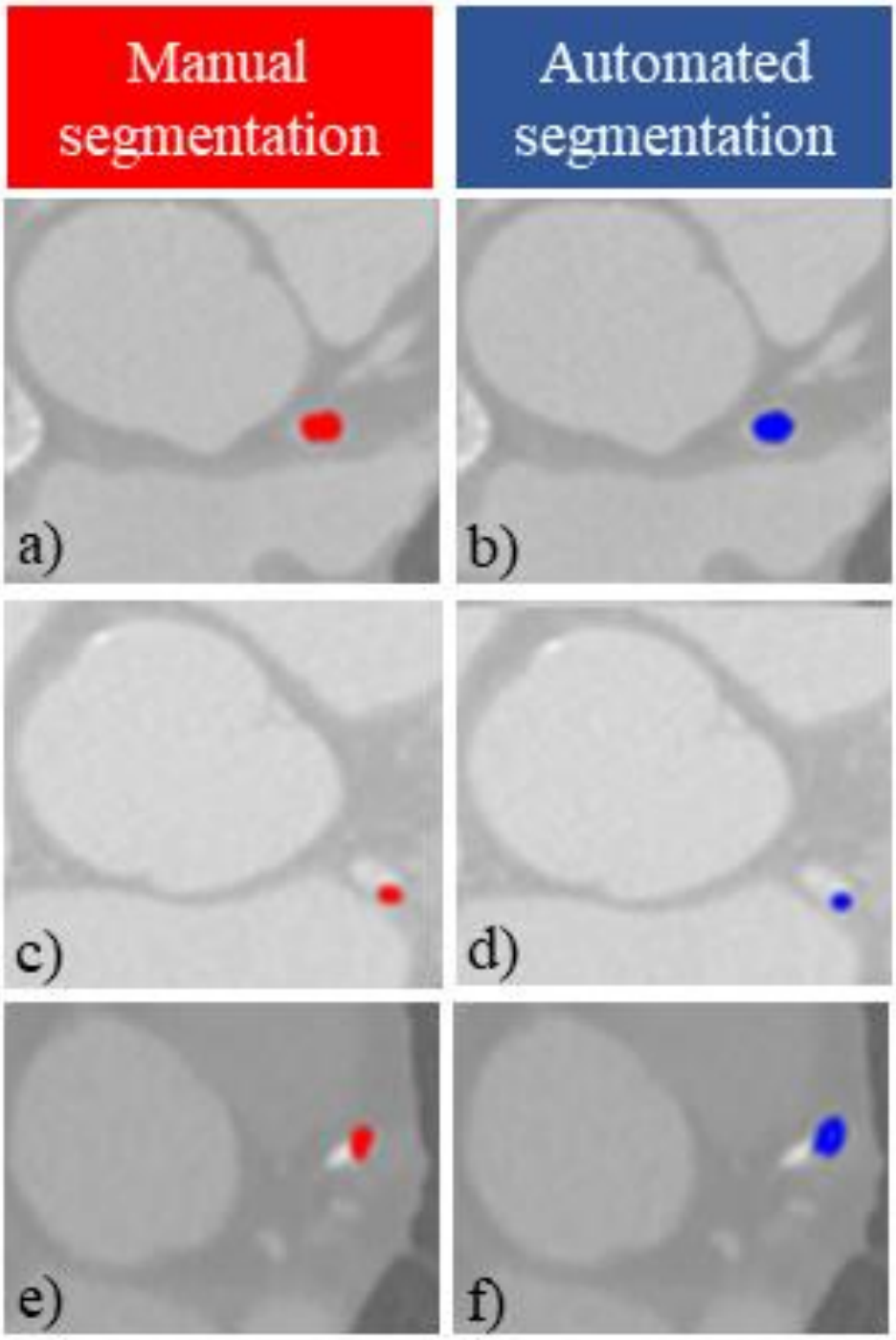

Figure 3.30 Manual and automated segmentation for the lumen (a, b), for the CP (c, d) and the NCP (e,f) [205].

3.5.2 Results for carotid artery reconstruction

Comparison with manual annotations

In the first validation strategy, shown in the Figure 3.31, we implement a validation procedure as far as the 2D segmentation is concerned, computing the DICE and the HD. Reconstruction of carotid arteries, both for the left common artery and the right common artery, was successfully obtained in 16 patients. Our results indicated a value of DICE is 0.83 and 0.8, whereas the HD is 1.55 and 1.44, as far as the inner wall and outer wall detection is concerned, respectively. The respective evaluation metrics for the identification of CP is 0.71 and 1.63, whereas the DICE and the HD for the NCP are 0.7 and, respectively.

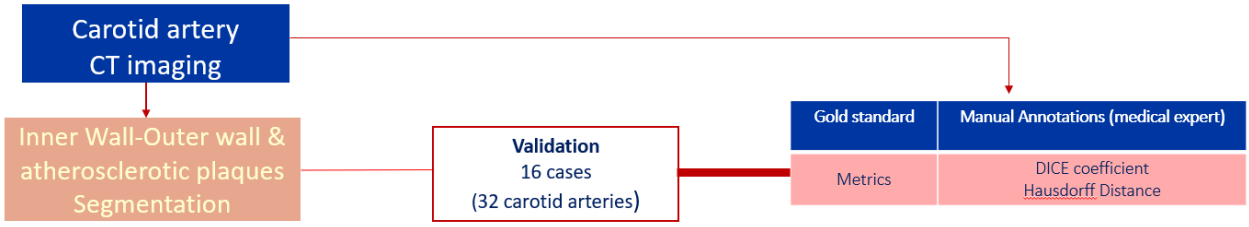

Figure 3.31 First Validation Strategy.

In the second validation strategy, shown in Figure 3.32, correlation analysis was implemented for the comparison of inner wall and outer wall area, derived by the automated proposed segmentation algorithm and these derived by manual expert's annotation. Correlation analysis for the inner wall is illustrated in Figure 3.33, whereas correlation analysis for the outer wall is illustrated in Figure 3.34.

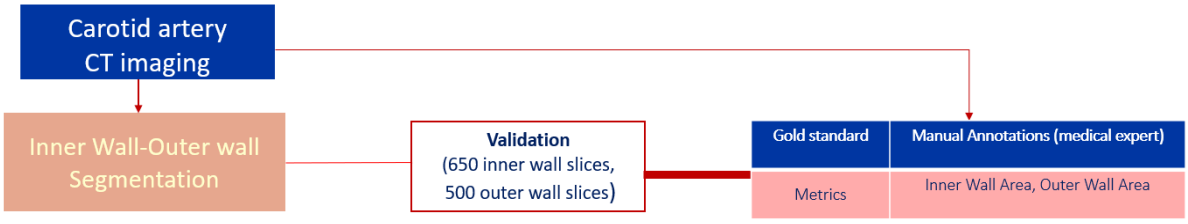

Figure 3.32 Second Validation Strategy.

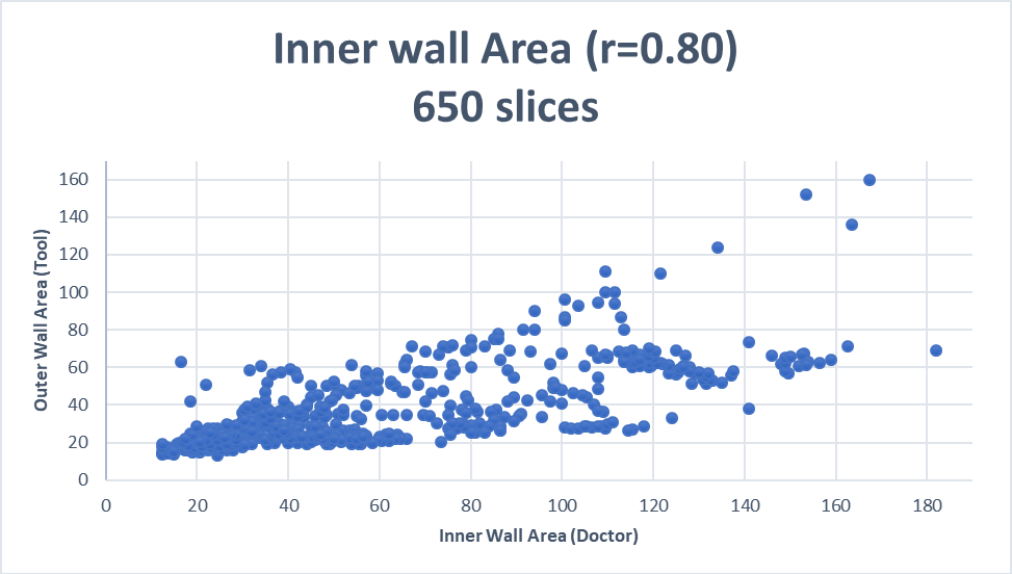

Figure 3.33 Correlation Analysis for the inner wall area.

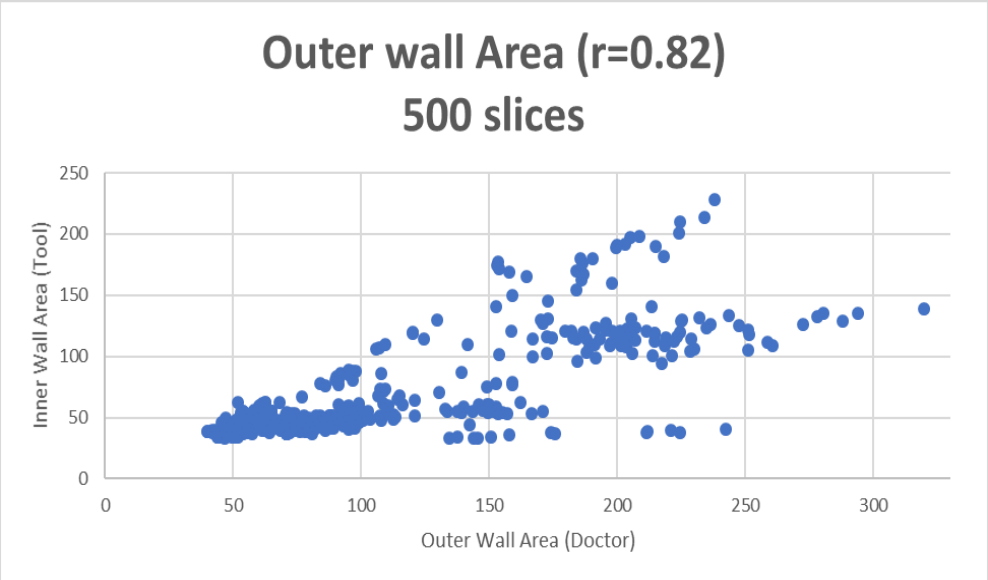

Figure 3.34 Correlation Analysis for the outer wall area.

3.6 Discussion

In this chapter, a semi-automated methodology for the reconstruction of the lumen and the outer wall of coronary and carotid arteries, using CTA images, was presented. The main innovation of this study is its automatic nature, since it only requires the starting and the ending point of the artery to

accurately reconstruct it. In addition to this, the presented methodology provides accurately the 3D reconstruction of the entire coronary and carotid arterial tree and the user needs only to annotate the starting point of the bifurcation and the ending points of the vessel branches. The presented methodology is mainly based on a level set segmentation model [215] for the segmentation of the lumen and the outer wall, which allows an accurate segmentation and its applicability is not limited by the low CTA images quality [214]. The unique preprocessing step of the acquired CTA images is to detect the vessel candidate regions. In addition to this, the extraction of the centerline using a minimum cost path based approach in combination with appropriate cost functions selection, ensures that the extracted centerline may be the globally optimal solution. Therefore, once the vessel centerline is successfully extracted, an appropriate initial vessel mask for the lumen segmentation is created. Furthermore, minimum cost path approaches are able to overcome problems related to overlapping pathologies, depicted in the image, as well as, issues of low image quality.

As for the coronary artery reconstruction, to our knowledge, our study is the only approach in the literature that allows a 3D reconstruction of coronary anatomy, which is compared both by a medical expert's annotations and IVUS findings. The results of the proposed methodology demonstrated that our approach is able to accurately segment the lumen and outer wall and provide geometrically correct 3D models. In the literature, several methodologies [73, 197, 228-230] have been presented for the segmentation of CTA cross sectional images and the classification of its plaque components. These approaches are time consuming, since corrections of the detected borders are sometimes required and the reconstruction of the coronary anatomy is not implemented in an automatic or semi-automatic manner. This limitation is overcome by the new proposed automatic methodology, which provides in a fast manner the segmentation of CTA images.

Furthermore, existing studies [64, 73, 80] are evaluated using only IVUS modality, whereas the presented study is validated using both medical expert annotation and IVUS 3D models. Although, the manual annotation requires a lot of effort and time, since the expert annotates in each slice the lumen and the outer wall, expert manual segmentation of real images is regarded as a practical gold standard against which new segmentation algorithms can be compared [223]. On the other hand, a comparison with IVUS 3D reconstructed models allow us to validate the geometry of the

reconstructed arteries and extract validation metrics, such as the coronary lumen stenosis, the PB, the MLA and the MLD.

In contrast to Athanasiou *et al.* existing study [64], which focused on comparison metrics, such as the comparison of the volume and areas of the ROI, the length and angle of the vessel, our validation process is dedicated to quantify the region of agreement, the overlap region between our proposed methodology and the manual segmentation, as well as validation metrics which demonstrate the 3D model accuracy. It is known that an objective validation of image segmentation is of great importance, but it is such a difficult task. Based on the literature [223, 225], the selected in this approach comparison metrics are two of the most common measures and their values indicate how accurately the segmentation algorithm performs. More specifically, the mean value of DICE is 0.749, while its standard deviation (SD) is 0.0787 and the mean value of HD is 1.746, while its SD is 0.573. The correlation coefficients for the DS1, the DS2, the PB, the MLA and MLD, when comparing the derived from the proposed methodology 3D models with the IVUS reconstructed models, were 0.79, 0.77, 0.75, 0.85, 0.81, respectively. In the proposed methodology, the validation dataset consists of data acquired from three different scanners. That means that our 3D reconstruction approach is applicable in different clinical environments, since each CT scanner is characterized by its different properties and settings.

Regarding the carotid artery reconstruction, different studies have been proposed in the literature for the carotid arteries CTA imaged analysis. Most of them are dedicated on the entire carotid vessel (outer wall) detection [86, 88, 89, 91, 97] and do not identify the inner wall (lumen) of the carotid arteries. On the other hand, Freiman *et al.* [93] in a study proposed in 2012 presented a graph-cut based segmentation algorithm for the accurate detection of carotid artery lumen (inner wall) and achieves an overall DICE of 84.5% ± 3.3%. Nevertheless, this study is validated against experimental existing studies [231] and not directly compared to experts' manual annotations, which is considered as the gold standard. Moreover, another carotid artery lumen segmentation methodology has been presented by Hemmati *et al.* [96] and achieves an overall DICE of 0.85. However, this study is not capable of identifying the vessel centerline of small branches and it is necessary to remove small branches in centerline extraction stage. Additionally, the methodology

fails to identify the vessel centerline in the regions where the vessel is occluded. Other existing studies, such as the study proposed by Guha *et al.* [98] and dos Santos *et al.* [95] have not been validated quantitatively.

Our proposed study for the segmentation of the entire carotid artery vasculature and the atherosclerotic plaques presented a complete methodology for the segmentation of the two anatomical layers of the carotid arteries, both the inner wall and the outer wall. The detection of both the inner and the outer wall of the carotid artery constitutes the novel aspect of the proposed methodology, since both layers have a clinical significance. More specifically, the DS, calculated by the segmentation of the inner wall, is considered as one of the most significant geometric features of the carotid artery and its value contributes to the clinical decision and the patient management. Additionally, the quantification of the overall PB region, computed by the detection of both the inner wall and outer wall, is also a significant measurement as far as the carotid artery disease prediction and its progression is concerned.

Regarding the atherosclerotic plaques detection based on CTA, this methodology presents an automated methodology for their reconstruction. To our knowledge, the proposed study constitutes a novel approach due to its automated nature and its validation process, as far as the CTA validation process is concerned. More specifically, the proposed study is validated utilizing VH-IVUS modality for the 3D distribution of CP and NCP, while it is also validated in terms of CP and NCP segmentation accuracy, by directly comparing medical doctor's annotations.

As far as the CP is concerned, its pixels can be easily identified using the CTA modality due to their high intensity values. On the other hand, the detection of NCP constitutes a challenging problem, since the NCP intensity is characterized by lower intensity values, close to the blood and the muscle tissues. The CTA modality is a promising non-invasive technique to accurately detect the NCP in the PB. Direct identification and quantification of NCP and PB is an important issue, since the NCP has been indicated as a significant indicator of acute coronary syndromes [35] and its presence is more likely to be associated with high-risk groups, such as those with elevated inflammatory biomarkers and those suffering from DM [36, 37].

In the proposed methodology, two 2D segmentation approaches are implemented for the detection of CP and NCP. The identification of CP relies on level set models, whereas the NCP detection is based on a dynamic thresholding technique. Although the level set models achieve an accurate segmentation, in the case of NCP this is not applicable, since the HU intensities of NCP proposed in the literature are close to those of the outer wall. Therefore, since an accurate inner and outer wall detection has been achieved, a validated ROI detection ensures the accurate CP and NCP detection. As a result, a good correlation was indicated and no significant differences of plaque volumes, lengths of coronary atherosclerotic plaque lesions and areas quantified by CTA and VH-IVUS were observed. In Table 3.12, a comparison of the NCP volumes derived by the proposed methodology and by other methodologies is presented. It is observed that the study proposed by Dey *et al.* [16] indicate a higher correlation. This study is concentrated only on the quantification and detection of NCP and does not provide any details about the vessel geometry and the distribution of the plaque in 3D space. In addition to this, the study proposed by Dey *et al*., [16] requires by the user the starting and the ending point of the plaque lesion to quantify the NCP lesion. Contrary to this study, our proposed methodology is able to identify either CP or NCP lesions in the entire coronary arterial tree, indicating only the starting point of the coronary bifurcation and the ending points of each coronary artery branch.

Furthermore, several studies have been conducted focusing on the ability of expert observers to detect and annotate manually NCP using CTA. These studies indicate a strong correlation with the NCP volumes derived by IVUS. However, these approaches are time consuming, since they are not implemented in an automatic or semi-automatic manner [39-42]. In our methodology, this limitation is overcome, since the proposed methodology requires low computation time for the 3D reconstruction of coronary arteries. In addition to this, it should be noted that the proposed methodology achieves an expedite 3D coronary reconstruction of the entire coronary arterial tree. In this way, the atherosclerotic plaque can be identified not only in coronary segments, but also in crucial regions, such as the coronary bifurcations.

Table 3.12 Comparison of the proposed methodology with the other methodologies for the NCP detection.

Study	**Volume by CTA**	**Volume by IVUS**	**Pearson's correlation (*r*)**	**Degree of correlation (R^2)**
Brodoefel *et al.* [17]	56.7±30	55.8±26	-	0.84
Dey *et al.*[16]	116±80.1	105.9±83.5	0.94	-
Leber *et al.* [18]	59.8±76.6	67.7±67.9	-	0.69
Schepis *et al.* [19]	89±66	90±73	0.89	-
Proposed methodology	120.59±83.11	129.28±101.3	0.92	0.85

Other studies have demonstrated that the mean luminal attenuation value varies, due to the different acquisition dose protocol. In addition to this, it has been indicated that the intra-arterial injection of iodinated contrast agent affects not only the luminal attenuation value, but also the atherosclerotic plaque attenuation value [43]. The luminal attenuation varies also between patients. In our study, the segmentation of NCP is clearly related to the mean luminal attenuation and more specifically the computed *ml* value varies from 276- 418 HU. As such, the proposed methodology can be adapted to all the CTA acquired images and acquisition protocols.

Apart from the selected HU threshold values for NCP, another important factor for the accurate identification of NCP is the reliable identification of ROI. Accurate lumen and outer wall area segmentation results in accurate PB region detection. In our study, the detection of the ROI has already been validated [38] using both manual expert's annotations and a 3D reconstruction methodology using angiographic data fuzzed with IVUS data [34]. Thus, the presented methodology for NCP detection is a promising technique allowing the accurate PB characterization.

Furthermore, it should be noted that the VH-IVUS permits the plaque characterization with a higher spatial resolution than the non-invasive CTA. Due to its low spatial resolution, CTA is limited to identify the different subtypes of the NCP, contrary to VH-IVUS and to extract in a reliable manner the smaller atherosclerotic plaque volumes. This approach was validated using large NCP volumes, ranging from 18.39 -263.38 mm^3, which are more prone to rupture.

The presented methodology incorporates the blooming effect removal, aiming to improve the CTA image visualization of calcifications. In this manner, an accurate 3D quantification of CP is achieved, without overestimating the CP volume, which results in the overestimation of the coronary lumen stenosis. The blooming effect remains a challenging artifact of CTA images, which can be limited by deconvolution and filtering techniques. In addition, we implement a deconvolution procedure on the high intensity image's pixels to quantify in a reliable manner the CP volume. Thus, no significance difference is observed between the CP volumes derived by the presented methodology and those extracted by VH-IVUS analysis, achieving a high Pearson's correlation ($r=0.93$).

Chapter 4 Coronary Artery Disease Prediction

4.1 Introduction

4.2 Coronary artery disease prediction based on imaging and non-imaging data

4.3 A Machine Learning Approach for the Prediction of the Progression of Cardiovascular Disease based on Clinical and Non-Invasive Imaging Data

4.4 Site specific prediction of atherosclerotic plaque progression

4.5 Site specific prediction of PCI stenting based on imaging and biomechanics data using gradient boosting tree ensembles

4.5 Conclusions

4.1 Introduction

The recent advances in coronary imaging techniques, either invasive or non-invasive, have enabled the identification of coronary vessels features, which are considered as CAD risk factors. Taking advantage of the overall pipeline of 3D reconstruction of coronary arteries, which was described in detail in chapter 3, in this chapter we aim to develop and present innovative ML models for the CAD prediction.

More specifically, the basic aim of this chapter is to:

i. develop a ML predictive model, which incorporates both non-invasive imaging data, derived by CTA and typical patient's baseline characteristics to predict the CAD risk and especially the obstructive disease [232].
ii. present a patient specific ML based model for the prediction of CAD progression using vasculature geometrical features based on CTA imaging and the conventional CAD risk factors [233].

iii. present a site specific ML tree- based model for the prediction of atherosclerosis progression [234].

iv. present a site specific model for the prediction of the coronary site specific Percutaneous Coronary Intervention (PCI) stenting placement, based on imaging derived predictors, implementing a gradient boosting tree ensemble technique [235].

4.2 Coronary artery disease prediction using machine learning

4.2.1 Methodology

4.2.1.1 CTA Image Analysis and Three-Dimensional Reconstruction

The first step of the development of the CAD risk prediction model was the analysis of the CTA images. This analysis was conducted by implementing an active contour based model for the segmentation of CTA images and aimed to provide a detailed geometry of the three major coronary arteries, the LAD, the LCX, and the RCA. This methodology is integrated in a dedicated software tool, which can semi-automatically provide the detailed 3D coronary artery anatomy [204, 205]. More details for the overall 3D reconstruction methodology can be found in the Chapter 3.

4.2.1.2 Calculation of the SMARTFFR index

In this study, except of the geometrical derived metrics. a blood-flow-based index, the SMARTFFR index [236] was utilized. In order to calculate SMARTFFR, blood flow finite element simulations are carried out on the reconstructed 3D models of the coronary arteries. The arterial lumen is discretized into tetrahedral finite elements of face size that ranges from 0.09 to 0.12 mm and the respective Navier-Stokes and continuity equations are then solved using the Finite Elements method. A transient blood flow simulation is performed on the 3D reconstructed artery. The flow is considered laminar and the blood is treated as a Newtonian fluid with density 1050 kg/m^3 and dynamic viscosity 0.0035 Pa·s. For each timestep, the Pd/Pa value is calculated to construct the Pd/Pa vs. flow curve. The calculated Pd/Pa values for every timestep are then connected to create the appropriate patient-specific curve. The patient-specific curve is constructed for a flow range of 0-4 ml/s and the SMARTFFR value is calculated by dividing the area under the patient specific curve to the respective AUC of the respective healthy arterial segment [237]. The overall procedure for SMARTFFR index calculation is shown in Figure 4.1.

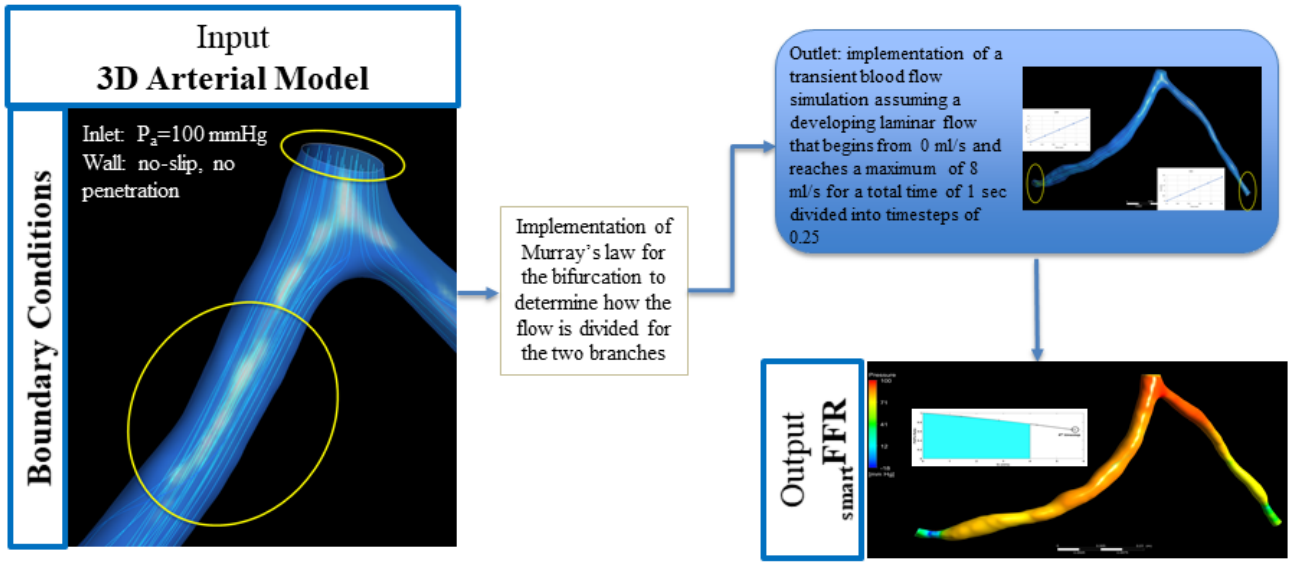

Figure 4. 1 Overall procedure for the SMARTFFR index calculation.

4.2.1.3 Problem Definition

In the presented approach, the CAD risk stratification problem has been formulated as a two-class classification problem based on the maximal coronary artery stenosis. This hypothesis is based on the findings of the Coronary Artery Disease Reporting and Data System (CAD-RADS) [238], which provides a standardized method to associate findings of the CTA imaging modality to facilitate decision making regarding further patient management. Figure 4.2 shows the distribution of the population across the two CAD-severity groups. More specifically, among the total 263 patients who underwent CTA imaging for clinical purposes, 55 patients underwent PCI stenting procedure and 10 patients coronary artery bypass grafting (CABG) procedure, whereas CTA images of 11 patients were considered as interpretable either at the baseline time step or at the follow-up time step. The annotation was based on the assessment of the obstructive disease: at least one major artery with stenosis > 50% and at the follow-up time step 125 participants were at Class 1— C_1, whereas 62 participants were at Class 2— C_2.

The definition of these two classes is based on the quantitative DS derived by the CTA imaging modality according to the society of CV CT guidelines committee [239]. More specifically, the first class, the no CAD—minimal CAD class (Class 1— C_1)), includes the grading scale 0, 1, and 2 (normal, minimal, and mild), whereas the obstructive CAD class (Class 2— C_2)) includes the grading scale 3, 4 ,and 5 (moderate, severe, and occluded), as it is shown in Table 4. 1. This classification was selected because we want to predict the obstructive CAD disease.

Table 4. 1 Definition of the utilized CAD risk classes.

Proposed Classes	Recommended Stenosis Grading Scale of CAD	Quantitative Stenosis
Class 1— C_1	0: Normal	No luminal stenosis
	1: Minimal	Plaque with <25% stenosis
	2: Mild	25–49% stenosis
Class 2— C_2	3: Moderate	50–69% stenosis
	4: Severe	70–99% stenosis
	5: Occluded	100% stenosis

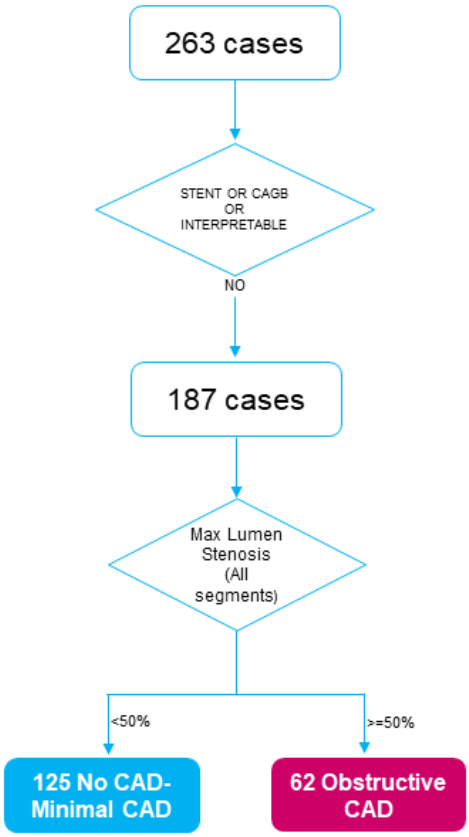

Figure 4.2 Flow chart depicting the distribution of the cohort in CAD-severity groups based on the CTA imaging at the follow-up step. in total, 287 patient imagings (125 in Class 1 and 62 in Class 2) were analyzed [232].

Baseline imaging and non-imaging characteristics were trained into a gradient boosting classification scheme, aiming to discriminate the patients at low risk (Class C_1) and those at high risk (Class C_2), concerning their follow-up time step. This predictive supervised learning approach aims to learn mapping from input features x to output Y given a labeled set of input output pairs $D = \{(x_i, Y_i)\}_{i=1}^{N}$, where D is the training set, and N is the number of training examples [240]. Each sample (x_i, Y_i)

associates the input features with the risk prediction of CAD severity, Y, where $Y \in \{C_1, C_2\}$, is estimated by a non-linear parameterized function (f) of input features $x \in R^d$, $x = [x_1, x_2, \dots, x_N]$. The goal of this supervised classification problem is to obtain an approximation $F(x)$ of the function $F^*(x)$ mapping the input x to output Y. The function $F^*(x)$ minimizes the expected value of some specified loss function $L\big(y, F(x)\big)$, whereas the procedure followed in this proposed study is to restrict the function $F(x)$ to be a member of parameterized class of functions $F(x; Y)$. In addition to this, in this paper, we constructed our model based on additive expansions of the form $F(x; \{\beta_m, a_m\}_1^M)$. $F^*(x)$ and $F(x; \{\beta_m, a_m\}_1^M)$, which are described in the following equations [240].

$$F^*(x) = \arg min\ E_{y,x} L\big(y, F(x)\big) = \arg\min E_x \left[\left(E_y \left(L\big(y, F(x)\big)\right) \middle| x\right)\right]. \quad (4.1)$$

$$F(x; \{\beta_m, a_m\}_1^M) = \sum_{m=1}^M \beta_m h(x; a_m), \quad (4.2)$$

where $h(x; a)$ is a simple parameterized function of the input variables x, characterized by the parameters $a = \{a_1, a_2, \dots\dots\}$, whereas $\{\beta_m, a_m\}_1^M$ denotes the entire parameter set [240].

The selected predictive model was nested into an easy ensemble classification scheme to overcome the class imbalance problem. To estimate the classification performance of the proposed method, an externally stratified 10-fold cross-validation was applied, with data pre-processing, a multivariate feature ranking, and a gradient boosting classification scheme being efficiently combined at each iteration of the procedure. The overall proposed model performs feature selection in the learning time since it achieves model fitting and feature selection simultaneously. Data-preprocessing and feature ranking follow the resampling procedure itself, which reduces the selection bias in the estimates of the model's performance, whereas stratification assures that each validation fold retains the class distribution in the dataset. In addition to these, randomized search optimization of the model's hyper-parameters over an internal 3-fold cross-validation contributes to the fine-tuning of the presented model.

4.2.1.4 Easy Ensemble Algorithm Implementation-Class Imbalance Handling

The easy ensemble algorithm [241] is a class imbalance handling approach in which P are the training instances of the minority class, whereas Q denotes the instances of the majority class. The idea of the easy ensemble algorithm is to employ random resampling to generate K subsets of $\{Q_1, Q_2, \dots\dots, Q_K\}$ from Q ($|Q_i| < |Q|$, i = 1,2,…K). Subsequently, each $Q_i \cup P$ is trained by the classifier, and the final decision is selected by majority voting. In the proposed predictive model approach, the easy ensemble approach is combined with the gradient boosting classifier.

4.2.1.5 Data Pre-Processing

In this step, one hot encoding procedure was implemented to represent all the categorical input features as binary vectors. In addition to this, a curation procedure was implemented to curate our dataset both for outliers and missing values. All the input features whose missing values were higher than 10% were removed from the dataset, whereas features with missing values lower than 10% were imputed by either the most frequent value (categorical type features) or the median value (numerical type features).

4.2.1.6 Recursive Feature Elimination

In this step, our aim is to reduce the dimensionality d of input features $x \in R^d$ to overcome the risk of overfitting, which basically arises when the number of d is comparatively large, and the number of the training patterns is small. In this study, a feature ranking technique with a SVM with recursive feature elimination (RFE) was implemented to rank the input features. The whole SMV RFE procedure is shown in Table 4.2 [242].

Table 4.2 SVM RFE Algorithm [242].

Input:

Training Examples: $X_0 = [x_1, x_2, \ldots., x_k, \ldots. x_l\ ]^T$

Class labels: $y = [y_1, y_2, \ldots., y_k, \ldots. y_l\ ]^T$

Initialize:

Subset of features: $s = [1,2, \ldots., d]$

Feature ranked list $r = [\quad]$

Repeat until $s = [\quad]$

Restrict training examples to good feature in DICEs: $X = X_0(:, s)$

Train the classifier: $a = SVM\_train(X, y)$

Minimize over a_k: $J = \frac{1}{2}\sum_{hk} y_h y_k a_h a_k (x_h \cdot x_k + \lambda\delta_{hk}) - \sum_k a_k$, subject to: $0 \le a_k \le C$ and $\sum_k a_k y_k = 0$

Outputs: Parameters: a_k

Compute the weight vector of dimension length(s): $w = \sum_k a_k y_k x_k$

Compute the ranking criteria: $c_i = (w_i)^2$, for all i

Find the feature with smallest ranking criterion: $f = argmin(c)$

Update feature ranked list: $r = [s(f), r]$

Eliminate the feature with smallest ranking criterion: $s = s(1: f-1, f+1: length(s))$

Output: Feature ranked list r

4.2.1.7 Gradient Boosting Classification

In the first step, the gradient boosting classification algorithm [243] implements a numerical implementation minimizing the equation of $F^*(x)$ (Equation 4.1). The Gradient Boosting classification algorithm [243] implements a numerical implementation, minimizing Equation 4.1 and yields an additive expansion of the form:

$$F^*(x) = \sum_{m=0}^{M} f_m(x), \tag{4.3}$$

where f_0 is an initial guess and $\{f_m\}_1^M$ are successive "boosts", each based on the sequence of preceding steps. More specifically, for the steepest-descent

$$f_m(x) = -p_m g_m(x), \tag{4.4}$$

with

$$g_m(x) = \left[\frac{\partial E_y\left[\left\langle L(y, F(x))\middle| x\right\rangle\right]}{\partial F(x)}\right]_{F(x)=F_{m-1}(x)}, \tag{4.5}$$

$$F_{m-1}(x) = \sum_{i=0}^{m-1} f_i(x), \tag{4.6}$$

and

$$p_m = argmin_p E_{y,x} L(y, F_{m-1}(x) - p g_m(x)), \tag{4.7}$$

Subsequently, based on the training dataset $D = \{(x_i, Y_i)\}_{i=1}^N$ and aiming to minimize Equation 4.1, we try a "greedy-stagewise" approach to obtain

$$(\beta_m, a_m) = argmin_{\beta,a} \sum_{i=1}^{N} L(y_i, F_{m-1}(x_i) + \beta h(x_i; a)), \tag{4.8}$$

and then

$$F_m(x) = F_{m-1}(x) + \beta_m h(x; a_m). \tag{4.9}$$

Given an approximation of $F_{m-1}(x)$, the function $\beta_m h(x; a_m)$ can be considered as the best greedy step towards the data based estimate of $F^*(x)$, whereas the data based analogue of the unconstrained negative gradient;

$$-g_m(x_i) = -\left[\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)}\right]_{F(x)=F_{m-1}(x)}, \tag{4.10}$$

gives the best steepest-descent step direction $-g_m = \{-g_m(x_i)\}_1^N$ in the N-dimensional data space at $F_{m-1}(x)$. The most highly correlated $h(x; a)$ with $-g_m(x)$ over the data distribution can be obtained from the solution

$$\alpha_m = argmin_{a,\beta} \sum_{i=1}^{N} [-g_m(x_i) - \beta h(x_i; a)]^2. \tag{4.11}$$

The line search is performed by

$$p_m = argmin_p \sum_{i=1}^{N} L\big(y_i, F_{m-1}(x_i) + p h(x_i; a_m)\big), \tag{4.12}$$

and the final approximation is given by:

$$F_m(x) = F_{m-1}(x) + p_m h(x; a_m). \tag{4.13}$$

The overall pipeline of the proposed ML methodology is shown in Figure 4.3, below.

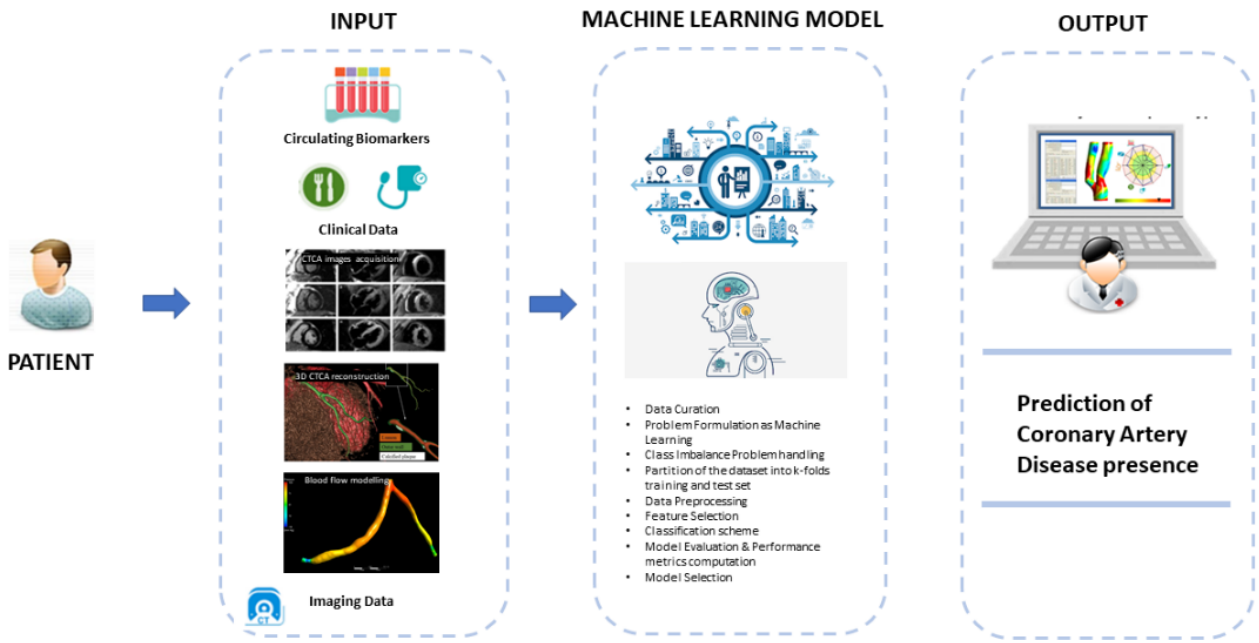

Figure 4.3 Overall pipeline of the proposed methodology [232].

4.2.2 Dataset

The proposed study is based on the EVINCI population [244], in which patient-specific information, both imaging and non-imaging, were collected for clinical purposes and utilized as the baseline information for the development of a CAD risk stratification methodology, whereas the follow-up data were collected after 6.22 ± 1.42 during the SMARTool project (September 2016–November 2017) [245]. More specifically, during the H2020 SMARTool project, a prospective, multicenter study in patients was conducted by 7 medical centers (Pisa, Turku, Zurich, Barcelona, Warsaw, Naples, Viareggio) from 5 European countries. All the participants signed informed consent to participate in the study and all the following procedures. Patients who previously underwent coronary CTA during the EVINCI (Evaluation of Integrated Cardiac Imaging for the Detection and Characterization of Ischemic Heart Disease; FP7-222915; n = 152-February 2009–June 2012) [12] and ARTreat (FP7-224297; n = 18) [13] clinical studies were prospectively included to undergo follow-up CTA. In addition to this, individuals (n = 32) who underwent CTA in the period from 2009 to 2012 were also prospectively included. A detailed list of inclusion and exclusion criteria is provided in the Appendix (Full list of eligibility, inclusion, exclusion and exit criteria).

Anonymized data were acquired from 187 patients, derived by different medical centers, and the cohort data were obtained under a data protection agreement fulling all the ethical and legal requirements for data sharing posed by the General Data Protection Regulation in a third-level care

setting. Table 4.3 below demonstrates the collected data types. The median age of the patients of our dataset is 61 years old (45–76), and at their first visit to the physician, all the participants underwent CTA imaging regardless of the presence of symptoms. More specifically, 45% and 25% of the participants had atypical and typical angina, respectively, whereas 12% of them had other symptoms, and 16% were asymptomatic. In addition to this, as for the pharmaceutical treatment of the participants, 18%, 28%, 13%, 40%, 13%, 10%, 3%, and 48% of them received angiotensin receptor blockers (ARBs), angiotensin converting enzyme inhibitors (ACEi), diuretics, beta blockers, calcium antagonists, oral antidiabetics, insulin, and statins, respectively, at the baseline time step.

Table 4.3 Imaging and non-imaging data utilized.

