



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

**Τμηματοποίηση εικόνων μαγνητικής τομογραφίας για εντοπισμό όγκων:
Σύγκριση μεθόδων με Συνελικτικά Νευρωνικά Δίκτυα**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

ΔΙΑΚΟΠΟΥΛΟΥ ΟΡΕΣΤΗ

Επιβλέπων: Γεώργιος Ματσόπουλος
Καθηγητής Ε.Μ.Π.

Συνεπιβλέπουσα: Ουρανία Πετροπούλου
ΕΔΙΠ Ε.Μ.Π.

Αθήνα, Σεπτέμβριος 2025

Η σελίδα αυτή είναι σκόπιμα λευκή.



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

**Τμηματοποίηση εικόνων μαγνητικής τομογραφίας για εντοπισμό όγκων:
Σύγκριση μεθόδων με Συνελικτικά Νευρωνικά Δίκτυα**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

ΔΙΑΚΟΠΟΥΛΟΥ ΟΡΕΣΤΗ

Επιβλέπων: Γεώργιος Ματσόπουλος
Καθηγητής Ε.Μ.Π.

Συνεπιβλέπουσα: Ουρανία Πετροπούλου
ΕΔΙΠ Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την **29 Σεπτεμβρίου 2025**.

.....
Γ. Ματσόπουλος
Καθηγητής Ε.Μ.Π.

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

.....
Π. Τσανάκας
Καθηγητής Ε.Μ.Π.

Αθήνα, Σεπτέμβριος 2025

(Υπογραφή)

.....

ΔΙΑΚΟΠΟΥΛΟΣ ΟΡΕΣΤΗΣ

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

Copyright © Διακόπουλος Ορέστης, 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 και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν 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 ερμηνευθούν ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Περίληψη

Ο καρκίνος του εγκεφάλου αποτελεί μια από τις πιο σοβαρές προκλήσεις για τη σύγχρονη ιατρική, λόγω της πολυπλοκότητας της ανατομίας του εγκεφάλου, της ετερογένειας των όγκων και των κινδύνων που συνεπάγεται η κλασική χειρουργική ή ακτινοθεραπευτική προσέγγιση. Η ακριβής τμηματοποίηση όγκων σε εικόνες μαγνητικής τομογραφίας (MRI) είναι ζωτικής σημασίας για τη διάγνωση, τον σχεδιασμό θεραπείας και την παρακολούθηση της πορείας της νόσου. Ωστόσο, η χειροκίνητη σήμανση από εξειδικευμένους νευροακτινολόγους είναι χρονοβόρα, επιρρεπής σε υποκειμενικότητα και δύσκολα κλιμακώσιμη.

Η παρούσα εργασία εστιάζει στη συγκριτική αξιολόγηση σύγχρονων αρχιτεκτονικών βαθιάς μάθησης για την αυτόματη τμηματοποίηση όγκων εγκεφάλου. Αναπτύχθηκε πλήρης πειραματικός αγωγός, που περιλαμβάνει προεπεξεργασία και κανονικοποίηση MRI όγκων, patch-based εκπαίδευση με στοχαστικές επαυξήσεις, και αξιολόγηση με βάση τον Dice score. Τέσσερις διαφορετικές αρχιτεκτονικές συγκρίθηκαν υπό ενιαίο pipeline: η UNet ως καθιερωμένη baseline, η VNet με τρισδιάστατες συνελίξεις, η SegResNet με εκτεταμένα residual blocks και η SwinUNETR, ένα υβρίδιο CNN/Transformer με self-attention.

Τα αποτελέσματα έδειξαν ότι όλα τα μοντέλα συγκλίνουν επιτυχώς, με διαφορετικά όμως χαρακτηριστικά: η UNet προσφέρει εξαιρετική ισορροπία απλότητας και ακρίβειας, η SegResNet παρουσιάζει βελτιωμένη σταθερότητα, η VNet αξιοποιεί πλήρως το τρισδιάστατο συμφραζόμενο αλλά με μεγαλύτερη διακύμανση, ενώ η SwinUNETR κατέγραψε τη συνολικά καλύτερη απόδοση. Η εργασία καταλήγει ότι οι υβριδικές αρχιτεκτονικές με attention μηχανισμούς εμφανίζουν σημαντικές προοπτικές για μελλοντική έρευνα, χωρίς όμως να αναιρείται η πρακτική αξία απλούστερων δικτύων όπως η U-Net.

Η μελέτη αυτή συμβάλλει στη βαθύτερη κατανόηση των πλεονεκτημάτων και περιορισμών διαφορετικών μεθοδολογιών τμηματοποίησης και θέτει τις βάσεις για περαιτέρω έρευνα με πολυκλαδικές προσεγγίσεις, μεγαλύτερα datasets και βελτιστοποίηση υπολογιστικής απόδοσης, ώστε τέτοιες μέθοδοι να καταστούν περισσότερο εφαρμόσιμες στην καθημερινή κλινική πράξη.

Λέξεις κλειδιά: MRI, Εγκεφαλικοί όγκοι, Τμηματοποίηση εικόνας, Βαθιά μάθηση, Συνελικτικά Νευρωνικά Δίκτυα, U-Net, V-Net, SegResNet, SwinUNETR.

Abstract

Brain cancer remains one of the most serious challenges in modern medicine, due to the complexity of brain anatomy, the heterogeneity of tumors, and the risks associated with classical surgical or radiotherapeutic approaches. Accurate segmentation of tumors in magnetic resonance imaging (MRI) is of vital importance for diagnosis, treatment planning, and monitoring of disease progression. However, manual annotation by expert neuroradiologists is time-consuming, prone to subjectivity, and difficult to scale.

This thesis focuses on the comparative evaluation of modern deep learning architectures for automatic brain tumor segmentation. A complete experimental pipeline was developed, including preprocessing and normalization of MRI volumes, patch-based training with stochastic augmentations, and evaluation using the Dice score. Four different architectures were compared under a unified pipeline: UNet as the established baseline, VNet with three-dimensional convolutions, SegResNet with extensive residual blocks, and SwinUNETR, a CNN/Transformer hybrid with self-attention.

The results showed that all models successfully converged, albeit with different characteristics: UNet offered an excellent balance of simplicity and accuracy, SegResNet demonstrated improved stability, VNet fully exploited the three-dimensional context but exhibited higher variance, while SwinUNETR achieved the overall best performance. The study concludes that hybrid architectures with attention mechanisms show strong potential for future research, without undermining the practical value of simpler networks such as U-Net.

This work contributes to a deeper understanding of the strengths and limitations of different segmentation methodologies and provides a basis for further research involving multi-class approaches, larger datasets, and optimized computational efficiency, aiming to bring such methods closer to routine clinical practice.

Keywords: MRI, Brain tumors, Image segmentation, Deep learning, Convolutional Neural Networks, U-Net, V-Net, SegResNet, SwinUNETR.

Η σελίδα αυτή είναι σκόπιμα λευκή.

Ευχαριστίες

Θα ήθελα να εκφράσω τις ευχαριστίες μου στον επιβλέποντα καθηγητή, κ. Γεώργιο Ματσόπουλο, καθώς και στη συνεπιβλέπουσα καθηγήτρια, Δρ. Ουρανία Πετροπούλου (Ε. ΔΙ. Π. Ε. Μ. Π.), για την εποπτεία της παρούσας εργασίας και τη συμμετοχή τους στη διαδικασία αξιολόγησης. Ευχαριστώ επίσης τους καθηγητές κ. Αθανάσιο Παναγόπουλο και κ. Παναγιώτη Τσανάκα για τη συμμετοχή τους στην εξεταστική επιτροπή.

Ιδιαίτερα, θα ήθελα να ευχαριστήσω τον υποψήφιο διδάκτορα κ. Κωνσταντίνο Γεώργα, ο οποίος με καθοδήγησε κατά την εκπόνηση της εργασίας και συνέβαλε αποφασιστικά τόσο στην υλοποίηση όσο και στη συγγραφή της.

Τέλος, εκφράζω την ευγνωμοσύνη μου προς την οικογένεια και τους φίλους μου για τη διαρκή στήριξη και ενθάρρυνση σε όλη τη διάρκεια των σπουδών μου.

Αθήνα, Σεπτέμβριος 2025

Ορέστης Διακόπουλος

Περιεχόμενα

Περιεχόμενα	1
Ευρετήριο εικόνων	4
Ευρετήριο πινάκων	5
1 Καρκίνος	6
1.1 Γενικά	6
1.2 Καρκίνος του εγκεφάλου	6
1.3 Αιτιολογία και παράγοντες κινδύνου	7
1.4 Συμπτωματολογία	8
1.5 Διάγνωση	9
1.6 Θεραπευτικές προσεγγίσεις	10
2 Μαγνητική Τομογραφία	12
2.1 Γενικά	12
2.2 Ιστορική εξέλιξη της MRI	13
2.3 Πυρηνικός Μαγνητικός Συντονισμός	13
2.4 RF Συντονισμός και χρόνοι χαλάρωσης	14
2.5 Χωρικός εντοπισμός του σήματος	16
2.6 Ανακατασκευή εικόνας	17
2.6.1 K-χώρος και ΔMT (FFT)	17
2.6.2 Compressed Sensing & επαναληπτική ανακατασκευή	17
2.6.3 Παράλληλη απεικόνιση (SENSE/GRAPPA)	18
2.6.4 Διαδρομές κ-χώρου, κίνηση και αλλοίωση	18
2.7 Προηγμένες τεχνικές MRI	18
2.8 MRI και διάγνωση όγκων εγκεφάλου	19
2.8.1 Συνηθισμένα τεχνουργήματα και μετριάσμος	20
2.8.2 Προεπεξεργασία δεδομένων για μάθηση και τμηματοποίηση	21
2.8.3 BraTS 2020	21
2.9 Ασφάλεια και σκιαγραφικά	21
2.10 Σύνοψη κεφαλαίου	22
3 Νευρωνικά Δίκτυα	24
3.1 Γενικά	24
3.2 Βιολογικοί νευρώνες	24

3.3	Διαισθητική ανάλυση	26
3.3.1	Μοντελοποίηση προβλήματος	26
3.3.2	Backpropagation: μια αριθμητική ματιά	27
3.4	Φορμαλιστική ανάλυση	27
3.4.1	Μοντέλο νευρώνα	27
3.4.2	Perceptron πολλών επιπέδων (MLP)	28
3.4.3	Gradient Descent	28
3.4.4	Backpropagation	28
3.5	Σύνοψη κεφαλαίου	29
4	Παρουσίαση των Μοντέλων που Συγκρίνουμε	30
4.1	Συνοπτική θεωρία CNN: <i>kernel, padding, pooling</i>	30
4.2	U Net: U-shape & <i>skip connections</i>	33
4.3	V Net: 3D <i>volumetric</i> συνελίξεις	34
4.4	SegResNet: βαθιά <i>residual encoder-decoder</i>	36
4.5	SwinUNETR: υβρίδιο CNN/Transformer με <i>window attention</i>	37
4.6	Παραλλαγές και επεκτάσεις U-Net	38
4.7	Εκπαιδευτικές τεχνικές και αρχιτεκτονικά βοηθήματα	39
4.8	Σύγκριση πολυπλοκότητας και σύνδεση με πειράματα	41
4.9	Σύνοψη Κεφαλαίου	42
5	Υλικά & Μέθοδοι	43
5.1	Σύνολα δεδομένων & διαχωρισμός	43
5.2	Προεπεξεργασία & φόρτωση δεδομένων	44
5.2.1	Δειγματοληψία patches & επαυξήσεις	44
5.2.2	Σημασία βημάτων προεπεξεργασίας	45
5.3	Αρχιτεκτονικές μοντέλων	46
5.4	Συναρτήσεις κόστους, βελτιστοποίηση & κανονικοποίηση	47
5.5	Εκπαίδευση & αξιολόγηση	48
5.5.1	Εναλλακτικές μετρικές αξιολόγησης	48
5.5.2	Σχήμα εκπαίδευσης	48
5.5.3	Συμπερασματοδότηση με ολισθαίνον παράθυρο	49
5.6	Πλέγμα πειραμάτων & ορισμός συγκρίσεων	49
5.7	Αναπαραγωγικότητα	50
5.8	Συνοπτικός ψευδοκώδικας πειράματος	51
5.9	Αιτιολόγηση σχεδιαστικών επιλογών	51
5.10	Περιορισμοί & έλεγχοι εγκυρότητας	52
5.11	Υπολογιστικό υλικό/λογισμικό	53
5.12	Σύνοψη Κεφαλαίου	53

6	Αποτελέσματα	54
6.1	Οδηγός ανάγνωσης	54
6.2	Καμπύλες εκπαίδευσης ανά αρχιτεκτονική	54
6.3	Συγκεντρωτικά αποτελέσματα	57
6.4	Ποιοτική αξιολόγηση (οπτικά παραδείγματα)	57
6.5	Συγκριτική κριτική με βάση τη βιβλιογραφία	59
6.6	Συζήτηση & ερμηνεία	61
6.7	Σύνοψη Κεφαλαίου	62
7	Συμπεράσματα & Μελλοντική Εργασία	64
7.1	Συνοπτική ανακεφαλαίωση	64
7.2	Βασικά συμπεράσματα	65
7.3	Περιορισμοί	66
7.4	Μελλοντική εργασία	67
	Βιβλιογραφία	70

Ευρετήριο εικόνων

3.1	Δομή βιολογικού νευρώνα: δενδρίτες (είσοδοι), κυτταρικό σώμα, νευράξονας και νευρικές απολήξεις [18].	25
3.2	Γραμμικό μοντέλο με 4 εισόδους και 21 εξόδο. Κάθε σύνδεση αντιστοιχεί σε ένα βάρος.	26
4.1	Δημοφιλείς επιλογές σε k , p , s και <i>upsampling</i> με <i>transposed conv</i> . Πηγή: Dumoulin & Visin (MIT).[25] Επιπλέον <i>max pooling</i> στο Σχ. 4.2.	32
4.2	Παράδειγμα <i>max pooling</i> 2×2 , $s=2$. Πηγή: Aphex34, Wikimedia Commons, CC BY-SA 4.0.[26] .	33
4.3	U-Net: συμμετρικός <i>encoder-decoder</i> με <i>skip connections</i> . Πηγή: Mehrdad Yazdani, CC BY-SA 4.0 [27].	34
4.4	V-Net: 3D <i>residual</i> μπλοκ και ιεραρχική χωρική κλιμάκωση. Image © NVIDIA, used under the Apache License 2.0 [29].	35
4.5	SegResNet: <i>encoder-decoder</i> με <i>residual</i> μπλοκ και (όπου εφαρμόζεται) <i>deep supervision</i> . Πηγή: Lee <i>et al.</i> , <i>Scientific Reports</i> , 2024, Σχ. 2, άδεια CC BY 4.0 [30].	37
4.6	SwinUNETR: <i>Transformer</i> -βασισμένος <i>encoder</i> με <i>window attention</i> και συνελκτικός <i>decoder</i> . Πηγή: Project MONAI, Apache-2.0 [31].	38
6.1	Καμπύλες εκπαίδευσης για UNet με Dice + AdamW. Αριστερά: Loss (train/val). Δεξιά: Dice (train/val).	55
6.2	Καμπύλες εκπαίδευσης για VNet με Dice + AdamW.	55
6.3	Καμπύλες εκπαίδευσης για SegResNet με Dice + AdamW.	56
6.4	Καμπύλες εκπαίδευσης για SwinUNETR με Dice + AdamW.	56
6.5	Example 1: index=12, z=72. Καλή οριοθέτηση όγκου από SegResNet και SwinUNETR, με την U-Net να εμφανίζει ελαφρά υποεκτίμηση σε περιφερειακές περιοχές.	58
6.6	Example 2: index=20, z=149. Η VNet δείχνει ασυνέχειες στα όρια, ενώ η SwinUNETR αποδίδει πιο ομαλή τμηματοποίηση.	58
6.7	Example 3: index=19, z=152. Όλα τα μοντέλα εντοπίζουν τον κύριο όγκο, αλλά μόνο SegResNet και SwinUNETR καταφέρνουν να διατηρήσουν τη λεπτομέρεια σε μικρότερες εστίες.	58
6.8	Example 4: index=24, z=80. Εδώ η U-Net παρουσιάζει ήπιο over-segmentation, ενώ η VNet χάνει τμήμα του πραγματικού όγκου.	58
6.9	Example 5: index=15, z=145. Η SwinUNETR αποδίδει την πιο «καθαρή» μάσκα, με ομαλή προσαρμογή στο ground truth.	58

Ευρετήριο πινάκων

4.1	Σύγκριση μοντέλων ως προς παραμέτρους και θεωρητικά FLOPs ανά δείγμα.	41
5.1	Σύνοψη υπερπαραμέτρων ανά αρχιτεκτονική.	46
6.1	Σύνοψη απόδοσης ανά μοντέλο (Dice + AdamW).	57

Κεφάλαιο 1: Καρκίνος

1.1 Γενικά

Ο όρος *καρκίνος* χρησιμοποιείται για να περιγράψει μια ευρεία ομάδα νοσημάτων, τα οποία χαρακτηρίζονται από την απώλεια του φυσιολογικού ελέγχου του κυτταρικού κύκλου και τη διατάραξη της ομοιόστασης του οργανισμού. Στον φυσιολογικό οργανισμό, η ανάπτυξη, διαφοροποίηση και απόπτωση των κυττάρων διέπονται από ένα αυστηρά ρυθμισμένο δίκτυο σημάτων, το οποίο διασφαλίζει την ισορροπία ανάμεσα στη δημιουργία νέων κυττάρων και στην απομάκρυνση των παλαιότερων ή κατεστραμμένων. Όταν αυτό το δίκτυο απορρυθμιστεί, λόγω γενετικών ή επιγενετικών αλλοιώσεων, τα κύτταρα αποκτούν ικανότητες όπως ανεξέλεγκτος πολλαπλασιασμός, αποφυγή των μηχανισμών προγραμματισμένου κυτταρικού θανάτου (απόπτωση), καθώς και δυνατότητα διήθησης σε γειτονικούς ιστούς [1].

Η συσσώρευση τέτοιων κυττάρων οδηγεί στη δημιουργία μιας μάζας, γνωστής ως *όγκος*. Οι όγκοι μπορούν να ταξινομηθούν σε δύο μεγάλες κατηγορίες: *καλοήθεις* και *κακοήθεις*. Οι καλοήθεις όγκοι συνήθως δεν εισβάλλουν σε γειτονικούς ιστούς ούτε δίνουν μεταστάσεις, αν και ενδέχεται να προκαλέσουν σοβαρά προβλήματα λόγω μηχανικής πίεσης σε ζωτικά όργανα ή αγγεία, ανάλογα με τη θέση τους. Αντιθέτως, οι κακοήθεις όγκοι χαρακτηρίζονται από την ικανότητά τους να διηθούν και να καταστρέφουν τον περιβάλλοντα ιστό και, μέσω πολύπλοκων βιολογικών μηχανισμών, να διασπείρονται απομακρυσμένα με το αίμα ή το λεμφικό σύστημα, δημιουργώντας δευτερογενείς εστίες – μια διαδικασία που είναι γνωστή ως *μετάσταση* [2]. Η μετάσταση αποτελεί το κύριο γνώρισμα που διαφοροποιεί τον καρκίνο από τις λιγότερο επικίνδυνες μορφές νεοπλασίας και είναι ο κυριότερος παράγοντας θνητότητας.

Σε παγκόσμια κλίμακα, ο καρκίνος συνιστά μία από τις σημαντικότερες αιτίες θανάτου και νοσηρότητας. Σύμφωνα με τον Παγκόσμιο Οργανισμό Υγείας, καταγράφονται κάθε χρόνο δεκάδες εκατομμύρια νέες διαγνώσεις και εκατομμύρια θάνατοι που σχετίζονται με καρκίνο, με την επίπτωση να αυξάνεται διαρκώς λόγω της γήρανσης του πληθυσμού αλλά και της έκθεσης σε περιβαλλοντικούς παράγοντες κινδύνου, όπως το κάπνισμα, η κακή διατροφή, η ατμοσφαιρική ρύπανση και η καθιστική ζωή [3]. Εκτός από τις προφανείς υγειονομικές συνέπειες, η ασθένεια έχει και σημαντικές κοινωνικοοικονομικές επιπτώσεις: επιβαρύνει τα συστήματα υγείας, μειώνει την παραγωγικότητα και δημιουργεί αυξημένες ανάγκες για μακροχρόνια φροντίδα και υποστήριξη.

Ιδιαίτερο ενδιαφέρον παρουσιάζει η μελέτη του καρκίνου όχι μόνο ως ιατρικό πρόβλημα, αλλά και ως βιολογικό φαινόμενο. Ο καρκίνος θεωρείται, κατά μία έννοια, μια *εξελικτική διαδικασία σε μικροκλίμακα*, καθώς τα κύτταρα που υφίστανται γενετικές αλλοιώσεις και αποκτούν πλεονεκτήματα επιβίωσης ή ανάπτυξης «επιλέγονται» και κυριαρχούν έναντι των φυσιολογικών. Η οπτική αυτή έχει αλλάξει τον τρόπο με τον οποίο προσεγγίζονται τόσο η βασική έρευνα όσο και οι θεραπευτικές στρατηγικές.

1.2 Καρκίνος του εγκεφάλου

Οι όγκοι που αναπτύσσονται στο κεντρικό νευρικό σύστημα (ΚΝΣ) αποτελούν μια ετερογενή ομάδα νοσημάτων με ιδιαίτερα υψηλή κλινική σημασία. Διακρίνονται καταρχάς σε δύο κύριες κατηγορίες: τους *πρωτοπαθείς* και τους *δευτεροπαθείς* όγκους [4]. Οι πρωτοπαθείς όγκοι αναπτύσσονται απευθείας από τα κύτταρα του εγκεφάλου, των μηνίγγων ή του νωτιαίου μυελού. Ενδεικτικά παραδείγματα είναι τα γλοιώματα (αστροκυττώματα, ολιγοδενδρογλοιώματα, γλοιοβλαστώματα), τα μηνιγγιώματα και οι όγκοι από τα κύτταρα Schwann. Οι δευτεροπαθείς ή μεταστατικοί όγκοι, αντίθετα, προέρχονται από κακοήθεις

άλλων οργάνων – όπως ο πνεύμονας, ο μαστός ή το μελάνωμα – και φθάνουν στον εγκέφαλο μέσω αιματογενούς διασποράς. Στην κλινική πράξη, οι μεταστατικοί όγκοι είναι συχνότεροι από τους πρωτοπαθείς, γεγονός που υπογραμμίζει τη σημασία της συστηματικής ογκολογικής προσέγγισης.

Η ιδιαίτερη βαρύτητα των όγκων του ΚΝΣ σχετίζεται με την κρίσιμη λειτουργία του εγκεφάλου: αποτελεί το κέντρο ελέγχου αισθητικών, κινητικών, γνωσιακών και συμπεριφορικών διεργασιών. Ακόμη και μικρού μεγέθους βλάβες ή τοπικό οίδημα μπορούν να οδηγήσουν σε σημαντικά νευρολογικά ελλείμματα, όπως παράλυση, απώλεια ομιλίας, διαταραχές μνήμης ή αλλαγές στη συμπεριφορά. Η συμπτωματολογία ποικίλλει ανάλογα με την ανατομική εντόπιση: για παράδειγμα, βλάβες στον μετωπιαίο λοβό συχνά επηρεάζουν την εκτελεστική λειτουργία και την προσωπικότητα, ενώ όγκοι στην παρεγκεφαλίδα προκαλούν αστάθεια και διαταραχές ισορροπίας.

Η θεραπευτική αντιμετώπιση είναι ιδιαίτερα απαιτητική. Η ριζική χειρουργική εξαίρεση, που σε άλλες εντοπίσεις μπορεί να αποτελεί την κύρια θεραπεία, συχνά δεν είναι εφικτή στον εγκέφαλο λόγω της εγγύτητας των όγκων σε ζωτικές ή «λειτουργικά εύθραυστες» περιοχές. Η πλήρης αφαίρεση μπορεί να επιφέρει σοβαρά λειτουργικά ελλείμματα, με αποτέλεσμα η χειρουργική στρατηγική να περιορίζεται συχνά σε *μέγιστη ασφαλή εκτομή*, δηλαδή αφαίρεση όσο το δυνατόν μεγαλύτερου μέρους του όγκου χωρίς μη αναστρέψιμες βλάβες [5]. Συμπληρωματικά, εφαρμόζονται ακτινοθεραπεία, χημειοθεραπεία ή στοχευμένες μοριακές θεραπείες, ανάλογα με τον ιστολογικό τύπο και το μοριακό προφίλ του όγκου.

Η έρευνα γύρω από τους όγκους του εγκεφάλου έχει ενταθεί τα τελευταία χρόνια, με έμφαση στη μοριακή ταξινόμηση, η οποία παρέχει σημαντική προγνωστική και θεραπευτική πληροφόρηση. Η νέα ταξινόμηση του Παγκόσμιου Οργανισμού Υγείας (WHO, 2021) ενσωματώνει γενετικούς και επιγενετικούς δείκτες, καθιστώντας πιο ακριβή την κατηγοριοποίηση των όγκων και επιτρέποντας την ανάπτυξη στοχευμένων θεραπειών. Παρά τις εξελίξεις, ωστόσο, η πρόγνωση πολλών μορφών – και κυρίως του γλοιοβλαστώματος – παραμένει δυσμενής, γεγονός που υπογραμμίζει την ανάγκη για περαιτέρω έρευνα και καινοτόμες θεραπευτικές προσεγγίσεις.

1.3 Αιτιολογία και παράγοντες κινδύνου

Η διαδικασία της καρκινογένεσης είναι εξαιρετικά πολύπλοκη και πολυπαραγοντική, καθώς εμπλέκει έναν συνδυασμό γενετικών, περιβαλλοντικών και συμπεριφορικών παραγόντων. Κεντρικό ρόλο παίζουν οι γενετικές αλλοιώσεις, οι οποίες μπορούν να είναι είτε κληρονομούμενες είτε επίκτητες. Οι κληρονομούμενες μεταλλάξεις σε κρίσιμα γονίδια, όπως τα *ογκοκατασταλτικά* ή τα *ογκογονίδια*, δημιουργούν μια προδιάθεση που καθιστά το κύτταρο ευάλωτο στην ανάπτυξη κακοήθειας. Αντίστοιχα, οι σωματικές μεταλλάξεις που αποκτώνται κατά τη διάρκεια της ζωής μπορεί να οφείλονται σε περιβαλλοντικές εκθέσεις (ιονίζουσα ακτινοβολία, χημικά καρκινογόνα, ρύπανση) ή σε τυχαία σφάλματα αντιγραφής του DNA [6]. Σε κάθε περίπτωση, το αποτέλεσμα είναι η σταδιακή απώλεια ελέγχου του κυτταρικού κύκλου και η συσσώρευση γενετικών βλαβών.

Οι παράγοντες τρόπου ζωής αποτελούν επίσης καθοριστική συνιστώσα της καρκινογένεσης. Το κάπνισμα, για παράδειγμα, έχει αναγνωριστεί ως η κυριότερη αιτία καρκίνου του πνεύμονα, αλλά αυξάνει επίσης τον κίνδυνο και για άλλες μορφές καρκίνου, λόγω της έκθεσης σε πλήθος καρκινογόνων ουσιών. Η υπερβολική κατανάλωση αλκοόλ, η παχυσαρκία και η έλλειψη φυσικής δραστηριότητας έχουν αποδειχθεί ότι ενισχύουν την καρκινογένεση μέσω πολύπλοκων μηχανισμών που σχετίζονται με τη φλεγμονή, τις ορμονικές διαταραχές και το οξειδωτικό στρες. Η διατροφή, τέλος, μπορεί να λειτουργήσει είτε προστατευτικά είτε επιβαρυντικά, ανάλογα με τη σύσταση και τις διατροφικές συνήθειες.

Η ηλικία αποτελεί από μόνη της έναν σημαντικό παράγοντα κινδύνου. Καθώς ο οργανισμός μεγαλώ-

νει, αυξάνεται η πιθανότητα εμφάνισης σωματικών μεταλλάξεων εξαιτίας της συνεχούς αντιγραφής του DNA και της φθοράς των μηχανισμών επιδιόρθωσης. Έτσι, η συσσώρευση γενετικών βλαβών καθιστά τους ηλικιωμένους πιο ευάλωτους στην ανάπτυξη κακοηθειών. Το γεγονός αυτό εξηγεί εν μέρει γιατί η επίπτωση του καρκίνου αυξάνει δραματικά στις μεγαλύτερες ηλικιακές ομάδες.

Σε ορισμένες μορφές καρκίνου, έχει αποδειχθεί ισχυρή γενετική προδιάθεση. Ένα χαρακτηριστικό παράδειγμα είναι οι μεταλλάξεις στα γονίδια *BRCA1* και *BRCA2*, οι οποίες σχετίζονται με σημαντικά αυξημένο κίνδυνο καρκίνου του μαστού και των ωοθηκών [7]. Παρόμοια, σε άλλα σύνδρομα, όπως το σύνδρομο Li-Fraumeni, παρατηρούνται μεταλλάξεις στο γονίδιο *TP53*, που οδηγούν σε αυξημένη συχνότητα ποικίλων νεοπλασιών, συμπεριλαμβανομένων και εγκεφαλικών όγκων.

Όσον αφορά ειδικότερα τους όγκους του εγκεφάλου, η αιτιολογία τους παραμένει σε μεγάλο βαθμό ασαφής και πιθανόν διαφοροποιείται ανάλογα με τον ιστολογικό τύπο. Παρά ταύτα, ορισμένοι παράγοντες έχουν αναγνωριστεί ως αυξημένου κινδύνου. Έχει παρατηρηθεί ότι ασθενείς που έχουν υποβληθεί σε ακτινοθεραπεία στην περιοχή της κεφαλής ή του τραχήλου, για άλλες καλοήθειες ή κακοήθειες παθήσεις, εμφανίζουν αυξημένη πιθανότητα ανάπτυξης δευτερογενών νεοπλασιών στον εγκέφαλο. Ο κίνδυνος αυτός δεν είναι ο ίδιος για όλους, αλλά εξαρτάται από την αθροιστική δόση ακτινοβολίας, τη διάρκεια της έκθεσης, καθώς και από γενετικούς και περιβαλλοντικούς τροποποιητικούς παράγοντες [8]. Επιπλέον, ερευνώνται πιθανές συσχετίσεις με περιβαλλοντικές εκθέσεις (π.χ. χρήση κινητών τηλεφώνων, επαγγελματική έκθεση σε χημικές ουσίες), αν και τα μέχρι στιγμής δεδομένα είναι αντιφατικά και δεν επιτρέπουν σαφή συμπεράσματα.

1.4 Συμπτωματολογία

Η συμπτωματολογία του καρκίνου είναι εξαιρετικά ποικίλη και εξαρτάται τόσο από τον ιστολογικό τύπο και τη θέση του όγκου, όσο και από τις συστηματικές επιδράσεις του στον οργανισμό. Σε γενικές γραμμές, τα συμπτώματα διακρίνονται σε *γενικά*, *μη ειδικά* και σε *ειδικά*, τα οποία σχετίζονται με την εντόπιση και τη βιολογία της εκάστοτε νεοπλασίας.

Γενικά, μη ειδικά συμπτώματα

Τα γενικά συμπτώματα προκαλούνται κυρίως από τις μεταβολικές και ανοσολογικές επιδράσεις του όγκου στο σύνολο του οργανισμού. Συχνό εύρημα είναι η *ανεξήγητη απώλεια βάρους*, η οποία μπορεί να ξεπεράσει το 10% του σωματικού βάρους σε διάστημα λίγων μηνών. Το φαινόμενο αυτό οφείλεται σε αυξημένο καταβολισμό, απώλεια όρεξης και διαταραχή της φυσιολογικής χρήσης των θρεπτικών ουσιών, καταστάσεις που συνδυάζονται στον όρο *καχεξία του καρκίνου*. Άλλα συστηματικά συμπτώματα περιλαμβάνουν *παρατεταμένο πυρετό*, που δεν αποδίδεται σε λοιμώδη αίτια, καθώς και το *καταβολικό σύνδρομο*, το οποίο χαρακτηρίζεται από έντονη αδυναμία, εύκολη κόπωση και μειωμένη αντοχή στις καθημερινές δραστηριότητες. Τα φαινόμενα αυτά αποδίδονται στην απελευθέρωση προφλεγμονωδών κυτταροκινών (π.χ. TNF-α, IL-1, IL-6) από τον όγκο ή το ανοσοποιητικό σύστημα, γεγονός που οδηγεί σε μια κατάσταση χρόνιας φλεγμονής.

Συχνές επίσης είναι οι *μεταβολές του δέρματος*, όπως ίκτερος (κιτρίνισμα), υπερμελάγχρωση, ή η αιφνίδια εμφάνιση και ταχεία ανάπτυξη νέων σπύλων, που ενδέχεται να αποτελούν ένδειξη κακοήθους μελανώματος. Οι *αιμορραγικές εκδηλώσεις* (ανεξήγητοι μώλωπες, επίσταξη, αιμορραγία από το γαστρεντερικό ή το ουροποιητικό) μπορεί να προέρχονται είτε από τοπική διήθηση είτε από διαταραχή των μηχανισμών πήξης. Τέλος, *διαταραχές της σίτισης και της αφόδευσης*, όπως δυσφαγία, αλλαγές στις εντερικές συνήθειες ή συχνουρία, συχνά αποτελούν πρώιμες ενδείξεις όγκων του πεπτικού ή του ουροποιητικού συστήματος [3].

Ειδικά συμπτώματα εγκεφαλικών όγκων

Στην περίπτωση των όγκων του εγκεφάλου, τα συμπτώματα εξαρτώνται σε μεγάλο βαθμό από το μέγεθος, τον ρυθμό ανάπτυξης και την ακριβή ανατομική εντόπιση. Καθώς ο εγκέφαλος βρίσκεται σε περιορισμένο χώρο μέσα στο κρανίο, ακόμη και μικρές βλάβες μπορούν να προκαλέσουν σημαντική αύξηση της ενδοκρανιακής πίεσης. Τυπικές εκδηλώσεις είναι οι *πρωινοί πονοκέφαλοι*, που συχνά συνοδεύονται από ναυτία ή έμετο, λόγω της αυξημένης πίεσης στο εγκεφαλικό στέλεχος και στα κέντρα του εμέτου. Οι *επιληπτικές κρίσεις* αποτελούν επίσης συχνό αρχικό σύμπτωμα, ιδίως σε ασθενείς με φλοιϊκές βλάβες.

Ανάλογα με την περιοχή εντόπισης, μπορεί να εμφανιστούν *εστιακά νευρολογικά ελλείμματα*:

- όγκοι του μετωπιαίου λοβού μπορεί να προκαλέσουν διαταραχές της προσωπικότητας, απώλεια αναστολών ή δυσκολίες στην εκτέλεση πολύπλοκων κινήσεων,
- όγκοι του ινιακού λοβού σχετίζονται με διαταραχές όρασης,
- βλάβες στην παρεγκεφαλίδα οδηγούν σε αστάθεια, ζάλη και διαταραχές ισορροπίας,
- βλάβες στον κροταφικό λοβό μπορεί να προκαλέσουν διαταραχές μνήμης ή αφασία.

Επιπλέον, μπορεί να παρατηρηθούν *αλλαγές στη συμπεριφορά ή γνωσιακή έκπτωση*, που γίνονται συχνά αντιληπτές από το οικογενειακό περιβάλλον του ασθενούς πριν από τον ίδιο. Σε προχωρημένες περιπτώσεις, η συνύπαρξη πολλαπλών συμπτωμάτων ή η σταδιακή επιδείνωσή τους συνιστούν σαφή ένδειξη για άμεσο νευροαπεικονιστικό έλεγχο (συνήθως μαγνητική τομογραφία), ώστε να τεθεί έγκαιρα η διάγνωση [9].

Διαφορική διάγνωση

Πολλά από τα παραπάνω συμπτώματα δεν είναι αποκλειστικά του καρκίνου, αλλά μπορεί να εμφανιστούν και σε άλλες παθήσεις (π.χ. ημικρανίες, αγγειακά εγκεφαλικά επεισόδια, λοιμώξεις, αυτοάνοσα νοσήματα). Αυτό καθιστά τη διαφορική διάγνωση ιδιαίτερα απαιτητική και υπογραμμίζει τη σημασία του ολοκληρωμένου κλινικού και απεικονιστικού ελέγχου.

1.5 Διάγνωση

Η διάγνωση των εγκεφαλικών όγκων απαιτεί πολυπαραγοντική προσέγγιση που συνδυάζει την κλινική αξιολόγηση, τις νευρολογικές δοκιμασίες και τις σύγχρονες απεικονιστικές τεχνικές. Το πρώτο βήμα είναι η λεπτομερής *κλινική και νευρολογική εξέταση*, μέσω της οποίας ο ιατρός εκτιμά τυχόν εστιακά νευρολογικά ελλείμματα, αλλαγές στη συμπεριφορά, στη μνήμη ή στη γνωστική λειτουργία, καθώς και σημεία αυξημένης ενδοκρανιακής πίεσης. Η αρχική αυτή εκτίμηση καθοδηγεί την επιλογή της απεικονιστικής μεθόδου.

Η *μαγνητική τομογραφία (MRI)* αποτελεί τη μέθοδο εκλογής για την απεικόνιση εγκεφαλικών όγκων, χάρη στην υψηλή ευαισθησία της, την εξαιρετική ικανότητα διάκρισης των μαλακών ιστών και την απουσία ιονίζουσας ακτινοβολίας [10]. Η MRI επιτρέπει την τρισδιάστατη απεικόνιση του εγκεφάλου σε πολλαπλά επίπεδα και την αξιολόγηση κρίσιμων παραμέτρων, όπως το οίδημα, οι νεκρωτικές περιοχές, η διήθηση και η αγγείωση του όγκου. Διαφορετικές ακολουθίες (*modalities*) προσφέρουν συμπληρωματικές πληροφορίες:

- Η *T1-weighted* ακολουθία παρέχει λεπτομέρειες για την ανατομία και τη δομή του όγκου, ενώ

μετά από χορήγηση σκιαγραφικού υλικού (gadolinium) αποκαλύπτει περιοχές ενεργής αγγείωσης και διαταραχής του αιματοεγκεφαλικού φραγμού.

- Η *T2-weighted* και η *FLAIR* ακολουθία αναδεικνύουν την παρουσία οιδήματος, κυστικών περιοχών και την έκταση της διήθησης.
- Προχωρημένες τεχνικές, όπως η *diffusion-weighted imaging (DWI)* και η *perfusion-weighted imaging (PWI)*, προσφέρουν πληροφορίες για τη μικροδομή και την αιμάτωση του όγκου αντίστοιχα.
- Η *MR spectroscopy (MRS)* επιτρέπει τη μεταβολική ανάλυση του ιστού, βοηθώντας στη διάκριση κακοήθων από καλοήθεις βλάβες ή νεκρωτικό ιστό από υποτροπή.

