



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ
ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ
ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

*Ραδιομική Ανάλυση Εικόνων Μαγνητικής
Τομογραφίας (MRI) με σκοπό την Ανάπτυξη
Μοντέλου Αναγνώρισης και Ταξινόμησης των
βλαβών στη Σπονδυλική Στήλη*

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Νικόλαου Φ. Ρομποτή

Επιβλέπων: Γεώργιος Ματσόπουλος

Καθηγητής ΕΜΠ

Αθήνα, Απρίλιος 2024

Η σελίδα αυτή είναι σκόπιμα λευκή.



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ
ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ
ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

*Ραδιομική Ανάλυση Εικόνων Μαγνητικής
Τομογραφίας (MRI) με σκοπό την Ανάπτυξη
Μοντέλου Αναγνώρισης και Ταξινόμησης των
βλαβών στη Σπονδυλική Στήλη*

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Νικόλαου Φ. Ρομποτή

Επιβλέπων: Γεώργιος Ματσόπουλος

Καθηγητής ΕΜΠ

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 23^η Απριλίου 2024

.....

Γεώργιος Ματσόπουλος

Καθηγητής ΕΜΠ

.....

Αθανάσιος Παναγόπουλος

Καθηγητής ΕΜΠ

.....

Παναγιώτης Τσανάκας

Καθηγητής ΕΜΠ

Αθήνα, Απρίλιος 2024

.....
Νικόλαος Φ. Ρομποτής

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών ΕΜΠ

Copyright ©Νικόλαος Ρομποτής, 2024

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα. Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Περίληψη

Η παρούσα διπλωματική εργασία επικεντρώνεται στην ανάλυση εικόνων Μαγνητικής Τομογραφίας (MRI) και την περαιτέρω ανάπτυξη ενός μοντέλου για την ταυτοποίηση και ταξινόμηση βλαβών στον μυελό των οστών της σπονδυλικής στήλης λόγω πολλαπλού μυελώματος. Η κύρια δομή της διπλωματικής εργασίας αποτελείται από τρία μέρη: το θεωρητικό υπόβαθρο, τα υλικά και μεθοδολογία, και τα αποτελέσματα, περιορισμοί, συμπεράσματα και προτάσεις για περαιτέρω έρευνα. Το πρώτο μέρος περιλαμβάνει την ανάλυση της τεχνολογίας MRI, τις βασικές αρχές λειτουργίας της και τις μεθόδους παραγωγής των εικόνων που χρησιμοποιούνται. Επιπλέον, περιγράφεται η παθοφυσιολογία του πολλαπλού μυελώματος, μαζί με τις τεχνικές απεικόνισης που χρησιμοποιούνται για τη διάγνωση. Επισημαίνεται ο ρόλος των νευρωνικών δικτύων και των μεθόδων τμηματοποίησης εικόνων στην ακριβή ερμηνεία των δεδομένων MRI, μαζί με τις μεθόδους εγγραφής και ταξινόμησης εικόνων. Το δεύτερο μέρος περιλαμβάνει την περιγραφή της ερευνητικής μεθοδολογίας από την περιγραφή του σετ δεδομένων και την προεπεξεργασία δεδομένων, μέχρι την επιλογή λογισμικού και την αναλυτική διαδικασία που ακολουθήθηκε για την απομόνωση του μυελού των οστών της σπονδυλικής στήλης μέσω δικτύων U-net, τη μεθοδολογία εγγραφής εικόνων, την απομόνωση βλαβών μέσω τεχνικών τμηματοποίησης και την ταξινόμηση των βλαβών της σπονδυλικής στήλης. Τα αποτελέσματα της ερευνητικής διαδικασίας αναλύονται, και συζητείται η αποτελεσματικότητα του αναπτυγμένου μοντέλου στην ταυτοποίηση και ταξινόμηση των βλαβών, με συστάσεις για περαιτέρω έρευνα και συμπεράσματα. Σκοπός της εργασίας είναι να συμβάλει στην βελτίωση των διαγνώσεων και των περιπτώσεων ταξινόμησης του πολλαπλού μυελώματος μέσω της ανάλυσης MRI.

Λέξεις κλειδιά: Μαγνητική τομογραφία, MRI, πολλαπλό μυέλωμα, βαθιά νευρωνικά δίκτυα, U-net, ταξινόμηση, support vector machine

Η σελίδα αυτή είναι σκόπιμα λευκή.

Abstract

The present thesis is oriented on the analysis of Magnetic Resonance Imaging (MRI) and further development of the model for lesion identification and classification in the bone marrow of the spinal column due to multiple myeloma. The main thesis structure consists of three parts: theoretical background, materials and methodology, and results, limitations, conclusions, and suggestions for further research. The first part includes the analysis of MRI technology, its basic principles of operation, and the methods of image production used. Moreover, the pathophysiology of multiple myeloma is described, along with the imaging techniques used for diagnosis. The role of neural networks and image segmentation methods in the accurate interpretation of MRI data is emphasized, together with image registration and classification methods. The second part includes the description of the research methodology starting from the dataset description and data preprocessing, up to the software selection and the analytical process followed for the isolation of the spinal column bone marrow via U-net networks, the methodology of image registration, lesion isolation through image segmentation techniques, and the classification of the spinal column lesions. The results of the investigating process are analysed, and the developed system's efficiency in lesion identification and classification is discussed, with recommendations for further research and conclusions. The aim of the work is to contribute to the improvement of multiple myeloma diagnoses and classification cases by MRI analysis.

Keywords: Magnetic resonance imaging, MRI, multiple myeloma, deep neural networks, U-net, classification, support vector machine

Η σελίδα αυτή είναι σκόπιμα λευκή.

Table of Contents

Περίληψη	5
Abstract	7
Ευρετήριο Εικόνων	13
Ευρετήριο Εξιιώσεων	16
Ευρετήριο Πινάκων	18
ΜΕΡΟΣ 1: ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ	20
Κεφάλαιο 1: Η μαγνητική τομογραφία.....	20
1.1 Αρχές Λειτουργίας	20
1.2 Χρόνοι Μαγνητικής Αποκατάστασης	22
1.3 Παραγωγή Ακολουθιών MRI.....	25
1.3.1 Παραγωγή της εικόνας.....	25
1.3.2 Επιλογή των χρόνων TR, TE:	26
1.4 Τα Μέρη του Τομογράφου	26
1.5 Ασφάλεια της Μαγνητικής Τομογραφίας.....	27
2 Κεφάλαιο 2: Παθοφυσιολογία του Πολλαπλού Μυελώματος και Απεικονιστικές τεχνικές.....	30
2.1 Φυσιολογία του αίματος και του ανοσιακού συστήματος.....	30
2.1.1 Αίμα.....	32
2.2 Πολλαπλό μυέλωμα.....	35
2.2.1 Εισαγωγή.....	35
2.2.2 Αίτια και διάγνωση ΠΜ.....	37
3 Κεφάλαιο 3: Τμηματοποίηση Εικόνας.....	46
3.1 Κλασικές Τεχνικές.....	46
3.1.1 Κατωφλιοποίηση Εικόνας.....	47
3.1.2 k-means segmentation	48
3.1.3 Αλγόριθμος Τμηματοποίησης Otsu	50
3.1.4 Πολλαπλή Κατωφλιοποίηση Otsu	51
3.2 Νευρωνικά Δίκτυα.....	53
3.2.1 Βασικά Συστατικά και Λειτουργικοί Μηχανισμοί.....	53
3.2.2 Συναρτήσεις Ενεργοποίησης: Εισαγωγή της Μη Γραμμικότητας στα CNNs	54

3.2.3	Προηγμένα Αρχιτεκτονικά Χαρακτηριστικά στα Νευρωνικά Δίκτυα.....	56
3.2.4	Εκπαίδευση και Στρατηγικές Βελτιστοποίησης στα Συνελκτικά Νευρωνικά Δίκτυα (CNNs) 57	
3.2.5	Βελτιστοποιητές: Ενίσχυση της Σύγκλισης στα Συνελκτικά Νευρωνικά Δίκτυα.....	59
3.2.6	Συναρτήσεις Απώλειας: Οδηγώντας τη Διαδικασία Εκπαίδευσης	59
3.2.7	Τρέχουσες Έρευνες και Μελλοντικές Τάσεις στις Αρχιτεκτονικές CNN	63
3.3	Σημαντικές Αρχιτεκτονικές Νευρωνικών Δικτύων.....	64
3.3.1	Το δίκτυο LeNet.....	64
3.3.2	Το δίκτυο AlexNet	65
3.3.3	Το δίκτυο U-NET.....	66
4	Κεφάλαιο 4: Image Registration.....	74
4.1	Εισαγωγή.....	74
4.2	Τεχνικές στην Εγγραφή Εικόνων 3D MRI.....	74
4.2.1	Μέθοδοι Βασισμένοι στην Ένταση.....	74
4.2.2	Μέθοδοι Βασισμένοι στη Δομή	75
4.2.3	Μέθοδοι Βασισμένοι στην Κλίση	75
4.2.4	Μέθοδοι Βασισμένοι στην Εκμάθηση	75
4.2.5	Εφαρμογές στην Ιατρική Απεικόνιση	76
4.3	Similarity Metrics.....	76
4.4	Elastix Registration	78
5	Κεφάλαιο 5: Classification	79
5.1	Τύποι Ταξινόμησης	79
5.2	Τύποι Αλγορίθμων Ταξινόμησης	80
5.3	Support Vector Machines (SVMs)	81
5.3.1	Τι είναι ένα SVM?.....	82
5.3.2	Η προέλευση των SVM.....	82
5.3.3	Τρόπος λειτουργίας των SVM	82
5.3.4	Εφαρμογές των SVM	83
5.4	Feature extraction	83
5.4.1	Feature extraction with Pyradiomics.....	84
5.4.2	Τύποι Χαρακτηριστικών που Εξάγονται.....	85
5.4.3	Μετεπεξεργασία και Ανάλυση.....	86
	ΜΕΡΟΣ 2: ΥΛΙΚΑ ΚΑΙ ΜΕΘΟΔΟΛΟΓΙΑ.....	86
6	Κεφάλαιο 6: Περιγραφή Dataset.....	86
6.1	Προεργασία	86
6.2	Διαθέσιμα Λογισμικά.....	86

7	Κεφάλαιο 7: Ανάλυση Pipeline	87
7.1	Εκπαίδευση U-net για την τμηματοποίηση των εικόνων	87
7.1.1	Προεπεξεργασία.....	87
7.1.2	Επιλογή του U-net.....	89
7.2	Μεθοδολογία Registration.....	89
7.2.1	Αξιολόγηση και επιλογή τεχνικής.....	91
7.3	Μεθοδολογία Απομόνωσης Βλαβών.....	93
7.3.1	Χρήση εικόνων b900.....	93
7.3.2	Συνδυασμός T1W και T2W	94
7.3.3	Χρήση ADC χάρτη.....	94
7.3.4	Δείκτης Fat Fraction.....	94
7.3.5	Τεχνικές Τμηματοποίησης	96
7.4	Διαδικασία Classification	96
7.4.1	Δυναδικός Ταξινομητής.....	96
7.4.2	Ταξινομητής ανά σπόνδυλο	98
7.4.3	Αξιολόγηση.....	98
7.5	Σύνοψη και ιδανική λειτουργία της διαδικασίας.....	98
	ΜΕΡΟΣ 3: ΑΠΟΤΕΛΕΣΜΑΤΑ, ΠΕΡΙΟΡΙΣΜΟΙ, ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΠΡΟΤΑΣΕΙΣ ΓΙΑ ΜΕΛΕΤΗ.....	100
8	Κεφάλαιο 8: Αποτελέσματα.....	100
8.1	Εκπαίδευση U-net.....	100
8.1.1	Main block	101
8.1.2	Inception block.....	105
8.1.3	Dense block.....	109
8.1.4	Καλύτερος Συνδυασμός: Main + Recurrent (t=3) + Residual connections	113
8.2	Image Registration.....	117
8.3	Component Analysis	118
8.4	Απομόνωση βλαβών.....	118
8.4.1	Φωτεινότητες b900.....	118
8.4.2	Συνδυασμός T1W και T2W	119
8.4.3	ADC Mapping.....	121
8.4.4	Fat Fraction	122
8.5	Διαδικασία Classification	122
8.5.1	Feature extraction.....	122
8.5.2	Εκπαίδευση του SVM.....	122
9	Κεφάλαιο 9: Περιορισμοί	124

9.1	Εκπαίδευση των U-net.....	124
9.2	Απομόνωση βλαβών.....	124
9.3	Ταξινόμηση	124
10	Κεφάλαιο 10: Συμπεράσματα και Προτάσεις για μελέτη.....	125
	Βιβλιογραφία	127

Ευρετήριο Εικόνων

ΕΙΚΟΝΑ 1.1-1: Ο ΠΥΡΗΝΑΣ ΤΟΥ ΥΔΡΟΓΟΝΟΥ ΣΥΜΠΕΡΙΦΕΡΕΤΑΙ ΩΣ ΜΑΓΝΗΤΙΚΟ ΔΙΠΟΛΟ, ΜΕ ΜΑΓΝΗΤΙΚΗ ΡΟΠΗ [1]	21
ΕΙΚΟΝΑ 1.1-2: Η ΟΛΙΚΗ ΜΑΓΝΗΤΙΣΗ Μ, ΕΙΝΑΙ Η ΣΥΝΙΣΤΑΜΕΝΗ ΤΩΝ ΠΛΕΟΝΑΖΟΥΣΩΝ ΜΑΓΝΗΤΙΚΩΝ ΡΟΠΩΝ ΤΩΝ ΠΑΡΑΛΛΗΛΩΝ ΠΥΡΗΝΩΝ ΚΑΙ ΕΧΟΥΝ ΔΙΕΥΘΥΝΣΗ ΠΑΡΑΛΛΗΛΗ ΜΕ ΤΟ ΕΞΩΤΕΡΙΚΟ ΠΕΔΙΟ Β [1]	22
ΕΙΚΟΝΑ 1.2-1: Η ΔΙΑΜΗΚΗΣ ΜΑΓΝΗΤΙΣΗ ΜΖ ΑΥΞΑΝΕΤΑΙ ΕΚΘΕΤΙΚΑ ΜΕ ΤΟΝ ΧΡΟΝΟ. Ο ΧΡΟΝΟΣ Τ1 ΕΙΝΑΙ Ο ΑΠΑΙΤΟΥΜΕΝΟΣ ΩΣΤΕ Η ΜΖ ΝΑ ΠΑΡΕΙ ΤΟ 63% ΤΗΣ ΜΕΓΙΣΤΗΣ ΤΙΜΗΣ ΤΗΣ [1]	23
ΕΙΚΟΝΑ 1.2-2: Η ΕΓΚΑΡΣΙΑ ΜΑΓΝΗΤΙΣΗ ΜΕΙΩΝΕΤΑΙ ΕΚΘΕΤΙΚΑ ΜΕ ΤΟΝ ΧΡΟΝΟ [1]	24
ΕΙΚΟΝΑ 2.1-1: ΠΡΟΕΛΕΥΣΗ ΤΩΝ ΔΙΑΦΟΡΩΝ ΚΥΤΤΑΡΙΚΩΝ ΤΥΠΩΝ ΣΤΟ ΑΙΜΑ [4]	30
ΕΙΚΟΝΑ 2.2-1: ΛΙΠΩΔΗΣ ΜΥΕΛΟΣ ΣΕ ΑΠΕΙΚΟΝΙΣΗ T1W [4]	40
ΕΙΚΟΝΑ 2.2-2: ΦΥΣΙΟΛΟΓΙΚΟΣ ΕΡΥΘΡΟΣ ΜΥΕΛΟΣ [4]	41
ΕΙΚΟΝΑ 2.2-3: ΠΑΘΟΛΟΓΙΚΟΣ ΜΥΕΛΟΣ [4]	41
ΕΙΚΟΝΑ 2.2-4: VARIEGATED PATTERN ΠΟΛΛΑΠΛΟΥ ΜΥΕΛΩΜΑΤΟΣ: Α) T1W ΕΙΚΟΝΑ ΠΟΛΛΩΝ FOCAL ΒΛΑΒΩΝ, Β) T2W FSE ΕΙΚΟΝΑ ΟΠΟΥ ΟΙ ΒΛΑΒΕΣ ΕΙΝΑΙ ΥΠΕΡΕΝΤΟΝΕΣ ΚΑΙ ΔΕΝ ΞΕΧΩΡΙΖΟΥΝ ΕΥΚΟΛΑ, Γ) STIR ΕΙΚΟΝΑ ΟΠΟΥ ΟΙ ΒΛΑΒΕΣ ΕΙΝΑΙ ΥΠΟΕΝΤΟΝΕΣ ΚΑΙ ΞΕΧΩΡΙΖΟΥΝ ΑΠΟ ΤΟ ΥΠΟΒΑΘΡΟ [4]	43
ΕΙΚΟΝΑ 2.2-5: ΕΓΚΑΡΣΙΑ ΤΟΜΗ DIXON ΕΙΚΟΝΑΣ ΟΠΟΥ ΔΙΑΚΡΙΝΟΝΤΑΙ ΟΙ ΔΙΑΦΟΡΕΣ ΜΕΤΑΞΥ ΤΩΝ ΙΣΤΩΝ ΑΝΑ ΚΑΝΑΛΙ [6]	44
ΕΙΚΟΝΑ 2.2-6: DWI ΕΙΚΟΝΑ ΜΥΕΛΟΥ ΥΓΕΙΟΥΣ ΝΕΑΡΟΥ ΕΝΗΛΙΚΑ: Α) ΕΙΚΟΝΕΣ ΜΕ B-VALUE = 0, 150, 250, 500, 750, Β) ΧΑΡΤΗΣ ADC [4]	46
ΕΙΚΟΝΑ 3.1-1: Η ΔΙΑΔΙΚΑΣΙΑ ΔΗΜΙΟΥΡΓΙΑΣ CLUSTERS ΑΠΟ ΤΟΝ ΑΛΓΟΡΙΘΜΟ K-MEANS [67]	48
ΕΙΚΟΝΑ 3.1-2: Η ΕΠΙΛΟΓΗ ΚΑΤΩΦΛΙΟΥ ΑΠΟ ΤΟΝ ΑΛΓΟΡΙΘΜΟ OTSU [68]	50
ΕΙΚΟΝΑ 3.1-3: Η ΕΠΙΛΟΓΗ ΠΟΛΛΩΝ ΚΑΤΩΦΛΙΩΝ ΑΠΟ ΤΟΝ ΑΛΓΟΡΙΘΜΟ OTSU MULTIPLE THRESHOLDING [7]	52
ΕΙΚΟΝΑ 3.2-1: ΕΞΕΛΙΞΗ ΚΑΙ ΒΑΣΙΚΑ ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΝΕΥΡΩΝΙΚΩΝ ΔΙΚΤΥΩΝ	53
ΕΙΚΟΝΑ 3.2-2: Η ΔΙΑΔΙΚΑΣΙΑ ΤΗΣ ΑΝΤΙΣΤΡΟΦΗΣ ΔΙΑΔΟΣΗΣ	58
ΕΙΚΟΝΑ 3.2-6: AI GENERATED	61
ΕΙΚΟΝΑ 3.2-7: AI GENERATED (DALL-E)	62
ΕΙΚΟΝΑ 3.2-8: AI GENERATED (DALL-E)	63
ΕΙΚΟΝΑ 3.3-1: ΑΡΧΙΤΕΚΤΟΝΙΚΗ LeNET	64
ΕΙΚΟΝΑ 3.3-2: ΑΡΧΙΤΕΚΤΟΝΙΚΗ ALEXNET	65
ΕΙΚΟΝΑ 3.3-3: ΒΑΣΙΚΗ ΑΡΧΙΤΕΚΤΟΝΙΚΗ U-NET [69]	67
ΕΙΚΟΝΑ 3.3-4: ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΠΥΛΗΣ ΠΡΟΣΟΧΗΣ [39]	69
ΕΙΚΟΝΑ 3.3-5: Α) ΑΡΧΙΚΗ INCEPTION ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΠΟΥ ΧΡΗΣΙΜΟΠΟΙΗΘΗΚΕ ΣΤΟ GOOGLENET, Β) ΒΕΛΤΙΩΜΕΝΗ INCEPTION ΑΡΧΙΤΕΚΤΟΝΙΚΗ [39]	70
ΕΙΚΟΝΑ 3.3-6: ΤΟ ΕΠΑΝΑΛΗΠΤΙΚΟ ΝΕΥΡΩΝΙΚΟ ΔΙΚΤΥΟ, ΕΝΑΣ ΑΠΛΟΣ ΤΥΠΟΣ ΔΙΚΤΥΟΥ ΣΤΟ ΟΠΟΙΟ ΤΟ ΔΕΥΤΕΡΟ ΚΑΙ ΤΟ ΤΡΙΤΟ ΣΤΡΩΜΑ ΕΙΝΑΙ ΕΠΑΝΑΛΗΠΤΙΚΑ ΣΤΡΩΜΑΤΑ. ΚΑΘΕ ΝΕΥΡΩΝΑΣ ΣΕ ΕΝΑ ΕΠΑΝΑΛΗΠΤΙΚΟ ΣΤΡΩΜΑ ΛΑΜΒΑΝΕΙ ΑΝΑΤΡΟΦΟΔΟΤΗΣΗ ΑΠΟ ΤΗ ΔΙΚΗ ΤΟΥ ΕΞΟΔΟ ΚΑΘΩΣ ΚΑΙ ΝΕΑ ΠΛΗΡΟΦΟΡΙΑ [39]	72
ΕΙΚΟΝΑ 3.3-7: Ένα ΠΕΝΤΑΣΤΡΩΜΟ ΠΥΚΝΟ ΜΠΛΟΚ. Η ΜΟΝΑΔΑ ΣΥΓΧΩΝΕΥΣΗΣ ΛΑΜΒΑΝΕΙ ΤΟΝ ΧΑΡΤΗ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ ΑΠΟ ΟΛΑ ΤΑ ΠΡΟΗΓΟΥΜΕΝΑ ΣΤΡΩΜΑΤΑ ΚΑΙ ΤΟ ΠΕΡΝΑ ΣΤΟ ΕΠΟΜΕΝΟ ΣΤΡΩΜΑ. [39]	73
ΕΙΚΟΝΑ 5.4-1: ΑΠΛΗ ΧΡΗΣΗ ΤΗΣ PYRADIOMICS ΜΕ ΡΥΘΜΟΝ, ΓΙΑ ΤΗΝ ΔΙΑΔΙΚΑΣΙΑ ΤΗΣ ΕΞΑΓΩΓΗΣ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ	85
ΕΙΚΟΝΑ 7.1-1: ΔΟΜΗ ΕΚΠΑΙΔΕΥΣΗΣ U-NET ΜΕ ΣΚΟΠΟ ΤΗΝ ΠΑΡΑΓΩΓΗ ΔΥΑΔΙΚΗΣ ΜΑΣΚΑΣ ΤΗΣ ΣΠΟΝΔΥΛΙΚΗΣ ΣΤΗΛΗΣ ΑΠΟ WHOLE-BODY MRI ΕΙΚΟΝΕΣ	88
ΕΙΚΟΝΑ 7.2-1: ΔΙΑΔΙΚΑΣΙΑ REGISTRATION ΜΕ ΤΟ ELASTIX	90
ΕΙΚΟΝΑ 7.2-2: ΧΡΗΣΗ TRANSFORMIX ΓΙΑ ΤΟ REGISTRATION ΤΩΝ ΕΝΑΠΟΜΕΙΝΑΝΤΩΝ ΚΑΝΑΛΙΩΝ	91
ΕΙΚΟΝΑ 7.3-1: ΔΙΑΓΡΑΜΜΑ ΤΕΧΝΙΚΩΝ ΑΠΟΜΟΝΩΣΗΣ ΒΛΑΒΩΝ	95
ΕΙΚΟΝΑ 7.5-1: ΔΙΑΓΡΑΜΜΑ ΟΛΙΚΟΥ PIPELINE ΓΙΑ ΕΚΤΙΜΗΣΗ ΑΣΘΕΝΟΥΣ	99
ΕΙΚΟΝΑ 8.1-1: ΟΙ ΚΑΛΥΤΕΡΟΙ ΣΥΝΔΥΑΣΜΟΙ U-NET, ΕΧΟΝΤΑΣ ΩΣ ΒΑΣΗ ΤΟ ΑΠΛΟ MAIN BLOCK	100
ΕΙΚΟΝΑ 8.1-2: TRAINING LOSS ΤΟΥ MAIN BLOCK ΓΙΑ 200 ΕΠΟΧΕΣ	101
ΕΙΚΟΝΑ 8.1-3: VALIDATION LOSS ΤΟΥ MAIN BLOCK ΓΙΑ 200 ΕΠΟΧΕΣ	102
ΕΙΚΟΝΑ 8.1-4: TRAINING DICE AND IOU SCORE ΓΙΑ ΤΟ MAIN BLOCK ΓΙΑ 200 ΕΠΟΧΕΣ	103
ΕΙΚΟΝΑ 8.1-5: ΑΠΟΤΕΛΕΣΜΑ ΤΟΥ MAIN MODEL ΜΕΤΑ ΤΗΝ ΕΚΠΑΙΔΕΥΣΗ	104
ΕΙΚΟΝΑ 8.1-6: TRAINING LOSS ΤΟΥ INCEPTION BLOCK ΓΙΑ 200 ΕΠΟΧΕΣ	105
ΕΙΚΟΝΑ 8.1-7: VALIDATION LOSS ΤΟΥ INCEPTION BLOCK ΓΙΑ 200 ΕΠΟΧΕΣ	106
ΕΙΚΟΝΑ 8.1-8: DICE AND IOU ΤΟΥ INCEPTION BLOCK ΓΙΑ 200 ΕΠΟΧΕΣ	107
ΕΙΚΟΝΑ 8.1-9: ΠΡΟΒΛΕΨΕΙΣ INCEPTION BLOCK	108

ΕΙΚΟΝΑ 8.1-10: TRAINING LOSS ΤΟΥ DENSE BLOCK ΓΙΑ 200 ΕΠΟΧΕΣ	109
ΕΙΚΟΝΑ 8.1-11: VALIDATION LOSS ΤΟΥ DENSE BLOCK ΓΙΑ 200 ΕΠΟΧΕΣ	110
ΕΙΚΟΝΑ 8.1-12: DICE AND IOU SCORE ΤΟΥ INCEPTION BLOCK ΓΙΑ 200 ΕΠΟΧΕΣ.....	111
ΕΙΚΟΝΑ 8.1-13: ΠΡΟΒΛΕΨΗ ΤΟΥ DENSE BLOCK.....	112
ΕΙΚΟΝΑ 8.1-14: TRAINING LOSS ΓΙΑ ΤΟ ΚΑΛΥΤΕΡΟ U-NET ΓΙΑ 200 ΕΠΟΧΕΣ.....	113
ΕΙΚΟΝΑ 8.1-15: VALIDATION LOSS ΓΙΑ ΤΟ ΚΑΛΥΤΕΡΟ U-NET ΓΙΑ 200 ΕΠΟΧΕΣ	114
ΕΙΚΟΝΑ 8.1-16: DICE AND IOU SCORE ΓΙΑ ΤΟ ΚΑΛΥΤΕΡΟ U-NET ΓΙΑ 200 ΕΠΟΧΕΣ	115
ΕΙΚΟΝΑ 8.1-17: ΠΡΟΒΛΕΨΗ ΤΟΥ ΚΑΛΥΤΕΡΟΥ U-NET.....	116
ΕΙΚΟΝΑ 8.2-1: DWI b50, ΠΑΡΑΜΟΡΦΩΜΕΝΗ ΜΕΤΑ ΤΗΝ ΕΠΕΞΕΡΓΑΣΙΑ ΑΠΟ ΤΟ ELASTIX	117
ΕΙΚΟΝΑ 8.3-1: ΠΑΡΑΔΕΙΓΜΑΤΑ ΑΠΟΜΟΝΩΣΗΣ ΣΠΟΝΔΥΛΩΝ.....	118
ΕΙΚΟΝΑ 8.4-1: ΙΣΤΟΓΡΑΜΜΑ ΟΜΑΛΟΠΟΙΗΜΕΝΗΣ b900 ΕΙΚΟΝΑΣ ΤΗΣ ΣΠΟΝΔΥΛΙΚΗΣ ΣΤΗΛΗΣ.....	118
ΕΙΚΟΝΑ 8.4-2: ΑΠΟΜΟΝΩΣΗ ΒΛΑΒΩΝ ΜΕΣΩ ΦΩΤΕΙΝΟΤΗΤΩΝ b900	119
ΕΙΚΟΝΑ 8.4-3: ΙΣΤΟΓΡΑΜΜΑ T1W ΕΙΚΟΝΑΣ ΑΠΟ ΠΑΛΑΙΟ ΤΥΠΟΥ ΜΑΓΝΗΤΗ.....	119
ΕΙΚΟΝΑ 8.4-4: ΙΣΤΟΓΡΑΜΜΑ T1W ΕΙΚΟΝΑΣ ΑΠΟ ΝΕΟΤΕΡΟΥ ΤΥΠΟΥ ΜΑΓΝΗΤΗ.....	119
ΕΙΚΟΝΑ 8.4-5: ΑΠΟΜΟΝΩΣΗ ΒΛΑΒΩΝ ΜΕ ΤΗ ΣΥΝΔΥΣΤΙΚΗ ΜΕΘΟΔΟ T1W/T2W	120
ΕΙΚΟΝΑ 8.4-6: ADC MASK ΒΛΑΒΩΝ.....	121
ΕΙΚΟΝΑ 8.4-7: ΑΠΟΜΟΝΩΣΗ ΒΛΑΒΩΝ ΜΕΣΩ ΤΟΥ ΔΕΙΚΤΗ FAT FRACTION	122
ΕΙΚΟΝΑ 8.5-1: ΑΠΟΤΕΛΕΣΜΑΤΑ ΕΚΠΑΙΔΕΥΣΗΣ ΤΟΥ SVM CLASSIFIER	123
ΕΙΚΟΝΑ 8.5-2: ΑΠΟΤΕΛΕΣΜΑΤΑ ΕΚΠΑΙΔΕΥΣΗΣ ΜΕ ΤΗ ΧΡΗΣΗ ΟΛΟΚΛΗΡΟΥ ΤΟΥ DATASET	123

Η σελίδα αυτή είναι σκόπιμα λευκή.

Ευρετήριο Εξισώσεων

EQUATION 1.1-1: ΜΑΓΝΗΤΙΚΗ ΡΟΠΗ	20
EQUATION 1.1-2: ΣΥΧΝΟΤΗΤΑ LARMOR	21
EQUATION 1.2-1: ΜΕΤΡΟ ΕΓΚΑΡΣΙΑΣ ΜΑΓΝΗΤΙΣΗΣ	23
EQUATION 2.2-1: ΈΝΤΑΣΗ ΣΗΜΑΤΟΣ IP	43
EQUATION 2.2-2: ΈΝΤΑΣΗ ΣΗΜΑΤΟΣ OP	43
EQUATION 2.2-3: ΈΝΤΑΣΗ ΣΗΜΑΤΟΣ ΝΕΡΟΥ	44
EQUATION 2.2-4: ΈΝΤΑΣΗ ΣΗΜΑΤΟΣ ΛΙΠΟΥΣ	44
EQUATION 2.2-5: ΥΠΟΛΟΓΙΣΜΟΣ ADC MAP.....	45
EQUATION 2.2-6:ΦΥΣΙΟΛΟΓΙΚΕΣ ΤΙΜΕΣ ADC	45
EQUATION 2.2-7:ΠΑΘΟΛΟΓΙΚΕΣ ΤΙΜΕΣ ADC	45
EQUATION 2.2-8: ΕΥΡΟΣ ΤΙΜΩΝ ADC ΟΠΟΥ ΥΠΑΡΧΕΙ ΕΠΙΚΑΛΥΨΗ ΜΕΤΑΞΥ ΥΓΕΙΟΥΣ ΚΑΙ ΠΑΘΟΛΟΓΙΚΟΥ ΜΥΕΛΟΥ	45
EQUATION 3.1-1: ΚΑΤΩΦΛΙΟΠΟΙΗΣΗ ΕΙΚΟΝΑΣ	47
} EQUATION 3.1-2: ΚΑΤΩΦΛΙΟΠΟΙΗΣΗ ΜΕ ΔΥΟ ΚΑΤΩΦΛΙΑ.....	47
EQUATION 3.1-3:ΕΥΡΕΣΗ ΒΕΛΤΙΣΤΟΥ ΚΑΤΩΦΛΙΟΥ ΣΤΗ ΜΕΘΟΔΟ OTSU	51
EQUATION 3.3-1:ΕΝΕΡΓΕΙΑ ΔΙΚΤΥΟΥ U-NET	68
EQUATION 3.3-2: ΣΥΝΑΡΤΗΣΗ SOFTMAX	68
EQUATION 3.3-3: ΕΞΙΣΩΣΗ ΠΡΟΣΘΕΤΙΚΗΣ ΠΥΛΗΣ.....	69
EQUATION 3.3-4: ΕΞΙΣΩΣΗ ΣΙΓΜΟΕΙΔΟΥΣ ΣΥΝΑΡΤΗΣΗΣ.....	70
EQUATION 3.3-5: ΖΕΥΓΟΣ ΣΧΕΣΩΝ ΥΠΟΛΕΙΠΟΜΕΝΩΝ ΜΟΝΑΔΩΝ ΣΤΑ RESNET	71
EQUATION 3.3-6: ΖΕΥΓΟΣ ΣΧΕΣΩΝ ΥΠΟΛΕΙΠΟΜΕΝΩΝ ΜΟΝΑΔΩΝ ΣΤΑ RESNET	71
EQUATION 3.3-7: ΕΞΙΣΩΣΗ ΕΞΟΔΟΥ ΤΟΥ RECURRENT U-NET	72
EQUATION 3.3-8: ΕΞΙΣΩΣΗ ΕΞΟΔΟΥ ΣΤΡΩΜΑΤΟΣ ΣΕ DENSE U-NET	73
EQUATION 7.3-1: ΕΞΙΣΩΣΗ FAT FRACTION	95

Η σελίδα αυτή είναι σκόπιμα λευκή.

Ευρετήριο Πινάκων

TABLE 2.2-1: ΕΠΙΔΡΑΣΗ ΑΥΞΗΜΕΝΩΝ ΕΠΙΠΕΔΩΝ ΚΥΤΤΑΡΩΝ ΣΤΟΝ ΟΣΤΙΚΟ ΜΥΕΛΟ.....	38
TABLE 7.1-1: : ΠΙΝΑΚΑΣ ΣΥΝΔΥΑΣΜΩΝ ΑΡΧΙΤΕΚΤΟΝΙΚΩΝ U-NET	89

Η σελίδα αυτή είναι σκόπιμα λευκή.

ΜΕΡΟΣ 1: ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ

Κεφάλαιο 1: Η μαγνητική τομογραφία

Η Απεικονιστική μέθοδος Μαγνητικού Συντονισμού MRI (Magnetic Resonance Imaging) είναι μια πρόσφατη σχετικά μέθοδος ιατρικής απεικόνισης η οποία βασίζεται στο φαινόμενο του μαγνητικού συντονισμού πυρήνων υδρογόνου για τη δημιουργία εικόνων του σώματος. Αν και το φαινόμενο του συντονισμού πυρήνων ήταν γνωστό από το 1940, οι πρώτες ιατρικές απεικονίσεις έγιναν το 1973 από τον Lauterbur. Από τότε μέχρι και σήμερα η MRI προτιμάται ως μέθοδος απεικόνισης, γιατί παρουσιάζει μεγάλη διακριτική ικανότητα μεταξύ των διαφορετικών μαλακών ιστών, παρέχει τη δυνατότητα για απεικόνιση σε οποιοδήποτε επίπεδο και δεν κάνει χρήση ionτίζουσας ακτινοβολίας[1].

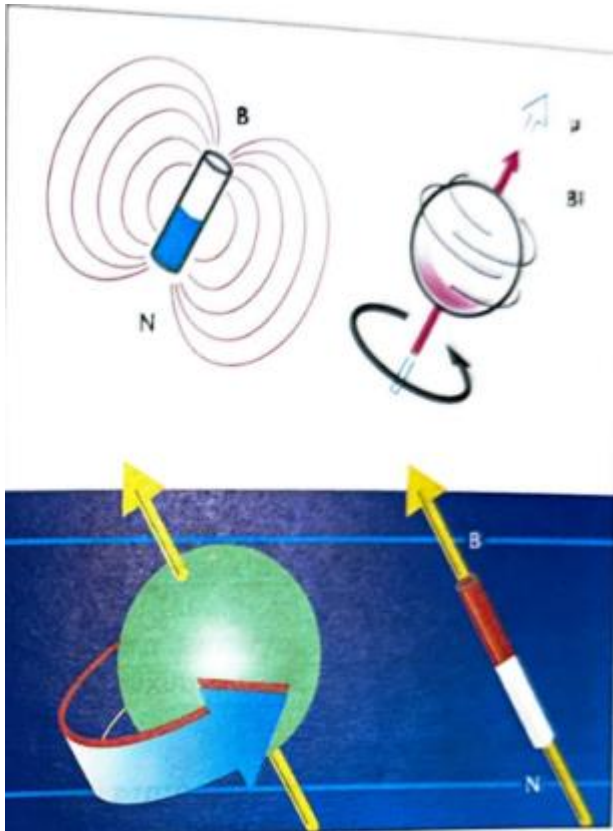
1.1 Αρχές Λειτουργίας

Από πλευράς φυσικής, οι πυρήνες υδρογόνου που βρίσκονται κατά κόρον στο σώμα μας στα μόρια νερού και σε συγκεκριμένα μόρια λιπιδίων, δρουν ως μικρά δίπολα-μαγνήτες με τυχαίο προσανατολισμό. Αυτό συμβαίνει διότι ο πυρήνας του υδρογόνου ο οποίος στοχεύεται κατά τη διαδικασία αποτελείται από ένα μόνο πρωτόνιο, το οποίο έχει ένα spin μη μηδενικό. Συνδυάζοντας το spin με το ηλεκτρικό φορτίο του πρωτονίου, ο πυρήνας αποκτά ένα μικρό μαγνητικό πεδίο και δρα ως μαγνήτης με πεδίο παράλληλο στον άξονα περιστροφής του και φορά καθορισμένη από τον κανόνα του δεξιού χεριού. Αποκτά έτσι μαγνητική ροπή

$$\mu = \gamma I$$

Equation 1.1-1: Μαγνητική Ροπή

όπου γ -> ο πυρηνικός γυρομαγνητικός λόγος και I -> η στροφορμή του πρωτονίου. Έπειτα, τα πρωτόνια εκτίθενται σε ισχυρό μαγνητικό πεδίο B_0 , έτσι ώστε να ξεφύγουν από τον τυχαίο προσανατολισμό τους και να τοποθετηθούν είτε παράλληλα είτε αντιπαράλληλα με τις μαγνητικές γραμμές του πεδίου B_0 [1].



Εικόνα 1.1-1: Ο πυρήνας του υδρογόνου συμπεριφέρεται ως μαγνητικό δίπολο, με μαγνητική ροπή [1]

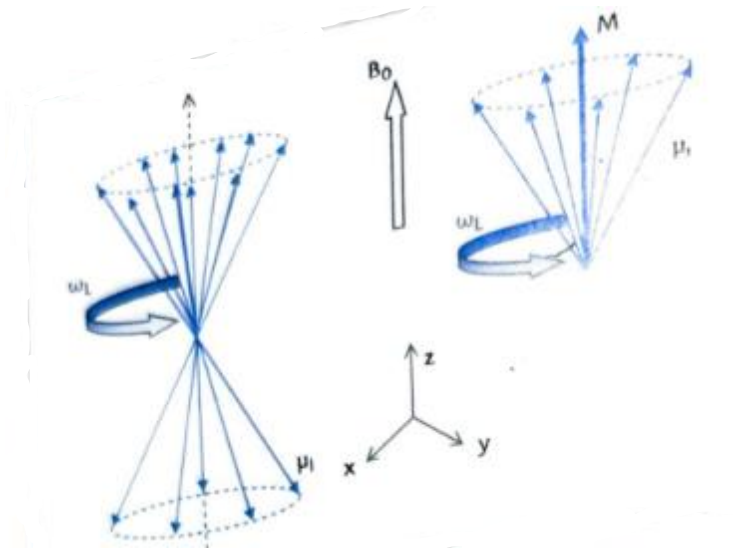
Ο αριθμός των παράλληλων πρωτονίων είναι πάντα μεγαλύτερος από αυτόν των αντιπαράλληλων, καθώς απαιτείται λιγότερη ενέργεια για τον παράλληλο προσανατολισμό. Τα παράλληλα με τα αντιπαράλληλα πρωτόνια εξουδετερώνονται, με αποτέλεσμα η συνισταμένη των μαγνητικών διπολικών ροών των περισσευόμενων παράλληλων πρωτονίων, που ονομάζεται Ολική Μαγνήτιση (M) να έχει τελικά φορά ίδια με αυτή του εξωτερικού ισχυρού πεδίου. Η ένταση αυτού του πεδίου B_0 , κυμαίνεται από 1-3 Tesla, δηλαδή από 2* έως φορές ισχυρότερο από το αντίστοιχο μαγνητικό πεδίο της Γης. Λόγω της επίδρασης του B_0 , οι ευθυγραμμισμένοι πλέον πυρήνες εκτελούν πέρα από το spin τους και μία κωνική περιστροφή γύρω από το διάνυσμα M, παρόμοια με την κίνηση μιας σβούρας. Η κωνική αυτή περιστροφή έχει μια συγκεκριμένη γωνιακή συχνότητα ω_L η οποία ονομάζεται συχνότητα Larmor και δίνεται από τον τύπο

$$\omega_L = \gamma B_0$$

Equation 1.1-2: Συχνότητα Larmor

Για να πραγματοποιηθεί επομένως ο πυρηνικός μαγνητικός συντονισμός και να μεταβούν οι πυρήνες σε υψηλότερη ενεργειακή στάθμη, πρέπει να αποδοθεί σε αυτούς ακτινοβολία στη συγκεκριμένη συχνότητα Larmor. Για τα πεδία B_0 που χρησιμοποιούνται στις εφαρμογές MRI (0,5-3 Tesla), η συχνότητα αυτή αντιστοιχεί στο φάσμα των ραδιοκυμάτων (RadioFrequencies-RF)[1]. Πρακτικά, εφαρμόζεται ένα χρονικά μεταβαλλόμενο μαγνητικό πεδίο B_1 κάθετο στο B_0 , το οποίο περιστρέφεται με την συχνότητα Larmor και συντονίζει τους πυρήνες, τους στρέφει δηλαδή από την αρχική τους θέση προς μία άλλη, συνήθως κατά 90° ή 180° , ανάλογα με την ένταση και τη χρονική διάρκεια του ραδιοπαλμού. Διανυσματικά, αυτό σημαίνει πως το

διάνυσμα ολικής μαγνήτισης M , το οποίο ήταν παράλληλο στο πεδίο B_0 πλέον έχει στραφεί κατά 90° ή 180° .



Εικόνα 1.1-2: Η ολική μαγνήτιση M , είναι η συνισταμένη των πλεοναζουσών μαγνητικών ροπών των παράλληλων πυρήνων και έχουν διεύθυνση παράλληλη με το εξωτερικό πεδίο B [1]

1.2 Χρόνοι Μαγνητικής Αποκατάστασης

Μετά το πέρας του ραδιοπαλμού, οι πυρήνες απελευθερώνουν την περίσσεια ενέργεια που έχουν λάβει και η οποία με κατάλληλη επεξεργασία χρησιμοποιείται για την κατασκευή των εικόνων και επανέρχονται στην αρχική κατάσταση ισορροπίας, παράλληλοι δηλαδή στο B_0 . Οι χρόνοι που χρειάζονται για να συμβεί αυτό χωρίζονται σε δύο κατηγορίες (T_1 και T_2) και είναι απαραίτητοι ώστε να οριστούν δύο άλλα μεγέθη, η Διαμήκης Μαγνητική Αποκατάσταση M_z και η Εγκάρσια Μαγνητική Αποκατάσταση M_{xy} . Ως M_z ορίζεται το μέτρο της μαγνήτισης που υπάρχει κάθε χρονική στιγμή στον άξονα Z , παράλληλα δηλαδή με το εξωτερικό πεδίο B_0 , ενώ ως M_{xy} ορίζεται το μέτρο της μαγνήτισης που υπάρχει κάθετα στο B_0 , δηλαδή παράλληλα στο επίπεδο XY .

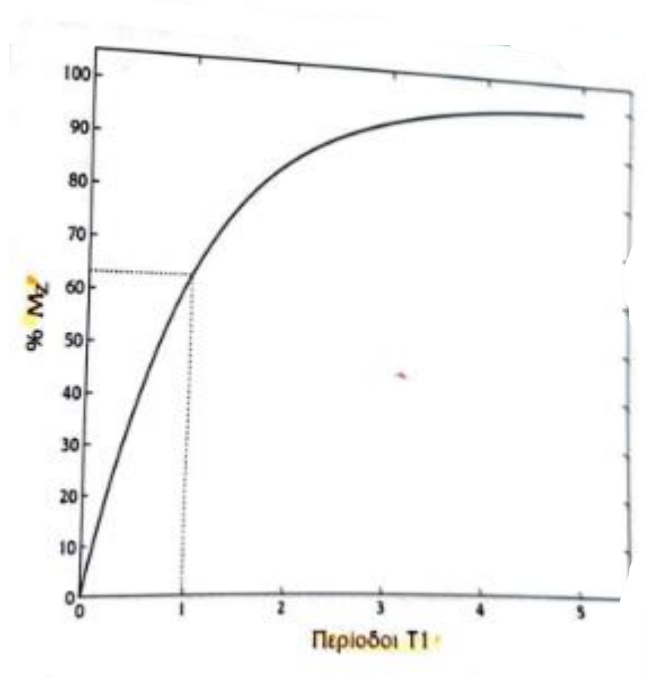
Χρόνος T_1 :

Ως χρόνος μαγνητικής αποκατάστασης T_1 ονομάζεται ο χρόνος που απαιτείται μετά το πέρας του ραδιοπαλμού, ώστε το σύστημα να επανέρθει στο 63% της αρχικής τιμής μαγνήτισης M και επομένως η τιμή του M_z κάθε χρονική στιγμή δίνεται από τον τύπο

$$M_z(t) = M \left(1 - e^{-\frac{t}{T_1}} \right)$$

Equation 1.2-1: Εξίσωση τιμής διαμήκους μαγνήτισης

Η αποκατάσταση αυτή οφείλεται στην αλληλεπίδραση των πυρήνων μεταξύ τους και για αυτό ονομάζεται και spin-lattice relaxation time, ενώ εξαρτάται άμεσα από την ένταση του εξωτερικού μαγνητικού πεδίου B_0 και την θερμοκρασία[1].



Εικόνα 1.2-1: Η διαμήκης μαγνήτιση M_z αυξάνεται εκθετικά με τον χρόνο. Ο χρόνος T_1 είναι ο απαιτούμενος ώστε η M_z να πάρει το 63% της μέγιστης τιμής της [1]

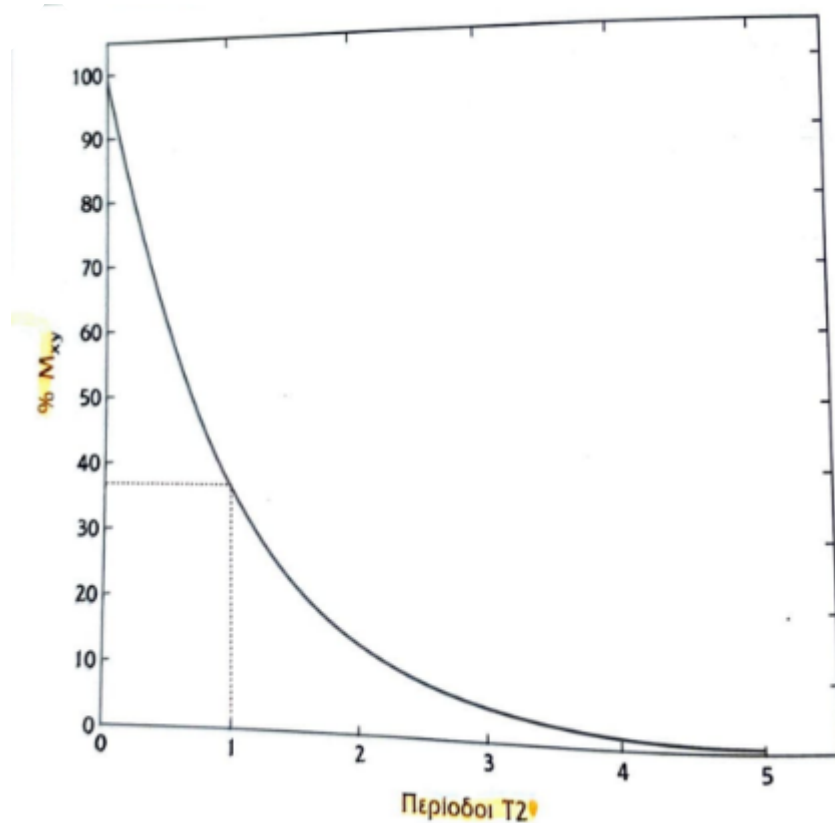
Χρόνος T_2 :

Ως χρόνος T_2 ονομάζεται ο χρόνος που απαιτείται ώστε το μέτρο της εγκάρσιας μαγνήτισης M_{XY} να μειωθεί στο 37% της αρχικής της τιμής μετά το πέρας του ραδιοπαλμού. Επομένως, για κάθε χρονική στιγμή, το μέτρο της M_{XY} δίνεται από τον τύπο

$$M_{XY}(t) = M_{XY}(0)e^{-t/T_2}$$

Equation 1.2-2: Μέτρο εγκάρσιας μαγνήτισης

Κατά την αποκατάσταση T_2 , σε αντίθεση με αυτήν της T_1 , οι πυρήνες δεν χάνουν ενέργεια, αλλά χάνεται η συμφωνία φάσεων που είχαν αποκτήσει κατά τον συντονισμό, μεταφέροντας το σύστημα από μία κατάσταση χαμηλής εντροπίας σε μία κατάσταση υψηλότερης. Καθώς οι φάσεις των σπιν αποσυντονίζονται γιατί τα σπιν αλληλεπιδρούν το ένα με το άλλο, η αποκατάσταση ονομάζεται επίσης σπιν-σπιν αποκατάσταση και εξαρτάται μόνο από τη θερμοκρασία. Στην πραγματικότητα, επειδή το πεδίο B_0 δεν είναι ομοιογενές, ο πραγματικός χρόνος εγκάρσιας απομαγνήτισης είναι $T_2^* < T_2$, γεγονός που λαμβάνεται υπ' όψιν στις τεχνικές spin-echo[2].



