



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

Σχολή Ηλεκτρολόγων Μηχανικών
και Μηχανικών Υπολογιστών

Τομέας Συστημάτων Μετάδοσης Πληροφορίας και
Τεχνολογίας Υλικών

**Αυτοματοποιημένη ανίχνευση αδενοκαρκινώματος του παγκρέατος σε
αξονικές τομογραφίες με χρήση βαθιάς μάθησης: διερεύνηση της επίδρασης
της προεπεξεργασίας των δεδομένων**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Κυριακή Καρατζούνη

Επιβλέπων : Γεώργιος Ματσόπουλος

Καθηγητής Ε.Μ.Π.

Αθήνα, Ιούλιος, 2025



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

Σχολή Ηλεκτρολόγων Μηχανικών
και Μηχανικών Υπολογιστών

Τομέας Συστημάτων Μετάδοσης Πληροφορίας και
Τεχνολογίας Υλικών

**Αυτοματοποιημένη ανίχνευση αδενοκαρκινώματος του παγκρέατος σε
αξονικές τομογραφίες με χρήση βαθιάς μάθησης: διερεύνηση της επίδρασης
της προεπεξεργασίας των δεδομένων**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Κυριακή Καρατζούνη

Επιβλέπων : Γεώργιος Ματσόπουλος

Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 2α Ιουλίου 2025

Αθανάσιος
Παναγόπουλος
Καθηγητής, Ε.Μ.Π.

Παναγιώτης
Τσανάκας
Καθηγητής, Ε.Μ.Π.

Ουρανία
Πετροπούλου
ΕΔΙΠ, Ε.Μ.Π.

Αθήνα, Ιούλιος, 2025

Κυριακή Καρατζούνη

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π

Copyright © Κυριακή Καρατζούνη, 2025

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

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

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

Περίληψη

Η παρούσα διπλωματική εργασία εξετάζει πώς οι επιλογές προεπεξεργασίας επηρεάζουν την ακρίβεια της αυτόματης τμηματοποίησης παγκρεατικού αδενοκαρκινώματος σε πολυφασικές αξονικές τομογραφίες. Χρησιμοποιήθηκε το δημόσιο σύνολο δεδομένων MSD Task 07 – Pancreas (420 3-D CT εξετάσεις) και το πλήρως αυτορρυθμιζόμενο πλαίσιο nnU-Net. Σχεδιάστηκαν 15 πειραματικά runs σε τρεις θεματικές:

1. Voxel spacing: ισοτροπικές ($1 \times 1 \times 1$, $1,5 \times 1,5 \times 1,5$, $2 \times 2 \times 2$ mm) και ανισοτροπικές ($1 \times 1 \times 2,5$, $1 \times 1 \times 5$ mm) επαναδειγματοληψίες.
2. Interpolation: spline τάξεων 1-5 (linear–quintic).
3. Intensity clipping: ποσοστημόρια 0–100 %, 0,1–99,9 %, 1–99 %, 5–95 %, 10–90 %.

Οι πειραματικές μελέτες έδειξαν ότι ο καθορισμός του κατάλληλου voxel spacing αποτελεί το κλειδί για τη βέλτιστη απόδοση του nnU-Net στην τμηματοποίηση του παγκρεατικού αδενοκαρκινώματος. Συγκεκριμένα, η χρήση ελαφρώς ανισοτροπικού spacing ($1 \times 1 \times 2,5$ mm) οδήγησε σε αύξηση του Dice coefficient για τον όγκο κατά σχεδόν 30% σε σχέση με ένα ισοτροπικό spacing 2 mm, ενώ η υπερβολική μείωση της ανάλυσης του άξονα z ($1 \times 1 \times 5$ mm) συρρίκνωσε την ευαισθησία ανίχνευσης. Αντίθετα, οι διαφοροποιήσεις στη μέθοδο παρεμβολής (linear έως quintic spline) απέφεραν μεταβολές στον Dice μικρότερου από το 2%, επιβεβαιώνοντας ότι η προεπιλεγμένη κυβική παρεμβολή είναι απόλυτα επαρκής. Πολύ περιοριστική κανονικοποίηση έντασης (clip 10–90%) αύξησε οριακά τον δείκτη Sensitivity, αλλά συνοδεύτηκε από ελαφρά πτώση του συνολικού Dice, ενώ η πιο ήπια αποκοπή (1–99%) διατήρησε τις βέλτιστες τιμές.

Συμπερασματικά, η βέλτιστη ρύθμιση του voxel spacing, σύμφωνα με τις ευριστικές μεθόδους του nnU-Net, είναι καθοριστική για την επίτευξη υψηλής ποιότητας segmentation του παγκρεατικού αδενοκαρκινώματος, ενώ οι προεπιλεγμένες τεχνικές interpolation και intensity normalization παρέχουν ήδη επαρκώς βέλτητα αποτελέσματα. Τα ευρήματα αυτά διαμορφώνουν ένα επαρκές πλαίσιο για την κλινική ενσωμάτωση συστημάτων CAD στο πάγκρεας και οριοθετούν σαφείς κατευθύνσεις για μελλοντική έρευνα, όπως η ανάπτυξη πολυβάθμιων (cascade) μοντέλων και η ενσωμάτωση πολυτροπικών απεικονιστικών δεδομένων.

Λέξεις Κλειδιά

Παγκρεατικό Αδενοκαρκίνωμα, Αξονική Τομογραφία, Βαθιά Μάθηση, nnU-Net, Τμηματοποίηση Εικόνας, Προεπεξεργασία Εικόνας, Voxel Spacing, Παρεμβολή, Κανονικοποίηση Έντασης

Abstract

This thesis investigates how preprocessing choices affect the accuracy of automated segmentation of pancreatic adenocarcinoma in multiphase computed tomography scans. The publicly available MSD Task 07 – Pancreas dataset (420 3-D CT scans) and the fully self-configuring nnU-Net framework were used. Fifteen experimental runs were designed, grouped into three thematic areas:

1. Voxel spacing: isotropic ($1 \times 1 \times 1$, $1.5 \times 1.5 \times 1.5$, $2 \times 2 \times 2$ mm) and anisotropic ($1 \times 1 \times 2.5$, $1 \times 1 \times 5$ mm) resampling.
2. Interpolation: spline interpolation of order 1–5 (linear–quintic).
3. Intensity clipping: percentiles 0–100%, 0.1–99.9%, 1–99%, 5–95%, 10–90%.

The experiments demonstrated that selecting appropriate voxel spacing is key for optimal nnU-Net performance in pancreatic adenocarcinoma segmentation. Specifically, the use of slightly anisotropic spacing ($1 \times 1 \times 2.5$ mm) resulted in an increase of almost 30% in Dice coefficient for volume compared to isotropic spacing of 2 mm, whereas excessive reduction of z-axis resolution ($1 \times 1 \times 5$ mm) led to decreased detection sensitivity. In contrast, variations in interpolation method (linear to quintic spline) produced changes in Dice of less than 2%, confirming that the default cubic interpolation is fully adequate. Highly restrictive intensity normalization (clip 10–90%) marginally increased the Sensitivity index, but was accompanied by a slight decrease in overall Dice, while milder clipping (1–99%) maintained optimal values.

In conclusion, optimal tuning of voxel spacing—following nnU-Net heuristics—is crucial for achieving high-quality segmentation of pancreatic adenocarcinoma, while the default interpolation and intensity normalization techniques already provide sufficiently optimal results. These findings establish a robust framework for the clinical integration of CAD systems for the pancreas and outline clear directions for future research, such as the development of cascade models and the incorporation of multimodal imaging data.

Key Words

Pancreatic Adenocarcinoma, Computed Tomography, Deep Learning, nnU-Net, Image Segmentation, Image Preprocessing, Voxel Spacing, Interpolation, Intensity Normalization

Ευχαριστίες

Για την εκπόνηση αυτής της διπλωματικής εργασίας θα ήθελα να ευχαριστήσω ιδιαίτερα τον επιβλέποντα καθηγητή κ. Γιώργο Ματσόπουλο και τον μεταδιδακτορικό ερευνητή κ. Ιωάννη Βεζάκη για τη συνεργασία και την υποστήριξη που μου παρείχανε καθ' όλη τη διάρκεια.

Επιπλέον, θα ήθελα να ευχαριστήσω θερμά την οικογένειά μου και τους ανθρώπους μου, οι οποίοι με στήριξαν όλα αυτά τα χρόνια και συνεχίζουν να στηρίζουν όλες μου τις προσπάθειες.

ΠΕΡΙΕΧΟΜΕΝΑ

1. Εισαγωγή.....	14
1.1. Σκοπός της εργασίας.....	14
1.2. Ιατρική απεικόνιση και τεχνικές segmentation.....	14
1.3. Εφαρμογές βαθιάς μάθησης στη διάγνωση καρκίνου.....	15
1.4. Διεθνή δεδομένα και benchmarks για παγκρεατικό καρκίνο.....	15
2. Ιατρικό Υπόβαθρο	16
2.1. Ανατομία και φυσιολογία του παγκρέατος	16
Σχήμα 1: Ανατομικές σχέσεις του παγκρέατος με γύρω όργανα και δομές	16
Σχήμα 2: Φυσιολογικό πάγκρεας σε ανατομή για να αποκαλυφθεί το σύστημα των πόρων	17
2.2. Επιδημιολογία και κατηγοριοποίηση παγκρεατικού καρκίνου.....	18
Σχήμα 3: Παγκόσμιοι χάρτες.....	19
2.3. Διαγνωστική προσέγγιση και ρόλος της αξονικής τομογραφίας.....	20
2.3.1. Βασικές αρχές αξονικής τομογραφίας	20
Σχήμα 4: Τοποθέτηση ROI	20
2.3.2. Πρωτόκολλα λήψης (contrast, phases).....	21
2.3.3. Χρονικές και χωρικές παράμετροι (spacing, slice thickness)	22
3. Θεωρία Επεξεργασίας Εικόνας.....	23
3.1. Τι είναι εικόνα; pixels, voxels, Hounsfield units	23
Σχήμα 5: Σχηματική αναπαράσταση.....	23
3.2. Interpolation και resampling.....	24
3.3. Spacing και anisotropy	25
3.4. Intensity normalization και windowing	26
Σχήμα 6: Σχηματικό βήμα-προς-βήμα επαναδειγματοληψίας (resampling) [13]....	26
4. Θεωρία Νευρωνικών Δικτύων	27
4.1. Βασικές έννοιες: νευρώνας, συνάρτηση ενεργοποίησης.....	27
Σχήμα 7: Ταξινόμηση των συναρτήσεων ενεργοποίησης	27
4.2. Δομή δικτύου και forward/backward propagation (διάδοση προς τα εμπρός και προς τα πίσω)	28
4.3. Εκπαίδευση: loss functions, βελτιστοποίηση, regularization	29
Σχήμα 8: Focal Loss	29

4.4. Συνελκτικά νευρωνικά δίκτυα (CNNs) και ειδικές αρχιτεκτονικές.....	30
Σχήμα 9: Υπολειμματική μάθηση	31
Σχήμα 10: Ένα πυκνό μπλοκ 5 επιπέδων με ρυθμό ανάπτυξης $k = 4$	31
Σχήμα 11: Σχηματική απεικόνιση της προτεινόμενης πύλης προσθετικής προσοχής (AG).	32
5. Το nnU-Net Framework.....	32
5.1. Ιστορικό και γενική φιλοσοφία του nnU-Net.....	32
5.2. Βασική αρχιτεκτονική (U-Net backbone, encoder-decoder).....	33
Σχήμα 12: Απεικόνιση της αρχιτεκτονικής U-Net.....	34
5.3. Αυτόματη προεπεξεργασία (spacing, intensity, patch-based).....	34
5.4. Δυναμική επιλογή υπερπαραμέτρων και augmentation	35
5.5. Σύγκριση με άλλες state-of-the-art μεθόδους.....	36
6. Μεθοδολογία Πειραμάτων	37
6.1. Περιγραφή dataset	37
6.2. Υπολογιστική υποδομή.....	37
6.2.1. GPU και hardware details.....	37
6.2.2. Περιβάλλον λογισμικού (OS, βιβλιοθήκες)	38
6.3. Σχεδιασμός πειραμάτων.....	38
6.3.1. Παράμετροι προς μεταβολή (spacing, interpolation, intensity)	38
6.3.2. Πρωτόκολλο εκπαίδευσης (15 runs $\times$ 1000 epochs).....	39
6.4. Μέτρα αξιολόγησης.....	39
6.4.1. Dice coefficient, sensitivity, specificity, precision.....	39
7. Αποτελέσματα	41
7.1. Πίνακας σύγκρισης πειραμάτων.....	41
Πίνακας 1: Αποτελέσματα για διαφορετικό spacing.....	41
Πίνακας 2: Αποτελέσματα για διαφορετικό Interpolation.....	41
Πίνακας 3: Αποτελέσματα για διαφορετικό Clipping	42
7.2. Διαγράμματα	43
7.3. Αξιολόγηση και ερμηνεία αποτελεσμάτων.....	58
7.3.1 Επίδραση Spacing.....	58
7.3.2 Επίδραση Interpolation.....	59
7.3.3 Επίδραση Intensity Normalization	59

7.3.4 Συμπεράσματα	60
8. Συζήτηση	63
8.1. Επίδραση spacing/interpolation/intensity στα αποτελέσματα	63
8.2. Συγκριτική ερμηνεία σε σχέση με τις συστάσεις του nnU-Net.....	65
8.3. Περιορισμοί της μελέτης	66
8.4. Προτάσεις βελτίωσης.....	68
9. Συμπεράσματα και Μελλοντική Εργασία	70
9.1. Κύρια συμπεράσματα.....	70
9.2. Ιδέες για επόμενες έρευνες.....	74
10. Βιβλιογραφία	77

1. Εισαγωγή

1.1. Σκοπός της εργασίας

Η παρούσα διπλωματική εργασία έχει ως στόχο να ερευνήσει την επίδραση διάφορων μεθόδων προεπεξεργασίας δεδομένων, πιο συγκεκριμένα της απόστασης μεταξύ των τομών (spacing), των μεθόδων παρεμβολής (interpolation) και της κανονικοποίησης έντασης (intensity normalization) στην απόδοση των συστημάτων βαθιάς μάθησης που χρησιμοποιούνται για την αυτόματη ανίχνευση αδενοκαρκινώματος του παγκρέατος σε εικόνες αξονικής τομογραφίας.

Για τον σκοπό αυτό, χρησιμοποιείται το nnU-Net, ένα αυτοματοποιημένο εργαλείο για τη δημιουργία μοντέλων βαθιάς μάθησης στον τομέα της ιατρικής τμηματοποίησης (segmentation). Ο στόχος είναι να εκτιμηθεί σε ποιον βαθμό οι τεχνικές προεπεξεργασίας επηρεάζουν την ακρίβεια του τελικού αποτελέσματος, με απώτερο σκοπό τη βελτιστοποίηση της εφαρμογής τους στην κλινική πράξη.

1.2. Ιατρική απεικόνιση και τεχνικές segmentation

Η ιατρική απεικόνιση μπορεί να παράγει λεπτομερείς εικόνες των εσωτερικών μερών του σώματος ενός ατόμου και αποτελεί ένα από τα ισχυρά εργαλεία διάγνωσης. Η αξονική τομογραφία, η μαγνητική τομογραφία και η τομογραφία εκπομπής ποζιτρονίων αποτελούν τη βάση της διάγνωσης και παρακολούθησης της νόσου, καθώς παρέχουν τρισδιάστατες εικόνες υψηλής ευκρίνειας που βοηθούν στην ακριβή αξιολόγηση όχι μόνο της ανατομίας αλλά και της λειτουργίας στους ιστούς [28].

Η τμηματοποίηση εικόνας είναι η διαδικασία της αυτόματης ή ημιαυτόματης ανίχνευσης και απομόνωσης συγκεκριμένων περιοχών ή δομών ενδιαφέροντος από ιατρικές εικόνες. Οι παραδοσιακές μέθοδοι τμηματοποίησης περιλαμβάνουν την κατωφλίωση, την ανίχνευση ακμών και την ανάπτυξη περιοχών που βασίζεται σε γεωμετρικά χαρακτηριστικά των περιοχών. Ωστόσο, οι τεχνικές αυτές περιορίζονται από το θόρυβο, τη χαμηλή αντίθεση και τους ανομοιογενείς ιστούς [18].

Οι πρόσφατες τεχνολογικές εξελίξεις, ιδίως η εφαρμογή τεχνικών βαθιάς μάθησης, έχουν βελτιώσει σημαντικά την ακρίβεια και την αποτελεσματικότητα της τμηματοποίησης [29]. Αρχιτεκτονικές όπως τα Συνελκτικά Νευρωνικά Δίκτυα (CNNs) και, ειδικότερα, το U-Net έχουν αποδειχθεί εξαιρετικά αποτελεσματικές στην τμηματοποίηση ιατρικών εικόνων λόγω της ικανότητάς τους να κατανοούν πολύπλοκες σχέσεις και λεπτές δομές. Αυτές οι αρχιτεκτονικές χρησιμοποιούνται πλέον σε ένα ευρύ φάσμα ιατρικών εφαρμογών, όπως η ανίχνευση και η αξιολόγηση νεοπλασιών, η μελέτη της ανάπτυξης παθολογικών δομών και η παρακολούθηση της ανταπόκρισης στη θεραπεία [38].

Ωστόσο, η εφαρμογή αυτών των μεθοδολογιών σε κλινικά δεδομένα εξακολουθεί να αντιμετωπίζει σημαντικές προκλήσεις. Η ενσωμάτωση αυτών των τεχνικών στην κλινική πρακτική παρεμποδίζεται από διάφορους παράγοντες. Σε αυτούς περιλαμβάνονται η ανάγκη

για επαρκώς μεγάλα σύνολα δεδομένων, η διαφοροποίηση των πρωτοκόλλων απεικόνισης μεταξύ διαφορετικών κλινικών κέντρων και η ανάγκη για ερμηνευσιμότητα των αποτελεσμάτων [18].

1.3. Εφαρμογές βαθιάς μάθησης στη διάγνωση καρκίνου

Οι τεχνικές βαθιάς μάθησης έχουν ανοίξει νέες προοπτικές στην άμεση και ακριβή διάγνωση του καρκίνου, αξιοποιώντας τεράστιες συλλογές ιατρικών εικόνων για την εκπαίδευση πολύπλοκων νευρωνικών δικτύων. Η ικανότητα των Συνελικτικών Νευρωνικών Δικτύων (CNNs) να μαθαίνουν ιεραρχικές αναπαραστάσεις από ακατέργαστες εικόνες επιτρέπει την αυτόματη αναγνώριση και περιγραφή των παθολογικών χαρακτηριστικών με ακρίβεια που συχνά ανταγωνίζεται ή ξεπερνά αυτή των ανθρώπινων επαγγελματιών [28].

Στη θωρακική απεικόνιση, τα μοντέλα 3D-CNN έχουν χρησιμοποιηθεί με επιτυχία για τον έλεγχο και την έγκαιρη ανίχνευση του καρκίνου του πνεύμονα με αξονική τομογραφία χαμηλής δόσης. Στη μελέτη των Ardila et al. το δίκτυο πέτυχε ROC AUC 0,944 στην αναγνώριση όγκων, συγκρίνοντάς το με τα αποτελέσματα που προέκυψαν από δύο ακτινολόγους, μειώνοντας τον χρόνο επεξεργασίας και τη διαγνωστική υποκειμενικότητα [2].

Στο επίπεδο της δερματολογίας, το μοντέλο που παρουσιάστηκε από τους Esteva et al. πέτυχε την ταξινόμηση εικόνων δερματικών βλαβών σε καλοήθεις και κακοήθεις με ακρίβεια ισοδύναμη ή υψηλότερη από εκείνη των εξειδικευμένων δερματολόγων. Η μελέτη χρησιμοποίησε περίπου 129.000 φωτογραφίες και κατέδειξε την πρακτική ανάγκη της εισαγωγής συμπληρωματικών εργαλείων στην κλινική πρακτική [12].

Η εφαρμογή της βαθιάς μάθησης στην παθολογία περιλαμβάνει την αναγνώριση μιτώσεων σε ιστολογικές τομές για τον υπολογισμό του δείκτη MIB-1 (Ki-67), ο οποίος συνδέεται με την πρόγνωση ορισμένων μορφών καρκίνου του μαστού. Το δίκτυο των Cireşan et al. πέτυχε υψηλό F1-score (μέτρο που συνδυάζει την ακρίβεια και την ευαισθησία της ανίχνευσης μιτώσεων) στην αυτόματη αναγνώριση μιτώσεων, ενισχύοντας τόσο την ταχύτητα όσο και την αντικειμενικότητα της μελέτης [6].

1.4. Διεθνή δεδομένα και benchmarks για παγκρεατικό καρκίνο

Το Medical Segmentation Decathlon (MSD) Task 07 «Pancreas» περιέχει 420 τρισδιάστατες αξονικές τομογραφίες από διάφορες τοποθεσίες, τόσο για παγκρεατικές όσο και για νεοπλασματικές βλάβες. Στον αρχικό διαγωνισμό, ο μέσος όρος των ομάδων έφθασε σε βαθμολογία Dice 0,69 για το πάγκρεας, ενώ η τεχνική nnU-Net σημείωσε βαθμολογία 0,93, υποδεικνύοντας τη χρησιμότητα αυτού του σημείου αναφοράς για την αξιολόγηση της γενικευσιμότητας των αλγορίθμων [1].

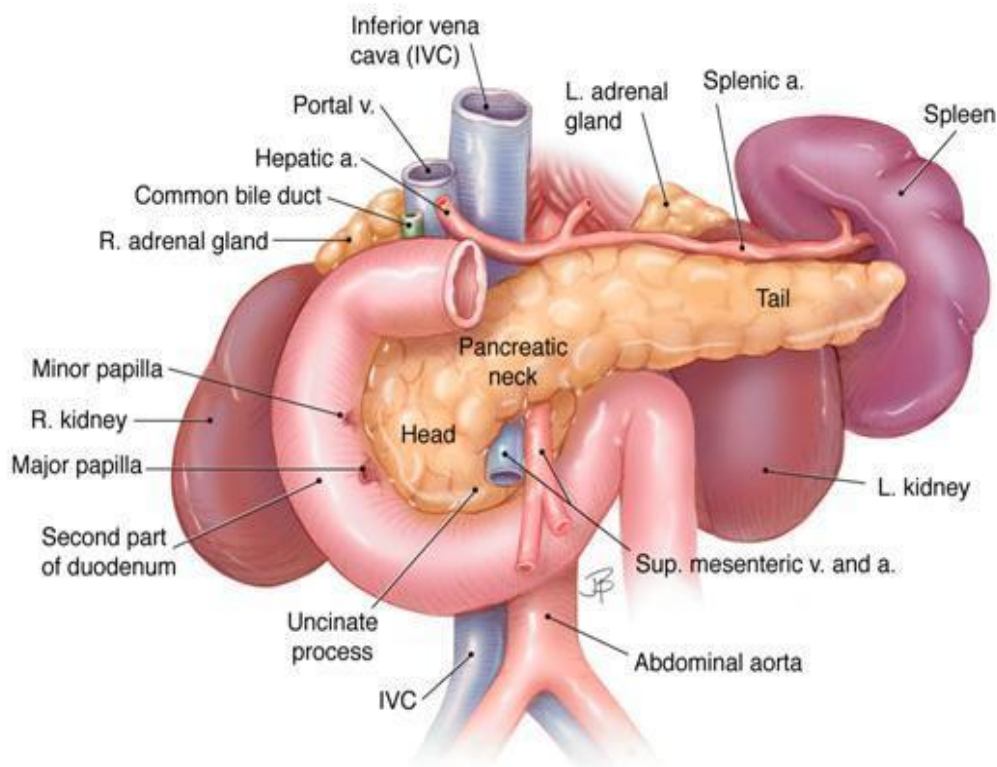
Πρόσφατα, η μελέτη των Mahmoudi et al. πρότεινε ένα υβριδικό πλαίσιο Attention U-Net ενισχυμένο με περιγραφείς υφής, το οποίο αξιολογήθηκε τόσο στο MSD Task 07 όσο και στη

συλλογή TCIA Pancreas-CT [39]. Η προσθήκη των attention μηχανισμών και των περιγραφέων υφής οδήγησε σε αύξηση 7,5 % στο ΔDSC (Διαφορά Dice Similarity Coefficient) για την κατάτμηση νεοπλασματικών ιστών, δείχνοντας την εξέλιξη από την απλή εντόπιση οργάνων προς την ακριβή διάκριση όγκων [31][1].

2. Ιατρικό Υπόβαθρο

2.1. Ανατομία και φυσιολογία του παγκρέατος

Το πάγκρεας είναι ένα όργανο μήκους 12-15 εκατοστών, που βρίσκεται στην οπισθοπεριτοναϊκή περιοχή της άνω κοιλίας, στο επίπεδο των σπονδύλων L1-L2, πίσω από το στομάχι. Ανατομικά διακρίνεται σε κεφαλή (η οποία «αγκαλιάζει» την καμπή του δωδεκαδακτύλου), λαιμό, σώμα και ουρά. Ο πρωτοπαθής παγκρεατικός πόρος διέρχεται από τον αδένα από την ουρά προς την κεφαλή και εκβάλλει στην αμπούλα του Vater μαζί με τον χοληδόχο πόρο [44].



Σχήμα 1: Ανατομικές σχέσεις του παγκρέατος με γύρω όργανα και δομές [30].

Η κεφαλή του παγκρέατος τοποθετείται στη θηλιά του δωδεκαδακτύλου καθώς εξέρχεται από το στομάχι.

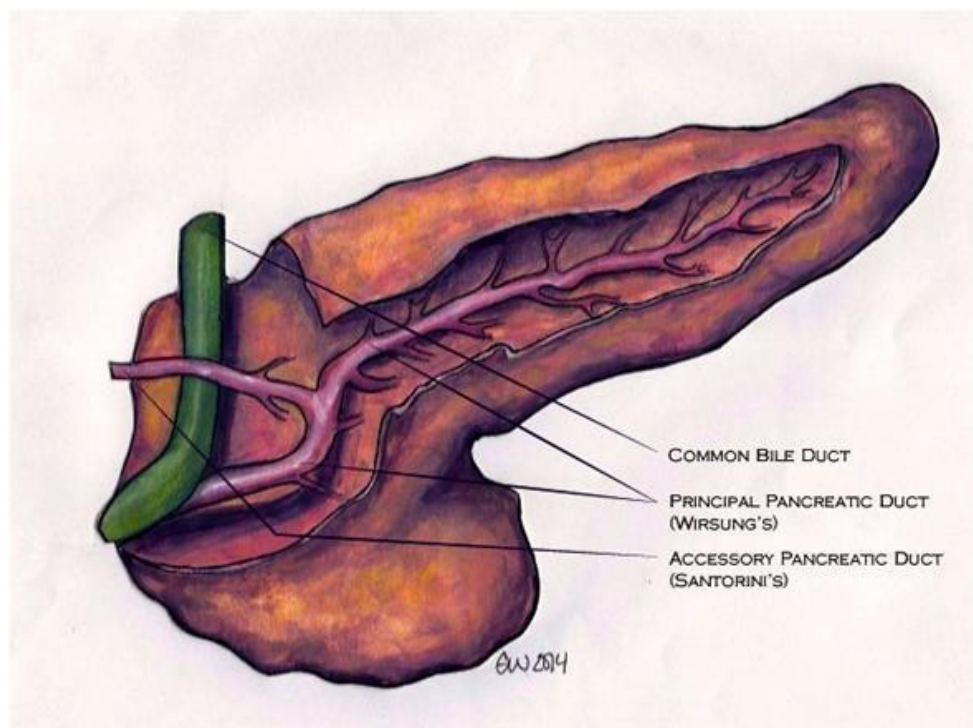
Η ουρά του παγκρέατος τοποθετείται κοντά στη σπληνική ράχη.

Το σώμα του παγκρέατος βρίσκεται πίσω από την άνω περιοχή του στομάχου μεταξύ της ουράς και του αυχένα και δεν σημειώνεται σε αυτό το σχέδιο.

Το τμήμα του παγκρέατος που βρίσκεται μπροστά από την αορτή είναι κάπως λεπτότερο από τα γειτονικά τμήματα της κεφαλής και του σώματος του παγκρέατος. Η περιοχή αυτή αναφέρεται συχνά ως λαιμός του παγκρέατος και υποδηλώνει τη συμβολή της κεφαλής με το σώμα.

Η στενή επαφή του αυχένα του παγκρέατος με τα κύρια αιμοφόρα αγγεία προς τα πίσω, συμπεριλαμβανομένης της άνω μεσεντέριας αρτηρίας, της άνω μεσεντέριας-παγκρεατικής φλέβας, της κάτω κοίλης φλέβας και της αορτής, περιορίζει τη δυνατότητα ευρέως χειρουργικού περιθωρίου κατά την παγκρεατεκτομή (χειρουργική αφαίρεση του παγκρέατος).

Το παγκρεατικό παρέγχυμα έχει λοβιακή αρχιτεκτονική: κάθε λοβός περιβάλλεται από συνδετικό ιστό που περιέχει τριχοειδή αγγεία και νεύρα. Τα εκκριτικά τμήματα (αδενοκύτταρα και πορογενή επιθηλιακά κύτταρα) αποτελούν το εξωκρινές ήμισυ (> 95 % της μάζας), ενώ τα νησίδια του Langerhans (1-2 %) είναι διασκορπισμένα ως μικροσκοπικές μορφές μεταξύ των λοβίων [30].



Σχήμα 2: Φυσιολογικό πάγκρεας σε ανατομή για να αποκαλυφθεί το σύστημα των πόρων [30].

Το εξωκρινές τμήμα παράγει παγκρεατικό υγρό, πλούσιο σε πεπτικά ένζυμα (αμυλάση, λιπάση, θρυψινογόνα) και διττανθρακικά, μέσω ενός δικτύου καναλιών που εκβάλλουν στον κύριο πόρο. Η έκκριση ελέγχεται ορμονικά (η χολοκυστοκίνη αυξάνει τα ένζυμα και η σεκρετίνη διεγείρει την έκκριση διττανθρακικών) και νευρολογικά.

Τα νησίδια Langerhans περιλαμβάνουν πέντε τύπους κυττάρων: β-κύτταρα (που παράγουν ινσουλίνη), α-κύτταρα (γλυκαγόνη), δ-κύτταρα (σωματοστατίνη), PP-κύτταρα (παγκρεατικό πεπτίδιο) και ε-κύτταρα (γκρελίνη). Η ινσουλίνη και η γλυκαγόνη διατηρούν την ομοιόσταση της γλυκόζης μέσω περίπλοκων βρόχων ανατροφοδότησης, επηρεάζοντας τις μεταβολικές διεργασίες σε πολλά όργανα [23].

Η αρτηριακή παροχή προέρχεται κυρίως από κλάδους του σπληνικού, άνω και κάτω παγκρεατοδωδεκαδακτυλικού δικτύου (κοινή σπληνική και σπονδυλική αρτηρία), με πολλές αναστομώσεις που εξασφαλίζουν ενισχυμένη αιμάτωση. Το νεύρο προέρχεται από το κοιλιακό πλέγμα, ενώ οι παρασυμπαθητικές ίνες ρέουν μέσω του πνευμονογαστρικού νεύρου. Η λεπτομερής γνώση αυτών των αγγειακών και νευρολογικών αλληλεπιδράσεων είναι ζωτικής σημασίας για τις χειρουργικές θεραπείες και τη μεταμόσχευση.

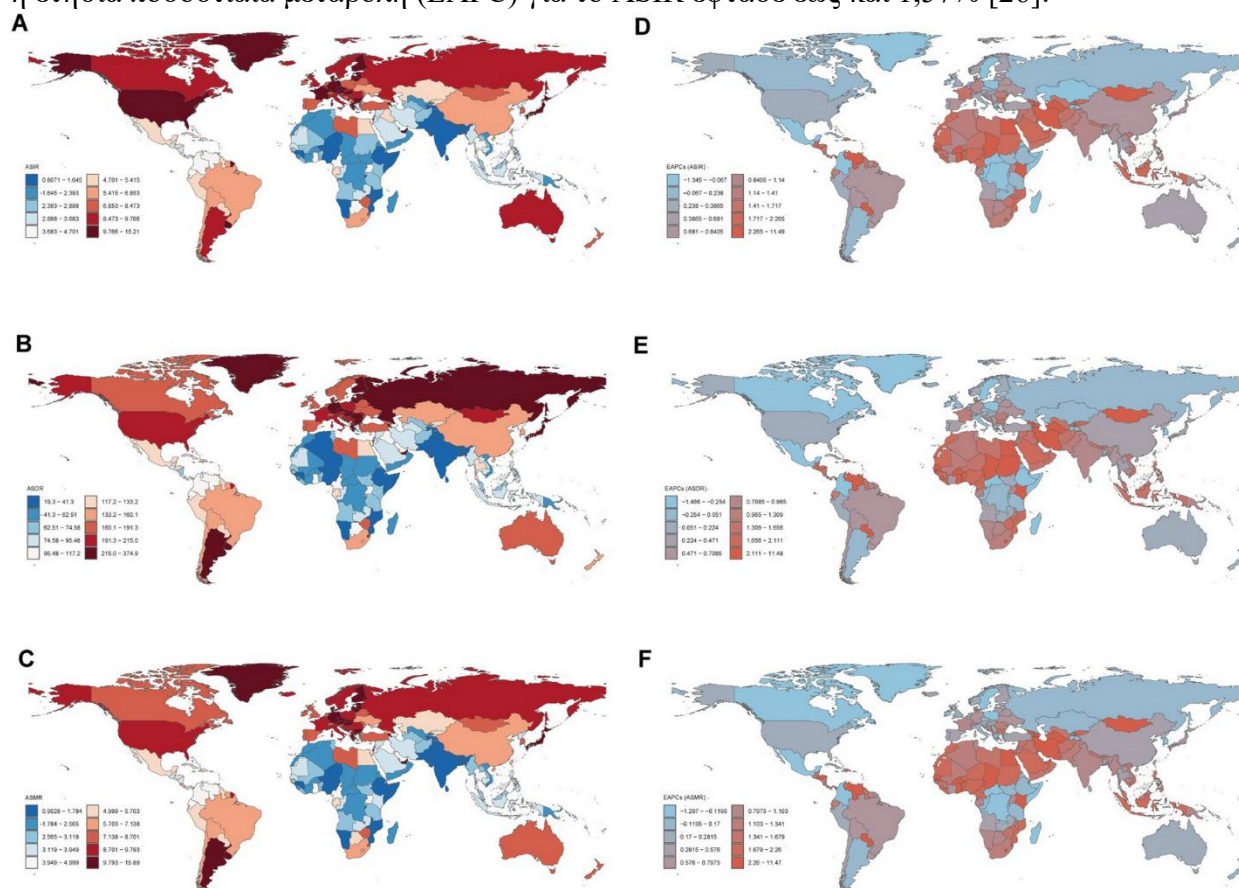
2.2. Επιδημιολογία και κατηγοριοποίηση παγκρεατικού καρκίνου

Η παγκόσμια συχνότητα εμφάνισης και θνησιμότητα του παγκρεατικού καρκίνου είναι η εξής:

Το 2020, καταγράφηκαν περίπου 495.773 νέες περιπτώσεις καρκίνου του παγκρέατος και 466.003 θάνατοι παγκοσμίως, καθιστώντας τον καρκίνο του παγκρέατος τον δωδέκατο πιο συχνά διαγιγνωσκόμενο καρκίνο, αλλά την έβδομη πιο κοινή αιτία θανάτου μεταξύ όλων των νεοπλασμάτων. Ο δείκτης επίπτωσης σταθμισμένος κατά ηλικία (ASIR) ήταν 4,9/100.000 και ο δείκτης θνησιμότητας σταθμισμένος κατά ηλικία (ASMR) ήταν 4,5/100.000. Ο ASIR και ο ASMR παρουσίασαν υψηλότερα ποσοστά στις ανεπτυγμένες χώρες (π.χ. ΗΠΑ και Ευρώπη) σε σύγκριση με τις περιοχές χαμηλού και μεσαίου εισοδήματος [43].

