



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ

Τμηματοποίηση Πλακούντα σε MRI με τεχνικές Βαθιάς Μάθησης

Μελέτη και υλοποίηση

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

ΓΡΗΓΟΡΙΟΥ ΤΣΕΝΟΥ

Επιβλέπων: Γεώργιος Ματσόπουλος
Καθηγητής, Ε.Μ.Π.

Αθήνα, Φεβρουάριος 2026



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ

Τμηματοποίηση Πλακούντα σε MRI με τεχνικές Βαθιάς Μάθησης

Μελέτη και υλοποίηση

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

ΓΡΗΓΟΡΙΟΥ ΤΣΕΝΟΥ

Επιβλέπων: Γεώργιος Ματσόπουλος
Καθηγητής, Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 22α Φεβρουαρίου 2026.

(Υπογραφή)

(Υπογραφή)

(Υπογραφή)

.....
Γεώργιος Ματσόπουλος
Καθηγητής, Ε.Μ.Π.

.....
Παναγιώτης Τσανάκας
Επίκουρος Καθηγητής

.....
Γεώργιος Γεωργίου

Αθήνα, Φεβρουάριος 2026



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ

Copyright © - All rights reserved. Με την επιφύλαξη παντός δικαιώματος.
Γρηγόριος Τσένος, 2026.

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

Το περιεχόμενο αυτής της εργασίας δεν απηχεί απαραίτητα τις απόψεις του Τμήματος, του Επιβλέποντα, ή της επιτροπής που την ενέκρινε.

ΔΗΛΩΣΗ ΜΗ ΛΟΓΟΚΛΟΠΗΣ ΚΑΙ ΑΝΑΛΗΨΗΣ ΠΡΟΣΩΠΙΚΗΣ ΕΥΘΥΝΗΣ

Με πλήρη επίγνωση των συνεπειών του νόμου περί πνευματικών δικαιωμάτων, δηλώνω ενυπογράφως ότι είμαι αποκλειστικός συγγραφέας της παρούσας Πτυχιακής Εργασίας, για την ολοκλήρωση της οποίας κάθε βοήθεια είναι πλήρως αναγνωρισμένη και αναφέρεται λεπτομερώς στην εργασία αυτή. Έχω αναφέρει πλήρως και με σαφείς αναφορές, όλες τις πηγές χρήσης δεδομένων, απόψεων, θέσεων και προτάσεων, ιδεών και λεκτικών αναφορών, είτε κατά κυριολεξία είτε βάσει επιστημονικής παράφρασης. Αναλαμβάνω την προσωπική και ατομική ευθύνη ότι σε περίπτωση αποτυχίας στην υλοποίηση των ανωτέρω δηλωθέντων στοιχείων, είμαι υπόλογος έναντι λογοκλοπής, γεγονός που σημαίνει αποτυχία στην Πτυχιακή μου Εργασία και κατά συνέπεια αποτυχία απόκτησης του Τίτλου Σπουδών, πέραν των λοιπών συνεπειών του νόμου περί πνευματικών δικαιωμάτων. Δηλώνω, συνεπώς, ότι αυτή η Πτυχιακή Εργασία προετοιμάστηκε και ολοκληρώθηκε από εμένα προσωπικά και αποκλειστικά και ότι, αναλαμβάνω πλήρως όλες τις συνέπειες του νόμου στην περίπτωση κατά την οποία αποδειχθεί, διαχρονικά, ότι η εργασία αυτή ή τμήμα της δεν μου ανήκει διότι είναι προϊόν λογοκλοπής άλλης πνευματικής ιδιοκτησίας.

(Υπογραφή)

.....
Γρηγόριος Τσένος

22 Φεβρουαρίου 2026

Περίληψη

Η αυτόματη τμηματοποίηση πλακούντα σε MRI αποτελεί κρίσιμο βήμα για ποσοτική ανάλυση της κύησης, καθώς μπορεί να υποστηρίξει πιο αντικειμενικές μετρήσεις και συγκρίσεις μεταξύ περιστατικών. Το πρόβλημα παραμένει απαιτητικό λόγω μεταβλητότητας στη μορφολογία, στην ένταση και στα όρια του οργάνου, καθώς και λόγω της σχετικά περιορισμένης διαθεσιμότητας επισημασμένων ιατρικών δεδομένων.

Αντικείμενο της παρούσας διπλωματικής είναι η συστηματική συγκριτική αξιολόγηση σύγχρονων αρχιτεκτονικών βαθιάς μάθησης για 3D τμηματοποίηση πλακούντα. Υλοποιήθηκε ενιαίο και αναπαραγώγιμο πειραματικό πλαίσιο σε MONAI/PyTorch, ώστε οι διαφορές στην απόδοση να αποδίδονται κυρίως στην αρχιτεκτονική και όχι σε ασυνεπείς ρυθμίσεις εκπαίδευσης. Μελετήθηκαν μοντέλα U-Net-οικογένειας (U-Net, Attention U-Net, DynUNet), Transformer-βασισμένες προσεγγίσεις (UNETR, SwinUNETR) και state-space-βασισμένα μοντέλα (SegMamba), με αξιολόγηση μέσω Dice και IoU.

Στο εξεταζόμενο validation split, τα καλύτερα αποτελέσματα επιτεύχθηκαν από το Seg-Mamba Heavier (Dice=0.8606, IoU=0.7566), πολύ κοντά στο SegResNet Heavier (Dice=0.8601). Το SwinUNETR ήταν το πιο ανταγωνιστικό Transformer-βασισμένο μοντέλο (Dice=0.849), ενώ το UNETR παρουσίασε τη χαμηλότερη επίδοση (Dice=0.772). Η ποιοτική αξιολόγηση ευθυγραμμίζεται με τις ποσοτικές μετρικές, αναδεικνύοντας πιο συνεκτικές μάσκες και λιγότερα ψευδοθετικά στα μοντέλα κορυφής.

Συνολικά, η εργασία δείχνει ότι αποδοτικές CNN- και SSM-βασισμένες αρχιτεκτονικές αποτελούν ισχυρές επιλογές για τμηματοποίηση πλακούντα σε MRI υπό ενιαίο πρωτόκολλο εκπαίδευσης. Παράλληλα, αναδεικνύεται ανάγκη για ισχυρότερη εκτίμηση γενίκευσης μέσω ανεξάρτητων συνόλων ελέγχου και εξωτερικής επικύρωσης.

Λέξεις Κλειδιά

MRI, Πλακούντας, Τμηματοποίηση ιατρικών εικόνων, Βαθιά Μάθηση, MONAI, U-Net, SegResNet, SegMamba, SwinUNETR, UNETR, Dice, IoU

Abstract

Placenta segmentation in MRI is a key step for quantitative pregnancy analysis, since it can support more objective measurements and comparisons across cases. The task remains challenging due to variability in morphology, intensity, and organ boundaries, as well as the limited availability of annotated medical data.

This thesis performs a systematic comparative evaluation of modern deep learning architectures for 3D placenta segmentation. A unified and reproducible MONAI/PyTorch framework was implemented, so performance differences can be attributed primarily to architecture rather than inconsistent training settings. The study includes U-Net-family models (U-Net, Attention U-Net, DynUNet), Transformer-based approaches (UNETR, SwinUNETR), and state-space-based models (SegMamba), evaluated with Dice and IoU.

On the evaluated validation split, the top result was achieved by SegMamba Heavier (Dice=0.8606, IoU=0.7566), closely followed by SegResNet Heavier (Dice=0.8601). SwinUNETR was the most competitive Transformer-based model (Dice=0.849), whereas UNETR showed the lowest performance (Dice=0.772). Qualitative inspection is consistent with the quantitative metrics, showing cleaner masks and fewer false positives for the top-performing models.

Overall, the results indicate that efficient CNN- and SSM-based architectures are strong choices for placenta segmentation in MRI under a common training protocol. At the same time, stronger generalization assessment through independent test sets and external validation remains an important next step.

Keywords

Placenta, MRI, Medical image segmentation, Deep Learning, MONAI, U-Net, SegResNet, SegMamba, SwinUNETR, UNETR, Dice, IoU

στον Πατέρα μου και στην Μητέρα μου

Ευχαριστίες

Θα ήθελα καταρχήν να ευχαριστήσω τον καθηγητή κ. Ματσόπουλο Γεώργιο για την επίβλεψη αυτής της διπλωματικής εργασίας και για την ευκαιρία που μου έδωσε να ερευνήσω ένα πεδίο με τόσο μεγάλο ερευνητικό ενδιαφέρον. Κάτα την εκπόνηση της εργασίας έλαβα πολύτιμη καθοδήγηση και σχόλια από τα μέλη του εργαστηρίου του κου Ματσόπουλου, και ειδικότερα από τον Υπ.Διδάκτορα Γεωργά Κωνσταντίνο και τον Διδάκτορα Μπρώμη Κωνσταντίνο. Τέλος θα ήθελα να ευχαριστήσω τους γονείς μου για την καθοδήγηση και την ηθική συμπαράσταση που μου προσέφεραν όλα αυτά τα χρόνια.

Αθήνα, Φεβρουάριος 2026

Γρηγόριος Τσένος

Περιεχόμενα

Περίληψη	1
Abstract	3
Ευχαριστίες	7
Πρόλογος	19
1 Εισαγωγή	21
1.1 Συνοπτική Εισαγωγή	21
1.2 Αντικείμενο της διπλωματικής	22
1.3 Οργάνωση του τόμου	22
I Θεωρητικό Μέρος	25
2 Τμηματοποίηση Ιατρικών Εικόνων και Μαγνητική Τομογραφία	27
2.1 Ο πλακούντας: ανατομία, ανάπτυξη και λειτουργίες	27
2.2 Η τμηματοποίηση ως υπολογιστικό πρόβλημα	28
2.3 Ιδιαιτερότητες τμηματοποίησης πλακούντα σε MRI	29
2.4 Μαγνητική Τομογραφία MRI	29
2.4.1 Ανομοιογένεια έντασης και διόρθωση bias field	30
2.5 Συνήθη βήματα προεπεξεργασίας για τμηματοποίηση MRI	30
2.6 Μετρικές αξιολόγησης τμηματοποίησης	31
2.6.1 Συντελεστής Dice (DSC)	31
2.6.2 Intersection over Union (IoU/Jaccard)	31
2.7 Σύνοψη	31
3 Μηχανική Μάθηση και Σύγχρονα Μοντέλα για Τμηματοποίηση	33
3.1 Επιβλεπόμενη μάθηση για τμηματοποίηση	33
3.2 Συνελικτικά Νευρωνικά Δίκτυα (CNNs)	33
3.2.1 Η συνέλιξη σε 2D/3D δεδομένα και η έννοια του receptive field	34
3.2.2 Αρχιτεκτονικές πυκνής πρόβλεψης: FCN, encoder-decoder και U-shaped σχεδίαση	34
3.2.3 Κανονικοποίηση και σταθεροποίηση εκπαίδευσης	34
3.2.4 Προσοχή και CNNs: Attention U-Net	35
3.2.5 Πλεονεκτήματα και περιορισμοί των CNNs	35

3.3 Μοντέλα Attention και Transformers	35
3.3.1 Βασικός μηχανισμός self-attention	35
3.3.2 Self-Attention και Εικόνες: προκλήσεις και λύσεις	36
3.3.3 Μείωση κόστους: ιεραρχία, τοπικά παράθυρα και προσεγγίσεις γραμμικής προσοχής	36
3.3.4 Υθριδικά Transformer-U-shape μοντέλα	36
3.4 Μοντέλα Χώρου Καταστάσεων (State Space Models) και Mamba	36
3.4.1 Δομημένα Μοντέλα Χώρου Καταστάσεων (Structured SSMs) και Μοντελοποίηση Μακρινών Εξαρτήσεων	37
3.4.2 Mamba: Επιλεκτικά Μοντέλα Χώρου Καταστάσεων και Γραμμική Κλιμάκωση	37
3.4.3 Εφαρμογή SSMs/Mamba σε 3D Τμηματοποίηση	38
3.4.4 Συγκριτική Επισκόπηση	38
3.5 Βελτιστοποίηση και πρακτικές εκπαίδευσης	38
4 Αρχιτεκτονικές Τμηματοποίησης που Χρησιμοποιήθηκαν	41
4.1 Το πλαίσιο MONAI	41
4.1.1 Υλοποίηση μοντέλων στην παρούσα εργασία	41
4.2 Οικογένειες μοντέλων που συγκρίνονται	42
4.3 UNet και 3D U-Net	42
4.4 DynUNet	43
4.5 SegResNet: υπολειμματικό encoder-decoder	43
4.6 AttentionUnet: προσοχή στις skip connections	43
4.7 UNETR: ViT encoder με U-shaped decoder	44
4.8 SwinUNETR: ιεραρχικός Swin encoder	44
4.9 SegMamba: Mamba/SSM blocks σε U-shape σχεδίαση	45
4.10 Συζήτηση: αναμενόμενοι συμβιβασμοί	45
II Πρακτικό Μέρος	47
5 Πειραματικό Πρωτόκολλο και Μεθοδολογία Αξιολόγησης	49
5.1 Στόχοι αξιολόγησης και αρχές σύγκρισης	49
5.2 Υλοποίηση και λογισμικό	50
5.3 Σύνολο δεδομένων και διαχωρισμός	50
5.4 Οργάνωση δεδομένων και ροή φόρτωσης	50
5.5 Ανάλυση στοιβας προεπεξεργασίας	51
5.5.1 Γεωμετρική εναρμόνιση και σταθεροποίηση εισόδου	51
5.5.2 Κανονικοποίηση εντάσεων	51
5.5.3 Περιορισμός πεδίου (ROI) με foreground cropping	51
5.5.4 Σταθεροποίηση διαστάσεων για συμβατότητα μοντέλων	51
5.6 Δειγματοληψία patches και επαύξηση δεδομένων	52
5.7 Αρχιτεκτονικές και πειραματικά σενάρια	52
5.7.1 Αρχιτεκτονικές υπερπαράμετροι ανά μοντέλο	52

5.8 Συνάρτηση κόστους και βελτιστοποίηση	54
5.8.1 Κριτήριο εκπαίδευσης	54
5.9 Επικύρωση και διαδικασία inference	54
5.9.1 Sliding Window Inference (SWI)	54
5.9.2 EMA (Exponential Moving Average)	54
5.10 Βελτιστοποίηση και Scheduler	54
5.10.1 Optimizer	54
5.10.2 Scheduler	54
6 Αποτελέσματα Πειραματικής Αξιολόγησης	57
6.1 Παρουσίαση αποτελεσμάτων	57
6.2 Συνοπτική κατάταξη και κύρια ευρήματα	58
6.3 Καμπύλες εκπαίδευσης και σταθερότητα σύγκλισης	59
6.4 Ποιοτική αξιολόγηση προβλέψεων	61
6.5 Σύγκριση με δημοσιευμένα αποτελέσματα	66
6.6 Συζήτηση και ερμηνεία	66
6.6.1 Ενδεικτικοί χρόνοι ολοκλήρωσης εκτέλεσης notebooks	68
III Επίλογος	71
7 Επίλογος	73
7.1 Συμπεράσματα	73
7.2 Μελλοντικές επεκτάσεις	74
Βιβλιογραφία	81

Κατάλογος Σχημάτων

2.1	Αριστερά : ενδεικτική δισδιάστατη τομή (slice) MRI με την επισήμανση (πράσινο) του πλακούντα. Δεξιά : τρισδιάστατη αναπαράσταση της αντίστοιχης μάσκας τμηματοποίησης (πράσινο) και της θέσης της τομής στον όγκο.	30
4.1	Δομή αρχιτεκτονικής U-net	42
4.2	Μηχανισμός residual connection στην αρχιτεκτονική ResNet [1]	43
4.3	Προτεινόμενος σχεδιασμός του attention gate	44
4.4	Αριστερά : Swin με ιεραρχικά stages (πολλαπλές κλίμακες μέσω patch merging). Δεξιά : ViT χωρίς ιεραρχία, με σταθερή ανάλυση tokens σε όλο τον encoder.	44
6.1	Καμπύλες εκπαίδευσης: Validation Dice (πάνω) και training/validation loss (κάτω) για Attention U-Net, DynUNet, SwinUNETR Heavier και SwinUNETR Lighter. Το αστέρι και η κατακόρυφη διακεκομμένη γραμμή δείχνουν το epoch του καλύτερου Dice.	59
6.2	Καμπύλες εκπαίδευσης: Validation Dice (πάνω) και training/validation loss (κάτω) για SegResNet Lighter, SegResNet Heavier και U-Net. Το αστέρι και η κατακόρυφη διακεκομμένη γραμμή δείχνουν το epoch του καλύτερου Dice.	60
6.3	Καμπύλες εκπαίδευσης: Validation Dice (πάνω) και training/validation loss (κάτω) για UNETR, SegMamba Heavier και SegMamba Lighter. Το αστέρι και η κατακόρυφη διακεκομμένη γραμμή δείχνουν το epoch του καλύτερου Dice.	60

Κατάλογος Εικόνων

6.1	Attention U-Net	62
6.2	DynUNet	63
6.3	SwinUNETR Heavier	63
6.4	SwinUNETR Lighter	63
6.5	SegResNet Heavier (No of Filters = 64)	64
6.6	SegResNet Lighter (No of Filters = 32)	64
6.7	U-Net	64
6.8	UNETR	65
6.9	SegMamba Heavier (5 layers)	65
6.10	SegMamba Lighter (4 layers)	65

Κατάλογος Πινάκων

6.1	Αποτελέσματα ανά αρχιτεκτονική στο validation set	58
6.2	Σύνοψη αποτελεσμάτων από σχετικές μελέτες τμηματοποίησης πλακούντα σε MRI.	66
6.3	Ενδεικτικοί χρόνοι ολοκλήρωσης ανά run.	68

Πρόλογος

Το υπολογιστικό μέρος της εργασίας υλοποιήθηκε και εκτελέστηκε στη πλατφόρμα Kaggle (Kaggle Notebooks). Το θεωρητικό υπόβαθρο βασίστηκε σε μελέτη της σχετικής βιβλιογραφίας.

Κεφάλαιο 1

Εισαγωγή

Ηιατρική απεικόνιση αποτελεί αναπόσπαστο μέρος της σύγχρονης κλινικής πρακτικής και της βιοϊατρικής έρευνας, καθώς επιτρέπει τη μη επεμβατική παρατήρηση ανατομικών δομών και φυσιολογικών διεργασιών. Πέρα όμως από την οπτική επιθεβαίωση ή περιγραφική αξιολόγηση, σε πολλές εφαρμογές είναι απαραίτητη η ποσοτικοποίηση: η εξαγωγή μετρήσιμων μεγεθών (π.χ. όγκοι, σχήματα, σχέσεις γειτονικών ιστών), που στη συνέχεια μπορούν να συσχετισθούν με κλινικά ευρήματα ή να αξιοποιηθούν σε συστήματα υποβοήθησης διάγνωσης.

1.1 Συνοπτική Εισαγωγή

Κεντρικό βήμα για την ποσοτικοποίηση αποτελεί η **τμηματοποίηση** (segmentation) μιας ιατρικής εικόνας ή ενός ογκομετρικού δεδομένου. Με τον όρο τμηματοποίηση εννοούμε την ανάθεση εικέτας σε κάθε (ογκομετρικό) εικονοστοιχείο (pixel/voxel), ώστε να διαχωριστεί η περιοχή ενδιαφέροντος (Region of Interest, ROI) από το υπόλοιπο υπόβαθρο. Στην πράξη, η τμηματοποίηση επιστρέφει μία **μάσκα** (mask) με την ίδια διάσταση όπως η είσοδος, η οποία κωδικοποιεί το ‘ανήκει/δεν ανήκει’ (δυαδική τμηματοποίηση) ή/και πολλαπλές κλάσεις (πολυκλασική τμηματοποίηση). Η τμηματοποίηση είναι η βάση για τον υπολογισμό όγκου, την εκτίμηση μορφολογικών χαρακτηριστικών, την ανάλυση χωρικής επερογένειας, αλλά και για πιο σύνθετες ροές εργασίας, όπως εξαγωγή χαρακτηριστικών (radiomics) ή καθοδήγηση επεμβάσεων.

Στο πλαίσιο της μαιευτικής απεικόνισης, ιδιαίτερο ενδιαφέρον παρουσιάζει ο **πλακούντας**, ένα λειτουργικά κρίσιμο όργανο που συνδέει μητέρα και έμβρυο. Η **αυτόματη τμηματοποίηση πλακούντα** σε δεδομένα MRI επιτρέπει επαναλήψιμες και αντικειμενικές μετρήσεις (π.χ. όγκος/σχήμα), καθώς και τη συστηματική σύγκριση μεταξύ εξετάσεων ή πρωτοκόλλων. Ωστόσο, η τμηματοποίηση πλακούντα σε MRI είναι απαιτητικό πρόβλημα: τα δεδομένα συχνά επηρεάζονται από κίνηση μητέρας/εμβρύου, ανομοιογένειες έντασης, περιορισμένη αντίθεση στα όρια οργάνου-υποβάθρου και μεγάλη διακύμανση σε μέγεθος, σχήμα και θέση. Προγενέστερες προσεγγίσεις περιλάμβαναν κλασικούς ταξινομητές ή ελάχιστα/μερικώς διαδραστικές μεθόδους, ενώ έχουν παρουσιαστεί και τρόποι για αντιμετώπιση σε περίπτωση που αλλοιωθεί η εικόνα λόγω κίνησης. [2][3][4].

