



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ
ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ
ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

**ΠΡΟΗΓΜΕΝΕΣ ΤΕΧΝΙΚΕΣ ΑΝΑΛΥΣΗΣ ΚΑΙ ΕΠΕΞΕΡΓΑΣΙΑΣ
ΚΑΡΔΙΟΛΟΓΙΚΩΝ ΔΕΔΟΜΕΝΩΝ**

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

ΧΡΗΣΤΟΥ Π. ΜΑΥΡΙΔΗ

Αθήνα, Δεκέμβριος, 2025



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ
ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ
ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

**ΠΡΟΗΓΜΕΝΕΣ ΤΕΧΝΙΚΕΣ ΑΝΑΛΥΣΗΣ ΚΑΙ ΕΠΕΞΕΡΓΑΣΙΑΣ
ΚΑΡΔΙΟΛΟΓΙΚΩΝ ΔΕΔΟΜΕΝΩΝ**

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ

ΧΡΗΣΤΟΥ Π. ΜΑΥΡΙΔΗ

Συμβουλευτική Επιτροπή: Ματσόπουλος Γεώργιος, Καθηγητής Η.Μ.Μ.Υ. Ε.Μ.Π.

Ουζούνoglou Νικόλαος, Ομ. Καθηγητής Η.Μ.Μ.Υ. Ε.Μ.Π.

Ασβεστάς Παντελεήμων Καθηγητής Τ.Μ.Β. ΠΑ.Δ.Α

Εγκρίθηκε από την επταμελή εξεταστική επιτροπή την 17^η Δεκεμβρίου 2025.

Ματσόπουλος Γεώργιος
Καθηγητής Ε.Μ.Π

Ουζούνoglou Νικόλαος
Ομ. Καθηγητής Ε.Μ.Π

Ασβεστάς Παντελεήμων
Καθηγητής ΠΑ.Δ.Α

Τσανάκας Παναγιώτης
Καθηγητής Ε.Μ.Π

Παναγόπουλος Αθανάσιος
Καθηγητής Ε.Μ.Π

Χριστοφόρου Ευάγγελος
Καθηγητής Ε.Μ.Π

Καλατζής Ιωάννης
Καθηγητής ΠΑ.Δ.Α

Αθήνα, Δεκέμβριος, 2025

.....

Χρήστος Π. Μαυρίδης

Διδάκτωρ Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright © Χρήστου Π. Μαυρίδη, 2025

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν στη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται στο συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευτεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Περίληψη

Η απαίτηση για έγκαιρη πρόληψη και διάγνωση βασικών καρδιολογικών παθήσεων του ανθρώπου, γέννησε την ανάγκη ανάπτυξης προηγμένων τεχνικών ανάλυσης και επεξεργασίας καρδιολογικών δεδομένων με στόχο την πρόβλεψη καρδιολογικών ασθενειών προτού εμφανιστούν σοβαρά συμπτώματα καθώς και την έγκαιρη διάγνωση για την βελτίωση της επιβίωσης και μείωσης της ανάγκης για επεμβάσεις ή επεμβατικές διαδικασίες. Εξίσου σημαντικό ρόλο στην ανάπτυξη προηγμένων τεχνικών ανάλυσης και επεξεργασίας, έπαιξε η αναγκαιότητα για αυξημένη ακρίβεια στη διάγνωση και ανάλυση του καρδιολογικού προφίλ ενός ασθενή, μειώνοντας το περιθώριο σφάλματος και εντοπίζοντας λεπτομέρειες που μπορεί να μην είναι ορατές με πιο συμβατικές μεθόδους.

Η παρούσα διδακτορική διατριβή επικεντρώνεται στην επεξεργασία καρδιολογικών δεδομένων, δηλαδή την ανάλυση εικόνων υψηλής ανάλυσης που προκύπτουν από Υπολογιστική Τομογραφία, CT, με στόχο την εξαγωγή χρήσιμων πληροφοριών για τη δομή και τη μορφολογία της καρδιάς. Στα πλαίσια της διατριβής, γίνεται ανάπτυξη και εφαρμογή εξειδικευμένων αλγορίθμων που βελτιώνουν την ακρίβεια και την αποτελεσματικότητα της τμηματοποίησης των εικόνων αυτών, καθώς και των δεδομένων που εξάγονται.

Αρχικά πραγματοποιείται εισαγωγή στην ανατομία και στις απεικονιστικές μεθόδους της καρδιάς. Η μελέτη της ανατομίας και των μεθόδων απεικόνισης της καρδιάς αποτελεί θεμελιώδη τομέα στην καρδιολογία και στην ιατρική απεικόνιση, καθώς προσφέρει σημαντικές πληροφορίες για τη δομή, τη λειτουργία και την υγεία της καρδιάς. Παράλληλα παρουσιάζονται και αναλύονται οι Μέθοδοι Απεικόνισης και Διαγνωστικής Αξιολόγησης της καρδιάς καθώς είναι ζωτικής σημασίας για την κατανόηση της δομής, της λειτουργίας και της υγείας της.

Στη συνέχεια παρουσιάζεται το σύστημα προ-επεξεργασίας και τμηματοποίησης εικόνων της αορτής καθώς είναι μια κρίσιμη διαδικασία στην καρδιολογία που επιτρέπει την ακριβή ανάλυση και αξιολόγηση της αορτής για τον εντοπισμό παθολογικών καταστάσεων, όπως ανευρύσματα και αθηρωματικές πλάκες. Συνδυάζοντας τεχνικές μηχανικής μάθησης και επεξεργασίας εικόνας, το σύστημα αυτό ενισχύει την ακρίβεια και την αποδοτικότητα των διαγνωστικών διαδικασιών.

Ακολούθως παρουσιάζεται η διαδικασία τμηματοποίησης της Αορτής σε 3Δ Δεδομένα Αξονικού Τομογράφου Συνδυάζοντας Τεχνικές Επεξεργασίας Εικόνας και Μηχανικής Μάθησης η οποία είναι εξαιρετικά χρήσιμη στην κλινική πράξη για τη διάγνωση πολυάριθμων παθολογιών όπως ο διαχωρισμός και το ανεύρυσμα της αορτής. Πιο συγκεκριμένα, υλοποιείται μια νέα πλήρως αυτοματοποιημένη μέθοδος 3Δ τμηματοποίησης σε δεδομένα ασθενών, η οποία χρησιμοποιεί αρχικά τη διαδικασία της κατωφλίωσης με μια σταθερή τιμή άνω και κάτω

κατωφλίου T σε όλη την εικόνα και η οποία ακολουθείται από μια προσέγγιση ταξινόμησης βασισμένη σε ένα δίκτυο τυχαίων πεδίων Markov (MRF). Η προτεινόμενη μεθοδολογία πέτυχε ανώτερη απόδοση τμηματοποίησης σε σύγκριση με όλες τις υπόλοιπες κλασσικές τεχνικές τμηματοποίησης, όσον αφορά την ακρίβεια του εξαγόμενου 3D μοντέλου αορτής. Αυτό είχε ως αποτέλεσμα, η προτεινόμενη μέθοδος τμηματοποίησης να μπορεί να χρησιμοποιηθεί στην κλινική πράξη, όπως στον σχεδιασμό και την αξιολόγηση της θεραπείας, καθώς μπορεί να επιταχύνει την αξιολόγηση των δεδομένων ιατρικής απεικόνισης, η οποία είναι συνήθως μια χρονοβόρα και κουραστική διαδικασία.

Στη συνέχεια παρέχουμε μια αναλυτική Συγκριτική Μελέτη, Αρχιτεκτονικών Βαθιάς Μάθησης για την Αυτόματη Τμηματοποίηση της Κοιλιακής Αορτής σε 3D Δεδομένα Αξονικού Τομογράφου. Η συγκεκριμένη μελέτη αξιολογεί τόσο ποιοτικά όσο και ποσοτικά τα εξαγόμενα 3D μοντέλα τα οποία εξήχθησαν από την αυτόματη τμηματοποίηση, χρησιμοποιώντας δεδομένα υγιών και ασθενών ατόμων. Τα συγκριτικά αποτελέσματα έδειξαν ότι η αρχιτεκτονική βαθιάς μάθησης U-Net, πέτυχε την υψηλότερη ακρίβεια τμηματοποίησης μεταξύ των άλλων τριών αρχιτεκτονικών που δοκιμάστηκαν. Επιπλέον η ικανότητα του U-Net να εξισορροπεί την υπολογιστική απόδοση μέσω της ακρίβειας τμηματοποίησης το καθιστά πολύτιμο εργαλείο σε κλινικά περιβάλλοντα όπου η έγκαιρη και ακριβής ανάλυση εικόνας είναι ζωτικής σημασίας.

Τέλος, στην παρούσα διδακτορική διατριβή, παρέχουμε βελτιωμένο το μοντέλο τμηματοποίησης μέσω των προαναφερθέντων Αρχιτεκτονικών Βαθιάς Μάθησης, αξιοποιώντας την ιδιότητα της προ-εκπαίδευσης (pre-training) και χρησιμοποιώντας τη μέθοδο η οποία βασίζεται στη μάσκα καθοδηγούμενη από την φωτεινότητα (Intensity Guided Masking, IGM) κατά την οποία αποκρύπτονται οι περιοχές της CT εικόνας εντός συγκεκριμένων εύρους εντάσεων φωτεινότητας ώστε το βελτιωμένο μοντέλο να καλείται να προβλέψει ή να ανακατασκευάσει την καλυμμένη περιοχή. Η προτεινόμενη βελτιωμένη μέθοδος αξιολογήθηκε τόσο ποιοτικά όσο και ποσοτικά, ξεπερνώντας τόσο τις προηγμένες εποπτευόμενες μεθόδους αναφοράς όσο και άλλες μεθόδους που βασίζονται σε προ-εκπαίδευση.

Λέξεις Κλειδιά: Τμηματοποίηση αορτής, Υπολογιστική τομογραφία, Μηχανική μάθηση, Βαθιά μάθηση, Επεξεργασία εικόνας, 3D μοντελοποίηση, Κοιλιακή αορτή, Ανάλυση ιατρικής εικόνας

Abstract

The demand for timely prevention and diagnosis of key human cardiovascular diseases has led to the need for developing advanced techniques for analyzing and processing cardiological data, in order to predict heart diseases before serious symptoms emerge, as well as enabling early diagnosis to improve survival rates and reduce the need for surgeries or invasive procedures. Equally important in advancing these analytical techniques has been the need for increased accuracy in diagnosing and analyzing a patient's cardiological profile, minimizing the margin of error and identifying details that may not be visible through more conventional methods.

This PhD thesis focuses on the cardiological data processing, specifically the analysis of high-resolution images derived from Computed Tomography (CT) scans, with the goal of extracting valuable information about the structure and morphology of the heart. This study involves the development and application of specialized algorithms that enhance the accuracy and effectiveness of image segmentation, as well as the data extracted from them.

Initially, an introduction of the heart anatomy and imaging methods is provided. Studying the anatomy and imaging techniques of the heart is a fundamental area in cardiology and medical imaging, as it offers crucial insights into the heart's structure, function, and health. Simultaneously, the Imaging and Diagnostic Evaluation Methods of the heart are presented and analyzed, as they are vital for understanding its structure, function, and health.

Following that, in this PhD thesis, we present an image pre-processing and segmentation system for the aorta which is a critical process in cardiology that allows precise analysis and evaluation of the aorta to identify pathological conditions, such as aneurysms and atherosclerotic plaques. By combining machine learning and image processing techniques, this system enhances the accuracy and efficiency of diagnostic procedures.

Afterwards, the process of segmenting the Aorta in 3D CT data is presented, combining Image Processing and Machine Learning Techniques. This method is highly useful in clinical practice for diagnosing various pathologies, such as aortic dissection and aneurysms. Specifically, a new fully automated 3D segmentation method is implemented on patient data. Initially a thresholding process with a fixed upper and lower threshold T across the image is being used, followed by a classification approach based on a Markov Random Field (MRF) network. The proposed methodology achieved superior segmentation performance compared to all other classical segmentation techniques regarding the accuracy of the extracted 3D aorta model. As a result, the proposed segmentation method can be applied in clinical practice, such as in treatment planning and evaluation, as it can accelerate the assessment of medical imaging data, which is usually a time-consuming and labor-intensive process.

Thereafter, this doctoral dissertation provides a detailed Comparative Study of Deep Learning Architectures for the Automatic Segmentation of the Abdominal Aorta in 3D CT data. This study evaluates qualitatively and quantitatively the 3D models results obtained from automatic segmentation using data from both healthy and patient subjects. Comparative results showed that the U-Net deep learning architecture achieved the highest segmentation accuracy among the other three architectures tested. Moreover, U-Net's ability to balance computational efficiency with segmentation accuracy makes it a valuable tool in clinical settings, where timely and precise image analysis is crucial.

Finally, in the present doctoral dissertation, we provide an improved segmentation model through the aforementioned Deep Learning Architectures, leveraging the property of pre-training and employing the Intensity Guided Masking (IGM) method, in which regions of the CT image within specific intensity ranges are masked so that the enhanced model is tasked with predicting or reconstructing the covered region. The proposed enhanced method was evaluated both qualitatively and quantitatively, surpassing both advanced supervised benchmark methods and other approaches based on pre-training.

Keywords: Aorta segmentation, Computed tomography, Machine learning, Deep learning, Image processing, 3D modeling, Abdominal Aorta, Medical Image Analysis

Ευχαριστίες

Η παρούσα διδακτορική διατριβή είναι το αποτέλεσμα της ερευνητικής μου προσπάθειας στο Εργαστήριο Βιοϊατρικής Μηχανικής του Τομέα Συστημάτων Μετάδοσης Πληροφορίας και Τεχνολογίας Υλικών της Σχολής Ηλεκτρολόγων Μηχανικών και Μηχανικών Ηλεκτρονικών Υπολογιστών του Εθνικού Μετσόβιου Πολυτεχνείου. Κατά τη διάρκεια αυτής της διαδρομής, είχα την τιμή και την τύχη να συνεργαστώ με άτομα που αποτέλεσαν θεμελιώδεις υποστηρικτές στην προσπάθειά μου για την ολοκλήρωση του ερευνητικού αυτού έργου. Θα ήθελα, σε αυτό το σημείο, να εκφράσω την ειλικρινή μου ευγνωμοσύνη προς όλους αυτούς για τη διαρκή τους στήριξη και καθοδήγηση.

Αρχικά, θα ήθελα να εκφράσω την ιδιαίτερη ευγνωμοσύνη μου προς τον κ. Γεώργιο Ματσόπουλο, Καθηγητή του Εθνικού Μετσόβιου Πολυτεχνείου, για την ευκαιρία που μου παρείχε να βιώσω μια μοναδική εμπειρία ή οποία παρά τις απαιτήσεις και τις προκλήσεις, αποδείχθηκε πλούσια σε πνευματικές και ψυχικές ανταμοιβές.

Χρυστάω ένα θερμό ευχαριστώ, στα υπόλοιπα μέλη της τριμελούς επιτροπής, συγκεκριμένα τον Ομ. Καθηγητή του Εθνικού Μετσόβιου Πολυτεχνείου κ. Νικόλαο Ουζούνογλου και τον Καθηγητή του Πανεπιστημίου Δυτικής Αττικής κ. Παντελεήμονα Ασβεστά.

Επίσης ένα θερμό ευχαριστώ στον Δρ. Θεόδωρο Οικονομόπουλο και τον γιατρό Δρ. Γεώργιο Ν. Μπενέτο για τη σημαντική συνεισφορά τους στην περάτωση της έρευνας του τετάρτου κεφαλαίου καθώς και τους Δρ. Ιωάννη Κάκκο και τον Υποψήφιο Διδάκτορα Θεόδωρο Π. Βαγενά για την πολύτιμη συμβολή τους στην ολοκλήρωση της έρευνας του πέμπτου και έκτου κεφαλαίου.

Τέλος, θα ήθελα να εκφράσω την πιο βαθιά μου ευγνωμοσύνη προς τον πατέρα μου και τη μητέρα μου, Παντελή και Εύχαρις, που με περιβάλλουν με αγάπη, κατανόηση και αδιάκοπη στήριξη. Μέσα από την καθοδήγηση και τα ψυχικά και πνευματικά εφόδια που μου έχουν προσφέρει, μου έδωσαν τη δύναμη να ανταποκρίνομαι στις απαιτήσεις της ζωής. Είναι πάντα παρόντες, με τη στήριξή τους να με συνοδεύει σε κάθε σημαντική στιγμή της ζωής μου.

ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

ΠΕΡΙΛΗΨΗ.....	5
ABSTRACT	8
ΕΥΧΑΡΙΣΤΙΕΣ	11
ΕΥΡΕΤΗΡΙΟ ΣΧΗΜΑΤΩΝ	15
ΕΥΡΕΤΗΡΙΟ ΠΙΝΑΚΩΝ	18
ΣΥΝΤΟΜΟΓΡΑΦΙΕΣ	18
ΚΕΦΑΛΑΙΟ 1: ΕΙΣΑΓΩΓΗ	21
1.1 Κίνητρο διδακτορικής διατριβής	21
1.2 Σκοπός διδακτορικής διατριβής	22
1.3 Δομή διδακτορικής διατριβής	23
ΚΕΦΑΛΑΙΟ 2: ΑΝΑΤΟΜΙΑ ΚΑΙ ΜΕΘΟΔΟΙ ΑΠΕΙΚΟΝΙΣΗΣ ΚΑΡΔΙΑΣ... 25	
2.1 Εισαγωγή στη Ανατομία της Καρδιάς	25
2.1.1 Βασική Μορφολογία	25
2.1.2 Σωματική κυκλοφορία	28
2.1.3 Δυναμικές Διαδικασίες Καρδιάς	30
2.1.4 Παθολογικές Καταστάσεις	31
2.2 Ανατομία και Λειτουργία της Αορτής	40
2.2.1 Εισαγωγή στην Ανατομία της Αορτής	41
2.2.2 Ιστολογία.....	42
2.2.3 Παθολογικές Καταστάσεις της Αορτής	43
2.3 Εισαγωγή στις Μεθόδους Απεικόνισης και Διαγνωστικής Αξιολόγησης.....	44
2.3.1 Ηλεκτροκαρδιογράφημα	44
2.3.2 Υπερηχοκαρδιογράφημα	47
2.3.3 Σπινθηρογράφημα Αιμάτωσης Μυοκαρδίου	48
2.3.4 Μαγνητική Τομογραφία	50
2.3.5 Αξονική Τομογραφία Καρδιάς (Αξονική Στεφανιογραφία)	52
ΚΕΦΑΛΑΙΟ 3: ΣΥΣΤΗΜΑ ΠΡΟ-ΕΠΕΞΕΡΓΑΣΙΑΣ ΚΑΙ ΤΜΗΜΑΤΟΠΟΙΗΣΗΣ ΕΙΚΟΝΩΝ	56
3.1 Εισαγωγή στα Βασικά Βήματα Προ-επεξεργασίας Εικόνας	56
3.1.1 Ισοστάθμιση Ιστογράμματος.....	56
3.1.2 Αφαίρεση Θορύβου και εξομάλυνση ακμών	63

3.1.3	Διορθώσεις Σχηματισμού	64
3.1.4	Ενίσχυση Ακμών	65
3.1.5	Ευθυγράμμιση Εικόνας.....	66
3.1.6	Ενίσχυση Δεδομένων.....	67
3.2	Εισαγωγή στις Μεθόδους Τμηματοποίησης Εικόνας.....	70
3.2.1	Κατωφλίωση (Thresholding)	70
3.2.2	Ανίχνευση Ακμών.....	71
3.2.3	Μέθοδοι Βασισμένες σε Περιοχές (Region Growing).....	75
3.2.4	Ενεργά Περιγράμματα (Active Contours).....	77
3.2.5	Μέθοδοι Ομαδοποίησης (K-Means)	86
3.2.6	Watershed	89
3.2.7	Μαρκοβιανό Τυχαίο Πεδίο (Markov Random Field).....	92
3.2.8	Μέθοδοι Μηχανικής & Βαθιάς Μάθησης.....	96

ΚΕΦΑΛΑΙΟ 4: ΤΜΗΜΑΤΟΠΟΙΗΣΗ ΤΗΣ ΑΟΡΤΗΣ ΣΕ 3Δ ΔΕΔΟΜΕΝΑ ΑΞΟΝΙΚΟΥ ΤΟΜΟΓΡΑΦΟΥ ΣΥΝΔΥΑΖΟΝΤΑΣ ΤΕΧΝΙΚΕΣ ΕΠΕΞΕΡΓΑΣΙΑΣ ΕΙΚΟΝΑΣ ΚΑΙ ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ..... 114

4.1	Εισαγωγή	114
4.2	Μεθοδολογία	117
4.2.1	Συλλογή Δεδομένων	118
4.2.2	Αντιστοίχιση Ιστογράμματος	119
4.2.3	Σχήμα Αυτόματης 3Δ Τμηματοποίησης	119
4.2.4	Κατωφλίωση (Thresholding)	120
4.2.5	Μαρκοβιανό Τυχαίο Πεδίο (Markov Random Field).....	121
4.2.6	Μορφολογικές Λειτουργίες (Εξομάλυνση)	126
4.2.7	3Δ Ανακατασκευή	127
4.2.8	Συγκρινόμενες Τεχνικές Τμηματοποίησης	128
4.2.9	Ευθυγράμμιση Μοντέλου	130
4.2.10	Υπολογισμός του Δείκτη Jaccard	131
4.2.11	Περιβάλλον υλοποίησης και δοκιμών.....	132
4.3	Αποτελέσματα.....	132
4.3.1	Ποιοτική Αξιολόγηση	132
4.3.2	Ποσοτική Αξιολόγηση.....	134
4.4	Συζήτηση	136
4.5	Συμπεράσματα.....	139

ΚΕΦΑΛΑΙΟ 5: ΣΥΓΚΡΙΤΙΚΗ ΜΕΛΕΤΗ ΑΡΧΙΤΕΚΤΟΝΙΚΩΝ ΒΑΘΙΑΣ ΜΑΘΗΣΗΣ ΓΙΑ ΤΗΝ ΑΥΤΟΜΑΤΗ ΤΜΗΜΑΤΟΠΟΙΗΣΗ ΤΗΣ ΚΟΙΛΙΑΚΗΣ ΑΟΡΤΗΣ ΣΕ 3Δ ΔΕΔΟΜΕΝΑ ΑΞΟΝΙΚΟΥ ΤΟΜΟΓΡΑΦΟΥ..... 140

5.1	Εισαγωγή	140
5.2	Μεθοδολογία	143
5.2.1	Συλλογή Δεδομένων	143
5.2.2	Προ-Επεξεργασία	144
5.2.3	Τμηματοποίηση μέσω Βαθιάς Μάθησης.....	144
5.2.4	Υλοποίηση και Υπολογιστικό Κόστος	153

5.2.5	3Δ Ανακατασκευή Πλέγματος.....	154
5.3	Αποτελέσματα.....	155
5.3.1	Ποσοτική Αξιολόγηση.....	155
5.3.2	Ποιοτική Αξιολόγηση	156
5.3.3	Αξιολόγηση 3Δ Ανακατασκευής.....	158
5.4	Συζήτηση	159
5.5	Συμπεράσματα	162

ΚΕΦΑΛΑΙΟ 6: ΑΥΤΟ-ΕΠΙΒΛΕΠΟΜΕΝΗ ΠΡΟ-ΕΚΠΑΙΔΕΥΣΗ ΜΕ ΜΑΣΚΑ ΚΑΘΟΔΗΓΟΥΜΕΝΗ ΑΠΟ ΤΗΝ ΦΩΤΕΙΝΟΤΗΤΑ ΓΙΑ ΒΕΛΤΙΩΜΕΝΗ ΤΜΗΜΑΤΟΠΟΙΗΣΗ ΤΗΣ ΑΟΡΤΗΣ ΣΕ ΑΞΟΝΙΚΗ ΤΟΜΟΓΡΑΦΙΑ 164

6.1	Εισαγωγή	164
6.2	Μεθοδολογία	165
6.2.1	Εισαγωγή	165
6.2.2	Προ-εκπαίδευση SSL	166
6.3	Αποτελέσματα.....	171
6.3.1	Ποιοτική Αξιολόγηση	172
6.3.2	Ποσοτική Αξιολόγηση.....	172
6.4	Συζήτηση	174
6.5	Συμπεράσματα	176

ΚΕΦΑΛΑΙΟ 7: ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΜΕΛΛΟΝΤΙΚΕΣ ΠΡΟΟΠΤΙΚΕΣ 177

BIBΛΙΟΓΡΑΦΙΑ.....	180
-------------------	-----

Ευρετήριο Σχημάτων

Εικόνα 2.1:	Ανατομία της καρδιάς [11]	26
Εικόνα 2.2:	Κόλποι και κοιλίες καρδιάς [11]	26
Εικόνα 2.3:	Διαφράγματα, κολποκοιλιακές και ημιτελικές βαλβίδες καρδιάς [14]	27
Εικόνα 2.4:	Στρώματα καρδιάς [16]	27
Εικόνα 2.5:	Ροή αίματος [16]	28
Εικόνα 2.6:	Αορτή και αορτική βαλβίδα [20].....	29
Εικόνα 2.7:	Τριχοειδή αγγεία [21]	29
Εικόνα 2.8:	Η καρδιά συμπεριφέρεται ως μια αντλία [22]	30
Εικόνα 2.9:	Ανατομία αορτής [51].....	41
Εικόνα 2.10:	Στρώματα αορτής [57].....	42
Εικόνα 2.11:	Ανευρύσματα αορτής [59]	43

Εικόνα 2.12: Το Ηλεκτροκαρδιογράφημα [69].....	45
Εικόνα 2.13: Το Υπερηχοκαρδιογράφημα [72].....	47
Εικόνα 2.14: Το Σπινθηρογράφημα Αιμάτωσης Μυοκαρδίου [81].....	49
Εικόνα 2.15: Μαγνητικός τομογράφος [85].....	50
Εικόνα 2.16: Απεικόνιση Μαγνητικής τομογραφίας καρδιάς [87]	51
Εικόνα 2.17: Αξονικός τομογράφος [88]	52
Εικόνα 2.18: Διαδικασία σάρωσης αξονικού τομογράφου [88]	53
Εικόνα 2.19: Αξονική τομογραφία καρδιάς	53
Εικόνα 3.1: Συνάρτηση μετασχηματισμού $T(g)$	58
Εικόνα 3.2: (α) Μορφές της $T(g)$ για διάφορες τιμές του p , (β) Αποτελέσματα του μετασχηματισμού δύναμης για διάφορες τιμές του p	60
Εικόνα 3.3: Καμπύλες λογαριθμικού μετασχηματισμού (α) για $b=45.886$, (β) $b = 255 \ln 1 + 255 \cdot a$	60
Εικόνα 3.4: Εφαρμογή του λογαριθμικού μετασχηματισμού (α) Αρχική εικόνα, (β) Τελική εικόνα με $a=1$	61
Εικόνα 3.5: Καμπύλες εκθετικού μετασχηματισμού για $b = 255(\ln 1 + 255 \cdot a)$	61
Εικόνα 3.6: Εφαρμογή του εκθετικού μετασχηματισμού (α) Αρχική εικόνα, (β) Τελική εικόνα για $a=1$	62
Εικόνα 3.7: Μορφή της συνάρτησης για διάφορες τιμές του a	62
Εικόνα 3.8: Εφαρμογή της Γκαουσιανής συνάρτησης σφάλματος (α) Αρχική εικόνα, (β) Τελική εικόνα με $a=0.02$	63
Εικόνα 3.9: (α) Αρχική εικόνα, (β) Καθολική εικόνα, (γ) Προσαρμοστική κατωφλίωση [162]	71
Εικόνα 3.10: Εφαρμογή του μετ/μού Hough. (α) Αρχική εικόνα, (β) Μάσκα ακμών, (γ) Δύο πρώτες γραμμές με μετ/μό Hough, (δ) Πλήρες αποτέλεσμα μετ/μού Hough [166]	75
Εικόνα 3.11: Η εξέλιξη του τυπικού αλγορίθμου region-growing [169].....	76
Εικόνα 3.12: Μεταβολή ενός ενεργού περιγράμματος υπό την επίδραση της καμπυλότητας ως μοναδικής εσωτερικής δύναμης [180].....	80
Εικόνα 3.13: Region-based εξέλιξη ενεργού περιγράμματος σε εικόνα κυττάρων. (α) Αρχικό περίγραμμα, (β-γ) Ενδιάμεσες καταστάσεις, (δ) Τελική τμηματοποίηση [191].....	85
Εικόνα 3.14: Διάγραμμα ροής K-Means.....	88
Εικόνα 3.15: Γραμμές διαχωρισμού, Λεκάνες απορροής, Ελάχιστα [206]	89
Εικόνα 3.16: Διάγραμμα ροής Watershed.....	91
Εικόνα 3.17: Μη κατευθυνόμενος γράφος [217]	93
Εικόνα 3.18: Διάγραμμα ροής της διαδικασίας	95
Εικόνα 3.19: Δυσδιάστατη συνέλιξη.....	99
Εικόνα 3.20: Η λειτουργία του Max Pooling [232].....	99
Εικόνα 3.21: FCNs αρχιτεκτονική για τη σημασιολογική κατάτμηση [154].....	100
Εικόνα 3.22: FCNs αρχιτεκτονική για τη σημασιολογική τμηματοποίηση μαζί με τις συνελκτικές και αποσυγκεντρωτικές συνιστώσες [234]	101
Εικόνα 3.23: Κλιμακωμένη προσοχή μέσω εσωτερικού γινομένου [238]	103
Εικόνα 3.24: Μηχανισμός προσοχής πολλαπλών κεφαλών [238]	105
Εικόνα 3.25: Μοντέλο μεταφοράς με κωδικοποίηση θέσης [238].....	107
Εικόνα 3.26: Αρχιτεκτονική του μοντέλου κωδικοποιητή-αποκωδικοποιητή [250].....	109
Εικόνα 4.1: Επισκόπηση της προτεινόμενης μεθοδολογίας για 3Δ αυτόματη τμηματοποίηση	118

Εικόνα 4.2: Σχήμα αυτόματης 3Δ τμηματοποίησης	120
Εικόνα 4.3: Σύστημα γειτονιάς πρώτης τάξης με τις αντίστοιχες κλίκες [217]	123
Εικόνα 4.4: Σύστημα γειτονιάς δεύτερης τάξης με τις αντίστοιχες κλίκες [217]	123
Εικόνα 4.5: Ορισμός των συνόλων εκπαίδευσης, με την επιλογή αντιπροσωπευτικών περιοχών για τις τέσσερις MRF κλάσεις	125
Εικόνα 4.6: Επισκόπηση των συγκρινόμενων τεχνικών τμηματοποίησης	130
Εικόνα 4.7: Συγχώνευση 3Δ μοντέλων των εκτιμώμενων και των πραγματικών περιοχών της αορτής για τρεις ασθενείς (α)-(δ), (ε)-(η), (θ)-(λ), Μετά την εφαρμογή της προτεινόμενης μεθόδου τμηματοποίησης (α), (ε), (θ), K-Means (β), (στ), (ι), Κατωφλίωση (γ), (ζ), (κ) και Μαρκοβιανών Τυχαίων Πεδίων (δ), (η), (λ)	134
Εικόνα 4.8: Τα 3Δ μοντέλα της αορτής που έχουν προκύψει από μία μόνο περίπτωση, κατά την εφαρμογή της προτεινόμενης μεθόδου τμηματοποίησης με χρήση: (α) τριών κλάσεων, (β) τεσσάρων κλάσεων και (γ) πέντε κλάσεων.....	137
Εικόνα 4.9: Καταγεγραμμένος χρόνος επεξεργασίας για την προτεινόμενη μεθοδολογία από ένα μόνο σύνολο δεδομένων, για διάφορο αριθμό κλάσεων στο δεύτερο στάδιο της διαδικασίας τμηματοποίησης.....	138
Εικόνα 5.1: Αρχιτεκτονική U-Net	149
Εικόνα 5.2: Αρχιτεκτονική UNETR	151
Εικόνα 5.3: Αρχιτεκτονική SwinUNETR.....	152
Εικόνα 5.4: Αρχιτεκτονική SegResNet	153
Εικόνα 5.5: Εντοπισμένες περιοχές ενδιαφέροντος (αορτή), συγχωνευμένες στις αρχικές τομές των απεικονιστικών δεδομένων για δύο κλινικές περιπτώσεις (σειρές 1 και 2) του δημόσιου συνόλου δεδομένων για: (α) Αρχική εικόνα, (β) Πραγματική περιοχή ενδιαφέροντος, (γ) Μοντέλο U-Net, (δ) Μοντέλο UNETR, (ε) Μοντέλο SwinUNETR, (στ) SegResNet.	157
Εικόνα 5.6: Εντοπισμένες περιοχές ενδιαφέροντος (αορτή), συγχωνευμένες στις αρχικές τομές των απεικονιστικών δεδομένων για δύο κλινικές περιπτώσεις (σειρές 1 και 2) του ιδιωτικού συνόλου δεδομένων για: (α) Αρχική εικόνα, (β) Πραγματική περιοχή ενδιαφέροντος, (γ) Μοντέλο U-Net, (δ) Μοντέλο UNETR, (ε) Μοντέλο SwinUNETR, (στ) SegResNet.	157
Εικόνα 5.7: 3Δ συγχωνευμένα μοντέλα της εκτιμώμενης αορτής (μπλε), επικαλυπτόμενα στην πραγματική περιοχή (κοραλλί) για έξι περιπτώσεις, χρησιμοποιώντας: (α) U-Net, (β) UNETR, (γ) SwinUNETR και (δ) SegResNet.....	158
Εικόνα 6.1: Το επάνω μέρος παρουσιάζει την προτεινόμενη στρατηγική προ-εκπαίδευσης βασισμένη στην μάσκα καθοδηγούμενη από την φωτεινότητα (IGM), ενώ το κάτω μέρος απεικονίζει τη διαδικασία βελτιστοποίησης για την τμηματοποίηση.....	166
Εικόνα 6.2: Σύγκριση της προτεινόμενης και της βασικής μεθόδου στο δημόσιο σύνολο δεδομένων. Τα δεδομένα αναφοράς (Ground Truth, GT) εμφανίζονται με κόκκινο, οι προβλέψεις με πράσινο και η επικάλυψη απεικονίζεται για κάθε μοντέλο	173
Εικόνα 6.3: Σύγκριση της προτεινόμενης και της βασικής μεθόδου στο ιδιωτικό σύνολο δεδομένων. Τα δεδομένα αναφοράς εμφανίζονται με κόκκινο, οι προβλέψεις με πράσινο και η επικάλυψη απεικονίζεται για κάθε μοντέλο	174

Ευρετήριο Πινάκων

Πίνακας 3.1: Πυρήνας φιλτραρίσματος74

Πίνακας 4.1: Ο δείκτης Jaccard της πραγματικής περιοχής της αορτής, σε σύγκριση με την εκτιμώμενη περιοχή της αορτής για όλες τις συγκρινόμενες μεθόδους τμηματοποίησης. ...135

Πίνακας 4.2: Συνολικός χρόνος επεξεργασίας τμηματοποίησης για όλα τα σύνολα δεδομένων σε λεπτά 139

Πίνακας 5.1: Υπερπαράμετροι αρχιτεκτονικών.....148

Πίνακας 5.2: Αριθμός συνολικών παραμέτρων για κάθε μοντέλο.....154

Πίνακας 5.3: Απόδοση των διαφόρων μοντέλων βαθιάς μάθησης στο δημόσιο σύνολο δεδομένων155

Πίνακας 5.4: Απόδοση των διαφόρων μοντέλων βαθιάς μάθησης στο ιδιωτικό σύνολο δεδομένων156

Πίνακας 5.5: Στατιστική αξιολόγηση της απόδοσης τμηματοποίησης.....156

Πίνακας 6.1: Σύγκριση απόδοσης μεθόδων για το δημόσιο σύνολο δεδομένων171

Πίνακας 6.2: Σύγκριση απόδοσης μεθόδων για το ιδιωτικό σύνολο δεδομένων173

Πίνακας 6.3: Πειράματα για διαφορετικά μεγέθη συνόλου εκπαίδευσης.....175

Συντομογραφίες

AI	Artificial Intelligence	Τεχνητή Νοημοσύνη
ML	Machine Learning	Μηχανική Μάθηση
CNNs	Convolutional Neural Networks	Συνελικτικά Νευρωνικά Δίκτυα
RNNs	Recurrent Neural Networks	Επαναλαμβανόμενα Νευρωνικά Δίκτυα
DTI	Diffusion Tensor Imaging	Απεικόνιση Τανυστή Διάχυσης
KA	Heart Failure	Καρδιακή Ανεπάρκεια
KEAK	Left Ventricular Ejection Fraction	Κλάσμα Εξώθησης Αριστερής Κοιλίας
AF	Atrial Fibrillation	Κολπική Μαρμαρυγή
ΗΚΓ	Electrocardiogram	Ηλεκτροκαρδιογράφημα
ICD	Implantable Cardioverter-Defibrillator	Εμφυτεύσιμος Καρδιομετατροπέας-Απινιδωτής
ΜΚΠ	Cardiomyopathies	Μυοκαρδιοπάθειες
ΤΔΠ	End-Diastolic Pressure	Τελοδιαστολική Πίεση
TAVI	Transcatheter Aortic Valve Implantation	Διαδερμική Αντικατάσταση Αορτικής Βαλβίδας
MRI	Magnetic Resonance Imaging	Απεικόνιση Μαγνητικού Συντονισμού
CT	Computed Tomography	Υπολογιστική Τομογραφία
DSC	Dice Similarity Coefficient	Συντελεστής Ομοιότητας Dice
MRF	Markov Random Fields	Μαρκοβιανό Τυχαίο Πεδίο
ΕΠ	Active Contours	Ενεργά Περιγράμματα
ESF	Edge Stopping Function	Συνάρτηση Φραγής Ακμών

GAC	Geodesic Active Contour	Γεωδαιτικό Ενεργό Περίγραμμα
FCNs	Fully Convolutional Networks	Πλήρως Συνελκτικά Νευρωνικά Δίκτυα
ReLU	Rectified Linear Unit	Διορθωμένη Γραμμική Μονάδα
AAA	Abdominal Aortic Aneurysm	Ανεύρυσμα Κοιλιακής Αορτής
CTA	Computed Tomography Angiography	Αξονική Τομογραφία Αγγειογραφία
DICOM	Digital Imaging and Communications in Medicine	Ψηφιακή Απεικόνιση και Επικοινωνίες στην Ιατρική
ICM	Iterated Conditional Modes	Επαναλαμβανόμενες Λειτουργίες υπό Όρους
ROI	Region of Interest	Περιοχή Ενδιαφέροντος
ANTs	Advanced Normalization Tools	Προηγμένα Εργαλεία Κανονικοποίησης
DNNs	Deep Neural Networks	Βαθιά Νευρωνικά Δίκτυα
ILT	Intraluminal Thrombus	Ενδοαυλικός Θρόμβος
ViT	Vision Transformer	Μετασχηματιστής Όρασης
IN	Instance Normalization	Κανονικοποίηση Στιγμιότυπου
W-MSA	Window-based Multi-head Self-Attention	Αυτό-Προσοχή Πολλαπλών κεφαλών με βάση το παράθυρο
SW-MSA	Shifted Window Multi-head Self-Attention	Αυτό-Προσοχή πολλαπλών κεφαλών με μετατοπισμένο παράθυρο
ASSD	Average Symmetric Surface Distance	Μέση Συμμετρική Απόσταση Επιφάνειας
SSL	Self-Supervised Learning	Αυτο-Εποπτευόμενη Μάθηση
IGM	Intensity Guided Masking	Μάσκας Καθοδηγούμενη από την Φωτεινότητα
MIM	Masked Image Modeling	Μοντελοποίηση Εικόνας με Μάσκα
MLP	Multi-Layer Perceptron	Πολυεπίπεδο Νευρωνικό Δίκτυο
LN	Layer Normalization	Επίπεδο Κανονικοποίησης
HU	Hounfield Units	Μονάδες Χάουνσφιλντ

Κεφάλαιο 1: Εισαγωγή

1.1 Κίνητρο διδακτορικής διατριβής

Η καρδιολογία αποτελεί έναν από τους πιο κρίσιμους τομείς της ιατρικής, καθώς οι καρδιαγγειακές παθήσεις παραμένουν κύρια αιτία θνησιμότητας παγκοσμίως. Οι εξελίξεις στις τεχνολογίες δεδομένων, όπως η τεχνητή νοημοσύνη (Artificial Intelligence, AI), η μηχανική μάθηση (Machine Learning, ML) και οι τεχνικές επεξεργασίας σήματος, προσφέρουν νέα εργαλεία για την ανάλυση και πρόβλεψη αυτών των παθήσεων, καθιστώντας την έρευνα στο συγκεκριμένο πεδίο περισσότερο επίκαιρη και απαραίτητη από ποτέ. Το κεφάλαιο αυτό επικεντρώνεται στα κίνητρα της διδακτορικής διατριβής με θέμα τις Προηγμένες Τεχνικές Ανάλυσης και Επεξεργασίας Καρδιολογικών Δεδομένων, αναδεικνύοντας τη σημασία και τον αντίκτυπο της έρευνας αυτής σε διαφορετικά επίπεδα.

Η επεξεργασία και ανάλυση καρδιολογικών δεδομένων έχει εξελιχθεί σε έναν από τους σημαντικότερους τομείς της ιατρικής τεχνολογίας, καθώς η καρδιακή υγεία συνδέεται άμεσα με την ανθρώπινη ευημερία και το προσδόκιμο ζωής [1]. Ο στόχος της χρήσης προηγμένων τεχνικών ανάλυσης σε δεδομένα από καρδιολογικές μετρήσεις δεν περιορίζεται μόνο στην ακριβή διάγνωση, αλλά περιλαμβάνει και την πρόληψη, την έγκαιρη παρέμβαση, καθώς και τη μείωση του κόστους ιατρικής περίθαλψης [2]. Παρακάτω αναλύονται τα κυριότερα κίνητρα για την έρευνα και ανάπτυξη στον τομέα αυτό.

Οι καρδιολογικές παθήσεις παραμένουν μία από τις βασικές αιτίες θανάτου παγκοσμίως, καθιστώντας την έγκαιρη και ακριβή διάγνωση κρίσιμης σημασίας (World Health Organization, 2022). Οι τεχνικές τεχνητής νοημοσύνης (AI) και μηχανικής μάθησης (ML) έχουν επιτρέψει τη δημιουργία μοντέλων που αναλύουν μεγάλα σύνολα δεδομένων και εντοπίζουν μοτίβα ή ανωμαλίες με ακρίβεια που υπερβαίνει τις παραδοσιακές μεθόδους [3]. Έρευνες έχουν δείξει ότι τα μοντέλα αυτά μπορούν να εντοπίζουν δύσκολες στην ανίχνευση καταστάσεις, όπως οι αρρυθμίες και η ισχαιμική καρδιοπάθεια, βελτιώνοντας τη συνολική ακρίβεια της διάγνωσης [4].

Η ανάλυση δεδομένων σε πραγματικό χρόνο και μεγάλης κλίμακας απαιτεί αυξημένες υπολογιστικές δυνατότητες, καθώς τα καρδιολογικά δεδομένα περιλαμβάνουν συχνά μεγάλα και σύνθετα σύνολα μετρήσεων. Στην προσπάθεια για αποδοτικότερη επεξεργασία και ανάλυση αυτών των δεδομένων, η έρευνα εστιάζει στη βελτίωση των αλγορίθμων επεξεργασίας σήματος και των μεθόδων υπολογιστικής νοημοσύνης. Για παράδειγμα, οι αλγόριθμοι συνελκτικών νευρωνικών δικτύων (CNNs) και τα μοντέλα αναδρομικής νευρωνικής δικτύωσης (RNNs) έχουν χρησιμοποιηθεί ευρέως για τη βελτίωση της ανάλυσης ECG και των προγνωστικών αποτελεσμάτων [5].

Τα παραπάνω κίνητρα καθιστούν σαφές ότι οι προηγμένες τεχνικές ανάλυσης και επεξεργασίας καρδιολογικών δεδομένων έχουν τη δυνατότητα να επαναπροσδιορίσουν τον τρόπο διάγνωσης, παρακολούθησης και πρόληψης καρδιακών παθήσεων. Η επένδυση σε έρευνα και ανάπτυξη στον τομέα αυτό, επομένως, δεν έχει μόνο ακαδημαϊκή σημασία αλλά και κοινωνικό όφελος, καθώς συμβάλλει στη βελτίωση της δημόσιας υγείας και στη μείωση του οικονομικού κόστους.

1.2 Σκοπός διδακτορικής διατριβής

Για το λόγο αυτό, σκοπός της παρούσας διδακτορικής διατριβής είναι η ανάπτυξη και αξιολόγηση προηγμένων τεχνικών ανάλυσης και επεξεργασίας καρδιολογικών δεδομένων, με στόχο τη βελτίωση της ακρίβειας διάγνωσης, την παροχή εξατομικευμένων θεραπειών και τη διευκόλυνση της πρόληψης καρδιολογικών κινδύνων. Η έρευνα φιλοδοξεί να συμβάλει στη δημιουργία καινοτόμων εργαλείων που θα ενισχύσουν τη λήψη κλινικών αποφάσεων και θα βελτιώσουν την ποιότητα ζωής των ασθενών.

Για την επίτευξη αυτού του στόχου, αναπτύχθηκαν προηγμένες τεχνικές ανάλυσης δεδομένων, που αξιοποιούν και συνδυάζουν σύγχρονες μεθόδους μηχανικής, βαθιάς μάθησης και επεξεργασίας δεδομένων που αποτελούν το βασικό κίνητρο για την παρούσα έρευνα. Αρχικά υλοποιήθηκε για το σκοπό αυτό ένας πλήρως αυτοματοποιημένος αλγόριθμος τμηματοποίησης 3Δ δεδομένων ασθενών ο οποίος συνδυάζει τεχνικές επεξεργασίας και μηχανικής μάθησης προσφέροντας ανώτερη απόδοση τμηματοποίησης από τις κλασσικές μεθόδους όσον αφορά την ακρίβεια των εξαγόμενων 3Δ μοντέλων αορτής.

Επίσης λόγω και των αυξανόμενων αναγκών ανάλυσης μεγάλου όγκου δεδομένων και βελτίωσης της ακρίβειας διάγνωσης για συγκεκριμένες περιοχές της αορτής, μας οδήγησε στην εφαρμογή τεχνικών αρχιτεκτονικής βαθιάς μάθησης με χρήση νευρωνικών δικτύων. Η ικανότητα των τεχνικών αρχιτεκτονικής βαθιάς μάθησης να επιτυγχάνουν μια ισορροπία μεταξύ υπολογιστικής απόδοσης και ακρίβειας τμηματοποίησης αποτελεί ένα από τα σημαντικότερα πλεονεκτήματά τους, ειδικά σε εφαρμογές που απαιτούν την ανάλυση ιατρικών εικόνων. Στα κλινικά περιβάλλοντα, όπου η ταχύτητα και η ακρίβεια είναι ζωτικής σημασίας, αυτές οι τεχνικές μπορούν να λειτουργήσουν ως αξιόπιστα εργαλεία, προσφέροντας κρίσιμες πληροφορίες για τη λήψη αποφάσεων.

1.3 Δομή διδακτορικής διατριβής

Στην παρούσα διδακτορική διατριβή, αναπτύχθηκε και παρουσιάζεται ένα σύνολο καινοτόμων υπολογιστικών τεχνικών για την ανάλυση και επεξεργασία καρδιολογικών δεδομένων. Η έρευνα επικεντρώνεται στην ανάπτυξη μεθόδων τμηματοποίησης μέσω της επεξεργασίας εικόνων και τεχνικών μηχανικής και βαθιάς μάθησης, οι οποίες εφαρμόζονται σε δεδομένα που παράγονται από την προ επεξεργασία αυτών των εικόνων. Κύριοι στόχοι της μελέτης είναι η αύξηση του επιπέδου αυτοματοποίησης, η βελτίωση της ακρίβειας και, όπου είναι εφικτό, η μείωση του απαιτούμενου χρόνου επεξεργασίας. Για την αξιολόγηση των μεθόδων που αναπτύχθηκαν, χρησιμοποιήθηκαν δεδομένα υγείων και ασθενών ατόμων από διάφορες πηγές. Η αξιολόγηση και απόδοση των μεθόδων αυτών, εκτιμήθηκε βάσει των ποσοτικών και ποιοτικών δεικτών, σε σύγκριση με αντίστοιχους εναλλακτικούς και ευρέως διαδεδομένους αλγόριθμους.

Στο δεύτερο κεφάλαιο παρουσιάζονται συνοπτικά τα κυριότερα στοιχεία της καρδιάς, αναλύεται η ανατομία και λειτουργία της και γίνεται μια εισαγωγή στις βασικές απεικονιστικές τεχνικές που χρησιμοποιούνται στο χώρο. Αρχικά πραγματοποιείται μία εισαγωγή στην ανατομία της καρδιάς, ξεκινώντας με τη θέση και το σχήμα της, συνεχίζοντας με τους καρδιακούς τοίχους και τις βαλβίδες και καταλήγοντας στην κυκλοφορία του αίματος. Στη συνέχεια περιγράφονται οι δυναμικές διαδικασίες που πραγματοποιούνται στην καρδιά καθώς είναι το θεμέλιο για την έναρξη και τον συγχρονισμό των συστολών της. Επίσης παρουσιάζονται οι παθολογικές καταστάσεις της καρδιάς οι οποίες επιδρούν είτε στη δομή είτε στη λειτουργία της, επηρεάζοντας τη φυσιολογική κυκλοφορία του αίματος και τη γενική υγεία του οργανισμού. Έχοντας περιγράψει πλήρως το θεωρητικό πλαίσιο της καρδιάς, ακολουθεί η παρουσίαση που αφορά την ανατομία και λειτουργία της αορτής η οποία είναι και η κύρια αρτηρία του ανθρώπινου σώματος που παίζει κεντρικό ρόλο στο καρδιαγγειακό σύστημα. Τέλος, παρουσιάζονται οι βασικές απεικονιστικές τεχνικές για την οπτικοποίηση της καρδιάς, οι οποίες αποτελούν το βασικό εργαλείο για τη διάγνωση, την παρακολούθηση και τη διαχείριση των καρδιαγγειακών παθήσεων.

Το τρίτο κεφάλαιο προσφέρει στον αναγνώστη μία εισαγωγή στα βασικά βήματα προ-επεξεργασίας, τις μεθόδους τμηματοποίησης εικόνας καθώς πραγματοποιείται και αποτίμηση των αλγορίθμων τμηματοποίησης. Το κεφάλαιο αυτό εστιάζει στο στάδιο προ-επεξεργασίας εικόνας το οποίο είναι κρίσιμο στην ανάλυση και επεξεργασία εικόνων, καθώς βελτιώνει την ποιότητα των δεδομένων μέσω εξειδικευμένων τεχνικών και εξασφαλίζει καλύτερα αποτελέσματα στις επόμενες φάσεις επεξεργασίας ή ανάλυσης. Στη συνέχεια παρουσιάζονται οι μέθοδοι τμηματοποίησης, οι οποίοι είναι απαραίτητοι στη διαδικασία επεξεργασίας της εικόνας και οι οποίοι αποσκοπούν στον διαχωρισμό μιας εικόνας σε επιμέρους περιοχές που μοιράζονται κοινά χαρακτηριστικά, όπως χρώμα, υφή ή φωτεινότητα. Κλείνοντας το κεφάλαιο,

γίνεται μια αποτίμηση των αλγορίθμων τμηματοποίησης εικόνας καθώς είναι μια απαραίτητη διαδικασία για την κατανόηση της απόδοσης, της ακρίβειας και της καταλληλότητάς τους σε διαφορετικά δεδομένα.

Στα τελευταία κεφάλαια της παρούσας διατριβής, παρουσιάζονται τα αποτελέσματα της έρευνας που αντλήθηκαν μέσω των προηγμένων τεχνικών ανάλυσης και επεξεργασίας. Στο τέταρτο κεφάλαιο, τα αποτελέσματα από την πλήρως αυτοματοποιημένη μέθοδο 3D τμηματοποίησης σε δεδομένα ασθενών, αξιολογούνται και συγκρίνονται με άλλες κλασσικές τεχνικές τμηματοποίησης, όσον αφορά την ακρίβεια του εξαγόμενου 3D μοντέλου αορτής. Στο πέμπτο κεφάλαιο, παρέχεται στον αναγνώστη μια αναλυτική συγκριτική μελέτη αρχιτεκτονικών βαθιάς μάθησης μέσω της οποίας επιτυγχάνεται η αυτόματη τμηματοποίηση της κοιλιακής αορτής σε δεδομένα υγιών και ασθενών ατόμων.

Στη συνέχεια στο έκτο κεφάλαιο παρουσιάζονται και αξιολογούνται τα αποτελέσματα που εξήχθησαν μέσω της βελτιωμένης μεθόδου αυτόματης τμηματοποίησης της κοιλιακής αορτής σε δεδομένα υγιών και ασθενών ατόμων, χρησιμοποιώντας αρχιτεκτονικές βαθιάς μάθησης. Κλείνοντας, το έβδομο κεφάλαιο αποτελεί τον επίλογο της διδακτορικής διατριβής, αναφέροντας τα συμπεράσματα των προτεινόμενων καινοτόμων τεχνικών ανάλυσης και επεξεργασίας δεδομένων, ενώ παράλληλα αναφέρονται οι πιθανές μελλοντικές κατευθύνσεις και προεκτάσεις των προτεινόμενων συστημάτων, με στόχο την περαιτέρω αναβάθμιση της επεξεργασίας των καρδιολογικών δεδομένων.

Κεφάλαιο 2: Ανατομία και Μέθοδοι Απεικόνισης Καρδιάς

2.1 Εισαγωγή στη Ανατομία της Καρδιάς

Η ανατομία της καρδιάς είναι θεμελιώδης για την κατανόηση των λειτουργιών του κυκλοφορικού συστήματος και της υγείας του ανθρώπινου σώματος. Η καρδιά, ως η κύρια αντλία αίματος του ανθρώπινου οργανισμού, αποτελείται από μια πολύπλοκη και εξαιρετικά εξελιγμένη δομή η οποία διασφαλίζει την αποτελεσματική κυκλοφορία του αίματος μέσω των αγγείων του σώματος και των πνευμόνων. Η μελέτη της ανατομίας της καρδιάς είναι καθοριστική για την κατανόηση της φυσιολογίας της καρδιακής λειτουργίας καθώς επίσης και για την ανάλυση των καρδιολογικών παθήσεων.

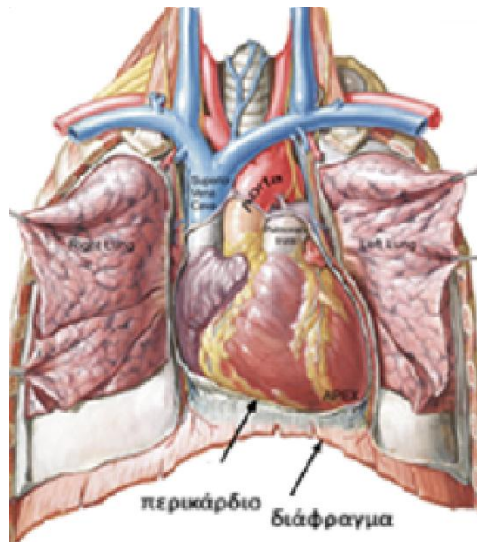
Βασισμένοι σε βιβλιογραφικές αναφορές, η καρδιά αποτελείται από τέσσερις κοιλότητες, δύο κόλπους και δύο κοιλίες, οι οποίες συνεργάζονται για να διασφαλίσουν τη συνεχιζόμενη ροή του αίματος [6]. Επιπλέον, η καρδιά περιέχει τέσσερις βασικές βαλβίδες (μιτροειδής, τριγλώχινη, αορτική, και πνευμονική), οι οποίες παίζουν σημαντικό ρόλο στη μονόδρομη ροή του αίματος, αποτρέποντας την αντίστροφη ροή και διασφαλίζοντας τη σωστή λειτουργία της [7].

Η τοποθέτηση της καρδιάς μέσα στη θωρακική κοιλότητα και η προστασία της από το περικάρδιο είναι εξίσου σημαντικές για τη διατήρηση της καρδιολογικής υγείας. Το περικάρδιο εκκρίνει λιπαντικό υγρό το οποίο βοηθά στην αποφυγή τριβής και επιπλέον προστατεύει την καρδιά από τραυματισμούς και φλεγμονές [8]. Η αιματική τροφοδοσία της καρδιάς πραγματοποιείται μέσω των στεφανιαίων αρτηριών, οι οποίες παίζουν σημαντικό ρόλο στη διατήρηση της ομαλής καρδιαγγειακής λειτουργίας [9].

Η κατανόηση της αναφερθείσας ανατομίας και σε συνδυασμό με τις σύγχρονες απεικονιστικές μεθόδους συνεχίζει να παρέχει πολύτιμες πληροφορίες για τη διάγνωση και την αντιμετώπιση των καρδιολογικών παθήσεων. Η συνεχιζόμενη έρευνα στον τομέα της καρδιαγγειακής φυσιολογίας και ανατομίας, αναδεικνύει την ανάγκη για μια ολοκληρωμένη κατανόηση των μοριακών, κυτταρικών και οργάνων-επιπέδων αλληλεπιδράσεων, έτσι ώστε να υπάρξει βελτίωση της θεραπείας και πρόληψης των καρδιαγγειακών ασθενειών [10].

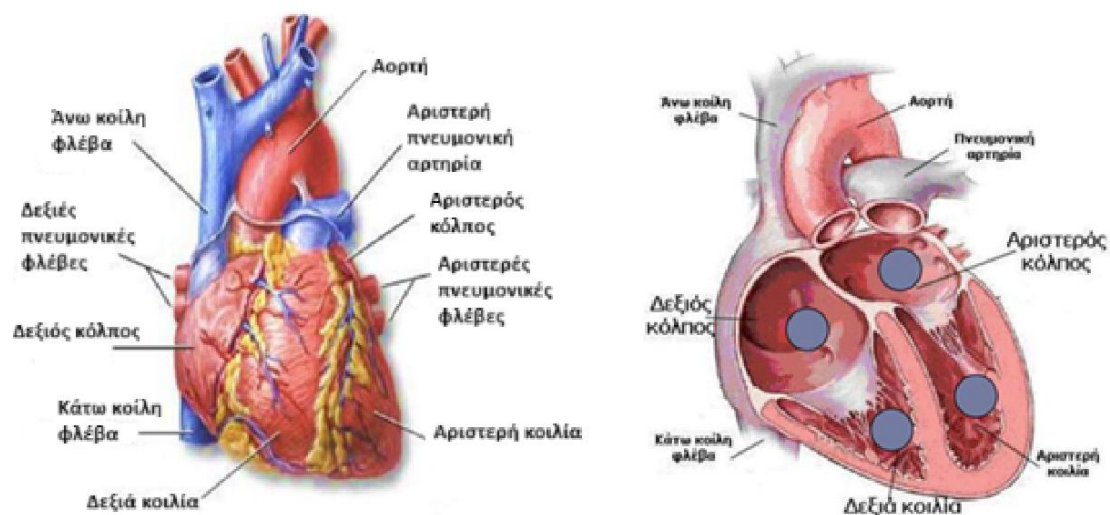
2.1.1 Βασική Μορφολογία

Η καρδιά είναι ένα κοίλο, μυώδες όργανο σε σχήμα κώνου το οποίο βρίσκεται στην θωρακική κοιλότητα ανάμεσα στους δυο πνεύμονες και είναι ελαφρώς προς τα αριστερά του μέσου γραμμικού άξονα (**Εικόνα 2.1**). Η κύρια λειτουργία της καρδιάς είναι να δέχεται το αίμα που προέρχεται από τις φλέβες και να το ωθεί προς τις αρτηρίες.



Εικόνα 2.1: Ανατομία της καρδιάς [11]

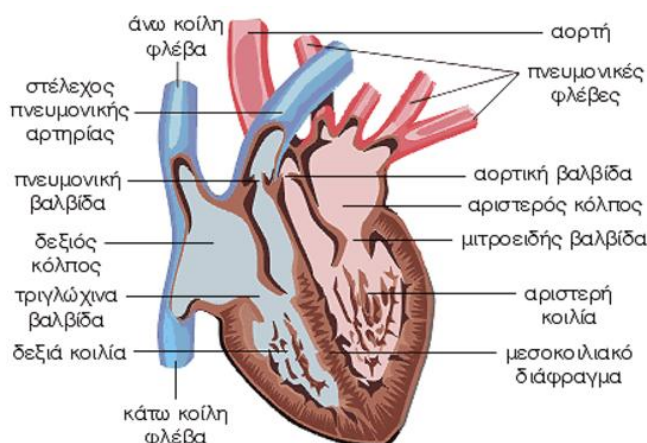
Η καρδιά κατά κύριο λόγο χωρίζεται σε δύο μέρη, κάθε ένα αποτελείται από έναν άνω θάλαμο (τον κόλπο) και έναν κάτω θάλαμο (κοιλία). Συνολικά διακρίνονται δύο κόλποι (δεξιός και αριστερός) και οι δύο κοιλίες (δεξιά και αριστερή) (**Εικόνα 2.2**). Οι κόλποι είναι υπεύθυνοι για την συλλογή του αίματος από τα διάφορα μέρη του σώματος ενώ οι κοιλίες προωθούν το αίμα έξω από την καρδιά.



Εικόνα 2.2: Κόλποι και κοιλίες καρδιάς [11]

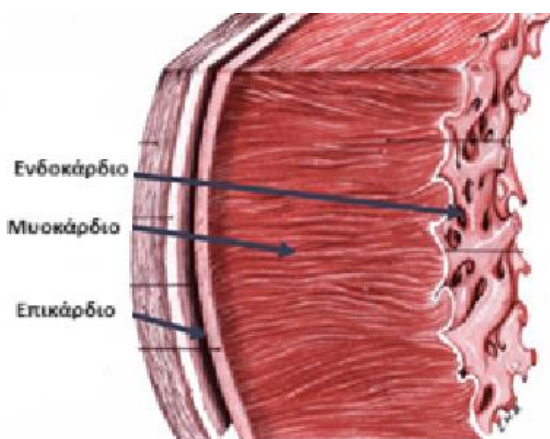
Οι τέσσερις κοιλότητες διαχωρίζονται από μεσοτοιχώματα, το μεσοκοιλιακό και μεσοκοιλιακό διάφραγμα. Οι κολποκοιλιακές βαλβίδες (μιτροειδής και τριγλώχινα) και οι ημιτελικές βαλβίδες (αορτική και πνευμονική) είναι υπεύθυνες για τη μονόδρομη ροή του αίματος. Πιο συγκεκριμένα, ο δεξιός κόλπος συνδέεται με τη δεξιά κοιλία μέσω του δεξιού

κολποκοιλιακού στομίου, το οποίο κλείνει με την τριγλώχινη ή μηνοειδή βαλβίδα. Αντιστοίχως ο αριστερός κόλπος συνδέεται με την αριστερή κοιλία μέσω του αριστερού κολποκοιλιακού στομίου το οποίο κλείνει με τη διγλώχινη ή μητροειδή βαλβίδα (**Εικόνα 2.3**). Σκοπός της λειτουργία των βαλβίδων, είναι η αποτροπή της παλινδρόμησης του αίματος από τις κοιλίες προς τους κόλπους [12]. Μελέτες δείχνουν ότι οι μορφολογικές διαφοροποιήσεις στις βαλβίδες ενδέχεται να επηρεάζουν την αιμοδυναμική λειτουργία της καρδιάς [13].



Εικόνα 2.3: Διαφράγματα, κολποκοιλιακές και ημιτελικές βαλβίδες καρδιάς [14]

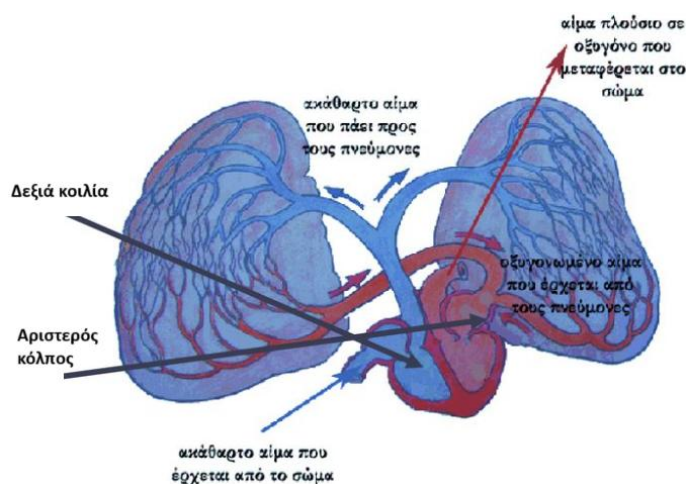
Η καρδιά περιβάλλεται από τρία στρώματα, το ενδοκάρδιο, το μυοκάρδιο και το επικάρδιο. Ανάμεσα στο ενδοκάρδιο και το επικάρδιο εντοπίζεται το παχύτερο τοίχωμα της καρδιάς το μυοκάρδιο το οποίο αποτελείται από μυϊκές ίνες που οργανώνονται σε περίπλοκες ελικοειδείς δομές (**Εικόνα 2.4**). Έρευνες χρησιμοποιώντας απεικονιστικές τεχνικές, όπως το Diffusion Tensor Imaging (DTI), έδειξαν ότι η διάταξη των ινών στο μυοκάρδιο βελτιστοποιεί σε μεγάλο βαθμό την αντλητική λειτουργία της καρδιάς [15].



Εικόνα 2.4: Στρώματα καρδιάς [16]

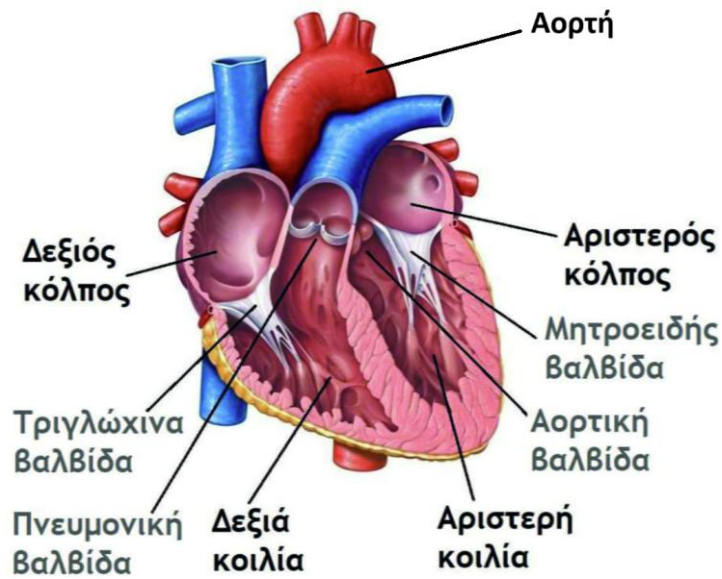
2.1.2 Σωματική κυκλοφορία

Η σωματική κυκλοφορία είναι ένα περίπλοκο δίκτυο αρτηριών, φλεβών και τριχοειδών αγγείων το οποίο είναι υπεύθυνο για την εξασφάλιση της μεταφοράς οξυγόνου, θρεπτικών ουσιών και αποβλήτων από και προς τους ιστούς του σώματος. Τα αγγεία που συνθέτουν αυτήν την κυκλοφορία του σώματος, σχετίζονται στενά με την καρδιακή λειτουργία, η οποία αντλεί το αίμα μέσω του συστήματος αυτού [17]. Πιο συγκεκριμένα, από τη δεξιά κοιλία εκφύεται η πνευμονική αρτηρία η οποία μεταφέρει φλεβικό αίμα στους πνεύμονες. Στη συνέχεια η μεταφορά του οξυγονωμένου αίματος στο σώμα γίνεται μέσω των πνευμονικών φλεβών που καταλήγουν στον αριστερό κόλπο (**Εικόνα 2.5**). Κύριο πλεονέκτημα της παρουσίας των βαλβίδων στο εσωτερικό των φλεβών είναι ότι αποτρέπει η παλινδρόμηση του αίματος.



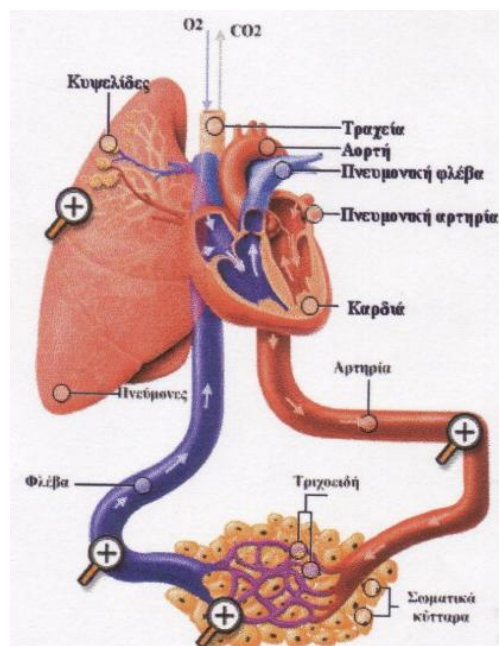
Εικόνα 2.5: Ροή αίματος [16]

Η αορτή, η οποία αποτελεί την κύρια αρτηρία της σωματικής κυκλοφορίας (**Εικόνα 2.6**), ξεκινάει από την αορτική βαλβίδα της καρδιάς και στη συνέχεια διακλαδίζεται σε μικρότερα αγγεία τα οποία με τη σειρά τους τροφοδοτούν τους ιστούς με οξυγονωμένο αίμα [18]. Έρευνες έχουν δείξει ότι οι ελαστικές ιδιότητες της αορτής είναι σημαντικές, καθώς η αυξημένη ακαμψία σχετίζεται άμεσα με καρδιαγγειακά νοσήματα και η αορτική σκλήρυνση οδηγεί στην αύξηση της κεντρικής αρτηριακής πίεσης, που προκαλεί με την σειρά της υπέρταση [19].



Εικόνα 2.6: Αορτή και αορτική βαλβίδα [20]

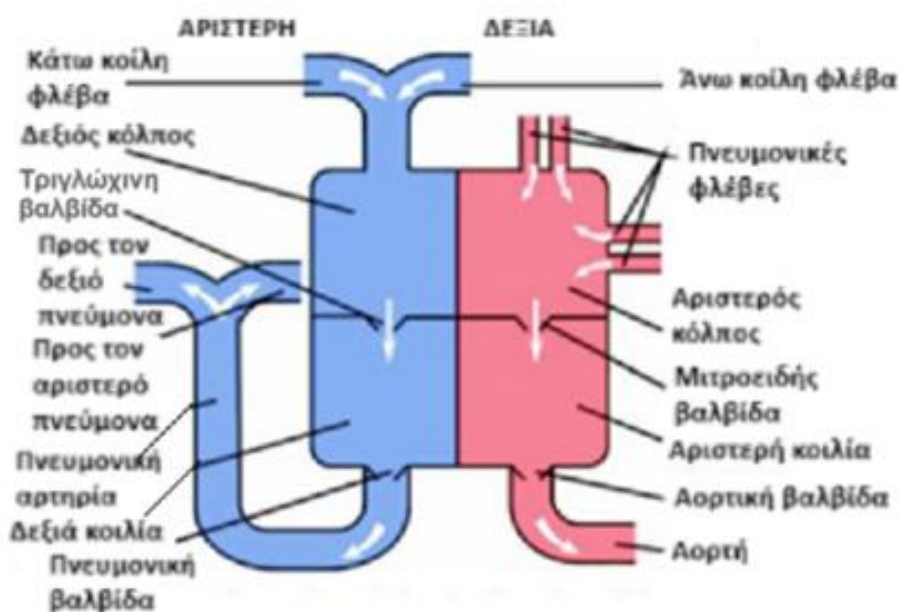
Τα τριχοειδή (Εικόνα 2.7) είναι τα μικρότερα αιμοφόρα αγγεία της σωματικής κυκλοφορίας και τα οποία είναι υπεύθυνα για την ανταλλαγή αερίων, θρεπτικών συστατικών και αποβλήτων. Η δομή τους, που αποτελείται από ένα μόνο στρώμα ενδοθηλιακών κυττάρων, διευκολύνει τη διάχυση και παράλληλα συνδέουν τις μικρότερες αρτηρίες με τις αντίστοιχες μικρότερες φλέβες [21].



Εικόνα 2.7: Τριχοειδή αγγεία [21]

2.1.3 Δυναμικές Διαδικασίες Καρδιάς

Οι Δυναμικές διαδικασίες της καρδιάς είναι ένα σύνθετο σύνολο μηχανισμών το οποίο καθορίζει, τη λειτουργία της καρδιάς ως αντλίας (Εικόνα 2.8), τη ρύθμιση του ρυθμού, τη μεταφορά ιόντων, τη μηχανική απόδοση και την αντίδραση της καρδιάς σε εξωτερικά και εσωτερικά ερεθίσματα. Οι ανωτέρω οι διεργασίες είναι κρίσιμες για τη διατήρηση της αιμοδυναμικής σταθερότητας της καρδιάς και επηρεάζονται άμεσα από την αλληλεπίδραση των φυσιολογικών και παθολογικών παραγόντων.



Εικόνα 2.8: Η καρδιά συμπεριφέρεται ως μια αντλία [22]

Η ηλεκτρική δραστηριότητα της καρδιάς είναι η βάση του καρδιακού ρυθμού και οι οποία ξεκινά από τον φλεβόκομβο (SA node) στον οποίο παράγονται τα ενεργειακά δυναμικά. Παράλληλα οι ιοντικές μεταφορές μέσω των διαύλων νατρίου (Na^+), καλίου (K^+), και ασβεστίου (Ca^{2+}) είναι υπεύθυνες και παίζουν σημαντικό ρόλο στη δημιουργία και μετάδοση των δυναμικών δράσης [23]. Σε ότι αφορά τις διαταραχές στα δυναμικά αυτά, σχετίζονται άμεσα με τις αρρυθμίες όπως είναι η κολπική μαρμαρυγή [24].

Η ρυθμιστική δραστηριότητα της καρδιάς αποτελείται από δύο βασικές φάσεις, τη συστολική και τη διαστολική, οι οποίες είναι αλληλένδετες και λειτουργούν διαδοχικά για την διοχέτευση του αίματος στο κυκλοφορικό σύστημα [25]. Αυτές οι δύο φάσεις αποτελούνται από σύνθετους μηχανισμούς οι οποίοι περιλαμβάνουν την ηλεκτρική διέγερση, τη μηχανική σύσπαση και τη ρύθμιση του ενδοκυττάριου ασβεστίου (Ca^{2+}).

Η καρδιακή συστολή είναι η ενεργή φάση όπου οι κοιλίες συστέλλονται έτσι ώστε να αντλήσουν αίμα στην πνευμονική και συστηματική κυκλοφορία. Η ενεργοποίηση της

συστολής γίνεται από την αποπόλωση των κυττάρων του μυοκαρδίου μέσω της ροής ιόντων Na^+ , K^+ και Ca^{2+} . Η ενδοκυτταρική αύξηση του Ca^{2+} είναι ο κύριος παράγοντας που ενεργοποιεί την αλληλεπίδραση της ακτίνης και της μυοσίνης. Η συστολή επιτυγχάνεται διαμέσου του συμπλέγματος τροπονίνης- Ca^{2+} καθώς αυτό απομακρύνει την τροπομυοσίνη από τη θέση δέσμευσης της ακτίνης.

Η διαστολή είναι η παθητική φάση κατά την οποία οι κοιλίες χαλαρώνουν και γεμίζουν με αίμα. Η χαλάρωση των μυοκαρδιοκυττάρων απαιτεί την απομάκρυνση του Ca^{2+} από το κυτταρόπλασμα μέσω μιας αντλίας ασβεστίου που ονομάζεται SERCA και του ανταλλάκτη $\text{Na}^+/\text{Ca}^{2+}$. Ως εκ τούτου η ενεργή μεταφορά του Ca^{2+} στο σαρκοπλασματικό δίκτυο μειώνει τη συγκέντρωση του Ca^{2+} μέσα στο κύτταρο, οδηγώντας στη χαλάρωση [23].

Η συστολική και η διαστολική φάση είναι αλληλεξαρτώμενες και στενά συνδεδεμένες καθώς η ελλιπής συστολή μπορεί να επηρεάσει τη διαστολή και το αντίστροφο. Η αλληλεξάρτηση των δύο φάσεων, επιβεβαιώνεται επίσης και μέσω της ανεπάρκειας της ενέργειας (ATP), καθώς αυτή επηρεάζει και τις δύο φάσεις μέσω διαταραχών στον κύκλο Ca^{2+} [26].

Η καρδιά προσαρμόζει την αιμοδυναμική της λειτουργία σύμφωνα με τις ανάγκες του σώματος, αξιοποιώντας μηχανισμούς όπως το φαινόμενο Frank-Starling και τη ρύθμιση του ρυθμού των χτύπων της. Πιο συγκεκριμένα η διάταση των μυοκαρδιακών ινών κατά τη διαστολή αυξάνει τη δύναμη συστολής η οποία διαταράσσεται στην καρδιακή ανεπάρκεια [27]. Επίσης, οι υποδοχείς στον καρδιακό ιστό οι οποίοι λειτουργούν ως μηχανικοί αισθητήρες, ανιχνεύουν μεταβολές στην πίεση και τη ροή, ρυθμίζοντας την αγγειοσυστολή και τη λειτουργία της καρδιάς [28].

Σχετικά με την ενεργειακή αποδοτικότητα της καρδιάς, αυτή καταναλώνει μεγάλες ποσότητες ενέργειας, οι οποίες παράγονται κυρίως μέσω οξειδωτικής φωσφορυλίωσης στα μιτοχόνδρια. Έρευνες δείχνουν ότι σε συνθήκες στρες ή καρδιακής ανεπάρκειας, υπάρχει έντονη διαταραχή του ενεργειακού μεταβολισμού [29].

Από τα ανωτέρω συμπεραίνουμε ότι οι δυναμικές διαδικασίες της καρδιάς είναι πολύπλευρες και συνδυάζουν φυσιολογικούς, βιοχημικούς και μηχανικούς μηχανισμούς οι οποίοι χαρακτηρίζουν τη λειτουργία της ως αντλίας. Η κατανόηση αυτών των διαδικασιών μέσω σύγχρονων ερευνών είναι η βάση για την ανάπτυξη νέων θεραπευτικών στρατηγικών.

2.1.4 Παθολογικές Καταστάσεις

Οι παθήσεις που επηρεάζουν τη λειτουργία της καρδιάς βρίσκονται στο επίκεντρο της επιστημονικής έρευνας εξαιτίας της πολυπλοκότητας των μηχανισμών τους και της σημαντικής τους επίδρασης στη δημόσια υγεία. Καταστάσεις όπως η καρδιακή ανεπάρκεια, οι διαταραχές του ρυθμού, οι μυοκαρδιοπάθειες και οι φλεγμονώδεις διαταραχές της καρδιάς μπορούν να προκληθούν από διάφορους παράγοντες όπως είναι οι γενετικές μεταλλάξεις, οι

περιβαλλοντικές επιδράσεις και οι μεταβολικές δυσλειτουργίες. Παρακάτω παρουσιάζονται οι κύριες παθολογικές καταστάσεις που σχετίζονται με την καρδιά.

2.1.4.1 Καρδιακή Ανεπάρκεια (Heart Failure)

Η καρδιακή ανεπάρκεια (ΚΑ) αποτελεί ένα πολύπλοκο κλινικό σύνδρομο στο οποίο η καρδιά δεν μπορεί να εξασφαλίσει επαρκή παροχή αίματος ώστε να καλύψει τις μεταβολικές ανάγκες του οργανισμού, ενώ παράλληλα δυσχεραίνεται η επιστροφή του φλεβικού αίματος από τους περιφερικούς ιστούς [30]. Αυτή η δυσλειτουργία οφείλεται κυρίως σε δομικές ή λειτουργικές βλάβες του μυοκαρδίου οι οποίες οδηγούν στην εμφάνιση χαρακτηριστικών συμπτωμάτων, όπως η δύσπνοια και η κόπωση καθώς και σημείων όπως το περιφερικό οίδημα και η αυξημένη φλεβική πίεση. Με την πρόοδο της ασθένειας τα συμπτώματα και οι επιπτώσεις γίνονται πιο σοβαρά, επιβαρύνοντας την πρόγνωση του ασθενούς.

Πριν την εκδήλωση των κλινικών συμπτωμάτων της ασθένειας, οι ασθενείς ενδέχεται να εμφανίζουν σημάδια καρδιακής δυσλειτουργίας όπως είναι οι διαταραχές στη συστολική ή διαστολική λειτουργία της αριστερής κοιλίας. Τα ευρήματα αυτά θεωρούνται πρώιμοι δείκτες καρδιακής ανεπάρκειας. Η ανίχνευση αυτών των ενδείξεων είναι κρίσιμη καθώς η έγκαιρη έναρξη της θεραπείας έχει αποδειχθεί ότι μπορεί να μειώσει τη θνησιμότητα σε άτομα με ασυμπτωματική συστολική δυσλειτουργία της αριστερής κοιλίας [31], [32].

Ο όρος που χρησιμοποιείται για την ταξινόμηση της καρδιακής ανεπάρκειας βασίζεται στον υπολογισμό του κλάσματος εξώθησης της αριστερής κοιλίας (ΚΕΑΚ). Με βάση αυτήν την προσέγγιση καλύπτεται ένα ευρύ φάσμα ασθενών. Ασθενείς με φυσιολογικό ΚΕΑΚ που συνήθως ορίζεται ως $>50\%$, παλαιότερα περιγράφονταν με τον όρο «διαστολική δυσλειτουργία», ενώ σήμερα αναφέρονται ως ασθενείς με διατηρημένο ΚΕΑΚ. Αντίθετα, οι ασθενείς με ΚΕΑΚ κάτω από το 40% κατατάσσονται στην κατηγορία της καρδιακής ανεπάρκειας με μειωμένο ΚΕΑΚ, όρος που αντικατέστησε τη συστολική δυσλειτουργία της αριστερής κοιλίας. Επιπλέον, έχει οριστεί μια ενδιάμεση περιοχή του ΚΕΑΚ, μεταξύ $40-49\%$, η οποία περιλαμβάνει ασθενείς με καρδιακή ανεπάρκεια και μετρίως μειωμένο ΚΕΑΚ. Αυτή η ομάδα φαίνεται να συνδυάζει χαρακτηριστικά τόσο συστολικής όσο και διαστολικής δυσλειτουργίας της αριστερής κοιλίας.

2.1.4.2 Ισχαιμική Καρδιοπάθεια

Η στεφανιαία νόσος, γνωστή και ως ισχαιμική καρδιοπάθεια, έχει αναδειχθεί ως η κύρια αιτία θανάτου παγκοσμίως και η οποία αποτελεί ένα σύνολο παθήσεων που περιλαμβάνουν τη σταθερή και ασταθή στηθάγχη, το έμφραγμα, το εγκεφαλικό επεισόδιο, τον αιφνίδιο ισχαιμικό θάνατο κ.ά. Τα τελευταία χρόνια έχει σημειωθεί αύξηση της θνησιμότητας που συνδέεται με

τη νόσο τόσο συνολικά όσο και σε άτομα νεότερης ηλικίας. Παρατηρείται επίσης συχνότερη εμφάνιση στους άνδρες σε σύγκριση με τις γυναίκες της ίδιας ηλικίας. Περίπου το ένα τρίτο των ενηλίκων που εισάγονται στο νοσοκομείο λόγω θωρακικού άλγους διαγιγνώσκονται με κάποια μορφή ισχαιμικής καρδιοπάθειας.

Οι παράγοντες κινδύνου για την εκδήλωση της ισχαιμικής καρδιοπάθειας διακρίνονται σε δύο κατηγορίες, τους τροποποιήσιμους και τους μη τροποποιήσιμους. Στην πρώτη κατηγορία ανήκουν παράγοντες όπως η υπέρταση, το κάπνισμα, ο σακχαρώδης διαβήτης, η έλλειψη φυσικής άσκησης, η παχυσαρκία, οι διαταραχές λιπιδίων και η υπερβολική κατανάλωση αλκοόλ. Οι μη τροποποιήσιμοι παράγοντες περιλαμβάνουν την ηλικία, το φύλο και το οικογενειακό ιστορικό.

Ο κύριος μηχανισμός που συμβάλλει στην εκδήλωση της νόσου αν και δεν είναι ο μοναδικός, αφορά την αθηροσκλήρωση των στεφανιαίων αρτηριών. Η κατάσταση αυτή εμφανίζεται συχνά σε άτομα που διαθέτουν πολλούς από τους προαναφερθέντες παράγοντες κινδύνου. Ωστόσο, υπάρχουν σπάνιες περιπτώσεις κατά τις οποίες η στεφανιαία νόσος μπορεί να παρουσιαστεί χωρίς την ύπαρξη στεφανιαίας αθηροσκλήρωσης. Ένα παράδειγμα τέτοιας περίπτωσης είναι το Σύνδρομο X το οποίο εκδηλώνεται με χαρακτηριστικό ισχαιμικό πόνο στο στήθος, ενώ στις μεγαλύτερες στεφανιαίες αρτηρίες δεν εντοπίζονται αθηρωματικές στενώσεις, γεγονός που επιβεβαιώνεται με στεφανιογραφία. Το Σύνδρομο X παρατηρείται συχνότερα σε γυναίκες ενώ η ακριβής αιτία του παραμένει ασαφής. Μια πιθανή εξήγηση που δεν έχει επιβεβαιωθεί πλήρως σχετίζεται με δυσλειτουργίες στη μικροκυκλοφορία και τη μικροαγγείωση.

Η πρόληψη της ισχαιμικής καρδιοπάθειας βασίζεται σε έναν υγιεινό τρόπο ζωής ο οποίος περιλαμβάνει τη σωστή διατροφή, τη συστηματική άσκηση, τη διακοπή του καπνίσματος και τη θεραπευτική αντιμετώπιση παραγόντων κινδύνου, όπως η υπέρταση, ο διαβήτης και η δυσλιπιδαιμία. Επιπλέον, διαγνωστικές εξετάσεις όπως το ηλεκτροκαρδιογράφημα, ο υπέρηχος καρδιάς, το τεστ κοπώσεως και η στεφανιογραφία αποτελούν σημαντικά εργαλεία για την ακριβή διάγνωση και την παρακολούθηση της εξέλιξης της νόσου.

Τα πιο διαδεδομένα συμπτώματα περιλαμβάνουν πόνο στο στήθος ο οποίος εντοπίζεται κυρίως πίσω από το στήννο και μπορεί να εξαπλωθεί στα χέρια, τους ώμους, την πλάτη, το σαγόνι και αλλού. Ο πόνος έχει διάφορες μορφές, όπως καυστικός, σφικτικός ή αίσθημα βάρους, διαρκεί λίγα λεπτά και συνήθως εμφανίζεται ή επιδεινώνεται κατά τη διάρκεια σωματικής ή συναισθηματικής κόπωσης [33], [34].

Σε ορισμένες περιπτώσεις ο πόνος ξεκινά από την περιοχή του επιγαστρίου, γεγονός που μπορεί να τον κάνει να συγχέεται με στομαχικές διαταραχές. Για τον λόγο αυτό, είναι απαραίτητη η διαφορική διάγνωση. Επιπλέον, δυσκολία στην αναπνοή αποτελεί συχνό συνοδό σύμπτωμα, ενώ δεν αποκλείεται να απουσιάζουν εντελώς τα τυπικά σημεία, όπως συμβαίνει σε ορισμένες περιπτώσεις οξέος εμφράγματος του μυοκαρδίου που ενδέχεται να προκαλέσει επικίνδυνες αρρυθμίες ή ακόμα και αιφνίδιο θάνατο.

Ο πόνος στο στήθος που προκύπτει κατά την εκτέλεση προβλέψιμων δραστηριοτήτων όπως ελαφριά άσκηση ή γεύματα, χαρακτηρίζεται ως σταθερή στηθάγχη. Αυτή η μορφή στηθάγχης συνδέεται με αυξημένη επιβάρυνση της καρδιάς καθώς η περιορισμένη ροή αίματος μέσω των στεφανιαίων αρτηριών δεν επαρκεί για την κάλυψη της αυξημένης ζήτησης κατά τη διάρκεια αυτών των δραστηριοτήτων.

Αντίθετα, όταν η ένταση, η συχνότητα ή ο χαρακτήρας του πόνου μεταβάλλεται τότε αναφερόμαστε για ασταθή στηθάγχη, που θεωρείται πρόδρομος του εμφράγματος του μυοκαρδίου. Σε κάθε περίπτωση, τα συμπτώματα απαιτούν λεπτομερή διερεύνηση η οποία μπορεί να περιλαμβάνει ακόμα και στεφανιογραφία. Αυτή η διαγνωστική μέθοδος αν και επεμβατική προσφέρει απόλυτη ακρίβεια στον εντοπισμό της στεφανιαίας νόσου [35].

Η θεραπεία περιλαμβάνει την αντιμετώπιση των στενώσεων που εντοπίζονται στις στεφανιαίες αρτηρίες, συνήθως μέσω αγγειοπλαστικής ή διαδερμικής επαναγγείωσης. Παράλληλα, η πρόληψη διαδραματίζει καθοριστικό ρόλο και η αξιολόγηση των παραγόντων κινδύνου για αθηροσκλήρωση και η τακτική διενέργεια προληπτικών εξετάσεων, ανάλογα με την ηλικία και την παρουσία προδιαθεσικών παραγόντων, συμβάλλουν καθοριστικά στη διατήρηση της καρδιακής υγείας.

