



ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ

ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΒΙΟΪΑΤΡΙΚΗΣ

Διπλωματική Εργασία

Τμηματοποίηση όγκων εγκεφάλου σε εικόνες
μαγνητικής τομογραφίας μέσω αλγορίθμων βαθιάς
μάθησης

Γεώργιος Ιάσων Χριστουλάκης

Αριθμός Μητρώου: 17107

Επιβλέπων Καθηγητής

Σπύρος Κωστόπουλος, Αν. Καθηγητής

Αθήνα, Σεπτέμβριος 2022

Η Τριμελής Εξεταστική Επιτροπή

Ο Επιβλέπων Καθηγητής

Σπύρος Κωστόπουλος

Αν. Καθηγητής

[ΨΗΦΙΑΚΗ ΥΠΟΓΡΑΦΗ]

Δημήτρης Γκλώτσος

Αν. Καθηγητής

[ΨΗΦΙΑΚΗ ΥΠΟΓΡΑΦΗ]

Παντελής Ασβεστάς

Αν. Καθηγητής

[ΨΗΦΙΑΚΗ ΥΠΟΓΡΑΦΗ]

ΔΗΛΩΣΗ ΣΥΓΓΡΑΦΕΑ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ

Ο υπογράφων Γεώργιος Ιάσων Χριστουλάκης του Κωνσταντίνου, με αριθμό μητρώου 48017107 φοιτητής του Τμήματος Μηχανικών Βιοϊατρικής της Σχολής Μηχανικών του Πανεπιστημίου Δυτικής Αττικής, δηλώνω υπεύθυνα ότι:

«Είμαι συγγραφέας αυτής της διπλωματικής εργασίας και κάθε βοήθεια την οποία είχα για την προετοιμασία της είναι πλήρως αναγνωρισμένη και αναφέρεται στην εργασία. Επίσης, οι όποιες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών ή λέξεων, είτε ακριβώς είτε παραφρασμένες, αναφέρονται στο σύνολό τους, με πλήρη αναφορά στους συγγραφείς, τον εκδοτικό οίκο ή το περιοδικό, συμπεριλαμβανομένων και των πηγών που ενδεχομένως χρησιμοποιήθηκαν από το διαδίκτυο. Επίσης, βεβαιώνω ότι αυτή η εργασία έχει συγγραφεί από μένα αποκλειστικά και αποτελεί προϊόν πνευματικής ιδιοκτησίας τόσο δικής μου, όσο και του Ιδρύματος.

Παράβαση της ανωτέρω ακαδημαϊκής μου ευθύνης αποτελεί ουσιώδη λόγο για την ανάκληση του διπλώματός μου».

Ημερομηνία

23-9-2022

Ο/Η Δηλών



Περίληψη

Εισαγωγή: Τα χαμηλού βαθμού γλοιώματα (LGG) στον εγκέφαλο είναι όγκοι του κεντρικού νευρικού συστήματος που παρουσιάζουν διακύμανση στην επιθετικότητα της ανάπτυξης. Ο μαγνητικός συντονισμός είναι η μέθοδος απεικόνισης εκλογής για τη διάγνωση των LGG. Τα χαρακτηριστικά του όγκου μπορεί να παρέχουν σημαντικές πληροφορίες πριν από τη χειρουργική επέμβαση και όταν η εκτομή δεν είναι δυνατή. Κατά συνέπεια, η αναγνώριση και η ακριβής κατάτμηση των όγκων παρέχουν τη βάση για περαιτέρω ανάλυση. Πρόσφατα, οι μέθοδοι βαθιάς μάθησης (DL) παρείχαν τα εργαλεία για την αυτόματη αναγνώριση και τμηματοποίηση των όγκων του εγκεφάλου.

Υλικό και μέθοδοι: Πειραματιστήκαμε με ένα σύνολο δεδομένων MRI εγκεφάλου LGG που λήφθηκε από το kaggle.com ως "LGG Segmentation Dataset" και αποτελείται από 3929 εικόνες MRI σε ακολουθία T2-FLAIR, που περιέχουν εικόνες με όγκους, μη όγκους και τις αντίστοιχες μάσκες με τους περιγεγραμμένους όγκους εγκεκριμένες από επιτροπή ραδιολόγων στο Πανεπιστήμιο Duke. Εφαρμόσαμε τρία διαφορετικά μοντέλα DL για να εκτελέσουμε τμηματοποίηση εικόνας βάσει pixel. τα μοντέλα U-NET, Residual U-NET και Attention U-NET. Τέλος, και προκειμένου να βελτιώσουμε την απόδοση τμηματοποίησης, χρησιμοποιήσαμε ένα μοντέλο συνόλου (EDL) χρησιμοποιώντας τεχνική hard voting. Οι εικόνες χωρίστηκαν για 70% εκπαίδευση και 10% test ενώ το 20% έμεινε εκτός για σκοπούς validaation, έτσι ώστε να εκτιμηθεί η ικανότητα γενίκευσης της προτεινόμενης μεθόδου. Όλα τα μοντέλα εκπαιδεύτηκαν σε cloud server με ισχυρές GPU.

Αποτελέσματα: Η μέθοδος μας εντοπίζει την παρουσία όγκου LGG κατά 97,6% ενώ υστερεί σε ποσοστό 1,5% στο validation set. Όσο αφορά την κατάτμηση, το μοντέλο EDL ξεπερνά τα υπόλοιπα μοντέλα που δοκιμάστηκαν όσο αφορά τα μετρικά Precision, Jaccard και Dice με 0.88 ± 0.15 , 0.74 ± 0.19 και 0.83 ± 0.16 αντίστοιχα, που είναι συγκρίσιμο, αν όχι ελαφρώς υψηλότερο, με τις προηγούμενες δημοσιευμένες μελέτες.

Συμπέρασμα: Παρά την υψηλή ετερογένεια των όγκων, η προτεινόμενη μέθοδος προσδιορίζει σωστά την παρουσία ενός όγκου LGG και τον τμηματοποιεί με υψηλή ακρίβεια σε άγνωστα δεδομένα.

Λέξεις Κλειδιά: χαμηλού βαθμού γλοιώμα; Μαγνητική Τομογραφία(MRI); Τμηματοποίηση; U-NET; Βαθιά Μάθηση

Abstract

Background: Low-grade gliomas (LGGs) in the brain are tumors of the central nervous system that exhibit variance in growth aggressiveness. Magnetic resonance is the imaging modality of choice for the diagnosis of LGGs. Tumor's characteristics may give important information before surgery and when resection is not possible. Consequently, identification and accurate segmentation of tumors provide the basis for further analysis. Recently, deep learning (DL) methods have provided the tools for automatic identification and segmentation of brain tumors.

Material and Methods: We have experimented with a LGG brain MRI dataset that was downloaded from kaggle.com as “LGG Segmentation Dataset”, and consists of 3929 MRI images in T2-FLAIR sequence, containing images with tumors, non-tumors and the corresponding manually delineated abnormality masks approved by a board-certified radiologist at Duke University. We applied three different DL models to perform pixel-based image segmentation; the U-NET, the Residual U-NET and the Attention U-NET DL models. Finally, and in order to improve segmentation performance we have employed an ensemble model (EDL) using hard voting. The images were split for 70% training and 10% testing, while 20% were left out for validation purposes, so as to estimate the generalization capability of the proposed method. All models were trained on cloud server with powerful GPUs.

Results: Our method identifies the presence of an LGG tumor by 97.6% while it misses by 1.5% rate in validation stage. Concerning segmentation, the EDL model outperforms the rest of the models tested in terms of the Precision, Jaccard and Dice metrics with 0.88 ± 0.15 , 0.74 ± 0.19 and 0.83 ± 0.16 respectively that is comparable, if not slightly higher, to the previously published studies.

Conclusion: Despite the high heterogeneity of the tumors, the proposed method identifies correctly the presence of an LGG tumor and it segmented it by high precision in unseen data.

https://github.com/iasonasxrist/MRI_Brain_Segmentation

Ευχαριστίες

Ευχαριστώ πολύ τον επιβλέποντα καθηγητή μου, κ. Σπύρο Κωστόπουλο για την καθοδήγηση του σε όλη την πορεία της εργασίας. Επίσης τον κ. Θωμά Μαρή του τμήματος Ιατρικής του Πανεπιστημίου Κρήτης για την πολύτιμη βοήθεια του.

Τέλος ευχαριστώ την οικογένεια μου και τους φίλους μου για την πνευματική υποστήριξη.

Περιεχόμενα

ΠΕΡΙΛΗΨΗ	- 4 -
ABSTRACT	- 5 -
ΠΕΡΙΕΧΟΜΕΝΑ	- 7 -
ΚΑΤΑΛΟΓΟΣ ΕΙΚΟΝΩΝ	- 9 -
ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ	- 12 -
ΘΕΩΡΗΤΙΚΟ ΜΕΡΟΣ	- 13 -
1. ΕΙΣΑΓΩΓΗ	- 13 -
2. ΕΓΚΕΦΑΛΟΣ	- 15 -
2.1 ΓΕΝΙΚΑ	- 15 -
2.2 ΑΝΑΤΟΜΙΑ ΕΓΚΕΦΑΛΟΥ	- 16 -
2.3 ΚΥΤΤΑΡΑ ΤΟΥ ΕΓΚΕΦΑΛΟΥ.....	- 17 -
2.4 ΕΓΚΕΦΑΛΙΚΟΙ ΌΓΚΟΙ	- 18 -
3. ΜΑΓΝΗΤΙΚΟΣ ΣΥΝΤΟΝΙΣΜΟΣ	- 20 -
3.1 ΑΡΧΗ ΛΕΙΤΟΥΡΓΙΑΣ	- 20 -
3.2 ΜΑΓΝΗΤΙΚΗ ΤΟΜΟΓΡΑΦΙΑ	- 21 -
3.3 ΠΡΩΤΟΚΟΛΛΑ	- 22 -
4. ΤΜΗΜΑΤΟΠΟΙΗΣΗ ΕΙΚΟΝΑΣ	- 27 -
4.1 ΜΕΘΟΔΟΙ ΕΠΕΞΕΡΓΑΣΙΑΣ ΙΑΤΡΙΚΗΣ ΕΙΚΟΝΑΣ	- 27 -
4.2 ΑΝΙΧΝΕΥΣΗ ΑΝΤΙΚΕΙΜΕΝΩΝ ΚΑΙ ΤΜΗΜΑΤΟΠΟΙΗΣΗ	- 29 -
4.3 ΜΕΤΡΑ ΑΞΙΟΛΟΓΗΣΗΣ ΤΜΗΜΑΤΟΠΟΙΗΣΗΣ.....	- 32 -
5. ΤΕΧΝΗΤΗ ΝΟΗΜΟΣΥΝΗ	- 35 -
5.1 ΚΑΤΗΓΟΡΙΕΣ ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ.....	- 35 -
5.2 ΕΠΟΠΤΕΥΟΜΕΝΗ ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ	- 36 -
5.3 ΜΗ-ΕΠΟΠΤΕΥΟΜΕΝΗ ΜΗΧΑΝΙΚΗ ΜΑΘΗΣΗ	- 37 -
5.4 ΒΑΘΙΑ ΜΑΘΗΣΗ ΚΑΙ ΣΥΝΕΛΙΚΤΙΚΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ.....	- 38 -
5.5 FINE-TUNING ΣΤΑ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ.....	- 43 -
6. ΝΕΥΡΩΝΙΚΟ ΔΙΚΤΥΟ UNET	- 49 -
6.1 U-NET	- 49 -
6.2 ATTENTION UNET	- 51 -

6.3 RESIDUAL U-NET	- 53 -
6.4 ENSEMBLE MODEL	- 55 -
6.5 ΣΥΝΑΡΤΗΣΗ ΑΠΩΛΕΙΑΣ	- 56 -
6.6 ΤΟ ΠΡΟΒΛΗΜΑ ΤΟΥ VANISHING GRADIENT	- 57 -
6.7 ΜΟΝΑΔΑ ΕΠΕΞΕΡΓΑΣΙΑΣ ΓΡΑΦΙΚΩΝ (GPU).....	- 59 -
ΠΕΙΡΑΜΑΤΙΚΟ ΜΕΡΟΣ	- 61 -
7. ΔΕΔΟΜΕΝΑ ΚΑΙ ΠΡΟΕΠΕΞΕΡΓΑΣΙΑ.....	- 61 -
7.1 ΜΟΡΦΕΣ ΙΑΤΡΙΚΩΝ ΔΕΔΟΜΕΝΩΝ.....	- 61 -
7.2 DATASET	- 61 -
7.3 ΠΡΟΕΠΕΞΕΡΓΑΣΙΑ	- 68 -
7.4 ΕΚΠΑΙΔΕΥΣΗ ΜΟΝΤΕΛΩΝ.....	- 70 -
8. ΑΠΟΤΕΛΕΣΜΑΤΑ.....	- 82 -
8.1 ΚΑΤΑΝΟΗΣΗ ΤΩΝ ΑΠΟΦΑΣΕΩΝ ΣΤΟ DEEP LEARNING	- 90 -
9. ΣΥΜΠΕΡΑΣΜΑΤΑ-ΣΥΖΗΤΗΣΗ	- 93 -
ΒΙΒΛΙΟΓΡΑΦΙΑ	- 94 -

Κατάλογος Εικόνων

ΕΙΚΟΝΑ 1. ΒΑΣΙΚΗ ΔΙΑΙΡΕΣΗ ΕΓΚΕΦΑΛΟΥ	- 15 -
ΕΙΚΟΝΑ 2. ΛΟΒΟΙ ΤΕΛΕΓΚΕΦΑΛΟΥ	- 16 -
ΕΙΚΟΝΑ 3. ΑΝΑΤΟΜΙΑ ΝΕΥΡΩΝΑ	- 18 -
ΕΙΚΟΝΑ 4. ΕΓΚΕΦΑΛΙΚΗ ΒΛΑΒΗ	- 18 -
ΕΙΚΟΝΑ 5. ΧΡΟΝΟΣ T1 ΣΕ ΔΙΑΦΟΡΕΤΙΚΟΥΣ ΙΣΤΟΥΣ	- 21 -
ΕΙΚΟΝΑ 6. Η ΠΡΩΤΗ ΛΗΨΗ ΑΠΟ ΜΑΓΝΗΤΙΚΟ ΤΟΜΟΓΡΑΦΟ ^[26]	- 22 -
ΕΙΚΟΝΑ 7. T1-W ΕΙΚΟΝΑ	- 22 -
ΕΙΚΟΝΑ 8. ΜΑΓΝΗΤΙΚΗ ΤΟΜΟΓΡΑΦΙΑ ΕΓΚΕΦΑΛΟΥ ΜΕ ΔΕΙΚΤΕΣ ΣΤΗΝ ΠΑΘΟΛΟΓΙΑ. Η T1 ΦΑΙΝΕΤΑΙ ΦΥΣΙΟΛΟΓΙΚΗ. ΠΑΡΟΛΑ ΑΥΤΑ ΣΤΗΝ T2 ΕΙΝΑΙ ΕΜΦΑΝΗΣ Η ΥΨΗΛΗ ΕΝΤΑΣΗ ΣΗΜΑΤΟΣ ΤΟΣΟ ΣΤΗΝ ΕΓΚΑΡΣΙΑ ΟΣΟ ΚΑΙ ΣΤΗ ΑΞΟΝΙΚΗ ΑΠΕΙΚΟΝΙΣΗ. ^[22]	- 23 -
ΕΙΚΟΝΑ 9. ΣΗΜΑ ΑΝΑΣΤΡΟΦΗΣ ΑΝΆΚΤΗΣΗΣ	- 26 -
ΕΙΚΟΝΑ 10. ΦΙΛΤΡΑ ΑΝΙΧΝΕΥΣΗΣ ΑΚΜΩΝ ROBERTS ΚΑΙ SOBEL	- 28 -
ΕΙΚΟΝΑ 11. ΦΙΛΤΡΑ ΑΝΙΧΝΕΥΣΗΣ ΑΚΜΩΝ SCHARR, PREWITT	- 28 -
ΕΙΚΟΝΑ 12. ΑΡΙΣΤΕΡΑ ΚΑΝΟΝΙΚΗ, ΔΕΞΙΑ ΦΙΛΤΡΟ ΑΝΙΧΝΕΥΣΗΣ ΑΚΜΩΝ CANNY	- 28 -
ΕΙΚΟΝΑ 13. ΜΕΘΟΔΟΣ ΚΑΤΩΦΛΙΩΣΗΣ OTSU ^[27]	- 29 -
ΕΙΚΟΝΑ 14. FULLY CONVOLUTIONAL LAYER FOR SEMANTIC SEGMENTATION ^[5]	- 30 -
ΕΙΚΟΝΑ 15. ΤΜΗΜΑΤΟΠΟΙΗΣΗ ΠΟΛΛΑΠΛΩΝ ΟΡΓΑΝΩΝ (MULTI-ORGAN SEGMENTATION) ^[6]	- 31 -
ΕΙΚΟΝΑ 16. ΠΑΡΑΔΕΙΓΜΑΤΑ ΕΦΑΡΜΟΓΗΣ ΤΜΗΜΑΤΟΠΟΙΗΣΗΣ ΣΤΗΝ ΙΑΤΡΙΚΗ ΑΠΕΙΚΟΝΙΣΗ ΜΕ ΤΗ ΧΡΗΣΗ DEEP LEARNING ΣΕ ΓΝΩΣΤΟΥΣ ΔΙΑΓΩΝΙΣΜΟΥΣ ΜΕ ΜΕΓΑΛΑ ΠΟΣΟΣΤΑ ΕΠΙΤΥΧΙΑΣ. ΜΕ ΤΗΝ ΣΕΙΡΑ ΑΠΟ ΠΑΝΩ ΠΡΟΣ ΤΑ ΚΑΤΩ ΚΑΙ ΑΠΟ ΑΡΙΣΤΕΡΑ ΠΡΟΣ ΤΑ ΔΕΞΙΑ ΤΑΞΙΝΟΜΗΣΗ ΜΑΖΑΣ ΜΑΣΤΟΓΡΑΦΙΑΣ ^[57] , ΤΜΗΜΑΤΟΠΟΙΗΣΗ ΒΛΑΒΩΝ ΕΓΚΕΦΑΛΟΥ (BRATS) ^[59] , ΑΝΙΧΝΕΥΣΗ ΔΙΑΡΡΟΗΣ ΣΤΗΝ ΚΑΤΑΤΜΗΣΗ ΔΕΝΤΡΩΝ ΑΕΡΑΓΩΓΩΝ ^[58] , ΔΙΑΒΗΤΙΚΗ ΑΜΦΙΒΛΗΣΤΡΟΕΙΔΟΠΑΘΕΙΑ (KAGGLE CHALLENGE 2015) ^[60] , ΤΜΗΜΑΤΟΠΟΙΗΣΗ ΑΔΕΝΑ ΠΡΟΣΤΑΤΗ (PROMISE12) ^[61] ΚΑΙ ΤΜΗΜΑΤΟΠΟΙΗΣΗ ΟΖΩΝ (LUNA16CHALLENGE) ^[62]	- 31 -
ΕΙΚΟΝΑ 17. ΟΠΤΙΚΟΠΟΙΗΣΗ ΤΟΥ ΠΙΝΑΚΑ ΑΛΗΘΕΙΑΣ ^[49]	- 32 -
ΕΙΚΟΝΑ 18. ΣΧΗΜΑΤΙΚΗ ΑΝΑΠΑΡΑΣΤΑΣΗ ΤΟΥ JACCARD ^[4]	- 34 -
ΕΙΚΟΝΑ 19. ΣΥΝΤΕΛΕΣΤΗΣ DICE ^{[45], [46]}	- 34 -
ΕΙΚΟΝΑ 20. ΥΠΟΚΑΤΗΓΟΡΙΕΣ ARTIFICIAL INTELLIGENCE ^[78]	- 35 -
ΕΙΚΟΝΑ 21. ΠΑΡΑΔΕΙΓΜΑΤΑ ΤΑΞΙΝΟΜΗΣΗΣ ΚΑΙ ΠΑΛΙΝΔΡΟΜΗΣΗΣ ^[3]	- 37 -
ΕΙΚΟΝΑ 22. ΠΑΡΑΔΕΙΓΜΑ ΣΥΣΤΑΔΟΠΟΙΗΣΗΣ	- 38 -
ΕΙΚΟΝΑ 23. ΑΠΛΟΣΤΕΡΟ ΝΕΥΡΩΝΙΚΟ ΔΙΚΤΥΟ PERCEPTRON ^[79]	- 40 -
ΕΙΚΟΝΑ 24. ΑΝΑΠΑΡΑΣΤΑΣΗ ΝΕΥΡΩΝΙΚΟΥ ΔΙΚΤΥΟΥ ^[36]	- 43 -
ΕΙΚΟΝΑ 25. IMAGE TRANSLATION T1-W TO T2-W	- 43 -
ΕΙΚΟΝΑ 26. ΑΝΑΠΑΡΑΣΤΑΣΗ ΒΑΘΜΩΤΗΣ ΣΥΓΚΛΙΣΗΣ ^[37]	- 45 -
ΕΙΚΟΝΑ 27. ΠΕΙΡΑΜΑΤΙΣΜΟΣ ΣΤΟ ΡΥΘΜΟΣ ΕΚΜΑΘΗΣΗΣ	- 46 -
ΕΙΚΟΝΑ 28. ΑΡΙΣΤΕΡΑ ΜΕ ΤΗΝ ΚΑΝΟΝΙΚΟΠΟΙΗΣΗ, ΔΕΞΙΑ ΧΩΡΙΣ ΚΑΝΟΝΙΚΟΠΟΙΗΣΗ	- 47 -
ΕΙΚΟΝΑ 29. ΑΝΑΠΑΡΑΣΤΑΣΗ ΤΟΥ BATCH NORMALIZATION ΣΤΗΝ ΕΚΠΑΙΔΕΥΣΗ ^[37]	- 47 -
ΕΙΚΟΝΑ 30. ΠΑΡΑΔΕΙΓΜΑ ΧΡΗΣΗΣ ΤΟΥ DROPOUT ΣΤΟ DATASET ΤΟΥ MNIST ^[24]	- 48 -
ΕΙΚΟΝΑ 31. UNET ARCHITECTURE ^{[14], [29]}	- 50 -