Η χειροκίνητη τμηματοποίηση, αν και θεωρείται σημείο αναφοράς, είναι χρονοβόρα και εμφανίζει διακύμανση μεταξύ παρατηρητών, κάτι που επηρεάζει τόσο την αξιολόγηση όσο και

την κλινική αξιοποίηση. Για τον λόγο αυτό, τα τελευταία χρόνια κυριαρχούν μέθοδοι **Βαθιάς Μάθησης** (Deep Learning) για τμηματοποίηση ιατρικών εικόνων, με αρχιτεκτονικές τύπου UNet να αποτελούν θεμελιώδη επιλογή [5]. Παράλληλα, μοντέλα με μηχανισμούς attention και Transformers έχουν δείξει ιδιαίτερα ισχυρή ικανότητα μοντελοποίησης διευκολύνοντας την συσχέτιση μακρινών εικονοστοιχείων (global context) [6, 7, 8]. Πρόσφατα, τα **Μοντέλα Χώρου και Κατάστασης** (State Space Models, SSMs) και ειδικότερα το Mamba εισάγουν μία εναλλακτική προσέγγιση μακράς εμβέλειας, με σόχο υψηλή αποδοτικότητα σε μνήμη και χρόνο [9], και έχουν προταθεί υθριδικά μοντέλα τμηματοποίησης όπως το SegMamba [10]. (Λεπτομέρειες στα κεφάλαια 2,3,4)

1.2 Αντικείμενο της διπλωματικής

Αντικείμενο της παρούσας διπλωματικής εργασίας είναι η **συστηματική μελέτη, υλοποίηση και συγκριτική αξιολόγηση** μοντέλων τμηματοποίησης πλακούντα από ογκομετρικά δεδομένα MRI, με χρήση της βιβλιοθήκης MONAI [11]. Η εργασία επικεντρώνεται σε ένα δίκαιο πειραματικό πρωτόκολλο, όπου τα μοντέλα εκπαιδεύονται και αξιολογούνται υπό *ιδιες* συνθήκες (ίδιο seed, ίδια προεπεξεργασία, ίδιοι βελτιστοποιητές, κριτήρια, σχήματα learning rate, αριθμός εποχών και συνολικό πείραμα), ώστε οι διαφορές στην επίδοση να αποδίδονται σε μεγαλύτερο βαθμό στην αρχιτεκτονική.

Συγκεκριμένα, εξετάζονται οι παρακάτω αρχιτεκτονικές:

- **Συνελικτικά μοντέλα τύπου UNet:** UNet [5], DynUNet (δυναμική παραλλαγή σχεδιασμένη με αρχές αυτο-διαμόρφωσης), SegResNet [12], καθώς και AttentionUnet [13].
- **Υθριδικά μοντέλα με Transformers:** UNETR [7] και SwinUNETR [8].
- **Μοντέλο βασισμένο σε SSM/Mamba:** SegMamba [10].

Για την αξιολόγηση χρησιμοποιείται κυρίως η **μετρική Dice** (γνωστή και ως DSC) και συμπληρωματικά ο **Intersection over Union** (IoU/Jaccard). Στόχος δεν είναι απλώς η καταγραφή μιας μέγιστης επίδοσης, αλλά και η **εξαγωγή συμπερασμάτων** για τα πλεονεκτήματα/μειονεκτήματα κάθε οικογένειας αρχιτεκτονικών (καθαρά συνελικτικές, Transformer-ενισχυμένες, SSM-βασισμένες) στο συγκεκριμένο πρόβλημα.

1.3 Οργάνωση του τόμου

Η διάρθρωση της διπλωματικής έχει ως εξής:

- Στο **Κεφάλαιο 2** παρουσιάζονται οι βασικές έννοιες που απαιτούνται για την ποσοτική αξιολόγηση του πλακούντα σε MRI: συνοπτικά στοιχεία ανατομίας/ανάπτυξης/λειτουργιών και κλινικής σημασίας του πλακούντα, η διατύπωση της τμηματοποίησης ως υπολογιστικού προβλήματος, οι ιδιαιτερότητες της τμηματοποίησης πλακούντα σε ογκομετρικά δεδομένα MRI, βασικές αρχές της MRI, συνήθη βήματα προεπεξεργασίας, καθώς και οι μετρικές αξιολόγησης (Dice/DSC, IoU).

- Στο **Κεφάλαιο 3** παρουσιάζεται το θεωρητικό υπόβαθρο της Μηχανικής Μάθησης και της Βαθιάς Μάθησης για τμηματοποίηση: επιβλεπόμενη μάθηση και συναρτήσεις κόστους, CNNs, Transformers και SSM/Mamba, καθώς και βασικές πρακτικές εκπαίδευσης/θελτιστοποίησης (π.χ. βελτιστοποιητές, learning rate schedulers, κανονικοποίηση).
- Στο **Κεφάλαιο 4** περιγράφονται οι αρχιτεκτονικές των επτά μοντέλων που χρησιμοποιήθηκαν (UNet, DynUNet, SegResNet, AttentionUnet, UNETR, SwinUNETR, Seg-Mamba), καθώς και οι κύριες σχεδιαστικές επιλογές και οι αναμενόμενοι συμβιβασμοί τους στην 3D τμηματοποίηση ιατρικών δεδομένων.
- Στο **Κεφάλαιο 5** παρουσιάζεται το πειραματικό πρωτόκολλο: το σύνολο δεδομένων, το pipeline προεπεξεργασίας και φόρτωσης, οι υπερπαράμετροι εκπαίδευσης και οι αρχές της δίκαιης συγκριτικής αξιολόγησης μεταξύ των αρχιτεκτονικών.
- Στο **Κεφάλαιο 6** παρατίθενται τα αποτελέσματα της πειραματικής αξιολόγησης, τόσο ποσοτικά (Dice/IoU, validation loss) όσο και ποιοτικά (οπτικές συγκρίσεις προβλέψεων), μαζί με τη συζήτηση και τη σύγκριση με σχετική βιβλιογραφία.
- Στο **Κεφάλαιο 7** διατυπώνονται τα τελικά συμπεράσματα της εργασίας, οι βασικοί περιορισμοί εγκυρότητας και οι κύριες μελλοντικές επεκτάσεις της μελέτης.

Μέρος I

Θεωρητικό Μέρος

Κεφάλαιο 2

Τμηματοποίηση Ιατρικών Εικόνων και Μαγνητική Τομογραφία

Ο πλακούντας αποτελεί ένα προσωρινό αλλά κρίσιμο όργανο της κύησης, το οποίο λειτουργεί ως επιφάνεια διεπαφής μεταξύ μητέρας και εμβρύου, επιτρέποντας την ανταλλαγή αερίων, θρεπτικών συστατικών και μεταβολισμού ουσιών, ενώ παράλληλα επιτελεί σημαντικούς ενδοκρινικούς και ανοσολογικούς ρόλους [14, 15, 16]. Ο πλακούντας, επίσης, κρατά προσκολλημένο το έμβρυο με το τοίχωμα της μήτρας όπου επιτελεί και αναπνευστική επιφάνεια για το έμβρυο. Η λειτουργία του πλακούντα σχετίζεται άμεσα με την ομαλή έκβαση της κύησης, και η δυσλειτουργία του έχει συσχετιστεί με σοβαρές μαιευτικές επιπλοκές, όπως προεκλαμψία, υπολειπόμενη ανάπτυξη εμβρύου (FGR), αποβολές και ενδομήτριο θάνατο του εμβρύου.

Σε αυτό το πλαίσιο, η ιατρική απεικόνιση – και ειδικότερα η Μαγνητική Τομογραφία (MRI) – μπορεί να υποστηρίξει την ποσοτική αξιολόγηση του πλακούντα. Ωστόσο, για να εξαχθούν αξιόπιστες μετρήσεις (π.χ. όγκος, σχήμα, μορφολογικοί δείκτες ή/και χαρακτηριστικά υφής), απαιτείται πρώτα ο ακριβής διαχωρισμός του πλακούντα από το υπόβαθρο: η **τμηματοποίηση** (segmentation). Στην παρούσα εργασία, στόχος είναι η τμηματοποίηση του πλακούντα σε ογκομετρικά δεδομένα MRI, δηλαδή η παραγωγή μιας δυαδικής μάσκας (πλακούντας έναντι υποβάθρου) για κάθε εικόνα.

2.1 Ο πλακούντας: ανατομία, ανάπτυξη και λειτουργίες

Ο πλακούντας αναπτύσσεται από την τροφόβλαστη της περιφέρειας της βλαστοκύστης, όπου συμβαίνει η εμφύτευση. Η εμφύτευση λαμβάνει χώρα εππά με δώδεκα ημέρες μετά τη γονιμοποίηση [17]. Από ανατομικής άποψης, στον πλακούντα διακρίνεται:

- το **χοριακό πέταλο** (chorionic plate) (εμβρυική πλευρά), από όπου εκφύονται οι χοριακές λάχνες και το δίκτυο αγγείων του εμβρύου,
- το **φθαρτικό πέταλο** (basal plate/decidua) (μητρική πλευρά), το οποίο συνδέεται με το τοίχωμα της μήτρας. Γίνεται εμφανές αφού κατα τον τοκετό πραγματοποιείται διαχωρισμός του πλακούντα από το σώμα της μήτρας μέσω αυτού,
- ο **μεσολάχνιος χώρος** (intervillous space), όπου κυκλοφορεί μητρικό αίμα, διαχωρίζεται η εμβρυική κυκλοφορία από τις δεξαμενές του μητρικού αίματος, και πραγμα-

τοποιείται η κύρια μεταφορά οξυγόνου, θρεπτικών συστατικών και μεταβολιτών μεταξύ μητέρας και εμβρύου.

Οι χοριακές λάχνες, με υψηλή πυκνότητα σε τροφοθλαστικούς πληθυσμούς (με κυρίαρχο το syncytiotrophoblast), αποτελούν τον κύριο συντελεστή στην ανταλλαγή συστατικών μεταξύ μητέρας και εμβρύου [16, 15].

Καθοριστικό βήμα στην ορθή πλακουντοποίηση είναι η **αναδιαμόρφωση των μητρικών σπειροειδών αρτηριών**, μέσω εισβολής εξωλάχνιων τροφοθλάστων (extravillous trophoblast) στον φθαρτό της μήτρας, ώστε να εξασφαλιστεί σταθερή αιμάτωση προς τον πλακούντα. Η ελαττωματική πλακουντοποίηση έχει προταθεί ως κοινός παθοφυσιολογικός μηχανισμός πίσω από τις διάφορες επιπλοκές της κύησης (π.χ. προεκλαμψία, αποκόλληση πλακούντα, fetal growth restriction) [18, 19].

Λειτουργικά, ο πλακούντας:

- εξασφαλίζει **μεταφορά** οξυγόνου, ηλεκτρολυτών και θρεπτικών ουσιών (νερό, γλυκόζη, αμινοξέα, λιπίδια, βιταμίνες) και συμβάλλει στην **απομάκρυνση** μεταβολικών προϊόντων όπως ουρεία, ουρικό οξύ και άλλων αποβλήτων [20],
- δρα ως **ενδοκρινικό όργανο** (παραγωγή ορμονών που ρυθμίζουν την κύηση και τον μεταβολισμό της μητέρας),
- συμμετέχει στην **ανοσολογική ανοχή** και στη ρύθμιση της διεπαφής μεταξύ μητέρας και εμβρύου.

Συνεπώς, ο πλακούντας δεν είναι απλώς «φίλτρο», αλλά ένα πολυλειτουργικό όργανο με δυναμική εξέλιξη και προσαρμογή κατά την κύηση [14, 15]. Η μεταφορά, ωστόσο, δεν αφορά μόνο χρήσιμες και απαραίτητες ουσίες. Επιβλαβείς παράγοντες μπορούν να περάσουν, όπως ψυχοδραστικές ουσίες, μονοξείδιο του άνθρακα, διάφοροι λοιμώδεις παράγοντες κ.α. [20].

Η παραπάνω φυσιολογική πολυπλοκότητα εξηγεί γιατί η **ποσοτική** αξιολόγηση του πλακούντα (δομή/μορφολογία, χωρική κατανομή, όγκος) αποτελεί ενεργό ερευνητικό πεδίο. Η Μαγνητική (MRI) προσφέρει υψηλή αντίθεση μαλακών ιστών και τρισδιάστατη πληροφορία, όμως η ποσοτικοποίηση προϋποθέτει αξιόπιστο ορισμό της περιοχής του πλακούντα μέσα στον όγκο, δηλαδή ακριβής τμηματοποίηση. Η ανάγκη αυτή αποτυπώνεται και σε μελέτες που αναπτύσσουν αυτοματοποιημένους αλγορίθμους για τμηματοποίηση πλακούντα/μήτρας σε MRI όπου αναφέρεται άμεσα η σχέση της τμηματοποίησης με ορθή και συνεπή υπολογισμό του όγκου [4].

2.2 Η τμηματοποίηση ως υπολογιστικό πρόβλημα

Σε ένα τυπικό σενάριο **δυαδικής τμηματοποίησης**, η είσοδος είναι ένας όγκος $X \in \mathbb{R}^{H \times W \times D}$ και η έξοδος είναι μια μάσκα $Y \in \{0, 1\}^{H \times W \times D}$ που υποδηλώνει για κάθε voxel αν ανήκει στην κλάση ενδιαφέροντος. Η μάθηση διατυπώνεται συνήθως ως **επιβλεπόμενη** μάθηση: διαθέτουμε σύνολο ζευγών (X_i, Y_i) και επιδιώκουμε μοντέλο f_θ ώστε $f_\theta(X_i) \approx Y_i$.

Η **μάσκα τμηματοποίησης** (segmentation mask) είναι μια εικόνα/όγκος *iδιων διαστάσεων* με την είσοδο, όπου κάθε pixel/voxel παίρνει τιμή 1 (λευκό) αν ανήκει στην κλάση ενδιαφέροντος (π.χ. πλακούντας) και 0 (μαύρο) διαφορετικά.

Η τμηματοποίηση διαφέρει από την ταξινόμηση εικόνας (Image Classification) στο γεγονός ότι η έξοδος είναι **πυκνή** πρόβλεψη (ανά pixel/voxel) και όχι μία ετικέτα ανά εικόνα.

Στην 3D τμηματοποίηση, οι υπολογιστικές απαιτήσεις αυξάνονται σημαντικά σε σχέση με το πλήθος των voxels. Έτσι, συχνά χρησιμοποιούνται στρατηγικές **patch-based** εκπαίδευσης και **sliding-window inference** (στο Κεφάλαιο 3 αναλύονται οι αντίστοιχες πρακτικές), έτσι ώστε να είναι εφικτή η εκπαίδευση/πρόβλεψη σε μνήμη GPU.

2.3 Ιδιαιτερότητες τμηματοποίησης πλακούντα σε MRI

Η τμηματοποίηση πλακούντα σε MRI είναι απαιτητική για λόγους που σχετίζονται τόσο με τη φυσιολογία όσο και με τη διαδικασία λήψης:

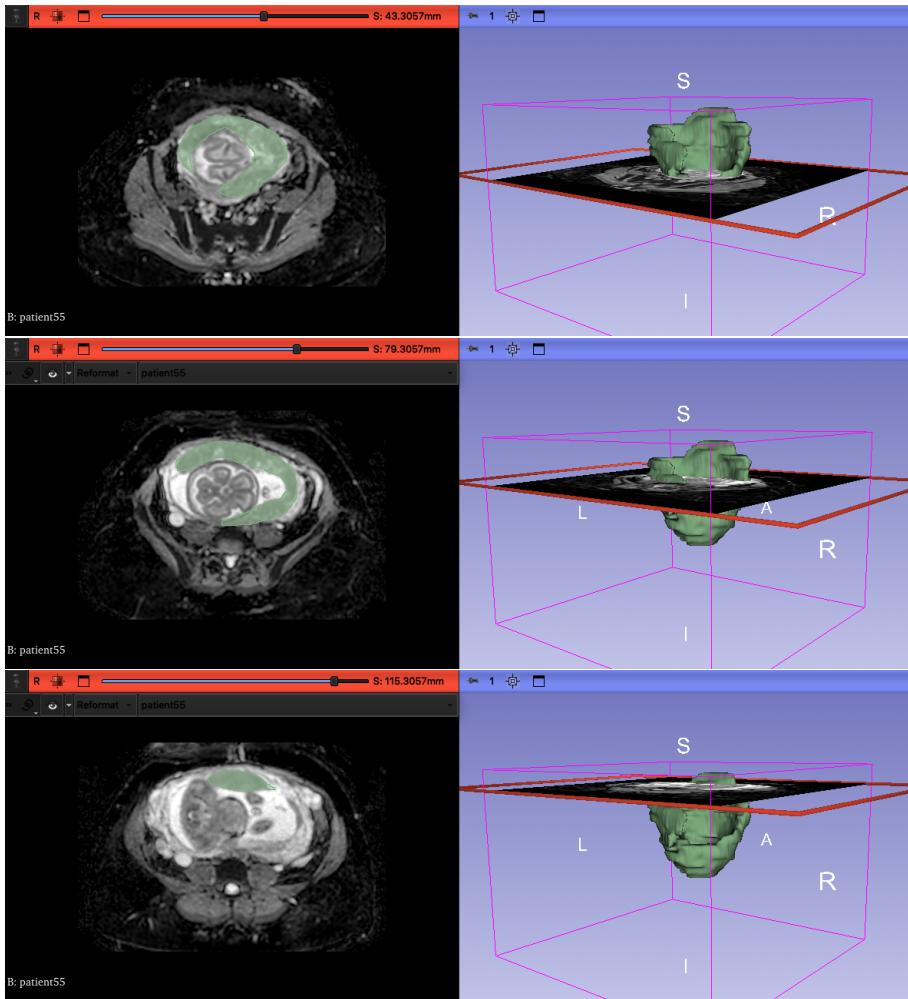
- **Μεγάλη μορφολογική μεταβλητότητα:** μέγεθος, σχήμα και θέση του πλακούντα μεταβάλλονται σημαντικά.
- **Περιορισμένη αντίθεση** σε ορισμένες ακολουθίες, ιδιαίτερα στα όρια πλακούντα και παρακείμενων ιστών.
- **Ανισορροπία κλάσεων:** στα περισσότερα slices ο πλακούντας καταλαμβάνει μικρό μέρος του όγκου, κάτι που δυσκολεύει τη μάθηση και φέρει την ανάγκη για στοχευμένη δειγματοληψία και μεθοδολογία (λεπτομερώς παρακάτω 2.5) [21, 22, 23].
- **Κίνηση** (μητέρα/εμβρύου) που οδηγεί σε artifacts, διαστρεβλώσεις και θολή εικόνα.
- **Ανομοιογένεια έντασης** (bias field) που επηρεάζει την εμφάνιση ίδιου ιστού σε διαφορετικές περιοχές.

Σχετικές εργασίες έχουν δείξει ότι ακόμη και με προσεκτική χειροκίνητη τμηματοποίηση η διαδικασία είναι χρονοβόρα και εμφανίζει ασυνέπεια ως προς διαφορετικούς παρατηρητές. Παράλληλα, έχουν παρουσιαστεί μέθοδοι αυτόματης/ημιαυτόματης τμηματοποίησης για πλακούντα σε MRI, συμπεριλαμβανομένων πλαισίων που αντιμετωπίζουν δεδομένα με θόρυβο [2, 3] καθώς και προσεγγίσεων βαθιάς μάθησης για τρισδιάστατη τμηματοποίηση πλακούντα και εκτίμηση όγκου [4, 24, 25].

2.4 Μαγνητική Τομογραφία MRI

Η Μαγνητική Τομογραφία (Magnetic Resonance Imaging, MRI) είναι τεχνική απεικόνισης που βασίζεται στο φαινόμενο του πυρηνικού μαγνητικού συντονισμού. Συνοπτικά, τα πρωτόνια (κυρίως του υδρογόνου) ευθυγραμμίζονται με το στατικό μαγνητικό πεδίο, διεγίρονται με ραδιοσυχνότητες, και η επιστροφή τους στην ισορροπία παράγει σήμα που καταγράφεται και αναδομείται σε εικόνα. Η αντίθεση μεταξύ ιστών εξαρτάται από παραμέτρους όπως οι χρόνοι χαλάρωσης T_1 και T_2 , καθώς και από τη διαμόρφωση των ακολουθιών λήψης. Αναλυτικότερη θεμελίωση παρέχεται σε κλασικά συγγράμματα [26, 27].

Για την τμηματοποίηση, η μαγνητική παρουσιάζει πλεονεκτήματα (υψηλή αντίθεση μαλακών ιστών, 3D πληροφορία), αλλά και προκλήσεις: θόρυβο, ανομοιογένειες έντασης, και



Σχήμα 2.1: Αριστερά: ενδεικτική δισδιάστατη τομή (slice) MRI με την επισήμανση (πράσινο) του πλακούντα. Δεξιά: τρισδιάστατη αναπαράσταση της αντίστοιχης μάσκας τμηματοποίησης (πράσινο) και της θέσης της τομής στον όγκο.

μεταβολές μεταξύ σαρωτών/πρωτοκόλλων. Ειδικά σε απεικόνιση κύησης, η κίνηση αποτελεί συχνό πρόβλημα, επιτρεάζοντας την ευκρίνεια και τη συνέπεια μεταξύ τομών [2].

2.4.1 Ανομοιογένεια έντασης και διόρθωση bias field

Η ανομοιογένεια έντασης (intensity non-uniformity) ή bias field είναι αργά μεταβαλλόμενη παραμόρφωση που οδηγεί σε διαφορετικές εντάσεις για τον ίδιο ιστό σε διαφορετικά σημεία του όγκου. Μία καθιερωμένη μέθοδος διόρθωσης είναι ο αλγόριθμος N4ITK [28], ο οποίος βελτιώνει προγενέστερες προσεγγίσεις και χρησιμοποιείται ευρέως ως προεπεξεργαστικό βήμα.

2.5 Συνήθη βήματα προεπεξεργασίας για τμηματοποίηση MRI

Αν και το ακριβές pipeline εξαρτάται από το πρόβλημα και το σύνολο δεδομένων, σε 3D MRI τμηματοποίηση εφαρμόζονται συχνά [29]:

- **Εναρμόνιση προσανατολισμού** (κοινός άξονας αναφοράς).