2.1.4.3 Αρρυθμίες

Οι καρδιακές αρρυθμίες είναι διαταραχές που επηρεάζουν τον φυσιολογικό ρυθμό της καρδιάς, χαρακτηριζόμενες από τον εντοπισμό της προέλευσής τους και τα συμπτώματα που εκδηλώνει ο κάθε ασθενής. Μπορεί να περιλαμβάνουν ακανόνιστους καρδιακούς παλμούς, ακανόνιστους ρυθμούς, γρήγορους καρδιακούς παλμούς (ταχυκαρδία) ή αργούς (βραδυκαρδία). Βάσει της πηγής προέλευσής τους, οι αρρυθμίες διακρίνονται σε κολπικές ή υπερκοιλιακές που προέρχονται από τους κόλπους και σε κοιλιακές που προκύπτουν από τις κοιλίες.

Οι αρρυθμίες μπορεί να εκδηλώνονται με διάφορους τρόπους, και τα συμπτώματα ποικίλλουν από ασθενή σε ασθενή. Ορισμένοι πάσχοντες εμφανίζουν αίσθημα παλμών, δυσφορία στο στήθος, δύσπνοια ή ακόμη και συγκοπή. Η συγκοπή μπορεί να οφείλεται σε αιτίες όπως η βραδυκαρδία, η ταχυκαρδία, οι βαλβιδικές παθήσεις ή η μυοκαρδιακή δυσλειτουργία, αλλά και σε μη αρρυθμικούς παράγοντες, όπως η αγγειοδιαστολή.

Η κολπική μαρμαρυγή (AF) είναι η συχνότερη μορφή καρδιακής αρρυθμίας και χαρακτηρίζεται από ακανόνιστους και ταχείς καρδιακούς παλμούς. Τα συμπτώματά της οποίας περιλαμβάνουν αίσθημα παλμών, ζάλη, άγχος και μειωμένη ικανότητα άσκησης, γεγονός που επηρεάζει σοβαρά την ποιότητα ζωής. Περίπου το ένα τρίτο των ασθενών παραμένουν ασυμπτωματικοί γεγονός που δυσχεραίνει τη διάγνωση και την έγκαιρη αντιμετώπιση.

Διάφοροι παράγοντες μπορούν να προκαλέσουν ή να επιδεινώσουν τις αρρυθμίες, όπως οι καρδιακές ή θωρακικές χειρουργικές επεμβάσεις, το σωματικό ή ψυχολογικό στρες, οι λοιμώξεις, η κατανάλωση αλκοόλ ή καφεΐνης, το κάπνισμα και η υπογλυκαιμία. Οι μηχανισμοί που εμπλέκονται στη γένεση και διατήρηση των αρρυθμιών σχετίζονται με τις ιοντικές διεργασίες που επηρεάζουν το ηλεκτρικό δυναμικό της καρδιάς. Το δυναμικό δράσης εξαρτάται από τη συντονισμένη λειτουργία πρωτεϊνών διαύλων ιόντων που ρυθμίζουν τη ροή ιόντων μέσω της κυτταρικής μεμβράνης. Διαφορές στην έκφραση ή τις ιδιότητες αυτών των διαύλων συμβάλλουν στις διακυμάνσεις του δυναμικού δράσης σε διαφορετικές περιοχές της καρδιάς.

Οι αρρυθμίες μπορούν να ταξινομηθούν με διάφορους τρόπους, όπως με βάση τη μηχανιστική τους προέλευση (μη αντανakλαστική δραστηριότητα, επανεισαγόμενη δραστηριότητα), το επίπεδο στο οποίο εμφανίζονται (κυτταρικό ή ιστικό) ή τις ηλεκτροφυσιολογικές τους αιτίες (αστάθεια ασβεστίου, μειωμένο ή υπερβολικό αποθεματικό επαναπόλωσης). Αυτές οι κατηγορίες βοηθούν στην κατανόηση και την εξατομίκευση της θεραπευτικής προσέγγισης για κάθε περίπτωση [36], [37], [38].

Η διάγνωση των αρρυθμιών συνήθως πραγματοποιείται μέσω καταγραφής ηλεκτροκαρδιογραφήματος (ΗΚΓ). Αυτό μπορεί να γίνει είτε με τη χρήση φορητών συσκευών, όπως η περιπατητική οθόνη Holter, είτε μέσω συστημάτων επείγουσας ιατρικής φροντίδας, όπως το ΗΚΓ 12 ηλεκτροδίων, που εφαρμόζονται άμεσα σε περιπτώσεις καρδιακής ανακοπής. Με τις μεθόδους αυτές, προσδιορίζεται η πηγή της αρρυθμίας και διαμορφώνεται το θεραπευτικό πλάνο [36], [39].

Οι επιλογές για την αντιμετώπιση των αρρυθμιών περιλαμβάνουν φαρμακευτική αγωγή, αφαίρεση καθετήρα, εμφύτευση βηματοδότη, χρήση απινιδωτή καρδιοανατάκτη (ICD), ή έναν συνδυασμό των παραπάνω. Η θεραπευτική προσέγγιση εξαρτάται από τη συγκεκριμένη αρρυθμία και τα χαρακτηριστικά του κάθε ασθενούς, ενώ απαιτείται συνεχής παρακολούθηση κατά τη διάρκεια της θεραπείας. Παρά την πρόοδο, η αποτελεσματικότητα των θεραπειών συχνά ποικίλλει, και η αντιμετώπιση των αρρυθμιών παραμένει πρόκληση.

Οι βηματοδότες αποτελούν μια αξιόπιστη λύση για τη βραδυκαρδία, ενώ οι ICD είναι αποτελεσματικοί για την αντιμετώπιση κοιλιακών ταχυαρρυθμιών, παρέχοντας ηλεκτρικά σοκ ή ειδικές τεχνικές βηματοδότησης. Ωστόσο, οι ICD δεν αποτρέπουν την εμφάνιση αρρυθμιών ή τη συγκοπή. Σχετικά με τα αντιαρρυθμικά φάρμακα, οι νέες εγκρίσεις είναι περιορισμένες, καθώς ορισμένα από αυτά μπορεί να προκαλέσουν σοβαρές παρενέργειες, όπως τοξικότητα σε όργανα ή καρδιακή ανακοπή σε ευαίσθητους ασθενείς.

Τα φάρμακα κατηγορίας I, που λειτουργούν ως αναστολείς ρευμάτων νατρίου (Na^+), μειώνουν τη διέγερση της καρδιάς, ενώ εκείνα της κατηγορίας III αποκλείουν διαύλους καλίου (K^+). Παρά ταύτα, η αποτελεσματικότητα της αντιαρρυθμικής φαρμακευτικής αγωγής παραμένει περιορισμένη, εντείνοντας την ανάγκη για την ανάπτυξη νέων θεραπειών.

Η αφαίρεση καθετήρα αν και χρήσιμη, ενέχει κινδύνους όπως τραυματισμούς στα αιμοφόρα αγγεία ή ακόμη και βλάβες στον καρδιακό μυ. Παρόλο που οι διαθέσιμες τεχνικές για την αντιμετώπιση κλινικών αρρυθμιών έχουν σημειώσει πρόοδο, εξακολουθούν να θεωρούνται περιορισμένες ως προς την αποτελεσματικότητα και την ακρίβεια.

Τα τελευταία χρόνια, η έρευνα εστιάζει στην ταυτοποίηση γονιδίων που σχετίζονται με μονογονιδιακά σύνδρομα αρρυθμίας, ενώ προηγμένα μοντέλα καρδιακών αρρυθμιών βασισμένα σε ανθρώπινα βλαστικά κύτταρα συμβάλλουν στην κατανόηση των παθολογικών μηχανισμών. Η χρήση γενετικής και γονιδιωματικής σε συνδυασμό με καινοτόμα πειραματικά μοντέλα, προσφέρει ελπίδα για την ανάπτυξη νέων φαρμάκων και μεθόδων θεραπείας καθιστώντας τη μελέτη των αρρυθμιών ένα πεδίο έντονης επιστημονικής προόδου [36], [37], [39], [40].

2.1.4.4 Μυοκαρδιοπάθειες

Οι μυοκαρδιοπάθειες (ΜΚΠ) αποτελούν πρωτογενείς διαταραχές του καρδιακού μυός, οι οποίες δεν αποδίδονται σε γνωστές αιτίες, όπως η στεφανιαία νόσος, η αρτηριακή υπέρταση, οι βαλβιδοπάθειες ή οι συγγενείς ανωμαλίες της καρδιάς. Διαχωρίζονται ανάλογα με τη δομική παθολογία του καρδιακού μυός και τους παθοφυσιολογικούς μηχανισμούς που εμπλέκονται, ενώ μια πιο εξειδικευμένη κατηγοριοποίηση γίνεται βάσει της αιτιολογίας τους, όταν αυτή είναι γνωστή.

Οι τέσσερις πιο συνήθεις μορφές μυοκαρδιοπαθειών είναι η διατατική, η υπερτροφική, η περιοριστική, καθώς και η αρρυθμογόνος δυσπλασία της δεξιάς κοιλίας. Συχνά, αυτές οι παθήσεις έχουν γενετική βάση και σε πολλές περιπτώσεις είναι κληρονομικές. Αποτελούν σημαντική αιτία αιφνίδιου καρδιακού θανάτου ειδικά σε νεαρούς ενήλικες και αθλητές. Για παράδειγμα, τόσο η υπερτροφική μυοκαρδιοπάθεια όσο και η αρρυθμογόνος δυσπλασία της δεξιάς κοιλίας αναγνωρίζονται ως κύριες αιτίες αιφνίδιου θανάτου σε αθλητές.

Η παθοφυσιολογία τους περιλαμβάνει δυσλειτουργία της συστολικής λειτουργίας της προσβεβλημένης κοιλίας (όπως στη διατατική και αρρυθμογόνος μορφή), δυσλειτουργία της διαστολικής φάσης (όπως στην υπερτροφική μορφή), ή έναν συνδυασμό αυτών των διαταραχών (όπως στην περιοριστική μορφή). Σε όλες τις περιπτώσεις παρατηρείται αύξηση της τελοδιαστολικής πίεσης (ΤΔΠ), η οποία οδηγεί στην εμφάνιση συμπτωμάτων καρδιακής ανεπάρκειας (ΚΑ). Η κλινική παρουσίαση αυτών των καταστάσεων περιλαμβάνει συμπτώματα όπως καρδιακή ανεπάρκεια, αρρυθμίες, αιφνίδιο καρδιακό θάνατο, συγκοπή, στηθαγχικό πόνο και θρομβοεμβολικά φαινόμενα.

Κατά τη φυσική εξέταση μπορεί να παρατηρηθούν ενδείξεις καρδιακής ανεπάρκειας καθώς και χαρακτηριστικά σημεία που σχετίζονται με τη συγκεκριμένη υποκείμενη παθολογία και παθοφυσιολογία. Για παράδειγμα, στη διατατική μυοκαρδιοπάθεια η κορυφαία καρδιακή ώση

είναι συχνά παρεκτοπισμένη, διάχυτη και μειωμένης έντασης, ενώ στην υπερτροφική μυοκαρδιοπάθεια παραμένει στη φυσιολογική θέση αλλά εμφανίζεται ισχυρή και μπορεί να είναι διπλή ή ακόμη και τριπλή.

Η διάγνωση της κύριας μορφής μυοκαρδιοπάθειας (ΜΚΠ) γίνεται στις περισσότερες περιπτώσεις μέσω απεικονιστικών μεθόδων, όπως η ακτινογραφία, το υπερηχογράφημα, η αξονική και μαγνητική τομογραφία, η αγγειογραφία και οι ραδιοϊσοτοπικές τεχνικές. Η θεραπευτική προσέγγιση περιλαμβάνει τη γενική αντιμετώπιση της καρδιακής ανεπάρκειας, ειδικές παρεμβάσεις για καταστάσεις όπως η υπερτροφική μυοκαρδιοπάθεια, την πρόληψη του αιφνίδιου θανάτου μέσω εμφύτευσης απινιδωτή, καθώς και τη μεταμόσχευση καρδιάς σε προχωρημένα στάδια. Δυστυχώς όμως δεν υπάρχουν ακόμα διαθέσιμες ειδικές θεραπείες που να αντιμετωπίζουν την αιτία της νόσου [41], [42], [43].

2.1.4.5 Φλεγμονώδεις Παθήσεις

Η μυοκαρδίτιδα και η περικαρδίτιδα αποτελούν φλεγμονώδεις παθήσεις που επηρεάζουν την καρδιά και μπορούν να προκληθούν από λοιμώδεις παράγοντες ή ανοσολογικές διαταραχές. Η μυοκαρδίτιδα είναι μια φλεγμονώδης αντίδραση που επηρεάζει τον καρδιακό μυ (μυοκάρδιο) και μπορεί να προκληθεί από διάφορα λοιμώδη ή μη λοιμώδη αίτια. Η διάγνωσή της βασίζεται σε ιστολογικά, ανοσολογικά και ανοσοϊστοχημικά ευρήματα. Αποτελεί μια από τις κύριες αιτίες αιφνίδιου καρδιακού θανάτου και διατακτικής μυοκαρδιοπάθειας. Ο όρος "φλεγμονώδης μυοκαρδιοπάθεια" έχει προταθεί πρόσφατα για να υπογραμμίσει τη στενή συσχέτιση μεταξύ της μυοκαρδίτιδας και της εξέλιξής της σε μυοκαρδιοπάθεια [44].

Σύμφωνα με τον Παγκόσμιο Οργανισμό Υγείας, η μυοκαρδίτιδα ορίζεται ως φλεγμονώδης πάθηση του μυοκαρδίου, και της οποίας η διάγνωση πραγματοποιείται μέσω βιοχημικών δεικτών, ηλεκτροκαρδιογραφικών αλλαγών, ηχοκαρδιογραφήματος και μαγνητικής τομογραφίας καρδιάς.

Η μυοκαρδίτιδα μπορεί να εκδηλωθεί σε διάφορες κλινικές μορφές, όπως οξεία, κεραυνοβόλος, υποξεία και χρόνια. Η οξεία μυοκαρδίτιδα ορίζεται ως η κατάσταση όπου η διάγνωση γίνεται μέσα σε διάστημα μικρότερο του ενός μήνα από την εμφάνιση των πρώτων συμπτωμάτων. Η κεραυνοβόλος μυοκαρδίτιδα αποτελεί μια ακραία και ταχέως εξελισσόμενη μορφή της οξείας μυοκαρδίτιδας, η οποία συνοδεύεται συνήθως από καρδιογενές σοκ και απαιτεί την εφαρμογή ενδογενούς ή μηχανικής υποστήριξης της κυκλοφορίας. Η υποξεία μυοκαρδίτιδα περιγράφει τη συνεχιζόμενη βλάβη του καρδιακού μυός λόγω της παρουσίας μιας επίμονης ή υποτροπιάζουσας φλεγμονώδους διαδικασίας, αλλά μπορεί επίσης να αναφέρεται στην ανάρρωση από μια προηγούμενη ενεργό μυοκαρδίτιδα. Η υποξεία μορφή ορίζεται συχνά ως η κατάσταση όπου τα συμπτώματα διαρκούν από 1 έως 3 μήνες από την έναρξή τους. Όταν τα συμπτώματα συνεχίζουν για διάστημα μεγαλύτερο του ενός μηνός, τότε

η κατάσταση χαρακτηρίζεται ως χρόνια φλεγμονώδης μυοκαρδιοπάθεια, ενώ υπάρχει και αλληλοκάλυψη με τη διάγνωση της υποξείας μυοκαρδίτιδας [45].

Η περικαρδίτιδα είναι μια φλεγμονώδης κατάσταση που επηρεάζει το περικάρδιο, τον λεπτό ιστό που περιβάλλει την καρδιά. Οι αιτίες της μπορεί να ποικίλλουν, περιλαμβάνοντας ιογενείς λοιμώξεις, αυτοάνοσες νόσους, τραυματισμούς ή ακόμη και εξάπλωση από άλλες περιοχές του σώματος. Μπορεί να εμφανιστεί είτε με οξεία είτε με χρόνια μορφή, και συχνά συνοδεύεται από πόνο στο στήθος, ο οποίος τείνει να επιδεινώνεται με την αναπνοή ή την κατάκλιση [46], [47].

Το πιο συχνό σύμπτωμα της οξείας περικαρδίτιδας, που εμφανίζεται ξαφνικά, είναι ο θωρακικός πόνος, ο οποίος συνήθως εντείνεται με τη βαθιά εισπνοή. Σε ορισμένες περιπτώσεις, ο πόνος είναι διάχυτος και μπορεί να θυμίζει τον πόνο που προκαλείται από έμφραγμα του μυοκαρδίου. Ο πόνος μπορεί να επιδεινώνεται όταν το άτομο ξαπλώνει και να ακτινοβολεί στους ώμους ή στο πάνω μέρος της πλάτης. Κάποιοι ασθενείς με περικαρδίτιδα μπορεί επίσης να εμφανίσουν πυρετό [48].

Η ανίχνευση της περικαρδίτιδας στηρίζεται στο ιατρικό ιστορικό του ασθενούς, στη φυσική εξέταση που πραγματοποιεί ο καρδιολόγος, καθώς και σε αιματολογικές και απεικονιστικές εξετάσεις, όπως αυτές είναι το ηλεκτροκαρδιογράφημα, το υπερηχοκαρδιογράφημα και η ακτινογραφία θώρακος [46], [47].

2.1.4.6 Βαλβιδοπάθειες

Οι βαλβίδες της καρδιάς, όπως και οι κοιλότητες, είναι τέσσερις, η αορτική, η μιτροειδής, η πνευμονική και η τριγλώχινα. Λειτουργούν σαν πόρτες, οι οποίες ανοίγουν και κλείνουν στον κατάλληλο χρόνο για να επιτρέψουν τη μεταφορά του αίματος από τη μια κοιλότητα στην άλλη. Αν δεν κλείσουν πλήρως ένα μέρος του αίματος επιστρέφει στην προηγούμενη κοιλότητα προκαλώντας ανεπάρκεια. Αν δεν ανοίγουν επαρκώς δυσχεραίνεται η ροή του αίματος προς την επόμενη κοιλότητα, οδηγώντας σε στένωση. Σε αμφότερες τις περιπτώσεις, η διαταραχή στη ροή του αίματος δημιουργεί έναν χαρακτηριστικό ήχο, γνωστό ως φύσημα κατά την ακρόαση. Τέτοιοι ήχοι μπορούν να δημιουργηθούν και από άλλες παθολογικές καταστάσεις, όπως η μεσοκοιλιακή επικοινωνία.

Οι παθήσεις των βαλβίδων κατατάσσονται με βάση τη σοβαρότητά τους σε ελαφρές, μέτριες και σοβαρές. Οι κυριότερες είναι, η στένωση και ανεπάρκεια της αορτικής βαλβίδας και η ανεπάρκεια και στένωση της μιτροειδούς βαλβίδας [49], [50].

2.1.4.6.1 Στένωση αορτικής βαλβίδας

Αν η βαλβίδα παρουσιάζει εκ γενετής δυσλειτουργία (δίπτυχη), η στένωση (με ή χωρίς ανεπάρκεια) εμφανίζεται συνήθως γύρω στην ηλικία των 50 ετών. Αντίθετα, σε περιπτώσεις φυσιολογικής (τρίπτυχης) βαλβίδας, η στένωση συνήθως εμφανίζεται αργότερα, συνήθως μετά τα 60 έτη. Το πρόβλημα στη βαλβίδα έγκειται στη συσσώρευση ασβεστίου και λιπιδίων, γεγονός που παρεμποδίζει το άνοιγμά της.

Τα τυπικά συμπτώματα της πάθησης περιλαμβάνουν στηθάγχη (πόνος στο στήθος) κατά τη σωματική άσκηση, δύσπνοια που μπορεί να εξελιχθεί σε πνευμονικό οίδημα, καθώς και λιποθυμικά επεισόδια. Αυτά τα συμπτώματα μοιάζουν με εκείνα της στεφανιαίας νόσου. Επειδή η στένωση εξελίσσεται σταδιακά, οι ασθενείς μειώνουν ασυνείδητα τη δραστηριότητά τους, με αποτέλεσμα να μην αναφέρουν σαφή συμπτώματα, γεγονός που δυσκολεύει τη διάγνωση. Η διάγνωση γίνεται συνήθως με την ακρόαση (χαρακτηριστικός ήχος) και επιβεβαιώνεται με το υπερηχογράφημα, το οποίο αξιολογεί και τη σοβαρότητα της κατάστασης.

Η θεραπεία μέχρι πρόσφατα ήταν αποκλειστικά χειρουργική, με αντικατάσταση της βαλβίδας με μεταλλική ή βιολογική (χοίρεια). Ωστόσο, μια νέα μέθοδος, γνωστή ως TAVI, έχει αναπτυχθεί. Αυτή η διαδικασία, που μοιάζει με την αγγειοπλαστική, περιλαμβάνει την εισαγωγή μιας ειδικής βαλβίδας μέσω της μηριαίας αρτηρίας, η οποία στη συνέχεια τοποθετείται και εκπτύσσεται στη θέση της παλιάς βαλβίδας. Η μέθοδος TAVI θεωρείται πλέον σχεδόν ισοδύναμη με τη χειρουργική επέμβαση και εφαρμόζεται σε ηλικιωμένους ή βαρέως πάσχοντες ασθενείς που πληρούν συγκεκριμένα κριτήρια [49], [50].

2.1.4.6.2 Ανεπάρκεια αορτικής βαλβίδας

Το πρόβλημα μπορεί να προέρχεται από δυσλειτουργία των ίδιων των γλωχίνων, όπως στη δίπτυχη βαλβίδα, την ενδοκαρδίτιδα ή τον ρευματικό πυρετό, ή από διαστολή ή ανεύρυσμα της ανιούσας αορτής, που συνδέονται με καταστάσεις όπως το σύνδρομο Marfan, την υπέρταση ή τον διαχωρισμό της αορτής. Η κατάσταση μπορεί να εκδηλωθεί είτε οξεία είτε χρόνια. Τα κύρια συμπτώματα περιλαμβάνουν δύσπνοια που μπορεί να οδηγήσει σε πνευμονικό οίδημα, στηθάγχη κατά την προσπάθεια, και έντονη αίσθηση παλμών. Η διάγνωση και η αξιολόγηση της σοβαρότητας πραγματοποιούνται μέσω υπερηχογραφήματος, ενώ η θεραπεία είναι κατά κύριο λόγο χειρουργική, με αντικατάσταση της βαλβίδας. Συχνά απαιτείται και η αντικατάσταση της ανιούσας αορτής με ειδικό μόσχευμα [49], [50].

2.1.4.6.3 Ανεπάρκεια μιτροειδούς βαλβίδας

Η μιτροειδής βαλβίδα αποτελεί έναν σύνθετο μηχανισμό. Μερικές πιθανές αιτίες για την ανεπάρκεια μπορεί να περιλαμβάνουν την πρόπτωση, την ενδοκαρδίτιδα, τον ρευματικό

πυρετό, την ασβέστωση του δακτυλίου, καθώς και τη διατακτική ή ισχαιμική μυοκαρδιοπάθεια. Η πρόπτωση της μιτροειδούς είναι σχετικά συχνή, επηρεάζοντας το 2.4% του πληθυσμού, με τις γυναίκες να παρουσιάζουν διπλάσια συχνότητα εμφάνισης σε σχέση με τους άνδρες. Ωστόσο, οι άνδρες φαίνεται να έχουν χειρότερη πρόγνωση.

Στις περισσότερες περιπτώσεις, οι γυναίκες διαγιγνώσκονται με ήπια μορφή πρόπτωσης σε ηλικίες μεταξύ 20 και 30 ετών, χωρίς να εξελίσσεται σε σοβαρή ανεπάρκεια. Αντίθετα, οι άνδρες (και ορισμένες γυναίκες) που διαγιγνώσκονται σε μεγαλύτερη ηλικία (40-70 ετών), με διαφορετική παθολογία της βαλβίδας, όπως μυξωματώδη εκφύλιση, διατρέχουν μεγαλύτερο κίνδυνο σοβαρής ανεπάρκειας. Η διάγνωση της πρόπτωσης βασίζεται στα ακουστικά ευρήματα και το υπερηχογράφημα, ενώ συνιστάται η αποφυγή βαρέων βαρών.

Τα συμπτώματα της χρόνιας σοβαρής ανεπάρκειας της μιτροειδούς περιλαμβάνουν δύσπνοια, οίδημα στα κάτω άκρα, πνευμονικό οίδημα και εύκολη κόπωση, καθώς και κολπική μαρμαρυγή. Η οξεία σοβαρή ανεπάρκεια, συχνά από έμφραγμα ή ενδοκαρδίτιδα, μπορεί να προκαλέσει καρδιογενές σοκ με υπόταση και σοβαρό πνευμονικό οίδημα.

Η διάγνωση της ανεπάρκειας της μιτροειδούς γίνεται με υπερηχογράφημα, το οποίο αξιολογεί και τη σοβαρότητα της κατάστασης. Η θεραπευτική προσέγγιση μπορεί να περιλαμβάνει χειρουργική αντικατάσταση της βαλβίδας με μεταλλική ή βιολογική βαλβίδα, καθώς και επιδιόρθωση είτε χειρουργικά είτε διαδερμικά με τεχνικές όπως το mitra-clip [49], [50].

2.1.4.6.4 Στένωση μιτροειδούς βαλβίδας

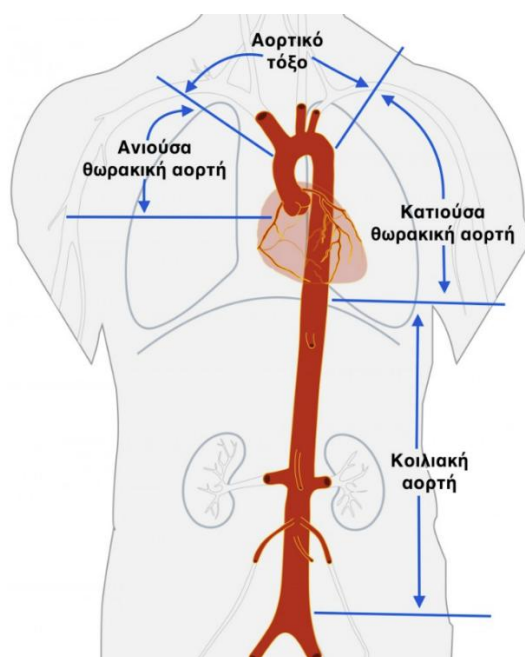
Η κυριότερη αιτία αυτής της κατάστασης είναι ο ρευματικός πυρετός, ο οποίος έχει γίνει σπάνιος στις σύγχρονες ανεπτυγμένες χώρες. Τα κύρια συμπτώματα περιλαμβάνουν δύσπνοια που μπορεί να φτάσει σε πνευμονικό οίδημα, έντονη κόπωση, λιποθυμικά επεισόδια, εμβολικά φαινόμενα (όπως εγκεφαλικό επεισόδιο) και κολπική μαρμαρυγή. Η διάγνωση γίνεται με υπερηχογράφημα, ενώ και η θεραπεία καθορίζεται με την ίδια μέθοδο. Ορισμένοι ασθενείς μπορεί να είναι κατάλληλοι για διαδερμική βαλβιδοπλαστική με μπαλόνι, ενώ άλλοι ενδέχεται να χρειάζονται χειρουργική αντικατάσταση της βαλβίδας [49], [50].

2.2 Ανατομία και Λειτουργία της Αορτής

Η αορτή ως η μεγαλύτερη αρτηρία στο ανθρώπινο σώμα, παίζει κρίσιμο ρόλο στη μεταφορά του οξυγονωμένου αίματος από την καρδιά προς τα υπόλοιπα όργανα και τους ιστούς. Η δομή της χαρακτηρίζεται ως πολύπλοκη καθώς με διάφορες προσαρμογές εξασφαλίζεται η αποτελεσματική κυκλοφορία του αίματος [18]. Παρακάτω αναλύεται η ανατομία και η ιστολογία της αορτής καθώς επίσης παρουσιάζονται οι παθολογικές της καταστάσεις.

2.2.1 Εισαγωγή στην Ανατομία της Αορτής

Η αορτή χωρίζεται στη θωρακική και κοιλιακή αορτή (**Εικόνα 2.9**). Η θωρακική αορτή περιλαμβάνει την ανιούσα αορτή, το αορτικό τόξο και την κατιούσα θωρακική αορτή [51]. Η ανιούσα θωρακική αορτή έχει μήκος περίπου 5 cm και εκκινεί από τον ινώδη δακτύλιο της αορτής, περιλαμβάνοντας τους κόλπους του Valsalva [52] και το σωληνώδες τμήμα που αρχίζει από αυτούς και φτάνει μέχρι την ανώνυμη αρτηρία. Αυτή η αορτική ενότητα βρίσκεται ολόκληρη μέσα στο ινώδες περικάρδιο, το οποίο τη διαχωρίζει από το αορτικό τόξο στο επίπεδο του άνω χείλους του δεύτερου δεξιού πλευρικού χόνδρου.



Εικόνα 2.9: Ανατομία αορτής [51]

Το αορτικό τόξο είναι το τμήμα της αορτής από το οποίο εκφύονται τα κύρια αγγεία, δηλαδή η ανώνυμη αρτηρία, η αριστερή καρωτίδα και η αριστερή υποκλείδια αρτηρία, που τροφοδοτούν το κεφάλι, τον τράχηλο και τα άνω άκρα [53]. Με μήκος περίπου 4 cm, το αορτικό τόξο εκτείνεται από την πρόσφυση του ινώδους περικαρδίου έως το αρτηριακό σύνδεσμο ή τον ισθμό της αορτής.

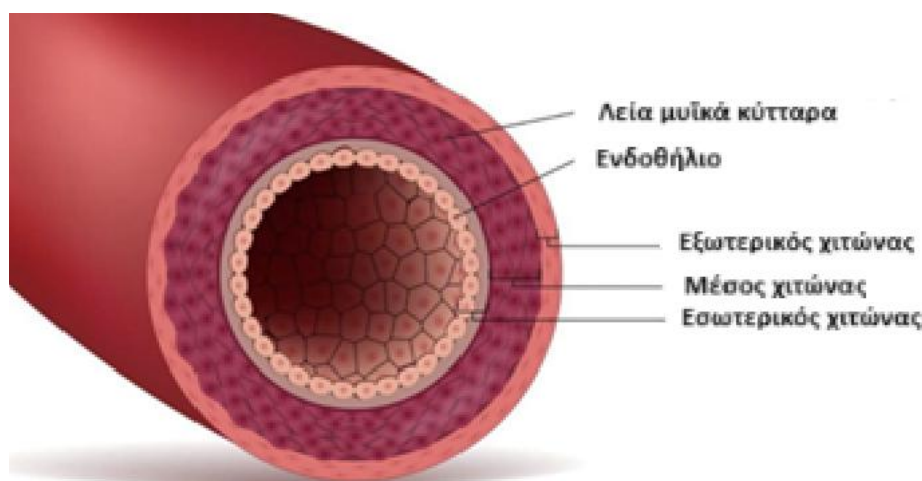
Η κατιούσα θωρακική αορτή ξεκινά στο επίπεδο της στερνικής γωνίας και του κάτω χείλους του τέταρτου θωρακικού σπονδύλου, αμέσως μετά την αριστερή υποκλείδια αρτηρία, και συνεχίζεται μέχρι το διάφραγμα, στο ύψος του δωδέκατου θωρακικού σπονδύλου. Από αυτήν εκφύονται σπλαχνικοί κλάδοι, όπως οι βρογχικές και οισοφαγικές αρτηρίες, και τοιχικοί κλάδοι, όπως οι μεσοπλεύριες αρτηρίες από την τρίτη έως τη δωδέκατη.

Η θωρακοκοιλιακή αορτή περιλαμβάνει τόσο την κατιούσα θωρακική όσο και την κοιλιακή αορτή. Η κοιλιακή αορτή ξεκινά από το αορτικό τρήμα του διαφράγματος και πορεύεται κατά μήκος της σπονδυλικής στήλης για περίπου 15 cm, μέχρι να διχαστεί στις κοινές λαγόνιες αρτηρίες στο ύψος του τέταρτου οσφυϊκού σπονδύλου.

Η ρίζα της αορτής περιλαμβάνει την έξοδο της αριστερής κοιλίας, την αορτική βαλβίδα και τους τρεις κόλπους του Valsalva, όπου εντοπίζονται τα στόμια των στεφανιαίων αρτηριών. Η φυσιολογική διάμετρος της αορτής ποικίλλει ανάλογα με την ηλικία, το φύλο και την επιφάνεια σώματος, με τους άνδρες να έχουν ελαφρώς μεγαλύτερη διάμετρο σε σύγκριση με τις γυναίκες. Η μέση διάμετρος της θωρακικής αορτής είναι περίπου 31 mm στη ρίζα της αορτής, 32 mm στην ανιούσα αορτή και 28 mm στην κατιούσα θωρακική αορτή [51], [52], [54], [55], [56].

2.2.2 Ιστολογία

Η αορτή είναι η μεγαλύτερη αρτηρία του σώματος και ανήκει στον τύπο των ελαστικών αρτηριών. Το τοίχωμά της αποτελείται από τρία στρώματα, τον έσω χιτώνα (tunica intima), τον μέσο χιτώνα (tunica media) και τον έξω χιτώνα (tunica externa) (**Εικόνα 2.10**). Το εσωτερικό τμήμα του έσω χιτώνα αποτελείται από τα ενδοθηλιακά κύτταρα που βρίσκονται πάνω σε μια λεπτή βασική μεμβράνη. Κάτω από το ενδοθήλιο υπάρχει ο υποενδοθηλιακός ιστός που περιλαμβάνει τις ινοβλάστες, τις κολλαγονώδη ίνες και τη βλεννώδη βασική ουσία. Το έσω ελαστικό πέταλο οριοθετεί το εξωτερικό άκρο του έσω χιτώνα.



Εικόνα 2.10: Στρώματα αορτής [57]

Ο μέσος χιτώνας έχει πάχος περίπου 1 mm και περιέχει την ελασίνη, τις λείες μυϊκές ίνες, το κολλαγόνο και τη βασική ουσία. Η παρουσία των ελαστικών ινών που είναι διατεταγμένες

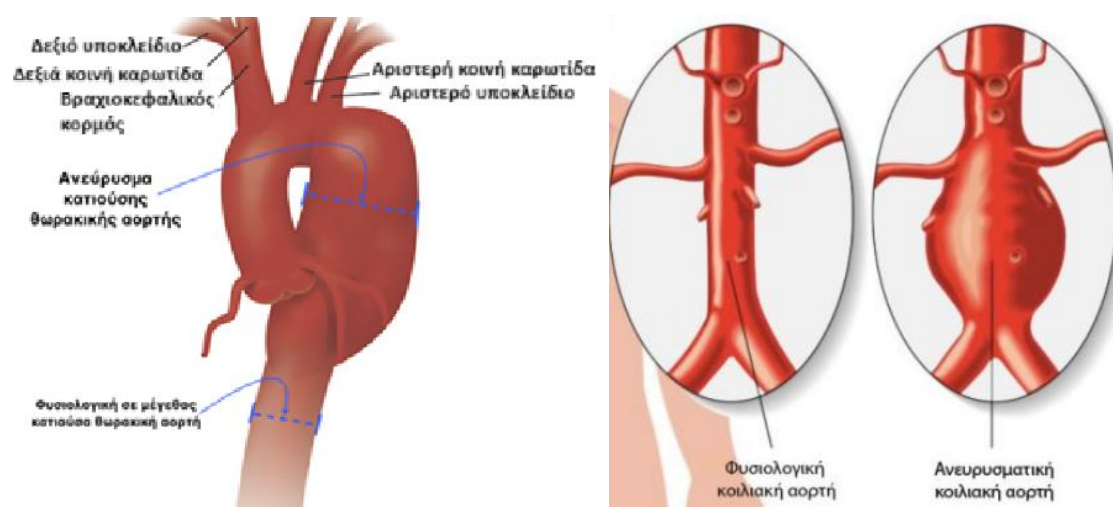
κυκλοτερώς είναι χαρακτηριστική της αορτής και τη διαφοροποιεί από τις μικρότερες αρτηρίες που είναι κυρίως μυϊκού τύπου.

Ο έξω χιτώνας (tunica externa) αποτελείται από συνδετικό ιστό με κύρια συστατικά τους ινοβλάστες, λίγες κολλαγονώδεις ίνες, ελαστίνη και τη βασική ουσία. Στο στρώμα αυτό βρίσκονται τα νεύρα των αγγείων (nervi vasorum) και τα αγγεία των αγγείων (vasa vasorum) τα οποία παρέχουν θρεπτικά συστατικά στον έξω και τον μέσο χιτώνα. Η διατροφή του έσω χιτώνα γίνεται μέσω διάχυσης από το αίμα που κυκλοφορεί στο εσωτερικό της αορτής ενώ το έξω ελαστικό πέταλο διαχωρίζει τον μέσο από τον έξω χιτώνα [58].

2.2.3 Παθολογικές Καταστάσεις της Αορτής

Οι παθήσεις της αορτής περιλαμβάνουν ανωμαλίες στη δομή, όπως η αύξηση της διαμέτρου, γνωστή και ως ανεύρυσμα, την παρουσία σοβαρών αλλοιώσεων όπως είναι η εκτεταμένη αθηρωμάτωση καθώς και οξείες καταστάσεις όπως η ρήξη ή ο διαχωρισμός της αορτής [59].

Ως ανεύρυσμα χαρακτηρίζεται η διόγκωση της αορτής (**Εικόνα 2.11**) όπου η διάμετρος της αυξάνεται κατά τουλάχιστον 1.5 φορά σε σύγκριση με το φυσιολογικό της μέγεθος. Εμφανίζεται σε οποιοδήποτε σημείο της αορτής και συνδέεται με σημαντικά αυξημένο κίνδυνο σοβαρών επιπλοκών όπως είναι ο διαχωρισμός ή η ρήξη της αορτής και καταστάσεις που έχουν υψηλή θνητότητα και απαιτούν άμεση ιατρική παρέμβαση. Από τα ανωτέρω συμπεραίνουμε ότι η έγκαιρη διάγνωση είναι κρίσιμη για την επιβίωση του ασθενούς. Για αυτόν τον λόγο οι σύγχρονες οδηγίες συνιστούν προληπτικό έλεγχο σε πληθυσμιακές ομάδες υψηλού κινδύνου όπως άτομα με υπέρταση, καπνιστές, παχύσαρκους και άτομα με οικογενειακό ιστορικό αορτικών παθήσεων.



Εικόνα 2.11: Ανευρύσματα αορτής [59]

Τα ανευρύσματα ταξινομούνται με βάση τη μορφολογία, τον τύπο, και την εντόπισή τους. Πιο συγκεκριμένα τα ανευρύσματα κατηγοριοποιούνται με βάση τη μορφή τους σε δύο τύπους, τα ατρακτοειδή που χαρακτηρίζονται από ομοιόμορφη κυκλική διάταση η οποία επηρεάζει ολόκληρη την περιφέρεια της αρτηρίας και τα σακοειδή όπου η διάταση είναι τοπική και ασύμμετρη επηρεάζοντας μόνο ένα τμήμα της περιφέρειας της αρτηρίας [60].

Η ζημιά που προκαλείται στο τοίχωμα των αληθινών ανευρυσμάτων επηρεάζει και τους τρεις χιτώνες του αορτικού τοιχώματος (tunica intima, tunica media, tunica externa). Τα εκφυλιστικά αυτά ανευρύσματα συνδέονται συχνά με την αθηρωμάτωση ενώ ορισμένα ανευρύσματα που οφείλονται σε λοιμώξεις κατατάσσονται επίσης σε αυτή την κατηγορία. Το τοίχωμα ενός ψευδοανευρύσματος αποτελείται εξολοκλήρου από τον έξω χιτώνα της αορτής και τον περιβάλλοντα ινώδη ιστό. Σε ορισμένες περιπτώσεις μπορεί να απουσιάζει ακόμη και ο έξω χιτώνας με αποτέλεσμα το ψευδοανεύρυσμα να συγκρατείται από εξαγγειωμένο αίμα που περιβάλλεται από τους γύρω ιστούς. Αυτού του τύπου τα ανευρύσματα μπορεί να προκληθούν από τραυματικές βλάβες της θωρακικής αορτής καθώς και από διαβρωτικές μάζες που οφείλονται σε νεοπλάσματα ή φλεγμονές [61].

Τα ανευρύσματα της θωρακικής αορτής ταξινομούνται επίσης και βάσει της ανατομικής τους θέσης, μπορεί δηλαδή να επηρεάζουν ένα ή περισσότερα τμήματα, όπως τη ρίζα της αορτής, την ανιούσα αορτή, το τόξο, την κατιούσα θωρακική αορτή ή την θωρακοκοιλιακή αορτή. Περίπου το 25% των ανευρυσμάτων της αορτής εντοπίζεται στη θωρακική περιοχή, με το 40-50% αυτών να αφορά την κατιούσα αορτή [62], [63], [64]. Επιπλέον αξίζει να σημειωθεί ότι το 20-25% των ασθενών με ανεύρυσμα θωρακικής αορτής παρουσιάζει επίσης ανεύρυσμα στην κοιλιακή αορτή [65], [66], [67].

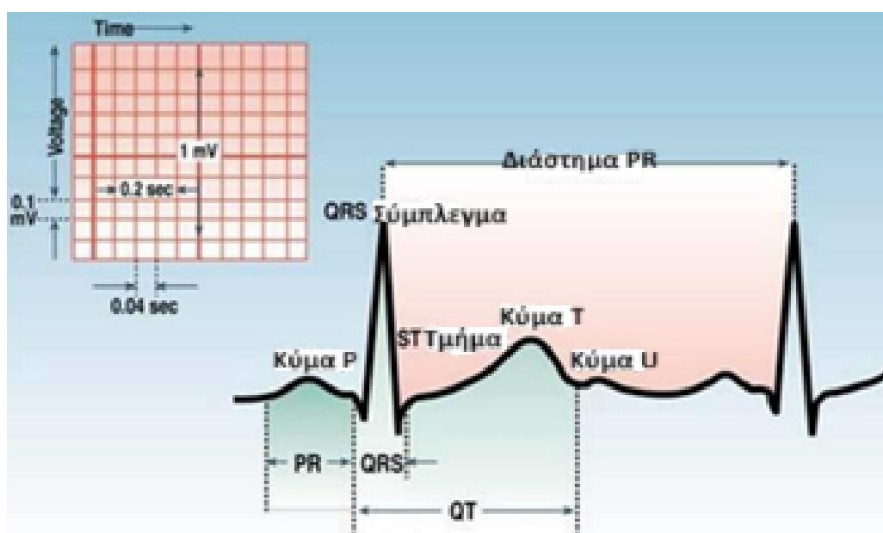
2.3 Εισαγωγή στις Μεθόδους Απεικόνισης και Διαγνωστικής Αξιολόγησης

Η καρδιά αποτελεί το κύριο όργανο του κυκλοφορικού συστήματος καθώς η σωστή λειτουργία της είναι κρίσιμη για τη διατήρηση της υγείας. Η εκτίμηση της καρδιακής απόδοσης είναι ουσιώδης τόσο για τη διάγνωση όσο και για την παρακολούθηση διαφόρων καρδιαγγειακών νοσημάτων. Η καρδιακή απεικόνιση παρέχει σημαντικά δεδομένα σχετικά με τη μορφολογία και τη λειτουργικότητα της καρδιάς, διευκολύνοντας τους ιατρούς να εντοπίζουν ανωμαλίες και να σχεδιάζουν τη βέλτιστη θεραπευτική αντιμετώπιση [68].

2.3.1 Ηλεκτροκαρδιογράφημα

Το ηλεκτροκαρδιογράφημα (**Εικόνα 2.12**) αποτελεί θεμελιώδες εργαλείο της καρδιολογικής αξιολόγησης καθώς για περισσότερο από έναν αιώνα παραμένει μία από τις πιο σημαντικές διαγνωστικές μεθόδους στην ιατρική. Μέσω αυτής της εξέτασης αναλύονται τα ηλεκτρικά δυναμικά που παράγονται κατά τη λειτουργία της καρδιάς και τα οποία παίζουν καθοριστικό

ρόλο στη διάγνωση παθήσεων όπως η στεφανιαία νόσος, το οξύ έμφραγμα του μυοκαρδίου, η κολπική μαρμαρυγή, διάφορες αρρυθμίες και άλλες καρδιολογικές διαταραχές. Σε πολλές περιπτώσεις η λεπτομερής λήψη ιατρικού ιστορικού, η προσεκτική κλινική εξέταση και η διενέργεια ηλεκτροκαρδιογραφήματος επαρκούν για την ακριβή διάγνωση.



Εικόνα 2.12: Το Ηλεκτροκαρδιογράφημα [69]

Το ηλεκτροκαρδιογράφημα είναι μια απόλυτα ασφαλής εξέταση η οποία χρησιμοποιεί μια συσκευή τον ηλεκτροκαρδιογράφο για να καταγράψει τα ηλεκτρικά σήματα που παράγονται κατά τη λειτουργία της καρδιάς. Η συσκευή αυτή περιλαμβάνει μια κεντρική μονάδα και ένα σύστημα καλωδίων με δέκα ηλεκτρόδια. Τα ηλεκτρόδια αυτά τοποθετούνται σε συγκεκριμένα σημεία του σώματος, τέσσερα στα άνω και κάτω άκρα, ενώ τα υπόλοιπα έξι στο θωρακικό τοίχωμα. Με αυτή τη διαδικασία παράγεται ένα πλήρες ηλεκτροκαρδιογράφημα 12 απαγωγών [69].

Η έξοδος του ηλεκτροκαρδιογράφου παρουσιάζεται ως μια γραφική αναπαράσταση όπου ο οριζόντιος άξονας (X) αντιπροσωπεύει τον χρόνο ενώ ο κατακόρυφος άξονας (Y) απεικονίζει τη διαφορά δυναμικού που καταγράφεται σε κάθε χρονική στιγμή. Το ηλεκτροκαρδιογράφημα χαρακτηρίζεται από επαναλαμβανόμενα και πανομοιότυπα μοτίβα απεικόνισης. Παρακάτω ακολουθεί η παρουσίαση των σημείων με κλινική σημασία στο ηλεκτροκαρδιογράφημα μαζί με τις αντίστοιχες φυσιολογικές τους τιμές.

2.3.1.1 Το κύμα P

Το κύμα P είναι ένα μικρό θετικό ή αρνητικό σήμα που εμφανίζεται αμέσως πριν από το σύμπλεγμα QRS. Σε φυσιολογικές συνθήκες το κύμα P έχει πλάτος που δεν ξεπερνά τα

0.25 mV και διάρκεια μικρότερη από 0.12 sec. Αποκλίσεις από αυτές τις τιμές είτε στο πλάτος είτε στη διάρκεια μπορεί να υποδηλώνουν παθολογικές καρδιακές καταστάσεις.

2.3.1.2 Το διάστημα PR

Το διάστημα PR αναφέρεται στο χρονικό διάστημα από την έναρξη του κύματος P έως την έναρξη του συμπλέγματος QRS. Στους ενήλικες το φυσιολογικό εύρος του διαστήματος PR κυμαίνεται μεταξύ 0.12 και 0.2 sec.

Το διάστημα αυτό αναφέρεται στον χρόνο που απαιτείται για τη μετάδοση του ηλεκτρικού σήματος από τους κόλπους στον κολποκοιλιακό κόμβο (AV κόμβο) και από εκεί στο δεμάτιο του Hiss και το δίκτυο Purkinje. Το διάστημα PR ξεκινά από την αρχή του επάρματος P και ολοκληρώνεται στην έναρξη του συμπλέγματος QRS. Η διάρκεια του διαστήματος PR έχει κρίσιμη σημασία στην ανάλυση του ηλεκτροκαρδιογραφήματος (ΗΚΓ) και κυμαίνεται μεταξύ 0.12 και 0.20 sec. Εάν ξεφεύγει από αυτά τα όρια, υποδεικνύει πιθανή παθολογία. Υπάρχει επίσης μία ζώνη που ονομάζεται τομέας PR, η οποία διαφέρει από το διάστημα PR. Αυτή ξεκινά στο σημείο όπου το έπαρμα P εφάπτεται με την ισοηλεκτρική γραμμή και φτάνει μέχρι το τέλος του οριζώντιου τμήματός της, πριν αρχίσει το σύμπλεγμα QRS. Το συγκεκριμένο τμήμα είναι σημαντικό για τη διάγνωση καταστάσεων όπως η περικαρδίτιδα.

2.3.1.3 Το σύμπλεγμα QRS

Το σύμπλεγμα QRS αποτελείται από τρία κύματα, το Q, το R και το S. Το σύμπλεγμα QRS έχει διάρκεια από 0.12 έως 0.20 sec και αντιπροσωπεύει την εκπόλωση των κοιλιών κατά την οποία πραγματοποιείται η εκπόλωση και των δύο κοιλιών. Το κύμα Q είναι η πρώτη αρνητική απόκλιση που εμφανίζεται μετά το διάστημα PR και το κύμα R αποτελεί την πρώτη θετική απόκλιση που εμφανίζεται μετά το κύμα Q ή αν δεν υπάρχει κύμα Q ακολουθεί αμέσως μετά το διάστημα PR. Τέλος το κύμα S είναι η πρώτη αρνητική απόκλιση που έπεται του R και εκτείνεται κάτω από την ισοηλεκτρική γραμμή.

2.3.1.4 Το τμήμα ST

Το τμήμα ST αντιπροσωπεύει το χρονικό διάστημα που μεσολαβεί από την ολοκλήρωση της κοιλιακής εκπόλωσης έως την έναρξη της κοιλιακής επαναπόλωσης. Ξεκινά από το τέλος του συμπλέγματος QRS και ολοκληρώνεται στην αρχή του επάρματος T. Η φυσιολογική διάρκεια αυτού του διαστήματος είναι μικρότερη από 0.2 δευτερόλεπτα.

2.3.1.5 Το κύμα

Το κύμα T αναπαριστά την επαναπόλωση των κοιλιών. Αρχίζει από το τέλος του διαστήματος ST και ολοκληρώνεται όταν συναντά την ισοηλεκτρική γραμμή. Το κύμα T εμφανίζει μικρότερη κλίση κατά το πρώτο του μισό μέχρι να φτάσει στο μέγιστο και μεγαλύτερη κλίση στο δεύτερο μισό. Χαρακτηρίζεται ως μη συμμετρικό κάτι που θεωρείται φυσιολογικό.

2.3.1.6 Το διάστημα QT

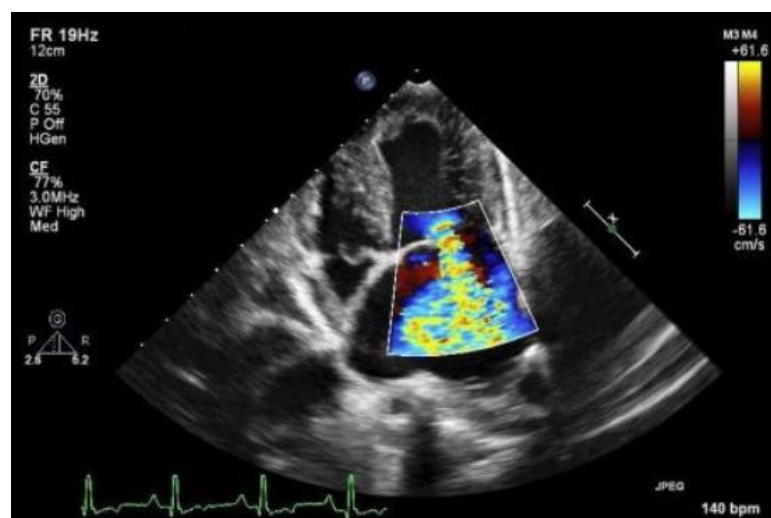
Το διάστημα QT αποτελεί δείκτη της συνολικής ηλεκτρικής δραστηριότητας των κοιλιών. Ξεκινά από την έναρξη του συμπλέγματος QRS και ολοκληρώνεται στο τέλος του κύματος T. Η διάρκεια του QT είναι μία από τις πιο σημαντικές παραμέτρους για την αξιολόγηση αρρυθμιών καθώς περιλαμβάνει και την ανερέθιστη περίοδο των κοιλιών.

2.3.1.7 Το κύμα U

Το κύμα U συνήθως εμφανίζεται ως μια ελαφριά ανωμαλία ή ύψωση στο τέλος ή αμέσως μετά το κύμα T. Δεν είναι πάντοτε ορατό, ενώ η αιτιολογία του παραμένει αδιευκρίνιστη [70].

2.3.2 Υπερηχοκαρδιογράφημα

Η υπερηχοκαρδιογραφία (ή ηχοκαρδιογραφία) (**Εικόνα 2.13**) αποτελεί ένα από τα πιο σημαντικά εργαλεία για την απεικόνιση και την αξιολόγηση της καρδιακής λειτουργίας σε πραγματικό χρόνο. Η αξία του είναι τόσο μεγάλη και η χρήση του τόσο διαδεδομένη ώστε θεωρείται αναπόσπαστο μέρος της καρδιολογικής αξιολόγησης [71].



Εικόνα 2.13: Το Υπερηχοκαρδιογράφημα [72]

Από τα ανωτέρω συμπεραίνουμε ότι η διάγνωση ενός καρδιακού νοσήματος είναι αδιανόητη χωρίς να έχει προηγηθεί μια λεπτομερής και πλήρως τεκμηριωμένη υπερηχοκαρδιογραφική αξιολόγηση. Παράλληλα, η χρήση της υπερηχοκαρδιογραφίας δεν περιορίζεται μόνο στη διάγνωση καθώς παρέχει κρίσιμες προγνωστικές πληροφορίες και διαδραματίζει κεντρικό ρόλο στη διαδικασία λήψης σημαντικών θεραπευτικών αποφάσεων. Ενδεικτικά η αξιολόγηση του κλάσματος εξώθησης επιτρέπει τον διαχωρισμό των ασθενών με καρδιακές παθήσεις σε ομάδες υψηλού ή χαμηλού κινδύνου καθοδηγώντας την επιλογή κατάλληλων θεραπευτικών στρατηγικών όπως εξειδικευμένα φαρμακευτικά σχήματα, εμφύτευση βηματοδοτών ή απινιδωτών καθώς και χειρουργικές επεμβάσεις, ανάλογα με τη συγκεκριμένη πάθηση [73], [74].

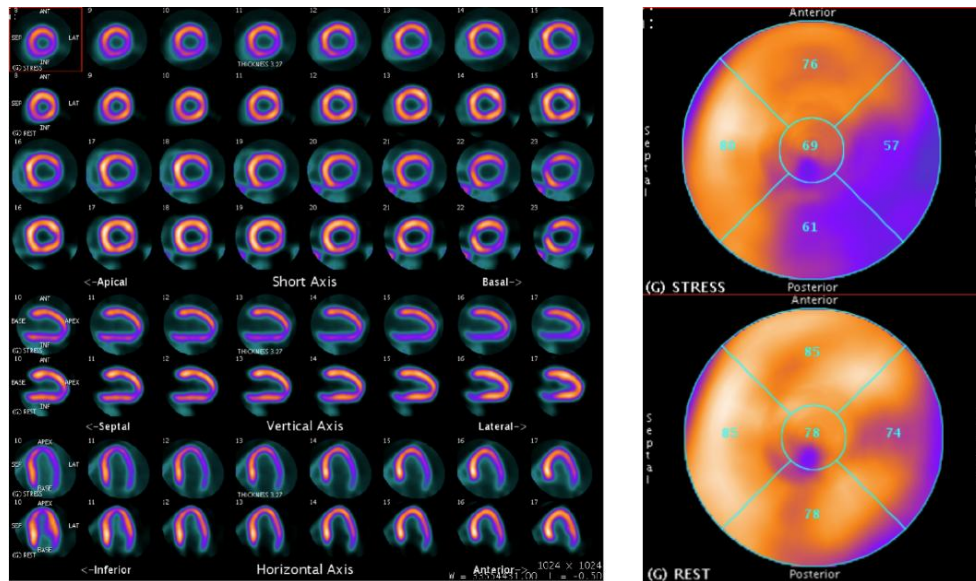
Η διαθωρακική υπερηχοκαρδιογραφία αποτελεί την κύρια διαγνωστική μέθοδο που εφαρμόζεται πριν από τη χρήση πιο σύγχρονων τεχνικών. Μέσω αυτής της εξέτασης, είναι εφικτή η αξιολόγηση των διαστάσεων των καρδιακών κοιλοτήτων, του πάχους των τοιχωμάτων, της κίνησης των επιμέρους τμημάτων της καρδιάς, καθώς και της λειτουργίας των βαλβίδων. Επιπλέον, δίνει τη δυνατότητα ανίχνευσης περικαρδιακής συλλογής [75], [76], [77].

Η εξέταση αυτή βασίζεται στη χρήση ηχητικών κυμάτων υψηλής συχνότητας προκειμένου να δημιουργηθούν εικόνες που απεικονίζουν τις βαλβίδες και τις δομές της καρδιάς. Μέσω αυτών ο καρδιολόγος μπορεί να αξιολογήσει τη λειτουργικότητα, τις διαστάσεις και τις πιέσεις της καρδιάς επιτρέποντας την ανίχνευση πιθανών προβλημάτων στις βαλβίδες, στις κοιλότητες και στην κινητικότητα των τοιχωμάτων ή στο περικάρδιο [78].

Κατά τη διάρκεια της εξέτασης ο ασθενής πρέπει να βρίσκεται ξαπλωμένος σε πλάγια θέση στηριζόμενος στην αριστερή πλευρά του. Στη συνέχεια ο γιατρός εφαρμόζει τη συσκευή του ηχοβολέα και την μετακινεί επάνω στο θωρακικό τοίχωμα του ασθενούς αφού πρώτα έχει τοποθετήσει την κατάλληλη ποσότητα gel [79].

2.3.3 Σπινθηρογράφημα Αιμάτωσης Μυοκαρδίου

Το τομογραφικό σπινθηρογράφημα αιμάτωσης του μυοκαρδίου (**Εικόνα 2.14**) αποτελεί μία από τις πλέον ευρέως χρησιμοποιούμενες μη επεμβατικές μεθόδους απεικόνισης για τη διάγνωση της στεφανιαίας νόσου. Η αποτελεσματικότητά του έχει αποδειχθεί μέσα από εκτενείς κλινικές μελέτες με διαγνωστική ακρίβεια που αγγίζει το 100%. Η διαδικασία αυτή περιλαμβάνει συνδυασμό δοκιμασίας κόπωσης και απεικόνισης της αιμάτωσης του μυοκαρδίου [80].



Εικόνα 2.14: Το Σπινθηρογράφημα Αιμάτωσης Μυοκαρδίου [81]

Στη διάρκεια της δοκιμασίας κόπωσης μόλις ο ασθενής φτάσει στο μέγιστο της άσκησης ή παρουσιάσει συμπτώματα και/ή ηλεκτροκαρδιογραφικές αλλαγές που απαιτούν τη διακοπή της δοκιμασίας γίνεται ενδοφλέβια χορήγηση μικρής ποσότητας ραδιοϊσοτόπου όπως τεχνήτιο ή θάλλιο. Το ραδιοϊσότοπο αυτό αναμειγνύεται με το αίμα και μέσω της ροής στις στεφανιαίες αρτηρίες εισέρχεται στα μυοκαρδιακά κύτταρα. Στη συνέχεια με τη χρήση γ-κάμερας η οποία περιστρέφεται γύρω από τον ασθενή λαμβάνονται εικόνες που απεικονίζουν την κατανομή της ουσίας στους ιστούς της καρδιάς [82].

Οι περιοχές που δεν αιματώνονται επαρκώς φαίνονται στην απεικόνιση με μειωμένη ή καθόλου πρόσληψη ραδιοφαρμάκου και εμφανίζονται ως σκοτεινότερες περιοχές σε αντίθεση με τις φυσιολογικά αιματούμενες περιοχές οι οποίες είναι πιο φωτεινές. Η βασική αρχή της μεθόδου βασίζεται στη διαίρεση της καρδιάς σε ζώνες αιμάτωσης που εξαρτώνται από συγκεκριμένες στεφανιαίες αρτηρίες. Αυτό έχει ως αποτέλεσμα η μειωμένη αιμάτωση μίας περιοχής στην έγχρωμη απεικόνιση να υποδεικνύει πιθανή στένωση ή απόφραξη στην αντίστοιχη αρτηρία [80].

Έρευνες έχουν δείξει ότι το σπινθηρογράφημα αιμάτωσης του μυοκαρδίου αποτελεί μια ασφαλή και αξιόπιστη διαγνωστική μέθοδο για την ανίχνευση της στεφανιαίας νόσου ακόμη και στα πρώιμα στάδιά της. Επιπλέον, σε περιπτώσεις ήδη διαγνωσμένης στεφανιαίας νόσου το σπινθηρογράφημα επιτρέπει τη διαχρονική παρακολούθηση της εξέλιξής της καθώς και τον έλεγχο για πιθανή επαναστένωση μετά από τοποθέτηση stent [83].

Η μέθοδος αυτή παρέχει πολύτιμες πληροφορίες που σχετίζονται με τα παρακάτω:

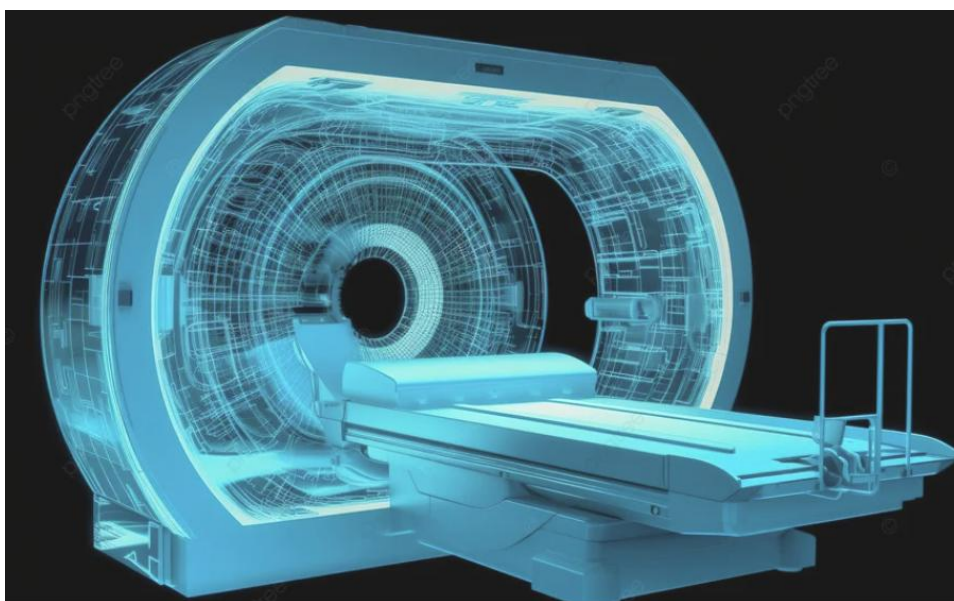
- Την κατάσταση των στεφανιαίων αγγείων (αν υπάρχει στένωση ή απόφραξη).

- Τη λειτουργικότητα και τη βιωσιμότητα του μυοκαρδίου (π.χ. ενδείξεις προηγούμενου εμφράγματος).
- Το μέγεθος και τη δομή των καρδιακών κοιλοτήτων.
- Την καρδιακή απόδοση, όπως το κλάσμα εξώθησης.
- Την αποτελεσματικότητα της τρέχουσας θεραπείας (π.χ. φαρμακευτική αγωγή, αγγειοπλαστική, ή καρδιοχειρουργική επέμβαση).

Αξίζει να σημειωθεί ότι οι ραγδαίες εξελίξεις στην τεχνολογία της καρδιαγγειακής απεικόνισης ανοίγουν νέους δρόμους στη διάγνωση και διαχείριση της στεφανιαίας νόσου. Αυτό μας οδηγεί στην εποχή της συνδυασμένης απεικόνισης όπου η διάγνωση μπορεί να επιτυγχάνεται μέσω της συνεργασίας λειτουργικών δοκιμασιών, όπως το σπινθηρογράφημα με απεικονιστικές μεθόδους της ανατομίας, όπως η αξονική τομογραφία καρδιάς. Ο συνδυασμός αυτών των τεχνικών αναμένεται να ενισχύσει σημαντικά την ακρίβεια της σύγχρονης Καρδιολογίας βελτιώνοντας τη διάγνωση, την παρακολούθηση και την εξατομικευμένη θεραπευτική προσέγγιση των ασθενών με στεφανιαία νόσο [81].

2.3.4 Μαγνητική Τομογραφία

Η απεικόνιση μαγνητικού συντονισμού (Magnetic Resonance Imaging - MRI) (**Εικόνα 2.15**) της καρδιάς αποτελεί μια προηγμένη τεχνική απεικόνισης δύο και τριών διαστάσεων η οποία χρησιμοποιείται ευρέως στην ιατρική ως διαγνωστικό εργαλείο. Χαρακτηρίζεται από την εξαιρετική χωρική ανάλυσή της, το μεγάλο οπτικό πεδίο και την υψηλή αντίθεση μεταξύ των μαλακών ιστών επιτρέποντας λεπτομερή αξιολόγηση της καρδιακής δομής και λειτουργίας [84].



Εικόνα 2.15: Μαγνητικός τομογράφος [85]

Αποτελεί την προτιμώμενη μέθοδο για την ανάλυση της συσταλτικότητας των κοιλιών (συμπεριλαμβανομένου του κλάσματος εξώθησης) καθώς και για τη μελέτη των κόλπων και των συγγενών καρδιοπαθειών. Επιπλέον, συμβάλλει στη διάγνωση και εκτίμηση καρδιακών όγκων, παθήσεων της αορτής και των μεγάλων αγγείων, καθώς και βαλβιδοπαθειών τόσο σε φυσικές όσο και σε προσθετικές βαλβίδες όπως στενώσεις και ανεπάρκειες [86]. Παράλληλα, χρησιμοποιείται στη διάγνωση μυοκαρδίτιδας, περικαρδίτιδας, ισχαιμικών βλαβών, στη μελέτη της βιωσιμότητας του μυοκαρδίου και στην ανίχνευση εναποθέσεων σιδήρου λόγω επαναλαμβανόμενων μεταγγίσεων [84].

Ως μία από τις πιο σύγχρονες απεικονιστικές μεθόδους, η απεικόνιση μαγνητικού συντονισμού πρόκειται για μια μη επεμβατική, ασφαλή και αναίμακτη διαγνωστική μέθοδο η οποία δεν εκθέτει τον ασθενή σε ιοντίζουσα ακτινοβολία. Διακρίνεται για την εξαιρετική ακρίβεια, την αξιόπιστη επαναληψιμότητα των μετρήσεων της και είναι κατάλληλη για τη συντριπτική πλειονότητα των ασθενών ανεξάρτητα από το βάρος ή τον σωματότυπό τους ενώ μπορεί να εφαρμοστεί ακόμα και σε παιδιά ή εγκύους. Παρέχει εικόνες υψηλής ευκρίνειας (**Εικόνα 2.16**) και επιτρέπει την τρισδιάστατη απεικόνιση των καρδιακών κοιλοτήτων και των μεγάλων αγγείων του θώρακα. Επιπλέον, διεισδύει στα καρδιακά τοιχώματα, προσφέροντας πολύτιμες πληροφορίες για τη δομή και τη σύστασή τους [87].



Εικόνα 2.16: Απεικόνιση Μαγνητικής τομογραφίας καρδιάς [87]

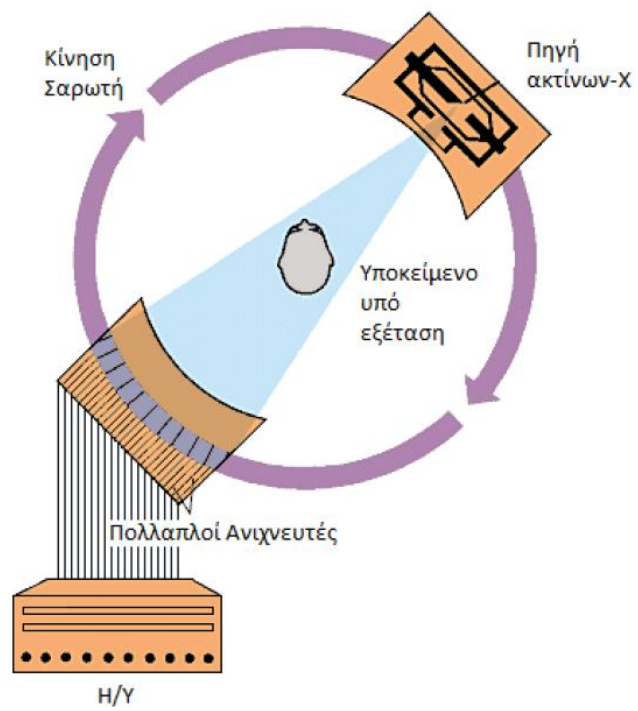
2.3.5 Αξονική Τομογραφία Καρδιάς (Αξονική Στεφανιογραφία)

Η αξονική τομογραφία (**Εικόνα 2.17**) αποτελεί μια διαγνωστική τεχνική απεικόνισης που βασίζεται στη χρήση ακτίνων X επιτρέποντας τη λήψη λεπτομερών εικόνων σχεδόν από κάθε σημείο του σώματος. Μέσω αυτής μπορούν να εξεταστούν τόσο οι μαλακοί ιστοί και τα οστά όσο και όργανα όπως ο εγκέφαλος, η καρδιά και τα αγγεία. Χρησιμοποιείται ευρέως τόσο για τη διάγνωση όσο και για την παρακολούθηση πορείας διάφορων παθολογικών καταστάσεων [88].

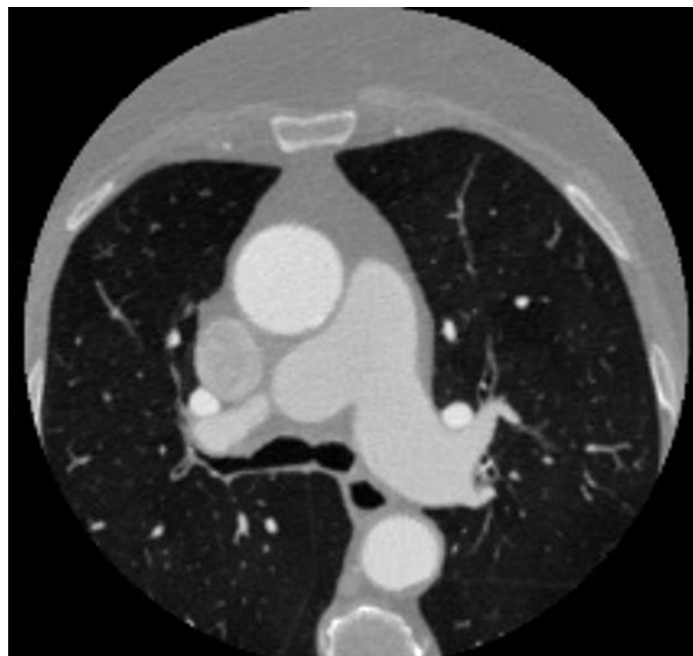


Εικόνα 2.17: Αξονικός τομογράφος [88]

Η υπολογιστική τομογραφία ακτίνων-X (X-ray computed tomography) βασίζεται στην εξασθένιση της ακτινοβολίας καθώς αυτή διέρχεται μέσα από ένα αντικείμενο (**Εικόνα 2.18**). Κατά τη διάρκεια της εξέτασης, μια πηγή ακτίνων-X εκπέμπει δέσμες φωτός προς το αντικείμενο-στόχο, το οποίο απορροφά την ακτινοβολία σε διαφορετικό βαθμό ανάλογα με τη σύστασή του. Το υπόλοιπο της ακτινοβολίας που δεν απορροφάται καταγράφεται από έναν ανιχνευτή, δημιουργώντας την αρχική εικόνα. Στη συνέχεια, χρησιμοποιούνται ειδικοί αλγόριθμοι επεξεργασίας δεδομένων για την ανακατασκευή του αντικειμένου (**Εικόνα 2.19**), επιλύοντας το αντίστροφο πρόβλημα της απεικόνισης [89].



Εικόνα 2.18: Διαδικασία σάρωσης αξονικού τομογράφου [88]



Εικόνα 2.19: Αξονική τομογραφία καρδιάς

Με την πάροδο του χρόνου, έχουν αναπτυχθεί διάφορες βελτιώσεις της τεχνικής αυτής, με στόχο τόσο την ευκολότερη συλλογή δεδομένων όσο και την επίτευξη εικόνων υψηλότερης ανάλυσης. Σημαντικές εξελίξεις περιλαμβάνουν τη χρήση κινούμενης πηγής ακτίνων-X, την τοποθέτηση κινούμενων ανιχνευτών και την προσθήκη επιπλέον συστημάτων συλλογής της ακτινοβολίας. Παρόλα αυτά, η βασική αρχή της μεθόδου παραμένει αμετάβλητη, καθώς στηρίζεται στη διαφορά των συντελεστών εξασθένησης που εμφανίζουν τα διάφορα υλικά [90].

Οι CT εικόνες αποτυπώνουν δισδιάστατες (2Δ) ή τρισδιάστατες (3Δ) κατανομές των συντελεστών γραμμικής εξασθένησης, γεγονός που σημαίνει ότι επηρεάζονται τόσο από την ενέργεια της δέσμης ακτίνων-X όσο και από τη χημική σύσταση των ιστών. Με κατάλληλη προσαρμογή των παραμέτρων, οι εικόνες αυτές μπορούν να μετατραπούν σε χάρτες πυκνότητας. Παρόλα αυτά, στα κλινικά CT συστήματα που χρησιμοποιούν πολυχρωματικές πηγές ακτίνων-X, το ευρύ ενεργειακό φάσμα μπορεί να οδηγήσει σε αύξηση του θορύβου, επηρεάζοντας την ποιότητα της απεικόνισης. Για την αντιμετώπιση αυτού του ζητήματος, έχουν αναπτυχθεί εξειδικευμένες λύσεις σε επίπεδο λογισμικού (software) και υλικού (hardware), οι οποίες συμβάλλουν στη βελτιστοποίηση της ρύθμισης του εξοπλισμού [88].

Ένα επιπλέον πρόβλημα που σχετίζεται με την ποιότητα της εικόνας είναι το φαινόμενο του μερικού όγκου. Αυτό προκύπτει όταν η διακριτική ικανότητα του τομογράφου δεν επαρκεί για να καταγράψει λεπτομέρειες μικρότερες από το όριό της, με αποτέλεσμα ορισμένες μικροσκοπικές δομές να μην αποτυπώνονται καθαρά ή να εμφανίζονται ως θόρυβος. Στα περισσότερα νοσοκομειακά συστήματα, η χωρική ανάλυση είναι περιορισμένη, με το πάχος κάθε τομής να φτάνει περίπου το 1 mm στις βέλτιστες συνθήκες [90].

Κατά τη διάρκεια της αξονικής στεφανιογραφίας καρδιάς, ο ασθενής είναι ξαπλωμένος και μετακινείται μέσα από έναν κυκλικό τομογράφο. Ένας καθετήρας εισάγεται σε φλέβα του χεριού, παρόμοια με τη διαδικασία αιμοληψίας, μέσω του οποίου χορηγείται το σκιαγραφικό υλικό. Στη συνέχεια, το τομογραφικό σύστημα, χρησιμοποιώντας χαμηλή δόση ακτινοβολίας, καταγράφει λεπτομερείς εικόνες των στεφανιαίων αρτηριών. Η συνολική διάρκεια της εξέτασης είναι περίπου 7-10 λεπτά [91].

Κάθε ιατρική εξέταση που περιλαμβάνει έκθεση σε διάφορους τύπους αξονικών τομογραφιών και στεφανιογραφιών πρέπει να πραγματοποιείται μόνο όταν κρίνεται ιατρικά αναγκαία. Αυτό διασφαλίζει ότι η έκθεση στην ακτινοβολία παραμένει στο ελάχιστο δυνατό επίπεδο. Παρόλα αυτά, τα οφέλη της εξέτασης είναι σημαντικά μεγαλύτερα από τον μικρό κίνδυνο που συνεπάγεται η ακτινοβολήση. Ο θεράπων ιατρός αξιολογεί πάντοτε την αναγκαιότητα της εξέτασης, λαμβάνοντας υπόψη τα οφέλη και τους ενδεχόμενους κινδύνους για κάθε ασθενή ξεχωριστά [92].

Μια σπάνια ανεπιθύμητη αντίδραση που μπορεί να εμφανιστεί κατά τη διάρκεια της εξέτασης είναι η αλλεργία στο σκιαγραφικό μέσο. Σε περιπτώσεις όπου υπάρχει ιστορικό αλλεργικής αντίδρασης σε προηγούμενη αξονική τομογραφία ή σε άλλες εξετάσεις που

απαιτούν τη χρήση σκιαγραφικού, όπως η πνευμογραφία, ακολουθείται ειδικό πρωτόκολλο προετοιμασίας με τη χορήγηση κατάλληλης φαρμακευτικής αγωγής, προκειμένου να μειωθεί ο κίνδυνος αλλεργικής εκδήλωσης [90].

Η αξονική στεφανιογραφία επιτρέπει την ακριβή απεικόνιση των στεφανιαίων αρτηριών, διευκολύνοντας τη διάγνωση αθηρωματικών πλακών ή στενώσεων. Παράλληλα, προσφέρει ταχεία και αξιόπιστη αξιολόγηση του βαθμού ασβεστοποίησης των στεφανιαίων αγγείων, ο οποίος αποτελεί σημαντικό δείκτη καρδιαγγειακού κινδύνου [93].

Εφόσον η εξέταση δεν αποκαλύψει στενώσεις, ο κίνδυνος εμφάνισης εμφράγματος του μυοκαρδίου λόγω αθηρωμάτωσης παραμένει χαμηλός για τουλάχιστον πέντε χρόνια. Αν, ωστόσο, εντοπιστούν στενώσεις, ο ασθενής μπορεί να υποβληθεί σε περαιτέρω έλεγχο, όπως σπινθηρογράφημα μυοκαρδίου ή άλλες διαγνωστικές διαδικασίες. Σε ορισμένες περιπτώσεις, μπορεί να κριθεί απαραίτητη η διενέργεια επεμβατικής στεφανιογραφίας, ανάλογα με την κλινική εικόνα και την εκτίμηση του θεράποντος ιατρού [93].

Κεφάλαιο 3: Σύστημα Προ-επεξεργασίας και Τμηματοποίησης Εικόνων

3.1 Εισαγωγή στα Βασικά Βήματα Προ-επεξεργασίας Εικόνας

Η προ-επεξεργασία εικόνας είναι ένα ουσιαστικό στάδιο στη βελτίωση της ποιότητάς της και στην προετοιμασία της για ανάλυση ή περαιτέρω επεξεργασία. Τα βασικά βήματα περιλαμβάνουν τη μετατροπή σε γκριζα κλίμακα, η οποία απλοποιεί την εικόνα διατηρώντας μόνο τις φωτεινότητες, τη μείωση του θορύβου μέσω φίλτρων όπως το Median filter, καθώς και την ενίσχυση της αντίθεσης με τεχνικές όπως η ισοστάθμιση ιστογράμματος [94]. Επιπλέον, εφαρμόζεται κανονικοποίηση για τη σταθεροποίηση των τιμών των εικονοστοιχείων (pixel), ενώ η αλλαγή μεγέθους και η αποκοπή συμβάλλουν στην προσαρμογή της εικόνας στις απαιτήσεις της εκάστοτε εφαρμογής. Οι παραπάνω διαδικασίες είναι απαραίτητες για την κατάλληλη προετοιμασία των ψηφιακών ιατρικών εικόνων με στόχο τη βελτίωση της ποιότητάς τους και την προσαρμογή τους στις ανάγκες των αλγορίθμων τμηματοποίησης ώστε να επιτευχθεί πιο ακριβής διάγνωση και να εξαχθούν χρήσιμες πληροφορίες για την υγεία του ασθενούς [95], [96].

Μια εικόνα μπορεί να περιγράφει ως μια δισδιάστατη συνάρτηση $f(x,y)$, όπου οι μεταβλητές x και y αντιπροσωπεύουν τις συντεταγμένες στον χώρο, ενώ η τιμή της συνάρτησης για κάθε συγκεκριμένο ζεύγος αυτών των συντεταγμένων εκφράζει τη φωτεινότητα στο αντίστοιχο σημείο. Στις ασπρόμαυρες εικόνες με τις οποίες θα επικεντρωθούμε η φωτεινότητα συχνά αναφέρεται ως «επίπεδο του γκρι» [97]. Μια αναλογική εικόνα χαρακτηρίζεται από συνεχείς τιμές τόσο στις χωρικές συντεταγμένες x,y όσο και στη φωτεινότητά της. Για να μπορέσει να υποστεί επεξεργασία από έναν υπολογιστή, είναι απαραίτητο να μετατραπεί σε ψηφιακή μορφή. Η διαδικασία αυτή περιλαμβάνει τη δειγματοληψία (sampling), δηλαδή τη διακριτοποίηση των συντεταγμένων, και την κβάντιση (quantization), που μετατρέπει τις συνεχείς τιμές φωτεινότητας σε ένα πεπερασμένο σύνολο διακριτών επιπέδων. Μέσω αυτής της μετατροπής, προκύπτει η ψηφιακή εικόνα [98].

3.1.1 Ισοστάθμιση Ιστογράμματος

Το ιστόγραμμα μιας εικόνας σε αποχρώσεις του γκρι αποτελεί πολύτιμη πηγή πληροφοριών σχετικά με το περιεχόμενό της και για αυτόν τον λόγο θεωρείται ένα από τα βασικά εργαλεία στην επεξεργασία ψηφιακών εικόνων. Η μέθοδος της ισοστάθμισης του ιστογράμματος, μπορεί να αξιοποιηθεί για τη βελτίωση της ποιότητας της εικόνας, την τροποποίηση των ιδιοτήτων της, τη μετατροπή της σε μορφή με μειωμένες αποχρώσεις του γκρι, καθώς και για την εξαγωγή χαρακτηριστικών που τη περιγράφουν, μεταξύ άλλων [97].

Ένα ιστόγραμμα είναι μια γραφική αναπαράσταση της κατανομής των δεδομένων. Στην επεξεργασία εικόνας, χρησιμοποιείται για να απεικονίσει την κατανομή των τιμών των pixel σε μια εικόνα. Το ιστόγραμμα μιας ψηφιακής εικόνας σε κλίμακα του γκρι, όπου οι τιμές κυμαίνονται στο εύρος $[0, L-1]$, μπορεί να περιγραφεί ως μια διακριτή συνάρτηση [99].

Το ιστόγραμμα μιας εικόνας μπορεί να υπολογιστεί με απλό τρόπο και περιγράφει πώς κατανέμονται οι διάφορες αποχρώσεις του γκρι σε αυτή. Συχνά, το ιστόγραμμα παρέχει ουσιαστικές πληροφορίες για τη μορφή της εικόνας. Πρόκειται για ένα διάγραμμα όπου στον οριζόντιο άξονα αποτυπώνονται οι τιμές φωτεινότητας (από 0 έως 255), ενώ στον κατακόρυφο άξονα παρουσιάζεται ο αριθμός των εικονοστοιχείων που αντιστοιχούν σε κάθε φωτεινότητα. Σε ορισμένες περιπτώσεις, οι τιμές στον κατακόρυφο άξονα προσαρμόζονται αναλογικά με βάση τη μέγιστη τιμή του ιστογράμματος [100].

Η ανάλυση της μεθόδου αυτής, αναλύεται θεωρώντας ότι έχουμε μια γρι εικόνα $A(k, m)$, με διαστάσεις $N \times M$, για να έχουμε $A(k, m) \{0, \dots, L-1\}$.

Θεωρώντας ότι $h(g)$, για $g = 0, \dots, L-1$, το ιστόγραμμα της I εικόνας. Ο υπολογισμός γίνεται:

$$P(g) = \frac{1}{N \cdot M} \sum_{i=0}^g h(i), g = 0, \dots, L-1$$

(3-1)

Η συνάρτηση μετασχηματισμού των φωτεινοτήτων δίνεται από τη σχέση :

$$T(g) = \text{int}[L \cdot P(g)]$$

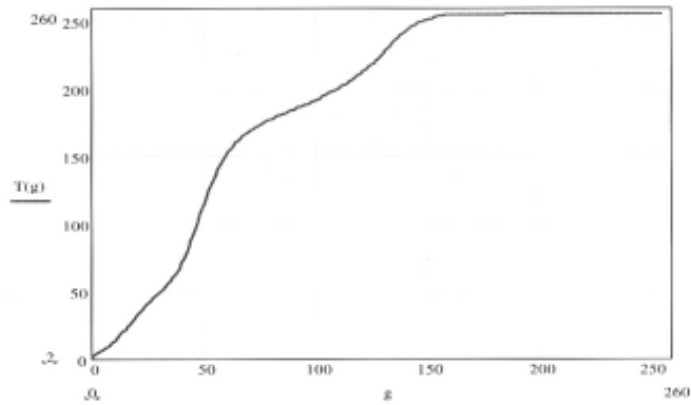
(3-2)

Η τιμή φωτεινότητας $A(k, m)$ της αρχικής εικόνας, μετατρέπεται στην αντίστοιχη φωτεινότητα $B(k, m)$ της επεξεργασμένης εικόνας με βάση τον εξής μετασχηματισμό:

$$B(k, m) = T\left(A(k, m)\right), k = 0, \dots, N - \frac{1}{m} = 0, \dots, M - 1$$

(3-3)

Η συνάρτηση μετασχηματισμού $T(g)$ έχει τη μορφή της **(Εικόνα 3.1)**, αφού έχει εφαρμοστεί η εξισορρόπηση ιστογράμματος στην προηγούμενη εικόνα.



Εικόνα 3.1: Συνάρτηση μετασχηματισμού $T(g)$

Η μέθοδος εξισορρόπησης ιστογράμματος που παρουσιάστηκε προηγουμένως, είναι γνωστή ως "ολική εξισορρόπηση ιστογράμματος" (global histogram equalization) και διαφέρει από τις μεθόδους που εφαρμόζουν την εξισορρόπηση σε τοπικές περιοχές της εικόνας [101], γνωστές ως "τοπική εξισορρόπηση ιστογράμματος" (local histogram equalization) [102].

3.1.1.1 Μετασχηματισμοί Φωτεινότητας

Η φωτεινότητα (brightness) μιας εικόνας μπορεί να οριστεί ως ο μέσος όρος της φωτεινότητας των εικονοστοιχείων της [103]. Συγκεκριμένα, για μια εικόνα με διαστάσεις $N \times M$, η συνολική φωτεινότητά της υπολογίζεται ως:

$$B = \frac{1}{N \cdot M} \sum_{n=1}^N \sum_{m=1}^M I(n, m) \quad (3-4)$$

Οι μέθοδοι μετασχηματισμού της φωτεινότητας στηρίζονται σε συναρτήσεις μετασχηματισμού $T(g)$, ακολουθώντας τη διαδικασία που περιγράφεται στη σχέση:

$$h' = T(h(k)), k = 0, \dots, L - 1 \quad (3-5)$$

Δεδομένου ότι η επεξεργασία κάθε pixel μιας εικόνας βασίζεται αποκλειστικά στη φωτεινότητά του, οι μέθοδοι αυτής της κατηγορίας αποκαλούνται και τεχνικές σημειακής επεξεργασίας (point processing).

Για το σκοπό αυτό, ακολουθεί παράδειγμα εικόνων κλίμακας τους γκρι, 8-bit, με εύρος τιμών φωτεινότητας $[0, 255]$, όπου μέσω της συνάρτησης μετασχηματισμού Εξ.(3-6), παράγεται το αρνητικό μιας εικόνας.

$$T(g) = 255 - g \quad (3-6)$$

Εικόνες με χαμηλή αντίθεση μπορεί να εμφανιστούν λόγω ανεπαρκούς φωτισμού, περιορισμένης δυναμικής εμβέλειας του οπτικού αισθητήρα ή λανθασμένων ρυθμίσεων κατά τη λήψη [104].

Για μια εικόνα σε κλίμακα του γκρι με διαστάσεις $N \times M$ και φωτεινότητα B , η αντίθεση της εικόνας δίνεται από την εξίσωση:

$$C = \sqrt{\frac{1}{N \cdot M} \sum_{n=1}^N \sum_{m=1}^M [I(n, m) - B]^2} \quad (3-7)$$

Σε πολλές περιπτώσεις, αυτό που μας ενδιαφέρει είναι η τοπική αντίθεση κάθε εικονοστοιχείου. Ως εκ τούτου η τοπική αντίθεση μπορεί να εκφραστεί ως:

$$C(n, m) = \frac{I(n, m) - \bar{I}}{I_{max} - 0} = \frac{I(n, m) - B}{255} \quad (3-8)$$

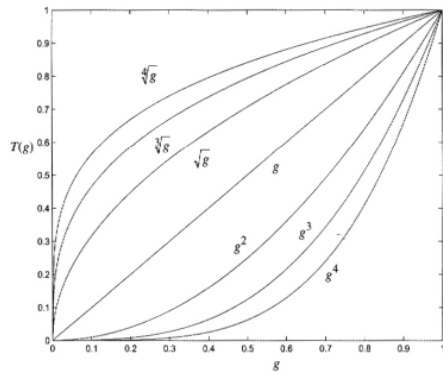
Με αυτόν τον τρόπο, η ενίσχυση της αντίθεσης θα εφαρμοστεί μόνο στις απαραίτητες περιοχές, χωρίς να επηρεαστεί ολόκληρη η εικόνα. Για παράδειγμα, μπορούμε να διατηρήσουμε σταθερή τη φωτεινότητα σε ομοιογενείς περιοχές ως προς τη φωτεινότητα, καθώς και σε περιοχές με ήδη υψηλή αντίθεση. Σε όλα τα υπόλοιπα εικονοστοιχεία της εικόνας, η αντίθεση αυξάνεται ακολουθώντας έναν από τους ακόλουθους μετασχηματισμούς [105].