Type		**Features**
Imaging data	**Geometrical vasculature**	DS, MLA, MLD, PB, CP Volume, NCP volume, $_{SMART}$FFR Index, Number of CP, Number of NCP
Non-imaging data	**Demographics**	Age, Gender
	Risk factors	Family History of CAD, Hypertension, DM, Dyslipidemia, Smoking, Obesity, Metabolic Syndrome, Past Smokers
	Biohumoral Markers	Creatinine, Uric Acid, Glucose, TC, HDL, LDL, Triglycerides, Insulin, Aspartate Aminotransferase, Alanine Aminotransferase, Alkaline Phosphatase, Gamma-glutamyl Transferase, hs-CRP, IL-6, TSH, fT3, fT4, Leptin, MMP2 Protein Plasma, MMP9 Protein Plasma, hs-cardiac Troponin T, N terminal Fragment of Pro-brain Natriuretic Peptide, Lipidomics, Metabolomics

4.2.3 Results

CAD Risk-Prediction Model Performance Evaluation

The utilized CAD risk-prediction model performance metrics are the balanced accuracy, the NPV, the PPV, the AUC, and the sensitivity and specificity. The values of the adopted performance metrics and their mean value and the 10-fold SD are given in Table 4.4. The average balanced accuracy of the selected predictive model is 0.81, while its sensitivity and specificity is 0.88 and 0.73, respectively. In Figure 4.4, we demonstrate the normalized confusion matrix regarding the selected gradient boosting classification algorithm combined with an SVM RFE feature selection technique, where the percentage of the true negative predicted cases is 73%, whereas the percentage of the true positives cases is 87%.

In addition to this, in Table 4.5 the respective performance metrics over the different folds and their mean and SD values using only non-imaging data are shown. The average balanced accuracy of the

predictive model trained only by non-imaging features is 0.69, while its sensitivity and specificity are both 0.69.

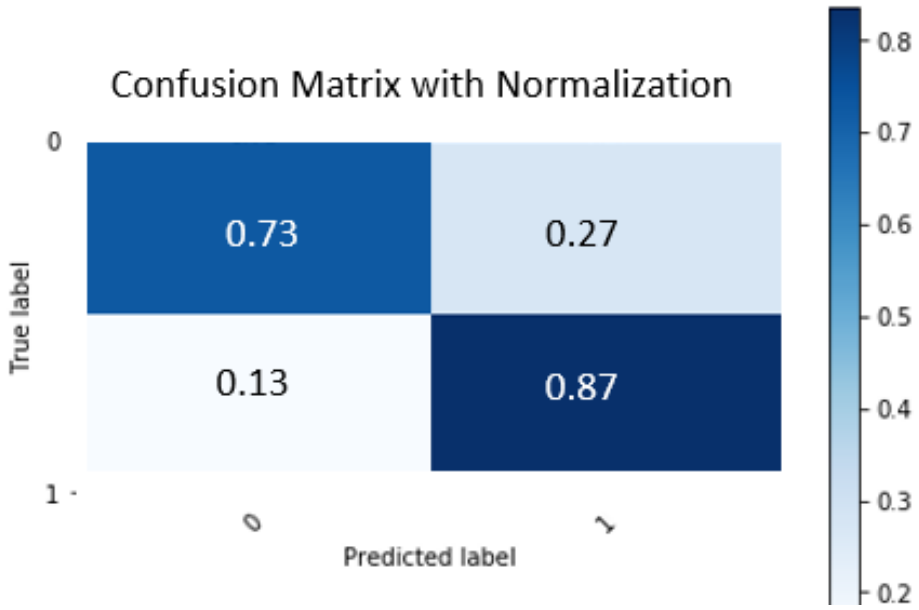

Figure 4. 4 Confusion Matrix

Table 4.4 Evaluation of the CAD risk-prediction problem over 10-fold using imaging and non-imaging data.

Folds	**Balanced Accuracy**	**NPV**	**PPV**	**ROC AUC**	**Sensitivity**	**Specificity**
Fold #0	0.73	0.78	0.67	0.60	0.67	0.78
Fold #1	0.75	0.86	0.63	0.82	0.84	0.67
Fold #2	0.89	1	0.72	0.92	1	0.78
Fold #3	0.69	0.78	0.6	0.72	0.6	0.78
Fold #4	0.84	1	0.63	0.83	1	0.67
Fold #5	0.84	1	0.63	0.89	1	0.67
Fold #6	0.95	1	0.84	1	1	0.89
Fold #7	0.78	1	0.56	0.78	1	0.56
Fold #8	0.8	0.86	0.72	0.88	0.84	0.75
Fold #9	0.8	0.86	0.72	0.75	0.84	0.75
Mean ± SD	0.81 ± 0.08	0.92 ± 0.1	0.68 ± 0.08	0.82 ± 0.11	0.88 ± 0.15	0.73 ± 0.09

Table 4.5 Evaluation of the CAD risk-prediction problem over 10-fold using only non-imaging data.

Folds	Balanced Accuracy	NPV	PPV	ROC AUC	Sensitivity	Specificity
Fold #0	0.5	0.6	0.4	0.33	0.33	0.67
Fold #1	0.56	0.64	0.5	0.65	0.33	0.78
Fold #2	0.72	1	0.5	0.89	1	0.44
Fold #3	0.69	0.78	0.6	0.69	0.6	0.78
Fold #4	0.79	0.88	0.67	0.87	0.8	0.78
Fold #5	0.78	1	0.56	0.78	1	0.56
Fold #6	0.79	0.88	0.67	0.8	0.8	0.78
Fold #7	0.72	1	0.5	0.76	1	0.44
Fold #8	0.6	0.64	0.67	0.79	0.33	0.88
Fold #9	0.71	0.75	0.67	0.83	0.67	0.75
Mean ± SD	0.69 ± 0.1	0.82 ± 0.16	0.57 ± 0.1	0.74 ± 0.16	0.69 ± 0.28	0.69 ± 0.15

Additionally, a SHAPley Additive exPlanations (SHAP) analysis was implemented for explaining the prediction of the proposed model by computing the contribution of each feature to the prediction [246]. The most important predictors of the proposed model are presented in Figure 4.5, below. Mean absolute SHAP values for the 10 most significant features are estimated to illustrate the global feature importance. As it is shown in Figure 4.5, the most significant feature is the number of the existing CP and the highest coronary DS at the baseline step. In addition to this, input features such as pro-brain natriuretic peptide (NT-proBNP), MMP-2 and 9, leptin, LDL, and patient characteristics such as weight, age, and height are highly ranked as significant features for the prognosis of CAD.

In addition to this, in Figure 4.6 below, we demonstrate the global interpretability of the proposed model by representing how much each input feature, either positively or negatively, contributes to the target variable. In Figure 4.6, we show with yellow columns the input features that contribute positively to the output target (detection of Class 2, CAD class). On the other hand, with blue columns, we indicate the input predictors that contribute negatively to the output target (detection of Class 1, no-CAD class). As it is shown in Figure 4.6, most of the input features contribute negatively to the output target and contribute to the prognosis of Class 1. Indicatively, the most significant features that contribute positively to the output target are thyroid stimulating hormone, medication therapy of beta blockers, aspartate aminotransferase, DM, and MLA. In the presented model, we observe that the most significant predictor for the prognosis of CAD is thyroid stimulating hormone, which confirms the effect of the thyroid hormones on the CV system [247, 248]. Thyroid hormone is considered as a significant regulator of CV system function and hemodynamics through different mechanisms. More specifically, inadequate thyroid hormone levels impair the relaxation of vascular SMCs and decrease cardiac contractility by regulating calcium uptake and the expression of

several contractile proteins in cardiomyocytes. Additionally, low thyroid hormone levels also increase systemic vascular resistance and induce endothelial dysfunction by reducing nitric oxide availability [249, 250]. As for the imaging-based input predictors, MLA has the most significant positive effect on the proposed model. As it is shown in Figure 4.6, the most significant feature with negative effect on the output is the number of the CP at the baseline analysis of patient imaging. Different studies in the literature have confirmed the prognostic capability of the presence of CP [251]. Calcification of the coronary arteries plays a key role in the pathophysiology of atherosclerosis, and these lesions are considered advanced lesions [252]. In addition to this, patient height contributes negatively to the prognosis of the output target, confirming the genetic relationship between height and CAD [253, 254]. As for the biochemical predictors for the no-CAD class (Class 1), we observed that pro-brain natriuretic peptide, LDL, and MMP- 2 have a high negative effect on the output target.

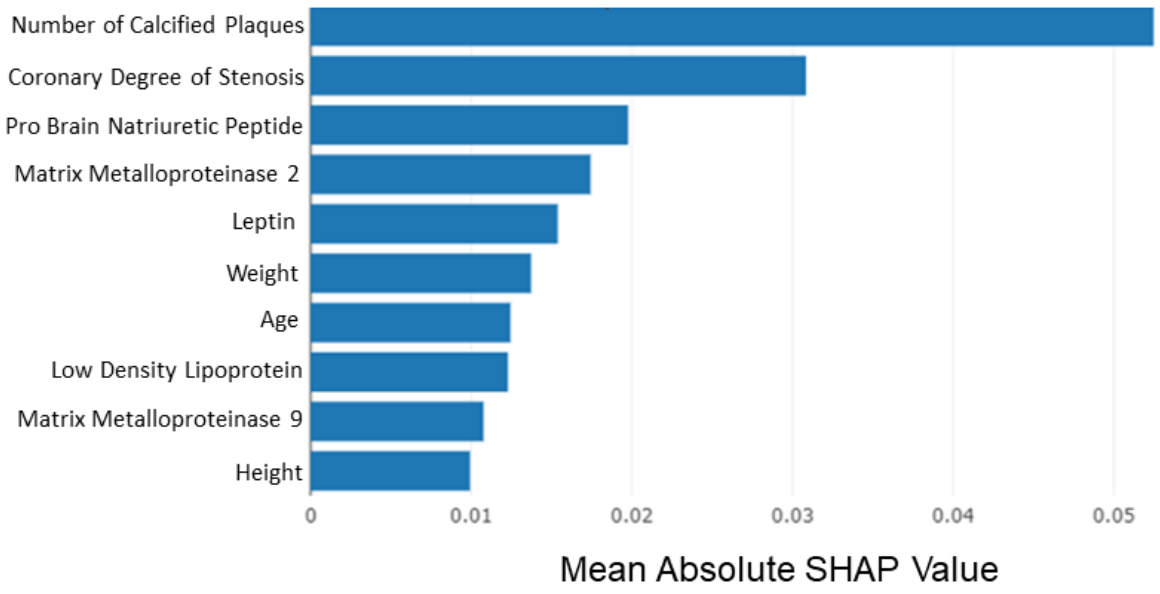

Figure 4.5 Feature importance based on mean SHAP values. The number of the existing CP and the highest coronary DS are indicated as the most significant features.

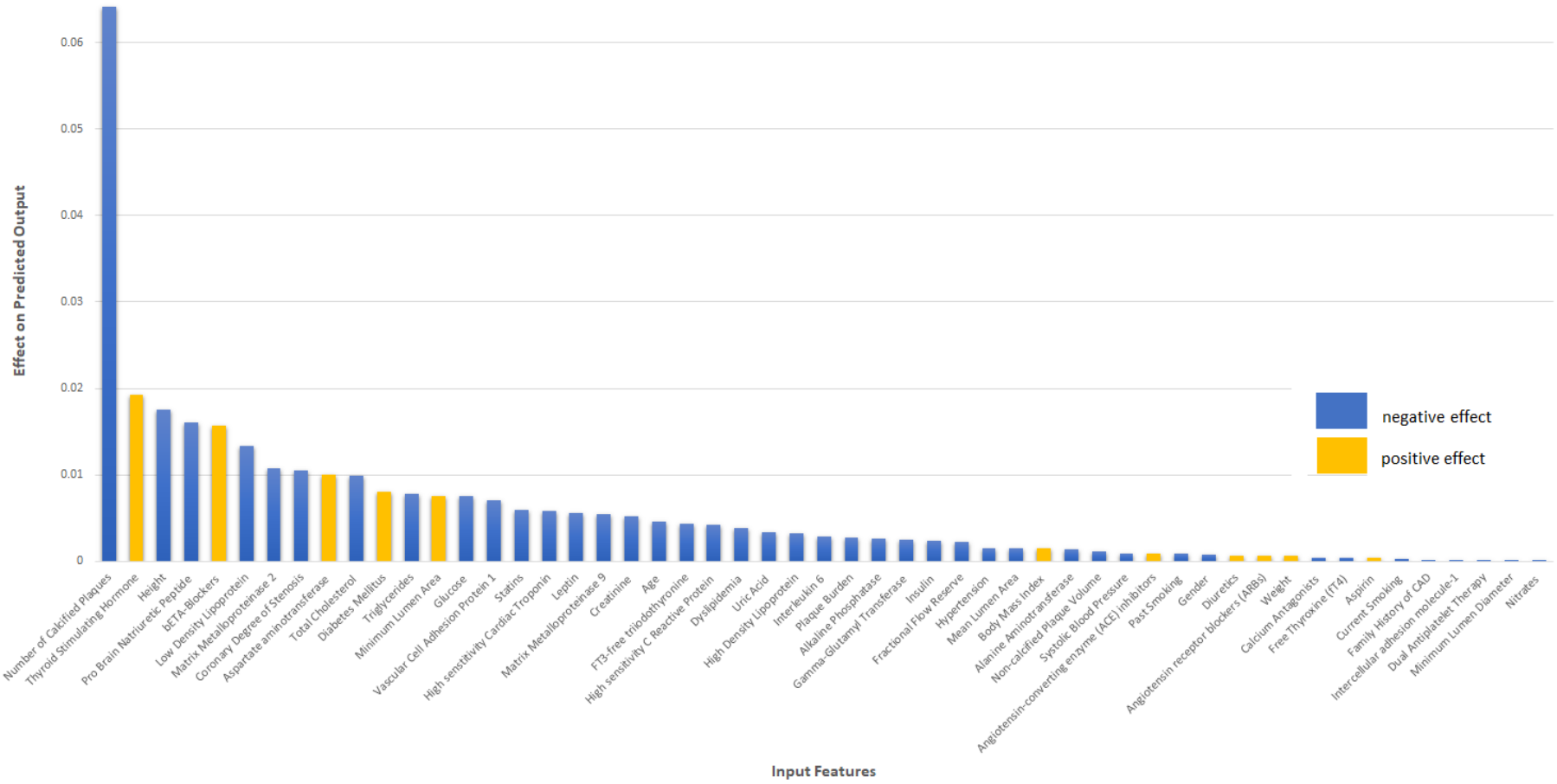

Figure 4.6 Input Features Contribution Table (blue, features with negative effect; yellow, features with positive effect). The most significant features that contribute positively to the output target are thyroid stimulating hormone, medication therapy of beta blockers, aspartate aminotransferase, diabetes, and minimum lumen area, whereas the most significant feature with negative effect on the output is the number of the CP at the baseline analysis of patient imaging.

4.3 Prediction of the progression of coronary artery disease using machine learning

4.3.1 Methodology

4.3.1.1 CTA image acquisition and analysis

The geometrical vasculature features utilized in the proposed ML based model are estimated based on an already published methodology for the 3D reconstruction of coronary arteries, which is described in detail in Chapter 3 [204, 205].

4.3.1.2 Problem Definition

The principal utilized classes C_i are namely "No CAD", "Non obstructive CAD" and "Obstructive CAD". C_0 is characterized by the absence of coronary stenosis, while C_1 is characterized by stenosis <30% at any vessel or stenosis between 30% and 50%. All the participants with at least one stenosis > 50% were considered of the obstructive CAD class C_2. At the follow-up time step 18 (37.5%) participants were at class C_0-No CAD, 17 (35.4%) at class C_1-non obstructive CAD and 16 (33.3%) were at class C_2-Obstructive CAD.

4.3.1.3 Feature Selection

In the proposed study, we employed feature selection techniques, aiming to reduce the dimension of input features and identify the redundant features. More specifically, we implemented the gain ratio algorithm (case 1), the principal components analysis (case 2) and the attribute evaluation technique (case 3). The progressive improvement of the sensitivity and specificity ratio is achieved through a proper customization of the input features, accompanied by a forward selection procedure [255-257].

4.3.1.4 Classification

After feature selection algorithms implementation, we examined four different ML algorithms; a parametric model [artificial neural network (ANN)], a non-parametric kernel-based model SVM, an ensemble model [(RF)] and J48 [255-257].

4.3.1.5 Evaluation

In terms of evaluation, we apply the ten-fold cross validation, which splits the initial dataset into ten subsets, whereby the nine subsets are used for training and the remaining one subset is used for testing. Moreover, it should be noted that the feature selection techniques are repeated in every ten-fold

repetition, ensuring that the feature selection procedure is based exclusively on the training dataset. In addition to this, it should be highlighted that our initial dataset is considered as a balance dataset, including almost same portions of patients of the three classes.

The performance of each classification scheme is quantified using as evaluation metrics the sensitivity and the specificity of each class, which denote the model's ability to quantify the positive and the negative results, respectively.

4.3.2 Dataset

Patient-specific information have been collected from retrospective data recorded during the EVINCI study [258] and are utilized as the baseline information. In the presented study, a dataset of 48 patients was used, who underwent CTA imaging to diagnose their risk of CAD and evaluate the percentage of coronary vessels' stenosis. This assessment has been reperformed after 5±2 years during SMARTool follow-up re-evaluation. Except for CTA imaging, in both time- slices, baseline and follow-up, a variety of data was collected and analysed, including clinical history and lifestyle of each patient, as well as molecular systemic variables and inflammatory and monocyte markers.

Based on the assembled dataset, we aim to identify the factors that affect the progression of atherosclerosis. More specifically, in this study due to the relatively small number of patients, we focused only on vessel's geometrical features based on the CTA imaging modality in the baseline step, as well as on the patient's medical history and lifestyle. In Table 4.6, we present the utilized feature set.

Table 4.6 Variables used in the current proposed model.

Category	**Features**
Geometrical vasculature	DS, MLA, MLD, PB, CP Volume, NCP Volume
Risk factors	Family History of CAD, Hypertension, DM, Dyslipidemia, Smoking, Obesity, Metabolic Syndrome, Past smokers

4.3.3 Results

In Table 4.7, we present the obtained sensitivity and specificity for each of the three different classes, after implementing the classification schemes. As far as the class of non-obstructive CAD is concerned, its sensitivity is observed lower than these of the other classes.

Table 4.7 Classification Performance.

		Class No CAD		**Class Nonobstructive CAD**		**Class Obstructive CAD**	
		Sensitivity	**Specificity**	**Sensitivity**	**Specificity**	**Sensitivity**	**Specificity**
Case 1	SVM	0.67	0.73	0.42	0.81	0.44	0.85
	ANN	0.6	0.79	0.49	0.65	0.42	0.78
	RF	0.69	0.85	0.48	0.71	0.47	0.75
	J48	0.67	0.73	0.42	0.81	0.63	0.75
Case 2	SVM	0.54	0.82	0.53	0.59	0.44	0.85
	ANN	0.8	0.79	0.48	0.88	0.57	0.75
	RF	0.6	0.82	0.53	0.78	0.63	0.79
	J48	0.6	0.82	0.58	0.81	0.59	0.82
Case 3	SVM	0.74	0.85	0.48	0.71	0.46	0.75
	ANN	1	0.5	0.42	0.91	0.32	0.97
	RF	0.74	0.76	0.36	0.81	0.63	0.79
	J48	0.87	0.79	0.36	0.91	0.82	0.82

It is noted that in case we implement the J48 classification algorithm, after the class attribute evaluation feature selection technique, we achieve a sensitivity of 0.87 and 0.82 for the healthy class, the No CAD class and the unhealthy class, the Obstructive class, respectively. The confusion matrix of the aforementioned approach is showed in Table 4.8. In addition to this, the ranking of the selected features, maintained by this classification scheme is illustrated in Table 4.9 and the best performance of the proposed model is achieved by maintaining the eight first ranking features.

Table 4.8 Confusion matrix of Case 3-J48.

		Prediction		
		Class No CAD	**Class Nonobstructive CAD**	**Class Obstructive CAD**
Annotation	**No CAD**	13	1	1
	Nonobstructive CAD	6	6	5
	Obstructive CAD	1	2	13

Table 4.9 Features ranking selection by Case 3-J48.

	Features
Class Attribute evaluator	Current symptoms [142]t
	volume of the most significant calfified plaque (mm^3)
	volume of the most significant NCP (mm^3)
	existence of NCP {0,1}
	Minimal lumen area (mm^2)
	Past smokers {0,1}
	existence of CP {0,1}
	MLD (mm)

4.4 Site specific prediction of atherosclerotic plaque progression using machine learning

4.4.1 Methodology

4.4.1.1 CTA analysis & 3D reconstruction

The coronary artery 3D reconstruction was performed using an already validated semi-automated methodology, which is described in detail in Chapter 3 [204, 205].

4.4.1.2 Blood Flow & Mass Transport Modeling

Blood behaves as a Newtonian fluid and its flow assumed to be laminar and incompressible, modeled by Navier-Stokes equations and the continuity equation as follows:

$$-\mu\nabla^2 u_l + \rho(u_l \cdot \nabla)u_l + \nabla p_l = 0 \tag{4.14}$$

$$\nabla \cdot u_l = 0 \tag{4.15}$$

where μ is the dynamic viscosity of blood, u_l is the blood velocity, ρ is the blood density , and p_l is the pressure. LDL was assumed not to influence the blood flow and LDL transport in the lumen was approximated by the convection-diffusion equations that follows:

$$\nabla \cdot (-D_l \nabla c_l + c_l u_l) = 0 \tag{4.16}$$

where D_l is the luminal LDL diffusivity, c_l is the luminal LDL concentration and u_l is the blood velocity, as previously described. Due to the oxidation of LDL particles, a term of degradation is added in its convection diffusion equation, as follows:

$$\nabla \cdot \left(-D_w \nabla c_w + K_{lag} c_w u_w\right) = r_w c_w \tag{4.17}$$

where D_w is the diffusivity of the solute in the arterial wall, c_w is the LDL concentration in the wall, K_{lag} the solute lag coefficient for the LDL, and $r_w$is the consumption rate constant.

In Figure 4.7, we illustrate the a) 3D luminal reconstruction, b) the endothelial shear stress (ESS) distribution and the c) normalized subendothelial LDL concentration.

4.4.1.3 Problem Definition

The atherosclerotic disease progression was formulated as a two-class problem, the "no significant progression class" ($Class_0$) and the "significant progression class" ($Class_1$). The factors which define the atheromatic plaque progression were the lumen change, the PB change and the plaque change. Significant progression was considered, where at least two of these factors increase higher than 20%. The proposed model was trained using 1018 artery segments, where 691 were at $Class_0$ and 327 were at $Class_1$.

4.4.1.4 Dataset Preprocessing

The preprocessing step, which refers to the transformation of data before feeding them to the algorithm, aims to enhance the quality of the input features in order to achieve a high prediction accuracy. In this step, we implement a class imbalance handling technique, the Synthetic Minority Oversampling Technique (SMOTE) algorithm, [259], which oversamples the minority class with a percentage of 20%. Due to synthetic samples, SMOTE method avoids over-fitting largely and does not avoid potentially useful samples.

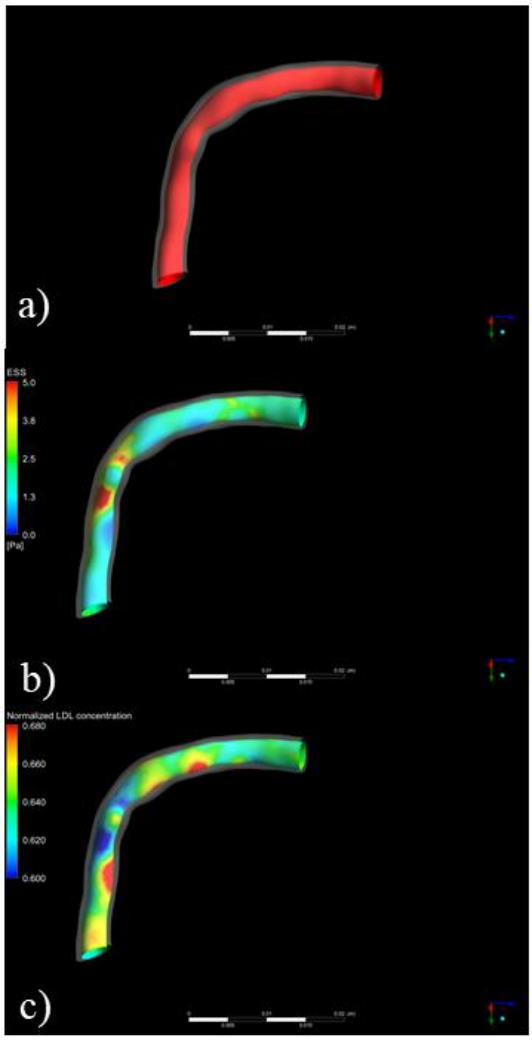

Figure 4.7 a) 3D reconstruction, b) ESS distribution at the luminal surface, c) normalized subendothelial LDL concentration [234].

4.4.1.5 Classification-Prediction

In this step, tree based ML schemes are implemented to predict the progression of atheromatic plaque in each coronary artery segment. Tree based prediction models are considered to be one of the best and mostly used supervised learning methods. In our approach 5 different modifications of tree based algorithms are implemented, the J48, the logistic model tree (LMT), the RF, the random tree (RT) and finally the RepTree algorithm.

In addition to this, ranking feature selection techniques are incorporated to the classification schemes in order to reduce the input dimensionality and identify the redundant features. More specifically, we implemented the ingogain ratio algorithm (Case 1), the principal components analysis (Case 2) and the relief ranking technique (Case 3).

4.4.1.6 Evaluation

The performance of the utilized ML based models is evaluated applying the ten-fold cross validation. Ten-fold cross validation divides the initial dataset into ten subsets in every ten-fold repetition and utilizes the nine subsets for training and the remaining one dataset for testing. The performance of each classification scheme is quantified using as evaluation metrics the sensitivity and the specificity of each class, which denote the model's ability to quantify the positive and the negative results, respectively.

4.4.2 Dataset

The utilized in this study baseline data was based on EVINCI study [244], in which patients underwent CTA imaging for clinical purposes and patient-specific data was collected, whereas the follow-up data have been assembled after 5±2 years during SMARTool project [260]. More specifically, 1018 coronary segments of 3 mm, derived from 40 reconstructed coronary arteries, where blood flow modeling was simulated. The utilized input features were imaging based and biomechanical features and included the lumen area (mm^2), the outer wall area (mm^2), the PB area, the maximum ESS (Pa), the minimum ESS (Pa), the mean ESS (Pa), the luminal area where ESS was lower than 1 Pa (mm^2), the maximum LDL concentration and the minimum LDL concentration. These values were calculated as the mean values of 6 0.5 mm cross-sections.

4.4.3 Results

In Table 4.10, we demonstrate the accuracy and the AUC for each classification scheme. The obtained results are very promising for all the different tree-based algorithms. However, RF outperforms the other models' performance and achieves an accuracy of 0.84 and an AUC of 0.91. Moreover, the most highly ranked input features after the implementation of Relief ranking technique [235] are the normalized maximum and minimum LDL concentration, and the luminal area where the WSS is lower than 1 Pa.

Table 4.10 Performance of the utilized prediction models.

		Tree based prediction models				
		J48	**LMT**	**RF**	**RT**	**RepTree**
No feature selection	Acc.	0.8	0.78	0.83	0.75	0.77
	AUC	0.83	0.83	0.91	0.75	0.82
Case 1	Acc.	0.789	0.78	0.84	0.76	0.78
	AUC	0.82	0.83	0.9	0.76	0.82
Case 2	Acc.	0.72	0.72	0.8	0.73	0.74
	AUC	0.74	0.77	0.87	0.73	0.78
Case 3	Acc.	0.79	0.79	0.84	0.76	0.77
	AUC	0.82	0.84	0.91	0.76	0.82

4.5 Site specific prediction of PCI stenting using machine learning

4.5.1 Methodology

4.5.1.1 CTA analysis & 3D reconstruction

CTA images were analyzed using a validated automated methodology, [204, 205], which is described in detail in Chapter 3. The whole procedure is incorporated in a user-friendly software tool, which requires the minimum user interaction and provides automatically not only the 3D models of coronary inner and outer wall and the CP and NCP in the 3D space, but also it provides geometrical artery features, such as the DS, and biomechanical features, such as the FFR index.

4.5.1.2 Calculation of SMARTFFR

The process of calculation of SMARTFFR index is described in detail in chapter 4, in section 4.2.1.2.

4.5.1.2 Problem Definition

In this work, the PCI risk problem is formulated as a two class problem, representing the PCI procedure as a nonlinear parametric function of two imaging based features $f(x) = C_i,\ x = [x_1, x_2]$, where x_1 represent the coronary artery DS and x_2 represents the Fractional Flow Reserve index (SMARTFFR). The proposed predictive model was trained using 481 coronary segments, where 445 were at $Class_0$ −No PCI placement and 36 were at $Class_1$-PCI placement.

4.5.1.3 Easy Ensemble algorithm implementation-Class Imbalance Handling

Imbalanced medical datasets constitute a problem often found in medical prediction problems. In the proposed predictive model, we utilize the Easy Ensemble algorithm [261] for handling the class imbalance problem. The idea behind Easy Ensemble algorithm is to employ random resampling to generate K subsets of $\{Q_1, Q_2, \ldots, Q_K\}$ from Q ($|Q_i| < |Q|$, i=1,2,…K). Subsequently, each $Q_i \cup P$ is trained by the classifier and the final decision is selected by majority voting.

In the proposed predictive modeling approach Easy Ensemble idea is combined with the Extreme Gradient Boosting classifier and each individual model is trained by the Equations (4.18)-(4.20) presented below.

4.5.1.4 Extreme gradient boosting using tree ensembles

The extreme gradient boosting (XGB) algorithm was used as an optimized classifier that uses the gradient methods to build a tree ensemble model consisting of a real set CARTs. These are used as base learners to additively combine multiple tree predictions, and thus yields higher performance over the conventional trees [262]. For a given set of N-observations $(x) = \{x_i, y_i\}, i = 1,2, \ldots, N$, the goal is to add at each step t, a tree function:

$$f_t(x) = w\big(q(x)\big), w \in R^L, q: R^d \to \{1,2, \ldots, L\} \tag{4.18}$$

which improves the prediction performance of the model by minimizing the following regularized objective:

$$E(t) = \sum_{i=1}^{N} l\big(y_i, \tilde{y}_{i,t-1} + f_t(x_i)\big) + r, \tag{4.19}$$

where $l(.)$ is the loss function at step t, $\tilde{y}_{i,t-1}$ is the estimated value at step $t-1$, q is a function that assigns values to the tree leaves, and r is the regularization term that is used to avoid overfitting and reduce the model complexity [17, 18]. The regularization term is defined as follows:

$$r = \gamma L + \frac{1}{2}\lambda \sum_{i=1}^{J} w_j^2 \tag{4.20}$$

where w is the weight on the leaves, γ is a constant value, and L is the total number of leaves in each tree learner.

4.5.2 Dataset

Patient-specific imaging and non-imaging data were collected during the EVINCI study [244]. In this study totally 263 patients have enrolled and underwent CTA imaging for clinical purposes and this imaging assessment has been reperformed after 5±2 years during the SMARTool project [15]. A PCI stenting procedure and CABG were performed in 55 patients and 10 patients, respectively. In addition to this, CTA images of 11 patients were considered as uninterpretable and thus, in the proposed study, baseline imaging of 187 patients was analyzed to predict the risk of PCI placement. In this approach, totally 481 coronary segments were utilized for the prediction of PCI placement and the input of the proposed model was the

4.5.3 Results

The XGB classification scheme was validated using 10-fold cross validation and for each fold we present in

Table 4.11 the accuracy and the AUC. The mean value of accuracy was 0.78, whereas the mean AUC was equal to 0.87. The receiver operating characteristic (ROC) curve is depicted in Figure 4.8, where the AUC score was evaluated on different thresholds across each fold and averaged across all folds.

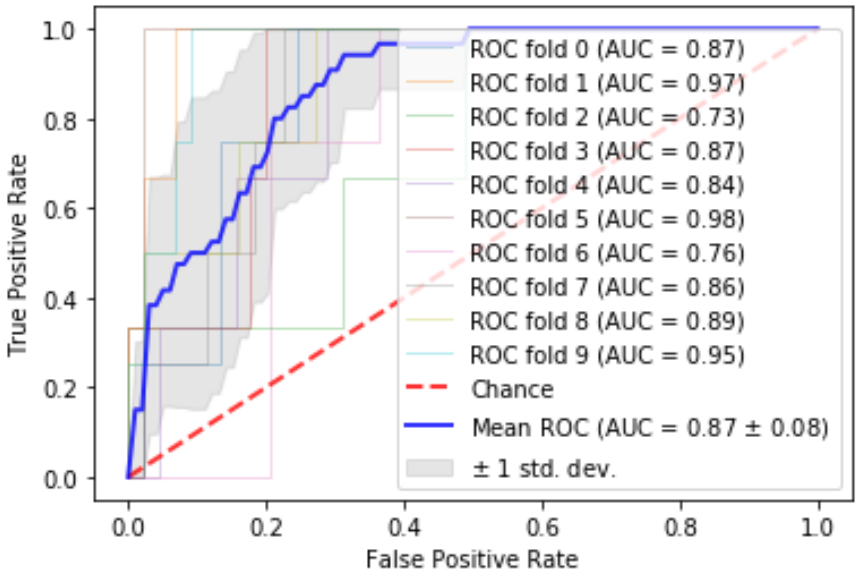

Figure 4.8 Prediction performance (sensitivity versus 1-specificity) of the XGB tree ensembles across each fold.

Table 4.11 Evaluation of the CAD risk prediction problem over the 10-folds.

Evaluation Metric		
Folds	***Accuracy***	***AUC***
Fold #0	0.88	0.87
Fold #1	0.88	0.97
Fold #2	0.52	0.73
Fold #3	0.85	0.87
Fold #4	0.82	0.84
Fold #5	0.88	0.98
Fold #6	0.40	0.76
Fold #7	0.83	0.86
Fold #8	0.77	0.89
Fold #9	0.89	0.95
Mean±std	0.78±0.17	0.87±0.08

4.6 Discussion

Regarding the first study, a novel approach for the prediction of obstructive CAD is presented. The aim of this study is to develop a ML model for the CAD risk prediction, which takes into account different types of data, including both imaging and non-imaging data. To our knowledge, our approach to combine the imaging and blood-flow-based characteristics with typical CAD risk factors constitutes the novelty of the presented study.

Different methodologies have been presented for the prediction of CAD and the identification of the major CAD risk factors. Most of these studies are concentrated on the different CAD-related risk-prediction outputs and are based either on statistical analysis [6,37,38] or ML classification schemes [263]. Our proposed study in comparison with these ones is more concentrated on the CAD risk prediction and its future presence and achieves a higher evaluation metrics, as it is shown in Table 4.12.

Additionally, recent studies have indicated that non-invasive CV imaging and especially the CTA imaging modality utilized in this study provides useful prognostic information of atherosclerosis progression since it permits the accurate quantification of luminal area and the detection of PB region and the characterization of its composition. Moreover, the overall PB, which can be provided by CTA

imaging, is highly relevant to the degree and characteristics of atherosclerosis [264]. In addition to this, the clinical relevance of the overall coronary PB has been also emphasized by studies showing that increased NCP volumes is directly linked with acute coronary syndrome patients [265]. Furthermore, the latest technological advancements in patient-specific blood-flow modeling have introduced alternative CAD progression risk factors, such as fractional flow reserve (FFR) index and WSS.

Table 4.12 Comparison of the proposed methodology with existing in the literature studies.

Study	**Methodology**	**Input**	**Results**
D'Agostino *et al.* [266]	Statistical analysis	Age, total cholesterol, HDL, Systolic BP, BP treatment, Smoking, Diabetes, Incident CAD events	C statistic: 0.763 (men), 0.793 (women)
Weng *et al*. [263]	ML classification schemes to predict CV events	Demographics, history of medical conditions, prescription drugs, acute medical outcomes, referrals to specialists, admissions to hospitals, biological results	sensitivity 67.5%, PPV 18.4%, specificity 70.7%, NPV 95.7%
Stone *et al.* [115]	Statistical analysis	Demographics and Clinical Characteristics, Medical Therapy, endothelial shear stress, atherosclerotic plaque characteristics, based on CA and IVUS	41% positive predictive value and 92% negative predictive value
Liu *et al*. [118]	Statistical analysis	WSS, von Mises stress (VMS) based on CTA imaging	-
Our study	Gradient Boosting Classification with RFE	Patient's Characteristics, CAD Risk Factors, biohumoral markers, imaging based features, SmartFFR index	81% accuracy, 88% sensitivity, 73% specificity

The prognostic capability of CTA imaging modality and its derived imaging features has also been confirmed by the proposed study, in which the overall accuracy of the proposed predictive model using both imaging and non-imaging data is 0.81. Moreover, the prognostic significance of imaging-derived features is also indicated by the collected results, shown in Table 4.4. More specifically, the predictive model trained by the non-imaging-based features achieved a comparatively lower accuracy of 0.69.
Furthermore, another notable point of the proposed CAD risk-predictive model is that the input geometrical features are derived by an automated CTA image analysis tool [204, 205], able to detect

accurately the inner and outer wall and atherosclerotic plaques and provide an accurate 3D model of coronary arteries and the atherosclerotic plaques distribution over the 3D space. As far as the SMARTFFR index is concerned, it is also calculated automatically by the developed software tool in the 3D reconstructed coronary artery.

In addition to this, another innovative aspect of the presented predictive model is the implementation of the easy ensemble algorithm, which constitutes a random resampling scheme, which mainly handles the class imbalance problem. Except for the class imbalance handling, the applied easy ensemble scheme allows the progressive correction of the model's decision hyperplane and subsequently the reduction of the classification error. Additionally, the predictive capability of the proposed model is evaluated based on nested stratified cross-validation, which provides unbiased estimation of the predictive model's capability. Moreover, except for the innate hyperparameters of the classification algorithm, the input features are also treated as a hyper-parameter, and SVM RFE feature selection technique is implemented to eliminate the input features' dimension. The particular ML algorithm was selected after the implementation of different classification schemes in combination with different feature selection techniques, and the highest accuracy was provided by the combination of the extreme gradient boosting algorithm and the SVM feature selection technique.

As for the second presented approach, multiple models were performed to predict the progression of CAD, taking into account non-invasive image-based features and traditional atherosclerosis risk factors, such as the medical history and lifestyle. The combination of the input features as well as the problem formulation as a three class problem, integrating the graduation of atherosclerosis, constitute a crucial novelty of the presented study. This study is considered as a preliminary approach and it is important to further validate our results in larger datasets.

After the implementation of different classification schemes, we conclude that the tree-based algorithms and more specifically C4.5 (J48) algorithm combined by a ranking feature selection method, outperforms the other classification models [257]. In general, tree-based algorithms constitutes a typical type of machine-learning approaches and are able to handle the non-linearities, the heterogeneous data, and many predictors, by searching through the input variables to find this one, which seperates the outcome into two groups [267].

In this point, it should be highlighted that the ability to foresee a future health condition of the most patients is a significant aspect of the proposed model. The sensitivity for the first class is 0.71±0.13, based on the different implementation, whereas for the most accurate model is 0.87. In this manner, the

patient may avoid undergo a multiple CTA imaging in the follow-up time slice step, considering the risks related to radiation from medical imaging. In addition to this, another aspect of the proposed study worth mentioning is the model's ability to accurately identify the non-healthy patients, those of the third class. In the last case of J48 algorithm, the sensitivity is 0.82. An accurate prediction of such a status contributes significantly to the clinical patient management, allowing the timely diagnosis of CAD and the safe selection of treatment.

Through the third approach, we investigate the role of the biomechanical forces acting within the artery and we investigate how these forces may affect the atheromatic plaque progression. More specifically, the developed model is based on imaging and biomechanical features, derived by a noninvasive imaging technique and ignoring the systematic atherosclerosis risk factors.

The basic novel aspect of the proposed approach is the implementation of ML based models in order to address the new analytic challenges in the CAD risk prediction. More specifically, the utilized tree based prediction schemes search among the utilized input predictors to find the variable that best separates the outcome. In this manner, tree based models are capable of handling the non-linearities, the heterogeneous data, and many predictors, by searching through the input variables to find this one, which seperates the outcome into two groups [267].

In comparison with the existing in the literature studies [114, 115], the data utilized in our approach derive by a non-invasive and thus, the double inference of the proposed study is the ability of CTA to visualize accurately the coronary artery anatomy and its clinical importance in the prediction of atheromatic plaque progression.

In the fourth study, we examine the predictability of two imaging based features, the coronary artery DS and the $_{SMART}$FFR index to identify the crucial coronary lesions. An overall accuracy 0.78±0.17 was achieved, implementing a class imbalance handling algorithm, the Easy ensemble technique and a ML classification scheme, the XGB. In general tree based classification algorithms are extensively implemented in medical prediction problems and constitute a typical classification scheme, able to handle non-linearities aiming to separate the outcome into two groups. More specifically, a tree ensemble method was used to develop a supervised learning model for the prediction of a site specific PCI stenting placement. The main idea behind gradient boosting trees is the use of first and second order gradient statistics to improve the overall performance of the prediction over a set of individual decision trees. In addition to this, the optimal solution across each split, that ensures that the most value split is maintained at each boosting step is based on an exact greedy algorithm.

Apart from the XGB algorithm predictability, through this study it is clear that CTA imaging and biomechanics based features have a significant prognostic value and its non-invasive nature of CTA imaging modality highly contributes to the CV research area. To our knowledge, the predicted outcome of the presented study constitutes the novelty of this study.

Chapter 5 Carotid artery disease prediction

5.1 Introduction

5.2 Methodology

5.3 Diagnostic Prediction of Carotid Artery Disease

5.4 Diagnostic Prediction of the vulnerability of Carotid Artery Plaques

5.5 Discussion- Beyond the state of the art

5.6 Conclusions

5.1 Introduction

Except of CAD risk prediction, carotid artery disease risk stratification plays also a crucial role for the prevention and management of CVDs. The management of carotid artery disease in the current clinical practice in based on the 2017 Clinical Practice Guidelines of the European Society for Vascular Surgery (ESVS) and the selection of management of carotid artery disease is based on the degree of CAS.

In this chapter, our aim is to present ML based models for the identification of carotid artery disease high risk patients, based on non-imaging data. Available data such as: (i) demographics, (ii) risk factors, (iii) circulating biomarkers, (iv) clinical data and (v) medication therapy will provide input to a series of ML risk stratification and prediction models including:

- Diagnosis of carotid artery disease presence
- Diagnosis of high risk vulnerable plaque

5.2 Methodology

5.2.1 Statistical analysis

Descriptive Statistics

The *per*-class sampling distribution of each continuous variable was described by the arithmetic mean and SD values, whereas the distribution of the categorical variables within each class was described by the corresponding percentages.