Από το 1990 έως το 2021, η παγκόσμια επίπτωση της πάθησης αυξήθηκε από περίπου 207.905 σε 508.533 περιπτώσεις (ASIR από 5,47 σε 5,96/100.000), ενώ ο αριθμός των θανάτων αυξήθηκε από 211.613 σε 505.752 (ASMR από 5,66 σε 5,95/100.000). Η αύξηση ήταν πιο έντονη σε περιοχές με χαμηλό-μέτριο SDI (Κοινωνικοδημογραφικός Δείκτης), όπου

η ετήσια ποσοστιαία μεταβολή (EAPC) για το ASIR έφτασε έως και 1,57% [26].



Σχήμα 3: Παγκόσμιοι χάρτες που απεικονίζουν το ποσοστό επίπτωσης σταθμισμένο κατά ηλικία (ASIR), το ποσοστό DALYs σταθμισμένο κατά ηλικία (ASDR) και το ποσοστό θνησιμότητας σταθμισμένο κατά ηλικία (ASMR) του καρκίνου του παγκρέατος το 2021 [26].

Συγκεκριμένα απεικονίζονται:

A. ASIR ανά 100.000 πληθυσμού, B. ASDR ανά 100.000 πληθυσμού και Γ. ASMR ανά 100.000 πληθυσμού. Τα διαγράμματα Δ. και ΣΤ. δείχνουν τις εκτιμώμενες ετήσιες ποσοστιαίες μεταβολές (EAPC) για τα ASIR, ASDR και ASMR αντίστοιχα, από το 1990 έως το 2021.

Οι σημαντικότεροι παράγοντες που συμβάλλουν στο παγκόσμιο βάρος της νόσου είναι η έναρξη του καπνίσματος και το έντονο κάπνισμα, ο υψηλός δείκτης μάζας σώματος (ΔΜΣ) και η χρόνια υπεργλυκαιμία (διαβήτης τύπου 2). Σύμφωνα με το BioMed Central, ο ανδρικός πληθυσμός παρουσιάζει ποσοστό DALY περίπου 1,2 φορές υψηλότερο από αυτό του γυναικείου πληθυσμού. Επιπλέον, τα άτομα άνω των 60 ετών αντιπροσωπεύουν περισσότερο από το 80% όλων των νέων διαγνώσεων [26].

Παρακάτω εξηγείται η κατηγοριοποίηση των όγκων:

- Οι παγκρεατικοί εξωκρινείς όγκοι αποτελούν περίπου το 90-95% των παρατηρούμενων περιπτώσεων. Το αδενοκαρκίνωμα του παγκρέατος (PDAC) είναι ο

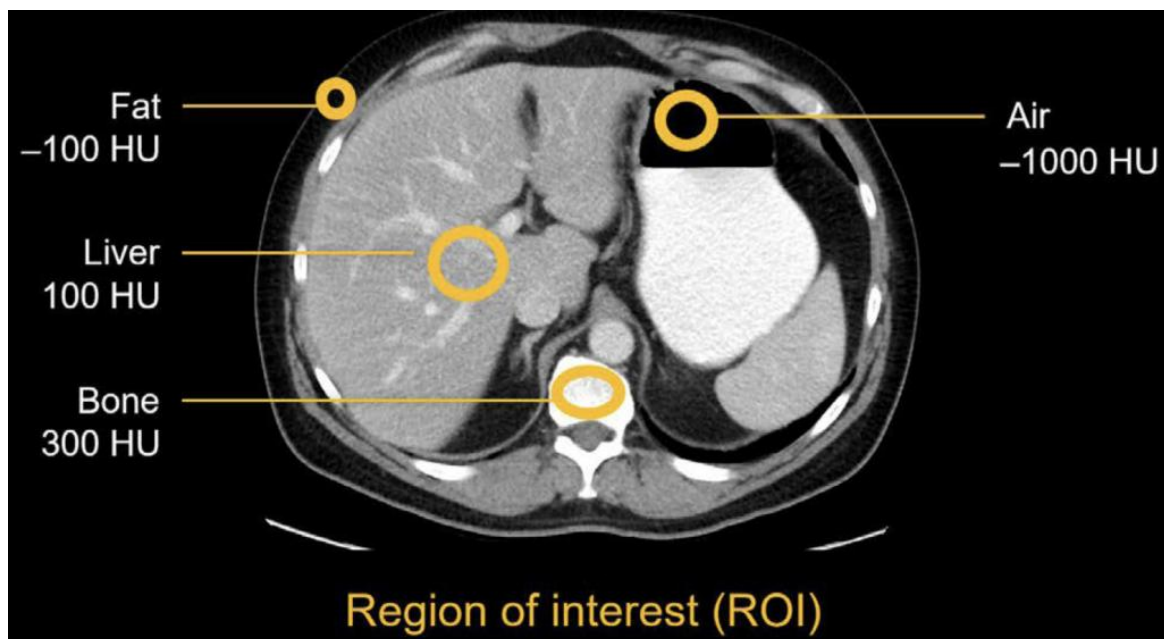
πιο διαδεδομένος ιστολογικός τύπος αδενοκαρκινώματος, με υποτύπους που περιλαμβάνουν κολλοειδές, αδενοσκαντίλλιο και αδιαφοροποίητο καρκίνωμα με γιγαντοκύτταρα οστεοκλαστών. Η ταξινόμησή τους καθορίζεται κυρίως από ιστοπαθολογικά κριτήρια και όχι από μοριακά χαρακτηριστικά.

- Οι παγκρεατικοί ενδοκρινείς όγκοι (pNETs) αντιπροσωπεύουν ποσοστό 5-10% των περιπτώσεων. Αυτά ταξινομούνται περαιτέρω σε δύο κατηγορίες: λειτουργικά (π.χ. ινσουλίνωμα, γαστρινώματα) και μη λειτουργικά. Το σύστημα βαθμολόγησης που χρησιμοποιείται βασίζεται στον δείκτη Ki-67 και στον αριθμό των μιτωτικών.

2.3. Διαγνωστική προσέγγιση και ρόλος της αξονικής τομογραφίας

2.3.1. Βασικές αρχές αξονικής τομογραφίας

Η αξονική τομογραφία (CT) είναι μια εξειδικευμένη μέθοδος απεικόνισης που χρησιμοποιεί ακτίνες X για τη δημιουργία διατομικών εικόνων των σωματικών δομών. Η βασική αρχή που διέπει αυτή τη διαδικασία βασίζεται στη διαφορική απορρόφηση της ιονίζουσας ακτινοβολίας από διαφορετικές ανατομικές δομές. Οι πληροφορίες που λαμβάνονται από τον ανιχνευτή μετατρέπονται στη συνέχεια σε ψηφιακή εικόνα μέσω της εφαρμογής του αντίστροφου μετασχηματισμού Radon (φιλτραρισμένη οπισθοπροβολή ή επαναληπτική ανακατασκευή). Η πυκνότητα κάθε εικονοστοιχείου εκφράζεται σε μονάδες Hounsfield (HU), με το νερό να έχει τιμή 0 και ο αέρας -1000 HU [22].



Σχήμα 4: Τοποθέτηση ROI (Region of Interest) για μέτρηση τιμών HU και κλίμακα βαθμολόγησης (αέρας -1000 HU, νερό 0 HU) [3].

Η εμφάνιση των τεχνολογικών εξελίξεων στους σαρωτές CT έχει επιταχύνει την εμφάνιση των σαρωτών πολυανιχνευτή CT (MDCT), οι οποίοι διευκολύνουν την ταχεία σάρωση ογκωδών δεδομένων με λεπτό πάχος τομών (0,5-1,25 mm) και δυνατότητες απεικόνισης πολλαπλών φάσεων. Η ταχύτητα περιστροφής του σωλήνα και η χωρική ανάλυση σε όλα τα

επίπεδα (αξονικό, στεφανιαίο, οβελιαίο) καθιστούν τη μέθοδο εξαιρετικά χρήσιμη για την απεικόνιση των κοιλιακών οργάνων, ειδικά του παγκρέατος.

Η χρήση της αξονικής τομογραφίας ως κύριας μεθόδου απεικόνισης του παγκρέατος οφείλεται κυρίως σε τρεις βασικούς παράγοντες. Πρώτον, η CT έχει καταστεί η προτιμώμενη μέθοδος λόγω της ευρείας προσβασιμότητάς της. Δεύτερον, η ικανότητά της για ανάλυση υψηλής αντίθεσης είναι αξιοσημείωτη. Τρίτον, η δυνατότητα πολυφασικής απεικόνισης, μετά τη χορήγηση ενός σκιαγραφικού παράγοντα, αποτελεί σημαντικό πλεονέκτημα. Η χρησιμότητά της στην ανίχνευση όγκων, στον προσδιορισμό της ανατομικής τους σχέσης με τα αιμοφόρα αγγεία και στην αξιολόγηση πιθανής μεταστατικής νόσου είναι καλά τεκμηριωμένη [45].

2.3.2. Πρωτόκολλα λήψης (contrast, phases)

Η χρήση ενδοφλέβιου ιωδιούχου σκιαγραφικού μέσου είναι απαραίτητη για τη βέλτιστη απεικόνιση του παγκρέατος, λόγω της ανάγκης διάκρισης μεταξύ παρεγχυματικών και αγγειακών δομών. Η χορήγηση σκιαγραφικού αυξάνει την αντίθεση μεταξύ φυσιολογικού ιστού και νεοπλασμάτων, ειδικά κατά την αρτηριακή και την παρεγχυματική φάση, διευκολύνοντας την ανίχνευση κακοήθων βλαβών [36].

Η πολυφασική αξονική τομογραφία (CT) αποτελείται συνήθως από τρεις φάσεις: (α) την πρώιμη αρτηριακή φάση (15-25 δευτερόλεπτα μετά την ένεση), (β) την φάση ενίσχυσης του παγκρέατος (ή παρεγχυματική φάση) (35-45 δευτερόλεπτα) και (γ) την φλεβική ή ηπατική φάση (60-70 δευτερόλεπτα). Η παρεγχυματική φάση είναι καθοριστική για τη διάγνωση του αδενοκαρκινώματος, καθώς ο όγκος εμφανίζεται ως υποπυκνή περιοχή λόγω μειωμένης αγγείωσης.

Η βέλτιστη απεικόνιση απαιτεί τη χορήγηση ιωδιούχου σκιαγραφικού παράγοντα (π.χ. ιωεξόλη) σε δόση 1,5 mL/kg με ρυθμό 3–5 mL/s με αντλία υψηλής πίεσης.

Ο συγχρονισμός φάσεων επιτυγχάνεται μέσω της εφαρμογής τεχνικών σταθερού χρονισμού ή παρακολούθησης bolus με ROI στην κοιλιακή αορτή.

Η αρτηριακή φάση διευκολύνει την αξιολόγηση της αγγειακής εμπλοκής, ενώ η φλεβική φάση είναι χρήσιμη για την επισημάνση των ηπατικών μεταστάσεων. Η ακρίβεια του προσδιορισμού της δυνατότητας εκτομής του όγκου εξαρτάται από την ολοκληρωμένη αξιολόγηση όλων των διαθέσιμων φάσεων.

Έχει αναπτυχθεί μια σειρά συγκεκριμένων πρωτοκόλλων για ομάδες υψηλού κινδύνου και έχουν χρησιμοποιηθεί τεχνικές χαμηλής δόσης.

Σε ασθενείς που θεωρούνται υψηλού κινδύνου ή σε ασθενείς που πάσχουν από χρόνιες παθήσεις, εφαρμόζονται πρωτόκολλα χαμηλής δόσης που χρησιμοποιούν επαναληπτική ανακατασκευή με στόχο την ελαχιστοποίηση της έκθεσης σε ακτινοβολία. Επιπλέον, η αρτηριακή φάση μπορεί να παραλειφθεί σε περιπτώσεις όπου η υποψία όγκου είναι χαμηλή, χωρίς σημαντική απώλεια της διαγνωστικής ακρίβειας.

2.3.3. Χρονικές και χωρικές παράμετροι (spacing, slice thickness)

1. Διάστημα (spacing) και πάχος τομής (slice thickness)

Η χωρική ανάλυση της υπολογιστικής τομογραφίας επηρεάζεται άμεσα από το πάχος της τομής, το οποίο με τη σειρά του καθορίζει την ικανότητα ανίχνευσης μικρών βλαβών. Στο πλαίσιο των πρωτοκόλλων απεικόνισης του παγκρέατος, συνιστάται η χρήση τομών πάχους όχι μεγαλύτερων από 3 χιλιοστά (συνήα 1–2,5 χιλιοστά) για την επίτευξη μιας πιο σαφούς εικόνας της υφής του παρεγχύματος και των αγγείων.

2. Isotropic voxel και ανασυνθέσεις πολλαπλών επιπέδων

Η χρήση λεπτών τομών, σε συνδυασμό με ιστροπικά voxels (π.χ. $1 \times 1 \times 1 \text{ mm}^3$), διευκολύνει τη δημιουργία πολυεπίπεδων αναδιαμορφώσεων (MPR) και τρισδιάστατων αναπαραστάσεων (3D volume rendering). Αυτές είναι ιδιαίτερα χρήσιμες για τον ακριβή εντοπισμό της σχέσης του όγκου με κρίσιμα αγγεία και για τον προεγχειρητικό σχεδιασμό.

3. Spacing και απόσταση μεταξύ τομών

Ο όρος «spacing» αναφέρεται στην απόσταση μεταξύ διαδοχικών τομών του όγκου και είναι κρίσιμος για την ακριβή ανακατασκευή του όγκου. Έχει αποδειχθεί ότι η αύξηση της απόστασης πέραν των 5 χιλιοστών έχει ως αποτέλεσμα σημαντική μείωση της ευαισθησίας στην ανίχνευση μικρών βλαβών. Αυτό, με τη σειρά του, περιπλέκει τη χρήση των δεδομένων σε τεχνικές τρισδιάστατης τμηματοποίησης. Ιδανικά, η απόσταση μεταξύ των τομών δεν πρέπει να υπερβαίνει τα 2 χιλιοστά και πρέπει να ταιριάζει με το πάχος της τομής, έτσι ώστε τα voxels να είναι ισότροπα [22].

4. Χρονικά χαρακτηριστικά και ταχύτητα σάρωσης

Η διάρκεια της σάρωσης σε σύγχρονους σαρωτές πολυαπεικονιστικής αξονικής τομογραφίας (MDCT) είναι συνήθως μεταξύ 0,5 και 1 δευτερολέπτου ανά περιστροφή του σωλήνα, επιτρέποντας την ολοκληρωμένη απεικόνιση της άνω κοιλιακής χώρας σε σύντομο χρονικό διάστημα. Η ταχύτητα σάρωσης είναι πολύ σημαντική για να εξασφαλιστεί η συνέπεια με τη φάση ενίσχυσης (αρτηριακή/παρεγχυματική) και να αποφευχθούν τεχνητά σφάλματα κίνησης, ειδικά σε μη συνεργάσιμους ασθενείς.

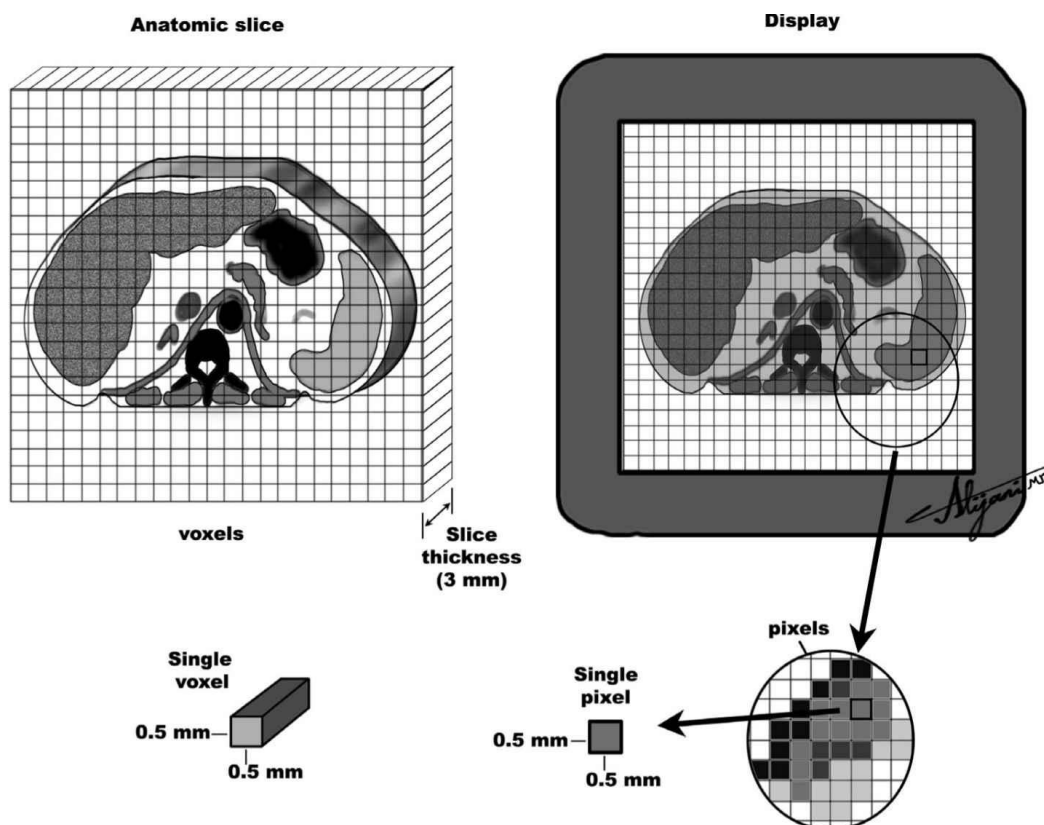
5. Επίδραση των παραμέτρων στην επεξεργασία εικόνας και τεχνητή νοημοσύνη

Η επιλογή του πάχους και του spacing των τομών έχει άμεση επίδραση στην ποιότητα των δεδομένων για την ανάλυση με τεχνικές μηχανικής μάθησης και βαθιάς μάθησης. Η ανομοιομορφία των χωρικών παραμέτρων δημιουργεί την ανάγκη για επαναδειγματοληψία, μια διαδικασία που ενδέχεται να εισαγάγει τεχνητά σφάλματα και να θέσει σε κίνδυνο την ακρίβεια των μοντέλων τμηματοποίησης. Κατά συνέπεια, η πλειονότητα των διεθνών συνόλων δεδομένων συνοδεύεται από τυποποιημένη επεξεργασία που ομαλοποιεί τις παραμέτρους σε ιστροπική ανάλυση 1 mm [1].

3. Θεωρία Επεξεργασίας Εικόνας

3.1. Τι είναι εικόνα; pixels, voxels, Hounsfield units

Μια ψηφιακή ιατρική τομογραφική εικόνα αποτελείται από βασικά «pixels» και «voxels». Το εικονοστοιχείο (picture element) αποτελεί το μικρότερο δομικό στοιχείο μιας δισδιάστατης τομής, καταγράφοντας την ένταση της απορρόφησης των ακτίνων Χ. Ένα εικονοκύτταρο (volume element) υποδηλώνει την τρισδιάστατη επέκταση ενός εικονοστοιχείου, με διαστάσεις πλάτος $\times$ ύψος $\times$ πάχος τομής, που αντιπροσωπεύει τον μέσο γραμμικό συντελεστή απορρόφησης εντός αυτού του όγκου. Σε κάθε voxel αποδίδεται μια αριθμητική τιμή που μετατρέπεται σε μονάδες Hounsfield (HU) χρησιμοποιώντας μια γραμμική κλίμακα, όπου 0 HU αντιπροσωπεύει το νερό και -1000 HU αντιπροσωπεύει τον αέρα, όπως αναφέρθηκε και παραπάνω. Αυτό επιτρέπει την ποιοτική και ποσοτική διαφοροποίηση των διαφόρων ιστών και βλαβών [5].



Σχήμα 5: Σχηματική αναπαράσταση μιας ανατομικής τομής (αριστερά) και η αντίστοιχη απεικόνιση της εικόνας (δεξιά) που δείχνει τυπικά εικονοστοιχεία και εικονοκύτταρα στην απεικόνιση CT [46].

Η μετατροπή των ακατέργαστων τιμών pixel (SV) σε Hounsfield Units (HU) επιτυγχάνεται με τη χρήση των σταθερών τιμών κλίσης (“Rescale Slope”, m) και τεμνόμενου (“Rescale Intercept”, b), που καταγράφονται στην κεφαλίδα DICOM (ετικέτες (0028,1053) και (0028,1052)), σύμφωνα με τον τύπο:

$$HU = m \times SV + b$$

Αυτός ο μετασχηματισμός διασφαλίζει ότι οι τιμές έντασης είναι συγκρίσιμες μεταξύ διαφορετικών κατασκευαστών και πρωτοκόλλων CT, διευκολύνοντας την ομαλή εισροή καθορισμένων δεδομένων σε αλγορίθμους μηχανικής μάθησης [32].

Η κλίμακα HU βασίζεται στον γραμμικό συντελεστή απορρόφησης ($\mu_{\text{μυμ}}$), όπου κάθε τιμή υπολογίζεται ως:

$$1000 \times \frac{\mu_{\text{ύλης}} - \mu_{\text{νερού}}}{\mu_{\text{νερού}} - \mu_{\text{αέρα}}}$$

επιτρέποντας κλινικά σημαντικές τιμές: π.χ. λίπος (~ -100 HU), μαλακοί ιστοί ($\sim 30-70$ HU), οστό (> 300 HU). Αυτή η κωδικοποίηση προσδίδει ευκρίνεια στις νεοπλασματικές δομές, οι οποίες συνήθως εμφανίζονται με HU τιμές διαφορετικές από το φυσιολογικό παρέγχυμα.

Η ποιότητα των CT εικόνων και η ακρίβεια στην επεξεργασία των voxels πρέπει να ελέγχονται διασφαλίζοντας την ομοιομορφία των τιμών HU πριν από οποιαδήποτε ανάλυση τμηματοποίησης ή ποσοτικοποίησης. Η κακή διόρθωση των ετικετών αναπροσαρμογής μπορεί να οδηγήσει σε σφάλματα στη μέτρηση του όγκου και στη στατιστική ανάλυση. Αντίθετα, η ρύθμιση παραθύρου (π.χ. πλάτος και επίπεδο παραθύρου) προσαρμόζει την οπτικοποίηση HU για να επισημάνει με τον βέλτιστο τρόπο συγκεκριμένους ιστούς ή παθολογίες.

Συνοπτικά, είναι πολύ σημαντικό να κατανοήσουμε τα pixel και τα voxel ως ποσοτικές δομικές μονάδες, που μετατρέπονται με ακρίβεια σε μονάδες Hounsfield μέσω παραμέτρων ανακλιμάκωσης DICOM, προκειμένου να διασφαλιστεί η αξιόπιστη επεξεργασία εικόνων, η εκπαίδευση αλγορίθμων βαθιάς μάθησης και η κλινική διάγνωση.

3.2. Interpolation και resampling

Στον τομέα της επεξεργασίας ιατρικών εικόνων, η επαναδειγματοληψία είναι μια διαδικασία που περιλαμβάνει την προσαρμογή χωρικών παραμέτρων, όπως το μέγεθος των εικονοστοιχείων (pixel spacing) και το διάστημα (spacing) μεταξύ των τομών, με στόχο να διασφαλιστεί ότι όλα τα δεδομένα εντός ενός δεδομένου συνόλου έχουν ομοιογενείς διαστάσεις εικονοστοιχείων. Αυτή η μετατροπή επιτυγχάνεται μέσω της εφαρμογής τεχνικών παρεμβολής, οι οποίες μετατρέπουν τα αρχικά voxels σε ένα νέο πλέγμα που χαρακτηρίζεται από προκαθορισμένες διαστάσεις, συνήθως ισοτροπικής φύσης (π.χ. $1 \times 1 \times 1 \text{ mm}^3$). Αυτή η τυποποίηση έχει αποδειχθεί κρίσιμη για τη σταθερότητα και την επαναληψιμότητα των ραδιομικών δεικτών, καθώς οι αλλαγές στο μέγεθος των voxels μπορούν να προκαλέσουν σημαντικές διακυμάνσεις στις τιμές των χαρακτηριστικών που εξάγονται από αλγόριθμους τμηματοποίησης ή ταξινόμησης [48].

Οι αλγόριθμοι παρεμβολής ταξινομούνται με βάση την πολυπλοκότητά τους και την επίδραση που έχουν στην τελική ποιότητα της εικόνας. Έχει αποδειχθεί ότι η μέθοδος παρεμβολής πλησιέστερου γείτονα είναι η ταχύτερη και διατηρεί την κατανομή του θορύβου.

Ωστόσο, έχει επίσης αποδειχθεί ότι αυτή η μέθοδος μπορεί να δημιουργήσει «μπλοκ» τεχνητά στοιχεία. Η γραμμική και η κυβική παρεμβολή spline προσφέρουν ομαλότερες μεταβάσεις έντασης, μειώνοντας τα ορατά τεχνητά στοιχεία. Αντίθετα, η παρεμβολή Lanczos (με βάση sinc) έχει αποδειχθεί ότι είναι η πιο ακριβής μέθοδος για ιατρικές εφαρμογές, καθώς περιορίζει την απώλεια ακρίβειας και τα σφάλματα σε σύγκριση με απλούστερες τεχνικές [34].

Η μέθοδος παρεμβολής που επιλέγουμε κατά τη μετατροπή ή την ανάλυση των ιατρικών εικόνων επηρεάζει άμεσα την ποιότητα και τις τιμές των voxels. Στη ραδιομική, από αυτές τις τιμές υπολογίζουμε χιλιάδες χαρακτηριστικά, όπως δείκτες υφής και μορφομετρικά μεγέθη. Διαφορετικοί αλγόριθμοι παρεμβολής (π.χ. γραμμική, κυβική) μπορούν να «μεταβάλλουν» αυτά τα χαρακτηριστικά, οδηγώντας σε διαφορετικά αποτελέσματα.

Για να αποφύγουμε τέτοιες ασυνεπείς διαφοροποιήσεις, η Image Biomarker Standardization Initiative (IBSI) προτείνει συγκεκριμένες ρυθμίσεις επεξεργασίας, όπως:

- Κυβική παρεμβολή (cubic interpolation)
- Τυποποιημένο διάστημα voxel (voxel spacing): $3 \times 3 \times 3 \text{ mm}^3$

Επίσης, κάθε δημοσίευση πρέπει να αναφέρει λεπτομερώς αυτές τις παραμέτρους. Έτσι, διασφαλίζεται ότι τα αποτελέσματα είναι διαφανή και μπορούν να αναπαραχθούν από άλλους ερευνητές [48].

Η εφαρμογή της ιστροπικής επαναδειγματοληψίας διευκολύνει την ανάπτυξη και την εκπαίδευση μοντέλων βαθιάς μάθησης, καθώς εξαλείφει την ανάγκη για πρόσθετες προσαρμογές κατά την εισαγωγή δεδομένων σε δίκτυα. Ωστόσο, είναι σημαντική η προσοχή κατά τη διάρκεια αυτής της διαδικασίας, καθώς η υπερβολική παρεμβολή μπορεί να οδηγήσει σε υπερβολική εξομάλυνση των λεπτομερειών της εικόνας, θέτοντας έτσι σε κίνδυνο την αναγνώριση λεπτών παθολογικών δομών. Στις σύγχρονες ροές εργασίας, συνιστάται η αξιολόγηση της αποτελεσματικότητας κάθε μεθόδου χρησιμοποιώντας μετρήσεις όπως ο Δείκτης Δομικής Ομοιότητας (SSIM) ή ο Δείκτης Σημείου προς Θόρυβο (PSNR) πριν από την εφαρμογή της σε ολόκληρο το σύνολο δεδομένων [47].

3.3. Spacing και anisotropy

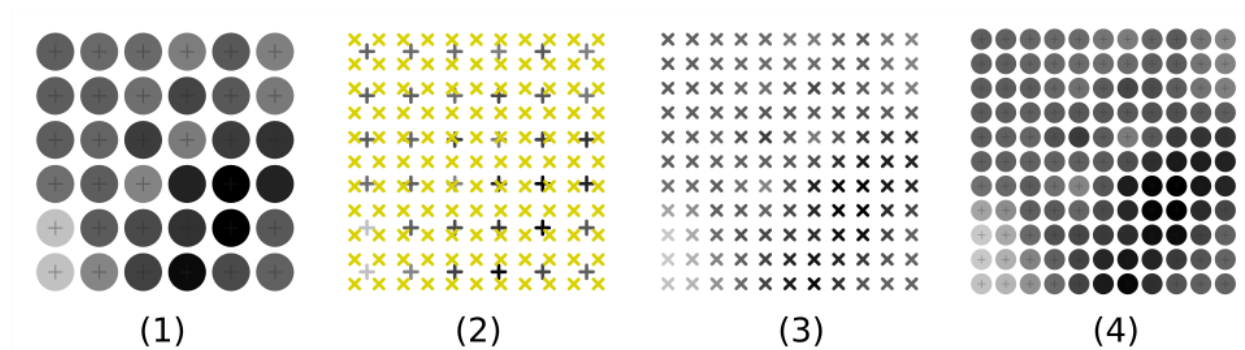
Το διάστημα (spacing) μεταξύ διαδοχικών τομών και η ανισοτροπία των εικονοστοιχείων είναι κρίσιμες παράμετροι για την ακρίβεια της επεξεργασίας ιατρικών εικόνων και την αξιοπιστία των ραδιομικών χαρακτηριστικών. Όταν το πάχος μιας τομής ή το διάστημα μεταξύ των τομών διαφέρει σημαντικά από την ανάλυση στο επίπεδο, δημιουργούνται συστηματικές αποκλίσεις στις μετρήσεις υφής και σχήματος. Μελέτες έχουν δείξει ότι οι αλλαγές στο πάχος της τομής από 2 mm σε 5 mm μπορούν να μεταβάλλουν τις σχετικές τιμές ορισμένων χαρακτηριστικών έως και 25%, ακόμη και μετά την βέλτιστη επιλογή των παραμέτρων ανακατασκευής της εικόνας [41]. Επιπλέον, η διάκριση μεταξύ ισότροπης και ανισότροπης ανάλυσης απαιτεί επαναδειγματοληψία σε ισότροπικά μεγέθη voxel (συνήθως $1 \times 1 \times 1 \text{ mm}^3$) πριν από την εξαγωγή χαρακτηριστικών, προκειμένου να διασφαλιστεί η

περιστροφική αμεταβλητότητα και να αποφευχθούν τεχνητά στοιχεία που υπονομεύουν την απόδοση των μοντέλων βαθιάς μάθησης. Ταυτόχρονα, έχει αποδειχθεί ότι η διακύμανση στην απόσταση μεταξύ των εικονοστοιχείων μεταξύ διαφορετικών σαρωτών CT επηρεάζει σημαντικά τα χαρακτηριστικά υφής. Αυτό απαιτεί την εναρμόνιση των εικόνων μέσω τυποποιημένων πρωτοκόλλων και συμπληρωματικών μεθόδων εναρμόνισης, όπως το ComBat, για την ελαχιστοποίηση της τεχνικής διακύμανσης και την ενίσχυση της γενικευσιμότητας των μοντέλων σε πολυκεντρικά σύνολα δεδομένων [4].

3.4. Intensity normalization και windowing

Στον τομέα της ιατρικής απεικόνισης, η κανονικοποίηση της έντασης και οι τεχνικές windowing είναι διαδικασίες ζωτικής σημασίας για τη διασφάλιση της σταθερότητας και της ακρίβειας των αλγορίθμων ανάλυσης δεδομένων. Πρώτον, οι τιμές Hounsfield Units (HU) σε κάθε εικόνα υποβάλλονται συχνά σε περικοπή εκατοστιαίων θέσεων για να μειωθεί η επίδραση των ακραίων τιμών και να αυξηθεί η σχετική αντίθεση των ιστών. Στη συνέχεια, εφαρμόζεται κανονικοποίηση z-score με βάση τον μέσο όρο και την τυπική απόκλιση των voxels εντός της περιοχής ενδιαφέροντος, διασφαλίζοντας ότι οι εικόνες έχουν κοινή κατανομή έντασης πριν από την εκπαίδευση. Επιπλέον, η εφαρμογή του windowing επιτρέπει την προσαρμογή του οπτικού δυναμικού εύρους (πλάτος παραθύρου) και του κέντρου έντασης (επίπεδο παραθύρου) για την αποτελεσματικότερη επισήμανση συγκεκριμένων ιστών ή παθολογιών. Για παράδειγμα, το «παγκρεατικό παράθυρο» (πλάτος: 300 HU, επίπεδο: 40 HU) βελτιστοποιεί την οπτική διακρίσιμότητα του παγκρεατικού παρεγχύματος σε σχέση με τους όγκους. Αυτή η προσέγγιση αποτελεί κρίσιμο στοιχείο στην κλινική ερμηνεία πριν από οποιαδήποτε ψηφιακή επεξεργασία [11].

Είναι προφανές ότι η συνεπής εφαρμογή αυτών των βημάτων είναι υψίστης σημασίας για να διασφαλιστεί ότι οι διαφορές μεταξύ σαρώσεων, μηχανημάτων και πρωτοκόλλων δεν θέτουν σε κίνδυνο την ακεραιότητα των μοντέλων βαθιάς μάθησης και των ποσοτικών ραδιομικών δεικτών.



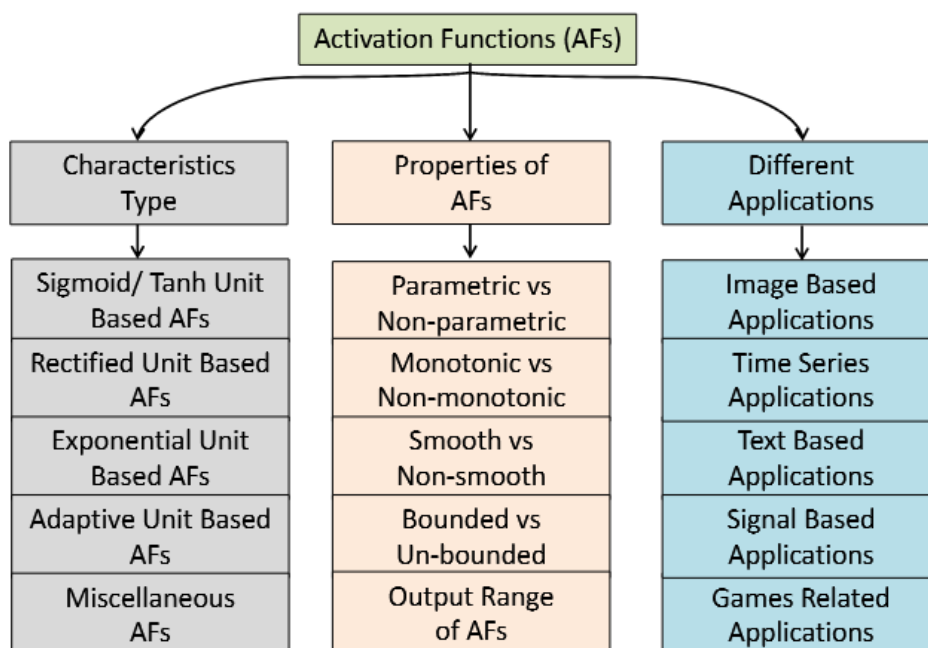
Σχήμα 6: Σχηματικό βήμα-προς-βήμα επαναδειγματοληψίας (resampling) [13].