ΕΙΚΟΝΑ 32. ΣΧΗΜΑΤΙΚΗ ΑΝΑΠΑΡΑΣΤΑΣΗ ΤΗΣ ΔΙΑΔΙΚΑΣΙΑΣ ΤΟΥ ATTENTION GATE.....	- 52 -
ΕΙΚΟΝΑ 33.ΑΝΑΠΑΡΑΣΤΑΣΗ ΤΟΥ ATTENTION GATE ^[73]	- 53 -
ΕΙΚΟΝΑ 34. ΔΙΑΦΟΡΕΣ ΠΑΡΑΛΛΑΓΕΣ ΤΟΥ RESIDUAL BLOCK, ΑΛΛΑΓΕΣ ΠΑΡΑΤΗΡΟΥΝΤΑΙ ΜΟΝΟ ΣΤΗΝ ΑΛΛΗΛΟΥΧΙΑ ΟΧΙ ΣΤΙΣ ΜΑΘΗΜΑΤΙΚΕΣ ΦΟΡΜΟΥΛΕΣ ΠΟΥ ΕΦΑΡΜΟΖΟΝΤΑΙ. ^[33]	- 55 -
ΕΙΚΟΝΑ 35. ΑΡΧΙΤΕΚΤΟΝΙΚΗ RESUNET	- 55 -
ΕΙΚΟΝΑ 36. ΑΝΑΠΑΡΑΣΤΑΣΗ ENSEMBLE MODEL	- 56 -
ΕΙΚΟΝΑ 37. Η ΤΜΗΜΑΤΟΠΟΙΗΣΗ ΤΗΣ ΠΕΡΙΟΧΗΣ ΕΝΔΙΑΦΕΡΟΝΤΟΣ ΓΙΝΕΤΑΙ ΒΑΣΕΙ ΤΗΣ ΠΙΘΑΝΟΤΗΤΑΣ ΤΟΥ PIXEL (PI) P: PREDICTION, G: GROUND ^{[35], [36]}	- 57 -
ΕΙΚΟΝΑ 38. ΠΑΡΑΓΩΓΙΣΗ ΤΗΣ ΛΟΓΙΣΤΙΚΗ ΣΥΝΑΡΤΗΣΗ ΠΟΥ ΟΔΗΓΕΙ ΣΤΟ ΦΑΙΝΟΜΕΝΟ VANISHING GRADIENT	- 58 -
ΕΙΚΟΝΑ 39. ΠΑΡΟΥΣΙΑΣΗ ΤΩΝ NVIDIA GPUS ^[48]	- 60 -
ΕΙΚΟΝΑ 40. BRAIN MR IMAGES ΤΟΥ DATASET	- 63 -
ΕΙΚΟΝΑ 41. COUNTRPLOT ΓΙΑ ΤΗΝ ΚΑΤΑΝΟΜΗ ΤΩΝ ΚΛΑΣΕΩΝ.....	- 64 -
ΕΙΚΟΝΑ 42. ΠΕΡΙΓΡΑΦΗ ΤΩΝ ΔΕΔΟΜΕΝΩΝ ΣΧΕΤΙΚΑ ΜΕ ΤΗΝ ΘΕΣΗ ΤΗΣ ΠΑΘΟΛΟΓΙΑΣ, ΤΟ ΦΥΛΟ, ΤΗΝ ΕΞΕΛΙΞΗ ΤΗΣ ΝΟΣΟΥ ΚΑΙ ΤΗΝ ΗΛΙΚΙΑ ΕΜΦΑΝΙΣΗΣ ΤΗΣ ΠΑΘΟΛΟΓΙΑ.	- 64 -
ΕΙΚΟΝΑ 43. ΠΡΟΣΕΓΓΙΣΗ ΣΤΡΑΤΗΓΙΚΗΣ ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ ^[66]	- 67 -
ΕΙΚΟΝΑ 44. ΠΕΙΡΑΜΑΤΙΚΑ ΔΕΔΟΜΕΝΑ ΑΡΙΣΤΕΡΑ Η ΤΟΜΗ ΤΗΣ ΜΑΓΝΗΤΙΚΗΣ, ΔΙΠΛΑ Η ΜΑΣΚΑ ΤΗΣ ΤΟΜΗΣ, ΔΕΞΙΟΤΕΡΑ Η ΠΡΟΒΛΕΨΗ ΤΟΥ ΜΟΝΤΕΛΟΥ ΤΟΥ UNET ΚΑΙ ΔΕΞΙΟΤΕΡΑ Η ΕΠΙΠΤΩΣΕΙΣ ΤΗΣ ΚΑΤΩΦΛΙΩΣΗΣ ΣΤΙΣ ΕΙΚΟΝΕΣ. ΜΠΟΡΕΙ ΚΑΠΟΙΟΣ ΝΑ ΠΑΡΑΤΗΡΗΣΕΙ ΤΗΝ ΑΠΟΥΣΙΑ ΕΙΚΟΝΟΣΤΟΙΧΕΙΩΝ ΣΤΟΝ ΠΥΡΗΝΑ ΤΟΥ ΟΓΚΟΥ ΣΤΗΝ ΕΙΚΟΝΑ ΤΗΣ ΜΑΓΝΗΤΙΚΗΣ ΤΟΜΟΓΡΑΦΙΑΣ.....	- 69 -
ΕΙΚΟΝΑ 45. ΛΕΠΤΟΜΕΡΗΣ ΑΡΧΙΤΕΚΤΟΝΙΚΗ ΤΟΥ UNET. ^[51]	- 72 -
ΕΙΚΟΝΑ 46. ΓΡΑΦΗΜΑ ΤΗΣ ΣΥΝΑΡΤΗΣΗΣ ΑΠΩΛΕΙΑΣ ΚΑΤΑ ΤΟ TRAINING	- 73 -
ΕΙΚΟΝΑ 47. ΔΙΑΓΡΑΜΜΑΤΑ ΤΩΝ ΜΕΤΡΙΚΩΝ JACCARD, PRECISION ΚΑΙ DICE.....	- 73 -
ΕΙΚΟΝΑ 48. ΠΡΟΒΛΕΨΕΙΣ ΣΤΟ VALIDATION SET.....	- 74 -
ΕΙΚΟΝΑ 49. ΠΑΡΑΔΕΙΓΜΑ ΚΑΚΗΣ ΠΡΟΒΛΕΨΗΣ ΠΟΥ ΟΦΕΙΛΕΤΑΙ ΣΤΗΝ ΠΡΟΠΕΞΕΡΓΑΣΙΑ ΜΕ ΕΓΧΡΩΜΑ ΚΑΝΑΛΙΑ, ΣΤΗΝ ΕΙΚΟΝΑ ΔΕΝ ΥΠΑΡΧΕΙ ΚΑΠΟΙΑ ΠΑΘΟΛΟΓΙΑ.....	- 75 -
ΕΙΚΟΝΑ 50. ΣΧΗΜΑΤΙΚΗ ΑΝΑΠΑΡΑΣΤΑΣΗ ATTENTION U-NET ^[29]	- 77 -
ΕΙΚΟΝΑ 51. ΓΡΑΦΗΜΑ ΤΗΣ ΣΥΝΑΡΤΗΣΗΣ ΑΠΩΛΕΙΑΣ ΚΑΤΑ ΤΟ TRAINING	- 77 -
ΕΙΚΟΝΑ 52. ΔΙΑΓΡΑΜΜΑΤΑ ΤΩΝ ΜΕΤΡΙΚΩΝ JACCARD, PRECISION ΚΑΙ DICE.....	- 78 -
ΕΙΚΟΝΑ 53.ΠΡΟΒΛΕΨΕΙΣ ΜΕ ΤΙΜΗ ΣΥΝΤΕΛΕΣΤΗ DICE ΜΕΓΑΛΥΤΕΡΗ ΤΟΥ 0.8	- 78 -
ΕΙΚΟΝΑ 54. ΑΝΑΠΑΡΑΣΤΑΣΗ ΤΗΣ ΑΡΧΙΤΕΚΤΟΝΙΚΗΣ ΤΟΥ RESUNET. ΜΕ ΓΑΛΑΖΙΑ ΔΙΑΚΕΚΟΜΜΕΝΗ ΤΑ RES BLOCKS. ^{[32], [33], [34]}	- 80 -
ΕΙΚΟΝΑ 55. ΓΡΑΦΗΜΑ ΤΗΣ ΣΥΝΑΡΤΗΣΗΣ ΑΠΩΛΕΙΑΣ ΚΑΤΑ ΤΟ TRAINING	- 80 -
ΕΙΚΟΝΑ 56. ΔΙΑΓΡΑΜΜΑΤΑ JACCARD , PRECISION, DICE	- 81 -
ΕΙΚΟΝΑ 57. ΠΡΟΒΛΕΨΕΙΣ ΣΤΟ VALIDATION SET ΓΙΑ ΤΟ ATTENTION UNET.....	- 81 -
ΕΙΚΟΝΑ 58. ΒΟΧPLOT UNET	- 83 -
ΕΙΚΟΝΑ 59. ΒΟΧPLOT ATTENTION UNET.....	- 83 -
ΕΙΚΟΝΑ 60. ΒΟΧPLOT RESUNET.....	- 84 -
ΕΙΚΟΝΑ 61. ΒΟΧPLOT ENSEMBLE MODEL	- 84 -
ΕΙΚΟΝΑ 62. (A-P) ΠΑΡΑΔΕΙΓΜΑΤΑ ΣΕ 16 ΕΙΚΟΝΕΣ ΜΕ ΤΑ ΑΝΤΙΣΤΟΙΧΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΜΗΜΑΤΟΠΟΙΗΣΗΣ ΓΙΑ ΚΑΘΕ ΕΝΑ ΑΠΟ ΤΑ ΔΟΚΙΜΑΖΟΜΕΝΑ ΔΙΚΤΥΑ.....	- 86 -

ΕΙΚΟΝΑ 63. (A-G) ΑΡΙΣΤΕΡΑ Η ΤΟΜΗ, ΣΤΗ ΕΠΟΜΕΝΗ ΕΧΟΥΜΕ ΥΠΕΡΘΕΣΕΙ ΤΗΝ ΜΑΣΚΑ ΓΙΑ ΣΑΦΗ ΠΡΟΒΟΛΗ, ΔΕΞΙΟΤΕΡΑ ΑΠΕΙΚΟΝΙΖΕΤΑΙ Η ΠΡΟΒΛΕΨΗ ΤΟΥ EDL ΜΟΝΤΕΛΟΥ ΜΑΣ ΚΑΙ ΤΕΛΟΣ Η ΔΙΑΦΟΡΑ ΜΕΤΑΞΥ ΠΡΟΒΛΕΠΟΜΕΝΗΣ ΚΑΙ ΑΛΗΘΙΝΗΣ ΩΣΤΕ ΝΑ ΚΑΤΑΝΟΗΘΕΙ Η ΑΠΟΔΟΣΗ ΤΟΥ ΜΟΝΤΕΛΟΥ (A-G)	- 89 -
ΕΙΚΟΝΑ 64. ΠΑΡΑΔΕΙΓΜΑ ΚΑΚΗΣ ΠΡΟΒΛΕΨΗΣ ΛΟΓΩ ΑΣΤΟΧΙΑΣ ΤΟΥ DATASET	- 89 -
ΕΙΚΟΝΑ 65. Η ΕΠΙΔΡΑΣΗ ΤΟΥ ENSEMBLING : ΑΡΧΙΚΗ ΕΙΚΟΝΑ ΟΠΟΥ ΔΕΝ ΕΙΧΕ ΔΟΘΕΙ ΜΑΣΚΑ ΠΑΘΟΛΟΓΙΑΣ (NORMAL) ΚΑΙ ΤΑ ΑΝΤΙΣΤΟΙΧΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΩΝ ΔΙΚΤΥΩΝ.	- 90 -
ΕΙΚΟΝΑ 66. (A-F) ΑΡΙΣΤΕΡΑ Η ΕΙΚΟΝΑ ΚΑΙ ΔΕΞΙΑ Ο ΧΑΡΤΗΣ ΠΥΚΝΟΤΗΤΑΣ ΠΙΘΑΝΟΤΗΤΑΣ (GRAD-CAM)	- 92 -

Κατάλογος Πινάκων

ΠΙΝΑΚΑΣ 3.1 ΣΥΓΚΡΙΣΗ ΑΠΟΧΡΩΣΕΩΝ ΣΤΗΝ ΑΠΕΙΚΟΝΙΣΗ T1 ΚΑΙ T2	- 23 -
ΠΙΝΑΚΑΣ 3.2 ΡΥΘΜΙΣΕΙΣ ΠΑΡΑΜΕΤΡΩΝ ΓΙΑ ΤΗΝ ΛΗΨΗ ΤΗΣ ΜΑΓΝΗΤΙΚΗΣ ΤΟΜΟΓΡΑΦΙΑΣ (ACQUISITION).....	- 24 -
ΠΙΝΑΚΑΣ 3.3 ΠΑΡΑΔΕΙΓΜΑ ΧΡΟΝΩΝ ΑΠΟΚΑΤΑΣΤΑΣΗΣ ΓΙΑ ΤΗΝ SPIN ECHO ΑΚΟΛΟΥΘΙΑ	- 25 -
ΠΙΝΑΚΑΣ 3.4 ΠΑΡΑΔΕΙΓΜΑ ΧΡΟΝΩΝ ΑΠΟΚΑΤΑΣΤΑΣΗΣ ΓΙΑ ΤΗΝ INVERSION RECOVERY ΑΚΟΛΟΥΘΙΑ	- 26 -
ΠΙΝΑΚΑΣ 7.1 ΔΕΔΟΜΕΝΑ ΣΧΕΤΙΚΑ ΜΕ ΤΟ ΦΥΛΟ ΚΑΙ ΤΗΝ ΗΛΙΚΙΑ ΤΩΝ ΠΕΡΙΣΤΑΤΙΚΩΝ	- 65 -
ΠΙΝΑΚΑΣ 7.2 ΑΝΑΛΥΣΗ ΤΩΝ ΙΣΤΟΛΟΓΙΚΩΝ ΥΠΟΤΥΠΩΝ ΣΤΟ LGG DATASET	- 65 -
ΠΙΝΑΚΑΣ 7.3 ΥΠΟΤΥΠΟΙ ΚΑΙ ΒΑΘΜΟΙ ΚΑΚΟΗΘΕΙΑΣ.....	- 66 -
ΠΙΝΑΚΑΣ 7.4 ΣΤΡΑΤΗΓΙΚΗ ΧΩΡΙΣΜΟΥ ΤΟΥ DATASET.....	- 66 -
ΠΙΝΑΚΑΣ 7.5 ΑΝΑΛΥΣΗ ΤΩΝ ΜΕΤΡΙΚΩΝ ΒΑΣΕΙ ΤΗΣ ΠΡΟΣΑΡΜΟΓΗΣ ΤΗΣ ΤΙΜΗΣ ΚΑΤΩΦΛΙΩΣΗΣ.....	- 70 -
ΠΙΝΑΚΑΣ 7.6 MODEL TRAINABLE PARAMETERS.....	- 70 -
ΠΙΝΑΚΑΣ 7.7 ΛΕΠΤΟΜΕΡΗΣ ΠΑΡΟΥΣΙΑΣΗ ΤΩΝ ΥΠΕΡΠΑΡΑΜΕΤΡΩΝ ΠΟΥ ΧΡΗΣΙΜΟΠΟΙΗΘΗΚΑΝ ΣΤΟ ΠΕΙΡΑΜΑ.....	- 72 -
ΠΙΝΑΚΑΣ 7.8 ΛΕΠΤΟΜΕΡΗΣ ΠΑΡΟΥΣΙΑΣΗ ΤΩΝ ΥΠΕΡΠΑΡΑΜΕΤΡΩΝ ΤΟΥ ΠΕΙΡΑΜΑΤΟΣ	- 77 -
ΠΙΝΑΚΑΣ 7.9 ΛΕΠΤΟΜΕΡΗΣ ΠΑΡΟΥΣΙΑΣΗ ΤΩΝ ΥΠΕΡΠΑΡΑΜΕΤΡΩΝ ΠΟΥ ΧΡΗΣΙΜΟΠΟΙΗΘΗΚΑΝ ΣΤΟ ΠΕΙΡΑΜΑ.....	- 80 -
ΠΙΝΑΚΑΣ 8.1 ΑΠΟΤΕΛΕΣΜΑΤΑ UNET	- 82 -
ΠΙΝΑΚΑΣ 8.2 ΑΠΟΤΕΛΕΣΜΑΤΑ ATTENTION UNET	- 82 -
ΠΙΝΑΚΑΣ 8.3 ΑΠΟΤΕΛΕΣΜΑΤΑ RESUNET.....	- 82 -
ΠΙΝΑΚΑΣ 8.4 ΑΠΟΤΕΛΕΣΜΑΤΑ ENSEMBLE MODEL	- 82 -
ΠΙΝΑΚΑΣ 8.5: ΤΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΟΥ DICE ΜΕΤΡΙΚΟΥ ΓΙΑ ΚΑΘΕ ΜΙΑ ΑΠΟ ΤΙΣ ΕΙΚΟΝΕΣ ΤΟΥ ΠΑΡΑΔΕΙΓΜΑΤΟΣ (ΕΙΚΟΝΑ 62 A-P)..	- 87 -

Θεωρητικό Μέρος

1. Εισαγωγή

Η εξέλιξη των ακτινοδιαγνωστικών συστημάτων όπως η υπολογιστική τομογραφία (CT), η υπερηχογραφία (US) και ο μαγνητικού τομογράφος (MRI) έδωσε την δυνατότητα της λεπτομερούς απεικόνισης των ανατομικών δομών του ανθρώπινου σώματος. Η συνεχής βελτίωση των συστημάτων ως προς τους χρόνους λήψης, την ποιότητα της απεικόνισης, την αποτελεσματικότητα της εξέτασης ενίσχυσε την ακρίβεια της διάγνωσης και της θεραπείας. Παρόλα αυτά, προβλήματα συνεχίζουν να εμφανίζονται που οφείλονται σε τεχνικά χαρακτηριστικά των συστημάτων όπως ο θόρυβος, η ομοιογένεια, η τυποποίηση και οι ψευδενδείξεις που συνεπάγονται αναξιόπιστη κλινική διάγνωση.

Ταυτόχρονα η ανάπτυξη του τομέας της πληροφορικής που ασχολείται με την ανάπτυξη ευφυών συστημάτων που σχετίζονται με ανώτερες νοητικές δεξιότητες όπως η αντίληψη, ο συλλογισμός και η ενέργεια, όπως όρισε ο Patrick Henry Winston την τεχνητή νοημοσύνη, ανοίγει τους ορίζοντες για καινοτόμες εφαρμογές.^[5]

Οι τεχνικές της τεχνητής νοημοσύνης αποτελούν εργαλείο χρήσης ποικίλων πεδίων, με έμφαση στην επεξεργασία της ιατρικής εικόνας στα σύγχρονα συστήματα ακτινοδιαγνωστικής. Οι μέθοδοι επεξεργασίας των ιατρικών εικόνων απαιτούν εξειδικευμένες στρατηγικές που βασίζονται κυρίως σε αλγορίθμους μηχανικής μάθησης. Το πεδίο της μηχανικής μάθησης και ειδικότερα της βαθιάς μάθησης έφερε μια επανάσταση στους αλγορίθμους που αναλύουν τις ιατρικές εικόνες και συνεπώς στην ανάπτυξη λογισμικών που υπόσχονται αυτοματοποιημένες λύσεις στην εξαγωγή χαρακτηριστικών όπως τα ευρέως διαδεδομένα συνελκτικά νευρωνικά δίκτυα (CNNs).

Η υπολογιστικά καθοδηγούμενη διάγνωση (CADx) επιτυγχάνει την ανάλυση των δεδομένων και την εξαγωγή ποσοτικών χαρακτηριστικών συνεισφέροντας στη διάγνωση και τη θεραπεία. Έχοντας αποκωδικοποιήσει σε ένα βαθμό τα απαραίτητα στάδια και έχοντας σχεδιάσει αλγορίθμους για να εκτελούνται από τους υπολογιστές επιλύσαμε ένα σημαντικό πρόβλημα. Πλέον μεθοδεύοντας την ανάπτυξη των συστημάτων υποστήριξης απόφασης, περιορίζουμε την υποκειμενικότητα και προσφέρουμε αυτοματοποιημένες λύσεις. Αναλογιζόμενοι τόσο της έλλειψη των ιατρικών πόρων για την παροχή των ιατρικών υπηρεσιών σε ακριτικές περιοχές όσο και το φόρτο εργασίας που απαιτεί μια ολοκληρωμένη ιατρική υπηρεσία.

Επομένως, συνδυάζοντας την όραση υπολογιστή με την βαθιά μάθηση μπορούμε να εκπαιδεύσουμε νευρωνικά δίκτυα που θα εκτελούν την εργασία της τμηματοποίησης σε εικόνες μαγνητικής τομογραφίας με μεγάλη ακρίβεια.

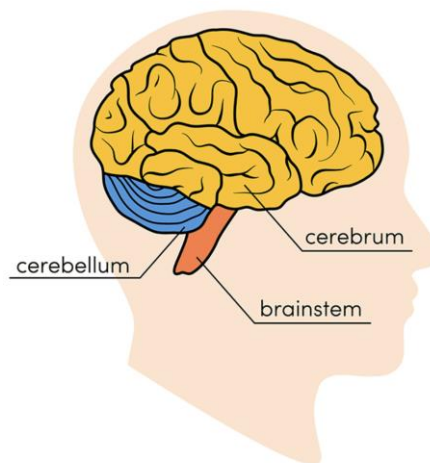
Εκπαιδεύοντας μεγάλο όγκο δεδομένων μπορούμε να δημιουργήσουμε ένα μοντέλο ικανό να τμηματοποιήσει μία περιοχή όπως κάποιος όγκος ανεξαρτήτως σχήματος ή διάστασης. Η ικανότητα του ως προς την τμηματοποίηση πρέπει να κριθεί σε ένα πραγματικό περιβάλλον και όχι στο «ιδανικό» περιβάλλον του εργαστηρίου. Οπότε χρησιμοποιούμε ποικίλες τεχνικές ώστε να επιτευχθεί ο παραπάνω στόχος, καθώς φαινόμενα υπερπροσαρμογής εμφανίζονται συχνά. Ποικίλα μοντέλα αναπτύχθηκαν, αξιολογώντας με διάφορα κριτήρια την αποτελεσματικότητά τους, καθώς τελικός σκοπός μας είναι η ανάπτυξη ενός εύρωστου μοντέλου για το κλινικό περιβάλλον. Για το σκοπό αυτό συχνά χρησιμοποιούμε μοντέλα συνδυαστικά προσεγγίζοντας ένα πρόβλημα από διαφορετικές οπτικές γωνίες, αυξάνοντας την πιθανότητα για μια ικανοποιητική προσέγγιση. ^[25]

2. Εγκέφαλος

2.1 Γενικά

Ο εγκέφαλος είναι το πιο πολύπλοκο μέρος του ανθρώπινου σώματος. Αποτελεί το κέντρο της νοημοσύνης, ο ερμηνευτής των αισθήσεων, ο εκκινητής των κινήσεων, και ο οργανωτής των διεργασιών του σώματος. Ευρισκόμενος μέσα στο κρανίο, περικλειόμενος από τρεις μήνιγγες και το εγκεφαλονωτιαίο υγρό αποτελεί την πηγή όλων των ανθρώπινων ιδιοτήτων. Θα μπορούσαμε να παρομοιάσουμε τον εγκέφαλο με μία επιτροπή από ειδικούς, καθώς όλα τα μέρη δρουν ταυτόχρονα και συνήθως συνδυαστικά βάσει των ιδιοτήτων τους. Ο εγκέφαλος διαιρείται σε τρία βασικά μέρη :α) τον πρόσθιο εγκέφαλο, β) τον μεσεγκέφαλο και γ) τον οπίσθιο εγκέφαλο (βλέπε Εικόνα 1). Ο οπίσθιος εγκέφαλος περιλαμβάνει το άνω μέρος του νωτιαίου μυελού, το εγκεφαλικό στέλεχος και την παρεγκεφαλίδα. Ο οπίσθιος εγκέφαλος είναι υπεύθυνος για τις ζωτικές λειτουργίες όπως ο ρυθμός αναπνοής, ο καρδιακός ρυθμός . Η παρεγκεφαλίδα συντονίζει την κίνηση και συμμετέχει σε διαδικασίες μάθησης. Παραδείγματος χάριν όταν αθλούμαστε, η παρεγκεφαλίδα είναι σε διέγερση. Το ανώτερο τμήμα του εγκεφαλικού στελέχους είναι ο μεσεγκέφαλος, εκεί εδράζεται το κέντρο των αντανακλαστικών, εμπλέκεται στο κύκλωμα για τη κίνηση των ματιών και γενικότερα συμβάλει μαζί με άλλα τμήματα για πλήθος εκκούσιων κινήσεων.

