

ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΚΑΙ ΗΛΕΚΤΡΟΝΙΚΩΝ ΣΥΣΤΗΜΑΤΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

«Επεξηγηματική Μηχανική Μάθηση για Ανίχνευση
καρκινικών όγκων μέσω εικόνων αξονικού
τομογράφου»



**Του φοιτητή
Τσεκερίδης Θεοφάνης
Αρ. Μητρώου: 185449**

**Επιβλέπων
Μπράτσας Χαράλαμπος
Βαθμίδα**

Ημερομηνία 31/5/2026

Τίτλος Δ.Ε. Επεξηγηματική Μηχανική Μάθηση για Ανίχνευση καρκινικών όγκων μέσω εικόνων

αξονικού τομογράφου

Κωδικός Δ.Ε. 23203

Όνοματεπώνυμο φοιτητή/των

Τσεκερίδης Θεοφάνης

Όνοματεπώνυμο εισηγητή Μπράτσας Χαράλαμπος

Ημερομηνία ανάληψης Δ.Ε. 02-04-2025

Ημερομηνία περάτωσης Δ.Ε. 31/5/2026

Βεβαιώνω ότι είμαι ο συγγραφέας αυτής της εργασίας και ότι κάθε βοήθεια την οποία είχα για την προετοιμασία της είναι πλήρως αναγνωρισμένη και αναφέρεται στην εργασία. Επίσης, έχω καταγράψει τις όποιες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών, εικόνων και κειμένου, είτε αυτές αναφέρονται ακριβώς είτε παραφρασμένες. Επιπλέον, βεβαιώνω ότι αυτή η εργασία προετοιμάστηκε από εμένα προσωπικά, ειδικά ως διπλωματική εργασία, στο Τμήμα Μηχανικών Πληροφορικής και Ηλεκτρονικών Συστημάτων του ΔΙ.ΠΑ.Ε.

Η παρούσα εργασία αποτελεί πνευματική ιδιοκτησία του φοιτητή Τσεκερίδη Θεοφάνη που την εκπόνησε. Στο πλαίσιο της πολιτικής ανοικτής πρόσβασης, ο συγγραφέας/δημιουργός εκχωρεί στο Διεθνές Πανεπιστήμιο της Ελλάδος άδεια χρήσης του δικαιώματος αναπαραγωγής, δανεισμού, παρουσίασης στο κοινό και ψηφιακής διάχυσης της εργασίας διεθνώς, σε ηλεκτρονική μορφή και σε οποιοδήποτε μέσο, για διδακτικούς και ερευνητικούς σκοπούς, άνευ ανταλλάγματος. Η ανοικτή πρόσβαση στο πλήρες κείμενο της εργασίας, δεν σημαίνει καθ' οιονδήποτε τρόπο παραχώρηση δικαιωμάτων διανοητικής ιδιοκτησίας του συγγραφέα/δημιουργού, ούτε επιτρέπει την αναπαραγωγή, αναδημοσίευση, αντιγραφή, πώληση, εμπορική χρήση, διανομή, έκδοση, μεταφόρτωση (downloading), ανάρτηση (uploading), μετάφραση, τροποποίηση με οποιονδήποτε τρόπο, τμηματικά ή περιληπτικά της εργασίας, χωρίς τη ρητή προηγούμενη έγγραφη συναίνεση του συγγραφέα/δημιουργού.

Η έγκριση της διπλωματικής εργασίας από το Τμήμα Μηχανικών Πληροφορικής και Ηλεκτρονικών Συστημάτων του Διεθνούς Πανεπιστημίου της Ελλάδος, δεν υποδηλώνει απαραίτητως και αποδοχή των απόψεων του συγγραφέα, εκ μέρους του Τμήματος.

«Αφιέρωση»

Πρόλογος

Μέσα από τις σπουδές μου σε αυτή τη σχολή κατάφερα να κάνω μια εισαγωγή στον κόσμο της μηχανικής μάθησης. Οι προοπτικές αυτού του κλάδου κέντρισαν το ενδιαφέρον μου από την αρχή, με αποτέλεσμα να επιλέξω ως θέμα διπλωματικής εργασίας την εφαρμογή τεχνικών μηχανικής μάθησης σε ιατρικά δεδομένα. Συνδυάζοντας το ενδιαφέρον μου για την ιατρική και την τεχνολογία, αξιοποίησα αξονικές τομογραφίες, αλλά και μαγνητικές, και σχεδίασα ένα 3D Convolutional Neural Network για δυαδική ταξινόμηση (καρκίνος / μη καρκίνος). Μέσα από τη διαδικασία βελτίωσα δεξιότητες στην προεπεξεργασία εικόνας, στην κατασκευή και αξιολόγηση μοντέλων και στην ανάλυση αποτελεσμάτων· ταυτόχρονα αντιλήφθηκα την αξία της διεπιστημονικής συνεργασίας. Η εργασία αυτή μου προσέφερε ουσιαστική εμπειρία έρευνας και πρακτική επάρκεια, ενώ με ενθάρρυνε να συνεχίσω την ενασχόλησή μου με εφαρμογές της τεχνητής νοημοσύνης στην υγεία.

Περίληψη

Στόχος της εργασίας ήταν η αξιοποίηση της επεξηγηματικής τεχνητής νοημοσύνης σε αξονικές τομογραφίες, ενώ προέκυψε και η αξιοποίηση μαγνητικών, για την αυτόματη διάκριση καρκίνου/μη-καρκίνου. Καθ' όλη τη διάρκεια της έρευνας κατέστησαν σαφές οι περιορισμοί: περιορισμένος όγκος διαθέσιμων δεδομένων και περιορισμένη υπολογιστική ισχύ για έναν φοιτητή. Για την αντιμετώπιση αυτών των περιορισμών εφαρμόστηκαν τεχνικές αύξησης δεδομένων (data augmentation) και προσεκτική προεπεξεργασία, ώστε να πολλαπλασιαστεί το σύνολο εκπαίδευσης. Το βασικό μοντέλο είναι ένα 3D Convolutional Neural Network το οποίο εκπαιδεύτηκε επαναληπτικά για τρεις διαφορετικούς τύπους ιστών: κάθε φορά το ίδιο δίκτυο εκπαιδεύθηκε σε ζεύγη παθολογικής και υγιούς κατάστασης του αντίστοιχου ιστού. Παράλληλα εφαρμόστηκαν μέθοδοι επεξηγηματικότητας (explainable AI) για να αναλυθούν τα χαρακτηριστικά και οι περιοχές των αξονικών που οδήγησαν στην απόφαση του μοντέλου. Το αποτέλεσμα ήταν ότι το δίκτυο πέτυχε μια σχετικά ικανοποιητική δυαδική ταξινόμηση παρά τους περιορισμούς στα δεδομένα, ενώ οι τεχνικές επεξηγηματικότητας ανέδειξαν επαναλαμβανόμενα χαρακτηριστικά/περιοχές που σχετίζονται με την απόφαση. Η μεθοδολογία αποδεικνύει τη δυνατότητα χρήσης 3D CNN σε ιατρικές εικόνες και προσφέρει πρακτικές ενδείξεις για το ποιες εικόνες και χαρακτηριστικά είναι πιο κρίσιμα. Ωστόσο συστήνονται μελλοντικές βελτιώσεις με μεγαλύτερα σύνολα δεδομένων και ισχυρότερη υπολογιστική υποδομή για περαιτέρω αξιολόγηση, κλινική επικύρωση και ανάδειξη της ποιότητας των αποτελεσμάτων που προσφέρει ο αναπτυγμένος κώδικας.

«Explainable Machine Learning for Cancer Tumor Detection Using CT Scan Images»

«Tsekeridis Theofanis»

Abstract

The aim of this thesis was to utilize explainable artificial intelligence on CT scans for the automatic classification of cancerous and non-cancerous cases. Throughout the research, the limitations became evident: a limited volume of available data and restricted computational resources for a student. To overcome these challenges, data augmentation techniques and careful preprocessing were applied to expand the training dataset. The core model is a 3D Convolutional Neural Network, which was trained iteratively on three different tissue types: in each case, the same network was trained on pairs of pathological and healthy tissue samples. In parallel, explainable AI methods were employed to analyze the features and regions of the scans that influenced the model's decisions. Results showed that the network achieved an almost satisfactory binary classification despite the limited data, while the explainability techniques highlighted recurring features and regions associated with the decisions. This methodology demonstrates the feasibility of using 3D CNNs on medical images with constrained resources and provides practical insights into which images and features are most critical. However, future improvements are recommended, including larger datasets and more powerful computational infrastructure for further evaluation, clinical validation, and to highlight the quality of the results produced by the code that has been developed.

Ευχαριστίες

Θα ήθελα να ευχαριστήσω την οικογένεια μου για την συμπαράσταση οποιουδήποτε είδους κατά την διάρκεια των σπουδών μου καθώς και όλους τους καθηγητές που συντέλεσαν στον να μου προσφέρουν τις απαραίτητες γνώσεις για να διεκπεραιώσω τη συγκεκριμένη διπλωματική εργασία μου αλλά και γενικότερα στο να είμαι έτοιμος για τις μελλοντικές προκλήσεις που χαρακτηρίζουν μια επιστήμη όπως την πληροφορική . Τέλος θα ήθελα να ευχαριστήσω την ομάδα του Cancer Imaging Archive η οποία κατέστησε εφικτή την συγκεκριμένη διπλωματική μέσω την παροχής δεδομένων .

Περιεχόμενα

Πρόλογος.....	vi
Περίληψη	vii
Abstract	viii
Ευχαριστίες	ix
Περιεχόμενα	x
Κατάλογος εικόνων	xii
Κατάλογος Σχημάτων	xiii
Συνομογραφίες.....	xiv
Κεφάλαιο 1ο: Εισαγωγή	15
1.1 Εισαγωγή	15
1.2 Στόχος Διπλωματικής	17
1.3 Δομή Διπλωματικής.....	17
Κεφάλαιο 2ο: Ιστορική αναδρομή στις 2 επιστήμες.....	19
2.1 Πως προέκυψε η μηχανική μάθηση.....	19
2.2 Μηχανική μάθηση	19
2.3 Neural networks	20
2.4 Οι πρώτες ακτίνες X.....	21
2.5 Σύγκλιση των δύο επιστημών	21
Κεφάλαιο 3ο: Θεωρητικές βάσεις	22
3.1 Τύποι προβλημάτων	22
3.2.1 Deep Learning	23
3.2.2 Artificial neural networks	24
3.2.3 Convolutional neural networks	28
3.2.3.1 Max-pooling layer	30
3.2.3.2 Flattening layer	31
3.3 Ιατρική.....	31
3.3.1 Ηλεκτρομαγνητικό φάσμα	32
3.3.2 Λυχνία ακτίνων X.....	32
3.3.3 Μηχανισμοί παραγωγής ακτίνων X.....	33
3.3.4 Αλληλεπίδραση ακτινοβολίας με το σώμα	33
3.3.5 Εξαγωγή μαθηματικών σχέσεων	34
3.3.6 Απόδοση και ανακατασκευή εικόνας	36

3.3.7 Υπολογιστική τομογραφία (CT).....	36
3.3.8 Μαγνητικές (MRI).....	37
Κεφάλαιο 4ο: Ανάλυση του πειράματος.....	39
4.1 Σύνολο δεδομένων.....	39
4.2 Επιλογή αρχείων	42
4.3 Προεπεξεργασία των CT-scans.....	43
4.4 Προεπεξεργασία των MRI-scans	49
4.5 Έμφαση στην διαφορά ανάμεσα σε CT-scans και MRI-scans.....	51
4.5 Δημιουργία του CNN	52
4.6 Επεξηγηματική τεχνητή νοημοσύνη	54
4.7 Εκτελέσιμος κώδικας	60
Κεφάλαιο 5ο: Αποτελέσματα του πειράματος	66
5.1 Αποτελέσματα επεξηγηματικής τεχνητής νοημοσύνης	66
5.2 Συνολικά αποτελέσματα	71
5.3 Σύνοψη	72
Βιβλιογραφία.....	74

Κατάλογος εικόνων

Εικόνα 1.1: Αριθμός αξονικών στις ΗΠΑ.....	16
Εικόνα 2.1: Το χέρι της συζύγου του Wilhelm Conrad Röntgen με την χρήση ακτίνων X.....	20
Εικόνα 4.1: PKG – BraTS-Africa MRI-Scan.....	40
Εικόνα 4.2: Spine-Mets CT-Scan.....	40
Εικόνα 4.3: LIDC-IDRI CT-Scan.....	41
Εικόνα 4.4: Healthy-Total-Body CT-Scan.....	41
Εικόνα 4.5: Lung_RIDER CT-Scan.....	41
Εικόνα 4.6: CFB-GBM MRI-Scan.....	41
Εικόνα 4.7: Vestibular-Schwannoma MRI-Scan.....	42
Εικόνα 4.8: Βασικές βιβλιοθήκες.....	43
Εικόνα 4.9: file system στο οποίο ανταποκρίνεται ο κώδικας.....	44
Εικόνα 4.10: Παράμετροι συνάρτησης διαχωρισμού μιας αξονικής σε υποσύνολα.....	44
Εικόνα 4.11: Δομή δεδομένων που περιέχει ποσοστά ομοιότητας ενός slice με δεδομένο όργανο.....	45
Εικόνα 4.12: file system που δημιουργεί η συνάρτηση διαχωρισμού.....	46
Εικόνα 4.13: Κώδικας μετατροπής τιμών φωτεινότητας των pixels σε HU.....	47
Εικόνα 4.14: Παράδειγμα robust clipping.....	50
Εικόνα 4.15: Κώδικας υπολογισμού των παραγώγων.....	55
Εικόνα 4.16: Κώδικας κατασκευής του heat-map.....	56
Εικόνα 4.17: Κώδικας upsampling του heat-map.....	59
Εικόνα 4.18: Κώδικας upsampling του heat-map.....	59
Εικόνα 4.19: Κώδικας χρωματισμού ενός 3d-volume με το αντίστοιχο heat-map.....	60
Εικόνα 4.20: Κλήση της συνάρτησης διαχωρισμού.....	61
Εικόνα 4.21: Κλήση συναρτήσεων φόρτωσης και προεπεξεργασίας των δεδομένων.....	62
Εικόνα 4.22: Ορίζοντας τον διαχωρισμό των δεδομένων.....	63
Εικόνα 4.23: Κλήση συνάρτησης εκπαίδευσης.....	63
Εικόνα 4.24: Δημιουργία ενός functional CNN μοντέλου.....	64
Εικόνα 4.25: Εκτέλεση της Grad-Cam λειτουργικότητας.....	64
Εικόνα 5.1: Εκπαίδευση του brain-CNN για fold 3.....	66
Εικόνα 5.2: slice_10&30_overlay_patient_0256_Lung_cancer_label_1.....	67
Εικόνα 5.3: XAI για patient 43 slices 10 & 22 με bile ducts και spine καρκίνο.....	68
Εικόνα 5.4: output για patient 43 fold 2.....	68
Εικόνα 5.5: αρίθμηση των σπόνδυλων.....	69
Εικόνα 5.6: XAI για patient 43 , slice 15 (T10) με spine καρκίνο.....	69
Εικόνα 5.7: Ο καρκινικός όγκος του patient_105.....	70
Εικόνα 5.8: XAI για τον patient_105_PKG_-_BraTS-Africa για fold_3 , slices = 3,11.....	70
Εικόνα 5.9: XAI για τον patient_105_PKG_-_BraTS-Africa για fold_3 , slices = 25,31.....	70
Εικόνα 5.10: Προβλέψεις του brain CNN.....	72
Εικόνα 5.10: Προβλέψεις του spine CNN.....	72

Κατάλογος Σχημάτων

Σχήμα 3.1: Francis Galton problem	21
Σχήμα 3.2: Βιολογικός νευρώνας	23
Σχήμα 3.3: Σύναψη μεταξύ βιολογικών νευρώνων	24
Σχήμα 3.4: Σύγκριση βιολογικού και τεχνητού νευρώνα	25
Σχήμα 3.5: Perceptron.....	26
Σχήμα 3.6: Συσχέτιση βιολογικής διαδικασίας μεταφοράς ηλεκτρικού σήματος με αυτή που συμβαίνει στον τεχνητό νευρώνα	27
Σχήμα 3.7: Διαισθητικό κομμάτι της συνέλιξης σε ένα 2d-Convolutional Layer	28
Σχήμα 3.8: Διαισθητικό κομμάτι της συνέλιξης σε ένα 3d-Convolutional Layer	29
Σχήμα 3.9: Ηλεκτρομαγνητικό φάσμα.....	32
Σχήμα 4.1: Οπτικοποίηση της επίδρασης του κώδικα πάνω στην αξονική	44
Σχήμα 4.2: Οπτικοποίηση της επαναδειγματοληψίας μιας αξονικής	48
Σχήμα 4.3: Οπτικοποίηση της αποθήκευσης των 3d volumes	48
Σχήμα 4.4: Ροή της επεξεργασίας των δεδομένων μου	52
Σχήμα 4.5: Δομή του CNN για την επεξηγηματική νοημοσύνη	54
Σχήμα 4.6: Πως υπολογίζεται η σημαντικότητα ενός φίλτρου	57
Σχήμα 4.7: Παράδειγμα υπολογισμού του χρώματος για ένα συγκεκριμένο σημείο στο heatmap που συντίθεται	58

Κατάλογος Πινάκων

Πίνακας 5.1: Learning curves των τριών CNN	71
Πίνακας 5.2: Αποτελέσματα των τριών CNN	72

Συντομογραφίες

Δ.Ε.	Διπλωματική Εργασία
ΔΙΠΙΑΕ	Διεθνές Πανεπιστήμιο Ελλάδος
CNN	Convolutional Neural Networks
ΜΜ	Μηχανική Μάθηση
ΧΑΙ	Explainable Artificial Intelligence
CT	Computed Tomography
MRI	Magnetic Resonance Imaging

Κεφάλαιο 1ο: Εισαγωγή

1.1 Εισαγωγή

Οι αξονικές τομογραφίες αποτελούν ένα ισχυρό μέσο εξαγωγής πληροφορίας από το εσωτερικό του ανθρώπινου σώματος: αποτυπώνουν δομές και αλλαγές που αντανακλούν τη φυσιολογία, την παθολογία ή την ανταπόκριση σε θεραπευτικές παρεμβάσεις. Αυτό σημαίνει ότι, μέσω της συστηματικής συλλογής και επεξεργασίας αξονικών εικόνων, μπορούν να εξαχθούν δεδομένα που συσχετίζονται με συγκεκριμένες κλινικές καταστάσεις ή χρονικά διαστήματα θεραπείας. Η συνεχής πρόοδος τόσο στις ιατρικές επιστήμες όσο και στους αλγορίθμους μηχανικής μάθησης έχει βελτιώσει σημαντικά την ποιότητα της πληροφορίας που εξάγεται, καθώς και την ικανότητα ανάλυσης και αξιοποίησής της [12]. Παράλληλα, η ευρεία χρήση απεικονιστικών εξετάσεων στη σύγχρονη κλινική πρακτική έχει οδηγήσει στη δημιουργία τεράστιων όγκων ιατρικών δεδομένων. Η εικόνα 1.1 που ακολουθεί αναδεικνύει τον όγκο δεδομένων που δημιουργήθηκε μόνο κατά το έτος 2023 στις Η.Π.Α. όπου θα μπορούσαν να αξιοποιηθούν για εφαρμογές νευρωνικών δικτύων στον ιατρικό τομέα, υπογραμμίζοντας έτσι ότι η σύγχρονη ιατρική απεικόνιση δεν χαρακτηρίζεται πλέον από έλλειψη δεδομένων αλλά, αντίθετα, από αφθονία πληροφορίας υψηλής διαστατικότητας. Η διαθεσιμότητα τόσο μεγάλων συνόλων δεδομένων δημιουργεί ιδιαίτερα ευνοϊκές συνθήκες για την ανάπτυξη συστημάτων βαθιάς μάθησης, καθώς τα μοντέλα μπορούν να εκπαιδευτούν σε ολοένα και πιο σύνθετα και αντιπροσωπευτικά παραδείγματα. Για τον λόγο αυτό, ο συνδυασμός ιατρικής απεικόνισης και μηχανικής μάθησης θεωρείται ένας από τους πλέον υποσχόμενους άξονες εξέλιξης της σύγχρονης ιατρικής.

Table 2. Projected Number of Future Cancers Overall and by Cancer Type Associated With CT Examinations Performed in the United States in 2023, by Age Group and Body Region^a

Cancer type	Projected future cancers, No. (90% UL)		
	All CT examinations (N = 93 000 000)	CT examinations in adults (n = 89 931 000)	CT examinations in children (n = 3 069 000)
Total	102 700 (96 400-109 500)	93 000 (86 900-99 600)	9700 (8100-11 600)
Projected cancer by type			
Lung	22 400 (20 200-25 000)	21 400 (19 200-24 000)	990 (870-1100)
Colon	8700 (7800-9700)	8400 (7500-9400)	330 (280-390)
Leukemia	7900 (6700-9500)	7400 (6100-8900)	550 (380-820)
Bladder	7100 (6000-8500)	6900 (5700-8200)	250 (200-320)
Stomach	7100 (5500-9100)	6800 (5200-8800)	280 (200-400)
Thyroid	7000 (5400-9200)	3500 (2700-4600)	3500 (2300-5500)
Breast	5700 (5000-6500)	5100 (4400-5900)	630 (550-730)
Liver	4100 (3400-5000)	4000 (3200-4900)	160 (130-200)
Kidney	3000 (2300-3900)	2900 (2200-3700)	130 (90-180)
Pancreas	2800 (2300-3500)	2700 (2200-3400)	100 (80-140)
Oral cavity or pharynx	2800 (2300-3400)	2300 (1900-2900)	450 (310-650)
Brain or CNS	1600 (1300-2000)	1200 (910-1500)	440 (320-620)
Esophagus	1500 (1300-1800)	1400 (1200-1700)	110 (90-150)
Prostate	1500 (820-2700)	1400 (760-2700)	70 (30-170)
Ovary	890 (670-1200)	850 (630-1100)	40 (30-70)
Rectum	560 (480-660)	540 (450-630)	30 (20-40)
Uterus	550 (400-760)	530 (380-730)	30 (16-50)
Other and ill-defined sites	17 400 (15 300-19 800)	15 800 (13 700-18 200)	1600 (1200-2000)
Projected cancer by CT examination body region			
Abdomen and pelvis	39 100 (34 600-44 200)	37 500 (32 900-42 600)	1600 (1300-2000)
Chest	22 700 (19 600-26 300)	21 500 (18 400-25 200)	1200 (960-1400)
Spine	12 900 (11 500-14 500)	11 600 (10 200-13 200)	1300 (1000-1600)
Head	12 500 (10 600-14 700)	7300 (6200-8700)	5100 (3700-7100)
Full body	8000 (7000-9100)	7600 (6600-8800)	320 (260-390)
Head and neck combined	4100 (3500-4800)	4100 (3500-4800)	NA
Cardiac	3400 (3200-3700)	3300 (3000-3600)	170 (140-210)
Extremity	80 (60-90)	70 (50-80)	9 (7-11)

Abbreviations: CNS, central nervous system; CT, computed tomography; NA, not applicable, meaning this category does not exist in this age group; UL, uncertainty limit.

^a More granular results by sex and cross-tabulation by body region and cancer type appear in eTable 3 in Supplement 1.

Εικόνα 1.1: Αριθμός αξονικών που συμβαίνουν στις Η.Π.Α

Ειδικότερα, τα Convolutional Neural Networks προσφέρουν ένα πλαίσιο για την αυτόματη ανίχνευση μοτίβων και χαρακτηριστικών μέσα σε τρισδιάστατα ιατρικά δεδομένα, επιτρέποντας την εξαγωγή συμπερασμάτων που μπορούν να απαντήσουν σε προχωρημένα ιατρικά ερωτήματα. Ωστόσο, η εφαρμογή τους στον ιατρικό τομέα απαιτεί προσεκτική προεπεξεργασία, διαχείριση περιορισμένων συνόλων δεδομένων και μεθόδους εξηγησιμότητας για την ερμηνεία των αποφάσεων[12]. Σκοπός της παρούσας εργασίας είναι να διερευνήσει τη χρήση 3D CNN σε αξονικές τομογραφίες και μαγνητικές τομογραφίες για δυαδική ταξινόμηση (καρκίνος/μη-καρκίνος) και να εφαρμόσει τεχνικές επεξηγηματικής τεχνητής νοημοσύνης ώστε να αναδειχθούν τα χαρακτηριστικά που οδηγούν στις προβλέψεις.

1.2 Στόχος Διπλωματικής

Βασικός στόχος της παρούσας διπλωματικής εργασίας ήταν η αξιοποίηση τεχνικών βαθιάς μάθησης για την ανάλυση ιατρικών δεδομένων και ειδικότερα αξονικών τομογραφιών, με σκοπό τη δυαδική ταξινόμηση ασθενών σε κατηγορίες καρκίνου και μη καρκίνου. Ωστόσο, η εργασία δεν περιορίστηκε μόνο στην παραγωγή μιας τελικής πρόβλεψης από το νευρωνικό δίκτυο. Σημαντικό μέρος της έρευνας αποτέλεσε η διερεύνηση της αξιοπιστίας και της ερμηνευσιμότητας των αποφάσεων του μοντέλου. Για τον λόγο αυτό εφαρμόστηκαν τεχνικές επεξηγηματικής τεχνητής νοημοσύνης – XAI, με στόχο την κατανόηση των χαρακτηριστικών και των μοτίβων που οδήγησαν το CNN στην τελική του απόφαση. Μέσα από αυτή τη διαδικασία επιχειρήθηκε όχι μόνο η αξιολόγηση της απόδοσης του μοντέλου, αλλά και η ερμηνεία του τρόπου με τον οποίο το δίκτυο αντιλαμβάνεται τις παθολογικές και υγιείς δομές μέσα στις ιατρικές εικόνες.

Στην παρούσα εργασία επιπλέον εξετάζεται διερευνητικά κατά πόσο ένας εμπειρικός στατιστικός κανόνας, όπως το $n \approx 30$, μπορεί να παρουσιαστεί με μια ενδεικτική αναλογία ως σημείο αναφοράς (lower bound) σε περιβάλλοντα βαθιάς μάθησης με περιορισμένα ιατρικά δεδομένα. Στη στατιστική ανάλυση, το μέγεθος δείγματος $n \geq 30$ χρησιμοποιείται συχνά ως πρακτικό όριο για την εφαρμογή προσεγγιστικών μεθόδων και την εξαγωγή συμπερασμάτων υπό την προϋπόθεση αντιπροσωπευτικότητας του δείγματος, επιτρέποντας τη γενίκευση προς έναν ευρύτερο πληθυσμό[42].

Αντίστοιχα, στο πλαίσιο της παρούσας προσέγγισης, το μέγεθος των 30 υγιών δειγμάτων εξετάζεται ως ενδεικτική συνθήκη χαμηλού ορίου δεδομένων, με στόχο την αξιολόγηση της επάρκειάς του για την παραγωγή σταθερών και αποδεκτών αποτελεσμάτων σε μοντέλα τρισδιάστατων συνελκτικών νευρωνικών δικτύων (3D-CNN) σε ιατρικές απεικονίσεις. Η σύγκριση αυτή δεν προϋποθέτει άμεση ισοδυναμία μεταξύ στατιστικής επάρκειας και απόδοσης μοντέλων μηχανικής μάθησης, αλλά λειτουργεί ως διερευνητικό πλαίσιο αξιολόγησης υπό συνθήκες περιορισμένης διαθεσιμότητας δεδομένων. Αυτή η προσέγγιση αποσκοπεί στο να ενισχύσει την αξιοπιστία των αποτελεσμάτων και αναδεικνύει τη σημασία της διαφάνειας στα συστήματα τεχνητής νοημοσύνης, ιδιαίτερα σε εφαρμογές που σχετίζονται με τον ιατρικό τομέα και τη λήψη κρίσιμων διαγνωστικών αποφάσεων.

1.3 Δομή Διπλωματικής

Η παρούσα διπλωματική εργασία οργανώνεται σε πέντε βασικά κεφάλαια, τα οποία παρουσιάζουν σταδιακά τόσο το θεωρητικό υπόβαθρο όσο και την πρακτική υλοποίηση της έρευνας. Στο πρώτο κεφάλαιο παρουσιάζεται το είδος των δεδομένων που αξιοποιούνται στην εργασία, δηλαδή οι αξονικές τομογραφίες, καθώς και η σημασία τους στον ιατρικό τομέα. Παράλληλα αναλύεται ο τρόπος με τον οποίο τα δεδομένα αυτά προσεγγίζονται και επεξεργάζονται στο πλαίσιο της παρούσας μελέτης. Στο δεύτερο κεφάλαιο γίνεται αναφορά στις βασικές αρχές και τις ρίζες των δύο επιστημονικών πεδίων που συνδυάζονται στην εργασία, δηλαδή της ιατρικής φυσικής και της πληροφορικής, με στόχο να αναδειχθεί η σύνδεση μεταξύ τους και η εξέλιξη των εφαρμογών τεχνητής νοημοσύνης στην ιατρική απεικόνιση. Στο τρίτο κεφάλαιο επιχειρείται μια εκτενέστερη θεωρητική προσέγγιση των δύο επιστημών, ώστε ο αναγνώστης να αποκτήσει το απαραίτητο υπόβαθρο για την καλύτερη κατανόηση της μεθοδολογίας και των τεχνικών που χρησιμοποιούνται στη συνέχεια της εργασίας. Στο τέταρτο κεφάλαιο παρουσιάζονται αναλυτικά τα σύνολα δεδομένων (datasets) που χρησιμοποιήθηκαν, η διαδικασία προεπεξεργασίας τους, καθώς και η ανάπτυξη του κώδικα και των μοντέλων 3D CNN που εφαρμόστηκαν. Τέλος, στο πέμπτο κεφάλαιο παρατίθενται τα αποτελέσματα των πειραμάτων, η

Κεφάλαιο 1

αξιολόγηση της απόδοσης των μοντέλων και η ερμηνεία μεμονωμένων αποτελεσμάτων μέσω τεχνικών επεξηγηματικής τεχνητής νοημοσύνης, αλλά και συνολικά.

Κεφάλαιο 2ο: Ιστορική αναδρομή στις 2 επιστήμες

2.1 Πως προέκυψε η μηχανική μάθηση

Το 1943 οι Warren S. McCulloch και Walter Pitts δημοσίευσαν το άρθρο «A Logical Calculus of Ideas Immanent in Nervous Activity» που θεωρείται η θεμελιώδης απόπειρα να περιγραφεί μαθηματικά η νευρωνική λειτουργία. Στην ουσία πρότειναν ένα αφηρημένο πρότυπο «νευρώνα» που δέχεται εισόδους, τις συνδυάζει και ενεργοποιείται αν το άθροισμα ξεπεράσει ένα κατώφλι — δηλαδή ένα δυαδικό, threshold unit. Με αυτό το απλό δομικό στοιχείο έδειξαν ότι δίκτυα τέτοιων μονάδων μπορούν να αναπαραστήσουν λογικές συναρτήσεις (AND, OR, NOT) και γενικά υπολογιστικές διαδικασίες[1]. Η ιδέα δεν προέκυψε από έναν μόνο τομέα: ο McCulloch έφερε στο τραπέζι γνώση από τη νευροφυσιολογία και τη φιλοσοφία της νόησης, ενώ ο Pitts — με μαθηματικό/λογικό υπόβαθρο — συνέβαλε με την μαθηματική και λογική θεμελίωση της ιδέας. Έτσι, η συνεργασία τους συνέδεσε νευροβιολογικά δεδομένα (πώς συμπεριφέρονται οι νευρώνες) με λογικά/αλγεβρικά εργαλεία και κατέστησε για πρώτη φορά σαφές ότι νευρικά δίκτυα μπορούν, θεωρητικά, να υλοποιήσουν υπολογισμούς — μια βάση πάνω στην οποία στηρίχτηκε, δεκαετίες αργότερα, όλο το πεδίο της τεχνητής νοημοσύνης και της μηχανικής μάθησης.

2.2 Μηχανική μάθηση

Η μηχανική μάθηση (MM) γεννήθηκε από την επιθυμία να κατανοηθεί και να αναπαραχθεί με υπολογιστικά μέσα την ικανότητα του ανθρώπου να μαθαίνει, να γενικεύει και να παίρνει αποφάσεις. Ιδέες που προέρχονται από τη νευροβιολογία —όπως η έννοια του νευρώνα και της σύναψης— αποτέλεσαν το θεωρητικό υπόβαθρο: οι πρώιμες μαθηματικές προσεγγίσεις προσπάθησαν να απλοποιήσουν τη βιολογική πολυπλοκότητα σε μοντέλα που ήταν δυνατόν να υπολογιστούν[9]. Η βασική υπόθεση είναι ότι η πολύπλοκη νοητική λειτουργία προκύπτει από το άθροισμα πολλών απλών στοιχείων και από την ποικιλία των συνδέσεων μεταξύ τους: διαφορετικές συναπτικές δομές παράγουν διαφορετικά «μονοπάτια» ενεργοποίησης, και το πλήθος και η ποιότητα αυτών των μονοπατιών καθορίζουν τον τρόπο που ένα σύστημα «σκέφτεται».

Από θεωρητική σκοπιά, τα πρώτα μοντέλα —όπως οι λογικοί νευρώνες των McCulloch και Pitts και ο Perceptron του Rosenblatt— έδειξαν ότι με απλούς κόμβους και ζυγισμένες συνδέσεις μπορούσαν να επιλυθούν βασικά προβλήματα ταξινόμησης[1],[2]. Η κρίσιμη εξέλιξη ήρθε με την έννοια της μάθησης από εμπειρία: μονάδες που προσαρμόζουν τις συνδέσεις τους βάσει σφαλμάτων ή εμπειρικών σχέσεων (ιδέες κοντά στην Hebbian plasticity) μπορούν, με την κατάλληλη δομή, να αποκτήσουν εκπληκτικά περίπλοκη συμπεριφορά[10]. Η μέθοδος της οπισθοδιαφορικής διάδοσης σφάλματος (backpropagation) κατέστησε πρακτική την εκπαίδευση πολυστρωματικών δικτύων, επιτρέποντας την ανάδυση ενδογενών αναπαραστάσεων από τα δεδομένα[11], δηλαδή το ίδιο το δίκτυο ανακαλύπτει μόνο του χρήσιμα χαρακτηριστικά ή μοτίβα από τα δεδομένα εισόδου.

Σταδιακά, η προσέγγιση «άθροιση πολλών τρόπων σκέψης» μετεξελίχθηκε σε βαθιά δίκτυα όπου κάθε στρώμα εξειδικεύεται σε διαφορετικά χαρακτηριστικά[6] — σαν ένα πολύεδρο δίκτυο συνάψεων που, όταν σχηματίζονται, οδηγούν σε μοναδικά μονοπάτια ενεργοποίησης για κάθε είσοδο. Η σύγχρονη MM συνδυάζει αυτή τη νευροβιολογική έμπνευση με μαθηματικά εργαλεία, μεγάλες ποσότητες δεδομένων και υπολογιστική ισχύ· έτσι, αντί για άμεση αντιγραφή του εγκεφάλου, δημιουργείται ένα «λογικό

άθροισμα» διαφορετικών υποσυστημάτων που, συνεργαζόμενα, πετυχαίνουν σύνθετες γνωστικές λειτουργίες.