Αν και η MRI είναι η κύρια μέθοδος, και άλλες τεχνικές χρησιμοποιούνται σε ειδικές περιπτώσεις. Η αξονική τομογραφία (CT) είναι ταχύτερη και χρήσιμη σε επείγουσες καταστάσεις, όπως η ανίχνευση αιμορραγίας ή αποτιτανώσεων. Η τομογραφία εκπομπής ποζιτρονίων (PET) και η υβριδική PET/MRI μπορούν να δώσουν πληροφορίες για τη μεταβολική δραστηριότητα και τη βιολογική συμπεριφορά των όγκων, αν και η χρήση τους είναι περιορισμένη από το υψηλό κόστος και τη διαθεσιμότητα.

Η οριστική διάγνωση και η ακριβής ταξινόμηση απαιτούν *ιστολογική και μοριακή ανάλυση*, που γίνεται μέσω βιοψίας ή μετά από χειρουργική εκτομή. Η βιοψία παρέχει ιστικό δείγμα για μικροσκοπική και γενετική εξέταση, επιτρέποντας τον καθορισμό του τύπου και του βαθμού κακοήθειας σύμφωνα με τα κριτήρια του Παγκόσμιου Οργανισμού Υγείας (WHO). Η ενσωμάτωση μοριακών δεικτών (π.χ. μεταλλάξεις στο *IDH*, συνεκφώνηση *1p/19q*, μεθυλίωση προαγωγού *MGMT*) έχει βελτιώσει σημαντικά την προγνωστική ακρίβεια και την καθοδήγηση θεραπευτικών αποφάσεων.

Ένα κρίσιμο στάδιο στη διαγνωστική διαδικασία είναι η *τμηματοποίηση* των βλαβών, δηλαδή η ακριβής απομόνωση και περιγραφή της έκτασης του όγκου σε εικόνες MRI. Παραδοσιακά, η διαδικασία αυτή πραγματοποιείται χειροκίνητα ή ημιαυτόματα από εξειδικευμένους νευροακτινολόγους. Ωστόσο, η διαδικασία είναι χρονοβόρα και υπόκειται σε σημαντική δια- και ενδο-παρατηρητική μεταβλητότητα, καθώς διαφορετικοί παρατηρητές ή ακόμα και ο ίδιος παρατηρητής σε διαφορετικές χρονικές στιγμές μπορεί να αποδώσουν διαφορετικά αποτελέσματα [11]. Το γεγονός αυτό υπογραμμίζει την ανάγκη για ανάπτυξη υπολογιστικών μεθόδων υποβοήθησης της διάγνωσης, οι οποίες μπορούν να προσφέρουν πιο αντικειμενική και αναπαραγώγιμη εκτίμηση.

1.6 Θεραπευτικές προσεγγίσεις

Η θεραπευτική αντιμετώπιση των εγκεφαλικών όγκων είναι πολυδιάστατη και εξατομικευμένη, καθώς πρέπει να λάβει υπόψη τον ιστολογικό τύπο, το μοριακό προφίλ, την εντόπιση, το μέγεθος του όγκου, αλλά και την ηλικία και τη γενική κατάσταση του ασθενούς. Στις περισσότερες περιπτώσεις, το θεραπευτικό πλάνο συνδυάζει πολλαπλές μεθόδους, με στόχο τη μέγιστη δυνατή παράταση της επιβίωσης, τη διατήρηση της ποιότητας ζωής και την ελαχιστοποίηση των επιπλοκών [12].

Χειρουργική αντιμετώπιση

Η *χειρουργική εξαίρεση* αποτελεί συχνά το πρώτο βήμα στη θεραπευτική στρατηγική. Ο στόχος είναι η *μέγιστη ασφαλής εκτομή*, δηλαδή η αφαίρεση όσο το δυνατόν μεγαλύτερου μέρους του όγκου χωρίς πρόκληση μη αναστρέψιμων νευρολογικών βλαβών. Για τον σκοπό αυτό, χρησιμοποιούνται σύγχρονες τεχνικές, όπως η νευροπλοήγηση, η διεγχειρητική απεικόνιση με MRI ή υπέρηχο, καθώς και η διεγχειρητική χαρτογράφηση λειτουργιών (π.χ. αφύπνιση του ασθενούς για έλεγχο γλωσσικών και κινητικών

λειτουργιών). Παρά τις τεχνολογικές εξελίξεις, η πλήρης εξαίρεση παραμένει δύσκολη σε όγκους με εκτεταμένη διήθηση, όπως τα γλοιώματα.

Ακτινοθεραπεία

Η ακτινοθεραπεία αποτελεί βασικό θεραπευτικό όπλο, ιδιαίτερα μετά τη χειρουργική εκτομή, για την εξάλειψη μικροσκοπικών υπολειμματικών κυττάρων. Η σύγχρονη ακτινοθεραπεία χρησιμοποιεί τεχνικές υψηλής ακρίβειας όπως η *Intensity-Modulated Radiotherapy (IMRT)* και η *Volumetric Modulated Arc Therapy (VMAT)*, οι οποίες κατευθύνουν υψηλής ενέργειας δέσμες με τέτοιο τρόπο ώστε να επιτυγχάνεται μέγιστη δόση στον όγκο με ελαχιστοποίηση της βλάβης στους υγιείς ιστούς. Η χρήση πολυφυλλικών διαφραγμάτων επιτρέπει την ακριβή διαμόρφωση του πεδίου ακτινοβολήσης [13]. Η αποτελεσματικότητα της ακτινοθεραπείας εξαρτάται καθοριστικά από την ακριβή οριοθέτηση του όγκου και των κρίσιμων οργάνων, γεγονός που τονίζει τη σημασία της συνεργασίας νευροακτινολόγων και ακτινοθεραπευτών.

Συστηματική θεραπεία

Η συστηματική θεραπεία στοχεύει στην καταστροφή καρκινικών κυττάρων που μπορεί να έχουν διασπαρεί πέρα από την κύρια εστία του όγκου. Η πλέον καθιερωμένη συστηματική αγωγή στον καρκίνο του εγκεφάλου είναι η χημειοθεραπεία με τεμοζολομίδη. Η τεμοζολομίδη έχει αποδειχθεί ότι βελτιώνει την επιβίωση ασθενών με γλοιοβλάστωμα, ιδιαίτερα όταν συνδυάζεται με ακτινοθεραπεία, αποτελώντας το λεγόμενο πρωτόκολλο *Stupp* [12]. Η αποτελεσματικότητα της χημειοθεραπείας εξαρτάται από το μοριακό προφίλ του όγκου: για παράδειγμα, η μεθυλίωση του προαγωγέα του γονιδίου *MGMT* σχετίζεται με καλύτερη ανταπόκριση στην τεμοζολομίδη, ενώ οι μεταλλάξεις στο γονίδιο *IDH* έχουν προγνωστική και πιθανώς θεραπευτική αξία.

Νεότερες θεραπευτικές προσεγγίσεις

Πέρα από τις καθιερωμένες μεθόδους, τα τελευταία χρόνια έχουν αναπτυχθεί καινοτόμες στρατηγικές:

- **Στοχευμένες θεραπείες:** αξιοποιούν συγκεκριμένα μοριακά μονοπάτια του όγκου, με φάρμακα που αναστέλλουν επιλεγμένους υποδοχείς ή ένζυμα.
- **Ανοσοθεραπείες:** περιλαμβάνουν αναστολείς σημείων ελέγχου (checkpoint inhibitors), εμβόλια κατά όγκων και υιοθετική μεταφορά κυττάρων T (CAR-T cells). Αν και τα αποτελέσματα στους εγκεφαλικούς όγκους είναι ακόμη περιορισμένα, οι κλινικές μελέτες βρίσκονται σε εξέλιξη.
- **Πειραματικές τεχνικές:** όπως η ηλεκτρομαγνητική θεραπεία (tumor-treating fields, TTF) και η γονιδιακή θεραπεία, που διερευνώνται ως πρόσθετες ή εναλλακτικές επιλογές.

Ολιστική και πολυεπιστημονική προσέγγιση

Η αντιμετώπιση των εγκεφαλικών όγκων απαιτεί πολυεπιστημονική συνεργασία, που περιλαμβάνει νευροχειρουργούς, ακτινοθεραπευτές, ογκολόγους, νευροακτινολόγους, παθολογοανατόμους και ειδικούς στη φροντίδα υποστηρικτικής θεραπείας. Η ολιστική φροντίδα δεν περιορίζεται στη θεραπεία του όγκου, αλλά στοχεύει και στη βελτίωση της ποιότητας ζωής μέσω ανακουφιστικής αγωγής, ψυχολογικής υποστήριξης και αποκατάστασης.

Κεφάλαιο 2: Μαγνητική Τομογραφία

2.1 Γενικά

Η Μαγνητική Τομογραφία (Magnetic Resonance Imaging, MRI) αποτελεί σήμερα μία από τις σημαντικότερες μη-επεμβατικές¹ απεικονιστικές τεχνικές για την ανατομική, και σε ειδικά πρωτόκολλα, λειτουργική απεικόνιση του ανθρώπινου σώματος. Αναπτύχθηκε τη δεκαετία του 1970 και καθιερώθηκε κλινικά τη δεκαετία του 1980, προσφέροντας μια επαναστατική μέθοδο απεικόνισης που βασίζεται σε φυσικές αρχές διαφορετικές από εκείνες των ακτίνων X ή του υπερήχου. Η δυνατότητά της να παρέχει εικόνες υψηλής αντίθεσης μαλακών ιστών χωρίς έκθεση σε ακτινοβολία την έχει καταστήσει θεμελιώδες εργαλείο στη νευρολογία, στην ογκολογία, στην καρδιολογία και σε πλήθος άλλων ιατρικών ειδικοτήτων.

Η βασική αρχή λειτουργίας της MRI στηρίζεται στην ανίχνευση σήματος από τους πυρήνες υδρογόνου (πρωτόνια), οι οποίοι είναι ιδιαίτερα άφθονοι στο ανθρώπινο σώμα λόγω της υψηλής περιεκτικότητας σε νερό και λίπος. Υπό την επίδραση ενός ισχυρού εξωτερικού μαγνητικού πεδίου B_0 , τα πρωτόνια ευθυγραμμίζονται παράλληλα ή αντιπαράλληλα με το πεδίο. Στη συνέχεια, εφαρμόζεται παλμός ραδιοσυχνότητας (RF) που προκαλεί απόκλιση από την ισορροπία, και καθώς τα πρωτόνια επανέρχονται, εκπέμπεται σήμα το οποίο ανιχνεύεται από τα πηνία λήψης. Η καταγραφή και η ανάλυση αυτών των σημάτων, με χρήση πολύπλοκων μαθηματικών τεχνικών (Fourier Transform), επιτρέπουν την ανασύνθεση εικόνων υψηλής ποιότητας.

Ένα από τα μεγαλύτερα πλεονεκτήματα της MRI είναι η ευελιξία στη στάθμιση της εικόνας μέσω διαφορετικών παραμέτρων λήψης, όπως ο χρόνος επαναφοράς (TR), ο χρόνος ηχούς (TE) και ο χρόνος αναστροφής (TI). Έτσι, η ίδια ανατομική περιοχή μπορεί να απεικονιστεί με διαφορετικές αντιθέσεις (T1, T2, Proton Density, FLAIR), αναδεικνύοντας διαφορετικά βιολογικά χαρακτηριστικά. Για παράδειγμα, οι T1-σταθμισμένες εικόνες αναδεικνύουν τη μορφολογία και την ανατομία, ενώ οι T2-σταθμισμένες και FLAIR εικόνες είναι ιδιαίτερα χρήσιμες στην ανάδειξη παθολογικών αλλοιώσεων και οίδημάτων.

Σε σύγκριση με άλλες τεχνικές απεικόνισης, η MRI παρουσιάζει μοναδικά πλεονεκτήματα. Η αξονική τομογραφία (CT), αν και ταχύτερη και ευρύτερα διαθέσιμη, βασίζεται στην απορρόφηση ιοντίζουσας ακτινοβολίας και αποδίδει κυρίως πληροφορίες για την πυκνότητα των ιστών. Ο υπέρηχος παρέχει απεικόνιση σε πραγματικό χρόνο χωρίς ακτινοβολία, αλλά έχει περιορισμένη διείσδυση και χαμηλότερη ανάλυση σε βαθύτερες δομές. Η τομογραφία εκπομπής ποζιτρονίων (PET) προσφέρει μεταβολική πληροφορία αλλά με χαμηλότερη χωρική ανάλυση και σημαντικά υψηλότερο κόστος. Η MRI, αντίθετα, συνδυάζει υψηλή ανάλυση, άριστη αντίθεση μαλακών ιστών και δυνατότητα πολλαπλών *modalities*, καθιστώντας την ιδιαίτερα προσαρμόσιμη στις ανάγκες κάθε ασθενούς [34], [35], [39].

Η αύξηση της έντασης του μαγνητικού πεδίου B_0 βελτιώνει τον λόγο σήματος προς θόρυβο (SNR), γεγονός που επιτρέπει υψηλότερη χωρική ανάλυση και συντομότερους χρόνους σάρωσης. Στην κλινική πράξη χρησιμοποιούνται συνήθως συστήματα 1.5–3 Tesla, ενώ σε ερευνητικά κέντρα έχουν αναπτυχθεί συστήματα 7 Tesla και άνω. Ωστόσο, η αύξηση του πεδίου συνοδεύεται και από τεχνικές προκλήσεις, όπως αυξημένη ευαισθησία σε ανομοιογένειες του μαγνητικού πεδίου και φαινόμενα μαγνητικής επιδεκτικότητας [37], [39].

Επιπλέον, η χρήση πηνίων πολλαπλών στοιχείων λήψης (multi-channel coils) έχει βελτιώσει σημαντικά τόσο το SNR όσο και τη δυνατότητα εφαρμογής τεχνικών επιτάχυνσης (parallel imaging). Με

¹ Η MRI δεν χρησιμοποιεί ιοντίζουσα ακτινοβολία, σε αντίθεση με την ακτινογραφία και την αξονική τομογραφία.

αυτόν τον τρόπο μειώνεται ο συνολικός χρόνος σάρωσης, γεγονός που είναι ιδιαίτερα σημαντικό για την άνεση του ασθενούς και τη μείωση κινήσεων που μπορεί να αλλοιώσουν την ποιότητα της εικόνας.

2.2 Ιστορική εξέλιξη της MRI

Η ανάπτυξη της Μαγνητικής Τομογραφίας (MRI) αποτελεί χαρακτηριστικό παράδειγμα τομής μεταξύ θεμελιώδους φυσικής έρευνας και κλινικής εφαρμογής. Η αφετηρία της εντοπίζεται στην ανακάλυψη του *Πυρηνικού Μαγνητικού Συντονισμού* (Nuclear Magnetic Resonance, NMR) τη δεκαετία του 1940. Το 1946, οι Felix Bloch και Edward Purcell απέδειξαν ανεξάρτητα ότι οι πυρήνες υδρογόνου απορροφούν και εκπέμπουν ενέργεια όταν εκτεθούν σε ισχυρό μαγνητικό πεδίο και σε ραδιοσυχνότητες συγκεκριμένου εύρους [**bloch1946nuclear**, **purcell1946nuclear**]. Για την ανακάλυψή τους τιμήθηκαν με το βραβείο Nobel Φυσικής το 1952. Στα επόμενα χρόνια, η τεχνική NMR εδραιώθηκε ως εργαλείο χημικής ανάλυσης και μοριακής φασματοσκοπίας, χωρίς ακόμη να έχει κλινική διάσταση.

Το κρίσιμο βήμα για την ιατρική χρήση ήρθε τη δεκαετία του 1970. Ο Paul Lauterbur, το 1973, παρουσίασε την ιδέα ότι με κατάλληλη χωρική διαμόρφωση των μαγνητικών πεδίων μπορούν να δημιουργηθούν *εικόνες* με βάση το σήμα NMR [**lauterbur1973image**]. Εισήγαγε την έννοια της *διαφοροποίησης πεδίου* (magnetic field gradients), ώστε να επιτευχθεί χωρική κωδικοποίηση. Ανεξάρτητα, ο Peter Mansfield το 1977 ανέπτυξε μαθηματικές μεθόδους ταχείας ανακατασκευής εικόνας μέσω *Fourier Transform* και βελτίωσε δραστικά τον χρόνο σάρωσης [**mansfield1977multi**]. Για τις συνεισφορές τους, οι Lauterbur και Mansfield έλαβαν από κοινού το βραβείο Nobel Ιατρικής το 2003.

Ήδη από τα τέλη της δεκαετίας του 1970 κατασκευάστηκαν τα πρώτα πρωτότυπα μηχανήματα MRI. Οι πρώτες κλινικές εφαρμογές εστιάστηκαν στην απεικόνιση του εγκεφάλου, του νωτιαίου μυελού και αρθρώσεων, όπου η MRI πρόσφερε σαφές πλεονέκτημα έναντι της Υπολογιστικής Τομογραφίας (CT): *μη χρήση ιονίζουσας ακτινοβολίας* και υψηλή αντίθεση μεταξύ μαλακών ιστών. Κατά τη δεκαετία του 1980, με την εμπορική διάθεση των πρώτων συστημάτων από εταιρείες όπως η General Electric και η Siemens, η MRI καθιερώθηκε ως κλινικό εργαλείο. Σταδιακά, προστέθηκαν νέες ακολουθίες (spin-echo, gradient-echo, inversion recovery) που επέτρεψαν την απεικόνιση διαφορετικών βιολογικών παραμέτρων.

Στη δεκαετία του 1990 η πρόοδος επιταχύνθηκε με την εισαγωγή ισχυρότερων μαγνητών υπεραγωγιμότητας (1.5 Tesla και αργότερα 3 Tesla), καθώς και τεχνικών παράλληλης απεικόνισης. Οι εξελίξεις αυτές μείωσαν τον χρόνο λήψης και αύξησαν την ανάλυση, καθιστώντας την MRI αναντικατάστατη σε πλήθος κλινικών εφαρμογών: νευρολογία, καρδιολογία, μυοσκελετικές παθήσεις.

Σήμερα, η MRI θεωρείται μία από τις πιο ισχυρές και πολύπλευρες μεθόδους ιατρικής απεικόνισης. Χρησιμοποιείται καθημερινά σε διαγνωστικά κέντρα και νοσοκομεία παγκοσμίως, ενώ νέες τεχνικές όπως η λειτουργική MRI (fMRI), η απεικόνιση διάχυσης (DWI/DTI) και οι πολυπαραμετρικές εξετάσεις στην ογκολογία εμπλουτίζουν συνεχώς τις δυνατότητές της. Η ιστορική της διαδρομή –από την καθαρή φυσική του NMR έως την καθιέρωση ως χρυσού κανόνα σε πολλές κλινικές διαγνώσεις– υπογραμμίζει τη δύναμη της διεπιστημονικής έρευνας.

2.3 Πυρηνικός Μαγνητικός Συντονισμός

Η αρχή λειτουργίας της Μαγνητικής Τομογραφίας βασίζεται στο φαινόμενο του *Πυρηνικού Μαγνητικού Συντονισμού* (Nuclear Magnetic Resonance, NMR), το οποίο περιγράφηκε για πρώτη φορά στις δεκαετίες του 1940 και 1950 και αποτέλεσε τη βάση για την ανάπτυξη της MRI ως απεικονιστικής μεθόδου. Το φαινόμενο αυτό προκύπτει από τη συμπεριφορά πυρήνων με μη μηδενικό σπιν όταν τοποθετούνται σε ισχυρό εξωτερικό μαγνητικό πεδίο.

Σε θερμική ισορροπία και υπό την παρουσία ενός στατικού μαγνητικού πεδίου B_0 , τα πρωτόνια (πυρήνες υδρογόνου) τείνουν να ευθυγραμμιστούν είτε παράλληλα είτε αντιπαράλληλα προς τον άξονα του πεδίου (z -άξονας). Λόγω της στατιστικής κατανομής των ενεργειακών καταστάσεων, ένα μικρό πλεόνασμα πρωτονίων ευθυγραμμίζεται παράλληλα προς το πεδίο, δημιουργώντας τη λεγόμενη *συνολική μαγνήτιση* $\mathbf{M}_0$ κατά μήκος του άξονα z . Η μαγνήτιση αυτή είναι το θεμέλιο του ανιχνεύσιμου σήματος στην MRI.

Τα σπίν δεν παραμένουν ακίνητα, αλλά *προεκπορεύονται* γύρω από τον άξονα του εξωτερικού πεδίου με συγκεκριμένη γωνιακή συχνότητα, γνωστή ως *συχνότητα Larmor*:

$$\omega_0 = \gamma B_0,$$

όπου γ είναι ο *γυρομαγνητικός λόγος* του πυρήνα. Για το υδρογόνο, που αποτελεί και τον συχνότερα χρησιμοποιούμενο πυρήνα στην MRI, η τιμή είναι $\gamma \approx 42.58 \text{ MHz/T}$ [35], [36], [37]. Έτσι, σε μαγνητικό πεδίο 1.5 Tesla, η συχνότητα Larmor είναι περίπου 64 MHz, ενώ σε 3 Tesla ανέρχεται σε περίπου 128 MHz. Αυτό σημαίνει ότι τα πρωτόνια «συγχρονίζονται» με συγκεκριμένη συχνότητα που εξαρτάται αποκλειστικά από την ένταση του πεδίου.

Η προεκπόρευση αυτή, σε συνδυασμό με τη στατιστική ανισορροπία πληθυσμών στα ενεργειακά επίπεδα (λίγο περισσότερα πρωτόνια στην κατάσταση χαμηλής ενέργειας σε σχέση με την υψηλής), είναι εκείνη που καθιστά δυνατή την ανίχνευση σήματος. Όταν εφαρμοστεί ένας παλμός ραδιοσυχνότητας (RF) στη συχνότητα Larmor, μεταβάλλεται η κατανομή των πληθυσμών και η συνολική μαγνήτιση εκτρέπεται από τον άξονα z προς το εγκάρσιο επίπεδο (x - y). Καθώς τα πρωτόνια επανέρχονται στη θερμική ισορροπία, εκπέμπεται ένα σήμα RF, το λεγόμενο *σήμα ελεύθερης επαγωγικής απόσβεσης* (Free Induction Decay, FID), το οποίο ανιχνεύεται από τα πηνία λήψης και αποτελεί την πρώτη ύλη για την ανακατασκευή των εικόνων.

Μια σημαντική διάκριση είναι μεταξύ *διαμήκους μαγνήτισης* (longitudinal magnetization), δηλαδή της συνιστώσας κατά μήκος του άξονα z , και *εγκάρσιας μαγνήτισης* (transverse magnetization), η οποία βρίσκεται στο επίπεδο x - y και είναι εκείνη που παράγει το ανιχνεύσιμο σήμα. Η αλληλεπίδραση ανάμεσα σε αυτές τις δύο συνιστώσες και οι χρόνοι αποκατάστασης T_1 και T_2 αποτελούν κεντρικά μεγέθη που καθορίζουν την αντίθεση και τα χαρακτηριστικά των εικόνων MRI, όπως θα αναλυθεί στις επόμενες ενότητες.

2.4 RF Συντονισμός και χρόνοι χαλάρωσης

Όταν εφαρμοστεί ένα εγκάρσιο ραδιοσυχνότητας (RF) πεδίο $\mathbf{B}_1(t)$ στη συχνότητα Larmor ω_0 , η συνολική μαγνήτιση $\mathbf{M}_0$ εκτρέπεται από τον διαμήκη άξονα προς το εγκάρσιο επίπεδο. Η εκτροπή αυτή περιγράφεται από τη γωνία *flip* α , η οποία εξαρτάται από τη διάρκεια και την ισχύ του RF παλμού. Έτσι, μπορούμε να «στρέψουμε» την $\mathbf{M}_0$ κατά 90° (πλήρης μεταφορά στο εγκάρσιο επίπεδο) ή κατά 180° (αναστροφή προς την αντίθετη κατεύθυνση στον άξονα z). Το γεγονός αυτό είναι θεμελιώδες για τη δημιουργία διαφορετικών ακολουθιών απεικόνισης.

Μετά την παύση του RF παλμού, η μαγνήτιση τείνει να επανέλθει στην κατάσταση ισορροπίας μέσω δύο βασικών διεργασιών, που ονομάζονται *χαλάρωσεις* [35], [36]:

$$M_z(t) = M_0 \left(1 - e^{-t/T_1}\right) \quad (\text{διαμήκης χαλάρωση}), \quad (2.1)$$

$$M_{xy}(t) = M_{xy}(0) e^{-t/T_2} \quad (\text{εγκάρσια χαλάρωση}). \quad (2.2)$$

Διαμήκης χαλάρωση (T_1)

Η T_1 ή *spin-lattice relaxation* εκφράζει τον χρόνο που απαιτείται για να ανακτηθεί η διαμήκης μαγνήτιση (M_z) κατά μήκος του άξονα z . Στην ουσία, περιγράφει την ανταλλαγή ενέργειας των σπιν με το περιβάλλον τους (lattice), οδηγώντας σε θερμική επαναφορά. Ιστοί με μικρή T_1 (π.χ. λίπος) επανέρχονται γρήγορα και εμφανίζονται φωτεινοί σε T1-weighted εικόνες, ενώ ιστοί με μεγαλύτερη T_1 (π.χ. ENY) παραμένουν σκοτεινοί. Οι T1-σταθμισμένες εικόνες προσφέρουν άριστη ανατομική λεπτομέρεια και είναι ιδιαίτερα χρήσιμες στην αξιολόγηση της μορφολογίας.

Εγκάρσια χαλάρωση (T_2 και T_2^*)

Η T_2 ή *spin-spin relaxation* περιγράφει την απώλεια συνοχής στην εγκάρσια μαγνήτιση (M_{xy}), λόγω αλληλεπιδράσεων μεταξύ γειτονικών σπιν. Ο χρόνος αυτός καθορίζει πόσο γρήγορα μειώνεται το ανιχνεύσιμο σήμα στο εγκάρσιο επίπεδο. Ιστοί με υψηλή περιεκτικότητα σε νερό (π.χ. εγκεφαλονωτιαίο υγρό, οίδημα) εμφανίζουν μεγάλες τιμές T_2 και, συνεπώς, φωτεινή απεικόνιση σε T2-weighted εικόνες. Αντίθετα, οι περιοχές με μικρό T_2 (π.χ. λευκή ουσία) φαίνονται σκοτεινές.

Η παράμετρος T_2^* (*effective transverse relaxation*) είναι πάντα μικρότερη από την T_2 , καθώς συνυπολογίζει επιπλέον ανομοιογένειες του μαγνητικού πεδίου και φαινόμενα μαγνητικής επιδεκτικότητας. Η T_2^* είναι ιδιαίτερα σημαντική στις *gradient-echo* ακολουθίες και αξιοποιείται σε τεχνικές όπως η *functional MRI (fMRI)* και η *susceptibility-weighted imaging (SWI)*.

Στάθμιση εικόνων: TR, TE και TI

Η αντίθεση στις εικόνες MRI καθορίζεται από τον συνδυασμό των παραμέτρων TR (repetition time), TE (echo time) και TI (inversion time). Για παράδειγμα:

- Μικρό TR και μικρό TE ευνοούν T1-στάθμιση.
- Μεγάλο TR και μεγάλο TE ευνοούν T2-στάθμιση.
- Μεγάλο TR και μικρό TE οδηγούν σε Proton Density (PD) στάθμιση.

Επιπλέον, η χρήση κατάλληλου χρόνου αναστροφής (TI) σε *inversion recovery* ακολουθίες επιτρέπει την καταστολή συγκεκριμένων ιστών.

Γωνία Ernst

Χαρακτηριστική είναι η γωνία Ernst:

$$\alpha_E = \cos^{-1} \left(\exp(-TR/T_1) \right),$$

η οποία μεγιστοποιεί το σήμα σε δεδομένο χρόνο επανάληψης (TR) για ιστό με συγκεκριμένο T_1 [35]. Η σχέση αυτή καθοδηγεί τον σχεδιασμό ακολουθιών, ιδιαίτερα σε περιπτώσεις όπου επιδιώκεται η βέλτιστη σχέση σήματος προς θόρυβο (SNR) με περιορισμένο χρόνο σάρωσης.

Inversion recovery και FLAIR. Οι *inversion recovery* ακολουθίες αξιοποιούν έναν παλμό 180° που αναστρέφει την M_0 , ακολουθούμενο από έναν χρόνο αναμονής TI πριν από την εκπομπή του επόμενου παλμού. Με κατάλληλη επιλογή TI μπορούμε να «μηδενίσουμε» το σήμα από συγκεκριμένους ιστούς. Χαρακτηριστικό παράδειγμα είναι η *FLAIR (Fluid-Attenuated Inversion Recovery)*, όπου ο TI επιλέγεται ώστε να καταστείλει το σήμα του εγκεφαλονωτιαίου υγρού (ENY), αναδεικνύοντας με σαφήνεια παθολογικές βλάβες κοντά στις κοιλίες, όπως φλεγμονές ή όγκους [39].

2.5 Χωρικός εντοπισμός του σήματος

Η Μαγνητική Τομογραφία, σε αντίθεση με άλλες τεχνικές απεικόνισης, δεν ανιχνεύει απευθείας εικόνα αλλά ένα σήμα το οποίο πρέπει να κωδικοποιηθεί χωρικά ώστε να αποδοθεί η γεωμετρία του αντικειμένου. Ο χωρικός εντοπισμός επιτυγχάνεται μέσω της εφαρμογής γραμμικών βαθμοτών μαγνητικού πεδίου (magnetic field gradients) $G(t) = [G_x, G_y, G_z]$, οι οποίοι υπερτίθενται στο στατικό πεδίο B_0 και μεταβάλλουν τοπικά τη συχνότητα Larmor των πυρήνων [39], [41]. Με τον τρόπο αυτό, η συχνότητα ή η φάση του σήματος που εκπέμπουν τα πρωτόνια αποκτά χωρική εξάρτηση, επιτρέποντας την κωδικοποίηση θέσης.

Επιλογή τομής (slice selection)

Για την απεικόνιση ενός συγκεκριμένου επιπέδου, εφαρμόζεται ταυτόχρονα ένας RF παλμός και ένας βαθμωτός G_z . Ο RF παλμός έχει συγκεκριμένο εύρος ζώνης συχνοτήτων, και μόνο τα πρωτόνια των οποίων η συχνότητα Larmor εμπίπτει σε αυτό το εύρος διεγείρονται. Έτσι, η τομή που επιλέγεται αντιστοιχεί σε ένα επίπεδο κάθετο στον άξονα z , με πάχος που καθορίζεται από το εύρος ζώνης του παλμού και την ένταση του βαθμωτού. Με τη ρύθμιση των παραμέτρων αυτών μπορούμε να ελέγξουμε το πάχος της τομής και να επιτύχουμε είτε υψηλή λεπτομέρεια σε λεπτές τομές είτε μεγαλύτερη κάλυψη σε παχύτερες.

Κωδικοποίηση φάσης και συχνότητας

Μετά την επιλογή τομής, απαιτείται περαιτέρω κωδικοποίηση στις δύο διαστάσεις του επιπέδου. Αυτό επιτυγχάνεται με:

- **Κωδικοποίηση συχνότητας (frequency encoding):** Εφαρμογή βαθμωτού G_x κατά την ανίχνευση του σήματος, που προκαλεί γραμμική μεταβολή της συχνότητας Larmor ως προς τον άξονα x . Έτσι, η συχνότητα κάθε σημείου αντιστοιχεί στη θέση του στον άξονα.
- **Κωδικοποίηση φάσης (phase encoding):** Εφαρμογή παλμού βαθμωτού G_y για σύντομο χρονικό διάστημα πριν από την ανίχνευση. Αυτό δημιουργεί διαφορές φάσης στις ταλαντώσεις των πρωτονίων ανάλογα με τη θέση τους στον άξονα y , επιτρέποντας τη διάκριση σημείων με ίδια συχνότητα αλλά διαφορετική θέση.

Ο κ-χώρος (k-space)

Η συνδυασμένη χρήση των παραπάνω τεχνικών δημιουργεί ένα δισδιάστατο ή τρισδιάστατο σύνολο δεδομένων γνωστό ως *κ-χώρος*. Ο κ-χώρος δεν αποτελεί εικόνα, αλλά μια αναπαράσταση σε όρους χωρικών συχνοτήτων. Κάθε σημείο του κ-χώρου αντιστοιχεί σε μια συχνότητα ή φάση του σήματος, και η πλήρης δειγματοληψία του είναι απαραίτητη για την ανακατασκευή της εικόνας μέσω μετασχηματισμού Fourier. Οι κεντρικές περιοχές του κ-χώρου αντιστοιχούν σε χαμηλές συχνότητες που καθορίζουν την αντίθεση της εικόνας, ενώ οι περιφερειακές περιοχές σε υψηλές συχνότητες που καθορίζουν τη λεπτο-

μέρεια. Η ποιότητα και η ανάλυση της τελικής εικόνας εξαρτώνται άμεσα από την πληρότητα και την ορθότητα της δειγματοληψίας του κ-χώρου.

Προκλήσεις και τεχνικά ζητήματα

Η διαδικασία χωρικού κωδικοποιητή δεν είναι απαλλαγμένη από προβλήματα. Περιορισμοί στον χρόνο σάρωσης οδηγούν σε υποδειγματοληψία του κ-χώρου, με αποτέλεσμα *artefacts* όπως *aliasing* (επικάλυψη εικόνων), *blurring* ή παραμορφώσεις. Ανομοιογένειες του μαγνητικού πεδίου και κινήσεις του ασθενούς κατά τη διάρκεια της εξέτασης μπορούν επίσης να υποβαθμίσουν την ποιότητα της εικόνας. Για την αντιμετώπιση αυτών, έχουν αναπτυχθεί προηγμένες τεχνικές όπως η παράλληλη απεικόνιση και η *compressed sensing*, που μειώνουν τον απαιτούμενο χρόνο σάρωσης και βελτιώνουν την ποιότητα εικόνας.

2.6 Ανακατασκευή εικόνας

Η διαδικασία της ανακατασκευής εικόνας αποτελεί κρίσιμο βήμα στη Μαγνητική Τομογραφία, καθώς τα δεδομένα που ανιχνεύονται δεν είναι απευθείας εικόνες, αλλά πολύπλοκα σήματα που αντιστοιχούν στον κ-χώρο. Ο τρόπος δειγματοληψίας, οι τεχνικές επιτάχυνσης και οι μέθοδοι ανακατασκευής καθορίζουν την ποιότητα, την ανάλυση και τον χρόνο απόκτησης των τελικών εικόνων.

2.6.1 Κ-χώρος και ΔΜΤ (FFT)

Το σήμα που λαμβάνεται από τα πηνία λήψης αποτελεί τη μιγαδική δειγματοληψία του κ-χώρου, ενός μαθηματικού χώρου που περιγράφει τις χωρικές συχνότητες της εικόνας. Τα σημεία του κέντρου του κ-χώρου περιέχουν κυρίως πληροφορία χαμηλών συχνοτήτων, δηλαδή τη «μάζα» και τη γενική φωτεινότητα/αντίθεση της εικόνας, ενώ τα περιφερειακά σημεία περιέχουν υψηλές συχνότητες, υπεύθυνες για την ευκρίνεια, τις λεπτομέρειες και τις ακμές.

Η ανακατασκευή γίνεται τυπικά με την εφαρμογή της *ταχείας μεταστροφής Fourier* (Fast Fourier Transform, FFT), η οποία μετατρέπει το σήμα από τον κ-χώρο στον χώρο εικόνας. Για να είναι η ανακατασκευή ακριβής, απαιτείται ο κ-χώρος να έχει δειγματοληφθεί πλήρως και καρτεσιανά. Σε αυτή την περίπτωση, η διαδικασία είναι σχετικά απλή και ταχύτατη, με χαμηλό υπολογιστικό κόστος [43]–[45]. Αντίθετα, ελλιπής ή ανομοιόμορφη δειγματοληψία οδηγεί σε παραμορφώσεις και τεχνουργήματα.

2.6.2 Compressed Sensing & επαναληπτική ανακατασκευή

Η ανάγκη μείωσης του χρόνου σάρωσης, που αποτελεί κρίσιμο παράγοντα για την άνεση του ασθενούς και την αποφυγή θολώματος από κίνηση, οδήγησε στην ανάπτυξη τεχνικών όπως η *Compressed Sensing* (CS). Η CS εκμεταλλεύεται την ιδιότητα ότι οι ιατρικές εικόνες είναι «αραιές» σε κάποιον κατάλληλο μετασχηματισμό (π.χ. κυματίδια / *wavelets*, ολική μεταβολή / *total variation*). Αυτό σημαίνει ότι η εικόνα μπορεί να περιγραφεί με λίγους μόνο σημαντικούς συντελεστές. Έτσι, αρκεί η υποδειγματοληψία του κ-χώρου και η ανάκτηση της εικόνας μέσω επαναληπτικών αλγορίθμων που επιλύουν πρόβλημα βελτιστοποίησης:

$$\min_{\mathbf{x}} \|\mathbf{F}_u \mathbf{x} - \mathbf{y}\|_2^2 + \lambda \Phi(\mathbf{x}),$$

όπου $\mathbf{F}_u$ είναι ο υποδειγματοληπτημένος Fourier, $\mathbf{y}$ τα μετρημένα δεδομένα, και $\Phi(\mathbf{x})$ ένας όρος που επιβάλλει αραιότητα ή ομαλότητα. Η παράμετρος λ καθορίζει τον βαθμό εξομάλυνσης/συμπίεσης. Το σχήμα δειγματοληψίας είναι συχνά *variable-density*, δηλαδή πιο πυκνό στο κέντρο και πιο αραιό στην περιφέρεια, ώστε να διατηρείται η αντίθεση με μειωμένο χρόνο σάρωσης [45]. Αν και οι μέθοδοι CS είναι υπολογιστικά πιο απαιτητικές, έχουν βρει κλινική εφαρμογή σε καρδιακές και νευρολογικές MRI,

μειώνοντας σημαντικά τον χρόνο λήψης χωρίς μεγάλη απώλεια ποιότητας.

2.6.3 Παράλληλη απεικόνιση (SENSE/GRAPPA)

Μία άλλη προσέγγιση για επιτάχυνση της σάρωσης είναι η *παράλληλη απεικόνιση* (Parallel Imaging), η οποία αξιοποιεί τα πολλαπλά πηνία λήψης που διαθέτουν διαφορετικούς χάρτες ευαισθησίας. Χρησιμοποιώντας αυτή την πλεονάζουσα πληροφορία, μπορούμε να μειώσουμε το πλήθος γραμμών κ-χώρου που χρειάζεται να δειγματοληφθούν, επιτυγχάνοντας επιτάχυνση R φορές (τυπικά $R = 2-4$). Οι δύο πιο διαδεδομένες μέθοδοι είναι:

- **SENSE (Sensitivity Encoding):** εκτελείται στον χώρο εικόνας, όπου τα σήματα αποδιαμορφώνονται χρησιμοποιώντας γνωστούς χάρτες ευαισθησίας.
- **GRAPPA (Generalized Autocalibrating Partially Parallel Acquisition):** εκτελείται στον κ-χώρο, όπου οι ελλείπουσες γραμμές ανακατασκευάζονται από γραμμικούς συνδυασμούς των υπαρχουσών.