Εικόνα 1.2-2: Η εγκάρσια μαγνήτιση μειώνεται εκθετικά με τον χρόνο [1]

Τεχνική spin-echo:

Έχει προσδιοριστεί πως μετά το πέρας του παλμού 90° , οι πυρήνες των ατόμων υδρογόνου έχουν στραφεί κατά 90° συγκριτικά με το διάνυσμα του εξωτερικού μαγνητικού πεδίου και το διάνυσμα M_{xy} έχει λάβει τη μέγιστη τιμή του. Τότε αρχίζει η απώλεια φάσης των σπιν κάθε πυρήνα, αφού ο κάθε ένας αρχίζει να περιστρέφεται με ελαφρώς διαφορετική συχνότητα, γεγονός που οφείλεται στις ανομοιομορφίες του εξωτερικού πεδίου και οδηγεί στον μηδενισμό του M_{xy} . Η διαδικασία αυτή συμβαίνει σε χρόνο T_2^* . Προκειμένου να εξαλειφθεί η επίδραση του T_2^* και εν τέλει η διαδικασία να εξαρτάται από τον T_2 , εφαρμόζεται ένας παλμός 180° κατά μήκος του x -άξονα, λίγο χρόνο μετά το πέρας του παλμού 90° . Ως αποτέλεσμα, τα σπιν των πυρήνων που είχαν ξεκινήσει να αποκλίνουν αρχίζουν να συγκλίνουν αντιδιαμετρικά και αθροιστικά δημιουργούν εκ νέου ένα μέγιστο του διανύσματος M_{xy} , το $M_{xy\text{-echo}}$, μικρότερο από το αρχικό M_{xy} , λόγω απωλειών [1]. Με τον τρόπο αυτό εξασφαλίζεται πως η απομαγνήτιση M_{xy} εξαρτάται αποκλειστικά από τον χρόνο T_2 .

Αντιθέσεις στις εικόνες MRI:

Οι αντιθέσεις στις εικόνες που παράγονται οφείλονται κυρίως στους χρόνους αποκατάστασης T_1 - και T_2 - και στην πυκνότητα πυρήνων υδρογόνου. Οι παράγοντες αυτοί διαφέρουν από ιστό σε ιστό, αλλά και μεταξύ διαφορετικών καταστάσεων του ίδιου ιστού. Έτσι, οι εικόνες, αν και θα εξαρτώνται και από τους τρεις παράγοντες, πάντα λαμβάνονται με τέτοιο τρόπο ώστε να έχουν μία μόνο κύρια πηγή αντίθεσης, με τις άλλες δύο παραμέτρους να έχουν παρασκηνιακό ρόλο.

1.3 Παραγωγή Ακολουθιών MRI

Η παραγωγή των εικόνων απαιτεί την επαναληπτική χρήση παλμών. Επομένως, ορίζεται ένας χρόνος επανάληψης TR, ο οποίος είναι ο χρόνος που μεσολαβεί μεταξύ δύο παλμών 90° . Ανάλογα με το ποια μορφή αντίθεσης είναι επιθυμητή, η τιμή αυτού του χρόνου μπορεί να αλλάζει. Επίσης ορίζεται ο χρόνος αντήχησης TE, ο οποίος είναι ο χρόνος που μεσολαβεί μεταξύ του παλμού 90° και της ανάγνωσης του παραγόμενου σήματος[1].

Σήμα Ελεύθερης Επαγωγικής Απόσβεσης:

Το παραγόμενο σήμα, το οποίο αναφέρθηκε προηγουμένως, είναι αυτό το οποίο θα επεξεργαστεί και θα δώσει την εικόνα. Ο μόνος τρόπος να παραχθεί είναι η δημιουργία της εγκάρσιας μαγνήτισης και έπειτα απομαγνήτισης M_{xy} , καθώς αυτή είναι χρονοεξαρτώμενη και μειώνεται εκθετικά με το χρόνο. Συγκεκριμένα, τοποθετείται ένα πηνίο με τον βασικό του άξονα παράλληλα στο εγκάρσιο επίπεδο-XY. Κατά την εγκάρσια απομαγνήτιση, το πεδίο B_1 που «βλέπει» το πηνίο μεταβάλλεται και άρα σύμφωνα με το νόμο Faraday παράγεται στα άκρα του επαγωγική τάση η οποία μειώνεται εκθετικά σύμφωνα με το χρόνο T_2^* , αφού το αρχικό εξωτερικό πεδίο B_0 δεν είναι στην πραγματικότητα ομοιογενές. Επειδή ο χρόνος αυτός δεν είναι αντιπροσωπευτικός, χρησιμοποιούνται οι προαναφερθείσες τεχνικές spin-echo με τον παλμό 180° , ώστε να υπάρξει εξάρτηση από τον χρόνο T_2 [1].

1.3.1 Παραγωγή της εικόνας

Το σήμα πληροφορίας στο χώρο:

Το επαγωγικό σήμα που λαμβάνεται πρέπει να κωδικοποιηθεί με τέτοιο τρόπο ώστε η πληροφορία του να τοποθετηθεί στο χώρο σε συγκεκριμένη θέση. Για να συμβεί αυτό, το εξωτερικό πεδίο B_0 πρέπει να μεταβάλλεται, όχι χρονικά, αλλά χωρικά. Αυτό επιτυγχάνεται με την εφαρμογή βαθμιδωτών πεδίων ταυτόχρονα με το B_0 , η ένταση των οποίων μεταβάλλεται γραμμικά κατά μήκος του άξονα στον οποίο εφαρμόζονται. Στην περίπτωση αυτή, οι πυρήνες υδρογόνου αποκτούν διαφορετικές συχνότητες Larmor ανάλογα με το πού βρίσκονται στο χώρο και έτσι ομάδες πυρήνων μπορούν να συντονιστούν ανεξάρτητα και επιλεκτικά, με τη χρήση κατάλληλης συχνότητας ραδιοπαλμού[1]. Ως εκ τούτου, δημιουργούνται εικόνες που παρουσιάζουν τα μαγνητικά χαρακτηριστικά των ιστών που ανταποκρίνονται στην εκάστοτε συχνότητα Larmor που επιλέγεται, με τη φωτεινότητα κάθε pixel της εικόνας να αντιστοιχεί στο μέσο όρο των σημάτων των πυρήνων που βρίσκονται στο αντίστοιχο voxel.

Διαχωρισμός Τομών:

Με την εφαρμογή βαθμιδωτού πεδίου κατά μήκος του άξονα-Z και κατάλληλα επιλεγμένου παλμού 90° , επιλέγεται το σημείο του σώματος στο οποίο θα παρθεί η τομή. Η κεντρική συχνότητα του παλμού προσδιορίζει τη θέση της τομής ενώ το εύρος ζώνης του το πάχος της. Παράλληλα, με την εφαρμογή αντίστοιχου βαθμιδωτού πεδίου στον άξονα-Y η τομή χωρίζεται σε οριζόντιες γραμμές. Αυτό επιτυγχάνεται διότι το πεδίο στον άξονα-Y προκαλεί την διαφοροποίηση στη συχνότητα spin αλλά και στη φάση των πυρήνων. Μετά την άρση του βαθμιδωτού πεδίου οι πυρήνες αποκτούν πάλι ίδιο spin, αλλά παραμένουν εκτός φάσης, με αποτέλεσμα τον διαχωρισμό σε γραμμές. Τέλος, με την εφαρμογή βαθμιδωτού πεδίου στον

άξονα- X , τα πρωτόνια στρέφονται με διαφορετικές συχνότητες Larmor, διατηρώντας όμως τη φάση που είχαν αποκτήσει από την προηγούμενη διαμόρφωση, χωρίζοντας έτσι την εικόνα και σε κάθετες γραμμές. Μια ειδική περίπτωση είναι αυτή της τρισδιάστατης απεικόνισης, στην οποία συντονίζεται όλος ο όγκος που είναι να απεικονιστεί. Η κωδικοποίηση σε αυτήν την περίπτωση γίνεται με την εφαρμογή βαθμιδωτού πεδίου στον άξονα- Y για το χωρισμό της εικόνας σε γραμμές και την εφαρμογή βαθμιδωτού πεδίου στον άξονα- Z αρχικά για την επιλογή της θέσης της τομής και έπειτα αλλαγή της έντασης του τόσες φορές όσες και οι τομές που απαιτούνται[1]. Η τεχνική αυτή προτιμάται καθώς επιτρέπει τη λήψη πολύ λεπτών τομών, την ανακατασκευή τους σε οποιοδήποτε επίπεδο και το υψηλό SNR που διαθέτει.

Χαρακτηριστικά της εικόνας:

Όπως έχει προαναφερθεί, η εικόνα του MRI εξαρτάται από τρεις παράγοντες, τους χρόνους T_1 και T_2 και την πυκνότητα πρωτονίων και κάθε φορά επιλέγεται ποιος από αυτούς θα έχει τη μεγαλύτερη βαρύτητα. Πέρα από αυτά όμως, η εικόνα χαρακτηρίζεται επίσης από την αντίθεση, τη χωρική διακριτική ικανότητα και το θόρυβό της. Ως αντίθεση, ορίζεται η διαφορά στην ένταση του σήματος, δηλαδή της φωτεινότητας, που διέπει διαφορετικούς ιστούς και εξαρτάται από τους χρόνους αποδιέγερσης και την πυκνότητα πρωτονίων. Η χωρική διακριτική ικανότητα είναι η ελάχιστη ικανή απόσταση ώστε δύο σημεία να μπορούν να αναγνωριστούν ως διαφορετικά. Εξαρτάται από το μέγεθος των voxel, με αντιστρόφως ανάλογη σχέση. Ο θόρυβος προκύπτει κυρίως από τα τυχαία σήματα που εκπέμπει το σώμα καθώς και από τον θερμικό θόρυβο των ηλεκτρονικών στοιχείων και δημιουργεί πρόβλημα στην ανάγνωση της εικόνας όσο μειώνεται το SNR. Εξαρτάται άμεσα από το μέγεθος του voxel, όπως και η διακριτική ικανότητα, αλλά με αντίθετη εξάρτηση. Εντούτοις, όσο πιο μικρό είναι το voxel τόσο χειρότερος είναι ο SNR, γιατί το σήμα πληροφορίας λαμβάνεται από λιγότερους πυρήνες και είναι ασθενέστερο.

1.3.2 Επιλογή των χρόνων TR, TE:

Ο χρόνος TR, όπως έχει αναφερθεί, είναι ο χρόνος που μεσολαβεί μεταξύ δύο διαδοχικών ραδιοπαλμών 90° και καθορίζει μαζί με το χρόνο T_1 την τιμή της διαμήκους μαγνήτισης M_z . Έτσι, για τη στάθμιση μιας εικόνας ως προς T_1 επιλέγεται μικρός TR για να αναδειχτούν οι διαφορές κάθε ιστού, αφού με μεγάλο TR όλοι οι ιστοί θα έχουν επανέρθει στη μέγιστη τιμή της M_z και άρα δεν θα παρατηρείται διαφορά. Αντίθετα, ο χρόνος TE, που μεσολαβεί μεταξύ του ραδιοπαλμού 90° και της ανάγνωσης του σήματος, επηρεάζει μαζί με τον χρόνο T_2 την τιμή της εγκάρσιας μαγνήτισης M_{xy} και για να υπάρξει στάθμιση ως προς T_2 πρέπει ο χρόνος TE να είναι αρκετά μεγάλος[1].

1.4 Τα Μέρη του Τομογράφου

Ο μαγνήτης:

Ο κεντρικός μαγνήτης αποτελεί το βασικό μέρος του τομογράφου, αφού δημιουργεί το εξωτερικό ισχυρό μαγνητικό πεδίο B_0 . Παλαιότερα γινόταν χρήση στατικού μαγνήτη, το

πεδίο όμως δεν ξεπερνούσε τα 0,5 Tesla. Πλέον προτιμώνται υπεραγώγιμοι μαγνήτες, οι οποίοι είναι ένα σωληνοειδές πηνίο μέσα από το οποίο διέρχεται ρεύμα σε θερμοκρασίες 4°K ώστε το υλικό του πηνίου, συνήθως Nb-Ti, να καταστεί υπεραγώγιμο, κάτι που επιτυγχάνεται με τη χρήση υγρού ηλίου και αζώτου. Τα πεδία αυτών των μαγνητών κυμαίνονται από 0,5-3 Tesla[1].

Τα βαθμιδωτά πεδία:

Τα βαθμιδωτά πεδία, που χρησιμοποιούνται για την χωρική μεταβολή του B_0 , εφαρμόζονται σε κάθε άξονα (X-Y-Z) και έχουν τη διεύθυνση του B_0 . Η έντασή τους μεταβάλλεται γραμμικά κατά μήκος του άξονα εφαρμογής τους και παράγονται από ειδικά πηνία, τα βαθμιδωτά πηνία. Επειδή κατά τη διάρκεια της εξέτασης τα βαθμιδωτά πεδία ενεργοποιούνται και απενεργοποιούνται με μεγάλη συχνότητα, δημιουργούν μικρά ρεύματα επαγωγής, τα οποία με τη σειρά τους παράγουν μικρά μαγνητικά πεδία που χαλούν την ομοιογένεια του ισχυρού εξωτερικού πεδίου B_0 [1].

Τα πηνία ραδιοσυχνότητας:

Τα πηνία αυτά χρησιμοποιούνται για την παραγωγή των ραδιοπαλμών 90° και 180°, καθώς και τη λήψη του επαγωγικού ηλεκτρικού σήματος που χρησιμοποιείται για την κατασκευή της εικόνας. Προκειμένου να ανιχνεύεται όσο το δυνατόν λιγότερος θόρυβος και άρα να προκύψει αύξηση του σηματοθορυβικού λόγου SNR, τα πηνία αυτά είναι τοπικά ώστε να καλύπτουν μόνο την επιφάνεια της τομής, με αποτέλεσμα ο θόρυβος από το υπόλοιπο σώμα να μην ανιχνεύεται[1].

Το σύστημα απεικόνισης:

Το σύστημα απεικόνισης του μαγνητικού τομογράφου δέχεται το σήμα, το ανακατασκευάζει και αναπαριστά την εικόνα με τα χαρακτηριστικά των ιστών. Πιο συγκεκριμένα, συνδέει τις μαγνητικές ιδιότητες του κάθε voxel με τη φωτεινότητα που πρέπει να έχει το αντίστοιχο pixel. Για να συμβεί αυτό, ορίζεται το μέγεθος του voxel, το πάχος της τομής και το μέγεθος της εικόνας, το οποίο είναι ανάλογο του πλήθους των voxel και συνήθως έχει μέγεθος 128x128, 256x256 ή 512x512. Κάθε ένα από αυτά τα pixel λαμβάνει την τιμή του μαγνητικού χαρακτηριστικού στο οποίο δίνεται έμφαση, δηλαδή των χρόνων T_1 , T_2 ή της πυκνότητας πρωτονίων του αντίστοιχου voxel και την κωδικοποιεί χρωματικά μέσα από μια κλίμακα με 256 αποχρώσεις του γκρι[1].

1.5 Ασφάλεια της Μαγνητικής Τομογραφίας

Η εξέταση MRI είναι μία από τις ασφαλέστερες μεθόδους απεικόνισης εφόσον τηρούνται όλες οι προδιαγραφές ασφαλείας. Οι πιθανές πηγές ανησυχίας είναι πέντε και βρίσκονται στην ύπαρξη:

- Του ισχυρού εξωτερικού μαγνητικού πεδίου
- Των βαθμιδωτών πεδίων
- Των πηνίων ραδιοσυχνότητας
- Ύπαρξη εμφυτευμάτων ή συσκευών στο σώμα του ασθενούς ή στο χώρο εξέτασης
- Των κρυογόνων

Το ισχυρό πεδίο:
Σύμφωνα με τις μέχρι στιγμής έρευνες, σύντομη έκθεση σε ισχυρό μαγνητικό πεδίο δεν έχει ανησυχητικές επιδράσεις στον ανθρώπινο οργανισμό. Συμπτώματα μπορεί να παρουσιαστούν κατά την έκθεση σε πεδίο ισχυρότερο των 4T και ενδέχεται να περιλαμβάνουν ναυτία, ίλιγγο, πονοκέφαλο, μούδιασμα οπτικές διαταραχές ή πόνο, σταματούν όμως μετά την έξοδο από το πεδίο. Ανησυχία δεν προκαλούν επίσης οποιεσδήποτε επιδράσεις μπορεί να έχει ο μαγνήτης σε υγρά του σώματος που είναι αγωγίμα, όπως το αίμα, αφού η επίδραση αυτή είναι μικρότερη από 0,2% της αρτηριακής πίεσης σε πεδία ισχυρότερα των 10T. Παρόλα αυτά, η Διεθνής Ηλεκτροτεχνική Επιτροπή (IEC) έχει επιβάλλει όρια σύμφωνα με δύο πρωτόκολλα. Το πρώτο, που ονομάζεται κανονική βαθμίδα λειτουργίας και στο οποίο εφαρμόζονται πεδία μέχρι 1,5T και το δεύτερο που ονομάζεται πρώτου επιπέδου ελεγχόμενη βαθμίδα και εφαρμόζεται σε ασθενείς υψηλού κινδύνου και στην οποία χρησιμοποιούνται πεδία μέχρι 3T. Όσον αφορά τις εγκύους, δεν υπάρχουν στοιχεία που να δείχνουν πως υπάρχει η οποιαδήποτε επιβάρυνση, αλλά παρόλα αυτά, η εξέταση MRI συνίσταται μόνο όταν δεν υπάρχουν άλλες ικανοποιητικότερες εξετάσεις μη-ιοντίζουσας ακτινοβολίες και κατά προτίμηση μετά το πρώτο τρίμηνο της εγκυμοσύνης[1].

Τα βαθμιδωτά πεδία:
Τα βαθμιδωτά πεδία ενεργοποιούνται και απενεργοποιούνται πολλές φορές κατά τη διάρκεια της εξέτασης και επειδή δημιουργούν και ηλεκτρικά πεδία, όταν η συχνότητα ενεργοποίησής τους είναι μεγάλη, ενδέχεται να διεγείρουν το νευρικό ή το μυϊκό σύστημα, προκαλώντας παροδικό πόνο όταν η συχνότητα υπερβαίνει τα 60T/s. Για πολύ υψηλές τιμές ενδέχεται να διεγερθεί η καρδιά και να προκληθούν αρρυθμίες. Επειδή η υψηλή συχνότητα είναι επιθυμητή, καθώς μειώνει το χρόνο εξέτασης, έχουν θεσπιστεί όρια ασφαλείας που προλαμβάνουν την προαναφερθείσα υψηλή έκθεση[1].

Επιπλέον, τα πηνία των βαθμιδωτών πεδίων προκαλούν αρκετό θόρυβο λόγω της σύγκρουσής τους με τα πλαίσια στήριξής τους και για αυτό γίνεται χρήση ωτοασπίδων ή ακουστικών που μειώνουν την έκθεση στο θόρυβο έως και 30dB, ενώ η μέγιστη επιτρεπτή στιγμιαία τιμή θορύβου είναι τα 140dB[1].

Τα πηνία ραδιοσυχνοτήτων:
Τα πηνία αυτά προσδίδουν ενέργεια στο σύστημα, δηλαδή τους ιστούς, η οποία πριν επανεκπεμφθεί απορροφάται ως θερμότητα. Τα ποσά θερμότητας που επιτρέπεται να απορροφηθούν ανά μονάδα μάζας ιστών, ορίζονται από τον ειδικό ρυθμό απορρόφησης SAR σε W/kg και τοποθετούνται στα 2 W/kg για τη συνήθη παρακολούθηση και στα 4 W/kg στους ασθενείς υψηλού κινδύνου. Επίσης, οι ραδιοσυχνότητες επάγουν ρεύματα και τάσεις σε συσκευές όπως βηματοδότες, αλλά και τατουάζ, καλλυντικά ή κοσμήματα που μπορεί να αποτελούνται από σιδηρομαγνητικά υλικά και τα οποία μπορεί να θερμανθούν και να προκαλέσουν οίδημα[1].

Εμφυτεύματα/συσκευές:

Οι περισσότερες συσκευές και τα εμφυτεύματα ιατρικής φύσεως δεν αλληλεπιδρούν με τον μαγνητικό τομογράφο, σε ειδικές περιπτώσεις όμως όπως με βηματοδότες ή αντλίες ινσουλίνης, όπου υπάρχει περίπτωση τα πεδία να διακόψουν την ομαλή λειτουργία τους, η εξέταση MRI απαγορεύεται. Αντίστοιχη προσοχή χρειάζεται και με οποιοδήποτε άλλο

μεταλλικό αντικείμενο στο σώμα του ασθενούς ή στο χώρο εξέτασης που ενδέχεται να μαγνητιστεί από το ισχυρό πεδίο. Τέτοιου τύπου αντικείμενα πρέπει να απαγορεύονται[1].

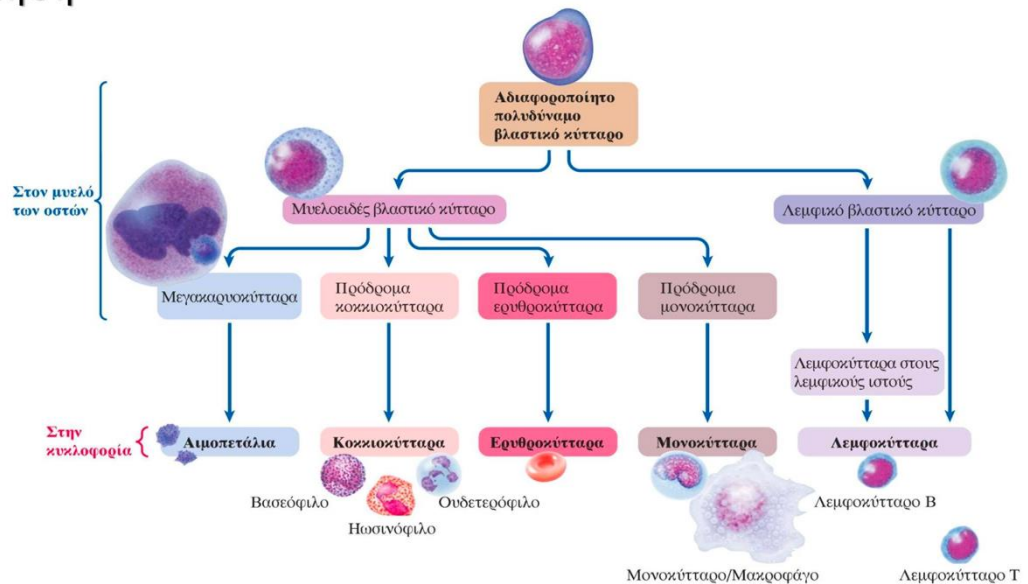
Τα

κρυογόνα:

Το υγρό ήλιο που χρησιμοποιείται για την ψύξη του πηνίου βρίσκεται σε θερμοκρασία 4°K. Αν το πηνίο χάσει την υπεραγωγιμότητά του, τότε μέσω μιας αντίδρασης και σε λιγότερο από 20sec, η ηλεκτρική ενέργεια του πηνίου μετατρέπεται σε θερμική και εξαερώνει το ήλιο, το οποίο είναι αρκετά ψυχρό για να υγροποιήσει το οξυγόνο και να αποτελέσει έτσι σοβαρό κίνδυνο πυρκαγιάς. Εκτός αυτού, το εξαερωμένο ήλιο διαστέλλεται εξαιρετικά γρήγορα αντικαθιστώντας το οξυγόνο στο χώρο και προκαλώντας κίνδυνο ασφυξίας. Για το λόγο αυτό στο χώρο της εξέτασης υπάρχουν ειδικά συστήματα εξαέρωσης[1].

2 Κεφάλαιο 2: Παθοφυσιολογία του Πολλαπλού Μυελώματος και Απεικονιστικές τεχνικές

2.1 Φυσιολογία του αίματος και του ανοσιακού συστήματος Αιμοποίηση



Εικόνα 2.1-1: Προέλευση των διαφόρων κυτταρικών τύπων στο αίμα[4]

Η **αιμοποίηση** Ξεκινά από το αδιαφοροποίητο πολυδύναμο βλαστικό κύτταρο

Στην μια οικογένεια δίνει γέννηση στο **λεμφικό βλαστικό κύτταρο** απόπου προκύπτουν τα B και T λεμφοκύτταρα.

Στην άλλη οικογένεια δίνει γέννηση στο **μυελοειδές βλαστικό κύτταρο** απόπου προκύπτουν τα **Μεγακαρυοκύτταρα** θραύσματα των οποίων δίνουν τα αιμοπετάλια που συμμετέχουν στην πήξη του αίματος.

Τα **πρόδρομα ερυθροκύτταρα** που δίνουν τα ερυθροκύτταρα

Τα **πρόδρομα μονοκύτταρα** που δίνουν τα μονοκύτταρα

Τα **πρόδρομα κοκκιοκύτταρα** που δίνουν τα κοκκιοκύτταρα (Βασεόφιλα, Ηωσινόφιλα, Ουδετερόφιλα)[4].

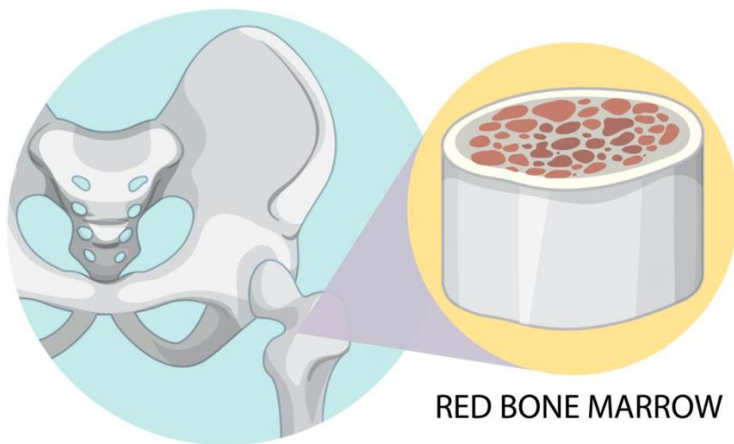
Όταν ένα πολυδύναμο αρχέγονο αιμοποιητικό κύτταρο διαιρείται, τα δύο θυγατρικά του κύτταρα είτε παραμένουν ως έχουν (αυτοανανέωση), είτε δεσμεύονται να ακολουθήσουν ένα συγκεκριμένο αναπτυξιακό μονοπάτι.

Μυελός των οστών.

Είναι το κύριο αιμοποιητικό όργανο στο οποίο παράγονται ερυθρά, αιμοπετάλια και λευκά αιμοσφαίρια. Περιέχει στις μυελοκυψέλες της σπογγώδους ουσίας των οστών και η σύστασή του εξαρτάται από την ηλικία.

Διακρίνεται σε ερυθρό μυελό και σε ωχρο ή λιπώδη μυελό. Στα νεογνά όλα τα οστά περιέχουν ερυθρό μυελό και καθώς ο οργανισμός ενηλικιώνεται ο ερυθρός μυελός υποστρέφεται σε ωχρο μυελό κατά μία συγκεκριμένη οδό όπου εν τέλει ο ερυθρός μυελός καταλήγει να υφίσταται κυρίως στον μυελό των σπονδύλων, στην πύελο καθώς και στα εγγύς άνω και κάτω άκρα[4].

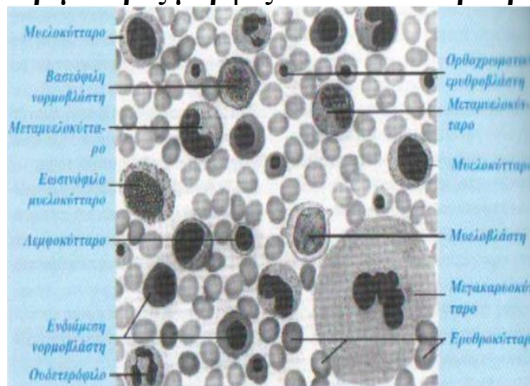
Ο πολλαπλασιασμός πραγματοποιείται με μιτωτική διαίρεση



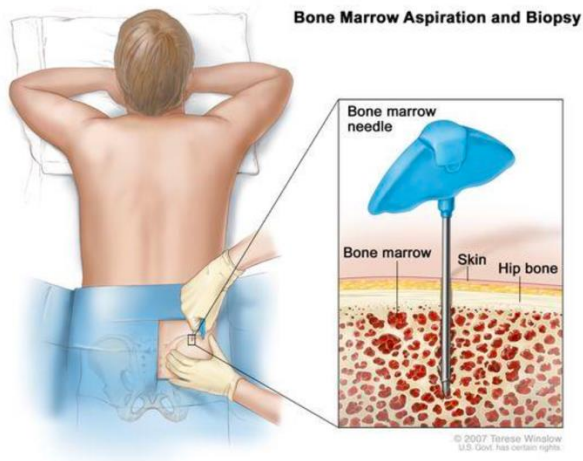
Στον φυσιολογικό μυελό επικρατούν:

Πρώιμες μορφές λευκών έναντι ερυθρών αιμοσφαιρίων , καθώς τα ερυθρά αιμοσφαίρια έχουν μακρότερο χρόνο ζωής (~ 3 μήνες) και αντικαθίστανται με βραδύτερο ρυθμό)

Ωριμότερες μορφές λευκών και ερυθρών



Μυελόγραμμα και διάγνωση αιματολογικών κακοηθειών



Το μυελόγραμμα δίνει πληροφορίες για την μορφολογία των κυττάρων του αίματος και την κυτταρική τους κατανομή στον ερυθρό μυελό.

Σε αιματολογικές παθήσεις όπως είναι το Πολλαπλούν Μυέλωμα, η Αναιμία και η Λευχαιμία παραδείγματος χάριν, εμφανίζονται **μεταβολές** στο μυελόγραμμα.

Η διενέργεια του μυελογράμματος προϋποθέτει την λήψη μυελικού πολφού η οποία συμβαίνει με **οστική παρακέντηση** από μία εκ των κάτωθεν πηγών:

Στέρνο,
Πλευρές,
Λαγόνια ακρολοφία

Η παρουσία αυξημένης συγκέντρωσης άωρων μορφών αιμοσφαιρίων αποτελεί σημείο εξαλλαγής του μυελού.

2.1.1 Αίμα

Τα κύτταρα του αίματος περιλαμβάνουν τα ερυθροκύτταρα, τα λευκοκύτταρα και τα αιμοπετάλια τα οποία είναι κυτταρικά θραύσματα.

Περισσότερο από το 99% των κυττάρων του αίματος είναι τα ερυθροκύτταρα, κύρια λειτουργία των οποίων είναι η μεταφορά του οξυγόνου στους ιστούς.

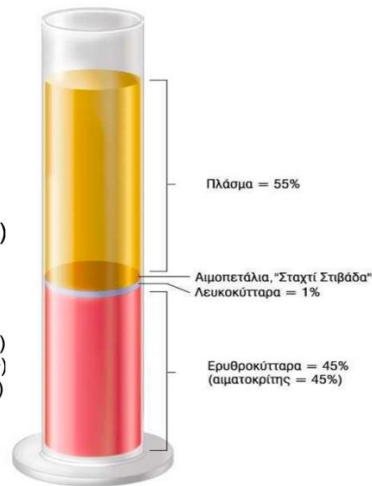
Τα λευκοκύτταρα προστατεύουν τον οργανισμό από τις μολύνσεις και τον καρκίνο, ενώ τα αιμοπετάλια συμβάλλουν στην αιμόσταση, δηλαδή στην πήξη του αίματος[4].

Αίμα Σύσταση

Πλάσμα (Εξωκυττάριο Υγρό)

Κυτταρικά στοιχεία

Ερυθρά αιμοσφαίρια (5×10^6)
Λευκά αιμοσφαίρια ($4-6 \times 10^3$)
Αιμοπετάλια ($0,15-0,3 \times 10^6$)



Πρωτεΐνες του πλάσματος

Ονομαστικά, οι πρωτεΐνες του πλάσματος είναι η Λευκωματίνη, η Αιμοσφαιρίνη η β1-Σφαιρίνη, η γ-Σφαιρίνη και το ινωδογόνο. Οι ρόλοι των άνωθεν πρωτεϊνών ληφθέντων ως σύνολο είναι οι εξής:

Συνεισφέρουν στην συγκράτηση του ύδατος εντός των αγγείων, στην μεταφορά ουσιών, όπου το πλάσμα είναι μια αποθήκη για τις ουσίες που ενωμένες με τις πρωτεΐνες του αποκτούν μια μεγαλομοριακή συγκρότηση και αποφεύγεται η νεφρική δίθηση όπως στην περίπτωση των βιταμινών, των ορμονών αλλά και των φαρμάκων.

Άμυνα: οι γ – σφαιρίνες είναι τα αντισώματα που βασίζεται η ανοσοποίηση του οργανισμού. Το ινωδογόνο αποτελεί βασικό πράγοντα του μηχανισμού της πήξης.

Διατήρηση οξεοβασικής ισορροπίας

Θρέψη κυττάρων: Οι πρωτεΐνες του πλάσματος ανανεώνονται συνεχώς και τα αινοξέα τους φέρονται στους ιστούς ως δομικά συστατικά των κυττάρων.

ΑΝΑΙΜΙΑ

Ως αναιμία ορίζεται:

Η μείωση του συνολικού αριθμού των ερυθροκυττάρων καθένα εκ των οποίων έχει φυσιολογική ποσότητα αιμοσφαιρίνης

Ή ως ελάττωση της συγκέντρωσης της αιμοσφαιρίνης ανά ερυθροκύτταρο ή ως ο συνδυασμός και των 2.

Οι κύριες αιτίες αναιμίας είναι η ανεπάρκεια σιδήρου, βιταμίνης B12, φυλλικού οξέος, ανεπάρκεια του μυελού των οστών, αιμορραγία, ανεπαρκής έκκρισης. Ερυθροποιητίνης, υπέρμετρη καταστροφή των ερυθροκυττάρων[4].

Λευκοκύτταρα:

Αποτελούν κινητές μονάδες άμυνας του οργανισμού που αναλαμβάνουν την καταστροφή μικροοργανισμών και παθολογικών κυττάρων, καθώς και την κάθαρση κυτταρικών υπολειμμάτων (φαγοκυττάρωση).

Συμβάλλουν στην αποκατάσταση της ιστικής βλάβης και οι διάφοροι τύποι λευκοκυττάρων ταξινομούνται σύμφωνα με τη δομή και τη χημική τους συγγένεια προς διάφορες χρωστικές ουσίες.

Παράγονται στον μυελό των οστών, και τα μονοκύτταρα καθώς και πολλά λεμφοκύτταρα υφίστανται περαιτέρω ανάπτυξη και κυτταρική διαίρεση σε ιστούς έξω από τον μυελό των οστών.

Το 30 – 40. % των λευκοκυττάρων παράγεται στον μυελό των οστών. Το 5% εξέρχεται από τον ερυθρό μυελό και κυκλοφορεί ελεύθερο στο αίμα είτε αποθηκεύεται εφεδρικά στο αίμα όπου λιμνάζει στα αγγεία του ήπατος του σπληνός και των πνευμόνων.

ΛΕΥΚΟΚΥΤΤΑΡΑ:

ΟΥΔΕΤΕΡΟΦΙΛΑ

Είναι το 50-70% ΟΛΩΝ ΤΩΝ ΛΕΥΚΟΚΥΤΤΑΡΩΝ, συμμετέχουν στις φλεγμονώδεις αντιδράσεις, ενεργοποιούν βακτηριοκτόνους μηχανισμούς και έχουν φαγοκυτταρική δράση.

ΗΩΣΙΝΟΦΙΛΑ

Αποτελούν 1-4% των λευκών αιμοσφαιρίων, λειτουργούν κατά των παρασίτων και συμμετέχουν στην παθογένεση των αλλεργικών αντιδράσεων

ΒΑΣΕΟΦΙΛΑ

0.5 % των λευκών αιμοσφαιρίων. Στο σημείο της φλεγμονής απελευθερώνουν **ισταμίνη** και **ηπαρίνη**. Συμμετέχουν σε αλλεργικές αντιδράσεις.

ΜΟΝΟΚΥΤΤΑΡΑ

Τα μονοκύτταρα αποτελούν 2-6% των λευκών αιμοσφαιρίων. Εγκαταλείπουν τα αγγεία (διαπήδηση) και φτάνουν σε ιστούς όπου μετατρέπονται σε μακροφάγα έχοντας φαγοκυτταρικές ιδιότητες.

ΛΕΜΦΟΚΥΤΤΑΡΑ

Τα λεμφοκύτταρα αποτελούν το 25-33% των λευκοκυττάρων. Τα Β-λεμφοκύτταρα παράγουν αντισώματα. Τα Τ- λεμφοκύτταρα καταστρέφουν άμεσα καρκινικά κύτταρα και κύτταρα που προσβάλλονται από ιούς. Ο λευκοκυτταρικός τύπος είναι η εκατοστιαία αναλογία μιας ομάδας λευκών αιμοσφαιρίων στον συνολικό αριθμό λευκών[4].

T – ΛΕΜΦΟΚΥΤΤΑΡΑ

Τα T-Λεμφοκύτταρα συμμετέχουν στην κυτταρική ανοσία και στην ανοσολογική μνήμη.

Παράγουν κυτταροτοξικές ουσίες και φέρουν αντιγόνα επιφανειακά. Παράγονται στον οστικό μυελό ως άωρα λεμφοκύτταρα και μεταναστεύουν στον θύμο, όπου ευαισθητοποιούνται στα αντιγονικά ερεθίσματα. Έπειτα εγκαθίστανται στο σπλήνα και στους λεμφαδένες απόπου εξορμούν στην κυκλοφορία και εγκαθίστανται στα λεμφαγγεία. Όταν συναντήσουν ειδικό αντιγόνο πολλαπλασιάζονται και σχηματίζουν κλώνους κυττάρων[4].

T-κυτταροτοξικά κύτταρα (T-cytotoxic/killer).

Είναι κύτταρα επίθεσης του οργανισμού προς τον στόχο. Εντοπίζουν τον στόχο, συνδέονται με αυτόν μέσω αντιγόνων και τον εξολοθρεύουν χωρίς την μεσολάβηση αντισωμάτων.

T-βοηθητικά κύτταρα (T-helper)

Ευοδώνουν την ενεργοποίηση και την λειτουργία των T-κυτταροτοξικών και των B-λεμφοκυττάρων.

B – ΛΕΜΦΟΚΥΤΤΑΡΑ

Διαφοροποιούνται σε **πλασματοκύτταρα** και εκκρίνουν αντισώματα. Διοχετεύονται στο σώμα για να συναντήσουν αντιγόνα όμοια με αυτά που προκάλεσαν την παραγωγή τους. Εγκαθίστανται σε σπλήνα και λεμφαδένες μετά την ευαισθητοποίηση[4].

2.2 Πολλαπλό μυέλωμα

2.2.1 Εισαγωγή

Στις Ηνωμένες Πολιτείες, σύμφωνα με δεδομένα από το τμήμα Επιδημιολογίας του Εθνικού Ινστιτούτου Καρκίνου, υπήρξαν περί τις 34470 νέες διαγνώσεις πολλαπλού μυελώματος (ΠΜ) το έτος 2022 οι οποίες κάλυψαν το 1,8% των συνολικών καρκινικών διαγνώσεων του έτους. Σύμφωνα με δεδομένα του ίδιου τμήματος, το 2019, κατά προσέγγιση 159.787 άνθρωποι ζούν με ΠΜ στις Ηνωμένες Πολιτείες[3].

Πέραν των Ηνωμένων Πολιτειών, παρατηρείται μη ενδεικνύμενη θεραπευτική προσέγγιση ανά τον κόσμο και κατά συνέπεια ανισότητα στην εξασφάλιση της ιατροφαρμακευτικής περίθαλψης των ασθενών όσον αφορά την διάγνωση και την αντιμετώπιση της νόσου. Εξίσου

σημαντική παράμετρος είναι και η ενημέρωση και η εκπαίδευση των ασθενών περί της νόσου για την συνολική διαχείριση του ΠΜ.

Η συνηθέστερη ηλικία των ασθενών που πρωτοδιαγιγνώσκονται με ΠΜ είναι 65-74 έτη, όμως σε αυτές τις νέες διαγνώσεις εμπíπτουν και άτομα ηλικία <50 ετών. Μόνο το 5-10% των ασθενών με ΠΜ είναι <40 έτη. Το ΠΜ στα παιδιά είναι εξαιρετικά σπάνιο.

Η νόσος εκδηλώνεται συχνότερα στους άρρενες και είναι συνηθέστερη σε άτομα Αφρικανικής καταγωγής ενώ η συχνότητα εμφάνισης ποικίλλει ανά φυλές.

Είναι σημαντικό να τονιστεί ότι το 5-7% των περιπτώσεων ΠΜ, συμβαίνει σε άτομα που έχουν στενή επαφή με συγγενείς σταδίου MGUS (βλ. **υποενότητα 1.2.2**), SMM ή ΠΜ.

Τα υγιή πλασματοκύτταρα συνιστούν απαραίτητη προϋπόθεση για την λειτουργία του ανοσιακού συστήματος (βλ. **υποενότητα 1.1**). Τα κύτταρα του μυελώματος είναι κακοήθη πλασματοκύτταρα που δεν δημιουργούν λειτουργικές ανοσοσφαιρίνες (αντισώματα) και εν αντιθέσει παράγουν μια ανώμαλη μονοκλωνική πρωτεΐνη, την M-protein [3].

Η ονομασία της νόσου έγκειται στο γεγονός ότι αφορά πολλαπλές περιοχές του σώματος, εκτός από μια ασυνήθη περίπτωση, το πλασματοκύττωμα. Οι περιοχές στις οποίες εκφράζεται το πολλαπλό μυέλωμα είναι ο μυελός των παρακάτω οστικών κοιλοτήτων:

1. Σπονδυλικής στήλης
2. Κρανίου
3. Πυέλου
4. Πλευρών
5. Όμων
6. Μηριαίων οστών

Στα μακρά οστά, επηρεάζονται κυρίως οι εγγύς αποφύσεις και όχι οι αποκακρυσμένες. Φαινοτυπικά, το πολλαπλό μυέλωμα (ΠΜ) εμφανίζεται ως κακοήθης μάζα ή/και ως περιοχή οστικής απώλειας όμως σε κάθε περίπτωση, οι περιοχές αυτές ονομάζονται κακοήθεις εξεργασίες.

Η παρουσία των κυττάρων του μυελώματος, δύνανται να δημιουργήσει ιατρικά προβλήματα τόσο στην μυελική κοιλότητα που αναπτύσσονται, όσο και στις περιοχές που αφορίζουν αυτές τις κοιλότητες.

Ενθαρρυντικό γεγονός για την πορεία της νόσου είναι ότι πολλές θεραπείες υψηλής αποτελεσματικότητας είναι εγκεκριμένες για την θεραπεία του ΠΜ από τον οργανισμό τροφίμων και φαρμάκων των Ηνωμένων Πολιτειών (FDA) καθώς και τον ευρωπαϊκό οργανισμό φαρμάκων (EMA). Παράλληλα, πολλές κλινικές δοκιμές που αφορούν την θεραπεία του ΠΜ βρίσκονται εν εξελίξει[3].

2.2.1.1 Πρώιμοι δείκτες έναρξης του ΠΜ

Το πιο πρώιμο στάδιο του μυελώματος είναι μια καλοήθης κατάσταση γνωστή και ως μονοκλωνική γ-πάθεια αδιευκρίνιστης σημασίας (MGUS). Όλοι οι ασθενείς με ΠΜ περνούν

πρώτα από αυτό το στάδιο αλλά μόνο το 20% αυτών των ασθενών που έχουν διαγνωσθεί με MGUS μεταβαίνουν στο στάδιο του ΠΜ.

Το στάδιο μεταξύ του σταδίου MGUS και του ενεργού ΠΜ είναι το ασυμπτωματικό μυέλωμα (smoldering multiple myeloma – SMM). Χαρακτηρίζεται από αυξημένα επίπεδα Μ – πρωτεΐνης από το στάδιο MGUS αλλά χωρίς ενδείξεις για ενεργό ΠΜ.

Οι ασθενείς με SMM χωρίς κάποιο προδιαθεσικό παράγοντα, κοινώς κανονικού ρίσκου, έχουν 10% πιθανότητα για να προχωρήσουν στο στάδιο του ενεργού ΠΜ για τα πρώτα 4 χρόνια, 3% για τα 5 επόμενα χρόνια και 1-2% πιθανότητα ανά χρόνο για τα επόμενα 10 χρόνια[3].

2.2.2 Αίτια και διάγνωση ΠΜ

Οποιαδήποτε συνθήκη που μπορεί να παρέμβει στο ανοσιακό σύστημα ή να το καταστρέφει, όπως πχ η μόλυνση με καρκινογενείς ιούς, μπορεί να προκαλέσει έναυσμα για το ΠΜ. Ορισμένα τοξικά χημικά που έχουν αναγνωρισθεί ως ανασοκατασταλτικά είναι τα:

1. Βενζένιο
2. Διοξίνες
3. Χημικά αγροτικών καλλιεργειών
4. Διαλύτες
5. Καύσιμα
6. Κανσαέρια
7. Καθαριστικά

Ενώ όσον αφορά τους ιούς που μπορούν να εκκινήσουν το ΠΜ, μέχρι στιγμής έχουν καταγραφεί οι εξής:

1. HIV
2. Ιός της ηπατίτιδας
3. Ιός του έρπη

Δεδομένου πως το ΠΜ είναι μια εξατομικευμένη ασθένεια που παρουσιάζει διαφορετικά φαινοτυπικά χαρακτηριστικά ανά οργανισμό, πολλές φορές μπορεί να εξελίσσεται αργά και άλλες φορές να είναι πολύ επιθετική νόσος[3].

Τα διαγνωστικά κριτήρια σύμφωνα με τον διεθνή ομάδα μελέτης του ΠΜ (IMWG) είναι τα εξής:

Μονοκλωνική γ-πάθεια αδιευκρίνιστης σημασίας (MGUS)

1. Μ-πρωτεΐνη χαμηλότερα απ'ότι επί Π.Μ. (Μ πρωτεΐνη < 3 g/dl)
2. Ποσοστό πλασματοκυττάρων στο μυελό <10%
3. Απουσία οστικής νόσου ή συμπτωμάτων συνδεδεμένων με μυέλωμα (νεφρική βλάβη, αναιμία) γνωστά και ως CRAB κριτήρια.
4. Απουσία ελάττωσης των μη κλωνικών ανοσοσφαιρινών

Ασυμπτωματικό μυέλωμα (SMM)

1. Μονοκλωνική πρωτεΐνη σε επίπεδα Π.Μ. (Μ πρωτεΐνη > 3 g/dl)
2. Ποσοστό πλασματοκυττάρων στο μυελό 10-60%
3. Ελάττωση των μη κλωνικών ανοσοσφαιρινών

4. Απουσία βλάβης τελικού οργάνου ή μεταβολικής διαταραχής

Συμπτωματικό μυέλωμα (πλασματοκύτταρα μυελού >10% και:)

1. Οποιαδήποτε βλάβη τελικού οργάνου σχετιζόμενη με μυέλωμα (αναιμία, υπερασβεστιαμία, οστική νόσος, νεφρική ανεπάρκεια, πλασμοκύττωμα κλπ)
2. Ένα ή περισσότερα εκ των κάτω προϋποθέσεων:
 - a. Παρουσία πλασματοκυττάρων στον οστικό μυελό > 60%
 - b. Λόγος των μη ελεύθερων προς ελεύθερων ελαφρών αλύσων των αντισωμάτων ≥ 100
 - c. ≥ 1 εστιακή αλλοίωση σε MRI, μεγέθους τουλάχιστον 5 mm
 - d. ≥ 1 οστεολυτική αλλοίωση σε CT ή PET-CT εξέταση

Τα στοιχεία a-c αποτελούν τα στοιχεία ορισμού του ΠΜ (MDE – myeloma defining events) και όταν παρατηρούνται ο ασθενής χρίζει την ανάγκη θεραπευτικής αγωγής.

Η λήψη πληροφοριών για τα MDE διενεργείται από τις παρακάτω εξετάσεις:

1. Βιοψία του οστικού μυελού
2. Τεστ ποσοτικοποίησης των ελαφρών αλύσων
3. MRI

Ιατρικά προβλήματα σχετιζόμενα με το ΠΜ:

Τα κύρια ιατρικά ζητήματα που σχετίζονται με το ΠΜ μπορούν να πινακοποιηθούν ως εξής [3]

Table 2.2-1: Επίδραση αυξημένων επιπέδων κυττάρων στον οστικό μυελό

Επιδράσεις των αυξημένων επιπέδων μυελώματος στον οστικό μυελό	ΑΙΤΙΑ	ΑΠΟΤΕΛΕΣΜΑ
Αυξημένα επίπεδα ασβεστίου στο αίμα	Απελευθέρωση ασβεστίου από τα καταστραμμένα οστά στο αίμα	Πνευματική σύγχυση Αφυδάτωση Πόνος Εξάντληση Αδυναμία Νεφρική βλάβη
Νεφρικά προβλήματα	Οι Μ πρωτεΐνες που παράγονται από τα μυελοκύτταρα, απελευθερώνονται στην αιματική κυκλοφορία, περνούν στα ούρα και προκαλούν νεφρική βλάβη. Επιπλέον,	Προβλήματα του κυκλοφορικού συστήματος Κόπωση Πνευματική σύγχυση
Αναιμία	Μείωση του αριθμού και της δραστηριότητας της παραγωγής των αρχέγονων ερυθροποιητικών κυττάρων στον οστικό μυελό	Κόπωση Αδυναμία
Οστικές Αλλοιώσεις	Τα κύτταρα του μυελώματος ενεργοποιούν τους οστεοκλάστες, που καταστρέφουν τα οστά και σταματούν την δράση των οστεοβλαστών που τα επαναδομούν.	Οστικά άλγη, Αυτόματα κατάγματα Οστική φλεγμονή Συμπίεση του νωτιαίου μυελού, και νευρική ζημία.
Ανώμαλη λειτουργία του ανοσιακού συστήματος	Τα κύτταρα του μυελώματος μειώνουν τον αριθμό και την δραστηριότητα των φυσιολογικών πλασματοκυττάρων που είναι ικανά να παράξουν αντισώματα εναντίον κάποιας λοίμωξης	Επιρρέπεια στην μόλυνση Καθυστερημένη επάνοδος από λοιμώξεις

2.2.2.1 MRI και Πολλαπλό Μυέλωμα

Το πρωτόκολλο που χρησιμοποιείται για την απεικόνιση του ΠΜ στην MRI συνίσταται από τις παρακάτω ακολουθίες[4] :

1. T1 Weighted-Turbo Spin Echo (T1W – TSE)
2. T2 Weighted Short Tau Inverse Recovery (T2W-STIR)
3. T1 – DIXON
4. T2 – DIXON
5. Diffusion Weighted Imaging (DWI)

Ακολουθούν ορισμένες πληροφορίες για τις άνωθεν ακολουθίες και την απεικόνιση του ΠΜ μέσω αυτών.

T1W-TSE:

Αποτελεί την καταλληλότερη ακολουθία για την αρχική μελέτη του οστικού μυελού. Ο λιπώδης μυελός εμφανίζεται με μεγάλο σήμα σε αυτήν την ακολουθία. Ο ερυθρός μυελός έχει μικρότερο ποσοστό, κατ' όγκον, λίπος και είναι κυτταροβριθής με αποτέλεσμα να έχει χαμηλότερο σήμα. Το καθαρό σήμα που λαμβάνεται από τις ακολουθίες αυτές εξαρτάται από την σύσταση σε λιπώδη και ερυθρό μυελό.

Ο παθολογικός οστικός μυελός, τυπικά έχει χαμηλότερο σήμα από τον λιπώδη καθότι είναι πιο κυτταροβριθής και περιέχει μεγαλύτερα ποσά ύδατος και λιγότερα ποσά λίπους (σε κατ' όγκον σύσταση).