Η αξία αυτής της προσέγγισης αποδεικνύεται στην πράξη: από την αναγνώριση εικόνας και τη φωνητική κατανόηση ως την ανίχνευση όγκων στην ιατρική απεικόνιση — όπου η ικανότητα ενός συστήματος να συνδυάζει πολλαπλά μοτίβα και να βελτιώνει τις συνδέσεις του μέσω μάθησης οδηγεί σε κλινικά χρήσιμα αποτελέσματα[12]. Έτσι, η ιστορία της μηχανικής μάθησης είναι ταυτόχρονα ιστορία βιολογικής έμπνευσης, μαθηματικής διατύπωσης και πρακτικής εφαρμογής — μια προσπάθεια αναπαραγωγής, με μαθηματικά μέσα, της ποικιλίας και του δυναμισμού των ανθρώπινων νευρωνικών μονοπατιών.

2.3 Neural networks

Τα τεχνητά νευρωνικά δίκτυα δεν αποτελούν σύγχρονη εφεύρεση· οι βασικές τους ιδέες εμφανίζονται ήδη από τα πρώτα βήματα της τεχνητής νοημοσύνης, ως απόπειρα μίμησης της λειτουργίας του ανθρώπινου εγκεφάλου μέσω απλοποιημένων υπολογιστικών μοντέλων [1], [2],[10]. Παρά τη θεωρητική τους δυναμική, τα πρώιμα αυτά συστήματα δεν κατόρθωσαν να ευδοκιμήσουν, κυρίως λόγω της έλλειψης επαρκούς υπολογιστικής ισχύος και των περιορισμένων δυνατοτήτων αποθήκευσης δεδομένων. Οι διαθέσιμοι πόροι δεν επέτρεπαν την εκπαίδευση μεγάλων και πολυστρωματικών δικτύων, γεγονός που οδήγησε σε περίοδο μειωμένου ενδιαφέροντος για τις νευρωνικές προσεγγίσεις [3].

Κατά τις πρώτες δεκαετίες της μηχανικής μάθησης, η συλλογή και αποθήκευση δεδομένων αποτελούσε σημαντικό εμπόδιο. Τα δεδομένα αποθηκεύονταν σε ογκώδη φυσικά μέσα, με υψηλό κόστος και περιορισμένη χωρητικότητα, γεγονός που περιόριζε τόσο την ποσότητα όσο και την ποικιλία των διαθέσιμων συνόλων εκπαίδευσης. Η εξέλιξη της τεχνολογίας αποθήκευσης υπήρξε καθοριστική για την πρόοδο του πεδίου, επιτρέποντας σταδιακά τη μετάβαση από σπάνια και μικρά σύνολα δεδομένων σε μεγάλης κλίμακας ψηφιακά αποθετήρια [5],[6].

Η επιστροφή των νευρωνικών δικτύων συνδέεται άμεσα με τη ραγδαία αύξηση της υπολογιστικής ισχύος, η οποία μπορεί να ερμηνευθεί και μέσα από τον νόμο του Moore[5], σύμφωνα με τον οποίο η υπολογιστική δυνατότητα των επεξεργαστών αυξάνεται εκθετικά με την πάροδο του χρόνου . Η συνεχής αυτή αύξηση κατέστησε εφικτή την εκπαίδευση βαθύτερων μοντέλων, τη χρήση πολύπλοκων αλγορίθμων μάθησης και την αξιοποίηση μεγάλων όγκων δεδομένων, οδηγώντας στη σύγχρονη αναγέννηση των νευρωνικών δικτύων και του deep learning [4],[6].

Στο ίδιο πλαίσιο τεχνολογικής εξέλιξης, η έννοια της αποθήκευσης δεδομένων έχει επεκταθεί πέρα από τα παραδοσιακά ψηφιακά μέσα. Ερευνητικές προσεγγίσεις έχουν δείξει ότι ακόμη και το DNA μπορεί να χρησιμοποιηθεί ως μέσο αποθήκευσης ψηφιακής πληροφορίας[7],[8], αξιοποιώντας τη φυσική του πυκνότητα και ανθεκτικότητα στον χρόνο . Αν και τέτοιες τεχνολογίες βρίσκονται σε πειραματικό στάδιο, αναδεικνύουν το πόσο δραστικά έχουν διευρυνθεί τα όρια της αποθήκευσης δεδομένων, στοιχείο που αποτελεί θεμελιώδη παράγοντα για τη συνεχιζόμενη ανάπτυξη των νευρωνικών δικτύων και της μηχανικής μάθησης γενικότερα[6].

2.4 Οι πρώτες ακτίνες X

Στα τέλη του 1895, ο Wilhelm Conrad Röntgen, πειραματιζόμενος με ακτινοβολίες καθοδικών σωλήνων, παρατήρησε ένα ανεξήγητο φαινόμενο — μια λάμψη σε μια φωσφορίζουσα οθόνη ακόμη και όταν ο σωλήνας ήταν καλυμμένος. Η περιέργεια τον οδήγησε σε συστηματικές δοκιμές και σύντομα ανακάλυψε ότι μπορούσε να παράγει εικόνες του εσωτερικού αντικειμένων, χωρίς να τα ανοίγει. Μία από τις πρώτες εφαρμογές των ακτίνων X ήταν η απεικόνιση του χεριού της συζύγου του Wilhelm Röntgen, Anna Bertha. Όπως φαίνεται και παρακάτω στην Εικόνα 2.1 διακρίνονται τα οστά του χεριού καθώς και το δαχτυλίδι της. Η συγκεκριμένη απεικόνιση ανέδειξε τις δυνατότητες των ακτίνων X στον τομέα της ιατρικής διάγνωσης και της μελέτης του ανθρώπινου σώματος. Ο Röntgen δεν είχε «προκαθορισμένο» σκοπό να φέρει επανάσταση στην ιατρική: η ανακάλυψη προέκυψε από πειραματισμό και παρατήρηση, αλλά η πρώτη εφαρμογή πάνω στην Anna Bertha έγινε σύμβολο: η εικόνα της χειρός της έμεινε ως η πρώτη ανθρώπινη ακτινογραφία — ένας απλός, ανθρώπινος δείκτης του πόσο γρήγορα μια φυσική ανακάλυψη μπορεί να μετατραπεί σε πρακτική τεχνολογία.[41]



Εικόνα 2.1: Το χέρι της συζύγου του Wilhelm Conrad Röntgen με την χρήση ακτίνων X

2.5 Σύγκλιση των δύο επιστημών

Από εκεί προκύπτει το εξής κρίσιμο σημείο: οι εικόνες ιατρικής απεικόνισης — CT, MRI, ακτινογραφίες — δεν είναι πλέον απλώς "εικόνες" αλλά ψηφιακά δεδομένα που αποθηκεύονται, αναζητώνται και αναλύονται ηλεκτρονικά (PACS, DICOM standards κ.ά.). Η μεγάλη, συνεχής ροή τέτοιων ψηφιακών εικόνων δίνει στην πληροφορική την πρώτη ύλη: εκατομμύρια slice-εικόνες, metadata, δόσεις ακτινοβολίας και κλινικά σχόλια που μπορούν να οργανωθούν και να προεπεξεργαστούν. Η σύγκλιση συμβαίνει επειδή η φυσική παρήγαγε μεθόδους (μη επεμβατικές απεικονίσεις) που δημιουργούν πολύπλοκα, ποσοτικοποιήσιμα data [41], και η πληροφορική — με εργαλεία αποθήκευσης, ανάλυσης και υπολογιστικής ισχύος — είναι πλέον ικανή να τα αξιοποιήσει σε μεγάλη κλίμακα.

Κεφάλαιο 3ο: Θεωρητικές βάσεις

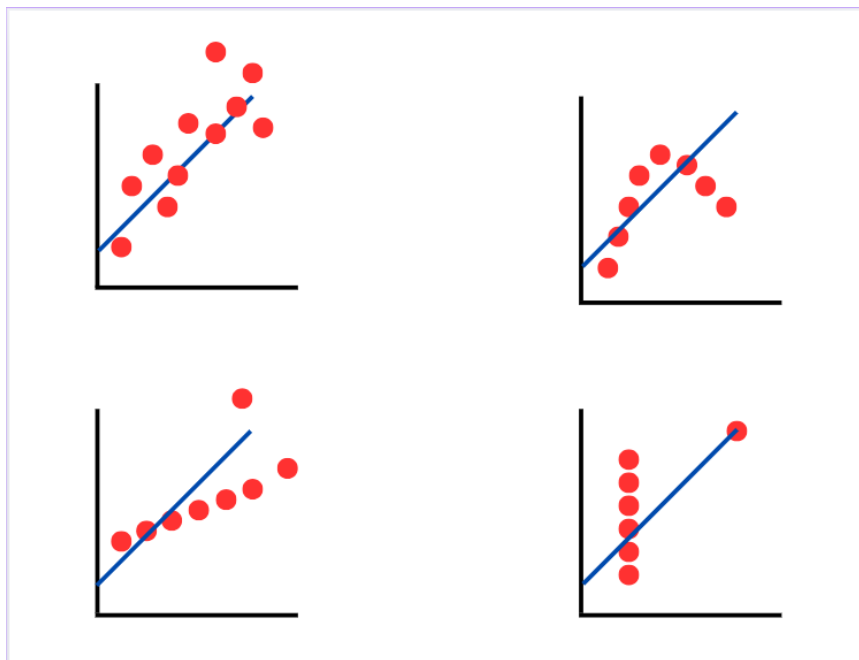
3.1 Τύποι προβλημάτων

3.1.1 Regression

Η regression (παλινδρόμηση) αφορά την εκτίμηση μιας συνεχούς ποσοτικής μεταβλητής ως συνάρτηση ενός ή περισσότερων χαρακτηριστικών. Στη μηχανική μάθηση αυτό σημαίνει ότι διαθέτουμε ένα σετ εκπαίδευσης όπου κάθε δείγμα συνοδεύεται από την πραγματική τιμή-στόχο (label) και ο σκοπός του αλγορίθμου είναι να μάθει μια απεικόνιση $f: X \rightarrow R$ που ελαχιστοποιεί κάποιο μέτρο σφάλματος (π.χ. MSE) [6],[19]. Η φύση του προβλήματος υπαγορεύει προτιμήσεις σε μοντέλα όπως γραμμικά μοντέλα τα οποία περιγράφονται από μαθηματικές σχέσεις όπως η σχέση 3.1 :

$$y = b_0 + b_1 * x_1 + b_2 * x_2 + \dots + b_n * x_n \quad (3.1)$$

και είναι κατάλληλα τα μοντέλα αυτά όταν η σχέση των δεδομένων είναι περίπου γραμμική. Παρακάτω στην εικόνα 3.1 παρουσιάζονται τέσσερις περιπτώσεις συνόλων δεδομένων που αναδεικνύουν την σημασία της αίσθησης των δεδομένων μέσω της οπτικοποίησης τους, με απώτερο σκοπό να διαπιστωθεί ότι στο σύνολο δεδομένων υφίσταται ένας ικανοποιητικός βαθμός γραμμικής συσχέτισης. Πέρα από την πρώτη περίπτωση η χρήση γραμμικού μοντέλου παλινδρόμησης στις υπόλοιπες τρεις περιπτώσεις είναι τουλάχιστον αποπροσανατολιστική .



Σχήμα 3.1: Francis Galton problem

Στην εικόνα 3.1 η οπτικοποίηση των δεδομένων συμβαίνει αναίμακτα λόγω του χαμηλού αριθμού διαστάσεων που χρειάζεται για να αποδοθούν σχηματικά αυτά. Όταν όμως τα δεδομένα χαρακτηρίζονται από ένα αριθμό διαστάσεων που καθιστούν μη εφικτή την οπτικοποίησή τους σε αυτό τον χώρο, τότε μπορούν να εφαρμοστούν μέθοδοι όπως Pearson correlation, R^2 , p-values, VIF, και PCA για να διακρίνουν την γραμμικότητα στον πολυδιάστατο χώρο. Ενώ μη-παραμετρικά μοντέλα όπως decision trees, kernel methods και νευρωνικά δίκτυα επιλέγονται για σύνθετες μη-γραμμικές συσχετίσεις [6]. Στην περίπτωση των ιατρικών εικόνων 3D (αξονικές τομογραφίες και MRI scans), τα

δεδομένα αντιπροσωπεύονται ως τρισδιάστατα volumes με εκατομμύρια voxels (π.χ. $512 \times 512 \times N$), δημιουργώντας έναν υψηλών διαστάσεων χώρο που καθιστά αδύνατη την άμεση οπτικοποίηση και ανάλυση. Για την διερεύνηση γραμμικών σχέσεων μεταξύ των εντάσεων voxels ή των εξαγόμενων radiomic χαρακτηριστικών, θα μπορούσαν να εφαρμοστούν οι κλασικές στατιστικές μέθοδοι που αναφέρθηκαν προηγουμένως για την ανίχνευση γραμμικότητας και την μείωση της πολυπλοκότητας στον πολυδιάστατο χώρο. Ωστόσο, λόγω της έμφυτης σύνθετης και μη-γραμμικής φύσης των ιατρικών εικόνων (όρια όγκων, υφές ιστών, χωρικές εξαρτήσεις και λεπτές διαφορές αντίθεσης), επιλέγονται μη-παραμετρικά μοντέλα όπως τα decision trees, kernel methods και κυρίως τα 3D Convolutional Neural Networks, τα οποία μπορούν να μάθουν αυτόματα ιεραρχικές μη-γραμμικές συσχετίσεις χωρίς να βασίζονται σε παραδοχές γραμμικότητας.

3.1.2 Classification

Το classification (κατηγοριοποίηση) αποτελεί ένα δεύτερο τύπο προβλήματος όπου ουσιαστικά προβαίνει κανείς σε μεθόδους ανίχνευσης της κατηγορίας στην οποία υπάγεται ένα συγκεκριμένο δεδομένο, ή αλλιώς observation, ενώ έχει προηγηθεί μια κάποια εκπαίδευση σε ένα γνωστό dataset. Πρόκειται για μάθηση με επίβλεψη άρα κάθε δείγμα εκπαίδευσης φέρει μία ή περισσότερες γνωστές κλάσεις για κάθε observation και ο αλγόριθμος μαθαίνει ένα διαχωριστικό (SVM, Perceptron) ή ένα μοντέλο πιθανοτήτων (logistic regression, Naive Bayes) που αποδίδει για νέα δείγματα είτε μια hard ετικέτα είτε μια κατανομή πιθανοτήτων πάνω στις κλάσεις [2],[3],[6]. Στην παρούσα διπλωματική, αυτός ο τύπος προβλήματος είναι που παρουσιάζει άμεσο ενδιαφέρον. Για την τελική απόφαση προτιμήθηκαν τα logits, καθώς όχι μόνο παρέχουν περισσότερη πληροφορία σχετικά με την «βεβαιότητα» της πρόβλεψης για κάθε κλάση, αλλά επίσης παράγουν εντονότερους υπολογισμούς παραγών κατά την εκπαίδευση. Η σημασία αυτής της επιλογής και οι λεπτομέρειες της εφαρμογής θα εξηγηθούν αναλυτικά στην ενότητα της πειραματικής αξιολόγησης.

3.1.3 Clustering

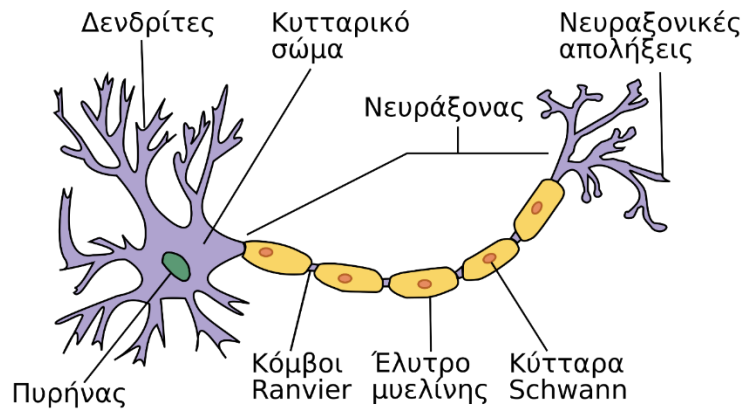
Στο clustering (ομαδοποίηση) ως τελευταίο σημαντικό τύπο προβλήματος δεν διαθέτουμε ετικέτες· ο στόχος είναι να ανακαλυφθούν δομές, ομάδες ή υπο-ομάδες μέσα στα δεδομένα που τις χαρακτηρίζει ένας βαθμός ομοιογένειας ή πυκνότητας στον χώρο [6]. Επομένως πρόκειται για μια πρώτη γνωριμία με αυτό που ονομάζεται μάθηση χωρίς επίβλεψη. Μέχρι τώρα αναφέρθηκαν μοντέλα που υπέστησαν μια εκπαίδευση από έτοιμα δεδομένα, ωστόσο τώρα σε αυτόν τον τύπο προβλήματος το μοντέλο καλείται να αποκαλύψει δομές από μόνο του δίχως να εκπαιδευτεί σε ήδη υπάρχουσες. Υπάρχουν διαφορετικές φιλοσοφίες clustering — centroid-based (π.χ. K-means), hierarchical (agglomerative/divisive), density-based (DBSCAN) και model-based (Gaussian Mixture Models + EM algorithm) — καθεμία με διαφορετικά πλεονεκτήματα και περιορισμούς.

3.2.1 Deep Learning

Η ιδέα του deep learning πηγάζει από την προσπάθεια να μιμηθούμε μαθησιακές ικανότητες του ανθρώπινου εγκεφάλου [9],[10]: το πώς συλλέγει ερεθίσματα, πώς ενισχύει ή αποδυναμώνει συνάψεις και πώς προσαρμόζει συμπεριφορές βάσει εμπειρίας. Αυτή η βιολογική έμπνευση δε σημαίνει άμεση αντιγραφή της ανατομίας, αλλά οδηγεί σε μαθηματικά μοντέλα που στοχεύουν στην ανάδυση ιεραρχικών αναπαραστάσεων από δεδομένα [6],[11].

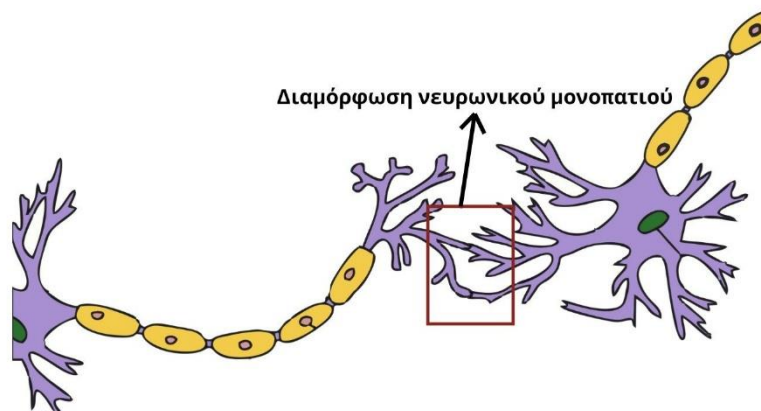
3.2.2 Artificial neural networks

Η δομή ενός βιολογικού νευρώνα όπως φαίνεται στην εικόνα 3.2 αποτελείται κυρίως από τους δενδρίτες, μέσω των οποίων δέχεται ηλεκτρικά σήματα, και τον νευράξονα, από τον οποίο τα μεταδίδει σε άλλους νευρώνες. Τα σήματα αυτά, γνωστά ως δυναμικά ενέργειας, διαδίδονται κατά μήκος πολύπλοκων ακολουθιών νευρώνων και ενεργοποιούνται ως απόκριση σε διάφορα είδη ερεθισμάτων. Τα ερεθίσματα αυτά μπορεί να είναι αισθητικά, όπως το φως, ο ήχος και η αφή, χημικά, όπως οι οσμές και οι γεύσεις, θερμικά, ή ακόμη και εσωτερικής προέλευσης, όπως χημικές μεταβολές του οργανισμού ή αυθόρμητη εγκεφαλική δραστηριότητα



Σχήμα 3.2: Βιολογικός νευρώνας

Ο ανθρώπινος εγκέφαλος μπορεί να θεωρηθεί ως ένα εξαιρετικά πολύπλοκο δίκτυο τέτοιων νευρώνων, συνδεδεμένων μέσω δισεκατομμυρίων συνάψεων όπως η σύναψη που φαίνεται στην εικόνα 3.3. Παρότι βρίσκεται «απομονωμένος» στο εσωτερικό του κρανίου, χωρίς άμεση επαφή με τον εξωτερικό κόσμο, καταφέρνει να αντιλαμβάνεται, να ερμηνεύει και να αντιδρά στο περιβάλλον βασιζόμενος αποκλειστικά στα ηλεκτρικά σήματα που του μεταφέρονται από τις αισθήσεις. Η λειτουργία του βασίζεται στη διαδοχική επεξεργασία αυτών των σημάτων μέσα από συγκεκριμένα νευρωνικά μονοπάτια, τα οποία κωδικοποιούν την πληροφορία και οδηγούν τελικά σε απόφαση ή δράση.



Σχήμα 3.3: Σύναψη μεταξύ βιολογικών νευρώνων

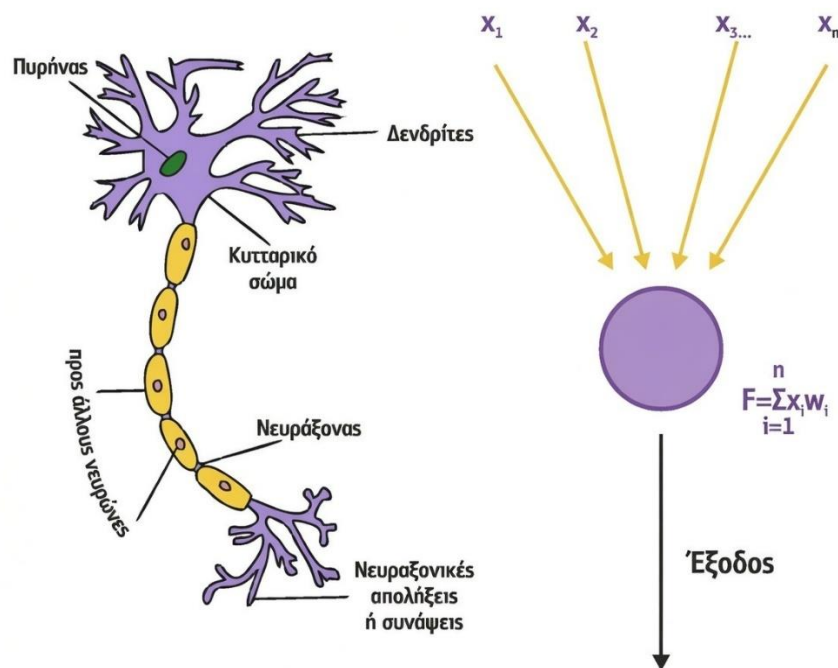
Τα τεχνητά νευρωνικά δίκτυα αντλούν έμπνευση από αυτήν ακριβώς τη βιολογική διαδικασία. Αν και δεν επιχειρούν να αναπαραστήσουν με ακρίβεια τη βιοφυσική λειτουργία του ανθρώπινου εγκεφάλου, υιοθετούν τη βασική φιλοσοφία της επεξεργασίας πληροφορίας μέσω διασυνδεδεμένων μονάδων. Στο πλαίσιο αυτό, τα ερεθίσματα του πραγματικού κόσμου μεταφράζονται σε αριθμητικές τιμές εισόδου

(input values), οι οποίες εισάγονται στο δίκτυο μέσω των νευρώνων εισόδου. Οι τιμές αυτές διαδίδονται στο εσωτερικό του δικτύου, υφίστανται μετασχηματισμούς και συνδυασμούς, και τελικά οδηγούν στην παραγωγή μιας εξόδου, η οποία αντιστοιχεί σε μια απόφαση, μια πρόβλεψη ή μια κατηγοριοποίηση.

Περνώντας πλέον στα τεχνητά νευρωνικά δίκτυα αξίζει να παρουσιαστεί αρχικά η σύλληψη του McCulloch–Pitts νευρώνα που είναι μια απλουστευμένη, μαθηματική απεικόνιση του βιολογικού νευρώνα που παρουσιάστηκε στο κλασικό άρθρο των W. S. McCulloch και W. Pitts το 1943 [1] (McCulloch & Pitts, *A logical calculus of the ideas immanent in nervous activity*). Στο επίπεδο της βιολογίας, ο νευρώνας λαμβάνει εισόδους μέσω των δενδριτών, τις ολοκληρώνει στο κυτταρικό σώμα και — εφόσον το άθροισμα των εισερχομένων σημάτων υπερβεί ένα κρίσιμο όριο στο άξονα (axon hillock) — παράγει έναν ηλεκτρικό παλμό που μεταφέρεται μέσω του νευράξονα προς τις συνάψεις. Η μαθηματική πρόταση των McCulloch–Pitts αναπαριστά αυτή τη διαδικασία με έναν απλό αλγόριθμο: κάθε είσοδος x_i πολλαπλασιάζεται με ένα βάρος w_i , οι σταθμισμένες εισόδους αθροίζονται και το άθροισμα συγκρίνεται με ένα κατώφλι θ . Η έξοδος του M–P νευρώνα είναι δυαδική και δίνεται από τη συνάρτηση:

$$y = H(\sum_i w_i x_i - \theta) \tag{3.2}$$

, όπου $H(\cdot)$ είναι η βηματική (Heaviside) συνάρτηση: παράγει 1 αν η είσοδος υπερβαίνει το όριο και 0 διαφορετικά. Με αυτή τη μορφή ο νευρώνας μπορεί να υλοποιήσει βασικές λογικές πράξεις — για παράδειγμα, με $w_1 = w_2 = 1, \theta = 2$ λειτουργεί ως AND, ενώ για $w_1 = w_2 = 1, \theta = 1$ ως OR — και οι συνδυασμοί τέτοιων μονάδων μπορούν να αναπαράγουν πολύπλοκες λογικές συναρτήσεις, όπως επιτεύχθηκε στο πρωτότυπο άρθρο[1].



Σχήμα 3.4: Σύγκριση βιολογικού και τεχνητού νευρώνα

Η διαδικασία μεταφοράς πληροφορίας στον νευρώνα McCulloch–Pitts, η οποία παρουσιάζεται συγκριτικά με τη μεταφορά ηλεκτρικού σήματος σε έναν βιολογικό νευρώνα στην εικόνα 3.4, περιγράφεται ως εξής: τα εισερχόμενα σήματα αναπαρίστανται ως δυαδικές τιμές στις εισόδους, κάθε τιμή σταθμίζεται και συμβάλλει στο συνολικό άθροισμα, και το αποτέλεσμα της σύγκρισης με το

κατώφλι αποφασίζει αν ένας «σπινθηρισμός» θα μεταφερθεί στην επόμενη στοιβάδα (hidden layer ή output layer) στο επόμενο χρονικό βήμα. Στο πρωτότυπο μοντέλο τα βάρη και το όριο θεωρούνται σταθερά, δηλαδή δεν υπάρχει ενδογενής μηχανισμός μάθησης· η ιδέα της προσαρμογής των βαρών (learning) εισήχθη αργότερα, με μεθόδους όπως ο αλγόριθμος του Perceptron και τελικά το backpropagation που επιτρέπει την αριθμητική βελτιστοποίηση των βαρών[2],[19].

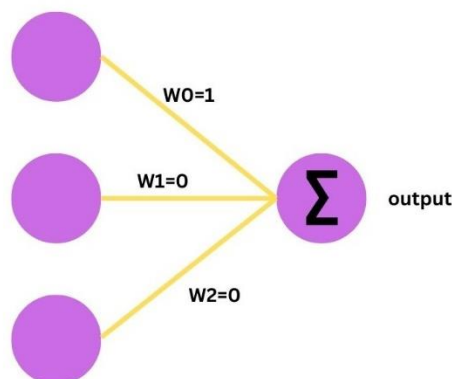
Η χρήση της βηματικής συνάρτησης ενεργοποίησης έχει προφανή περιορισμό: δεν είναι διαφορίσιμη στο όριο, επομένως δεν επιτρέπει εύκολα την εφαρμογή gradient-based μεθόδων μάθησης. Γι' αυτό, στα σύγχρονα τεχνητά νευρωνικά δίκτυα, η ιδέα του McCulloch–Pitts επεκτείνεται στη μορφή της σχέσης 3.3 :

$$y = \phi (\sum_i w_i x_i + b), \quad (3.3)$$

όπου το bias b αντιστοιχεί αρνητικά στο θ , και ϕ είναι μια ομαλή, διαφορίσιμη συνάρτηση ενεργοποίησης (π.χ. sigmoid, tanh ή πιο σύγχρονα ReLU). Αυτές οι συνεχείς συναρτήσεις διατηρούν την έννοια «ενεργοποίησης» αλλά καθιστούν εφικτό τον υπολογισμό παραγώγων και, κατά συνέπεια, τη μαθησιακή προσαρμογή των παραμέτρων με αριθμητικούς αλγορίθμους [11],[19].

Παρά τη σαφή θεωρητική απλότητα και τη σημασία του ως εννοιολογικό θεμέλιο, το μοντέλο McCulloch–Pitts έχει βιολογικούς και μαθηματικούς περιορισμούς: η δυαδική και διακριτή του φύση απλοποιεί πολύ την πραγματική δυναμική των νευρώνων (graded potentials, χρονική πλαίσιο-εξάρτηση, πλαστικότητα συναπτικής ισχύος), δεν περιλαμβάνει μηχανισμούς μάθησης ούτε χρονική λεπτομέρεια για spike timing, και δεν αναπαριστά βιοφυσικές ιδιότητες όπως η μυελίνη ή οι μεταβλητές συναπτικές προσαρμογές. Ωστόσο, ο McCulloch–Pitts νευρώνας υπήρξε καθοριστικός επειδή ενέπνευσε την ιδέα της μονάδας που ομαδοποιεί σταθμισμένες εισόδους και εφαρμόζει μη γραμμική συνάρτηση — την κεντρική δομή κάθε τεχνητού νευρώνα στις σύγχρονες αρχιτεκτονικές [1],[9].

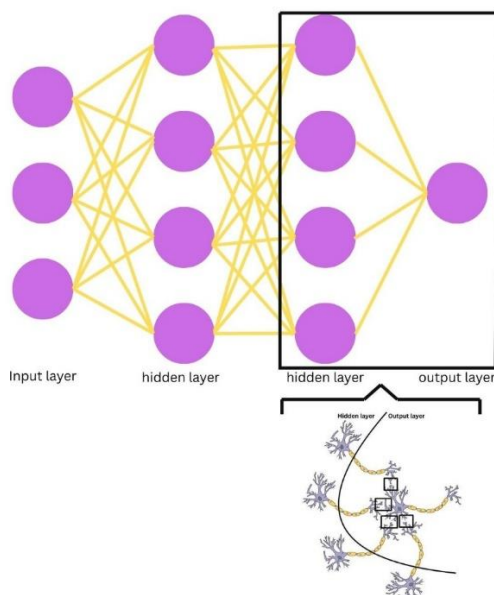
Μετά τη διατύπωση του νευρώνα McCulloch–Pitts, ο οποίος λειτουργούσε με σταθερά βάρη και χωρίς δυνατότητα προσαρμογής, προέκυψε φυσικά η ανάγκη εισαγωγής ενός μηχανισμού μάθησης που να επιτρέπει στο μοντέλο να βελτιώνει τη συμπεριφορά του βάσει εμπειρίας. Το πρώτο ουσιαστικό βήμα προς αυτή την κατεύθυνση έγινε με τον αλγόριθμο του Perceptron από τον Frank Rosenblatt στα τέλη της δεκαετίας του 1950. Η βασική δομή του αντίστοιχου νευρωνικού δικτύου απεικονίζεται στην εικόνα 3.5 .



Σχήμα 3.5: Perceptron

Στο μοντέλο αυτό, τα βάρη του νευρώνα δεν θεωρούνταν πλέον σταθερά, αλλά ενημερώνονταν με βάση το σφάλμα μεταξύ της προβλεπόμενης και της πραγματικής εξόδου. Η βασική ιδέα ήταν απλή: όταν το

δίκτυο κάνει λάθος πρόβλεψη, τα βάρη προσαρμόζονται αναλογικά με την είσοδο και το σφάλμα, ενισχύοντας ή αποδυναμώνοντας τη συμβολή κάθε εισόδου στη μελλοντική απόφαση. Αυτή η διαδικασία αποτελεί μια μαθηματική αναλογία της βιολογικής πλαστικότητας των συνάψεων[2],[10], σύμφωνα με την οποία οι συνδέσεις μεταξύ νευρώνων ενδυναμώνονται ή εξασθενούν ανάλογα με τη δραστηριότητά τους. Αργότερα, με την εισαγωγή διαφορίσιμων συναρτήσεων ενεργοποίησης και του αλγορίθμου οπισθοδιάδοσης του σφάλματος (backpropagation), κατέστη δυνατή η εκπαίδευση πολυεπίπεδων νευρωνικών δικτύων, όπως το ρηχό νευρωνικό δίκτυο της εικόνας 3.6, επιτρέποντας την αυτόματη ρύθμιση μεγάλου αριθμού βαρών μέσω μεθόδων βελτιστοποίησης βασισμένων στην κλίση. Έτσι, τα νευρωνικά δίκτυα απέκτησαν τη δυνατότητα να μαθαίνουν από δεδομένα, μεταβαίνοντας από στατικά λογικά μοντέλα σε προσαρμοστικά συστήματα ικανά να γενικεύονται σε νέα, άγνωστα παραδείγματα[2],[10],[11]. Με άλλα λόγια τα νευρωνικά δίκτυα δεν βασίζονται σε σταθερούς λογικούς κανόνες, αλλά προσαρμόζονται μέσω μάθησης από δεδομένα και μπορούν να γενικεύουν τη γνώση τους σε νέα παραδείγματα.



Σχήμα 3.6: Συσχέτιση βιολογικής διαδικασίας μεταφοράς ηλεκτρικού σήματος με αυτή που συμβαίνει στον τεχνητό νευρώνα .