- **Επαναδειγματοληψία** (resampling) σε επιθυμητή χωρική ανάλυση (voxel spacing).
- **Κανονικοποίηση έντασης** (κλιμάκωση σε $[0, 1]$).
- **Περικοπή εικόνας** (π.χ. crop foreground) ώστε να μειωθεί το περιττό υπόβαθρο και η ανισορροπία κλάσεων.
- **Εμπλουτισμός δεδομένων** (augmentation) με τυχαίες περιστροφές/αντιστροφές/παραμορφώσεις/θόρυβο.

Στόχος των παραπάνω είναι (α) η διευκόλυνση της εκπαίδευσης με σταθερότερα στατιστικά, (β) η κανονικοποίηση των δεδομένων και (γ) η βελτίωση της γενίκευσης του μοντέλου.

2.6 Μετρικές αξιολόγησης τμηματοποίησης

Για την αξιολόγηση δυαδικής τμηματοποίησης χρησιμοποιούνται μετρικές επικάλυψης μεταξύ πρόβλεψης P και αλήθειας G . Οι πιο διαδεδομένες είναι:

2.6.1 Συντελεστής Dice (DSC)

Ο συντελεστής Dice ορίζεται ως:

$$\text{DSC}(P, G) = \frac{2|P \cap G|}{|P| + |G|}. \quad ^1$$
 (2.1)

Η τιμή του βρίσκεται στο $[0, 1]$, όπου 1 δηλώνει τέλεια επικάλυψη.

2.6.2 Intersection over Union (IoU/Jaccard)

Ο δείκτης Jaccard ορίζεται ως:

$$\text{IoU}(P, G) = \frac{|P \cap G|}{|P \cup G|}.$$
 (2.2)

Και εδώ οι τιμές είναι στο $[0, 1]$. Οι δύο μετρικές σχετίζονται μονοτονικά και συχνά αναφέρονται μαζί.

2.7 Σύνοψη

Στο κεφάλαιο αυτό παρουσιάστηκαν:

- βασικές έννοιες φυσιολογίας/ανάπτυξης του πλακούντα και η κλινική του σημασία,
- ο ρόλος της τμηματοποίησης ως υπολογιστικό πρόβλημα,
- οι ιδιαίτερες δυσκολίες της τμηματοποίησης πλακούντα σε MRI,
- βασικές αρχές της MRI που σχετίζονται με τη φύση των δεδομένων,
- και οι μετρικές αξιολόγησης που χρησιμοποιούνται στην παρούσα εργασία.

¹Ο τελεστής \cap υπολογίζει την τομή δύο συνόλων A και B που αποτελέται από τα κοινά στοιχεία τους.

Κεφάλαιο 3

Μηχανική Μάθηση και Σύγχρονα Μοντέλα για Τμηματοποίηση

H αυτόματη τμηματοποίηση με σύγχρονες προσεγγίσεις βασίζεται κυρίως σε **επιβλεπόμενη** Μηχανική Μάθηση, όπου το μοντέλο εκπαιδεύεται σε ζεύγη ιατρικών εικόνων και των αντίστοιχων **μασκών τμηματοποίησης** που έχουν δημιουργηθεί χειροκίνητα από ειδικούς [30]. Η επιτυχία της Βαθιάς Μάθησης στην τμηματοποίηση οφείλεται στην ικανότητα της να μαθαίνει πλούσιες ιεραρχικές αναπαραστάσεις (από απλά τοπικά χαρακτηριστικά έως πολύπλοκα χωρικά μοτίβα) από μεγάλα σύνολα δεδομένων, καθώς και στην αποδοτική βελτιστοποίηση με backpropagation. Για γενικότερο πλαίσιο στη Βαθιά Μάθηση, παραπέμπουμε σε κλασικές πηγές [30, 31, 32].

3.1 Επιβλεπόμενη μάθηση για τμηματοποίηση

Έστω ότι διαθέτουμε σύνολο εκπαίδευσης $\mathcal{D} = \{(X_i, Y_i)\}_{i=1}^N$, όπου X_i είναι MRI όγκος και Y_i η αντίστοιχη δυαδική μάσκα πλακούντα. Θέλουμε να βρούμε παραμέτρους θ ώστε το μοντέλο f_θ να ελαχιστοποιεί μία συνάρτηση απώλειας:

$$\hat{\theta}^* = \arg \min_{\theta} \sum_{i=1}^N \mathcal{L}(f_\theta(X_i), Y_i). \quad (3.1)$$

Στην τμηματοποίηση, η επιλογή \mathcal{L} είναι καθοριστική. Η απλή binary cross-entropy είναι χρήσιμη ως per-voxel ταξινόμηση, όμως συχνά συνδυάζεται με απώλειες επικάλυψης (π.χ. Dice loss), ώστε να μετριαστεί η ανισορροπία κλάσεων (το υπόβαθρο καταλαμβάνει μεγάλο μέρος των voxels) [33]. Μία γνωστή 3D αρχιτεκτονική που ανέδειξε τη χρησιμότητα απωλειών επικάλυψης είναι το V-Net [34].

3.2 Συνελικτικά Νευρωνικά Δίκτυα (CNNs)

Τα CNNs αποτελούν την κυρίαρχη οικογένεια μοντέλων για εικόνες, καθώς επεξεργάζονται εισόδους δομημένες σε πλέγμα και αξιοποιούν τη διαδικασία της **συνέλιξης** (convolution). Με κοινή χρήση βαρών σε όλο το πεδίο εισόδου, η ίδια ανίχνευση προτύπων εφαρμόζεται σε διαφορετικές χωρικές θέσεις, δηλαδή παρουσιάζουν **ισοδυναμία ως προς μετατοπίσεις** (translation equivariance) [35]. Οι συνελίξεις λειτουργούν ως φίλτρα που ανιχνεύουν

τοπικά πρότυπα (άκρα, υφές, σχήματα), ενώ η ιεραρχική στοιβαξή πολλών στρωμάτων επιτρέπει την αναπαράσταση ολοένα και πιο σύνθετων δομών.

Στην πράξη, τα CNNs χρησιμοποιούνται ευρέως σε προβλήματα όρασης, επειδή μαθαίνουν ιεραρχικά χαρακτηριστικά και κλιμακώνονται αποτελεσματικά με το μέγεθος των δεδομένων εκπαίδευσης [36]. Στη συνέχεια, αρχιτεκτονικές με **υπολειμματικές συνδέσεις** (residual connections) επέτρεψαν την εκπαίδευση πολύ βαθιών δικτύων [1].

3.2.1 Η συνέλιξη σε 2D/3D δεδομένα και η έννοια του receptive field

Η βασική πράξη της συνέλιξης μπορεί να θεωρηθεί ως γραμμικός μετασχηματισμός με **κοινή χρήση βαρών** σε όλα τα χωρικά σημεία. Για 3D όγκους, μια 3D συνέλιξη εφαρμόζει πυρήνα K διαστάσεων $k_x \times k_y \times k_z$ και παράγει χαρακτηριστικά που κωδικοποιούν τοπικές χωρικές σχέσεις μέσα στον όγκο. Η τοπικότητα αυτή είναι ιδιαίτερα χρήσιμη στην τμηματοποίηση, όπου τα όρια ενός οργάνου καθορίζονται από τοπικές μεταβολές έντασης και υφής.

Καθώς το βάθος του δικτύου αυξάνεται, το **receptive field** ενός νευρώνα μεγαλώνει (μέσω διαδοχικών συνελίξεων και υποδειγματοληψιών), επιτρέποντας στο μοντέλο να ενσωματώνει ευρύτερο συμφραζόμενο (wider context). Παρ' όλα αυτά, η αύξηση αυτή είναι έμμεση και εξαρτάται από το βάθος και τη σχεδίαση του δικτύου.

3.2.2 Αρχιτεκτονικές πυκνής πρόβλεψης: FCN, encoder-decoder και U-shaped σχεδίαση

Η τμηματοποίηση απαιτεί **πυκνή** πρόβλεψη (ετικέτα ανά pixel/voxel). Μια θεμελιώδης ιδέα για αυτόν τον σκοπό είναι τα Fully Convolutional Networks (FCNs), τα οποία αντικαθιστούν τα πλήρως συνδεδεμένα στρώματα με συνελικτικά, ώστε το δίκτυο να παράγει χάρτες χαρακτηριστικών (feature maps) [37].

Στην ιατρική τμηματοποίηση, η αρχιτεκτονική U-Net είναι ιδιαίτερα επιδραστική, καθώς συνδυάζει:

- έναν **κωδικοποιητή** (encoder) που συμπιέζει την πληροφορία σε πολλαπλές κλίμακες (κωδικοποιόντας το «τι»),
- έναν **αποκωδικοποιητή** (decoder) που ανακατασκευάζει την ανάλυση (βρίσκοντας το «πού»),
- **συνδέσεις παράκαμψης** (skip connections) που μεταφέρουν λεπτομέρεια υψηλής ανάλυσης από τον κωδικοποιητή στον αποκωδικοποιητή, βελτιώνοντας την χωρική ακρίβεια [5].

Για ογκομετρικά δεδομένα, η λογική αυτή επεκτάθηκε σε 3D U-Net με 3D συνελίξεις, επιτρέποντας αξιοποίηση της πληροφορίας κατά μήκος όλων των αξόνων του όγκου [38].

3.2.3 Κανονικοποίηση και σταθεροποίηση εκπαίδευσης

Σε βαθιά δίκτυα, η σταθερότητα εκπαίδευσης επηρεάζεται από την κατανομή ενεργοποίησεων. Η Batch Normalization είναι κλασική τεχνική επιτάχυνσης και σταθεροποίησης [39],

όμως στη 3D τμηματοποίηση το batch size είναι συχνά πολύ μικρό λόγω μνήμης. Σε τέτοιες περιπτώσεις, τεχνικές όπως η Group Normalization μπορεί να είναι πρακτικά πιο σταθερές [40].

3.2.4 Προσοχή και CNNs: Attention U-Net

Η συγκεκριμένη ‘προσοχή’ δεν ταυτίζεται απαραίτητα με τους Transformers. Στα CNN-βασισμένα μοντέλα, έχουν προταθεί **attention gates** που «φιλτράρουν» τις skip connections ώστε να τονίζουν περιοχές σχετικές με το ενδιαφερόμενο όργανο. Ένα κλασικό παράδειγμα είναι το Attention U-Net [13], το οποίο μπορεί να βελτιώσει την εστίαση σε μικρές ή πολύπλοκες δομές χωρίς να απαιτεί πλήρη self-attention σε όλο τον όγκο.

3.2.5 Πλεονεκτήματα και περιορισμοί των CNNs

Τα κύρια πλεονεκτήματα για τμηματοποίηση είναι:

- Υψηλή **αποδοτικότητα** (η συνέλιξη είναι υπολογιστικά ευνοϊκή).
- Έμφαση σε **τοπικά μορφολογικά στοιχεία**, χρήσιμα για ακριβή όρια.
- Ταχύτερα σε **μικρότερα σύνολα δεδομένων** σε σχέση με καθαρά Transformer μοντέλα.

Κύριος περιορισμός είναι ότι η **παγκόσμια πληροφορία** (global context) δεν είναι άμεσα διαθέσιμη: οι μακρινές χωρικές συσχετίσεις μοντελοποιούνται έμμεσα μέσω βάθους/υποδειγματοληψιών, και η αποτελεσματικότητα αυτής της διαδικασίας εξαρτάται έντονα από τη σχεδίαση του δικτύου (βάθος, μέγεθος φίλτρου (kernels), κ. λπ.).

3.3 Μοντέλα Attention και Transformers

Οι Transformers εισήχθησαν ως αρχιτεκτονική ακολουθιών με μηχανισμό **αυτο-προσοχής** (self-attention) [6]. Η ιδέα είναι ότι κάθε στοιχείο της εισόδου μπορεί να «συσχετιστεί» **άμεσα** με άλλα στοιχεία, επιτρέποντας μοντελοποίηση μακρινών εξαρτήσεων χωρίς να απαιτείται πολύ μεγάλο βάθος όπως στα προαναφερθέντα CNNs. Το πλεονέκτημα αυτό συνοδεύεται από τετραγωνικό κόστος ως προς το πλήθος tokens, κάτι που γίνεται κρίσιμο σε υψηλής ανάλυσης 2D/3D εισόδους.

3.3.1 Βασικός μηχανισμός self-attention

Για μια ακολουθία n tokens (π. χ. patches εικόνας), το self-attention υπολογίζει:

$$\text{Att}(Q, K, V) = \text{softmax}\left(\frac{QK^\top}{\sqrt{d_k}}\right)V, \quad (3.2)$$

όπου Q, K, V (queries, keys, values) είναι γραμμικοί μετασχηματισμοί των tokens. Η multi-head εκδοχή εφαρμόζει πολλαπλές τέτοιες προσοχές παράλληλα, ώστε το μοντέλο να μαθαίνει διαφορετικούς τύπους συσχετίσεων.

3.3.2 Self-Attention και Εικόνες: προκλήσεις και λύσεις

Ο κλασικός self-attention έχει $O(n^2)$ υπολογιστική και χωρική πολυπλοκότητα ως προς το πλήθος tokens n . Στην όραση, όμως, το n προκύπτει από το πλήθος patches: για 2D εικόνα $n \approx (H \cdot W)/P^2$, ενώ για 3D όγκο $n \approx (H \cdot W \cdot D)/P^3$, όπου P είναι η ακμή του patch. Ακόμη και με patching, το n παραμένει μεγάλο για πρακτική 3D εκπαίδευση.

Ο Vision Transformer (ViT) [41] διατύπωσε τη 2D όραση ως ακολουθία patch tokens με positional embeddings. Ωστόσο, για ιατρικά δεδομένα (συχνά μικρότερα σύνολα και 3D όγκοι), απαιτούνται πρόσθετες σχεδιαστικές επιλογές ώστε να είναι πρακτική η εκπαίδευση.

3.3.3 Μείωση κόστους: ιεραρχία, τοπικά παράθυρα και προσεγγίσεις γραμμικής προσοχής

Κύριες κατευθύνσεις για να καταστεί ο attention πρακτικός σε μεγάλες εισόδους είναι:

- **Ιεραρχική αναπαράσταση** με υποδειγματοληψίες, ώστε να μειώνεται σταδιακά το πλήθος tokens.
- **Τοπική προσοχή σε παράθυρα** (window attention): ο Swin Transformer υπολογίζει προσοχή σε τοπικά παράθυρα και τα μετατοπίζει (shifted windows) για ανταλλαγή πληροφορίας μεταξύ τους, επιτυγχάνοντας αποδοτική κλιμάκωση [42].
- **Προσεγγίσεις υπο-τετραγωνικής προσοχής**: π. χ. χαμηλόβαθμη προβολή (Lformer) [43] ή στοχαστικές προσεγγίσεις με random features (Performer) [44].

Στο πλαίσιο ιατρικής τμηματοποίησης, συχνά προτιμώνται οι δύο πρώτες στρατηγικές (ιεραρχία/παράθυρα), επειδή διατηρούν καλύτερα τη χωρική δομή και είναι πιο ‘φιλικές’ σε 3D δεδομένα.

3.3.4 Υθριδικά Transformer-U-shape μοντέλα

Υθριδικά μοντέλα όπως τα UNETR [7] και SwinUNETR [8] χρησιμοποιούν Transformer-ενισχυμένους κωδικοποιητές για καλύτερη κατανόηση συμφραζόμενων, διατηρώντας ταυτόχρονα decoder που ανακτά λεπτομέρεια υψηλής ανάλυσης, όπως στο U-Net. Έτσι επιχειρείται συνδυασμός:

- **παγκόσμιας/μακρινής συσχέτισης** (ισχυρό σημείο των Transformers)
- με **ακριβή χωρική ανακατασκευή** (ισχυρό σημείο της U-shaped αποκωδικοποίησης).

3.4 Μοντέλα Χώρου Καταστάσεων (State Space Models) και Mamba

Τα Μοντέλα Χώρου Καταστάσεων (State Space Models - SSMS) αποτελούν μια κλασική κατηγορία δυναμικών συστημάτων που περιγράφουν την εξέλιξη μιας κρυφής κατάστασης $h_t \in \mathbb{R}^N$ μέσω μιας ακολουθίας εισόδων $x_t \in \mathbb{R}^M$. Σε διακριτή μορφή, ένα γραμμικό, χρονικά

αναλλοίωτο (Linear Time-Invariant - LTI) SSM περιγράφεται από:

$$h_t = \bar{A}h_{t-1} + \bar{B}x_t, \quad (3.3)$$

$$y_t = Ch_t + Dx_t, \quad (3.4)$$

όπου \bar{A}, \bar{B}, C, D είναι οι παράμετροι του μοντέλου. Αυτή η αναδρομική μορφή επιτρέπει θεωρητικά ‘άπειρης’ μνήμη, αλλά για μεγάλες ακολουθίες ο υπολογισμός είναι αυστηρά διαδοχικός και, κατά συνέπεια, αργός. Ωστόσο, αξιοποιώντας την LTI ιδιότητα, η εξέλιξη του συστήματος μπορεί να υπολογιστεί ισοδύναμα ως γραμμική συνέλιξη (global convolution), επιτρέποντας τον πλήρη παραλληλισμό [45].

Παρότι τα LTI SSMs είναι υπολογιστικά αποδοτικά (π.χ. μέσω της ισοδυναμίας με γραμμική συνέλιξη), στην απλή γραμμική τους μορφή συχνά υστερούν σε εκφραστικότητα σε σύγκριση με Transformers [9, 45]. Ένας βασικός λόγος είναι ότι οι παράμετροι (π.χ. \bar{A} και \bar{B}) δεν εξαρτώνται από την είσοδο, άρα απουσιάζει εγγενής content-based selection [9]. Αυτό οδήγησε σε δομημένα SSMs (π.χ. S4) και σε επιλεκτικά SSMs όπως το Mamba, όπου κρίσιμες παράμετροι επηρεάζονται από την είσοδο.

3.4.1 Δομημένα Μοντέλα Χώρου Καταστάσεων (Structured SSMs) και Μοντελοποίηση Μακρινών Εξαρτήσεων

Για να καταστούν τα SSMs ανταγωνιστικά σε προβλήματα με πολύ μεγάλες ακολουθίες (π.χ. ήχος, DNA), εισήχθησαν τα **Δομημένα Μοντέλα Χώρου Καταστάσεων (Structured SSMs)**. Μοντέλα όπως το **S4** επιβάλλουν συγκεκριμένη δομή στον πίνακα κατάστασης, επιτρέποντας αποδοτικό υπολογισμό και στις δύο μορφές μέσω ειδικών αλγορίθμων [45]. Το S4 επέδειξε ισχυρή απόδοση σε εργασίες που απαιτούν μοντελοποίηση μακρινών εξαρτήσεων [45, 46].

3.4.2 Mamba: Επιλεκτικά Μοντέλα Χώρου Καταστάσεων και Γραμμική Κλιμάκωση

Παρόλο που τα δομημένα SSMs (όπως το S4) είναι αποτελεσματικά σε συνεχή σήματα, αντιμετωπίζουν δυσκολίες σε διακριτά δεδομένα όπως το κείμενο, όπου η ικανότητα **επιλογής με βάση το περιεχόμενο** (content-based selection) είναι κρίσιμη [9]. Το μοντέλο **Mamba** επιλύει αυτόν τον περιορισμό εισάγοντας εκλεκτικούς χώρους καταστάσεων (selective state spaces): κρίσιμες παράμετροι (π.χ. \bar{B}, C και το βήμα διακριτοποίησης Δ) γίνονται συναρτήσεις της τρέχουσας εισόδου x_t , επιτρέποντας στο μοντέλο να ‘αποφασίζει’ δυναμικά ποια πληροφορία θα διατηρήσει.

Η αλλαγή αυτή δυσκολεύει τη χρήση καθαρής συνελικτικής μορφής, οπότε το Mamba βασίζεται σε **υλοποίηση βελτιστοποιημένη για GPU** (hardware-aware scan), ώστε να διατηρεί **γραμμική πολυπλοκότητα** $O(L)$ ως προς το μήκος L της ακολουθίας, με υψηλή πρακτική αποδοτικότητα [9].

3.4.3 Εφαρμογή SSMS/Mamba σε 3D Τμηματοποίηση

Για την εφαρμογή των SSMS σε ογκομετρικά δεδομένα, απαιτείται η σειριοποίηση του όγκου σε ακολουθία (serialization). Μια τυπική μεθοδολογία περιλαμβάνει:

1. Διαχωρισμό του 3D όγκου σε patches και προβολή σε tokens.
2. Ιεραρχική αρχιτεκτονική τύπου U-shaped για multi-scaling processing.
3. Ενσωμάτωση Mamba blocks που αντικαθιστούν τα στρώματα προσοχής ή μέρος των συνελικτικών blocks.

Με αυτόν τον τρόπο, το μοντέλο μπορεί να συλλάβει **μακρινές χωρικές εξαρτήσεις** σε ολόκληρο τον όγκο με σημαντικά χαμηλότερο υπολογιστικό κόστος σε σχέση με πλήρη self-attention. Πρόσφατες προσεγγίσεις όπως το SegMamba [10] δείχνουν ότι η ιδέα αυτή μπορεί να δώσει ανταγωνιστική ακρίβεια σε 3D τμηματοποίηση.

3.4.4 Συγκριτική Επισκόπηση

Συνοψίζοντας τη θεωρητική προοπτική για την τμηματοποίηση πλακούντα:

- **CNNs / U-Net:** Αποδοτικά για τοπικό σήμα, αλλά η ‘καθολική’ πληροφορία εισάγεται έμμεσα (μέσω βάθους).
- **Transformers:** Παρέχουν άμεση καθολική συσχέτιση (global context), αλλά το τετραγωνικό κόστος $O(L^2)$ είναι συχνά απαγορευτικό για υψηλής ανάλυσης 3D δεδομένα.
- **SSMs / Mamba:** Στοχεύουν σε **γραμμική κλιμάκωση** $O(n)$ με ικανότητα μοντελοποίησης μακρινών εξαρτήσεων, αποτελώντας υποσχόμενη μέση λύση.