Null Hypothesis Significance Testing

The statistical significance of the utilized feature sets, with respect to the defined problems, was evaluated by appropriate parametric and nonparametric univariate two-sample statistical tests, based on the type (continuous numeric or categorical) and the *per* class distribution (Gaussian or non-Gaussian) of the dependent variables, as it is shown in Table 5. 1. The assumption of normality was verified using the Shapiro-Wilk test. In case of k-sample statistical testing, a multiple pairwise comparison of the group means is subsequently applied using the Tukey's honestly significant difference criterion.

5.2.2 Machine learning predictive modeling

ML modeling utilizes two different techniques: the supervised learning and the unsupervised learning. Supervised learning trains a model based on known input and output data to predict the future output given new input data, whereas unsupervised learning aims to investigate hidden patterns or intrinsic structures based on input data.

More specifically, in this section we present a supervised ML model that makes predictions, taking into consideration set of input data and known responses to the data output and trains a model to generate reasonable predictions for the response to new data.

The development of the predictive models overall training and evaluation procedure based on ML techniques is shown in Figure 5.1.

Table 5. 1 Univariate Null hypothesis Significance Tests.

Statistical Test		**Independent Variable Type**	**Dependent Variable Type**	**Null Hypothesis**
Fisher's exact test	Two-sample test	Binary	Binary	Independence (no non-random associations) between two binary variables by estimating exactly the conditional distributions (2×2 contingency table).
Pearson Chi-square test (χ^2) with one degree of freedom	Two-sample test	Binary	Categorical	Independence (no non-random associations) between two categorical variables when the comparing groups are independent.
Student t-test	Two-sample parametric test	Binary	Numerical	Equality of means and variances of two independent data samples; a Gaussian distribution is assumed.
Mann-Whitney U test	Two-sample non-parametric test	Binary	Numerical	Equality of medians of two independent data samples; normality of data is not assumed.
Pearson Chi-square test with more than one degrees of freedom	k-sample test	Categorical	Categorical	Independence (no non-random associations) between two categorical variables when the comparing groups $(k > 2)$ independently
One-way analysis of variance (ANOVA)	k-sample parametric test	Categorical	Numerical	Equality of means of $k > 2$ independent data samples; a Gaussian distribution is assumed.
Kruskal-Wallis test	k-sample non-parametric test	Categorical	Numerical	Equality of distributions of $k > 2$ independent data samples; normality of data is not assumed.

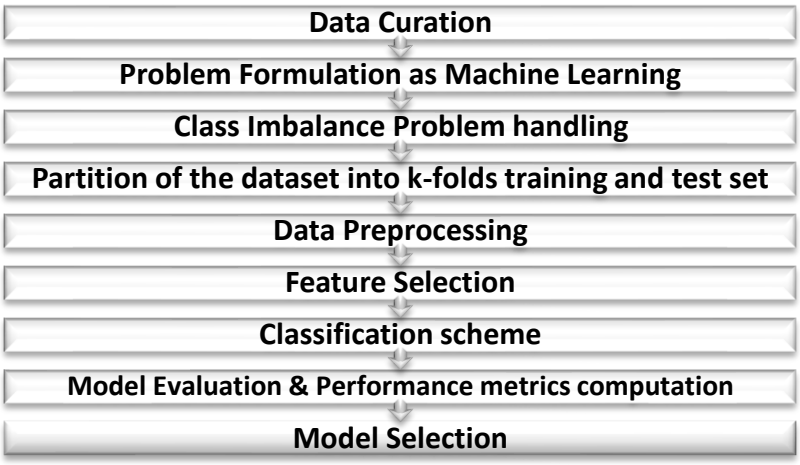

Figure 5. 1 Pipeline of Predictive Models Overall Training and Evaluation Procedure based on Machine Learning Techniques.

5.2.2.1 Data Curation

In the first step, a data quality assessment procedure is implemented to effectively manage the quality of the data. Data cleaning procedure, also known as data curation procedure, is considered as a key aspect, prior to the development of any data analytics services. In this step, the data curation framework proposed by Pezoulas *et al*. [268] is implemented in each utilized dataset and the framework outputs two different documents: (i) the data quality assessment report, and (ii) the curated dataset. In Figure 5.2, we show the output of the implemented data curation framework.

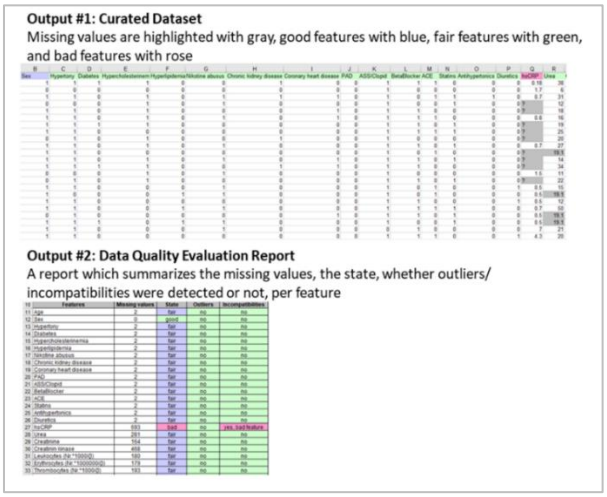

Figure 5.2 Curation Procedure Output.

5.2.2.2 Problem Formulation as Machine Learning

The carotid artery disease risk stratification problem has been formulated as a two class classification problem, based on the presence of CAS, individuals' symptoms and the vulnerability of the carotid atherosclerotic plaques.

This overall predictive supervised learning approach aims to learn a mapping from input features x to output Y, given a labeled set of input output pairs $D = \{(x_i, y_i)\}_{i=1}^{N}$, where D is the training set and N is the number of training examples [240]. Each sample (x_i, y_i) associates the input features with the carotid artery disease risk prediction Y, where $Y \in \{C_1, C_2\}$, is estimated by a non-linear parameterized function (f) of input features $x \in R^d$, $x = [x_1, x_2, \dots, x_d]$. The goal of this supervised classification problem is to obtain an approximation $F(x)$ of the function $F^*(x)$ mapping the input x to output Y. Table 5.2 outlines the model selection, identification and evaluation process given the model M and the dataset Z. An external cross-validation is applied to obtain a less biased estimation of model's performance.

5.2.2.3 Class Imbalance Problem handling

Medical datasets are often not balanced in their class labels. Most existing classification methods tend to perform poorly on minority class examples when the dataset is extremely imbalanced. In the development of this predictive model, we utilize the Easy ensemble technique. The Easy Ensemble algorithm [241] is a class imbalance handling approach, in which P are the training instances of the minority class, whereas Q denotes the instances of the majority class. The idea of Easy Ensemble algorithm is to employ random resampling to generate K subsets of $\{Q_1, Q_2, \dots, Q_K\}$ from Q ($|Q_i| < |Q|$, i=1,2,…K). Subsequently, each $Q_i \cup P$ is trained by the classifier and the final decision is selected by majority voting.

Let us denote by P and Q the training samples corresponding to the minority and majority class, respectively. EasyEnsemble algorithm: (i) employs bootstrap resampling to generate T subsets Q_1, Q_2, K, Q_T from Q ($|Q_i| < |Q|, i = 1, K, T$), (ii) trains a classifier M_i for each $Q_i \cup P$, $i = 1, K, T$, and (iii) combines via majority voting their individual decisions [241] (Table 5.3).

Table 5.2 CAD Risk Stratification Model Training and Evaluation Procedure.

Input: X, Model M

1. Data Curation

Cross-validation:

2. if the dataset X exhibits the Class Imbalance Problem

   $\boldsymbol{M} \equiv$ classifier encapsulated into an Easy Ensemble repeated under-sampling scheme

   else

   $\boldsymbol{M} \equiv$ selected classifier

3. Randomly partition X into 10 disjoint stratified folds X_k, with $k = 1, k, 10$, of equal size (i.e. $N/10$).
4. For $k = 1, k, 10$:
   - **a.** Let $X_{train} = X - X_k$ be the training set and $X_{test} = X_k$ the test set.
   - **b.** Data-preprocessing
     - i) One Hot encoding of categorical features.
     - ii) Imputation of missing values.
   - **c.** Feature Ranking
     - i) Apply the balanced random forest algorithm to X_{train} so as to produce a ranked list of features R_k.
   - **d.** Model Selection and Identification
     - i) Optimize model's M_k hyper-parameters on X_{train} by grid-search over 3-fold cross-validation.
     - ii) Compute $\theta_k \in R^m$ on X_{train} based on the selected classifier.
   - **e.** Model Evaluation
     - i) Evaluate model's M_k performance on X_{test} by classification performance metrics (e.g. sensitivity, specificity, negative predictive value).
5. Compute the average performance metrics over the 10-fold cross-validation for the training dataset and the external dataset.

Table 5.3 The Easy Ensemble Algorithm [241].

Input: P, Q, T

1. for $i = 1, K, T$
   - a. Randomly sample a subset Q_i from Q, $|Q_i| = |P|$.
   - b. Learn M_i using P and Q_i. M_i is a classifier with M base learners $F_{m,i},\ m = 1, K, M$.

Output: A bag of T balanced classifiers, $\{M_i\}_{i=1,K,T}$.

5.2.2.4 Data Preprocessing

In this step, a one hot encoding procedure was implemented to represent all the categorical input features as binary vectors. In addition to this, a curation procedure was implemented to check our dataset both for outliers and missing values. All the input features whose missing values were higher than 10% were removed from the dataset, whereas features with missing values lower than 10% have been imputed by either the most frequent value (categorical type features) or the median value (numerical type features).

5.2.2.5 Feature Selection

Feature selection techniques aim to identify and remove the irrelevant and redundant attributes from the overall dataset that do not contribute to the accuracy of the predictive model. Different feature selection techniques, as described in detail in Table 5.4, have been implemented and tested for the development of the most accurate predictive model.

5.2.2.6 Classification scheme

Classification schemes implementation constitutes one of the most important aspect in supervised learning. In the overall pipeline, we implement different classification schemes, such as the Gradient Boosting, the XGB, the Adaboost, the C-Support Vector kernel based, the k-nearest neighbors, the Decision tree classifier, as described in detail in Table 5.4.

5.2.2.7 Model Evaluation & Selection

The evaluation of the expected generalization performance was based on nested stratified cross-validation which provides an unbiased estimation of the model's classification performance, with data-pre-processing and feature evaluation following the resampling procedure itself in order to reduce the selection bias in the estimates of the model's classification performance. We should mention that, besides the innate hyper-parameters of the classification algorithm, the input size is also treated as a hyper-parameter, which yields the minimum input size optimizing a specific criterion.

Table 5.4: Development of Predictive Models.

1) Data Curation	Data quality evaluation report (missing values, range values, types, outliers/incompatibilities, data standardization)	
2) Problem Formulation as Machine Learning	Classes $C \in \{C_1, C_2, \dots, C_n\}$ estimated by a non-linear parameterized function, F, of a confined set of features $x \in R$, such that $F(x) = C$	
3) Class Imbalance Problem handling	Easy Ensemble Algorithm (Randomly repeated under-sampling scheme)	
4) Partition of the dataset into k-folds training and test set	K stratified disjoint folds of equal size	
5) Data Preprocessing	• Encoding of categorical features • Imputation of missing values (<10%)	
6) Feature Selection	• RFE • RFE and cross-validated selection of the best number of features • Select From Model (considered unimportant and removed features under a specific feature_importances_ values) • RF, Select KBest (Select features according to the k highest scores)	• Select Fpr scheme (Select the pvalues below alpha based on a False Positive Rate test) • Select I (Select the p-values corresponding to Family-wise error rate) • Sparse Principal Components Analysis • Nonnegative Matrix Factorization feature selection • Linear Discriminant Analysis
7) Classification scheme	• Gradient Boosting • XGBoost • AdaBoost	• C-Support Vector kernel based • k-nearest neighbors • Decision tree classifier
8) Model Evaluation and Performance metrics computation	• Accuracy • Sensitivity • Specificity • PPV	• NPV • ROC AUC • True Negative • False Positive
9) Model Selection	• Exhaustive search over specified parameter values for an estimator • Randomized search on hyper parameters	

5.2.3 Tuning of the cut off threshold values

The proposed binary classification model, expect for the prediction of the Class (C_0, C_1), provides the probability of each patient to be targeted as pathological (C_1), by assuming that instances with probabilities over 0.5 are considered in C_1. The main idea of the tuning of the cut off threshold is to find a threshold according to an optimization procedure of some performance metric. More specifically, in this proposed model and based on the training dataset, we tune the cut off threshold based on i) ROC curve and on the ii) balanced accuracy. The idea of using the ROC curve for tuning the threshold is to identify that threshold that gives us the upper-left corner of the curve [269]. Mathematically speaking, that threshold p that satisfies the following equations:

$$TPR(p) = 1 - FPR(p) \tag{5.1}$$

$$p^* = \arg min_p |TPR(p) + FPR(p) - 1| \tag{5.2}$$

where TPR is the true positive rate and FPR is the false positive rate.

5.3 Diagnostic prediction of carotid artery disease

5.3.1 Problem definition

The diagnostic prediction of carotid artery disease problem aims to identify the individuals of high risk for the presence of at least one of the carotid arteries with CAS > 50%. The development of this model is based on retrospective dataset, with different clinical views of input features. The detection of these individuals has been formulated as a multivariate 2 class classification problem based on the CAS, as it is defined by carotid US. More specifically, high risk group (Class 1) included the individuals whose at least one of the carotid arteries (common, internal or external, left or right) has a CAS above 50%, whereas low risk group (Class 0) included the individuals whose CAS were below 50%. In Figure 5.3, we demonstrate the concept of diagnostic prediction of CAS >50% clinical problem.

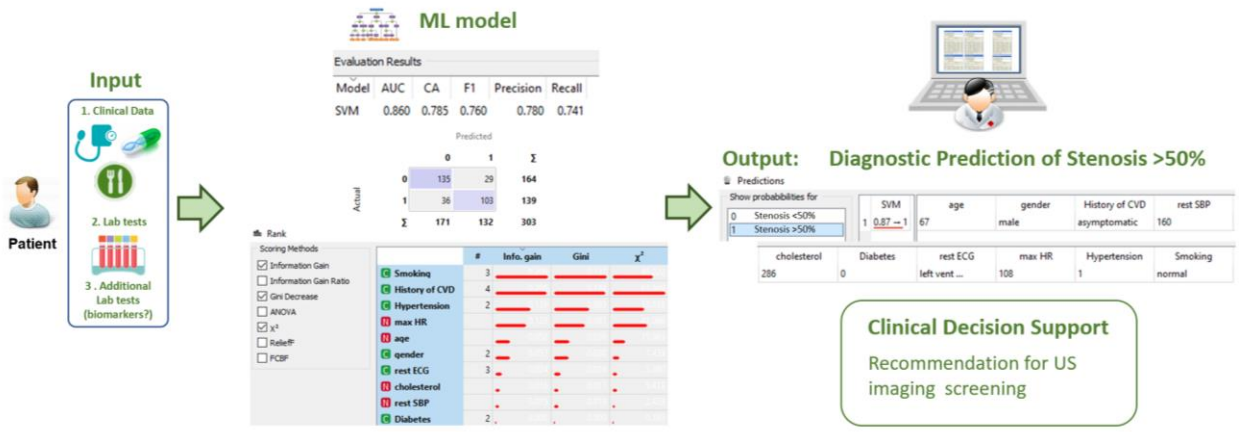

Figure 5.3 Diagnostic Prediction of CAS>50% clinical problem.

5.3.2 Dataset

The training dataset, utilized for the presented carotid artery disease risk stratification model, was collected from individuals admitted to the Clinic for Vascular and endovascular surgery, in the University Clinical Center of Serbia during Taxinomisis project (European Union's Horizon 2020 research and innovation program under grant agreement No 755320). The data were anonymized and

were obtained under a data protection agreement, fulling all the ethical and legal requirements for data sharing posed by the General Data Protection Regulation. The patients were included no matter the indications and the only criterion for the inclusion was the age over 50 years old. The time period collected the dataset was from 01.03.2021 to 01.01.2022. More specifically, the non imaging based input features views were Demographics, Clinical Data, Risk Factors, and Medication Therapy. For the training process, 881 individuals were utilized and 438 of them were at high risk group (Class 1).

Regarding the validation cohort a total 140 participants were utilized for the external validation of the proposed model. These participants were collected by Clinic for Vascular and endovascular surgery, in the University Clinical Center of Serbia during Taxinomisis project. Table 5.5 demonstrates the baseline characteristics (mean-SD for continuous features, presence percentage for categorical features) for the different features' types the training cohort.

Table 5.5 Baseline Characteristics of training and external validation cohort.

Baseline Input features	**Overall Cohort (N=881)**
Age (years), mean (SD)	66.99 (8.11)
Gender (male), n (%)	642 (73.32%)
Height (cm), mean (SD)	173.33 (9)
Weight (kg), mean (SD)	81.16 (15.75)
Body Mass Index (BMI) (kg/m^2), mean (SD)	26.89 (4.09)
Smoking, n (%)	429 (48.69%)
Alcohol Consumption, n (%)	175 (19.86%)
DM, n (%)	249 (28.26%)
Hypertension, n (%)	741 (84.11)
Hyperlipoproteinemia, n (%)	424 (48.13)
CAD, n (%)	250 (28.39%)
Aneurysm Disease, n (%)	288 (32.69%)
Antiaggregant Therapy, n (%)	707 (80.25%)
Antihypertensive Therapy, n (%)	707 (80.25%)
Anticoagulant Therapy, n (%)	68 (7.72%)
Statin Therapy, n (%)	505 (57.32%)
At least one CAS>50%, n (%)	438 (49.72%)

5.3.3 Results

Statistical analysis results

Table 5.6 provides the basic descriptive statistics of the distribution of numerical and categorical features of the utilized training dataset in each class. Statistical analysis was performed to the whole dataset and continuous variables were analyzed with independent-samples t-test, whereas categorical data were compared with Chi-square test. Compared with the healthy group, CAS group was in older age (CAS group vs Non-CAS group: 68.18±7.52 vs 65.81±8.49 years old), had higher weight (CAS group vs Non-CAS group: 95.92±38.01 vs 83.98±16.55 kg) and lower height (CAS group vs Non-CAS group: 154.58±39.2 vs 174.26±10.02 cm). Additionally, CAS group had a higher proportion of females and individuals with hyperlipoproteinemia, DM, hypertension, smoking, CAD, antiaggregant, antihypertensive and statin therapy. Amongst the risk factors, hyperlipoproteinemia, DM, hypertension, smoking, alcohol consumption, CAD and aneurysm disease show statistically significant different proportions between Class 0 and Class 1. Regarding the medication therapy, the use of antiaggregant, antihypertensive and statin therapy were identified as statistically different between Class 0 and Class 1

Table 5. 6 Subjects characteristics in CAS group and non-CAS group.

Variables	Non CAS group (N=443)	CAS group (N=438)	P value
Age	65.81±8.49	68.18±7.52	0.261
Height	174.47±8.72	172.11±9.15	0.132
Weight	83.76±16.11	78.4±14.88	0.261
BMI	27.4±4.16	26.36±3.96	0.609
Gender (male)	110 (24.83%)	125 (28.53%)	0.214
HLP	169 (38.14%)	255 (58.21%)	**<0.001**
DM	93 (20.99%)	156 (35.62%)	**<0.001**
HTA	342 (77.2%)	399 (91.1%)	**<0.001**
Smoking	180 (40.63%)	249 (56.85%)	**<0.001**
Alcohol consumption	107 (24.15%)	68 (15.52%)	**0.001**
Coronary disease	93 (20.99%)	157 (35.84%)	**<0.001**
Aneurysm Disease	231 (52.14%)	57 (13.01%)	**<0.001**
Antiaggregant	317 (71.56%)	390 (89.04%)	**<0.001**
Antihypertensive	304 (68.62%)	362 (82.65%)	**<0.001**
Anticoagulant	37 (8.35%)	31 (7.08%)	0.473
Statin	225 (50.79%)	280 (63.93%)	**<0.001**

Model performance evaluation results

The proposed machine learning model was evaluated using as performance metrics the balanced accuracy, negative predictive value (NPV), positive predictive value (PPV), AUC, the sensitivity and the specificity. In Table 5.7, we demonstrate the values for each utilized metric per fold, their mean values and SD for training dataset, where the mean balanced accuracy, sensitivity and specificity of the proposed model is 0.78, 0.82 and 0.74, respectively, while in Figure 5.4 and Figure 5.5, we demonstrate the normalized confusion matrix and the ROC curve for the proposed model, respectively. Except of the 10-fold cross validation for the evaluation of the proposed models, we utilized also an external validation dataset, including 521 individuals. Regarding the external dataset cohort, we show in Table 5.8 the obtained results in three different cases (Case 1-Cut-off: 0.5, Case 2-Cut-off: 0.4 and Case 3-Cut-off: 0.56). The highest sensitivity (0.93) and specificity (0.88) is achieved in case 2 and 3, respectively.

Table 5.7 Evaluation of the carotid artery disease risk prediction problem over the 10-folds.

Folds	Balanced Accuracy	NPV	PPV	ROC AUC	Sensitivity	Specificity
Fold #0	0.730556	0.744186	0.717391	0.832828	0.75	0.711111
Fold #1	0.830233	0.857143	0.804348	0.903876	0.860465	0.8
Fold #2	0.875711	0.904762	0.847826	0.946253	0.906977	0.844444
Fold #3	0.693182	0.666667	0.72973	0.786674	0.613636	0.772727
Fold #4	0.727273	0.857143	0.666667	0.860537	0.909091	0.545455
Fold #5	0.840909	0.894737	0.8	0.872934	0.909091	0.772727
Fold #6	0.772727	0.772727	0.772727	0.829545	0.772727	0.772727
Fold #7	0.840909	0.857143	0.826087	0.907541	0.863636	0.818182
Fold #8	0.693182	0.707317	0.680851	0.779959	0.727273	0.659091
Fold #9	0.75	0.805556	0.711538	0.782025	0.840909	0.659091
Mean	0.775468	0.806738	0.755717	0.850217	0.815381	0.735556
SD	0.066792	0.081343	0.062966	0.058085	0.097482	0.091537

Table 5.8 Validation of the carotid artery disease risk problem using an external validation dataset.

		Case 1 Cut-off: 0.5	Case 2 Cut-off: 0.4	Case 3 Cut-off: 0.56
N=521	True Positive (TP)	388	407	360
	True Negative (TN)	68	63	71
	False Positive (FP)	13	18	10
	False Negative (FN)	52	33	80
	Sensitivity	0.88	0.93	0.82
	Specificity	0.84	0.78	0.88
	PPV	0.97	0.96	0.97
	NPV	0.57	0.66	0.47
	Accuracy	0.88	0.9	0.83

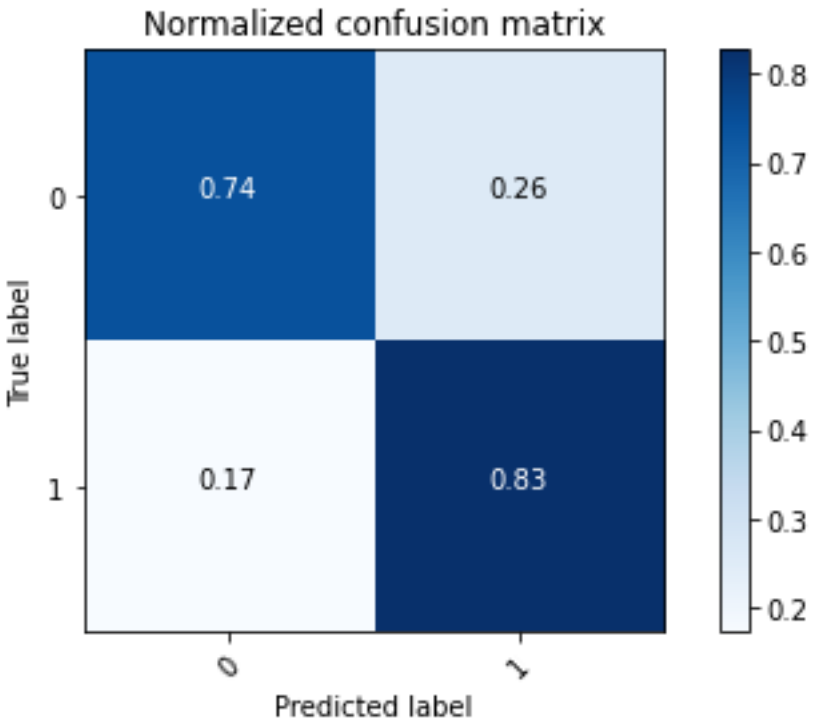

Figure 5.4 Normalized confusion matrix regarding the diagnosis of carotid artery disease presence.

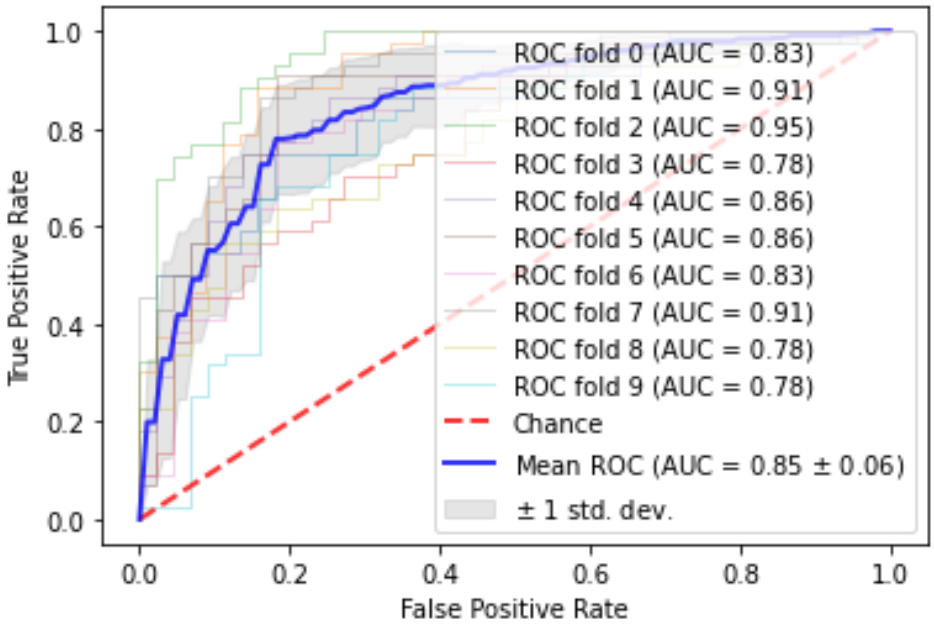

Figure 5.5 ROC curve analysis regarding the diagnosis of carotid artery disease presence.

Tuning of cut-off threshold values

Two different approaches have been followed for the definition of cut-off probability threshold values, using the ROC analysis and the balanced accuracy maximization approach. In Figure 5.6 and Figure 5.7, we show these two approaches, concluding that 0.57 and 0.56 are the best cut-off values for the definition of low risk class (Class 0) and high risk class (Class 1).

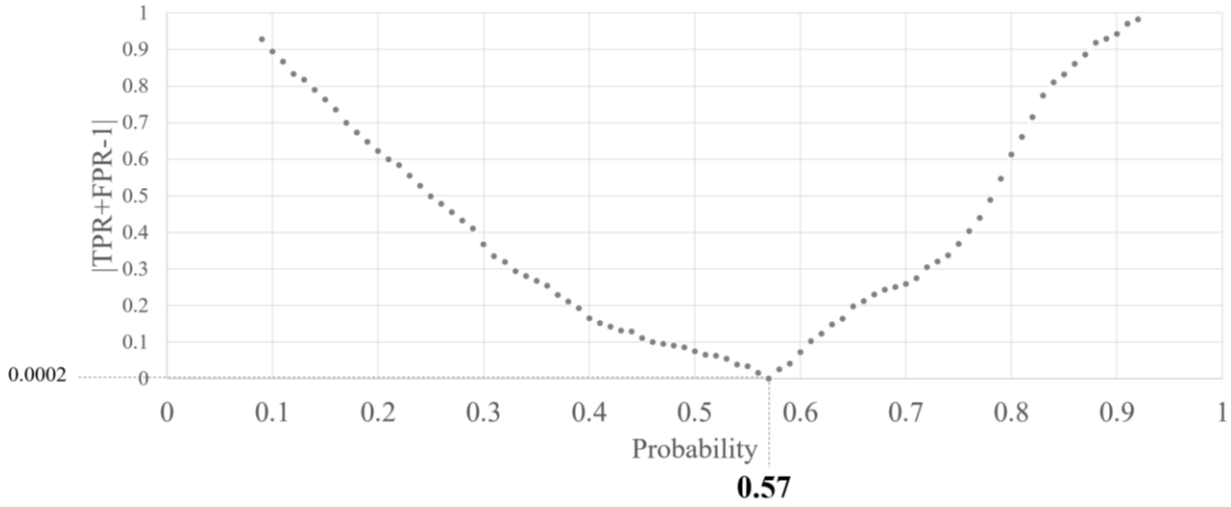

Figure 5.6 Tuning of threshold value based on ROC analysis regarding the diagnosis of carotid artery disease presence.

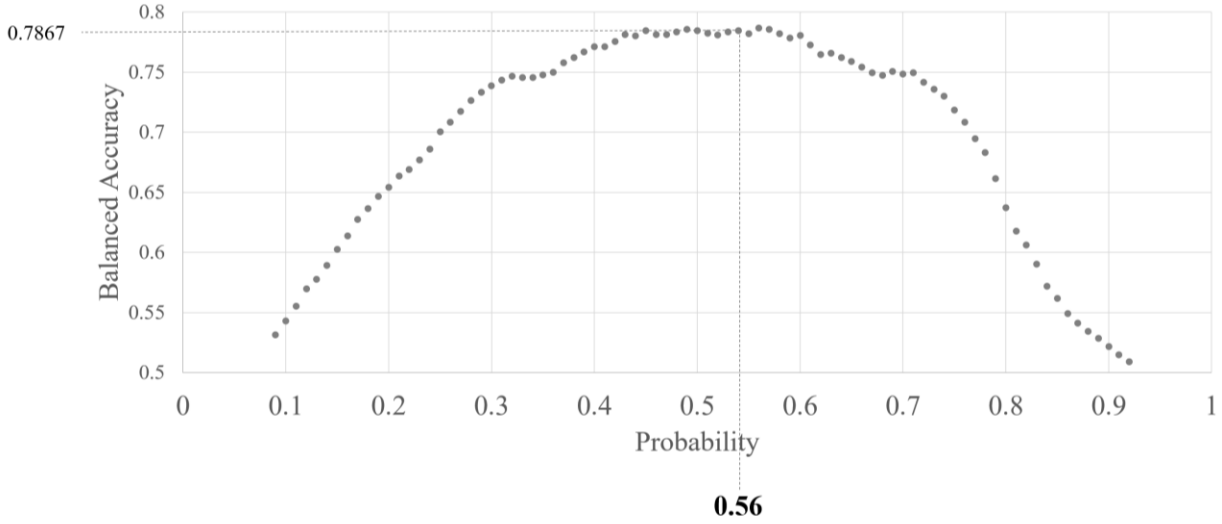

Figure 5.7 Tuning of threshold value based on balanced accuracy regarding the diagnosis of carotid artery disease presence.

5.4 Diagnostic prediction of the vulnerability of carotid artery plaques

5.4.1 Problem definition

The aim of the this proposed second approach is to identify individuals of high risk carotid plaque, as it is shown in Figure 5.8. Different datasets have been utilized for the development of this model, where high risk plaques have been defined based either on the histological composition of atherosclerotic plaques or on the presence of symptoms. The carotid artery disease risk stratification problem based on the vulnerability of plaques has been formulated as multivariate 2-class classification problem of the presence of high risk plaques. On this basis, seven subproblems were defined aiming at providing a multilevel characterization of one individual's status.

i. Subproblem 1:-USMI Dataset based on the composition of total collagen
    a. Class 0: Low Risk Plaques (cases 107)
    b. Class 1: High Risk Plaques (cases 101)
ii. Subproblem 2:-USMI Dataset based on the composition of SMCs
    a. Class 0: Low Risk Plaques (cases 157)
    b. Class 1: High Risk Plaques (cases 52)
iii. Subproblem 3:-USMI Dataset based on the composition of lipid
    a. Class 0: Low Risk Plaques (cases 111)
    b. Class 1: High Risk Plaques (cases 96)
iv. Subproblem 4:-USMI Dataset based on the composition of neutrophils
    a. Class 0: Low Risk Plaques (cases 113)
    b. Class 1: High Risk Plaques (cases 85)
v. Subproblem 5:-USMI Dataset based on the composition of macrophages
    a. Class 0: Low Risk Plaques (cases 113)
    b. Class 1: High Risk Plaques (cases 86)
vi. Subproblem 6- USMI whole dataset based on the presence of symptomatic disease
    a. Class 0: Low Risk Plaques (cases 241)
    b. Class 1: High Risk Plaques (cases 44)
vii. Subproblem 7- Atheroexpress dataset based on the presence of symptomatic disease
    a. Class 0: Low Risk Plaques (cases 64)
    b. Class 1: High Risk Plaques (cases 310)

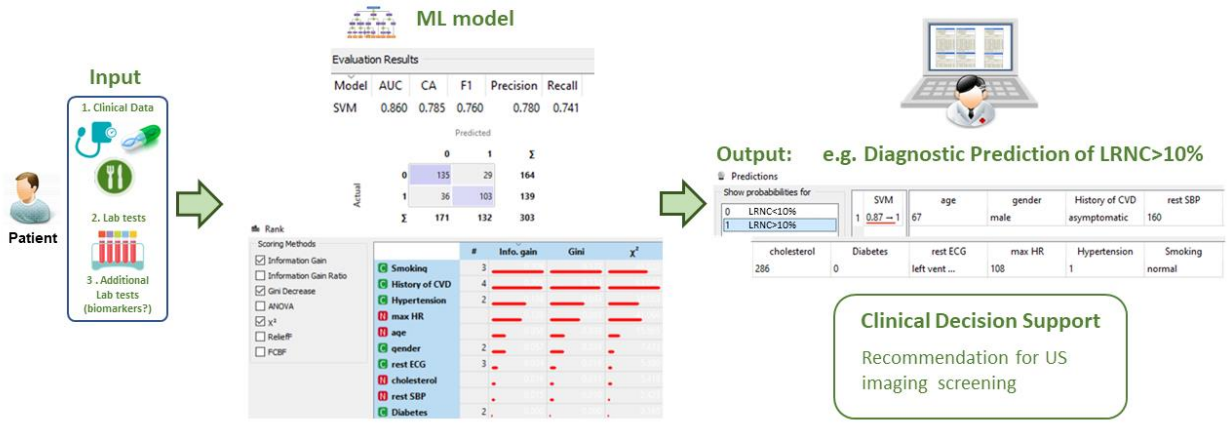

Figure 5.8 Diagnostic Prediction of vulnerable plaque components.

5.4.2 Dataset

Input features for the diagnostic prediction of vulnerable plaque components were retrospective dataset, collected during the Taxinomisis project and were provided by USMI and UMC clinical center. Subproblem 1-6 were formulated based on a dataset provided by USMI center, whereas Subproblem 7 was formulated based on a dataset provided by UMC [270].

More specifically, Subproblem 1-5 were defined as a multivariate 2 class classification problem based on 5 different plaque histology related features. Thus, 5 different subproblems (Subproblem 1, Subproblem 2, Subproblem 3, Subproblem 4 and Subproblem 5) have been defined and their input features and baseline characteristics for each subproblem are illustrated in Table 5.10. Clinical doctors from USMI have identified the most significant histological-plaque related features, which are directly associated with the vulnerability of the atherosclerotic plaques. For each one of these five plaque related features, a cut off value has been defined, in order to binarize the output of our proposed models. Clinical doctors have statistically analyzed USMI dataset to identify accurate cut-off values for the classification of non-vulnerable-low risk plaques (Class 0) and vulnerable-high risk plaques (Class 1). Subproblem 1-5 have been defined based on the concentration of total collagen, SMCs, lipid, neutrophils and macrophages, respectively, whereas the implemented cut-off values for the definition of the output of our proposed models are shown in Table 5.9, below.

Regarding the subproblems 6 and 7, the output has been defined based on the presence of symptoms. In Table 5.11 and Table 5.12, we show the input features and their basic statistics for Subproblem 6 and

Subproblem 7, respectively, including as input views demographics, clinical data, risk factors, heamatological, biochemical and serum markers.

Table 5.9 Roc curve analysis identifying symptomatic patients according to intraplaque parameters.

Overall content	**AUC (95% CI)**	***p*-value**	**Youden index**	**Cut-off value**	**Sensitivity (%)**	**Specificity (%)**	**+ LR**	**- LR**
Total collagen (%)	0.808 (0.743-0.863)	<0.001	0.55	**≦ 22.68**	86.48	68.06	2.71	0.20
SMCs (%)	0.748 (678-0.809)	<0.001	0.43	**≦ 2.82**	59.46	81.94	3.29	0.49
Lipid	0.705 (0.633-0.771)	<0.001	0.39	**>5.61**	78.38	60.42	1.98	0.36
Neutrophil (cell/mm^2)	0.757 (0.688-0.818)	<0.001	0.48	**>5.44**	83.78	64.58	2.37	0.25
Macrophages	0.628 (0.553-0.699)	0.006	0.27	**>4.46**	89.19	37.50	1.43	0.29

Table 5.10 Baseline Characteristics of training datasets regarding Subproblems 1-5.

	Subproblem 1 (<22.68 threshold of total collagen) (N=208)	**Subproblem 2 (<2.82 of SMCs) (N=209)**	**Subproblem 3 (>5.61 threshold of Lipid) (N=207)**	**Subproblem 4 (>5.44 threshold Neutrophils) (N=199)**	**Subproblem 5 (>4.46 of Macrophages) (N=202)**
Demographics					
Age (years), mean (SD)	72.61 (8.26)	72.51 (8.24)	72.52 (8.29)	72.53 (8.21)	72.59 (8.17)
Gender (male), n (%)	140 (67.31%)	141 (67.46%)	138 (66.67%)	135 (67.84%)	137 (67.82%)
Clinical Data					
SBP (mmHg), mean (SD)	137 (12.59)	136.82 (12.48)	136.80 (12.57)	137.32 (12.5)	137.11 (12.53)
DBP (mmHg), mean (SD)	81.87 (7.29)	81.91 (7.29)	81.88 (7.34)	81.75 (7.38)	81.79 (7.35)
Risk Factors					
Smoking, n (%)	53 (25.48%)	52 (24.88%)	52 (25.12%)	49 (24.62%)	49 (24.26%)
Hypertension, n (%)	150 (72.12%)	150 (71.77%)	148 (71.5%)	144 (72.36%)	147 (72.77%)
DM, n (%)	46 (22.12%)	47 (22.49%)	46 (22.22%)	44 (22.11%)	45 (22.23%)
Chronic CAD, n (%)	41 (19.71%)	42 (20.1%)	41 (19.81%)	39 (19.6%)	40 (19.8%)
Medication Therapy					
Statin, n (%)	112 (53.85%)	113 (54.07%)	112 (54.11%)	109 (54.77%)	110 (54.46%)
Clopidogrel, n (%)	48 (23.08%)	48 (22.97%)	48 (23.19%)	46 (23.12%)	47 (23.27%)
Ca blockers, n (%)	66 (31.73%)	65 (31.1%)	65 (31.4%)	64 (32.16%)	65 (32.18%)
Anticoagulants, n (%)	13 (6.25%)	13 (6.22%)	13 (6.28%)	12 (6.03%)	13 (6.43%)
Diuretics, n (%)	30 (14.42%)	29 (13.88%)	29 (14.01%)	27 (13.57%)	27 (13.37%)
β-blockers, n (%)	58 (27.88%)	57 (27.27%)	56 (27.05%)	54 (27.14%)	55 (27.23%)
ACEi, n (%)	9 (4.33%)	9 (4.31%)	9 (4.35%)	9 (4.52%)	6 (2.97%)
ARBs, n (%)	98 (47.12%)	100 (47.85%)	98 (47.34%)	94 (47.24%)	96 (47.52%)
Lipid Biomarkers					
TC, mean (SD)	192.38 (44.3)4	191.97 (44.39)	192.1 (44.56)	191.71 (44.86)	191.64 (44.55)
HDL, mean (SD)	50.96 (14.97)	50.99 (15.21)	51.02 (15.21)	51.15 (15.32)	51.25 (15.22)
LDL, mean (SD)	114.73 (39.75)	114.11 (39.94)	114.28 (40.15)	114.19 (40.48)	114.07 (40.19)
Triglycerides, mean (SD)	141 (74.42)	141.91 (74.51)	141.55 (74.62)	139.75 (72.42)	139.36 (72.25)
High Risk Plaques	101 (48.56%)	52 (24.88%)	96 (46.38%)	86 (43.22%)	135 (66.83%)

Table 5.11 Baseline Characteristics of training datasets regarding Subproblem 6.