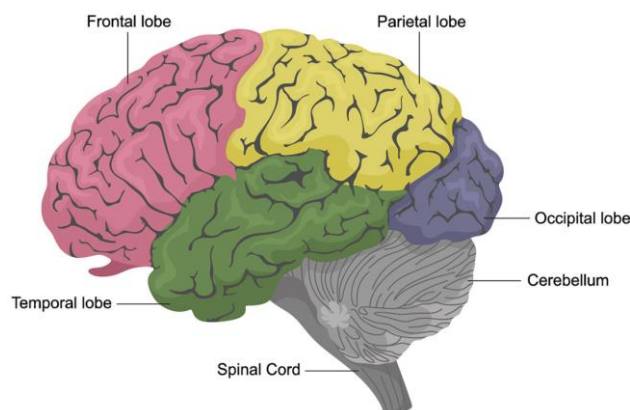
Τέλος, το μεγαλύτερο τμήμα και πιο αναπτυγμένο αποτελεί ο πρόσθιος εγκέφαλος, ο οποίο αποτελείται από τον εγκέφαλο, ο οποίο αντιπροσωπεύει τα 2/3 της μάζας του εγκέφαλου και καλύπτει τις περισσότερες δομές του. ^{[41],[42]}



Εικόνα 1. Βασική διαίρεση εγκεφάλου

2.2 Ανατομία Εγκεφάλου

Ο πρόσθιος εγκέφαλος ως το μεγαλύτερο τμήμα του εγκεφάλου χωρίζεται περαιτέρω σε τελεγκέφαλο και διεγκέφαλο. Ο τελεγκέφαλος αποτελεί τον εγκεφαλικό φλοιό ο οποίος χωρίζεται περαιτέρω σε λοβούς, τους μετωπιαίους λοβούς, τους βρεγματικούς λοβούς, τους ινιακούς λοβούς και τους κροταφικούς λοβούς (βλέπε Εικόνα 2). Ο μετωπιαίος λοβός (frontal lobe) είναι υπεύθυνος για την μνήμη, την κίνηση, το συναίσθημα, και την λήψη αποφάσεων. Ο βρεγματικός λοβός (parietal lobe) ερμηνεύει τις λέξεις, την αίσθηση της αφής και του πόνου, την χωρική-οπτική αντίληψη και ερμηνεύει κινητικά ή αισθητικά σήματα. Ο κροταφικός λοβός (Temporal lobe), κατανοεί την γλώσσα, συνεισφέρει στην ακοή, την μνήμη, την ακολουθία και την οργάνωση. Τέλος, ο ινιακός λοβός (occipital lobe) αποτελεί κέντρο λήψης και επεξεργασίας των οπτικών σημάτων από τον αμφιβληστροειδή. Οι λοβοί δρουν συνδυαστικά οπότε είναι λογικό να συμβάλλουν στον έλεγχο κάποιων διεργασιών.



Εικόνα 2. Λοβοί τελεγκέφαλου

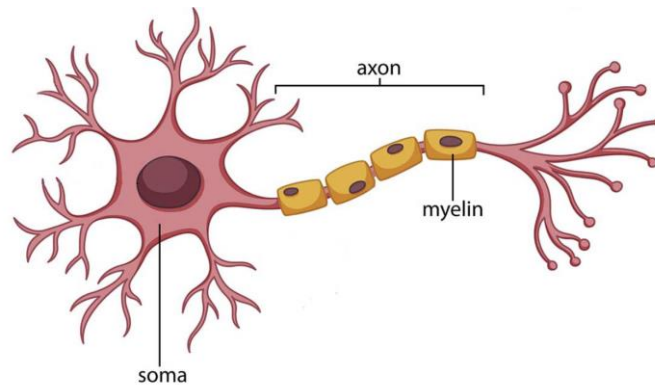
Ο εγκεφαλικός φλοιός περιέχει 16 δισεκατομμύρια νευρώνες (70 δισεκατομμύρια νευρώνες έχει η παρεγκεφαλίδα άρα συνολικά 86 δις) που είναι διατεταγμένοι σε επίπεδα. Τα σώματα των νευρώνων προσδίδουν το γκριζό χρώμα στο φλοιό, στη φαιά ουσία. Στα ενδότερα υπάρχουν μακρές νευρικές ίνες που συνδέουν τις περιοχές μεταξύ τους προσδίδονται την αλληλεπίδραση σε όλες το μήκος του εγκέφαλου, η λευκή ουσία. Ο εγκεφαλικός φλοιός περιέχει αυλακώσεις και έλικες, οι οποίες του προσδίδουν μια αναδιπλωμένη εμφάνιση. Αυτή η ιδιότητα αυξάνει την επιφάνεια του εγκεφάλου, άρα την χωρητικότητα του.

Ο διεγκέφαλος αποτελεί την γέφυρα διασύνδεσης του ενδοκρινικού συστήματος με το νευρικό σύστημα και μεταβιβάζει αισθητηριακές πληροφορίες. Ρυθμίζει ένα σύνολο λειτουργιών όπως οι αυτόνομες, οι ενδοκρινείς και οι κινητικές.

Αποτελείται από πολύ σημαντικές εν τω βάθι δομές όπως ο θάλαμος, ο υποθάλαμος και η επίφυση. Ο θάλαμος, αποτελεί σταθμό αναμετάδοσης για όλες σχεδόν τις πληροφορίες που καταλήγουν στο φλοιό. Επίσης παίζει ρόλο στον έλεγχο των κύκλων του ύπνου και στην αφύπνιση. Ο υποθάλαμος, αποτελεί περιοχή ελέγχου πολλών ζωτικών λειτουργιών όπως της αναπνοής, της αρτηριακής πίεσης και της ρύθμισης της θερμοκρασίας. Εκκρίνει ορμόνες που δρουν καταλυτικά για την ρύθμιση βιολογικών διεργασιών στην υπόφυση. Επίσης, επηρεάζει διάφορες συναισθηματικές αποκρίσεις μέσω την επίδρασης του στην υπόφυση και στο αυτόνομο νευρικό σύστημα. Τέλος, η επίφυση συνδέεται άμεσα με τον υποθάλαμο και ελέγχει ενδοκρινείς αδένες, εκκρίνει ορμόνες που ελέγχουν την ανάπτυξη των οστών και των μυών και την ανταπόκριση στο στρες. ^{[41], [42], [43]}

2.3 Κύτταρα του Εγκεφάλου

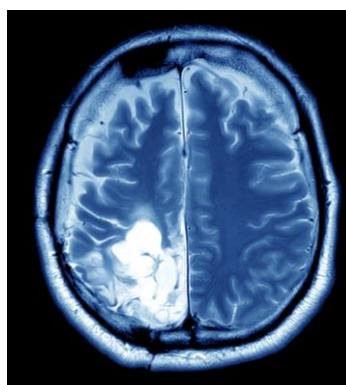
Τα κύτταρα του εγκεφάλου είναι δύο κατηγοριών τα νευρικά κύτταρα και τα κύτταρα γλοίας ή νευρογλοιακά. Τα νευρικά κύτταρα αποτελούν την κύρια λειτουργική μονάδα του κεντρικού νευρικού συστήματος και κατ' επέκταση του εγκεφάλου. Οι αναμνήσεις, οι κινήσεις, οι σκέψεις και τα συναισθήματα αποτελούν ηλεκτρικά σήματα τα οποία παράγονται, ενισχύονται και μεταβιβάζονται μέσω των νευρώνων. Η δομή του αποτελείται από τρία διακριτά μέρη: α) το σώμα, που περιέχει όλα τα απαραίτητα συστατικά για την επιβίωση και τη λειτουργία του ο νευρώνας, β) οι δενδρίτες, που λειτουργούν σαν κλαδιά ώστε να δέχονται μηνύματα από άλλους νευρώνες (προσυναπτικός νευρώνας) και γ) ο νευράξονας, που μεταβιβάζει αναλλοίωτα τα ηλεκτρικά σήματα στο μετασυναπτικό νευρώνα ώστε να συνεχιστεί η διάδοση του. Σε όλο το μήκος του νευράξονα απαντώνται τα υποστηρικτικά κύτταρα που αναφέρθηκαν παραπάνω, τα νευρογλοιακά. Είναι υψίστης σημασίας καθώς προσδίδουν στους νευρώνες μόνωση μέσω την τύλιξης τους γύρω από το νευράξονα, εξασφαλίζουν την διατήρηση της έντασης της νευρικής ώσης σε μεγάλες αποστάσεις. Το λιπώδες μόριο που τυλίγεται ονομάζεται μυελίνη. Τέλος, υπάρχουν κάποιες ασυνέχειες στις θήκες μυελίνης, οι κόμβοι Ranvier, οι οποίοι συμβάλλουν στην αναγέννηση της νευρικής ώσης. (βλέπε Εικόνα 3) ^{[42], [43]}



Εικόνα 3. Ανατομία Νευρώνα

2.4 Εγκεφαλικοί Όγκοι

Ένας εγκεφαλικός όγκος προκύπτει από εκφυλισμό των εγκεφαλικών κυττάρων ο οποίος συνεπάγεται τον ανεξέλεγκτο πολλαπλασιασμό τους. Η ανάπτυξη τους ποικίλει σε σχήμα και μέγεθος και εξαρτάται από το σημείο. Οι όγκοι είναι πολύ πυκνοί και συνήθως πιέζουν δομές ή μειώνουν τον διάμεσο χώρο εμποδίζοντας τη διάχυση του νερού. Οι κύριοι τύποι των όγκων είναι οι καλοήθεις (μη-καρκινικοί) και οι κακοήθεις (καρκινικοί). Οι καλοήθεις συνήθως προέρχονται από άλλη περιοχή του σώματος εκτός εγκεφάλου, μεταστατικοί όγκοι, ενώ οι κακοήθεις αναπτύσσονται στον εγκέφαλο (βλέπε Εικόνα 4). Τα συμπτώματα συνήθως περιλαμβάνουν πονοκεφάλους, προβλήματα όρασης, επιληπτικές κρίσεις, εμετός. Άλλα συμπτώματα είναι η δυσκολίες στο περπάτημα ή την ομιλία. Παρόλα αυτά, ποικίλλουν και εξαρτώνται από το σημείο ανάπτυξης του όγκου. [28]



Εικόνα 4. Εγκεφαλική Βλάβη

2.4.1 Γλοίωμα

Το γλοίωμα είναι όγκος νευροεπιθηλιακής προέλευσης που προέρχονται από νευρογλοιακά κύτταρα ή πρόδρομες κυτταρικές ουσίες. Το γλοίωμα είναι μία περιεκτική ομαδοποίηση καθώς περιλαμβάνει δύο κύριους ιστοπαθολογικούς υποτύπους: το αστροκύτωμα και το ολιγοδενδρογλοίωμα. Για παράδειγμα, το αστροκύτωμα ταξινομείται σε 5 διαβαθμίσεις όπου στο Grade IV ο όγκος μετατρέπεται σε μεταστατικό Γλοιοβλάστωμα (GBM).

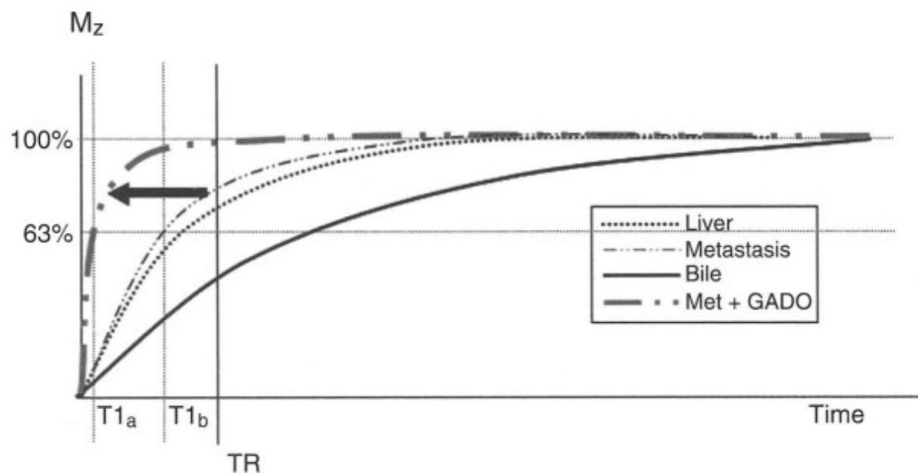
Τα γλοιώματα χωρίζονται στα υψηλού βαθμού (HGG) και στα χαμηλού βαθμού (LGG). Η διαφορά του έγκειται στον ρυθμό ανάπτυξης και την συχνότητα εμφάνισης σε νεότερους ενήλικες.

Τα αίτια εμφάνισης τους είναι άρρηκτα συνδεδεμένα με μεταλλάξεις γονιδίων. Η πιο συχνή μετάλλαξη είναι η IDH1/2. Επίσης συχνές μεταλλάξεις που συναντώνται είναι οι IDH, TP53 και ATRX στους χρωμοσωμικούς βραχίονες 1p/16q. [69] [70]

3. Μαγνητικός Συντονισμός

3.1 Αρχή Λειτουργίας

Ο μαγνητικός συντονισμός στηρίζεται στην εφαρμογή ηλεκτρομαγνητικών παλμών που οδηγούν στη διέγερση των πυρήνων υδρογόνου των ιστών. Αυτό οδηγεί στην εξαιρετικά λεπτομερή απεικόνιση των ιστών σε ατομικό επίπεδο αφού το ανθρώπινο σώμα αποτελείται από 60-70% νερό. Για να επιτραπεί η λήψη εικόνας απαιτείται μια ηλεκτρομαγνητική πηγή διέγερσης ώστε να μετατοπιστεί η παράλληλη μαγνήτιση με το πεδίο του υπεραγώγιμου μαγνήτη B. Μέσα στο μαγνητικό πεδίο B, το άθροισμα των spins των πυρήνων υδρογόνου σχηματίζουν την παράλληλη μαγνήτιση. Η πρόσθετη διέγερση ονομάζεται παλμός ραδιοσυχνότητας (RF) B₁. Η χρήση του RF παλμού επιτρέπει την αλληλεπίδραση του διανύσματος της μαγνήτισης με το πηνίο λήψης μέσω ηλεκτρομαγνητικής επαγωγής, το ρεύμα που παράγεται μας παρέχει πληροφορία για την ανατομική δομή του υπό εξέταση ασθενή. Η αρχή λειτουργίας του μαγνητικού συντονισμού NMR έγκειται στη χρήση του πρόσθετου πεδίου ώστε να στρέφεται με τη ιδιοσυχνότητα Larmor των πρωτονίων, στο εξωτερικό πεδίο. Η συχνότητα Larmor είναι ανάλογη της έντασης του πεδίου. Ο χρόνος εφαρμογής του B₁ είναι μικρός έως ότου στραφεί το διάνυσμα της μαγνήτισης κατά 90° ή 180°. Γνωστοί ως παλμοί 90° και 180°. Χρόνος μαγνητικής αποκατάστασης είναι μια δυναμική φυσική διαδικασία κατά την οποία το σύστημα των spins επανέρχεται σε κατάσταση ισορροπίας. Η μέτρηση της μαγνητικής αποκατάστασης γίνεται από τους χρόνους T1 και T2. Ο χρόνος T1 εκφράζει την ταχύτητα ανάκτησης της μέγιστης διαμήκουσ μαγνήτισης. Πρακτικά αναφέρεται στην ανάκτηση του 63% της τιμής του παλμού των 90°. Ενώ ο χρόνος T2 εκφράζει την ταχύτητα με την οποία η εγκάρσια μαγνήτιση μηδενίζεται. Ο χρόνος T2 αναφέρεται στην απώλεια του 37% της αρχικής μαγνήτισης. Ο πραγματικός χρόνος αποκατάστασης T2* χρησιμοποιείται πιο αραιά. Οι χρόνοι T1 και T2 ακολουθούν εκθετική αύξηση και μείωση αντίστοιχα και συμβαίνουν ταυτόχρονα ως συμπληρωματικά φαινόμενα. Ο χρόνος T1 είναι συνήθως 10 φορές μεγαλύτερος από τον T2. Τέλος, οι χρόνοι χαλάρωσης είναι χαρακτηριστικοί για τα spin κάθε ιστού (βλέπε Εικόνα 5). [38], [39], [40], [80]



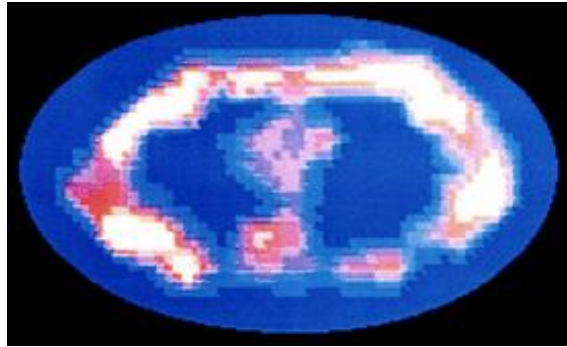
Εικόνα 5. Χρόνος T1 σε διαφορετικούς ιστούς

3.2 Μαγνητική Τομογραφία

3.2.1 Ιστορική Αναδρομή

Όλα ξεκίνησαν από δύο φυσικούς τον Edward Purcell και τον Felix Bloch που πειραματίστηκαν ανεξάρτητα στον ατομικό-μοριακό μαγνητικό συντονισμό στερεών και υγρών. Η συνεισφορά τους για την χρήση του μορίου του νερού λόγω των μοναδικών ιδιοτήτων του, οδήγησε στο Νόμπελ φυσικής το 1952. Έπειτα, ένας ακόμη φυσικός ο Raymond Damadian έρχεται το 1969 να επιβεβαιώσει την υποψία του ότι η διαφοροποίηση των καρκινικών κυττάρων από τα υγιή είναι εφικτή λόγω των μεγαλύτερων χρόνων χαλάρωσης. Ο πειραματισμός του λοιπόν με τους χρόνους χαλάρωσης σε ατομικό επίπεδο οδήγησε στην πρόταση του για ένα σαρωτή σώματος. Η πρόταση ευρεσιτεχνίας εγκρίνεται το 1974 και το πρώτο σύστημα μαγνητικής τομογραφίας ονομάστηκε Indomitable (Ακαταμάχητος). Το έτος 1977 η πρώτη απεικόνιση μαγνητικού συντονισμού είναι γεγονός και αποτελεί μια τομή του στήθους του βοηθού του Damadian, Larry Minkoff (βλέπε Εικόνα 6). Παράλληλα ο φυσικός Sir Peter Mansfield δούλεψε πάνω στην προσθήκη κλίσης στο μαγνητικό πεδίο ώστε να υπάρξει η δυνατότητα εξέτασης μιας δομής σε επίπεδο ατόμου. Αυτή η καινοτομία σε συνδυασμό με την μέθοδο του χημικού Lauterbur της ανασυγκρότησης προβολής, εισάγει την δυνατότητα για τρισδιάστατη απεικόνιση. Το 1974 πετυχαίνει να απεικονίσει το δάκτυλο ενός μαθητή σε 15-23 λεπτά, ενώ αναπτύσσει την τεχνική επίπεδης ηχούς (Echo Planar Imaging-EPI). Η υιοθέτηση του μαγνητικού τομογράφου ως απεικονιστική τεχνική είναι ζήτημα χρόνου καθώς τα πρώτα συστήματα εγκαθίστανται την Αμερική το 1980.

[39]



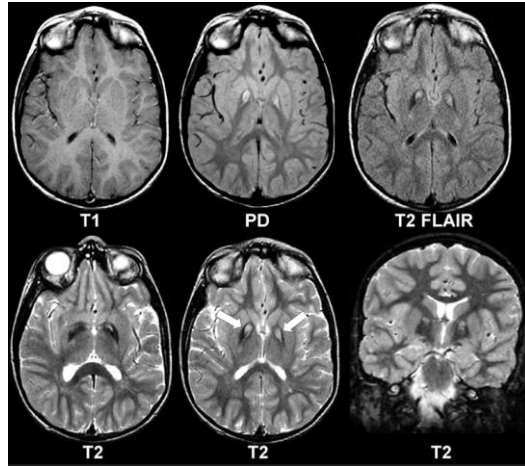
Εικόνα 6. Η πρώτη λήψη από μαγνητικό τομογράφο [26]

3.3 Πρωτόκολλα

Η μαγνητική τομογραφία αποτελεί μια από τις πιο δημοφιλείς απεικονιστικές τεχνικές που χρησιμοποιείται στην ακτινολογία για την λήψη εικόνων σχετικά με την ανατομία του σώματος. Χάρης στο ισχυρό μαγνητικό πεδίο και τα βαθμωτά πηνία παράγονται εικόνες υψηλής αντίθεσης των εσωτερικών δομών (οργάνων). Οι μαγνητικοί τομογράφοι έχουν χαρακτηριστική τιμή μαγνητικού πεδίου λόγω της σταθερής τιμής του υπεραγωγίμου μαγνήτη. Συνήθως απαντώνται μαγνήτες του 1.5T όπου η ποιότητα της εικόνας είναι υψηλή. Σπανιότερα λαμβάνονται απεικονίσεις από μαγνήτες των 3T όπου η ποιότητα της εικόνας είναι ανώτερη. Οι απεικονίσεις του συστήματος λαμβάνονται από διαφορετικές ακολουθίες και ανάλογα την ανατομική δομή ή τον τύπο της εξέτασης λαμβάνονται υπόψιν κάποιες παράμετροι. Οι δύο τύποι απεικόνισης είναι οι T1 και T2 εικόνες. Οι εικόνες απεικονίζουν τους ιστούς με χαρακτηριστικές αποχρώσεις που μας διευκολύνουν στην ανίχνευση των δομών (βλέπε Πίνακα 3.1). Αυτή η ιδιότητα της μαγνητικής που μας επιτρέπει να δούμε πολλές διαφορετικές δυσμορφίες οφείλεται στα διαφορετικά πρωτόκολλα που θα αναλύσουμε παρακάτω. [81], [82]



Εικόνα 7. T1-W εικόνα



Εικόνα 8. Μαγνητική Τομογραφία Εγκεφάλου με δείκτες στην παθολογία. Η T1 φαίνεται φυσιολογική. Παρόλα αυτά στην T2 είναι εμφανής η υψηλή ένταση σήματος τόσο στην εγκάρσια όσο και στη αξονική απεικόνιση. [22]

Πίνακας 3.1 Σύγκριση αποχρώσεων στην απεικόνιση T1 και T2

Τύπος Ιστός	T1 Απεικόνιση	T2 Απεικόνιση
Νερό	Μαύρο	Πολύ Λευκό
Λίπος	Πολύ Λευκό	Μαύρο
Κόκκαλο	Μαύρο	Μαύρο
Μυς	Ενδιάμεσο	Μαύρο
Όγκος	Ενδιάμεσο	Λευκό

Κάθε ακολουθία είναι ένας συνδυασμός παλμών των ραδιοσυχνοτήτων και των βαθμίδων. Παρόλα αυτά κάθε ακολουθία παλμών στοχεύει στην λήψη σήματος από τον παρακείμενο ιστό (αντίθεση), στο μικρότερο δυνατό χρόνο ενώ γίνεται προσπάθεια περιορισμού των ψευδενδείξεων και διατήρηση του σηματοθορυβικός λόγος (SNR). Άρα σε κάθε παλμό γίνεται ένας συμβιβασμός των παραμέτρων ώστε να εξασφαλιστεί το καλύτερο δυνατό αποτέλεσμα.

Η αρχιτεκτονική της ακολουθίας αποτελείται από παλμούς διέγερσης RF που απαιτούνται για το φαινόμενο του μαγνητικού συντονισμού, τις βαθμώσεις για την χωρική κωδικοποίηση (2D ή 3D) που θα διαμορφώσουν το k-space και ένα signal reading. Αυτά αποτελούν τα

ουσιαστικά τμήματα της απεικόνισης, υπάρχουν και επιλογές ώστε να τροποποιούνται οι αντιθέσεις, να επιταχύνονται οι ακολουθίες ή να μειώνονται οι ψευδενδείξεις.

Τέλος, λαμβάνονται υπόψιν οι παράμετροι των ακολουθιών, χρόνοι αποκατάστασης, TR, TE, Flip Angle, turbo factor, field of view (FOV) και acquired matrix ώστε να βρεθεί ο καλύτερος συμβιβασμός ταχύτητας, αντίθεσης και χωρική διακριτικής ικανότητας (βλέπε Πίνακα 3.2).^{[39], [40]}

Πίνακας 3.2 Ρυθμίσεις παραμέτρων για την λήψη της μαγνητικής τομογραφίας (Acquisition)

Parameter	T1	T2
TE	23 ms	100 ms
TR	2100 ms	3000 ms
TI	1000 ms	–
Acquired matrix	288 × 197	512 × 336
FOV	220 mm	220 mm
Voxel size	0.76 × 0.76 × 1.12 mm	0.43 × 0.43 × 0.65 mm
Scan time [min]	2.12	2.36

TE — time to Echo; TR — repetition time; TI — inversion time; FOV — field-of-view

3.3.1 Βασικές Ακολουθίες στο MRI

Οι δύο μεγάλες κατηγορίες ακολουθιών είναι οι Gradient Echo και οι Spin Echo. Από αυτές έχουν προκύψει κάποιες ακολουθίες για να αυξηθεί η ταχύτητα λήψης των εικόνων.