3.5 Βελτιστοποίηση και πρακτικές εκπαίδευσης

Η εκπαίδευση βαθιών δικτύων πραγματοποιείται συνήθως με στοχαστική κλίση κατάβασης και παραλλαγές της. Στο πλαίσιο αυτό, ο AdamW αποτελεί διαδεδομένη επιλογή λόγω του αποσυνδεδεμένου (decoupled) weight decay [47].

Στον Adam, οι εκθετικοί κινητοί μέσοι πρώτης και δεύτερης ροπής υπολογίζονται ως:

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1)g_t, \quad v_t = \beta_2 v_{t-1} + (1 - \beta_2)g_t^2.$$

Όταν η κανονικοποίηση εφαρμόζεται ως L^2 -norm μέσα στο gradient, ισχύει:

$$g_t = \nabla_{\theta} \mathcal{L}(\theta_t) + \lambda \theta_t.$$

Σε αυτή την περίπτωση, ο όρος κανονικοποίησης αλληλεπιδρά με την προσαρμοστική κλίμακα του Adam. Αντίθετα, στον AdamW η συρρίκνωση των βαρών εφαρμόζεται ξεχωριστά από τον προσαρμοστικό όρο ενημέρωσης:

$$\hat{\theta}_{t+1} = \theta_t - \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon} - \eta \lambda \theta_t.$$

Η αποσύνδεση αυτή καθιστά τη ρύθμιση του weight decay πιο προβλέψιμη ως υπερπαράμετρο και συχνά συνδέεται με καλύτερη γενίκευση σε πρακτικές εφαρμογές [47].

Στην 3D τμηματοποίηση ακολουθούνται κατά κανόνα οι εξής πρακτικές:

- **Μικρό batch size** λόγω περιορισμών στη μνήμη της GPU.
- **Εκπαίδευση με patches** και ισορροπία θετικών/αρνητικών δειγμάτων.
- **Μεταβαλλόμενος ρυθμός μάθησης** (learning rate schedules) για επίτευξη βέλτιστης σύγκλισης (π.χ. CyclicLR, Cosine Annealing [48]).
- **Κανονικοποίηση** (π.χ. GroupNorm ή InstanceNorm) όταν το batch size είναι πολύ μικρό.

Στην παρούσα εργασία, οι παραπάνω επιλογές εφαρμόζονται ενιαία στα πειράματα, με στόχο τη διασφάλιση μιας **δίκαιης σύγκρισης** μεταξύ διαφορετικών αρχιτεκτονικών. Με τον τρόπο αυτό, απομονώνεται η επίδραση της σχεδίασης του μοντέλου στην ακρίβεια τμηματοποίησης πλακούντα σε εικόνες MRI, ελαχιστοποιώντας την επιρροή των υπερπαραμέτρων εκπαίδευσης [11].

Κεφάλαιο 4

Αρχιτεκτονικές Τμηματοποίησης που Χρησιμοποιήθηκαν

Στο κεφάλαιο αυτό παρουσιάζονται οι αρχιτεκτονικές που χρησιμοποιήθηκαν στα πειράματα της παρούσας εργασίας. Η ανάλυση εστιάζει στον τρόπο με τον οποίο κάθε μοντέλο συλλέγει context (τοπικό ή μακρινής εμβέλειας), καθώς και στους συμβιβασμούς ακρίβειας, μνήμης και υπολογιστικού κόστους στην 3D ιατρική τμηματοποίηση.

4.1 Το πλαίσιο MONAI

Το MONAI (Medical Open Network for AI) είναι ένα ανοικτού κώδικα, PyTorch-βασεδ πλαίσιο που στοχεύει στη διευκόλυνση ανάπτυξης και αξιολόγησης μοντέλων Βαθιάς Μάθησης για εφαρμογές υγείας, με έμφαση στην ιατρική απεικόνιση [11]. Παρέχει:

- **Μετασχηματισμούς** (transforms) για 3D δεδομένα (φόρτωση, resampling, κανονικοποίηση, augmentation).
- **Έτοιμες αρχιτεκτονικές** τμηματοποίησης και building blocks (συνελικτικά/υπολειμματικά βλοσκς, κανονικοποίησεις, upsampling).
- **Απώλειες και μετρικές** (π. χ. Dice-based losses και Dice/IoU μετρικές) και τυποποιημένες ροές εκπαίδευσης.
- **Μηχανισμούς inference** όπως sliding-window inference για μεγάλους 3D όγκους.

Η χρήση του MONAI επιτρέπει πιο **αναπαραγώγιμο** και **συγκρίσιμο** πειραματικό πρωτόκολλο, καθώς πολλές επιλογές υλοποίησης είναι κοινές και ελεγχόμενες [11]. Για περαιτέρω τεχνικές λεπτομέρειες παραπέμπουμε στην επίσημη τεκμηρίωση και στο αποθετήριο του έργου [49, 50].

4.1.1 Υλοποίηση μοντέλων στην παρούσα εργασία

Στην παρούσα εργασία, όλες οι αρχιτεκτονικές UNet, DynUNet, SegResNet, AttentionUnet, UNETR και SwinUNETR χρησιμοποιούνται μέσω των αντίστοιχων υλοποιήσεων του MONAI [11]. Εξαίρεση αποτελεί το SegMamba, το οποίο **δεν διατίθεται στο MONAI** και ενσωματώνεται μέσω της **επίσημης υλοποίησης των συγγραφέων** [10, 51]. Για να διατηρηθεί

δίκαιη σύγκριση, το SegMamba εκπαιδεύεται και αξιολογείται με το ίδιο πρωτόκολλο προ-επεξεργασίας, δειγματοληψίας patches, απώλειας, μετρικών και sliding-window inference που εφαρμόζεται και στα MONAI-μοντέλα.

4.2 Οικογένειες μοντέλων που συγκρίνονται

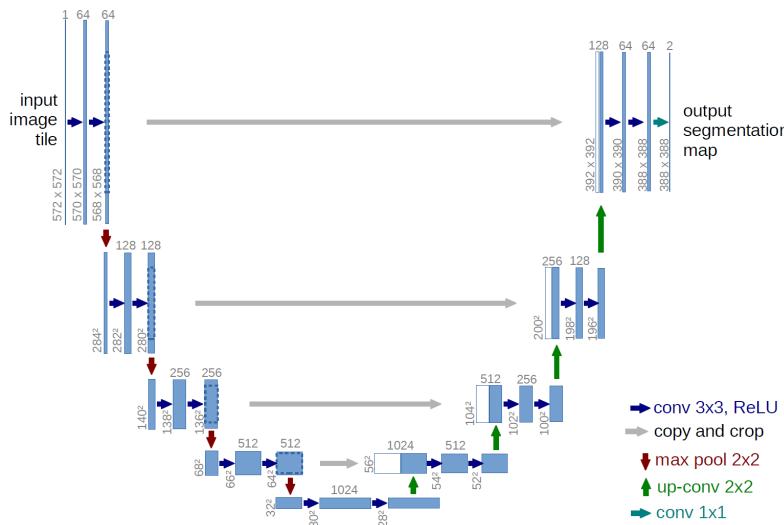
Η αρχιτεκτονικές ομαδοποιούνται ως εξής:

- **CNN-based:** UNet, DynUNet, SegResNet, AttentionUnet.
- **Transformer-ενισχυμένα:** UNETR, SwinUNETR.
- **SSM/Mamba-based:** SegMamba (εκτός MONAI, μέσω επίσημου κώδικα).

4.3 UNet και 3D U-Net

Το U-Net [5] αποτελεί σημείο αναφοράς στην ιατρική τμηματοποίηση. Πρόκειται για συμμετρική αρχιτεκτονική 4.1 encoder-decoder με skip connections:

- Ο **κωδικοποιητής** (encoder) μειώνει την ανάλυση (downsampling), αυξάνοντας τα κανάλια, ώστε να κωδικοποιεί υψηλού επιπέδου χαρακτηριστικά.
- Ο **αποκωδικοποιητής** (decoder) ανακτά την ανάλυση (upsampling) και συνδυάζει πληροφορία υψηλής ανάλυσης από τον encoder, ώστε να βελτιώνει την ακρίβεια εντοπισμού ορίων.



Σχήμα 4.1: Δομή αρχιτεκτονικής U-net

Για ογκομετρικά δεδομένα, η λογική επεκτείνεται σε 3D U-Net με 3D συνελίξεις [38]. Τα skip connections είναι κρίσιμα για ακριβή localization, καθώς μετριάζουν την απώλεια λεπτομέρειας στα στάδια υποδειγματοληψίας.

Σχόλιο. Το UNet τείνει να αποδίδει πολύ καλά σε τοπικά μορφολογικά μοτίβα και οριακές μεταβολές. Ο κύριος περιορισμός είναι ότι το global context εισάγεται έμμεσα μέσω βάθους και multi-scaling συμπίεσης.

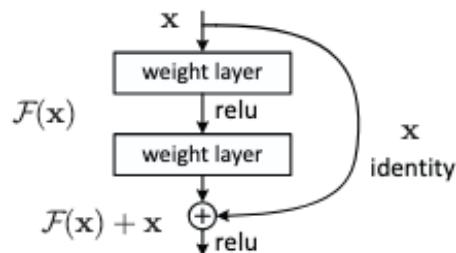
4.4 DynUNet

Το DynUNet (στη MONAI υλοποίηση) είναι παραλλαγή 3D UNet που επιτρέπει **δυναμική** προσαρμογή της τοπολογίας (στάδια, strides, μεγέθη μυρήνων) με βάση τις διαστάσεις/ανάλυση εισόδου και το patch size. Η φιλοσοφία συνδέεται με το nnU-Net, όπου έδειξε ότι συστηματικές επιλογές προεπεξεργασίας και ‘εύλογες’ αρχιτεκτονικές αποφάσεις μπορούν να δώσουν ισχυρά αποτελέσματα σε πολλά datasets [52].

Κεντρική ιδέα. Η προσαρμογή επιδιώκει επαρκές receptive field με ελεγχόμενο κόστος, ώστε να μην χάνεται υπερβολικά χωρική πληροφορία σε 3D δεδομένα. Συχνά υποστηρίζεται και deep supervision, που μπορεί να βοηθήσει στη σύγκλιση σε ρυθμίσεις μικρού batch size [52].

4.5 SegResNet: υπολειμματικό encoder-decoder

Το SegResNet χρησιμοποιεί υπολειμματικά blocks τύπου ResNet [1] μέσα σε encoder-decoder σχήμα. Μία επιδραστική μορφή αυτού του στυλ για 3D MRI τμηματοποίηση είναι η προσέγγιση του Myronenko [12], όπου η υπολειμματική δομή συνδυάζεται και με autoencoder regularization.

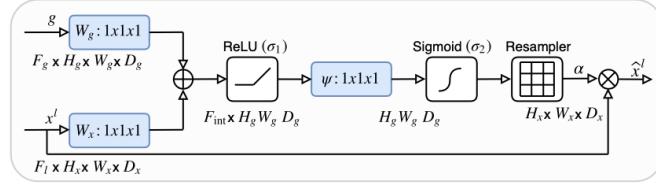


Σχήμα 4.2: Μηχανισμός residual connection στην αρχιτεκτονική ResNet [1]

Γιατί υπολείμματα. Οι υπολειμματικές συνδέσεις (4.1) διευκολύνουν τη ροή κλίσεων και επιτρέπουν εκπαίδευση βαθύτερων δικτύων με σταθερότερη σύγκλιση [1]. Αυτό είναι χρήσιμο όταν απαιτείται μεγαλύτερο βάθος για ενσωμάτωση συμφραζομένων σε 3D όγκους.

4.6 AttentionUnet: προσοχή στις skip connections

Η Attention U-Net εισάγει attention gates 4.3 στις skip connections, ώστε το δίκτυο να ενισχύει περιοχές σχετικές με το ROI και να καταστέλλει άσχετο υπόβαθρο [13]. Το gate συνδυάζει πληροφορία από ρηχότερα επίπεδα (λεπτομέρεια) και βαθύτερα επίπεδα (συμφραζόμενο), υπολογίζοντας βάρη εστίασης.



Σχήμα 4.3: Προτεινόμενος σχεδιασμός του *attention gate*

Πότε βοηθά. Η προσοχή είναι ιδιαίτερα χρήσιμη όταν το ROI είναι μικρό, υπάρχει ισχυρή ανισορροπία κλάσεων ή όταν δομές έχουν παρόμοια ένταση με το υπόβαθρο.

4.7 UNETR: ViT encoder με U-shaped decoder

To UNETR [7] αποτελεί υβριδικό σχήμα, όπου ο encoder βασίζεται σε Transformer (ViT-like) και ο decoder διατηρεί τη λογική ανακατασκευής τύπου U-Net. Η είσοδος χωρίζεται σε μη επικαλυπτόμενα **3D patches** που προβάλλονται σε tokens, ενώ το self-attention επιτρέπει άμεση μοντελοποίηση μακρινών συσχετίσεων.

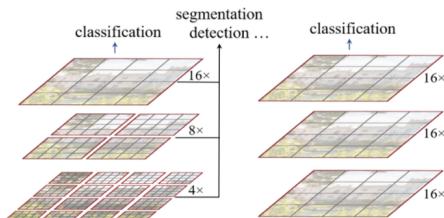
Σχόλιο κόστους. Το πλεονέκτημα του global context συνοδεύεται από αυξημένο κόστος μνήμης/χρόνου, το οποίο εξαρτάται έντονα από το patch grid. Για αυτό, η επιλογή patch size και ο τρόπος ινφερενς (π.χ. sliding-window) είναι κρίσιμα πρακτικά σημεία [7].

4.8 SwinUNETR: ιεραρχικός Swin encoder

To SwinUNETR βασίζεται σε Swin Transformer encoder με ιεραρχία και window-based attention [8, 42].

Αντί για πλήρη attention σε όλα τα tokens, η προσοχή υπολογίζεται μέσα σε τοπικά παράθυρα. Τα παράθυρα μετατοπίζονται (shifted windows) μεταξύ διαδοχικών blocks, επιτρέποντας ανταλλαγή πληροφορίας μεταξύ γειτονικών περιοχών.

Ιεραρχία έναντι ViT. Σε ένα κλασικό ViT ο αριθμός και η διάταξη των tokens παραμένουν σταθερά σε όλο τον encoder (μονο-κλίμακη αναπαράσταση). Ο Swin χωρίζεται σε διακριτά stages και εφαρμόζει patch merging, μειώνοντας προοδευτικά τη χωρική ανάλυση και δημιουργώντας πυραμίδα χαρακτηριστικών πολλαπλών κλιμάκων (4.4).



Σχήμα 4.4: Αριστερά: Swin με ιεραρχικά stages (πολλαπλές κλίμακες μέσω patch merging). Δεξιά: ViT χωρίς ιεραρχία, με σταδερή ανάλυση tokens σε όλο τον encoder.

Γιατί ταιριάζει σε 3D. Η τοπική προσοχή μειώνει το υπολογιστικό κόστος σε μεγάλους όγκους, ενώ η ιεραρχία (μέσω patch merging) μειώνει προοδευτικά το πλήθος tokens. Τα ενδιάμεσα επίπεδα πολλαπλών κλιμάκων τροφοδοτούν έναν U-shaped decoder με skip connections, διατηρώντας λεπτομέρεια και χωρική ακρίβεια.

4.9 SegMamba: Mamba/SSM blocks σε U-shape σχεδίαση

To SegMamba προτείνει μια 3D αρχιτεκτονική τμηματοποίησης που ενσωματώνει Mamba-βασισμένα SSM βλοσκς για μοντελοποίηση μακρινών εξαρτήσεων με ευνοϊκή κλιμάκωση σε σχέση με πλήρες self-attention [10]. Ο μηχανισμός Mamba εισάγει selective state spaces, επιτρέποντας content-dependent παραμετροποίηση που λειτουργεί ως εναλλακτική της attention σε μεγάλες ακολουθίες [9].

Σημείωση υλοποίησης. Σε αντίθεση με τα υπόλοιπα μοντέλα που προέρχονται από MONAI, η αρχιτεκτονική SegMamba υλοποιείται στην παρούσα εργασία μέσω του **επίσημου αποθετηρίου** των συγγραφέων [51]. Η ενσωμάτωση γίνεται με τρόπο που να επιτρέπει κοινό πειραματικό πρωτόκολλο (προεπεξεργασία, patch sampling, loss, metrics και inference), ώστε η σύγκριση να παραμένει συνεπής.

Αναμενόμενη συμπεριφορά. Η σχεδίαση στοχεύει να συνδυάσει:

- συλλογή **μακρινής** εμβέλειας πληροφοριών σε πολλαπλές κλίμακες,
- U-shaped αποκαθικοποίηση για διατήρηση χωρικής ακρίβειας,
- βελτιωμένη διαχείριση πόρων σε 3D δεδομένα, σε σύγκριση με Transformer encoders αντίστοιχης ανάλυσης.

4.10 Συζήτηση: αναμενόμενοι συμβιβασμοί

Συνοψίζοντας, οι οικογένειες μοντέλων διαφέρουν ως προς το **πώς** συλλέγουν συμφραζόμενο:

- **CNN-based:** ισχυρά σε όρια/υφές και αποδοτικά, αλλά το global context εισάγεται έμμεσα.
- **Transformer-based:** ισχυρό global context, αλλά αυξημένες απαιτήσεις μνήμης/χρόνου.
- **SSM/Mamba-based:** στόχος η μακρινής εμβέλειας μοντελοποίηση με καλύτερη κλιμάκωση, κάτι ιδιαίτερα ελκυστικό σε 3D τμηματοποίηση.

Στο πειραματικό μέρος παρουσιάζεται το κοινό πρωτόκολλο εκπαίδευσης/αξιολόγησης που επιτρέπει να αποτιμήθουν εμπειρικά οι παραπάνω συμβιβασμοί για την τμηματοποίηση πλακούντα σε MRI.

Μέρος II

Πρακτικό Μέρος

Κεφάλαιο 5

Πειραματικό Πρωτόκολλο και Μεθοδολογία Αξιολόγησης

Στο κεφάλαιο αυτό παρουσιάζεται η πειραματική διαδικασία που θεσπίστηκε με σκοπό την εκπαίδευση και τη συγκριτική αξιολόγηση διαφορετικών αρχιτεκτονικών, με γνώμονα τη διεργασία της τιμηματοποίησης σε 3D MRI εικόνες.

Η επιλογή των μοντέλων προέκυψε έπειτα από συστηματική μελέτη της βιβλιογραφίας: αρχικά εξετάστηκαν τα πρωτότυπα δίκτυα σχήματος U (U-Net) [5] και στη συνέχεια παραλαγές/επεκτάσεις όπως το Attention U-Net [13], σύγχρονες U-Net μορφές όπως το DynUNet [52] και όμοιες CNN-based encoder-decoder μορφές που υλοποιούν υπολειμματικές συνδέσεις (SegResNet) [12]. Τέλος, αξιολογήθηκαν νεότερες προσεγγίσεις που μοντελοποιούν μακρινές εξαρτήσεις, όπως τα Transformer-based μοντέλα (UNETR [7], SwinUNETR [8]) και State Space-based μοντέλα (SegMamba [10], επέκταση του Mamba [9]).

Κεντρικός στόχος είναι η **δίκαιη και αναπαραγώγιμη σύγκριση** των αρχιτεκτονικών. Για τον λόγο αυτό υλοποιήθηκε ένα **ενιαίο πειραματικό πλαίσιο** όπου όλες οι αρχιτεκτονικές εκπαιδεύονται με κοινή ροή (pipeline) προεπεξεργασίας, δειγματοληψίας patches, εκπαίδευσης, αξιολόγησης και καταγραφής μετρικών. Στην πράξη, κάθε αρχιτεκτονική αντιστοιχεί σε ξεχωριστό notebook, όμως η δομή του κώδικα παραμένει σταθερή, με κύρια διαφοροποίηση την αρχικοποίηση του μοντέλου και ελάχιστες αναγκαίες ρυθμίσεις συμβατότητας.

5.1 Στόχοι αξιολόγησης και αρχές σύγκρισης

Στόχος της πειραματικής αξιολόγησης είναι να αποτυπωθεί με σαφήνεια η απόδοση διαφορετικών αρχιτεκτονικών στη τιμηματοποίηση πλακούντα σε 3D MRI. Η σύγκριση σχεδιάστηκε ώστε να είναι δίκαιη, με κοινό τρόπο αναφοράς αποτελεσμάτων και συνεπείς επιλογές αξιολόγησης, έτοι ώστε οι παρατηρούμενες διαφορές να αποδίδονται όσο δυνατόν αποκλειστικά στην αρχιτεκτονική.

Η αναφορά βασίζεται κυρίως σε Dice και IoU, μετρικές που είναι κατάλληλες για προβλήματα έντονης ανισορροπίας κλάσεων, ενώ παράλληλα συμπληρώνεται από ποιοτική επιθεώρηση και τυπική σύγκριση με αποτελέσματα από άλλες έρευνες ομοίου σκοπού (βλ. Κεφάλαιο 6).

5.2 Υλοποίηση και λογισμικό

Τα πειράματα υλοποιήθηκαν κυρίως με χρήση των βιβλιοθηκών PyTorch και MONAI [11] σε γλώσσα Python. Η εκπαίδευση εκτελέστηκε σε περιβάλλον cloud μέσω της πλατφόρμας Kaggle Notebooks, αξιοποιώντας GPU επιτάχυνση (NVIDIA T4 σε διαθέσιμη διαμόρφωση) εντός του εβδομαδιαίου ορίου χρήσης που παρέχεται από την πλατφόρμα.¹

Οι εκτελέσεις πραγματοποιήθηκαν ως *non-interactive runs* (με καταγραφή logs/μετρικών), ώστε να διατηρείται σταθερή η διαδικασία μεταξύ μοντέλων και να αποφεύγονται χειροκίνητες παρεμβάσεις. Για αναπαραγωγιμότητα χρησιμοποιήθηκαν σταθερά seeds σε NumPy και PyTorch, καθώς και κοινή δομή κώδικα για όλα τα μοντέλα (με διαφοροποίηση κυρίως στην αρχικοποίηση της αρχιτεκτονικής).