Ο μετασχηματισμός δύναμης (**Εικόνα 3.2**) αποτελεί την πρώτη κατηγορία τέτοιων συναρτήσεων και δίνεται από την ακόλουθη σχέση:

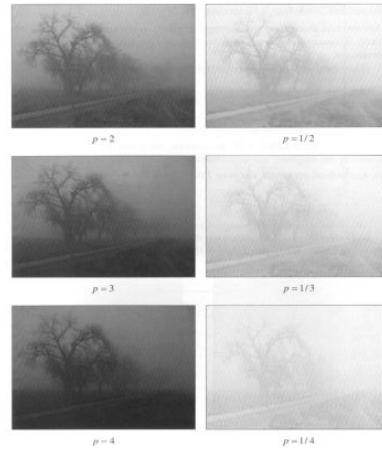
$$T(g) = 255 \cdot g^p \quad (3-9)$$

για $0 \leq g \leq 1$ και ο p ο εκθέτης. Άρα για μια εικόνα $I(i, j) \in [0, 255]$ θα έχουμε:

$$I(i, j) = \text{int}((255 \cdot g(i, j)^p)), g(i, j) = \frac{I(i, j)}{255} \quad (3-10)$$



(α)



(β)

Εικόνα 3.2: (α) Μορφές της $T(g)$ για διάφορες τιμές του p , (β) Αποτελέσματα του μετασχηματισμού δύναμης για διάφορες τιμές του p

Στην δεύτερη κατηγορία ανήκει ο λογαριθμικός μετασχηματισμός (**Εικόνα 3.3**), (**Εικόνα 3.4**), Εξ.(3-11) ο οποίος αξιοποιείται για την βελτίωση σκοτεινών και φωτεινών εικόνων αντίστοιχα.

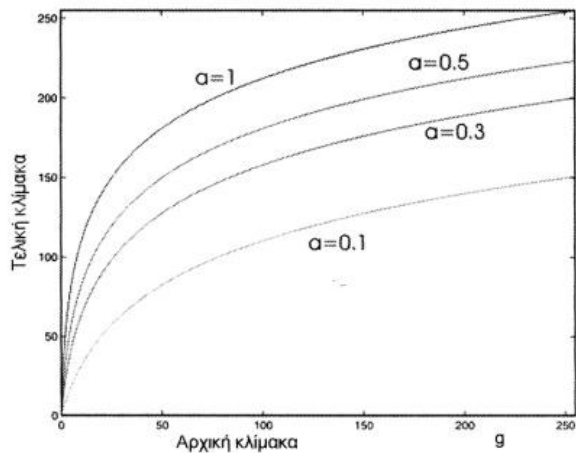
$$T(g) = b \cdot \ln(1 + a \cdot g)$$

(3-11)

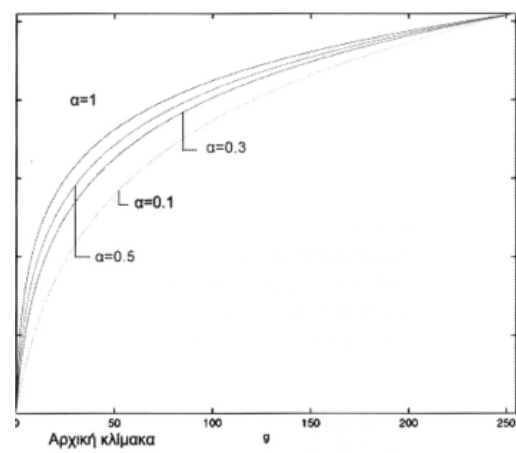
Θέτοντας ότι $T(0) = 0$ και $T(255) = 255$ τότε έχουμε ότι :

$$b = \frac{255}{\ln(1 + 255 \cdot a)}$$

(3-12)

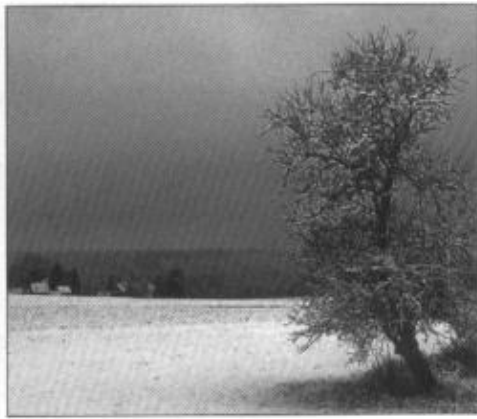


(α)



(β)

Εικόνα 3.3: Καμπύλες λογαριθμικού μετασχηματισμού (α) για $b=45.886$, (β) $b = 255(\ln(1 + 255 \cdot a))$



(α)



(β)

Εικόνα 3.4: Εφαρμογή του λογαριθμικού μετασχηματισμού (α) Αρχική εικόνα, (β) Τελική εικόνα με $a=1$

Ακολουθεί εκθετικός μετασχηματισμός (**Εικόνα 3.5**), (**Εικόνα 3.6**) μέσω της παρακάτω σχέσης.

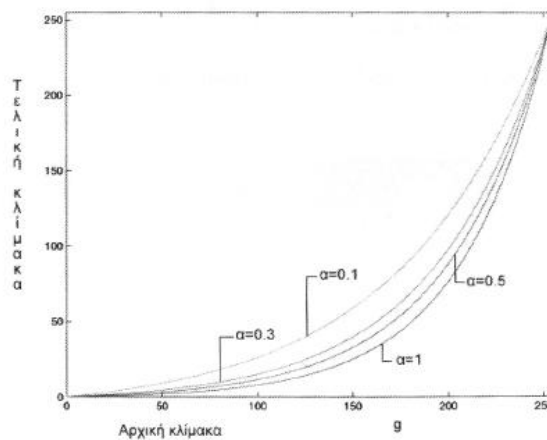
$$T(g) = \frac{1}{a} (e^{\frac{g}{b}} - 1)$$

(3-13)

Θέτοντας $T(0) = 0$ και $T(255) = 255$ τότε το b υπολογίζεται από μέσω της εξίσωσης:

$$b = \frac{255}{\ln(1 + 255 \cdot a)}$$

(3-14)



Εικόνα 3.5: Καμπύλες εκθετικού μετασχηματισμού για $b = 255 / (\ln 1 + 255 \cdot a)$



(α)



(β)

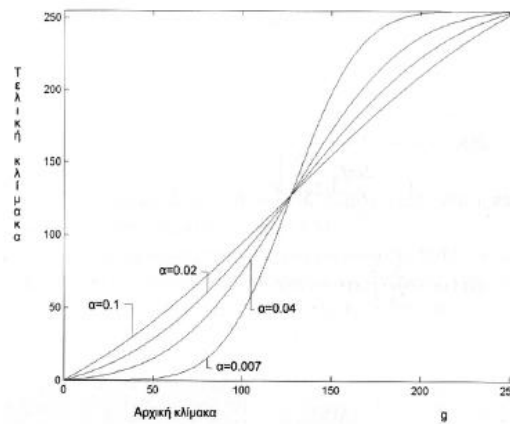
Εικόνα 3.6: Εφαρμογή του εκθετικού μετασχηματισμού (α) Αρχική εικόνα, (β) Τελική εικόνα για $a=1$

Στην τελευταία κατηγορία εντάσσεται ο μετασχηματισμός που βασίζεται στην Γκαουσιανή συνάρτηση σφάλματος (**Εικόνα 3.7**), (**Εικόνα 3.8**) και η οποία εκφράζεται από τις παρακάτω σχέσεις:

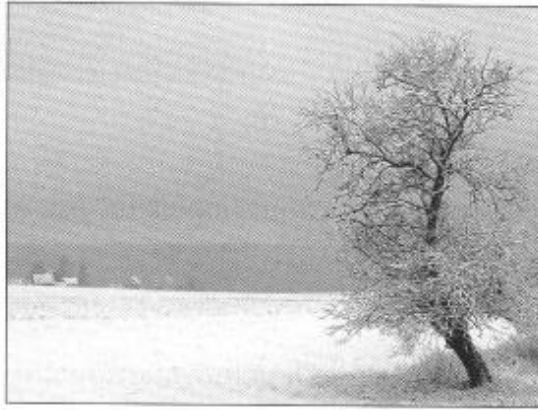
$$T(g) = 255 \frac{\operatorname{erf}\left(\frac{\left(\frac{g}{255}\right) - 0.5}{2\sqrt{a}}\right) + \operatorname{erf}\left(\frac{0.5}{2\sqrt{a}}\right)}{2 \cdot \operatorname{erf}\left(\frac{0.5}{2\sqrt{a}}\right)} \quad (3-15)$$

όπου

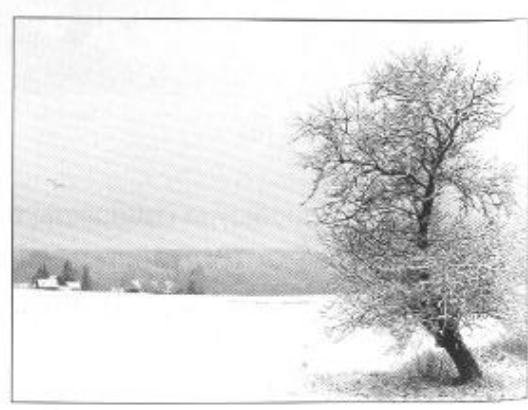
$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-y^2) dy \quad (3-16)$$



Εικόνα 3.7: Μορφή της συνάρτησης για διάφορες τιμές του a



(α)



(β)

Εικόνα 3.8: Εφαρμογή της Γκαουσιανής συνάρτησης σφάλματος (α) Αρχική εικόνα, (β) Τελική εικόνα με $a=0.02$

3.1.2 Αφαίρεση Θορύβου και εξομάλυνση ακμών

Η χρήση του φίλτρου μέσης τιμής (median filter) αποτελεί μία μη γραμμική μέθοδο επεξεργασίας ψηφιακών εικόνων, η οποία εφαρμόζεται κυρίως για την απομάκρυνση θορύβου και την εξομάλυνση (smoothing) των ακμών. Η μέση (median) τιμή ενός συνόλου A αντιστοιχεί στο στοιχείο που βρίσκεται στη μέση του συνόλου όταν αυτό έχει ταξινομηθεί σε αύξουσα σειρά [106] [107].

Ειδικότερα, σε ένα σύνολο A με περιττό αριθμό στοιχείων, η μέση τιμή προκύπτει εύκολα ως το μεσαίο στοιχείο της ταξινομημένης λίστας. Αντίθετα όταν το σύνολο περιέχει άρτιο αριθμό στοιχείων, τότε η μέση τιμή προκύπτει από τον μέσο όρο των δύο γειτονικών στοιχείων μετά την ταξινόμηση σε αύξουσα σειρά [108] [109].

Έστω ότι έχουμε το ακόλουθο σύνολο στοιχείων :

$$A = \{a_1, a_2, \dots, a_n\}$$

(3-17)

με $a_1 \leq a_2 \leq \dots \leq a_n \in \mathbb{R}$, το median του συνόλου A θα είναι ίσο με :

$$A = \begin{cases} \frac{a_{n+1}}{2}, & \text{αν } n \text{ περιττός} \\ \frac{1}{2} \left(a_{\frac{n}{2}} + a_{\frac{n}{2}+1} \right), & \text{αν } n \text{ άρτιος} \end{cases}$$

(3-18)

όπου n ο αριθμός των στοιχείων του συνόλου A .

Για να εφαρμοστεί το φίλτρο median σε μια εικόνα, αρχικά πρέπει να οριστεί η περιοχή γύρω από κάθε εικονοστοιχείο που θα ληφθεί υπόψη. Συνήθως, αυτή η περιοχή καθορίζεται με βάση μια ακτίνα απόστασης από το κεντρικό pixel. Για παράδειγμα, αν η ακτίνα είναι 1, τότε η γειτονική περιοχή σχηματίζει ένα παράθυρο 3x3. Αντίστοιχα, για ακτίνα 2, προκύπτει ένα παράθυρο 5x5, και ούτω καθεξής. Ανεξαρτήτως του μεγέθους της γειτονιάς, η τιμή του κεντρικού εικονοστοιχείου αντικαθίσταται από τη διάμεση τιμή των τιμών της περιοχής αυτής, μέσω της λειτουργίας του median φίλτρου. Τα φίλτρα median ανήκουν στην ευρύτερη κατηγορία των φίλτρων κατάταξης (rank filters), τα οποία βασίζονται στην ταξινόμηση των τιμών της γειτονιάς σύμφωνα με ορισμένους κανόνες [110].

3.1.3 Διορθώσεις Σχηματισμού

Στην προ-επεξεργασία εικόνας, η μεταβολή μεγέθους και η αποκοπή (cropping) συγκαταλέγονται στις βασικές τεχνικές που χρησιμοποιούνται για την κατάλληλη προετοιμασία δεδομένων προς εισαγωγή σε μοντέλα μηχανικής μάθησης [111], [112]. Η εφαρμογή αυτών των διαδικασιών εξασφαλίζει συνεπή διαστασιολόγηση των εικόνων κάτι το οποίο είναι κρίσιμο καθώς τα περισσότερα μοντέλα βαθιάς μάθησης, όπως τα Συνελικτικά Νευρωνικά Δίκτυα (CNNs), απαιτούν εισόδους σταθερών διαστάσεων για αποτελεσματική εκπαίδευση και πρόβλεψη [113].

Η μεταβολή του μεγέθους προσαρμόζει την ανάλυση μιας εικόνας, είτε μέσω μεγέθυνσης είτε μέσω σμίκρυνσης, ώστε να αντιστοιχεί σε προκαθορισμένες διαστάσεις. Αυτή η ενέργεια είναι ιδιαίτερα σημαντική όταν τα δεδομένα προέρχονται από ετερογενείς πηγές με διαφορετικά μεγέθη εικόνων. Παρόλα αυτά, η μη ορθολογική αλλαγή μεγέθους ενδέχεται να αλλοιώσει τις αναλογίες της εικόνας, προκαλώντας πιθανές απώλειες κρίσιμων οπτικών χαρακτηριστικών [114].

Η τεχνική της αποκοπής εστιάζει στην επιλογή και εξαγωγή ενός συγκεκριμένου τμήματος της εικόνας. Μέσω αυτής της διαδικασίας, επιτυγχάνεται τόσο η τυποποίηση του μεγέθους όσο και η ανάδειξη των σημαντικότερων περιοχών που ενδιαφέρουν το μοντέλο [115]. Η αποκοπή μπορεί να πραγματοποιηθεί κεντραρισμένα, τυχαία ή βάσει της θέσης του αντικειμένου και, όταν χρησιμοποιείται στρατηγικά, μπορεί να λειτουργήσει και ως τεχνική ενίσχυσης δεδομένων, παρέχοντας στο σύστημα ποικίλες απεικονίσεις του ίδιου στοιχείου [116].

Η συνδυασμένη εφαρμογή αλλαγής μεγέθους και αποκοπής αποτελεί αναπόσπαστο μέρος των σύγχρονων πρακτικών προ-επεξεργασίας εικόνων ενισχύοντας την αποδοτικότητα και την αποτελεσματικότητα της εκπαίδευσης βαθέων αλγορίθμων.

3.1.4 Ενίσχυση Ακμών

Η ενίσχυση ακμών αποτελεί ένα από τα θεμελιώδη στάδια στην προ-επεξεργασία εικόνων καθώς στοχεύει στην ανάδειξη των σημείων όπου παρατηρούνται έντονες μεταβολές στη φωτεινότητα. Οι ακμές αυτές συχνά αντανακλούν τα όρια αντικειμένων, τη δομή της υφής ή άλλες κρίσιμες λεπτομέρειες [117]. Μέσα από την ανάδειξη αυτών των μεταβάσεων διευκολύνεται η εργασία μεταγενέστερων αλγορίθμων όπως η τμηματοποίηση εικόνας, η αναγνώριση αντικειμένων και τα μοντέλα ταξινόμησης, ενισχύοντας την ακρίβεια στην ανάλυση οπτικών δεδομένων.

Σε πεδία όπως η υπολογιστική όραση και η μηχανική μάθηση η διατήρηση και η προβολή των δομικών χαρακτηριστικών μιας εικόνας είναι ιδιαίτερα σημαντική. Η διαδικασία ενίσχυσης ακμών συμβάλλει ουσιαστικά σε αυτόν τον στόχο εντοπίζοντας και ενισχύοντας τα σημεία όπου οι αλλαγές φωτεινότητας είναι απότομες διευκολύνοντας έτσι την καλύτερη αναγνώριση σημαντικών στοιχείων [118]. Ελλείψει αποτελεσματικής ενίσχυσης τα όρια των αντικειμένων μπορεί να αποδίδονται ασαφή ή θολά γεγονός που επηρεάζει αρνητικά την απόδοση των αναλυτικών εφαρμογών.

Διάφορες μέθοδοι έχουν αναπτυχθεί για τη βελτίωση των ακμών σε εικόνες κάθε μία με τα δικά της πλεονεκτήματα ανάλογα με τις απαιτήσεις της εκάστοτε εφαρμογής:

- Gradient-Based Filters: Τελεστές όπως οι Sobel και Prewitt ανιχνεύουν ακμές μέσω του υπολογισμού των μεταβολών στην ένταση φωτεινότητας, ενισχύοντας περιοχές με απότομες αλλαγές [119].
- Second Derivative Methods: Τεχνικές όπως το φίλτρο Laplacian βασίζονται στις δεύτερες παραγώγους για να ανιχνεύσουν περιοχές όπου η μεταβολή της έντασης παρουσιάζει αιχμηρές διακυμάνσεις ανιχνεύοντας ακόμα πιο λεπτές λεπτομέρειες αλλά με αυξημένη ευαισθησία στον θόρυβο [118].
- Canny Edge Detector: Πρόκειται για έναν από τους πιο διαδεδομένους αλγορίθμους, ο οποίος συνδυάζει Gaussian φίλτρα για εξομάλυνση θορύβου, υπολογισμό διαβαθμίσεων, καταστολή μη-μεγίστων (non-maximum suppression) και τεχνικές υστέρησης κατωφλίου (hysteresis thresholding) για τη διατήρηση των ουσιαστικών ακμών [120].
- High-Pass Filters: Μέσα από τη διέλευση μόνο των υψηλής συχνότητας σημάτων — δηλαδή των απότομων αλλαγών φωτεινότητας — αυτά τα φίλτρα αναδεικνύουν αποτελεσματικά τις ακμές, περιορίζοντας παράλληλα την επίδραση ομαλών περιοχών [121].

3.1.5 Ευθυγράμμιση Εικόνας

Η ευθυγράμμιση εικόνων (image registration) είναι μια βασική διαδικασία στην επεξεργασία εικόνας, την υπολογιστική όραση και την ιατρική απεικόνιση. Σκοπός της είναι η ακριβής στοίχιση δύο ή περισσότερων απεικονίσεων της ίδιας σκηνής οι οποίες ενδέχεται να έχουν ληφθεί σε διαφορετικές χρονικές στιγμές από ποικίλες γωνίες ή με τη χρήση διαφορετικών αισθητήρων. Μέσω αυτής της διαδικασίας επιτυγχάνεται η χωρική συσχέτιση μεταξύ των εικόνων επιτρέποντας τη συνένωση δεδομένων, τη διενέργεια συγκριτικών μελετών και την παραγωγή πιο πλούσιων ή ακριβέστερων αναπαραστάσεων [122].

Η διαδικασία της ευθυγράμμισης ανεξαρτήτως εφαρμογής συνήθως περιλαμβάνει τέσσερα κύρια βήματα: εντοπισμό χαρακτηριστικών σημείων, αντιστοίχισή τους, εκτίμηση του μετασχηματισμού που ευθυγραμμίζει τις εικόνες και τέλος, επαναδειγματοληψία και εφαρμογή του μετασχηματισμού στην εικόνα [123].

Οι μέθοδοι που στηρίζονται σε χαρακτηριστικά (feature-based) επικεντρώνονται στην ανίχνευση διακριτών και επαναλαμβανόμενων δομών μέσα στην εικόνα όπως σημεία ενδιαφέροντος, άκρα ή περιοχές. Τεχνικές όπως το SIFT [124] και το SURF [125] έχουν καθιερωθεί χάρη στην αξιοπιστία τους απέναντι σε μεταβολές κλίμακας, περιστροφής και συνθηκών φωτισμού.

Αντίθετα, οι μέθοδοι που στηρίζονται στις τιμές φωτεινότητας (intensity-based ή area-based) επεξεργάζονται άμεσα τις εντάσεις των pixel της εικόνας, χρησιμοποιώντας συνήθως κριτήρια ομοιότητας όπως η αμοιβαία πληροφορία και η κανονικοποιημένη συσχέτιση [126].

Ανάλογα με τον τύπο του μετασχηματισμού που εφαρμόζεται η ευθυγράμμιση εικόνας, μπορεί να διαχωριστεί σε τρεις κατηγορίες: άκαμπτη (όπου επιτρέπονται μόνο η μετάθεση και η περιστροφή), αφηρημένη (που επιτρέπει επίσης την κλίμακα και την κλίση) και μη άκαμπτη (γνωστή και ως ελαστική) καταχώρηση. Η μη άκαμπτη ευθυγράμμιση είναι ιδιαίτερα σημαντική για την ιατρική απεικόνιση, καθώς οι ανατομικές δομές ενδέχεται να υποστούν σύνθετες παραμορφώσεις [127].

Τα τελευταία χρόνια, οι τεχνικές βαθιάς μάθησης έχουν φέρει επανάσταση στον τομέα της ευθυγράμμισης εικόνας επιτρέποντας τη δημιουργία μοντέλων που βασίζονται σε δεδομένα και μπορούν να προβλέψουν άμεσα τους μετασχηματισμούς. Προσεγγίσεις όπως το VoxelMorph [128] και το AIRNet [129] έχουν δείξει ότι τα συνελκτικά νευρωνικά δίκτυα μπορούν να μάθουν να ευθυγραμμίζουν σύνθετα ιατρικά δεδομένα, όπως εικόνες όγκων, χωρίς την ανάγκη ρητής εποπτείας.

Ένας άλλος σημαντικός παράγοντας κατηγοριοποίησης είναι ο τύπος των εικόνων που συμμετέχουν στην ευθυγράμμιση:

- Η ενδο-μοναδική ευθυγράμμιση αφορά εικόνες που έχουν ληφθεί από τον ίδιο αισθητήρα ή συσκευή (π.χ., MRI-MRI) [130].

- Η δια-μοναδική ευθυγράμμιση (όπως CT-MRI, RGB-Infrared) παρουσιάζει μεγαλύτερες προκλήσεις λόγω της έλλειψης ενός ενιαίου χάρτη έντασης και απαιτεί ισχυρούς δείκτες ομοιότητας, όπως η αμοιβαία πληροφορία [131].

Η αξιολόγηση της απόδοσης της ευθυγράμμισης εξακολουθεί να αποτελεί ενεργό πεδίο έρευνας. Δείκτες όπως το σφάλμα ευθυγράμμισης στόχου (Target Registration Error - TRE) [132] και ο συντελεστής ομοιότητας Dice (Dice Similarity Coefficient - DSC) [133] χρησιμοποιούνται ευρέως.

Παρά την εκτενή έρευνα, η ευθυγράμμιση εικόνας παραμένει μια δύσκολη πρόκληση λόγω παραμέτρων όπως οι μεγάλες διαφορές στην εμφάνιση των εικόνων, η απόκρυψη πληροφοριών, οι ανισότητες στη φωτεινότητα και τα artifacts. Η σύγχρονη έρευνα επικεντρώνεται όλο και περισσότερο σε μη εποπτευόμενες, ερμηνεύσιμες και υπολογιστικά αποδοτικές μεθόδους ευθυγράμμισης, οι οποίες είναι ικανές να γενικεύσουν σε διαφορετικά σύνολα δεδομένων [134].

Συνοψίζοντας, η ευθυγράμμιση εικόνας είναι ένας δυναμικός τομέας ο οποίος συνεχώς προσαρμόζεται στις νέες προκλήσεις που θέτουν οι αναδυόμενες εφαρμογές και τα νέα σύνολα δεδομένων. Η εξέλιξη του συνδέεται στενά με τις προόδους στη μηχανική μάθηση, τη θεωρία βελτιστοποίησης και την υψηλών επιδόσεων υπολογιστική.

3.1.6 Ενίσχυση Δεδομένων

Η Ενίσχυση Δεδομένων περιλαμβάνει μια σειρά από τεχνικές που στοχεύουν στην τεχνητή αύξηση του όγκου και της ποικιλομορφίας ενός δεδομένου συνόλου, δημιουργώντας παραλλαγμένες μορφές των ήδη υπαρχόντων δειγμάτων. Αποτελεί θεμελιώδη διαδικασία στην προ-επεξεργασία εικόνας ιδίως σε περιπτώσεις όπου τα διαθέσιμα δεδομένα είναι περιορισμένα ή παρουσιάζουν ανισορροπία [135].

Οι παραδοσιακές μέθοδοι ενίσχυσης δεδομένων χωρίζονται στους Γεωμετρικούς και Φωτομετρικούς μετασχηματισμούς [136].

3.1.6.1 Γεωμετρικοί μετασχηματισμοί

Οι γεωμετρικοί μετασχηματισμοί συνιστούν μια ομάδα τεχνικών ενίσχυσης δεδομένων εικόνας οι οποίες τροποποιούν τη διάταξη των εικονοστοιχείων χωρίς να επηρεάζουν τις τιμές τους. Μέσα από τη διαδικασία αυτή τα εικονοστοιχεία μετακινούνται από τις αρχικές τους θέσεις σε νέες αλλά η χρωματική ή φωτεινή πληροφορία τους παραμένει αναλλοίωτη. Σκοπός της εφαρμογής τέτοιων μετασχηματισμών είναι να εμπλουτιστεί το σύνολο εκπαίδευσης με εικόνες που προσομοιώνουν μεταβολές όπως αλλαγές στην οπτική γωνία, παραμορφώσεις,

διαφοροποιήσεις προοπτικής και κλίμακας, καταστάσεις δηλαδή που απαντώνται φυσικά στον πραγματικό κόσμο [137].

Η υλοποίηση των γεωμετρικών μετασχηματισμών μπορεί να γίνει απευθείας μέσω πολυωνυμικών συναρτήσεων. Σε αυτό το πλαίσιο θα αναφερθούμε στις βασικές αρχές και στις πιο συνηθισμένες εφαρμογές τους. Γενικά, οι γεωμετρικοί μετασχηματισμοί ταξινομούνται στους αφινικούς μετασχηματισμούς (affine) και στους μη αφινικούς (non-affine).

Οι αφινικοί μετασχηματισμοί περιγράφονται μέσω γενικευμένων γραμμικών εξισώσεων πρώτου βαθμού και απαιτούν περιορισμένο αριθμό παραμέτρων για την πραγματοποίηση συγκεκριμένων ενισχυτικών λειτουργιών. Αντίθετα, οι μη αφινικοί μετασχηματισμοί βασίζονται σε πιο περίπλοκες μαθηματικές σχέσεις όπου εμπλέκονται πολυώνυμα υψηλότερων βαθμών [138].

3.1.6.1.1 Αφινικοί μετασχηματισμοί

Οι βασικότερες και πιο ευρέως χρησιμοποιούμενες τεχνικές γεωμετρικών μετασχηματισμών είναι οι αφινικοί μετασχηματισμοί, όπως η περιστροφή, η διάτμηση, η μετατόπιση, η κλιμάκωση (μεταβολή μεγέθους χωρίς μεγέθυνση ή περικοπή) και η κατοπτρική αντανάκλαση ή αναστροφή. Αν και η μεγέθυνση και η περικοπή χρησιμοποιούνται συχνά για την κλιμάκωση εικόνων, δεν εμπίπτουν στους αυστηρούς ορισμούς των αφινικών μετασχηματισμών. Ένα υποσύνολο των αφινικών μετασχηματισμών αποτελούν οι Ευκλείδειοι μετασχηματισμοί [139], στους οποίους ανήκουν οι περιστροφές, οι μετατοπίσεις και οι ανακλάσεις. Παρά την απλότητα που τους χαρακτηρίζει, αρκετές έρευνες [140], [141], έχουν αποδείξει την αποτελεσματικότητά τους σε διάφορες εφαρμογές όρασης υπολογιστή. Για τον λόγο αυτό, συχνά εφαρμόζονται ως αρχικό στάδιο ενίσχυσης δεδομένων, πριν από την αξιοποίηση πιο σύνθετων τεχνικών [142], [143].

3.1.6.1.2 Μη αφινικοί μετασχηματισμοί

Αυτή η κατηγορία μετασχηματισμών προσφέρει μια μέθοδο για την προσομοίωση πιο περίπλοκων γεωμετρικών χαρακτηριστικών. Σε πολλές περιπτώσεις όπως στην ιατρική απεικόνιση [144] και την ανάλυση εγγράφων [145], απαιτούνται εξελεγμένοι μετασχηματισμοί για να αποδοθούν οι περίπλοκες γεωμετρικές αλλαγές που παρατηρούνται σε τέτοιες εφαρμογές.

Συνηθισμένοι τύποι μη αφινικών μετασχηματισμών περιλαμβάνουν τους προβολικούς ή προοπτικούς μετασχηματισμούς οι οποίοι στοχεύουν στην εισαγωγή νέων δειγμάτων που αναπαριστούν διαφορετικές γωνίες θέασης ή προοπτικές του παρατηρητή. Αυτό συνήθως συνεπάγεται την αντιστοίχιση σημείων από την αρχική εικόνα σε αντίστοιχα σημεία σε ένα νέο σύστημα αναφοράς με βάση μια προκαθορισμένη σχέση. Τεχνικές ενίσχυσης δεδομένων που

βασίζονται σε προβολικούς μετασχηματισμούς χρησιμοποιούνται συχνά για την επέκταση μοντέλων εκπαιδευμένων σε τυπικά σύνολα εικόνας, προκειμένου να διαχειρίζονται γεωμετρικές παραμορφώσεις που σχετίζονται με κάμερες ευρέως πεδίου όρασης. Τέτοιες κάμερες χρησιμοποιούνται συνήθως σε εφαρμογές όπως η αναγνώριση δορυφορικών εικόνων και η παρακολούθηση με UAV (Αναγνωριστικά Μη Επανδρωμένα Αεροσκάφη), καθώς και σε συστήματα παν-κατευθυντικού πεδίου όρασης [146], [147].

Η μη-γραμμική παραμόρφωση [144], [145], αποτελεί έναν άλλο σημαντικό μετασχηματισμό που χρησιμοποιείται συχνά για την ενίσχυση δεδομένων. Αυτή η μετατροπή μπορεί να επιτευχθεί προσθέτοντας περισσότερους βαθμούς ελευθερίας, επιτρέποντας στις βασικές (π.χ., αφινικές) γεωμετρικές μετατροπές να διαφέρουν ως προς το μέγεθος σε διαφορετικές τοπικές περιοχές της εικόνας. Η μη-γραμμική παραμόρφωση είναι ιδιαίτερα χρήσιμη για την αντιστάθμιση διαφορών στην εμφάνιση που προκαλούνται από φυσικά φαινόμενα (π.χ., παραμορφώσεις μη-άκαμπτων σωμάτων) ή από διάφορους παράγοντες που επηρεάζουν τα αρχεία εικόνας.

3.1.6.2 Φωτομετρικοί μετασχηματισμοί

Οι πραγματικές εργασίες αναγνώρισης εικόνας σε υπολογιστές ενδέχεται να επηρεαστούν αρνητικά από διάφορους φωτομετρικούς παράγοντες οι οποίοι προκύπτουν από τις ιδιότητες των καμερών και τις συνθήκες λήψης (όπως θόλωση λόγω κίνησης, οπτικός θόρυβος, παραμορφώσεις), χρωματικές ανακρίβειες και διαταραχές στα δεδομένα εικόνας. Αυτά τα φαινόμενα μπορούν να προσομοιωθούν επιτυχώς με μεθόδους φωτομετρικού μετασχηματισμού οι οποίες βασίζονται σε γνωστές τεχνικές επεξεργασίας ψηφιακής εικόνας επιτρέποντας τη δημιουργία εκπαιδευτικών εικόνων που διευκολύνουν την ακριβή και ανθεκτική οπτική κατανόηση. Σε αντίθεση με τις γεωμετρικές προσεγγίσεις, οι μέθοδοι φωτομετρικής ενίσχυσης τροποποιούν τα pixel της εικόνας χωρίς να επηρεάζουν τη χωρική της διάταξη. Συνήθως, αυτές οι τεχνικές μεταβάλλουν τη σύνθεση των καναλιών RGB μέσω καθορισμένων συναρτήσεων χρησιμοποιώντας γνωστές μεθόδους επεξεργασίας εικόνας [148].

Στο διαισθητικό επίπεδο, ο φωτομετρικός μετασχηματισμός επιτυγχάνεται με τη μεταβολή οπτικών χαρακτηριστικών της εικόνας όπως το χρώμα, η φωτεινότητα, η ευκρίνεια, η αντίθεση και ο κορεσμός. Επιπλέον, η φωτομετρική ενίσχυση μπορεί να ενισχύσει την ανθεκτικότητα των μοντέλων βαθιάς μάθησης προσαρμόζοντας τα δεδομένα στις παραλλαγές που προκύπτουν από περιβαλλοντικούς παράγοντες όπως φωτισμός, καιρικές συνθήκες, ώρες της ημέρας, αλλά και από σφάλματα που προκαλούνται από τις συσκευές εικόνας, τον θόρυβο και τις ρυθμίσεις της κάμερας. Οι πιο διαδεδομένες μέθοδοι φωτομετρικής ενίσχυσης περιλαμβάνουν την αλλαγή των χρωμάτων, την τροποποίηση του χώρου χρωμάτων και τεχνικές ενίσχυσης ή παραμόρφωσης της εικόνας [149].

3.2 Εισαγωγή στις Μεθόδους Τμηματοποίησης Εικόνας

Η τμηματοποίηση εικόνας αποτελεί ένα από τα πιο κρίσιμα βήματα στην επεξεργασία και ανάλυση εικόνων καθώς αποσκοπεί στον διαχωρισμό μιας εικόνας σε σημασιολογικά ή οπτικά συνεκτικές περιοχές. Πρόκειται για μια διαδικασία που επιτρέπει τον εντοπισμό και απομόνωση αντικειμένων ή περιοχών ενδιαφέροντος, διευκολύνοντας σημαντικά μεταγενέστερες διεργασίες όπως η αναγνώριση αντικειμένων, η ιατρική διάγνωση ή η αυτόνομη πλοήγηση [150].

Οι μέθοδοι τμηματοποίησης διακρίνονται σε διάφορες κατηγορίες με βάση τη φιλοσοφία τους και τα χαρακτηριστικά της εικόνας που αξιοποιούν. Οι παραδοσιακές τεχνικές όπως η κατωφλίωση (thresholding), οι αλγόριθμοι ομαδοποίησης (π.χ., K-Means) και οι μέθοδοι βασισμένες σε ακμές και περιγράμματα έχουν χρησιμοποιηθεί ευρέως σε εφαρμογές με απλές συνθήκες ή καλά ορισμένες δομές [151]. Ωστόσο, οι μέθοδοι αυτές συχνά αποτυγχάνουν όταν οι εικόνες περιέχουν θόρυβο, ασαφή όρια ή χαμηλή αντίθεση.

Πιο εξελιγμένες προσεγγίσεις όπως οι μέθοδοι βασισμένες σε γραφήματα (graph-based segmentation), τα μοντέλα ενεργών περιγραμμάτων (active contours), καθώς και τεχνικές βασισμένες σε πιθανοτικά μοντέλα, όπως το μοντέλο Markov Random Fields (MRF), προσφέρουν καλύτερα αποτελέσματα σε πιο απαιτητικά σενάρια [152].

Με την άνοδο της μηχανικής μάθησης και ιδιαίτερα της βαθιάς μάθησης, η τμηματοποίηση εικόνας γνώρισε σημαντικές εξελίξεις. Τα συνελκτικά νευρωνικά δίκτυα (CNNs), και πιο πρόσφατα τα encoder-decoder μοντέλα όπως το U-Net [153], επέτρεψαν την αυτόματη εκμάθηση πολύπλοκων χαρακτηριστικών και τη βελτιστοποίηση της ακρίβειας της τμηματοποίησης ακόμη και σε δύσκολα περιβάλλοντα, όπως η ιατρική απεικόνιση ή οι δορυφορικές εικόνες.

Οι σύγχρονες έρευνες στρέφονται προς μεθόδους που είναι όχι μόνο ακριβείς αλλά και ερμηνεύσιμες, υπολογιστικά αποδοτικές και ικανές να γενικεύουν σε διαφορετικούς τύπους δεδομένων. Παράλληλα, εξετάζεται η συνδυασμένη χρήση δεδομένων πολλαπλών καναλιών ή αισθητήρων, καθώς και η ενσωμάτωση χωρικών και συμφραζόμενων πληροφοριών [154].

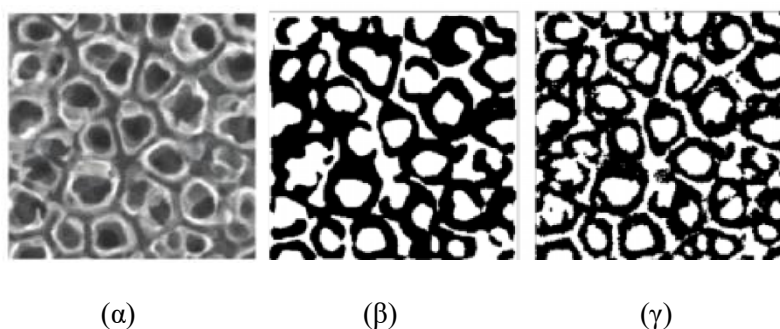
3.2.1 Κατωφλίωση (Thresholding)

Η κατωφλίωση η οποία χωρίζεται στην καθολική και την προσαρμοστική [118], αποτελεί μια τεχνική επεξεργασίας εικόνας κατά την οποία εφαρμόζονται ένα ή περισσότερα όρια φωτεινότητας (κατώφλια) προκειμένου να ταξινομηθούν τα εικονοστοιχεία (pixels) βάσει της έντασης φωτός τους, εάν δηλαδή αυτή υπερβαίνει ή υπολείπεται ενός καθορισμένου ορίου. Η μέθοδος αυτή είναι ιδιαίτερα χρήσιμη όταν τα εικονοστοιχεία που αντιστοιχούν στο ζητούμενο αντικείμενο έχουν παρόμοιες τιμές φωτεινότητας και ξεχωρίζουν σαφώς από το φόντο της

εικόνας. Αξίζει να σημειωθεί ότι αντικείμενα με αξιόλογο μέγεθος, τείνουν να σχηματίζουν συγκεντρώσεις τιμών γύρω από μια μέση φωτεινότητα στο ιστόγραμμα της εικόνας [155], [156], [157].

Η καθολική κατωφλίωση (global thresholding) (**Εικόνα 3.9**) βασίζεται στη χρήση μιας σταθερής τιμής κατωφλίου T , η οποία εφαρμόζεται σε ολόκληρη την εικόνα. Σε αυτή την περίπτωση, τα pixels με τιμές μικρότερες του T αντικαθίστανται με 0, ενώ όσα έχουν μεγαλύτερη ή ίση τιμή, αντικαθίστανται με 1. Παρόλο που η μέθοδος αυτή είναι ιδιαίτερα γρήγορη και απλή στην εφαρμογή της, εμφανίζει ορισμένα μειονεκτήματα. Δυσκολεύεται να διαχειριστεί εικόνες που περιέχουν περισσότερα από ένα αντικείμενα, είναι ευάλωτη στις διακυμάνσεις της φωτεινότητας και τον θόρυβο, και συχνά προκαλεί ασυνέχειες στις περιοχές ενδιαφέροντος. Ορισμένες συνηθισμένες συναρτήσεις κατωφλίωσης είναι το Binary Threshold, Inverse Binary Threshold, Truncate Threshold, To Zero Threshold και Inverse To Zero Threshold [158].

Αντίθετα, η προσαρμοστική κατωφλίωση (adaptive thresholding) (**Εικόνα 3.9**) χρησιμοποιεί μεταβλητά κατώφλια τα οποία προσαρμόζονται ανά περιοχή της εικόνας και ανάλογα με τις τοπικές τιμές φωτεινότητας των pixels. Η επιλογή του κατωφλίου γίνεται με βάση μια αρχική τιμή η οποία προσαρμόζεται προσθέτοντας την ένταση φωτεινότητας από ένα φιλτραρισμένο (συνήθως θολωμένο) αντίγραφο της εικόνας. Αυτή η προσέγγιση επιτρέπει τη χρήση χαμηλότερων κατωφλίων σε σκοτεινές περιοχές και υψηλότερων σε φωτεινές, οδηγώντας σε πιο ακριβή αποτελέσματα, ιδίως σε εικόνες με έντονο θόρυβο. Μεταξύ των δημοφιλών μεθόδων προσαρμοστικής κατωφλίωσης περιλαμβάνονται η τεχνική του Otsu [159], τα Mixture Models [160] και η μέθοδος Multispectral Thresholding [161].



Εικόνα 3.9: (α) Αρχική εικόνα, (β) Καθολική εικόνα, (γ) Προσαρμοστική κατωφλίωση [162]

3.2.2 Ανίχνευση Ακμών

Ακμές θεωρούνται οι περιοχές μιας εικόνας όπου παρατηρούνται τοπικές μεταβολές στην τιμή της φωτεινότητας [156]. Στην περίπτωση απεικόνισης πραγματικών αντικειμένων, οι ακμές μπορούν είτε να εντοπίζονται στο εξωτερικό τους περίγραμμα είτε να σχετίζονται με

εσωτερικά χαρακτηριστικά και μοτίβα. Συνήθως, το μεγαλύτερο ενδιαφέρον παρουσιάζουν οι εξωτερικές ακμές, καθώς αυτές προσδιορίζουν και τα όρια μεταξύ διαφορετικών περιοχών ή αντικειμένων της εικόνας. Παρ' όλα αυτά, η παρουσία θορύβου καθώς και η περιορισμένη ικανότητα των μεθόδων να διακρίνουν σταθερά τις εξωτερικές από τις εσωτερικές ακμές δυσχεραίνουν σημαντικά την ανίχνευσή τους, ιδιαίτερα σε εικόνες του πραγματικού κόσμου [118].

Υπάρχει ευρεία ποικιλία μεθόδων εντοπισμού ακμών που επιλέγονται ανάλογα με τη φύση της εικόνας και τις ειδικές απαιτήσεις της εκάστοτε εφαρμογής. Δεδομένου ότι μια ακμή συσχετίζεται με έντονη αλλαγή στις τιμές της εικόνας, σημαντικό ρόλο στη διαδικασία παίζουν τα φίλτρα παραγωγίσης, τα οποία εφαρμόζονται για την ανίχνευση αυτών των μεταβολών.

3.2.2.1 Τελεστής κλίσης (gradient)

Για μια συνεχή συνάρτηση δύο μεταβλητών, το διάνυσμα της κλίσης προσδιορίζεται μέσω των μερικών παραγώγων της ως προς τις συντεταγμένες x και y . Ο συνηθέστερος τρόπος εντοπισμού των ακμών σε μια εικόνα περιλαμβάνει τον υπολογισμό της πρώτης παραγώγου, δηλαδή της κλίσης και στη συνέχεια την εφαρμογή μιας διαδικασίας κατωφλίωσης για την ανάδειξη των σημείων όπου παρατηρούνται οι μέγιστες μεταβολές. Το διάνυσμα της κλίσης ουσιαστικά υποδεικνύει την κατεύθυνση και το μέγεθος της μεγαλύτερης αλλαγής στη φωτεινότητα της εικόνας $f(x, y)$ [163].

$$\vec{\nabla} f = \frac{\partial f}{\partial x} \vec{i} + \frac{\partial f}{\partial y} \vec{j} \quad (3-19)$$

με κατεύθυνση :

$$\theta = \tan^{-1} \left\{ \frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial y}} \right\} \quad (3-20)$$

και μέτρο :

$$A = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2} \quad (3-21)$$

Ο διακριτός τελεστής κλίσης υπολογίζεται μέσω των :

$$G_x = \Delta_x f(i, j) = f(i, j + 1) - f(i, j) \quad (3-22)$$

$$G_y = \Delta_y f(i, j) = f(i, j) - f(i + 1, j) \quad (3-23)$$

3.2.2.2 Τελεστές δεύτερης τάξης παραγώγου

Ο εντοπισμός των περιοχών όπου η δεύτερη παράγωγος της εικόνας αλλάζει πρόσημο, δηλαδή όπου η πρώτη παράγωγος γίνεται μηδέν, αποκαλύπτει πιθανά σημεία ασυνέχειας που μπορεί να αντιστοιχούν σε ακμές. Η μαθηματική αναπαράσταση της δεύτερης παραγώγου μιας συνάρτησης δύο μεταβλητών $f(x, y)$ έχει την εξής μορφή:

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} \vec{i} + \frac{\partial^2 f}{\partial y^2} \vec{j} \quad (3-24)$$

Ο διακριτός τελεστής Laplace υπολογίζεται από την ακόλουθη σχέση:

$$\nabla^2 f(i, j) = f(i + 1, j) + f(i - 1, j) + f(i, j + 1) + f(i, j - 1) - 4f(i, j) \quad (3-25)$$

και ο αντίστοιχος πυρήνας φιλτραρίσματος Laplace υπολογίζεται από:

$$\nabla^2 f(i, j) \approx \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad (3-26)$$

Η εφαρμογή φίλτρων δεύτερης παραγώγου προσφέρει αυξημένη ακρίβεια στην ανίχνευση ακμών. Ωστόσο παρουσιάζει αυξημένη ευαισθησία στον θόρυβο της εικόνας. Για την αντιμετώπιση αυτού του μειονεκτήματος εφαρμόζεται μια συνδυαστική τεχνική που ενσωματώνει Gaussian φιλτράρισμα, το οποίο μειώνει τον θόρυβο με την ανίχνευση ακμών μέσω του Laplace. Η προσέγγιση αυτή είναι γνωστή ως Laplacian of Gaussian (LoG). Στην πράξη η μέθοδος LoG υλοποιείται σε ένα ενιαίο βήμα μέσω της χρήσης κατάλληλου φίλτρου:

$$\nabla^2 G(r) = -\frac{1}{\pi\sigma^4} \left(1 - \frac{r^2}{2\sigma^2}\right) \exp\left\{-\frac{r^2}{2\sigma^2}\right\} \quad (3-27)$$

Αναφορικά με τον διακριτό υπολογισμό του LoG, ο αντίστοιχος πυρήνας φιλτραρίσματος (**Πίνακας 3.1**) είναι:

0	0	-1	0	0
0	-1	-2	-1	0
-1	-2	16	-2	-1
0	-1	-2	-1	0
0	0	-1	0	0

Πίνακας 3.1: Πυρήνας φιλτραρίσματος

Η τιμή της παραμέτρου σ καθορίζει τον βαθμό με τον οποίο εξομαλύνονται οι ακμές. Συνήθως προτιμάται η χρήση χαμηλών τιμών σ , ωστόσο είναι καλό να δοκιμάζονται διαφορετικά φίλτρα με μεταβαλλόμενο σ ώστε να εντοπιστεί η βέλτιστη τιμή για κάθε συγκεκριμένη περίπτωση [164].

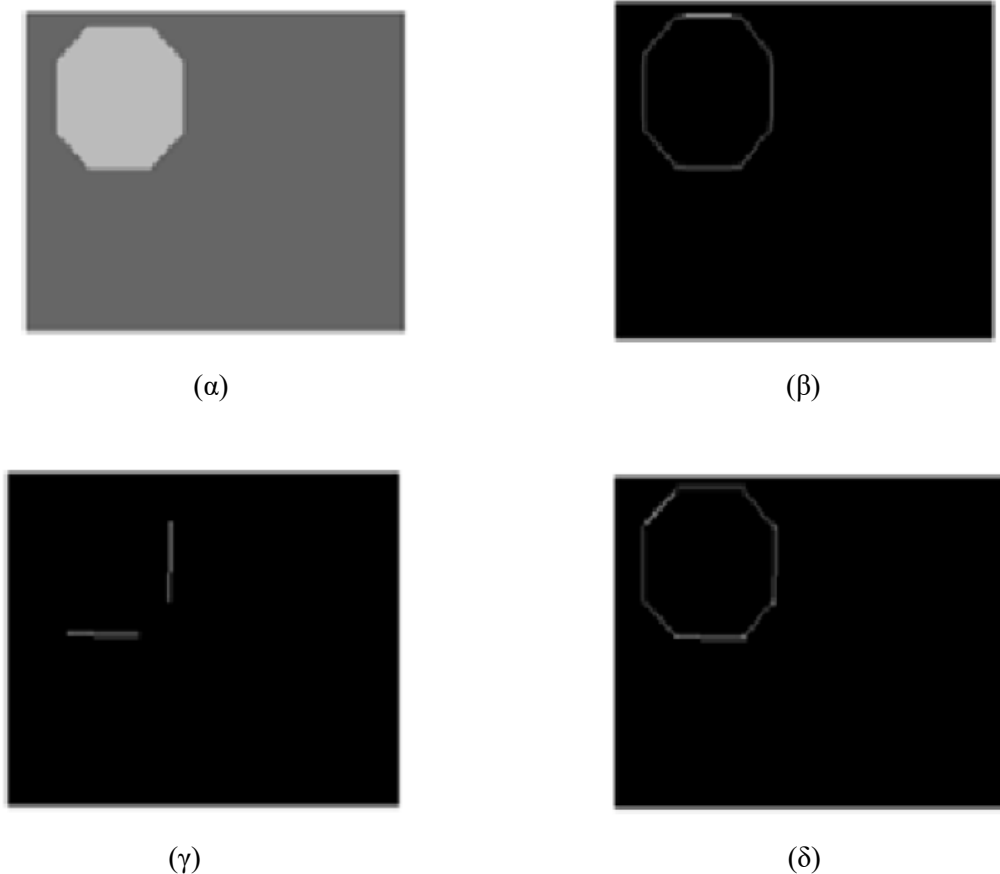
3.2.2.3 Μετασχηματισμός του Hough

Αξιοποιώντας τα εικονοστοιχεία που αντιστοιχούν σε ακμές, γίνεται προσπάθεια εντοπισμού αντικειμένων τα οποία μπορούν να περιγραφούν μέσω κλειστών μαθηματικών εκφράσεων, όπως γραμμές, κύκλοι ή ελλείψεις. Μέσω κατάλληλων συνδυασμών γραμμών μπορούν να προσεγγιστούν και πιο σύνθετα σχήματα. Μια από τις πιο ευρέως χρησιμοποιούμενες τεχνικές για την ανίχνευση ακμών και τη δημιουργία ορίων σε εικόνες είναι ο μετασχηματισμός Hough, ο οποίος αποδεικνύεται ιδιαίτερα αποτελεσματικός ακόμη και όταν η εικόνα παρουσιάζει ατέλειες ή ορισμένα σημεία απουσιάζουν [156]. Η μαθηματική αναπαράσταση μιας ευθείας σε πολικές συντεταγμένες δίνεται από την παρακάτω εξίσωση:

$$x\cos\theta + y\sin\theta = \rho \quad (3-28)$$

με ρ και θ να καθορίζουν μοναδικά μια γραμμή.

Για τον εντοπισμό γραμμών μέσω του μετασχηματισμού Hough, αρχικά υπολογίζονται όλες οι πιθανές ευθείες που θα μπορούσαν να διέρχονται από κάθε pixel που ανήκει σε ακμή. Κατόπιν, ενημερώνεται ένας δισδιάστατος πίνακας (θ, ρ) , όπου κάθε θέση του αντιστοιχεί σε μια συγκεκριμένη ευθεία. Στη συνέχεια εντοπίζονται τα κελιά του πίνακα που περιέχουν τις υψηλότερες τιμές, τα οποία αντιπροσωπεύουν τις πιο "ισχυρές" γραμμές της εικόνας. Τα αντίστοιχα ζεύγη (θ, ρ) περιγράφουν τις γραμμές που ανιχνεύθηκαν. Ένα παράδειγμα εφαρμογής του μετασχηματισμού Hough παρουσιάζεται στην **(Εικόνα 3.10)** [165].



Εικόνα 3.10: Εφαρμογή του μετ/μού Hough. (α) Αρχική εικόνα, (β) Μάσκα ακμών, (γ) Δύο πρώτες γραμμές με μετ/μό Hough, (δ) Πλήρες αποτέλεσμα μετ/μού Hough [166]

3.2.3 Μέθοδοι Βασισμένες σε Περιοχές (Region Growing)

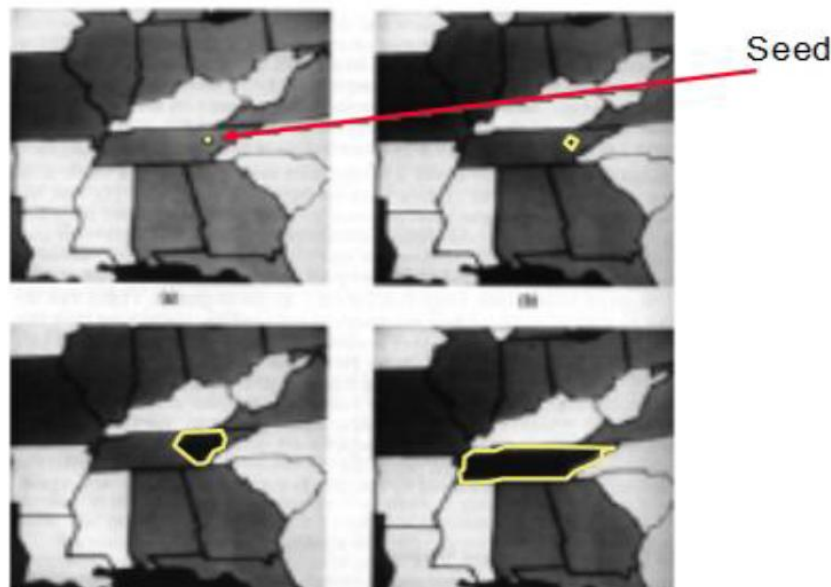
Η μέθοδος ανάπτυξης περιοχών (region growing) αποτελεί μια εναλλακτική προσέγγιση για την τμηματοποίηση εικόνας, με βασικό στόχο τη διαίρεση της εικόνας R σε μικρότερες επιμέρους περιοχές R_i , $i = 1, 2, \dots, N$, οι οποίες να πληρούν τα ακόλουθα κριτήρια:

- $U_{i=1}^n R_i = R$
- R_i είναι συνεκτική περιοχή για $i = 1, 2, \dots, N$
- $R_i \cap R_j = \emptyset$ για $i \neq j$
- $H(R_i) = \text{True}$ για $i = 1, 2, \dots, N$
- $H(R_i \cup R_j) = \text{False}$, για κάθε γειτονική περιοχή R_i, R_j

Η συνάρτηση $H(R_i)$ αποτελεί μία λογική συνάρτηση που χρησιμοποιείται ως κριτήριο ομοιογένειας για τον προσδιορισμό των εικονοστοιχείων που ανήκουν στην ίδια περιοχή. Τα εικονοστοιχεία μιας περιοχής πρέπει να πληρούν αυτό το κριτήριο, ενώ εκείνα που βρίσκονται σε διαφορετικές περιοχές δεν πρέπει να το ικανοποιούν. Το κριτήριο ομοιογένειας μπορεί να βασίζεται σε διάφορα χαρακτηριστικά, όπως η ένταση του γκρι, το χρώμα, η υφή ή το σχήμα [167].

Ακολουθεί η περιγραφή της διαδικασίας υλοποίησης του κλασικού αλγορίθμου region growing. Αρχικά, δημιουργείται ένα σύνολο από μικρές, αρχικές περιοχές που ορίζονται μέσω ενός αντίστοιχου συνόλου σημείων, γνωστών ως seeds (**Εικόνα 3.11**). Τα σημεία αυτά επιλέγονται έτσι ώστε να ικανοποιούν το κριτήριο ομοιογένειας. Στη συνέχεια, κάθε αρχική περιοχή επεκτείνεται προς τα γειτονικά εικονοστοιχεία που επίσης πληρούν το ίδιο κριτήριο και συγχωνεύονται με την περιοχή με αποτέλεσμα τη σταδιακή αύξηση του μεγέθους της.

Μόλις ολοκληρωθεί η ανάπτυξη μιας περιοχής επιλέγεται ένα νέο seed που δεν ανήκει ακόμα σε κάποια περιοχή και η διαδικασία επαναλαμβάνεται. Η διαδικασία αυτή συνεχίζεται μέχρις ότου όλα τα εικονοστοιχεία της εικόνας να έχουν καταταχιστούν σε κάποια περιοχή [168].



Εικόνα 3.11: Η εξέλιξη του τυπικού αλγορίθμου region-growing [169]

Τα πλεονεκτήματα της τεχνικής Region Growing περιλαμβάνουν την αποτελεσματική τμηματοποίηση εικόνων που διαθέτουν καθαρές και διακριτές ακμές, την απλή και ευέλικτη υλοποίηση, καθώς και την ικανότητα να χρησιμοποιεί ποικίλα κριτήρια ομοιότητας, ακόμα και να συνδυάζει περισσότερα από ένα. Επιπλέον, η μέθοδος παρουσιάζει ανθεκτικότητα στον θόρυβο. Ωστόσο, ανάμεσα στα μειονεκτήματα συγκαταλέγονται ο σχετικά μεγάλος χρόνος επεξεργασίας, η έντονη εξάρτηση από την αρχική επιλογή των σημείων εκκίνησης, καθώς και η πιθανότητα παραγωγής περιοχών που δεν είναι πλήρως τμηματοποιημένες. Για την αντιμετώπιση αυτών των δυσκολιών, έχουν προταθεί διάφορες προσεγγίσεις οι οποίες αξιοποιούν χαρακτηριστικά της εικόνας, όπως τη σύγκριση και την ομοιότητα μεταξύ γειτονικών περιοχών, την εφαρμογή περιορισμών στην εξέλιξη της διαδικασίας, όπως τον περιορισμό ώστε να μην επικρατεί μόνο μια περιοχή ή την ταυτόχρονη ανάπτυξη πολλαπλών περιοχών, καθώς και την αξιοποίηση παράλληλων υπολογιστικών πόρων [169].

3.2.4 Ενεργά Περιγράμματα (Active Contours)

Η μέθοδος των Ενεργών Περιγραμμάτων (Active Contours - AC), βασίζεται στις αρχές της δυναμικής εξέλιξης καμπυλών [157], [158], [159]. Ο βασικός σκοπός αυτών των τεχνικών είναι η ελαχιστοποίηση μιας συνάρτησης ενέργειας η οποία υπόκειται σε συγκεκριμένους περιορισμούς. Οι περιορισμοί αυτοί διακρίνονται σε δύο τύπους, τους εξωτερικούς που σχετίζονται με χαρακτηριστικά της εικόνας όπως η ένταση των εικονοστοιχείων και τους εσωτερικούς που συνδέονται με ιδιότητες της ίδιας της καμπύλης όπως το μήκος ή η ελαστικότητα της.

Ανάλογα με τον τρόπο με τον οποίο η καμπύλη ενσωματώνεται στη συνάρτηση ενέργειας, οι τεχνικές αυτές κατηγοριοποιούνται σε δύο κύριες ομάδες:

- Ευθέως ορισμένα ενεργά περιγράμματα (Explicit Active Contours)
- Έμμεσα ενεργά περιγράμματα (Implicit Active Contours)

Στην πρώτη κατηγορία, τα ευθέως ορισμένα ενεργά περιγράμματα (explicit AC), γνωστά και ως παραμετρικά ενεργά περιγράμματα (parametric AC), η διαδικασία τμηματοποίησης πραγματοποιείται μέσω της μετακίνησης σημείων που ανήκουν σε μια ρητά ορισμένη καμπύλη C_0 [173], [174]. Ο στόχος είναι η ελαχιστοποίηση μιας κατάλληλα διαμορφωμένης συνάρτησης ενέργειας (Energy Functional), η οποία έχει σχεδιαστεί ώστε να λαμβάνει τις χαμηλότερες τιμές στις ακμές των αντικειμένων ενώ παράλληλα εξαρτάται άμεσα από την παραμετροποίηση της καμπύλης.

Το αρχικό μοντέλο αποτελείται από δύο βασικά μέρη, την εσωτερική ενέργεια (Internal Energy) και την εξωτερική ενέργεια (External Energy). Η εσωτερική ενέργεια καθορίζει την

ομαλότητα του περιγράμματος ενώ η εξωτερική ενέργεια κατευθύνει την καμπύλη C_0 προς τις ακμές της εικόνας. Η συνολική συνάρτηση ενέργειας διαμορφώνεται ως εξής:

$$E(C) = \alpha \int_0^1 \left| \frac{dC(q)}{dq} \right|^2 dq + \beta \int_0^1 \left| \frac{d^{(2)}C(q)}{dq^2} \right|^2 dq - \lambda \int_0^1 |\nabla I(C(q))| dq \quad (3-29)$$

όπου:

$$E_{internal} = \alpha \int_0^1 \left| \frac{dC(q)}{dq} \right|^2 dq + \beta \int_0^1 \left| \frac{d^{(2)}C(q)}{dq^2} \right|^2 dq \quad (3-30)$$

$$E_{external} = \lambda \int_0^1 |\nabla I(C(q))| dq \quad (3-31)$$

Επιπλέον, για την καμπύλη C ισχύει ότι $C(q): [0,1] \rightarrow \mathbb{R}^2$, ενώ για την εικόνα I που πρόκειται από την τμηματοποίηση έχουμε ότι $I(x, y): [0, a] \times [0, b] \rightarrow \mathbb{R}^+$. Οι παράμετροι α , β και λ θεωρούνται σταθερές και ανήκουν στο σύνολο των πραγματικών αριθμών, δηλαδή ισχύει $\alpha, \beta, \lambda \in \mathbb{R}$. Συνεπώς, η ενεργειακή συνάρτηση μπορεί να ελαχιστοποιηθεί αποκλειστικά με την κατάλληλη επιλογή της μορφής ή της τοπολογίας της καμπύλης C .

Ωστόσο, τα άμεσα καθορισμένα ενεργά περιγράμματα εμφανίζουν ορισμένα ουσιαστικά μειονεκτήματα. Πρώτον, δεν είναι εφικτός ο ταυτόχρονος εντοπισμός περισσότερων από ενός αντικειμένων, ενώ η χρήση πολλών περιγραμμάτων οδηγεί σε σημαντική αύξηση του υπολογιστικού φόρτου, χωρίς να διασφαλίζεται απαραίτητα η πλήρης αντιμετώπιση του προβλήματος. Επιπλέον, το μοντέλο δεν διαθέτει την ικανότητα να προσαρμόζεται με ακρίβεια στις ιδιαιτερότητες του πραγματικού περιγράμματος, με αποτέλεσμα η τελική μορφή να εξαρτάται έντονα από τις παραμέτρους που ορίζουν την καμπύλη.

Για να ξεπεραστούν οι περιορισμοί που παρουσιάζουν τα ευθέως ορισμένα Ενεργά Περιγράμματα (ΕΠ), εισήχθησαν νέα υπολογιστικά μοντέλα, τα οποία αξιοποιούν τη θεωρία της εξέλιξης καμπυλών και τις γεωμετρικές ροές [157], [158]. Η προσέγγιση αυτή επέφερε μια ουσιαστική αλλαγή στον τρόπο προσδιορισμού της αρχικής καμπύλης C_0 , η οποία αντί να ορίζεται άμεσα, χρησιμοποιείται πλέον ένας έμμεσος τρόπος, μέσω μιας βοηθητικής συνάρτησης φ . Η συνάρτηση αυτή γνωστή ως συνάρτηση Ίσων Επιπέδων (Level Set function), συνδέεται με την καμπύλη C_0 μέσω της ακόλουθης σχέσης:

$$C_0: \varphi = 0 \quad (3-32)$$

Τα σημεία όπου η συνάρτηση φ αποκτά μηδενική τιμή θεωρούνται ότι καθορίζουν την καμπύλη C_0 . Για τον λόγο αυτό, τα μοντέλα Ενεργών Περιγραμμάτων (ΕΠ) που βασίζονται σε αυτή την προσέγγιση αποκαλούνται Έμμεσα ή Γεωμετρικά ΕΠ. Η καμπύλη C_0 εξελίσσεται μέσω μιας δύναμης που ασκείται κάθετα στην επιφάνειά της με αποτέλεσμα η μαθηματική περιγραφή της κίνησης να γίνεται μέσω μιας μερικής διαφορικής εξίσωσης η οποία συχνά βασίζεται στη μέση καμπυλότητα (geometric flow). Η διαδικασία αυτή συνεχίζεται έως ότου η καμπύλη φτάσει σε σημείο ισορροπίας που θεωρείται η τελική της μορφή.

Αυτή η κατηγορία μοντέλων προσφέρει σειρά πλεονεκτημάτων συγκριτικά με τα παραδοσιακά ευθέως ορισμένα ΕΠ. Κυριότερα εξ αυτών είναι η ικανότητά τους να διαχωρίζουν πολλαπλά αντικείμενα με φυσικό τρόπο, να προσαρμόζονται στο περιεχόμενο της εικόνας χωρίς προ-υπάρχουσες πληροφορίες, καθώς και η ευκολία με την οποία μπορούν να επεκταθούν σε προβλήματα υψηλότερης διάστασης, όπως η απευθείας τρισδιάστατη τμηματοποίηση ή η παρακολούθηση στο χρόνο (όπως σε ακολουθίες εικόνων ή βίντεο).

Ωστόσο, τα συμβατικά έμμεσα ΕΠ παρουσιάζουν και ορισμένα μειονεκτήματα. Η σταδιακή μεταβολή της καμπύλης με την πάροδο του χρόνου μπορεί να οδηγήσει σε παραμορφώσεις, γεγονός που απαιτεί την εκ νέου αρχικοποίησή της σε τακτά διαστήματα. Επιπλέον, η απόδοση του αλγορίθμου είναι ευαίσθητη στις αρχικές συνθήκες ενώ η ύπαρξη πολλών παραμέτρων απαιτεί προσεκτική ρύθμιση ώστε να προσαρμόζεται στις ιδιαιτερότητες κάθε περίπτωσης για βέλτιστη απόδοση [175], [176]. Για την αντιμετώπιση αυτών των ζητημάτων έχουν αναπτυχθεί διάφορες παραλλαγές που στοχεύουν στη βελτίωση της απόδοσης. Στην ενότητα που ακολουθεί περιγράφονται τα βασικά χαρακτηριστικά και οι λειτουργίες που διέπουν τα έμμεσα ΕΠ (Εικόνα 3.12).

3.2.4.1 Κίνηση έμμεσου περιγράμματος

Οι παράγοντες που επηρεάζουν την πορεία της καμπύλης αναφέρονται συνήθως ως όροι που σχετίζονται με τα δεδομένα (data terms). Η μεταβολή της καμπύλης μηδενικού ισοεπιπέδου πραγματοποιείται προς την κατεύθυνση της εξωτερικής κάθετης σε αυτή, δηλαδή κατά μήκος του διανύσματος $\nabla\varphi$. Παρόλο αυτά, γίνεται χρήση του μοναδιαίου διανύσματος της εξωτερικής κάθετου $\vec{N} = \frac{\nabla\varphi}{|\nabla\varphi|}$.

Όταν υπάρχει ένα εξωτερικό διανυσματικό πεδίο $\vec{F}_{ext}$, η μετατόπιση της καμπύλης του μηδενικού ισοεπιπέδου πραγματοποιείται κάθετα προς αυτήν και περιγράφεται μέσω της παρακάτω μερικής διαφορικής εξίσωσης [177], [178].

$$\frac{\partial\varphi}{\partial t} = -\vec{F}_{ext} \cdot \nabla\varphi \quad (3-33)$$

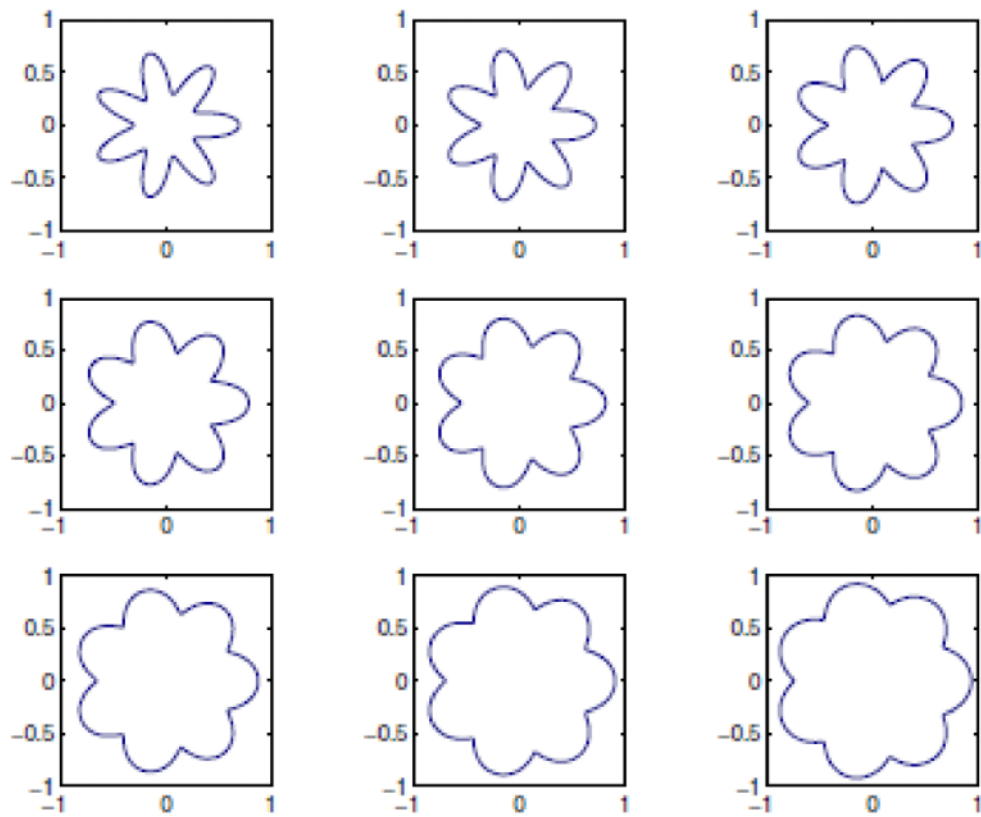
Εφόσον υφίστανται εσωτερικές δυνάμεις F_{int} , τροποποιείται η αρχική εξίσωση με την προσθήκη ενός μοναδιαίου κάθετου διανύσματος προκειμένου να διασφαλιστεί η μαθηματική της συνέπεια και να επιτρέπεται ο ορισμός του εσωτερικού γινομένου μεταξύ διανυσμάτων [179].

Η απλούστερη μορφή μιας τέτοιας εσωτερικής δύναμης βασίζεται στην καμπυλότητα κ , πολλαπλασιασμένη με έναν συντελεστή b σε κάθε σημείο $F_{int} = -b \cdot \kappa$. Η εξέλιξη της καμπύλης μηδενικού ισοεπιπέδου περιγράφεται τότε από την ακόλουθη μερική διαφορική εξίσωση:

$$\frac{\partial \varphi}{\partial t} = -(F_{int} - \vec{N}) \cdot \nabla \varphi = -F_{int} \cdot |\nabla \varphi| \quad (3-34)$$

Τέλος, όταν συνυπάρχουν τόσο εσωτερικές όσο και εξωτερικές δυνάμεις, η συνολική κίνηση προκύπτει όπως φαίνεται παρακάτω από το άθροισμα των δυο επιμέρους συναρτήσεων:

$$\frac{\partial \varphi}{\partial t} = -\vec{F}_{ext} \cdot \nabla \varphi - F_{int} \cdot |\nabla \varphi| \quad (3-35)$$



Εικόνα 3.12: Μεταβολή ενός ενεργού περιγράμματος υπό την επίδραση της καμπυλότητας ως μοναδικής εσωτερικής δύναμης [180]

3.2.4.2 Όροι Εξομάλυνσης

Κατά τη διάρκεια της διαδικασίας εξέλιξης της καμπύλης μηδενικού ισοεπιπέδου, είναι απαραίτητο να πληρούνται ορισμένοι περιορισμοί που συμβάλλουν στη διατήρηση της ομαλής μορφής της [167]. Οι περιορισμοί αυτοί περιγράφονται μέσω μαθηματικών εκφράσεων γνωστών ως όροι εξομάλυνσης (regularization terms) και συνδυάζονται με τους όρους που προέρχονται από τα δεδομένα, ώστε να καθορίσουν με ακρίβεια την εξέλιξη της καμπύλης. Ένας από τους πιο κοινούς όρους εξομάλυνσης είναι αυτός που επιδιώκει τη μείωση του μήκους της καμπύλης, αναγκάζοντάς την να περιβάλλει την αντίστοιχη περιοχή με το μικρότερο δυνατό περίγραμμα. Ο συγκεκριμένος όρος περιγράφεται από τη σχέση:

$$R_{Length} = \text{div}\left(\frac{\nabla\varphi}{|\nabla\varphi|}\right) \quad (3-36)$$

Με αυτόν τον τρόπο αποτρέπεται η υπερβολική χαλαρότητα της καμπύλης, η οποία μπορεί να προκαλέσει ταλαντώσεις γύρω από τα πραγματικά σημεία ισορροπίας. Μία ακόμη συνθήκη που συναντάται συχνά αφορά στη διατήρηση της φ ως προσημασμένης συνάρτησης απόστασης και εκφράζεται μέσω της [181]:

$$R_{SDF} = \nabla^2\varphi - \text{div}\left(\frac{\nabla\varphi}{|\nabla\varphi|}\right) \quad (3-37)$$

Η συγκεκριμένη ιδιότητα είναι απαραίτητο να ικανοποιείται σε κάθε περίπτωση, ακόμη και κατά προσέγγιση, προκειμένου το μαθηματικό μοντέλο να παραμένει έγκυρο. Πιο αναλυτικά, η ιδανική συνάρτηση των ισοεπιπέδων φ θα πρέπει να παρουσιάζει τα εξής χαρακτηριστικά:

- $|\nabla\varphi| = 1$
- $\varphi < 0$, $x \in \Omega^-$, εντός της περιοχής του προσκηνίου
- $\varphi > 0$, $x \in \Omega^+$, εκτός της περιοχής του προσκηνίου (εντός του φόντου)

Η συνάρτηση φ εφαρμόζεται πρακτικά μέσω της συνάρτησης προ-σημασμένης απόστασης η οποία καταφέρνει να προσεγγίσει με ακρίβεια τις επιθυμητές ιδιότητες. Με την πάροδο του χρόνου όμως η ιδιότητα $|\nabla\varphi| = 1$ τείνει να αλλοιώνεται εξαιτίας της δυναμικής μεταβολής της καμπύλης. Για την αποκατάσταση αυτής της συνθήκης εφαρμόζεται η διαδικασία της επαναρχικοποίησης όπου μέσω αυτής, οι τιμές της φ επαναυπολογίζονται λαμβάνοντας ως σημείο εκκίνησης την καμπύλη $\varphi = 0$. Η διαδικασία αυτή είναι απαραίτητο να επαναλαμβάνεται ανά τακτά χρονικά διαστήματα με τη συχνότητά της να εξαρτάται από τη

συγκεκριμένη μέθοδο που εφαρμόζεται και τα χαρακτηριστικά της υπό επεξεργασία εικόνας. Η μαθηματική εξίσωση που χρησιμοποιείται πιο συχνά για το σκοπό αυτό είναι η εξίσωση [171].

$$\frac{\partial \phi}{\partial t} = \text{sign}(\phi_0) \cdot (1 - |\nabla \phi|) \quad (3-38)$$

Παρόλα αυτά, η προσέγγιση αυτή εισάγει έναν νέο παράγοντα που επηρεάζει το τελικό αποτέλεσμα και αυτός είναι η συχνότητα επαναρχικοποίησης. Για να περιοριστεί αυτή η εξωτερική επίδραση και να καταστεί η διαδικασία επαναρχικοποίησης πιο αποδοτική, προτάθηκε η ενσωμάτωση ενός επιπλέον όρου εξομάλυνσης στη συνάρτηση ενέργειας. Ο ρόλος του όρου είναι να ελέγχει τον βαθμό απόκλισης της συνάρτησης ϕ από την ιδανική υπογεγραμμένη απόσταση (Signed Distance Function - SDF) [181]. Η μαθηματική μορφή του όρου αυτού είναι η ακόλουθη:

$$P(\phi) = \int \frac{1}{2} (|\nabla \phi| - 1)^2 dx \quad (3-39)$$

Η μερική παραγωγή της Εξ.(3-39), οδηγεί στην Εξ.(3-38).

Τα γενικά χαρακτηριστικά που αναφέρθηκαν προηγουμένως ισχύουν για όλα τα είδη των έμμεσων AC. Παρ' όλα αυτά, ανάλογα με τη φύση των δεδομένων που χρησιμοποιούνται, προκύπτουν ορισμένες επιμέρους κατηγορίες. Ειδικότερα, όταν το μοντέλο αξιοποιεί πληροφορίες σχετικές με ακμές, τότε κατατάσσεται στα edge-based μοντέλα [182], [183], [184]. Αντίθετα, όταν γίνεται χρήση δεδομένων περιοχής, προκύπτουν τα λεγόμενα region-based μοντέλα [178], [185], [186], [187]. Στην ενότητα που ακολουθεί, παρουσιάζονται συνοπτικά οι κύριοι τύποι αυτών των μοντέλων καθώς και οι πιο αντιπροσωπευτικοί αλγόριθμοι που σχετίζονται με το καθένα.

3.2.4.3 Edge-Based Μοντέλα

Τα Active Contours που βασίζονται στις ακμές (edge-based AC) αξιοποιούν τις ακμές της εικόνας ως φράγματα που εμποδίζουν την περαιτέρω εξέλιξη της καμπύλης. Πρόκειται για την πρώτη κατηγορία έμμεσων Active Contours που προτάθηκε στη σχετική βιβλιογραφία. Ο εντοπισμός των ακμών επιτυγχάνεται μέσω μιας συνάρτησης $g(I)$, γνωστής ως συνάρτηση φραγής ακμών (edge stopping function – ESF), η οποία καθορίζει την κατεύθυνση και τον ρυθμό μετακίνησης του περιγράμματος. Η συνάρτηση αυτή συνήθως λαμβάνει την εξής μορφή:

$$g(I) = \frac{1}{1 + |\nabla I|} \quad (3-40)$$

Η καμπύλη εξελίσσεται με τη βοήθεια μιας σταθερής δύναμης, γνωστής ως δύναμη μπαλονιού (balloon force), επειδή κατά την επανάληψη φουσκώνει την καμπύλη με τρόπο παρόμοιο με τον αέρα που φουσκώνει ένα μπαλόνι. Ωστόσο, αυτό το μοντέλο είναι ευαίσθητο σε θόρυβο και απαιτεί χειροκίνητη αρχικοποίηση και ρύθμιση της δύναμης ανά περίπτωση. Για το λόγο αυτό δεν μπορεί να εφαρμοστεί ευρέως στην αρχική του μορφή.

Το γεωδαιτικό μοντέλο Ενεργού Περιγράμματος (geodesic active contour, GAC) αποτελεί μια φυσική εξέλιξη του αρχικού έμμεσου Ενεργού Περιγράμματος, όπου επίσης χρησιμοποιείται η μέθοδος ESF [183], [188]. Το όνομά του προέρχεται από την έννοια της γεωδαιτικής απόστασης που ορίζεται ως η μικρότερη απόσταση μεταξύ δύο σημείων σε έναν πολυδιάστατο χώρο, δηλαδή σε έναν χώρο με διάσταση τρία ή μεγαλύτερη. Η συνάρτηση $g_{GAC}(I)$ εφαρμόζει εξομάλυνση της αρχικής εικόνας μέσω μιας συνάρτησης Gauss που συνήθως παίρνει τη μορφή:

$$g_{GAC} = \frac{1}{1 + |\nabla G \cdot I|^p} \quad (3-41)$$

με $p = 1$ ή 2 . Η συνάρτηση της δύναμης είναι η ακόλουθη:

$$\begin{aligned} \frac{\partial \varphi}{\partial t} &= |\nabla \varphi| \cdot \operatorname{div} \left(g_{GAC}(I) \cdot \frac{\nabla \varphi}{|\nabla \varphi|} \right) = g_{GAC}(I) \cdot |\nabla \varphi| \cdot \operatorname{div} \left(\frac{\nabla \varphi}{|\nabla \varphi|} \right) + \nabla g_{GAC}(I) \cdot \nabla \varphi \\ &= g_{GAC}(I) \cdot |\nabla \varphi| \cdot \kappa + \nabla g_{GAC}(I) \cdot \nabla \varphi \end{aligned} \quad (3-42)$$

όπου κ είναι η καμπυλότητα και δίνεται από την ακόλουθη σχέση :

$$\kappa = \operatorname{div} \left(\frac{\nabla \varphi}{|\nabla \varphi|} \right) \quad (3-43)$$

Το γεωδαιτικό μοντέλο ΕΠ αποτελεί επέκταση του κλασικού γεωμετρικού μοντέλου, εισάγοντας τον όρο $c \cdot g_{GAC}(I) \cdot |\nabla \varphi|$. Αυτός ο όρος προσελκύει την καμπύλη προς τις ακμές της εικόνας και παράλληλα δεν εξαρτάται από την καμπυλότητα κ . Η τελική έκφραση της δύναμης δίνεται από την ακόλουθη σχέση:

$$\begin{aligned}
\frac{\partial \varphi}{\partial t} &= |\nabla \varphi| \cdot \operatorname{div} \left(g_{GAC}(I) \cdot \frac{\nabla \varphi}{|\nabla \varphi|} \right) + c \cdot g_{GAC}(I) \cdot |\nabla \varphi| \\
&= g_{GAC}(I) \cdot |\nabla \varphi| \cdot (\kappa + c) + \nabla g_{GAC}(I) \cdot \nabla \varphi
\end{aligned}
\tag{3-44}$$

Το GAC διευκολύνει την ανίχνευση ακμών με σημαντικές διαφορές στις τιμές τους και καταργεί την ανάγκη για τον προκαθορισμό της σταθεράς δύναμης του μπαλονιού, η οποία συνήθως εκφράζεται μέσω της παραμέτρου c . Ωστόσο, η μέθοδος αυτή είναι ευαίσθητη στην αρχική θέση του περιγράμματος καθώς και στη μορφή της συνάρτησης ESF, $g_{GAC}(I)$ [189].

3.2.4.4 Region-Based Μοντέλα

Οι region-based AC (**Εικόνα 3.13**) αξιοποιούν στατιστικά στοιχεία τόσο από το εσωτερικό όσο και από το εξωτερικό της καμπύλης ώστε να καθοδηγήσουν την εξέλιξή της. Ένα από τα πιο χαρακτηριστικά και σημαντικά παραδείγματα αυτής της κατηγορίας είναι το μοντέλο Chan-Vese (CV) [186], το οποίο ενσωματώνει την ενεργειακή προσέγγιση του Mumford-Shah [190] με τη μεθοδολογία των Level Sets [170] για την παρακολούθηση της καμπύλης.

Κατά τη διάρκεια κάθε επανάληψης της εξέλιξης της καμπύλης το μοντέλο βασίζεται στη διαφορά μεταξύ της τοπικής τιμής έντασης κάθε εικονοστοιχείου και της μέσης έντασης των περιοχών που ορίζονται ως εσωτερικές (foreground) και εξωτερικές (background) της καμπύλης.

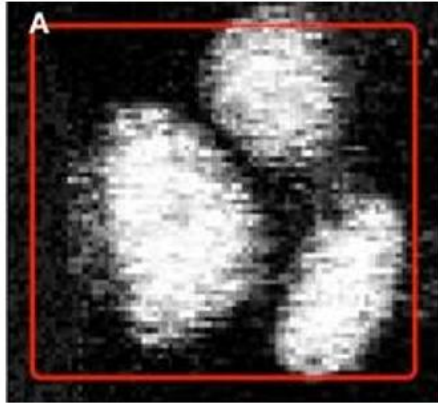
Το CV μοντέλο προσφέρει αρκετά πλεονεκτήματα σε σύγκριση με προηγούμενες μεθόδους, όπως την εξάλειψη της ανάγκης επαναρχικοποίησης μετά από κάθε βήμα, τη δυνατότητα ανίχνευσης πολλών υποπεριοχών, την υπολογιστική του απλότητα, καθώς και την ευκολία γενίκευσης σε πολυδιάστατα δεδομένα.

Παρά τα παραπάνω, το μοντέλο παρουσιάζει συγκεκριμένα μειονεκτήματα κυρίως λόγω της αποκλειστικής χρήσης ολικής πληροφορίας (global information) και της εξάρτησης από ένα μόνο χαρακτηριστικό δηλαδή την ένταση των εικονοστοιχείων. Αυτό έχει ως αποτέλεσμα τη μείωση της αποτελεσματικότητας σε περιβάλλοντα με έντονο θόρυβο ή ανομοιογενή ένταση αλλά και σε περιπτώσεις όπου οι διαφορές έντασης μεταξύ διαφορετικών περιοχών foreground είναι περιορισμένες.

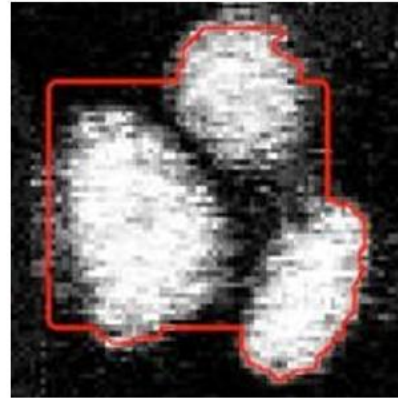
Επιπλέον, παρότι λιγότερο έντονη, παραμένει η ευαισθησία στην αρχικοποίηση κάτι που επιβάλλει ακόμη την ανάγκη χειροκίνητης οριστικοποίησης της αρχικής μάσκας, ιδίως σε πραγματικές εικόνες. Η ενεργειακή συνάρτηση που χρησιμοποιείται στο πλαίσιο του μοντέλου CV έχει την εξής μορφή:

$$E_{CV}(c_1, c_2, \varphi) = \mu \cdot \text{Length}(\varphi) + v \cdot \text{Area}(\text{inside}(\varphi)) + \lambda_1 \cdot \int_{\varphi \leq 0} |I(x, y) - c_1|^2 dx dy - \lambda_2 \cdot \int_{\varphi > 0} |I(x, y) - c_2|^2 dx dy \quad (3-45)$$

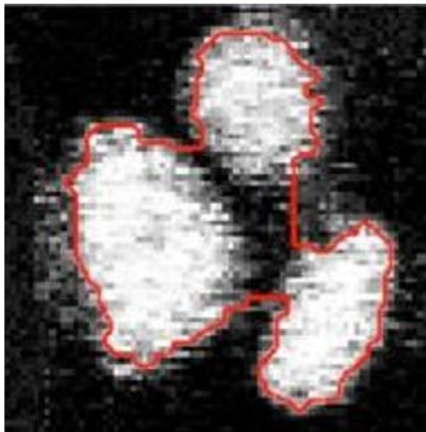
όπου φ είναι η συνάρτηση που παριστάνει τη λειτουργία των ισοεπιπέδων, ενώ οι τιμές c_1 και c_2 αφορούν την μέση τιμή της έντασης εντός της καμπύλης για $\varphi \leq 0$ και εκτός αυτής για $\varphi > 0$ αντίστοιχα. Ο συντελεστής μ χρησιμοποιείται για την κανονικοποίηση του μήκους της καμπύλης και ο συντελεστής v εκφράζει τη βαρύτητα που αποδίδεται στο εμβαδόν της περιοχής που περικλείεται από την καμπύλη. Τέλος, οι παράμετροι λ_1 και λ_2 προσδιορίζουν τους συντελεστές βαρύτητας της εξωτερικής και εσωτερικής δύναμης αντίστοιχα.



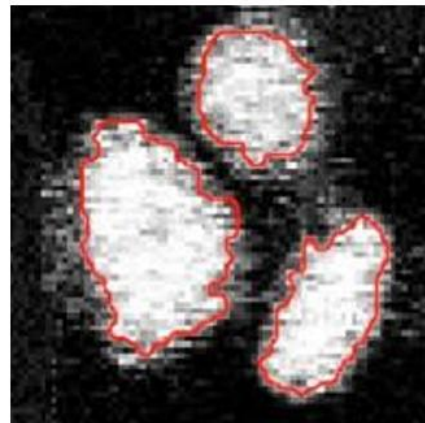
(α)



(β)



(γ)



(δ)

Εικόνα 3.13: Region-based εξέλιξη ενεργού περιγράμματος σε εικόνα κυττάρων. (α) Αρχικό περίγραμμα, (β-γ) Ενδιάμεσες καταστάσεις, (δ) Τελική τμηματοποίηση [191]

Μια ουσιώδης τροποποίηση σε σχέση με το αρχικό μοντέλο είναι η αντικατάσταση του όρου κανονικοποίησης $\mu \cdot Length(\varphi)$ με τη διαδικασία συνέλιξης της καμπύλης φ με μια καμπύλη Gauss στο τέλος κάθε επανάληψης [192], [193]. Η μεθοδολογία αυτή στηρίζεται στη θεωρία του χώρου κλίμακας (scale-space) [194] μέσω της οποίας επιτυγχάνεται σημαντική μείωση του υπολογιστικού φορτίου κατά την εξέλιξη της καμπύλης, γεγονός που συνεπάγεται ταχύτερους χρόνους εκτέλεσης και διευκολύνει την εφαρμογή της μεθόδου σε πραγματικά περιβάλλοντα.