Η εισαγωγή του αλγορίθμου οπισθοδιάδοσης (1986 από τους Rumelhart, Hinton και Williams.) αποτέλεσε μεν το θεωρητικό υπόβαθρο για την εκπαίδευση πολυεπίπεδων νευρωνικών δικτύων, ωστόσο η πρακτική του αξιοποίηση σε πραγματικά προβλήματα έγινε ουσιαστικά εφικτή μέσα από το έργο του Yann LeCun στα τέλη της δεκαετίας του 1980 και στις αρχές του 1990. Ο LeCun προχώρησε ένα βήμα παραπέρα, συνδυάζοντας το backpropagation με αρχιτεκτονικές συνελκτικών νευρωνικών δικτύων (Convolutional Neural Networks), οι οποίες εκμεταλλεύονται την τοπική συσχέτιση των δεδομένων και τη συμμετρία μετατόπισης, ιδιαίτερα σε προβλήματα εικόνας. Μέσα από αυτό το πλαίσιο, το backpropagation δεν χρησιμοποιήθηκε απλώς για την προσαρμογή βαρών σε πλήρως συνδεδεμένα δίκτυα, αλλά για την εκπαίδευση φίλτρων συνέλιξης που μαθαίνουν αυτόματα χωρικά χαρακτηριστικά από τα δεδομένα. Η προσέγγιση αυτή μείωσε δραστικά τον αριθμό των παραμέτρων, βελτίωσε τη γενίκευση και κατέστησε εφικτή την εκπαίδευση βαθύτερων μοντέλων με πρακτική σημασία. Η επιτυχία των μοντέλων του LeCun, όπως το LeNet για αναγνώριση χειρόγραφων ψηφίων,

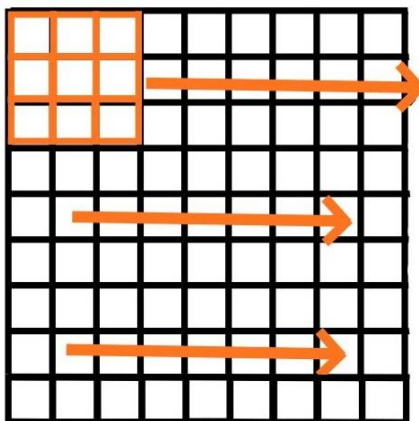
απέδειξε ότι ο συνδυασμός κατάλληλης αρχιτεκτονικής και οπισθοδιάδοσης μπορεί να οδηγήσει σε αποδοτικά και αξιόπιστα συστήματα μάθησης, ανοίγοντας τον δρόμο για τη σύγχρονη εξέλιξη της βαθιάς μάθησης. Η απουσία προσαρμοστικότητας κατέστησε σαφές ότι, για να αποκτήσουν πρακτική αξία, τα μοντέλα αυτά έπρεπε να ενσωματώσουν μηχανισμούς μάθησης[11],[13],[6]

Σε απλούστερα λόγια, όταν τροφοδοτούνται στο δίκτυο αξονικές τομές, κάθε pixel σε κάθε slice λειτουργεί ως μία είσοδος (input) στο μοντέλο· τα pixels που περιέχουν «σήμα» σχετικό με τον όγκο (π.χ. όγκος/βλάβη) θα επηρεάσουν ισχυρότερα την έξοδο επειδή οι συνδυασμοί τους ενεργοποιούν συγκεκριμένα φίλτρα/νευρώνες μέσα στις συνελκτικές στρώσεις. Κατά τη διάρκεια της εκπαίδευσης, ο αλγόριθμος backpropagation υπολογίζει πόσο κάθε είσοδος συνεισφέρει στο σφάλμα και προσαρμόζει αναλόγως τα βάρη· έτσι, οι συνδέσεις (weights) που συνδέονται με περιοχές όπου εμφανίζεται επανειλημμένα το χαρακτηριστικό ενδιαφέροντος αυξάνουν την τιμή τους, χαρακτηρίζοντας αυτές τις θέσεις ως «σημαντικές» για την τελική απόφαση. Όσο μεγαλύτερο και πιο ποικίλο είναι το σύνολο δεδομένων στο οποίο εκτίθεται το μοντέλο —δηλαδή όσο πιο ρεαλιστικά και αντιπροσωπευτικά δείγματα του πληθυσμού περιλαμβάνονται— τόσο πιο αξιόπιστη γίνεται αυτή η εκμάθηση βαρών: τα μοτίβα που επαναλαμβάνονται σε πολλαπλά δείγματα θα ισχυροποιηθούν, ενώ τα τυχαία θορυβώδη μοτίβα δεν θα αποκτήσουν σταθερή επίδραση. Με άλλα λόγια, περισσότερα και πιο αντιπροσωπευτικά δεδομένα βελτιώνουν τη γενίκευση του δικτύου και μειώνουν την τάση για overfitting, καθώς το μοντέλο μαθαίνει πραγματικές, σταθερές συσχετίσεις αντί για συμπτώσεις του συγκεκριμένου training set[4],[6]. Γι' αυτό η σωστή δειγματοληψία, η αύξηση δεδομένων και οι τεχνικές regularization είναι κρίσιμες — διασφαλίζουν ότι τα βάρη που αποκτούν υψηλή τιμή αντανακλούν πραγματική, κλινικά σημαντική πληροφορία[12].

3.2.3 Convolutional neural networks

Η βασική ιδέα πίσω από τα Convolutional Neural Networks (CNNs) μπορεί να περιγραφεί με όρους αναλογίας προς τον τρόπο που λειτουργεί η όραση: το ανθρώπινο μάτι δεν «βλέπει» απλώς μια εικόνα ως σύνολο pixel, αλλά αναγνωρίζει σταδιακά χαρακτηριστικά — γραμμές, ακμές, σκιές — και στη συνέχεια συνθέτει αυτές τις πρώτες ενδείξεις σε όλο και πιο σύνθετες δομές (π.χ. ένα μάτι, ένα πρόσωπο, ένα αντικείμενο). Τα CNN υλοποιούν ακριβώς αυτή τη σταδιακή εξαγωγή χαρακτηριστικών μέσω διαδοχικών επιπέδων, ώστε στο τελικό στάδιο το δίκτυο να μπορεί να τοποθετήσει την εικόνα σε μια κατάλληλη κατηγορία[13],[9],[6].

Το κύριο «εργαλείο» με το οποίο λειτουργεί το CNN είναι το συνελκτικό (convolutional) φίλτρο: ένα μικρό kernel που «γλιστράει» πάνω στην εικόνα και για κάθε θέση υπολογίζει ένα εσωτερικό γινόμενο μεταξύ των τιμών του φίλτρου και των τιμών των pixel της τοπικής περιοχής. Η διαδικασία της ολίσθησης του kernel πάνω σε μια εικόνα απεικονίζεται σχηματικά στην Εικόνα 3.7 για την περίπτωση μιας δισδιάστατης (2D) απεικόνισης. Το αποτέλεσμα αυτής της ολίσθησης σχηματίζει ένα feature map — έναν χάρτη που δείχνει πόσο έντονα γίνεται αντιληπτό το μοτίβο που αντιπροσωπεύει το φίλτρο σε κάθε θέση. Με αυτόν τον τρόπο ένα φίλτρο μπορεί να εξειδικευτεί στην αναγνώριση συγκεκριμένων μοτίβων (π.χ. οριζόντιες ακμές), ενώ διαφορετικά φίλτρα μαθαίνουν διαφορετικά μοτίβα[13],[4].



Σχήμα 3.7: Διαισθητικό κομμάτι της συνέλιξης σε ένα 2d-Convolutional Layer

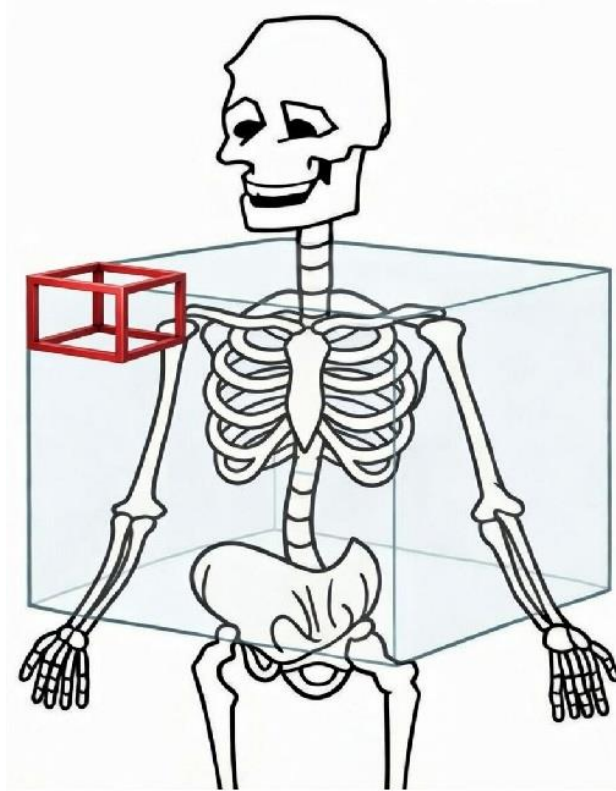
Δύο από τα πιο θεμελιώδη χαρακτηριστικά των Convolutional Neural Networks είναι η κοινή χρήση βαρών (weight sharing) και η τοπική σύζευξη (local connectivity). Στην κοινή χρήση βαρών, το ίδιο συνελκτικό φίλτρο — δηλαδή το ίδιο σύνολο εκπαιδευσίμων παραμέτρων — εφαρμόζεται σε όλες τις χωρικές θέσεις της εικόνας. Αν και οι τιμές των βαρών προσαρμόζονται κατά τη διάρκεια της εκπαίδευσης μέσω του αλγορίθμου οπισθοδιάδοσης, παραμένουν κοινές για κάθε θέση εφαρμογής του φίλτρου. Με αυτόν τον τρόπο, το δίκτυο μπορεί να ανιχνεύει το ίδιο χαρακτηριστικό ανεξάρτητα από τη θέση του, μειώνοντας ταυτόχρονα τον αριθμό των παραμέτρων και τον κίνδυνο υπερπροσαρμογής (overfitting).

Παράλληλα, η τοπική σύζευξη αναφέρεται στο γεγονός ότι κάθε έξοδος ενός συνελκτικού επιπέδου εξαρτάται μόνο από ένα μικρό, τοπικό τμήμα της εισόδου, το οποίο ορίζεται από το μέγεθος του φίλτρου. Έτσι, το δίκτυο δεν επεξεργάζεται ολόκληρη την εικόνα ταυτόχρονα, αλλά εστιάζει σε μικρά χωρικά patches, επιτρέποντας την ανίχνευση τοπικών χαρακτηριστικών χαμηλού επιπέδου. Αυτή η συμπεριφορά αποτελεί αναλογία του βιολογικού «πεδίου πρόσληψης» (*receptive field*) του ανθρώπινου οπτικού συστήματος, όπου κάθε νευρώνας ανταποκρίνεται σε περιορισμένη περιοχή του οπτικού πεδίου. Μέσα από τη διαδοχική στοιβάζση συνελκτικών επιπέδων, τα τοπικά αυτά χαρακτηριστικά συνδυάζονται προοδευτικά σε πιο σύνθετες και αφηρημένες αναπαραστάσεις, οδηγώντας στη συνολική κατανόηση της εικόνας από το δίκτυο[13],[9],[6].

Στην πράξη οι αρχιτεκτονικές χρησιμοποιούν πολλαπλά συνελκτικά επίπεδα: τα πρώτα επίπεδα εντοπίζουν απλά, χαμηλού επιπέδου χαρακτηριστικά· τα μεσαία επίπεδα συνθέτουν αυτές τις ενδείξεις σε πιο σύνθετα μοτίβα· και τα βαθύτερα επίπεδα αντιστοιχούν σε αφηρημένες αναπαραστάσεις που είναι χρήσιμες για το τελικό καθήκον (π.χ. ταξινόμηση). Ενδιάμεσα συχνά ενσωματώνονται επίπεδα pooling που μειώνουν διαστάσεις και παρέχουν ανοχή σε μικρές μετατοπίσεις, καθώς και μη γραμμικές συναρτήσεις ενεργοποίησης που επιτρέπουν τη μάθηση πολύπλοκων συσχετίσεων[13],[4],[6].

Η εκμάθηση των φίλτρων γίνεται δεδομένων κάποιων παραδειγμάτων· μέσω backpropagation[11] ο αλγόριθμος προσαρμόζει τα βάρη έτσι ώστε τα ενεργοποιούμενα μοτίβα να γίνουν όσα πράγματι ενδιαφέρουν για την επίδοση στο πρόβλημα. Όσο πιο πολλά και πιο αντιπροσωπευτικά τα δεδομένα εκπαίδευσης[19], τόσο πιο σταθερές και χρήσιμες γίνονται αυτές οι συναρτήσεις φίλτρων (και, αντίστροφα, τόσο λιγότερο πιθανό είναι να προκύψει overfitting σε τυχαία μοτίβα)[6].

Τέλος, αξίζει να σημειωθεί ότι η ίδια φιλοσοφία εφαρμόζεται και σε τρισδιάστατα δεδομένα (όπως όγκοι CT): εκεί χρησιμοποιούνται 3D convolution kernels [12] που κινούνται και κατά τον άξονα του βάθους, όπως συμβαίνει στην Εικόνα 3.8, ώστε το δίκτυο να μάθει χωρικο-διαχρονικά μοτίβα μέσα στον όγκο και όχι μόνο σε επιμέρους τομές. Με αυτόν τον τρόπο τα CNN προσφέρουν μια ισχυρή, επεκτάσιμη προσέγγιση για την αυτόματη εξαγωγή πολυεπίπεδων χαρακτηριστικών σε προβλήματα όρασης και ιατρικής απεικόνισης[17].



Σχήμα 3.8: Διαισθητικό κομμάτι της συνέλιξης σε ένα 3d-Convolutional Layer

3.2.3.1 Max-pooling layer

Το επίπεδο max-pooling παίζει κρίσιμο ρόλο στην αύξηση της ευρωστίας του δικτύου απέναντι σε διαφορές που προέρχονται από τα πρωτόκολλα απεικόνισης ή ακόμα και από τη διαφορά modality (CT vs MRI). Με απλά λόγια, το pooling «ψάχνει» την ισχυρότερη (ή την πιο αντιπροσωπευτική) ενεργοποίηση μέσα σε κάθε τοπικό παράθυρο(feature map) και την κρατάει, ενώ «πετάει» το υπόλοιπο[13]. Στην πράξη αυτό σημαίνει ότι ακόμη κι αν δύο εικόνες του ίδιου ανατομικού τμήματος διαφέρουν λόγω διαφορετικής παραμετροποίησης της λυχνίας ή λόγω διαφορετικού μηχανήματος, τα τοπικά, επαναλαμβανόμενα χαρακτηριστικά — δηλαδή η εγγύτητα σε κάποιο σημαντικό μοτίβο — θα επιβιώσουν στο pooled feature map[6]. Με κάθε pooling βήμα «αποβάλλεται» ένα σημαντικό ποσοστό της λεπτομέρειας που δεν σχετίζεται άμεσα με το χαρακτηριστικό ενδιαφέροντος: αυτό αφενός μειώνει τον όγκο των υπό επεξεργασία δεδομένων και άρα τον αριθμό των παραμέτρων σε επόμενα layers, αφετέρου μειώνει την ευπάθεια του μοντέλου στο να «μαθαίνει» πολύ ειδικές λεπτομέρειες του training set (overfitting)[19]. Το pooled feature map, δηλαδή η έξοδος του max-pooling layer, δεν είναι παρά μια αριθμητική αναπαράσταση — για τον άνθρωπο μεμονωμένα αυτά τα νούμερα δεν έχουν νόημα — αλλά αποτελούν το πυρήνα πάνω στον οποίο η εξηγησιμότητα (XAI) μπορεί να εργαστεί: με τεχνικές επεξηγησιμότητας επιχειρείτε η χαρτογράφηση αυτών ενεργοποιήσεων σε εικόνες/θερμικούς χάρτες

που ο τελικός αναγνώστης (κλινικός ή ερευνητής) να μπορεί να κατανοήσει και να συνδέσει με πραγματικά ανατομικά/παθολογικά ευρήματα[12].

Συμπληρωματικά, αξίζει να επισημανθεί ότι αυτή η συμπεριφορά του max-pooling μιμείται με έναν πρόχειρο αλλά αποτελεσματικό τρόπο την ανθρώπινη οπτική στρατηγική: ο άνθρωπος δεν περιμένει να δει ένα χαρακτηριστικό μόνο από μία συγκεκριμένη γωνία, υπό συγκεκριμένο φωτισμό ή σε μία μοναδική κλίμακα για να το αναγνωρίσει — αντίθετα συχνά «ρίχνει μια γενική, πιο καθολική ματιά» και διακρίνει τα κρίσιμα μοτίβα παρά τις τοπικές διακυμάνσεις[9]. Το max-pooling, με την επιλογή της ισχυρότερης ενεργοποίησης ανά τοπικό παράθυρο, συμπύσσει τις τοπικές λεπτομέρειες και διατηρεί την πιο αντιπροσωπευτική πληροφορία· έτσι το δίκτυο αποκτά ανθεκτικότητα σε αλλαγές πρωτοκόλλου απεικόνισης ή σε διαφοροποιήσεις μεταξύ CT και MRI[13]. Με άλλα λόγια, «δίνουμε μάτια» στο νευρωνικό δίκτυο — το κάνουμε ικανό να βλέπει πιο ολιστικά, κατά τρόπο ανάλογο με την ανθρώπινη οπτική αντίληψη κάτω από διαφορετικές συνθήκες.

3.2.3.2 Flattening layer

Στη συνέχεια ακολουθεί το στάδιο του flattening: το pooled feature map — δηλαδή η συλλογή των ενεργοποιήσεων που προέκυψαν από την εφαρμογή των φίλτρων και του pooling — «ξεδιπλώνεται» σε ένα ενιαίο μονοδιάστατο διάνυσμα. Αυτό το διάνυσμα εισάγεται στο πλήρως συνδεδεμένο (dense / artificial neural) επίπεδο, όπου κάθε νευρώνας υπολογίζει ένα σταθμισμένο άθροισμα των εισόδων του και εφαρμόζει μια μη γραμμική συνάρτηση ενεργοποίησης[11]. Με άλλα λόγια, εδώ συνδυάζονται με διαφορετικούς τρόπους τα χαρακτηριστικά που εξήχθησαν προηγουμένως: κάποιιοι συνδυασμοί (κάποιοι «τρόποι σκέψης») θα έχουν μεγαλύτερη βαρύτητα και μεγαλύτερη επίδραση στην τελική απόφαση, ενώ άλλοι θα αποδυναμωθούν. Στο μαθηματικό επίπεδο αυτό υλοποιείται ως dot-product μεταξύ του διανύσματος χαρακτηριστικών και του διανύσματος βαρών κάθε νευρώνα, ακολουθούμενο από πιθανή προσθήκη bias και εφαρμογή activation (π.χ. ReLU ή sigmoid για την έξοδο)[13]. Τελικά, τα μεμονωμένα συνεισφέροντα «συγκεντρώνονται» — αθροίζονται/συνδυάζονται — και προκύπτει ένα τελικό score ή μια πιθανότητα, η οποία ερμηνεύεται ως η απόφαση του δικτύου (π.χ. «καρκίνος» ή «μη καρκίνος»).

3.3 Ιατρική

Στο πλαίσιο της παρούσας διπλωματικής εργασίας, ο όρος «ιατρική» δεν αναφέρεται αποκλειστικά στην κλινική πρακτική ή στη βιολογική διάσταση της υγείας, αλλά κυρίως στο επιστημονικό πεδίο στο οποίο εφαρμόζονται φυσικές αρχές και τεχνολογικές μέθοδοι για την παραγωγή, επεξεργασία και ερμηνεία ιατρικής πληροφορίας. Με αυτή την έννοια, η ιατρική προσεγγίζεται ως ένα κατεξοχήν διεπιστημονικό πεδίο, στο οποίο η φυσική αποτελεί θεμελιώδη συνιστώσα της εξέλιξής της, οδηγώντας στη διαμόρφωση του κλάδου της Ιατρικής Φυσικής. Η Ιατρική Φυσική περιλαμβάνει την εφαρμογή αρχών της φυσικής σε ιατρικές τεχνολογίες, όπως η ιατρική απεικόνιση (π.χ. αξονική και μαγνητική τομογραφία), η ακτινοθεραπεία και οι διαγνωστικές τεχνικές, και αποτελεί τον βασικό μηχανισμό μέσω του οποίου η φυσική συμβάλλει άμεσα στη βελτίωση της διάγνωσης και της θεραπείας. Υπό αυτή την οπτική, η πρόοδος της ιατρικής δεν προκύπτει μόνο από βιολογικές ή κλινικές ανακαλύψεις, αλλά σε μεγάλο βαθμό από την εξέλιξη φυσικών μεθόδων και υπολογιστικών τεχνολογιών που επιτρέπουν την ακριβέστερη απεικόνιση και ανάλυση του ανθρώπινου σώματος. Συνεπώς, στην παρούσα εργασία, η ιατρική αντιμετωπίζεται ως το πεδίο εφαρμογής της Ιατρικής Φυσικής, το οποίο σε συνδυασμό με τη μηχανική μάθηση αποτελεί το βασικό υπόβαθρο για την ανάπτυξη σύγχρονων υπολογιστικών συστημάτων ανάλυσης ιατρικών δεδομένων.[41]

3.3.1 Ηλεκτρομαγνητικό φάσμα

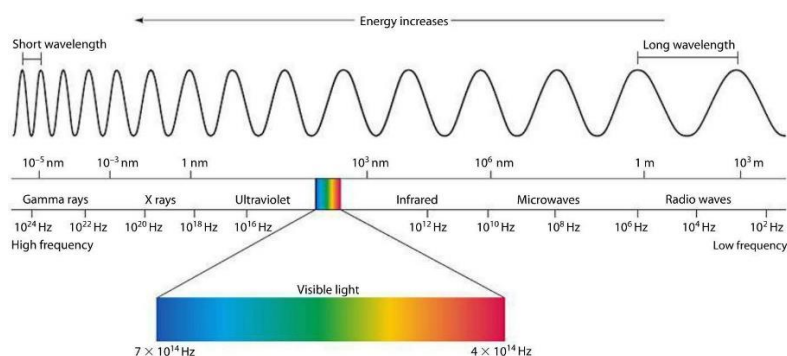
Η πιο θεμελιώδης γνώση για την κατανόηση της έννοιας των ακτίνων X είναι το ηλεκτρομαγνητικό φάσμα. Η ηλεκτρομαγνητική ακτινοβολία αποτελεί μορφή ενέργειας που διαδίδεται στο χώρο με τη μορφή ηλεκτρικών και μαγνητικών πεδίων, τα οποία ταλαντώνονται κάθετα μεταξύ τους και κάθετα προς την κατεύθυνση διάδοσης, σχηματίζοντας εγκάρσια κύματα [20],[21]. Κάθε κύμα χαρακτηρίζεται από συγκεκριμένη συχνότητα f , μήκος κύματος λ και ενέργεια E . Η ενέργεια ενός φωτονίου δίνεται από τη σχέση 3.1 :

$$E = hf \quad (3.4)$$

, όπου h η σταθερά του Planck, ενώ η ταχύτητα διάδοσης σχετίζεται με το μήκος κύματος και τη συχνότητα μέσω $c = \lambda f$ [21],[22]. Επιπλέον, η ορμή ενός φωτονίου μπορεί να εκφραστεί ως $p = h/\lambda$, συνδέοντας την ενέργεια με τη σωματιδιακή φύση του φωτός[22].

Η ηλεκτρομαγνητική ακτινοβολία παρουσιάζει διττή φύση (wave-particle duality), καθώς εκδηλώνει κυματικές ιδιότητες, όπως παρεμβολή και περίθλαση, αλλά και σωματιδιακές συμπεριφορές, όπως το φωτοηλεκτρικό φαινόμενο [22]. Το φωτοηλεκτρικό πείραμα, εξηγούμενο κβαντικά από τον Einstein το 1905, απέδειξε ότι η εκπομπή ηλεκτρονίων από μέταλλα εξαρτάται από την ενέργεια των φωτονίων και όχι από την ένταση του φωτός, επιβεβαιώνοντας τη σωματιδιακή φύση τους[22]. Η διττή αυτή συμπεριφορά αξιοποιείται στην πράξη, τόσο στην ανίχνευση ακτινοβολίας μέσω φωτοηλεκτρικών ανιχνευτών όσο και στις εφαρμογές περίθλασης για απεικονιστικές τεχνικές [23],[24].

Κατά μήκος του ηλεκτρομαγνητικού φάσματος όπως φαίνεται στην εικόνα 3.9, όσο αυξάνεται η συχνότητα, το μήκος κύματος μειώνεται και η ενέργεια ανά φωτόνιο αυξάνεται. Στο τμήμα υψηλών συχνοτήτων και μικρών μηκών κύματος βρίσκονται οι ακτίνες X, οι οποίες λόγω της υψηλής ενέργειας των φωτονίων τους παρουσιάζουν αυξημένη ικανότητα διείσδυσης στην ύλη και αξιοποιούνται σε εφαρμογές ιατρικής απεικόνισης [23],[26]. Οι ακτίνες X βρίσκονται μεταξύ της υπεριώδους ακτινοβολίας και της περιοχής των γ-ακτίνων, με μήκη κύματος συνήθως μεταξύ 0,01–10 nm και ενέργειες από 0,1 έως 120 keV [23],[25].



Σχήμα 3.9: Ηλεκτρομαγνητικό φάσμα

3.3.2 Λυχνία ακτίνων X

Η σύγχρονη τεχνολογία παραγωγής ακτίνων X βασίζεται στη θεμελιώδη συμβολή του William David Coolidge, ο οποίος στις αρχές του 20ού αιώνα εισήγαγε τη λυχνία ακτίνων X με θερμαινόμενη κάθοδο (Coolidge tube). Η καινοτομία αυτή επέτρεψε τον ακριβή έλεγχο της εκπομπής ηλεκτρονίων μέσω θερμιονικής εκπομπής, καθιστώντας δυνατή τη σταθερή, επαναλήψιμη και ασφαλή παραγωγή

ακτινοβολίας X, σε αντίθεση με τις πρώιμες λυχνίες ψυχρής καθόδου [24]. Η βασική αρχή λειτουργίας της λυχνίας Coolidge παραμένει έως σήμερα ο ακρογωνιαίος λίθος των ιατρικών συστημάτων ακτινολογίας και υπολογιστικής τομογραφίας[26].

Σε μια τυπική λυχνία ακτίνων X, η κάθοδος αποτελείται από ένα λεπτό νήμα, το οποίο θερμαίνεται μέσω εφαρμογής ηλεκτρικού ρεύματος. Η θέρμανση αυτή προκαλεί την εκπομπή ηλεκτρονίων από την επιφάνεια του μετάλλου μέσω του φαινομένου της θερμιονικής εκπομπής. Τα εκπεμπόμενα ηλεκτρόνια σχηματίζουν ηλεκτρονική δέσμη, η οποία επιταχύνεται προς τον άνοδο (στόχο) μέσω εφαρμογής υψηλής διαφοράς δυναμικού, συνήθως της τάξης δεκάδων έως εκατοντάδων kilovolt σε ιατρικές εφαρμογές [24],[25]. Η ενέργεια που αποκτούν τα ηλεκτρόνια εξαρτάται άμεσα από την εφαρμοζόμενη τάση και καθορίζει το ενεργειακό φάσμα της παραγόμενης ακτινοβολίας[25].

Η μέγιστη δυνατή ενέργεια των παραγόμενων φωτονίων ακτίνων X δίνεται από τη σχέση Duane–Hunt(3.3)[25],

$$h\nu_{\max} = eV, \quad (3.5)$$

όπου e είναι το στοιχειώδες φορτίο και V η διαφορά δυναμικού μεταξύ καθόδου και ανόδου. Η σχέση αυτή θέτει το κατώτερο όριο μήκους κύματος του φάσματος ($\lambda_{\min} = hc/eV$) και αποτελεί θεμελιώδη περιορισμό στη φασματική κατανομή της ακτινοβολίας X [25]. Στην πράξη, μόνο ένα μικρό ποσοστό της κινητικής ενέργειας των ηλεκτρονίων μετατρέπεται σε ακτινοβολία, ενώ το μεγαλύτερο μέρος αποδίδεται ως θερμότητα στον στόχο, γεγονός που καθιστά αναγκαία τη χρήση ανόδων υψηλής θερμικής αντοχής και, συχνά, περιστρεφόμενων ανόδων για την αποδοτική απαγωγή θερμότητας [24],[25].

3.3.3 Μηχανισμοί παραγωγής ακτίνων X

Η ακτινοβολία bremsstrahlung (γερμανικός όρος που σημαίνει «φρενάρισμα») παράγεται όταν τα επιταχυνόμενα ηλεκτρόνια αλληλεπιδρούν με το ηλεκτρικό πεδίο του πυρήνα των ατόμων του στόχου. Κατά την προσέγγιση του πυρήνα, η τροχιά και η ταχύτητα των ηλεκτρονίων μεταβάλλονται απότομα, με αποτέλεσμα την εκπομπή φωτονίων ακτίνων X συνεχούς ενεργειακής κατανομής. Το παραγόμενο φάσμα εκτείνεται από χαμηλές ενέργειες έως το μέγιστο όριο που καθορίζεται από τη σχέση Duane–Hunt(3.3) και εξαρτάται ισχυρά από την εφαρμοζόμενη τάση και το ατομικό αριθμό Z του υλικού στόχου [25],[26]. Στις περισσότερες ιατρικές λυχνίες, η bremsstrahlung αποτελεί το κυρίαρχο συστατικό του παραγόμενου φάσματος[26],[27].

Η χαρακτηριστική ακτινοβολία προκύπτει όταν ένα προσπίπτον ηλεκτρόνιο ιονίζει ηλεκτρόνιο εσωτερικής στιβάδας (π.χ. K ή L) του ατόμου του στόχου. Το κενό που δημιουργείται πληρώνεται από ηλεκτρόνιο ανώτερης ενεργειακής στάθμης, και η διαφορά ενέργειας μεταξύ των δύο σταθμών εκπέμπεται υπό μορφή φωτονίου ακτίνων X με διακριτή ενέργεια. Οι ενέργειες αυτές είναι χαρακτηριστικές του στοιχείου του στόχου και εμφανίζονται στο φάσμα ως αιχμές (π.χ. $K\alpha$, $K\beta$). Η χαρακτηριστική ακτινοβολία παίζει σημαντικό ρόλο σε εξειδικευμένες εφαρμογές, όπως η μαστογραφία, όπου επιλέγονται συγκεκριμένα υλικά στόχου (π.χ. Mo ή Rh) για τη βελτιστοποίηση της αντίθεσης εικόνας και της δόσης [24],[25],[29].

3.3.4 Αλληλεπίδραση ακτινοβολίας με το σώμα

Όταν μια δέσμη φωτονίων, όπως οι ακτίνες X, διέρχεται μέσα από ένα υλικό ή έναν βιολογικό ιστό, παρατηρείται μείωση της έντασής της, φαινόμενο γνωστό ως εξασθένηση. Η εξασθένηση περιγράφει

τη μείωση του αριθμού των φωτονίων που συνεχίζουν στην αρχική τους κατεύθυνση λόγω αλληλεπιδράσεων με τα άτομα του μέσου και εξαρτάται τόσο από την ενέργεια των φωτονίων όσο και από τις φυσικές ιδιότητες του υλικού, όπως η πυκνότητα και το αποτελεσματικό ατομικό του νούμερο [27], [30],[33]. Κατά τη διέλευση μιας δέσμης ακτίνων X μέσα από την ύλη, κάθε φωτόνιο μπορεί να υποστεί μία από τις ακόλουθες τρεις περιπτώσεις: (α) να απορροφηθεί πλήρως, (β) να σκεδαστεί, αλλάζοντας διεύθυνση και ενδεχομένως ενέργεια, ή (γ) να διαπεράσει το υλικό χωρίς καμία αλληλεπίδραση [27],[36],[37].

Οι ακτίνες X αποτελούν μορφή ιοντίζουσας ακτινοβολίας, καθώς οι αλληλεπιδράσεις τους με την ύλη μπορούν να προκαλέσουν ιοντισμό των ατόμων ή μορίων του ιστού, δηλαδή απομάκρυνση ηλεκτρονίων από τα ατομικά τους τροχιακά [28],[31]. Στο εύρος ενεργειών που χρησιμοποιείται στη διαγνωστική ακτινολογία, οι κυριότεροι μηχανισμοί αλληλεπίδρασης της ακτινοβολίας με το ανθρώπινο σώμα είναι το φωτοηλεκτρικό φαινόμενο, η σκέδαση Compton και η σκέδαση Rayleigh[27],[32].

Στο φωτοηλεκτρικό φαινόμενο, το φωτόνιο απορροφάται πλήρως από ένα άτομο, μεταφέροντας όλη την ενέργειά του σε ένα ηλεκτρόνιο εσωτερικής στιβάδας, το οποίο εκτινάσσεται από το άτομο. Η πιθανότητα εμφάνισης του φαινομένου εξαρτάται έντονα από το ατομικό αριθμό του υλικού και μειώνεται με την αύξηση της ενέργειας των φωτονίων, γεγονός που το καθιστά ιδιαίτερα σημαντικό για υλικά υψηλού Z και χαμηλότερες ενέργειες ακτίνων X [27], [30],[33].

Η σκέδαση Compton συμβαίνει όταν ένα φωτόνιο αλληλεπιδρά με ένα χαλαρά δεσμευμένο ηλεκτρόνιο εξωτερικής στιβάδας, μεταβιβάζοντάς του μέρος της ενέργειάς του. Το φωτόνιο συνεχίζει την πορεία του με μειωμένη ενέργεια και διαφορετική κατεύθυνση. Το φαινόμενο Compton είναι ο κυρίαρχος μηχανισμός αλληλεπίδρασης στις μεσαίες ενέργειες που χρησιμοποιούνται στην ιατρική απεικόνιση και αποτελεί κύρια αιτία υποβάθμισης της αντίθεσης της εικόνας λόγω της παραγωγής σκεδαζόμενης ακτινοβολίας [27], [36],[37].

Η σκέδαση Rayleigh, ή συνεκτική σκέδαση, αποτελεί ελαστική αλληλεπίδραση κατά την οποία το φωτόνιο αλληλεπιδρά με ολόκληρο το άτομο, αλλάζει διεύθυνση χωρίς απώλεια ενέργειας και χωρίς να προκαλεί ιοντισμό. Παρότι η συνεισφορά της είναι μικρότερη σε σχέση με τη σκέδαση Compton, η Rayleigh σκέδαση συμβάλλει επίσης στη μείωση της ποιότητας εικόνας, ιδιαίτερα σε χαμηλές ενέργειες φωτονίων και πυκνά υλικά [27],[36].

Η συνολική εξασθένηση της ακτινοβολίας προκύπτει από το συνδυασμένο αποτέλεσμα των παραπάνω μηχανισμών και παίζει καθοριστικό ρόλο τόσο στην ποιότητα της παραγόμενης εικόνας όσο και στη δόση ακτινοβολίας που λαμβάνει ο ασθενής[26],[28],[39]. Η κατανόηση των επιμέρους αλληλεπιδράσεων αποτελεί βασική προϋπόθεση για την ποσοτική περιγραφή της εξασθένησης, η οποία θα αναλυθεί μαθηματικά στην επόμενη ενότητα[24],[29].