Η διαφοροδιάγνωση ανάμεσα σε παθολογικό και ερυθρό μυελό, καθότι και οι δύο όντες εμφανιζόμενοι με χαμηλότερο σήμα στις εν λόγω ακολουθίες, έγκειται στο γεγονός ότι ο παθολογικός μυελός έχει χαμηλότερο σήμα από αυτό των μεσοσπονδύλιων δίσκων και των παρακείμενων μυών ενώ ο ερυθρός υψηλότερο[4].



Εικόνα 2.2-1: Λιπώδης μυελός σε απεικόνιση T1W[4]



Εικόνα 2.2-2: Φυσιολογικός ερυθρός μυελός [4]



Εικόνα 2.2-3: Παθολογικός μυελός [4]

T2W

Τα πρωτόνια στα λίπους έχουν μεγαλύτερο χρόνο αποδιέγερσης T2 συγκριτικά με αυτόν του ύδατος. Για τον λόγο αυτό, η ένταση του σήματος του λίπους είναι χαμηλότερη από αυτή των πλουσίων σε ύδωρ δομών όπως πχ των αλλοιώσεων του ΠΜ[4].

Στις εν λόγω ακολουθίες, η ένταση του σήματος του λίπους είναι παρόμοια ή ελαφρώς χαμηλότερη από αυτήν του υποδόριου λίπους. Ο ερυθρός μυελός έχει μια ενδιάμεση ένταση σήματος που είναι χαμηλότερη από αυτήν των αλλοιώσεων του οστικού μυελού. Η

διαφοροποίηση του ερυθρού από τον λιπώδη μυελό δεν είναι τόσο εύκολη όσο στις T1W ακολουθίες.

Εάν δεν χρησιμοποιηθούν τεχνικές καταστολής του σήματος του λίπους, η ένταση του λίπους παραμένει υψηλή. Το γεγονός αυτό δυσχαιρένει την μελέτη του μυελού καθότι τόσο ο λιπώδης μυελός όσο και μια αλλοίωση μεγάλης περιεκτικότητας κατ' όγκον σε ύδωρ θα εμφανίζονται με υψηλό σήμα. Κατά συνέπεια, η καταστολή του σήματος που προέρχεται από τον λιπώδη μυελό είναι αναγκαία συνθήκη για την απεικόνιση των αλλοιώσεων[4].

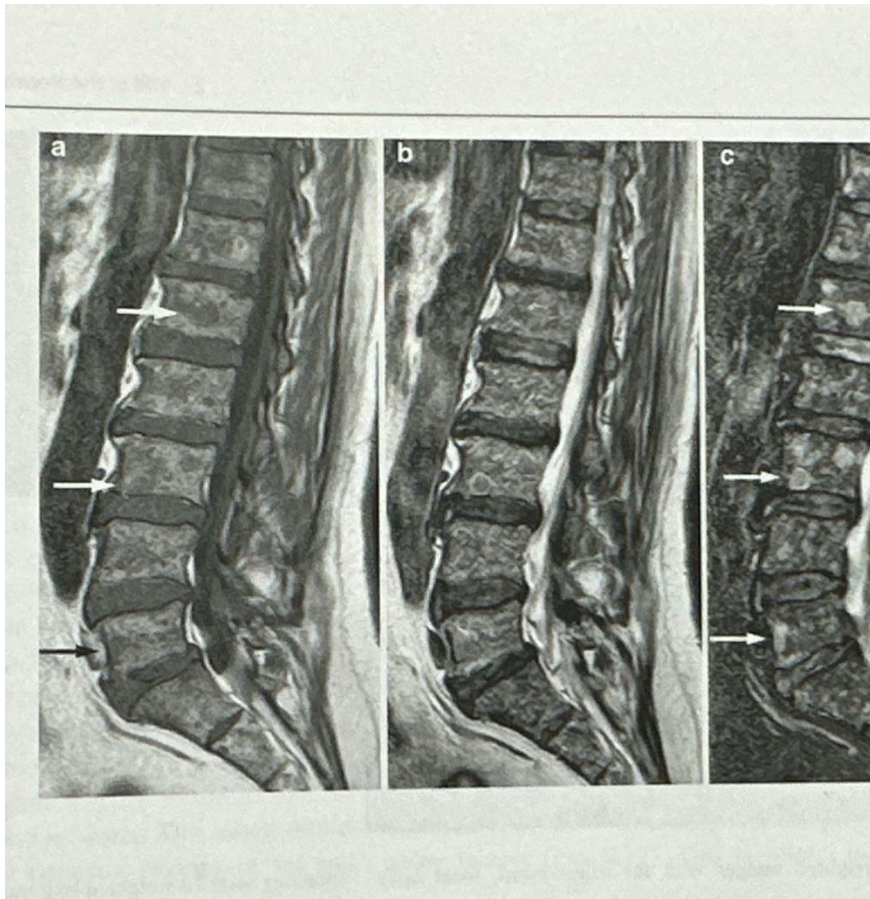
Η καταστολή του λίπους μπορεί να συμβεί με:

1. χημικώς επιλεκτική μέθοδο καταστολής
2. την τεχνική της μικρού χρονικού διαστήματος ανάστροφης αποδιέγερσης (short tau inversion recovery technique).

Στην 1^η περίπτωση εφαρμόζεται ραδιοπαλμός που καταστέλλει επιλεκτικά το σήμα των πρωτονίων του λίπους. Βέβαια, όσο μικρότερη είναι η διαφορά ανάμεσα στην χημική μετατόπιση (η χημική μετατόπιση είναι ανάλογη του κύριου μαγνητικού πεδίου του τομογράφου) των πρωτονίων του ύδατος και του λίπους, τόσο δυσκολότερη είναι η διαφοροποίησή τους. Εάν εφαρμοσθεί η 1^η τεχνική επιτυχώς τότε:

- A) Ο λιπώδης μυελός θα έχει χαμηλότερο σήμα από τους παρακείμενους μύες
- B) Ο ερυθρός μυελός θα έχει προσεγγιστικά την ίδια ένταση σήματος με τους παρακείμενους μύες
- C) Οι αλλοιώσεις θα έχουν υψηλότερο σήμα από αυτό των παρακείμενων μυών.

Στην 2^η περίπτωση, εφαρμόζεται ένας ραδιοπαλμός 180 μοιρών για την αναστροφή της μαγνήτισης όλων των απεικονιζομένων ιστών. Εφόσον εφαρμοσθεί ο ραδιοπαλμός, μετά του περατος χρόνο T1 (χρόνος αναστροφής, inversion time), εφαρμόζεται ένας ραδιοπαλμός 90 μοιρών για την γέννεση σήματος από όλους, εκτός του λίπους, ιστών. Καθότι οι ιστοί που δεν περιέχουν λίπος επανακτούν την εγκάρσια μαγνήτισή τους με μικρότερο ρυθμό από αυτόν των λιπώδων ιστών, ο λόγος σήματος προς θόρυβο (signal to noise ratio - SNR) είναι υψηλότερος για την τεχνική αυτή[4]. Το αποτέλεσμα είναι οι ιστοί που έχουν μακρό χρόνο αποδιέγερσης T1 και T2 να απεικονίζονται με υψηλό σήμα



Εικόνα 2.2-4: *Variegated pattern* πολλαπλού μυελώματος; α) T1W εικόνα πολλών focal βλαβών, β) T2W FSE εικόνα όπου οι βλάβες είναι υπερέντονες και δεν ξεχωρίζουν εύκολα, γ) STIR εικόνα όπου οι βλάβες είναι υποέντονες και ξεχωρίζουν από το υπόβαθρο [4]

DIXON

Οι ακολουθίες DIXON λαμβάνονται με GRE τεχνικές [1] αν και τα τελευταία χρόνια η λήψη τους συμβαίνει και με spin-echo μεθόδους. Συνίστανται από την λήψη εικόνων **σύμφωνης φάσης (in-phase ή IP)** και από την λήψη εικόνων **αντίθετης φάσης (opposed-phase ή OP)**[5].

Η ένταση σήματος (signal intensity ή SI) από τις IP εικόνες προέρχεται εκ ορισμού τόσο από τα πρωτόνια του ύδατος όσο και από τα πρωτόνια του λίπους του απεικονιζόμενου ιστού[6].

$$IP = W + F$$

Equation 2.2-1: Ένταση Σήματος IP

Αντίστοιχα, η SI από τις OP εικόνες προκύπτει από την αφαίρεση του σήματος των πρωτονίων του λίπους από το σήμα των πρωτονίων του ύδατος ενός απεικονιζόμενου ιστού.

$$OP = W - F$$

Equation 2.2-2: Ένταση Σήματος OP

Από τις σχέσεις 2.2-1 και 2.2-2 προκύπτουν δύο επιπρόσθετα κανάλια, αυτά του ύδατος (Water only image ή WO) και του λίπους (Fat only image ή FO), σχέσεις 2.2-3 και 2.2-4

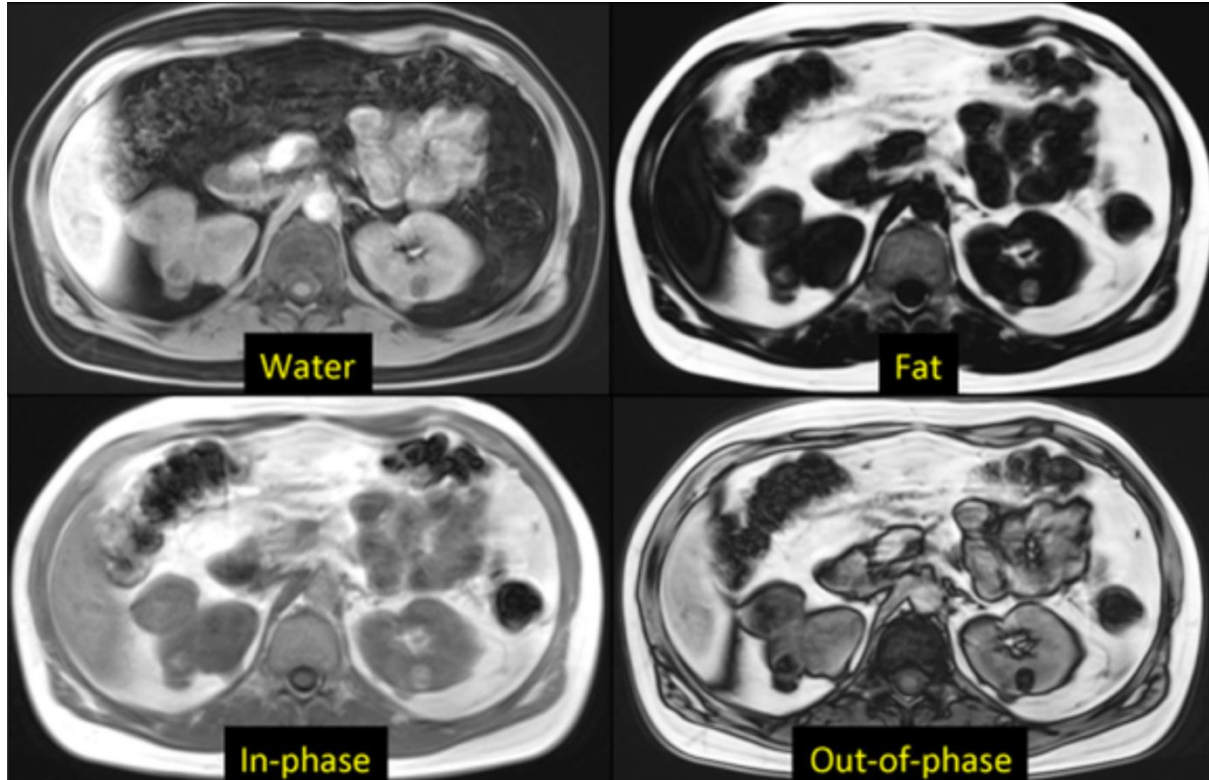
$$1/2 (IP + OP) = 1/2 [(W+F) + (W-F)] = W$$

Equation 2.2-3: Ένταση Σήματος νερού

$$1/2 (IP - OP) = 1/2 [(W+F) - (W-F)] = F$$

Equation 2.2-4: Ένταση Σήματος λίπους

Στην εικόνα 2.2-6 φαίνεται μια εγκάρσια τομή ακολουθίας DIXON όπου μπορούν να παρατηρηθούν οι φαινοτυπικές διαφορές διαφορές των ιστών ανά κανάλι.



Εικόνα 2.2-5: Εγκάρσια τομή DIXON εικόνας όπου διακρίνονται οι διαφορές μεταξύ των ιστών ανά κανάλι[6]

Οι αλλοιώσεις που υποκαθιστούν τον φυσιολογικό μυελό, στις ακολουθίες DIXON, θα έχουν μηδενική ένταση σήματος στο κανάλι FO. Οι ακολουθίες DIXON έχουν υψηλότερο λόγο αντίθεσης προς σήμα (contrast to noise ratio) συγκριτικά με τις ακολουθίες T1W, κατά συνέπεια οι αλλοιώσεις είναι πιο ευδιάκριτες και τα ψευδώς αρνητικά αποτελέσματα μειωμένα[5].

DWI

Η αντίθεση στις ακολουθίες διάχυσης μπορεί να ερμηνευθεί τόσο **ποσοτικά** όσο και **ποιοτικά**:

Ποιοτική διερμηνεία:

Στις ακολουθίες διάχυσης που έχουν ληφθεί με τεχνικές SSFP (υποκατηγορία των GRE ακολουθιών), η διαφοροποίηση μεταξύ καλοηθών και κακοηθών αλλοιώσεων συμβαίνει βάση οπτικής παρατήρησης.

Ποσοτική διερμηνεία:

Μπορεί να επιτευχθεί μέσω της λήψης εικόνων DWI χρησιμοποιώντας ≥ 2 τιμές βάρους διάχυσης και υπολογίζοντας τους χάρτες φαινομενικού δείκτη διάχυσης (apparent diffusion coefficient – ADC) από αυτές μέσω του τύπου:

$$ADC = \frac{\ln\left(\frac{S_{b_1}}{S_{b_0}}\right)}{(b_0 - b_1)}$$

Equation 2.2-5: Υπολογισμός ADC map

Οι συνηθέστερες τεχνικές για την λήψη των εικόνων διάχυσης είναι οι τεχνικές EPI και TSE.

Οι τιμές ADC:

1. Οι τιμές του λιπώδους μυελού είναι χαμηλότερες από αυτές του ερυθρού.
2. Θεωρείται ότι οι κύριοι παράγοντες που επηρεάζουν τον ADC είναι η σύσταση του μυελού σε λίπος και οι δοκίδες του σπογγώδους οστού.
3. Εκ των άνωθεν σημείων, προκύπτει πως παθολογικές διεργασίες που υποκαθιστούν τον φυσιολογικό μυελό θα εμφανίζονται ως περιοχές με αυξημένο ADC. Βέβαια, αυτό ισχύει μόνο για τον μυελό καθότι σε περιοχές μαλακών μορίων όπως είναι το ήπαρ και ο εγκέφαλος, οι αλλοιώσεις εμφανίζονται ως περιοχές μειωμένου ADC.
4. Του φυσιολογικού μυελού ανήκουν στο εύρος $\frac{0.2 - 0.5 * 10^{-3} mm^2}{s}$

Equation 2.2-6: Φυσιολογικές τιμές ADC

5. Του παθολογικού μυελού ανήκουν στο εύρος $\frac{0.7 - 1.2 * 10^{-3} mm^2}{s}$

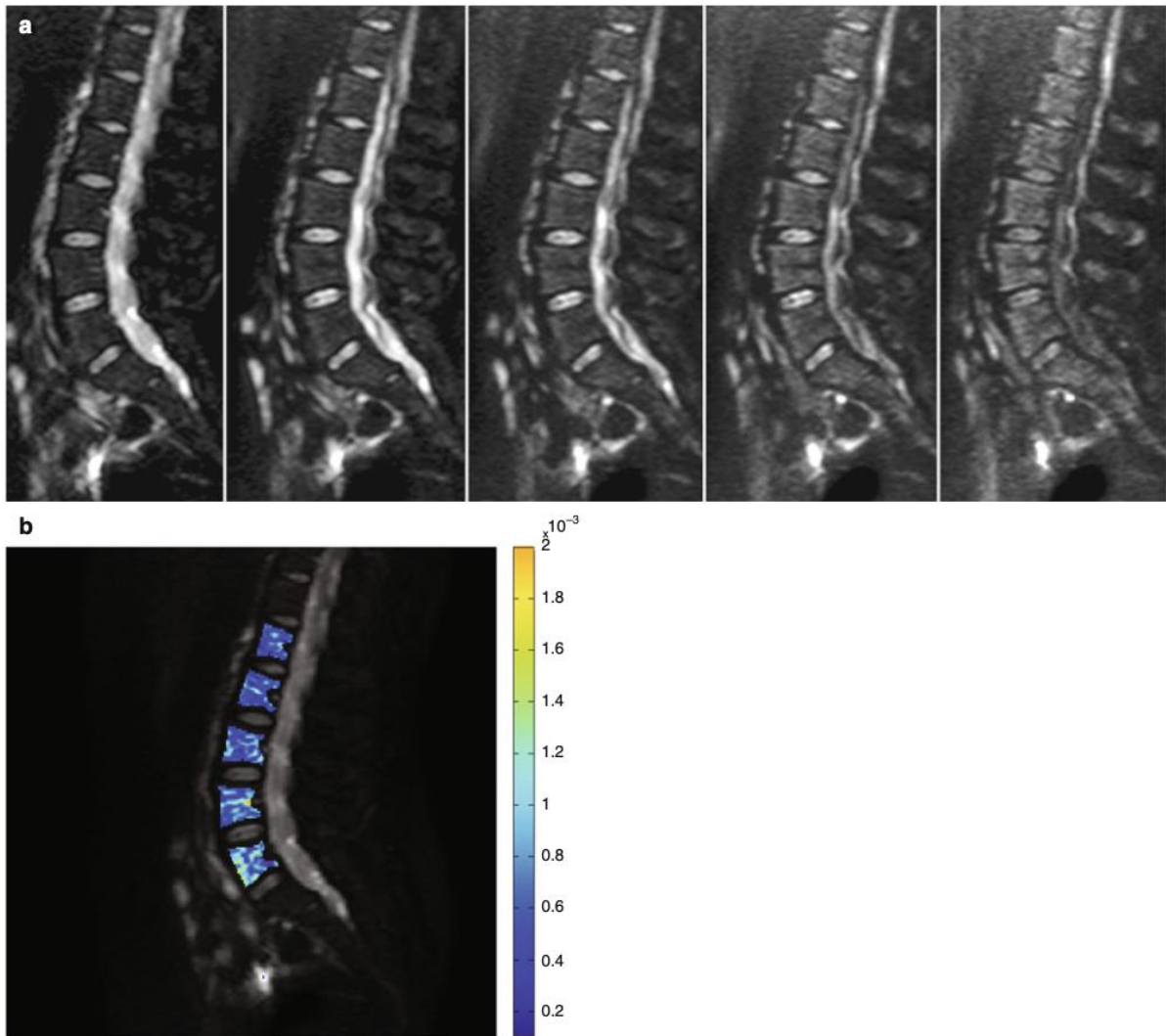
Equation 2.2-7: Παθολογικές τιμές ADC

6. Σε αυτό το σημείο τονίζεται πως στο διάχυτο πρότυπο του ΠΜ, οι τιμές του ADC μπορούν να επικαλύπτονται με αυτές του φυσιολογικού μυελού. Το εύρος τιμών στο οποίο μπορεί να παρατηρηθεί επικάλυψη είναι $\frac{0.5 - 0.7 * 10^{-3} mm^2}{s}$

Equation 2.2-8: Εύρος τιμών ADC όπου υπάρχει επικάλυψη μεταξύ υγιούς και παθολογικού μυελού

Συνοπτική ταξινόμηση φυσιολογικού και παθολογικού μυελού κατά αύξουσα σειρά τιμών ADC:

Λιπώδης μυελός < Ερυθρό μυελός < Διάχυτα παθολογικά πρότυπα < Εστιακά παθολογικά πρότυπα < Οξεία καλοήθη κατάγματα



Εικόνα 2.2-6: DWI εικόνα μυελού υγιούς νεαρού ενήλικα: α) Εικόνες με b -value = 0, 150, 250, 500, 750, β) Χάρτης ADC[4]

3 Κεφάλαιο 3: Τμηματοποίηση Εικόνας

3.1 Κλασικές Τεχνικές

Η τμηματοποίηση εικόνας είναι μια θεμελιώδης διαδικασία στην ανάλυση εικόνας και στην υπολογιστική όραση που περιλαμβάνει τον διαχωρισμό μιας εικόνας σε πολλαπλά τμήματα ή περιοχές, συχνά για να απομονωθούν και να ταυτοποιηθούν αντικείμενα ή όρια μέσα στην εικόνα. Ο στόχος είναι να απλοποιηθεί ή να αλλάξει η αναπαράσταση μιας εικόνας σε κάτι πιο σημαντικό και πιο εύκολο στην ανάλυση. Η τμηματοποίηση είναι κρίσιμη σε διάφορες εφαρμογές, συμπεριλαμβανομένης της ιατρικής απεικόνισης, όπου μπορεί να βοηθήσει στον εντοπισμό περιοχών ενδιαφέροντος, όπως όγκοι ή όργανα, για περαιτέρω ανάλυση, διάγνωση ή σχεδιασμό θεραπείας.

3.1.1 Κατωφλιοποίηση Εικόνας

Η τεχνική των κατωφλίων είναι μία από τις απλούστερες αλλά παράλληλα πιο αποτελεσματικές τεχνικές που χρησιμοποιούνται για την τμηματοποίηση εικόνων. Λειτουργεί επιλέγοντας μια συγκεκριμένη τιμή κατωφλίου και στη συνέχεια όλα τα εικονοστοιχεία στην εικόνα με ένταση πάνω (ή κάτω) από αυτή την τιμή κατατάσσονται σε ένα τμήμα, ενώ τα υπόλοιπα σε ένα άλλο. Η επιλογή της τιμής του κατωφλίου είναι κρίσιμη και μπορεί να καθοριστεί με βάση την ανάλυση ιστογραμμάτων, δοκιμή και σφάλμα, ή πιο προχωρημένες μεθόδους που βρίσκουν αυτόματα το βέλτιστο κατώφλι με βάση τα χαρακτηριστικά της εικόνας[70].

Έστω ότι ένα ιστόγραμμα έντασης αντιστοιχεί σε μια εικόνα, $f(x, y)$, που αποτελείται από φωτεινά αντικείμενα σε ένα σκοτεινό φόντο, με τέτοιο τρόπο ώστε τα εικονοστοιχεία του αντικειμένου και του φόντου να έχουν τιμές έντασης ομαδοποιημένες σε δύο κυρίαρχες λειτουργίες. Ένας προφανής τρόπος για να απομονώσετε τα αντικείμενα από το φόντο είναι να επιλέξετε ένα κατώφλι, T , που να χωρίζει αυτές τις λειτουργίες. Στη συνέχεια, κάθε σημείο (x, y) στην εικόνα στο οποίο $f(x, y) > T$ ονομάζεται σημείο αντικειμένου. Διαφορετικά, το σημείο ονομάζεται σημείο φόντου. Με άλλα λόγια, η τμηματοποιημένη εικόνα, που συμβολίζεται με $g(x, y)$, δίνεται από[70]:

$$g(x, y) = \begin{cases} 1 & \text{if } f(x, y) > T \\ 0 & \text{if } f(x, y) \leq T \end{cases}$$

Equation 3.1-1: Κατωφλιοποίηση εικόνας

Όταν το T είναι μια σταθερή τιμή που εφαρμόζεται σε όλη την εικόνα, η διαδικασία που δίνεται σε αυτή την εξίσωση αναφέρεται ως παγκόσμια τμηματοποίηση με κατώφλια. Όταν η τιμή του T αλλάζει σε μια εικόνα, χρησιμοποιούμε τον όρο μεταβλητή τμηματοποίηση με κατώφλια. Οι όροι τοπική ή περιφερειακή τμηματοποίηση με κατώφλια χρησιμοποιούνται μερικές φορές για να δηλώσουν μεταβλητή τμηματοποίηση στην οποία η τιμή του T σε οποιοδήποτε σημείο (x, y) σε μια εικόνα εξαρτάται από τις ιδιότητες μιας γειτονιάς του (x, y) (για παράδειγμα, τη μέση ένταση των εικονοστοιχείων στη γειτονιά). Αν το T εξαρτάται από τις χωρικές συντεταγμένες (x, y) καθεαυτές, τότε η μεταβλητή τμηματοποίηση με κατώφλια συχνά αναφέρεται ως δυναμική ή προσαρμοστική τμηματοποίηση με κατώφλια. Η χρήση αυτών των όρων δεν είναι παντού καθολική[70].

Ένα πιο δύσκολο πρόβλημα τμηματοποίησης, θα ήταν ένα πρόβλημα με κατώφλια που περιλαμβάνει ένα ιστόγραμμα με τρεις κυρίαρχες λειτουργίες που αντιστοιχούν, για παράδειγμα, σε δύο τύπους φωτεινών αντικειμένων σε ένα σκοτεινό φόντο. Εδώ, η πολλαπλή τμηματοποίηση με κατώφλια κατατάσσει ένα σημείο (x, y) ως ανήκον στο φόντο αν $f(x, y) \leq T_1$, σε μία κλάση αντικειμένου αν $T_1 < f(x, y) \leq T_2$, και στην άλλη κλάση αντικειμένου αν $f(x, y) > T_2$. Δηλαδή, η τμηματοποιημένη εικόνα δίνεται από[70]:

$$g(x, y) = \begin{cases} a & \text{if } f(x, y) > T_2 \\ b & \text{if } T_1 < f(x, y) \leq T_2 \\ c & \text{if } f(x, y) \leq T_1 \end{cases}$$

Equation 3.1-2: Κατωφλιοποίηση με δύο κατώφλια

όπου a , b , και c είναι οποιεσδήποτε τρεις διακριτές τιμές έντασης. Προβλήματα τμηματοποίησης που απαιτούν περισσότερα από δύο κατώφλια είναι δύσκολα (ή συχνά αδύνατα) να επιλυθούν, και καλύτερα αποτελέσματα συνήθως επιτυγχάνονται χρησιμοποιώντας άλλες μεθόδους, όπως η μεταβλητή τμηματοποίηση με κατώφλια ή η ανάπτυξη περιοχής[70].

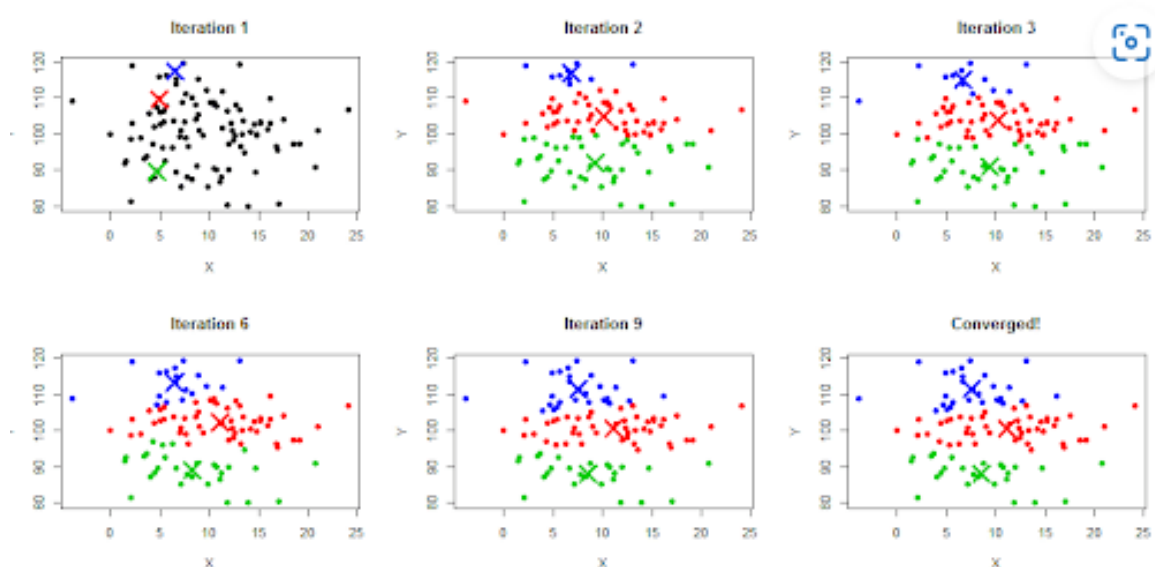
Με βάση τα παραπάνω, μπορούμε να συμπεράνουμε διαισθητικά ότι η επιτυχία της τμηματοποίησης με κατώφλια σχετίζεται άμεσα με το πλάτος και το βάθος της(των) κοιλάδας(ων) που χωρίζει τις λειτουργίες του ιστογράμματος.

3.1.2 k-means segmentation

Η συσταδοποίηση k-means είναι μια δημοφιλής μέθοδος που χρησιμοποιείται για την τμηματοποίηση εικόνων, η οποία διαιρεί μια εικόνα σε K διακριτές συστάδες. Σε αντίθεση με την τμηματοποίηση με κατώφλια, η οποία βασίζεται στην ένταση των εικονοστοιχείων, η k-means λαμβάνει υπόψη τα χαρακτηριστικά των εικονοστοιχείων, όπως το χρώμα ή την ένταση, για να τα ομαδοποιήσει σε συστάδες. Αυτή η μέθοδος μπορεί να είναι ιδιαίτερα αποτελεσματική όταν μια εικόνα περιέχει πολλαπλά αντικείμενα ενδιαφέροντος ή όταν η αντίθεση μεταξύ των αντικειμένων και του φόντου δεν είναι ομοιόμορφη.

Η k-means είναι ένας επαναληπτικός αλγόριθμος που διαιρεί ένα σύνολο δεδομένων (σε αυτή την περίπτωση, εικονοστοιχεία) σε K μη-επικαλυπτόμενα υποσύνολα (συστάδες) με βάση την ομοιότητα των χαρακτηριστικών τους. Κάθε συστάδα χαρακτηρίζεται από το κεντροειδές της, το οποίο είναι ο μέσος όρος των σημείων (εικονοστοιχείων) που έχουν ανατεθεί στη συστάδα. Ο στόχος της συσταδοποίησης k-means στην τμηματοποίηση εικόνας είναι να διαιρέσει την εικόνα έτσι ώστε τα εικονοστοιχεία εντός της ίδιας συστάδας να είναι πιο παρόμοια μεταξύ τους απ' ό,τι με εικονοστοιχεία σε άλλες συστάδες, με βάση μια προκαθορισμένη μέτρηση ομοιότητας, συχνά την Ευκλείδεια απόσταση[72].

Η διαδικασία χρήσης της k-means για την τμηματοποίηση εικόνας συνήθως περιλαμβάνει τα ακόλουθα βήματα:



Εικόνα 3.1-1: Η διαδικασία δημιουργίας clusters από τον αλγόριθμο k-means[67]

Αρχικοποίηση: Επιλογή K αρχικών κεντροειδών, είτε τυχαία είτε με βάση κάποια ευριστική. Αυτά τα κεντροειδή θα μπορούσαν να αντιπροσωπεύουν το μέσο χρώμα (στον χώρο RGB ή σε κάποιον άλλο χώρο χρωμάτων) ή την ένταση διαφορετικών τμημάτων που αναμένετε να βρείτε στην εικόνα.

Ανάθεση: Ανάθεση κάθε εικονοστοιχείου της εικόνας στο κοντινότερο κεντροειδές, με βάση την Ευκλείδεια απόσταση μεταξύ του χαρακτηριστικού του εικονοστοιχείου (π.χ. του διανύσματος χρώματος του στον χώρο RGB) και του χαρακτηριστικού του κεντροειδούς. Αυτό το βήμα διαιρεί την εικόνα σε συστάδες, με κάθε συστάδα να αντιστοιχεί σε ένα τμήμα.

Ενημέρωση: Επανυπολογισμός των κεντροειδών των συστάδων παίρνοντας το μέσο όρο όλων των εικονοστοιχείων που είναι τρέχοντα ανατεθειμένα σε κάθε συστάδα.

Επανάληψη: Επανάληψη των βημάτων ανάθεσης και ενημέρωσης μέχρι τα κεντροειδή να μην αλλάζουν πλέον σημαντικά ή να επιτευχθεί ένας προκαθορισμένος αριθμός επαναλήψεων. Αυτή η επαναληπτική διαδικασία διασφαλίζει ότι οι συστάδες είναι περισσότερο διευκρινισμένες και αντιπροσωπεύουν καλύτερα την τμηματοποίηση της εικόνας.

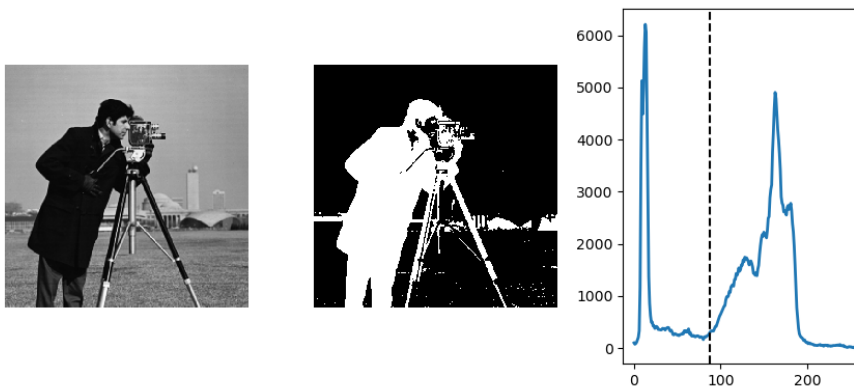
Αποτέλεσμα: Κάθε συστάδα αντιπροσωπεύει ένα τμήμα της εικόνας. Τα εικονοστοιχεία εντός της ίδιας συστάδας μπορούν να θεωρηθούν ως ανήκοντα στο ίδιο τμήμα ή αντικείμενο μέσα στην εικόνα[72].

Η K -means είναι μια δημοφιλής μέθοδος για την τμηματοποίηση εικόνας λόγω της απλότητας, ευελιξίας και αποδοτικότητάς της. Αυτή η μέθοδος είναι απλή στην κατανόηση και στην υλοποίηση, απαιτώντας ελάχιστη προεπεξεργασία και ρύθμιση παραμέτρων, γεγονός που την καθιστά προσβάσιμη για διάφορες εφαρμογές. Η ευελιξία της στην τμηματοποίηση εικόνων βασισμένη σε διάφορα χαρακτηριστικά, όχι μόνο στην ένταση, επιτρέπει μια πιο λεπτομερή τμηματοποίηση που μπορεί να αποτυπώσει χρωματικές υφές ή μοτίβα. Πέρα από αυτό, η K -means είναι υπολογιστικά αποδοτική για σύνολα δεδομένων μέτριου μεγέθους, καθιστώντας την κατάλληλη για εργασίες τμηματοποίησης εικόνας σε πραγματικό χρόνο ή σχεδόν σε πραγματικό χρόνο, προσφέροντας μια ισχυρή λύση για την ανάλυση και επεξεργασία εικόνων.

Η χρήση της K -means για την τμηματοποίηση εικόνας φέρνει μαζί της κάποιους περιορισμούς και σημαντικές σκέψεις που πρέπει να ληφθούν υπόψη. Ένας από τους βασικούς παράγοντες είναι η επιλογή του αριθμού των συστάδων K , η οποία πρέπει να γίνει εκ των προτέρων και μπορεί να είναι δύσκολη αν ο αριθμός των τμημάτων στην εικόνα δεν είναι γνωστός από πριν. Ευριστικές μέθοδοι, όπως η Μέθοδος του Αγκώνα ή η Μέθοδος του Silhouette, μπορούν να βοηθήσουν στην εκτίμηση του βέλτιστου K . Η αρχική επιλογή των κεντροειδών μπορεί επίσης να έχει σημαντικό αντίκτυπο στην τελική τμηματοποίηση, και η κακή αρχικοποίηση μπορεί να οδηγήσει σε υποβέλτιστες συστάδες. Τεχνικές όπως η k -means++ μπορούν να βελτιώσουν αυτήν την αρχικοποίηση. Επιπλέον, η K -means μπορεί να είναι ευαίσθητη στον θόρυβο και στα ακραία δεδομένα, καθώς μπορούν να παραμορφώσουν τον υπολογισμό των κεντροειδών. Τέλος, η επιλογή του χώρου χρώματος, όπως RGB ή HSV, μπορεί να επηρεάσει τα αποτελέσματα της τμηματοποίησης, καθώς μερικοί χώροι μπορεί να διαχωρίζουν καλύτερα τα χαρακτηριστικά ενδιαφέροντος σε μια εικόνα.

3.1.3 Αλγόριθμος Τμηματοποίησης Otsu

Η μέθοδος τμηματοποίησης ή διακριτοποίησης εικόνας του Otsu είναι μια παγκοσμίως χρησιμοποιούμενη τεχνική για την αυτόματη τμηματοποίηση ή διακριτοποίηση εικόνων, βασισμένη στην τμηματοποίηση με κατώφλια μέσω συσταδοποίησης. Αναπτύχθηκε από τον Nobuyuki Otsu το 1979. Αυτή η μέθοδος επιλέγει το κατώφλι ώστε να ελαχιστοποιεί την ενδοκλάσης διακύμανση των μαύρων και λευκών εικονοστοιχείων, το οποίο ισοδυναμεί με τη μεγιστοποίηση της μεταξύ κλάσεων διακύμανσης. Η μέθοδος του Otsu υποθέτει ότι η εικόνα περιέχει δύο κλάσεις εικονοστοιχείων (προσκήνιο και φόντο) και ότι το ιστόγραμμα των εντάσεων των εικονοστοιχείων μέσα στην εικόνα μπορεί να μοντελοποιηθεί ως μια διμορφική κατανομή. Αναζητά να βρει ένα κατώφλι που διαχωρίζει αυτές τις δύο κλάσεις έτσι ώστε η συνδυασμένη διασπορά τους (ενδοκλάσης διακύμανση) να είναι ελάχιστη, ή, ισοδύναμα, η απόσταση μεταξύ των κλάσεων (μεταξύ κλάσεων διακύμανση) να είναι μέγιστη[73].



Εικόνα 3.1-2: Η επιλογή κατωφλίου από τον αλγόριθμο Otsu[68]

Η διαδικασία περιλαμβάνει αρκετά βήματα:

Υπολογισμός Ιστογράμματος: Αρχικά, υπολογίζεται το ιστόγραμμα της εικόνας σε κλίμακα του γκρι, όπου καταγράφεται η συχνότητα κάθε εντάσεως εικονοστοιχείου.

Κανονικοποίηση: Κανονικοποιείται το ιστόγραμμα έτσι ώστε να αντιπροσωπεύει μια κατανομή πιθανότητας των εντάσεων των εικονοστοιχείων.

Συσσωρευτικά Αθροίσματα: Υπολογίζεται το συσσωρευτικό άθροισμα των πιθανοτήτων, το οποίο λειτουργεί ως βάρη για το προσκήνιο και το φόντο.

Συσσωρευτικά Μέσα: Υπολογίζονται τα συσσωρευτικά μέσα μέχρι το επίπεδο του κατωφλίου για το προσκήνιο και το φόντο.

Διακύμανση Κλάσης: Υπολογίζεται η διακύμανση και για τις δύο κλάσεις, προσκήνιο και φόντο.

Μεταξύ Κλάσεων Διακύμανση: Για κάθε τιμή κατωφλίου, υπολογίζεται η διακύμανση μεταξύ κλάσεων, η οποία είναι το σταθμισμένο άθροισμα των διακυμάνσεων των κλάσεων.

Βέλτιστο Κατώφλι: Το βέλτιστο κατώφλι είναι εκείνο που μεγιστοποιεί τη διακύμανση μεταξύ κλάσεων. Με μαθηματικούς όρους, αν συμβολίσουμε τις κλάσεις ως C_0 (φόντο) και C_1 (προσκήνιο), και τις πιθανότητές τους ως ω_0 και ω_1 , και τα μέσα τους ως μ_0 και μ_1 , ο στόχος είναι να μεγιστοποιηθεί η συνάρτηση:

$$\sigma_B^2(t) = \omega_0(t)\omega_1(t)[\mu_0(t) - \mu_1(t)]^2$$

Equation 3.1-3: Εύρεση βέλτιστου κατώφλιου στη μέθοδο Otsu

όπου t είναι το κατώφλι, $\omega(t)$ είναι η διακύμανση μεταξύ κλάσεων για το κατώφλι t , ω αντιπροσωπεύει τις πιθανότητες των κλάσεων, και μ είναι τα μέσα των κλάσεων. Το κατώφλι που μεγιστοποιεί το $\omega(t)$ επιλέγεται ως το βέλτιστο κατώφλι για την τμηματοποίηση[73].

Η υλοποίηση και εφαρμογή της μεθόδου Otsu περιλαμβάνει την ευκολία χρήσης ως ένα από τα κύρια πλεονεκτήματά της, καθώς είναι αυτόματη και δεν απαιτεί χειροκίνητη επιλογή του κατώφλιου. Αυτό την καθιστά ιδιαίτερα χρήσιμη σε εφαρμογές όπου η αντίθεση της εικόνας διαφέρει ή κατά την επεξεργασία μεγάλου όγκου εικόνων.

Η μέθοδος του Otsu έχει βρει εφαρμογές σε διάφορους τομείς, συμπεριλαμβανομένων της σάρωσης εγγράφων, της ιατρικής απεικόνισης για την ανίχνευση όγκων και της προεπεξεργασίας σε εργασίες υπολογιστικής όρασης.

Ωστόσο, η μέθοδος αντιμετωπίζει ορισμένους περιορισμούς:

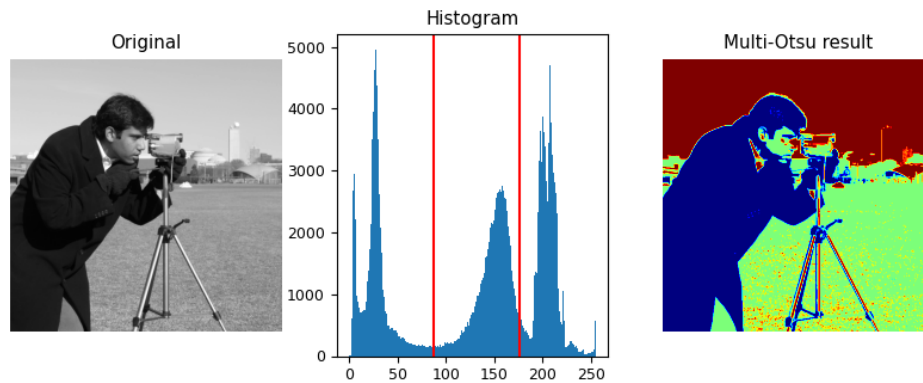
Υπόθεση Διμορφικότητας: Η μέθοδος του Otsu υποθέτει μια διμορφική κατανομή των εντάσεων των εικονοστοιχείων, η οποία μπορεί να μην υπάρχει πάντα σε πολύπλοκες εικόνες.

Ομοιογένεια: Λειτουργεί καλύτερα όταν η εικόνα έχει μεγάλες ομοιογενείς περιοχές προσκηνίου και φόντου.

Ευαισθησία στον Θόρυβο: Όπως και άλλες μέθοδοι παγκόσμιας τμηματοποίησης με κατώφλια, η μέθοδος του Otsu μπορεί να είναι ευαίσθητη στον θόρυβο και στη μη ομοιόμορφη φωτισμό της εικόνας.

3.1.4 Πολλαπλή Κατωφλιοποίηση Otsu

Ένας από τους τρόπους να επιλυθεί το πρόβλημα της διμορφικότητας είναι η μέθοδος Otsu πολλαπλών κατωφλίων. Η μέθοδος του Otsu για την τμηματοποίηση εικόνας, γνωστή για την απλότητα και την αποτελεσματικότητά της στη διακριτοποίηση με δυαδικά κατώφλια, μπορεί να επεκταθεί ώστε να υποστηρίζει πολλαπλά κατώφλια, επιτρέποντας έτσι την τμηματοποίηση μιας εικόνας σε περισσότερες από δύο κλάσεις. Αυτή η επέκταση, γνωστή ως πολλαπλή τμηματοποίηση με τη μέθοδο Otsu, είναι ιδιαίτερα χρήσιμη για εικόνες που περιέχουν αρκετές διακριτές περιοχές ενδιαφέροντος, οι οποίες χαρακτηρίζονται από διαφορετικά επίπεδα έντασης[7].



Εικόνα 3.1-3: Η επιλογή πολλών κατωφλίων από τον αλγόριθμο Otsu Multiple Thresholding[7]

Η πολλαπλή τμηματοποίηση με τη μέθοδο Otsu λειτουργεί με βάση την ίδια αρχή με την αρχική μέθοδο, στοχεύοντας να βρει κατωφλιακές τιμές που ελαχιστοποιούν την ενδοκλάσης διακύμανση (ή, ισοδύναμα, μεγιστοποιούν τη μεταξύ κλάσεων διακύμανση) των προκύπτουσων τμημάτων. Ωστόσο, αντί να αναζητά μία μόνο τιμή κατωφλίου για να διαιρέσει την εικόνα σε δύο κλάσεις (προσκήνιο και φόντο), η τεχνική της πολλαπλής τμηματοποίησης αναζητά να ταυτοποιήσει αρκετές τιμές κατωφλίου, διαιρώντας έτσι την εικόνα σε πολλαπλά τμήματα[7].

Η διαδικασία για τον υπολογισμό πολλαπλών κατωφλίων περιλαμβάνει την επανάληψη μέσα από συνδυασμούς πιθανών τιμών κατωφλίου για να προσδιοριστεί το σετ που μεγιστοποιεί τη μεταξύ κλάσεων διακύμανση μεταξύ των προκύπτουσων κλάσεων. Για μια εικόνα που απαιτεί N κατωφλιακές τιμές, αυτό περιλαμβάνει:

Υπολογισμός Ιστογράμματος: Αρχικά, υπολογίζεται και κανονικοποιείται το ιστόγραμμα της εικόνας για να αντιπροσωπεύει την κατανομή πιθανότητας των εντάσεων των εικονοστοιχείων.

Συσσωρευτικά Αθροίσματα και Μέσοι Όροι: Υπολογίζονται τα συσσωρευτικά αθροίσματα και οι συσσωρευτικοί μέσοι όροι μέχρι κάθε επίπεδο έντασης στο ιστόγραμμα. Αυτά θα χρησιμοποιηθούν για τον υπολογισμό των μέσων όρων και των διακυμάνσεων των τμημάτων που ορίζονται από τις πιθανές τιμές κατωφλίου.

Υπολογισμοί Διακύμανσης: Για κάθε συνδυασμό τιμών κατωφλίου, υπολογίζεται η ενδοκλάσης διακύμανση για κάθε τμήμα και η μεταξύ κλάσεων διακύμανση μεταξύ των τμημάτων. Η ενδοκλάσης διακύμανση είναι το σταθμισμένο άθροισμα των διακυμάνσεων εντός κάθε τμήματος, και η μεταξύ κλάσεων διακύμανση είναι μια μέτρηση της διαχωριστικότητας μεταξύ των τμημάτων.

Βελτιστοποίηση: Το σετ των κατωφλίων που μεγιστοποιεί τη μεταξύ κλάσεων διακύμανση (και ελαχιστοποιεί την ενδοκλάσης διακύμανση) σε όλα τα τμήματα επιλέγεται ως τα βέλτιστα κατωφλιακές τιμές. Αυτό μπορεί να είναι υπολογιστικά απαιτητικό για ένα μεγάλο αριθμό κατωφλίων.

Η μέθοδος του Otsu με πολλαπλά κατωφλιακές τιμές χρησιμοποιείται σε διάφορες εφαρμογές όπου οι εικόνες περιέχουν περισσότερους από δύο διακριτούς τύπους περιοχών ή αντικειμένων, όπως στην ιατρική απεικόνιση για την τμηματοποίηση διαφορετικών τύπων ιστών, στην απομακρυσμένη ανίχνευση για την ταξινόμηση καλύψεων εδάφους και στην

ανάλυση εικόνων εγγράφων. Αυτή η μέθοδος προσφέρει μια σημαντική λύση για την αποδοτική και ακριβή τμηματοποίηση εικόνων με πολλαπλές περιοχές ενδιαφέροντος, παρά τις προκλήσεις που συνδέονται με την υπολογιστική της πολυπλοκότητα και την επιλογή του αριθμού των κατωφλίων[7].

3.2 Νευρωνικά Δίκτυα

Η εξέλιξη των CNNs ξεκινά από την αντιγραφή βιολογικών συστημάτων όρασης. Το οπτικό σύστημα των ζώων, και ειδικότερα οι νευρωνικές δομές του οπτικού φλοιού, προσέφεραν την έμπνευση για την αρχική ιδέα πίσω από τα CNNs. Σημαντικά βήματα στην ανάπτυξη των CNNs περιλαμβάνουν το Neocognitron[8] του Kunihiko Fukushima στη δεκαετία του 1980 και το LeNet-5[8] του Yann LeCun στα τέλη της δεκαετίας του 1990, το οποίο αποτέλεσε μία από τις πρώτες επιτυχημένες εφαρμογές των CNNs στην αναγνώριση αριθμών.

3.2.1 Βασικά Συστατικά και Λειτουργικοί Μηχανισμοί

Τα στρώματα συνέλιξης αποτελούν την καρδιά των CNNs[9], [10]. Μέσω της χρήσης φίλτρων ή πυρήνων, αυτά τα στρώματα εκτελούν λειτουργίες συνέλιξης, εξάγοντας χαρακτηριστικά όπως ακμές, γωνίες και υφές από τις εισαχθείσες εικόνες. Η διαδικασία εκμάθησης των φίλτρων είναι κρίσιμη και συντελείται μέσω τεχνικών όπως η αντίστροφη διάδοση και η καταβαράθρωση της κλίσης[11]. Επιπλέον, τα στρώματα pooling παίζουν σημαντικό ρόλο στη μείωση του χωρικού μεγέθους της εικόνας, μειώνοντας την πολυπλοκότητα και τον όγκο των υπολογισμών και βοηθώντας στην πρόληψη της υπερπροσαρμογής.



Εικόνα 3.2-1: Εξέλιξη και βασικά χαρακτηριστικά νευρωνικών δικτύων

3.2.1.1 *Επισκόπηση του Pooling*

Το pooling, επίσης γνωστό ως υποδειγματοληψία ή μείωση διαστηματικότητας, αποτελεί ένα ζωτικής σημασίας στάδιο στη διαδικασία των συνελκτικών νευρωνικών δικτύων (CNNs). Η βασική του λειτουργία είναι η μείωση της διαστατικότητας των χαρτογραφημάτων χαρακτηριστικών, διατηρώντας ταυτόχρονα τις πιο σημαντικές πληροφορίες.[12] Αυτό επιτρέπει στο δίκτυο να αποκτά μια μορφή ανεξαρτησίας από μικρομεταβολές στη θέση των χαρακτηριστικών στην εικόνα εισόδου, ενισχύοντας έτσι την ανθεκτικότητα και την απόδοσή του.