Μια εναλλακτική μέθοδος προσέγγισης είναι η αξιοποίηση συναρτήσεων που είναι τμηματικά ομαλές [195] με στόχο την καλύτερη διαχείριση της ανομοιογένειας η οποία αποτελεί συχνό χαρακτηριστικό των πραγματικών και ιατρικών εικόνων. Ωστόσο, αυτή η μέθοδος συνοδεύεται από αυξημένες υπολογιστικές απαιτήσεις, γεγονός που περιορίζει την πρακτική της απόδοση.

Μια διαφορετική στρατηγική περιλαμβάνει τη χρήση τοπικής πληροφορίας [178], [192], [196], [197], [198], ώστε να ενισχυθεί η προσαρμοστικότητα του μοντέλου στις μεταβολές της εικόνας. Συνήθως, οι τεχνικές αυτές εφαρμόζουν μια global δύναμη όταν η εξελισσόμενη καμπύλη είναι απομακρυσμένη από τα όρια των αντικειμένων, ενώ επιστρατεύουν local δυνάμεις όταν πλησιάζει τα όρια με στόχο την καλύτερη αποτύπωση των λεπτομερειών. Αν και αυτές οι μέθοδοι είναι πιο κατάλληλες για την αντιμετώπιση φαινομένων ανομοιογένειας, παρουσιάζουν ευαισθησία στο θόρυβο και στις αρχικές συνθήκες γεγονός που δυσκολεύει την εφαρμογή τους σε πραγματικά δεδομένα.

Η φυσική εξέλιξη των Active Contours περιλαμβάνει την ενοποίηση των προσεγγίσεων που βασίζονται σε όρια (edge-based) και σε περιοχές (region-based) σε ένα ενιαίο μοντέλο [199], [200]. Ένα χαρακτηριστικό παράδειγμα αυτής της προσπάθειας είναι το μοντέλο Selective Binary and Gaussian Filtering Regularized Level Set (SBGFRLS), το οποίο συνδυάζει τις ιδέες των GAC και CV μοντέλων [201]. Το SBGFRLS εφαρμόζει μια region-based επιφανειακή δύναμη (SPF), σε αντίθεση με το αρχικό GAC που στηριζόταν σε πληροφορία ορίων και εστιάζει αποκλειστικά σε global στατιστικά στοιχεία. Παρά την καλή του απόδοση σε εικόνες με ομοιογενή χαρακτηριστικά, το μοντέλο δεν ανταποκρίνεται εξίσου καλά σε περιπτώσεις με έντονη ανομοιογένεια.

3.2.5 Μέθοδοι Ομαδοποίησης (K-Means)

Ο αλγόριθμος K-Means συγκαταλέγεται στις πιο ευρέως χρησιμοποιούμενες μεθόδους μη επιβλεπόμενης ομαδοποίησης δεδομένων και έχει αποδειχθεί ιδιαίτερα χρήσιμος στον χώρο της ιατρικής απεικόνισης. Η μέθοδος αυτή διαχωρίζει τα εικονοστοιχεία (pixels) μιας εικόνας σε k ομάδες (clusters), βασιζόμενη σε n κοινά χαρακτηριστικά. Η διαδικασία είναι επαναληπτική και λειτουργεί τοποθετώντας κάθε pixel στην ομάδα της οποίας το κέντρο υπολογίζεται ελαχιστοποιώντας την απόσταση:

$$V = \sum_{i=1}^k \sum_{x_j \in S_i} (x_j - \mu_i)^2 \quad (3-46)$$

όπου k δηλώνει το πλήθος των συστάδων S_i , $i = 1, 2, \dots, k$ και το μ_i αντιπροσωπεύει το κεντρικό pixel όλων των σημείων $x_j \in S_i$. Η διαδικασία που ακολουθεί είναι επαναληπτική και περιλαμβάνει τα εξής βήματα:

1. Υπολογίζεται η κατανομή της φωτεινότητας των pixels, δηλαδή δημιουργείται το αντίστοιχο ιστόγραμμα.
2. Τα αρχικά κέντρα των συστάδων επιλέγονται τυχαία από τις διαθέσιμες τιμές k φωτεινότητων. Κάθε κέντρο αντιστοιχεί σε μια τυχαία φωτεινότητα.
3. Εκτελείται επανάληψη των επόμενων δύο σταδίων μέχρι να σταθεροποιηθούν οι συστάδες (δηλαδή να μην αλλάζουν πλέον).
4. Η ομαδοποίηση των σημείων πραγματοποιείται με κριτήριο την απόσταση των τιμών φωτεινότητας από το κεντρικό σημείο κάθε συστάδας, σύμφωνα με τον ακόλουθο τύπο:

$$c^{(i)} := \underset{j}{\operatorname{argmin}} \|x^{(i)} - \mu_j\|^2 \quad (3-47)$$

5. Επαναπροσδιορισμός του κέντρου κάθε συστάδας χρησιμοποιώντας τη σχέση :

$$\mu_i := \frac{\sum_{i=1}^m 1\{c_{(i)} = j\} x^{(i)}}{\sum_{i=1}^m 1\{c_{(i)} = j\}} \quad (3-48)$$

όπου i αναπαριστά τον δείκτη επανάληψης για όλες τις τιμές φωτεινότητας k , το j αντιστοιχεί στον δείκτη για όλα τα κέντρα των συστάδων και το μ_i συμβολίζει τις αντίστοιχες τιμές φωτεινότητας των κέντρων αυτών [202].

Προκειμένου να μπορεί να υλοποιηθεί ο αλγόριθμος του K-Means, ορίστηκαν οι παρακάτω παράμετροι :

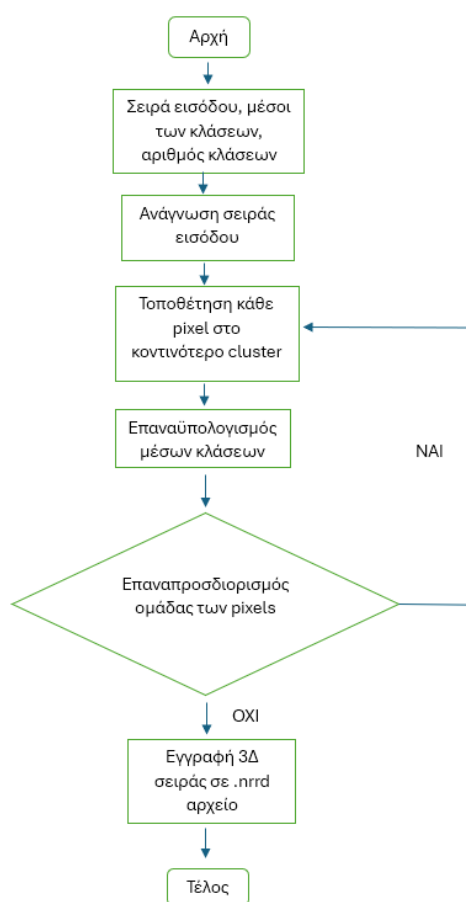
1. Non contiguous labels
2. Number of classes N
3. mean1
4. mean2....
5.meanN

Ο αλγόριθμος K-Means, υλοποιεί ένα μοντέλο ταξινόμησης με σκοπό την κατηγοριοποίηση των pixels της εισερχόμενης σειράς σε ομάδες σύμφωνα με τη μεθοδολογία που περιγράφεται στην προηγούμενη υποενότητα 3.2.5. Αρχικά ο reader αναλαμβάνει την ανάγνωση των δεδομένων εισόδου τα οποία στη συνέχεια μεταβιβάζονται στον K-Means αλγόριθμο για επεξεργασία ώστε στο τέλος το αποτέλεσμα να εξάγεται μέσω του writer σε αρχείο τύπου .nrrd.

Η πρώτη παράμετρος του μοντέλου καθορίζει αν θα χρησιμοποιηθεί λογική κατανομή των ετικετών εξόδου. Αν και η προεπιλεγμένη τιμή είναι false, συνήθως επιλέγεται η ρύθμιση true (1) ώστε οι ετικέτες να καλύπτουν όλο το φάσμα φωτεινότητας της εικόνας εξόδου. Σε διαφορετική περίπτωση, οι τιμές των ετικετών θα ήταν διαδοχικές (π.χ. 0, 1, 2, ..., N), με αποτέλεσμα η παραγόμενη εικόνα να εμφανίζεται σκοτεινή [203].

Ακολουθεί ο καθορισμός του αριθμού των κλάσεων και η ανάθεση αρχικών μέσων όρων για κάθε μία. Η επιλογή των μέσων τιμών και του αριθμού των ομάδων γίνεται κατόπιν μελέτης της κατανομής φωτεινότητας των pixels σε κάθε γραμμή με στόχο τη βέλτιστη κάλυψη του δυναμικού εύρους τους.

Τέλος, η συνάρτηση GetFinalMeans() χρησιμοποιείται για την ανάκτηση των τελικών μέσων όρων που προέκυψαν για κάθε κλάση μετά την εφαρμογή του αλγορίθμου. Το διάγραμμα ροής (Εικόνα 3.14) που ακολουθεί, απεικονίζει τη συνολική εκτέλεση της παραπάνω διαδικασίας [204], [205].



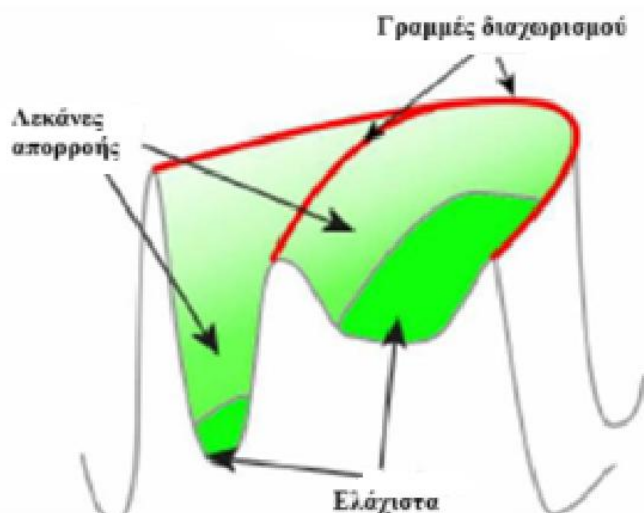
Εικόνα 3.14: Διάγραμμα ροής K-Means

3.2.6 Watershed

Ο ονομασία του αλγορίθμου watershed προέρχεται από τη μαθηματική μορφολογία και χρησιμοποιείται για να αποδώσει τη βασική ιδέα της μεθόδου. Για να γίνει πιο εύκολα κατανοητός ο τρόπος λειτουργίας του συχνά παρουσιάζεται με αναλογία σε γεωγραφικές και τοπογραφικές έννοιες.

Η τεχνική αυτή αντιμετωπίζει την εικόνα ως ένα είδος τοπογραφικού ανάγλυφου όπου διακρίνονται υψώματα και κοιλάδες. Σε αυτή την αναπαράσταση, κάθε σημείο στο χάρτη συσχετίζεται με μια φωτεινότητα σε κλίμακα του γκρι. Η βασική σύλληψη είναι ότι το νερό της βροχής συγκεντρώνεται στις χαμηλότερες περιοχές και σταδιακά τις πλημμυρίζει. Όταν το νερό γεμίσει αυτές τις περιοχές και αρχίσει να επικοινωνεί μεταξύ τους, δημιουργούνται φυσικά όρια που στον αλγόριθμο αντιστοιχούν στις γραμμές διαχωρισμού, δηλαδή τα λεγόμενα watersheds (**Εικόνα 3.15**) και στις περιοχές που περικλείονται από αυτές γνωστές ως λεκάνες απορροής ή catchment basins [206]. Οι λεκάνες αυτές αποτελούν τις τελικές τμηματοποιημένες περιοχές της εικόνας.

Μια εναλλακτική θεώρηση της ίδιας μεθόδου περιγράφει τη διαδικασία ως εξής: οι κοιλάδες θεωρούνται ότι έχουν ανοίγματα στη βάση τους, τα τοπικά ελάχιστα, από τα οποία αρχίζει να αναβλύζει νερό. Με αυτόν τον τρόπο, η πλημμύρα δεν έρχεται από πάνω, όπως στην πρώτη αναλογία, αλλά ξεκινά από τα χαμηλά σημεία και κινείται προς τα πάνω, διαμορφώνοντας διαφορετικά τα όρια των περιοχών [207], [208], [209].



Εικόνα 3.15: Γραμμές διαχωρισμού, Λεκάνες απορροής, Ελάχιστα [206]

Ο συγκεκριμένος αλγόριθμος ξεχωρίζει από τους υπολοίπους καθώς είναι ο μοναδικός που παράγει ως έξοδο μια ακολουθία εικόνων σε RGB κλίμακα [210]. Οι βασικές παράμετροι που ρυθμίζονται είναι οι εξής:

1. Threshold
2. Level

Αμφότερες καθορίζονται ως ποσοστά επί του μέγιστου ύψους της εικόνας και επηρεάζουν σε μεγάλο βαθμό το τελικό αποτέλεσμα, ενώ ακόμα και μικρές τροποποιήσεις στις τιμές τους μπορούν να έχουν αισθητή επίδραση. Η παράμετρος threshold καθορίζει το ελάχιστο ύψος που λαμβάνεται υπόψη, δηλαδή το σημείο εκκίνησης για την άνοδο της στάθμης του νερού. Αν κάποια pixel με χαμηλές τιμές δεν είναι σημαντικά για τον χρήστη, μπορεί να ορίσει ένα threshold που τα αγνοεί κατά την τμηματοποίηση. Έτσι, περιορίζονται τα τοπικά ελάχιστα και κατ' επέκταση τα λιγότερο σημαντικά σημεία, με αποτέλεσμα τη μείωση των περιοχών ενδιαφέροντος όσο αυξάνεται η τιμή του threshold [211].

Η παράμετρος level, καθορίζει τη μέγιστη στάθμη που φτάνει το υποθετικό νερό. Αν η τιμή της είναι 1 (δηλαδή 100% του μέγιστου ύψους) τότε η στάθμη φτάνει μέχρι τις κορυφές και δημιουργείται μία ενιαία περιοχή, γνωστή ως υποκατάτμηση. Αντιθέτως, με τιμή 0, το νερό σταματά σε πρώιμο στάδιο με αποτέλεσμα να σχηματίζονται πολλές μικρές περιοχές, γνωστή ως υπερκατάτμηση. Συνήθως, για μια ισορροπημένη κατάτμηση επιλέγεται τιμή γύρω στο 40% του μέγιστου ύψους [212].

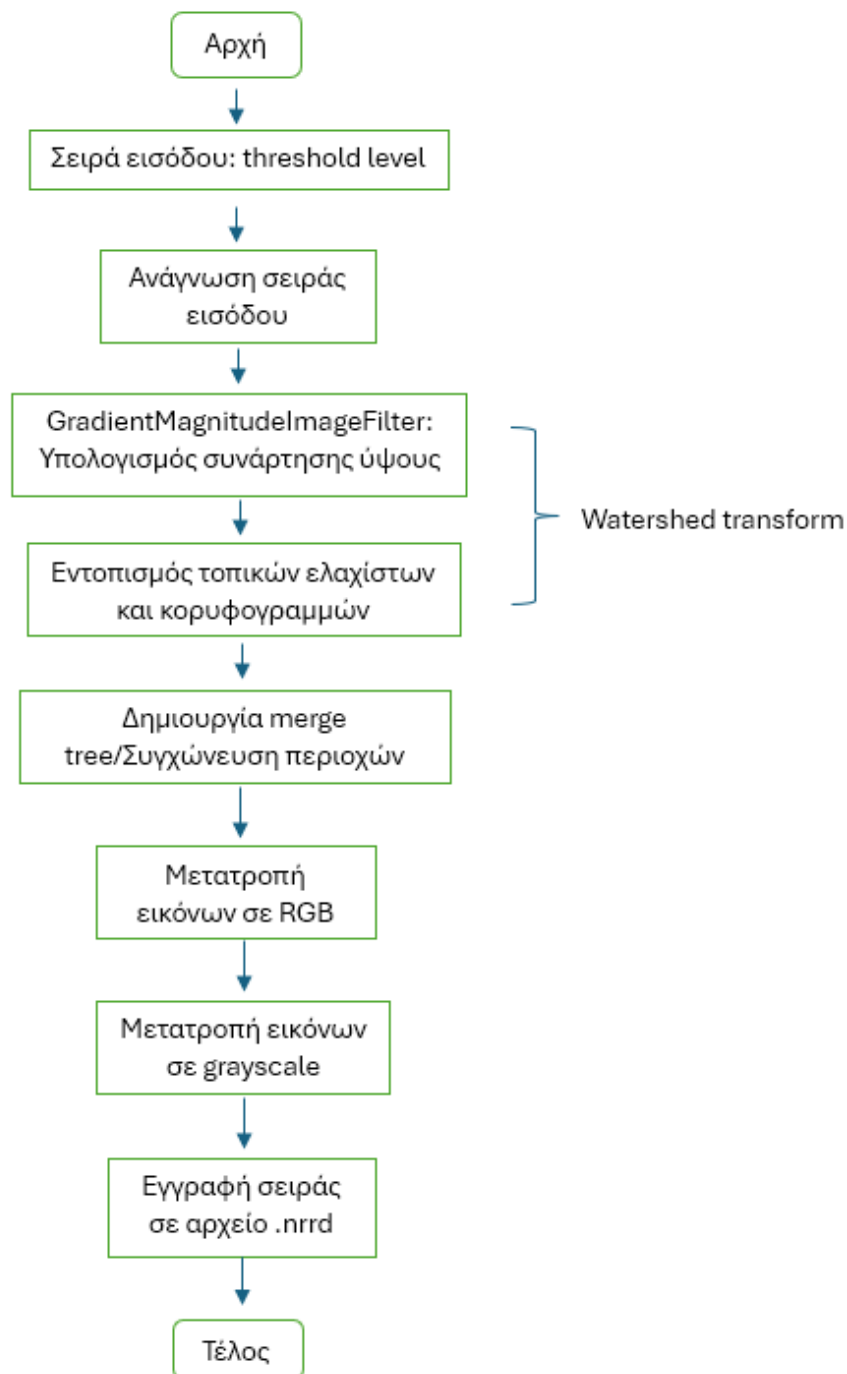
Εάν θεωρήσουμε την εικόνα ως συνάρτηση ύψους f , τότε οι λεκάνες απορροής ορίζονται ως τα σημεία που οδηγούν με τη μέγιστη κλίση προς ένα τοπικό ελάχιστο. Αυτό αποτελεί και το αρχικό στάδιο του αλγορίθμου watershed, κατά το οποίο η εικόνα διαχωρίζεται σε λεκάνες απορροής ακολουθώντας τα μονοπάτια που προκύπτουν. Στη συνέχεια, οι γειτονικές περιοχές αναλύονται βάσει ενός ποσοτικού χαρακτηριστικού με σκοπό τη δημιουργία ενός merge tree που περιγράφει τις συγχωνεύσεις περιοχών ανάλογα με το συγκεκριμένο κριτήριο. Όσο υψηλότερη είναι η τιμή του level τόσο περισσότερα όρια καταργούνται και οι περιοχές ενώνονται μεταξύ τους [213].

Στην υλοποίηση που ακολουθεί, παρουσιάζονται τα βήματα του αλγορίθμου: αρχικά, ο reader διαβάζει την τρισδιάστατη είσοδο και τη μεταφέρει μέσω του pipeline στο φίλτρο Gradient Magnitude Image Filter. Επειδή τα δεδομένα βασίζονται στη φωτεινότητα των pixels (intensity-based), το βήμα αυτό είναι αναγκαίο για τον υπολογισμό της κλίσης, δηλαδή της υψομετρικής συνάρτησης. Ακολουθεί η εφαρμογή του watershed φίλτρου. Στη συνέχεια, χρησιμοποιείται το φίλτρο Scalar to RGB Colormap Image Filter που μετατρέπει κάθε τιμή κλίσης σε αντίστοιχο RGB χρώμα μέσω της συνάρτησης SetColorMap(). Έπειτα, εφαρμόζεται το φίλτρο RGB to Luminance Image Filter για τη μετατροπή της εικόνας σε αποχρώσεις του γκρι ώστε τα αποτελέσματα να είναι συγκρίσιμα με εκείνα άλλων μεθόδων. Τέλος, η εικόνα αποθηκεύεται σε αρχείο .nrrd μέσω του writer [214].

Η επιλογή κατάλληλων τιμών για τις παραμέτρους αποδεικνύεται ιδιαίτερα απαιτητική και χρονοβόρα. Η βασική δυσκολία εντοπίζεται στην απουσία καθαρών ορίων ανάμεσα στους

ιστούς σε εικόνες μαγνητικής τομογραφίας, με αποτέλεσμα ο watershed αλγόριθμος συχνά να οδηγεί σε υποκατάτμηση. Παρόλα αυτά, σε εικόνες όπου οι δομές είναι καθαρά διαχωρισμένες από το περιβάλλον τους, ο αλγόριθμος παρουσιάζει ικανοποιητικά αποτελέσματα [215].

Ακολουθεί διάγραμμα ροής (Εικόνα 3.16) που αποτυπώνει την εκτέλεση του κώδικα με χρήση του watershed αλγορίθμου.



Εικόνα 3.16: Διάγραμμα ροής Watershed

3.2.7 Μαρκοβιανό Τυχαίο Πεδίο (Markov Random Field)

3.2.7.1 Πιθανοτική προσέγγιση

Ο Markov Random Field (MRF) [216] είναι από τους ανωτέρω ο πιο σύνθετος και πολύπλοκος αλγόριθμος. Πρόκειται για μια έννοια που εξετάζεται διεξοδικά στο πεδίο της θεωρίας πιθανοτήτων και βρίσκει ευρεία εφαρμογή στην Υπολογιστική Όραση και την ανάλυση εικόνας.

Στο πλαίσιο αυτό, η τμηματοποίηση εικόνας μπορεί να ερμηνευτεί ως η διαδικασία εκχώρησης μιας ετικέτας σε κάθε pixel. Άρα κάθε pixel s , αντιπροσωπεύει ένα χαρακτηριστικό διάνυσμα και επομένως για το σύνολο της εικόνας έχουμε:

$$f = \{\vec{f}_s : s \in S\} \quad (3-49)$$

Κάθε pixel s έχει μια ετικέτα $\omega_s \in \Lambda$ και επομένως για την συνολική εικόνα έχουμε:

$$\omega = \{\omega_s, s \in S\} \quad (3-50)$$

Επίσης υπό την παρακάτω προϋπόθεση, ορίζεται το πεδίο των ετικετών X ως Μαρκοβιανό:

$$\forall \omega \in \Omega: P(X = \omega) > 0 \quad (3-51)$$

και

$$P(\omega_s | \omega_r, r \neq s) = P(\omega_s | \omega_r, r \in N_s) \quad (3-52)$$

με N_s τα γειτονικά εικονοστοιχεία του s .

Κατ' αυτόν τον τρόπο, καθίσταται δυνατό να ορίσουμε την έννοια του Μαρκοβιανού Τυχαίου Πεδίου. Σύμφωνα με το θεώρημα Hammersley-Clifford, ένα τυχαίο πεδίο χαρακτηρίζεται ως Μαρκοβιανό αν και μόνο αν η κατανομή πιθανότητάς του είναι της ακόλουθης μορφής Gibbs :

$$P(\omega) = \frac{1}{Z} \exp(-U(\omega)) = \frac{1}{Z} \exp(-\sum_{c \in C} V_c(\omega)) \quad (3-53)$$

με $Z = \sum_{\omega \in \Omega} \exp(-U(\omega))$ να είναι η σταθερά κανονικοποίησης.

Μια εικόνα μπορεί να διαχωριστεί σε περιοχές με διάφορους τρόπους. Στο πλαίσιο μιας πιθανοτικής προσέγγισης με χρήση Μαρκοβιανών Τυχαίων Πεδίων (MRF), και δεδομένης της πιθανότητας $P(\omega|f)$, η βέλτιστη τμηματοποίηση ορίζεται ως εκείνη που μεγιστοποιεί αυτήν την πιθανότητα. Συνεπώς, επιδιώκουμε να εντοπίσουμε την κατάλληλη εκδοχή του $\hat{\omega}$, ώστε να ισχύει:

$$\hat{\omega}^{MAP} = \arg \max_{\omega \in \Omega} P(\omega|F) \quad (3-54)$$

όπου MAP η Maximum a Posteriori εκτίμηση.

Ο υπολογισμός αυτής της πιθανότητας βασίζεται στο θεώρημα του Bayes :

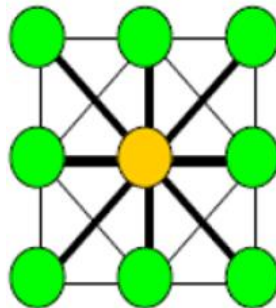
$$P(\omega|f) = \frac{P(f|\omega)P(\omega)}{P(f)} \propto P(f|\omega)P(\omega), P(f) \text{ σταθερά} \quad (3-55)$$

3.2.7.2 Μοντελοποίηση Μαρκοβιανού Τυχαίου Πεδίου

Ένα τυχαίο Μαρκοβιανό Πεδίο μπορεί να αναπαρασταθεί με έναν μη κατευθυνόμενο γράφο (**Εικόνα 3.17**), $G = (V, E)$ όπου το σύνολο V περιλαμβάνει τους κόμβους που αντιστοιχούν σε τυχαίες μεταβλητές, ενώ το σύνολο των ακμών E περιγράφει τις σχέσεις τοπικής ανεξαρτησίας. Οι σχέσεις αυτές γνωστές και ως τοπικές ιδιότητες του Markov, δηλώνουν ότι κάθε κόμβος είναι στατιστικά ανεξάρτητος από όλους τους υπόλοιπους, αρκεί να γνωρίζουμε τις τιμές των γειτονικών του κόμβων.

$$\forall_i \in V, X_i \perp X_{V-\{i\}} | X_{N_i} \quad (3-56)$$

με $N_i = \{j | \{i, j\} \in E\}$ συμβολίζει το σύνολο των κόμβων που είναι γείτονες του κόμβου i σε έναν γράφο G , το $X_i \perp X_j | X_k$ να δηλώνει ότι οι μεταβλητές X_i και X_j ανεξάρτητες υπό συνθήκη όταν γνωρίζουμε τις τιμές των μεταβλητών που αντιστοιχούν στους κόμβους X_k .



Εικόνα 3.17: Μη κατευθυνόμενος γράφος [217]

Αξίζει να αναφερθεί και ο ορισμός της κλίκας. Πρόκειται για ένα υποσύνολο κόμβων μέσα σε έναν γράφο, όπου κάθε κόμβος συνδέεται άμεσα με όλους τους υπόλοιπους. Μια τέτοια κλίκα χαρακτηρίζεται ως μέγιστη (maximal) όταν δεν μπορεί να επεκταθεί περαιτέρω, δηλαδή δεν αποτελεί τμήμα κάποιας μεγαλύτερης κλίκας. Σε αναλογία με μια εικόνα, μπορούμε να φανταστούμε τους κόμβους του γράφου ως εικονοστοιχεία (pixels), ενώ οι ακμές αναπαριστούν τις σχέσεις ή συσχετίσεις μεταξύ αυτών [218].

Στόχος της εφαρμογής του αλγορίθμου Markov Random Field (MRF), είναι η κατηγοριοποίηση των εικονοστοιχείων της εικόνας εισόδου. Η προσέγγιση αυτή αξιοποιεί την εκτίμηση Maximum A Posteriori (MAP) ώστε να αποδώσει κάθε pixel στην κατάλληλη κλάση με βάση το στατιστικό μοντέλο που έχει προηγουμένως εκπαιδευτεί. Αρχικά, εφαρμόζεται ένας ταξινομητής που υπολογίζει την απόσταση Mahalanobis για τα pixels, ώστε να προσδιοριστεί η αρχική κατανομή. Στη συνέχεια, ο MRF αλγόριθμος επεξεργάζεται όλα τα εικονοστοιχεία λαμβάνοντας υπόψη τις παραπάνω αποστάσεις και ενσωματώνει τη χωρική συσχέτιση μεταξύ των pixels όπως αυτή εκφράζεται μέσω του μοντέλου. Το τελικό στάδιο περιλαμβάνει την αντιστοίχιση κάθε pixel στην κλάση με την οποία έχει τη μικρότερη απόσταση. Η επιρροή των γειτονικών pixels υπολογίζεται ως ένα σταθμισμένο άθροισμα της κατανομής των κλάσεων σε μια τρισδιάστατη περιοχή γύρω από το pixel με την παραδοχή ότι ένα εικονοστοιχείο είναι πιθανότερο να ανήκει σε μια συγκεκριμένη κλάση αν η πλειοψηφία των γειτόνων του ανήκει επίσης σε αυτή.

Οι μοναδικές παράμετροι που τροποποιούνται ανά σύνολο δεδομένων, είναι ο αριθμός των κατηγοριών και οι αντίστοιχες μέσες τιμές τους. Όπως και σε κάθε άλλη εφαρμογή, το pipeline ξεκινά με τον reader. Στη συνέχεια, τα δεδομένα προωθούνται στο φίλτρο ComposeImageFilter καθώς ο αλγόριθμος MRF απαιτεί είσοδο με διανυσματική μορφή. Ωστόσο, τα pixels αντιστοιχούν σε μονοδιάστατες τιμές αφού περιγράφονται αποκλειστικά από τη φωτεινότητά τους. Το συγκεκριμένο φίλτρο αναλαμβάνει να τα μετατρέψει σε διανύσματα ώστε να είναι συμβατά με τον αλγόριθμο.

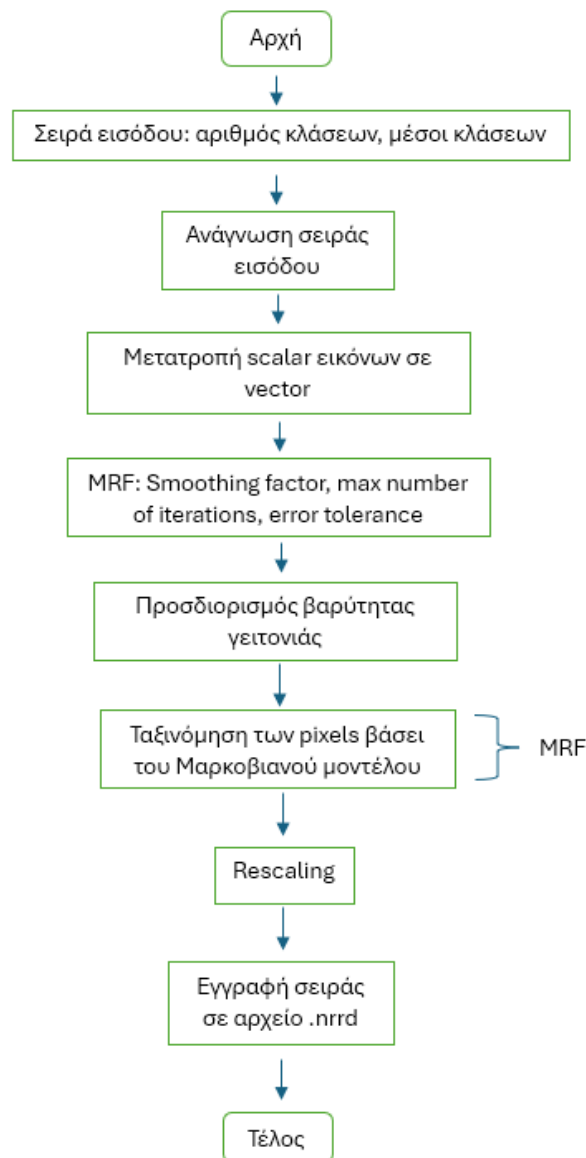
Το επόμενο στάδιο (**Εικόνα 3.18**) στην ακολουθία επεξεργασίας είναι το MRF. Σε αυτό το σημείο εισάγεται ο αριθμός των κατηγοριών όπως έχει καθοριστεί από τον χρήστη μέσω κατάλληλης συνάρτησης. Επιπλέον, ορισμένες σταθερές τιμές καθορίζονται εκ των προτέρων για κάθε σύνολο δεδομένων:

- Μέγιστος Αριθμός Επαναλήψεων: Καθορίζει το ανώτατο όριο για τη διαδικασία εκτέλεσης του φίλτρου.
- Ανοχή Σφάλματος: Χρησιμοποιείται ως κριτήριο για την επίτευξη σύγκλισης μεταξύ των περιοχών.

- Συντελεστής Εξομάλυνσης: Ρυθμίζει την ισορροπία ανάμεσα στην αρχική εικόνα και το αποτέλεσμα της τμηματοποίησης. Πολλαπλασιάζεται με τα βάρη των γειτονικών εικονοστοιχείων επηρεάζοντας τον βαθμό ομοιομορφίας των παραγόμενων περιοχών, όσο μεγαλύτερη η τιμή του, τόσο πιο ομοιόμορφο είναι το αποτέλεσμα.

Στη συνέχεια καθορίζεται ο τρόπος ορισμού της «γειτονιάς». Στην αρχή, επιλέγεται ακτίνα γειτονιάς ίση με 1. Αυτό συνεπάγεται ότι, δεδομένου πως τα εισερχόμενα δεδομένα είναι τρισδιάστατα, η κλίκα που σχηματίζεται έχει διαστάσεις 3x3x3. Στη συνέχεια, αποδίδονται τιμές βαρών στα εικονοστοιχεία που ανήκουν στην κλίκα οι οποίες συμβάλλουν στον προσδιορισμό της τελικής κατηγορίας κάθε εικονοστοιχείου.

Έπειτα από τους απαραίτητους υπολογισμούς για τον ακριβή προσδιορισμό της επίδρασης της γειτονιάς, ακολουθεί το rescaling της τελικής εικόνας με την ελάχιστη φωτεινότητα να ορίζεται στο 0 και τη μέγιστη στο 1000 [219].



Εικόνα 3.18: Διάγραμμα ροής της διαδικασίας

3.2.8 Μέθοδοι Μηχανικής & Βαθιάς Μάθησης

3.2.8.1 Εισαγωγή στη Μηχανική & Βαθιά μάθηση

Η μηχανική μάθηση [220] αποτελεί έναν σημαντικό τομέα της τεχνητής νοημοσύνης ο οποίος δίνει τη δυνατότητα στους υπολογιστές να μαθαίνουν από τα δεδομένα και να λαμβάνουν αποφάσεις ή να κάνουν προβλέψεις χωρίς να απαιτείται ρητός προγραμματισμός. Ο βασικός στόχος είναι η δημιουργία μοντέλων και αλγορίθμων που μπορούν να εντοπίζουν και να αναλύουν πρότυπα μέσα σε σύνολα δεδομένων ώστε να καταλήγουν σε χρήσιμα συμπεράσματα.

Οι κύριες μορφές μηχανικής μάθησης είναι τρεις: μάθηση με επίβλεψη, χωρίς επίβλεψη και ενισχυτική μάθηση [221]. Η μάθηση με επίβλεψη είναι η πιο ευρέως διαδεδομένη και στηρίζεται στη χρήση χαραγμένων δεδομένων στα οποία είναι γνωστή η ζητούμενη έξοδος. Το μοντέλο εκπαιδεύεται πάνω σε αυτά και στη συνέχεια, αξιολογείται με νέα μη οικεία δεδομένα. Ενδεικτικές εφαρμογές περιλαμβάνουν την κατηγοριοποίηση εικόνων, την αναγνώριση φωνής και την ανάλυση φυσικής γλώσσας [222].

Αντίθετα, στη μη επιβλεπόμενη μάθηση τα δεδομένα που χρησιμοποιούνται δεν διαθέτουν προκαθορισμένες εξόδους. Σκοπός της είναι να αποκαλύψει κρυφές σχέσεις ή ομάδες μέσα στα δεδομένα όπως στην ομαδοποίηση ή τη μείωση διαστάσεων. Χρήσεις αυτής της μεθόδου περιλαμβάνουν την ανίχνευση ανωμαλιών, την τμηματοποίηση της αγοράς και την εξαγωγή χρήσιμων χαρακτηριστικών [221].

Η ενισχυτική μάθηση διαφέρει, καθώς επικεντρώνεται στη διαδικασία λήψης αποφάσεων μέσω αλληλεπίδρασης με το περιβάλλον και βασίζεται στην έννοια της επιβράβευσης. Τέτοιες μέθοδοι χρησιμοποιούνται ευρέως σε ρομποτικά συστήματα, παιχνίδια και αυτόματα συστήματα ελέγχου όπου οι καταστάσεις μεταβάλλονται δυναμικά και οι αποφάσεις πρέπει να προσαρμόζονται [223].

Η διαδικασία μηχανικής μάθησης ξεκινά με τη συλλογή και τον καθαρισμό των δεδομένων, ένα κρίσιμο στάδιο αφού η ποιότητα των δεδομένων επηρεάζει άμεσα την ακρίβεια των προβλέψεων. Τα δεδομένα διαχωρίζονται σε σύνολα εκπαίδευσης και δοκιμής. Το πρώτο χρησιμοποιείται για την εκπαίδευση του μοντέλου, ενώ το δεύτερο για την αξιολόγηση της απόδοσής του.

Ακολουθεί η επιλογή κατάλληλου αλγορίθμου ανάλογα με τη φύση του προβλήματος και τα χαρακτηριστικά των δεδομένων. Υπάρχει πληθώρα επιλογών, όπως δέντρα αποφάσεων, νευρωνικά δίκτυα, υποστηρικτικές μηχανές διανυσμάτων και άλλα. Κατά την εκπαίδευση, το μοντέλο προσαρμόζει τις παραμέτρους του ώστε να ελαχιστοποιήσει το σφάλμα και μετά αξιολογείται για να διασφαλιστεί ότι δεν υπερπροσαρμόζεται στα δεδομένα εκπαίδευσης [224]. Μόλις το μοντέλο αποδώσει ικανοποιητικά, μπορεί να εφαρμοστεί σε νέα δεδομένα, ένα στάδιο γνωστό ως ανάπτυξη ή συμπερασματική φάση.

Η μηχανική μάθηση έχει ευρεία εφαρμογή σε πολλούς τομείς, όπως η υγεία, οι οικονομικές υπηρεσίες, οι μεταφορές και η βιομηχανία. Στην ιατρική, αξιοποιείται για την πρόβλεψη εξελίξεων υγείας, τη διάγνωση νοσημάτων και την εκτίμηση κινδύνων.

Παρά τα σημαντικά οφέλη, υπάρχουν και σημαντικές προκλήσεις που δεν πρέπει να αγνοούνται. Η προστασία της ιδιωτικότητας και η ασφάλεια των δεδομένων είναι κομβικά ζητήματα ιδιαίτερα όταν πρόκειται για προσωπικά ή ευαίσθητα δεδομένα. Οι οργανισμοί καλούνται να υιοθετήσουν αυστηρές πρακτικές διαχείρισης για την αποτροπή παραβιάσεων και διαρροών [225].

Επιπλέον, τίθενται ηθικά ερωτήματα, ειδικά όταν τα μοντέλα λαμβάνουν αποφάσεις που επηρεάζουν ανθρώπινες ζωές. Είναι κρίσιμο τα μοντέλα να είναι δίκαια και αμερόληπτα χωρίς να αναπαράγουν κοινωνικές προκαταλήψεις. Τέλος, για την επιτυχία της μηχανικής μάθησης απαιτείται μεγάλος όγκος δεδομένων υψηλής ποιότητας, η συλλογή και επεξεργασία των οποίων μπορεί να είναι χρονοβόρα και απαιτητική [226].

Συνοψίζοντας, η μηχανική μάθηση προσφέρει ισχυρά εργαλεία για την υποστήριξη αποφάσεων και τη βελτιστοποίηση λειτουργιών. Ωστόσο, η αποτελεσματική και υπεύθυνη αξιοποίησή της προϋποθέτει επίγνωση των περιορισμών της και την εφαρμογή κατάλληλων στρατηγικών αντιμετώπισης των σχετικών προκλήσεων.

Η βαθιά μάθηση [227] αποτελεί έναν εξειδικευμένο κλάδο της μηχανικής μάθησης, ο οποίος βασίζεται στη χρήση πολυεπίπεδων τεχνητών νευρωνικών δικτύων για την κατανόηση και αντιμετώπιση πολύπλοκων προβλημάτων. Η ιδέα της βαθιάς μάθησης προέρχεται από το πεδίο της τεχνητής νοημοσύνης και τα τελευταία χρόνια έχει γνωρίσει εκρηκτική ανάπτυξη χάρη στην ικανότητά της να επιτυγχάνει εξαιρετικά αποτελέσματα σε πολλές και διαφορετικές εφαρμογές, όπως η αναγνώριση εικόνας, η μετατροπή ομιλίας σε κείμενο και η κατανόηση φυσικής γλώσσας [228].

Τα δίκτυα βαθιάς μάθησης, ή αλλιώς βαθιά νευρωνικά δίκτυα, απαρτίζονται από πολλά διαδοχικά επίπεδα τεχνητών νευρώνων με κάθε επίπεδο να επεξεργάζεται την πληροφορία που του μεταβιβάζεται από το προηγούμενο και να παράγει δεδομένα για το επόμενο. Αυτού του τύπου η αρχιτεκτονική δίνει στο σύστημα τη δυνατότητα να εντοπίζει αυτόματα χαρακτηριστικά από ανεπεξέργαστα δεδομένα και να πραγματοποιεί σύνθετους μη γραμμικούς μετασχηματισμούς. Σε αντίθεση με τις κλασικές μεθόδους μηχανικής μάθησης, όπου απαιτείται ανθρώπινη παρέμβαση για την επιλογή κατάλληλων χαρακτηριστικών, η βαθιά μάθηση βασίζεται στην αυτόματη εξαγωγή αυτών από τα ίδια τα δεδομένα.

Η εκπαίδευση των μοντέλων βαθιάς μάθησης μπορεί να γίνει με διάφορες τεχνικές ανάλογα με την εκάστοτε ανάγκη. Η εποπτευόμενη μάθηση είναι η πιο διαδεδομένη προσέγγιση όπου το σύστημα εκπαιδεύεται με χρήση δεδομένων που φέρουν ετικέτες και επιχειρεί να μειώσει το σφάλμα μεταξύ της προβλεπόμενης και της πραγματικής απάντησης. Επιπλέον, υπάρχουν μέθοδοι χωρίς επίβλεψη στις οποίες το μοντέλο μαθαίνει από μη χαραγμένα δεδομένα, καθώς

και τεχνικές ενισχυτικής μάθησης όπου η μάθηση βασίζεται σε ανταμοιβές ή ανατροφοδότηση που λαμβάνει το σύστημα κατά την εκτέλεση ενέργειας [229].

Η βαθιά μάθηση έχει αξιοποιηθεί με εντυπωσιακά αποτελέσματα σε πληθώρα εφαρμογών. Στην ταξινόμηση εικόνων για παράδειγμα έχει επιτρέψει τον αυτόματο εντοπισμό αντικειμένων ή σκηνών με ακρίβεια που σε πολλές περιπτώσεις υπερβαίνει εκείνη του ανθρώπου.

Το πεδίο εξελίσσεται ραγδαία με νέες αρχιτεκτονικές όπως τα συνελκτικά, τα επαναλαμβανόμενα και τα μετασχηματιστικά δίκτυα να προσφέρουν αυξημένες δυνατότητες και επιδόσεις. Επιπλέον, τεχνικές όπως η μεταφορά μάθησης και η προσαρμογή σε διαφορετικά πεδία έχουν συμβάλει στην αποτελεσματικότερη εφαρμογή των μοντέλων σε νέες ή περιορισμένες περιοχές [230].

Συνολικά, η βαθιά μάθηση αποτελεί ένα ισχυρό εργαλείο στην επίλυση σύνθετων προβλημάτων και μπορεί να φέρει σημαντικές αλλαγές σε ποικίλους τομείς. Παρά τα πλεονεκτήματά της, ωστόσο υπάρχουν και προκλήσεις όπως η ανάγκη για μεγάλες ποσότητες ποιοτικών δεδομένων, η αυξημένη υπολογιστική ισχύς και η πολυπλοκότητα στην ανάπτυξη των μοντέλων. Παρ' όλα αυτά, η συμβολή της στην τεχνητή νοημοσύνη αναμένεται να παραμείνει καθοριστική με τη βαθιά μάθηση να οδηγεί σε καινοτομίες και πρωτοποριακές εφαρμογές τα επόμενα χρόνια.

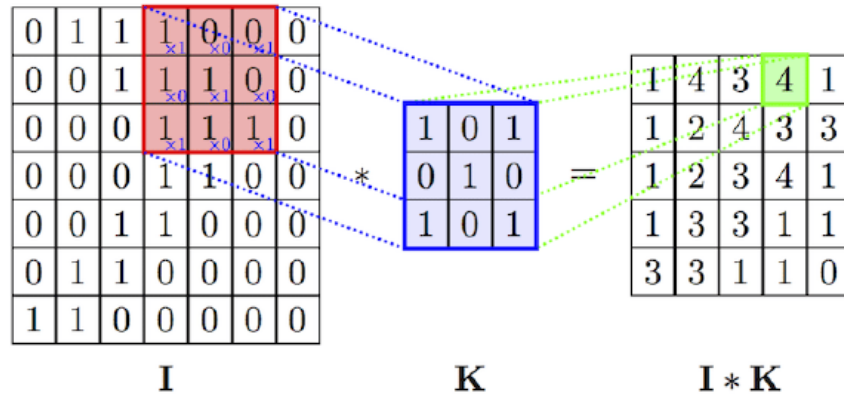
3.2.8.2 Συνελκτικά Νευρωνικά Δίκτυα (CNN)

Τα Συνελκτικά Νευρωνικά Δίκτυα (CNNs) [231] ανήκουν στην κατηγορία των βαθιών νευρωνικών δικτύων και έχουν αναπτυχθεί με σκοπό την επεξεργασία δεδομένων που εμφανίζουν δομή πλέγματος όπως είναι οι εικόνες, τα βίντεο και τα ηχητικά σήματα. Αυτού του τύπου τα δίκτυα αποτελούνται από διαδοχικά επίπεδα, καθένα από τα οποία εφαρμόζει συγκεκριμένα φίλτρα στην είσοδο και παράγει ως έξοδο έναν χάρτη χαρακτηριστικών. Κατά τη διάρκεια της διαδικασίας εκπαίδευσης, τα φίλτρα αυτά μαθαίνουν και έχουν ως στόχο την αυτόματη ανάδειξη σημαντικών χαρακτηριστικών από τα δεδομένα.

Το αρχικό επίπεδο, γνωστό ως επίπεδο εισόδου, επεξεργάζεται τα ακατέργαστα δεδομένα όπως αυτά μπορεί να είναι μια εικόνα. Στη συνέχεια, τα επόμενα επίπεδα που καλούνται και συνελκτικά επίπεδα εφαρμόζουν φίλτρα στην είσοδο ώστε να προκύψει μια νέα αναπαράσταση των δεδομένων (χάρτης χαρακτηριστικών). Αυτή η διαδικασία επιτρέπει την αυτόματη ανίχνευση κρίσιμων στοιχείων μέσω της μάθησης των φίλτρων.

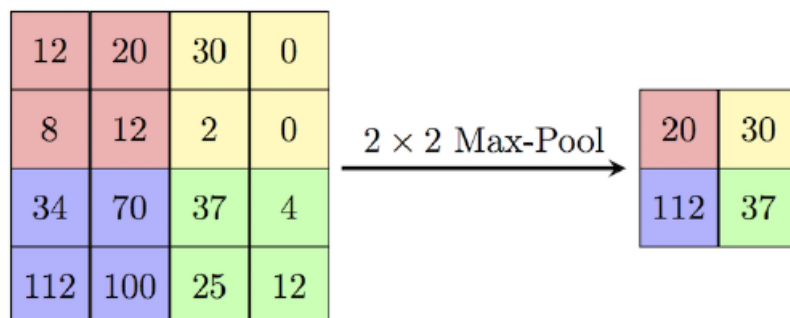
Τα φίλτρα συνήθως έχουν μικρές διαστάσεις και επικεντρώνονται σε τοπικές περιοχές των δεδομένων, δίνοντάς τους τη δυνατότητα να εντοπίζουν τοπικά πρότυπα. Η εφαρμογή των φίλτρων στα δεδομένα εισόδου πραγματοποιείται μέσω της διαδικασίας της συνέλιξης. Σε αυτή, το φίλτρο μετατοπίζεται πάνω στα δεδομένα και σε κάθε θέση υπολογίζεται το εσωτερικό

γινόμενο μεταξύ των δύο. Το αποτέλεσμα αυτής της διαδικασίας (**Εικόνα 3.19**) είναι ένας χάρτης χαρακτηριστικών ο οποίος προσφέρει μια πιο αφαιρετική και συνοπτική απεικόνιση των αρχικών δεδομένων.



Εικόνα 3.19: Δυσδιάστατη συνέλιξη

Μετά τα επίπεδα της συνελκτικής επεξεργασίας, ακολουθούν τα επίπεδα συγκέντρωσης τα οποία χρησιμοποιούνται για να περιορίσουν τη χωρική ανάλυση των χαρτών χαρακτηριστικών και να ενισχύσουν την ανθεκτικότητά τους απέναντι σε μικρές μεταβολές στα δεδομένα εισόδου. Η πιο συχνά εφαρμοζόμενη τεχνική συγκέντρωσης είναι αυτή της μέγιστης τιμής (max pooling) (**Εικόνα 3.20**) κατά την οποία επιλέγεται η μεγαλύτερη τιμή μέσα σε μια μικρή περιοχή του χάρτη χαρακτηριστικών ως η αντιπροσωπευτική έξοδος για εκείνη την περιοχή. Αυτή η διαδικασία μειώνει τον αριθμό των δειγμάτων και συμβάλλει στην ενίσχυση της σταθερότητας των χαρτών χαρακτηριστικών έναντι μικρών αλλαγών στα δεδομένα εισόδου.

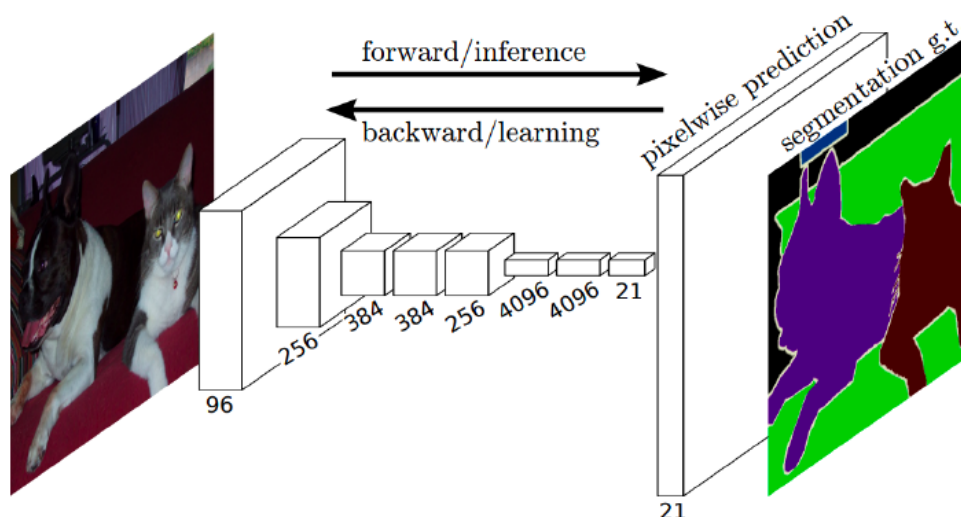


Εικόνα 3.20: Η λειτουργία του Max Pooling [232]

Αφού ολοκληρωθεί η διαδικασία συγκέντρωσης, οι χάρτες χαρακτηριστικών προωθούνται σε στρώματα πλήρους συνδεσιμότητας. Τα στρώματα αυτά συμβάλλουν στην παραγωγή των τελικών προβλέψεων του μοντέλου. Παρόμοια με τα πλήρως συνδεδεμένα στρώματα που

συναντώνται στα κλασικά νευρωνικά δίκτυα, χρησιμοποιούνται επίσης για την εκμάθηση μιας μη γραμμικής σχέσης που συσχετίζει τα χαρακτηριστικά των εισόδων με τις αντίστοιχες εξόδους.

Τα πλήρως συνελκτικά νευρωνικά δίκτυα (Fully Convolutional Networks - FCNs), **(Εικόνα 3.21)**, αποτελούν μία παραλλαγή των συμβατικών συνελκτικών νευρωνικών δικτύων (CNNs) σχεδιασμένη να διαχειρίζεται εισόδους μεταβλητού μεγέθους αντί για σταθερών διαστάσεων, όπως απαιτείται στα τυπικά CNNs. Αυτό καθίσταται δυνατό με την αντικατάσταση των πλήρως συνδεδεμένων στρωμάτων με συνελκτικά στρώματα, επιτρέποντας στο δίκτυο να διατηρεί σταθερό αριθμό παραμέτρων ανεξάρτητα από το μέγεθος της εισόδου. Χάρη σε αυτήν τη δομή, τα FCNs είναι ιδιαίτερα κατάλληλα για εφαρμογές όπως η σημασιολογική κατάτμηση εικόνων, όπου απαιτείται η εκχώρηση μιας νοηματικής ετικέτας σε κάθε εικονοστοιχείο της εικόνας.

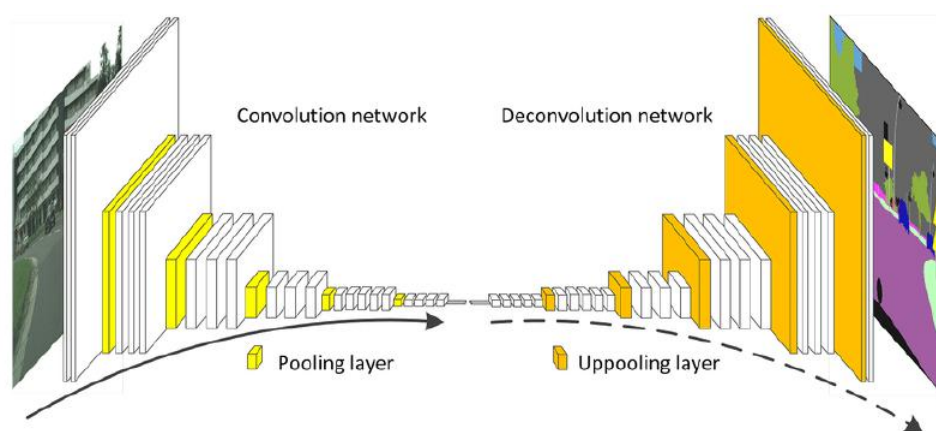


Εικόνα 3.21: FCNs αρχιτεκτονική για τη σημασιολογική κατάτμηση [154]

Στα Δίκτυα Πλήρους Συνέλιξης (Fully Convolutional Networks - FCN), τα τελικά επίπεδα ενός τυπικού Συνελκτικού Νευρωνικού Δικτύου (CNN) αντικαθίστανται από συνελκτικά στρώματα διαστάσεων 1×1 . Αυτά τα στρώματα εξυπηρετούν διπλό σκοπό: μειώνουν τον αριθμό των χαρτών χαρακτηριστικών και βελτιώνουν τη χωρική ανάλυση των δεδομένων. Η αναβάθμιση της ανάλυσης επιτυγχάνεται μέσω μιας τεχνικής που ονομάζεται μετασυντεταγμένη συνέλιξη (transposed convolution), γνωστή και ως αποσυμβολισμός (deconvolution), όπως παρουσιάζεται στην **(Εικόνα 3.22)**. Η τεχνική αυτή αξιοποιεί φίλτρα τα οποία εκπαιδεύονται κατά τη διάρκεια της μάθησης του μοντέλου.

Η έξοδος από το τελευταίο στρώμα του δικτύου εισέρχεται σε μια συνάρτηση ενεργοποίησης softmax [233], η οποία αποδίδει πιθανότητες για κάθε πιθανή κατηγορία. Για κάθε

εικονοστοιχείο η κατηγορία με τη μεγαλύτερη πιθανότητα επιλέγεται ως η τελική πρόβλεψη του συστήματος.



Εικόνα 3.22: FCNs αρχιτεκτονική για τη σημασιολογική τμηματοποίηση μαζί με τις συνελκτικές και αποσυγκεντρωτικές συνιστώσες [234]

Συνοπτικά, τόσο τα CNN όσο και τα FCN είναι προηγμένες αρχιτεκτονικές ειδικά σχεδιασμένες για την ανάλυση δεδομένων με δομή πλέγματος, όπως εικόνες, βίντεο και ακουστικά σήματα. Τα CNN βασίζονται σε διαδοχικά στρώματα, στα οποία εφαρμόζονται φίλτρα για την παραγωγή χαρτών χαρακτηριστικών από τα δεδομένα εισόδου. Τα φίλτρα αυτά εκπαιδεύονται αυτόματα και εξάγουν σημαντικές πληροφορίες. Από την άλλη, τα FCN [235] επεκτείνουν τα CNN αντικαθιστώντας τα πλήρως συνδεδεμένα στρώματα με συνελκτικά, κάτι που τους επιτρέπει να επεξεργάζονται εισόδους μεταβλητού μεγέθους χωρίς αύξηση των παραμέτρων. Οι δύο αυτές αρχιτεκτονικές έχουν καταγράψει εντυπωσιακά αποτελέσματα σε εφαρμογές όπως η κατηγοριοποίηση εικόνων, ο εντοπισμός αντικειμένων και η σημασιολογική τμηματοποίηση.

3.2.8.3 Μηχανισμοί Attention

Ο μηχανισμός attention (προσοχής), αποτελεί μια τεχνική που εφαρμόζεται σε νευρωνικά δίκτυα επιτρέποντας την εστίαση του μοντέλου σε συγκεκριμένα σημεία των δεδομένων εισόδου. Μέσω αυτής της μεθόδου, το δίκτυο μπορεί να εκτιμήσει τη σχετική σημασία κάθε στοιχείου της εισόδου και να δημιουργήσει ένα σταθμισμένο άθροισμα το οποίο χρησιμεύει ως είσοδος για τα επόμενα επίπεδα του μοντέλου. Οι μηχανισμοί attention έχουν βρει εκτεταμένη εφαρμογή σε διάφορους τομείς, όπως η επεξεργασία φυσικής γλώσσας, η αυτόματη παραγωγή περιγραφών εικόνων και η αναγνώριση φωνής [236].

Υπάρχουν πολλοί τύποι μηχανισμών attention, καθένας από τους οποίους παρουσιάζει ιδιαίτερα χαρακτηριστικά και πλεονεκτήματα. Μία από τις πιο θεμελιώδεις παραλλαγές είναι

η additive attention, ή αλλιώς Bahdanau attention [237], η οποία υπολογίζει τα βάρη προσοχής (attention weights) μεταξύ των στοιχείων της εισόδου συνδυάζοντας τις τιμές της εισόδου με ένα σύνολο παραμέτρων που μαθαίνονται κατά την εκπαίδευση. Τα βάρη αυτά προκύπτουν μέσω ενός απλού νευρωνικού δικτύου εμπρόσθιας τροφοδότησης και χρησιμοποιούνται για την παραγωγή ενός σταθμισμένου αθροίσματος της εισόδου, το οποίο αποτελεί το νέο αναπαραστατικό σήμα για το επόμενο στάδιο του μοντέλου. Ο μηχανισμός Bahdanau attention αποτελεί βασικό εργαλείο σε εφαρμογές που βασίζονται στην επεξεργασία ακολουθιών, όπως η μηχανική μετάφραση.

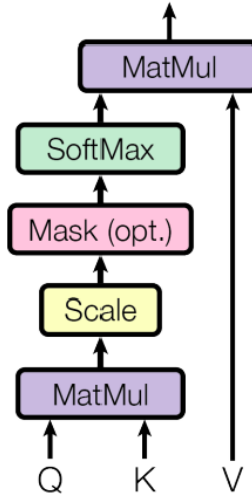
Ο μαθηματικός τύπος για τον υπολογισμό των attention weights στον μηχανισμό Bahdanau παρουσιάζεται στη συνέχεια:

$$a_t = \text{softmax}(v_a^T \tanh(W_a^T h_t + U_a S_{t-1} + b_a)) \quad (3-57)$$

$$c_t = \sum_{i=1}^T a_{t,i} h_i \quad (3-58)$$

- με h_t να είναι η t-οστή κρυφή κατάσταση του encoder (κωδικοποιητής)
- S_{t-1} να είναι η προηγούμενη κρυφή κατάσταση του decoder (αποκωδικοποιητής)
- W_a, U_a, v_a και b_a να είναι οι μαθησιακές παράμετροι του attention μηχανισμού
- Το a_t αποτελεί ένα διάνυσμα προσοχής που εκφράζει τη βαρύτητα κάθε κρυφής κατάστασης του encoder. Προκύπτει μέσω της εφαρμογής της συνάρτησης softmax σε έναν συνδυασμό ανάμεσα στην προηγούμενη κρυφή κατάσταση του decoder S_{t-1} και στις κρυφές καταστάσεις του κωδικοποιητή h_i
- Το c_t αναπαριστά το διάνυσμα πλαισίου στη χρονική στιγμή t και υπολογίζεται ως ένα σταθμισμένο άθροισμα των κρυφών καταστάσεων του encoder όπου οι σταθμίσεις καθορίζονται από τους συντελεστές προσοχής a_t

Μια διαφορετική εκδοχή του μηχανισμού προσοχής είναι η Dot-Product Attention, γνωστή και ως Scaled Dot-Product Attention (**Εικόνα 3.23**). Σε αυτή την προσέγγιση, ο υπολογισμός των βαρών προσοχής βασίζεται στο εσωτερικό γινόμενο μεταξύ των διανυσμάτων εισόδου και των αντίστοιχων βαρών. Το αποτέλεσμα του γινομένου κλιμακώνεται διαιρώντας το με την τετραγωνική ρίζα της διαστασιμότητας των διανυσμάτων εισόδου ώστε να ελέγχεται το μέγεθος των παραγόμενων τιμών.



Εικόνα 3.23: Κλιμακωμένη προσοχή μέσω εσωτερικού γινομένου [238]

Ο συγκεκριμένος τύπος προσοχής είναι πιο αποδοτικός από τον additive attention [239] από πλευράς υπολογιστικής πολυπλοκότητας και αποτελεί βασικό στοιχείο των transformer αρχιτεκτονικών [240]. Συγκεκριμένα, όταν δίνεται ένα σύνολο από ερωτήματα $Q \in R^{n \times d_k}$, κλειδιά $K \in R^{m \times d_k}$ και τιμές $V \in R^{m \times d_v}$, ο μηχανισμός προσοχής παράγει ένα σύνολο τιμών προσοχής $a(i, j)$, οι οποίες εκφράζουν τη συσχέτιση ανάμεσα σε κάθε ερώτημα $Q(i)$ και του αντίστοιχου κλειδιού $K(j)$, με τον εξής τρόπο:

$$a(i, j) = \frac{Q(i) \cdot K(j)}{\sqrt{d_k}} \quad (3-59)$$

όπου το d_k αντιπροσωπεύει τη διαστατικότητα των κλειδιών, ενώ το $\sqrt{d_k}$ λειτουργεί ως παράγοντας κλιμάκωσης που βοηθά στη συγκράτηση των τιμών από τα γινόμενα τελείας σε λογικά όρια. Ακολούθως, οι τιμές προσοχής υποβάλλονται σε κανονικοποίηση μέσω της συνάρτησης softmax [241], η οποία μετατρέπει αυτές τις τιμές σε βάρη $w(i, j)$ που το άθροισμά τους είναι ίσο με 1:

$$w(i, j) = \text{softmax}(a(i, j)) \quad (3-60)$$

Στη συνέχεια, ο υπολογισμός του σταθμισμένου αθροίσματος των τιμών V γίνεται με τη χρήση των βαρών προσοχής, σύμφωνα με τον ακόλουθο τύπο:

$$output(i) = \sum_j w(i, j) \cdot V(j) \quad (3-61)$$

Το άθροισμα υπολογίζεται για κάθε πιθανό κλειδί j . Ο κλιμακωτός μηχανισμός προσοχής που βασίζεται σε τετραγωνικό γινόμενο, μπορεί να υλοποιηθεί μέσω πράξεων γραμμικής άλγεβρας και συγκεκριμένα με πολλαπλασιασμούς πινάκων. Σε αυτή την περίπτωση, οι πίνακες Q, K και V χρησιμοποιούνται για να αναπαραστήσουν τα ερωτήματα, τα κλειδιά και τις τιμές αντίστοιχα. Οι βαθμολογίες και τα αντίστοιχα βάρη υπολογίζονται εφαρμόζοντας συναρτήσεις τετραγωνικού γινομένου και softmax πάνω σε αυτούς τους πίνακες. Συγκεκριμένα, οι βαθμολογίες προσοχής προσδιορίζονται με τον εξής τρόπο:

$$A = \frac{Q \cdot K^T}{\sqrt{d_k}} \quad (3-62)$$

όπου η πράξη μετατόπισης αναπαρίσταται με το σύμβολο T , ενώ τα βάρη προκύπτουν από την εφαρμογή της συνάρτησης softmax στις γραμμές του πίνακα A . Έπειτα, το τελικό αποτέλεσμα υπολογίζεται ως σταθμισμένο άθροισμα των τιμών μέσω ενός πολλαπλασιασμού πινάκων:

$$output = softmax(A) \cdot V \quad (3-63)$$

Η συνάρτηση softmax εφαρμόζεται σε κάθε γραμμή του πίνακα A και στη συνέχεια γίνεται πολλαπλασιασμός μεταξύ των προκύπτοντων βαρών και του πίνακα τιμών V .

Μία παραλλαγή του μηχανισμού attention είναι ο μηχανισμός multi-head attention [242]. Σε αυτή την προσέγγιση, υπολογίζονται πολλαπλά σύνολα attention βαρών για τα δεδομένα εισόδου μέσω χρήσης διαφορετικών συνόλων παραμέτρων. Ο μηχανισμός αυτός δίνει τη δυνατότητα στο μοντέλο να επικεντρώνεται σε διάφορες πτυχές των δεδομένων ταυτόχρονα, παρακολουθώντας διαφορετικά σημεία της εισόδου. Αυτό είναι ιδιαίτερα χρήσιμο σε εφαρμογές όπως η κατανόηση φυσικής γλώσσας όπου η σειρά των λέξεων έχει σημασία.

Ο multi-head attention αποτελεί βασικό στοιχείο των μοντέλων τύπου transformer καθώς χρησιμοποιείται για τον υπολογισμό του self-attention [243] μεταξύ των στοιχείων της ακολουθίας εισόδου. Δεδομένων των πινάκων Q, K και V των αντίστοιχων γραμμικών μετασχηματισμών μέσω των W_q, W_k και W_v , η διαδικασία προσοχής με H διαφορετικές κεφαλές ορίζεται ως εξής:

$$MultiHead(Q, K, V) = Concat(head_1, head_2, \dots, head_H)W_o \quad (3-64)$$

όπου το κάθε $head$ ορίζεται ως:

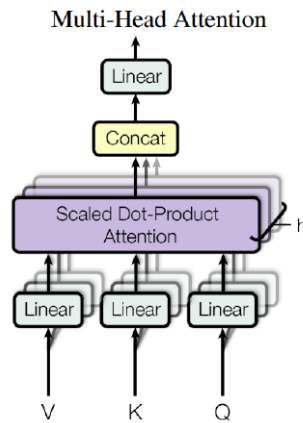
$$head_i = Attention(QW_q, i, KW_{k,i}, VW_{v,i}) \quad (3-65)$$

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (3-66)$$

η τιμή d_k αναφέρεται στη διαστατικότητα των διανυσμάτων κλειδιών και η λειτουργία softmax εφαρμόζεται κατά μήκος κάθε σειράς.

Στις προηγούμενες εξισώσεις, ο τελεστής *Concat* υποδηλώνει τη συνένωση κατά μήκος της διάστασης των χαρακτηριστικών, ενώ ο W_o αποτελεί έναν πίνακα βαρών που μαθαίνεται κατά την εκπαίδευση και εφαρμόζεται στο ενοποιημένο αποτέλεσμα. Οι πίνακες Q, K και V μπορούν να θεωρηθούν ως τα βασικά συστατικά του μηχανισμού προσοχής, τα ερωτήματα, τα κλειδιά και οι τιμές αντίστοιχα. Οι πίνακες W_q, W_k και W_v είναι μετασχηματιστικοί πίνακες βαρών οι οποίοι προβάλλουν τα Q, K και V σε έναν κοινό διανυσματικό χώρο.

Ο μηχανισμός προσοχής πολλαπλών κεφαλών (multi-head attention) (**Εικόνα 3.24**) επιτρέπει στο μοντέλο να εστιάζει ταυτόχρονα σε διαφορετικά τμήματα της εισόδου ενισχύοντας την ικανότητά του να αναγνωρίζει πολύπλοκα πρότυπα στα δεδομένα. Εκτός από τις βασικές μορφές προσοχής, υπάρχουν και πιο εξειδικευμένες παραλλαγές, όπως ο *sparse attention* [244], που περιορίζει την προσοχή σε ένα επιλεγμένο υποσύνολο στοιχείων της εισόδου καθώς και ο *axial attention* [245] ο οποίος επικεντρώνεται σε συγκεκριμένους άξονες των δεδομένων κάτι που τον καθιστά ιδιαίτερα χρήσιμο για επεξεργασία πολυδιάστατων εισόδων. Αυτές οι τεχνικές αναπτύχθηκαν ώστε να ενισχύσουν την απόδοση σε πιο απαιτητικές ή εξειδικευμένες εργασίες.



Εικόνα 3.24: Μηχανισμός προσοχής πολλαπλών κεφαλών [238]

Οι μηχανισμοί προσοχής έχουν αποδειχθεί ιδιαίτερα αποδοτικοί σε πληθώρα εφαρμογών όπως η κατανόηση και επεξεργασία φυσικής γλώσσας, η αυτόματη παραγωγή λεζάντας εικόνας, καθώς και η αναγνώριση ομιλίας. Έχουν αξιοποιηθεί για τη βελτίωση της απόδοσης διαφόρων τύπων νευρωνικών δικτύων, όπως τα επαναλαμβανόμενα (RNN) και τα συνελκτικά (CNN) δίκτυα.

Ένα πρόσθετο πλεονέκτημα των μηχανισμών προσοχής είναι η ερμηνευσιμότητα. Οι χάρτες προσοχής που προκύπτουν μπορούν να οπτικοποιηθούν ως χάρτες θερμότητας αποκαλύπτοντας τα σημεία της εισόδου που επηρεάζουν περισσότερο τις αποφάσεις του μοντέλου. Αυτή η δυνατότητα συμβάλλει στη βαθύτερη κατανόηση της συμπεριφοράς του μοντέλου και διευκολύνει την αποσφαλμάτωση και την ανάλυση [246].

Συνοψίζοντας, οι μηχανισμοί προσοχής αποτελούν έναν θεμελιώδη πυλώνα των σύγχρονων αρχιτεκτονικών τεχνητής νοημοσύνης επιτρέποντας στα μοντέλα να επικεντρώνονται επιλεκτικά σε σημαντικά μέρη των εισόδων. Λειτουργούν μέσω της εκτίμησης της σχετικής σημασίας των στοιχείων εισόδου και παράγουν σταθμισμένα αθροίσματα που προωθούνται στα επόμενα επίπεδα του δικτύου. Διακρίνονται διάφοροι τύποι, όπως οι additive, dot-product και multi-head attention, οι οποίοι έχουν ευρεία εφαρμογή και αναμένεται να διαδραματίσουν πρωταγωνιστικό ρόλο στις μελλοντικές εξελίξεις της μηχανικής μάθησης και της τεχνητής νοημοσύνης.

3.2.8.4 Transformers

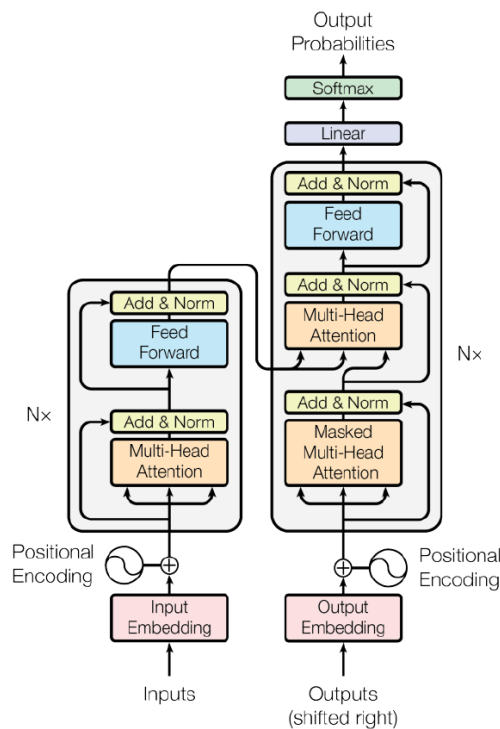
Οι Transformers (μετασχηματιστές) αποτελούν μια κατηγορία αρχιτεκτονικών νευρωνικών δικτύων ειδικά σχεδιασμένη για εφαρμογές στην επεξεργασία φυσικής γλώσσας, και εικόνας [247]. Η βασική δομή του Transformer περιλαμβάνει δύο κύρια τμήματα: έναν κωδικοποιητή (encoder) που αναλύει την είσοδο και έναν αποκωδικοποιητή (decoder) που δημιουργεί την έξοδο. Και τα δύο μέρη αποτελούνται από πολλαπλά παρόμοια στρώματα, το καθένα εκ των οποίων περιλαμβάνει δύο βασικά στοιχεία: έναν μηχανισμό πολλαπλών κεφαλών αυτό-προσοχής (multi-head self-attention) και ένα πλήρως συνδεδεμένο feed-forward νευρωνικό δίκτυο το οποίο εφαρμόζεται ξεχωριστά σε κάθε θέση της ακολουθίας [240].

Η αυτό-προσοχή (self-attention) [243] επιτρέπει στο μοντέλο να υπολογίζει δυναμικά τη σημαντικότητα κάθε στοιχείου της εισόδου σε σχέση με τα υπόλοιπα μέσω των βαρών προσοχής. Αυτά τα βάρη καθορίζονται λαμβάνοντας υπόψη τις θέσεις και τις τιμές των στοιχείων και αξιοποιούνται για τη δημιουργία μιας νέας σταθμισμένης αναπαράστασης της εισόδου.

Το feed-forward [248] δίκτυο που ακολουθεί κάθε μηχανισμό αυτό-προσοχής ενισχύει την ικανότητα του μοντέλου να συλλαμβάνει πιο πολύπλοκες αναπαραστάσεις μέσω δύο διαδοχικών γραμμικών μετασχηματισμών με συνάρτηση ενεργοποίηση ReLU [233] ανάμεσα

τους. Για να μπορέσει το μοντέλο να κατανοήσει τη σειρά των στοιχείων στην είσοδο το οποίο είναι ένα κρίσιμο στοιχείο για τις γλωσσικές εργασίες, χρησιμοποιείται η τεχνική της κωδικοποίησης θέσης (positional encoding), (**Εικόνα 3.25**). Οι αντίστοιχες πληροφορίες θέσης προστίθενται στις ενσωματώσεις των εισόδων δίνοντας στο δίκτυο τη δυνατότητα να διακρίνει τη σχετική διάταξη των λέξεων.

Ένα σημαντικό πλεονέκτημα του Transformer είναι η δυνατότητα του να επεξεργάζεται ακολουθίες μεταβλητού μήκους χάρη στον μηχανισμό αυτό-προσοχής, ο οποίος επιτρέπει στο μοντέλο να εστιάζει σε διαφορετικά σημεία της εισόδου ανάλογα με τη σημασία τους. Επίσης, επειδή η αυτό-προσοχή μπορεί να παραλληλιστεί, επιτυγχάνεται σημαντική επιτάχυνση στην εκπαίδευση και την εκτέλεση του μοντέλου. Ακόμη, το μοντέλο είναι σε θέση να παράγει εξόδους μεταβλητού μήκους μέσω του αποκωδικοποιητή ο οποίος χρησιμοποιεί τις κωδικοποιημένες πληροφορίες της εισόδου και τον μηχανισμό προσοχής για τη δημιουργία της τελικής εξόδου [249].



Εικόνα 3.25: Μοντέλο μεταφοράς με κωδικοποίηση θέσης [238]

Συνολικά, η αρχιτεκτονική των Transformers έχει αποδειχθεί εξαιρετικά αποτελεσματική σε πλήθος εργασιών επεξεργασία εικόνας. Η ευελιξία, η δυνατότητα παραλληλισμού και η αξιοποίηση της κωδικοποίησης θέσης συγκαταλέγονται στους βασικούς λόγους για την επιτυχία της. Η ευρεία υιοθέτηση και η διαρκής εξέλιξη της αναμένεται να συνεχίσουν να διαμορφώνουν τον τομέα της επεξεργασίας φυσικής γλώσσας τα επόμενα χρόνια.

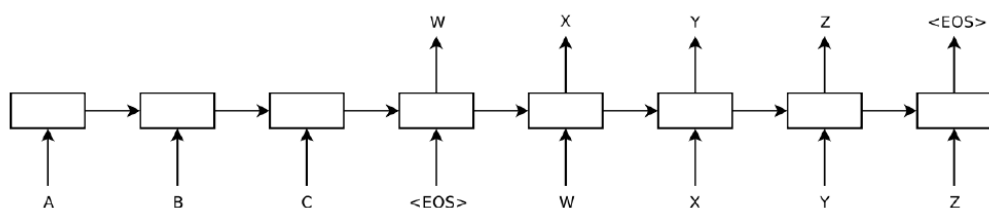
3.2.8.5 Encoder-Decoder

Η αρχιτεκτονική κωδικοποιητή-αποκωδικοποιητή (encoder-decoder), (**Εικόνα 3.26**) αποτελεί μία από τις πιο διαδεδομένες μεθόδους στη βαθιά μάθηση, ιδιαίτερα στον τομέα της υπολογιστικής όρασης. Χρησιμοποιείται ευρέως για εφαρμογές επεξεργασίας εικόνας, όπως η ταξινόμηση, ο εντοπισμός αντικειμένων και η τμηματοποίηση εικόνων. Η δομή αυτής της αρχιτεκτονικής περιλαμβάνει δύο βασικά τμήματα: το δίκτυο κωδικοποιητή το οποίο εξάγει κρίσιμα χαρακτηριστικά από την εικόνα εισόδου, και το δίκτυο αποκωδικοποιητή το οποίο ανασυνθέτει την τελική έξοδο με βάση τα εξαγόμενα χαρακτηριστικά [251].

Ο κωδικοποιητής είναι συνήθως ένα βαθύ συνελκτικό νευρωνικό δίκτυο (CNN) σχεδιασμένο για να εντοπίζει και να απομονώνει σημαντικά χαρακτηριστικά υψηλού επιπέδου από την είσοδο. Η είσοδος διέρχεται από μια σειρά συνελκτικών επιπέδων όπου κάθε επίπεδο εφαρμόζει εκπαιδευμένα φίλτρα ώστε να ανιχνεύει πρότυπα και λεπτομέρειες. Η έξοδος από κάθε επίπεδο περνάει μέσα από μια μη γραμμική συνάρτηση ενεργοποίησης, όπως η ReLU (Rectified Linear Unit) η οποία ενισχύει τη δυνατότητα του μοντέλου να συλλαμβάνει σύνθετες συσχετίσεις. Το τελικό επίπεδο του κωδικοποιητή δημιουργεί μια συνοπτική αλλά ουσιαστική αναπαράσταση της εικόνας η οποία φέρει τις απαραίτητες πληροφορίες για την επεξεργασία που ακολουθεί.

Αντίθετα, ο αποκωδικοποιητής έχει ως αποστολή να μεταφράσει αυτή τη συμπιεσμένη πληροφορία σε μια τελική έξοδο. Αποτελείται από στρώματα που εκτελούν την αντίστροφη λειτουργία των συνελκτικών, τα συχνά ονομαζόμενα αποσυμβολικά ή ανοδικής συμβολής. Μέσω αυτών των στρωμάτων, επεκτείνεται ο χάρτης χαρακτηριστικών ανακτώντας τις αρχικές διαστάσεις της εικόνας. Ένα σημαντικό στοιχείο αυτής της φάσης είναι οι συνδέσεις παράκαμψης, οι οποίες μεταφέρουν χαρακτηριστικά από τα πρώιμα στάδια του κωδικοποιητή απευθείας στον αποκωδικοποιητή. Αυτές οι συνδέσεις υλοποιούνται με συνένωση των χαρτών χαρακτηριστικών, προσφέροντας στο μοντέλο πρόσβαση σε λεπτομέρειες χαμηλού επιπέδου που ενισχύουν την ακρίβεια της εξόδου [252].

Κάθε επίπεδο του αποκωδικοποιητή περιλαμβάνει και αυτό συνάρτηση ενεργοποίησης με πιο κοινές επιλογές τη σιγμοειδή και τη softmax ανάλογα με τη φύση του προβλήματος. Στην περίπτωση της κατάτμησης εικόνας για παράδειγμα, η έξοδος είναι ένας χάρτης πιθανοτήτων σε επίπεδο εικονοστοιχείου ο οποίος υποδεικνύει σε ποια κατηγορία ή αντικείμενο ανήκει κάθε σημείο της εικόνας.



Εικόνα 3.26: Αρχιτεκτονική του μοντέλου κωδικοποιητή-αποκωδικοποιητή [250]

Το δίκτυο του κωδικοποιητή παρουσιάζεται στην αριστερή πλευρά της (**Εικόνα 3.26**) και αποτελείται από ένα επαναλαμβανόμενο νευρωνικό δίκτυο (RNN) [253] με πολλά επίπεδα. Η είσοδος δίνεται στον κωδικοποιητή διαδοχικά, ένα σύμβολο τη φορά και κάθε σύμβολο μετατρέπεται σε διάνυσμα πραγματικών αριθμών. Σε κάθε χρονική στιγμή, το RNN ενημερώνει την εσωτερική του κατάσταση με βάση το τρέχον σύμβολο της εισόδου και την προηγούμενη εσωτερική του κατάσταση. Η τελική εσωτερική κατάσταση του RNN αξιοποιείται ως μια αναπαράσταση της συνολικής ακολουθίας εισόδου με σταθερό μήκος και στη συνέχεια μεταβιβάζεται στον αποκωδικοποιητή [254].

Αντίστοιχα, το δίκτυο του αποκωδικοποιητή το οποίο φαίνεται στη δεξιά πλευρά του ίδιου σχήματος (**Εικόνα 3.26**) είναι επίσης ένα πολυεπίπεδο αναδρομικό νευρωνικό δίκτυο. Ο αποκωδικοποιητής χρησιμοποιεί την αναπαράσταση σταθερού μήκους της εισόδου που έχει παραχθεί από τον κωδικοποιητή ως αρχική εσωτερική του κατάσταση. Έπειτα, παράγει την ακολουθία εξόδου ένα σύμβολο κάθε φορά. Σε κάθε χρονικό βήμα, ενημερώνει την εσωτερική του κατάσταση με βάση το τρέχον σύμβολο εισόδου και την προηγούμενη κατάστασή του. Ακολουθώς, υπολογίζει μια κατανομή πιθανότητας για τα πιθανά σύμβολα εξόδου μέσω της συνάρτησης softmax. Το σύμβολο με τη μεγαλύτερη πιθανότητα επιλέγεται ως το επόμενο στην έξοδο και η διαδικασία επαναλαμβάνεται μέχρι να παραχθεί το ειδικό σύμβολο που σηματοδοτεί το τέλος της ακολουθίας.

Το σύστημα κωδικοποιητή-αποκωδικοποιητή εκπαιδεύεται συνολικά από την είσοδο έως την έξοδο με τη βοήθεια μιας τροποποιημένης μορφής του αλγορίθμου της οπισθοδιάδοσης που είναι γνωστή ως οπισθοδιάδοση μέσω χρόνου (BPTT) [255]. Κατά τη διαδικασία εκπαίδευσης, το μοντέλο μαθαίνει να συσχετίζει τις ακολουθίες εισόδου με τις αντίστοιχες ακολουθίες εξόδου επιδιώκοντας την ελαχιστοποίηση της συνάρτησης απώλειας η οποία βασίζεται στη διασταυρούμενη εντροπία μεταξύ των προβλεπόμενων εξόδων και των επιθυμητών αποτελεσμάτων.

Ας υποθέσουμε ότι έχουμε μια ακολουθία εισόδου μήκους T και μια αντίστοιχη ακολουθία εξόδου μήκους L . Ο ρόλος του κωδικοποιητή είναι να επεξεργαστεί την είσοδο και να παραγάγει ένα διάνυσμα αναπαράστασης σταθερού μεγέθους c το οποίο συνοψίζει τις πληροφορίες της ακολουθίας πηγής. Πιο συγκεκριμένα, για κάθε χρονική στιγμή t της εισόδου,

το αντίστοιχο διάνυσμα x_t εισάγεται σε ένα μονοστρωματικό αναδρομικό νευρωνικό δίκτυο (RNN) με κρυφό μέγεθος h . Η εσωτερική (ή κρυφή) κατάσταση του RNN στο χρόνο t συμβολίζεται με h_t :

$$h_t = f(x_t, h_{t-1}) \quad (3-67)$$

Το πλαίσιο c προκύπτει από έναν υπολογισμό που βασίζεται στις κρυφές καταστάσεις του RNN σε όλα τα χρονικά σημεία. Μία συνηθισμένη προσέγγιση για τον προσδιορισμό του c είναι η χρήση της τελικής κρυφής κατάστασης h_T του RNN κωδικοποιητή αν και υπάρχουν και εναλλακτικές μέθοδοι που μπορούν να εφαρμοστούν:

$$c = g(h_1, \dots, h_T) \quad (3-68)$$

Ο αποκωδικοποιητής αξιοποιεί το διάνυσμα περιβάλλοντος c σε συνδυασμό με την έως τώρα παραχθείσα ακολουθία στόχου, προκειμένου να δημιουργήσει το επόμενο στοιχείο της εξόδου σε κάθε χρονική στιγμή. Πιο συγκεκριμένα, για κάθε χρονική στιγμή l κατά τη δημιουργία της ακολουθίας στόχου, ο αποκωδικοποιητής δέχεται ως είσοδο την προηγούμενη έξοδο y_{l-1} , καθώς και το πλαίσιο c . Αυτές οι είσοδοι δίνονται σε ένα μονόστρωμα RNN με κρυφή διάσταση h . Στο χρονικό βήμα l , η εσωτερική κατάσταση του RNN αναπαρίσταται με το σύμβολο s_l :

$$s_l = f(y_{l-1}, c, s_{l-1}) \quad (3-69)$$

Ο αποκωδικοποιητής υπολογίζει το σύμβολο εξόδου y_l εφαρμόζοντας ένα στρώμα softmax στην κρυφή κατάσταση s_l :

$$y_l = \text{softmax}(W_{s_l} + b) \quad (3-70)$$

Οι συντελεστές του RNN κωδικοποιητή και αποκωδικοποιητή καθώς και εκείνοι του στρώματος softmax εκπαιδεύονται ταυτόχρονα με στόχο την ελαχιστοποίηση της διασταυρούμενης εντροπίας μεταξύ της προβλεπόμενης και της πραγματικής ακολουθίας στόχου.

$$\text{Encoder: } h_t = f(x_t, h_{t-1}) \quad c = g(h_1, \dots, h_T) \quad (3-71)$$

$$\text{Decoder: } s_l = f(y_{l-1}, c, s_{l-1}) \quad y_l = \text{softmax}(W_{s_l} + b) \quad (3-72)$$

Ένα βασικό πλεονέκτημα της αρχιτεκτονικής τύπου κωδικοποιητή-αποκωδικοποιητή είναι η ικανότητά της να αποτυπώνει τόσο το ευρύτερο όσο και το τοπικό συμφραζόμενο από την είσοδο της εικόνας. Συγκεκριμένα, το τμήμα του κωδικοποιητή εξάγει χαρακτηριστικά υψηλότερου επιπέδου, συλλαμβάνοντας πληροφορίες που σχετίζονται με το γενικό περιεχόμενο της εικόνας, ενώ το αποκωδικοποιητικό μέρος επικεντρώνεται στις λεπτομέρειες μέσω της αξιοποίησης συνδέσεων παράκαμψης ώστε να διατηρείται η πρόσβαση σε χαρακτηριστικά χαμηλού επιπέδου με πλούσιες λεπτομέρειες [256].

Επιπλέον, η συγκεκριμένη αρχιτεκτονική είναι ευέλικτη ως προς το μέγεθος των δεδομένων εισόδου. Δεδομένου ότι οι κωδικοποιητές και οι αποκωδικοποιητές υλοποιούνται συχνά ως πλήρως συνελκτικά νευρωνικά δίκτυα είναι σε θέση να διαχειρίζονται εικόνες ποικίλων διαστάσεων. Αυτό την καθιστά ιδανική για εφαρμογές όπως η κατάτμηση εικόνων όπου οι εισοδοί ενδέχεται να διαφέρουν σε μέγεθος.

