



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ
ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

Κατάτμηση τρισδιάστατων ιατρικών εικόνων με σύγχρονες μεθόδους μηχανικής και βαθιάς μάθησης

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

ΤΖΑΝΑΚΑΚΗ ΑΛΕΞΑΝΔΡΟΥ

Επιβλέπων : Γεώργιος Ματσόπουλος
Καθηγητής Ε.Μ.Π.
Συνεπιβλέπουσα : Ουρανία Πετροπούλου
ΕΔΙΠ Ε.Μ.Π.

Αθήνα, Ιούνιος 2025

Η σελίδα αυτή είναι σκόπιμα λευκή.



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ
ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

Κατάτμηση τρισδιάστατων ιατρικών εικόνων με σύγχρονες μεθόδους μηχανικής και βαθιάς μάθησης

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ
του
ΤΖΑΝΑΚΑΚΗ ΑΛΕΞΑΝΔΡΟΥ

Επιβλέπων : Γεώργιος Ματσόπουλος
Καθηγητής Ε.Μ.Π.
Συνεπιβλέπουσα : Ουρανία Πετροπούλου
ΕΔΙΠ Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 27^η Ιουνίου 2025.

(Υπογραφή)

(Υπογραφή)

(Υπογραφή)

.....
Γ. Ματσόπουλος
Καθηγητής Ε.Μ.Π.

.....
Α. Παναγόπουλος
Καθηγητής Ε.Μ.Π.

.....
Π. Τσανάκας
Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούνιος 2025

(Υπογραφή)

.....

ΤΖΑΝΑΚΑΚΗΣ ΑΛΕΞΑΝΔΡΟΣ

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

Copyright © Τζανακάκης Αλέξανδρος, 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 τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται στον συγγραφέα.

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

Περίληψη

Η τμηματοποίηση ιατρικών εικόνων συνιστά μία από τις πιο κρίσιμες διαδικασίες στην ανάλυση δεδομένων μαγνητικής τομογραφίας, καθώς επιτρέπει την ακριβή εντόπιση παθολογικών περιοχών και διευκολύνει τη διάγνωση, την πρόγνωση και την παρακολούθηση της εξέλιξης ασθενειών. Στην περίπτωση των μηνιγγιωμάτων, που αποτελούν τον πιο συχνό καλοήγη ενδοκρανιακό όγκο, η έγκαιρη και ακριβής τμηματοποίηση του όγκου μπορεί να συμβάλλει ουσιαστικά στη βελτιστοποίηση της θεραπευτικής αντιμετώπισης και στην αποφυγή επιπλοκών.

Η παρούσα εργασία εστιάζει στην αξιολόγηση διαφορετικών αρχιτεκτονικών βαθιάς μάθησης για το πρόβλημα της τρισδιάστατης τμηματοποίησης μηνιγγιωμάτων σε μαγνητικές τομογραφίες. Αξιοποιήθηκε το σύνολο δεδομένων Meningioma-SEG-CLASS και αναπτύχθηκε ένα κοινό πειραματικό πλαίσιο με ενιαίες παραμέτρους εκπαίδευσης, τεχνικές προεπεξεργασίας και διαδικασίες αξιολόγησης. Δοκιμάστηκαν παραδοσιακές συνελκτικές αρχιτεκτονικές, πιο σύγχρονα μοντέλα με μηχανισμούς προσοχής, καθώς και πρόσφατες υβριδικές προσεγγίσεις που βασίζονται σε νέα μοντέλα ακολουθιακής επεξεργασίας.

Ιδιαίτερη έμφαση δόθηκε στην ορθή διαχείριση των δεδομένων, στην επιλογή τεχνικών επαύξησης, στον καθορισμό κατάλληλων συναρτήσεων κόστους και βελτιστοποιητών, καθώς και στη χρήση μετα-επεξεργασίας κατά το στάδιο της εξαγωγής προβλέψεων (inference). Τα αποτελέσματα δείχνουν ότι, ακόμη και υπό συνθήκες περιορισμένων δεδομένων και υπολογιστικών πόρων, είναι δυνατή η ανάπτυξη μοντέλων με αξιόλογη απόδοση, γεγονός που αναδεικνύει τη σημασία ενός καλά σχεδιασμένου πειραματικού pipeline.

Λέξεις-κλειδιά: Τμηματοποίηση, Μαγνητική Τομογραφία, Μηνιγγίωμα, Τεχνητή Νοημοσύνη, Βαθιά Μάθηση.

Abstract

Medical image segmentation plays a vital role in processing MRI data, enabling the precise identification of pathological regions and assisting in diagnosis, prognosis, and treatment planning. In the case of meningiomas — the most common type of intracranial tumor — accurate segmentation can significantly impact the quality of patient care, supporting better surgical planning and reducing risks.

This thesis investigates the use of deep learning architectures for the task of 3D meningioma segmentation on magnetic resonance images. Using the Meningioma-SEG-CLASS dataset, a unified experimental framework was developed to ensure consistent comparison across models. The selected models include classic convolutional neural networks, modern architectures incorporating attention mechanisms, and state-of-the-art hybrid approaches based on sequential modeling paradigms.

Beyond model selection, emphasis was placed on the complete pipeline: from data preparation and augmentation to loss function design, optimizer selection, and post-processing at inference time. This comprehensive approach allowed for meaningful evaluation under realistic constraints, such as limited annotated data and moderate computational resources. The findings demonstrate that, through careful experimental design, competitive segmentation performance can be achieved even in resource-constrained environments.

Keywords: Segmentation, Magnetic Resonance Imaging, Meningioma, Artificial Intelligence, Deep Learning.

Η σελίδα αυτή είναι σκόπιμα λευκή.

Ευχαριστίες

Θα ήθελα να εκφράσω τις θερμές μου ευχαριστίες στον καθηγητή κ. Γεώργιο Ματσόπουλο για την πολύτιμη επίβλεψη αυτής της διπλωματικής εργασίας και για την εμπιστοσύνη που μου έδειξε, δίνοντάς μου τη δυνατότητα να εμβαθύνω σε ένα θεματικό πεδίο με αυξανόμενο ερευνητικό και πρακτικό ενδιαφέρον.

Ιδιαίτερες ευχαριστίες οφείλω στον Υπ. Διδάκτορα Κωνσταντίνο Γεώργα για τη συνεχή καθοδήγηση, την υποστήριξη και τις ουσιαστικές παρατηρήσεις του σε όλα τα στάδια της εργασίας. Η συνεργασία μας υπήρξε ιδιαίτερα εποικοδομητική και συνέβαλε ουσιαστικά στην ολοκλήρωση του παρόντος έργου.

Τέλος, θα ήθελα να ευχαριστήσω την οικογένειά μου — τους γονείς μου και τον αδερφό μου — καθώς και τους φίλους μου, για τη σταθερή υποστήριξή τους καθ' όλη τη διάρκεια των σπουδών μου.

Αθήνα, Ιούνιος 2025
Τζανακάκης Αλέξανδρος

Πίνακας Περιεχομένων

Ευρετήριο Εικόνων	13
Ευρετήριο Πινάκων	15
Εισαγωγή	16
Κεφάλαιο 1: Καρκίνος.....	17
1.1 Γενικά.....	17
1.2 Καρκίνος του Εγκεφάλου	17
1.3 Μηνιγγίωμα.....	18
1.4 Αιτίες του Καρκίνου	19
1.5 Συμπτώματα του Καρκίνου	19
1.5.1 Καρκίνος	19
1.5.2 Καρκίνος Εγκεφάλου και Μηνιγγίωμα.....	20
1.6 Διάγνωση	20
1.7 Θεραπεία	21
Κεφάλαιο 2: Μαγνητική Τομογραφία	23
2.1 Γενικά.....	23
2.2 Πυρηνικός Μαγνητικός Συντονισμός.....	23
2.3 RF Συντονισμός	25
2.4 Χωρικός Εντοπισμός του Σήματος	27
2.5 Μαγνητικός Τομογράφος.....	28
Κεφάλαιο 3: Μηχανική Μάθηση και Βαθιά Μάθηση	29
3.1 Μηχανική Μάθηση	29
3.1.1 Εισαγωγή	29
3.1.2 Κατηγορίες Μηχανικής Μάθησης	30
3.1.2.1 Επιβλεπόμενη Μάθηση.....	30
3.1.2.2 Μη Επιβλεπόμενη Μάθηση	31
3.1.2.3 Ενισχυτική Μάθηση.....	33
3.1.3 Μηχανική Μάθηση στην Υγεία	34
3.2 Βαθιά Μάθηση	35
3.2.1 Εισαγωγή	35
3.2.2 Συνελικτικά Νευρωνικά Δίκτυα	36
3.2.2.1 Επίπεδα Συνέλιξης	36
3.2.2.2 Επίπεδα Υποδειγματοληψίας.....	37
3.2.2.3 Πλήρως Συνδεδεμένα Επίπεδα	38
3.2.2.4 Συναρτήσεις Ενεργοποίησης	38
3.2.2.5 Συναρτήσεις Σφάλματος	39
3.2.2.6 Αλγόριθμοι Βελτιστοποίησης.....	39

3.2.3 Εκπαίδευση Συνελκτικών Νευρωνικών Δικτύων	40
3.2.3.1 Προετοιμασία των Δεδομένων	40
3.2.3.2 Σχεδιασμός της Αρχιτεκτονικής	41
3.2.3.3 Αξιολόγηση και Βελτιστοποίηση	42
3.2.4 Μηχανισμοί Προσοχής και Αρχιτεκτονικές Τύπου Transformer	43
3.2.5 Αρχιτεκτονικές Δικτύων για επεξεργασία εικόνων	44
3.2.5.1 U-Net.....	45
3.2.5.2 SegResNet.....	46
3.2.5.3 SwinUNETR.....	47
3.2.5.4 SegMamba	48
Κεφάλαιο 4: Υλοποίηση και Πειραματική Αξιολόγηση	50
4.1 Εισαγωγή	50
4.2 Περιγραφή Συνόλου Δεδομένων	50
4.3 Επαύξηση Δεδομένων και Μετασχηματισμοί	52
4.4 Εκπαίδευση και Βασικές Υπερπαράμετροι	53
4.4.1 Συναρτήσεις Απώλειας	53
4.4.2 Βελτιστοποιητές και Προγραμματιστές Μάθησης (Schedulers)	54
4.4.3 Υπερπαράμετροι Εκπαίδευσης	54
4.5 Αποτελέσματα.....	55
4.5.1 Βελτιστοποίηση Υπερπαραμέτρων.....	55
4.5.1.1 Επιλογή Συνάρτησης Απώλειας.....	55
4.5.1.2 Βελτιστοποίηση Learning Rate Scheduler.....	56
4.5.1.3 Τελική Επιλογή Υπερπαραμέτρων	56
4.5.2 Συγκριτική Αξιολόγηση Αρχιτεκτονικών	56
4.5.2.1 Αποτελέσματα Αρχικής Αξιολόγησης.....	57
4.5.2.2 Ανάλυση Απόδοσης SegMamba.....	58
4.5.2.3 Περιορισμοί και Προκλήσεις.....	58
4.5.3 Ποιοτική Ανάλυση Αποτελεσμάτων.....	59
4.5.3.1 Στατιστική Ανάλυση Κατανομής Dice Scores	59
4.5.3.2 Ανάλυση Περιπτώσεων με Μηδενικό Dice Score	60
4.5.3.3 Το Πρόβλημα του Dice Score σε Μικρούς Όγκους	61
4.5.4 Πρόταση Αναθεωρημένης Αξιολόγησης.....	61
4.5.4.1 Αιτιολόγηση για Τροποποίηση της Μετρικής	61
4.5.4.2 Τροποποιημένη Μετρική Αξιολόγησης.....	61
4.5.4.3 Επαναξιολόγηση με Βελτιωμένη Μετρική	62
4.5.5 Συνολική Αξιολόγηση και Συμπεράσματα	64
4.5.5.1 Αποτελεσματικότητα των Μοντέλων	64

4.5.5.2 Σημασία της Στατιστικής Ανάλυσης	64
4.5.5.3 Συγκριτική Απόδοση και Ερευνητικές Κατευθύνσεις.....	64
4.5.5.4 Ανάλυση Υπεροχής του U-Net.....	64
Συμπεράσματα-Μελλοντικές Προεκτάσεις	66
Βιβλιογραφία	67

Ευρετήριο Εικόνων

Εικόνα 1: Κύριες ανατομικές περιοχές εντόπισης μηνιγγιωμάτων στον ανθρώπινο εγκέφαλο [21].	18
Εικόνα 2: Το πρωτόνιο του υδρογόνου συμπεριφέρεται όπως ένας μικρός ραβδόμορφος μαγνήτης. Υπό κανονικές συνθήκες, αυτοί οι "μαγνήτες" πρωτονίων υδρογόνου περιστρέφονται μέσα στο σώμα με τους άξονές τους ευθυγραμμισμένους τυχαία [43].	25
Εικόνα 3: Όταν το σώμα τοποθετείται μέσα σε ένα ισχυρό μαγνητικό πεδίο, όπως αυτό ενός μαγνητικού τομογράφου, οι άξονες των πρωτονίων ευθυγραμμίζονται. Αυτή η ομοιόμορφη ευθυγράμμιση δημιουργεί ένα μαγνητικό διάνυσμα προσανατολισμένο κατά μήκος του άξονα του μαγνητικού τομογράφου [43].	26
Εικόνα 4: Μαγνητικός τομογράφος τύπου MRI, ενδεικτικός του εξοπλισμού που χρησιμοποιείται στην τμηματοποίηση εγκεφαλικών εικόνων [47].	28
Εικόνα 5: Γραφική αναπαράσταση γραμμικού ταξινομητή SVM. Το βέλτιστο υπερπίπεδο (optimal hyperplane) διαχωρίζει τις δύο κλάσεις (κόκκινοι κύκλοι και μπλε τετράγωνα) μεγιστοποιώντας το περιθώριο μεταξύ των πιο κοντινών δειγμάτων, γνωστών ως support vectors. [50].	31
Εικόνα 6: Παραδείγματα διαφορετικών αλγορίθμων μη επιβλεπόμενης ομαδοποίησης (clustering): K-means clustering, Gaussian mixture models, ιεραρχική ομαδοποίηση (hierarchical clustering) και ομαδοποίηση βασισμένη σε γράφους (graph-based clustering). Κάθε μέθοδος χρησιμοποιεί διαφορετικά κριτήρια για τον διαχωρισμό των δεδομένων σε ομάδες.	32
Εικόνα 7: Αρχιτεκτονική ενισχυτικής μάθησης (Reinforcement Learning). Ο πράκτορας (agent) αλληλεπιδρά με το περιβάλλον (environment) εκτελώντας ενέργειες (actions) και λαμβάνοντας ως ανατροφοδότηση καταστάσεις (states) και ανταμοιβές ή τιμωρίες (rewards/punishments), με στόχο τη βελτιστοποίηση της στρατηγικής του [56].	34
Εικόνα 8: Σχηματική σύγκριση μεταξύ ενός απλού νευρωνικού δικτύου και ενός βαθιού νευρωνικού δικτύου (deep neural network) [60].	36
Εικόνα 9: Αρχιτεκτονική ενός συνελκτικού νευρωνικού δικτύου (CNN). Το δίκτυο περιλαμβάνει στάδια εξαγωγής χαρακτηριστικών (convolution και pooling) και στη συνέχεια πλήρως συνδεδεμένα στρώματα για την τελική ταξινόμηση [61].	36
Εικόνα 10: Παράδειγμα λειτουργίας συνελκτικού νευρωνικού δικτύου: η αρχική εικόνα (Input Vector) υφίσταται υποδειγματοληψία (Pooling), ακολουθεί η εφαρμογή πυρήνα (Kernel) και τελικά υπολογίζεται η έξοδος (Destination Pixel) με συνέλιξη [63].	37
Εικόνα 11: Παράδειγμα εφαρμογής υποδειγματοληψίας μέγιστης τιμής (max pooling) με παράθυρο 2×2 και βήμα 2. Από κάθε τοπική περιοχή του αρχικού πίνακα εισόδου (4×4) επιλέγεται η μέγιστη τιμή, παράγοντας έναν συνοπτικότερο πίνακα εξόδου (2×2) [65].	38
Εικόνα 12: Τυπική ροή ανάπτυξης και αξιολόγησης ενός συστήματος μηχανικής μάθησης. Η διαδικασία περιλαμβάνει διαδοχικά στάδια επιλογής και προεπεξεργασίας του συνόλου δεδομένων, σχεδιασμού αρχιτεκτονικής (π.χ. CNN), εκπαίδευσης, αξιολόγησης και τελικής βελτιστοποίησης. Τα βέλη υποδεικνύουν τον επαναληπτικό χαρακτήρα της διαδικασίας, έως ότου επιτευχθεί η επιθυμητή απόδοση [68].	43
Εικόνα 13: Αρχιτεκτονική του μοντέλου Transformer. Στο (Α) παρουσιάζονται οι δομικές μονάδες του Encoder (αριστερά) και του Decoder (δεξιά), που αποτελούνται από στρώσεις πολλαπλών κεφαλών προσοχής (Multi-Head Attention), κανονικοποίησης (Add & Norm) και τροφοδότησης προώθησης (Feed Forward). Στο (Β) απεικονίζεται ο μηχανισμός προσοχής (Attention Mechanism), όπου κάθε είσοδος μετατρέπεται σε τριάδες (query, key, value) και υπολογίζεται ένα σταθμισμένο άθροισμα βάσει της ομοιότητας μεταξύ query και key [71].	44
Εικόνα 14: Σχηματική απεικόνιση της αρχιτεκτονικής U-Net. Διακρίνονται τα στάδια κωδικοποίησης και αποκωδικοποίησης, καθώς και τα skip connections που συνδέουν	

συμμετρικά επίπεδα του encoder και του decoder. Απεικονίζονται επίσης τα βασικά βήματα επεξεργασίας: συνέλιξη, υποδειγματοληψία, υπερδειγματοληψία και τελική παραγωγή του χάρτη τμηματοποίησης [72].	46
Εικόνα 15: Διάγραμμα αρχιτεκτονικής του SegResNet για ιατρική τμηματοποίηση. Το δίκτυο ακολουθεί συμμετρική διάταξη τύπου encoder–decoder, ενσωματώνοντας residual blocks και skip connections για ενίσχυση της πληροφορίας και σταθερότητα κατά την εκπαίδευση. Η είσοδος υποβάλλεται σε επεξεργασία μέσω διαδοχικών συνελκτικών επιπέδων με group normalization και ενεργοποίηση ReLU, ενώ η έξοδος παράγεται μέσω σταδιακής ανακατασκευής με upsampling και συνδυασμό χαρακτηριστικών [75].	47
Εικόνα 16: Δομική απεικόνιση της αρχιτεκτονικής SwinUNETR. Το μοντέλο ακολουθεί διάταξη τύπου encoder-decoder, με Swin Transformer blocks να εφαρμόζονται στο encoding path για την εξαγωγή χαρακτηριστικών, ενώ τα skip connections και το decoding path οδηγούν στην τελική έξοδο τμηματοποίησης [76].	48
Εικόνα 17: Σχηματική αναπαράσταση της αρχιτεκτονικής SegMamba. Ο encoder αποτελείται από TSMamba blocks και μονάδες υποδειγματοληψίας, ενώ ο decoder εφαρμόζει transposed convolutions και skip connections. Περιλαμβάνονται επίσης residual blocks και μηχανισμοί εκτίμησης αβεβαιότητας για ενίσχυση της εκφραστικότητας και της σταθερότητας του μοντέλου [77].	49
Εικόνα 18: Ενδεικτικές εικόνες MRI από το σύνολο δεδομένων Meningioma-SEG-CLASS. Παρουσιάζονται περιπτώσεις Grade I και Grade II μηνιγγιωμάτων, με τις περιοχές του όγκου επισημασμένες με κόκκινο περίγραμμα, τόσο σε ακολουθίες T1-CE όσο και σε T2-FLAIR (μόνο για λόγους οπτικής απεικόνισης· η παρούσα μελέτη χρησιμοποιεί αποκλειστικά τις T1-CE) [4].	51
Εικόνα 19: Οπτική απεικόνιση επιλεγμένων στοχαστικών μετασχηματισμών που εφαρμόστηκαν σε εικόνες MRI κατά την εκπαίδευση. Από αριστερά προς τα δεξιά και πάνω προς τα κάτω: αρχική εικόνα, RandFlip, RandAffine, RandScaleIntensity και RandShiftIntensity. Οι μετασχηματισμοί υλοποιήθηκαν μέσω της βιβλιοθήκης MONAI.	53
Εικόνα 20: Καμπύλες εκπαίδευσης για το μοντέλο U-Net. Απεικονίζονται η τιμή του Dice score στο validation set και η τιμή της συνάρτησης κόστους (loss) στο training set ανά εποχή. Παρατηρείται σταθερή βελτίωση της απόδοσης μέχρι τη σύγκλιση.	57
Εικόνα 21: Εξέλιξη του Dice Score στο validation set κατά την εκπαίδευση του μοντέλου SegMamba. Παρατηρείται καλή απόδοση από τις πρώτες εποχές (0.65), με το σκορ να συγκλίνει στο 0.80 στα τελευταία στάδια της εκπαίδευσης.	58
Εικόνα 22: Παράδειγμα οπτικής αξιολόγησης αποτελεσμάτων τμηματοποίησης. Εμφανίζεται η αρχική εικόνα MRI, η ground truth μάσκα και η πρόβλεψη του μοντέλου, αναδεικνύοντας τη γενικά υψηλή ακρίβεια της πρόβλεψης παρά τη σχετικώς χαμηλή μέση Dice τιμή.	59
Εικόνα 23: Κατανομή του Dice score στο σύνολο επικύρωσης. Το γράφημα αποτυπώνει την απόδοση του μοντέλου σε κάθε δείγμα ξεχωριστά, αναδεικνύοντας την ύπαρξη outliers με πολύ χαμηλό score.	60
Εικόνα 24: Παράδειγμα περίπτωσης με μηδενικό Dice score. Παρά την ύπαρξη μικρών όγκων-στόχων στη ground truth, το μοντέλο δεν εντοπίζει την περιοχή ενδιαφέροντος, με αποτέλεσμα μηδενική επικάλυψη και αντίστοιχα μηδενικό Dice score.	61
Εικόνα 25: Καμπύλες εκπαίδευσης για το μοντέλο U-Net μετά τον αποκλεισμό των outliers. Απεικονίζονται η τιμή του Dice score στο validation set και η τιμή της συνάρτησης κόστους (loss) στο training set ανά εποχή.	63
Εικόνα 26: Καμπύλες εκπαίδευσης για το μοντέλο SegResNet μετά τον αποκλεισμό των outliers. Απεικονίζονται η τιμή του Dice score στο validation set και η τιμή της συνάρτησης κόστους (loss) στο training set ανά εποχή.	63

Ευρετήριο Πινάκων

Πίνακας 1: Συγκριτική αξιολόγηση συναρτήσεων απώλειας (U-Net baseline).....	56
Πίνακας 2: Τιμές Dice Score για τα τέσσερα εξεταζόμενα μοντέλα Από τα αρχικά αποτελέσματα προκύπτει ότι το SegMamba παρουσιάζει τη βέλτιστη απόδοση με Dice score 0.8, ακολουθούμενο από το SwinUNETR (0.65), το U-Net (0.64) και το SegResNet (0.61).57	
Πίνακας 3: Συγκριτική αξιολόγηση Dice Score με και χωρίς τον αποκλεισμό ακραίων τιμών (outliers) στο σύνολο επικύρωσης	62

Εισαγωγή

Η ιατρική απεικόνιση αποτελεί έναν από τους σημαντικότερους τομείς εφαρμογής της τεχνητής νοημοσύνης, με τις τεχνικές βαθιάς μάθησης (deep learning) να επιτυγχάνουν εντυπωσιακά αποτελέσματα σε καθήκοντα όπως η ταξινόμηση, η ανίχνευση ανωμαλιών και η τμηματοποίηση οργάνων ή παθολογικών περιοχών [1]. Μεταξύ αυτών, η αυτόματη τμηματοποίηση εγκεφαλικών όγκων σε εικόνες μαγνητικής τομογραφίας (MRI) αναδεικνύεται ως ιδιαίτερα απαιτητική, καθώς περιλαμβάνει σημαντικές διαφορές σε μορφολογία, μέγεθος και εντόπιση μεταξύ διαφορετικών ασθενών [2].

Τα μηνιγγιώματα είναι ο πιο συχνός τύπος πρωτοπαθούς εγκεφαλικού όγκου, παρουσιάζοντας όμως μεγάλη ετερογένεια ως προς τα κλινικά χαρακτηριστικά και την εικόνα τους στη μαγνητική τομογραφία [3]. Η έγκαιρη και ακριβής χαρτογράφηση της χωρικής τους έκτασης παίζει κρίσιμο ρόλο τόσο στη διάγνωση όσο και στον σχεδιασμό της θεραπείας. Παρά την κλινική σημασία της, η χειρωνακτική τμηματοποίηση παραμένει χρονοβόρα και επιρρεπής σε υποκειμενικές αποκλίσεις, αναδεικνύοντας την ανάγκη για αξιόπιστες αυτοματοποιημένες μεθόδους.

Στο πλαίσιο αυτό, αναπτύσσονται και αξιολογούνται αρχιτεκτονικές βαθιάς μάθησης για τρισδιάστατη τμηματοποίηση όγκων εγκεφάλου, οι οποίες επιχειρούν να εξισορροπήσουν την ακρίβεια με την αποδοτική χρήση υπολογιστικών πόρων. Στην παρούσα εργασία μελετάται η απόδοση διαφορετικών κατηγοριών νευρωνικών δικτύων—από συνελκτικά μοντέλα (CNNs) όπως το U-Net, έως σύγχρονες αρχιτεκτονικές με μηχανισμούς προσοχής ή state space modeling—στο πρόσφατα δημοσιευμένο σύνολο δεδομένων Meningioma-SEG-CLASS [4].

Ιδιαίτερη έμφαση δίνεται στην πειραματική διερεύνηση του τρόπου με τον οποίο οι περιορισμοί σε υπολογιστικούς πόρους και δεδομένα μπορούν να επηρεάσουν την απόδοση των μοντέλων, καθώς και σε τεχνικές προεπεξεργασίας και βελτιστοποίησης που μπορούν να μειώσουν αυτούς τους περιορισμούς. Η εργασία στοχεύει να συνεισφέρει στην αξιολόγηση των νέων προσεγγίσεων για την τμηματοποίηση μηνιγγιωμάτων, αναδεικνύοντας τα πλεονεκτήματα και τις προκλήσεις κάθε αρχιτεκτονικής.

Κεφάλαιο 1: Καρκίνος

1.1 Γενικά

Ο καρκίνος περιγράφει ένα σύνολο ασθενειών με κοινό χαρακτηριστικό τον ανεξέλεγκτο κυτταρικό πολλαπλασιασμό, ο οποίος προκαλείται από γενετικές μεταλλάξεις. Τα κύτταρα που υφίστανται αυτήν τη μεταβολή συγκεντρώνονται, σχηματίζοντας όγκους. Οι όγκοι δεν είναι απαραίτητα απειλητικοί: στην περίπτωση των καλοηθών μορφών, δεν παρατηρείται επιθετική συμπεριφορά ούτε άμεσος κίνδυνος για τη ζωή, εκτός αν υπάρξει πίεση σε γειτονικά όργανα [5].

Η σοβαρότητα εμφανίζεται όταν τα παθολογικά κύτταρα αποκτούν τη δυνατότητα να εισχωρούν σε υγιείς ιστούς, καταστρέφοντάς τους [6]. Αυτή η διείσδυση μπορεί να εκδηλωθεί τοπικά ή να εξελιχθεί σε συστηματική, με τα κύτταρα να διασπείρονται μέσω της κυκλοφορίας του αίματος και του λεμφικού συστήματος [7], οδηγώντας τελικά σε εκτεταμένη προσβολή του οργανισμού [8]. Τα καρκινικά κύτταρα, σε αντίθεση με τα φυσιολογικά, αγνοούν τους εσωτερικούς ρυθμιστικούς μηχανισμούς του κυτταρικού κύκλου, πολλαπλασιαζόμενα χωρίς περιορισμούς [9].

Οι επιπτώσεις της νόσου ποικίλλουν, καθώς εξαρτώνται τόσο από το σημείο εντόπισης του όγκου όσο και από το στάδιο ανάπτυξής του. Σε ορισμένες περιπτώσεις, ο όγκος επιφέρει φυσική παρεμπόδιση της λειτουργίας ενός οργάνου, όπως συμβαίνει στον καρκίνο του γαστρεντερικού συστήματος [10]. Σε άλλες, προκαλεί κυτταρική καταστροφή, όπως παρατηρείται στον καρκίνο του πνεύμονα [11].

Ο καρκίνος αποτελεί την πρωταρχική αιτία θανάτου στις αναπτυγμένες χώρες [12]. Η ποικιλομορφία των τύπων της νόσου και η πολυπλοκότητα της βιολογίας της έχουν καταστήσει την ανακάλυψη αποτελεσματικής θεραπείας μείζονα προτεραιότητα. Ενδεικτικά, η Ευρωπαϊκή Ένωση έχει επενδύσει 4 δισεκατομμύρια ευρώ για την ενίσχυση δράσεων πρόληψης και θεραπείας του καρκίνου [13].

1.2 Καρκίνος του Εγκεφάλου

Όταν ένας όγκος εντοπίζεται στον εγκέφαλο, η επικινδυνότητά του αυξάνεται σημαντικά, όχι μόνο λόγω της πιθανής επιθετικής συμπεριφοράς των κυττάρων, αλλά και εξαιτίας της ανατομικής και λειτουργικής ευαισθησίας της περιοχής. Οι όγκοι του εγκεφάλου διακρίνονται σε πρωτοπαθείς — που εμφανίζονται απευθείας στον εγκέφαλο — και σε μεταστατικούς, οι οποίοι προέρχονται από άλλες περιοχές του σώματος και μεταφέρονται στον εγκέφαλο μέσω της κυκλοφορίας [14].

Η πολυπλοκότητα του εγκεφαλικού καρκίνου έγκειται στο γεγονός ότι κάθε τμήμα του εγκεφάλου επιτελεί λειτουργίες ζωτικής σημασίας για τον άνθρωπο [15]. Η προσβολή έστω και μικρής περιοχής μπορεί να οδηγήσει σε σοβαρές νευρολογικές επιπλοκές. Αυτό καθιστά τη χειρουργική αντιμετώπιση ιδιαίτερα περίπλοκη και επικίνδυνη [16], σε πλήρη αντίθεση με

επεμβάσεις όπως η μαστεκτομή, όπου είναι δυνατή η πλήρης αφαίρεση του προσβεβλημένου ιστού [17].