Μέγιστο Pooling

Στην τεχνική του μέγιστου pooling, το δίκτυο επιλέγει τη μέγιστη τιμή από κάθε τμήμα της εικόνας που καλύπτεται από τον πυρήνα. Για παράδειγμα, σε ένα επίπεδο μέγιστου pooling 2x2, η μέγιστη τιμή από κάθε τετράγωνο 2x2 της εικόνας εισόδου επιλέγεται και χρησιμοποιείται. Αυτή η μέθοδος είναι δημοφιλής για την αποτελεσματικότητά της στη μείωση των υπολογισμών και στην εξαγωγή ισχυρών χαρακτηριστικών, ενώ παράλληλα παρέχει ανθεκτικότητα σε μικρές μετατοπίσεις. Ωστόσο, ένα μειονέκτημα είναι ότι μπορεί να οδηγήσει σε απώλεια σημαντικών πληροφοριών, καθώς διατηρεί μόνο τη μέγιστη τιμή από κάθε τμήμα.[13]

3.2.2 *Συναρτήσεις Ενεργοποίησης: Εισαγωγή της Μη Γραμμικότητας στα CNNs*

Στα Συνελκτικά Νευρωνικά Δίκτυα (CNNs), οι συναρτήσεις ενεργοποίησης είναι ζωτικής σημασίας για την εισαγωγή μη γραμμικών ιδιοτήτων. Αυτή η μη γραμμικότητα επιτρέπει στο δίκτυο να αναλύσει και να μάθει πιο πολύπλοκα πρότυπα από τα δεδομένα[14]. Στο πλαίσιο αυτό, εξετάζουμε διάφορες συναρτήσεις ενεργοποίησης, τις λειτουργίες τους, τις εφαρμογές τους και άλλες σχετικές λεπτομερείες.

3.2.2.1 *Επισκόπηση των Συναρτήσεων Ενεργοποίησης*

Οι συναρτήσεις ενεργοποίησης λαμβάνουν μία είσοδο και εκτελούν μια συγκεκριμένη μαθηματική λειτουργία πάνω σε αυτήν. Στα νευρικά δίκτυα, αυτές εφαρμόζονται στην έξοδο κάθε νευρώνα και επηρεάζουν τη μετάδοση του σήματος στο επόμενο επίπεδο. Η επιλογή της συνάρτησης ενεργοποίησης είναι καθοριστική για την ικανότητα του δικτύου να μαθαίνει αποτελεσματικά και για την ταχύτητα της σύγκλισής του.[15]

3.2.2.2 *Κοινές Συναρτήσεις Ενεργοποίησης στα CNNs*

ReLU (Rectified Linear Unit): Η ReLU είναι μία από τις πιο διαδεδομένες συναρτήσεις ενεργοποίησης και λειτουργεί επιστρέφοντας την είσοδο αν είναι θετική, ενώ για αρνητικές τιμές επιστρέφει μηδέν. Μαθηματικά, ορίζεται ως $f(x) = \max(0, x)$. Η ReLU και οι παραλλαγές της, όπως η Leaky ReLU και η Parametric ReLU, χρησιμοποιούνται ευρέως στα σύγχρονα νευρωνικά δίκτυα. Το κύριο πλεονέκτημα της ReLU είναι ότι βοηθά στην αντιμετώπιση του προβλήματος του εξαφανιζόμενου gradient, διευκολύνοντας την εκμάθηση. Ωστόσο, μπορεί να εμφανιστεί το πρόβλημα της «πεθαίνουσας ReLU», όπου νευρώνες σταματούν να είναι ενεργοί και να εξάγουν μόνο το μηδέν.[16]

Sigmoid: Η συνάρτηση sigmoid αποτυπώνει τις εισόδους σε μια τιμή μεταξύ 0 και 1 και ορίζεται ως $f(x) = 1 / (1 + e^{(-x)})$. Ήταν παλαιότερα πολύ δημοφιλής, ιδιαίτερα για δυαδική ταξινόμηση, καθώς μπορεί να ερμηνεύσει την έξοδο ως πιθανότητα. Έχει ομαλή κλίση, αποτρέποντας απότομες αλλαγές στις τιμές εξόδου[17]. Ωστόσο, είναι επιρρεπής στο πρόβλημα του εξαφανιζόμενου gradient, το οποίο μπορεί να επιβραδύνει την εκπαίδευση ή να δυσκολέψει το δίκτυο να μάθει.

Tanh (Hyperbolic Tangent): Η συνάρτηση Tanh, που είναι παρόμοια με την sigmoid, αποτυπώνει τις τιμές εισόδου σε ένα εύρος μεταξύ -1 και 1, και ορίζεται ως $f(x) = \tanh(x)$. Χρησιμοποιείται συχνά στα κρυφά στρώματα ενός δικτύου καθώς κεντράρει τα δεδομένα, κάνοντας την εκμάθηση πιο ευκολότερη για τα επόμενα στρώματα. Σε σύγκριση με την sigmoid, έχει καλύτερη απόδοση επειδή κεντράρει φυσικά τα δεδομένα. Ωστόσο, παρουσιάζει ακόμα το πρόβλημα του εξαφανιζόμενου gradient, αν και σε μικρότερο βαθμό σε σύγκριση με την συνάρτηση sigmoid.[17]

Leaky ReLU: Η Leaky ReLU είναι μια παραλλαγή της ReLU και επιτρέπει μια μικρή, μη μηδενική κλίση όταν η μονάδα δεν είναι ενεργή. Ορίζεται ως $f(x) = x$ αν $x > 0$, αλλιώς $\alpha * x$ όπου α είναι μια μικρή σταθερά. Χρησιμοποιείται για την αντιμετώπιση του προβλήματος της πεθαίνουσας ReLU. Αποτρέπει τον θάνατο των νευρώνων, κάτι που μπορεί να είναι κρίσιμο σε βαθιά δίκτυα.[16] Ωστόσο, η επιπλέον παράμετρος (α) μπορεί να χρειάζεται ρύθμιση.

Softmax: Η συνάρτηση softmax εξάγει ένα διάνυσμα που αντιπροσωπεύει τις πιθανοτικές κατανομές μιας λίστας πιθανών αποτελεσμάτων. Είναι η εκθετική της δοθείσας εισόδου κανονικοποιημένη από το άθροισμα των εκθετικών τιμών όλων των εισόδων στο διάνυσμα. Συνηθίζεται στο εξωτερικό στρώμα ενός νευρωνικού δικτύου για προβλήματα ταξινόμησης πολλαπλών κλάσεων και είναι χρήσιμη για την εξαγωγή μιας πιθανοτικής κατανομής για πολυκλασικά προβλήματα. Δεν είναι κατάλληλη για κρυφά στρώματα ή για δυαδικά προβλήματα ταξινόμησης.

3.2.2.3 Προηγμένες Συναρτήσεις Ενεργοποίησης

Swish: Είναι μια νεότερη συνάρτηση ενεργοποίησης, ορίζεται ως $f(x) = x / (1 + e^{(-x)})$. Έχει δείξει υποσχόμενα αποτελέσματα σε διάφορα μοντέλα βαθιάς μάθησης και φαίνεται να λειτουργεί καλύτερα από το ReLU σε βαθύτερα μοντέλα, αν και είναι πιο υπολογιστικά δαπανηρή[17].

Mish: Ορίζεται ως $f(x) = x * \tanh(\text{softplus}(x))$ όπου $\text{softplus}(x) = \log(1 + e^x)$. Έχει αποδειχθεί ότι ξεπερνά το ReLU σε πολλά σενάρια και είναι λεία και μη μονότονη, βοηθώντας στην αντιμετώπιση των προβλημάτων των "πεθαίνοντων" νευρώνων[17]. Παρ' όλα αυτά, είναι υπολογιστικά πιο απαιτητική από το ReLU και τις απλούστερες παραλλαγές του.

3.2.2.4 Σημασία των Συναρτήσεων Ενεργοποίησης στα Νευρωνικά Δίκτυα

Οι συναρτήσεις ενεργοποίησης παίζουν καθοριστικό ρόλο στα νευρωνικά δίκτυα, καθώς εισάγουν μη γραμμικότητα. Χωρίς αυτές, ένα νευρωνικό δίκτυο θα ήταν απλά ένα γραμμικό μοντέλο, ανίκανο να μάθει πολύπλοκα μοτίβα. Με την ενσωμάτωση συναρτήσεων

ενεργοποίησης, τα νευρωνικά δίκτυα μπορούν να αποκτήσουν βάθος και να επεξεργάζονται περίπλοκα χαρακτηριστικά[14].

Η επιλογή της σωστής συνάρτησης ενεργοποίησης εξαρτάται από διάφορους παράγοντες, όπως τον τύπο του προβλήματος που αντιμετωπίζει το δίκτυο, τους διαθέσιμους υπολογιστικούς πόρους και την αρχιτεκτονική του δικτύου[18]. Για παράδειγμα, βαθύτερα δίκτυα μπορεί να απαιτούν συναρτήσεις ενεργοποίησης που αντιμετωπίζουν το πρόβλημα της εξαφανιζόμενης κλίσης.

Στο τέλος της διαδικασίας συνέλιξης και ομαδοποίησης, τα νευρωνικά δίκτυα χρησιμοποιούν πλήρως συνδεδεμένα επίπεδα για να ενσωματώσουν τα μαθημένα χαρακτηριστικά στην τελική τους έξοδο, όπως είναι οι βαθμολογίες ταξινόμησης. Συνήθως, σε αυτά τα επίπεδα χρησιμοποιείται η συνάρτηση ενεργοποίησης softmax σε εργασίες ταξινόμησης για να μετατρέψουν την έξοδο σε πιθανοτικές βαθμολογίες.[19]

3.2.3 Προηγμένα Αρχιτεκτονικά Χαρακτηριστικά στα Νευρωνικά Δίκτυα

Στον κόσμο των Συνελκτικών Νευρωνικών Δικτύων (CNNs), η αρχιτεκτονική βαθιών επιπέδων έχει διαδραματίσει σημαντικό ρόλο στην ενίσχυση της μάθησης χαρακτηριστικών. Αρχιτεκτονικές όπως το AlexNet, VGGNet, GoogLeNet και ResNet[20] ανήκουν στις πιο επιδραστικές στον τομέα, με κάθε μία να εισάγει σημαντικές καινοτομίες. Ιδιαίτερα, τα Δίκτυα Υπολοίπων (ResNets) χρησιμοποιούν συνδέσεις παράκαμψης για να πηδήξουν μερικά επίπεδα, αντιμετωπίζοντας το πρόβλημα της εξαφανιζόμενης κλίσης και επιτρέποντας την εκπαίδευση πολύ βαθύτερων δικτύων.

3.2.3.1 Τεχνικές Ρύθμισης για Αντιμετώπιση της Υπερπροσαρμογής

Στα CNNs, οι τεχνικές ρύθμισης είναι κρίσιμες για την πρόληψη της υπερπροσαρμογής, ενισχύοντας την γενίκευση του μοντέλου και βελτιώνοντας την απόδοση σε αδιάκριτα δεδομένα. Η ρύθμιση αναφέρεται σε τεχνικές που μειώνουν την υπερπροσαρμογή, που συμβαίνει όταν ένα μοντέλο μαθαίνει μοτίβα ειδικά για τα δεδομένα εκπαίδευσης αλλά αποτυγχάνει να γενικεύσει σε νέα δεδομένα[21]. Στο πλαίσιο των CNNs, η ρύθμιση είναι ουσιαστική για τη διαχείριση της υψηλής χωρητικότητας του δικτύου, λόγω των βαθιών και πολύπλοκων αρχιτεκτονικών τους.

3.2.3.1.1 Κοινές Τεχνικές Ρύθμισης στα Συνελκτικά Νευρωνικά Δίκτυα (CNNs)

Ρύθμιση L1 και L2: Η ρύθμιση L1 και L2 είναι δύο βασικές τεχνικές που χρησιμοποιούνται στα νευρωνικά δίκτυα. Η ρύθμιση L1 (Lasso) προσθέτει την απόλυτη τιμή των βαρών στη συνάρτηση απώλειας, ενώ η ρύθμιση L2 (Ridge) προσθέτει την τετραγωνική τιμή των βαρών[22]. Αυτές οι ποινές ενθαρρύνουν τα βάρη να είναι μικρά, κάτι που απλοποιεί το μοντέλο και είναι χρήσιμο για την πρόληψη της υπερπροσαρμογής.

Διακοπή (Dropout): Η τεχνική της Διακοπής (Dropout) συνίσταται στην τυχαία απόρριψη νευρώνων κατά τη διάρκεια της εκπαίδευσης, οδηγώντας σε ένα μοντέλο που δεν εξαρτάται

υπερβολικά από κανένα μεμονωμένο νευρώνα.[23] Αυτό βοηθά στην αποτροπή της υπερπροσαρμογής και ενθαρρύνει την ανθεκτικότητα του δικτύου.

Επαύξηση Δεδομένων: Η επαύξηση δεδομένων περιλαμβάνει τη δημιουργία τροποποιημένων εκδοχών των υπαρχόντων δεδομένων εκπαίδευσης για να αυξηθεί η ποικιλομορφία τους. Αυτό μπορεί να περιλαμβάνει τεχνικές όπως αντιστροφή, περιστροφή, κλιμάκωση και την παραλλαγή χρωμάτων. Είναι ιδιαίτερα χρήσιμη σε εργασίες επεξεργασίας εικόνων και βοηθά στη βελτίωση της γενίκευσης του μοντέλου.

Πρόωρη Διακοπή στα Συνελκτικά Νευρωνικά Δίκτυα (CNNs): Η πρόωρη διακοπή αποτελεί μια τεχνική για την πρόληψη της υπερπροσαρμογής μέσω της διακοπής της διαδικασίας εκπαίδευσης πριν το μοντέλο φτάσει σε ένα σημείο μειωμένης απόδοσης. Αυτό γίνεται παρακολουθώντας την απόδοση του μοντέλου σε ένα σύνολο επικύρωσης και είναι μια δημοφιλής πρακτική σε σενάρια όπου η παρατεταμένη εκπαίδευση μπορεί να οδηγήσει σε υπερπροσαρμογή.[24] Η τεχνική αυτή είναι απλή στην υλοποίηση και δεν απαιτεί τροποποίηση του υποκείμενου μοντέλου, αλλά απαιτεί προσεκτική επιλογή κριτηρίων διακοπής για να αποφευχθεί η υποπροσαρμογή.

Κανονικοποίηση Δέσμης: Η κανονικοποίηση δέσμης είναι μια τεχνική που βελτιώνει τη σταθερότητα της διαδικασίας εκπαίδευσης μέσω της ρύθμισης και της κλιμάκωσης των ενεργοποιήσεων σε κάθε επίπεδο εισόδου. Αυτή η τεχνική είναι σημαντική για την εξασφάλιση ότι η κατανομή των εισόδων σε ένα επίπεδο δεν αλλάζει σημαντικά με τον καιρό, οδηγώντας σε ταχύτερη σύγκλιση και σταθεροποίηση της διαδικασίας εκπαίδευσης[24]. Παρόλο που είναι μια χρήσιμη τεχνική, μπορεί να είναι υπολογιστικά δαπανηρή και να μη λειτουργεί καλά με ορισμένα μεγέθη δέσμης.

Εισαγωγή Θορύβου: Η εισαγωγή θορύβου αποτελεί μια τεχνική ρύθμισης που περιλαμβάνει την προσθήκη θορύβου στις εισόδους, τα βάρη, ή ακόμα και στις εξόδους κατά τη διάρκεια της εκπαίδευσης. Αυτή η πρακτική μπορεί να λειτουργήσει ως μορφή ρύθμισης, ενθαρρύνοντας το μοντέλο να μάθει πιο ανθεκτικά χαρακτηριστικά. Παρόλο που μπορεί να βελτιώσει την ανθεκτικότητα του μοντέλου, αν δεν ελεγχθεί προσεκτικά, η εισαγωγή θορύβου μπορεί να διαταράξει τη διαδικασία μάθησης.

3.2.4 Εκπαίδευση και Στρατηγικές Βελτιστοποίησης στα Συνελκτικά Νευρωνικά Δίκτυα (CNNs)

3.2.4.1 Αντίστροφη Διάδοση: Η Ραχοκοκαλιά της Μάθησης

Η αντίστροφη διάδοση είναι μια θεμελιώδης μέθοδος στην εκπαίδευση των νευρωνικών δικτύων, και ιδιαίτερα στα CNNs. Αποτελεί τον βασικό μηχανισμό για την ρύθμιση των βαρών του δικτύου, με στόχο την ελαχιστοποίηση της συνάρτησης απώλειας[25]. Ας εξερευνήσουμε περαιτέρω τη λειτουργικότητα, τις πολυπλοκότητες, τις εφαρμογές και άλλες σημαντικές πτυχές της αντίστροφης διάδοσης.

3.2.4.1.1 Επισκόπηση της Αντίστροφης Διάδοσης

Η αντίστροφη διάδοση, ή αντίστροφη διάδοση λαθών, είναι ένας αλγόριθμος που χρησιμοποιείται για τον υπολογισμό της κλίσης της συνάρτησης απώλειας σε σχέση με τα

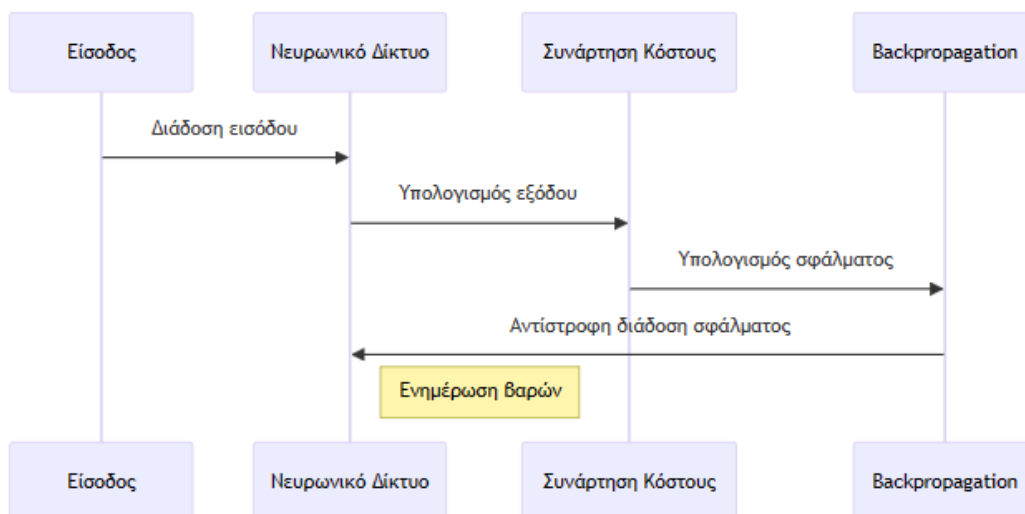
βάρη του δικτύου. Αυτή η διαδικασία είναι ζωτικής σημασίας για τη μάθηση στα νευρωνικά δίκτυα, καθώς παρέχει μια δομημένη μέθοδο για την ενημέρωση των βαρών στο δίκτυο, με βάση την απόκλιση της πραγματικής εξόδου από την επιθυμητή έξοδο. Η διαδικασία αυτή περιλαμβάνει δύο βασικά στάδια: την προώθηση και την αντίστροφη προώθηση.

Στην προώθηση, τα δεδομένα εισόδου περνούν μέσα από το δίκτυο, ενεργοποιώντας τα διάφορα στρώματα και παράγοντας προβλέψεις εξόδου. Αυτό το στάδιο είναι απαραίτητο για να καθοριστεί η αρχική απόδοση του μοντέλου πριν από την εφαρμογή της αντίστροφης διάδοσης.

Η συνάρτηση απώλειας, στη συνέχεια, μετρά τη διαφορά μεταξύ των πραγματικών τιμών και των προβλέψεων του δικτύου. Αυτή η διαφορά είναι ουσιαστικά ένα μέτρο του λάθους που διαπράττει το δίκτυο και είναι η βάση για την επόμενη φάση της αντίστροφης διάδοσης.

Στην αντίστροφη προώθηση, το λάθος διαδίδεται πίσω στο δίκτυο, υπολογίζοντας την κλίση της συνάρτησης απώλειας σε σχέση με κάθε βάρος. Αυτός ο υπολογισμός δίνει έναν δείκτη του πώς πρέπει να τροποποιηθούν τα βάρη για να μειωθεί το συνολικό σφάλμα του δικτύου.

Τέλος, η ενημέρωση των βαρών γίνεται με βάση τον ρυθμό μάθησης, μια κρίσιμη υπερπαραμέτρο που διαμορφώνει το πόσο μεγάλη θα είναι η προσαρμογή των βαρών σε κάθε επανάληψη της διαδικασίας εκπαίδευσης.



Εικόνα 3.2-2: Η διαδικασία της αντίστροφης διάδοσης

3.2.4.1.2 Σημασία της Αντίστροφης Διάδοσης στα CNNs

Στα CNNs, η αντίστροφη διάδοση είναι κρίσιμη για την αποτελεσματική εκμάθηση σύνθετων χαρακτηριστικών από τα δεδομένα. Μέσω αυτής της διαδικασίας, το δίκτυο μαθαίνει να διορθώνει τα βάρη του, βελτιώνοντας σταδιακά την απόδοσή του σε εργασίες όπως η αναγνώριση εικόνων και η κατηγοριοποίηση[25], [26].

3.2.4.1.3 Προκλήσεις στην Αντίστροφη Διάδοση

- **Πρόβλημα Εξαφανιζόμενης Κλίσης:** Σε βαθιά δίκτυα, οι κλίσεις μπορεί να γίνουν πολύ μικρές καθώς διαδίδονται προς τα πίσω, καθιστώντας την εκπαίδευση αργή ή ακόμη και αδύνατη.
- **Πρόβλημα Εκρηκτικής Κλίσης:** Αντίθετα, οι κλίσεις μπορεί να αυξηθούν εκθετικά, οδηγώντας σε ασταθή δίκτυα.

3.2.5 Βελτιστοποιητές: Ενίσχυση της Σύγκλισης στα Συνελκτικά Νευρωνικά Δίκτυα

Η επιλογή του σωστού βελτιστοποιητή είναι κρίσιμη στην εκπαίδευση των Συνελκτικών Νευρωνικών Δικτύων (CNNs), καθώς επηρεάζει την ταχύτητα και την αποδοτικότητα της σύγκλισης προς το ελάχιστο της συνάρτησης απώλειας.

3.2.5.1 Προηγμένοι Βελτιστοποιητές

Βελτιστοποιητές όπως ο Adam (Adaptive Moment Estimation), ο RMSprop (Root Mean Square Propagation) και ο Adagrad (Adaptive Gradient Algorithm) αποτελούν προηγμένες επιλογές στην βελτιστοποίηση νευρωνικών δικτύων[29]. Αυτοί οι βελτιστοποιητές προσαρμόζουν δυναμικά τον ρυθμό μάθησης κατά τη διάρκεια της εκπαίδευσης, βελτιώνοντας τους ρυθμούς σύγκλισης και προσφέροντας μεγαλύτερη σταθερότητα στην εκπαίδευση.

3.2.6 Συναρτήσεις Απώλειας: Οδηγώντας τη Διαδικασία Εκπαίδευσης

Στα CNNs, οι συναρτήσεις απώλειας, ή συναρτήσεις κόστους, καθορίζουν πόσο αποτελεσματικά το μοντέλο προβλέπει τις επιθυμητές εξόδους. Αυτές οι συναρτήσεις ποσοτικοποιούν τη διαφορά μεταξύ των προβλέψεων του μοντέλου και των πραγματικών στοχευόμενων τιμών και είναι θεμελιώδεις στην οδήγηση της διαδικασίας βελτιστοποίησης μέσω της αντίστροφης διάδοσης.

3.2.6.1 Επισκόπηση Συναρτήσεων Απώλειας στα Συνελκτικά Νευρωνικά Δίκτυα

Στην εκπαίδευση των Συνελκτικών Νευρωνικών Δικτύων, οι συναρτήσεις απώλειας διαδραματίζουν έναν πρωταγωνιστικό ρόλο, καθώς προσδιορίζουν το πώς το δίκτυο αξιολογεί τη διαφορά μεταξύ των προβλεπόμενων και των πραγματικών τιμών. Η επιλογή της συνάρτησης απώλειας ποικίλλει ανάλογα με τον τύπο της εργασίας - είτε πρόκειται για παλινδρόμηση, είτε για ταξινόμηση - και μπορεί να έχει σημαντική επίδραση στην απόδοση και την ταχύτητα εκπαίδευσης του μοντέλου.

Στις εργασίες παλινδρόμησης, το Μέσο Τετραγωνικό Σφάλμα (MSE) είναι μια δημοφιλής επιλογή. Αυτή η συνάρτηση υπολογίζει τον μέσο όρο των τετραγώνων των διαφορών μεταξύ των προβλεπόμενων και των πραγματικών τιμών. Η δύναμη της συνάρτησης MSE έγκειται στην απλότητά της και στην ικανότητά της να τιμωρεί περισσότερο τα μεγαλύτερα σφάλματα. Ωστόσο, αυτή η ίδια ιδιότητα μπορεί να την καταστήσει ευαίσθητη σε ακραίες τιμές.

Στον τομέα της ταξινόμησης, η Απώλεια Διασταυρωμένης Εντροπίας είναι ευρέως χρησιμοποιούμενη. Αυτή η συνάρτηση είναι αποτελεσματική στη μέτρηση της διαφοράς μεταξύ των προβλεπόμενων πιθανοτήτων εξόδου και των πραγματικών στοχευόμενων τιμών. Η διασταυρωμένη εντροπία είναι ιδιαίτερα χρήσιμη σε προβλήματα με δύο ή περισσότερες τάξεις και επιτρέπει την αποτελεσματική εκτίμηση της απόδοσης των μοντέλων ταξινόμησης. Ωστόσο, αυτή η συνάρτηση απαιτεί προσοχή κατά την υλοποίηση, καθώς η αριθμητική αστάθεια (π.χ., λογάριθμος του μηδενός) μπορεί να προκαλέσει προβλήματα.

Δυαδική Διασταυρωμένη Εντροπία: Η Δυαδική Διασταυρωμένη Εντροπία είναι μια ειδική περίπτωση της διασταυρωμένης εντροπίας που εφαρμόζεται σε δυαδικά προβλήματα ταξινόμησης[30]. Είναι ιδανική για καταστάσεις όπου τα αποτελέσματα μοντελοποιούνται ως πιθανότητες για δύο κατηγορίες. Αυτή η συνάρτηση είναι κατάλληλη για μοντέλα που αξιολογούν πιθανότητες σε δυαδικό πλαίσιο, αλλά δεν είναι κατάλληλη για πολυκλασικά προβλήματα ταξινόμησης.

Κατηγορική Διασταυρωμένη Εντροπία: Από την άλλη πλευρά, η Κατηγορική Διασταυρωμένη Εντροπία είναι ιδανική για πολυκλασικές ταξινομήσεις, όπου κάθε δείγμα ανήκει σε μία μόνο από πολλές τάξεις[30]. Αυτή η συνάρτηση απαιτεί τις ετικέτες στόχου να βρίσκονται σε μορφή one-hot κωδικοποίησης, κάτι που μπορεί να προσθέσει μια επιπλέον πολυπλοκότητα στην προετοιμασία των δεδομένων.

Απώλεια Hinge: Η Απώλεια Hinge, από την άλλη πλευρά, συνδέεται στενά με τα μηχανήματα διανυσματικής υποστήριξης (SVMs) και χρησιμοποιείται κυρίως σε εργασίες δυαδικής ταξινόμησης. Η λειτουργικότητά της εστιάζει στη διατήρηση ενός περιθωρίου μεταξύ των κατηγοριών και ενθαρρύνει το μοντέλο να είναι σίγουρο για τις προβλέψεις του[31]. Παρόλα αυτά, η εφαρμογή της περιορίζεται κυρίως σε δυαδικά προβλήματα ταξινόμησης και δεν είναι τόσο προτιμητέα για πολυκλασικές καταστάσεις.

3.2.6.2 Προηγμένες Συναρτήσεις Απώλειας στα Νευρωνικά Δίκτυα

Οι προηγμένες συναρτήσεις απώλειας αναπτύσσονται για να ανταποκρίνονται σε ειδικά προβλήματα και προκλήσεις που συναντώνται στην εκπαίδευση νευρωνικών δικτύων. Αυτές οι συναρτήσεις ενσωματώνουν πιο πολύπλοκες δομές και αρχές για να βελτιστοποιήσουν την απόδοση του μοντέλου σε συγκεκριμένες συνθήκες.

Απώλεια Huber: Η Απώλεια Huber συνδυάζει στοιχεία του Μέσου Τετραγωνικού Σφάλματος και του Μέσου Απόλυτου Σφάλματος, προσφέροντας έναν συμβιβασμό μεταξύ των δύο. Είναι χρήσιμη σε εργασίες παλινδρόμησης, όπου τα ακραία σφάλματα μπορούν να επηρεάσουν δυσανάλογα το συνολικό σφάλμα.[32] Η λιγότερη ευαισθησία της σε ακραίες τιμές σε σύγκριση με το MSE την καθιστά μια προτιμητέα επιλογή σε πολλές περιπτώσεις.

Απώλεια Εστίασης: Η Απώλεια Εστίασης εστιάζει στην αντιμετώπιση της ανισοροπίας τάξεων, ιδιαίτερα σημαντική σε εργασίες ανίχνευσης αντικειμένων. Με την επικέντρωση σε δύσκολα προς ταξινόμηση παραδείγματα, βοηθά το μοντέλο να βελτιώσει την απόδοσή του σε στοιχεία που είναι πιο πιθανό να καταλήξουν σε λανθασμένη ταξινόμηση, αυξάνοντας την αποτελεσματικότητά του σε πιο προκλητικά δεδομένα.[33]

Απώλεια Dice: Η Απώλεια Dice είναι ιδιαίτερα χρήσιμη σε εργασίες τμηματοποίησης, όπως στην τμηματοποίηση ιατρικών εικόνων. Εστιάζει στη μέτρηση της επικάλυψης μεταξύ της προβλεπόμενης τμηματοποίησης και της πραγματικής, προσφέροντας μια διαισθητική μετρική για την αξιολόγηση της ακρίβειας της τμηματοποίησης. Η καλή προσαρμογή της σε προβλήματα τμηματοποίησης την καθιστά μια πολύτιμη εργαλειοθήκη σε αυτό τον τομέα.[34]

3.2.6.3 Πέρα από τις Οπτικές Εφαρμογές

Υπολογιστική Όραση: Ο Πρωταρχικός Τομέας

Τα Συνελκτικά Νευρωνικά Δίκτυα (CNNs) έχουν αποδειχθεί ιδιαίτερα αποτελεσματικά στην υπολογιστική όραση, από την ταξινόμηση εικόνων έως την ανίχνευση και την τμηματοποίηση αντικειμένων.



Εικόνα 3.2-3: AI Generated

Επεξεργασία Φυσικής Γλώσσας (NLP)

Αν και τα CNNs είναι λιγότερο διαδεδομένα στο NLP σε σύγκριση με τα RNNs και τα Transformers, έχουν βρει εφαρμογή σε εργασίες όπως η ταξινόμηση προτάσεων και η μοντελοποίηση θεμάτων.



Εικόνα 3.2-4: AI Generated (DALL-E)

Επεξεργασία και Ανάλυση Ήχου

Σε αυτό τον τομέα, τα CNNs μπορούν να επεξεργαστούν και να ταξινομήσουν ηχητικά σήματα, όπως σε συστήματα αναγνώρισης ομιλίας, προσφέροντας μοναδικές δυνατότητες στην επεξεργασία ήχου.



Εικόνα 3.2-5: AI Generated (DALL-E)

3.2.7 Τρέχουσες Έρευνες και Μελλοντικές Τάσεις στις Αρχιτεκτονικές CNN

Καινοτομίες στις Αρχιτεκτονικές CNN

Οι νέες αρχιτεκτονικές CNN ενσωματώνουν μηχανισμούς προσοχής και transformers για βελτιωμένη εστίαση σε σημαντικά μέρη των εισόδων. Δίκτυα Capsule εξερευνούν πιο προηγμένες αναπαραστάσεις χωρικών ιεραρχιών για αυξημένη ακρίβεια.

Υπολογιστική Αποδοτικότητα και Βιωσιμότητα

Η έρευνα επικεντρώνεται στη συμπίεση μοντέλων και κλάδεμα για μείωση του μεγέθους των CNNs, καθιστώντας τα πιο αποδοτικά για φορητές συσκευές. Επιπλέον, η ενεργειακή αποδοτικότητα αποτελεί προτεραιότητα στο πλαίσιο των περιβαλλοντικών ανησυχιών.

Ηθικές και Κοινωνικές Συνέπειες

Η αντιμετώπιση προκαταλήψεων και η διασφάλιση της δικαιοσύνης στα μοντέλα TN είναι ζωτικής σημασίας. Επιπρόσθετα, η εξηγήσιμη Τεχνητή Νοημοσύνη γίνεται όλο και πιο σημαντική για την αύξηση της διαφάνειας και της αξιοπιστίας των συστημάτων.

Τα CNNs συνεχίζουν να εξελίσσονται, προσφέροντας σημαντικές εφαρμογές στην Τεχνητή Νοημοσύνη. Οι μελλοντικές εξελίξεις θα διαμορφώσουν την πορεία τους, λαμβάνοντας υπόψη τόσο τεχνικές προκλήσεις όσο και κοινωνικο-ηθικές ανησυχίες.

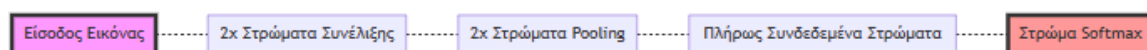
3.3 Σημαντικές Αρχιτεκτονικές Νευρωνικών Δικτύων

3.3.1 Το δίκτυο LeNet

Το LeNet θεωρείται ένα από τα πρωτοποριακά μοντέλα στον τομέα των συνελκτικών νευρωνικών δικτύων (CNNs) και έχει συνεισφέρει σημαντικά στην ανάπτυξη της βαθιάς μάθησης.[35] Το LeNet, ως ένα από τα πρώτα συνελκτικά νευρωνικά δίκτυα, έχει συμβάλει σημαντικά στην εξέλιξη της βαθιάς μάθησης και των νευρωνικών δικτύων. Αναπτύχθηκε από τον Yann LeCun και άλλους στα τέλη της δεκαετίας του 1980 και αρχές της δεκαετίας του 1990, με κύριο σκοπό την αναγνώριση αριθμών και χαρακτήρων, όπως τους αριθμούς επιταγών τραπεζών και τους ταχυδρομικούς κώδικες.

3.3.1.1 Πρωτοποριακή Αρχιτεκτονική

Η αρχιτεκτονική του LeNet, με την εναλλαγή μεταξύ συνελκτικών και υποδειγματοληπτικών επιπέδων, θεμελίωσε τις βάσεις για τα σύγχρονα CNNs. Εισήγαγε την ιδέα της τοπικής χωρικής συνέλιξης και της σταδιακής μείωσης του χωρικού μεγέθους των χαρακτηριστικών.[35] Η αρχιτεκτονική του LeNet, γνωστή ως LeNet-5, αποτελείται από επτά επίπεδα εκμάθησης (εξαιρουμένου του εισαγωγικού επιπέδου). Αυτά τα επίπεδα περιλαμβάνουν συνελκτικά επίπεδα, επίπεδα υποδειγματοληψίας ή μέσης συγκέντρωσης και πλήρως συνδεδεμένα επίπεδα. Αυτή η δομή επέτρεψε στο LeNet να αναγνωρίζει αποτελεσματικά χωρικές ιεραρχίες και μοτίβα στα δεδομένα, ενώ η χρήση των κοινόχρηστων βαρών στα επίπεδα συνέλιξης έκανε το δίκτυο πιο αποδοτικό στη χρήση των παραμέτρων του.



Εικόνα 3.3-1: Αρχιτεκτονική LeNet

3.3.1.2 Εφαρμογή σε Αναγνώριση Χαρακτήρων

Το LeNet βρήκε εκτεταμένη εφαρμογή στην αναγνώριση χειρόγραφων και μηχανογραφημένων χαρακτήρων, ιδιαίτερα σε τραπεζικά και ταχυδρομικά συστήματα. Η ικανότητά του να αναγνωρίζει ποικίλους χαρακτήρες με υψηλή ακρίβεια αποδείχθηκε πολύτιμη για αυτοματοποιημένες διαδικασίες.

3.3.1.3 Επίδραση στην Έρευνα και Ανάπτυξη

Το LeNet ενέπνευσε μια νέα γενιά ερευνητών και αναπτυξιακών προσπαθειών στον τομέα της τεχνητής νοημοσύνης και της υπολογιστικής όρασης.[36] Συνέβαλε στην αναγνώριση της σημασίας των βαθιών νευρωνικών δικτύων και της αποτελεσματικότητας των συνελκτικών δομών.[37]

3.3.2 Το δίκτυο AlexNet

Το AlexNet, μία κομβική αρχιτεκτονική στον τομέα των Συνελικτικών Νευρωνικών Δικτύων (CNNs), έχει παίξει καθοριστικό ρόλο στην ανάπτυξη της βαθιάς μάθησης, ειδικά στην υπολογιστική όραση.

3.3.2.1 Εισαγωγή στο AlexNet

Το AlexNet αναπτύχθηκε από τον Alex Krizhevsky, τον Ilya Sutskever, και τον Geoffrey Hinton και αναδείχθηκε στον διαγωνισμό ImageNet Large Scale Visual Recognition Challenge (ILSVRC) το 2012, σημειώνοντας μεγάλη επιτυχία με τη σημαντική μείωση του ποσοστού σφάλματος top-5 σε σύγκριση με τα προηγούμενα μοντέλα[38].

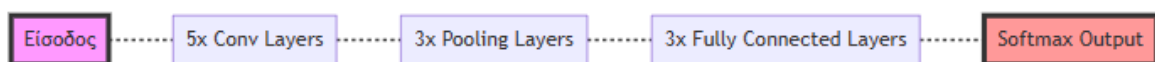
3.3.2.2 Αρχιτεκτονική του AlexNet

Αποτελείται από οκτώ επίπεδα εκμάθησης, με τα πέντε να είναι συνελικτικά και τα τρία να είναι πλήρως συνδεδεμένα. Η αρχιτεκτονική του χρησιμοποιεί τη συνάρτηση ενεργοποίησης Rectified Linear Unit (ReLU), η οποία συνέβαλε στην ταχύτερη εκπαίδευση σε σύγκριση με προηγούμενες συναρτήσεις όπως η tanh.[38]

3.3.2.3 Κύρια Χαρακτηριστικά του AlexNet

Το πρώτο συνελικτικό επίπεδο χρησιμοποιεί 96 πυρήνες διαστάσεων 11x11x3 με βήμα 4, ενώ το δεύτερο χρησιμοποιεί 256 πυρήνες διαστάσεων 5x5x48.[38] Το max pooling εφαρμόζεται μετά από ορισμένα συνελικτικά επίπεδα για τη μείωση της χωρικής διάστασης. Η Τοπική Απόκριση Κανονικοποίησης (LRN) ενισχύει τη γενίκευση μετά τα πρώτα δύο συνελικτικά επίπεδα. Τα τρία πλήρως συνδεδεμένα επίπεδα έχουν 4096, 4096, και 1000 κόμβους, με το τελευταίο επίπεδο να χρησιμοποιεί softmax για ταξινόμηση. Το dropout εφαρμόζεται στα πρώτα δύο πλήρως συνδεδεμένα επίπεδα για την αποφυγή υπερπροσαρμογής.[38]

Το AlexNet αποτελεί πρωτοπόρο στον τομέα των CNNs, παρουσιάζοντας ουσιαστικές προσεγγίσεις όπως η χρήση της ReLU, το max pooling, και το dropout, οι οποίες εξακολουθούν να αποτελούν βασικά στοιχεία στην αρχιτεκτονική των σύγχρονων νευρωνικών δικτύων.



Εικόνα 3.3-2: Αρχιτεκτονική AlexNet

3.3.2.4 Επίδραση του AlexNet στη Βαθιά Μάθηση

Το AlexNet, με τις καινοτομίες του στην αρχιτεκτονική των CNNs, έχει προκαλέσει σημαντική επιρροή στον τομέα της βαθιάς μάθησης.

Χρήση GPU

Η χρήση των Γραφικών Επεξεργαστικών Μονάδων (GPUs) για την επιτάχυνση της διαδικασίας εκπαίδευσης ήταν μία από τις κύριες καινοτομίες του AlexNet. Το δίκτυο εκπαιδεύτηκε σε δύο NVIDIA GTX 580 GPUs για περίπου πέντε έως έξι ημέρες, αποδεικνύοντας την αποδοτικότητα της χρήσης GPUs στην εκπαίδευση βαθιών μοντέλων[38].

Επικαλυπτόμενη Συγκέντρωση

Το AlexNet εισήγαγε την επικαλυπτόμενη συγκέντρωση, μία τεχνική που μειώνει το μέγεθος του δικτύου και βοηθά στην αποφυγή υπερπροσαρμογής, αυξάνοντας την απόδοση και τη γενίκευση.[38]

Επανάσταση στην Επίδοση

Η επιτυχία του AlexNet στον διαγωνισμό ImageNet επιβεβαίωσε το δυναμικό των CNNs σε εργασίες αναγνώρισης εικόνων, αναζωπυρώνοντας το ενδιαφέρον για τη βαθιά μάθηση και ανοίγοντας νέους δρόμους έρευνας.

Κληρονομιά και Σημασία

Το AlexNet έχει εμπνεύσει πληθώρα επόμενων μοντέλων και προσεγγίσεων στη βαθιά μάθηση, μεταβάλλοντας την πορεία της έρευνας στην τεχνητή νοημοσύνη και στην υπολογιστική όραση.

Η ανάδυση του AlexNet σηματοδότησε μια εποχή μετασχηματισμού στη βαθιά μάθηση. Με την πρωτοποριακή χρήση της ReLU, του dropout, και των GPUs, το AlexNet έθεσε τις βάσεις για την εκτεταμένη ανάπτυξη και εφαρμογή των CNNs σε διάφορους τομείς, ενισχύοντας συνεχώς τις δυνατότητες της τεχνητής νοημοσύνης.

3.3.3 Το δίκτυο U-NET

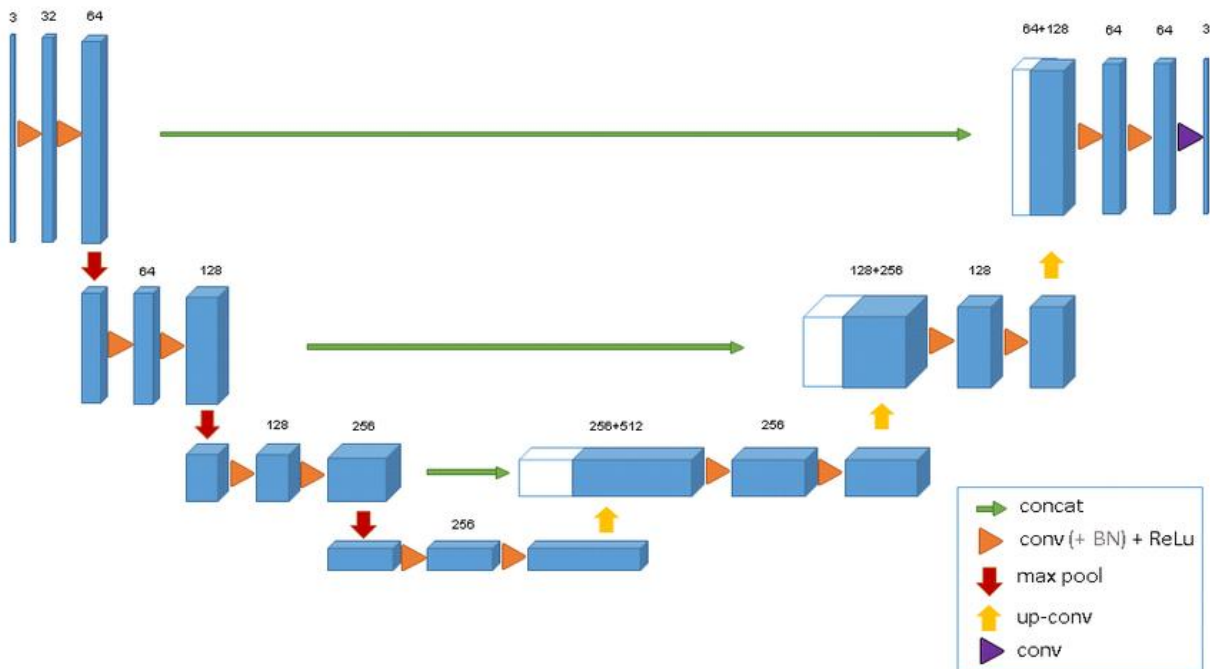
Τα U-Nets αποτελούν έναν τύπο αρχιτεκτονικής νευρωνικών δικτύων συνέλιξης (CNN), τα οποία χρησιμοποιούνται κυρίως για εργασίες τμηματοποίησης εικόνας, ιδίως στον τομέα της βιοϊατρικής τμηματοποίησης εικόνας. Είναι γνωστά για το χαρακτηριστικό τους σχήμα "U", το οποίο τους δίνει και το όνομά τους. Παρακάτω ακολουθεί μια εκτεταμένη επισκόπηση των U-Nets και του τρόπου δημιουργίας και εκπαίδευσής τους χρησιμοποιώντας το PyTorch.

3.3.3.1 Επισκόπηση των U-Nets

Βασική Δομή

- Συμμετρική Αρχιτεκτονική: Η αρχιτεκτονική U-Net αποτελείται από δύο διαδρομές: μία συσταλτική (downsampling) διαδρομή και μία διευρυντική (upsampling) διαδρομή, σχηματίζοντας μια δομή σε σχήμα U.
- Συσταλτική Διαδρομή: Η διαδρομή μείωσης δειγματοληψίας είναι ένα τυπικό συνελκτικό δίκτυο που περιλαμβάνει επαναλαμβανόμενη εφαρμογή συνελίξεων, ακολουθούμενη από συναρτήσεις ενεργοποίησης rectified linear unit (ReLU) και λειτουργίες max pooling. Αυτό το μέρος του δικτύου αποτυπώνει το πλαίσιο στην εικόνα.
- Διευρυντική Διαδρομή: Η διαδρομή αύξησης δειγματοληψίας περιλαμβάνει την ανάπτυξη του χάρτη χαρακτηριστικών ακολουθούμενη από συνελκτικές λειτουργίες.

Αυτή η διαδρομή στοχεύει στην επίτευξη ακριβούς τοπικοποίησης χρησιμοποιώντας αντίστροφες συνελίξεις.[39]



Εικόνα 3.3-3: Βασική αρχιτεκτονική U-net [69]

Συνδέσεις Παράκαμψης

- **Λειτουργικότητα:** Οι συνδέσεις παράκαμψης αποτελούν ένα κρίσιμο χαρακτηριστικό της αρχιτεκτονικής U-Net. Περιλαμβάνουν τη μεταφορά χαρτών χαρακτηριστικών από τη συσταλτική διαδρομή στη διευρυντική, συνδυάζοντας σημασιολογικές πληροφορίες από βαθιά, χονδρικά επίπεδα με πληροφορίες εμφάνισης από ρηχά, λεπτά επίπεδα.
- **Σκοπός:** Αυτές οι συνδέσεις βοηθούν το δίκτυο να ανακτήσει το χωρικό πλαίσιο που χάνεται κατά τις λειτουργίες μείωσης δειγματοληψίας[39]

Εφαρμογές

- **Τμηματοποίηση Βιοϊατρικής Εικόνας:** Τα U-Nets χρησιμοποιούνται ευρέως για εργασίες όπως η τμηματοποίηση κυττάρων και η ανίχνευση όγκων στην ιατρική απεικόνιση[40]
- **Γενικές Εργασίες Τμηματοποίησης Εικόνας:** Η αποτελεσματικότητά τους στις εργασίες τμηματοποίησης εκτείνεται πέρα από την ιατρική απεικόνιση σε τομείς όπως η ανάλυση δορυφορικών εικόνων και η τμηματοποίηση δρόμων για αυτόνομη οδήγηση.

3.3.3.2 Αρχιτεκτονικές U-net

1. Βασικό U-net

Όπως αναφέρθηκε νωρίτερα, το δίκτυο U-net μπορεί να χωριστεί σε δύο μέρη: Το πρώτο είναι ο συμπίεστικός δρόμος που χρησιμοποιεί μια τυπική αρχιτεκτονική CNN. Κάθε μπλοκ στον συμπίεστικό δρόμο αποτελείται από δύο συνεχόμενες συνελίξεις 3×3 ακολουθούμενες από

μονάδα ενεργοποίησης ReLU και ένα στρώμα max-pooling. Αυτή η διάταξη επαναλαμβάνεται πολλές φορές.[39] Η καινοτομία του U-net έρχεται στο δεύτερο μέρος, το λεγόμενο εκτεταμένο δρόμο, στο οποίο κάθε στάδιο αυξάνει την ανάλυση του χάρτη χαρακτηριστικών χρησιμοποιώντας επικονίαση 2×2 . Στη συνέχεια, ο χάρτης χαρακτηριστικών από το αντίστοιχο στρώμα στον συμπιεστικό δρόμο κόβεται και συγκολλάται πάνω στον χάρτη χαρακτηριστικών με αυξημένη ανάλυση. Αυτό ακολουθείται από δύο συνεχόμενες συνελίξεις 3×3 και ενεργοποίηση ReLU. Στο τελικό στάδιο, εφαρμόζεται μια επιπλέον συνέλιξη 1×1 για να μειωθεί ο χάρτης χαρακτηριστικών στον απαιτούμενο αριθμό καναλιών και να παραχθεί η τμηματοποιημένη εικόνα. [39]Το κόψιμο είναι απαραίτητο καθώς τα χαρακτηριστικά των εικονοστοιχείων στις άκρες έχουν την λιγότερη ποσότητα πληροφοριακού περιεχομένου και επομένως πρέπει να απορριφθούν. Αυτό οδηγεί σε ένα δίκτυο που μοιάζει με το σχήμα του γράμματος u και, πιο σημαντικό, διαδίδει πληροφοριακό περιεχόμενο κατά μήκος του δικτύου, που του επιτρέπει να τμηματοποιεί αντικείμενα σε μια περιοχή χρησιμοποιώντας πληροφορίες από μια μεγαλύτερη επικαλυπτόμενη περιοχή.