Υπάρχουν αρκετές παραλλαγές αυτής της αρχιτεκτονικής, με πιο γνωστές τα U-Net [153], SegNet [257] και Fully Convolutional Network (FCN) [154]. Το U-Net, ιδιαίτερα, έχει ευρεία χρήση σε προβλήματα τμηματοποίησης εικόνων. Η δομή του περιλαμβάνει ένα συρρικνούμενο μονοπάτι, δηλαδή το τμήμα του κωδικοποιητή και ένα επεκτεινόμενο μονοπάτι, δηλαδή τον αποκωδικοποιητή. Στο συρρικνούμενο τμήμα πραγματοποιούνται πολλαπλές συνελκτικές διαδικασίες ακολουθούμενες από επίπεδα μέγιστης συγκέντρωσης που μειώνουν τις διαστάσεις των χαρτών χαρακτηριστικών. Στο επεκτεινόμενο μέρος χρησιμοποιούνται αποσυμπίεστικά στρώματα τα οποία συνενώνονται με τα αντίστοιχα χαρακτηριστικά από το συρρικνούμενο μονοπάτι μέσω συνδέσεων παράκαμψης. Αυτή η σύνδεση επιτρέπει στον αποκωδικοποιητή να ενσωματώνει πληροφορίες λεπτομερειών από τα πρώιμα στάδια της επεξεργασίας.

Η αρχιτεκτονική κωδικοποιητή-αποκωδικοποιητή αποτελεί ένα ισχυρό μοντέλο βαθιάς μάθησης, το οποίο έχει βρει εκτεταμένη εφαρμογή στην υπολογιστική όραση, κυρίως για εργασίες που σχετίζονται με την επεξεργασία εικόνας. Το μοντέλο αυτό περιλαμβάνει ένα τμήμα κωδικοποιητή, υπεύθυνο για την εξαγωγή σημαντικών χαρακτηριστικών από την αρχική εικόνα, και ένα τμήμα αποκωδικοποιητή, το οποίο αξιοποιεί αυτά τα χαρακτηριστικά για να παράγει την τελική έξοδο.

Η συγκεκριμένη αρχιτεκτονική διακρίνεται για την ικανότητά της να αποτυπώνει πληροφορίες τόσο σε παγκόσμιο όσο και σε τοπικό επίπεδο από τα δεδομένα εισόδου, ενώ ταυτόχρονα μπορεί να προσαρμόζεται σε εικόνες διαφόρων διαστάσεων. Αυτό την καθιστά ιδιαίτερα κατάλληλη για μια ευρεία γκάμα εφαρμογών στην επεξεργασία εικόνας [258].

Το δίκτυο κωδικοποιητή επικεντρώνεται στην απομόνωση των πιο σημαντικών χαρακτηριστικών της εικόνας, ενώ ο αποκωδικοποιητής επεξεργάζεται αυτά τα χαρακτηριστικά ώστε να δημιουργήσει το τελικό αποτέλεσμα. Η ικανότητα του μοντέλου να

χειρίζεται διαφορετικά επίπεδα πληροφορίας το καθιστά ευέλικτο και αποτελεσματικό σε πολλές περιπτώσεις.

Υπάρχουν πολλές εκδοχές της αρχιτεκτονικής αυτής, όπως τα μοντέλα U-Net, SegNet και FCN, καθεμία από τις οποίες προσφέρει διαφορετικά πλεονεκτήματα και περιορισμούς όσον αφορά την υπολογιστική απόδοση, την ακρίβεια και την αποδοτικότητα. Η επιλογή του κατάλληλου μοντέλου εξαρτάται από τις απαιτήσεις της εκάστοτε εφαρμογής καθώς κάθε παραλλαγή ισορροπεί διαφορετικά τα τεχνικά χαρακτηριστικά και τις ανάγκες του προβλήματος.

3.2.8.6 Transfer Learning

Το transfer learning αποτελεί μια ιδιαίτερα ισχυρή τεχνική στον χώρο της βαθιάς μάθησης, συμβάλλοντας ουσιαστικά στη βελτίωση της απόδοσης των μοντέλων. Η βασική της αρχή έγκειται στην αξιοποίηση προ-εκπαιδευμένων μοντέλων, τα οποία έχουν εκπαιδευτεί σε μεγάλης κλίμακας σύνολα δεδομένων. Μέσω αυτών, είναι δυνατή η άντληση γενικών και αφηρημένων χαρακτηριστικών που μπορούν να χρησιμοποιηθούν σε διαφορετικά είδη προβλημάτων [259].

Τα μοντέλα αυτά συλλαμβάνουν σύνθετες δομές και μοτίβα, γεγονός που τα καθιστά κατάλληλα για επαναχρησιμοποίηση σε πολλές εφαρμογές. Η ενσωμάτωσή τους σε νέα προβλήματα επιφέρει σημαντικά οφέλη, όπως η μείωση του απαιτούμενου όγκου δεδομένων και του κόστους επισημάνσεων που συχνά αποτελούν σημαντικό εμπόδιο στην ανάπτυξη μοντέλων μηχανικής μάθησης. Μέσω της μεταφοράς γνώσης από μια σχετική εργασία (πηγή) σε μια εργασία (στόχο), το μοντέλο μπορεί να λειτουργήσει αποτελεσματικά ακόμα και με περιορισμένα χαραγμένα δεδομένα [260].

Ένα από τα πιο καθοριστικά πλεονεκτήματα αυτής της προσέγγισης είναι η ταχύτερη εκπαίδευση και η βελτιωμένη σύγκλιση. Καθώς τα προ-εκπαιδευμένα μοντέλα έχουν ήδη μάθει βασικές αναπαραστάσεις από τεράστιους όγκους δεδομένων, όταν χρησιμοποιούνται ως βάση για νέα έργα προσφέρουν ένα προχωρημένο σημείο εκκίνησης. Αυτό επιταχύνει τη διαδικασία μάθησης και ενισχύει την αποδοτικότητα του τελικού μοντέλου.

Η μέθοδος αυτή είναι ιδιαίτερα αποτελεσματική όταν υπάρχει ομοιότητα μεταξύ της εργασίας (πηγής) και της εργασίας (στόχου) καθώς το μοντέλο μπορεί να μεταφέρει κρίσιμες πληροφορίες με μεγάλη σχετικότητα. Επιπλέον, η ικανότητα του συστήματος να προσαρμόζει τις γνώσεις του στις ιδιαιτερότητες του νέου προβλήματος, προσφέρει καλύτερη εξειδίκευση χωρίς να θυσιάζεται η αποδοτικότητα [261].

Ένα ακόμη κρίσιμο όφελος είναι η ενισχυμένη δυνατότητα γενίκευσης σε δεδομένα που δεν έχει δει το μοντέλο στο παρελθόν. Η ευρεία εκπαίδευση του αρχικού μοντέλου του επιτρέπει

να αναγνωρίζει μοτίβα και χαρακτηριστικά που ξεπερνούν συγκεκριμένες περιπτώσεις οδηγώντας σε υψηλή απόδοση ακόμα και σε άγνωστα σενάρια.

Συνολικά, το transfer learning αποτελεί πολύτιμο εργαλείο στη βαθιά μάθηση, προσφέροντας λύσεις σε σημαντικές προκλήσεις όπως η έλλειψη δεδομένων, το υψηλό κόστος επισήμανσης και η ανάγκη για ταχεία εκπαίδευση. Με την κατάλληλη αξιοποίηση της υπάρχουσας γνώσης, τα μοντέλα μπορούν να γίνουν πιο ευέλικτα, αποδοτικά και ισχυρά απέναντι σε μια πληθώρα εφαρμογών [262].

Κεφάλαιο 4: Τμηματοποίηση της Αορτής σε 3Δ Δεδομένα Αξονικού Τομογράφου Συνδυάζοντας Τεχνικές Επεξεργασίας Εικόνας και Μηχανικής Μάθησης

4.1 Εισαγωγή

Οι τεχνικές οπτικοποίησης βοηθούν τους ειδικούς της ιατρικής στην καθημερινή τους εργασία να αντιλαμβάνονται τα ιατρικά απεικονιστικά δεδομένα με πιο αποτελεσματικό τρόπο. Στην περίπτωση τρισδιάστατων (3Δ) δεδομένων, κατάλληλες τεχνικές οπτικοποίησης μπορούν να χρησιμοποιηθούν για τη μοντελοποίηση και την απεικόνιση των απαιτούμενων πληροφοριών με υψηλή ακρίβεια. Η 3Δ οπτικοποίηση προσφέρει μια πλήρη ανατομική απεικόνιση της περιοχής κλινικού ενδιαφέροντος και έτσι βοηθά τους ειδικούς της ιατρικής να εξετάζουν τα ιατρικά δεδομένα από οπτικές γωνίες που θα ήταν αδύνατο να επιτευχθούν μέσω των συνηθισμένων δισδιάστατων (2Δ) απεικονιστικών συστημάτων [263]. Ιδιαίτερα για ιατρικές παθήσεις που σχετίζονται με το καρδιαγγειακό σύστημα, η 3Δ οπτικοποίηση μπορεί να είναι ιδιαίτερα ωφέλιμη καθώς προσφέρει έναν αποτελεσματικό και μη επεμβατικό τρόπο εξέτασης των περιοχών ενδιαφέροντος που έχουν επηρεαστεί. Τέτοιες ιατρικές παθήσεις περιλαμβάνουν τη συσσώρευση πλάκας στις στεφανιαίες αρτηρίες, τις διαρθρωτικές καρδιοπάθειες και τα ανευρύσματα της αορτής [264].

Ωστόσο, πριν από την εφαρμογή οποιασδήποτε τεχνικής οπτικοποίησης, είναι απαραίτητο να οριστούν και να εξαχθούν με σαφήνεια οι ανατομικές περιοχές ενδιαφέροντος από τα υπόλοιπα δεδομένα απεικόνισης. Αυτό επιτυγχάνεται συνήθως μέσω της χρήσης μιας μεθόδου τμηματοποίησης εικόνας. Η τμηματοποίηση εικόνας αποτελεί μία από τις σημαντικότερες τεχνικές υπολογιστικής οπτικοποίησης στην ψηφιακή επεξεργασία και ανάλυση εικόνας. Γενικά, οι περισσότερες τεχνικές τμηματοποίησης βασίζονται σε διάφορα χαρακτηριστικά των επεξεργαζόμενων δεδομένων εικόνας, όπως το χρώμα ή η υφή και χρησιμοποιούνται ευρέως στην αναγνώριση εικόνας και τον εντοπισμό αντικειμένων [265], [266]. Ο σκοπός της τμηματοποίησης εικόνας είναι η διαίρεση μιας ψηφιακής εικόνας σε συνεχείς και ομοιογενείς υποομάδες οι οποίες ονομάζονται τμήματα εικόνας [267]. Τεχνικά, η τμηματοποίηση συμβάλλει στη μείωση της πολυπλοκότητας της εικόνας και επιτρέπει την περαιτέρω επεξεργασία ή ανάλυση κάθε τμήματος [268], κάτι που είναι κρίσιμο για πιο πολύπλοκες τεχνικές επεξεργασίας ή αξιολόγησης δεδομένων [269].

Στην περίπτωση μας, ο στόχος της τμηματοποίησης είναι η αορτή. Η αορτή είναι το σημαντικότερο αρτηριακό τμήμα της συστηματικής κυκλοφορίας του αίματος το οποίο συνδέεται άμεσα με την καρδιά και είναι υπεύθυνο για τη μεταφορά του αίματος. Η αορτή είναι η μεγαλύτερη αρτηρία του ανθρώπινου σώματος και μπορεί να ταξινομηθεί βάσει της

ανατομικής της θέσης, σε θωρακική και κοιλιακή αορτή με διαμέτρους 2.5 cm και 1.82 cm αντίστοιχα. Τα διάφορα μέρη της αορτής μπορούν επίσης να ταξινομηθούν ως ανιούσα αορτή, αορτικό τόξο και κατιούσα αορτή ανάλογα με τη μορφολογία της και την κατεύθυνση της ροής του αίματος [56], [270], [271].

Οι παθήσεις και οι νόσοι της αορτής όπως τα ανευρύσματα αορτής, το σύνδρομο του αορτικού τόξου, η αορτική ανεπάρκεια, το ενδοτοιχωματικό αιμάτωμα της αορτής, η αποφρακτική νόσος της αορτής, το διατιτραίνον έλκος της αορτής και η νόσος των αορτικών βαλβίδων, μπορούν να διαγνωστούν μέσω υποκλινικής έρευνας. Το οξύ αορτικό σύνδρομο, αποτελεί βασικά το αρχικό σημάδι της αορτικής νόσου και είναι μια επείγουσα κατάσταση που απαιτεί άμεση αντιμετώπιση [272]. Ως εκ τούτου, η τμηματοποίηση της αορτής είναι κρίσιμη για την αξιολόγηση των προαναφερθεισών καρδιαγγειακών παθήσεων, καθώς επιτρέπει την εφαρμογή αρκετών πιο προηγμένων επακόλουθων τεχνικών αξιολόγησης βασισμένων στην τμηματοποιημένη κλινική περιοχή ενδιαφέροντος.

Όπως προαναφέρθηκε και στην ενότητα 3.2, οι πιο βασικές προσεγγίσεις των τεχνικών τμηματοποίησης είναι να διαχωρίζουν την ψηφιακή εικόνα σε συνεχείς υποομάδες (τμήματα) με βάση την ένταση ή τη φωτεινότητα των εικονοστοιχείων (pixels) της επεξεργαζόμενης εικόνας [268]. Συγκεκριμένα, η τμηματοποίηση βάσει κατωφλίου (threshold-based) αποτελεί την πιο βασική και ευρέως χρησιμοποιούμενη τεχνική στην τμηματοποίηση εικόνας για τη διάκριση πληροφορίας προσκηνίου και παρασκηνίου. Το κύριο πλεονέκτημα της μεθόδου κατωφλίου είναι η ταχύτητα επεξεργασίας καθώς οι υπολογισμοί που απαιτούνται είναι απλοί και δεν απαιτούν προηγούμενη γνώση σχετικά με τη μορφολογία της ιατρικής εικόνας. Δεδομένου ότι οι τεχνικές κατωφλίου βασίζονται αποκλειστικά στην ένταση της εικόνας, είναι ζωτικής σημασίας οι περιοχές που διαχωρίζονται να παρουσιάζουν επαρκή αντίθεση. Εικόνες με χαμηλή αντίθεση χρειάζονται προ-επεξεργασία για την ενίσχυση της αντίθεσής τους πριν εφαρμοστεί η τμηματοποίηση βάσει κατωφλίου [273]. Η μέθοδος του Otsu είναι ο πιο διαδεδομένος αλγόριθμος τμηματοποίησης βάσει κατωφλίου και βασίζεται στην επιλογή ενός καθολικά βέλτιστου κατωφλίου μέσω της μεγιστοποίησης της διακύμανσης μεταξύ των τμημάτων της εικόνας [274].

Μια πιο εξελιγμένη προσέγγιση στην τμηματοποίηση βασισμένη στη διαίρεση της εικόνας σε προκαθορισμένα συμπλέγματα είναι ο αλγόριθμος K-Means, ο οποίος αποτελεί έναν μη επιβλεπόμενο αλγόριθμο που αποσκοπεί στη βελτίωση μιας αυθαίρετης ομαδοποίησης K-Means. Τα κύρια πλεονεκτήματα της συγκεκριμένης προσέγγισης είναι η απλότητα και η ταχύτητα επεξεργασίας της [275]. Με την πάροδο των χρόνων έχουν προταθεί αρκετές εναλλακτικές τεχνικές για την τμηματοποίηση της αορτής, όπως η μέθοδος flood-fill [276], οι μέθοδοι βασισμένες στο graph-cut [277], [278], [279], η grow cut [280] και η watershed [281]. Στην πράξη, οι προαναφερθείσες μέθοδοι τείνουν να παρουσιάζουν υποδεέστερες επιδόσεις σε σύγκριση ακόμα και με απλές τεχνικές τμηματοποίησης εικόνας ιδιαίτερα σε ανατομικές

περιοχές όπου η αορτή εφάπτεται με κάποιο όργανο παρόμοιας φωτεινότητας. Πέρα από τις βασικές και γενικές μεθόδους για την τμηματοποίηση της αορτής, η απόδοση των τεχνικών τμηματοποίησης εικόνας που βασίζονται στη βαθιά μάθηση έχει βελτιωθεί σημαντικά με την πάροδο του χρόνου. Ως αποτέλεσμα, αρκετές από αυτές τις τεχνικές μπορούν πλέον να χρησιμοποιηθούν άμεσα στην κλινική πράξη για την επεξεργασία ιατρικών απεικονιστικών δεδομένων.

Πέρα από την ανίχνευση ολόκληρης της αορτικής δομής, έχουν προταθεί διάφορες μέθοδοι για τον εντοπισμό συγκεκριμένων αορτικών δομών ή για την αξιολόγηση μιας συγκεκριμένης περιοχής της αορτής. Οι Lareyre, F. και συν. πρότειναν μια πλήρως αυτοματοποιημένη διαδικασία για την ανίχνευση ανευρύσματος κοιλιακής αορτής χρησιμοποιώντας τμηματοποίηση εικόνας. Σε αυτή τη μελέτη, μια μέθοδος ανίχνευσης περιγράμματος βασισμένη σε κατώφλι εφαρμόστηκε επιτυχώς για την αυτόματη ανίχνευση των βασικών χαρακτηριστικών του ανευρύσματος κοιλιακής αορτής (abdominal aortic aneurysm - AAA), όπως η απόσταση από τις νεφρικές και λαγόνιες αρτηρίες και ο εντοπισμός του ανευρύσματος στην αορτή [282].

Είναι συνηθισμένη πρακτική σε πολλές κλινικές εφαρμογές που αφορούν την ανίχνευση της αορτής, να πραγματοποιείται αρχικά μια χειροκίνητη ή ημι-αυτόματη διαδικασία ανίχνευσης. Ωστόσο, τόσο η χειροκίνητη όσο και η ημι-αυτόματη προσέγγιση έχουν ορισμένους περιορισμούς που προέρχονται κυρίως από το γεγονός ότι απαιτείται ανθρώπινη παρέμβαση σε μικρότερο ή μεγαλύτερο βαθμό. Επιπλέον, η απόδοση εξαρτάται από τη ρύθμιση πολλών παραμέτρων γεγονός που επηρεάζει την εφαρμοσιμότητά τους στην κλινική πρακτική. Για να εξαλειφθούν αυτοί οι περιορισμοί προτάθηκαν τα συνελκτικά νευρωνικά δίκτυα (CNN) για την αδρή αποκοπή και τον εντοπισμό της αορτής στα απεικονιστικά δεδομένα. Οι López-Linares, K., και συν. [283] επιχείρησαν πλήρως αυτοματοποιημένη ανίχνευση και αποκοπή θρόμβου στην κοιλιακή αορτή σε μετεγχειρητικές εικόνες αξονικής τομογραφίας αγγειογραφίας (CTA) χρησιμοποιώντας βαθιά συνελκτικά νευρωνικά δίκτυα (CNN). Επιπλέον, οι Larsson, και συν. [284] πρότειναν μια πλήρως αυτόματη μέθοδο για την αποκοπή κοιλιακών οργάνων χρησιμοποιώντας βαθιά συνελκτικά νευρωνικά δίκτυα. Η συγκεκριμένη μέθοδος αρχικά λειτουργεί ως ισχυρός εντοπιστής για το στόχο-όργανο και στη συνέχεια ενεργοποιείται η χρήση ενός βαθιού συνελκτικού νευρωνικού δικτύου (CNN) για την ταξινόμηση σε επίπεδο εικονοστοιχείων. Οι Julia M. H. Noothout και συν. [285] πρότειναν μια αυτόματη μέθοδο αποκοπής για την ανίχνευση της θωρακικής αορτής σε χαμηλής δόσης αξονική τομογραφία θώρακα. Σε αυτήν την περίπτωση, χρησιμοποιήθηκε ένα διαστελλόμενο συνελκτικό νευρωνικό δίκτυο (CNN), το οποίο αναλύθηκε σε εγκάρσιο, μετωπιαίο και οβελιαίο επίπεδο. Τα αποτελέσματα από κάθε επίπεδο εικόνας συγχωνεύονται για να παρασχεθεί η τελική εκτίμηση της τμηματοποίησης. Τέλος, οι A. Fantazzini και συν. [286] πρότειναν μια 3D αυτόματη τμηματοποίηση για δεδομένα CTA αορτής, συνδυάζοντας

πολλαπλές προβολές 2Δ συνελικτικών νευρωνικών δικτύων. Η συγκεκριμένη μελέτη εφαρμόζει μια πλήρως αυτόματη διαδικασία για την προεγχειρητική αποκοπή της αορτής, αρχικά εφαρμόζοντας ένα συνελικτικό νευρωνικό δίκτυο (CNN) βασισμένο στη φωτεινότητα των CTA δεδομένων ακολουθούμενο από επιπλέον CNN για την εκτίμηση των τελικά τμηματοποιημένων περιοχών.

Εκτός από τα δεδομένα αξονικής τομογραφίας (CT), έχουν χρησιμοποιηθεί και άλλες απεικονιστικές μέθοδοι για την τμηματοποίηση της αορτής. Οι M. Müller-Eschner και συν. [287] επιχειρούν να συγκρίνουν τις ακολουθίες αγγειογραφίας μαγνητικού συντονισμού (MR) με φυσική και ενισχυμένη αντίθεση μέσω τρισδιάστατης αυτόματης τμηματοποίησης του αυλού. Η μέθοδος αυτή επέτρεψε την ποσοτική αξιολόγηση (εκτίμηση των διαφορών στις διαμέτρους και στο μήκος του αορτικού αυλού) και για τις δύο τεχνικές λήψης, δηλαδή για τη φυσική και ενισχυμένη αντίθεση αγγειογραφίας μαγνητικού συντονισμού (MR).

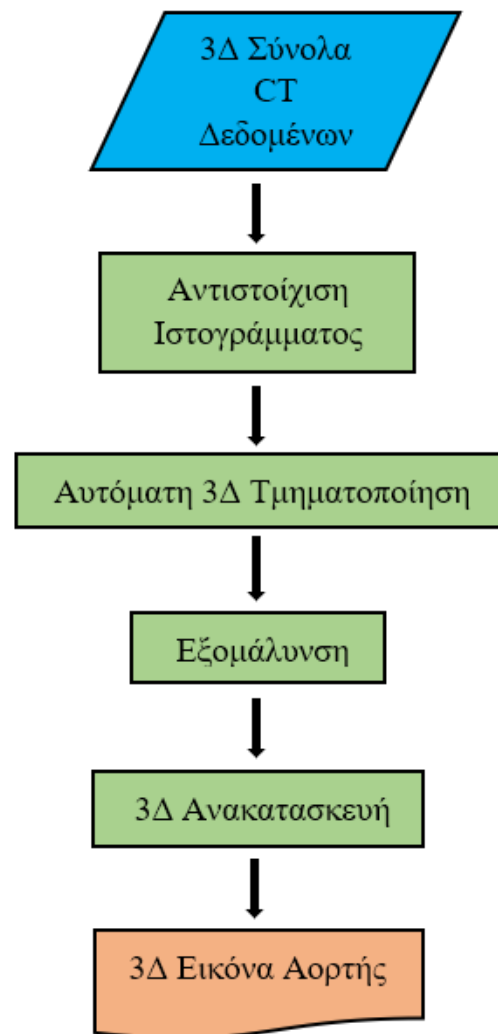
Λαμβάνοντας υπόψη όλα τα παραπάνω, στόχος της παρούσας μελέτης είναι η παροχή μιας πλήρως αυτόματης μεθόδου τρισδιάστατης τμηματοποίησης για την ανίχνευση και απεικόνιση ολόκληρης της αορτής ως τρισδιάστατο μοντέλο χρησιμοποιώντας έναν συνδυασμό απλών τεχνικών επεξεργασίας εικόνας και πιο προηγμένων μεθόδων μηχανικής μάθησης. Συγκεκριμένα, μια διαδικασία τμηματοποίησης βασισμένη σε κατώφλι (threshold-based) [273], [288] και ένα δίκτυο Μαρκοβιανού Τυχαίου Πεδίου (Markov Random Field - MRF) συνδυάζονται ώστε να παραχθεί μια πλήρης διαδικασία τμηματοποίησης δύο σταδίων, η οποία εφαρμόστηκε σε τρισδιάστατα σύνολα δεδομένων αξονικής τομογραφίας (CT).

Το υπόλοιπο της μελέτης οργανώνεται ως εξής: στην ενότητα «Μεθοδολογία», παρουσιάζεται η προτεινόμενη μεθοδολογία τμηματοποίησης, καθώς και οι τεχνικές τμηματοποίησης με τις οποίες συγκρίνεται. Στην ενότητα «Αποτελέσματα» παρουσιάζονται τα συγκριτικά αποτελέσματα, στην ενότητα «Συζήτηση» εξετάζονται βασικές πτυχές της προτεινόμενης μεθοδολογίας και διερευνάται η ενδεχόμενη χρήση της σε ποικίλες ιατρικές εφαρμογές και τέλος στην ενότητα «Συμπεράσματα» συνοψίζονται τα κύρια ευρήματα της μελέτης, υπογραμμίζονται τα πλεονεκτήματα της προτεινόμενης μεθοδολογίας και προτείνονται κατευθύνσεις για μελλοντική έρευνα.

4.2 Μεθοδολογία

Η παρούσα μελέτη ασχολείται με τον σχεδιασμό και την υλοποίηση ενός πλήρως αυτόματου αλγορίθμου 3Δ τμηματοποίησης για την εξαγωγή, μοντελοποίηση και οπτικοποίηση της αορτής από εικόνες αξονικής τομογραφίας (CT). Πιο συγκεκριμένα, η προτεινόμενη μεθοδολογία αποτελείται από τα εξής διακριτά στάδια: (α) προ-επεξεργασία των απεικονιστικών δεδομένων με διόρθωση της φωτεινότητας και της αντίθεσης μέσω εφαρμογής φίλτρου αντιστοίχισης ιστογράμματος, (β) αυτόματη 3Δ τμηματοποίηση των εικόνων μέσω

μιας διαδικασίας τμηματοποίησης δύο σταδίων, (γ) εξομάλυνση των τμηματοποιημένων δεδομένων με χρήση διακριτού Γκαούσιαν φίλτρου (Discrete Gaussian Filter) [289] και (δ) 3D ανακατασκευή και οπτικοποίηση του προκύπτοντος τρισδιάστατου μοντέλου της αορτής. Μια επισκόπηση της προτεινόμενης μεθοδολογίας παρουσιάζεται στην **(Εικόνα 4.1)**.



Εικόνα 4.1: Επισκόπηση της προτεινόμενης μεθοδολογίας για 3D αυτόματη τμηματοποίηση

4.2.1 Συλλογή Δεδομένων

Για τους σκοπούς αυτής της μελέτης, παραχωρήθηκαν από την Κλινική Λευκός Σταυρός, 16 3D σύνολα δεδομένων αξονικής τομογραφίας από ασθενείς με σοβαρή στένωση αορτικής βαλβίδας, οι οποίοι υποβλήθηκαν σε αξονική τομογραφία για προ-διακαθετηριακή εμφύτευση αορτικής βαλβίδας (TAVI).

Το πρωτόκολλο απεικόνισης περιλάμβανε ελικοειδή σάρωση με αναδρομική συγχρονισμένη καταγραφή με το ηλεκτροκαρδιογράφημα στην περιοχή του αορτικού δακτυλίου μετά από

χορήγηση σκιαγραφικού μέσου (Iobitridol, 350 mg/ml). Όλα τα σύνολα δεδομένων αποκτήθηκαν χρησιμοποιώντας τρισδιάστατο αξονικό τομογράφο (Phillips, Brilliance 64, Ολλανδία). Όλα τα παραπάνω δεδομένα ήταν σε μορφή Digital Imaging and Communications in Medicine (DICOM), ως σειρές εικόνων γκρι κλίμακας 16-bit, αποτελούμενες από 406 έως 921 τομές ανάλογα με τον ασθενή. Το μέγεθος όλων των εικόνων ήταν 512x512 pixels με διαστάσεις εικονοστοιχείου 0.4 x 0.4 mm. Το πάχος των τομών ήταν σταθερό.

4.2.2 Αντιστοίχιση Ιστογράμματος

Πριν από την εφαρμογή του ίδιου του αλγορίθμου τμηματοποίησης, τα ληφθέντα δεδομένα αξονικής τομογραφίας (CT) πρέπει να υποβληθούν σε προ-επεξεργασία, λόγω των εμφανών αποκλίσεων στη φωτεινότητα και την αντίθεση σε κάθε περίπτωση. Αυτό το βήμα διασφαλίζει ότι θα χρησιμοποιηθούν κοινές παράμετροι για τη διαδικασία τμηματοποίησης που ακολουθεί. Ο κύριος στόχος της προ-επεξεργασίας είναι η κανονικοποίηση των τιμών της κλίμακας του γκρι μιας εικόνας πηγής, με βάση τις τιμές της κλίμακας του γκρι μιας εικόνας αναφοράς, χρησιμοποιώντας το φίλτρο Histogram Matching Image Filter της ITK [290]. Στην περίπτωση μας, η εικόνα αναφοράς επιλέχθηκε τυχαία από τα διαθέσιμα τρισδιάστατα δεδομένα αξονικής τομογραφίας που χρησιμοποιήθηκαν στη μελέτη αυτή.

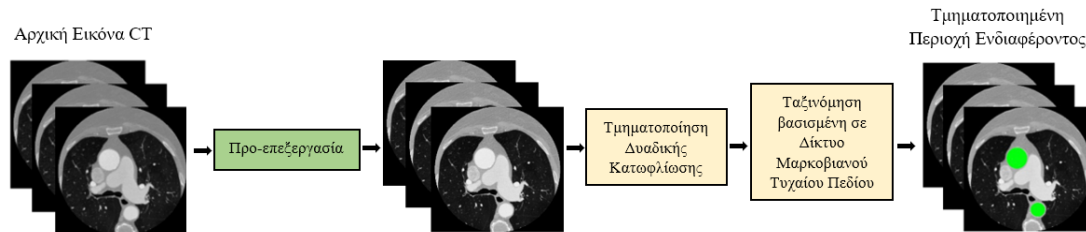
Το συγκεκριμένο φίλτρο εφαρμόζει μια τεχνική αντιστοίχισης ιστογραμμάτων, αντιστοιχίζοντας τα ιστογράμματα της εικόνας αναφοράς και της εικόνας πηγής μόνο σε έναν συγκεκριμένο αριθμό ποσοστιαίων τιμών. Η βέλτιστη λειτουργία του αλγορίθμου επιτυγχάνεται όταν τα εικονοστοιχεία του υποβάθρου αποκλείονται από τα ιστογράμματα τόσο της εικόνας αναφοράς όσο και της εικόνας πηγής.

Η υλοποίηση του φίλτρου Histogram Matching Image Filter απαιτεί τον καθορισμό μιας σειράς παραμέτρων ελέγχου, δηλαδή τον αριθμό των επιπέδων του ιστογράμματος που θα χρησιμοποιήσει η τελική εικόνα ως προς την ένταση, καθώς και τον αριθμό των ποσοστιαίων σημείων που θα αντιστοιχιστούν μεταξύ της εικόνας αναφοράς και της επεξεργασμένης εικόνας. Μετά από μια σειρά δοκιμών, επιλέχθηκαν οι τιμές 128 και 10 για τα επίπεδα ιστογράμματος και τα ποσοστιαία σημεία αντίστοιχα. Τέλος, τα ογκοστοιχεία του φόντου αποκλείστηκαν από τη διαδικασία αντιστοίχισης ιστογραμμάτων [290].

4.2.3 Σχήμα Αυτόματης 3Δ Τμηματοποίησης

Μετά την εξισορρόπηση της αντίθεσης και της συνολικής έντασης των επεξεργασμένων συνόλων δεδομένων, μπορεί να εφαρμοστεί το πραγματικό σχήμα τμηματοποίησης. Το προτεινόμενο σχήμα τμηματοποίησης το οποίο συνοψίζεται στην **(Εικόνα 4.2)**, βασίζεται σε μια διαδικασία τμηματοποίησης δύο σταδίων. Αρχικά εφαρμόζεται ένα πρώτο βήμα

τμηματοποίησης βασισμένο σε κατώφλι (threshold-based) το οποίο ακολουθείται από μια πιο προηγμένη ταξινόμηση βασισμένη σε δίκτυο Μαρκοβιανού Τυχαίου Πεδίου (Markov Random Field). Τα δύο βήματα της προτεινόμενης μεθοδολογίας τμηματοποίησης περιγράφονται με περισσότερες λεπτομέρειες στις επόμενες ενότητες. Οι βέλτιστες παράμετροι και για τα δύο βήματα της διαδικασίας τμηματοποίησης δύο σταδίων προσδιορίστηκαν μετά από αρκετές δοκιμές χρησιμοποιώντας το επιλεγμένο σύνολο δεδομένων αναφοράς.



Εικόνα 4.2: Σχήμα αυτόματης 3Δ τμηματοποίησης

4.2.4 Κατωφλίωση (Thresholding)

Η πρώτη φάση του προτεινόμενου αλγορίθμου τμηματοποίησης βασίζεται σε μια απλή προσέγγιση κατωφλίου. Γενικά, η τεχνική του κατωφλίου χρησιμοποιείται στην ιατρική απεικόνιση για τον διαχωρισμό του προσκήνιου από το φόντο. Το πλεονέκτημα αυτής της μεθόδου είναι ότι είναι πολύ γρήγορη και ως εκ τούτου χρησιμοποιείται συχνά για την προ-επεξεργασία των δεδομένων πριν από την εφαρμογή πιο σύνθετων αλγορίθμων.

Στη δική μας περίπτωση, το αποτέλεσμα αυτού του πρώτου σταδίου που βασίζεται στο κατώφλι είναι η δυαδικοποίηση των αρχικών δεδομένων απεικόνισης με το προσκήνιο να αντιστοιχεί στη γενική δομή της αορτής. Αυτό είναι ιδιαίτερα χρήσιμο για την εκτίμηση του σχήματος και της θέσης της αορτής που αποτελεί την ανατομική περιοχή ενδιαφέροντος για περαιτέρω επεξεργασία. Η συγκεκριμένη τεχνική που χρησιμοποιείται σε αυτή τη μελέτη εφαρμόζει δύο διακριτά κατώφλια: Κατώτερο και Ανώτερο. Ουσιαστικά, το προσκήνιο αντιστοιχεί στα σημεία (εικονοστοιχεία) των αρχικών δεδομένων απεικόνισης των οποίων η ένταση βρίσκεται μεταξύ των δύο αυτών τιμών.

Αυτό παρουσιάζεται και στην Εξ.(4-1), όπου, εάν η ένταση x_i του εικονοστοιχείου i βρίσκεται στο εύρος που ορίζεται από το Κατώτερο Κατώφλι L_T και το Ανώτερο Κατώφλι U_T , τότε η τιμή της εξόδου στην αντίστοιχη θέση της εικόνας ορίζεται ίση με 1 που αντιστοιχεί στο προσκήνιο. Αντίθετα, αν η ένταση βρίσκεται εκτός του διαστήματος που ορίζεται από τις δύο αυτές τιμές, θεωρείται πληροφορία φόντου και λαμβάνει την τιμή 0. Στη δική μας περίπτωση, το πρώτο στάδιο της προτεινόμενης τεχνικής τμηματοποίησης χρησιμοποίησε ένα ευρύ φάσμα κατωφλίων το οποίο προσδιορίστηκε μετά από σειρά δοκιμών [251], [276], [291].

$$Output(x_i) = \begin{cases} 1 & \text{if } L_T \leq x_i \leq U_T \\ 0, & \text{otherwise} \end{cases}$$

(4-1)

4.2.5 Μαρκοβιανό Τυχαίο Πεδίο (Markov Random Field)

Σε συνέχεια του αρχικού σταδίου τμηματοποίησης βασισμένου στην κατωφλίωση, το οποίο υπολογίζει κατά προσέγγιση τη δομή της αορτής, εφαρμόζεται ένα δεύτερο πιο λεπτομερές στάδιο τμηματοποίησης για την ακριβή εξαγωγή της αορτής. Για τον σκοπό αυτό, χρησιμοποιήθηκε η προσέγγιση των Μαρκοβιανών Τυχαίων Πεδίων (Markov Random Fields - MRF). Γενικά, τα MRF είναι πιθανοτικά μοντέλα που προσδιορίζουν την περιοχή του αντικειμένου με βάση τη συσχέτιση χαρακτηριστικών (χρώματος και υφής) μεταξύ των εικονοστοιχείων σε μια γειτονιά. Τα συγκεκριμένα μοντέλα μπορούν να χρησιμοποιηθούν για τη διαμέριση μιας εικόνας σε διακριτές περιοχές. Η τμηματοποίηση της εικόνας επιτυγχάνεται μέσω της ταξινόμησης των εικονοστοιχείων σε διαφορετικές κατηγορίες οι οποίες χαρακτηρίζονται από πολυμεταβλητές Γκαουσιανές κατανομές. Οι παράμετροι των Γκαουσιανών αυτών υπολογίζονται βάσει ενός συνόλου δεδομένων εκπαίδευσης ή εκτιμώνται από την είσοδο της εικόνας. Στην περίπτωση μας, οι Γκαουσιανές παράμετροι υπολογίζονται με βάση την εικόνα εξόδου του αρχικού σταδίου τμηματοποίησης μέσω της κατωφλίωσης.

Γενικά, το πρόβλημα της τμηματοποίησης εικόνας θεωρείται ως έργο επισήμανσης (labeling) των εικονοστοιχείων. Κατά συνέπεια, αρχικά εξάγονται τα χαρακτηριστικά από τη δυαδική εικόνα κατωφλίωσης όπου κάθε εικονοστοιχείο s της εικόνας μπορεί να αναπαρασταθεί από ένα διάνυσμα χαρακτηριστικών $\vec{f}_s$. Για ολόκληρη την εικόνα, έχουμε το σύνολο $F = \{\vec{f}_s : s \in S\}$, το οποίο αποτελεί την είσοδο του αλγορίθμου τμηματοποίησης με χρήση Μαρκοβιανού Τυχαίου Πεδίου. Στη συνέχεια, ορίζεται ένα σύνολο ετικετών όπου κάθε εικονοστοιχείο s λαμβάνει μια κλάση ετικέτας ω_s , η οποία θεωρείται διακριτή τυχαία μεταβλητή. Οι τιμές L που μπορεί να λάβει η ω_s βρίσκονται στο εύρος $L = \{1, 2, \dots, L\}$, με $L = 4$ στην περίπτωση μας, αντιστοιχώντας σε τέσσερις κλάσεις τμηματοποίησης. Αυτό μπορεί να γενικευτεί για ολόκληρη την εικόνα ως $\omega = \{\omega_s, s \in S\}$ και αποκαλείται διαδικασία ετικετοθέτησης (label process). Η επεξεργασία της εικόνας F αναπαριστά το αποτέλεσμα της βασικής διαδικασίας ετικετοθέτησης ω .

Στη συνέχεια, απαιτείται να προσδιοριστεί ένα μέτρο πιθανότητας πάνω στο σύνολο όλων των πιθανών ετικετοθετήσεων και να επιλεγεί η πιο πιθανή. Η επικαιροποιημένη πιθανότητα ενός δεδομένου συμβάντος αφού ληφθούν υπόψη νέες πληροφορίες ονομάζεται εκ των υστέρων πιθανότητα (posterior probability) $P(\omega|F)$. Η εκ των υστέρων πιθανότητα μετρά την πιθανότητα μιας ετικετοθέτησης δεδομένων των παρατηρούμενων χαρακτηριστικών F .

Ο στόχος είναι να βρεθεί η βέλτιστη ετικετοθέτηση $\hat{\omega}$ που μεγιστοποιεί την εκ των υστέρων πιθανότητα $P(\omega|F)$. Η διαδικασία αυτή ονομάζεται μέγιστη εκ των υστέρων εκτίμηση (maximum a posteriori – MAP) και ορίζεται ως εξής:

$$\hat{\omega} = \arg \max_{\omega \in \Omega} P(\omega|F) \quad (4-2)$$

όπου το Ω δηλώνει το σύνολο όλων των δυνατών επισημάνσεων.

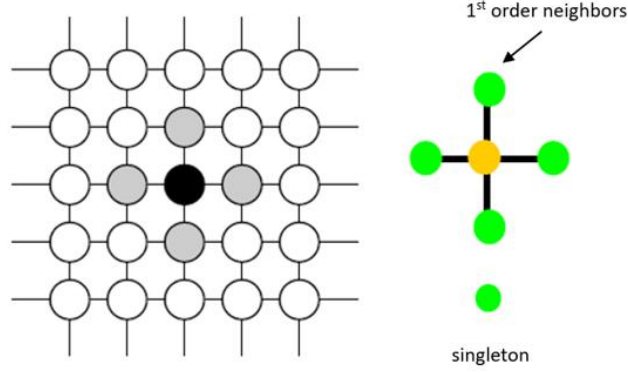
Σύμφωνα με το θεώρημα Hammersley–Clifford [292], ένα τυχαίο πεδίο είναι Μαρκοβιανό τυχαίο πεδίο αν και μόνο αν η $P(\omega)$ ακολουθεί την κατανομή Gibbs. Όπως φαίνεται στην Εξ.(4-3), η $P(\omega)$ ακολουθεί την κατανομή Gibbs, όπου η $U(\omega)$ ονομάζεται ενεργειακή συνάρτηση και η $Z = \sum_{\omega \in \Omega} \exp(-U(\omega))$ είναι η σταθερά κανονικοποίησης.

Το θεώρημα Hammersley–Clifford ορίζει τα μοντέλα Τυχαίων Πεδίων Markov μέσω των δυναμικών των κλικών. Κλίκα είναι ένα υποσύνολο $C \subseteq S$, εάν κάθε ζεύγος εικονοστοιχείων (pixels) σε αυτό το υποσύνολο είναι γείτονες. Το σύνολο των κλικών σε μία εικόνα αναπαρίσταται ως $C = C_1 \cup C_2 \cup \dots \cup C_k$. Η ανάθεση μίας τιμής $V_C(\omega)$ σε κάθε κλίκα της εικόνας, ονομάζεται δυναμικό της κλικας c , όπου το ω αντιπροσωπεύει τη διαμόρφωση του πεδίου επισημάνσεων. Όπως φαίνεται στην Εξ.(4-3), η ενέργεια $U(\omega)$ της διαμόρφωσης ω δίνεται από το άθροισμα των δυναμικών όλων των κλικών.

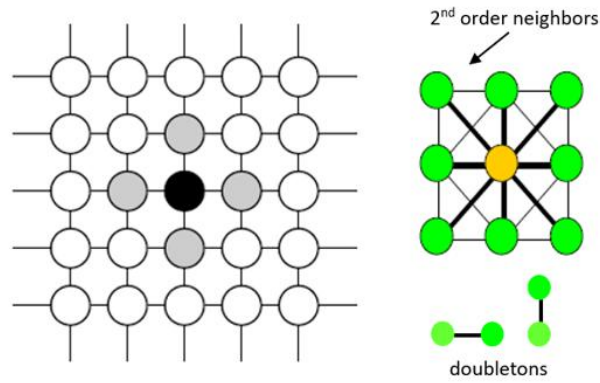
$$U(\omega) = \sum_{C \in \mathcal{C}} V_C(\omega) = \sum_{i \in C_1} V_{C_1}(\omega_i) + \sum_{(i,j) \in C_2} V_{C_2}(\omega_i, \omega_j) + \dots \quad (4-3)$$

$$P(\omega) = \frac{1}{Z} \exp(-U(\omega)) = \frac{1}{Z} \exp\left(-\sum_{C \in \mathcal{C}} V_C(\omega)\right) \quad (4-4)$$

Όπως απεικονίζεται στην **(Εικόνα 4.3)** και **(Εικόνα 4.4)**, υπάρχουν δύο συστήματα γειτνίασης με τις αντίστοιχες κλίκες.



Εικόνα 4.3: Σύστημα γειτονιάς πρώτης τάξης με τις αντίστοιχες κλίκες [217]



Εικόνα 4.4: Σύστημα γειτονιάς δεύτερης τάξης με τις αντίστοιχες κλίκες [217]

Όπως φαίνεται στην Εξ.(4-5), οι ετικέτες των εικονοστοιχείων (ή κλάσεις) αναπαρίστανται από Γκαουσιανές κατανομές.

$$P(f_s|\omega_s) = \frac{1}{\sqrt{2\pi\sigma_{\omega_s}}} \exp\left(-\frac{(f_s - \mu_{\omega_s})^2}{2\sigma_{\omega_s}^2}\right) \quad (4-5)$$

Όπως φαίνεται στην Εξ.(4-6), τα δυναμικά των μοναδιαίων κλικών είναι ανάλογα με τη πιθανότητα των χαρακτηριστικών που δίνονται. Όπως εκφράζεται στην Εξ.(4-7), τα δυναμικά των διπλών κλικών ευνοούν παρόμοιες κλάσεις σε γειτονικά εικονοστοιχεία καθώς όπως έχει οριστεί προηγουμένως η $P(\omega)$ δείχνει ότι οι τμηματοποιήσεις θα πρέπει να είναι τοπικά ομοιογενείς.

$$\omega: \log(P(f|\omega)) \quad (4-6)$$

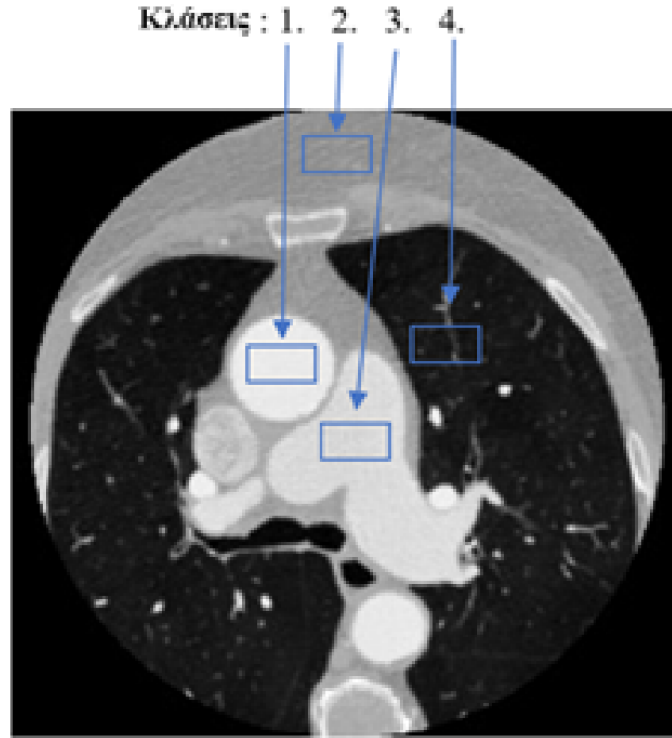
$$V_c = \beta \delta(\omega_s, \omega_r) = \begin{cases} +\beta & \text{if } \omega_s \neq \omega_r \\ -\beta & \text{otherwise} \end{cases} \quad (4-7)$$

Όπου το β είναι η παράμετρος των βαρών, της οποίας η σπουδαιότητα ελέγχεται εκ των προτέρων, με $\beta = 3$ για την δική μας περίπτωση. Καθώς η τιμή της παραμέτρου β αυξάνεται, οι προκύπτουσες περιοχές γίνονται πιο ομοιογενείς. Κάθε κλάση φωτεινότητας των εικονοστοιχείων $\lambda \in \Lambda = \{1, 2, \dots, L\}$, εκφράζεται μέσω μιας ομάδας διανυσμάτων χαρακτηριστικών $\vec{f}_s$, του διανύσματος μέσης τιμής $\vec{\mu}_\lambda$ [293], [294], [295] και του πίνακα συνδιακύμανσης σ_λ . Επιπλέον, κάθε ετικέτα $\lambda \in \Lambda$ αναπαρίσταται από μία Γκαουσιανή κατανομή $N = (\mu_\lambda, \sigma_\lambda)$, η οποία εκτιμάται με βάση την εικόνα εισόδου.

Οι στατιστικές μέσου όρου και διασποράς των τάξεων μπορούν να εκτιμηθούν μέσω του εμπειρικού μέσου και της εμπειρικής διασποράς:

$$\begin{aligned} \mu_\lambda &= \frac{1}{|S_\lambda|} \sum_{s \in S_\lambda} f_s \\ \forall \lambda \in \Lambda : \quad \sigma_\lambda^2 &= \frac{1}{|S_\lambda|} \sum_{s \in S_\lambda} (f_s - \mu_\lambda)^2 \end{aligned} \quad (4-8)$$

Όπου το S_λ δηλώνει το σύνολο των εικονοστοιχείων που ανήκουν στην κλάση εκπαίδευσης λ . Στη δική μας περίπτωση, η εκπαίδευση πραγματοποιείται με χειροκίνητη απόδοση ετικετών κλάσεων που αντιστοιχούν σε τιμές έντασης των ογκοστοιχείων. Με αυτή την έννοια, ένα σύνολο εκπαίδευσης μπορεί να οριστεί με τη χειροκίνητη επιλογή αντιπροσωπευτικών περιοχών στην εικόνα εισόδου. Η συγκεκριμένη διαδικασία εκτελέστηκε από έμπειρο καρδιολόγο όπως φαίνεται στην **(Εικόνα 4.5)**, όπου έχουν σημειωθεί οι γενικές περιοχές που αντιστοιχούν στις τέσσερις κλάσεις που χρησιμοποιούνται στη προτεινόμενη μεθοδολογία. Τα κέντρα των κλάσεων σε κάθε περίπτωση υπολογίστηκαν λαμβάνοντας υπόψη τη μέση ένταση των καθορισμένων περιοχών.



Εικόνα 4.5: Ορισμός των συνόλων εκπαίδευσης, με την επιλογή αντιπροσωπευτικών περιοχών για τις τέσσερις MRF κλάσεις

Χρησιμοποιώντας την Εξ.(4-6) και Εξ.(4-7), μπορούμε να ορίσουμε τη συνάρτηση ενέργειας του MRF μοντέλου μας:

$$U(\omega) = \sum_s \left(\log(\sqrt{2\pi}\sigma_{\omega_s}) + \frac{(f_s - \mu_{\omega_s})^2}{2\sigma_{\omega_s}^2} \right) + \sum_{s,r} \beta \delta(\omega_s, \omega_r) \quad (4-9)$$

Επομένως, η αρχική εκτίμηση MAP στην Εξ. (4-2), ισοδυναμεί με το ακόλουθο πρόβλημα ελαχιστοποίησης ενέργειας:

$$\hat{\omega} = \arg \max_{\omega \in \Omega} P(\omega|F) = \arg \min_{\omega \in \Omega} U(\omega) \quad (4-10)$$

Εφόσον το πρόβλημα μειώνεται στη ελαχιστοποίηση μιας μη-κυρτής συνάρτησης ενέργειας, απαιτούνται τεχνικές βελτιστοποίησης για να βρεθεί το παγκόσμιο ελάχιστο. Για αυτό το σκοπό, στη μελέτη μας χρησιμοποιήσαμε τη μέθοδο των Επαναληπτικές Λειτουργίες Υπό Συνθήκες (Iterated Conditional Modes - ICM) [296]. Η ICM είναι ο πιο δημοφιλής ντετερμινιστικός αλγόριθμος ο οποίος επιτρέπει την επιτάχυνση της σύγκλισης. Επιπλέον, ο

αλγόριθμος ICM χρειάζεται μια σχετικά καλή αρχική διαμόρφωση ώστε να επιτευχθεί πολύ γρήγορη σύγκλιση. Επειδή η μέθοδος ICM εκτελεί μόνο φθίνουσα πορεία προς την πλησιέστερη ενεργειακή περιοχή, η ποιότητα του τελικού αποτελέσματος εξαρτάται σε μεγάλο βαθμό από την αρχικοποίηση. Προφανώς, το ελάχιστο που βρίσκει είναι μόνο τοπικό, όμως η σύγκλιση προς αυτό το αποτέλεσμα επιτυγχάνεται συνήθως μετά από πεπερασμένο αριθμό επαναλήψεων (με μέγιστο αριθμό 50 επαναλήψεων στην περίπτωση μας).

Αρχικά, ο αλγόριθμος ICM ξεκινά με μια καλή αρχική διαμόρφωση, με ω^0 and $k = 0$. Έπειτα, υπολογίζουμε την ενέργεια $U(\eta)$ ($\eta \in N_{\omega^k}$) για κάθε διαμόρφωση που διαφέρει το πολύ σε ένα στοιχείο από την τρέχουσα διαμόρφωση ω^k (όπου αναπαρίστανται με N_{ω^k}), όπου $\omega, \eta, \dots$ είναι οι διαμορφώσεις του αλγορίθμου βελτιστοποίησής μας. Στη συνέχεια, από τις διαμορφώσεις N_{ω^k} , επιλέγουμε αυτή που έχει την ελάχιστη ενέργεια όπως φαίνεται στην Εξ.(4-11). Τέλος, η ενέργεια $U(\eta)$ υπολογίζεται επαναληπτικά για k έως ότου επιτευχθεί σύγκλιση.

$$\omega^{k+1} = \arg \min_{\eta \in N_{\omega^k}} U(\eta) \quad (4-11)$$

4.2.6 Μορφολογικές Λειτουργίες (Εξομάλυνση)

Το αποτέλεσμα του αλγορίθμου τμηματοποίησης δύο σταδίων είναι μια λεπτομερής εκτίμηση του αορτικού συστήματος. Ωστόσο, λόγω της φύσης των επεξεργασμένων δεδομένων (αξονική τομογραφία - CT), η εκτίμηση της τμηματοποίησης ενδέχεται να περιλαμβάνει χαμηλής ανάλυσης ή θορυβώδεις περιοχές οι οποίες πρέπει να εξαλειφθούν πριν τη δημιουργία του τρισδιάστατου μοντέλου της τμηματοποιημένης αορτής. Αυτό μπορεί να επιτευχθεί με την εφαρμογή ορισμένων μορφολογικών λειτουργιών (φιλτράρισμα). Επομένως, ο στόχος αυτού του σταδίου επεξεργασίας είναι η ελαχιστοποίηση της τραχύτητας της επιφάνειας μέσω μιας επαναληπτικής διαδικασίας εξομάλυνσης. Για τον σκοπό αυτό, χρησιμοποιήθηκε ένα Διακριτό Γκαουσιανό Φίλτρο Εικόνας (Discrete Gaussian Image Filter) [290]. Η Γκαουσιανή εξομάλυνση είναι μια γνωστή τεχνική για την αφαίρεση θορύβου από εικόνες ενώ παράλληλα διατηρείται η πληροφορία. Η εξομάλυνση επιτυγχάνεται με την εφαρμογή συνέλιξης ενός πυρήνα, ο οποίος ορίζεται από μια Γκαουσιανή συνάρτηση πάνω στην εικόνα [297]. Οι Γκαουσιανοί πυρήνες κατασκευάζονται σύμφωνα με τη θεωρία του Tony Lindeberg [298] ώστε οι λειτουργίες εξομάλυνσης και παραγώγων να είναι αντιμεταθετικές πριν και μετά τη διακριτοποίηση.

Αυτό το στάδιο απαιτεί μια τιμή για τη διασπορά που σχετίζεται με τον Γκαουσιανό πυρήνα. Στην περίπτωση μας, χρησιμοποιήσαμε τιμή ίση με 5, η οποία προσδιορίστηκε μετά από μια σειρά δοκιμών. Οι υπόλοιπες παράμετροι του παραπάνω φίλτρου διατηρήθηκαν στις

προεπιλεγμένες τους τιμές. Γενικά, για μικρές τιμές διασποράς, το Γκαουσιανό φίλτρο τείνει να είναι ταχύτερο, ενώ μεγαλύτερες τιμές διασποράς προκαλούν παραμόρφωση στα όρια της αορτής [297].

4.2.7 3Δ Ανακατασκευή

Μετά το βήμα της εξομάλυνσης, η τμηματοποιημένη αορτή μπορεί να μοντελοποιηθεί μέσω μιας διαδικασίας τρισδιάστατης (3Δ) ανακατασκευής η οποία θα επιτρέψει την οπτικοποίηση ολόκληρης της δομής της αορτής. Η τρισδιάστατη (3Δ) ανακατασκευή αποτελεί ένα χρήσιμο εργαλείο για τους ειδικούς της ιατρικής καθώς προσφέρει μια πιο ακριβή και ταχύτερη προσέγγιση στη διάγνωση και θεραπεία επιτρέποντας την εξέταση των περιοχών κλινικού ενδιαφέροντος από διάφορες γωνίες και οπτικές κάτι που αλλιώς θα ήταν αδύνατο να επιτευχθεί μη επεμβατικά.

Γενικά, υπάρχουν δύο μέθοδοι για την 3Δ ανακατασκευή και οπτικοποίηση: η Απεικόνιση Επιφανείας (Surface Rendering) και η Απεικόνιση Όγκου (Volume Rendering). Η Απεικόνιση Επιφανείας είναι μια αποτελεσματική και γρήγορη μέθοδος για την αποτύπωση των επιφανειών των αντικειμένων αλλά περιορίζεται στην ανακατασκευή των περιγραμμάτων των απεικονιζόμενων αντικειμένων απορρίπτοντας οποιαδήποτε εσωτερική πληροφορία [299]. Αντίθετα, η Απεικόνιση Όγκου μοντελοποιεί κάθε σημείο δεδομένων των ανακατασκευασμένων αντικειμένων. Ως αποτέλεσμα, παρέχει πλήρη λεπτομέρεια αλλά γενικά είναι πολύ πιο αργή από την Απεικόνιση Επιφανείας. Δεδομένου ότι η ακριβής απεικόνιση της περιοχής ενδιαφέροντος έχει υψίστη σημασία στις ιατρικές εφαρμογές, η Απεικόνιση Όγκου χρησιμοποιείται συχνά στην ιατρική απεικόνιση [300], [301].

Το προτεινόμενο σχήμα ανακατασκευής και μοντελοποίησης βασίζεται σε μια πολύ δημοφιλή προσέγγιση Απεικόνισης Όγκου, γνωστή ως Ray-casting. Στο Ray-casting, η ανακατασκευή επιτυγχάνεται μέσω της ακόλουθης διαδικασίας: Αρχικά, τα 3Δ δεδομένα σαρώνονται και σε κάθε σημείο δεδομένων (εικονοστοιχείο) αποδίδονται τιμές χρώματος και αδιαφάνειας. Στη συνέχεια, σχεδιάζεται μια εικονική ακτίνα η οποία ξεκινά από κάθε σημείο της δισδιάστατης (2Δ) συσκευής απεικόνισης (οθόνη) και διέρχεται μέσα από τον χώρο των 3Δ δεδομένων σύμφωνα με την κατεύθυνση της γραμμής θέασης σε κάθε περίπτωση. Η εικονική ακτίνα δειγματοληπτείται σε ίσα διαστήματα, παράγοντας k σημεία δειγματοληψίας. Για κάθε σημείο δειγματοληψίας υπολογίζονται οι τιμές χρώματος και αδιαφάνειας με χρήση φίλτρου τριγωνικής παρεμβολής το οποίο λαμβάνει υπόψη τις κορυφές των οκτώ γειτονικών σημείων του κάθε δείγματος. Στη συνέχεια, οι τιμές χρώματος και αδιαφάνειας για όλα τα σημεία της 2Δ συσκευής απεικόνισης (οθόνης) υπολογίζονται με συσσώρευση των τιμών των αντίστοιχων σημείων δειγματοληψίας με σειρά από μπροστά προς τα πίσω. Τέλος, απεικονίζονται το χρώμα και η αδιαφάνεια κάθε σημείου της συσκευής απεικόνισης [302]. Η τελική 3Δ ανακατασκευή

της περιοχής της αορτής παρουσιάζεται στην **(Εικόνα 4.7)**. Η υλοποίηση του προαναφερθέντος σχήματος ανακατασκευής και οπτικοποίησης βασίστηκε στο Visualization Toolkit (VTK) [303].

4.2.8 Συγκρινόμενες Τεχνικές Τμηματοποίησης

Το τμηματοποιημένο κομμάτι της προτεινόμενης μεθοδολογίας, αξιολογήθηκε σε σύγκριση με τρεις άλλες τεχνικές τμηματοποίησης ως προς την ακρίβεια της εκτιμώμενης περιοχής της αορτής. Όπως φαίνεται στην **(Εικόνα 4.6)**, όλα τα διαθέσιμα σύνολα δεδομένων είχαν υποστεί διόρθωση έντασης και αντίθεσης μέσω της προαναφερθείσας μεθόδου εξομείωσης ιστογράμματος. Η προτεινόμενη τεχνική τμηματοποίησης δύο σταδίων καθώς και οι τρεις συγκρινόμενες μέθοδοι, η προσέγγιση βασισμένη στην κατωφλίωση, η μέθοδος K-Means και ο αρχικός αλγόριθμος Μαρκοβιανών Τυχαίων Πεδίων (Markov Random Fields) εφαρμόστηκαν στα διορθωμένα σύνολα δεδομένων.

Η τεχνική βασισμένη στην κατωφλίωση και το μοντέλο Μαρκοβιανού Τυχαίου Πεδίου αποτελούν συστατικά της προτεινόμενης μεθοδολογίας και περιγράφηκαν με κάποια λεπτομέρεια στις προηγούμενες ενότητες. Από την άλλη πλευρά, η μέθοδος τμηματοποίησης K-Means είναι μια αυτόνομη τεχνική ομαδοποίησης (clustering), η οποία βρίσκει πολλές εφαρμογές στην ψηφιακή επεξεργασία εικόνας και ιδιαίτερα στην τμηματοποίηση εικόνων. Σε γενικές γραμμές η ομαδοποίηση είναι μια μέθοδος διαχωρισμού ενός συνόλου δεδομένων σε έναν συγκεκριμένο αριθμό ομάδων έτσι ώστε τα σημεία κάθε ομάδας να τείνουν να είναι κοντά μεταξύ τους. Ο αλγόριθμος K-Means είναι μια μη επιβλεπόμενη τεχνική καθώς η κατάτμηση των δεδομένων επιτυγχάνεται χωρίς τη χρήση εξωτερικής ταξινόμησης. Στην περίπτωση μας, η ιατρική εικόνα θα τμηματοποιηθεί σύμφωνα με τις τιμές του επιπέδου του γκρι (gray level values) οι οποίες χρησιμοποιούνται για τον καθορισμό των κεντρικών σημείων των ομάδων. Για να εντοπιστεί η Περιοχή Ενδιαφέροντος (Region of Interest - ROI) και τα όρια της στην εικόνα, πραγματοποιείται τμηματοποίηση της περιοχής με βάση το κέντρο κάθε ομάδας. Αυτό συνεπάγεται ότι απαιτείται εκ των προτέρων η γνώση του αριθμού των ομάδων για τον αλγόριθμο K-Means.

Έστω $D = \{d_i | i = 1, \dots, n\}$ ένα σύνολο δεδομένων με προκαθορισμένους K-ομάδες (clusters), και $C = \{c_k | k = 1, \dots, K\}$ ένα σύνολο από K κέντρα ομάδων και $S_j = \{d | d \in c_k\}$ είναι το σύνολο των δειγμάτων που ανήκουν στην k-οστή ομάδα. Τότε, ο αλγόριθμος ομαδοποίησης K-Means ελαχιστοποιεί τη συνάρτηση κόστους που παρουσιάζεται στην Εξ.(4-12), όπου $dist(d_i, c_k)$ είναι η Ευκλείδεια απόσταση μεταξύ ενός προτύπου d_i και της ομάδας του c_k . Τα κέντρα των ομάδων υπολογίζονται μέσω των παρακάτω βημάτων: (α) τα κέντρα αρχικοποιούνται στα c_k μέσω τυχαίας δειγματοληψίας, (β) η ανάθεση των προτύπων σε μία από τις K-ομάδες βασίζεται στην ελαχιστοποίηση της απόστασης από το κεντρικό σημείο της

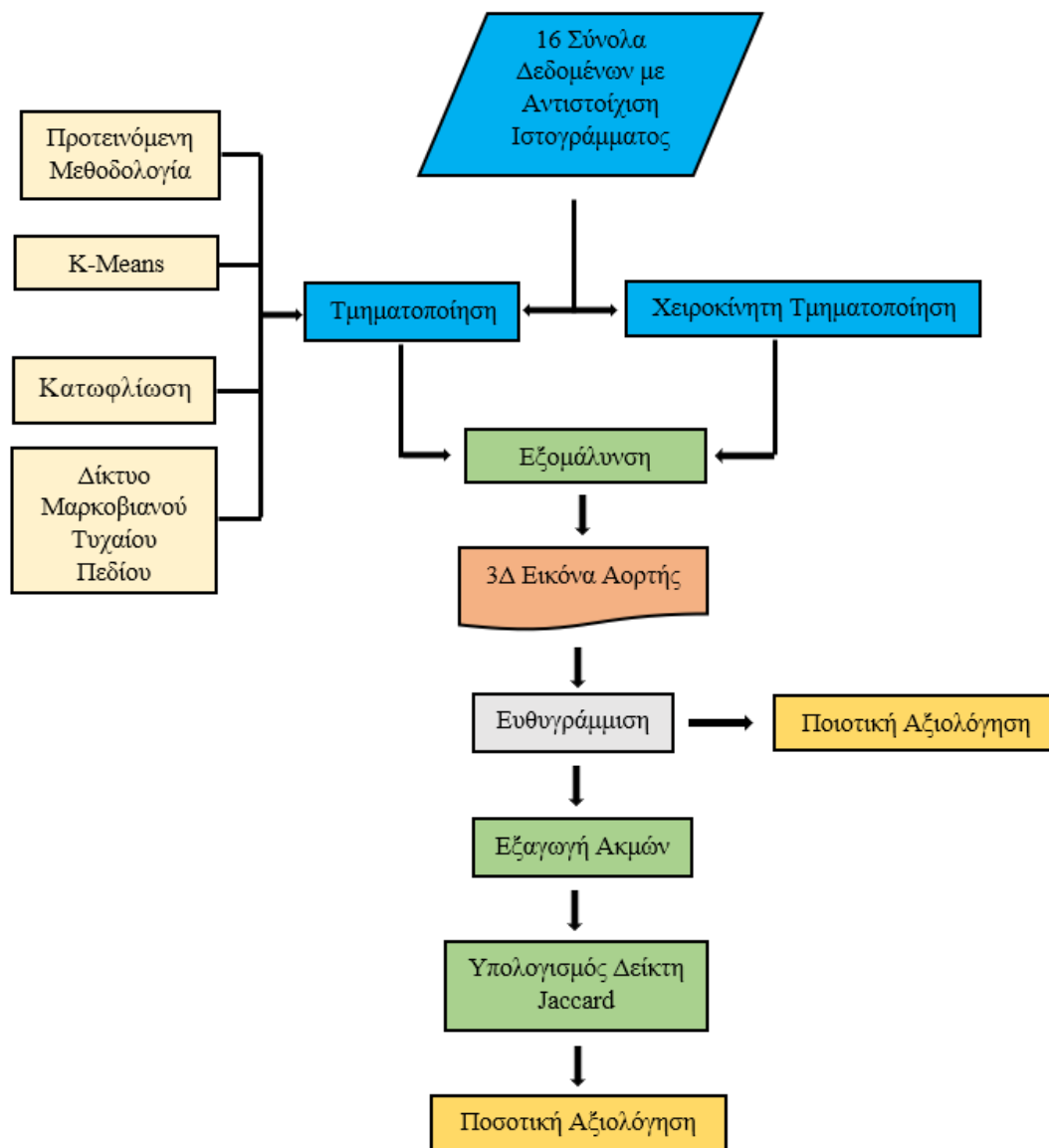
κάθε ομάδας, (γ) μια νέα εκτίμηση για το c_k παρέχεται μέσω της Εξ.(4-13), όπου το $|S_k|$ στην περίπτωση μας είναι ο αριθμός των εικονοστοιχείων που ταξινομήθηκαν στην k-οστή ομάδα, (δ) τα βήματα (β) και (γ) επαναλαμβάνονται μέχρις ότου η απόσταση μεταξύ των εκτιμώμενων κέντρων των ομάδων ελαχιστοποιηθεί.

$$Cost = \sum_{i=1}^n dist(d_i, c_k) \quad (4-12)$$

$$c_k = \frac{\sum_{d_i \in S_k} d_i}{|S_k|} \quad (4-13)$$

Μετά από μια σειρά δοκιμών στα διαθέσιμα σύνολα δεδομένων, διαπιστώθηκε ότι η βέλτιστη απόδοση στην τμηματοποίηση επιτεύχθηκε για τρεις ομάδες με κεντρικά σημεία έντασης τα 0, 300 και 500 [304].

Σε καθεμία από τις δύο εναπομείνουσες αλγοριθμικές μεθόδους τμηματοποίησης που συγκρίθηκαν, οι βέλτιστες παράμετροι προσδιορίστηκαν επίσης μετά από μια σειρά δοκιμών. Πιο συγκεκριμένα, η μέθοδος που βασίζεται στην κατωφλίωση, εφαρμόστηκε χρησιμοποιώντας κατώφλι 350 για το κάτω όριο και 450 για το άνω όριο. Για την αρχική τεχνική Δικτύου Τυχαίου Πεδίου Markov, χρησιμοποιήθηκαν τιμές έντασης για τα κέντρα των τάξεων ίσες με 0, 300, 350 και 400. Μετά την εξαγωγή της αορτής σε κάθε περίπτωση και για κάθε τεχνική που συγκρίθηκε, δημιουργήθηκαν τρισδιάστατα μοντέλα για την αξιολόγηση της απόδοσης όλων των αλγορίθμων. Επιπλέον, όλα τα τρισδιάστατα μοντέλα της τμηματοποιημένης αορτής περιλαμβανομένης της πραγματικής αορτής όπως αυτή ορίστηκε χειροκίνητα από ειδικούς, ευθυγραμμίστηκαν για περαιτέρω επεξεργασία και τα αποτελέσματα της ποιοτικής και ποσοτικής αξιολόγησης παρουσιάζονται στην ενότητα «Αποτελέσματα».



Εικόνα 4.6: Επισκόπηση των συγκρινόμενων τεχνικών τμηματοποίησης

4.2.9 Ευθυγράμμιση Μοντέλου

Όπως αναφέρθηκε, η προτεινόμενη μεθοδολογία αξιολογήθηκε σε σύγκριση με τρεις άλλες τεχνικές τμηματοποίησης με βάση την ομοιότητα της εκτιμώμενης δομής της αορτής με την πραγματική αορτή όπως αυτή ορίστηκε από ειδικούς. Αυτό το σχήμα αξιολόγησης απαιτεί την τέλεια ευθυγράμμιση των συγκρινόμενων δεδομένων απεικόνισης, δηλαδή των εκτιμώμενων και των πραγματικών περιοχών της αορτής.

Η πραγματική αορτή ορίστηκε χειροκίνητα από ειδικούς, χρησιμοποιώντας το εργαλείο ITK-SNAP [305], το οποίο είναι μια ανοιχτού κώδικα πολυπλατφορμική εφαρμογή για την οπτικοποίηση εικόνων. Το συγκεκριμένο εργαλείο χρησιμοποιεί διαφορετικό σύστημα συντεταγμένων από αυτό που λαμβάνεται υπόψη από τον αλγόριθμο τμηματοποίησής μας.

Συνεπώς, η ευθυγράμμιση των δύο συγκρινόμενων συνόλων δεδομένων (εκτιμώμενων και χειροκίνητα ορισμένων) είναι κρίσιμη πριν από την αξιολόγηση των αποτελεσμάτων.

Στην περίπτωση μας, η αυτόματη ευθυγράμμιση των τρισδιάστατων μοντέλων της αορτής πραγματοποιήθηκε χρησιμοποιώντας το εργαλείο Advanced Normalization Tools (ANTs) του 3D Slicer και πιο συγκεκριμένα μέσω της λειτουργίας του General Registration [306]. Η ευθυγράμμιση επιτυγχάνεται μέσω μιας επαναληπτικής διαδικασίας δοκιμών και σφαλμάτων, όπου ένα σύνολο δεδομένων εικόνας (κινούμενη εικόνα) ευθυγραμμίζεται χωρικά ως προς ένα σύνολο αναφοράς (σταθερή εικόνα) σύμφωνα με ένα επιλεγμένο μοντέλο μετασχηματισμού (παραγόμενη εικόνα), μεθόδου βελτιστοποίησης και συνάρτηση κόστους. Μετά από μια σειρά δοκιμών, το 3D Slicer πέτυχε βέλτιστα αποτελέσματα ευθυγράμμισης χρησιμοποιώντας τον μετασχηματισμό Rigid, ο οποίος μπορεί να αναλυθεί σε περιστροφή και μετατόπιση. Πιο συγκεκριμένα, η ευθυγράμμιση επιτεύχθηκε με βάση τον μετασχηματισμό των εικόνων ως προς τα γεωμετρικά τους κέντρα και με τη χρήση των προκαθορισμένων παραμέτρων (τρισδιάστατη απεικόνιση, γραμμική παρεμβολή εξόδου και περιορισμός εντάσεων εικόνας [0.005, 0.995]).

4.2.10 Υπολογισμός του Δείκτη Jaccard

Αφού ευθυγραμμίστηκαν τα συγκρινόμενα σύνολα απεικονιστικών δεδομένων, αξιολογήθηκαν οι διαφορές τους χρησιμοποιώντας τον δείκτη Jaccard ως μέτρο ποσοτικοποίησης της ομοιότητας των συγκρινόμενων συνόλων. Δεδομένου ότι πρέπει να συγκριθούν επιφάνειες, ο δείκτης Jaccard εφαρμόστηκε στις εξωτερικές επιφάνειες των συγκρινόμενων μοντέλων. Συνεπώς, ήταν απαραίτητο να υπολογιστούν και να εξαχθούν οι ακμές των εκτιμώμενων μοντέλων και των πραγματικών περιοχών της αορτής σε κάθε περίπτωση. Για τον σκοπό αυτό χρησιμοποιήθηκε το ParaView, μια μηχανή ανοιχτού κώδικα για οπτικοποίηση [307].

Αρχικά, οι ακμές για κάθε σύνολο δεδομένων εξήχθησαν χρησιμοποιώντας το φίλτρο Extract Edges, ακολουθούμενο από μια διαδικασία επιλογής σημείων επιφάνειας. Το φίλτρο Extract Edges δημιουργεί μια προσχεδιασμένη εκδοχή του αρχικού συνόλου δεδομένων, εξαγοντας όλες τις ακμές των κελιών των δεδομένων του ως γραμμές. Τα επιλεγμένα σημεία της επιφάνειας χρησιμοποιήθηκαν ως είσοδος για τον υπολογισμό του δείκτη Jaccard, ο οποίος πραγματοποιήθηκε στο περιβάλλον της MATLAB®.

Αρχικά, ο δείκτης Jaccard $J(W_1, W_2)$ μεταξύ των δύο συγκρινόμενων συνόλων σε τρισδιάστατο καρτεσιανό σύστημα υπολογίστηκε με βάση την Εξ.(4-14), όπου W_1 είναι ο αριθμός των σημείων της πραγματικής αορτής (χειροκίνητη τμηματοποίηση από τους ειδικούς) και W_2 είναι ο αριθμός των σημείων που προέκυψαν από την εφαρμογή των μεθόδων K-Means, Κατωφλίωσης και MRF σε καθένα από τα 16 σύνολα δεδομένων. Επιπλέον, τα $W_1 \cap W_2$ και

$W_1 \cup W_2$, αντιπροσωπεύουν τα σημεία που ανήκουν και στα δύο σύνολα και τα σημεία που ανήκουν σε τουλάχιστον ένα από τα δύο σύνολα αντίστοιχα.

$$J(W_1, W_2) = \frac{|W_1 \cap W_2|}{|W_1 \cup W_2|} \quad (4-14)$$

Ο δείκτης Jaccard μπορεί να λάβει τιμές μεταξύ 0 και 1, όπου το 0 αντιστοιχεί σε καθόλου επικάλυψη και το 1 σε πλήρη επικάλυψη μεταξύ των συγκρινόμενων μοντέλων [308].

4.2.11 Περιβάλλον υλοποίησης και δοκιμών

Όλοι οι αλγόριθμοι τμηματοποίησης, η αντιστοίχιση ιστογραμμάτων, τα μορφολογικά στάδια και η τρισδιάστατη απεικόνιση του μοντέλου της αορτής, υλοποιήθηκαν ως προσαρμοσμένες εφαρμογές λογισμικού σε γλώσσα προγραμματισμού C++. Όπως αναφέρθηκε, η διαδικασία ευθυγράμμισης πραγματοποιήθηκε χρησιμοποιώντας το 3D Slicer. Η εξαγωγή των ακμών έγινε με την χρήση του ParaView και ο δείκτης Jaccard υπολογίστηκε στο περιβάλλον της MATLAB® (MathWorks®, Natick, MA). Όλες οι δοκιμές πραγματοποιήθηκαν σε κοινό σύστημα αναφοράς (AMD Ryzen 3 PRO 2200G στα 3,5 GHz, 16 GB RAM, με λειτουργικό σύστημα Linux 64-bit).

4.3 Αποτελέσματα

Το προτεινόμενο σχήμα τμηματοποίησης, εφαρμόστηκε σε 16 σύνολα δεδομένων από αξονικές τομογραφίες (CT) και αξιολογήθηκε σε σύγκριση με τρεις τεχνικές τμηματοποίησης εικόνας, την K-Means, Markov Random Fields και την μέθοδο κατωφλίωσης. Πραγματοποιήθηκε τόσο ποιοτική όσο και ποσοτική αξιολόγηση, συγκρίνοντας τα εκτιμώμενα αποτελέσματα τμηματοποίησης με την πραγματική αορτή για κάθε περίπτωση. Για την επίτευξη της παραπάνω αξιολόγησης όλα τα διαθέσιμα σύνολα δεδομένων ευθυγραμμίστηκαν χρησιμοποιώντας το 3D Slicer. Όλα τα αποτελέσματα της αξιολόγησης παρουσιάζονται στις επόμενες υποενότητες.

4.3.1 Ποιοτική Αξιολόγηση

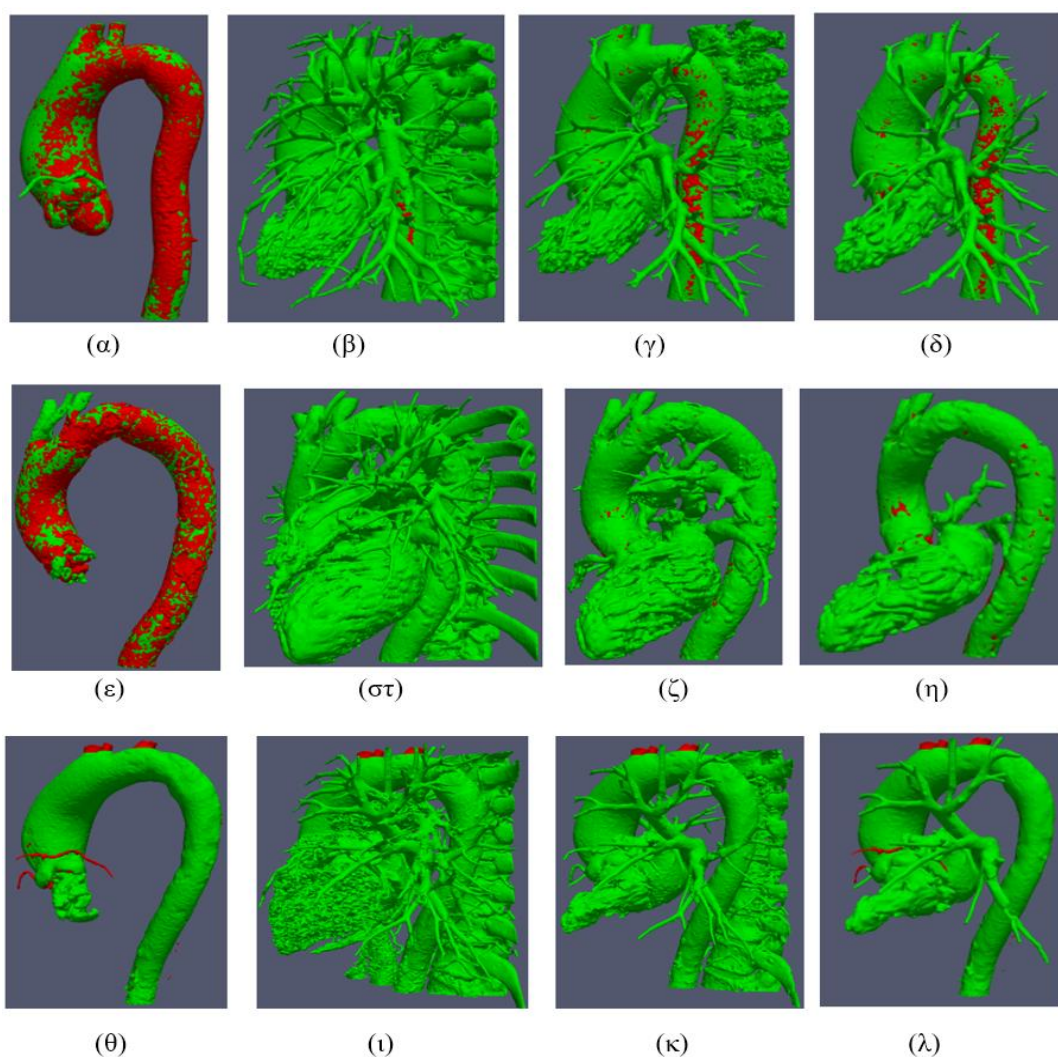
Η ποιοτική αξιολόγηση πραγματοποιήθηκε μέσω οπτικής εκτίμησης. Για κάθε σύνολο δεδομένων τα αποτελέσματα της τμηματοποίησης για όλες τις τέσσερις συγκρινόμενες μεθόδους τμηματοποίησης ευθυγραμμίστηκαν χωρικά σύμφωνα με το αντίστοιχο σύνολο απεικόνισης της αορτής. Σε όλες τις περιπτώσεις και για όλες τις συγκρινόμενες τεχνικές τμηματοποίησης χρησιμοποιήθηκαν οι βέλτιστες παράμετροι, όπως αυτές προσδιορίστηκαν

μετά από σειρά δοκιμών. Πιο συγκεκριμένα, για την προτεινόμενη πλήρως αυτόματη μέθοδο τμηματοποίησης, οι αρχικές κατώτερες και ανώτερες τιμές κατωφλίου ορίστηκαν στις 300 και 500 αντίστοιχα, ενώ για το επόμενο στάδιο του δικτύου Μαρκοβιανού Τυχαίου Πεδίου (Markov Random Field), οι τιμές 0, 300, 350 και 440 χρησιμοποιήθηκαν ως κεντρικά σημεία έντασης των κλάσεων.

Όσον αφορά τον αλγόριθμο ταξινόμησης K-Means, ο αριθμός των κλάσεων ορίστηκε σε 3 και επιλέχθηκαν οι τιμές 0, 300 και 500 ως τα κέντρα των σημείων κάθε κλάσης. Για τη μέθοδο τμηματοποίησης που βασίζεται στην κατωφλίωση, χρησιμοποιήθηκαν οι τιμές έντασης 350 και 450 ως κατώτερο και ανώτερο κατώφλι αντίστοιχα. Τέλος, για το δίκτυο Μαρκοβιανού Τυχαίου Πεδίου, χρησιμοποιήθηκαν οι τιμές έντασης 0, 300, 350 και 400 ως τα κέντρα των σημείων των τεσσάρων κλάσεων.

Στην **(Εικόνα 4.7)** παρουσιάζονται τρία παραδείγματα του προαναφερθέντος σχήματος οπτικής αξιολόγησης, η οποία απεικονίζει μια συγχώνευση των 3Δ μοντέλων που παρήχθησαν μετά την ανακατασκευή του εκτιμώμενου τμήματος της αορτής και της πραγματικής περιοχής της αορτής. Πιο συγκεκριμένα, οι **(Εικόνα 4.7 (α))**, **(Εικόνα 4.7 (ε))** και **(Εικόνα 4.7 (θ))**, αναπαριστούν τα 3Δ ανακατασκευασμένα μοντέλα από τη προτεινόμενη μέθοδο (με πράσινο χρώμα), οι **(Εικόνα 4.7 (β))**, **(Εικόνα 4.7 (στ))** και **(Εικόνα 4.7 (ι))**, παρουσιάζουν τα μοντέλα που προέκυψαν μετά την εφαρμογή της μεθόδου K-Means (πράσινο χρώμα), οι **(Εικόνα 4.7 (γ))**, **(Εικόνα 4.7 (ζ))** και **(Εικόνα 4.7 (κ))**, απεικονίζουν τα μοντέλα που προέκυψαν μετά την εφαρμογή της μεθόδου τμηματοποίησης μέσω της κατωφλίωσης (πράσινο χρώμα), και οι **(Εικόνα 4.7 (δ))**, **(Εικόνα 4.7 (η))** και **(Εικόνα 4.7 (λ))**, δείχνουν τα 3Δ μοντέλα της εκτιμώμενης αορτής που προέκυψαν με τη μέθοδο του Δικτύου Μαρκοβιανού Τυχαίου Πεδίου (επίσης πράσινο χρώμα). Σε όλες τις περιπτώσεις, η πραγματική περιοχή της αορτής όπως αυτή ορίστηκε από τους ειδικούς (με κόκκινο χρώμα), συγχωνεύτηκε με τα εκτιμώμενα 3Δ μοντέλα της αορτής.

Όπως φαίνεται στην **(Εικόνα 4.7)**, η καλύτερη ακρίβεια επιτεύχθηκε με την εφαρμογή της προτεινόμενης μεθόδου τμηματοποίησης δύο σταδίων καθώς τα συγχωνευμένα μοντέλα της εκτιμώμενης και της πραγματικής περιοχής της αορτής παρουσιάζουν ελάχιστες αποκλίσεις. Αντίθετα, τα 3Δ μοντέλα που ανακατασκευάστηκαν με τις μεθόδους K-Means, Κατωφλίωσης και Δικτύου Μαρκοβιανού Τυχαίου Πεδίου παρουσίασαν διάφορους βαθμούς υπο-τμηματοποίησης σε σύγκριση με την πραγματική περιοχή της αορτής. Γενικά, η υπο-τμηματοποίηση δεν είναι επιθυμητή καθώς είναι δυνατόν να συγχωνευτούν υπερ-τμηματοποιημένες περιοχές μιας εικόνας, αλλά από την άλλη πλευρά είναι αρκετά δύσκολο να διαχωριστούν μεγάλες περιοχές που έχουν υποστεί υπο-τμηματοποίηση [309].



Εικόνα 4.7: Συγχώνευση 3Δ μοντέλων των εκτιμώμενων και των πραγματικών περιοχών της αορτής για τρεις ασθενείς (α)-(δ), (ε)-(η), (θ)-(λ), Μετά την εφαρμογή της προτεινόμενης μεθόδου τμηματοποίησης (α), (ε), (θ), K-Means (β), (στ), (ι), Κατωφλίωση (γ), (ζ), (κ) και Μαρκοβιανών Τυχαίων Πεδίων (δ), (η), (λ)

4.3.2 Ποσοτική Αξιολόγηση

Όπως αναφέρθηκε, η ποσοτική ανάλυση των αποτελεσμάτων βασίστηκε στη μετρική του δείκτη Jaccard η οποία αξιολογεί την ομοιότητα μεταξύ των συνόλων της εκτιμώμενης τμηματοποιημένης αορτής και της πραγματικής όπως αυτή ορίστηκε χειροκίνητα από τους ειδικούς. Για να επιτευχθεί αυτό, οι συγκρινόμενες δομές σε κάθε περίπτωση έπρεπε πρώτα να ευθυγραμμιστούν πριν εφαρμοστεί η Εξ.(4-14) για τον υπολογισμό του δείκτη Jaccard. Οι τιμές του δείκτη Jaccard για όλες τις συγκρινόμενες τεχνικές παρουσιάζονται στον (Πίνακας 4.1). Δεδομένου ότι πρόκειται αποκλειστικά για μια μετρική ομοιότητας, προκύπτει ότι όσο υψηλότερη είναι η τιμή του δείκτη Jaccard τόσο καλύτερη είναι η αντιστοιχία μεταξύ των συγκρινόμενων μοντέλων και επομένως τόσο πιο ακριβής είναι η τμηματοποιημένη περιοχή.

Όπως φαίνεται στον (Πίνακας 4.1), οι τιμές του δείκτη Jaccard που αντιστοιχούν στην προτεινόμενη μεθοδολογία είναι σημαντικά υψηλότερες από εκείνες που αντιστοιχούν στις

άλλες τρεις τεχνικές τμηματοποίησης. Το γεγονός αυτό υποστηρίζει τα ευρήματα της ποιοτικής αξιολόγησης και πρακτικά υποδηλώνει ότι η προτεινόμενη μέθοδος προσφέρει ανώτερη ακρίβεια τμηματοποίησης σε σύγκριση με όλες τις άλλες τεχνικές που εξετάστηκαν στη μελέτη αυτή. Αντίθετα, σύμφωνα με τα αποτελέσματα που παρουσιάστηκαν, οι προσεγγίσεις βασισμένες στο K-Means και στην Κατωφλίωση (threshold-based) φαίνεται να υστερούν, όσον αφορά τις καταγεγραμμένες τιμές του δείκτη Jaccard οι οποίες παρουσιάζονται μειωμένες και στις δύο περιπτώσεις. Από την άλλη πλευρά, για την τμηματοποίηση βασισμένη στα Μαρκοβιανά Τυχαία Πεδία (Markov Random Fields), η ακρίβεια είναι ικανοποιητική αν και η συγκεκριμένη τεχνική εξακολουθεί να υπολείπεται της προτεινόμενης μεθοδολογίας όσον αφορά τις τιμές του δείκτη Jaccard μεταξύ των εκτιμώμενων και των πραγματικών περιοχών της αορτής.