3.3.5 Εξαγωγή μαθηματικών σχέσεων

Στην μελέτη της αλληλεπίδρασης των ακτίνων-X με την ύλη, δεν αρκεί απλώς να κατανοήσουμε διαισθητικά την εξασθένηση της δέσμης κατά μήκος της διέλευσής της από ένα αντικείμενο· για την ποσοτικοποίηση, τον σχεδιασμό πειραμάτων, την επεξεργασία σήματος και την ανακατασκευή εικόνων είναι απαραίτητο να εκφράσουμε αυτά που παρατηρούμε με σαφείς μαθηματικές σχέσεις[32],[34]. Μία τέτοια θεμελιώδης μαθηματική σχέση είναι η εκθετική μείωση της έντασης $I(x)$ μιας στενής, μονοχρωματικής δέσμης ακτίνων-X καθώς διαπερνά υλικό με συνολικό πάχος x . Η εξασθένηση αυτή περιγράφεται από την διαφορική σχέση :

$$dI(x) = -\mu I(x) dx, \quad (3.6)$$

όπου $I(x)$ είναι η στιγμιαία ένταση της δέσμης στο βάθος x , και μ γραμμικός συντελεστής αποσβεστικότητας του υλικού (με μονάδα 1/μήκος). Ο στοιχειώδης όρος μdx αντιπροσωπεύει την πιθανότητα ενός τυπικού φωτονίου να αλληλεπιδράσει και να «αφαιρεθεί» από τη δέσμη στο στοιχειώδες στρώμα πάχους dx . Ολοκληρώνοντας την παραπάνω διαφορική εξίσωση με αρχική συνθήκη $I(0) = I_0$, λαμβάνουμε την εκθετική μορφή :

$$I(x) = I_0 e^{-\mu x}, \quad (3.7)$$

η οποία δείχνει καθαρά ότι η ένταση μειώνεται με εκθετικό ρυθμό όσο το πάχος του υλικού αυξάνει[32],[34],[40]. Αυτή αποτελεί την πρακτική μορφή του νόμου Beer–Lambert για στενές, μονοχρωματικές δέσμες και αποτελεί την βάση για τις γραμμικές ολοκληρωματικές σχέσεις που χρησιμοποιούνται στην αξονική τομογραφία (CT) και στην επεξεργασία των μετρήσεων της εισόδου–εξόδου της δέσμης για ανακατασκευή εικόνας [32], [34].

Όταν μια ακτίνα διαπερνά διαδοχικά δύο ή περισσότερα στρώματα υλικών με γραμμικούς συντελεστές μ_1, μ_2, \dots και αντίστοιχα πάχη x_1, x_2, \dots , τότε η συνολική μετάδοση εκφράζεται ως το γινόμενο των επιμέρους μεταδόσεων :

$$I_{\text{trans}} = I_0 e^{-\mu_1 x_1} e^{-\mu_2 x_2} \dots = I_0 e^{-(\mu_1 x_1 + \mu_2 x_2 + \dots)}. \quad (3.8)$$

Γενικά, για ένα συνεχή, ετερογενές υλικό όπου ο γραμμικός συντελεστής εξαρτάται από τη θέση $\mu(x)$, η αντίστοιχη σχέση είναι :

$$I_{\text{trans}} = I_0 \exp\left(-\int_0^L \mu(x) dx\right). \quad (3.9)$$

Αυτή η γενίκευση είναι ουσιαστική για την αξονική τομογραφία, όπου η πληροφορία που λαμβάνεται σε κάθε προβολή είναι μια γραμμική προβολή του $\mu(x)$ κατά μήκος γραμμών που διασχίζουν τον όγκο του εξεταζόμενου αντικειμένου [32],[34],[40].

Ο νόμος $I = I_0 e^{-\mu x}$ είναι στατιστικός στην ουσία του: η I που μετράτε σε έναν ανιχνευτή δεν είναι μια σταθερή τιμή αλλά μια τυχαία μεταβλητή που αντικατοπτρίζει την πιθανοτική φύση των αλληλεπιδράσεων αφαίρεσης φωτονίων από τη δέσμη[32],[34],[40]. Σε πολλές εφαρμογές CT με photon-counting ανιχνευτές, ο αριθμός των μετρημένων φωτονίων N σε δεδομένο χρονικό παράθυρο μπορεί να θεωρηθεί ως μια μεταβλητή Poisson, με μέση τιμή ανάλογη του I_0 : οι αποκλίσεις και η διασπορά των μετρήσεων έχουν σημαντικό ρόλο στην εκτίμηση του συντελεστή μ και στον υπολογισμό σφαλμάτων των εκτιμήσεων αυτών [32],[34],[40]. Για πιο προχωρημένα στατιστικά μοντέλα, όπου λαμβάνονται υπ' όψη και ενεργειακές εξαρτήσεις του φάσματος της δέσμης, χρησιμοποιούνται κατανομές compound Poisson ή προσεγγίσεις με Gamma ή Gaussian κατανομές ανάλογα με την ενεργειακή πολυπλοκότητα και το σήμα-θόρυβος [19].

Ο γραμμικός συντελεστής μ έχει μικροσκοπική ερμηνεία μέσω της σχέσης

$$\mu = n \sigma_{\text{tot}}, \quad (3.10)$$

όπου n είναι ο αριθμός των ατόμων/στόχων ανά μονάδα όγκου και σ_{tot} η συνολική διατομή για όλες τις διεργασίες αφαίρεσης (φωτοηλεκτρική απορρόφηση, Compton, Rayleigh κ.λπ.). Με δεδομένη την

πυκνότητα ρ και το ατομικό/μοριακό βάρος A , αυτή συνδέεται με τον μαζικό συντελεστή αποσβεστικότητας μ/ρ ως

$$\frac{\mu}{\rho} = \frac{N_A}{A} \sigma_{\text{tot}}, \quad (3.11)$$

όπου N_A είναι η σταθερά του Avogadro. Η χρήση του μ/ρ κάνει δυνατή την εύκολη σύγκριση υλικών με διαφορετικές πυκνότητες, καθώς ο μαζικός συντελεστής αποσβεστικότητας είναι ουσιαστικά ανεξάρτητος από την πυκνότητα ρ του μέσου και προέρχεται συνήθως από πειραματικά/θεωρητικά δεδομένα (π.χ. πίνακες NIST XCOM για στοιχεία και ενώσεις) .

3.3.6 Απόδοση και ανακατασκευή εικόνας

Η απόδοση ενός συστήματος ιατρικής απεικόνισης δεν εξαρτάται μόνο από την ποιότητα της παραγόμενης ακτινοβολίας, αλλά και από τον τρόπο με τον οποίο οι ανιχνευτές μετατρέπουν την προσπίπτουσα ενέργεια σε μετρήσιμο σήμα, καθώς και από τη διαδικασία ανακατασκευής της εικόνας από τις μετρήσεις αυτές [24],[38],[39]. Στην πράξη, κάθε σύστημα απεικόνισης χαρακτηρίζεται από συγκεκριμένα όρια χωρικής ανάλυσης, αντίθεσης, θορύβου και ευαισθησίας, τα οποία καθορίζουν το πόσο αξιόπιστα μπορεί να αναδείξει λεπτές ανατομικές ή παθολογικές δομές. Η έννοια της απόδοσης συνδέεται επομένως άμεσα με τη δυνατότητα διάκρισης ουσιαστικά διαφορετικών ιστών, αλλά και με την ελαχιστοποίηση της δόσης στον ασθενή χωρίς σημαντική απώλεια διαγνωστικής πληροφορίας [39].

Η ανακατασκευή εικόνας αποτελεί το μαθηματικό στάδιο κατά το οποίο τα ακατέργαστα δεδομένα των προβολών μετατρέπονται σε χωρική απεικόνιση του εξεταζόμενου αντικειμένου. Για την κλασική αξονική τομογραφία, το πρόβλημα ανάγεται στην εκτίμηση της κατανομής του γραμμικού συντελεστή εξασθένησης από ένα σύνολο προβολών υπό διαφορετικές γωνίες, κάτι που συνδέεται άμεσα με τον μετασχηματισμό Radon και τις αντίστροφες μεθόδους ανακατασκευής [32],[34]. Οι απλούστερες μέθοδοι, όπως η οπισθοπροβολή, συνοδεύονται από θόλωση, ενώ η φιλτραρισμένη οπισθοπροβολή και οι επαναληπτικές, στατιστικές προσεγγίσεις βελτιώνουν ουσιαστικά την ποιότητα της ανακατασκευής, ιδίως σε περιπτώσεις χαμηλής δόσης ή ατελών μετρήσεων [32],[34],[39],[40].

3.3.7 Υπολογιστική τομογραφία (CT)

Η υπολογιστική τομογραφία (Computed Tomography, CT) αποτελεί μία από τις σημαντικότερες εφαρμογές των ακτίνων X στην ιατρική, καθώς επιτρέπει τη δημιουργία τομών υψηλής χωρικής ανάλυσης από πολλαπλές λήψεις γύρω από τον ασθενή [29],[34]. Σε αντίθεση με τη συμβατική ακτινογραφία, η οποία παράγει διδιάστατη προβολή όλων των ανατομικών δομών πάνω στο ίδιο επίπεδο, η CT ανασυνθέτει τρισδιάστατη πληροφορία και έτσι μειώνει σημαντικά την επικάλυψη ιστών [24],[29]. Το αποτέλεσμα είναι η καλύτερη απεικόνιση τόσο οστικών δομών όσο και μαλακών ιστών, με δυνατότητα ποσοτικής εκτίμησης της εξασθένησης μέσω των μονάδων Hounsfield [34],[39].

Η λειτουργία της CT βασίζεται στη μέτρηση της εξασθένησης της δέσμης ακτίνων X για μεγάλο αριθμό γωνιών περιστροφής, ενώ η τελική εικόνα παράγεται μέσω αλγορίθμων ανακατασκευής που επιλύουν το αντίστροφο πρόβλημα της απεικόνισης [32],[34]. Στην κλινική πράξη, η ποιότητα της εικόνας επηρεάζεται από φαινόμενα όπως η σκλήρυνση δέσμης, η σκέδαση και ο θόρυβος των μετρήσεων, τα οποία μπορούν να προκαλέσουν παραμορφώσεις και σφάλματα στην εκτίμηση της πραγματικής κατανομής του συντελεστή εξασθένησης [31],[37],[40]. Για τον λόγο αυτό, η σύγχρονη CT αξιοποιεί τόσο τεχνικές διόρθωσης όσο και στατιστικές μεθόδους ανακατασκευής, οι οποίες βελτιώνουν την ακρίβεια και επιτρέπουν την αξιοποίηση χαμηλότερων δόσεων ακτινοβολίας [32],[39],[40].

3.3.8 Μαγνητικές (MRI)

Η μαγνητική τομογραφία (Magnetic Resonance Imaging, MRI) αποτελεί διαφορετική απεικονιστική τεχνική από την CT, καθώς δεν βασίζεται σε ιοντίζουσα ακτινοβολία αλλά στα φαινόμενα του πυρηνικού μαγνητικού συντονισμού [41]. Η μέθοδος χρησιμοποιεί ισχυρό στατικό μαγνητικό πεδίο και ραδιοσυχνότητες ώστε να διεγείρει τους πυρήνες υδρογόνου του σώματος, οι οποίοι στη συνέχεια εκπέμπουν σήμα κατά την επαναφορά τους στη θεμελιώδη κατάσταση [41]. Το λαμβανόμενο σήμα επεξεργάζεται μαθηματικά ώστε να παραχθούν εικόνες υψηλής αντίθεσης, ιδιαίτερα χρήσιμες στην απεικόνιση μαλακών ιστών, του νωτιαίου μυελού, των μεσοσπονδυλίων δίσκων και παθολογικών αλλοιώσεων των οστών και των παρακείμενων δομών [41].

Σε σχέση με τις ακτίνες X, η MRI προσφέρει το πλεονέκτημα της απουσίας ιοντίζουσας έκθεσης, γεγονός που την καθιστά ιδιαίτερα σημαντική σε εξετάσεις όπου απαιτείται επαναλαμβανόμενη απεικόνιση ή όπου η δόση ακτινοβολίας πρέπει να αποφεύγεται [41]. Παράλληλα, η εικόνα στη MRI εξαρτάται από φυσιολογικές παραμέτρους όπως ο χρόνος χαλάρωσης T1 και T2, η πυκνότητα πρωτονίων και οι επιλεγμένες ακολουθίες παλμών, παρέχοντας έτσι διαφορετικό τύπο αντίθεσης σε σύγκριση με την CT [41]. Ως εκ τούτου, οι δύο μέθοδοι δεν είναι ανταγωνιστικές αλλά συμπληρωματικές, με την CT να υπερέχει στην ταχεία και ποσοτική απεικόνιση της εξασθένησης και τη MRI να υπερτερεί στην ανάδειξη λεπτομερειών μαλακών ιστών και λειτουργικών χαρακτηριστικών [24],[29],[41].

Κεφάλαιο 4ο: Ανάλυση του πειράματος

Όπως παραδόξως, όλη αυτή η έμπνευση από το ανθρώπινο σώμα — που οδήγησε στην αναπαράσταση της βιολογίας του εγκεφάλου προκειμένου να κατασκευαστεί ένα τεχνητό νευρωνικό δίκτυο που λαμβάνει αποφάσεις — με επανενώνει με την πρώτη επιστήμη: αρχικώς αξιοποίησα τη βιολογία για να προχωρήσω τη γνώση στην πληροφορική· πλέον, αξιοποιείται η προχωρημένη γνώση στην πληροφορική για να επεξεργαστώ μεγάλα σύνολα δεδομένων με σύγχρονους αλγορίθμους και, τελικώς, να προωθήσω την επιστήμη που αποτέλεσε την αφετηρία των νευρωνικών δικτύων. Το παράδειγμα αυτό καταδεικνύει πόσο στενά συνδεδεμένες είναι οι επιστήμες μεταξύ τους: η μία τροφοδοτεί και επιταχύνει την πρόοδο της άλλης, και μέσα από αυτή την αμφίδρομη σχέση προκύπτει νέα γνώση.

Στο παρόν πείραμα η αρχική εστίαση σε μια συλλογή αποκλειστικά από CT-scans οδήγησε πολύ σύντομα σε αδιέξοδο, εξαιτίας της περιορισμένης διαθεσιμότητας και της δυσκολίας άμεσης αξιοποίησης επαρκούς αριθμού υγιών δεδομένων. Πέραν της πολύπλοκης θεωρίας που διέπει την παραγωγή ιατρικών απεικονίσεων, πρέπει να ληφθεί υπόψη και ο κίνδυνος για την υγεία του ασθενούς λόγω επαναλαμβανόμενης έκθεσης σε ιοντίζουσα ακτινοβολία. Ως αποτέλεσμα, στην παρούσα εργασία συμπεριλήφθηκαν επιπλέον MRI-scans, τα οποία αύξησαν τον όγκο και τα ποιοτικά χαρακτηριστικά των δεδομένων και βελτίωσαν την αντιπροσωπευτικότητα της μελέτης σε σχέση με τον πληθυσμό αναφοράς (βλ. σχετική βιβλιογραφία ιατρικής στατιστικής). Η ενσωμάτωση πολυτροπικών δεδομένων επέτρεψε την ανάπτυξη διαφορετικών εφαρμογών πάνω στο ίδιο συνελκτικό νευρωνικό δίκτυο, μεγιστοποιώντας την αξιοποίηση των διαθέσιμων εικόνων και παρέχοντας πλούσιες πληροφορίες για τη διερεύνηση πολλαπλών ερευνητικών ερωτημάτων, χωρίς να υπονομεύεται η αξιοπιστία και η εγκυρότητα των αποτελεσμάτων.

Πιο συγκεκριμένα, προκειμένου να αποφευχθεί η σύγχυση μεταξύ ετερογενών ανατομικών δομών και λαμβάνοντας υπόψη το περιορισμένο πλήθος των διαθέσιμων δεδομένων, καθώς και τη μειωμένη υπολογιστική ισχύ του συστήματος που χρησιμοποιήθηκε για την εκπαίδευση, επιλέχθηκε η ανεξάρτητη εκπαίδευση τρισδιάστατων συνελκτικών νευρωνικών δικτύων (3D CNN) για κάθε ανατομική περιοχή (πνεύμονες, εγκέφαλος, σπονδυλική στήλη), με στόχο τη δυαδική ταξινόμηση μεταξύ υγιούς και καρκινικού ιστού.

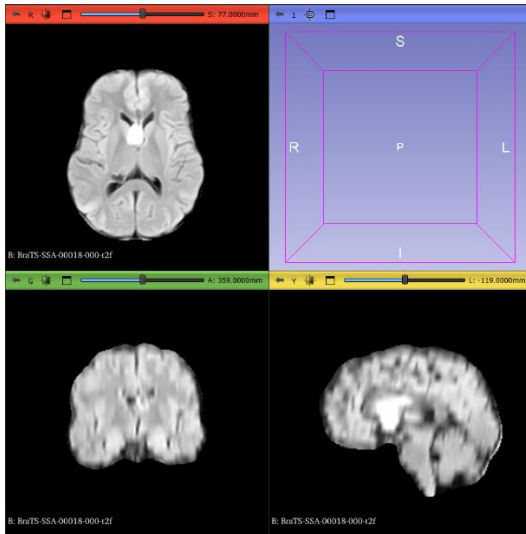
4.1 Σύνολο δεδομένων

Για την υλοποίηση της παρούσας εργασίας επιλέχθηκαν επτά δημόσια διαθέσιμα σύνολα δεδομένων ιατρικών απεικονίσεων (CT και MRI), τα οποία καλύπτουν τόσο περιπτώσεις κακοήθων νεοπλασιών όσο και υγιείς ασθενείς. Η επιλογή των συγκεκριμένων συνόλων έγινε με στόχο την κάλυψη διαφορετικών ανατομικών περιοχών, απεικονιστικών μεθόδων και πληθυσμιακών χαρακτηριστικών, καθώς και τη διερεύνηση της γενίκευσης των μοντέλων σε ετερογενή δεδομένα. Συγκεκριμένα, χρησιμοποιήθηκαν τα εξής:

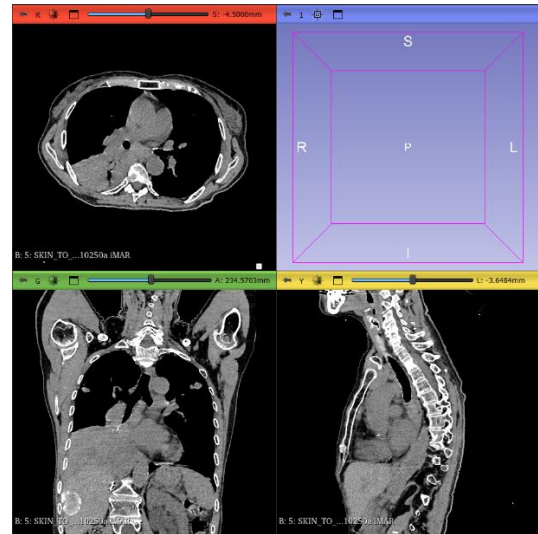
1. PKG – BraTS-Africa:

Πρόκειται για σύνολο δεδομένων μαγνητικών τομογραφιών (MRI) εγκεφάλου, το οποίο περιλαμβάνει ασθενείς με γλοιώματα και προέρχεται από αφρικανικούς πληθυσμούς. Το dataset έχει ως σκοπό τη συγκέντρωση νευρο-ογκολογικών απεικονίσεων από υποεκπροσωπούμενες γεωγραφικές και πληθυσμιακές ομάδες, παρουσιάζοντας σημαντική ποικιλομορφία ως προς φυλετικά, γεωγραφικά και τεχνολογικά χαρακτηριστικά, καθώς και ως προς τους διαφορετικούς σαρωτές και τα πρωτόκολλα

λήψης. Το μέγεθος του συνόλου ανέρχεται σε 146 ασθενείς, εκ των οποίων κρατήθηκαν οι 127, ενώ ως προς τις ετικέτες περιλαμβάνει 76 ασθενείς με γλοιώματα και 51 ασθενείς με άλλου τύπου ενδοκρανιακές νεοπλασματικές αλλοιώσεις του κεντρικού νευρικού συστήματος.



Εικόνα 4.1: PKG – BraTS-Africa MRI-Scan



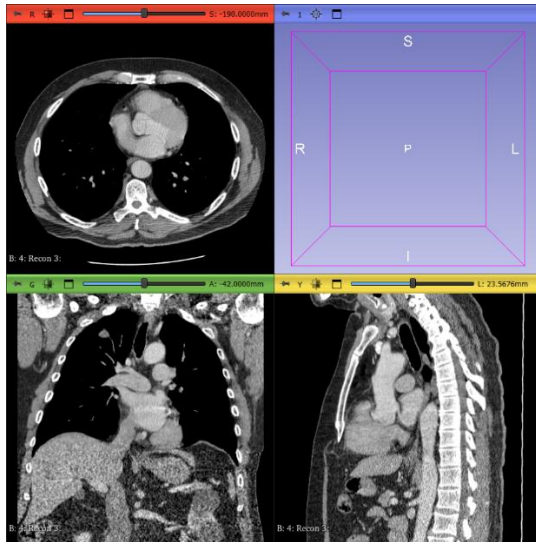
Εικόνα 4.2: Spine-Mets CT-Scan

2. Spine-Mets-CT-SEG:

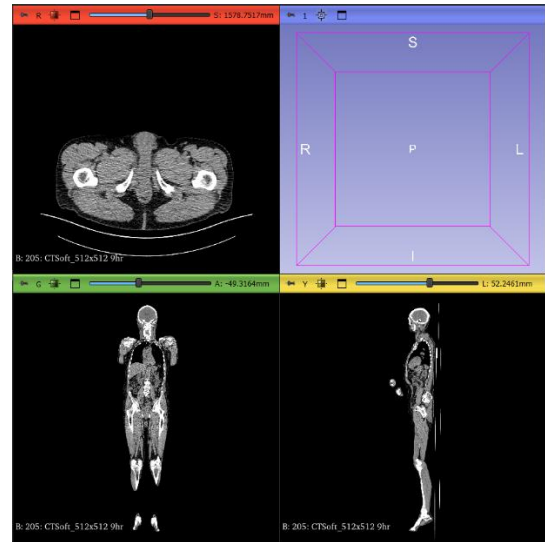
Πρόκειται για σύνολο δεδομένων αξονικών τομογραφιών (CT) της σπονδυλικής στήλης, το οποίο περιλαμβάνει μεταστατικές αλλοιώσεις και αφορά περιπτώσεις μεταστατικού καρκίνου της σπονδυλικής στήλης. Το dataset επικεντρώνεται σε παθολογίες που χαρακτηρίζονται από υψηλή ετερογένεια των βλαβών και συχνή συνύπαρξη μεταστάσεων σε άλλα όργανα. Το σύνολο αποτελείται από 55 ασθενείς, οι οποίοι συμπεριλαμβάνονται όλοι χωρίς εξαίρεση, ενώ οι ετικέτες αντιστοιχούν σε επισημασμένες μεταστατικές αλλοιώσεις.

3. LIDC-IDRI (manifest-1600709154662):

Πρόκειται για ευρέως χρησιμοποιούμενο σύνολο δεδομένων αξονικών τομογραφιών (CT) θώρακος, το οποίο αφορά την ανίχνευση και ανάλυση πνευμονικών όζων. Το dataset χαρακτηρίζεται από πολλαπλές, ανεξάρτητες επισημάνσεις από διαφορετικούς ακτινολόγους, στοιχείο που επιτρέπει τη μελέτη της δια-παρατηρητικής μεταβλητότητας και της αβεβαιότητας στις ετικέτες. Περιλαμβάνει περίπου 1000 ασθενείς, εκ των οποίων στο παρόν σύνολο έχουν επιλεγεί 26 ασθενείς με διαγνωσμένο καρκίνο στους πνεύμονες, ενώ οι ετικέτες αντιστοιχούν σε πολλαπλές ιατρικές επισημάνσεις για πνευμονικούς όζους.



Εικόνα 4.3: LIDC-IDRI CT-Scan



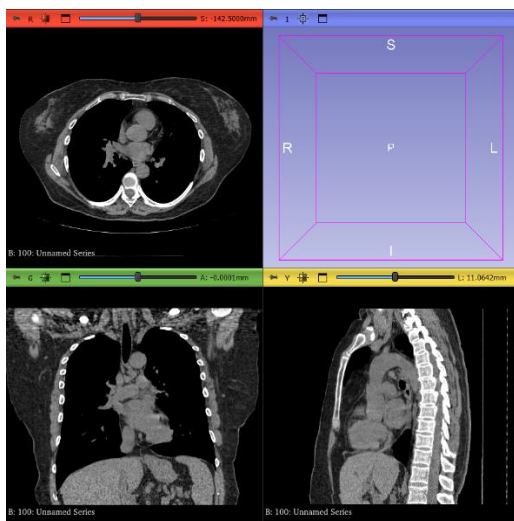
Εικόνα 4.4: Healthy-Total-Body CT-Scan

4. Healthy-Total-Body-CTs:

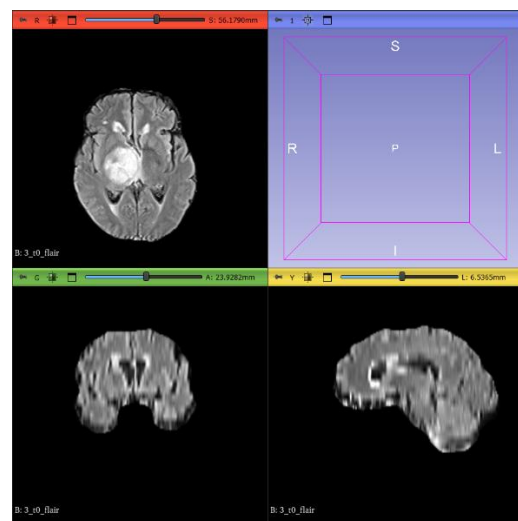
Πρόκειται για σύνολο δεδομένων ολικών αξονικών τομογραφιών (CT) υγιών εθελοντών, το οποίο αποτελεί ιδιαίτερα δυσεύρετη πηγή δεδομένων λόγω των ηθικών και ιατρικών περιορισμών που σχετίζονται με την έκθεση υγιών ατόμων σε ιοντίζουσα ακτινοβολία. Το μέγεθος του συνόλου ανέρχεται σε $N = 30$ ασθενείς και περιλαμβάνει απεικονίσεις χωρίς παθολογικά ευρήματα, αποτελώντας σύνολο υγιών δειγμάτων χωρίς επιμέρους κλινικές ετικέτες.

5. Lung_RIDER (Reference Image Database to Evaluate Response):

Πρόκειται για σύνολο δεδομένων αξονικών τομογραφιών (CT) πνευμόνων, το οποίο περιλαμβάνει επαναλαμβανόμενες απεικονίσεις των ίδιων ασθενών με διαφοροποιήσεις στις παραμέτρους λήψης και ανακατασκευής. Το dataset χρησιμοποιείται για τη μελέτη της επαναληψιμότητας των απεικονιστικών χαρακτηριστικών και της μεταβλητότητας που προκύπτει από διαφορετικά πρωτόκολλα CT. Αποτελείται από 32 ασθενείς με περιπτώσεις πνευμονικών όγκων και περιλαμβάνει διαχρονικές και επαναλαμβανόμενες εξετάσεις, χωρίς ρητά ορισμένα κλινικά labels στο παρόν πλαίσιο περιγραφής.



Εικόνα 4.5: Lung_RIDER CT-Scan



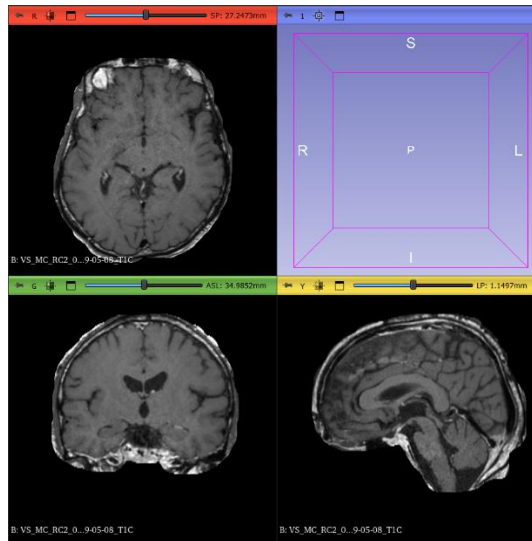
Εικόνα 4.6: CFB-GBM MRI-Scan

6. CFB-GBM (Centre François Baclesse – Glioblastoma Dataset):

Πρόκειται για σύνολο δεδομένων μαγνητικών τομογραφιών (MRI) εγκεφάλου ασθενών με γλοιοβλάστωμα, το οποίο προέρχεται από πραγματικές κλινικές ροές και χαρακτηρίζεται από υψηλή ετερογένεια ως προς τα πρωτόκολλα απεικόνισης και τις συνθήκες λήψης. Το dataset περιλαμβάνει 264 δείγματα και αντιστοιχεί σε κλινικά επιβεβαιωμένες περιπτώσεις γλοιοβλαστώματος.

7. Vestibular-Schwannoma:

Πρόκειται για σύνολο δεδομένων μαγνητικών τομογραφιών (MRI) ασθενών με αιθουσαίο σβάννωμα, συχνά οργανωμένο σε διαχρονική (longitudinal) μορφή, γεγονός που επιτρέπει την καταγραφή επαναλαμβανόμενων εξετάσεων για τους ίδιους ασθενείς. Το dataset αφορά την περιοχή του έσω ακουστικού πόρου και της εγκεφαλοπαραεγκεφαλιδικής γωνίας και περιλαμβάνει απεικονίσεις καλοήθων όγκων. Οι ετικέτες αντιστοιχούν στη διάγνωση αιθουσαίου σβαννώματος, ενώ το σύνολο των 242 δειγμάτων χαρακτηρίζεται από υψηλή σπανιότητα λόγω της διαχρονικής φύσης των δεδομένων.



Εικόνα 4.7: Vestibular-Schwannoma MRI-Scan

4.2 Επιλογή αρχείων

Στη φάση προεπεξεργασίας εφαρμόστηκε στοχευμένη «διαλογή» των εισερχόμενων εικόνων με σκοπό την ελαχιστοποίηση της ετερογένειας που θα μπορούσε να παραπλανήσει το δίκτυο κατά την εκπαίδευση. Συγκεκριμένα, προτιμήθηκαν αρχεία και πρωτόκολλα των οποίων οι χωρικές και εντασιακές ιδιότητες συμπίπτουν με το convolution-kernel `b_soft_a` του συνόλου `Healthy-Total-Body` και με το «standard» convolution kernel. Στην περίπτωση των MRI, επιλέχθηκαν σειρές που δίνουν τιμές εντασης και ανατομική απεικόνιση πλησιέστερη στις HU των CT (π.χ. `6_t2_t1gd.nii.gz`, `VS_MC_RC2_013_1999-02-26_T1C.nii.gz`), ώστε να ελαττωθεί η απόκλιση στην «φιλοσοφία» των αναπαριστώμενων pixels μεταξύ modalities.

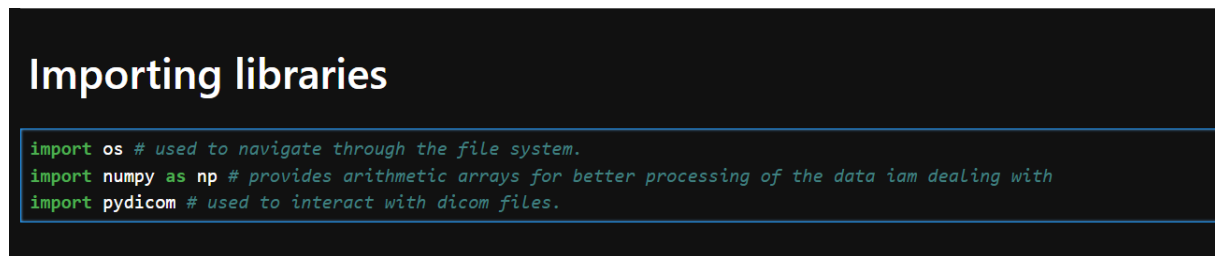
Ο σκοπός αυτής της προσέγγισης ήταν, πρώτον να εξασφαλιστεί ότι τα δεδομένα που τροφοδοτούν το μοντέλο «συγκλίνουν» ως προς την κλίμακα και την κατανομή εντάσεων (όχι να περιέχουν ακραία, μη συγκρίσιμα patterns), και δεύτερον, να αποφευχθεί η εκπαίδευση σε χαρακτηριστικά που προκύπτουν

αποκλειστικά από το πρωτόκολλο απεικόνισης ή από ειδικούς kernel effects. Για παράδειγμα, σειρές όπως το FLAIR εμφανίζουν συστηματικά υψηλότερες φωτεινότητες σε ορισμένα ανατομικά στοιχεία· αντίστοιχα, κάποιοι CT-kernel (π.χ. lung kernel) τονίζουν πολύ την εμφάνιση του όγκου σε σχέση με το standard kernel. Η ανάμειξη τέτοιων πηγών χωρίς φροντισμένο φιλτράρισμα θα μπορούσε να οδηγήσει το CNN στο να μάθει proxy-features (scanner/sequence artifacts) αντί για κλινικά ουσιώδη μοτίβα.

Συνεπώς, με την επιλεκτική διατήρηση μόνο των σειρών που ικανοποιούν τα παραπάνω κριτήρια (χωρική ανάλυση, κατανομή εντάσεων, συμβατότητα με τα HU-like χαρακτηριστικά όπου αυτό ήταν επιθυμητό), μειώθηκε ο domain shift μεταξύ δειγμάτων και αυξήθηκε η πιθανότητα το δίκτυο να μάθει γενικεύσιμα, ανατομικά και παθολογικά χαρακτηριστικά. Αυτή η προσέγγιση δεν αντικαθιστά την ανάγκη για περαιτέρω normalization και harmonization (π.χ. windowing, rescaling, standardization), αλλά λειτουργεί ως προληπτικό βήμα ώστε τα δεδομένα εκκίνησης να είναι πιο ομοιογενή και κατά συνέπεια η εκπαίδευση πιο αξιόπιστη.

4.3 Προεπεξεργασία των CT-scans

Η υλοποίηση βασίστηκε κυρίως στις βιβλιοθήκες os, numpy και pydicom της Εικόνας 4.8, οι οποίες επέτρεψαν την οργάνωση της δομής των αρχείων, την αριθμητική επεξεργασία των απεικονιστικών δεδομένων και την ανάγνωση των DICOM αρχείων αντίστοιχα. Η συνάρτηση `split_total_body_by_slice_stats()` αποτελεί το βασικό στάδιο προεπεξεργασίας των ολικών υγιών αξονικών τομογραφιών και στοχεύει στον εντοπισμό ανατομικά συνεκτικών τμημάτων μέσα σε κάθε εξέταση, ώστε να εξαχθούν επιμέρους υποσύνολα τομών που αντιστοιχούν σε συγκεκριμένες περιοχές ενδιαφέροντος. Η λειτουργία αυτή εφαρμόζεται αποκλειστικά σε CT δεδομένα, καθώς βασίζεται σε τιμές Hounsfield Units (HU), οι οποίες αποτυπώνουν τις ιδιότητες εξασθένησης της ακτινοβολίας στους ιστούς.