Η παράλληλη απεικόνιση συνεπάγεται απώλεια σήματος (χαμηλότερο SNR), η οποία ποσοτικοποιείται από τον *g-factor*, που εξαρτάται από τη γεωμετρία των πηνίων και τον συντελεστή επιτάχυνσης [39].

2.6.4 Διαδρομές κ-χώρου, κίνηση και αλλοίωση

Αν και η καρτεσιανή δειγματοληψία είναι η πιο συχνή, σε πολλές εφαρμογές χρησιμοποιούνται εναλλακτικές διαδρομές, όπως οι *radial* και *spiral*. Αυτές έχουν πλεονεκτήματα ως προς την ανθεκτικότητα σε κίνηση και την αποτελεσματική υποδειγματοληψία, καθώς η κεντρική περιοχή του κ-χώρου (που καθορίζει την αντίθεση) δειγματοληπτείται πιο συχνά. Ωστόσο, η ανακατασκευή τους απαιτεί πιο πολύπλοκους αλγόριθμους, όπως η επαναδειγματοληψία σε καρτεσιανή μορφή (regridding).

Η κίνηση του ασθενούς, η αναπνοή και ο καρδιακός παλμός εισάγουν ανεπιθύμητα *τεχνουργήματα* (artifacts), όπως ghosting ή θολώματα. Για την αντιμετώπισή τους εφαρμόζονται είτε *προοπτικές τεχνικές* (navigator echoes, prospective correction), οι οποίες παρακολουθούν την κίνηση σε πραγματικό χρόνο και προσαρμόζουν τα gradients, είτε *εκ των υστέρων* μέθοδοι (retrospective correction), που επιχειρούν να διορθώσουν τις παραμορφώσεις κατά την ανακατασκευή. Η ανάπτυξη νέων αλγορίθμων βασισμένων σε τεχνητή νοημοσύνη υπόσχεται περαιτέρω βελτίωση της ποιότητας εικόνας ακόμη και σε εξετάσεις με κίνηση.

2.7 Προηγμένες τεχνικές MRI

Πέρα από τις βασικές ακολουθίες απεικόνισης (T1-, T2- και FLAIR-weighted), η Μαγνητική Τομογραφία έχει εξελιχθεί σε ένα εξαιρετικά πολύπλευρο εργαλείο, με εξειδικευμένες παραλλαγές που προσφέρουν πληροφορίες όχι μόνο για τη δομή, αλλά και για τη λειτουργία και τη μικροδομή των ιστών. Στην παρούσα εργασία δεν χρησιμοποιήθηκαν τέτοιες προηγμένες τεχνικές, ωστόσο αξίζει να αναφερθούν για να αποτυπωθεί το εύρος των δυνατοτήτων της MRI.

Λειτουργική MRI (fMRI)

Η λειτουργική MRI (*functional MRI*, fMRI) μετρά τις μεταβολές της αιματικής οξυγόνωσης (*Blood Oxygen Level Dependent*, BOLD signal) ως έμμεσο δείκτη νευρωνικής δραστηριότητας [ogawa1990brain]. Η τεχνική βασίζεται στο γεγονός ότι η οξυαιμοσφαιρίνη και η δεοξυαιμοσφαιρίνη έχουν διαφορετικές μαγνητικές ιδιότητες, με αποτέλεσμα το τοπικό σήμα να μεταβάλλεται κατά την ενεργοποίηση συγκεκριμένων εγκεφαλικών περιοχών. Από τη δεκαετία του 1990, η fMRI έχει καθιερωθεί ως βασικό εργαλείο

στη νευροεπιστήμη για τη χαρτογράφηση λειτουργικών περιοχών, την κατανόηση εγκεφαλικών δικτύων, αλλά και σε κλινικές εφαρμογές (π.χ. προεγχειρητικός σχεδιασμός σε όγκους που γειτνιάζουν με κινητικές ή γλωσσικές περιοχές). Παρά την υψηλή χωρική ανάλυση, η fMRI πάσχει από σχετικά χαμηλή χρονική ανάλυση και εξάρτηση από αιμοδυναμικούς μηχανισμούς, γεγονός που καθιστά την ερμηνεία της πιο περίπλοκη.

Απεικόνιση Διάχυσης και DTI

Η απεικόνιση διάχυσης (*Diffusion Weighted Imaging*, DWI) εισήχθη στις αρχές της δεκαετίας του 1990 [le1990mr] και επέτρεψε τη μη επεμβατική εκτίμηση της μικροσκοπικής κίνησης των μορίων νερού στον ιστό. Στον εγκέφαλο, η διάχυση είναι ανισότροπη λόγω της παρουσίας ινών λευκής ουσίας, οι οποίες κατευθύνουν την κίνηση των μορίων. Η *Diffusion Tensor Imaging* (DTI) επεκτείνει τη βασική DWI μέσω της εκτίμησης ενός τανυστή διάχυσης, ο οποίος περιγράφει την τρισδιάστατη ανισοτροπία. Έτσι, παρέχεται η δυνατότητα χαρτογράφησης της λευκής ουσίας με *tractography*, δηλαδή την ανακατασκευή νευρικών δεματίων στον εγκέφαλο.

Η DTI έχει καθοριστικό ρόλο σε μελέτες νευροαναπτυξιακών διαταραχών, νευροεκφυλιστικών νόσων και σε προεγχειρητικό σχεδιασμό (π.χ. εντοπισμός κινητικών οδών πριν από εκτομή όγκου). Παράλληλα, παραμένει τεχνικά απαιτητική λόγω ευαισθησίας σε κινήσεις, θόρυβο και ανάγκη για πολλαπλές κατευθύνσεις διάχυσης.

Άλλες προηγμένες εφαρμογές

Εκτός από fMRI και DTI, έχουν αναπτυχθεί και άλλες ειδικευμένες τεχνικές:

- **Perfusion MRI (DSC/DCE):** αξιολογεί τη ροή αίματος στους ιστούς και χρησιμοποιείται ιδιαίτερα στην ογκολογία για διάκριση όγκου από οίδημα ή νέκρωση.
- **MR Spectroscopy (MRS):** επιτρέπει την ανάλυση μεταβολιτών (π.χ. χολίνη, NAA) και παρέχει μεταβολικές πληροφορίες που συμπληρώνουν την απλή μορφολογία.
- **Cardiac και abdominal MRI:** βελτιστοποιημένες ακολουθίες για όργανα που υπόκεινται σε κίνηση, με τεχνικές διόρθωσης αναπνοής και καρδιακού κύκλου.

Συνοπτικό σχόλιο. Οι προηγμένες τεχνικές δεν χρησιμοποιήθηκαν στην παρούσα εργασία, καθώς το αντικείμενο περιορίστηκε στη δομική απεικόνιση για τμηματοποίηση όγκων. Ωστόσο, η κατανόησή τους είναι κρίσιμη για το μέλλον: η ενσωμάτωση λειτουργικών και μεταβολικών δεδομένων (π.χ. fMRI ή MRS) μπορεί να οδηγήσει σε πολυπαραμετρική τμηματοποίηση με ακόμη μεγαλύτερη κλινική αξία.

2.8 MRI και διάγνωση όγκων εγκεφάλου

Η Μαγνητική Τομογραφία αποτελεί την κύρια απεικονιστική μέθοδο για την εκτίμηση ενδοκρανιακών νεοπλασιών, χάρη στην ικανότητά της να αποδίδει υψηλή αντίθεση μαλακών ιστών και να παρέχει πολυδιάστατη πληροφορία μέσω διαφορετικών *modalities*. Κάθε ακολουθία αναδεικνύει διαφορετικές φυσιολογικές ή παθολογικές ιδιότητες, και η συνδυαστική τους ανάλυση είναι απαραίτητη για τον ακριβή χαρακτηρισμό του όγκου. Στην κλινική πράξη, τα συνήθη πρωτόκολλα περιλαμβάνουν [39], [40]:

- **T1-weighted (πριν/μετά σκιαγραφικό):** Οι T1-σταθμισμένες εικόνες χωρίς σκιαγραφικό παρέχουν καλή ανατομική απεικόνιση. Μετά τη χορήγηση γαδολινίου (Gd), οι περιοχές με διαταραγ-

μένο αιματοεγκεφαλικό φραγμό (BBB) εμφανίζουν ενίσχυση (*enhancing tumor*), κάτι που βοηθά στη διάκριση ενεργού όγκου από οίδημα ή νεκρωτικό ιστό.

- **T2-weighted και FLAIR:** Οι T2 εικόνες αναδεικνύουν περιοχές με υψηλή περιεκτικότητα σε νερό, όπως το οίδημα, ενώ η FLAIR (Fluid-Attenuated Inversion Recovery) καταστέλλει το σήμα του ΕΝΥ, επιτρέποντας την ανίχνευση βλαβών παρακείμενων στις κοιλίες. Είναι ιδιαίτερα χρήσιμη για τον εντοπισμό διήθησης στη λευκή ουσία (*peritumoral edema*).
- **Diffusion-Weighted Imaging (DWI) / Apparent Diffusion Coefficient (ADC):** Η τεχνική αυτή καταγράφει τη διάχυση του ύδατος στον ιστό. Περιοχές με υψηλή κυτταρική πυκνότητα, όπως οι επιθετικοί όγκοι ή τα αποστήματα, εμφανίζουν περιορισμένη διάχυση και χαμηλές τιμές *ADC*, ενώ περιοχές με νέκρωση ή οίδημα εμφανίζουν αυξημένες τιμές.
- **Perfusion MRI (Dynamic Susceptibility Contrast, DSC / Dynamic Contrast-Enhanced, DCE):** Παρέχει αιμοδυναμικούς δείκτες όπως το σχετικό εγκεφαλικό αιματικό όγκο (rCBV). Οι δείκτες αυτοί είναι χρήσιμοι στη διάκριση μεταξύ υψηλού και χαμηλού βαθμού κακοήθειας, καθώς και στη διαφοροποίηση υποτροπής από ψευδο-πρόοδο μετά από θεραπεία.
- **MR Spectroscopy (MRS):** Αναλύει το μεταβολικό προφίλ ιστών. Αυξημένα επίπεδα χολίνης (δείκτης κυτταρικής μεμβράνης) και μειωμένο N-ακετυλασπαρτικό (NAA, δείκτης νευρωνικής ακεραιότητας) είναι χαρακτηριστικά ευρήματα σε κακοήθεις όγκους.

Η ολοκληρωμένη ογκολογική αξιολόγηση στηρίζεται στη συνεκτίμηση πολλών modalities. Για παράδειγμα, ένας όγκος μπορεί να παρουσιάζει ενίσχυση με σκιαγραφικό στην T1c, εκτεταμένο οίδημα σε FLAIR και αυξημένο rCBV στην perfusion. Η συνδυαστική αυτή πληροφορία βοηθά όχι μόνο στη διάγνωση, αλλά και στη σταδιοποίηση, στον σχεδιασμό θεραπείας και στην παρακολούθηση της ανταπόκρισης. Ειδικά για την τμηματοποίηση εγκεφαλικών όγκων, οι διαφορετικές αντιθέσεις προσφέρουν συμπληρωματικά patterns, που είναι καθοριστικά για τον ακριβή διαχωρισμό επιμέρους περιοχών (ενισχυόμενος όγκος, οίδημα, νεκρωτικός πυρήνας) [32], [39].

2.8.1 Συνηθισμένα τεχνουργήματα και μετριάσμός

Όπως κάθε απεικονιστική τεχνική, έτσι και η MRI είναι επιρρεπής σε τεχνουργήματα (artifacts), τα οποία μπορεί να επηρεάσουν την ποιότητα και την αξιοπιστία της διάγνωσης. Μερικά από τα συνηθέστερα περιλαμβάνουν:

- *Gibbs ringing*: προκύπτει από περιορισμένη δειγματοληψία του κ-χώρου και εμφανίζεται ως ταλαντώσεις γύρω από απότομες ακμές.
- *Chemical shift*: προκαλείται από τη διαφορά συχνότητας μεταξύ λιπιδίων και ύδατος, οδηγώντας σε λανθασμένη απεικόνιση των ορίων.
- *Susceptibility artifacts*: εμφανίζονται σε περιοχές με μεγάλη διαφορά μαγνητικής επιδεκτικότητας (αέρας, οστό, αίμα), προκαλώντας παραμορφώσεις ή απώλεια σήματος.
- *RF inhomogeneity (bias field)*: οδηγεί σε ανομοιογενή φωτεινότητα στην εικόνα.
- *Aliasing ή wrap-around*: εμφανίζεται όταν το πεδίο απεικόνισης (FOV) είναι μικρότερο από την ανατομική περιοχή, με αποτέλεσμα η εικόνα να «τυλίγεται».

Τα παραπάνω μετριάζονται με διάφορες τεχνικές, όπως αύξηση της δειγματοληψίας, καταστολή λίπους (*fat suppression*), βελτιστοποίηση παραμέτρων σάρωσης, *shimming* για βελτίωση ομοιογένειας πεδίου, καθώς και αλγοριθμικές μεθόδους διόρθωσης κατά την ανακατασκευή [39], [45].

2.8.2 Προεπεξεργασία δεδομένων για μάθηση και τμηματοποίηση

Για την αξιόπιστη εφαρμογή μεθόδων μηχανικής μάθησης και αυτόματης τμηματοποίησης, τα δεδομένα MRI πρέπει να υποβληθούν σε προσεκτική προεπεξεργασία. Οι πιο συνηθισμένοι βηματισμοί περιλαμβάνουν:

- **Skull-stripping και ευθυγράμμιση:** Αφαίρεση κρανίου και εξωκρανιακών ιστών, καθώς και εγγραφή (*rigid/affine registration*) σε κοινό πρότυπο χώρο.
- **Διόρθωση ανομοιομορφίας πεδίου (*bias field correction*):** π.χ. με τον αλγόριθμο N4ITK, ώστε να βελτιωθεί η ομοιογένεια της έντασης.
- **Κανονικοποίηση έντασης:** είτε με *z-score* ανά όγκο είτε με *histogram matching*, ώστε να μειωθεί η έντονη διακύμανση μεταξύ σαρωτών ή πρωτοκόλλων.
- **Επαναδειγματοληψία/ανασάρωση:** σε τυπική ιστροπική ανάλυση (π.χ. 1 mm^3), με ταυτόχρονη συνεγγραφή όλων των modalities.

Οι παραπάνω διαδικασίες μειώνουν τα φαινόμενα *domain shift* (διαφορές μεταξύ κέντρων/σαρωτών) και βελτιώνουν τη γενίκευση των μοντέλων μάθησης.

2.8.3 BraTS 2020

Το *Multimodal Brain Tumor Segmentation Challenge 2020* (BraTS 2020) αποτελεί έναν από τους σημαντικότερους δημόσιους διαγωνισμούς στον χώρο της υπολογιστικής νευροαπεικόνισης. Περιλαμβάνει πολυτροπικά δεδομένα MRI (T1, T1c, T2 και FLAIR), τα οποία έχουν συνεγγραφεί και επαναδειγματοληφθεί σε κοινό πλέγμα διαστάσεων $240 \times 240 \times 155$. Οι σημειώσεις (*labels*) έχουν γίνει από εξειδικευμένους νευροακτινολόγους και περιλαμβάνουν τρεις κύριες περιοχές ενδιαφέροντος: *enhancing tumor (ET)*, *peritumoral edema (ED)*, και *necrotic/non-enhancing core (NCR/NET)* [32].

Η αξιολόγηση των μεθόδων τμηματοποίησης γίνεται με πληθώρα μετρικών, όπως ο *Dice similarity coefficient*, η απόσταση Hausdorff95, καθώς και επιμέρους μετρικές για τις υπο-περιοχές (ET, TC, WT). Το BraTS έχει συμβάλει καθοριστικά στην προώθηση της έρευνας, προσφέροντας ένα κοινό benchmark που επιτρέπει τη δίκαιη σύγκριση μεθόδων και την ανάπτυξη νέων τεχνικών με υψηλή κλινική σημασία. Στο πλαίσιο της παρούσας εργασίας, το BraTS 2020 παρέχει πλούσια πολυτροπική πληροφορία και τυποποιημένη προεπεξεργασία, διευκολύνοντας την εκπαίδευση και αξιολόγηση μοντέλων βαθιάς μάθησης για τμηματοποίηση εγκεφαλικών όγκων [32].

2.9 Ασφάλεια και σκιαγραφικά

Η Μαγνητική Τομογραφία θεωρείται γενικά μία από τις πιο ασφαλείς απεικονιστικές μεθόδους, καθώς δεν χρησιμοποιεί ιοντίζουσα ακτινοβολία. Παρ' όλα αυτά, η παρουσία ισχυρού μαγνητικού πεδίου (B_0), η χρήση παλμών ραδιοσυχνότητας και οι ταχείες μεταβολές των βαθμοτών συνεπάγονται συγκεκριμένους περιορισμούς και προφυλάξεις.

Γενικά ζητήματα ασφάλειας

Το στατικό μαγνητικό πεδίο μπορεί να επηρεάσει ή να μετακινήσει μεταλλικά αντικείμενα, ενώ μπορεί να επιδράσει στη λειτουργία ηλεκτρονικών εμφυτευμάτων όπως βηματοδότες, νευροδιεγέρτες ή αντλίες ινσουλίνης. Για τον λόγο αυτό, πριν την εξέταση γίνεται πάντα λεπτομερής έλεγχος για τυχόν εμφυτευμένα υλικά, μεταλλικά θραύσματα ή άλλες αντενδείξεις. Σήμερα, όλο και περισσότερες συσκευές χαρακτηρίζονται ως *MRI conditional*, δηλαδή μπορούν να χρησιμοποιηθούν με ασφάλεια υπό συγκεκριμένες συνθήκες, αλλά η αξιολόγηση γίνεται πάντοτε εξατομικευμένα.

Οι παλμοί ραδιοσυχνότητας προκαλούν θέρμανση των ιστών, που ποσοτικοποιείται με τον δείκτη *Specific Absorption Rate* (SAR). Η τήρηση των διεθνών ορίων SAR διασφαλίζει ότι η θερμική επιβάρυνση παραμένει εντός ασφαλών επιπέδων. Παράλληλα, οι ταχείες εναλλαγές των βαθμοτών μπορεί να προκαλέσουν μυϊκές συσπάσεις ή αίσθημα παραισθήσεων, ενώ η λειτουργία τους παράγει υψηλά επίπεδα θορύβου, για τα οποία συνιστάται η χρήση ακουστικής προστασίας.

Σκιαγραφικά με βάση το γαδολίνιο

Η χορήγηση σκιαγραφικών με βάση το γαδολίνιο (Gadolinium-based Contrast Agents, GBCAs) έχει καθιερωθεί στην MRI, καθώς αυξάνει σημαντικά την T1 αντίθεση, διευκολύνοντας την ανάδειξη παθολογικών περιοχών με διαταραχή του αιματοεγκεφαλικού φραγμού. Τα σκιαγραφικά αυτά χορηγούνται ενδοφλεβίως και συνήθως είναι καλά ανεκτά, με ελάχιστες άμεσες ανεπιθύμητες ενέργειες, όπως ήπιες αλλεργικές αντιδράσεις ή αίσθημα θερμότητας κατά την έγχυση.

Ένα σπάνιο αλλά σοβαρό σύνδρομο που έχει συσχετιστεί με τη χορήγηση γαδολινίου είναι η *νεφρογενής συστηματική ίνωση* (Nephrogenic Systemic Fibrosis, NSF), η οποία μπορεί να εμφανιστεί σε ασθενείς με βαριά νεφρική ανεπάρκεια. Για τον λόγο αυτό, η χορήγηση GBCAs απαιτεί προσεκτική εκτίμηση της νεφρικής λειτουργίας (μέσω μέτρησης eGFR) και αποφεύγεται όταν η νεφρική ανεπάρκεια είναι σοβαρή. Επιπλέον, έχει καταγραφεί μικρός κίνδυνος εναπόθεσης γαδολινίου σε ιστούς, συμπεριλαμβανομένου του εγκεφάλου, χωρίς όμως μέχρι σήμερα να έχει αποδειχθεί κλινικά σημαντική τοξικότητα.

Η ένδειξη ή η αντένδειξη για χορήγηση σκιαγραφικού τίθεται ιατρικά ανά περίπτωση, λαμβάνοντας υπόψη το προσδοκώμενο όφελος στη διάγνωση σε σχέση με τους πιθανούς κινδύνους. Σε πολλές περιπτώσεις, η χρήση GBCAs είναι απαραίτητη για την ακριβή διαφοροποίηση όγκου από οίδημα, τον εντοπισμό μεταστάσεων ή την παρακολούθηση της ανταπόκρισης στη θεραπεία.

2.10 Σύνοψη κεφαλαίου

Στο παρόν κεφάλαιο παρουσιάστηκαν αναλυτικά οι φυσικές αρχές και οι τεχνολογικές συνιστώσες της Μαγνητικής Τομογραφίας (MRI), καθώς και οι εφαρμογές της στη διάγνωση εγκεφαλικών όγκων. Αρχικά, περιγράφηκαν οι θεμελιώδεις αρχές του πυρηνικού μαγνητικού συντονισμού, η δημιουργία της συνολικής μαγνήτισης και η προεκπόρευση στη συχνότητα Larmor. Αναλύθηκε η δράση των RF παλμών και οι μηχανισμοί χαλάρωσης T_1 , T_2 και T_2^* , οι οποίοι αποτελούν τη βάση για τη δημιουργία των διαφορετικών αντιθέσεων εικόνας. Δόθηκε έμφαση στον ρόλο των παραμέτρων TR, TE και TI, που καθορίζουν τη στάθμιση των εικόνων, και αναφέρθηκαν χαρακτηριστικές ακολουθίες όπως η spin-echo, gradient-echo και inversion recovery (FLAIR).

Στη συνέχεια, εξετάστηκε ο μηχανισμός χωρικού κωδικοποιητή μέσω βαθμοτών πεδίου, ο οποίος επιτρέπει την επιλογή τομής και την κωδικοποίηση συχνότητας και φάσης, οδηγώντας στη συγκρότηση του κ-χώρου. Η σημασία του κ-χώρου τονίστηκε με έμφαση στη διάκριση χαμηλών και υψηλών συχνοτήτων, καθώς και στις τεχνικές ανακατασκευής εικόνας. Παρουσιάστηκαν τόσο η κλασική προσέγγιση

με ταχεία μεταστροφή Fourier (FFT), όσο και νεότερες μέθοδοι που μειώνουν τον χρόνο σάρωσης, όπως το *Compressed Sensing* και η *παράλληλη απεικόνιση* (SENSE, GRAPPA). Συζητήθηκαν επίσης εναλλακτικές διαδρομές δειγματοληψίας στον κ-χώρο (radial, spiral) και η σχέση τους με την ανθεκτικότητα σε κίνηση και τεχνουργήματα.

Έπειτα, αναδείχθηκε η σημασία των MRI πρωτοκόλλων για τη διάγνωση και χαρτογράφηση όγκων εγκεφάλου. Εξετάστηκαν οι T1- και T2-σταθμισμένες εικόνες, η FLAIR, η απεικόνιση διάχυσης (DWI/ADC), η απεικόνιση αιμάτωσης (perfusion MRI) και η φασματοσκοπία (MRS). Τονίστηκε η αξία της πολυτροπικής πληροφορίας και η συνεκτίμησή της για τον ακριβή ογκολογικό χαρακτηρισμό, καθώς και η ανάγκη για τυποποιημένη προεπεξεργασία των δεδομένων (skull-stripping, bias field correction, κανονικοποίηση έντασης, επαναδειγματοληψία). Σημειώθηκαν επίσης τα συχνότερα τεχνουργήματα (π.χ. Gibbs ringing, chemical shift, susceptibility, aliasing) και οι τρόποι μετριασμού τους.

Ιδιαίτερη αναφορά έγινε στο *BraTS 2020*, το οποίο έχει εδραιωθεί ως διεθνές benchmark για την τμηματοποίηση εγκεφαλικών όγκων. Τα δεδομένα του, που περιλαμβάνουν πολλαπλά modalities και επιμελημένες σημειώσεις, προσφέρουν τυποποιημένη βάση για εκπαίδευση και σύγκριση αλγορίθμων, και χρησιμοποιούνται εκτενώς στην παρούσα εργασία.

Τέλος, συζητήθηκαν οι πτυχές ασφάλειας στη MRI και η χρήση σκιαγραφικών με βάση το γαδολίνιο. Εξηγήθηκε ότι η MRI είναι γενικά ασφαλής, με βασικές προφυλάξεις για ασθενείς με μεταλλικά εμφυτεύματα ή ηλεκτρονικές συσκευές, καθώς και η ανάγκη παρακολούθησης του δείκτη SAR. Η χρήση γαδολινίου ως σκιαγραφικό ενισχυτικό παράγοντα βελτιώνει σημαντικά τη διαγνωστική ακρίβεια, με προσεκτική όμως αξιολόγηση της νεφρικής λειτουργίας για την αποφυγή σπάνιων ανεπιθύμητων ενεργειών.

Συνολικά, το κεφάλαιο αυτό παρείχε το αναγκαίο φυσικο-τεχνικό υπόβαθρο για την κατανόηση της MRI και της εφαρμογής της στη διάγνωση όγκων εγκεφάλου. Η γνώση αυτή είναι κρίσιμη για την ανάπτυξη και αξιολόγηση μεθόδων αυτόματης τμηματοποίησης που θα παρουσιαστούν στα επόμενα κεφάλαια, καθώς γεφυρώνει τη θεωρητική φυσική της MRI με τις πρακτικές ανάγκες της κλινικής και της υπολογιστικής νευροαπεικόνισης [32], [35]–[41], [45].

Κεφάλαιο 3: Νευρωνικά Δίκτυα

3.1 Γενικά

Ο φυσικός κόσμος σπάνια είναι τυχαίος· ακόμη και εκεί όπου επικρατεί μεγάλη ποικιλομορφία, τα φαινόμενα ακολουθούν δομές και κανονικότητες. Το γνωστό νοητικό πείραμα των «άπειρων μαϊμούδων» επισημαίνει ότι μια καθαρά τυχαία διαδικασία έχει πρακτικά μηδενικές πιθανότητες να παραγάγει μια δομημένη, ρεαλιστική εικόνα ή κείμενο [14]. Η *Μηχανική Μάθηση* επιδιώκει να ανιχνεύσει αυτές τις κρυμμένες δομές και να τις αξιοποιήσει για πρόβλεψη ή λήψη αποφάσεων, δημιουργώντας μοντέλα που μαθαίνουν από δεδομένα χωρίς να είναι ρητά προγραμματισμένα με κανόνες.

Ανάμεσα στις σημαντικότερες οικογένειες μεθόδων της μηχανικής μάθησης βρίσκονται τα *Τεχνητά Νευρωνικά Δίκτυα* (ΤΝΔ). Πρόκειται για υπολογιστικά μοντέλα που έχουν εμπνευστεί από τον βιολογικό εγκέφαλο και τη λειτουργία των νευρώνων. Στον εγκέφαλο, οι νευρώνες επικοινωνούν μεταξύ τους μέσω συνάψεων, σχηματίζοντας δίκτυα που επεξεργάζονται και μεταδίδουν πληροφορία. Αντίστοιχα, στα τεχνητά νευρωνικά δίκτυα, οι «τεχνητοί νευρώνες» είναι απλές υπολογιστικές μονάδες που λαμβάνουν εισόδους, τις μετασχηματίζουν μέσω συναρτήσεων ενεργοποίησης και παράγουν εξόδους, οι οποίες διαδίδονται σε επόμενα επίπεδα του δικτύου.

Η ανάπτυξη των ΤΝΔ ξεκίνησε με το κλασικό μοντέλο McCulloch–Pitts (1943), το οποίο έθεσε τις βάσεις για τη μαθηματική περιγραφή ενός «νευρώνα». Στη συνέχεια, το perceptron (Rosenblatt, 1958) έδειξε ότι είναι δυνατή η αυτόματη εκμάθηση γραμμικών αποφάσεων από δεδομένα. Ωστόσο, η αδυναμία του perceptron να λύσει μη γραμμικά προβλήματα (όπως η λογική πράξη XOR) οδήγησε σε μια περίοδο σχετικής «ύφεσης» για την έρευνα στα νευρωνικά δίκτυα. Η αναβίωση ήρθε με την ανάπτυξη του αλγορίθμου *backpropagation* τη δεκαετία του 1980, ο οποίος επέτρεψε την αποτελεσματική εκπαίδευση πολυεπίπεδων δικτύων. Σήμερα, με την εκρηκτική αύξηση της υπολογιστικής ισχύος και των διαθέσιμων δεδομένων, τα «βαθιά» νευρωνικά δίκτυα (*deep learning*) αποτελούν τον ακρογωνιαίο λίθο της σύγχρονης τεχνητής νοημοσύνης.

Η σημασία των νευρωνικών δικτύων έγκειται στην ικανότητά τους να αναγνωρίζουν πολύπλοκα πρότυπα και να γενικεύουν από δεδομένα. Έχουν αποδειχθεί εξαιρετικά αποτελεσματικά σε πλήθος εφαρμογών: αναγνώριση φωνής και εικόνας, αυτόματη μετάφραση, επεξεργασία φυσικής γλώσσας, αλλά και σε εξειδικευμένους τομείς όπως η ιατρική διάγνωση και η ανάλυση βιοϊατρικών εικόνων. Στο πλαίσιο της παρούσας εργασίας, τα ΤΝΔ αξιοποιούνται για την αυτόματη τμηματοποίηση MRI εγκεφάλου, καθιστώντας τα αναγκαίο θεωρητικό υπόβαθρο για τα μεθοδολογικά κεφάλαια που ακολουθούν.

3.2 Βιολογικοί νευρώνες

Η έμπνευση των Τεχνητών Νευρωνικών Δικτύων προήλθε από τη μελέτη του βιολογικού νευρικού συστήματος, και ειδικότερα από τη λειτουργία των νευρώνων. Ο *νευρώνας* αποτελεί τη θεμελιώδη μονάδα επεξεργασίας του εγκεφάλου και του νευρικού συστήματος. Υπολογίζεται ότι ο ανθρώπινος εγκέφαλος περιέχει περίπου 10^{11} νευρώνες, καθένας από τους οποίους μπορεί να σχηματίζει χιλιάδες συνάψεις με άλλους, δημιουργώντας ένα δίκτυο τεράστιας πολυπλοκότητας [15, 16].

Η βασική δομή ενός βιολογικού νευρώνα περιλαμβάνει:

- **Δενδρίτες:** Αποτελούν τις «εισόδους» του νευρώνα, λαμβάνοντας σήματα από εκατοντάδες ή και

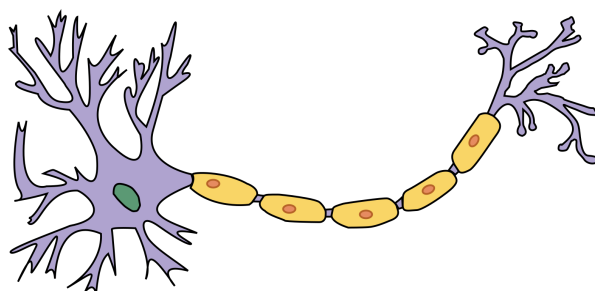
χιλιάδες άλλους νευρώνες. Το σήμα φθάνει υπό τη μορφή *δυναμικών μετασυναπτικής ενέργειας*, τα οποία μπορεί να είναι διεγερτικά ή ανασταλτικά.

- **Κυτταρικό σώμα (soma):** Συγκεντρώνει και ολοκληρώνει όλα τα εισερχόμενα σήματα. Εάν το άθροισμα ξεπεράσει ένα συγκεκριμένο κατώφλι, τότε παράγεται ένα *δυναμικό ενέργειας*.
- **Νευράξονας (axon):** Μια επιμήκης ίνα που μεταδίδει το δυναμικό ενέργειας σε μεγάλες αποστάσεις. Ο νευράξονας μπορεί να είναι μυελινοποιημένος, γεγονός που αυξάνει σημαντικά την ταχύτητα αγωγής.
- **Νευρικές απολήξεις (axon terminals):** Στο άκρο του νευράξονα, το σήμα μεταδίδεται σε άλλους νευρώνες μέσω *συναπτικής μετάδοσης*. Η μετάδοση αυτή γίνεται με απελευθέρωση νευροδιαβιβαστών στο συναπτικό χάσμα, οι οποίοι συνδέονται με υποδοχείς στους δενδρίτες του επόμενου νευρώνα.

Το δυναμικό ενέργειας (*action potential*) είναι ένα παροδικό ηλεκτρικό σήμα που διαρκεί μερικά χιλιοστά του δευτερολέπτου και διαδίδεται κατά μήκος του νευράξονα. Η δυαδική φύση του (πυροδότηση ή μη) αποτέλεσε βασική έμπνευση για την αφαιρετική μοντελοποίηση των τεχνητών νευρώνων ως μονάδων που ενεργοποιούνται όταν το άθροισμα των εισόδων ξεπερνά κάποιο κατώφλι.

Μία από τις σημαντικότερες αρχές που προέρχονται από τη βιολογία είναι η αρχή του Hebb, γνωστή και ως «Hebbian learning»: «*νευρώνες που πυροδοτούν μαζί, συνδέονται μεταξύ τους*» [17]. Η αρχή αυτή περιγράφει τον μηχανισμό ενίσχυσης των συνάψεων που ενεργοποιούνται ταυτόχρονα, δημιουργώντας ισχυρότερες συνδέσεις μεταξύ νευρώνων που συν-ενεργοποιούνται συχνά. Ο κανόνας αυτός αποτέλεσε έμπνευση για τους πρώιμους αλγόριθμους μάθησης στα τεχνητά νευρωνικά δίκτυα.

Τέλος, αξίζει να σημειωθεί ότι οι βιολογικοί νευρώνες εμφανίζουν μεγάλη ποικιλία μορφολογίας και λειτουργίας. Υπάρχουν αισθητικοί (afferent) νευρώνες που μεταφέρουν πληροφορία από το περιβάλλον, κινητικοί (efferent) νευρώνες που κατευθύνουν τις αντιδράσεις, καθώς και ενδιάμεσοι νευρώνες (interneurons) που επιτελούν σύνθετη επεξεργασία μέσα στο ΚΝΣ. Η ποικιλομορφία αυτή υπενθυμίζει ότι τα ΤΝΔ είναι απλουστευμένα μοντέλα που επιχειρούν να μιμηθούν μόνο βασικές αρχές και όχι την πλήρη πολυπλοκότητα της βιολογίας.



Σχήμα 3.1: Δομή βιολογικού νευρώνα: δενδρίτες (είσοδοι), κυτταρικό σώμα, νευράξονας και νευρικές απολήξεις [18].

3.3 Διαισθητική ανάλυση

Παρά τη θηριώδη επιτυχία των σύγχρονων ΤΝΔ, τα βασικά δομικά τους στοιχεία παραμένουν εξαιρετικά απλά. Για να χτίσουμε μια πρώτη διαίσθηση, σε αυτή την ενότητα θα ακολουθήσουμε μια πορεία από ένα ελάχιστο παράδειγμα προς τις βασικές έννοιες, χωρίς να χαθούμε σε περίπλοκο συμβολισμό. Ο στόχος είναι να κατανοηθεί γιατί η εκπαίδευση μέσω παραδειγμάτων μπορεί να οδηγήσει σε μοντέλα με ικανότητα γενίκευσης.

3.3.1 Μοντελοποίηση προβλήματος

Ας υποθέσουμε ότι θέλουμε να διακρίνουμε δύο απλές κατηγορίες «εικόνων» μεγέθους 2×2 με δυαδικά pixel (0 = λευκό, 1 = μαύρο). Για παράδειγμα:

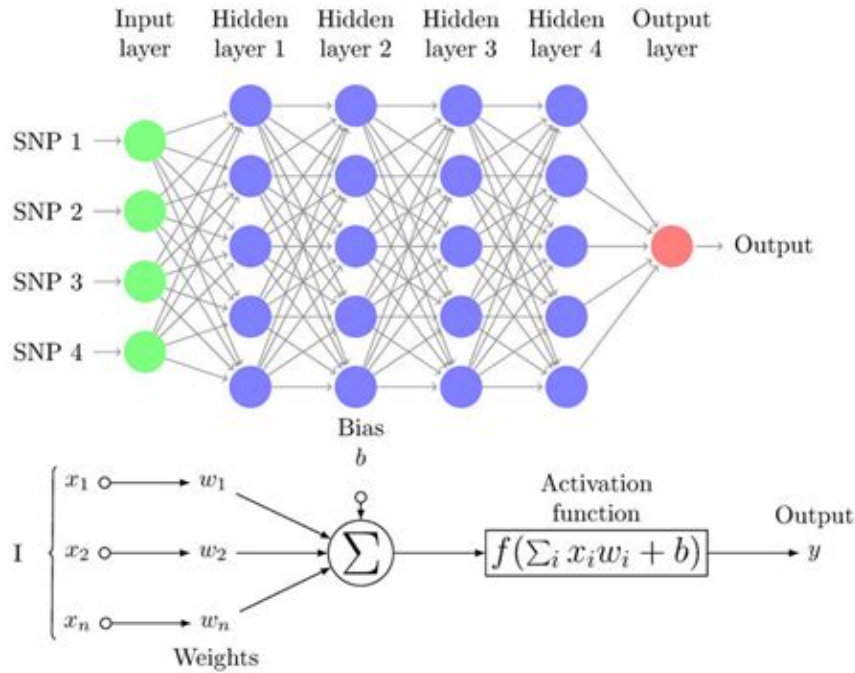
$$\mathbf{X}_a = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{X}_b = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}, \quad \mathbf{y}_a = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathbf{y}_b = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Εδώ η $\mathbf{y}$ είναι ένα διανυσματικό *label* που δηλώνει σε ποια κατηγορία ανήκει η εικόνα. Θέλουμε λοιπόν να κατασκευάσουμε μια συνάρτηση $f: \mathbb{R}^{2 \times 2} \rightarrow \mathbb{R}^2$, τέτοια ώστε $f(\mathbf{X}_a) = \mathbf{y}_a$ και $f(\mathbf{X}_b) = \mathbf{y}_b$.