Η υπεροχή της προτεινόμενης μεθόδου τμηματοποίησης για όλα τα διαθέσιμα σύνολα δεδομένων επιβεβαιώθηκε με τη διενέργεια στατιστικού ελέγχου δύο δειγμάτων (two-tailed paired t-test), ο οποίος χρησιμοποιήθηκε για να αποδειχθεί ότι η διαφορά μεταξύ των συγκρινόμενων τεχνικών όσον αφορά τη μετρική του δείκτη Jaccard που παρουσιάζεται στον **(Πίνακας 4.1)** είναι στατιστικά σημαντική και συνεπώς όχι τυχαία. Στην περίπτωση μας, η μηδενική υπόθεση (H_0), σύμφωνα με την οποία οι συγκρινόμενες τεχνικές δεν παράγουν στατιστικά σημαντικά διαφορετικές τιμές του δείκτη Jaccard, απορρίφθηκε σε όλες τις περιπτώσεις, εξετάζοντας τη προτεινόμενη μεθοδολογία έναντι των τριών τεχνικών σύγκρισης (με $p = 0$, $p = 2.975 \times 10^{-14}$ και $p = 6.063 \times 10^{-10}$ για τις τεχνικές K-Means, Κατωφλίωση και Μαρκοβιανά Τυχαία Πεδία, αντίστοιχα).

Σύνολα Δεδομένων	Προτεινόμενη Μεθοδολογία	K-Means	Κατωφλίωση	Δίκτυο Μαρκοβιανού Τυχαίου Πεδίου
1	0.887	0.276	0.518	0.709
2	0.896	0.322	0.562	0.747
3	0.889	0.274	0.518	0.731
4	0.888	0.291	0.515	0.718
5	0.868	0.332	0.561	0.779
6	0.871	0.270	0.530	0.764
7	0.872	0.319	0.501	0.752
8	0.876	0.328	0.515	0.789
9	0.889	0.327	0.520	0.797
10	0.881	0.319	0.490	0.784
11	0.886	0.327	0.580	0.788
12	0.852	0.331	0.630	0.780
13	0.884	0.326	0.590	0.786
14	0.885	0.320	0.625	0.795
15	0.883	0.294	0.534	0.751
16	0.882	0.335	0.601	0.792
$\mu \pm \sigma$	0.881 ± 0.011	0.312 ± 0.023	0.549 ± 0.044	0.766 ± 0.028

Πίνακας 4.1: Ο δείκτης Jaccard της πραγματικής περιοχής της αορτής, σε σύγκριση με την εκτιμώμενη περιοχή της αορτής για όλες τις συγκρινόμενες μεθόδους τμηματοποίησης.

4.4 Συζήτηση

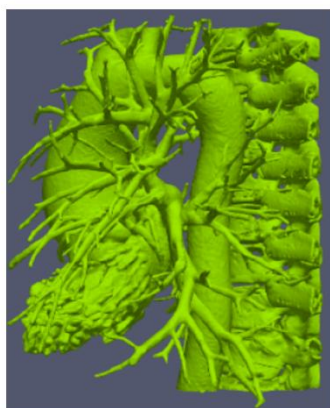
Η ακρίβεια και η ταχύτητα αποτελούν σημαντικούς παράγοντες στην τμηματοποίηση της αορτής παρέχοντας ακριβείς μετρήσεις για τον προγραμματισμό επεμβάσεων σε ανευρύσματα αορτής, διαδικασίες TAVI και οξέα αορτικά σύνδρομα. Πιο συγκεκριμένα, η αυτοματοποίηση για την εξασφάλιση, επιτάχυνση και χαρτογράφηση της τοποθέτησης της προσθετικής βαλβίδας θα μπορούσε να βελτιώσει σημαντικά τις διαδικασίες προγραμματισμού για επεμβάσεις TAVI [310]. Επιπλέον, μέσω της αυτόματης τμηματοποίησης, η μείωση του χρόνου ανάλυσης και η αναπαραγωγικότητα θα μπορούσαν να διαδραματίσουν καθοριστικό ρόλο στην αποτύπωση του ανευρύσματος κοιλιακής αορτής (AAA) [311]. Τέλος, οι αλγόριθμοι αυτόματης 3D τμηματοποίησης αποσκοπούν στο να επιτρέψουν την αυτόματη εκτίμηση της διαμέτρου της αορτής η οποία αποτελεί ένα σημαντικό μορφολογικό απεικονιστικό χαρακτηριστικό στην αξιολόγηση ασθενών με αορτική διάσχιση [312].

Όπως αναφέρθηκε προηγουμένως, η χειροκίνητη τμηματοποίηση δεν μπορεί να εγγυηθεί ποιότητα και αναπαραγωγικότητα, καθώς η ανθρώπινη παρέμβαση μπορεί να οδηγήσει σε ανακριβή αποτελέσματα εισάγοντας σφάλματα λόγω υποκειμενικής κρίσης [313]. Επιπλέον, τα εμπορικά εργαλεία τμηματοποίησης είναι κατά κύριο λόγο ημιαυτόματα, απαιτώντας συνήθως αρχικοποίηση από τον χρήστη για την ανάλυση της αορτής. Η χρήση συνελκτικών νευρωνικών δικτύων (CNNs) είναι επίσης δύσκολη, διότι απαιτείται μεγάλος όγκος εικονικών δεδομένων που να έχουν επισημανθεί με τις αντίστοιχες κατηγορίες [314]. Αντίθετα, η πλήρως αυτόματη μέθοδος τρισδιάστατης τμηματοποίησης της αορτής που προτείνουμε, θα μπορούσε να διευκολύνει και να τυποποιήσει τη μορφολογική ανάλυση της αορτής προσφέροντας αξιόπιστες και αναπαραγώγιμες μετρήσεις.

Στην παρούσα μελέτη προτείνεται μια πλήρως αυτόματη μέθοδος τρισδιάστατης τμηματοποίησης για την ανίχνευση και μοντελοποίηση της αορτής από δεδομένα τρισδιάστατης απεικόνισης αξονικής τομογραφίας (CT). Στη προτεινόμενη μεθοδολογία, αρχικά χρησιμοποιήθηκε ένα στάδιο προ-επεξεργασίας για τη διόρθωση της φωτεινότητας και αντίθεσης των απεικονιστικών δεδομένων, ακολουθούμενο από μια πλήρως αυτόματη μέθοδο τρισδιάστατης τμηματοποίησης η οποία περιλάμβανε δύο στάδια: μια αρχική διαδικασία τμηματοποίησης βασισμένη στην κατωφλίωση και μια επακόλουθη προσέγγιση τμηματοποίησης βασισμένη σε ταξινόμηση, η οποία βασιζόταν σε δίκτυο Μαρκοβιανού Τυχαίου Πεδίου (Markov Random Field). Στη συνέχεια, τα τμηματοποιημένα δεδομένα εξομαλύνθηκαν χρησιμοποιώντας Διακριτό Γκαουσιανό Φίλτρο και ακολούθησε η διαδικασία τρισδιάστατης ανακατασκευής για την απεικόνιση του τρισδιάστατου μοντέλου της αορτής. Το προτεινόμενο σχήμα τμηματοποίησης εφαρμόστηκε σε 16 σύνολα δεδομένων και τα αποτελέσματα της τμηματοποίησης αξιολογήθηκαν ποιοτικά και ποσοτικά σε σύγκριση με τη μέθοδο K-Means, την Κατωφλίωση και ένα σχήμα τμηματοποίησης βασισμένο σε

Μαρκοβιανά Τυχαία Πεδία. Σε όλες τις περιπτώσεις, τα παραγόμενα τρισδιάστατα μοντέλα συγκρίθηκαν με το πραγματικό μοντέλο της αορτής το οποίο είχε οριστεί χειροκίνητα από ειδικούς.

Η απόδοση και η αποτελεσματικότητα της προτεινόμενης μεθοδολογίας τμηματοποίησης και ανακατασκευής εξαρτώνται από ορισμένες βασικές παραμέτρους οι οποίες έπρεπε να βελτιστοποιηθούν για να εξασφαλιστεί η μέγιστη δυνατή ακρίβεια. Ένας από τους πιο καθοριστικούς παράγοντες που επηρεάζουν την απόδοση της προτεινόμενης μεθόδου τμηματοποίησης είναι ο αριθμός των κατηγοριών που χρησιμοποιούνται στο δεύτερο στάδιο, στο δίκτυο Μαρκοβιανού Τυχαίου Πεδίου. Ο προτεινόμενος αλγόριθμος αξιολογήθηκε χρησιμοποιώντας διαφορετικό αριθμό κατηγοριών και παρουσιάζονται τα αντίστοιχα ανακατασκευασμένα τρισδιάστατα μοντέλα για κάθε περίπτωση. Πιο συγκεκριμένα, στην **(Εικόνα 4.8 (α))** παρουσιάζεται το τρισδιάστατο μοντέλο που παράχθηκε μετά την ανακατασκευή της εκτιμώμενης αορτής χρησιμοποιώντας τρεις κλάσεις, η **(Εικόνα 4.8 (β))** δείχνει το μοντέλο από την πλήρως αυτόματη μέθοδο τρισδιάστατης τμηματοποίησης χρησιμοποιώντας τέσσερις κλάσεις και τέλος η **(Εικόνα 4.8 (γ))** απεικονίζει το μοντέλο που προέκυψε όταν χρησιμοποιήθηκαν πέντε κλάσεις.



(α)



(β)

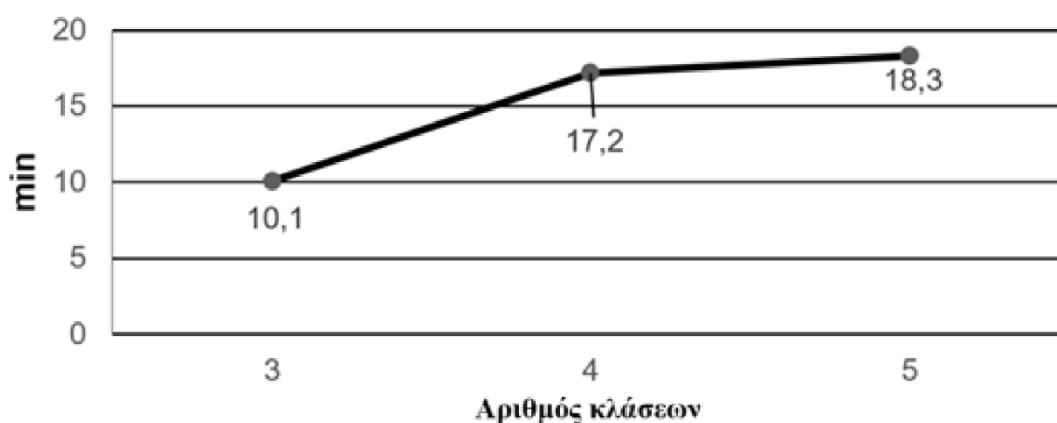


(γ)

Εικόνα 4.8: Τα 3Δ μοντέλα της αορτής που έχουν προκύψει από μία μόνο περίπτωση, κατά την εφαρμογή της προτεινόμενης μεθόδου τμηματοποίησης με χρήση: (α) τριών κλάσεων, (β) τεσσάρων κλάσεων και (γ) πέντε κλάσεων.

Ο αριθμός των υπολογισμών που απαιτούνται και κατά συνέπεια ο χρόνος επεξεργασίας, εξαρτάται άμεσα από τον αριθμό των κλάσεων που χρησιμοποιούνται στο στάδιο του δικτύου Μαρκοβιανών Τυχαίων Πεδίων (Markov Random Fields) του προτεινόμενου αλγορίθμου τμηματοποίησης. Αυτό φαίνεται καθαρά στην **(Εικόνα 4.9)**, όπου καταγράφεται ο χρόνος επεξεργασίας για 3, 4 και 5 κλάσεις. Όπως φαίνεται στην **(Εικόνα 4.9)**, ο χρόνος επεξεργασίας αυξάνεται όσο αυξάνεται ο αριθμός των κλάσεων. Ο χρόνος που απαιτείται για την

τμηματοποίηση αποτελεί κρίσιμο παράγοντα στην επεξεργασία εικόνας. Όπως φαίνεται και πάλι στην **(Εικόνα 4.9)**, ο αριθμός των κλάσεων παίζει καθοριστικό ρόλο στον χρόνο τμηματοποίησης. Καθώς αυξάνεται ο αριθμός των κλάσεων, αυξάνεται με τον ίδιο ρυθμό και ο απαιτούμενος χρόνος για την τμηματοποίηση.



Εικόνα 4.9: Καταγεγραμμένος χρόνος επεξεργασίας για την προτεινόμενη μεθοδολογία από ένα μόνο σύνολο δεδομένων, για διάφορο αριθμό κλάσεων στο δεύτερο στάδιο της διαδικασίας τμηματοποίησης.

Σύμφωνα με τα παρουσιαζόμενα ποιοτικά αποτελέσματα, το προτεινόμενο πλήρως αυτόματο σύστημα τμηματοποίησης παρείχε πιο ακριβή απεικόνιση των ανακατασκευασμένων 3Δ μοντέλων της αορτής σε σύγκριση με τις τεχνικές που συγκρίθηκαν λόγω της αποδοτικότητας της διαδικασίας τμηματοποίησης δύο σταδίων. Όπως φαίνεται στον **(Πίνακα 4.2)**, οι ταχύτερες προσεγγίσεις για την τμηματοποίηση της αορτής είναι οι τεχνικές που βασίζονται σε κατώφλια έντασης και στον αλγόριθμο K-Means. Ωστόσο, όπως απεικονίζεται στην **(Εικόνα 4.7)**, τα πραγματικά 3Δ μοντέλα της αορτής που παράγονται από αυτές τις δύο τεχνικές δεν μπορούν να χρησιμοποιηθούν αποδοτικά στην κλινική πράξη, καθώς αποτυγχάνουν να διαχωρίσουν την αορτή από τους γύρω ιστούς. Επιπλέον, παρόλο που η προτεινόμενη μέθοδος πέτυχε καλύτερη ακρίβεια σε σύγκριση με την τεχνική των Μαρκοβιανών Τυχαίων Πεδίων (Markov Random Fields) όταν εφαρμόζεται μεμονωμένα, ο **(Πίνακας 4.2)** δείχνει ότι ο συνολικός χρόνος επεξεργασίας και για τις δύο τεχνικές είναι παρόμοιος.

Σύνολα Δεδομένων	Προτεινόμενη Μεθοδολογία	K-Means	Κατωφλίωση	Δίκτυο Μαρκοβιανού Τυχαίου Πεδίου
1	17.200	5.280	0.067	17.900
2	17.280	5.810	0.100	17.160
3	17.560	6.850	0.100	17.600
4	14.360	4.330	0.133	14.510
5	18.610	6.100	0.167	18.660
6	15.830	4.810	0.150	16.250
7	13.450	4.180	0.083	13.530
8	15.400	5.060	0.150	15.383
9	16.410	5.330	0.150	16.466
10	14.630	4.560	0.150	14.700
11	28.950	14.630	0.250	29.150
12	15.360	4.600	0.050	15.450
13	14.830	4.260	0.217	14.610
14	25.400	15.330	0.533	28.850
15	24.780	15.010	0.717	34.666
16	16.350	8.600	0.300	25.133
$\mu \pm \sigma$	17.900 ± 4.481	7.171 ± 4.039	0.207 ± 0.179	19.376 ± 6.403

Πίνακας 4.2: Συνολικός χρόνος επεξεργασίας τμηματοποίησης για όλα τα σύνολα δεδομένων σε λεπτά

4.5 Συμπεράσματα

Σε αυτό το άρθρο παρουσιάστηκε μια νέα πλήρως αυτόματη μέθοδος τρισδιάστατης (3Δ) τμηματοποίησης βασισμένη σε μια διαδικασία τμηματοποίησης δύο σταδίων για την ανίχνευση και μοντελοποίηση της αορτής σε δεδομένα 3Δ αξονικής τομογραφίας (CT). Η προτεινόμενη πλήρως αυτόματη μέθοδος 3Δ τμηματοποίησης και μοντελοποίησης συγκρίθηκε άμεσα με τις μεθόδους K-Means, Κατωφλίωσης και τα Δίκτυα Μαρκοβιανών Τυχαίων Πεδίων (Markov Random Field Network) όσον αφορά την ακρίβεια του ανακατασκευασμένου 3Δ μοντέλου της αορτής σε κάθε περίπτωση. Όλες οι συγκρινόμενες τεχνικές εφαρμόστηκαν σε 16 σύνολα δεδομένων 3Δ CT και αξιολογήθηκαν τόσο ποιοτικά όσο και ποσοτικά. Η προτεινόμενη μέθοδος τμηματοποίησης δύο σταδίων πέτυχε ανώτερη απόδοση σε σύγκριση με τις άλλες τρεις τεχνικές τμηματοποίησης γεγονός που την καθιστά κατάλληλη για πιθανή χρήση σε πολυάριθμες εφαρμογές στην κλινική πρακτική, όπως ο προγραμματισμός και η αξιολόγηση θεραπειών.

Κεφάλαιο 5: Συγκριτική Μελέτη Αρχιτεκτονικών Βαθιάς Μάθησης για την Αυτόματη Τμηματοποίηση της Κοιλιακής Αορτής σε 3Δ Δεδομένα Αξονικού Τομογράφου

5.1 Εισαγωγή

Το ανεύρυσμα κοιλιακής αορτής (Abdominal Aortic Aneurysm - AAA) είναι μια αγγειακή νόσος που χαρακτηρίζεται από εντοπισμένη διάταση του αγγείου (όπου η μέγιστη διάμετρος μεταξύ των νεφρικών και λαγονίων αρτηριών υπερβαίνει κατά περισσότερο από 50% τη φυσιολογική του διάμετρο), η οποία αν δεν αντιμετωπιστεί, μπορεί να οδηγήσει σε καταστροφική ρήξη [315], [316]. Η έγκαιρη ανίχνευση του AAA και άλλων συναφών καταστάσεων όπως ο ενδοαυλικός θρόμβος και η αγγειακή ασβέστωση είναι ζωτικής σημασίας για την πορεία βελτίωσης των ασθενών [317]. Ως εκ τούτου, η γνώση της γεωμετρίας της κοιλιακής αορτής (ξεκινώντας από το διάφραγμα και μέχρι τη διακλάδωσή της στις κοινές λαγόνιες αρτηρίες), συμπεριλαμβανομένης της διαμέτρου και του μήκους της, είναι κρίσιμη για τη διαγνωστική αξιολόγηση του AAA (καθώς η κοιλιακή αορτή διαδραματίζει σημαντικό ρόλο στην αιμάτωση πολλών οργάνων) [318], [319].

Η ιατρική απεικόνιση αποτελεί βασικό στοιχείο στη διάγνωση και διαχείριση του AAA. Μεταξύ των διαφόρων τεχνικών απεικόνισης, η υπολογιστική τομογραφία (CT) είναι ιδιαίτερα πολύτιμη για την ικανότητά της να παρέχει εικόνες υψηλής ανάλυσης των ανατομικών δομών. Οι τεχνικές απεικόνισης και ιδιαίτερα η τρισδιάστατη (3Δ) απεικόνιση, είναι απαραίτητες για να επιτρέψουν στους επαγγελματίες υγείας να αξιολογούν με ακρίβεια τις ανατομικές περιοχές ενδιαφέροντος (ROIs) πέρα από τις δυνατότητες των δισδιάστατων (2Δ) μεθόδων απεικόνισης [263]. Μέσω της 3Δ απεικόνισης, οι ειδικοί στον τομέα της ιατρικής μπορούν να αξιολογήσουν ποιοτικά και ποσοτικά γεωμετρικά χαρακτηριστικά και ανωμαλίες παρέχοντας έναν αξιόπιστο τρόπο εξέτασης περιοχών ενδιαφέροντος, όπως η κοιλιακή αορτή [264].

Η τμηματοποίηση εικόνας αποτελεί απαραίτητο βήμα πριν από την τρισδιάστατη απεικόνιση, καθώς ορίζει με σαφήνεια και απομονώνει την ανατομική περιοχή ενδιαφέροντος (ROI) από τα γύρω δεδομένα [265], [286]. Παρόλο που οι χειροκίνητες ή ημι-αυτόματες μέθοδοι τμηματοποίησης χρησιμοποιούνται συνήθως στην κλινική πρακτική, η εξάρτησή τους από την ανθρώπινη παρέμβαση μπορεί να περιορίσει την επεκτασιμότητα και την ακρίβειά τους. Οι αυτοματοποιημένες προσεγγίσεις τμηματοποίησης που βασίζονται στη μηχανική μάθηση και σε αρχιτεκτονικές βαθιάς μάθησης αποτελούν μια ελπιδοφόρα εναλλακτική εξαλείφοντας πολλά από τα εγγενή μειονεκτήματα των χειροκίνητων μεθόδων [205], [320].

Τα τελευταία χρόνια, τα βαθιά νευρωνικά δίκτυα (DNNs) έχουν αποδειχθεί ιδιαίτερα αποτελεσματικά σε εργασίες ιατρικής απεικόνισης [321], [322]. Η ικανότητά τους να εξάγουν

ζωτικής σημασίας χαρακτηριστικά και μοτίβα από τα δεδομένα εικόνων, αξιοποιώντας μεγάλο όγκο δεδομένων, έχει αποδειχθεί καθοριστική για την παροχή πιο ακριβών και αποτελεσματικών αποτελεσμάτων ανάλυσης [323]. Επιπλέον, μέσω ποσοτικής ανάλυσης και ακριβών μετρήσεων, τα DNNs μπορούν να προσφέρουν ακριβή διάκριση μεταξύ γειτονικών οργάνων ή άλλων περιοχών ενδιαφέροντος, γεγονός που είναι χρήσιμο στην κατηγοριοποίηση ανατομικών χαρακτηριστικών, ανωμαλιών και παθήσεων όπως το ανεύρυσμα κοιλιακής αορτής (AAA) [324].

Σε αυτό το πλαίσιο, τα συνελκτικά νευρωνικά δίκτυα (CNNs) έχουν καθιερωθεί ως αποτελεσματικά εργαλεία για την τμηματοποίηση ιατρικών εικόνων αξιοποιώντας διαφορετικά επίπεδα για εξαγωγή και απεικόνιση χαρακτηριστικών [325]. Συγκεκριμένα, οι Lopez-Linares και συν. [283] πρότειναν μια πλήρως αυτοματοποιημένη τεχνική βασισμένη στα CNNs για την ανίχνευση και τμηματοποίηση ενδοαυλικής θρόμβωσης (Intraluminal Thrombus - ILT) σε μετεγχειρητικές αγγειογραφίες υπολογιστικής τομογραφίας (CTA). Επιπλέον, οι Lareyre και συν. [18] παρουσίασαν μια αυτόματη μέθοδο ανίχνευσης περιγραμμάτων με βάση την κατωφλίωση για την ανίχνευση και αξιολόγηση των βασικών χαρακτηριστικών του AAA.

Μεταξύ των αρχιτεκτονικών που βασίζονται στα CNNs, τα 3D U-Net και SegResNet χρησιμοποιούνται ευρέως για την τρισδιάστατη τμηματοποίηση ιατρικών εικόνων λόγω της ικανότητάς τους να επεξεργάζονται ογκομετρικά δεδομένα [326]. Ωστόσο, το σταθερό πεδίο λήψης (receptive field) των επιπέδων των CNNs περιορίζει την ικανότητά τους να συλλαμβάνουν αποτελεσματικά μακρινές συσχετίσεις, κάτι που μπορεί να οδηγήσει σε υποβέλτιστη απόδοση ως προς τη μοντελοποίηση της συνολικής πληροφορίας περιβάλλοντος [327]. Αυτός ο περιορισμός μπορεί να αντιμετωπιστεί με τη χρήση μοντέλων τύπου transformer, τα οποία υπερέχουν στη μοντελοποίηση μακρινών σχέσεων μέσω μηχανισμών αυτό-προσοχής (self-attention), βελτιώνοντας την ακρίβεια της τμηματοποίησης [328]. Έτσι, τα UNETR και SwinUNETR είναι αρχιτεκτονικές DNNs (Deep Neural Networks) βασισμένες σε transformers που αξιοποιούν μηχανισμούς αυτό-προσοχής για να ξεπεράσουν τους περιορισμούς των παραδοσιακών CNNs [329]. Το UNETR ενσωματώνει τους transformers στη δομή του U-Net, ενώ το SwinUNETR βασίζεται σε αυτή την προσέγγιση με έναν βελτιωμένο σχεδιασμό ειδικά προσαρμοσμένο για τμηματοποίηση εικόνων [330].

Παρά την αποδεδειγμένη αποτελεσματικότητα της μηχανικής μάθησης στην τμηματοποίηση περιοχών της κοιλιακής αορτής η πληθώρα αρχιτεκτονικών βαθιάς μάθησης και οι διαφορές στα δεδομένα εισόδου, καθιστούν δύσκολη τη σύγκριση των αποτελεσμάτων. Για να αντιμετωπιστεί αυτό, είναι απαραίτητες οι συγκριτικές μελέτες για τη συστηματική αξιολόγηση της απόδοσης διαφορετικών αρχιτεκτονικών υπό ενοποιημένες συνθήκες και σύνολα δεδομένων, επιτρέποντας πιο αξιόπιστα συμπεράσματα όσον αφορά την ακρίβεια, την ανθεκτικότητα και την κλινική εφαρμοσιμότητα στην τμηματοποίηση της κοιλιακής αορτής.

Από αυτή την άποψη, οι Camara και συν. [331] χρησιμοποίησαν συνελκτικά νευρωνικά δίκτυα (CNNs) για την ανίχνευση και τον εντοπισμό αποτελεσμάτων CTA από υπονεφρικά ανευρύσματα κοιλιακής αορτής (AAAs), τυχαία κατανεμημένα σε σύνολα 60%, 10% και 30% για εκπαίδευση, επικύρωση και δοκιμή του μοντέλου αντίστοιχα. Τα αποτελέσματά τους έδειξαν ακρίβεια 99.1% για το τελικό εξατομικευμένο μοντέλο CNN και μια περιοχή κάτω από την χαρακτηριστική καμπύλη λειτουργίας με 0.99. Με παρόμοιο σχεδιασμό, οι Cao και συν. [332] υλοποίησαν τρία CNNs βασισμένα στην αρχιτεκτονική U-Net για την τμηματοποίηση της αορτικής διαχωριστικής νόσου τύπου B, και με το CNN3 πέτυχαν τις καλύτερες τιμές του δείκτη Dice με $93\% \pm 0.01$ για ολόκληρη την αορτή, $93\% \pm 0.01$ για τον πραγματικό αυλό και $91\% \pm 0.02$ για τον ψευδή αυλό αντίστοιχα.

Αν και η υλοποίηση των διαφορετικών μοντέλων μηχανικής μάθησης προσφέρει ένα πλαίσιο για την ομοιόμορφη αξιολόγηση δεικτών απόδοσης στο ίδιο σύνολο δεδομένων, η επίδραση διαφορετικών τύπων δεδομένων εισόδου (π.χ., ανάλυση αξονικής τομογραφίας ή επίπεδα θορύβου) στην απόδοση των μοντέλων βαθιάς μάθησης για τμηματοποίηση παραμένει σε μεγάλο βαθμό ανεξερεύνητη. Από αυτή την άποψη η σύγκριση διαφορετικών αλγορίθμων τμηματοποίησης όταν εκπαιδεύονται και αξιολογούνται σε ένα (περιορισμένο ή ομοιογενές) σύνολο δεδομένων, περιορίζει τη γενικευσιμότητα σε διαφορετικούς πληθυσμούς ασθενών, μειώνοντας τη δυνατότητά τους να εφαρμοστούν σε πραγματικά κλινικά περιβάλλοντα.

Επιπλέον, δεδομένου ότι τα ανευρύσματα κοιλιακής αορτής προκαλούν δομικές και μορφολογικές αλλαγές στην αορτή όπως διάταση και ακανόνιστο περίγραμμα τοιχώματος (που διαφέρουν σημαντικά από την ομοιόμορφη εμφάνιση υγιών αορτών), τα μοντέλα βαθιάς μάθησης που εκπαιδεύονται αποκλειστικά σε εικόνες υγιών αορτών ενδέχεται να δυσκολεύονται να γενικεύσουν σε περιπτώσεις με AAA [333]. Αυτή η αναντιστοιχία μπορεί να οδηγήσει σε ανακριβή τμηματοποίηση σε ανευρυσματικές αορτές καθώς το μοντέλο μπορεί λανθασμένα να υποθέτει πρότυπα χαρακτηριστικά, όπως η σταθερή διάμετρος και το σχήμα του αγγείου, μη έχοντας εκτεθεί σε αυτές τις παθολογικές παραλλαγές κατά την εκπαίδευση.

Λαμβάνοντας υπόψη όλα τα παραπάνω, στόχος αυτής της μελέτης είναι η διεξαγωγή μιας ολοκληρωμένης αξιολόγησης, τόσο ποιοτικής όσο και ποσοτικής για την αποτίμηση και σύγκριση της απόδοσης διαφόρων αρχιτεκτονικών βαθιάς μάθησης ενσωματώνοντας δεδομένα από υγιή και ασθενή άτομα από διαφορετικά διακριτά σύνολα δεδομένων. Συγκεκριμένα, υλοποιούνται τέσσερα συγκεκριμένα νευρωνικά δίκτυα (δηλ. U-Net, SegResNet, UNETR και SwinUNETR) για την αυτόματη τμηματοποίηση της κοιλιακής αορτής, συνδυάζοντας και αναλύοντας λεπτομερώς συνελκτικές και μετασχηματιστικές (transformer) αρχιτεκτονικές. Αυτό περιλαμβάνει όλα τα στάδια της προ-επεξεργασίας, της εκπαίδευσης και της δημιουργίας πλέγματος, με εκτεταμένη εκπαίδευση και δοκιμή σε πάνω από 56 εικόνες αορτής από υγιή άτομα από διαφορετικά κέντρα (δημόσιο σύνολο δεδομένων),

και επιπλέον επικύρωση με 20 εικόνες κοιλιακής αορτής από ασθενείς με AAA (ιδιωτικό σύνολο δεδομένων).

Η υπόλοιπη εργασία οργανώνεται ως εξής: στην ενότητα «Μεθοδολογία», περιγράφονται οι τέσσερις αρχιτεκτονικές βαθιάς μάθησης που χρησιμοποιήθηκαν για την εκπαίδευση των δεδομένων και την εξαγωγή των προβλέψεων τμηματοποίησης για τις αόρατες εικόνες. Αναλύεται επίσης το διάγραμμα ροής των αρχιτεκτονικών CNN και transformer. Στην ενότητα «Αποτελέσματα» παρουσιάζονται τα συγκριτικά αποτελέσματα, στην ενότητα «Συζήτηση» εξετάζονται οι βασικές πτυχές των τεσσάρων προαναφερθέντων συνελκτικών νευρωνικών δικτύων και αξιολογεί τη δυναμική χρήση τους στην κλινική πράξη, συμπεριλαμβανομένου του σχεδιασμού και της αποτίμησης της θεραπείας και τέλος στην ενότητα «Συμπεράσματα» συνοψίζονται τα κύρια ευρήματα της μελέτης, υπογραμμίζονται τα πλεονεκτήματα της προτεινόμενης μεθοδολογίας και προτείνονται κατευθύνσεις για μελλοντική έρευνα.

5.2 Μεθοδολογία

5.2.1 Συλλογή Δεδομένων

Σε αυτή τη μελέτη, χρησιμοποιήθηκαν εβδομήντα έξι (76) εικόνες αξονικής τομογραφίας (CT) για την τμηματοποίηση της κοιλιακής αορτής. Από αυτές, πενήντα έξι (56) εικόνες (δημόσιο σύνολο δεδομένων) προήλθαν από υγιή άτομα και οι οποίες προέρχονται από τρεις διαφορετικές βάσεις δεδομένων: το KiTS19 Grand Challenge, το Rider Lung CT dataset και περιστατικά από το Νοσοκομείο Dongyang [334], [335]. Το πρωτόκολλο λήψης περιλάμβανε απεικόνιση με αγγειογραφία αξονικής τομογραφίας (CTA), με παραμέτρους εικόνας 512×512 εικονοστοιχεία για τα KiTS19 Grand Challenge και Rider Lung CT datasets και 512×666 εικονοστοιχεία για τα δεδομένα από το Νοσοκομείο Dongyang. Το διάμεσο πάχος τομής και ο αριθμός των αξονικών τομών ήταν 5 mm, 0.625 mm, 3 mm, 146, 1008 και 149 αντίστοιχα, για τα σύνολα δεδομένων KiTS19 Grand Challenge, Rider Lung CT και του Νοσοκομείου Dongyang.

Επιπλέον, είκοσι (20) εικόνες CT λήφθηκαν από ασθενείς με ανεύρυσμα κοιλιακής αορτής (AAA), οι οποίες παρασχέθηκαν από το Τμήμα Αγγειοχειρουργικής του Γενικού Πανεπιστημιακού Νοσοκομείου ΑΤΤΙΚΟΝ (ιδιωτικό σύνολο δεδομένων). Όλα τα ιδιωτικά δεδομένα αποκτήθηκαν με τη χρήση 3Δ αξονικού τομογράφου (Philips Brilliance 16-τομών), με τις ακόλουθες παραμέτρους μεγέθους για τις ψηφιακές εικόνες: μέσες διαστάσεις πίνακα: $512 \times 512 \times 553,3$ εικονοστοιχεία, μέση απόσταση ογκοστοιχείων: $0.84 \times 0.84 \times 2.31$ mm, διάμεσες διαστάσεις πίνακα: $512 \times 512 \times 448$ εικονοστοιχεία και διάμεση απόσταση ογκοστοιχείων: $0.88 \times 0.88 \times 1.13$ mm.

5.2.2 Προ-Επεξεργασία

Για να διασφαλιστεί η ομοιογένεια σε όλα τα σύνολα δεδομένων και να μειωθεί η επίδραση πιθανών ακραίων τιμών, εφαρμόστηκε περιορισμός έντασης με εύρος $[-275, 1900]$ [336]. Το συγκεκριμένο εύρος καθορίστηκε μέσω επαναληπτικών διερευνητικών δοκιμών που πραγματοποιήθηκαν στο σύνολο των δεδομένων με στόχο τη διατήρηση των βασικών χαρακτηριστικών της περιοχής της αορτής, ενώ ταυτόχρονα αποκλείονταν μη σχετικές μεταβολές. Επιπλέον, το στενότερο αυτό εύρος μπορεί δυνητικά να περιλαμβάνει τις τυπικές τιμές για τους μαλακούς ιστούς, συμπεριλαμβανομένων των αγγειακών δομών οι οποίες είναι κρίσιμες για την ακριβή αποτύπωση της αορτής. Στη συνέχεια, οι εντάσεις κανονικοποιήθηκαν στο εύρος $[0,1]$, μειώνοντας περαιτέρω την αριθμητική αστάθεια κατά την εκπαίδευση του μοντέλου αποτρέποντας πολύ μεγάλες ή πολύ μικρές τιμές εισόδου που θα μπορούσαν να εμποδίσουν τη διαδικασία μάθησης.

Για την αύξηση των δεδομένων εισόδου και κατά συνέπεια την αποδοτικότητα του μοντέλου, εφαρμόστηκε η διαδικασία εμπλουτισμού των δεδομένων, η οποία περιλάμβανε τεχνικές τυχαίας αναστροφής, περιστροφής και μετατόπισης της έντασης [337]. Συγκεκριμένα, εφαρμόστηκε πιθανότητα 25% για όλες τις τεχνικές εμπλουτισμού, αυξάνοντας την ποικιλία των δεδομένων εκπαίδευσης και προσδίδοντας στο μοντέλο αυξημένη ικανότητα γενίκευσης. Μετά τον εμπλουτισμό των δεδομένων, οι εικόνες επαναδειγματολήθηκαν ομοιόμορφα σε προκαθορισμένη ανάλυση με διαστάσεις εικονοστοιχείων $[1.0, 1.0, 1.5]$, διασφαλίζοντας έτσι τη συνέπεια μεταξύ των δύο συνόλων δεδομένων.

5.2.3 Τμηματοποίηση μέσω Βαθιάς Μάθησης

Για να επιτύχουμε αποτελεσματική τμηματοποίηση και να διεξάγουμε μια ολοκληρωμένη σύγκριση, υλοποιήσαμε τέσσερις αρχιτεκτονικές βαθιάς μάθησης: U-Net, SegResNet, UNETR και SwinUNETR. Η απόδοση κάθε μοντέλου αξιολογήθηκε σε σύγκριση με την πραγματική απεικόνιση της αορτής (ground truth) για κάθε περίπτωση η οποία είχε χειροκίνητα αποτυπωθεί από ιατρούς ειδικούς στις καρδιαγγειακές παθήσεις. Η διαδικασία αυτή πραγματοποιήθηκε και για τα δύο σύνολα δεδομένων.

Βασισμένοι στη συλλογιστική μας ότι οι μέθοδοι βαθιάς μάθησης θα πρέπει να επιδεικνύουν υψηλή απόδοση ανεξαρτήτως εισόδου (δηλαδή ανεξαρτήτως τύπου δεδομένων), εκπαιδεύσαμε τα μοντέλα χρησιμοποιώντας το μεγαλύτερο δημόσιο σύνολο δεδομένων. Στο πλαίσιο αυτό, εάν τα μοντέλα κατόρθωναν να τμηματοποιήσουν επιτυχώς την κοιλιακή αορτή σε ασθενείς (ιδιωτικό σύνολο δεδομένων), αυτό θα υποδήλωνε την ικανότητά τους να τμηματοποιούν με ακρίβεια τόσο σε δεδομένα από ασθενείς όσο και από υγιή άτομα.

Έτσι, εφαρμόσαμε έναν τυχαίο διαχωρισμό 80–20 στο δημόσιο (υγιές) σύνολο δεδομένων για τις φάσεις εκπαίδευσης και δοκιμής (δηλαδή, το 80% του συνόλου διατέθηκε για εκπαίδευση, ώστε τα μοντέλα να μάθουν από σημαντικό όγκο δεδομένων, ενώ το υπόλοιπο 20% χρησιμοποιήθηκε ως σύνολο δοκιμής, προκειμένου να αξιολογηθεί η απόδοση και η γενικευσιμότητα των εκπαιδευμένων μοντέλων). Κατά τη διαδικασία της εκπαίδευσης, όλες οι εικόνες εισήχθησαν στα μοντέλα βαθιάς μάθησης σε μορφή μη επικαλυπτόμενων παραθύρων εικόνας $64 \times 64 \times 64$, κεντραρισμένων σε τυχαία σημεία εντός της μάσκας τμηματοποίησης με στόχο τη μείωση της χρήσης μνήμης GPU.

Αφού έγινε η εκπαίδευση και δοκιμή στο δημόσιο σύνολο δεδομένων, χρησιμοποιήθηκε μια μέθοδος διασταυρούμενης επικύρωσης δύο πτυχών (two-fold cross-validation) για την αξιολόγηση της απόδοσης του μοντέλου στο ιδιωτικό σύνολο δεδομένων (AAA). Επιπλέον, χρησιμοποιήσαμε τη συνάρτηση απώλειας Dice σε συνδυασμό με τη διασταυρούμενη εντροπία (cross-entropy), προκειμένου να ποσοτικοποιηθεί η επικάλυψη μεταξύ των προβλεπόμενων και των πραγματικών τμηματοποιήσεων λόγω της αποτελεσματικότητάς της στην ανίχνευση μικρότερων αγγείων [338]. Η συνάρτηση απώλειας που εφαρμόστηκε σε όλα τα μοντέλα βαθιάς μάθησης υπολογίστηκε όπως περιγράφεται παρακάτω.

Για τα βήματα της εκπαίδευσης, χρησιμοποιήθηκε μια σύνθετη συνάρτηση απώλειας που συνδυάζει την απώλεια διασταύρωσης εντροπίας (cross-entropy loss) και την απώλεια Dice. Αυτή η υβριδική συνάρτηση απώλειας σχεδιάστηκε ώστε να αξιοποιεί τα πλεονεκτήματα τόσο της ακρίβειας κατά ταξινόμηση σε επίπεδο εικονοστοιχείου όσο και των μετρήσεων ομοιότητας με βάση την επικάλυψη κάτι που είναι ιδιαίτερα ωφέλιμο για τη βελτίωση της ποιότητας τμηματοποίησης της δομής της αορτής με παρακλάδια που είναι μικρότερα από τη βασική δομή του αυλού. Η συνδυασμένη συνάρτηση απώλειας ορίστηκε ως εξής:

$$Loss = \alpha L_{CE} + \beta L_{DSC} \quad (5-1)$$

όπου το L_{CE} είναι η απώλεια διασταύρωσης εντροπίας, το L_{DSC} είναι η απώλεια Dice και α, β είναι οι συντελεστές βαρών ορισμένοι στο 0.5.

Η απώλεια διασταύρωσης εντροπίας υπολογίστηκε μέσω της ακόλουθης εξίσωσης :

$$L_{CE} = - \sum_{i=1}^N \sum_{c=1}^C y_{ic} \log(p_{ic}) \quad (5-2)$$

όπου N είναι ο συνολικός αριθμός των εικονοστοιχείων (pixels) σε κάθε παρτίδα, C είναι ο αριθμός των κλάσεων και y_{ic} είναι η πραγματική ετικέτα για το εικονοστοιχείο i και την κλάση c . Συνεπώς, το p_{ic} είναι η προβλεπόμενη πιθανότητα ότι το εικονοστοιχείο i ανήκει στην κλάση c , με $y_{ic} = 1$ αν το εικονοστοιχείο i ανήκει στην κλάση c , και $y_{ic} = 0$ αν δεν ανήκει. Η απώλεια Dice υπολογίστηκε ως εξής:

$$L_{Dice} = 1 - \frac{2 \sum_{i=1}^N y_i p_i + \epsilon}{\sum_{i=1}^N y_i + \sum_{i=1}^N p_i + \epsilon} \quad (5-3)$$

όπου το y_i είναι η δυαδική ετικέτα αλήθειας εδάφους για το εικονοστοιχείο i , το p_i είναι η προβλεπόμενη πιθανότητα για το εικονοστοιχείο i , και το ϵ είναι μία μικρή σταθερά που προστίθεται για να αποτραπεί η διαίρεση με το μηδέν.

Κάθε αρχιτεκτονική βαθιάς μάθησης παρουσιάζεται συνοπτικά παρακάτω ενώ οι υπερπαραμέτροι κάθε αρχιτεκτονικής παρουσιάζονται στον **(Πίνακας 5.1)**. Οι τιμές των υπερπαραμέτρων επιλέχθηκαν με βάση τις ακόλουθες διαδικασίες. Ο ρυθμός εκμάθησης καθορίστηκε μέσω μίας αναζήτησης πλέγματος στο σύνολο επικύρωσης εξερευνώντας τιμές στο εύρος από 10^{-5} έως 10^{-1} . Επιπλέον, ο ρυθμός απόρριψης περιλήφθηκε επίσης στην αναζήτηση πλέγματος με δοκιμές σε τιμές 0.01, 0.1 και 0.35. Ο αριθμός εποχών εκπαίδευσης ορίστηκε σταθερά σε x , καθώς παρατηρήθηκε ότι οι απώλειες εκπαίδευσης και επικύρωσης όλων των μοντέλων είχαν σταθεροποιηθεί και το μέγεθος δέσμης ορίστηκε στη μέγιστη τιμή που μπορούσε να χωρέσει στη διαθέσιμη μνήμη της GPU.

Όσον αφορά τον σχεδιασμό των αρχιτεκτονικών για παραμέτρους όπως τα χαρακτηριστικά ανά επίπεδο, το μέγεθος χαρακτηριστικών, οι πυρήνες συνέλιξης και ο αριθμός των κεφαλών, υιοθετήθηκαν οι αρχικά προτεινόμενες τιμές από τα αντίστοιχα άρθρα και πραγματοποιήθηκαν μικρές τροποποιήσεις προκειμένου να επιτευχθεί η βέλτιστη τιμή DSC στο σύνολο επικύρωσης. Αναφορικά με τις τεχνικές κανονικοποίησης, η ομαδική κανονικοποίηση αποτέλεσε μέρος της αρχιτεκτονικής SegResNet, ενώ τα UNETR και SwinUNETR απέδωσαν καλύτερα στο σύνολο επικύρωσης με κανονικοποίηση ανά δείγμα και η κανονικοποίηση κατά παρτίδα έκανε την εκπαίδευση του U-Net πιο σταθερή. Ο βελτιστοποιητής Adam και η συνάρτηση απώλειας που αποτελείται από το άθροισμα του δείκτη DSC και της διασταυρούμενης εντροπίας, διατηρήθηκαν ίδια για όλα τα μοντέλα ώστε να εξασφαλιστεί η ομοιομορφία και η συνέπεια και να δοθεί έμφαση σε βασικές συγκρίσεις απόδοσης.

Για την αξιολόγηση των τμηματοποιήσεων χρησιμοποιήθηκαν τέσσερις βασικές μετρικές, ο συντελεστής Dice (Sorensen–Dice coefficient, DSC), η ανάκληση (recall), η ακρίβεια (precision) και η μέση συμμετρική επιφανειακή απόσταση (ASSD) [339], [340].

Συγκεκριμένα, ο Δείκτης Ομοιότητας Dice (DSC) ποσοτικοποιεί τον βαθμό ομοιότητας ή επικάλυψης μεταξύ της πραγματικής τμηματοποίησης και του αποτελέσματος που παράγεται από τον αλγόριθμο. Ο DSC υπολογίστηκε με βάση τον ακόλουθο τύπο:

$$DSC = \frac{2|A \cap B|}{|A| + |B|} \quad (5-4)$$

όπου τα A και B αντιπροσωπεύουν τα σύνολα ογκοστοιχείων των δύο τμηματοποιήσεων και τα |A|, |B| απεικονίζουν τον συνολικό αριθμό των ογκοστοιχείων στα A και B, αντίστοιχα.

Η μετρική ανάκλησης ορίζεται ως ο λόγος των σωστά αναγνωρισμένων θετικών ογκοστοιχείων προς το συνολικό αριθμό των πραγματικών θετικών περιπτώσεων και υπολογίζεται μέσω της ακόλουθης εξίσωσης:

$$Recall = \frac{TP}{TP + FN} \quad (5-5)$$

όπου τα TP και FN δηλώνουν τον αριθμό των αληθώς θετικών και ψευδώς αρνητικών αντίστοιχα.

Η ακρίβεια αντιπροσωπεύει τον λόγο των πραγματικών θετικών προς το συνολικό πλήθος των φαινομενικά θετικών φάσεων και ορίζεται ως εξής:

$$Precision = \frac{TP}{TP + FP} \quad (5-6)$$

Παρόμοια με την ανάκληση, τα TP και FP υποδηλώνουν τον αριθμό των αληθώς θετικών και ψευδώς θετικών αντίστοιχα.

Το ASSD είναι ο μέσος όρος των αποστάσεων από τα σημεία του παραγόμενου περιγράμματος προς το πραγματικό περίγραμμα και αντίστροφα. Το ASSD υπολογίζεται ως εξής:

$$ASSD(A, B) = \frac{1}{|S(A)| + |S(B)|} \sum_{s_A \in S(A)} d(s_A, S(B)) + \sum_{s_B \in S(B)} d(s_B, S(A)) \quad (5-7)$$

όπου τα S(A) και S(B) δηλώνουν το σύνολο των επιφανειακών ογκοστοιχείων των A και B αντίστοιχα, το $d(s_A, S(B))$ δηλώνει τη μικρότερη απόσταση των ογκοστοιχείων του A από το S(B) και υπολογίζεται ως εξής:

$$d(s_A, S(B)) = \min_{S_B \in S(B)} \|A - S_B\|$$

(5-8)

όπου το $\|\cdot\|$ δηλώνει την Ευκλείδεια απόσταση.

Υπερπαράμετροι	U-Net	UNETR	SwinUNETR	SegResNet
Features per layer	(16,32,64,128,256)	(16,32,64,128)	(48,96,192,384,768)	(8,16,32,64)
Number of residual connections	2	-	-	8
Conv Kernel size	3	3	3	3
Normalization	Batch	Instance	Instance	Group
Dropout rate	0.1	0.01	0.01	0.35
Learning rate	0.0001	0.0001	0.0001	0.0001
Optimizer	Adam	Adam	Adam	Adam
Feature size	0	16	24	-
Hidden size	-	768	-	-
Number of Heads	-	12	(3, 6, 12, 24)	-

Πίνακας 5.1: Υπερπαράμετροι αρχιτεκτονικών

5.2.3.1 Αρχιτεκτονική U-Net

Συνοπτικά, η δομή του U-Net αποτελείται από τον κωδικοποιητή, ο οποίος είναι υπεύθυνος για την εξαγωγή χαρακτηριστικών με πληροφορία πλαισίου και τον αποκωδικοποιητή, ο οποίος μπορεί να ανακατασκευάσει τη μάσκα τμηματοποίησης σε πλήρη ανάλυση [341]. Ο κωδικοποιητής (συμβατική διαδρομή) περιλαμβάνει συνελκτικά μπλοκς τα οποία περιέχουν συνελκτικά επίπεδα, τη συνάρτηση ενεργοποίησης ReLU και επίπεδα μέγιστης υποδειγματοληψίας, με σκοπό τη μείωση της ανάλυσης των χαρτών χαρακτηριστικών κατά 2, ενώ διπλασιάζεται ο αριθμός των χαρακτηριστικών σε κάθε στάδιο.

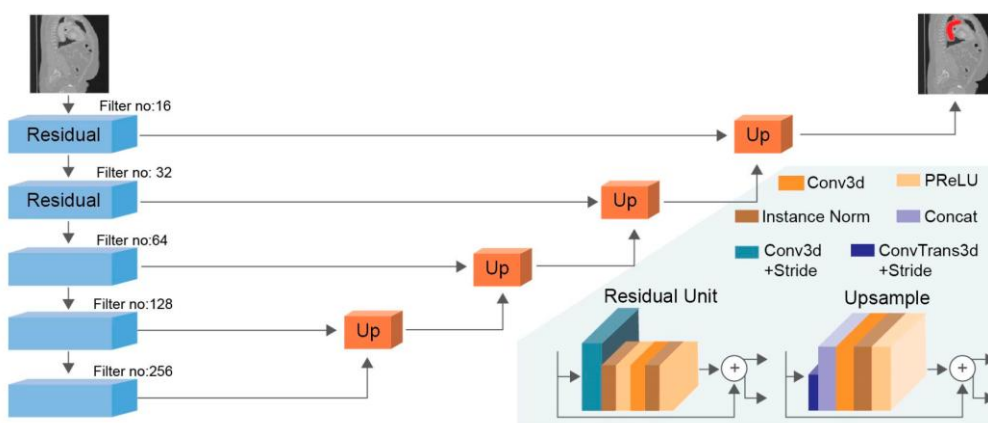
Σε αυτή την εργασία, προσθέσαμε επιπλέον υπολειμματικές συνδέσεις μέσα στα συνελκτικά μπλοκς τόσο του κωδικοποιητή όσο και του αποκωδικοποιητή, ώστε να διευκολυνθεί η διάδοση των βαθμίδων. Συγκεκριμένα, κάθε μπλοκ του κωδικοποιητή περιλάμβανε δύο συνελκτικά επίπεδα, με το πρώτο να έχει βήμα 2 για την υποδειγματοληψία

του χάρτη χαρακτηριστικών, ακολουθούμενη από την κανονικοποίηση ανα δείγμα και τη συνάρτηση ενεργοποίησης PReLU. Ο αριθμός των συνελκτικών φίλτρων για τα 5 επίπεδα του κωδικοποιητή ορίστηκε σε (16, 32, 64, 128, 256). Το τελικό επίπεδο του κωδικοποιητή ήταν το επίπεδο bottleneck μεταξύ του κωδικοποιητή και του αποκωδικοποιητή.

Σε κάθε στάδιο του αποκωδικοποιητή, οι χάρτες χαρακτηριστικών από το προηγούμενο στάδιο αυξήθηκαν κατά 2 μέσω των συνελίξεων μεταφοράς και στη συνέχεια συνενώθηκαν με τους αντίστοιχους χάρτες χαρακτηριστικών από τον κωδικοποιητή μέσω των συνδέσεων παράκαμψης. Πιο συγκεκριμένα, σε κάθε μπλοκ του αποκωδικοποιητή τα χαρακτηριστικά υπερδειγματοληψίας, συνενώθηκαν με τα χαρακτηριστικά από τις συνδέσεις παράκαμψης κατά μήκος της διάστασης των καναλιών και στη συνέχεια πέρασαν από ένα συνελκτικό επίπεδο, την κανονικοποίηση ανα δείγμα και τη συνάρτηση ενεργοποίησης PReLU.

Τέλος, η υπολειμματική σύνδεση πρόσθετε την συνενωμένη είσοδο στην έξοδο της PReLU, διαμορφώνοντας την τελική έξοδο του αντίστοιχου επιπέδου του αποκωδικοποιητή. Το τελικό συνελκτικό επίπεδο ακολουθήθηκε από μία συνάρτηση ενεργοποίησης softmax για την παραγωγή της μάσκας τμηματοποίησης σε πλήρη ανάλυση.

Ο αρχιτεκτονικός σχεδιασμός του U-Net παρουσιάζεται στην **(Εικόνα 5.1)**, ενώ οι υπερπαραμέτροί του παρατίθενται στον **(Πίνακα 5.1)**.



Εικόνα 5.1: Αρχιτεκτονική U-Net

5.2.3.2 Αρχιτεκτονική UNETR

Το μοντέλο UNETR συνδυάζει την αρχιτεκτονική τύπου U-Net, η οποία αποτελείται από έναν κωδικοποιητή για την εξαγωγή χαρακτηριστικών και έναν αποκωδικοποιητή για την ανακατασκευή της τμηματοποίησης με ένα δίκτυο μετασχηματιστή (transformer) [342]. Ο συμβατικός συνελκτικός κωδικοποιητής αντικαθίσταται από έναν μετασχηματιστή ο οποίος έχει την ικανότητα να μοντελοποιεί ευρείες χωρικές εξαρτήσεις. Οι συνδέσεις παράκαμψης

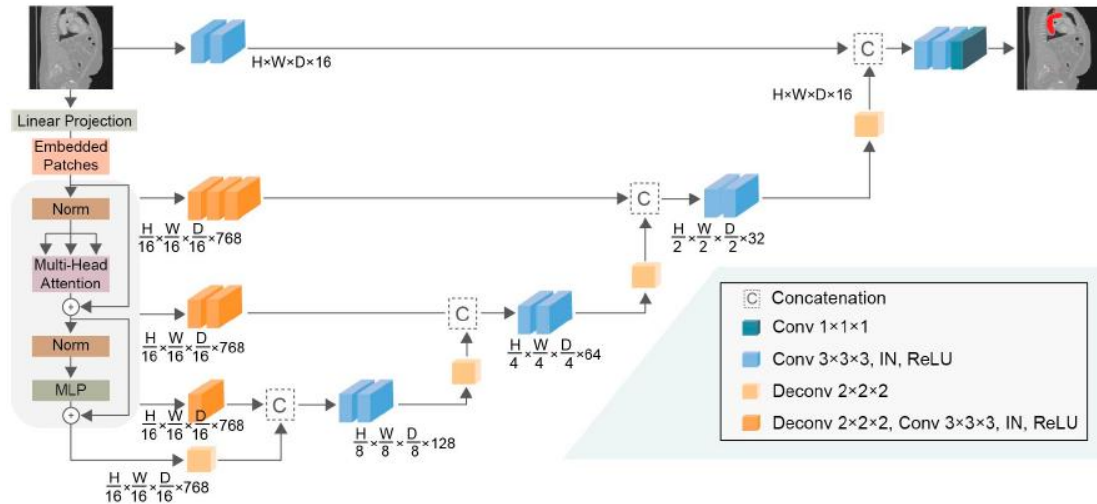
εφαρμόζονται για τη μεταφορά της πληροφορίας από τον κωδικοποιητή σε διαφορετικές αναλύσεις στον αποκωδικοποιητή.

Σε αυτή τη μελέτη, ο κωδικοποιητής του UNETR βασίστηκε σε έναν οπτικό μετασχηματιστή (Vision Transformer - ViT), ο οποίος απαιτεί τη διαίρεση της εισόδου εικόνας σε επιθέματα σταθερού μεγέθους. Αυτά τα μη επικαλυπτόμενα τρισδιάστατα επιθέματα μετατράπηκαν σε μονοδιάστατα διανύσματα x_v , τα οποία χρησιμοποιούνται ως είσοδος στον ViT. Στο επόμενο βήμα, εφαρμόστηκε ένα γραμμικό επίπεδο για να μετατρέψει τα επιθέματα σε διανύσματα μικρότερης διάστασης K . Σε αυτά τα διανύσματα προστέθηκε μία μονοδιάστατη εκμαθούμενη ενσωμάτωση θέσης, ώστε να διατηρηθούν οι αρχικές χωρικές πληροφορίες, οδηγώντας στις τελικές ενσωματώσεις z που θα εισαχθούν στον ViT.

Η δημιουργία των ενσωματώσεων ακολουθείται από μια ακολουθία μπλοκς μετασχηματιστών, που ενσωματώνουν πολλαπλές κεφαλές αυτό-προσοχής (MSA) και πολυεπίδεδους πολυπλέκτες (MLP). Στα επίπεδα 3, 6, 9 και 12 του μετασχηματιστή, εξήχθησαν αναπαραστάσεις ακολουθίας. Αυτά τα επίπεδα αντιστοιχούν σε διαφορετικές αναλύσεις στον κωδικοποιητή, αποτυπώνοντας διαφορετικά επίπεδα αφηρημένων χαρακτηριστικών. Οι αναπαραστάσεις αυτές είναι αρχικά με τη μορφή μονοδιάστατων ενσωματώσεων επιθεμάτων και στη συνέχεια αναδιαμορφώνονται σε τρισδιάστατους τανυστές της απαιτούμενης διάστασης για τον αποκωδικοποιητή. Οι αναδιαμορφωμένοι τανυστές από τον κωδικοποιητή μεταφέρονται άμεσα στον αποκωδικοποιητή μέσω των συνδέσεων παράκαμψης.

Στο επίπεδο bottleneck εφαρμόστηκε μια ανεστραμμένη συνέλιξη για την αύξηση της χωρικής ανάλυσης της εξόδου του τελευταίου επιπέδου του μετασχηματιστή κατά έναν παράγοντα 2. Στη συνέχεια, ο χάρτης χαρακτηριστικών με αλλαγμένες τις διαστάσεις, συνενώθηκε με τον προηγούμενο χάρτη εξόδου του μετασχηματιστή. Ο αποκωδικοποιητής του UNETR περιελάμβανε συνελκτικά επίπεδα καθώς και μια διαδικασία κανονικοποίησης δειγμάτων (Instance Normalization - IN), σχεδιασμένη να πραγματοποιεί σταδιακή ανοδική δειγματοληψία (up-sampling) των κωδικοποιημένων αναπαραστάσεων μέχρι την αρχική ανάλυση της εικόνας. Οι συνδέσεις παράκαμψης, συγχώνευαν χαρακτηριστικά από διαφορετικά επίπεδα του μετασχηματιστή με αντίστοιχα επίπεδα του αποκωδικοποιητή διατηρώντας τις χωρικές λεπτομέρειες. Σε κάθε στάδιο, οι αναδιαμορφωμένοι τανυστές επεξεργάζονταν από συνελκτικά επίπεδα $3 \times 3 \times 3$, ακολουθούμενα από κανονικοποίηση, διασφαλίζοντας ομαλή μετάβαση από τον χώρο ενσωματώσεων στον χώρο της εισόδου. Στο τελικό επίπεδο εξόδου, μια συνέλιξη $1 \times 1 \times 1$ παρήγαγε τον χάρτη τμηματοποίησης.

Η αρχιτεκτονική σχεδίαση του UNETR παρουσιάζεται στην **(Εικόνα 5.2)**, ενώ οι υπερπαράμετροί του αναφέρονται στον **(Πίνακας 5.1)**.



Εικόνα 5.2: Αρχιτεκτονική UNETR

5.2.3.3 Αρχιτεκτονική SwinUNETR

Η αρχιτεκτονική SwinUNETR ακολουθεί επίσης ένα σχήμα τύπου U-Net με κωδικοποιητή-αποκωδικοποιητή με τη βασική διαφορά ότι ο κωδικοποιητής βασίζεται στη μέθοδο Shifted-Windows Self-Attention [329]. Αυτός ο κωδικοποιητής αξιοποιεί το ιεραρχικό σχήμα των μετασχηματιστών Swin για να αποτυπώνει αποδοτικά τόσο τοπικές όσο και καθολικές συμφραζόμενες πληροφορίες σε τρισδιάστατες ιατρικές εικόνες. Αρχικά, η αρχική είσοδος χωρίζεται σε επιμέρους περιοχές για να δημιουργηθεί μια σειρά από τρισδιάστατα κουτιά (tokens).

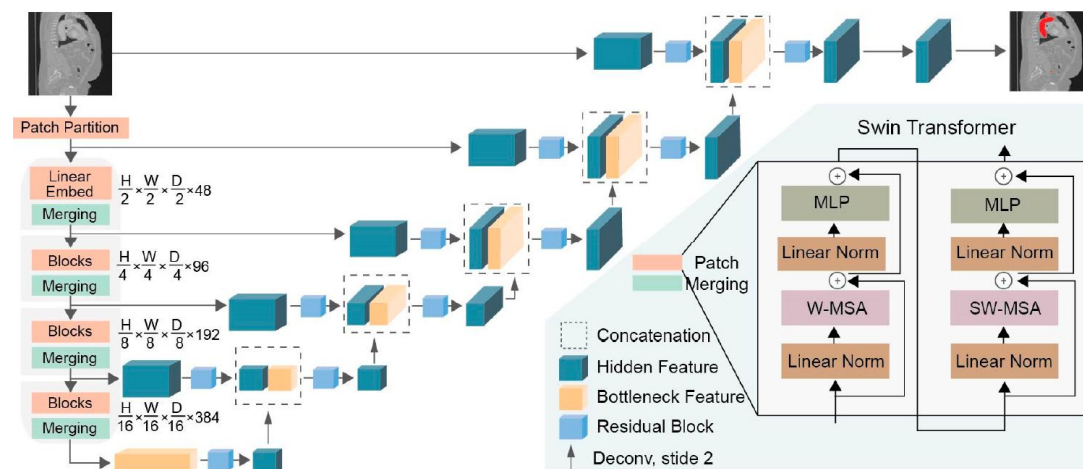
Αυτές οι περιοχές προβάλλονται γραμμικά σε διανύσματα διάστασης C . Χρησιμοποιώντας τα μετατοπισμένα παράθυρα (shifted windows), δημιουργούνται μη επικαλυπτόμενες περιοχές που χρησιμοποιούνται στα επίπεδα αυτο-προσοχής (self-attention). Ένα τρισδιάστατο παράθυρο με μέγεθος M σε κάθε διάσταση χωρίζει τα αρχικά τρισδιάστατα κουτιά σε μικρότερες περιοχές.

Στο επόμενο επίπεδο, η διαμέριση μετακινείται κατά $(\frac{M}{2}, \frac{M}{2}, \frac{M}{2})$ ογκοστοιχεία, οδηγώντας στη δημιουργία νέων επιθεμάτων που χρησιμοποιούνται ως κουτιά για τον μηχανισμό αυτό-προσοχής. Ο κωδικοποιητής του μετασχηματιστή Swin εξάγει χαρακτηριστικά σε πέντε διαφορετικές αναλύσεις αξιοποιώντας μετατοπισμένα παράθυρα για τον υπολογισμό της αυτό-προσοχής και συνδέεται με έναν αποκωδικοποιητή βασισμένο σε FCNN (πλήρως συνελκτικό νευρωνικό δίκτυο) σε κάθε ανάλυση μέσω συνδέσεων παράκαμψης. Στην παρούσα μελέτη, ο κωδικοποιητής περιλαμβάνει 4 στάδια με 2 μετασχηματιστικά μπλοκς το καθένα. Στην αρχή, εφαρμόζεται ένα επίπεδο διαμέρισης (όπου το μέγεθος επιθέματος ορίστηκε σε $2 \times 2 \times 2$ με διάσταση χαρακτηριστικών 32) για τη μετατροπή των εισερχόμενων μπλοκς σε μια ακολουθία ενσωματώσεων χαρακτηριστικών.

Σε κάθε στάδιο, αυτά τα μπλοκς χρησιμοποιούν έναν συνδυασμό των μηχανισμών Window-based Multi-head Self-Attention (W-MSA) και Shifted Window Multi-head Self-Attention (SW-MSA) για τον υπολογισμό της αυτό-προσοχής εντός μη επικαλυπτόμενων παραθύρων. Καθώς οι χάρτες χαρακτηριστικών μετακινούνται μέσα από τα στάδια του κωδικοποιητή, μια λειτουργία συγχώνευσης επιθεμάτων μειώνει τη χωρική ανάλυση κατά συντελεστή 2, ενώ αυξάνει τη διάσταση χαρακτηριστικών δημιουργώντας έτσι πολλαπλές κλίμακες αναπαράστασης χαρακτηριστικών.

Ο αποκωδικοποιητής περιλαμβάνει πολλαπλά υπολειμματικά μπλοκς σε κάθε στάδιο, όπου κάθε μπλοκ περιέχει δύο συνελκτικά επίπεδα (με πυρήνα μεγέθους 3), κανονικοποίηση ανά παράδειγμα (instance normalization) και ενεργοποίηση ReLU με σκοπό την ενίσχυση των χαρακτηριστικών. Μετά το υπολειμματικό μπλοκ, οι χάρτες χαρακτηριστικών υποβάλλονται σε υπερδειγματοληψία κατά συντελεστή 2 μέσω ενός επιπέδου αποσυνέλιξης.

Οι υπερδειγματοληπτικοί χάρτες χαρακτηριστικών έπειτα συγχωνεύονται με τις συνδέσεις παράκαμψης από τον κωδικοποιητή. Οι τελικοί χάρτες τμηματοποίησης παράγονται από ένα συνελκτικό επίπεδο με πυρήνα μεγέθους 1 και λειτουργία ενεργοποίησης softmax. Ο αρχιτεκτονικός σχεδιασμός του SwinUNETR παρουσιάζεται στην (Εικόνα 5.3), ενώ οι υπερπαραμέτροι του αναφέρονται στον (Πίνακας 5.1).



Εικόνα 5.3: Αρχιτεκτονική SwinUNETR

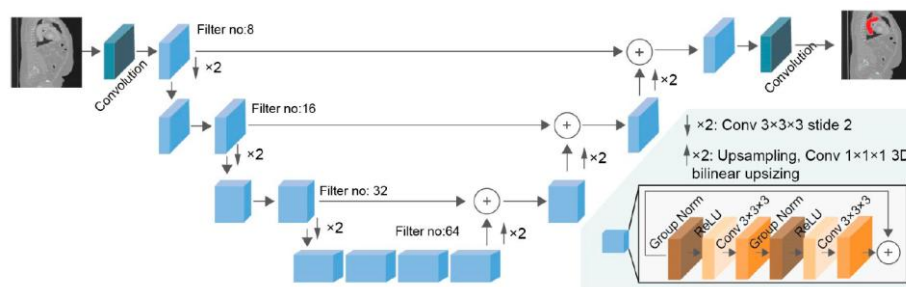
5.2.3.4 Αρχιτεκτονική SegResNet

Η αρχιτεκτονική SegResNet ακολουθεί τη δομή του U-Net, αποτελούμενη από τον κωδικοποιητή, το επίπεδο bottleneck και τον αποκωδικοποιητή [343]. Το τμήμα του παραμετρικού αυτόματου-κωδικοποιητή του αποκωδικοποιητή, παραλείπεται για τη μείωση

του υπολογιστικού κόστους [344]. Η διαδρομή του κωδικοποιητή αποτελείται από συνελκτικά ResNet μπλοκ τα οποία περιλαμβάνουν υπολειμματικές συνδέσεις προκειμένου να μετριαστεί το πρόβλημα της εξαφανιζόμενης κλίσης. Κάθε μπλοκ του κωδικοποιητή αποτελείται από δύο συνελκτικά μπλοκ τα οποία περιλαμβάνουν ομαδοποιημένη κανονικοποίηση, ενεργοποίηση ReLU και 3Δ συνέλιξη με μία παράκαμψη σύνδεσης για τη διευκόλυνση της διάδοσης των κλίσεων.

Σε αυτή τη μελέτη, στη διαδρομή του κωδικοποιητή, κάθε μπλοκ διπλασιάζει τον αριθμό των φίλτρων και υποδειγματοληπτεί τους χάρτες χαρακτηριστικών εισόδου κατά 2 χρησιμοποιώντας συνέλιξεις με βήμα 2 προκειμένου να καταγραφεί το περιεχόμενο της εικόνας σε διαφορετικά επίπεδα αφαίρεσης. Ένα επίπεδο bottleneck που αποτελείται από 256 φίλτρα βρίσκεται μεταξύ του κωδικοποιητή και του αποκωδικοποιητή. Η μονάδα του αποκωδικοποιητή ανακατασκευάζει τη μάσκα τμηματοποίησης αυξάνοντας σταδιακά την κλίμακα της αναπαράστασης των χαρακτηριστικών από τα διάφορα επίπεδα του κωδικοποιητή.

Σε κάθε μπλοκ αποκωδικοποιητή οι χάρτες χαρακτηριστικών από το προηγούμενο επίπεδο του αποκωδικοποιητή (οι οποίοι υποβλήθηκαν σε υπερδειγματοληψία κατά 2 μέσω ανάστροφης συνέλιξης) προστίθενται στους χάρτες χαρακτηριστικών από την αντίστοιχη σύνδεση παράκαμψης και εισάγονται σε συνέλιξεις $1 \times 1 \times 1$ για να βελτιωθεί ο αριθμός των χαρακτηριστικών. Σε αντίθεση με τα μπλοκ του κωδικοποιητή, τα μπλοκ του αποκωδικοποιητή περιλαμβάνουν ένα μόνο συνελκτικό επίπεδο ανά μπλοκ. Το τελευταίο επίπεδο εφαρμόζει μια λειτουργία ενεργοποίησης softmax με 2 κανάλια για την τελική μάσκα τμηματοποίησης. Ο αρχιτεκτονικός σχεδιασμός του SegResNet παρουσιάζεται στην (Εικόνα 5.4), ενώ οι υπερπαραμέτροί του παρατίθενται στον (Πίνακα 5.1).



Εικόνα 5.4: Αρχιτεκτονική SegResNet

5.2.4 Υλοποίηση και Υπολογιστικό Κόστος

Τα απαιτούμενα πειράματα πραγματοποιήθηκαν σε έναν επιτραπέζιο υπολογιστή εξοπλισμένο με κάρτα γραφικών Nvidia 4090 η οποία διαθέτει 24GB μνήμης GPU. Όλες οι συγκριτικές μέθοδοι βαθιάς μάθησης υλοποιήθηκαν χρησιμοποιώντας την έκδοση

PyTorch 2.1.0 και τη βιβλιοθήκη MONAI [39]. Ωστόσο, το υπολογιστικό κόστος κάθε ενός από τα τέσσερα διαφορετικά μοντέλα βαθιάς μάθησης διέφερε σημαντικά λόγω του διαφορετικού αριθμού παραμέτρων (ο οποίος αντιπροσωπεύει το άθροισμα όλων των βαρών και μεταβλητών μεροληψίας εντός των επιπέδων του δικτύου). Ο αριθμός αυτός αποτελεί κρίσιμο παράγοντα που επηρεάζει την πολυπλοκότητα ενός μοντέλου, τις απαιτήσεις σε μνήμη και τις απαιτήσεις υπολογιστικών πόρων. Στον (Πίνακας 5.2) παρουσιάζεται ο συνολικός αριθμός παραμέτρων για κάθε μοντέλο (ο οποίος εκτιμήθηκε με ενσωματωμένες συναρτήσεις του PyTorch για τον υπολογισμό παραμέτρων ανά επίπεδο και συνολικά για το μοντέλο).

Μοντέλο	Παράμετροι
U-Net	4.808.917
UNETR	92.667.106
SwinUNETR	62.186.708
SegResNet	1.186.994

Πίνακας 5.2: Αριθμός συνολικών παραμέτρων για κάθε μοντέλο

5.2.5 3Δ Ανακατασκευή Πλέγματος

Ακολουθώντας τις μεθόδους τμηματοποίησης με μηχανική μάθηση, η οριοθετημένη κοιλιακή αορτή αξιολογήθηκε περαιτέρω μέσω μοντελοποίησης επιφανειακής γεωμετρίας. Στο πλαίσιο αυτό χρησιμοποιήθηκε ο αλγόριθμος marching cubes [345] για τη δημιουργία ανακατασκευασμένων επιφανειακών γεωμετριών που χρησίμευσαν ως 3Δ επιφανειακό μοντέλο τόσο για τους σκοπούς οπτικοποίησης όσο και για την επεξεργασία πλέγματος.

Συνοπτικά, ο αλγόριθμος marching cubes είναι μια ευρέως χρησιμοποιούμενη τεχνική για την εξαγωγή πολυγωνικού πλέγματος από ένα τρισδιάστατο πεδίο βαθμωτών τιμών (το οποίο σε αυτή τη μελέτη χρησιμοποιήθηκε για τη δημιουργία επιφανειακού πλέγματος από ογκομετρικά δεδομένα της εξαγόμενης μάσκας τμηματοποίησης). Συγκεκριμένα, το τρισδιάστατο πεδίο βαθμωτών τιμών χωρίστηκε σε μικρότερους κύβους με τις βαθμωτές τιμές να υπολογίζονται σε κάθε κορυφή για να καταστεί δυνατή η παρεμβολή και να διευκολυνθεί η δημιουργία τριγώνων που προσεγγίζουν με ακρίβεια την ισοεπιφάνεια εντός κάθε κύβου.

Στη συνέχεια, τα τρίγωνα από τους γειτονικούς κύβους συγχωνεύτηκαν για να σχηματίσουν ένα ομαλό συνεχές επιφανειακό πλέγμα ενώ εφαρμόστηκε ένα Laplacian φίλτρο εξομάλυνσης εικόνας στην αρχικά εξαγόμενη γεωμετρία για την περαιτέρω μείωση της τραχύτητας της επιφάνειας του εξαγόμενου πλέγματος [346]. Οι παράμετροι εξομάλυνσης του φίλτρου Laplace, ρυθμίστηκαν με έναν συντελεστή εξομάλυνσης λ ίσο με 0.5 και συνολικά 10 επαναλήψεις.

Για την αξιολόγηση της 3Δ ανακατασκευής πλέγματος, εφαρμόστηκαν μικρές χειροκίνητες διορθώσεις στο παραγόμενο πλέγμα για την εξάλειψη των ιδιομορφιών, των αυτοτομών και εκφυλισμένων στοιχείων εξασφαλίζοντας αδιαπερατότητα στα ρευστά. Για τις διαδικασίες δημιουργίας και εξομάλυνσης του πλέγματος, χρησιμοποιήθηκε η βιβλιοθήκη Trimesh σε Python 3.7+ [347]. Η τελική 3Δ ανακατασκευή της περιοχής της κοιλιακής αορτής βασίστηκε στο Visualization Toolkit (VTK) [303] με χρήση της OpenGL 3Δ απόδοσης.

5.3 Αποτελέσματα

5.3.1 Ποσοτική Αξιολόγηση

Για την ποσοτική αξιολόγηση των αποτελεσμάτων, η απόδοση αξιολογήθηκε συγκρίνοντας την εκτιμώμενη περιοχή ενδιαφέροντος με την πραγματική. Οι μετρικές που χρησιμοποιήθηκαν περιλάμβαναν τον συντελεστή ομοιότητας Dice (Dice Similarity Coefficient – DSC), την ανάκληση (Recall), την ακρίβεια (Precision) και την μέση συμμετρική απόσταση επιφάνειας (Average Symmetric Surface Distance – ASSD). Για την ASSD, η βέλτιστη απόδοση αντιπροσωπεύεται από χαμηλότερες τιμές καθώς μετρά τη διαφορά μεταξύ της εκτιμώμενης τμηματοποίησης και της πραγματικής. Αντίθετα, για όλες τις άλλες μετρικές, υψηλότερες τιμές υποδηλώνουν καλύτερη απόδοση. Τα αποτελέσματα των διαφορετικών αρχιτεκτονικών βαθιάς μάθησης παρουσιάζονται παρακάτω στον (Πίνακας 5.3) (για το δημόσιο σύνολο δεδομένων) και στον (Πίνακας 5.4) (για το ιδιωτικό σύνολο δεδομένων).

Μοντέλο	DSC*	Recall*	Precision*	ASSD*
U-Net	0.89 ± 0.05	0.90 ± 0.06	0.89 ± 0.05	0.08 ± 0.04
UNETR	0.75 ± 0.16	0.77 ± 0.12	0.74 ± 0.20	0.10 ± 0.06
SwinUNETR	0.88 ± 0.08	0.87 ± 0.08	0.90 ± 0.09	0.08 ± 0.04
SegResNet	0.88 ± 0.08	0.85 ± 0.10	0.91 ± 0.07	0.08 ± 0.04

*Η τυπική απόκλιση κάθε μετρικής παρουσιάζεται δίπλα σε κάθε μέτρηση μετά το σύμβολο ±

Πίνακας 5.3: Απόδοση των διαφόρων μοντέλων βαθιάς μάθησης στο δημόσιο σύνολο δεδομένων

Όσον αφορά την απόδοση του μοντέλου κατά τη χρήση του δημόσιου συνόλου δεδομένων, το U-Net πέτυχε τη συνολικά υψηλότερη τιμή DSC με 0.89 σε σύγκριση με τις άλλες τρεις αρχιτεκτονικές. Επιπλέον, οι μετρικές Recall και Precision για τη μέθοδο U-Net ήταν επίσης υψηλές φτάνοντας τα 0.90 και 0.89 αντίστοιχα. Ενδιαφέρον παρουσιάζει το γεγονός ότι τα U-Net, SwinUNETR και SegResNet είχαν παρόμοια απόδοση ως προς την ASSD (0.08), ελαφρώς καλύτερη από τη μέθοδο UNETR (0.10).

Μοντέλο	DSC*	Recall*	Precision*	ASSD*
U-Net	0.89 ± 0.07	0.89 ± 0.10	0.89 ± 0.05	0.04 ± 0.02
UNETR	0.80 ± 0.13	0.84 ± 0.11	0.77 ± 0.17	0.05 ± 0.02
SwinUNETR	0.85 ± 0.09	0.86 ± 0.12	0.85 ± 0.08	0.05 ± 0.02
SegResNet	0.81 ± 0.13	0.75 ± 0.18	0.92 ± 0.04	0.06 ± 0.02

*Η τυπική απόκλιση κάθε μετρικής παρουσιάζεται δίπλα σε κάθε μέτρηση μετά το σύμβολο $\pm$

Πίνακας 5.4: Απόδοση των διαφόρων μοντέλων βαθιάς μάθησης στο ιδιωτικό σύνολο δεδομένων

Παρόμοια με το δημόσιο σύνολο δεδομένων, κατά τη χρήση του ιδιωτικού συνόλου δεδομένων επιτεύχθηκε συνολικά υψηλότερη απόδοση με την αρχιτεκτονική U-Net. Αυτό ισχύει για τις τιμές DSC, Recall και ASSD (0.89, 0.89 και 0.04 αντίστοιχα) σε σύγκριση με τα υπόλοιπα τρία μοντέλα. Αντίθετα, το SegResNet παρουσίασε συνολικά την υψηλότερη τιμή Precision (0.92) σε σύγκριση με τα U-Net, UNETR και SwinUNETR (0.89, 0.77 και 0.85 αντίστοιχα).

Για περαιτέρω αξιολόγηση της απόδοσης των μοντέλων τμηματοποίησης, διεξάχθηκε ένα ζεύγος t-test χρησιμοποιώντας τον συντελεστή Dice ως μετρική αξιολόγησης μεταξύ των εικόνων. Αυτή η δοκιμή λαμβάνει υπόψη τη συζευγμένη φύση των δεδομένων, μειώνοντας τη μεταβλητότητα λόγω ατομικών διαφορών και εστιάζοντας αποκλειστικά στη διακύμανση της απόδοσης των μοντέλων. Η ανάλυση πραγματοποιήθηκε με κατώφλι σημαντικότητας $p = 0.05$, αξιολογώντας αν οι παρατηρούμενες διαφορές στις τιμές Dice μεταξύ ζευγών μοντέλων ήταν στατιστικά σημαντικές (**Πίνακας 5.5**).

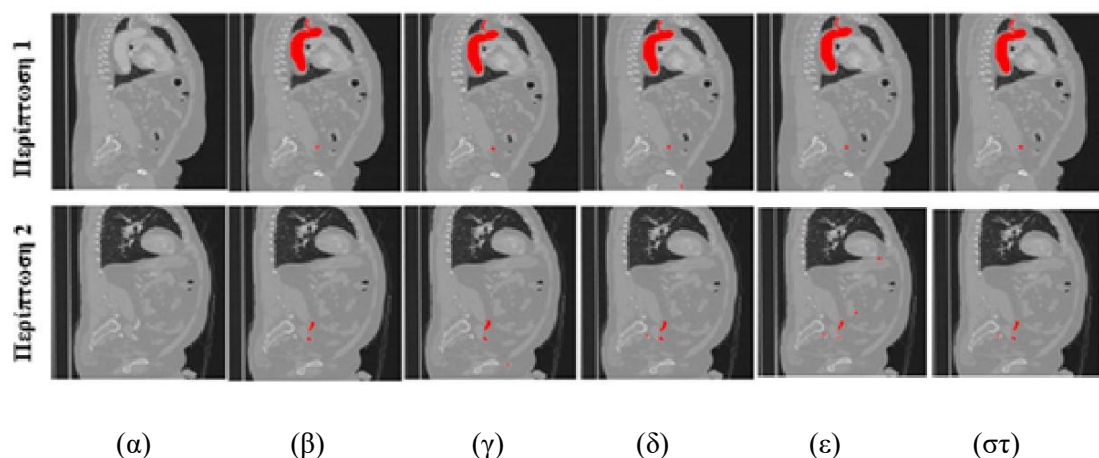
Μοντέλα	U-Net	UNETR	SwinUNETR	SegResNet
U-Net	-	5.01×10^{-5}	0.0318	0.0566
UNETR	5.01×10^{-5}	-	4.5×10^{-5}	0.0271
SwinUNETR	0.0318	4.5×10^{-5}	-	0.3881
SegResNet	0.0566	0.0271	0.3881	-

Πίνακας 5.5: Στατιστική αξιολόγηση της απόδοσης τμηματοποίησης

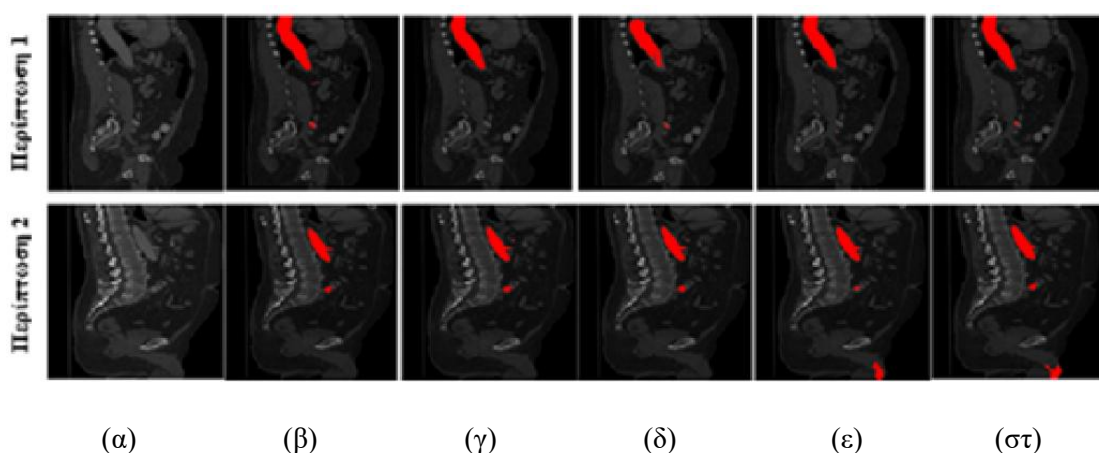
5.3.2 Ποιοτική Αξιολόγηση

Για τη ποιοτική αξιολόγηση των τεσσάρων αρχιτεκτονικών βαθιάς μάθησης, η ανιχνευθείσα δομή της αορτής, επικαλύφθηκε στην αρχική εικόνα σε κάθε περίπτωση. Όπως αναφέρθηκε στις προηγούμενες ενότητες, το αρχικό πρότυπο ήταν μια μάσκα της πραγματικής κοιλιακής αορτής η οποία ορίστηκε χειροκίνητα από ειδικούς (χρησιμοποιώντας την εφαρμογή 3D Slicer). Αντιπροσωπευτικά παραδείγματα παρουσιάζονται στις (**Εικόνα 5.5**) και (**Εικόνα 5.6**)

(για τα δημόσια και ιδιωτικά σύνολα δεδομένων αντίστοιχα), όπου η τμηματοποιημένη περιοχή ενδιαφέροντος επισημαίνεται με κόκκινο χρώμα για σκοπούς απεικόνισης.



Εικόνα 5.5: Εντοπισμένες περιοχές ενδιαφέροντος (αορτή), συγχωνευμένες στις αρχικές τομές των απεικονιστικών δεδομένων για δύο κλινικές περιπτώσεις (σειρές 1 και 2) του δημόσιου συνόλου δεδομένων για: (α) Αρχική εικόνα, (β) Πραγματική περιοχή ενδιαφέροντος, (γ) Μοντέλο U-Net, (δ) Μοντέλο UNETR, (ε) Μοντέλο SwinUNETR, (στ) SegResNet.



Εικόνα 5.6: Εντοπισμένες περιοχές ενδιαφέροντος (αορτή), συγχωνευμένες στις αρχικές τομές των απεικονιστικών δεδομένων για δύο κλινικές περιπτώσεις (σειρές 1 και 2) του ιδιωτικού συνόλου δεδομένων για: (α) Αρχική εικόνα, (β) Πραγματική περιοχή ενδιαφέροντος, (γ) Μοντέλο U-Net, (δ) Μοντέλο UNETR, (ε) Μοντέλο SwinUNETR, (στ) SegResNet.

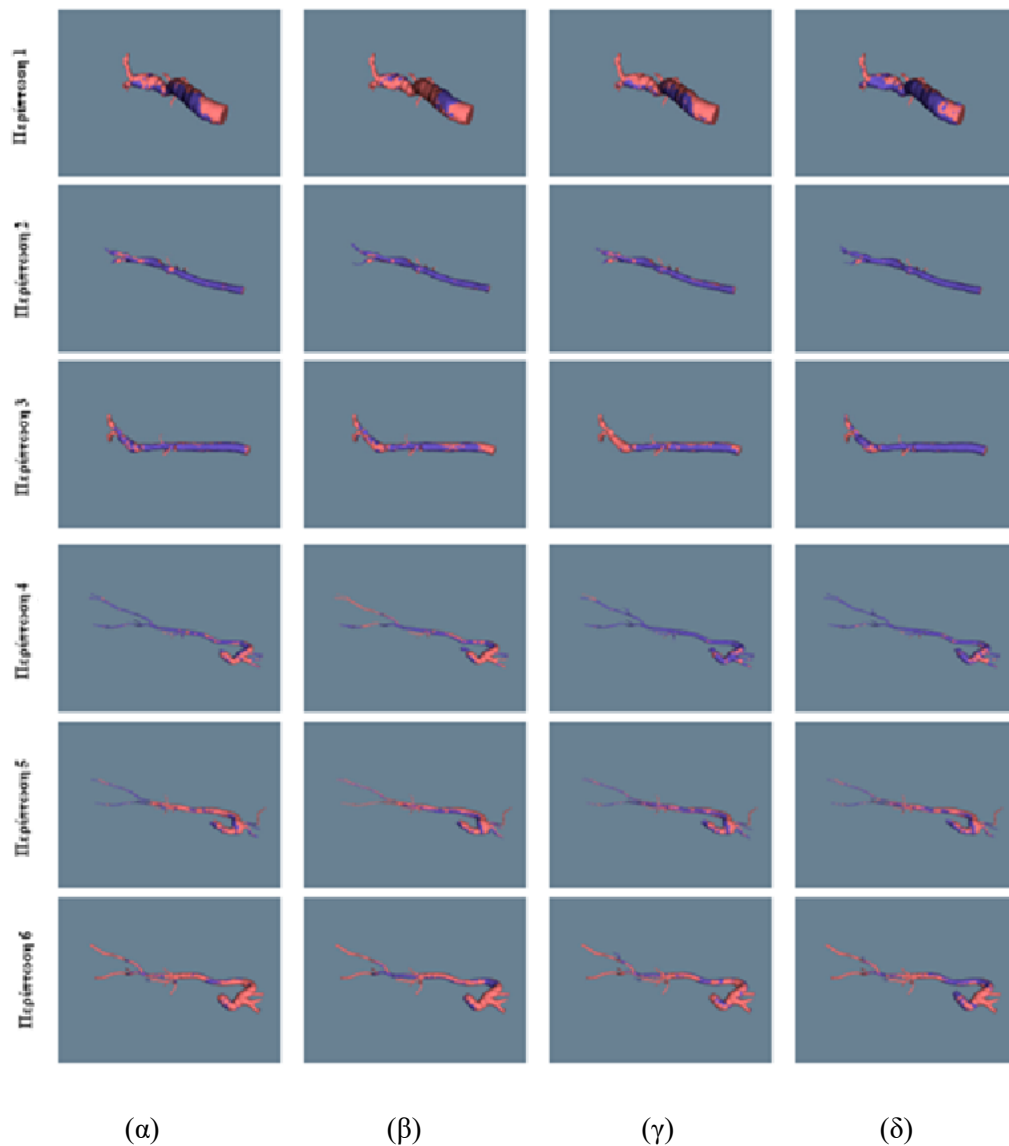
Όπως απεικονίζεται στην (**Εικόνα 5.5**), η αρχιτεκτονική U-Net παρουσιάζει αισθητά υψηλότερη ακρίβεια σε σύγκριση με τα άλλα τρία μοντέλα βαθιάς μάθησης. Επιπλέον, στη δεύτερη περίπτωση (**Εικόνα 5.5**, κάτω σειρά), το U-Net ανιχνεύει με επιτυχία τις περισσότερες από τις μικρότερες περιοχές μέσα στη μάσκα της κοιλιακής αορτής ενώ τα υπόλοιπα μοντέλα αποκλίνουν από την πραγματική τμηματοποίηση.