Η Spin Echo (SE) ή ακολουθία ηχούς περιστροφής αποτελείται από δυο διαδοχικούς παλμούς 90° pulse – 180° rephasing pulse στο σημείο TE/2 και ένα reading signal στο TE.

Η ακολουθία επαναλαμβάνεται για TR (Χρόνος επανάληψης). Με κάθε επανάληψη συμπληρώνεται μία γραμμή στο k-space. Ο παλμός 180° χρησιμοποιείται για να εξασφαλιστεί η T2 ηχώ αντί της T2*. Η αντίθεση στη διάρκεια του spin echo καθορίζεται από τον χρόνο TR. Δηλαδή το χρόνο μεταξύ δύο διαδοχικών 90° παλμών. Όσο πιο μεγάλο TR τόσο μεγαλύτερη η αποκατάσταση της αξονικής μαγνήτισης (βλέπε Πίνακα 3.3).

Πίνακας 3.3 Παράδειγμα χρόνων αποκατάστασης για την spin echo ακολουθία

Spin Echo	TR	TE
T1	621	11
T2	3013	106
PD	3013	11

Η Gradient Echo (GE) ή διαβαθμισμένη ακολουθία ηχούς διαφέρει από την ακολουθία ηχούς περιστροφής στην παράμετρο της flip angle καθώς η γωνία είναι κάτω από τις 90° και επίσης δεν χρησιμοποιείται ο παλμός 180° . Η μικρότερη γωνία ανάκτησης επάγει μικρότερη ποσότητα μαγνήτισης να ανατρέπεται στο εγκάρσιο επίπεδο. Η συνέπεια είναι μια ταχύτερη ανάκτηση της διαμήκους μαγνήτισης που επιτρέπει μικρότερες τιμές TR και TE και άρα μικρότερο χρόνο σάρωσης.

Τέλος, η Inversion recovery (IR) είναι μία ακολουθία προπαρασκευής καθώς έπειτα ακολουθεί συνήθως μια ακολουθία απεικόνισης spin echo. Πρακτικά προστίθεται ένας παλμός 180° πριν την ακολουθία spin echo, ο οποίος αντιστρέφει τη διαμήκη μαγνήτιση M_z στην αντίστροφη κατεύθυνση. Λόγω της αντιστροφής η διαμήκης μαγνήτιση θα αυξηθεί για να επιστρέψει στην αρχική της τιμή. Όταν η μαγνήτιση M_z περνά από το μηδέν τότε εφαρμόζεται ο παλμός 90° .

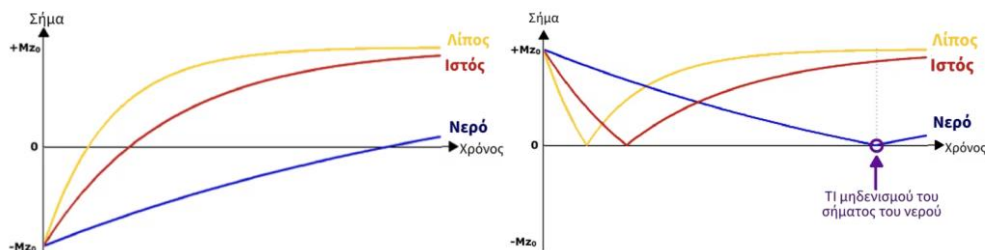
Η καθυστέρηση μεταξύ του παλμού αναστροφής 180° και του παλμού διέγερσης 90° ονομάζεται χρόνος αναστροφής (Time Inversion). Καθώς η διαμήκης αποκατάσταση χαρακτηρίζεται από τον χρόνο αποκατάστασης T1, αυτές οι αλληλουχίες σταθμίζονται σύμφωνα με τον χρόνο T1. Επομένως κάποια είδη ιστών θα έχουν αρνητικό σήμα, δηλαδή δεν θα λάβουμε σήμα για αυτούς. Με αυτό τον τρόπο καταστέλλουμε το ανεπιθύμητο σήμα. Η IR ακολουθία έχει εφαρμογή στην καταστολή υγρών (νερού ή λίπους), την διαδεδομένη FLAIR ακολουθία (βλέπε Πίνακα 3.4).

Άρα γνωρίζοντας τους χρόνους μαγνητικής αποκατάστασης ενός ιστού μπορούμε να προσαρμόσουμε το χρόνο T1. Το νερό έχει μεγάλο χρόνο T1, ο μηδενισμός του σήματος του νερού συμβαίνει για χρόνο ανάκτησης 2000-2500 χιλιοστών του δευτερολέπτου. (βλέπε Εικόνα 9)

Συνήθως στην εγκεφαλική μαγνητική τομογραφία, χρησιμοποιείται η ακολουθία της ηχούς ταχείας περιστροφής (Fast Spin Echo) ώστε να αντισταθμιστεί ο χρόνος αντιστροφής και ο χρόνος ανάκτησης που απαιτεί ο μεγάλος χρόνος TR. ^{[39], [44]}

Πίνακας 3.4 Παράδειγμα χρόνων αποκατάστασης για την inversion recovery ακολουθία

IR	TR	TE	TI
T2	3013	106	1006



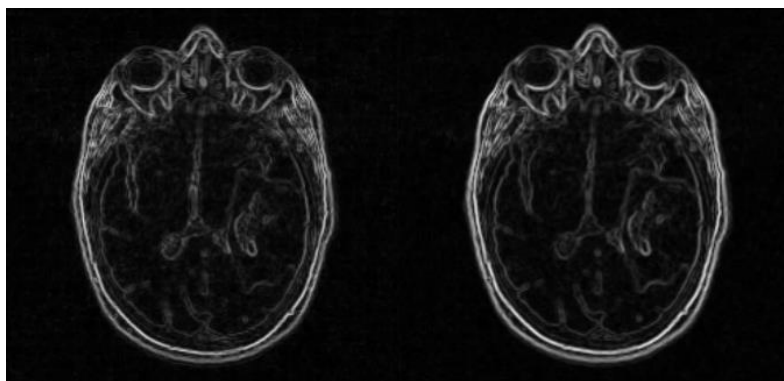
Εικόνα 9. Σήμα Ανάστροφης Ανάκτησης

4. Τμηματοποίηση Εικόνας

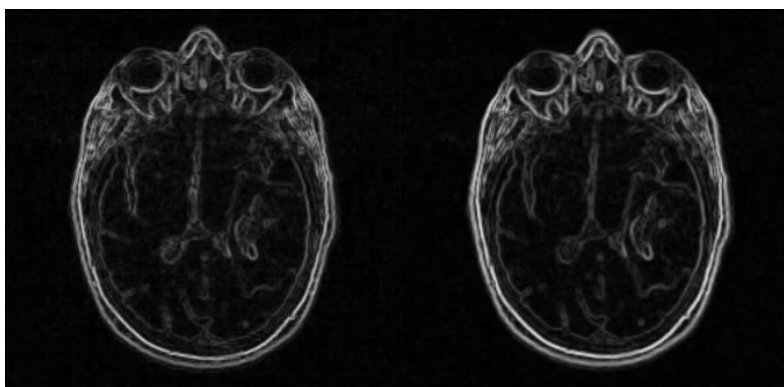
4.1 Μέθοδοι Επεξεργασίας Ιατρικής Εικόνας

Στην ψηφιακή επεξεργασία εικόνας και την ρομποτική όραση εφαρμόζουμε μεθόδους για να χωρίσουμε την εικόνα σε πολλά τμήματα. Ο στόχος της τμηματοποίησης είναι ο μετασχηματισμός της εικόνας σε κάτι πιο ουσιαστικό και πιο εύχρηστο. Στην ιατρική η τμηματοποίηση της εικόνας είναι συχνά αναγκαίο στάδιο καθώς κάποια περιοχή της εικόνας μας ενδιαφέρει να αναλυθεί περαιτέρω είτε κλινικά είτε να ληφθεί για την ανάπτυξη κάποιου συστήματος υποστήριξης απόφασης. Οι βασικές τεχνικές χωρίζονται σε αυτές που αξιοποιούν την χωρική πληροφορία και τις *blackbox*. Οι τεχνικές περιλαμβάνουν κατωφλίωση, ομαδοποίηση, ανίχνευση ακμών, μορφολογική επεξεργασία, ανάπτυξη περιοχών. Μερικές από τις τεχνικές εμφανίζουν πολύ καλά αποτελέσματα σε συγκεκριμένες εργασίες. Παρόλα αυτά έχουν μεγάλη ευαισθησία σε διακυμάνσεις τιμών και δεν έχουν κάποιο αντικειμενικό κριτήριο στην εφαρμογή τους παρά μόνο την τιμή της απόχρωσης του γκρι, δημιουργώντας εικόνες οι οποίες δεν μας παρέχουν χρήσιμη πληροφορία. Η προσέγγιση της τμηματοποίησης είναι μια πολύπλοκη διαδικασία και συχνά την συναντάμε στο τέλος του σταδίου της προεπεξεργασίας. Μια εικόνα ενδέχεται να έχει μετασχηματιστεί μέσω φίλτρων όπως το φίλτρο μέσης τιμής για την απομάκρυνση θορύβου ή κάποιας διαδικασίας για την τροποποίηση του ιστογράμματος της. Συνεπώς ανεξάρτητα από την προσέγγιση η μεθοδολογία που ακολουθεί κανείς στο στάδιο της προεπεξεργασίας έχει καθοριστικό ρόλο. (GIGO-Garbage In/Garbage Out).

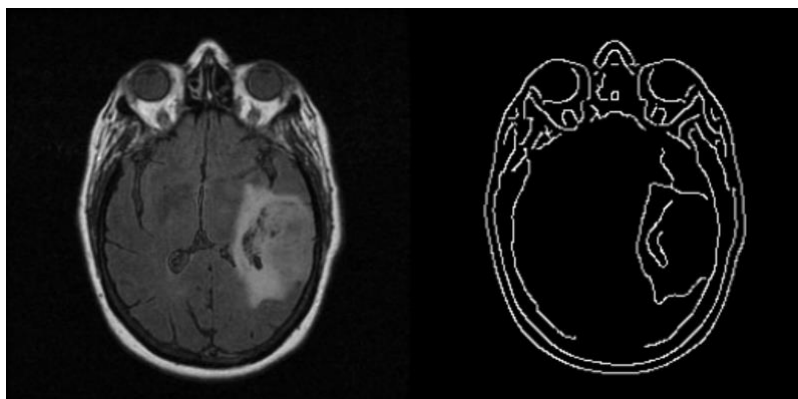
Παρακάτω απεικονίζονται κάποιες από τις πιο δημοφιλείς τεχνικές που εφαρμόσαμε στα δεδομένα μας. Συγκεκριμένα εφαρμόσαμε φίλτρα ανίχνευσης ακμών και την τεχνική της κατωφλίωσης Otsu (βλέπε Εικόνα 10-13).



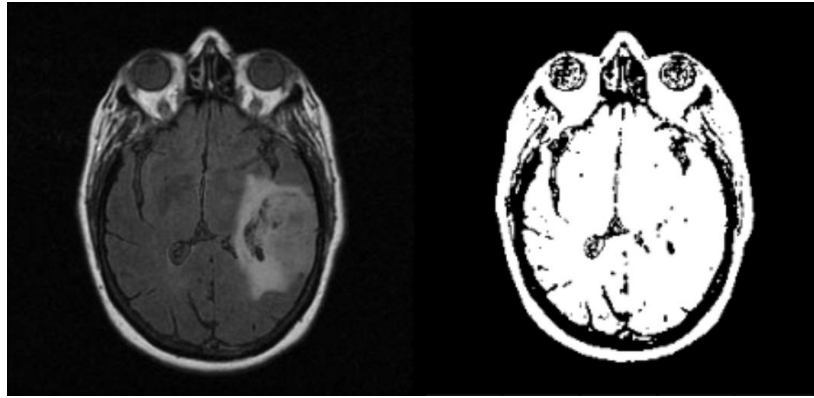
Εικόνα 10. Φίλτρα ανίχνευσης ακμών Roberts και Sobel



Εικόνα 11. Φίλτρα ανίχνευσης ακμών Scharr, Prewitt



Εικόνα 12. Αριστερά Κανονική, Δεξιά Φίλτρο ανίχνευσης ακμών Canny



Εικόνα 13. Μέθοδος Κατωφλίωσης Otsu ^[27]

Παρατηρούμε ότι οι τεχνικές ανίχνευσης ακμών δεν εμφανίζουν ιδιαίτερες διαφορές μεταξύ τους. Προσεγγίζουν τα όρια κάποιων δομών της εικόνας χωρίς κάποιο κριτήριο. Αυτό τις καθιστά αναποτελεσματικές για την τμηματοποίηση σε εικόνες μαγνητικής τομογραφίας.

Στην μέθοδο κατωφλίωσης Otsu, το αποτέλεσμα μας επιβεβαιώνει την ανάγκη για ανάπτυξη ειδικών τεχνικών που να μπορούν να αναγνωρίσουν την επιθυμητή δομή στην ιατρική εικόνα. Αυτό επιβεβαιώνει την αρχική μας πρόταση, καθώς οι παραπάνω μέθοδοι δεν έχουν κάποιο κριτήριο πέραν των αποχρώσεων του γκρι για να στοχεύσουν την επιθυμητή περιοχή.

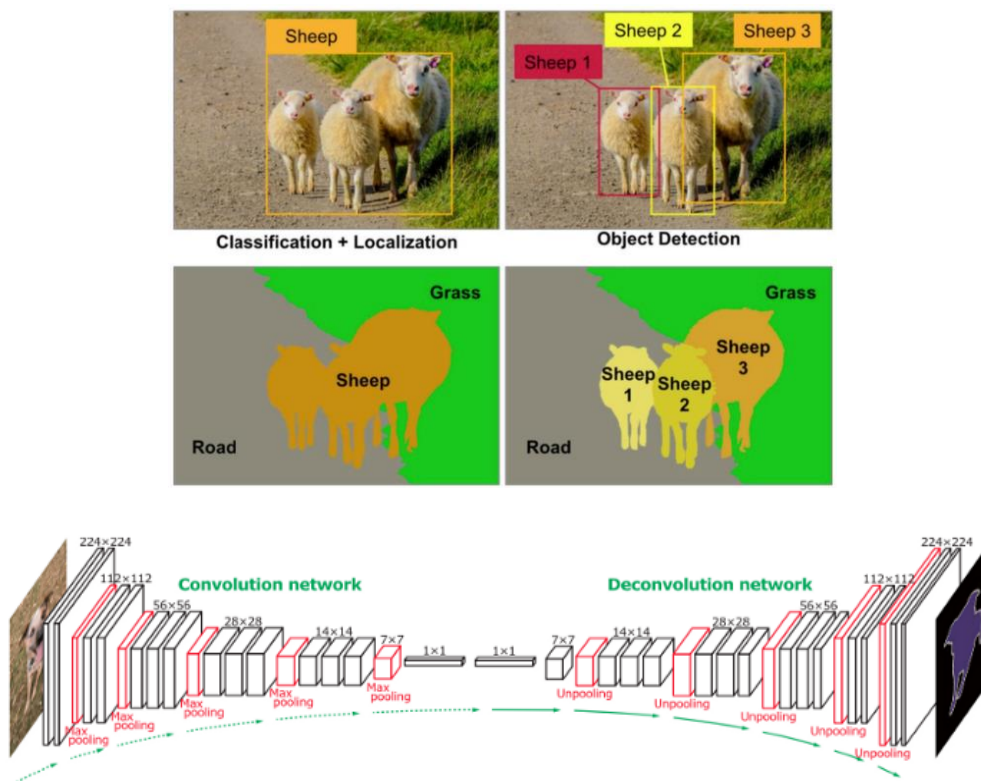
4.2 Ανίχνευση Αντικειμένων και Τμηματοποίηση

Τα νευρωνικά δίκτυα όπως θα εξετάσουμε στη συνέχεια έχουν ένα ευρύ φάσμα εφαρμογών στον τομέα της μηχανικής μάθησης. Ένα πεδίο στο οποίο τα εφαρμόζουμε ονομάζεται σημασιολογική τμηματοποίηση (semantic segmentation). Χρησιμοποιούμε εικόνες ως δεδομένα εισόδου και ως έξοδο κάποιες κατηγορίες στις οποίες θα ταξινομούνται τα εικονοστοιχεία. Έτσι λοιπόν χωρίς να διαφοροποιείται σημαντικά η λειτουργία τους, ταξινομούν τα εικονοστοιχεία σε μια από τις πιθανές κατηγορίες. Για παράδειγμα αυτό μπορεί κάλλιστα να αποτελεί μια διαφοροποίηση υγιούς ιστού από μια κακοήθεια (όγκο, ανωμαλία κ.ά.). Άρα μπορούμε εύκολα να μιλήσουμε για μια μεγάλης κλίμακας ταξινόμηση κυλιόμενου παραθύρου όπου πολλά εικονοστοιχεία θα ανήκουν σε μια κατηγορία σε κάθε εικόνα. Με άλλα λόγια μπορούμε να αναπτύξουμε δίκτυα ικανά να ταξινομούν τα pixels σε πολλαπλές κατηγορίες (multi-class segmentation). Παρόλο που ακούγεται εύκολη και οικεία σαν διαδικασία, είναι αρκετά απαιτητική σε υπολογιστική ισχύ. Αυτό που καλούμαστε να κάνουμε

είναι η μείωση του αριθμού των περασμάτων που τροφοδοτούμε το CNN μας με κάθε εικόνα, με ιδανικό στόχο τον περιορισμό σε μία φορά. Ευτυχώς βρίσκονται διαθέσιμα τέτοια δίκτυα με δυνατότητα πρόβλεψης για όλα τα εικονοστοιχεία ταυτόχρονα με ένα μόνο πέρασμα.

Ένα δίκτυο ικανό να επιτελέσει την εργασία της τμηματοποίησης τυπικά αποτελείται από πολλαπλά επίπεδα κατερχόμενης δειγματοληψίας, ένα bottleneck και τέλος πολλαπλά επίπεδα ανερχόμενης δειγματοληψίας. Θα μπορούσαμε χάριν πληρότητας να χωρίσουμε το νευρωνικό δίκτυο σε image encoder και image decoder. Αυτή η αρχιτεκτονική encoder-decoder έχει σημαντικό πλεονέκτημα, την διατήρηση της χωρικής πληροφορίας. Η κατερχόμενη δειγματοληψία πραγματοποιείται μέσω τεχνικών striding και pooling ενώ η αντίστροφη διαδικασία αξιοποιεί τεχνικές όπως αυτή του πλησιέστερου γείτονα και τη μεταφερόμενη συνέλιξη (transposed convolution).

Έπειτα εφαρμόζοντας την συνάρτηση απώλειας την διασταυρούμενης εντροπίας σε επίπεδο pixel εκπαιδεύουμε το νευρωνικό μας δίκτυο μέσω της αντίστροφης διάδοσης (backprop). Αναλυτικά θα αναφερθούμε σε όλες τις μαθηματικές διαδικασίες στο επόμενο κεφάλαιο. [5], [7]

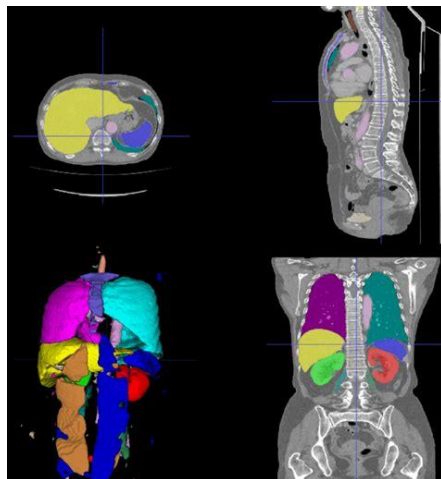


Εικόνα 14. Fully convolutional layer for semantic segmentation [5]

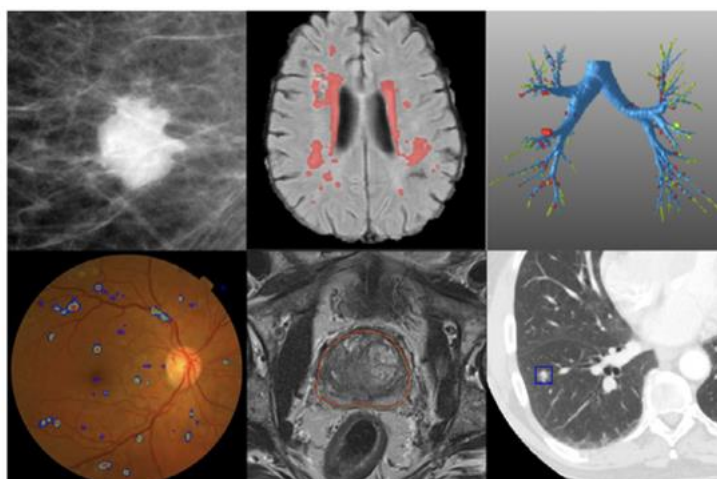
Με ανάλογο τρόπο μπορούμε να χρησιμοποιήσουμε τα νευρωνικά δίκτυα και να τμηματοποιήσουμε πολλαπλά αντικείμενα όπως όργανα ή ιστούς χρησιμοποιώντας εικόνες ιατρικού περιεχομένου. Ιδιαίτερο ενδιαφέρον έχουν οι εφαρμογές τμηματοποίησης εικόνας

αξονική και μαγνητικής τομογραφίας όπως και η παρούσα εργασία, όπου η ανάλυση της εικόνας μας επιτρέπει να προσεγγίσουμε τους διάφορους ιστούς με μεγάλη ακρίβεια.

Η ανάγκη για ακριβή τμηματοποίηση είναι μεγάλη καθώς οι παραγόμενες μάσκες είναι ιδιαίτερα χρήσιμες για πλάνα θεραπείας στην ακτινοθεραπεία ή για περαιτέρω επεξεργασία. Παρακάτω παρατίθεται ένα παράδειγμα πολλαπλής τμηματοποίησης οργάνων με εικόνες από CT και μια συλλογή από διαφορετικά modalities για να κατανοηθεί η σπουδαιότητα της τμηματοποίησης σαν μέθοδο ανάλυσης της ιατρικής εικόνας (βλέπε Εικόνα 15-16).



Εικόνα 15. Τμηματοποίηση πολλαπλών οργάνων (Multi-Organ Segmentation) [6]

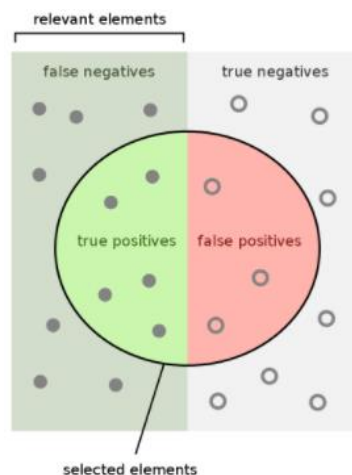


Εικόνα 16. Παραδείγματα εφαρμογής τμηματοποίησης στην ιατρική απεικόνιση με τη χρήση deep learning σε γνωστούς διαγωνισμούς με μεγάλα ποσοστά επιτυχίας. Με την σειρά από πάνω προς τα κάτω και από αριστερά προς τα δεξιά ταξινομήση μάζας μαστογραφίας [57], τμηματοποίηση βλαβών εγκεφάλου (BRATS) [59], ανίχνευση διαρροής στην κατάτμηση δέντρων αεραγωγών [58], διαβητική αμφιβληστροειδοπάθεια (Kaggle challenge 2015) [60], τμηματοποίηση αδένος προστάτη (PROMISE12) [61] και τμηματοποίηση όζων (LUNA16challenge)[62].