4. Θεωρία Νευρωνικών Δικτύων

4.1. Βασικές έννοιες: νευρώνας, συνάρτηση ενεργοποίησης

Στο πλαίσιο των τεχνητών νευρωνικών δικτύων, ένας τεχνητός νευρώνας λειτουργεί ως η βασική μονάδα επεξεργασίας. Λαμβάνει τιμές εισόδου ($x_1, x_2, \dots, x_n$), τις πολλαπλασιάζει με τους αντίστοιχους συντελεστές στάθμισης $w_1, w_2, \dots, w_n$, προσθέτει έναν όρο μεροληψίας (b) (bias) και εφαρμόζει μια μη γραμμική συνάρτηση ενεργοποίησης για να παράγει την έξοδό του. Η μη γραμμική συνάρτηση ενεργοποίησης (σ) είναι κρίσιμη για τον προσδιορισμό της συμπεριφοράς του νευρώνα, επιτρέποντάς του έτσι να μαθαίνει σύνθετα, μη γραμμικά μοτίβα από τα δεδομένα. Εάν δεν υπήρχε αυτή η μη γραμμικότητα, το δίκτυο θα ήταν ισοδύναμο με ένα απλό γραμμικό μοντέλο, ανεξάρτητα από τον αριθμό των επιπέδων που περιέχει [42].

Η εξέλιξη των συναρτήσεων ενεργοποίησης έχει καθοδηγηθεί από την ανάγκη να αντιμετωπιστούν διάφορες απαιτήσεις εκπαίδευσης. Οι σιγμοειδείς και οι Tanh ήταν μεταξύ των αρχικών επιλογών που εξετάστηκαν, ωστόσο παρουσιάζουν προκλήσεις όπως η εξαφάνιση της κλίσης σε βαθιά δίκτυα. Οι διορθωμένες γραμμικές μονάδες (ReLU) εισήχθησαν για να αντιμετωπίσουν αυτά τα ζητήματα, προσφέροντας επιταχυνόμενη σύγκλιση και πιο απλό υπολογισμό $\text{ReLU}(x) = (\max(0, x))$. Πιο πρόσφατες συναρτήσεις, όπως ELU, Swish και Mish, επιδιώκουν να συνδυάσουν την ταχύτητα της ReLU με ομαλή κορεσμό και βελτιωμένη ροή κλίσης. Μια μελέτη 18 διαφορετικών συναρτήσεων ενεργοποίησης έδειξε ότι η επιλογή της βέλτιστης συνάρτησης μπορεί να βελτιώσει την απόδοση και τη σταθερότητα της εκπαίδευσης, ανάλογα με την αρχιτεκτονική και το σύνολο δεδομένων [9].



Σχήμα 7: Ταξινόμηση των συναρτήσεων ενεργοποίησης [10].

Μια ιστορική αναδρομή στα τεχνητά νευρωνικά δίκτυα δείχνει ότι αρχικά σχεδιάστηκαν ως μαθηματικό μοντέλο βιολογικών νευρώνων, με απλές εισόδους και εξόδους. Η εισαγωγή του μηχανισμού οπισθοδρόμησης σφάλματος (backpropagation) επέτρεψε την αυτόματη ρύθμιση των βαρών μέσω διαφορίσιμων συναρτήσεων ενεργοποίησης. Αρχικά αναπτύχθηκαν ρηχές αρχιτεκτονικές (shallow networks). Με την αύξηση του αριθμού των στρωμάτων, προέκυψαν βαθιές αρχιτεκτονικές (deep networks). Τα βαθιά δίκτυα μπορούν πλέον να αναγνωρίζουν ολοένα και πιο περίπλοκα χαρακτηριστικά σε πολλαπλά επίπεδα αφαίρεσης: από χαμηλού επιπέδου ακμές και υφές έως υψηλού επιπέδου έννοιες και αντικείμενα [42].

4.2. Δομή δικτύου και forward/backward propagation (διάδοση προς τα εμπρός και προς τα πίσω)

Σε ένα τυπικό πολυεπίπεδο δίκτυο, οι νευρώνες οργανώνονται σε διαδοχικά επίπεδα. Το πρώτο στρώμα, που ονομάζεται επίπεδο εισόδου, λαμβάνει ακατέργαστα δεδομένα. Τα ενδιάμεσα στρώματα, γνωστά και ως κρυφά επίπεδα, εξάγουν όλο και πιο αφηρημένα μοτίβα. Τέλος, το επίπεδο εξόδου παράγει προβλέψεις. Κάθε επίπεδο αποτελείται από πολλαπλούς νευρώνες που είναι πλήρως ή τοπικά συνδεδεμένοι με το επόμενο επίπεδο μέσω σταθμισμένων συνδέσεων. Κατά τη διάρκεια της προώθησης προς τα εμπρός, οι είσοδοι πολλαπλασιάζονται με τα τρέχοντα βάρη, προστίθεται ο όρος μεροληψίας (b) και εφαρμόζεται μια συνάρτηση ενεργοποίησης. Αυτό έχει ως αποτέλεσμα η έξοδος κάθε επιπέδου να τροφοδοτεί το επόμενο, επιτρέποντας στα δίκτυα να καταγράφουν πολύπλοκες μη γραμμικές σχέσεις στα δεδομένα.

Στο πλαίσιο της εκπαίδευσης, το δίκτυο χρησιμοποιεί μια συνάρτηση κόστους για να συγκρίνει την τελική έξοδο με την επιθυμητή απόκριση, υπολογίζοντας έτσι το συνολικό σφάλμα. Η οπισθοδρόμηση είναι μια διαδικασία που χρησιμοποιεί τον κανόνα της αλυσίδας για να διαδώσει το σφάλμα από το επίπεδο εξόδου στο επίπεδο εισόδου. Περιλαμβάνει τον υπολογισμό των μερικών παραγώγων του σφάλματος σε σχέση με κάθε βάρος. Στη συνέχεια, τα βάρη ενημερώνονται με βάση τη σχέση:

$w \leftarrow w - \eta \cdot \partial L / \partial w$, όπου η είναι ο ρυθμός μάθησης. Το πρωτοποριακό έργο των Rumelhart et al. ήταν το πρώτο που χρησιμοποίησε δίκτυα οπισθοδρόμησης, αποδεικνύοντας την αποτελεσματικότητα της αυτόματης ρύθμισης των βαρών με βάση τις κλίσεις, θέτοντας έτσι τα θεμέλια για αξιόπιστη σύγκλιση [40].

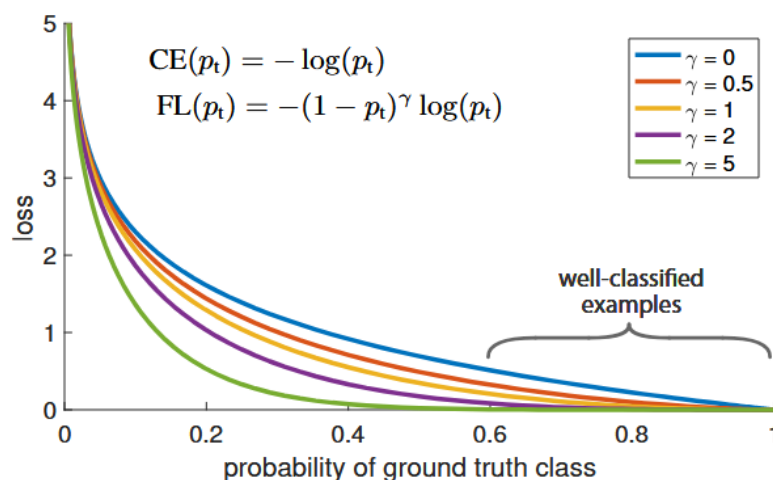
Ωστόσο, τα βαθιά δίκτυα αντιμετωπίζουν προκλήσεις, όπως το φαινόμενο της εξαφάνισης των κλίσεων, όπου οι κλίσεις στα αρχικά στρώματα πλησιάζουν το μηδέν, με αποτέλεσμα την εξασθένιση των μαθησιακών ικανοτήτων ή ακόμη και την αδυναμία μάθησης. Η προσέγγιση Efficient BackProp των LeCun et al. εισήγαγε μια σειρά βελτιστοποιήσεων, όπως η κατάλληλη αρχικοποίηση βαρών (Xavier/He init), ο προσαρμοστικός προγραμματισμός του ρυθμού μάθησης και οι τεχνικές ορμής. Αυτές οι τροποποιήσεις οδήγησαν σε πολλά σημαντικά οφέλη, όπως επιταχυνόμενη εκπαίδευση, βελτιωμένη

σταθερότητα σύγκλισης και μειωμένη ευπάθεια σε απώλειες πληροφοριών κατά την ανάστροφη διέλευση [25].

Τα σύγχρονα πλαίσια βασίζονται σε αυτά τα θεμέλια, συνδυάζοντας κανονικοποίηση ανά παρτίδα (batch normalization), τεχνικές όπως το dropout για ρύθμιση (regularization) και προηγμένους βελτιστοποιητές, ώστε τα μοντέλα να επιτυγχάνουν υψηλή απόδοση σε μεγάλα, σύνθετα σύνολα δεδομένων. Υπάρχει ανάγκη να κατανοήσουμε την αλληλεπίδραση μεταξύ της αρχιτεκτονικής του δικτύου, της προώθησης/οπισθοδρόμησης και των τεχνικών βελτιστοποίησης, ώστε να σχεδιάσουμε αποτελεσματικά νευρωνικά δίκτυα για εφαρμογές όπως η ιατρική απεικόνιση.

4.3. Εκπαίδευση: loss functions, βελτιστοποίηση, regularization

Στο πλαίσιο της εκπαίδευσης νευρωνικών δικτύων για ιατρική τμηματοποίηση, η επιλογή και ο σχεδιασμός της συνάρτησης κόστους είναι σημαντικά για την επίτευξη υψηλής ακρίβειας. Πέρα από την ποσοτικοποίηση των διαφορών σε επίπεδο εικονοστοιχείων, η απώλεια Dice δίνει προτεραιότητα στην ομοιότητα μεταξύ της προβλεπόμενης και της πραγματικής μάσκας, αντιμετωπίζοντας αποτελεσματικά την ανισορροπία των κλάσεων. Η απώλεια Focal, από την άλλη πλευρά, δίνει έμφαση στην εκπαίδευση σε δείγματα που είναι δύσκολο να ταξινομηθούν [27].



Σχήμα 8: Focal Loss, η οποία προσθέτει έναν συντελεστή $(1 - p_t)^\gamma$ στο τυπικό κριτήριο διασταυρούμενης εντροπίας. Ο καθορισμός του $\gamma > 0$ μειώνει τη σχετική απώλεια για τα καλά ταξινομημένα παραδείγματα ($p_t > .5$), δίνοντας μεγαλύτερη έμφαση στα δύσκολα, λανθασμένα ταξινομημένα παραδείγματα [14].

Τα βάρη προσαρμόζονται με εξελιγμένους αλγόριθμους βελτιστοποίησης. Ο αλγόριθμος Adam ενσωματώνει προσαρμοστικούς ρυθμούς μάθησης με στοιχεία ορμής, επιταχύνοντας έτσι τη σύγκλιση και ελαχιστοποιώντας την ανάγκη για χειροκίνητη προσαρμογή του ρυθμού μάθησης. Παραλλαγές όπως ο αλγόριθμος AdamW διαχωρίζει την αποσβεστική ποινή

βάρους από την ενημέρωση των βαρών, ώστε να έχει πιο προβλέψιμη συμπεριφορά regularization [24].

Ταυτόχρονα, η χρήση τεχνικών κανονικοποίησης είναι απαραίτητη για την αποφυγή της υπερπροσαρμογής και την ενίσχυση της γενικευσιμότητας των αποτελεσμάτων του μοντέλου σε κλινικά περιβάλλοντα. Ένα χαρακτηριστικό παράδειγμα είναι η μέθοδος dropout, η οποία περιλαμβάνει την τυχαία απενεργοποίηση ενός ποσοστού νευρώνων κατά τη φάση της εκπαίδευσης. Αυτή η διαδικασία ενισχύει το δίκτυο έναντι εξωτερικών διαταραχών ή θορύβου. Η κανονικοποίηση βαρών χρησιμεύει για να επιβάλει όρια στις λιγιστές τιμές των βαρών, εξασφαλίζοντας μια ισορροπημένη κατανομή σε όλο το δίκτυο. Η ομαλοποίηση παρτίδων, από την άλλη πλευρά, λειτουργεί για να σταθεροποιήσει την κατανομή των εισροών σε κάθε στρώμα, επιταχύνοντας έτσι τη διαδικασία μάθησης και μειώνοντας την ευαισθησία του μοντέλου στις αρχικές ρυθμίσεις. Η εφαρμογή της πρόωρης διακοπής και της αύξησης των δεδομένων ενισχύει περαιτέρω τη σταθερότητα της απόδοσης του μοντέλου όταν αξιολογείται σε ανεξάρτητα σύνολα δοκιμών.

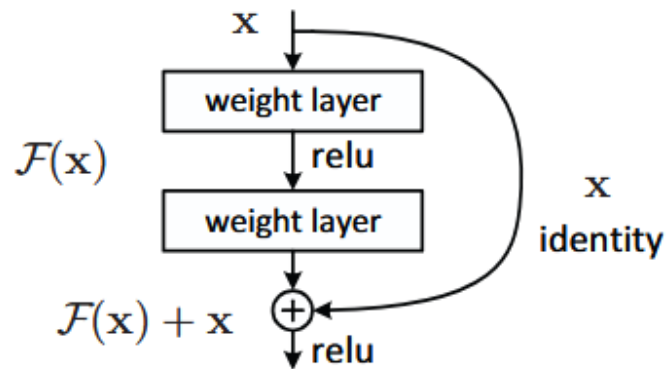
4.4. Συνελικτικά νευρωνικά δίκτυα (CNNs) και ειδικές αρχιτεκτονικές

Τα συνελικτικά νευρωνικά δίκτυα (CNNs) χρησιμοποιούν κοινή χρήση βαρών (weight sharing) για να εξάγουν χαρακτηριστικά από εικόνες. Συγκεκριμένα, κάθε συνελικτικό φίλτρο μετακινείται κατά μήκος της εισόδου, ανιχνεύοντας μοτίβα χαμηλού επιπέδου, όπως άκρα και υφές. Οι χάρτες χαρακτηριστικών που προκύπτουν τροφοδοτούνται στη συνέχεια σε επόμενα επίπεδα για να δημιουργήσουν ιεραρχικές αναπαραστάσεις υψηλού επιπέδου [15]. Η ενσωμάτωση στρωμάτων συγκέντρωσης, όπως η μέγιστη συγκέντρωση, έχει αποδειχθεί ότι μειώνει τη διαστατικότητα των χαρακτηριστικών και ενισχύει την ανθεκτικότητα σε μικρές μετατοπίσεις.

Τα πλήρως συνελικτικά δίκτυα (FCN) μετατρέπουν την κλασική δομή CNN σε δίκτυα τμηματοποίησης από άκρο σε άκρο. Αυτή η μετατροπή περιλαμβάνει την αντικατάσταση πλήρως συνδεδεμένων επιπέδων με συνελικτικά επίπεδα, επιτρέποντας έτσι την πρόβλεψη σε επίπεδο εικονοστοιχείων, διατηρώντας παράλληλα τις παραμέτρους των χωρικών πληροφοριών. Οι Ronneberger et al. ανέπτυξαν το U-Net, ένα συμμετρικό δίκτυο κωδικοποιητή αποκωδικοποιητή με συνδέσεις παράκαμψης (skip connections) που διατηρούν τις τοπικές λεπτομέρειες κατά την ανακατασκευή χαρακτηριστικών στη διαδρομή του αποκωδικοποιητή. Το U-Net αποδεικνύει αξιοσημείωτη αποτελεσματικότητα ακόμη και με περιορισμένο αριθμό δειγμάτων εκπαίδευσης, επιτυγχάνοντας μια σημαντική επιτυχία στην πρόκληση τμηματοποίησης ISBI [38].

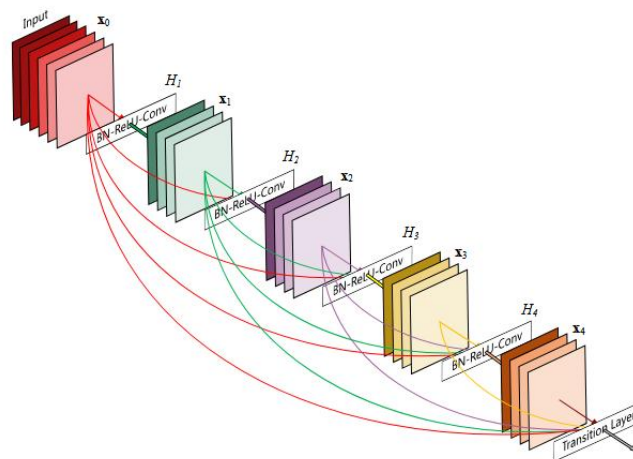
Η ανάγκη για πιο εκτεταμένα δίκτυα οδήγησε στην ανάπτυξη υπολειμματικών μπλοκ, όπως το ResNet, το οποίο ενσωματώνει συντομεύσεις ταυτότητας (απλές συνδέσεις που παρακάμπτουν ένα ή περισσότερα ενδιάμεσα στρώματα) για να επιτρέπει στο δίκτυο να μαθαίνει υπολειμματικές λειτουργίες αντί για πλήρεις μετασχηματισμούς. Αυτή η απλή

επέκταση επέτρεψε την εκπαίδευση δικτύων με εκατοντάδες στρώματα χωρίς το πρόβλημα της εξαφάνισης της κλίσης, επιτυγχάνοντας ακρίβεια στο ImageNet [7].



Σχήμα 9: Υπολειμματική μάθηση: ένα δομικό στοιχείο [17].

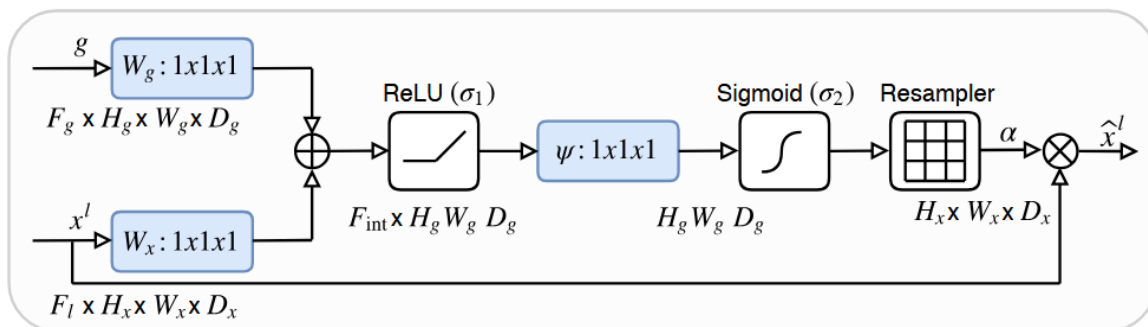
Για να βελτιώσει περαιτέρω τη ροή πληροφοριών και την επαναχρησιμοποίηση χαρακτηριστικών, το DenseNet εισήγαγε πυκνές συνδέσεις μεταξύ όλων των επιπέδων ενός μπλοκ, επιτυγχάνοντας σημαντική βελτίωση της ροής κλίσης, μείωση των παραμέτρων και βελτιωμένη απόδοση σε βασικά σύνολα αναφοράς όπως το CIFAR και το ImageNet [19].



Σχήμα 10: Ένα πυκνό μπλοκ 5 επιπέδων με ρυθμό ανάπτυξης $k = 4$. Κάθε επίπεδο λαμβάνει ως είσοδο όλους τους προηγούμενους χάρτες χαρακτηριστικών [8].

Επιπλέον, οι αρχιτεκτονικές που χρησιμοποιούν μηχανισμούς προσοχής, όπως το Attention U-Net, ενσωματώνουν πύλες προσοχής που έχουν εκπαιδευτεί να εντοπίζουν δυναμικά

περιοχές ενδιαφέροντος σε χάρτες χαρακτηριστικών, ενώ ταυτόχρονα καταστέλλουν τις άσχετες περιοχές χωρίς να αυξάνουν σημαντικά το υπολογιστικό κόστος. Αυτή η προσέγγιση έχει αποδειχθεί ιδιαίτερα αποτελεσματική στην αντιμετώπιση των προκλήσεων της τμηματοποίησης στην ιατρική απεικόνιση, βελτιώνοντας τόσο την ευαισθησία όσο και την ακρίβεια των προβλέψεων.



Σχήμα 11: Σχηματική απεικόνιση της προτεινόμενης πύλης προσθετικής προσοχής (AG). Τα χαρακτηριστικά εισόδου (x^l) κλιμακώνονται με συντελεστές προσοχής (α) που υπολογίζονται στην AG. Οι χωρικές περιοχές επιλέγονται με την ανάλυση τόσο των ενεργοποιήσεων όσο και των πληροφοριών συμφοραζομένων που παρέχονται από το σήμα πύλης (g), το οποίο συλλέγεται από μια πιο χονδροειδή κλίμακα. Η επαναδειγματοληψία πλέγματος των συντελεστών προσοχής πραγματοποιείται με τριγραμμική παρεμβολή [35].

5. To nnU-Net Framework

5.1. Ιστορικό και γενική φιλοσοφία του nnU-Net

Το 2020, ο Fabian Isensee και η ομάδα του στο Πανεπιστήμιο της Χαϊδελβέργης παρουσίασαν το nnU-Net (no-new-Net) ως μια πλήρως αυτόνομη λύση για την τμηματοποίηση ιατρικών εικόνων. Αυτή η λύση βασίζεται στην αρχιτεκτονική U-Net. Σε αυτό το πλαίσιο γίνεται εμφανές ότι δεν είναι απαραίτητες νέες ή πιο σύνθετες αρχιτεκτονικές για την επίτευξη καλής απόδοσης. Αντιθέτως, απαιτείται μια ενοποιημένη, ισχυρά αυτορυθμιζόμενη διαδικασία, η οποία προσαρμόζει τη διαμόρφωση του μοντέλου, την προεπεξεργασία, την εκπαίδευση και την αξιολόγηση με βάση τα χαρακτηριστικά κάθε συνόλου δεδομένων [21].

Η φιλοσοφία του nnU-Net αμφισβητεί την επικρατούσα τάση της υπερβολικής τεχνικής εξειδίκευσης, προτείνοντας έναν νέο κανόνα σχεδιασμού για κάθε σύνολο δεδομένων. Το πλαίσιο χρησιμοποιεί αυτόματη ανάλυση παραμέτρων όπως η απόσταση, οι διαστάσεις της εικόνας, η στατιστική κατανομή έντασης και η χωρική ανάλυση, επιτρέποντας την προσαρμογή σε πραγματικό χρόνο της χωρικής ανάλυσης (απόσταση στόχου), του μεγέθους του patch, των επιλογών μεγέθους παρτίδας και των τεχνικών επαύξησης δεδομένων. Το nnU-Net δημιουργεί βέλτιστες ρυθμίσεις εκπαίδευσης χωρίς την ανάγκη ανθρώπινης

παρέμβασης. Αυτές οι ρυθμίσεις προέρχονται από εμπειρικά καθορισμένους κανόνες και ευριστικές μεθόδους που έχουν γενικευτεί από πολλά σύνολα δεδομένων από το Medical Segmentation Decathlon (MSD) και άλλες προκλήσεις.

Η επιτυχία του δεν οφείλεται στην καινοτομία της αρχιτεκτονικής του, η οποία βασίζεται στο U-Net με μικρές παραλλαγές, αλλά μάλλον στη συστηματική εναρμόνιση όλων των σταδίων της διαδικασίας, καθιστώντας το έτσι μια ισχυρή βάση για οποιοδήποτε νέο σύνολο δεδομένων. Η απλότητα της φιλοσοφίας του nnU-Net και η πλήρης αυτοματοποίησή του εξηγούν γιατί έχει καθιερωθεί ως το *de facto* σημείο αναφοράς στην ιατρική τμηματοποίηση, με περισσότερες από 100 εφαρμογές σε διαφορετικές ανατομικές περιοχές και μορφές μέσα σε λίγα μόνο χρόνια από την κυκλοφορία του [21].

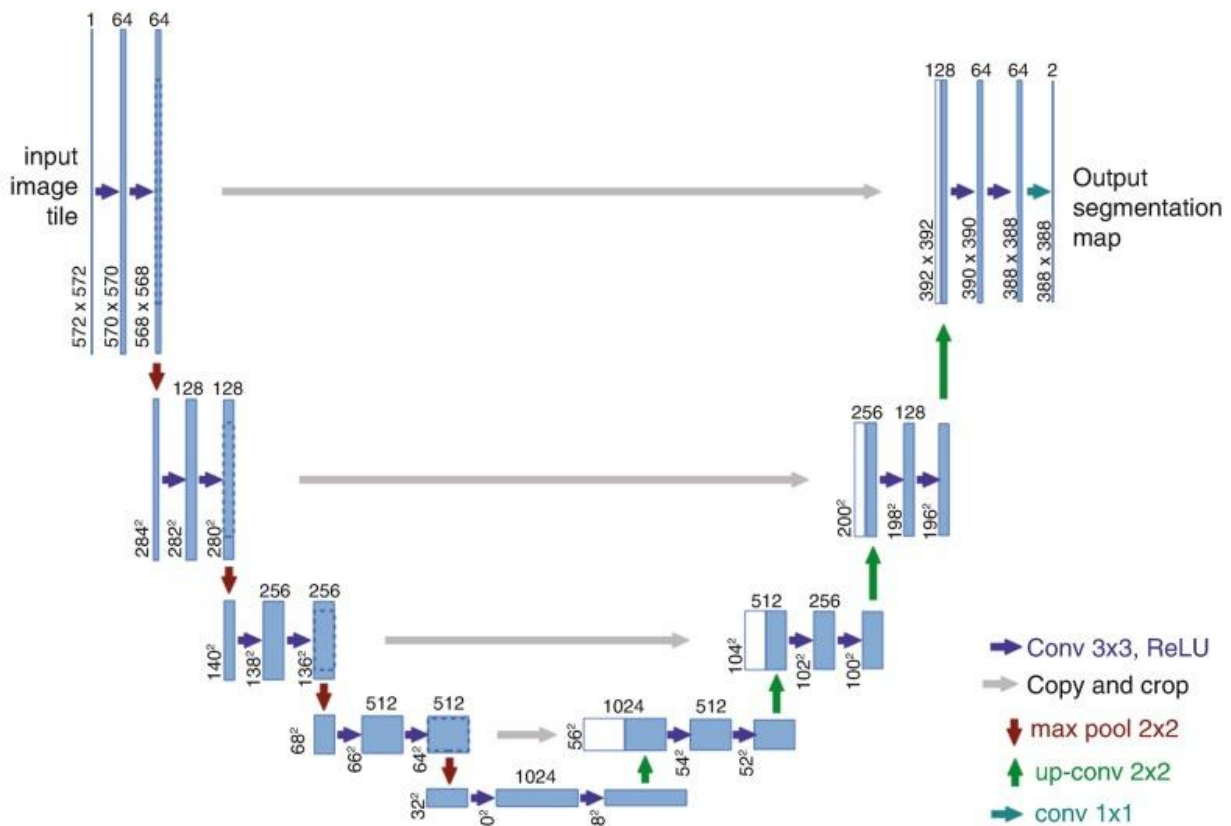
5.2. Βασική αρχιτεκτονική (U-Net backbone, encoder-decoder)

Η βασική αρχιτεκτονική του nnU-Net βασίζεται στην κλασική αρχιτεκτονική U-Net που εισήγαγαν οι Ronneberger et al. το 2015. Το U-Net έχει καταστεί το πρότυπο αναφοράς για τα προβλήματα ιατρικής τμηματοποίησης. Η αρχιτεκτονική του U-Net αποτελείται από δύο συμμετρικά μέρη: τον κωδικοποιητή, ο οποίος εξάγει ιεραρχικά χαρακτηριστικά μέσω διαδοχικών συνελίξεων και υποδειγματοληψίας (*downsampling*), και τον αποκωδικοποιητή, ο οποίος ανακατασκευάζει τον χάρτη τμηματοποίησης στην αρχική χωρική ανάλυση μέσω αντελίξεων (*upsampling*) και συνδέσεων παράκαμψης που μεταφέρουν πληροφορίες υψηλής ανάλυσης από τα αντίστοιχα επίπεδα του κωδικοποιητή [38].

Το nnU-Net εφαρμόζει παραλλαγές αυτής της αρχιτεκτονικής και προσαρμόζει αυτόματα το βάθος του δικτύου, τον αριθμό των φίλτρων και τη χρήση της κανονικοποίησης περιπτώσεων ανάλογα με τις διαστάσεις των δεδομένων εισόδου. Ο αριθμός των επιπέδων υποδειγματοληψίας καθορίζεται έτσι ώστε η τελική ανάλυση να μην είναι μικρότερη από 4×4 (2D) ή $4 \times 4 \times 4$ (3D) voxels για τη διατήρηση πληροφοριών υψηλής ανάλυσης. Οι συνελκτικές μονάδες χρησιμοποιούν πυρήνες 3×3 (ή πυρήνες $3 \times 3 \times 3$ για δεδομένα 3D), padding για τη διατήρηση των χωρικών διαστάσεων και τη συνάρτηση Leaky ReLU ως συνάρτηση ενεργοποίησης.

Στον αποκωδικοποιητή, κάθε βήμα ανάλυσης συνοδεύεται από συνένωση με τον αντίστοιχο χάρτη χαρακτηριστικών του κωδικοποιητή (μέσω συνδέσεων παράκαμψης) και ακολουθούνται από δύο διαδοχικές συνελίξεις. Αυτό επιτρέπει τη συνδυασμένη εκμάθηση τοπικών λεπτομερειών και περιφερειακών πληροφοριών. Το τελικό επίπεδο προβολής χρησιμοποιεί συνελίξεις 1×1 (ή $1 \times 1 \times 1$) για να παράγει την τελική μάσκα τμηματοποίησης με τον κατάλληλο αριθμό καναλιών ίσο με τις κατηγορίες-στόχους.

Αν και η αρχιτεκτονική του nnU-Net διατηρεί το ίδιο λειτουργικό πλαίσιο σε όλες τις παραλλαγές (2D, 3D πλήρης ανάλυση, 3D χαμηλή ανάλυση + καταρράκτης), παρουσιάζει παραλλαγές σε παραμέτρους όπως το βάθος, η ανάλυση, το μέγεθος της παρτίδας και άλλες, ανάλογα με το «αποτύπωμα» κάθε συνόλου δεδομένων. Αυτή η ευελιξία, σε συνδυασμό με την απλότητα και την αποδεδειγμένη αποτελεσματικότητα της αρχιτεκτονικής U-Net, συμβάλλει στην εξαιρετική απόδοσή της σε ένα ευρύ φάσμα εφαρμογών [21].



Σχήμα 12: Απεικόνιση της αρχιτεκτονικής U-Net [38].

5.3. Αυτόματη προεπεξεργασία (spacing, intensity, patch-based)

Ένα θεμελιώδες χαρακτηριστικό του πλαισίου nnU-Net αφορά την πλήρως αυτόματη προεπεξεργασία των δεδομένων. Αυτή η διαδικασία παρουσιάζει ικανότητα δυναμικής προσαρμογής σε κάθε σύνολο δεδομένων, με τους υποκείμενους μηχανισμούς της να ενημερώνονται από ένα σύστημα ευριστικών μεθόδων που προέρχονται από την ολοκληρωμένη ανάλυση του «αποτυπώματος» ολόκληρου του συνόλου δεδομένων. Οι συμβατικές προσεγγίσεις απαιτούν χειροκίνητη ρύθμιση κατά την προεπεξεργασία, ενώ το nnU-Net αυτοματοποιεί κρίσιμα βήματα, όπως η χωρική ανάλυση, η επαναδειγματοληψία, η κανονικοποίηση της έντασης και η εκπαίδευση εικόνων με βάση patches. Αυτές οι διαδικασίες βασίζονται στην ανάλυση και στους διαθέσιμους υπολογιστικούς πόρους [21].

Στο πλαίσιο των χωρικών παραμέτρων, το nnU-Net υπολογίζει αυτόνομα εκ νέου την χωρική ανάλυση-στόχο, γνωστή και ως απόσταση-στόχος, ενσωματώνοντας τα αρχικά χαρακτηριστικά ισοτροπίας ή ανισοτροπίας του συνόλου δεδομένων. Για παράδειγμα, σε σύνολα δεδομένων με σημαντικό πάχος τομών (>3 mm), εφαρμόζεται σχολαστική επαναδειγματοληψία 3D για να αποφευχθεί η παραμόρφωση των ανατομικών ορίων που μπορεί να προκύψει από υπερβολική παρεμβολή. Σε σύνολα δεδομένων που χαρακτηρίζονται

από ισοτροπικά μεγέθη voxel, το μοντέλο επιλέγει την άμεση εκπαίδευση 3D πλήρους ανάλυσης. Αντίθετα, σε σύνολα δεδομένων με υψηλή ανάλυση αλλά περιορισμένη μνήμη GPU, το μοντέλο επιλέγει χαμηλή ανάλυση σε συνδυασμό με την παραλλαγή cascade. Αυτή η διαδικασία έχει ως στόχο να διασφαλίσει ότι η ανάλυση παραμένει κλινικά σχετική χωρίς να υπερβαίνει τους διαθέσιμους πόρους.

Στο στάδιο της κανονικοποίησης της έντασης, το nnU-Net εφαρμόζει διαφορετικές στρατηγικές για κάθε μορφή. Στις εικόνες CT, χρησιμοποιείται κλιπ με βάση τα συνολικά εκατοστημόρια (π.χ. 0,5–99,5%), ακολουθούμενο από κανονικοποίηση z-score (με βάση τον μέσο όρο και την τυπική απόκλιση). Για τις εικόνες MRI, οι οποίες έχουν διαφορετικές κλίμακες και δυναμικά εύρη ανά λήψη, εφαρμόζεται κανονικοποίηση z ανά εικόνα. Η ομογενοποίηση της έντασης είναι κρίσιμη τόσο για την κανονικοποίηση της εκπαίδευσης όσο και για τη σταθερότητα των χαρακτηριστικών που μαθαίνει το δίκτυο.

Τέλος, όσον αφορά την εκπαίδευση με βάση τα patch, το nnU-Net δεν μεταβάλλει το συνολικό μέγεθος των εικόνων, αλλά επιλέγει δυναμικά ένα μέγεθος patch που καλύπτει επαρκώς την ανατομική περιοχή που μας ενδιαφέρει, εξισορροπώντας τις πληροφορίες σχετικά με το περιβάλλον και τις λεπτομέρειες (ανάλυση). Το μέγεθος του patch, το μέγεθος της παρτίδας και ο αριθμός των επιπέδων υποδειγματοληψίας καθορίζονται αυτόματα, έτσι ώστε το δίκτυο να μεγιστοποιεί τη χρήση της διαθέσιμης μνήμης GPU χωρίς τον κίνδυνο σφαλμάτων εξάντλησης μνήμης. Αυτή η δυναμική προσέγγιση εξαλείφει εντελώς την ανάγκη για ρύθμιση και εξασφαλίζει τη γενικευσιμότητα της διαδικασίας σε οποιοδήποτε νέο σύνολο δεδομένων.

5.4. Δυναμική επιλογή υπερπαραμέτρων και augmentation

Στο nnU-Net, η επιλογή των υπερπαραμέτρων δεν πραγματοποιείται χειροκίνητα. Αντίθετα, εφαρμόζονται μέσω ενός δυναμικού μηχανισμού αυτορύθμισης που εξαρτάται από τα χαρακτηριστικά κάθε συνόλου δεδομένων. Αντί να χρησιμοποιεί σταθερές τιμές για το ρυθμό μάθησης, το μέγεθος της παρτίδας ή το μέγεθος του patch, το nnU-Net δημιουργεί μια προσαρμοσμένη διαμόρφωση με βάση το χαρακτηριστικό αποτύπωμα των δεδομένων. Το χαρακτηριστικό αποτύπωμα των δεδομένων περιλαμβάνει χωρικές διαστάσεις, στατιστικά στοιχεία έντασης, τον αριθμό των καναλιών, το μέγεθος της κλάσης και τη διαθέσιμη μνήμη GPU. Αυτή η στρατηγική επιτρέπει στο πλαίσιο να βελτιστοποιεί την αποτελεσματικότητα της εκπαίδευσης και να παρακάμπτει την ανάγκη για χειροκίνητη ρύθμιση, επιβεβαιώνοντας έτσι την εννοιολόγησή του ως «μέθοδος αυτοδιαμόρφωσης» [21].