Εικόνα 4.8: Βασικές βιβλιοθήκες

Αρχικά πραγματοποιείται ο έλεγχος και η οργάνωση της ιεραρχίας των φακέλων εισόδου και εξόδου, ώστε η επεξεργασία να ανταποκρίνεται στο file system που παρουσιάζεται στην Εικόνα 4.9 . Στη συνέχεια, τα DICOM αρχεία κάθε σειράς ταξινομούνται με βάση τη χωρική τους θέση στον άξονα z, αξιοποιώντας τα διαθέσιμα μεταδεδομένα, ώστε να διατηρείται η σωστή ανατομική ακολουθία των τομών. Η διαδικασία αυτή είναι σημαντική, καθώς επιτρέπει τη διατήρηση της χωρικής συνοχής του όγκου πριν από οποιαδήποτε περαιτέρω ανάλυση.

```

# Saving the actual paths to each dcm file to the file_paths variable :
file_paths = [os.path.join(series_path, f) for f in dcm_files]
# Process of collecting the dcm files ends .

#INPUT
#root_dir/
#├── non_cancer/
#   ├── Healthy-Total-Body-CTs/
#   │   ├── Patient_01/
#   │   │   ├── Series_A/
#   │   │   │   ├── slice001.dcm
#   │   │   │   ├── slice002.dcm
#   │   │   │   └── slice003.dcm
#   │   └── ...
#   └── ...
# read headers first to get z positions (fast)
headers = []

```

Εικόνα 4.9: file system στο οποίο ανταποκρίνεται ο κώδικας

Ακολουθεί η μετατροπή των αρχικών τιμών των pixels σε μονάδες HU μέσω των αντίστοιχων συντελεστών ανακλιμάκωσης ($HU = \text{pixel_value} \times \text{slope} + \text{intercept}$.) του DICOM αρχείου. Το βήμα αυτό είναι ουσιώδες για τις CT εικόνες, επειδή τα HU αποτελούν το φυσικό μέτρο με το οποίο διαφοροποιούνται οι ιστοί ως προς την πυκνότητα και την ακτινολογική τους συμπεριφορά. Πάνω σε αυτή τη βάση εφαρμόζονται προκαθορισμένα παράθυρα τιμών που εμφανίζονται ως παράμετροι στην Εικόνα 4.10 για τρεις ανατομικές κατηγορίες: πνευμονικός ιστός, μαλακός ιστός/εγκέφαλος και οστό. Τα όρια αυτά δεν είναι αυθαίρετα, αλλά αντανακλούν τη γνωστή συμπεριφορά των ιστών στις αξονικές τομογραφίες, όπου ο πνεύμονας εμφανίζει χαμηλές και αρνητικές HU λόγω της υψηλής περιεκτικότητας σε αέρα, το οστό υψηλές HU λόγω αυξημένης απορρόφησης ακτίνων X, ενώ ο μαλακός ιστός καταλαμβάνει ενδιάμεσο εύρος τιμών.

```

# 4.)
def split_total_body_by_slice_stats(root_dir,
    output_root,
    main_folder_name='non_cancer',
    subfolder_name='Healthy-Total-Body-CTs',
    organs=('lung', 'brain', 'spine'),
    # The HU values that represent a certain organ :
    lung_hu_window=(-1000, -200), # it contains a lot of air so the hu will be low
    brain_hu_window=(15, 60), # it contains soft-tissue so the hu will be close to 0
    bone_hu_thresh=170, # it absorbs a lot of x-rays so the hu will be the highest of all.
    # What percentage is gonna persuade me into thinking that a slice looks like a
    # lung or a brain or a bone correspondingly :
    lung_frac_thresh=0.30, # i want the percentage of the slice to be lung-like atleast by 30% to be acceptable
    brain_frac_thresh=0.30, # i want the percentage of the slice to be brain-like atleast by 30% to be acceptable
    bone_frac_thresh=0.10, # i want the percentage of the slice to be bone-like atleast by 10% to be acceptable
    # What iam willing to accept regarding the blocks which are gonna be created :
    min_block_len=25, # minimum sequence of slices that can be recognised as an undivided block
    gap_tolerance=5, # I will merge smaller blocks with a certain tolerance
    # for instance : True - True - False - False - True - True , with gap_tolerance=2 this will never become a block.
    verbose=True):

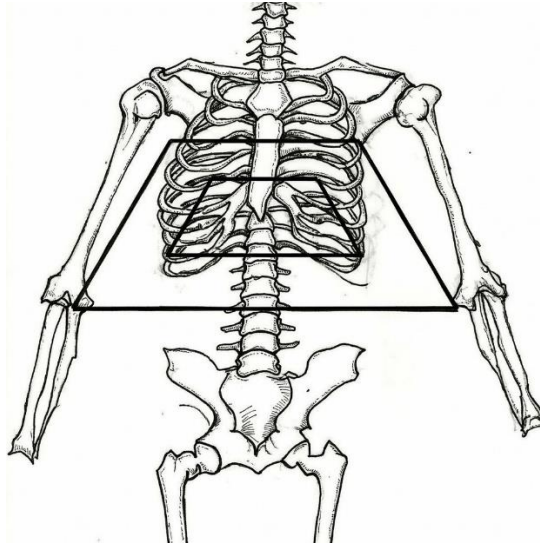
```

Εικόνα 4.10: Παράμετροι συνάρτησης διαχωρισμού μιας αξονικής σε υποσύνολα

Για κάθε τομή εκτιμάται το ποσοστό των pixels (π.χ `lung_frac_thresh=0.30`) που εμπίπτουν στα αντίστοιχα HU κριτήρια, με στόχο να προσδιοριστεί αν η τομή παρουσιάζει κυρίαρχα χαρακτηριστικά μιας συγκεκριμένης ανατομικής περιοχής. Η επιλογή των κατωφλίων βασίζεται σε δύο παραμέτρους: το εύρος HU που ορίζει τη σχετική «ταυτότητα» κάθε ιστού και το ελάχιστο ποσοστό εμφάνισης που απαιτείται ώστε η τομή να θεωρηθεί αντιπροσωπευτική της συγκεκριμένης περιοχής. Αυτή η λογική παρέχει ευελιξία: μεγαλύτερα thresholds απαιτούν πιο «καθαρό» σήμα για αναγνώριση, μικρότερα thresholds αυξάνουν την ευαισθησία αλλά και τον κίνδυνο ψευδώς θετικών.

Στο στάδιο αυτό, από κάθε τομή των CT εξετάσεων εξάγονται τρία συμπληρωματικά ανατομικά κριτήρια, με σκοπό τον εντοπισμό περιοχών που αντιστοιχούν στη σπονδυλική στήλη, στους πνεύμονες και στον εγκέφαλο. Για τη σπονδυλική περιοχή δίνεται έμφαση στο κεντρικό τμήμα της εικόνας, ώστε να αναγνωρίζονται οστικές δομές και να αποφεύγεται η επιρροή περιφερικών οστικών περιοχών που

δεν σχετίζονται με το ζητούμενο ανατομικό πεδίο. Η εστίαση αυτή αναδεικνύεται στο Σχήμα 4.1. Αντίστοιχα, για τους πνεύμονες εξετάζονται οι αμφοτερόπλευρες περιοχές της τομής, έτσι ώστε να απομονώνεται η πνευμονική μορφολογία και να αποφεύγονται κεντρικές περιοχές που περιέχουν αέρα και μπορεί ψευδώς να θεωρηθούν ως πνεύμονες. Για τον εγκέφαλο, τέλος, επιλέγεται το άνω-κεντρικό τμήμα της τομής, καθώς εκεί αναμένεται να εμφανιστεί πρώιμα και πιο συνεκτικά ο εγκεφαλικός μαλακός ιστός σε κεφαλικές αξονικές απεικονίσεις.



Σχήμα 4.2: Οπτικοποίηση της έμφαση στο κεντρικό τμήμα μιας τομής

Η αναγνώριση κάθε περιοχής βασίζεται στη σύγκριση των τιμών HU με προκαθορισμένα όρια και στην εκτίμηση του ποσοστού των pixels που ικανοποιούν τα αντίστοιχα κριτήρια που εμφανίζονται ως παράμετροι στην Εικόνα 4.10. Με αυτόν τον τρόπο, κάθε τομή αξιολογείται ως προς το κατά πόσο παρουσιάζει χαρακτηριστικά οστού, πνευμονικού ιστού ή εγκεφαλικού/μαλακού ιστού, και η συνεχής αυτή εκτίμηση που αποθηκεύεται στην δομή δεδομένων της Εικόνας 4.11, μετατρέπεται στη συνέχεια σε δυαδική απόφαση για κάθε ανατομική κατηγορία. Το αποτέλεσμα είναι η δημιουργία σταθερών και ανατομικά συνεκτικών masks (True-True-False-False-True-True), τα οποία καθορίζουν ποιες τομές ανήκουν σε κάθε περιοχή ενδιαφέροντος και αποτελούν τη βάση για το επόμενο στάδιο της επεξεργασίας.

```
lung_frac.append(lf) # here i'll be storing ALL the different M.O percentages that tell me what percentage of each
# slice looks like lungs.
#Lung_frac = [
#0.00, # slice 0
#0.02, # slice 1
#0.18, # slice 2
#0.21, # slice 3 # this one is the most lung-like slice for instance.
#0.16, # slice 4
#0.01 # slice 5
# ]
```

Εικόνα 4.31: Δομή δεδομένων που περιέχει ποσοστά ομοιότητας ενός slice με δεδομένο όργανο

Στη συνέχεια, τα διαδοχικά θετικά ευρήματα ομαδοποιούνται σε συνεχή blocks, με χρήση δύο βασικών κριτηρίων: του ελάχιστου μήκους block και της ανεκτικότητας σε μικρά κενά μεταξύ διαδοχικών τομών όπου και αυτά δίνονται ως παράμετροι στην Εικόνα 4.10. Η λογική αυτή επιτρέπει την απόρριψη μικρών, θορυβωδών ή ανατομικά ασταθών τμημάτων και τη διατήρηση μεγαλύτερων, συνεκτικών ακολουθιών που αντιστοιχούν πιο αξιόπιστα στην εκάστοτε περιοχή ενδιαφέροντος. Ειδικά για το

οστικό τμήμα, προτιμάται η επιλογή του επικρατέστερου block, εφόσον αυτό έχει επαρκές μήκος, ώστε να απομονώνεται με μεγαλύτερη ακρίβεια η σπονδυλική περιοχή. Αντίστοιχα, για τον πνευμονικό ιστό επιλέγεται το block που βρίσκεται χωρικά πλησιέστερα στη σπονδυλική στήλη, ενώ για τον εγκέφαλο εφαρμόζεται λογική ανίχνευσης του πρώτου συνεχούς -και κεντρικού- τμήματος μαλακού ιστού μιας και το πρώτο ανατομικά συνεκτικό block που προκύπτει είναι ο εγκέφαλος σε head-first scans.

Τέλος, τα επιλεγμένα blocks αντιστοιχίζονται σε ξεχωριστές ιεραρχίες φακέλων όπως αυτή της Εικόνας 4.12, και οι αντίστοιχες τομές αποθηκεύονται οργανωμένα σε αυτές, ώστε να προκύπτουν διακριτά υποσύνολα για κάθε ανατομική περιοχή. Η διαδικασία αυτή παράγει τρία επιμέρους σύνολα τομών, ένα για τους πνεύμονες, ένα για τη σπονδυλική στήλη και ένα για τον εγκέφαλο, διατηρώντας τη δομή των δεδομένων καθαρή, αναγνώσιμη και κατάλληλη για τα επόμενα στάδια της μεθοδολογίας .

```
# Now i begin the process of creating a file structure system for the output as a final task for my function :
blocks = {'lung': lung_blocks, 'spine': bone_blocks, 'brain': brain_blocks}

# copy files for each organ block
for organ_name, organ_blocks in blocks.items():
    for (s_idx, e_idx) in organ_blocks:
        # Creating an output folder :
        # CHANGED: output dir mapping to e.g. Healthy-Total-Body-CTs_Lungs inside the existing non_cancer (output_root)
        organ_folder = organ_output_map.get(organ_name, f"{subfolder_name}_{organ_name}")
        dst_dir = os.path.join(output_root, organ_folder, patient, series)
        os.makedirs(dst_dir, exist_ok=True)

#OUTPUT
#output_root/
#├── non_cancer/
#│   └── Lung/
#│       └── Patient_01/
#│           └── Series_A/
#│               ├── slice010.dcm
#│               ├── slice011.dcm
#│               └── slice012.dcm
#
```

Εικόνα 4.12: file system που δημιουργεί η συνάρτηση διαχωρισμού

Με τον τρόπο αυτό, το αρχικό σύνολο των 30 υγιών ασθενών κατακερματίστηκε σε τρία υποσύνολα (ένα ανά όργανο), επιτρέποντας την εκπαίδευση τριών ανεξάρτητων μοντέλων – ουσιαστικά την επανάληψη του πειράματος τρεις φορές με διαφορετικά δεδομένα, ενισχύοντας την εγκυρότητα των αποτελεσμάτων.

Η δεύτερη βασική λειτουργικότητα αφορά τη συνάρτηση `load_dicom_files`, η οποία είναι υπεύθυνη για τη φόρτωση των DICOM αρχείων και τη μετατροπή τους σε τρισδιάστατους όγκους με ομοιόμορφη μορφή, ώστε να είναι κατάλληλοι για περαιτέρω ανάλυση. Η συνάρτηση διατηρεί παράλληλα και την πληροφορία της κλάσης κάθε όγκου, διασφαλίζοντας ότι κάθε 3D volume αντιστοιχεί σωστά είτε σε καρκινικό είτε σε μη καρκινικό δείγμα. Η αντιστοίχιση αυτή αποθηκεύεται σε ξεχωριστή δομή με κοινή αρίθμηση, ώστε να εξασφαλίζεται η πλήρης συμφωνία μεταξύ όγκου και ετικέτας. Με τον τρόπο αυτό, η χρήση του ίδιου index και στις δύο δομές εξασφαλίζει ότι γίνεται αναφορά στην ίδια αξονική εξέταση.

Η διαδικασία ξεκινά από τη βασική ιεραρχία φακέλων του dataset, στην οποία κάθε κλάση αντιστοιχεί σε συγκεκριμένο υποφάκελο. Με αυτόν τον τρόπο, το σύστημα, δουλεύει με ένα συγκεκριμένο πρότυπο ιεραρχίας φακέλων, και οργανώνει συστηματικά τα δεδομένα ανά κατηγορία και αναζητά τις επιμέρους εξετάσεις των ασθενών μέσα στη δομή αποθήκευσης. Κατά τη φόρτωση κάθε σειράς, αξιολογείται η εγκυρότητα του DICOM υλικού και συλλέγονται τα απαραίτητα μεταδεδομένα μαζί με τις απεικονιστικές τιμές των pixels. Η προσέγγιση αυτή επιτρέπει την κατασκευή ενός συνεκτικού συνόλου από 2D τομές, οι οποίες στη συνέχεια συντίθενται σε ενιαίο 3D volume.

Αφού ολοκληρωθεί η συλλογή των τομών μιας εξέτασης, ελέγχεται η ομοιομορφία των διαστάσεων ώστε οι επιμέρους εικόνες να μπορούν να στοιβαχθούν σε τρισδιάστατη μορφή χωρίς ασυμφωνίες στο

σχήμα. Στη συνέχεια, πραγματοποιείται ένας βασικός ποιοτικός έλεγχος ως προς την κλίμακα των τιμών, προκειμένου να διαπιστωθεί αν τα δεδομένα βρίσκονται ήδη σε κλίμακα HU ή αν απαιτείται μετατροπή από τις αρχικές τιμές των pixels. Η απόφαση αυτή βασίζεται στη γενικότερη κατανομή των τιμών φωτεινότητας του όγκου και στην παρουσία αρνητικών τιμών, οι οποίες αποτελούν ισχυρή ένδειξη ότι τα δεδομένα έχουν ήδη μετασηματιστεί σε Hounsfield Units. Εάν αυτό δεν ισχύει, εφαρμόζεται η κατάλληλη γραμμική μετατροπή με βάση τους συντελεστές ανακλιμάκωσης του DICOM αρχείου που φαίνεται στην Εικόνα 4.13, ώστε όλα τα volumes να εκφράζονται σε ενιαία, φυσικά ερμηνεύσιμη κλίμακα.

```

"""Start Turning Raw pixel values to HU :"""
# A better way to do it :
neg_count = np.sum(volume < 0)
total_voxels = volume.size
neg_ratio = neg_count / total_voxels
print(f"Negative voxels: {neg_count} ({neg_ratio:.4f})")

volume_min = volume.min()
print(f"Volume before HU conversion - Min: {volume_min:.2f}, Max: {volume.max():.2f}")
if neg_ratio>0.02:
    print("Min value < 0 -> assuming volume already in HU. Skipping HU conversion.")
else:
    # ensure we operate in float to avoid uint16 clipping :
    volume = volume.astype(np.float32, copy=False)
    depth=volume.shape[0]
    if len(slopes) != depth:
        print("Warning: slopes length != depth. Falling back to single slope/intercept if available.")
        common_slope = slopes[0] if slopes else 1.0
        common_inter = intercepts[0] if intercepts else 0.0
        volume = volume.astype(np.float32) * float(common_slope) + float(common_inter)
    else:
        # Looping through the depth axis :
        for i in range(depth):
            s = float(slopes[i])
            inter = float(intercepts[i])
            volume[i] = volume[i].astype(np.float32) * s + inter
    print(f"HU conversion applied. Volume after conversion - Min: {volume.min():.2f}, Max: {volume.max():.2f}")
    print(f"3D volume created for {patient_folder}, shape: {volume.shape}")
    print(f"Volume - Min: {volume.min()}, Max: {volume.max()}, Sample (center slice, 5x5):\n {volume[volume.shape[0]//2, slice_shape[0]//2:
    """End of turning raw pixel values to HU. """

```

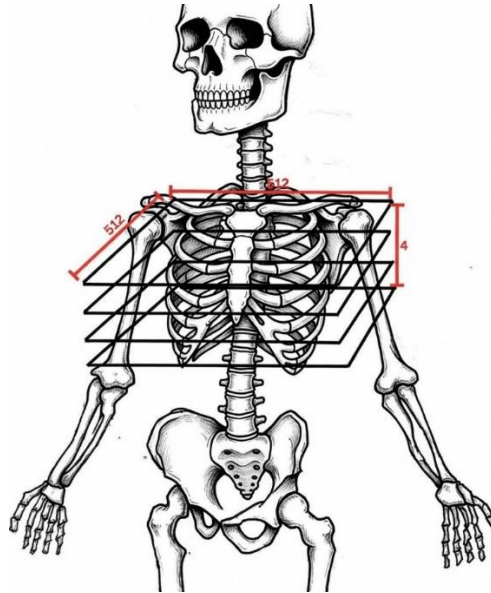
Εικόνα 4.13: Κώδικας μετατροπής τιμών φωτεινότητας των pixels σε HU

Στη συνέχεια εφαρμόζεται επαναδειγματοληψία στον άξονα του βάθους, όταν ο αριθμός των slices διαφέρει από το προκαθορισμένο επιθυμητό βάθος (depth = 32), με στόχο την ομογενοποίηση της τρισδιάστατης δομής όλων των όγκων και τη διασφάλιση ότι κάθε δείγμα έχει κοινή αναπαράσταση εισόδου. Η διαδικασία αυτή υλοποιείται με γραμμική παρεμβολή κατά μήκος του άξονα του βάθους, ώστε να αφαιρούνται περιττά ή πλεονάζοντα slices χωρίς να χάνεται η ουσία της ανατομικής πληροφορίας και χωρίς να προκαλούνται απότομες αλλοιώσεις στην απεικόνιση. Πιο συγκεκριμένα, για κάθε θέση του δισδιάστατου επιπέδου λαμβάνεται η ακολουθία τιμών κατά μήκος όλων των αρχικών slices και αυτή επαναδειγματοληπτείται σε έναν νέο, προκαθορισμένο αριθμό επιπέδων, διατηρώντας τη χωρική συνέχεια των εντάσεων και συμπυκνώνοντας με συνεπή τρόπο την πληροφορία της αρχικής εξέτασης σε ενιαίο βάθος. Το αποτέλεσμα είναι ένα νέο 3D-volume με το ίδιο χωρικό μέγεθος (H×W) αλλά με διαφορετικό βάθος, όπου κάθε voxel στον νέο όγκο αποτελεί μια συνεκτική, αριθμητικά συμβατή «περίληψη» της αρχικής κατανομής εντάσεων στην αντίστοιχη κάθετη στήλη του αρχικού όγκου.

Ακολούθως, όπου απαιτείται, εφαρμόζεται και χωρική επαναδειγματοληψία στους άξονες ύψους και πλάτους, έτσι ώστε κάθε slice να έχει σταθερές διαστάσεις(256x256). Η επιλογή αυτή εξυπηρετεί την ανάγκη για ομοιομορφία μεταξύ των εξετάσεων και επιτρέπει τη δημιουργία όγκων με κοινό σχήμα, ανεξάρτητα από τις επιμέρους διαφορές στη διαδικασία λήψης. Η συγκεκριμένη προσαρμογή διασφαλίζει ότι οι όγκοι εισάγονται σε ενιαία μορφή και μπορούν να χρησιμοποιηθούν συνεκτικά στα επόμενα στάδια της ανάλυσης.

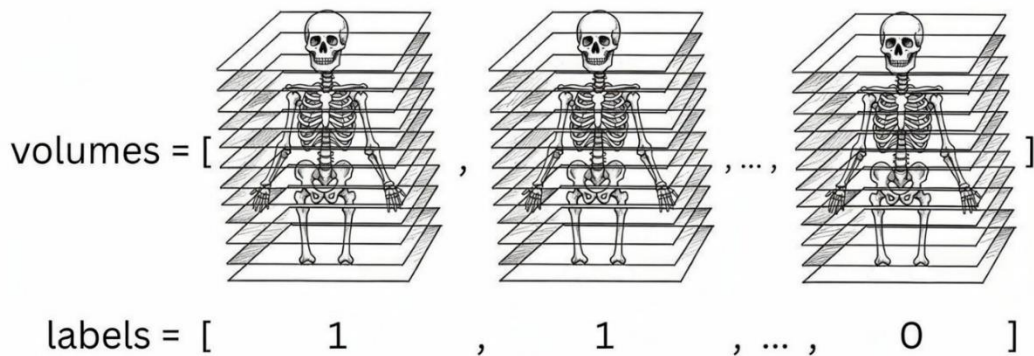
Η προσέγγιση αυτή έχει δύο βασικά πλεονεκτήματα: διατηρεί την ανατομική συνέπεια κάθε χωρικής θέσης (γιατί η επαναδειγματοληψία γίνεται ανεξάρτητα για κάθε pixel) και αποφεύγει απότομες

αλλοιώσεις που θα προέκυπταν με απλό downsampling χωρίς interpolation. Από την άλλη, η γραμμική παρέμβαση εισάγει κάποιο smoothing κατά μήκος του βάθους και μπορεί να αμβλύνει πολύ λεπτές, απότομες δομές· εναλλακτικές στρατηγικές (π.χ. μέσοι όροι, μέγιστα, ή πιο εξελιγμένες τεχνικές με σεβασμό στο voxel spacing) μπορούν να προταθούν αν χρειαστεί πιο «σκληρή» διατήρηση μικροδομών. Στην παρούσα εργασία η γραμμική επαναδειγματοληψία κρίθηκε ως ασφαλής μέθοδος για την προετοιμασία των όγκων προς εισαγωγή στο δίκτυο.



Σχήμα 4.2: Οπτικοποίηση της επαναδειγματοληψίας μιας αξονικής

Τέλος, οι όγκοι και οι αντίστοιχες ετικέτες τους συγκεντρώνονται σε τελικές δομές δεδομένων όπως στο Σχήμα 4.3, και επιστρέφονται ως ενιαίο σύνολο, ώστε κάθε τρισδιάστατη απεικόνιση να συνδέεται με τη σωστή κλάση και η διαδικασία να παραμένει οργανωμένη και συμβατή με τα επόμενα στάδια επεξεργασίας. Σε περιπτώσεις όπου το dataset περιλαμβάνει πολλαπλές εξετάσεις για τον ίδιο ασθενή σε διαφορετικές χρονικές στιγμές, η ονοματοδοσία στους φακέλους του file system προσαρμόζεται κατάλληλα ώστε να διατηρείται η συμβατότητα με τον τρόπο που εκτελείται ο κώδικας, όπως για παράδειγμα μέσω επισημάνσεων τύπου patient_1.1, patient_1.2 κ.ο.κ.



Σχήμα 4.3: Οπτικοποίηση της αποθήκευσης των 3d volumes

Η επόμενη λειτουργικότητα αφορά την ομαδοποιημένη προεπεξεργασία των τρισδιάστατων CT όγκων που προέκυψαν από τα DICOM αρχεία. Η συνάρτηση `preprocess_CT_volumes` εφαρμόζεται στο σύνολο των όγκων και επιστρέφει τα προεπεξεργασμένα δεδομένα σε ενιαία δομή παρόμοια με αυτή του Σχήματος 4.3, διατηρώντας τη σειρά εισόδου ώστε να παραμένει η αντιστοιχία με τυχόν εξωτερικές ετικέτες ή μεταδεδομένα. Στο πλαίσιο της παρούσας υλοποίησης, η προτεραιότητα δίνεται στη ρύθμιση της κλίμακας των CT που αφορούν εγκεφαλικές 3d αναπαραστάσεις, ώστε να καταστούν πιο συμβατές με την ανάλυση μαλακών ιστών, ειδικά σε περιπτώσεις όπου απαιτείται συγκριτική επεξεργασία με MRI δεδομένα.

Για κάθε όγκο, η διαδικασία ξεκινά με έλεγχο της τρισδιάστατης μορφής του, καθώς η συνάρτηση αναμένει είσοδο σε μορφή `depth–height–width`. Στη συνέχεια εφαρμόζεται περιορισμός της δυναμικής περιοχής των τιμών HU, όταν αυτό είναι απαραίτητο, δηλαδή για 3d volumes εγκεφάλων, ώστε να απομακρυνθούν ακραίες εντάσεις που αντιστοιχούν κυρίως σε οστικές δομές και δεν εξυπηρετούν την ανάδειξη του `soft tissue`. Η επιλογή των ορίων `clipping` βασίζεται στη γνωστή ακτινολογική συμπεριφορά των CT απεικονίσεων, όπου οι πολύ υψηλές HU τιμές αντιστοιχούν σε ιστούς με αυξημένη πυκνότητα και έντονη απορρόφηση ακτινοβολίας. Έτσι τα -προερχόμενα από CT-scans - 3d volumes ανάγονται σε μια κοινή κλίμακα τιμών με αυτά που θα προκύψουν από τα MRI scans.

Ακολούθως, οι τιμές μετατρέπονται σε μορφή κατάλληλη για αριθμητική επεξεργασία και εφαρμόζεται γραμμική κανονικοποίηση στην κλίμακα $[0, 1]$. Η κανονικοποίηση αυτή έχει ως στόχο την ομογενοποίηση των τιμών μεταξύ διαφορετικών όγκων, ώστε να μειωθεί η επίδραση των διαφορών έντασης και να διασφαλιστεί συγκρίσιμη συμπεριφορά κατά τα επόμενα στάδια του pipeline. Μετά την επεξεργασία κάθε όγκου, τα αποτελέσματα συλλέγονται σε ενιαία δομή και τελικά επιστρέφονται ως τετραδιάστατος πίνακας $(N, \text{depth}, \text{width}, \text{height})$, έτοιμος για τα επόμενα βήματα της μεθοδολογίας, όπως η ενίσχυση δεδομένων και η είσοδος στο μοντέλο.

4.4 Προεπεξεργασία των MRI-scans

Τα data όπως έχει ήδη αναφερθεί, εξακολουθούν να μην επαρκούν, κάτι το οποίο οδήγησε στην εξέταση της συνδυασμένης χρήση CT-scan και MRI-scan για ένα από τα τρία CNN, αλλά παράλληλα προέκυψε ένα νέο αδιέξοδο, μιας και στην προσπάθειά αύξησης του data set προέκυψε μια σημαντική δυσκολία, η οποία σχετίζεται με την ετερογένεια των modalities των συνόλων δεδομένων. Ενώ το σύνολο `Healthy-Total-Body-CTs` παρέχει αξονικές τομογραφίες CT υγιών ατόμων, τα σύνολα για όγκους εγκεφάλου (`PKG - BraTS-Africa`, `CFB-GBM`, `Vestibular-Schwannoma-MC-RC2`) βασίζονται αποκλειστικά σε MRI απεικονίσεις. Αυτή η ασυμβατότητα αντιβαίνει στην αρχική φιλοσοφία του πειράματος, που προέβλεπε ομοιογένεια modality (CT για όλες τις κατηγορίες), και δημιουργεί κίνδυνο `domain shift`, όπου το μοντέλο μπορεί να μάθει χαρακτηριστικά συγκεκριμένου τύπου απεικόνισης αντί πραγματικών παθολογικών patterns.

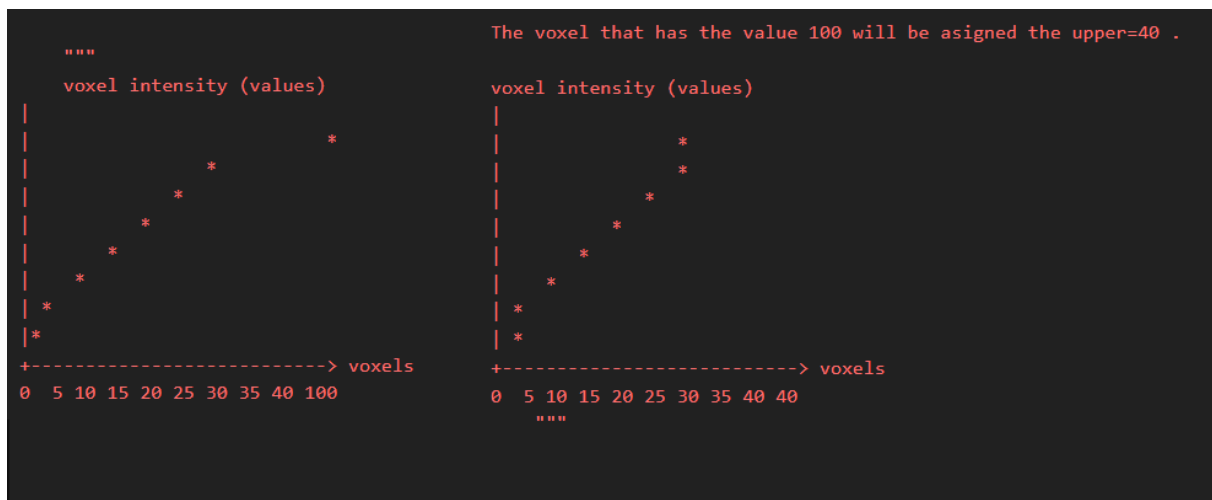
Γι' αυτό η επόμενη λειτουργικότητα αφορά την προεπεξεργασία των MRI δεδομένων και αποτελεί το αντίστοιχο workflow της CT ροής, προσαρμοσμένο όμως στις ιδιαιτερότητες της μαγνητικής τομογραφίας. Η ανάγκη για τη χρήση MRI προέκυψε από τη δομική ετερογένεια των διαθέσιμων datasets, καθώς τα σύνολα που αφορούν εγκεφαλικούς όγκους βασίζονται αποκλειστικά σε μαγνητικές τομογραφίες και όχι σε αξονικές, γεγονός που εισάγει διαφορετική φύση απεικονιστικής πληροφορίας. Σε αντίθεση με τα CT-scans, όπου οι τιμές εκφράζονται σε σταθερή φυσική κλίμακα HU, οι MRI εντάσεις είναι σχετικές και εξαρτώνται από το πρωτόκολλο λήψης και το απεικονιστικό σύστημα. Για τον λόγο αυτό απαιτείται ξεχωριστή προεπεξεργασία, με στόχο την εξομάλυνση των διαφορών έντασης και την εξασφάλιση συνεπούς χωρικής αναπαράστασης.

Η φόρτωση των MRI όγκων πραγματοποιείται από αρχεία NIfTI, τα οποία αποτελούν την καθιερωμένη μορφή αποθήκευσης για τρισδιάστατες μαγνητικές τομογραφίες. Κατά τη φόρτωση, κάθε όγκος μετατρέπεται σε κοινό ανατομικό προσανατολισμό, ώστε να διασφαλίζεται ότι όλα τα δείγματα έχουν συνεπή χωρική ευθυγράμμιση ανεξάρτητα από την πηγή προέλευσής τους. Η ομοιομορφία του προσανατολισμού είναι ουσιώδης, ιδιαίτερα όταν τα δεδομένα προέρχονται από διαφορετικά κέντρα ή πρωτόκολλα λήψης, καθώς αποτρέπει ασυμφωνίες στην ανατομική αντιστοίχιση των τομών.

Στη συνέχεια εφαρμόζεται σταθεροποίηση της έντασης μέσω robust clipping με βάση ποσοστημόρια, όπως φαίνεται στην Εικόνα 4.14, ώστε να απομακρύνονται ακραίες τιμές και να περιορίζεται η επίδραση θορύβου ή ακραίων εντάσεων που δεν αντιπροσωπεύουν τον κύριο απεικονιστικό όγκο του ιστού. Ακολουθεί κανονικοποίηση των τιμών στην κλίμακα $[0,1]$,

$$vol_{norm} = \frac{vol - clip_min}{clip_max - clip_min}, \quad (4.1)$$

ώστε οι διαφορετικοί όγκοι να αποκτούν συγκρίσιμη αριθμητική συμπεριφορά. Η επιλογή percentile-based clipping σε συνδυασμό με min-max normalization είναι ιδιαίτερα κατάλληλη για MRI, επειδή οι εντάσεις τους δεν είναι από μόνες τους ποσοτικοποιημένες όπως στις CT, αλλά παρουσιάζουν σημαντικές διακυμάνσεις μεταξύ εξετάσεων.



Εικόνα 4.14: Παράδειγμα robust clipping

Έπειτα, οι όγκοι επαναπροσανατολίζονται έτσι ώστε ο άξονας του βάθους να βρίσκεται πρώτος και εφαρμόζεται τρισδιάστατη επαναδειγματοληψία, όχι μόνο στον άξονα των τομών αλλά και στους χωρικούς άξονες, ώστε όλα τα δείγματα να αποκτούν κοινό σχήμα. Η διαδικασία αυτή είναι αναγκαία, επειδή τα MRI datasets εμφανίζουν σημαντική μεταβλητότητα ως προς το voxel spacing και τις χωρικές διαστάσεις, και συνεπώς απαιτείται ενοποίηση της γεωμετρίας πριν από οποιαδήποτε περαιτέρω ανάλυση. Με τον τρόπο αυτό τα volumes αποκτούν σταθερό τετραδιάστατο σχήμα εισόδου (N,D,H,W) και γίνονται συμβατά με τα επόμενα στάδια του pipeline.