Τα συμπτώματα που συνοδεύουν έναν εγκεφαλικό όγκο δεν είναι σταθερά· διαμορφώνονται από παράγοντες όπως το σημείο εμφάνισης, το μέγεθος και ο ρυθμός ανάπτυξης του νεοπλασματος. Δεδομένου ότι πλήττεται το κεντρικό νευρικό σύστημα, σχεδόν κάθε σωματική ή νοητική λειτουργία μπορεί να επηρεαστεί [6].

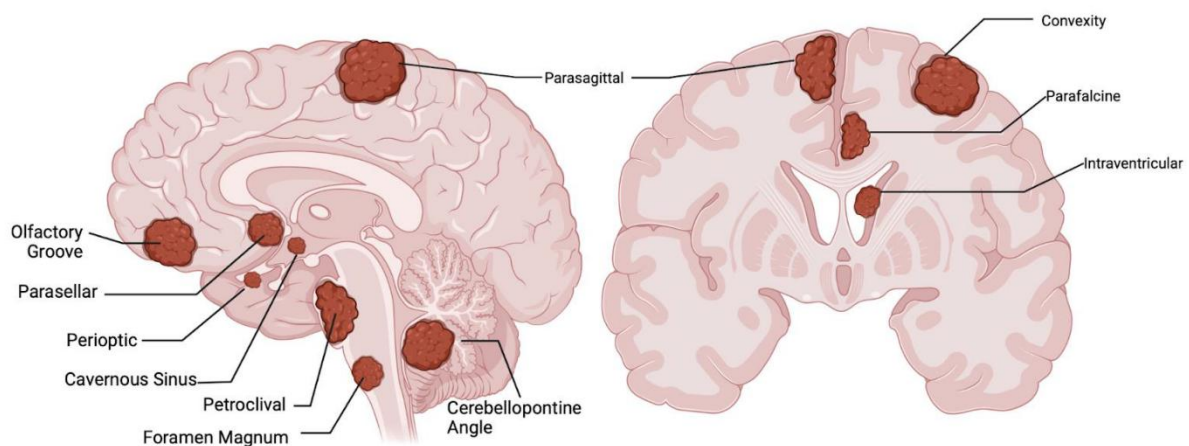
1.3 Μηνιγγίωμα

Το μηνιγγίωμα αποτελεί τον πιο συχνό πρωτοπαθή όγκο του κεντρικού νευρικού συστήματος, με προέλευση από τα μηνίγγια, δηλαδή τις προστατευτικές μεμβράνες που περιβάλλουν τον εγκέφαλο και τον νωτιαίο μυελό [3]. Στην πλειονότητά τους είναι καλοήγη (WHO Grade I), με αργή ανάπτυξη και σαφή όρια. Ωστόσο, υφίστανται και πιο επιθετικοί ιστολογικοί τύποι, οι οποίοι κατατάσσονται στους βαθμούς II (άτυπα) και III (αναπλαστικά ή κακοήγη) κατά την ταξινόμηση του Παγκόσμιου Οργανισμού Υγείας [3].

Από επιδημιολογική άποψη, τα μηνιγγιώματα εμφανίζονται συχνότερα στις γυναίκες, με τη μεγαλύτερη επίπτωση να παρατηρείται σε ηλικίες άνω των 60 ετών [18]. Η αιτία αυτής της διαφοράς ως προς το φύλο παραμένει υπό μελέτη, με ενδείξεις ότι ενδογενείς ορμονικοί παράγοντες ενδέχεται να διαδραματίζουν ρόλο στην παθογένεια. Σύμφωνα με στατιστικά δεδομένα από τις ΗΠΑ, τα μηνιγγιώματα αντιστοιχούν σε περίπου το 39% όλων των πρωτοπαθών όγκων του εγκεφάλου και του νωτιαίου μυελού [19].

Αξίζει να σημειωθεί ότι ένα ποσοστό των μηνιγγιωμάτων ανιχνεύεται τυχαία, καθώς δεν προκαλούν πάντα εμφανή κλινικά σημεία. Σε άλλες περιπτώσεις, μπορεί να απαιτείται περαιτέρω διερεύνηση και παρακολούθηση, ιδιαίτερα σε όγκους με ατυπικά χαρακτηριστικά ή ιστορικό ταχείας ανάπτυξης [20].

Meningioma Regions



Εικόνα 1: Κύριες ανατομικές περιοχές εντόπισης μηνιγγιωμάτων στον ανθρώπινο εγκέφαλο [21].

1.4 Αιτίες του Καρκίνου

Η προέλευση του καρκίνου είναι πολυπαραγοντική και περιλαμβάνει έναν σύνθετο μηχανισμό όπου αλληλεπιδρούν γενετικοί, περιβαλλοντικοί και συμπεριφορικοί παράγοντες. Ο κοινός παρονομαστής όλων των μορφών καρκίνου είναι η μεταβολή του γενετικού υλικού των κυττάρων, η οποία διαταράσσει τη φυσιολογική ρύθμιση της κυτταρικής διαίρεσης, οδηγώντας σε ανεξέλεγκτη κυτταρική αναπαραγωγή. Οι ακριβείς βιολογικοί μηχανισμοί πίσω από τις μεταλλάξεις αυτές δεν είναι ακόμη πλήρως κατανοητοί, ωστόσο έχουν εντοπιστεί σαφείς αιτιολογικοί συσχετισμοί.

Η κληρονομικότητα αποτελεί έναν από τους σημαντικότερους ενδογενείς παράγοντες. Ορισμένα είδη καρκίνου, όπως αυτός του παγκρέατος, σχετίζονται άμεσα ή σε σημαντικό βαθμό με γενετικές προδιαθέσεις [22]. Συγκεκριμένες μεταλλάξεις σε ογκοκατασταλτικά γονίδια, όπως τα BRCA1 και BRCA2 [23], έχουν συνδεθεί με ιδιαίτερα αυξημένο κίνδυνο ανάπτυξης καρκίνου, καθώς περιορίζουν σημαντικά την ικανότητα του κυττάρου να αποτρέπει την εξέλιξη της νόσου [24].

Παράλληλα, καθοριστικό ρόλο διαδραματίζει και ο τρόπος ζωής. Σημαντικοί εξωγενείς παράγοντες κινδύνου περιλαμβάνουν το κάπνισμα [25], την υπερβολική κατανάλωση αλκοόλ, την παχυσαρκία, τη μειωμένη σωματική δραστηριότητα και τις ανθυγιεινές διατροφικές συνήθειες. Επιπλέον, η χρόνια φλεγμονή, το ψυχολογικό στρες και οι δυσλειτουργίες του ανοσοποιητικού συστήματος έχουν επίσης συσχετιστεί με αυξημένη πιθανότητα καρκινογένεσης [26].

Εκτιμάται ότι έως και το 50% των περιστατικών καρκίνου στις αναπτυγμένες χώρες θα μπορούσαν να έχουν αποφευχθεί μέσω αλλαγών στον τρόπο ζωής [27]. Χαρακτηριστικό παράδειγμα αποτελεί ο καρκίνος του μαστού, όπου η μη απομάκρυνση του μητρικού γάλακτος μετά τον τοκετό έχει συνδεθεί με αυξημένο κίνδυνο [27].

Τέλος, η ηλικία αποτελεί ανεξάρτητο παράγοντα κινδύνου, καθώς η συχνότητα εμφάνισης καρκίνου αυξάνεται με την πάροδο των ετών [28]. Η έκθεση σε ιοντίζουσα ακτινοβολία, η οποία μπορεί να προκαλέσει αλλοιώσεις στο DNA, θεωρείται επίσης σημαντικός αιτιολογικός μηχανισμός για την εμφάνιση διαφόρων μορφών καρκίνου [23].

1.5 Συμπτώματα του Καρκίνου

1.5.1 Καρκίνος

Η συμπτωματολογία του καρκίνου εξαρτάται από τον ιστολογικό τύπο, τη θέση εμφάνισης και το στάδιο εξέλιξης της νόσου. Πολλοί τύποι καρκίνου στα αρχικά στάδια δεν εμφανίζουν καθόλου συμπτώματα ή συνοδεύονται από μη ειδικά σημεία που συχνά αγνοούνται από τους ασθενείς ή αποδίδονται σε άλλες καταστάσεις [29]. Καθώς η νόσος εξελίσσεται, μπορεί να προκύψουν τοπικά ή συστηματικά συμπτώματα, ανάλογα με την επίδραση του όγκου στα όργανα ή στους ιστούς.

Κοινές εκδηλώσεις της νόσου περιλαμβάνουν ανεξήγητη απώλεια βάρους, αίσθημα κόπωσης, αναιμία, πυρετό χωρίς προφανή αιτία, ή αιμορραγίες. Η παρουσία ενός αναπτυσσόμενου όγκου μπορεί να προκαλέσει μηχανική παρεμπόδιση ή λειτουργική δυσλειτουργία, ανάλογα με την τοποθεσία. Για παράδειγμα, όγκοι στο πεπτικό σύστημα μπορεί να επηρεάσουν την

πέψη ή την απορρόφηση θρεπτικών συστατικών [10], ενώ στον καρκίνο του πνεύμονα μπορεί να εμφανιστεί δύσπνοια ή επίμονος βήχας με αιμόπτυση [11].

Η ποικιλομορφία των συμπτωμάτων και η κατά περιπτώσεις γενική φύση τους συχνά καθυστερούν τη διάγνωση. Για τον λόγο αυτό, η έγκαιρη ανίχνευση της νόσου μέσω στοχευμένων προληπτικών ελέγχων έχει ιδιαίτερη σημασία. Μελέτες δείχνουν ότι τα προγράμματα προσυμπτωματικού ελέγχου (screening) συμβάλλουν σε ουσιαστικό βαθμό στη μείωση της θνησιμότητας σε πληθυσμούς υψηλού κινδύνου [30].

1.5.2 Καρκίνος Εγκεφάλου και Μηνιγγίωμα

Οι όγκοι του εγκεφάλου, είτε πρόκειται για πρωτοπαθείς είτε για μεταστατικούς, μπορούν να προκαλέσουν ένα ευρύ φάσμα νευρολογικών συμπτωμάτων. Τα συμπτώματα εξαρτώνται από τη θέση, το μέγεθος και τον ρυθμό ανάπτυξης του όγκου, καθώς και από τη φύση του (καλοήθης ή κακοήθης). Η παρουσία όγκου σε οποιοδήποτε σημείο του εγκεφάλου μπορεί να οδηγήσει σε αύξηση της ενδοκρανιακής πίεσης, διαταραχές στη λειτουργία των γειτονικών νευρικών ιστών και μεταβολές στην εγκεφαλική αιμάτωση [31].

Τα συχνότερα γενικά συμπτώματα περιλαμβάνουν πονοκεφάλους, ναυτία, εμέτους, διαταραχές όρασης ή ισορροπίας και αλλαγές στην ψυχική κατάσταση ή στη συμπεριφορά. Επιπλέον, η εμφάνιση επιληπτικών κρίσεων χωρίς προηγούμενο ιστορικό αποτελεί συχνά μία από τις πρώτες ενδείξεις ενδοκρανιακής μάζας [6].

Στην περίπτωση των μηνιγγιωμάτων, τα συμπτώματα οφείλονται συνήθως στην πίεση που ασκείται στον παρακείμενο εγκεφαλικό ιστό, δεδομένου ότι τα νεοπλάσματα αυτά αναπτύσσονται αργά και σπανίως τον διηθούν [20]. Η συμπτωματολογία ποικίλλει ανάλογα με την τοποθεσία εμφάνισης: όγκοι στο μετωπιαίο λοβό μπορεί να προκαλέσουν αλλαγές στη συμπεριφορά και δυσκολίες συγκέντρωσης, ενώ όγκοι στον κροταφικό λοβό συχνά σχετίζονται με επιληπτικές κρίσεις ή διαταραχές μνήμης [18].

Σε ορισμένες περιπτώσεις, ιδιαίτερα όταν ο όγκος βρίσκεται σε μη λειτουργικά ή περιφερειακά σημεία, το μηνιγγίωμα μπορεί να παραμείνει ασυμπτωματικό για μεγάλο χρονικό διάστημα και να εντοπιστεί τυχαία σε κάποιον απεικονιστικό έλεγχο [3].

1.6 Διάγνωση

Η ερμηνεία των συμπτωμάτων που σχετίζονται με τον καρκίνο αποτελεί μια ιδιαίτερα απαιτητική διαδικασία, η οποία προϋποθέτει εμπειρία, ιατρική εξειδίκευση και προσεκτική κλινική αξιολόγηση [32]. Ένα από τα βασικά προβλήματα στη διάγνωση είναι ότι τα συμπτώματα δεν είναι πάντοτε χαρακτηριστικά και συχνά μοιάζουν με εκείνα άλλων, λιγότερο σοβαρών παθήσεων. Για παράδειγμα, η επίμονη κόπωση και η απώλεια σωματικού βάρους μπορεί να αποδοθούν σε άγχος, ψυχολογική καταπόνηση ή διατροφικές διαταραχές, ενώ ο πονοκέφαλος συχνά θεωρείται απλό αποτέλεσμα έντασης ή υπερκόπωσης.

Το στοιχείο που διαφοροποιεί τη σοβαρή παθολογία από μια κοινή ενόχληση είναι κυρίως η **παρατεταμένη διάρκεια και η σταδιακή επιδείνωση των συμπτωμάτων**. Όταν αυτά επιμένουν στον χρόνο ή εμφανίζουν αυξανόμενη ένταση, είναι ένδειξη ότι απαιτείται περαιτέρω ιατρική διερεύνηση. Επιπλέον, η ταυτόχρονη παρουσία πολλών συμπτωμάτων — ακόμη και αν καθένα ξεχωριστά δεν είναι ανησυχητικό — αυξάνει την πιθανότητα να υποκρύπτεται μια σοβαρότερη κατάσταση.

Η διαφοροδιάγνωση πρέπει πάντα να λαμβάνει υπόψη το οικογενειακό ιστορικό και τους προσωπικούς παράγοντες κινδύνου του ατόμου, καθώς αυτοί επηρεάζουν τη συνολική κατάσταση. Η έγκαιρη αναγνώριση και ορθή ερμηνεία των κλινικών ενδείξεων μπορεί να οδηγήσει σε πρόωπη διάγνωση, γεγονός που είναι κρίσιμο για την επιτυχή αντιμετώπιση της νόσου [6].

Όταν υπάρχει βάσιμη υποψία εμφάνισης καρκίνου, ακολουθείται συνήθως ένας συνδυασμός εξειδικευμένων εξετάσεων. Στην περίπτωση του καρκίνου του εγκεφάλου, συχνά προηγείται **νευρολογικός έλεγχος** για την αξιολόγηση της λειτουργικότητας του νευρικού συστήματος [32]. Η **Μαγνητική Τομογραφία (MRI)** αποτελεί τη βασική διαγνωστική μέθοδο [16], [6], [32], καθώς παρέχει λεπτομερείς εικόνες χωρίς να επιβαρύνει τον οργανισμό με ακτινοβολία, όπως συμβαίνει με την αξονική τομογραφία.

Μετά τη λήψη των απεικονιστικών δεδομένων, ακολουθεί η διαδικασία της **τμηματοποίησης εικόνων**, κατά την οποία ο ακτινολόγος εντοπίζει και απομονώνει τις περιοχές που παρουσιάζουν αλλοιώσεις. Η διαδικασία αυτή είναι ιδιαίτερα **χρονοβόρα και απαιτητική** [33], καθώς απαιτεί υψηλό επίπεδο συγκέντρωσης και ακρίβειας. Ωστόσο, είναι επιρρεπής σε σφάλματα, τα οποία ενδέχεται να οφείλονται όχι μόνο στην εμπειρία του ιατρού, αλλά και στην ψυχολογική ή σωματική του κατάσταση τη στιγμή της εξέτασης [34]. Αυτοί οι παράγοντες καθιστούν σαφές ότι η **δυνατότητα μείωσης της επίδρασης του ανθρώπινου παράγοντα** στη διαδικασία της διάγνωσης θα μπορούσε να βελτιώσει σημαντικά την ακρίβεια και την ταχύτητα της αξιολόγησης.

1.7 Θεραπεία

Η αντιμετώπιση του καρκίνου αποτελεί μια εξαιρετικά πολύπλοκη διαδικασία, καθώς πρόκειται για μια νόσο με πολλαπλές μορφές, βιολογικές ιδιαιτερότητες και μεταβλητότητα από ασθενή σε ασθενή. Δεν υπάρχει μία καθολικά αποδεκτή θεραπευτική προσέγγιση που να εφαρμόζεται σε όλες τις περιπτώσεις. Η επιλογή μεθόδου εξαρτάται από πολλούς παράγοντες, όπως ο τύπος και το στάδιο του όγκου, η γενική κατάσταση της υγείας του ασθενούς, η διαθεσιμότητα των ιατρικών μέσων και, φυσικά, η συναίνεση και η προσωπική βούληση του ίδιου του ασθενούς [35].

Οι κύριες θεραπευτικές μέθοδοι που χρησιμοποιούνται για την αντιμετώπιση του καρκίνου είναι οι εξής:

- Χειρουργική αφαίρεση (αφαιρετική επέμβαση)
- Ακτινοθεραπεία
- Χημειοθεραπεία

Στην πράξη, τις περισσότερες φορές ακολουθείται ένας **συνδυασμός των παραπάνω μεθόδων**, προκειμένου να αυξηθεί η αποτελεσματικότητα της θεραπείας. Η χειρουργική προσέγγιση αποτελεί βασική επιλογή σε πολλές περιπτώσεις, ιδιαίτερα όταν ο όγκος μπορεί να αφαιρεθεί με ασφάλεια. Στόχος της επέμβασης είναι όχι μόνο η απομάκρυνση του ίδιου του όγκου, αλλά και ενός μικρού περιθωρίου υγιούς ιστού γύρω του, ώστε να μειωθεί ο κίνδυνος υποτροπής [17], [36]. Η χειρουργική αφαίρεση περιλαμβάνεται σχεδόν πάντοτε στα θεραπευτικά πρωτόκολλα για όγκους που είναι τεχνικά προσπελάσιμοι [17].

Η ακτινοθεραπεία εφαρμόζεται ως κύρια ή επικουρική μέθοδος, κυρίως όταν η χειρουργική επέμβαση δεν είναι εφικτή ή για την εξάλειψη υπολειπόμενων καρκινικών κυττάρων. Βασίζεται στη στόχευση του όγκου με ακτίνες Χ υψηλής ενέργειας, οι οποίες καταστρέφουν τα καρκινικά κύτταρα, ενέχοντας όμως τον κίνδυνο πρόκλησης βλαβών και στους παρακείμενους υγιείς ιστούς. Για την αντιμετώπιση αυτού του περιορισμού, η τεχνολογία των ακτινοθεραπευτικών συστημάτων έχει εξελιχθεί σημαντικά, επιτρέποντας μεγαλύτερη ακρίβεια στην εφαρμογή της ακτινοβολίας μέσω δυναμικής στόχευσης και βελτιστοποιημένης κατανομής της δόσης, μειώνοντας έτσι την επιβάρυνση στους υγιείς ιστούς [37].

Η χημειοθεραπεία είναι μια συστηματική θεραπεία που χρησιμοποιείται για την καταστροφή των καρκινικών κυττάρων, κυρίως όταν υπάρχει κίνδυνος να έχουν εξαπλωθεί σε άλλα σημεία του σώματος. Συχνά χορηγείται μετά το χειρουργείο για να μειωθεί η πιθανότητα υποτροπής ή, σε ορισμένες περιπτώσεις, πριν την επέμβαση ώστε να μειωθεί το μέγεθος του όγκου. Η απόφαση για χημειοθεραπεία εξαρτάται από τον τύπο και την επιθετικότητα του όγκου, καθώς και από τη συνολική κατάσταση του ασθενούς [36].

Κεφάλαιο 2: Μαγνητική Τομογραφία

2.1 Γενικά

Η μαγνητική τομογραφία (MT) αποτελεί μία από τις βασικότερες μεθόδους για τη μη επεμβατική απεικόνιση του εσωτερικού του ανθρώπινου σώματος, και χρησιμοποιείται ευρέως τόσο για τη διάγνωση όσο και για την αξιολόγηση της θεραπευτικής πορείας [38]. Η τεχνολογία αυτή βασίζεται στο φαινόμενο του **πυρηνικού μαγνητικού συντονισμού**, το οποίο χρησιμοποιήθηκε πρώτη φορά για απεικονιστικούς σκοπούς το 1973 από τον Paul Lauterbur, ενώ η ευρύτερη εφαρμογή της εγκρίθηκε το 1985 [38]. Έκτοτε, η MT θεωρείται μία από τις πιο σημαντικές εξελίξεις στην απεικονιστική ιατρική, ενδεχομένως η σημαντικότερη μετά την ανακάλυψη της ακτινογραφίας [39].

Σε αντίθεση με άλλες απεικονιστικές τεχνικές, όπως η αξονική τομογραφία, η μαγνητική τομογραφία **δεν χρησιμοποιεί ιονίζουσα ακτινοβολία**. Αντίθετα, βασίζεται στη χρήση **ραδιοσυχνοτήτων και ισχυρού μαγνητικού πεδίου**, γεγονός που την καθιστά ασφαλέστερη για τον οργανισμό [40]. Επίσης, η μέθοδος είναι **μη επεμβατική**, καθώς δεν απαιτεί την εισαγωγή κάποιου αισθητήρα ή οργάνου στο σώμα του ασθενούς [41].

Η MT παρέχει τη δυνατότητα **τριδιάστατης απεικόνισης** των ιστών, με υψηλή χωρική ανάλυση και αντίθεση, κάτι που διευκολύνει τον διαχωρισμό των επιμέρους ανατομικών δομών. Σε πολλές περιπτώσεις, η χρήση **σκιαγραφικού φαρμάκου** ενισχύει περαιτέρω την απεικονιστική διακριτικότητα. Τα σκιαγραφικά βασίζονται συνήθως σε παραμαγνητικά στοιχεία, όπως το γαδολίνιο, το οποίο αυξάνει την αντίθεση μεταξύ των ιστών, βελτιώνοντας την ευκρίνεια της εικόνας [38]. Ωστόσο, η χρήση τους δεν είναι πάντα αναγκαία, και έτσι η MT μπορεί να εκτελεστεί και χωρίς καμία φαρμακευτική παρέμβαση.

Ένα επιπλέον πλεονέκτημα της MT είναι η **προσαρμοστικότητα της**. Οι παραμέτροι της εξέτασης μπορούν να ρυθμιστούν ανάλογα με την πάθηση και την περιοχή ενδιαφέροντος, γεγονός που επιτρέπει την ευρεία εφαρμογή της τεχνικής σε διάφορα ιατρικά πεδία [42].

2.2 Πυρηνικός Μαγνητικός Συντονισμός

Το φαινόμενο του Πυρηνικού Μαγνητικού Συντονισμού (Nuclear Magnetic Resonance, NMR) αποτελεί τη θεμελιώδη φυσική αρχή πίσω από τη μαγνητική τομογραφία και έχει τις ρίζες του στην θεωρία κβαντομηχανικής [40]. Παρόλο που η πλήρης εξήγηση του φαινομένου απαιτεί την βαθύτερη γνώση κβαντομηχανικών εννοιών, εάν πρόκειται για την περιγραφή των βασικών αρχών της μαγνητικής τομογραφίας μπορούν να γίνουν ορισμένες κλασικές παραδοχές, που προσφέρουν επαρκή κατανόηση.

Οι πυρήνες των ατόμων αποτελούνται από πρωτόνια και νετρόνια, τα οποία διαθέτουν ιδιότητες όπως το **spin** (στροφορμή), που τους επιτρέπει να λειτουργούν σαν μικροσκοπικοί μαγνήτες. Ιδιαίτερο ενδιαφέρον παρουσιάζει το **πρωτόνιο του υδρογόνου**, καθώς είναι το απλούστερο άτομο (αποτελείται μόνο από ένα πρωτόνιο) και απαντά σε ιδιαίτερα υψηλή συγκέντρωση στους ιστούς του ανθρώπινου σώματος [41], [40]. Το πρωτόνιο του υδρογόνου, λόγω της στροφορμής και του θετικού φορτίου του, αναπτύσσει **μαγνητική ροπή**, με αποτέλεσμα να συμπεριφέρεται σαν μαγνητικό δίπολο [41].

Όταν ένα τέτοιο σωματίδιο τοποθετείται σε **στατικό εξωτερικό μαγνητικό πεδίο B_0** , οι δυνατοί προσανατολισμοί του είναι δύο: είτε παράλληλος είτε αντιπαράλληλος προς το πεδίο. Οι καταστάσεις αυτές αντιστοιχούν σε διαφορετικά ενεργειακά επίπεδα, με την **παράλληλη ευθυγράμμιση** να είναι **ενεργειακά πιο σταθερή** (ευνοϊκή) [40]. Αν και η διαφορά μεταξύ των δύο καταστάσεων είναι πολύ μικρή, αυτή η **υπεροχή** των πρωτονίων στη χαμηλότερη ενεργειακή κατάσταση δημιουργεί μια **καθαρή μαγνήτιση** προς την κατεύθυνση του πεδίου, δηλαδή στον άξονα z [42].

Επιπλέον, λόγω της στροφορμής τους, οι πυρήνες δεν ευθυγραμμίζονται αμέσως με το εξωτερικό πεδίο, αλλά εκτελούν μία χαρακτηριστική **κυκλική κίνηση γύρω από τον άξονα του πεδίου**. Η κίνηση αυτή ονομάζεται **μετάπτωση Larmor** και η συχνότητά της, γνωστή ως **Larmor frequency**, δίνεται από την εξίσωση:

$$\omega_0 = \gamma B_0$$

όπου:

- ω_0 : η συχνότητα Larmor (σε rad/s),
- γ : ο γυρομαγνητικός λόγος, χαρακτηριστικός για κάθε πυρήνα (για το υδρογόνο είναι περίπου 42.58 MHz/Tesla),
- B_0 : η ένταση του μαγνητικού πεδίου [42].

Η συχνότητα αυτή καθορίζει τη συμπεριφορά των πυρήνων όταν διεγείρονται από **παλμούς ραδιοσυχνοτήτων (RF)**, στοιχείο που είναι κρίσιμο για την παραγωγή σήματος στη Μαγνητική Τομογραφία.

Καθώς η φάση προσανατολισμού των πυρήνων είναι τυχαία, οι συνιστώσες του συνολικού μαγνητικού πεδίου στον άξονα x και y αλληλοαναιρούνται, με αποτέλεσμα η συνολική μαγνήτιση να εμφανίζεται μόνο στον άξονα z [42]. Αυτό το σταθερό διάνυσμα μαγνήτισης είναι το μετρήσιμο φυσικό μέγεθος που αξιοποιείται στην MRI. Το γεγονός ότι το υδρογόνο υπάρχει σε μεγάλη αφθονία στο ανθρώπινο σώμα το καθιστά ιδανικό στόχο για την απεικόνιση μέσω Μαγνητικής Τομογραφίας.

Η ενεργειακή διαφορά μεταξύ των δύο καταστάσεων προσανατολισμού των πυρήνων δίνεται από τη σχέση:

$$\Delta E = \hbar \gamma B_0$$

όπου:

- $\hbar$: η σταθερά του Planck διαιρεμένη με 2π ,
- γ : ο γυρομαγνητικός λόγος,
- B_0 : η ένταση του εξωτερικού μαγνητικού πεδίου.

Η σχέση αυτή εξηγεί γιατί όσο ισχυρότερο είναι το εξωτερικό πεδίο, τόσο μεγαλύτερη είναι η ενεργειακή διαφορά και, επομένως, τόσο πιο ευδιάκριτο το σήμα που μπορεί να ληφθεί.

2.3 RF Συντονισμός

Όπως σε κάθε φυσικό σύστημα ταλάντωσης, έτσι και στο φαινόμενο του Πυρηνικού Μαγνητικού Συντονισμού, υπάρχει μία **ιδιοσυχνότητα**, η οποία ονομάζεται **συχνότητα Larmor** και αντιστοιχεί στη συχνότητα με την οποία περιστρέφεται (εκτελεί μετάπτωση) το διάνυσμα μαγνήτισης ενός πυρήνα γύρω από το εξωτερικό μαγνητικό πεδίο B_0 . Όταν εφαρμόζεται ένα επιπλέον εναλλασσόμενο μαγνητικό πεδίο B_1 , κάθετο προς το B_0 , το οποίο **ταλαντώνεται ακριβώς με τη συχνότητα Larmor**, τότε ο πυρήνας **βρίσκεται σε συντονισμό** και είναι δυνατόν να απορροφήσει ενέργεια από αυτό το πεδίο [41], [40], [42].

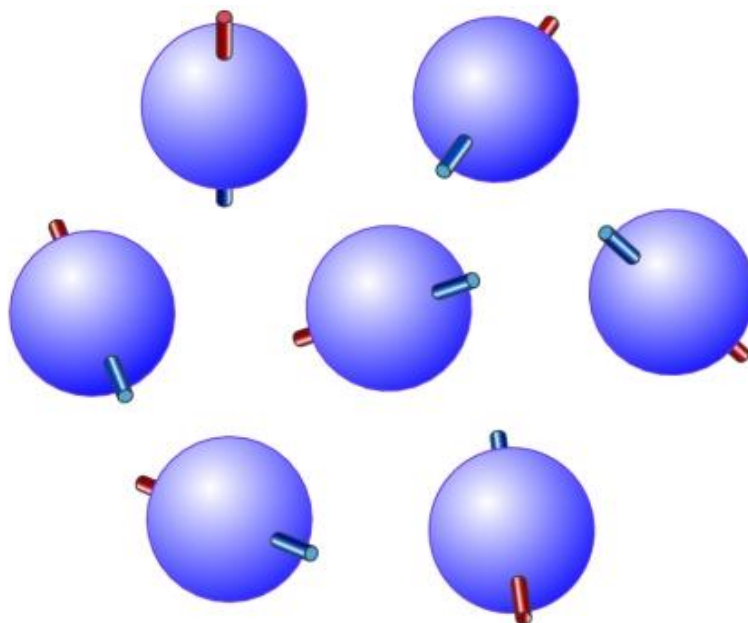
Για την πληρέστερη κατανόηση του φαινομένου, υιοθετείται ένα **περιστρεφόμενο σύστημα αναφοράς**. Σε αυτό το σύστημα, που περιστρέφεται με την ίδια συχνότητα με τη μετάπτωση Larmor των πυρήνων, το στατικό πεδίο B_0 παρατηρείται ως σταθερό και θεωρείται μηδενικό, ενώ το διάνυσμα μαγνήτισης εμφανίζεται ακίνητο στον άξονα z [40]. Σε αυτό το πλαίσιο εφαρμόζεται το εναλλασσόμενο πεδίο B_1 , το οποίο **στο νέο σύστημα αναφοράς εμφανίζεται ως σταθερό διάνυσμα** στο επίπεδο $x-y$, κάθετο προς τον άξονα z [42].