Overall Cohort (N=285)			
Clinical Data		**Serum Markers**	
Age, mean (SD)	72.73 (8.2)	CCL2, mean (SD)	90.94 (123.74)
Gender (male), n (%)	189 (66.32%)	OPG, mean (SD)	883.62 (859.56)
SBP, mean (SD)	134.91 (12.67)	hs-CRP, mean (SD)	3.82 (5.26)
DBP, mean (SD)	81.24 (7.21)	P-selectin, mean (SD)	97.67 (88.18)
waist circumference, mean (SD)	91.4 (7.925)	ICAM-1, mean (SD)	221.94 (100.56)
Risk Factors		VCAM-1, mean (SD)	420.62 (265.39)
Hypertension, n (%)	207 (72.63%)	Adiponectin, mean (SD)	4.63 (3.35)
DM, n (%)	67 (23.51%)	E-selectin, mean (SD)	21.28 (13.64)
Dyslipidemia, n (%)	151 (52.98%)	Resistin, mean (SD)	11.47 (42.97)
Chronic CAD, n (%)	64 (22.46%)	Leptin, mean (SD)	25.76 (40.52)
Smoking, n (%)	69 (24.21%)	IL-α, mean (SD)	11.09 (22.99)
Medication		TNF-α, mean (SD)	43.59 (136.08)
Statins, n (%)	158 (55.44%)	CCL5, mean (SD)	70.85 (58.36)
Biochemical		CCL4, mean (SD)	41.08 (96.71)
D-dimer, mean (SD)	571.97 (1331.38)	CCL3, mean (SD)	69.23 (309.38)
Fibrinogenemia, mean (SD)	3.94 (1.42)	CD40L, mean (SD)	602.75 (1277.29)
Fasting Insulinemia, mean (SD)	10.98 (8.88)	IL-6, mean (SD)	2.81 (3.38)
Fasting C-peptidemia, mean (SD)	3.09 (1.43)	sIL-6R, mean (SD)	36.62 (17.73)
Lpa, mean (SD)	0.12 (0.19)	IGF-1, mean (SD)	1.39 (2.47)
TC, mean (SD)	195.02 (45.71)	L-selectin, mean (SD)	3018.34 (1437.18)
HDL, mean (SD)	51.8 (15.77)	MMP-9, mean (SD)	579.38 (529.52)
LDL, mean (SD)	116.69 (41.14)	proMMP-9, mean (SD)	22.16 (9.45)
Triglyceridemia, mean (SD)	138.41 (68.9)	MMP-8, mean (SD)	15.77 (20.11)
Fasting Glycemia, mean (SD)	113.07 (33.78)	TIMP-1, mean (SD)	410.02 (431.72)
Heamatological		TIMP-2, mean (SD)	115.42 (57.28)
WBC, mean (SD)	7.18 (1.65)	TIMP-3, mean (SD)	14.76 (17.18)
Neutrophils, mean (SD)	64.12 (7.68)	TIMP-4, mean (SD)	6.01 (15.24)
Monocytes, mean (SD)	6.47 (1.87)	MMP-9/TIMP-1, mean (SD)	17.14 (14.51)
Lymphocytes, mean (SD)	25.2 (6.7)	FAP, mean (SD)	142.05 (89.99)
Platelets, mean (SD)	233.88 (65.05)	OPN, mean (SD)	69.58 (75.8)
RBC, mean (SD)	4.64 (0.5)	RANKL, mean (SD)	1915.6 (1924.73)
		Neutrophil elastase, mean (SD)	310.73 (377.69)
		Total Vitamin D, mean (SD)	42.75 (31.95)
		PCSK-9, mean (SD)	294.73 (128.17)

Table 5.12 Baseline Characteristics of training datasets regarding Subproblem 7.

Overall Cohort (N=374)			
Demographics		***Serum markers***	
Age, mean (SD)	68.51 (8.84)	FABP4, mean (SD)	12984.2 (11711.48)
Gender (male), n (%)	94 (25.13%)	Cystatin C, mean (SD)	337610.88 (131515.87)
Clinical		Lp-PLA(2), mean (SD)	113.9 (47.63)
SBP, mean (SD)	153.55 (25.76)	PCSK-9, mean (SD)	31896.72 (18998.12)
DBP, mean (SD)	81.70 (13.22)	GDF15, mean (SD)	3361.78 (7608.91)
MAP, mean (SD)	105.65 (15.65)	RANTES, mean (SD)	3.63 (4.29)
BMI, mean (SD)	26.59 (3.88)	hs-CRP, mean (SD)	31.96 (385.22)
Heamatological		TAT, mean (SD)	116.12 (770.8)
Hb, mean (SD)	8.69 (0.97)	Myeloperoxidase, mean (SD)	44.03 (42.04)
Haematocrit, mean (SD)	0.41 (0.05)	Nt-pro-b, mean (SD)	234.19 (2365.15)
Risk factors		PDGF, mean (SD)	0.25 (0.65)
CAD, n (%)	122 (32.62%)	OPG, mean (SD)	1.75 (0.72)
PAOD, n (%)	80 (21.39%)	VEGF-A, mean (SD)	8.53 (6.68)
peripheral intervention, n (%)	61 (16.31%)	VWF, mean (SD)	57.6 (50.44)
DM, n (%)	66 (17.65%)	***Biochemical***	
treatment of DM, n (%)	64 (17.11%)	Creatinine, mean (SD)	95.46 (29.8)
smoking current, n (%)	138 (36.89 %)	GFR, mean (SD)	76.2 (23.14)
Hypertension, n (%)	255 (68.18%)	Glucose, mean (SD)	6.64 (2.09)
		TC, mean (SD)	176.26 (46.24)
Symptomatic Disease	310 (82.89%)	LDL, mean (SD)	105.95 (39.91)
		HDL, mean (SD)	43.4 (14.31)
		Triglycerides, mean (SD)	138.1 (74.68)

5.4.3 Results

Statistical Analysis Results

In Table 5.13-Table 5.17, we show the statistical analysis results for class of low risk plaques, in comparison with class of high risk plaques for Subproblems 1-5. Individuals, who receive anticoagulants therapy (Subproblem 1-p:0.035, Subproblem 1-p:0.013, Subproblem 4-p:0.014), Clopidogrel (Subproblem 1- p:0.002, Subproblem 5- p:0.007), ACEi (Subproblem 3- p:0.009) and Ca blockers (Subproblem 3- p:0.036) are more susceptible of the presence of high risk plaques.

Table 5.13 Subjects characteristics in high risk plaques group and low risk plaques group regarding Subproblem 1.

	Low risk plaques (N=107)	**High risk plaques (N=101)**	**P value**
Age (years), mean (SD)	71.21 (8.02)	74.09 (8.3)	0.899
Gender (male), n (%)	66 (61.11)	74 (73.27)	0.076
SBP (mmHg), mean (SD)	139.48 (15.54)	134.4 (10.98)	0.305
DBP (mmHg), mean (SD)	81.98 (6.54)	81.76 (8.04)	0.097
Smoking, n (%)	29 (27.1%)	24 (23.76%)	0.837
Hypertension, n (%)	80 (74.76%)	70 (69%)	0.381
DM, n (%)	25 (23.36%)	21 (21%)	0.656
Chronic CAD, n (%)	20 (18.69%)	21 (21%)	0.704
Statin, n (%)	58 (54.21%)	54 (53%)	0.915
Clopidogrel, n (%)	20 (18.69%)	28 (28%)	0.132
Ca blockers, n (%)	36 (33.64%)	30 (30%)	0.543
Anticoagulants, n (%)	3 (2.8%)	10 (10%)	**0.035**
Diuretics, n (%)	14 (13.08%)	16 (16%)	0.572
β-blockers, n (%)	33 (30.84%)	25 (25%)	0.329
ACEi, n (%)	5 (4.67%)	4 (4%)	0.801
ARBs, n (%)	54 (50.47%)	44 (44%)	0.32
TC, mean (SD)	194.81 (50.4)	189.75 (36.75)	0.084
HDL, mean (SD)	51.74 (14.95)	50.12 (15.02)	0.474
LDL, mean (SD)	115.06 (45.18)	114.39 (33.13)	0.076
Triglycerides, mean (SD)	144.55 (85.1)	137.17 (61.01)	0.239

Table 5.14 Subjects characteristics in high risk plaques group and low risk plaques group regarding Subproblem 2.

	Low risk plaques (N=157)	**High risk plaques (N=52)**	**P value**
Age (years), mean (SD)	72.36 (7.86)	72.96 (9.37)	0.145
Gender (male), n (%)	106 (67.52)	35 (67.31)	0.978
SBP (mmHg), mean (SD)	136.3 (12.48)	138.43 (12.47)	0.844
DBP (mmHg), mean (SD)	81.83 (7.51)	82.18 (6.65)	0.23
Smoking, n (%)	40 (25.48%)	12 (23.08%)	0.277
Hypertension, n (%)	115 (73.25%)	35 (67.31%)	0.411
DM, n (%)	37 (23.57%)	10 (19.23%)	0.517
Chronic CAD, n (%)	31 (19.75%)	11 (21.15%)	0.827
Statin, n (%)	87 (55.41%)	26 (50%)	0.498
Clopidogrel, n (%)	28 (17.83%)	20 (38.46%)	**0.002**
Ca blockers, n (%)	47 (29.94%)	18 (34.62%)	0.529

	Low risk plaques (N=157)	**High risk plaques (N=52)**	**P value**
Anticoagulants, n (%),	6 (3.82%)	7 (13.46%)	**0.013**
Diuretics, n (%)	23 (14.65%)	6 (11.54%)	0.575
β-blockers, n (%)	37 (23.57%)	20 (38.46%)	0.037
ACEi, n (%)	7 (4.46%)	2 (3.85%)	0.851
ARBs, n (%)	78 (49.68%)	22 (42.31%)	0.357
TC, mean (SD)	192.89 (45.63)	189.14 (40.65)	0.99
HDL, mean (SD)	52.29 (14.93)	47 (15.49)	0.684
LDL, mean (SD)	113.78 (40.68)	115.1 (37.94)	0.766
Triglycerides, mean (SD)	139.16 (68.75)	150.24 (89.96)	0.196

Table 5.15 Subjects characteristics in high risk plaques group and low risk plaques group regarding Subproblem 3.

	Low risk plaques (N=111)	**High risk plaques (N=96)**	**P value**
Age (years), mean (SD)	72.47 (7.63)	72.57 (9.03)	0.143
Gender (male), n (%)	70 (63.06%)	68 (70.83%)	0.238
SBP (mmHg), mean (SD)	133.73 (11.1)	140.39 (13.28)	0.661
DBP (mmHg), mean (SD)	81.55 (7.74%)	82.26 (6.86)	0.219
Smoking, n (%)	33 (29.73%)	19 (19.79%)	0.831
Hypertension, n (%)	79 (71.17%)	69 (71.88%)	0.723
DM, n (%)	23 (20.72%)	23 (23.96%)	0.577
Chronic CAD, n (%)	21 (18.92%)	20 (20.83%)	0.731
Statin, n (%)	59 (53.15%)	53 (55.21%)	0.768
Clopidogrel, n (%)	25 (22.52%)	23 (23.96%)	0.808
Ca blockers, n (%)	29 (26.13%)	36 (37.5%)	0.079
Anticoagulants, n (%)	5 (4.5%)	8 (8.33%)	0.259
Diuretics, n (%)	14 (12.61%)	15 (15.63%)	0.534
β-blockers, n (%)	30 (27.03%)	26 (27.08%)	0.993
ACEi, n (%)	1 (0.9%)	8 (8.33%)	**0.009**
ARBs, n (%)	55 (49.55%)	43 (44.79%)	0.495
TC, mean (SD)	186.22 (42.96)	198.77 (45.62)	0.582
HDL, mean (SD)	51.65 (14.85)	50.3 (15.64)	0.741
LDL, mean (SD)	108.88 (38.52)	120.41 (41.27)	0.946
Triglycerides, mean (SD)	134.14 (65.02)	149.97 (83.78)	0.339

Table 5.16 Subjects characteristics in high risk plaques group and low risk plaques group regarding Subproblem 4.

	Low risk plaques (N=113)	**High risk plaques (N=86)**	**P value**
Age (years), mean (SD)	71.99 (7.68)	73.24 (8.85)	0.412
Gender (male), n (%)	75 (66.37%)	60 (69.77%)	0.612
SBP (mmHg), mean (SD)	135.53 (12.06)	139.65 (12.74)	0574
DBP (mmHg), mean (SD)	81.54 (7.24)	82.03 (7.6)	0.59
Smoking, n (%)	27 (23.89%)	22 (25.58%)	0.641
Hypertension, n (%)	80 (70.8%)	64 (74.42%)	0.572
DM, n (%)	27 (23.89%)	17 (19.77%)	0.488
Chronic CAD, n (%)	20 (17.7%)	19 (22.09%)	0.44
Statin, n (%)	69 (61.06%)	40 (46.51%)	0.042
Clopidogrel, n (%)	23 (20.35%)	23 (26.74%)	0.291
Ca blockers, n (%)	30 (26.55%)	34 (39.53%)	0.053
Anticoagulants, n (%)	3 (2.65%)	9 (10.47%)	**0.022**
Diuretics, n (%)	14 (12.39%)	13 (15.12%)	0.579
β-blockers, n (%)	33 (29.2%)	21 (24.42%)	0.453
ACEi, n (%)	5 (4.42%)	4 (4.65%)	0.939
ARBs, n (%)	54 (47.79%)	40 (46.51%)	0.859
TC, mean (SD)	188.15 (42.09)	196.3 (48.06)	0.798
HDL, mean (SD)	51.2 (14.21)	51.09 (16.72)	0.488
LDL, mean (SD)	110.71 (37.82)	118.69 (43.48)	0.576
Triglycerides, mean (SD)	134.72 (58.48)	146.27 (87.17)	0.066

Table 5.17 Subjects characteristics in high risk plaques group and low risk plaques group regarding Subproblem 5.

	Low risk plaques (N=67)	**High risk plaques (N=135)**	**P value**
Age (years), mean (SD)	71.78 (7.3)	73 (8.57)	0.246
Gender (male), n (%)	44 (65.679%)	93 (68.89%)	0.646
SBP (mmHg), mean (SD)	136.13 (12.19)	137.6 (12.72)	0.759
DBP (mmHg), mean (SD)	81.78 (7.2)	81.8 (7.45)	0.425
Smoking, n (%)	14 (20.9%)	35 (25.93%)	0.866
Hypertension, n (%)	48 (71.64%)	99 (73.33%)	0.8
DM, n (%)	16 (23.88%)	29 (21.48%)	0.7
Chronic CAD, n (%)	11 (16.42%)	29 (21.48%)	0.396
Statin, n (%)	36 (53.73%)	74 (54.81%)	0.885
Clopidogrel, n (%)	8 (11.94%)	39 (28.89%)	**0.007**
Ca blockers, n (%)	15 (22.39%)	50 (37.04%)	**0.036**

	Low risk plaques (N=67)	**High risk plaques (N=135)**	**P value**
Anticoagulants, n (%)	1 (1.49%)	12 (8.89%)	0.044
Diuretics, n (%)	11 (16.42%)	16 (11.85%)	0.37
β-blockers, n (%)	18 (26.87%)	37 (27.41%)	0.935
ACEi, n (%)	2 (2.99%)	4 (2.96%)	0.993
ARBs, n (%)	35 (52.24%)	61 (45.19%)	0.346
TC, mean (SD)	186.94 (42.7)	193.95 (45.42)	0.908
HDL, mean (SD)	49.92 (14.57)	51.90 (15.55)	0.283
LDL, mean (SD)	108.38 (37.07)	116.88 (41.48)	0.794
Triglycerides, mean (SD)	142 (83)	138.07 (66.67)	0.603

Table 5.18 and Table 5.19 provide the basic descriptive statistics of the distribution of numerical and categorical features with respect to Subproblem 6 and Subproblem 7, respectively. Regarding Subproblem 6, the presence of carotid artery disease symptoms is significantly correlated with higher mean value of monocytes % (p=0.046), hs-CRP (p=0.026), leptin p=0.013), TNF-α (p=0.021), CCL5 (p=0.023), MMP-8 (p=0.032), MMP-9/ TIMP-1 (p=0.001) and OPN (p=0.006). On the other hand, carotid artery disease symptoms are significantly correlated with lower mean values of Neutrophils (%) (p=0.045), RBCs (p=0.038), Lpa (t-test p=0.004), TIMP2 (p=0.014), Total Vitamin D (p=0.017) and PCSK9 (p=0.034). As for Subproblem 7, carotid artery disease symptoms are significantly associated with higher mean values of Serum FABP4 (p=0.006), of RANTES (p=0.032) and age (p=0.003), whereas GDF15 exhibits lower mean values (p=0.04). In addition to this, individuals with history of CAD are more susceptible to present carotid artery symptoms.

Table 5.18 Subjects characteristics in high risk plaques group and low risk plaques group regarding Subproblem 6.

Subproblem 6 Variables	Low risk plaques (N=241)	High risk plaques (N=44)	P value
Age, mean (SD)	72.39 (8.14	74.59 (8.48	0.885
Gender	65.1	72.7	0.328
waist circumference, mean (SD)	91.57 (8.55	90.3 (8.7	0.736
Hypertension, (%)	73%	70.5%	0.725
DM, (%)	23.7%	22.7%	0.894
Dyslipidemia, (%)	51.5%	61.4%	0.226

Subproblem 6 Variables	Low risk plaques (N=241)	High risk plaques (N=44)	P value
Chronic CAD, (%)	23.7%	15.9%	0.258
Total WBC, mean (SD)	7.15 (1.71)	7.33 (1.35)	0.101
Neutrophils, mean (SD)	64.41 (7.37)	62.5 (9.29)	**0.045**
Monocytes, mean (SD)	6.39 (1.91)	6.8 (1.64)	**0.046**
Lymphocytes, mean (SD)	25.01 (6.4)	26.23 (8.28)	0.059
Platelets, mean (SD)	232.26 (66.73)	242.66 (56.6)3	0.58
RBC, mean (SD)	4.67 (0.49)	4.47 (0.55)	**0.038**
D-dimer, mean (SD)	610.6 (4500.1)	353.56 (470.29)	0.168
Fibrinogenemia, mean (SD)	3.93 (1.5)	4.01 (1.04)	0.877
Fasting Insulinemia, mean (SD)	11.1 (9.97)	10.31 (5.14)	0.23
Fasting C peptidemia, mean (SD)	3.08 (1.53)	3.17 (1.41)	0.175
Lpa, mean (SD)	0.13 (0.24)	0.06 (0.09)	**0.004**
TC, mean (SD)	195.57 (47.63)	191.98 (37.03)	0.292
HDL, mean (SD)	52.52 (15.68)	47.79 (16.78)	0.793
LDL, mean (SD)	116.32 (43.16)	118.77 (31.19)	0.177
Triglyceridemia, mean (SD)	137.48 (72.15)	143.49 (5.36)	0.283
Fasting Glycemia, mean (SD)	112.81 (34.45)	114.52 (31.51)	0.867
CCL2, mean (SD)	94.62 (133.68)	71 (65.85)	0.099
OPG, mean (SD)	853.94 (925.82)	1037.64 (743.15)	**0.033**
hs-CRP, mean (SD)	3.69 (5.76)	4.43 (4.07)	**0.026**
P-selectin, mean (SD)	98.05 (90.12)	95.73 (104.28)	0.48
ICAM-1, mean (SD)	221.3 (97.45)	225.41 (127.03)	0.238
VCAM-1, mean (SD)	427.31 (267.29)	384.58 (285.92)	0.701
Adiponectin, mean (SD)	4.7 (3.61)	4.28 (3.09)	0.412
E-selectin, mean (SD)	21.1 (14.96)	22.19 (11.07)	0.674
Resistin, mean (SD)	12.23 (49.12)	7.58 (7.66)	0.41
Leptin, mean (SD)	24.08 (29.56)	34.37 (81.98)	**0.013**
IL-α, mean (SD)	12.86 (28.57)	16.21 (36.09)	0.368
TNF-α, mean (SD)	36.14 (135.44)	82.29 (171.3)	**0.021**
CCL5, mean (SD)	68.32 (60.04)	84.21 (58.43)	**0.023**
CCL4, mean (SD)	41.58 (109.54)	38.47 (35.53)	0.412
CCL3, mean (SD)	70.22 (340.99)	64.08 (218.09)	0.882
CD40L, mean (SD)	589.77 (1328.57)	670.12 (1397.59)	0.625
IL-6, mean (SD)	2.82 (3.51)	2.76 (4.2)	0.802
sIL-6R, mean (SD)	37.14 (19.18)	33.93 (15.45)	0.323
IGF-1, mean (SD)	1.5 (3.57)	0.92 (0.82)	0.284
L-selectin, mean (SD)	3068.96 (1562.11)	2757.99 (1195.13)	0.342
MMP-9, mean (SD)	569.65 (525.82)	631.65 (615.29)	0.125
pro-MMP-9, mean (SD)	21.62 (11.96)	24.47 (14.78)	0.108
MMP-8, mean (SD)	14.31 (19.8)	23.67 (22.46)	**0.032**
TIMP-1, mean (SD)	415.52 (457.3)	380.58 (345.01)	0.478

Subproblem 6 Variables	Low risk plaques (N=241)	High risk plaques (N=44)	P value
TIMP-2, mean (SD)	118.99 (61.89)	96.88 (45.13)	**0.014**
TIMP-3, mean (SD)	14.67 (17.97)	15.2 (18.31)	0.614
TIMP-4, mean (SD)	6.32 (17.69)	4.41 (1.78)	0.363
MMP-9/TIMP-1, mean (SD)	16.55 (15.77)	19.99 (24.23)	**0.001**
FAP, mean (SD)	142.93 (97.75)	137.52 (77.3)	0.923
OPN, mean (SD)	64.4 (76.14)	97.28 (79.4)	**0.006**
RANKL, mean (SD)	1911.25 (2253.08)	1935.97 (2037.27)	0.677
MPO, mean (SD)	318.16 (542.5)	398.08 (589.36)	0.098
Neutrophil elastase, mean (SD)	317.59 (494.7)	279.52 (230.29)	0.524
Total Vitamin D, mean (SD)	44.04 (35.87)	36.17 (22.41)	**0.017**
PCSK-9, mean (SD)	301.07 (130.91)	260.95 (129.08)	**0.034**

Table 5.19 Subjects characteristics in high risk plaques group and low risk plaques group regarding Subproblem 7.

Subproblem 7 Variables	Low risk plaques (N=64)	High risk plaques (N=310)	P value
Age, mean (SD)	65.52 (9.35)	69.12 (8.62)	**0.003**
Gender (male), n (%)	62.1	74.7	0.431
SBP, mean (SD)	157.2 (23.77)	152.8 (26.13)	0.249
DBP, mean (SD)	83.78 (13.17)	81.27 (13.21)	0.2
MAP, mean (SD)	108.25 (14.46)	105.11 (15.85)	0.176
BMI, mean (SD)	27.11 (3.56)	26.49 (3.94)	0.254
Hb, mean (SD)	8.83 (0.85)	8.66 (0.99)	0.23
Haematocrit, mean (SD)	0.42 (0.04)	0.41 (0.05)	0.162
CAD, n (%)	37.5%	31.6%	0.36
PAOD, n (%)	25.0%	20.6%	0.439
peripheral intervention, n (%)	20.3%	15.5%	0.341
DM, n (%)	14.1%	18.4%	0.448
treatment of DM, n (%)	14.1%	17.7%	0.525
smoking current, n (%)	31.2%	38.1%	0.287
Hypertension, n (%)	73.4%	67.1%	0.349
FABP4, mean (SD)	10665.36 (4736.71)	13482.39 (12672.09)	**0.006**
Cystatin C, mean (SD)	331507.69 (112363.14)	338922.11 (135438.71)	0.705
Lp-PLA(2), mean (SD)	110.52 (42.49)	114.65 (48.72)	0.541
PCSK-9, mean (SD)	34973.27 (23839.7)	31204.49 (17706.34)	0.155
GDF15, mean (SD)	3809.32 (7751.99)	3261.44 (7586.86)	**0.04**
RANTES, mean (SD)	3.11 (3.99)	3.74 (4.35)	**0.032**
hs-CRP, mean (SD)	3.79 (5.73)	37.97 (424.2)	0.52
TAT, mean (SD)	94.047 (186.91)	120.81 (844.64)	0.801

Subproblem 7 Variables	Low risk plaques (N=64)	High risk plaques (N=310)	P value
Myeloperoxidase, mean (SD)	37.59 (35.82)	45.39 (43.17)	0.178
Nt-pro-b, mean (SD)	77.11 (74.9)	267.48 (2603.03)	0.559
PDGF, mean (SD)	0.22 (0.33)	0.26 (0.7)	0.708
OPG, mean (SD)	1.67 (0.66)	1.77 (0.74)	0.315
VEGF-A, mean (SD)	7.18 (4.14)	8.84 (7.11)	0.09
VWF, mean (SD)	48.05 (26.77)	59.62 (53.85)	0.14
Creatinine, mean (SD)	95.89 (25.86)	95.37 (30.58)	0.902
GFR, mean (SD)	78.67 (26.88)	75.4 (26.14)	0.38
Glucose, mean (SD)	6.7 (2.35)	6.62 (2.04)	0.832
TC, mean (SD)	181.27 (48.34)	175.18 (45.81)	0.381
LDL, mean (SD)	109.07 (40.47)	105.29 (39.84)	0.532
HDL, mean (SD)	43.87 (13.88)	43.29 (14.42)	0.79
Triglycerides, mean (SD)	151.95 (72.92)	135.12 (74.86)	0.133

Model performance-evaluation results

Balanced accuracy, sensitivity, specificity, PPV, NPV and AUC were utilized for the evaluation of the proposed models. In Table 5.20-Table 5.26, we demonstrate the values for each utilized performance metric per fold, their mean values and SD, for Subproblems 1-7, respectively.

The achieved accuracies were 0.62±0.13, 0.64±0.11, 0.63±0.07, 0.58±0.12, 0.53±0.12, 0.65± 0.09, 0.79±0.12, whereas the AUC values were 0.64, 0.67, 0.65, 0.64, 0.54, 0.57, 0.77, for Subproblems 1-7, respectively. ROC curves analyses are shown in Figure 5.9- Figure 5.15.

Table 5.20 Evaluation of the carotid artery disease risk prediction problem based on the vulnerability of plaques over the 10-folds, regarding Subproblem 1.

	Subproblem 1					
	Accuracy	**Sensitivity**	**Specificity**	**PPV**	**NPV**	**AUC**
fold #0	0.768182	0.9	0.636364	0.692308	0.875	0.881818
fold #1	0.472727	0.4	0.545455	0.444444	0.5	0.445455
fold #2	0.809091	0.8	0.818182	0.8	0.818182	0.809091
fold #3	0.659091	0.5	0.818182	0.714286	0.642857	0.827273
fold #4	0.627273	0.8	0.454545	0.571429	0.714286	0.627273
fold #5	0.563636	0.4	0.727273	0.571429	0.571429	0.527273
fold #6	0.477273	0.5	0.454545	0.454545	0.5	0.563636
fold #7	0.763636	0.727273	0.8	0.8	0.727273	0.745455

	Subproblem 1					
	Accuracy	**Sensitivity**	**Specificity**	**PPV**	**NPV**	**AUC**
fold #8	0.5	0.6	0.4	0.5	0.5	0.42
fold #9	0.55	0.6	0.5	0.545455	0.555556	0.57
Mean	0.6190909	0.622727	0.615455	0.60939	0.640458	0.641727
SD	0.126690605	0.176709	0.165203	0.133609	0.137739	0.164343

Table 5. 21 Evaluation of the carotid artery disease risk prediction problem based on the vulnerability of plaques over the 10-folds, regarding Subproblem 2.

	Subproblem 2					
	Accuracy	**Sensitivity**	**Specificity**	**PPV**	**NPV**	**AUC**
fold #0	0.475	0.2	0.75	0.2	0.75	0.35
fold #1	0.7125	0.8	0.625	0.4	0.909091	0.725
fold #2	0.70625	0.6	0.8125	0.5	0.866667	0.875
fold #3	0.60625	0.4	0.8125	0.4	0.8125	0.675
fold #4	0.64375	0.6	0.6875	0.375	0.846154	0.5875
fold #5	0.7125	0.8	0.625	0.4	0.909091	0.8375
fold #6	0.44375	0.2	0.6875	0.166667	0.733333	0.475
fold #7	0.8	0.666667	0.933333	0.8	0.875	0.966667
fold #8	0.616667	0.5	0.733333	0.428571	0.785714	0.566667
fold #9	0.633333	0.6	0.666667	0.375	0.833333	0.68
Mean	0.635	0.536667	0.733333	0.404524	0.832088	0.673833
SD	0.10940883	0.214562	0.097143	0.172016	0.061573	0.188268

Table 5.22 Evaluation of the carotid artery disease risk prediction problem based on the vulnerability of plaques over the 10-folds, regarding Subproblem 3.

	Subproblem 3					
	Accuracy	**Sensitivity**	**Specificity**	**PPV**	**NPV**	**AUC**
fold #0	0.694444	0.888889	0.5	0.571429	0.857143	0.796296
fold #1	0.568182	0.5	0.636364	0.555556	0.583333	0.5
fold #2	0.668182	0.7	0.636364	0.636364	0.7	0.718182
fold #3	0.613636	0.5	0.727273	0.625	0.615385	0.554545
fold #4	0.718182	0.8	0.636364	0.666667	0.777778	0.763636
fold #5	0.527273	0.6	0.454545	0.5	0.555556	0.554545
fold #6	0.527273	0.6	0.454545	0.5	0.555556	0.690909
fold #7	0.661616	0.777778	0.545455	0.583333	0.75	0.606061
fold #8	0.59596	0.555556	0.636364	0.555556	0.636364	0.656566
fold #9	0.69697	0.666667	0.727273	0.666667	0.727273	0.646465
Mean	0.6271718	0.658889	0.595455	0.586057	0.675839	0.648721
SD	0.070824639	0.131994	0.101527	0.061407	0.102577	0.096224

Table 5.23 Evaluation of the carotid artery disease risk prediction problem based on the vulnerability of plaques over the 10-folds, regarding Subproblem 4.

	Subproblem 4					
	Accuracy	**Sensitivity**	**Specificity**	**PPV**	**NPV**	**AUC**
fold #0	0.474747	0.222222	0.727273	0.4	0.533333	0.676768
fold #1	0.560606	0.666667	0.454545	0.5	0.625	0.757576
fold #2	0.59596	0.555556	0.636364	0.555556	0.636364	0.656566
fold #3	0.383838	0.222222	0.545455	0.285714	0.461538	0.40404
fold #4	0.530303	0.333333	0.727273	0.5	0.571429	0.575758
fold #5	0.540404	0.444444	0.636364	0.5	0.583333	0.434343
fold #6	0.729167	0.625	0.833333	0.714286	0.769231	0.729167
fold #7	0.541667	0.5	0.583333	0.444444	0.636364	0.645833
fold #8	0.75	0.5	1	1	0.75	0.75
fold #9	0.693182	0.75	0.636364	0.6	0.777778	0.761364
Mean	0.5799874	0.481944	0.67803	0.55	0.634437	0.639142
SD	0.11530471	0.179788	0.15432	0.195223	0.104809	0.130109

Table 5.24 Evaluation of the carotid artery disease risk prediction problem based on the vulnerability of plaques over the 10-folds, regarding Subproblem 5.

	Subproblem 5					
	Accuracy	**Sensitivity**	**Specificity**	**PPV**	**NPV**	**AUC**
fold #0	0.535714	0.357143	0.714286	0.714286	0.357143	0.530612
fold #1	0.607143	0.5	0.714286	0.777778	0.416667	0.612245
fold #2	0.452381	0.571429	0.333333	0.666667	0.25	0.607143
fold #3	0.47619	0.285714	0.666667	0.666667	0.285714	0.440476
fold #4	0.380952	0.428571	0.333333	0.6	0.2	0.416667
fold #5	0.554945	0.538462	0.571429	0.7	0.4	0.659341
fold #6	0.664835	0.615385	0.714286	0.8	0.5	0.747253
fold #7	0.730769	0.461538	1	1	0.5	0.78022
fold #8	0.543956	0.230769	0.857143	0.75	0.375	0.373626
fold #9	0.368132	0.307692	0.428571	0.5	0.25	0.241758
Mean	0.5315017	0.42967	0.633333	0.71754	0.353452	0.540934
SD	0.116715996	0.130302	0.218852	0.132455	0.104952	0.172064

Table 5.25 Evaluation of the carotid artery disease risk prediction problem based on the vulnerability of plaques over the 10-folds, regarding Subproblem 6.

	Subproblem 6			
	Accuracy	**Sensitivity**	**Specificity**	**AUC**
fold #0	0.66	1	0.32	0.83
fold #1	0.675	0.6	0.75	0.7
fold #2	0.754167	0.8	0.708333	0.75
fold #3	0.429167	0.4	0.458333	0.466667
fold #4	0.654167	0.6	0.708333	0.708333
fold #5	0.729167	0.75	0.708333	0.645833
fold #6	0.6875	0.75	0.625	0.71875
fold #7	0.708333	0.75	0.666667	0.75
fold #8	0.604167	0.5	0.708333	0.53125
fold #9	0.625	0.75	0.5	0.614583
Mean	0.652667	0.69	0.615333	0.671542
SD	0.090732	0.16964	0.141669	0.109226

Table 5.26 Evaluation of the carotid artery disease risk prediction problem based on the vulnerability of plaques over the 10-folds, regarding Subproblem 7.

	Subproblem 6			
	Accuracy	**Sensitivity**	**Specificity**	**AUC**
fold #0	0.75	0.75	0.75	0.775
fold #1	0.75	1	0.5	0.75
fold #2	0.825	0.9	0.75	0.7
fold #3	0.7	0.9	0.5	0.725
fold #4	0.95	0.9	1	0.975
fold #5	1	1	1	1
fold #6	0.616667	0.9	0.333333	0.366667
fold #7	0.783333	0.9	0.666667	0.883333
fold #8	0.697368	0.894737	0.5	0.776316
fold #9	0.796053	0.842105	0.75	0.710526
Mean	0.786842	0.898684	0.675	0.766184
SD	0.115869	0.071385	0.220304	0.17677

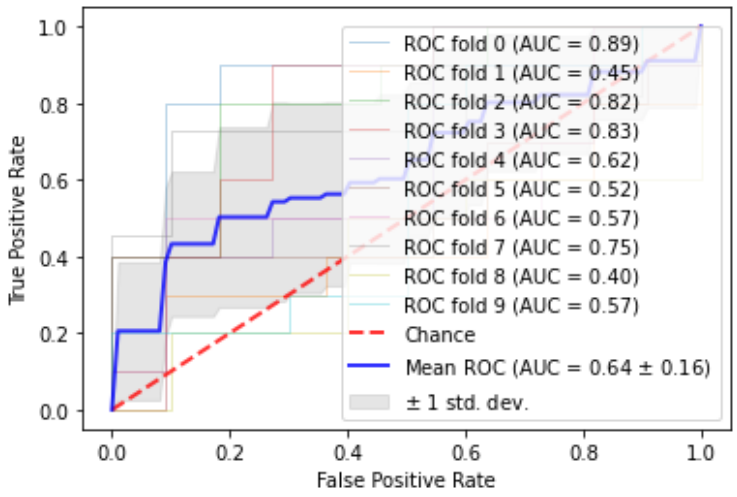

Figure 5. 9 ROC curve analysis regarding Subproblem 1.

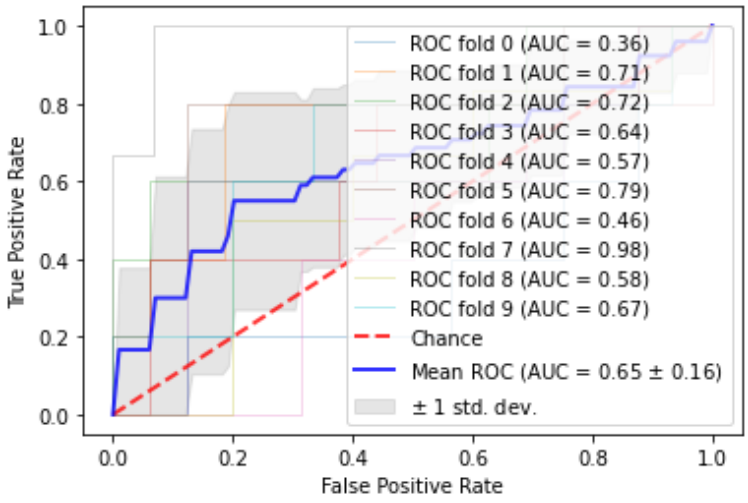

Figure 5.10 ROC curve analysis regarding Subproblem 2.

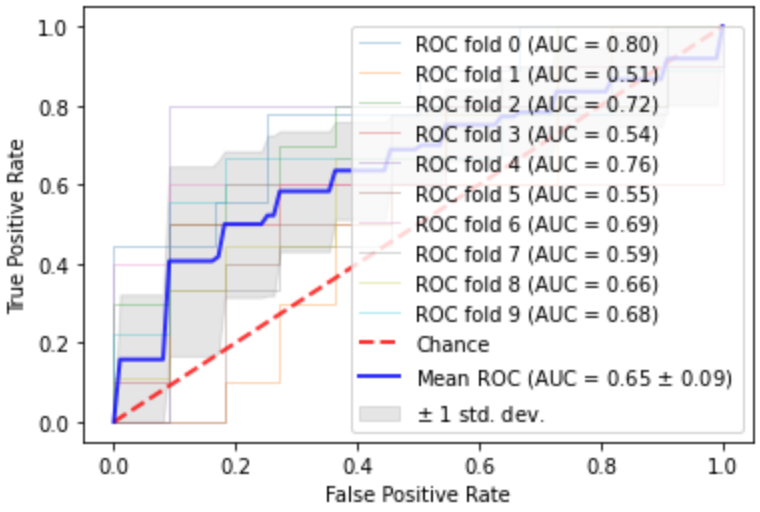

Figure 5.11 ROC curve analysis regarding Subproblem 3.

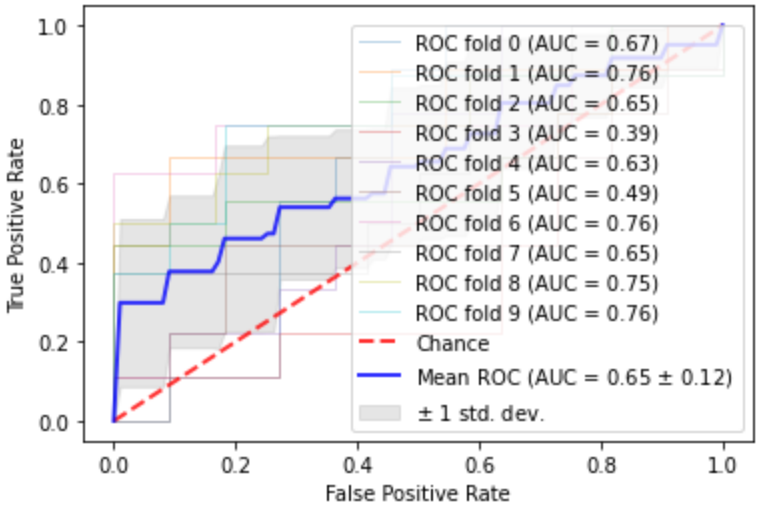

Figure 5.12 ROC curve analysis regarding Subproblem 4.

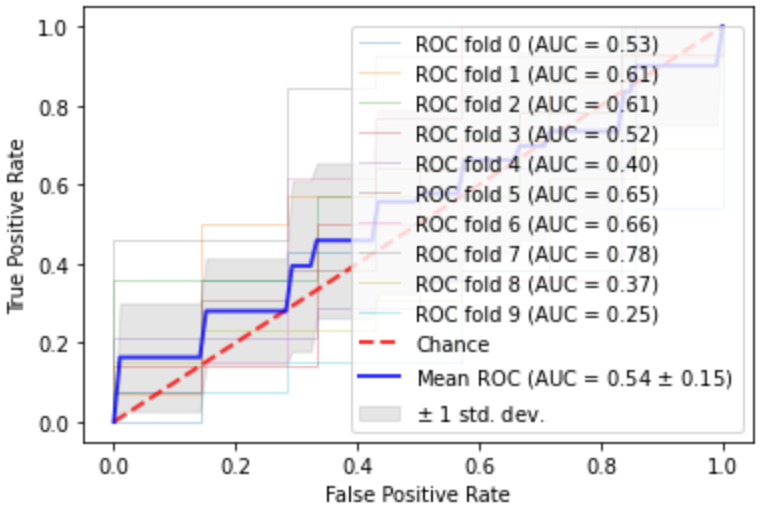

Figure 5. 13 ROC curve analysis regarding Subproblem 5.

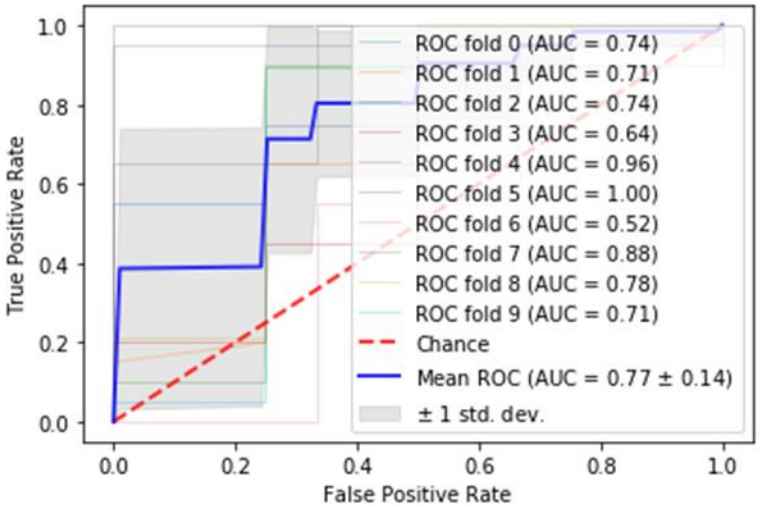

Figure 5.14 ROC curve analysis regarding Subproblem 6.

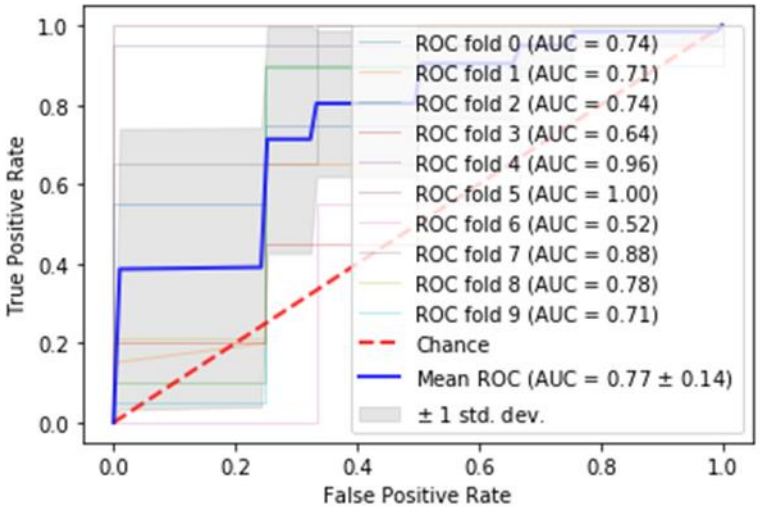

Figure 5. 15 ROC curve analysis regarding Subproblem 7.

Tuning of cut-off threshold values

In Figure 5.16- Figure 5.25, we show the tuning process of threshold values for Subproblem 1-5. Regarding the first approach and the second approach (described in section 5.2.3), cut-off threshold values were defined 0.5, 0.45, 0.52, 0.47, 0.47 and 0.51, 0.54, 0.52, 0.65, 0.48 for Subproblem 1-5, respectively.