Η ενέργεια του δικτύου δίνεται από τον τύπο[39]:

$$E = \sum w(x) \log(\rho_k(x))$$

Equation 3.3-1: Ενέργεια δικτύου U-net

όπου ρ_k είναι η συνάρτηση SoftMax ανά εικονοστοιχείο που εφαρμόζεται πάνω στον τελικό χάρτη χαρακτηριστικών, ορισμένη ως:

$$\rho_k = \exp(a_k(x)) / \sum_{k'=1}^K \exp(a_{k'}(x))$$

Equation 3.3-2: Συνάρτηση Softmax

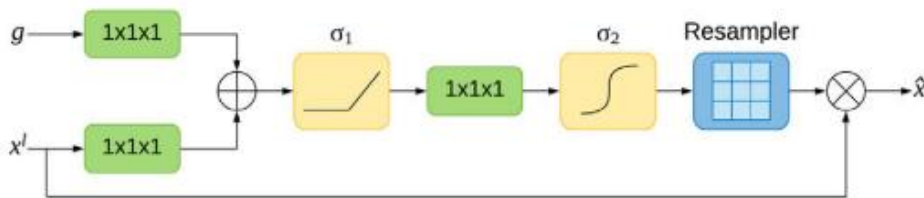
2. 3D U-NET

Το 3D U-Net είναι μια επέκταση του βασικού πλαισίου U-Net που επιτρέπει την τρισδιάστατη τομογραφική τομή [3]. Η κεντρική δομή παραμένει αποτελείται από ένα μονοπάτι σύσφιξης και εκτάσεως. Ωστόσο, όλες οι 2D λειτουργίες αντικαθίστανται με αντίστοιχες 3D λειτουργίες, δηλαδή 3D συνελίξεις, 3D μέγιστης συγκέντρωσης και 3D αντίστροφων συνελίξεων, με αποτέλεσμα μια τρισδιάστατη τμηματοποιημένη εικόνα. Αυτό το δίκτυο είναι σε θέση να τμηματοποιεί εικόνες χρησιμοποιώντας ελάχιστα σχολιασμένα παραδείγματα. Αυτό οφείλεται στο γεγονός ότι οι 3D εικόνες έχουν πολλές επαναλαμβανόμενες δομές και σχήματα, επιτρέποντας έτσι ένα ταχύτερο διαδικασία εκπαίδευσης ακόμη και με δεδομένα που είναι ελάχιστα επισημασμένα[39]. Το 3D U-net έχει χρησιμοποιηθεί εκτενώς σε εφαρμογές τμηματοποίησης εικόνων όγκου CT και MR, συμπεριλαμβανομένων της διάγνωσης των καρδιακών δομών, των οστικών δομών, της σπονδυλικής στήλης, των όγκων του εγκεφάλου, των όγκων του ήπατος, των πνευμονικών λεκέδων, του καρκίνου του νασοφάρυγγα, της τμηματοποίησης πολλαπλών οργάνων, της αξιολόγησης οργάνων κινδύνου στο κεφάλι και τον λαιμό, και της τμηματοποίησης λευκών ουσιών. Το 3D U-net έχει χρησιμοποιηθεί με μεγάλη επιτυχία σε πολλές βιοϊατρικές εφαρμογές. Ο Zeng και οι συνεργάτες του δημιούργησαν ένα δίκτυο που παρήγαγε εικόνες τμηματοποίησης πολλαπλών επιπέδων που επιτρέπουν την αφαίρεση κατά τη διάγνωση[39].

3. Attention U-net

Ένα συχνά επιθυμητό χαρακτηριστικό σε ένα δίκτυο επεξεργασίας εικόνων είναι η ικανότητα εστίασης σε συγκεκριμένα αντικείμενα που είναι σημαντικά ενώ παράλληλα αγνοούνται οι ανεπιθύμητες περιοχές. Το attention U-net επιτυγχάνει αυτό κάνοντας χρήση της πύλης προσοχής (attention gate). Μια πύλη προσοχής είναι μια μονάδα που περικλύπει χαρακτηριστικά που δεν είναι σχετικά με το εκάστοτε έργο. Κάθε στρώμα στον εκτεταμένο δρόμο έχει μια πύλη προσοχής μέσω της οποίας τα αντίστοιχα χαρακτηριστικά από τον συμπίεστικό δρόμο πρέπει να περάσουν πριν τα χαρακτηριστικά συγχωνευτούν με τα χαρακτηριστικά με αυξημένη ανάλυση στον εκτεταμένο δρόμο. Οι επαναλαμβανόμενες χρήσεις της πύλης προσοχής μετά από κάθε στρώμα βελτιώνουν σημαντικά την απόδοση της τμηματοποίησης χωρίς να προσθέτουν υπερβολική υπολογιστική πολυπλοκότητα στο μοντέλο[39].

Η μονάδα προσοχής είναι χρήσιμη σε μοντέλα κωδικοποιητή-αποκωδικοποιητή όπως το U-net, καθώς μπορεί να παρέχει τοπικοποιημένη πληροφορία ταξινόμησης σε αντίθεση με την παγκόσμια ταξινόμηση. Στο U-net, αυτό επιτρέπει σε διαφορετικά μέρη του δικτύου να εστιάζουν στην τμηματοποίηση διαφορετικών αντικειμένων. Επιπλέον, με κατάλληλα επισημασμένα δεδομένα εκπαίδευσης, το δίκτυο μπορεί να προσαρμοστεί σε συγκεκριμένα αντικείμενα σε μια εικόνα. Η πύλη προσοχής εφαρμόζει μια λειτουργία στην οποία ο χάρτης χαρακτηριστικών βαραίνεται σύμφωνα με κάθε κλάση, και το δίκτυο μπορεί να προσαρμοστεί ώστε να εστιάζει σε μια συγκεκριμένη κλάση και, επομένως, να δίνει προσοχή σε συγκεκριμένα αντικείμενα σε μια εικόνα. Ενώ υπάρχουν διαφορετικοί τύποι πυλών προσοχής, η προσθετική προσοχή είναι πιο δημοφιλής στην επεξεργασία εικόνας λόγω της παραγωγής υψηλότερης ακρίβειας[39]. Η προσθετική προσοχή λειτουργεί προσθέτοντας τα χαρακτηριστικά του εισόδου με έναν μετασχηματισμό των χαρακτηριστικών που αποκτήθηκαν από μια μονάδα προσοχής, βοηθώντας το δίκτυο να "εστιάσει" στις πιο σημαντικές πληροφορίες για την τρέχουσα εργασία.



Εικόνα 3.3-4: Αρχιτεκτονική πύλης προσοχής [39]

Αυτός ο τύπος προσοχής έχει αποδειχθεί ιδιαίτερα ωφέλιμος σε εφαρμογές όπου η ποιότητα και η λεπτομέρεια της τμηματοποίησης είναι κρίσιμης σημασίας, όπως στην ιατρική απεικόνιση.

Η προσθετική πύλη περιγράφεται από την εξίσωση:

$$q_{att}^l = \psi^T \left(\sigma_1 \left(W_x^T x_i^l + W_g^T g_i + b_g \right) \right) + b_\psi$$

$$\alpha_i^l = \sigma_2(q_{att}^l(x_i^l, g_i; \Theta_{att}))$$

Equation 3.3-3: Εξίσωση προσθετικής πύλης

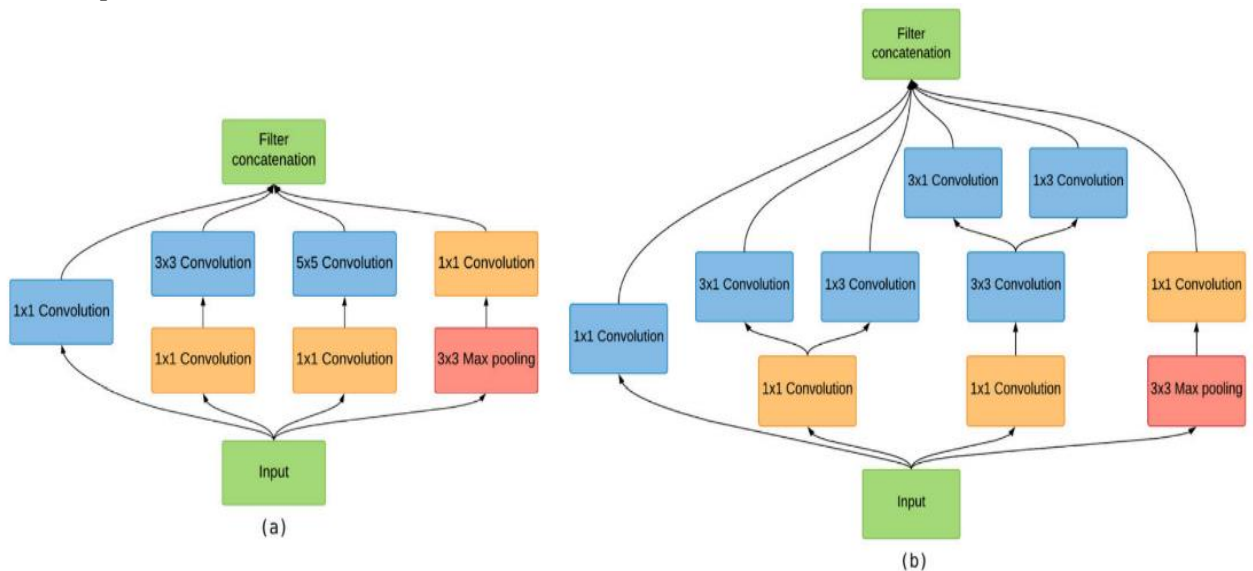
Όπου x^l είναι τα χαρακτηριστικά από τον συμπίεστικό δρόμο και g είναι το σήμα της πύλης. Ο όρος σ αντιπροσωπεύει την σιγμοειδή συνάρτηση:

$$\sigma_2(x_{i,c}) = \frac{1}{1 + \exp(-x_{i,c})}$$

Equation 3.3-4: Εξίσωση στιγμειδούς συνάρτησης

Το Attention U-net έχει εφαρμοστεί με επιτυχία σε προβλήματα όπως η διάγνωση οφθαλμικών νοσημάτων, το μελάνωμα, ο πνευμονικός καρκίνος, ο καρκίνος του τραχήλου, η τμηματοποίηση ανατομικών δομών της κοιλίας, η ανάπτυξη του εμβρύου και η ποσοτικοποίηση των ιστών του εγκεφάλου. Η προσθήκη της λειτουργίας προσοχής στο U-net βοηθά το δίκτυο να εστιάσει σε σημαντικές περιοχές της εικόνας για ακριβέστερη τμηματοποίηση και διάγνωση, ενισχύοντας την αποτελεσματικότητα των βιοϊατρικών εφαρμογών[39].

4. Inception U-net



Εικόνα 3.3-5: a) Αρχική inception αρχιτεκτονική που χρησιμοποιήθηκε στο GoogleNet, b) Βελτιωμένη inception αρχιτεκτονική [39]

Οι περισσότεροι αλγόριθμοι επεξεργασίας εικόνας τείνουν να χρησιμοποιούν φίλτρα σταθερού μεγέθους για τις συνελίξεις. Ωστόσο, η ρύθμιση του μοντέλου για να βρεθεί το σωστό μέγεθος φίλτρου μπορεί συχνά να είναι δύσκολη. Επιπλέον, τα φίλτρα σταθερού μεγέθους είναι κατάλληλα μόνο για εικόνες με παρόμοιου μεγέθους σημαντικά μέρη. Σε πολλές εφαρμογές, η ανάλυση περιλαμβάνει εικόνες με μεγάλες διακυμάνσεις σε σχήματα και μεγέθη στη σημαντική περιοχή. Μία λύση σε αυτό το πρόβλημα θα ήταν η χρήση πιο βαθιών δικτύων που μπορούν να διαβάσουν λεπτομέρειες υψηλού επιπέδου σε ένα φάσμα μεγεθών και σχημάτων. Ωστόσο, τέτοια βαθιά δίκτυα είναι αρκετά υπολογιστικά ακριβά. Μια εναλλακτική λύση, που ονομάζεται δίκτυο inception, χρησιμοποιεί φίλτρα πολλαπλών μεγεθών στο ίδιο στρώμα του δικτύου. Οι εξόδοι από τα διαφορετικά φίλτρα συγκολλώνται και μεταφέρονται στο επόμενο στρώμα. Το δίκτυο inception μπορεί να αναλύσει αποτελεσματικά εικόνες με διαφορετικές σημαντικές περιοχές λόγω των διαφορετικών μεγεθών φίλτρων[39]. Για τη μείωση της υπολογιστικής πολυπλοκότητας, το δίκτυο inception

προσθέτει μια συνέλιξη 1×1 πριν από κάθε φίλτρο 3×3 ή μεγαλύτερο για τη μείωση της διαστασιμότητας. Επιπλέον, στρώματα pooling μπορούν επίσης να προστεθούν παράλληλα σε κάθε μονάδα inception[39].

Το αρχικό δίκτυο inception, γνωστό ως GoogLeNet, πέτυχε τα καλύτερα αποτελέσματα στον διαγωνισμό ILSVRC14. Αυτό ακολουθήθηκε σύντομα από περαιτέρω βελτιώσεις στο δίκτυο, συμπεριλαμβανομένης της εφαρμογής μεθόδων παραγοντοποίησης και της αντικατάστασης της συνέλιξης 5×5 με δύο διαδοχικές συνέλιξεις 3×3 . Στη δεύτερη περίπτωση, μια μόνη συνέλιξη 5×5 είναι 2,78 φορές πιο υπολογιστικά ακριβή από δύο ισοδύναμες συνέλιξεις 3×3 . Επιπλέον παραγοντοποίηση μπορεί να εφαρμοστεί διαχωρίζοντας φίλτρα $n \times n$ σε ένα φίλτρο $1 \times n$ και ένα φίλτρο $n \times 1$ αντίστοιχα. Παραγοντοποιώντας ένα φίλτρο 3×3 με αυτόν τον τρόπο καθιστά το δίκτυο 33% λιγότερο δαπανηρό. Μονάδες inception διαφόρων διαμορφώσεων έχουν εφαρμοστεί σε πλήθος εφαρμογών U-net, συμπεριλαμβανομένων της ανίχνευσης όγκων του εγκεφάλου, της χαρτογράφησης των ιστών του εγκεφάλου, της τμηματοποίησης της καρδιάς, της ανίχνευσης πνευμονικών λεκέδων, της τμηματοποίησης ανθρώπινων εμβρύων και της τμηματοποίησης νευρών μέσω υπερήχων[39].

5. *Residual U-Net*

Αυτή η παραλλαγή του U-net βασίζεται στην αρχιτεκτονική ResNet. Η κίνηση πίσω από το ResNet ήταν να ξεπεραστεί η δυσκολία στην εκπαίδευση πολύ βαθιών νευρωνικών δικτύων. Είναι γνωστό ότι τα νευρωνικά δίκτυα μπορούν να συγκλίνουν γρηγορότερα σε μια λύση όταν υπάρχουν περισσότερα στρώματα. Ωστόσο, τα πειραματικά αποτελέσματα έχουν δείξει ότι η αύξηση του αριθμού των στρωμάτων οδηγεί σε κορεσμό, και περαιτέρω αυξήσεις μπορούν να προκαλέσουν υποβάθμιση της απόδοσης. Αυτή η υποβάθμιση προκύπτει λόγω της απώλειας ταυτοτήτων χαρακτηριστικών σε βαθύτερα νευρωνικά δίκτυα που προκαλείται από την μείωση των κλίσεων στο διάνυμα βάρους. Το ResNet μειώνει αυτό το πρόβλημα χρησιμοποιώντας συνδέσεις παράκαμψης, οι οποίες παίρνουν τον χάρτη χαρακτηριστικών από ένα στρώμα και το προσθέτουν σε ένα άλλο στρώμα πιο βαθιά στο δίκτυο. Αυτή η συμπεριφορά επιτρέπει στο δίκτυο να διατηρήσει καλύτερα τους χάρτες χαρακτηριστικών σε βαθύτερα νευρωνικά δίκτυα και να παρέχει βελτιωμένη απόδοση για τέτοια πιο βαθιά δίκτυα[39].

Στο residual U-net, σε κάθε μπλοκ του δικτύου, η είσοδος στο πρώτο στρώμα συνέλιξης προστίθεται στην έξοδο από το δεύτερο στρώμα συνέλιξης χρησιμοποιώντας μια σύνδεση παράκαμψης. Αυτή η υπολειπόμενη σύνδεση εφαρμόζεται πριν από την υποδειγματοληψία ή την υπερδειγματοληψία στις αντίστοιχες διαδρομές στο U-net. Η χρήση των υπολειπόμενων συνδέσεων παράκαμψης βοηθά στην ανακούφιση του προβλήματος της εξαφανιζόμενης κλίσης, επιτρέποντας έτσι το σχεδιασμό μοντέλων U-net με πιο βαθιά νευρωνικά δίκτυα. Κάθε υπολειπόμενη μονάδα μπορεί να σημειωθεί ως:

$$y_l = h(x_l) + F(x_l, w_l)$$

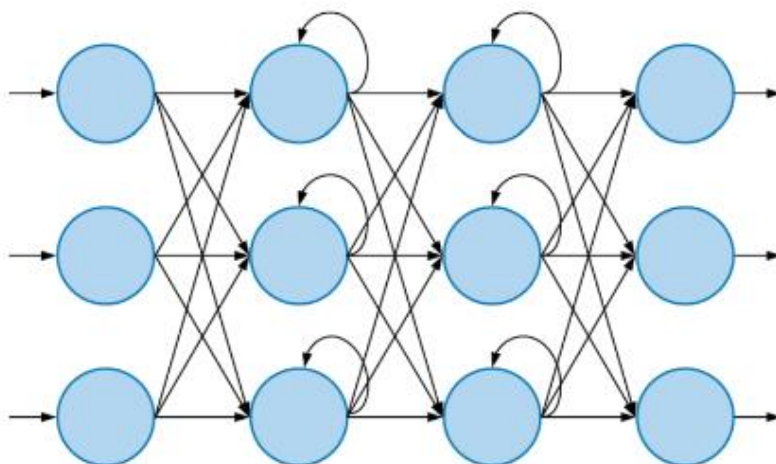
Equation 3.3-5: Ζεύγος σχέσεων υπολειπόμενων μονάδων στα ResNet

$$x_{l+1} = f(y_l)$$

Equation 3.3-6: Ζεύγος σχέσεων υπολειπόμενων μονάδων στα ResNet

όπου x_l και x_{l+1} αντιστοιχούν στην είσοδο και την έξοδο της υπολειπόμενης μονάδας αντίστοιχα, $F(\cdot)$ αντιστοιχεί στην υπολειπόμενη συνάρτηση, $f(\cdot)$ είναι η συνάρτηση ενεργοποίησης, και $h(\cdot)$ είναι η συνάρτηση ταυτότητας[39].

6. Recurrent U-nets



Εικόνα 3.3-6: Το Επαναληπτικό Νευρωνικό Δίκτυο, ένας απλός τύπος δικτύου στο οποίο το δεύτερο και το τρίτο στρώμα είναι επαναληπτικά στρώματα. Κάθε νευρώνας σε ένα επαναληπτικό στρώμα λαμβάνει ανατροφοδότηση από τη δική του έξοδο καθώς και νέα πληροφορία [39]

Τα επαναλαμβανόμενα νευρωνικά δίκτυα (Recurrent Neural Networks - RNN) είναι ένας τύπος νευρωνικού δικτύου που σχεδιάστηκε αρχικά για την ανάλυση διαδοχικών δεδομένων όπως τα κείμενα ή τα ηχητικά δεδομένα. Το δίκτυο σχεδιάζεται με τέτοιο τρόπο ώστε η έξοδος ενός κόμβου να αλλάζει με βάση την προηγούμενη έξοδο από τους ίδιους κόμβους, δηλαδή, ένας βρόχος ανάδρασης αντίθετα με ένα παραδοσιακό δίκτυο feed-forward. Αυτός ο βρόχος ανάδρασης, που αποκαλείται επίσης επαναλαμβανόμενη σύνδεση, δημιουργεί μια εσωτερική κατάσταση ή μνήμη που παρέχει στον κόμβο χρονικές ιδιότητες που αλλάζουν την έξοδο σε διακριτά χρονικά βήματα. Όταν επεκταθεί σε ολόκληρο το στρώμα, αυτό επιτρέπει στο δίκτυο να επεξεργάζεται πληροφορία πλαισίου από τα προηγούμενα δεδομένα[39].

Το επαναλαμβανόμενο U-net κάνει χρήση των επαναλαμβανόμενων συνελκτικών νευρωνικών δικτύων (RCNN) ενσωματώνοντας τους βρόχους ανάδρασης σε ένα συνελκτικό στρώμα. Η ανάδραση εφαρμόζεται μετά τη συνέλιξη και μια συνάρτηση ενεργοποίησης και τροφοδοτεί πίσω στο συσχετισμένο στρώμα τον χάρτη χαρακτηριστικών που παράγεται από ένα φίλτρο.

Η ιδιότητα της ανάδρασης επιτρέπει στις μονάδες να ενημερώνουν τους χάρτες χαρακτηριστικών τους με βάση το πλαίσιο από τις γειτονικές μονάδες, παρέχοντας καλύτερη ακρίβεια και απόδοση. Η έξοδος y του επαναλαμβανόμενου συνελκτικού νευρωνικού δικτύου μπορεί να εκφραστεί ως:

$$y_{ijk}^l(t) = (w_k^f)^T x_l^{f(i,j)}(t) + (w_k^r)^T x_l^{r(i,j)}(t-1) + b_k$$

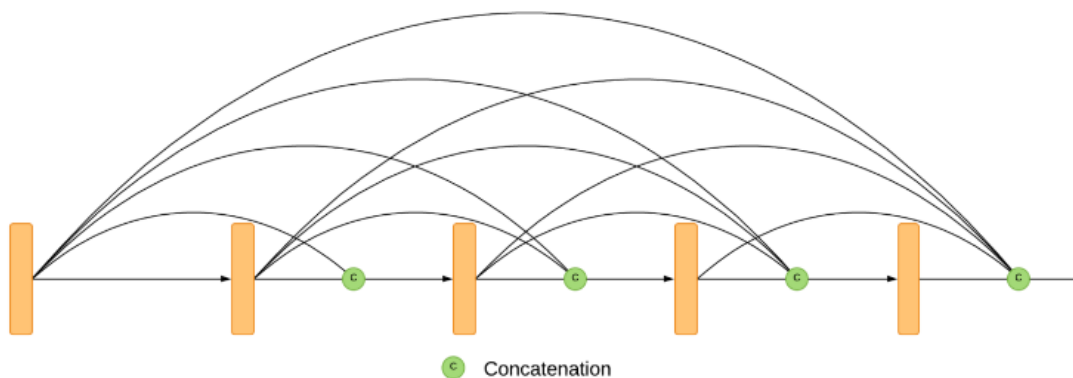
Equation 3.3-7: Εξίσωση εξόδου του Recurrent U-net

Όπου $x_l^f(t)$ είναι η είσοδος από το feedforward και $x_l^r(t-1)$ είναι η επαναλαμβανόμενη είσοδος για το l-οστό στρώμα, w_k^f είναι το βάρος του feedforward, w_k^r είναι το επαναλαμβανόμενο βάρος, και b_k είναι η πόλωση (bias) του k-οστού χάρτη χαρακτηριστικών. Ο Alom και οι συνεργάτες του σχεδίασαν ένα μοντέλο U-net που περιέχει και επαναλαμβανόμενες συνδέσεις και υπολειπόμενες συνδέσεις[39]. Το προκύπτον δίκτυο ξεπέρασε τα μοντέλα U-net που

χρησιμοποιούν μόνο υπολειπόμενες ή επαναλαμβανόμενες συνδέσεις, καθώς και προηγούμενες κορυφαίες αρχιτεκτονικές μεθόδους, χρησιμοποιώντας παρόμοιο αριθμό παραμέτρων.

7. *Dense U-nets*

Τα Dense U-nets χρησιμοποιούν μπλοκ DenseNet αντί για κανονικά στρώματα. Ενώ το μοντέλο ResNet επιτρέπει βαθύτερα νευρωνικά δίκτυα, δεν εξαλείφει το πρόβλημα των εξαφανιζόμενων κλίσεων. Η αρχιτεκτονική ResNet επίσης τελικά υποβαθμίζεται σε απόδοση με την αύξηση των στρωμάτων. Για να διορθωθεί αυτό, το DenseNet είναι μια αρχιτεκτονική βαθιάς μάθησης που έχει κατασκευαστεί πάνω στο ResNet με δύο κύριες αλλαγές. Πρώτον, κάθε στρώμα σε ένα μπλοκ λαμβάνει τον χάρτη χαρακτηριστικών ή την ταυτότητα από όλα τα προηγούμενα στρώματά του. Η δεύτερη μεγάλη αλλαγή είναι ότι οι χάρτες ταυτότητας συνδυάζονται μέσω συγκόλλησης κατά κανάλι σε τανυστές, σε αντίθεση με το ResNet, στο οποίο οι χάρτες ταυτότητας συνοψίζονται μέσω πρόσθεσης στοιχείο προς στοιχείο. Έτσι, η αναγνώριση ταυτότητας κάθε στρώματος εξαρτάται όχι μόνο από το προηγούμενο στρώμα αλλά και από όλα τα προηγούμενα στρώματα στο μπλοκ[39].



Εικόνα 3.3-7: Ένα πεντάστρωμο πυκνό μπλοκ. Η μονάδα συγχώνευσης λαμβάνει τον χάρτη χαρακτηριστικών από όλα τα προηγούμενα στρώματα και το περνά στο επόμενο στρώμα. [39]

Αυτό επιτρέπει στο DenseNet να διατηρεί όλους τους χάρτες ταυτότητας από προηγούμενα στρώματα και να προωθεί σημαντικά τη διάδοση της κλίσης. Η συνέπεια είναι ότι κάθε στρώμα μπορεί να έχει λιγότερα κανάλια, καθώς οι πληροφορίες διατηρούνται πιο εύκολα μεταξύ των στρωμάτων, με αποτέλεσμα υψηλότερη ακρίβεια με λιγότερους υπολογισμούς, που με τη σειρά του επιτρέπει στα μοντέλα βαθιάς μάθησης με μεγαλύτερο αριθμό στρωμάτων. Η έξοδος για κάθε στρώμα σε ένα πυκνό μπλοκ περιγράφεται ως:

$$x_i = H_i ([x_0, x_1, x_2, \dots, x_{i-1}])$$

Equation 3.3-8: Εξίσωση εξόδου στρώματος σε Dense U-net

όπου $H_i(\cdot)$ αντιπροσωπεύει την πυκνή συνάρτηση απεικόνισης, που συνήθως περιλαμβάνει κανονικοποίηση παρτίδας, ενεργοποίηση ReLU και ένα στρώμα συνέλιξης ενώ $[\cdot]$ σημαίνει συγκόλληση κατά κανάλι[39]. Κατά την υλοποίηση ενός U-net, κάθε παραδοσιακό μπλοκ U-

net αντικαθίσταται με ένα πυκνό μπλοκ δύο ή περισσότερων συνελκτικών στρωμάτων. Η υιοθέτηση πυκνών μπλοκ επιτρέπει βαθύτερα μοντέλα U-net, τα οποία μπορούν να τμηματοποιούν αντικείμενα σε μια εικόνα με μεγαλύτερη διάκριση. Αυτό το χαρακτηριστικό των πυκνών U-nets είναι ιδιαίτερα επιθυμητό στην ανάλυση ιατρικών εικόνων λόγω της υψηλής πυκνότητας των αντικειμένων σε τέτοιες εικόνες, πολλές φορές σε σημείο επικάλυψης. Εφαρμογές του πυκνού U-net έχουν βρεθεί στην ανάλυση όγκων του εγκεφάλου, στην τμηματοποίηση αιμοφόρων αγγείων του αμφιβληστροειδούς, στην τμηματοποίηση εγκεφαλικών αιμοφόρων αγγείων, στο μελάνωμα, στον πνευμονικό καρκίνο, στον ηπατικό καρκίνο και στην τμηματοποίηση πολλαπλών οργάνων[39].

4 Κεφάλαιο 4: Image Registration

4.1 Εισαγωγή

Η τρισδιάστατη (3D) απεικόνιση με μαγνητική τομογραφία (MRI) διαδραματίζει κρίσιμο ρόλο στην ιατρική διαγνωστική και στον σχεδιασμό θεραπείας. Η εγγραφή εικόνων 3D MRI είναι μια θεμελιώδης διαδικασία που συγχρονίζει διαφορετικές εικόνες σε ένα κοινό χωρικό πλαίσιο για ανάλυση και σύγκριση. Αυτή η ενότητα εξετάζει πρόσφατες προόδους στην εγγραφή εικόνων 3D MRI, επικεντρώνοντας σε διάφορες μεθοδολογίες και τις εφαρμογές τους στην ιατρική διαγνωστική, στον σχεδιασμό θεραπείας και στην έρευνα[41]. Τονίζονται βασικές τεχνικές, συμπεριλαμβανομένων των μεθόδων βασισμένων στην ένταση, στη δομή, στην κλίση και στην εκμάθηση, και συζητείται η αποτελεσματικότητά τους στην αντιμετώπιση των προκλήσεων σχετικά με την ακρίβεια, την αποδοτικότητα και την ευκολία στην εφαρμογή σε κλινικά και ερευνητικά περιβάλλοντα.[42]

Η εγγραφή εικόνων 3D MRI αποτελεί ένα ζωτικό βήμα στην ιατρική απεικόνιση, επιτρέποντας τη σύγκριση, την ένταξη και την ανάλυση εικόνων από διαφορετικές χρονικές στιγμές, μεθόδους ή άτομα. Η ακριβής εγγραφή διευκολύνει την ακριβή διάγνωση νόσων, τον χειρουργικό σχεδιασμό και την μετεγχειρητική αξιολόγηση, ενισχύοντας την αποτελεσματικότητα των ιατρικών παρεμβάσεων.

4.2 Τεχνικές στην Εγγραφή Εικόνων 3D MRI

4.2.1 Μέθοδοι Βασισμένοι στην Ένταση

Η εγγραφή βασισμένη στην ένταση συγχρονίζει τις εικόνες συγκρίνοντας τα μοτίβα έντασης μεταξύ τους. Αυτές οι μέθοδοι δεν βασίζονται σε εξαγωγές χαρακτηριστικών ή ορόσημα αλλά απευθείας στις τιμές των εικονοστοιχείων ή των βοξέλ. Κοινώς χρησιμοποιούμενες μετρικές για την εγγραφή βασισμένη στην ένταση περιλαμβάνουν την αμοιβαία πληροφορία, την διασταύρωση συσχέτισης και την κανονικοποιημένη αμοιβαία πληροφορία[43]. Παρόλο που

αυτές οι τεχνικές είναι ευρέως εφαρμόσιμες και δεν απαιτούν ρητή εξαγωγή χαρακτηριστικών, μπορούν να είναι υπολογιστικά απαιτητικές και μπορεί να επηρεαστούν από αλλαγές στην ένταση λόγω παραγόντων όπως οι ρυθμίσεις του σαρωτή ή αντιθετικοί παράγοντες αντί για πραγματικές ανατομικές διαφορές.

4.2.2 Μέθοδοι Βασισμένοι στη Δομή

Οι μέθοδοι βασισμένοι στη δομή (ή βασισμένοι σε χαρακτηριστικά) περιλαμβάνουν την ευθυγράμμιση εικόνων με βάση τα αναγνωρισμένα χαρακτηριστικά ή δομές εντός των εικόνων, όπως άκρα, περιγράμματα ή συγκεκριμένα ανατομικά ορόσημα. Αυτές οι τεχνικές συχνά περιλαμβάνουν ένα αρχικό βήμα ανίχνευσης χαρακτηριστικών, ακολουθούμενο από μια διαδικασία αντιστοίχισης όπου τα αντίστοιχα χαρακτηριστικά στις στοχευμένες και αναφερόμενες εικόνες ευθυγραμμίζονται[44]. Αυτή η προσέγγιση μπορεί να προσφέρει βελτιωμένη ακρίβεια σε περιπτώσεις όπου οι εικόνες έχουν σημαντικές δομικές διαφορές ή όταν οι εικόνες που πρέπει να εγγραφούν έχουν διαφορετικά modalities. Ωστόσο, η επιτυχία των μεθόδων βασισμένων στη δομή εξαρτάται σημαντικά από την ακριβή ανίχνευση και αντιστοίχιση των χαρακτηριστικών[45], που μπορεί να αποτελέσει πρόκληση σε εικόνες χαμηλής αντίθεσης ή όταν τα χαρακτηριστικά ενδιαφέροντος δεν είναι καλά ορισμένα.

4.2.3 Μέθοδοι Βασισμένοι στην Κλίση

Οι τεχνικές εγγραφής βασισμένες στην κλίση χρησιμοποιούν τις κλίσεις των εντάσεων της εικόνας για να καθοδηγήσουν τη διαδικασία εγγραφής. Επικεντρώνοντας στις αλλαγές εντάσεων, αντί για τις απόλυτες τιμές εντάσεων, αυτές οι μέθοδοι στοχεύουν στην ευθυγράμμιση περιοχών με υψηλές πληροφορίες χωρικής συχνότητας, όπως τα άκρα[46]. Αυτό μπορεί να κάνει τις προσεγγίσεις βασισμένες στην κλίση ιδιαίτερα αποτελεσματικές για εικόνες όπου η περιοχή ενδιαφέροντος ορίζεται από σαφή όρια. Αυτές οι μέθοδοι είναι λιγότερο ευαίσθητες σε ομοιόμορφες διακυμάνσεις εντάσεων που δεν επηρεάζουν τις πληροφορίες κλίσης. Ωστόσο, η απόδοσή τους μπορεί να περιοριστεί σε εικόνες με ομαλές διακυμάνσεις εντάσεων ή όπου οι πληροφορίες κλίσης δεν είναι ξεκάθαρες.

4.2.4 Μέθοδοι Βασισμένοι στην Εκμάθηση

Οι προσεγγίσεις βασισμένες στην εκμάθηση, ιδιαίτερα αυτές που χρησιμοποιούν τη βαθιά μάθηση, αποτελούν την πρωτοπορία στην εγγραφή εικόνων 3D MRI. Αυτές οι μέθοδοι εκπαιδεύουν νευρωνικά δίκτυα σε ένα σύνολο εικόνων για να μάθουν τη λειτουργία εγγραφής απευθείας από τα δεδομένα. Αυτό μπορεί να γίνει με επιβλεπόμενο τρόπο, χρησιμοποιώντας ζευγάρια εικόνων με γνωστές εγγραφές ως δεδομένα εκπαίδευσης[47], ή μη επιβλεπόμενο, όπου το δίκτυο μαθαίνει να βελτιστοποιεί μια μετρική εγγραφής χωρίς ρητές συμφωνίες αλήθειας εδάφους. Το τελευταίο είναι ιδιαίτερα σημαντικό για την ιατρική απεικόνιση, όπου οι πραγματικές εγγραφές συχνά δεν είναι διαθέσιμες. Οι μέθοδοι βασισμένοι στην εκμάθηση μπορούν να επιτύχουν υψηλά επίπεδα ακρίβειας και αποδοτικότητας,[48] επεξεργάζονται εικόνες γρήγορα μόλις το δίκτυο εκπαιδευτεί. Επίσης, είναι ικανά να διαχειριστούν μια ευρεία ποικιλία εργασιών εγγραφής χωρίς την ανάγκη να προσαρμόσουν τον αλγόριθμο εγγραφής για συγκεκριμένα χαρακτηριστικά εικόνων. Ωστόσο, η επιτυχία αυτών των μεθόδων εξαρτάται από τη διαθεσιμότητα μεγάλων, διαφορετικών συνόλων εκπαίδευσης δεδομένων και μπορεί

να επηρεαστεί από την επιλογή της αρχιτεκτονικής του δικτύου και των συναρτήσεων απώλειας.

4.2.5 Εφαρμογές στην Ιατρική Απεικόνιση

Η εγγραφή εικόνων 3D MRI βρίσκει εφαρμογές σε διάφορους ιατρικούς τομείς[49], συμπεριλαμβανομένων της νευρολογίας, της ογκολογίας και της ορθοπαιδικής. Βοηθά σε μακροχρόνιες μελέτες, την παρακολούθηση θεραπείας και την ένταξη δεδομένων από πολυτρόπους απεικονιστικούς τρόπους, συμβάλλοντας στη βελτιωμένη φροντίδα των ασθενών[49]. Η επιλογή της μεθόδου εγγραφής εξαρτάται από τις συγκεκριμένες απαιτήσεις της εφαρμογής, συμπεριλαμβανομένων των τύπων των εικόνων που πρέπει να εγγραφούν, της επιθυμητής ακρίβειας και ανθεκτικότητας, καθώς και των υπολογιστικών περιορισμών. Καθώς ο τομέας προοδεύει, οι υβριδικές προσεγγίσεις που συνδυάζουν τα δυνατά σημεία διαφόρων μεθόδων, καθώς και οι πρόοδοι στη μηχανική μάθηση, είναι πιθανό να διαδραματίσουν σημαντικό ρόλο στην περαιτέρω βελτίωση της απόδοσης και της ευχρηστίας των τεχνικών εγγραφής εικόνων 3D MRI.

4.3 Similarity Metrics

Στην εγγραφή εικόνας, ιδιαίτερα στην ιατρική απεικόνιση, οι μετρικές ομοιότητας είναι κρίσιμες για τον καθορισμό του πόσο στενά συμφωνούν δύο εικόνες[50] (μια σταθερή εικόνα και μια κινούμενη ή στοχευμένη εικόνα) μεταξύ τους κατά τη διάρκεια της διαδικασίας εγγραφής. Η επιλογή της μετρικής ομοιότητας εξαρτάται από τη φύση των εικόνων που εγγράφονται, τον τύπο της μετατροπής (σταθερή, αφινική, παραμορφωτική) και την συγκεκριμένη εφαρμογή.[50] Ακολουθεί μια ανάλυση κάποιων από τις πιο συνηθισμένα χρησιμοποιούμενες μετρικές ομοιότητας στην εγγραφή εικόνας:

Αμοιβαία Πληροφορία (MI) και Κανονικοποιημένη Αμοιβαία Πληροφορία (NMI)

Η Αμοιβαία Πληροφορία είναι μια μετρική που προέρχεται από τη θεωρία πληροφορίας, και ποσοτικοποιεί την "πληροφορία" που η μία εικόνα περιέχει για την άλλη[51]. Είναι ιδιαίτερα αποτελεσματική για την πολυ-τροπική εγγραφή, όπου οι δύο εικόνες μπορεί να έχουν διαφορετικές κατανομές εντάσεων (π.χ., από MRI σε CT).

Η MI υπολογίζει πόσο η γνώση της μιας εικόνας μειώνει την αβεβαιότητα για την άλλη[51]. Η NMI κανονικοποιεί αυτή την τιμή, καθιστώντας τη λιγότερο ευαίσθητη σε αλλαγές του μεγέθους επικάλυψης των εικόνων.

Χρησιμοποιείται ευρέως σε πολυτροπικά σενάρια, όπως η εγγραφή εικόνων από διαφορετικές μεθόδους απεικόνισης.[52]

Διασταυρούμενη Συσχέτιση (CC)

Η Διασταυρούμενη Συσχέτιση μετρά τη γραμμική σχέση μεταξύ των εντάσεων των δύο εικόνων. Είναι πιο αποτελεσματική όταν οι εικόνες έχουν μια γραμμική σχέση εντάσεων.

Η CC υπολογίζει έναν συντελεστή συσχέτισης που αντιπροσωπεύει τον βαθμό στον οποίο οι εικόνες είναι γραμμικά συσχετισμένες[53].

Ιδιαίτερα χρήσιμη στην εγγραφή εικόνων single modality, όπου οι διακυμάνσεις εντάσεων μεταξύ των εικόνων είναι γραμμικές.

Άθροισμα Τετραγωνικών Διαφορών (SSD)

Το Άθροισμα Τετραγωνικών Διαφορών είναι μια απλή και διαισθητική μετρική, που υπολογίζει το άθροισμα των τετραγωνικών διαφορών εντάσεων μεταξύ των αντίστοιχων βοξέλ στις δύο εικόνες. [54]

Απαιτεί το τετράγωνο της διαφοράς εντάσεως για κάθε ζεύγος αντίστοιχων βοξέλ και την άθροιση αυτών των τιμών σε ολόκληρη την εικόνα.

Ιδανικό για εγγραφή εντός της ίδιας modality όπου οι εικόνες έχουν παρόμοια προφίλ εντάσεων.

Κανονικοποιημένη Διασταυρούμενη Συσχέτιση (NCC)

Μια επέκταση της CC, η Κανονικοποιημένη Διασταυρούμενη Συσχέτιση, λαμβάνει υπόψη αλλαγές στη φωτεινότητα και την αντίθεση της εικόνας, παρέχοντας μια πιο ανθεκτική μέτρηση ομοιότητας.

Η NCC κανονικοποιεί το γινόμενο των διαφορών εντάσεων με τις τυπικές αποκλίσεις των εντάσεων της εικόνας, καθιστώντας την ανεπηρέαστη από γραμμικές αλλαγές στην ένταση της εικόνας. [55]

Χρήσιμη σε σενάρια όπου οι εικόνες μπορεί να διαφέρουν σε φωτεινότητα ή αντίθεση αλλά είναι κατά τα άλλα παρόμοιες.

Λόγος Συσχέτισης (CR)

Ο Λόγος Συσχέτισης μετρά τη στατιστική εξάρτηση μεταξύ των εντάσεων των αντίστοιχων βοξέλ στις δύο εικόνες. Είναι αποτελεσματικός στο να αποτυπώνει μη γραμμικές σχέσεις.

Ο CR αξιολογεί πόσο καλά η ένταση της μιας εικόνας μπορεί να προβλέψει την ένταση της άλλης, λαμβάνοντας υπόψη μη γραμμικές διακυμάνσεις εντάσεων.[56] Εφαρμόσιμος τόσο σε multimodal όσο και σε intramodal εγγραφή, ιδίως όταν αναμένονται μη γραμμικές διακυμάνσεις εντάσεων.

Δείκτης Δομικής Ομοιότητας (SSIM)

Ο Δείκτης Δομικής Ομοιότητας μετρά την αντιληπτή διαφορά ποιότητας μεταξύ δύο εικόνων. Λαμβάνει υπόψη αλλαγές στη φωτεινότητα, την αντίθεση και τη δομή. Ο SSIM αξιολογεί την ομοιότητα μεταξύ παραθύρων των εικόνων, λαμβάνοντας υπόψη ξεχωριστά τη φωτεινότητα, την αντίθεση και τα δομικά στοιχεία, και στη συνέχεια συνδυάζοντάς τα[57].

Χρησιμοποιείται σε μελέτες όπου η διατήρηση της δομικής ακεραιότητας της εικόνας είναι κρίσιμη, όπως σε κάποιες εφαρμογές νευροαπεικόνισης.

4.4 Elastix Registration

Το Elastix είναι ένα λογισμικό ανοιχτού κώδικα που χρησιμοποιείται ευρέως στον τομέα της εγγραφής ιατρικών εικόνων, γνωστό για την ευελιξία και την ανθεκτικότητά του στην ευθυγράμμιση διαφορετικών εικόνων μέσω του καθορισμού βέλτιστων χωρικών μετασχηματισμών[58]. Η σχεδίασή του είναι κατ' ουσίαν ενόητος, επιτρέποντας εύκολα τα πειράματα και τις ρυθμίσεις σε διάφορα στοιχεία της διαδικασίας εγγραφής, συμπεριλαμβανομένων των μοντέλων μετατροπής, των μετρικών ομοιότητας, των μεθόδων βελτιστοποίησης, και των πυραμίδων εικόνων. Αυτή η ευελιξία καθιστά το Elastix κατάλληλο για μια ευρεία γκάμα εργασιών εγγραφής, από σταθερές και αφινικές μετατροπές που απαιτούνται για την παγκόσμια ευθυγράμμιση έως μη σταθερές μετατροπές B-spline που αντιμετωπίζουν τοπικές προσαρμογές[58].

Μία από τις δυνάμεις του Elastix βρίσκεται στην ικανότητά του να χρησιμοποιεί πολλαπλές μετρικές ομοιότητας, όπως η αμοιβαία πληροφορία και η κανονικοποιημένη διασταυρούμενη συσχέτιση. Αυτό το καθιστά ικανό να ανταποκρίνεται τόσο σε μονο-τροπικές όσο και σε πολυ-τροπικές σενάρια εγγραφής εικόνων αποτελεσματικά. Επιπλέον, το λογισμικό υλοποιεί αρκετές στρατηγικές βελτιστοποίησης, συμπεριλαμβανομένων της καθοδικής κλίσης και των μεθόδων παραπλήσιων του Newton. Αυτές οι μέθοδοι στοχεύουν στον εντοπισμό των παραμέτρων του μοντέλου μετατροπής που μεγιστοποιούν την επιλεγμένη μετρική ομοιότητας, εξασφαλίζοντας ακριβή ευθυγράμμιση μεταξύ των εικόνων.

Το Elastix χρησιμοποιεί στρατηγικές πολλαπλών αναλύσεων μέσω πυραμίδων εικόνων, οι οποίες βελτιώνουν σταδιακά την εγγραφή από χονδρικές σε λεπτομερείς αναλύσεις. Αυτή η προσέγγιση ενισχύει την αποδοτικότητα και την ανθεκτικότητα της διαδικασίας εγγραφής, καθιστώντας το Elastix ιδιαίτερα αποτελεσματικό σε κλινικές και ερευνητικές εφαρμογές.

Για παράδειγμα, η χρήση του στην προσαρμοστική ακτινοθεραπεία για τον καρκίνο του προστάτη έχει δείξει γρήγορη και ανθεκτική απόδοση, τονίζοντας την καταλληλότητά του για κλινικές εφαρμογές. Επιπλέον, σε σύγκριση με άλλα εργαλεία σε μελέτες που αφορούν την εγγραφή εικόνων εγκεφάλου υψηλής ανάλυσης 3D, το Elastix έχει αποδειχθεί ότι παρέχει σημαντική χρησιμότητα και αποδοτικότητα.

Πρόσφατες μελέτες έχουν επίσης εξερευνήσει τον συνδυασμό του Elastix με τεχνικές βαθιάς μάθησης, παρουσιάζοντας το δυναμικό του στη βελτίωση της ακρίβειας και της αποδοτικότητας της εγγραφής[58]. Για παράδειγμα, η εφαρμογή του στη μη επιβλεπόμενη μάθηση για την εγγραφή εικόνων CT θώρακα ενός αντικειμένου και η ενσωμάτωσή του σε έναν σωλήνα εγγραφής για τον σχεδιασμό θεραπείας καρκίνου του προστάτη υπογραμμίζουν την ικανότητα του λογισμικού να ενσωματωθεί με προηγμένες υπολογιστικές τεχνικές. Αυτή η ενσωμάτωση δεν βελτιώνει μόνο την ανάλυση ιατρικών εικόνων αλλά προωθεί επίσης την περίθαλψη ασθενών, διευκολύνοντας πιο ακριβή και προσωποποιημένο σχεδιασμό θεραπείας. Συνολικά, οι περιεκτικές δυνατότητες και η ευελιξία του Elastix, μαζί με την απόδοσή του σε διάφορα πλαίσια ιατρικής απεικόνισης και το δυναμικό του για ενσωμάτωση με τη βαθιά μάθηση, αποδεικνύουν τον καθοριστικό του ρόλο στην πρόωθηση της ανάλυσης ιατρικών εικόνων και στη βελτίωση των αποτελεσμάτων για τους ασθενείς.

5 Κεφάλαιο 5: Classification

Η ταξινόμηση, στο πλαίσιο της μηχανικής μάθησης και της στατιστικής, αποτελεί μια θεμελιώδη διαδικασία που χρησιμοποιείται για την κατηγοριοποίηση ή την επισήμανση σημείων δεδομένων σε προκαθορισμένες κλάσεις ή κατηγορίες. Περιλαμβάνει την ανάλυση των χαρακτηριστικών ενός συνόλου δεδομένων για την αναγνώριση προτύπων που μπορούν να χρησιμοποιηθούν για τον καθορισμό των ετικετών κλάσης νέων, αδιαφανών περιπτώσεων. Η ταξινόμηση είναι ένας τύπος επιβλεπόμενης μάθησης, που σημαίνει ότι βασίζεται σε ένα επισημειωμένο σύνολο δεδομένων όπου κάθε παράδειγμα συνδέεται με μια σωστή ετικέτα.

Ο πρωταρχικός στόχος της ταξινόμησης είναι να προβλέψει με ακρίβεια την κατηγορία ή την ετικέτα κλάσης νέων περιπτώσεων βάσει προηγούμενων παρατηρήσεων. Χρησιμοποιείται ευρέως σε διάφορους τομείς, συμπεριλαμβανομένης της ανίχνευσης ανεπιθύμητων μηνυμάτων στα email, της ανάλυσης συναισθήματος πελατών, της ιατρικής διάγνωσης, της αναγνώρισης εικόνας και πολλών άλλων.

Οι εργασίες ταξινόμησης στη μηχανική μάθηση μπορούν να κατηγοριοποιηθούν ευρέως με βάση τη φύση της διαδικασίας μάθησης και τον τύπο της παραγόμενης εξόδου.

5.1 Τύποι Ταξινόμησης

Επιβλεπόμενη Ταξινόμηση

Η επιβλεπόμενη ταξινόμηση είναι η πιο διαδεδομένη μορφή ταξινόμησης, όπου το μοντέλο μαθαίνει από ένα επισημειωμένο σύνολο δεδομένων. Αυτό το σύνολο δεδομένων περιέχει παραδείγματα εισόδων μαζί με τις αντίστοιχες ετικέτες στόχου. Ο κύριος στόχος είναι να εκπαιδευτεί ένα μοντέλο που μπορεί να προβλέψει με ακρίβεια τις ετικέτες νέων, αδιαφανών περιπτώσεων με βάση τα μαθημένα πρότυπα. Η επιβλεπόμενη ταξινόμηση μπορεί να διαιρεθεί περαιτέρω σε:

Διαδική Ταξινόμηση: Η στοχευόμενη μεταβλητή έχει δύο κατηγορίες. Κοινά παραδείγματα περιλαμβάνουν την ανίχνευση ανεπιθύμητων μηνυμάτων (spam ή όχι spam) και τη διάγνωση νόσων (άρρωστος ή υγιής).

Πολυκλασική Ταξινόμηση: Η στοχευόμενη μεταβλητή έχει περισσότερες από δύο κατηγορίες. Παραδείγματα περιλαμβάνουν την αναγνώριση ψηφίων (0 έως 9) και την πρόβλεψη τύπων κουζίνας από συστατικά[59].

Μη Επιβλεπόμενη Ταξινόμηση (Συσταδοποίηση)

Η μη επιβλεπόμενη ταξινόμηση, που συχνά αναφέρεται ως συσταδοποίηση, περιλαμβάνει την ομαδοποίηση σημείων δεδομένων με βάση την ομοιότητα χωρίς προκαθορισμένα επισημειωμένα παραδείγματα. Ο στόχος είναι να ανακαλυφθούν εγγενείς ομαδοποιήσεις μέσα στα δεδομένα, όπου τα μέλη κάθε συστάδας μοιράζονται κοινά χαρακτηριστικά. Η συσταδοποίηση είναι χρήσιμη για την εξερευνητική ανάλυση δεδομένων, τη διαφοροποίηση πελατών και την ανίχνευση ανωμαλιών. Κοινοί αλγόριθμοι συσταδοποίησης περιλαμβάνουν τον K-means, την ιεραρχική συσταδοποίηση και τον DBSCAN[60].

Ημι-Επιβλεπόμενη Ταξινόμηση

Η ημι-επιβλεπόμενη ταξινόμηση βρίσκεται μεταξύ της επιβλεπόμενης και μη επιβλεπόμενης μάθησης. Χρησιμοποιεί μια μικρή ποσότητα επισημειωμένων δεδομένων μαζί με ένα μεγαλύτερο σύνολο μη επισημειωμένων δεδομένων κατά τη διάρκεια της εκπαίδευσης. Αυτή η προσέγγιση είναι ιδιαίτερα χρήσιμη όταν η επισήμανση δεδομένων είναι δαπανηρή ή χρονοβόρα. Η ημι-επιβλεπόμενη μάθηση στοχεύει να εκμεταλλευτεί τα μη επισημειωμένα δεδομένα για να αποτυπώσει καλύτερα την υποκείμενη δομή της κατανομής των δεδομένων, βελτιώνοντας έτσι την απόδοση του μοντέλου σε άγνωστα δεδομένα[61].