Στην αξιολόγηση του ιδιωτικού συνόλου δεδομένων (**Εικόνα 5.6**) και οι τέσσερις συγκρινόμενες αρχιτεκτονικές βαθιάς μάθησης πέτυχαν επαρκώς ακριβή μάσκα της κοιλιακής αορτής αν και σε ορισμένες περιπτώσεις παρατηρήθηκαν υπερ-τμηματοποιήσεις ή υπο-τμηματοποιήσεις. Συγκεκριμένα, στην πρώτη περίπτωση (επάνω σειρά της **Εικόνα 5.6**), το SegResNet παρήγαγε την πιο ομαλή μάσκα της κοιλιακής αορτής σε σύγκριση με τα U-Net,

UNETR και SwinUNETR (υπο-τμηματοποίηση). Ομοίως, στη δεύτερη περίπτωση οι αρχιτεκτονικές U-Net και UNETR παρουσίασαν καλύτερη ακρίβεια σε σύγκριση με τα SwinUNETR και SegResNet (υπερ-τμηματοποίηση).

5.3.3 Αξιολόγηση 3Δ Ανακατασκευής

Για την περαιτέρω αξιολόγηση της απόδοσης του μοντέλου, η 3Δ ανακατασκευή εξετάστηκε από κλινικούς ιατρούς προκειμένου να μελετηθεί η περιοχή της κοιλιακής αορτής από διάφορες γωνίες και οπτικές κάτι που δεν είναι εφικτό με τη δισδιάστατη ιατρική απεικόνιση. Τα 3Δ μοντέλα ανακατασκευής ενδεικτικών περιστατικών μαζί με τα πλέγματα των τεσσάρων αρχιτεκτονικών βαθιάς μάθησης παρουσιάζονται στην (Εικόνα 5.7).



Εικόνα 5.7: 3Δ συγχωνευμένα μοντέλα της εκτιμώμενης αορτής (μπλε), επικαλυπτόμενα στην πραγματική περιοχή (κοραλλί) για έξι περιπτώσεις, χρησιμοποιώντας: (α) U-Net, (β) UNETR, (γ) SwinUNETR και (δ) SegResNet.

Όπως φαίνεται στην **(Εικόνα 5.7)**, το συγχωνευμένο μοντέλο από τον συνδυασμό της πραγματικής περιοχής της αορτής και των πλεγμάτων U-Net πέτυχε τη μεγαλύτερη ακρίβεια. Σε σύγκριση, η συγχώνευση 3Δ μοντέλων που παράγεται από τα UNETR **(Εικόνα 5.7 (β))**, SwinUNETR **(Εικόνα 5.7 (γ))** και SegResNet **(Εικόνα 5.7 (δ))**, παρουσίασε διαφορετικού βαθμού υπο-τμηματοποίησης όταν συγκρίθηκαν με την πραγματική περιοχή της κοιλιακής αορτής. Επιπλέον, ο συνδυασμός του αλγορίθμου Marching Cubes με τη συμπληρωματική προσέγγιση πλέγματος και τις τοπικές διορθώσεις απέδειξε την ανώτερη ποιότητα των πλεγμάτων με αδιαπερατότητα σε ροή.

5.4 Συζήτηση

Σε αυτή τη μελέτη, υλοποιήσαμε τέσσερις διαφορετικές αρχιτεκτονικές βαθιάς μάθησης για την αυτόματη τμηματοποίηση της κοιλιακής αορτής αξιοποιώντας διαφορετικά σύνολα δεδομένων για μια αξιόπιστη συγκριτική ανάλυση. Συγκεκριμένα, οι τέσσερις μέθοδοι αξιολογήθηκαν χρησιμοποιώντας έναν μεγάλο όγκο τρισδιάστατων δεδομένων αξονικής τομογραφίας (CT) που προέρχονταν από δημόσια και ιδιωτικά σύνολα δεδομένων ακολουθώντας μια πλήρη ποιοτική και ποσοτική αξιολόγηση. Τα ευρήματά μας δείχνουν ότι το μοντέλο U-Net υπερίσχυσε σταθερά των SegResNet, UNETR και SwinUNETR στις περισσότερες περιπτώσεις τόσο ως προς τους βασικούς δείκτες απόδοσης (που χρησιμοποιήθηκαν στη μελέτη) όσο και από ποιοτική άποψη.

Συγκεκριμένα, όσον αφορά την ποσοτική απόδοση των μοντέλων βαθιάς μάθησης, το U-Net υπερέχει όλων των άλλων αρχιτεκτονικών βάσει του δείκτη DSC σε όρους ακρίβειας τμηματοποίησης. Αυτό ήταν εμφανές τόσο στα δεδομένα υγιών ατόμων (δημόσια) όσο και στα δεδομένα ασθενών (ιδιωτικά) όπως φαίνεται στους **(Πίνακας 5.3)** και **(Πίνακας 5.4)**. Ενδιαφέρον παρουσιάζει το γεγονός ότι παρόλο που τα μοντέλα βαθιάς μάθησης δεν είχαν εκτεθεί σε παραλλαγές του AAA κατά την εκπαίδευση, η απόδοσή τους στην ολική αποτύπωση της αορτής ήταν συγκρίσιμη με εκείνη στα δεδομένα υγιών ατόμων. Η στατιστική αξιολόγηση των τιμών DSC **(Πίνακας 5.5)** ανέδειξε επίσης σημαντική βελτίωση της απόδοσης μεταξύ του U-Net και των UNETR/Swin UNETR ($p < 0.05$), επιβεβαιώνοντας ότι το U-Net παρουσίασε σταθερά καλύτερη απόδοση από ορισμένα μοντέλα βασισμένα σε μετασχηματιστές. Αντιθέτως, οι στατιστικές διαφορές μεταξύ του U-Net και του SegResNet δεν ήταν σημαντικές ($p = 0.0566$), όπως επίσης και μεταξύ SegResNet και Swin UNETR ($p = 0.3881$). Αξίζει να σημειωθεί ότι παρόλο που η καλύτερη απόδοση στην τμηματοποίηση εκτιμήθηκε υπό τους περιορισμούς περιορισμένου αριθμού χαραγμένων εικόνων, τα λεπτότερα ή πιο ανατομικά περίπλοκα αγγειακά παρακλάδια της αορτής δεν αποτυπώθηκαν πλήρως. Αυτό μπορεί να αποδοθεί στην περιορισμένη ικανότητα των CNN να καταγράφουν εξαρτήσεις μεγάλου εύρους οι οποίες είναι κρίσιμες για την ακριβή αναπαράσταση της πλήρους ανατομικής

πολυπλοκότητας της αορτής [348]. Από την άλλη πλευρά, η βέλτιστη ακρίβεια επιτεύχθηκε από την αρχιτεκτονική SegResNet. Ωστόσο, η απόκλιση μεταξύ ακρίβειας και ανάκλησης η οποία ήταν ιδιαίτερα αισθητή στο δημόσιο σύνολο δεδομένων υποδηλώνει μια τάση του μοντέλου SegResNet για υπο-τμηματοποίηση, γεγονός που οδηγεί στη μη ανίχνευση μικρών ή λεπτών αγγείων [349].

Μια ακόμη σημαντική παράμετρος που πρέπει να ληφθεί υπόψη είναι το υπολογιστικό κόστος κάθε μοντέλου σε σύγκριση με την απόδοση των αλγορίθμων. Με βάση αυτό το σκεπτικό και οι τέσσερις αρχιτεκτονικές βασίζονται σε μια ποικιλία παραμέτρων που ρυθμίζουν βασικές πτυχές των αλγορίθμων τους (**Πίνακας 5.2**). Από αυτή την άποψη, το SegResNet ήταν η πιο ελαφριά αρχιτεκτονική που χρησιμοποιήθηκε, με περίπου 1.18 εκατομμύρια παραμέτρους, ακολουθούμενο από το U-Net με περίπου 4.8 εκατομμύρια παραμέτρους. Αντίθετα, τα μοντέλα που βασίζονται σε transformers, όπως το UNETR και το Swin UNETR, έχουν σημαντικά μεγαλύτερο αριθμό παραμέτρων, περίπου 92.6 εκατομμύρια και 62.1 εκατομμύρια αντίστοιχα. Ως εκ τούτου, οι αυξημένες υπολογιστικές απαιτήσεις για εκπαίδευση και πρόβλεψη απαιτούν πιο ισχυρό υλικό με μεγαλύτερη μνήμη. Το αυξημένο υπολογιστικό κόστος των UNETR και Swin UNETR υπογραμμίζει τον συμβιβασμό μεταξύ πολυπλοκότητας μοντέλου και βιωσιμότητας πόρων αναδεικνύοντας την πρακτικότητα του U-Net στην παρούσα μελέτη, όπου οι υπολογιστικοί πόροι ήταν περιορισμένοι [350].

Παρά το γεγονός ότι, όποτε ήταν εφικτό χρησιμοποιήθηκαν οι ίδιες τιμές υπερπαραμέτρων για κοινά χαρακτηριστικά και στις τέσσερις μεθόδους, οι τελικές τιμές καθορίστηκαν μετά από πολλαπλές δοκιμές με εικόνες από όλα τα διαθέσιμα σύνολα δεδομένων (εξασφαλίζοντας έτσι τη βέλτιστη απόδοση) [351]. Συνεπώς, μπορεί να υποτεθεί ότι διαφορετικά (μεγαλύτερα) σύνολα δεδομένων θα επηρέαζαν τις τιμές των υπερπαραμέτρων επηρεάζοντας σημαντικά τη συνολική απόδοση κάθε μοντέλου [352]. Ομοίως, ο αριθμός των εποχών που χρησιμοποιήθηκαν για κάθε μέθοδο σχετίζεται άμεσα με τον χρόνο επεξεργασίας ο οποίος αυξάνεται όσο αυξάνεται ο αριθμός των εποχών [353].

Οι υλοποιημένες αρχιτεκτονικές U-Net φαίνεται να επιτυγχάνουν μια ισορροπία μεταξύ υπολογιστικής αποδοτικότητας και ακρίβειας στην τμηματοποίηση προσφέροντας ανταγωνιστική απόδοση χωρίς σημαντική αύξηση του χρόνου επεξεργασίας ακόμη και όταν αυξάνεται ο αριθμός των εποχών [354]. Αυτό τις καθιστά ιδιαίτερα κατάλληλες για εφαρμογές ενσωμάτωσης στην κλινική πράξη όπου η ταχύτητα και η ακρίβεια είναι κρίσιμες [355].

Στην πραγματικότητα, η ανθεκτικότητα των αρχιτεκτονικών βαθιάς μάθησης και ειδικά του U-Net στη γενίκευση από υγιείς ελέγχους σε περιπτώσεις ανευρύσματος, αναδεικνύει το δυναμικό των τεχνικών τμηματοποίησης με βάση τη βαθιά μάθηση να βελτιώσουν τη λήψη κλινικών αποφάσεων, να απλοποιήσουν τις ροές εργασίας και να ενισχύσουν τα αποτελέσματα για τους ασθενείς παρέχοντας μια βάση για την ενσωμάτωση αυτοματοποιημένων εργαλείων στις κλινικές διαδικασίες [356].

Με τη σημαντική μείωση του χρόνου και της μεταβλητότητας που συνδέονται με τη χειροκίνητη τμηματοποίηση, η υπολογιστική αποδοτικότητα και η υψηλή ακρίβεια της βαθιάς μάθησης την καθιστούν κατάλληλη για εφαρμογές απεικόνισης σε πραγματικό χρόνο, όπως η ενδοεγχειρητική 3Δ απεικόνιση ή τα φορητά διαγνωστικά συστήματα [357]. Η αυτοματοποιημένη τμηματοποίηση μπορεί επίσης να ενισχύσει τη διαγνωστική συνέπεια, να υποστηρίξει την εξατομικευμένη θεραπευτική προσέγγιση παρέχοντας αξιόπιστες μετρήσεις της μορφολογίας της αορτής και να βοηθήσει στον προσδιορισμό της καταλληλότητας του ασθενούς για επεμβάσεις όπως η τοποθέτηση ενδομοσχεύματος [358].

Αναφορικά με την ποιοτική αξιολόγηση, όλες οι αρχιτεκτονικές βαθιάς μάθησης παρουσίασαν παρόμοια απόδοση με την U-Net να είναι εμφανώς πιο ακριβής από τις άλλες τρεις αρχιτεκτονικές (**Εικόνα 5.5**) και (**Εικόνα 5.6**). Ιδιαίτερα κατά την παρατήρηση των αποτελεσμάτων της 3Δ ανακατασκευής, η U-Net ισορροπεί αποτελεσματικά μεταξύ των κινδύνων υπερ-τμηματοποίησης και υπο-τμηματοποίησης (**Εικόνα 5.7**). Παρ' όλα αυτά, όλα τα μοντέλα βαθιάς μάθησης παρουσίασαν δυσκολίες στην ανίχνευση πολύ μικρών αγγείων της αορτής. Αυτό σχετίζεται με το γεγονός ότι τα μικρότερα αγγεία συχνά εμφανίζουν υπο-τμηματοποίηση λόγω της μειωμένης επιφάνειάς τους κάτι που αποτελεί πρόκληση για τα μοντέλα τμηματοποίησης [359]. Σε αυτές τις περιπτώσεις, το μοντέλο τείνει να δίνει προτεραιότητα στον μεγαλύτερο, πιο εμφανή αυλό παραλείποντας τις λεπτομέρειες των μικρότερων αρτηριών. Το ζήτημα αυτό είναι ιδιαίτερα εμφανές σε εργασίες που αφορούν πολύπλοκες αγγειακές δομές όπου οι μικρότεροι κλάδοι είναι πιο επιρρεπείς στο να μην ανιχνεύονται από το μοντέλο [360]. Για την αντιμετώπιση αυτού του περιορισμού χρησιμοποιήσαμε τη συνάρτηση απώλειας Dice η οποία είναι αποτελεσματική στη μέτρηση της επικάλυψης μεταξύ της προβλεπόμενης τμηματοποίησης και της πραγματικής [338]. Ωστόσο, λόγω της ανισομερούς φύσης των μεγεθών των αγγείων η αποκλειστική χρήση της συνάρτησης απώλειας Dice αποδείχθηκε ανεπαρκής για την αποτύπωση των μικρότερων αρτηριών.

Για την αντιμετώπιση του προβλήματος της υπο-τμηματοποίησης, ενσωματώσαμε έναν όρο διασταυρούμενης εντροπίας στη συνάρτηση απώλειας, ο οποίος προσέφερε καλύτερη ισορροπία εστιάζοντας στην ακρίβεια ταξινόμησης ανά εικονοστοιχείο. Αυτή η συνδυαστική προσέγγιση βοήθησε στη βελτίωση της ικανότητας του μοντέλου να τμηματοποιεί τόσο τα μεγαλύτερα όσο και τα μικρότερα αγγεία οδηγώντας σε πιο ακριβή αναπαράσταση της αγγειακής δομής.

Σε μελλοντικές εργασίες σκοπεύουμε να επεκτείνουμε τη μελέτη ενσωματώνοντας μεγαλύτερα δείγματα και εξερευνώντας επιπλέον μοντέλα ώστε να διερευνήσουμε αν ο συνδυασμός της διασταυρούμενης εντροπίας με τη συνάρτηση Dice μπορεί να προσφέρει μια αποτελεσματική λύση για τη μείωση της υπο-τμηματοποίησης των μικρότερων αγγείων

(ενθαρρύνοντας το μοντέλο να δίνει μεγαλύτερη έμφαση σε αυτές τις μικρότερες δομές, διατηρώντας ταυτόχρονα την ακρίβεια τμηματοποίησης των μεγαλύτερων αγγείων).

5.5 Συμπεράσματα

Η παρούσα μελέτη πραγματοποίησε μια εκτενή σύγκριση τεσσάρων ευρέως χρησιμοποιούμενων αρχιτεκτονικών βαθιάς μάθησης (U-Net, SegResNet, UNETR και SwinUNETR) για την αυτόματη τμηματοποίηση της κοιλιακής αορτής χρησιμοποιώντας δεδομένα 3D αξονικής τομογραφίας (CT). Με τη χρήση δύο διαφορετικών συνόλων δεδομένων (ένα που περιλάμβανε υγιή άτομα και ένα που περιλάμβανε ασθενείς με ανεύρυσμα κοιλιακής αορτής - AAA), προσφέρθηκε μια ολοκληρωμένη ανάλυση της απόδοσης των μοντέλων. Η εκπαίδευση και οι δοκιμές πραγματοποιήθηκαν στο σύνολο δεδομένων με υγιή άτομα, ενώ η πρόσθετη επικύρωση έγινε στο σύνολο δεδομένων με ασθενείς AAA για την αξιολόγηση της γενικευσιμότητας των μοντέλων.

Τα αποτελέσματα έδειξαν ότι οι αρχιτεκτονικές βαθιάς μάθησης είναι ικανές να προσδιορίσουν με ακρίβεια την κοιλιακή αορτή ακόμη και όταν εκπαιδεύονται αποκλειστικά σε δεδομένα από υγιείς. Ανάμεσα στα μοντέλα που αξιολογήθηκαν, το U-Net υπερίσχυσε σταθερά έναντι των υπολοίπων με μέσο DSC 0.89 ± 0.05 στο δημόσιο σύνολο δεδομένων και 0.89 ± 0.07 στο ιδιωτικό σύνολο. Η στατιστική ανάλυση αποκάλυψε σημαντικές διαφορές ανάμεσα στο U-Net και τα μοντέλα βασισμένα σε μετασχηματιστές UNETR ($p = 5.01 \times 10^{-5}$) και SwinUNETR ($p = 0.0318$). Αντίθετα, το SegResNet πέτυχε τη μεγαλύτερη ακρίβεια (0.92 ± 0.04), αλλά η σχετικά χαμηλότερη ανάκληση (0.75 ± 0.18) ανέδειξε μια τάση για υπο-τμηματοποίηση ιδίως σε μικρότερα αγγεία.

Παρά την απουσία δεδομένων εκπαίδευσης ειδικά για AAA, όλα τα μοντέλα έδειξαν υψηλό βαθμό γενίκευσης σε περιπτώσεις με ανεύρυσμα. Οι ποιοτικές αξιολογήσεις ανέδειξαν περαιτέρω την ανώτερη απόδοση του U-Net το οποίο ισορρόπησε αποτελεσματικά την υπερ και υπο-τμηματοποίηση παράγοντας πιο ομαλές και ακριβείς 3D ανακατασκευές της αορτής. Ωστόσο, οι μικρότερες αγγειακές δομές παρουσίασαν προκλήσεις για όλα τα μοντέλα με την υπο-τμηματοποίηση να αποδίδεται στο μικρό εμβαδόν της επιφάνειάς τους.

Για την αντιμετώπιση αυτού του ζητήματος εφαρμόστηκε μια υβριδική συνάρτηση κόστους που συνδυάζει την απώλεια Dice και την διασταυρούμενη εντροπία, οδηγώντας σε βελτίωση της ακρίβειας τμηματοποίησης για αυτές τις μικρότερες δομές χωρίς να επηρεάζεται η ακρίβεια για τα μεγαλύτερα αγγεία. Πέρα από την ανώτερη απόδοσή του, το U-Net επέδειξε και υπολογιστική αποδοτικότητα με σημαντικά μικρότερο αριθμό παραμέτρων σε σχέση με τα μοντέλα βασισμένα σε μετασχηματιστές όπως τα UNETR και SwinUNETR. Αυτή η αποδοτικότητα σε συνδυασμό με την ακρίβεια υπογραμμίζει την πρακτικότητα του U-Net για

εφαρμογές σε πραγματικό χρόνο στον κλινικό τομέα όπου οι υπολογιστικοί πόροι μπορεί να είναι περιορισμένοι.

Κεφάλαιο 6: Αυτο-επιβλεπόμενη Προ-εκπαίδευση με Μάσκα Καθοδηγούμενη από την Φωτεινότητα για Βελτιωμένη Τμηματοποίηση της Αορτής σε Αξονική Τομογραφία

6.1 Εισαγωγή

Όπως προαναφέρθηκε στην Εισαγωγή του πέμπτου κεφαλαίου, το ανεύρυσμα κοιλιακής αορτής (AAA) είναι σοβαρή αγγειακή πάθηση που χαρακτηρίζεται από εντοπισμένη διάταση της κοιλιακής αορτής [316] και, αν δεν αντιμετωπιστεί, ενέχει υψηλό κίνδυνο ρήξης με σοβαρή νοσηρότητα ή θνησιμότητα [317]. Η ακριβής εκτίμηση της γεωμετρίας της αορτής είναι απαραίτητη για διαγνωστικούς σκοπούς, με την υπολογιστική τομογραφία (CT) να αποτελεί μέθοδο αναφοράς. Η τμηματοποίηση εικόνας είναι κρίσιμη για τον σαφή καθορισμό των ανατομικών δομών [286] και, ειδικά στο AAA, επιτρέπει τη μέτρηση της διαμέτρου, του πάχους τοιχώματος και του μεγέθους του ενδοαυλικού θρόμβου (ILT), παρέχοντας σημαντικές πληροφορίες για την εξέλιξη και τον κίνδυνο ρήξης.

Ιστορικά, η τμηματοποίηση στηριζόταν σε χειροκίνητη αποτύπωση, μια διαδικασία χρονοβόρα και επιρρεπή σε μεταβλητότητα μεταξύ παρατηρητών [361]. Οι ημι-αυτόματες μέθοδοι βελτίωσαν την αποδοτικότητα, αλλά εξακολουθούν να απαιτούν σημαντική ανθρώπινη παρέμβαση. Αυτό οδήγησε στην ανάπτυξη πλήρως αυτοματοποιημένων προσεγγίσεων, βασισμένων σε μηχανική και βαθιά μάθηση, οι οποίες έχουν μετασχηματίσει την ιατρική τμηματοποίηση [362]. Σήμερα, αυτοματοποιημένοι αλγόριθμοι αξιοποιούν μεγάλα χαραγμένα σύνολα δεδομένων για εκπαίδευση μοντέλων που αναγνωρίζουν πολύπλοκα πρότυπα και ανατομικές δομές [363]. Τα συνελκτικά νευρωνικά δίκτυα (CNNs), όπως το U-Net και οι παραλλαγές του χρησιμοποιούν την αρχιτεκτονική encoder–decoder (κωδικοποιητή – αποκωδικοποιητή) για ακριβή τμηματοποίηση δομών όπως αγγεία, όγκοι και όργανα [364]. Επιπλέον, αρχιτεκτονικές με transformers, όπως τα UNETR [365] και SwinUNETR [328], αξιοποιούν μηχανισμούς αυτο-προσοχής για αποδοτική μοντελοποίηση μακρινών χωρικών συσχετίσεων ακόμη και για τμηματοποίηση της αορτής [327], [366].

Οι μέθοδοι βαθιάς μάθησης χρησιμοποιούνται για την αυτόματη τμηματοποίηση οργάνων από ιατρικές εικόνες αλλά η έλλειψη μεγάλων σχολιασμένων συνόλων δεδομένων περιορίζει την εφαρμογή τους. Το πρόβλημα είναι ιδιαίτερα έντονο στην τμηματοποίηση της αορτής και του ανευρύσματος (AAA) όπου η δημιουργία ακριβών δεδομένων αναφοράς για πολύπλοκες δομές απαιτεί χρόνο και εξειδικευμένη γνώση.

Οι τεχνικές αυτο-εποπτευόμενης μάθησης (Self-Supervised Learning, SSL) στοχεύουν στην άμβλυνση αυτής της απαίτησης προ-εκπαιδύοντας τα μοντέλα σε μη χαραγμένα σύνολα δεδομένων με τη χρήση βοηθητικών διεργασιών [367]. Αυτές οι εργασίες επιτρέπουν στα

δίκτυα να μαθαίνουν ουσιαστικές και ανθεκτικές αναπαραστάσεις των δεδομένων απεικόνισης χωρίς την ανάγκη χειροκίνητων επισημάνσεων. Κάποιες συνήθειες βοηθητικές διεργασίες περιλαμβάνουν τον χρωματισμό εικόνας [368], τη μάσκα ανακατασκευής τμήματος εικόνας που λείπει [369], την επίλυση παζλ τύπου Jigsaw [370] και την πρόβλεψη περιστροφής [371]. Οι μέθοδοι αντιθετικής μάθησης όπως το SimCLR [372], εκπαιδεύουν μοντέλα να διακρίνουν μεταξύ παρόμοιων και διαφορετικών εικόνων στον χώρο χαρακτηριστικών, ενώ οι μη-αντιθετικές προσεγγίσεις, όπως τα BYOL [373], DINO [374] και MOCO [375], βελτιώνουν περαιτέρω τις αναπαραστάσεις χρησιμοποιώντας ενισχυμένες εικόνες χωρίς αρνητικά δείγματα ή αξιοποιώντας μονάδες μνήμης, αρχιτεκτονικές μαθητή-δασκάλου ή μεθόδους ορμής.

Οι μέθοδοι SSL έχουν δείξει ενθαρρυντικά αποτελέσματα σε εργασίες ιατρικής απεικόνισης συμπεριλαμβανομένης της ταξινόμησης, της τμηματοποίησης και της αποθορυβοποίησης με εφαρμογές στην CT απεικόνιση να παρουσιάζουν ιδιαίτερα θετικά αποτελέσματα [376], [377], [378], [379], [380]. Στην τμηματοποίηση της αορτής, η προ-εκπαίδευση μέσω SSL καταγράφει τις δομικές λεπτομέρειες και τις παθολογικές παραλλαγές (π.χ. ανευρυσματική διάταση, ILT), ενισχύοντας την ακρίβεια ακόμη και με περιορισμένα χαραγμένα δεδομένα.

Στην παρούσα εργασία προτείνουμε μια βοηθητική διεργασία αυτο-εποπτευόμενης μάθησης που αξιοποιεί τις εγγενείς ιδιότητες των τρισδιάστατων CT εικόνων όπου διαφορετικά εύρη εντάσεων φωτεινότητας αντιστοιχούν σε διαφορετικές ανατομικές δομές. Συγκεκριμένα, προτείνουμε τη μονάδα της μάσκας καθοδηγούμενη από την φωτεινότητα (Intensity Guided Masking - IGM), η οποία τυχαία αποκρύπτει περιοχές της εικόνας εντός συγκεκριμένων εύρων φωτεινότητας αναγκάζοντας το δίκτυο να προβλέψει αυτές τις περιοχές και να μάθει αποτελεσματικά από μη χαραγμένα δεδομένα εστιάζοντας ταυτόχρονα σε κρίσιμες ανατομικές δομές.

Επιπλέον, ο συνδυασμός της μονάδας IGM με αντιθετική μάθηση, ενίσχυση με περιστροφή και τον κωδικοποιητή Transformer ο οποίος μπορεί να μάθει μακρινές χωρικές εξαρτήσεις, παράγει πλούσιες αναπαραστάσεις ακόμη και με περιορισμένα σχολιασμένα δεδομένα. Η μεθόδός μας αξιολογείται τόσο σε δημόσια όσο και σε ιδιωτικά σύνολα δεδομένων CT με χειροκίνητη τμηματοποίηση της χαραγμένης αορτής, αποδεικνύοντας την ανθεκτικότητα και την ευελιξία της σε κλινικό και ερευνητικό πλαίσιο.

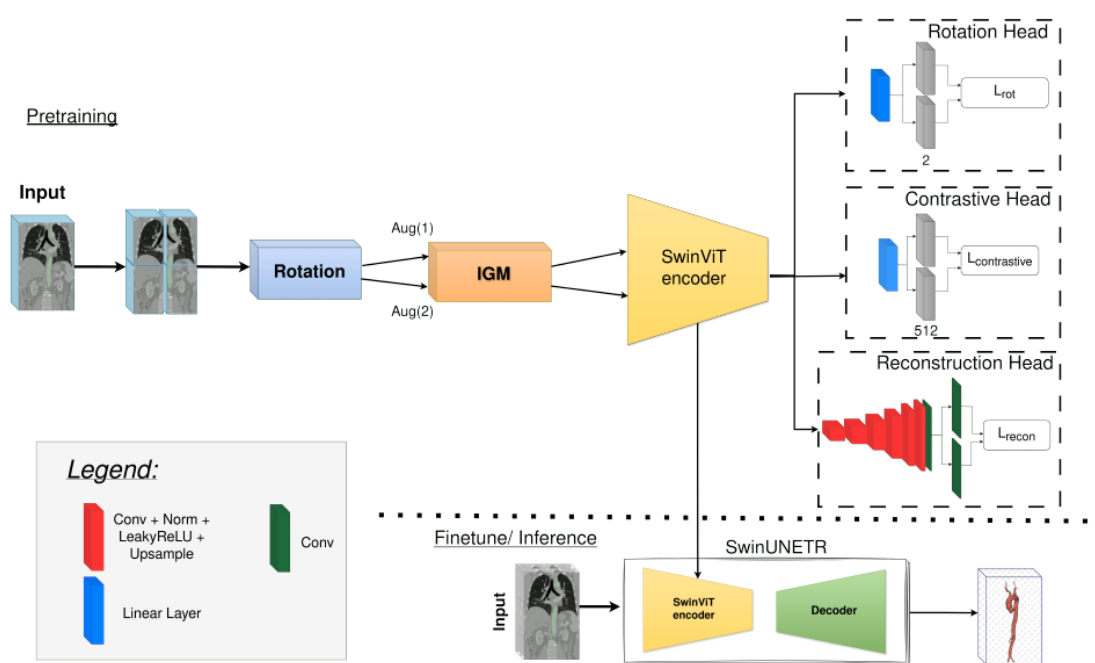
6.2 Μεθοδολογία

6.2.1 Εισαγωγή

Η βασική αρχιτεκτονική της προτεινόμενης μεθόδου παρουσιάζεται στην **(Εικόνα 6.1)**. Το ανώτερο τμήμα απεικονίζει τη στρατηγική προ-εκπαίδευσης όπου ένα αρχικό ζεύγος όγκων υφίσταται ενίσχυση βασισμένη στην περιστροφή πριν εισέλθει στη μονάδα Intensity Guided Masking (IGM) η οποία επιλέγει τυχαία ένα κέντρο και ένα παράθυρο εντάσεων και δημιουργεί

τη μάσκα όγκου. Στη συνέχεια, οι μάσκες των όγκων εισάγονται στον κωδικοποιητή SwinViT για την παραγωγή ενός διάνυσματος χαρακτηριστικών το οποίο θα χρησιμοποιηθεί στα επόμενα SSL τμήματα.

Τα τμήματα περιστροφής και αντίθεσης περιλαμβάνουν ένα γραμμικό επίπεδο που αποδίδει την κλάση περιστροφής και το διάνυσμα χαρακτηριστικών για τις αντίστοιχες απώλειες (L_{rot}, L_{cont}). Το τμήμα ανακατασκευής αποτελείται από έναν αποκωδικοποιητή με συνελκτικά μπλοκς και μονάδες υπερδειγματοληψίας για την ανακατασκευή των όγκων χωρίς μάσκα για την απώλεια L_{recon} . Ο κωδικοποιητής εκπαιδεύεται με τον συνδυασμό των τριών συνιστωσών απώλειας. Το δεύτερο τμήμα παρουσιάζει τη διαδικασία βελτιστοποίησης η οποία φορτώνει τον προ-εκπαιδευμένο κωδικοποιητή και εκπαιδεύει ολόκληρο το δίκτυο SwinUNETR [328] για την τμηματοποίηση της αορτής.



Εικόνα 6.1: Το επάνω μέρος παρουσιάζει την προτεινόμενη στρατηγική προ-εκπαίδευσης βασισμένη στην μάσκα καθοδηγούμενη από την φωτεινότητα (IGM), ενώ το κάτω μέρος απεικονίζει τη διαδικασία βελτιστοποίησης για την τμηματοποίηση

6.2.2 Προ-εκπαίδευση SSL

6.2.2.1 Μονάδα μάσκας καθοδηγούμενη από την φωτεινότητα

Η μοντελοποίηση εικόνων με μάσκα (Masked Image Modeling – MIM) [369] έχει χρησιμοποιηθεί σε φυσικές και ιατρικές εικόνες με το να αποκρύπτει τμήματα μέσα στις εικόνες και να στοχεύει στην ανακατασκευή αυτών των τμημάτων μέσω του κορμού του

νευρωνικού δικτύου. Ωστόσο, η αποτελεσματικότητά της εξαρτάται σε μεγάλο βαθμό από τη στρατηγική εφαρμογής της μάσκας οδηγώντας σε μεγάλη μεταβλητότητα στην απόδοση [381].

Στη δική μας εργασία λαμβάνοντας υπόψη ότι στις εικόνες αξονικής τομογραφίας (CT) το εύρος των μονάδων Hounsfield μπορούν να συσχετιστούν με συγκεκριμένα ανατομικά μέρη του σώματος, π.χ. τιμές για οστά, μαλακούς ιστούς και όργανα (π.χ. ήπαρ ≈ 50 , φλοιώδες οστό > 1000) [382], χρησιμοποιούμε μια μονάδα που αποκρύπτει τμήματα της εικόνας όπου οι τιμές φωτεινότητας αντιστοιχούν σε ένα τυχαία επιλεγμένο εύρος.

Η μονάδα λαμβάνει ως είσοδο το κέντρο C και το πλάτος W του παραθύρου, το οποίο ορίζει το εύρος των εντάσεων στις οποίες θα εφαρμοστεί η μάσκα. Η εφαρμογή της μάσκας ορίζεται από την ακόλουθη εξίσωση:

$$M(I) = \begin{cases} 0 & \text{if } (C - \frac{W}{2}) \leq I \leq (C + \frac{W}{2}) \\ 1 & \text{otherwise} \end{cases} \quad (6-1)$$

και η εικόνα έπειτα από την εφαρμογή της μάσκας παράγεται από την ακόλουθη σχέση:

$$I_{masked} = M^{\circ} I \quad (6-2)$$

όπου το $\circ$ δηλώνει τον πολλαπλασιασμό ανά στοιχείο που χρησιμοποιείται για την εφαρμογή της μάσκας.

6.2.2.2 Συνάρτηση απώλειας SSL

Η πλήρης συνάρτηση απώλειας για την προ-εκπαίδευση SSL είναι ένας συνδυασμός απωλειών που περιλαμβάνει την απώλεια ανακατασκευής, την απώλεια αντίθεσης και την απώλεια βασισμένη στην περιστροφή. Οι δύο τελευταίες απώλειες βασίζονται στην εργασία [383]. Το τμήμα ανακατασκευής αποτελείται από έναν αποκωδικοποιητή με συνελκτικά μπλοκ και τις μονάδες υπερδειγματοληψίας (5 μπλοκς όπως φαίνεται στην **(Εικόνα 6.1)**), με σκοπό την ανακατασκευή των όγκων χωρίς μάσκα για την απώλεια ανακατασκευής (Σφάλμα Μέσου Τετραγώνου — Mean Squared Error (MSE)):

$$L_{recon} = \frac{1}{N} \sum_{i=1}^N (I_{rec} - I_{true})^2 \quad (6-3)$$

Για την απώλεια αντίθεσης, οι όγκοι εισόδου και οι μετατροπές τους, εισάγονται στο δίκτυο του κωδικοποιητή. Οι αναπαραστάσεις περνούν μέσα από ένα γραμμικό επίπεδο και τα

διανύσματα εξόδου συγκρίνονται. Η απώλεια επιβάλλει υψηλή ομοιότητα ανάμεσα σε μια εικόνα και την μετασχηματισμένη εκδοχή της ενώ διατηρεί χαμηλότερη ομοιότητα με άλλες εικόνες. Υπολογίζεται με συνημίτονο ομοιότητας σύμφωνα με την παρακάτω εξίσωση:

$$L_{cont} = -\log \frac{\exp(\text{sim}(z_i, z_j)/t)}{\sum_k^{2N} 1_{k \neq i} \exp(\text{sim}(z_i, z_k)/t)} \quad (6-4)$$

όπου το z_i δηλώνει την αναπαράσταση του τμήματος αντίθεσης, το t δηλώνει την παράμετρο θερμοκρασίας και το 1 είναι μία συνάρτηση-δείκτης που παίρνει την τιμή 1 μόνο αν $k \neq i$.

Για τη απώλεια βασισμένη στην περιστροφή, ο όγκος εισόδου περιστρέφεται τυχαία γύρω από τον άξονα z κατά μία από τέσσερις γωνίες (0° , 90° , 180° ή 270°). Στη συνέχεια, ένα γραμμικό επίπεδο συνδέεται με τον κωδικοποιητή για να προβλέψει τις πιθανότητες της γωνίας περιστροφής και εφαρμόζεται μια απώλεια διασταυρούμενης εντροπίας. Η συνολική απώλεια SSL είναι το άθροισμα των προηγούμενων απωλειών :

$$L = L_{recon} + L_{contr} + L_{rot} \quad (6-5)$$

6.2.2.3 Βασικό τμήμα δικτύου - SwinUNETR

Το SwinUNETR [328] χρησιμοποιείται στην προτεινόμενη μέθοδο ως το βασικό δίκτυο τμηματοποίησης που θα εκπαιδευτεί για την τμηματοποίηση της αορτής. Το τμήμα του κωδικοποιητή, προ-εκπαιδεύεται μέσω του προτεινόμενου σχήματος. Η αρχιτεκτονική βασίζεται στο σχήμα κωδικοποιητή-αποκωδικοποιητή, παρόμοιο με το U-Net. Η αρχιτεκτονική του κωδικοποιητή χρησιμοποιεί τις μονάδες Swin μετασχηματιστών [384] σε μια ιεραρχική δομή για την επεξεργασία τρισδιάστατων κομματιών εισόδου συλλαμβάνοντας αποτελεσματικά τόσο τοπικά όσο και συνολικά χαρακτηριστικά.

Ένα επίπεδο διαχωρισμού σε τμήματα χρησιμοποιείται για να χωρίσει τον αρχικό τρισδιάστατο όγκο, $X \in R^{H \times W \times D \times S}$ σε μη επικαλυπτόμενα παράθυρα με αποτέλεσμα μια ακολουθία από τμήματα (με μέγεθος $\left\lceil \frac{H}{H'} \right\rceil \times \left\lceil \frac{W}{W'} \right\rceil \times \left\lceil \frac{D}{D'} \right\rceil$). Αυτά τα τμήματα προβάλλονται γραμμικά για να δημιουργήσουν τις ενσωματώσεις της διάστασης C . Ένα τρισδιάστατο παράθυρο μεγέθους P χρησιμοποιείται για να χωρίσει ομοιόμορφα το τρισδιάστατο διακριτό στοιχείο σε περιοχές $\left\lceil \frac{H'}{P} \right\rceil \times \left\lceil \frac{W'}{P} \right\rceil \times \left\lceil \frac{D'}{P} \right\rceil$ σε κάθε επίπεδο του μετασχηματιστή κωδικοποιητή. Στο επόμενο επίπεδο, τα παράθυρα μετατοπίζονται κατά $\left\lfloor \frac{P}{2} \right\rfloor$ σε κάθε διάσταση και οι εξισώσεις για τα επίπεδα l και $l + 1$ είναι οι ακόλουθες:

$$\begin{aligned}
\hat{z}^l &= W - MSA\left(LN(z^{l-1})\right) + z^{l-1} \\
z^l &= MLP\left(LN(\hat{z}^l)\right) + \hat{z}^l \\
\hat{z}^{l+1} &= SW - MSA\left(LN(z^l)\right) + z^l \\
z^{l+1} &= MLP\left(LN(\hat{z}^{l+1})\right) + \hat{z}^{l+1}
\end{aligned}
\tag{6-6}$$

Στις εξισώσεις αυτές, το W-MSA αντιπροσωπεύει το τυπικό και το SW-MSA αντιπροσωπεύει το πολυ-τμηματικό μηχανισμό αυτο-προσοχής βασισμένο σε παράθυρα. Οι έξοδοι των επιπέδων αυτο-προσοχής απεικονίζονται με το $\hat{z}$. Επιπλέον, το MLP και το LN αναφέρονται στα στοιχεία Πολυεπίπεδου Νευρωνικού Δικτύου (Multi-Layer Perceptron) και Επίπεδου Κανονικοποίησης (Layer Normalization) αντίστοιχα. Η αυτο-προσοχή υπολογίζεται ως εξής:

$$Attention(Q, K, V) = Softmax\left(\frac{QK^T}{\sqrt{d}}\right)V
\tag{6-7}$$

όπου τα Q, K και V αναφέρονται αντίστοιχα σε ερώτημα (query), κλειδί (key) τιμή (value), και το d είναι το μέγεθος των Q και K . Ο κωδικοποιητής αποτελείται από τέσσερα στάδια, το καθένα με δύο μπλοκ μετασχηματιστών (σύνολο οκτώ επιπέδων) και μειώνει προοδευτικά την χωρική ανάλυση (ενώ διπλασιάζει τα κανάλια) στο τέλος κάθε σταδίου μέσω συγχώνευσης τμημάτων εικόνας με αποτέλεσμα την ιεραρχική αναπαράσταση χαρακτηριστικών σε όλα τα στάδια. Το τμήμα αποκωδικοποιητή του δικτύου αποτελείται από υπολειμματικά μπλοκς με δύο διαδοχικά συνελκτικά επίπεδα και συνάρτηση ενεργοποίησης ReLU. Η έξοδος αυτών των υπολειμματικών μπλοκς περνά μέσα από ένα επίπεδο αντίστροφης συνέλιξης για να αυξηθεί η ανάλυση κατά 2 φορές. Τα αντίστοιχα χαρακτηριστικά από τον κωδικοποιητή μεταφέρονται αρχικά μέσω συνδέσεων παράκαμψης και στη συνέχεια συνενώνονται με τους αυξημένης ανάλυσης προαναφερθέντες χάρτες χαρακτηριστικών. Αυτή η διαδικασία επαναλαμβάνεται σε κάθε στάδιο του αποκωδικοποιητή και στο τελικό βήμα ένα συνελκτικό επίπεδο με συνάρτηση ενεργοποίησης softmax παράγει τη μάσκα τμηματοποίησης.

6.2.2.4 Σύνολα δεδομένων

Η παρούσα μελέτη χρησιμοποιεί τρία σύνολα δεδομένων, δύο δημόσια και ένα ιδιωτικό. Το δημόσιο σύνολο δεδομένων από το Multi-Atlas Labeling Beyond the Cranial Vault – Workshop and Challenge [385] (αναφερόμενο ως SYNAPSE) το οποίο περιλάμβανε 50 αξονικές τομογραφίες κοιλίας και το οποίο χρησιμοποιήθηκε μόνο για προ-εκπαίδευση. Οι 50 αξονικές τομογραφίες αποκτήθηκαν με ποικίλες διαστάσεις όγκου που κυμαίνονταν από $512 \times 512 \times 85$ έως $512 \times 512 \times 198$ mm³ και μετρήσεις πεδίου όρασης μεταξύ περίπου

$280 \times 280 \times 280$ και $500 \times 500 \times 650$ mm³. Η ενδοεπίπεδη ανάλυση κυμαινόταν από 0.54×0.54 έως 0.98×0.98 mm² ενώ το πάχος τομής κυμαινόταν από 2.5 έως 5.0 mm.

Το δεύτερο δημόσιο σύνολο δεδομένων (αναφέρεται ως PUBLIC-AORTA) περιλαμβάνει 56 εικόνες αξονικής τομογραφίας (CT) που συλλέχθηκαν από υγιή άτομα προερχόμενες αρχικά από τρία διαφορετικά σύνολο δεδομένων: το KiTS19 Grand Challenge, το Rider Lung CT σύνολο δεδομένων και περιστατικά από το Dongyang Hospital [334]. Οι εικόνες αυτές αποκτήθηκαν με χρήση αξονικής αγγειογραφίας (CTA). Τα σύνολα δεδομένων KiTS19 και Rider Lung CT έχουν ανάλυση 512×512 pixels με διάμεσο πάχος τομής 5 mm και 0.625 mm και 146 και 1008 αξονικές τομές αντίστοιχα. Το σύνολο δεδομένων του Dongyang Hospital έχει ανάλυση 512×666 pixels, πάχος τομής 3 mm και 149 αξονικές τομές.

Σχετικά με το ιδιωτικό σύνολο δεδομένων (PRIVATE-AORTA), συνολικά 20 αξονικές τομογραφίες (CT) από ασθενείς με διάγνωση ανευρύσματος κοιλιακής αορτής (AAA) συλλέχθηκαν από το Τμήμα Αγγειοχειρουργικής του Γενικού Πανεπιστημιακού Νοσοκομείου ΑΤΤΙΚΟΝ. Οι εικόνες αποκτήθηκαν με τη χρήση 3Δ αξονικού τομογράφου Philips Brilliance 16 τομών. Η ηθική ανασκόπηση και έγκριση παραλείφθηκαν για τη συγκεκριμένη μελέτη λόγω της ανωνυμοποίησης των τελικών δεδομένων που χρησιμοποιήθηκαν από τον προτεινόμενο αλγόριθμο. Οι ψηφιακές εικόνες είχαν τα εξής χαρακτηριστικά: διάμεσες διαστάσεις πλέγματος $512 \times 512 \times 448$ pixels, με διάμεση ανάλυση ογκοστοιχείων $0.88 \times 0.88 \times 1.13$ mm.

6.2.2.5 Λεπτομέρειες υλοποίησης

Οι υλοποιήσεις βαθιάς μάθησης βασίστηκαν στο Pytorch και στη βιβλιοθήκη MONAI [386]. Τα πειράματα εκτελέστηκαν σε σταθερό υπολογιστή με κάρτα γραφικών Nvidia 4090 εξοπλισμένη με 24 GB μνήμης. Λόγω περιορισμών μνήμης GPU, οι όγκοι εισόδου χωρίστηκαν σε υπο-όγκους μεγέθους (64,64,64) μέσω τυχαίας περικοπής. Οι εικόνες CT αρχικά περιορίστηκαν σε εύρος Hounsfield Units (HU) (-275,1900) ώστε να συμπεριληφθούν οι μαλακοί ιστοί και οι εντάσεις που βελτιώνουν τη τμηματοποίηση της αορτικής δομής [382]. Στη συνέχεια εφαρμόστηκε κανονικοποίηση εντάσεων στο (0.1) και επαναδειγματοληψία σε διάστημα (1.0,1.0,1.5) για τη σταθεροποίηση της διαδικασίας εκπαίδευσης.

Για την προ-εκπαίδευση αλλά και για την εκ των υστέρων εργασία τμηματοποίησης της αορτής χρησιμοποιήθηκε ο βελτιστοποιητής AdamW με ρυθμό μάθησης 10^{-4} . Η τμηματοποίηση εκπαιδεύτηκε με συνδυασμένη συνάρτηση απώλειας που περιλάμβανε τη διασταυρούμενη εντροπία και τη μετρική DSC για 800 εποχές με επιλογή πρόωρης διακοπής.

Όσον αφορά τις υπερπαραμέτρους του δικτύου, το SwinUNETR περιλαμβάνει 4 στάδια στον κωδικοποιητή, η διάσταση χαρακτηριστικών ορίζεται σε 48 και ο αριθμός των τμημάτων σε (3, 6, 12, 24). Ο αποκωδικοποιητής περιλαμβάνει 5 συνελκτικά μπλοκς με στρώματα υπερδειγματοληψίας σύμφωνα με την [328].

6.3 Αποτελέσματα

Οι πειραματικές συγκρίσεις πραγματοποιήθηκαν χρησιμοποιώντας 3 σύνολα δεδομένων. Χωρίσαμε το PUBLIC-AORTA σε σύνολα εκπαίδευσης/ελέγχου με αναλογία 80%/20%. Για την προ-εκπαίδευση των δικτύων χρησιμοποιήσαμε το σύνολο δεδομένων SYNAPSE καθώς και το τμήμα εκπαίδευσης του PUBLIC-AORTA. Για την βελτιστοποίηση χρησιμοποιήσαμε το εκπαιδευτικό σύνολο του PUBLIC-AORTA και παρουσιάσαμε τα αποτελέσματα στο μη-ορατό σύνολο ελέγχου.

Για το ιδιωτικό σύνολο δεδομένων τα μοντέλα που είχαν προ-εκπαιδευτεί στο PUBLIC-AORTA βελτιστοποιήθηκαν περαιτέρω χρησιμοποιώντας το 50% των δεδομένων και δοκιμάστηκαν στο υπόλοιπο μισό. Για την αξιολόγηση της τμηματοποίησης υπολογίστηκαν ο συντελεστής DSC, η Ανάκληση (Recall), η Ακρίβεια (Precision) και η Μέση Συμμετρική Απόσταση Επιφανείας (ASSD).

Στον (Πίνακας 6.1) συγκρίναμε τη μέθοδό μας τόσο με τα πλέον προηγμένα μοντέλα τμηματοποίησης ιατρικών εικόνων όπως το U-Net [387] με αριθμό φίλτρων (16, 32, 64, 128, 256), το SegResNet (τύπου U-Net με υπολειμματικές συνδέσεις) [388] με φίλτρα (8, 16, 32, 64), το UNETR [365] με φίλτρα (16, 32, 64, 128), και το SwinUNETR [328], όσο και με μεθόδους προ-εκπαίδευσης ώστε να καταδειχθούν τα πλεονεκτήματα και η ανώτερη επίδοση της προτεινόμενης προσέγγισης. Για τη δίκαιη σύγκριση όλων των μεθόδων προ-εκπαίδευσης ο κωδικοποιητής αξιοποιήθηκε στη συνέχεια στην αρχιτεκτονική SwinUNETR για βελτιστοποίηση της τμηματοποίησης.

Μέθοδος	DSC (%)	Recall(%)	Prec(%)	ASSD(mm)
<i>Επιβλεπόμενη χωρίς προ-εκπαίδευση</i>				
U-Net	90.66±3.64	90.52±3.19	90.93±5.26	0.06±0.03
SegResNet	88.70±6.86	91.69±4.57	86.26±9.60	0.05±0.03
UNETR	85.46±8.61	93.43±3.30	79.60±12.85	0.06±0.04
SwinUNETR	88.92±4.97	94.31±4.05	84.57±8.42	0.05±0.03
<i>Αυτό-επιβλεπόμενη με προ-εκπαίδευση</i>				
SwinMM	88.18±6.98	94.62±3.28	83.38±11.22	0.05±0.03
SwinUNETR*	90.11±4.91	91.68±6.00	89.01±6.74	0.05±0.03
Προτεινόμενη	91.20±3.75	93.18±4.03	89.46±5.09	0.05±0.03

±: υποδεικνύει την τιμή της τυπικής απόκλισης

*: SwinUNETR προ-εκπαιδευμένο σύμφωνα με [383]

Πίνακας 6.1: Σύγκριση απόδοσης μεθόδων για το δημόσιο σύνολο δεδομένων

6.3.1 Ποιοτική Αξιολόγηση

Στην **(Εικόνα 6.2)** παρουσιάζονται τα αποτελέσματα τμηματοποίησης από μια εικόνα του συνόλου δεδομένων PUBLIC-AORTA. Σε κάθε υπο-εικόνα απεικονίζεται η επικάλυψη μεταξύ των δεδομένα αναφοράς (Ground Truth – GT (κόκκινο χρώμα)) και της προβλεπόμενης μάσκας (πράσινο χρώμα). Στα σημεία όπου εμφανίζεται μόνο το κόκκινο χρώμα έχουμε υπο-τμηματοποίηση, ενώ στις περιοχές με μόνο πράσινο χρώμα έχουμε υπερ-τμηματοποίηση. Μπορεί να παρατηρηθεί ότι η προτεινόμενη μέθοδος και το SwinUNETR με προ-εκπαίδευση πέτυχαν τα καλύτερα αποτελέσματα χωρίς να χαθούν οι βασικές ανατομίες ενώ το SwinUNETR χωρίς προ-εκπαίδευση και το U-Net δεν τμηματοποίησαν έναν κλάδο και οι υπόλοιποι αλγόριθμοι έχασαν και τους δύο κλάδους.

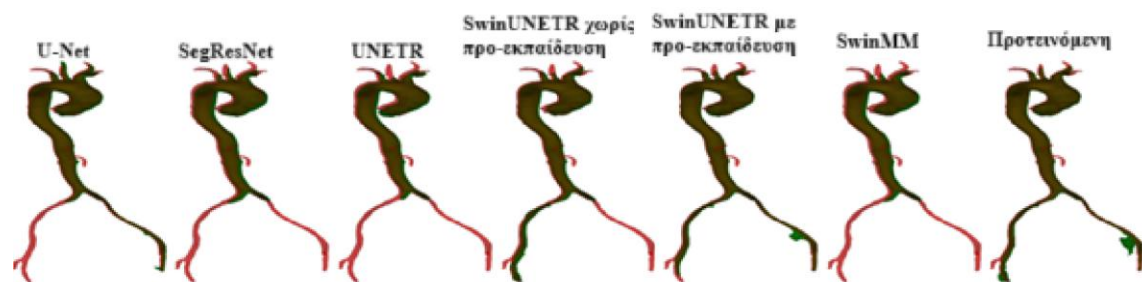
Στο ιδιωτικό σύνολο δεδομένων **(Εικόνα 6.3)** η τμηματοποίηση της αορτής εστιάζει σε μια συγκεκριμένη περιοχή, γεγονός που καθιστά το έργο πιο απαιτητικό λόγω της αυξημένης μεταβλητότητας σε σύγκριση με το δημόσιο σύνολο δεδομένων στο οποίο είχε προ-εκπαιδευτεί το μοντέλο. Επιπλέον, μπορούμε να παρατηρήσουμε ότι σε μια βασική περίπτωση αναφοράς η ακρίβεια τμηματοποίησης παραμένει υψηλή με τα περισσότερα μοντέλα να περιγράφουν το μεγαλύτερο μέρος της πραγματικής αορτής. Στο συγκεκριμένο δείγμα η προτεινόμενη μέθοδος και το SwinUNETR παρουσίασαν λιγότερες περιοχές ψευδώς θετικών αποτελεσμάτων.

6.3.2 Ποσοτική Αξιολόγηση

Οι συγκρίσεις μεταξύ των βασικών εποπτευόμενων μοντέλων και των μοντέλων που βασίζονται σε προ-εκπαίδευση παρουσιάζονται στον **(Πίνακας 6.1)**. Ο πίνακας εμφανίζει τις μέσες τιμές και την τυπική απόκλιση σε παρένθεση. Η προτεινόμενη μέθοδος πέτυχε την υψηλότερη τιμή DSC με 91.20% ξεπερνώντας το δεύτερο καλύτερο U-Net με 90.66% και το SegResNet με 88.70%. Τα μοντέλα που βασίζονται σε μετασχηματιστές, UNETR και SwinUNETR, πέτυχαν 85.46% και 88.92% DSC κάτι που μπορεί επίσης να σχετίζεται με την απαίτηση των μετασχηματιστών για μεγαλύτερο σύνολο εκπαίδευσης. Όσον αφορά την απόσταση (ASSD), τα U-Net και UNETR απέδωσαν χειρότερα από τα υπόλοιπα.

Στο δεύτερο μέρος του **(Πίνακας 6.1)**, η προτεινόμενη μέθοδος προ-εκπαίδευσης συγκρίθηκε με τη βασική προ-εκπαίδευση SwinUNETR [383] και τις μεθόδους SwinMM [380] δείχνοντας υψηλότερη ακρίβεια και από τις δύο, με αύξηση 1.09% και 3.02% στο DSC αντίστοιχα. Ο συνδυασμός υψηλών τιμών Recall και Precision της μεθόδου μας δείχνει ότι εντοπίσαμε μεγάλο μέρος της αορτής αποφεύγοντας παράλληλα τα ψευδώς θετικά. Οι μέθοδοί μας έδειξαν επίσης σταθερά αποτελέσματα παρουσιάζοντας συνολικά χαμηλή τυπική απόκλιση μεταξύ των μετρικών. Η στατιστική σημαντικότητα επιβεβαιώθηκε με χρήση ενός ζεύγους t-test ($p\text{-value} < 0.05$) μεταξύ των μοντέλων με την καλύτερη απόδοση που παρουσιάζονται στους πίνακες.

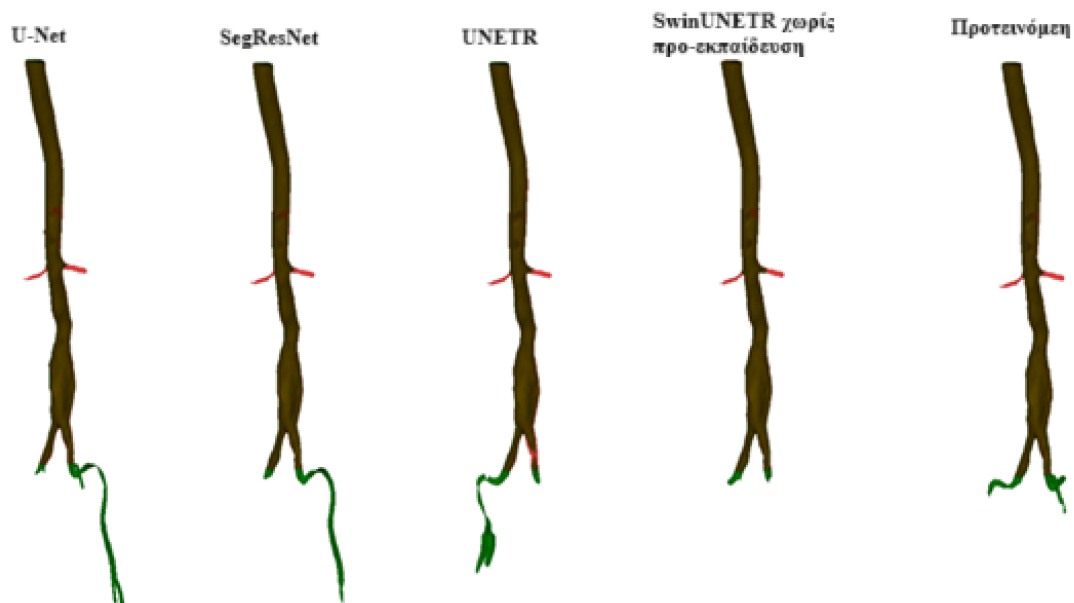
Ο (Πίνακας 6.2) παρουσιάζει την απόδοση τμηματοποίησης των επιβλεπόμενων μοντέλων και της προτεινόμενης προσέγγισης στο ιδιωτικό σύνολο δεδομένων. Σε αντίθεση με το δημόσιο σύνολο δεδομένων το οποίο αποτελείται από υγιή άτομα, το ιδιωτικό σύνολο περιλαμβάνει μόνο ασθενείς με ανεύρυσμα κοιλιακής αορτής (AAA) γεγονός που καθιστά πιο απαιτητική τη μεταφερσιμότητα του μοντέλου. Ως εκ τούτου, αναμένονταν χαμηλότερες συνολικές τιμές DSC λόγω και του περιορισμένου μεγέθους του δείγματος. Όσον αφορά τις αποστάσεις οι μέθοδοι απέδωσαν παρόμοια. Μπορούμε να παρατηρήσουμε ότι οι τιμές DSC παρέμειναν επίσης υψηλές (81%-85%) και ότι το προτεινόμενο μοντέλο το οποίο είχε προηγουμένως εκπαιδευτεί με χρήση IGM ξεπέρασε τις επιβλεπόμενες βάσεις σε όλες τις μετρικές.



Εικόνα 6.2: Σύγκριση της προτεινόμενης και της βασικής μεθόδου στο δημόσιο σύνολο δεδομένων. Τα δεδομένα αναφοράς (Ground Truth, GT) εμφανίζονται με κόκκινο, οι προβλέψεις με πράσινο και η επικάλυψη απεικονίζεται για κάθε μοντέλο

Μέθοδος	DSC (%)	Recall(%)	Prec(%)	ASSD(mm)
U-Net	83.68±17.80	90.01±11.65	79.48±21.30	0.04±0.02
SegResNet	82.72±17.52	87.75±10.91	79.87±21.97	0.05±0.02
UNETR	81.67±18.69	92.97±4.87	76.47±22.25	0.05±0.02
SwinUNETR	83.22±18.74	90.91±11.34	78.74±23.58	0.04±0.02
Προτεινόμενη	85.00±14.36	93.48±6.80	80.17±19.42	0.04±0.02

Πίνακας 6.2: Σύγκριση απόδοσης μεθόδων για το ιδιωτικό σύνολο δεδομένων



Εικόνα 6.3: Σύγκριση της προτεινόμενης και της βασικής μεθόδου στο ιδιωτικό σύνολο δεδομένων. Τα δεδομένα αναφοράς εμφανίζονται με κόκκινο, οι προβλέψεις με πράσινο και η επικάλυψη απεικονίζεται για κάθε μοντέλο

6.4 Συζήτηση

Η ακριβής τμηματοποίηση της αορτής είναι κρίσιμη για τη διάγνωση και την παρακολούθηση παθήσεων όπως το ανεύρυσμα κοιλιακής αορτής (AAA) καθώς επιτρέπει ακριβείς μετρήσεις για την εξέλιξη της νόσου και τον σχεδιασμό της θεραπείας. Στην παρούσα εργασία προτείνουμε μια μέθοδο προ-εκπαίδευσης όπου οι ανατομικές ιδιότητες της αξονικής τομογραφίας αξιοποιούνται για τον σχεδιασμό και την υλοποίηση μίας βοηθητικής διεργασίας, η οποία επιτρέπει στο δίκτυο να μαθαίνει από μη χαραγμένα δεδομένα. Αυτό επιτυγχάνεται μέσω της προτεινόμενης μονάδας Intensity Guided Masking (IGM) και της βοηθητικής λειτουργίας που στοχεύει στην πρόβλεψη αποκρυμμένων περιοχών από CT με βάση το εύρος HU χωρίς να απαιτούνται χειροκίνητες σημειώσεις. Τα προ-εκπαιδευμένα βάρη του κωδικοποιητή μεταφέρονται στη συνέχεια σε ένα μοντέλο SwinUNETR τμηματοποίησης για να διευκολύνουν την αποτύπωση της δομής της αορτής.

Η μέθοδός μας δοκιμάστηκε στην τμηματοποίηση της αορτής τόσο σε ένα δημόσιο σύνολο όσο και σε ένα μικρό ιδιωτικό σύνολο 3Δ CT εικόνων προκειμένου να αξιολογηθεί η αποτελεσματικότητά της και τα οφέλη της. Το προτεινόμενο προ-εκπαιδευμένο δίκτυο ξεπέρασε σε απόδοση δύο ευρέως χρησιμοποιούμενες μεθόδους προ-εκπαίδευσης επιτυγχάνοντας παράλληλα ανώτερη απόδοση σε σύγκριση με τα τρέχοντα πρωτοποριακά μοντέλα τμηματοποίησης Βαθιάς Μάθησης. Η σύγκριση του προτεινόμενου προ-εκπαιδευμένου SwinUNETR με το βασικό U-Net, μια παραλλαγή του U-Net με

υπολειμματικές συνδέσεις (SegResNet), τον ευρέως χρησιμοποιούμενο μετασχηματιστή UNETR και το εποπτευόμενο SwinUNETR δείχνει την υψηλή ακρίβειά του απέναντι σε διαφορετικές αρχιτεκτονικές νευρωνικών δικτύων για τμηματοποίηση. Αν και οι μέθοδοι προ-εκπαίδευσης όπως το προ-εκπαιδευμένο SwinUNETR [369] και το SwinMM [380] έχουν αξιοποιηθεί για CT εικόνες, δεν αξιοποιούν αποτελεσματικά τις έμφυτες ιδιότητες των CT με αποτέλεσμα να έχουν υποδεέστερη απόδοση. Όταν πραγματοποιήθηκε η ενσωμάτωση με το απαιτητικό σε δεδομένα SwinUNETR, η μέθοδός μας πέτυχε ανώτερη DSC επικάλυψη ξεπερνώντας τις μεθόδους προ-εκπαίδευσης και βελτιώνοντας την απλή εποπτευόμενη έκδοση χωρίς προ-εκπαίδευση κατά 2.28 μονάδες.

Επιπλέον, στον **(Πίνακας 6.3)** η προτεινόμενη μέθοδος με προ-εκπαίδευση συγκρίθηκε με το ίδιο SwinUNETR με τη συμβατική εποπτευόμενη εκπαίδευση (χωρίς προ-εκπαίδευση) για διαφορετικά ποσοστά των δειγμάτων εκπαίδευσης. Όταν χρησιμοποιήθηκε μόνο το 15% και το 30% του δημόσιου συνόλου δεδομένων αορτής, το προτεινόμενο μοντέλο πέτυχε 80.40% και 86.53% DSC επικάλυψη ξεπερνώντας την εποπτευόμενη έκδοσή κατά περίπου 3%. Όπως ήταν αναμενόμενο, καθώς το σύνολο εκπαίδευσης αυξάνεται η διαφορά μειώνεται επειδή η έκδοχή χωρίς προ-εκπαίδευση βρίσκει περισσότερα δεδομένα για εκπαίδευση. Ωστόσο, η προτεινόμενη μέθοδος παραμένει καλύτερη κατά περίπου 1% στο DSC, ενώ η απόσταση παραμένει σταθερή σε όλα τα δείγματα.

Ποσοστό	SwinUNETR χωρίς προ-εκπαίδευση		Προτεινόμενη	
	DSC(%)	ASSD(mm)	DSC(%)	ASSD(mm)
15%	77.40±25.21	0.06±0.04	80.40±25.77	0.05±0.04
30%	83.13±8.79	0.07±0.04	86.53±9.71	0.05±0.03
50%	88.68±7.96	0.05±0.03	89.62±4.70	0.05±0.03
100%	88.92±4.97	0.05±0.03	91.20±3.75	0.05±0.03

SwinUNETR χωρίς προ-εκπαίδευση: Απλή εποπτευόμενη έκδοση χωρίς προ-εκπαίδευση

Πίνακας 6.3: Πειράματα για διαφορετικά μεγέθη συνόλου εκπαίδευσης

Παρόλο που η μέθοδός μας έχει επικυρωθεί σε δύο δημόσια και ένα ιδιωτικό σύνολο δεδομένων, η αξιοποίηση ενός ακόμη μεγαλύτερου συνόλου δεδομένων CT θα μπορούσε να βελτιώσει περαιτέρω την απόδοση και να συμβάλει στην ανάπτυξη ενός ισχυρού θεμελιακού μοντέλου για την CT απεικόνιση. Επιπλέον, λόγω περιορισμών στη μνήμη GPU, τα δίκτυα τροφοδοτήθηκαν με όγκους μεγέθους (64,64,64) κάτι που μπορεί να επιβάλει περιορισμούς στο πεδίο όρασης και κατά συνέπεια στην τμηματοποίηση της δομής της αορτής. Η χρήση ενός μεγαλύτερου παραθύρου θα επιτρέψει στο δίκτυο να μοντελοποιήσει καλύτερα τις μακροχρόνιες εξαρτήσεις. Τέλος, η ενίσχυση του μοντέλου μέσω πολυ-εργασιακής μάθησης

ώστε να τμηματοποιεί επίσης τον ενδοαυλικό θρόμβο (ILT) και τις αγγειακές αποτιτανώσεις σε αορτικές παθήσεις μπορεί να βελτιώσει σημαντικά τη ροή εργασιών στην κλινική πράξη.

6.5 Συμπεράσματα

Η χαρτογράφηση της αορτής είναι ζωτικής σημασίας για τη διάγνωση και την ποσοτικοποίηση αγγειακών παθήσεων όπως τα ανευρύσματα της κοιλιακής αορτής. Ωστόσο, η χειροκίνητη τμηματοποίηση κομμάτι-κομμάτι σε 3D απεικονίσεις CT είναι τόσο χρονοβόρα όσο και ευαίσθητο σε διαφορές μεταξύ παρατηρητών. Αν και οι μέθοδοι βαθιάς μάθησης έχουν μεγάλη επιτυχία στην τμηματοποίηση ιατρικών εικόνων, η ανάγκη τους για μεγάλες ανεπτυγμένες βάσεις δεδομένων περιορίζει την εφαρμοσιμότητά τους. Σε αυτή τη μελέτη, προτείνουμε μια αυτοεποπτευόμενη βοηθητική διεργασία για προ-εκπαίδευση σε μη χαραγμένα δεδομένα, το οποίο προβλέπει κρυφές περιοχές στις CT εικόνες με βάση τη συσχέτιση μεταξύ των μονάδων Hounsfield και των ανατομικών δομών. Το μοντέλο που προ-εκπαιδεύτηκε με την προτεινόμενη μέθοδο Intensity Guided Masking (IGM) πέτυχε υψηλή ακρίβεια τμηματοποίησης και αποδοτικότητα δεδομένων τόσο στα δημόσια όσο και ιδιωτικά σύνολα δεδομένων CT υπερβαίνοντας τις πρωτοποριακές μεθόδους.

Κεφάλαιο 7: Συμπεράσματα και μελλοντικές προοπτικές

Η παρούσα διδακτορική διατριβή παρουσιάζει την ανάπτυξη και εφαρμογή καινοτόμων υπολογιστικών τεχνικών για την ανάλυση και επεξεργασία καρδιολογικών δεδομένων. Εστιάζει στη χρήση μεθόδων τμηματοποίησης με εργαλεία επεξεργασίας εικόνας, μηχανικής και βαθιάς μάθησης με σκοπό την αυτοματοποίηση, τη βελτίωση της ακρίβειας και τη μείωση του χρόνου επεξεργασίας. Η αξιολόγηση πραγματοποιήθηκε με δεδομένα υγιών και ασθενών ατόμων από διάφορες πηγές ενώ η απόδοση των προτεινόμενων μεθόδων συγκρίθηκε ποσοτικά και ποιοτικά με καθιερωμένους ευρέως διαδεδομένους αλγόριθμους.

Στο δεύτερο κεφάλαιο παρουσιάζονται συνοπτικά τα θεμελιώδη χαρακτηριστικά της καρδιάς με έμφαση στην ανατομική της δομή και τις λειτουργικές της ιδιότητες ενώ παράλληλα εισάγονται οι βασικές απεικονιστικές μέθοδοι που εφαρμόζονται στον κλινικό και ερευνητικό χώρο. Αρχικά εξετάζεται η ανατομία της καρδιάς ξεκινώντας από τη θέση και τη μορφολογία της, προχωρώντας στην περιγραφή των καρδιακών τοιχωμάτων και των βαλβιδικών μηχανισμών και ολοκληρώνοντας με την παρουσίαση της αιματικής κυκλοφορίας. Στη συνέχεια αναλύονται οι δυναμικές διεργασίες που λαμβάνουν χώρα στο μυοκάρδιο οι οποίες αποτελούν τη βάση για την εκκίνηση και τον συγχρονισμό των συστολικών κινήσεων. Επιπλέον, γίνεται αναφορά στις παθολογικές καταστάσεις που επηρεάζουν είτε τη μορφολογία είτε τη λειτουργικότητα της καρδιάς διαταράσσοντας τη φυσιολογική αιμάτωση και κατ' επέκταση την ομοιότητα του οργανισμού. Αφού διαμορφωθεί το θεωρητικό πλαίσιο της καρδιακής λειτουργίας, ακολουθεί η ανάλυση της ανατομίας και του ρόλου της αορτής η οποία ως κύρια αρτηρία του ανθρώπινου σώματος διαδραματίζει καθοριστικό ρόλο στο καρδιαγγειακό σύστημα. Τέλος, εξετάζονται οι κύριες απεικονιστικές τεχνικές για την οπτικοποίηση της καρδιάς οι οποίες αποτελούν αναπόσπαστο διαγνωστικό και κλινικό εργαλείο για την εκτίμηση, παρακολούθηση και θεραπευτική αντιμετώπιση των καρδιαγγειακών νοσημάτων.

Το τρίτο κεφάλαιο παρέχει στον αναγνώστη μία συστηματική εισαγωγή στα θεμελιώδη στάδια προ-επεξεργασίας εικόνων, στις μεθόδους τμηματοποίησης, καθώς και σε διαδικασίες αξιολόγησης της απόδοσης των αντίστοιχων αλγορίθμων. Ειδικότερα, δίνεται έμφαση στο στάδιο της προ-επεξεργασίας, το οποίο συνιστά καθοριστική φάση στην ανάλυση και επεξεργασία εικόνων, καθώς μέσω εξειδικευμένων τεχνικών συμβάλλει στη βελτίωση της ποιότητας των δεδομένων και διασφαλίζει την αποτελεσματικότητα των επόμενων σταδίων επεξεργασίας ή ανάλυσης. Ακολουθώντας, εξετάζονται οι μέθοδοι τμηματοποίησης, οι οποίες αποτελούν κρίσιμο βήμα στη διαδικασία ανάλυσης εικόνων, με στόχο τον διαχωρισμό τους σε διακριτές περιοχές που εμφανίζουν ομοιογένεια ως προς χαρακτηριστικά όπως το χρώμα, η υφή ή η φωτεινότητα. Το κεφάλαιο ολοκληρώνεται με την παρουσίαση διαδικασιών

αποτίμησης των αλγορίθμων τμηματοποίησης, οι οποίες είναι αναγκαίες για την εκτίμηση της ακρίβειας, της αποδοτικότητας και της καταλληλότητάς τους σε διαφορετικά είδη δεδομένων.

Στο τέταρτο κεφάλαιο παρουσιάζεται μία νέα, πλήρως αυτοματοποιημένη μέθοδος τρισδιάστατης (3D) τμηματοποίησης, η οποία βασίζεται σε μία διαδικασία δύο σταδίων για την ανίχνευση και τη μοντελοποίηση της αορτής σε δεδομένα τρισδιάστατης αξονικής τομογραφίας (CT). Η προτεινόμενη μέθοδος τμηματοποίησης και μοντελοποίησης συγκρίθηκε άμεσα με τις μεθόδους K-Means, Κατωφλίωσης και τα Δίκτυα Μαρκοβιαίων Τυχαίων Πεδίων (Markov Random Field Networks), με κριτήριο την ακρίβεια του παραγόμενου τρισδιάστατου ανακατασκευασμένου μοντέλου της αορτής σε κάθε περίπτωση. Οι συγκριτικές τεχνικές εφαρμόστηκαν σε 16 σύνολα τρισδιάστατων δεδομένων CT και αξιολογήθηκαν τόσο ποιοτικά όσο και ποσοτικά. Τα αποτελέσματα κατέδειξαν ότι η προτεινόμενη μέθοδος δύο σταδίων υπερέχει σε απόδοση έναντι των άλλων τριών τεχνικών, γεγονός που την καθιστά ισχυρή υποψήφια για εφαρμογή σε πληθώρα κλινικών σεναρίων, όπως ο προεγχειρητικός σχεδιασμός και η αξιολόγηση θεραπευτικών παρεμβάσεων.

Η μελέτη που παρουσιάζεται στο πέμπτο κεφάλαιο, συγκρίνει τέσσερις δημοφιλείς αρχιτεκτονικές βαθιάς μάθησης (U-Net, SegResNet, UNETR, SwinUNETR) για την αυτόματη τμηματοποίηση της κοιλιακής αορτής σε 3D CT δεδομένα χρησιμοποιώντας σύνολα δεδομένων από υγιείς και από ασθενείς με ανεύρυσμα (AAA). Τα αποτελέσματα δείχνουν ότι τα μοντέλα μπορούν να γενικεύσουν αποτελεσματικά σε περιπτώσεις με ανεύρυσμα, παρότι εκπαιδεύτηκαν μόνο σε υγιείς. Το U-Net αναδείχθηκε ως το πιο ισορροπημένο και αποδοτικό μοντέλο πετυχαίνοντας υψηλή ακρίβεια DSC 0.89 ± 0.05 στο δημόσιο σύνολο δεδομένων και 0.89 ± 0.07 στο ιδιωτικό. Το SegResNet εμφάνισε υψηλή ακρίβεια αλλά χαμηλότερη ανάκληση με τάση για υπο-τμηματοποίηση μικρών αγγείων. Η χρήση υβριδικής συνάρτησης κόστους που συνδυάζει την απώλεια Dice και την διασταυρούμενη εντροπία βελτίωσε την απόδοση σε μικρότερες δομές. Συνολικά, το U-Net ξεχώρισε όχι μόνο για την ακρίβεια αλλά και για την υπολογιστική του αποδοτικότητα, καθιστώντας το πρακτική λύση για κλινικές εφαρμογές με περιορισμένους πόρους.

Στην μελέτη του έκτου κεφαλαίου, παρουσιάζεται και αναλύεται μια αυτοεποπτευόμενη βοηθητική διεργασία προ-εκπαίδευσης σε μη χαραγμένα δεδομένα η οποία βασίζεται στην πρόβλεψη καλυμμένων περιοχών σε εικόνες CT αξιοποιώντας τη συσχέτιση μεταξύ των τιμών Hounsfield και των αντίστοιχων ανατομικών δομών. Το μοντέλο που προ-εκπαιδεύτηκε μέσω της προτεινόμενης μεθοδολογίας Intensity Guided Masking (IGM) κατέγραψε υψηλή ακρίβεια τμηματοποίησης και αυξημένη αποδοτικότητα δεδομένων σε δημόσια και ιδιωτικά σύνολα CT υπερβαίνοντας τις υφιστάμενες πρωτοποριακές προσεγγίσεις.

Συνοψίζοντας, η παρούσα μελέτη κατάφερε σε σημαντικό βαθμό να επιτύχει τους κύριους στόχους της που αφορούσαν την αποδοτική και αξιόπιστη αυτοματοποίηση της τμηματοποίησης 3D CT δεδομένων αορτής καθώς και τη βελτίωση τόσο της ακρίβειας όσο και

της ταχύτητας της επεξεργασίας τους. Αυτό είχε ως αποτέλεσμα τη βελτίωση λήψης κλινικών αποφάσεων ώστε να απλοποιούν τις ροές εργασίας και να ενισχύουν τα αποτελέσματα για τους ασθενείς, παρέχοντας μια βάση για την ενσωμάτωση αυτοματοποιημένων εργαλείων στις κλινικές διαδικασίες όπως τον προγραμματισμό για επεμβάσεις TAVI και την αποτύπωση του ανευρύσματος κοιλιακής αορτής (AAA). Για να το πετύχει, εισαγάγαμε και συνδυάσαμε τόσο τεχνικές επεξεργασίας εικόνας και μηχανικής μάθησης όσο και αρχιτεκτονικές βαθιάς μάθησης χρησιμοποιώντας σύνολα δεδομένων από υγιείς και ασθενείς.

Η διερεύνηση βελτιωμένων μεθοδολογιών και προηγμένων τεχνικών για την ανάλυση και επεξεργασία καρδιολογικών δεδομένων δεν περιορίζεται στο πλαίσιο της παρούσας διατριβής. Αντιθέτως, αναδεικνύονται προοπτικές και κατευθύνσεις με ιδιαίτερο ερευνητικό ενδιαφέρον, οι οποίες δύνανται να συμβάλουν ουσιαστικά στη μελλοντική εξέλιξη και περαιτέρω βελτιστοποίηση των εν λόγω τεχνικών. Μια τέτοια περίπτωση αποτελεί η εκτενέστερη έρευνα που σκοπεύουμε να υλοποιήσουμε στη μελέτη του πέμπτου κεφαλαίου ενσωματώνοντας μεγαλύτερα δείγματα και εξερευνώντας επιπλέον μοντέλα ώστε να διερευνήσουμε αν ο συνδυασμός της διασταυρούμενης εντροπίας με τη συνάρτηση Dice μπορεί να προσφέρει μια αποτελεσματική λύση για τη μείωση της υπο-τμηματοποίησης των μικρότερων αγγείων. Επίσης θα είχε ενδιαφέρον η επέκταση του ιδιωτικού συνόλου δεδομένων του έκτου κεφαλαίου προκειμένου να συμπεριληφθούν σημάνσεις ενδοαυλικού θρόμβου και ασβεστοποίησης επιτρέποντας την τμηματοποίηση πολλαπλών κατηγοριών για πιο ολοκληρωμένη διάγνωση και ποσοτικοποίηση της αορτής.

BIBΛΙΟΓΡΑΦΙΑ

- [1] N. Zhu *et al.*, “A Novel Coronavirus from Patients with Pneumonia in China, 2019.,” *N Engl J Med*, vol. 382, no. 8, pp. 727–733, Feb. 2020, doi: 10.1056/NEJMoa2001017.
- [2] Y. Liu, Z. Zhang, X. Liu, L. Wang, and X. Xia, “Efficient image segmentation based on deep learning for mineral image classification,” *Advanced Powder Technology*, vol. 32, no. 10, pp. 3885–3903, Oct. 2021, doi: 10.1016/j.appt.2021.08.038.
- [3] s Isabella and S. Srinivasan, “An understanding of machine learning techniques in big data analytics: A survey,” *International Journal of Engineering and Technology(UAE)*, vol. 7, pp. 666–669, Jun. 2018, doi: 10.14419/ijet.v7i2.33.15471.
- [4] P. Rajpurkar *et al.*, “Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists.,” *PLoS Med*, vol. 15, no. 11, p. e1002686, Nov. 2018, doi: 10.1371/journal.pmed.1002686.
- [5] S. Kiranyaz, J. Malik, H. B. Abdallah, T. Ince, A. Iosifidis, and M. Gabbouj, “Self-organized Operational Neural Networks with Generative Neurons,” *Neural Networks*, vol. 140, pp. 294–308, Aug. 2021, doi: 10.1016/j.neunet.2021.02.028.
- [6] M. Gavaghan, “Cardiac Anatomy and Physiology: A Review,” *AORN Journal*, vol. 67, no. 4, pp. 800–822, Apr. 1998, doi: 10.1016/S0001-2092(06)62644-6.
- [7] F. J. Schoen, “Evolving Concepts of Cardiac Valve Dynamics,” *Circulation*, vol. 118, no. 18, pp. 1864–1880, Oct. 2008, doi: 10.1161/CIRCULATIONAHA.108.805911.
- [8] R. Reyaldeén *et al.*, “Pericardial Anatomy, Interventions and Therapeutics: A Contemporary Review,” *Structural Heart*, vol. 5, no. 6, pp. 556–569, Nov. 2021, doi: 10.1080/24748706.2021.1989531.
- [9] A. P. Fishman and D. W. Richards, “Coronary Circulation and Cardiac Metabolism,” in *Circulation of the Blood: Men and Ideas*, A. P. Fishman and D. W. Richards, Eds., New York, NY: Springer New York, 1982, pp. 199–264. doi: 10.1007/978-1-4614-7546-0_4.
- [10] Kumar and Clark’s *Clinical Medicine*, 10th ed. 2020.
- [11] I. Rehman, C. Kerndt, and A. Rehman, “Anatomy, Thorax, Heart Left Anterior Descending (LAD) Artery: A Chapter Review,” 2020, p. PMID: 29493997.
- [12] A. J. Weinhaus and K. P. Roberts, “Anatomy of the Human Heart,” in *Handbook of Cardiac Anatomy, Physiology, and Devices*, P. A. Iaizzo, Ed., Totowa, NJ: Humana Press, 2005, pp. 51–79. doi: 10.1007/978-1-59259-835-9_4.
- [13] R. B. Hinton and K. E. Yutzey, “Heart valve structure and function in development and disease.,” *Annu Rev Physiol*, vol. 73, pp. 29–46, 2011, doi: 10.1146/annurev-physiol-012110-142145.
- [14] H. Hosseinzadeh, “Cardiovascular System,” in *Fundamentals of Medicine for Biomedical Engineering*, H. Hosseinzadeh, Ed., Cham: Springer Nature Switzerland, 2025, pp. 435–514. doi: 10.1007/978-3-031-74085-5_13.
- [15] P. P. Sengupta, A. J. Tajik, K. Chandrasekaran, and B. K. Khandheria, “Twist mechanics of the left ventricle: principles and application.,” *JACC Cardiovasc Imaging*, vol. 1, no. 3, pp. 366–376, May 2008, doi: 10.1016/j.jcmg.2008.02.006.
- [16] N. I.A., “The cardiovascular system,” *Texas Journal of Medical Science*, vol. 25, pp. 39–41, Oct. 2023, doi: 10.62480/tjms.2023.vol25.pp39-41.
- [17] L. Gonzalez and A. Dardik, “Chapter 21 - Normal vascular identity (arteries, veins, and lymphatics) and malformations,” in *The Vasculome*, Z. S. Galis, Ed., Academic Press, 2022, pp. 251–263. doi: 10.1016/B978-0-12-822546-2.00016-2.
- [18] C. R. di Gioia, A. Ascione, R. Carletti, and C. Giordano, “Thoracic Aorta: Anatomy and Pathology,” *Diagnostics*, vol. 13, no. 13, 2023, doi: 10.3390/diagnostics13132166.
- [19] S. Laurent *et al.*, “Expert consensus document on arterial stiffness: methodological issues and clinical applications.,” *Eur Heart J*, vol. 27, no. 21, pp. 2588–2605, Nov. 2006, doi: 10.1093/eurheartj/ehl254.
- [20] V. Katsi, N. Magkas, A. Antonopoulos, G. Trantalis, K. Toutouzas, and D. Tousoulis, “Aortic valve: anatomy and structure and the role of vasculature in the degenerative process,” *Acta Cardiologica*, vol. 76, pp. 1–14, Jun. 2020, doi: 10.1080/00015385.2020.1746053.

- [21] S. Z. Li and A. Jain, Eds., “Capillary Blood Vessel,” in *Encyclopedia of Biometrics*, Boston, MA: Springer US, 2009, pp. 178–178. doi: 10.1007/978-0-387-73003-5_133.
- [22] R. M. Anderson, J. M. Fritz, and J. E. O’Hare, “The mechanical nature of the heart as a pump,” *American Heart Journal*, vol. 73, no. 1, pp. 92–105, Jan. 1967, doi: 10.1016/0002-8703(67)90313-4.
- [23] D. M. Bers, “Cardiac excitation-contraction coupling,” *Nature*, vol. 415, no. 6868, pp. 198–205, Jan. 2002, doi: 10.1038/415198a.
- [24] D. P. Zipes, J. Jalife, and W. G. Stevenson, *Cardiac Electrophysiology: From Cell to Bedside: Seventh Edition*. 2017, p. 1389.
- [25] U. Ackermann, “Regulation of arterial blood pressure,” *Surgery (Oxford)*, vol. 22, no. 5, pp. 120a–120f, May 2004, doi: 10.1383/surg.22.5.120a.33383.
- [26] D. A. Kass, J. G. F. Bronzwaer, and W. J. Paulus, “What mechanisms underlie diastolic dysfunction in heart failure?,” *Circ Res*, vol. 94, no. 12, pp. 1533–1542, Jun. 2004, doi: 10.1161/01.RES.0000129254.25507.d6.
- [27] V. Sequeira and J. van der Velden, “Historical perspective on heart function: the Frank-Starling Law,” *Biophys Rev*, vol. 7, no. 4, pp. 421–447, Dec. 2015, doi: 10.1007/s12551-015-0184-4.
- [28] A. A. Gibb, M. P. Lazaropoulos, and J. W. Elrod, “Myofibroblasts and Fibrosis,” *Circulation Research*, vol. 127, no. 3, pp. 427–447, Jul. 2020, doi: 10.1161/CIRCRESAHA.120.316958.
- [29] S. Neubauer, “The failing heart--an engine out of fuel,” *N Engl J Med*, vol. 356, no. 11, pp. 1140–1151, Mar. 2007, doi: 10.1056/NEJMra063052.
- [30] P. Ponikowski *et al.*, “2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: The Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC) Developed with the special contribution of the Heart Failure Association (HFA) of the ESC,” *Eur Heart J*, vol. 37, no. 27, pp. 2129–2200, Jul. 2016, doi: 10.1093/eurheartj/ehw128.
- [31] T. J. Wang, J. C. Evans, E. J. Benjamin, D. Levy, E. C. LeRoy, and R. S. Vasan, “Natural history of asymptomatic left ventricular systolic dysfunction in the community,” *Circulation*, vol. 108, no. 8, pp. 977–982, Aug. 2003, doi: 10.1161/01.CIR.0000085166.44904.79.
- [32] S. Yusuf, B. Pitt, C. E. Davis, W. B. J. Hood, and J. N. Cohn, “Effect of enalapril on mortality and the development of heart failure in asymptomatic patients with reduced left ventricular ejection fractions,” *N Engl J Med*, vol. 327, no. 10, pp. 685–691, Sep. 1992, doi: 10.1056/NEJM199209033271003.
- [33] R. S. Ross, “Ischemic heart disease: An overview,” *American Journal of Cardiology*, vol. 36, no. 4, pp. 496–505, Oct. 1975, doi: 10.1016/0002-9149(75)90901-7.
- [34] R. V. Jensen, M. V. Hjortbak, and H. E. Bøtker, “Ischemic Heart Disease: An Update,” *Semin Nucl Med*, vol. 50, no. 3, pp. 195–207, May 2020, doi: 10.1053/j.semnuclmed.2020.02.007.
- [35] B. Ibanez *et al.*, “2017 ESC Guidelines for the management of acute myocardial infarction in patients presenting with ST-segment elevation: The Task Force for the management of acute myocardial infarction in patients presenting with ST-segment elevation of the European Society of Cardiology (ESC),” *Eur Heart J*, vol. 39, no. 2, pp. 119–177, Jan. 2018, doi: 10.1093/eurheartj/ehx393.
- [36] *Cardiovascular Disability: Updating the Social Security Listings*. Washington (DC), 2010. doi: 10.17226/12940.
- [37] C. Schmidt, J. Kisselbach, P. A. Schweizer, H. A. Katus, and D. Thomas, “The pathology and treatment of cardiac arrhythmias: focus on atrial fibrillation,” *Vasc Health Risk Manag*, vol. 7, pp. 193–202, 2011, doi: 10.2147/VHRM.S10758.
- [38] G. Tse, “Mechanisms of cardiac arrhythmias,” *J Arrhythm*, vol. 32, no. 2, pp. 75–81, Apr. 2016, doi: 10.1016/j.joa.2015.11.003.
- [39] L. Jordaens, “A clinical approach to arrhythmias revisited in 2018: From ECG over noninvasive and invasive electrophysiology to advanced imaging,” *Neth Heart J*, vol. 26, no. 4, pp. 182–189, Apr. 2018, doi: 10.1007/s12471-018-1089-1.