5.3 Σύνολο δεδομένων και διαχωρισμός

Το σύνολο δεδομένων περιλαμβάνει $N = 137$ περιστατικά, με ένα ζεύγος αρχείων .nii.gz ανά περίπτωση: την εικόνα καταγραφής του ασθενή στον Μαγνητικό Τομογράφο και την αντίστοιχη δυαδική μάσκα τμηματοποίησης πλακούντα. Η επέκταση αρχείου nii προέρχεται από την λέξη NIfTI όπου είναι ακρωνύμιο των λέξεων **Neuroimaging Informatics Technology Initiative**. Με αυτή τη δομή μαζί με την εικόνα αποθηκεύονται και στατιστικά, ρυθμίσεις φωτισμού και διάφορες άλλες χρήσιμες πληροφορίες για την εικόνα (metadata).

Ως βασική μονάδα εκπαίδευσης και αξιολόγησης ορίζεται το case (δηλαδή ολόκληρος ο 3D όγκος), ώστε ο διαχωρισμός να γίνεται σε επίπεδο ασθενή/περίπτωσης και να αποφεύγεται διαρροή πληροφορίας.

Ο διαχωρισμός σε σύνολα εκπαίδευσης και επικύρωσης πραγματοποιείται με σταθερό seed (121) για λόγους αναπαραγωγιμότητας. Στα πειράματα της παρούσας εργασίας χρησιμοποιήθηκε διαχωρισμός 80/20, που αντιστοιχεί περίπου σε 109 περιστατικά εκπαίδευσης και 28 επικύρωσης (validation).

5.4 Οργάνωση δεδομένων και ροή φόρτωσης

Κάθε περιστατικό περιλαμβάνει έναν 3D MRI όγκο και την αντίστοιχη δυαδική μάσκα πλακούντα σε μορφή NIfTI (.nii). Τα δεδομένα μετατρέπονται σε λίστα από dictionaries (το καθένα περιέχει ένα κλειδί για image με τιμή το μονοπάτι προς την εικόνα και ένα κλειδί label αντίστοιχα με τιμή το μονοπάτι προς την μάσκα), η οποία τροφοδοτείται σε Dataset του MONAI.

Για επιτάχυνση, το αιτιοκρατικό τμήμα της προεπεξεργασίας υπολογίζεται μια φορά και αποθηκεύεται σε disk cache μέσω **PersistentDataset**. Αντιθέτως, οι στοχαστικοί μετασχηματισμοί (Random Augmentations, εφαρμόζονται μόνο στο training set) εκτελούνται κάθε εποχή κατά τη φόρτωση των δεδομένων μέσω DataLoader. Αυτό συμβαίνει ώστε να έχουμε διαφορετικές παραλλαγές των ίδιων δειγμάτων χωρίς **επαναλαμβανόμενο** κόστος σταθερής προεπεξεργασίας.

¹Οι ακριβείς διαθέσιμοι πόροι και τα όρια χρήσης μπορεί να μεταβάλλονται με τον χρόνο και εξαρτώνται από την πολιτική της πλατφόρμας.

5.5 Ανάλυση στοίβας προεπεξεργασίας

Η στοίβα προεπεξεργασίας σχεδιάστηκε έχοντας υπόψη δύο βασικά σημεία: (α) υψηλό background to foreground ratio (ο πλακούντας καταλαμβάνει μικρό μέρος του όγκου) και (β) την ανάγκη γεωμετρικής συνέπειας μεταξύ περιστατικών. Οι μετασχηματισμοί επιλέχθηκαν από το MONAI [53].

5.5.1 Γεωμετρική εναρμόνιση και σταθεροποίηση εισόδου

Το πρώτο στάδιο εξασφαλίζει ότι το μοντέλο διαβάζει τα δεδομένα σε κοινό γεωμετρικό πλαίσιο:

- **Φόρτωση και τυποποίηση:** ανάγνωση NIfTI, μεταφορά σε tensor και εξασφάλιση σωστού καναλιού (channel-first), ώστε οι επόμενοι μετασχηματισμοί να εφαρμόζονται ομοιόμορφα.
- **Επαναπροσανατολισμός (Orientation):** μετατροπή σε κοινό σύστημα αξόνων (RAS), ώστε να αποφεύγονται ασυνέπειες σε left/right, anterior/posterior κ.λπ.
- **Αναδειγματοληψία (Spacing):** μετασχηματισμός σε κοινό voxel spacing ((2.0, 2.0, 2.0) mm), ώστε η μάθηση να μην επηρεάζεται από διαφορετικές φυσικές κλίμακες και ανομοιογενείς αναλύσεις μεταξύ εξετάσεων.

5.5.2 Κανονικοποίηση εντάσεων

Εφαρμόζεται κανονικοποίηση με percentiles (π.χ. 2.0-99.9) και χαρτογράφηση στο [0, 1]. Η πρακτική αυτή (a) περιορίζει την επίδραση outliers/θορύβου και (β) διευκολύνει τη σταθερότητα του optimization.

5.5.3 Περιορισμός πεδίου (ROI) με foreground cropping

Δεδομένης της μεγάλης έκτασης background, εφαρμόστηκε περικοπή σε πλαίσιο με την CropForeground με source_key=label και μικρό περιθώριο (margin= 8) ώστε να περιορίζεται η εικόνα γύρω από την περιοχή ενδιαφέροντος. Με τον τρόπο αυτό μειώνεται δραστικά το μέγεθος του υπολογιστικού κόστους και παράλληλα αυξάνονται την πιθανότητα τα patches να περιέχουν χρήσιμη πληροφορία.

5.5.4 Σταθεροποίηση διαστάσεων για συμβατότητα μοντέλων

Τέλος, εφαρμόζεται:

- **Padding σε σταθερό roi_size ((96, 96, 64))**, για ομοιόμορφη εκπαίδευση με patch-based στρατηγική.
- **Divisible padding** ώστε οι τελικές διαστάσεις να είναι πολλαπλάσια συγκεκριμένων παραγόντων (π.χ. (32, 32, 16)), κάτι που διευκολύνει αρχιτεκτονικές με downsample/patch merging.

5.6 Δειγματοληψία patches και επαύξηση δεδομένων

Λόγω του μεγέθους των 3D όγκων και περιορισμών μνήμης, η εκπαίδευση πραγματοποιείται σε patches σταθερού μεγέθους (roi_size= (96, 96, 64)).

Για την αντιμετώπιση της έντονης ανισορροπίας background/foreground χρησιμοποιήθηκε RandCropByPosNegLabeld, ώστε να ελέγχεται η πιθανότητα επιλογής patches που περιέχουν foreground. Επιπλέον εφαρμόστηκε μεταβολή στην αναλογία pos/neg κατά τη διάρκεια της εκπαίδευσης, με στόχο αρχικά να σταθεροποιηθεί η μάθηση πάνω στο foreground και στη συνέχεια να μειωθούν τα false positives μέσω σταδιακής εισαγωγής αρνητικών δειγμάτων. Για τα πρώτα 40 epochs είχαμε αποκλειστικά foreground δεδομένα, έπειτα για 30 epochs υπήρχε αναλογία 3/1 foreground to background δεδομένα, και για τις υπόλοιπες εποχές ισάξια αναλογία θετικών προς αρνητικών δειγμάτων.

Για βελτίωση της γενίκευσης, εφαρμόστηκαν στοχαστικές επαυξήσεις (random flips, ήπιες περιστροφές, διαστρεβλώσεις όπως stretching κ.λ.π., Gaussian Θόρυβος/εξομάλυνση), οι οποίες προσομοιώνουν ρεαλιστική μεταβλητότητα χωρίς να αλλοιώνουν τη σημασιολογία της μάσκας.

5.7 Αρχιτεκτονικές και πειραματικά σενάρια

Αξιολογήθηκαν οι παρακάτω αρχιτεκτονικές:

- U-Net [5]
- Attention U-Net [13]
- DynUNet [52]
- UNETR [7]
- SwinUNETR [8]
- SegResNet [12]
- SegMamba [10]

Οικογένειες μοντέλων. Οι αρχιτεκτονικές που αξιολογούνται καλύπτουν κλασικές CNN προσεγγίσεις (U-Net και παραλλαγές), νεότερες Transformer-based δομές (UNETR, SwinUNETR) και αποδοτικές εναλλακτικές για μοντελοποίηση μακρινών εξαρτήσεων (state space-based, SegMamba). Με αυτόν τον τρόπο εξετάζεται τόσο η επίδραση ισχυρής τοπικής επαγωγικής μεροληψίας (inductive bias) όσο και η συμβολή μηχανισμών καθολικών συμφράζομένων (global context).

5.7.1 Αρχιτεκτονικές υπερπαράμετροι ανά μοντέλο

Παρότι το πειραματικό πρωτόκολλο (προεπεξεργασία, δειγματοληψία, optimizer, κ.λπ.) διατηρήθηκε κοινό, κάθε αρχιτεκτονική απαιτεί ορισμένες ειδικές υπερπαραμέτρους που καθορίζουν κυρίως τη χωρητικότητα (capacity) και τη δομή του encoder/decoder. Το dropout ορίστηκε 0.0 σε κάθε περίπτωση. Συνοπτικά, χρησιμοποιήθηκαν:

- **U-Net:** channels (16, 32, 64, 128, 256), strides (2, 2, 2, 2), num_res_units = 2.
- **Attention U-Net:** channels (16, 32, 64, 128, 256), strides (2, 2, 2, 2), kernel_size = 3, up_kernel_size = 3.
- **DynUNet:** kernel_size = [3, 3, 3, 3], strides = [1, 2, 2, 2], upsample_kernel_size = [2, 2, 2].
- **UNETR:** img_size (96, 96, 64), feature_size = 32, hidden_size = 768, mlp_dim = 3072, num_heads = 24, proj_type="conv", res_block=True.
- **SwinUNETR:** εκτελέστηκαν δύο ρυθμίσεις:
 - SwinUNETR Heavier: img_size (96, 96, 64), feature_size = 60, patch_size = 2, window_size = 5, depths (2, 2, 2, 2), use_checkpoint=True.
 - SwinUNETR Lighter: img_size (96, 96, 64), feature_size = 48, patch_size = 2, window_size = 4, depths (2, 2, 2, 2), use_checkpoint=True.
- **SegResNet:** βασικές παράμετροι όπως προαρμογή πλάτους επιπέδων (init_filters) και η κατανομή blocks στο down/up path. Εκτελέστηκαν δύο ρυθμίσεις:
 - SegResNet Heavier: init_filters= 64, blocks_down= [1, 2, 2, 4], blocks_up= [1, 1, 1].
 - SegResNet Lighter: init_filters= 32, blocks_down= [1, 2, 2, 4], blocks_up= [1, 1, 1].
- **SegMamba:** βασικές παράμετροι το 'πλάτος' χαρακτηριστικών (feat_size) και το βάθος (depths). Εκτελέστηκαν δύο ρυθμίσεις:
 - SegMamba Heavier: feat_size= 32, depths= [2, 2, 2, 2, 2].
 - SegMamba Lighter: feat_size= 32, depths= [2, 2, 2, 2].

Οι παραπάνω ρυθμίσεις αφορούν αποκλειστικά τη **χωρητικότητα της αρχιτεκτονικής** και όχι το εκπαιδευτικό πρωτόκολλο, το οποίο παρέμεινε κοινό για όλα τα μοντέλα.

Για κάθε αρχιτεκτονική εκτελέστηκε μία βασική εκπαίδευση (baseline run) με το κοινό pipeline. Επιπλέον πραγματοποιήθηκαν πειράματα εναλλαγής χωρητικότητας για επιλεγμένα μοντέλα (SegResNet, SegMamba και SwinUNETR), με στόχο να διερευνηθεί η επίδραση ρυθμίσεων που αυξάνουν την ικανότητα του δικτύου ή/και βελτιώνουν τη σταθερότητα σύγκλισης.

Το ενισχυμένο πείραμα του SegResNet είχε μεγαλύτερο πλάτος το δίκτυο, ενώ του SegMamba μεγαλύτερο βάθος (περισσότερα layers). Για το SwinUNETR συγκρίθηκαν η αρχική ρύθμιση (Heavier) και μία ελαφρύτερη παραλλαγή (Lighter) με μικρότερα feature_size και window_size. Οι αποκλίσεις αυτές από τη βασική ρύθμιση καταγράφονται ρητά στις αντίστοιχες υποενότητες αποτελεσμάτων ([6.2 - 6.4](#), [6.6](#)).

5.8 Συνάρτηση κόστους και βελτιστοποίηση

5.8.1 Κριτήριο εκπαίδευσης

Ως βασικό κριτήριο χρησιμοποιήθηκε η DiceCELoss (Dice + Cross-Entropy), με sigmoid έξodo και include_background=False. Ο συνδυασμός Dice και CE είναι πρακτικός σε έντονα ανισόρροπες τμηματοποιήσεις, καθώς ο Dice στοχεύει άμεσα στην επικάλυψη, ενώ το CE συνεισφέρει σταθερότητα και καλύτερη συμπεριφορά γραδιεντς σε πρώιμα στάδια σύγκλισης.

5.9 Επικύρωση και διαδικασία inference

5.9.1 Sliding Window Inference (SWI)

Η επικύρωση σε πλήρεις 3D όγκους είναι χρονοβόρα αλλά και ακριβή σε μνήμη, ειδικά για βαριά μοντέλα. Για τον λόγο αυτό εφαρμόστηκε sliding-window inference με επικάλυψη (overlap) και gaussian blending, ώστε να συντίθεται τελική πρόβλεψη σε όλο τον όγκο χωρίς να θυσιάζεται η ποιότητα στα όρια των patches.

5.9.2 EMA (Exponential Moving Average)

Παράλληλα, χρησιμοποιήθηκε EMA, δηλαδή εκθετικός κινητός μέσος των παραμέτρων του μοντέλου. Η αξιολόγηση με EMA τείνει να είναι πιο σταθερή από το στιγμιαίο checkpoint, ειδικά όταν ο ρυθμός μάθησης κυμαίνεται σε εύρος.

5.10 Βελτιστοποίηση και Scheduler

5.10.1 Optimizer

Για όλες τις υλοποιήσεις χρησιμοποιήθηκε ο AdamW με τις ίδιες ρυθμίσεις.

Στα πειράματα τέθηκαν base_lr = 2.5e-5 και weight_decay = 2e-5. Η τιμή lr=base_lr λειτουργεί ως βασικός ρυθμός μάθησης, πάνω στον οποίο εφαρμόζεται ο scheduler της επόμενης υποενότητας.

Οι παράμετροι betas=(0.9, 0.999) αντιστοιχούν στους συντελεστές β_1 και β_2 του Adam-τύπου κινητού μέσου: η $\beta_1 = 0.9$ ελέγχει την εξομάλυνση της πρώτης ροπής (μέση κλίση, τύπου momentum), ενώ η $\beta_2 = 0.999$ την εξομάλυνση της δεύτερης ροπής (μέσο τετραγωνικών κλίσεων).

Οι συγκεκριμένες επιλογές ήταν σταθερές σε όλες τις αρχιτεκτονικές για δίκαιη σύγκριση. Η θεωρητική ανάλυση του AdamW δίνεται στην Ενότητα 3.5.

5.10.2 Scheduler

Για λόγους αποφυγής τοπικών ελαχίστων και άλλους, έγινε χρήση Scheduler ώστε ο ρυθμός μάθησης να (learning rate) μεταβαλλόταν διαρκώς κατά τη διάρκεια της εκπαίδευσης με CyclicLR [54] (σχήμα triangular2). Στο συγκεκριμένο σχήμα ο ρυθμός μάθησης αυξάνεται

γραμμικά από `base_lr` έως `max_lr` και στη συνέχεια μειώνεται πάλι, ενώ σε κάθε επόμενο κύκλο το εύρος ταλάντωσης μειώνεται, οδηγώντας σε πιο «ήπια» εξερεύνηση στα όψιμα στάδια.

Η επιλογή κυκλικού χρονοπρογραμματισμού εξυπηρετεί δύο στόχους: (i) μειώνει την ευαισθησία σε μία μοναδική επιλογή learning rate, και (ii) λειτουργεί ως μηχανισμός εξερεύνησης του χώρου λύσεων, κάτι που συχνά βοηθά σε πιο σταθερή σύγκλιση όταν συνδυάζεται με έντονες στοχαστικές διεργασίες (όπως data augmentation και patch sampling).

Σημειώνεται ότι στα διαγράμματα ιστορικού εκπαίδευσης αποτυπώνεται ο κυκλικός χαρακτήρας του LR, καθώς και η χρονική στιγμή στην οποία επιτυγχάνεται η καλύτερη επίδοση στο validation (π.χ. μέγιστο Dice), που χρησιμοποιείται και ως κριτήριο επιλογής checkpoint.

Κεφάλαιο 6

Αποτελέσματα Πειραματικής Αξιολόγησης

Στο κεφάλαιο αυτό παρουσιάζονται τα αποτελέσματα της συγκριτικής αξιολόγησης των αρχιτεκτονικών τμηματοποίησης με τη δομή που περιγράφηκε στο Κεφάλαιο 5. Όλα τα μοντέλα εκπαιδεύτηκαν και αξιολογήθηκαν με **ενιαίο** πειραματικό πρωτόκολλο (προεπεξεργασία, δειγματοληψία, βελτιστοποίηση, επικύρωση), ώστε οι διαφορές στο τέλος να αποδίδονται όσο το δυνατόν περισσότερο στην αρχιτεκτονική και όχι σε ασυνεπείς ρυθμίσεις ή διαδικασίες.

Πρώτα, παρατίθενται τα ποσοτικά αποτελέσματα στα δεδομένα αξιολόγησης μέσω μετρικών επικάλυψης (Dice, IoU) και της τιμής της απώλειας επικύρωσης (validation loss). Επειτα, συμπληρώνονται με ποιοτική αξιολόγηση μέσω οπτικών προβλέψεων έναντι ground truth (τμηματοποίηση παρεχόμενη από ειδικούς υγείας), και πραγματοποιείται και έλεγχος ως προς τη συμφωνία των ποσοτικών αποτελεσμάτων με τις εικόνες.

Τέλος, παρουσιάζονται και σχετικές δημοσιευμένες εργασίες με όμοιο στόχο, μαζί με τα τελικά τους αποτελέσματα, ώστε να γίνει μια ενδεικτική (μη αυστηρή) και πιο ελεύθερη αντιπαραβολή με τα ευρήματα της παρούσας εργασίας.

6.1 Παρουσίαση αποτελεσμάτων

Τα αποτελέσματα παρουσιάζονται σε δύο επίπεδα:

- **Ποσοτικά:** μετρικές Dice/IoU και απώλειες (train/val loss), με σύγκριση μεταξύ μοντέλων.
- **Ποιοτικά:** ενδεικτικές οπτικοποιήσεις προβλέψεων/σφαλμάτων σε αντιπροσωπευτικά περιστατικά.

Ο Πίνακας 6.1 συνοψίζει τη **καλύτερη** επίδοση κάθε μοντέλου στο validation set, όπως προέκυψε σε επίπεδο 3D όγκου σε 120 epochs. Οι τιμές Dice χρησιμοποιούνται ως σημείο αναφοράς για την αναλυτικότερη ποιοτική και συγκριτική συζήτηση που ακολουθεί και αντιστοιχούν στη μέγιστη τιμή του mean Dice που επιτεύχθηκε σε κάποιο epoch εντός της εκπαίδευσης.

Σημειώνεται ότι οι μετρικές επικάλυψης αποτυπώνουν κυρίως τη συμφωνία ως προς το εμβαδό/όγκο της πρόβλεψης και δεν εγγυώνται πάντα οπτικά σωστά όρια σε όλες τις περιπτώσεις.

Πίνακας 6.1: Αποτελέσματα ανά αρχιτεκτονική στο validation set.

Μοντέλο	Dice	IoU	Val Loss
U-Net	0.823	0.7002	0.2177
Attention U-Net	0.846	0.735	0.2004
DynUNet	0.846	0.7342	0.1897
UNETR	0.772	0.6345	0.2842
SwinUNETR Heavier	0.8497	0.7401	0.1838
SwinUNETR Lighter	0.8490	0.7387	0.1838
SegResNet Heavier	0.8601	0.7558	0.1678
SegResNet Lighter	0.859	0.7544	0.1806
SegMamba Heavier	0.8606	0.7566	0.1685
SegMamba Lighter	0.8581	0.7526	0.1697

6.2 Συνοπτική κατάταξη και κύρια ευρήματα

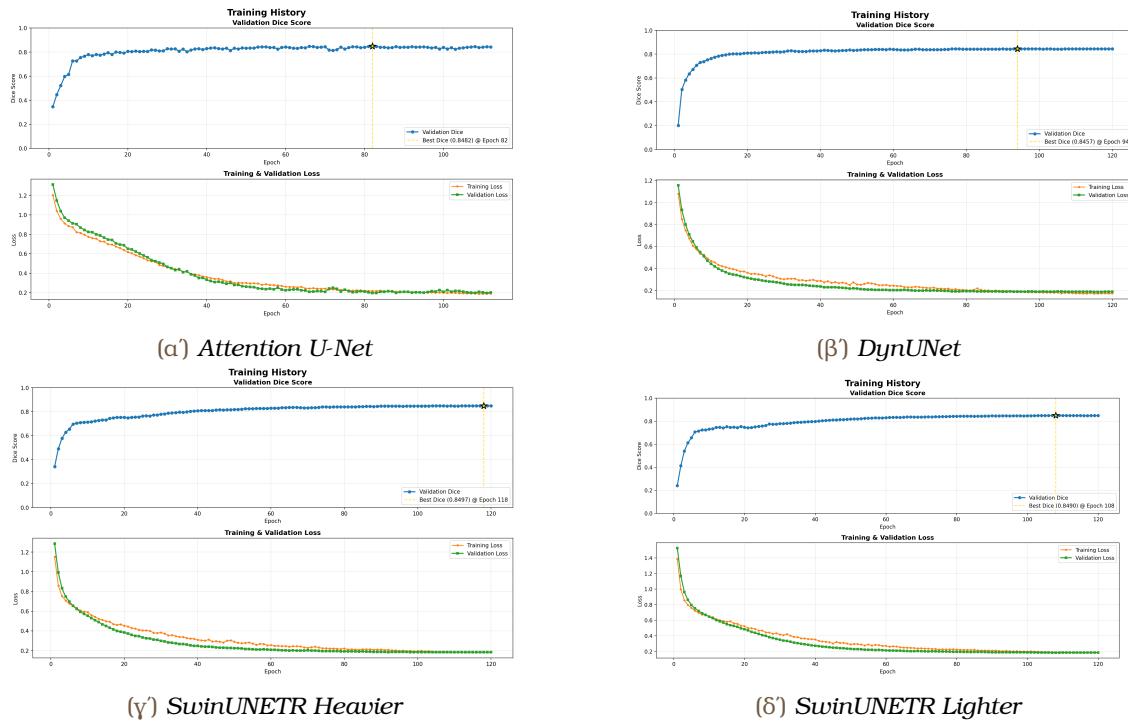
Θεκινώντας από τον Πίνακα 6.1, η κατάταξη των μοντέλων με βάση τη μέγιστη μέση τιμή του Dice δείχνει ότι η καλύτερη επίδοση επιτυγχάνεται από το SegMamba Heavier (Dice= 0.8606) και έπειτα τα SegResNet Heavier (Dice= 0.8601) και SegResNet Lighter (Dice= 0.859). Οι διαφορές μέσα στα προαναφερθέντα είναι της τάξης του $\approx 0.002\text{--}0.007$ σε Dice.