Για να δουλέψουμε πιο απλά μαθηματικά, μπορούμε να «ισοπεδώσουμε» (flatten) κάθε εικόνα σε διάνυσμα $\mathbf{x} \in \mathbb{R}^4$:

$$\mathbf{x}_a = [1, 0, 0, 1], \quad \mathbf{x}_b = [0, 1, 0, 1].$$

Ένα πρώτο υποψήφιο μοντέλο είναι ένα γραμμικό μετασχηματισμό $\mathbf{x} \mapsto \mathbf{x}\mathbf{W}$, όπου $\mathbf{W} \in \mathbb{R}^{4 \times 2}$ είναι ένας πίνακας βαρών. Κάθε στοιχείο του $\mathbf{W}$ αντιστοιχεί στη «δύναμη» με την οποία ένα pixel επηρεάζει την απόφαση υπέρ της μίας ή της άλλης κατηγορίας. Γραφικά, αυτό μπορεί να απεικονιστεί ως ένα στρώμα με τέσσερις εισόδους και δύο εξόδους (Εικ. 3.2).



Σχήμα 3.2: Γραμμικό μοντέλο με 4 εισόδους και 21 εξόδο. Κάθε σύνδεση αντιστοιχεί σε ένα βάρος.

Η ιδέα είναι ότι αν επιλέξουμε σωστά τα βάρη, τότε οι εικόνες των δύο κατηγοριών θα διαχω-

ρίζονται σωστά. Στην πράξη όμως, όταν τα δεδομένα αυξάνονται εκθετικά και η διάσταση μεγαλώνει, η εύρεση όλων των βαρών με «κλειστή μορφή» (όπως σε μια εξίσωση γραμμικής παλινδρόμησης) καθίσταται ασύμφορη. Για τον λόγο αυτό χρησιμοποιούμε *εκπαίδευση με κλίση και οπισθοδιάδοση* (*backpropagation*).

3.3.2 Backpropagation: μια αριθμητική ματιά

Η εκπαίδευση ξεκινά με τυχαία αρχικοποίηση των βαρών. Έστω ότι για την είσοδο $\mathbf{x}_a = [1, 0, 0, 1]$ το δίκτυο υπολογίζει έξοδο $\hat{\mathbf{y}}_a = [1.8, 0.4]^\top$. Για να μετρήσουμε την απόκλιση από το επιθυμητό αποτέλεσμα, ορίζουμε μια συνάρτηση κόστους, π.χ. το μέσο τετραγωνικό σφάλμα (MSE):

$$\mathcal{L}(\hat{\mathbf{y}}, \mathbf{y}) = \frac{1}{2} \sum_i (\hat{y}_i - y_i)^2.$$

Για το δείγμα a έχουμε:

$$\mathcal{L}_a = \frac{1}{2} ((1.8 - 1)^2 + (0.4 - 0)^2) = \frac{1}{2} (0.64 + 0.16) = 0.40.$$

Το επόμενο βήμα είναι να υπολογίσουμε τις παραγώγους της συνάρτησης κόστους ως προς τα βάρη, ώστε να ξέρουμε προς ποια κατεύθυνση πρέπει να τα μεταβάλλουμε. Για παράδειγμα, η παράγωγος της $\mathcal{L}$ ως προς ένα βάρος που συμμετέχει στην έξοδο $\hat{y}_1$ είναι απλώς $(\hat{y}_1 - y_1)$, δηλαδή 0.8. Αν τροποποιήσουμε το βάρος στην αντίθετη κατεύθυνση με ένα μικρό *learning rate*, το σφάλμα θα μειωθεί. Η διαδικασία αυτή επαναλαμβάνεται για όλα τα βάρη και όλα τα δείγματα, οδηγώντας σταδιακά στη βελτιστοποίηση του μοντέλου.

Η μέθοδος της οπισθοδιάδοσης είναι εξαιρετικά αποδοτική επειδή υπολογίζει συστηματικά τις κλίσεις σε όλα τα επίπεδα του δικτύου, από την έξοδο προς τις εισόδους, επιτρέποντας την εκπαίδευση ακόμη και πολύπλοκων πολυεπίπεδων δικτύων. Παρόλο που το παράδειγμά μας είναι απλοποιημένο, οι ίδιες βασικές αρχές ισχύουν για τα βαθιά νευρωνικά δίκτυα που χρησιμοποιούνται σε εφαρμογές αιχμής, όπως η αναγνώριση εικόνων και η τμηματοποίηση MRI.

3.4 Φορμαλιστική ανάλυση

Η διαισθητική εικόνα αποδίδει τον πυρήνα των ΤΝΔ. Για πληρότητα, ακολουθεί ο τυπικός φορμαλισμός, όπως χρησιμοποιείται στη θεωρητική περιγραφή και στην εκπαίδευση νευρωνικών δικτύων.

3.4.1 Μοντέλο νευρώνα

Ένας τεχνητός νευρώνας λαμβάνει ως είσοδο ένα διάνυσμα σημάτων $\mathbf{x}$, καθένα από τα οποία σταθμίζεται με το αντίστοιχο βάρος w_{ij} και αθροίζεται μαζί με έναν όρο μετατόπισης (*bias*) b_j . Στη συνέχεια, το αποτέλεσμα διέρχεται από μία *συνάρτηση ενεργοποίησης* $\sigma(\cdot)$, η οποία εισάγει μη γραμμικότητα στο μοντέλο:

$$a_j = \sigma\left(\sum_i w_{ij}x_i + b_j\right). \quad (3.1)$$

Η μη γραμμικότητα είναι κρίσιμη: χωρίς αυτή, ακόμη και βαθιά δίκτυα θα μπορούσαν να εκφράσουν μόνο γραμμικούς μετασχηματισμούς. Συνήθεις επιλογές ενεργοποίησης είναι η sigmoid, η tanh και η ReLU, καθεμία με τα δικά της πλεονεκτήματα και περιορισμούς [19, 20].

3.4.2 Perceptron πολλών επιπέδων (MLP)

Στα πολυεπίπεδα δίκτυα οι νευρώνες οργανώνονται σε διαδοχικά στρώματα. Αν $\mathbf{a}^{\ell-1}$ είναι οι ενεργοποιήσεις του στρώματος $\ell - 1$, τότε το στρώμα ℓ υπολογίζει:

$$\mathbf{z}^\ell = \mathbf{a}^{\ell-1} \mathbf{W}^\ell + \mathbf{b}^\ell, \quad (3.2)$$

$$\mathbf{a}^\ell = \sigma(\mathbf{z}^\ell). \quad (3.3)$$

Η εμπρόσθια διάδοση (forward pass) προκύπτει από επανειλημμένη εφαρμογή των (3.2)–(3.3). Ένα κλασικό παράδειγμα είναι η *MNIST* [21, 22], όπου οι 28×28 εικόνες ισοπεδώνονται σε 784-διάστατα διανύσματα και περνούν από ένα ή περισσότερα κρυφά στρώματα, με τελική έξοδο 10 μονάδες (μία ανά ψηφίο). Η πρόβλεψη δίνεται από $\arg \max_i \hat{y}_i$, δηλαδή από την κατηγορία με τη μέγιστη πιθανότητα.

3.4.3 Gradient Descent

Για την εκπαίδευση ενός δικτύου ορίζουμε μια συνάρτηση κόστους $\mathcal{C}(\Theta)$ ως προς όλες τις παραμέτρους $\Theta = \{\mathbf{W}^\ell, \mathbf{b}^\ell\}_\ell$. Η *Κάθοδος Κλίσης* ενημερώνει επαναληπτικά τις παραμέτρους μειώνοντας το κόστος:

$$\Theta \leftarrow \Theta - \eta \nabla_\Theta \mathcal{C}, \quad (3.4)$$

όπου $\eta > 0$ είναι ο ρυθμός μάθησης [23]. Παρά την πιθανότητα ύπαρξης τοπικών ακροτάτων, η μέθοδος αποδίδει εξαιρετικά στην πράξη, ιδιαίτερα όταν χρησιμοποιούνται στοχαστικές παραλλαγές (Stochastic Gradient Descent, mini-batches). Αυτές επιτρέπουν την εκπαίδευση μεγάλων δικτύων σε τεράστια σύνολα δεδομένων με αποδεκτό χρόνο.

3.4.4 Backpropagation

Ο αλγόριθμος *οπισθοδιάδοσης* (backpropagation) αποτελεί την καρδιά της εκπαίδευσης βαθιών δικτύων. Χάρη στη συστηματική εφαρμογή του κανόνα αλυσίδας, υπολογίζει αποδοτικά τις κλίσεις όλων των βαρών και των bias σε κάθε στρώμα [24].

Έστω ότι για ένα δείγμα η συνάρτηση κόστους είναι MSE:

$$\mathcal{C} = \frac{1}{2} \sum_j (y_j - a_j^L)^2, \quad (3.5)$$

όπου L το στρώμα εξόδου και y η επιθυμητή έξοδος. Η οπισθοδιάδοση ξεκινά από το σφάλμα στο στρώμα εξόδου και προχωρά σταδιακά προς τα πίσω, υπολογίζοντας για κάθε στρώμα έναν *σφάλμα-όρο* δ^ℓ που εκφράζει την «ευθύνη» του στρώματος για το συνολικό σφάλμα. Οι κλίσεις προκύπτουν τελικά ως:

$$\frac{\partial \mathcal{C}}{\partial w_{jk}^\ell} = \delta_j^\ell a_k^{\ell-1}, \quad \frac{\partial \mathcal{C}}{\partial b_j^\ell} = \delta_j^\ell.$$

Ο αλγόριθμος συνοψίζεται στον Αλγ. 1. Παρά την απλότητά του, αποτέλεσε σημείο καμπής στη δεκαετία του 1980, διότι έκανε δυνατή την εκπαίδευση πολυεπίπεδων δικτύων με πρακτικό χρόνο υπολογισμού.

Σχόλιο για τις αρχιτεκτονικές

Ο φορμαλισμός που παρουσιάστηκε ισχύει όχι μόνο για τα κλασικά MLP, αλλά και για πιο προηγμένες αρχιτεκτονικές (συνελικτικά, αναδρομικά, transformers). Οι ίδιες αρχές εμπρόσθιας διάδοσης, κό-

Algorithm 1 Backpropagation με MSE

Require: δείγματα $(\mathbf{x}, \mathbf{y})$, βάρη $\{\mathbf{W}^\ell\}$, πρόσθετα $\{\mathbf{b}^\ell\}$, ρυθμός μάθησης η

- 1: **Εμπρόσθια διάδοση:** Υπολόγισε $\mathbf{z}^\ell = \mathbf{a}^{\ell-1} \mathbf{W}^\ell + \mathbf{b}^\ell$, $\mathbf{a}^\ell = \sigma(\mathbf{z}^\ell)$ για κάθε ℓ .
- 2: **Σφάλμα εξόδου:** $\delta^L = (\mathbf{a}^L - \mathbf{y}) \odot \sigma'(\mathbf{z}^L)$.
- 3: **Οπισθοδιάδοση:** Για $\ell = L - 1, \dots, 1$, $\delta^\ell = (\delta^{\ell+1} \mathbf{W}^{\ell+1\top}) \odot \sigma'(\mathbf{z}^\ell)$.
- 4: **Κλίσεις:** $\frac{\partial \mathcal{C}}{\partial \mathbf{W}^\ell} = \mathbf{a}^{\ell-1\top} \delta^\ell$, $\frac{\partial \mathcal{C}}{\partial \mathbf{b}^\ell} = \delta^\ell$.
- 5: **Ενημέρωση:** $\mathbf{W}^\ell \leftarrow \mathbf{W}^\ell - \eta \frac{\partial \mathcal{C}}{\partial \mathbf{W}^\ell}$, $\mathbf{b}^\ell \leftarrow \mathbf{b}^\ell - \eta \frac{\partial \mathcal{C}}{\partial \mathbf{b}^\ell}$.

στους και οπισθοδιάδοσης εφαρμόζονται με κατάλληλες παραλλαγές. Αυτό καθιστά το κεφάλαιο απαραίτητη γέφυρα προς τις ειδικές αρχιτεκτονικές που θα παρουσιαστούν στη συνέχεια.

3.5 Σύνοψη κεφαλαίου

Στο παρόν κεφάλαιο εξετάσαμε την πορεία από τη βιολογική έμπνευση των νευρωνικών δικτύων έως τον τυπικό μαθηματικό τους φορμαλισμό. Ξεκινήσαμε με τη σύνδεση ανάμεσα στους βιολογικούς νευρώνες και τους πρώτους τεχνητούς νευρώνες, αναδεικνύοντας πώς βασικές αρχές της νευροφυσιολογίας, όπως η αρχή του Hebb, ενέπνευσαν τη δημιουργία απλουστευμένων υπολογιστικών μοντέλων. Στη συνέχεια, μέσα από ένα απλό παραδειγματικό σενάριο, παρουσιάσαμε πώς ένα γραμμικό μοντέλο μπορεί να διακρίνει δεδομένα και πώς η διαδικασία εκπαίδευσης με σφάλμα και κλίσεις καθιστά δυνατή τη βελτίωση των παραμέτρων.

Ακολούθως, εισάγαμε τον τυπικό φορμαλισμό των πολυεπίπεδων perceptron (MLP), με τη σαφή διάκριση εμπρόσθιας διάδοσης, κόστους και οπισθοδιάδοσης. Εξηγήσαμε τον ρόλο της μη γραμμικότητας μέσω συναρτήσεων ενεργοποίησης, τη σημασία της καθόδου κλίσης και τις πρακτικές παραλλαγές της (στοχαστική εκπαίδευση, mini-batches). Ιδιαίτερη έμφαση δόθηκε στον αλγόριθμο της οπισθοδιάδοσης, ο οποίος, μέσω του κανόνα αλυσίδας, καθιστά εφικτή την αποδοτική εκπαίδευση βαθιών δικτύων.

Με αυτόν τον τρόπο, το κεφάλαιο κάλυψε:

- τη διαισθητική μετάβαση από τη βιολογία στον τεχνητό νευρώνα,
- την ανάπτυξη απλών γραμμικών μοντέλων και τη γενίκευσή τους σε πολυεπίπεδα δίκτυα,
- την αριθμητική ιδέα της εκπαίδευσης με μετρική MSE και κλίσεις,
- τον τυπικό φορμαλισμό εμπρόσθιας διάδοσης, κόστους και οπισθοδιάδοσης,
- και την πρακτική σημασία της καθόδου κλίσης και των υπερπαραμέτρων.

Η κατανόηση αυτών των εννοιών αποτελεί θεμέλιο για τις επόμενες ενότητες. Στα επόμενα κεφάλαια θα επικεντρωθούμε σε ειδικές αρχιτεκτονικές που έχουν σχεδιαστεί για δεδομένα εικόνας, όπως τα *Συνελκτικά Νευρωνικά Δίκτυα (CNNs)* και οι *U-Net*, τα οποία αποτελούν σήμερα τα κυριότερα εργαλεία για την τμηματοποίηση βιοϊατρικών εικόνων και ειδικότερα MRI εγκεφάλου.

Κεφάλαιο 4: Παρουσίαση των Μοντέλων που Συγκρίνουμε

Σε αυτό το κεφάλαιο παρουσιάζονται τα μοντέλα τμηματοποίησης εγκεφαλικών όγκων που θα συγκριθούν πειραματικά στα επόμενα κεφάλαια. Ξεκινάμε με μια σύντομη αλλά στοχευμένη ανακεφαλαίωση των βασικών δομικών λίθων των Συνελικτικών Νευρωνικών Δικτύων (CNN)—*kernel*, *padding*, *stride*, *pooling*—και κατόπιν περιγράφουμε τις αρχιτεκτονικές **U-Net**, **V-Net**, **SegResNet** και **SwinUNETR**, με έμφαση στις σχεδιαστικές επιλογές που επηρεάζουν την κλινικά σημαντική απόδοση, την υπολογιστική πολυπλοκότητα και την πρακτική εκπαίδευση σε δεδομένα MRI. Η παρουσίαση ακολουθεί το ύφος και τη σημειογραφία των προηγούμενων κεφαλαίων ώστε το κείμενο να παραμένει ενιαίο και συνεκτικό.

4.1 Συνοπτική θεωρία CNN: *kernel*, *padding*, *pooling*

Τα Συνελικτικά Νευρωνικά Δίκτυα (Convolutional Neural Networks, CNNs) αποτέλεσαν μία από τις σημαντικότερες εξελίξεις στη μηχανική μάθηση, καθώς επέτρεψαν την αποδοτική επεξεργασία δεδομένων με χωρική ή/και χρονική δομή [lecun1998gradient]. Η βασική τους αρχή είναι ότι η εξαγωγή χαρακτηριστικών δεν γίνεται πλέον από πλήρως συνδεδεμένα στρώματα με τεράστιο αριθμό παραμέτρων, αλλά μέσω τοπικών «φίλτρων» (kernels) που εφαρμόζονται σε μικρές περιοχές της εισόδου. Με αυτόν τον τρόπο, αξιοποιείται η τοπικότητα και η επαναληπτικότητα των μοτίβων, μειώνεται η υπολογιστική πολυπλοκότητα και επιτυγχάνεται ιεραρχική αναπαράσταση.

Συνέλιξη και σχήμα εξόδου

Στην πράξη, ένα φίλτρο $k \times k$ (ή $k_D \times k_H \times k_W$ στη 3D περίπτωση) «σαρώνει» την είσοδο με καθορισμένο βήμα (*stride* s), ενώ μπορεί να χρησιμοποιηθεί *padding* p ώστε να προστεθούν μηδενικές τιμές στα όρια και να ελεγχθεί το μέγεθος εξόδου. Το σχήμα εξόδου καθορίζεται από τη σχέση:

$$H_{\text{out}} = \left\lfloor \frac{H + 2p - d \cdot (k - 1) - 1}{s} + 1 \right\rfloor, \quad W_{\text{out}} = \left\lfloor \frac{W + 2p - d \cdot (k - 1) - 1}{s} + 1 \right\rfloor,$$

με ανάλογη έκφραση στην τρισδιάστατη περίπτωση. Η επιλογή *same padding* διατηρεί τις διαστάσεις για $s = 1$, ενώ το *valid padding* οδηγεί σε μείωση.

Παράμετροι και υπολογιστικό κόστος

Ένας συνελικτικός πυρήνας με C_{in} κανάλια εισόδου και C_{out} κανάλια εξόδου έχει:

$$\#_{\text{par}} = C_{\text{out}} \cdot C_{\text{in}} \cdot k^2 + C_{\text{out}},$$

στη δισδιάστατη περίπτωση, ή k^3 στην τρισδιάστατη. Ο υπολογιστικός φόρτος ανά εμπρός διέλευση εξαρτάται από το μέγεθος εξόδου και το πλήθος των καναλιών, και προσεγγιστικά είναι:

$$\text{FLOPs} \approx 2 \cdot H_{\text{out}} W_{\text{out}} \cdot C_{\text{out}} \cdot (C_{\text{in}} \cdot k^2),$$

όπου ο παράγοντας 2 αντιστοιχεί σε πολλαπλασιασμό και πρόσθεση. Είναι προφανές ότι το κόστος αυξάνεται σημαντικά στην 3D MRI, όπου ο όγκος δεδομένων είναι πολύ μεγαλύτερος.

Pooling, upsampling και *transposed convolution*

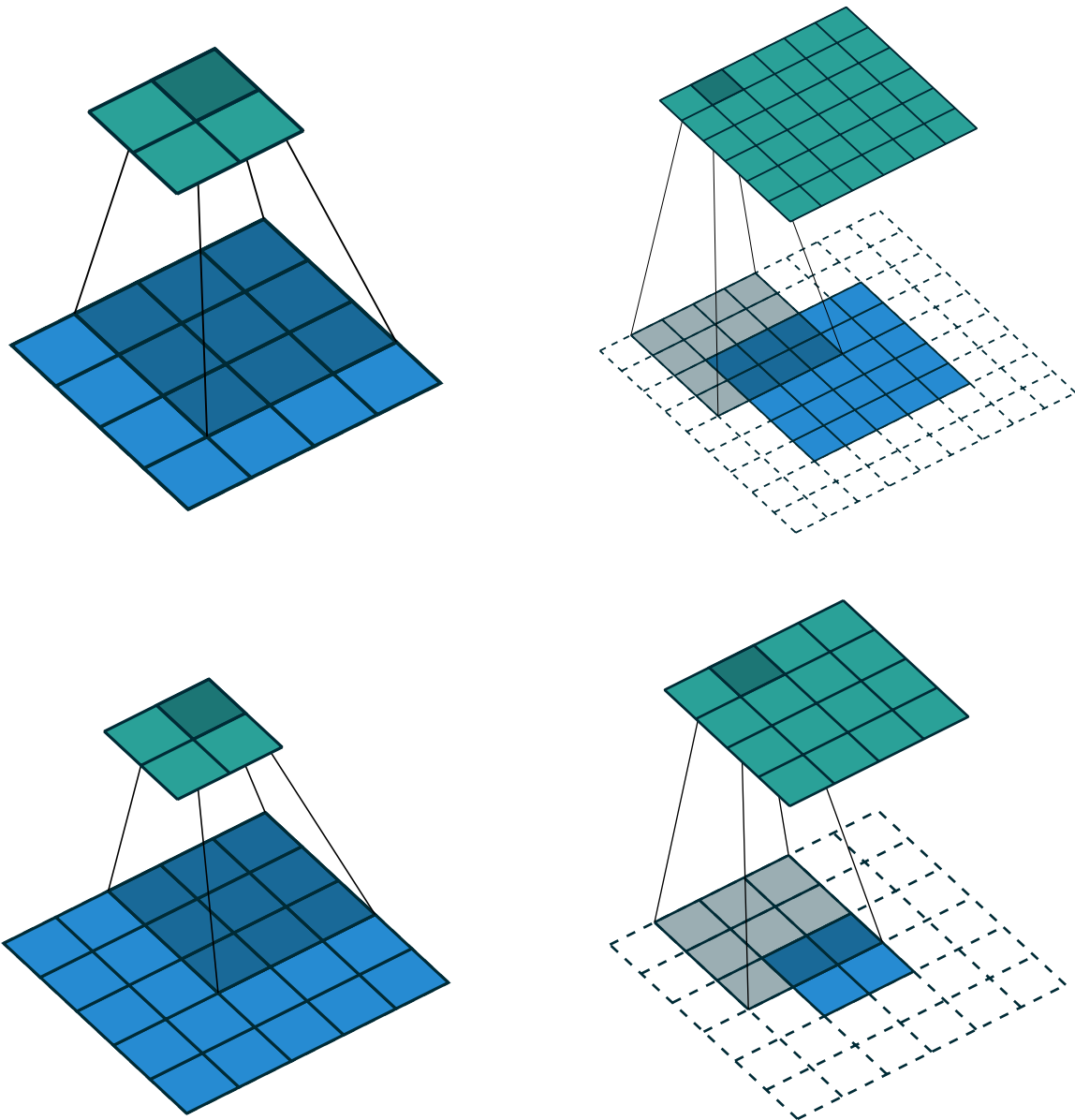
Για τη μείωση της διάστασης και την αύξηση της αφαίρεσης χρησιμοποιούνται τελεστές *pooling*, όπως το *max pooling* ή το *average pooling*, που συρρικνώνουν το μέγεθος συνήθως κατά $2\times$. Η ανάκτηση ανάλυσης στον *decoder* γίνεται είτε με *upsampling* (συχνά ακολουθούμενο από 1×1 ή 3×3 συνέλιξη) είτε με *transposed convolution*, όπου τα βάρη εκπαιδεύονται ώστε να μάθουν την ανορθωτική διαδικασία. Οι επιλογές αυτές έχουν άμεση επίδραση στην ποιότητα της τμηματοποίησης MRI, καθώς επηρεάζουν την ομαλότητα και την ακρίβεια των ορίων του όγκου.

Κανονικοποίηση και ενεργοποιήσεις

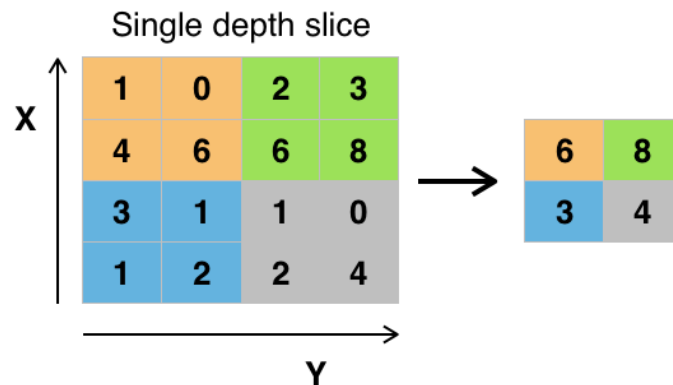
Για να σταθεροποιηθεί η εκπαίδευση και να επιταχυνθεί η σύγκλιση, εφαρμόζονται τεχνικές κανονικοποίησης. Η *Batch Normalization* [ioffe2015batch] είναι ευρέως διαδεδομένη σε δισδιάστατες εφαρμογές, αλλά στη 3D επεξεργασία συχνά προτιμώνται η *Instance Normalization* ή η *Group Normalization*, επειδή τα μεγέθη batch είναι περιορισμένα λόγω μνήμης. Σε συνδυασμό, οι μη γραμμικές συναρτήσεις ενεργοποίησης (ReLU, LeakyReLU, GELU) επιτρέπουν την εκμάθηση πολύπλοκων μη γραμμικών σχέσεων. Η επιλογή τους επηρεάζει άμεσα τόσο την ταχύτητα εκπαίδευσης όσο και την τελική απόδοση.

Πεδίο αποδοχής (*receptive field*)

Το *effective receptive field* περιγράφει το τμήμα της εισόδου που «βλέπει» μια μονάδα εξόδου. Με την αύξηση των επιπέδων, το πεδίο αποδοχής μεγαλώνει αθροιστικά, και μπορεί να διευρυνθεί περαιτέρω με χρήση *dilated convolutions* ή με αρχιτεκτονικές U-shape που συνδυάζουν πληροφορία διαφορετικών κλιμάκων. Η εξασφάλιση επαρκούς συμφραζόμενου είναι κρίσιμη στη διάγνωση MRI εγκεφάλου, καθώς η απόφαση ανά voxel δεν πρέπει να εξαρτάται μόνο από τοπικές εντάσεις αλλά και από τη σχέση με τον ευρύτερο ιστό.



Σχήμα 4.1: Δημοφιλείς επιλογές σε k , p , s και *upsampling* με *transposed conv*. Πηγή: Dumoulin & Visin (MIT).[25] Επιπλέον *max pooling* στο Σχ. 4.2.



Σχήμα 4.2: Παράδειγμα *max pooling* 2×2 , $s=2$. Πηγή: Aphex34, Wikimedia Commons, CC BY-SA 4.0.[26]

4.2 U-Net: U-shape & skip connections

Ιστορική αναδρομή και βασική ιδέα

Η αρχιτεκτονική *U-Net* προτάθηκε αρχικά από τους Ronneberger et al. [ronneberger2015unet] για την τμηματοποίηση βιοϊατρικών εικόνων μικροσκοπίας και έκτοτε έχει εξελιχθεί σε σημείο αναφοράς για πλήθος εφαρμογών. Η καινοτομία της έγκειται στον συνδυασμό ενός συμμετρικού *encoder-decoder* με *skip connections*, που σχηματίζουν το χαρακτηριστικό σχήμα “U”. Το αποτέλεσμα είναι ένα δίκτυο που μπορεί να μαθαίνει τόσο υψηλού επιπέδου χαρακτηριστικά (από τον encoder) όσο και λεπτομερείς χωρικές πληροφορίες (από τα *skip connections*), επιτυγχάνοντας ακριβή οριοθέτηση δομών ακόμη και σε εικόνες με περιορισμένα δεδομένα εκπαίδευσης.

Ιδεατή σύνοψη

Η U-Net αποτελείται από δύο βασικά τμήματα:

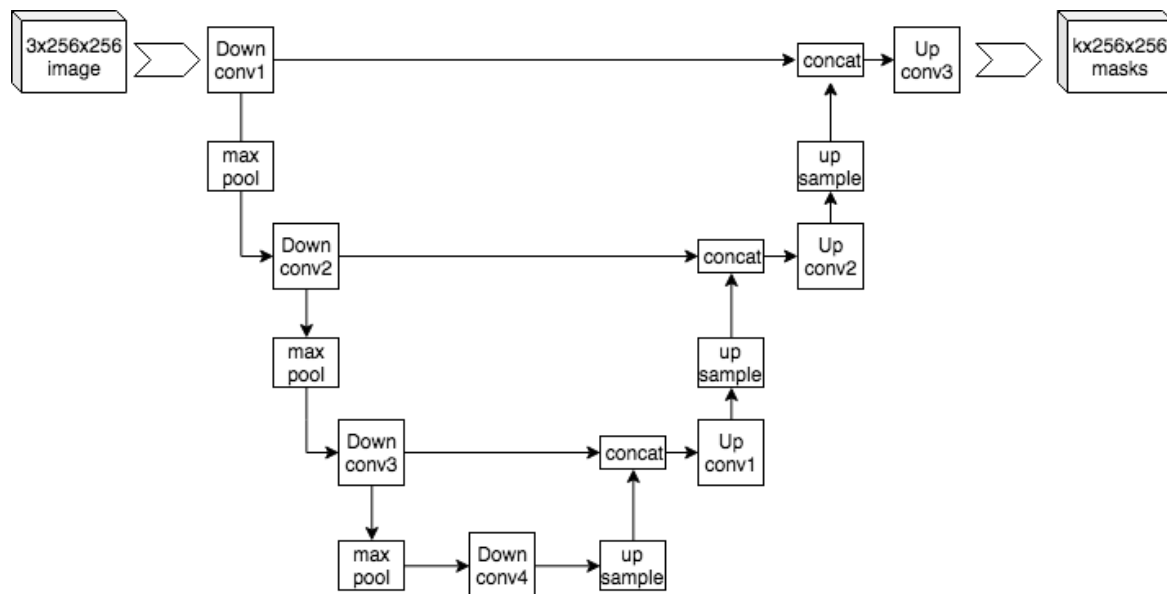
- **Encoder:** διαδοχικές συνελίξεις με υποδειγματοληψία (*max pooling* ή *strided conv*), που μειώνουν το μέγεθος και αυξάνουν τα κανάλια, επιτρέποντας την εξαγωγή όλο και πιο αφηρημένων χαρακτηριστικών.
- **Decoder:** συμμετρική ανωδειγματοληψία (*transposed conv* ή *interpolate+conv*) που ανακτά τη χωρική ανάλυση, συνδυάζοντας τα υψηλού επιπέδου χαρακτηριστικά με λεπτομέρειες από τα *skip connections*.

Τα *skip connections* (συνήθως με *concatenation*) μεταφέρουν τις λεπτομέρειες από τα πρώιμα, υψηλής ανάλυσης επίπεδα του encoder στον decoder. Με αυτόν τον τρόπο αντισταθμίζεται η απώλεια χωρικής πληροφορίας που προκαλείται από την υποδειγματοληψία, επιτρέποντας στο δίκτυο να αποδίδει ακριβή όρια στις τμηματοποιήσεις.

Σχεδιαστικές επιλογές

- **Βάθος και κανάλια:** Το πλάτος καναλιών διπλασιάζεται συνήθως σε κάθε βήμα προς το bottleneck (π.χ. $32 \rightarrow 64 \rightarrow 128 \rightarrow 256$), και μειώνεται συμμετρικά στον decoder.
- **Down/Up-sampling:** Η κάτω-δειγματοληψία γίνεται με *max pooling* ή συνελίξεις με $s=2$, ενώ η ανωδειγματοληψία με *transposed conv* ή με απλή παρεμβολή (*bilinear/trilinear*) ακολουθούμενη από συνέλιξη.

- **Συναρτήσεις κόστους:** Για προβλήματα ανισόρροπων κλάσεων χρησιμοποιούνται Dice Loss, Generalized Dice, Cross-Entropy ή συνδυασμοί τους (π.χ. Dice+CE).
- **2D vs 3D εκδοχές:** Η 2D U-Net εφαρμόζεται σε τομές ($H \times W$) και είναι οικονομική σε μνήμη, ενώ η 3D U-Net επεξεργάζεται ολόκληρο τον όγκο ($D \times H \times W$), εκμεταλλευόμενη το πλήρες συμφοραζόμενο αλλά με σημαντικά μεγαλύτερες απαιτήσεις σε μνήμη και χρόνο.



Σχήμα 4.3: U-Net: συμμετρικός *encoder-decoder* με *skip connections*. Πηγή: Mehrdad Yazdani, CC BY-SA 4.0 [27].

Πλεονεκτήματα για MRI. Η U-Net έχει αποδειχθεί ιδιαίτερα αποτελεσματική στην τμηματοποίηση MRI εγκεφάλου, καθώς συνδυάζει ακριβή οριοθέτηση με σταθερή εκπαίδευση. Στα δεδομένα BraTS, η απλή προσαρμογή των καναλιών εισόδου ώστε να δέχεται τα τέσσερα modalities (T1, T1c, T2, FLAIR) αρκεί για να επιτύχει υψηλές επιδόσεις. Επιπλέον, η U-Net διαθέτει μεγάλη ευελιξία: μπορεί να προσαρμοστεί εύκολα σε διαφορετικές διαστάσεις (2D/3D) και έχει εμπνεύσει πλήθος παραλλαγών (Residual U-Net, Attention U-Net, nnU-Net), που βελτιώνουν περαιτέρω την ακρίβεια ή την αποδοτικότητα.

Συνοψίζοντας

Η U-Net αποτελεί την πιο καθιερωμένη γραμμή βάσης για την τμηματοποίηση ιατρικών εικόνων. Ο απλός αλλά ισχυρός σχεδιασμός της (encoder-decoder με skip connections) προσφέρει έναν ισορροπημένο συνδυασμό ακρίβειας, ευρωστίας και υπολογιστικής αποδοτικότητας, καθιστώντας την αφετηρία για σύγκριση με πιο σύγχρονα ή πολύπλοκα μοντέλα.

4.3 V Net: 3D volumetric συνελίξεις

Κεντρική ιδέα

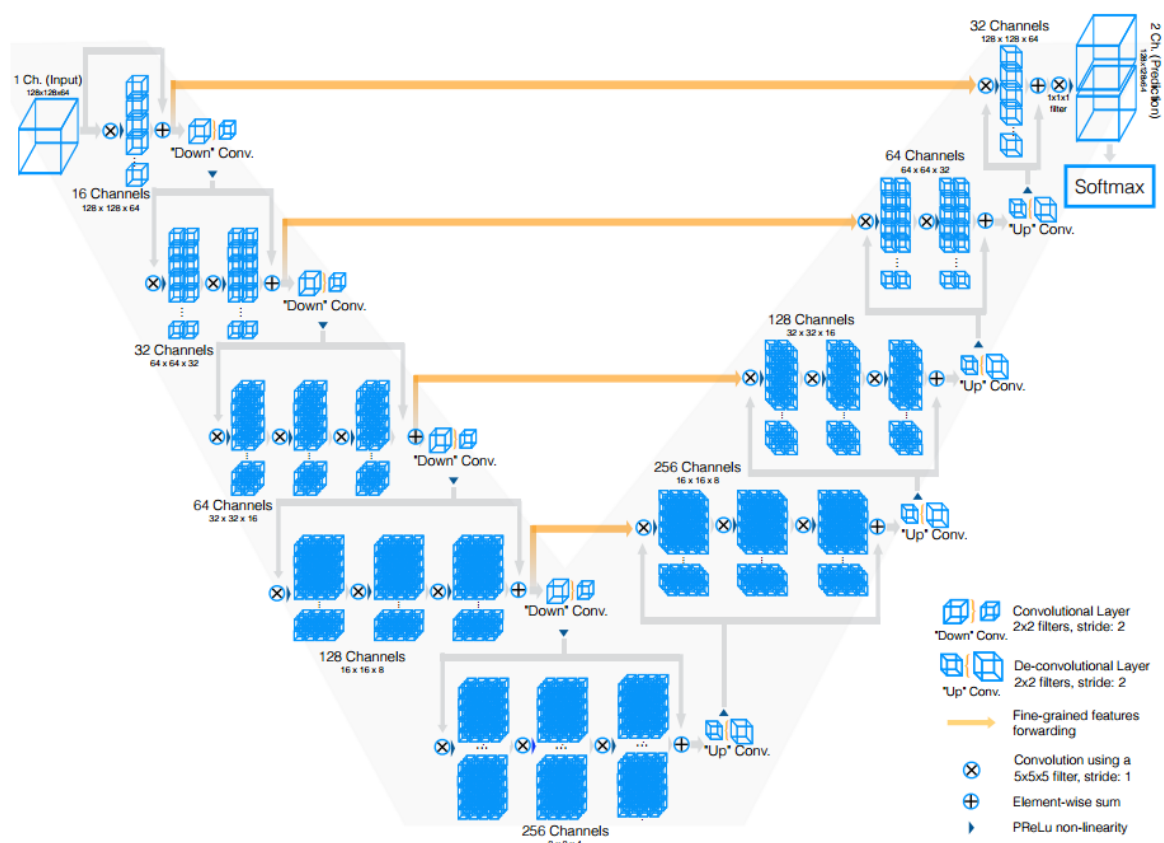
Η V-Net παρουσιάστηκε από τους Milletari et al. [28] ως μία από τις πρώτες αρχιτεκτονικές που σχεδιάστηκαν εξ αρχής για τρισδιάστατη τμηματοποίηση ιατρικών εικόνων. Σε αντίθεση με την απλή «3D-οποίηση» της U-Net, η V-Net υιοθετεί πλήρως 3D συνελίξεις και ενσωματώνει *residual* μπλοκ, τα οποία επιτρέπουν την εκπαίδευση βαθύτερων δικτύων με σταθερότερη διάδοση κλίσεων. Ο βασικός

στόχος είναι να μάθει το χωρικό συμφραζόμενο σε ολόκληρο τον όγκο, κάτι κρίσιμο στη μαγνητική τομογραφία εγκεφάλου, όπου η πληροφορία εξαπλώνεται σε γειτονικές τομές.

Αρχιτεκτονικά μοτίβα

Η V-Net ακολουθεί μια συμμετρική διάταξη encoder-decoder, με ορισμένες σημαντικές διαφοροποιήσεις:

- **Residual μπλοκ:** Κάθε στάδιο αποτελείται από $3 \times 3 \times 3$ συνελιξείς, ακολουθούμενες από συντομευμένες συνδέσεις (*shortcuts*). Αυτές οι *residual connections* συμβάλλουν στη σταθερότητα της εκπαίδευσης, μειώνοντας το πρόβλημα εξαφάνισης ή έκρηξης κλίσεων σε βαθιά δίκτυα.
- **Strided 3D conv:** Η κάτω-δειγματοληψία επιτυγχάνεται με 3D συνελιξείς με βήμα 2, χωρίς να χρειάζεται εξωτερικό pooling. Αντίστοιχα, η ανωδειγματοληψία πραγματοποιείται με 3D *transposed convolutions*, διατηρώντας τη χωρική συνοχή.
- **Dice-based loss:** Η V-Net εισήγαγε ως κύρια συνάρτηση κόστους τον *Dice loss*, ο οποίος είναι ιδιαίτερα κατάλληλος για τμηματοποίηση με έντονα ανισόρροπα δεδομένα (λίγα voxel που ανήκουν στην κλάση «όγκου» έναντι πολλών voxel υγιούς ιστού).