Τέλος, η φόρτωση των MRI δεδομένων οργανώνεται αναδρομικά μέσα από τη δομή των φακέλων, με αντιστοίχιση κάθε δείγματος στη σωστή κατηγορία και με επιλογή ενός αντιπροσωπευτικού όγκου ανά ασθενή. Τα segmentation files παραλείπονται, ώστε να διατηρούνται μόνο οι πρωτογενείς απεικονίσεις. Οι τελικοί όγκοι και οι αντίστοιχες ετικέτες συγκεντρώνονται σε ενιαία δομή δεδομένων και

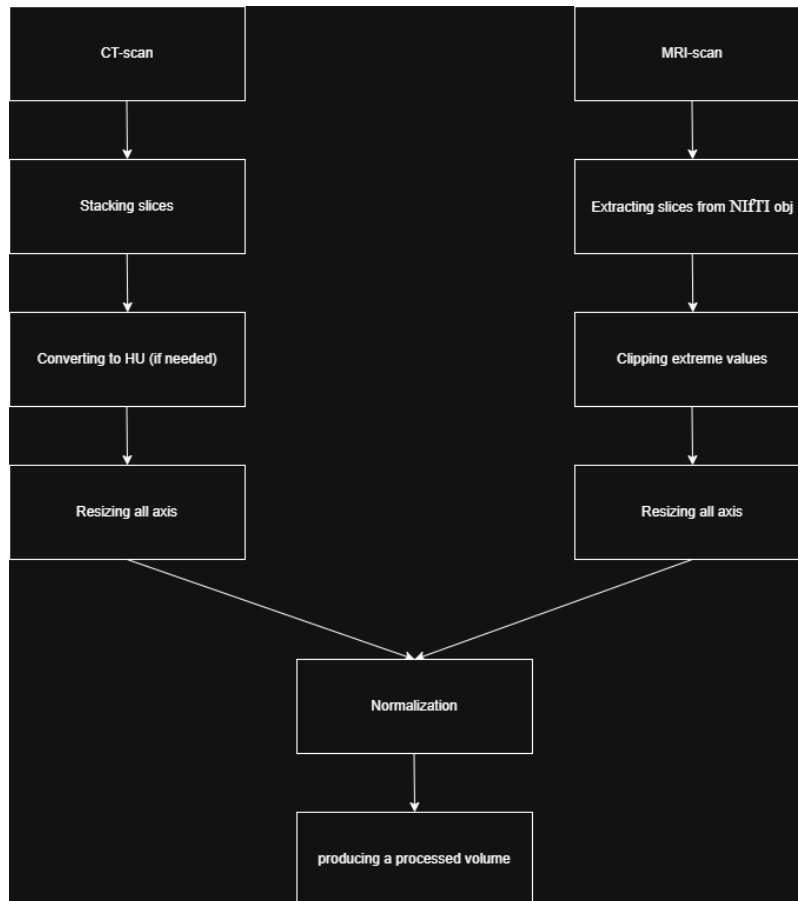
επιστρέφονται ως τελικό σύνολο, έτοιμο για τα επόμενα στάδια της επεξεργασίας όπως ακριβώς στο Σχήμα 4.3.

4.5 Έμφαση στην διαφορά ανάμεσα σε CT-scans και MRI-scans

Οι CT και MRI απεικονίσεις παρουσιάζουν σημαντικές διαφορές τόσο ως προς τη φύση των εντάσεων όσο και ως προς τη χωρική τους συνέπεια, γεγονός που επηρεάζει άμεσα τη διαδικασία προεπεξεργασίας του καθενός. Οι CT απεικονίσεις βασίζονται σε Hounsfield Units (HU), οι οποίες παρέχουν μια πιο συνεπή και φυσικά βαθμονομημένη κλίμακα εντάσεων, επιτρέποντας την εφαρμογή thresholding και HU-based παραθύρων για την αναγνώριση διαφορετικών ιστών. Παρόλα αυτά, οι CT εξετάσεις εξακολουθούν να επηρεάζονται από διαφοροποιήσεις reconstruction kernels, slice thickness και acquisition parameters. Αντίθετα, στις MRI απεικονίσεις οι τιμές έντασης δεν αντιστοιχούν σε σταθερή φυσική κλίμακα και εξαρτώνται σημαντικά από την ακολουθία λήψης, τον κατασκευαστή και τις παραμέτρους σάρωσης, γεγονός που καθιστά αναγκαία την εφαρμογή percentile clipping και normalization αντί για HU-based κανόνες.

Παράλληλα, τα MRI δεδομένα που προέρχονται από διαφορετικά κέντρα ενδέχεται να εμφανίζουν διαφοροποιήσεις στον χωρικό προσανατολισμό και στη γεωμετρική αναπαράσταση των όγκων. Για τον λόγο αυτό εφαρμόζεται μετατροπή σε canonical orientation μέσω της διαδικασίας as_closest_canonical, ώστε να επιτευχθεί μεγαλύτερη συνέπεια ως προς τον προσανατολισμό των αξόνων πριν από την περαιτέρω επεξεργασία. Ωστόσο, η διαδικασία αυτή δεν εξαλείφει άλλες πηγές μεταβλητότητας, όπως διαφοροποιήσεις voxel spacing, intensity bias ή geometric distortions. Επιπλέον, οι MRI εξετάσεις παρουσιάζουν συχνά μεγάλη ποικιλία ως προς τα voxel sizes, το spatial resolution και τον αριθμό των slices, γεγονός που καθιστά απαραίτητη την επαναδειγματοληψία σε όλους τους άξονες ώστε να επιτευχθεί κοινό target shape. Αντίστοιχα, στις CT εξετάσεις, παρότι υπάρχουν διαφοροποιήσεις στις παραμέτρους ανακατασκευής, οι τομές προέρχονται συνήθως από πιο συνεπείς volumetric σειρές, με αποτέλεσμα η κύρια ανάγκη επαναδειγματοληψίας να εντοπίζεται κυρίως στον άξονα του βάθους.

Η διαδικασία προεπεξεργασίας που ακολουθήθηκε για την εκπαίδευση των 3D CNN συνοψίζεται σχηματικά στο Σχήμα 4.4. Όπως φαίνεται, η ροή διαφοροποιείται αρχικά ανάλογα με τον τύπο της ιατρικής απεικόνισης (CT ή MRI), καθώς τα δύο modalities παρουσιάζουν διαφορετικά φυσικά και υπολογιστικά χαρακτηριστικά. Στην περίπτωση των CT-scans, οι επιμέρους DICOM τομές στοιβάζονται αρχικά ώστε να σχηματιστεί ένας ενιαίος τρισδιάστατος όγκος (stacking slices), ενώ στη συνέχεια πραγματοποιείται μετατροπή των intensities σε Hounsfield Units (HU), όπου αυτό απαιτείται, προκειμένου οι τιμές έντασης να αποκτήσουν φυσική σημασία και συγκρισιμότητα μεταξύ διαφορετικών εξετάσεων. Αντίθετα, στα MRI-scans η διαδικασία ξεκινά από την εξαγωγή των slices από αρχεία μορφής NIfTI και ακολουθεί clipping ακραίων τιμών intensities, ώστε να περιοριστεί η επίδραση θορύβου και ακραίων αποκλίσεων που σχετίζονται με τη μεταβλητότητα των MRI πρωτοκόλλων. Και στις δύο περιπτώσεις εφαρμόζεται στη συνέχεια επαναδειγματοληψία/ανακλιμάκωση (resizing-resampling) σε όλους τους άξονες, με στόχο την παραγωγή volumes σταθερών διαστάσεων και χωρικής ανάλυσης, κατάλληλων για είσοδο στο συνελκτικό νευρωνικό δίκτυο. Τέλος, οι δύο ροές συγκλίνουν σε κοινό στάδιο normalization, όπου οι τιμές intensities κανονικοποιούνται σε κοινή αριθμητική κλίμακα, οδηγώντας τελικά στην παραγωγή ενός επεξεργασμένου τρισδιάστατου volume (processed volume), το οποίο χρησιμοποιείται ως τελική είσοδος για την εκπαίδευση και αξιολόγηση του μοντέλου.



Σχήμα 4.4: Ροή της επεξεργασίας των δεδομένων μου

Το τελευταίο στάδιο του preprocessing αφορά την αύξηση των δεδομένων -γενικότερα, είτε είναι CT, είτε είναι MRI- με τεχνητά μέσα, μέσω στοχευμένων μετασχηματισμών πάνω στους τρισδιάστατους όγκους, ώστε να δημιουργηθούν επιπλέον παραλλαγές που διατηρούν την ανατομική πληροφορία αλλά ενισχύουν την ποικιλία του συνόλου. Η διαδικασία περιλαμβάνει τυχαίες μικρές περιστροφές, μετατοπίσεις, πιθανή αναστροφή της απεικόνισης και προσθήκη ήπιου Gaussian θορύβου, με στόχο να αυξηθεί η ανθεκτικότητα του μοντέλου σε μικρές γεωμετρικές και αριθμητικές μεταβολές των εισόδων. Επειδή το augmentation εφαρμόζεται σε ήδη κανονικοποιημένα volumes, οι τιμές παραμένουν εντός του επιθυμητού εύρους, ενώ η χρήση του αποσκοπεί κυρίως στη μείωση του overfitting και στη βελτίωση της γενίκευσης, ιδιαίτερα σε περιπτώσεις περιορισμένου ή ετερογενούς πλήθους δεδομένων. Η εφαρμογή του πρέπει να γίνεται αποκλειστικά στο training set και μετά τον διαχωρισμό των δεδομένων, ώστε να αποφεύγεται το φαινόμενο data leakage και να διατηρείται το validation set ως πραγματικά ανεξάρτητο σύνολο αξιολόγησης.

4.5 Δημιουργία του CNN

Η αρχιτεκτονική του συνελκτικού νευρωνικού δικτύου που αναπτύχθηκε βασίζεται σε τρισδιάστατες συνελίξεις (3D Convolutional Neural Network), ώστε να αξιοποιείται η χωρική πληροφορία των ιατρικών δεδομένων, τα οποία αναπαρίστανται ως τρισδιάστατοι όγκοι (volumes) και όχι ως μεμονωμένες δισδιάστατες εικόνες. Με αυτόν τον τρόπο, το μοντέλο μπορεί να λαμβάνει υπόψη όχι μόνο τη μορφολογία εντός ενός slice, αλλά και τις σχέσεις μεταξύ διαδοχικών slices κατά μήκος του άξονα βάθους. Η επιλογή τριών convolutional layers επιδιώκει μια ισορροπία μεταξύ εκφραστικής ικανότητας και υπολογιστικής πολυπλοκότητας.

Η είσοδος του δικτύου ορίζεται ως τετραδιάστατος πίνακας με διαστάσεις (βάθος \times ύψος \times πλάτος \times κανάλι), όπου το βάθος αντιστοιχεί στον αριθμό των slices του όγκου, το ύψος και το πλάτος στις χωρικές διαστάσεις κάθε slice, ενώ το κανάλι λαμβάνει την τιμή 1, καθώς πρόκειται για δεδομένα μίας μόνο έντασης. Όλα τα δεδομένα έχουν προηγουμένως κανονικοποιηθεί και επαναδειγματοληφθεί, ώστε να διαθέτουν κοινές διαστάσεις και συγκρίσιμα αριθμητικά χαρακτηριστικά.

Το πρώτο τμήμα του δικτύου αποτελείται από διαδοχικά 3D συνελκτικά επίπεδα, τα οποία λειτουργούν ως μηχανισμός εξαγωγής χαρακτηριστικών. Το αρχικό συνελκτικό επίπεδο εφαρμόζει φίλτρα μικρού μεγέθους ($3 \times 3 \times 3$) στον όγκο εισόδου και στοχεύει στην ανίχνευση απλών τοπικών δομών, όπως ακμές, μεταβολές έντασης και βασικά χωρικά μοτίβα. Ακολουθεί επίπεδο μέγιστης υποδειγματοληψίας (MaxPooling3D), το οποίο μειώνει τις διαστάσεις των παραγόμενων χαρακτηριστικών χαρτών, περιορίζοντας την υπολογιστική πολυπλοκότητα και ενισχύοντας την ανθεκτικότητα του μοντέλου σε μικρές χωρικές μετατοπίσεις.

Στη συνέχεια, το δίκτυο περιλαμβάνει επιπλέον συνελκτικά επίπεδα με αυξημένο αριθμό φίλτρων, επιτρέποντας την εξαγωγή πιο σύνθετων και αφηρημένων χαρακτηριστικών. Καθώς το βάθος του δικτύου αυξάνεται, τα φίλτρα παύουν να ανταποκρίνονται αποκλειστικά σε απλές τοπικές μεταβολές και αρχίζουν να κωδικοποιούν πιο υψηλού επιπέδου χωρικές δομές, οι οποίες σχετίζονται με την ανατομία και τις παθολογικές διαφοροποιήσεις των εξεταζόμενων όγκων. Μεταξύ των συνελκτικών επιπέδων εφαρμόζεται εκ νέου υποδειγματοληψία, ώστε η πληροφορία να συμπυκνώνεται προοδευτικά.

Μετά την ολοκλήρωση της συνελκτικής επεξεργασίας, οι τριδιάστατοι χάρτες χαρακτηριστικών μετατρέπονται σε μονοδιάστατο διάνυσμα μέσω του επιπέδου Flatten. Το διάνυσμα αυτό τροφοδοτείται σε ένα πλήρως συνδεδεμένο επίπεδο (Dense layer), το οποίο λειτουργεί ως ταξινομητής και συνδυάζει τα εξαγόμενα χαρακτηριστικά για τη λήψη της τελικής απόφασης. Η χρήση μη γραμμικής ενεργοποίησης στο στάδιο αυτό επιτρέπει τη μοντελοποίηση σύνθετων συσχετίσεων μεταξύ των χαρακτηριστικών.

Η έξοδος του δικτύου υλοποιείται μέσω ενός μονού νευρώνα χωρίς ενεργοποίηση (linear output), ο οποίος παράγει μια πραγματική τιμή (logit) που αντιπροσωπεύει τη μη κανονικοποιημένη ισχύ της πρόβλεψης υπέρ της κλάσης του καρκινικού όγκου. Η τελική ερμηνεία αυτής της τιμής ως πιθανότητα πραγματοποιείται έμμεσα μέσω της συνάρτησης απώλειας binary cross-entropy with logits, η οποία ενσωματώνει εσωτερικά τη sigmoid κατά τον υπολογισμό του σφάλματος και παρέχει αριθμητική σταθερότητα. Το δίκτυο βελτιστοποιείται με τον αλγόριθμο Adam, ο οποίος προσαρμόζει δυναμικά τον ρυθμό μάθησης και συμβάλλει σε σταθερή και αποδοτική σύγκλιση κατά την εκπαίδευση.

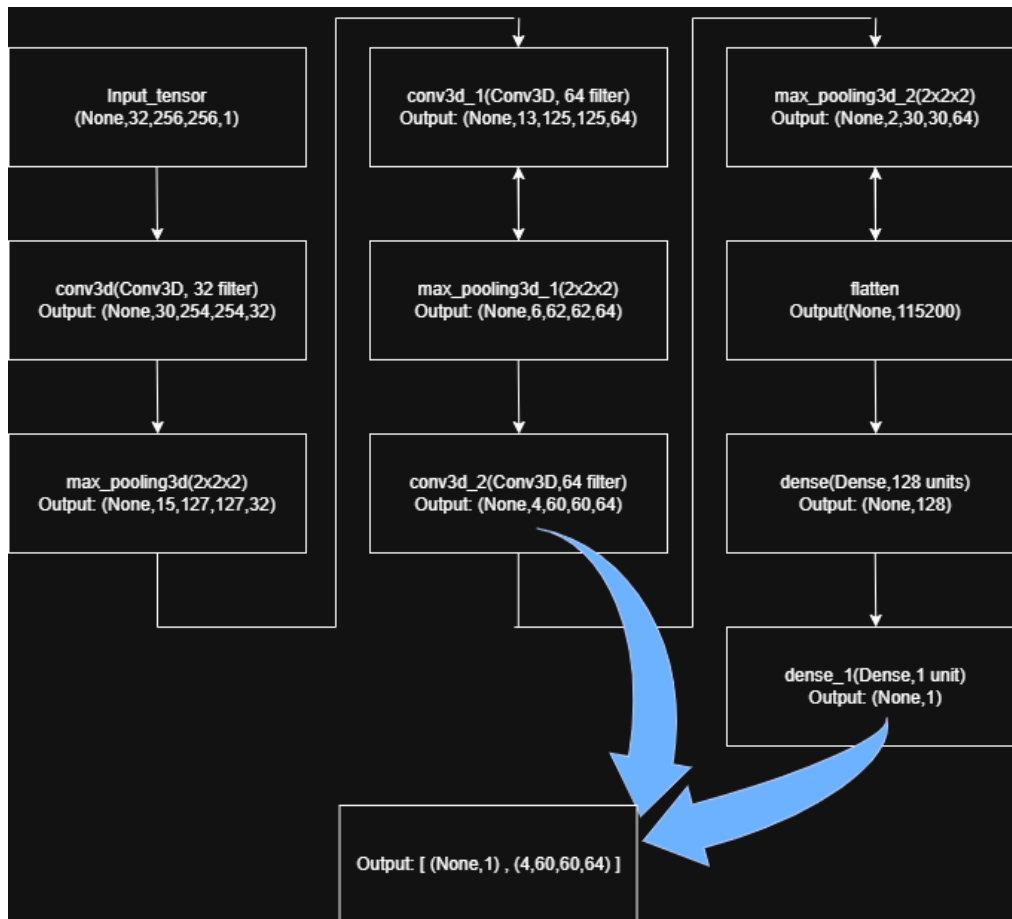
Επιπλέον, η επιλογή της εξόδου σε μορφή logits εξυπηρετεί και πρακτικούς σκοπούς εξηγησιμότητας, καθώς οι πιθανότητες στο διάστημα $[0,1]$ μπορεί να κορεστούν και να οδηγήσουν σε πολύ μικρά gradients. Η εργασία σε logits διατηρεί συνήθως πιο ευαίσθητες παραγώγους, γεγονός που διευκολύνει την παραγωγή πιο διακριτών heatmaps σε τεχνικές εξηγησιμότητας, όπως το Grad-CAM, στα σημεία στα οποία εστιάζει το CNN. Αντίθετα, σε αρχιτεκτονικές τύπου segmentation, όπως το U-Net, η έξοδος είναι χωρικά κατανομημένη και όχι μία ενιαία κλιμακωτή πρόβλεψη ανά κλάση, με αποτέλεσμα ο υπολογισμός και η ερμηνεία των gradients να μην είναι το ίδιο άμεσοι για τον ίδιο τύπο heatmap. Για τον λόγο αυτό, η συγκεκριμένη επιλογή εξόδου είναι ιδιαίτερα κατάλληλη για το παρόν classification setup.

Τέλος, για τον περιορισμό φαινομένων υπερπροσαρμογής εφαρμόζεται μηχανισμός πρόωρης διακοπής της εκπαίδευσης (early stopping), ο οποίος παρακολουθεί τη συμπεριφορά του σφάλματος στο σύνολο επικύρωσης (validation_set) και διακόπτει την εκπαίδευση όταν δεν παρατηρείται περαιτέρω βελτίωση.

Η προσέγγιση αυτή επιδιώκει να διατηρήσει το μοντέλο σε σημείο όπου επιτυγχάνεται καλή προσαρμογή στα διαθέσιμα δεδομένα χωρίς περιττή συνέχιση της εκπαίδευσης.

4.6 Επεξηγηματική τεχνητή νοημοσύνη

Η επεξηγηματική συνιστώσα βασίζεται σε ένα λειτουργικό (functional) Keras μοντέλο που επιστρέφει δύο τιμές: έναν δεκαδικό αριθμό (logit) και τον χάρτη χαρακτηριστικών (feature map) του τελευταίου Conv3D επιπέδου, ακριβώς όπως φαίνεται στο Σχήμα 4.5. Η επιλογή για γραμμική έξοδο (logits) — αντί για πρότερη εφαρμογή sigmoid — διευκολύνει τον υπολογισμό ισχυρότερων και μη-κορεσμένων παραγώγων, βελτιώνοντας την ευαισθησία των τεχνικών επεξηγησιμότητας (π.χ. Grad-CAM) και παράγοντας πιο έντονα, ερμηνεύσιμα heatmaps που μπορούν να λειτουργήσουν ως ενδεικτικοί χάρτες επεξήγησης της απόφασης του μοντέλου. Το λειτουργικό μοντέλο κατασκευάζεται επαναχρησιμοποιώντας τα layers του εκπαιδευμένου sequential μοντέλου: ένα μοναδικό `Input(shape=(D,H,W,1))` τροφοδοτεί σειριακά όλα τα layers, ο δε τελευταίος Conv3D εξάγεται ως το `target_tensor`, ενώ το τελικό αποτέλεσμα του δικτύου διατηρείται ως `out`. Το `conv_model` που προκύπτει επιστρέφει ταυτόχρονα `[out, target_tensor]` για κάθε πρόβλεψη, παρέχοντας έτσι τις απαραίτητες πληροφορίες για την επεξηγησιμότητα.



Σχήμα 4.5: Δομή του CNN για την επεξηγηματική νοημοσύνη

Το κύριο βήμα εξήγησης είναι ο υπολογισμός ενός 3-D Grad-CAM heatmap για κάθε όγκο. Δίνοντας έναν όγκο volume σε μορφή (D,H,W), στην Εικόνα 4.15, ο όγκος πρώτα μετατρέπεται σε tf-tensor με σχήμα (1,D,H,W,1). Στη συνέχεια, με `tf.GradientTape()` προωθούνται τα δεδομένα μέσα στο `conv_model` ώστε να προκύψει το `pred` (logit) και το `conv_outputs` (τελευταίος 3D feature map). Η τιμή

στόχου (score) λαμβάνεται από το output του μοντέλου ως μία βαθμωτή ποσότητα(logit) και στη συνέχεια υπολογίζονται οι παράγωγοί της ως προς τα στοιχεία του conv_outputs . Τα προκύπτοντα gradients εκφράζουν πόσο ευαίσθητη είναι η τελική απόφαση του δικτύου σε κάθε φίλτρο και σε κάθε χωρική θέση του feature map.

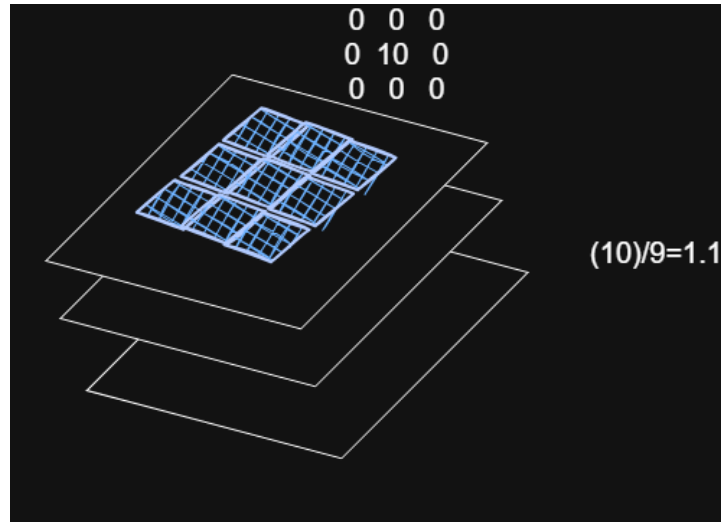
```
# 1) build concrete input tensor (1,D,H,W,1)
input_tensor=getting_an_input_tensor_out_of_a_volume(volume)

# Find the target layer if not provided
# if target_layer is None:
#     target_layer = find_last_conv3d_layer(model)
# elif isinstance(target_layer, str):
#     target_layer = model.get_layer(target_layer)

# Forward + gradient computation
with tf.GradientTape() as tape:
    print(f"Forward pass through the CNN :")
    # Actually getting the logit and the feature map for a single sample(from the input_tensor
    # that i produced out of the volume parameter) :
    preds, conv_outputs = model(input_tensor, training=False) # preds=probability, conv_outputs=feature map
    print(f"Got prediction={preds} and feature_map={conv_outputs.shape} from forwarding")
    # Ensure tape watches the conv layer outputs by forwarding once and getting conv output
    #tape.watch(conv_outputs)
    print(f"tape.watch(model.inputs worked...)")
    # If model returns tuple/List for the preds variable i handle that by getting the first element out of the list which is the one
    # iam interested at :
    if isinstance(preds, (list, tuple)):
        out0 = preds[0]
    else:
        out0 = preds
    print(f"Got the logit : {preds}") #and the feature map : {conv_outputs} out of a single sample with the format of an input tensor : {input_tensor}
```

Εικόνα 4.15: Κώδικας υπολογισμού των παραγώγων

Οι τελικές «βάσεις» (weights) κάθε καναλιού προκύπτουν από την εξαγωγή ενός μέσου όρου των gradients πάνω στις χωρικές διαστάσεις (batch, D', H', W'), όπως φαίνεται στην Εικόνα 4.16, δηλαδή $w_c = \text{mean}_{b,d,h,w}(\partial \text{score} / \partial \text{feature}_{b,d,h,w,c})$. Τα βάρη αυτά κλιμακώνονται για ορατότητα και πολλαπλασιάζονται κατόπιν με τα αντίστοιχα feature maps(1,4,60,60,64) , όπου τα feature maps είναι τόσα όσα και τα φίλτρα , δηλαδή 64. Το Grad-CAM 3-D προκύπτει ως άθροισμα κατά το κανάλι: $\text{cam}(d, h, w) = \sum_c w_c \cdot \text{feature}_{0,d,h,w,c}$, και εφαρμόζεται ReLU ώστε να διατηρηθούν μόνο οι θετικές συνεισφορές που ενισχύουν την προβλεψη της θετικής κλάσης. Το αποτέλεσμα είναι ένας χαρτοχρωματικός όγκος (D',H',W') όπου οι υψηλές τιμές υποδεικνύουν περιοχές που το μοντέλο φαίνεται να χρησιμοποιεί εντονότερα για την απόφαση. Το κοκκινωπό χρώμα αντιστοιχεί σε περιοχές αυξημένης συνεισφοράς στο μοντέλο, οι οποίες ενδέχεται να συσχετίζονται με παθολογικά ευρήματα.



Σχήμα 4.6: Πως υπολογίζεται η σημαντικότητα ενός φίλτρου

Ο μέσος όρος προφανώς δεν είναι 10 αλλά $10/9 \approx 1.11$, επειδή η τιμή κατανέμεται σε όλα τα στοιχεία του πλέγματος. Με αυτόν τον τρόπο, ένα φίλτρο αποκτά μεγάλο βάρος μόνο όταν παρουσιάζει συστηματικά υψηλές τιμές σε πολλές θέσεις και όχι όταν εμφανίζει μια μεμονωμένη έντονη κορυφή. Ο υπολογισμός του mean λειτουργεί επομένως ως μηχανισμός εξομάλυνσης που αποτυπώνει τη συνολική σημασία του φίλτρου για την απόφαση του δικτύου, αποφεύγοντας να υπερεκτιμηθούν τυχαίες τοπικές ενεργοποιήσεις.

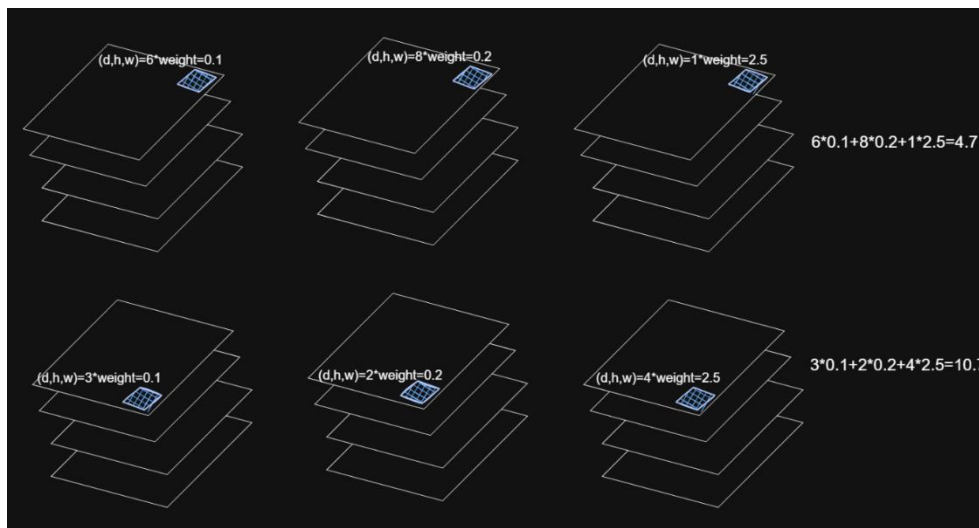
Ο τελικός ανεπεξέργαστος χάρτης επιρροής προκύπτει ως σταθμισμένο άθροισμα κατά κανάλι:

$$CAM(d, h, w) = \sum_c w_c A_c(d, h, w), \quad (4.3)$$

και ακολουθεί εφαρμογή ReLU ώστε να διατηρηθούν μόνο οι θετικές (ενισχυτικές) συνεισφορές:

$$CAM^+(d, h, w) = \text{ReLU}(CAM(d, h, w)). \quad (4.4)$$

Για οπτική αναπαράσταση το αποτέλεσμα κανονικοποιείται στο διάστημα $[0, 1]$, προσαρμόζεται στην ανάλυση του αρχικού όγκου (spatial resize + depth resample) και επιστρώνεται χρωματικά (colormap) πάνω στην κανονικοποιημένη ακτινολογική εικόνα για παραγωγή του overlay.



Σχήμα 4.7: Παράδειγμα υπολογισμού του χρώματος για ένα συγκεκριμένο σημείο στο heatmap που συντίθεται. Στο παράδειγμα του παραπάνω Σχήματος 4.7 αποτυπώνεται καθαρά η αλληλεπίδραση μεταξύ τοπικής ενεργοποίησης και σημαντικότητας φίλτρων στο Grad-CAM. Σε ένα σημείο P_1 τα φίλτρα c_1, c_2, c_3 εμφανίζουν σχετικά υψηλές ενεργοποιήσεις (π.χ. 6, 8, 1), όμως τα δύο πρώτα φίλτρα έχουν μικρά βάρη (0.1 και 0.2), ενώ μόνο το τρίτο έχει μεγάλο βάρος (2.5). Το σταθμισμένο άθροισμα στο σημείο αυτό γίνεται $6 \cdot 0.1 + 8 \cdot 0.2 + 1 \cdot 2.5 = 4.7$. Σε ένα άλλο σημείο P_2 οι ενεργοποιήσεις είναι γενικά μέτριες (π.χ. 3, 2, 4), αλλά η ενεργοποίηση στο φίλτρο με το μεγάλο βάρος είναι σημαντικά υψηλότερη· έτσι το σταθμισμένο άθροισμα που προκύπτει είναι $3 \cdot 0.1 + 2 \cdot 0.2 + 4 \cdot 2.5 = 10.7$. Συνεπώς, παρότι στο P_1 ορισμένα φίλτρα ανταποκρίνονται εντονότερα συνολικά, το P_2 αναδεικνύεται πιο έντονα στο τελικό heatmap, επειδή η ενεργοποίηση αφορά φίλτρο που το δίκτυο θεωρεί πολύ πιο καθοριστικό για την τελική πρόβλεψη.

Ως άμεση συνέπεια του παραπάνω παραδείγματος, προκύπτει ότι το Grad-CAM heatmap δεν αποτυπώνει απλώς το πόσο έντονα ενεργοποιούνται τα τοπικά φίλτρα, αλλά κυρίως τον συνδυασμό έντασης ενεργοποίησης και σημασίας του κάθε φίλτρου για την τελική απόφαση του δικτύου. Έτσι, μια περιοχή που εμφανίζεται πιο «κόκκινη» δεν αντιστοιχεί απαραίτητα στο σημείο με τη μεγαλύτερη ακατέργαστη ενεργοποίηση, αλλά ενδέχεται να οφείλεται σε μέτρια ενεργοποίηση φίλτρων τα οποία έχουν υψηλή βαρύτητα για την πρόβλεψη. Για τον λόγο αυτό, η ερμηνεία των χαρτών τύπου heatmap πρέπει να γίνεται με προσοχή και με συγκεκριμένη μεθοδολογία: απαιτείται σταθερή επιλογή του επιπέδου στόχου (target layer), ορθή κανονικοποίηση των heat-maps και επιβεβαίωση των ευρημάτων με κλινικά κριτήρια, καθώς η τελική ένταση του overlay εξαρτάται άμεσα τόσο από την κλίμακα των βαρών w_c όσο και από τις ενεργοποιήσεις A_c , καθώς και από τις διαδικασίες κανονικοποίησης και επαναδειγματοληψίας (resizing).

Οπότε επειδή το feature map έχει συνήθως μικρότερη χωρική ανάλυση από τον αρχικό όγκο, όπως ειπώθηκε και προηγουμένως ακολουθεί η διαδικασία επεξεργασίας που διατυπώνεται στις Εικόνες 4.17 και 4.18: αρχικά κανονικοποίηση σε $[0,1]$, μετά ανά-επίπεδη (per-slice) χωρική επανεκτίμηση (cv2.resize) ώστε κάθε slice να αποκτήσει το επιθυμητό (H,W) , και τελικά τρισδιάστατο resampling κατά τον άξονα βάθους (π.χ. scipy.ndimage.zoom) για να προσαρμοστεί το βάθος $D' \rightarrow D$. Μετά από δεύτερη κανονικοποίηση προκύπτει το τελικό heatmap cam_final με σχήμα (D,H,W) και τιμές στο $[0,1]$. Η χρήση δύο σταδίων κανονικοποίησης εξασφαλίζει ότι οι αλλαγές κλίμακας κατά το resizing δεν δημιουργούν ψηφιακή αλλοίωση των σχετικών εντάσεων.

```

"""Start of processing the heatmap(D=4,H=124,W=124) as a numpy array so i can bring it to the same state(D=32,H=512,W=512)
as the rest of the volumes :"""
# Normalize then resize each slice in depth to target H,W, and resample depth via simple np.repeat or linear interpolation
D_out, H_out, W_out = cam_np.shape
D, H_t, W_t = volume.shape

# First normalize cam to [0,1]
print(f"Before first normalization to [0,1] : min={cam_np.min()},max={cam_np.max()}")
mn, mx = cam_np.min(), cam_np.max()
if mx - mn > 1e-6:
    cam_norm = (cam_np - mn) / (mx - mn)
else:
    cam_norm = np.zeros_like(cam_np)
print(f"After first normalization to [0,1] : min={cam_norm.min()},max={cam_norm.max()}")

# Resize spatially each depth slice to target H,W using cv2.resize :
print(f"Before adjusting the Height,Width : H={H_out},W={W_out}")
cam_resized_slices = []
for d in range(D_out):
    slice_ = cam_norm[d]
    slice_u8 = np.uint8(np.clip(slice_ * 255.0, 0, 255))
    # cv2.resize expects (W,H) order -> pass (W_t,H_t)
    resized = cv2.resize(slice_u8, (W_t, H_t), interpolation=cv2.INTER_LINEAR) # returns (H_t, W_t)
    cam_resized_slices.append(resized.astype(np.float32) / 255.0)
cam_spatial_resized = np.stack(cam_resized_slices, axis=0) # (D_out, H_t, W_t)
print(f"After adjusting the Height,Width : H={cam_spatial_resized.shape[1]}, W={cam_spatial_resized.shape[2]}")

```

Εικόνα 4.17: Κώδικας upsampling του heat-map

```

# Now handle depth resizing: if D_out != D, resample along depth axis :
print(f"Before adjusting the depth of the heat-map : depth={D_out}")
if D_out != D:
    # simple linear interpolation along depth
    from scipy.ndimage import zoom
    zoom_factor = (D / D_out, 1.0, 1.0)
    cam_final = zoom(cam_spatial_resized, zoom=zoom_factor, order=1) # trilinear-ish
else:
    cam_final = cam_spatial_resized
print(f"After adjusting the depth of the heat-map : depth={cam_final.shape[0]}")

# Final normalization to [0,1] since the resample might have affected the values.
print(f"Before second normalization to [0,1] : min={cam_final.min()},max={cam_final.max()}")
mn, mx = cam_final.min(), cam_final.max()
if mx - mn > 1e-6:
    cam_final = (cam_final - mn) / (mx - mn)
else:
    cam_final = np.zeros_like(cam_final)
print(f"After second normalization to [0,1] : min={cam_final.min()},max={cam_final.max()}")
print(f"In the end iam returning this : {cam_final.shape}") # this prints the final heat-map that went from (4,124,124) to (32,512,512).
return cam_final # (D=32,H=512,W=512) float in [0,1]
"""End of processing the heatmap as a numpy array ."""