Για παράδειγμα, ο ρυθμός μάθησης αρχικοποιείται γενικά στο 0,01 χρησιμοποιώντας πολυωνυμική απόσβεση του ρυθμού μάθησης, ενώ το μέγεθος της παρτίδας επιλέγεται αυτόματα με βάση την υπολογιστική ικανότητα της GPU, ώστε να χρησιμοποιείται η μέγιστη δυνατή κάλυψη patch χωρίς να προκύπτουν σφάλματα εξάντλησης μνήμης. Το μέγεθος του patch καθορίζεται επίσης δυναμικά, ώστε να καλύπτει επαρκώς τις ανατομικές πληροφορίες κάθε συνόλου δεδομένων, με έμφαση στην ισορροπία μεταξύ τοπικών λεπτομερειών και συνολικού πλαισίου. Ταυτόχρονα, η βαθιά εποπτεία με βοηθητικούς κλάδους απώλειας

εφαρμόζεται σε ενδιάμεσα επίπεδα του αποκωδικοποιητή, ενισχύοντας έτσι τη σύγκλιση και προωθώντας μια πιο σταθερή ροή προς τα πίσω.

Η φάση αύξησης των δεδομένων στο nnU-Net είναι εξίσου κρίσιμη για τη γενίκευση του μοντέλου και βασίζεται σε ένα πλούσιο σύνολο τυχαίων (στοχαστικών) μετασχηματισμών που εφαρμόζονται σε πραγματικό χρόνο κατά τη διάρκεια της εκπαίδευσης. Το λογισμικό περιλαμβάνει μια σειρά λειτουργιών, όπως μετατοπίσεις, περιστροφές, μετασχηματισμούς κλίμακας, ελαστικές παραμορφώσεις, μετατόπιση έντασης, διόρθωση γάμμα, αύξηση της αντίθεσης και εφαρμογή γκαουσιανού θορύβου. Αυτοί οι μετασχηματισμοί δεν είναι σταθεροί, αλλά προσαρμόζονται δυναμικά με βάση το μέγεθος και τη φύση των δεδομένων. Ο σκοπός αυτών των προσαρμογών είναι να διατηρηθεί η ανατομική ακεραιότητα των δομών που μας ενδιαφέρουν χωρίς να προκληθούν παραμορφώσεις. Αυτή η έξυπνη χρήση της επαύξησης έχει αποδειχθεί κρίσιμη για την αποτελεσματική εκπαίδευση σε σύνολα δεδομένων με περιορισμένο αριθμό παραδειγμάτων, όπως συμβαίνει συχνά στην ιατρική απεικόνιση.

Η ενσωμάτωση της δυναμικής επιλογής υπερπαραμέτρων και των προσαρμοστικών αυξήσεων είναι ένας βασικός παράγοντας που συμβάλλει στην ανώτερη απόδοση του nnU-Net σε σύγκριση με άλλα συστήματα τμηματοποίησης, εξαλείφοντας έτσι την ανάγκη για ανθρώπινη παρέμβαση ή τεχνική ρύθμιση.

5.5. Σύγκριση με άλλες state-of-the-art μεθόδους

Το nnU-Net έχει επανειλημμένα αποδείξει την ανωτερότητά του σε σχέση με μια σειρά από προηγμένες μεθόδους τμηματοποίησης. Αυτό δεν οφείλεται σε κάποια ριζική καινοτομία στην αρχιτεκτονική του, αλλά στην συστηματική και αυτοματοποιημένη προσέγγιση που εφαρμόζει σε ολόκληρη τη διαδικασία. Κατά τη διάρκεια του Medical Segmentation Decathlon (MSD), το nnU-Net απέδειξε την υπεροχή του κατακτώντας την πρώτη θέση στην τελική πρόκληση, ξεπερνώντας έτσι πολλά εξειδικευμένα μοντέλα και στις δέκα εργασίες. Αξίζει να σημειωθεί ότι αυτό το επίτευγμα πραγματοποιήθηκε χωρίς την ανάγκη ad hoc σχεδιασμού ή λεπτομερούς ρύθμισης, κάτι που αποδεικνύει τις αυτόνομες δυνατότητες μάθησης του μοντέλου [21].

Σε αντίθεση με άλλες προηγμένες προσεγγίσεις που βασίζονται σε μηχανισμούς προσοχής, αρχιτεκτονικές πολλαπλών κλάδων ή στρατηγικές βασισμένες στη σύντηξη, όπως το Attention U-Net, το V-Net και το DeepLabV3+, το nnU-Net αποφεύγει τον πειραματισμό με πολύπλοκες αρχιτεκτονικές, επιτυγχάνοντας κορυφαία απόδοση μόνο με ένα τυπικό backbone U-Net, το οποίο έχει διαμορφωθεί δυναμικά με βάση τις ιδιαιτερότητες του συνόλου δεδομένων. Αυτό το εύρημα συνοψίζει ένα κρίσιμο συμπέρασμα: η συνεπής και προσεκτική διαμόρφωση του pipeline συχνά υπερτερεί της αρχιτεκτονικής πολυπλοκότητας στις περισσότερες εφαρμογές ιατρικής τμηματοποίησης.

Μια επακόλουθη σύγκριση με εξειδικευμένα δίκτυα, όπως τα nn-V-Net, 3D ResUNet και SegResNet, αποκάλυψε ότι, υπό βέλτιστες συνθήκες, το nnU-Net παρουσιάζει σταθερά συγκρίσιμη ή ανώτερη απόδοση. Αυτή η ανωτερότητα μπορεί να αποδοθεί στην αυτόνομη

επιλογή της ανάλυσης, της κανονικοποίησης, του μεγέθους του patch και της επαύξεσης, ένα χαρακτηριστικό που συνήθως καθορίζεται χειροκίνητα σε άλλα πλαίσια. Ταυτόχρονα, έχει αποδειχθεί ότι η προσαρμογή του nnU-Net σε σύνολα δεδομένων εκτός του MSD, συμπεριλαμβανομένων των BraTS, LiTS και KiTS, έχει ως αποτέλεσμα υψηλή γενικευσιμότητα, διατηρώντας παράλληλα την ευκολία χρήσης και την αναπαραγωγικότητα χωρίς την ανάγκη τροποποιήσεων [21].

Συνοψίζοντας, η σύγκριση του nnU-Net με άλλες μεθόδους δείχνει ότι η επίτευξη υψηλής απόδοσης στην ιατρική απεικόνιση δεν απαιτεί αρχιτεκτονική καινοτομία. Αντίθετα, είναι πιο σημαντικό να δοθεί έμφαση στη συνέπεια, την αυτοματοποίηση και την επαναχρησιμοποίηση αποδεδειγμένων πρακτικών. Το nnU-Net έχει μετατραπεί από ένα απλό μοντέλο σε πλαίσιο αναφοράς, το οποίο χρησιμοποιείται εκτενώς τόσο για τη συγκριτική αξιολόγηση όσο και για την παραγωγή βάσεων δεδομένων και εφαρμογών στον πραγματικό κόσμο.

6. Μεθοδολογία Πειραμάτων

6.1. Περιγραφή dataset

Το σύνολο δεδομένων που χρησιμοποιείται προέρχεται από το Medical Segmentation Decathlon (MSD), Task07 – Pancreas, το οποίο περιλαμβάνει αξονικές τομογραφίες κοιλίας με στόχο την ταυτόχρονη τμηματοποίηση του παγκρέατος και των νεοπλασμάτων. Τα δεδομένα παρέχονται σε μορφή NIfTI (.nii.gz) και είναι δημοσίως διαθέσιμα μέσω του [MSD challenge](#). Οι εικόνες προέρχονται από πολυκεντρικές κλινικές μελέτες, ενώ η ανεξάρτητη σχολιασμός των μασκών ground truth πραγματοποιήθηκε από εξειδικευμένους ακτινολόγους και ιατρικούς ερευνητές, εξασφαλίζοντας υψηλό βαθμό αξιοπιστίας και κλινικής εγκυρότητας. Οι μάσκες περιλαμβάνουν τρεις διακριτές κατηγορίες: 0 (φόντο), 1 (παγκρεατικό παρέγχυμα) και 2 (όγκος).

6.2. Υπολογιστική υποδομή

6.2.1. GPU και hardware details

Τα πειράματα διεξήχθησαν σε έναν προσωπικό σταθμό εργασίας με λειτουργικό σύστημα Linux (kernel 6.8.0-59-generic) και όνομα κεντρικού υπολογιστή phorever. Το σύστημα ήταν εξοπλισμένο με επεξεργαστή 6 πυρήνων/12 νημάτων και κάρτα γραφικών NVIDIA GeForce RTX 4080 με μνήμη GPU GDDR6X 16 GB. Η επιλογή αυτής της GPU διευκόλυνε τη διαδικασία εκπαίδευσης σε πλήρη ανάλυση 3D για πολλαπλές παραλλαγές του nnU-Net, εξαλείφοντας την ανάγκη για μειωμένη ανάλυση ή κασκαδική μάθηση. Η διαχείριση των πόρων GPU και η παρακολούθηση της κατανάλωσης μνήμης πραγματοποιήθηκαν μέσω της χρήσης του PyTorch CUDA profiling και των εργαλείων NVIDIA driver stack.

Η διάρκεια ενός πλήρους πειράματος ($15 \text{ εκτελέσεις} \times 1000 \text{ εποχές}$ σε πολλαπλές ρυθμίσεις διαστήματος/παρεμβολής) κυμάνθηκε από 24 έως 36 ώρες ανά εκτέλεση, ανάλογα με το μέγεθος των patches, τον αριθμό των καναλιών και τις απαιτήσεις του αγωγού επαύξησης.

6.2.2. Περιβάλλον λογισμικού (OS, βιβλιοθήκες)

Το σύστημα εκτέλεσης υλοποιήθηκε σε περιβάλλον Linux μέσω του Windows Subsystem for Linux (WSL) και περιελάμβανε ένα περιβάλλον Python 3.12.9, το οποίο διαχειριζόταν το Miniconda. Το pipeline υλοποιήθηκε στο κύριο έργο panorama, με το train.py να χρησιμεύει ως το κύριο αρχείο εκπαίδευσης. Το περιβάλλον της εικονικής μηχανής δημιουργήθηκε χρησιμοποιώντας ένα περιβάλλον conda με την ονομασία «panorama».

Οι κύριες βιβλιοθήκες που χρησιμοποιήθηκαν σε αυτή τη μελέτη παρατίθενται παρακάτω:

- PyTorch 2.x με CUDA 12 υποστήριξη
- MONAI για medical image utilities και loss functions
- SimpleITK και NiBabel για προεπεξεργασία ιατρικών εικόνων
- NumPy, SciPy, scikit-image για αριθμητικούς μετασχηματισμούς
- Weights & Biases (wandb v0.19.8) για καταγραφή και παρακολούθηση πειραμάτων

Η παρακολούθηση των runs γινόταν μέσω της πλατφόρμας Weights & Biases, με πλήρη καταγραφή hyperparameters, hardware usage, validation metrics και visualization heatmaps. Αυτό το ολοκληρωμένο σύστημα παρακολούθησης διευκόλυνε την ανάλυση της απόδοσης ανά παραμετροποίηση παρεμβολής/διαστήματος.

6.3. Σχεδιασμός πειραμάτων

6.3.1. Παράμετροι προς μεταβολή (spacing, interpolation, intensity)

Στο πλαίσιο της μελέτης εξετάστηκαν τρεις βασικές κατηγορίες παραμέτρων προεπεξεργασίας:

- Spatial resolution (target spacing)
- Interpolation method
- Intensity normalization (clipping ranges)

Κάθε run αντιστοιχεί σε μία μεταβολή των παραπάνω παραμέτρων, με στόχο την αξιολόγηση της επίδρασής τους στην ακρίβεια της τμηματοποίησης. Συγκεκριμένα:

Target spacing runs:

- UNet_spacing_1.0_1.0_1.0 → ισοτροπικό spacing 1 mm
- UNet_spacing_1.5_1.5_1.5 → ισοτροπικό spacing 1.5 mm
- UNet_spacing_2.0_2.0_2.0 → ισοτροπικό spacing 2 mm

- UNet_spacing_1.0_1.0_2.5 → ανισοτροπικό spacing (ελαφρά υποδειγματική z-axis ανάλυση)
- UNet_spacing_1.0_1.0_5.0 → ισχυρά ανισοτροπικό spacing (χαμηλή z-axis ανάλυση)

Interpolation runs:

- UNet_order1_interp → linear interpolation
- UNet_order2_interp → quadratic interpolation
- UNet_order3_interp → cubic spline
- UNet_order4_interp → higher-order
- UNet_order5_interp → πειραματική υψηλότερη τάξη

Intensity clipping/normalization runs:

- UNet_clip_none → χωρίς clipping
- UNet_clip_0.1_99.9 → clipping στα ποσοστημόρια 0.1-99.9%
- UNet_clip_1_99 → clipping στα ποσοστημόρια 1-99%
- UNet_clip_5_95 → clipping στα ποσοστημόρια 5-95%
- UNet_clip_10_90 → clipping στα ποσοστημόρια 10-90%

6.3.2. Πρωτόκολλο εκπαίδευσης (15 runs × 1000 epochs)

Κάθε πείραμα ακολουθεί σταθερό πρωτόκολλο εκπαίδευσης:

- Εποχές & Early Stopping: Η εκπαίδευση πραγματοποιείται για 1000 εποχές, με ενεργοποιημένο μηχανισμό early stopping.
- Συνάρτηση κόστους: Χρησιμοποιείται ο συνδυασμός Dice loss και Cross-Entropy loss.
- Βελτιστοποιητής & Σχήμα εκμάθησης ρυθμού: Εφαρμόζεται Stochastic Gradient Descent (SGD) με αρχικό learning rate = 0.01 και πολωνυμική απόσβεση.
- Βαθιά επίβλεψη & Κανονικοποίηση στιγμιαίων τιμών: Αξιοποιούνται τα modules του nnU-Net για deep supervision και instance normalization.
- Μέγεθος παρτίδας & patch: Ορίζονται αυτόματα από το nnU-Net, ανάλογα με το spacing των δεδομένων και τη διαθέσιμη μνήμη.

6.4. Μέτρα αξιολόγησης

6.4.1. Dice coefficient, sensitivity, specificity, precision

Η απόδοση των μοντέλων τμηματοποίησης αξιολογήθηκε χρησιμοποιώντας καθιερωμένους ποσοτικούς δείκτες ακρίβειας. Ο στόχος ήταν να αξιολογηθεί συνολικά η μορφολογική ομοιότητα μεταξύ των προβλεπόμενων και των πραγματικών масκών, καθώς και να αναλυθεί η διαγνωστική συμπεριφορά του μοντέλου ανά κατηγορία.

Dice Similarity Coefficient (DSC)

Θεμελιώδης μετρική για segmentation με ανισόρροπες κλάσεις, καθώς μετρά την επικαλυπτόμενη περιοχή σε σχέση με το συνολικό μέγεθος των δύο μασκών. Υπολογίζεται ξεχωριστά για κάθε κλάση και συνοψίζεται επίσης ως μέσος όρος των επιμέρους τιμών.

Η μαθηματική μορφή της είναι: $DSC = [(2TP)/(2TP+FP+FN)]$

Ακρίβεια (Precision) $[TP / (TP + FP)]$

Εκτιμά την ακρίβεια των θετικών προβλέψεων για κάθε κλάση, δηλαδή το ποσοστό των pixels που προβλέφθηκαν ως θετικά και όντως ανήκουν στην αντίστοιχη δομή.

Ευαισθησία (Recall) $[TP / (TP + FN)]$

Καταγράφει την ικανότητα ανίχνευσης των πραγματικών θετικών pixels, απεικονίζοντας το ποσοστό των αληθινά θετικών περιπτώσεων που αναγνωρίστηκαν από το μοντέλο.

Ειδικότητα (Specificity) $[TN / (TN + FP)]$

Δείχνει την ικανότητα απόρριψης των αρνητικών pixels, μετρώντας το ποσοστό των αληθινά αρνητικών περιπτώσεων που δεν εσφαλμένα ταξινομήθηκαν ως θετικές.

7. Αποτελέσματα

7.1. Πίνακας σύγκρισης πειραμάτων

Spacing	Total Dice	Dice ₁	Dice ₂	Prec ₁	Prec ₂	Sens ₁	Sens ₂	Spec ₁	Spec ₂	Train Loss
1.0 × 1.0 × 1.0	0.463	0.753	0.173	0.692	0.416	0.842	0.052	0.99957	0.99997	0.445
1.5 × 1.5 × 1.5	0.461	0.749	0.173	0.703	0.418	0.823	0.038	0.99960	0.99998	1.044
2.0 × 2.0 × 2.0	0.447	0.742	0.152	0.716	0.442	0.795	0.039	0.99966	0.99998	0.710
1.0 × 1.0 × 2.5	0.465	0.757	0.173	0.711	0.395	0.831	0.041	0.99962	0.99997	0.798
1.0 × 1.0 × 5.0	0.449	0.748	0.151	0.699	0.440	0.828	0.036	0.99959	0.99998	0.416

Πίνακας 1: Αποτελέσματα για διαφορετικό spacing

Interpolation	Total Dice	Dice ₁	Dice ₂	Prec ₁	Prec ₂	Sens ₁	Sens ₂	Spec ₁	Spec ₂	Train Loss
Order 1	0.474	0.753	0.194	0.707	0.468	0.831	0.065	0.99961	0.99997	0.398
Order 2	0.463	0.758	0.169	0.721	0.502	0.826	0.041	0.99964	0.99998	0.414
Order 3	0.460	0.755	0.166	0.714	0.431	0.826	0.046	0.99962	0.99997	0.429
Order 4	0.453	0.758	0.148	0.717	0.172	0.825	0.045	0.99963	0.99990	0.409
Order 5	0.472	0.761	0.183	0.734	0.488	0.812	0.046	0.99966	0.99998	0.481

Πίνακας 2: Αποτελέσματα για διαφορετικό Interpolation

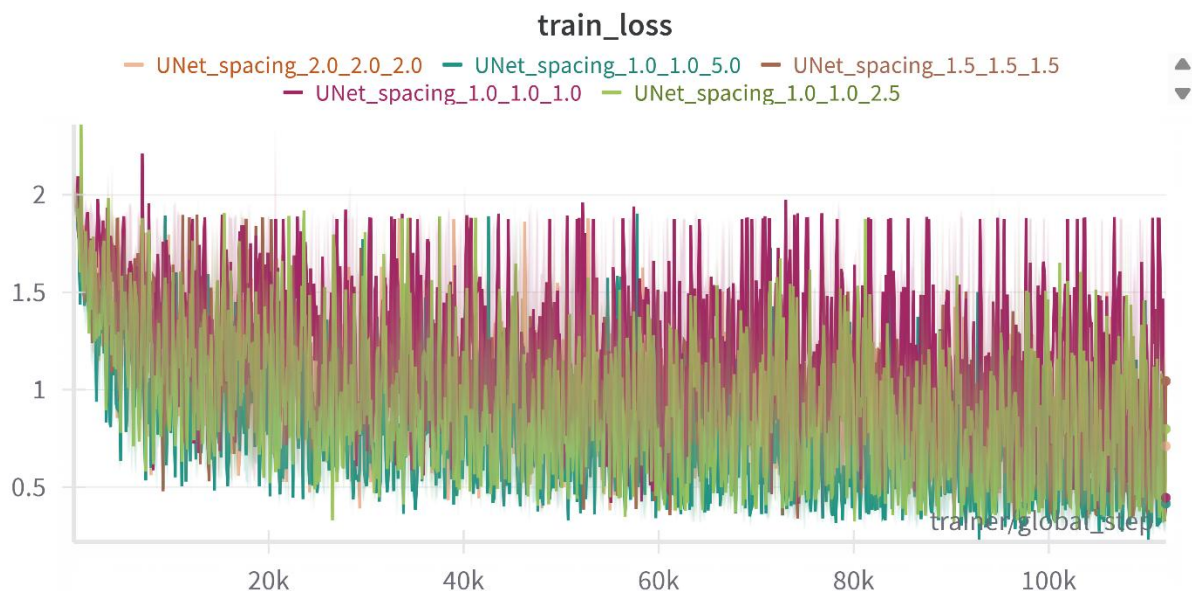
Clipping	Total Dice	Dice ₁	Dice ₂	Prec ₁	Prec ₂	Sens ₁	Sens ₂	Spec ₁	Spec ₂	Train Loss
None	0.479	0.760	0.199	0.710	0.433	0.843	0.054	0.9996 1	0.9999 7	1.079
0.1– 99.9 %	0.471	0.752	0.189	0.692	0.449	0.843	0.052	0.9995 7	0.9999 7	1.103
1–99 %	0.472	0.746	0.198	0.688	0.450	0.837	0.051	0.9995 7	0.9999 7	1.079
5–95 %	0.461	0.754	0.169	0.695	0.418	0.844	0.046	0.9995 8	0.9999 7	1.114
10–90 %	0.438	0.713	0.164	0.659	0.396	0.792	0.059	0.9995 3	0.9999 6	0.360

Πίνακας 3: Αποτελέσματα για διαφορετικό Clipping

7.2. Διαγράμματα

Παρακάτω φαίνονται τα training και validation curves για κάθε ομάδα πειραμάτων.

A. Spacing runs:



Διάγραμμα 1: Διάγραμμα Train Loss για τις διαφορετικές τιμές Spacing

Σχολιασμός:

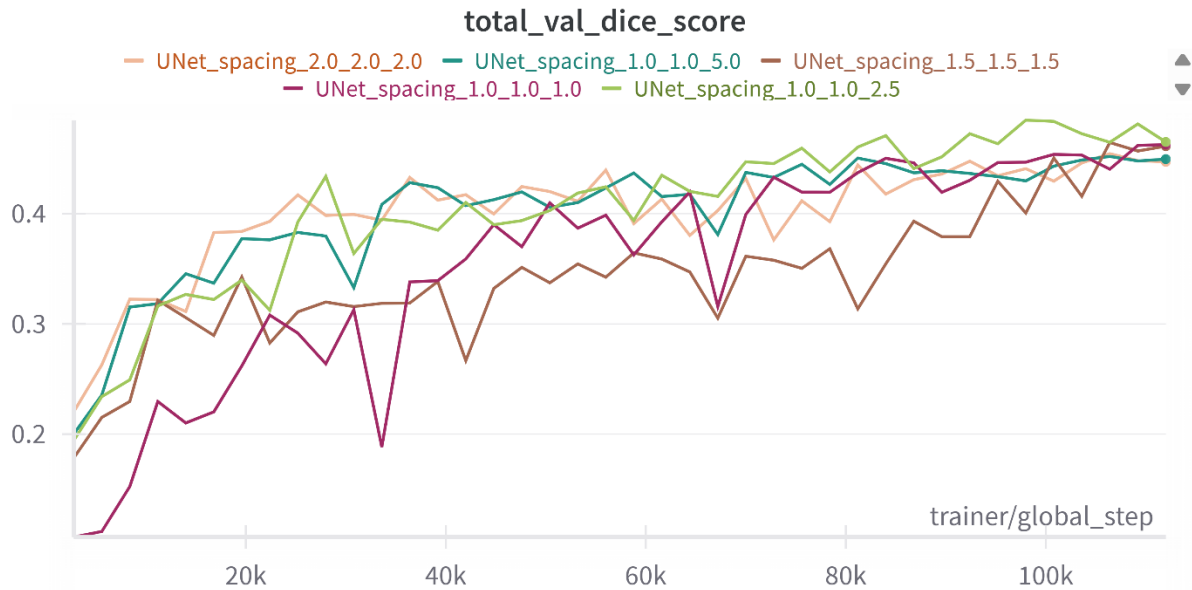
Το run με spacing $1 \times 1 \times 5$ mm έχει τη χαμηλότερη τιμή train loss καθ' όλη τη διάρκεια, φτάνοντας γρήγορα σε σταθεροποίηση (plateau) $\sim 0.4-0.45$.

Το spacing $1 \times 1 \times 1$ mm συγκλίνει επίσης γρήγορα σε plateau ~ 0.45 , ελαφρώς πάνω από το 5 mm spacing.

Τα spacings $2 \times 2 \times 2$ mm και $1.5 \times 1.5 \times 1.5$ mm εμφανίζουν υψηλότερο train loss (0.7 για 2 mm και ~ 1.04 για 1.5 mm), υποδεικνύοντας πιο αργή και λιγότερο σταθερή μάθηση.

Το $1 \times 1 \times 2.5$ mm plateau εμφανίζεται περίπου στο 0.8, καλύτερα από τα coarse ιστροπικά spacings αλλά υψηλότερα από τα άλλα ελαφρώς ανισοτροπικά.

Αυτή η συμπεριφορά δείχνει ότι η πολύ χονδροειδής (coarse) δειγματοληψία (2 mm ή 1.5 mm) δυσχεραίνει την εκπαίδευση, ενώ η ελαφρώς ανισοτροπική ($1.0 \times 1.0 \times 2.5$ mm) δίνει μια ισορροπία μεταξύ λεπτομέρειας και ταχύτητας σύγκλισης. Τα ακραία coarse (5 mm) αφενός απλοποιούν πολύ το πρόβλημα και έτσι επιτυγχάνουν πολύ χαμηλή train loss, αφετέρου οδηγούν σε υπο-εκπαίδευση του validation set.



Διάγραμμα 2: Διάγραμμα Total Val Dice Score για τις διαφορετικές τιμές Spacing

Σχολιασμός:

Όλες οι καμπύλες ανεβαίνουν από 0.20 → 0.30 μέσα στα πρώτα 15 – 20 k βήματα, αλλά η ποιότητά τους διαχωρίζεται μετά τα 30 k.

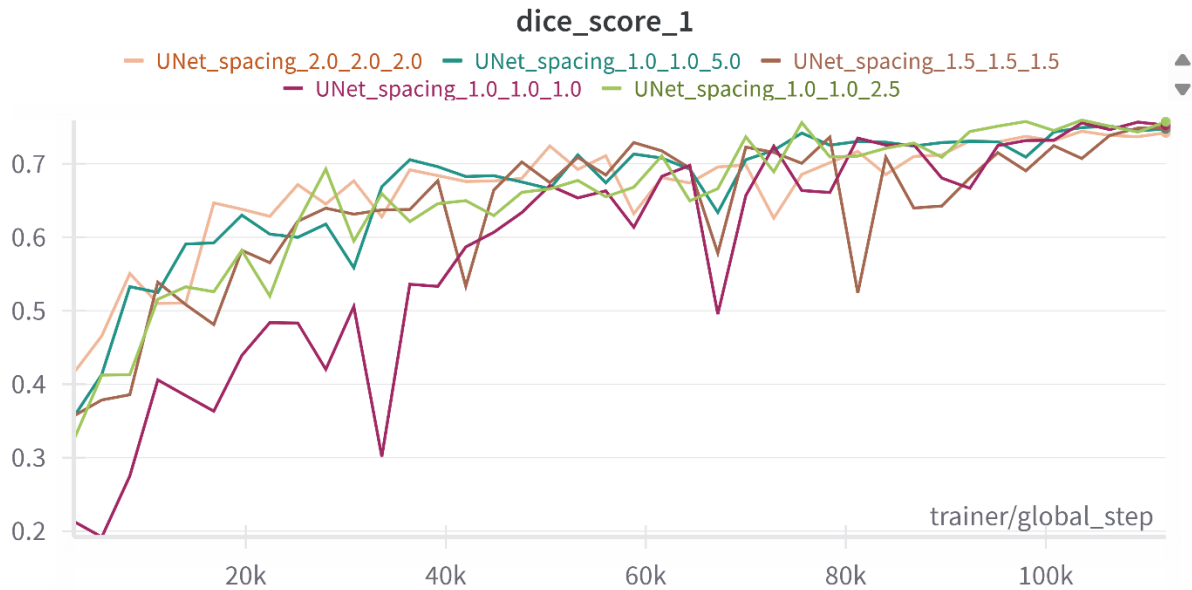
Η καμπύλη $1 \times 1 \times 2.5$ mm διατηρεί τη μεγαλύτερη κλίση και φτάνει στο υψηλότερο plateau (~0.48). Δείχνει ότι ένα ελαφρά ανισοτροπικό spacing προσφέρει καλύτερη ισορροπία ανάμεσα στη διατήρηση της χωρικής λεπτομέρειας και στο ευρύτερο πλαίσιο (context).

Η καμπύλη $1 \times 1 \times 5$ mm εκτοξεύεται γρήγορα → 0.45, αλλά σταθεροποιείται λίγο χαμηλότερα από το 2.5 mm.

Η καμπύλη $1 \times 1 \times 1.0$ mm ανεβαίνει πιο αργά και τελειώνει στο ~0.45, δηλαδή είναι αξιοπρεπής, χωρίς υπερβολικό θόρυβο.

Η καμπύλη $1.5 \times 1.5 \times 1.5$ mm έχει τη χειρότερη κλίση και plateau (~0.42).

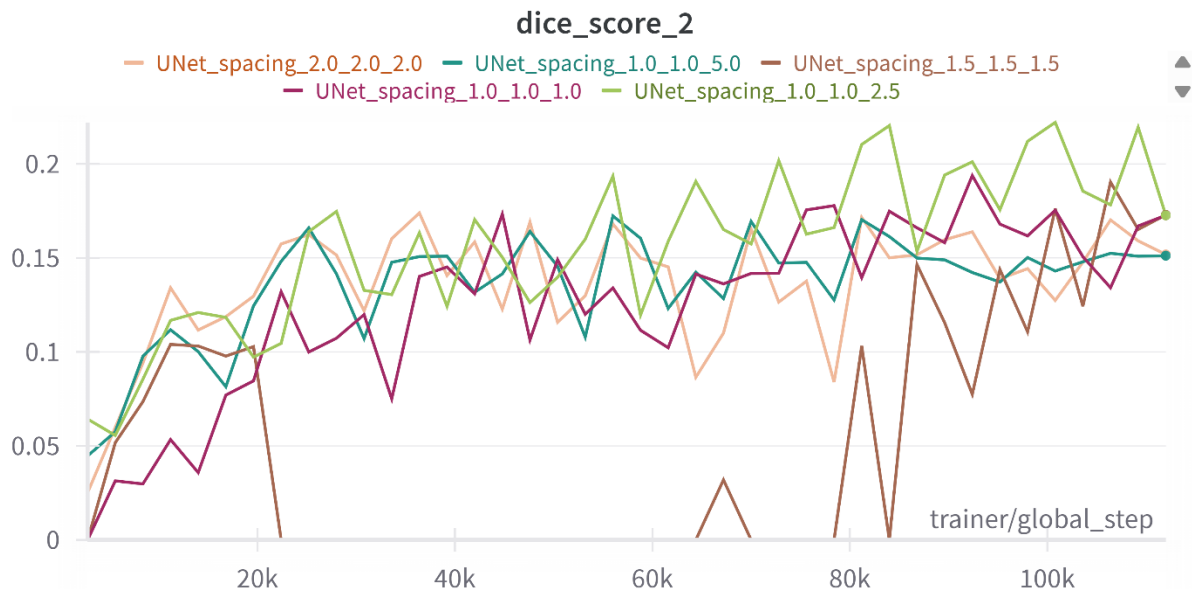
Η καμπύλη $2 \times 2 \times 2$ mm ανεβαίνει γρήγορα αρχικά αλλά «σβήνει» στο ~0.44.



Διάγραμμα 3: Διάγραμμα Dice Score 1 για τις διαφορετικές τιμές Spacing

Σχολιασμός:

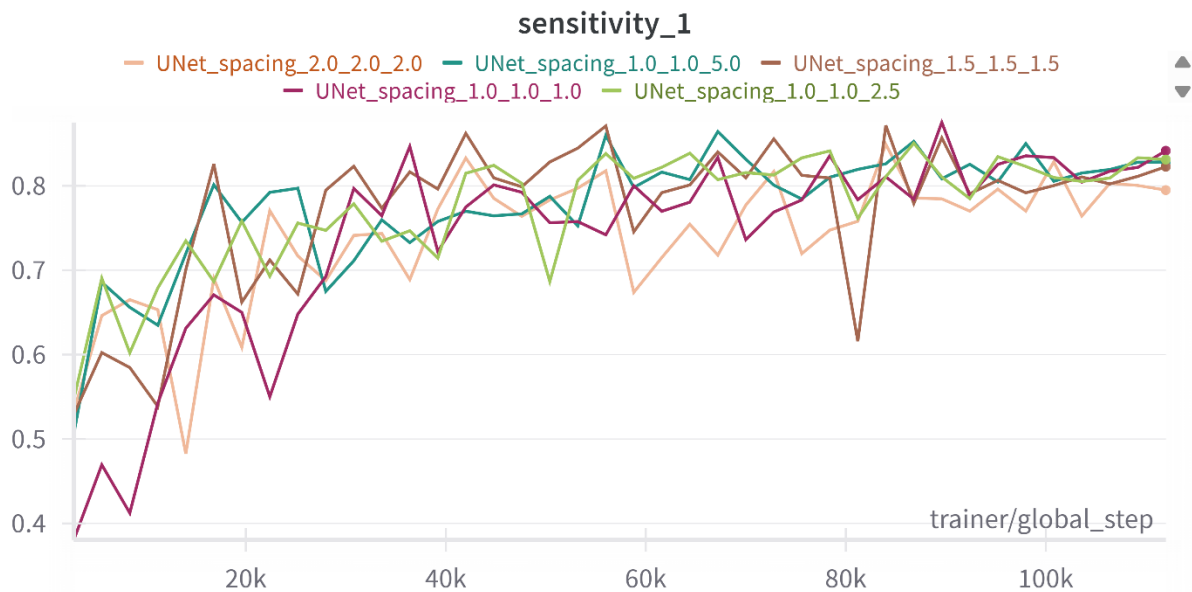
Όλες οι παραμετροποιήσεις συγκλίνουν σε παρόμοιο επίπεδο Dice₁ ($\approx 0.72\text{--}0.75$). Το spacing $1 \times 1 \times 2.5$ mm παρουσιάζει ελαφρό πλεονέκτημα: φτάνει ταχύτερα σε τιμές > 0.70 και σταθεροποιείται γύρω στο 0.75. Συνεπώς, η επιλογή spacing επηρεάζει ελάχιστα την τελική ακρίβεια, αφού μόλις το μοντέλο “μάθει” το ολικό σχήμα, η λεπτομέρεια του z-άξονα παίζει δευτερεύοντα ρόλο.



Διάγραμμα 4: Διάγραμμα Dice Score 2 για τις διαφορετικές τιμές Spacing

Σχολιασμός:

Τα αποτελέσματα διαφοροποιούνται έντονα. Το $1 \times 1 \times 2.5$ mm πετυχαίνει το υψηλότερο plateau ($\approx 0.20\text{--}0.22$), ακολουθούμενο από το $1 \times 1 \times 5$ mm (≈ 0.18). Το ιστροπικό 1 mm και το 2 mm σταθεροποιούνται χαμηλότερα ($\approx 0.15\text{--}0.17$), ενώ το 1.5 mm εμφανίζει αστάθεια και τη χαμηλότερη τελική τιμή. Η υπεροχή του spacing 2.5 mm οφείλεται στον επιτυχημένο συμβιβασμό μεταξύ υψηλής εγκάρσιας ανάλυσης (1 mm στο xy) και επαρκούς «πάχους» κατά τον z-άξονα, που παρέχει πρόσθετο όγκο χωρίς να εξαλείφει λεπτά μορφολογικά στοιχεία.