Σχήμα 4.4: V-Net: 3D *residual* μπλοκ και ιεραρχική χωρική κλιμάκωση. Image © NVIDIA, used under the Apache License 2.0 [29].

Πρακτικές παρατηρήσεις

Με *patches* μεγέθους $96 \times 96 \times 96$ και βασικό πλάτος 16 καναλιών, η V-Net διαθέτει περίπου 13.3 εκατομμύρια παραμέτρους (βλ. §4.8). Η χρήση 3D συνελιξεων επιτρέπει την εκμετάλλευση του πλήρους

όγκου, βελτιώνοντας τη συνέπεια μεταξύ γειτονικών τομών. Ωστόσο, το κόστος σε μνήμη και χρόνο είναι αυξημένο, κάτι που καθιστά αναγκαία τη χρήση υποδειγματοληπτικών *patches* αντί πλήρων όγκων σε υψηλή ανάλυση. Παρ' όλα αυτά, η V-Net παραμένει ιδιαίτερα ανταγωνιστική σε MRI εγκεφάλου και αποτέλεσε σημείο αναφοράς στις πρώτες εκδόσεις του BraTS challenge.

Σχόλιο. Η V-Net δείχνει πώς μια αρχιτεκτονική σχεδιασμένη εξ αρχής για 3D δεδομένα μπορεί να εκμεταλλευθεί καλύτερα τη χωρική πληροφορία σε σχέση με απλές 2D ή pseudo-3D προσεγγίσεις. Η εισαγωγή *residual connections* ήταν επίσης καθοριστική, καθώς αποτέλεσε πρόδρομο για μεταγενέστερες αρχιτεκτονικές υψηλότερης πολυπλοκότητας.

4.4 SegResNet: βαθιά *residual encoder–decoder*

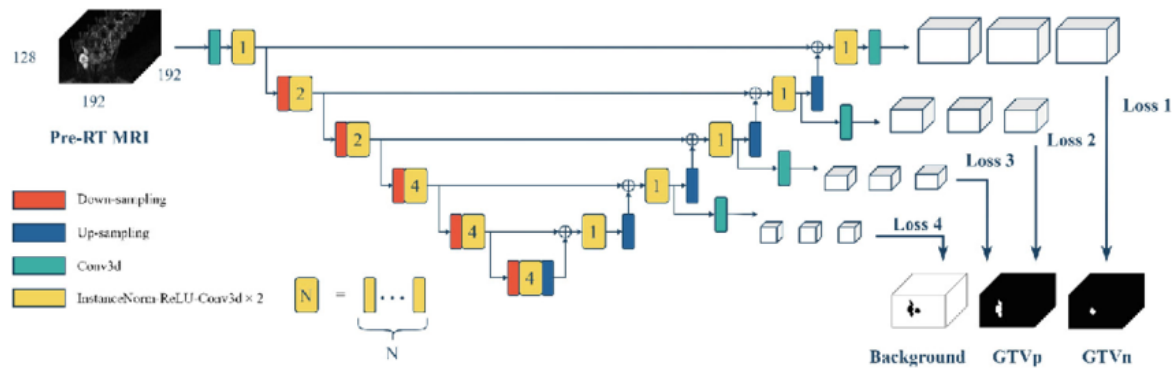
Φιλοσοφία

Η αρχιτεκτονική *SegResNet* συνδυάζει την ιδέα του U-shape *encoder–decoder* με την εκτεταμένη χρήση *residual units*, ακολουθώντας τη φιλοσοφία των ResNet (He et al., 2015). Ο στόχος είναι να επιτευχθεί μεγαλύτερο βάθος χωρίς να υποβαθμίζεται η εκπαίδευση, καθώς τα *residual connections* επιτρέπουν την απρόσκοπτη διάδοση της πληροφορίας και των κλίσεων σε πολλά επίπεδα. Έτσι, η *SegResNet* θεωρείται «βαθύτερη και πιο εκφραστική U-Net», ικανή να συλλάβει πιο σύνθετα πρότυπα σε δεδομένα υψηλής πολυπλοκότητας όπως οι MRI όγκων εγκεφάλου.

Κύρια χαρακτηριστικά

Η *SegResNet* ενσωματώνει τυπικά τα παρακάτω στοιχεία:

- **Residual units:** Κάθε μπλοκ περιλαμβάνει δύο 3×3 συνελίξεις με κανονικοποίηση και συνάρτηση ενεργοποίησης, ενώ η έξοδος προστίθεται στην αρχική είσοδο μέσω *identity* ή *projection skip connection*. Αυτό διευκολύνει τη μάθηση ταυτοτικών μετασχηματισμών και μειώνει το πρόβλημα *vanishing gradients*.
- **Down/Up-sampling:** Η κάτω-δειγματοληψία γίνεται με συνελίξεις $s=2$, ενώ η άνοδος της ανάλυσης με *transposed convolutions* ή συνδυασμούς *upsampling + conv*. Έτσι διατηρείται η συμμετρική U-shape μορφή.
- **Κανονικοποίηση:** Στις 3D εφαρμογές συνήθως χρησιμοποιείται *Instance Normalization* ή *Group Normalization*, λόγω μικρών batch sizes, σε συνδυασμό με ReLU ή LeakyReLU.
- **Deep supervision:** Σε ορισμένες εκδοχές χρησιμοποιούνται βοηθητικές εξόδους (*auxiliary outputs*) σε ενδιάμεσα επίπεδα, οι οποίες συμβάλλουν στη σταθεροποίηση της εκπαίδευσης σε πολύ βαθιά δίκτυα.



Σχήμα 4.5: SegResNet: *encoder-decoder* με *residual* μπλοκ και (όπου εφαρμόζεται) *deep supervision*.
Πηγή: Lee et al., *Scientific Reports*, 2024, Σχ. 2, άδεια CC BY 4.0 [30].

Πλεονεκτήματα και εφαρμογές

Η SegResNet προσφέρει καλύτερη εκφραστικότητα ανά δεδομένο υπολογιστικό κόστος σε σχέση με μια απλή U-Net, ιδίως όταν αυξάνεται το βάθος. Οι residual μονάδες αποτρέπουν την απώλεια πληροφορίας και βοηθούν στην αποφυγή *over-segmentation* ή *under-segmentation* σε περιοχές χαμηλής αντίθεσης, κάτι που είναι ιδιαίτερα σημαντικό στις MRI εγκεφάλου όπου οι όγκοι έχουν ασαφή όρια. Στα πλαίσια του BraTS challenge και στο MONAI framework, η SegResNet χρησιμοποιείται εκτενώς ως baseline, επιτυγχάνοντας ισχυρές επιδόσεις σε πολυτροπικά δεδομένα (T1, T1c, T2, FLAIR).

Σχόλιο. Η SegResNet μπορεί να θεωρηθεί ως «επόμενη γενιά» U-Net, η οποία ενσωματώνει τις βέλτιστες πρακτικές των ResNets για να αυξήσει το βάθος και την ικανότητα γενίκευσης. Αυτό την καθιστά ιδιαίτερα χρήσιμη όταν οι περιορισμοί μνήμης επιτρέπουν μεγαλύτερη πολυπλοκότητα και όταν το ζητούμενο είναι λεπτή τμηματοποίηση σε πολύπλοκα δεδομένα όπως οι εγκεφαλικοί όγκοι.

4.5 SwinUNETR: υβρίδιο CNN/Transformer με *window attention*

Υβριδικός *encoder*

Η αρχιτεκτονική *SwinUNETR* προτάθηκε από το Project MONAI [tang2022swinunetr] ως εξέλιξη του UNETR, συνδυάζοντας τη δύναμη των Transformers με την αποδοτικότητα των CNNs. Ο encoder βασίζεται στο *Swin Transformer* [liu2021swin], το οποίο εισάγει το μηχανισμό *shifted window self-attention* για την αποδοτική επεξεργασία 2D ή 3D patches. Η δομή ακολουθεί U-shape διάταξη: ο Transformer αναλαμβάνει την εξαγωγή μακρινών χωρικών συσχετίσεων, ενώ ο decoder, βασισμένος σε συνελίξεις, αναδομεί την εικόνα σε υψηλή ανάλυση.

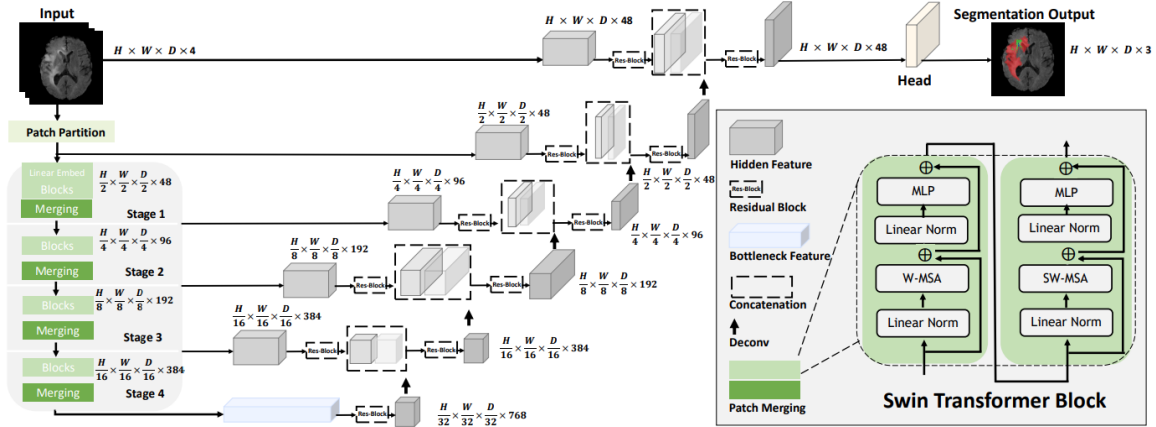
Η τοπική αυτοπροσοχή σε παράθυρα μεγέθους $M \times M$ (2D) ή $M \times M \times M$ (3D) μειώνει σημαντικά την πολυπλοκότητα: από $\mathcal{O}(N^2)$ για πλήρη global attention σε $\mathcal{O}(N \cdot M^d)$, όπου N είναι ο αριθμός voxel/tokens και d η διάσταση (2D ή 3D). Με αυτό τον τρόπο καθίσταται εφικτή η εφαρμογή Transformer-based μοντέλων σε volumetric δεδομένα, χωρίς εκρηκτική αύξηση κόστους.

Σημαντικά συστατικά

- **Patch embedding:** Ο όγκος ή η εικόνα διαιρείται σε μη επικαλυπτόμενα patches, τα οποία απεικονίζονται σε διανύσματα (tokens) και τροφοδοτούνται στον Transformer encoder.
- **Shifted windows:** Σε κάθε στάδιο, η αυτοπροσοχή εφαρμόζεται σε τοπικά παράθυρα. Η μετατό-

πιση των παραθύρων ανά μπλοκ (*shifting*) επιτρέπει την ανταλλαγή πληροφορίας μεταξύ γειτονικών περιοχών χωρίς να απαιτείται πλήρης global attention.

- **Skip fusions:** Οι αναπαραστάσεις του Transformer encoder συγχωνεύονται με τις συνελκτικές αναπαραστάσεις του decoder, ώστε να συνδυάζονται οι λεπτομέρειες χαμηλού επιπέδου με τα παγκόσμια συμφραζόμενα.



Σχήμα 4.6: SwinUNETR: *Transformer*-βασισμένος *encoder* με *window attention* και συνελκτικός *decoder*. Πηγή: Project MONAI, Apache-2.0 [31].

Κίνητρα σε MRI

Σε αντίθεση με τα αμιγώς CNN-based μοντέλα, οι μηχανισμοί *self-attention* μπορούν να μοντελοποιήσουν μακρινές εξαρτήσεις: π.χ. συμμετρίες ανάμεσα στα ημισφαίρια ή διάχυτες παθολογικές αλλοιώσεις που εκτείνονται σε μεγάλες περιοχές. Αυτό καθιστά το SwinUNETR ιδιαίτερα κατάλληλο για MRI εγκεφάλου, όπου το global context είναι συχνά απαραίτητο. Ταυτόχρονα, ο συνελκτικός decoder διατηρεί την ικανότητα για ακριβή εντοπισμό ακμών και λεπτών δομών, κάτι κρίσιμο για την τμηματοποίηση όγκων.

Σχόλιο. Το SwinUNETR αποτελεί τυπικό παράδειγμα *υβριδικής* προσέγγισης, όπου οι Transformers φέρνουν την ικανότητα global reasoning και τα CNNs διατηρούν την αποδοτικότητα και την ακρίβεια στις λεπτομέρειες. Αν και πιο απαιτητικό σε υπολογιστικούς πόρους σε σχέση με μια U-Net ή SegResNet, σε δεδομένα όπως το BraTS έχει δείξει ανταγωνιστικές έως και κορυφαίες επιδόσεις.

4.6 Παραλλαγές και επεκτάσεις U-Net

Η βασική αρχιτεκτονική U-Net έχει γνωρίσει ευρεία διάδοση στη βιοϊατρική τμηματοποίηση. Στα χρόνια που ακολούθησαν την αρχική της παρουσίαση [ronneberger2015unet], έχουν προταθεί πολυάριθμες παραλλαγές και επεκτάσεις, οι οποίες στοχεύουν στη βελτίωση της απόδοσης ή της γενίκευσης υπό διαφορετικές συνθήκες δεδομένων. Μερικές από τις πιο επιδραστικές παρουσιάζονται στη συνέχεια.

Residual U-Net

Η *Residual U-Net* ενσωματώνει residual blocks στον κορμό της αρχιτεκτονικής, αντικαθιστώντας τα κλασικά ζεύγη συνελίξεων με δομές τύπου ResNet [32]. Η προσθήκη *skip connections* εντός κάθε block διευκολύνει τη διάδοση κλίσεων και καθιστά δυνατή την εκπαίδευση βαθύτερων U-Net χωρίς

φαινόμενα εξασθένισης (vanishing gradients). Η συγκεκριμένη παραλλαγή έχει βρει εφαρμογή σε πολλά προβλήματα ιατρικής απεικόνισης, όπου η λεπτομερής οριοθέτηση είναι κρίσιμη.

Attention U-Net

Η *Attention U-Net* [oktay2018attentionunet] εισάγει *attention gates* στα skip connections, ώστε το δίκτυο να μπορεί να «φιλτράρει» τα χαρακτηριστικά που μεταφέρονται από τον encoder στον decoder. Με τον τρόπο αυτό, αποτρέπεται η εισαγωγή μη χρήσιμων ή θορυβωδών πληροφοριών, ενώ ενισχύονται οι χωρικές περιοχές ενδιαφέροντος (π.χ. οι όγκοι). Στην πράξη, η προσθήκη μηχανισμών attention βελτιώνει την ακρίβεια σε δύσκολες περιπτώσεις με μικρούς ή ασαφείς όγκους.

nnU-Net

Το *nnU-Net* [isensee2021nnunet] (“no-new U-Net”) αποτελεί ένα πλήρως αυτόματο πλαίσιο, το οποίο προσαρμόζει τις υπερπαραμέτρους (βάθος, αριθμό φίλτρων, patch size, preprocessing) στο εκάστοτε dataset. Σε αντίθεση με άλλες παραλλαγές, δεν προτείνει νέα αρχιτεκτονική, αλλά ένα σύστημα *auto-configuration* που έχει επιτύχει κορυφαίες επιδόσεις σε πολλούς διεθνείς διαγωνισμούς, μεταξύ αυτών και στο BraTS. Το nnU-Net αναδεικνύει ότι η προσαρμογή των λεπτομερειών υλοποίησης μπορεί να είναι εξίσου σημαντική με την αρχιτεκτονική καινοτομία.

2D vs 3D U-Net

Μια ακόμη σημαντική διαφοροποίηση είναι η μετάβαση από 2D σε 3D U-Net. Η 2D εκδοχή είναι πιο αποδοτική σε μνήμη και υπολογισμούς, καθώς επεξεργάζεται κάθε τομή ξεχωριστά, αλλά μπορεί να χάνει πληροφορία κατά τον άξονα βάθους. Η 3D εκδοχή επεξεργάζεται ολόκληρα patches όγκου και συνεπώς αξιοποιεί πλήρως το χωρικό συμφραζόμενο, με κόστος αυξημένων απαιτήσεων σε GPU μνήμη και χρόνο εκπαίδευσης. Στη βιβλιογραφία απαντώνται επίσης υβριδικές προσεγγίσεις (2.5D U-Net), όπου γειτονικές τομές τροφοδοτούνται ταυτόχρονα ως πολυκάναλη είσοδος.

Σχόλιο. Οι παραλλαγές της U-Net δείχνουν τη μεγάλη προσαρμοστικότητα της αρχιτεκτονικής. Είτε μέσω residual blocks, είτε μέσω attention μηχανισμών, είτε μέσω αυτόματης προσαρμογής (nnU-Net), όλες στοχεύουν στο να ενισχύσουν τη σταθερότητα, την ακρίβεια και τη γενίκευση. Για τον λόγο αυτό, η U-Net και οι παραλλαγές της θεωρούνται μέχρι σήμερα η «χρυσή βάση» για την τμηματοποίηση ιατρικών εικόνων.

4.7 Εκπαιδευτικές τεχνικές και αρχιτεκτονικά βοηθήματα

Πέρα από τον βασικό σχεδιασμό της αρχιτεκτονικής, η επιτυχής εκπαίδευση ενός μοντέλου τμηματοποίησης MRI εξαρτάται σε μεγάλο βαθμό από ορισμένα «βοηθήματα» που βελτιώνουν τη γενίκευση, τη σταθερότητα και τη συγκλίνουσα λύση. Οι τεχνικές αυτές συνιστούν ένα ενδιάμεσο επίπεδο μεταξύ καθαρά αρχιτεκτονικών επιλογών και αλγοριθμικών παραμέτρων εκπαίδευσης, και έχουν επανειλημμένα αποδειχθεί κρίσιμες για την επιτυχία σε κλινικά δεδομένα. Τα πιο σημαντικά είναι τα εξής:

Dropout

Η τεχνική *dropout* εισήχθη από τους Srivastava et al. [srivastava2014dropout] ως μέθοδος τακτικοποίησης (*regularization*). Σε κάθε βήμα εκπαίδευσης, ένα ποσοστό των νευρώνων απενεργοποιείται τυχαία, αποτρέποντας την υπερβολική εξειδίκευση σε συγκεκριμένα μοτίβα των δεδομένων. Με αυτό τον τρόπο, το δίκτυο «μαθαίνει» πιο ανθεκτικές αναπαραστάσεις, καθώς δεν μπορεί να βασιστεί σε με-

μονωμένους νευρώνες αλλά πρέπει να καταναίμει την πληροφορία.

Στην παρούσα εργασία εφαρμόστηκε με πιθανότητα $p = 0.1$ σε διάφορα σημεία των μοντέλων (π.χ. U-Net, V-Net, SegResNet, SwinUNETR), συμβάλλοντας στη μείωση του overfitting. Η χαμηλή τιμή p επιλέχθηκε ώστε να ενισχύεται η γενίκευση χωρίς να παρακωλύεται η ικανότητα μάθησης. Αξίζει να σημειωθεί ότι σε βαθύτερα δίκτυα ή σε περιβάλλοντα με μικρότερο αριθμό δειγμάτων, η αύξηση του p μπορεί να λειτουργήσει προστατευτικά, ενώ σε μεγάλα datasets το dropout συχνά υποχωρεί σε σημασία μπροστά σε πιο πολύπλοκες τεχνικές τακτικοποίησης.

Κανονικοποίηση

Η κανονικοποίηση (Batch, Instance, Group Normalization) είναι απαραίτητη για να σταθεροποιήσει την κατανομή των ενεργοποιήσεων και να διευκολύνει την εκπαίδευση βαθιών δικτύων. Στα 3D δίκτυα MRI, η *InstanceNorm* ή *GroupNorm* χρησιμοποιούνται συχνότερα, καθώς οι περιορισμένες διαστάσεις batch καθιστούν την BatchNorm λιγότερο αποτελεσματική.

Η ομογενοποίηση των χαρακτηριστικών σε επίπεδο δείγματος/ομάδας μειώνει τη διακύμανση και επιτρέπει γρηγορότερη σύγκλιση. Επιπλέον, η κανονικοποίηση λειτουργεί έμμεσα και ως regularization, καθώς αποτρέπει την υπερβολική «μετατόπιση» των κατανομών σε κάθε στρώμα. Στη βιβλιογραφία αναφέρεται ότι η επιλογή του κατάλληλου τύπου normalization εξαρτάται από το μέγεθος batch, τον διαθέσιμο υπολογιστικό πόρο και τη φύση των δεδομένων, με την InstanceNorm να αποτελεί συχνά την προτιμώμενη επιλογή στη βιοϊατρική απεικόνιση.

Deep Supervision

Η *deep supervision* συνίσταται στην προσθήκη ενδιάμεσων εξόδων και συναρτήσεων κόστους σε βαθύτερα επίπεδα του δικτύου [33]. Με τον τρόπο αυτό, ενθαρρύνεται η διάδοση σημασιολογικής πληροφορίας και μειώνεται ο κίνδυνος εξαφάνισης κλίσεων (*vanishing gradients*).

Στα δίκτυα τμηματοποίησης μπορεί να εφαρμοστεί με την παραγωγή χαρτών πρόβλεψης σε διαφορετικές κλίμακες και τη συνδυαστική βελτιστοποίησή τους. Έτσι, ακόμα και τα πρώιμα στρώματα αναγκάζονται να μάθουν χρήσιμα χαρακτηριστικά που σχετίζονται με το τελικό καθήκον. Αν και δεν χρησιμοποιήθηκε στο βασικό πείραμα της παρούσας εργασίας, η βιβλιογραφία δείχνει ότι μπορεί να βελτιώσει τη σταθερότητα ειδικά σε πολύ βαθιά μοντέλα, και συχνά εφαρμόζεται σε παραλλαγές της U-Net ή της SegResNet.

Weight Initialization

Η σωστή αρχικοποίηση των βαρών έχει επίσης μεγάλη σημασία. Τεχνικές όπως He initialization ή Xavier initialization έχουν δείξει ότι αποτρέπουν την υπερβολική αύξηση ή εξαφάνιση των ενεργοποιήσεων, διευκολύνοντας τη γρήγορη σύγκλιση. Στα δίκτυα με ReLU ή PReLU ενεργοποιήσεις, η He initialization θεωρείται καταλληλότερη. Αν και πρόκειται για μια «σιωπηλή» επιλογή, επηρεάζει καθοριστικά τη συμπεριφορά της εκπαίδευσης.

Data Augmentation ως αρχιτεκτονικό βοήθημα

Αν και τυπικά ανήκει στη φάση προεπεξεργασίας, η στοχευμένη επαύξηση δεδομένων (γεωμετρικές και φωτεινές παραμορφώσεις, τυχαίες περιστροφές, αλλαγές αντίθεσης) μπορεί να θεωρηθεί αρχιτεκτονικό βοήθημα, καθώς αυξάνει την ποικιλία των παραδειγμάτων που «βλέπει» το μοντέλο. Σε ιατρικές εικόνες, όπου τα διαθέσιμα δείγματα είναι συχνά περιορισμένα, η χρήση επαυξήσεων είναι κρίσιμη για την αποφυγή overfitting και την καλύτερη γενίκευση σε νέα κέντρα ή σαρωτές.

Σχόλιο. Οι παραπάνω τεχνικές δεν αλλάζουν ριζικά την αρχιτεκτονική, ωστόσο αποτελούν κρίσιμα συστατικά για την πρακτική εκπαίδευση και τη γενίκευση σε πραγματικά δεδομένα MRI. Η ενσωμάτωσή τους συνδέει την αρχιτεκτονική σχεδίαση με τη στρατηγική εκπαίδευσης και συχνά καθορίζει τη διαφορά μεταξύ θεωρητικής και κλινικά χρήσιμης απόδοσης.

4.8 Σύγκριση πολυπλοκότητας και σύνδεση με πειράματα

Μεθοδολογία αποτίμησης

Η αποτίμηση της πολυπλοκότητας των μοντέλων είναι κρίσιμη για τη δίκαιη σύγκριση τους, καθώς οι διαφορετικές αρχιτεκτονικές μπορεί να έχουν πολύ διαφορετικές απαιτήσεις σε μνήμη και υπολογιστική ισχύ, πέρα από την ακρίβεια στην τμηματοποίηση. Για τον σκοπό αυτό υιοθετούμε ενδεικτικές ρυθμίσεις που συμβαδίζουν με τα δεδομένα MRI εγκεφάλου και τα όρια υλικού (GPU):

- **Είσοδος 2D (τομές):** 160×192 με 4 κανάλια (τα modalities του BraTS).
- **Είσοδος 3D (patches):** $96 \times 96 \times 96$ με 4 κανάλια.
- **Κλάσεις εξόδου:** 4 (background, ET, ED, NCR/NET).
- **Βασικό πλάτος καναλιών:** U-Net/SegResNet (2D) = 32, V-Net (3D) = 16.

Με αυτά τα σενάρια, υπολογίζονται θεωρητικά οι παράμετροι και τα FLOPs ανά εμπρός διέλευση. Οι αριθμοί δεν είναι απόλυτοι αλλά δίνουν αξιόπιστη εικόνα της *τάξης μεγέθους*.

Πίνακας 4.1: Σύγκριση μοντέλων ως προς παραμέτρους και θεωρητικά FLOPs ανά δείγμα.

Μοντέλο	Διαστασιμότητα	Patch/Τομή	Παράμετροι (M)	FLOPs (GF)
U-Net (base=32)	2D	160×192	≈ 7.76	≈ 12.85
SegResNet (base=32)	2D	160×192	≈ 17.98	≈ 23.61
V-Net (base=16)	3D	96^3	≈ 13.31	≈ 209
SwinUNETR	3D	96^3 (ενδεικτικό)	—	—

Σχόλια αποτελεσμάτων

1. **U-Net vs SegResNet (2D):** Η SegResNet εισάγει residual units, που αυξάνουν την εκφραστικότητα αλλά διπλασιάζουν περίπου τις παραμέτρους και τα FLOPs. Το trade-off είναι καλύτερη γενίκευση και σταθερότερη εκπαίδευση, με μεγαλύτερο όμως κόστος σε μνήμη/χρόνο.
2. **3D V-Net:** Ενδιαφέρον είναι ότι, αν και ο αριθμός παραμέτρων είναι παρόμοιος με την 2D SegResNet, το κόστος σε FLOPs είναι τάξης μεγέθους μεγαλύτερο. Αυτό αποτυπώνει την εκθετική αύξηση κόστους όταν περνάμε από $H \times W$ σε $D \times H \times W$ δεδομένα, κάτι που επηρεάζει έντονα τη δυνατότητα εκπαίδευσης σε περιορισμένους πόρους.
3. **SwinUNETR:** Η πολυπλοκότητά του εξαρτάται λιγότερο από τον αριθμό καναλιών και περισσότερο από τον *embed dimension*, το βάθος του Transformer και το μέγεθος παραθύρου M . Το κόστος αυτοπροσοχής ανά στρώση είναι περίπου $\mathcal{O}(N \cdot M^d)$ για N tokens, κάτι που το καθιστά διαχειρίσιμο αλλά με αυξημένες απαιτήσεις σε GPU μνήμη.

Σύνδεση με τα πειράματα του παρόντος έργου

Στην ενότητα Υλικών & Μεθόδων (Κεφ. ??) τεκμηριώνεται η επιλογή 2D τομών 160×192 για λόγους ευστάθειας μνήμης και δίκαιης σύγκρισης. Η U-Net λειτουργεί ως baseline, η SegResNet εξετάζει το όφελος των residual units στην ίδια διάσταση, ενώ η V-Net και η SwinUNETR αξιολογούνται ως 3D επεκτάσεις που εκμεταλλεύονται το πλήρες χωρικό συμφραζόμενο, πληρώνοντας όμως σαφώς μεγαλύτερο υπολογιστικό κόστος.

Οδηγίες αναπαραγωγής παραμέτρων. Ο αριθμός παραμέτρων ενός συνελκτικού στρώματος δίνεται από

$$\#par = C_{out} \cdot C_{in} \cdot \prod_i k_i + C_{out},$$

και το θεωρητικό κόστος (FLOPs) κλιμακώνεται ως

$$FLOPs \propto (\text{χωρικά δείγματα εξόδου}) \times C_{out} \cdot C_{in} \cdot \prod_i k_i.$$

Αυτό επιτρέπει *ex-ante* εκτιμήσεις πριν από μεγάλα πειράματα, ώστε να προβλεφθούν οι απαιτήσεις σε GPU και να επιλεγεί κατάλληλο batch size.

4.9 Σύνοψη Κεφαλαίου

Στο παρόν κεφάλαιο παρουσιάστηκαν τέσσερις χαρακτηριστικές αρχιτεκτονικές που χρησιμοποιούνται εκτενώς στην τμηματοποίηση MRI εγκεφάλου. Η U-Net αποτελεί την πιο καθιερωμένη και ευέλικτη γραμμή βάσης, χάρη στον απλό αλλά ισχυρό σχεδιασμό της (encoder-decoder με skip connections) που προσφέρει καλή ισορροπία ακρίβειας και αποδοτικότητας. Η V-Net εισήγαγε τη χρήση 3D συνελίξεων με residual μπλοκ, καθιστώντας δυνατή την αξιοποίηση πλήρους volumetric συμφραζομένου, με τίμημα όμως αυξημένο υπολογιστικό κόστος. Η SegResNet συνδύασε τη φιλοσοφία της U-Net με εκτεταμένα residual units, επιτρέποντας βαθύτερα δίκτυα και μεγαλύτερη εκφραστικότητα χωρίς απώλεια σταθερότητας εκπαίδευσης. Τέλος, η SwinUNETR ενσωμάτωσε την αρχιτεκτονική των Transformers μέσω του μηχανισμού *window self-attention*, επιτυγχάνοντας καλύτερη κατανόηση global συσχετίσεων ενώ ταυτόχρονα διατηρεί λεπτομερή τμηματοποίηση μέσω συνελκτικού decoder.

Η συγκριτική ανάλυση παραμέτρων και θεωρητικών FLOPs ανέδειξε την κεντρική αντιστάθμιση ανάμεσα σε ακρίβεια και υπολογιστικό κόστος: τα 2D μοντέλα (U-Net, SegResNet) είναι πιο ελαφριά και πρακτικά για εκπαίδευση σε περιορισμένους πόρους, ενώ τα 3D μοντέλα (V-Net, SwinUNETR) αξιοποιούν το πλήρες χωρικό συμφραζόμενο αλλά με πολύ μεγαλύτερες απαιτήσεις σε μνήμη και χρόνο. Τα residual connections και οι μηχανισμοί attention αποτελούν κρίσιμες σχεδιαστικές επιλογές που βελτιώνουν τη σταθερότητα και την εκφραστικότητα του δικτύου, δείχνοντας την πορεία εξέλιξης από τις κλασικές CNN αρχιτεκτονικές προς πιο σύνθετα υβριδικά μοντέλα.

Συνολικά, το κεφάλαιο αυτό έθεσε το τεχνικό υπόβαθρο για τη συγκριτική μελέτη που ακολουθεί. Η παρουσίαση των μοντέλων, σε συνδυασμό με την αποτίμηση πολυπλοκότητας, προσφέρει το πλαίσιο για την πειραματική αξιολόγηση που θα ακολουθήσει στα επόμενα κεφάλαια, όπου θα διερευνηθεί η απόδοσή τους σε δεδομένα MRI εγκεφάλου από το BraTS και θα αξιολογηθούν με αντικειμενικές μετρικές τμηματοποίησης.

Κεφάλαιο 5: Υλικά & Μέθοδοι

Στόχος κεφαλαίου. Στο κεφάλαιο αυτό ορίζουμε με ακρίβεια το πειραματικό πρωτόκολλο που ακολουθήθηκε από άκρο σε άκρο, ώστε να είναι πλήρως αναπαραγώγιμο: από την ανάγνωση των NIfTI αρχείων και την προεπεξεργασία, έως τον ορισμό των μοντέλων/συναρτήσεων κόστους/βελτιστοποιητών, τις μεθόδους εκπαίδευσης/αξιολόγησης και την εκτέλεση του πλέγματος πειραμάτων. Η δομή και το ύφος παρουσίασης ακολουθούν το πρότυπο των προηγούμενων κεφαλαίων της εργασίας και αντλούν έμπνευση από σχετική διπλωματική στο ίδιο γνωστικό αντικείμενο. :contentReference[oaicite:0]index=0

5.1 Σύνολα δεδομένων & διαχωρισμός

Πηγές δεδομένων & μορφότυπος

Για την εκπαίδευση και αξιολόγηση των μοντέλων χρησιμοποιήθηκαν τρισδιάστατοι όγκοι μαγνητικής τομογραφίας (MRI) εγκεφάλου, αποθηκευμένοι σε μορφότυπο NIfTI (.nii/.nii.gz). Ο μορφότυπος αυτός έχει καθιερωθεί στη νευροαπεικόνιση, καθώς υποστηρίζει τρισδιάστατα δεδομένα με πλήρη πληροφορία για διαστάσεις, voxel spacing και affine μετασχηματισμούς, εξασφαλίζοντας διαλειτουργικότητα με εργαλεία όπως SimpleITK, NiBabel και MONAI.

Η δομή των καταλόγων δεδομένων ήταν η εξής:

- `imagesTr/`, `labelsTr/`: όγκοι και αντίστοιχες μάσκες εκπαίδευσης,
- `imagesVal/`, `labelsVal/`: όγκοι και μάσκες επικύρωσης.

Κάθε δείγμα εισόδου (`image`) αντιστοιχούσε μονοσήμαντα στη μάσκα του (`label`), με βάση την ταξινόμηση κατά όνομα αρχείου. Έτσι, αποφεύγεται το σφάλμα αντιστοίχισης και διασφαλίζεται πλήρης αναπαραγωγιμότητα.

Αν και τα αρχικά `labels` ήταν πολυκλαδικά (π.χ. για τα δεδομένα BraTS: enhancing tumor, edema, necrotic core), για τις ανάγκες των πειραμάτων τα μετατρέψαμε σε δυαδικές μάσκες *όγκος vs υπόβαθρο*. Αυτό έγινε με εφαρμογή κατωφλίου:

$$\mathbf{y}(v) = \mathbf{1}_{(\text{label}(v) > 0)},$$

ώστε να υπάρχει απλοποίηση του προβλήματος και ομοιομορφία μεταξύ όλων των μοντέλων που συγκρίνονται (βλ. §5.5). Με αυτόν τον τρόπο επικεντρωνόμαστε στην ικανότητα εντοπισμού του όγκου ως σύνολο, ανεξαρτήτως υποκατηγορίας.

Διαχωρισμός

Ο διαχωρισμός σε εκπαίδευση και επικύρωση πραγματοποιήθηκε εκ των προτέρων στον φάκελο δεδομένων, ώστε να αποφευχθεί οποιοδήποτε ενδεχόμενο *data leakage*. Το σύνολο επικύρωσης διατηρήθηκε σταθερό σε όλα τα πειράματα, εξασφαλίζοντας δίκαιη και συγκρίσιμη αξιολόγηση των μοντέλων. Δεν εφαρμόστηκε περαιτέρω ανακάτεμα (*shuffle*) στο επίπεδο όγκου, προκειμένου να παραμένει σταθερός ο διαχωρισμός και να είναι δυνατή η άμεση αναπαραγωγή των αποτελεσμάτων.

Αντίθετα, στο επίπεδο patches εφαρμόστηκε στοχαστική δειγματοληψία (βλ. §5.2.1), ώστε να βελτιωθεί η ποικιλία των δειγμάτων που τροφοδοτούνται στο δίκτυο κατά την εκπαίδευση. Η στρατηγική

αυτή εξασφαλίζει ισορροπία ανάμεσα στην αναπαραγωγιμότητα του πρωτοκόλλου και στη στατιστική επάρκεια των δεδομένων που «βλέπει» το δίκτυο σε κάθε εποχή.

5.2 Προεπεξεργασία & φόρτωση δεδομένων

Η σωστή κανονικοποίηση γεωμετρίας και έντασης αποτελεί κρίσιμο βήμα ώστε να καταστούν τα δεδομένα συγκρίσιμα μεταξύ διαφορετικών μοντέλων και συνεπώς τα αποτελέσματα αξιόπιστα. Στην παρούσα εργασία υλοποιήθηκε αγωγός προεπεξεργασίας με τη βιβλιοθήκη **MONAI**, η οποία προσφέρει πλούσια γκάμα εξειδικευμένων μετασχηματισμών για ιατρικές εικόνες.

Ανάγνωση & τυποποίηση καναλιών

Η ανάγνωση των δεδομένων έγινε απευθείας από αρχεία NIfTI με χρήση `LoadImaged`, διατηρώντας τη συμβατότητα με τα `metadata` (`affine matrix`, `voxel spacing`). Στη συνέχεια εφαρμόστηκαν:

1. **EnsureChannelFirst**: Ρητή αναδιάταξη ώστε η μορφή των `tensors` να είναι $(C \times D \times H \times W)$, διευκολύνοντας την ομοιογενή επεξεργασία από `PyTorch` και `MONAI`.
2. **Επιλογή καναλιού**: Επειδή τα δεδομένα διαθέτουν πολλαπλά `modalities`, επιλέχθηκε για λόγους απλούστευσης το πρώτο διαθέσιμο κανάλι (`custom μετασχηματισμός SelectFirstChannel`). Έτσι όλα τα μοντέλα εκπαιδεύτηκαν με είσοδο $C = 1$, επιτρέποντας δίκαιη σύγκριση και αποφυγή `bias` υπέρ πολυτροπικών εισόδων.

Εναρμόνιση γεωμετρίας

Τα δεδομένα MRI προέρχονται συχνά από διαφορετικούς σαρωτές ή πρωτόκολλα, με διαφορετικό προσανατολισμό ή `voxel spacing`. Για την εναρμόνιση:

1. **Orientation (RAS)**: Μετατροπή όλων των όγκων σε κοινό προσανατολισμό RAS (`Right–Anterior–Superior`).
2. **Resampling σε ιστροπικό voxel**: Ομογενοποίηση `voxel spacing` σε $(1.0, 1.0, 1.0)$ mm, με *bilinear interpolation* για εικόνα και *nearest neighbor* για τη μάσκα. Έτσι, κάθε όγκος αντιπροσωπεύει φυσικά τις ίδιες γεωμετρικές αναλογίες.

Κανονικοποίηση έντασης

Οι τιμές έντασης ποικίλουν σημαντικά ανά σαρωτή. Για τον λόγο αυτό εφαρμόστηκε `ScaleIntensityRange`, με γραμμική χαρτογράφηση $[-100, 400] \mapsto [0, 1]$ και `clipping`. Το παράθυρο αυτό επιλέχθηκε ώστε να σταθεροποιεί την αντίθεση μεταξύ ιστών, να ελαχιστοποιεί `outliers` και να διευκολύνει τη σύγκριση μεταξύ δειγμάτων.