Ενισχυτική Μάθηση

Αν και δεν κατατάσσεται παραδοσιακά υπό την ομπρέλα των εργασιών ταξινόμησης, η ενισχυτική μάθηση περιλαμβάνει τη μάθηση για τη λήψη αποφάσεων σε σειρά. Μαθαίνει μια πολιτική για να δρα σε ένα περιβάλλον με στόχο τη μεγιστοποίηση κάποιας έννοιας της συσσωρευτικής ανταμοιβής. Η ενισχυτική μάθηση μπορεί να χρησιμοποιηθεί για εργασίες λήψης αποφάσεων όπου οι αποφάσεις (ή οι ενέργειες) μπορούν να θεωρηθούν ως ταξινόμηση καταστάσεων σε ενέργειες[62].

Πολυ-ετικετική Ταξινόμηση

Στην πολυ-ετικετική ταξινόμηση, σε κάθε περίπτωση μπορούν να αποδοθούν πολλαπλές ετικέτες αντί για μία μόνο. Αυτό διαφέρει από την πολυκλασική ταξινόμηση, όπου κάθε περίπτωση ανατίθεται σε μία από αρκετές κλάσεις. Η πολυ-ετικετική ταξινόμηση είναι χρήσιμη σε σενάρια όπου οι κατηγορίες δεν είναι αμοιβαία αποκλειστικές, όπως η επισήμανση φωτογραφιών με πολλαπλές ετικέτες (π.χ., παραλία, ηλιοβασίλεμα, διακοπές)[63].

Ταξινόμηση Πολλαπλής Εξόδου

Η ταξινόμηση πολλαπλής εξόδου ή πολλαπλού στόχου περιλαμβάνει την πρόβλεψη πολλαπλών εξαρτημένων μεταβλητών για κάθε δείγμα εισόδου. Αποτελεί μια επέκταση της πολυ-ετικετικής ταξινόμησης, όπου κάθε ετικέτα μπορεί να είναι πολυκλασική και υπάρχει συσχέτιση μεταξύ των εξόδων ετικετών. Είναι εφαρμόσιμη σε σενάρια όπως η πρόβλεψη πολλαπλών ιδιοτήτων αντικειμένων σε εικόνες, όπου κάθε ιδιότητα (π.χ., χρώμα, σχήμα, μέγεθος) μπορεί να έχει πολλαπλές κατηγορίες[64].

5.2 Τύποι Αλγορίθμων Ταξινόμησης

Δέντρα Αποφάσεων

Διαχωρίζουν τα δεδομένα σε υποσύνολα χρησιμοποιώντας απλούς κανόνες αποφάσεων που συνάγονται από τα χαρακτηριστικά. Είναι διαισθητικό και εύκολο στην απεικόνιση αλλά μπορεί να είναι επιρρεπές στην υπερπροσαρμογή[65].

Τυχαία Δάση:

Μια μέθοδος συνόλου που δημιουργεί πολλαπλά δέντρα αποφάσεων και τα συγχωνεύει μαζί για να λάβει μια πιο ακριβή και σταθερή πρόβλεψη. Είναι ανθεκτική έναντι της

υπερπροσαρμογής

και

πολύ

ευέλικτη[65].

Μηχανές Διανυσματικής Υποστήριξης (SVM)

Βρίσκει το υπερεπίπεδο που διαχωρίζει καλύτερα τις διαφορετικές κλάσεις στο χώρο χαρακτηριστικών. Το SVM είναι αποτελεσματικό σε χώρους υψηλής διάστασης και για περιπτώσεις όπου ο αριθμός των διαστάσεων υπερβαίνει τον αριθμό των δειγμάτων[65].

K-Κοντινότεροι

Γείτονες

(K-NN)

Ταξινομεί περιπτώσεις με βάση την πλειοψηφία των ψήφων των πλησιέστερων γειτόνων τους. Το K-NN είναι απλό και αποτελεσματικό αλλά υπολογιστικά απαιτητικό για μεγάλα σύνολα δεδομένων[65].

Naive

Bayes

Ένας πιθανοκρατικός ταξινομητής που βασίζεται στην εφαρμογή του θεωρήματος του Bayes με την υπόθεση της ανεξαρτησίας μεταξύ των χαρακτηριστικών. Είναι αποδοτικός και αποδίδει καλά σε πολλά σενάρια, ιδιαίτερα για την ταξινόμηση κειμένου[65].

Λογιστική Παλινδρόμηση:

Παρά το όνομά της, χρησιμοποιείται για προβλήματα δυαδικής ταξινόμησης. Μοντελοποιεί τις πιθανότητες της προεπιλεγμένης κλάσης μέσω μιας λογιστικής συνάρτησης[65].

Νευρωνικά Δίκτυα και Βαθιά Μάθηση:

Αποτελούνται από στρώματα νευρώνων που μαθαίνουν να κάνουν ταξινομήσεις από πολύπλοκα δεδομένα. Μοντέλα βαθιάς μάθησης, όπως τα Συνελκτικά Νευρωνικά Δίκτυα (CNNs) για ταξινόμηση εικόνων και τα Επαναληπτικά Νευρωνικά Δίκτυα (RNNs) για δεδομένα ακολουθιών, είναι ισχυρά για την αντιμετώπιση πολύπλοκων προτύπων αλλά απαιτούν μεγάλες ποσότητες δεδομένων και υπολογιστικούς πόρους.

Μηχανές Ενίσχυσης Κλίσης (GBM): Μια τεχνική συνόλου που κατασκευάζει μοντέλα διαδοχικά, με κάθε νέο μοντέλο να διορθώνει λάθη που έγιναν από τα προηγούμενα. Τεχνικές όπως το XGBoost, LightGBM και CatBoost είναι δημοφιλείς λόγω της απόδοσης και της ταχύτητάς τους[65].

5.3 Support Vector Machines (SVMs)

Οι Μηχανές Διανυσματικής Υποστήριξης (SVMs) αποτελούν ένα ισχυρό σύνολο μεθόδων επιβλεπόμενης μάθησης που χρησιμοποιούνται για ταξινόμηση, παλινδρόμηση και ανίχνευση ακραίων τιμών. Οι Μηχανές Διανυσματικής Υποστήριξης (SVMs) αποτελούν ένα βασικό στοιχείο της μηχανικής μάθησης, προσφέροντας έναν αξιόπιστο τρόπο για την ταξινόμηση δεδομένων. Προερχόμενες από τη θεωρία στατιστικής μάθησης του Vapnik, οι SVMs έχουν εφαρμοστεί με επιτυχία σε μια ευρεία γκάμα προβλημάτων του πραγματικού κόσμου, από την κατηγοριοποίηση κειμένου έως τη βιοπληροφορική, αποδεικνύοντας την πολυπλοκότητα και την αποτελεσματικότητά τους.

5.3.1 Τι είναι ένα SVM?

Οι Μηχανές Διανυσματικής Υποστήριξης (SVMs) είναι μοντέλα επιβλεπόμενης μάθησης που αναλύουν δεδομένα για ταξινόμηση και ανάλυση παλινδρόμησης. Στον πυρήνα τους, οι SVMs στοχεύουν στην εύρεση του καλύτερου ορίου (ή υπερεπίπεδου) που χωρίζει τα σετ δεδομένων σε κλάσεις. Οι "διανύσματα υποστήριξης" είναι τα σημεία δεδομένων που βρίσκονται πιο κοντά στο υπερεπίπεδο και είναι κρίσιμα για τον ορισμό του περιθωρίου μεταξύ των διαφορετικών κλάσεων. Οι SVMs είναι ιδιαίτερα κατάλληλες για δεδομένα υψηλής διάστασης, όπου οι σχέσεις μεταξύ των κλάσεων δεν είναι εύκολα αναγνωρίσιμες[2].

5.3.2 Η προέλευση των SVM

Η έννοια των Μηχανών Διανυσματικής Υποστήριξης (SVMs) εισήχθη το 1963 από τον Vladimir Vapnik και τον Alexey Chervonenkis κατά τη διάρκεια της εργασίας τους στα εργαστήρια AT&T Bell. Ο αρχικός αλγόριθμος SVM σχεδιάστηκε για γραμμική ταξινόμηση. Μέσα στα χρόνια, ο αλγόριθμος έχει επεκταθεί ώστε να υποστηρίζει μη-γραμμική ταξινόμηση, καθιστώντας τον ευέλικτο για μια ευρεία γκάμα εφαρμογών. Αυτή η επέκταση διευκολύνθηκε σε μεγάλο βαθμό από την εισαγωγή του κόλπου του πυρήνα, που επιτρέπει στο SVM να ταξινομεί μη-γραμμικά διαχωρίσιμα δεδομένα με τον χάρτη τους σε έναν χώρο υψηλότερης διάστασης όπου μπορεί να βρεθεί ένας γραμμικός διαχωριστής[2].

5.3.3 Τρόπος λειτουργίας των SVM

Οι Μηχανές Διανυσματικής Υποστήριξης (SVMs) λειτουργούν εντοπίζοντας το υπερεπίπεδο που διαχωρίζει καλύτερα τις κλάσεις στον χώρο χαρακτηριστικών. Ο στόχος είναι να μεγιστοποιηθεί το περιθώριο μεταξύ των πλησιέστερων σημείων των κλάσεων, τα οποία είναι τα διανύσματα υποστήριξης. Αυτό επιτυγχάνεται μέσω τεχνικών βελτιστοποίησης που επιδιώκουν την ελαχιστοποίηση μιας συνάρτησης κόστους. Η διαδικασία περιλαμβάνει αρκετά κύρια στοιχεία[2]:

1. **Γραμμική Διαχωρισιμότητα:** Με απλούς όρους, οι SVMs προσπαθούν να σχεδιάσουν μια ευθεία γραμμή (σε δύο διαστάσεις) ή ένα υπερεπίπεδο (σε υψηλότερες διαστάσεις) για να διαχωρίσουν διαφορετικές κλάσεις στον χώρο. Αυτό είναι απλό όταν τα δεδομένα είναι γραμμικά διαχωρίσιμα, δηλαδή υπάρχει ένα σαφές κενό που μπορεί να σχεδιαστεί χωρίς κανένα σημείο διαφορετικής κλάσης να βρίσκεται μέσα σε αυτό.
2. **Το Κόλπο του Πυρήνα:** Για δεδομένα που δεν είναι γραμμικά διαχωρίσιμα, οι SVMs χρησιμοποιούν το κόλπο του πυρήνα για να μετασχηματίσουν τον χώρο εισόδου σε έναν χώρο υψηλότερης διάστασης όπου τα δεδομένα μπορεί να γίνουν γραμμικά διαχωρίσιμα. Κοινοί πυρήνες περιλαμβάνουν τον γραμμικό, πολυωνυμικό, συνάρτηση βάσης ακτινικής (RBF), και σιγμοειδή.
3. **Βελτιστοποίηση:** Ο αλγόριθμος SVM χρησιμοποιεί μαθηματική βελτιστοποίηση για να βρει το υπερεπίπεδο που μεγιστοποιεί το περιθώριο μεταξύ των κλάσεων. Αυτό περιλαμβάνει την επίλυση ενός κυρτού προβλήματος βελτιστοποίησης, το οποίο ευτυχώς εγγυάται ένα παγκόσμιο ελάχιστο - μια μοναδική καλύτερη λύση.

5.3.4 Εφαρμογές των SVM

Λόγω της ανθεκτικότητάς τους και της αποτελεσματικότητάς τους, οι SVMs έχουν εφαρμοστεί σε διάφορους τομείς[2]:

1. **Κατηγοριοποίηση Κειμένου και Υπερκειμένου:** Η ικανότητα των SVMs να χειρίζονται δεδομένα υψηλής διάστασης τα καθιστά ιδανικά για την ταξινόμηση κειμένου και υπερκειμένου για την ανάκτηση πληροφοριών και την κατηγοριοποίηση περιεχομένου ιστού.
2. **Κατηγοριοποίηση Εικόνας:** Οι SVMs έχουν χρησιμοποιηθεί σε εργασίες αναγνώρισης εικόνων, συχνά σε συνδυασμό με άλλες τεχνικές όπως τα νευρωνικά δίκτυα.
3. **Βιοπληροφορική:** Από την ταξινόμηση πρωτεϊνών έως την ταξινόμηση καρκίνου με βάση δεδομένα έκφρασης γονιδίων, οι SVMs διαδραματίζουν κρίσιμο ρόλο στη βιοπληροφορική για την ταξινόμηση περίπλοκων βιολογικών δεδομένων.
4. **Οικονομικά και Οικονομία:** Στον χρηματοοικονομικό τομέα, οι SVMs χρησιμοποιούνται για την πρόβλεψη τάσεων της αγοράς μετοχών, την πρόβλεψη πτώχευσης και την ανάλυση πιστοληπτικής αξιολόγησης.

5.4 Feature extraction

Αφού μια εικόνα έχει τμηματοποιηθεί σε περιοχές ή στα όριά τους με μεθόδους όπως εκείνες που περιγράφονται παραπάνω, τα προκύπτοντα σύνολα τμηματοποιημένων εικονοστοιχείων συνήθως πρέπει να μετατραπούν σε μορφή κατάλληλη για περαιτέρω επεξεργασία από τον υπολογιστή. Συνήθως, το βήμα μετά την τμηματοποίηση είναι η εξαγωγή χαρακτηριστικών, που αποτελείται από την ανίχνευση χαρακτηριστικών και την περιγραφή χαρακτηριστικών. Η ανίχνευση χαρακτηριστικών αναφέρεται στο εύρημα των χαρακτηριστικών σε μια εικόνα, περιοχή ή όριο. Η περιγραφή χαρακτηριστικών αποδίδει ποσοτικά χαρακτηριστικά στα ανιχνευμένα χαρακτηριστικά. Για παράδειγμα, μπορεί να ανιχνεύσουμε γωνίες σε ένα όριο περιοχής και να περιγράψουμε αυτές τις γωνίες με τον προσανατολισμό και την τοποθεσία τους, τα οποία είναι και τα δύο ποσοτικά χαρακτηριστικά. Οι μέθοδοι επεξεργασίας χαρακτηριστικών που συζητούνται σε αυτήν την ενότητα, χωρίζονται σε τρεις κύριες κατηγορίες, ανάλογα με το αν εφαρμόζονται σε όρια, περιοχές ή ολόκληρες εικόνες. Ορισμένα χαρακτηριστικά εφαρμόζονται σε περισσότερες από μία κατηγορίες. Οι περιγραφείς χαρακτηριστικών θα πρέπει να είναι όσο το δυνατόν λιγότερο ευαίσθητοι σε μεταβολές παραμέτρων όπως κλίμακα, μετάφραση, περιστροφή, φωτισμός και σημείο θέας. Οι περιγραφείς που συζητούνται σε αυτό το κεφάλαιο είτε είναι ανευαίσθητοι σε, είτε μπορούν να κανονικοποιηθούν για να αντισταθμίσουν, μεταβολές σε ένα ή περισσότερους από αυτούς τους παράγοντες[70].

Υπάρχουν δύο κύριες πτυχές της εξαγωγής χαρακτηριστικών εικόνας: η ανίχνευση χαρακτηριστικών και η περιγραφή τους. Δηλαδή, όταν αναφερόμαστε στην εξαγωγή

χαρακτηριστικών, αναφερόμαστε τόσο στην ανίχνευση των χαρακτηριστικών όσο και στη συνέχεια στην περιγραφή τους. Για να είναι χρήσιμη, η διαδικασία εξαγωγής πρέπει να περιλαμβάνει και τα δύο. Οι όροι που είναι πιθανό να συναντηθούν στην επεξεργασία και ανάλυση εικόνας για να περιγράψουν την ανίχνευση και περιγραφή χαρακτηριστικών διαφέρουν, αλλά ένα απλό παράδειγμα θα βοηθήσει να κατανοήσουμε τη χρήση αυτών των όρων. Έστω ότι χρησιμοποιούμε τις γωνίες αντικειμένων ως χαρακτηριστικά για κάποια εργασία επεξεργασίας εικόνας. Σε αυτή την ενότητα, η ανίχνευση αναφέρεται στον εντοπισμό των γωνιών σε μια περιοχή ή εικόνα. Η περιγραφή, από την άλλη πλευρά, αναφέρεται στην απόδοση ποσοτικών (ή μερικές φορές ποιοτικών) χαρακτηριστικών στα ανιχνευμένα χαρακτηριστικά, όπως ο προσανατολισμός της γωνίας και η τοποθεσία σε σχέση με άλλες γωνίες. Με άλλα λόγια, η γνώση ότι υπάρχουν γωνίες σε μια εικόνα έχει περιορισμένη χρήση χωρίς πρόσθετες πληροφορίες που μπορούν να μας βοηθήσουν να διακρίνουμε μεταξύ αντικειμένων σε μια εικόνα, ή μεταξύ εικόνων, με βάση τις γωνίες και τα χαρακτηριστικά τους[70].

Γενικά, τα χαρακτηριστικά θα πρέπει να είναι ανεξάρτητα από την τοποθεσία, την περιστροφή και την κλίμακα. Άλλοι παράγοντες, όπως η ανεξαρτησία από τα επίπεδα φωτισμού και οι αλλαγές που προκαλούνται από την οπτική γωνία μεταξύ του αισθητήρα απεικόνισης και της σκηνής, είναι επίσης σημαντικοί. Όποτε είναι δυνατόν, η προεπεξεργασία θα πρέπει να χρησιμοποιείται για να κανονικοποιήσει τις εισαγόμενες εικόνες πριν από την εξαγωγή χαρακτηριστικών. Για παράδειγμα, σε καταστάσεις όπου οι αλλαγές στο φωτισμό είναι αρκετά έντονες ώστε να προκαλέσουν δυσκολίες στην ανίχνευση χαρακτηριστικών, θα ήταν λογικό να προεπεξεργαστεί μια εικόνα για να αντισταθμίσει αυτές τις αλλαγές. Η ισοστάθμιση ιστογράμματος ή η διαδικασία προσδιορισμού είναι τεχνικές που έρχονται στο μυαλό ως αυτόματες μέθοδοι που γνωρίζουμε ότι είναι χρήσιμες σε αυτό το πλαίσιο[70]. Η ιδέα είναι να χρησιμοποιηθεί όσο το δυνατόν περισσότερη προκαταρκτική πληροφορία για να προεπεξεργαστούμε τις εικόνες με σκοπό να βελτιωθούν οι πιθανότητες ακριβούς εξαγωγής χαρακτηριστικών.

5.4.1 Feature extraction with Pyradiomics

Το Pyradiomics είναι ένα ανοιχτού κώδικα πακέτο Python που σχεδιάστηκε για την εξαγωγή χαρακτηριστικών ραδιομικής από ιατρικές εικόνες. Η ραδιομική είναι ένας τομέας που χρησιμοποιεί την ποσοτική ανάλυση εικόνων για την εξαγωγή ενός μεγάλου αριθμού χαρακτηριστικών από ιατρικές εικόνες χρησιμοποιώντας αλγορίθμους χαρακτηρισμού δεδομένων. Αυτά τα χαρακτηριστικά, που συχνά δεν είναι αντιληπτά από το ανθρώπινο μάτι, μπορούν να περιγράψουν την ετερογένεια του όγκου και έχουν πιθανές εφαρμογές στη διάγνωση, την πρόγνωση και την πρόβλεψη των αποτελεσμάτων της θεραπείας.

Το Pyradiomics μπορεί να εξάγει χαρακτηριστικά από διάφορες ιατρικές εικόνες (π.χ., CT, MRI, PET). Για την εξαγωγή, χρειάζονται τα δεδομένα απεικόνισης και ένα αντίστοιχο αρχείο μάσκας που διαγράφει την περιοχή ενδιαφέροντος (ROI) εντός της εικόνας, από την οποία θα εξαχθούν τα χαρακτηριστικά.

Ένας απλοποιημένος ενδεικτικός τρόπος εξαγωγής χαρακτηριστικών με την pyradiomics σε python είναι ο εξής:

```
python Copy code  
  
from radiomics import featureextractor  
import SimpleITK as sitk  
  
# Load your image and mask  
imagePath = 'path/to/image.nii'  
maskPath = 'path/to/mask.nii'  
image = sitk.ReadImage(imagePath)  
mask = sitk.ReadImage(maskPath)  
  
# Initialize feature extractor  
extractor = featureextractor.RadiomicsFeatureExtractor()  
  
# Optionally, load parameter file  
# extractor.loadParams('path/to/params.yaml')  
  
# Extract features  
result = extractor.execute(image, mask)  
  
# Print the extracted features  
for key, value in result.items():  
    print(f"{key}: {value}")
```

Εικόνα 5.4-1: Απλή χρήση της pyradiomics με python, για την διαδικασία της εξαγωγής χαρακτηριστικών

Το Pyradiomics επιτρέπει εκτενή προσαρμογή μέσω αρχείων παραμέτρων. Αυτά τα αρχεία YAML ή JSON μπορούν να καθορίσουν ποια χαρακτηριστικά να εξαχθούν, τους τύπους εικόνων που θα προκύψουν από την αρχική εικόνα (π.χ. εικόνες μετασχηματισμένες με wavelet) και τα βήματα προ-επεξεργασίας (π.χ. επαναδειγματοληψία, κανονικοποίηση).

5.4.2 Τύποι Χαρακτηριστικών που Εξάγονται

Το Pyradiomics μπορεί να εξάγει μια ευρεία γκάμα χαρακτηριστικών, περιλαμβανομένων αλλά όχι περιορισμένων στα:

Στατιστικά Πρώτης Τάξης: Περιγραφικά στατιστικά των εντάσεων των εικονοστοιχείων εντός της περιοχής ενδιαφέροντος (ROI).

Χαρακτηριστικά Βασισμένα στο Σχήμα: Περιγραφές του σχήματος και του μεγέθους της ROI.

Χαρακτηριστικά Υφής: Ποσοτικοποιούν την υφή αναλύοντας την κατανομή και τη σχέση των εντάσεων των εικονοστοιχείων (π.χ., Μήτρα Συν-εμφάνισης Επιπέδων Γκρι (GLCM), Μήτρα Διάρκειας Επιπέδων Γκρι (GLRLM)).

Χαρακτηριστικά Υψηλότερης Τάξης: Προέρχονται από φιλτραρισμένες ή μετασχηματισμένες εικόνες, αποτυπώνοντας επιπλέον πτυχές της υφής και των μοτίβων.

5.4.3 Μετεπεξεργασία και Ανάλυση

Τα εξαγόμενα χαρακτηριστικά μπορούν να αναλυθούν για την αναγνώριση προτύπων που συνδέονται με διαφορετικά κλινικά αποτελέσματα ή για τη δημιουργία προγνωστικών μοντέλων. Αυτό συνήθως περιλαμβάνει στατιστική ανάλυση ή μηχανική μάθηση και μπορεί να απαιτήσει περαιτέρω επιλογή χαρακτηριστικών ή μείωση της διαστασιμότητας για να επικεντρωθεί στα πιο πληροφοριακά χαρακτηριστικά.

ΜΕΡΟΣ 2: ΥΛΙΚΑ ΚΑΙ ΜΕΘΟΔΟΛΟΓΙΑ

6 Κεφάλαιο 6: Περιγραφή Dataset

6.1 Προεργασία

Το διαθέσιμο dataset αποτελούταν από 51 ασθενείς, με πέντε διαθέσιμες ακολουθίες ανά ασθενή, τις T1W, T2W, T1MDIXON, T2MDIXON και DWI. Από τους 51 ασθενείς, οι πρώτοι 27 είχαν εμφανή παθολογία της νόσου, ενώ οι υπόλοιποι δεν είχαν εμφανή συμπτώματα. Έχοντας υπ' όψιν πως κάποιες από τις ακολουθίες ήταν ελλιπείς, με κύριο παράδειγμα την T2MDIXON, το dataset χωρίστηκε, με 41 εικόνες να αποτελούν το training dataset και 10 να απομονώνονται για χρήση testing. Για τις 41 αυτές εικόνες επιλέχτηκε, για την χειροκίνητη χάραξη των ground truth μασκών της σπονδυλικής στήλης στο 3D Slicer, η ακολουθία T1W, στην οποία είναι πιο εμφανείς οι ανατομικές δομές.

6.2 Διαθέσιμα Λογισμικά

Το βασικό λογισμικό που χρησιμοποιήθηκε για τον χειρισμό και την επεξεργασία των εικόνων ήταν το 3D Slicer, ενώ για τον χειρισμό του dataset, έγινε χρήση της γλώσσας python και συγκεκριμένα των βιβλιοθηκών pytorch για την δημιουργία και εκπαίδευση των U-net, pyradiomics και scikit για την εξαγωγή χαρακτηριστικών και την εκπαίδευση του classifier, και SiTK για την διαχείριση (loading/saving/filtering/normalisation) των .nifti αρχείων.

7 Κεφάλαιο 7: Ανάλυση Pipeline

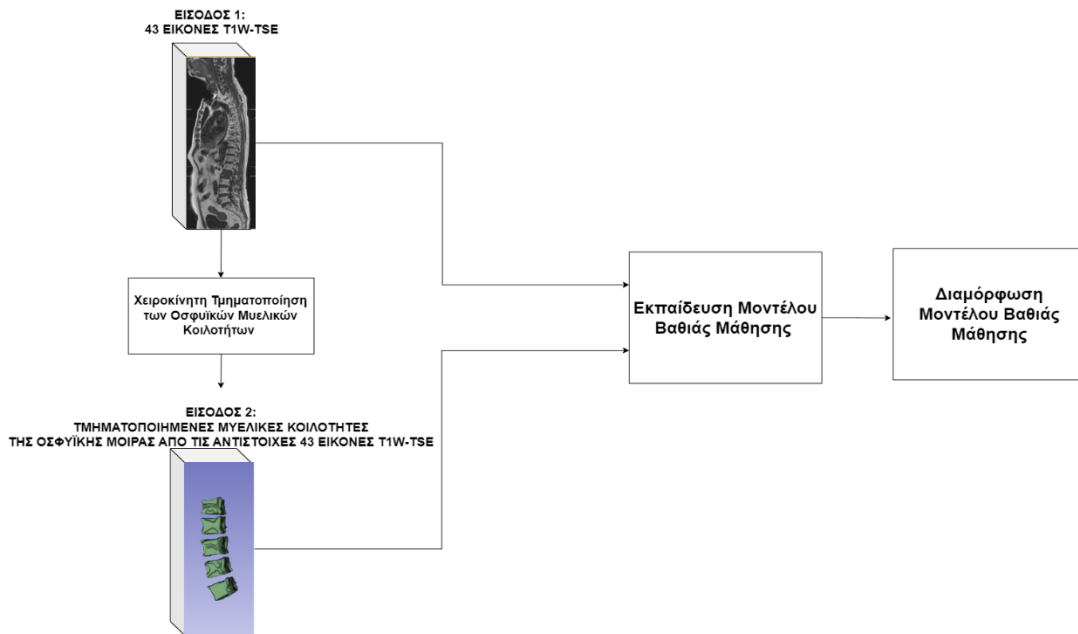
7.1 Εκπαίδευση U-net για την τμηματοποίηση των εικόνων

Στόχος υπήρξε η εκπαίδευση ενός U-net που από την whole-body MRI εικόνα, θα μπορούσε να απομονώσει με ικανοποιητική ακρίβεια την σπονδυλική στήλη ως δυαδική μάσκα.

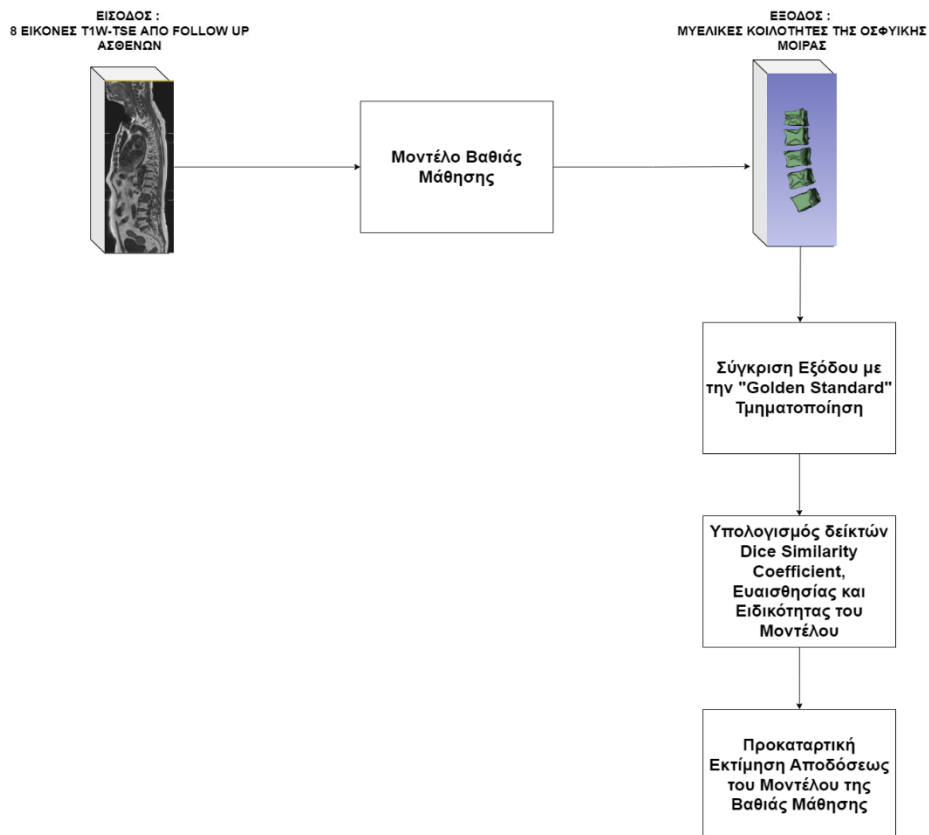
7.1.1 Προεπεξεργασία

Οι εικόνες και οι αντίστοιχες μάσκες τους ‘φορτώθηκαν’ με τη βοήθεια της κλάσης `NiftiDatasetTransforms` της βιβλιοθήκης `pytorch`. Στην αρχή κανονικοποιήθηκαν μέσω της συνάρτησης `tie.ZNormalisation`. Στη συνέχεια πραγματοποιήθηκε `rescale` των εικόνων στο μέγεθος (256, 512, 32), καθώς οι διαστάσεις αυτές ήταν οι καταλληλότερες διαστάσεις της μορφής 2^y που διατηρούσαν την αναλογία των αρχικών διαστάσεων των εικόνων, προς αποφυγή παραμορφώσεων. Έπειτα, για να αντιμετωπιστεί το πρόβλημα του λιγοστού dataset, εφαρμόστηκαν οι τεχνικές `data augmentation`. Οι τεχνικές αυτές περιελάμβαναν `Random Elastic Deformation` και `Random Flip` γύρω από τον άξονα 0. Τέλος το dataset εικόνων-μασκών χωρίστηκε στη μορφή (0.7, 0.2, 0.1) για `training dataset`, `validation dataset` και `test dataset` αντίστοιχα.

ΣΤΑΔΙΟ 1: ΕΚΠΑΙΔΕΥΣΗ ΜΟΝΤΕΛΟΥ ΒΑΘΙΑΣ ΜΑΘΗΣΗΣ



ΣΤΑΔΙΟ 2: ΔΟΚΙΜΗ ΜΟΝΤΕΛΟΥ ΒΑΘΙΑΣ ΜΑΘΗΣΗΣ



Εικόνα 7.1-1: Δομή εκπαίδευσης U-net με σκοπό την παραγωγή δυαδικής μάσκας της σπονδυλικής στήλης από whole-body MRI εικόνες

7.1.2 Επιλογή του U-net

Table 7.1-1: Πίνακας συνδυασμών αρχιτεκτονικών U-net

200 epochs training	REC,t =2,3	ATT	RES	REC, +ATT t=2,3	REC, =2,3+RES t	ATT+RES	REC, 2,3+ATT+RES t =	DEN
DEN	2	1	1	2	2	1	2	1
INCEPTION	2	1	1	2	2	1	2	1
MAIN	2	1	1	2	2	1	2	1
total runs	36							

Τα τρεξίματα οργανώθηκαν ως εξής:

Διαθέσιμα είναι τρία διαφορετικά blocks, όπως φαίνεται στον παραπάνω πίνακα. Το βασικό main block είναι μια κλασική αρχιτεκτονική 3D U-net με διπλές συνελίξεις, οι παράμετροι του οποίου ορίζονται από ένα configuration file.

Το inception block είναι μια αρχιτεκτονική με μονές συνελίξεις, η μια με πυρήνα 3*3 και η δεύτερη με πυρήνα 7*7. Μετά την επεξεργασία γίνεται concatenation του αποτελέσματος, το οποίο έπειτα διέρχεται μέσω μιας μονής συνέλιξης με 1*1 πυρήνα, φτάνοντας έτσι στις διαστάσεις που θα είχε η εικόνα μετά από απλή διπλή συνέλιξη.

Το dense block ακολουθεί πιστά την αρχιτεκτονική που περιγράφεται στο θεωρητικό υπόβαθρο.

Για κάθε μία από τις τρεις αρχιτεκτονικές, μπορούν να γίνουν οι εξής παραλλαγές:

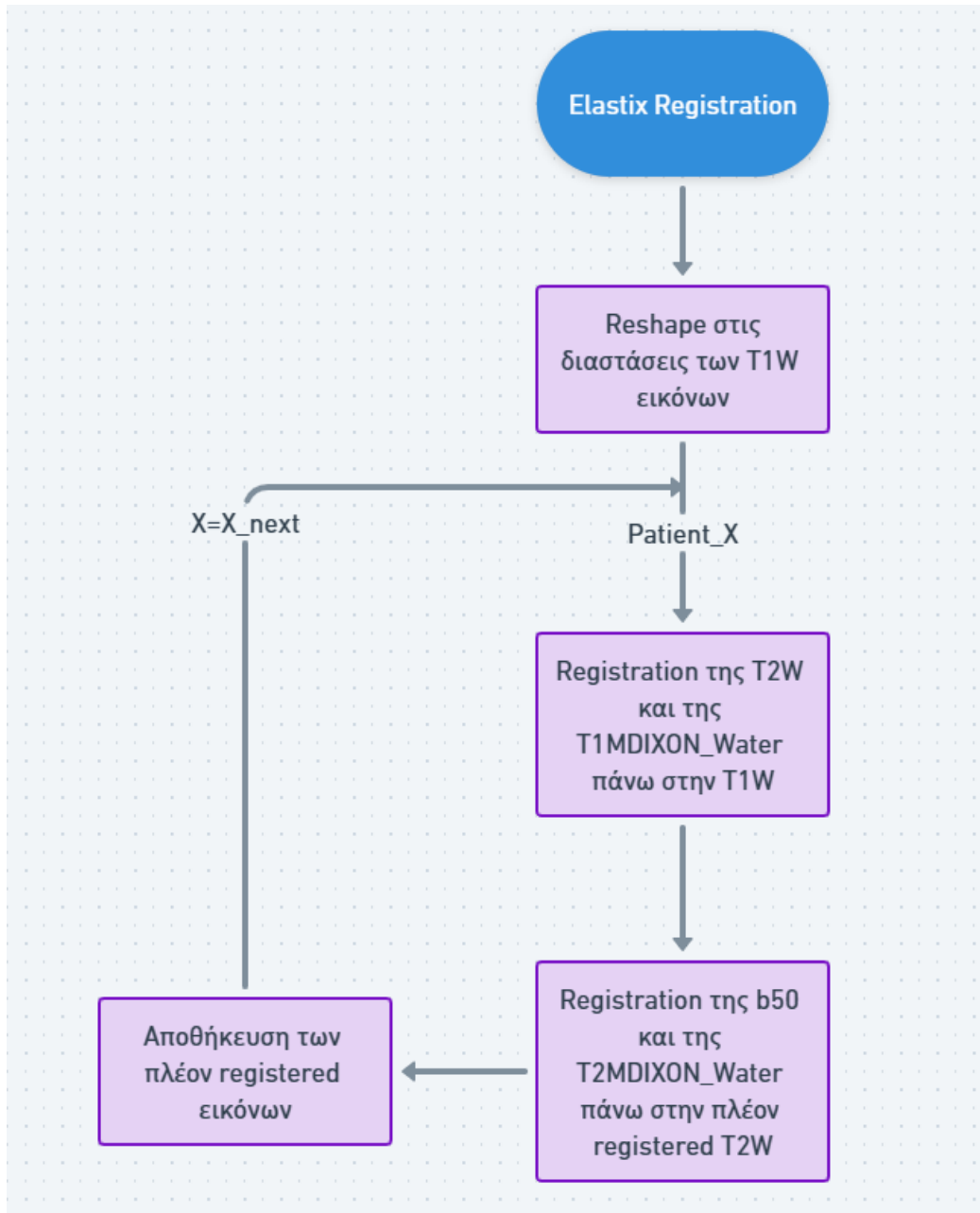
1. Δύναται να προστεθούν recurrent connections, σύμφωνα με τα οποία η έξοδος του κάθε νευρώνα εισέρχεται ξανά ως είσοδος είτε δύο είτε τρεις φορές (t=2 ή t=3).
2. Δύναται να προστεθούν attention connections
3. Δύναται να προστεθούν residual connections

Ως εκ τούτου, υπάρχουν τέσσερις δυνατοί συνδυασμοί ανά block, ήτοι τριάντα έξι συνολικοί συνδυασμοί. Από αυτούς, επιλέχθηκε ο αποδοτικότερος συνδυασμός για την απομόνωση της σπονδυλικής στήλης από τις εικόνες Whole-Body MRI.

7.2 Μεθοδολογία Registration

Η διαδικασία ξεκίνησε με το reshape των εικόνων ώστε να ταιριάζουν στις διαστάσεις των T1W εικόνων, δημιουργώντας μια ενιαία βάση για το επόμενο στάδιο. Στη συνέχεια, χρησιμοποιήθηκε το εργαλείο elastix για την πραγματοποίηση του registration. Καθώς οι σειρές DIXON και DWI περιλαμβάνουν πολλά κανάλια (fat, water, ip, op για τις DIXON και

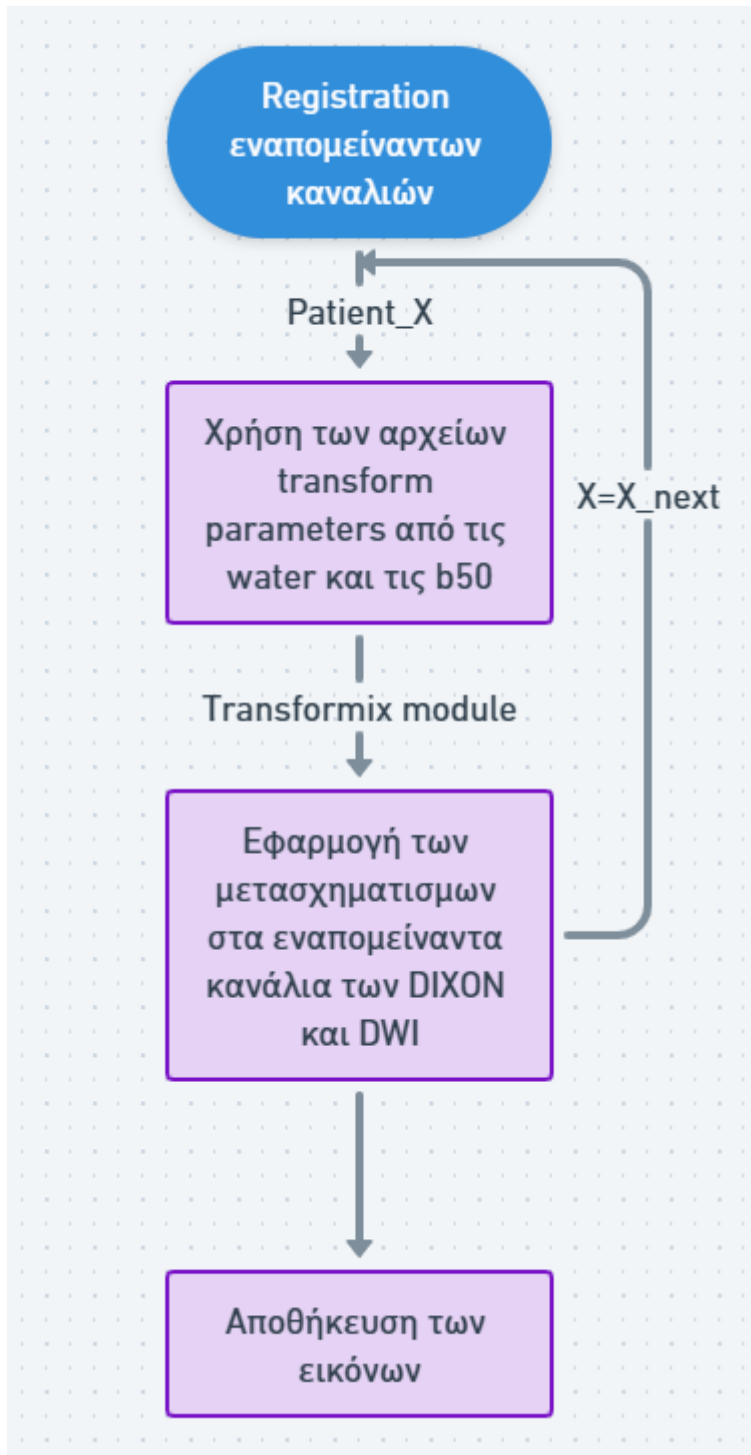
b50, b400, b900 για τις DWI), αποφασίστηκε να χρησιμοποιηθεί ένα κανάλι από την κάθε ακολουθία για την διαδικασία της ευθυγράμμισης με το elastix, συγκεκριμένα το κανάλι water για τις DIXON και το κανάλι b50 για τις DWI, λόγω του υψηλότερου SNR. Έτσι, πρώτα ευθυγραμμίζονταν οι T2W και οι T1MDIXON πάνω στην T1W, λόγω των περισσότερων κοινών τους χαρακτηριστικών και έπειτα οι b50 και οι T2MDIXON πάνω στις πλέον ευθυγραμμισμένες T2W, πάλι λόγω των περισσότερων κοινών χαρακτηριστικών.



Εικόνα 7.2-1: Διαδικασία registration με το elastix

Για τα υπόλοιπα κανάλια, η διαδικασία registration πραγματοποιήθηκε με τη χρήση του module transformix, χρησιμοποιώντας τα κατάλληλα αρχεία παραμέτρων που προέκυψαν από την αρχική ευθυγράμμιση με το elastix. Αυτή η προσέγγιση επιτρέπει την εφαρμογή των υπολογισμένων μετασχηματισμών σε όλα τα κανάλια χωρίς την ανάγκη επανάληψης του

πλήρους υπολογισμού του registration, εξασφαλίζοντας ομοιογένεια και εξοικονόμηση χρόνου.



Εικόνα 7.2-2: Χρήση transformix για το registration των εναπομείναντων καναλιών

7.2.1 Αξιολόγηση και επιλογή τεχνικής

Πριν την επιλογή του elastix, δοκιμάστηκαν πολλές τεχνικές registration, συμπεριλαμβανομένων του ANTs registration και της συνδυαστικής χρήσης rigid και elastic μεθόδων. Η τελική επιλογή του elastix στηρίχθηκε στην αποδοτικότητα, την ακρίβεια και την

ικανότητα της μεθόδου να χειρίζεται πολυδιάστατα δεδομένα με επιτυχία, καθώς και την ευκολία εφαρμογής της σε διάφορα κανάλια με την χρήση του transformix για τον μετασχηματισμό των υπολοίπων καναλιών.

Όπως προαναφέρθηκε, το elastix παρέχει μεγάλη ευελιξία στην επιλογή παραμέτρων σύμφωνα με τις οποίες μπορεί να γίνει η εγγραφή των εικόνων. Η επιλογή αυτή πραγματοποιείται με την κατάλληλη επιλογή και τον συνδυασμό των αρχείων παραμέτρων (parameter files) που συμπεριλαμβάνονται στο module του elastix. Σύμφωνα με την μελέτη [71], τα αρχεία παραμέτρων που συνδυάστηκαν ήταν σε σειρά το Rigid.txt και το Deformable.txt. Σύμφωνα με αυτά, ορίζονται τα εξής:

Η στρατηγική εγγραφής, η MultiResolutionRegistration, υποδηλώνει ότι η διαδικασία εγγραφής θα συμβεί σε πολλαπλές αναλύσεις για αποδοτικότητα και ακρίβεια. Οι μέθοδοι παρεμβολής (BSplineInterpolator και FinalBSplineInterpolator) καθορίζουν πώς υπολογίζονται οι τιμές των εικονοστοιχείων κατά τη διάρκεια της μετατροπής και της επαναδειγματοληψίας. Η παρεμβολή B-spline χρησιμοποιείται για την ομαλότητά της και την υπολογιστική αποδοτικότητα. Ορίζεται ο DefaultResampler για να επαναδειγματοληφτείται η κινούμενη εικόνα μόλις βρεθεί ο βέλτιστος μετασχηματισμός. Οι FixedSmoothingImagePyramid και MovingSmoothingImagePyramid μειώνουν την ανάλυση των εικόνων με βηματικό τρόπο για την εγγραφή πολλαπλών αναλύσεων. Αυτή η προσέγγιση βοηθά στη διαχείριση των υπολογιστικών πόρων και βελτιώνει την ανθεκτικότητα της εγγραφής.

Το EulerTransform υποδηλώνει ένα σταθερό μοντέλο μετατροπής που περιλαμβάνει περιστροφή και μεταφορά, αλλά όχι κλιμάκωση ή παραμόρφωση. Είναι κατάλληλο για εφαρμογές όπου το σχήμα του αντικειμένου δεν αλλάζει. Η Αυτόματη Εκτίμηση Κλιμάκων ενεργοποιεί την αυτόματη εκτίμηση κλιμάκων για τις παραμέτρους του μετατροπής, βελτιώνοντας τη σύγκλιση της βελτιστοποίησης. Η Αυτόματη Αρχική Μετατροπή εκτιμά αυτόματα μια αρχική μετατροπή, διευκολύνοντας ένα καλό σημείο εκκίνησης για την βελτιστοποίηση. Το Compose υποδεικνύει ότι οι μετατροπές σε κάθε ανάλυση θα συνδυαστούν μαζί για να σχηματίσουν την τελική μετατροπή. Η μετρική που χρησιμοποιήθηκε στον Rigid μετασχηματισμό είναι η AdvancedMattesMutualInformation, μια μετρική ομοιότητας που βασίζεται στην αμοιβαία πληροφορία, αποτελεσματική για την πολυ-τροπική εγγραφή εικόνων όπου η σχέση εντάσεων μεταξύ των εικόνων είναι μη γραμμική, ενώ στον Deformable χρησιμοποιήθηκε επιπλέον η TransformBendingEnergyPenalty, ο στόχος της οποίας δεν είναι μόνο να ευθυγραμμίσει δύο εικόνες βρίσκοντας τον βέλτιστο χωρικό μετασχηματισμό, αλλά και να διασφαλίσει ότι ο μετασχηματισμός είναι φυσικά πιθανός. Αυτή η απαίτηση είναι το σημείο όπου οι στρατηγικές κανονικοποίησης, συμπεριλαμβανομένης της ποινής ενέργειας κάμψης, έρχονται σε εφαρμογή. Η ποινή ενέργειας κάμψης είναι μια μορφή κανονικοποίησης που διασφαλίζει την ομαλότητα στο αποτέλεσμα του πεδίου παραμόρφωσης, αποτρέποντας την ανεργμάνιση κάμψης ή δίπλωμα της εικόνας που εγγράφεται.

Για την ακρίβεια, η ποινή ενέργειας κάμψης χρησιμοποιείται για να κανονικοποιήσει τη διαδικασία παραμόρφωσης, διασφαλίζοντας ότι οι μετασχηματισμοί που εφαρμόζονται στην

κινούμενη εικόνα δεν οδηγούν σε φυσικά αδύνατες ή ιδιαίτερα απίθανες μορφές. Επιβάλλει ποινή σε υπερβολική κάμψη, ενθαρρύνοντας έτσι ομαλότερες και πιο ρεαλιστικές παραμορφώσεις.

Η πολυανάλυση, `NumberOfResolutions`, υποδηλώνει ότι η εγγραφή εκτελείται σε τρεις αναλύσεις, ενώ η `ImagePyramidSchedule` ορίζει τους παράγοντες μείωσης δειγματοληψίας για κάθε επίπεδο, μειώνοντας το μέγεθος της εικόνας για να επιταχύνει τη διαδικασία σε χονδρικότερες αναλύσεις. Το μέγεθος των πυρήνων για κάθε βαθμίδα της πυραμίδας ήταν:

- Πυρήνας_1: (8,8,2)
- Πυρήνας_2: (4,4,1)
- Πυρήνας_3: (1,1,1)

Ως βελτιστοποιητής (Optimizer), επιλέγεται ο `AdaptiveStochasticGradientDescent` για τον εντοπισμό των βέλτιστων παραμέτρων του μοντέλου μετατροπής. Είναι ανθεκτικός και αποδοτικός για μια ευρεία γκάμα εργασιών εγγραφής. Τα `MaximumNumberOfIterations` και `MaximumStepLength` ελέγχουν τη διάρκεια της διαδικασίας βελτιστοποίησης και τα μεγέθη των βημάτων, αντίστοιχα, ενώ το `RequiredRatioOfValidSamples` ορίζει το κλάσμα των δειγμάτων που πρέπει να συμβάλλουν στον υπολογισμό της συνάρτησης κόστους, επηρεάζοντας την ανθεκτικότητα και την ταχύτητα της βελτιστοποίησης.

7.3 Μεθοδολογία Απομόνωσης Βλαβών

Διάφορες προσεγγίσεις για την απομόνωση των βλαβών και τη δημιουργία ground truth μασκών υιοθετήθηκαν βάσει των χαρακτηριστικών των εικόνων και των δεικτών που προτάθηκαν από ακτινολόγους:

7.3.1 Χρήση εικόνων b900

Η πρώτη προσέγγιση εστίασε στην επιλογή των φωτεινών σημείων στις εικόνες b900, καθώς οι βλάβες συχνά φαίνονται πιο φωτεινές σε αυτές τις λήψεις. Η DWI είναι ιδιαίτερα ευαίσθητη σε αλλαγές στο μικροπεριβάλλον των μορίων νερού, καθιστώντας την ισχυρό εργαλείο στην ανίχνευση και χαρακτηρισμό βλαβών σε διάφορα μέρη του σώματος. Τα "φωτεινά μέρη" των εικόνων DWI αναφέρονται σε περιοχές όπου υπάρχει περιορισμένη διάχυση, συχνά υποδηλώνοντας παθολογικές καταστάσεις όπως όγκοι, φλεγμονές ή έμφραξη. Για την πραγματοποίηση αυτής της μεθόδου, η κάθε εικόνα b900 'φορτώνεται' και φιλτράρεται από ένα `tv_cambolle` φίλτρο με `weight = 5`, με σκοπό να αποθρονοποιηθεί χωρίς να υπάρξει σημαντική απώλεια πληροφορίας. Στη συνέχεια γίνεται `resampling` στην κλίμακα (0-1), ώστε όλες οι εικόνες να υποστούν επεξεργασία υπό την ίδια κλίμακα. Έπειτα, η εικόνα περνά από ένα `Otsu Multiple Threshold Filter` με όρισμα δύο κατώφλια, ώστε να χωριστούν κατάλληλα οι φωτεινότητες. Ως μάσκα αποθηκεύεται η εικόνα που αντιστοιχεί στο μεγαλύτερο κατώφλι.