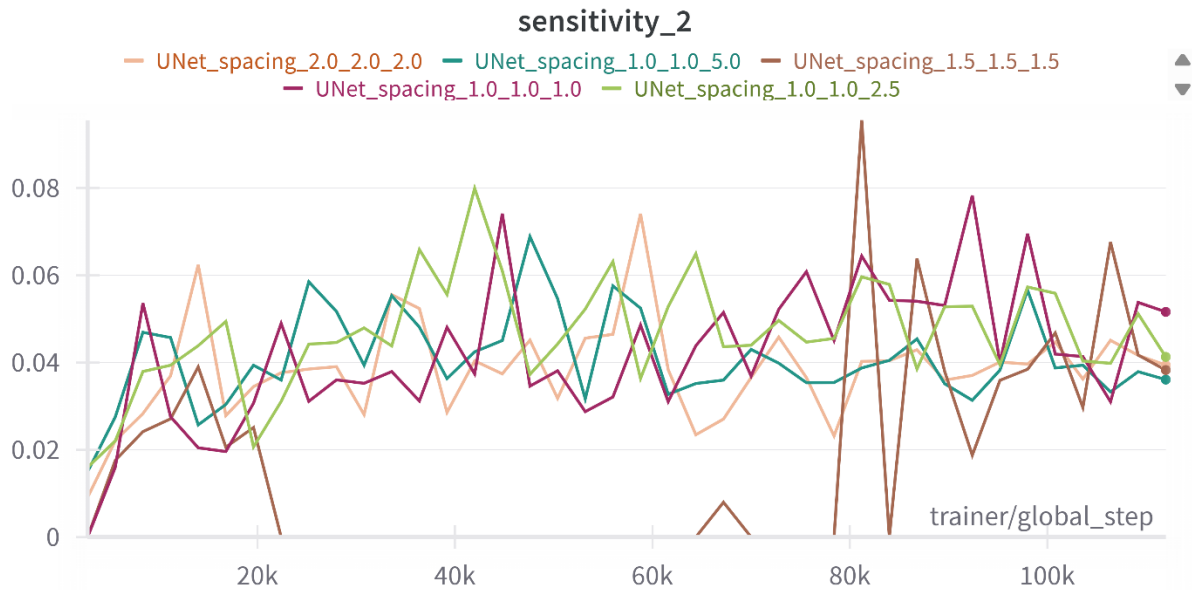
Διάγραμμα 5: Διάγραμμα Sensitivity 1 για τις διαφορετικές τιμές Spacing

Σχολιασμός:

Η ιστροπική δειγματοληψία 1 mm επιτυγχάνει τη μέγιστη ευαισθησία (0,842).

Η τιμή μειώνεται σταδιακά όσο αυξάνεται το βήμα voxel και η χαμηλότερη καταγράφεται στο 2 mm (0,795).

Παρ' όλα αυτά, οι διαφορές είναι μικρές ($< 5\%$), γεγονός που δείχνει ότι η αναγνώριση της μεγάλης ανατομικής δομής δεν επηρεάζεται δραματικά από το spacing.



Διάγραμμα 6: Διάγραμμα Sensitivity 2 για τις διαφορετικές τιμές Spacing

Σχολιασμός:

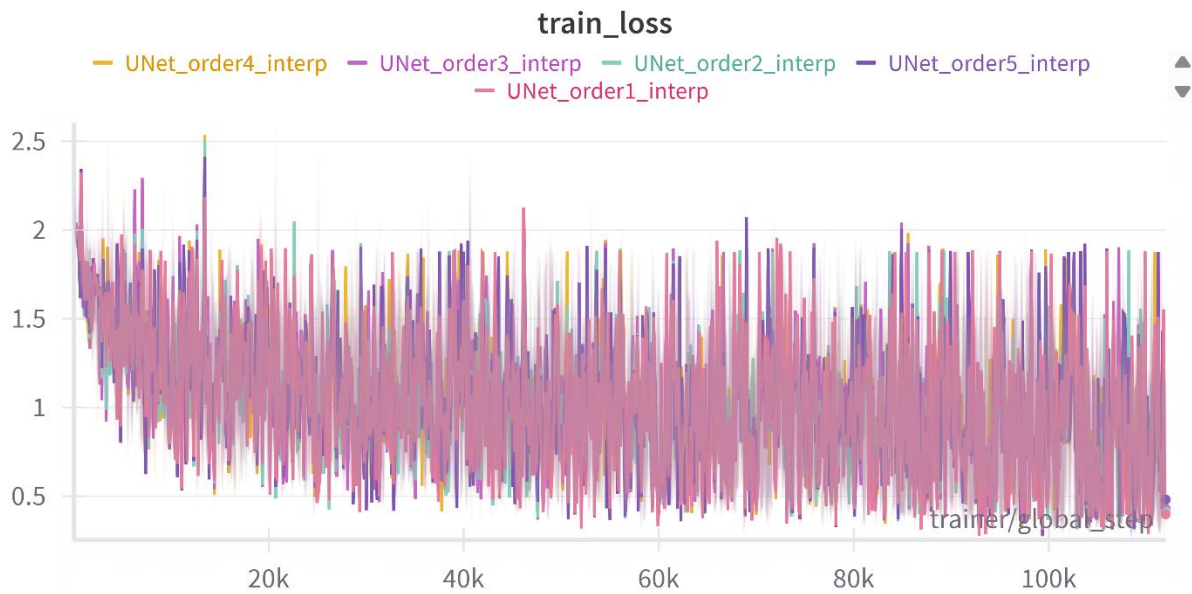
Οι τιμές είναι χαμηλές εξαιτίας του έντονου class-imbalance, ωστόσο η τάση είναι ευδιάκριτη:

Με spacing 1 mm ισοτροπικό έχουμε υψηλότερο recall (0,052).

Το spacing $1 \times 1 \times 2,5$ mm ακολουθεί (0,041) προσφέροντας αξιοσημείωτη βελτίωση έναντι των coarse isotropic επιλογών.

Το $1 \times 1 \times 5$ mm, παρά τη χαμηλή τιμή train-loss, έχει το χαμηλότερο Sensitivity₂ (0,036), το οποίο είναι ένδειξη ότι «χάνει» λεπτές βλάβες εξαιτίας της πολύ χαμηλής z-ανάλυσης.

B. Interpolation runs



Διάγραμμα 7: Διάγραμμα Train Loss για τις διαφορετικές τιμές Interpolation

Σχολιασμός:

Ταχύτητα σύγκλισης:

Όλες οι καμπύλες πέφτουν από >2 σε ≈ 1 μέσα στα πρώτα 10 k steps, δείγμα ότι το σχήμα παρεμβολής δεν επηρεάζει ουσιωδώς την αρχική μάθηση.

Τελικό επίπεδο (plateau):

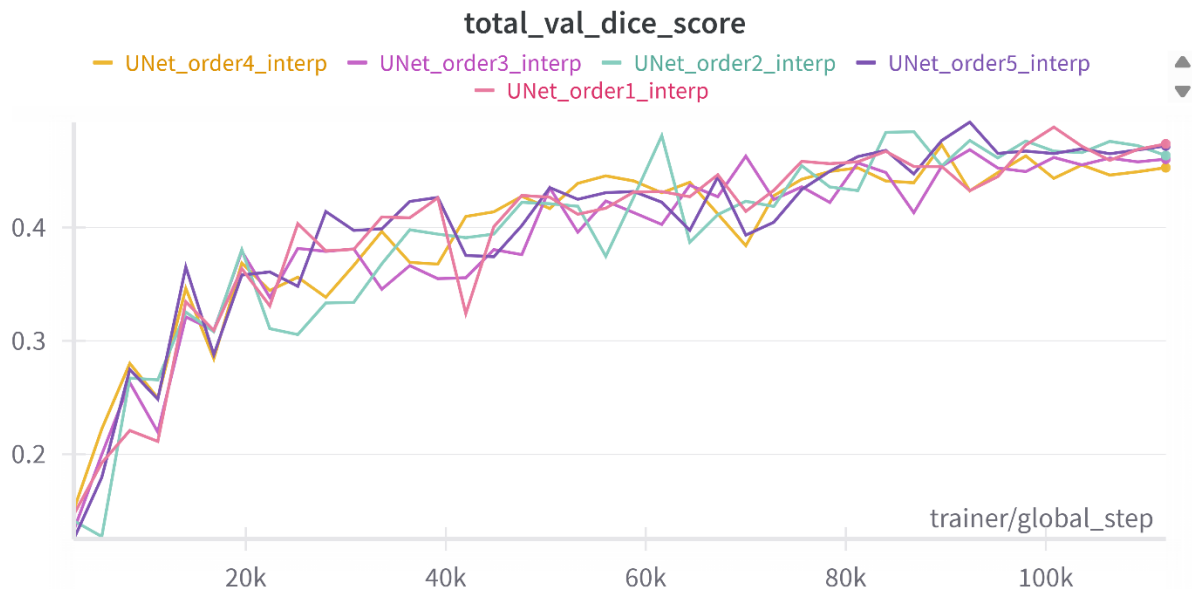
Linear (Order 1) επιτυγχάνει το χαμηλότερο plateau (~ 0.40).

Τα Orders 2–4 διαφέρουν < 0.03 μονάδες, δηλαδή πρακτικά ισοδυναμούν σε δυσκολία εκπαίδευσης.

Το πειραματικό υψηλής τάξης (Order 5) διατηρεί αισθητά υψηλότερο train loss (~ 0.48), ένδειξη ότι οι επιθετικότερες συναρτήσεις παρεμβολής δυσκολεύουν το optimisation.

Θόρυβος καμπύλης:

Όλες παρουσιάζουν «αιχμές» έως 2.0 σε μεμονωμένα βήματα. Αυτό οφείλεται στον στοχαστικό χαρακτήρα των augmentations και στο μικρό batch size.



Διάγραμμα 8: Διάγραμμα Total Val Dice Score για τις διαφορετικές τιμές Interpolation

Σχολιασμός:

Αρχική άνοδος (0–20 k steps):

Όλα τα μοντέλα εκτοξεύονται από $\text{Dice} \approx 0,15 \rightarrow 0,30\text{--}0,35$ χωρίς ουσιαστικές διαφορές· ο τύπος παρεμβολής δεν επηρεάζει το “ξεκίνημα”.

Στάδιο σύγκλισης (20 k–60 k):

Η καμπύλη Order 1 ανεβαίνει σταθερά και πρώτη ξεπερνά το 0,40.

Τα Orders 2–4 εμφανίζουν μικρές “βουτιές”, ειδικά το Order 3 γύρω στα 40 k, χάνοντας momentum.

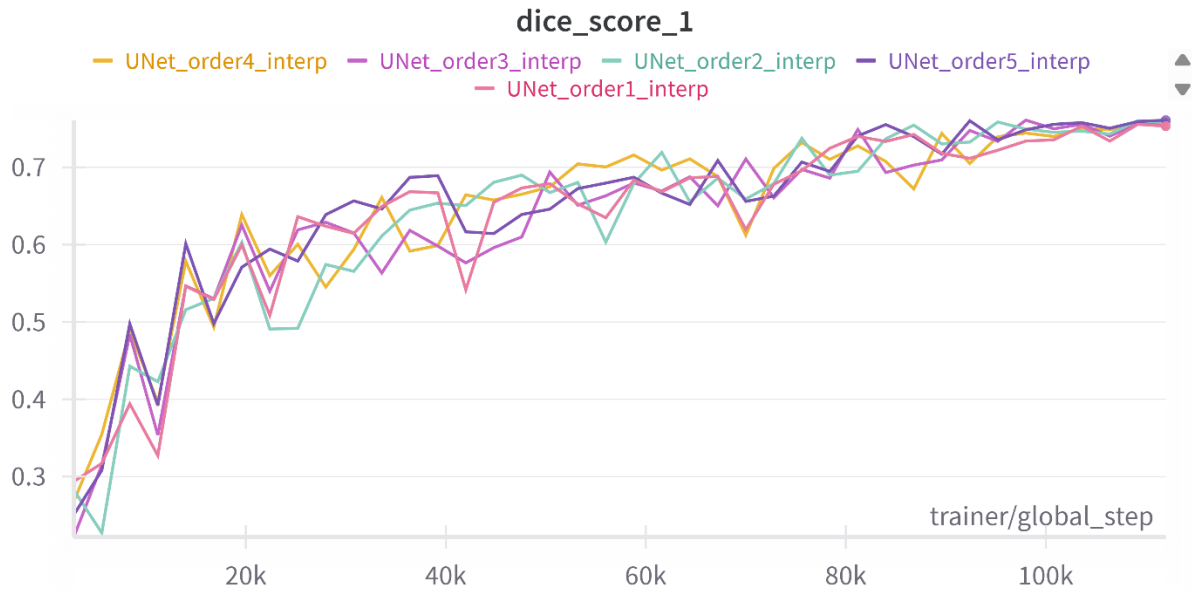
Το Order 5 ακολουθεί στενά το Order 1, αλλά με μεγαλύτερο θόρυβο.

Τελικό plateau (≥ 80 k):

Order 1 σταθεροποιείται στο υψηλότερο επίπεδο, 0,47–0,48.

Order 5 αποδίδει σχεδόν το ίδιο (0,47 +) αλλά χρειάζεται λίγα χιλιάδες βήματα παραπάνω για να φτάσει εκεί.

Οι ενδιάμεσες τάξεις (2–4) κλείνουν σημαντικά χαμηλότερα (0,453–0,463).



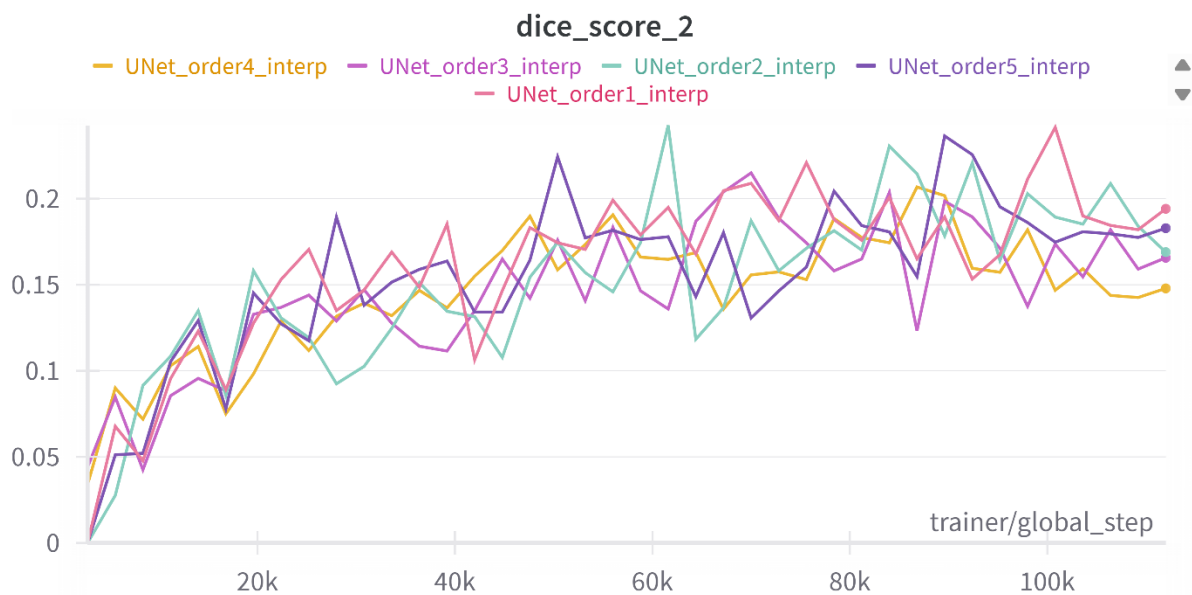
Διάγραμμα 9: Διάγραμμα Dice Score 1 για τις διαφορετικές τιμές Interpolation

Σχολιασμός:

Η απόδοση είναι υψηλή σε όλα τα orders (0,753 – 0,761).

Υπάρχει οριακό κέρδος +0,005–0,008 όταν περνάμε από linear (Order 1) σε πιο πολύπλοκες παρεμβολές (Orders 2-5).

Το γραφικό δείχνει ότι, μετά τα 60 k steps, όλες οι καμπύλες συγκλίνουν σε κοινό plateau ~0,75, δηλαδή οι διαφορές είναι αμελητέες σε πρακτικό επίπεδο.



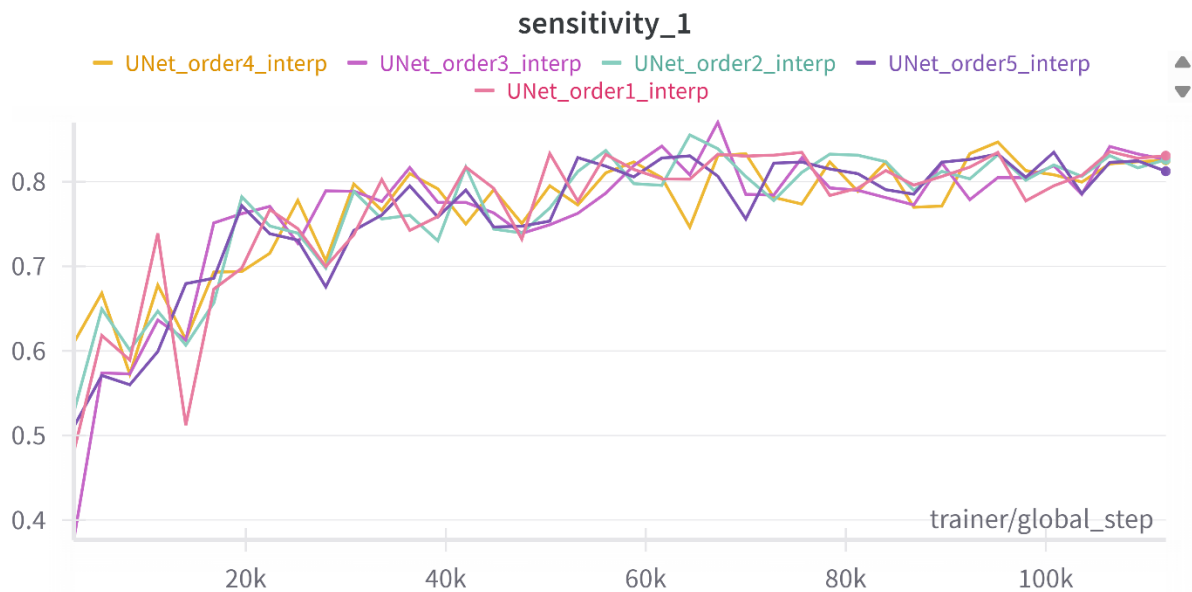
Διάγραμμα 10: Διάγραμμα Dice Score 2 για τις διαφορετικές τιμές Interpolation

Σχολιασμός:

Η Linear (Order 1) πετυχαίνει το υψηλότερο $\text{Dice}_2 = 0,194$, διατηρώντας λεπτές δομές.

Οι spline παρεμβολές εξομαλύνουν υπερβολικά τις οριακές εντάσεις και “κόβουν” μικρο-βλάβες· το Dice_2 πέφτει στα 0,166-0,148.

Το πειραματικό high-order (Order 5) είναι λίγο βελτιωμένο σε σχέση με τα Orders 2-4 (0,183), αλλά δεν φτάνει το linear.



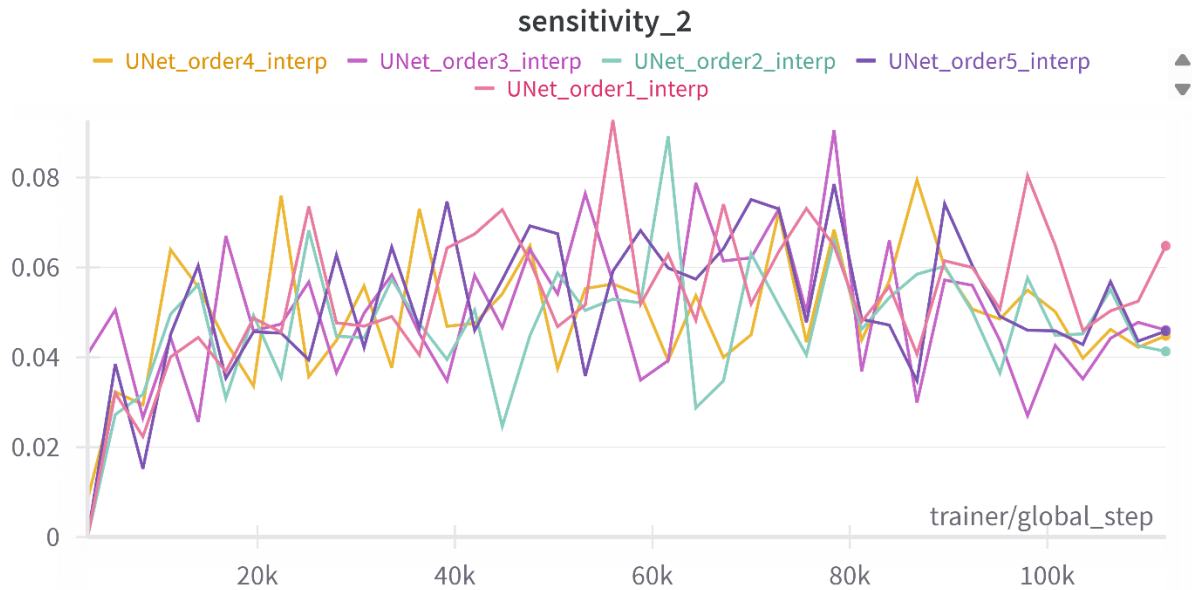
Διάγραμμα 11: Διάγραμμα Sensitivity 1 για τις διαφορετικές τιμές Interpolation

Σχολιασμός:

Όλα τα μοντέλα καταλήγουν σε υψηλό recall ($\geq 0,81$).

Το linear interpolation (Order 1) πετυχαίνει την κορυφαία τιμή (0,831) και φτάνει εκεί ταχύτερα, όπως φαίνεται στο γράφημα.

Τα Orders 2-4 απέχουν μόλις $\sim 0,005$ μονάδες, διαφορά πρακτικά αμελητέα· ωστόσο το Order 5 υποχωρεί αισθητά (0,812).



Διάγραμμα 12: Διάγραμμα Sensitivity 2 για τις διαφορετικές τιμές Interpolation

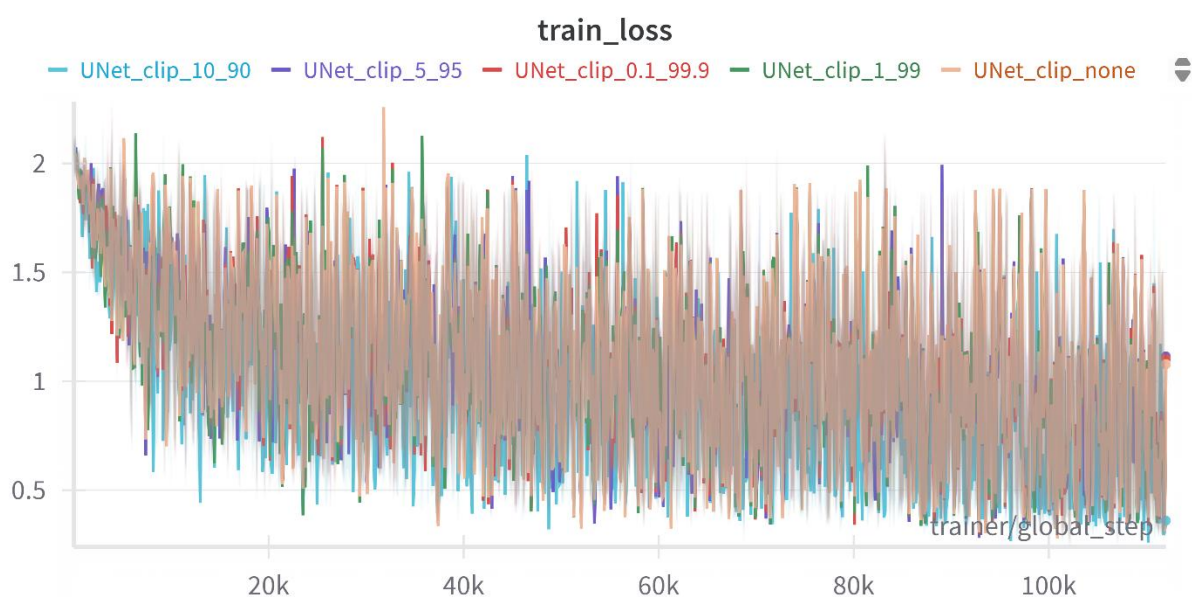
Σχολιασμός:

Linear (Order 1) διατηρεί το υψηλότερο recall (0,065), δηλαδή ~50 % περισσότερο από το quadratic.

Τα spline orders (2-4) μειώνουν αισθητά το Sens₂, λόγω υπερ-εξομάλυνσης που «κόβει» τις λεπτές βλάβες.

Το high-order φίλτρο (Order 5) βελτιώνει ελαφρά σε σχέση με τα Orders 2-4 (0,046) αλλά παραμένει πολύ κάτω από το linear.

Γ. Clipping runs



Διάγραμμα 13: Διάγραμμα Train Loss για τις διαφορετικές τιμές Clipping

Σχολιασμός:

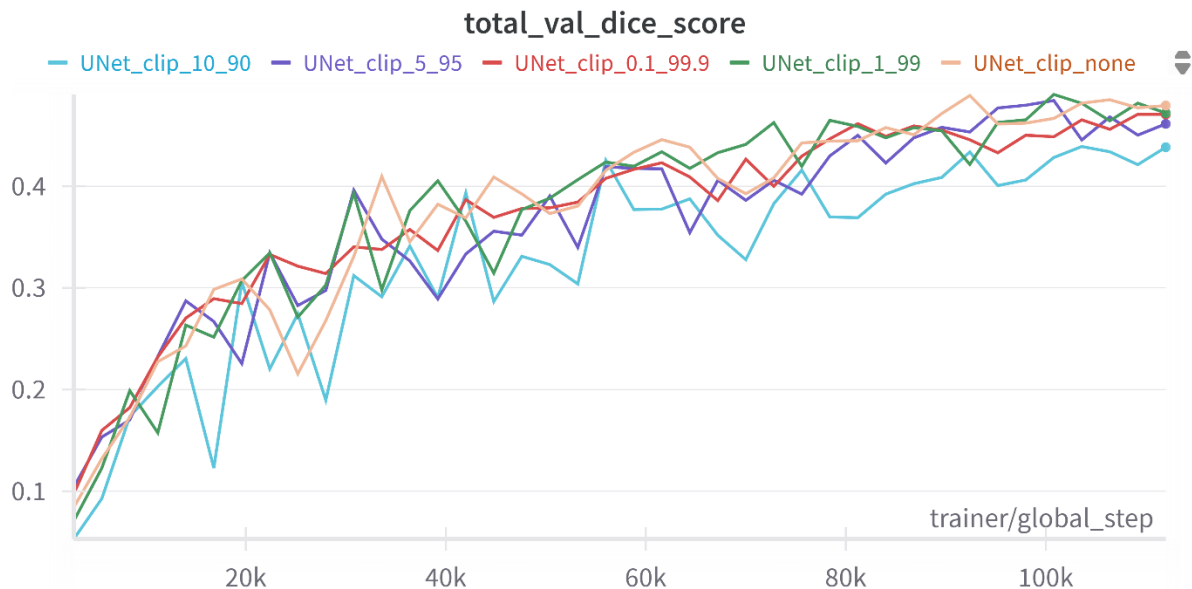
Αρχικό στάδιο (0–15 k steps):

Όλες οι καμπύλες ξεκινούν από $\text{loss} \approx 2,0$ και κατεβαίνουν γρήγορα κάτω από 1,0. Η επιλογή clipping δεν επηρεάζει τη φάση «γρήγορης» μάθησης.

Σύγκλιση (μετά τα 20 k):

Τα runs χωρίς clipping ή με ήπια αποκοπή (0.1–99.9 %, 1–99 %, 5–95 %) σταθεροποιούνται σε plateau $\approx 1,05 - 1,15$.

Αντίθετα, το 10–90 % ακολουθεί διαφορετική τροχιά: το loss συνεχίζει να μειώνεται και καταλήγει στο 0,36.



Διάγραμμα 14: Διάγραμμα Total Val Dice Score για τις διαφορετικές τιμές Clipping

Σχολιασμός:

Αρχική φάση (0–20 k steps):

Όλα τα runs ανεβαίνουν γρήγορα από Dice $\approx 0,05$ σε 0,28–0,32, δηλαδή το clipping δεν επηρεάζει το πρώτο κύμα μάθησης.

Μεσαίο στάδιο (20 k–60 k):

Οι καμπύλες χωρίς clipping, 1–99 % και 0,1–99,9 % διατηρούν υψηλότερη κλίση, ξεπερνώντας το 0,40.

Το 5–95 % ακολουθεί με πιο ήπιους ρυθμούς.

Το 10–90 % αρχίζει να «αποκλίνει» και ο Dice κυμαίνεται χαμηλότερα ($\sim 0,32$ –0,35).

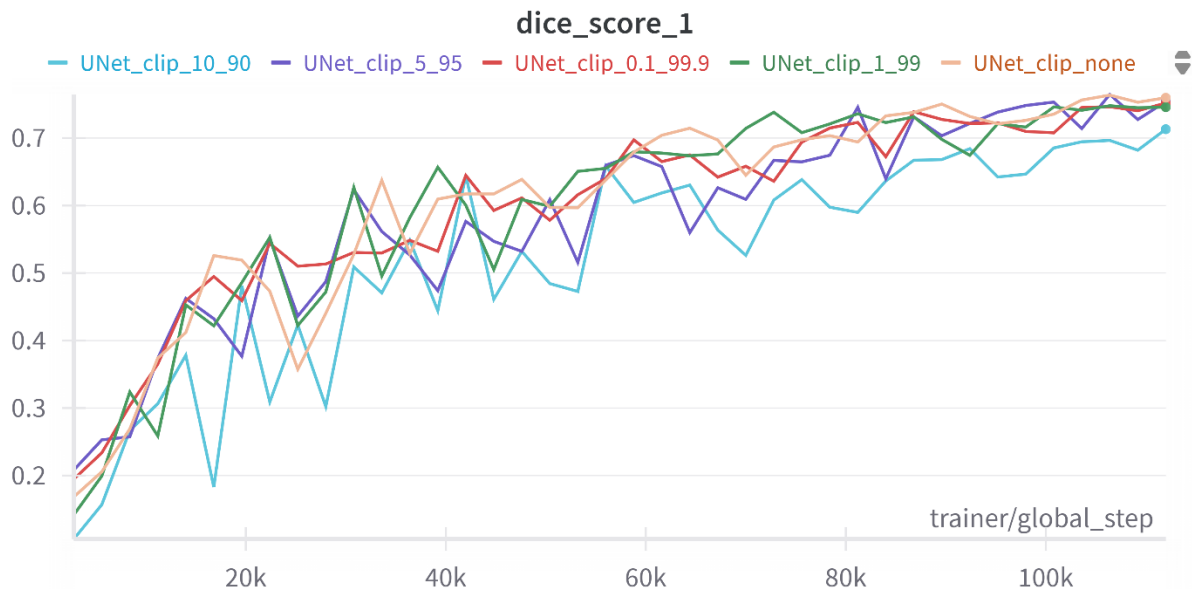
Plateau (≥ 80 k):

Καλύτερο plateau: καμία αποκοπή – Dice $\approx 0,48$.

Πολύ κοντά βρίσκονται τα 1–99 % και 0,1–99,9 % ($\approx 0,47$).

Το 5–95 % σταθεροποιείται χαμηλότερα ($\approx 0,46$).

10–90 % μένει αισθητά πίσω ($\approx 0,44$).

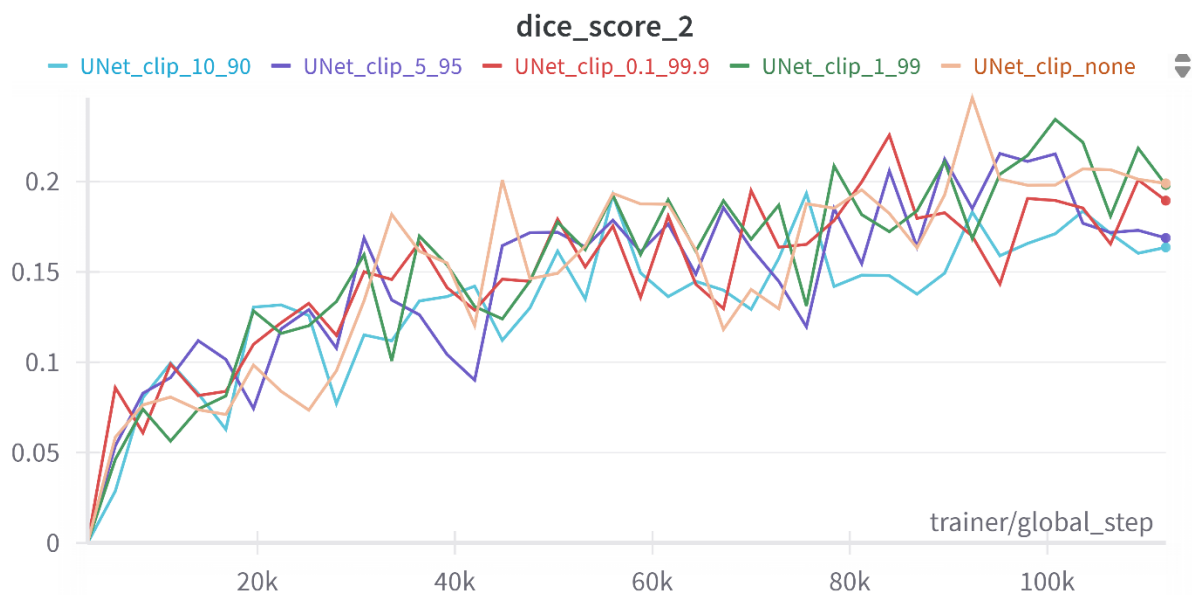


Διάγραμμα 15: Διάγραμμα Dice Score 1 για τις διαφορετικές τιμές Clipping

Σχολιασμός:

Η αφαίρεση ακραίων εντάσεων $< 5\%$ επηρεάζει ελάχιστα τον Dice1: όλες οι ρυθμίσεις 0,1–99,9 %, 1–99 % και 5–95 % παραμένουν εντός $\pm 0,01$ από το «None».

Επιθετικό clipping 10–90 % μειώνει αισθητά τον Dice1 (-0,04), επιβεβαιώνοντας ότι η απώλεια δυναμικού εύρους οδηγεί σε under-segmentation ακόμη και στη μεγάλη δομή.



Διάγραμμα 16: Διάγραμμα Dice Score 2 για τις διαφορετικές τιμές Clipping

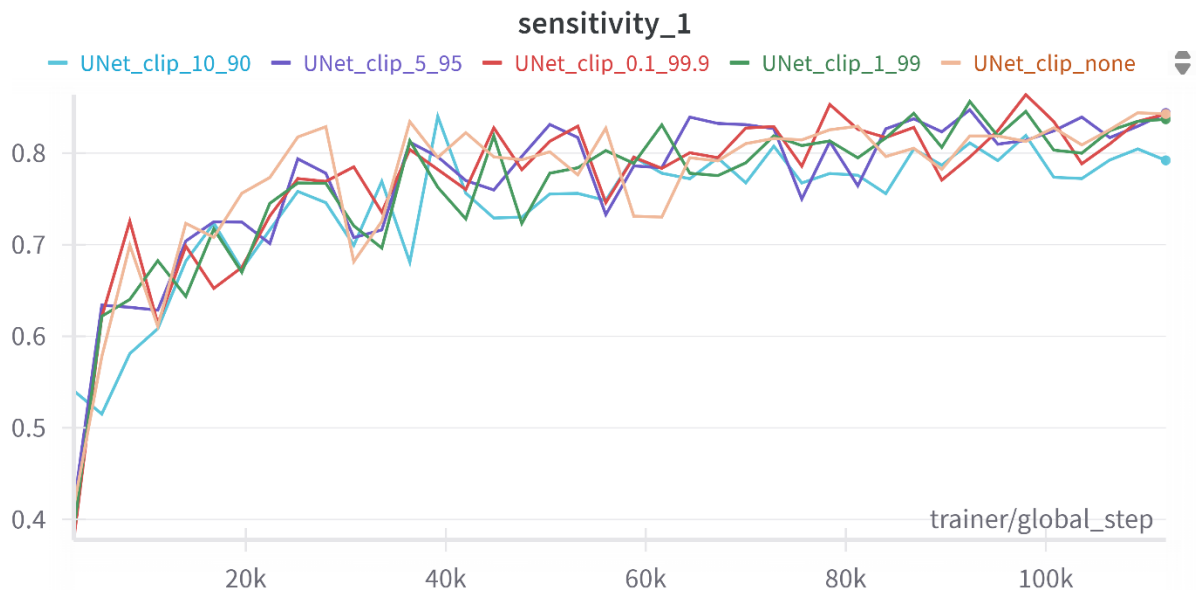
Σχολιασμός:

Χωρίς clipping αποδίδει το υψηλότερο Dice₂ (0,199).