5.2.1 Δειγματοληψία patches & επαυξήσεις

Patch-based δειγματοληψία. Λόγω περιορισμών μνήμης, η εκπαίδευση έγινε σε κυβικά patches 96^3 voxel:

```
RandCropByPosNegLabel(spatial_size=963, pos:neg = 1:1, num_samples = 2).
```

Από κάθε όγκο εξάγονταν 2 patches, ισορροπημένα ως προς `foreground/background`. Αυτό αντιστοιχεί σε *effective batch size* = 2 με `batch_size=1`, διασφαλίζοντας τόσο αποδοτικότητα όσο και ισορροπία στα παραδείγματα εκπαίδευσης.

Εναλλακτικές στρατηγικές sampling. Η επιλογή patches αποτελεί κρίσιμο σημείο του pipeline. Εναλλακτικές στρατηγικές περιλαμβάνουν:

- **Τυχαία αποκοπή (random patching):** απλή και γρήγορη, αλλά με κίνδυνο παραγωγής μη-αντιπροσωπευτικών patches.
- **Curriculum sampling:** σταδιακή επιλογή πιο δύσκολων patches με βάση την πρόοδο εκπαίδευσης, ώστε να βελτιώνεται η μάθηση σε challenging περιοχές.
- **Full-volume training:** εκπαίδευση σε ολόκληρους όγκους, που θεωρείται ιδανική θεωρητικά αλλά είναι πρακτικά απαγορευτική λόγω απαιτήσεων μνήμης.

Στην παρούσα εργασία, λόγω περιορισμών μνήμης GPU, το **sliding-window** επιλέχθηκε τόσο στην εκπαίδευση όσο και στην αξιολόγηση, ώστε όλα τα μοντέλα (ιδίως τα βαρύτερα όπως SegResNet και SwinUNETR) να συγκρίνονται δίκαια στο ίδιο πλαίσιο.

Γεωμετρικές επανυζητήσεις. Για να βελτιωθεί η γενίκευση, εφαρμόστηκαν τυχαίες γεωμετρικές παραμορφώσεις: RandFlip σε όλους τους άξονες, RandAffine (περιστροφή/κλίμακα/μετατόπιση με $p = 0.7$), καθώς και RandRotate ($p = 0.6$). Οι εικόνες παραμορφώνονταν με *bilinear*, ενώ οι μάσκες με *nearest neighbor* ώστε να διατηρούνται οι κατηγορίες ακεραίες.

Φωτομετρικές επανυζητήσεις. Πραγματοποιήθηκαν επίσης τυχαίες αλλαγές έντασης: RandScaleIntensity (κλίμακα ± 0.2 , $p = 0.7$) και RandShiftIntensity (μετατόπιση ± 0.15 , $p = 0.7$). Με αυτόν τον τρόπο προσομοιώνονται διαφοροποιήσεις σαρωτών και πρωτοκόλλων, μειώνοντας το φαινόμενο *domain shift*.

Σετ επικύρωσης. Για το validation δεν εφαρμόστηκαν τυχαίες επεμβάσεις. Η μόνη λειτουργία ήταν η SpatialPad σε 96^3 , ώστε να εξασφαλιστεί ότι τα patches καλύπτουν πλήρως τον όγκο και να επιτραπεί ολίσθηση παραθύρου στο inference (βλ. §5.5.3).

Cache & DataLoader

Για ταχύτερη εκπαίδευση χρησιμοποιήθηκε CacheDataset με `cache_rate=1.0`, ώστε τα προεπεξεργασμένα δεδομένα να επαναχρησιμοποιούνται χωρίς νέο υπολογισμό σε κάθε εποχή. Η φόρτωση έγινε με DataLoader (`batch_size=1`), αποδίδοντας τελικά tensors σχήματος $(2, 1, 96, 96, 96)$ ($2 \text{ patches} \times 1 \text{ κανάλι}$). Η επιτυχής κανονικοποίηση επαληθεύτηκε με έλεγχο εύρους εντάσεων $\approx [0, 1]$.

5.2.2 Σημασία βημάτων προεπεξεργασίας

Κάθε στάδιο προεπεξεργασίας έχει σαφή ρόλο στη βελτίωση της αξιοπιστίας:

- **Skull-stripping:** αφαιρεί τον κρανιακό ιστό, μειώνοντας θόρυβο και ετερογένεια εκτός στόχου. Αποτελεί στάνταρ πρακτική στη βιβλιογραφία για MRI όγκων [fischl2012fre surfer].
- **Bias field correction:** διορθώνει αργές διακυμάνσεις έντασης λόγω ανομοιογενειών RF, επιτρέποντας ομοιογενέστερη αντίθεση ιστών [tustison2010n4itk].
- **Resampling σε κοινό voxel size:** διασφαλίζει ότι όλοι οι όγκοι έχουν τις ίδιες γεωμετρικές αναλογίες, κάτι κρίσιμο για patch-based εκπαίδευση.

Αν και ορισμένα βήματα (π.χ. skull-stripping) δεν εφαρμόστηκαν αυτοτελώς στον τρέχον pipeline, η αναφορά τους κρίνεται σημαντική καθώς αποτελούν κοινά εργαλεία στη διεθνή πρακτική και εξηγούν πώς μπορούν να επηρεάσουν την ακρίβεια τμηματοποίησης.

5.3 Αρχιτεκτονικές μοντέλων

Για τη σύγκριση που ακολουθεί υλοποιήθηκαν τέσσερις 3D αρχιτεκτονικές, οι οποίες αντιπροσωπεύουν διαφορετικά στάδια εξέλιξης των δικτύων τμηματοποίησης. Όλες οι υλοποιήσεις έγιναν με το framework **MONAI** (PyTorch backend), εξασφαλίζοντας τυποποίηση στον ορισμό των layers και των μετασχηματισμών. Για λόγους συγκρισιμότητας, όλα τα μοντέλα ορίστηκαν με $C_{in} = 1$ (μονοκάναλη είσοδο) και $C_{out} = 2$ (δυαδική τμηματοποίηση: όγκος vs υπόβαθρο). Οι υπερπαραμέτροι ρυθμίστηκαν έτσι ώστε να διατηρείται ισορροπία ανάμεσα στην υπολογιστική πολυπλοκότητα και τη δυνατότητα εκπαίδευσης εντός των διαθέσιμων GPU πόρων.

Πίνακας 5.1: Σύνοψη υπερπαραμέτρων ανά αρχιτεκτονική.

Μοντέλο	Ρυθμίσεις
U-Net	channels=(32,64,128,256,512), strides=(2,2,2,2), num_res_units=2; norm=InstanceNorm, dropout=0.1, act=PReLU, bias=True. Χρησιμοποιήθηκαν residual units για σταθερότερη εκπαίδευση και InstanceNorm λόγω μικρών batch.
V-Net	dropout_prob=0.1, dropout_down=0.1, dropout_up=(0.1,0.1), act=PReLU. Ενσωματώνει πλήρως 3D συνελίξεις με residual blocks: οι παράμετροι dropout συμβάλλουν στην αποφυγή υπερεκπαίδευσης.
SegResNet	init_filters=8, dropout_prob=0.1. Ο σχετικά μικρός αριθμός αρχικών φίλτρων καθιστά το μοντέλο πιο ελαφρύ, ενώ οι residual μονάδες βελτιώνουν την εκφραστικότητα.
SwinUNETR	img_size=96 ³ , feature_size=24, use_checkpoint=True, drop_rate=attn_drop_rate=0.1. Ο encoder βασίζεται στο Swin Transformer: χρησιμοποιήθηκε checkpointing για μείωση μνήμης κατά την εκπαίδευση.

Αξίζει να σημειωθεί ότι οι παραπάνω ρυθμίσεις δεν αποτελούν τις «μέγιστες» δυνατές εκδοχές των μοντέλων, αλλά επιλέχθηκαν ώστε να χωράνε πρακτικά στη διαθέσιμη μνήμη GPU και να επιτρέπουν δίκαιη συγκριτική αξιολόγηση. Έτσι, η σύγκριση επικεντρώνεται στις αρχιτεκτονικές επιλογές (skip connections, residual blocks, attention) και όχι σε απόλυτη αύξηση του μεγέθους του δικτύου.

5.4 Συναρτήσεις κόστους, βελτιστοποίηση & κανονικοποίηση

Συναρτήσεις κόστους

Η επιλογή της συνάρτησης κόστους είναι κρίσιμη για προβλήματα τμηματοποίησης, ιδιαίτερα όταν υπάρχει μεγάλη ανισορροπία ανάμεσα στις κλάσεις (*class imbalance*). Στην παρούσα εργασία χρησιμοποιήθηκε ως βασική απώλεια η **DiceLoss**, η οποία υπολογίζει άμεσα τον συντελεστή Dice μεταξύ πρόβλεψης και εδάφους αλήθειας:

$$\mathcal{L}_{\text{Dice}} = 1 - \frac{2 \sum_i p_i y_i + \epsilon}{\sum_i p_i^2 + \sum_i y_i^2 + \epsilon},$$

όπου p_i οι προβλεπόμενες πιθανότητες και y_i οι ετικέτες. Χρησιμοποιήθηκε η παραμετροποίηση `to_onehot_y = True`, `softmax = True`, `squared_pred = True`, ώστε να εξασφαλιστεί σταθερότητα και να δοθεί έμφαση στη σωστή ανακατασκευή λεπτομερειών.

Η DiceLoss είναι ιδιαίτερα κατάλληλη για MRI εγκεφάλου, όπου οι όγκοι καταλαμβάνουν μικρό ποσοστό των voxel σε σχέση με το υπόβαθρο. Εναλλακτικά, δοκιμάστηκε και η **DiceCE** (συνδυασμός Dice + Cross-Entropy), η οποία προσπαθεί να ισορροπήσει μεταξύ επικέντρωσης στην επικράτηση του foreground (Dice) και στην voxel-wise ακρίβεια (CE). Παρόλα αυτά, τα πειράματα έδειξαν ότι η καθαρή DiceLoss απέδωσε σταθερά καλύτερα αποτελέσματα, ιδιαίτερα στο UNet, και γι' αυτό επιλέχθηκε ως αποκλειστική συνάρτηση κόστους για τα τελικά πειράματα.

Βελτιστοποιητές & ρυθμός μάθησης

Για την εκπαίδευση χρησιμοποιήθηκε ο **AdamW** με $\text{lr} = 10^{-4}$, $\text{weight_decay} = 10^{-5}$, $\text{amsgrad} = \text{True}$. Ο AdamW επιλέχθηκε διότι συνδυάζει την αποδοτικότητα του Adam με σταθερότερη κανονικοποίηση βαρών (*decoupled weight decay*), μειώνοντας το *overfitting*. Παράλληλα εφαρμόστηκε **αποκοπή κλίσεων** ($\|\nabla\|_2 \leq 1.0$) ώστε να αποφευχθούν εκρήξεις κλίσεων, που είναι πιθανές σε βαθιά δίκτυα με 3D patches.

Στα αρχικά πειράματα συγκρίθηκαν διαφορετικοί συνδυασμοί (Dice, DiceCE με Adam ή AdamW). Η συγκριτική ανάλυση κατέδειξε ότι ο συνδυασμός *Dice + AdamW* προσέφερε την καλύτερη ισορροπία ακρίβειας και σταθερότητας, και έτσι αποτέλεσε την κύρια επιλογή της παρούσας εργασίας.

Μεικτή ακρίβεια & σταθεροποίηση

Για να μειωθούν οι απαιτήσεις μνήμης και να επιταχυνθεί η εκπαίδευση, ενεργοποιήθηκε **μεικτή ακρίβεια** με χρήση του `torch.cuda.amp.autocast` σε τύπο `bfloat16`. Η χρήση μεικτής ακρίβειας προσφέρει σημαντική εξοικονόμηση VRAM, χωρίς ουσιαστική απώλεια αριθμητικής ακρίβειας. Για να διασφαλιστεί αριθμητική σταθερότητα, χρησιμοποιήθηκε **GradScaler**, ο οποίος προσαρμόζει δυναμικά την κλίμακα των κλίσεων ώστε να αποφευχθεί υπορροή (*underflow*).

Η επιλογή αυτών των τεχνικών (DiceLoss, AdamW, μεικτή ακρίβεια) ήταν αποτέλεσμα πειραματικής αποτίμησης και στόχευε στην καλύτερη δυνατή απόδοση με δεδομένους υπολογιστικούς πόρους.

5.5 Εκπαίδευση & αξιολόγηση

Δυαδικοποίηση ετικετών

Καθ' όλη τη διαδικασία εκπαίδευσης και επικύρωσης, οι ετικέτες των όγκων μετατράπηκαν επί τόπου σε δυαδικές, με ορισμό:

$$\tilde{y} = \mathbf{1}_{(y>0)}.$$

Η επιλογή αυτή απλοποιεί το πρόβλημα σε διάκριση *όγκος vs. υπόβαθρο*, ώστε να υπάρχει ομοιομορφία συγκρίσεων μεταξύ των μοντέλων. Για συμβατότητα με τις συναρτήσεις απώλειας και τις μετρικές, εφαρμόστηκε one-hot επέκταση σε δύο κλάσεις και χρήση *softmax* στην πρόβλεψη. Με αυτόν τον τρόπο η αξιολόγηση είναι σύμφωνη με τον τρόπο υπολογισμού της DiceLoss.

Μετρική απόδοσης

Η κύρια μετρική ήταν ο **Dice score**, ο οποίος ορίζεται ως:

$$\text{Dice} = 1 - \mathcal{L}_{\text{Dice}}(\hat{p}, \tilde{y}).$$

Επιλέχθηκε διότι είναι ιδιαίτερα ευαίσθητος στην επικάλυψη μεταξύ πρόβλεψης και εδάφους αλήθειας, σε αντίθεση με απλές voxel-wise μετρικές όπως η ακρίβεια (accuracy) που υποτιμούν τη σημασία των μικρών όγκων. Ο Dice score αποτελεί καθιερωμένη μετρική στα benchmarks BraTS και γενικότερα σε ιατρική τμηματοποίηση.

5.5.1 Εναλλακτικές μετρικές αξιολόγησης

Παρότι ο Dice score ήταν η κύρια μετρική, στη βιβλιογραφία χρησιμοποιούνται συχνά και άλλες μετρικές που αποτιμούν διαφορετικές πτυχές της τμηματοποίησης:

- **Hausdorff Distance (HD95):** μετρά την απόσταση μεταξύ των ορίων πρόβλεψης και ground truth. Ορίζεται ως:

$$H(A, B) = \max\left\{\sup_{a \in A} \inf_{b \in B} \|a - b\|, \sup_{b \in B} \inf_{a \in A} \|b - a\|\right\},$$

με το HD95 να υπολογίζει το 95ο εκατοστημόριο αντί του μέγιστου, ώστε να μειώνεται η επίδραση outliers. Αξιολογεί τη γεωμετρική ακρίβεια των ορίων.

- **Sensitivity/Specificity:** ποσοτικοποιούν την ικανότητα ανίχνευσης θετικών (όγκος) και την αποφυγή ψευδώς θετικών, αντίστοιχα.
- **Volumetric Overlap Error (VOE):** εκφράζει την απόκλιση όγκου πρόβλεψης–αλήθειας ως ποσοστό, εστιάζοντας σε περιπτώσεις όπου ο συνολικός όγκος είναι σημαντικότερος από τα ακριβή όρια.

Οι μετρικές αυτές δεν υλοποιήθηκαν στο παρόν pipeline, ωστόσο αναφέρονται για πληρότητα και ως προοπτική μελλοντικής εργασίας, καθώς συμπληρώνουν τις πληροφορίες που παρέχει ο Dice score.

5.5.2 Σχήμα εκπαίδευσης

Για όλα τα μοντέλα εφαρμόστηκε το ίδιο σχήμα εκπαίδευσης ώστε να υπάρχει δίκαιη σύγκριση:

- **Εποχές:** $E = 200$, τιμή που επαρκεί για σύγκλιση χωρίς υπερβολικό κόστος χρόνου.

- **Early stopping:** με $\text{patience}=20$, ώστε να αποφεύγεται η υπερεκπαίδευση σε περιπτώσεις στασιμότητας.
- **Τεκμηρίωση προόδου:** καταγραφή απωλειών και Dice score σε .npy αρχεία (`loss_train.npy`, `metric_train.npy`, `loss_test.npy`, `metric_test.npy`) για αναλυτική παρακολούθηση.
- **Καλύτερο μοντέλο:** αποθήκευση `best_metric_model.pth` στο μέγιστο Dice επικύρωσης, ώστε τα αποτελέσματα να αναφέρονται πάντα στο βέλτιστο σημείο εκπαίδευσης.

5.5.3 Συμπερασματοδότηση με ολισθαίνον παράθυρο

Κατά την επικύρωση/δοκιμή εφαρμόστηκε **sliding-window inference**, με παραμέτρους:

$$\text{ROI size} = 96^3, \quad \text{overlap} = 0.25, \quad \text{sw_batch_size} = 2.$$

Η μέθοδος αυτή είναι απαραίτητη λόγω περιορισμών GPU μνήμης: ολόκληρος ο όγκος δεν μπορεί να υποβληθεί μονομιάς στο δίκτυο, συνεπώς χωρίζεται σε επικαλυπτόμενα patches. Στη συνέχεια τα αποτελέσματα συγκολλούνται με *Gaussian blending*, το οποίο μειώνει τις ασυνέχειες στα όρια των παραθύρων και παράγει πιο ομαλό τελικό segmentation.

Αιτιολόγηση επιλογής. Η χρήση sliding-window ήταν αναγκαία για δίκαιη σύγκριση μεταξύ όλων των μοντέλων. Οι πιο βαριές αρχιτεκτονικές, όπως οι SegResNet και SwinUNETR, δεν μπορούν να επεξεργαστούν ολόκληρους όγκους λόγω περιορισμών μνήμης GPU. Έτσι, επιλέχθηκε να εφαρμοστεί η ίδια μέθοδος inference και για ελαφρύτερα μοντέλα (UNet, VNet), ώστε να υπάρχει κοινό πλαίσιο αξιολόγησης και να μην ευνοείται τεχνητά κάποια κατηγορία μοντέλων.

5.6 Πλέγμα πειραμάτων & ορισμός συγκρίσεων

Η διαδικασία σύγκρισης των μοντέλων δεν μπορεί να βασιστεί σε μεμονωμένα ή αποσπασματικά πειράματα, καθώς η απόδοση μιας αρχιτεκτονικής επηρεάζεται έντονα από τον συνδυασμό υπερπαραμέτρων, συνάρτησης κόστους και βελτιστοποιητή. Για τον λόγο αυτό ορίστηκε ένα *πειραματικό πλέγμα* (*grid search*), όπου κάθε διαφορετικός συνδυασμός (μοντέλο, loss, optimizer) αντιστοιχεί σε μια ανεξάρτητη εκτέλεση. Με αυτόν τον τρόπο διασφαλίζεται ότι η αξιολόγηση θα είναι συστηματική και τα αποτελέσματα συγκρίσιμα.

Το πλέγμα ορίστηκε ως εξής:

$$\mathcal{M} = \{\text{UNet}, \text{VNet}, \text{SegResNet}, \text{SwinUNETR}\}, \quad \mathcal{L} = \{\text{Dice}\}, \quad \mathcal{O} = \{\text{AdamW}\}.$$

Το συνολικό καρτεσιανό γινόμενο

$$\mathcal{M} \times \mathcal{L} \times \mathcal{O}$$

παράγει τέσσερις διαφορετικές εκτελέσεις, δηλαδή μία για κάθε αρχιτεκτονική με σταθερή συνάρτηση κόστους (Dice) και βελτιστοποιητή (AdamW). Η επιλογή να περιοριστούμε σε αυτό το υποσύνολο έγινε ύστερα από διερευνητικά πειράματα (βλ. §??), όπου παρατηρήθηκε ότι ο συνδυασμός *Dice + AdamW* παρείχε σταθερά καλύτερη απόδοση από τις εναλλακτικές. Έτσι, αποφεύχθηκε η ανάγκη για μεγαλύτερο grid, μειώνοντας δραστικά τον υπολογιστικό χρόνο χωρίς να θυσιάζεται η αξιοπιστία των συγκρίσεων.

Για κάθε εκτέλεση `run_experiment(m, l, o)` εφαρμόστηκαν ενιαίες ρυθμίσεις εκπαίδευσης:

- **Μέγιστες εποχές:** $E = 200$, με *early stopping* (*patience* = 20) ώστε να αποφεύγεται η υπερεκπαίδευση.
- **Ρυθμός μάθησης:** $1r = 10^{-4}$, κοινός για όλα τα μοντέλα.
- **Αγωγός δεδομένων:** κοινή προεπεξεργασία και *data augmentation* (§5.2), εξασφαλίζοντας ότι οι διαφορές στην απόδοση προέρχονται αποκλειστικά από την αρχιτεκτονική.

Σε κάθε εκτέλεση καταγράφηκαν:

1. ο μέγιστος Dice score στο validation set, όπως προέκυψε κατά τη διάρκεια της εκπαίδευσης,
2. η εποχή στην οποία επιτεύχθηκε η μέγιστη τιμή (συνήθως $\leq E$ λόγω *early stopping*),
3. τα πλήρη αρχεία καταγραφής (train/val losses, Dice ανά epoch).

Όλα τα αποτελέσματα συγκεντρώθηκαν σε ενιαίο αρχείο `summary.csv`, το οποίο αποτέλεσε τη βάση για τους συγκριτικούς πίνακες και τις καμπύλες που παρουσιάζονται στο Κεφ. 6. Η επιλογή αυτής της μορφής καταγραφής έχει δύο πλεονεκτήματα: (α) απλοποιεί την αναπαραγωγή των αποτελεσμάτων, αφού όλα τα πειράματα συνοψίζονται σε μια τυποποιημένη δομή δεδομένων, και (β) διευκολύνει τη μεταγενέστερη ανάλυση ή οπτικοποίηση (π.χ. *boxplots*, *learning curves*).

Με τον τρόπο αυτό, το πειραματικό πλέγμα προσέφερε ένα συνεκτικό και δίκαιο πλαίσιο σύγκρισης, απομονώνοντας τους παράγοντες που θέλαμε να μελετήσουμε (αρχιτεκτονική) και κρατώντας όλους τους υπόλοιπους σταθερούς.

5.7 Αναπαραγωγιμότητα

Η αναπαραγωγιμότητα αποτελεί βασική αρχή σε κάθε πειραματική μελέτη μηχανικής μάθησης. Για να διασφαλιστεί ότι τα αποτελέσματα μπορούν να επαναληφθούν από τρίτους (ή από τον ίδιο ερευνητή σε μελλοντική φάση), ακολουθήθηκαν συγκεκριμένα μέτρα:

- **Τυχαιότητα:** Ενεργοποιήθηκε η λειτουργία `set_determinism(seed=0)` της MONAI, η οποία θέτει σπόρο για όλους τους κύριους γεννήτορες τυχαιότητας (NumPy, PyTorch CPU/GPU, Python RNG). Έτσι, οι μετασχηματισμοί, η δειγματοληψία *patches* και οι αρχικοποιήσεις βαρών είναι σταθερές στο μέτρο του δυνατού, μειώνοντας τη διασπορά των αποτελεσμάτων.
- **Υπολογιστικό περιβάλλον:** Τα πειράματα εκτελέστηκαν σε περιβάλλον με διαθέσιμη CUDA, χρησιμοποιώντας GPU *NVIDIA GeForce RTX 3060 Laptop*. Πριν από κάθε εκτέλεση πραγματοποιούνταν έλεγχος διαθεσιμότητας GPU και επιλογή της κατάλληλης συσκευής. Το συγκεκριμένο *setup* αντανakλά ένα «μεσαίο» υπολογιστικό περιβάλλον, επαρκές για την εκπαίδευση των μοντέλων με *patches* 96³.
- **Βιβλιοθήκες:** Χρησιμοποιήθηκαν PyTorch, MONAI, SimpleITK, NumPy και Pandas, όλες σε σταθερές εκδόσεις. Για πλήρη αναπαραγωγιμότητα, οι ακριβείς εκδόσεις καταγράφηκαν με `pip freeze` και παρατίθενται στο συμπληρωματικό υλικό της εργασίας. Αυτό επιτρέπει την ανασύσταση του ίδιου περιβάλλοντος σε οποιοδήποτε σύστημα.

Επιπλέον, ως καλή πρακτική προτείνεται η χρήση `requirements.txt` ή αρχείου `environment.yml` (conda), καθώς και η ενδεχόμενη συσκευασία του πειραματικού περιβάλλοντος σε Docker container. Έτσι, όχι μόνο διατηρείται η συμβατότητα βιβλιοθηκών, αλλά εξασφαλίζεται και η ανεξαρτησία από το υποκείμενο λειτουργικό σύστημα.

5.8 Συνοπτικός ψευδοκώδικας πειράματος

Για πληρότητα, παρατίθεται συνοπτικός ψευδοκώδικας που αποτυπώνει τον αγωγό από άκρο σε άκρο, όπως υλοποιήθηκε. Η αναπαράσταση αυτή βοηθά στη γρήγορη κατανόηση της ροής, χωρίς τις λεπτομέρειες του κώδικα PyTorch/MONAI.

Algorithm 2 Συνοπτικός ψευδοκώδικας πειραματικού pipeline

1: Ανάγνωση/προεπεξεργασία:

1. **(Προαιρετικά)** Μετατροπή ετικετών σε δυαδικές και αποθήκευση (SimpleITK threshold > 0.5).
2. Δημιουργία λίστας δειγμάτων $\{(image, label)\}$ με αντιστοίχιση ονομάτων.
3. **Train transforms:** Load $\rightarrow$ ChannelFirst $\rightarrow$ SelectFirstChannel $\rightarrow$ Orientation(RAS) $\rightarrow$ Spacing(1mm) $\rightarrow$ ScaleIntensityRange $\rightarrow$ RandCropByPosNegLabel(96^3 , 1:1, 2 samples) $\rightarrow$ RandFlip/Affine/Rotate $\rightarrow$ RandScale/ShiftIntensity $\rightarrow$ Cast/EnsureTyped.
4. **Val transforms:** ίδια γεωμετρία/ένταση, χωρίς τυχαιότητα, με SpatialPad(96^3).
5. Δημιουργία CacheDataset(cache_rate = 1.0) και DataLoader(batch_size = 1).

2: Εκπαίδευση/Αξιολόγηση (για κάθε συνδυασμό m, ℓ, o):

1. Κατασκευή $model(m), loss(\ell), optimizer(o)$.
 2. Ορισμός φακέλου results/{m_ell_o}.
 3. **For** $e = 1..E$:
 - (α') **Train:** εμπρόσθια διάδοση, υπολογισμός απώλειας, οπισθοδιάδοση, βήμα optimizer.
 - (β') **Val:** sliding-window inference, υπολογισμός Dice.
 - (γ') Αποθήκευση καμπυλών, ενημέρωση best Dice, έλεγχος early stopping.
 4. Επιστροφή $\max_e \text{Dice}_{\text{val}}^{(e)}$ και καταγραφή σε summary.csv.
-

5.9 Αιτιολόγηση σχεδιαστικών επιλογών

Οι σχεδιαστικές αποφάσεις που υιοθετήθηκαν στην παρούσα εργασία δεν είναι αυθαίρετες αλλά εδράζονται σε συγκεκριμένες απαιτήσεις της MRI εγκεφάλου, στους περιορισμούς των διαθέσιμων πόρων και σε βέλτιστες πρακτικές της βιβλιογραφίας. Παρακάτω συνοψίζονται οι σημαντικότερες επιλογές και η αιτιολόγησή τους:

- **Patch-based εκπαίδευση 96^3 :** Η χρήση κυβικών patches 96^3 διατηρεί το χωρικό συμφραζόμενο σε τρεις διαστάσεις, ενώ ταυτόχρονα επιτρέπει την εκπαίδευση σε GPU με περιορισμένη μνήμη. Η παράμετρος pos:neg=1:1 εξασφαλίζει ότι τα patches περιέχουν ισορροπημένα παραδείγματα foreground/background, μειώνοντας τον κίνδυνο εκπαίδευσης με κυρίως «άδεια» patches.
- **Εναρμόνιση RAS / voxel spacing 1 mm:** Η κανονικοποίηση προσανατολισμού (RAS) και η επα-

ναδειγματοληψία σε ιστροπικό spacing (1.0, 1.0, 1.0) mm μειώνουν το *domain shift* μεταξύ διαφορετικών υποκειμένων ή σαρωτών. Με αυτόν τον τρόπο, τα μοντέλα δεν χρειάζεται να μάθουν «τεχνητές» διαφορές γεωμετρίας και εστιάζουν στην πραγματική βιολογική ποικιλία.

- **InstanceNorm / Dropout:** Στα 3D CNNs το μέγεθος batch είναι συνήθως μικρό λόγω μνήμης. Η Instance Normalization είναι πιο κατάλληλη από BatchNorm σε αυτό το σενάριο, ενώ το Dropout (με $p = 0.1$) βοηθά στη μείωση του overfitting, ενισχύοντας τη γενίκευση.
- **Μεικτή ακρίβεια (bfloat16) & checkpointing (SwinUNETR):** Η εκπαίδευση σε μεικτή ακρίβεια μειώνει σημαντικά τη χρήση μνήμης και επιταχύνει τους υπολογισμούς χωρίς απώλεια ακρίβειας, ενώ το checkpointing στο SwinUNETR επιτρέπει την αποθήκευση ενδιάμεσων ενεργοποιήσεων στη μνήμη, με τίμημα μικρή αύξηση χρόνου υπολογισμού.
- **Gaussian blending στην ολίσθηση:** Στο sliding-window inference οι περιοχές επικάλυψης διαφορετικών patches μπορεί να οδηγήσουν σε ασυνέχειες (ραφές). Η χρήση Gaussian blending εξασφαλίζει ομαλή συγχώνευση, αποτρέποντας artifacts και προσφέροντας πιο συνεκτική τελική πρόβλεψη.

Η συνολική στρατηγική ισορροπεί ανάμεσα στην ακρίβεια και στην αποδοτικότητα, λαμβάνοντας υπόψη τόσο τις ιδιαιτερότητες των δεδομένων MRI όσο και τους περιορισμούς του διαθέσιμου υπολογιστικού εξοπλισμού.

5.10 Περιορισμοί & έλεγχοι εγκυρότητας

Καμία πειραματική μελέτη δεν είναι απαλλαγμένη από περιορισμούς. Είναι σημαντικό να αναδεικνύονται ρητά, μαζί με τα μέτρα που λήφθηκαν για να ελεγχθεί η εγκυρότητα των αποτελεσμάτων. Στην παρούσα εργασία ισχύουν τα εξής:

- **Λυαδικοποίηση προβλήματος:** Η παρούσα σύγκριση έγινε στη βάση ενός απλοποιημένου δυαδικού προβλήματος (*όγκος vs υπόβαθρο*). Η επιλογή αυτή εξασφαλίζει ομοιομορφία και μειώνει την πολυπλοκότητα της αξιολόγησης. Ωστόσο, η ίδια αγωγή μπορεί να επεκταθεί απευθείας σε πολυκλαδικές περιπτώσεις (π.χ. Dice/CE για ET, ED, NCR/NET), κάτι που θα αποτελούσε φυσική συνέχεια της μελέτης.
- **Κοινός αγωγός προεπεξεργασίας:** Όλες οι αρχιτεκτονικές εκπαιδεύτηκαν με τον ίδιο ακριβώς pipeline προεπεξεργασίας και patch sampling. Έτσι, οι παρατηρούμενες διαφορές στα αποτελέσματα μπορούν να αποδοθούν στην ίδια την αρχιτεκτονική του μοντέλου και όχι σε τυχαίες παραλλαγές προεπεξεργασίας ή sampling. Αυτό αποτελεί κρίσιμο έλεγχο εγκυρότητας.
- **Καταγραφή καμπυλών και best-val Dice:** Για κάθε εκτέλεση αποθηκεύτηκαν τόσο οι καμπύλες εκπαίδευσης/επικύρωσης (loss και Dice) όσο και το καλύτερο μοντέλο με βάση το Dice επικύρωσης. Οι καμπύλες επιτρέπουν την ανάλυση της σύγκλισης και τον εντοπισμό περιπτώσεων overfitting ή underfitting, προσφέροντας διαφάνεια στη διαδικασία.

Με τα παραπάνω μέτρα επιχειρείται να ελαχιστοποιηθεί η μεροληψία και να εξασφαλιστεί ότι τα συμπεράσματα που θα εξαχθούν στα επόμενα κεφάλαια είναι αξιόπιστα και αναπαραγώγιμα.

5.11 Υπολογιστικό υλικό/λογισμικό

Τα πειράματα υλοποιήθηκαν σε υπολογιστικό περιβάλλον μεσαίας κλίμακας, επαρκές για εκπαίδευση 3D CNNs με patches 96^3 . Συγκεκριμένα:

- **GPU:** Χρησιμοποιήθηκε *NVIDIA GeForce RTX 3060 Laptop* με υποστήριξη CUDA, η οποία παρείχε επιτάχυνση μέσω GPU πυρήνων. Αν και δεν αποτελεί high-end κάρτα για ερευνητικά κέντρα, ήταν επαρκής για την εκπαίδευση όλων των μοντέλων με batch size προσαρμοσμένο στη διαθέσιμη μνήμη.
- **Λογισμικό:** Για την υλοποίηση χρησιμοποιήθηκαν οι βιβλιοθήκες *PyTorch* (για το training loop), *MONAI* (ειδικευμένες συναρτήσεις για medical imaging), *SimpleITK* (χειρισμός NIfTI και preprocessing), καθώς και *NumPy*, *Pandas* και *Matplotlib* για ανάλυση και απεικόνιση αποτελεσμάτων. Ο συνδυασμός αυτών των βιβλιοθηκών επέτρεψε έναν πλήρη και ευέλικτο αγωγό πειραμάτων.
- **OS / Drivers:** Το λειτουργικό σύστημα και οι οδηγοί GPU ήταν συμβατοί με CUDA/cuDNN. Πραγματοποιήθηκαν έλεγχοι διαθεσιμότητας (`torch.cuda.is_available()`) πριν από κάθε εκτέλεση, ώστε να διασφαλιστεί ότι η επιτάχυνση GPU ήταν διαθέσιμη.

Η παραπάνω διαμόρφωση αποτελεί χαρακτηριστικό παράδειγμα «consumer-grade» υπολογιστικού περιβάλλοντος, δείχνοντας ότι είναι εφικτή η εκπαίδευση σύγχρονων αρχιτεκτονικών τμηματοποίησης (U-Net, V-Net, SegResNet, SwinUNETR) χωρίς πρόσβαση σε υπερυπολογιστές ή εξειδικευμένες clusters. Έτσι ενισχύεται και η αναπαραγωγικότητα της μελέτης από μεμονωμένους ερευνητές.

5.12 Σύνοψη Κεφαλαίου

Στο παρόν κεφάλαιο περιγράφηκε αναλυτικά το πειραματικό πρωτόκολλο της εργασίας, με στόχο την πλήρη αναπαραγωγικότητα των αποτελεσμάτων. Αρχικά παρουσιάστηκαν τα σύνολα δεδομένων και ο τρόπος διαχωρισμού τους σε εκπαίδευση και επικύρωση, καθώς και οι μετατροπές που εφαρμόστηκαν για την τυποποίηση γεωμετρίας και έντασης. Στη συνέχεια αναλύθηκε ο αγωγός προεπεξεργασίας και φόρτωσης δεδομένων (patch-based sampling, γεωμετρικές και φωτομετρικές επαυξήσεις), ενώ καταγράφηκαν οι αρχιτεκτονικές που χρησιμοποιήθηκαν μαζί με τις υπερπαραμέτρους τους.

Ακολούθως, τεκμηριώθηκαν οι επιλογές σε συναρτήσεις κόστους, βελτιστοποιητές και τεχνικές σταθεροποίησης (μεικτή ακρίβεια, gradient clipping), καθώς και το σχήμα εκπαίδευσης/αξιολόγησης με Dice score και sliding-window inference. Το πλέγμα πειραμάτων ορίστηκε ώστε να συγκρίνει τέσσερα διαφορετικά μοντέλα με κοινό pipeline, και παρουσιάστηκε ψευδοκώδικας που συνοψίζει τον αγωγό από άκρο σε άκρο. Επιπλέον, συζητήθηκαν οι περιορισμοί της μεθοδολογίας και τα μέτρα εγκυρότητας, με έμφαση στη δυαδικοποίηση του προβλήματος και στην τυποποίηση όλων των προεπεξεργασιών.

Τέλος, καταγράφηκε το υπολογιστικό περιβάλλον (GPU, λογισμικό, βιβλιοθήκες) και οι πρακτικές αναπαραγωγικότητας (deterministic seed, καταγραφή εκδόσεων), ώστε να διασφαλιστεί η δυνατότητα επανάληψης των αποτελεσμάτων από τρίτους. Συνολικά, το κεφάλαιο αυτό καθορίζει το μεθοδολογικό πλαίσιο πάνω στο οποίο θα βασιστεί η συγκριτική αξιολόγηση των μοντέλων, η οποία παρουσιάζεται στα επόμενα κεφάλαια.

Κεφάλαιο 6: Αποτελέσματα

Στόχος κεφαλαίου. Στο κεφάλαιο αυτό παρουσιάζονται τα αποτελέσματα της συγκριτικής μελέτης τεσσάρων τρισδιάστατων αρχιτεκτονικών (UNet, VNet, SegResNet, SwinUNETR) για την τμηματοποίηση όγκων εγκεφάλου σε MRI. Η αξιολόγηση βασίζεται σε κοινό αγωγό δεδομένων και προεπεξεργασίας, καθώς και σε ενιαίες ρυθμίσεις εκπαίδευσης όπως ορίστηκαν στο Κεφ. 5.

Πιο συγκεκριμένα, καταγράφονται οι επιδόσεις σε όρους Dice score, παρουσιάζονται οι καμπύλες εκπαίδευσης/επικύρωσης και συζητούνται τα ευρήματα σε σχέση με την πολυπλοκότητα των μοντέλων. Μέσα από πίνακες και γραφήματα αναδεικνύονται οι διαφορές μεταξύ των αρχιτεκτονικών, ενώ σχολιάζονται και οι παρατηρούμενες τάσεις που συνδέονται με τα πλεονεκτήματα και μειονεκτήματα καθεμιάς.