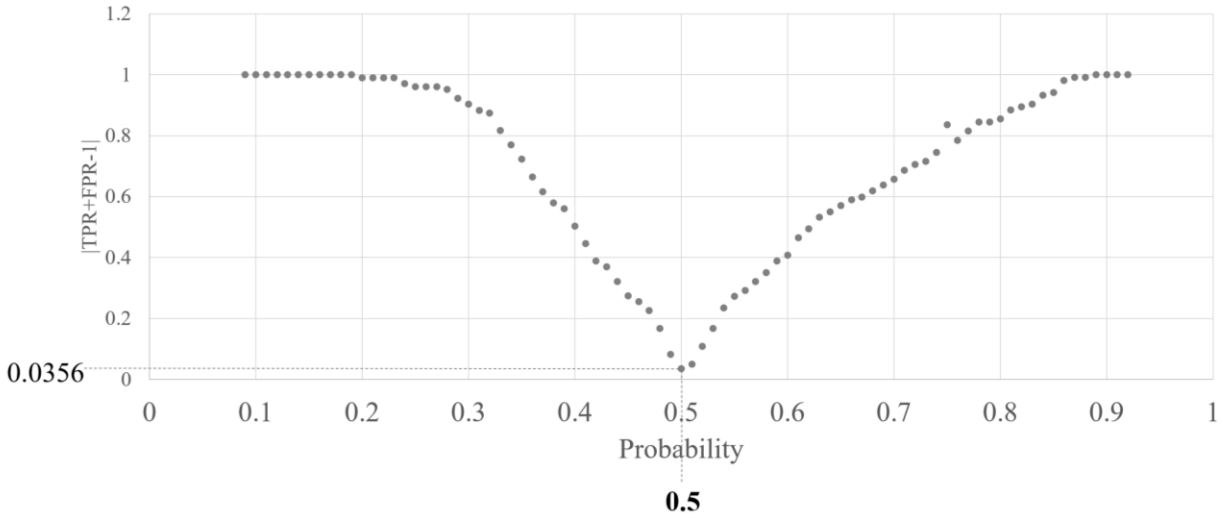

Figure 5.16 Tuning of threshold value based on ROC analysis regarding Subproblem 1.

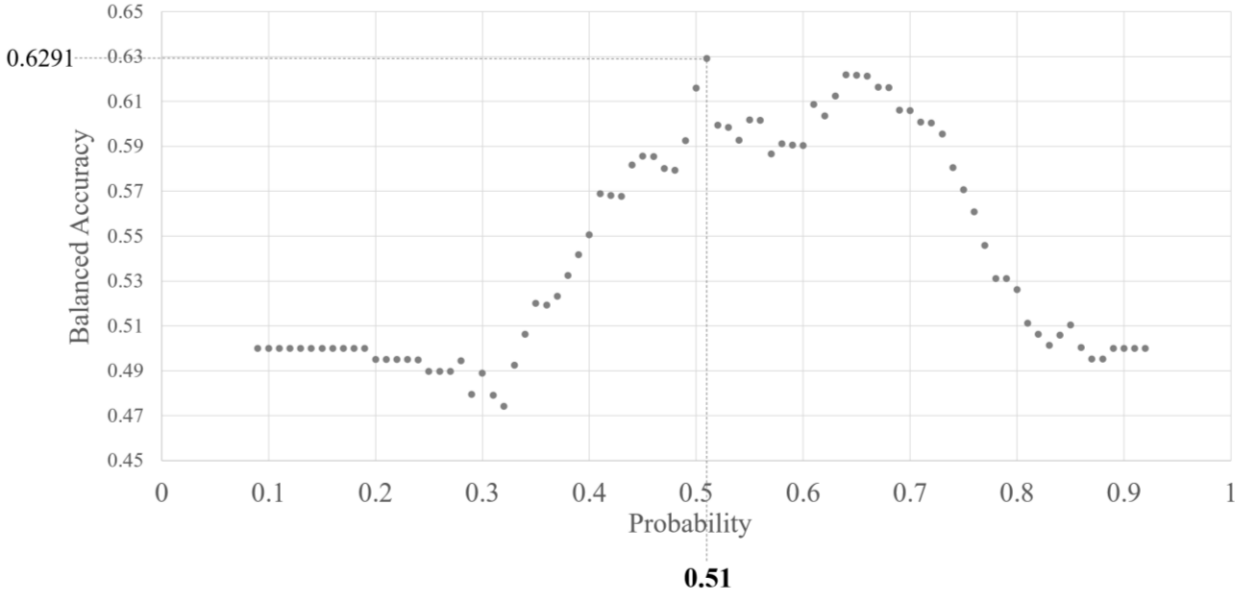

Figure 5.17 Tuning of threshold value based on balanced accuracy regarding Subproblem 1.

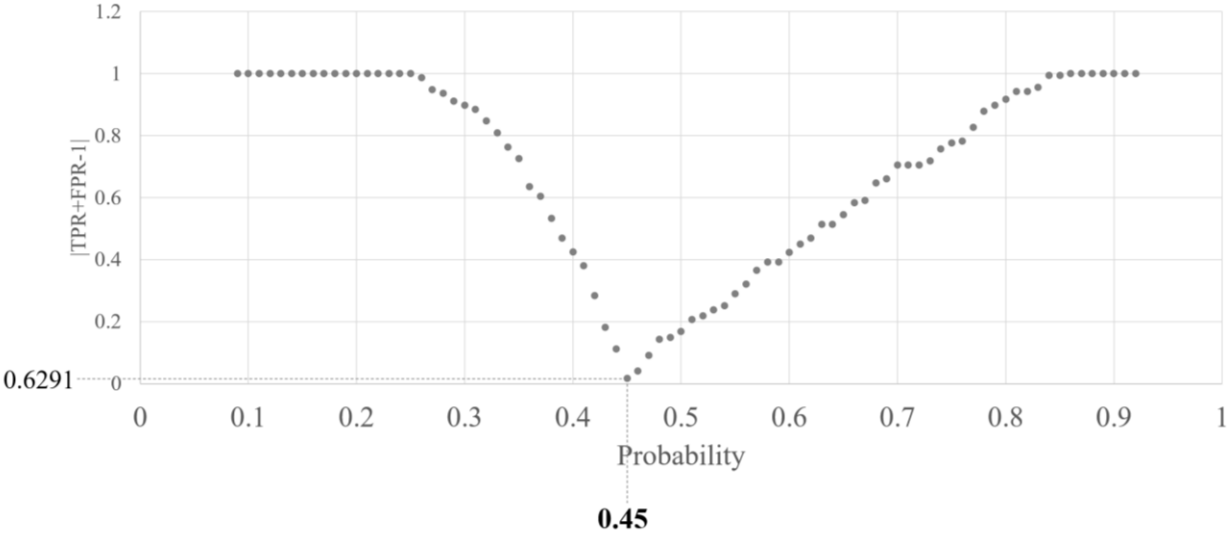

Figure 5. 18 Tuning of threshold value based on ROC analysis regarding Subproblem 2.

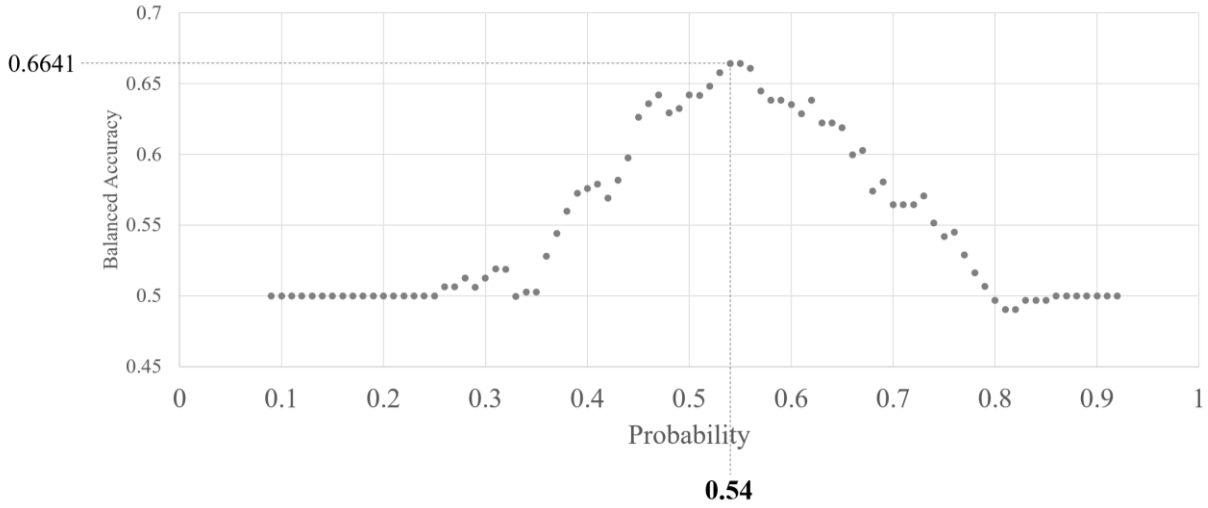

Figure 5. 19 Tuning of threshold value based on balanced accuracy regarding Subproblem 2.

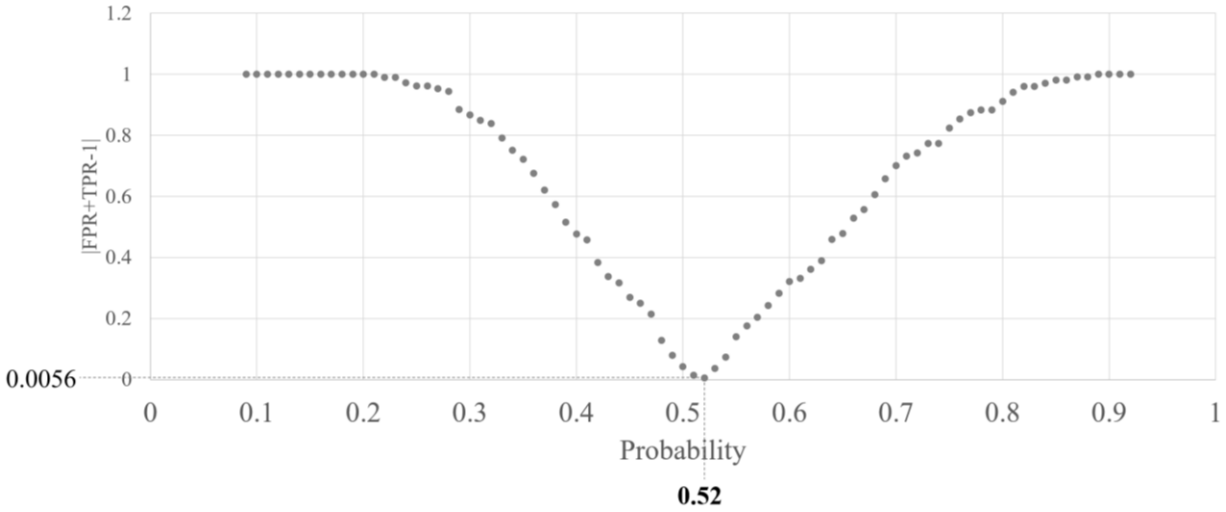

Figure 5. 20 Tuning of threshold value based on ROC analysis regarding Subproblem 3.

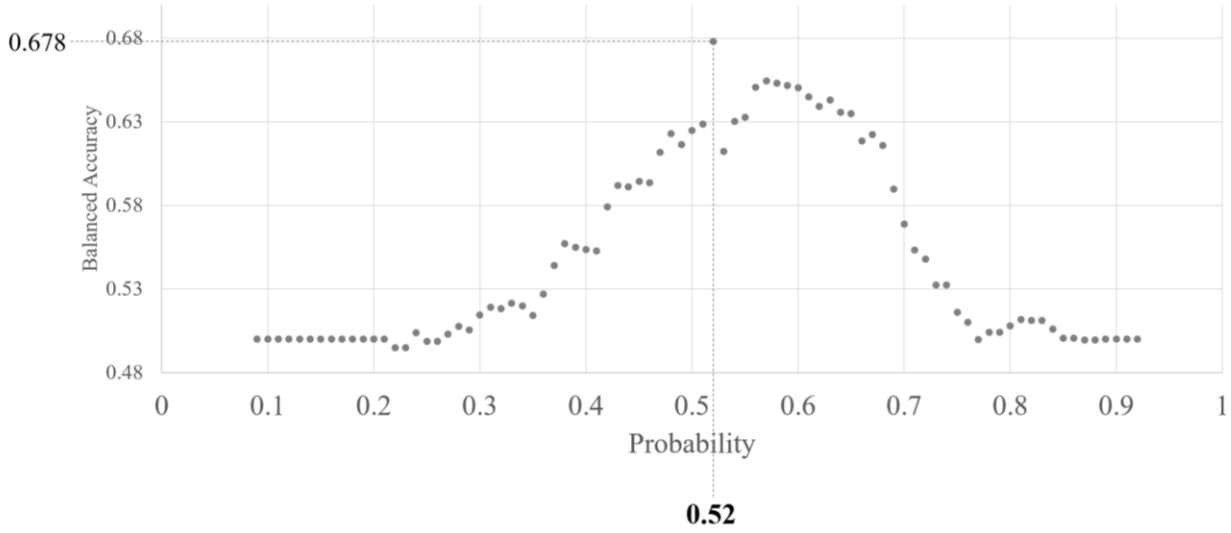

Figure 5. 21 Tuning of threshold value based on balanced accuracy regarding Subproblem 3.

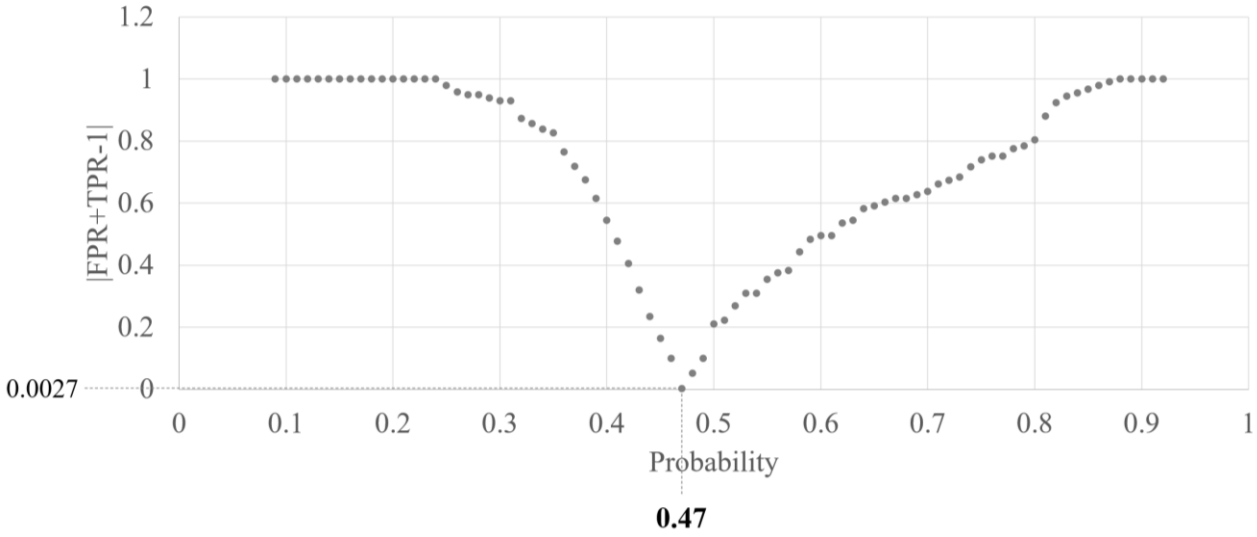

Figure 5. 22 Tuning of threshold value based on ROC analysis regarding Subproblem 4.

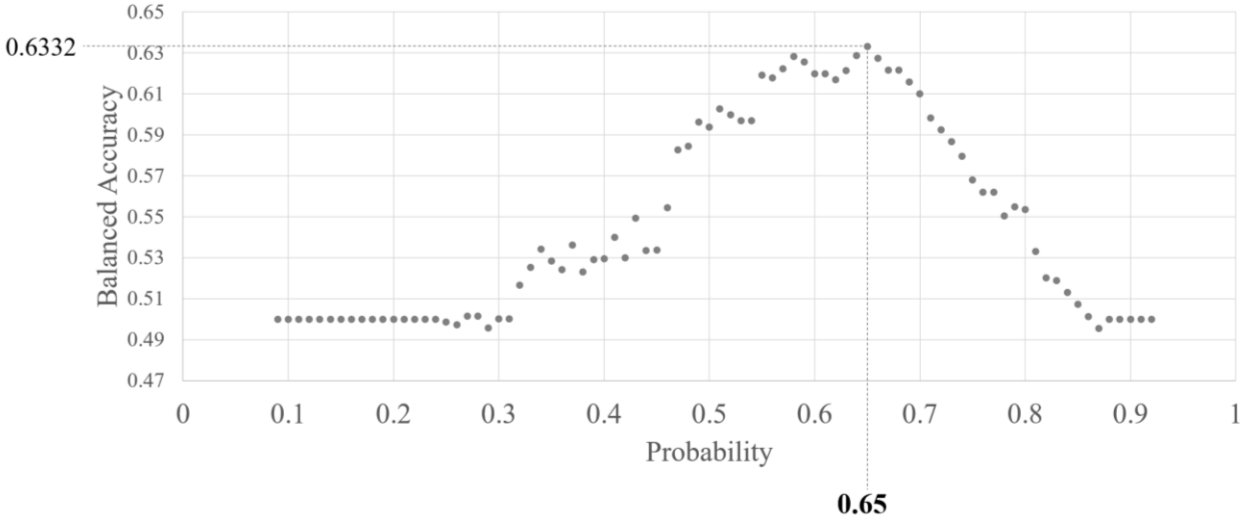

Figure 5. 23 Tuning of threshold value based on balanced accuracy regarding Subproblem 4.

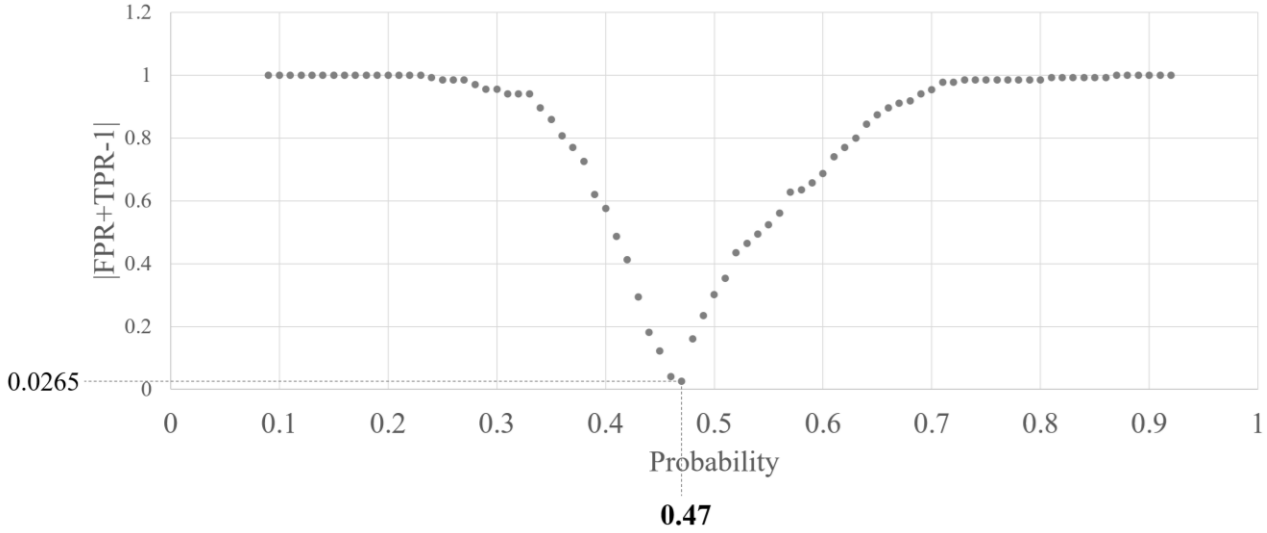

Figure 5. 24 Tuning of threshold value based on ROC analysis regarding Subproblem 5.

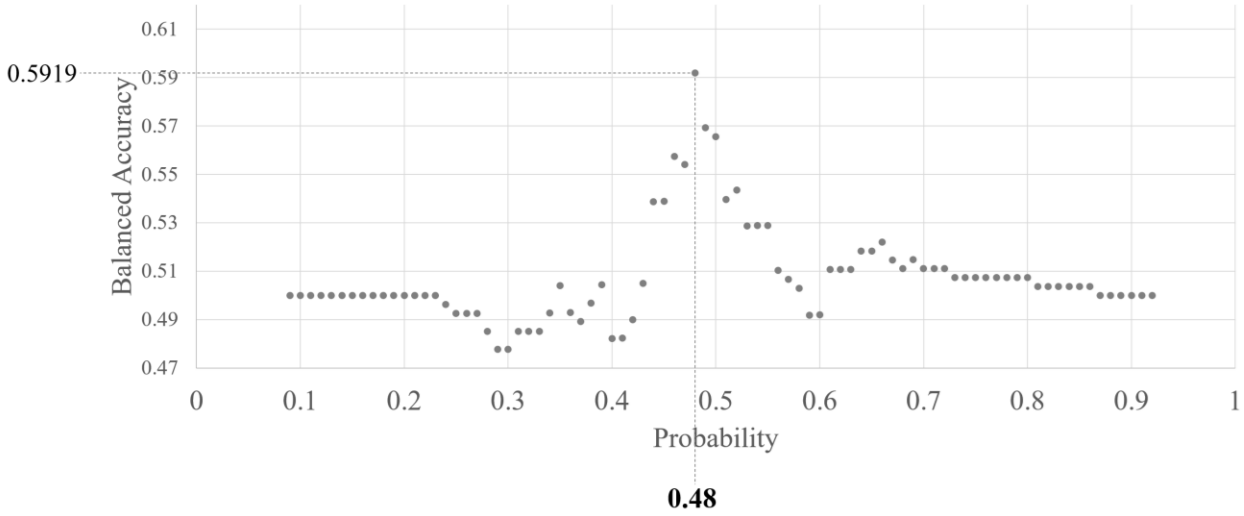

Figure 5. 25 Tuning of threshold value based on balanced accuracy regarding Subproblem 5.

5.5 Discussion

In this section, we have focused on the design and evaluation of ML based models for the prediction of carotid artery disease. Having formulated the specified problems as binary classification problems according to the presence of CAS over 50%, the vulnerability and the stability of the atherosclerotic plaques, we methodically developed a carotid artery disease risk stratification methodology encompassing data pre-processing, feature evaluation, class imbalance handling and classification steps. Different feature selection techniques have been implemented for the identification of the most informative input features, aiming to achieve the highest classification accuracy. In this direction, we conclude that most of the selected predictive models incorporate RF-based feature ranking, which was able to effectively detect conditional dependencies between input and output variables, taking also into account the underlying feature inter-correlations. On top of these, a complete data pre-processing procedure was also applied to resolve issues concerning missing data, unbalanced classes, and, as a result, improve the quality of the dataset.

In addition to this, the input data of all the proposed machine learning models are typical health records, such as demographics data, clinical data, risk factors and medical therapy data. Thus, an innovative concept of the proposed models is their ability to either identify individuals with the presence of CAS or with high risk of unstable plaques with the minimum cost and safety.

In addition to this, all the presented models have been integrated in a cloud-based platform, the Taxinomisis platform, and except for their easy applicability, the proposed models provide an explainability to the users. More specifically, a SHAP analysis was implemented to identify the most significant features in each developed model.

Except for the studies aiming to identify the biomarkers associated with carotid artery disease diagnosis, there have been presented also studies aiming to identify CAS based on ML techniques. More specifically, Jun Xiong Yin *et al.* [271] concluded that family history of dyslipidemia, high level of LDL, low level of HDL, aging, and low BMI are the most significant risk factors of asymptomatic CAS. Yu *et al.* [272] implemented typical ML classifiers to identify subjects with CAS, using 17 candidate input features and achieved the highest accuracy of 0.748. In a same way, Fan *et al*. [273] implemented six different ML models to predict the asymptomatic CAS patients and LR showed the optimal performance in predicting asymptomatic

CAS, with an accuracy of 0.747. On the other hand, Poorthuis *et al*. [274] presented a model for the detection of CAS, including as input features age, sex, smoking, hypertension, hypercholesterolemia, DM, vascular and cerebrovascular disease, measured blood pressure, and blood lipids and achieving an area under the receiver operating characteristic curve of 0.75 (95% CI, 0.74-0.75) for ≥50% asymptomatic CAS.

In the overall proposed pipeline for the detection of asymptomatic carotid artery disease, the highest accuracy and AUC were 0.78 and 0.85, respectively and were achieved implementing the RF selection technique and the Gradient boosting classification scheme. In addition to the results using the training dataset, an external validation dataset of 521 individuals was used for evaluation of the proposed model. Accuracy, sensitivity and specificity were 0.88, 0.88 and 0.84, respectively. In Table 5.27, we summarize the existing machine learning based studies for the diagnosis of CAS, in comparison with our overall proposed approach. As it is shown, the overall evaluation metrics of the proposed pipeline were higher than those provided in other recent studies in the literature [271-274].

As for the diagnosis of high risk plaques, seven different ML models have been developed for the early detection of individuals with high risk plaques. The output of each model has been defined based on different histological based features and presence of symptoms.

The majority of the existing studies take into account imaging based features to characterize the high risk plaques and their aim is to associate the high risk plaques with different biomarkers [150] [149] [170]. However, in the presented ML models, we aim not to identify a relation between the output and specific biomarkers, but to identify the output (high risk plaques), based on defined input features. The average accuracy of the proposed seven models was 0.64.

Regarding the diagnostic prediction of symptomatic carotid artery disease (Subproblem 6, Subproblem 7), the input includes atherosclerosis related serum markers and the achieved accuracy was 0.65 and 0.79 for Subproblem 6 and 7, respectively.

Table 5. 27 Summary of the existing in the literature machine learning based studies in comparison with our proposed approach.

Study	**Methodology**	**Dataset**	**Output-Results**
Jun Xiong Yin *et al.* [271]	ML	2841 high risk individual of stroke enrolled, 326 (11.6%) were diagnosed as ACS by ultrasonography	concluded that family history of dyslipidemia, high level of LDL, low level of high HDL, aging, and low BMI are the most significant risk factors of ACS, AUC=0.87
Yu *et al.* [272]	ML	17 candidate input features, 2732 asymptomatic subjects	Accuracy= 0.748
Fan *et al.* [273]	ML	18,441 subjects, 6553 were diagnosed with asymptomatic CAS, input risk factors & biomarkers	Accuracy= 0.747
Poorthuis *et al.* [274]	Systematic review	input risk factors & biomarkers	AUC=0.75
Our Appproach (subproblem #1)	ML	Demographics, Clinical Data, Risk Factors, Medication Therapy, 881 patients	Accuracy= 0.78, AUC=0.85

Chapter 6 Silent brain lesions detection

6.1 Introduction

6.2 Association of carotid artery disease with silent brain lesions

6.3 Discussion

6.1 Introduction

Despite the association of the presence of CAS with stroke manifestation, the relation between the carotid territory condition and the presence of clinically asymptomatic ischemic events remains unclear. More specifically, these clinically asymptomatic lesions are small, radiologically-detected infarctions, detected by brain imaging without a corresponding clinical manifestation. Their prevalence is estimated about 10–20% with a yearly increase of 3–4% in population-based cohorts, while magnetic MRI-defined SBIs increased the risk of incident stroke, ischemic stroke, intracerebral hemorrhage, and death [275]. In a meta-analysis study conducted by Finn *et al*. [275], it was found that both carotid IMT and CAS are both are significantly associated with silent brain infarcts (SBIs). In another study, SBIs and white matter brain lesions were investigated in patients with asymptomatic CAS >50% and they concluded that DS may be relevant in the association between asymptomatic CAS and SBIs, and DS ≥ 70% may pose a risk of SBI development [276]. Similarly, Rudolph *et al.* [277] found that SBI is a significant risk factor for future stroke, with odds ratio (OR) 4.6 (95% CI: 3.0–7.2;$p < 0.0001$), with an estimated prevalence in asymptomatic carotid patients of 17–33.3%.

In this chapter our aim is to study the association of the presence of SBIs with the carotid territory, as it is defined by US-based parameters and to associate demographics, clinical data and CVD-related risk factors with the presence of SBIs.

6.2 Association of carotid artery disease with silent brain lesions

6.2.1 Brain MRI analysis

Brain MRI imaging was performed in individuals, participating in Taxinomisis clinical study. More specifically, individuals with moderate to severe extracranial CAS, both asymptomatic and symptomatic, were enrolled in the prospective observational multi-center trial in six European vascular centers (Athens-NKUA, Barcelona-FCRB, Belgrade-UBEO, Genoa-USMI, Munich-TUM, and Utrecht-UMC). Inclusion lasted from 30.3.2018 to 31.12.2019. All study participants have underwent MRI imaging of brain and carotid arterial tree from aortic arch up to the circle of Willis. Diffusion weighted imaging has been performed additionally to detect acute ischemic brain lesions.

Regarding the MRI acquisition protocol, 2D and 3D Time-of-flight (TOF) imaging was performed to identify the carotid bifurcation and disease affected section of artery, as it is shown in Figure 6.1. Axial T1-weighted imaging is performed through this segment of artery (Figure 6. 2). Selected matched T2- and proton density-weighted imaging was also performed. Phase contrast imaging is carried out at locations in the CCA and the ICA to provide inlet and outlet flow profiles.

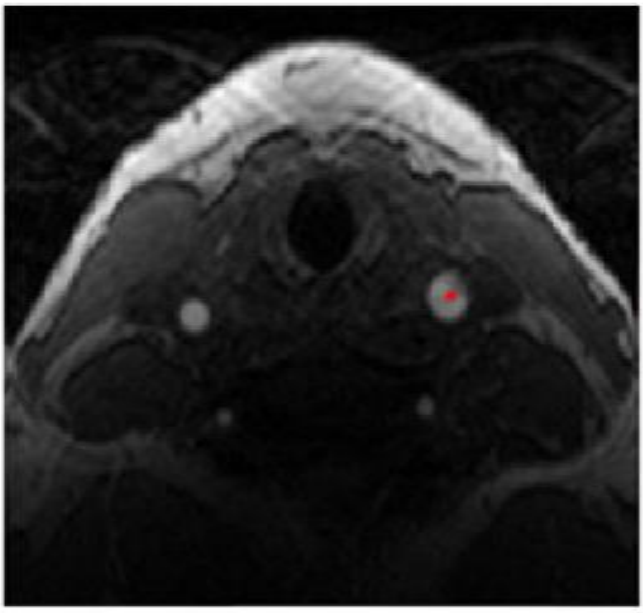

Figure 6.1: Example of 3D TOF image, where the red point depicts the lumen of the CCA.

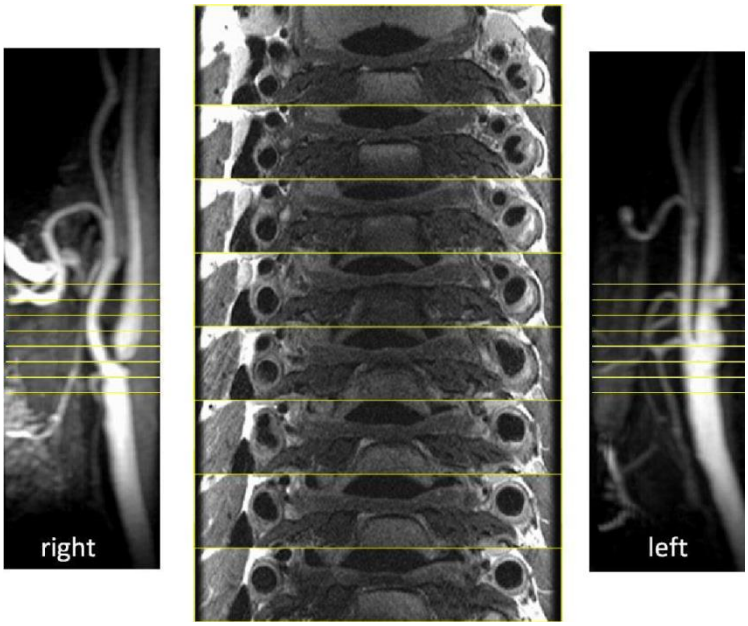

Figure 6. 2: 2D TOF images of the right and left carotid arteries are shown at the sides of the image with the axial slice locations for the T1-weighted sequence illustrated in yellow. The corresponding T1-weighted slices, from CCA to ICA, are shown in the centre of the image.

The acquisition parameters for the MRI sequences were as follows: (i) TOF images: repetition time: 23 ms, effective echo time: 3.2 ms, field of view (FOV): 160 mm; slice thickness 0.5 mm (ii) fast-spin echo double-inversion recovery prepared sequences (T1-weighted): repetition time: 1428.57 ms, effective echo time: 7.672 ms, FOV: 100 mm, slice thickness: 2.5 mm (or lower); (iii) T2-weighted : repetition time: 1379.31 ms, effective echo time: 99.74 ms, FOV: 100 mm, slice thickness: 2.5 mm (or lower); and (iv) proton density-weighted sequences: repetition time: 1379.31 ms, effective echo time: 7.67 ms. Phase-contrast images were acquired in CCA and the ICA. The acquired data were stored in DICOM format and then transferred to a workstation for further analysis.

Brain MRI analysis consisted of identification of different brain lesions in terms of their localization, time of occurrence, and correlation with symptoms and carotid plaque features seen both on MRI and US. More specifically, an expert neuroradiologist annotated different brain lesions categories:

1. Chronic white matter ischemia, which were defined as a white matter hyperintensity of presumed vascular origin and were recognized on MRI as a signal abnormality of variable size in the white matter. These lesions have the following characteristics: hyperintensity on T2-weighted images

such as fluid-attenuated inversion recovery, without cavitation (signal different from CSF). Lesions in the subcortical grey matter or brainstem are not included in this category unless explicitly stated. The utilized scale for this type of lesions was the Modified Fazekas scale (1987) (Figure 6.3)

2. Cortical infarcts, which affect the cerebral cortex and are typically presented with deficits such as neglect and aphasia.
3. Lacunar infarcts, which occur when an artery to the deep part of the brain, containing structures like the thalamus or basal ganglia, is blocked. These arteries are very small and branch off directly from a larger artery, making them particularly vulnerable to blockages.
4. Subcortical infarcts, which affect the small vessels deep in the brain, and typically present with purely motor hemiparesis affecting the face, arm, and leg (Figure 6. 4)
5. Other data that were collected related to posterior circulation stroke or watershed territory lesions in patients (either cortical or lacunar). (Figure 6.5)

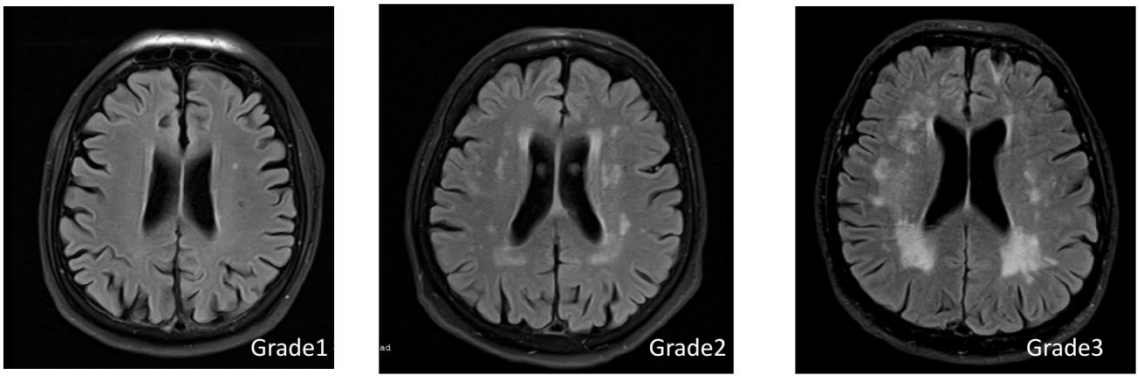

Figure 6.3 Chronic white matter ischemia.

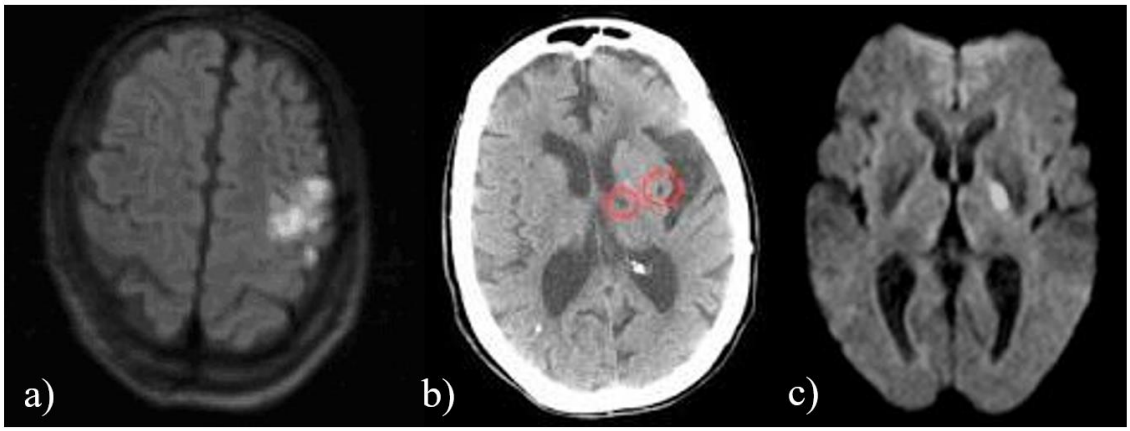

Figure 6. 4 Examples of a) cortical, b) lacunar and c) small subcortical infarcts.

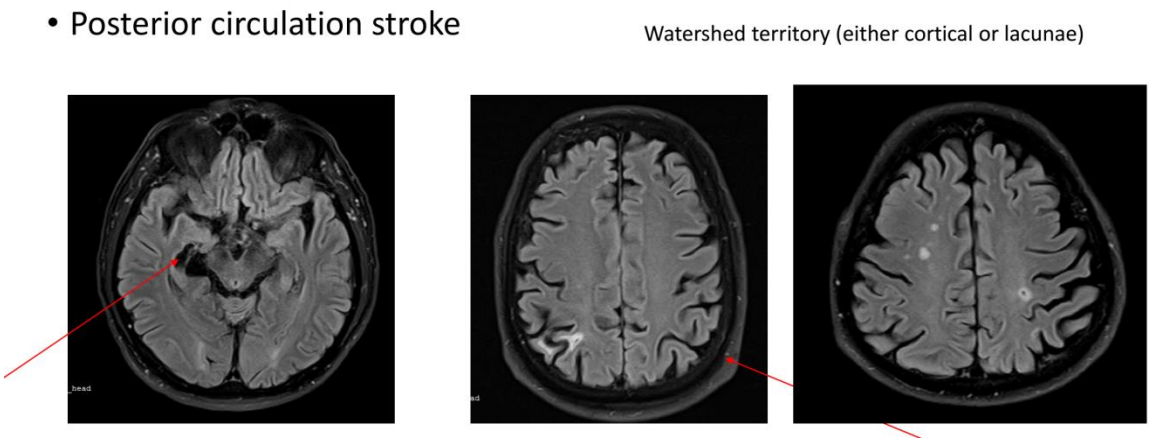

Figure 6.5 Posterior circulation stroke.

6.2.2 Methodology

The SPSS 23.0 program (Statistical Package for the Social Sciences, version 23.0 for Windows) was used for the analyses. Normal distribution of continuous variables was tested using the Kolmogorov–Smirnov and Shapiro–Wilk tests. Descriptive statistics were presented as mean (SD) for continuous variables and as number (%) for categorical variables. Mann–Whitney U test was used for continuous variables Pearson's chi-squared test and OR analysis were used for categorical variables. The statistical significance level was accepted as $p<0.05$.

6.2.3 Dataset

Assessment of all baseline brain MRI images has been performed by expert radiologist from Bologna group. Analysis of the baseline brain MRI images was performed in cases provided by 4 clinical centers (UBEO, FCRB, USMI and TUM), participating in Taxinomisis clinical study, while brain images from UMC and NKUA are not suitable for brain tissue assessment. Regarding the location of brain lesions, neuroradiologists annotated cortical, lacunar and subcortical infarcts. Active lesions were reported as active if the stenosis and plaque morphology corresponded to the brain lesions, as shown in Figure 6.6. All brain images of the baseline were assessed by automatic segmentation and the quantification of "white matter lesion of presumed vascular origin" was performed using BIANCA software [278].

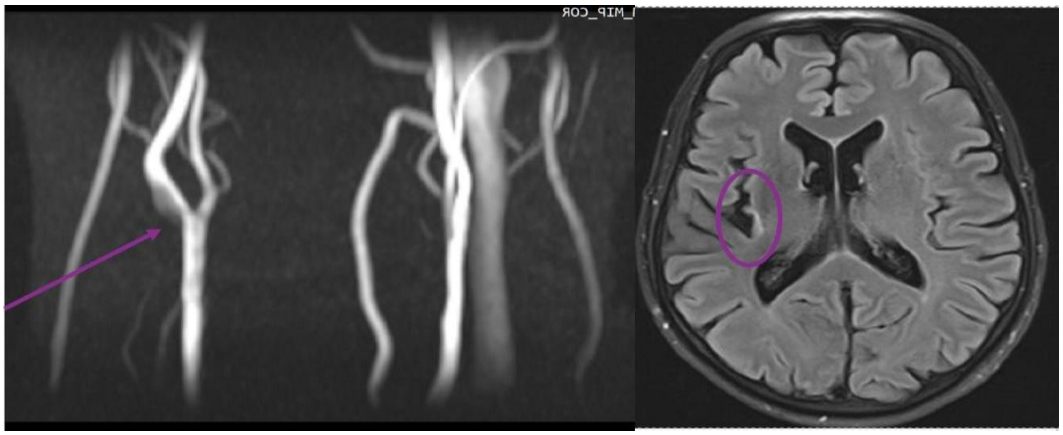

Figure 6.6 Plaque stenosis (arrow) on the left side that corresponds to the brain lesion (encircled) on the right side.

In Table 6.1-6.3, we show 3 different datasets and their baseline characteristics (mean value for continuous features and percentages for categorical ones), which were used in the proposed approach. The 1st dataset and the 2nd dataset are site specific datasets and aim to associate US based features with the presence of SBIs. On the other hand, the 3rd dataset is a patient specific dataset and aims to associate demographics, clinical data, risk factors and medication therapy with the presence of SBIs. Regarding the Dataset 1, it includes 296 carotid arteries (right and left) and their ipsilateral presence of SBIs. The percentages of brain lesions for cortical, lacunar and subcortical infracts are 13.9%, 1.7% and 7.1%, respectively. As for Dataset 2, we use the CAS of 389 carotid arteries and their ipsilateral presence of brain lesions. 50.4 % of the utilized carotid arteries have a DS over 50%, whereas 14.7% of carotid arteries have a DS over 70%, as it is shown in Table 6. As far as Dataset 3 is concerned, in Table 6.3, we

demonstrate baseline characteristics of clinical data, hematological, biochemical data and CVD related risk factors for 211 patients from 4 different clinical centers.