Η εικόνα αυτή επιβεβαιώνεται και από τη συμπληρωματική μετρική IoU: τα υψηλότερα IoU παρατηρούνται επίσης στα SegMamba Heavier (0.7566) και SegResNet Heavier (0.7558), με τα υπόλοιπα μοντέλα της κορυφής να ακολουθούν πολύ κοντά. Παράλληλα, οι τιμές validation loss είναι χαμηλότερες για τα μοντέλα με καλύτερο Dice (π.χ. SegResNet Heavier: 0.1678, SegMamba Heavier: 0.1685), γεγονός που είναι συμβατό με καλύτερη συνολική βελτιστοποίηση. Ωστόσο, επειδή το loss αποτελεί σύνθετο κριτήριο και δεν αντιστοιχεί μονοσήμαντα σε μια overlap μετρική, η τελική κατάταξη στηρίζεται πρωτίστως στο Dice (και υποστηρίζεται από IoU).

Σε επίπεδο οικογενειών αρχιτεκτονικής, παρατηρείται ότι οι αποδοτικές CNN-based προσεγγίσεις (SegResNet) και τα State Space-based μοντέλα (SegMamba) υπερέχουν συνολικά των Transformer-based αρχιτεκτονικών σε αυτή τη πειραματική αξιολόγηση. Ειδικότερα, το UNETR εμφανίζει τη χαμηλότερη επίδοση (Dice= 0.772, IoU= 0.6345), αποτελώντας σαφή outlier. Αντιθέτως, και οι δύο εκτελέσεις του SwinUNETR αποδίδουν εξαιρετικά: SwinUNETR Heavier (Dice= 0.8497, IoU= 0.7401) και SwinUNETR Lighter (Dice= 0.8490, IoU= 0.7387), με τη Lighter να φτάνει Val Loss= 0.1838 στο epoch 108. Οι τιμές αυτές απέχουν κατά περίπου 0.01 σε Dice από την κορυφή.

Αξίζει να σχολιασουμε το U-Net (ο πρωτοπόρος του αντικειμένου) απέδωσε αρκετά καλά με Dice Score= 0.823, συμφωνώντας με την τυπική απόδοση σε σχετικές εργασίες όπως θα σχολιαστεί στη συνέχεια.

Τέλος, οι παραλλαγές χωρητικότητας (SegResNet Heavier έναντι SegResNet Lighter, SegMamba Heavier έναντι SegMamba Lighter και SwinUNETR Heavier έναντι SwinUNETR Lighter) δείχνουν ότι η «βαρύτερη» ρύθμιση προσφέρει αμελητέο πλεονέκτημα (π.χ. SegResNet Heavier: 0.8601 έναντι 0.859), συνεπώς θα πρέπει να αναλογιστεί η επιβάρυνση στο υπολογιστικό κόστος που επιφέρει η επιπρόσθετη χωρητικότητα, με την απόδοση που



Σχήμα 6.1: Καμπύλες εκπαίδευσης: *Validation Dice* (πάνω) και *training/validation loss* (κάτω) για *Attention U-Net*, *DynUNet*, *SwinUNETR Heavier* και *SwinUNETR Lighter*. Το αστέρι και η κατακόρυφη διακεκομμένη γραμμή δείχνουν το epoch του καλύτερου Dice.

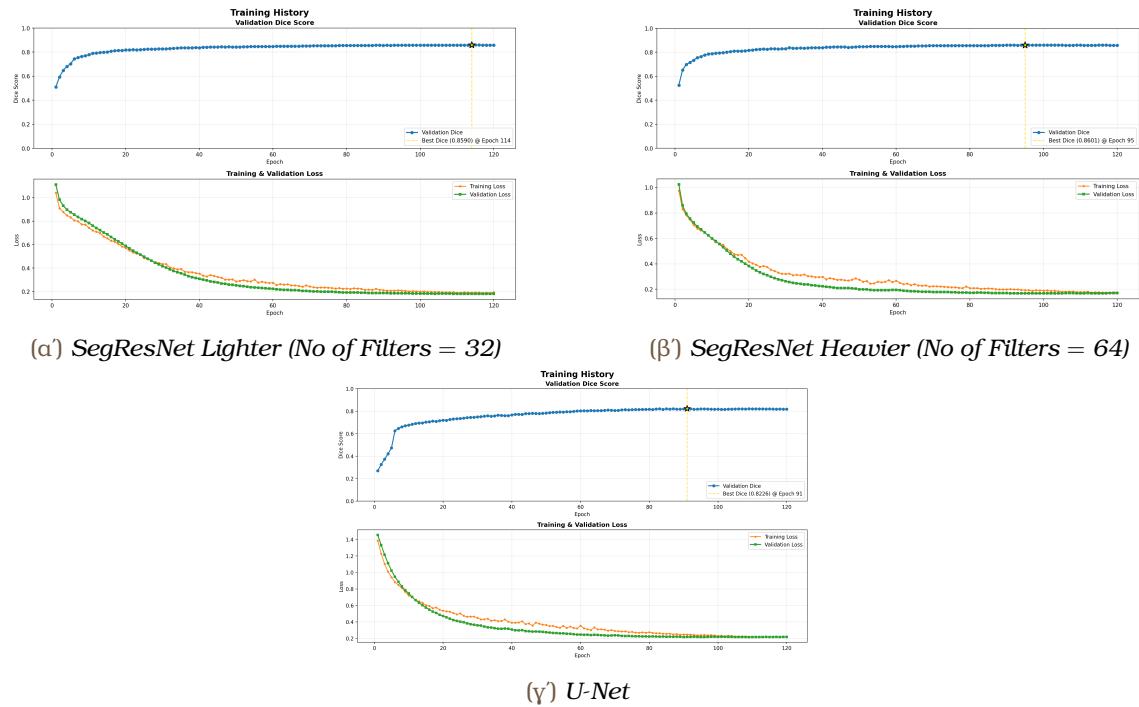
προσφέρεται λόγω αυτής.

6.3 Καμπύλες εκπαίδευσης και σταθερότητα σύγκλισης

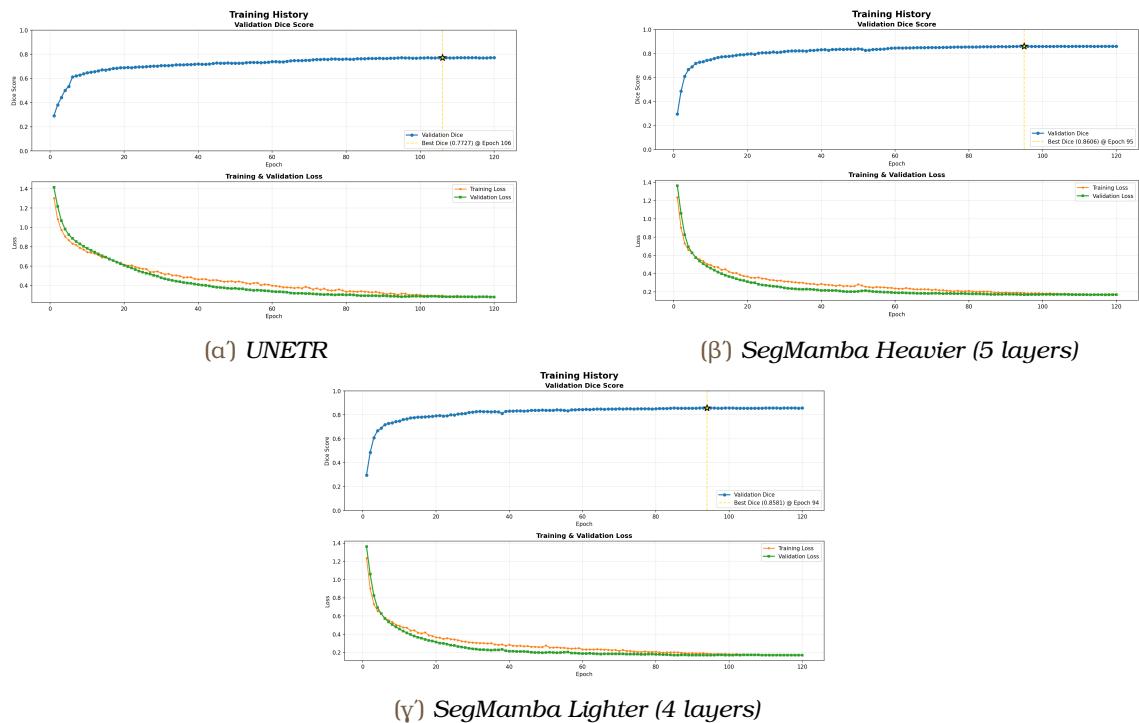
Στα σχήματα 6.1, 6.2 και 6.3 παρουσιάζονται οι καμπύλες εκπαίδευσης για κάθε αρχιτεκτονική. Σε κάθε περίπτωση, πρώτη καμπύλη παρουσιάζει τη μετρική Dice και η δεύτερη τις απώλειες training και validation κατά τη διάρκεια των epochs.

Γενική εικόνα σύγκλισης. Σε όλες τις CNN-βασισμένες αρχιτεκτονικές παρατηρείται **ταχεία αρχική βελτίωση** του validation Dice στα πρώτα 5-15 epochs, συγκλίνοντας σχετικά νωρίς, ακολουθούμενη από σταδιακή βελτίωση μέχρι το τέλος της εκπαίδευσης. Αυτό υποδηλώνει ότι το μοντέλο μαθαίνει γρήγορα τη χονδρική μορφολογία του πλακούντα, ενώ οι επιπλέον εποχές συμβάλλουν κυρίως σε λεπτότερη προσαρμογή (όρια, μικρές ασυνέχειες, μείωση false positives).

Συγκριτική ερμηνεία καμπυλών. Όλες οι προσωμοιώσεις εμφανίζουν απότομη άνοδο του validation Dice στα πρώτα ~10-20 epochs και στη συνέχεια σταθεροποίηση (plateau). Επίσης, όλες οι καμπύλες είναι σχεδόν πανομοιότυπες σε μορφή, κάτι που αναμένοταν δεδομένου ότι το πειραματικό πλαίσιο ήταν κοινό για όλες τις προσωμοιώσεις. Ομοίως, οι παραλλαγές SegMamba συγκλίνουν γρήγορα σε υψηλές τιμές Dice και παραμένουν σταθερά στην κορυφή.



Σχήμα 6.2: Καμπύλες εκπαίδευσης: Validation Dice (πάνω) και training/validation loss (κάτω) για SegResNet Lighter, SegResNet Heavier και U-Net. Το αστέρι και η κατακόρυφη διακεκομένη γραμμή δείχνουν το epoch του καλύτερου Dice.



Σχήμα 6.3: Καμπύλες εκπαίδευσης: Validation Dice (πάνω) και training/validation loss (κάτω) για UNETR, SegMamba Heavier και SegMamba Lighter. Το αστέρι και η κατακόρυφη διακεκομένη γραμμή δείχνουν το epoch του καλύτερου Dice.

Transformer-based συμπεριφορά. Το UNETR λειτουργεί ως outlier, επιτυγχάνοντας χαμηλότερο best Dice, γεγονός που συμφωνεί με τη συνολική του υστέρηση στον Πίνακα 6.1. Αντίθετα, και οι δύο διαμορφώσεις SwinUNETR Heavier/Lighter παρουσιάζουν πιο ομαλή και σταδιακή βελτίωση και, ενώ δεν ξεπερνούν την κορυφή (SegResNet/SegMamba), αποδίδουν καλύτερα από το απλό UNETR. Η Lighter εκδοχή φτάνει Dice= 0.8490, IoU= 0.7387 και Val Loss= 0.1838 στο epoch 108, πολύ κοντά στη Heavier. Το αποτέλεσμα είναι ενδεικτικό ότι η ιεραρχική προσοχή με τοπικά παράθυρα του Swin είναι πιο κατάλληλη από έναν ViT-style encoder [8] σε προβλήματα με όμοιες συνθήκες.

Σταθερότητα και γενίκευση. Οι καμπύλες training και validation loss μειώνονται παράλληλα και δεν παρατηρείται έντονη απόκλιση που να υποδεικνύει ισχυρή υπερπροσαρμογή στο εξεταζόμενο εύρος epochs. Το καλύτερο checkpoint ορίζεται ως το epoch μέγιστου validation Dice (κατακόρυφη γραμμή στα γραφήματα), το οποίο εμφανίζεται συνήθως σε μεταγενέστερο epoch, ενώ οι μεγαλύτερες βελτιώσεις έχουν ήδη πραγματοποιηθεί νωρίτερα.

Σύνδεση με ποιοτική αξιολόγηση. Στην επόμενη ενότητα εξετάζονται οπτικά παραδείγματα προβλέψεων και τυπικών σφαλμάτων, ώστε να αξιολογηθεί κατά πόσο οι διαφορές στο Dice/IoU αντιστοιχούν σε ουσιαστικές διαφορές στην ποιότητα ορίων.

6.4 Ποιοτική αξιολόγηση προβλέψεων

Η ποιοτική αξιολόγηση συμπληρώνει τις ποσοτικές μετρικές του Πίνακα 6.1, καθώς οι μετρικές επικάλυψης (Dice/IoU) αποτυπώνουν κυρίως τη συμφωνία ως προς τον όγκο και δεν εγγυώνται πάντα αντίστοιχη ακρίβεια στα όρια ή στα λεπτά τμήματα της δομής. Στα σχήματα 6.1 έως 6.4 που ακολουθούν παρουσιάζονται **ενδεικτικές τομές** από δύο περιστατικά (Sample 1/2) και η σύγκριση: MRI slice (αριστερά), ground truth μάσκα (κέντρο) και η πρόβλεψη που επετεύχθη από το μοντέλο (δεξιά). Ο πλακούντας και η πρόβλεψη απεικονίζονται με πράσινο και το υπόβαθρο με κόκκινο.

Κατά την οπική επιθεώρηση εξετάζονται κυρίως:

- **Υπο- ή υπερ-τμηματοποίηση,**
- **συνέχεια/συνοχή** της μάσκας,
- **ψευδοθετικά** (false positives) ως απομονωμένες περιοχές,
- **ποιότητα ορίων** (ομαλότητα και μορφολογική συνέπεια).

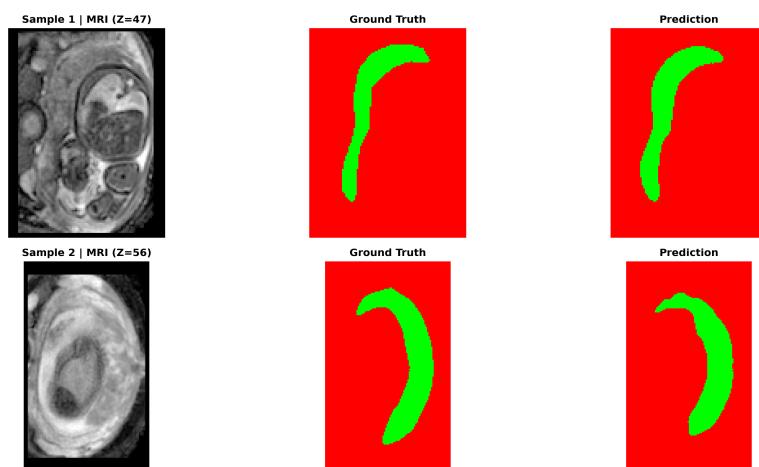
Συμφωνία ποιοτικών και ποσοτικών αποτελεσμάτων. Σε γενικές γραμμές, η οπική εικόνα επιβεβαιώνει την κατάταξη της Ενότητας 6.2. Τα μοντέλα με την υψηλότερη επίδοση σε Dice/IoU παράγουν πιο συνεπείς και ‘καθαρές’ μάσκες, με όρια που ακολουθούν καλύτερα τη μορφολογία του πλακούντα και με ελάχιστα ψευδοθετικά. Αντίθετα, το UNETR, που είναι ποσοτικά ο ασθενέστερος outlier, εμφανίζει ορατές αστοχίες (ανομοιόμορφα περιγράμματα, τμηματοποιήσεις εκτός στόχου και μικρές απομονωμένες περιοχές), οι οποίες εξηγούν τη χαμηλή του IoU και το υψηλότερο validation loss.

CNN-based μοντέλα. Τα U-Net, Attention U-Net και DynUNet αποδίδουν γενικά σωστή γεωμετρία και συνεχείς μάσκες, με αποκλίσεις που εμφανίζονται κυρίως στα άκρα και σε λεπτές περιοχές (ελαφρά υπο- ή υπερ-τμηματοποίηση). Η Attention U-Net και το DynUNet δείχνουν πιο σταθερή συμπεριφορά από το βασικό U-Net, συμβατό με το μικρό αλλά μετρήσιμο κέρδος στις μετρικές ($\text{Dice} \approx 0.846$ έναντι 0.829). Παράλληλα, το U-Net παρότι παρουσιάζει αποδεκτή επικάλυψη, εμφανίζει μεγαλύτερο validation loss (Πίνακας 6.1), κάτι που είναι ενδεικτικό ότι το loss ($\text{Dice} + \text{CE}$) είναι πιο ευαίσθητο σε αθεβαιότητα/λάθη κοντά στα όρια, ακόμη και όταν η τελική δυαδική μάσκα (μετά το thresholding) είναι οπτικά ικανοποιητική.

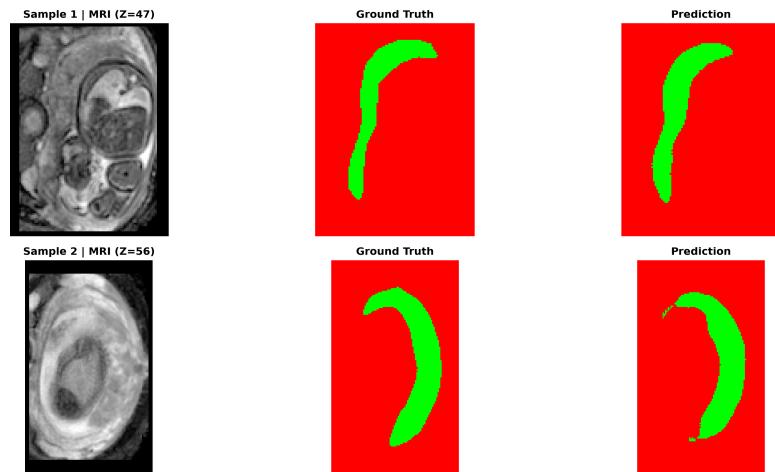
Transformer-based μοντέλα. Τα SwinUNETR Heavier/Lighter παράγουν συνεπείς προβλέψεις με καλό περιορισμό των ψευδοθετικών, κάτι που συμφωνεί με το ότι αποτελούν τις καλύτερες Transformer-based διαμορφώσεις της σύγκρισης ($\text{Dice} = 0.8497/0.8490$, $\text{IoU} = 0.7401/0.7387$). Η διαφορά τους είναι πολύ μικρή, με τη Lighter εκδοχή να διατηρεί $\text{Val Loss} = 0.1838$ στο epoch 108. Ωστόσο, σε ορισμένα σημεία η πρόβλεψη είναι πιο ‘συντηρητική’ (ελαφρά υπο-τμηματοποίηση σε λεπτές προεξοχές), γεγονός που μπορεί να συμβάλλει στη διαφορά ~ 0.01 σε Dice από την κορυφή. Αντίθετα, το UNETR εμφανίζει πιο ασταθή συμπεριφορά, με εμφανή false positives και ανομοιόμορφα περιγράμματα, τα οποία υποβαθμίζουν τόσο την IoU όσο και την αντιληπτή ποιότητα ορίων.

Μοντέλα κορυφής: SegResNet και SegMamba. Τα SegResNet και SegMamba δίνουν, στις ενδεικτικές περιπτώσεις, προβλέψεις πολύ κοντά στο ground truth: η μάσκα είναι συνεχής (χωρίς ‘τρύπες’) και τα όρια ακολουθούν πιο πιστά τη μορφολογία. Οι διαφορές μεταξύ ‘θαριάς’ και ‘ελαφριάς’ παραλλαγής είναι μικρές και συμφωνούν με τις αντίστοιχα μικρές διαφορές στις μετρικές (Dice στην τρίτη δεκαδική).