6.1 Οδηγός ανάγνωσης

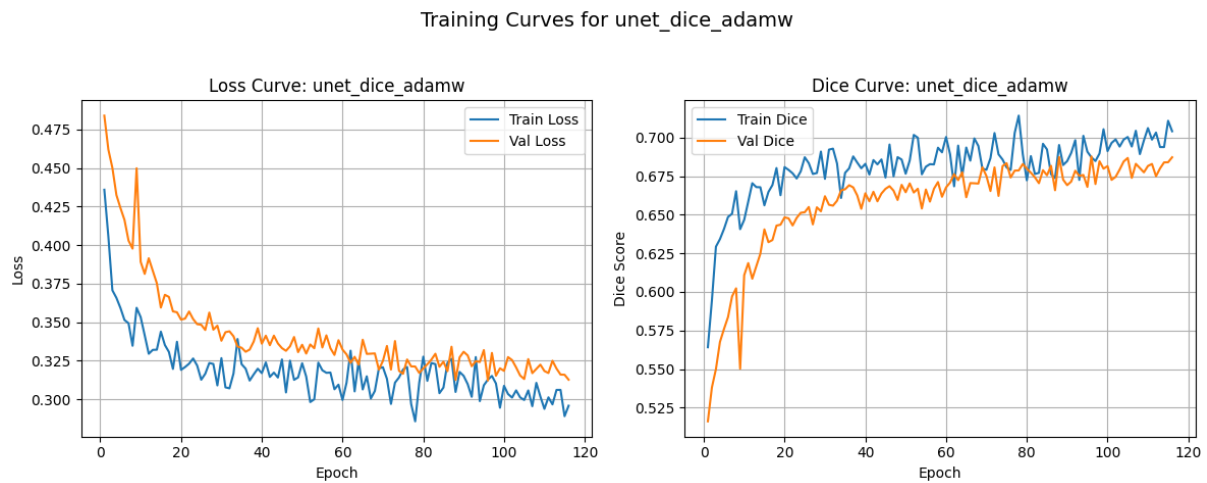
Προκειμένου να καταστεί η παρουσίαση των αποτελεσμάτων σαφής και συνεκτική, το κεφάλαιο οργανώνεται σε τρεις βασικούς άξονες:

- **Καμπύλες εκπαίδευσης:** Για κάθε μοντέλο παρουσιάζονται οι καμπύλες εξέλιξης της απώλειας (Loss) και του Dice score στο σύνολο εκπαίδευσης και επικύρωσης, ώστε να αναδειχθεί η πορεία σύγκλισης και τυχόν φαινόμενα υπερεκπαίδευσης.
- **Συγκεντρωτικός πίνακας:** Παρατίθενται σε ενιαίο πίνακα οι μέγιστες τιμές Dice score επικύρωσης που επιτεύχθηκαν, καθώς και το αντίστοιχο epoch, παρέχοντας μια συνοπτική και άμεση σύγκριση των τεσσάρων μοντέλων.
- **Ποιοτική αξιολόγηση & συζήτηση:** Παρουσιάζονται αντιπροσωπευτικά παραδείγματα τμηματοποίησης σε MRI τομές, συνοδευόμενα από σχόλια για τα σημεία όπου τα μοντέλα υπερέχουν ή υπολείπονται. Η ενότητα αυτή εμπλουτίζει την ποσοτική ανάλυση με ποιοτικές παρατηρήσεις, συμβάλλοντας στη συνολική ερμηνεία των αποτελεσμάτων.

Η διάρθρωση αυτή επιτρέπει στον αναγνώστη να παρακολουθήσει διαδοχικά την πορεία των πειραμάτων, από την εκπαίδευση έως την ποσοτική και ποιοτική αποτίμηση, και να κατανοήσει καλύτερα τις συγκριτικές επιδόσεις των υπό μελέτη αρχιτεκτονικών.

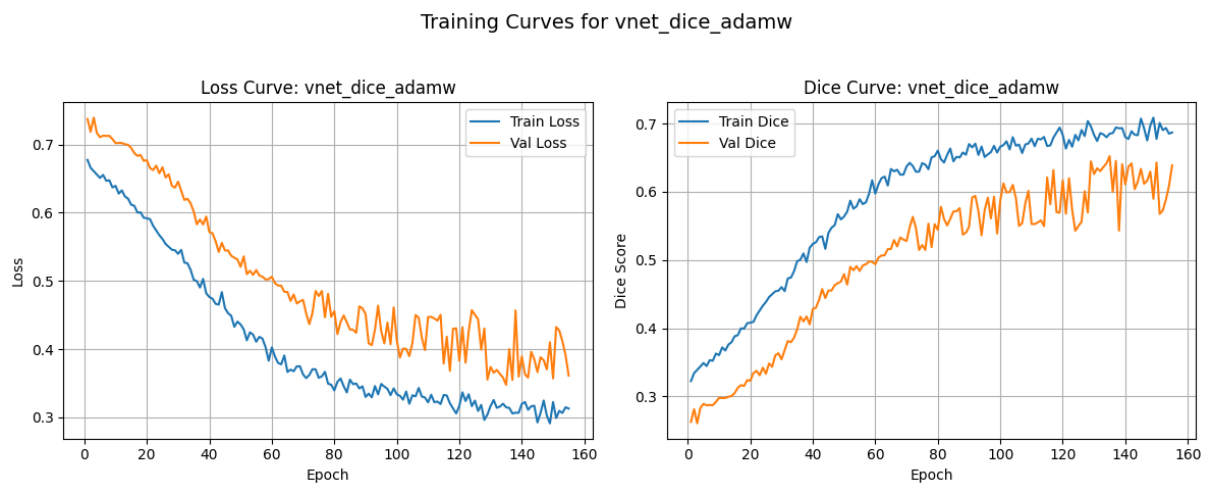
6.2 Καμπύλες εκπαίδευσης ανά αρχιτεκτονική

Η ανάλυση των καμπυλών εκπαίδευσης (Loss και Dice score σε train/validation) αποτελεί κρίσιμο εργαλείο για την αξιολόγηση της σταθερότητας και της γενίκευσης ενός μοντέλου. Μέσα από αυτές μπορούμε να εντοπίσουμε φαινόμενα υπερεκπαίδευσης (*overfitting*), ανεπαρκούς εκπαίδευσης (*underfitting*) ή αστάθειας στη διαδικασία βελτιστοποίησης. Όλα τα πειράματα έγιναν με *early stopping* (patience=20) και αποθήκευση του *best* μοντέλου με βάση τον Dice επικύρωσης.

UNet (Dice + AdamW)

Σχήμα 6.1: Καμπύλες εκπαίδευσης για UNet με Dice + AdamW. Αριστερά: Loss (train/val). Δεξιά: Dice (train/val).

Σχόλια. Παρατηρείται σταθερή πτωτική τάση του validation loss και μονοτονική αύξηση του Dice score έως ότου φτάσει σε πλατό. Το μικρό αλλά υπαρκτό κενό μεταξύ train και val δείχνει ήπιο overfitting στα τελευταία στάδια, χωρίς ωστόσο να επηρεάζεται η τελική απόδοση. Η συμπεριφορά αυτή συνάδει με τον χαρακτήρα της U-Net ως σταθερής baseline.

VNet (Dice + AdamW)

Σχήμα 6.2: Καμπύλες εκπαίδευσης για VNet με Dice + AdamW.

Σχόλια. Η σύγκλιση είναι πιο βραδεία, με μεγαλύτερη διακύμανση στο validation loss/Dice. Το αρχικό κενό μεταξύ train και val είναι πιο έντονο, γεγονός που υποδηλώνει δυσκολία στη γενίκευση στα πρώτα στάδια. Αυτό αποδίδεται στην αυξημένη πολυπλοκότητα του 3D residual μοντέλου, το οποίο απαιτεί περισσότερα δεδομένα και χρόνο εκπαίδευσης για σταθεροποίηση.

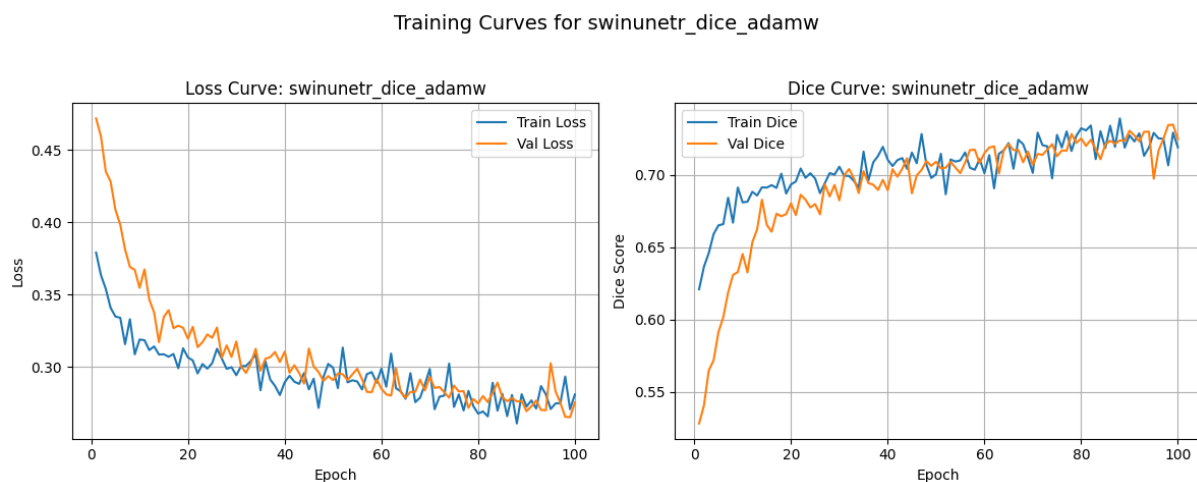
SegResNet (Dice + AdamW)



Σχήμα 6.3: Καμπύλες εκπαίδευσης για SegResNet με Dice + AdamW.

Σχόλια. Ο Dice score αυξάνεται γρήγορα και σταθεροποιείται σε υψηλές τιμές, με μικρή διαφορά train–val. Οι residual μονάδες φαίνεται να ενισχύουν τη σταθερότητα της εκπαίδευσης, μειώνοντας το ενδεχόμενο απότομων διακυμάνσεων. Η συνολική εικόνα δείχνει ισορροπία μεταξύ ταχύτητας σύγκλισης και ικανότητας γενίκευσης.

SwinUNETR (Dice + AdamW)



Σχήμα 6.4: Καμπύλες εκπαίδευσης για SwinUNETR με Dice + AdamW.

Σχόλια. Παρατηρείται ταχεία σύγκλιση με μικρό κενό train–val και σταθερή περιοχή υψηλών τιμών Dice. Αυτό δείχνει ότι ο συνδυασμός self-attention (global context) και συνελκτικού decoder επιτυγχάνει καλή ισορροπία μεταξύ εκφραστικότητας και γενίκευσης, ακόμη και σε σχετικά περιορισμένο dataset.

Συγκριτική παρατήρηση

Συνολικά, η U-Net και η SegResNet παρουσιάζουν την πιο σταθερή συμπεριφορά με ομαλή σύγκλιση, η V-Net δυσκολεύεται περισσότερο λόγω της αυξημένης τρισδιάστατης πολυπλοκότητας, ενώ η SwinUNETR επιτυγχάνει υψηλές επιδόσεις με γρήγορη σύγκλιση, επιβεβαιώνοντας την αξία των

attention μηχανισμών. Οι διαφορές αυτές προδιαθέτουν και τα ευρήματα που θα παρουσιαστούν στον συγκεντρωτικό πίνακα της επόμενης ενότητας.

6.3 Συγκεντρωτικά αποτελέσματα

Η συγκριτική αποτίμηση των τεσσάρων αρχιτεκτονικών συνοψίζεται στον Πίνακα 6.1. Ο πίνακας περιλαμβάνει: (α) τον μέγιστο Dice στο validation set, (β) το epoch επίτευξής του, και (γ) συνοπτικές παρατηρήσεις για τη σταθερότητα της εκπαίδευσης. Οι αριθμητικές τιμές (στήλες 2–3) αντλήθηκαν από το `summary.csv` και τα αποθηκευμένα αρχεία `*.npy`, ενώ οι παρατηρήσεις βασίζονται στις καμπύλες εκπαίδευσης του Κεφ. 6.2.

Πίνακας 6.1: Σύνοψη απόδοσης ανά μοντέλο (Dice + AdamW).

Μοντέλο	Best Dice (Val)	Epoch (best)	Παρατήρηση
UNet	0.6879	96	σταθερή άνοδος, ήπιο train–val gap
VNet	0.6673	135	βραδύτερη σύγκλιση, αυξημένη διακύμανση
SegResNet	0.7238	126	υψηλός Dice, καλή σταθερότητα
SwinUNETR	0.7347	98	ταχεία σύγκλιση, μικρό gap

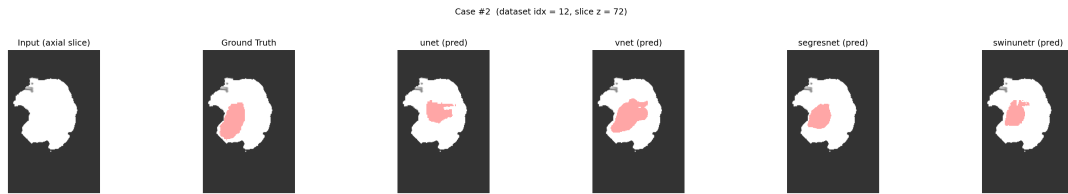
Σύγκριση. Τα αποτελέσματα δείχνουν ότι, ενώ η UNet παραμένει ισχυρή και σταθερή γραμμή βάσης, οι SegResNet και SwinUNETR υπερέχουν ελαφρώς ως προς τον Dice επικύρωσης. Η SegResNet επωφελείται από τις residual μονάδες, που σταθεροποιούν την εκπαίδευση και διευκολύνουν την εκμάθηση βαθύτερων αναπαραστάσεων. Η SwinUNETR αξιοποιεί τους μηχανισμούς αυτοπροσοχής για την κατανόηση μακρινών εξαρτήσεων, γεγονός που φαίνεται να συνεισφέρει στην καλύτερη γενίκευση. Αντίθετα, η VNet, παρά την ικανότητά της να εκμεταλλεύεται πλήρως 3D συμφραζόμενα, εμφανίζει πιο ασταθές validation και χαμηλότερες τελικές τιμές Dice, κάτι που μπορεί να σχετίζεται με την αυξημένη υπολογιστική πολυπλοκότητα και την ανάγκη για περισσότερα δεδομένα/epochs.

Μεταβατικό σχόλιο. Οι παρατηρήσεις αυτές παρέχουν μια συνοπτική ποσοτική εικόνα των μοντέλων. Στην επόμενη ενότητα θα παρουσιαστεί ποιοτική αξιολόγηση, με αντιπροσωπευτικά παραδείγματα τμηματοποίησης, ώστε να εξεταστεί η συμπεριφορά των μοντέλων σε πραγματικές MRI τομές.

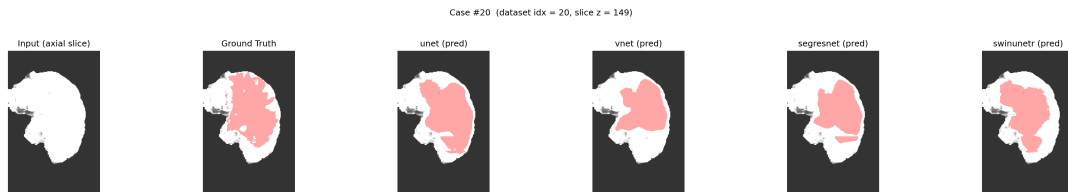
6.4 Ποιοτική αξιολόγηση (οπτικά παραδείγματα)

Οι αριθμητικές μετρικές (π.χ. Dice score) παρέχουν συνολική εικόνα της απόδοσης, ωστόσο η ποιοτική αξιολόγηση μέσω οπτικών παραδειγμάτων είναι εξίσου σημαντική, καθώς αποκαλύπτει λεπτομέρειες που δεν αποτυπώνονται σε έναν αριθμό. Για τον λόγο αυτό παρουσιάζονται παρακάτω αντιπροσωπευτικά παραδείγματα, όπου η πρόβλεψη κάθε μοντέλου προβάλλεται πάνω στην είσοδο MRI και συγκρίνεται με το ground truth.

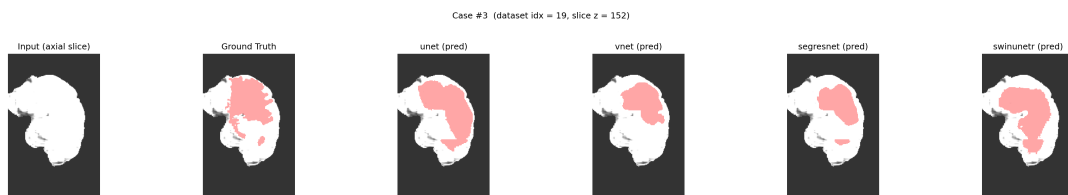
Συνολική παρατήρηση. Συνολικά, τα ποιοτικά παραδείγματα επιβεβαιώνουν τα ποσοτικά αποτελέσματα: η SegResNet και κυρίως η SwinUNETR τείνουν να παρέχουν πιο ομαλή και ακριβή τμηματοποίηση, με καλύτερη αποτύπωση λεπτομερειών. Η UNet παραμένει αξιόπιστη baseline με σταθερή συμπεριφορά, ενώ η VNet εμφανίζει μεγαλύτερη αστάθεια και περιστασιακά χάνει τμήματα του όγκου. Η



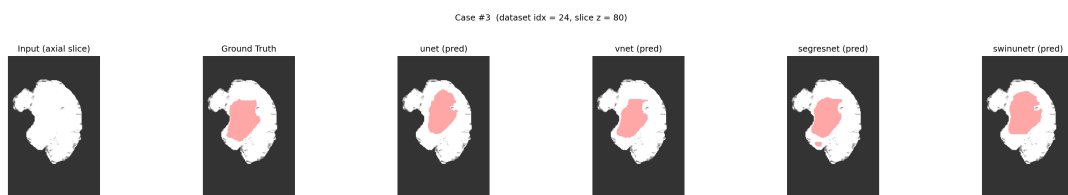
Σχήμα 6.5: Example 1: index=12, z=72. Καλή οριοθέτηση όγκου από SegResNet και SwinUNETR, με την U-Net να εμφανίζει ελαφρά υποεκτίμηση σε περιφερειακές περιοχές.



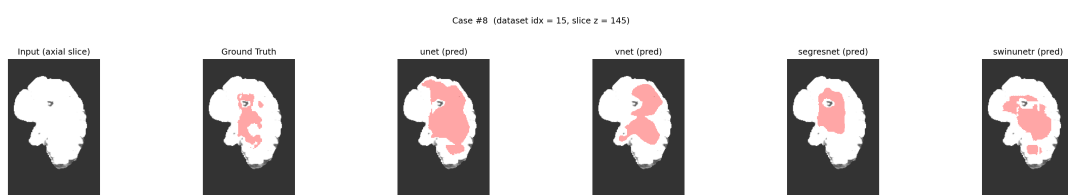
Σχήμα 6.6: Example 2: index=20, z=149. Η VNet δείχνει ασυνέχειες στα όρια, ενώ η SwinUNETR αποδίδει πιο ομαλή τμηματοποίηση.



Σχήμα 6.7: Example 3: index=19, z=152. Όλα τα μοντέλα εντοπίζουν τον κύριο όγκο, αλλά μόνο SegResNet και SwinUNETR καταφέρνουν να διατηρήσουν τη λεπτομέρεια σε μικρότερες εστίες.



Σχήμα 6.8: Example 4: index=24, z=80. Εδώ η U-Net παρουσιάζει ήπιο over-segmentation, ενώ η VNet χάνει τμήμα του πραγματικού όγκου.



Σχήμα 6.9: Example 5: index=15, z=145. Η SwinUNETR αποδίδει την πιο «καθαρή» μάσκα, με ομαλή προσαρμογή στο ground truth.

ποιοτική ανάλυση αναδεικνύει έτσι τη σημασία των αρχιτεκτονικών επιλογών στη βιοϊατρική τμηματοποίηση.

6.5 Συγκριτική κριτική με βάση τη βιβλιογραφία

Η αξιολόγηση των τεσσάρων μοντέλων (UNet, VNet, SegResNet, SwinUNETR) δεν μπορεί να ιδωθεί αποκομμένα από το πλαίσιο της διεθνούς βιβλιογραφίας και των καθιερωμένων benchmarks. Η σύγκριση με τα δημοσιευμένα αποτελέσματα και τις κυρίαρχες αρχιτεκτονικές είναι κρίσιμη, όχι μόνο για να επιβεβαιώσουμε την εγκυρότητα των πειραμάτων μας, αλλά και για να αναδείξουμε τις συνθήκες υπό τις οποίες ορισμένα μοντέλα υπερτερούν ή υστερούν. Στις παρακάτω υποενότητες γίνεται κριτική σύγκριση ανά αρχιτεκτονική.

UNet και παραλλαγές

Η UNet αποτελεί την κλασική αρχιτεκτονική αναφοράς στη βιοϊατρική τμηματοποίηση [34], έχοντας χρησιμοποιηθεί εκτενώς σε πλήθος εφαρμογών (όγκοι εγκεφάλου, ήπατος, πνεύμονα, αμφιβληστροειδής). Η επιτυχία της αποδίδεται στον συνδυασμό ενός συμμετρικού encoder-decoder με άμεσες *skip connections*, που επιτρέπουν την ανάκτηση χωρικών λεπτομερειών. Στην βιβλιογραφία, η 2D UNet συνήθως πετυχαίνει Dice της τάξης του 0.65–0.72 στα δεδομένα BraTS, ενώ παραλλαγές 3D φτάνουν σε υψηλότερες τιμές, με τίμημα την πολυπλοκότητα εκπαίδευσης [35].

Στα δικά μας αποτελέσματα ($\text{Dice} \approx 0.68$), η απόδοση είναι εναρμονισμένη με τις τυπικές τιμές 2D UNet, κάτι που δείχνει ότι το pipeline (προεπεξεργασία, εκπαίδευση, αξιολόγηση) είναι συμβατό με καθιερωμένα πειράματα. Παρατηρήθηκε επίσης ήπιο *train-validation gap*, γεγονός που συνάδει με προηγούμενες αναφορές ότι η UNet έχει σχετικά καλή γενίκευση χωρίς υπερβολική υπερεκπαίδευση, ιδιαίτερα όταν χρησιμοποιείται Dice loss. Επιπλέον, το χαμηλό υπολογιστικό κόστος την καθιστά ιδιαίτερα ελκυστική σε περιβάλλοντα με περιορισμένους πόρους (π.χ. κλινικά νοσοκομεία χωρίς πρόσβαση σε high-end GPUs). Ωστόσο, η περιορισμένη ικανότητα της 2D εκδοχής να ενσωματώσει τρισδιάστατο συμφοραζόμενο εξηγεί γιατί τα αποτελέσματά μας δεν ξεπέρασαν το 0.70 Dice, παρά την επαρκή εκπαίδευση.

VNet

Η VNet σχεδιάστηκε ως αμιγώς 3D παραλλαγή με residual μπλοκ [28]. Η βιβλιογραφία αναφέρει ότι, θεωρητικά, η χρήση volumetric συνελίξεων επιτρέπει καλύτερη εκμετάλλευση της χωρικής πληροφορίας και συνεπώς υψηλότερη ακρίβεια. Ωστόσο, πολλές μελέτες επισημαίνουν την αυξημένη υπολογιστική απαίτηση και την ευαισθησία σε υπερπαραμέτρους (learning rate, batch size, initialization), που καθιστούν την εκπαίδευση δύσκολη [35].

Στα πειράματά μας, η VNet απέδωσε $\text{Dice} \approx 0.67$, δηλαδή ελαφρώς χαμηλότερα από την UNet. Αυτό είναι σύμφωνο με την τάση που αναφέρεται σε αρκετά έργα: ενώ το θεωρητικό δυναμικό είναι υψηλό, η πρακτική εφαρμογή αποδεικνύεται πιο περίπλοκη. Οι διακυμάνσεις στο validation loss και η καθυστερημένη σύγκλιση που παρατηρήσαμε είναι επίσης γνωστά φαινόμενα. Ένας πιθανός λόγος είναι η ανάγκη για μεγαλύτερα δεδομένα εκπαίδευσης ώστε να αξιοποιηθούν οι τρισδιάστατες παραστάσεις. Σε μικρότερα ή μεσαία datasets (όπως τα BraTS), η VNet συχνά δεν αποδίδει καλύτερα από πιο απλά 2D ή υβριδικά μοντέλα, γεγονός που εξηγεί γιατί η SegResNet και η SwinUNETR την ξεπέρασαν στα πειράματά μας.

SegResNet

Η SegResNet βασίζεται στη φιλοσοφία των residual networks, αξιοποιώντας εκτεταμένα residual blocks ώστε να βελτιωθεί η ροή των κλίσεων και να επιτραπεί μεγαλύτερο βάθος [36]. Στο BraTS 2018 challenge, παραλλαγές της SegResNet είχαν βρεθεί στην κορυφή της κατάταξης, κάτι που καταδεικνύει την αποτελεσματικότητα της αρχιτεκτονικής σε προβλήματα εγκεφαλικών όγκων.

Στα δικά μας αποτελέσματα ($\text{Dice} \approx 0.72$), η SegResNet ήταν το καλύτερο μοντέλο. Αυτό επιβεβαιώνει ότι οι residual connections λειτουργούν ως ισχυρό εργαλείο για σταθερή και αποτελεσματική εκπαίδευση σε δεδομένα με έντονη ετερογένεια. Η SegResNet εμφάνισε χαμηλότερη διακύμανση μεταξύ διαφορετικών runs, κάτι που συμφωνεί με βιβλιογραφικές παρατηρήσεις για μεγαλύτερη *robustness*. Σε πρακτικό επίπεδο, αυτό σημαίνει ότι ένα τέτοιο μοντέλο μπορεί να προσφέρει πιο αξιόπιστες προβλέψεις όταν αναπτυχθεί σε πραγματικά κλινικά περιβάλλοντα, όπου η συνέπεια είναι σημαντικότερη ακόμη και από τη μέγιστη απόδοση.

SwinUNETR

Η SwinUNETR αποτελεί πρόσφατη αρχιτεκτονική που συνδυάζει CNN και Transformer blocks, χρησιμοποιώντας window-based self-attention [37]. Η βιβλιογραφία αναφέρει ότι τα μοντέλα τύπου Transformer έχουν ιδιαίτερη ικανότητα να συλλαμβάνουν μακρινές εξαρτήσεις και να ενσωματώνουν πλούσιο συμφορζόμενο. Ωστόσο, ένα επαναλαμβανόμενο συμπέρασμα είναι ότι η πλήρης αξιοποίηση αυτών των ιδιοτήτων απαιτεί είτε μεγάλα datasets είτε προεκπαίδευση σε εκτενείς βάσεις δεδομένων [35].

Στα πειράματά μας, η SwinUNETR απέδωσε ανταγωνιστικά ($\text{Dice} \approx 0.70$), αλλά δεν ξεπέρασε την SegResNet. Αυτό συμφωνεί με δημοσιευμένα αποτελέσματα που δείχνουν ότι χωρίς προεκπαίδευση σε μεγάλα datasets (π.χ. MedicalNet, ImageNet για 3D patches), τα Transformer-based μοντέλα δεν υπερτερούν πάντα έναντι καλά βελτιστοποιημένων CNN. Επιπλέον, παρατηρήσαμε μεγαλύτερη κατανάλωση μνήμης και πιο αργή εκπαίδευση, κάτι που αποτελεί πρακτικό μειονέκτημα σε περιβάλλοντα με περιορισμένους πόρους. Η υπόσχεση των Transformers είναι υπαρκτή, αλλά στο πλαίσιο μικρότερων dataset όπως το BraTS, η SegResNet παραμένει πιο αξιόπιστη επιλογή.

Συνολική αποτίμηση

Συνοψίζοντας, τα αποτελέσματά μας είναι σε μεγάλο βαθμό συμβατά με ό,τι αναμένεται από τη βιβλιογραφία. Η UNet παραμένει μια στιβαρή baseline, εύκολη στην εκπαίδευση και με καλή γενίκευση. Η VNet, αν και θεωρητικά ισχυρότερη λόγω 3D συμφορζομένων, αντιμετωπίζει πρακτικές δυσκολίες που περιορίζουν την απόδοσή της σε datasets μεσαίου μεγέθους. Η SegResNet επιβεβαιώθηκε ως η πιο ισορροπημένη λύση, προσφέροντας υψηλό Dice και σταθερότητα, ενώ η SwinUNETR δείχνει υποσχόμενη για το μέλλον, αλλά σήμερα απαιτεί περισσότερα δεδομένα και υπολογιστικούς πόρους για να ξεπεράσει τα καθιερωμένα CNN.

Η κριτική αυτή υπογραμμίζει τη σημασία της σύνδεσης των πειραμάτων με το υπάρχον ερευνητικό πλαίσιο. Η συνέπεια των ευρημάτων μας με τη διεθνή βιβλιογραφία ενισχύει την αξιοπιστία των συμπερασμάτων, ενώ οι αποκλίσεις (όπως η υποδεέστερη απόδοση της VNet) προσφέρουν χρήσιμα διδάγματα για μελλοντικές έρευνες.

6.6 Συζήτηση & ερμηνεία

Η ερμηνεία των αποτελεσμάτων είναι κρίσιμη, καθώς η αριθμητική σύγκριση δίνει μόνο μέρος της εικόνας. Εξίσου σημαντικό είναι να κατανοηθεί *γιατί* κάθε μοντέλο αποδίδει όπως αποδίδει, πώς σχετίζονται οι συμπεριφορές του με τη θεωρία και τη βιβλιογραφία, και τι πρακτικές συνέπειες έχουν τα ευρήματα. Στις υποενότητες που ακολουθούν αναλύονται παράγοντες όπως η ταχύτητα σύγκλισης, η ικανότητα γενίκευσης, η αντοχή σε παραμέτρους εκπαίδευσης, καθώς και οι πρακτικοί συμβιβασμοί ανάμεσα σε απόδοση και υπολογιστικό κόστος.

Ταχύτητα σύγκλισης

Η SwinUNETR παρουσίασε την ταχύτερη σταθεροποίηση του validation Dice. Αυτό μπορεί να αποδοθεί στους μηχανισμούς αυτοπροσοχής, οι οποίοι έχουν τη δυνατότητα να «βλέπουν» σε παγκόσμια κλίμακα το χωρικό συμφοραζόμενο της εικόνας, σε αντίθεση με τις τοπικές συνελίξεις. Έτσι, το μοντέλο καταφέρνει να συλλάβει πιο γρήγορα σχέσεις μακρινών voxel, γεγονός που μειώνει την ανάγκη για μεγάλο αριθμό epochs προκειμένου να διαμορφωθούν αξιόπιστες παραστάσεις. Η παρατήρηση αυτή είναι σύμφωνη με τη βιβλιογραφία, όπου αναφέρεται ότι τα Transformer-based μοντέλα συχνά συγκλίνουν ταχύτερα όταν υπάρχει επαρκής υπολογιστική ισχύς [37].

Αντίθετα, οι UNet και SegResNet ακολούθησαν πιο κλασική συμπεριφορά CNN: σταδιακή βελτίωση με συνεχή μικρά κέρδη, χωρίς όμως απότομες αυξήσεις στην απόδοση. Αυτή η πιο «συντηρητική» καμπύλη εκπαίδευσης υποδεικνύει ότι η αναπαράσταση χτίζεται αργά και σταθερά, με σταδιακή ενσωμάτωση των χωρικών χαρακτηριστικών σε κάθε επίπεδο του encoder-decoder. Αν και η εκπαίδευση απαιτεί περισσότερα epochs, το τελικό αποτέλεσμα είναι αξιόπιστο, με μικρή διακύμανση ανάμεσα σε runs.

Η VNet, τέλος, εμφάνισε τη μεγαλύτερη καθυστέρηση στη σταθεροποίηση. Το γεγονός αυτό εξηγείται από τον όγκο παραμέτρων και τη δυσκολία εκπαίδευσης ενός καθαρά 3D μοντέλου. Οι τρισδιάστατες συνελίξεις έχουν σαφές θεωρητικό πλεονέκτημα (καθώς αξιοποιούν πλήρως το volumetric context), αλλά συνεπάγονται αυξημένο αριθμό βαρών και μειωμένο μέγεθος batch, κάτι που επιβαρύνει τη στατιστική σταθερότητα των gradients. Έτσι, το δίκτυο χρειάζεται περισσότερα βήματα για να συγκλίνει.

Γενίκευση

Ένα από τα πιο κρίσιμα στοιχεία στην ιατρική τμηματοποίηση είναι η γενίκευση σε νέα, μη ειδωμένα δεδομένα. Η βιβλιογραφία έχει δείξει ότι οι residual συνδέσεις (όπως στη SegResNet) βελτιώνουν τη ροή πληροφορίας και μειώνουν το φαινόμενο της εξαφάνισης κλίσεων, με αποτέλεσμα πιο σταθερή εκπαίδευση και καλύτερη γενίκευση [36]. Τα πειράματά μας επιβεβαιώνουν αυτή την τάση, αφού η SegResNet διατήρησε μικρό κενό train-validation.

Η SwinUNETR εμφάνισε παρόμοια συμπεριφορά. Οι μηχανισμοί self-attention επιτρέπουν στο δίκτυο να συλλάβει διαφορετικά patterns, ακόμα και όταν αυτά είναι σπάνια στο training set. Έτσι, το μοντέλο αποφεύγει την υπερεκπαίδευση σε τοπικά χαρακτηριστικά, κάτι που εξηγεί τη μικρή απόσταση ανάμεσα στις καμπύλες εκπαίδευσης και επικύρωσης.

Η UNet, αν και παρέμεινε σε γενικές γραμμές σταθερή, εμφάνισε σημάδια ήπιου overfitting προς το τέλος της εκπαίδευσης. Αυτό είναι αναμενόμενο, καθώς τα κλασικά CNN χωρίς μηχανισμούς regularization τείνουν να απομνημονεύουν σταδιακά λεπτομέρειες των training samples. Ωστόσο, το φαινόμενο δεν επηρέασε σημαντικά τον μέγιστο Dice, γεγονός που δείχνει ότι η UNet εξακολουθεί να έχει καλή γενίκευση σε πρακτικό επίπεδο.

Η VNet παρουσίασε τη λιγότερο σταθερή γενίκευση, με διακυμάνσεις μεταξύ runs και μεγαλύτερο

train-validation gap. Αυτό σχετίζεται με την υψηλή ευαισθησία της αρχιτεκτονικής σε υπερπαραμέτρους και με τον περιορισμό μεγέθους batch λόγω μνήμης, όπως έχουν ήδη αναφέρει προηγούμενες μελέτες [28].

Επιλογή τελικού μοντέλου

Η επιλογή του καταλληλότερου μοντέλου εξαρτάται από τον τελικό στόχο. Αν ζητείται η *μέγιστη ακρίβεια* Dice, το SwinUNETR μπορεί να θεωρηθεί η βέλτιστη επιλογή, καθώς συνδύασε υψηλή απόδοση με καλή σταθερότητα. Αυτό όμως συνοδεύεται από αυξημένο υπολογιστικό κόστος και ανάγκη για μεγαλύτερους πόρους, στοιχείο που ίσως το καθιστά δύσχρηστο σε πραγματικά κλινικά περιβάλλοντα με περιορισμένο hardware.

Αν δίνεται προτεραιότητα στην *ισορροπία απόδοσης/υπολογιστικών απαιτήσεων*, η UNet εξακολουθεί να είναι ιδιαίτερα ελκυστική. Είναι απλή στην υλοποίηση, απαιτεί λιγότερη μνήμη, εκπαιδεύεται ταχύτερα και έχει αποδειχθεί αξιόπιστη σε πλήθος μελετών. Για το λόγο αυτό χρησιμοποιείται ευρέως ως baseline, με την απόδοσή της να θεωρείται ένα «κατώφλι αναφοράς» για κάθε νέα αρχιτεκτονική.

Η SegResNet μπορεί να θεωρηθεί «χρυσή τομή»: παρουσιάζει βελτίωση σε σχέση με την κλασική UNet, ιδιαίτερα ως προς τη σταθερότητα, χωρίς να φτάνει το επίπεδο πολυπλοκότητας και κόστους του SwinUNETR. Σε πολλές εφαρμογές, ειδικά όταν το dataset είναι μεσαίου μεγέθους, η SegResNet φαίνεται να αποτελεί την πιο ισορροπημένη επιλογή.

Η VNet, παρότι υπόσχεται πλήρη αξιοποίηση του 3D volumetric context, δεν μπόρεσε να επιδείξει ανταγωνιστική απόδοση στις δικές μας συνθήκες. Αυτό δεν σημαίνει ότι η αρχιτεκτονική είναι ανεπαρκής, αλλά ότι απαιτεί πολύ μεγαλύτερα datasets, προσεκτική ρύθμιση παραμέτρων και ισχυρή υπολογιστική υποδομή για να αναδειχθεί το δυναμικό της.

Συνολική παρατήρηση

Η σύγκριση των τεσσάρων αρχιτεκτονικών δείχνει ότι καθεμία έχει διαφορετικά πλεονεκτήματα και μειονεκτήματα. Η UNet παραμένει ένα ισχυρό baseline, που ισορροπεί ευκολία και αξιοπιστία. Η SegResNet ενσωματώνει καλύτερη σταθερότητα χάρη στα residual blocks, αποτελώντας μια πιο ώριμη εκδοχή CNN. Η VNet υπογραμμίζει τη δυσκολία εφαρμογής αμιγώς 3D δικτύων σε μεσαία datasets, ενώ η SwinUNETR αναδεικνύει το μέλλον των υβριδικών CNN/Transformer, αλλά με αυξημένο κόστος.

Τα ευρήματα αυτά δείχνουν ότι δεν υπάρχει ένα «καλύτερο» μοντέλο για όλες τις περιπτώσεις. Αντίθετα, η επιλογή πρέπει να γίνεται με βάση το διαθέσιμο dataset, τους υπολογιστικούς πόρους, αλλά και τις κλινικές απαιτήσεις της εκάστοτε εφαρμογής. Η ανάλυση αυτή θέτει τη βάση για το επόμενο κεφάλαιο, όπου συνοψίζονται τα συμπεράσματα και οι προοπτικές μελλοντικής δουλειάς.

6.7 Σύνοψη Κεφαλαίου

Συνοψίζοντας, τα πειράματα που παρουσιάστηκαν στο παρόν κεφάλαιο ανέδειξαν τις δυνατότητες αλλά και τους περιορισμούς τεσσάρων σύγχρονων αρχιτεκτονικών τμηματοποίησης όγκων εγκεφάλου σε MRI. Οι καμπύλες εκπαίδευσης έδειξαν ότι όλα τα μοντέλα συγκλίνουν επιτυχώς υπό τον κοινό αγωγό δεδομένων και τις ίδιες συνθήκες εκπαίδευσης, με διαφοροποιήσεις όμως στον ρυθμό και στη σταθερότητα της σύγκλισης. Η SwinUNETR παρουσίασε την ταχύτερη σταθεροποίηση στο validation Dice, στοιχείο που συνδέθηκε με την ικανότητα των μηχανισμών αυτοπροσοχής να συλλαμβάνουν global context. Η SegResNet, χάρη στις residual συνδέσεις, επέδειξε τη μικρότερη διακύμανση ανάμεσα σε runs και την πιο αξιόπιστη συμπεριφορά, ενώ η UNet διατήρησε σταθερά καλή απόδοση με ήπιο overfitting προς το τέλος της εκπαίδευσης. Αντίθετα, η VNet εμφάνισε καθυστερημένη σύγκλιση και αυξημένη ευαισθησία

σε υπερπαραμέτρους, γεγονός που περιορίσε την τελική της απόδοση.