Το ήπιο 0,1–99,9 % υστερεί μόλις κατά 0,01.

Το ενδιάμεσο 1–99 % κρατά σχεδόν αμετάβλητο τον Dice₂ (0,198), θυσιάζοντας ελάχιστο recall.

Ο Dice₂ μειώνεται σημαντικά στο clipping 5–95 % (-0,03) και ακόμη περισσότερο στο 10–90 % (-0,035), ένδειξη ότι χάνονται συστηματικά οι λεπτές βλάβες.



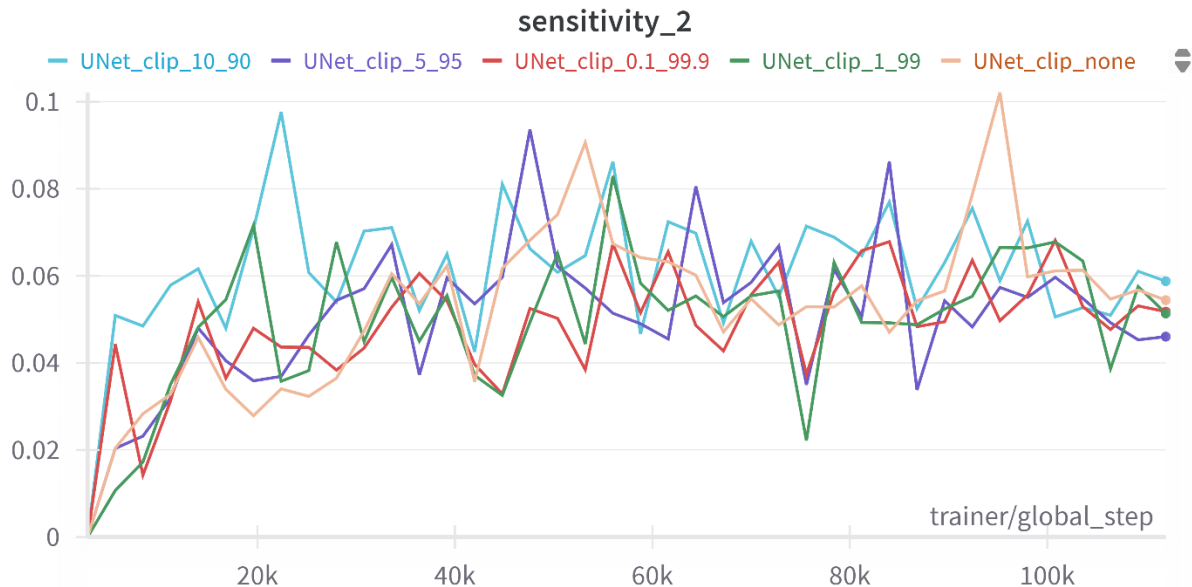
Διάγραμμα 17: Διάγραμμα Sensitivity 1 για τις διαφορετικές τιμές Clipping

Σχολιασμός:

Χωρίς clipping ή με ήπια αποκοπή (0,1–99,9 % και 1–99 %) η ευαισθησία παραμένει σταθερά υψηλή (~0,84).

Το 5–95 % διατηρεί το υψηλότερο Sens₁ (0,844), ένδειξη ότι η ελαφριά εξάλειψη των εντελώς ακραίων voxels βοηθά το μοντέλο να εντοπίσει τα όρια της μεγάλης δομής χωρίς να χάνει πληροφορία.

Το 10–90 % ρίχνει το Sens₁ στο 0,792· δηλαδή χάνονται περίπου 6 % από τα πραγματικά θετικά voxels της μεγάλης κλάσης, συνεπώς το μοντέλο απορρίπτει τμήματα της δομής.



Διάγραμμα 18: Διάγραμμα Sensitivity 2 για τις διαφορετικές τιμές Clipping

Σχολιασμός:

Οι τιμές είναι γενικά χαμηλές λόγω class-imbalance, όμως:

Το 10–90 % εμφανίζει τη μεγαλύτερη τιμή (0,059). Αυτό οφείλεται στο ότι το clipping καταστέλλει artifactual outliers, καθιστώντας τις μικρές βλάβες πιο «ορατές».

Η διαφορά από το None είναι μικρή (+0,005), όμως, όπως φάνηκε στον Dice₂, αυτή η αύξηση συνοδεύεται από αρκετά περισσότερα false positives (χαμηλή precision, άρα τελικό Dice μένει χαμηλά).

Το 5–95 % έχει τη χαμηλότερη ευαισθησία (0,046), επιβεβαιώνοντας ότι η περικοπή 5 % των άκρων αρχίζει να αφαιρεί κλινικά χρήσιμες πληροφορίες.

7.3. Αξιολόγηση και ερμηνεία αποτελεσμάτων

Αυτή η ενότητα προσφέρει μια ολοκληρωμένη αξιολόγηση των πειραματικών αποτελεσμάτων, ερμηνεύοντάς τα στο πλαίσιο της υπάρχουσας βιβλιογραφίας. Η απόδοση του μοντέλου τονίζεται ανά παραμετροποίηση προεπεξεργασίας και αναλύονται οι διαφορές στις μετρήσεις Dice, ευαισθησίας, ειδικότητας και θετικής προγνωστικής αξίας. Επιπλέον, η συμπεριφορά του μοντέλου κατά τη διάρκεια της εκπαίδευσης εξετάζεται στο πλαίσιο των φαινομένων υπερβολικής ή ανεπαρκούς εκπαίδευσης, γενίκευσης και ευαισθησίας σε μικρές δομές.

7.3.1 Επίδραση Spacing

Τα ευρήματα δείχνουν ότι η χωρική ανάλυση των δεδομένων έχει σημαντική επίδραση στην ανίχνευση όγκων (κατηγορία 2). Η χρήση λεπτότερων τομών, που χαρακτηρίζονται από μειωμένη απόσταση και εγγύτητα στην ισοτροπική ανάλυση, οδήγησε σε σημαντική βελτίωση του δείκτη Dice για τον όγκο. Συγκεκριμένα, η επαναδειγματοληψία σε σχεδόν ισότροπο voxel (π.χ. $1,5 \times 1,5 \times 1,5$ mm) απέδωσε Dice₂ ~0,17, το οποίο είναι σημαντικά υψηλότερο σε σύγκριση με μια πιο αραιή ανάλυση, όπως πάχος τομής 5 mm (Dice₂ ~0,15). Αυτό, δείχνει ότι οι λεπτομερείς χωρικές πληροφορίες είναι σημαντικές για τη διάκριση μικρών, χαμηλής αντίθεσης βλαβών αδενοκαρκινώματος εντός του παγκρεατικού παρεγχύματος. Αντίθετα, η απόδοση στην οριοθέτηση του παγκρέατος (κατηγορία 1) έδειξε το σκορ Dice να κυμαίνεται από περίπου 0,74 έως 0,75 σε όλες τις επαναλήψεις, διατηρώντας υψηλή ευαισθησία ακόμη και με μειωμένη απόσταση. Το μοντέλο παρουσιάζει ικανοποιητική αναπαράσταση της μορφολογίας του μεγαλύτερου οργάνου ακόμη και σε χαμηλότερη ανάλυση, ενώ ο μικρός όγκος απαιτεί πυκνότερα δεδομένα για αξιόπιστη ανίχνευση. Το παραπάνω συμπέρασμα είναι σύμφωνο με τη βιβλιογραφία. Σε μελέτες σύγκρισης μεθόδων τμηματοποίησης, η διάκριση μεγάλων οργάνων όπως το ήπαρ ή ο σπλήνας θεωρείται ευκολότερη, ενώ οι μικροί όγκοι του παγκρέατος συγκαταλέγονται μεταξύ των πιο δύσκολων περιπτώσεων. Ακόμη και οι πιο προηγμένες μέθοδοι έχουν επιτύχει μέσο όρο βαθμολογίας Dice μόνο ~0,21 για τον όγκο και ~0,69 για το ίδιο το πάγκρεας [1].

Η μείωση της απόστασης οδήγησε σε βελτίωση της ευαισθησίας της ανίχνευσης μικρών μάζων, καθώς είχε ως αποτέλεσμα την εμφάνιση πιο λεπτομερών χαρακτηριστικών στο δίκτυο. Ωστόσο, παρατηρήθηκε επίσης μια φθίνουσα απόδοση: η μετάβαση από 1,5 χιλιοστά σε πλήρως ισότροπη απόσταση 1,0 χιλιοστών δεν αύξησε σημαντικά τη βαθμολογία Dice του όγκου. Αυτό το εύρημα υποδηλώνει την ύπαρξη ενός ορίου πέρα από το οποίο η περαιτέρω αύξηση της ανάλυσης δεν οδηγεί σε αντίστοιχη βελτίωση. Αυτό το όριο μπορεί να προκύπτει από το γεγονός ότι η βελτίωση της λεπτομέρειας έχει ως κόστος την αύξηση του θορύβου ή, εναλλακτικά, η διαφορά στην ανάλυση μπορεί να μην έχει σημαντική βαρύτητα από πλευράς πληροφοριών. Δεδομένου του υπολογιστικού κόστους, το οποίο περιλαμβάνει σημαντική αύξηση του όγκου των δεδομένων και του χρόνου εκπαίδευσης για απόσταση 1,0 mm, η επιλογή απόστασης περίπου $1,5 \times 1,5 \times 1,5$ mm κρίθηκε ως η βέλτιστη ισορροπία μεταξύ ακρίβειας και απλότητας. Αυτή η παραμετροποίηση επιτυγχάνει ακρίβεια τμηματοποίησης

που είναι σχεδόν ισοδύναμη με εκείνη του πλήρως ισότροπου μοντέλου, ενώ παρουσιάζει σημαντικά μειωμένες απαιτήσεις μνήμης και χρόνου. Αντίθετα, η χρήση διαστήματος όπως $1,0 \times 1,0 \times 5,0$ mm (που παρουσιάζει ένα πολύ παχύ στρώμα), ενώ είναι πιο απλή από υπολογιστική άποψη, είχε ως αποτέλεσμα μια αισθητή μείωση της ευαισθησίας ανίχνευσης όγκων (υποεκπαίδευση σε σχέση με μικρές δομές). Αυτό αποδίδεται στο γεγονός ότι οι χωρικές πληροφορίες σε κάθε τομή είναι ανεπαρκείς για την καταγραφή των περιπλοκών του όγκου.

7.3.2 Επίδραση Interpolation

Η σύγκριση διαφορετικών μεθόδων interpolation κατά τη διάρκεια της επαναδειγματοληψίας αποκάλυψε επίσης σημαντικές διαφορές στην απόδοση, ειδικά όσον αφορά τον όγκο. Η διαδικασία αξιολόγησης περιελάμβανε την εφαρμογή γραμμικής παρεμβολής (1ου βαθμού) και παρεμβολών υψηλότερης τάξης (2ου έως και 5ου βαθμού). Τα ευρήματα έδειξαν ότι η εφαρμογή απλής γραμμικής παρεμβολής στις εικόνες είχε ως αποτέλεσμα την υψηλότερη βαθμολογία Dice για τον όγκο (π.χ. Dice₂ ~0,19), ξεπερνώντας έτσι την απόδοση των μεθόδων υψηλότερης τάξης που αξιολογήθηκαν. Αντίθετα, η παρεμβολή 3ου-4ου βαθμού δεν βελτίωσε την ακρίβεια. Στην πραγματικότητα, παρατηρήθηκε ελαφρώς χαμηλότερη βαθμολογία Dice (~0,15–0,17) για τον όγκο σε σύγκριση με τη γραμμική παρεμβολή. Αυτό το φαινόμενο μπορεί να αποδοθεί στα εγγενή χαρακτηριστικά των δεδομένων και στη φύση της εκάστοτε εργασίας. Οι παρεμβολές υψηλότερης τάξης εισάγουν επιπλέον εξομάλυνση και κανονικοποίηση στις εντάσεις των εικονοστοιχείων, γεγονός που μπορεί να οδηγήσει σε απώλεια τοπικής αντίθεσης στα όρια μικρών δομών. Ο όγκος στο πάγκρεας παρουσιάζει λεπτές διακυμάνσεις φωτεινότητας σε σχέση με τον περιβάλλοντα υγιή ιστό. Μια πιο επιθετική παρεμβολή μπορεί να εξομαλύνει αυτές τις διακυμάνσεις, καθιστώντας τον όγκο λιγότερο διακριτό στο δίκτυο. Αντίθετα, η γραμμική παρεμβολή έχει αποδειχθεί ότι διατηρεί επαρκώς τα όρια και τις σχετικές αντιθέσεις χωρίς να εισάγει σημαντικά τεχνητά στοιχεία. Είναι σημαντικό να σημειωθεί ότι η χρήση της μεθόδου του πλησιέστερου γείτονα (0ου βαθμού) για την επεξεργασία εικόνων μπορεί να οδηγήσει σε aliasing και απώλεια συνεχούς πληροφορίας. Αυτό αναμένεται να κάνει τη διαδικασία εκπαίδευσης πιο δύσκολη. Όπως αναφέρεται στη διαθέσιμη βιβλιογραφία, οι βέλτιστες προσεγγίσεις τμηματοποίησης (π.χ. το πλαίσιο nnU-Net) προτείνουν την κατάλληλη επιλογή μεθόδου παρεμβολής με βάση τα χαρακτηριστικά των δεδομένων [16].

Στην περίπτωση μας, η γραμμική παρεμβολή αποδείχθηκε επαρκής και μάλιστα ανώτερη από πιο σύνθετους παρεμβολείς, πιθανώς επειδή επιτυγχάνει τη βέλτιστη ισορροπία μεταξύ της διατήρησης κρίσιμων λεπτομερειών και της αποφυγής εισαγωγής θορύβου.

7.3.3 Επίδραση Intensity Normalization

Εξετάστηκε η επίδραση διαφόρων στρατηγικών για την κανονικοποίηση του δυναμικού εύρους των εικόνων (κανονικοποίηση έντασης), συμπεριλαμβανομένης της αποκοπής ακραίων τιμών σε συγκεκριμένα ποσοστιαία όρια. Οι συγκρίσεις περιλάμβαναν δύο

διαμορφώσεις: χωρίς αποκοπή, η οποία χρησιμοποιούσε το πλήρες φάσμα HU, και αποκοπή σε διαδοχικά αυξανόμενα εκατοστημόρια (99–1%, 99,9–0,1%, 95–5%, 90–10%). Η αξιολόγηση έδειξε ότι η υπερβολικά επιθετική κανονικοποίηση επηρεάζει αρνητικά την απόδοση ανίχνευσης όγκων. Για παράδειγμα, η αποκοπή του 10% των εντάσεων στα άκρα (10–90ο εκατοστημόριο) είχε ως αποτέλεσμα χαμηλότερο Dice₂ (~0,16) σε σύγκριση με την απουσία αποκοπής (~0,20). Παρατηρήθηκε ότι ακόμη και η πιο μέτρια αποκοπή (5–95ο εκατοστημόριο) υστερούσε, χωρίς αποκοπή ακραίων τιμών. Αυτό υποδηλώνει ότι το δίκτυο επωφελήθηκε από τη διατήρηση του πλήρους εύρους των εντάσεων CT (που αντιστοιχούν σε μονάδες Hounsfield). Οι ακραίες τιμές πιθανώς περιέχουν πληροφορίες σχετικά με δομές υψηλής ή χαμηλής απορρόφησης που σχετίζονται με τον όγκο. Η αφαίρεση αυτών των δομών μπορεί να οδήγησε στην απώλεια διακριτικών χαρακτηριστικών που διαφορετικά θα βοηθούσαν στη διάκριση μεταξύ όγκου και μη όγκου ιστού.

Αντίθετα, η πιο ήπια κανονικοποίηση (π.χ. clipping 1–99%) δεν παρουσίασε καμία αξιοσημείωτη αρνητική επίδραση. Το Dice₂ παρέμεινε ουσιαστικά αμετάβλητο (~0,198 με 1–99% έναντι ~0,199 χωρίς περικοπή). Αυτό το εύρημα δείχνει ότι η εξάλειψη μόνο των λίγων πραγματικά ακραίων τιμών (που ενδεχομένως περιλαμβάνουν θόρυβο, μεταλλικά εμφυτεύματα ή άλλα τεχνητά στοιχεία) μπορεί να επιτευχθεί χωρίς να διακυβεύονται κρίσιμες πληροφορίες.

Συνοψίζοντας, η καταλληλότερη στρατηγική προεπεξεργασίας όσον αφορά την ένταση ήταν η ελάχιστη κανονικοποίηση, η οποία συνεπαγόταν είτε την απουσία αποκοπής είτε την αποκοπή ενός ελάχιστου ποσοστού ακραίων τιμών. Αυτά τα ευρήματα είναι συνεπή με τις συστάσεις μεθόδων όπως αυτές των Isensee et al. (nnU-Net), οι οποίες τονίζουν την κατάλληλη κανονικοποίηση της έντασης και την χωρική κανονικοποίηση ως βασικά βήματα για τη βελτίωση της γενίκευσης [1].

7.3.4 Συμπεράσματα

Έχει διατυπωθεί ότι η σωστή προεπεξεργασία των δεδομένων (preprocessing) μπορεί να είναι σημαντικότερη από την ίδια την αρχιτεκτονική του δικτύου [1]. Αυτό το φαινόμενο είναι επίσης εμφανές στα αποτελέσματα, καθώς οι διακυμάνσεις στην απόδοση μεταξύ των επαναλήψεων αποδίδονται σε αλλαγές στην προεπεξεργασία, ενώ το βασικό μοντέλο παρέμεινε αμετάβλητο.

Μια ποιοτική ανάλυση των καμπυλών εκπαίδευσης/επικύρωσης (train/val loss και Dice) διευκολύνει μια πιο εμπεριστατωμένη ερμηνεία της απόδοσης. Γενικά, παρατηρήθηκε ότι το μοντέλο παρουσίαζε υπερπροσαρμογή σε περιπτώσεις όπου το σύνολο εκπαίδευσης μπορούσε να προσαρμοστεί πολύ καλά, αλλά το σύνολο επικύρωσης δεν μπορούσε να προσαρμοστεί εξίσου καλά. Για παράδειγμα, στις εκτελέσεις με την υψηλότερη ανάλυση (απόσταση 1,0 mm), η απώλεια εκπαίδευσης συνέχισε να μειώνεται ακόμη και μετά τη σταθεροποίηση του Dice επικύρωσης, υποδηλώνοντας ότι το δίκτυο άρχισε να μαθαίνει λεπτομέρειες συγκεκριμένες για το σύνολο εκπαίδευσης (ενδεχομένως θόρυβος ή ιδιαιτερότητες συγκεκριμένων ασθενών) που δεν είναι γενικεύσιμες. Αντίθετα, σε πιο μέτριες

ρυθμίσεις (π.χ. απόσταση 1,5 mm), οι καμπύλες εκπαίδευσης/επικύρωσης παρουσίαζαν μεγαλύτερη συνέπεια, υποδηλώνοντας ότι η απόδοση επικύρωσης πλησίαζε την απόδοση εκπαίδευσης, υποδηλώνοντας ανώτερη γενίκευση. Αντίθετα, η υπερβολική κανονικοποίηση έντασης (π.χ. κλιπ 10-90%) είχε ως αποτέλεσμα την αδυναμία του μοντέλου να επιτύχει υψηλή βαθμολογία Dice για τον όγκο, ακόμη και εντός του συνόλου εκπαίδευσης, υποδηλώνοντας ανεπαρκή εκπαίδευση. Ένα διακριτικό χαρακτηριστικό αυτής της περίπτωσης είναι η παράλληλη πορεία των καμπυλών εκπαίδευσης και επικύρωσης σε χαμηλότερα επίπεδα απόδοσης, υποδηλώνοντας ότι το μοντέλο δεν είναι σε θέση να αποδώσει καλύτερα λόγω έλλειψης πληροφοριών στα δεδομένα εισόδου. Αντίθετα, απουσία κλιπ, το μοντέλο παρουσίασε σημαντικά μειωμένη απώλεια εκπαίδευσης, υποδηλώνοντας ενισχυμένη ικανότητα να καταγράφει περίπλοκες λεπτομέρειες. Ωστόσο, η απώλεια επικύρωσης παρουσίασε ασυνεπή μείωση, υποδηλώνοντας τάση υπερπροσαρμογής. Συγκριτικά, η βέλτιστη παραμετροποίηση (απόσταση 1,5 mm, γραμμική παρεμβολή, ελάχιστο κούρεμα) είχε ως αποτέλεσμα πιο σταθερή σύγκλιση, όπως αποδεικνύεται από τη συνεπή βελτίωση του μοντέλου τόσο στο σύνολο εκπαίδευσης όσο και στο σύνολο επικύρωσης, χωρίς απότομες διακυμάνσεις. Αυτή η παραμετροποίηση πέτυχε επίσης υψηλή βαθμολογία Dice στην επικύρωση, διατηρώντας παράλληλα μια αμελητέα απόκλιση από την απόδοση στην εκπαίδευση. Αυτό το εύρημα δείχνει ότι η συγκεκριμένη ρύθμιση διευκόλυνε την απόκτηση από το δίκτυο των γενικών χαρακτηριστικών του προβλήματος, όπως η διαμόρφωση του παγκρέατος και τα μοτίβα εκδήλωσης του όγκου, χωρίς να αποσπάται η προσοχή από εξειδικευμένες γνώσεις που στερούνται γενικευσιμότητας. Αυτό έχει ιδιαίτερη σημασία στην παρούσα εφαρμογή, επειδή το σύνολο δεδομένων είναι σχετικά μικρό και υπάρχει κίνδυνος υπερπροσαρμογής. Όπως έχουν επισημάνει και άλλοι ερευνητές, η επιλογή της σωστής προεπεξεργασίας λειτουργεί εν μέρει ως τακτική κανονικοποίησης [1].

Ένα σημαντικό εύρημα είναι η διάκριση μεταξύ ευαισθησίας και ειδικότητας στην ανίχνευση όγκων. Σε όλα τα πειράματα, η ειδικότητα για την κατηγορία 2 ήταν σταθερά υψηλότερη από την ευαισθησία. Κατά συνέπεια, το μοντέλο εμφανίζει μια τάση να παράγει χαμηλό αριθμό ψευδών θετικών, δηλαδή σπάνια επισημαίνει φυσιολογικό ιστό ως όγκο, επιτυγχάνοντας έτσι υψηλή ειδικότητα. Ωστόσο, ταυτόχρονα παραλείπει αρκετές πραγματικές περιοχές όγκων, με αποτέλεσμα χαμηλή ευαισθησία. Για παράδειγμα, σε αρκετές περιπτώσεις, το δίκτυο ανίχνευσε μόνο μέρος του όγκου ή καθόλου όγκο σε δύσκολες περιπτώσεις, αλλά απέφυγε να «χρωματίσει» περιοχές εκτός του παγκρέατος ή εντός του παγκρέατος όπου η παρουσία όγκου ήταν αβέβαιη. Αυτό το φαινόμενο αντικατοπτρίζεται επίσης στους δείκτες ακρίβειας: η θετική προγνωστική αξία για την κατηγορία 2 ήταν συγκριτικά υψηλότερη από την ευαισθησία, υποδηλώνοντας ότι το δίκτυο υιοθέτησε μια πιο συντηρητική προσέγγιση. Δηλαδή, κάθε φορά που προβλέπεται η παρουσία όγκου, η πρόβλεψη είναι πιθανώς σωστή (λίγα ψευδώς θετικά αποτελέσματα), αλλά το δίκτυο συχνά αποτυγχάνει να ανιχνεύσει πλήρως όλες τις εστίες του όγκου. Αυτό το φαινόμενο μπορεί να αποδοθεί στη σημαντική ανισορροπία των κατηγοριών, όπου ο όγκος καταλαμβάνει ένα αμελητέο ποσοστό του όγκου των δεδομένων σε σύγκριση με το υπόβαθρο. Το μοντέλο εκπαιδεύτηκε χρησιμοποιώντας ένα συνδυασμό απώλειας Dice και για τις δύο κατηγορίες. Εάν ο όγκος είναι μικροσκοπικός, μια μερική πρόβλεψη δεν τιμωρείται τόσο αυστηρά όσο ένα εκτεταμένο σφάλμα στην μεγάλη κατηγορία (πάγκρεας ή φόντο). Κατά συνέπεια, το δίκτυο φαίνεται να έχει αναπτύξει

μια στρατηγική για να παρακάμψει την υπερβολική πρόβλεψη, μια διαδικασία που θα έθετε σε κίνδυνο την ακρίβειά του, θέτοντας σε κίνδυνο την πλήρη ανάκληση όλων των όγκων.

Στην πράξη, αυτό σημαίνει ότι χωρίς πρόσθετη επεξεργασία, το μοντέλο θα παραλείψει ορισμένους όγκους (χαμηλότερη ευαισθησία), αλλά είναι απίθανο να ενεργοποιήσει συναγερμό σε περιπτώσεις όπου δεν υπάρχει όγκος (υψηλό specificity). Μια σειρά ερευνών έχει δείξει ότι η πρόκληση που αποτελεί ο προσδιορισμός των ακριβών ορίων των όγκων του παγκρέατος δημιουργεί ασυμφωνίες ακόμη και μεταξύ ακτινολόγων με εξειδικευμένη εκπαίδευση [1]. Αυτό το φαινόμενο εξηγεί γιατί ένα αυτόματο μοντέλο είναι πιο επιρρεπές στο να διακρίνει μόνο τον πιο σαφή πυρήνα του όγκου, σε αντίθεση με την επιθετική πρόβλεψη όλων των πιθανών ορίων διήθησης (η οποία μπορεί να συνεπάγεται αυξημένο κίνδυνο ψευδών θετικών αποτελεσμάτων). Μια πιθανή βελτίωση σε αυτόν τον τομέα θα μπορούσε να είναι η εφαρμογή μεθοδολογιών μετα-επεξεργασίας για την αύξηση της ευαισθησίας. Για παράδειγμα, θα μπορούσε να χρησιμοποιηθεί ένα χαμηλότερο όριο πιθανότητας για την κατηγορία του όγκου και/ή να υιοθετηθεί μια προσέγγιση δύο σταδίων, όπου αρχικά προσδιορίζεται η περιοχή του παγκρέατος και στη συνέχεια γίνεται μια πιο ακριβής ανίχνευση μικρών εστιών εντός αυτής. Πράγματι, οι κορυφαίες προσεγγίσεις στο MSD Challenge χρησιμοποίησαν μετα-επεξεργασία για την εξάλειψη προφανών ψευδών θετικών αποτελεσμάτων ή χρησιμοποίησαν σύνολα μοντέλων για να επιτύχουν μια ανώτερη ισορροπία ευαισθησίας και ειδικότητας [1].

Στη μελέτη αυτή δόθηκε προτεραιότητα στην υψηλή ειδικότητα και, ως εκ τούτου, τα αποτελέσματα πρέπει να ερμηνευθούν υπό αυτό το πρίσμα: μια υψηλή ειδικότητα κοντά στο 1,0 υποδηλώνει ότι ο αλγόριθμος είναι αξιόπιστος, καθώς αυτό που ορίζει ως «όγκο» είναι κατά πάσα πιθανότητα όγκος. Ωστόσο, το σχετικά χαμηλό ποσοστό ευαισθησίας υποδηλώνει ότι ενδέχεται να απαιτούνται περαιτέρω βελτιώσεις για την ανίχνευση όλων των όγκων (ή η αποδοχή ορισμένων ψευδών θετικών αποτελεσμάτων προκειμένου να αυξηθεί η ευαισθησία, ανάλογα με τις κλινικές απαιτήσεις).

Ακολουθεί μια σύνοψη των ευρημάτων: Η αξιολόγηση των αποτελεσμάτων έδειξε ότι η σωστή προεπεξεργασία των δεδομένων είναι ένας καθοριστικός παράγοντας για την επιτυχία στην αυτοματοποιημένη ανίχνευση του αδενοκαρκινώματος του παγκρέατος. Το πιο σημαντικό βήμα ήταν η επαναδειγματοληψία σε βέλτιστη ανάλυση, η οποία συνεπαγόταν την υιοθέτηση μιας ομοιόμορφης ισοτροπικής απόστασης περίπου 1,5 χιλιοστών. Αυτό είχε ως αποτέλεσμα τη σημαντική βελτίωση της τμηματοποίησης του όγκου χωρίς να επιβαρύνεται υπερβολικά το μοντέλο. Επιπλέον, η εφαρμογή γραμμικής παρεμβολής για τις εικόνες είχε ως αποτέλεσμα τη διατήρηση κρίσιμων λεπτομερειών, αποδίδοντας έτσι μεγαλύτερο όγκο Dice σε σύγκριση με πιο σύνθετες μεθόδους παρεμβολής. Όσον αφορά την ομαλοποίηση της έντασης, παρατηρήθηκε ότι αρκούσε η αφαίρεση μόνο των ακραίων τιμών (αντί για τον δραστικό περιορισμό του εύρους). Αυτή η προσέγγιση εξασφαλίζει ότι το δίκτυο εκτίθεται στο σύνολο των πληροφοριών αντίθεσης που υπάρχουν στην αρχική εικόνα. Η απόδοση για την κατηγορία 1 (πάγκρεας) ήταν σταθερά υψηλή υπό όλες τις συνθήκες, υποδηλώνοντας ότι το μοντέλο μαθαίνει αποτελεσματικά το σχήμα και τη θέση του παγκρέατος. Αντίθετα, η κατηγορία 2 (όγκος) συνεχίζει να αποτελεί σημαντική πρόκληση, όπως αποδεικνύεται από την παρατήρηση χαμηλότερου Dice και ιδιαίτερα χαμηλής

ευαισθησίας. Αυτά τα ευρήματα συνάδουν με τη διεθνή βιβλιογραφία, η οποία υπογραμμίζει την πολυπλοκότητα αυτής της συγκεκριμένης εργασίας τμηματοποίησης [1].

Αυτά τα ευρήματα θα συζητηθούν περαιτέρω στην επόμενη ενότητα, με τίτλο «Συζήτηση», στην οποία θα τεθούν σε πλαίσιο με ανάλογες μελέτες και θα προταθούν πιθανές οδοί βελτίωσης, όπως η εφαρμογή πιο περίπλοκων αρχιτεκτονικών ή διαδικασιών, η αύξηση των δεδομένων και η ενσωμάτωση της γνώσης των ειδικών. Συνοπτικά, η μελέτη κατέδειξε ότι με τη χρήση κατάλληλων επιλογών προεπεξεργασίας, μπορεί να επιτευχθεί σημαντική βελτίωση στην απόδοση ανίχνευσης όγκων του παγκρέατος, επιβεβαιώνοντας έτσι τη σημασία της προεπεξεργασίας, όπως έχει υπογραμμιστεί προηγουμένως από άλλους ερευνητές [1]. Αυτή η σειρά αποτελεσμάτων και συμπερασμάτων θα χρησιμεύσει ως σημείο αναφοράς για περαιτέρω ανάλυση στη συζήτηση, όπου η μέθοδος θα αξιολογηθεί κριτικά και θα διερευνηθεί το πεδίο βελτίωσης και μελλοντικής εργασίας.

8. Συζήτηση

8.1. Επίδραση spacing/interpolation/intensity στα αποτελέσματα

Τα πειράματα που διεξήχθησαν αποδεικνύουν ότι η επιλογή του διαστήματος των εικονοστοιχείων, ή της ανάλυσης της εικόνας, ασκεί σημαντική επίδραση στην ακρίβεια της τμηματοποίησης. Συγκεκριμένα, η εκπαίδευση με μειωμένη δειγματοληψία (π.χ. ισότροπο $2,0 \times 2,0 \times 2,0$ mm) παρουσίασε σημαντικά χαμηλότερη απόδοση στην ανίχνευση όγκων σε σύγκριση με τα λεπτότερα διαστήματα εικονοστοιχείων. Το δίκτυο έδειξε αδυναμία ανίχνευσης πολλών μικρών όγκων όταν τα δεδομένα υποβλήθηκαν σε υπερδειγματοληψία (χονδροειδής επαναδειγματοληψία), ένα αποτέλεσμα που ήταν αναμενόμενο δεδομένου του ελάχιστου αριθμού voxels που καταλάμβανε η παγκρεατική βλάβη σε τόσο χαμηλή ανάλυση. Αντίθετα, η χρήση απόστασης κοντά στη μέση ανάλυση εικόνας (π.χ. $1,0 \times 1,0 \times 2,5$ mm, που προσεγγίζει την επικρατούσα ανάλυση των τομών CT) είχε ως αποτέλεσμα βελτιωμένη απόδοση. Συγκεκριμένα, ο δείκτης Dice για τον όγκο παρουσίασε βελτίωση έως και ~50% σε σύγκριση με την υπερβολικά χονδροειδή απόσταση, συνοδευόμενη από αυξημένη ευαισθησία στην ανίχνευση μικροσκοπικών βλαβών. Ωστόσο, η περαιτέρω αύξηση της ανάλυσης σε τέλεια ισότροπη $1,0 \times 1,0 \times 1,0$ mm δεν οδήγησε σε ανάλογη βελτίωση των αποτελεσμάτων. Στην πραγματικότητα, παρατηρήθηκε μια ελαφρά μείωση της απόδοσης. Αυτό το φαινόμενο μπορεί να αποδοθεί, πρώτον, στο γεγονός ότι η υπερδειγματοληψία δεν συμβάλλει σε σημαντικές νέες πληροφορίες, καθώς δημιουργεί τεχνητά τμήματα μέσω παρεμβολής. Δεύτερον, η ανάγκη για μειωμένο μέγεθος patch κατά τη διάρκεια της εκπαίδευσης, που προκλήθηκε από τον αυξημένο όγκο δεδομένων, περιόρισε το οπτικό πεδίο του μοντέλου. Κατά συνέπεια, η βέλτιστη απόσταση προσδιορίστηκε σε ένα ενδιάμεσο επίπεδο, επιτυγχάνοντας μια ισορροπία μεταξύ περίπλοκων λεπτομερειών και επαρκούς οπτικού πεδίου. Αυτό το εύρημα είναι σύμφωνο με τη στρατηγική του nnU-Net για επαναδειγματοληψία όλων των εικόνων στην μέση απόσταση voxel του καθορισμένου συνόλου [20].