Η εφαρμογή του πεδίου B_1 οδηγεί στη **μετατόπιση του διανύσματος μαγνήτισης** από τον άξονα z προς το επίπεδο $x-y$. Η κλίση αυτή ορίζεται από τη γωνία **περιστροφής** της μαγνήτισης:

$$\beta = \omega_L \cdot t_p$$

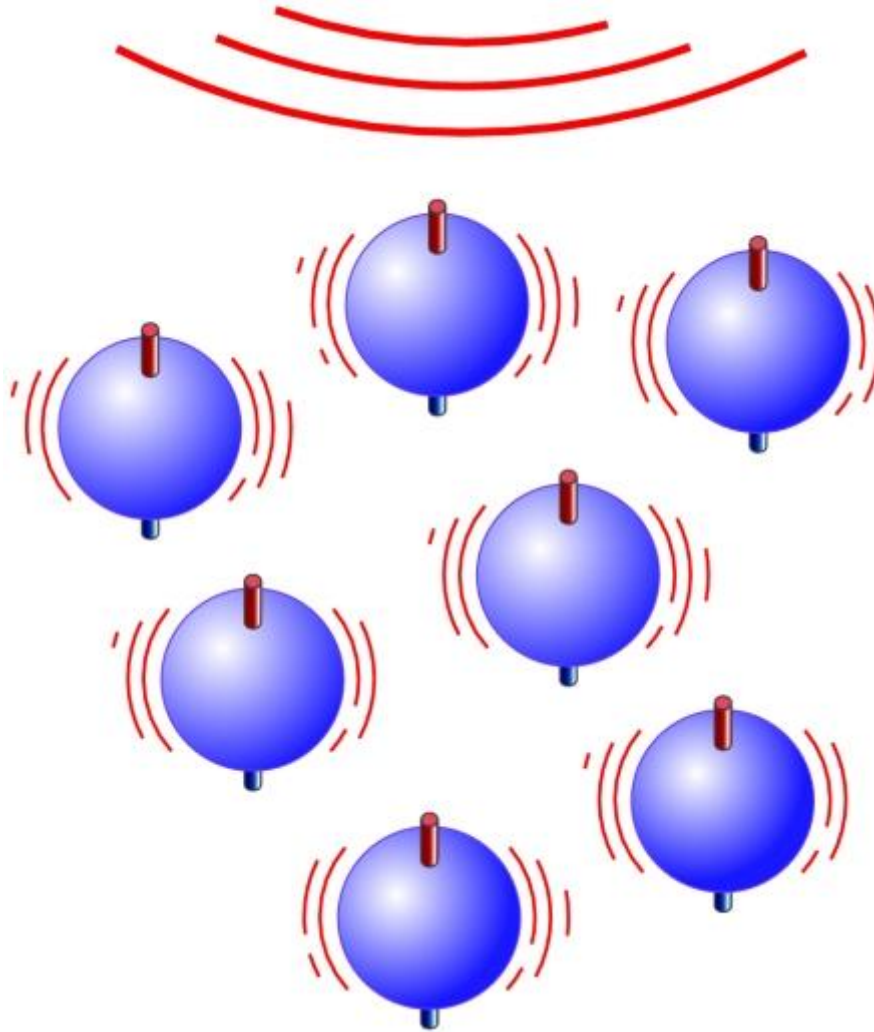
όπου:

- β : η γωνία που σχηματίζεται με τον άξονα z ,
- ω_L : η συχνότητα Larmor,
- t_p : ο χρόνος εφαρμογής του παλμού RF.



Εικόνα 2: Το πρωτόνιο του υδρογόνου συμπεριφέρεται όπως ένας μικρός ραβδόμορφος μαγνήτης. Υπό κανονικές συνθήκες, αυτοί οι "μαγνήτες" πρωτονίων υδρογόνου περιστρέφονται μέσα στο σώμα με τους άξονές τους ευθυγραμμισμένους τυχαία [43].

Το αποτέλεσμα είναι η **ανάδυση ενός εγκάρσιου (transverse) μαγνητικού πεδίου** στο επίπεδο $x-y$ το οποίο προηγουμένως μηδενιζόταν λόγω της αταξίας φάσης των πυρήνων. Το πεδίο αυτό είναι **παρατηρήσιμο** και παρέχει τη βάση για τη λήψη σήματος στη μαγνητική τομογραφία [41], [42].



Εικόνα 3: Όταν το σώμα τοποθετείται μέσα σε ένα ισχυρό μαγνητικό πεδίο, όπως αυτό ενός μαγνητικού τομογράφου, οι άξονες των πρωτονίων ευθυγραμμίζονται. Αυτή η ομοιόμορφη ευθυγράμμιση δημιουργεί ένα μαγνητικό διάνυσμα προσανατολισμένο κατά μήκος του άξονα του μαγνητικού τομογράφου [43].

Αμέσως μετά την παύση του παλμού RF, το μαγνητικό διάνυσμα αρχίζει να επανέρχεται στην αρχική κατάσταση ισορροπίας, δηλαδή να επαναπροσανατολίζεται παράλληλα με το B_0 . Η διαδικασία αυτή ονομάζεται **επαναφορά ή αποκατάσταση (relaxation)**, και περιγράφεται μέσω δύο βασικών χρονικών σταθερών: τον χρόνο διαμήκους χαλάρωσης T_1 και τον χρόνο εγκάρσιας χαλάρωσης T_2 [42].

- Ο χρόνος T_1 αφορά την **αποκατάσταση της συνιστώσας της μαγνήτισης στον άξονα z** . Αντιπροσωπεύει το χρονικό διάστημα που απαιτείται για να επιστρέψει η μαγνήτιση

στον άξονα του εξωτερικού πεδίου, μέσω **μεταφοράς ενέργειας** από τους πυρήνες στο περιβάλλον τους.

- Ο χρόνος T_2 αναφέρεται στην **απώλεια της συνοχής των φάσεων των πυρήνων** στο επίπεδο $x-y$, οδηγώντας στη βαθμιαία εξαφάνιση της εγκάρσιας μαγνήτισης. Αυτό οφείλεται κυρίως σε τοπικές ανομοιογένειες του πεδίου και διακυμάνσεις στο μικροπεριβάλλον των πυρήνων [42].

Οι χρόνοι αυτοί διαφέρουν από ιστό σε ιστό, καθώς εξαρτώνται από τη χημική σύσταση και την αλληλεπίδραση των μορίων. Αυτές οι διαφορές αποτελούν τη βασική αρχή πίσω από τη δημιουργία αντίθεσης εικόνας στη Μαγνητική Τομογραφία [45].

2.4 Χωρικός Εντοπισμός του Σήματος

Η διαδικασία του πυρηνικού μαγνητικού συντονισμού παρέχει πληροφορίες για ιδιότητες των ιστών, όπως οι χρόνοι χαλάρωσης και η πυκνότητα των πυρήνων υδρογόνου. Ένα κρίσιμο ζήτημα, ωστόσο, είναι ο **προσδιορισμός της χωρικής προέλευσης του σήματος**. Επειδή το μαγνητικό πεδίο B_0 είναι ομοιογενές και διαχέεται σε όλο τον χώρο, κάθε σημείο του σώματος συμμετέχει στην παραγωγή σήματος ταυτόχρονα, καθιστώντας αδύνατη τη διάκριση από ποιο σημείο του σώματος προέρχεται κάθε σήμα.

Για την επίλυση αυτού του προβλήματος, χρησιμοποιούνται **βαθμωτά μαγνητικά πεδία** (magnetic field gradients) κατά μήκος των τριών καρτεσιανών αξόνων. Τα βαθμωτά αυτά συναθροίζονται στο κύριο μαγνητικό πεδίο B_0 , μεταβάλλοντας την έντασή του κατά τρόπο γραμμικό σε κάθε άξονα. Αυτό έχει ως αποτέλεσμα η **συχνότητα Larmor των πυρήνων να διαφοροποιείται χωρικά**, σύμφωνα με την εξίσωση:

$$\omega_0 = \gamma \cdot B_0$$

Ωστόσο το B_0 πλέον διαφέρει κατά θέση λόγω των βαθμωτών και επομένως η συχνότητα Larmor γίνεται **μοναδική για κάθε σημείο** στο σώμα. Έτσι, όταν εφαρμοστεί ένας RF παλμός συγκεκριμένης συχνότητας, διεγείρονται μόνο οι πυρήνες που βρίσκονται στο σημείο με την αντίστοιχη συχνότητα Larmor [44], [46]. Με αυτόν τον τρόπο επιτυγχάνεται **χωρική επιλεκτικότητα** και κατά συνέπεια, προσδιορίζεται η θέση από την οποία προήλθε το σήμα [45].

Αναλόγως με τη διεύθυνση των βαθμωτών, επιτυγχάνεται η **επιλογή τομής (slice selection)**, η **κωδικοποίηση φάσης** και η **κωδικοποίηση συχνότητας**, που είναι απαραίτητες για την απεικόνιση στον τρισδιάστατο χώρο. Η δυνατότητα επιλεκτικής διέγερσης σημείων στο σώμα μέσω των βαθμωτών αποτελεί το θεμέλιο της χωρικής απεικόνισης στη μαγνητική τομογραφία.

Τέλος, η συνολική πληροφορία που συλλέγεται αποθηκεύεται σε έναν πίνακα δεδομένων γνωστό ως **k-χώρος**, στον οποίο καταγράφονται τα σήματα με βάση τις **φασικές και συχνοτικές** τους συνιστώσες. Η εφαρμογή του **Μετασχηματισμού Fourier** σε αυτό το σύνολο δεδομένων επιτρέπει την **ανακατασκευή της τελικής εικόνας**, με ακριβή **χωρική αντιστοίχιση** των σημάτων στους ιστούς.

2.5 Μαγνητικός Τομογράφος

Ο βασικός εξοπλισμός για τη μαγνητική τομογραφία περιλαμβάνει τον **μαγνήτη**, ο οποίος δημιουργεί το σταθερό εξωτερικό μαγνητικό πεδίο B_0 . Ανάλογα με την τεχνολογία, ο μαγνήτης μπορεί να είναι **μόνιμος, ηλεκτρομαγνήτης ή υπεραγώγιμος**. Οι δύο τελευταίοι τύποι στηρίζονται σε αγωγίμες περιελίξεις (πηνία) που διαρρέονται από ισχυρό ηλεκτρικό ρεύμα. Στην περίπτωση των απλών αγωγών, ένα σημαντικό μέρος της ενέργειας μετατρέπεται σε θερμότητα, περιορίζοντας την ένταση του παραγόμενου πεδίου. Αντίθετα, οι **υπεραγώγιμοι μαγνήτες** λειτουργούν σε πολύ χαμηλές θερμοκρασίες, επιτρέποντας την παραγωγή ισχυρών μαγνητικών πεδίων χωρίς απώλειες, καθιστώντας τους κατάλληλους για χρήση σε τομογράφους υψηλής απόδοσης.

Το δεύτερο βασικό στοιχείο είναι το **σύστημα RF (Radio Frequency)**, το οποίο αποτελείται από πηνία που παράγουν το μεταβλητό μαγνητικό πεδίο B_1 , απαραίτητο για τη διέγερση των πυρήνων. Τα πηνία αυτά τοποθετούνται γύρω από την υπό εξέταση περιοχή και λειτουργούν σε συχότητες ραδιοκυμάτων (15–80 MHz), με πεδία έντασης της τάξης των μTesla .

Τέλος, το σύστημα συμπληρώνεται από **ηλεκτρονικά κυκλώματα υψηλής ακρίβειας**, τα οποία είναι υπεύθυνα για:

- τη δημιουργία των παλμών RF,
- την ανίχνευση του σήματος που εκπέμπεται από τους διεγερμένους πυρήνες,
- την ψηφιακή επεξεργασία του σήματος και,
- την τελική παρουσίαση της εικόνας στον χρήστη μέσω εξειδικευμένου λογισμικού.

Η ποιότητα της εικόνας που εξάγεται εξαρτάται από το είδος και την ακρίβεια αυτών των υποσυστημάτων, καθιστώντας την επιλογή και τη ρύθμιση του εξοπλισμού κρίσιμο παράγοντα για την κλινική και ερευνητική εφαρμογή της μαγνητικής τομογραφίας.



Εικόνα 4: Μαγνητικός τομογράφος τύπου MRI, ενδεικτικός του εξοπλισμού που χρησιμοποιείται στην τμηματοποίηση εγκεφαλικών εικόνων [47].

Κεφάλαιο 3: Μηχανική Μάθηση και Βαθιά Μάθηση

3.1 Μηχανική Μάθηση

3.1.1 Εισαγωγή

Η Μηχανική Μάθηση (Machine Learning – ML) αποτελεί έναν από τους βασικούς πυλώνες της σύγχρονης Τεχνητής Νοημοσύνης. Πρόκειται για το πεδίο της επιστήμης των υπολογιστών που εστιάζει στην ανάπτυξη αλγορίθμων ικανών να εντοπίζουν μοτίβα σε δεδομένα και να βελτιώνονται μέσω εκπαίδευσης, χωρίς να έχουν προγραμματιστεί ρητά για την εκτέλεση συγκεκριμένων εργασιών.

Η Μηχανική Μάθηση διακρίνεται από την ικανότητά της να ενσωματώνει στατιστική, θεωρία πληροφορίας και υπολογιστική νοημοσύνη, με στόχο τη δημιουργία προγνωστικών ή περιγραφικών μοντέλων. Σε πρακτικό επίπεδο, αυτό επιτρέπει σε έναν αλγόριθμο να αναλύει δεδομένα, να εντοπίζει τάσεις ή υποκείμενες δομές και να γενικεύει αυτή τη γνώση για την επίλυση μελλοντικών προβλημάτων.

Υπάρχουν τρεις βασικές κατηγορίες Μηχανικής Μάθησης:

- **Επιβλεπόμενη Μάθηση (Supervised Learning)**, όπου το μοντέλο εκπαιδεύεται σε δεδομένα εισόδου και τις αντίστοιχες επιθυμητές εξόδους (ετικέτες).
- **Μη Επιβλεπόμενη Μάθηση (Unsupervised Learning)**, που στοχεύει στην εξαγωγή μοτίβων ή δομών από δεδομένα χωρίς ετικέτες.
- **Ενισχυτική Μάθηση (Reinforcement Learning)**, όπου ο αλγόριθμος μαθαίνει μέσω αλληλεπίδρασης με το περιβάλλον, λαμβάνοντας ενισχύσεις (ανταμοιβές ή ποινές).

Η πρόοδος σε αυτό το πεδίο έχει οδηγήσει στην ανάπτυξη μεθόδων που ξεπερνούν την παραδοσιακή στατιστική ανάλυση, προσφέροντας εργαλεία για την αντιμετώπιση σύνθετων προβλημάτων σε ποικίλες επιστήμες — από την οικονομία και τη ρομποτική έως την ιατρική απεικόνιση και τη βιοπληροφορική.

Στο πλαίσιο της ιατρικής, και ειδικότερα της επεξεργασίας ιατρικών εικόνων, η Μηχανική Μάθηση προσφέρει τη δυνατότητα αυτόματης αναγνώρισης παθολογικών περιοχών, πρόγνωσης ασθενειών, βελτίωσης της ακρίβειας της διάγνωσης και επιτάχυνσης του κλινικού έργου. Η ικανότητά της να εντοπίζει μορφολογικά και στατιστικά χαρακτηριστικά σε μεγάλες ποσότητες δεδομένων καθιστά τη Μηχανική Μάθηση ένα ιδιαίτερα χρήσιμο εργαλείο για την υποστήριξη των αποφάσεων των ειδικών.

Η ενσωμάτωση της Μηχανικής Μάθησης σε ιατρικές εφαρμογές, και ειδικότερα η αξιοποίησή της για την τμηματοποίηση εγκεφαλικών όγκων σε δεδομένα μαγνητικής τομογραφίας, αποτελεί το βασικό αντικείμενο της παρούσας διπλωματικής εργασίας. Στις επόμενες ενότητες παρουσιάζονται οι βασικές αρχές των σχετικών μεθόδων, καθώς και τα χαρακτηριστικά τους σε σχέση με την πρακτική εφαρμογή στην ανάλυση ιατρικών εικόνων.

3.1.2 Κατηγορίες Μηχανικής Μάθησης

Οι αλγόριθμοι Μηχανικής Μάθησης κατατάσσονται σε τρεις βασικές κατηγορίες, ανάλογα με τον τρόπο με τον οποίο μαθαίνουν από τα δεδομένα: **επιβλεπόμενη μάθηση**, **μη επιβλεπόμενη μάθηση** και **ενισχυτική μάθηση**. Η κατηγοριοποίηση αυτή αντικατοπτρίζει το είδος της πληροφορίας που παρέχεται στο μοντέλο κατά την εκπαίδευσή του, καθώς και το είδος της εργασίας που καλείται να εκτελέσει.

3.1.2.1 Επιβλεπόμενη Μάθηση

Η επιβλεπόμενη μάθηση (*supervised learning*) βασίζεται στην υπόθεση ότι για κάθε παράδειγμα εισόδου παρέχεται και η αντίστοιχη επιθυμητή έξοδος. Στόχος του αλγορίθμου είναι να μάθει μια **συνάρτηση** που να συσχετίζει τα δεδομένα εισόδου με τις αντίστοιχες εξόδους, έτσι ώστε να μπορεί να προβλέπει σωστά και για νέα, άγνωστα δεδομένα [48].

Η επιβλεπόμενη μάθηση εφαρμόζεται κυρίως σε δύο είδη προβλημάτων:

- **Ταξινόμηση (Classification):** όταν η έξοδος ανήκει σε διακριτές κατηγορίες. Παραδείγματα περιλαμβάνουν την αναγνώριση ψηφίων, την ανίχνευση ανεπιθύμητης αλληλογραφίας ή την ταξινόμηση παθολογικών εικόνων.
- **Παλινδρόμηση (Regression):** όταν η έξοδος είναι συνεχής αριθμητική τιμή, όπως στην πρόβλεψη τιμών χρηματιστηρίου ή στην πρόβλεψη της ηλικίας του εγκεφάλου [49].

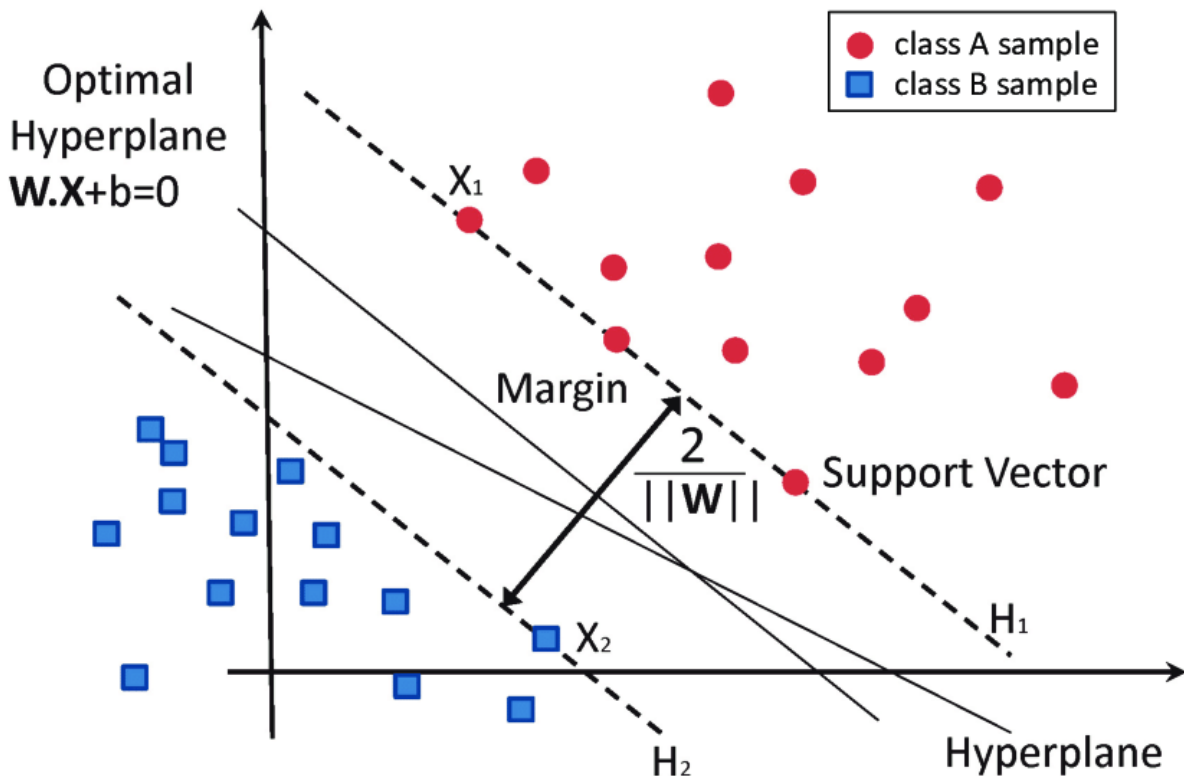
Η εκπαίδευση γίνεται πάνω σε ένα σύνολο ζευγών (x, y) , όπου x είναι η είσοδος και y η επιθυμητή έξοδος. Ο αλγόριθμος επιχειρεί να μάθει μια συνάρτηση $h(x)$ που προσεγγίζει τη συνάρτηση στόχο $f(x)$. Η εκμάθηση επιτυγχάνεται με την ελαχιστοποίηση μιας **συνάρτησης κόστους** $L(f(x), h(x))$, που μετρά την απόκλιση της πρόβλεψης από την πραγματική τιμή.

Η **γενίκευση**, δηλαδή η ικανότητα να αποδίδει σωστά σε άγνωστα δεδομένα, είναι ο βασικός στόχος. Για τον λόγο αυτό χρησιμοποιούνται τεχνικές όπως **cross-validation**, **regularization**, **feature selection** και **early stopping**.

Ακολουθούν διαδεδομένοι αλγόριθμοι επιβλεπόμενης μάθησης:

- **Δένδρα Απόφασης (Decision Trees),**
- **Νευρωνικά Δίκτυα (Neural Networks),**
- **Μηχανές Διανυσμάτων Υποστήριξης (Support Vector Machines),**
- **K-Πλησιέστεροι Γείτονες (K-Nearest Neighbors),**
- **Γραμμικές μέθοδοι παλινδρόμησης (Linear & Logistic Regression),**
- **Συνδυαστικά Μοντέλα (Ensemble models)** όπως τα *Random Forests* και το *Gradient Boosting*.

Στον χώρο της **ιατρικής απεικόνισης**, η επιβλεπόμενη μάθηση έχει κυρίαρχη θέση για εφαρμογές όπως **ταξινόμηση ιστών**, **εντοπισμό όγκων**, και **τμηματοποίηση** ανατομικών δομών.



Εικόνα 5: Γραφική αναπαράσταση γραμμικού ταξινομητή SVM. Το βέλτιστο υπερεπίπεδο (optimal hyperplane) διαχωρίζει τις δύο κλάσεις (κόκκινοι κύκλοι και μπλε τετράγωνα) μεγιστοποιώντας το περιθώριο μεταξύ των πιο κοντινών δειγμάτων, γνωστών ως support vectors. [50]

3.1.2.2 Μη Επιβλεπόμενη Μάθηση

Η **Μη Επιβλεπόμενη Μάθηση (Unsupervised Learning)** περιλαμβάνει μεθόδους κατά τις οποίες τα δεδομένα εισόδου δεν συνοδεύονται από αντίστοιχες ετικέτες εξόδου. Σε αυτό το πλαίσιο, το μοντέλο δεν εκπαιδεύεται βάσει προκαθορισμένων τιμών-στόχων, αλλά επιχειρεί να εντοπίσει υποκείμενες δομές, πρότυπα ή κανονικότητες που προκύπτουν από τα ίδια τα δεδομένα.

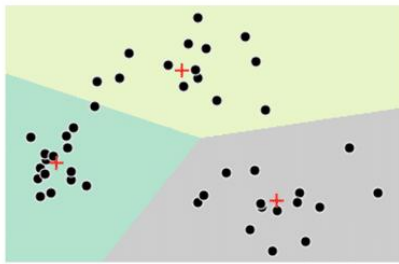
Η μη επιβλεπόμενη μάθηση είναι χρήσιμη σε περιπτώσεις όπου η επισήμανση δεδομένων είναι ακριβή ή ανέφικτη — χαρακτηριστική πρόκληση σε εφαρμογές βιοϊατρικών δεδομένων και εικόνων μεγάλης κλίμακας. Οι πιο ευρέως χρησιμοποιούμενες τεχνικές είναι η συσταδοποίηση, η μείωση διαστάσεων και η ανίχνευση αποκλίσεων ή ανωμαλιών.

► Συσταδοποίηση (Clustering)

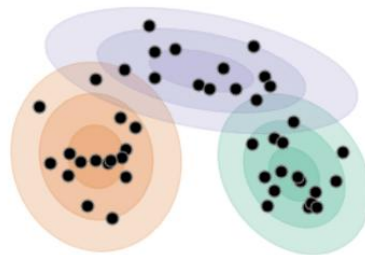
Η συσταδοποίηση (clustering) αποσκοπεί στην ομαδοποίηση των δεδομένων σε κατηγορίες (clusters) με βάση την εσωτερική τους ομοιότητα. Τα δείγματα που ανήκουν στην ίδια συστάδα παρουσιάζουν παρόμοια χαρακτηριστικά, ενώ οι συστάδες μεταξύ τους είναι διακριτές. Μεταξύ των πιο διαδεδομένων αλγορίθμων συγκαταλέγονται:

- **K-means**, που ελαχιστοποιεί την διασπορά εντός του ίδιου cluster,
- **Ιεραρχική Συσταδοποίηση (Hierarchical Clustering)**, η οποία δημιουργεί δενδρικές ιεραρχίες,

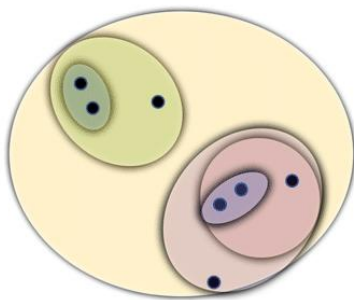
- **DBSCAN**, που βασίζεται στην πυκνότητα για την εύρεση περιοχών με υψηλή συγκέντρωση [51].



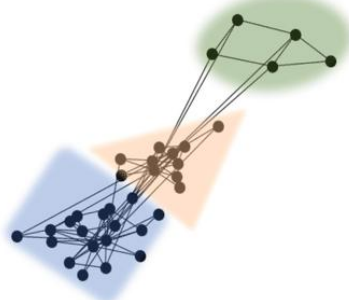
K-means clustering



Mixture model (Gaussian)



Hierarchical clustering



Graph based clustering

Εικόνα 6: Παραδείγματα διαφορετικών αλγορίθμων μη επιβλεπόμενης ομαδοποίησης (clustering): *K-means clustering*, *Gaussian mixture models*, *ιεραρχική ομαδοποίηση (hierarchical clustering)* και *ομαδοποίηση βασισμένη σε γράφους (graph-based clustering)*. Κάθε μέθοδος χρησιμοποιεί διαφορετικά κριτήρια για τον διαχωρισμό των δεδομένων σε ομάδες

► Μείωση Διαστάσεων (Dimensionality Reduction)

Όταν τα δεδομένα διαθέτουν μεγάλο πλήθος χαρακτηριστικών, ενδέχεται να περιέχουν θόρυβο ή πλεονάζουσα πληροφορία. Η μείωση διαστάσεων βοηθά στην απλοποίηση του διανύσματος εισόδου μέσω επιλογής των πιο κρίσιμων χαρακτηριστικών, διατηρώντας τα πιο σημαντικά, ενισχύοντας την αποδοτικότητα και την ερμηνευσιμότητα του μοντέλου.

Δύο κλασικές μέθοδοι είναι:

- **PCA (Principal Component Analysis)**, η οποία προβάλλει τα δεδομένα σε ορθογώνιες συνιστώσες μέγιστης διασποράς,
- **SVD (Singular Value Decomposition)**, που αποσυνθέτει πίνακες σε πιο απλές μορφές για αποδοτικότερη αναπαράσταση [53].

► Ανίχνευση Ανωμαλιών (Anomaly Detection)

Η ανίχνευση ανωμαλιών αφορά τον εντοπισμό σπάνιων ή ασυνήθιστων δειγμάτων που αποκλίνουν από το γενικό μοτίβο των δεδομένων. Αποτελεί κρίσιμη διαδικασία σε εφαρμογές όπως η **ανίχνευση απάτης**, η **πρόβλεψη βλαβών** και η **αναγνώριση παθολογικών ευρημάτων** σε ιατρικά δεδομένα.

Δημοφιλείς αλγόριθμοι περιλαμβάνουν:

- **Isolation Forest**, ο οποίος εντοπίζει αποκλίνοντα σημεία απομονώνοντάς τα στα αρχικά στάδια της κατασκευής δέντρων, αξιοποιώντας το γεγονός ότι απαιτούνται λιγότερα βήματα για τη διάκρισή τους.
- **Local Outlier Factor (LOF)**, που υπολογίζει την τοπική πυκνότητα κάθε δείγματος και εντοπίζει εκείνα που παρουσιάζουν σημαντικά χαμηλότερη πυκνότητα σε σχέση με τα γειτονικά τους [54].

3.1.2.3 Ενισχυτική Μάθηση

Η **Ενισχυτική Μάθηση (Reinforcement Learning, RL)** αποτελεί ένα διακριτό υποπεδίο της Μηχανικής Μάθησης, στο οποίο ένας πράκτορας (agent) μαθαίνει να λαμβάνει αποφάσεις μέσω της αλληλεπίδρασής του με ένα περιβάλλον. Η βασική αρχή είναι ότι ο πράκτορας μαθαίνει από την εμπειρία του, μέσα από ένα σύστημα **ανταμοιβών (rewards)** και **ποινών (penalties)**, ώστε να βελτιώσει σταδιακά τη συμπεριφορά του με στόχο τη **μέγιστη μακροπρόθεσμη απόδοση** [55].

Σε αντίθεση με την επιβλεπόμενη μάθηση, όπου παρέχονται ετικέτες για κάθε είσοδο, στην ενισχυτική μάθηση ο πράκτορας πρέπει να **ανακαλύψει μόνος του** ποιες ενέργειες οδηγούν σε επιθυμητά αποτελέσματα, μέσω δοκιμής και σφάλματος.

► Το Μοντέλο Απόφασης Markov

Η μαθησιακή διαδικασία βασίζεται συνήθως στο μαθηματικό πλαίσιο των **Markov Decision Processes (MDPs)**, το οποίο περιγράφει ένα δυναμικό σύστημα με τις εξής βασικές συνιστώσες:

- **Καταστάσεις (states) $s \in S$** : Η περιγραφή της τρέχουσας κατάστασης του περιβάλλοντος.
- **Ενέργειες (actions) $a \in A$** : Οι δυνατές επιλογές του πράκτορα.
- **Μεταβάσεις (transitions)**: Πιθανότητες μετάβασης από μία κατάσταση σε άλλη.
- **Συνάρτηση Ανταμοιβής (reward function)**: Ποσοτική ανατροφοδότηση για κάθε ενέργεια.
- **Πολιτική (policy)**: Η στρατηγική του πράκτορα για την επιλογή ενεργειών.
- **Συνάρτηση Αξίας (value function)**: Εκτίμηση της “χρησιμότητας” μιας κατάστασης ή δράσης μακροπρόθεσμα.