4.3 Μέτρα Αξιολόγησης Τμηματοποίησης

Μιλώντας για μέτρα αξιολόγησης, παρακάτω θα αναπτύξουμε μερικούς σημαντικούς δείκτες που χρησιμοποιήσουμε για την αξιολόγηση των μοντέλων τεχνητής νοημοσύνης. Αυτοί είναι η ευαισθησία (sensitivity ή recall) και η ειδικότητα (specificity) όπως χρησιμοποιούν οι Karabacak et al. [68]. Αυτοί οι όροι είναι επίσης γνωστοί ως true positive rate (TPR) και true negative rate (TNR). Η ευαισθησία είναι η πιθανότητα το μοντέλο να ταξινομήσει έναν ασθενή ως πάσχοντα από τη νόσο δεδομένου ότι είναι ασθενής. Η ειδικότητα είναι η πιθανότητα το μοντέλο να ταξινομήσει έναν ασθενή ως φυσιολογικό δεδομένου ότι είναι δεν πάσχει. Το precision είναι το κλάσμα των αληθώς θετικών (TP) ασθενών ως προς το σύνολο αυτών που ταξινομήθηκαν ως ασθενείς.

Έπειτα ακόμα ένα μετρικό είναι η ακρίβεια εικονοστοιχείου (pixel accuracy) αφού η τμηματοποίηση είναι μπορεί να είναι βασισμένη στα εικονοστοιχεία (pixel-wise). Προκύπτει από το άθροισμα των TP και TN pixels δια τον συνολικό αριθμό των pixels. Παρόλα αυτά η ακρίβεια δεν αποτελεί το πιο αξιόπιστο μετρικό καθώς το πλεόνασμα των pixels αποτελεί background ενώ η επιθυμητή περιοχή (δυσπλασία) αποτελεί ένα μικρό τμήμα της εικόνας. Αυτό είναι γνωστό ως το πρόβλημα class unbalanced. Με άλλα λόγια μια κλάση κυριαρχεί στην εικόνα, αυτό είναι κοινό πρόβλημα στα datasets, επομένως χρειαζόμαστε κάποια εναλλακτικά metrics ώστε να διαχειριστούμε το παραπάνω πρόβλημα. [8],[23]



Εικόνα 17. Οπτικοποίηση του πίνακα αλήθειας^[49]

$$\text{Sensitivity} = \frac{TP}{TP + FN}$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{PA} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \text{ and } \in [0,1]$$

4.4 Δείκτες αξιολόγησης Jaccard και Dice

Το μοντέλο μας εξετάζει πολυδιάστατες εικόνες οι οποίες έχουν έως και 90% πληροφορία υποβάθρου. Αυτό οδηγεί στο πρόβλημα ταξινόμησης των unbalanced dataset όπου η πληροφορία για τις κατηγορίες ενδιαφέροντος είναι σημαντικά μικρή. Για παράδειγμα μια εικόνα με χιλιάδες pixels έχει μερικές εκατοντάδες ή δεκάδες pixels που σχετίζονται με την πληροφορία των labels. Η ακρίβεια των pixels ως μετρικό που χρησιμοποιούμε για την αξιολόγηση της απόδοσης του ταξινομητή δεν είναι έγκυρη. Ας σκεφτούμε την πρώτη πρόταση, κατανοούμε ότι το μοντέλο μας έχει πολλή πληροφορία για το background. Άρα του είναι αρκετά εύκολο να το προβλέψει σε σχέση με το label. Αυτό αυξάνει την ακρίβεια χωρίς όμως να συνεισφέρει ουσιαστικά στις προβλέψεις μας. Είναι ένα τυπικό παράδειγμα bias. Έτσι προσεγγίζουμε άλλους δείκτες ώστε να αξιολογήσουμε το περιεχόμενο των κλάσεων. Ο πρώτος λέγεται IoU (τομή έναντι της ένωσης) ή αλλιώς Jaccard. Ουσιαστικά εξετάζει την τομή της πραγματικής κλάσης και της προβλέψιμης από το μοντέλο και διαιρείται από την ένωση αυτών. Πολλές φορές η προβλέψιμη απέχει από την πραγματική και αυτό είναι εμφανές τόσο από την απεικόνιση όσο και από το αποτέλεσμα του μετρικού αυτού.

Όσο πιο πολύ αυξήσουμε τον αριθμητή, τόσο πιο κοντά θα φέρουμε την προβλέψιμη στην πραγματική κλάση. Ενώ όσο πιο πολύ μειώσουμε τον παρανομαστή, τόσο πιο λίγα pixels θα έχουν προβλεπτεί λάθος (FP, FN).

Άρα λαμβάνοντας υπόψιν τα παραπάνω μπορούμε να κατανοήσουμε πόσο κοντά στην αλήθεια είναι η πρόβλεψη μας.

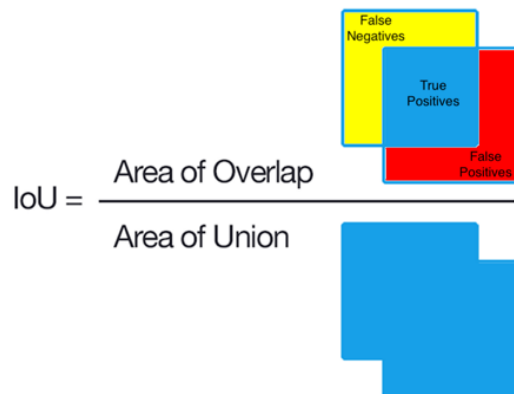
Ένα ακόμα μετρικό το οποίο ορίζεται παρόμοια με το jaccard καλείται συντελεστής Dice.^[63] Το μετρικό αυτό αποτελεί ένα ακόμα πιο αξιόπιστο δείκτη καθώς είναι ο αρμονικός μέσος. Η αξιοπιστία του Dice έγκειται στην μέση τιμή μεταξύ recall και precision όπου πάντα επιστρέφει την χαμηλότερη τιμή.

Γι' αυτό κάνουμε λόγο για αρμονικό μέσο, καθώς υπάρχει μια μορφή bias στην μικρότερη τιμή. Αυτό όμως τον κάνει λιγότερο διερμηνευόμενο στην πράξη, και πολλές φορές αυξάνεται

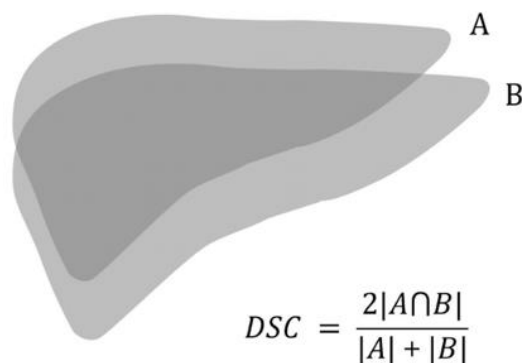
Τμηματοποίηση όγκων εγκεφάλου σε εικόνες μαγνητικής τομογραφίας μέσω αλγορίθμων βαθιάς μάθησης

δύσκολα. Τέλος, προκειμένου να αυξηθεί ο Dice γίνεται ένας συμβιβασμός μεταξύ precision και recall και χρησιμοποιούνται ποινές (penalties) ώστε να μειώσουμε το bias.

Παρόλα αυτά, αποτελούν state-of-art δείκτες μαζί με το Jaccard και εφαρμόζονται συμπληρωματικά στα προβλήματα τμηματοποίησης εικόνας. [2], [3]



Εικόνα 18. Σχηματική αναπαράσταση του Jaccard [4]



Εικόνα 19. Συντελεστής Dice [45], [46]

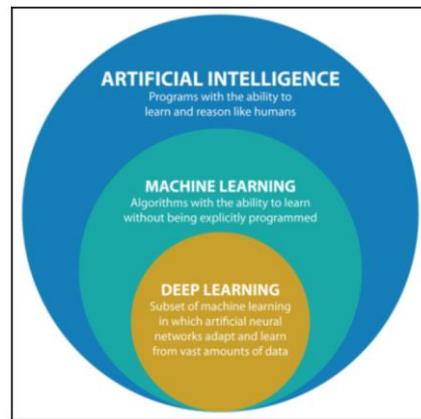
$$IoU = \frac{TP}{TP + FP + FN}$$

$$Dice = \frac{2 \times TP}{(TP + FP) + (TP + FN)}$$

5. Τεχνητή Νοημοσύνη

Μηχανική Μάθηση: Τομέας της τεχνητής νοημοσύνης που εκπαιδεύει μηχανές να επιτελούν συγκεκριμένες διαδικασίες με τρόπο που θυμίζει την ανθρώπινη ευφυΐα.

Βαθιά Μάθηση: Υποκατηγορία της μηχανικής μάθησης η οποία στηρίζεται στην μάθηση μέσα από μεγάλο όγκο δεδομένων για την κατανόηση κάποιας συγκεκριμένης εργασίας.



Εικόνα 20.Υποκατηγορίες Artificial Intelligence ^[78]

5.1 Κατηγορίες Μηχανικής Μάθησης

Εποπτευόμενη Μάθηση: Μοντέλα που παρουσιάζονται με τα δεδομένα εισόδου και τα επιθυμητά αποτελέσματα. Το μοντέλο θα προσπαθήσει στη συνέχεια να μάθει κανόνες που αντιστοιχούν τα δεδομένα εισόδου στα επιθυμητά αποτελέσματα. Όταν χρησιμοποιούμε τον όρο εποπτευόμενη μάθηση, συνήθως αναφερόμαστε σε ταξινόμηση δεδομένων σε κατηγορίες. (π.χ. γάτα-σκύλος) ^[12]

Μη-εποπτευόμενη Μάθηση: Μοντέλα όπου τα δεδομένα τους παρουσιάζονται χωρίς κατηγορίες ή προκαθορισμένα πρότυπα, και αυτά τα μοντέλα θα προσπαθήσουν να συμπεράνουν τις υποκείμενες δομές από το σύνολο των δεδομένων. Χαρακτηριστικό παράδειγμα μη εποπτευόμενης μάθησης είναι η γενετική τεχνητή νοημοσύνη. (π.χ. GANs για μια άγνωστη διαδικασία που θέλουμε να βελτιστοποιήσουμε)

Ενισχυτική Μάθηση: Το μοντέλο ή ο πράκτορας θα αλληλεπιδράσει με ένα δυναμικό κόσμο ώστε να επιτύχει ένα συγκεκριμένο στόχο. Ο δυναμικός κόσμος θα επιβραβεύει ή θα τιμωρεί τον πράκτορα βάσει των ενεργειών του. Με την πάροδο του χρόνου, ο πράκτορας θα μαθαίνει να περιηγείται στον δυναμικό κόσμο και να επιτυγχάνει τους στόχους του βάσει των ανταμοιβών και των τιμωριών που έχει λάβει. (π.χ. AI Robot)

Ημι-εποπτευόμενη Μάθηση: Η κατηγορία αυτή είναι υβριδική καθώς αποτελείται από ένα δείγμα δεδομένων που φέρουν τις κατηγορίες στις οποίες ανήκουν αλλά το μεγαλύτερο μέρος των δεδομένων δεν φέρει τις επιθυμητές κατηγορίες. Είναι κατά κάποιο τρόπο η σύμφυση εποπτευόμενης και μη-εποπτευόμενης.

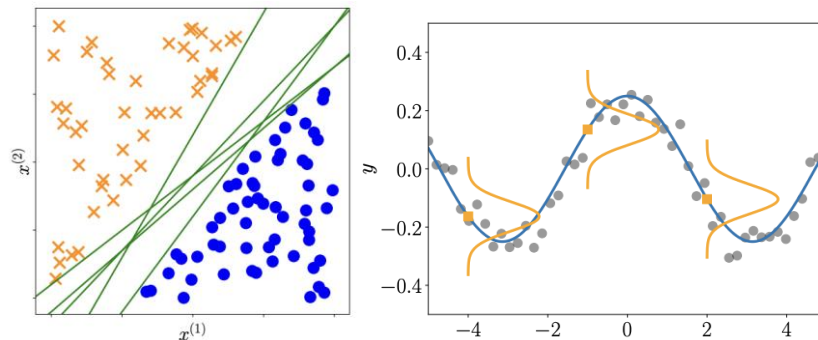
Γενικευμένη Τεχνητή Νοημοσύνη AGI

Ως μεγάλος στόχος του AI ανεξάρτητα από τις κατηγορίες του είναι μια μορφή γενικευμένης νοημοσύνης. Ανοίγοντας τον δρόμο για ένα κόσμο από πιθανές εφαρμογές βασισμένο στην δημιουργική συνεργασία ανθρώπου και υπολογιστή. Το AGI ή αλλιώς ισχυρή τεχνητή νοημοσύνη, έχει αναλυθεί σε υποπροβλήματα ανάλογα τα χαρακτηριστικά του συστήματος. Έτσι δίνεται έμφαση στην μάθηση των χαρακτηριστικών ώστε να μετασχηματιστεί ένα σύστημα σε ευφυές. Άρα η διαδικασία της μάθησης είναι καταλυτική για την αναγνώριση των ουσιαστικών χαρακτηριστικών καθώς χωρίς αυτά η γενίκευση είναι αδύνατη. Αυτό αποτελεί τη βασική πρόκληση στο δρόμο προς πραγματικά ευφυή συστήματα. [1], [2]

5.2 Εποπτευόμενη Μηχανική Μάθηση

Σε αυτή την κατηγορία της μηχανικής μάθησης, τα δεδομένα που διαθέτουμε έρχονται μαζί με μία περιγραφή, την οποία ονομάζουμε κατηγορία (label). Οι κατηγορίες που ανήκουν τα δεδομένα μας συνεισφέρουν στην διαδικασία μάθησης, εξ ου και όρος εποπτευόμενη μάθηση. Κάθε λοιπόν δεδομένο χρησιμοποιείται στην διαδικασία εκπαίδευσης μαζί με την κατηγορία του. Εξάγουμε μετρήσιμα χαρακτηριστικά που αντιπροσωπεύουν την εκάστοτε κατηγορία τα οποία χρησιμοποιούμε στους υπολογισμούς μας. Ουσιαστικά, το μοντέλο μαθαίνει πως να δημιουργεί μια αντιστοίχιση μεταξύ δεδομένων εισόδου και εξόδου βασισμένο στην πληροφορία των χαρακτηριστικών των κατηγοριών.

Άρα ο στόχος είναι η εύρεση μιας σχέσης για τα δεδομένα που διαθέτουμε. Τα πιο συχνά καθήκοντα ενός μοντέλου εποπτευόμενης μάθησης είναι η ταξινόμηση και η παλινδρόμηση. Η λογική πάνω στην οποία αναπτύσσουμε τους αλγορίθμους μας είναι η ίδια απλώς αλλάζει ο τύπος των δεδομένων. Εκτός από πίνακες με χαρακτηριστικά έχουμε πολλές φορές και εικόνες. Η ταξινόμηση είναι ο διαχωρισμός των δεδομένων με τρόπο που να καταστεί σαφές το όριο μεταξύ των κλάσεων. Τυπικό παράδειγμα ο διαχωρισμός εικόνων με διαφορετικά ζώα γάτα-σκύλος. Στην παλινδρόμηση επιτελείται προσπάθεια να βρούμε την καλύτερη δυνατή συνάρτηση που να μας πληροφορεί για τον τρόπο που μεταβάλλονται οι μετρήσεις μας (βλέπε Εικόνα 21). Οι συναρτήσεις αυτές είναι από απλές γραμμικές μέχρι πολύπλοκες πολυωνυμικές. Τυπικό παράδειγμα η πρόβλεψη για την τιμή των σπιτιών βάσει του αριθμού των δωματίων ή η πρόβλεψη της θερμοκρασίας εξετάζοντάς τα επίπεδα υγρασίας. [3], [8], [9], [11]



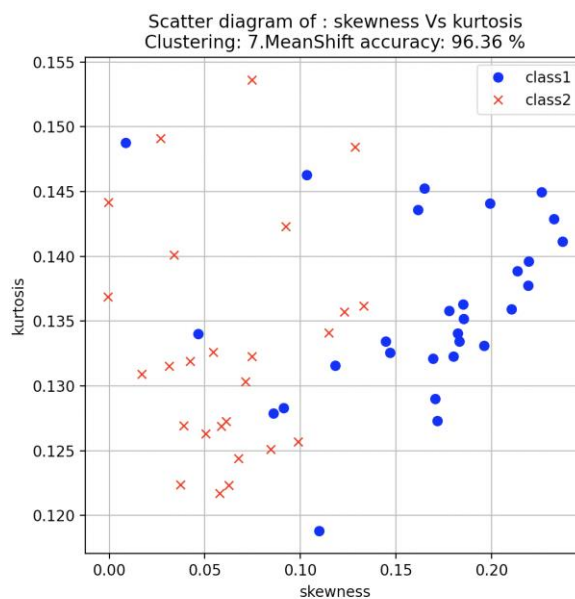
Εικόνα 21. Παραδείγματα Ταξινόμησης και Παλινδρόμησης [3]

5.3 Μη-Εποπτευόμενη Μηχανική Μάθηση

Στην μη εποπτευόμενη ή αυτο-οργανούμενη μάθηση, δεν υπάρχει επόπτευση. Αυτό πρακτικά σημαίνει ότι τα δεδομένα μας εισέρχονται για εκπαίδευση χωρίς κάποια περιγραφή για την κατηγορία τους. Άρα ένα πρώτο σχόλιο είναι ότι κάνουμε λόγο για την αντίθετη διαδικασία από την εποπτευόμενη μάθηση. Τα δεδομένα μας δεν έχουν κατηγορίες, συνεπώς η διαδικασία της εκπαίδευσης λειτουργεί διαφορετικά αφού παράμετροι όπως η ακρίβεια δεν έχουν πια λόγο χρήσης. Έτσι το μοντέλο βασισμένο στα δεδομένα προσπαθεί να μάθει μια κατανομή, από την οποία εξάγει χρήσιμες πληροφορίες ή χαρακτηριστικά. Θα μπορούσαμε να πούμε ότι ανάλογα τα δεδομένα που έχουμε στη διάθεση μας, δημιουργούμε ένα μοντέλο που λειτουργεί σαν γεννήτρια. Τυπικές υποκατηγορίες μη-εποπτευόμενης μάθησης είναι η συσταδοποίηση (clustering) και τα Generative models (GANs, VAEs). [19], [20]

5.3.1 Συσταδοποίηση

Ένας κλασικός αλγόριθμος με πολλές εφαρμογές είναι η συσταδοποίηση που στοχεύει στην ομαδοποίηση δεδομένων. Χρησιμοποιώντας δυο ή περισσότερα χαρακτηριστικά, μπορούμε οπτικά να κατατάξουμε σε ομάδες τα δεδομένα που εμφανίζουν κοντινές τιμές μεταξύ τους. Για παράδειγμα λαμβάνοντας τα χαρακτηριστικά κύρτωση και λοξότητα από δείγματα μη επιβιωσάντων και επιβιωσάντων καρκίνων του τραχήλου της μήτρας ομαδοποιούμαι με αρκετά μεγάλη ακρίβεια στα outliers (βλέπε Εικόνα 22). Το μοντέλο της συσταδοποίησης μαθαίνει από την κατανομή των δεδομένων (kernel density estimation). [13], [19], [20]



Εικόνα 22. Παράδειγμα συσταδοποίησης

5.4 Βαθιά Μάθηση και Συνελκτικά Νευρωνικά Δίκτυα

5.4.1 Εισαγωγή στη Βαθιά Μάθηση

Η βαθιά μάθηση αποτελεί μια υποκατηγορία της Μηχανικής Μάθησης όπου τα τεχνητά νευρωνικά δίκτυα μαθαίνουν και προσαρμόζονται μέσα από μεγάλο όγκο δεδομένων. Τα δεδομένα έχουν κάποια κοινά χαρακτηριστικά που εμφανίζονται σε όλο το dataset, αυτά στοχεύει το μοντέλο να εξάγει σε ένα χάρτη χαρακτηριστικών. Έτσι κάθε φορά που μία είσοδος έχει χάρτη χαρακτηριστικών σχετιζόμενο με αυτόν που έχει εκπαιδευτεί το μοντέλο, θα εμφανίζει αυξημένη πιθανότητα σωστής πρόβλεψης.

5.4.2 Μοντέλο νευρώνα

Το βασικό μαθηματικό μοντέλο των νευρωνικών δικτύων προτάθηκε από τους McCulloch and Pitts το 1943 και αποτελείται από τρία βασικά components:

- Ένα σετ από weights w_i που αντιστοιχούν στις τιμές εισόδου x_i (συνάψεις)
- Ένα αθροιστή για το άθροισμα των weights με τις τιμές των νευρώνων
- Μια συνάρτηση ενεργοποίησης που θα ορίσει την ενεργοποίηση των νευρώνων με βάση το άθροισμα των βαρών.^[79]

$$h = \sum_i w_i x_i \quad (1.1)$$

$$y = g(h) = \begin{cases} 1, & \text{if } h > \theta \\ 0, & \text{if } h \leq \theta \end{cases} \quad (1.2)$$

Εξίσωση. Συνάρτηση Ενεργοποίησης

Η εξίσωση δέχεται σαν είσοδο τα $\{x_i\}$, $\{w_i\}$ και εκφράζονται ως διανύσματα. Τα βάρη ρυθμίζουν την ακρίβεια της εξόδου. Η παράμετρος θ καθορίζει εάν ο νευρώνας θα πυροδοτηθεί ή όχι. Εάν ενεργοποιηθεί τότε $y=1$, αλλιώς $y=0$.

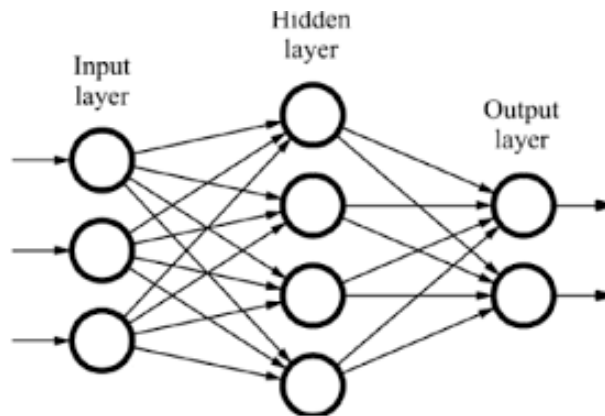
Αυτό αποτελεί την μαθηματική μοντελοποίηση του νευρώνα, όμως τα νευρωνικά δίκτυα έχουν χιλιάδες ώστε να συγκλίνουν στα επιθυμητά αποτελέσματα. Συνεπώς τα μαθηματικά γίνονται πολύπλοκα, και οι παράμετροι των νευρώνων αρκετά εκατομμύρια.