Όσον αφορά την παράμετρο παρεμβολής, τα αποτελέσματα έδειξαν μόνο μικρές διακυμάνσεις στην απόδοση τμηματοποίησης, παρά τις σημαντικές διαφορές στη μέθοδο επαναδειγματοληψίας. Διεξήχθη ένα πείραμα στο οποίο αξιολογήθηκαν διάφορες σειρές παρεμβολής σπλάν (από την πρώτη έως την πέμπτη σειρά) κατά τη διαδικασία προεπεξεργασίας της εικόνας. Οι καμπύλες μάθησης και οι τελικές βαθμολογίες Dice ήταν συγκρίσιμες για όλες τις περιπτώσεις, με αποκλίσεις 1-2%. Η γραμμική παρεμβολή (τάξη 1) πέτυχε μια ελαφρώς υψηλότερη μέση τιμή Dice (περίπου 0,47), ενώ οι κυβικές (τάξη 3, προεπιλογή nnU-Net) και οι σπλάν υψηλότερης τάξης (τάξη 4-5) είχαν πολύ παρόμοιες επιδόσεις (0,45–0,47). Η παρατηρούμενη απόκλιση δεν θεωρήθηκε στατιστικά σημαντική, υποδηλώνοντας ότι το δίκτυο επιδεικνύει υψηλό βαθμό ανθεκτικότητας στις διακυμάνσεις των μεθόδων παρεμβολής, υπό την προϋπόθεση ότι χρησιμοποιείται μια ικανοποιητική τεχνική για τη διατήρηση των λεπτομερειών. Αυτό είναι λογικό, καθώς όλοι οι αλγόριθμοι παρεμβολής που χρησιμοποιήθηκαν (τριγωνικός, δικυβικός κ.λπ.) διατηρούν σε μεγάλο βαθμό τις χωρικές πληροφορίες του όγκου. Συνεπώς, μπορεί να συναχθεί το συμπέρασμα ότι η προεπιλεγμένη χρήση της κυβικής παρεμβολής από το nnU-Net είναι επαρκής και ότι η χρήση πιο εξελιγμένων μεθόδων δεν οδηγεί σε ουσιαστική βελτίωση των αποτελεσμάτων της τμηματοποίησης.

Συμπερασματικά, όσον αφορά τη διαδικασία ομαλοποίησης της έντασης, η επίδραση στα αποτελέσματα αποδείχθηκε λιγότερο σημαντική από ό,τι αρχικά αναμενόταν. Αξιολογήθηκε μια σειρά στρατηγικών αποκοπής για τις τιμές Hounsfield των voxels, που κυμαίνονταν από την απουσία περιορισμού (clip_none) έως τον επιθετικό περιορισμό εντός του εύρους έντασης 10–90%. Η διακύμανση στους δείκτες Dice και στην ευαισθησία ανίχνευσης όγκων ήταν ελάχιστη, με αλλαγές που κυμαίνονταν σε μερικά ποσοστιαία σημεία. Για παράδειγμα, η μείωση του δυναμικού εύρους εντάσεων στο 5–95% οδήγησε σε μια μικρή μείωση της ευαισθησίας, που εκτιμάται σε περίπου 0,01–0,02. Αυτή η μείωση μπορεί να αποδοθεί στην αφαίρεση πληροφοριών από τα άκρα που σχετίζονται με τον όγκο, όπως περιοχές υψηλής αντίθεσης που περιβάλλουν έναν ενισχυμένο όγκο. Αντίθετα, η πλήρης απουσία αποκοπής παρουσίασε απόδοση παρόμοια ή ελαφρώς ανώτερη από την προεπιλεγμένη ήπια αποκοπή (1–99%), υποδηλώνοντας ότι η διατήρηση του πλήρους φάσματος HU δεν προκαλεί σημαντική ζημιά στο μοντέλο. Συνοπτικά, η προεπιλεγμένη προσέγγιση του nnU-Net — δηλαδή, ο περιορισμός των ακραίων τιμών έντασης σε περίπου 0,5–99,5% ανά περίπτωση και η επακόλουθη κανονικοποίηση (z-score) — κρίθηκε λογική και επαρκής [20]. Η λογική που υποκρύπτει αυτό το φαινόμενο είναι ότι τόσο ο φυσιολογικός ιστός του παγκρέατος όσο και τα αδενοκαρκινώματα εκδηλώνονται σε συγκρίσιμο εύρος έντασης (μαλακά μόρια, διακριτές ενισχύσεις), αποκλείοντας έτσι την αποτελεσματικότητα του επιθετικού κλιπινγκ στην ενίσχυση του διαχωρισμού τους. Το μοντέλο εκπαιδεύτηκε ώστε να δίνει προτεραιότητα στο σχετικό εύρος έντασης των δομών που παρουσιάζουν ενδιαφέρον, ενώ οι ακραίες τιμές (π.χ. οστά, αέρας) είτε αφαιρέθηκαν είτε τους αποδόθηκε ελάχιστο βάρος κατά τη διάρκεια της διαδικασίας εκπαίδευσης. Συνοψίζοντας, από τις τρεις παραμέτρους που εξετάστηκαν, η γεωμετρική προεπεξεργασία (απόσταση) είχε τη σημαντικότερη επίδραση στην απόδοση της τμηματοποίησης, ενώ οι ρυθμίσεις παρεμβολής και κανονικοποίησης έντασης άσκησαν πιο μέτρια επίδραση, υπό την προϋπόθεση ότι παρέμειναν εντός λογικών ορίων.

8.2. Συγκριτική ερμηνεία σε σχέση με τις συστάσεις του nnU-Net

Τα προαναφερθέντα ευρήματα αξιολογούνται σε σύγκριση με τις συστάσεις και τις ευριστικές μεθόδους που προτείνει το ίδιο το πλαίσιο nnU-Net [21]. Ένα από τα σημαντικότερα χαρακτηριστικά του nnU-Net είναι η ικανότητά του για αυτόνομη διαμόρφωση του βέλτιστου αγωγού, μια διαδικασία που βασίζεται σε ένα σύνολο εμπειρικών κανόνων. Για παράδειγμα, το nnU-Net είναι ένα πλαίσιο βαθιάς μάθησης που αναλύει το σύνολο δεδομένων και ορίζει ένα στόχο απόστασης ίσο με τη μέση απόσταση voxel ανά άξονα (ή το 10ο εκατοστημόριο για άξονες με ισχυρή ανισοτροπία). Στην παρούσα μελέτη, η βέλτιστη απόδοση επιτεύχθηκε όταν χρησιμοποιήθηκε απόσταση που ήταν πολύ κοντά στη μέση ανάλυση των διαθέσιμων αξονικών τομογραφιών (πάχος τομής περίπου 2,5 mm), επιβεβαιώνοντας έτσι την εγκυρότητα αυτής της σύστασης. Αντίθετα, έχει παρατηρηθεί ότι οι αποκλίσεις από αυτό το πρότυπο (π.χ. αυξημένη ή μειωμένη πυκνότητα δειγματοληψίας) υποβαθμίζουν την αποτελεσματικότητα των αποτελεσμάτων, επιβεβαιώνοντας τα ευρήματα των Isensee et al. ότι η διατήρηση της αρχικής ανάλυσης βελτιώνει την ακρίβεια σε σύγκριση με την υπερβολική υποδειγματοληψία.

Επιπλέον, το nnU-Net, από προεπιλογή, εφαρμόζει τριγωνική (3ης τάξης) παρεμβολή για την επαναδειγματοληψία ιατρικών εικόνων και την πλησιέστερη γειτονική για μάσκες, θεωρώντας τις επαρκείς. Τα πειραματικά αποτελέσματα έδειξαν ότι η κυβική παρεμβολή παρουσίασε επίπεδο απόδοσης συγκρίσιμο με αυτό των εναλλακτικών λύσεων υψηλότερης τάξης. Επιπλέον, η γραμμική παρεμβολή δεν παρουσίασε καμία σημαντική καθυστέρηση. Κατά συνέπεια, δεν υπάρχει επιτακτικός λόγος να αποκλίνουμε από τις προεπιλεγμένες ρυθμίσεις σχετικά με την παρεμβολή, επιβεβαιώνοντας έτσι την αξιοπιστία της προεπιλεγμένης επιλογής του πλαισίου σε πρακτικές εφαρμογές.

Στο πλαίσιο της κανονικοποίησης της έντασης, το nnU-Net χρησιμοποιεί μια προσέγγιση βασισμένη σε δεδομένα που περιλαμβάνει τον υπολογισμό του μέσου όρου και της τυπικής απόκλισης των εντάσεων, συνήθως για τα βόξελ του προσκηνίου, σε ολόκληρο το σύνολο εκπαίδευσης για τις αξονικές τομογραφίες. Στη συνέχεια, κάθε εικόνα κανονικοποιείται ως $((I - \mu)/\sigma)$, με τις ακραίες τιμές να αποκόπτονται πρώτα πέρα από τα 99,5% και 0,5% εκατοστημόρια [20]. Αυτή η προεπιλεγμένη τακτική αντιστοιχεί στο πείραμα clip 1–99 + z-score που περιγράφηκε προηγουμένως. Τα πειραματικά αποτελέσματα έδειξαν ότι η εφαρμογή μιας πιο αυστηρής παρέμβασης, όπως η χρήση clip 5–95 ή 10–90, δεν οδήγησε σε σημαντική βελτίωση του διαχωρισμού μεταξύ του όγκου και του υγιούς ιστού. Αντίθετα, η απουσία αποκοπής μπορεί να οδηγήσει σε αύξηση του θορύβου χωρίς ουσιαστική αλλαγή στην ακρίβεια. Συνεπώς, συνάγεται το συμπέρασμα ότι οι προεπιλεγμένες ρυθμίσεις του nnU-Net για την προεπεξεργασία έντασης είναι επαρκώς κατάλληλες για το πρόβλημα και δεν απαιτείται παρέμβαση. Αξίζει να σημειωθεί ότι το nnU-Net αναπτύχθηκε ακριβώς με στόχο την βέλτιστη αυτοματοποίηση αυτών των επιλογών, και τα αποτελέσματα αυτής της μελέτης είναι σύμφωνα με αυτή τη φιλοσοφία. Οποιοσδήποτε χειροκίνητες τροποποιήσεις στις προεπιλεγμένες ευριστικές μεθόδους δεν απέφεραν σημαντικά οφέλη, επιβεβαιώνοντας ότι η προσεκτική διαμόρφωση της μεθόδου είναι πιο σημαντική από την τροποποίηση της αρχιτεκτονικής του μοντέλου [33].

Μια αξιοσημείωτη παρατήρηση είναι η οριακή αύξηση της ευαισθησίας ανίχνευσης όγκων που παρατηρήθηκε με την εφαρμογή επιθετικής περικοπής στο 10-90%. Αν και αυτή η ρύθμιση δεν αποτελεί σύσταση του nnU-Net, παρατηρήθηκε ότι το δίκτυο με κανονικοποίηση 10-90% ανίχνευσε μερικούς επιπλέον μικρούς όγκους που δεν είχαν εντοπιστεί σε άλλες ρυθμίσεις, όπως υποδεικνύεται από ορισμένους δείκτες (π.χ. ευαισθησία ανά περίπτωση). Ωστόσο, αυτό συνοδεύτηκε από μια μικρή μείωση του μέσου όρου της βαθμολογίας Dice (λόγω υπερεκτίμησης του θορύβου ως όγκου), καθιστώντας την βελτίωση μάλλον μη επαναλαμβανόμενη. Σε κάθε περίπτωση, οι εν λόγω αποκλίσεις ήταν αμελητέες. Κατά συνέπεια, δεν συνιστάται καμία αλλαγή στις επικρατούσες πρακτικές του nnU-Net, όπως τεκμηριώνεται από τα ευρήματα της παρούσας μελέτης. Στην πραγματικότητα, τα αποτελέσματα επιβεβαιώνουν τις προεπιλεγμένες διαμορφώσεις του πλαισίου, οι οποίες θεωρούνται σχεδόν βέλτιστες για αυτό το συγκεκριμένο σύνολο δεδομένων.

8.3. Περιορισμοί της μελέτης

Παρά τα ευνοϊκά αποτελέσματα, είναι επιτακτική ανάγκη να αναγνωριστούν οι εγγενείς περιορισμοί της μελέτης, οι οποίοι πρέπει να ληφθούν υπόψη κατά την ερμηνεία των αποτελεσμάτων.

Το μέγεθος του δείγματος και η ποικιλία του συνόλου των δεδομένων είναι κρίσιμοι παράγοντες για τον προσδιορισμό της αξιοπιστίας και της εγκυρότητας των αποτελεσμάτων. Ο αριθμός των αξονικών τομογραφιών που χρησιμοποιήθηκαν για σκοπούς εκπαίδευσης και αξιολόγησης ήταν περιορισμένος. Παρά την εφαρμογή της διασταυρούμενης επικύρωσης, το περιορισμένο μέγεθος του συνόλου δεδομένων μπορεί να αυξήσει τον κίνδυνο υπερπροσαρμογής και ενδέχεται να μην καλύπτει όλο το φάσμα των μορφολογιών του παγκρέατος και των όγκων. Η αντιπροσωπευτικότητα του μοντέλου όσον αφορά τις διαφορές στους σαρωτές, τα πρωτόκολλα απεικόνισης και τις φάσεις αντίθεσης είναι περιορισμένη, γεγονός που εγείρει ερωτήματα σχετικά με τη γενικευσιμότητα του μοντέλου σε ανεξάρτητα δεδομένα. Προκειμένου να εξαχθούν πιο αξιόπιστα συμπεράσματα, θα απαιτηθεί ένα μεγαλύτερο, πολυκεντρικό δείγμα.

Έχει παρατηρηθεί μια έντονη ανισότητα στην κατηγορία και σημαντικές διαστάσεις όγκων. Οι όγκοι του παγκρέατος, συγκεκριμένα τα αδενοκαρκινώματα, καταλαμβάνουν συνήθως ένα αμελητέο μέρος του όγκου δεδομένων μιας αξονικής τομογραφίας της κοιλιακής χώρας, που συχνά ανέρχεται σε λιγότερο από το 0,1% του συνολικού όγκου voxels [33]. Αυτή η σοβαρή ανισορροπία στην κατηγορία έχει ως αποτέλεσμα την έλλειψη διαθέσιμων θετικών δειγμάτων κατηγορίας για εκπαίδευση, επηρεάζοντας έτσι αρνητικά κοινές μετρήσεις όπως το Dice και το IoU. Παρά την ικανότητα ενός μοντέλου να ανιχνεύει με ακρίβεια πολλαπλούς όγκους, ο μέσος όρος της βαθμολογίας Dice παραμένει χαμηλός λόγω του αποκλεισμού ενός μεμονωμένου, μικρού όγκου, ο οποίος έχει ως αποτέλεσμα βαθμολογία Dice 0 για τη συγκεκριμένη περίπτωση. Στην παρούσα μελέτη, παρατηρήθηκαν αρκετές περιπτώσεις στις οποίες ο όγκος δεν ανιχνεύθηκε, με αποτέλεσμα βαθμολογία Dice μηδέν για τις περιπτώσεις αυτές. Αυτή η σημαντική μείωση της μέσης βαθμολογίας Dice υπογραμμίζει τη σημασία των ολοκληρωμένων μεθόδων ανίχνευσης στην ογκολογική απεικόνιση. Αυτός

είναι ένας γνωστός περιορισμός και σε άλλες μελέτες: συνολικά, οι αλγόριθμοι στο MSD Challenge είχαν μέση τιμή DSC μόνο $\sim 0,21$ για όγκους του παγκρέατος [1], αντανακλώντας με ακρίβεια τη μεγάλη δυσκολία του προβλήματος.

Έχει παρατηρηθεί χαμηλή αντίθεση και αβεβαιότητα στα όρια. Συχνά παρατηρείται αξιοσημείωτη ομοιότητα στην ακτινολογική πυκνότητα μεταξύ του παγκρεατικού παρεγχύματος και του αδενοκαρκινώματος στις αξονικές τομογραφίες, ιδιαίτερα κατά τη διάρκεια της πυλαίας φλεβικής φάσης. Η διαφορά στην ένταση μεταξύ του υγιούς παγκρέατος και του όγκου είναι ελάχιστη, με αποτέλεσμα τα όρια να είναι ασαφή [1]. Αυτό περιπλέκει τις διαδικασίες αυτόματης ανίχνευσης και τη δημιουργία πραγματικών δεδομένων από τους ακτινολόγους. Στην βιβλιογραφία επισημαίνεται ότι οι μικροί παγκρεατικοί όγκοι χάνουν τον διακριτό τους χαρακτήρα μέσα στον αδένα και συχνά δεν υπάρχει απόλυτη συμφωνία μεταξύ εμπειρογνομόνων για την έκτασή τους [37]. Κατά συνέπεια, ένα μέρος του παρατηρούμενου σφάλματος μπορεί να αποδοθεί στην παρουσία θορύβου στις σημειώσεις. Συγκεκριμένα, το μοντέλο μπορεί να ορίσει περιοχές που δεν περιλαμβάνονται στην μάσκα αναφοράς. Ωστόσο, είναι δύσκολο να διαπιστωθεί εάν αυτές οι περιοχές περιέχουν πράγματι διηθημένο ιστό. Η ασάφεια στην οριοθέτηση δημιουργεί δύο βασικά προβλήματα: ψευδώς θετικά αποτελέσματα, που χαρακτηρίζονται από υπερβολική τμηματοποίηση του φυσιολογικού παγκρεατικού ιστού, και ψευδώς αρνητικά αποτελέσματα, που χαρακτηρίζονται από την αποτυχία τμηματοποίησης των υποεντατικών τμημάτων του όγκου.

Η αξιολόγηση είναι περιορισμένη και δεν υπάρχει εξωτερική επικύρωση. Η αξιολόγηση της απόδοσης βασίστηκε σε διασταυρωτική επικύρωση εντός του διαθέσιμου συνόλου δεδομένων. Η απουσία ενός ανεξάρτητου συνόλου δοκιμών από διαφορετικό κέντρο αποκλείει την επιβεβαίωση του γενικού συμπεράσματος. Αυτό αποτελεί περιορισμό, καθώς η πραγματική κλινική απόδοση ενός μοντέλου είναι εμφανής μόνο όταν δοκιμάζεται σε εντελώς νέα δεδομένα. Επιπλέον, η αξιολόγηση πραγματοποιήθηκε κυρίως μέσω της χρήσης μετρήσεων όπως το Dice και η ευαισθησία σε επίπεδο voxel. Παρά την καθιερωμένη φύση τους, αυτοί οι δείκτες δεν αντικατοπτρίζουν πλήρως την κλινική χρησιμότητα. Από την οπτική γωνία ενός ογκολόγου, είναι επιτακτική ανάγκη να ανιχνεύονται όλοι οι όγκοι (δηλ. ευαισθησία ανά περίπτωση), ακόμη και αν η ακριβής τμηματοποίηση είναι δευτερεύουσα. Κατά τη διάρκεια της έρευνάς μας, συναντήσαμε περιπτώσεις όπου ένα μερικώς τμηματοποιημένο εύρημα, που χαρακτηριζόταν από ατέλεια, απέδωσε χαμηλή βαθμολογία Dice, αλλά θα μπορούσε να θεωρηθεί επιτυχές από την άποψη της ανίχνευσης. Είναι επιτακτική ανάγκη να αναγνωριστούν οι διαφορές στην αντίληψη μεταξύ των μετρήσεων και της κλινικής αξιολόγησης κατά την ερμηνεία αυτών των ευρημάτων.

Συνοψίζοντας, οι προαναφερθείσες περιορισμοί υποδηλώνουν ότι τα αποτελέσματα, παρά τον ενθαρρυντικό τους χαρακτήρα, πρέπει να θεωρηθούν προκαταρκτικά. Απαιτείται περαιτέρω μελέτη με περισσότερα δεδομένα και ευρύτερη αξιολόγηση για να επιβεβαιωθεί η γενική εγκυρότητα των συμπερασμάτων και η σταθερότητα της μεθόδου σε ρεαλιστικές συνθήκες.

8.4. Προτάσεις βελτίωσης

Λαμβάνοντας υπόψη τα σημεία, προτείνονται οι παρακάτω συμβουλές για μελλοντικές βελτιώσεις. Αυτές οι βελτιώσεις μπορεί να γίνουν κρατώντας την δομή του nnU-Net, αποφεύγοντας έτσι μεγάλες αλλαγές στο αρχιτεκτονικό του σχέδιο.

Η προσαρμογή των ευριστικών μεθόδων και των υπερπαραμέτρων είναι μια κρίσιμη πτυχή αυτής της διαδικασίας. Μια πιθανή προσέγγιση περιλαμβάνει τη βελτίωση των αυτοματοποιημένων ρυθμίσεων του nnU-Net. Για παράδειγμα, η συχνότητα δειγματοληψίας των θετικών voxels θα μπορούσε να αυξηθεί κατά την εξαγωγή των patch εκπαίδευσης. Το nnU-Net έχει αποδειχθεί ότι εγγυάται την παρουσία τουλάχιστον ενός βόξελ προσκηνίου σε κάθε patch. Ωστόσο, μπορεί να είναι απαραίτητη μια ακόμη πιο επιθετική υπερκάλυψη των περιοχών του όγκου, ώστε να διασφαλιστεί ότι το δίκτυο συναντά σπάνια παραδείγματα με μεγαλύτερη συχνότητα. Επιπλέον, η προεπιλεγμένη συνάρτηση κόστους (απώλεια Dice+CE) θα μπορούσε να εμπλουτιστεί με όρους που στοχεύουν συγκεκριμένα στην ανίχνευση μικρών αντικειμένων, όπως η χρήση της απώλειας Focal Loss ή η αύξηση του βάρους της κατηγορίας όγκου στην απώλεια. Αυτές οι τροποποιήσεις εμπίπτουν στο πεδίο της βελτιστοποίησης της ίδιας διαδικασίας χωρίς να αλλάζουν το μοντέλο. Επιπλέον, θα μπορούσε να εξεταστεί ο προγραμματισμός του ρυθμού μάθησης και του αριθμού των εποχών εκπαίδευσης. Είναι πιθανό να απαιτείται περισσότερη εκπαίδευση ή χαμηλότερος ρυθμός μάθησης στο τέλος (πλατό/ανόπτηση) για να μάθει το δίκτυο τις λεπτές δομές.

Βέλτιστη επιλογή και αύξηση του μεγέθους του patch: Όπως αναφέρθηκε προηγουμένως, το nnU-Net επέλεξε ένα σχετικά περιορισμένο μέγεθος εισόδου (patch) για να χωρέσει σε μια GPU ~10GB. Σε περίπτωση που υπάρχουν στη διάθεση κάποιου πιο σημαντικοί υπολογιστικοί πόροι, θα ήταν πλεονεκτικό να αυξηθεί το μέγεθος του patch (δηλ. ο υποόγκος της CT που επεξεργάζεται το δίκτυο σε κάθε επανάληψη) [20].

Ένα πιο εκτεταμένο patch μπορεί να περιλαμβάνει ολόκληρο το πάγκρεας και τα γύρω όργανα, παρέχοντας έτσι στο δίκτυο πιο ουσιαστικό βάθος συμφραζομένων. Αυτό πιθανότατα θα διευκόλυνε τη διάκριση του όγκου από τις γειτονικές δομές, για παράδειγμα, αναγνωρίζοντας ότι μια περιοχή χαμηλής αντίθεσης εντός του παγκρέατος είναι ενδεικτική κακοήθειας, σε συνδυασμό με τη συνολική μορφολογία του οργάνου. Επομένως, συνιστάται η διερεύνηση μεγαλύτερων patch (αν όχι κατά την εκπαίδευση, τουλάχιστον κατά τη διάρκεια της εξαγωγής συμπερασμάτων με παράθυρα υψηλής ανάλυσης). Αυτή η βελτίωση παραμένει εντός του οικοσυστήματος nnU-Net, καθώς το πλαίσιο επιτρέπει στον χρήστη να καθορίσει προσαρμοσμένα μεγέθη patch και παρτίδων, υπό την προϋπόθεση ότι υπάρχει επαρκής μνήμη.

Η παρούσα μελέτη προτείνει μια στρατηγική δύο σταδίων (καταρράκτη) εντός του nnU-Net. Λαμβάνοντας υπόψη τη φύση του προβλήματος, που χαρακτηρίζεται από ένα μικρό όγκο ενσωματωμένο σε ένα μεγαλύτερο όργανο, μια προσέγγιση δύο σταδίων μπορεί να αποδειχθεί πιο αποτελεσματική. Πράγματι, η τρισδιάστατη έκδοση του nnU-Net υποστηρίζει μοντέλα καταρράκτη, που περιλαμβάνουν ένα πρώτο δίκτυο με χαμηλότερη ανάλυση και ένα δεύτερο με υψηλότερη ανάλυση, με το δεύτερο δίκτυο να εστιάζει στην περιοχή ενδιαφέροντος. Στο μέλλον, θα μπορούσαμε να εκπαιδεύσουμε ένα πρώτο nnU-Net για τον

κατακερματισμό του παγκρέατος (κατηγορία 1) ώστε να αναγνωρίζει αυτόματα την περιοχή του οργάνου σε κάθε τομή. Στη συνέχεια, ένα δευτερεύον nnU-Net θα εκπαιδεύονταν για να αναγνωρίζει τον όγκο (κατηγορία 2) εντός ενός τμήματος του συνόλου δεδομένων που δίνει προτεραιότητα στο πάγκρεας. Κατά συνέπεια, το δεύτερο στάδιο θα επεξεργαζόταν έναν σημαντικά μειωμένο όγκο (μόνο το πάγκρεας αντί για ολόκληρη την κοιλιακή χώρα). Αυτό θα επέτρεπε τη χρήση μικρότερης απόστασης και/ή υψηλότερης ανάλυσης χωρίς να ξεπεραστούν τα όρια μνήμης. Αυτή η ιδέα είναι σύμφωνη με τη μεθοδολογία που χρησιμοποιούν οι ακτινολόγοι για την αντιμετώπιση του προβλήματος. Η προσέγγισή τους ξεκινά με την αναγνώριση του εν λόγω οργάνου, ακολουθούμενη από μια σχολαστική εξέταση με στόχο τον εντοπισμό τυχόν μικροσκοπικών όγκων [37].

Διεπιστημονικές μεθοδολογίες U-Net δύο σταδίων για την ανίχνευση όγκων του παγκρέατος έχουν τεκμηριωθεί στη βιβλιογραφία, με ευνοϊκά αποτελέσματα. Ένα αξιοσημείωτο πλεονέκτημα αυτής της προσέγγισης είναι η μείωση των παρεμβολών του φόντου. Στο στάδιο 2, ένα σημαντικό μέρος του σύνθετου φόντου εξαλείφεται, επιτρέποντας στο δίκτυο να επικεντρωθεί στις λεπτές τοπικές διακυμάνσεις εντός του οργάνου.

Η ενσωμάτωση προγνωστικών μοντέλων (σύνολα) είναι ένα κρίσιμο στοιχείο του nnU-Net, το οποίο υπογραμμίζει τη σημασία του συνδυασμού πολλαπλών μοντέλων για τη βελτιστοποίηση της ακρίβειας [33]. Σε αυτή τη μελέτη, εκπαιδεύσαμε ένα ενιαίο μοντέλο 3D U-Net. Για να βελτιώσουμε περαιτέρω την απόδοση του μοντέλου, θα μπορούσαμε να εξετάσουμε το ενδεχόμενο εκπαίδευσης πολλαπλών μοντέλων και συνδυασμού των προβλέψεών τους. Μια προσέγγιση περιλαμβάνει την εφαρμογή ενός συνόλου 5-πλάσιας διασταυρούμενης επικύρωσης, το οποίο συνεπάγεται την εκπαίδευση πέντε μοντέλων (ένα μοντέλο σε κάθε πτυχή). Κατά τη φάση πρόβλεψης, τα αποτελέσματα από αυτά τα μοντέλα υπολογίζονται κατά μέσο όρο ή συγκεντρώνονται μέσω μιας διαδικασίας ψηφοφορίας. Αυτή η προσέγγιση συχνά οδηγεί σε πιο σταθερές και γενικές προβλέψεις, εξαλείφοντας έτσι τα σφάλματα που είναι εγγενή σε ένα μεμονωμένο μοντέλο. Επιπλέον, το nnU-Net έχει τη δυνατότητα να ενσωματώνει ένα συνδυασμό μοντέλων 2D και 3D. Το πλαίσιο έχει τη δυνατότητα να εκπαιδεύσει τόσο το καθαρό 2D U-Net (σε διατομές) όσο και το 3D U-Net. Η ομάδα του nnU-Net αναφέρει συχνά ότι το σύνολο 2D+3D βελτιώνει τα αποτελέσματα σε συγκεκριμένα σύνολα δεδομένων [33]. Η λογική πίσω από αυτήν την προσέγγιση είναι ότι το μοντέλο 2D είναι ικανό να μαθαίνει συγκεκριμένες τοπικές υφές με μεγαλύτερη ακρίβεια, ενώ το μοντέλο 3D καταγράφει με επιδεξιότητα το περιβάλλον 3D. Η ενσωμάτωση αυτών των μοντέλων, όπως προτείνεται, υποτίθεται ότι αποδίδει βέλτιστη απόδοση αξιοποιώντας τα πλεονεκτήματα και των δύο.

Αύξηση δεδομένων: Αν και το nnU-Net είναι γνωστό ότι εφαρμόζει μια ολοκληρωμένη σειρά προεπιλεγμένων αυξήσεων (π.χ. τυχαίες περιστροφές/μετατοπίσεις, κλιμάκωση, οριζόντιες αναστροφές, αλλαγές έντασης γάμμα, ελαστικές παραμορφώσεις κ.λπ. [1]), θα μπορούσαν να ενσωματωθούν εξειδικευμένες τεχνικές δεδομένων για την αντιμετώπιση του συγκεκριμένου προβλήματός μας. Για παράδειγμα, η δημιουργία συνθετικών δειγμάτων μικροσκοπικών όγκων σε ένα υγιές πάγκρεας (σύνθεση δεδομένων) έχει τη δυνατότητα να εμπλουτίσει το σύνολο εκπαίδευσης με μια πιο ποικίλη σειρά παραδειγμάτων. Επιπλέον, η ενσωμάτωση ισχυρότερης τυχαιότητας στην αντίθεση των εικόνων, για παράδειγμα, μέσω της

προσομοίωσης διαφόρων βαθμών ενίσχυσης της αντίθεσης, έχει τη δυνατότητα να ενισχύσει την ανθεκτικότητα του μοντέλου στις διαφορές στην ποιότητα απεικόνισης μεταξύ διαφορετικών εξετάσεων. Η ενσωμάτωση αυτών των προσεγγίσεων στην pipeline του nnU-Net μπορεί να επιτευχθεί μέσω τροποποιήσεων στον γεννήτρια δεδομένων του. Τέλος, αξίζει να εξεταστεί η πιθανή εφαρμογή της ημι-εποπτευόμενης μάθησης. Αυτή η προσέγγιση περιλαμβάνει τη χρήση μη επισημασμένων αξονικών τομογραφιών, οι οποίες στη συνέχεια επισημαίνονται μέσω της δημιουργίας ψευδοεπισημάνσεων από ένα ήδη εκπαιδευμένο μοντέλο. Αυτή η διαδικασία εξηγείται με το παράδειγμα της επισήμανσης περίπου 2.000 μη επισημασμένων εξετάσεων. Στη συνέχεια, το μοντέλο υποβάλλεται σε μια διαδικασία επανεκπαίδευσης που ενσωματώνει τόσο εποπτευόμενα όσο και μη εποπτευόμενα σύνολα.

Πρόσφατες μελέτες έχουν δείξει ότι αυτή η στρατηγική μπορεί να βελτιώσει την απόδοση χωρίς να απαιτείται επιπλέον χειροκίνητη επισήμανση δεδομένων. Οι προαναφερθείσες έννοιες δεν απαιτούν ένα νέο αρχιτεκτονικό σχεδιασμό, αλλά εκτελούνται στο πλαίσιο της παραμετροποίησης και της χρήσης του nnU-Net, αξιοποιώντας την επεκτασιμότητά του.

Οι προαναφερθείσες προτάσεις αποσκοπούν στην αντιμετώπιση των εντοπισμένων περιορισμών, δηλαδή στην ενίσχυση της ευαισθησίας στην ανίχνευση μικροσκοπικών ελαττωμάτων, στην ελαχιστοποίηση των ψευδών αρνητικών και θετικών σφαλμάτων και στην αύξηση της γενικευσιμότητας του μοντέλου. Η ευελιξία του πλαισίου nnU-Net επιτρέπει τέτοια πειράματα, διατηρώντας παράλληλα τα σημαντικά πλεονεκτήματά του όσον αφορά την αυτοματοποιημένη διαμόρφωση και την αξιόπιστη βασική απόδοση.

9. Συμπεράσματα και Μελλοντική Εργασία

9.1. Κύρια συμπεράσματα

Στη μελέτη αυτή, αναπτύχθηκε ένα αυτοματοποιημένο σύστημα βαθιάς μάθησης για την ανίχνευση και τον διαχωρισμό του αδενοκαρκινώματος του παγκρέατος σε εικόνες αξονικής τομογραφίας, χρησιμοποιώντας αποκλειστικά το πλαίσιο nnU-Net. Τα κύρια συμπεράσματα που προέκυψαν από την έρευνα αυτή περιγράφονται παρακάτω:

Η ακόλουθη έρευνα θα εξετάσει τη σκοπιμότητα του αυτόματου διαχωρισμού. Αποδείχθηκε η αποτελεσματικότητα της εκπαίδευσης ενός μοντέλου ικανό να εντοπίζει την πρωτοπαθή μάζα στο πάγκρεας με ελάχιστη ανθρώπινη παρέμβαση. Το τελικό μοντέλο επέδειξε αξιοσημείωτη ακρίβεια, ιδίως λαμβάνοντας υπόψη την πολυπλοκότητα του θέματος. Οι μέσες βαθμολογίες Dice για τον όγκο προσέγγισαν ή ξεπέρασαν τα σημεία αναφοράς που είχαν τεθεί από προηγούμενες μελέτες που χρησιμοποίησαν παρόμοια σύνολα δεδομένων (π.χ., η μέση τιμή DSC όλων των μεθόδων στο MSD Challenge ήταν περίπου 0.21 [1]). Παρά τη μέτρια απόλυτη βαθμολογία Dice (~20-30%), είναι σημαντικό να αναγνωριστεί ότι αυτό οφείλεται στον ελάχιστο αριθμό βλαβών. Στην πράξη, το μοντέλο μας έδειξε αξιοσημείωτη ικανότητα ανίχνευσης σημαντικού ποσοστού όγκων, κάτι που θεωρείται ευρέως επιτυχία για ένα τόσο απαιτητικό κλινικό περιβάλλον.

Η σημασία της διαμόρφωσης του μοντέλου: Μια ολοκληρωμένη έρευνα σχετικά με τις ρυθμίσεις προεπεξεργασίας αποκάλυψε μια ισχυρή συσχέτιση μεταξύ της απόδοσης του συστήματος και της συνετής επιλογής παραμέτρων όπως η απόσταση και η κανονικοποίηση. Συνοπτικά, το αποτέλεσμα εξαρτάται από τον αγωγό. Οι προεπιλεγμένοι κανόνες του nnU-Net, όπως η επαναδειγματοληψία στην κατάλληλη ανάλυση και η σωστή κανονικοποίηση της έντασης, έχουν αποδειχθεί σωστοί και δύσκολο να βελτιωθούν με χειροκίνητες παρεμβάσεις. Αυτό το εύρημα συνάδει με την υποκείμενη φιλοσοφία του nnU-Net, η οποία υποστηρίζει ότι οι βέλτιστες διαμορφώσεις δικτύου υπερτερούν της πολυπλοκότητας και των ad-hoc τροποποιήσεων στο δίκτυο [33]. Κατά συνέπεια, στην περίπτωση μας επιβεβαιώνεται ότι η έμφαση πρέπει να δοθεί στην κατάλληλη προεπεξεργασία και στα δεδομένα, και όχι στην επινόηση νέων αρχιτεκτονικών.