```

Εικόνα 4.18: Κώδικας upsampling του heat-map

Για ανθρώπινη ανάγνωση, κάθε slice χρωματίζεται με την λειτουργικότητα της Εικόνας 4.19: ο προεπεξεργασμένος γκρι όγκος ($\text{gray_norm} = \text{clip}((\text{vol} - (-1000))/2000, 0, 1)$) αναπαράγεται ως RGB, το cam_map μετατρέπεται σε heatmap με cv2.applyColorMap και τα δύο συνδυάζονται με cv2.addWeighted για να παραχθεί ένα εύκολα αναγνώσιμο overlay. Τα τελικά αποτελέσματα αποθηκεύονται ανά δείγμα/φάκελο: το πρωτότυπο .npy heatmap διατηρείται για περαιτέρω ποσοτική ανάλυση και κάθε slice γράφεται ως .png με το αντίστοιχο overlay ώστε να είναι άμεσα διαθέσιμα για οπτική επιθεώρηση από κλινικούς και ερευνητές.

Η επιλογή της χωρικής ανάλυσης 256×256 προέκυψε κυρίως ως αναγκαίος συμβιβασμός λόγω των περιορισμών υπολογιστικής ισχύος, καθώς η χρήση 512×512 οδήγησε σε αδυναμία εκτέλεσης λόγω εξάντλησης μνήμης. Αντίθετα, το βάθος των 32 slices επιλέχθηκε συνειδητά, ώστε να διατηρείται μια επαρκής τρισδιάστατη αναπαράσταση του όγκου χωρίς να αλλοιώνεται υπερβολικά η χωρική αντιστοίχιση κατά το upsampling του Grad-CAM, διασφαλίζοντας έτσι ότι η σημασία των περιοχών που αναδεικνύει το heatmap παραμένει ουσιαστικά συνεπής πάνω στο αρχικό 3D volume. Με τον τρόπο αυτό επιτυγχάνεται ισορροπία ανάμεσα στη διαχειρίσιμη υπολογιστική απαίτηση και στη διατήρηση της ερμηνευτικής αξίας των εξηγηματικών χαρτών.

```

# here by using the logic of a heat-map i color a slice out of the 3d-volume with this logic :
def overlay_cam_on_slice_gray(vol, cam_map, slice_idx=None, cmap=cv2.COLORMAP_JET):
    """
    vol: (D,H,W) float in [0,1] , this is the one i produce and process from the DICOM or NIFTI files.
    cam_map: (D,H,W) float in [0,1] , this one is the output from the compute_3d_gradcam_for_volume() , the way i get cam_map
    is this one : vol → CNN → Grad-CAM → cam_map .
    returns overlay uint8 (H,W,3)
    Obviously the vol and the cam_map have to originate from the same 3d-volume in order for the coloring to have meaning
    """
    #print(f"Iam gonna color the slice_{slice_idx} by using the heatmap i just produced")
    D = vol.shape[0] # iam taking the depth of the volume which i know is the first axis.
    # In case i didnt define a slice i'll give by default the middle slice :
    if slice_idx is None:
        slice_idx = D // 2
    # Then iam selecting the slice i choosed from the gray-scale volume :
    img = vol[slice_idx] # the index of the slice is actually the depth value ! that makes sense ...
    # Then iam also selecting the corresponding slice from the heatmap :
    cam = cam_map[slice_idx]
    # Then iam creating an image with 3 channels so i can color it(i need it to have 3 channels so i can color it), out of the gray-scale image :
    img_3c = np.stack([img, img, img], axis=-1).astype(np.float32) # (H,W,3), float [0,1]
    img_rgb_uint8 = np.uint8(np.clip(img_3c * 255.0, 0, 255))
    # I color the image by using the heatmap i gave as a parameter :
    # Prepare heatmap
    heatmap = np.uint8(np.clip(cam * 255.0, 0, 255))
    heatmap_colored = cv2.applyColorMap(heatmap, cmap) # BGR # preparing the heatmap by turning it to 3d from 2d
    # by adding a channel dimension for the rgb logic : (H,W,3)
    heatmap_colored = cv2.cvtColor(heatmap_colored, cv2.COLOR_BGR2RGB) # BRG to RGB

    # Blend heatmap and original image (img_3c scaled to 0-255)
    overlay = cv2.addWeighted(img_rgb_uint8, 0.5, heatmap_colored, 0.5, 0)
    return overlay # uint8 (H,W,3)

```

Εικόνα 4.19: Κώδικας χρωματισμού ενός 3d-volume με το αντίστοιχο heat-map

Παράλληλα, η βοηθητική συνάρτηση `predict_on_test_set` αξιοποιεί το ίδιο `conv_model` για παραγωγή πιθανοτήτων: μετατρέπει τα logits σε πιθανότητες με sigmoid και επιστρέφει το διάνυσμα probs και τα binary preds. Τελικά, η συντονιστική συνάρτηση `run_and_save_gradcams_on_set` διενεργεί επαναληπτικά τη διαδικασία για όλα τα δείγματα ενός συνόλου (test/validation), αποθηκεύει heatmaps και overlays σε φακέλους ανά fold και παράγει εκτυπώσεις κατάστασης για debugging και ποιοτικό έλεγχο.

4.7 Εκτελέσιμος κώδικας

Η συνολική πειραματική ροή οργανώνεται σε τρία διακριτά στάδια: προετοιμασία δεδομένων, εκπαίδευση και αξιολόγηση του μοντέλου, και επεξηγηματική ανάλυση των αποτελεσμάτων. Στο πρώτο στάδιο, η κλήση της `split_total_body_by_slice_stats()` παράγει τα μη παθολογικά CT δείγματα που αξιοποιούνται ως υγιής αναφορά για τα επιμέρους μοντέλα.

Η εκτέλεση των λειτουργικοτήτων που αναπτύχθηκαν ξεκινάει, στην Εικόνα 4.20, με την κλήση της `split_total_body_by_slice_stats()` ή οποία παρέχει τα μη-ευπαθείς data που θα αποτελέσουν την τροφή για καθένα από τα 3 CNN έτσι ώστε να αντιληφθούν τα μοτίβα στις τιμές των pixels που καθιστούν ένα ιστό φυσιολογικό.

```

# Root directory that contains the non_cancer folder
root_dir = 'E:/DATA' # this destination has to contain a non_cancer/Healthy-Total-Body-CTs path
# in order for the split_total_body_by_slice_stats() to work ...

# Output directory where i'll save the results
output_root = "E:/DATA/non_cancer/"

# Calling the main functionality :
results = split_total_body_by_slice_stats(
    root_dir=root_dir,
    output_root=output_root,
    main_folder_name="non_cancer",
    subfolder_name="Healthy-Total-Body-CTs",
    organs=("lung", "brain", "spine"),

    # HU windows
    lung_hu_window=(-1000, -200),
    brain_hu_window=(15, 80),
    bone_hu_thresh=150,

    # Fraction thresholds
    lung_frac_thresh=0.30,
    brain_frac_thresh=0.04,
    bone_frac_thresh=0.03,

    # Block rules
    min_block_len=50,
    gap_tolerance=3,

    # Debug prints
    verbose=True
)

```

Εικόνα 4.20: Κλήση της συνάρτησης διαχωρισμού

Το μόνο στάδιο που απομένει είναι ο σχεδιασμός μιας συστηματικής διαδικασίας αξιοποίησης και ενοποίησης των επιμέρους συναρτήσεων που συγκροτούν το συνελκτικό νευρωνικό δίκτυο. Στο σημείο αυτό παρουσιάζεται αναλυτικά η αρχιτεκτονική και η συνολική ροή λειτουργίας του μεγαλύτερου και υπολογιστικά πιο απαιτητικού μοντέλου της εργασίας, δηλαδή του Brain CNN, το οποίο αναπτύχθηκε για την ανίχνευση εγκεφαλικών όγκων μέσω τρισδιάστατων ιατρικών απεικονίσεων.

Αρχικά, όπως φαίνεται στην Εικόνα 4.21, ορίστηκαν δύο δομές οργάνωσης των δεδομένων στο σύστημα αρχείων, οι οποίες αντιστοιχούν στις επιμέρους πηγές απεικονίσεων για το brain cohort. Η πρώτη αφορά τα MRI δεδομένα και περιλαμβάνει ξεχωριστούς υποφακέλους για τις καρκινικές και μη καρκινικές περιπτώσεις, ενώ η δεύτερη χρησιμοποιείται για την εισαγωγή συμπληρωματικών DICOM δεδομένων μη καρκινικών εγκεφαλικών απεικονίσεων. Η ιεράρχηση αυτή επιτρέπει την αυτοματοποιημένη φόρτωση των δεδομένων από τον ριζικό κατάλογο του πειράματος, σύμφωνα με μια προκαθορισμένη δομή αποθήκευσης.

Στη συνέχεια πραγματοποιείται η φόρτωση και προεπεξεργασία των MRI όγκων, καθώς και η αντίστοιχη φόρτωση των DICOM σειρών που θα χρησιμοποιηθούν ως συμπληρωματικά μη καρκινικά δείγματα. Κατά τη φόρτωση των MRI δεδομένων εφαρμόζονται οι απαραίτητες μετατροπές ως προς τον προσανατολισμό, την κανονικοποίηση της έντασης και την επαναδειγματοληψία, ώστε όλα τα volumes να αποκτούν κοινή αναπαράσταση εισόδου. Αντίστοιχα, τα DICOM δεδομένα προεπεξεργάζονται ώστε να καταστούν συμβατά ως προς τη δομή και την κλίμακα τιμών, με στόχο να είναι εφικτή η συνδυαστική αξιοποίησή τους στο ίδιο πειραματικό πλαίσιο. Η προσέγγιση αυτή υιοθετεί ενιαία κριτήρια μορφοποίησης των δεδομένων, χωρίς να υπονοεί ότι οι διαφορετικές modalities είναι εκ των προτέρων ισοδύναμες. Έτσι εξασφαλίζεται ότι η ετερογένεια των αρχικών δεδομένων δεν εισάγει ασυνέπειες στη μεταγενέστερη εκπαίδευση του μοντέλου.

```

"""Brain """
# Creating a dictionary that represents the 2 folders in my file system :
main_folders = {
    'cancer': ['PKG - BraTS-Africa', 'PKG - CFB-GBM_brain'],
    'non_cancer': 'Vestibular-Schwannoma-MC-RC2_Oct2025'
}
main_folders_2 = {
    'cancer': '',
    'non_cancer': 'Healthy-Total-Body-CTs_brain'
}
directory = 'E:/DATA'

print("Loading MRI files for brain CNN...")
print()
mri_volumes, mri_labels, mri_metadata = load_mri_files(directory, main_folders)
print("Loading DICOM files for brain CNN...\n")
print()
dicom_volumes, dicom_labels, dicom_metadata = load_dicom_files(directory, main_folders_2)

print("Preprocessing healthy brain volumes coming from CT-scans...\n")
processed_dicom_volumes = preprocess_CT_volumes(
    dicom_volumes,
    target_size=(32, 256, 256),
    clip_min=-20,
    clip_max=100,
    brain_volume=True
)
# processed_dicom_volumes doesn't change order, so metadata stays aligned:
processed_dicom_metadata = dicom_metadata

```

Εικόνα 4.21: Κλήση συναρτήσεων φόρτωσης και προεπεξεργασίας των δεδομένων

Αφού ολοκληρωθεί η προεπεξεργασία των δύο ομάδων δεδομένων, οι τελικοί όγκοι και οι αντίστοιχες ετικέτες τους ενοποιούνται σε ενιαία δομή, ώστε να προκύψει ένα κοινό σύνολο εκπαίδευσης με σαφή αντιστοιχισμός μεταξύ κάθε απεικόνισης και της κλάσης της. Η συνένωση αυτή πραγματοποιείται με διατήρηση της σειράς εισαγωγής των δεδομένων και παράγει δύο τελικά σύνολα, τα `brain_volumes` και `brain_labels`, όπου κάθε όγκος αντιστοιχεί σε ετικέτα καρκινικού ή μη καρκινικού δείγματος. Για την επιβεβαίωση της ορθής ολοκλήρωσης της διαδικασίας πραγματοποιούνται έλεγχοι ακεραιότητας, οι οποίοι διασφαλίζουν ότι το πλήθος των τελικών όγκων ισούται με το άθροισμα των επιμέρους συνόλων, ότι οι ετικέτες ακολουθούν σωστά το πλήθος των αντιστοιχών δειγμάτων και ότι δεν προκύπτει ασυμφωνία μεταξύ όγκων και labels. Με αυτόν τον τρόπο διασφαλίζεται η συνοχή του τελικού dataset και αποτρέπεται η εμφάνιση σφαλμάτων αντιστοιχισμός που θα μπορούσαν να επηρεάσουν αρνητικά την εκπαίδευση και την αξιολόγηση του μοντέλου.

Στο δεύτερο στάδιο, το ενιαίο σύνολο δεδομένων περνά σε διαδικασία εκπαίδευσης και αξιολόγησης με `StratifiedKFold`, ώστε σε κάθε fold να διατηρείται η ίδια αναλογία κλάσεων. Η ροή αυτή διαχωρίζεται καθαρά σε training και test υποσύνολα πριν από οποιαδήποτε μορφή augmentation. Το augmentation εφαρμόζεται αποκλειστικά στο training set, ακριβώς μετά τον διαχωρισμό, έτσι ώστε να αποφεύγεται το data leakage και να παραμένει το test set πραγματικά ανεξάρτητο. Με τον τρόπο αυτό, το μοντέλο εκπαιδεύεται σε εμπλουτισμένα δεδομένα χωρίς να αποκτά έμμεση πρόσβαση στα δείγματα αξιολόγησης. Η εκπαίδευση ολοκληρώνεται με `early stopping`, ώστε να περιορίζεται ο κίνδυνος υπερπροσαρμογής, ενώ τα βάρη αποθηκεύονται ανά fold για μελλοντική σύγκριση και επαναληψιμότητα. Ακολουθεί ο βασικός βρόχος του πειράματος, στην Εικόνα 4.22, όπου για κάθε fold εκτελούνται διαδοχικά όλα τα στάδια. Τα δεδομένα διαχωρίζονται σε training και test sets με βάση τα indices που επιστρέφει το `StratifiedKFold`. Σε αυτό το σημείο τα δεδομένα εξακολουθούν να βρίσκονται στη μορφή (N, D, H, W), δηλαδή ως καθαροί τρισδιάστατοι όγκοι.

```

#####
# From now on starts the Logic of the main Loop :#
#####

fold_num=1
for train_idx, test_idx in skf.split(brain_volumes,brain_labels): # Now i do the split with a certain StratifiedKFold object!
    """
    train_idx  -> numpy array with indices(position in the array) for TRAIN_SET
    test_idx   -> numpy array with indices(position in the array) for TEST_SET

    Assume i have N=10 , then i'll get :
    labels = [1,0,1,0,1,0,0,1,0,1]
    train_idx = [0,1,2,4,6,7,8,9]
    test_idx  = [3,5]
    """

    """Splitting the data :"""
    #split data according to the indices i got from skf.split() :
    # Here i still work with this format : (N_train, D, H, W) , (N_test, D, H, W)
    x_train, x_test = brain_volumes[train_idx],brain_volumes[test_idx] # python syntax
    y_train,y_test = brain_labels[train_idx],brain_labels[test_idx] # python syntax

    # --- NEW: split the metadata in the same way (metadata is a list of dicts) ---
    # train_meta, test_meta will travel along with the volumes/Labels
    train_meta = [ brain_metadata[idx] for idx in train_idx ]
    test_meta  = [ brain_metadata[idx] for idx in test_idx ]

```

Εικόνα 4.22: Ορίζοντας τον διαχωρισμό των δεδομένων

Στη συνέχεια, τόσο το training όσο και το test set μετασχηματίζονται στη μορφή που απαιτείται από το Keras, δηλαδή (BatchSize, Depth, Height, Width, Channels), με την προσθήκη ενός καναλιού (channels = 1). Οι ετικέτες αναδιαμορφώνονται αντίστοιχα ώστε να είναι συμβατές με τη δυαδική ταξινόμηση.

Αφού τα δεδομένα αποκτήσουν τη σωστή μορφή, στην Εικόνα 4.23, ξεκινά η εκπαίδευση του CNN για το συγκεκριμένο fold. Το μοντέλο εκπαιδεύεται με early stopping ώστε να αποφεύγεται η περιττή εκπαίδευση όταν η απόδοση στο validation set παύει να βελτιώνεται. Μετά το τέλος της εκπαίδευσης, τα βάρη του μοντέλου αποθηκεύονται ξεχωριστά για κάθε fold. Αυτό επιτρέπει μελλοντική ανάλυση, σύγκριση μεταξύ folds ή επαναχρησιμοποίηση των μοντέλων χωρίς επανεκπαίδευση.

```

"""Training of the model for this fold :"""
print(f"=====")
print(f"Start training :")
model = train_and_evaluate_model(x_train_aug,x_test,y_train_aug,y_test,epochs=EPOCHS,batch_size=BATCH_SIZE)
#model.to(device) # TensorFlow / Keras uses automatically the GPU if there is one so this line is Unnecessary .

# Save whatever weights i ended up having from the training for later analysis :
fold_weights_dir = r"E:\DATA\models_by_fold_brain"
os.makedirs(fold_weights_dir,exist_ok=True)
model.save(os.path.join(fold_weights_dir, f"model_fold_{fold_num}.keras"))
"""End of Training of the model for this fold ."""

```

Εικόνα 4.23: Κλήση συνάρτησης εκπαίδευσης

Στη συνέχεια υλοποιείται ένα κρίσιμο βήμα, στην Εικόνα 4.24, για την επεξηγηματική τεχνητή νοημοσύνη. Δημιουργείται ένα λειτουργικό (functional) μοντέλο το οποίο, πέρα από την τελική πρόβλεψη, επιστρέφει και το feature map της τελευταίας Conv3D στρώσης. Αυτό επιτυγχάνεται επαναχρησιμοποιώντας τα layers του αρχικού Sequential μοντέλου και καταγράφοντας το tensor που αντιστοιχεί στην τελευταία συνελκτική στρώση. Το αποτέλεσμα είναι ένα νέο μοντέλο που δίνει δύο εξόδους: την πρόβλεψη (probability για cancer) και τον χάρτη χαρακτηριστικών.

```

"""Creating a functional model that returns 2 outputs , the prediction and the feature map :"""
# Creating a new unique input tensor from which i will produce a unique computational graph with :
D,H,W= x_test.shape[1], x_test.shape[2],x_test.shape[3] # where x_test.shape[1]=32 , x_test.shape[2]=512 and x_test.shape[3]=512
inp = Input(shape=(D,H,W,1)) # inp = (32,512,512,1) , the unique input tensor
#####
# Using the layers of the sequential model as functions to process the input tensor till i reach the output tensor .
x=inp # initializing the x variable with the unique input tensor
target_tensor=None # intializing a variable in which iam gonna store the keras tensor of the final conv3d layer.
# Iterating through the Layers of the sequential model :
for layer in model.layers:
    x = layer(x) # feeding a keras tensor in each iteration to the next layer till i reach the output layer.
    if isinstance(layer,Conv3D):
        target_tensor = x # storing the output tensor of a Conv3 layer
if target_tensor is None: # just in case
    raise RuntimeError("No Conv3d layer found in the Sequential model.")
out = x # storing the final output tensor after applying all layers .
#####
# Creating the second model , a functional one with the unique input tensor and the output tensor that came out of the input one.
functional_model = Model(inputs=inp,outputs=out)
print("functional_model.output =", functional_model.output)
#####
#Combining last conv3d layer tensor and output tensor into a third model , this way i'll be gettin 2 results from a prediction , the prediction 1 or 0
# and the feature map of the last conv3d layer . Exactly what iam Looking for in a model for the Grad-Cam !
conv_model = Model(inputs=inp, outputs=[out, target_tensor]) # the model i'll be using with the Grad-Cam functionality .
""" End of Creating a functional model that returns 2 outputs , the prediction and the feature map ."""

```

Εικόνα 4.24: Δημιουργία ενός functional CNN μοντέλου

Ακολουθεί πρόβλεψη στο test set τόσο με την ενσωματωμένη μέθοδο predict, όσο και με τη χειροποίητη συνάρτηση predict_on_test_set. Η δεύτερη χρησιμοποιείται ώστε να υπάρχει πλήρης έλεγχος του τρόπου με τον οποίο εξάγονται οι πιθανότητες και οι τελικές ετικέτες, αλλά και για να διασφαλιστεί ότι η διαδικασία πρόβλεψης είναι συμβατή με τη λογική του Grad-CAM.

Τέλος, για κάθε fold εκτελείται η διαδικασία Grad-CAM στην Εικόνα 4.25. Για κάθε δείγμα του test set υπολογίζεται ένας τρισδιάστατος χάρτης ενεργοποίησης, ο οποίος αποτυπώνει ποιες περιοχές του εγκεφαλικού όγκου συνέβαλαν περισσότερο στη λήψη της απόφασης από το μοντέλο. Οι χάρτες αυτοί αποθηκεύονται τόσο ως ακατέργαστα .npy αρχεία όσο και ως έγχρωμες επικαλύψεις (overlays) πάνω σε επιλεγμένες τομές, επιτρέποντας οπτική ερμηνεία των αποτελεσμάτων. Κάθε fold έχει τον δικό του φάκελο αποτελεσμάτων, γεγονός που διευκολύνει τη σύγκριση και την οργάνωση των πειραματικών δεδομένων.

```

"""Grad-Cam :"""
print(f"=====")
print(f"Start the Grad-Cam :")
# Run GRAD-CAM on the test set and save results :
out_dir = os.path.join(GRADCAM_OUT_ROOT,f"fold_{fold_num}") # creating a different Grad-Cam folder for each fold .
os.makedirs(out_dir,exist_ok=True) # actually creating the folder .
try:
    print(f"The shape of the data iam passing to the run_and_save_gradcams_on_set :")
    print(f"{x_test.shape}")
    print(f"{y_test.shape}")
    # Running the Grad-Cam functionality :
    # --- NEW: pass test_meta so filenames & saved metadata include patient_id / dataset ---
    run_and_save_gradcams_on_set(x_test, y_test, conv_model, out_dir,
                                max_examples=GRADCAM_MAX_EXAMPLES,
                                number_of_fold=fold_num,
                                target_layer=None,
                                metadata=test_meta,
                                brain_sample=True)
except Exception as e:
    print(f"Error while running Grad-CAM for fold {fold_num}: {e}")
fold_num +=1

#####
# End of the main loop's logic #
#####

```

Εικόνα 4.25: Εκτέλεση της Grad-Cam λειτουργικότητας

Συνολικά, η παραπάνω διαδικασία ολοκληρώνει έναν πλήρη κύκλο: από τον διαχωρισμό των δεδομένων, την ενίσχυση του training set, την εκπαίδευση και αξιολόγηση του CNN, έως και την παραγωγή επεξηγηματικών χαρτών Grad-CAM. Παρότι η υλοποίηση παραμένει σχετικά απλή και σε ορισμένα σημεία «χειροποίητη», αντικατοπτρίζει μια συνειδητή προσπάθεια κατανόησης όχι μόνο του

τι προβλέπει το μοντέλο, αλλά και του *γιατί* καταλήγει σε κάθε απόφαση — στοιχείο ιδιαίτερα σημαντικό για εφαρμογές τεχνητής νοημοσύνης στον ιατρικό τομέα.

Κεφάλαιο 5ο: Αποτελέσματα του πειράματος

5.1 Αποτελέσματα επεξηγηματικής τεχνητής νοημοσύνης

Κατά την πειραματική αξιολόγηση παρατηρήθηκε υψηλή απόδοση του μοντέλου στις περισσότερες περιπτώσεις που εξετάστηκαν όπως συμβαίνει στην Εικόνα 5.1. Παράλληλα, φαίνεται ότι η αύξηση της πολυπλοκότητας του εκάστοτε προβλήματος συνδέεται με σταδιακή μείωση της απόδοσης του μοντέλου, με το δυσκολότερο σενάριο να αφορά την κατηγοριοποίηση όγκων μεταστατικού καρκίνου της σπονδυλικής στήλης. Η περιορισμένη διαθεσιμότητα δεδομένων ενδέχεται να λειτουργήσει περιοριστικά ως προς τη δυνατότητα του 3D CNN να αποτυπώσει επαρκώς τις ιδιαιτερότητες των πιο απαιτητικών περιπτώσεων, χωρίς ωστόσο να παρατηρηθεί έντονη υπερεκπαίδευση. Με άλλα λόγια, είναι πιθανό ότι η ίδια η πολυπλοκότητα των επιμέρους περιπτώσεων, σε συνδυασμό με το μικρό μέγεθος του διαθέσιμου συνόλου δεδομένων, επηρέασε καθοριστικά το τελικό accuracy, το οποίο παρέμεινε σχετικά χαμηλότερο σε αυτά τα δυσκολότερα σενάρια. Ωστόσο, ανεξαρτήτως των μετρικών απόδοσης που προέκυψαν, και στις τρεις περιπτώσεις προβλημάτων παρατηρήθηκε αξιοσημείωτη σταθερότητα στα αποτελέσματα της επεξηγηματικής τεχνητής νοημοσύνης, καθώς ο αλγόριθμος ανέδειξε επαναλαμβανόμενα παρόμοια πρότυπα και περιοχές ενδιαφέροντος σε κάθε επανεκτέλεση των τριών πειραμάτων. Ακόμα κι αν οι περιοχές αυτές δεν σχετίζονταν με καρκινικό όγκο.

```

Train on 351 samples, validate on 25 samples
Epoch 1/6
351/351 [=====] - 76s 217ms/sample - loss: 0.3174 - accuracy: 0.8746 - val_loss: 0.9101 - val_accuracy: 0.9200

Epoch 2/6
351/351 [=====] - 70s 201ms/sample - loss: 0.2311 - accuracy: 0.9630 - val_loss: 4.2493 - val_accuracy: 0.9200

Epoch 3/6
351/351 [=====] - 71s 203ms/sample - loss: 0.1575 - accuracy: 0.9772 - val_loss: 4.3046 - val_accuracy: 0.9200

Epoch 4/6
351/351 [=====] - 71s 204ms/sample - loss: 0.0829 - accuracy: 0.9829 - val_loss: 5.9694 - val_accuracy: 0.9200
preds.shape: (70, 1) fmap.shape: (70, 4, 60, 60, 64)
Confusion Matrix:
[[29  2]
 [ 0 39]]
Accuracy: 0.9714285714285714

```

	precision	recall	f1-score	support
Non-Cancer	1.00	0.94	0.97	31
Cancer	0.95	1.00	0.97	39
accuracy			0.97	70
macro avg	0.98	0.97	0.97	70
weighted avg	0.97	0.97	0.97	70

Εικόνα 5.1: Εκπαίδευση του brain-CNN_b για fold 2

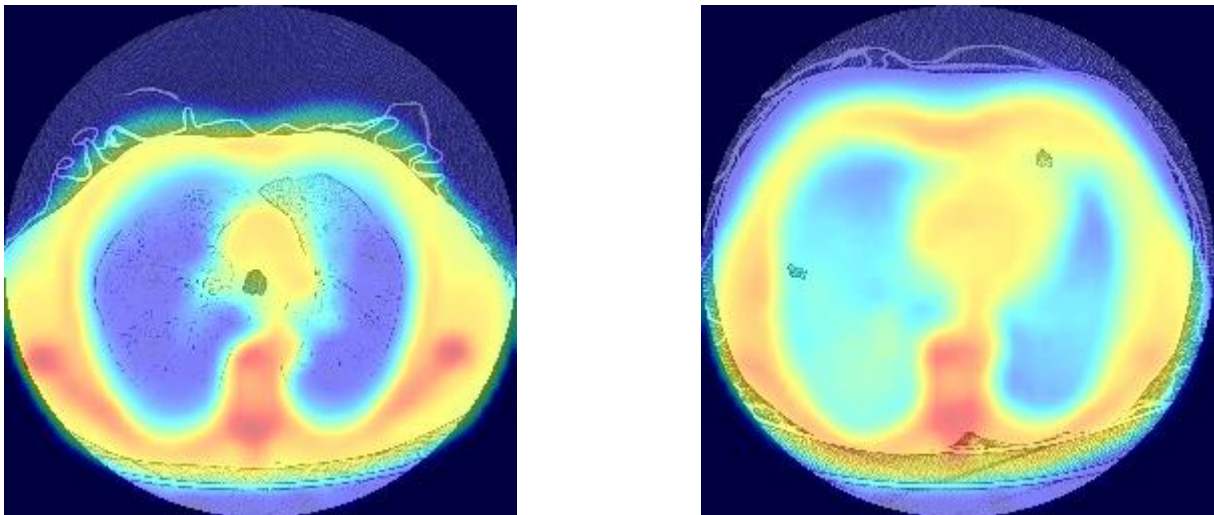
Ο έλεγχος της γενίκευσης βασίστηκε κυρίως στον ορθό διαχωρισμό των δεδομένων και στη στοχευμένη εφαρμογή data augmentation αποκλειστικά στο σύνολο εκπαίδευσης. Η διαδικασία augmentation εφαρμόστηκε μία φορά, επιτυγχάνοντας πρακτικά τον διπλασιασμό του πλήθους των δειγμάτων του training set, με στόχο την ενίσχυση της ποικιλίας χωρίς να δημιουργηθεί υπερβολική τεχνητή ομοιότητα μεταξύ των δειγμάτων. Η αξιοποίηση περισσότερων από μια αξονική ή μαγνητική για ένα δείγμα εγκαταλείφθηκε σύντομα ως ιδέα μιας και οδηγούσε σε υπερεκπαίδευση πολύ σύντομα.

Επιπλέον, λόγω της περιορισμένης διαθεσιμότητας υγιών δειγμάτων, δεν κρίθηκε σκόπιμη η εισαγωγή επιπλέον καρκινικών περιστατικών στο dataset, καθώς μια τέτοια επιλογή θα οδηγούσε σε σημαντική ανισορροπία μεταξύ των κατηγοριών και ενδεχομένως θα επηρέαζε αρνητικά τη στατιστική αξιοπιστία

και τη σταθερότητα της εκπαίδευσης. Κάτι το οποίο δοκιμάστηκε και απορρίφθηκε μιας και οδηγούσε σε υπερεκπαίδευση.

Στο πείραμα για τους πνεύμονες, η διαδικασία αξιολόγησης σχεδιάστηκε ώστε να προσεγγίζει την κλινική πραγματικότητα: το dataset (LIDC-IDRI) περιείχε ακτινολογικές εκτιμήσεις από τουλάχιστον τέσσερις ανεξάρτητους ακτινολόγους ανά εξέταση, και αυτές οι πολλαπλές εκτιμήσεις χρησιμοποιήθηκαν ως σημείο αναφοράς για την ποιοτική σύγκριση των προβλέψεων του δικτύου. Η ύπαρξη αυτής της συλλογής πολλαπλών αναγνώσεων συμβάλλει στην αύξηση της αξιοπιστίας της αξιολόγησης και επιτρέπει μια πιο ουσιαστική αντιπαραβολή (και όχι μόνο έναν αμιγώς αριθμητικό έλεγχο), καθώς το μοντέλο δεν συγκρίνεται με μία μεμονωμένη κρίση αλλά με μια μορφή κλινικής «συναίνεσης» των ειδικών.

Στην πράξη, το Lung-CNN φαίνεται να επιτυγχάνει σωστή ταξινόμηση των περιστατικών που περιείχαν καρκινικό όγκο στο συγκεκριμένο σύνολο αξιολόγησης. Ωστόσο, οι χωρικές περιοχές που αναδείχθηκαν από το Grad-CAM, όπως αυτές στην Εικόνα 5.2, δεν παρουσιάζουν πάντοτε σαφή και άμεση αντιστοίχιση με τα σημεία που είχαν επισημανθεί από τους ακτινολόγους, γεγονός που καθιστά την ερμηνεία των heatmaps σε αρκετές περιπτώσεις μη απολύτως ξεκάθαρη.

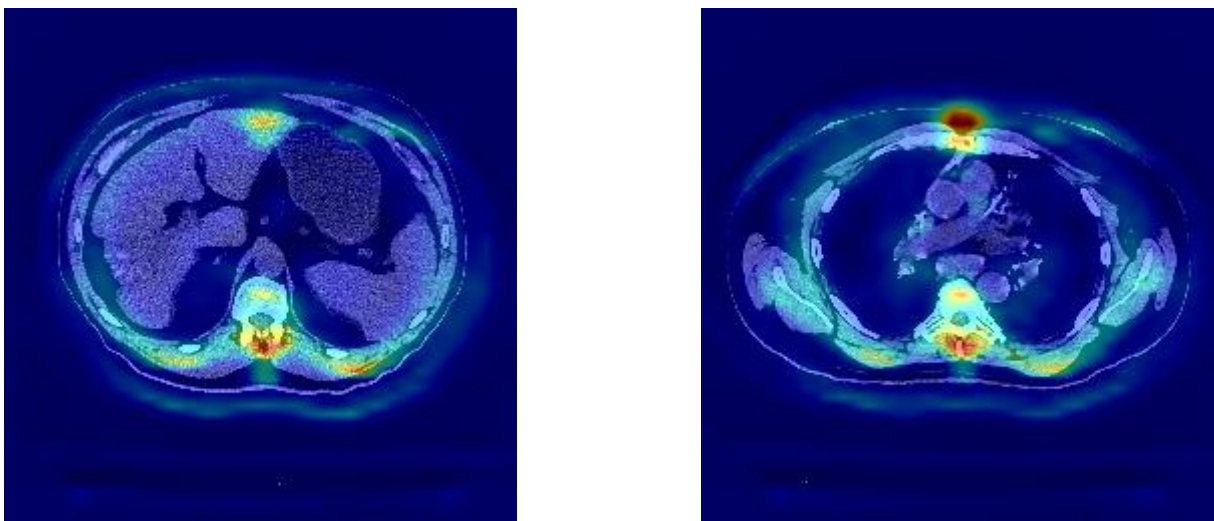


Εικόνα 5.2: slice_10&30_overlay_patient_0256_Lung_cancer_label_1

Επιπλέον, είναι πιθανό πιο λεπτομερή και τοπικά μορφολογικά μοτίβα του όγκου να ανιχνεύονται σε αρχικά συνελκτικά επίπεδα, όπου τα feature maps διατηρούν υψηλότερη χωρική ανάλυση, ενώ τα βαθύτερα επίπεδα — από τα οποία συνήθως εξάγεται το Grad-CAM — κωδικοποιούν πιο αφηρημένες και χωρικά πιο «διάχυτες» αναπαραστάσεις. Το φαινόμενο αυτό είναι σύμφωνο με τη σχετική βιβλιογραφία των CNN και δεν συνεπάγεται απαραίτητα λανθασμένη λειτουργία του μοντέλου, αλλά αναδεικνύει ότι τα heatmaps αποτυπώνουν περιοχές που συμβάλλουν περισσότερο στην τελική απόφαση και όχι κατ' ανάγκη το ακριβές περίγραμμα της βλάβης.