7.3.2 Συνδυασμός T1W και T2W

Η δεύτερη προσέγγιση αφορούσε την επιλογή σημείων που είναι ταυτόχρονα σκοτεινά στις T1W και φωτεινά στις T2W, εκμεταλλευόμενη τις αντίθετες αντιδράσεις των βλαβών σε αυτές τις λήψεις. Οι εικόνες T1W και T2W παρέχουν συμπληρωματικές πληροφορίες λόγω των διαφορετικών ευαισθησιών τους στις φυσικές και χημικές ιδιότητες των ιστών, ιδιαίτερα όσον αφορά το περιεχόμενο σε νερό και λίπος. Στις εικόνες T1W, οι ιστοί με σύντομο χρόνο χαλάρωσης T1 φαίνονται φωτεινοί. Το λίπος είναι συνήθως φωτεινό στις εικόνες T1W, ενώ οι δομές που περιέχουν νερό φαίνονται σκούρες. Οι εικόνες T2W τονίζουν τους ιστούς με μακρύ χρόνο χαλάρωσης T2, κάνοντάς τους να φαίνονται φωτεινοί. Δομές γεμάτες με υγρά και περιοχές με υψηλό περιεχόμενο σε νερό, όπως το οίδημα ή η φλεγμονή, φαίνονται φωτεινές στις εικόνες T2W.

Στην δεύτερη αυτή μέθοδο, ακολουθήθηκε η ίδια διαδικασία με την πρώτη, αρχικά για τις T1W εικόνες και μετά για τις T2W εικόνες. Ως τελικό αποτέλεσμα πάρθηκε η τομή των μασκών για κάθε ασθενή.

7.3.3 Χρήση ADC χάρτη

Η τρίτη προσέγγιση βασίστηκε στη χρήση του ADC χάρτη, ο οποίος υπολογίζεται από τον συνδυασμό των b50 και b900 εικόνων, μέσω του τύπου 2.2-5 :

όπου $S_{b1} \rightarrow$ η b900 εικόνα,

$S_{b0} \rightarrow$ η b50 εικόνα,

$b_1 = 900,$

$b_2 = 50$

Για την τελική τμηματοποίηση του χάρτη χρησιμοποιήθηκε το εύρος τιμών μεταξύ $0.6-1.4 * 10^{-3} \text{ sec/mm}^2$, μια τιμή που τυπικά συνδέεται με τις βλάβες. Η διαδικασία απομόνωσης βλαβών χρησιμοποιώντας χάρτες ADC περιλαμβάνει τη σύγκριση τιμών ADC διαφορετικών περιοχών εντός του πεδίου απεικόνισης. Βλάβες, όπως αυτές από πολλαπλό μυέλωμα, συχνά εμφανίζουν περιορισμένη διάχυση σε σύγκριση με τον περιβάλλοντα κανονικό ιστό, οδηγώντας σε χαμηλότερες τιμές ADC. Οι μυελωματικές βλάβες συνήθως εμφανίζουν περιορισμένη διάχυση (χαμηλές τιμές ADC) λόγω της υψηλής κυτταρικής πυκνότητας των κακοήθων κυττάρων πλάσματος.

Στην περίπτωση αυτή, για την παραγωγή της μάσκας χρησιμοποιήθηκαν οι εικόνες b50 και b900, καθώς και οι μάσκες που παρήχθησαν από την τεχνική 1. Μετά την αποθορυβοποίησή τους όπως περιγράφηκε νωρίτερα, υπολογίστηκε ο χάρτης ADC σύμφωνα με τον τύπο 2.2-5,

Στη συνέχεια η τελική μάσκα προέκυψε από την απομόνωση των βόξελ που ανήκουν στο επιθυμητό εύρος τιμών που και ταυτόχρονα στη μάσκα της μεθόδου 1.

7.3.4 Δείκτης Fat Fraction

Η τελευταία προσέγγιση χρησιμοποίησε τον δείκτη fat fraction των εικόνων, υπολογιζόμενο από τον τύπο:

$$FF = \frac{(IP - OP)}{2IP}$$

Equation 7.3-1: Εξίσωση Fat Fraction

όπου

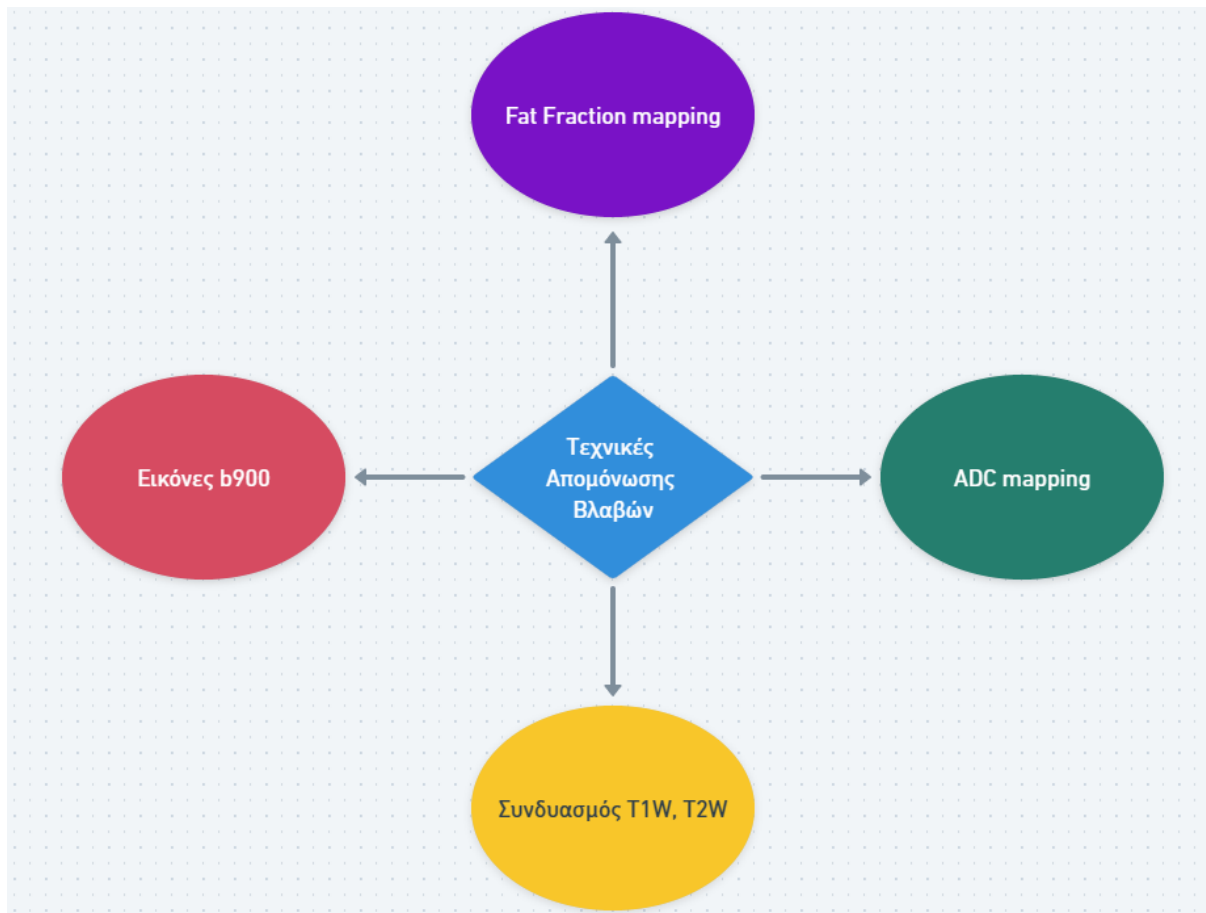
IP → η T1MDIXON IP εικόνα

OP → η T1MDIXON OP εικόνα

προσπαθώντας να εκμεταλλευτεί τις διαφορές στη σύνθεση του λίπους στις βλάβες.

Στην τελευταία μέθοδο που δοκιμάστηκε, αξιοποιήθηκαν οι εικόνες T1MDIXON_IP και T1MDIXON_OP. Μετά την αποθρομβοποίησή τους υπολογίστηκε ο χάρτης fat fraction σύμφωνα με τον τύπο 7.3-1,

και στη συνέχεια η τελική μάσκα προέκυψε από την απομόνωση των βόξελ που βρίσκονται πάνω από το επιθυμητό κατώφλι



Εικόνα 7.3-1: Διάγραμμα τεχνικών απομόνωσης βλαβών

7.3.5 Τεχνικές Τμηματοποίησης

Για την απομόνωση των βλαβών, δοκιμάστηκαν οι τεχνικές τμηματοποίησης, k-means clustering και Otsu's multiple thresholding. Αμφότερες απέδωσαν εξίσου καλά, η μέθοδος Otsu όμως υπήρξε οικονομικότερη χρονικά και για αυτόν το λόγο επιλέχθηκε για την τμηματοποίηση στις περιπτώσεις 7.3.1 και 7.3.2.

7.4 Διαδικασία Classification

Για την διαδικασία του classification ακολουθήθηκαν δύο διαφορετικές προσεγγίσεις, οι οποίες μπορούν να αντιμετωπιστούν ανεξάρτητα, αλλά αποδίδουν καλύτερα αν χρησιμοποιηθούν συμπληρωματικά. Η πρώτη προσέγγιση αφορά στην εκπαίδευση ενός μοντέλου μηχανικής μάθησης, συγκεκριμένα ενός Support Vector Machine (SVM), για την κατηγοριοποίηση της σπονδυλικής στήλης σε δύο κατηγορίες: με μυέλωμα (label = 1) ή χωρίς μυέλωμα (label = 0). Η προσέγγιση αυτή είναι κρίσιμη για την αυτόματη διάγνωση και την εκτίμηση της παρουσίας του μυελώματος με βάση τις MRI εικόνες. Η δεύτερη διαδικασία αποτελεί μια επέκταση της πρώτης, αφού έχει ως στόχο την εκπαίδευση ενός SVM μοντέλου ανά σπόνδυλο. Η διαδικασία ενσωματώνει μια πιο λεπτομερή ανάλυση, δίνοντας έμφαση στην εκτίμηση της παρουσίας και του τύπου του μυελώματος σε κάθε σπόνδυλο. Η βασική διαφορά με την ανάλυση ολόκληρης της σπονδυλικής, είναι η διαδικασία απομόνωσης των σπονδύλων και η κατηγοριοποίηση των βλαβών ως εξής:

- 0: Όχι μυέλωμα
- 1: Focal
- 2: Variegated
- 3: Diffuse

Συγκεκριμένα, η διαδικασία ξεκίνησε με την ανάλυση κάθε σπονδύλου ξεχωριστά, όπου τόσο οι MRI εικόνες όσο και οι αντίστοιχες μάσκες χωρίστηκαν ανά σπόνδυλο (από C2 έως L5) και ονοματίστηκαν καταλλήλως. Αυτό επέτρεψε την ακριβή εστίαση σε κάθε σπόνδυλο για την εξαγωγή χαρακτηριστικών και την εκπαίδευση του μοντέλου.

7.4.1 Δυαδικός Ταξινομητής

7.4.1.1 Ετοιμασία Δεδομένων

Χρησιμοποιήθηκαν εικόνες από όλες τις διαθέσιμες ακολουθίες MRI. Από τις whole-body mri έπρεπε να απομονωθεί μόνο το τμήμα της σπονδυλικής στήλης και για τον σκοπό αυτό αξιοποιήθηκαν τα αποτελέσματα του μοντέλου Unet. Μέσω της συνάρτησης `sitk.Mask(image, mask)`, όπου στο όρισμα `image` βρισκόταν η WHMRI εικόνα και στο όρισμα `mask` η δυαδική μάσκα της σπονδυλικής που παρήγαγε το εκπαιδευμένο U-net, απομονώθηκε το επιθυμητό τμήμα της κάθε εικόνας.

Για τη διαδικασία εξαγωγής χαρακτηριστικών αποφασίστηκε να αξιοποιηθούν μαζί με τις αντίστοιχες εικόνες, οι μάσκες τόσο βλαβών όσο και ολόκληρης της σπονδυλικής στήλης, ώστε ο ταξινομητής να εκπαιδευτεί και σε contextual δεδομένα. Μέσω της χρήσης της pyradiomics, πραγματοποιήθηκε εξαγωγή χαρακτηριστικών (feature extraction) για κάθε εικόνα.

7.4.1.2 Δημιουργία Πίνακα Χαρακτηριστικών

Οι πληροφορίες από την εξαγωγή χαρακτηριστικών συγχωνεύθηκαν σε έναν τρισδιάστατο πίνακα, όπου οι διαστάσεις αντιπροσωπεύουν τους ασθενείς, τις ακολουθίες και τα συνολικά χαρακτηριστικά. Ακολούθησε συγχώνευση σε δύο διαστάσεις (ασθενείς/χαρακτηριστικά) και κανονικοποίηση των χαρακτηριστικών στο διάστημα (0-1).

7.4.1.3 Εκπαίδευση SVM Μοντέλου

Η εκπαίδευση του SVM μοντέλου περιελάμβανε τα εξής βήματα: Προετοιμασία και Επιλογή Χαρακτηριστικών: Το dataset χωρίστηκε ξανά σε training και testing dataset σε αναλογία (0.8, 0.2). Έγινε χρήση της τεχνικής imputation μέσης τιμής για την αντιμετώπιση των απουσιαζουσών τιμών, καθώς δεν υπήρχαν διαθέσιμες εικόνες σε όλες τις ακολουθίες για όλους τους ασθενείς. Ακολούθησε κανονικοποίηση, επιλογή χαρακτηριστικών με βάση την διακύμανση και εφαρμογή κριτηρίων για την επιλογή των καλύτερων χαρακτηριστικών. Πιο συγκεκριμένα απορρίπτονται τα χαρακτηριστικά με διακύμανση = 0, καθώς είναι κοινά και δεν προσφέρουν χρήσιμη πληροφορία στον classifier.

Ορισμός και Εφαρμογή του SVM Classifier: Χρήση του SVC της sklearn με παραμέτρους όπως το C, το kernel και το gamma. Συγκεκριμένα:

```
# Define the SVM and grid search parameters
```

```
svm_classifier = SVC()
```

```
param_grid = {
```

```
'svm__C': [0.1, 1, 10, 100],
```

```
'svm__kernel': ['linear', 'rbf', 'poly', 'sigmoid'],
```

```
'svm__gamma': ['scale', 'auto']
```

```
}
```

Χρήση Grid Search για την Εύρεση των Καλύτερων Παραμέτρων: Αυτό περιελάμβανε την εκπαίδευση του μοντέλου με διάφορες παραμέτρους και την επιλογή του συνδυασμού που προσέφερε την καλύτερη απόδοση.

7.4.2 Ταξινομητής ανά σπόνδυλο

Η λογική των ταξινομητών ανά σπόνδυλο, είναι πανομοιότυπη με αυτή του δυαδικού ταξινομητή. Η διαφορά τους έγκειται στο γεγονός ότι, ενώ ο δυαδικός ταξινομητής χρησιμοποιεί ολόκληρη τη σπονδυλική, στην προέκταση των ανά σπόνδυλο ταξινομητών εκπαιδεύεται ένας ταξινομητής για κάθε σπόνδυλο. Επομένως το μέρος της διαδικασίας που αλλάζει εντοπίζεται στην ετοιμασία των δεδομένων.

7.4.2.1 Ετοιμασία δεδομένων

Η διαφοροποίηση ξεκινά μετά το βήμα της απομόνωσης της σπονδυλικής από τις whole-body εικόνες. Ακολουθείται διαδικασία component analysis για τη σπονδυλική στήλη και για τις αντίστοιχες μάσκες, ώστε να χωριστούν και να ονοματιστούν οι σπόνδυλοι. Έπειτα ακολουθούνται οι τεχνικές εντοπισμού βλαβών για κάθε σπόνδυλο και άρα η δημιουργία της αντίστοιχης μάσκας βλαβών. Τέλος, η εικόνα σπονδύλου, η μάσκα σπονδύλου και η μάσκα βλαβών σπονδύλου χρησιμοποιούνται για την εκπαίδευση του ταξινομητή σπονδύλου. Η διαδικασία επαναλαμβάνεται για κάθε σπόνδυλο, έως ότου εκπαιδευτούν 24 ταξινομητές, ένας για κάθε σπόνδυλο.

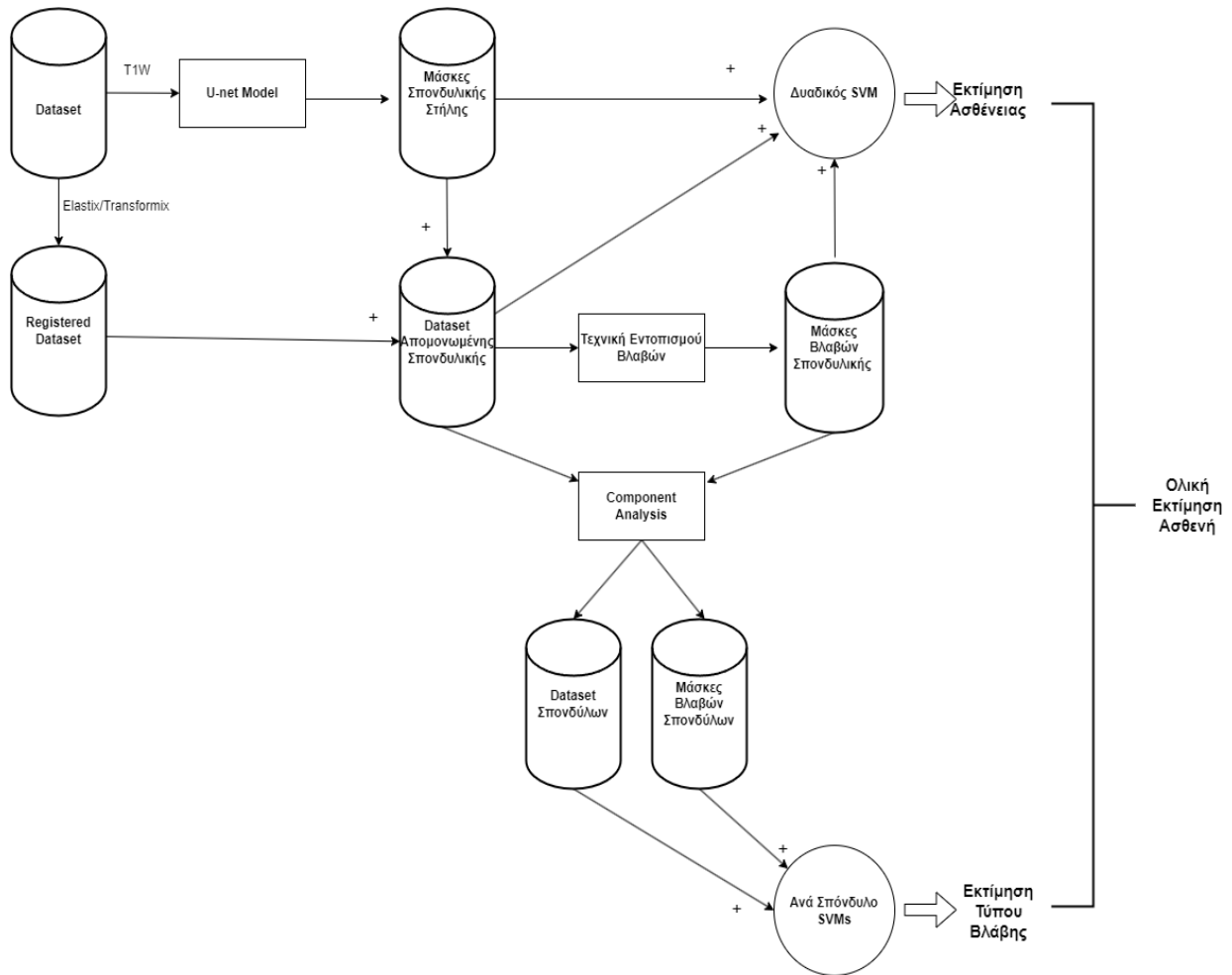
7.4.3 Αξιολόγηση

Η αποτελεσματικότητα του τελικού μοντέλου αξιολογήθηκε με βάση την ακρίβεια, την ανάκληση και την F1 σκορ, παρέχοντας μια σφαιρική εικόνα της απόδοσης του μοντέλου στην κατηγοριοποίηση των σπονδυλικών στηλών. Η διαδικασία αυτή είναι κρίσιμη για την αξιολόγηση της ικανότητας του μοντέλου να διαχωρίζει με ακρίβεια τις περιπτώσεις με μύελωμα από εκείνες χωρίς, συμβάλλοντας σημαντικά στην πρόωμη διάγνωση και στην εκτίμηση της ανάγκης για περαιτέρω διαγνωστικές εξετάσεις ή θεραπείες. Όσον αφορά την ανάλυση ανά σπόνδυλο, εκεί παρουσιάστηκαν περαιτέρω δυσκολίες λόγω της φύσης των διαθέσιμων δεδομένων. Η συντριπτική πλειοψηφία των δεδομένων να ανήκει στην κατηγορία 1 (Focal) προκαλεί έναν σημαντικό προβληματισμό στην εκπαίδευση του μοντέλου, καθώς μπορεί να οδηγήσει σε μια προκατάληψη του μοντέλου προς την πλειοψηφική κλάση. Αυτό συχνά συμβαίνει σε περιπτώσεις ανισορροπίας δεδομένων και μπορεί να επηρεάσει την ικανότητα του μοντέλου να αναγνωρίζει σωστά τις λιγότερο συχνές κατηγορίες.

7.5 Σύνοψη και ιδανική λειτουργία της διαδικασίας

Για την σωστή λειτουργία ολόκληρης της διαδικασίας τα βήματα που αναλύθηκαν σε αυτό το κεφάλαιο πρέπει να συνδεθούν κατάλληλα ως εξής: Αρχικά, συλλογή του dataset και κανονικοποίηση των εικόνων όλων των ακολουθιών στο διάστημα (0-1) καθώς και αποθορυβοποίησή τους. Στη συνέχεια, χρήση των T1W εικόνων για την παραγωγή των масκών της σπονδυλικής στήλης, κάνοντας χρήση του εκπαιδευμένου U-net όπως αναλύθηκε στην ενότητα 7.1. Έπειτα, registration (εγγραφή) των εικόνων ως προς τις εικόνες T1W, όπως περιγράφεται στην ενότητα 7.2. Ακολουθεί απομόνωση του τμήματος της σπονδυλικής στήλης για όλες τις ακολουθίες, χάρη στις μάσκες που παρήχθησαν από το νευρωνικό δίκτυο και χρήση της κατάλληλης μεθόδου απομόνωσης βλαβών, ώστε να παραχθούν οι μάσκες των βλαβών (ενότητα 7.3). Επιπλέον, απομόνωση των σπονδύλων και παραγωγή масκών βλαβών για τους σπονδύλους. Τέλος, χρήση των εικόνων της σπονδυλικής,

των μασκών σπονδυλικής και των μασκών βλαβών ως είσοδο στον εκπαιδευμένο SVM ταξινομητή (ενότητα 7.4) για την τελική εκτίμηση της ύπαρξης ή μη, πολλαπλού μυελώματος, καθώς και χρήση των αντίστοιχων εικόνων και μασκών για κάθε σπόνδυλο, για την εκτίμηση του είδους της βλάβης ανά σπόνδυλο.



Εικόνα 7.5-1: Διάγραμμα ολικού pipeline για εκτίμηση ασθενούς

ΜΕΡΟΣ 3: ΑΠΟΤΕΛΕΣΜΑΤΑ, ΠΕΡΙΟΡΙΣΜΟΙ, ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΠΡΟΤΑΣΕΙΣ ΓΙΑ ΜΕΛΕΤΗ

8 Κεφάλαιο 8: Αποτελέσματα

8.1 Εκπαίδευση U-net

Τα αποτελέσματα των πιο αποδοτικών τρεξιμάτων φαίνονται στον παρακάτω πίνακα:

Subdirectory	Dice	IoU	Specificity	Accuracy	F1_Score	Recall	Precision	Aggregated Score
MAIN+REC_t=3+RES+MODEL	0.7909	0.655	0.9974636	0.991794205	0.716563386	0.996803267	0.98973128	0.72295
MAIN+REC_t=3+ATT+MODEL	0.7736	0.6343	0.9980557	0.989657402	0.697058712	0.997492982	0.98680694	0.70395
MAIN+REC_t=2+RES+MODEL	0.7508	0.6059	0.99758769	0.991355133	0.670612103	0.996797308	0.98861686	0.67835
MAIN+REC_t=2+ATT+RES+MODEL	0.7481	0.6012	0.99678911	0.99083786	0.666653405	0.995726283	0.98790097	0.67465
MAIN+REC_MODEL	0.7184	0.5724	0.99739194	0.99102459	0.637143105	0.996382758	0.98766055	0.6454
MAIN+REC2_MODEL	0.7202	0.5691	0.9968693	0.990153694	0.635795889	0.99567183	0.98651277	0.64465
MAIN+DEN_MODEL	0.6614	0.5103	0.99190798	0.98666954	0.576107229	0.987913191	0.98024328	0.58585
MAIN+REC_t=3+ATT+RES+MODEL	0.6438	0.4874	0.9893975	0.98398571	0.554788048	0.983798201	0.9757291	0.5656
MAIN+DEN+REC_t=2+RES+MODEL	0.6311	0.4672	0.99038675	0.984640503	0.53692055	0.984996016	0.97624059	0.54915
MAIN+REC_t=2+DEN_MODEL	0.6309	0.4645	0.98524677	0.980410385	0.535061256	0.97714991	0.96988482	0.5477
MAIN+REC_t=2+ATT+MODEL	0.6195	0.4571	0.99828308	0.987530708	0.526051365	0.997236193	0.98026915	0.5383
MAIN+DEN+REC_t=3_ATT+RES+MODEL	0.6174	0.447	0.9845053	0.981847382	0.518560316	0.975517722	0.97143806	0.5322
MAIN+REC_t=3+DEN_MODEL	0.595	0.4254	0.98142822	0.97867527	0.496105449	0.969731688	0.96540017	0.5102
MAIN+INC+REC_t=2+MODEL	0.5746	0.416	0.97279139	0.970522118	0.482603675	0.954788593	0.95120185	0.4953
MAIN+DEN+ATT+RES+MODEL	0.5486	0.3855	0.97575035	0.973294258	0.452810834	0.957668388	0.95357991	0.46705
MAIN+DEN+ATT+MODEL	0.5472	0.3794	0.9748728	0.971748924	0.448106367	0.956096449	0.95090621	0.4633
MAIN_MODEL	0.5258	0.3695	0.95866652	0.957562637	0.434006702	0.927118596	0.92531754	0.44765

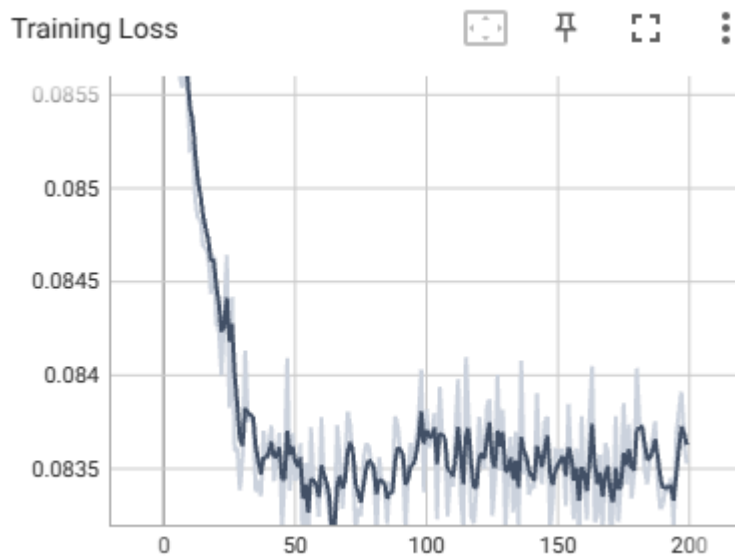
Εικόνα 8.1-1: Οι καλύτεροι συνδυασμοί U-net, έχοντας ως βάση το απλό main block

Φαίνεται πως όλοι οι συνδυασμοί παρουσιάζουν υψηλό accuracy, recall, precision και specificity. Όμως οι δείκτες Dice, IoU και F1_Score παραμένουν αρκετά χαμηλοί. Το απλό main model παρουσιάζει το χειρότερο συνδυαστικό σκορ, μόλις 0.44, ενώ ο πιο αποδοτικός συνδυασμός υπήρξε αυτός του main model με προσθήκη recurrent connections τρεις φορές για κάθε ναυρώνα, μαζί με την προσθήκη residual connections. Το συνδυαστικό σκορ αυτού του συνδυασμού ήταν το 0.72, με σαφώς βελτιωμένο Dice και F1_Score, αλλά ακόμη αρκετά μέτριο IoU.

Παρακάτω θα αναλυθούν και θα συγκριθούν τα αποτελέσματα των τριών βασικών modules, main, dense, inception, καθώς και του βέλτιστου συνδυασμού.

8.1.1 Main block

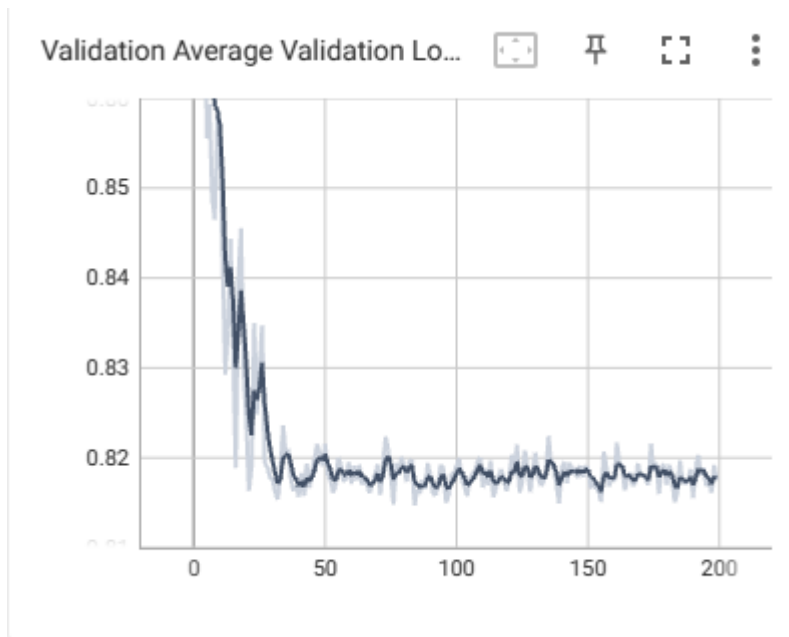
Training loss:



Εικόνα 8.1-2: Training loss του main block για 200 εποχές

Στην πρώτη αυτή περίπτωση, το σφάλμα εκπαίδευσης παρουσιάζει την φθίνουσα καμπύλη που υποδηλώνει εκπαίδευση του μοντέλου, αλλά υπάρχουν μεγάλες διακυμάνσεις γύρω από την περιοχή του ελαχίστου. Πιθανά αίτια αυτής της συμπεριφοράς είναι το learning rate, που ενδεχομένως είναι μεγάλο και προκύπτει overshooting κοντά στην περιοχή της σύγκλισης και το ενδεχόμενο το module να είναι αρκετά απλό και να υπάρχει underfitting.

Validation loss:



Εικόνα 8.1-3: Validation loss του main block για 200 εποχές

Το διάγραμμα της απώλειας επικύρωσης (validation loss) δείχνει επίσης μια πτωτική τάση, αλλά όπως και η απώλεια εκπαίδευσης, εμφανίζει κάποιες διακυμάνσεις. Ωστόσο, η απώλεια επικύρωσης φαίνεται να σταθεροποιείται καθώς προχωρά η εκπαίδευση. Αυτή η συμπεριφορά υποδηλώνει μερικά πράγματα. Το μοντέλο γενικεύει καλά. Η πτωτική τάση στην απώλεια επικύρωσης δείχνει ότι το μοντέλο μαθαίνει μοτίβα από τα δεδομένα εκπαίδευσης που είναι εφαρμόσιμα στα άγνωστα δεδομένα επικύρωσης. Οι διακυμάνσεις στο πρώτο μέρος της εκπαίδευσης θα μπορούσαν ακόμα να υποδηλώνουν ότι υπερπαραμέτροι όπως ο ρυθμός μάθησης ίσως χρειάζονται περαιτέρω ρύθμιση. Το επίπεδο θορύβου φαίνεται να μειώνεται με τον καιρό, κάτι που συχνά συμβαίνει όταν χρησιμοποιούνται βελτιστοποιητές που μειώνουν τον ρυθμό μάθησης με τον καιρό ή όταν το μοντέλο αρχίζει να συγκλίνει. Υπάρχει όμως σημαντική διαφορά μεταξύ της απώλειας εκπαίδευσης και επικύρωσης που υποδηλώνει ότι το μοντέλο μάλλον υπερπροσαρμόζεται.

Training Dice και IoU Score:



Εικόνα 8.1-4: Training Dice and IoU Score για το main block για 200 εποχές

Και οι δύο βαθμολογίες, Dice και IoU, φαίνεται να είναι αρκετά χαμηλές. Αυτοί οι δείκτες κυμαίνονται από 0 έως 1, όπου 1 υποδηλώνει τέλεια επικάλυψη μεταξύ των προβλεπόμενων και αληθινών τμημάτων και 0 υποδηλώνει καθόλου επικάλυψη. Οι τιμές που πλησιάζουν το 1 είναι επιθυμητές. Δεδομένου ότι αυτές οι βαθμολογίες είναι στο εύρος του 0.01 έως 0.026, αυτό υποδηλώνει ότι οι προβλεπόμενες τμηματοποιήσεις του μοντέλου έχουν πολύ λίγη επικάλυψη με την πραγματικότητα. Και τα δύο διαγράμματα δείχνουν βελτίωση με τον καιρό, που σημαίνει ότι το μοντέλο μαθαίνει και οι προβλέψεις πλησιάζουν περισσότερο στην πραγματικότητα καθώς προχωρά η εκπαίδευση. Ωστόσο, οι τελικές τιμές παραμένουν ακόμη χαμηλές, πράγμα που δείχνει ότι το μοντέλο μπορεί να μην επιδίδεται καλά στην εργασία τμηματοποίησης. Οι βαθμολογίες φτάνουν σε ένα πλατό μετά από έναν ορισμένο αριθμό εποχών, που θα μπορούσε να υπονοεί ότι το μοντέλο έχει φτάσει στη μέγιστη δυνατότητα όσον αφορά τη μάθηση από τα δεδομένα εκπαίδευσης που παρέχονται.

Προβλέψεις:

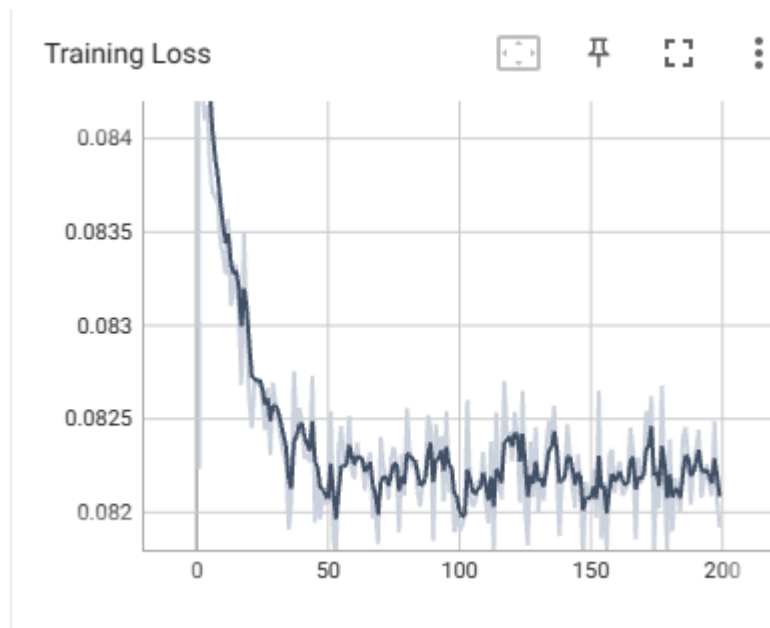
Το μοντέλο μετά την εκπαίδευση είχε τα εξής αποτελέσματα:



Εικόνα 8.1-5: Αποτέλεσμα του main model μετά την εκπαίδευση

8.1.2 Inception block

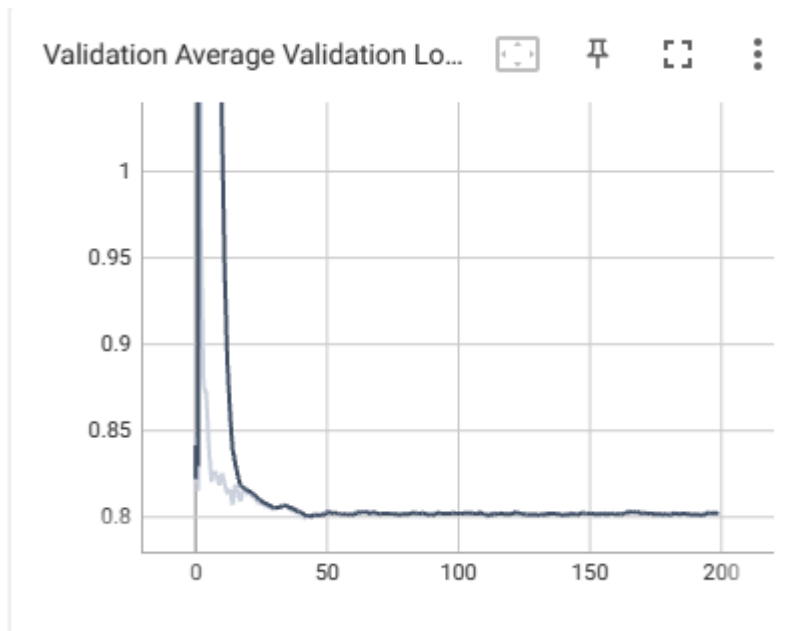
Training loss:



Εικόνα 8.1-6: Training loss του inception block για 200 εποχές

Αυτό το διάγραμμα απώλειας εκπαίδευσης είναι αρκετά σταθερό και δείχνει την απώλεια να μειώνεται μέχρι ένα σημείο σχετικής σταθερότητας. Ωστόσο, οι τιμές της απώλειας διακυμαίνονται εντός ενός στενού εύρους καθ' όλη τη διάρκεια της διαδικασίας εκπαίδευσης. Το μοντέλο φαίνεται να έχει συγκλίνει σε έναν βαθμό, όπως υποδηλώνεται από την εξομάλυνση της καμπύλης απώλειας. Το μοντέλο δεν είναι πιθανό να βελτιωθεί σημαντικά από πρόσθετη εκπαίδευση χωρίς αλλαγές στο μοντέλο ή στη διαδικασία εκπαίδευσης. Οι διακυμάνσεις εντός ενός στενού εύρους υποδηλώνουν ότι το μοντέλο είναι σταθερό αλλά μπορεί να βιώνει ελαφρές παραλλαγές στην μάθηση από παρτίδα σε παρτίδα. Ο ρυθμός μάθησης μπορεί να είναι κατάλληλος, καθώς το μοντέλο δεν εμφανίζει έντονες διακυμάνσεις στην απώλεια, αλλά ίσως να βρίσκεται στο ανώτερο άκρο του βέλτιστου εύρους. Συνήθως απαιτείται κάποια πειραματική διερεύνηση για τη λεπτομερή ρύθμιση αυτού.

Validation loss:

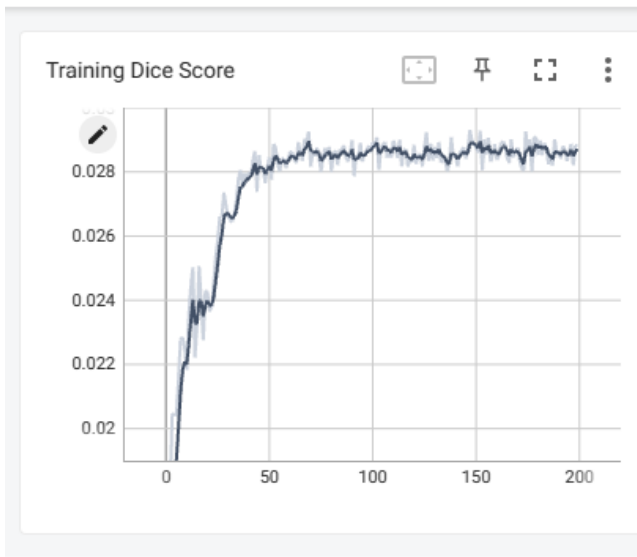


Εικόνα 8.1-7: Validation loss του inception block για 200 εποχές

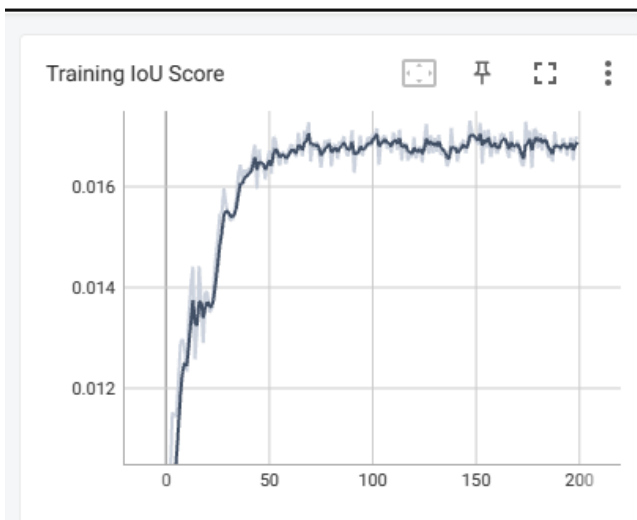
Η καμπύλη απώλειας επικύρωσης δείχνει μια γρήγορη μείωση στην αρχή, η οποία στη συνέχεια σταθεροποιείται σε μια σχετικά επίπεδη γραμμή. Αυτό είναι ένα τυπικό και ευνοϊκό μοτίβο. Το μοντέλο μαθαίνει γρήγορα από τα δεδομένα εκπαίδευσης, όπως υποδηλώνεται από την απότομη πτώση της απώλειας στην αρχή της εκπαίδευσης. Μετά από αυτή την αρχική φάση, το μοντέλο φαίνεται να έχει σταθεροποιηθεί, με την απώλεια επικύρωσης να φτάνει και να διατηρεί μια χαμηλή διακύμανση. Όμως και εδώ το πλατό της απώλειας επικύρωσης παραμένει δέκα περίπου φορές υψηλότερο από το αντίστοιχο της απώλειας εκπαίδευσης, παρέχοντας έτσι δείγματα πιθανής υπερπροσαρμογής.

Dice και IoU Score

Training Dice Score



Training IoU Score



Εικόνα 8.1-8: Dice and IoU του inception block για 200 εποχές

Και οι δύο βαθμολογίες, Dice και IoU, είναι χαμηλές, υποδηλώνοντας ότι οι προβλέψεις του μοντέλου έχουν περιορισμένη επικάλυψη με την πραγματικότητα. Για τις εργασίες τμηματοποίησης, συνήθως επιθυμούνται υψηλότερες βαθμολογίες καθώς υποδηλώνουν καλύτερη απόδοση. Το γεγονός ότι αυτές οι βαθμολογίες είναι χαμηλές μπορεί να δείχνει ένα μοντέλο που δεν αποτυπώνει πλήρως την πολυπλοκότητα της εργασίας τμηματοποίησης. Οι βαθμολογίες βελτιώνονται σημαντικά στην αρχική φάση της εκπαίδευσης, πράγμα που είναι θετικό. Ωστόσο, φαίνεται να φτάνουν σε ένα πλατό σχετικά νωρίς στη διαδικασία εκπαίδευσης, πράγμα που θα μπορούσε να υποδηλώνει ότι το μοντέλο δεν μαθαίνει περαιτέρω ή δεν μπορεί να βελτιωθεί δεδομένης της τρέχουσας αρχιτεκτονικής ή αναπαράστασης δεδομένων.

Προβλέψεις

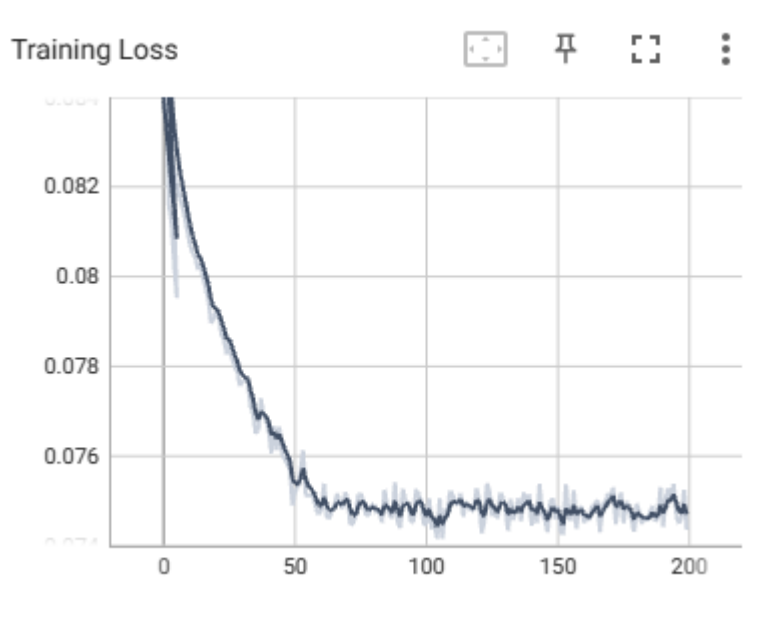
Οι προβλέψεις του inception block ήταν της εξής μορφής:



Εικόνα 8.1-9: Προβλέψεις inception block

8.1.3 Dense block

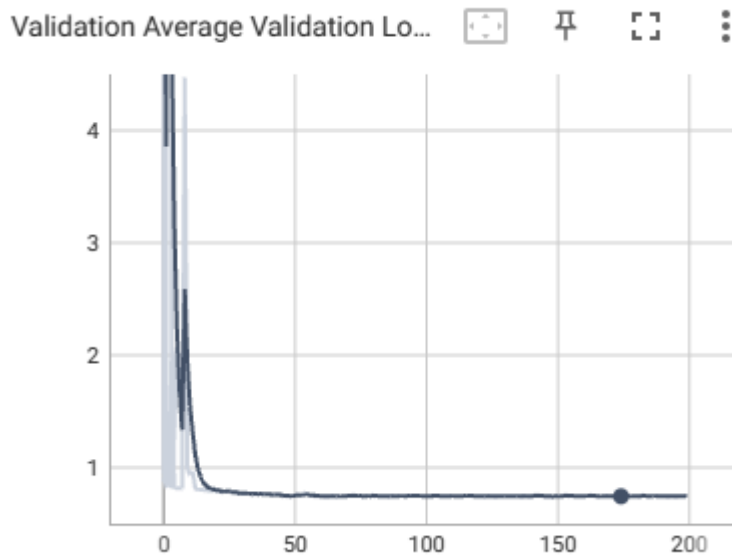
Training loss:



Εικόνα 8.1-10: Training loss του dense block για 200 εποχές

Για το διάγραμμα απώλειας εκπαίδευσης για το dense block, η συμπεριφορά είναι αρκετά παρόμοια με την προηγούμενη, αλλά με μερικές αξιοσημείωτες διαφορές. Υπάρχει μια σταθερή μείωση στην απώλεια, δείχνοντας ότι το μοντέλο μαθαίνει από τα δεδομένα εκπαίδευσης. Η απώλεια σταθεροποιείται σε μια χαμηλότερη τιμή σε σύγκριση με το προηγούμενο μοντέλο, υποδηλώνοντας καλύτερη απόδοση όσον αφορά την ελαχιστοποίηση της συνάρτησης απώλειας. Οι διακυμάνσεις φαίνονται λιγότερο έντονες σε αυτό το μοντέλο σε σύγκριση με τα προηγούμενα, πράγμα που θα μπορούσε να υποδηλώνει καλύτερη ρύθμιση του ρυθμού μάθησης ή ότι η χωρητικότητα του μοντέλου είναι πιο κατάλληλη για την πολυπλοκότητα της εργασίας. Δεδομένων αυτών των πληροφοριών, το μοντέλο φαίνεται να είναι σχετικά καλά ρυθμισμένο, τουλάχιστον όσον αφορά την απώλεια εκπαίδευσης.

Validation loss:



Εικόνα 8.1-11: Validation loss του dense block για 200 εποχές

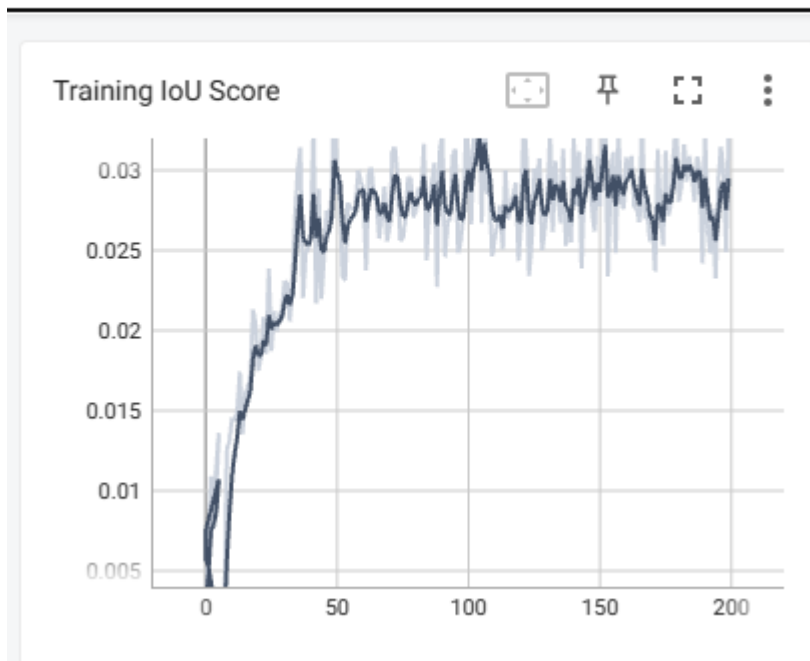
Το διάγραμμα απώλειας επικύρωσης δείχνει μια έντονη μείωση στην αρχή, η οποία σταθεροποιείται σε μια επίπεδη γραμμή αρκετά νωρίς στη διαδικασία εκπαίδευσης. Αυτό είναι γενικά ένα καλό σημάδι, καθώς μπορεί να υποδηλώνει ότι το μοντέλο δεν υπερπροσαρμόζεται και γενικεύει καλά σε άγνωστα δεδομένα. Ωστόσο, υπάρχουν κάποια αξιοσημείωτα σημεία. Η έντονη πτώση υποδηλώνει ότι το μοντέλο μαθαίνει γρήγορα από τα δεδομένα εκπαίδευσης και κάνει σημαντικές βελτιώσεις στις αρχικές εποχές. Μετά την αρχική έντονη πτώση, η απώλεια επικύρωσης σταθεροποιείται, γεγονός που αποτελεί ένδειξη ότι οι προβλέψεις του μοντέλου είναι συνεπείς κατά τη διάρκεια των εποχών μόλις το μοντέλο μάθει από τα δεδομένα. Η επίπεδη γραμμή υποδηλώνει ότι πρόσθετες εποχές εκπαίδευσης πέρα από ένα σημείο δεν οδηγούν σε περαιτέρω βελτιώσεις της απόδοσης του μοντέλου στο σετ επικύρωσης. Αυτό θα μπορούσε να είναι μια ένδειξη για την εφαρμογή πρόωρης διακοπής προκειμένου να αποφευχθεί η άσκοπη υπολογιστική διαδικασία και η πιθανή υπερεκπαίδευση. Δεδομένου ότι η απώλεια εκπαίδευσης είχε χαμηλότερο πλατό από την απώλεια επικύρωσης, πιθανή είναι επίσης η περίπτωση το μοντέλο να υπεπροσαρμόζεται.