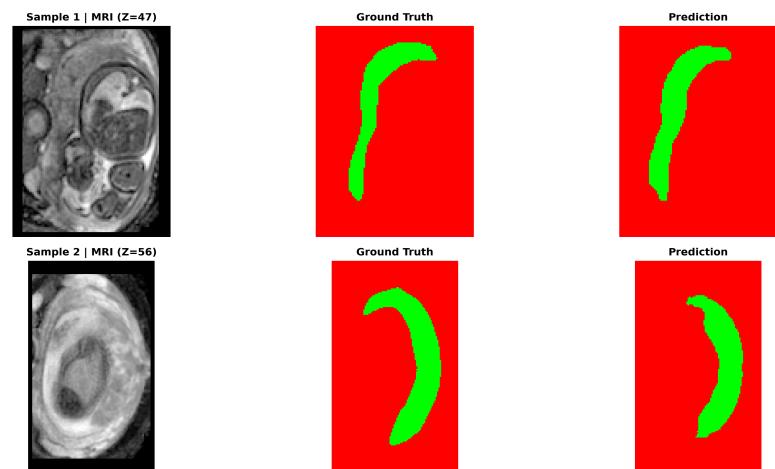
Περιορισμοί της οπτικής σύγκρισης. Οι απεικονίσεις αφορούν μεμονωμένες 2D τομές, ενώ οι μετρικές υπολογίζονται σε 3D όγκους. Επομένως, μικρές τοπικές αποκλίσεις μπορεί να μην είναι αντιπροσωπευτικές για το σύνολο του όγκου, αλλά παραμένουν χρήσιμες για την κατανόηση τυπικών σφαλμάτων και την ερμηνεία των ποσοτικών δεικτών.



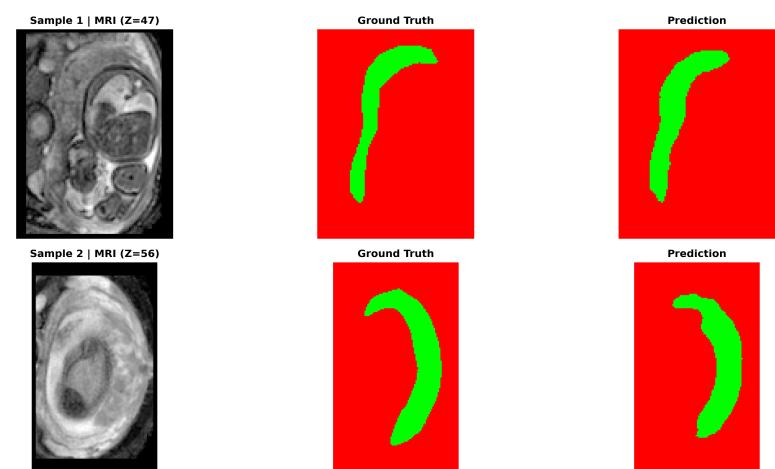
Εικόνα 6.1: *Attention U-Net*



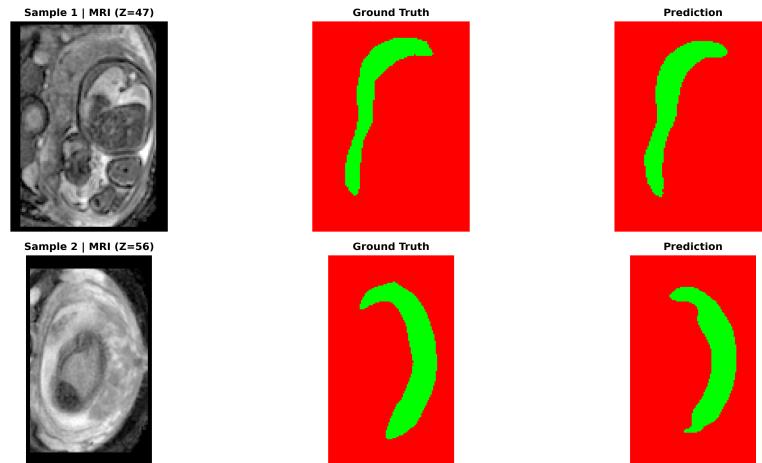
Εικόνα 6.2: *DynUNet*



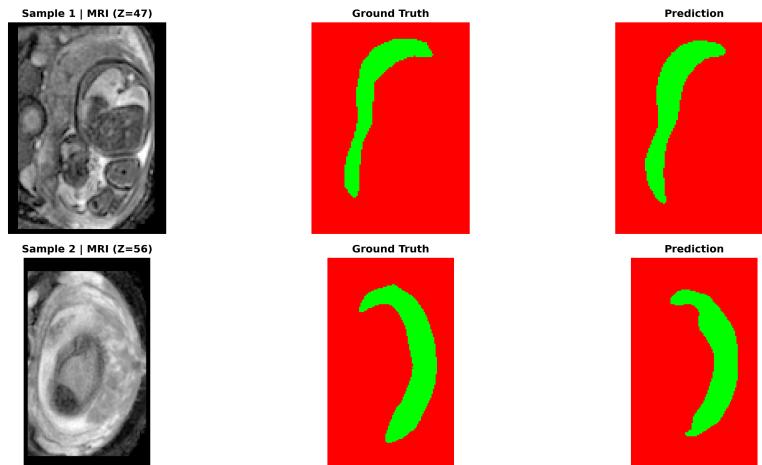
Εικόνα 6.3: *SwinUNETR Heavier*



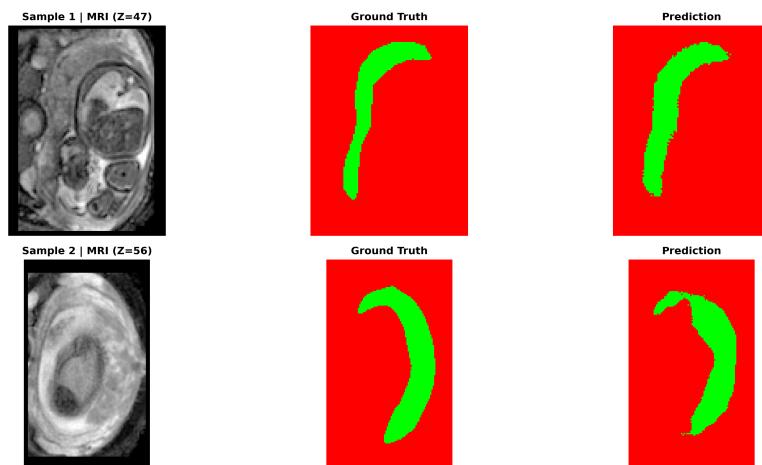
Εικόνα 6.4: *SwinUNETR Lighter*



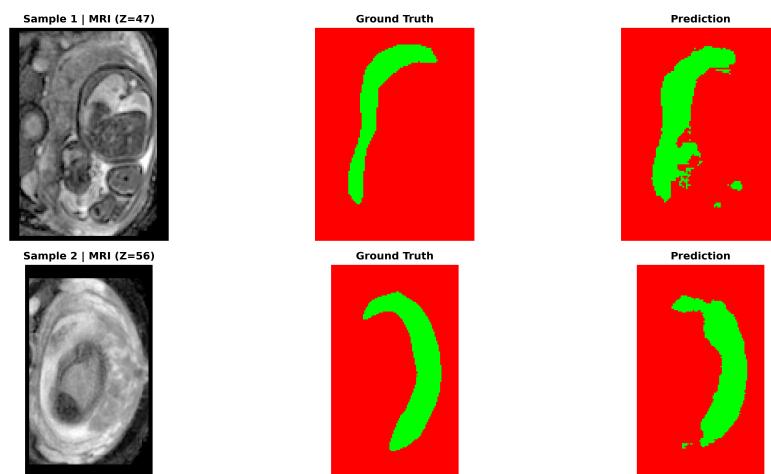
Εικόνα 6.5: *SegResNet Heavier (No of Filters = 64)*



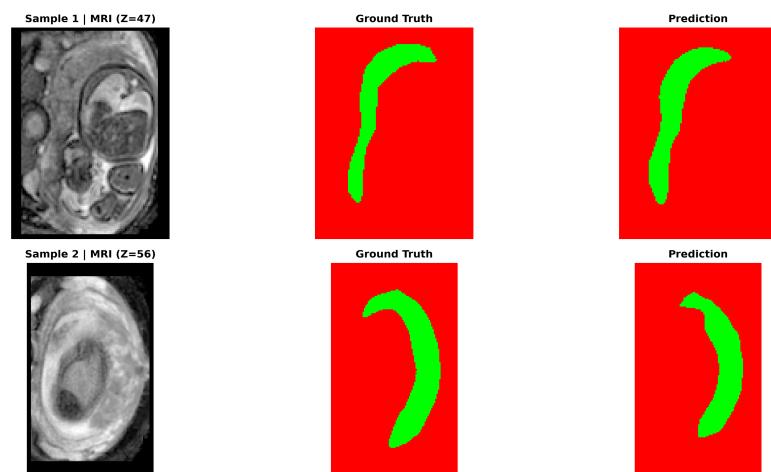
Εικόνα 6.6: *SegResNet Lighter (No of Filters = 32)*



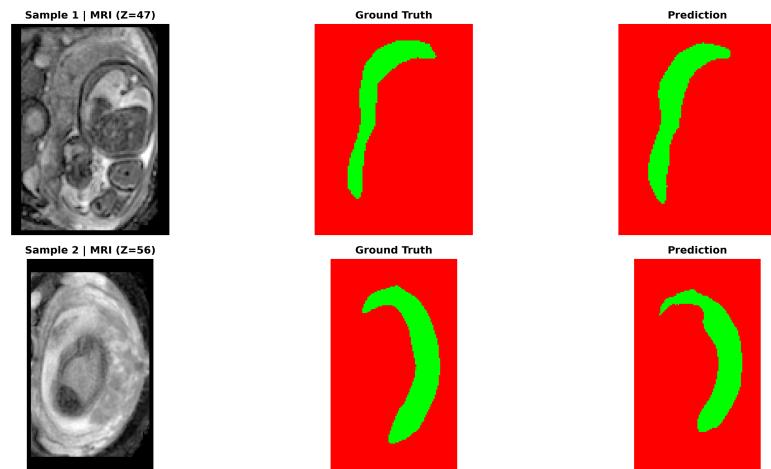
Εικόνα 6.7: *U-Net*



Εικόνα 6.8: *UNETR*



Εικόνα 6.9: *SegMamba Heavier (5 layers)*



Εικόνα 6.10: *SegMamba Lighter (4 layers)*

6.5 Σύγκριση με δημοσιευμένα αποτελέσματα

Για να τοποθετηθούν τα ευρήματα της παρούσας εργασίας σε ευρύτερο πλαίσιο, παρατίθεται στον Πίνακα 6.2 συνοπτική σύγκριση με σχετικές μελέτες τμηματοποίησης πλακούντα σε MRI [4, 24, 25, 55, 56, 57].

Πίνακας 6.2: Σύνοψη αποτελεσμάτων από σχετικές μελέτες τμηματοποίησης πλακούντα σε MRI.

Μελέτη	Μοντέλο	Δεδομένα / split	Μετρική πλακούντα	Σχόλιο
Shahedi et al. 2022 [24]	CNN-based	181 total (train=157, val=24) + test=60	DSC=0.80	Μητρική κοιλότητα: DSC=0.92.
Shahedi et al. 2021 [4]	Δύο CNN-based	70 total, έλεγχος σε 20 (normal) και 50 (normal+PAS)	DSC=0.82 (normal), 0.83 (+PAS)	Μητρική κοιλότητα: DSC=0.92 και 0.88 αντίστοιχα.
Liu et al. 2023 [25]	SADL	154 total, train=108, val=15, test=31	DSC=0.83±0.06, 0.84±0.05	UNet achieved: (0.77, 0.76).
Huang et al. 2023 [55]	CNN-based	241 (axial), και 101 (sagittal) MRI	DSC=0.82 (axial), 0.82 (sagittal)	Μητρική κοιλότητα: DSC=0.87 (axial), 0.92 (sagittal).
Lee et al. 2023 [56]	MRF-MAS vs UNet	390 total, train=312, test=78	Dice=0.8992 (MRF-MAS), U-Net Dice=0.8632	Αναφορά και IoU=0.8169 για το πρωτεινόμενο μοντέλο.
Saito et al. 2025 [57]	PlaNet-S (UNet+SegNeXt)	1090 total, 875 train-215 test	IoU=0.78±0.10 (PlaNet-S)	U-Net IoU=0.73±0.13, DS-transUNet IoU=0.64±0.16.
Παρούσα εργασία	UNet	137 περιπτώσεις, train=109, val=28	Dice=0.829, IoU=0.7109	Ενιαίο πρωτόκολλο σε 10 διαμορφώσεις μοντέλων και αξιολόγηση σε 3D όγκους.
Παρούσα εργασία	SegMamba Heavier		Dice=0.8606, IoU=0.7566	Απέδωσε καλύτερα από τις άλλες διαμορφώσεις

Η παραπάνω σύγκριση είναι **ενδεικτική** και όχι αυστηρά αντικειμενική, επειδή τα πειράματα δεν έχουν εκτελεστεί σε κοινό σύνολο δεδομένων ή ενιαίο πρωτόκολλο. Παρά τα παραπάνω, οι τιμές επικάλυψης της παρούσας εργασίας (Dice= 0.8606, IoU= 0.7566) κινούνται στο ίδιο εύρος με τη σύγχρονη βιβλιογραφία για τμηματοποίηση πλακούντα σε MRI, κάτι που υποστηρίζει τη συνολική εγκυρότητα της πειραματικής μεθοδολογίας.

6.6 Συζήτηση και ερμηνεία

Στην ενότητα αυτή συντίθενται τα ποσοτικά ευρήματα (Πίνακας 6.1 και καμπύλες εκπαίδευσης) με την ποιοτική επιθεώρηση (Ενότητα 6.4), με στόχο να δοθεί συνεκτική ερμηνεία των διαφορών μεταξύ των αρχιτεκτονικών και να αποσαφηνιστεί ποια συμπεράσματα είναι σταθερά, αλλά και ποιοι παράγοντες επηρεάζουν την αξιοπιστία/γενίκευση των αποτελεσμάτων.

Σύνοψη επιδόσεων και κλίμακα διαφορών. Η συνολική εικόνα δείχνει ότι οι καλύτερες επιδόσεις στο validation set συγκεντρώνονται σε μια ‘ομάδα κορυφής’ γύρω από Dice 0.86: SegMamba Heavier (Dice= 0.8606) και SegResNet Heavier/Lighter (Dice= 0.8601/0.859), με πολύ μικρές μεταξύ τους διαφορές (περίπου 0.002–0.003). Το SwinUNETR Heavier/-Lighter (Dice= 0.8497/0.8490) είναι ανταγωνιστικό, αλλά παραμένει χαμηλότερα από την κορυφή κατά περίπου 0.01 σε Dice, αλλά κάτα 0.02 παραπάνω από το base UNet, ενώ το UNETR αποτελεί σαφή outlier με αισθητά χαμηλότερη επικάλυψη (Dice= 0.772). Η μικρή απόσταση εντός της κορυφής υποδηλώνει ότι, για το συγκεκριμένο σύνολο δεδομένων και πρωτόκολλο, η επίδοση τείνει να κορεστεί: περαιτέρω βελτιώσεις αναμένονται να είναι

οριακές και πιθανόν να εξαρτώνται περισσότερο από λεπτομέρειες pre-processing, δειγματοληψίας patches ή ποιότητας επισημειώσεων, παρά από ριζικά διαφορετική αρχιτεκτονική.

Ερμηνεία υπεροχής CNN/SSM προσεγγίσεων. Τα αποτελέσματα υποδεικνύουν ότι οι CNN-βασισμένες λύσεις (SegResNet, DynUNet, Attention U-Net) και τα SSM-based μοντέλα (SegMamba) είναι ιδιαίτερα κατάλληλα για τμηματοποίηση πλακούντα σε 3D MRI. Μια εύλογη ερμηνεία είναι ότι το πρόβλημα κυριαρχείται από τοπική μορφολογία και συνέπεια ορίων, όπου ο inductive bias των συνελίξεων και η αποδοτική μοντελοποίηση εξαρτήσεων των SSM αρκούν για να ανακτηθεί το απαραίτητο context, χωρίς self-attention σε όλο τον όγκο. Η ποιοτική αξιολόγηση συμφωνεί με το παραπάνω: τα μοντέλα κορυφής εμφανίζουν πιο συνεκτικές μάσκες και περιορισμένα false positives, με όρια που ακολουθούν καλύτερα τη μορφολογία του ground truth.

Σχόλια για Transformer μοντέλα: UNETR έναντι SwinUNETR. Η αισθητά χαμηλότερη επίδοση του UNETR είναι συμβατή με τη γενικότερη παρατήρηση ότι ViT-τύπου encoders τείνουν να απαιτούν μεγαλύτερο όγκο δεδομένων, προσεκτικότερο regularization και/ή πιο πλούσια προεκπαίδευση, ώστε να αποδώσουν ανταγωνιστικά σε ιατρική τμηματοποίηση. Αντίθετα, το SwinUNETR εμφανίζει πιο σταθερή και καλύτερη συμπεριφορά, πιθανώς επειδή ενσωματώνει ισχυρότερη τοπικότητα (ιεραρχικά windows και patch merging), κάτι που λειτουργεί ως πιο κατάλληλος prior για 3D δεδομένα με περιορισμένο αριθμό περιπτώσεων. Οπτικά, αυτό αποτυπώνεται σε πιο ‘καθαρά’ όρια και σε περιορισμό των σποραδικών ψευδο-θετικών περιοχών, σε σύγκριση με το UNETR. Επιπλέον, οι δύο διαμορφώσεις SwinUNETR Heavier/Lighter έχουν πρακτικά ισοδύναμη επίδοση (Dice= 0.8497/0.8490), άρα η μείωση χωρητικότητας δεν επιφέρει ουσιαστική ποιοτική υποβάθμιση.

Σχέση Dice/IoU και validation loss. Παρότι το validation loss ακολουθεί γενικά την κατάταξη των μοντέλων, δεν είναι πλήρως μονοσήμαντο ως προς τις μετρικές επικάλυψης. Η περίπτωση του U-Net (ικανοποιητικό Dice αλλά υψηλότερο loss) υποδηλώνει ότι ο συνδυασμός Dice + CE επιτρέαζεται από τη βαθμονόμηση της εξόδου και από λάθη κοντά στα όρια που μπορεί να μην αλλάζουν δραστικά το τελικό overlap μετά το thresholding. Έτσι, οι μετρικές επικάλυψης πρέπει να ερμηνεύονται συμπληρωματικά με το loss και, ιδανικά, με boundary-ευαίσθητες μετρικές όταν το ζητούμενο είναι ακριβής εντοπισμός ορίων.

Ακρίβεια έναντι υπολογιστικού κόστους. Οι παραλλαγές χωρητικότητας (SegResNet Heavier έναντι SegResNet Lighter και SegMamba Heavier έναντι SegMamba Lighter και SwinUNETR Heavier έναντι SwinUNETR Lighter) προσφέρουν μόνο μικρό κέρδος στην επικάλυψη. Αυτό πρακτικά σημαίνει ότι, σε περιβάλλοντα με περιορισμένους πόρους (VRAM/χρόνος), μια ελαφρύτερη διαμόρφωση μπορεί να αποτελεί ελκυστική επιλογή, χωρίς σημαντική θυσία ακρίβειας. Επίσης, παρότι το SwinUNETR (και στις δύο εκδοχές) επιτυγχάνει ανταγωνιστική επίδοση (σε Dice) και βρίσκεται πολύ κοντά στα κορυφαία CNN-based μοντέλα, η επίτευξη αυτής της επίδοσης συνοδεύεται από σημαντικά αυξημένο υπολογιστικό κόστος. Η Transformer-based δομή του σε 3D patches και ιεραρχικό patch merging οδηγεί σε

μεγαλύτερο χρόνο άνα εποχή και αυξημένες απαιτήσεις μνήμης σε σύγκριση με αποδοτικότερες CNN επιλογές όπως το SegResNet ή ακόμα το DynUNet. Συνεπώς, το SwinUNETR αποτελεί ισχυρή επιλογή από πλευράς απόδοσης, αλλά έχοντας το υπολογιστικό κόστος υπόψιν, οι εναλλακτικές καθίστανται προτιμότερες.

6.6.1 Ενδεικτικοί χρόνοι ολοκλήρωσης εκτέλεσης notebooks

Για να αποτυπωθεί πρακτικά το κόστος εκτέλεσης, καταγράφηκε ο συνολικός wall-clock χρόνος μέχρι την ολοκλήρωση κάθε run στο cloud περιβάλλον εκπαίδευσης. Στον Πίνακα 6.3 παρατίθενται οι αντίστοιχοι χρόνοι, μαζί με τον επιπλέον χρόνο φόρτωσης Mamba για τις δύο εκτελέσεις SegMamba.

Πίνακας 6.3: Ενδεικτικοί χρόνοι ολοκλήρωσης ανά run.

Μοντέλο	Συνολικός χρόνος	Χρόνος φόρτωσης Mamba
SegMamba Lighter	05:07:21	01:14:47
SegMamba Heavier	05:01:31	01:13:02
U-Net	00:46:01	–
UNETR	02:53:41	–
SwinUNETR Heavier	08:42:00	–
SwinUNETR Lighter	06:00:00	–
SegResNet Lighter	04:28:34	–
SegResNet Heavier	09:46:43	–
DynUNet	02:52:00	–
Attention U-Net	01:37:13	–

Οι παραπάνω χρόνοι **δεν αποτελούν επίσημο benchmark** απόδοσης. Παρότι οι εκτελέσεις έγιναν με τον ίδιο τύπο GPU στο cloud (άρα χωρίς άμεση επίδραση του τοπικού δικτύου), υπάρχουν πολλοί παράγοντες που δεν είναι πάντα σταθεροί και επηρεάζουν τον συνολικό χρόνο. Επομένως, οι τιμές χρησιμοποιούνται ως **ενδεικτική αναφορά σχετικής ταχύτητας** και χρόνου ολοκλήρωσης των ίδιων εργασιών μεταξύ μοντέλων.