Η **ιδιότητα Markov** συνεπάγεται ότι η επόμενη κατάσταση εξαρτάται μόνο από την παρούσα κατάσταση και ενέργεια, όχι από το ιστορικό παρελθόν.

► Εξερεύνηση και Εκμετάλλευση

Ένα κρίσιμο στοιχείο στην ενισχυτική μάθηση είναι η ισορροπία μεταξύ:

- **Εξερεύνησης (exploration)**: Η αναζήτηση νέων ενδεχομένων για κατανόηση του περιβάλλοντος.

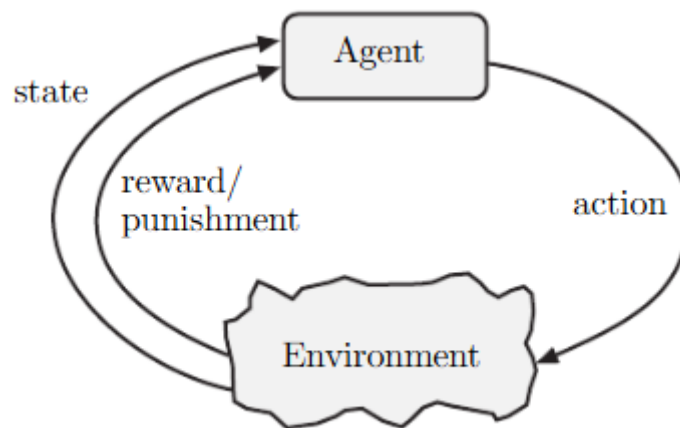
- **Εκμετάλλευσης (exploitation):** Η αξιοποίηση της ήδη αποκτηθείσας γνώσης για την επίτευξη υψηλών ανταμοιβών.

Η βέλτιστη πολιτική επιτυγχάνεται όταν ο πράκτορας έχει επαρκώς εξερευνήσει το περιβάλλον ώστε να μπορεί να εκμεταλλευτεί τις καλύτερες στρατηγικές.

► Εφαρμογές

Η ενισχυτική μάθηση έχει χρησιμοποιηθεί με επιτυχία σε προβλήματα όπως:

- **Αυτόνομα συστήματα και ρομποτική** (έλεγχος κίνησης),
- **Προβλήματα στρατηγικής και παιχνιδιών** (π.χ. AlphaGo),
- **Βελτιστοποίηση πόρων και scheduling**



Εικόνα 7: Αρχιτεκτονική ενισχυτικής μάθησης (Reinforcement Learning). Ο πράκτορας (agent) αλληλεπιδρά με το περιβάλλον (environment) εκτελώντας ενέργειες (actions) και λαμβάνοντας ως ανατροφοδότηση καταστάσεις (states) και ανταμοιβές ή τιμωρίες (rewards/punishments), με στόχο τη βελτιστοποίηση της στρατηγικής του [56].

3.1.3 Μηχανική Μάθηση στην Υγεία

Η Μηχανική Μάθηση έχει αναδειχθεί ως μία από τις πιο καθοριστικές τεχνολογίες στον τομέα της υγείας, προσφέροντας καινοτόμες λύσεις σε κρίσιμες προκλήσεις που σχετίζονται με τη διάγνωση, την πρόγνωση και τη θεραπεία ασθενειών. Η δυνατότητα των αλγορίθμων να μαθαίνουν από δεδομένα και να εξάγουν χρήσιμα πρότυπα επιτρέπει την υποστήριξη των ιατρικών αποφάσεων και τη βελτιστοποίηση της φροντίδας του ασθενούς.

Μία από τις βασικότερες εφαρμογές της Μηχανικής Μάθησης είναι η **ανάλυση ιατρικών εικόνων**, όπως οι ακτινογραφίες, οι αξονικές και μαγνητικές τομογραφίες. Μέσω τεχνικών ταξινόμησης και τμηματοποίησης, τα συστήματα Μηχανικής Μάθησης μπορούν να εντοπίζουν παθολογικές περιοχές, να διαχωρίζουν ιστούς και να υποστηρίζουν στη διάγνωση [57]. Επιπλέον, αλγόριθμοι πρόβλεψης μπορούν να εκτιμούν την εξέλιξη μιας πάθησης ή την πιθανή απόκριση ενός ασθενούς σε συγκεκριμένες θεραπείες.

Παράλληλα, σημαντική είναι η συμβολή της Μηχανικής Μάθησης στην **εξατομικευμένη ιατρική**, με την αξιοποίηση γονιδιωματικών δεδομένων, βιοδεικτών ή ιστορικών ασθενών. Τα μοντέλα μάθησης υποστηρίζουν την ανάπτυξη θεραπευτικών πλάνων που προσαρμόζονται στις ιδιαιτερότητες του κάθε ασθενή [58].

Τέλος, η χρήση τέτοιων τεχνικών σε **προληπτικές εφαρμογές**, όπως η πρόβλεψη κινδύνου εμφάνισης νοσημάτων ή η ανίχνευση σπάνιων ανωμαλιών σε μεγάλα πληθυσμιακά δείγματα, καθιστά τη Μηχανική Μάθηση πολύτιμο εργαλείο για τη δημόσια υγεία και τη λήψη αποφάσεων σε συστήματα υγειονομικής περίθαλψης.

3.2 Βαθιά Μάθηση

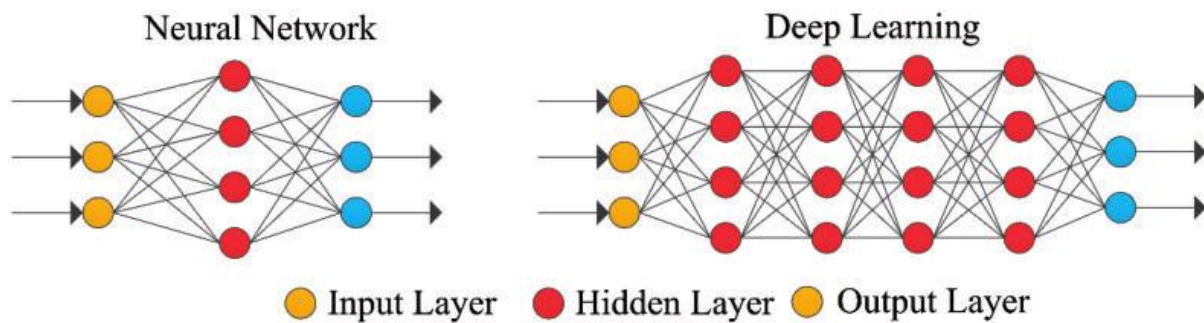
3.2.1 Εισαγωγή

Η βαθιά μάθηση (Deep Learning) αποτελεί έναν ταχέως αναπτυσσόμενο τομέα της μηχανικής μάθησης, ο οποίος βασίζεται στη χρήση πολυεπίπεδων νευρωνικών δικτύων για την αυτόματη εξαγωγή αναπαραστάσεων και την εκμάθηση σύνθετων λειτουργικών σχέσεων από τα δεδομένα. Σε αντίθεση με τις παραδοσιακές μεθόδους μηχανικής μάθησης, οι οποίες συχνά απαιτούν χειροκίνητη εξαγωγή χαρακτηριστικών, τα μοντέλα βαθιάς μάθησης είναι σε θέση να αντλούν ιεραρχικές δομές πληροφορίας απευθείας από τα ακατέργαστα δεδομένα, επιτρέποντας υψηλής ακρίβειας λύσεις σε προβλήματα όπως η ανάλυση εικόνας, η επεξεργασία φυσικής γλώσσας και η πρόβλεψη χρονοσειρών .

Ο όρος "βαθιά" αναφέρεται στο πλήθος των επιπέδων (layers) που συγκροτούν το νευρωνικό δίκτυο. Αυτά τα επίπεδα επιτρέπουν στο δίκτυο να μάθει διαδοχικά αφηρημένες αναπαραστάσεις των εισόδων: για παράδειγμα, σε μια εφαρμογή ταξινόμησης ιατρικών εικόνων, τα πρώτα επίπεδα μπορούν να εντοπίζουν απλές δομές όπως ακμές ή υφές, ενώ τα βαθύτερα επίπεδα αναγνωρίζουν πιο πολύπλοκα πρότυπα, όπως σχήματα όγκων ή χαρακτηριστικά παθολογίας.

Η βαθιά μάθηση υποστηρίζεται κυρίως από την εκθετική αύξηση των διαθέσιμων δεδομένων, την εξέλιξη των υπολογιστικών πόρων (π.χ. GPUs), καθώς και από την ανάπτυξη νέων αρχιτεκτονικών και τεχνικών εκπαίδευσης, όπως η κανονικοποίηση, οι βελτιστοποιητές τύπου Adam και οι τεχνικές κανονικοποίησης (regularization), που αντιμετωπίζουν ζητήματα υπερεκπαίδευσης [59].

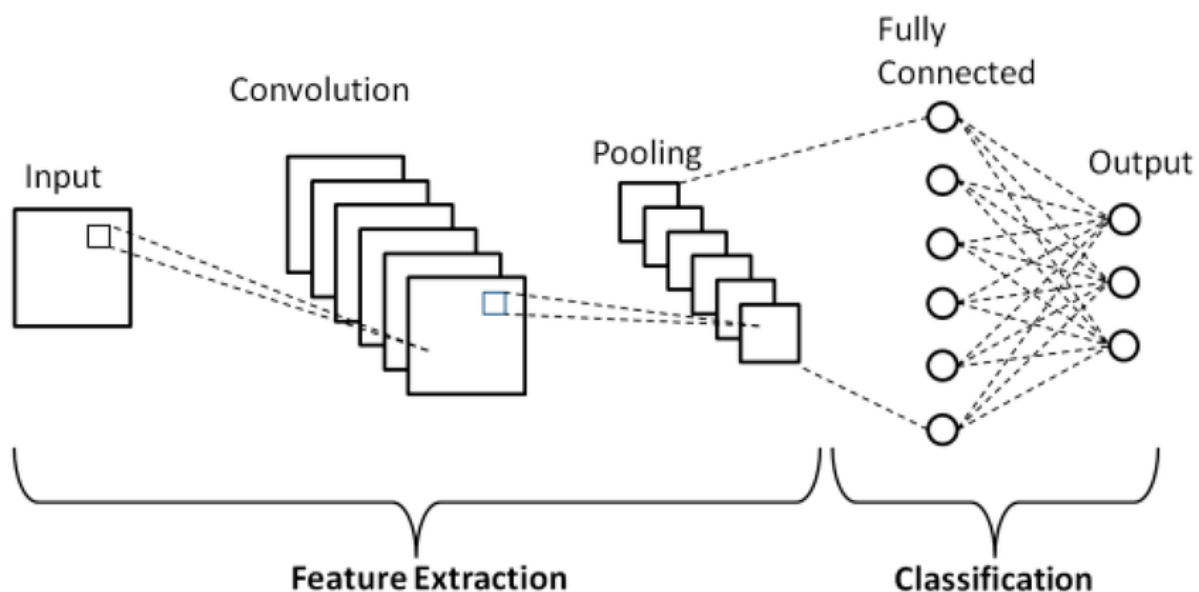
Στο πεδίο της ιατρικής απεικόνισης, η βαθιά μάθηση έχει οδηγήσει σε σημαντικές εξελίξεις σε εφαρμογές όπως η ταξινόμηση παθολογιών, η ανίχνευση όγκων, και η τμηματοποίηση ανατομικών δομών, με τις επιδόσεις πολλών μοντέλων να ανταγωνίζονται ή και να ξεπερνούν την ανθρώπινη ακρίβεια σε συγκεκριμένες εφαρμογές .



Εικόνα 8: Σχηματική σύγκριση μεταξύ ενός απλού νευρωνικού δικτύου και ενός βαθιού νευρωνικού δικτύου (deep neural network) [60].

3.2.2 Συνελικτικά Νευρωνικά Δίκτυα

Τα Συνελικτικά Νευρωνικά Δίκτυα (Convolutional Neural Networks, CNNs) αποτελούν μια από τις πιο επιτυχημένες και διαδεδομένες κατηγορίες αρχιτεκτονικών βαθιάς μάθησης, ειδικά στον τομέα της υπολογιστικής όρασης. Χρησιμοποιούνται για εργασίες όπως η ταξινόμηση εικόνων, η αναγνώριση αντικειμένων, η ανίχνευση αντικειμένων και η τμηματοποίηση. Ο πυρήνας της λειτουργίας τους είναι η μαθηματική πράξη της συνέλιξης, η οποία επιτρέπει στο δίκτυο να εντοπίζει τοπικά πρότυπα και δομές στην εικόνα εισόδου. Μια τυπική αρχιτεκτονική CNN αποτελείται από βασικά συστατικά, όπως συνελικτικά και υποδειγματοληπτικά επίπεδα, πλήρως συνδεδεμένες μονάδες, συναρτήσεις ενεργοποίησης, καθώς και κατάλληλη συνάρτηση κόστους και αλγόριθμο βελτιστοποίησης.

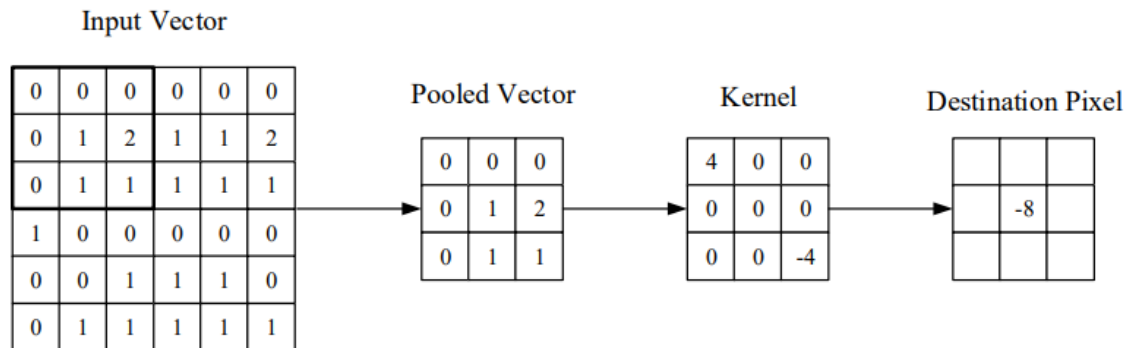


Εικόνα 9: Αρχιτεκτονική ενός συνελικτικού νευρωνικού δικτύου (CNN). Το δίκτυο περιλαμβάνει στάδια εξαγωγής χαρακτηριστικών (convolution και pooling) και στη συνέχεια πλήρως συνδεδεμένα στρώματα για την τελική ταξινόμηση [61].

3.2.2.1 Επίπεδα Συνέλιξης

Στα επίπεδα συνέλιξης (convolutional layers), εφαρμόζονται συνελικτικοί πυρήνες (ή φίλτρα) πάνω στα δεδομένα εισόδου, με σκοπό την εξαγωγή χαρακτηριστικών όπως ακμές, γωνίες ή

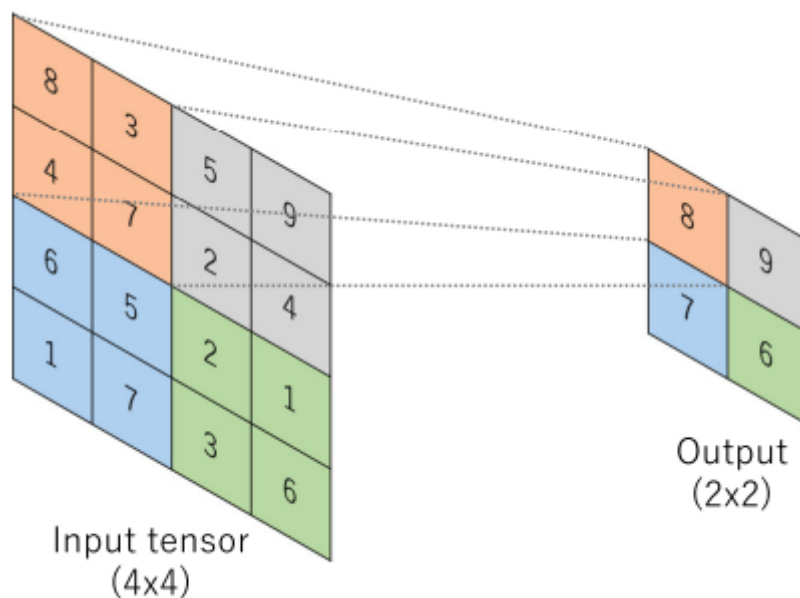
πιο σύνθετες μορφές. Κάθε φίλτρο ολισθαίνει πάνω στον χάρτη εισόδου με συγκεκριμένο βήμα (stride), ενώ συχνά εφαρμόζεται κατάλληλο γέμισμα (padding) ώστε να διατηρηθούν οι χωρικές διαστάσεις. Οι τιμές των πυρήνων προσδιορίζονται κατά τη διάρκεια της εκπαίδευσης μέσω της διαδικασίας οπισθοδιάδοσης σφάλματος (backpropagation) [62].



Εικόνα 10: Παράδειγμα λειτουργίας συνελκτικού νευρωνικού δικτύου: η αρχική εικόνα (Input Vector) υφίσταται υποδειγματοληψία (Pooling), ακολουθεί η εφαρμογή πυρήνα (Kernel) και τελικά υπολογίζεται η έξοδος (Destination Pixel) με συνέλιξη [63].

3.2.2.2 Επίπεδα Υποδειγματοληψίας

Τα επίπεδα υποδειγματοληψίας (pooling layers) χρησιμοποιούνται για τη μείωση της χωρικής διάστασης των χαρτών χαρακτηριστικών, γεγονός που οδηγεί σε μικρότερο υπολογιστικό κόστος και σε περιορισμό του κινδύνου υπερπροσαρμογής. Οι δύο πιο συχνά χρησιμοποιούμενοι τύποι είναι η υποδειγματοληψία μέγιστης τιμής (max pooling) και μέσης τιμής (average pooling). Κατά την εφαρμογή της, ένα μικρό παράθυρο (π.χ. 2×2) ολισθαίνει πάνω από τον χάρτη χαρακτηριστικών και αντικαθιστά κάθε τοπική περιοχή με μία ενιαία τιμή — είτε τη μέγιστη είτε τη μέση — ανάλογα με τον τύπο pooling. Αυτό έχει ως αποτέλεσμα τη συμπίεση των χωρικών διαστάσεων, διατηρώντας παράλληλα τα πιο αντιπροσωπευτικά στοιχεία της πληροφορίας [64].



Εικόνα 11: Παράδειγμα εφαρμογής υποδειγματοληψίας μέγιστης τιμής (*max pooling*) με παράθυρο 2×2 και βήμα 2. Από κάθε τοπική περιοχή του αρχικού πίνακα εισόδου (4×4) επιλέγεται η μέγιστη τιμή, παράγοντας έναν πιο συμπαγές πίνακα εξόδου (2×2) [65].

3.2.2.3 Πλήρως Συνδεδεμένα Επίπεδα

Τα πλήρως συνδεδεμένα επίπεδα (*fully connected layers*) βρίσκονται συνήθως στα τελευταία στάδια ενός νευρωνικού δικτύου και λειτουργούν ως ο τελικός επεξεργαστής της πληροφορίας. Κάθε νευρώνας σε αυτά τα επίπεδα είναι συνδεδεμένος με όλους τους νευρώνες του προηγούμενου επιπέδου, επιτρέποντας την συγκέντρωση και σύνθεση πληροφορίας από όλο το δίκτυο.

Πριν από την εφαρμογή ενός τέτοιου επιπέδου, τα δεδομένα εξόδου από τα συνελκτικά ή υποδειγματοληπτικά επίπεδα μετατρέπονται σε μονοδιάστατο διάνυσμα μέσω της διαδικασίας *flattening*, ώστε να είναι συμβατά με τη γραμμική δομή του επιπέδου. Τα πλήρως συνδεδεμένα επίπεδα εκτελούν τον τελικό υπολογισμό για την εκάστοτε πρόβλεψη όπως η ταξινόμηση ή η παλινδρόμηση, βασιζόμενα στον συνδυασμό των πληροφοριών που εξήχθησαν στα προηγούμενα στάδια.

3.2.2.4 Συναρτήσεις Ενεργοποίησης

Οι συναρτήσεις ενεργοποίησης προσδίδουν στο δίκτυο μη-γραμμικότητα, επιτρέποντάς του να μοντελοποιήσει πολύπλοκες σχέσεις μεταξύ εισόδου και εξόδου. Η πιο συχνά χρησιμοποιούμενη συνάρτηση στα CNNs είναι η **ReLU** (*Rectified Linear Unit*), η οποία ορίζεται ως:

$$ReLU(x) = \max(0, x)$$

Η ReLU είναι υπολογιστικά αποδοτική και περιορίζει το πρόβλημα των **εξαφανιζόμενων κλίσεων** (*vanishing gradients*), το οποίο επηρεάζει αρνητικά τη μάθηση σε βαθιά δίκτυα.

Παραλλαγές της ReLU είναι:

- **Leaky ReLU (LReLU):** επιτρέπει μικρές αρνητικές τιμές, ώστε να αποφεύγεται το φαινόμενο των «νεκρών νευρώνων».

$$LReLU(x) = x, \text{ για } x \geq 0 \text{ και } \alpha x, \text{ για } x < 0$$

- **Parametric ReLU (PReLU):** εισάγει την παράμετρο α ως **εκπαιδευσιμη μεταβλητή**, επιτρέποντας στο δίκτυο να προσαρμόσει τη μορφή της συνάρτησης ενεργοποίησης κατά τη διάρκεια της εκπαίδευσης [66].

Άλλες διαδεδομένες συναρτήσεις περιλαμβάνουν τη **sigmoid** και την **tanh**, οι οποίες χρησιμοποιούνταν ευρέως σε παλαιότερες αρχιτεκτονικές. Σήμερα αποφεύγονται σε βαθιά CNNs λόγω των προβλημάτων κορεσμού και εξαφάνισης των κλίσεων, αλλά εξακολουθούν να βρίσκουν εφαρμογή σε ειδικές περιπτώσεις, όπως σε εξόδους δυαδικής πρόβλεψης ή σε επαναληπτικές αρχιτεκτονικές.

3.2.2.5 Συναρτήσεις Σφάλματος

Οι συναρτήσεις σφάλματος (loss functions) αποτελούν θεμελιώδες στοιχείο στην εκπαίδευση νευρωνικών δικτύων. Στόχος τους είναι να ποσοτικοποιήσουν το πόσο αποκλίνουν οι προβλέψεις του μοντέλου από τις πραγματικές τιμές των δεδομένων εκπαίδευσης. Όσο μικρότερη είναι η τιμή της συνάρτησης σφάλματος, τόσο πιο «σωστό» θεωρείται το μοντέλο.

Οι συναρτήσεις σφάλματος διαφέρουν ανάλογα με το είδος του προβλήματος. Για παράδειγμα:

- Η **μέση τετραγωνική απόκλιση (Mean Squared Error, MSE)** είναι συνήθης επιλογή για προβλήματα παλινδρόμησης και δίνεται από τον τύπο:

$$MSE(y, \hat{y}) = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

- Για προβλήματα ταξινόμησης, χρησιμοποιείται συνήθως η **διασταυρούμενη εντροπία (Cross-Entropy Loss)**:

$$CE(y, \hat{y}) = - \sum_{i=1}^N \sum_{j=1}^C y_{ij} \log(\hat{y}_{ij})$$

Στην πράξη, η επιλογή της συνάρτησης σφάλματος επηρεάζει τη διαδικασία εκπαίδευσης και την τελική απόδοση του μοντέλου, καθώς καθορίζει ποιο σφάλμα θεωρείται «σημαντικό» και σε ποιο βαθμό πρέπει να διορθωθεί κατά την ενημέρωση των παραμέτρων του δικτύου.

3.2.2.6 Αλγόριθμοι Βελτιστοποίησης

Οι αλγόριθμοι βελτιστοποίησης αποτελούν τον μηχανισμό μέσω του οποίου ένα νευρωνικό δίκτυο «μαθαίνει», δηλαδή προσαρμόζει τις παραμέτρους του ώστε να ελαχιστοποιήσει τη συνάρτηση σφάλματος. Η βασική αρχή που διέπει την πλειονότητα των αλγορίθμων αυτών είναι η **μέθοδος κλίσης (Gradient Descent)**.

Η μέθοδος αυτή βασίζεται στον υπολογισμό της παραγώγου της συνάρτησης σφάλματος ως προς κάθε παράμετρο του δικτύου, και στην αντίστοιχη ενημέρωσή της προς την κατεύθυνση της μεγαλύτερης μείωσης του σφάλματος:

$$\theta_{t+1} = \theta_t - \eta \nabla L(\theta_t)$$

Όπου:

- θ είναι το διάνυσμα των παραμέτρων του μοντέλου,
- η είναι ο ρυθμός εκμάθησης (learning rate),
- L είναι η συνάρτηση σφάλματος.

Υπάρχουν διαφορετικές εκδοχές της **Gradient Descent**, όπως η **Batch Gradient Descent**, η **Stochastic Gradient Descent (SGD)** και η **Mini-Batch Gradient Descent**, οι οποίες διαφέρουν στον αριθμό των δειγμάτων που χρησιμοποιούνται για τον υπολογισμό των παραγώγων σε κάθε επανάληψη.

Πάνω σε αυτήν τη θεμελιώδη ιδέα έχουν αναπτυχθεί και πιο εξελιγμένοι αλγόριθμοι που ενσωματώνουν τεχνικές όπως ροπές, κανονικοποίηση ή προσαρμογή ρυθμού εκμάθησης. Αυτοί θα αναλυθούν πιο αναλυτικά στο Κεφάλαιο 4, στο πλαίσιο των πειραμάτων και των συγκριτικών αξιολογήσεων που πραγματοποιήθηκαν.

3.2.3 Εκπαίδευση Συνελκτικών Νευρωνικών Δικτύων

Η διαδικασία της εκπαίδευσης των συνελκτικών νευρωνικών δικτύων περιλαμβάνει διαδοχικά στάδια που στοχεύουν στην αξιοποίηση της ικανότητας των δικτύων αυτών να εξάγουν και να γενικεύουν σημαντικά χαρακτηριστικά από τα δεδομένα εισόδου. Από την αρχική προετοιμασία των δεδομένων μέχρι τον σχεδιασμό της αρχιτεκτονικής και την αξιολόγηση της απόδοσης του μοντέλου, κάθε στάδιο διαδραματίζει καθοριστικό ρόλο στην τελική αποτελεσματικότητα του συστήματος. Στη συνέχεια παρουσιάζονται συνοπτικά τα βασικά βήματα αυτής της διαδικασίας.

3.2.3.1 Προετοιμασία των Δεδομένων

Η αποτελεσματικότητα των συνελκτικών νευρωνικών δικτύων εξαρτάται σε μεγάλο βαθμό από την ποιότητα και την κατάλληλη προετοιμασία των δεδομένων που χρησιμοποιούνται κατά την εκπαίδευσή τους. Η διαδικασία αυτή περιλαμβάνει μια σειρά από κρίσιμα στάδια, τα οποία στοχεύουν τόσο στην αύξηση της ποικιλομορφίας όσο και στη βελτίωση της αντιπροσωπευτικότητας του συνόλου των δεδομένων.

Σύνολο και Επισημάνσεις Δεδομένων

Η συλλογή δεδομένων πρέπει να καλύπτει ένα ευρύ φάσμα περιπτώσεων ώστε να ενισχύεται η ικανότητα γενίκευσης του μοντέλου. Σε προβλήματα επιβλεπόμενης μάθησης, τα δεδομένα πρέπει να συνοδεύονται από τις αντίστοιχες ετικέτες (labels), οι οποίες ιδανικά θα έχουν επισημανθεί από ειδικούς του πεδίου για ελαχιστοποίηση του σφάλματος.

Προεπεξεργασία και Κανονικοποίηση

Η προεπεξεργασία περιλαμβάνει βήματα όπως:

- **Καθαρισμός** των δεδομένων από θόρυβο, λανθασμένες ή ελλιπείς τιμές.
- **Κανονικοποίηση** (normalization), ώστε τα δεδομένα να προσαρμόζονται σε κοινή κλίμακα, κάτι που επιταχύνει τη διαδικασία εκπαίδευσης και μειώνει την ευαισθησία του μοντέλου σε χαρακτηριστικά με μεγάλες αποκλίσεις τιμών.
- **Αλλαγή μεγέθους** (resizing) εικόνων ώστε να είναι συμβατές με τις απαιτήσεις εισόδου του εκάστοτε μοντέλου.

Επαύξηση Δεδομένων (Data Augmentation)

Για να ενισχυθεί η γενίκευση του μοντέλου και να αντιμετωπιστεί το φαινόμενο της υπερπροσαρμογής (overfitting), εφαρμόζονται τεχνικές επαύξησης δεδομένων. Αυτές περιλαμβάνουν:

- **Γεωμετρικούς μετασχηματισμούς** όπως περιστροφή, αναστροφή, μετατόπιση και κλιμάκωση.
- **Προσθήκη θορύβου, εφαρμογή φίλτρων θόλωσης (blur), ή γεωμετρικές παραμορφώσεις (π.χ. elastic deformations).**
- **Χρωματικούς μετασχηματισμούς** σε εικόνες RGB (όπου εφαρμόζεται).

Οι τεχνικές αυτές έχουν καθιερωθεί ως βασικές στρατηγικές ενίσχυσης της απόδοσης σε προβλήματα υπολογιστικής όρασης, όπως επισημαίνεται και από τους Shorten και Khoshgoftaar (2019) [67], οι οποίοι παρουσιάζουν εκτενή επισκόπηση στις πιο διαδεδομένες μεθόδους data augmentation και τη θετική τους επίδραση στη γενίκευση μοντέλων CNN.

Διαχωρισμός Δεδομένων

Τα δεδομένα διαχωρίζονται συνήθως σε τρία σύνολα:

- **Εκπαίδευσης (training set):** χρησιμοποιείται για την εκπαίδευση του μοντέλου.
- **Επικύρωσης (validation set):** αξιολογεί την απόδοση κατά την εκπαίδευση και συμβάλλει στη ρύθμιση των υπερπαραμέτρων.
- **Δοκιμής (test set):** χρησιμοποιείται για την τελική αξιολόγηση μετά την ολοκλήρωση της εκπαίδευσης.

Ο σωστός διαχωρισμός είναι καίριας σημασίας για την αξιόπιστη εκτίμηση της απόδοσης του μοντέλου σε νέα δεδομένα.