Table 6.1 Baseline characteristics of Dataset 1.

Overall Cohort (N=296)	
ICA stenosis, mean (SD)	53.63 (25.961)
PSV, mean (SD)	183.630 (97.6744)
ICA/PSV CCA, mean (SD)	3.59 (2.11)
St Mary's ratio, mean (SD)	13.74 (8.59)
Cortical, n (%)	41 (13.9%)
Lacunar, n (%)	5 (1.7%)
Small subcortical, n (%)	21 (7.1%)
Active, n (%)	59 (19.9%)

Table 6.2 Baseline characteristics of Dataset 2.

Overall Cohort (N=389)	
Presence of CAS>50%	196 (50.4%)
Presence of CAS>70%	57 (14.7%)
Cortical , n (%)	47 (12.1%)
Lacunar, n (%)	9 (2.3%)
Small subcortical, n (%)	29 (7.5%)
Active, n (%)	77 (19.8%)

Table 6. 3 Baseline characteristics of Dataset 3.

Overall Cohort (N=211)			
Height, mean (SD)	169.83 (9.15)	Sex (male), n (%)	131 (62.1%)
Weight, mean (SD)	76.36 (11.57)	Alcohol abuse, n(%)	21 (10%)
BMI, mean (SD)	26.23 (3.19)	DM, n (%)	62 (29.4%)
SBP, mean (SD)	131.31 (16.36)	Hypertension, n(%)	180 (85.3%)
DBP, mean (SD)	79.88 (8.62)	Hypercholesterolemia, n(%)	125 (59.2%)
Puls rate, mean (SD)	70.56 (7.15)	COPD, n(%)	25 (11.8%)
Age, mean (SD)	69.86 (7.63)	Coronary disease, n(%)	48 (22.7%)
Hb, mean (SD)	12.16 (2.79)	Angina pectoris, n(%)	50 (23.7%)
Hct, mean (SD)	40.61 (4.69)	Previous MI, n(%)	20 (9.5%)
Creatinine, mean (SD)	86.01 (25.33)	Previous CABG/PCI, n(%)	29 (13.7%)
Cholesterol, mean (SD)	4.61 (1.1)	Atherosclerosis of aortiliac segment or femoropoplitealcrural, n(%)	52 (24.6%)
LDL, mean (SD)	2.56 (0.98)	Aneurysm, n(%)	8 (3.8%)
HDL, mean (SD)	1.49 (0.56)	Statin, n(%)	190 (90%)
Triglycerides, mean (SD)	1.53 (0.81)	Cortical , n (%)	46 (21.8%)
Glucose, mean (SD)	6.44 (2.48)	Lacunar, n (%)	8 (3.8%)
CRP, mean (SD)	3.53 (3.63)	Small subcortical, n (%)	26 (12.3%)
HbA1c, mean (SD)	6.07 (0.89)	Active, n (%)	66 (31.3%)

6.2.4 Results

Regarding the analysis of Dataset 1, as it is shown in Table 6.4, the higher mean values of ICA stenosis and St Mary's ratio are significantly associated with the presence of Cortical infarcts (p=0.014, p=0.004 for stenosis and St. Mary's ratio, respectively), small subcortical infarcts (p=0.018, p=0.013 for stenosis and St. Mary's ratio, respectively) and Active lesions (p=0.002 for both stenosis and St. Mary's ratio).

In Table 6.5, it is clearly shown that the CAS in the ipsilateral carotid artery is considered as a statistically significant risk factor for the presence of cortical infarcts. More specifically, we have observed in 47 (12.1%) out of 389 cortical infarcts, in which 30 of them had an over 50 % DS in their ipsilateral carotid artery (OR: 1.860, CI:0.989-3.499, p=0.05) and 18 of them had an over 70 % DS in their ipsilateral carotid artery (OR: 4.822, CI:2.453-9.481, p<0.001).

As for the 3rd dataset, the mean age of the study's participants was 69.86 ±7.63 in the 211 patients (131 males, 80 females) enrolled in the study. Positive risk factors for the presence of brain lesion was the previous MI (OR: 2.685, CI: 1.022-7.056, p value:0.039 for cortical infarct association and OR:2.481, CI: 0.975-6.311, p value: 0.05 for active lesions association), the alcohol abuse (OR: 3.474 , CI: 1.204-10.018, p value:0.015 for small subcortical infarct association) and the previous CABG/PCI placement (OR: 2.882, CI: 1.074-7.734, p value:0.03 for small subcortical infarct association). In addition to this, statistically significant higher mean values have been reported in the Class of brain lesion presence for the SBP (lacunar infarct, p value: 0.019), triglycerides (lacunar infarct, p value<0.001), Creatinine (small subcortical infarct, p value: 0.001) and HbA1c (small subcortical infarct, p value: 0.008).

Table 6.4 Subjects characteristics in the Class 0 (Absence of Brain lesion) and Class 1 (Presence of Brain lesion) for cortical, lacunar, small subcortical and active infarct, regarding Dataset 1.

	Class 0 (Absence of Cortical) N=248	**Class 0 (Presence of Cortical) N=41**	**P value**	**Class 0 (Absence of Lacunar) N=291**	**Class 0 (Presence of Lacunar) N=6**	**P value**
ICA stenosis, mean (SD)	52.22	59.44	**0.014**	53.64 (26.07)	52.80 (20.413)	0.572
PSV, mean (SD)	178.88	212.533	0.674	184.15 (97.35)	158.25(126.43)	0.59
PSV ICA/ PSV CCA, mean (SD)	3.4 (2.07)	4.53 (2.2)	0.222	3.6 (2.1)	2.89 (2.78)	0.536
St Mary's ratio, mean (SD)	13.09 (7.78)	16.96 (12.13)	**0.004**	13.76 (8.53)	12.71 (12.78)	0.331
	Class 0 (Absence of small subcortical) N=275	Class 0 (Presence small subcortical) N=21	P value	Class 0 (Absence of Active N=237	Class 0 (Presence of Active N=59	P value
ICA stenosis, mean (SD)	51.90 (33.37)	53.76 (25.38)	**0.018**	53.45 (24.46)	54.34 (31.52)	**0.002**
PSV, mean (SD)	181.35(97.69)	219.08(94.37)	0.79	179.33 (98.77)	201.28 (92.16)	0.619
PSV ICA/ PSV CCA, mean (SD)	3.55 (2.09)	4.15 (2.45)	0.381	3.47 (2.08)	4.11 (2.18)	0.266
St Mary's ratio, mean (SD)	13.46 (8.2)	18.05 (13.04)	**0.013**	13.17 (7.59)	16.06 (11.7)	**0.002**

Table 6.5 Odds Ratio analysis for the association of DS and the presence of Brain lesion, regarding Dataset 2.

	Presence of Cortical Infarct	**Odds Ration –CI**	**P value**	**Presence of Lacunar Infarct**	**Odds Ration –CI**	**P value**
Presence of CAS >50%, n (%)	30 (15.3%)	1.860 (989-3.499)	**0.05**	4 (2%)	0.779 (0.206-2.946)	0.712
Absence of CAS >50%, n (%)	17 (8.9%)			5 (2.6%)		
Presence of CAS >70%, n (%)	18 (31.6%)	4.822 (2.453-9.481)	**<0.001**	1 (1.8%)	0.723 (0.089-5.895)	0.761
Absence of CAS >70%, n (%)	29 (8.7%)			8 (2.4%)		
	Presence of Small subcortical Infarct	**Odds Ration –CI**	**P value**	**Presence of Active Lesion**	**Odds Ration –CI**	**P value**
Presence of CAS >50%, n (%)	12 (6.1%)	0.671 (0.312-1.446)	0.306	38 (19.8%)	1.007 (0.611-1.658)	0.979
Absence of CAS >50%, n (%)	17 (8.9%			39 (19.9%)		
Presence of CAS >70%, n (%)	6 (10.5%)	1.581 (0.614-4.071)	0.339	19 (33.3%)	2.362 (1.272-4.388)	
Absence of CAS >70%, n (%)	23 (6.9%)			58 (17.5%)		

Table 6.6 Subjects characteristics for continuous input features in the Class 0 (Absence of Brain lesion) and Class 1 (Presence of Brain lesion) for cortical, lacunar, small subcortical and active infarct, regarding Dataset 2.

	Class 0 (Absence of Cortical) N=165	**Class 0 (Presence of Cortical) N=46**	**P value**	**Class 0 (Absence of Lacunar) N=203**	**Class 0 (Presence of Lacunar) N=8**	**P value**
Height	169.42 (8.85)	171.16 (10.07)	0.209	169.81 (9.197)	170.57 (8.26)	0.757
Weight	76.03 (12.08)	77.47 (9.74)	0.111	76.009 (11.49)	85.14 (9.84)	0.392
BMI	26.31 (3.37)	25.96 (2.45)	0.132	26.16 (3.16)	28.37 (3.73)	0.346
SBP	132.18 (16.71)	128.16 (14.8)	0.325	131.16 (16.54)	134.75 (14.64)	**0.019**
DBP	80.46 (8.09)	77.81 (10.13)	0.055	79.73 (8.66)	81.50 (7.65)	0.473
Puls rate	70.66 (7.54)	70.20 (5.6)	0.255	70.43 (7.26)	72 (2.24)	0.118
Age	70.25 (7.91)	68.48 (6.43)	0.242	69.82 (7.55)	73.38 (8.62)	0.552
Hb	11.98 (2.8)	12.76 (2.67)	0.141	12.14 (2.79)	11.84 (2.92)	0.931
Hct	40.49 (4.85)	41.02 (4.13)	0.72	40.48 (4.62)	43.3 (6.43)	0.491
Creatinine	86.53 (24.74)	84.03 (27.7)	0.677	85.58 (25.31)	98.11 (26.52)	0.639
Cholesterol	4.6 (1.08)	4.62 (1.2)	0.355	4.61 (1.1)	4.15 (1.22)	0.928
LDL	2.54 (0.99)	2.63 (0.94)	0.811	2.57 (0.99)	2.18 (0.74)	0.447
HDL	1.5 (0.56)	1.42 (0.56)	0.901	1.5 (0.57)	1.22 (0.44)	0.363
Triglycerides	1.52 (0.8)	1.56 (0.83)	0.726	1.49 (0.73)	2.44 (1.75)	**<0.001**
Glucose	6.56 (2.35)	6.03 (2.88)	0.835	6.46 (2.5)	6.12 (2.43)	0.962
CRP	3.55 (3.68)	3.46 (3.49)	0.864	3.55 (3.7)	2.96 (1.77)	0.481
HbA1c	6.09 (0.85)	6.03 (1.03)	0.594	6.08 (0.89)	6.15 (1.48)	0.371
	Class 0 (Absence of Small subcortical) N=185	**Class 0 (Presence of Small subcortical) N=26**	**P value**	**Class 0 (Absence of Active) N=145**	**Class 0 (Presence of Active) N=66**	**P value**
Height	170.06 (8.83)	168.74 (11.22)	**0.048**	169.78 (8.64)	169.93 (10.24)	**0.037**
Weight	76.13 (11.55)	78.09 (12.02)	0.723	75.86 (11.88)	77.42 (10.92)	0.483
BMI	26.09 (3.17)	27.07 (3.31)	0.516	26.18 (3.28)	26.36 (2.99)4	0.875
SBP	130.74 (16.28)	135.64 (16.81)	0.504	131.08 (16.29)	131.81 (16.64)	0.448
DBP	79.60 (8.55)	81.44 (9.05)	0.879	80.13 (8.08)	79.34 (9.75)	0.134
Puls rate	70.28 (7.45)	72.38 (4.64)	0.182	70.51 (7.85)	70.67 (5.3)	0.056
Age	69.92 (7.75)	69.42 (7.03)	0.517	70.12 (7.85)	69.30 (7.16)	0.585
Hb	12.07 (2.82)	12.74 (2.6)	0.227	11.98 (2.82)	12.54 (2.69)	0.276
Hct	40.53 (4.65)	41.42 (4.94)	0.377	40.34 (4.65)	41.22 (4.78)	0.457
Creatinine	84.71 (22.95)	96.71 (38.09)	**0.001**	85.72 (24.46)	86.69 (27.47)	0.686
Cholesterol	4.65 (1.11)	4.29 (1.05)	0.85	4.64 (1.08)	4.53 (1.17)	0.431
LDL	2.59 (0.99)	2.3 (0.85)	0.404	2.58 (1)	2.51 (0.94)	0.773
HDL	1.51 (0.57)	1.34 (0.47)	0.711	1.53 (0.57)	1.38 (0.53)	0.724
Triglycerides	1.52 (0.81)	1.59 (0.82)	0.967	1.45 (0.72)	1.71 (0.97)	0.058
Glucose	6.42 (2.45)	6.73 (2.73)	0.383	6.5 (2.33)	6.32 (2.79)	0.518
CRP	3.59 (3.77)	3.18 (2.61)	0.437	3.71 (4.1)	3.15 (2.31)	0.178
HbA1c	6.03 (0.79)	6.45 (1.28)	**0.008**	6.03 (0.83)	6.18 (1.01)	0.263

Table 6. 7 Odds Ratio analysis for the association of CVD related risk factors and the presence of Brain lesions, regarding Dataset 3.

	Presence of Cortical Infarct	**Odds Ration - CI**	**P value**	**Presence of Lacunar Infarct**	**Odds Ration - CI**	**P value**
Sex (male), n(%)	32 (24.4%)	0.656(0.326-1.323)	0.237	8 (6.2%)	0.938 (0.898-0.981)	**0.025**
Sex (female), n(%)	14 (17.5%)			0 (0%)		
Alcohol abuse (no), n(%)				7 (3.8%)	1.257 (0.147-10.746)	0.834
Alcohol abuse (yes), n(%)				1 (4.8%)		
Diabetes mellitus (no), n(%)	36 (24.2%)	0.604 (0.278-1.309)	0.198	6 (4.1%)	0.783 0.154-3.992	0.768
Diabetes mellitus (yes), n(%)	10 (16.1 %)			2 (3.2%)		
Hypertension (no), n(%)	7 (23.3%)	0.909 (0.363-2.274)	0.838	0 (0%)	1.047 (1.014-1.080)	0.246
Hypertension (yes), n(%)	39 (21.7%)			8 (4.5%)		
COPD (no), n(%)	43 (24.2%)	0.131 (0.017-0.996)	**0.022**	7 (4%)	1.006 (0.119-8.537)	0.996
COPD (yes), n(%)	1 (4%)			1 (4%)		
Coronary disease (no), n(%)	31 (20%)		0.298	5 (3.3%)	1.973 (0.454-8.58)	0.357
Coronary disease (yes), n(%)	13 (27.1%)			3 (6.3%)		
Angina pectoris (no), n(%)	26 (26.8%)	0.372(0.142-0.976)	**0.039**	2 (2.1%)	2.968 (0.479-18.378)	0.222
Angina pectoris (yes), n(%)	6 (12%)			3 (6%)		
Previous MI (no), n(%)	36 (19.9%)	2.685 (1.022-7.056)	**0.039**	8 (4.5%)	0.955 (0.926-0.986)	0.335
Previous MI (yes), n(%)	8 (40%)			0 (0%)		
Previous Cabg/pci (no), n(%)	36 (20.9%)	1.439 (0.589-3.516)	0.423	6 (3.5%)	2.025 (0.388-10.556)	0.394
Previous Cabg/pci (yes), n(%)	8 (27.6%)			2 (6.9%)		
Atherosclerosis of aortiliac segment or femoropoplitealcrural (no), n (%)	34 (23.4%)	0.683 (0.303-1.543)	0.358	6 (4.2 %)	0.913 (0.178-4.674)	0.913
Atherosclerosis of aortiliac segment or femoropoplitealcrural (yes), n(%)	9 (17.3%)			2 (3.8%)		
Aneurysm (no), n(%)	40 (21.2%)	1.242 (0.241-6.388)	0.795	8 (4.3%)	0.957 (0.929-0.987)	0.55
Aneurysm (yes), n(%)	2 (25%)			0 (0%)		
Statin (no), n(%)	6 (31.6%)	0.578 (0.207-1.616)	0.381	1 (5.3%)	0.696 (0.081-5.98)	0.74
Statin (yes), n(%)	40 (21.1%)			7 (3.7%)		

	Presence of Small subcortical Infarct	**Odds Ration - CI**	**P value**	**Presence of Active Infarct**	**Odds Ration - CI**	**P value**
Sex (male), n(%)	17 (13%)	0.862 (0.364-2.04)	0.736	44 (33.6%)	0.75 (0.407-1.381)	0.355
Sex (female), n(%)	9 (11.4%)			22 (27.5%)		
Alcohol abuse (no), n(%)	19 (10.3%)	3.474 (1.204-10.018)	**0.015**	59 (31.9%)	0.854 (0.316-2.313)	0.756
Alcohol abuse (yes), n(%)	6 (28.6%)			6 (28.6%)		
Diabetes mellitus (no), n(%)	19 (12.8%)	0.864 (0.344-2.173)	0.756	49 (32.9%)	0.771 (0.401-1.483)	0.435
Diabetes mellitus (yes), n(%)	7 (11.3%)			17 (27.4%)		
Hypertension (no), n(%)	2 (6.9%)	2.077 (0.464-9.301)	0.33	6 (20%)	2 (0.776-5.155)	0.145
Hypertension (yes), n(%)	24 (13.3%)			60 (33.3%)		
COPD (no), n(%)	22 (12.4%)	0.613 (0.135-2.78)	0.522	58 (32.6%)	0.394 (0.129-1.201)	0.092
COPD (yes), n(%)	2 (8%)			4 (16%)		
Coronary disease (no), n(%)	15 (9.7%)	2.138 (0.87-5.257)	0.092	43 (27.7%)	1.706 (0.867-3.358)	0.12
Coronary disease (yes), n(%)	9 (18.8%)			19 (39.6%)		
Angina pectoris (no), n(%)	15 (15.6%)	0.6 (0.205-1.759)	0.348	33 (34%)	0.681 (0.319-1.455)	0.32
Angina pectoris (yes), n(%)	5 (10%)			13 (26%)		
Previous MI (no), n(%)	20 (11.1%)	2 (0.608-6.575)	0.246	52 (28.7%)	2.481 (0.975-6.311)	**0.05**
Previous MI (yes), n(%)	4 (20%)			10 (50%)		
Previous Cabg/pci (no), n(%)	17 (9.9%)	2.882 (1.074-7.734)	**0.03**	50 (29.1%)	1.722 (0.767-3.868)	0.184
Previous Cabg/pci (yes), n(%)	7 (24.1%)			12 (41.4%)		
Atherosclerosis of aortiliac segment or femoropoplitealcrural (no), n(%)	18 (12.5%)	0.745 (0.262-2.119)	0.58	49 (33.8%)	0.588 (0.283-1.221)	0.152
Atherosclerosis of aortiliac segment or femoropoplitealcrural (yes), n(%)	5 (9.6%)			12 (23.1%)		
Aneurysm (no), n(%)	22 (11.7%)	1.078 (0.127-9.18)	0.945	58 (30.7%)	0.753 (0.148-3.842)	0.732
Aneurysm (yes), n(%)	1 (12.5%)			2 (25%)		
Statin (no), n(%)	1 (5.3%)	2.744 (0.351-21.468)	0.317	5 (26.3%)	1.324 (0.456-3.843)	0.605
Statin (yes), n(%)	25 (13.2%)			61 (32.1%)		

6.3 Discussion

According to existing in the literature studies and due to the rapid development of MRI scans, there are not specific diagnostic criteria for the detection of SBIs. Accurate diagnosis of SBIs is related to the utilized scan methodology, the size and location of the lesions, the implemented acquisition protocol and modality's resolution. Despite MRI is considered as the gold standard for the detection of SBIs, there are still clinical centers that identify SBIs through the use of CT. This leads to the underestimations of SBIs and to the increase of SBI-related cerebral events [277].

In the literature, there are different approaches that attempted to associate the CAS with the presence of SBIs either in the ipsilateral site or in the contralateral. More specifically, Baradaran *et al.* [279] concluded that the prevalence of SBIs ipsilateral to ICA disease (33%) compared with the contralateral side was higher and the prevalence of cortical SBIs occurring downstream from ICA disease was also higher. This was also confirmed by a study proposed by Müjdat Deniz Benli *et al.* [276], who found that the number of SBIs was in the ipsilateral hemisphere compared to that in the contralateral hemisphere ($p = 0.022$). Based on this conclusion, we investigated the role of US based features of each carotid artery with the SBIs in the ipsilateral hemisphere in brain. The main innovative aspect of our approach is the proposed concept, since there are no other similar approaches in the literature, which are dedicated to associate typical US based metrics with the presence of SBIs. Through our study, it is clear that there is a strong relation between the DS and St. Mary's ratio of carotid artery with the presence of cortical, lacunar and active brain lesions. Additionally, another innovation of our study is the attempt that has been undertaken to characterize the brain lesions according to their localization (cortical, lacunar and small subcortical infarcts). Thus, we have concluded that CAS is a statistically significant measure for all the types of SBIs, except of lacunar lesions. Lacunar lesions have traditionally been considered to arise from microcirculatory disturbances at the level of the small perforating arteries and published literature does not support extracranial carotid disease as an etiological factor for carotid artery disease [280]. Similarly to other published in the literature studies, our findings have confirmed that the presence of CAS in lacunar infarct is an incidental finding and routinely evaluating for carotid artery disease is not required in patients presenting with radiologically confirmed lacunar infarcts [281].

Except for the per site analysis of association between carotid territory and SBIs' presence, a per patient analysis has been performed, aiming to investigate more systemic factors for the pathogenesis of SBIs. Several studies have examined the incidence of SBIs and its relation to risk factors for stroke, with an

increased two- to three-fold risk in the presence of SBI on MRI in elderly populations. Risk factors for SBIs are considered to be comparable to those for stroke and CVD; therefore, CV patients may also be at high risk of silent infarcts with similar risk factors such as ageing, hypertension and DM. Based on this hypothesis, SBI is prevalent following involvement of atherosclerosis in systemic vascular disease and often unnoticed and untreated, this condition may account for undetected cognitive and functional decline as well as unexpected ischaemic stroke [282].

Based on the presence of systemic risk factors for SBIs similar to these of vascular disease, risk factors such as the gender, alcohol abuse, the DM, hypertension, hypercholesterolemia, COPD, CAD, angina pectoris, previous MI, previous CABG/PCI placement, atherosclerosis, aneurysm disease and statin therapy, were investigate with the association of SBIs. Through our study, it was found that the previous MI, the alcohol abuse and previous CABG/PCI placement are strongly related to the presence of SBIs, confirming the relation of systemic vascular disease risk factors with the risk factors for SBIs. Except of risk factors, it was also found that high values of clinical data, such as the SBP, and biochemical data, such as triglycerides, creatine and HbA1c, are associated with the presence of SBIs, assuming that high values of SBP, triglycerides, creatine and HbA1c correspond to hypertension, hypertriglyceridemia, renal dysfunction and DM.

In Table 6.8, we demonstrate a summary of similar in the literature studies, showing that the utilized dataset of our approach has more participants and the output of our study is not limited to the assessment of the contribution of CAS presence to the ipsilateral and contralateral SBIs presence.

Table 6. 8 Comparison of our approach with existing in the literature studies.

Study	**Dataset**	**Methodology**	**Results**
Baradaran *et al.* [279]	104 patients with Asymptomatic Carotid Artery	McNemar test to compare the prevalence of any SBI Wilcoxon signed-rank test to compare the total number of SBIs ipsilateral vs contralateral to extracranial internal carotid artery stenosis	-higher prevalence of SBIs ipsilateral to ICA disease (33%) compared with the contralateral side -no significant difference in the prevalence of lacunar SBIs between hemispheres (P=0.109), -significantly higher prevalence of cortical SBIs occurring downstream from ICA disease (P=0.0045)
Claudina Rudolph et al. [277]		Meta analysis	- The prevalence of SBI in asymptomatic carotid patients is 17–33.3%. -SBI is a significant risk factor for future stroke, OR 4.6 (95% CI: 3.0–7.2; $p < 0.0001$)

Study	Dataset	Methodology	Results
Müjdat Deniz Benli et al. [276]	69 patients (35 females, 34 males) , (15.9%) →SBIs	-Mann–Whitney U test was used for continuous variables -Fisher's exact test or Pearson's chi-squared test was used for categorical variables	-frequency of Fazekas grade 1 DWMLs was lower in the hemisphere ipsilateral to the ICA stenosis compared to the contralateral hemisphere ($p = 0.035$) -number of SBIs was also higher in the ipsilateral hemisphere compared to that in the contralateral hemisphere ($p = 0.022$).
Our study	296 carotid arteries 389 carotid arteries 211 patients	Statistical analysis	Association of CAS, previous MI, alcohol abuse and previous CABG/PCI with presence of SBIs High values of SBP, triglycerides, creatine and HbA1c are associated with the presence of SBIs

Chapter 7 Conclusions and future work

7.1 Conclusions

7.2 Future work

7.1 Conclusions

The aim of this thesis is to present different methodologies for the predictive modeling of atherosclerosis, developing models both for the CAD and carotid artery disease prediction. Overall pipelines based on ML techniques have been proposed for CAD and carotid disease prediction.

Based on the lack of existing in the literature studies for imaging based CAD prediction, our idea was to collect both imaging and non-imaging features for the disease's prediction. In the literature, the association of imaging and biomechanical related features with the CAD progression, is well established. Different studies have confirmed the association of low WSS with the CAD progression, either analyzing IVUS or CTA images. However, most of these studies do not take into account other typical CAD risk factors and the new concept of geometrical based CAD risk factors. Thus, the first objective of our thesis was to collect geometry based CAD risk factors by analyzing CTA images., including features, such as DS, MLA, MLD, PB, CP and NCP volume.

An overall methodology for the 3D coronary reconstruction and plaque characterization using CTA images, was developed and evaluated through this thesis. This overall methodology includes the following steps: (i) CTA images preprocessing, (ii) blooming effect removal, (iii) centerline extraction, (iv) inner wall and outer wall segmentation, (v) CP and NCP segmentation and (vi) 3D surface construction and provides accurately 3D models of the inner wall, outer wall, CP and NCP. Active contour models technique was implemented for the inner, outer wall and CP segmentation, whereas dynamic thresholding technique was used for NCP segmentation. The proposed methodology was validated using manual annotations for inner wall, CP and NCP segmentation and was compared also with IVUS and VH-IVUS modality. DICE for inner wall, CP and NCP was 0.75, 0.7 and 0.62,

respectively, whereas degree of correlation for CP and NCP volume was 0.93 and 0.92, respectively. Regarding the carotid artery reconstruction, manual annotations were used for the technique evaluation and it was found that DICE for the inner wall, outer wall, CP and NCP was 0.83, 0.8, 0.71 and 0.7, respectively [204-206].

The technical innovations of the proposed methodology were the integration of the blooming effect removal in cases of high intensity objects and the integration of vessel centerline extraction, which allows the accurate lumen segmentation in cases of fully occluded vessels. In addition to this, both the detection of CP and the NCP constitute a novelty of the proposed methodology. Based on the concept that the absolute HU range of both CP and NCP plaques is significantly affected by the dose protocol selection and the luminal density [221], our approach is fully adaptive on each acquisition dose protocol methodology for the detection of CP and NCP, as well as their 3D models construction. Overall proposed pipeline was integrated into a user-friendly platform which requires the minimal user interaction and provides 3D models for inner wall, outer wall, CP and NCP and all the geometrical based features by annotating the starting and the ending point of the vessel [194].

Promising results of the proposed 3D reconstruction methodology allow the calculation of geometry based features using the provided by the methodology 3D models. In particular, features such as the DS, the MLA, MLD, PB, CP volume, NCP volume and biomechanical modeling features, such as the $_{smart}$FFR index, the maximum ESS, minimum ESS, mean ESS, the luminal area where ESS was lower than 1 Pa, the maximum LDL concentration and the minimum LDL concentration, were utilized for ML CAD risk prediction models. The overall methodology for these models development was the dataset collection, the preprocessing of the data, the classification and feature selection scheme implementation and finally the model evaluation. Four different approaches have been developed for CAD prediction having as clinical outcomes, the prediction of obstructive CAD (as a 2-class and 3-class problem), the CAD progression and the prediction of PCI placement. Achieved accuracies of the proposed models were 0.81, 0.67, 0.74 and 0.78, respectively, indicating promising results in the CAD prediction. Additionally, high predictability of imaging based features was also confirmed through our first approach (Chapter 4.2), in which accuracy of the model using imaging and non imaging features was 0.81, whereas accuracy using only non imaging data was 0.69 [232-235].

However, except for the prediction of obstructive CAD presence, CAD progression and PCI placement, the prediction of the CAD-related events is also a very important task both for the clinical research area

and for patients' management. However, all the proposed ML models were trained using an existing dataset of 187 participants, in which there were only few CAD-related events. This is a low–medium-risk population, and we have few major CAD events to use for the development of such an event-prediction model. On the other hand, thanks to the advantage of our intermediate CAD risk population, we were able to build a model that can be used as a prognostic decision-support tool by clinicians to properly monitor and manage patients of intermediate CAD risk for the next years after a first imaging is available. In addition to this, number of patients utilized for the development of proposed models was limited, compared to some existing in the literature models.

Another limitation of the proposed methodologies for CAD prediction is that the evaluation of the proposed models was applied with the use of the training dataset and not with the use of an external dataset.

In addition to CAD prediction, in this thesis, prognostic modeling of atherosclerosis progression was developed for the carotid artery disease risk stratification. Based on the latest clinical guidelines, the management of asymptomatic carotid artery disease is based on the CAS, and not take into account the presence of vulnerable atherosclerotic plaques [283]. Thus, in the fifth chapter of our thesis we aim to develop ML models for the diagnostic prediction of CAS presence, high risk plaques presence and symptomatic disease presence. The overall ML based pipeline includes the data collection and curation, the definition of the clinical outcome, the class imbalance handling, the implementation of feature selection and classification techniques, the evaluation of the model and finally the integration of them into a cloud-based platform.

Regarding the detection of CAS presence only demographics, clinical data, risk factors and medication therapy were used as input and an overall 0.78 accuracy was achieved, using as training dataset 881 individuals. The training dataset included only asymptomatic individuals and was also externally validated using 521 asymptomatic individuals. The achieved accuracy of the proposed model using the external dataset was 0.88, whereas the sensitivity and specificity was 0.88 and 0.84, respectively. The main limitation of this external validation dataset was the class imbalance, since the majority of individuals (n=440-84.5%) were at class of CAS presence and only 81 of the individuals were at class of CAS absence.

On the other hand, ML models for the atherosclerotic high risk plaques detection have used as input clinical, demographics, risk factors, biochemical and medication therapy data and their clinical outcome was defined based on the atherosclerotic plaque histology.

A first limitation of these models was the use of imbalance retrospective training datasets and the inclusion of both symptomatic and asymptomatic individuals for the training of the proposed models. In case of symptomatic individuals, the carotid artery disease risk stratification has not clinical significance, since according to the current guidelines, surgery is recommended. In addition to this, the proposed models for diagnostic prediction of plaque vulnerability have not been externally validated.

In this thesis, attempts have been also undertaken to associate carotid artery disease with silent brain lesions. Despite the known relationship between CAS and brain lesions, there are very limited studies that aim to associate CAS with asymptomatic brain lesions. In the last chapter of this PhD thesis, we aim to associate both the carotid condition and patient's comorbidities with the presence of SBIs in the ipsilateral hemisphere. The detection of SBIs has been performed by experts neuroradiologists and typical statistical analysis techniques have been implemented for the detection of the most statistically significant features.

Through this thesis, analyzing brain MRI of 211 patients, it was indicated that high mean values of ICA stenosis and St Mary's ratio are associated with the presence of cortical, small subcortical and active infarcts. Additionally, it was found that CAS in the ipsilateral carotid artery is considered as a statistically significant risk factor for the presence of cortical infarcts, whereas positive risk factors for the presence of silent brain lesions was the previous MI, the alcohol abuse and the previous CABG/PCI placement. Moreover, higher mean values of SBP, triglycerides, creatinine and HbA1c are also associated with silent brain lesions presence.

Undoubtedly, as we enter the age of precision medicine, risk assessment and prediction models are considered more notable. In this thesis, our principal aim is to present innovative ML based models for the CAD and carotid artery disease prediction, taking into account both imaging and non-imaging based features. The capability of ML models, combined with a detailed input set of parameters and a balanced dataset of patients may provide novel and promising stratification approaches, contributing to the clinical and research CV area.

7.2 Future work

Regarding the methodology for the 3D arteries reconstruction and atherosclerotic plaques detection, a future step could be the implementation of deep learning techniques in CTA images to improve segmentation performance, using more CTA cases both for training and testing. In addition to this, another future important analysis is the use of the proposed methodology to compare the CTA performance in the field of CAD diagnosis with PET and SPECT.

In spite of the high predictability of proposed studies for the CAD risk prediction, a future step for their deployment could be the integration of new features, concerning molecular systemic variables, inflammatory and monocyte markers, the lipid profile, exposome as well as mRNA sequencing. Therefore, a more detailed input space and a larger dataset of patients ensure a more effective multimodal prediction scheme and potentially a refined formulation of the classification problem. Additionally, development of ML model for CAD related events prediction is also a future step for the CAD risk prediction problem, using as input both imaging and non imaging data. Collection of external validation datasets for the CAD risk prediction developed models is also our next step for their broadly evaluation and to check how accurately they generally perform.

The work performed in this thesis is not finished in terms of carotid artery disease prediction. More specifically, our first aim is to use more external individuals, to test the diagnostic presence of CAS, especially of healthy (Class 0) individuals and to collect data for external validation of the diagnostic high risk plaques presence. This task is challenging since high risk plaques are difficult to be defined and are basically based either on histology of plaque features or on image related features, such as large, echolucent plaques, intraplaque haemorrhage and presence of lipid-rich necrotic core. Another future work for the carotid artery disease prediction is to develop a model for the prediction of CVD related events, such as stroke, TIA, MI, using imaging and non imaging features.

Final, regarding the brain MRI analysis, a future step is the analysis of the brain MRI images at a follow up time step, in order to evaluate the progression of the lesions. In addition to this, except of the implementation of statistical analysis for the data analysis, another step is to perform feature selection techniques for the detection of significant features for the presence of brain lesions.

References

[1] J. Moini, Ed. *Anatomy and Physiology of the Cardiovascular System* (Phlebotomy: Principles and Practice. 2013, pp. 35-52.

[2] J. Sobotta, J. P. McMurrich, and W. H. Thomas, *Atlas of human anatomy*. Philadelphia and London,: W. B. Saunders company, 1906.

[3] H. Gray and W. H. Lewis, *Anatomy of the human body*. Lea & Febiger Philadelphia, 1942.

[4] A. C. Burris and M. Shoukfeh, "Basic Coronary Artery Anatomy and Histology," in *Interventional Cardiology Imaging: An Essential Guide*, A. E. Abbas Ed. London: Springer London, 2015, pp. 1-12.

[5] V. Fuster, *Hurst's the heart*. McGraw-Hill, Medical Pub. Division, 2004.

[6] H. Gray and C. M. Goss, "Anatomy of the human body," *American Journal of Physical Medicine & Rehabilitation,* vol. 53, no. 6, p. 293, 1974.

[7] M. Borysenko, & Beringer, T, "Functional histology," *Boston: Little, Brown and Company,* pp. 195-208, 1984.

[8] G. Millonig, H. Niederegger, and G. Wick, "Analysis of the cellular composition of the arterial intima with modified en face techniques," *Laboratory investigation; a journal of technical methods and pathology,* vol. 81, no. 4, pp. 639-41, Apr 2001. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/11304584.

[9] P. B. Dobrin, "Distribution of lamellar deformations - Implications for properties of the arterial media," (in English), *Hypertension,* vol. 33, no. 3, pp. 806-810, Mar 1999.

[10] J. M. Pearce, "Henry Gray's Anatomy," *Clinical anatomy,* vol. 22, no. 3, pp. 291-5, Apr 2009.

[11] J. D Humphrey, *Mechanics of the Arterial Wall: Review and Directions*. 1995, pp. 1-162.

[12] K. L. Moore and A. F. Dalley, *Clinically oriented anatomy*. Wolters kluwer india Pvt Ltd, 2018.

[13] G. G. Wind and R. J. Valentine, *Anatomic exposures in vascular surgery*. Lippincott Williams & Wilkins, 2013.

[14] H. Ashrafian, "Anatomically specific clinical examination of the carotid arterial tree," *Anatomical science international,* vol. 82, pp. 16-23, 2007.

[15] R. B. Singh, S. A. Mengi, Y.-J. Xu, A. S. Arneja, and N. S. Dhalla, "Pathogenesis of atherosclerosis: A multifactorial process," *Experimental & Clinical Cardiology,* vol. 7, no. 1, pp. 40-53, Spring 2002. [Online]. Available: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2716189/.

[16] A. Tonkin, *Atherosclerosis and heart disease*. CRC Press, 2003.

[17] P. C. Choy, Y. L. Siow, D. Mymin, and K. O, "Lipids and atherosclerosis," *Biochemistry and Cell Biology,* vol. 82, no. 1, pp. 212-224, 2004/02/01 2004.

[18] M. Aziz and K. Yadav, *Pathogenesis of Atherosclerosis A Review*. 2016.

[19] P. Libby and P. Theroux, "Pathophysiology of coronary artery disease," *Circulation,* vol. 111, no. 25, pp. 3481-3488, 2005.

[20] P. Libby, P. M. Ridker, and G. K. Hansson, "Progress and challenges in translating the biology of atherosclerosis," *Nature,* vol. 473, no. 7347, p. 317, 2011.

[21] N. Townsend, L. Wilson, P. Bhatnagar, K. Wickramasinghe, M. Rayner, and M. Nichols, "Cardiovascular disease in Europe: epidemiological update 2016," *European Heart Journal,* vol. 37, no. 42, pp. 3232-3245, 2016, doi: 10.1093/eurheartj/ehw334.

[22] T. He *et al.*, "Diagnostic models of the pre-test probability of stable coronary artery disease: A systematic review," (in eng), *Clinics (Sao Paulo, Brazil),* vol. 72, no. 3, pp. 188-196, 2017, doi: 10.6061/clinics/2017(03)10.

[23] S. D. Fihn *et al.*, "2012 ACCF/AHA/ACP/AATS/PCNA/SCAI/STS guideline for the diagnosis and management of patients with stable ischemic heart disease: executive summary: a report of the American College of Cardiology Foundation/American Heart Association task force on practice guidelines, and the American College of Physicians, American Association for Thoracic Surgery, Preventive Cardiovascular Nurses Association, Society for Cardiovascular Angiography and Interventions, and Society of Thoracic Surgeons," (in eng), *Circulation,* vol. 126, no. 25, pp. 3097-137, Dec 18 2012, doi: 10.1161/CIR.0b013e3182776f83.

[24] G. Montalescot *et al.*, "2013 ESC guidelines on the management of stable coronary artery disease: the Task Force on the management of stable coronary artery disease of the European Society of Cardiology," (in eng), *Eur Heart J,* vol. 34, no. 38, pp. 2949-3003, Oct 2013, doi: 10.1093/eurheartj/eht296.

[25] G. A. Diamond and J. S. Forrester, "Analysis of Probability as an Aid in the Clinical Diagnosis of Coronary-Artery Disease," *New England Journal of Medicine,* vol. 300, no. 24, pp. 1350-1358, 1979, doi: 10.1056/nejm197906143002402.

[26] T. S. S. Genders *et al.*, "Prediction model to estimate presence of coronary artery disease: retrospective pooled analysis of existing cohorts," *BMJ : British Medical Journal,* vol. 344, p. e3485, 2012, doi: 10.1136/bmj.e3485.

[27] T. C. Consortium *et al.*, "A clinical prediction rule for the diagnosis of coronary artery disease: validation, updating, and extension," *European Heart Journal,* vol. 32, no. 11, pp. 1316-1330, 2011, doi: 10.1093/eurheartj/ehr014.

[28] E. Messas *et al.*, "Management of carotid stenosis for primary and secondary prevention of stroke: state-of-the-art 2020: a critical review," *European Heart Journal Supplements,* vol. 22, no. Supplement_M, pp. M35-M42, 2020.

[29] E. V. Ratchford and N. S. Evans, "Carotid artery disease," *Vascular medicine,* vol. 19, no. 6, pp. 512-515, 2014.

[30] V. Aboyans *et al.*, "2017 ESC guidelines on the diagnosis and treatment of peripheral arterial diseases, in collaboration with the European Society for Vascular Surgery (ESVS)," *Kardiologia Polska (Polish Heart Journal),* vol. 75, no. 11, pp. 1065-1160, 2017.

[31] P. W. Wilson and P. S. Douglas, "Epidemiology of coronary heart disease," *UpToDate, Basow, DS (Ed), UpToDate, Waltham, MA,* 2010.

[32] J. M. Tarkin *et al.*, "Imaging atherosclerosis," *Circulation research,* vol. 118, no. 4, pp. 750-769, 2016.

[33] A. DE ROOS, L. J. KROFT, M. J. SCHALIJ, and J. REIBER, "Imaging of atherosclerosis: invasive and noninvasive techniques," *Hellenic J Cardiol,* vol. 50, pp. 245-263, 2009.

[34] L. S. Athanasiou, D. I. Fotiadis, and L. K. Michalis, *Atherosclerotic Plaque Characterization Methods Based on Coronary Imaging*. Academic Press, 2017.

[35] S. K. Mehta, J. R. McCrary, A. D. Frutkin, W. J. Dolla, and S. P. Marso, "Intravascular ultrasound radiofrequency analysis of coronary atherosclerosis: an emerging technology for the assessment of vulnerable plaque," *European heart journal,* vol. 28, no. 11, pp. 1283-1288, 2007.

[36] S. E. Nissen and P. Yock, "Intravascular ultrasound: novel pathophysiological insights and current clinical applications," *Circulation,* vol. 103, no. 4, pp. 604-616, 2001.

[37] P. N. Naka KK, "Intravascular Image Interpretation," *Intravascular Imaging: Current Applications and Research Developments: Current Applications and Research Developments,* 2011.

[38] L. Athanasiou, N. Bruining, F. Prati, and D. Koutsouris, "Optical coherence tomography: basic principles of image acquisition," *Intravascular Imaging: Current Applications and Research Developments,* pp. 180-194, 2011.