Dice και IoU Score:

Training Dice Score



Training IoU Score



Εικόνα 8.1-12: Dice and IoU score του inception block για 200 εποχές

Και οι δύο βαθμολογίες, Dice και IoU, είναι πάλι αρκετά χαμηλές. Οι βαθμολογίες δείχνουν ότι οι προβλέψεις του μοντέλου έχουν λίγη επικάλυψη με την πραγματικότητα. Υπάρχει βελτίωση και στους δύο δείκτες καθώς προχωρά η εκπαίδευση, πράγμα που υποδηλώνει ότι το μοντέλο μαθαίνει να τμηματοποιεί καλύτερα τις εικόνες με τον καιρό. Οι μετρικές είναι κάπως θορυβώδεις, πράγμα που είναι κοινό κατά την εκπαίδευση, αλλά δείχνουν μια ανοδική

τάση, υποδηλώνοντας ότι το μοντέλο βελτιώνεται. Ωστόσο, η ποσότητα του θορύβου και ο βαθμός της βελτίωσης μπορεί να διαφέρουν ανάλογα με παράγοντες όπως το μέγεθος της παρτίδας, ο ρυθμός μάθησης ή η πολυπλοκότητα του συνόλου δεδομένων. Οι βαθμολογίες φαίνεται να σταθεροποιούνται καθώς προχωρά η εκπαίδευση, υποδηλώνοντας ότι το μοντέλο μπορεί να πλησιάζει το όριο σε ό,τι μπορεί να μάθει από τα δεδομένα στην τρέχουσα διαμόρφωσή του.

Προβλέψεις

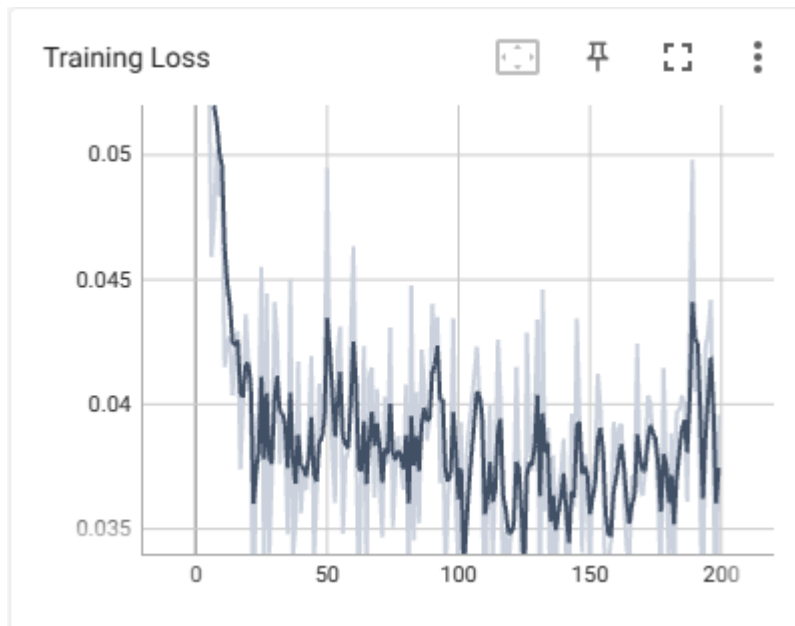
Ενδεικτική πρόβλεψη του dense block:



Εικόνα 8.1-13: Πρόβλεψη του dense block

8.1.4 Καλύτερος Συνδυασμός: Main + Recurrent (t=3) + Residual connections

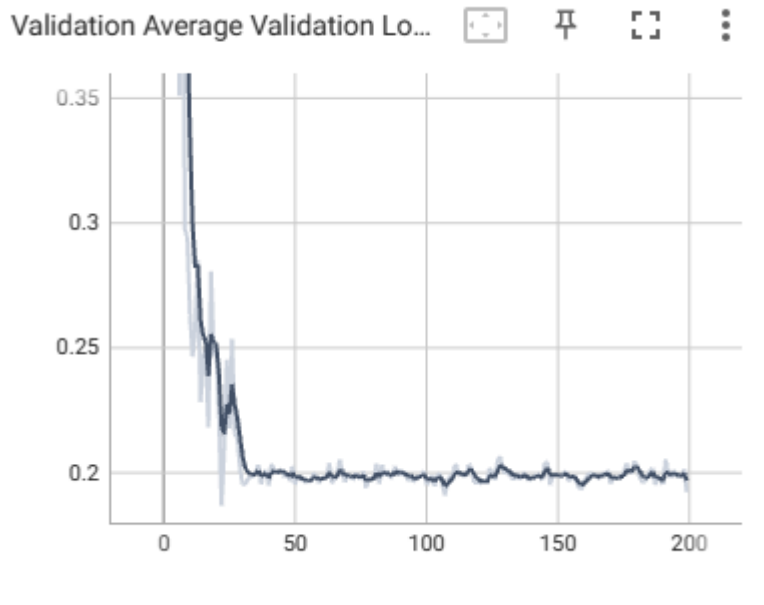
Training loss



Εικόνα 8.1-14: Training loss για το καλύτερο U-net για 200 εποχές

Το διάγραμμα απώλειας εκπαίδευσης για αυτό το τελευταίο μοντέλο δείχνει μερικά αξιοσημείωτα χαρακτηριστικά. Υπάρχει μια σαφής τάση μείωσης της απώλειας εκπαίδευσης, που υποδηλώνει ότι το μοντέλο μαθαίνει και βελτιώνει τις προβλέψεις του με τον καιρό. Η απώλεια εμφανίζει σημαντικές διακυμάνσεις καθ' όλη τη διάρκεια της εκπαίδευσης. Αυτή η μεταβλητότητα ίσως οφείλεται σε διάφορους παράγοντες, όπως μικρό μέγεθος παρτίδας, υψηλός ρυθμός μάθησης ή θορυβώδη δεδομένα. Σε αντίθεση με κάποια από τα προηγούμενα μοντέλα, αυτό δεν εμφανίζει ένα σαφές πλατό. Οι διακυμάνσεις καθιστούν πιο δύσκολο να καθοριστεί εάν η απώλεια έχει σταθεροποιηθεί σε ένα ελάχιστο τιμή, που θα μπορούσε να υποδηλώνει ότι το μοντέλο θα μπορούσε να επωφεληθεί από συνεχιζόμενη εκπαίδευση ή ρύθμιση υπερπαραμέτρων. Ο βαθμός της διακύμανσης στην απώλεια, ειδικά εάν αυξάνεται με τον καιρό, θα μπορούσε να δηλώνει ότι το μοντέλο αρχίζει να υπερπροσαρμόζεται.

Validation loss:



Εικόνα 8.1-15: Validation loss για το καλύτερο U-net για 200 εποχές

Η καμπύλη απώλειας επικύρωσης για το τελευταίο μοντέλο δείχνει ένα υγιές χαρακτηριστικό εκπαίδευσης. Αρχικά, υπάρχει μια απότομη μείωση στην απώλεια επικύρωσης, υποδηλώνοντας ότι το μοντέλο έμαθε γρήγορα τα σχετικά μοτίβα από τα δεδομένα. Μετά την αρχική κατάβαση, η απώλεια επικύρωσης σταθεροποιείται και διατηρεί μια χαμηλή διακύμανση, υποδηλώνοντας ότι το μοντέλο γενικεύει καλά χωρίς υπερπροσαρμογή. Η απώλεια επικύρωσης αυτού του μοντέλου δεν δείχνει την ανώμαλη πτώση που παρατηρήθηκε στην απώλεια επικύρωσης του Μοντέλου 2, καθιστώντας την μια πιο αξιόπιστη ένδειξη σταθερής βελτίωσης του μοντέλου. Όταν συγκρίνεται η απώλεια επικύρωσης του τελευταίου μοντέλου με την απώλεια εκπαίδευσης παρατηρείται το εξής:

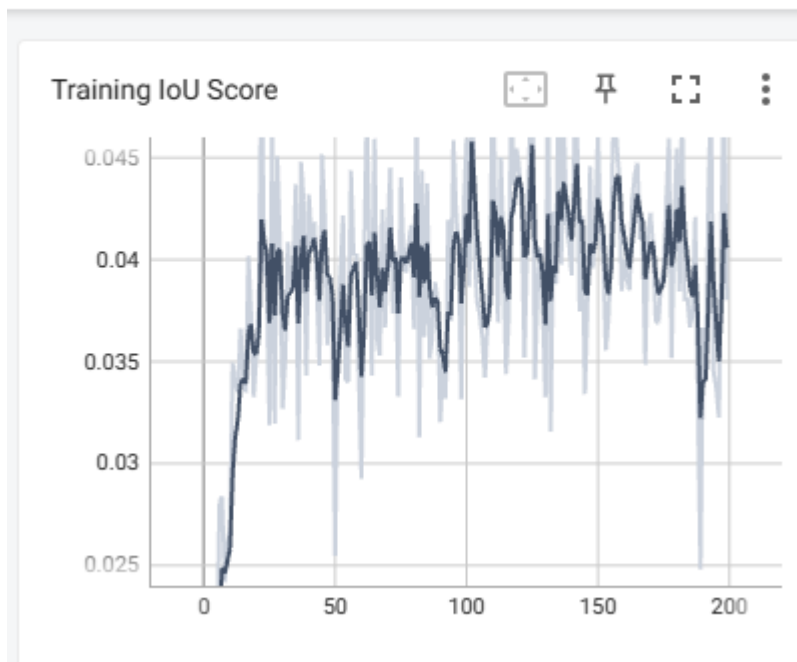
Η απώλεια εκπαίδευσης και η απώλεια επικύρωσης εμφανίζουν ένα συνεπές μοτίβο γρήγορης βελτίωσης ακολουθούμενο από σταθεροποίηση. Αυτή η ταύτιση υποδηλώνει ότι το μοντέλο μαθαίνει γενικεύσιμα μοτίβα αντί για υπερπροσαρμογή στα δεδομένα εκπαίδευσης. Παρά το γεγονός ότι η απώλεια εκπαίδευσης δείχνει περισσότερη διακύμανση από την απώλεια επικύρωσης, αυτό δεν υποδηλώνει απαραίτητα υπερπροσαρμογή, καθώς η απώλεια επικύρωσης παραμένει σταθερή και χαμηλή. Όμως και σε αυτήν την περίπτωση η απώλεια επικύρωσης κάνει πλατό σε σημαντικά υψηλότερες τιμές από το αντίστοιχο της εκπαίδευσης, γεγονός που πάλι δείχνει στοιχεία υπερπροσαρμογής.

Dice και IoU Score:

Training Dice Score



Training IoU Score



Εικόνα 8.1-16: Dice and IoU score για το καλύτερο U-net για 200 εποχές

Οι βαθμολογίες Dice και IoU για το τελευταίο μοντέλο στα δεδομένα εκπαίδευσης επισημαίνουν μερικά σημεία που αξίζει να σημειωθούν. Όπως και στα προηγούμενα μοντέλα, και οι δύο βαθμολογίες, Dice και IoU, είναι αρκετά χαμηλές. Αυτές οι μετρικές κυμαίνονται από 0 έως 1, όπου οι τιμές που πλησιάζουν το 1 υποδηλώνουν καλύτερη απόδοση τμηματοποίησης. Οι βαθμολογίες γύρω από το 0.04 έως 0.05 υποδηλώνουν ότι οι

προβλεπόμενες τμηματοποιήσεις του μοντέλου δεν συμφωνούν καλά με την πραγματικότητα. Υπάρχει μια αρχική βελτίωση καθώς οι βαθμολογίες αυξάνονται από τις αρχικές τους τιμές, αλλά φτάνουν σχετικά γρήγορα σε ένα πλατό και παραμένουν χαμηλές. Οι βαθμολογίες διακυμαίνονται σημαντικά κατά τη διάρκεια της εκπαίδευσης. Αυτό θα μπορούσε να οφείλεται στη διακύμανση των δεδομένων εκπαίδευσης, στη στοχαστική φύση του αλγορίθμου βελτιστοποίησης, ή ίσως στην ευαισθησία του μοντέλου σε διαφορετικές παρτίδες δεδομένων.

Προβλέψεις:

Ενδεικτική πρόβλεψη:



Εικόνα 8.1-17: Πρόβλεψη του καλύτερου U-net

8.2 Image Registration

Η διαδικασία του registration είναι κρίσιμη για την ακριβή ανάλυση των MRI εικόνων και την επιτυχή αναγνώριση πολλαπλού μυελώματος στη σπονδυλική στήλη. Η επιλογή της κατάλληλης μεθόδου και η εφαρμογή της σε όλα τα διαθέσιμα δεδομένα με τρόπο που εξοικονομεί χρόνο και πόρους αποτελεί σημαντικό βήμα προς την βελτιστοποίηση της διαδικασίας αυτής. Η προσέγγιση αυτή διασφαλίζει την υψηλή ποιότητα των δεδομένων που θα χρησιμοποιηθούν για την τελική ανάλυση και αναγνώριση των μορφωμάτων.

Η διαδικασία ήταν σχεδόν εξ' ολοκλήρου επιτυχής, με τις εικόνες να εγγράφονται με πολύ καλή ακρίβεια πάνω στην ακολουθία T1W. Η διαδικασία του elastix είναι αρκετά χρονοβόρα, διαρκώντας αρκετές ώρες, όμως η μέθοδος με τη χρήση του transformix συντομεύει αισθητά την ολοκλήρωση της διαδικασίας, εφόσον δεν χρειάζεται να επανυπολογιστούν οι επιμέρους παράμετροι μετασχηματισμού. Μοναδική εξαίρεση στην επιτυχία της μεθόδου υπήρξε ένα μόνο δείγμα, όπου ο μετασχημός elastix οδήγησε σε παραμόρφωση της DWI b50 εικόνας, ως εξής:



Εικόνα 8.2-1: DWI b50, παραμορφωμένη μετά την επεξεργασία από το elastix

8.3 Component Analysis

Η διαδικασία διαχωρισμού των σπονδύλων ήταν επιτυχής και διεκπεραιώθηκε χωρίς προβλήματα. Ενδεικτικά παρατίθενται κάποια σπόνδυλοι, μαζί με τις αντίστοιχες μάσκες τους:

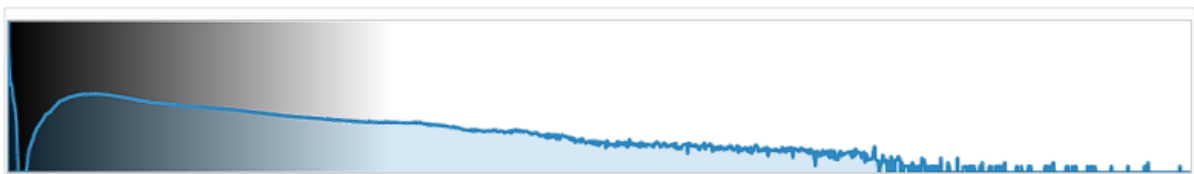


Εικόνα 8.3-1: Παραδείγματα απομόνωσης σπονδύλων

8.4 Απομόνωση βλαβών

8.4.1 Φωτεινότητες b900

Το κριτήριο των φωτεινότητων στις εικόνες της ακολουθίας DWI στο κανάλι b900 ήταν εύκολο στην υλοποίηση, καθώς υπάρχει σαφής διαχωρισμός των βόξελ στο ιστόγραμμα των εικόνων



Εικόνα 8.4-1: Ιστόγραμμα ομαλοποιημένης b900 εικόνας της σπονδυλικής στήλης

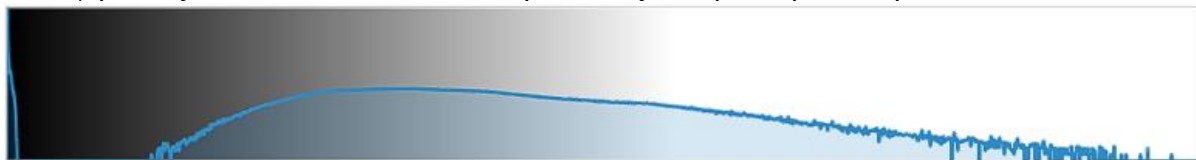
Αν και επιτυχές κριτήριο στον εντοπισμό των βλαβών, δημιουργεί πολλές ψευδώς αληθείς περιπτώσεις βλαβών, καθώς τα έντονα σημεία στις εικόνες b900 αποτελούν μεν αναγκαία αλλά όχι και ικανή συνθήκη για την διάγνωση. Ως εκ τούτου, τα αποτελέσματα των масκών είναι ανάλογα του παρακάτω δείγματος:



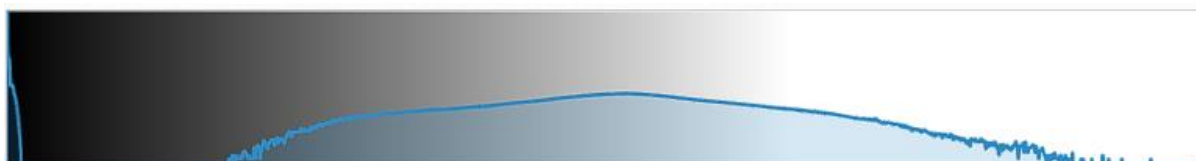
Εικόνα 8.4-2: Απομόνωση βλαβών μέσω φωτεινότητας b900

8.4.2 Συνδυασμός T1W και T2W

Το κριτήριο αυτό είναι μια επέκταση της παραπάνω μεθόδου και καθώς συνδυάζει ενδεικτικές φωτεινότητες από δύο ακολουθίες, είναι πιο ακριβές στον εντοπισμό των βλαβών. Το βασικό μειονέκτημα και εμπόδιο, υπήρξε το γεγονός πως στις εικόνες που προέρχονται από παλαιότερους μαγνήτες, η ανάλυση, ο θόρυβος και τα λοιπά artifacts παρεμποδίζουν την επιλογή ενός ιδανικού threshold, πρωτίστως στην περίπτωση των T1W εικόνων.

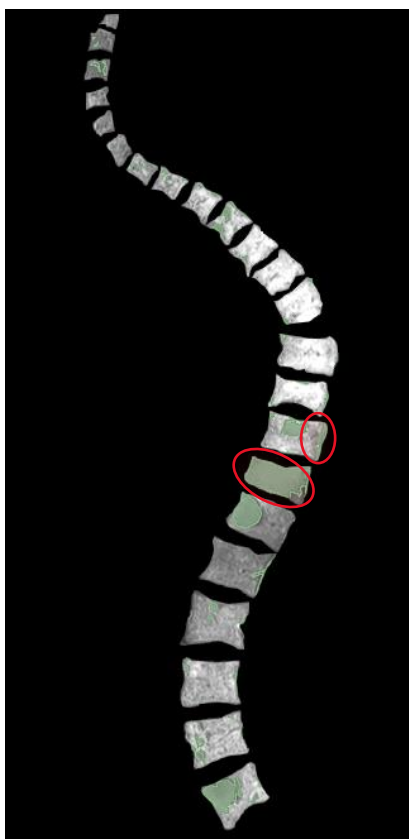


Εικόνα 8.4-3: Ιστόγραμμα T1W εικόνας από παλαιό τύπου μαγνήτη



Εικόνα 8.4-4: Ιστόγραμμα T1W εικόνας από νεότερο τύπου μαγνήτη

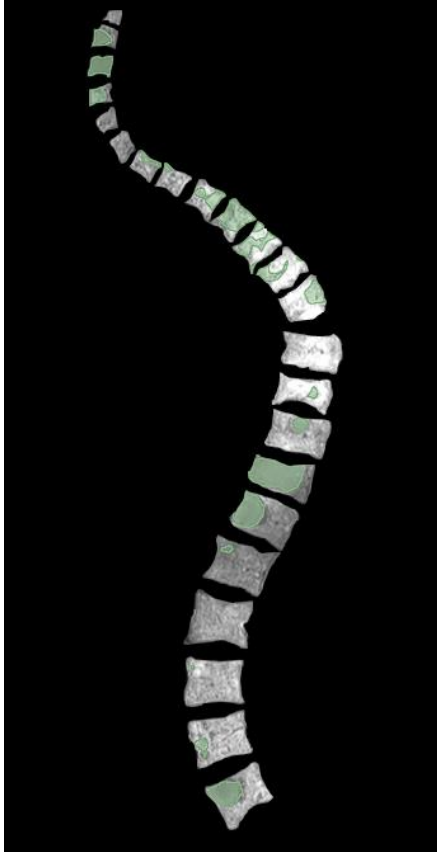
Συγκρίνοντας τις εικόνες 8.4-3 και 8.4-4, φαίνεται πως στην πρώτη περίπτωση δεν υπάρχει ξεκάθαρος διαχωρισμός στην περιοχή των γκριζων βόξελ- που θα μπορούσαν να ξεχωρίσουν τις βλάβες από τον γύρω ιστό. Έτσι, αν και υπάρχει βελτίωση συγκριτικά με την πιο απλοϊκή προσέγγιση των εικόνων b900, και σε αυτή την περίπτωση παρατηρούνται αρκετά ψευδώς αληθή δείγματα.



Εικόνα 8.4-5: Απομόνωση βλαβών με τη συνδυαστική μέθοδο T1W/T2W

8.4.3 ADC Mapping

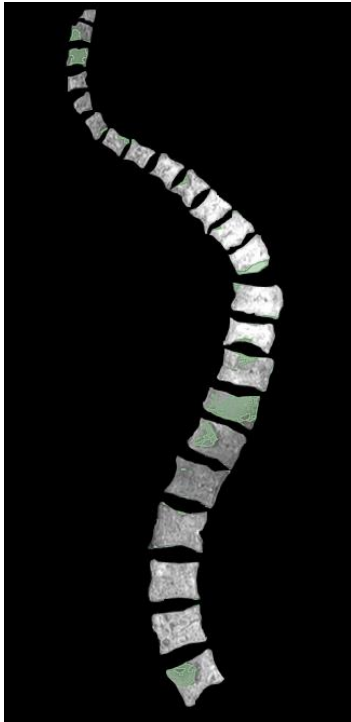
Το κριτήριο των χαρτών ADC παράγαγε ενδιαφέροντα αποτελέσματα. Ενώ αναμενόταν να είναι το πιο ακριβές από τα τέσσερα κριτήρια, τα αποτελέσματα δεν διαφέρουν σημαντικά από αυτά του κριτηρίου φωτεινότητας των εικόνων b900. Σύμφωνα με ακτινολόγο που μελέτησε το αποτέλεσμα, εικάζεται πως το γεγονός αυτό οφείλεται στο ότι αρκετά από τα περιστατικά του dataset είναι ιδιάζοντα και πιθανότατα παρουσιάζονται πολλές τιμές στο εύρος επικάλυψης φυσιολογικού και παθολογικού ιστού στους χάρτες ADC.



Εικόνα 8.4-6: ADC mask βλαβών

8.4.4 Fat Fraction

Η τεχνική fat fraction φάνηκε να έχει τον μικρότερο αριθμό ψευδώς αληθών δειγμάτων από τις τεχνικές που ακολουθήθηκαν. Παρόλαυτα, η απομόνωση των βλαβών δεν ήταν ακόμη ικανοποιητικά ακριβής, αφού σε σημεία του μυελού όπου υπήρχε βλάβη, ο χάρτης fat fraction δεν κάλυπτε ολόκληρη την περιοχή. Επομένως αν και υπήρξε βελτίωση του προβλήματος των ψευδώς αληθών μασκών, εισήχθη το ενδεχόμενο παράληψης παθολογικών δομών από τις τελικές μάσκες.



Εικόνα 8.4-7: Απομόνωση βλαβών μέσω του δείκτη Fat Fraction

8.5 Διαδικασία Classification

8.5.1 Feature extraction

Η διαδικασία της επιλογής χαρακτηριστικών ήταν επιτυχής. Εξαγόntonταν 110 χρήσιμα χαρακτηριστικά από κάθε εικόνα, συνολικά 660 χαρακτηριστικά για κάθε ασθενή, από όλες τις ακολουθίες. Σε αυτά τα χαρακτηριστικά εκπαιδεύτηκε ο δυαδικός ταξινομητής SVM.

8.5.2 Εκπαίδευση του SVM

Η εκπαίδευση του SVM μοντέλου με τον αρχικό διαμοιρασμό του dataset χαρακτηρίστηκε από τον περιορισμένο αριθμό των δεδομένων. Αν και το μοντέλο ολοκλήρωσε την εκπαίδευση χωρίς σφάλματα, παρουσιάστηκε τυπική περίπτωση overfitting, όπως φαίνεται και από τα αποτελέσματα της διαδικασίας.

```

Best parameters found: {'svm_C': 1, 'svm_gamma': 'scale',
'svm_kernel': 'rbf'}
Classification report for the best classifier:
      precision    recall  f1-score   support

   0           1.00     1.00     1.00         2
   1           1.00     1.00     1.00         2

 accuracy          1.00
 macro avg          1.00
 weighted avg       1.00

Accuracy score for the best classifier: 1.0

```

Εικόνα 8.5-1: Αποτελέσματα εκπαίδευσης του SVM classifier

Στην προσπάθεια να αντιμετωπιστεί αυτό το πρόβλημα, αποφασίστηκε να χρησιμοποιηθεί για εκ νέου εκπαίδευση, ολόκληρο το dataset. Στην δεύτερη αυτή περίπτωση, εκπαιδεύτηκε ένας SVM ταξινομητής με τα εξής χαρακτηριστικά:

```

Fitting 5 folds for each of 32 candidates, totalling 160 fits
Best parameters found: {'svm_C': 0.1, 'svm_gamma': 'scale', 'svm_kernel': 'linear'}
Classification report for the best classifier:
      precision    recall  f1-score   support

   0           0.80     0.80     0.80         5
   1           0.75     0.75     0.75         4

 accuracy          0.78
 macro avg          0.78
 weighted avg       0.78

Accuracy score for the best classifier: 0.7777777777777778

```

Εικόνα 8.5-2: Αποτελέσματα εκπαίδευσης με τη χρήση ολόκληρου του dataset

Τα καλύτερα αποτελέσματα βρέθηκαν με παραμέτρους

- C = 0.1
- Gamma = scale
- Kernel = linear

Και accuracy score = 0.77778.

9 Κεφάλαιο 9: Περιορισμοί

Κατά τη διάρκεια της διαδικασίας που ακολουθήθηκε, υπήρχαν ορισμένοι περιορισμοί που δεν επέτρεψαν την ύπαρξη πλήρως αντικειμενικών συμπερασμάτων. Στη συνέχεια αυτού του κεφαλαίου θα αναλυθούν αυτοί οι περιορισμοί, ανά τμήμα του εργασιακού pipeline.

9.1 Εκπαίδευση των U-net

Ο πυρήνας των περιορισμών στο συγκεκριμένο βήμα ήταν ο περιορισμένος αριθμός δειγμάτων, καθώς τον νούμερο των 41ός δειγμάτων είναι αρκετά μικρό για διαδικασίες εκπαίδευσης νευρωνικών δικτύων και μοντέλων μηχανικής μάθησης. Στο κομμάτι των νευρωνικών δικτύων U-net, ο περιορισμός αυτός αντιμετωπίστηκε με αρκετή επιτυχία χρησιμοποιώντας τεχνικές data augmentation, σε συνδυασμό με προσεκτική επιλογή των εποχών εκπαίδευσης και early stopping κριτηρίων, επιτρέποντας έτσι σε μεγάλο βαθμό την αποφυγή περιπτώσεων overfitting. Επιπλέον περιορισμό στη δυνατότητα δοκιμής διαφορετικών, πιο σύνθετων δικτύων, προκάλεσε το διαθέσιμο hardware, τόσο από άποψη μνήμης RAM, όσο και από άποψη κάρτας γραφικών που θα μπορούσε να υποστηρίξει το υπολογιστικό βάρος αυτών των δικτύων.

9.2 Απομόνωση βλαβών

Στο βήμα αυτό, αντιμετωπίστηκαν αρκετές δυσκολίες αναφορικά με τον ακριβή εντοπισμό των βλαβών, καθώς λόγω χρονοτριβής των ακτινολόγων δεν υπήρξε η δυνατότητα απόκτησης μασκών αλήθειας για τις βλάβες εγκαίρως. Οι τέσσερις μέθοδοι που δοκιμάστηκαν, αν και αρκετά ενδεικτικοί, δεν είναι αλάθητοι, όπως φάνηκε στο κεφάλαιο των αποτελεσμάτων. Σε πολλές εικόνες που είχαν προκύψει από παλαιότερους μαγνήτες, η ύπαρξη θορύβου και artifacts καθιστούσε τις αλλαγές στις φωτεινότητες των βόξελ μεταξύ βλαβών και γύρω ιστού σχετικά μικρές, με αποτέλεσμα οι αλγόριθμοι απομόνωσης που ακολουθήθηκαν να συμπεριλαμβάνουν false positive δείγματα, με αποτέλεσμα τα δεδομένα που χρησιμοποιήθηκαν στο βήμα της ταξινόμησης να περιέχουν σφάλματα, επηρεάζοντας σημαντικά την αξιοπιστία του ταξινομητή.

9.3 Ταξινόμηση

Πέραν του περιορισμού που μόλις αναφέρθηκε, αναφορικά με την αξιοπιστία των δεδομένων αλήθειας, σημαντικός περιορισμός για την ταξινόμηση υπήρξε ο περιορισμένος αριθμός δειγμάτων. Σε αντίθεση με τα νευρωνικά δίκτυα, όπου οι μέθοδοι augmentation βοήθησαν στην επίλυση αυτού του θέματος, στην περίπτωση αυτή η έλλειψη δεδομένων φάνηκε καθοριστική. Ο δυαδικός ταξινομητής εκπαιδεύτηκε μεν χωρίς σφάλματα, αλλά παρουσίασε overfitting, όπως φαίνεται στην εικόνα 8.3-9. Κατά την δεύτερη εκπαίδευση, αντιμετωπίστηκε το πρόβλημα του overfitting, δεν υπήρχαν όμως νέα ανεξάρτητα δεδομένα ώστε να ελεγχθεί η απόδοση του ταξινομητή σε μη γνωστά δεδομένα. Οι ανά σπόνδυλο ταξινομητές ζημιώθηκαν ακόμη περισσότερο από την έλλειψη δειγμάτων. Πιο συγκεκριμένα, στα 27 δείγματα

παθογένειας της νόσου, variegated και diffusion τύπους παρουσίαζε ένα δείγμα για την κάθε περίπτωση, με τα υπόλοιπα 25 να είναι δείγματα τύπου focal και τα εναπομείναντα να είναι μη-παθογενείς περιπτώσεις. Ως εκ τούτου, λόγω του bias που προκύπτει από την ανισορροπία αυτή του dataset, δεν κατέστη δυνατή η επιτυχής εκπαίδευση αυτών των επί μέρους ταξινομητών.

10 Κεφάλαιο 10: Συμπεράσματα και Προτάσεις για μελέτη

Το αντικείμενο της παρούσης εργασίας ήταν η ανάπτυξη ενός pipeline που θα μπορούσε να διαχειριστεί τρισδιάστατα δεδομένα μαγνητικής τομογραφίας και εν τέλει να παράξει πρόβλεψη τόσο για την ύπαρξη ή μη πολλαπλού μυελώματος στην σπονδυλική στήλη, όσο και για το είδος της βλάβης. Για τον σκοπό αυτό, επιστρατεύτηκαν νευρωνικά δίκτυα U-net, τεχνικές εγγραφής (registration) εικόνας, τεχνικές τμηματοποίησης για την αυτόματη ανίχνευση των βλαβών και ταξινομητές SVM. Όσον αφορά τα U-net, η απόδοσή τους δεν ήταν η προβλεπόμενη. Παρά την εκπαίδευση 36 μοντέλων και την χρήση πολλών συνδυασμών αρχιτεκτονικών, ο μετρητής dice-loss, που είναι ενδεικτικός στις περιπτώσεις τμηματοποίησης, παρέμεινε χαμηλός, με το καλύτερο σκορ να είναι 0.79, μετρική όμως που προκύπτει από ένα πάρα πολύ μικρό test dataset 5 εικόνων, ενώ το dice-loss κατά τη διαδικασία της εκπαίδευσης ήταν πολύ χαμηλότερο. Εκτιμάται πως λόγω των περιορισμένων υπολογιστικών πόρων, το βάθος των δικτύων και το επιτρεπτό batch size δεν επαρκούσαν έτσι ώστε να μάθουν τα δίκτυα ικανοποιητικά από το ήδη περιορισμένο dataset.

Η ακριβής απομόνωση των βλαβών επίσης είναι μια σύνθετη διαδικασία, που λόγω και της φύσης των δεδομένων (τρόπου λήψης, τύπο μαγνήτη, παλαιότητα μαγνήτη), γίνεται ακόμη πιο απαιτητική. Αυτοματοποιημένοι αλγόριθμοι που βασίζονται σε ένα ή δύο biomarkers αποτελούν καλές ενδείξεις, αλλά όχι επαρκείς ώστε να εξασφαλίζεται η απαραίτητη ακρίβεια. Απαραίτητη φαίνεται να είναι η σύμπραξη έμπειρων ακτινολόγων και πιο σύνθετων αλγορίθμων.

Η ταξινόμηση με τη χρήση SVM εναρμονίζεται με την διαδικασία, λόγω της πολυδιάστατης φύσης του τελικού προβλήματος ταξινόμησης. Η τελική ακρίβεια του δυαδικού ταξινομητή (0.7778) δεν είναι ιδανική, αλλά υποδεικνύει πως με μεγαλύτερο αρχικό dataset και ακριβέστερη απομόνωση βλαβών, δύναται να εκπαιδευτεί ένα κατά πολύ ακριβέστερο εργαλείο ταξινόμησης. Αντιθέτως, οι ταξινομητές ανά σπόνδυλο δεν μπόρεσαν να εκπαιδευτούν λόγω της έλλειψης δεδομένων, και άρα δεν μπορούν να εξαχθούν συμπεράσματα για τη λειτουργία τους.

Παρά τους περιορισμούς που συζητήθηκαν, η συγκεκριμένη εργασία μπορεί να αποτελέσει ένα υπόβαθρο μελλοντικών προεκτάσεων. Με την σύμπραξη ακτινολόγων, μπορούν να βελτιστοποιηθούν οι αλγόριθμοι εντοπισμού βλαβών και να παρέχουν ακριβή δεδομένα για όλες τις περιπτώσεις. Επιπλέον, απαραίτητη θα πρέπει να είναι και η διαθεσιμότητα ενός αρκετά εκτενέστερου dataset, τόσο για τις ανάγκες των νευρωνικών δικτύων, όσο και για τους

ταξινομητές. Οι κώδικες εκπαίδευσης και για τους δύο τύπους ταξινομητών, τον δυαδικό και τους ανά σπόνδυλο, είναι διαθέσιμοι, και με τον κατάλληλο όγκο δεδομένων μπορούν να εκπαιδευτούν ακριβείς ταξινομητές που θα είναι σε θέση αποτελέσουν βοηθητικά εργαλεία για την διάγνωση της νόσου. Τέλος, χρήσιμη θα μπορούσε να φανεί και η εκπαίδευση των U-net σε ισχυρότερους υπολογιστές που θα είναι σε θέση να διαχειριστούν πολυπλοκότερα δίκτυα, μεγαλύτερο batch size και περεταίρω συνδυασμούς, ώστε να βρεθεί ένα ακόμη πιο βελτιωμένο δίκτυο για την αρχική τμηματοποίηση της σπονδυλικής στήλης.

Βιβλιογραφία

- [1] Γεωργίου Ευάγγελος, *ΙΑΤΡΙΚΗ ΦΥΣΙΚΗ (ΕΠΙΤΟΜΟ) ΔΙΑΓΝΩΣΤΙΚΕΣ ΚΑΙ ΘΕΡΑΠΕΥΤΙΚΕΣ ΕΦΑΡΜΟΓΕΣ ΤΩΝ ΑΚΤΙΝΟΒΟΛΙΩΝ*.
- [2] “https://en.wikipedia.org/wiki/Support_vector_machine.”
- [3] “<https://www.myeloma.org/what-is-multiple-myeloma>.”
- [4] Lia Angela Mouloupoulos and Vassilis Koutoulidis, *Bone Marrow MRI: A Pattern-Based Approach*.
- [5] P. Omoumi, “The Dixon method in musculoskeletal MRI: from fat-sensitive to fat-specific imaging,” *Skeletal Radiol*, vol. 51, no. 7, pp. 1365–1369, Jul. 2022, doi: 10.1007/S00256-021-03950-1/FIGURES/1.
- [6] “Dixon method - Questions and Answers in MRI.” Accessed: Mar. 28, 2024. [Online]. Available: <https://mriquestions.com/dixon-method.html>
- [7] “https://scikit-image.org/docs/stable/auto_examples/segmentation/plot_multiotsu.html.”
- [8] Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, “A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects,” *IEEE Trans Neural Netw Learn Syst*, vol. 33, no. 12, pp. 6999–7019, Dec. 2022, doi: 10.1109/TNNLS.2021.3084827.
- [9] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, “Convolutional neural networks: an overview and application in radiology,” *Insights Imaging*, vol. 9, no. 4, pp. 611–629, Aug. 2018, doi: 10.1007/S13244-018-0639-9/FIGURES/15.
- [10] “Sci-Hub | A survey of the recent architectures of deep convolutional neural networks. Artificial Intelligence Review, 53(8), 5455–5516 | 10.1007/s10462-020-09825-6.” Accessed: Mar. 24, 2024. [Online]. Available: <https://sci-hub.se/https://link.springer.com/article/10.1007/s10462-020-09825-6>
- [11] D. X. Zhou, “Universality of deep convolutional neural networks,” *Appl Comput Harmon Anal*, vol. 48, no. 2, pp. 787–794, Mar. 2020, doi: 10.1016/J.ACHA.2019.06.004.
- [12] H. Gholamalizhad and H. Khosravi, “Pooling Methods in Deep Neural Networks, a Review”.
- [13] M. Nikzad, Y. Gao, and J. Zhou, “Gradient-based pooling for convolutional neural networks,” *2019 IEEE International Conference on Visual Communications and Image Processing, VCIP 2019*, Dec. 2019, doi: 10.1109/VCIP47243.2019.8966049.
- [14] S. Qian, H. Liu, C. Liu, S. Wu, and H. S. Wong, “Adaptive activation functions in convolutional neural networks,” *Neurocomputing*, vol. 272, pp. 204–212, Jan. 2018, doi: 10.1016/J.NEUCOM.2017.06.070.
- [15] H. H. Xiu, “Research on Activation Function in Deep Convolutional Neural Network,” *Proceedings of the 2020 Conference on Artificial Intelligence and Healthcare*, pp. 19–24, Oct. 2020, doi: 10.1145/3433996.3434001.
- [16] R. ZahediNasab and H. Mohseni, “Neuroevolutionary based convolutional neural network with adaptive activation functions,” *Neurocomputing*, vol. 381, pp. 306–313, Mar. 2020, doi: 10.1016/J.NEUCOM.2019.11.090.
- [17] R. Zahedinasab and H. Mohseni, “Enhancement of CT Brain Images Classification Based on Deep Learning Network with Adaptive Activation Functions,” *2018 8th International Conference on Computer and Knowledge Engineering (ICCKE)*, pp. 182–187, Dec. 2018, doi: 10.1109/ICCKE.2018.8566362.
- [18] D. C. Marcu and C. Grava, “The impact of activation functions on training and performance of a deep neural network,” *2021 16th International Conference on Engineering of Modern Electric Systems (EMES)*, pp. 1–4, Jun. 2021, doi: 10.1109/EMES52337.2021.9484108.
- [19] A. A. Alkhoully, A. Mohammed, and H. A. Hefny, “Improving the Performance of Deep Neural Networks Using Two Proposed Activation Functions,” *IEEE Access*, vol. 9, pp. 82249–82271, 2021, doi: 10.1109/ACCESS.2021.3085855.

- [20] M. M. Bejani and M. Ghatee, “A systematic review on overfitting control in shallow and deep neural networks,” *Artif Intell Rev*, vol. 54, no. 8, pp. 6391–6438, Dec. 2021, doi: 10.1007/S10462-021-09975-1.
- [21] C. Shorten and T. M. Khoshgoftaar, “A survey on Image Data Augmentation for Deep Learning,” *J Big Data*, vol. 6, no. 1, pp. 1–48, Dec. 2019, doi: 10.1186/S40537-019-0197-0.
- [22] A. Badola, V. P. Nair, and R. P. Lal, “An Analysis of Regularization Methods in Deep Neural Networks,” *2020 IEEE 17th India Council International Conference (INDICON)*, pp. 1–6, Dec. 2020, doi: 10.1109/INDICON49873.2020.9342192.
- [23] G. H. de Rosa, J. P. Papa, and X. S. Yang, “Handling dropout probability estimation in convolution neural networks using meta-heuristics,” *Soft comput*, vol. 22, no. 18, pp. 6147–6156, Sep. 2018, doi: 10.1007/S00500-017-2678-4.
- [24] Y. Wang, Z. P. Bian, J. Hou, and L. P. Chau, “Convolutional Neural Networks With Dynamic Regularization,” *IEEE Trans Neural Netw Learn Syst*, vol. 32, no. 5, pp. 2299–2304, May 2021, doi: 10.1109/TNNLS.2020.2997044.
- [25] A. Agnes Lydia and F. Sagayaraj Francis, “Convolutional Neural Network with an Optimized Backpropagation Technique,” *2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN)*, pp. 1–5, Mar. 2019, doi: 10.1109/ICSCAN.2019.8878719.
- [26] T. P. Lillicrap, A. Santoro, L. Marris, C. J. Akerman, and G. Hinton, “Backpropagation and the brain,” *Nat Rev Neurosci*, vol. 21, no. 6, pp. 335–346, Jun. 2020, doi: 10.1038/S41583-020-0277-3.
- [27] Z. Liu, R. Feng, X. Li, W. Wang, and X. Wu, “Gradient-Sensitive Optimization for Convolutional Neural Networks,” *Comput. Intell. Neurosci.*, vol. 2021, pp. 66718301–667183016, 2021, doi: 10.1155/2021/6671830.
- [28] Y. Sun *et al.*, “Convolutional Neural Network Based Models for Improving Super-Resolution Imaging,” *IEEE Access*, vol. 7, pp. 43042–43051, 2019, doi: 10.1109/ACCESS.2019.2908501.
- [29] H. Lee, “Analysis on performances of the optimization algorithms in CNN speech noise attenuator,” *Indonesian Journal of Electrical Engineering and Informatics (IJEI)*, vol. 9, no. 4, pp. 827–833, Dec. 2021, doi: 10.52549/IJEI.V9I4.3245.
- [30] G. J. Duarte, T. A. Pereira, E. J. Nascimento, D. Mesquita, and A. H. Souza Junior, “How do loss functions impact the performance of graph neural networks?,” *Anais do 15. Congresso Brasileiro de Inteligência Computacional*, pp. 1–7, Nov. 2021, doi: 10.21528/CBIC2021-161.
- [31] F. Gao, C. H. Chen, H. Zhang, and S. Y. Hsu, “Improving Robustness of CNN on Noisy Labels with Hinge Loss Function,” *2023 15th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, pp. 32–35, 2023, doi: 10.1109/IHMSC58761.2023.00016.
- [32] S. Balasundaram and S. C. Prasad, “Robust twin support vector regression based on Huber loss function,” *Neural Comput Appl*, vol. 32, no. 15, pp. 11285–11309, Aug. 2020, doi: 10.1007/S00521-019-04625-8.
- [33] F. Z. El Biach, I. Iala, H. Laanaya, and K. Minaoui, “Efficient balanced focal loss function for manipulated images detection,” *2021 Fifth International Conference On Intelligent Computing in Data Sciences (ICDS)*, pp. 1–4, 2021, doi: 10.1109/ICDS53782.2021.9626750.
- [34] P. Wang and A. C. S. Chung, “Focal Dice Loss and Image Dilation for Brain Tumor Segmentation,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11045 LNCS, pp. 119–127, 2018, doi: 10.1007/978-3-030-00889-5_14.
- [35] Y. Bengio, “Learning Deep Architectures for AI,” *Foundations and Trends® in Machine Learning*, vol. 2, no. 1, pp. 1–127, Nov. 2009, doi: 10.1561/22000000006.
- [36] Y. Amalia, M. Khairoh, A. B., and B. Santoso, “The Implementation of Deep Learning for White Blood Cell Subtype Classification from Microscopic Images,” *Proceedings of the Proceedings of the 1st Asian Conference on Humanities, Industry, and Technology for Society, ACHITS 2019, 30-31 July 2019, Surabaya, Indonesia*, Oct. 2019, doi: 10.4108/EAI.30-7-2019.2287610.

- [37] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature* 2015 521:7553, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.
- [38] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun ACM*, vol. 60, no. 6, pp. 84–90, Jun. 2017, doi: 10.1145/3065386.
- [39] N. Siddique, S. Paheding, C. P. Elkin, and V. Devabhaktuni, "U-Net and Its Variants for Medical Image Segmentation: A Review of Theory and Applications," *IEEE Access*, vol. 9, pp. 82031–82057, 2021, doi: 10.1109/ACCESS.2021.3086020.
- [40] X. X. Yin, L. Sun, Y. Fu, R. Lu, and Y. Zhang, "U-Net-Based Medical Image Segmentation," *J Healthc Eng*, vol. 2022, 2022, doi: 10.1155/2022/4189781.
- [41] P. Markelj, D. Tomaževič, B. Likar, and F. Pernuš, "A review of 3D/2D registration methods for image-guided interventions," *Med Image Anal*, vol. 16 3, no. 3, pp. 642–61, Apr. 2012, doi: 10.1016/J.MEDIA.2010.03.005.
- [42] H. Song and P. Qiu, "Intensity-based 3D local image registration," *Pattern Recognit. Lett.*, vol. 94, pp. 15–21, Jul. 2017, doi: 10.1016/J.PATREC.2017.04.021.
- [43] P. Devadas, G. Kalaiarasi, and M. Selvi, "Intensity based Image Registration on Brain MRI Images," *2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA)*, pp. 257–262, Jul. 2020, doi: 10.1109/ICIRCA48905.2020.9183191.
- [44] J. Zheng, Z. Ji, K. Yu, Q. An, Z. Guo, and Z. Wu, "A feature-based solution for 3D registration of CT and MRI images of human knee," *Signal Image Video Process*, vol. 9, no. 8, pp. 1815–1824, Nov. 2015, doi: 10.1007/S11760-014-0660-5.
- [45] S. Yang *et al.*, "Multiscale feature fusion network for 3D head MRI image registration," *Med Phys*, vol. 50, no. 9, pp. 5609–5620, Sep. 2023, doi: 10.1002/MP.16387.
- [46] S. Ghafurian, I. Hacihaliloglu, D. N. Metaxas, V. Tan, and K. Li, "A computationally efficient 3D/2D registration method based on image gradient direction probability density function," *Neurocomputing*, vol. 229, pp. 100–108, Mar. 2017, doi: 10.1016/J.NEUCOM.2016.07.070.
- [47] D. Shen, G. Wu, and H. Il Suk, "Deep Learning in Medical Image Analysis," *Annu Rev Biomed Eng*, vol. 19, pp. 221–248, Jun. 2017, doi: 10.1146/ANNUREV-BIOENG-071516-044442.
- [48] J. Liu *et al.*, "Applications of deep learning to MRI images: A survey," *Big Data Min. Anal.*, vol. 1, no. 1, pp. 1–18, Mar. 2018, doi: 10.26599/BDMA.2018.9020001.
- [49] H. Sun, "A Review of 3D-2D Registration Methods and Applications based on Medical Images," *Highlights in Science, Engineering and Technology*, vol. 35, pp. 200–224, Apr. 2023, doi: 10.54097/HSET.V35I.7055.
- [50] J. Jin and Y.-F. Yang, "Similarity metric in medical image registration," pp. 688–691, 2013, doi: 10.2991/ICCSEE.2013.175.
- [51] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Trans Med Imaging*, vol. 16, no. 2, pp. 187–198, 1997, doi: 10.1109/42.563664.
- [52] W. R. Crum, D. L. G. Hill, and D. J. Hawkes, "Information Theoretic Similarity Measures in Non-rigid Registration," *Inf Process Med Imaging*, vol. 18, pp. 378–87, 2003, doi: 10.1007/978-3-540-45087-0_32.
- [53] E. Peli, R. A. Augliere, and G. T. Timberlake, "Feature-Based Registration of Retinal Images," *IEEE Trans Med Imaging*, vol. 6, no. 3, pp. 272–278, 1987, doi: 10.1109/TMI.1987.4307837.
- [54] G. Bieszczad, T. Sosnowski, and H. Madura, "Improved sum-of-squared-differences tracking algorithm for thermal vision systems," *International Symposium on Photoelectronic Detection and Imaging 2011: Advances in Infrared Imaging and Applications*, vol. 8193, p. 81932R, Jun. 2011, doi: 10.1117/12.900540.
- [55] J. N. Sarvaiya, S. Patnaik, and S. Bombaywala, "Image Registration by Template Matching Using Normalized Cross-Correlation," *2009 International Conference on Advances in Computing, Control, and Telecommunication Technologies*, pp. 819–822, 2009, doi: 10.1109/ACT.2009.207.

- [56] A. Li, A. Kumar, Y. Ha, and H. Corporaal, “Correlation ratio based volume image registration on GPUs,” *Microprocess. Microsystems*, vol. 39, no. 8, pp. 998–1011, Nov. 2015, doi: 10.1016/J.MICPRO.2015.04.002.
- [57] D. Brunet, E. R. Vrscay, and Z. Wang, “On the Mathematical Properties of the Structural Similarity Index,” *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 1488–1499, Apr. 2012, doi: 10.1109/TIP.2011.2173206.
- [58] “[https://en.wikipedia.org/wiki/Elastix_\(image_registration\)](https://en.wikipedia.org/wiki/Elastix_(image_registration)).”
- [59] “https://en.wikipedia.org/wiki/Supervised_learning.”
- [60] “https://en.wikipedia.org/wiki/Unsupervised_learning.”
- [61] “https://en.wikipedia.org/wiki/Weak_supervision.”
- [62] “<https://spinningup.openai.com/en/latest/>.”
- [63] “https://en.wikipedia.org/wiki/Multi-label_classification.”
- [64] “https://en.wikipedia.org/wiki/Multi-label_classification.”
- [65] “https://en.wikipedia.org/wiki/Category:Classification_algorithms.”
- [66]
- [67] “<https://stackoverflow.com/questions/60312401/when-using-the-k-means-clustering-algorithm-is-it-possible-to-have-a-set-of-dat>”
- [68] “https://scipy-lectures.org/packages/scikit-image/auto_examples/plot_threshold.html”
- [69] 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation”Ozgun Annotation”ozgün A, İçek C, [...] Ronneberger O
- [70] Rafael C. Gonzalez and Richard E. Woods, *Digital Image Processing*.
- [71] Acta Oncologica A strategy for multimodal deformable image registration to integrate PET/MR into radiotherapy treatment planning) A strategy for multimodal deformable image registration to integrate PET/MR into radiotherapy treatment planning
Leibfarth S, Mönnich D, [...] Thorwarth D
Acta Oncologica (2013) 52(7) 1353-1359
- [72] “[Image segmentation - Wikipedia](#)”
- [73] “[Otsu's method - Wikipedia](#)”