3.2.3.2 Σχεδιασμός της Αρχιτεκτονικής

Ο σχεδιασμός της αρχιτεκτονικής αποτελεί ένα κρίσιμο στάδιο στην ανάπτυξη ενός μοντέλου βαθιάς μάθησης, καθώς επηρεάζει καθοριστικά την ικανότητά του να εξάγει χρήσιμες αναπαραστάσεις από τα δεδομένα και να επιτυγχάνει υψηλή ακρίβεια στο εκάστοτε πρόβλημα. Η διαδικασία περιλαμβάνει μια σειρά από επιλογές που σχετίζονται με τα επιμέρους επίπεδα του δικτύου, τις υπερπαραμέτρους, καθώς και τις μεθόδους εκπαίδευσης. Συγκεκριμένα:

Επίπεδα Συνέλιξης:

- **Αριθμός επιπέδων:** Περισσότερα συνελικτικά επίπεδα επιτρέπουν στο μοντέλο να εξάγει πιο σύνθετα και αφηρημένα χαρακτηριστικά από τα δεδομένα εισόδου.
- **Φίλτρα (kernels):** Ο αριθμός και το μέγεθος των φίλτρων σε κάθε επίπεδο επηρεάζουν την ποιότητα και την ποικιλία των χαρακτηριστικών που ανιχνεύονται.
- **Ολίσθηση (stride):** Καθορίζει το μέγεθος του βήματος με το οποίο το φίλτρο μετακινείται πάνω στα δεδομένα εισόδου. Επηρεάζει τόσο τη χωρική ανάλυση των εξαγόμενων χαρακτηριστικών όσο και τις τελικές διαστάσεις του χάρτη εξόδου.
- **Συμπλήρωση (padding):** Χρησιμοποιείται για να αποτραπεί η απώλεια πληροφορίας στα άκρα της εισόδου και να διατηρηθούν σταθερές οι χωρικές διαστάσεις των χαρακτηριστικών μετά τη συνέλιξη.

Επίπεδα Υποδειγματοληψίας:

- **Αριθμός επιπέδων:** Η υποδειγματοληψία μειώνει τις διαστάσεις των χαρτών χαρακτηριστικών, συμβάλλοντας στη μείωση της υπολογιστικής πολυπλοκότητας και στην αποφυγή υπερπροσαρμογής.
- **Τύπος:** Συνηθέστερες επιλογές είναι το **max pooling** και το **average pooling**, με το πρώτο να επιλέγει το μέγιστο και το δεύτερο τον μέσο όρο των τιμών σε κάθε περιοχή.
- **Μέγεθος φίλτρων και ολίσθηση:** Αντίστοιχα με τα συνελκτικά επίπεδα, οι διαστάσεις και η ολίσθηση του παραθύρου υποδειγματοληψίας καθορίζουν τον βαθμό μείωσης της χωρικής διάστασης και τη συμπίκνωση της πληροφορίας.

Πλήρως Συνδεδεμένα Επίπεδα:

- **Αριθμός επιπέδων και νευρώνων:** Αυτά τα επίπεδα συνδυάζουν τα εξαγόμενα χαρακτηριστικά σε μία τελική προβλεπτική απόφαση. Ο αριθμός των νευρώνων ρυθμίζει την υπολογιστική ισχύ και την εκφραστική ικανότητα του μοντέλου.

Συνάρτηση Σφάλματος:

Η επιλογή της κατάλληλης συνάρτησης σφάλματος (όπως MSE για παλινδρόμηση ή Cross Entropy για ταξινόμηση) είναι καθοριστική για τη σωστή ποσοτικοποίηση της απόδοσης του μοντέλου και κατευθύνει την εκπαιδευτική διαδικασία.

Αλγόριθμος Βελτιστοποίησης:

Ο βελτιστοποιητής ρυθμίζει τον τρόπο με τον οποίο προσαρμόζονται οι παράμετροι του δικτύου ώστε να ελαχιστοποιείται η συνάρτηση σφάλματος. Η επιλογή κατάλληλου αλγορίθμου (όπως SGD, BGD, κ.ά.) εξαρτάται από τη φύση του προβλήματος και τη δυναμική των δεδομένων.

3.2.3.3 Αξιολόγηση και Βελτιστοποίηση

Η προσεκτική αξιολόγηση της απόδοσης ενός συνελκτικού νευρωνικού δικτύου είναι απαραίτητη προκειμένου να εξασφαλιστεί ότι το μοντέλο όχι μόνο έχει μάθει αποτελεσματικά από τα δεδομένα εκπαίδευσης, αλλά μπορεί επίσης να γενικεύει σωστά σε μη-ορατά δεδομένα. Η διαδικασία αυτή πραγματοποιείται μέσω του συνόλου επαλήθευσης (validation set), στο οποίο το μοντέλο δεν έχει εκτεθεί κατά την εκπαίδευση.

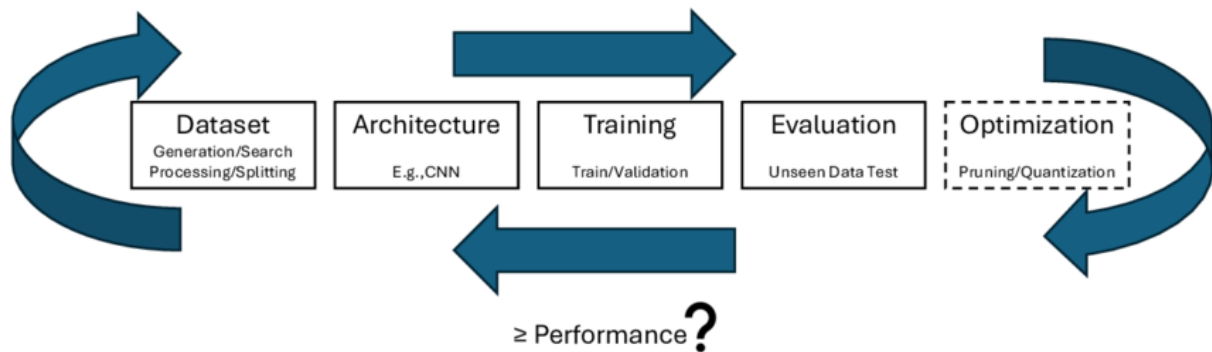
Μέσω της αξιολόγησης, μπορούν να εντοπιστούν φαινόμενα **υποπροσαρμογής** (underfitting) και **υπερπροσαρμογής** (overfitting). Στην πρώτη περίπτωση, το μοντέλο αποτυγχάνει να «κατανοήσει» τη δομή των δεδομένων, λόγω απλής αρχιτεκτονικής ή ανεπαρκούς εκπαίδευσης. Στη δεύτερη, προσαρμόζεται υπερβολικά στις ιδιαιτερότητες των δεδομένων εκπαίδευσης, εις βάρος της ικανότητάς του να γενικεύει σε νέα δείγματα.

Η βελτιστοποίηση επιτυγχάνεται μέσω στρατηγικών όπως:

- **Αλλαγή της αρχιτεκτονικής**, ώστε να ενισχυθεί ή να απλοποιηθεί το δίκτυο.
- **Ρύθμιση υπερπαραμέτρων**, όπως ο ρυθμός μάθησης, το πλήθος φίλτρων ή ο αριθμός των επιπέδων.
- **Εφαρμογή τεχνικών κανονικοποίησης**, όπως Dropout ή Weight Decay, που δρουν ως ρυθμιστές της πολυπλοκότητας του μοντέλου, περιορίζοντας την υπερπροσαρμογή.

- Χρήση τεχνικών **early stopping**, για τη διακοπή της εκπαίδευσης μόλις η απόδοση στο validation set παύσει να βελτιώνεται.

Η συνεχής παρακολούθηση των μετρικών απόδοσης και η ανατροφοδότηση στη διαδικασία εκπαίδευσης αποτελούν βασικές προϋποθέσεις για την επιτυχή εφαρμογή ενός μοντέλου σε πραγματικά προβλήματα.



Εικόνα 12: Τυπική ροή ανάπτυξης και αξιολόγησης ενός συστήματος μηχανικής μάθησης. Η διαδικασία περιλαμβάνει διαδοχικά στάδια επιλογής και προεπεξεργασίας του συνόλου δεδομένων, σχεδιασμού αρχιτεκτονικής (π.χ. CNN), εκπαίδευσης, αξιολόγησης και τελικής βελτιστοποίησης. Τα βέλη υποδεικνύουν τον επαναληπτικό χαρακτήρα της διαδικασίας, έως ότου επιτευχθεί η επιθυμητή απόδοση [68].

3.2.4 Μηχανισμοί Προσοχής και Αρχιτεκτονικές Τύπου Transformer

Οι μηχανισμοί προσοχής (attention mechanisms) αποτελούν μια από τις πιο ριζοσπαστικές καινοτομίες στην βαθιά μάθηση της τελευταίας δεκαετίας. Εισήχθησαν αρχικά στον τομέα της επεξεργασίας φυσικής γλώσσας, με στόχο τη μοντελοποίηση εξαρτήσεων μεταξύ στοιχείων ακολουθιών ανεξαρτήτως θέσης, σε αντίθεση με τις παραδοσιακές αναδρομικές αρχιτεκτονικές που έχουν εγγενή περιορισμό στο πώς διαχειρίζονται μεγάλες αποστάσεις. Η πιο emblematicή εφαρμογή τους ήταν το άρθρο “*Attention is All You Need*” των Vaswani et al. (2017), το οποίο εισήγαγε την αρχιτεκτονική Transformer και έθεσε τα θεμέλια για μια σειρά από μοντέλα αιχμής σε πολλαπλούς τομείς της τεχνητής νοημοσύνης [69].

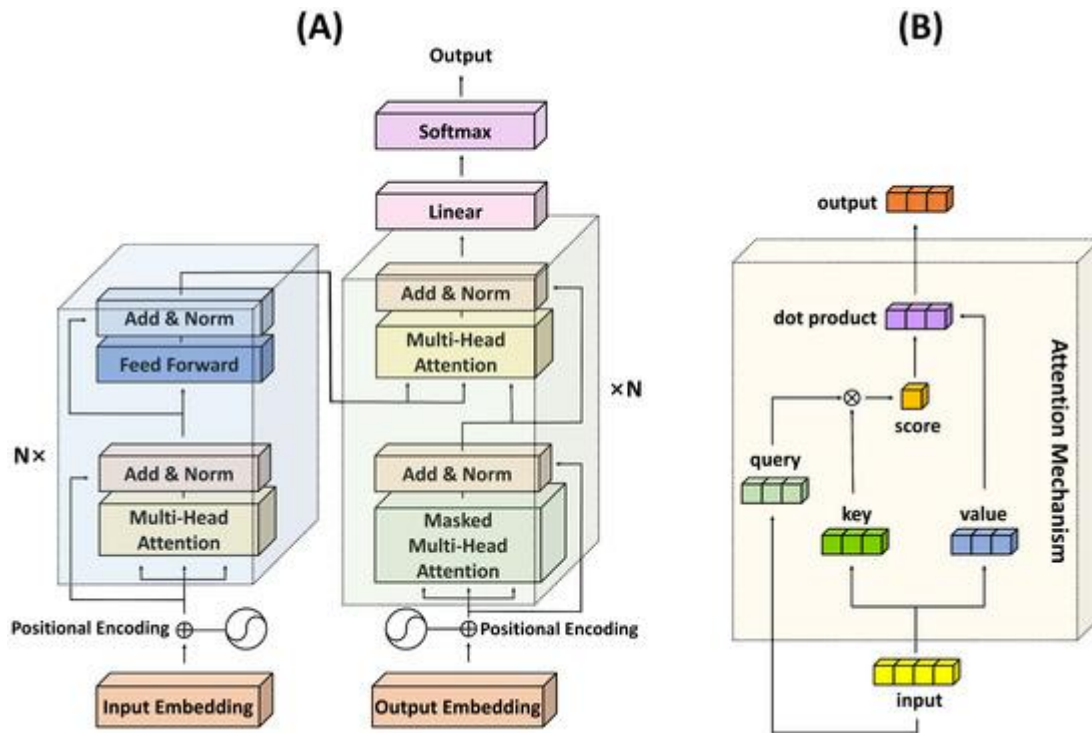
Η βασική ιδέα πίσω από τους μηχανισμούς προσοχής είναι ότι κάθε στοιχείο εισόδου μαθαίνει να «εστιάζει» επιλεκτικά στα πιο σχετικά στοιχεία, με έναν τρόπο που μαθαίνεται κατά τη διάρκεια της εκπαίδευσης. Ο υπολογισμός της προσοχής γίνεται με την παραγωγή τριών διανυσμάτων (queries, keys, values), από τα οποία προκύπτει ένας σταθμισμένος μέσος όρος των values με βάση τη συσχέτιση των queries με τα keys.

Η επιτυχία των μηχανισμών προσοχής στην επεξεργασία ακολουθιών οδήγησε στην υιοθέτησή τους και σε άλλους τομείς πέραν της φυσικής γλώσσας, όπως η **υπολογιστική όραση**. Σε αυτό το πλαίσιο, αναπτύχθηκαν παραλλαγές όπως οι **Vision Transformers (ViTs)** [70], που αντιμετωπίζουν τις εικόνες ως ακολουθίες patches αντί για πλέγματα χαρακτηριστικών που εξάγονται από συνελκτικά επίπεδα. Οι αρχιτεκτονικές ViT κατέδειξαν συγκρίσιμη ή και ανώτερη απόδοση σε προβλήματα ταξινόμησης εικόνων σε σύγκριση με τα κλασικά CNNs, με την προϋπόθεση επαρκών δεδομένων και υπολογιστικών πόρων.

Επιπλέον, έχουν προταθεί υβριδικά μοντέλα που συνδυάζουν τις δυνατότητες των συνελκτικών δικτύων με *attention modules*, όπως τα **Swin Transformers** και τα **U-Net με**

Attention Blocks, με στόχο τη διατήρηση της ιεραρχικής επεξεργασίας και την ενίσχυση της προσοχής σε κρίσιμες περιοχές των δεδομένων.

Η εισαγωγή *attention* μηχανισμών αποτελεί πλέον ένα θεμελιώδες εργαλείο στην ανάπτυξη μοντέλων για εφαρμογές που απαιτούν κατανόηση χωρικών ή εννοιολογικών συσχετίσεων, και συνεπώς υιοθετείται ευρέως και σε εργασίες ιατρικής απεικόνισης, όπως αυτές που εξετάζονται στην παρούσα εργασία.



Εικόνα 13: Αρχιτεκτονική του μοντέλου Transformer. Στο (A) παρουσιάζονται οι δομικές μονάδες του Encoder (αριστερά) και του Decoder (δεξιά), που αποτελούνται από στρώσεις πολλαπλών κεφαλών προσοχής (Multi-Head Attention), κανονικοποίησης (Add & Norm) και τροφοδότησης προώθησης (Feed Forward). Στο (B) απεικονίζεται ο μηχανισμός προσοχής (Attention Mechanism), όπου κάθε είσοδος μετατρέπεται σε τριάδες (query, key, value) και υπολογίζεται ένα σταθμισμένο άθροισμα βάσει της ομοιότητας μεταξύ query και key [71].

3.2.5 Αρχιτεκτονικές Δικτύων για επεξεργασία εικόνων

Τα νευρωνικά δίκτυα που εφαρμόζονται σε προβλήματα επεξεργασίας εικόνων έχουν εξελιχθεί ραγδαία, ξεκινώντας από καθαρά συνελκτικές αρχιτεκτονικές και φτάνοντας σε πιο πρόσφατες μορφές που ενσωματώνουν μηχανισμούς προσοχής και δομές τύπου *Transformer*. Αυτές οι αρχιτεκτονικές έχουν επιτρέψει σημαντικές βελτιώσεις στην ακρίβεια και τη γενίκευση των μοντέλων, ειδικά σε προκλήσεις όπως η ταξινόμηση εικόνων, η ανίχνευση αντικειμένων, και κυρίως η τμηματοποίηση ιατρικών εικόνων, που αποτελεί και το επίκεντρο της παρούσας εργασίας.

Η επιλογή αρχιτεκτονικής είναι στενά συνδεδεμένη με:

- Το είδος του προβλήματος (π.χ. ταξινόμηση, τμηματοποίηση),
- Το μέγεθος και τη φύση των δεδομένων, αλλά και

- **Τους διαθέσιμους υπολογιστικούς πόρους**, αφού μοντέλα όπως το Swin-UNETR απαιτούν σημαντικά περισσότερη μνήμη και ισχύ σε σύγκριση με πιο «ελαφριές» συνελκτικές δομές όπως το U-Net ή το SegResNet.

Στις επόμενες ενότητες παρουσιάζονται επιλεγμένες αρχιτεκτονικές που χρησιμοποιήθηκαν ή εξετάστηκαν στην παρούσα μελέτη, ταξινομημένες με βάση το είδος τους και την τεχνολογική τους βάση — είτε συνελκτική είτε βασισμένη σε μηχανισμούς προσοχής.

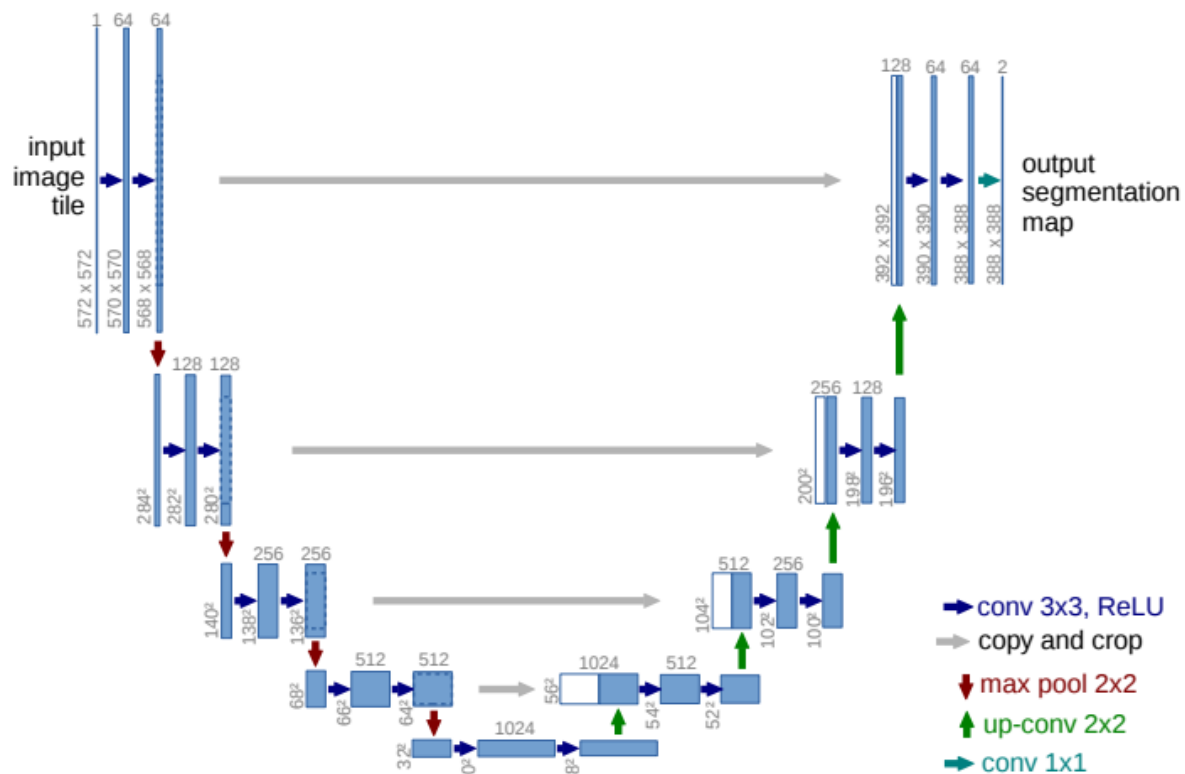
3.2.5.1 U-Net

Η αρχιτεκτονική **U-Net** παρουσιάστηκε για πρώτη φορά το 2015 από τους Ronneberger et al. και αποτελεί μια από τις πιο επιδραστικές προσεγγίσεις στην ιατρική απεικόνιση για προβλήματα τμηματοποίησης [72]. Σχεδιάστηκε ειδικά για τη βιοϊατρική τμηματοποίηση όπου συχνά διατίθενται περιορισμένα δεδομένα εκπαίδευσης. Η κύρια συνεισφορά του U-Net έγκειται στη συμμετρική του δομή τύπου "U", η οποία συνδυάζει ένα **συνελκτικό κατηφορικό σκέλος (encoder)** και ένα **ανοδικό σκέλος υπερδειγματοληψίας (decoder)**, ενισχυμένο από **skip connections**.

Η αρχιτεκτονική λειτουργεί ως εξής:

- Στον *encoder*, εφαρμόζονται διαδοχικά επίπεδα συνέλιξης και υποδειγματοληψίας, τα οποία μειώνουν τις χωρικές διαστάσεις και επιτρέπουν την εξαγωγή πλούσιων χαρακτηριστικών υψηλού επιπέδου από την είσοδο.
- Στο *decoder*, τα χαρακτηριστικά επανέρχονται σταδιακά στις αρχικές διαστάσεις μέσω επιπέδων υπερδειγματοληψίας (*upsampling*), με σκοπό την ακριβή χωρική ανακατασκευή.
- Τα *skip connections* μεταξύ συμμετρικών επιπέδων του *encoder* και του *decoder* μεταφέρουν λεπτομερείς χωρικές πληροφορίες, τις οποίες τα βαθύτερα επίπεδα ενδέχεται να έχουν απολέσει.

Αυτός ο σχεδιασμός καθιστά το U-Net ιδιαίτερα αποτελεσματικό στην ανακατασκευή μασκών τμηματοποίησης με ακρίβεια, ακόμη και με περιορισμένα δεδομένα, διατηρώντας τόσο την τοπική όσο και την παγκόσμια χωρική πληροφορία.



Εικόνα 14: Σχηματική απεικόνιση της αρχιτεκτονικής U-Net. Διακρίνονται τα στάδια κωδικοποίησης και αποκωδικοποίησης, καθώς και τα *skip connections* που συνδέουν συμμετρικά επίπεδα του *encoder* και του *decoder*. Απεικονίζονται επίσης τα βασικά βήματα επεξεργασίας: συνέλιξη, υποδειγματοληψία, υπερδειγματοληψία και τελική παραγωγή του χάρτη τμηματοποίησης [72].

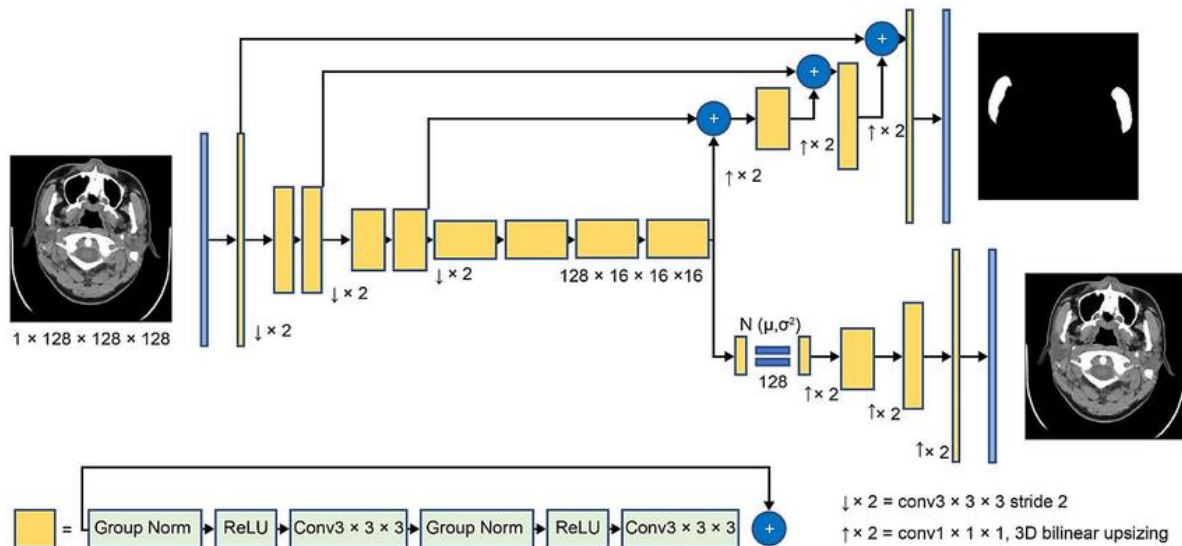
3.2.5.2 SegResNet

Το **SegResNet** αποτελεί μια αρχιτεκτονική συνελικτικού νευρωνικού δικτύου που βασίζεται στην ιδέα των **residual blocks** - δομών που επιτρέπουν την απευθείας μετάδοση της εισόδου στην έξοδο μέσω παράκαμψης ενός ή περισσότερων επιπέδων -όπως εισήχθησαν αρχικά στα ResNet [73], και προσαρμόστηκε για προβλήματα ιατρικής τμηματοποίησης εικόνας. Η συγκεκριμένη αρχιτεκτονική παρουσιάστηκε στο πλαίσιο του **Medical Open Network for AI (MONAI)**, μιας βιβλιοθήκης ανοικτού κώδικα για την ανάπτυξη και πειραματισμό με μοντέλα μηχανικής μάθησης στον τομέα της υγείας [74].

Η αρχιτεκτονική του SegResNet ακολουθεί την προσέγγιση **encoder-decoder** με την προσθήκη **residual connections** σε όλα τα επίπεδα, επιτρέποντας έτσι την εκπαίδευση βαθύτερων δικτύων χωρίς το πρόβλημα της εξαφάνισης ή εκρηκτικότητας των κλίσεων (*vanishing/exploding gradients*). Συγκεκριμένα:

- Ο *encoder* αποτελείται από διαδοχικά *residual blocks* που εφαρμόζουν συνέλιξη, κανονικοποίηση και μη γραμμικότητα.
- Ο *decoder* ακολουθεί αντίστροφη διαδικασία ανακατασκευής της ανάλυσης μέσω *transposed convolutions* και *skip connections*.
- Τα *residual connections* επιτρέπουν την απευθείας μετάδοση της πληροφορίας μεταξύ επιπέδων, ενισχύοντας τη σταθερότητα και την αποτελεσματικότητα της εκπαίδευσης.

Το SegResNet έχει αποδειχθεί ιδιαίτερα αποτελεσματικό σε διάφορα *benchmarks* ιατρικής τμηματοποίησης, όπως σε δεδομένα από MRI και CT, και παρέχεται ως προκαθορισμένο μοντέλο στη βιβλιοθήκη MONAI, καθιστώντας το ιδιαίτερα προσβάσιμο για ερευνητές και επαγγελματίες του χώρου.



Εικόνα 15: Διάγραμμα αρχιτεκτονικής του SegResNet για ιατρική τμηματοποίηση. Το δίκτυο ακολουθεί συμμετρική διάταξη τύπου *encoder-decoder*, ενσωματώνοντας *residual blocks* και *skip connections* για ενίσχυση της πληροφορίας και σταθερότητα κατά την εκπαίδευση. Η είσοδος υποβάλλεται σε επεξεργασία μέσω διαδοχικών συνελκτικών επιπέδων με *group normalization* και ενεργοποίηση *ReLU*, ενώ η έξοδος παράγεται μέσω σταδιακής ανακατασκευής με *upsampling* και συνδυασμό χαρακτηριστικών [75].

3.2.5.3 SwinUNETR

Το **SwinUNETR** αποτελεί ένα προηγμένο μοντέλο για την τρισδιάστατη τμηματοποίηση ιατρικών εικόνων, το οποίο συνδυάζει την κλασική αρχιτεκτονική **U-Net** με τον μηχανισμό **μετασχηματιστών με παράθυρο (Swin Transformers)**. Η αρχιτεκτονική προτάθηκε από τους Hatamizadeh et al. (2022) [76] στο πλαίσιο του Medical Open Network for AI (MONAI) και προσφέρει σημαντική βελτίωση στην ακρίβεια της τμηματοποίησης σε σχέση με παραδοσιακά CNN-based μοντέλα.

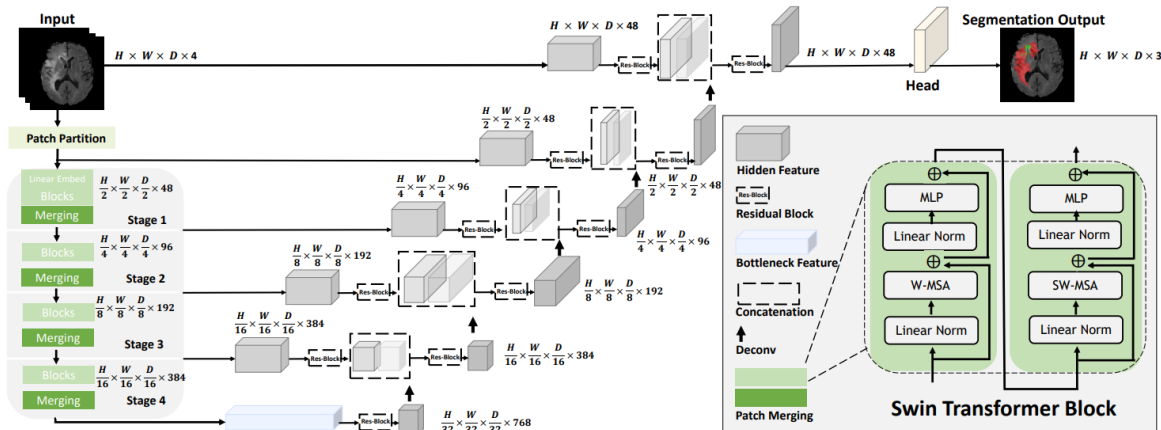
Το **SwinUNETR** χρησιμοποιεί έναν **3D Swin Transformer encoder**, ο οποίος εφαρμόζει *self-attention* σε μη-επικαλυπτόμενα τοπικά παράθυρα (windows) εντός του όγκου εισόδου. Αυτό επιτρέπει στο δίκτυο να εξάγει **ιεραρχικά χαρακτηριστικά** με τοπική και σταδιακά αυξανόμενη χωρική κατανόηση, διατηρώντας παράλληλα αποδοτική υπολογιστική πολυπλοκότητα.

Η δομή του ακολουθεί την τυπική λογική *encoder-decoder*:

- Ο *encoder* βασίζεται σε *Swin Transformer blocks* που παράγουν χαρακτηριστικά σε πολλαπλά επίπεδα ανάλυσης.

- Ο *decoder* συνθέτει την έξοδο μέσω *skip connections* που συνδέονται με τα αντίστοιχα στάδια του *encoder*, διατηρώντας έτσι την πληροφορία χαμηλού επιπέδου.
- Η τελική έξοδος παράγεται μέσω *3D upsampling*, ενοποίησης χαρακτηριστικών και ενός *output head*.

Το SwinUNETR έχει επιδείξει **state-of-the-art αποτελέσματα** σε διάφορα *benchmarks* τρισδιάστατης τμηματοποίησης όπως τα BraTS και MSD, συνδυάζοντας την **παγκόσμια κατανόηση** των μετασχηματιστών με την τοπική λεπτομέρεια που προσφέρουν συνδέσεις τύπου CNN (*skip connections*).