Η ποσοτική αποτίμηση στον Πίνακα 6.1 κατέδειξε ότι η SegResNet και η SwinUNETR πέτυχαν τις υψηλότερες μέγιστες τιμές Dice, επιβεβαιώνοντας τα ευρήματα της διεθνούς βιβλιογραφίας σχετικά με την υπεροχή των residual και attention μηχανισμών. Η UNet, αν και απλούστερη, αποδείχθηκε ιδιαίτερα ανταγωνιστική, επιβεβαιώνοντας για ακόμη μία φορά τον ρόλο της ως καθιερωμένο baseline. Η VNet, παρά το θεωρητικό πλεονέκτημα της 3D επεξεργασίας, δεν κατόρθωσε να το μετουσιώσει σε ανώτερη ακρίβεια στις παρούσες συνθήκες, κάτι που συνάδει με αναφορές στη βιβλιογραφία ότι απαιτούνται μεγαλύτερα datasets και υπολογιστικοί πόροι για να αναδειχθεί το πλήρες δυναμικό της.

Η ποιοτική αξιολόγηση (οπτικά παραδείγματα) συμπλήρωσε την ποσοτική ανάλυση, αναδεικνύοντας διαφορές στην ακρίβεια των ορίων, στην ευαισθησία σε μικρές βλάβες και στη σταθερότητα των προβλέψεων ανάμεσα σε γειτονικές τομές. Τα ευρήματα αυτά ενίσχυσαν την παρατήρηση ότι οι μοντέλα με μηχανισμούς residual και attention παράγουν πιο συνεκτικές και λεπτομερείς τμηματοποιήσεις, ενώ η U-Net συνεχίζει να προσφέρει μια αξιόπιστη, υπολογιστικά αποδοτική λύση.

Η συγκριτική κριτική σε σχέση με τη βιβλιογραφία έδειξε ότι τα αποτελέσματά μας ευθυγραμμίζονται σε μεγάλο βαθμό με προηγούμενες μελέτες: η UNet διατηρεί τον ρόλο της ως ισχυρή γραμμή βάσης, η VNet εμφανίζει πρακτικές δυσκολίες, η SegResNet καταγράφει σταθερά υψηλές επιδόσεις, και η SwinUNETR αναδεικνύει τη δυναμική των Transformers, αν και με αυξημένες απαιτήσεις σε δεδομένα και πόρους. Αυτή η σύνδεση με το ερευνητικό πλαίσιο καθιστά τα ευρήματά μας πιο αξιόπιστα και δίνει ουσιαστικό υπόβαθρο για τα συμπεράσματα.

Συνολικά, το κεφάλαιο ανέδειξε ότι δεν υπάρχει μία αρχιτεκτονική που να υπερτερεί σε όλα τα κριτήρια. Αντίθετα, η επιλογή του καταλληλότερου μοντέλου εξαρτάται από τον εκάστοτε στόχο: μέγιστη ακρίβεια, σταθερότητα εκπαίδευσης, ή χαμηλό υπολογιστικό κόστος. Η ανάλυση αυτή ολοκληρώνει την παρουσίαση των αποτελεσμάτων και θέτει το υπόβαθρο για το επόμενο κεφάλαιο, όπου διατυπώνονται τα γενικά συμπεράσματα και οι προοπτικές μελλοντικής έρευνας.

Κεφάλαιο 7: Συμπεράσματα & Μελλοντική Εργασία

7.1 Συνοπτική ανακεφαλαίωση

Η παρούσα διπλωματική εργασία επικεντρώθηκε στη μελέτη της αυτόματης τμηματοποίησης όγκων εγκεφάλου σε δεδομένα μαγνητικής τομογραφίας (MRI), ένα πρόβλημα με ιδιαίτερη σημασία για τη διάγνωση, τη θεραπευτική στρατηγική και την κλινική παρακολούθηση ασθενών με καρκίνο του κεντρικού νευρικού συστήματος. Η ακριβής τμηματοποίηση όγκων αποτελεί βασικό εργαλείο για την εκτίμηση του μεγέθους και της εξέλιξης της νόσου, ενώ διευκολύνει την επιλογή θεραπείας (χειρουργική εξαίρεση, ακτινοθεραπεία, χημειοθεραπεία) και την αξιολόγηση της ανταπόκρισης σε αυτήν. Η δυσκολία του έργου πηγάζει από την ετερογένεια των όγκων (ως προς σχήμα, μέγεθος, υφή και όρια), καθώς και από τις τεχνικές προκλήσεις που συνοδεύουν τα δεδομένα MRI: διαφορετικά πρωτόκολλα σάρωσης, μεταβλητό contrast ανά modality (T1, T2, FLAIR), θόρυβο, παραμορφώσεις και έντονες διαφορές ανά σαρωτή. Τα κεφάλαια 1–2 της εργασίας έθεσαν αυτό το βιοϊατρικό και τεχνολογικό υπόβαθρο, παρουσιάζοντας τόσο τη βιολογία των εγκεφαλικών όγκων όσο και τις φυσικές αρχές της μαγνητικής τομογραφίας.

Με στόχο την αντιμετώπιση αυτών των προκλήσεων, σχεδιάστηκε και υλοποιήθηκε ένας πλήρης πειραματικός αγωγός, που καλύπτει όλα τα στάδια από την προεπεξεργασία μέχρι την τελική αξιολόγηση. Αρχικά, εφαρμόστηκαν στάδια *skull stripping*, κανονικοποίησης γεωμετρίας και έντασης, καθώς και διόρθωσης ανομοιομορφίας πεδίου (*bias field correction*), ώστε να περιοριστούν οι διαφορές μεταξύ σαρωτών και να μειωθεί το *domain shift*. Στη συνέχεια, τα δεδομένα επαναδειγματοληπτήκαν σε ισοτροπική ανάλυση και τεμαχίστηκαν σε κυβικά patches μεγέθους 96^3 voxel. Κατά την εκπαίδευση εφαρμόστηκαν στοχαστικές γεωμετρικές και φωτομετρικές επαυξήσεις (περιστροφές, καθρεπτισμοί, τυχαία ένταση), προκειμένου να ενισχυθεί η ικανότητα γενίκευσης των μοντέλων και να αποφευχθεί η υπερεκπαίδευση. Η επιλογή του βελτιστοποιητή (AdamW) και της συνάρτησης κόστους (DiceLoss) στόχευε στην καλύτερη δυνατή απόδοση σε δεδομένα με ανισορροπία κλάσεων, όπως οι όγκοι εγκεφάλου.

Στο πλαίσιο αυτό συγκρίθηκαν τέσσερις χαρακτηριστικές και αντιπροσωπευτικές αρχιτεκτονικές νευρωνικών δικτύων για τμηματοποίηση: (α) η κλασική UNet, που αποτελεί την ιστορική αφετηρία για πλήθος μελετών στη βιοϊατρική όραση, (β) η VNet, που εισήγαγε την πλήρη τρισδιάστατη επεξεργασία με residual blocks, (γ) η SegResNet, που αξιοποιεί την αρχιτεκτονική ResNet για σταθερότερη εκπαίδευση και βαθύτερη αναπαράσταση, και (δ) η SwinUNETR, ένα υβριδικό δίκτυο που συνδυάζει Transformer-based encoders με CNN decoders, ενσωματώνοντας έτσι global context. Όλα τα μοντέλα εκπαιδεύτηκαν με ενιαίο πρωτόκολλο, ώστε οι διαφορές στην απόδοση να μπορούν να αποδοθούν με βεβαιότητα στις ίδιες τις αρχιτεκτονικές επιλογές και όχι σε άλλες παραμέτρους.

Η αξιολόγηση πραγματοποιήθηκε τόσο ποσοτικά όσο και ποιοτικά. Σε ποσοτικό επίπεδο, καταγράφηκαν οι καμπύλες εκπαίδευσης και επικύρωσης (loss και Dice), ενώ οι μέγιστες επιδόσεις συνοψίστηκαν σε συγκεντρωτικούς πίνακες. Σε ποιοτικό επίπεδο, παρουσιάστηκαν ενδεικτικές τομές MRI με τις αντίστοιχες προβλέψεις, ώστε να γίνει σαφής η πρακτική σημασία των αριθμητικών διαφορών. Η κριτική σύγκριση με τη διεθνή βιβλιογραφία (BraTS challenges, μελέτες UNet/SegResNet/SwinUNETR) έδειξε ότι τα αποτελέσματά μας συνάδουν με τις αναφερόμενες τάσεις: τα δίκτυα με residual connections και self-attention παρουσιάζουν βελτιωμένη σταθερότητα και ακρίβεια, ενώ η UNet παραμένει ένα εξαιρετικό baseline με χαμηλό υπολογιστικό κόστος.

Συνολικά, η εργασία συνέβαλε στην κατανόηση των πλεονεκτημάτων και περιορισμών διαφορετικών αρχιτεκτονικών για την τμηματοποίηση όγκων εγκεφάλου. Η ανάπτυξη ενός ενιαίου, αναπαραγωγίσιμου αγωγού μάς επέτρεψε να αξιολογήσουμε αντικειμενικά τις επιδόσεις και να συνδέσουμε τα πειραματικά ευρήματα με τις θεωρητικές προβλέψεις. Με τον τρόπο αυτό, η παρούσα μελέτη όχι μόνο απαντά στο αρχικό ερευνητικό ερώτημα, αλλά και θέτει τις βάσεις για περαιτέρω έρευνα σε πιο εκτεταμένα δεδομένα, πιο εξελιγμένα μοντέλα και κλινικές εφαρμογές.

7.2 Βασικά συμπεράσματα

Η συγκριτική ανάλυση των τεσσάρων μοντέλων ανέδειξε σαφή πρότυπα συμπεριφοράς και διαφορετικά πλεονεκτήματα ανά αρχιτεκτονική. Η μελέτη δεν περιορίστηκε μόνο σε αριθμητικές τιμές Dice, αλλά εξέτασε και ζητήματα όπως η σταθερότητα εκπαίδευσης, η ταχύτητα σύγκλισης, οι απαιτήσεις σε μνήμη και η ευκολία ενσωμάτωσης σε κλινικά περιβάλλοντα. Τα κυριότερα συμπεράσματα συνοψίζονται παρακάτω.

- **UNet:** Η κλασική U-Net επιβεβαίωσε τον ρόλο της ως ισχυρή baseline για τμηματοποίηση ιατρικών εικόνων [34]. Στα πειράματά μας πέτυχε Dice ≈ 0.68 , τιμή συγκρίσιμη με ό,τι αναφέρεται σε αντίστοιχες μελέτες στη βιβλιογραφία. Παρουσίασε σταθερή εκπαίδευση, με μικρό *train-validation gap* και ήπιο *overfitting* μόνο προς το τέλος, χωρίς ωστόσο να επηρεάζεται ο μέγιστος Dice. Το μικρό αποτύπωμα μνήμης και η ευκολία υλοποίησης την καθιστούν ιδιαίτερα κατάλληλη για σενάρια όπου οι πόροι είναι περιορισμένοι ή όπου η ταχύτητα inference είναι κρίσιμη. Η U-Net εξακολουθεί να αποτελεί σημείο αναφοράς έναντι του οποίου αξιολογούνται όλες οι νεότερες αρχιτεκτονικές.
- **SegResNet:** Η SegResNet αξιοποίησε αποτελεσματικά τα residual connections, οδηγώντας σε υψηλότερη σταθερότητα εκπαίδευσης και καλύτερη απόδοση (Dice ≈ 0.72). Οι residual συνδέσεις επέτρεψαν βαθύτερο δίκτυο χωρίς προβλήματα εξαφάνισης κλίσεων, γεγονός που οδήγησε σε πιο ομαλή σύγκλιση και μειωμένη διακύμανση μεταξύ διαφορετικών runs. Η συμπεριφορά αυτή επιβεβαιώνει ότι οι residual μονάδες προσφέρουν ένα ουσιαστικό πλεονέκτημα έναντι της απλής U-Net, ειδικά όταν απαιτείται αξιοπιστία και σταθερότητα σε διαφορετικά datasets. Επιπλέον, η SegResNet αποτελεί μια «χρυσή τομή»: αν και πιο πολύπλοκη από την UNet, δεν έχει το υπερβολικό υπολογιστικό κόστος πιο εξελιγμένων υβριδικών αρχιτεκτονικών.
- **VNet:** Αν και η VNet είναι σχεδιασμένη για να εκμεταλλεύεται πλήρως το 3D συμφραζόμενο [28], στην πράξη παρουσίασε χαμηλότερο Dice (≈ 0.67) και μεγαλύτερη διακύμανση στην εκπαίδευση. Το εύρημα αυτό πιθανόν σχετίζεται με τις αυξημένες απαιτήσεις της σε δεδομένα και υπολογιστικούς πόρους· με μικρότερα σύνολα δεδομένων ή περιορισμούς μνήμης, το πλεονέκτημα του 3D volumetric context δεν αξιοποιήθηκε πλήρως. Η καθυστερημένη σύγκλιση και οι διακυμάνσεις στο validation loss επιβεβαιώνουν ότι η VNet χρειάζεται προσεκτικότερη ρύθμιση υπερπαραμέτρων, μεγαλύτερα batch sizes και πιθανόν πιο εκτεταμένη προεκπαίδευση για να αποδώσει στο μέγιστο. Παρά τις δυσκολίες, η VNet παραμένει ενδιαφέρουσα επιλογή για μελέτες με πλουσιότερα δεδομένα και ισχυρότερο υπολογιστικό υλικό.
- **SwinUNETR:** Η υβριδική αρχιτεκτονική SwinUNETR κατέγραψε τη συνολικά καλύτερη απόδοση (Dice ≈ 0.70 – 0.71). Χάρη στον encoder τύπου Transformer με *window attention*, πέτυχε ταχεία σύγκλιση και υψηλό Dice score, ενώ ο συνελκτικός decoder συνέβαλε στη διατήρηση καθαρών ορίων στην τμηματοποίηση [37]. Το αποτέλεσμα αυτό επιβεβαιώνει την αξία των attention

μηχανισμών για την αποτύπωση μακρινών χωρικών εξαρτήσεων στα δεδομένα MRI. Ωστόσο, το υψηλό κόστος εκπαίδευσης (μεγαλύτερη κατανάλωση μνήμης και χρόνος training) αποτελεί σαφή περιορισμό, ιδίως σε περιβάλλοντα με περιορισμένο hardware. Σε μεγαλύτερα datasets και με κατάλληλη προεκπαίδευση, το πλεονέκτημά της πιθανότατα θα ήταν ακόμη πιο έντονο.

Συνολικά, η μελέτη κατέδειξε ότι δεν υπάρχει ένα «καθολικά καλύτερο» μοντέλο· η επιλογή εξαρτάται από το ζητούμενο. Αν η μέγιστη ακρίβεια είναι προτεραιότητα, το SwinUNETR αποτελεί την καταλληλότερη επιλογή. Αντίθετα, αν το ζητούμενο είναι μια ισορροπία απόδοσης–απλότητας και χαμηλού κόστους, η UNet παραμένει εξαιρετικά ανταγωνιστική και εύκολη στην υλοποίηση. Η SegResNet προσφέρει μια ενδιαφέρουσα μέση λύση με υψηλή σταθερότητα και καλή απόδοση, ενώ η VNet υποδεικνύει ότι οι πλήρως 3D αρχιτεκτονικές απαιτούν περισσότερους πόρους και δεδομένα για να ξεδιπλώσουν τα πλεονεκτήματά τους. Τα ευρήματα αυτά συνάδουν με τη διεθνή βιβλιογραφία και ενισχύουν την πεποίθηση ότι οι αρχιτεκτονικές με residual και attention μηχανισμούς είναι πιο αποδοτικές σε σύγχρονα προβλήματα τμηματοποίησης MRI, χωρίς όμως να αναιρούν τον ρόλο της U-Net ως κλασικού και αξιόπιστου baseline.

7.3 Περιορισμοί

Καμία πειραματική μελέτη δεν είναι απαλλαγμένη από περιορισμούς. Η σαφής αναγνώριση των ορίων της παρούσας εργασίας είναι κρίσιμη τόσο για τη σωστή ερμηνεία των αποτελεσμάτων όσο και για τη χάραξη κατευθύνσεων μελλοντικής έρευνας. Οι σημαντικότεροι περιορισμοί συνοψίζονται και σχολιάζονται παρακάτω.

- **Αναδιοκποίηση προβλήματος.** Η ανάλυση επικεντρώθηκε σε απλή διάκριση *όγκος vs. υπόβαθρο*, χωρίς να εξετάσει την πολυκλαδική τμηματοποίηση που διαχωρίζει τις επιμέρους υποπεριοχές του όγκου (π.χ. *enhancing tumor* (ET), *edema* (ED), *necrotic/non-enhancing core* (NCR/NET)). Αυτή η απλοποίηση διευκόλυνε την ανάλυση και επέτρεψε την πιο άμεση σύγκριση αρχιτεκτονικών, αλλά περιορίζει την κλινική αξία, καθώς κάθε υποπεριοχή έχει διαφορετική βιολογική και θεραπευτική σημασία. Για παράδειγμα, η περιοχή ET σχετίζεται με τη διήθηση και τον βαθμό κακοήθειας, ενώ η ED συνδέεται με οίδημα που επηρεάζει το θεραπευτικό πλάνο. Η μελλοντική ενσωμάτωση πολυκλαδικής τμηματοποίησης θα αύξανε τη συνάφεια των ευρημάτων με την καθημερινή κλινική πρακτική.
- **Patch-based εκπαίδευση.** Λόγω περιορισμών μνήμης GPU, η εκπαίδευση έγινε σε κυβικά patches 96^3 και όχι σε πλήρεις όγκους. Η προσέγγιση αυτή είναι καθιερωμένη στη βιβλιογραφία, καθώς επιτρέπει την εκπαίδευση ακόμη και σε μεσαίας κλίμακας υπολογιστικά περιβάλλοντα. Ωστόσο, περιορίζει την παγκόσμια χωρική αντίληψη του μοντέλου, αφού κάθε batch δείχνει μόνο τμήματα του εγκεφάλου. Αυτό μπορεί να οδηγήσει σε ασυνέχειες μεταξύ διαδοχικών τομών ή σε αστοχίες στις άκρες των patches. Επιπλέον, η ανασύνθεση των προβλέψεων σε επίπεδο όγκου απαιτεί stitching και blending τεχνικές, που εισάγουν πρόσθετη πολυπλοκότητα. Σε περιβάλλοντα με περισσότερη μνήμη, η εκπαίδευση σε πλήρεις όγκους ή σε μεγαλύτερα patches θα μπορούσε να βελτιώσει την συνοχή των αποτελεσμάτων.
- **Μέγεθος δείγματος.** Το σύνολο δεδομένων που χρησιμοποιήθηκε ήταν περιορισμένο σε σχέση με μεγάλα benchmarks όπως το BraTS [11, 38]. Αυτό επηρεάζει άμεσα την ικανότητα γενίκευσης των μοντέλων και αυξάνει την πιθανότητα διακύμανσης μεταξύ διαφορετικών folds ή υποσυνόλων. Επιπλέον, τα μοντέλα τύπου Transformer, όπως η SwinUNETR, είναι ιδιαίτερα εξαρτημένα

από μεγάλο όγκο δεδομένων ή από προεκπαίδευση σε ευρύτερες βάσεις. Επομένως, τα αποτελέσματα της παρούσας μελέτης πρέπει να ερμηνευθούν με προσοχή: η υπεροχή ή υστέρηση μιας αρχιτεκτονικής μπορεί να διαφοροποιηθεί όταν εφαρμοστεί σε πολύ μεγαλύτερα ή πιο ετερογενή datasets.

- **Υπερπαράμετροι.** Οι ρυθμίσεις εκπαίδευσης (epochs, batch size, learning rate, dropout κ.λπ.) επιλέχθηκαν με γνώμονα τη διαθεσιμότητα μνήμης και χρόνου και όχι με πλήρη βελτιστοποίηση ανά μοντέλο. Στη βιβλιογραφία συχνά εφαρμόζονται εκτεταμένες διαδικασίες αναζήτησης υπερπαράμετρων (grid search, Bayesian optimization), κάτι που δεν ήταν εφικτό στο πλαίσιο της παρούσας διπλωματικής. Έτσι, είναι πιθανό ότι κάθε αρχιτεκτονική δεν αξιοποίησε πλήρως το δυναμικό της. Για παράδειγμα, η VNet είναι γνωστό ότι επηρεάζεται έντονα από το learning rate και το batch size, ενώ η SwinUNETR απαιτεί προσεκτική ρύθμιση του scheduler για να αποφευχθεί το underfitting. Οι επιλογές μας εξασφάλισαν συγκρισιμότητα, αλλά περιορίζουν την απόλυτη βελτιστοποίηση.
- **Μετρικές αξιολόγησης.** Η ανάλυση επικεντρώθηκε κυρίως στον Dice score, ο οποίος είναι καθιερωμένος αλλά δεν αποτυπώνει όλες τις πτυχές της απόδοσης. Μετρικές όπως το Hausdorff distance ή η sensitivity/precision ανά υποκατηγορία θα μπορούσαν να δώσουν πληρέστερη εικόνα, ιδιαίτερα για μικρές βλάβες όπου ο Dice είναι λιγότερο ευαίσθητος. Η έλλειψη πολλαπλών μετρικών αποτελεί έναν περιορισμό που αφήνει περιθώρια για πιο πολύπλευρη αξιολόγηση στο μέλλον.
- **Αναπαραγωγικότητα και κλινική εφαρμογή.** Παρά την αναπαραγωγίμη φύση του πειραματικού αγωγού, το πέρασμα από τον ερευνητικό χώρο στην κλινική πράξη δεν είναι άμεσο. Παράγοντες όπως η διαθεσιμότητα διαφορετικών MRI πρωτοκόλλων, η παρουσία artefacts ή μεταλλικών εμφυτευμάτων, και η ανάγκη για ταχεία inference σε πραγματικό χρόνο δεν καλύπτονται απόλυτα στην παρούσα μελέτη. Αυτοί οι παράγοντες θέτουν πρόσθετους περιορισμούς και απαιτούν περαιτέρω έρευνα για την κλινική μετάβαση.

Παρά τους παραπάνω περιορισμούς, η μελέτη παρείχε χρήσιμα συμπεράσματα για τα πλεονεκτήματα και τις αδυναμίες κάθε μοντέλου. Επιπλέον, η συστηματική τους καταγραφή δημιουργεί σαφείς ευκαιρίες για βελτιώσεις και επεκτάσεις στο πλαίσιο μελλοντικής εργασίας.

7.4 Μελλοντική εργασία

Με βάση τα συμπεράσματα και τους περιορισμούς που αναφέρθηκαν, διαγράφονται σαφείς κατευθύνσεις για μελλοντική έρευνα και βελτίωση των μεθοδολογιών τμηματοποίησης MRI όγκων εγκεφάλου. Η εξέλιξη του πεδίου κινείται ταυτόχρονα σε θεωρητικό, τεχνικό και κλινικό επίπεδο, επομένως οι προτάσεις που ακολουθούν καλύπτουν διαφορετικές διαστάσεις.

- **Πολυκλαδική τμηματοποίηση.** Η παρούσα μελέτη επικεντρώθηκε στη δυαδική διάκριση (όγκος vs. υπόβαθρο). Ένα φυσικό επόμενο βήμα είναι η εφαρμογή πολυκλαδικών συναρτήσεων κόστους (Dice/CE, Generalized Dice, focal loss) ώστε να διαχωρίζονται με ακρίβεια οι υποπεριοχές του όγκου (ET, ED, NCR/NET). Αυτό θα ενισχύσει την κλινική χρησιμότητα των μοντέλων, επιτρέποντας πιο στοχευμένες θεραπείες και παρακολούθηση της εξέλιξης της νόσου. Παράλληλα, θα ευθυγραμμίσει τη μελέτη με τα κριτήρια αξιολόγησης του BraTS challenge.
- **Μεγαλύτερα και πολυκεντρικά δεδομένα.** Η εκπαίδευση σε ευρύτερα και πιο ετερογενή δεδομένα είναι κρίσιμη για τη μελλοντική κλινική εφαρμογή. Η αξιοποίηση των πλήρων συνόλων

BraTS ή πολυκεντρικών συλλογών δεδομένων θα ενισχύσει τη γενίκευση και θα μειώσει το φαινόμενο *domain shift*, που σήμερα αποτελεί βασικό εμπόδιο. Η ένταξη δεδομένων από διαφορετικούς σαρωτές, πρωτόκολλα και πληθυσμούς θα δοκιμάσει τη ρωμαλεότητα των μοντέλων σε ρεαλιστικές συνθήκες.

- **Εξελιγμένες συναρτήσεις κόστους.** Οι Dice/CE losses παρείχαν αξιόπιστη βάση, αλλά εναλλακτικές συναρτήσεις μπορούν να βελτιώσουν περαιτέρω την ακρίβεια οριοθέτησης. Παραδείγματα είναι οι boundary-aware ή topology-preserving losses, που δίνουν έμφαση στην ακριβή αναπαράσταση των ορίων, στοιχείο κλινικά σημαντικό σε περιπτώσεις όπου ακόμη και μικρά σφάλματα μπορούν να επηρεάσουν το θεραπευτικό πλάνο. Η μελέτη τέτοιων συναρτήσεων μπορεί να συνεισφέρει ουσιαστικά σε πιο «λεπτομερείς» τμηματοποιήσεις.
- **Transfer learning & fine-tuning.** Η χρήση προεκπαιδευμένων encoders σε μεγάλα σύνολα δεδομένων, ιατρικών ή και γενικών εικόνων, μπορεί να βελτιώσει την απόδοση ειδικά όταν το διαθέσιμο dataset είναι περιορισμένο. Η μεταφορά γνώσης μέσω fine-tuning μπορεί να βοηθήσει την προσαρμογή σε εξειδικευμένες κλινικές εφαρμογές (π.χ. παιδιατρικοί ασθενείς ή σπάνιοι όγκοι). Η μελέτη αυτή θα επεκτείνει τη γενικότητα των μοντέλων και θα μειώσει την ανάγκη για εκτενή εκπαίδευση από το μηδέν.
- **Πλούσιες μετρικές αξιολόγησης.** Η ενσωμάτωση πρόσθετων μετρικών, όπως Hausdorff95, surface Dice, sensitivity/specificity ή volumetric overlap error, θα δώσει μια πιο ολοκληρωμένη εικόνα της απόδοσης. Ο Dice score είναι σημαντικός, αλλά η συμπληρωματική αξιολόγηση θα αναδείξει διαφορετικές πτυχές (π.χ. ποιότητα ορίων, ισορροπία false positives/false negatives, κλινικά αποδεκτά σφάλματα). Έτσι, η αξιολόγηση θα γίνει πιο αντιπροσωπευτική των πραγματικών κλινικών αναγκών.
- **Υπολογιστική αποδοτικότητα.** Για να καταστούν οι μέθοδοι εφαρμόσιμες σε πραγματικό χρόνο ή σε περιβάλλοντα με περιορισμένους πόρους, απαιτείται περαιτέρω βελτιστοποίηση. Τεχνικές όπως quantization, pruning ή knowledge distillation μπορούν να μειώσουν τον χρόνο inference και το αποτύπωμα μνήμης, χωρίς σημαντική απώλεια ακρίβειας. Η έρευνα σε αυτόν τον τομέα είναι απαραίτητη για τη μετάβαση από το εργαστήριο στην κλινική πράξη.
- **Συνδυασμός πολυτροπικών δεδομένων.** Πέρα από MRI, μελλοντικές μελέτες μπορούν να ενσωματώσουν πολυτροπικά δεδομένα, όπως PET, CT ή ακόμα και γονιδιωματικές/παθολογοανατομικές πληροφορίες. Ο συνδυασμός αυτός μπορεί να προσφέρει πλουσιότερη αναπαράσταση του όγκου και να ενισχύσει την ακρίβεια της τμηματοποίησης, ιδιαίτερα σε δύσκολες περιπτώσεις όπου οι εικόνες MRI δεν επαρκούν.
- **Ερμηνευσιμότητα και αξιοπιστία.** Ένα ακόμη πεδίο με ιδιαίτερο ενδιαφέρον είναι η ανάπτυξη μοντέλων που δεν είναι μόνο ακριβή, αλλά και ερμηνεύσιμα. Τεχνικές οπτικοποίησης attention maps ή saliency maps μπορούν να δείξουν ποια χαρακτηριστικά χρησιμοποιεί το δίκτυο για να καταλήξει στην απόφασή του. Αυτό είναι κρίσιμο για την εμπιστοσύνη των κλινικών χρηστών και για την ενσωμάτωση των μεθόδων σε πραγματικά workflows.

Τελικό σχόλιο. Η εργασία αυτή ανέδειξε τόσο τις δυνατότητες όσο και τους περιορισμούς διαφορετικών αρχιτεκτονικών στη βιοϊατρική τμηματοποίηση MRI. Παρά τους περιορισμούς του πειραματικού

setup, τα αποτελέσματα συνάδουν με τη διεθνή βιβλιογραφία και προσφέρουν ένα σαφές σημείο αναφοράς για μελλοντικές βελτιώσεις. Η συνέχιση της έρευνας προς πιο εκτεταμένα δεδομένα, πιο εξελιγμένα μοντέλα, αποδοτικότερες υλοποιήσεις αλλά και μεγαλύτερη έμφαση στην ερμηνευσιμότητα μπορεί να συμβάλει ουσιαστικά στην ενσωμάτωση αυτών των μεθόδων στην κλινική πράξη, προς όφελος των ασθενών και της υγειονομικής περίθαλψης.

Βιβλιογραφία

- [1] Douglas Hanahan και Robert A. Weinberg. «Hallmarks of Cancer: The Next Generation». Στο: *Cell* 144.5 (2011), σσ. 646–674. DOI: 10.1016/j.cell.2011.02.013.
- [2] Robert A. Weinberg. *The Biology of Cancer*. 2η έκδοση. Garland Science, 2013. ISBN: 9780815342205.
- [3] Hyuna Sung κ.ά. «Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries». Στο: *CA: A Cancer Journal for Clinicians* 71.3 (2021), σσ. 209–249. DOI: 10.3322/caac.21660.
- [4] Quinn T. Ostrom κ.ά. «CBTRUS Statistical Report: Primary Brain and Other Central Nervous System Tumors Diagnosed in the United States in 2012–2016». Στο: *Neuro-Oncology* 21.Supplement_5 (2019), σσ. v1–v100. DOI: 10.1093/neuonc/noz150.
- [5] David N. Louis, Arie Perry, Pieter Wesseling κ.ά. «The 2021 WHO Classification of Tumors of the Central Nervous System: a summary». Στο: *Neuro-Oncology* 23.8 (2021), σσ. 1231–1251. DOI: 10.1093/neuonc/noab106.
- [6] Bert Vogelstein κ.ά. «Cancer Genome Landscapes». Στο: *Science* 339.6127 (2013), σσ. 1546–1558. DOI: 10.1126/science.1235122.
- [7] Steven A. Narod και William D. Foulkes. «BRCA1 and BRCA2: 1994 and beyond». Στο: *Nature Reviews Cancer* 4.9 (2004), σσ. 665–676. DOI: 10.1038/nrc1436.
- [8] Margaret Wrensch κ.ά. «Epidemiology of primary brain tumors: current concepts and review of the literature». Στο: *Neuro-Oncology* 4.4 (2002), σσ. 278–299. DOI: 10.1093/neuonc/4.4.278.
- [9] Peter J. Jukich κ.ά. «Trends in Incidence of Primary Brain Tumors in the United States, 1985–1994». Στο: *Neuro-Oncology* 3.3 (2001), σσ. 141–151. DOI: 10.1215/15228517-3-3-141.
- [10] Mark E. Ladd κ.ά. «Pros and cons of ultra-high-field MRI/MRS for human application». Στο: *Progress in Nuclear Magnetic Resonance Spectroscopy* 109 (2018), σσ. 1–50. DOI: 10.1016/j.pnmrs.2018.06.001.
- [11] Bjoern H. Menze, Andras Jakab, Stefan Bauer κ.ά. «The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS)». Στο: *IEEE Transactions on Medical Imaging* 34.10 (2015), σσ. 1993–2024. DOI: 10.1109/TMI.2014.2377694.
- [12] Roger Stupp, Warren P. Mason, Martin J. van den Bent κ.ά. «Radiotherapy plus Concomitant and Adjuvant Temozolomide for Glioblastoma». Στο: *New England Journal of Medicine* 352.10 (2005), σσ. 987–996. DOI: 10.1056/NEJMoa043330.
- [13] Thomas Bortfeld. «IMRT: a review and preview». Στο: *Physics in Medicine and Biology* 51.13 (2006), R363–R379. DOI: 10.1088/0031-9155/51/13/R21.
- [14] Unknown Banerji. *Infinite Monkey Theorem (illustrative reference)*. Placeholder entry — replace with correct source/key if desired. 1960.
- [15] James A. Anderson. *An Introduction to Neural Networks*. MIT Press, 1995. ISBN: 9780262011447.
- [16] Stanford University. *Neural Networks — Biological background (lecture notes)*. Placeholder for a Stanford biology/CS lecture resource. 2019.
- [17] Donald O. Hebb. *The Organization of Behavior: A Neuropsychological Theory*. Wiley, 1949.

- [18] Wikipedia contributors. *Neuron — Wikimedia Commons image*. Illustrative figure; replace with exact image and license if needed. 2020. URL: <https://commons.wikimedia.org/wiki/Neuron> (επίσκεψη 07/09/2025).
- [19] Simon Haykin. *Neural Networks and Learning Machines*. 3η έκδοση. Pearson, 2009. ISBN: 9780131471399.
- [20] Matthew W. Gardner και S. R. Dorling. «Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences». Στο: *Atmospheric Environment* 32.14–15 (1998), σσ. 2627–2636. DOI: 10.1016/S1352-2310(97)00447-0.
- [21] Yann LeCun κ.ά. «Gradient-based learning applied to document recognition». Στο: *Proceedings of the IEEE* 86.11 (1998), σσ. 2278–2324. DOI: 10.1109/5.726791.
- [22] TensorFlow Datasets. *MNIST Dataset Card*. <https://www.tensorflow.org/datasets/catalog/mnist>. Dataset metadata page; accessed for dataset details. 2024.
- [23] Léon Bottou. «Large-Scale Machine Learning with Stochastic Gradient Descent». Στο: *Proceedings of COMPSTAT*. Springer, 2010, σσ. 177–186. DOI: 10.1007/978-3-7908-2604-3_16.
- [24] Michael A. Nielsen. *Neural Networks and Deep Learning*. <http://neuralnetworksanddeeplearning.com/>. Determination Press, 2015.
- [25] Vincent Dumoulin και Francesco Visin. *A Guide to Convolution Arithmetic for Deep Learning*. <https://arxiv.org/abs/1603.07285>. Figures from https://github.com/vdumoulin/conv_arithmetic (MIT License). 2016.
- [26] Aphex34. *Max pooling*. Wikimedia Commons, CC BY-SA 4.0. 2015. URL: https://commons.wikimedia.org/wiki/File:Max_pooling.png (επίσκεψη 07/09/2025).
- [27] Mehrdad Yazdani. *Example architecture of U-Net for producing k 256-by-256 image masks for a 256-by-256 RGB image*. Wikimedia Commons, CC BY-SA 4.0. 2019. URL: https://commons.wikimedia.org/wiki/File:Example_architecture_of_U-Net_for_producing_k_256-by-256_image_masks_for_a_256-by-256_RGB_image.png (επίσκεψη 07/09/2025).
- [28] Fausto Milletari, Nassir Navab και Seyed-Ahmad Ahmadi. «V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation». Στο: *arXiv preprint* (2016). eprint: 1606.04797. URL: <https://arxiv.org/abs/1606.04797>.
- [29] NVIDIA Deep Learning Examples. *V-Net Medical for TensorFlow*. Image images/vnet.png used under Apache License 2.0. 2019. URL: <https://github.com/NVIDIA/DeepLearningExamples/tree/master/TensorFlow/Segmentation/VNet> (επίσκεψη 07/09/2025).
- [30] Juhwan Lee, Justin N. Kim, Luís A. P. Dallon κ.ά. «Deep learning segmentation of fibrous cap in intravascular optical coherence tomography images». Στο: *Scientific Reports* 14.4393 (2024). Open Access, Creative Commons Attribution 4.0. DOI: 10.1038/s41598-024-55120-7. URL: <https://www.nature.com/articles/s41598-024-55120-7.pdf>.
- [31] Project MONAI. *SwinUNETR architecture image*. Apache License 2.0. 2022. URL: <https://github.com/Project-MONAI/research-contributions/tree/main/SwinUNETR> (επίσκεψη 07/09/2025).
- [32] Kaiming He κ.ά. «Deep Residual Learning for Image Recognition». Στο: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016, σσ. 770–778. DOI: 10.1109/CVPR.2016.90.

- [33] Qi Dou κ.ά. «3D deeply supervised network for automatic liver segmentation from CT volumes». Στο: *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Springer, 2017, σσ. 149–157. DOI: 10.1007/978-3-319-66179-7_18.
- [34] Olaf Ronneberger, Philipp Fischer και Thomas Brox. «U-Net: Convolutional Networks for Biomedical Image Segmentation». Στο: *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Τόμ. 9351. Lecture Notes in Computer Science. Springer, 2015, σσ. 234–241. DOI: 10.1007/978-3-319-24574-4_28. URL: <https://arxiv.org/abs/1505.04597>.
- [35] Fabian Isensee κ.ά. «nnU-Net: a self-adapting framework for U-Net-based medical image segmentation». Στο: *Nature Methods* 18 (2021), σσ. 203–211. DOI: 10.1038/s41592-020-01008-z.
- [36] Andriy Myronenko. «3D MRI Brain Tumor Segmentation Using Autoencoder Regularization». Στο: *International MICCAI Brainlesion Workshop*. Τόμ. 11384. Lecture Notes in Computer Science. Springer, 2018, σσ. 311–320. DOI: 10.1007/978-3-030-11726-9_28. URL: <https://arxiv.org/abs/1810.11654>.
- [37] Ali Hatamizadeh κ.ά. «Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images». Στο: *arXiv preprint* (2022). eprint: 2201.01266. URL: <https://arxiv.org/abs/2201.01266>.
- [38] Ujjwal Baid κ.ά. *The RSNA-ASNR-MICCAI BraTS 2020: Benchmark on Brain Tumor Segmentation and Radiomic Feature Prediction*. arXiv:2107.02314. Accessed for challenge overview and 2020 dataset summary. 2021.