[39] T. Kume *et al.*, "Assessment of coronary arterial plaque by optical coherence tomography," *American Journal of Cardiology,* vol. 97, no. 8, pp. 1172-1175, 2006.

[40] S. Brugaletta *et al.*, "NIRS and IVUS for characterization of atherosclerosis in patients undergoing coronary angiography," *JACC: Cardiovascular Imaging,* vol. 4, no. 6, pp. 647-655, 2011.

[41] V. D. Tsakanikas, L. K. Michalis, D. I. Fotiadis, K. K. Naka, and C. V. Bourantas, *Intravascular imaging: current applications and research developments*. IGI Global, 2011.

[42] Z. Sun, G. H. Choo, and K. H. Ng, "Coronary CT angiography: current status and continuing challenges," *The British Journal of Radiology,* vol. 85, no. 1013, pp. 495-510, 10/29/received.

[43] E. Maffei *et al.*, "Plaque imaging with CT coronary angiography: effect of intra-vascular attenuation on plaque type classification," *World journal of radiology,* vol. 4, no. 6, p. 265, 2012.

[44] A. W. Leber *et al.*, "Accuracy of multidetector spiral computed tomography in identifying and differentiating the composition of coronary atherosclerotic plaques," *Journal of the American College of Cardiology,* vol. 43, no. 7, pp. 1241-1247, 2004/04/07/, doi: http://dx.doi.org/10.1016/j.jacc.2003.10.059.

[45] I. Springer and M. Dewey, "Comparison of multislice computed tomography with intravascular ultrasound for detection and characterization of coronary artery plaques: a systematic review," *European journal of radiology,* vol. 71, no. 2, pp. 275-282, 2009.

[46] L. Husmann *et al.*, "Feasibility of low-dose coronary CT angiography: first experience with prospective ECG-gating," *European heart journal,* vol. 29, no. 2, pp. 191-197, 2007.

[47] A. Saxena, E. Y. K. Ng, and S. T. Lim, "Imaging modalities to diagnose carotid artery stenosis: progress and prospect," *Biomedical engineering online,* vol. 18, no. 1, p. 66, 2019.

[48] G. B. Anderson, R. Ashforth, D. E. Steinke, R. Ferdinandy, and J. M. Findlay, "CT angiography for the detection and characterization of carotid artery bifurcation disease," *Stroke,* vol. 31, no. 9, pp. 2168-2174, 2000.

[49] J. Eckert, M. Schmidt, A. Magedanz, T. Voigtländer, and A. Schmermund, "Coronary CT Angiography in Managing Atherosclerosis," *International Journal of Molecular Sciences,* vol. 16, no. 2, p. 3740, 2015. [Online]. Available: http://www.mdpi.com/1422-0067/16/2/3740.

[50] J. K. Min *et al.*, "Prognostic value of multidetector coronary computed tomographic angiography for prediction of all-cause mortality," *Journal of the American College of Cardiology,* vol. 50, no. 12, pp. 1161-1170, 2007.

[51] Z. Chen and S. Molloi, "Automatic 3D vascular tree construction in CT angiography," *Computerized Medical Imaging and Graphics,* vol. 27, no. 6, pp. 469-479, 2003.

[52] C. Florin, N. Paragios, and J. Williams, "Particle filters, a quasi-monte carlo solution for segmentation of coronaries," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2005: Springer, pp. 246-253.

[53] A. Szymczak, A. Stillman, A. Tannenbaum, and K. Mischaikow, "Coronary vessel trees from 3D imagery: a topological approach," *Medical image analysis,* vol. 10, no. 4, pp. 548-559, 2006.

[54] J. Lee, P. Beighley, E. Ritman, and N. Smith, "Automatic segmentation of 3D micro-CT coronary vascular images," *Medical image analysis,* vol. 11, no. 6, pp. 630-647, 2007.

[55] S. D. Olabarriaga, M. Breeuwer, and W. J. Niessen, "Minimum cost path algorithm for coronary artery central axis tracking in CT images," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2003: Springer, pp. 687-694.

[56] C. Wang and Ö. Smedby, "Coronary artery segmentation and skeletonization based on competing fuzzy connectedness tree," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2007: Springer, pp. 311-318.

[57] B. Bouraoui, C. Ronse, J. Baruthio, N. Passat, and P. Germain, "3D segmentation of coronary arteries based on advanced mathematical morphology techniques," *Computerized medical imaging and graphics,* vol. 34, no. 5, pp. 377-387, 2010.

[58] R. Shahzad *et al.*, "Automatic segmentation, detection and quantification of coronary artery stenoses on CTA," *The international journal of cardiovascular imaging,* vol. 29, no. 8, pp. 1847-1859, 2013.

[59] S.-T. Chen *et al.*, "DWT-Based Segmentation Method for Coronary Arteries," *Journal of Medical Systems,* journal article vol. 38, no. 6, p. 55, May 09 2014, doi: 10.1007/s10916-014-0055-8.

[60] Y. Yang, A. R. Tannenbaum, D. P. Giddens, and A. Stillman, "Automatic segmentation of coronary arteries using bayesian driven implicit surfaces," 2007: Georgia Institute of Technology.

[61] Y. Kitamura, Y. Li, W. Ito, and H. Ishikawa, "Coronary lumen and plaque segmentation from CTA using higher-order shape prior," in *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2014: Springer, pp. 339-347.

[62] S.-T. Chen *et al.*, "Coronary arteries segmentation based on the 3D discrete wavelet transform and 3D neutrosophic transform," *BioMed research international,* vol. 2015, 2015.

[63] Y. Cheng, X. Hu, J. Wang, Y. Wang, and S. Tamura, "Accurate vessel segmentation with constrained B-snake," *IEEE Trans. Image Processing*, vol. 24, no. 8, pp. 2440-2455, 2015.

[64] L. Athanasiou *et al.*, "Three-dimensional reconstruction of coronary arteries and plaque morphology using CT angiography – comparison and registration with IVUS," *BMC Medical Imaging*, vol. 16, p. 9, accepted 2016, doi: 10.1186/s12880-016-0111-6.

[65] H. Zhu, S. Song, L. Xu, A. Song, and B. Yang, "Segmentation of coronary arteries images using spatio-temporal feature fusion network with combo loss," *Cardiovascular Engineering and Technology*, pp. 1-12, 2022.

[66] P. Mirunalini, C. Aravindan, A. T. Nambi, S. Poorvaja, and V. P. Priya, "Segmentation of Coronary Arteries from CTA axial slices using Deep Learning techniques," in *TENCON 2019-2019 IEEE Region 10 Conference (TENCON)*, 2019: IEEE, pp. 2074-2080.

[67] Y. Arad, L. A. Spadaro, K. Goodman, D. Newstein, and A. D. Guerci, "Prediction of coronary events with electron beam computed tomography," *Journal of the American College of Cardiology*, vol. 36, no. 4, pp. 1253-1260, 2000.

[68] R. C. Detrano and E. V. Romero, "Coronary artery calcium screening: where do we go from here?," *Italian heart journal: official journal of the Italian Federation of Cardiology*, vol. 5, no. 6, p. 421, 2004.

[69] S. Schroeder *et al.*, "Noninvasive detection and evaluation of atherosclerotic coronary plaques with multislice computed tomography," *Journal of the American College of Cardiology*, vol. 37, no. 5, pp. 1430-1435, 2001.

[70] J. Hur *et al.*, "Quantification and characterization of obstructive coronary plaques using 64-slice computed tomography: a comparison with intravascular ultrasound," *Journal of computer assisted tomography*, vol. 33, no. 2, pp. 186-192, 2009.

[71] S.-L. Papadopoulou *et al., Detection and quantification of coronary atherosclerotic plaque by 64-slice multidetector CT: a systematic head-to-head comparison with Intravascular Ultrasound*. 2011, pp. 163-70.

[72] A. W. Leber *et al.*, "Accuracy of 64-Slice Computed Tomography to Classify and Quantify Plaque Volumes in the Proximal Coronary System: A Comparative Study Using Intravascular Ultrasound," *Journal of the American College of Cardiology*, vol. 47, no. 3, pp. 672-677, 2006/02/07/ 2006, doi: https://doi.org/10.1016/j.jacc.2005.10.058.

[73] S. Voros *et al.*, "Prospective validation of standardized, 3-dimensional, quantitative coronary computed tomographic plaque measurements using radiofrequency backscatter intravascular ultrasound as reference standard in intermediate coronary arterial lesions: results from the ATLANTA (assessment of tissue characteristics, lesion morphology, and hemodynamics by angiography with fractional flow reserve, intravascular ultrasound and virtual histology, and noninvasive computed tomography in atherosclerotic plaques) I study," *JACC: Cardiovascular Interventions,* vol. 4, no. 2, pp. 198-208, 2011.

[74] K. Akram, S. Rinehart, and S. Voros, "Coronary arterial atherosclerotic plaque imaging by contrast-enhanced computed tomography: fantasy or reality?," *Journal of Nuclear Cardiology,* vol. 15, no. 6, pp. 818-829, 2008.

[75] S. Rinehart, G. Vazquez, Z. Qian, and S. Voros, "Coronary plaque imaging with multi-slice computed tomographic angiography and intravascular ultrasound: a close look inside and out," *The Journal of invasive cardiology,* vol. 21, no. 7, pp. 367-372, 2009.

[76] D. Dey, T. Schepis, M. Marwan, P. J. Slomka, D. S. Berman, and S. Achenbach, "Automated Three-dimensional Quantification of Noncalcified Coronary Plaque from Coronary CT Angiography: Comparison with Intravascular US," *Radiology,* vol. 257, no. 2, pp. 516-522, 2010, doi: 10.1148/radiol.10100681.

[77] D. Dey *et al.*, "Automated 3-dimensional quantification of noncalcified and calcified coronary plaque from coronary CT angiography," *Journal of cardiovascular computed tomography,* vol. 3, no. 6, pp. 372-382, 2009.

[78] H. Brodoefel *et al.*, "Accuracy of dual-source CT in the characterisation of non-calcified plaque: use of a colour-coded analysis compared with virtual histology intravascular ultrasound," *The British Journal of Radiology,* vol. 82, no. 982, pp. 805-812, 03/30 2009, doi: 10.1259/bjr/35768497.

[79] M. J. Boogers *et al.*, "Automated quantification of coronary plaque with computed tomography: comparison with intravascular ultrasound using a dedicated registration algorithm for fusion-based quantification," *European heart journal,* vol. 33, no. 8, pp. 1007-1016, 2012.

[80] M. A. de Graaf *et al.*, "Automatic quantification and characterization of coronary atherosclerosis with computed tomography coronary angiography: cross-correlation with intravascular ultrasound virtual histology," *The international journal of cardiovascular imaging,* vol. 29, no. 5, pp. 1177-1190, 2013.

[81] H. Brodoefel *et al.*, "Characterization of coronary atherosclerosis by dual-source computed tomography and HU-based color mapping: a pilot study," *European radiology,* vol. 18, no. 11, pp. 2466-2474, 2008.

[82] S. Voros *et al.*, "Coronary Atherosclerosis Imaging by Coronary CT Angiography: Current Status, Correlation With Intravascular Interrogation and Meta-Analysis," *JACC: Cardiovascular Imaging,* vol. 4, no. 5, pp. 537-548, 2011/05/01/ 2011, doi: https://doi.org/10.1016/j.jcmg.2011.03.006.

[83] M. Utsunomiya, H. Hara, M. Moroi, K. Sugi, and M. Nakamura, "Relationship between tissue characterization with 40MHz intravascular ultrasound imaging and 64-slice computed tomography," *Journal of Cardiology,* vol. 57, no. 3, pp. 297-302, 2011/05/01/ 2011, doi: https://doi.org/10.1016/j.jjcc.2011.01.016.

[84] M. M. Jawaid, A. Riaz, R. Rajani, C. C. Reyes-Aldasoro, and G. Slabaugh, "Framework for detection and localization of coronary non-calcified plaques in cardiac CTA using mean radial profiles," *Computers in biology and medicine,* vol. 89, pp. 84-95, 2017.

[85] A. Huibers *et al.*, "Non-invasive carotid artery imaging to identify the vulnerable plaque: current status and future goals," *European Journal of Vascular and Endovascular Surgery,* vol. 50, no. 5, pp. 563-572, 2015.

[86] M. Sanderse, H. A. Marquering, E. A. Hendriks, A. van der Lugt, and J. H. C. Reiber, "Automatic Initialization Algorithm for Carotid Artery Segmentation in CTA Images," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2005*, Berlin, Heidelberg, J. S. Duncan and G. Gerig, Eds., 2005// 2005: Springer Berlin Heidelberg, pp. 846-853.

[87] E. F. S. van Velsen *et al.*, "Evaluation of an improved technique for lumen path definition and lumen segmentation of atherosclerotic vessels in CT angiography," *European Radiology,* journal article vol. 17, no. 7, pp. 1738-1745, July 01 2007, doi: 10.1007/s00330-006-0469-x.

[88] R. Manniesing and W. J. Niessen, "Automatic segmentation of the internal carotid arteries through the skull base," in *Medical Imaging 2007: Image Processing*, 2007, vol. 6512: International Society for Optics and Photonics, p. 65120I.

[89] O. Cuisenaire, S. Virmani, M. E. Olszewski, and R. Ardon, "Fully automated segmentation of carotid and vertebral arteries from contrast enhanced CTA," in *Medical Imaging 2008: Image Processing*, 2008, vol. 6914: International Society for Optics and Photonics, p. 69143R.

[90] R. Manniesing *et al.*, "Robust CTA lumen segmentation of the atherosclerotic carotid artery bifurcation in a large patient population," *Medical image analysis*, vol. 14, no. 6, pp. 759-769, 2010.

[91] D. Vukadinovic, T. van Walsum, R. Manniesing, S. Rozie, A. van der Lugt, and W. J. Niessen, "Region based level set segmentation of the outer wall of the carotid bifurcation in CTA," in *Medical Imaging 2011: Image Processing*, 2011, vol. 7962: International Society for Optics and Photonics, p. 79623P.

[92] R. Manniesing, B. K. Velthuis, M. S. van Leeuwen, I. Van Der Schaaf, P. Van Laar, and W. J. Niessen, "Level set based cerebral vasculature segmentation and diameter quantification in CT angiography," *Medical image analysis*, vol. 10, no. 2, pp. 200-214, 2006.

[93] M. Freiman *et al.*, "Carotid vasculature modeling from patient CT angiography studies for interventional procedures simulation," *International Journal of Computer Assisted Radiology and Surgery*, vol. 7, no. 5, pp. 799-812, 2012/09/01 2012, doi: 10.1007/s11548-012-0673-x.

[94] H. Tang, T. van Walsum, R. Hameeteman, R. Shahzad, L. J. van Vliet, and W. J. Niessen, "Lumen segmentation and stenosis quantification of atherosclerotic carotid arteries in CTA utilizing a centerline intensity prior," *Medical physics*, vol. 40, no. 5, 2013.

[95] F. L. C. dos Santos, A. Joutsen, M. Terada, J. Salenius, and H. Eskola, "A semi-automatic segmentation method for the structural analysis of carotid atherosclerotic plaques by computed tomography angiography," *Journal of atherosclerosis and thrombosis*, vol. 21, no. 9, pp. 930-940, 2014.

[96] H. Hemmati, A. Kamli-Asl, A. Talebpour, and S. Shirani, "Semi-automatic 3D segmentation of carotid lumen in contrast-enhanced computed tomography angiography images," *Physica Medica*, vol. 31, no. 8, pp. 1098-1104, 2015.

[97] F. Bozkurt, C. Köse, and A. Sarı, "An inverse approach for automatic segmentation of carotid and vertebral arteries in CTA," *Expert Systems with Applications*, vol. 93, pp. 358-375, 2018.

[98] I. Guha, N. Das, P. Rakshit, M. Nasipuri, P. K. Saha, and S. Basu, "A semiautomatic approach for segmentation of carotid vasculature from patients' CTA images," *Innovations in Systems and Software Engineering*, vol. 13, no. 4, pp. 243-250, 2017.

[99] L. Saba, G. Argiolas, P. Siotto, and M. Piga, "Carotid artery plaque characterization using CT multienergy imaging," *American Journal of Neuroradiology*, vol. 34, no. 4, pp. 855-859, 2013.

[100] E. I. Georga *et al.*, "Artificial Intelligence and Data Mining Methods for Cardiovascular Risk Prediction," in *Cardiovascular Computing — Methodologies and Clinical Applications*, S. Golemati and K. S. Nikita Eds., (Series in BioEngineering: Springer Singapore, 2019, ch. Artificial Intelligence and Data Mining Methods for Cardiovascular Risk Prediction.

[101] I. Kurt, M. Ture, and A. T. Kurum, "Comparing performances of logistic regression, classification and regression tree, and neural networks for predicting coronary artery disease," *Expert Syst Appl*, vol. 34, no. 1, pp. 366-374, 1// 2008, doi: https://doi.org/10.1016/j.eswa.2006.09.004.

[102] L. Verma, S. Srivastava, and P. C. Negi, "A Hybrid Data Mining Model to Predict Coronary Artery Disease Cases Using Non-Invasive Clinical Data," (in eng), *J Med Syst*, vol. 40, no. 7, p. 178, Jul 2016, doi: 10.1007/s10916-016-0536-z.

[103] R. Alizadehsani *et al.*, "A data mining approach for diagnosis of coronary artery disease," *Computer methods and programs in biomedicine*, vol. 111, no. 1, pp. 52-61, Jul 2013, doi: 10.1016/j.cmpb.2013.03.004.

[104] R. Alizadehsani *et al.*, "Coronary artery disease detection using computational intelligence methods," *Knowledge-Based Systems*, vol. 109, pp. 187-197, 2016.

[105] J. Nahar, T. Imam, K. S. Tickle, and Y.-P. P. Chen, "Computational intelligence for heart disease diagnosis: A medical knowledge driven approach," *Expert Syst Appl*, vol. 40, no. 1, pp. 96-104, 2013/01/01/ 2013, doi: https://doi.org/10.1016/j.eswa.2012.07.032.

[106] E. M. Karabulut and T. Ibrikci, "Effective diagnosis of coronary artery disease using the rotation forest ensemble method," (in eng), *J Med Syst*, vol. 36, no. 5, pp. 3011-8, Oct 2012, doi: 10.1007/s10916-011-9778-y.

[107] M. G. Tsipouras *et al.*, "Automated diagnosis of coronary artery disease based on data mining and fuzzy modeling," *IEEE transactions on information technology in biomedicine : a publication of the IEEE Engineering in Medicine and Biology Society*, vol. 12, no. 4, pp. 447-58, Jul 2008, doi: 10.1109/TITB.2007.907985.

[108] M. R. Elashoff *et al.*, "Development of a blood-based gene expression algorithm for assessment of obstructive coronary artery disease in non-diabetic patients," *BMC Medical Genomics*, journal article vol. 4, no. 1, p. 26, March 28 2011, doi: 10.1186/1755-8794-4-26.

[109] M. V. Dogan, I. M. Grumbach, J. J. Michaelson, and R. A. Philibert, "Integrated genetic and epigenetic prediction of coronary heart disease in the Framingham Heart Study," (in eng), *Plos One,* vol. 13, no. 1, p. e0190549, 2018, doi: 10.1371/journal.pone.0190549.

[110] S. Voros, M. R. Elashoff, J. A. Wingrove, M. J. Budoff, G. S. Thomas, and S. Rosenberg, "A peripheral blood gene expression score is associated with atherosclerotic Plaque Burden and Stenosis by cardiovascular CT-angiography: results from the PREDICT and COMPASS studies," (in eng), *Atherosclerosis,* vol. 233, no. 1, pp. 284-90, Mar 2014, doi: 10.1016/j.atherosclerosis.2013.12.045.

[111] H. Zhu, Z. Ding, R. N. Piana, T. R. Gehrig, and M. H. Friedman, "Cataloguing the geometry of the human coronary arteries: a potential tool for predicting risk of coronary artery disease," *International journal of cardiology,* vol. 135, no. 1, pp. 43-52, 07/01 2009, doi: 10.1016/j.ijcard.2008.03.087.

[112] P. H. Stone *et al.,* "Effect of endothelial shear stress on the progression of coronary artery disease, vascular remodeling, and in-stent restenosis in humans: in vivo 6-month follow-up study," *Circulation,* vol. 108, no. 4, pp. 438-444, 2003.

[113] P. H. Stone *et al.,* "Regions of low endothelial shear stress are the sites where coronary plaque progresses and vascular remodelling occurs in humans: an in vivo serial study," *European heart journal,* vol. 28, no. 6, pp. 705-710, 2007.

[114] H. Samady *et al.,* "Coronary artery wall shear stress is associated with progression and transformation of atherosclerotic plaque and arterial remodeling in patients with coronary artery disease," *Circulation,* p. CIRCULATIONAHA. 111.021824, 2011.

[115] P. H. Stone *et al.,* "Prediction of progression of coronary artery disease and clinical outcomes using vascular profiling of endothelial shear stress and arterial plaque characteristics: the PREDICTION Study," *Circulation,* p. CIRCULATIONAHA. 112.096438, 2012.

[116] H. Hetterich *et al.,* "Coronary computed tomography angiography based assessment of endothelial shear stress and its association with atherosclerotic plaque distribution in-vivo," *PLoS One,* vol. 10, no. 1, p. e0115408, 2015.

[117] M. Papafaklis, M. Mavrogiannis, and P. Stone, "Identifying the progression of coronary artery disease: prediction of cardiac events," *Continuing Cardiology Education,* vol. 2, no. 2, pp. 105-114, 2016.

[118] X. Liu *et al.,* "Prediction of coronary plaque progression using biomechanical factors and vascular characteristics based on computed tomography angiography," *Computer Assisted Surgery,* vol. 22, no. sup1, pp. 286-294, 2017.

[119] M. Kolossváry *et al.,* "Radiomic features are superior to conventional quantitative computed tomographic metrics to identify coronary plaques with napkin-ring sign," *Circulation: Cardiovascular Imaging,* vol. 10, no. 12, p. e006843, 2017.

[120] H. J. Aerts *et al.,* "Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach," *Nature communications,* vol. 5, no. 1, pp. 1-9, 2014.

[121] M. Mannil, J. von Spiczak, R. Manka, and H. Alkadhi, "Texture analysis and machine learning for detecting myocardial infarction in noncontrast low-dose computed tomography: unveiling the invisible," *Investigative radiology,* vol. 53, no. 6, pp. 338-343, 2018.

[122] M. Motwani *et al.,* "Machine learning for prediction of all-cause mortality in patients with suspected coronary artery disease: a 5-year multicentre prospective registry analysis," *European heart journal,* vol. 38, no. 7, pp. 500-507, 2017.

[123] A. R. van Rosendael *et al.,* "Maximization of the usage of coronary CTA derived plaque information using a machine learning based algorithm to improve risk stratification; insights from the CONFIRM registry," *Journal of cardiovascular computed tomography,* vol. 12, no. 3, pp. 204-209, 2018.

[124] A. Jamthikar *et al.,* "A low-cost machine learning-based cardiovascular/stroke risk assessment system: integration of conventional factors with image phenotypes," *Cardiovascular Diagnosis and Therapy,* vol. 9, no. 5, pp. 420-430, 2019.

[125] L. Verde and G. De Pietro, "A Machine Learning Approach for Carotid Diseases using Heart Rate Variability Features," in *HEALTHINF*, 2018, pp. 658-664.

[126] G. Greco *et al.,* "A model for predicting the risk of carotid artery disease," *Annals of surgery,* vol. 257, no. 6, pp. 1168-1173, 2013.

[127] M. de Weerd *et al.,* "Prediction of asymptomatic carotid artery stenosis in the general population: identification of high-risk groups," *Stroke,* vol. 45, no. 8, pp. 2366-2371, 2014.

[128] N. N. Khanna *et al.,* "Nonlinear model for the carotid artery disease 10-year risk prediction by fusing conventional cardiovascular factors to carotid ultrasound image phenotypes: A Japanese diabetes cohort study," *Echocardiography,* vol. 36, no. 2, pp. 345-361, 2019.

[129] T. Araki *et al.*, "Stroke risk stratification and its validation using ultrasonic echolucent carotid wall plaque morphology: a machine learning paradigm," *Computers in biology and medicine*, vol. 80, pp. 77-96, 2017.

[130] H.-w. Li *et al.*, "Association between ADAMTS7 polymorphism and carotid artery plaque vulnerability," *Medicine*, vol. 98, no. 43, p. e17438, 2019.

[131] S. A. Nasser, E. A. Afify, F. Kobeissy, B. Hamam, A. H Eid, and M. M. El-Mas, "Inflammatory basis of atherosclerosis: Modulation by sex hormones," *Current Pharmaceutical Design*, 2021.

[132] P. Puz, A. Lasek-Bal, D. Ziaja, Z. Kazibutowska, and K. Ziaja, "Inflammatory markers in patients with internal carotid artery stenosis," *Archives of medical science: AMS*, vol. 9, no. 2, p. 254, 2013.

[133] E. Debing, E. Peeters, C. Demanet, M. De Waele, and P. Van den Brande, "Markers of inflammation in patients with symptomatic and asymptomatic carotid artery stenosis: a case-control study," *Vascular and endovascular surgery*, vol. 42, no. 2, pp. 122-127, 2008.

[134] C. S. Horn *et al.*, "High-sensitivity C-reactive protein at different stages of atherosclerosis: results of the INVADE study," *Journal of neurology*, vol. 256, no. 5, pp. 783-791, 2009.

[135] H. Yamagami *et al.*, "Higher levels of interleukin-6 are associated with lower echogenicity of carotid artery plaques," *Stroke*, vol. 35, no. 3, pp. 677-681, 2004.

[136] A. Shindo *et al.*, "Inflammatory biomarkers in atherosclerosis: pentraxin 3 can become a novel marker of plaque vulnerability," *PloS one*, vol. 9, no. 6, p. e100045, 2014.

[137] Y. Zhou, W. Han, D. Gong, C. Man, and Y. Fan, "Hs-CRP in stroke: a meta-analysis," *Clinica chimica acta*, vol. 453, pp. 21-27, 2016.

[138] Y. Yang *et al.*, "Serum lipoprotein-associated phospholipase A2 predicts the formation of carotid artery plaque and its vulnerability in anterior circulation cerebral infarction," *Clinical neurology and neurosurgery*, vol. 160, pp. 40-45, 2017.

[139] M. Knoflach *et al.*, "Pentraxin-3 as a marker of advanced atherosclerosis results from the Bruneck, ARMY and ARFY Studies," *PloS one*, vol. 7, no. 2, p. e31474, 2012.

[140] L. Yi *et al.*, "Pentraxin 3, TNF-α, and LDL-C are associated with carotid artery stenosis in patients with ischemic stroke," *Frontiers in neurology*, vol. 10, p. 1365, 2020.

[141] J. Andersson *et al.*, "The carotid artery plaque size and echogenicity are related to different cardiovascular risk factors in the elderly: the Prospective Investigation of the Vasculature in Uppsala Seniors (PIVUS) study," *Lipids*, vol. 44, no. 5, p. 397, 2009.

[142] A. Abbas *et al.,* "Matrix metalloproteinase 7 is associated with symptomatic lesions and adverse events in patients with carotid atherosclerosis," *PloS one,* vol. 9, no. 1, p. e84935, 2014.

[143] J. Schneiderman *et al.,* "Leptin locally synthesized in carotid atherosclerotic plaques could be associated with lesion instability and cerebral emboli," *Journal of the American Heart Association,* vol. 1, no. 5, p. e001727, 2012.

[144] I. Mitroulis, V. I. Alexaki, I. Kourtzelis, A. Ziogas, G. Hajishengallis, and T. Chavakis, "Leukocyte integrins: role in leukocyte recruitment and as therapeutic targets in inflammatory disease," *Pharmacology & therapeutics,* vol. 147, pp. 123-135, 2015.

[145] M. Hoke *et al.,* "The impact of selectins on mortality in stable carotid atherosclerosis," *Thrombosis and Haemostasis,* vol. 114, no. 09, pp. 632-638, 2015.

[146] F. Biscetti *et al.,* "Identification of a potential proinflammatory genetic profile influencing carotid plaque vulnerability," *Journal of Vascular Surgery,* vol. 61, no. 2, pp. 374-381, 2015.

[147] W. Eilenberg *et al.,* "Neutrophil gelatinase associated lipocalin (NGAL) for identification of unstable plaques in patients with asymptomatic carotid stenosis," *European Journal of Vascular and Endovascular Surgery,* 2019.

[148] J. Pelisek *et al.,* "Multiple biological predictors for vulnerable carotid lesions," *Cerebrovascular Diseases,* vol. 28, no. 6, pp. 601-610, 2009.

[149] B. Alvarez, C. Ruiz, P. Chacón, J. Alvarez-Sabin, and M. Matas, "Serum values of metalloproteinase-2 and metalloproteinase-9 as related to unstable plaque and inflammatory cells in patients with greater than 70% carotid artery stenosis," *Journal of vascular surgery,* vol. 40, no. 3, pp. 469-475, 2004.

[150] Z. Y. Guo *et al.,* "Specific matrix metalloproteinases and calcification factors are associated with the vulnerability of human carotid plaque," *Experimental and therapeutic medicine,* vol. 16, no. 3, pp. 2071-2079, 2018.

[151] F. Sigala *et al.,* "Oxidized LDL in human carotid plaques is related to symptomatic carotid disease and lesion instability," *Journal of vascular surgery,* vol. 52, no. 3, pp. 704-713, 2010.

[152] C. Zhong *et al.,* "Serum matrix metalloproteinase-9 levels and prognosis of acute ischemic stroke," *Neurology,* vol. 89, no. 8, pp. 805-812, 2017.

[153] L. Fernández-Friera *et al.,* "Normal LDL-cholesterol levels are associated with subclinical atherosclerosis in the absence of risk factors," *Journal of the American College of Cardiology,* vol. 70, no. 24, pp. 2979-2991, 2017.

[154] G. Hindy *et al.*, "Role of blood lipids in the development of ischemic stroke and its subtypes: a Mendelian randomization study," *Stroke,* vol. 49, no. 4, pp. 820-827, 2018.

[155] Y. Shingai *et al.*, "Effect of preoperative administration of proprotein convertase subtilisin/kexin type 9 inhibitor on carotid artery stenting," *World neurosurgery,* vol. 135, pp. e36-e42, 2020.

[156] X. Gu *et al.*, "Association of lipids with ischemic and hemorrhagic stroke: a prospective cohort study among 267 500 Chinese," *Stroke,* vol. 50, no. 12, pp. 3376-3384, 2019.

[157] A. V. Poznyak *et al.*, "Overview of OxLDL and its impact on cardiovascular health: focus on atherosclerosis," *Frontiers in Pharmacology,* p. 2248, 2021.

[158] Y. Ishigaki, Y. Oka, and H. Katagiri, "Circulating oxidized LDL: a biomarker and a pathogenic factor," *Current opinion in lipidology,* vol. 20, no. 5, pp. 363-369, 2009.

[159] K. Nishi *et al.*, "Oxidized LDL in carotid plaques and plasma associates with plaque instability," *Arteriosclerosis, thrombosis, and vascular biology,* vol. 22, no. 10, pp. 1649-1654, 2002.

[160] H. Markstad *et al.*, "High levels of soluble lectinlike oxidized low-density lipoprotein receptor-1 are associated with carotid plaque inflammation and increased risk of ischemic stroke," *Journal of the American Heart Association,* vol. 8, no. 4, p. e009874, 2019.

[161] A. Wang *et al.*, "Association of oxidized low-density lipoprotein with prognosis of stroke and stroke subtypes," *Stroke,* vol. 48, no. 1, pp. 91-97, 2017.

[162] O. Meilhac, "High-density lipoproteins in stroke," *High Density Lipoproteins,* pp. 509-526, 2015.

[163] E. B. Mathiesen, K. H. Bønaa, and O. Joakimsen, "Low levels of high-density lipoprotein cholesterol are associated with echolucent carotid artery plaques: the Tromsø study," *Stroke,* vol. 32, no. 9, pp. 1960-1965, 2001.

[164] S. Peters *et al.*, "Increased age, high body mass index and low HDL-C levels are related to an echolucent carotid intima–media: the METEOR study," *Journal of internal medicine,* vol. 272, no. 3, pp. 257-266, 2012.

[165] B. G. Nordestgaard, M.-L. M. Grønholdt, and H. Sillesen, "Echolucent rupture-prone plaques," *Current opinion in lipidology,* vol. 14, no. 5, pp. 505-512, 2003.

[166] P.-S. Yeh *et al.*, "Low levels of high-density lipoprotein cholesterol in patients with atherosclerotic stroke: a prospective cohort study," *Atherosclerosis,* vol. 228, no. 2, pp. 472-477, 2013.

[167] B. G. Talayero and F. M. Sacks, "The role of triglycerides in atherosclerosis," *Current cardiology reports,* vol. 13, no. 6, pp. 544-552, 2011.

[168] S. C. Kofoed, M.-L. M. Grønholdt, J. Bismuth, J. E. Wilhjelm, H. Sillesen, and B. G. Nordestgaard, "Echolucent, rupture-prone carotid plaques associated with elevated triglyceride-rich lipoproteins, particularly in women," *Journal of vascular surgery,* vol. 36, no. 4, pp. 783-792, 2002.

[169] D. Mannheim *et al.*, "Enhanced expression of Lp-PLA2 and lysophosphatidylcholine in symptomatic carotid atherosclerotic plaques," *Stroke,* vol. 39, no. 5, pp. 1448-1455, 2008.

[170] G. Sarlon-Bartoli *et al.*, "Circulating lipoprotein-associated phospholipase A2 in high-grade carotid stenosis: a new biomarker for predicting unstable plaque," *European Journal of Vascular and Endovascular Surgery,* vol. 43, no. 2, pp. 154-159, 2012.

[171] H.-H. S. Oei *et al.*, "Lipoprotein-associated phospholipase A2 activity is associated with risk of coronary heart disease and ischemic stroke: the Rotterdam Study," *Circulation,* vol. 111, no. 5, pp. 570-575, 2005.

[172] M. S. Elkind, W. Tai, K. Coates, M. C. Paik, and R. L. Sacco, "Lipoprotein-associated phospholipase A2 activity and risk of recurrent stroke," *Cerebrovascular Diseases,* vol. 27, no. 1, pp. 42-50, 2009.

[173] K. B. Walsh *et al.*, "Apolipoprotein AI and paraoxonase-1 are potential blood biomarkers for ischemic stroke diagnosis," *Journal of Stroke and Cerebrovascular Diseases,* vol. 25, no. 6, pp. 1360-1365, 2016.

[174] H. Dong *et al.*, "Apolipoprotein A1, B levels, and their ratio and the risk of a first stroke: a meta-analysis and case–control study," *Metabolic brain disease,* vol. 30, no. 6, pp. 1319-1330, 2015.

[175] L. Paternoster, N. A. n. Martínez González, S. Lewis, and C. Sudlow, "Association between apolipoprotein E genotype and carotid intima-media thickness may suggest a specific effect on large artery atherothrombotic stroke," *Stroke,* vol. 39, no. 1, pp. 48-54, 2008.

[176] N. G. Seidah, Z. Awan, M. Chrétien, and M. Mbikay, "PCSK9: a key modulator of cardiovascular health," *Circulation research,* vol. 114, no. 6, pp. 1022-1036, 2014.

[177] D. C. Chan *et al.*, "Plasma proprotein convertase subtilisin kexin type 9 as a predictor of carotid atherosclerosis in asymptomatic adults," *Heart, Lung and Circulation,* vol. 25, no. 5, pp. 520-525, 2016.

[178] W. Xie *et al.*, "Association between plasma PCSK9 levels and 10-year progression of carotid atherosclerosis beyond LDL-C: a cohort study," *International journal of cardiology,* vol. 215, pp. 293-298, 2016.

[179] K. Gasbarrino, C. Mantzoros, J. Gorgui, J. P. Veinot, C. Lai, and S. S. Daskalopoulou, "Circulating chemerin is associated with carotid plaque instability, whereas resistin is related to cerebrovascular symptomatology," *Arteriosclerosis, thrombosis, and vascular biology,* vol. 36, no. 8, pp. 1670-1678, 2016.

[180] J. Gorgui *et al.*, "Circulating adiponectin levels in relation to carotid atherosclerotic plaque presence, ischemic stroke risk, and mortality: a systematic review and meta-analyses," *Metabolism,* vol. 69, pp. 51-66, 2017.

[181] J. Gairolla, R. Kler, M. Modi, and D. Khurana, "Leptin and adiponectin: pathophysiological role and possible therapeutic target of inflammation in ischemic stroke," *Reviews in the Neurosciences,* vol. 28, no. 3, pp. 295-306, 2017.

[182] S. Gustafsson, L. Lind, S. Soderberg, and E. Ingelsson, "Associations of circulating adiponectin with measures of vascular function and morphology," *The Journal of Clinical Endocrinology & Metabolism,* vol. 95, no. 6, pp. 2927-2934, 2010.

[183] L. A. Saarikoski *et al.*, "Adiponectin is related with carotid artery intima-media thickness and brachial flow-mediated dilatation in young adults—the Cardiovascular Risk in Young Finns Study," *Annals of medicine,* vol. 42, no. 8, pp. 603-611, 2010.

[184] S. Holm *et al.*, "Fatty Acid binding protein 4 is associated with carotid atherosclerosis and outcome in patients with acute ischemic stroke," *PLoS One,* vol. 6, no. 12, p. e28785, 2011.

[185] H. Agardh *et al.*, "Expression of fatty acid–binding protein 4/aP2 is correlated with plaque instability in carotid atherosclerosis," *Journal of internal medicine,* vol. 269, no. 2, pp. 200-210, 2011.

[186] S. Alsulaimani, H. Gardener, M. S. Elkind, K. Cheung, R. L. Sacco, and T. Rundek, "Elevated homocysteine and carotid plaque area and densitometry in the Northern Manhattan Study," *Stroke,* vol. 44, no. 2, pp. 457-461, 2013.

[187] B. Alvarez, X. Yugueros, E. Fernández, F. Luccini, A. Gené, and M. Matas, "Relationship between plasma homocysteine and the morphological and immunohistochemical study of carotid plaques in patients with carotid stenosis over 70%," *Annals of vascular surgery,* vol. 26, no. 4, pp. 500-505, 2012.

[188] J. Jia *et al.*, "Homocysteine and its relationship to asymptomatic carotid stenosis in a Chinese community population," *Scientific reports,* vol. 6, no. 1, pp. 1-7, 2016.

[189] J.-M. Davaine *et al.*, "Osteoprotegerin, pericytes and bone-like vascular calcification are associated with carotid plaque stability," *PLoS One*, vol. 9, no. 9, p. e107642, 2014.

[190] A. Schiro, F. L. Wilkinson, R. Weston, J. V. Smyth, F. Serracino-Inglott, and M. Y. Alexander, "Elevated levels of endothelial-derived microparticles and serum CXCL9 and SCGF-β are associated with unstable asymptomatic carotid plaques," *Scientific reports*, vol. 5, no. 1, pp. 1-12, 2015.

[191] A. Saxena, E. Y. K. Ng, and S. T. Lim, "Imaging modalities to diagnose carotid artery stenosis: progress and prospect," *Biomedical engineering online*, vol. 18, no. 1, pp. 1-23, 2019.

[192] Z. Ma, Y. Yue, Y. Luo, W. Wang, Y. Cao, and Q. Fang, "Clinical utility of the inflammatory factors combined with lipid markers in the diagnostic and prognostic assessment of ischemic stroke: Based on logistic regression models," *Journal of Stroke and Cerebrovascular Diseases*, vol. 29, no. 4, p. 104653, 2020.

[193] F. J. Mayer *et al.*, "Combined effects of inflammatory status and carotid atherosclerosis: a 12-year follow-up study," *Stroke*, vol. 47, no. 12, pp. 2952-2958, 2016.

[194] S. Kyriakidis *et al.*, "An All-in-One Tool for 2D Atherosclerotic Disease Assessment and 3D Coronary Artery Reconstruction," *Journal of Cardiovascular Development and Disease*, vol. 10, no. 3, p. 130, 2023.

[195] J. A. Damen *et al.*, "Prediction models for cardiovascular disease risk in the general population: systematic review," *BMJ (Clinical research ed.)*, vol. 353, p. i2416, May 16 2016, doi: 10.1136/bmj.i2416.

[196] S. Gauss, S. Achenbach, T. Pflederer, A. Schuhbäck, W. G. Daniel, and M. Marwan, "Assessment of coronary artery remodelling by dual-source CT: a head-to-head comparison with intravascular ultrasound," *Heart*, vol. 97, no. 12, pp. 991-997, 2011.

[197] S.-L. Papadopoulou *et al.*, "Detection and quantification of coronary atherosclerotic plaque by 64-slice multidetector CT: A systematic head-to-head comparison with intravascular ultrasound," *Atherosclerosis*, vol. 219, no. 1, pp. 163-170, doi: 10.1016/j.atherosclerosis.2011.07.005.

[198] F. Saremi and S. Achenbach, "Coronary plaque characterization using CT," *American Journal of Roentgenology*, vol. 204, no. 3, pp. W249-W260, 2015.

[199] A. P. Carnicelli *et al.*, "Cross-sectional area for the calculation of carotid artery stenosis on computed tomographic angiography," *Journal of Vascular Surgery*, vol. 58, no. 3, pp. 659-665, 2013.

[200] D. Andreini *et al.*, "A long-term prognostic value of coronary CT angiography in suspected coronary artery disease," *JACC: Cardiovascular Imaging*, vol. 5, no. 7, pp. 690-701, 2012.

[201] A. Sakellarios *et al.*, "Prediction of Atherosclerotic Plaque Development in an in Vivo Coronary Arterial Segment Based on a Multi-level Modeling Approach," *IEEE Transactions on Biomedical Engineering*, 2016.

[202] C. V. Bourantas *et al.*, "Noninvasive prediction of atherosclerotic progression: The PROSPECT-MSCT study," *JACC: Cardiovascular Imaging*, vol. 9, no. 8, pp. 1009-1011, 2016.

[203] K. Dakis *et al.*, "Carotid Plaque Vulnerability Diagnosis by CTA versus MRA: A Systematic Review," *Diagnostics*, vol. 13, no. 4, p. 646, 2023.

[204] V. I. Kigka *et al.*, "3D reconstruction of coronary arteries and atherosclerotic plaques based on computed tomography angiography images," *Biomedical Signal Processing and Control*, vol. 40, pp. 286-294, 2018/02/01/ 2018, doi: https://doi.org/10.1016/j.bspc.2017.09.009.