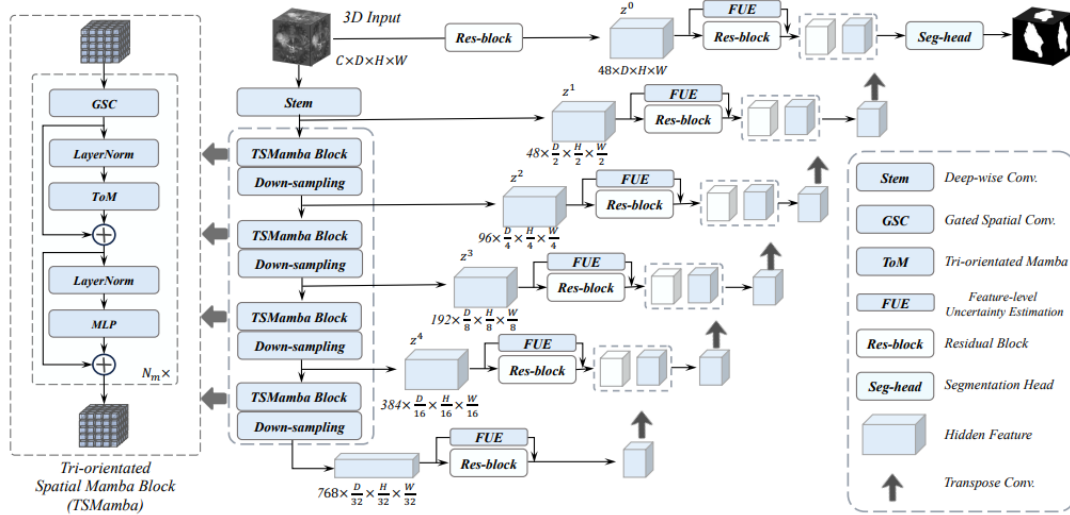
Εικόνα 16: Δομική απεικόνιση της αρχιτεκτονικής SwinUNETR. Το μοντέλο ακολουθεί διάταξη τύπου *encoder-decoder*, με *Swin Transformer blocks* να εφαρμόζονται στο *encoding path* για την εξαγωγή χαρακτηριστικών, ενώ τα *skip connections* και το *decoding path* οδηγούν στην τελική έξοδο τμηματοποίησης [76].

3.2.5.4 SegMamba

Το **SegMamba** είναι μια πρόσφατη αρχιτεκτονική τμηματοποίησης που ανήκει στην ευρύτερη κατηγορία των **state space models (SSMs)** και αξιοποιεί τον μηχανισμό **Mamba**, ο οποίος προτάθηκε ως εναλλακτική των μετασχηματιστών (Transformers) για εφαρμογές σε ακολουθιακά δεδομένα [77]. Τα **State Space Models** είναι μαθηματικά μοντέλα που περιγράφουν δυναμικά συστήματα με βάση την εσωτερική τους κατάσταση, και όταν ενσωματώνονται σε *deep learning* αρχιτεκτονικές, έχουν τη δυνατότητα να μαθαίνουν **μακροπρόθεσμες εξαρτήσεις** σε **γραμμική υπολογιστική πολυπλοκότητα**. Ο μηχανισμός **Mamba** επεκτείνει αυτή την προσέγγιση με τη χρήση *gating*, δηλαδή μηχανισμών ελέγχου ροής πληροφορίας, βελτιώνοντας σημαντικά την εκφραστικότητα του μοντέλου. Το SegMamba εφαρμόζει αυτή τη λογική στον τομέα της ιατρικής τμηματοποίησης, προτείνοντας ένα **υπολογιστικά ελαφρύ και αποδοτικό μοντέλο** για δισδιάστατα και τρισδιάστατα δεδομένα.

Ο πυρήνας του SegMamba είναι το **Mamba block**, το οποίο βασίζεται στη μοντελοποίηση γραμμικών και μη γραμμικών σχέσεων σε μεγάλα χρονικά (ή χωρικά) διαστήματα, μέσω επιλεκτικών **SSMs** με *gating*, αντί του *self-attention*. Αυτό επιτρέπει στο μοντέλο να έχει **χρονική/χωρική μνήμη** σε γραμμική υπολογιστική πολυπλοκότητα και να διατηρεί **long-range dependencies**, παρόμοια με τους μετασχηματιστές, αλλά με καλύτερη αποδοτικότητα [78].

Η γενική δομή του **SegMamba** βασίζεται σε **αρχιτεκτονική τύπου U-Net**, όπου ο *encoder* περιλαμβάνει **Mamba blocks** και ο *decoder* ακολουθεί παραδοσιακή προσέγγιση *upsampling* με *skip connections*. Το μοντέλο έχει αποδειχθεί ιδιαίτερα αποτελεσματικό σε πολλά *benchmarks* τμηματοποίησης, προσφέροντας **συγκρίσιμη ή ανώτερη ακρίβεια** σε σχέση με πιο βαριά *transformer-based* μοντέλα όπως το SwinUNETR, ενώ διατηρεί **χαμηλότερες απαιτήσεις σε πόρους και χαμηλότερο latency**.



Εικόνα 17: Σχηματική αναπαράσταση της αρχιτεκτονικής SegMamba. Ο *encoder* αποτελείται από TSMamba blocks και μονάδες υποδειγματοληψίας, ενώ ο *decoder* εφαρμόζει *transposed convolutions* και *skip connections*. Περιλαμβάνονται επίσης *residual blocks* και μηχανισμοί εκτίμησης αβεβαιότητας για ενίσχυση της εκφραστικότητας και της σταθερότητας του μοντέλου [77].

Κεφάλαιο 4: Υλοποίηση και Πειραματική Αξιολόγηση

4.1 Εισαγωγή

Στο παρόν κεφάλαιο περιγράφεται η πειραματική διαδικασία που ακολουθήθηκε για την εκπαίδευση και αξιολόγηση διαφορετικών μοντέλων βαθιάς μάθησης για το πρόβλημα της τρισδιάστατης τμηματοποίησης εγκεφάλου. Πιο συγκεκριμένα, αξιοποιήθηκε το διαθέσιμο σύνολο δεδομένων με ελεγχόμενη πρόσβαση **Meningioma-SEG-CLASS** [4], το οποίο περιλαμβάνει εικόνες μαγνητικής τομογραφίας εγκεφάλου με χειροκίνητες επισημάνσεις (ground truth annotations) μηνιγγιώματος.

Η ανάλυση επικεντρώνεται στη συγκριτική αξιολόγηση διαφορετικών αρχιτεκτονικών:

- παραδοσιακών συνελκτικών μοντέλων (όπως **U-Net** και **SegResNet**),
- αρχιτεκτονικών βασισμένων σε μηχανισμούς προσοχής (όπως **SwinUNETR**),
- καθώς και πιο πρόσφατων υβριδικών προσεγγίσεων (όπως το **SegMamba**).

Στόχος της παρούσας μελέτης είναι η εμπειριστατωμένη αξιολόγηση της ακρίβειας και της γενικευσιμότητας σύγχρονων αρχιτεκτονικών βαθιάς μάθησης για την τρισδιάστατη τμηματοποίηση μηνιγγιώματος, μέσα από τον σχεδιασμό και την υλοποίηση ενός ολοκληρωμένου πειραματικού pipeline.

Η ανάλυση καλύπτει όχι μόνο τη συγκριτική αξιολόγηση διαφορετικών μοντέλων (U-Net, SegResNet, SwinUNETR, SegMamba), αλλά και τη **συστηματική διερεύνηση κρίσιμων παραμέτρων**, όπως η επιλογή συνάρτησης κόστους, optimizer, learning rate scheduler και τεχνικών *data augmentation*.

Επιπλέον, δίνεται έμφαση στη **διεξοδική ανάλυση των πειραματικών αποτελεσμάτων**, η οποία οδήγησε σε **επαναληπτικές βελτιώσεις και τροποποιήσεις του πειραματικού pipeline**. Η αξιολόγηση βασίζεται κυρίως στον **Dice Similarity Coefficient** και αποσκοπεί στην τεκμηριωμένη ανάδειξη των δυνατοτήτων και περιορισμών κάθε προσέγγισης σε πραγματικά ιατρικά δεδομένα.

4.2 Περιγραφή Συνόλου Δεδομένων

Για την πειραματική αξιολόγηση των αρχιτεκτονικών τρισδιάστατης τμηματοποίησης εγκεφαλικών όγκων αξιοποιήθηκε το δημόσια διαθέσιμο σύνολο δεδομένων **Meningioma-SEG-CLASS** [4]. Πρόκειται για ένα πολυκεντρικό και προτυποποιημένο *dataset* που δημιουργήθηκε ειδικά για την υποστήριξη αλγορίθμων προεγχειρητικής διάγνωσης και επεξεργασίας μηνιγγιωμάτων. Το σύνολο δεδομένων περιλαμβάνει τρισδιάστατες εικόνες μαγνητικής τομογραφίας (MRI) του εγκεφάλου, συνοδευόμενες από χειροκίνητα δημιουργημένες μάσκες που καθορίζουν με ακρίβεια τη χωρική εντόπιση των όγκων.

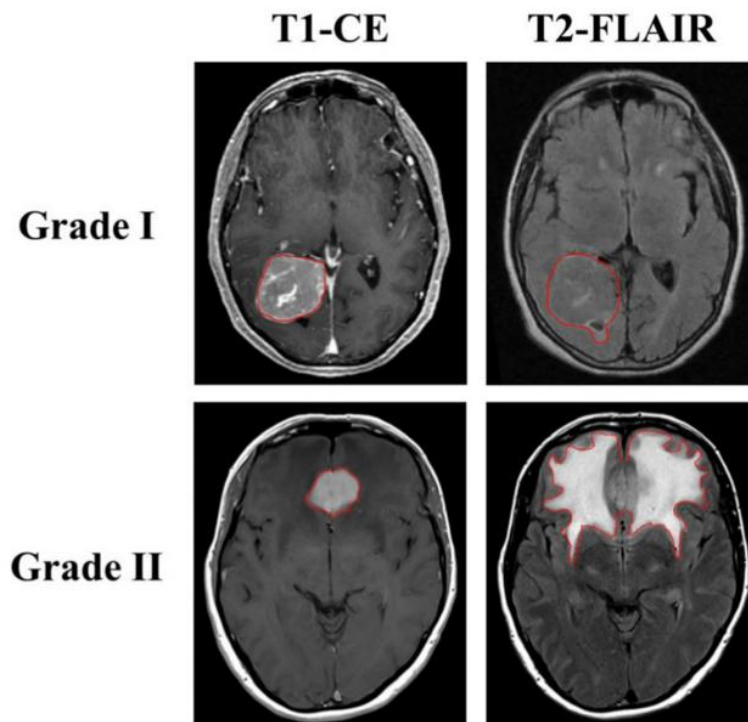
Στο πλαίσιο της παρούσας εργασίας, χρησιμοποιήθηκαν αποκλειστικά οι εικόνες **T1-weighted με σκιαγραφικό** (T1-CE - T1-weighted Contrast Enhanced). Αυτές οι εικόνες λαμβάνονται μετά από χορήγηση σκιαγραφικής ουσίας που συσσωρεύεται στις περιοχές όπου ο

αιματοεγκεφαλικός φραγμός είναι διαταραγμένος, όπως συμβαίνει στις περιπτώσεις όγκων. Η τεχνική αυτή παρέχει βελτιωμένη αντίθεση μεταξύ του όγκου και των παρακείμενων υγιών ιστών, διευκολύνοντας σημαντικά την ακριβή τμηματοποίηση.

Από τις συνολικά 96 περιπτώσεις του *dataset*, επιλέχθηκαν και αξιοποιήθηκαν **76 πλήρως διαθέσιμες περιπτώσεις** που ήταν κατάλληλες για την εκπαίδευση μοντέλων τμηματοποίησης. Σημειώνεται ότι δεν χρησιμοποιήθηκαν οι πληροφορίες κατηγοριοποίησης του μηνιγγιώματος, καθώς η παρούσα μελέτη επικεντρώνεται αποκλειστικά στην εργασία της τρισδιάστατης τμηματοποίησης του όγκου.

Η μορφή αποθήκευσης των δεδομένων (.nii.gz) εξασφαλίζει τη συμβατότητα με ειδικές ιατρικές βιβλιοθήκες όπως το **MONAI**, το οποίο αξιοποιήθηκε για την υλοποίηση της πειραματικής διαδικασίας. Η κατανομή των 76 δειγμάτων έγινε με αναλογία **80%-20%** σε εκπαιδευτικό και επικυρωτικό σύνολο αντίστοιχα (61 και 15 περιπτώσεις), εξασφαλίζοντας τη συγκρίσιμη αξιολόγηση της απόδοσης των διαφορετικών μοντέλων.

Τέλος, μετά από ανάλυση του *dataset* και αρχικούς πειραματισμούς, αποφασίστηκε η εφαρμογή συγκεκριμένων τεχνικών προεπεξεργασίας εικόνας. Αυτές περιλαμβάνουν προσαρμογή του *voxel spacing*, μετασχηματισμούς των τιμών έντασης των *voxels* και κανονικοποίηση του μεγέθους των δειγμάτων. Οι τεχνικές αυτές εφαρμόστηκαν με στόχο τη διασφάλιση της ομοιομορφίας και της αποδοτικότητας της εκπαίδευσης, λαμβάνοντας υπόψη τους περιορισμούς των διαθέσιμων υπολογιστικών πόρων.



Εικόνα 18: Ενδεικτικές εικόνες MRI από το σύνολο δεδομένων *Meningioma-SEG-CLASS*. Παρουσιάζονται περιπτώσεις Grade I και Grade II μηνιγγιωμάτων, με τις περιοχές του όγκου

επισημασμένες με κόκκινο περίγραμμα, τόσο σε ακολουθίες T1-CE όσο και σε T2-FLAIR (μόνο για λόγους οπτικής απεικόνισης· η παρούσα μελέτη χρησιμοποιεί αποκλειστικά τις T1-CE) [4].

4.3 Επαύξηση Δεδομένων και Μετασχηματισμοί

Η προεπεξεργασία και η επαύξηση των δεδομένων συνιστά καθοριστικό παράγοντα για την επιτυχή λειτουργία ενός συστήματος τμηματοποίησης ιατρικών εικόνων. Στο πλαίσιο της παρούσας εργασίας, εφαρμόστηκε ένα ολοκληρωμένο σύνολο μετασχηματισμών με διττό στόχο: την **εξομάλυνση των δεδομένων εισόδου** και την **τεχνητή επέκταση του συνόλου εκπαίδευσης**. Με τον τρόπο αυτό επιδιώχθηκε η ενίσχυση της ικανότητας γενίκευσης των μοντέλων και η μείωση του φαινομένου της υπερπροσαρμογής.

Αρχικά, όλες οι εικόνες και οι αντίστοιχες μάσκες **μετατράπηκαν σε ενιαίο προσανατολισμό (RAS)**. Στις περιπτώσεις όπου απαιτήθηκε περαιτέρω σταθεροποίηση των διαστάσεων, εφαρμόστηκαν είτε τεχνικές **επαναδειγματοληψίας με συμπλήρωση ή περικοπή** είτε μέθοδοι **τυχαίας δειγματοληψίας**. Ιδιαίτερα σημαντικό βήμα στην προεπεξεργασία των εικόνων είναι η **κανονικοποίηση των εντάσεων (NormalizeIntensity)**, η οποία επιτρέπει στα μοντέλα να εστιάζουν στις μορφολογικές διαφοροποιήσεις.

Για την αντιμετώπιση του προβλήματος της σπανιότητας των θετικών δειγμάτων (δηλαδή των περιοχών που περιέχουν όγκο) και των περιορισμένων υπολογιστικών πόρων, εφαρμόστηκε ο μετασχηματισμός **RandCropByPosNegLabeld**. Αυτός επιλέγει τυχαία υποπεριοχές της εικόνας με ελεγχόμενη αναλογία μεταξύ θετικών και αρνητικών περιπτώσεων. Η συγκεκριμένη στρατηγική εξασφαλίζει ότι κάθε ομάδα εκπαίδευσης περιλαμβάνει αντιπροσωπευτικές περιοχές του όγκου, επιταχύνοντας έτσι τη σύγκλιση και βελτιώνοντας την ακρίβεια του μοντέλου.

Για την τεχνητή αύξηση της ποικιλομορφίας των δεδομένων, εφαρμόστηκαν οι ακόλουθοι στοχαστικοί χωρικοί μετασχηματισμοί:

Τυχαίες αναστροφές (RandFlipd) κατά μήκος των αξόνων x, y και z, με πιθανότητα εμφάνισης 50%. Αυτές ενισχύουν την ανθεκτικότητα του μοντέλου σε συμμετρικές μεταβολές.

Τυχαίοι γραμμικοί μετασχηματισμοί (RandAffined) που περιλαμβάνουν μικρές περιστροφές της τάξης των $\pm 5^\circ$ και κλιμακώσεις $\pm 5\%$. Οι μετασχηματισμοί αυτοί προσδίδουν στο μοντέλο ανθεκτικότητα έναντι παραμορφώσεων ή διαφορετικών συνθηκών λήψης των μαγνητικών τομογραφιών.

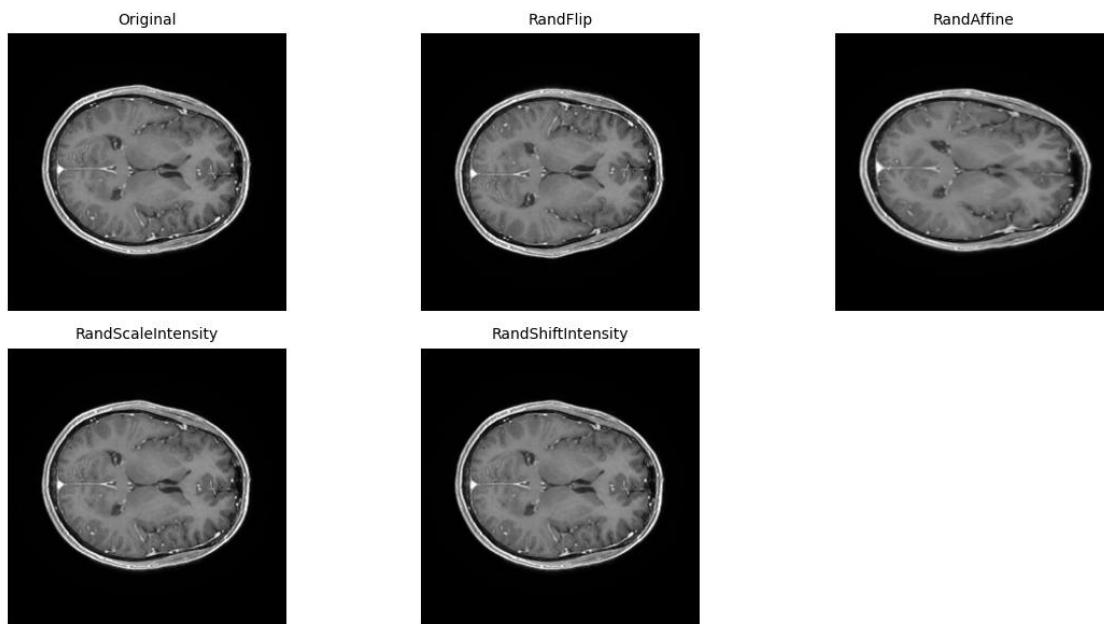
Σε ορισμένες παραλλαγές των πειραμάτων, αξιοποιήθηκαν επιπρόσθετοι μετασχηματισμοί έντασης:

Στοχαστικές μετατοπίσεις ή κλιμακώσεις έντασης (RandShiftIntensityd, RandScaleIntensityd) που χρησιμοποιήθηκαν πειραματικά με σκοπό την περαιτέρω αύξηση της ποικιλοότητας των τιμών έντασης.

Για το σύνολο επικύρωσης, εφαρμόστηκαν αποκλειστικά **μη στοχαστικοί μετασχηματισμοί**: σταθερός προσανατολισμός, αποστάσεις, περικοπή ή συμπλήρωση σε προκαθορισμένες διαστάσεις, καθώς και κανονικοποίηση της έντασης. Η επιλογή αυτή διασφαλίζει τη συνοχή

και τη συγκρισιμότητα των προβλέψεων ανεξάρτητα από την αρχιτεκτονική που χρησιμοποιείται.

Οι παραπάνω τεχνικές βασίζονται σε **καθιερωμένες βέλτιστες πρακτικές** που έχουν αποδειχθεί αποτελεσματικές στη διεθνή βιβλιογραφία. Η αξιοποίηση της βιβλιοθήκης **MONAI** καθιστά τη διαχείριση και τη σύνθεση των προαναφερθέντων μετασχηματισμών ιδιαίτερα ευέλικτη και αποδοτική, διευκολύνοντας την πειραματική τροποποίηση και τον συνδυασμό τους.



Εικόνα 19: Οπτική απεικόνιση επιλεγμένων στοχαστικών μετασχηματισμών που εφαρμόστηκαν σε εικόνες MRI κατά την εκπαίδευση. Από αριστερά προς τα δεξιά και πάνω προς τα κάτω: αρχική εικόνα, RandFlip, RandAffine, RandScaleIntensity και RandShiftIntensity. Οι μετασχηματισμοί υλοποιήθηκαν μέσω της βιβλιοθήκης MONAI.

4.4 Εκπαίδευση και Βασικές Υπερπαράμετροι

Η διαδικασία εκπαίδευσης των μοντέλων σχεδιάστηκε με στόχο τη μεγιστοποίηση της ικανότητας γενίκευσης σε άγνωστα δεδομένα και την αποτελεσματική σύγκριση διαφορετικών αρχιτεκτονικών στο πεδίο της τρισδιάστατης τμηματοποίησης μηνιγγιωμάτων. Για την επίτευξη αυτού του στόχου, εφαρμόστηκε ένα συνεκτικό πλαίσιο τεχνικών επιλογών αναφορικά με τις συναρτήσεις απώλειας, τους αλγορίθμους βελτιστοποίησης, τις μετρικές αξιολόγησης και τις στρατηγικές εκπαίδευσης.

4.4.1 Συναρτήσεις Απώλειας

Κατά την εκπαίδευση των μοντέλων εξετάστηκαν διαφορετικές συναρτήσεις απώλειας, καθεμία με τα δικά της χαρακτηριστικά και πλεονεκτήματα:

Dice Loss: Ιδιαίτερα κατάλληλη για προβλήματα τμηματοποίησης με ανισόρροπες κλάσεις, καθώς εστιάζει στην επικάλυψη μεταξύ της προβλεπόμενης και της πραγματικής μάσκας. Η συνάρτηση αυτή είναι ανεξάρτητη από την κατανομή των κλάσεων και ιδιαίτερα αποτελεσματική όταν οι θετικές περιοχές αποτελούν μικρό ποσοστό της συνολικής εικόνας.

DiceCE Loss: Συνδυάζει το Dice Loss με το Cross-Entropy Loss, αξιοποιώντας τόσο τη γεωμετρική επικάλυψη όσο και τη σημειακή ακρίβεια της πρόβλεψης. Ο συνδυασμός αυτός παρέχει πιο σταθερή και γρήγορη σύγκλιση, ενώ διατηρεί την ευαισθησία στις μικρές δομές.

DiceFocal Loss: Επεκτείνει το Dice Loss ενσωματώνοντας το μηχανισμό εστίασης της Focal Loss. Η Focal Loss αποτελεί μια παραλλαγή του Cross-Entropy Loss που εισάγει έναν παράγοντα εστίασης για τη μείωση της σχετικής απώλειας των εύκολα ταξινομημένων δειγμάτων. Με τον τρόπο αυτό, το DiceFocal Loss μειώνει την επίδραση των εύκολων δειγμάτων και ενισχύει την εκμάθηση από τις δυσκολότερες περιοχές, βελτιώνοντας την απόδοση σε περιπτώσεις με σημαντική ανισορροπία κλάσεων.

Η επιλογή της κατάλληλης συνάρτησης απώλειας πραγματοποιήθηκε με βάση τη δομή και τα χαρακτηριστικά του εκάστοτε μοντέλου, καθώς και την επιθυμητή συμπεριφορά κατά την εκπαίδευση.

4.4.2 Βελτιστοποιητές και Προγραμματιστές Μάθησης (Schedulers)

Για την ελαχιστοποίηση της συνάρτησης απώλειας εφαρμόστηκαν προηγμένοι αλγόριθμοι βελτιστοποίησης που έχουν αποδειχθεί αποτελεσματικοί στη βαθιά εκμάθηση:

Adam: Αποτελεί έναν ευρέως διαδεδομένο αλγόριθμο βελτιστοποίησης που συνδυάζει τα πλεονεκτήματα των μεθόδων momentum και προσαρμοστικού ρυθμού εκμάθησης. Ο Adam διατηρεί κινητούς μέσους όρους τόσο των κλίσεων όσο και των τετραγώνων τους, επιτυγχάνοντας γρήγορη σύγκλιση και αποτελεσματική διαχείριση των παραμέτρων. [59].

AdamW: Παραλλαγή του Adam που ενσωματώνει βελτιωμένη κανονικοποίηση βάρους (decoupled weight decay). Ο AdamW διαχωρίζει την κανονικοποίηση βάρους από τον υπολογισμό των κλίσεων, οδηγώντας σε καλύτερη γενίκευση και έχει αποδειχθεί ιδιαίτερα αποτελεσματικός σε σύγχρονες αρχιτεκτονικές νευρωνικών δικτύων [79].

Για τη δυναμική προσαρμογή του ρυθμού εκμάθησης εφαρμόστηκαν δύο διαφορετικές στρατηγικές προγραμματισμού:

ReduceLROnPlateau: Μειώνει τον ρυθμό εκμάθησης όταν η απόδοση στο σύνολο επικύρωσης παύει να βελτιώνεται για έναν προκαθορισμένο αριθμό εποχών. Η προσέγγιση αυτή είναι προσαρμοστική και βασίζεται στην πραγματική απόδοση του μοντέλου, επιτρέποντας λεπτότερη ρύθμιση κατά τα τελικά στάδια της εκπαίδευσης.

CosineAnnealing: Ακολουθεί μια συνημιτονοειδή μείωση του ρυθμού εκμάθησης με δυνατότητα επαναλήψεων (restarts). Η μέθοδος αυτή ενισχύει την εξερεύνηση του χώρου παραμέτρων μέσω περιοδικών αυξήσεων του ρυθμού εκμάθησης, αποφεύγοντας την παγίδευση σε τοπικά ελάχιστα και βελτιώνοντας την ικανότητα εύρεσης καλύτερων λύσεων.

4.4.3 Υπερπαραμέτροι Εκπαίδευσης

Λόγω της πειραματικής φύσης της εργασίας και των περιορισμένων υπολογιστικών πόρων, η επιλογή των υπερπαραμέτρων έγινε στρατηγικά για τη μεγιστοποίηση της αποδοτικότητας και της ποιότητας των αποτελεσμάτων.

Αρχιτεκτονικές Παράμετροι: Για την εκμετάλλευση στο έπακρο των διαθέσιμων υπολογιστικών πόρων, δοκιμάστηκαν διαφορετικές διαστάσεις μοντέλων. Συγκεκριμένα, εξετάστηκαν διάφορα βάθη δικτύων, αριθμοί φίλτρων ανά επίπεδο, και παράμετροι εξειδικευμένων αρχιτεκτονικών όπως ο αριθμός κεφαλών προσοχής (attention heads) στα Transformer-based μοντέλα. Η προσέγγιση αυτή επέτρεψε τον εντοπισμό της βέλτιστης ισορροπίας μεταξύ πολυπλοκότητας μοντέλου και υπολογιστικής απαίτησης.

Εποχές Εκπαίδευσης: Δοκιμάστηκαν διάφοροι αριθμοί εποχών με στόχο την εξισορρόπηση μεταξύ ικανότητας εκμάθησης και υπολογιστικής αποδοτικότητας. Η διάρκεια της εκπαίδευσης προσαρμόστηκε ανάλογα με τη συμπεριφορά σύγκλισης του εκάστοτε μοντέλου και τη διαθεσιμότητα των πόρων.

Μέγεθος Παρτίδας: Το *batch size* καθορίστηκε με βάση τους περιορισμούς της διαθέσιμης μνήμης GPU, διατηρώντας παράλληλα επαρκή στατιστική κάλυψη παραδειγμάτων ανά ενημέρωση των βαρών. Η επιλογή αυτή εξασφάλισε σταθερή σύγκλιση και αποτελεσματική εκμετάλλευση των υπολογιστικών πόρων.

Μετρικές Αξιολόγησης και Παρακολούθηση: Η κύρια μετρική αξιολόγησης ήταν ο *Validation Dice Score*, η οποία αποτυπώνει άμεσα την ποιότητα της παραγόμενης τμηματοποίησης. Κατά την εκπαίδευση, καταγράφονταν και απεικονίζονταν σε πραγματικό χρόνο οι τιμές των συναρτήσεων απώλειας εκπαίδευσης και επικύρωσης ανά εποχή, παρέχοντας άμεση ανατροφοδότηση για την πρόοδο της εκμάθησης.

Οπτική Επιθεώρηση: Πραγματοποιήθηκε συστηματική οπτική επιθεώρηση των προβλεπόμενων μασκών μέσω τυχαίων τομών από τους τρεις ανατομικούς άξονες, με παράλληλη σύγκριση προς τις πραγματικές ετικέτες. Η πρακτική αυτή παρείχε πολύτιμες ποιοτικές πληροφορίες για τη συμπεριφορά των μοντέλων, επιτρέποντας τον εντοπισμό τόσο των επιτυχημένων όσο και των προβληματικών περιπτώσεων, καθώς και την κατανόηση των χαρακτηριστικών λαθών κάθε αρχιτεκτονικής.

4.5 Αποτελέσματα

4.5.1 Βελτιστοποίηση Υπερπαραμέτρων

Πριν από τη σύγκριση των διαφορετικών αρχιτεκτονικών, πραγματοποιήθηκε εκτενής αξιολόγηση των κύριων υπερπαραμέτρων του συστήματος με στόχο τον προσδιορισμό της βέλτιστης πειραματικής διάταξης. Η διαδικασία αυτή περιλάμβανε τη συστηματική εξέταση διαφορετικών συναρτήσεων απώλειας, *schedulers* και άλλων κρίσιμων παραμέτρων.

4.5.1.1 Επιλογή Συνάρτησης Απώλειας

Αρχικά εξετάστηκε η επίδραση διαφορετικών συναρτήσεων απώλειας στην απόδοση των μοντέλων. Για την αξιολόγηση αυτή χρησιμοποιήθηκε το U-Net ως baseline αρχιτεκτονική, διατηρώντας σταθερές όλες τις υπόλοιπες υπερπαραμέτρους.