Η διαδικασία μάθησης είναι απλή σε αντίθεση με τις μαθηματική μοντελοποίηση που περικλείει κάθε βήμα της. Αναλυτικότερα τα βήμα συνοψίζονται ακολούθως:

- Τροφοδοτούμε το νευρωνικό δίκτυο μας με δεδομένα εισόδου και υπολογίζουμε ποσότητες που συνεισφέρουν στην εκπαίδευση.
- Υπολογίζουμε τη συνάρτηση κόστους ώστε να παρακολουθούμε πόσο κοντά είναι η πρόβλεψη μας στη πραγματική τιμή
- Χρησιμοποιώντας τους υπολογισμούς που έκανε το δίκτυο και τον μαθηματικό κανόνα της αλυσίδας υπολογίζουμε πόσο επηρεάζει κάθε παράμετρος των δεδομένων τη συνάρτηση κόστους
- Ενημερώνουμε τις τιμές των παραμέτρων (weights) ώστε να μειωθεί η απώλεια με την πάροδο των δεδομένων από το μοντέλο (backpropagation) ^{[3], [11]}

5.4.3 Μοντέλο Perceptron

Η απλούστερη δομή ενός νευρωνικού δικτύου είναι το δίκτυο του perceptron που αποτελείται από ένα κρυφό επίπεδο και πολλαπλούς νευρώνες (Εικόνα 23) . Το πρώτο επίπεδο είναι η είσοδος κάθε κόμβος λαμβάνει μια είσοδο και την μεταβιβάζει την έξοδο διαμέσου του κρυφού επιπέδου. Οι συνδέσεις επιτρέπονται μόνο σε νευρώνες διαδοχικών επιπέδων, και όχι σε συνδέσεις στο ίδιο επίπεδο. Η έξοδος παράγεται από τους νευρώνες του τελευταίου επιπέδου. Το κρυφό επίπεδο, δηλαδή το ενδιάμεσο στρώμα, δεν συνδέεται άμεσα με το επίπεδο εξόδου και οι νευρώνες του ενεργοποιούνται με τους κόμβους του προηγούμενου επιπέδου. Ένας μεμονωμένος νευρώνας δεν είναι χρήσιμος από αυστηρά μαθησιακή σκοπιά αλλά μπορεί να μας δώσει την βασική αρχή για την λειτουργία του μέσα στο νευρωνικό δίκτυο. Πρακτικά η μάθηση είναι η αλλαγή των weights, των biases και των παραμέτρων των συναρτήσεων ενεργοποίησης μεταξύ των συνδέσεων των νευρώνων του δικτύου.^{[3], [7]}



Εικόνα 23. Απλούστερο Νευρωνικό Δίκτυο Perceptron ^[79]

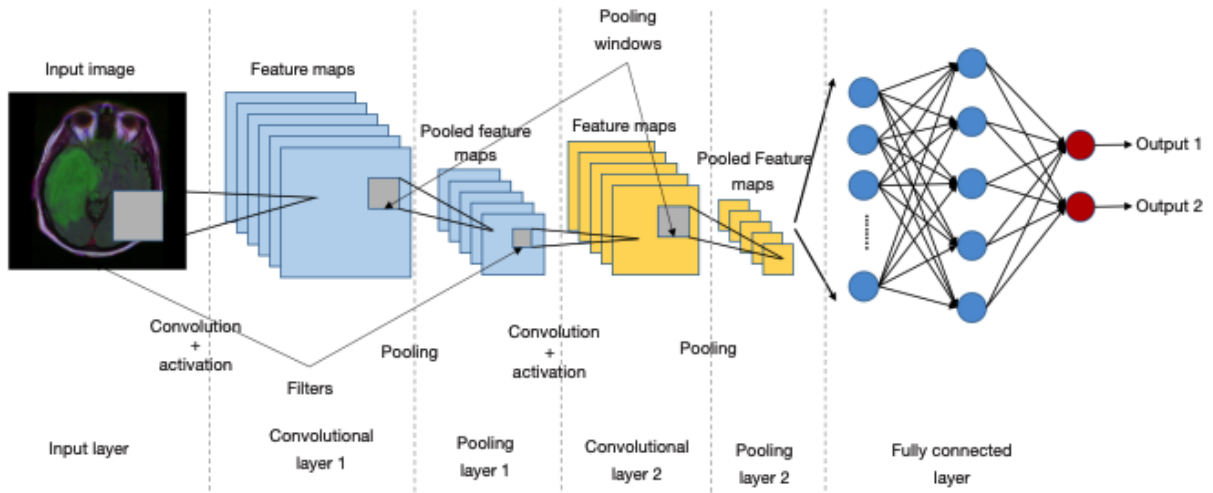
5.4.4 Βασική Αρχή Λειτουργίας Συνελκτικού Νευρωνικού Δικτύου

Ένα συνελκτικό νευρωνικό δίκτυο CNN, δέχεται στην είσοδο του δυσδιάστατους πίνακες δηλαδή εικόνες τις οποίες καλείται να επεξεργαστεί καταλλήλως ώστε να εξάγει κάποιο αποτέλεσμα. Έτσι λοιπόν η επεξεργασία που καλείται να εφαρμόσει έχει να κάνει με χωρική πληροφορία την οποία θα αναδειξεί. Ο τύπος της επεξεργασίας ονομάζεται χωρική συνέλιξη και είναι πρακτικά το άθροισμα των γινομένων των εικονοστοιχείων της εικόνας και ενός συνόλου βαρών ανάλογα με τον πυρήνα (φίλτρο) που χρησιμοποιούμε. Αυτή η διαδικασία επαναλαμβάνεται πολλές φορές με διαφορετικούς πυρήνες. Η συνέλιξη είναι η πρώτη πράξη που πραγματοποιείται από ένα δίκτυο CNN, οι τιμές της οποίας σχηματίζονται, μετακινώντας κατά μήκος της εικόνας το πεδίο υποδοχής, το οποίο είναι πρακτικά ένας πίνακας 2×2 ή 3×3 ώστε να μπορεί να επιτευχθεί η συνέλιξη. Ο τρόπος που το πεδίο υποδοχής μετατοπίζεται πάνω στην εικόνα καλείται βηματισμός (padding). Η τιμή του βηματισμού παίρνει συνήθως την τιμή 1, ωστόσο συναντάμε και διαφορετικές τιμές. Η αύξηση της τιμή του βηματισμού έχει να κάνει με την μείωση του όγκου των δεδομένων που επιθυμούμε να διατηρήσουμε. Έπειτα η τιμή της συνέλιξης αθροίζεται με μια σταθερά πόλωσης (bias) και το αποτέλεσμα περνάει μέσα από μία συνάρτηση ενεργοποίησης. Η τιμή αυτή θα αποθηκευτεί σε μια θέση ενός πίνακα και θα αποτελέσει είσοδο για το επόμενο επίπεδο. Όταν η διαδικασία επαναληφθεί για όλες τις θέσεις της εικόνας εισόδου, σχηματίζεται ένας δυσδιάστατος πίνακας ο οποίος ονομάζεται χάρτης χαρακτηριστικών. Αυτοί οι χάρτες μας δίνουν χρήσιμες πληροφορίες για τις εικόνες εισόδου όπως οι ακμές, διάφορα σημεία και περιοχές. Μην ξεχνάμε ότι η συνέλιξη αποτελεί τη βάση για το φιλτράρισμα των εικόνων με τα διάφορα φίλτρα όξυνσης, εξομάλυνσης κλπ. Στη συνέχεια σειρά έχει η συνάθροιση (pooling), μια μέθοδος κατά την οποία μειώνεται η ανάλυση των εικόνων συνεπώς και ο όγκος των προς επεξεργασία δεδομένων. Αυτού του είδους η μοντελοποίηση εξυπηρετεί πολύ λόγω των Big Data που καλούμαστε να διαχειριστούμε. Η διαδικασία αυτή θυμίζει πολύ την συνέλιξη καθώς πάλι χρησιμοποιούμε μικρούς πίνακες (2×2 ή 3×3) και αντικαθιστούμε τους πίνακες από μία μόνο τιμή, συνήθως την μεγαλύτερη (min, max, avg). Υπάρχουν πολλές επιλογές τις οποίες συναντάμε στα CNN όπως η μέγιστη συνάθροιση, η ελάχιστη συνάθροιση και η μέση συνάθροιση. (maxpooling, minpooling, avgpooling κ.ά) Έτσι μπορούμε να ονομάσουμε τους πίνακες που προκύπτουν συγκεντρωτικούς χάρτες χαρακτηριστικών. Αυτοί οι πίνακες αποτελούν τις εισόδους του δικτύου για το επόμενο επίπεδο.

Χρησιμοποιώντας τους συγκεντρωτικούς χάρτες ως εισόδους για το δεύτερο επίπεδο θα επαναλάβουμε την ίδια διαδικασία όπως περιγράφεται παραπάνω. Συνέλιξη, προσθήκη

βαρών, συνάρτηση ενεργοποίησης και συνάθροιση. Αυτό γίνεται για κάθε μια από τις εικόνες εισόδου με την μόνη διαφορά ότι αλλάζουν οι πυρήνες και οι τιμές πόλωσής (biases). Αρά από κάθε χάρτη θα προκύψουν πολλαπλοί διαφορετικοί χάρτες ανάλογα των αριθμό των πυρήνων που εφαρμόζουμε. Αν για παράδειγμα έχω τρεις εικόνες εισόδου στο δεύτερο επίπεδο και 5 πυρήνες τότε θα προκύψουν 15 διαφορετικοί χάρτες χαρακτηριστικών. Οι 5 νέοι χάρτες χαρακτηριστικών που έχουν παραχθεί για κάθε είσοδο μπορούν να συνδυαστούν μεταξύ τους και να προκύψει ένας τελικός. Αυτό είναι εφικτό καθώς η συνέλιξη είναι μια γραμμική διαδικασία η οποία στηρίζεται στην αρχή της επαλληλίας. Παρόλα αυτά, ως στόχος έχουμε την χρήση των χαρακτηριστικών για την ταξινόμηση. Άρα χρειαζόμαστε ένα ταξινομητή. Αφού λοιπόν έχουμε περάσει μέσα από τα κρυμμένα επίπεδα, τροφοδοτούμε την τιμή του τελευταίου συγκεντρωτικού χάρτη χαρακτηριστικών σε ένα πλήρως συνδεδεμένο νευρωνικό δίκτυο το οποίο θα τον μετατρέψει σε ένα διάνυσμα (flatten). Το διάνυσμα αυτό είναι μια μεγάλη κατακόρυφη στήλη η οποία περιέχει όλες τις τιμές του δυσδιάστατου πίνακα. Τέλος, σε κάθε εφαρμογή το πλήθος των νευρώνων εξόδου είναι ίσο με το πλήθος των κλάσεων των προτύπων που ταξινομούνται. Έτσι χρησιμοποιούμε το διάνυσμα στήλης ώστε να εξάγουμε μέσω μιας συνάρτησης την πιθανότητα ένα πρότυπο να ταξινομείται σε μια από τις κλάσεις εξόδου. Η συνάρτηση αυτή ονομάζεται λογιστική (logistic regression) και πρακτικά μετατρέπει το διάνυσμα σε μια πιθανότητα. Σε πολλά πακέτα μηχανικής μάθησης η λογιστική συνάρτηση καλείται ως Softmax (εξίσωση 1.3). Έτσι η έξοδος με την μεγαλύτερη τιμή πιθανότητας, καθορίζει την κλάση της εισόδου. Εν κατακλείδι, αυτή ήταν μια γενική περιγραφή της λειτουργίας ενός συνελκτικού νευρωνικού δικτύου χωρίς να αποτελεί πανάκεια για όλα τα νευρωνικά δίκτυα, παρακάτω παρατίθεται μία σχηματική αναπαράσταση όσων περιγράφηκαν παραπάνω (Εικόνα 24). [2], [8], [21]

$$\widehat{p}_k = \frac{e^{-x^T \theta^{(k)}}}{\sum_{j=1}^K e^{-x^T \theta^{(j)}}} \quad (1.3)$$

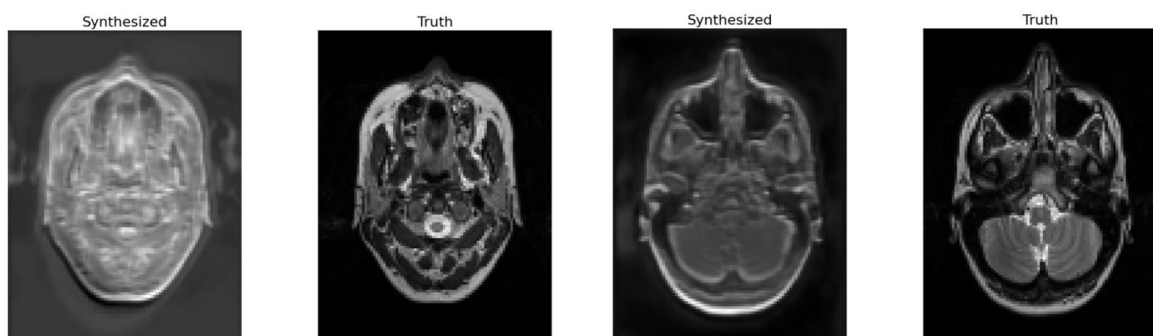


Εικόνα 24. Αναπαράσταση Νευρωνικού Δικτύου [36]

5.5 Fine-tuning στα Νευρωνικά δίκτυα

Η διαδικασία του fine-tuning είναι κρίσιμη για τα νευρωνικά δίκτυα. Η ρύθμιση των υπερπαραμέτρων αποτελεί μια πανίσχυρη διαδικασία καθώς μας οδηγεί σε εξαιρετικές αποδόσεις υπό την προϋπόθεση ότι επιλέχθηκαν οι σωστές τιμές για τις υπερπαραμέτρους που αναφέραμε παραπάνω.

Η εκπαίδευση είναι η κύρια διαδικασία που επιτελείται στα νευρωνικά δίκτυα. Παρακάτω ακολουθεί ένα παράδειγμα Image Translation όμοιο με αυτό των Platscher et al ώστε να αναδείξουμε την σημασία του fine-tuning των υπερπαραμέτρων για την εκπαίδευση (βλέπε Εικόνα 25). [18]



Εικόνα 25. Image Translation T1-w to T2-w

Μπορούμε να παρατηρήσουμε με γυμνό μάτι την διαφορά στην ανασύνθεση των εικόνων χρησιμοποιώντας διαφορετικό συνδυασμό υπερπαραμέτρων. Αυτό είναι ένα παράδειγμα ώστε

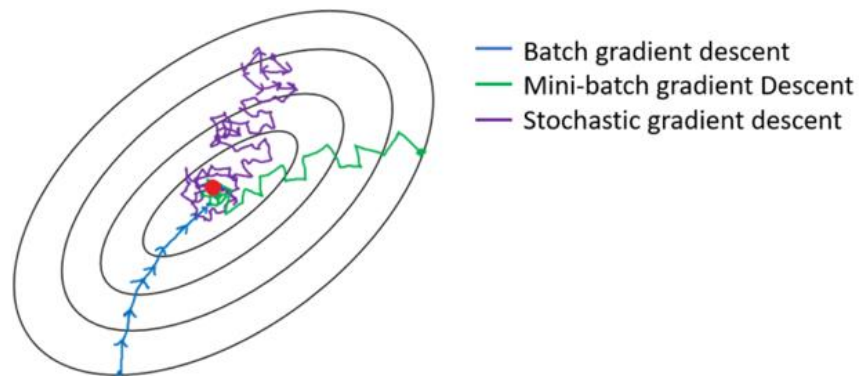
να αναδείξουμε την αναγκαιότητα της καλής ρύθμισης του μοντέλου (fine-tuning) μας κατά την εκπαίδευση των δεδομένων. Η παραπάνω εικόνες είναι αποτελέσματα από προηγούμενη εργασία χρησιμοποιώντας GANs και UNET. Η τιμή της υπερπαραμέτρου number of jobs διαφοροποιεί τα αποτελέσματα. ^{[16], [17], [19]}

5.5.1 Αλγόριθμος Βελτιστοποίησης Gradient Descent

Η τεχνική της βαθμωτής σύγκλισης (Gradient Descent) είναι state-of-the-art στη μηχανική μάθηση και ειδικότερα στα νευρωνικά δίκτυα. Η τεχνική αυτή έγκειται στον υπολογισμό των μερικών παραγώγων των πινάκων των βαρών στα κρυφά επίπεδα του δικτύου. Έτσι προσεγγίζεται η ακρίβεια του μοντέλου στο πέρας των επαναλήψεων που ανανεώνονται τα βάρη. Έως ότου η συνάρτηση απώλειας είναι κοντά ή ίση με μηδέν, το μοντέλο θα συνεχίσει να προσαρμόζει τις παραμέτρους του για να αποφεύγει όσο το δυνατόν μεγαλύτερο σφάλμα. Για παράδειγμα, όσο πιο κοντά είναι η προβλεπόμενη με την πραγματική τιμή τόσο πιο μικρό είναι το μέσο τετραγωνικό σφάλμα, άρα τόσο πιο μεγάλη η ακρίβεια του μοντέλου μας. Ανάλογα λειτουργεί η βαθμωτή σύγκλιση (gradient descent) όταν τα βάρη προσαρμοστούν στις βέλτιστες τιμές. Η κλίση (gradient) μας δίνει την κατεύθυνση που κινείται η συνάρτηση απώλειας.

Βήματα που ακολουθούνται στην βαθμωτή σύγκλιση:

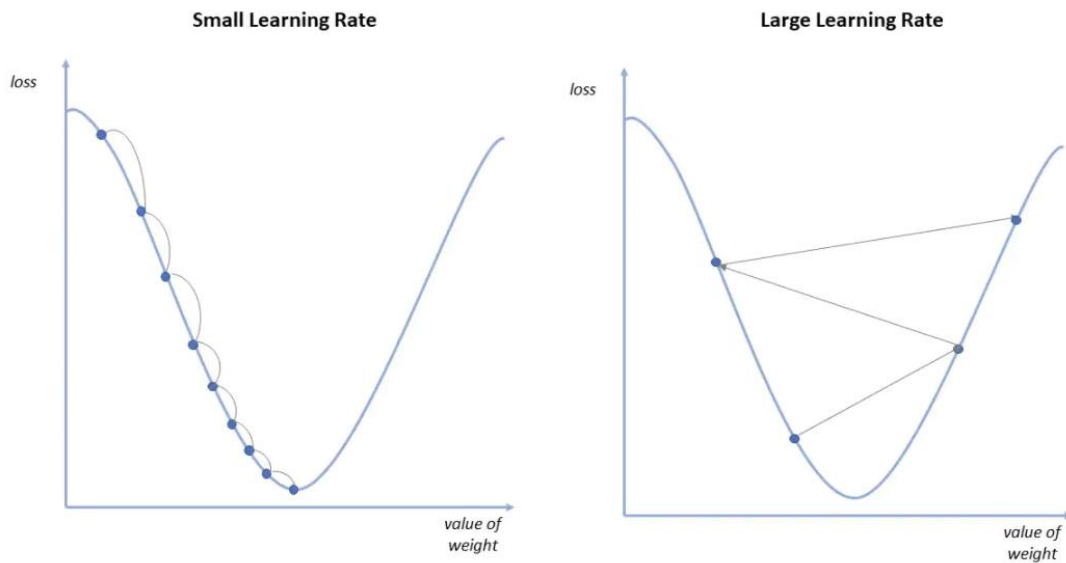
- Αρχικοποίηση: $\theta = 0$ (αρχικοποίηση βαρών)
- Υπολογισμός Συνάρτηση Απώλειας: $J(\theta) = \text{evaluate}(f(\text{βαρών}))$
- Παράγωγος της συνάρτησης : $\frac{\partial}{\partial \theta_j} J(\theta)$
- Ενημέρωση coefficients : $\theta_j = \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta)$ (1.4)
- Η διαδικασία επαναλαμβάνεται έως ότου βρεθεί local minimum.



Εικόνα 26. Αναπαράσταση βαθμωτής σύγκλισης^[37]

Το σημείο εκκίνησης (starting point) είναι αυθαίρετο για να αξιολογήσουμε την επίδραση των βαρών στην συνάρτηση απώλειας. Από εκείνο το σημείο ξεκινά η χρήση των παραγώγων (gradients) ώστε να υποδειχθεί η κατεύθυνση προς την οποία πρέπει να ενημερωθούν οι παράμετροι του δικτύου.^[3]

Η παράμετρος α στην (1.3) είναι ο ρυθμός εκμάθησης ή αλλιώς το μέγεθος των βημάτων που γίνονται για φτάσει το τοπικό ελάχιστο. Η τιμή είναι συνήθως μικρή ώστε να επιφέρει προοδευτικά αλλαγές στα βάρη. Τυπικά κυμαίνεται σε ένα εύρος $[10^{-6}, 10^{-2}]$, αυτό μεταφράζεται σε μικρά βήματα και αξιολογείται από την συμπεριφορά της συνάρτησης απώλειας. Η βέλτιστη τιμή αποτελεί πρόκληση που χρήζει διερεύνησης. Αυτό οφείλεται στις διαφοροποιήσεις των μοντέλων αφού συχνά υπάρχουν τοπικά ελάχιστα που μπορούν να παγιδέψουν την διαδικασία της βελτιστοποίησης. Επίσης, αυτές οι διαδικασίες έχουν απαιτήσεις σε χρόνο και υπολογιστική ισχύ. Ακολουθούν εικόνες για την καλύτερη κατανόηση της παραμέτρου.^[3]

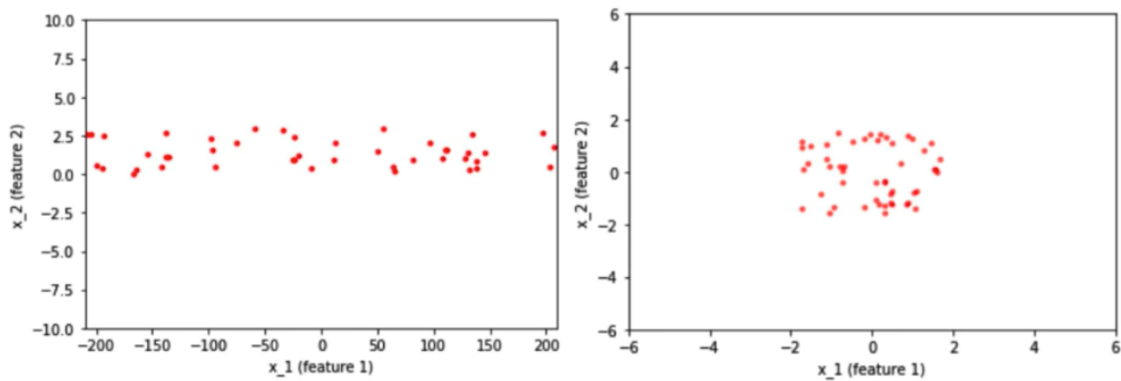


Εικόνα 27. Πειραματισμός στο ρυθμός εκμάθησης

Η κανονικοποίηση αποτελεί μια σημαντική τεχνική στην μηχανική και τη βαθιά μάθηση καθώς επιταχύνει την διαδικασία της εκπαίδευσης.

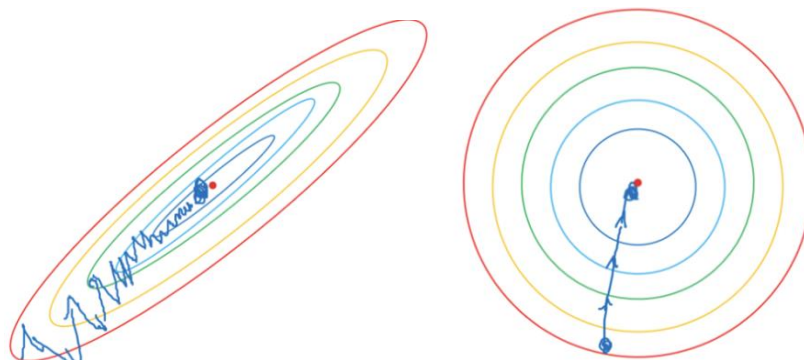
$$x_i = \frac{x_i - \mu}{\sigma} \quad (1.5)$$

Ένα απλό παράδειγμα απεικονίζεται παρακάτω, εφαρμόζοντας την εξίσωση της κανονικοποίησης (1.5) έχουμε μία πολύ διαφορετική εικόνα, η οποία επηρεάζει πολύ την διαδικασία της εκπαίδευσης που θα ακολουθήσει. Συγκεκριμένα, τα βάρη που θα είναι συνδεδεμένα με αυτά τα δεδομένα θα έχουν πολύ μεγάλο εύρος λόγω των τιμών $[-200, -200]$ σε σχέση με τις τιμές $[-2, 2]$. Άρα τα βάρη θα είχαν πολύ μεγάλες τιμές, έτσι η ενημέρωση των βαρών που θα επιφέρει η οπισθοδρόμηση θα αναγκαστεί να επιφέρει μεγάλες αλλαγές και ούτω καθεξής. Αυτό συνεπάγεται ότι ο αλγόριθμος βελτιστοποίησης θα εγκλωβιστεί στο τοπικό μέγιστο αντί να κατευθύνεται προς το ολικό ελάχιστο.