- [40] A. A. Grace and D. M. Roden, "Systems biology and cardiac arrhythmias.," *Lancet*, vol. 380, no. 9852, pp. 1498–1508, Oct. 2012, doi: 10.1016/S0140-6736(12)61462-7.
- [41] T. Ciarambino, G. Menna, G. Sansone, and M. Giordano, "Cardiomyopathies: An Overview.," *Int J Mol Sci*, vol. 22, no. 14, Jul. 2021, doi: 10.3390/ijms22147722.
- [42] E. Braunwald, "Cardiomyopathies," *Circulation Research*, vol. 121, no. 7, pp. 711–721, Sep. 2017, doi: 10.1161/CIRCRESAHA.117.311812.
- [43] R. K. Wexler, T. Elton, A. Pleister, and D. Feldman, "Cardiomyopathy: an overview.," *Am Fam Physician*, vol. 79, no. 9, pp. 778–784, May 2009.
- [44] J. Sanz, D. Sánchez-Quintana, E. Bossone, H. J. Bogaard, and R. Naeije, "Anatomy, Function, and Dysfunction of the Right Ventricle: JACC State-of-the-Art Review.," *J Am Coll Cardiol*, vol. 73, no. 12, pp. 1463–1482, Apr. 2019, doi: 10.1016/j.jacc.2018.12.076.
- [45] T. Lampejo, S. M. Durkin, N. Bhatt, and O. Guttmann, "Acute myocarditis: aetiology, diagnosis and management.," *Clin Med (Lond)*, vol. 21, no. 5, pp. e505–e510, Sep. 2021, doi: 10.7861/clinmed.2021-0121.
- [46] A. L. Klein *et al.*, "Pericardial Diseases: International Position Statement on New Concepts and Advances in Multimodality Cardiac Imaging," *JACC: Cardiovascular Imaging*, vol. 17, no. 8, pp. 937–988, Aug. 2024, doi: 10.1016/j.jcmg.2024.04.010.
- [47] M. H. Khandaker *et al.*, "Pericardial disease: diagnosis and management.," *Mayo Clin Proc*, vol. 85, no. 6, pp. 572–593, Jun. 2010, doi: 10.4065/mcp.2010.0046.
- [48] T. F. Ismail, "Acute pericarditis: Update on diagnosis and management.," *Clin Med (Lond)*, vol. 20, no. 1, pp. 48–51, Jan. 2020, doi: 10.7861/clinmed.cme.20.1.4.
- [49] K. Maganti, V. H. Rigolin, M. E. Sarano, and R. O. Bonow, "Valvular heart disease: diagnosis and management.," *Mayo Clin Proc*, vol. 85, no. 5, pp. 483–500, May 2010, doi: 10.4065/mcp.2009.0706.
- [50] F. Praz, F. Beyersdorf, K. Haugaa, and B. Prendergast, "Valvular heart disease: from mechanisms to management," *The Lancet*, vol. 403, no. 10436, pp. 1576–1589, Apr. 2024, doi: 10.1016/S0140-6736(23)02755-1.
- [51] A. Vegas, "Aorta," in *Perioperative Two-Dimensional Transesophageal Echocardiography: A Practical Handbook*, A. Vegas, Ed., New York, NY: Springer New York, 2012, pp. 137–149. doi: 10.1007/978-1-4419-9952-8_6.
- [52] G. Ironi, G. Brembilla, G. Benedetti, and F. De Cobelli, "Anatomical Overview and Imaging of the Aorta and Visceral Arteries," in *Visceral Vessels and Aortic Repair: Challenges and Difficult Cases*, Y. Tshomba, D. Baccellieri, and R. Chiesa, Eds., Cham: Springer International Publishing, 2019, pp. 3–17. doi: 10.1007/978-3-319-94761-7_1.
- [53] B. Rylski *et al.*, "Aortic Arch Anatomy in Candidates for Aortic Arch Repair.," *Semin Thorac Cardiovasc Surg*, vol. 34, no. 1, pp. 19–26, Spring 2022, doi: 10.1053/j.semthvs.2021.03.001.
- [54] M. Lansakara and S. Unai, "An overview of aortic valve anatomy: the current understanding," *Indian Journal of Thoracic and Cardiovascular Surgery*, vol. 39, no. 2, pp. 246–252, Dec. 2023, doi: 10.1007/s12055-023-01645-x.
- [55] R. R. DeMartino, "Normal Aortic Anatomy and Variations of Its Branches," in *Endovascular Aortic Repair: Current Techniques with Fenestrated, Branched and Parallel Stent-Grafts*, G. S. Oderich, Ed., Cham: Springer International Publishing, 2017, pp. 63–71. doi: 10.1007/978-3-319-15192-2_4.
- [56] T. C. Gasser, "Chapter 8 - Aorta," in *Biomechanics of Living Organs*, vol. 1, Y. Payan and J. Ohayon, Eds., Oxford: Academic Press, 2017, pp. 169–191. doi: 10.1016/B978-0-12-804009-6.00008-0.
- [57] P. Berillis, "The Role of Collagen in the Aorta's Structure," *The Open Circulation & Vascular Journal*, vol. 6, Feb. 2013, doi: 10.2174/1877382601306010001.
- [58] Z.-M. Fang, X. Feng, Y. Chen, H. Luo, D.-S. Jiang, and X. Yi, "Targeting autophagy in aortic aneurysm and dissection," *Biomedicine & Pharmacotherapy*, vol. 153, p. 113547, Sep. 2022, doi: 10.1016/j.biopha.2022.113547.
- [59] M. Halushka, "Pathology of the Aorta," *Surgical Pathology Clinics*, vol. 5, no. 2, pp. 417–433, Jun. 2012, doi: 10.1016/j.path.2012.04.005.

- [60] A. Mathur, V. Mohan, D. Ameta, B. Gaurav, and P. Haranahalli, "Aortic aneurysm," *J Transl Int Med*, vol. 4, no. 1, pp. 35–41, Apr. 2016, doi: 10.1515/jtim-2016-0008.
- [61] K. W. Johnston, R. B. Rutherford, M. D. Tilson, D. M. Shah, L. Hollier, and J. C. Stanley, "Suggested standards for reporting on arterial aneurysms. Subcommittee on Reporting Standards for Arterial Aneurysms, Ad Hoc Committee on Reporting Standards, Society for Vascular Surgery and North American Chapter, International Society for Cardiovascular Surgery," *J Vasc Surg*, vol. 13, no. 3, pp. 452–458, Mar. 1991, doi: 10.1067/mva.1991.26737.
- [62] M. A. Bettmann *et al.*, "Atherosclerotic Vascular Disease Conference: Writing Group VI: revascularization," *Circulation*, vol. 109, no. 21, pp. 2643–2650, Jun. 2004, doi: 10.1161/01.CIR.0000128526.35982.9A.
- [63] L. K. Bickerstaff *et al.*, "Thoracic aortic aneurysms: a population-based study," *Surgery*, vol. 92, no. 6, pp. 1103–1108, Dec. 1982.
- [64] W. D. Clouse, J. W. J. Hallett, H. V. Schaff, M. M. Gayari, D. M. Ilstrup, and L. J. 3rd Melton, "Improved prognosis of thoracic aortic aneurysms: a population-based study," *JAMA*, vol. 280, no. 22, pp. 1926–1929, Dec. 1998, doi: 10.1001/jama.280.22.1926.
- [65] E. S. Crawford and E. S. Cohen, "Aortic aneurysm: a multifocal disease. Presidential address," *Arch Surg*, vol. 117, no. 11, pp. 1393–1400, Nov. 1982, doi: 10.1001/archsurg.1982.01380350001001.
- [66] V. Pressler and J. J. McNamara, "Aneurysm of the thoracic aorta. Review of 260 cases," *J Thorac Cardiovasc Surg*, vol. 89, no. 1, pp. 50–54, Jan. 1985.
- [67] M. E. de Bakey, C. H. McCollum, and J. M. Graham, "Surgical treatment of aneurysms of the descending thoracic aorta: long-term results in 500 patients," *J Cardiovasc Surg (Torino)*, vol. 19, no. 6, pp. 571–576, Dec. 1978.
- [68] Q. Counseller and Y. Aboelkassem, "Recent technologies in cardiac imaging," *Frontiers in Medical Technology*, vol. 4, 2023, doi: 10.3389/fmedt.2022.984492.
- [69] A. Kalra, A. Lowe, and A. Al-Jumaily, "Critical review of electrocardiography measurement systems and technology," *Measurement Science and Technology*, vol. 30, no. 1, p. 012001, Dec. 2018, doi: 10.1088/1361-6501/aaf2b7.
- [70] P. T. Weinfurt, "Electrocardiographic monitoring: An overview," *Journal of Clinical Monitoring*, vol. 6, no. 2, pp. 132–138, Apr. 1990, doi: 10.1007/BF02828290.
- [71] H. Heß, S. Eibel, C. Mukherjee, U. X. Kaisers, and J. Ender, "Quantification of mitral valve regurgitation with color flow Doppler using baseline shift," *Int J Cardiovasc Imaging*, vol. 29, no. 2, pp. 267–274, Feb. 2013, doi: 10.1007/s10554-012-0084-7.
- [72] J. M. Thijssen and C. L. de Korte, "Cardiological Ultrasound Imaging," *Curr Pharm Des*, vol. 20, no. 39, pp. 6150–6161, 2014, doi: 10.2174/1381612820666140417113304.
- [73] R. García-Orta *et al.*, "Three-dimensional versus two-dimensional transesophageal echocardiography in mitral valve repair," *J Am Soc Echocardiogr*, vol. 20, no. 1, pp. 4–12, Jan. 2007, doi: 10.1016/j.echo.2006.07.005.
- [74] J. C. George, V. Varghese, G. Dangas, and T. E. Feldman, "Percutaneous mitral valve repair: lessons from the EVEREST II (Endovascular Valve Edge-to-Edge REpair Study) and beyond," *JACC Cardiovasc Interv*, vol. 4, no. 7, pp. 825–827, Jul. 2011, doi: 10.1016/j.jcin.2011.05.010.
- [75] J. E. Ellis *et al.*, "Right heart dysfunction, pulmonary embolism, and paradoxical embolization during liver transplantation. A transesophageal two-dimensional echocardiographic study," *Anesth Analg*, vol. 68, no. 6, pp. 777–782, Jun. 1989.
- [76] J. E. Ellis and E. M. Bender, "Intraoperative transesophageal echocardiography in blunt thoracic trauma," *J Cardiothorac Vasc Anesth*, vol. 5, no. 4, pp. 373–376, Aug. 1991, doi: 10.1016/1053-0770(91)90163-n.
- [77] H. K. Eltzschig, P. Rosenberger, M. Löffler, J. A. Fox, S. F. Aranki, and S. K. Shernan, "Impact of intraoperative transesophageal echocardiography on surgical decisions in 12,566 patients undergoing cardiac surgery," *Ann Thorac Surg*, vol. 85, no. 3, pp. 845–852, Mar. 2008, doi: 10.1016/j.athoracsur.2007.11.015.
- [78] A. S. Pearlman, "Celebrating (More Than) 50 Years of Doppler Echocardiography," *J Am Soc Echocardiogr*, vol. 31, no. 12, p. 1307, Dec. 2018, doi: 10.1016/j.echo.2018.08.002.

- [79]N. K. Brown, P.-N. Jone, S. Patel, and A. L. Hauck, "Echocardiography," in *Pediatric Cardiology: Fetal, Pediatric, and Adult Congenital Heart Diseases*, R. Abdulla, S. Berger, C. Backer, R. Anderson, R. Holzer, N. Blom, and J. Robinson, Eds., Cham: Springer International Publishing, 2020, pp. 1–65. doi: 10.1007/978-3-030-42937-9_33-1.
- [80]E. Prvulovich, "Myocardial perfusion scintigraphy," *Clinical Medicine*, vol. 6, no. 3, pp. 263–266, May 2006, doi: 10.7861/clinmedicine.6-3-263.
- [81]P. Knaapen, "Myocardial perfusion scintigraphy: interpretation of a normal scan," *Netherlands Heart Journal*, vol. 22, no. 4, pp. 148–150, Apr. 2014, doi: 10.1007/s12471-014-0541-0.
- [82]A. Fathala, "Myocardial perfusion scintigraphy: techniques, interpretation, indications and reporting.," *Ann Saudi Med*, vol. 31, no. 6, pp. 625–634, Dec. 2011, doi: 10.4103/0256-4947.87101.
- [83]S. R. Underwood *et al.*, "Myocardial perfusion scintigraphy: the evidence.," *Eur J Nucl Med Mol Imaging*, vol. 31, no. 2, pp. 261–291, Feb. 2004, doi: 10.1007/s00259-003-1344-5.
- [84]P. S. Rajiah, C. J. François, and T. Leiner, "Cardiac MRI: State of the Art.," *Radiology*, vol. 307, no. 3, p. e223008, May 2023, doi: 10.1148/radiol.223008.
- [85]G. Katti, S. Ara, and D. Shireen, "Magnetic resonance imaging (MRI) - A review," *Intl J Dental Clin*, vol. 3, Mar. 2011.
- [86]J. A. C. Lima and M. Y. Desai, "Cardiovascular magnetic resonance imaging: Current and emerging applications," *Journal of the American College of Cardiology*, vol. 44, no. 6, pp. 1164–1171, Sep. 2004, doi: 10.1016/j.jacc.2004.06.033.
- [87]C. B. Higgins, "Magnetic Resonance Imaging of the Heart," in *Radiology of the Heart: Cardiac Imaging in Infants, Children, and Adults*, H. Spindola-Franco and B. G. Fish, Eds., New York, NY: Springer New York, 1985, pp. 671–679. doi: 10.1007/978-1-4613-8205-8_17.
- [88]A. M. den Harder, A. M. R. Schilham, and M. J. Willemink, "Cardiovascular CT: Image Reconstruction," in *CT of the Heart*, U. J. Schoepf, Ed., Totowa, NJ: Humana Press, 2019, pp. 117–123. doi: 10.1007/978-1-60327-237-7_11.
- [89]M. van der Glas, "Principles of Computerized Tomographic Imaging," 2000. [Online]. Available: <https://api.semanticscholar.org/CorpusID:14508620>
- [90]G. Pontone *et al.*, "Clinical applications of cardiac computed tomography: a consensus paper of the European Association of Cardiovascular Imaging-part II.," *Eur Heart J Cardiovasc Imaging*, vol. 23, no. 4, pp. e136–e161, Mar. 2022, doi: 10.1093/ehjci/jeab292.
- [91]S. S. Halliburton and S. Abbara, "Practical tips and tricks in cardiovascular computed tomography: Patient preparation for optimization of cardiovascular CT data acquisition," *Journal of Cardiovascular Computed Tomography*, vol. 1, no. 1, pp. 62–65, Jul. 2007, doi: 10.1016/j.jcct.2007.04.004.
- [92]G. M. Feuchtnner *et al.*, "Cardiac Computed Tomography: State of the Art and Future Horizons," *Journal of Clinical Medicine*, vol. 11, no. 15, 2022, doi: 10.3390/jcm11154429.
- [93]Z. Sun, "Cardiac CT imaging in coronary artery disease: Current status and future directions.," *Quant Imaging Med Surg*, vol. 2, no. 2, pp. 98–105, Jun. 2012, doi: 10.3978/j.issn.2223-4292.2012.05.02.
- [94]N. T. Singh, C. Dhadli, A. Chaudhary, and S. Goyal, *Preprocessing of Medical Images using Deep Learning: A Comprehensive Review*. 2023, p. 527. doi: 10.1109/ICAISS58487.2023.10250462.
- [95]K. C. Prabu Shankar and S. Prayla Shyry, "RETRACTED: A Survey of image pre-processing techniques for medical images," *Journal of Physics: Conference Series*, vol. 1911, no. 1, p. 012003, May 2021, doi: 10.1088/1742-6596/1911/1/012003.
- [96]T. B. Nguyen-Tat, T. Q. Hung, P. T. Nam, and V. M. Ngo, "Evaluating pre-processing and deep learning methods in medical imaging: Combined effectiveness across multiple modalities," *Alexandria Engineering Journal*, vol. 119, pp. 558–586, Apr. 2025, doi: 10.1016/j.aej.2025.01.090.

- [97] A. C. Bovik, "2.1 - Basic Gray-Level Image Processing," in *Handbook of Image and Video Processing (Second Edition)*, A. BOVIK, Ed., Burlington: Academic Press, 2005, pp. 21–37. doi: 10.1016/B978-012119792-6/50066-8.
- [98] A. Kour, V. Yv, V. Maheshwari, and D. Prashar, "A Review on Image Processing," *International Journal of Electronics Communication and Computer Engineering*, vol. 4, pp. 2278–4209, Jan. 2012.
- [99] N. Salem, H. Malik, and A. Shams, "Medical image enhancement based on histogram algorithms," *Procedia Computer Science*, vol. 163, pp. 300–311, Jan. 2019, doi: 10.1016/j.procs.2019.12.112.
- [100] O. Patel, P. Maravi, and S. Sharma, "A Comparative Study of Histogram Equalization Based Image Enhancement Techniques for Brightness Preservation and Contrast Enhancement," *Signal & Image Processing: An International Journal*, vol. 4, Nov. 2013, doi: 10.5121/sipij.2013.4502.
- [101] W. Zhihong and X. Xiaohong, "Study on Histogram Equalization," in *2011 2nd International Symposium on Intelligence Information Processing and Trusted Computing*, Oct. 2011, pp. 177–179. doi: 10.1109/IPTC.2011.52.
- [102] H. Zhu, F. H. Y. Chan, and F. K. Lam, "Image Contrast Enhancement by Constrained Local Histogram Equalization," *Computer Vision and Image Understanding*, vol. 73, no. 2, pp. 281–290, Feb. 1999, doi: 10.1006/cviu.1998.0723.
- [103] S. Bezryadin, P. Bourov, and D. Ilinih, "Brightness Calculation in Digital Image Processing," *International Symposium on Technologies for Digital Photo Fulfillment*, vol. 2007, pp. 10–15, Jan. 2007, doi: 10.2352/ISSN.2169-4672.2007.1.0.10.
- [104] N. Singh and A. Bhandari, "Image Contrast Enhancement with Brightness Preservation Using an Optimal Gamma and Logarithmic Approach," *IET Image Processing*, vol. 14, Feb. 2020, doi: 10.1049/iet-ipr.2019.0921.
- [105] A. Singh, "A COMPREHENSIVE METHOD FOR IMAGE CONTRAST ENHANCEMENT BASED ON GLOBAL –LOCAL CONTRAST AND LOCAL STANDARD DEVIATION," *International Journal of Research in Engineering and Technology*, vol. 03, pp. 413–416, Aug. 2014, doi: 10.15623/ijret.2014.0308064.
- [106] H. H. Draz, N. E. Elashker, and M. M. A. Mahmoud, "Optimized Algorithms and Hardware Implementation of Median Filter for Image Processing," *Circuits, Systems, and Signal Processing*, vol. 42, no. 9, pp. 5545–5558, Sep. 2023, doi: 10.1007/s00034-023-02370-x.
- [107] P. Li, H. Wang, M. Yu, and Y. Li, "Overview of Image Smoothing Algorithms," *Journal of Physics: Conference Series*, vol. 1883, p. 012024, Apr. 2021, doi: 10.1088/1742-6596/1883/1/012024.
- [108] Y. Zhu and C. Huang, "An Improved Median Filtering Algorithm for Image Noise Reduction," *Physics Procedia*, vol. 25, pp. 609–616, Dec. 2012, doi: 10.1016/j.phpro.2012.03.133.
- [109] G. Dong and K. Palaniappan, "A Robust Method for Edge-Preserving Image Smoothing," in *Advanced Concepts for Intelligent Vision Systems*, J. Blanc-Talon, S. Bourennane, W. Philips, D. Popescu, and P. Scheunders, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 390–399.
- [110] R. Raju, T. Maul, and A. Bargiela, "New image processing pipelines for membrane detection," *Journal of the Institute of Industrial Applications Engineers*, vol. 3, pp. 15–23, Jan. 2015, doi: 10.12792/JIAE.3.15.
- [111] S. Saponara and A. Elhanashi, "Impact of Image Resizing on Deep Learning Detectors for Training Time and Model Performance," 2022, pp. 10–17. doi: 10.1007/978-3-030-95498-7_2.
- [112] E. M. McKenzie, N. Tong, D. Ruan, M. Cao, R. K. Chin, and K. Sheng, "Using neural networks to extend cropped medical images for deformable registration among images with differing scan extents.," *Med Phys*, vol. 48, no. 8, pp. 4459–4471, Aug. 2021, doi: 10.1002/mp.15039.

- [113] P. Tarasiuk and P. S. Szczepaniak, "Novel convolutional neural networks for efficient classification of rotated and scaled images," *Neural Computing and Applications*, vol. 34, no. 13, pp. 10519–10532, Jul. 2022, doi: 10.1007/s00521-021-06645-9.
- [114] J. Deng, W. Dong, R. Socher, L. -J. Li, Kai Li, and Li Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, Jun. 2009, pp. 248–255. doi: 10.1109/CVPR.2009.5206848.
- [115] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," *Neural Information Processing Systems*, vol. 25, Jan. 2012, doi: 10.1145/3065386.
- [116] E. Goceri, "Medical image data augmentation: techniques, comparisons and interpretations," *Artificial Intelligence Review*, vol. 56, no. 11, pp. 12561–12605, Nov. 2023, doi: 10.1007/s10462-023-10453-z.
- [117] J. L. Flores, J. Ferrari, A. Fernández, J. Alonso, G. Ayubi, and M. Di Martino, "Methods for edge enhancement in color images based on derivative operations," *Proceedings of SPIE - The International Society for Optical Engineering*, vol. 8511, p. 15, Oct. 2012, doi: 10.1117/12.935618.
- [118] R. Gonzalez and Z. Faisal, *Digital Image Processing Second Edition*. 2019.
- [119] J. A. M. Saif, M. Hammad, and I. Alqubati, "Gradient Based Image Edge Detection," *International Journal of Engineering and Technology*, vol. 8, pp. 153–156, Mar. 2016, doi: 10.7763/IJET.2016.V6.876.
- [120] L. Ding and A. Goshtasby, "On the Canny edge detector," *Pattern Recognition*, vol. 34, no. 3, pp. 721–725, Mar. 2001, doi: 10.1016/S0031-3203(00)00023-6.
- [121] W.-C. Lin and J.-W. Wang, "Edge detection in medical images with quasi high-pass filter based on local statistics," *Biomedical Signal Processing and Control*, vol. 39, pp. 294–302, Jan. 2018, doi: 10.1016/j.bspc.2017.08.011.
- [122] D. Mistry and A. Banerjee, "Review: Image Registration," *International Journal of Graphics and Image Processing (ISSN 2249 – 5452)*, vol. II, pp. 18–22, Feb. 2012.
- [123] B. Zitová and J. Flusser, "Image registration methods: a survey," *Image and Vision Computing*, vol. 21, no. 11, pp. 977–1000, Oct. 2003, doi: 10.1016/S0262-8856(03)00137-9.
- [124] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, Nov. 2004, doi: 10.1023/B:VISI.0000029664.99615.94.
- [125] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-Up Robust Features (SURF)," *Computer Vision and Image Understanding*, vol. 110, no. 3, pp. 346–359, Jun. 2008, doi: 10.1016/j.cviu.2007.09.014.
- [126] P. Viola and W. M. Wells III, "Alignment by Maximization of Mutual Information," *International Journal of Computer Vision*, vol. 24, no. 2, pp. 137–154, Sep. 1997, doi: 10.1023/A:1007958904918.
- [127] D. Rueckert, L. Sonoda, C. Hayes, D. Hill, M. O. Leach, and D. J. Hawkes, "Nonrigid Registration Using Free-Form Deformations: Application to Breast MR Images," *Medical Imaging, IEEE Transactions on*, vol. 18, pp. 712–721, Sep. 1999, doi: 10.1109/42.796284.
- [128] G. Balakrishnan, A. Zhao, M. R. Sabuncu, J. Guttag, and A. V. Dalca, "VoxelMorph: A Learning Framework for Deformable Medical Image Registration," *IEEE Transactions on Medical Imaging*, vol. 38, no. 8, pp. 1788–1800, Aug. 2019, doi: 10.1109/TMI.2019.2897538.
- [129] P. Yan, S. Xu, A. Rastinehad, and B. Wood, "Adversarial Image Registration with Application for MR and TRUS Image Fusion," 2018, pp. 197–204. doi: 10.1007/978-3-030-00919-9_23.
- [130] Z. Król, "Registration of Intramodal Medical Images Using a Novel s-Distance Approach," in *Bildverarbeitung für die Medizin 2000*, A. Horsch and T. Lehmann, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2000, pp. 63–67.
- [131] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Mutual-information-based registration of medical images: a survey," *IEEE Trans Med Imaging*, vol. 22, no. 8, pp. 986–1004, Aug. 2003, doi: 10.1109/TMI.2003.815867.

- [132] J. M. Fitzpatrick and J. B. West, "The distribution of target registration error in rigid-body point-based registration.," *IEEE Trans Med Imaging*, vol. 20, no. 9, pp. 917–927, Sep. 2001, doi: 10.1109/42.952729.
- [133] Y. M. Wong *et al.*, "Machine learning prediction of Dice similarity coefficient for validation of deformable image registration," *Intelligence-Based Medicine*, vol. 10, p. 100163, Jan. 2024, doi: 10.1016/j.ibmed.2024.100163.
- [134] G. Haskins, U. Kruger, and P. Yan, "Deep learning in medical image registration: a survey," *Machine Vision and Applications*, vol. 31, Jan. 2020, doi: 10.1007/s00138-020-01060-x.
- [135] Y. N. Wu, "Data Augmentation," in *Computer Vision: A Reference Guide*, K. Ikeuchi, Ed., Cham: Springer International Publishing, 2021, pp. 239–240. doi: 10.1007/978-3-030-63416-2_741.
- [136] S. Volkova, "An Overview on Data Augmentation for Machine Learning," in *Digital and Information Technologies in Economics and Management*, A. Gibadullin, Ed., Cham: Springer Nature Switzerland, 2024, pp. 143–154.
- [137] J. Gallier, *Geometric Methods and Applications For Computer Science and Engineering*, vol. 38. 2011. doi: 10.1007/978-1-4419-9961-0.
- [138] A. Mumuni and F. Mumuni, "Data augmentation: A comprehensive survey of modern approaches," *Array*, vol. 16, p. 100258, Dec. 2022, doi: 10.1016/j.array.2022.100258.
- [139] P. Ryan, *Euclidean and Non-Euclidean Geometry: An Analytic Approach*. 2012. doi: 10.1017/CBO9780511806209.
- [140] Y. xu *et al.*, *Improved Relation Classification by Deep Recurrent Neural Networks with Data Augmentation*. 2016.
- [141] S. Wong, A. Gatt, V. Stamatescu, and M. McDonnell, *Understanding Data Augmentation for Classification: When to Warp?* 2016, p. 6. doi: 10.1109/DICTA.2016.7797091.
- [142] M. I. Bellil and X. Xu, "Affine and Magnitude Transforms Based Data Augmentation Comparison for SAR Automatic Target Recognition," in *2018 IEEE 4th International Conference on Computer and Communications (ICCC)*, Dec. 2018, pp. 1649–1653. doi: 10.1109/CompComm.2018.8780691.
- [143] X. Dong, M. Potter, G. Kumar, Y.-C. Tsai, and V. Saripalli, *Automating Augmentation Through Random Unidimensional Search*. 2021. doi: 10.48550/arXiv.2106.08756.
- [144] F. Milletari, N. Navab, and S. -A. Ahmadi, "V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation," in *2016 Fourth International Conference on 3D Vision (3DV)*, Oct. 2016, pp. 565–571. doi: 10.1109/3DV.2016.79.
- [145] P. Simard, D. Steinkraus, and J. Platt, *Best Practices for Convolutional Neural Networks Applied to Visual Document Analysis*. 2003, p. 962. doi: 10.1109/ICDAR.2003.1227801.
- [146] K. Wang, B. Fang, J. Qian, S. Yang, X. Zhou, and J. Zhou, "Perspective Transformation Data Augmentation for Object Detection," *IEEE Access*, vol. PP, pp. 1–1, Dec. 2019, doi: 10.1109/ACCESS.2019.2962572.
- [147] M. Franke, V. Gopinath, C. Reddy, D. Ristic-Durrant, and K. Michels, *Bounding Box Dataset Augmentation for Long-range Object Distance Estimation*. 2021, p. 1677. doi: 10.1109/ICCVW54120.2021.00192.
- [148] N. E. Khalifa, M. Loey, and S. Mirjalili, "A comprehensive survey of recent trends in deep learning for digital images augmentation," *Artificial Intelligence Review*, vol. 55, no. 3, pp. 2351–2377, Mar. 2022, doi: 10.1007/s10462-021-10066-4.
- [149] C. Gormond, "Influence of Photometric Data Augmentation," Sep. 2021.
- [150] N. R. Pal and S. K. Pal, "A review on image segmentation techniques," *Pattern Recognition*, vol. 26, no. 9, pp. 1277–1294, Sep. 1993, doi: 10.1016/0031-3203(93)90135-J.
- [151] S. Prabu and J. M. Gnanasekar, "A Study on Image Segmentation Method for Image Processing," 2021. doi: 10.3233/APC210223.

- [152] S. C. Zhu and A. Yuille, "Region Competition: Unifying Snakes, Region Growing, and Bayes/MDL for Multiband Image Segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 9, pp. 884–900, Sep. 1996, doi: 10.1109/34.537343.
- [153] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds., Cham: Springer International Publishing, 2015, pp. 234–241.
- [154] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2015, pp. 3431–3440. doi: 10.1109/CVPR.2015.7298965.
- [155] J. Russ, "The image processing handbook. 6th ed," Apr. 2016, doi: 10.1201/b10720.
- [156] E. Davies, *Computer and Machine Vision: Theory, Algorithms, Practicalities*. 2012.
- [157] F. A. and J. Ponce, "Computer Vision: A Modern Approach Prentice Hall," Jan. 2002.
- [158] J. -w. Jang, S. Lee, H. -J. Hwang, and K. -R. Baek, "Global thresholding algorithm based on boundary selection," in *2013 13th International Conference on Control, Automation and Systems (ICCAS 2013)*, Oct. 2013, pp. 704–706. doi: 10.1109/ICCAS.2013.6703961.
- [159] N. Otsu, "A Threshold Selection Method from Gray-Level Histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, Jan. 1979, doi: 10.1109/TSMC.1979.4310076.
- [160] Z.-K. Huang and K.-W. Chau, "A new image thresholding method based on Gaussian mixture model," *Applied Mathematics and Computation*, vol. 205, no. 2, pp. 899–907, Nov. 2008, doi: 10.1016/j.amc.2008.05.130.
- [161] F. Bovolo and L. Bruzzone, *An adaptive thresholding approach to multiple-change detection in multispectral images*. 2011, p. 236. doi: 10.1109/IGARSS.2011.6048935.
- [162] P. Slavíková, M. Mudrová, J. Petrová, J. Fojt, L. Joska, and A. Procházka, "Automatic characterization of titanium dioxide nanotubes by image processing of scanning electron microscopic images," *Nanomaterials and Nanotechnology*, vol. 6, pp. 1–13, Nov. 2016, doi: 10.1177/1847980416673784.
- [163] C. Zheng and D.-W. Sun, "2 - Image Segmentation Techniques," in *Computer Vision Technology for Food Quality Evaluation*, D.-W. Sun, Ed., Amsterdam: Academic Press, 2008, pp. 37–56. doi: 10.1016/B978-012373642-0.50005-3.
- [164] W. Fechner, E. Gselmann, and A. Świątczak-Kolenda, "Characterizations of second-order differential operators," *Bulletin of the Malaysian Mathematical Sciences Society*, vol. 48, no. 2, p. 47, Jan. 2025, doi: 10.1007/s40840-025-01823-7.
- [165] A. Shehata, S. Mohammad, M. Abdallah, and M. Ragab, "A Survey on Hough Transform, Theory, Techniques and Applications," Feb. 2015.
- [166] V. Cantoni and E. Mattia, "Hough Transform," in *Encyclopedia of Systems Biology*, W. Dubitzky, O. Wolkenhauer, K.-H. Cho, and H. Yokota, Eds., New York, NY: Springer New York, 2013, pp. 917–918. doi: 10.1007/978-1-4419-9863-7_1310.
- [167] N. Ikonomatakis, K. Plataniotis, M. Zervakis, and A. N. Venetsanopoulos, *Region growing and region merging image segmentation*. 1997, p. 302 vol.1. doi: 10.1109/ICDSP.1997.628077.
- [168] E. Sobotnicka, A. Sobotnicki, K. Horoba, and P. Porwik, *The Application of the Region Growing Method to the Determination of Arterial Changes*, vol. 2. 2016, p. 471. doi: 10.1007/978-3-319-45246-3_44.
- [169] H. Rahman, F. Din, S. Ūr Rahman, and S. Sehatullah, "Performance evaluation of region-growing based segmentation algorithms for segmenting the aorta," *Jurnal Teknologi*, vol. 78, Apr. 2016, doi: 10.11113/jt.v78.8226.
- [170] S. Osher and J. A. Sethian, "Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations," *Journal of Computational Physics*, vol. 79, no. 1, pp. 12–49, Nov. 1988, doi: 10.1016/0021-9991(88)90002-2.
- [171] S. Osher and R. Fedkiw, "The Level Set Methods and Dynamic Implicit Surfaces," in *Applied Mechanics Reviews*, vol. 57, 2004, p. xiv+273. doi: 10.1115/1.1760520.

- [172] D. Cremers, M. Rousson, and R. Deriche, “A Review of Statistical Approaches to Level Set Segmentation: Integrating Color, Texture, Motion and Shape,” *International Journal of Computer Vision*, vol. 72, no. 2, pp. 195–215, Apr. 2007, doi: 10.1007/s11263-006-8711-1.
- [173] M. Kass, A. Witkin, and D. Terzopoulos, “Snakes: Active contour models,” *International Journal of Computer Vision*, vol. 1, no. 4, pp. 321–331, Jan. 1988, doi: 10.1007/BF00133570.
- [174] R. Szeliski, *Computer Vision: Algorithms and Applications*. 2022. doi: 10.1007/978-3-030-34372-9.
- [175] Y. Chen, P. Ge, G. Wang, G. Weng, and H. Chen, “An overview of intelligent image segmentation using active contour models,” *Intelligence & Robotics*, vol. 3, pp. 23–55, Feb. 2023, doi: 10.20517/ir.2023.02.
- [176] X. Shan, X. Gong, Y. Ren, and A. Nandi, “Image Segmentation Using an Active Contour Model Based on the Difference Between Local Intensity Averages and Actual Image Intensities,” *IEEE Access*, vol. PP, pp. 1–1, Feb. 2020, doi: 10.1109/ACCESS.2020.2975854.
- [177] G. Kühne, J. Weickert, M. Beier, and W. Effelsberg, *Fast Implicit Active Contour Models*, vol. 2449. 2002, p. 140. doi: 10.1007/3-540-45783-6_17.
- [178] C. Li, C.-Y. Kao, J. Gore, and Z. Ding, *Implicit Active Contours Driven by Local Binary Fitting Energy*, vol. 1–7. 2007, p. 7. doi: 10.1109/CVPR.2007.383014.
- [179] H.-R. Jung and M.-R. Jung, “An Implicit Active Contour Model for Feature Regions and Lines,” in *Advances in Multimedia Modeling*, S. Satoh, F. Nack, and M. Etoh, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 35–44.
- [180] M. Savelonas, D. Maroulis, D. Iakovidis, and S. Karkanis, *Image Segmentation Using Local Binary Pattern and Active Contour Model without Edges*. 2006.
- [181] C. Li, C. Xu, C. Gui, and M. D. Fox, *Level set evolution without re-initialization: A new variational formulation*, vol. 1. 2005, p. 436 vol. 1. doi: 10.1109/CVPR.2005.213.
- [182] S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A. Yezzi, *Gradient flows and geometric active contour models*. 1995, p. 815. doi: 10.1109/ICCV.1995.466855.
- [183] V. Caselles, R. Kimmel, and G. Sapiro, “Geodesic Active Contours,” *International Journal of Computer Vision*, vol. 22, no. 1, pp. 61–79, Feb. 1997, doi: 10.1023/A:1007979827043.
- [184] Chenyang Xu and J. L. Prince, “Snakes, shapes, and gradient vector flow,” *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 359–369, Mar. 1998, doi: 10.1109/83.661186.
- [185] A. Tsai, A. Yezzi, and A. S. Willsky, “Curve evolution implementation of the Mumford-Shah functional for image segmentation, denoising, interpolation, and magnification,” *IEEE Transactions on Image Processing*, vol. 10, no. 8, pp. 1169–1186, Aug. 2001, doi: 10.1109/83.935033.
- [186] L. A. Vese and T. F. Chan, “A Multiphase Level Set Framework for Image Segmentation Using the Mumford and Shah Model,” *International Journal of Computer Vision*, vol. 50, no. 3, pp. 271–293, Dec. 2002, doi: 10.1023/A:1020874308076.
- [187] M. Allili and D. Ziou, “Globally adaptive region information for automatic color–texture image segmentation,” *Pattern Recognition Letters*, vol. 28, pp. 1946–1956, Nov. 2007, doi: 10.1016/j.patrec.2007.05.002.
- [188] Chenyang Xu, A. Yezzi, and J. L. Prince, “On the relationship between parametric and geometric active contours,” in *Conference Record of the Thirty-Fourth Asilomar Conference on Signals, Systems and Computers (Cat. No.00CH37154)*, Nov. 2000, pp. 483–489 vol.1. doi: 10.1109/ACSSC.2000.911003.
- [189] W. Zhou, X. Du, and S. Wang, *Techniques for Image Segmentation Based on Edge Detection*. 2021, p. 403. doi: 10.1109/CEI52496.2021.9574569.
- [190] I. Tamanini, “Optimal Approximation by Piecewise Constant Functions,” in *Variational Methods for Discontinuous Structures*, R. Serapioni and F. Tomarelli, Eds., Basel: Birkhäuser Basel, 1996, pp. 73–85.

- [191] R. Arulmurugan and A. Haldorai, "Region-based seed point cell segmentation and detection for biomedical image analysis," *International Journal of Biomedical Engineering and Technology*, vol. 27, p. 273, Jan. 2018, doi: 10.1504/IJBET.2018.094296.
- [192] K. Zhang, L. Zhang, H. Song, and W. Zhou, "Active contours with selective local or global segmentation: A new formulation and level set method," *Image and Vision Computing*, vol. 28, pp. 668–676, Apr. 2010, doi: 10.1016/j.imavis.2009.10.009.
- [193] K. Zhang, L. Zhang, and S. Zhang, *A variational multiphase level set approach to simultaneous segmentation and bias correction*. 2010, p. 4108. doi: 10.1109/ICIP.2010.5651554.
- [194] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 7, pp. 629–639, Jul. 1990, doi: 10.1109/34.56205.
- [195] J. Piovano, M. Rousson, and T. Papadopoulos, *Efficient Segmentation of Piecewise Smooth Images*, vol. 44885. 2007, p. 720. doi: 10.1007/978-3-540-72823-8_61.
- [196] K.-Y. Ni, X. Bresson, T. Chan, and S. Esedoglu, "Local Histogram Based Segmentation Using the Wasserstein Distance," *International Journal of Computer Vision*, vol. 84, pp. 97–111, Aug. 2009, doi: 10.1007/s11263-009-0234-0.
- [197] K. Zhang, H. Song, and L. Zhang, "Active contours driven by local image fitting energy," *Pattern Recognition*, vol. 43, pp. 1199–1206, Apr. 2010, doi: 10.1016/j.patcog.2009.10.010.
- [198] D. Li, W. Li, and Q. Liao, "Active contours driven by local probability distributions," *Journal of Visual Communication and Image Representation*, vol. 24, pp. 522–533, Jul. 2013, doi: 10.1016/j.jvcir.2013.03.007.
- [199] S. Osher and N. Paragios, *Geometric Level Set Methods in Imaging, Vision, and Graphics*, vol. xxix. 2003. doi: 10.1007/b97541.
- [200] Y. Tian, F. Duan, M. Zhou, and zk wu, "Active contour model combining region and edge information," *Machine Vision and Applications - MVA*, vol. 24, pp. 1–15, Aug. 2011, doi: 10.1007/s00138-011-0363-7.
- [201] S. Banerjee and M. Bhattacharya, "Segmentation of medical images using Selective Binary and Gaussian Filtering regularized level set (SBGFRLS) method," in *2010 3rd International Conference on Biomedical Engineering and Informatics*, Oct. 2010, pp. 541–545. doi: 10.1109/BMEI.2010.5639990.
- [202] S. Tatiraju and A. Mehta, "Image Segmentation using k-means clustering, EM and Normalized Cuts," *University of California Irvine*, Jan. 2008.
- [203] N. Dhanachandra, K. Manglem, and Y. J. Chanu, "Image Segmentation Using K - means Clustering Algorithm and Subtractive Clustering Algorithm," *Procedia Computer Science*, vol. 54, pp. 764–771, Jan. 2015, doi: 10.1016/j.procs.2015.06.090.
- [204] K. K. D. Ramesh, K. GUTHIKONDA, Dr. K. SWAPNA, D. Datta, and S. Rajest, "A Review of Medical Image Segmentation Algorithms," *EAI Endorsed Transactions on Pervasive Health and Technology*, Apr. 2021, doi: 10.4108/eai.12-4-2021.169184.
- [205] C. Mavridis, T. L. Economopoulos, G. Benetos, and G. K. Matsopoulos, "Aorta Segmentation in 3D CT Images by Combining Image Processing and Machine Learning Techniques.," *Cardiovasc Eng Technol*, vol. 15, no. 3, pp. 359–373, Jun. 2024, doi: 10.1007/s13239-024-00720-7.
- [206] L. Čomić, L. De Floriani, and L. Papaleo, "Morse-Smale Decompositions for Modeling Terrain Knowledge," in *Spatial Information Theory*, A. G. Cohn and D. M. Mark, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, pp. 426–444.
- [207] V. Grau, A. U. J. Mewes, M. Alcañiz Raya, R. Kikinis, and S. Warfield, "Improved Watershed Transform for Medical Image Segmentation Using Prior Information," *IEEE transactions on medical imaging*, vol. 23, pp. 447–58, May 2004, doi: 10.1109/TMI.2004.824224.
- [208] K. Nallaperumal and K. Krishnaveni, "A Multiscale Morphological Watershed Segmentation using Color Composite Gradient and Marker Extraction," 2008. [Online]. Available: <https://api.semanticscholar.org/CorpusID:5720353>

- [209] D. Goshal and P. P. Acharjya, "MRI Image Segmentation Using Watershed Transform," 2012. [Online]. Available: <https://api.semanticscholar.org/CorpusID:16720063>
- [210] J. Zhou, Y. Yin, and S. Wang, "Image Segmentation Based on Watershed Algorithm," in *2021 International Conference on Intelligent Computing, Automation and Applications (ICAA)*, Jun. 2021, pp. 10–13. doi: 10.1109/ICAA53760.2021.00010.
- [211] L. K V, A. .R, S. Michahial, and D. M, "Implementation of Watershed Segmentation," *IJARCCCE*, vol. 5, pp. 196–199, Dec. 2016, doi: 10.17148/IJARCCCE.2016.51243.
- [212] X. Zhang, Y. Shan, W. Wei, and Z. Zhu, "An Image Segmentation Method Based on Improved Watershed Algorithm," in *2010 International Conference on Computational and Information Sciences*, Dec. 2010, pp. 258–261. doi: 10.1109/ICCIS.2010.69.
- [213] S. Beucher and C. Mathmatique, "The Watershed Transformation Applied To Image Segmentation," *Scanning. Microsc.*, vol. 6, Jul. 2000.
- [214] J. Smółka and M. Skublewska-Paszkowska, "Watershed merging method for color images," *Annales UMCS Informatica*, vol. AI 8, pp. 111–121, Jan. 2008, doi: 10.2478/v10065-008-0011-7.
- [215] Y. Wu and Q. Li, "The Algorithm of Watershed Color Image Segmentation Based on Morphological Gradient," *Sensors*, vol. 22, no. 21, 2022, doi: 10.3390/s22218202.
- [216] Z. Kato, *Markov Random Fields in Image Segmentation*. 2011. doi: 10.1561/9781601985897.
- [217] Z. Kato and T.-C. Pong, "A Markov random field image segmentation model for color textured images," *Image and Vision Computing*, vol. 24, pp. 1103–1114, Oct. 2006, doi: 10.1016/j.imavis.2006.03.005.
- [218] C. Wang, N. Komodakis, and N. Paragios, "Markov Random Field modeling, inference & learning in computer vision & image understanding: A survey," *Computer Vision and Image Understanding*, vol. 117, no. 11, pp. 1610–1627, Nov. 2013, doi: 10.1016/j.cviu.2013.07.004.
- [219] *ITK: itk::MRImageFilter< TInputImage, TClassifiedImage > Class Template Reference. (n.d.). Insight ToolKit. Retrieved September 13, 2021, from https://itk.org/Doxygen/html/classitk_1_1MRImageFilter.html.*
- [220] O. Simeone, "A Very Brief Introduction to Machine Learning With Applications to Communication Systems," *IEEE Transactions on Cognitive Communications and Networking*, vol. 4, no. 4, pp. 648–664, Dec. 2018, doi: 10.1109/TCCN.2018.2881442.
- [221] S. Naem, A. Ali, S. Anam, and M. Ahmed, "An Unsupervised Machine Learning Algorithms: Comprehensive Review," *IJCDS Journal*, vol. 13, pp. 911–921, Apr. 2023, doi: 10.12785/ijcds/130172.
- [222] V. Nasteski, "An overview of the supervised machine learning methods," *HORIZONS.B*, vol. 4, pp. 51–62, Dec. 2017, doi: 10.20544/HORIZONS.B.04.1.17.P05.
- [223] A. Srinivasan, "Reinforcement Learning: Advancements, Limitations, and Real-world Applications," *INTERANTIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT*, vol. 07, Aug. 2023, doi: 10.55041/IJSREM25118.
- [224] J. G. Carbonell, R. S. Michalski, and T. M. Mitchell, "1 - AN OVERVIEW OF MACHINE LEARNING," in *Machine Learning*, R. S. Michalski, J. G. Carbonell, and T. M. Mitchell, Eds., San Francisco (CA): Morgan Kaufmann, 1983, pp. 3–23. doi: 10.1016/B978-0-08-051054-5.50005-4.
- [225] T. Ayodele, "Machine Learning Overview," 2010. doi: 10.5772/9374.
- [226] I. H. Sarker, "Machine Learning: Algorithms, Real-World Applications and Research Directions," *SN Computer Science*, vol. 2, no. 3, p. 160, Mar. 2021, doi: 10.1007/s42979-021-00592-x.
- [227] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Networks*, vol. 61, pp. 85–117, Jan. 2015, doi: 10.1016/j.neunet.2014.09.003.
- [228] I. H. Sarker, "Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions," *SN Computer Science*, vol. 2, no. 6, p. 420, Aug. 2021, doi: 10.1007/s42979-021-00815-1.
- [229] H. Shahinzadeh, A. Mahmoudi, A. Asilian, H. Sadrarhami, M. Hemmati, and Y. Saberi, "Deep Learning: A Overview of Theory and Architectures," in *2024 20th CSI International*

- Symposium on Artificial Intelligence and Signal Processing (AISP)*, Feb. 2024, pp. 1–11. doi: 10.1109/AISP61396.2024.10475265.
- [230] Z. Hao, “Deep learning review and discussion of its future development,” *MATEC Web of Conferences*, vol. 277, p. 02035, Jan. 2019, doi: 10.1051/mateconf/201927702035.
 - [231] K. O’Shea and R. Nash, “An Introduction to Convolutional Neural Networks,” *ArXiv e-prints*, Nov. 2015.
 - [232] J. Nagi *et al.*, *Max-pooling convolutional neural networks for vision-based hand gesture recognition*. 2011, p. 347. doi: 10.1109/ICSIPA.2011.6144164.
 - [233] S. R. Dubey, S. K. Singh, and B. Chaudhuri, “Activation Functions in Deep Learning: A comprehensive Survey and Benchmark,” *Neurocomputing*, vol. 503, Jul. 2022, doi: 10.1016/j.neucom.2022.06.111.
 - [234] X. Liu, Z. Deng, and Y. Yang, “Recent progress in semantic image segmentation,” *Artif. Intell. Rev.*, vol. 52, no. 2, pp. 1089–1106, Aug. 2019, doi: 10.1007/s10462-018-9641-3.
 - [235] S.-Y. Huang, W.-L. Hsu, R.-J. Hsu, and D.-W. Liu, “Fully Convolutional Network for the Semantic Segmentation of Medical Images: A Survey,” *Diagnostics*, vol. 12, no. 11, 2022, doi: 10.3390/diagnostics12112765.
 - [236] Z. Niu, G. Zhong, and H. Yu, “A review on the attention mechanism of deep learning,” *Neurocomputing*, vol. 452, pp. 48–62, Sep. 2021, doi: 10.1016/j.neucom.2021.03.091.
 - [237] D. Bahdanau, K. Cho, and Y. Bengio, “Neural Machine Translation by Jointly Learning to Align and Translate,” *ArXiv*, vol. 1409, Sep. 2014.
 - [238] A. Vaswani *et al.*, “Attention is All you Need,” in *Neural Information Processing Systems*, 2017. [Online]. Available: <https://api.semanticscholar.org/CorpusID:13756489>
 - [239] Y. Liu, N. Yang, S. Yao, and X. Wang, “Additive Attention for Medical Visual Question Answering,” in *2023 5th International Conference on Frontiers Technology of Information and Computer (ICFTIC)*, Nov. 2023, pp. 932–935. doi: 10.1109/ICFTIC59930.2023.10456252.
 - [240] K. T. Chitty-Venkata, M. Emani, V. Vishwanath, and A. K. Somani, “Neural Architecture Search for Transformers: A Survey,” *IEEE Access*, vol. 10, pp. 108374–108412, 2022, doi: 10.1109/ACCESS.2022.3212767.
 - [241] D. Soydaner, “Attention mechanism in neural networks: where it comes and where it goes,” *Neural Computing and Applications*, vol. 34, May 2022, doi: 10.1007/s00521-022-07366-3.
 - [242] Z. Yang, J. Cen, X. Liu, J. Xiong, and H. Chen, “Research on bearing fault diagnosis method based on transformer neural network,” *Measurement Science and Technology*, vol. 33, May 2022, doi: 10.1088/1361-6501/ac66c4.
 - [243] J. Zhu *et al.*, “Efficient self-attention mechanism and structural distilling model for Alzheimer’s disease diagnosis,” *Computers in Biology and Medicine*, vol. 147, p. 105737, Aug. 2022, doi: 10.1016/j.compbiomed.2022.105737.
 - [244] T. Isomura, R. Shimizu, and M. Goto, “Sparse attention is all you need for pre-training on tabular data,” *Neural Computing and Applications*, vol. 37, no. 3, pp. 1509–1522, Jan. 2025, doi: 10.1007/s00521-024-10698-x.
 - [245] J. Ho, N. Kalchbrenner, D. Weißenborn, and T. Salimans, *Axial Attention in Multidimensional Transformers*. 2019. doi: 10.48550/arXiv.1912.12180.
 - [246] A. Hernández and J. Amigó, “Attention Mechanisms and Their Applications to Complex Systems,” *Entropy*, vol. 23, p. 283, Feb. 2021, doi: 10.3390/e23030283.
 - [247] S. Islam *et al.*, “A comprehensive survey on applications of transformers for deep learning tasks,” *Expert Systems with Applications*, vol. 241, p. 122666, May 2024, doi: 10.1016/j.eswa.2023.122666.
 - [248] G. Bebis and M. Georgiopoulos, “Feed-forward neural networks,” *IEEE Potentials*, vol. 13, no. 4, pp. 27–31, Nov. 1994, doi: 10.1109/45.329294.
 - [249] T. T. W. Haryono, A. Zailani, Y. Djaksana, N. Rosmawarni, and N. Arianti, “Transformers in Machine Learning: Literature Review,” *Jurnal Penelitian Pendidikan IPA*, vol. 9, pp. 604–610, Sep. 2023, doi: 10.29303/jppipa.v9i9.5040.
 - [250] I. Sutskever, O. Vinyals, and Q. Le, “Sequence to Sequence Learning with Neural Networks,” *Advances in Neural Information Processing Systems*, vol. 4, Sep. 2014.

- [251] M. Costa-jussa, Á. Nuez, and C. Segura, “Experimental Research on Encoder-Decoder Architectures with Attention for Chatbots,” *Computación y Sistemas*, vol. 22, Dec. 2018, doi: 10.13053/cys-22-4-3060.
- [252] T. Theodoropoulos, A. -C. Maroudis, J. Violos, and K. Tserpes, “An Encoder-Decoder Deep Learning Approach for Multistep Service Traffic Prediction,” in *2021 IEEE Seventh International Conference on Big Data Computing Service and Applications (BigDataService)*, Aug. 2021, pp. 33–40. doi: 10.1109/BigDataService52369.2021.00010.
- [253] Z. Lipton, “A Critical Review of Recurrent Neural Networks for Sequence Learning,” May 2015.
- [254] S. Das, A. Tariq, T. Santos, S. S. Kantareddy, and I. Banerjee, “Recurrent Neural Networks (RNNs): Architectures, Training Tricks, and Introduction to Influential Research,” in *Machine Learning for Brain Disorders*, O. Colliot, Ed., New York, NY: Springer US, 2023, pp. 117–138. doi: 10.1007/978-1-0716-3195-9_4.
- [255] G. Bird and M. Polivoda, *Backpropagation Through Time For Networks With Long-Term Dependencies*. 2021. doi: 10.48550/arXiv.2103.15589.
- [256] B. Sistaninejhad, H. Rasi, and P. Nayeri, “A Review Paper about Deep Learning for Medical Image Analysis.,” *Comput Math Methods Med*, vol. 2023, p. 7091301, 2023, doi: 10.1155/2023/7091301.
- [257] V. Badrinarayanan, A. Kendall, and R. Cipolla, “SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation,” Dec. 2017, doi: 10.17863/CAM.17966.
- [258] I. Ullah, F. Ali, B. Shah, S. El-Sappagh, T. Abuhmed, and S. H. Park, “A deep learning based dual encoder–decoder framework for anatomical structure segmentation in chest X-ray images,” *Scientific Reports*, vol. 13, no. 1, p. 791, Jan. 2023, doi: 10.1038/s41598-023-27815-w.
- [259] F. Raza, *Transfer Learning in Deep Neural Networks*. 2023. doi: 10.13140/RG.2.2.25635.63529.
- [260] A. Ali, M. G. Yaseen, M. Aljanabi, S. Abed, and C. GPT, “Transfer Learning: A New Promising Techniques,” *Mesopotamian Journal of Big Data*, vol. 2023, pp. 31–32, Feb. 2023, doi: 10.58496/MJBD/2023/004.
- [261] A. Hosna, E. Merry, J. Gyalmo, Z. Alom, Z. Aung, and M. A. Azim, “Transfer learning: a friendly introduction,” *Journal of Big Data*, vol. 9, no. 1, p. 102, Oct. 2022, doi: 10.1186/s40537-022-00652-w.
- [262] A. Cosa, N. Santhanam, M. Jannesari, M. Maros, and T. Ganslandt, “Transfer learning for medical image classification: a literature review,” *BMC Medical Imaging*, vol. 22, Apr. 2022, doi: 10.1186/s12880-022-00793-7.
- [263] L. Zhou, M. Fan, C. Hansen, C. R. Johnson, and D. Weiskopf, “A Review of Three-Dimensional Medical Image Visualization,” *Health Data Science*, vol. 2022, Jan. 2022, doi: 10.34133/2022/9840519.
- [264] Y. Ma *et al.*, “Three-dimensional printing for heart diseases: clinical application review,” *Bio-Design and Manufacturing*, vol. 4, no. 3, pp. 675–687, Sep. 2021, doi: 10.1007/s42242-021-00125-8.
- [265] “Shashi P, R S. Review Study on Digital Image Processing and Segmentation. Am. J. Comput. Sci. Technol. 2019; <https://doi.org/10.11648/j.ajcst.20190204.14>”.
- [266] S. Khalid Abdulateef, S. R. Ahmed AHMED, and M. Dawood Salman, “A Novel Food Image Segmentation Based on Homogeneity Test of K-Means Clustering,” *IOP Conference Series: Materials Science and Engineering*, vol. 928, no. 3, p. 032059, Nov. 2020, doi: 10.1088/1757-899X/928/3/032059.
- [267] “Khan W. A Survey: Image Segmentation Techniques. Int. j. future comput. commun. 2014; <https://doi.org/10.7763/IJFCC.2014.V3.274>”.
- [268] F. Adolfi, T. Wareham, and I. van Rooij, “A Computational Complexity Perspective on Segmentation as a Cognitive Subcomputation,” *Topics in Cognitive Science*, vol. 15, no. 2, pp. 255–273, Apr. 2023, doi: 10.1111/tops.12629.
- [269] N. Hussain Dar, *Image segmentation Techniques and its application*. 2020.

- [270] Tortora GJ, Derrickson BH. *Principles of Anatomy and Physiology*. 15th ed. John Wiley & Sons: Hoboken; 2016.
- [271] Tortora GJ, Nielsen M. *Principles of Human Anatomy*. 14th ed. John Wiley & Sons: Hoboken; 2016.
- [272] Authors/Task Force members *et al.*, “2014 ESC Guidelines on the diagnosis and treatment of aortic diseases: Document covering acute and chronic aortic diseases of the thoracic and abdominal aorta of the adult The Task Force for the Diagnosis and Treatment of Aortic Diseases of the European Society of Cardiology (ESC),” *European Heart Journal*, vol. 35, no. 41, pp. 2873–2926, Nov. 2014, doi: 10.1093/eurheartj/ehu281.
- [273] T. Xu and I. Gondra, “A simple and effective texture characterization for image segmentation,” *Signal, Image and Video Processing*, vol. 6, no. 2, pp. 231–245, Jun. 2012, doi: 10.1007/s11760-010-0182-8.
- [274] D. Liu and J. Yu, “Otsu Method and K-means,” in *2009 Ninth International Conference on Hybrid Intelligent Systems*, Aug. 2009, pp. 344–349. doi: 10.1109/HIS.2009.74.
- [275] D. Arthur and S. Vassilvitskii, *K-Means++: The Advantages of Careful Seeding*, vol. 8. 2007, p. 1035. doi: 10.1145/1283383.1283494.
- [276] E. T. Lee, Y. J. Pan, and P. Chu, “An algorithm for region filling using two-dimensional grammars,” *International Journal of Intelligent Systems*, vol. 2, no. 3, pp. 255–263, Sep. 1987, doi: 10.1002/int.4550020302.
- [277] N. Xu, N. Ahuja, and R. Bansal, “Object segmentation using graph cuts based active contours,” *Computer Vision and Image Understanding*, vol. 107, no. 3, pp. 210–224, Sep. 2007, doi: 10.1016/j.cviu.2006.11.004.
- [278] Y. Y. Boykov and M. . -P. Jolly, “Interactive graph cuts for optimal boundary & region segmentation of objects in N-D images,” in *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, Jul. 2001, pp. 105–112 vol.1. doi: 10.1109/ICCV.2001.937505.
- [279] Y. Boykov and M.-P. Jolly, “Interactive Organ Segmentation Using Graph Cuts,” in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2000*, S. L. Delp, A. M. DiGoia, and B. Jaramaz, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2000, pp. 276–286.
- [280] V. Vezhnevets and V. Konushin, “‘GrowCut’ - Interactive Multi-Label ND Image Segmentation By Cellular Automata,” *Graphicon*, vol. 1, Nov. 2004.
- [281] A. Bieniek and A. Moga, “An efficient watershed algorithm based on connected components,” *Pattern Recognition*, vol. 33, no. 6, pp. 907–916, Jun. 2000, doi: 10.1016/S0031-3203(99)00154-5.
- [282] F. Lareyre, C. Adam, M. Carrier, C. Dommerc, C. Mialhe, and J. Raffort, “A fully automated pipeline for mining abdominal aortic aneurysm using image segmentation,” *Scientific Reports*, vol. 9, no. 1, p. 13750, Sep. 2019, doi: 10.1038/s41598-019-50251-8.
- [283] K. López-Linares *et al.*, “Fully automatic detection and segmentation of abdominal aortic thrombus in post-operative CTA images using Deep Convolutional Neural Networks,” *Medical Image Analysis*, vol. 46, pp. 202–214, May 2018, doi: 10.1016/j.media.2018.03.010.
- [284] M. Larsson, Y. Zhang, and F. Kahl, “Deepseg: Abdominal Organ Segmentation Using Deep Convolutional Neural Networks,” 2016.
- [285] Julia M. H. Noothout, Bob D. de Vos, Jelmer M. Wolterink, and Ivana Išgum, “Automatic segmentation of thoracic aorta segments in low-dose chest CT,” presented at the Proc.SPIE, Mar. 2018, p. 105741S. doi: 10.1117/12.2293114.
- [286] A. Fantazzini *et al.*, “3D Automatic Segmentation of Aortic Computed Tomography Angiography Combining Multi-View 2D Convolutional Neural Networks,” *Cardiovascular Engineering and Technology*, vol. 11, no. 5, pp. 576–586, Oct. 2020, doi: 10.1007/s13239-020-00481-z.
- [287] M. Müller-Eschner *et al.*, “3D morphometry using automated aortic segmentation in native MR angiography: an alternative to contrast enhanced MRA?,” *Cardiovasc Diagn Ther*, vol. 4, no. 2, pp. 80–87, Apr. 2014, doi: 10.3978/j.issn.2223-3652.2013.10.06.

- [288] L. S. Davis, A. Rosenfeld, and J. S. Weszka, "Region Extraction by Averaging and Thresholding," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-5, no. 3, pp. 383–388, May 1975, doi: 10.1109/TSMC.1975.5408419.
- [289] T. Lindeberg, "Discrete Scale-Space Theory and the Scale-Space Primal Sketch," 1991.
- [290] Johnson HJ, McCormick MM, Ibanez L. *The ITK Software Guide: Design and Functionality*. 4th ed. 2021.
- [291] P. Singh and R. S. Chadha, "A Novel Approach to Image Segmentation," 2013.
- [292] S. Geman and D. Geman, "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-6, no. 6, pp. 721–741, Nov. 1984, doi: 10.1109/TPAMI.1984.4767596.
- [293] H. Permuter, J. Francos, and I. Jermyn, "A study of Gaussian mixture models of color and texture features for image classification and segmentation," *Pattern Recognition*, vol. 39, no. 4, pp. 695–706, Apr. 2006, doi: 10.1016/j.patcog.2005.10.028.
- [294] A. K. Jain and F. Farrokhnia, "Unsupervised texture segmentation using Gabor filters," *Pattern Recognition*, vol. 24, no. 12, pp. 1167–1186, Jan. 1991, doi: 10.1016/0031-3203(91)90143-S.
- [295] Sangwine SJ, Horne REN. *The colour image processing handbook*. London: Chapman & Hall; 1998.
- [296] J. Besag, "On the Statistical Analysis of Dirty Pictures," *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 48, no. 3, pp. 259–279, Jul. 1986, doi: 10.1111/j.2517-6161.1986.tb01412.x.
- [297] E. Gedraite and M. Hadad, *Investigation on the effect of a Gaussian Blur in image filtering and segmentation*. 2011, p. 396.
- [298] T. Lindeberg, *Scale-Space Theory in Computer Vision*. 1993. doi: 10.1007/978-1-4757-6465-9.
- [299] M. Levoy, "Display of surfaces from volume data," *IEEE Computer Graphics and Applications*, vol. 8, no. 3, pp. 29–37, May 1988, doi: 10.1109/38.511.
- [300] Wenlian Cheng, Chunxiao Chen, Zhiyu Qian, and Lina Lu, "Research On Medical Image Three Dimensional Visualization System," in *2007 IEEE/ICME International Conference on Complex Medical Engineering*, May 2007, pp. 919–923. doi: 10.1109/ICCME.2007.4381874.
- [301] M. Ogata, T. Ohkami, H. C. Lauer, and H. Pfister, "A real-time volume rendering architecture using an adaptive resampling scheme for parallel and perspective projections," in *IEEE Symposium on Volume Visualization (Cat. No.989EX300)*, Oct. 1998, pp. 31–38. doi: 10.1109/SVV.1998.729582.
- [302] J. Sun, H. Li, P. Gao, and L. Wu, "Research on High Efficient Ray Casting Algorithm Based on VTK," in *2015 7th International Conference on Information Technology in Medicine and Education (ITME)*, Nov. 2015, pp. 212–214. doi: 10.1109/ITME.2015.118.
- [303] W. D. C. Advincula *et al.*, "Development and future trends in the application of visualization toolkit (VTK): The case for medical image 3D reconstruction," *AIP Conference Proceedings*, vol. 2092, no. 1, p. 020022, Apr. 2019, doi: 10.1063/1.5096690.
- [304] S. S. Khan and A. Ahmad, "Cluster center initialization algorithm for K-means clustering," *Pattern Recognition Letters*, vol. 25, no. 11, pp. 1293–1302, Aug. 2004, doi: 10.1016/j.patrec.2004.04.007.
- [305] P. A. Yushkevich, Y. Gao, and G. Gerig, "ITK-SNAP: An interactive tool for semi-automatic segmentation of multi-modality biomedical images," in *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Aug. 2016, pp. 3342–3345. doi: 10.1109/EMBC.2016.7591443.
- [306] B. Avants, N. Tustison, and G. Song, "Advanced normalization tools (ANTs)," *Insight J*, vol. 1–35, Nov. 2008, doi: 10.54294/uvnhin.
- [307] J. P. Ahrens, B. Geveci, and C. C. Law, "ParaView: An End-User Tool for Large-Data Visualization," in *The Visualization Handbook*, 2005.
- [308] "Costa LF. Further Generalizations of the Jaccard Index, Mach. Learn. 2021; <https://doi.org/10.48550/arXiv.2110.09619>".

- [309] F. C. Monteiro and A. C. Campilho, "Performance Evaluation of Image Segmentation," in *Image Analysis and Recognition*, A. Campilho and M. S. Kamel, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 248–259.
- [310] F. Lalys *et al.*, "Automatic aortic root segmentation and anatomical landmarks detection for TAVI procedure planning," *Minimally Invasive Therapy & Allied Technologies*, vol. 28, no. 3, pp. 157–164, May 2019, doi: 10.1080/13645706.2018.1488734.
- [311] M. de Bruijne, B. van Ginneken, M. A. Viergever, and W. J. Niessen, "Interactive segmentation of abdominal aortic aneurysms in CTA images," *Medical Image Analysis*, vol. 8, no. 2, pp. 127–138, Jun. 2004, doi: 10.1016/j.media.2004.01.001.
- [312] D. Mastrodicasa *et al.*, "Artificial Intelligence Applications in Aortic Dissection Imaging," *Seminars in Roentgenology*, vol. 57, no. 4, pp. 357–363, Oct. 2022, doi: 10.1053/j.ro.2022.07.001.
- [313] M. Bertolini, G. Luraghi, I. Belicchi, F. Migliavacca, and G. Colombo, "Evaluation of segmentation accuracy and its impact on patient-specific CFD analysis," *International Journal on Interactive Design and Manufacturing (IJIDeM)*, vol. 16, no. 2, pp. 545–556, Jun. 2022, doi: 10.1007/s12008-021-00802-z.
- [314] N. Tajbakhsh *et al.*, "Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?," *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1299–1312, May 2016, doi: 10.1109/TMI.2016.2535302.
- [315] L. F. Rodella *et al.*, "Abdominal aortic aneurysm and histological, clinical, radiological correlation," *Acta Histochemica*, vol. 118, no. 3, pp. 256–262, Apr. 2016, doi: 10.1016/j.acthis.2016.01.007.
- [316] Shaw, P.M.; Loree, J.; Gibbons, R.C. *Abdominal Aortic Aneurysm. In StatPearls; StatPearls Publishing: St. Petersburg, FL, USA, 2024.*
- [317] A. Sethi *et al.*, "Calcification of the abdominal aorta is an under-appreciated cardiovascular disease risk factor in the general population," *Frontiers in Cardiovascular Medicine*, vol. Volume 9-2022, 2022, doi: 10.3389/fcvm.2022.1003246.
- [318] M. Gameraddin, "Normal abdominal aorta diameter on abdominal sonography in healthy asymptomatic adults: impact of age and gender," *Journal of Radiation Research and Applied Sciences*, vol. 12, no. 1, pp. 186–191, Jan. 2019, doi: 10.1080/16878507.2019.1617553.
- [319] Tran, C.T.; Wu, C.Y.; Bordes, S.J.; Lui, F. *Anatomy, Abdomen and Pelvis: Abdominal Aorta. In StatPearls; StatPearls Publishing: St. Petersburg, FL, USA, 2024.*
- [320] A. Abdolmanafi, A. Forneris, R. D. Moore, and E. S. Di Martino, "Deep-learning method for fully automatic segmentation of the abdominal aortic aneurysm from computed tomography imaging," *Frontiers in Cardiovascular Medicine*, vol. Volume 9-2022, 2023, doi: 10.3389/fcvm.2022.1040053.
- [321] A. Vezakis, I. Vezakis, T. P. Vagenas, I. Kakkos, and G. K. Matsopoulos, "A Multidimensional Framework Incorporating 2D U-Net and 3D Attention U-Net for the Segmentation of Organs from 3D Fluorodeoxyglucose-Positron Emission Tomography Images," *Electronics*, vol. 13, no. 17, 2024, doi: 10.3390/electronics13173526.
- [322] T. Lyu *et al.*, "Dissected aorta segmentation using convolutional neural networks," *Computer Methods and Programs in Biomedicine*, vol. 211, p. 106417, Nov. 2021, doi: 10.1016/j.cmpb.2021.106417.
- [323] V. Chang, V. R. Bhavani, A. Q. Xu, and M. Hossain, "An artificial intelligence model for heart disease detection using machine learning algorithms," *Healthcare Analytics*, vol. 2, p. 100016, Nov. 2022, doi: 10.1016/j.health.2022.100016.
- [324] M. Izadikhah, "A fuzzy stochastic slacks-based data envelopment analysis model with application to healthcare efficiency," *Healthcare Analytics*, vol. 2, p. 100038, Nov. 2022, doi: 10.1016/j.health.2022.100038.
- [325] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural networks: an overview and application in radiology.," *Insights Imaging*, vol. 9, no. 4, pp. 611–629, Aug. 2018, doi: 10.1007/s13244-018-0639-9.

- [326] Maria-Pavlina Kalla, Theodoros P. Vagenas, Theodore L. Economopoulos, and George K. Matsopoulos, "Deep learning-based registration of two-dimensional dental images with edge specific loss," *Journal of Medical Imaging*, vol. 10, no. 3, p. 034002, Jun. 2023, doi: 10.1117/1.JMI.10.3.034002.
- [327] F. Shamshad *et al.*, "Transformers in medical imaging: A survey," *Medical Image Analysis*, vol. 88, p. 102802, Aug. 2023, doi: 10.1016/j.media.2023.102802.
- [328] "Hatamizadeh, A.; Nath, V.; Tang, Y.; Yang, D.; Roth, H.R.; Xu, D. Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images. In *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*; Crimi, A., Bakas, S., Eds.; Springer International Publishing: Cham, Switzerland, 2022; pp. 272–284."
- [329] T. Yang, G. Zhu, L. Cai, J. H. Yeo, Y. Mao, and J. Yang, "A benchmark study of convolutional neural networks in fully automatic segmentation of aortic root," *Frontiers in Bioengineering and Biotechnology*, vol. Volume 11-2023, 2023, doi: 10.3389/fbioe.2023.1171868.
- [330] T. Kim, S. On, J. G. Gwon, and N. Kim, "Computed tomography-based automated measurement of abdominal aortic aneurysm using semantic segmentation with active learning," *Scientific Reports*, vol. 14, no. 1, p. 8924, Apr. 2024, doi: 10.1038/s41598-024-59735-8.
- [331] J. R. Camara *et al.*, "Development of a convolutional neural network to detect abdominal aortic aneurysms," *Journal of Vascular Surgery Cases, Innovations and Techniques*, vol. 8, no. 2, pp. 305–311, Jun. 2022, doi: 10.1016/j.jvscit.2022.04.003.
- [332] L. Cao *et al.*, "Fully automatic segmentation of type B aortic dissection from CTA images enabled by deep learning," *European Journal of Radiology*, vol. 121, Dec. 2019, doi: 10.1016/j.ejrad.2019.108713.
- [333] C. Zhang, S. Bengio, M. Hardt, B. Recht, and O. Vinyals, "Understanding deep learning (still) requires rethinking generalization," *Commun. ACM*, vol. 64, no. 3, pp. 107–115, Feb. 2021, doi: 10.1145/3446776.
- [334] L. Radl *et al.*, "AVT: Multicenter aortic vessel tree CTA dataset collection with ground truth segmentation masks," *Data in Brief*, vol. 40, p. 107801, Feb. 2022, doi: 10.1016/j.dib.2022.107801.
- [335] "Radl, Lukas; Jin, Yuan; Pepe, Antonio; Li, Jianning; Gsaxner, Christina; Zhao, Fenhua; et al. (2022). Aortic Vessel Tree (AVT) CTA Datasets and Segmentations. figshare. Dataset. <https://doi.org/10.6084/m9.figshare.14806362.v1>".
- [336] F. Peyrin and K. Engelke, "CT Imaging: Basics and New Trends," 2021, pp. 1173–1215. doi: 10.1007/978-3-319-93785-4_36.
- [337] C. Shorten and T. M. Khoshgoftaar, "A survey on Image Data Augmentation for Deep Learning," *Journal of Big Data*, vol. 6, no. 1, p. 60, Jul. 2019, doi: 10.1186/s40537-019-0197-0.
- [338] S. Jadon, "A survey of loss functions for semantic segmentation," in *2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, Oct. 2020, pp. 1–7. doi: 10.1109/CIBCB48159.2020.9277638.
- [339] 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, Aug. 2015, doi: 10.1186/s12880-015-0068-x.
- [340] Varduhi Yeghiazaryan and Irina D. Voiculescu, "Family of boundary overlap metrics for the evaluation of medical image segmentation," *Journal of Medical Imaging*, vol. 5, no. 1, p. 015006, Feb. 2018, doi: 10.1117/1.JMI.5.1.015006.
- [341] K. Sanjar, O. Bekhzod, J. Kim, J. Kim, A. Paul, and J. Kim, "Improved U-Net: Fully Convolutional Network Model for Skin-Lesion Segmentation," *Applied Sciences*, vol. 10, no. 10, 2020, doi: 10.3390/app10103658.
- [342] M. Gillot *et al.*, "Automatic multi-anatomical skull structure segmentation of cone-beam computed tomography scans using 3D UNETR," *PLOS ONE*, vol. 17, no. 10, p. e0275033, Oct. 2022, doi: 10.1371/journal.pone.0275033.

- [343] J. Lee *et al.*, “Deep learning segmentation of fibrous cap in intravascular optical coherence tomography images,” *Scientific Reports*, vol. 14, no. 1, p. 4393, Feb. 2024, doi: 10.1038/s41598-024-55120-7.
- [344] T. P. Vagenas, K. Georgas, and G. Matsopoulos, *Deep Learning-Based Segmentation and Mesh Reconstruction of the Aortic Vessel Tree from CTA Images*. 2024, p. 94. doi: 10.1007/978-3-031-53241-2_7.
- [345] W. E. Lorensen and H. E. Cline, “Marching cubes: A high resolution 3D surface construction algorithm,” *SIGGRAPH Comput. Graph.*, vol. 21, no. 4, pp. 163–169, Aug. 1987, doi: 10.1145/37402.37422.
- [346] J. Vollmer, R. Mencl, and H. Müller, “Improved Laplacian Smoothing of Noisy Surface Meshes,” *Computer Graphics Forum*, vol. 18, no. 3, pp. 131–138, Sep. 1999, doi: 10.1111/1467-8659.00334.
- [347] M. Attene, “A lightweight approach to repairing digitized polygon meshes,” *The Visual Computer*, vol. 26, no. 11, pp. 1393–1406, Nov. 2010, doi: 10.1007/s00371-010-0416-3.
- [348] “Yang, Y.; Jiang, P.; Cai, X.; Xue, Z.; Shen, D. Integrating Convolutional Neural Network and Transformer for Lumen Prediction Along the Aorta Sections. In Machine Learning in Medical Imaging; Xu, X., Cui, Z., Rekik, I., Ouyang, X., Sun, K., Eds.; Springer Nature: Cham, Switzerland, 2024; pp. 340–349.”
- [349] Q. Yang, C. Wang, K. Pan, B. Xia, R. Xie, and J. Shi, “An improved 3D-UNet-based brain hippocampus segmentation model based on MR images,” *BMC Medical Imaging*, vol. 24, no. 1, p. 166, Jul. 2024, doi: 10.1186/s12880-024-01346-w.
- [350] R. Azad *et al.*, “Medical Image Segmentation Review: The Success of U-Net,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 46, no. 12, pp. 10076–10095, Dec. 2024, doi: 10.1109/TPAMI.2024.3435571.
- [351] A. W. Salehi *et al.*, “A Study of CNN and Transfer Learning in Medical Imaging: Advantages, Challenges, Future Scope,” *Sustainability*, vol. 15, no. 7, 2023, doi: 10.3390/su15075930.
- [352] H. -C. Shin *et al.*, “Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning,” *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1285–1298, May 2016, doi: 10.1109/TMI.2016.2528162.
- [353] L. Ali, F. Alnajjar, H. A. Jassmi, M. Gocho, W. Khan, and M. A. Serhani, “Performance Evaluation of Deep CNN-Based Crack Detection and Localization Techniques for Concrete Structures,” *Sensors*, vol. 21, no. 5, 2021, doi: 10.3390/s21051688.
- [354] G. Chen, L. Li, J. Zhang, and Y. Dai, “Rethinking the unpretentious U-net for medical ultrasound image segmentation,” *Pattern Recognition*, vol. 142, p. 109728, Oct. 2023, doi: 10.1016/j.patcog.2023.109728.
- [355] L. Pinto-Coelho, “How Artificial Intelligence Is Shaping Medical Imaging Technology: A Survey of Innovations and Applications,” *Bioengineering*, vol. 10, no. 12, 2023, doi: 10.3390/bioengineering10121435.
- [356] M. Elhaddad and S. Hamam, “AI-Driven Clinical Decision Support Systems: An Ongoing Pursuit of Potential,” *Cureus*, vol. 16, no. 4, p. e57728, Apr. 2024, doi: 10.7759/cureus.57728.
- [357] H. Rahman, A. R. Khan, T. Sadiq, A. H. Farooqi, I. U. Khan, and W. H. Lim, “A Systematic Literature Review of 3D Deep Learning Techniques in Computed Tomography Reconstruction,” *Tomography*, vol. 9, no. 6, pp. 2158–2189, 2023, doi: 10.3390/tomography9060169.
- [358] K. O. Kappe, S. P. M. Smorenburg, A. W. J. Hoksbergen, J. M. Wolterink, and K. K. Yeung, “Deep Learning–Based Intraoperative Stent Graft Segmentation on Completion Digital Subtraction Angiography During Endovascular Aneurysm Repair,” *J Endovasc Ther*, vol. 30, no. 6, pp. 822–827, Dec. 2023, doi: 10.1177/15266028221105840.
- [359] E. E. Ntiri *et al.*, “Improved Segmentation of the Intracranial and Ventricular Volumes in Populations with Cerebrovascular Lesions and Atrophy Using 3D CNNs,” *Neuroinformatics*, vol. 19, no. 4, pp. 597–618, Oct. 2021, doi: 10.1007/s12021-021-09510-1.

- [360] D. Lesage, E. D. Angelini, I. Bloch, and G. Funka-Lea, "A review of 3D vessel lumen segmentation techniques: Models, features and extraction schemes," *Medical Image Analysis*, vol. 13, no. 6, pp. 819–845, Dec. 2009, doi: 10.1016/j.media.2009.07.011.
- [361] B. Preim and C. Botha, "Image Analysis for Medical Visualization," 2014, pp. 111–175. doi: 10.1016/B978-0-12-415873-3.00004-3.
- [362] A. Abdolmanafi, A. Forneris, R. D. Moore, and E. S. Di Martino, "Deep-learning method for fully automatic segmentation of the abdominal aortic aneurysm from computed tomography imaging," *Frontiers in Cardiovascular Medicine*, vol. 9, p. 1040053, 2022, doi: 10.3389/fcvm.2022.1040053.
- [363] I. Kakkos, T. P. Vagenas, A. Zygogianni, and G. K. Matsopoulos, "Towards Automation in Radiotherapy Planning: A Deep Learning Approach for the Delineation of Parotid Glands in Head and Neck Cancer," *Bioengineering*, vol. 11, no. 3, p. 214, Mar. 2024, doi: 10.3390/bioengineering11030214.
- [364] S. Niyas, S. J. Pawan, M. A. Kumar, and J. Rajan, "Medical image segmentation with 3D convolutional neural networks: A survey," *Neurocomputing*, vol. 493, pp. 397–413, 2022, doi: <https://doi.org/10.1016/j.neucom.2022.04.065>.
- [365] A. Hatamizadeh *et al.*, "UNETR: Transformers for 3D Medical Image Segmentation," in *2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, Jan. 2022, pp. 1748–1758. doi: 10.1109/WACV51458.2022.00181.
- [366] C. Mavridis *et al.*, "Automatic Segmentation in 3D CT Images: A Comparative Study of Deep Learning Architectures for the Automatic Segmentation of the Abdominal Aorta," *Electronics*, vol. 13, no. 24, p. 4919, 2024.
- [367] V. Rani, S. T. Nabi, M. Kumar, A. Mittal, and K. Kumar, "Self-supervised Learning: A Succinct Review," *Archives of Computational Methods in Engineering*, vol. 30, no. 4, pp. 2761–2775, May 2023, doi: 10.1007/s11831-023-09884-2.
- [368] C. Vondrick, A. Shrivastava, A. Fathi, S. Guadarrama, and K. Murphy, "Tracking Emerges by Colorizing Videos," in *Computer Vision – ECCV 2018*, vol. 11217, V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss, Eds., Cham: Springer International Publishing, 2018, pp. 402–419. doi: 10.1007/978-3-030-01261-8_24.
- [369] Z. Xie *et al.*, "SimMIM: A Simple Framework for Masked Image Modeling," 2022, pp. 9653–9663.
- [370] P. Chen, S. Liu, and J. Jia, "Jigsaw Clustering for Unsupervised Visual Representation Learning," 2021, pp. 11526–11535.
- [371] A. Taleb *et al.*, "3D Self-Supervised Methods for Medical Imaging," in *Advances in Neural Information Processing Systems*, Curran Associates, Inc., 2020, pp. 18158–18172.
- [372] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A Simple Framework for Contrastive Learning of Visual Representations," *37th International Conference on Machine Learning, ICML 2020*, vol. PartF168147-3, pp. 1575–1585, Feb. 2020.
- [373] J.-B. Grill *et al.*, "Bootstrap your own latent-a new approach to self-supervised learning," *Advances in neural information processing systems*, vol. 33, pp. 21271–21284, 2020.
- [374] M. Caron *et al.*, "Emerging properties in self-supervised vision transformers," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 9650–9660.
- [375] K. He, H. Fan, Y. Wu, S. Xie, and R. Girshick, "Momentum Contrast for Unsupervised Visual Representation Learning," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 9726–9735, Nov. 2019, doi: 10.1109/CVPR42600.2020.00975.
- [376] S.-C. Huang, A. Pareek, M. Jensen, M. P. Lungren, S. Yeung, and A. S. Chaudhari, "Self-supervised learning for medical image classification: a systematic review and implementation guidelines," *npj Digital Medicine*, vol. 6, no. 1, pp. 1–16, Apr. 2023, doi: 10.1038/s41746-023-00811-0.
- [377] M. R. H. Taher, M. Ikuta, and R. Soni, "Curriculum Self-Supervised Learning for 3D CT Cardiac Image Segmentation," in *Proceedings of the 3rd Machine Learning for Health Symposium*, PMLR, Dec. 2023, pp. 145–156.

- [378] E. Özbay, F. A. Özbay, and F. S. Gharehchopogh, “Kidney Tumor Classification on CT images using Self-supervised Learning,” *Computers in Biology and Medicine*, vol. 176, p. 108554, Jun. 2024, doi: 10.1016/j.compbimed.2024.108554.
- [379] C. Ouyang, C. Biffi, C. Chen, T. Kart, H. Qiu, and D. Rueckert, “Self-Supervised Learning for Few-Shot Medical Image Segmentation,” *IEEE Transactions on Medical Imaging*, vol. 41, no. 7, pp. 1837–1848, Jul. 2022, doi: 10.1109/TMI.2022.3150682.
- [380] Y. Wang *et al.*, “SwinMM: Masked Multi-view with Swin Transformers for 3D Medical Image Segmentation,” in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2023: 26th International Conference, Vancouver, BC, Canada, October 8–12, 2023, Proceedings, Part III*, Berlin, Heidelberg: Springer-Verlag, Oct. 2023, pp. 486–496. doi: 10.1007/978-3-031-43898-1_47.
- [381] Z. Zhou and X. Liu, “Masked autoencoders in computer vision: A comprehensive survey,” *IEEE Access*, 2023.
- [382] R. Bibb, D. Eggbeer, and A. Paterson, “2 - Medical imaging,” in *Medical Modelling (Second Edition)*, R. Bibb, D. Eggbeer, and A. Paterson, Eds., Oxford: Woodhead Publishing, 2015, pp. 7–34. doi: 10.1016/B978-1-78242-300-3.00002-0.
- [383] Y. Tang *et al.*, “Self-Supervised Pre-Training of Swin Transformers for 3D Medical Image Analysis,” in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2022, pp. 20698–20708. doi: 10.1109/CVPR52688.2022.02007.
- [384] Z. Liu *et al.*, “Swin Transformer: Hierarchical Vision Transformer using Shifted Windows,” in *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, Oct. 2021, pp. 9992–10002. doi: 10.1109/ICCV48922.2021.00986.
- [385] B. Landman, Z. Xu, J. Igelsias, M. Styner, T. Langerak, and A. Klein, “Miccai multi-atlas labeling beyond the cranial vault—workshop and challenge,” in *Proc. MICCAI Multi-Atlas Labeling Beyond Cranial Vault—Workshop Challenge*, 2015, p. 12.
- [386] M. J. Cardoso *et al.*, “MONAI: An open-source framework for deep learning in healthcare.” arXiv, Nov. 2022. doi: 10.48550/arXiv.2211.02701.
- [387] T. Falk *et al.*, “U-Net: deep learning for cell counting, detection, and morphometry,” *Nature Methods*, vol. 16, no. 1, pp. 67–70, Jan. 2019, doi: 10.1038/s41592-018-0261-2.
- [388] A. Myronenko, “3D MRI Brain Tumor Segmentation Using Autoencoder Regularization,” in *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*, A. Crimi, S. Bakas, H. Kuijf, F. Keyvan, M. Reyes, and T. van Walsum, Eds., Cham: Springer International Publishing, 2019, pp. 311–320.