	Dice-CrossEntropy Loss	Dice Loss	DiceFocal Loss
Dice Score	0.8092	0.7554	0.5670

Πίνακας 1: Συγκριτική αξιολόγηση συναρτήσεων απώλειας (U-Net baseline)

Από τα αποτελέσματα προκύπτει ότι η συνδυασμένη συνάρτηση απώλειας Dice-CrossEntropy με αναλογία 0.7:0.3 παρουσιάζει τη βέλτιστη απόδοση. Η επιλογή αυτή αιτιολογείται από το γεγονός ότι το Dice loss αντιμετωπίζει αποτελεσματικά το πρόβλημα της ανισορροπίας κλάσεων που χαρακτηρίζει τα ιατρικά δεδομένα, ενώ το CrossEntropy loss συνεισφέρει στη σταθεροποίηση της εκπαίδευσης και τη βελτίωση της ακρίβειας σε επίπεδο *voxel*.

4.5.1.2 Βελτιστοποίηση Learning Rate Scheduler

Η αξιολόγηση των scheduler algorithms έδειξε διαφορετικά πρότυπα ανάλογα με τη διάρκεια της εκπαίδευσης:

- **Για μακροχρόνια εκπαίδευση (>50 epochs):** Ο Cosine Annealing scheduler παρουσίασε ανώτερη απόδοση, επιτρέποντας στο μοντέλο να εξερευνήσει αποτελεσματικότερα τον χώρο των παραμέτρων μέσω της ομαλής μείωσης του learning rate.
- **Για βραχυχρόνια εκπαίδευση (<50 epochs):** Ο ReduceLROnPlateau scheduler αποδείχθηκε καταλληλότερος, παρέχοντας γρήγορη σύγκλιση όταν οι υπολογιστικοί περιορισμοί απαιτούν συντομότερη εκπαίδευση.

4.5.1.3 Τελική Επιλογή Υπερπαραμέτρων

Βάσει των πειραματικών αποτελεσμάτων, το βέλτιστο σύνολο υπερπαραμέτρων που υιοθετήθηκε για όλες τις επόμενες συγκρίσεις ήταν:

- **Loss Function:** Dice-CrossEntropy (αναλογία 0.7:0.3)
- **Optimizer:** AdamW με αρχικό *learning rate* $1e-3$ και τελικό $1e-6$
- **Scheduler:** Cosine Annealing για εκπαιδεύσεις 50-100 epochs
- **Batch Size:** 2 (λόγω περιορισμών μνήμης GPU) (1 για SwinUNETR)
- **Weight Decay:** $5e-6$

4.5.2 Συγκριτική Αξιολόγηση Αρχιτεκτονικών

Αξιοποιώντας τη βελτιστοποιημένη ρύθμιση που καθορίστηκε στην προηγούμενη ενότητα, πραγματοποιήθηκε συγκριτική αξιολόγηση τεσσάρων διαφορετικών αρχιτεκτονικών βαθιάς μάθησης για το πρόβλημα της τμηματοποίησης. Οι αρχιτεκτονικές που εξετάστηκαν περιλαμβάνουν:

- **U-Net:** Η κλασική αρχιτεκτονική *encoder-decoder* για ιατρική τμηματοποίηση
- **SegResNet:** Βελτιωμένη εκδοχή που ενσωματώνει *residual connections*
- **SwinUNETR:** *Transformer-based* αρχιτεκτονική που αξιοποιεί το Swin Transformer
- **SegMamba:** *State-of-the-art* αρχιτεκτονική βασισμένη στο Mamba architecture

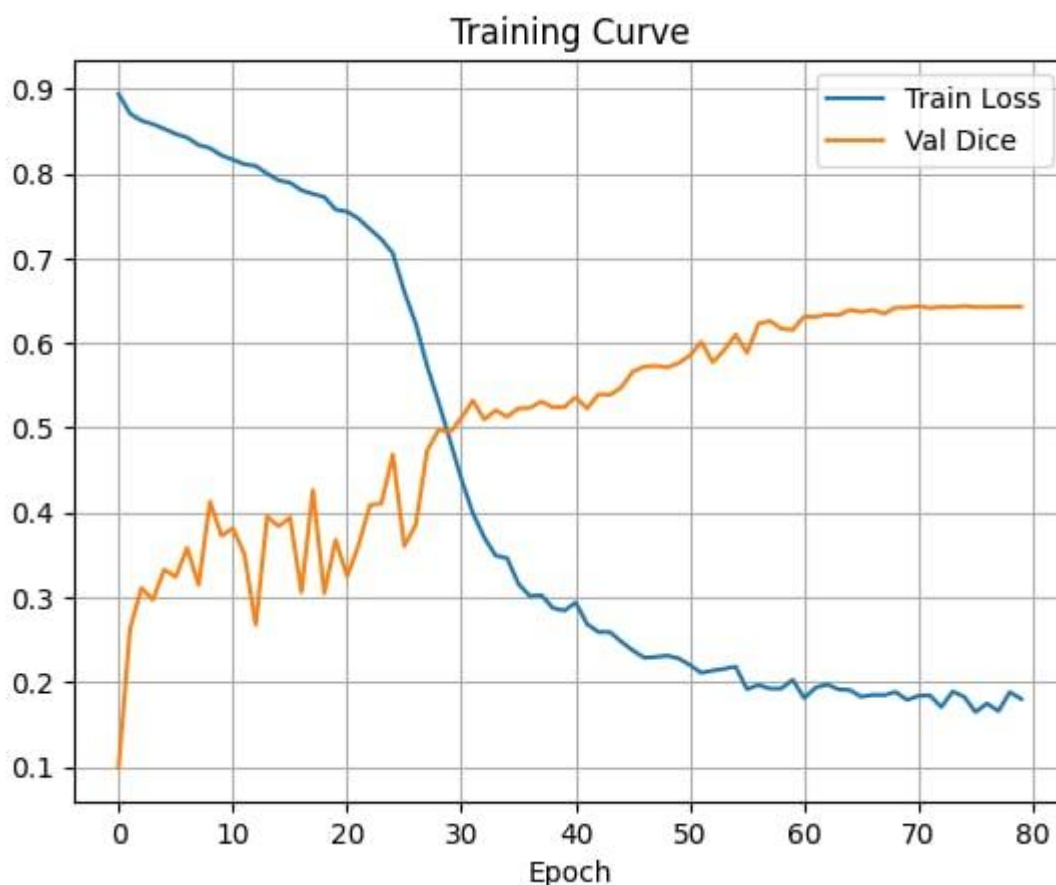
4.5.2.1 Αποτελέσματα Αρχικής Αξιολόγησης

Διευκρίνιση: Τα αποτελέσματα που παρουσιάζονται στον παρακάτω πίνακα, καθώς και σε όλες τις επόμενες ενότητες, αντιστοιχούν στο μέγιστο Dice Similarity Coefficient που παρατηρήθηκε στο *validation set* ανά μοντέλο κατά την εκπαίδευση. Το εκάστοτε καλύτερο μοντέλο αποθηκεύτηκε και αξιοποιήθηκε για το τελικό *inference*, εξασφαλίζοντας συνεπή και δίκαιη συγκριτική αξιολόγηση.

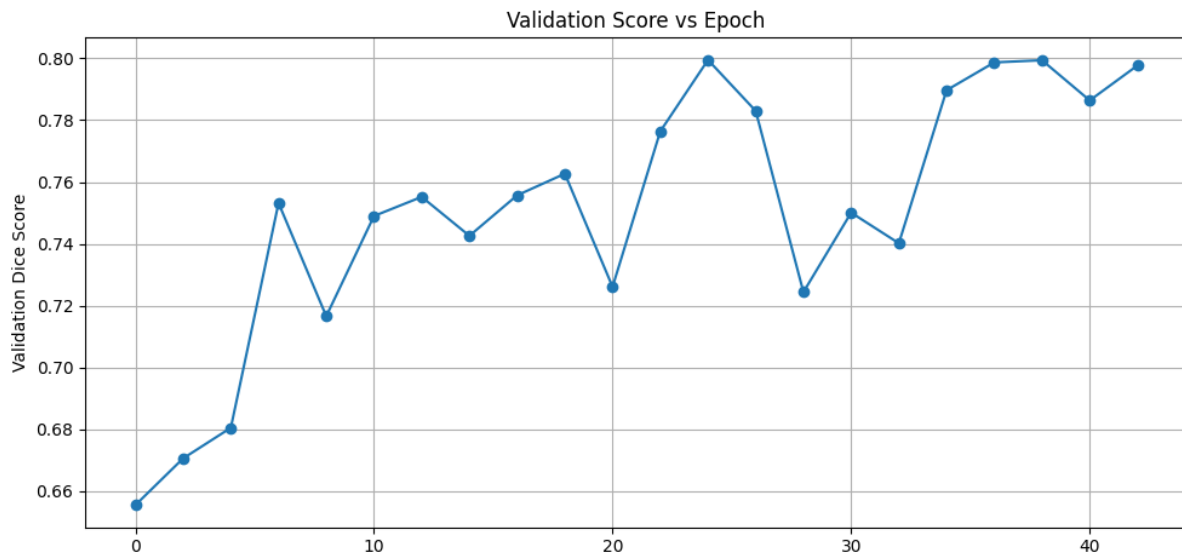
	U-Net	SegResNet	SwinUNETR	SegMamba
Dice Score	0.6438	0.6041	0.6557	0.7995

Πίνακας 2: Τιμές Dice Score για τα τέσσερα εξεταζόμενα μοντέλα

Από τα αρχικά αποτελέσματα προκύπτει ότι το SegMamba παρουσιάζει τη βέλτιστη απόδοση με Dice score 0.8, ακολουθούμενο από το SwinUNETR (0.65), το U-Net (0.64) και το SegResNet (0.61).



Εικόνα 20: Καμπύλες εκπαίδευσης για το μοντέλο U-Net. Απεικονίζονται η τιμή του Dice score στο *validation set* και η τιμή της συνάρτησης κόστους (*loss*) στο *training set* ανά εποχή. Παρατηρείται σταθερή βελτίωση της απόδοσης μέχρι τη σύγκλιση.



Εικόνα 21: Εξέλιξη του Dice Score στο validation set κατά την εκπαίδευση του μοντέλου SegMamba. Παρατηρείται καλή απόδοση από τις πρώτες εποχές (0.65), με το σκορ να συγκλίνει στο 0.80 στα τελευταία στάδια της εκπαίδευσης.

4.5.2.2 Ανάλυση Απόδοσης SegMamba

Η ανώτερη απόδοση του SegMamba μπορεί να αποδοθεί στα ακόλουθα χαρακτηριστικά της αρχιτεκτονικής:

Αποδοτικότητα σε Περιορισμένους Πόρους: Το Mamba architecture χαρακτηρίζεται από γραμμική πολυπλοκότητα ως προς το μήκος της ακολουθίας, σε αντίθεση με την τετραγωνική πολυπλοκότητα των παραδοσιακών Transformers. Αυτό το καθιστά ιδιαίτερα αποδοτικό σε περιβάλλοντα με περιορισμένους υπολογιστικούς πόρους, όπως στην παρούσα εργασία.

Βέλτιστη Αξιοποίηση Μικρών Patches: Λόγω των υπολογιστικών περιορισμών, η εκπαίδευση πραγματοποιήθηκε με σχετικά μικρά patches. Το SegMamba κατάφερε να αξιοποιήσει αποτελεσματικότερα τις τοπικές πληροφορίες σε σύγκριση με το SwinUNETR, το οποίο απαιτεί μεγαλύτερα patches για τη βέλτιστη απόδοσή του.

Robust Feature Learning: Η ικανότητα του Mamba να διατηρεί μακροπρόθεσμες εξαρτήσεις (long-range dependencies) παρέχει πλεονεκτήματα στην τμηματοποίηση πολύπλοκων ανατομικών δομών, ακόμη και όταν η εκπαίδευση γίνεται με περιορισμένα δεδομένα.

4.5.2.3 Περιορισμοί και Προκλήσεις

Παρά τα ενθαρρυντικά αποτελέσματα, τα παραπάνω scores θεωρούνται σχετικά χαμηλά για πρακτική εφαρμογή. Οι κύριοι περιορισμοί που επηρέασαν την απόδοση περιλαμβάνουν:

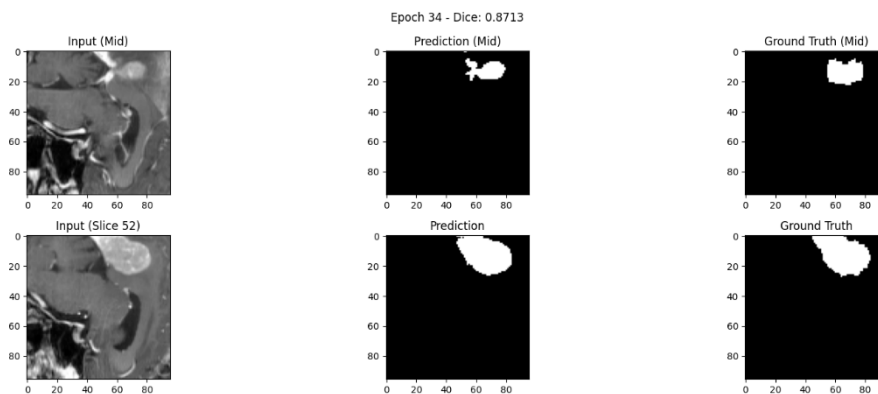
Μέγεθος Dataset: Το σύνολο εκπαίδευσης αποτελούνταν από σχετικά μικρό αριθμό δειγμάτων, γεγονός που περιορίσε την ικανότητα γενίκευσης των μοντέλων.

Δυσκολία Περιπτώσεων: Το dataset περιλάμβανε σημαντικό αριθμό απαιτητικών περιπτώσεων με πολύ μικρούς όγκους-στόχους, οι οποίοι αποτελούν εγγενώς δύσκολο πρόβλημα τμηματοποίησης.

Υπολογιστικοί Περιορισμοί: Οι περιορισμένοι υπολογιστικοί πόροι κατέστησαν αναγκαία τη χρήση μικρών *batch sizes* (1–2), μειωμένων *patch sizes* και αρχιτεκτονικών περιορισμένης πολυπλοκότητας, αποκλίνοντας από τις προτεινόμενες παραμέτρους για βέλτιστη απόδοση. Η επίδραση αυτών των περιορισμών ήταν εντονότερη στο SwinUNETR, του οποίου η αποδοτικότητα εξαρτάται σε μεγάλο βαθμό από την επεξεργασία εκτεταμένων χωρικών αποσπασμάτων (large patches).

4.5.3 Ποιοτική Ανάλυση Αποτελεσμάτων

Παρά τα αρχικά μέτρια ποσοτικά αποτελέσματα, η οπτική επιθεώρηση των αποτελεσμάτων τμηματοποίησης αποκάλυψε ένα ενδιαφέρον φαινόμενο. Οι περισσότερες από τις παραγόμενες μάσκες τμηματοποίησης παρουσίαζαν οπτικά πολύ καλή ποιότητα, με ακριβή οριοθέτηση των στοχευόμενων ανατομικών δομών και υψηλή ανταπόκριση στις πραγματικές μάσκες αναφοράς.

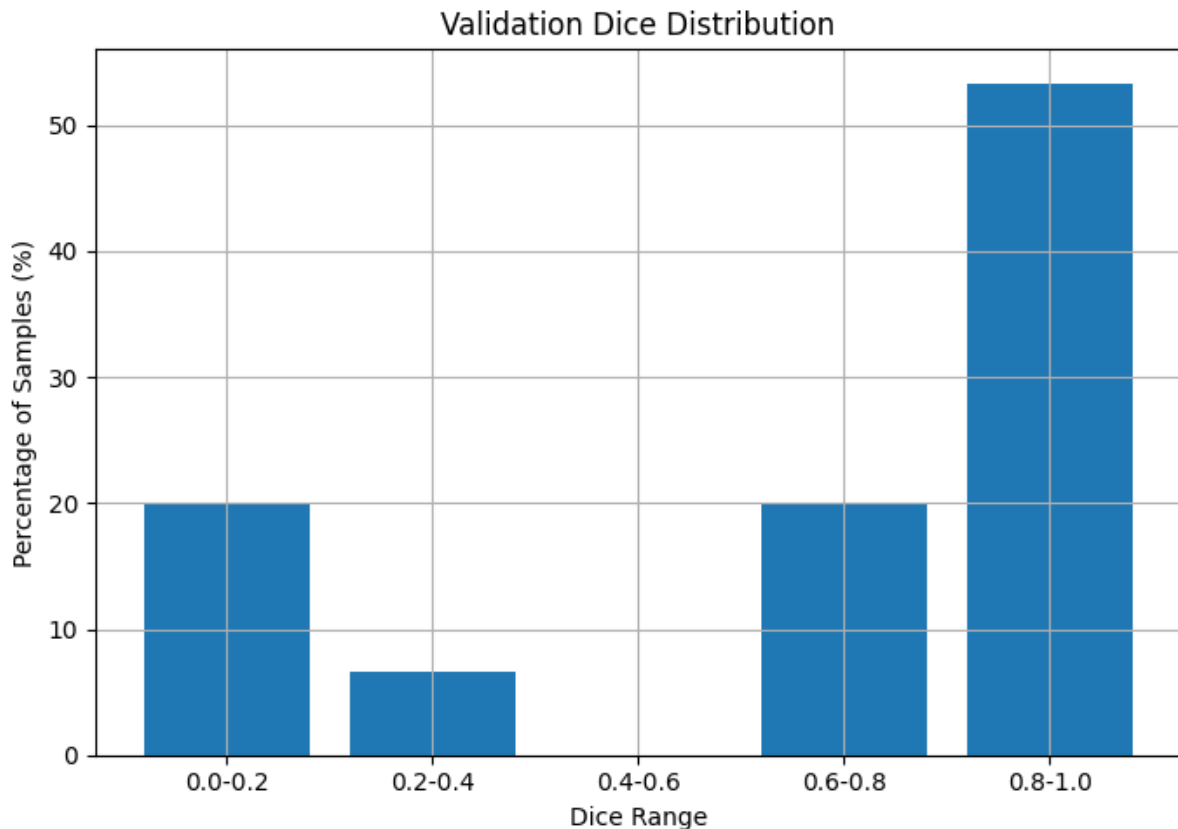


Εικόνα 22: Παράδειγμα οπτικής αξιολόγησης αποτελεσμάτων τμηματοποίησης. Εμφανίζεται η αρχική εικόνα MRI, η *ground truth* μάσκα και η πρόβλεψη του μοντέλου, αναδεικνύοντας τη γενικά υψηλή ακρίβεια της πρόβλεψης παρά τη σχετικά χαμηλή μέση *Dice* τιμή.

Αυτή η αντίφαση μεταξύ των χαμηλών μέσων τιμών *Dice scores* και της καλής οπτικής απόδοσης οδήγησε στην υπόθεση ότι τα συνολικά αποτελέσματα επηρεάζονταν δυσανάλογα από συγκεκριμένες προβληματικές περιπτώσεις, παρά από γενικευμένη χαμηλή απόδοση των μοντέλων.

4.5.3.1 Στατιστική Ανάλυση Κατανομής *Dice Scores*

Για τη διερεύνηση αυτής της υπόθεσης, πραγματοποιήθηκε λεπτομερής στατιστική ανάλυση της κατανομής των *Dice scores* σε επίπεδο μεμονωμένων εικόνων. Τα αποτελέσματα της ανάλυσης παρουσιάζονται στο παρακάτω ιστόγραμμα:



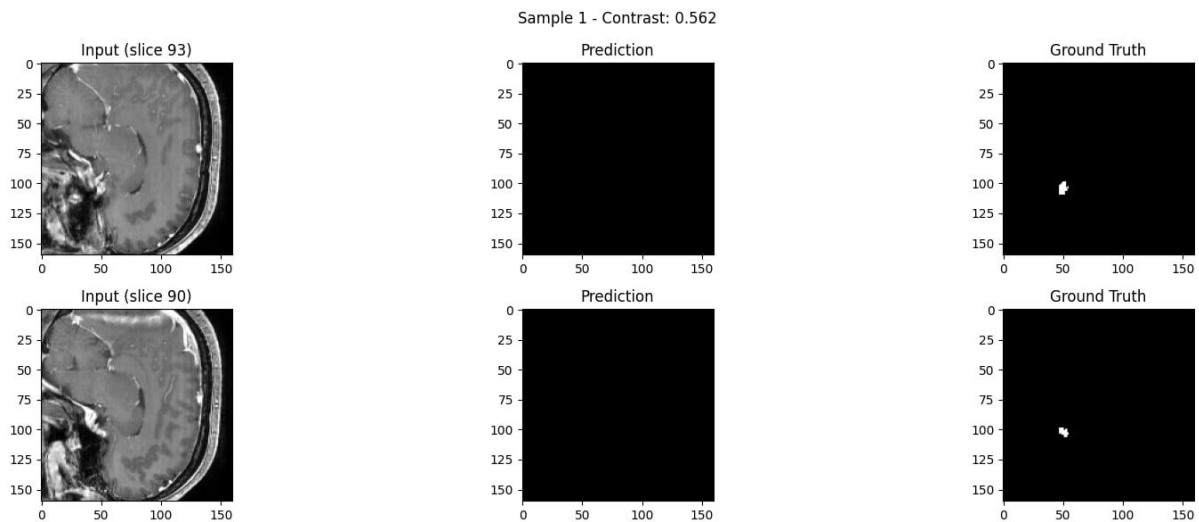
Εικόνα 23: Κατανομή του Dice score στο σύνολο επικύρωσης. Το γράφημα αποτυπώνει την απόδοση του μοντέλου σε κάθε δείγμα ξεχωριστά, αναδεικνύοντας την ύπαρξη outliers με πολύ χαμηλό score.

Τα αποτελέσματα της ανάλυσης αποκάλυψαν μια χαρακτηριστική κατανομή:

- **Κυρίαρχη συγκέντρωση υψηλών scores:** Το 53.3% των εικόνων παρουσίαζε Dice scores στο εύρος 0.8-1.0, επιβεβαιώνοντας την υψηλή ποιότητα της τμηματοποίησης για την πλειοψηφία των περιπτώσεων.
- **Ενδιάμεσες επιδόσεις:** Το 20% των εικόνων εμφάνιζε μέτρια απόδοση στο εύρος 0.6-0.8, υποδεικνύοντας περιπτώσεις με αυξημένη δυσκολία αλλά όχι πλήρη αποτυχία.
- **Χαμηλές επιδόσεις:** Το 6.7% των περιπτώσεων κυμαινόταν στο εύρος 0.2-0.4, ενώ το 20% παρουσίαζε πολύ χαμηλά scores (0.0-0.2).
- **Ασύμμετρη κατανομή:** Η κατανομή χαρακτηρίζεται από έντονη συγκέντρωση στις υψηλές τιμές, με σταδιακή μείωση προς τις χαμηλότερες κατηγορίες, υποδεικνύοντας ότι τα μοντέλα έχουν συνεπή καλή απόδοση για την πλειοψηφία των δεδομένων.

4.5.3.2 Ανάλυση Περιπτώσεων με Μηδενικό Dice Score

Η εξέταση των περιπτώσεων που παρουσίαζαν Dice score ίσο με μηδέν αποκάλυψε ένα κοινό χαρακτηριστικό: όλες αφορούσαν εικόνες με εξαιρετικά μικρούς όγκους-στόχους, συχνά αποτελούμενους από λιγότερα από μερικές εκατοντάδες *voxel*.



Εικόνα 24: Παράδειγμα περίπτωσης με μηδενικό *Dice score*. Παρά την ύπαρξη μικρών όγκων-στόχων στη *ground truth*, το μοντέλο δεν εντοπίζει την περιοχή ενδιαφέροντος, με αποτέλεσμα μηδενική επικάλυψη και αντίστοιχα μηδενικό *Dice score*.

4.5.3.3 Το Πρόβλημα του Dice Score σε Μικρούς Όγκους

Το Dice score, ως μέτρο που βασίζεται στην επικάλυψη μεταξύ πρόβλεψης *label* και πραγματικού *label*, παρουσιάζει εγγενή ευαισθησία σε μικρές ανακρίβειες όταν οι στοχευόμενοι όγκοι είναι πολύ μικροί. Συγκεκριμένα:

Μαθηματική Ευαισθησία: Όταν οι όγκοι-στόχοι είναι πολύ μικροί, ακόμη και μικρές αποκλίσεις στην πρόβλεψη ενδέχεται να οδηγούν σε πλήρη απουσία επικάλυψης, με αποτέλεσμα εξαιρετικά χαμηλό Dice score, παρά την κατά τα άλλα σχετικά καλή εντοπιστική ακρίβεια.

Φάσμα Λάθους: Σε αυτές τις περιπτώσεις, ενώ ένα μεγάλο ποσοστό των *voxels* της εικόνας ταξινομούνται σωστά (background), η μικρή απόκλιση στην περιοχή ενδιαφέροντος καθιστά το Dice score μη αντιπροσωπευτικό της πραγματικής απόδοσης.

Επίδραση στο Validation Set: Με το *validation dataset* να αποτελείται από μόλις 15 εικόνες, η παρουσία 2-3 περιπτώσεων με μηδενικό *score* επηρεάζει δραματικά τον συνολικό μέσο όρο.

4.5.4 Πρόταση Αναθεωρημένης Αξιολόγησης

4.5.4.1 Αιτιολόγηση για Τροποποίηση της Μετρικής

Βάσει της παραπάνω ανάλυσης, διαπιστώθηκε ότι η παραδοσιακή προσέγγιση υπολογισμού του μέσου Dice score δεν αντικατοπτρίζει επαρκώς την πραγματική απόδοση των μοντέλων στο συγκεκριμένο *dataset*. Οι περιπτώσεις με μηδενικό Dice score αποτελούν στατιστικά *outliers* που παραμορφώνουν την εικόνα της συνολικής απόδοσης.

4.5.4.2 Τροποποιημένη Μετρική Αξιολόγησης

Για την αντιμετώπιση αυτού του προβλήματος, τροποποιήθηκε η συνάρτηση υπολογισμού του *validation score* ώστε να αποκλείει από την συσσώρευση (aggregation) τις περιπτώσεις με Dice score ίσο με μηδέν. Αυτή η προσέγγιση:

- **Διατηρεί την αυστηρότητα:** Δεν αλλάζει τον τρόπο υπολογισμού του Dice score για κάθε εικόνα
- **Εξαλείφει τα *outliers*:** Αποκλείει τις στατιστικά ακραίες περιπτώσεις που δεν αντιπροσωπεύουν την τυπική απόδοση
- **Βελτιώνει την αντιπροσωπευτικότητα:** Παρέχει μια πιο ρεαλιστική εικόνα της απόδοσης για την πλειοψηφία των περιπτώσεων

4.5.4.3 Επαναξιολόγηση με Βελτιωμένη Μετρική

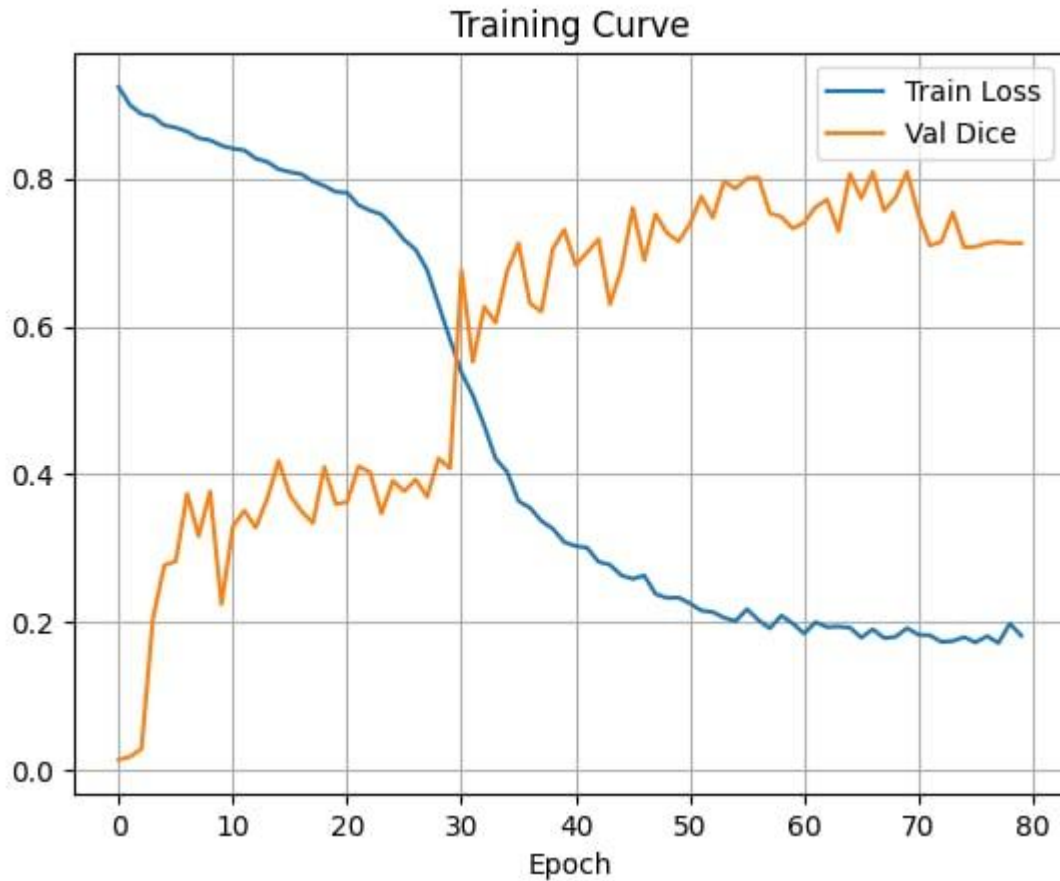
Η εφαρμογή της τροποποιημένης μετρικής στα αρχικά αποτελέσματα παρήγαγε τα ακόλουθα βελτιωμένα scores:

	U-Net	SegResNet	SwinUnetr
Συνολικό Dataset	0.6438	0.6041	0.6557
Dataset χωρίς outliers	0.8092	0.7554	0.7605

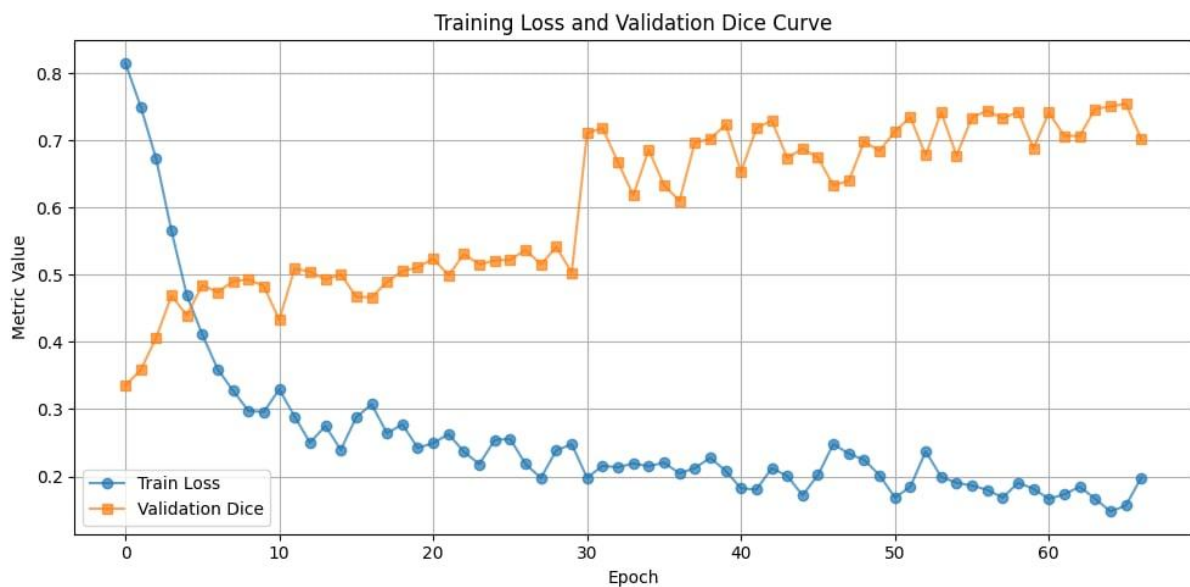
Πίνακας 3: Συγκριτική αξιολόγηση Dice Score με και χωρίς τον αποκλεισμό ακραίων τιμών (*outliers*) στο σύνολο επικύρωσης

Τα βελτιωμένα αποτελέσματα αποκαλύπτουν μια εντελώς διαφορετική εικόνα της σχετικής απόδοσης των μοντέλων:

- **U-Net (0.81):** Αναδεικνύεται ως το βέλτιστο μοντέλο, επιβεβαιώνοντας την αξιοπιστία και αποτελεσματικότητα της κλασικής αρχιτεκτονικής για την πλειοψηφία των περιπτώσεων.
- **SwinUNETR (0.76):** Παρουσιάζει σημαντική βελτίωση σε σχέση με τα αρχικά αποτελέσματα, αλλά υπολείπεται του U-Net, γεγονός που μπορεί να αποδοθεί σε αρχιτεκτονικούς περιορισμούς.
- **SegResNet (0.76):** Διατηρεί παρόμοια απόδοση με το SwinUNETR, υποδεικνύοντας ότι οι *residual connections* προσφέρουν περιορισμένο πλεονέκτημα στο συγκεκριμένο πρόβλημα.