Εικόνα 28. Αριστερά με την κανονικοποίηση, Δεξιά χωρίς κανονικοποίηση

Αυτό το πρόβλημα λύνει η τεχνική batch normalization προσαρμόζοντας τα δεδομένα στην ίδια κλίμακα. Τα νευρωνικά δίκτυα θα τροφοδοτήσουν τα δεδομένα στους επιμέρους νευρώνες των επιπέδων και θα οδηγηθούν σε πολύπλοκες μαθηματικές πράξεις όσο πηγαίνουμε σε βαθύτερα επίπεδα, οι τιμές των νευρώνων θα διαφέρουν ανεξάρτητα από τις τιμές εισόδου. Οπότε μπορούμε να κανονικοποιούμε κάθε φορά τις τιμές στην είσοδο των νευρώνων ώστε να επιταχύνουμε την εκπαίδευση, αποφεύγοντας επιπλέον υπολογισμούς. Παρακάτω αναπαρίσταται σχηματικά η κατεύθυνση της συνάρτησης απώλειας του δικτύου εφαρμόζοντας την batch normalization (βλέπε Εικόνα 29). Η μείωση είναι σχεδόν εκθετική λόγω της μετατόπισης των τιμών των κρυφών επιπέδων για κάθε μερίδα δεδομένων που τροφοδοτούνται στο δίκτυο. Αυτό βοηθά στην ταχύτερη σύγκλιση ταυτόχρονα με την μείωση του χρόνου εκπαίδευσης.

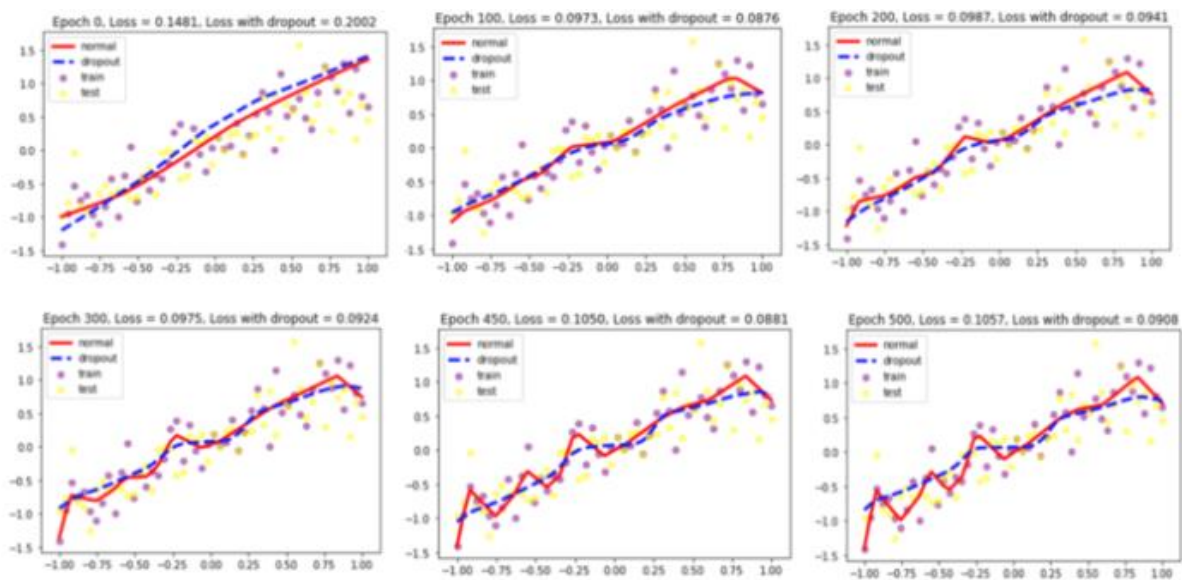


Εικόνα 29. Αναπαράσταση του Batch normalization στην εκπαίδευση ^[37]

Μία ακόμη τεχνική συστηματοποίησης που συνεισφέρει στη μείωση της απώλειας και του overfitting είναι η τεχνική dropout ή pruning. Η ονομασία “κλάδεμα” αποδίδει απόλυτα την λειτουργία της που δεν είναι άλλη από την απενεργοποίηση μέρους των νευρώνων του

δικτύου. Ουσιαστικά δοκιμάζει διαφορετικά δίκτυα απενεργοποιώντας κάθε φορά ένα διαφορετικό συνδυασμό νευρώνων ανά επίπεδο.

Deep neural networks με εκατομμύρια παραμέτρους μπορεί να οδηγήσουν το δίκτυο σε overfitting, αυτό θα μπορούσε να αποφευχθεί χρησιμοποιώντας πολλαπλά δίκτυα με διαφορετικές διαμορφώσεις. Έπειτα λαμβάνοντας μια μέση τιμή για τις προβλέψεις που θα έκαναν θα μπορούσαμε να αξιολογήσουμε το δίκτυο. Πρακτικά όμως αυτό δεν είναι εφικτό. Έτσι χρησιμοποιούμε την dropout, απενεργοποιώντας τυχαία νευρώνες σε κάθε βήμα κάθε επανάληψης της εκπαίδευσης. Ο τρόπος που επιλέγονται οι νευρώνες που δεν θα συνεισφέρουν συνδέεται με πιθανότητες. Διαφορετικές εκδοχές του δικτύου μας δοκιμάζονται ώστε η εκπαίδευση να μην έχει μια οπτική γωνία. Συνεπώς είναι μια τεχνική που μας παρέχει ένα εύρος δοκιμών, μας εξοικονομεί χρόνο και συνεισφέρει στο generalizability του μοντέλου (βλέπε Εικόνα 30).^[11]



Εικόνα 30. Παράδειγμα χρήσης του Dropout στο dataset του MNIST^[24]

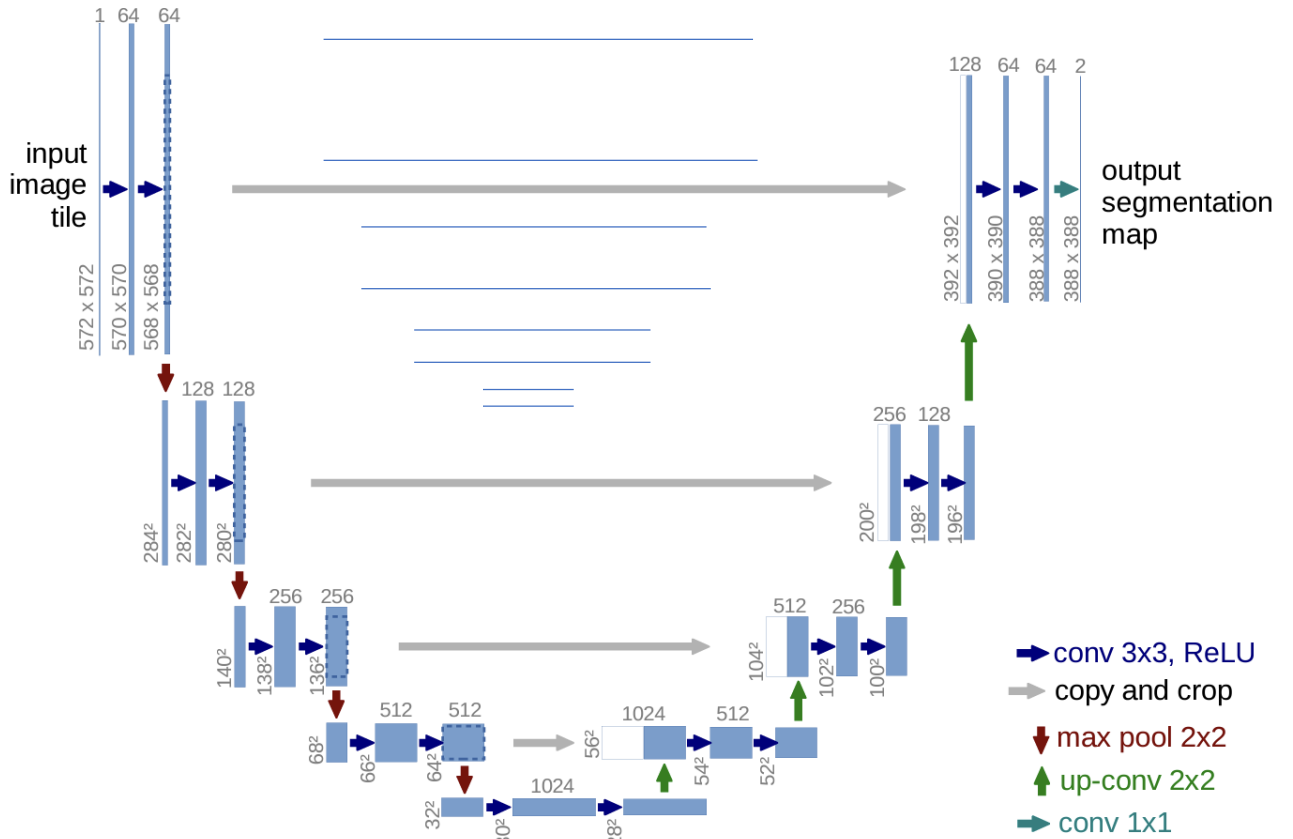
6. Νευρωνικό Δίκτυο UNET

6.1 U-NET

Τα συνελκτικά νευρωνικά δίκτυα όπως αναφέραμε έχουν μια εξαιρετική ικανότητα να εξάγουν χαρακτηριστικά, βοηθώντας καταλυτικά στην ταξινόμηση και την τμηματοποίηση. Έτσι λοιπόν, αναδύθηκαν σε συνδυασμό με τις τεχνικές παράλληλης επεξεργασίας δεδομένων. Παρόλα αυτά όσο εύκολη κι αν είναι η διαδικασία μετάβασης από μια εικόνα σε ένα χάρτη χαρακτηριστικών (convolution), η διαδικασία ανακατασκευής της εικόνας από ένα χάρτη χαρακτηριστικών (deconvolution) αποτελεί μια πρόκληση. Την πρόκληση που αφορά την απαιτητική διαδικασία της ανακατασκευής της εικόνας έρχεται να επιλύσει ένα δίκτυο, το UNET. Είναι ένα δίκτυο το οποίο συναντάμε συχνά σε image segmentation εργασίες μιας και αναπτύχθηκε για να δώσει λύσεις στην ιατρική απεικόνιση. Συχνά απαντάτε στο MRI και στο CT, ειδικά για τμηματοποίηση εγκεφαλικών όγκων όπως θα δούμε παρακάτω.

Η αρχιτεκτονική του μοιάζει με U εξ ου και το όνομα του (βλέπε Εικόνα 31). Χαρακτηριστικά αποτελείται από τρία μέρη: τη συστολή, το bottleneck και το μέρος της διαστολής. Το τμήμα της συστολής αποτελείται από πολλαπλά μπλοκ, η εικόνα εισέρχεται στο δίκτυο και ακολουθούν δύο ακολουθίες συνέλιξης με kernel 3X3 και συνάθροισης με kernel 2X2. Με αυτή την διαδικασία οι χάρτες χαρακτηριστικών διπλασιάζονται διαδοχικά σε αντίθεση με την εικόνα που υποδιπλασιάζεται. Αυτή η διαδικασία είναι το κλειδί ώστε να μάθει το δίκτυο αποτελεσματικά τις πολύπλοκες δομές. Έπειτα το κατώτατο στρώμα συνδέει τα τμήματα της συστολής και της διαστολής. Εδώ απαντώνται δύο επίπεδα συνέλιξης με kernel 3X3 και ένα επίπεδο με kernel 2X2 όπου ξεκινάει η ανακατασκευή της εικόνας (up sampling) οπότε αυτό το επίπεδο ονομάζεται up-conv. Η διαδικασία που ακολουθήσαμε στο τμήμα της συστολής επαναλαμβάνεται στο τμήμα της διαστολής. Κάθε μπλοκ λοιπόν περιλαμβάνει δύο ακολουθίες CNN με kernel 3X3 και μια συνέλιξη ανερχόμενης δειγματοληψίας με kernel 2X2. Το αξιοθαύμαστο σημείο του δικτύου είναι κατά την διαστολή όπου σε κάθε επίπεδο συνέλιξης στο δίκτυο εφαρμόζει copy_concat ώστε να μεταφέρει τα feature maps μέσω των skip connections από το επίπεδο της συστολής. Αυτό εξασφαλίζει την απόδοση του δικτύου, καθώς η διαδικασία μάθησης στο επίπεδο της συστολής χρησιμοποιείται για την ανοικοδόμηση της εικόνας εξασφαλίζονται την διατήρηση της χωρικής πληροφορίας.

Τέλος, η προκύπτουσα εικόνα περνά μέσα από ακόμη ένα επίπεδο CNN 3X3 όπου ο αριθμός των χαρτών χαρακτηριστικών είναι ίσος με τον αριθμό των επιθυμητών τμηματοποιημένων περιοχών. [14],[15],[30],[77]



Εικόνα 31. UNET Architecture [14],[29]

6.2 Attention UNET

Η αρχιτεκτονική του UNET αποτελεί state-of-art για πλήθος εργασιών στην όραση του υπολογιστή. Όπως ήδη αναφέραμε παραπάνω η ιατρική απεικόνιση έχει προχωρήσει σημαντικά χάρις σε αυτό.

Παρόλα αυτά η ανάγκη για αυτοματοποιημένες λύσεις απαιτεί ένα σαφή διαχωρισμό των κατηγοριών της εικόνας ιδιαίτερα στα όρια αυτών. Ένας όγκος για παράδειγμα πρέπει να διαχωρισθεί πλήρως από κάποια παρακείμενη δομή την οποία μπορεί οριακά να προσεγγίσει. Άρα χρειαζόμαστε ένα μηχανισμό που να εστιάζει όσο το δυνατόν σε οριακές περιοχές αγνοώντας ταυτόχρονα την άσχετη πληροφορία.

Ο μηχανισμός που εκπληρώνει το σκοπό αυτό ονομάζεται Attention και χρησιμοποιήθηκε αρχικά στο πεδίο του NLP σε εφαρμογές language-to-language translation για την ακριβή μετάφραση των προτάσεων.

Η λειτουργία του έγκειται στην στόχευση πεδίων που έχουν αυξημένη πιθανότητα να περιέχουν χρήσιμη πληροφορία, αποφεύγοντας άσχετες ενεργοποιήσεις σε νευρώνες από περιοχές της εικόνας που δεν συνεισφέρουν σε πληροφορία. Επίσης, εξασφαλίζεται η μείωση των υπολογιστικών απαιτήσεων αφού το πλήθος των learnable parameters μειώνεται σημαντικά.

Η στόχευση της χρήσιμης πληροφορίας είναι η βασική μας επιδίωξη, καθώς μας επιτρέπει να παρέχουμε καλύτερη γενίκευση στο μοντέλο μας, αυξάνοντας την ευαισθησία του.

Ειδικότερα, ο μηχανισμός του Attention εισάγεται σε συγκεκριμένα σημεία της αρχιτεκτονικής του UNET.

Τα Attention gates τοποθετούνται στο μέρος των skip connections και ακολουθεί μια διαδικασία μεταξύ των δύο εισόδων που δέχεται. Η πρώτη είσοδος είναι ο x πίνακας που έρχεται από το μέρος της συστολής και ο g έρχεται από το αμέσως προηγούμενο επίπεδο στο μέρος της διαστολής όπου η αναπαράσταση των χαρακτηριστικών είναι αρκετά καλύτερη λόγω της προέλευσης της από βαθιά επίπεδα του δικτύου (βλέπε Εικόνα 33).

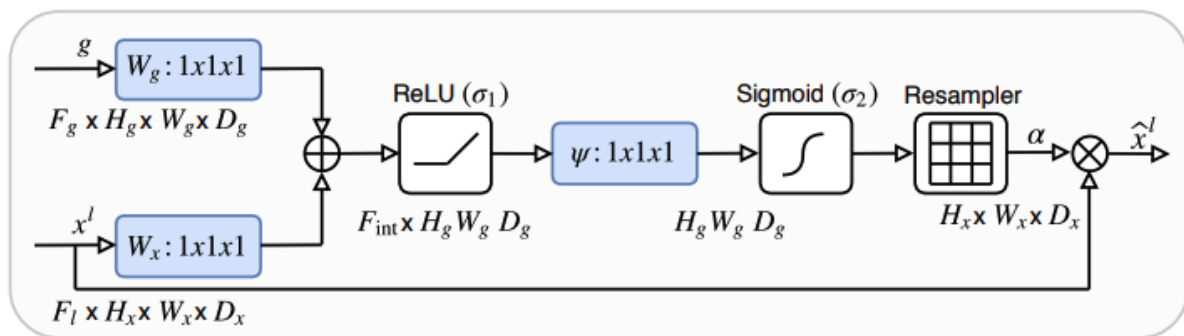
Η διαδικασία που ακολουθείται αποτελείται από:

1. Συνέλιξη με kernel 1×1 ,
2. Άθροισμα $x+g$,
3. ReLU συνάρτηση ενεργοποίησης,
4. Συνέλιξη με kernel 1×1 ,
5. Softmax συνάρτηση ενεργοποίησης,

6. Πολλαπλασιασμό των συντελεστές του Attention με τον πίνακα x

Ο στόχος του αθροίσματος μεταξύ των πινάκων x και g , αποσκοπεί την ανάδειξη των βαρών που είναι ευθυγραμμισμένα και άρα θα γίνουν μεγαλύτερα ενώ το αντίθετο θα συμβεί για τα μη-ευθυγραμμισμένα.

Άρα αναδεικνύουμε την χρήσιμη πληροφορία και υποβαθμίζουμε την ανεπιθύμητη. Επίσης οι συντελεστές του attention είναι βάρη που αποκτούν τιμές $[0,1]$ έπειτα από την συνάρτηση Softmax (βλέπε Εικόνα 32).



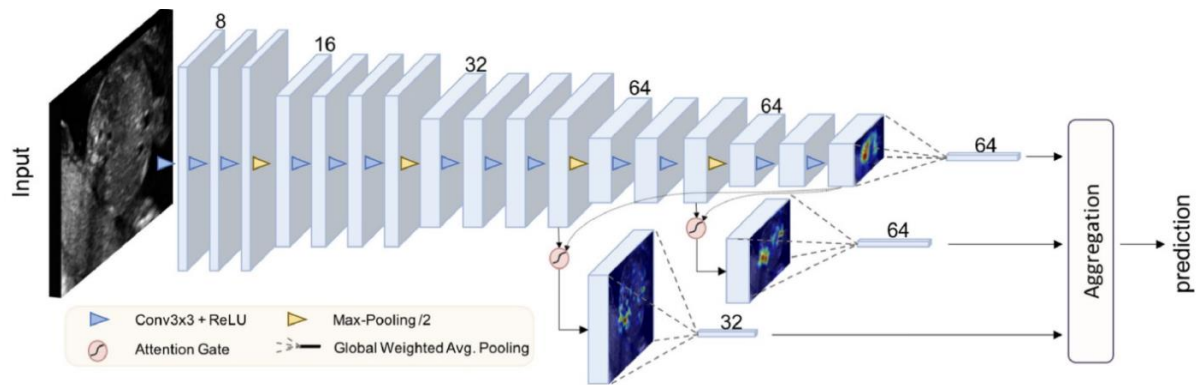
Εικόνα 32. Σχηματική αναπαράσταση της διαδικασίας του Attention gate.

Οι συντελεστές του attention συμβολίζονται με (a) .

Ακολουθεί η αρχιτεκτονική του Attention U-Net με επισυνημμένα τα attention gates. Είναι αποδεδειγμένη η απόδοση του Attention U-Net έναντι του standard U-Net.

Τέλος, παρατηρούμε ότι ο μηχανισμός του Attention που αναπτύξαμε στη παρούσα εργασία υπάρχει στα βαθιά νευρωνικά δίκτυα σιωπηρά καθώς το βάθος των επιπέδων επιτρέπει μόνο στην χρήσιμη πληροφορία να εξαχθεί. Ζητώντας λοιπόν από το δίκτυο να σταθμίσει την εικόνα βάσει των προηγούμενων εισόδων (από μνήμης), το «αναγκάζουμε» να δώσει προσοχή. Αυτό έχει εκφράσει ο Alex Graves το 2020 σε μία διάλεξη λέγοντας «Memory is attention through time. Alex Graves 2020». ^[31]

Άρα ο αναγνώστης θα μπορούσε να συμπεράνει ότι μηχανισμός υπάρχει διαισθητικά. ^{[29], [30]}.



Εικόνα 33.Αναπαράσταση του Attention gate [73]

6.3 Residual U-NET

Ακόμη μία παραλλαγή που υπόσχεται εξαιρετική απόδοση στην τμηματοποίηση είναι ο συνδυασμός του UNET και του residual block. Το Residual UNET αποτελεί μια ακόμη αρχιτεκτονική, η οποία αναπτύχθηκε αρχικά για την εξαγωγή δρόμων από εναέριες εικόνες υψηλής ανάλυσης στον τομέα της τηλεπισκόπησης (βλέπε Εικόνα 35).

Υπάρχει η πεποίθηση ότι τα βαθιά νευρωνικά δίκτυα αποδίδουν καλύτερα σε σχέση με τα ρηγά, στην πράξη όμως αυτό δεν ισχύει. Όσο προσθέτουμε επίπεδα οδηγούμαστε την αύξηση του σφάλματος εκπαίδευσης. Έτσι προκύπτει το πρόβλημα της υποβάθμισης. Σε αυτό το πρόβλημα καλούνται να επιλύσουν τα residual blocks που προσθέτουμε στα νευρωνικά δίκτυα.

Στα παραδοσιακά νευρωνικά δίκτυα κάθε επίπεδο προωθεί την έξοδο του στον επόμενο επίπεδο, το επόμενο επίπεδο την δέχεται σαν είσοδο και ούτω κάθε εξής. Όμως, σε ένα δίκτυο με residual blocks ταυτόχρονα με την προώθηση της εισόδου κατά μήκος του επιπέδου, η είσοδος παραλείπει το επίπεδα και στην έξοδο αθροίζεται με την είσοδο που πέρασε μέσα από αυτά. Το skip connection συνεισφέρει στην διατήρηση της πληροφορίας κατά μήκος του νευρωνικού δικτύου διασφαλίζοντας την απουσία φαινομένων όπως vanishing/exploding gradients. Στην παρούσα εργασία αναφέρεται αναλυτικά το φαινόμενο παρακάτω.

Αυτό καθιστά εφικτό την ανάπτυξη αρχιτεκτονικών με εκατοντάδες επίπεδα. Η διατήρηση της πληροφορίας επιτυγχάνεται λόγω του skip connection, καθώς από την εξίσωση 1 μπορούμε να διαπιστώσουμε ότι αν το $F(x_i, W_i)$ είναι μηδέν τότε η έξοδος ισούται με την identity function (είσοδο). Άρα η είσοδος μπορεί να μεταφερθεί αναλλοίωτη στο επόμενο επίπεδο παρόλο που δεν βελτιώνεται η απόδοση.

Η παραπάνω διαπίστωση αποτελεί παράδειγμα, στην πράξη οι παραμέτροι του μοντέλου αποδίδουν καλύτερα από $F(x_l, W_l) = 0$. Άρα το μοντέλο είναι αποτελεσματικό σε μεγάλο βαθμό. Ιδιαίτερα σε συνελκτικά δίκτυα που έχουν κάποιες εκατοντάδες επίπεδα οπότε η δυσκολία των νευρώνων να μάθουν αυξάνεται σημαντικά. Έτσι σε συνδυασμό με την τεχνική του gradient descent η απόδοση είναι εγγυημένη.

Τέλος, οι παραλλαγές των residual blocks είναι ποικίλλες. Παρακάτω θα εξηγήσουμε την εσωτερική δομή του residual unit που επιλέξαμε.