[205] V. I. Kigka *et al.*, "A three-dimensional quantification of calcified and non-calcified plaques in coronary arteries based on computed tomography coronary angiography images: Comparison with expert's annotations and virtual histology intravascular ultrasound," *Computers in biology and medicine*, vol. 113, p. 103409, 2019.

[206] V. I. Kigka *et al.*, "Three-Dimensional Reconstruction of Carotid Arteries Using Computed Tomography Angiography," in *European Medical and Biological Engineering Conference*, 2020: Springer, pp. 1130-1136.

[207] A. F. Frangi, W. J. Niessen, K. L. Vincken, and M. A. Viergever, "Multiscale vessel enhancement filtering," in *Medical Image Computing and Computer-Assisted Intervention — MICCAI'98: First International Conference Cambridge, MA, USA, October 11–13, 1998 Proceedings*, W. M. Wells, A. Colchester, and S. Delp Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 1998, pp. 130-137.

[208] G. Ayers and J. C. Dainty, "Iterative blind deconvolution method and its applications," *Optics letters*, vol. 13, no. 7, pp. 547-549, 1988.

[209] A. M. Castillo-Amor, C. A. Navarro-Navia, A. J. Cadena-Bonfanti, and S. H. Contreras-Ortiz, "Reduction of blooming artifacts in cardiac CT images by blind deconvolution and anisotropic diffusion filtering," in *11th International Symposium on Medical Information Processing and Analysis*, 2015, vol. 9681: International Society for Optics and Photonics, p. 96810P.

[210] C. Metz, M. Schaap, A. Weustink, N. Mollet, T. van Walsum, and W. Niessen, "Coronary centerline extraction from CT coronary angiography images using a minimum cost path approach," *Medical physics*, vol. 36, no. 12, pp. 5568-5579, 2009.

[211] J. A. Bærentzen, "On the implementation of fast marching methods for 3D lattices," 2001.

[212] S. Leschka *et al.*, "Ex vivo evaluation of coronary atherosclerotic plaques: Characterization with dual-source CT in comparison with histopathology," *Journal of Cardiovascular Computed Tomography*, vol. 4, no. 5, pp. 301-308, doi: 10.1016/j.jcct.2010.05.016.

[213] M. Galonska, F. Ducke, T. Kertesz-Zborilova, R. Meyer, H. Guski, and F. D. Knollmann, "Characterization of Atherosclerotic Plaques in Human Coronary Arteries With 16-Slice Multidetector Row Computed Tomography by Analysis of Attenuation Profiles," *Academic Radiology*, vol. 15, no. 2, pp. 222-230, doi: 10.1016/j.acra.2007.09.007.

[214] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Transactions on Image Processing*, vol. 10, no. 2, pp. 266-277, 2001, doi: 10.1109/83.902291.

[215] T. Chan and W. Zhu, "Level set based shape prior segmentation," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, 2005, vol. 2: IEEE, pp. 1164-1170.

[216] D. Cremers, N. Sochen, and C. Schnörr, "Towards recognition-based variational segmentation using shape priors and dynamic labeling," in *International Conference on Scale-Space Theories in Computer Vision*, 2003: Springer, pp. 388-400.

[217] R. T. Whitaker, "A Level-Set Approach to 3D Reconstruction from Range Data," *International Journal of Computer Vision*, vol. 29, no. 3, pp. 203-231, 1998/09/01 1998, doi: 10.1023/A:1008036829907.

[218] W. E. Lorensen and H. E. Cline, "Marching cubes: A high resolution 3D surface construction algorithm," in *ACM siggraph computer graphics*, 1987, vol. 21, no. 4: ACM, pp. 163-169.

[219] T. Chan and Z. Wei, "Level set based shape prior segmentation," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, 20-25 June 2005 2005, vol. 2, pp. 1164-1170 vol. 2, doi: 10.1109/CVPR.2005.212.

[220] D. Cremers, N. Sochen, and C. Schnörr, "Towards Recognition-Based Variational Segmentation Using Shape Priors and Dynamic Labeling," in *Scale Space Methods in Computer Vision: 4th International Conference, Scale Space 2003 Isle of Skye, UK, June 10–12, 2003 Proceedings*, L.

D. Griffin and M. Lillholm Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003, pp. 388-400.

[221] M. G. Dalager *et al.*, "Impact of luminal density on plaque classification by CT coronary angiography," *The International Journal of Cardiovascular Imaging,* journal article vol. 27, no. 4, pp. 593-600, April 01 2011, doi: 10.1007/s10554-010-9695-z.

[222] W. E. Lorensen and H. E. Cline, "Marching cubes: A high resolution 3D surface construction algorithm," presented at the Proceedings of the 14th annual conference on Computer graphics and interactive techniques, 1987.

[223] W. R. Crum, O. Camara, and D. L. Hill, "Generalized overlap measures for evaluation and validation in medical image analysis," *IEEE transactions on medical imaging,* vol. 25, no. 11, pp. 1451-1461, 2006.

[224] C. V. Bourantas *et al.*, "A method for 3D reconstruction of coronary arteries using biplane angiography and intravascular ultrasound images," *Computerized Medical Imaging and Graphics,* vol. 29, no. 8, pp. 597-606, 12// 2005, doi: http://dx.doi.org/10.1016/j.compmedimag.2005.07.001.

[225] A. A. Taha and A. Hanbury, "Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool," *BMC medical imaging,* vol. 15, no. 1, p. 29, 2015.

[226] K. H. Zou *et al.*, "Statistical Validation of Image Segmentation Quality Based on a Spatial Overlap Index: Scientific Reports," *Academic radiology,* vol. 11, no. 2, pp. 178-189, 2004, doi: 10.1016/S1076-6332(03)00671-8.

[227] J. M. K. S. U-King-Im *et al.*, "Measuring Carotid Stenosis on Contrast-Enhanced Magnetic Resonance Angiography," *Diagnostic Performance and Reproducibility of 3 Different Methods,* vol. 35, no. 9, pp. 2083-2088, 2004, doi: 10.1161/01.STR.0000136722.30008.b1.

[228] A. W. Leber *et al.*, "Accuracy of 64-slice computed tomography to classify and quantify plaque volumes in the proximal coronary system: a comparative study using intravascular ultrasound," *Journal of the American College of Cardiology,* vol. 47, no. 3, pp. 672-677, 2006.

[229] H. Brodoefel *et al.*, "Accuracy of dual-source CT in the characterisation of non-calcified plaque: use of a colour-coded analysis compared with virtual histology intravascular ultrasound," *The British journal of radiology,* 2014.

[230] M. Utsunomiya, H. Hara, M. Moroi, K. Sugi, and M. Nakamura, "Relationship between tissue characterization with 40 MHz intravascular ultrasound imaging and 64-slice computed

tomography," *Journal of Cardiology,* vol. 57, no. 3, pp. 297-302, 5// 2011, doi: http://dx.doi.org/10.1016/j.jjcc.2011.01.016.

[231] K. Hameeteman *et al.*, "Evaluation framework for carotid bifurcation lumen segmentation and stenosis grading," *Medical image analysis,* vol. 15, no. 4, pp. 477-488, 2011.

[232] V. I. Kigka *et al.*, "Machine Learning Coronary Artery Disease Prediction Based on Imaging and Non-Imaging Data," *Diagnostics,* vol. 12, no. 6, p. 1466, 2022.

[233] V. I. Kigka *et al.*, "A Machine Learning Approach for the Prediction of the Progression of Cardiovascular Disease based on Clinical and Non-Invasive Imaging Data," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2018: IEEE, pp. 6108-6111.

[234] V. I. Kigka *et al.*, "Site specific prediction of atherosclerotic plaque progression using computational biomechanics and machine learning," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2019: IEEE, pp. 6998-7001.

[235] V. I. Kigka *et al.*, "Site specific prediction of PCI stenting based on imaging and biomechanics data using gradient boosting tree ensembles," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 2020: IEEE, pp. 2812-2815.

[236] P. K. Siogkas *et al.*, "Noninvasive CT-based hemodynamic assessment of coronary lesions derived from fast computational analysis: a comparison against fractional flow reserve," *European Radiology,* journal article vol. 29, no. 4, pp. 2117-2126, April 01 2019, doi: 10.1007/s00330-018-5781-8.

[237] M. I. Papafaklis *et al.*, "Fast virtual functional assessment of intermediate coronary lesions using routine angiographic data and blood flow simulation in humans: comparison with pressure wire-fractional flow reserve," *EuroIntervention: journal of EuroPCR in collaboration with the Working Group on Interventional Cardiology of the European Society of Cardiology,* vol. 10, no. 5, pp. 574-583, 2014.

[238] R. C. Cury *et al.*, "CAD-RADSTM coronary artery disease–reporting and data system. An expert consensus document of the Society of Cardiovascular Computed Tomography (SCCT), the American College of Radiology (ACR) and the North American Society for Cardiovascular Imaging (NASCI). Endorsed by the American College of Cardiology," *Journal of cardiovascular computed tomography,* vol. 10, no. 4, pp. 269-281, 2016.

[239] G. L. Raff *et al.*, "SCCT guidelines for the interpretation and reporting of coronary computed tomographic angiography," *Journal of cardiovascular computed tomography*, vol. 3, no. 2, pp. 122-136, 2009.

[240] C. Robert, "Machine learning, a probabilistic perspective," ed: Taylor & Francis, 2014.

[241] X. Y. Liu, J. Wu, and Z. H. Zhou, "Exploratory undersampling for class-imbalance learning," (in eng), *IEEE transactions on systems, man, and cybernetics. Part B, Cybernetics : a publication of the IEEE Systems, Man, and Cybernetics Society*, vol. 39, no. 2, pp. 539-50, Apr 2009, doi: 10.1109/tsmcb.2008.2007853.

[242] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, "Gene selection for cancer classification using support vector machines," *Machine learning*, vol. 46, no. 1-3, pp. 389-422, 2002.

[243] J. H. Friedman, "Greedy Function Approximation: A Gradient Boosting Machine," *The Annals of Statistics*, vol. 29, no. 5, pp. 1189-1232, 2001. [Online]. Available: http://www.jstor.org/stable/2699986.

[244] R. Liga *et al.*, "Multicentre multi-device hybrid imaging study of coronary artery disease: results from the EValuation of INtegrated Cardiac Imaging for the Detection and Characterization of Ischaemic Heart Disease (EVINCI) hybrid imaging population," *European Heart Journal-Cardiovascular Imaging*, vol. 17, no. 9, pp. 951-960, 2016.

[245] A. I. Sakellarios *et al.*, "SMARTool: A tool for clinical decision support for the management of patients with coronary artery disease based on modeling of atherosclerotic plaque process," in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2017: IEEE, pp. 96-99.

[246] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," *Advances in neural information processing systems*, vol. 30, 2017.

[247] A. R. Cappola *et al.*, "Thyroid and cardiovascular disease: research agenda for enhancing knowledge, prevention, and treatment," *Circulation*, vol. 139, no. 25, pp. 2892-2909, 2019.

[248] K. Inoue, B. Ritz, G. A. Brent, R. Ebrahimi, C. M. Rhee, and A. M. Leung, "Association of subclinical hypothyroidism and cardiovascular disease with mortality," *JAMA Network Open*, vol. 3, no. 2, pp. e1920745-e1920745, 2020.

[249] I. Klein and S. Danzi, "Thyroid disease and the heart," *Circulation*, vol. 116, no. 15, pp. 1725-1735, 2007.

[250] E. Galli, A. Pingitore, and G. Iervasi, "The role of thyroid hormone in the pathophysiology of heart failure: clinical evidence," *Heart failure reviews,* vol. 15, no. 2, pp. 155-169, 2010.

[251] H.-Y. Jin *et al.*, "The relationship between coronary calcification and the natural history of coronary artery disease," *Cardiovascular Imaging,* vol. 14, no. 1, pp. 233-242, 2021.

[252] R. Virmani, A. P. Burke, A. Farb, and F. D. Kolodgie, "Pathology of the vulnerable plaque," *Journal of the American College of Cardiology,* vol. 47, no. 8S, pp. C13-C18, 2006.

[253] C. P. Nelson *et al.*, "Genetically determined height and coronary artery disease," *New England Journal of Medicine,* vol. 372, no. 17, pp. 1608-1618, 2015.

[254] J. Moon and I. C. Hwang, "The link between height and cardiovascular disease: to be deciphered," *Cardiology,* vol. 143, no. 3-4, pp. 114-115, 2019.

[255] P. N. Tan, M. Steinbach, and V. Kumar, *Introduction to Data Mining*. Pearson Addison Wesley, 2006.

[256] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA data mining software: an update," *SIGKDD Explor. Newsl.,* vol. 11, no. 1, pp. 10-18, 2009, doi: 10.1145/1656274.1656278.

[257] S. L. Salzberg, "C4.5: Programs for Machine Learning by J. Ross Quinlan. Morgan Kaufmann Publishers, Inc., 1993," *Machine Learning,* journal article vol. 16, no. 3, pp. 235-240, September 01 1994, doi: 10.1007/bf00993309.

[258] R. Liga *et al.*, "Multicentre multi-device hybrid imaging study of coronary artery disease: results from the EValuation of INtegrated Cardiac Imaging for the Detection and Characterization of Ischaemic Heart Disease (EVINCI) hybrid imaging population," *European Heart Journal - Cardiovascular Imaging,* vol. 17, no. 9, pp. 951-960, 2016, doi: 10.1093/ehjci/jew038.

[259] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," *Journal of artificial intelligence research,* vol. 16, pp. 321-357, 2002.

[260] A. I. Sakellarios *et al.*, "SMARTool: A tool for clinical decision support for the management of patients with coronary artery disease based on modeling of atherosclerotic plaque process," in *Engineering in Medicine and Biology Society (EMBC), 2017 39th Annual International Conference of the IEEE*, 2017: IEEE, pp. 96-99.

[261] X.-Y. Liu, J. Wu, and Z.-H. Zhou, "Exploratory undersampling for class-imbalance learning," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics),* vol. 39, no. 2, pp. 539-550, 2008.

[262] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016, pp. 785-794.

[263] S. F. Weng, J. Reps, J. Kai, J. M. Garibaldi, and N. Qureshi, "Can machine-learning improve cardiovascular risk prediction using routine clinical data?," *PloS one,* vol. 12, no. 4, p. e0174944, 2017.

[264] M. S. Bittencourt *et al.*, "Prognostic value of nonobstructive and obstructive coronary artery disease detected by coronary computed tomography angiography to identify cardiovascular events," *Circulation: Cardiovascular Imaging,* vol. 7, no. 2, pp. 282-291, 2014.

[265] D. Dey *et al.*, "Comparison of quantitative atherosclerotic plaque burden from coronary CT angiography in patients with first acute coronary syndrome and stable coronary artery disease," *Journal of cardiovascular computed tomography,* vol. 8, no. 5, pp. 368-374, 2014.

[266] R. B. D'agostino *et al.*, "General cardiovascular risk profile for use in primary care: the Framingham Heart Study," *Circulation,* vol. 117, no. 6, pp. 743-753, 2008.

[267] B. A. Goldstein, A. M. Navar, and R. E. Carter, "Moving beyond regression techniques in cardiovascular risk prediction: applying machine learning to address analytic challenges," *European Heart Journal,* vol. 38, no. 23, pp. 1805-1814, 2017, doi: 10.1093/eurheartj/ehw302.

[268] V. C. Pezoulas *et al.*, "Medical data quality assessment: On the development of an automated framework for medical data curation," *Computers in biology and medicine,* vol. 107, pp. 270-283, 2019.

[269] K. Hajian-Tilaki, "Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation," *Caspian journal of internal medicine,* vol. 4, no. 2, p. 627, 2013.

[270] N. Timmerman *et al.*, "The TAXINOMISIS Project: A multidisciplinary approach for the development of a new risk stratification model for patients with asymptomatic carotid artery stenosis," *European Journal of Clinical Investigation,* vol. 50, no. 12, p. e13411, 2020.

[271] J. Yin *et al.*, "Detection of Asymptomatic Carotid Artery Stenosis in High-Risk Individuals of Stroke Using a Machine-Learning Algorithm," *Chinese Medical Sciences Journal,* vol. 35, no. 4, pp. 297-305, 2020.

[272] J. Yu *et al.*, "Machine learning models for screening carotid atherosclerosis in asymptomatic adults," *Scientific reports,* vol. 11, no. 1, pp. 1-8, 2021.

[273] J. Fan *et al.*, "The prediction of asymptomatic carotid atherosclerosis with electronic health records: a comparative study of six machine learning models," *BMC medical informatics and decision making,* vol. 21, no. 1, pp. 1-9, 2021.

[274] M. H. Poorthuis *et al.*, "Validation of risk prediction models to detect asymptomatic carotid stenosis," *Journal of the American Heart Association,* vol. 9, no. 8, p. e014766, 2020.

[275] C. Finn *et al.*, "The association between carotid artery atherosclerosis and silent brain infarction: a systematic review and meta-analysis," *Journal of Stroke and Cerebrovascular Diseases,* vol. 26, no. 7, pp. 1594-1601, 2017.

[276] M. D. Benli, B. Güven, H. Güven, and I. Conkbayır, "Silent brain infarcts and white matter lesions in patients with asymptomatic carotid stenosis," *Acta Neurologica Belgica,* vol. 121, pp. 983-991, 2021.

[277] C. Rudolph and N. Eldrup, "Asymptomatic carotid stenosis and concomitant silent brain infarctions," *Vascular,* vol. 28, no. 1, pp. 7-15, 2020.

[278] Y. Ling, E. Jouvent, L. Cousyn, H. Chabriat, and F. De Guio, "Validation and optimization of BIANCA for the segmentation of extensive white matter hyperintensities," *Neuroinformatics,* vol. 16, no. 2, pp. 269-281, 2018.

[279] H. Baradaran, G. Gialdini, E. Mtui, G. Askin, H. Kamel, and A. Gupta, "Silent brain infarction in patients with asymptomatic carotid artery atherosclerotic disease," *Stroke,* vol. 47, no. 5, pp. 1368-1370, 2016.

[280] D. Inzitari, M. Eliasziw, B. Sharpe, A. Fox, H. Barnett, and N. A. S. C. E. T. Group, "Risk factors and outcome of patients with carotid artery stenosis presenting with lacunar stroke," *Neurology,* vol. 54, no. 3, pp. 660-660, 2000.

[281] H. Singh and G. Mohan, "To Study the Incidence of Lacunar Infarcts in Patients with Acute Ischemic Stroke and its Correlation with Carotid Artery Stenosis," *Current Trends in Diagnosis and Treatment,* vol. 2, no. 2, 2018.

[282] J. Slark, P. Bentley, and P. Sharma, "Silent brain infarction in the presence of systemic vascular disease," *JRSM Cardiovascular Disease,* vol. 1, no. 1, pp. 1-9, 2012.

[283] A. F. AbuRahma *et al.*, "Society for Vascular Surgery clinical practice guidelines for management of extracranial cerebrovascular disease," *Journal of vascular surgery,* vol. 75, no. 1, pp. 4S-22S, 2022.

[284] O. Klass *et al.*, "Coronary plaque imaging with 256-slice multidetector computed tomography: interobserver variability of volumetric lesion parameters with semiautomatic plaque analysis software," *The international journal of cardiovascular imaging,* vol. 26, no. 6, pp. 711-720, 2010.

[285] C. R. Becker *et al.*, "Ex vivo coronary atherosclerotic plaque characterization with multi-detector-row CT," *European Radiology,* journal article vol. 13, no. 9, pp. 2094-2098, September 01 2003, doi: 10.1007/s00330-003-1889-5.

[286] R. Chopard *et al.*, "How reliable are 40 MHz IVUS and 64-slice MDCT in characterizing coronary plaque composition? An ex vivo study with histopathological comparison," *The International Journal of Cardiovascular Imaging,* journal article vol. 26, no. 4, pp. 373-383, April 01 2010, doi: 10.1007/s10554-009-9562-y.

[287] T. Kitagawa *et al.*, "Comprehensive evaluation of noncalcified coronary plaque characteristics detected using 64-slice computed tomography in patients with proven or suspected coronary artery disease," *American Heart Journal,* vol. 154, no. 6, pp. 1191-1198, 2007/12/01/ 2007, doi: http://dx.doi.org/10.1016/j.ahj.2007.07.020.

[288] S. Leschka *et al.*, "Ex vivo evaluation of coronary atherosclerotic plaques: Characterization with dual-source CT in comparison with histopathology," *Journal of Cardiovascular Computed Tomography,* vol. 4, no. 5, pp. 301-308, 2010/09/01/ 2010, doi: http://dx.doi.org/10.1016/j.jcct.2010.05.016.

[289] M. Marwan *et al.*, "In vivo CT detection of lipid-rich coronary artery atherosclerotic plaques using quantitative histogram analysis: A head to head comparison with IVUS," *Atherosclerosis,* vol. 215, no. 1, pp. 110-115, 2011/03/01/ 2011, doi: http://dx.doi.org/10.1016/j.atherosclerosis.2010.12.006.

[290] M. Galonska, F. Ducke, T. Kertesz-Zborilova, R. Meyer, H. Guski, and F. D. Knollmann, "Characterization of Atherosclerotic Plaques in Human Coronary Arteries With 16-Slice Multidetector Row Computed Tomography by Analysis of Attenuation Profiles," *Academic Radiology,* vol. 15, no. 2, pp. 222-230, 2008/02/01/ 2008, doi: http://dx.doi.org/10.1016/j.acra.2007.09.007.

Appendix

Threshold selection (Chapter 3)

The selection of the intensity thresholds of the lumen and the CP were based on the current literature. More specifically, due to the variety of CT scanners, the average range of lumen intensity is 200 HU–500 HU [221], whereas the intensity of CP is higher than 500 HU [284]. In this methodology, we have to identify the optimal threshold values for the discrimination of the lumen, the outer wall and the CP. These threshold values are extracted based on the calculation of the mean lumen intensity (I_{lumen}), in combination with the known ranges of the lumen and CP intensities. Thus, as far as the lumen threshold is concerned and considering that the intensity of the lumen is affected by the acquisition dose protocol, we selected a relative small l_{thres} (l_{thres} =80HU). As it is demonstrated in Table 1 (page 33), in order to find the lower limit of the lumen intensity value, the l_{thres} is subtracted from the mean lumen intensity and the value of 80 HU is a good approximation in order to agree with the lumen range proposed from the literature (around 500HU). As it is demonstrated in the Figure 10, the HU values which are possible to match with the lumen are about 500 HU. Furthermore, several experiments have been implemented using l_{thres} values close to 80HU in order to examine the algorithm effectiveness for the whole range of the membership function demonstrated in Figure 10 (page 33). A similar approach has been implemented for the CP. More specifically, according to current literature HU>500 may identify CP. For this purpose our approach was implemented after the detection of the lumen border and especially at the region out of the lumen border. In a similar approach the cp_{thres} should be 400HU for optimal accuracy in CP detection.

As far as the attenuation HU value for NCP is concerned, it depends also on the contrast protocol. However, based on the literature a potential range for NCP is from 0 to100 HU [228, 285-290]. Thus, the threshold value of NCP is defined to be 50 HU, in order to include HU values around 50 HU, depending on the density of the lumen. This selected threshold value is considered as an indicative value for the NCP, which is adapted to the lumen density.

Full list of eligibility, inclusion, exclusion and exit criteria (Chapter 4)

Eligibility criteria:

A. Clinical history and lifestyle data records available at one-time point.

B. At least one previous CCTA examination performed for suspected CHD and of good quality to allow for: a) Non-invasive FFR-CT assessment b) Quantitative (automated) 17 segments (AHA) analysis and measurement with ≤10% error of MLA (mm2), lumen area stenosis (%), mean PB (mm3), PB at MLA (%), and remodeling index, c) Plaque phenotype assessment: HU based classification in CP, NCP and mixed, napkin-ring sign, CAC score.

C. Previous blood and plasma sample available for retrospective analysis

Inclusion criteria:

1) male and female subjects

2) aged 45-82 years

3) Caucasian population

4) submitted to CCTA for suspected CHD between 2009 and 2012 (in the context of EVINCI and ARTreat FPVII studies) at the Hospitals reported in "SMARTool Clinical Center" document and satisfying the elegibility criteria reported above

5) submitted to clinical Follow-up in the last 6 months with stable clinical conditions and documented CHD or persistent intermediate/high probability of CHD

6) Signed informed consents (clinical and genetic)

Exclusion criteria:

1) Multi-vessel severe disease (3 vessels and/or LM disease with >90% stenosis).

2) Severe coronary calcification (CAC score > 600).

3) Having undergone surgical procedures related to heart diseases (valve replacement, CRT or CRTD treatment, any surgery of the heart or arteries).

4) Documented MACE at history (MI, severe heart failure, recurrent angina) in the last 6 months with/without revascularization

5) Documented severe peripheral vascular disease (carotid, femoral)

6) Surgery of carotid and/or peripheral arteries or cerebral ischemic attack

7) History/surgery of Abdominal Aortic Aneurysm(AAA).

8) Severe Heart failure (NYHA Class III-IV)

9) LV dysfunction (left ventricle EF <40%).

10) Atrial fibrillation.

11) Lack of written informed consent (clinical consent and/or genetic consent)

12) Pregnancy (evaluated by urine test) and breastfeeding

13) Active Cancer

14) Asthma

15) Cardiomyopathy or congenital heart disease

16) Significant valvular disease (hemodynamically significant valvular stenosis or insufficiency by echoDoppler)

17) Renal dysfunction (creatinine > 1.3 mg/dL)

18) Chronic Kidney Disease (eGFR < 30 ml/min/1.73 m2)

19) Hepatic failure (at least 3 of the following: albumin < 3.5 g/dL; prolonged prothrombin time–PT; jaunDICE; ascites)

20) Waldenstrom disease

21) Multiple myeloma

22) Autoimmune/Acute inflammatory disease

23) Previous severe adverse reaction to iodine contrast agent

24) Positivity at blood tests for HIV, Hepatitis B and C (CRF number 1-clinical evaluation)

Exit Criteria:

A) Informed consent retired by the patient (genetic or clinical)

B) Adverse events to contrast medium during

Author's Publications

Journal Publications

1. **V. I. Kigka**, G. Rigas, A. Sakellarios, P. Siogkas, I. O. Andrikos, T. Exarchos, D. Loggitsi, C. Anagnostopoulos, L. Michalis, D. Neglia, G. Pelosi, O. Parodi, D. Fotiadis, "3D Reconstruction of Coronary Arteries and Atherosclerotic Plaquesbased on Computed Tomography Angiography images", Biomedical Signal Processing and Control, Elsevier, 2018.
2. **V. I. Kigka**, A. Sakellarios, G. Rigas, L. Athanasiou, P. Tsompou, G. Karanasiou, P. Lemos, L. Michalis, D. Fotiadis, "A three-dimensional quantification of calcified and non-calcified plaques in coronary arteries based on computed tomography coronary angiography images: Comparison with expert's annotations and virtual histology intravascular ultrasound", Computers in Biology and Medicine, 2019.
3. **V. I. Kigka**, V. Potsika, M. Mantzaris, V. Tsakanikas, I. Koncar, D. I Fotiadis, Serum Biomarkers in Carotid Artery Disease, Diagnsotics, 2021.
4. **Vassiliki I Kigka**, Eleni Georga, Vassilis Tsakanikas, Savvas Kyriakidis, Panagiota Tsompou, Panagiotis Siogkas, Lampros K Michalis, Katerina K Naka, Danilo Neglia, Silvia Rocchiccioli, Gualtiero Pelosi, Dimitrios I Fotiadis, Antonis Sakellarios, Machine learning coronary artery disease prediction based on imaging and non-imaging data, Diagnostis, MDPI, 2022.
5. Vassiliki I. Kigka *et al.*, Estimating the risk for asymptomatic carotid artery disease using machine learning, (to be submitted)
6. Savvas Kyriakidis, George Rigas, **Vassiliki Kigka**, Dimitris Zaridis, Georgia Karanasiou, Panagiota Tsompou, Gianna Karanasiou, Lampros Lakkas, Sotirios Nikopoulos, Katerina K Naka, Lampros K Michalis, Dimitrios I Fotiadis, Antonis I Sakellarios, An All-in-One Tool for 2D Atherosclerotic Disease Assessment and 3D Coronary Artery Reconstruction, Journal of Cardiovascular Development and Disease, MDPI, Journal of Cardiovascular Development and Disease, 2023.
7. Antonis I Sakellarios, Panagiotis Siogkas, **Vassiliki Kigka**, Panagiota Tsompou, Dimitrios Pleouras, Savvas Kyriakidis, Georgia Karanasiou, Gualtiero Pelosi, Sotirios Nikopoulos, Katerina K Naka, Silvia Rocchiccioli, Lampros K Michalis, Dimitrios I Fotiadis, Error Propagation in the Simulation of Atherosclerotic Plaque Growth and the Prediction of Atherosclerotic Disease Progression, Diagnostics, MDPI, 2021.

8. Antonis I Sakellarios, Panagiota Tsompou, **Vassiliki Kigka**, Panagiotis Siogkas, Savvas Kyriakidis, Nikolaos Tachos, Georgia Karanasiou, Arthur Scholte, Alberto Clemente, Danilo Neglia, Oberdan Parodi, Juhani Knuuti, Lampros K Michalis, Gualtiero Pelosi, Silvia Rocchiccioli, Dimitrios I Fotiadis, Non-invasive prediction of site-specific coronary atherosclerotic plaque progression using lipidomics, blood flow, and LDL transport modeling, Applied Sciences, MDPI, 2021.
9. Dimitrios S Pleouras, Antonis I Sakellarios, Panagiota Tsompou, **Vassiliki Kigka**, Savvas Kyriakidis, Silvia Rocchiccioli, Danilo Neglia, Juhani Knuuti, Gualtiero Pelosi, Lampros K Michalis, Dimitrios I Fotiadis, Simulation of atherosclerotic plaque growth using computational biomechanics and patient-specific data, Scientific Reports, 2020.
10. Dimitrios S Pleouras, Antonis I Sakellarios, George Rigas, Georgia Karanasiou, Panagiota Tsompou, Gianna Karanasiou, **Vassiliki Kigka**, Savvas Kyriakidis, Vasileios Pezoulas, George Gois, Nikolaos Tachos, Aidonis Ramos, Gualtiero Pelosi, Silvia Rocchiccioli, Lampros K Michalis, Dimitrios I Fotiadis, A Novel Approach to Generate a Virtual Population of Human Coronary Arteries for In Silico Clinical Trials of Stent Design, IEEE Open Journal of Engineering in Medicine and Biology, 2021.
11. Theofilos Karasavvidis, Vasileios Bouris, William Xiang, Georgios Tzavellas, Nektarios Charisis, Leonidas Palaiodimos, **Vassiliki Kigka**, Christos V Bourantas, Ioannis Gkiatas, Prophylaxis for Venous Thromboembolic Events in Elective Total Hip and Total Knee Arthroplasty, Current Pharmaceutical Design, 2022.

Conference Publications

1. **V. I. Kigka**, A. Sakellarios, G. Rigas, P. Tsobou, I. O. Andrikos, L. K. Michalis, D. I. Fotiadis," A three-dimensional quantification of non-calcified plaque based on computed tomography coronary angiography images: comparison with virtual histology intravascular ultrasound", World Congress on Medical Physics & Biomedical Engineering (IUPESM), Prague, Czech Republic, 3-8 June, 2018.
2. **V. I. Kigka**, E. I. Georga, A. I. Sakellarios, N. S. Tachos, I. Andrikos, P. Tsompou, S. Rocchiccioli, G. Pelosi, O. Parodi, L. K. Michalis, D. I. Fotiadis, "A Machine Learning Approach for the Prediction of the Progression of Cardiovascular Disease based on Clinical and Non-Invasive Imaging Data", 40th International Engineering in Medicine and Biology Conference (EMBC), Honolulu, Hawaii, July 17-21, 2018.
3. **Vassiliki I Kigka**, Antonis I Sakellarios, Eleni I Georga, Panagiotis Siogkas, Panagiota Tsompou, Savvas Kyriakidis, Silvia Rocchiccioli, Gualtiero Pelosi, Katerina Naka, Lampros K Michalis,

Dimitrios I Fotiadis, Site specific prediction of PCI stenting based on imaging and biomechanics data using gradient boosting tree ensembles, 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2020.

4. **Vassiliki I Kigka**, Antonis I Sakellarios, Michalis D Mantzaris, Vassilis D Tsakanikas, Vassiliki T Potsika, Domenico Palombo, Fabrizio Montecucco, Dimitrios I Fotiadis, A Machine Learning Model for the Identification of High risk Carotid Atherosclerotic Plaques., 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2021.
5. **Vassiliki I Kigka**, Savvas Kyriakidis, Antonis Sakellarios, Vassiliki Potsika, Vasilis Tsakanikas, Dimitra Loggitsi, Lampros K Michalis, Dimitrios I Fotiadis, Three-Dimensional Reconstruction of Carotid Arteries Using Computed Tomography Angiography, 8th European Medical and Biological Engineering Conference: Proceedings of the EMBEC 2020, November 29–December 3, 2020 Portorož, Slovenia, 2021.
6. **Vassiliki I Kigka**, Antonis I Sakellarios, Vassilis D Tsakanikas, Vassiliki T Potsika, Igor Koncar, Dimitrios I Fotiadis, Detection of asymptomatic carotid artery stenosis through machine learning, 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2022.
7. **Vassiliki Kigka**, Antonis Sakellarios, Vasilis D. Tsakanikas, Vassiliki Potsika, Igor Koncar, Dimitrios I. Fotiadis, Early Diagnosis of Carotid Artery Disease based on non-imaging data, IEEE EMBS International Conference on Data Science and Engineering in Healthcare, Medicine & Biology, 2023 (accepted)
8. P. Siogkas, **V. Kigka**, A. Sakellarios, T. Exarchos, D. Fotiadis, "Analysis of Coronary Computed Tomography Angiography for 3D Reconstruction of Arterial Trees and Plaque Detection", IEEE International Conference on Biomedical and Health Informatics (BHI), Orlando, Florida, 16-19 February, 2017.
9. SC Tassi, **Vassiliki Kigka**, Panagiotis Siogkas, Silvia Rocchiccioli, Gualtriero Pelosi, Dimitrios I Fotiadis, Antonis I Sakellarios, Graph-guided gaussian process-based diagnosis of cvd severity with uncertainty measures, Proccedings of the 45th EMBC Conference, 2023.
10. A. Sakellarios, G. Rigas, **V. Kigka**, P. Siogkas, P. Tsompou, G. Karanasiou , T. P. Exarchos, I. Andrikos, N. Tachos, Gualtiero Pelosi, Oberdan Parodi, Dimitrios I. Fotiadis, "SMARTool: A Tool for Clinical Decision Support for the Management of Patients with Coronary Artery Disease Based

on Modeling of Atherosclerotic Plaque Process", International Conference of the IEEE Engineering in Medicine and Biology Society , Korea, 11-15 July 2017

11. Antonis I Sakellarios, Panagiota Tsompou, **Vassiliki Kigka**, Gianna Karanasiou, Konstantina Tsarapatsani, Savvas Kyriakidis, Georgia Karanasiou, Panagiotis Siogkas, Sotiris Nikopoulos, Silvia Rocchiccioli, Gualtiero Pelosi, Lampros K Michalis, Dimitrios I Fotiadis, A proof-of-concept study for the prediction of the de-novo atherosclerotic plaque development using finite elements*, 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2021.
12. Antonis I Sakellarios, Panagiota Tsompou, Panagiotis Siogkas, **Vassiliki Kigka**, Ioannis Andrikos, Nikolaos Tachos, Elena Georga, Savvas Kyriakidis, Silvia Rocchiccioli, Gualtriero Pelosi, Dimitrios I Fotiadis, Predictive models of coronary artery disease based on computational modeling: the SMARTool system, 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2019.
13. Dimitrios Pleouras, Antonis I Sakellarios, Savvas Kyriakidis, **Vassiliki Kigka**, Panagiotis Siogkas, Panagiota Tsompou, Nikolaos Tachos, Elena Georga, Ioannis Andrikos, Silvia Rocchiccioli, Gualtriero Pelosi, Lampros K Michalis, Dimitrios I Fotiadis, A computational multi-level atherosclerotic plaque growth model for coronary arteries, 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2019.
14. Michalis Mantzaris, Vassiliki Potsika, Panagiotis Siogkas, **Vassiliki Kigka**, Vasileios Pezoulas, Ioannis Pappas, Themis Exarchos, Igor Koncar, Jaroslav Pelisek, Evangelos Andreakos, Dimitrios Fotiadis, “A Multimodal Advanced Approach for the Stratification of Carotid Artery Disease”, IEEE 19th International Conference on Bioinformatics and Bioengineering (BIBE), 2019.
15. A. I. Sakellarios, N. Tachos, E. Georga, G. Rigas, **V. I. Kigka**, P. Siogkas, S. Kyriakidis, G. Karanasiou, P. Tsompou, I. Andrikos, S. Rocchiccioli, G. Pelosi, O. Parodi, D. I. Fotiadis, "A novel concept of the management of coronary artery disease patients based on machine learning risk stratification and computational biomechanics ", World Congress on Medical Physics & Biomedical Engineering (IUPESM), Prague, Czech Republic, 3-8 June, 2018.
16. I. O. Andrikos, A. I Sakellarios, P. K. Siogkas, P. I. Tsompou, **V. I. Kigka**, L. K. Michalis, D. I. Fotiadis, "A Novel Method for 3D Reconstruction of Coronary Bifurcation Using Quantitative Coronary Angiography", World Congress on Medical Physics & Biomedical Engineering (IUPESM), Prague, Czech Republic, 3-8 June, 2018.

17. P. I. Tsompou, A. I. Sakellarios, P. K. Siogkas, I. O. Andrikos, **V. I. Kigka**, P. A. Lemos, L. K. Michalis, D. I. Fotiadis, "Comparison of 3D reconstruction methods based on different cardiovascular imaging: a study of multimodality reconstruction method", 40th International Engineering in Medicine and Biology Conference (EMBC), Honolulu, Hawaii, July 17-21, 2018.

18. A. Sakellarios, P. Siogkas, E. Georga, N. Tachos, **V. I. Kigka**, P. Tsompou, I. Andrikos, G. S. Karanasiou, S. Rocchiccioli, J. Correia, G. Pelosi, P. Stofella, N. Filipovic, O. Parodi, D. I. Fotiadis, "A Clinical Decision Support Platform for the Risk Stratification, Diagnosis, and Prediction of Coronary Artery Disease Evolution", 40th International Engineering in Medicine and Biology Conference (EMBC), Honolulu, Hawaii, July 17-21, 2018.

19. Ioannis O Andrikos, Antonis I Sakellarios, Panagiotis K Siogkas, Panagiota I Tsompou, **Vassiliki I Kigka**, Lampros K Michalis, Dimitrios I Fotiadis, "A new method for the 3D reconstruction of coronary bifurcations pre and post the angioplasty procedure using the QCA", 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2019.
20. Panagiota I Tsompou, Panagiotis K Siogkas, Antonis I Sakellarios, Ioannis O Andrikos, **Vassiliki I Kigka**, Pedro A Lemos, Lampros K Michalis, Dimitrios I Fotiadis, "A comparison of three multimodality coronary 3D reconstruction methods", 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2019.
21. Vassilis Tsakanikas, Panagiotis Siogkas, Michalis Mantzaris, Vassiliki Potsika, **Vassiliki Kigka**, Themis Exarchos, Igor Koncar, Marija Jovanović, Aleksandra Vujčić, Stefan Dučić, Jaroslav Pelisek, Dimitrios Fotiadis, "A deep learning oriented method for automated 3D reconstruction of carotid arterial trees from MR imaging", 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2020.
22. Dimitris Pleouras, Antonis Sakellarios, Georgia Karanasiou, Savvas Kyriakidis, Panagiota Tsompou, **Vassiliki Kigka**, Dimitrios I Fotiadis, "Atherosclerotic Plaque Growth Prediction in Coronary Arteries using a Computational Multi-level Model: The Effect of Diabetes", 19th International Conference on Bioinformatics and Bioengineering (BIBE), 2019.
23. Panagiota I Tsompou, Ioannis O Andrikos, Georgia S Karanasiou, Antonis I Sakellarios, Nikolaos Tsigkas, **Vassiliki I Kigka**, Savvas Kyriakidis, Lampros K Michalis, Dimitrios I Fotiadis, Validation study of a novel method for the 3D reconstruction of coronary bifurcations, 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2020.

Book Chapters

1. **V. I. Kigka**, T. Exarchos, G. Rigas, A. Sakellarios, P. Siogkas, L.K. Michalis and D.I. Fotiadis, "IVUS tracking: Advantages and Disadvantage s of Intravascular Ultrasound in the Detection of Artery Geometrical Features and Plaque Type Morphology", 'Handbook of Speckle Filtering and Tracking in Cardiovascular Ultrasound Imaging and Video, IET, 2017, Editors: Christos P. Loizou, Constantinos S. Pattichis, Jan D'hooge.
2. A. Sakellarios, G. Karanasiou, P. Siogkas, **V. I. Kigka**, T. Exarchos, G. Rigas, L. K. Michalis, D. I. Fotiadis,"Available Computational Techniques to Model Atherosclerotic Plaque Progression Implementing a Multi-Level Approach", Computational Biomechanics for Medicine, Springer, Cham, 2017, Editors: Karol Miller, Poul Nielsen.
3. E. I. Georga, N. S. Tachos, A. I. Sakellarios, **V. I. Kigka**, T. P. Exarchos, G. Pelosi, O. Parodi, L. K. Michalis, D. I. Fotiadis, "Artificial Intelligence and Data Mining Methods for Cardiovascular Risk Prediction", Cardiovascular Computing- Methodologies and Clinical Applications, Series in BioEngineering , Springer, Editors: Golemati Spyretta, Nikita Konstantina.

Short CV

Vassiliki I. Kigka was born in Ioannina, Greece, in 1992. She received the Diploma degree (diploma of Computer Engineer- 5 years equivalent to Meng.) in Computer Engineering and Informatics from the University of Patras, Greece, in 2015. She holds a Master of science in Advanced Materials with Specialisation in Biomaterials & Biomedical Engineering from the Department of Material Science, University of Ioannina, Greece. She received Medical Doctor degree in July 2023, from Medical school, University of Ioannina, Greece. In May 2022, she obtained a scholarship from the National Scholarships' Foundation (IKY, Greece) for the finalization of her PhD thesis, entitled Predictive Modeling of atheromatic plaque growth. She is currently a PhD candidate in the Department of Materials Science and Engineering of University of Ioannina.

Since 2016 she is a research assistant at the Unit of Medical Technology and Intelligent Information Systems at the Department of Computer Science of the University of Ioannina and has participated in various European projects. Her research interests include medical image processing, medical data mining and biomedical engineering.