Για την περίπτωση της σπονδυλικής στήλης επιλέχθηκε ένα dataset με επιβεβαιωμένες μεταστάσεις, προκειμένου να διερευνηθεί κατά πόσο το CNN μπορεί να ανταποκρίνεται σε διαφορετικούς τύπους μεταστατικών εστιών. Παρότι το μοντέλο εμφάνισε την χαμηλότερη απόδοση στο training set, η αξιολόγησή του σε ανεξάρτητα δείγματα του test set έδειξε ότι μπορούσε να αναγνωρίσει μεταστάσεις που σχετίζονταν με διαφορετικά πρωτογενή νεοπλασμάτα, όπως χοληδόχο κύστη, στο παράδειγμα patient_43_bile_ducts. Στο δείγμα patient_43, για το οποίο τα αριθμητικά αποτελέσματα

παρουσιάζονται στην Εικόνα 5.4, η κλινική αναφορά τοποθετεί τη βλάβη στο επίπεδο T10 και το Grad-CAM ανέδειξε περιοχή ενδιαφέροντος κοντά σε αυτό το σημείο. Σημειώνεται ότι η αρίθμηση των slices ξεκινά από το ανώτερο προς το κατώτερο τμήμα της απεικόνισης. Ενδεικτικά, στον ασθενή patient_43 ο καρκινικός εντοπισμός εμφανίζεται στο 15ο slice, όπως φαίνεται στην Εικόνα 5.6, σε σύνολο 32 slices ανά ασθενή. Σύμφωνα με τη γνωμάτευση των ακτινολόγων, η βλάβη εντοπίζεται στον σπόνδυλο T10, ο οποίος αντιστοιχεί περίπου στο κεντρικό τμήμα της εξεταζόμενης περιοχής της σπονδυλικής στήλης. Στον χάρτη Grad-CAM, όπου παρουσιάζεται στην Εικόνα 5.3, παρατηρείται η ισχυρότερη ενεργοποίηση στην ίδια γενική περιοχή, με την εντονότερη απόκριση να εμφανίζεται στο 10ο και να κορυφώνεται στο 15ο slice, και να μειώνεται σταδιακά έως περίπου το 22ο slice. Η παρατήρηση αυτή δεν αποδεικνύει από μόνη της ακριβή κλινική αντιστοίχιση, ωστόσο υποδηλώνει ότι η απόφαση του μοντέλου σχετίζεται με χωρικά συνεπή περιοχή γύρω από τη βλάβη.

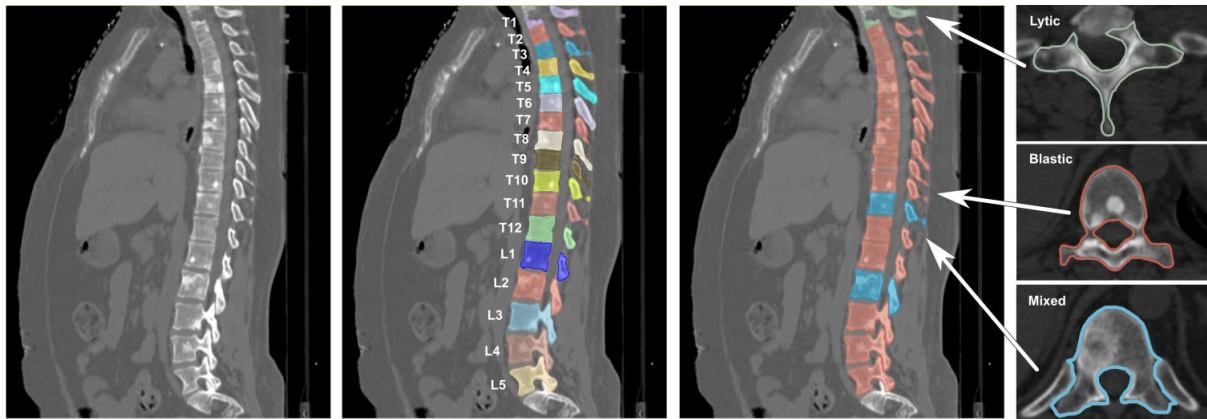


Εικόνα 5.3: XAI για patient 43 , slices 10 & 22 με bile ducts και spine καρκίνο

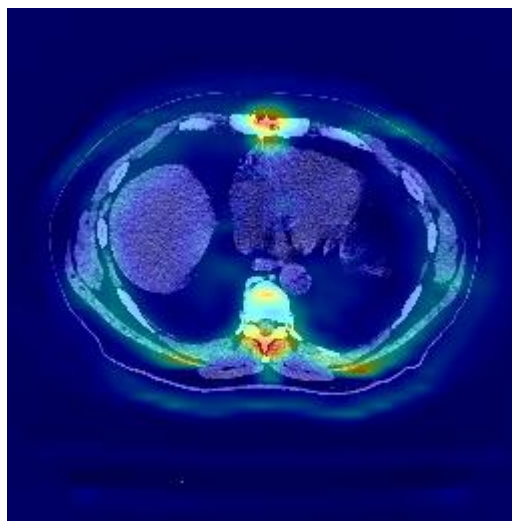
```
SAMPLE-7 :

Turned sample 7 shape from (32,256,256,1) to (32, 256, 256), label: 1
=====
Iam executing the code inside of the compute_3d_gradcam_for_volume :
Forward pass through the CNN :
Got prediction=[[54.50567]] and feature_map=(1, 4, 60, 60, 64) from forwarding
tape.watch(model.inputs worked...)
Got the logit : [[54.50567]]
=====
Computing Grads :
GRADS mean: 0.0012398815
GRADS max: 0.14124086
=====
Taking the average value of grads for a certain feature map meaning the weights :
weights min/max: -0.37431994 1.0
Feature map min/max: 0.0 2.9495363 dtype: <dtype: 'float32'>
Multiplying every spot(1, 4, 60, 60, 64) with each weight(64,) that comes from a channel(64):
raw_heat_map min/max (before ReLU): 0.083573624 4.08443
CAM pre-ReLU values range:0.08357362449169159 to 4.08443021774292
Iam gonna apply relu on the volume
CAM post-ReLU values range:0.08357362449169159 to 4.08443021774292
conv_outputs.shape = (1, 4, 60, 60, 64)
=====
weights any nan? False
Heatmap value type before processing it :float32
=====
```

Εικόνα 5.4: output για patient 43 fold 2

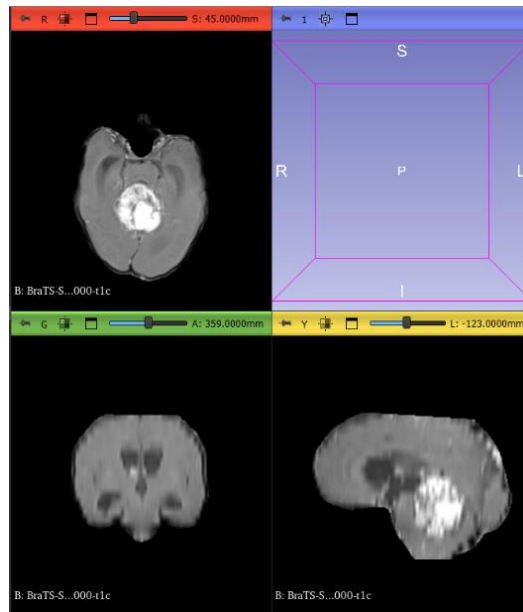


Εικόνα 5.5: αρίθμηση των σπόνδυλων

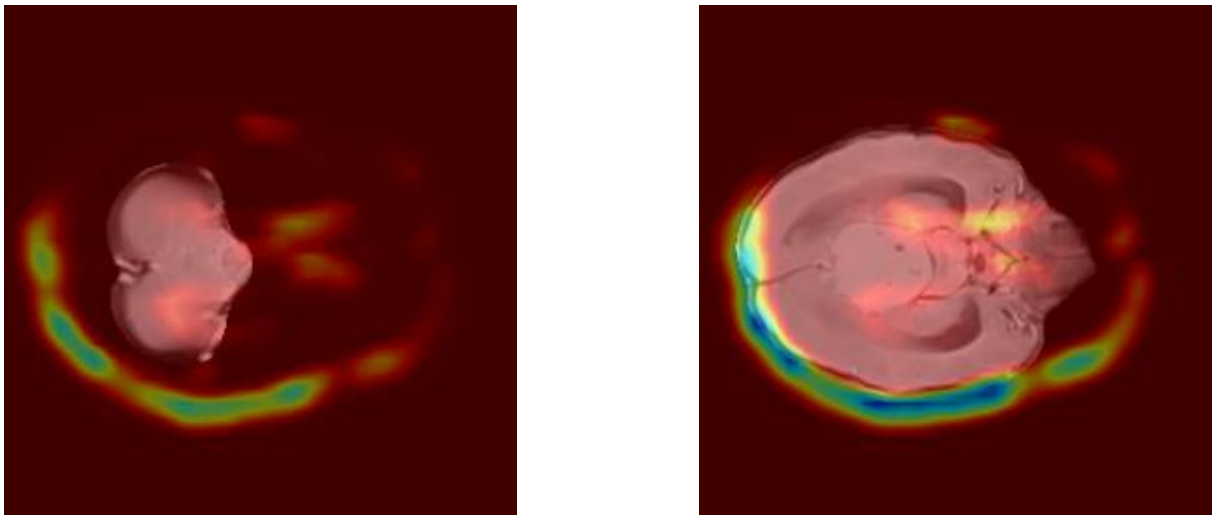


Εικόνα 5.6: XAI για patient 43 , slice 15 (T10) με spine καρκίνο

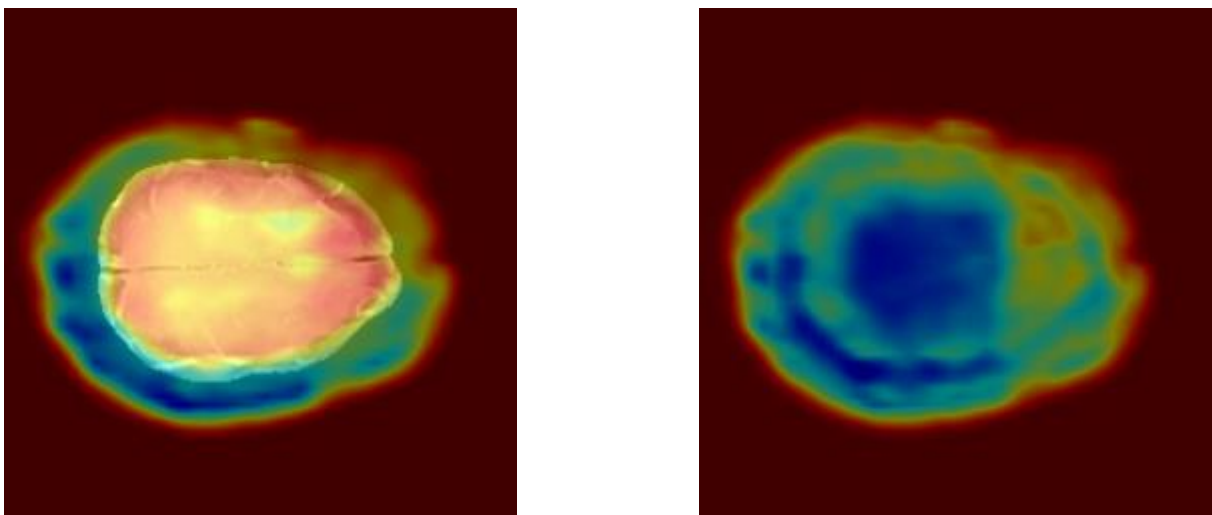
Στην περίπτωση του εγκεφάλου παρατηρήθηκε διαφορετική συμπεριφορά σε σχέση με τα υπόλοιπα πειράματα, καθώς το CNN εκπαιδεύτηκε σε μεγαλύτερο πλήθος δεδομένων, το οποίο έφτασε κοντά στα διαθέσιμα υπολογιστικά όρια. Τα heatmaps των Εικόνων 5.8 και 5.9 ανέδειξαν επίσης ένα ενδιαφέρον, αλλά όχι απόλυτα ερμηνεύσιμο, μοτίβο: σε ορισμένα δείγματα η περιοχή του όγκου δεν εμφανιζόταν ως η εντονότερα ενεργοποιημένη, ενώ η ισχυρότερη απόκριση εντοπιζόταν γύρω από τον περιβάλλοντα ιστό. Μια πιθανή ερμηνεία είναι ότι δεν αποκλείεται μέρος της συμπεριφοράς αυτής να σχετίζεται με το γεγονός ότι τα gradients και οι εσωτερικές αναπαραστάσεις του δικτύου μπορούν να έχουν διαφορετικό πρόσημο πριν από μη γραμμικές ενεργοποιήσεις, όπως η ReLU, γεγονός που επηρεάζει το πώς αποτυπώνεται η ενεργοποίηση στα heatmaps. Για λόγους σύγκρισης, στην Εικόνα 5.7 παρουσιάζεται ο καρκινικός όγκος του ασθενή 105.



Εικόνα 5.7: Ο καρκινικός όγκος του patient_105



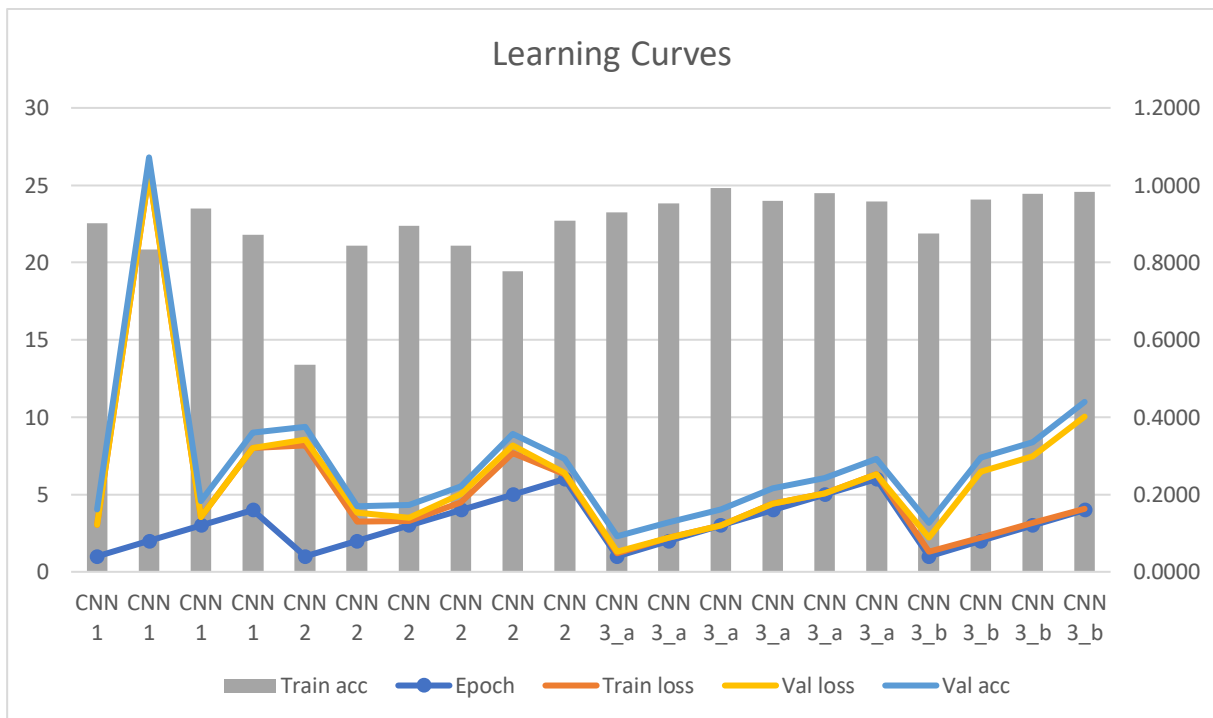
Σχήμα 5.8: ΧΑΙ για τον patient_105_PKG_-_BraTS-Africa για fold_3 , slices = 3,11



Εικόνα 5.9: XAI για τον patient_105_PKG_-_BraTS-Africa για fold_3 , slices = 25,31

5.2 Συνολικά αποτελέσματα

Στον Πίνακα 5.1 των καμπυλών μάθησης παρατηρείται ότι το CNN επιτυγχάνει γενικά υψηλές επιδόσεις κατά την εκπαίδευση σε όλες τις εξεταζόμενες περιπτώσεις, γεγονός που υποδηλώνει ότι το μοντέλο προσαρμόζεται αποτελεσματικά στα δεδομένα εκπαίδευσης. Ωστόσο, οι τιμές επικύρωσης παρουσιάζουν διαφοροποιήσεις μεταξύ των περιπτώσεων, με ορισμένα σύνολα δεδομένων να οδηγούν σε πιο σταθερή συμπεριφορά και άλλα σε μεγαλύτερη απόκλιση μεταξύ των επιδόσεων εκπαίδευσης και επικύρωσης. Το εύρημα αυτό υποδεικνύει ότι η ικανότητα γενίκευσης του ίδιου μοντέλου επηρεάζεται από τα χαρακτηριστικά και τη δυσκολία των δεδομένων που χρησιμοποιούνται σε κάθε περίπτωση.



Πίνακας 5.1: Learning curves των τριών CNN

Ο Πίνακας 5.2 παρουσιάζει τις βασικές μετρικές αξιολόγησης για τις τέσσερις αρχιτεκτονικές CNN που χρησιμοποιήθηκαν στα επιμέρους προβλήματα ταξινόμησης. Πιο συγκεκριμένα, το CNN1 εμφάνισε την υψηλότερη συνολική απόδοση, επιτυγχάνοντας accuracy ίσο με 1, καθώς και μηδενικές τιμές ψευδώς θετικών (FP) και ψευδώς αρνητικών (FN) προβλέψεων. Το αποτέλεσμα αυτό είναι πιθανό να σχετίζεται με το γεγονός ότι το συγκεκριμένο μοντέλο κλήθηκε να επιλύσει το απλούστερο πρόβλημα ταξινόμησης, δηλαδή την ανίχνευση παρουσίας όγκου ανεξαρτήτως κακοήθειας. Αντίθετα, το CNN2 παρουσίασε αισθητά χαμηλότερη απόδοση, με accuracy 0,77 και recall 0,7143, γεγονός που υποδηλώνει μεγαλύτερη δυσκολία στην ορθή αναγνώριση όλων των θετικών περιπτώσεων. Η εικόνα αυτή φαίνεται να είναι συμβατή με τη φύση του προβλήματος που εξετάστηκε, καθώς το συγκεκριμένο μοντέλο εφαρμόστηκε στην κατηγοριοποίηση όγκων μεταστατικού καρκίνου της σπονδυλικής στήλης, ένα σαφώς πιο σύνθετο και απαιτητικό διαγνωστικό σενάριο. Τα CNN3_a και CNN3_b, τα οποία χρησιμοποιήθηκαν για τη διάκριση καρκινικών όγκων στον εγκέφαλο και για την διάκριση καλοήθων και κακοήθων βλαβών αντίστοιχα, παρουσίασαν επίσης ιδιαίτερα υψηλές επιδόσεις. Το CNN3_a κατέγραψε accuracy 0,989 με μηδενικά ψευδώς θετικά αποτελέσματα, ενώ το CNN3_b εμφάνισε

accuracy 0,971 και την υψηλότερη τιμή precision (0,9512), γεγονός που υποδηλώνει αυξημένη αξιοπιστία στις θετικές προβλέψεις του μοντέλου.

Το Spine CNN φαίνεται να είναι το πιο ισορροπημένο από τα τρία, καθώς το σύνολο των κλάσεων ήταν πιο συμμετρικό σε σχέση με τα άλλα πειράματα και η ποσότητα της πληροφορίας ήταν περισσότερη μιας και ένα αρχικό 3d-volume του Spine-mets dataset αποτελείται από περίπου 1100 slices κατά μέσο όρο ενώ ένα μέσο 3d-volume αποδίδεται με 800 slices από τα υπόλοιπα datasets. Αυτό πιθανόν συνέβαλε στη σταθερότερη συμπεριφορά του μοντέλου και στην πιο αξιόπιστη ερμηνεία των αποτελεσμάτων του σε σχέση με τα υπόλοιπα δύο CNN. Συνεπώς, η συνολική αξιολόγηση δεν πρέπει να βασιστεί μόνο στις απόλυτες τιμές των μετρικών, αλλά να εξεταστεί σε συνδυασμό με τη δομή των δεδομένων, τον τρόπο δημιουργίας των splits και τα αποτελέσματα της επεξηγηματικής ανάλυσης.

	TP	TN	FP	FN	Accuracy	Precision	Recall	F1-score	Specificity
CNN 1	17	10	0	0	1	1	1	1	1
CNN 2	10	7	1	4	0,77	0,9091	0,7143	0,8	0,875
CNN 3_a	49	42	0	1	0,989	0,9091	0,7143	0,8	0,875
CNN 3_b	39	29	2	0	0,971	0,9512	1	0,975	0,9355

Πίνακας 5.2: Αποτελέσματα των τριών CNN

5.3 Σύνοψη

Στο πλαίσιο της παρούσας διπλωματικής εργασίας, αναπτύχθηκαν και αξιολογήθηκαν τρία τρισδιάστατα CNN μοντέλα για διαφορετικά ιατρικά σύνολα δεδομένων, με σκοπό την κατηγοριοποίηση και την επεξηγηματική ανάλυση των αποτελεσμάτων μέσω Grad-CAM. Συγκεκριμένα, για το CNN 1 χρησιμοποιήθηκαν 135 δείγματα εκπαίδευσης(78,49%), μετά από augmentation, 10 δείγματα για validation(5,81%) και 27 δείγματα αξιολόγησης(15,70%), για το CNN 2 153 δείγματα εκπαίδευσης(81,82%), μετά από augmentation, 12 δείγματα για validation(6,42%), και 22 δείγματα αξιολόγησης(11,76%), ενώ για το CNN 3_a 453 δείγματα εκπαίδευσης(78,37%), μετά από augmentation, 33 δείγματα για validation(5,71%), και 92 δείγματα αξιολόγησης(15,92%), και για το CNN 3_b 351 δείγματα εκπαίδευσης(78,70%), μετά από augmentation, 25 δείγματα για validation(5,61%), και 70 δείγματα αξιολόγησης(15,70%). Η ποσότητα των δεδομένων ήταν ανάλογη και για τις δύο κλάσεις πράγμα που επέφερε τα καλύτερα αποτελέσματα. Προηγουμένως συνέβησαν απόπειρες να αυξηθούν τα δεδομένα με την εισαγωγή επιπρόσθετων αξονικών για ένα δείγμα καθώς και χρήση άνισων ποσοτικά δεδομένων για τις δυο κλάσεις κάτι το οποίο έφερε μη επιθυμητά αποτελέσματα.

Τα αριθμητικά αποτελέσματα έδειξαν υψηλή απόδοση σε όλα τα πειράματα, όμως η ερμηνεία τους δεν θα πρέπει να περιοριστεί μόνο στις τιμές της ακρίβειας. Όπως προκύπτει και από την ανάλυση των learning curves, τα διαθέσιμα δεδομένα επιτρέπουν να γίνει λόγος για σταθερή εκπαίδευση και καλή συμπεριφορά στα συγκεκριμένα splits, χωρίς όμως να τεκμηριώνεται από μόνα τους με απόλυτη βεβαιότητα η απουσία υπερεκπαίδευσης. Για τον λόγο αυτό, η αξιολόγηση των μοντέλων συνδυάστηκε με την επεξηγηματική ανάλυση των Grad-CAM heatmaps, ώστε να εξεταστεί κατά πόσο οι περιοχές που αναδεικνύονται από το δίκτυο παρουσιάζουν μια εύλογη αντιστοίχιση με τις κλινικά αναμενόμενες περιοχές ενδιαφέροντος. Η ανάλυση αυτή προσφέρει μια πιο ουσιαστική εικόνα της λειτουργίας των

μοντέλων, χωρίς όμως να επιτρέπει από μόνη της τον απόλυτο χαρακτηρισμό τους ως «σωστών» ή «κλινικά επιβεβαιωμένων».

Ανάμεσα στα τρία πειράματα, το spine CNN εμφανίζεται ως το πιο ισορροπημένο ως προς τη συνολική συμπεριφορά του, καθώς το accuracy του παρέμεινε στο 0.77, κάτι που μπορεί να θεωρηθεί ενδεικτικό μιας πιο ρεαλιστικής απόδοσης σε σχέση με περιπτώσεις τέλει ταξινόμησης. Παράλληλα, τα αντίστοιχα Grad-CAM αποτελέσματα φαίνεται να προσφέρουν πιο πειστική ποιοτική ερμηνεία σε σχέση με τα υπόλοιπα μοντέλα, καθώς οι περιοχές ενεργοποίησης παρουσιάζουν σε αρκετές περιπτώσεις μεγαλύτερη εγγύτητα με τις περιοχές ενδιαφέροντος. Ωστόσο, το γεγονός αυτό πρέπει να διατυπωθεί με προσοχή: δεν αποδεικνύει ότι το μοντέλο «μαθαίνει» αποκλειστικά κλινικά χρήσιμα χαρακτηριστικά, αλλά υποδηλώνει ότι η τελική του απόφαση μπορεί να σχετίζεται με πληροφορία που έχει κάποια ερμηνευτική αξία.

Συνολικά, η παρούσα εργασία καταδεικνύει ότι τα 3D CNN μπορούν να επιτύχουν πολύ υψηλή ταξινομητική απόδοση στα εξεταζόμενα σύνολα δεδομένων, ενώ η χρήση Grad-CAM συμβάλλει στην καλύτερη κατανόηση της λειτουργίας τους. Παράλληλα, η ανάλυση ανέδειξε ότι η υψηλή ακρίβεια δεν αρκεί από μόνη της για να τεκμηριώσει πλήρως τη γενικευσιμότητα ή την κλινική αξιοπιστία ενός μοντέλου. Η πιο συνετή ερμηνεία των αποτελεσμάτων δείχνει ότι η συνδυαστική αξιολόγηση ποσοτικών μετρικών και επεξηγηματικών τεχνικών είναι απαραίτητη για μια πιο ισορροπημένη και επιστημονικά τεκμηριωμένη αποτίμηση των CNN σε ιατρικές εφαρμογές.

Βιβλιογραφία

- [1] W. S. McCulloch and W. Pitts, “A logical calculus of ideas immanent in nervous activity,” *Bull. Math. Biophys.*, vol. 5, pp. 115–133, 1943.
- [2] F. Rosenblatt, “The perceptron: A probabilistic model for information storage and organization in the brain,” *Psychol. Rev.*, vol. 65, no. 6, pp. 386–408, Nov. 1958.
- [3] M. Minsky and S. Papert, *Perceptrons*. Cambridge, MA: MIT Press, 1969.
- [4] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” in *Proc. Advances in Neural Information Processing Systems (NIPS)*, 2012, pp. 1097–1105.
- [5] G. E. Moore, “Cramming more components onto integrated circuits,” *Electronics*, vol. 38, pp. 114–117, 1965.
- [6] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.
- [7] G. M. Church, Y. Gao, and S. Kosuri, “Next-generation digital information storage in DNA,” *Science*, vol. 337, no. 6102, Aug. 2012. DOI: 10.1126/science.1226355.
- [8] N. Goldman, P. Bertone, S. Chen, C. Dessimoz, E. M. LeProust, B. Sipos, and E. Birney, “Towards practical, high-capacity, low-maintenance information storage in synthesized DNA,” *Nature*, vol. 494, no. 7435, pp. 77–80, Feb. 2013.
- [9] A. H. Marblestone, G. Wayne, and K. P. Kording, “Toward an integration of deep learning and neuroscience,” *Front. Comput. Neurosci.*, vol. 10, art. 94, Sept. 2016. DOI: 10.3389/fncom.2016.00094.
- [10] D. O. Hebb, *The Organization of Behavior: A Neuropsychological Theory*. New York: John Wiley & Sons, 1949.
- [11] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, “Learning representations by back-propagating errors,” *Nature*, vol. 323, pp. 533–536, Oct. 1986.
- [12] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghahfarokian, J. A. W. M. van der Laak, B. van Ginneken, and C. I. Sánchez, “A survey on deep learning in medical image analysis,” *Med. Image Anal.*, vol. 42, pp. 60–88, 2017.
- [13] Y. LeCun, L. Bottou, Y. Bengio and P. Haffner, “Gradient-based learning applied to document recognition,” *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.
- [14] G. N. Hounsfield, “Computerized transverse axial scanning (tomography). Part I. Description of system,” *Br. J. Radiol.*, vol. 46, no. 552, pp. 1016–1022, Dec. 1973.
- [15] S. K. Zhou, H. Greenspan, and D. Shen (eds.), *A review of deep learning in medical imaging — (review articles & collection)*, 2021. (See: review/overview articles on DL for medical imaging.)
- [16] L. R. Koetzier, D. Mastrodicasa, T. P. Szczykutowicz, N. R. van der Werf, A. S. Wang, V. Sandfort, A. J. van der Molen, D. Fleischmann, and M. J. Willeminck, “Deep learning image reconstruction for

CT: Technical principles and clinical prospects,” *Radiology*, vol. 306, no. 3, p. e221257, Mar. 2023, doi: 10.1148/radiol.221257.

[17] Menze, B. H., et al., “The Multimodal Brain Tumor Image Segmentation Benchmark (BRATS),” *IEEE Transactions on Medical Imaging*, 2015.

[18] S. G. Armato III et al., “The lung image database consortium (LIDC) and image database resource initiative (IDRI): A completed reference database of lung nodules on CT,” *Medical Physics*, vol. 38, no. 2, pp. 915–931, Feb. 2011, doi: 10.1118/1.3528204.

[19] Y. A. LeCun, L. Bottou, G. B. Orr, and K.-R. Müller, “Efficient BackProp,” in *Neural Networks: Tricks of the Trade*, G. Montavon, G. B. Orr, and K.-R. Müller, Eds., *Lecture Notes in Computer Science*, vol. 7700, Springer, 2012, pp. 9–48.

[20] L. Gavrilas and K. T. Kotsis, “Electromagnetic radiation: A comprehensive review of misconceptions,” *Eurasian Journal of Science and Environmental Education*, vol. 4, no. 2, pp. 19–38, 2024, doi: 10.30935/ejsee/15719.

[21] OpenStax, *College Physics 2e*, Chapter “Photon Energies and the Electromagnetic Spectrum.” [Online]. Accessed: Feb. 9, 2026.

[22] A. Aiello, “A probabilistic view of wave-particle duality for single photons,” arXiv:2303.15185 [quant-ph], Oct. 2023. [Online]. Available: <https://arxiv.org/abs/2303.15185>

[23] X. Ou et al., “Recent Development in X-Ray Imaging Technology: Future and Challenges,” *Research* (Wash. D C), 2021.

[24] H. H. Barrett and W. Swindell, *Radiological Imaging: The Theory of Image Formation, Detection, and Processing*, 2nd ed. San Diego, CA, USA: Academic Press, 2010.

[25] A. Lipovka and I. Cardenas, “Duane-Hunt relation improved,” arXiv:1506.08783 [physics.gen-ph], Jun. 2015. [Online]. Available: <https://arxiv.org/abs/1506.08783>

[26] J. A. Bushberg, J. M. Boone, and E. J. Seibert, “The essential physics of medical imaging revisited,” *Med. Phys.*, vol. 47, no. 3, pp. e1–e20, 2020.

[27] D. Banks et al., “X-ray interaction and the electronic, atomic cross-sections and Compton mass-attenuation coefficients of human blood, breasts, eye lens, ovaries, and testis,” *Radiation*, vol. 5, Art. no. 24, 2025, doi: 10.3390/radiation5030024.

[28] B. B. Balázs, D. Laczkó, D. Gergő, P. Hegyi, B. Merkely, I. F. Édes, et al., “Radiation exposure risk for various personnel in the catheterization laboratory: A systematic review and meta-analysis,” *Research*, vol. 137, Art. no. 105096, Sep. 2025, doi: 10.34133/2025/105096.

[29] M. Berger, Q. Yang, and A. Maier, “X-ray Imaging,” in *Medical Imaging Systems: An Introductory Guide* [Internet]. [Online]. Available: [X-ray Imaging - Medical Imaging Systems - NCBI Bookshelf](#) Published online: Aug. 3, 2018. Accessed: Feb. 23, 2026.

[30] National Institute of Standards and Technology (NIST), “X-Ray Mass Attenuation Coefficients,” *NIST Standard Reference Database 126*, 2025. [Online]. Accessed: Feb. 9, 2026.

[31] M. Baur, N. Uhlmann, T. Pöschel, and M. Schröter, “Correction of beam hardening in X-ray radiograms,” arXiv:1811.04102 [physics.ins-det], Nov. 2018. [Online]. Available: <https://arxiv.org/abs/1811.04102>. Accessed: Feb. 9, 2026.

- [32] I. A. Elbakri, Y. Zhang, L. Chen, N. H. Clinthorne, and J. A. Fessler, “Statistical Reconstruction for Quantitative CT Applications,” in Proc. IEEE Nuclear Science Symp. Conf. Record, 2003.
- [33] H. C. Manjunatha and L. Seenappa, “Pocket formula for mass attenuation coefficient, effective atomic number, and electron density of human tissues,” Nuclear Science and Techniques, vol. 30, no. 36, 2019.
- [34] A. C. Kak and M. Slaney, Principles of Computerized Tomographic Imaging. New York, NY, USA: IEEE Press, 1988.
- [35] Gevdeep Bhabra, Samuel Mackeith, Pedro Monteiro, and David D. Pothier, “An experimental comparison of handover methods,” Annals of the Royal College of Surgeons of England, vol. 89, no. 3, pp. 298–300, Apr. 2007, doi: 10.1308/003588407X168352.
- [36] J. Anthony Seibert, “Flat-panel detectors: how much better are they?,” Pediatric Radiology, vol. 36, Suppl. 2, pp. 173–181, Jul. 2006, doi: 10.1007/s00247-006-0208-0.
- [37] J. H. Siewerdsen, M. J. Daly, B. Bakhtiar, D. J. Moseley, S. Richard, H. Keller, and D. A. Jaffray, “A simple, direct method for x-ray scatter estimation and correction in digital radiography and cone-beam CT,” Medical Physics, vol. 33, no. 1, pp. 187–197, Jan. 2006, doi: 10.1118/1.2148916.
- [38] Arthur Burgess, “Signal Detection Theory: A Brief History,” in Part I — Historical Reflections and Theoretical Foundations, E. Samei and E. A. Krupinski, Eds., Cambridge University Press, published online Dec. 20, 2018. Accessed: Feb. 9, 2026.
- [39] E. Samei et al., Performance Evaluation of Computed Tomography Systems, Report of AAPM Task Group 233, Alexandria, VA, USA: American Association of Physicists in Medicine, AAPM Report No. 233, Apr. 2019.
- [40] G. Vegas-Sánchez-Ferrero and R. San José Estépar, “Statistical characterization of the linear attenuation coefficient in polychromatic CT scans,” Medical Physics, vol. 47, no. 11, pp. 5568–5581, Nov. 2020, doi: 10.1002/mp.14384.
- [41] S. A. Kane και B. A. Gelman, Εισαγωγή στη Φυσική της Σύγχρονης Ιατρικής, 3η έκδ., Αθήνα: Εκδόσεις Ροτόντα, 2022.
- [42] T. Arvanitidou-Vayiona and A.-B. Haidits, Ιατρική Στατιστική: Βασικές Αρχές, Thessaloniki, Greece: University Studio Press, 2013.