Απόδοση και τρέχουσα τεχνολογία αιχμής: Σε σύγκριση με τις μεθόδους αιχμής, το μοντέλο μας (ένα ενιαίο 3D U-Net χωρίς σύνολο μοντέλων) παρουσίασε συγκρίσιμο επίπεδο απόδοσης. Ωστόσο, δεν έφτασε το επίπεδο απόδοσης που παρουσίασαν οι πιο προηγμένες μέθοδοι, οι οποίες χρησιμοποίησαν πιο εκτεταμένες τεχνικές. Για παράδειγμα, η πιο αποτελεσματική προσέγγιση που καταγράφηκε στο Nature Communications (2022) πέτυχε μέση βαθμολογία Dice περίπου 0,52 για όγκους του παγκρέατος μέσω της εφαρμογής του nnU-Net στην πλήρη φάση ανάπτυξής του, συνοδευόμενη από στρατηγικές συνόλου και βελτιστοποίησης [1]. Τα ευρήματα καταδεικνύουν ότι η απόδοση του μοντέλου, όπως υποδεικνύεται από τις βαθμολογίες Dice που κυμαίνονται από 0,2 έως 0,3 και τις ευαισθησίες σε επίπεδο voxel από 0,5 έως 0,6, τεκμηριώνει την αποτελεσματικότητα του nnU-Net ως θεμελιώδους πλαισίου. Ωστόσο, υπάρχει περιθώριο βελτίωσης για την επίτευξη του ανώτατου ορίου ~0,5 Dice που αναφέρεται στη βιβλιογραφία για ανάλογα προβλήματα. Αξίζει να σημειωθεί ότι κανένα σφάλμα του μοντέλου δεν αποκλίνει από τις προσδοκίες. Οι αποτυχίες του συνάδουν με γνωστές δυσκολίες (μικροί όγκοι, χαμηλό SNR), χωρίς να εμφανίζονται απροσδόκητα σφάλματα.

Η κλινική σημασία αυτών των ευρημάτων είναι η εξής: Ενώ οι αριθμητικοί δείκτες είναι ζωτικής σημασίας για τη μηχανική μάθηση, από ιατρική άποψη, αξίζει να σημειωθεί ότι το σύστημα ανίχνευσε αυτόματα αρκετές παθολογικές αλλοιώσεις που διαφορετικά θα απαιτούσαν επίπονη ανθρώπινη σήμανση. Σε ορισμένες περιπτώσεις, ο αλγόριθμος ήταν σε θέση να ανιχνεύσει μικροσκοπικές αλλοιώσεις στον ιστό του παγκρέατος που διαφορετικά θα ήταν δύσκολο να εντοπιστούν, υποδηλώνοντας ότι τέτοια εργαλεία μπορούν να χρησιμεύσουν ως «δεύτερο ζευγάρι ματιών» για τον ακτινολόγο. Ωστόσο, υπήρξαν και περιπτώσεις πλήρους αποτυχίας στην ανίχνευση βλαβών, γεγονός που υποδηλώνει ότι το σύστημα δεν είναι ακόμη αξιόπιστο για αυτόνομη διάγνωση. Συνοψίζοντας, τα αποτελέσματα της μελέτης δείχνουν ένα βαθμό αισιοδοξίας όσον αφορά τη δυνατότητα ενσωμάτωσης αυτών των μοντέλων στην κλινική πρακτική ως βοηθητικά εργαλεία, υπό την προϋπόθεση ότι θα γίνουν περαιτέρω βελτιώσεις.

Συνοψίζοντας, η παρούσα μελέτη καταδεικνύει το δυναμικό της βαθιάς μάθησης στην ανίχνευση του καρκίνου του παγκρέατος και επισημαίνει τις παραμέτρους που πρέπει να βελτιστοποιηθούν για την επίτευξη υψηλότερης ευαισθησίας και ακρίβειας. Τα κύρια εμπόδια είναι εγγενή στο ίδιο το πρόβλημα (π.χ. ελάχιστο μέγεθος όγκου, χαμηλή αντίθεση),

ωστόσο η σχολαστική μηχανική χαρακτηριστικών του nnU-Net έθεσε μια αξιόπιστη βάση. Έχει επιτευχθεί μια πιο ακριβής κατανόηση των υφιστάμενων κενών, καθώς και των απαραίτητων βημάτων για την αντιμετώπισή τους.

Τα ευρήματα και το μοντέλο που αναπτύχθηκε έχουν τη δυνατότητα να εφαρμοστούν σε διάφορες πραγματικές ιατρικές απεικονιστικές συνθήκες, συμβάλλοντας στη βελτίωση της διάγνωσης και της θεραπείας.

Η παροχή διαγνωστικής βοήθειας και δεύτερης γνώμης αποτελεί αναπόσπαστο μέρος του συστήματος υγειονομικής περίθαλψης. Ένα εγκεκριμένο σύστημα όπως το δικό μας θα μπορούσε να χρησιμεύσει ως βοηθητικό εργαλείο για τους ακτινολόγους κατά την ανάγνωση των αξονικών τομογραφιών της κοιλιακής χώρας. Συγκεκριμένα, το μοντέλο έχει σχεδιαστεί για να αναγνωρίζει αυτόματα ύποπτες περιοχές εντός ή γύρω από το πάγκρεας που μπορεί να αντιστοιχούν σε όγκους. Αυτό είναι ιδιαίτερα ωφέλιμο σε περιπτώσεις μικρών όγκων σε πρώιμο στάδιο που μπορεί να παραβλεφθούν κατά τη διάρκεια μιας επιπόλαιης εξέτασης. Έρευνες έχουν δείξει ότι η εφαρμογή αυτοματοποιημένων σημάνσεων CAD έχει τη δυνατότητα να ενισχύσει την ευαισθησία των ακτινολόγων, χρησιμεύοντας έτσι ως συμπληρωματικό διαγνωστικό εργαλείο. Στο πλαίσιο του καρκίνου του παγκρέατος, μιας νόσου για την οποία η έγκαιρη διάγνωση είναι επιτακτική, δεδομένου του ποσοστού επιβίωσης 5 ετών που ανέρχεται σε 26% στα εντοπισμένα στάδια έναντι 2% στα μεταστατικά στάδια [37], η εφαρμογή ενός τέτοιου διαγνωστικού εργαλείου έχει σημαντικές δυνατότητες. Το εργαλείο αυτό, όταν χρησιμοποιείται, μπορεί να χρησιμεύσει για την πρόληψη της παράβλεψης ύποπτων ευρημάτων, ενισχύοντας έτσι την πιθανότητα έγκαιρης ανίχνευσης.

Η μείωση του φόρτου εργασίας και του χρόνου ανάγνωσης είναι ένα αξιοσημείωτο πλεονέκτημα. Η εφαρμογή της αυτόματης τμηματοποίησης έχει αποδειχθεί ότι οδηγεί σε σημαντική βελτίωση της αποδοτικότητας της ροής εργασίας. Στην τρέχουσα πρακτική, η χειροκίνητη οριοθέτηση του παγκρέατος και του όγκου σε 100-200 τομές ανά εξέταση (τυπικός αριθμός τομών με πάχος ~5 mm) είναι μια χρονοβόρα και επίπονη διαδικασία [37]. Η λειτουργικότητα του εργαλείου περιλαμβάνει τη γρήγορη δημιουργία μιας αρχικής μάσκας και του περιγράμματος του όγκου, μια διαδικασία που ολοκληρώνεται σε λίγα δευτερόλεπτα. Αυτό όχι μόνο μειώνει το φόρτο εργασίας του ακτινολόγου, αλλά και τυποποιεί τη διαδικασία, μειώνοντας έτσι τη μεταβλητότητα μεταξύ διαφορετικών παρατηρητών. Οι ακτινολόγοι θα μπορούν να επικεντρωθούν στην επικύρωση και τη διόρθωση της αυτόματης μάσκας αντί να ξοδεύουν σημαντικό χρόνο και προσπάθεια για τη δημιουργία της από το μηδέν. Κατά συνέπεια, ο αριθμός των εξετάσεων που μπορούν να αξιολογηθούν σε μικρότερο χρονικό διάστημα μπορεί να αυξηθεί, ενισχύοντας έτσι την αποτελεσματικότητα του ακτινολογικού τμήματος.

Η ποσοτικοποίηση των όγκων και ο επακόλουθος σχεδιασμός της θεραπείας τους Η αξιόπιστη τμηματοποίηση διευκολύνει την αυτόματη εξαγωγή ποσοτικών δεδομένων, συμπεριλαμβανομένου του συνολικού όγκου του όγκου (σε cm^3), της διαμέτρου του και της σχέσης/επαφής του με γειτονικές ανατομικές δομές. Στον τομέα της ογκολογίας, αυτά τα δεδομένα είναι ζωτικής σημασίας για τον σχεδιασμό της θεραπείας (π.χ. χειρουργική εκτομή

ή ακτινοθεραπεία) και για την αξιολόγηση της ανταπόκρισης στη θεραπεία. Στο σύγχρονο πλαίσιο, η μέτρηση των όγκων πραγματοποιείται συχνά με πιο αυθαίρετο τρόπο, όπως αποδεικνύεται από τη χρήση μετρήσεων όπως η μέγιστη διάμετρος σε ένα μόνο τμήμα σύμφωνα με τα κριτήρια RECIST. Η εφαρμογή της αυτόματης τμηματοποίησης διευκολύνει την απόκτηση πιο ακριβών τρισδιάστατων μετρήσεων των όγκων και την επακόλουθη παρακολούθηση των μεταβολών με την πάροδο του χρόνου, όπως αυτές που συμβαίνουν σε απόκριση σε χημειοθεραπευτικές αγωγές. Αυτή η αυτοματοποίηση ενισχύει την αντικειμενικότητα της παρακολούθησης της νόσου.

Σχεδιασμός ακτινοθεραπείας και χειρουργικών επεμβάσεων: Σε περιπτώσεις όπου ένας ασθενής χρειάζεται ακτινοθεραπεία, ο όγκος-στόχος μπορεί να καθοριστεί στον αξονικό τομογράφο σχεδιασμού με τη βοήθεια του προτεινόμενου συστήματος. Το μοντέλο nnU-Net έχει τη δυνατότητα να ενσωματωθεί στο σταθμό σχεδιασμού, διευκολύνοντας έτσι την πρόταση μιας περιοχής θεραπείας (CTV) γύρω από τον όγκο του παγκρέατος. Κατά συνέπεια, ο γιατρός θα διαθέτει ένα προκαθορισμένο περίγραμμα που θα μπορεί να τροποποιηθεί σε περιορισμένο βαθμό, εάν είναι απαραίτητο, αντί να ξεκινά από ένα κενό καμβά. Στον τομέα του χειρουργικού σχεδιασμού, για παράδειγμα, στις επεμβάσεις Whipple, η οριοθέτηση της θέσης του όγκου και η επέκτασή του στα αγγεία μπορεί να προσδιοριστεί με μεγαλύτερη σαφήνεια με τρισδιάστατες εικόνες βασισμένες στην τμηματοποίηση. Έχουν γίνει ερευνητικές προσπάθειες για τη χρήση αυτόματων τμηματοποιήσεων στη δημιουργία τρισδιάστατων μοντέλων οργάνων και όγκων. Αυτά τα μοντέλα βοηθούν τους χειρουργούς να κατανοήσουν τα ανατομικά χαρακτηριστικά του ασθενούς.

Ενσωμάτωση σε συστήματα PACS και CADx: Η τεχνική σκοπιμότητα της ενσωμάτωσης του εργαλείου σε ένα νοσοκομειακό σύστημα PACS/RIS είναι μια βιώσιμη πρόταση. Καθώς νέες αξονικές τομογραφίες εισάγονται στο αρχείο, ένας διακομιστής με το εκπαιδευμένο μοντέλο nnU-Net θα είναι σε θέση να δημιουργεί αυτόματα τμηματοποιήσεις και να τις επισυνάπτει στις εικόνες ως δευτερεύουσες καταγραφές ή DICOM RTSTRUCT. Κατά τη διάρκεια της διαδικασίας εξέτασης, ο ακτινολόγος θα είναι σε θέση να διακρίνει τα ανατομικά περιγράμματα του παγκρέατος, εκτός από τυχόν όγκους που υπερτίθενται στις αντίστοιχες τομές. Επιπλέον, στον τομέα της διαγνωστικής με τη βοήθεια υπολογιστή (CADx), τα ευρήματα του μοντέλου μπορούν να ενσωματωθούν σε άλλα σύνολα κλινικών δεδομένων. Για παράδειγμα, η παρουσία ενός ύποπτου όγκου μπορεί να οδηγήσει σε σύσταση για περαιτέρω διερεύνηση (π.χ. μαγνητική τομογραφία, ενδοσκόπηση) μέσω αυτοματοποιημένων ειδοποιήσεων. Κατά συνέπεια, το σύστημα λειτουργεί ως μέτρο ασφαλείας στο πλαίσιο της ροής εργασίας, μειώνοντας έτσι την πιθανότητα να παραβλεφθεί ένας μικρός καρκίνος του παγκρέατος κατά την αρχική αξιολόγηση.

Φυσικά, είναι ανάγκη να υπογραμμιστεί ότι ένα τέτοιο σύστημα θα λειτουργούσε ως βοηθητικός μηχανισμός και όχι ως αυτόνομη οντότητα. Η τελική απόφαση και αξιολόγηση θα εναπόκειται στον θεράποντα ιατρό. Τα ευρήματα που παράγονται αυτόματα πρέπει να επιβεβαιώνονται και, όταν είναι απαραίτητο, να διορθώνονται. Επιπλέον, το μοντέλο πρέπει να υποβληθεί σε αυστηρή επικύρωση σε πολλαπλά κέντρα και να λάβει τις απαραίτητες κανονιστικές εγκρίσεις πριν από τη χρήση του σε κλινικό περιβάλλον. Ωστόσο, μόλις πληρωθούν αυτές οι προϋποθέσεις, οι προαναφερθείσες εφαρμογές είναι απολύτως

ρεαλιστικές και αντιπροσωπεύουν το μέλλον της ακτινολογίας με τη βοήθεια της τεχνητής νοημοσύνης.

9.2. Ιδέες για επόμενες έρευνες

Η έρευνα αυτή μπορεί να επεκταθεί σε διάφορες κατευθύνσεις, αξιοποιώντας τόσο τις δυνατότητες του nnU-Net όσο και τις νεότερες εξελίξεις για την περαιτέρω βελτίωση της απόδοσης και της γενικευσιμότητας.

Η συλλογή ενός πιο εκτεταμένου και πολυκεντρικού συνόλου δεδομένων είναι επιτακτική ανάγκη. Όπως αναφέρεται στην ενότητα των περιορισμών, ένα επόμενο βήμα που είναι προφανές αλλά δεν έχει ακόμη πραγματοποιηθεί είναι η συλλογή πρόσθετων δεδομένων από πολλαπλά κέντρα που χρησιμοποιούν διαφορετικούς σαρωτές και πρωτόκολλα. Η αρχιτεκτονική του nnU-Net μπορεί εύκολα να επανεκπαιδευτεί σε ένα νέο σύνολο ή σε ένα συνδυασμό συνόλων δεδομένων. Υποτίθεται ότι, με ένα επαυξημένο σύνολο εκπαίδευσης, το μοντέλο θα επιδείξει βελτιωμένες δυνατότητες γενίκευσης και θα βελτιώσει την ευαισθησία του. Ταυτόχρονα, η μέθοδος θα αξιολογηθεί ως προς την αποτελεσματικότητά της σε σενάρια που δεν περιλαμβάνονταν στο αρχικό δείγμα, όπως πολύ μικρά όγκους μεγέθους μικρότερου από 1 εκατοστό ή όγκους που παρουσιάζουν άτυπη ενίσχυση.

Η εφαρμογή της ημι-εποπτευόμενης μάθησης και της μεταφοράς μάθησης: Εκτός από την συμβατική ενσωμάτωση δεδομένων, ένας πολλά υποσχόμενος τομέας είναι η χρήση μη επισημασμένων εξετάσεων. Όπως αναφέρθηκε προηγουμένως, η χρήση τεχνικών ψευδο-επισήμανσης έχει τη δυνατότητα να αυξήσει το αποτελεσματικό μέγεθος των δεδομένων χωρίς να επιβαρύνει με επιπλέον έξοδα επισήμανσης [20]. Επιπλέον, θα μπορούσε να αξιολογηθεί η μεταφορά μάθησης. Το nnU-Net επιτρέπει στην έκδοση v2 να προ-εκπαιδεύσει ένα μοντέλο σε ένα σημαντικό γενικό σύνολο δεδομένων (π.χ. όλα τα κοιλιακά όργανα) και στη συνέχεια να το τελειοποιήσει σε έναν στόχο (π.χ. όγκος του παγκρέατος) [33]. Αυτή η προσέγγιση έχει τη δυνατότητα να βελτιώσει την απόδοση, ιδιαίτερα στο πλαίσιο μικρών συνόλων δεδομένων, αξιοποιώντας τις γνώσεις από σχετικές εργασίες τμηματοποίησης.

Η ενσωμάτωση πολυφασικών και πολυτροπικών δεδομένων αποτελεί κρίσιμο στοιχείο της ανάλυσης. Η παρούσα μελέτη επικεντρώθηκε αποκλειστικά σε μία μόνο φάση CT, πιθανότατα την πυλαία φλεβική φάση. Ωστόσο, στην κλινική πρακτική, συχνά χρησιμοποιούνται πολλαπλές φάσεις ενδοφλέβιας σκιαγραφικής (αρτηριακή, φλεβική) και/ή άλλες απεικονιστικές μέθοδοι (π.χ. CT, MRI, PET) για την ανάδειξη των όγκων του παγκρέατος. Μια πιθανή επέκταση του παρόντος πλαισίου θα ήταν η τροποποίηση της αρχιτεκτονικής του nnU-Net ώστε να μπορεί να φιλοξενήσει πολυκαναλικές εισόδους, όπως δύο κανάλια για αρτηριακές και φλεβικές αξονικές τομογραφίες του ίδιου ασθενούς ή αξονικές τομογραφίες και PET. Πρέπει να σημειωθεί ότι το nnU-Net έχει σχεδιαστεί ώστε να υποστηρίζει ήδη πολλαπλά κανάλια εισόδου. Ουσιαστικά, το μοντέλο θα εκπαιδεύεται με εικόνες που περιέχουν πρόσθετες πληροφορίες σε επίπεδο voxel. Υποτίθεται ότι η χρήση ενός συνδυασμού φάσεων θα ενισχύσει την ευαισθησία, καθώς πολλοί μικροί όγκοι είναι πιο

ορατοί στην αρτηριακή φάση λόγω ταχύτερης εμπλουτισμού. Επιπλέον, η PET παρέχει λειτουργικές πληροφορίες (συσσώρευση 18-FDG) που θα μπορούσαν να διευκολύνουν τη διάκριση μεταξύ όγκων και φλεγμονής. Είναι σημαντικό να σημειωθεί ότι όλες αυτές οι επεκτάσεις μπορούν να επιτευχθούν διατηρώντας την αρχιτεκτονική U-Net και απλώς επεκτείνοντας τα δεδομένα εισόδου. Αυτή η επέκταση είναι κάτι που υποστηρίζει άμεσα το nnU-Net.

Η εξερεύνηση νέων εκδόσεων και βελτιώσεων του nnU-Net είναι δικαιολογημένη, καθώς πρόκειται για ένα εργαλείο που εξελίσσεται δυναμικά. Μια πρόσφατη μελέτη, υπογραμμίζει την ανάγκη για σχολαστική αξιολόγηση και προτείνει βελτιώσεις στη μεθοδολογία που χρησιμοποιείται για την επικύρωση μοντέλων [33]. Επιπλέον, έχει κυκλοφορήσει το nnU-Net v2, το οποίο εισάγει αρκετές τροποποιήσεις και βελτιστοποιήσεις, συμπεριλαμβανομένων, μεταξύ άλλων, καλύτερης διαχείρισης μνήμης, υποστήριξης για νέες αυξήσεις και πιο αποτελεσματικής υλοποίησης. Στο μέλλον, θα ήταν σκόπιμο να μεταφέρουμε τον αγωγό μας σε αυτή τη νεότερη έκδοση και να αξιολογήσουμε εάν έχει καλύτερη απόδοση. Ταυτόχρονα, διερευνώνται μεθοδολογίες για την ενσωμάτωση μονάδων προσοχής ή υβριδικών αρχιτεκτονικών (π.χ. U-Net με μπλοκ Transformer) στο πλαίσιο του nnU-Net [33]. Αυτή η υπόθεση υποστηρίζεται από τις πρόσφατες δοκιμές μοντέλων όπως το MedSAM (Segment Anything Model για ιατρικές εικόνες), τα οποία έχουν επιδείξει καινοτόμες προσεγγίσεις στην τμηματοποίηση και ανταγωνιστικά αποτελέσματα, όπως αναφέρεται στο πάγκρεας [33]. Ωστόσο, δεδομένης της αποδεδειγμένης αποτελεσματικότητας του nnU-Net στην υπεροχή ή την προσέγγιση της απόδοσης πιο σύνθετων μοντέλων, μια πολλά υποσχόμενη προσέγγιση περιλαμβάνει την ενσωμάτωση των βέλτιστων πρακτικών αυτών των νέων μεθόδων στην υπάρχουσα, αποτελεσματική pipeline μας. Για να το επεξηγήσουμε, θα μπορούσαμε να διεξάγουμε ένα πείραμα στο οποίο θα εγκαθιστούσαμε plug-in attention modules στον κορμό του U-Net, διατηρώντας όλες τις άλλες μεταβλητές σταθερές. Αυτό θα επέτρεπε μια δίκαιη σύγκριση.

Η εφαρμογή της μεθόδου σε συναφή προβλήματα: Τέλος, μια προφανής επέκταση θα ήταν η εφαρμογή του προαναφερθέντος πλαισίου σε άλλες συναφείς κλινικές εφαρμογές. Το nnU-Net έχει ήδη επιδείξει βέλτιστη απόδοση σε μια σειρά από σύνολα δεδομένων και υπάρχει ενδιαφέρον για εφαρμογές πέραν του παγκρέατος. Για παράδειγμα, το μοντέλο έχει τη δυνατότητα να εκπαιδευτεί ώστε να ανιχνεύει αυτόματα κύστες του παγκρέατος ή μεταστατικούς λεμφαδένες στην κοιλιακή χώρα [33]. Η βασική φιλοσοφία είναι ανάλογη: το δίκτυο τροφοδοτείται με σχετικά δεδομένα και αυτόματες μάσκες (π.χ. λεμφαδένες >1 cm), ενώ το nnU-Net προσαρμόζει τη ροή του. Μια άλλη επέκταση αφορά τη συνδυασμένη πρόβλεψη πολλαπλών οντοτήτων. Για παράδειγμα, μια πολυκατηγοριακή τμηματοποίηση θα επέτρεπε στο nnU-Net να μάθει ταυτόχρονα να τμηματοποιεί το πάγκρεας, τον όγκο και άλλα όργανα (ή αγγεία) που το περιβάλλουν. Ένα μοντέλο πολλαπλής τμηματοποίησης μπορεί να είναι επωφελές για την ολοκληρωμένη ανατομική χαρτογράφηση, όπως αποδεικνύεται από την εφαρμογή του στην ογκολογική χειρουργική.

Συνοψίζοντας, προκύπτουν πολλές δυνατότητες για μελλοντική έρευνα. Το κοινό στοιχείο μεταξύ όλων αυτών είναι ότι η έρευνα διεξάγεται εντός του οικοσυστήματος του nnU-Net, αξιοποιώντας την ευελιξία και τον αυτοματοποιημένο χαρακτήρα του. Σε αυτή τη μελέτη, το

nnU-Net χρησίμευσε ως αξιόπιστο σημείο εκκίνησης (έτοιμο προς χρήση, τελευταίας τεχνολογίας) και, με τις προτάσεις που αναφέρθηκαν, μπορεί να επεκταθεί περαιτέρω. Η εφαρμογή μελλοντικών βημάτων αναμένεται να αποφέρει όχι μόνο βελτίωση των αριθμητικών δεικτών, αλλά, πιο σημαντικά, αύξηση της κλινικής χρησιμότητας του τελικού συστήματος, φέρνοντάς το έτσι πιο κοντά στην εφαρμογή στον πραγματικό κόσμο και ενισχύοντας τα οφέλη του για τους ασθενείς.

10. Βιβλιογραφία

- [1] Antonelli, M., Reinke, A., Bakas, S., Farahani, K., Kopp-Schneider, A., Landman, B. A., Litjens, G., Menze, B., Ronneberger, O., Summers, R. M., Van Ginneken, B., Bilello, M., Bilic, P., Christ, P. F., G, R. K., DO, Gollub, M. J., Heckers, S. H., Huisman, H., Jarnagin, W. R., . . . Cardoso, M. J. (2022). The Medical Segmentation Decathlon. *Nature Communications*, 13(1). <https://doi.org/10.1038/s41467-022-30695-9>
- [2] Ardila, D., Kiraly, A. P., Bharadwaj, S., Choi, B., Reicher, J. J., Peng, L., Tse, D., Etemadi, M., Ye, W., Corrado, G., Naidich, D. P., & Shetty, S. (2019). End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. *Nature Medicine*, 25(6), 954–961. <https://doi.org/10.1038/s41591-019-0447-x>
- [3] Brown, M. P. H. (2024, August 14). *Abdominal CT: attenuation*. Life in the Fast Lane • LITFL. <https://litfl.com/abdominal-ct-attenuation/>
- [4] Cherezov, D., Viswanathan, V. S., Fu, P., Gupta, A., & Madabhushi, A. (2023). Rank acquisition impact on radiomics estimation (AcquIRE) in chest CT imaging: A retrospective multi-site, multi-use-case study. *Computer Methods and Programs in Biomedicine*, 244, 107990. <https://doi.org/10.1016/j.cmpb.2023.107990>
- [5] Chin, M. (2016). Pixels from digital photos to medical images. www.marychin.org. <https://www.marychin.org/pixvox.html>
- [6] Cireşan, D. C., Giusti, A., Gambardella, L. M., & Schmidhuber, J. (2013). Mitosis Detection in Breast Cancer Histology Images with Deep Neural Networks. *Lecture Notes in Computer Science*, 411–418. https://doi.org/10.1007/978-3-642-40763-5_51
- [7] *Deep residual learning for image recognition*. (2016, June 1). IEEE Conference Publication | IEEE Xplore. <https://ieeexplore.ieee.org/document/7780459>
- [8] *Densely connected convolutional networks*. (2017, July 1). IEEE Conference Publication | IEEE Xplore. <https://ieeexplore.ieee.org/document/8099726>

[9] Dubey, S. R., Singh, S. K., & Chaudhuri, B. B. (2021). Activation Functions in Deep Learning: A comprehensive survey and benchmark. *arXiv (Cornell University)*.

<https://doi.org/10.48550/arxiv.2109.14545>

[10] Dubey, S. R., Singh, S. K., & Chaudhuri, B. B. (2022). Activation functions in deep learning: A comprehensive survey and benchmark. *Neurocomputing*, 503, 92–108.

<https://doi.org/10.1016/j.neucom.2022.06.111>

[11] Engelson, S., Ehrhardt, J., Kepp, T., Niemeijer, J., & Handels, H. (2024). LNQ Challenge 2023: Learning Mediastinal Lymph Node Segmentation with a Probabilistic Lymph Node Atlas. *The Journal of Machine Learning for Biomedical Imaging*, 2(MICCAI 2023 LNQ challenge), 817–833. <https://doi.org/10.59275/j.melba.2024-f95c>

[12] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639), 115–118. <https://doi.org/10.1038/nature21056>

[13] *File:Image resampling.svg - Wikimedia Commons*. (n.d.). https://commons.wikimedia.org/wiki/File:Image_resampling.svg

[14] *Focal loss for dense object detection*. (2017, October 1). IEEE Conference Publication | IEEE Xplore. <https://ieeexplore.ieee.org/document/8237586>

[15] *Gradient-based learning applied to document recognition*. (1998, November 1). IEEE Journals & Magazine | IEEE Xplore. <https://ieeexplore.ieee.org/document/726791>

[16] Gupta, P. (2021, December 15). nnU-Net : The no-new-UNet for automatic segmentation. *Medium*. <https://medium.com/miccai-educational-initiative/nnu-net-the-no-new-unet-for-automatic-segmentation-8d655f3f6d2a>

[17] He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep residual learning for image recognition. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.1512.03385>

- [18] Hesamian, M. H., Jia, W., He, X., & Kennedy, P. (2019). Deep Learning Techniques for Medical Image Segmentation: Achievements and challenges. *Journal of Digital Imaging*, 32(4), 582–596. <https://doi.org/10.1007/s10278-019-00227-x>
- [19] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2016). Densely connected convolutional networks. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.1608.06993>
- [20] Huang, Z., Wang, H., Ye, J., Niu, J., Tu, C., Yang, Y., Du, S., Deng, Z., Gu, L., & He, J. (2022). Revisiting NNU-Net for iterative pseudo labeling and efficient sliding window inference. In *Lecture notes in computer science* (pp. 178–189). https://doi.org/10.1007/978-3-031-23911-3_16
- [21] Isensee, F., Jaeger, P.F., Kohl, S.A.A., et al. (2021) nnU-Net A Self-Configuring Method for Deep Learning-Based Biomedical Image Segmentation. *Nature Methods*, 18, 203-211. - References - Scientific Research Publishing. (n.d.). <https://www.scirp.org/reference/referencespapers?referenceid=3700031>
- [22] Kalra, M. K., Maher, M. M., Toth, T. L., Schmidt, B., Westerman, B. L., Morgan, H. T., & Saini, S. (2004). Techniques and applications of Automatic tube current modulation for CT. *Radiology*, 233(3), 649–657. <https://doi.org/10.1148/radiol.2333031150>
- [23] Karpińska, M., & Czauderna, M. (2022). Pancreas—Its functions, disorders, and physiological impact on the mammals' organism. *Frontiers in Physiology*, 13. <https://doi.org/10.3389/fphys.2022.807632>
- [24] Kingma, D. P., & Ba, J. (2014, December 22). *Adam: A method for stochastic optimization*. arXiv.org. <https://arxiv.org/abs/1412.6980>
- [25] LeCun, Y. A., Bottou, L., Orr, G. B., & Müller, K. (2012). Efficient BackProp. In *Lecture notes in computer science* (pp. 9–48). https://doi.org/10.1007/978-3-642-35289-8_3

[26] Li, T., Lin, C., & Wang, W. (2025). Global, regional, and national burden of pancreatic cancer from 1990 to 2021, its attributable risk factors, and projections to 2050: a systematic analysis of the global burden of disease study 2021. *BMC Cancer*, 25(1).

<https://doi.org/10.1186/s12885-025-13597-z>

[27] Lin, T., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal loss for dense object detection. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.1708.02002>

[28] Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. a. A., Ciompi, F., Ghafoorian, M., Van Der Laak, J. A., Van Ginneken, B., & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42, 60–88.

<https://doi.org/10.1016/j.media.2017.07.005>

[29] Liu, H., Hu, D., Li, H., & Oguz, I. (2023). Medical image segmentation using deep learning. In *Neuromethods* (pp. 391–434). https://doi.org/10.1007/978-1-0716-3195-9_13

[30] Longnecker, D. (2021). Anatomy and histology of the pancreas [Dataset]. In *Pancreapedia: The Exocrine Pancreas Knowledge Base*.

<https://doi.org/10.3998/panc.2021.01>

[31] Mahmoudi, T., Kouzahkanan, Z. M., Radmard, A. R., Kafieh, R., Salehnia, A., Davarpanah, A. H., Arabalibeik, H., & Ahmadian, A. (2022). Segmentation of pancreatic ductal adenocarcinoma (PDAC) and surrounding vessels in CT images using deep convolutional neural networks and texture descriptors. *Scientific Reports*, 12(1).

<https://doi.org/10.1038/s41598-022-07111-9>

[32] McCormick, M. (n.d.). *DICOM Rescale Intercept / Rescale Slope and ITK*. <https://www.kitware.com/dicom-rescale-intercept-rescale-slope-and-itk/>

- [33] Moglia, A., Cavicchioli, M., Mainardi, L., & Cerveri, P. (2025). Deep learning for pancreas segmentation on computed tomography: a systematic review. *Artificial Intelligence Review*, 58(8). <https://doi.org/10.1007/s10462-024-11050-4>
- [34] Moraes, T., Amorim, P., Da Silva, J. V., & Pedrini, H. (2019). Medical image interpolation based on 3D Lanczos filtering. *Computer Methods in Biomechanics and Biomedical Engineering Imaging & Visualization*, 8(3), 294–300. <https://doi.org/10.1080/21681163.2019.1683469>
- [35] Oktay, O., Schlemper, J., Folgoc, L. L., Lee, M. C. H., Heinrich, M. P., Misawa, K., Mori, K., McDonagh, S. G., Hammerla, N. Y., Kainz, B., Glocker, B., & Rueckert, D. (2018). Attention U-Net: Learning where to look for the pancreas. *arXiv (Cornell University)*. <https://doi.org/10.48550/arxiv.1804.03999>
- [36] Prokesch, R. W., Schima, W., Chow, L. C., & Jeffrey, R. B. (2003). Multidetector CT of pancreatic adenocarcinoma: diagnostic advances and therapeutic relevance. *European Radiology*, 13(9), 2147–2154. <https://doi.org/10.1007/s00330-003-1926-4>
- [37] Qiu, D., Ju, J., Ren, S., Zhang, T., Tu, H., Tan, X., & Xie, F. (2024). A deep learning-based cascade algorithm for pancreatic tumor segmentation. *Frontiers in Oncology*, 14. <https://doi.org/10.3389/fonc.2024.1328146>
- [38] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-NET: Convolutional Networks for Biomedical Image Segmentation. In *Lecture notes in computer science* (pp. 234–241). https://doi.org/10.1007/978-3-319-24574-4_28
- [39] Roth, H. R., Farag, A., Lu, L., Turkbey, E. B., & Summers, R. M. (2015). Deep convolutional networks for pancreas segmentation in CT imaging. *Proceedings of SPIE, the International Society for Optical Engineering/Proceedings of SPIE*. <https://doi.org/10.1117/12.2081420>

- [40] Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1986). Learning representations by back-propagating errors. *Nature*, 323(6088), 533–536.
<https://doi.org/10.1038/323533a0>
- [41] Sanchez, L. E., Rundo, L., Gill, A. B., Hoare, M., Serrao, E. M., & Sala, E. (2021). Robustness of radiomic features in CT images with different slice thickness, comparing liver tumour and muscle. *Scientific Reports*, 11(1).
<https://doi.org/10.1038/s41598-021-87598-w>
- [42] Schmidhuber, J. (2014). Deep learning in neural networks: An overview. *Neural Networks*, 61, 85–117. <https://doi.org/10.1016/j.neunet.2014.09.003>
- [43] Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F. (2021). Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA a Cancer Journal for Clinicians*, 71(3), 209–249. <https://doi.org/10.3322/caac.21660>
- [44] Talathi, S. S., Zimmerman, R., & Young, M. (2023, April 5). *Anatomy, abdomen and pelvis, pancreas*. StatPearls - NCBI Bookshelf.
<https://www.ncbi.nlm.nih.gov/books/NBK532912/>
- [45] Tempero, M. A., Malafa, M. P., Al-Hawary, M., Behrman, S. W., Benson, A. B., Cardin, D. B., Chiorean, E. G., Chung, V., Czito, B., Del Chiaro, M., Dillhoff, M., Donahue, T. R., Dotan, E., Ferrone, C. R., Fountzilas, C., Hardacre, J., Hawkins, W. G., Klute, K., Ko, A. H., . . . George, G. V. (2021). Pancreatic adenocarcinoma, Version 2.2021, NCCN Clinical Practice Guidelines in Oncology. *Journal of the National Comprehensive Cancer Network*, 19(4), 439–457. <https://doi.org/10.6004/jnccn.2021.0017>
- [46] Themes, U. (2017, February 12). *Computed tomography*. Radiology Key.
<https://radiologykey.com/computed-tomography-12/>

[47] Trojani, V., Bassi, M. C., Verzellesi, L., & Bertolini, M. (2024). Impact of preprocessing Parameters in Medical Imaging-Based Radiomic Studies: A Systematic review. *Cancers*, 16(15), 2668. <https://doi.org/10.3390/cancers16152668>

[48] Zwanenburg, A., Vallières, M., Abdalah, M. A., Aerts, H. J. W. L., Andrearczyk, V., Apte, A., Ashrafinia, S., Bakas, S., Beukinga, R. J., Boellaard, R., Bogowicz, M., Boldrini, L., Buvat, I., Cook, G. J. R., Davatzikos, C., Depeursinge, A., Desseroit, M., Dinapoli, N., Dinh, C. V., . . . Löck, S. (2020). The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. *Radiology*, 295(2), 328–338. <https://doi.org/10.1148/radiol.2020191145>