Η μεγάλη διαφορά στον χρόνο εκτέλεσης σε σχέση με την μικρή απόκλιση στην τελική εικόνα και ολική απόδοση μεταξύ των παραλλαγών του SegResNet αναδεικνύει το μικρό του κέρδος εις βάρος παραπάνω από διπλάσιου χρόνου εκτέλεσης. Σε μικρότερο βαθμό το ίδιο ισχύει και για τις παραλλαγές του SwinUNETR.

Περιορισμοί και παράγοντες γενίκευσης. Με τα παραπάνω συμπεράσματα πρέπει να μη ξεχνάμε και τα ακόλουθα:

- οι μετρικές προέρχονται από ένα validation split και όχι από ανεξάρτητο test set ή cross-validation,
- οι ποιοτικές απεικονίσεις βασίζονται σε μεμονωμένες 2D τομές, ενώ η αξιολόγηση είναι 3D,

Παρόλα αυτά, επειδή όλες οι αρχιτεκτονικές αξιολογούνται με κοινό πρωτόκολλο, τα συγκριτικά συμπεράσματα (η σχετική κατάταξη και τα τυπικά σφάλματα) είναι ισχυρότερα από την απόλυτη τιμή της επίδοσης.

Σύνοψη ενότητας. Η παρούσα ενότητα αποσαφήνισε τη σχετική συμπεριφορά των αρχιτεκτονικών, συνδυάζοντας ποσοτικά και ποιοτικά ευρήματα υπό κοινό πρωτόκολλο. Η τελική σύνθεση των συμπερασμάτων της εργασίας και η συστηματική διατύπωση των μελλοντικών επεκτάσεων παρουσιάζονται στο επόμενο κεφάλαιο.

Μέρος 

Επίλογος

Κεφάλαιο 7

Επίλογος

Στο κεφάλαιο αυτό συνοψίζονται τα βασικά συμπεράσματα της εργασίας και διατυπώνονται οι κύριες κατευθύνσεις για συνέχεια της μελέτης. Η αποτίμηση βασίζεται στο ενιαίο πρωτόκολλο του Κεφαλαίου 5 και στα αποτελέσματα του Κεφαλαίου 6.

7.1 Συμπεράσματα

Μεθοδολογική σύνοψη. Η εργασία ανέπτυξε ένα ενιαίο και αναπαραγώγιμο pipeline για τημηματοποίηση πλακούντα σε 3D MRI, με κοινή προεπεξεργασία, κοινή διαδικασία εκπαίδευσης/επικύρωσης και ενιαίο τρόπο αναφοράς μετρικών. Με τον τρόπο αυτό, οι διαφορές μεταξύ μοντέλων ερμηνεύονται κυρίως ως αρχιτεκτονικές διαφορές και όχι ως τεχνητές αποκλίσεις ρυθμίσεων.

Συγκριτική αξιολόγηση μοντέλων. Η συνοπτική κατάταξη της Ενότητας 6.2 και η αναλυτική συζήτηση της Ενότητας 6.6 οδηγούν στα εξής κύρια ευρήματα:

- Τα SegMamba Heavier και SegResNet Heavier πέτυχαν τις υψηλότερες τιμές επικάλυψης ($\text{Dice} \approx 0.86$), με πολύ μικρή μεταξύ τους διαφορά.
- Το SwinUNETR αποτέλεσε το πιο ανταγωνιστικό Transformer-βασισμένο μοντέλο, αλλά παρέμεινε χαμηλότερα από την ομάδα κορυφής.
- Το UNETR λειτούργησε ως σαφής outlier με αισθητά χαμηλότερη επίδοση σε Dice/IoU στο συγκεκριμένο split.
- Οι παραλλαγές μεγαλύτερης χωρητικότητας (SegResNet Heavier, SegMamba Heavier) έδωσαν μικρό αλλά μετρήσιμο όφελος έναντι των ελαφρύτερων εκδόσεων, χωρίς δραστική μεταβολή της συνολικής εικόνας.

Συνέπεια ποσοτικής και ποιοτικής ανάλυσης. Η ποιοτική αξιολόγηση της Ενότητας 6.4 ευθυγραμμίζεται με τις μετρικές Dice/IoU. Τα μοντέλα κορυφής παρήγαγαν πιο συνεκτικές μάσκες, καλύτερη μορφολογική συμφωνία και λιγότερα ψευδοθετικά, ενώ τα μοντέλα με χαμηλότερες μετρικές εμφάνισαν αστάθειες στα όρια και συχνότερα σφάλματα εκτός στόχου.

Σύνδεση με τη βιβλιογραφία. Η σύγκριση της Ενότητας 6.5 δείχνει ότι οι επιδόσεις της παρούσας μελέτης κινούνται σε ανταγωνιστικό εύρος σε σχέση με πρόσφατες εργασίες τμηματοποίησης πλακούντα. Παράλληλα, η χρήση ενιαίου πρωτοκόλλου ενισχύει τη σαφήνεια της σύγκρισης μεταξύ οικογενειών αρχιτεκτονικών.

Κλινική συνάφεια και πρακτική αξιοποίηση. Τα αποτελέσματα υποστηρίζουν ότι σύγχρονες CNN- και SSM-βασισμένες προσεγγίσεις μπορούν να προσφέρουν αξιόπιστη αυτόματη τμηματοποίηση πλακούντα, χρήσιμη για ποσοτική ανάλυση και συγκρίσιμες μετρήσεις σε MRI. Η σταθερότητα των κορυφαίων μοντέλων αποτελεί θετική ένδειξη για ενσωμάτωση σε research-oriented ροές εργασίας.

Περιορισμοί εγκυρότητας. Η ερμηνεία των ευρημάτων οφείλει να λάβει υπόψη ότι:

- η αξιολόγηση βασίζεται σε validation split και όχι σε ανεξάρτητο test set,
- δεν πραγματοποιήθηκε εξωτερική επικύρωση σε δεδομένα άλλου κέντρου ή διαφορετικού πρωτοκόλλου,
- οι διαφορές στην ομάδα κορυφής είναι μικρές (τρίτη δεκαδική του Dice) και δεν πρέπει να υπερερμηνεύονται.

Τελική αποτίμηση. Ο βασικός στόχος της διπλωματικής, δηλαδή η τεκμηριωμένη και δίκαιη συγκριτική αξιολόγηση ετερογενών αρχιτεκτονικών για τμηματοποίηση πλακούντα σε 3D MRI, επιτεύχθηκε. Με τα παρόντα δεδομένα και υπό το ίδιο πρωτόκολλο εκπαίδευσης, τα SegMamba και SegResNet αποτελούν τις ισχυρότερες πρακτικές επιλογές.

7.2 Μελλοντικές επεκτάσεις

Ισχυρότερη εκτίμηση γενίκευσης. Κρίσιμη συνέχεια αποτελεί η αξιολόγηση σε ανεξάρτητο test set και η χρήση cross-validation, ώστε τα συμπεράσματα να υποστηρίζονται από πιο σταθερή στατιστική τεκμηρίωση.

Εξωτερική επικύρωση και domain shift. Απαιτείται έλεγχος των μοντέλων σε δεδομένα από διαφορετικά κέντρα, scanners και πρωτόκολλα απεικόνισης, για καλύτερη αποτίμηση ανθεκτικότητας σε μεταβολές κατανομής.

Μετρικές ορίων και ανάλυση σφαλμάτων. Η προσθήκη boundary-sensitive μετρικών (π. χ. HD95) και η συστηματική κατηγοριοποίηση απουχιών ανά μορφολογικό μοτίβο θα προσφέρουν πληρέστερη εικόνα πέρα από τις overlap μετρικές.

Αποδοτικότητα και επιχειρησιακή αξιολόγηση. Η καταγραφή χρόνου εκπαίδευσης/inference, απαιτήσεων VRAM και συνολικού κόστους θα βοηθήσει στην επιλογή μοντέλων με ισορροπία ακρίβειας και υπολογιστικής βιωσιμότητας.

Εκτίμηση αβεβαιότητας και *human-in-the-loop*. Η ενσωμάτωση χαρτών αβεβαιότητας και στρατηγικών ημι-αυτόματης διόρθωσης μπορεί να αυξήσει τη διαφάνεια και την αξιοπιστία σε απαιτητικές περιπτώσεις.

Πολυτροπική και πολυ-εργασιακή μοντελοποίηση. Πιθανές επεκτάσεις περιλαμβάνουν αξιοποίηση επιπλέον ακολουθιών MRI όπου διατίθενται, καθώς και πολυ-εργασιακή μάθηση που συνδυάζει τμηματοποίηση με παράλληλη πρόβλεψη κλινικών δεικτών.

Βιβλιογραφία

- [1] Kaiming He, Xiangyu Zhang, Shaoqing Ren και Jian Sun. *Deep Residual Learning for Image Recognition*. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, σελίδες 770–778, 2016.
- [2] A Alansary, K Kamnitsas, A Davidson, R Khlebnikov, M Rajchl, C Malamateniou, M Rutherford, JV Hajnal, B Glocker, D Rueckert και et al. *Fast Fully Automatic Segmentation of the Human Placenta from Motion Corrupted MRI*. 2016.
- [3] Guotai Wang, Maria A. Zuluaga, Rosalind Pratt, Michael Aertsen, Tom Doel, Maria Klusmann, Anna L. David, Jan Deprest, Tom Vercauteren και Sébastien Ourselin. *Slic-Seg: A minimally interactive segmentation of the placenta from sparse and motion-corrupted fetal MRI in multiple views*. *Medical Image Analysis*, 34:137–147, 2016. Σπεσιαλ Ισσυε ον τηε 2015 ονφερενςε ον Μεδισαλ Ιμαγε θμπυτινγ ανδ θμπυτερ Ασσιστεδ Ιντερεντιον.
- [4] Maysam Shahedi, Catherine Y. Spong, James D. Dormer, Quyen N. Do, Yin Xi, Matthew A. Lewis, Christina Herrera, Ananth J. Madhuranthakam, Diane M. Twickler και Baowei Fei. *Deep learning-based segmentation of the placenta and uterus on MR images*. *Journal of Medical Imaging*, 8(5):054001, 2021.
- [5] Olaf Ronneberger, Philipp Fischer και Thomas Brox. *U-Net: Convolutional Networks for Biomedical Image Segmentation*. *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015*, τόμος 9351 στο *Lecture Notes in Computer Science*, σελίδες 234–241. Springer, 2015.
- [6] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser και Illia Polosukhin. *Attention Is All You Need*. arXiv preprint, 2017.
- [7] Ali Hatamizadeh, Yucheng Tang, Vishwesh Nath, Dong Yang, Andriy Myronenko, Holger R. Roth και Daguang Xu. *UNETR: Transformers for 3D Medical Image Segmentation*. *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, σελίδες 574–584, 2022.
- [8] Ali Hatamizadeh, Vishwesh Nath, Yucheng Tang, Dong Yang, Holger R. Roth και Daguang Xu. *Swin UNETR: Swin Transformers for Semantic Segmentation of Brain Tumors in MRI Images*. *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries (BrainLes 2021)*, τόμος 12962 στο *Lecture Notes in Computer Science*, σελίδες 272–284. Springer, 2022.

- [9] Albert Gu, Tri Dao και others. *Mamba: Linear-Time Sequence Modeling with Selective State Spaces*. arXiv preprint, 2023.
- [10] Zhaohu Xing, Tian Ye, Yijun Yang, Guang Liu και Lei Zhu. *SegMamba: Long-range Sequential Modeling Mamba For 3D Medical Image Segmentation*. *Medical Image Computing and Computer Assisted Intervention - MICCAI 2024*, τόμος 15008 στο *Lecture Notes in Computer Science*, σελίδες 578–588. Springer, 2024.
- [11] M. Jorge Cardoso, Wenqi Li, Richard Brown, Nic Ma, Eric Kerfoot, Yiheng Wang, Benjamin Murray, Andriy Myronenko, Can Zhao, Dong Yang και others. *MONAI: An open-source framework for deep learning in healthcare*. arXiv preprint, 2022.
- [12] Andriy Myronenko. *3D MRI brain tumor segmentation using autoencoder regularization*. arXiv preprint, 2018.
- [13] Ozan Oktay, Jo Schlemper, Loic Le Folgoc, Matthew Lee, Mattias Heinrich, Kazunari Misawa, Kensaku Mori, Steven McDonagh, Nils Y. Hammerla, Bernhard Kainz, Ben Glocker και Daniel Rueckert. *Attention U-Net: Learning Where to Look for the Pancreas*. arXiv preprint, 2018.
- [14] Graham J. Burton και Eric Jauniaux. *What is the placenta? A guide for clinicians*. *American Journal of Obstetrics and Gynecology*, 2015.
- [15] Emin Maltepe και Susan J. Fisher. *Placenta: the forgotten organ*. *Annual Review of Cell and Developmental Biology*, 31:523–552, 2015.
- [16] Naveen M. Gude, Clare T. Roberts, Bill Kalionis και Roger G. King. *Growth and function of the normal human placenta*. *Thrombosis Research*, 114(5-6):397–407, 2004.
- [17] Charlotte H. J. R. Jansen, Arnoud W. Kastelein, C. Emily Kleinrouweler, Elisabeth Van Leeuwen, Kees H. De Jong, Eva Pajkrt και Cornelis J. F. Van Noorden. *Development of placental abnormalities in location and anatomy*. *Acta Obstetricia et Gynecologica Scandinavica*, 99(8):983–993, 2020.
- [18] Ivo Brosens, Robert Pijnenborg, Liesbeth Vercruyse και Roberto Romero. *The “Great Obstetrical Syndromes” are associated with disorders of deep placentation*. *American Journal of Obstetrics and Gynecology*, 204(3):193–201, 2011.
- [19] Emin Maltepe, Anna I. Bakardjiev και Susan J. Fisher. *The placenta: transcriptional, epigenetic, and physiological integration during development*. *The Journal of Clinical Investigation*, 120(4):1016–1025, 2010.
- [20] Bordoni B. Herrick EJ. *Embryology, Placenta*.
- [21] Giuseppe Morgese, Samuele Salti, Lukas Bednar και Daniel Soukup. *Mitigating the effects of severe imbalance in multi-class semantic segmentation*. Μεταπτυχιακή διπλωματική εργασία, University of Bologna, 2024.

- [22] Konstantinos Kamnitsas, Christian Ledig, Virginia F.J. Newcombe, Joanna P. Simpson, Andrew D. Kane, David K. Menon, Daniel Rueckert και Ben Glocker. *Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation.* *Medical Image Analysis*, 36:61–78, 2017.
- [23] Fabian Isensee, Paul F. Jaeger, Simon A. A. Kohl, Jens Petersen και Klaus H. Maier-Hein. *Nature Methods*, (2):203–211, 2020.
- [24] Maysam Shahedi, James D Dormer, Quyen N Do, Yin Xi, Matthew A Lewis, Christina L Herrera, Catherine Y Spong, Ananth J Madhuranthakam, Diane M Twickler και Baowei Fei. *Automatic segmentation of uterine cavity and placenta on MR images using deep learning.* *Proceedings of SPIE-the International Society for Optical Engineering*, τόμος 12036, σελίδα 1203611, 2022.
- [25] Yongkai Liu, Fatemeh Zabihollahy, Ran Yan, Brian Lee, Carla Janzen, Sherin U Devaskar και Kyunghyun Sung. *Evaluation of spatial attentive deep learning for automatic placental segmentation on longitudinal MRI.* *Journal of Magnetic Resonance Imaging*, 57(5):1533–1540, 2023.
- [26] E. Mark Haacke, Robert W. Brown, Michael R. Thompson και Ramesh Venkatesan. *Magnetic Resonance Imaging: Physical Principles and Sequence Design.* Wiley-Liss, 1999.
- [27] Donald W. McRobbie, Elizabeth A. Moore, Martin J. Graves και Martin R. Prince. *MRI from Picture to Proton.* Cambridge University Press, 2η έκδοση, 2006.
- [28] Nicholas J. Tustison, Brian B. Avants, Philip A. Cook, Yuanjie Zheng, Alexander Egan, Paul A. Yushkevich και James C. Gee. *N4ITK: Improved N3 Bias Correction.* *IEEE Transactions on Medical Imaging*, 29(6):1310–1320, 2010.
- [29] Sabir Muhammad και Jing Zhang. *Segmentation of Liver Tumors by Monai and PyTorch in CT Images with Deep Learning Techniques.* *Applied Sciences*, 14(12), 2024.
- [30] Geert Litjens, Thijs Kooi, Babak Ehteshami Bejnordi, Arnaud Arindra Adiyoso Setio, Francesco Ciompi, Mohsen Ghafoorian, Jeroen A. W. M. van der Laak, Bramvan Ginneken και Clara I. Sánchez. *A survey on deep learning in medical image analysis.* *Medical Image Analysis*, 42:60–88, 2017.
- [31] Ian Goodfellow, Yoshua Bengio και Aaron Courville. *Deep Learning.* MIT Press, 2016.
- [32] Christopher M. Bishop. *Pattern Recognition and Machine Learning.* Springer, 2006.
- [33] Carole H. Sudre, Wenqi Li, Tom Vercauteren, Sébastien Ourselin και M. Jorge Cardoso. *Generalised Dice Overlap as a Deep Learning Loss Function for Highly Unbalanced Segmentations.* *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support (DLMIA 2017)*, τόμος 10553 στο *Lecture Notes in Computer Science*, σελίδες 240–248. Springer, 2017.

- [34] Fausto Milletari, Nassir Navab και Seyed Ahmad Ahmadi. *V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation*. arXiv preprint, 2016.
- [35] Charu C Aggarwal και others. *Neural networks and deep learning*, τόμος 10. Springer, 2018.
- [36] Alex Krizhevsky, Ilya Sutskever και Geoffrey E. Hinton. *ImageNet Classification with Deep Convolutional Neural Networks*. *Advances in Neural Information Processing Systems (NeurIPS)*, 2012.
- [37] Jonathan Long, Evan Shelhamer και Trevor Darrell. *Fully Convolutional Networks for Semantic Segmentation*. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, σελίδες 3431–3440, 2015.
- [38] Özgün Cicek, Ahmed Abdulkadir, Soeren S. Lienkamp, Thomas Brox και Olaf Ronneberger. *3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation*. *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2016*, τόμος 9901 στο *Lecture Notes in Computer Science*, σελίδες 424–432. Springer, 2016.
- [39] Sergey Ioffe και Christian Szegedy. *Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift*. 2015.
- [40] Yuxin Wu και Kaiming He. *Group Normalization*. 2018.
- [41] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, Jakob Uszkoreit και Neil Houlsby. *An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale*. arXiv preprint, 2021.
- [42] Ze Liu, Yutong Lin, Yue Cao, Han Hu, Yixuan Wei, Zheng Zhang, Stephen Lin και Baining Guo. *Swin Transformer: Hierarchical Vision Transformer using Shifted Windows*. arXiv preprint, 2021.
- [43] Sinong Wang, Belinda Z. Li, Madian Khabsa, Han Fang και Hao Ma. *Linformer: Self-Attention with Linear Complexity*. 2020.
- [44] Krzysztof Choromanski και others. *Rethinking Attention with Performers*. 2021.
- [45] Albert Gu, Karan Goel και Christopher Ré. *Efficiently Modeling Long Sequences with Structured State Spaces*. *International Conference on Learning Representations (ICLR)*, 2022.
- [46] Karan Goel, Albert Gu, Chris Donahue και Christopher Ré. *It's Raw! Audio Generation with State-Space Models*. *International Conference on Machine Learning (ICML)*, 2022.
- [47] Ilya Loshchilov και Frank Hutter. *Decoupled Weight Decay Regularization*. arXiv preprint, 2019.

-
- [48] *Learning Rate Schedulers in PyTorch*. Ассеосеδ: 2024-06-10.
 - [49] Project MONAI. *MONAI Documentation*. <https://monai-dev.readthedocs.io/en/new-doc-theme/api.html>, 2025. Ассеосеδ: 2026-01-25.
 - [50] Project MONAI. *Project MONAI: MONAI (GitHub Repository)*. <https://github.com/Project-MONAI/MONAI>, 2025. Ассеосеδ: 2026-01-25.
 - [51] Zhaohu Xing, Tian Ye, Yijun Yang, Guang Liu και Lei Zhu. *SegMamba: Official Implementation (GitHub Repository)*. <https://github.com/ge-xing/SegMamba>, 2024. Ассеосеδ: 2026-01-25.
 - [52] Fabian Isensee, Paul F. Jaeger, Simon A. A. Kohl, Jens Petersen και Klaus H. Maier-Hein. *nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation*. *Nature Methods*, 18(2):203–211, 2021.
 - [53] Project MONAI. *MONAI Transforms Documentation*. <https://monai-dev.readthedocs.io/en/new-doc-theme/transforms.html#dictionary-transforms>, 2024. Ассеосеδ: 2024-06-10.
 - [54] Leslie N. Smith. *Cyclical Learning Rates for Training Neural Networks*, 2017.
 - [55] James Huang, Quyen N. Do, Maysam Shahedi, Yin Xi, Matthew A. Lewis, Christina L. Herrera, David Owen, Catherine Y. Spong, Ananth J. Madhuranthakam, Diane M. Twickler και Baowei Fei. *Deep learning based automatic segmentation of placenta and uterine cavity on prenatal MR images*. *Medical Imaging 2023: Computer-Aided Diagnosis*, τόμος 12465, σελίδα 124650N. SPIE, 2023.
 - [56] Cong Lee, Zhifang Liao, Yuanzhe Li, Qingquan Lai, Yingying Guo, Jing Huang, Shuting Li, Yi Wang και Ruizheng Shi. *Placental magnetic resonance imaging segmentation based on multi-receptive field and mixed attention separation mechanism*. *Computer Methods and Programs in Biomedicine*, 242:107699, 2023.
 - [57] Isso Saito, Shinnosuke Yamamoto, Eichi Takaya, Ayaka Harigai, Tomomi Sato, Tomoya Kobayashi, Kei Takase και Takuya Ueda. *PlaNet-S: an automatic semantic segmentation model for placenta using U-Net and SegNeXt*. *Journal of Imaging Informatics in Medicine*, 2025.