Εικόνα 25: Καμπύλες εκπαίδευσης για το μοντέλο U-Net μετά τον αποκλεισμό των outliers. Απεικονίζονται η τιμή του Dice score στο validation set και η τιμή της συνάρτησης κόστους (loss) στο training set ανά εποχή.



Εικόνα 26: Καμπύλες εκπαίδευσης για το μοντέλο SegResNet μετά τον αποκλεισμό των outliers. Απεικονίζονται η τιμή του Dice score στο validation set και η τιμή της συνάρτησης κόστους (loss) στο training set ανά εποχή.

4.5.5 Συνολική Αξιολόγηση και Συμπεράσματα

4.5.5.1 Αποτελεσματικότητα των Μοντέλων

Τα τελικά αποτελέσματα επιβεβαιώνουν ότι, παρά τους τεχνικούς περιορισμούς (μικρό dataset, περιορισμένοι υπολογιστικοί πόροι), τα εκπαιδευμένα μοντέλα επιτυγχάνουν **υψηλής ποιότητας απόδοση** τμηματοποίησης για την πλειοψηφία των περιπτώσεων. Dice scores στο εύρος **0.75–0.81** θεωρούνται **ανταγωνιστικά** σε σχέση με τα καλύτερα δημοσιευμένα αποτελέσματα στο συγκεκριμένο dataset (0.83–0.85) [80], υποδεικνύοντας την αποτελεσματικότητα του pipeline.

4.5.5.2 Σημασία της Στατιστικής Ανάλυσης

Η παρούσα μελέτη υπογραμμίζει τη σημασία της εις βάθος στατιστικής ανάλυσης των αποτελεσμάτων πέρα από τον απλό υπολογισμό μέσων τιμών. Η εξέταση της κατανομής των *scores* και η αναγνώριση των *outliers* αποδείχθηκε κρίσιμη για την ορθή ερμηνεία της απόδοσης των μοντέλων.

4.5.5.3 Συγκριτική Απόδοση και Ερευνητικές Κατευθύνσεις

Αξίζει να σημειωθεί ότι δεδομένου ότι το SegMamba απέδωσε αρκετά καλά (Dice score 0.8) ακόμη και στο αρχικό dataset με τα *outliers*, επιτυγχάνοντας *scores* κοντά σε *state-of-the-art* επίπεδα, δεν κρίθηκε αναγκαία η περαιτέρω ανάλυσή του με την τροποποιημένη μετρική.

Η εις βάθος ανάλυση εστίασε στα τρία υπόλοιπα μοντέλα (U-Net, SegResNet, SwinUNETR) καθώς για αυτά αναπτύχθηκε εξ αρχής πλήρες *pipeline* επεξεργασίας, συμπεριλαμβανομένης της ανάλυσης δεδομένων, του *preprocessing*, των *data augmentation* τεχνικών και του πειραματισμού με διάφορες υπερπαραμέτρους και διαστάσεις. Αντίθετα, το SegMamba παρέχει ένα ολοκληρωμένο αυτοματοποιημένο *pipeline* που διαχειρίζεται εσωτερικά όλες τις παραπάνω διεργασίες, περιορίζοντας τη δυνατότητα παρέμβασης του ερευνητή κυρίως στις υπερπαραμέτρους εκπαίδευσης. Αυτός ο περιορισμός καθιστά την εις βάθος ανάλυση του SegMamba λιγότερο ενδιαφέρουσα από ερευνητικής άποψης.

4.5.5.4 Ανάλυση Υπεροχής του U-Net

Από τα τρία custom μοντέλα, το **U-Net αναδείχθηκε ως το βέλτιστο** μετά την εφαρμογή της τροποποιημένης μετρικής. Αυτό το αποτέλεσμα μπορεί να αποδοθεί σε πολλαπλούς παράγοντες:

Αρχιτεκτονικοί Περιορισμοί του SwinUNETR: Λόγω των περιορισμένων υπολογιστικών πόρων, το SwinUNETR αναγκάστηκε να λειτουργήσει με σημαντικούς αρχιτεκτονικούς περιορισμούς:

- Χρήση εξαιρετικά μικρών *patch sizes* που μειώνουν την ικανότητα αναγνώρισης μεγάλων χωρικών εξαρτήσεων
- Ρηχή αρχιτεκτονική με περιορισμένο αριθμό *transformer layers*
- Μειωμένη διαστατικότητα των *attention heads* λόγω περιορισμών μνήμης

Πλεονεκτήματα του U-Net στο Συγκεκριμένο Πλαίσιο: Το U-Net, ως απλούστερη και πιο αποδοτική αρχιτεκτονική, κατάφερε να αξιοποιήσει καλύτερα τους διαθέσιμους υπολογιστικούς πόρους:

- Δυνατότητα εκπαίδευσης αρχιτεκτονικής επαρκούς μεγέθους με βαθύτερα *layers*
- Αποτελεσματική χρήση των *convolutional operations* για τα περιορισμένα δεδομένα
- Σταθερότητα εκπαίδευσης που επιτρέπει καλύτερη σύγκλιση

Επίδραση του Μεγέθους Dataset: Το μικρό μέγεθος του *dataset* ενδεχομένως ευνοεί τις πιο απλές αρχιτεκτονικές όπως το U-Net, καθώς τα *transformer-based* μοντέλα τυπικά απαιτούν μεγαλύτερες ποσότητες δεδομένων για να εκπαιδευτούν αποτελεσματικά.

Συμπέρασμα: Τα αποτελέσματα υποδεικνύουν ότι σε περιβάλλοντα με περιορισμένους υπολογιστικούς πόρους και μικρά *datasets*, οι κλασικές CNN αρχιτεκτονικές όπως το U-Net μπορεί να υπερτερούν των πιο σύνθετων *transformer-based* μοντέλων, ιδιαίτερα όταν τα τελευταία αναγκάζονται να λειτουργήσουν με σημαντικούς αρχιτεκτονικούς συμβιβασμούς. Αυτό το εύρημα τονίζει τη σημασία της προσεκτικής επιλογής αρχιτεκτονικής βάσει των διαθέσιμων πόρων και όχι μόνο των θεωρητικών δυνατοτήτων του κάθε μοντέλου.

Συμπεράσματα-Μελλοντικές Προεκτάσεις

Στην παρούσα εργασία παρουσιάστηκε μία ολοκληρωμένη διαδικασία αξιολόγησης τεσσάρων διαφορετικών αρχιτεκτονικών βαθιάς μάθησης για την τρισδιάστατη τμηματοποίηση μηνιγγιωμάτων σε δεδομένα μαγνητικής τομογραφίας. Η αξιολόγηση έγινε με κοινό πειραματικό πλαίσιο, αξιοποιώντας το σύνολο δεδομένων Meningioma-SEG-CLASS και εφαρμόζοντας συνεπή στρατηγική προεπεξεργασίας, παραμετροποίησης και αξιολόγησης.

Τα αποτελέσματα ανέδειξαν την υπεροχή της αρχιτεκτονικής SegMamba έναντι των υπολοίπων, ενώ παράλληλα αναδείχθηκε και η ανταγωνιστική επίδοση των SwinUNETR και U-Net, υπό την προϋπόθεση σωστού συνδυασμού παραμέτρων και τεχνικών *augmentation*. Ιδιαίτερη βαρύτητα δόθηκε στην επιλογή κατάλληλων *loss functions*, *schedulers* και τεχνικών *regularization*, με σκοπό τη βελτίωση της γενίκευσης και της σταθερότητας κατά την εκπαίδευση.

Η εργασία αυτή επιβεβαιώνει τη σημασία του ολοκληρωμένου σχεδιασμού ενός pipeline ιατρικής τμηματοποίησης, καθώς και τη δυνατότητα επίτευξης υψηλής απόδοσης ακόμη και σε περιορισμένα δεδομένα, μέσω τεχνικών προσεκτικής βελτιστοποίησης.

Στο μέλλον, η παρούσα εργασία μπορεί να επεκταθεί με τη διερεύνηση πιο πολύπλοκων τεχνικών βελτιστοποίησης, όπως **self-supervised pretraining** [81] ή **meta-learning** για προσαρμογή σε περιορισμένα δεδομένα [82], [83]. Επιπλέον, η **εφαρμογή ensemble μεθόδων** ή **test-time augmentation** [84] μπορεί να ενισχύσει περαιτέρω την απόδοση. Τέλος, ενδιαφέρον παρουσιάζει η διερεύνηση των μοντέλων σε πολυκεντρικά *datasets* για αξιολόγηση της γενίκευσης σε διαφορετικά πρωτόκολλα μαγνητικής τομογραφίας και χαρακτηριστικά ασθενών.

Βιβλιογραφία

- [1] G. Litjens *et al.*, “A survey on deep learning in medical image analysis,” *Medical Image Analysis*, vol. 42, pp. 60–88, 2017, doi: 10.1016/j.media.2017.07.005.
- [2] B. H. Menze *et al.*, “The multimodal brain tumor image segmentation benchmark (BRATS),” *IEEE Trans. Med. Imaging*, vol. 99, pp. 1–12, 2014, doi: 10.1109/TMI.2014.2377694.
- [3] D. N. Louis *et al.*, “The 2021 WHO classification of tumors of the central nervous system: A summary,” *Neuro-Oncology*, vol. 23, no. 8, pp. 1231–1251, 2021, doi: 10.1093/neuonc/noab106.
- [4] A. Vassantachart *et al.*, “A repository of grade 1 and 2 meningioma MRIs in a public dataset for radiomics reproducibility tests,” *Med. Phys.*, vol. 51, no. 3, pp. 2334–2344, 2024, doi: 10.1002/mp.16763.
- [5] J. Boutry *et al.*, “The evolution and ecology of benign tumors,” *Biochim. Biophys. Acta Rev. Cancer*, vol. 1875, no. 2, p. 188643, 2021, doi: 10.1016/j.bbcan.2021.188643.
- [6] L. M. DeAngelis, “Brain tumors,” *N. Engl. J. Med.*, vol. 344, no. 2, pp. 114–123, Jan. 2001, doi: 10.1056/NEJM200101113440207.
- [7] M. Kanda *et al.*, “Pattern of lymph node metastasis spread in pancreatic cancer,” *Pancreas*, vol. 40, no. 6, pp. 951–955, 2011, doi: 10.1097/MPA.0b013e3182148342.
- [8] S. H. Jafari *et al.*, “Breast cancer diagnosis: Imaging techniques and biochemical markers,” *J. Cell. Physiol.*, vol. 233, no. 7, pp. 5200–5213, Jul. 2018, doi: 10.1002/jcp.26379.
- [9] R. L. Siegel *et al.*, “Cancer statistics, 2022,” *CA Cancer J. Clin.*, vol. 72, no. 1, pp. 7–33, Jan. 2022, doi: 10.3322/caac.21708.
- [10] S. Yang *et al.*, “Immune cell infiltration as a biomarker for the diagnosis and prognosis of digestive system cancer,” *Cancer Sci.*, vol. 110, no. 12, pp. 3639–3649, Dec. 2019, doi: 10.1111/cas.14216.
- [11] J. D. Minna *et al.*, “Focus on lung cancer,” *Cancer Cell*, vol. 1, no. 1, pp. 49–52, Feb. 2002, doi: 10.1016/S1535-6108(02)00027-2.
- [12] The Institute for Health Metrics and Evaluation (2020). Global Burden of Disease (GBD). [online] www.healthdata.org. Available at: <https://www.healthdata.org/research-analysis/gbd>.
- [13] commission.europa.eu. (n.d.). A cancer plan for Europe. [online] Available at: https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/promoting-our-european-way-life/european-health-union/cancer-plan_europe_en.
- [14] T. C. W. Mok and A. C. S. Chung, “Learning Data Augmentation for Brain Tumor Segmentation with Coarse-to-Fine Generative Adversarial Networks,” in *Lecture notes in computer science*, 2019, pp. 70–80. doi: 10.1007/978-3-030-11723-8_7.

- [15] Y. KOO et al., “Brain cancer diagnosis and therapy with nanoplatforms☆,” *Adv Drug Deliv Rev*, vol. 58, no. 14, pp. 1556–1577, Dec. 2006, doi: 10.1016/j.addr.2006.09.012.
- [16] M. Lara-Velazquez et al., “Advances in Brain Tumor Surgery for Glioblastoma in Adults,” *Brain Sci*, vol. 7, no. 12, p. 166, Dec. 2017, doi: 10.3390/brainsci7120166.
- [17] M. Akram, M. Iqbal, M. Daniyal, and A. U. Khan, “Awareness and current knowledge of breast cancer,” *Biol Res*, vol. 50, no. 1, p. 33, Dec. 2017, doi: 10.1186/s40659-017-0140-9.
- [18] J. Wiemels, M. Wrensch, and E. B. Claus, “Epidemiology and etiology of meningioma,” *Journal of Neuro-Oncology*, vol. 99, no. 3, pp. 307–314, Sep. 2010, doi: 10.1007/s11060-010-0386-3.
- [19] Q. T. Ostrom *et al.*, “CBTRUS Statistical Report: Primary brain and other central nervous system tumors diagnosed in the United States in 2016—2020,” *Neuro-Oncology*, vol. 25, no. Supplement_4, pp. iv1–iv99, Oct. 2023, doi: 10.1093/neuonc/noad149.
- [20] E. B. Claus *et al.*, “Survival and low-grade glioma: the emergence of genetic information,” *Neurosurgical FOCUS*, vol. 38, no. 1, p. E6, Jan. 2015, doi: 10.3171/2014.10.focus12367.
- [21] M. J. Goldman, B. S. Teh, S. S. Lo, E. B. Butler, and D. S. Baskin, “Radiosurgery for Intracranial Meningiomas: A Review of Anatomical Challenges and an Update on the Evidence,” *Cancers*, vol. 15, no. 6, p. 1713, Mar. 2023. doi: 10.3390/cancers15061713
- [22] R. H. Hruban, G. M. Petersen, P. K. Ha, and S. E. Kern, “Genetics of Pancreatic Cancer,” *Surg Oncol Clin N Am*, vol. 7, no. 1, pp. 1–23, Jan. 1998, doi: 10.1016/S1055-3207(18)30282-5.
- [23] K. N. Nathanson, R. Wooster, and B. L. Weber, “Breast cancer genetics: What we know and what we need,” *Nat Med*, vol. 7, no. 5, pp. 552–556, May 2001, doi: 10.1038/87876.
- [24] A. R. Venkitaraman, “How do mutations affecting the breast cancer genes BRCA1 and BRCA2 cause cancer susceptibility?,” *DNA Repair (Amst)*, vol. 81, p. 102668, Sep. 2019, doi: 10.1016/j.dnarep.2019.102668
- [25] A. J. Sasco, M. B. Secretan, and K. Straif, “Tobacco smoking and cancer: a brief review of recent epidemiological evidence,” *Lung Cancer*, vol. 45, pp. S3–S9, Aug. 2004, doi: 10.1016/j.lungcan.2004.07.998.
- [26] E. Weiderpass, “Lifestyle and Cancer Risk,” *Journal of Preventive Medicine and Public Health*, vol. 43, no. 6, p. 459, 2010, doi: 10.3961/jpmph.2010.43.6.459.
- [27] V. A. Katzke, R. Kaaks, and T. Kühn, “Lifestyle and Cancer Risk,” *The Cancer Journal*, vol. 21, no. 2, pp. 104–110, Mar. 2015, doi: 10.1097/PPO.0000000000000101.
- [28] J. Campisi, J. K. Andersen, P. Kapahi, and S. Melov, “Cellular senescence: A link between cancer and age-related degenerative disease?,” *Semin Cancer Biol*, vol. 21, no. 6, pp. 354–359, Dec. 2011, doi: 10.1016/J.SEMCANCER.2011.09.001.

- [29] American Cancer Society – *Signs and Symptoms of Cancer*.
(<https://www.cancer.org/cancer/cancer-basics/signs-and-symptoms-of-cancer.html>)
- [30] M. G. Marmot, D. G. Altman, D. A. Cameron, J. A. Dewar, S. G. Thompson, and M. Wilcox, “The benefits and harms of breast cancer screening: an independent review,” *British Journal of Cancer*, vol. 108, no. 11, pp. 2205–2240, 2013, doi: 10.1038/bjc.2013.177.
- [31] M. Weller *et al.*, “European Association for Neuro-Oncology (EANO) guideline on the diagnosis and treatment of adult astrocytic and oligodendroglial gliomas,” *The Lancet Oncology*, vol. 18, no. 6, pp. e315–e329, May 2017, doi: 10.1016/s1470-2045(17)30194-8.
- [32] Mayo Clinic (2019). Brain Tumor - Diagnosis and Treatment - Mayo Clinic. [online] Mayoclinic.org. Available at: <https://www.mayoclinic.org/diseasesconditions/brain-tumor/diagnosis-treatment/drc-20350088>.
- [33] M. G. Linguraru *et al.*, “Tumor Burden Analysis on Computed Tomography by Automated Liver and Tumor Segmentation,” *IEEE Trans Med Imaging*, vol. 31, no. 10, pp. 1965–1976, Oct. 2012, doi: 10.1109/TMI.2012.2211887.
- [34] F. Yepes-Calderon and J. Gordon McComb, “Manual Segmentation Errors in Medical Imaging. Proposing a Reliable Gold Standard,” 2019, pp. 230–241. doi: 10.1007/978-3-030-32475-9_17.
- [35] L. Wang, “Early Diagnosis of Breast Cancer,” *Sensors*, vol. 17, no. 7, p. 1572, Jul. 2017, doi: 10.3390/s17071572
- [36] E. S. McDonald, A. S. Clark, J. Tchou, P. Zhang, and G. M. Freedman, “Clinical Diagnosis and Management of Breast Cancer,” *Journal of Nuclear Medicine*, vol. 57, no. Supplement 1, pp. 9S-16S, Feb. 2016, doi: 10.2967/jnumed.115.157834.
- [37] D. M. Aleman, A. Kumar, R. K. Ahuja, H. E. Romeijn, and J. F. Dempsey, “Neighborhood search approaches to beam orientation optimization in intensity modulated radiation therapy treatment planning,” *Journal of Global Optimization*, vol. 42, no. 4, pp. 587–607, Dec. 2008, doi: 10.1007/s10898-008-9286-x.
- [38] G.-P. Yan, L. Robinson, and P. Hogg, “Magnetic resonance imaging contrast agents: Overview and perspectives,” *Radiography*, vol. 13, pp. e5–e19, Dec. 2007, doi: 10.1016/j.radi.2006.07.005
- [39] W. G. B. C. J. L. Ray Hashman Hashemi, *MRI: The Basics: The Basics* .
- [40] Catherine Westbrook and John Talbot, *MRI in Practice*.
- [41] D. B. Plewes and W. Kucharczyk, “Physics of MRI: A primer,” *Journal of Magnetic Resonance Imaging*, vol. 35, no. 5, pp. 1038–1054, May 2012, doi: 10.1002/jmri.23642.
- [42] H. A. Sharma, “MRI physics–basic principles,” *Acta Neuropsychiatr*, vol. 21, no. 4, pp. 200–201, Aug. 2009, doi: 10.1111/j.1601-5215.2009.00404.x.
- [43] A. Berger, “How does it work?: Magnetic resonance imaging,” *BMJ*, vol. 324, no. 7328, p. 35, Jan. 2002, doi: 10.1136/bmj.324.7328.35.

- [44] T. Yousaf, G. Dervenoulas, and M. Politis, “Advances in MRI Methodology,” *Int Rev Neurobiol*, vol. 141, pp. 31–76, Jan. 2018, doi: 10.1016/BS.IRN.2018.08.008
- [45] I. Young, “Nuclear magnetic resonance imaging,” *Electronics and Power*, vol. 30, no. 3, p. 205, 1984, doi: 10.1049/ep.1984.0112.
- [46] V. Mlynárik, “Introduction to nuclear magnetic resonance,” *Anal Biochem*, vol. 529, pp. 4–9, Jul. 2017, doi: 10.1016/J.AB.2016.05.006
- [47] Ματσόπουλος Γεώργιος, Σημειώσεις Μαθήματος Έπεξεργασία και Ανάλυση Ιατρικών Σημάτων. EMPI, 2024.
- [48] N. Agarwal and D. Yadav, “A comprehensive analysis of classical machine learning and modern deep learning methodologies,” *Int. J. Eng. Res. Technol.*, vol. 13, 2024, doi: 10.17577/IJERTV13IS050275.
- [49] B. A. Jonsson *et al.*, “Brain age prediction using deep learning uncovers associated sequence variants,” *Nature Communications*, vol. 10, no. 1, Nov. 2019, doi: 10.1038/s41467-019-13163-9.
- [50] E. García-Gonzalo, Z. Fernández-Muñiz, P. J. G. Nieto, A. B. Sánchez, and M. M. Fernández, “Hard-Rock Stability Analysis for Span Design in Entry-Type Excavations with Learning Classifiers,” *Materials*, vol. 9, no. 7, p. 531, Jun. 2016, doi: 10.3390/ma9070531.
- [51] H. Yin, A. Aryani, S. Petrie, A. Nambissan, A. Astudillo, and S. Cao, “A rapid review of clustering algorithms,” *arXiv (Cornell University)*, Jan. 2024, doi: 10.48550/arxiv.2401.07389.
- [52] M. Khosla, K. Jamison, G. H. Ngo, A. Kuceyeski, and M. R. Sabuncu, “Machine learning in resting-state fMRI analysis,” *Magnetic Resonance Imaging*, vol. 64, pp. 101–121, Jun. 2019, doi: 10.1016/j.mri.2019.05.031.
- [53] C. O. S. Sorzano, J. Vargas, and A. D. Pascual-Montano, “A survey of dimensionality reduction techniques,” *arXiv (Cornell University)*, Jan. 2014, doi: 10.48550/arxiv.1403.2877.
- [54] T. Lu, L. Wang, and X. Zhao, “Review of anomaly Detection Algorithms for Data streams,” *Applied Sciences*, vol. 13, no. 10, p. 6353, May 2023, doi: 10.3390/app13106353.
- [55] ZS. K. Zhou, H. N. Le, K. Luu, H. Nguyen V., and N. Ayache, “Deep reinforcement learning in medical imaging: A literature review,” *arXiv (Cornell University)*, Jan. 2021, doi: 10.48550/arxiv.2103.05115.
- [56] H. R. Tizhoosh and G. W. Taylor, “REINFORCED CONTRAST ADAPTATION,” *International Journal of Image and Graphics*, vol. 06, no. 03, pp. 377–392, Jul. 2006, doi: 10.1142/s0219467806002379.
- [57] A. Esteva *et al.*, “A guide to deep learning in healthcare,” *Nature Medicine*, vol. 25, no. 1, pp. 24–29, Dec. 2018, doi: 10.1038/s41591-018-0316-z.

- [58] E. J. Topol, “High-performance medicine: the convergence of human and artificial intelligence,” *Nature Medicine*, vol. 25, no. 1, pp. 44–56, Dec. 2018, doi: 10.1038/s41591-018-0300-7.
- [59] D. P. Kingma and J. L. Ba, “Adam: A method for stochastic optimization,” *arXiv (Cornell University)*, Jan. 2014, doi: 10.48550/arxiv.1412.6980.
- [60] W. Xing and D. Du, “Dropout prediction in MOOCs: Using deep learning for personalized intervention,” *Journal of Educational Computing Research*, vol. 57, no. 3, pp. 547–570, Mar. 2018, doi: 10.1177/0735633118757015.
- [61] V. H. Phung and E. J. Rhee, “A High-Accuracy Model average ensemble of convolutional neural networks for classification of cloud image patches on small datasets,” *Applied Sciences*, vol. 9, no. 21, p. 4500, Oct. 2019, doi: 10.3390/app9214500.
- [62] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, Jan. 1998, doi: 10.1109/5.726791.
- [63] O'Shea, K., & Nash, R. (2015). An introduction to convolutional neural networks. *arXiv preprint arXiv:1511.08458*.
- [64] D. Scherer, A. Müller, and S. Behnke, “Evaluation of pooling operations in convolutional architectures for object recognition,” in *Lecture notes in computer science*, 2010, pp. 92–101. doi: 10.1007/978-3-642-15825-4_10.
- [65] R. Yamashita, M. Nishio, R. K. Gian DO, and K. Togashi, “Convolutional neural networks: an overview and application in radiology,” *Insights Into Imaging*, vol. 9, no. 4, pp. 611–629, Jun. 2018, doi: 10.1007/s13244-018-0639-9.
- [66] K. He *et al.*, “Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification,” in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2015, pp. 1026–1034, doi: 10.1109/ICCV.2015.123.
- [67] C. Shorten and T. M. Khoshgoftaar, “A survey on Image Data Augmentation for Deep Learning,” *Journal of Big Data*, vol. 6, no. 1, Jul. 2019, doi: 10.1186/s40537-019-0197-0.
- [68] H. Bello *et al.*, “Towards certifiable AI in aviation: Landscape, challenges, and opportunities,” *arXiv preprint*, arXiv:2409.08666, 2024.
- [69] A. Vaswani *et al.*, “Attention is all you need,” in *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 30, 2017.
- [70] A. Dosovitskiy *et al.*, “An image is worth 16×16 words: Transformers for image recognition at scale,” *arXiv preprint*, arXiv:2010.11929, 2020.
- [71] S. R. Choi and M. Lee, “Transformer Architecture and Attention Mechanisms in Genome Data Analysis: A Comprehensive review,” *Biology*, vol. 12, no. 7, p. 1033, Jul. 2023, doi: 10.3390/biology12071033.

- [72] O. Ronneberger, P. Fischer, and T. Brox, “U-NET: Convolutional Networks for Biomedical Image Segmentation,” in *Lecture notes in computer science*, 2015, pp. 234–241. doi: 10.1007/978-3-319-24574-4_28.
- [73] K. He *et al.*, “Deep residual learning for image recognition,” *arXiv preprint*, arXiv:1512.03385, 2015.
- [74] M. J. Cardoso *et al.*, “MONAI: An open-source framework for deep learning in healthcare,” *arXiv (Cornell University)*, Jan. 2022, doi: 10.48550/arxiv.2211.02701.
- [75] K. Sakamoto *et al.*, “Automated evaluation of masseter muscle volume: deep learning prognostic approach in oral cancer,” *BMC Cancer*, vol. 24, no. 1, Jan. 2024, doi: 10.1186/s12885-024-11873-y.
- [76] A. Hatamizadeh, V. Nath, Y. Tang, D. Yang, H. R. Roth, and D. Xu, “SWIN UNETR: SWIN Transformers for Semantic Segmentation of Brain Tumors in MRI images,” in *Lecture notes in computer science*, 2022, pp. 272–284. doi: 10.1007/978-3-031-08999-2_22.
- [77] Z. Xing, T. Ye, Y. Yang, G. Liu, and L. Zhu, “SEGMAMBA: Long-Range Sequential Modeling MAMBA for 3D medical Image segmentation,” in *Lecture notes in computer science*, 2024, pp. 578–588. doi: 10.1007/978-3-031-72111-3_54.
- [78] A. Gu and T. Dao, “Mamba: Linear-time sequence modeling with selective state spaces,” *arXiv preprint*, arXiv:2312.00752, 2023.
- [79] I. Loshchilov and F. Hutter, “Decoupled weight decay regularization,” in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2017.
- [80] Z. Song *et al.*, “REHRSeg: Unleashing the power of self-supervised super-resolution for Resource-Efficient 3D MRI Segmentation,” *Neurocomputing*, p. 129425, Jan. 2025, doi: 10.1016/j.neucom.2025.129425.
- [81] Z. Liu *et al.*, “A review of self-supervised, generative, and few-shot deep learning methods for data-limited magnetic resonance imaging segmentation,” *NMR in Biomedicine*, vol. 37, no. 8, Mar. 2024, doi: 10.1002/nbm.5143.
- [82] R. Khadka *et al.*, “Meta-learning with implicit gradients in a few-shot setting for medical image segmentation,” *Computers in Biology and Medicine*, vol. 143, p. 105227, Jan. 2022, doi: 10.1016/j.compbio.2022.105227.
- [83] A. M. Alsaleh, E. Albalawi, A. Algosaibi, S. S. Albakheet, and S. B. Khan, “Few-Shot learning for medical image segmentation using 3D U-Net and Model-Agnostic Meta-Learning (MAML),” *Diagnostics*, vol. 14, no. 12, p. 1213, Jun. 2024, doi: 10.3390/diagnostics14121213.
- [84] X. Ma, Y. Tao, Y. Zhang, Z. Ji, Y. Zhang, and Q. Chen, “Test-Time generative augmentation for medical image segmentation,” *arXiv (Cornell University)*, Jun. 2024, doi: 10.48550/arxiv.2406.17608.