Η γενική μορφή του residual unit:

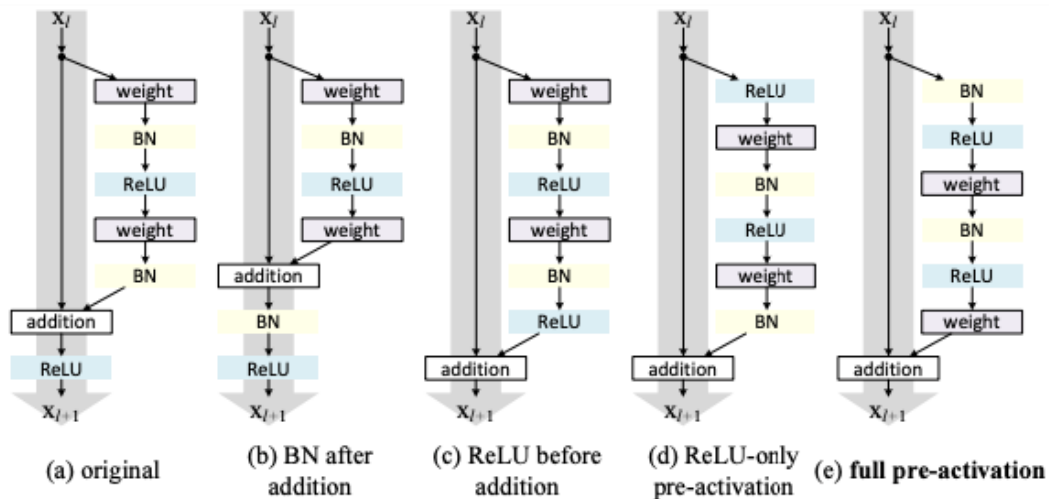
$$y_l = h(x_l) + F(x_l, W_l), x_{l+1} = f(y_l),$$

Όπου που x_l και x_{l+1} είναι η είσοδος και η έξοδος του l-th residual unit, F η residual συνάρτηση, f συνάρτηση ενεργοποίησης, και $h(x_l)$ είναι ο identity map δηλαδή η είσοδος x_l χωρίς να έχει αλλάξει. ^{[32], [33], [34], [64]}

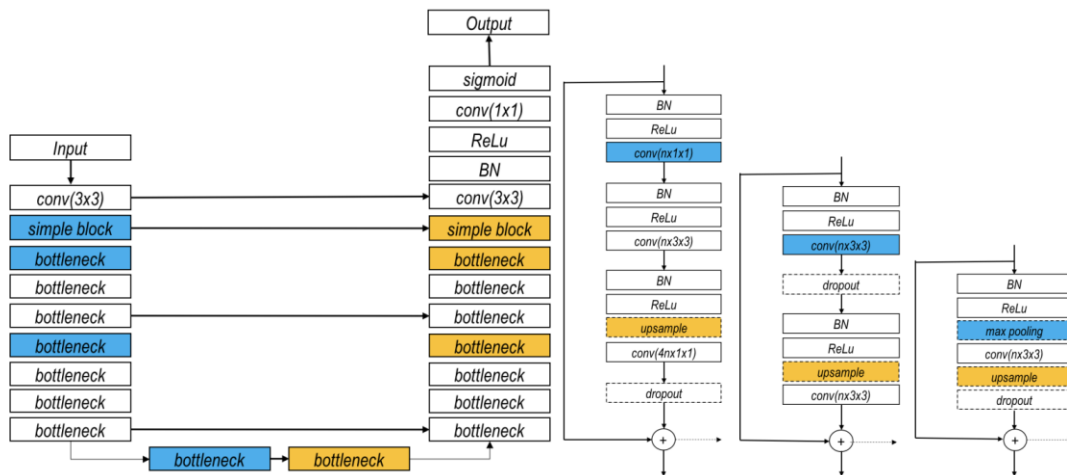
$$\text{Άρα } h(x_l) = x_l$$

6.3.1 Residual Block

Ένα residual block αποτελείται από μια σειρά επιπέδων batch normalization, convolution και ReLU. Οι μαθηματικές φόρμουλες δεν αλλάζουν, όμως αλλάζει η σειρά εφαρμογής τους καθώς έχει ιδιαίτερη σημασία. Η Εικόνα 34 παρακάτω απεικονίζει διαφορετικές προσεγγίσεις για την υλοποίηση του residual unit. Πειραματικά αποδείχθηκε ότι η αλληλουχία σύμφωνα με τη full pre-activation προσέγγιση αποδίδει καλύτερα σε σχέση με τους υπόλοιπους τρόπους, όπως προτείνουν οι He et al. ^[33] Έπειτα το αποτέλεσμα της αλληλουχίας των επιμέρους επιπέδων αθροίζεται με το αποτέλεσμα του νευρωνικού δικτύου και προκύπτει η τελική τιμή για το συγκεκριμένο επίπεδο. Στην παρούσα εργασία δεν πειραματιστήκαμε με την αρχιτεκτονική του residual unit πέρα από την υιοθέτηση της παραπάνω προσέγγισης (βλέπε Εικόνα 34).



Εικόνα 34. Διάφορες παραλλαγές του residual block, αλλαγές παρατηρούνται μόνο στην αλληλουχία όχι στις μαθηματικές φόρμουλες που εφαρμόζονται. [33]



Εικόνα 35. Αρχιτεκτονική ResUNET

6.4 Ensemble Model

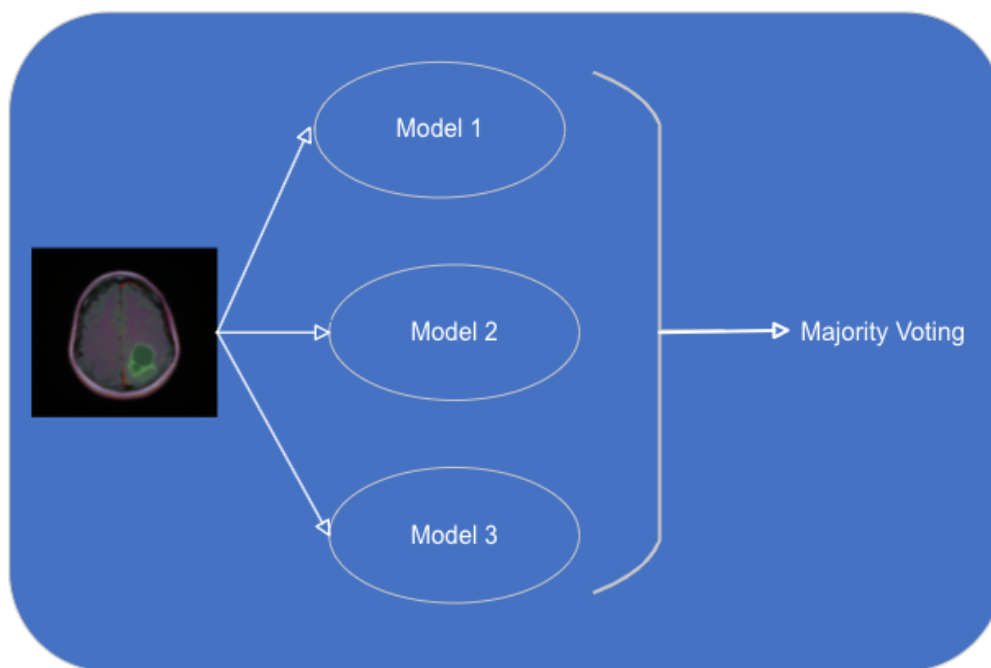
Η τεχνική του ensemble learning αποτελεί ένα συνδυασμό αλγορίθμων στην μηχανική μάθηση ώστε να ληφθεί μια συνεργατική/πλειοψηφική απόφαση. Το πρώτο ensemble system πρωτάθηκε από τους Zhou και Feng [50] αποδεικνύοντας την χρησιμότητα του για την βελτίωση των προβλέψεων καθώς περισσότερα του ενός μοντέλα συνδυάζουν την οπτική τους.

Ο σχεδιασμός του ensemble model αποτελείται από δύο στάδια, την παραγωγή των επιμέρους μοντέλων και την ενσωμάτωση των μοντέλων ώστε να παράξουμε τις προβλέψεις. Οι τρόποι σύντηξης των μοντέλων σε επίπεδο απόφασης είναι οι soft voting δηλαδή λαμβάνουμε τον μέσο όρο των πολλαπλών αλγορίθμων, οι hard voting όπου λαμβάνουμε

υπόψιν μας την πλειοψηφία των μοντέλων (majority vote) και οι adaptive voting που σταθμίζονται τα διάφορα μοντέλα κατά το δοκούν, συχνά τοποθετείται νευρωνικό δίκτυο στο fully-connected layer για να γίνουν οι προβλέψεις.

Οι τεχνικές ensembling χρησιμοποιούνται συχνά στο κλασσικό machine learning καθώς σύμφωνα με την Free Lunch Theorem κανένα μοντέλο δεν είναι καλύτερο από κάποιο άλλο οπότε υλοποιούμε μία γκάμα μοντέλων ούτως ώστε να τα συγκρίνουμε μεταξύ τους.

Στην παρούσα εργασία χρησιμοποιούμε τρία νευρωνικά δίκτυα με τεχνική τύπου hard voting ώστε να λάβουμε καλύτερα αποτελέσματα για την τμηματοποίηση οπότε επί της ουσίας συνδυάζουμε αρχιτεκτονικές μοντέλων και όχι αλγόριθμο (βλέπε Εικόνα 36).^{[65], [66]}



Εικόνα 36. Αναπαράσταση Ensemble Model

6.5 Συνάρτηση Απώλειας

Εκτός από την αρχιτεκτονική, σημαντικό ρόλο έχει η συνάρτηση απώλειας που εφαρμόζουμε. Η έννοια της εντροπίας εισήχθη στην φυσική για να ορίσει την τυχαιότητα σε ένα σύστημα ενώ η διασταυρούμενη εντροπία είναι το μέτρο της διαφοράς μεταξύ δύο κατανομών πιθανοτήτων την οποία και αξιοποιούμε σε αρκετές εφαρμογές στο machine learning. Χρησιμοποιείται ευρέως και σε εργασίες τμηματοποίησης ως συνάρτηση απώλειας καθώς βοήθησε τα blocks του μοντέλου U-NET να προσαρμοστούν σε βιοϊατρικές εικόνες με ασυνεχή τρόπο, έτσι ώστε τα μεμονωμένα pixels να μπορούν εύκολα να αναγνωριστούν μέσα

στον δυαδικό χάρτη τμηματοποίησης. Η ιδέα είναι ότι ακόμη και στην τμηματοποίηση κάθε pixel πρέπει να βρίσκεται σε κάποια κατηγορία και πρέπει απλώς να βεβαιωθούμε ότι το κάνουν. Έτσι απλώς μετατρέψαμε ένα πρόβλημα τμηματοποίησης σε ένα ταξινόμησης πολλών κλάσεων και απέδωσε πολύ καλά σε σύγκριση με τις παραδοσιακές συναρτήσεις απώλειας. (βλέπε Εικόνα 37)

Η παρακάτω εξίσωση (1.6) δίνει τον γενικό ορισμό, ενώ για $C=2$ δηλαδή για δύο κατηγορίες, για παράδειγμα εικόνων (cats-dogs), η εξίσωση (1.6) παίρνει την μορφή της εξίσωσης 1.7.

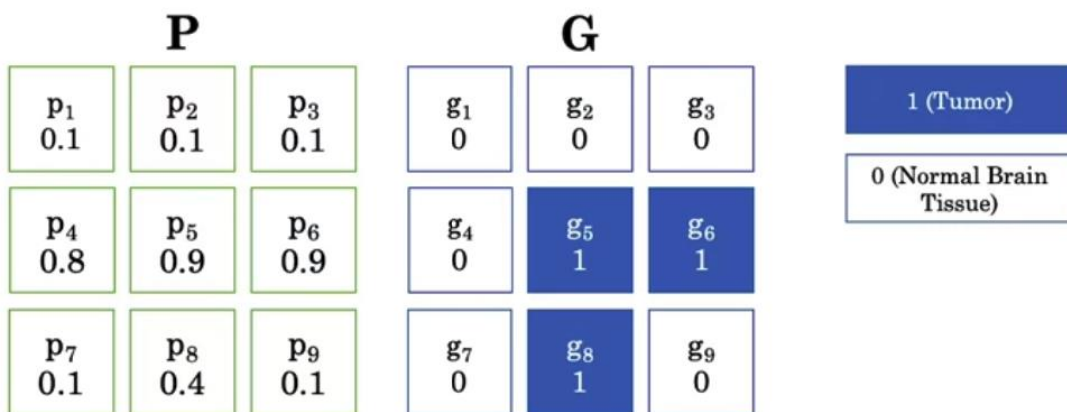
Στην εξίσωση 1.7 ως y καλείται η πραγματική τιμή ενώ \hat{y} η τιμή της πρόβλεψης.

Χρησιμοποιείται ως συνάρτηση απώλειας στο training.

Πέρα από την παρακάτω συνάρτηση υπάρχουν αρκετές που έχουν πετύχει υψηλά ποσοστά όπως η dice similarity loss που χρησιμοποιούν οι Huang et al .^[72] Στην παρούσα διπλωματική δεν την αναλύσαμε καθώς εστίασαμε σε άλλα τμήματα της στρατηγικής σχεδίασης ενός συστήματος βαθιάς μάθησης.

$$\text{Cross Entropy}(y, \hat{y}) = - \sum_i^C y \log(\hat{y}) \quad (1.6)$$

$$\text{Binary Cross Entropy}(y, \hat{y}) = -y \log \hat{y} - (1 - y) \log(1 - \hat{y}) \quad (1.7)$$

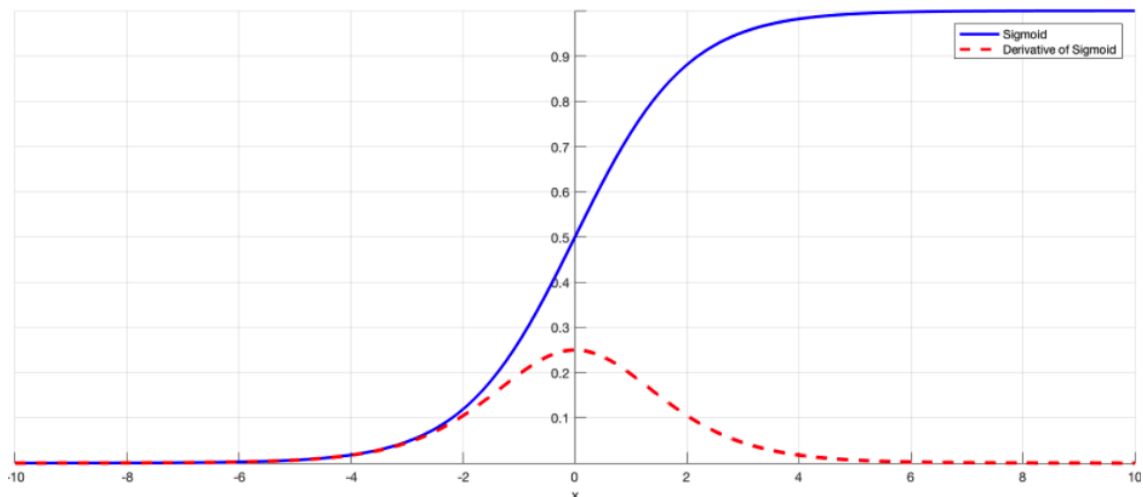


Εικόνα 37. Η τμηματοποίηση της περιοχής ενδιαφέροντος γίνεται βάσει της πιθανότητας του pixel (p_i) P: Prediction, G: Ground^{[35], [36]}

6.6 Το πρόβλημα του Vanishing Gradient

Τα βάρη των νευρωνικών δικτύων είναι ο στόχος κάθε προσπάθειας εκπαίδευσης (training) οποιουδήποτε μοντέλου. Χρησιμοποιούμε πολλές τεχνικές για να κάνουμε fine-tuning το

νευρωνικό μας δίκτυο, ξεκινώντας από τους optimizers μέχρι τις τεχνικές αρχικοποίησης των βαρών. Υπάρχουν πολλά benchmarks όπου οι βελτιστοποιητές (optimizers) συγκρίνονται μεταξύ τους ως προς την ικανότητα τους να ενημερώνουν τα βάρη έως ότου οδηγήσουν τις παραμέτρους του δικτύου στο βέλτιστο σημείο (ολικό ελάχιστο), συνήθως βάσει της ακρίβειας, σε γνωστά datasets. Επίσης, βιβλιογραφικά συναντάμε τεχνικές όπου τα βάρη των νευρώνων αρχικοποιούνται υπό συνθήκες ώστε να εξασφαλιστεί η αξιοπιστία του δικτύου. Αρκετά δημοφιλής είναι η τεχνική Xavier Initialization για την διατήρηση της διακύμανσης των βαρών σε όλα τα κρυφά επίπεδα, η απλή κανονικοποίηση γύρω από το μηδέν. Όμως αυτά που αναφέραμε συμβαίνουν καθώς συναντάμε ένα πρόβλημα το οποίο οφείλεται στη φύση της λογιστική συνάρτηση την οποία χρησιμοποιούμε ευρεία ως συνάρτηση ενεργοποίησης των νευρώνων. Χαρακτηριστικά, πραγματοποιείται συμπίεση των αλλαγών στις τιμές εισόδου προκαλώντας μείωση στις τιμές εξόδου. Αυτό συνεπάγεται οι τιμές εισόδου να επηρεάζουν σταθερά λιγότερο τα βάρη των επιπέδων όσο μετακινούμαστε σε βαθύτερα επίπεδα μέσα στο νευρωνικό δίκτυο. Έτσι αναπόφευκτη είναι η αύξηση του σφάλματος κατά την εκπαίδευση, χωρίς να υπολογίσουμε τις επιπτώσεις στο test set. Χάριν παραδείγματος ας σκεφτούμε μία ακολουθία διαδοχικών πολλαπλασιασμών, στο τέλος της σειράς αν προσθέσουμε ακόμη ένα παρατηρούμε ότι η επίδραση που θα έχει στο αποτέλεσμα είναι ελάχιστη έως και ανεπαίσθητη. (βλέπε Εικόνα 38) [3], [52]



Εικόνα 38. Παραγωγή της λογιστική συνάρτηση που οδηγεί στο φαινόμενο vanishing gradient

6.7 Μονάδα Επεξεργασίας Γραφικών (GPU)

Η GPU είναι ένας επεξεργαστής ο οποίος είναι καλός στον χειρισμό εξειδικευμένων υπολογισμών εν αντιθέσει με μια κεντρική μονάδα επεξεργασίας (CPU) η οποία είναι ένας επεξεργαστής ο οποίος χειρίζεται γενικού τύπου υπολογισμούς. CPUs είναι οι επεξεργαστές που διεξάγουν τους περισσότερους από τους τυπικούς υπολογισμούς στις ηλεκτρονικές μας συσκευές. Παρόλο που , μια GPU είναι αρκετά γρηγορότερη στον υπολογισμό σε σχέση με μια CPU, αυτό δεν μας απασχολεί πάντα. Η ταχύτητα μίας GPU σχετικά με την CPU εξαρτάται από τον τύπο των υπολογισμών που πραγματοποιούνται . Χαρακτηριστικά , ο τύπος των υπολογισμών που διαφοροποιούν τις δύο μονάδες επεξεργασίας είναι αυτές που μπορούν να γίνουν παράλληλα. Αυτή η τεχνική ονομάζεται παράλληλο υπολογισμό (parallel computing).

Ο παράλληλος υπολογισμός είναι ένας τύπος υπολογισμού όπου ένας συγκεκριμένος υπολογισμός διασπάται σε ανεξάρτητους μικρότερους υπολογισμούς που μπορούν να πραγματοποιηθούν ταυτόχρονα. Οι υπολογισμοί που προκύπτουν στη συνέχεια ανασυνδυάζονται ή συγχρονίζονται για να σχηματίσουν το αποτέλεσμα του αρχικού μεγαλύτερου υπολογισμού τον αριθμό.

Οι εργασίες στις οποίες μπορεί να σπάσει ένας μεγαλύτερος υπολογισμός εξαρτώνται από τον αριθμό των πυρήνων που περιέχονται σε ένα συγκεκριμένο κομμάτι πυρήνων υλικού. Αυτές είναι οι μονάδες που πραγματικά κάνουν τον υπολογισμό σε έναν δεδομένο επεξεργαστή και οι CPU έχουν συνήθως 4 ή 16 πυρήνες (cores) ενώ οι GPU έχουν δυνητικά χιλιάδες πυρήνες. Φυσικά υπάρχουν άλλες τεχνικές προδιαγραφές που έχουν σημασία, αλλά αυτή η περιγραφή προορίζεται να οδηγήσει τη γενική ιδέα ώστε να συμπεράνουμε ότι ο παράλληλος υπολογισμός γίνεται χρησιμοποιώντας GPU.

Επίσης μπορούμε να συμπεράνουμε ότι διάφορες ρουτίνες (tasks) μπορούν να εκτελεστούν ταχύτερα σε μια GPU , με βασική προϋπόθεση να μπορούν να σπάσουν σε μικρότερα κομμάτια ώστε να μπορούν να επιταχυνθούν οι πράξεις μέσω των τεχνικών του parallel computing.

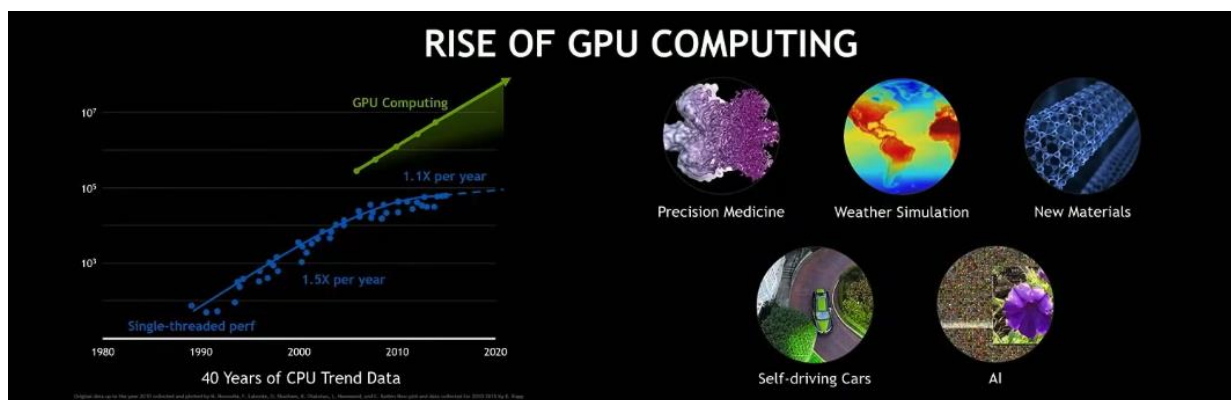
6.7.1 Η GPU στη βαθιά μάθηση

Παραπάνω αναφέραμε το λόγο που οι GPUs είναι κατάλληλες για παράλληλο υπολογισμό και αυτό το γεγονός είναι ο λόγος για τον οποίο η βαθιά μάθηση τα αξιοποιεί. Τα νευρωνικά δίκτυα είναι “ενοχλητικά” παράλληλα. Ο υπολογισμός μιας παράλληλης εργασίας είναι ένα πρόβλημα όπου πρακτικά χρειάζεται από ελάχιστη έως καθόλου προσπάθεια για να χωριστεί η εργασία σε ένα ανεξάρτητο σύνολο μικρότερων ρουτίνων.

Τα νευρωνικά δίκτυα είναι ενοχλητικά παράλληλα και οι GPUs έχουν συνήθως 3000 υψηλής τεχνολογίας πυρήνες που μπορούν να εκτελέσουν υπολογισμούς παράλληλα. Πολλοί από τους υπολογισμούς που καλούμαστε να κάνουμε στα νευρωνικά δίκτυα μπορούν πράγματι να χωριστούν εύκολα σε μικρότερους υπολογισμούς που είναι ανεξάρτητοι μεταξύ τους.

Άρα είναι ξεκάθαρο ότι εφαρμόζοντας τεχνικές παράλληλου υπολογισμού σε ένα από τη φύση του παράλληλο δίκτυο υπολογισμών, η επιτάχυνση της εκπαίδευσης γίνεται με την ταχύτητα του φωτός. Αυτό εξηγεί γιατί νευρωνικά δίκτυα με εκατοντάδες εκατομμύρια παραμέτρους βελτιστοποιούνται σε μικρούς χρόνους εκπαίδευσης.

Κοιτώντας βαθύτερα μέσα στους υπολογισμούς που συμβαίνουν διαδοχικά παρατηρούμε ότι κάθε υπολογισμός είναι ανεξάρτητος από τους άλλους. Αυτό σημαίνει ότι ένας από τους υπολογισμούς εξαρτάται από τα αποτελέσματα των υπολοίπων αλλά όχι από την διαδικασία υπολογισμού του καθενός καθώς αυτοί συμβαίνουν παράλληλα και ταυτόχρονα. Με άλλα λόγια τα αποτελέσματα των υπολογισμών συνδυάζονται ώστε να προκύψει ο τελικός υπολογισμός στο κανάλι εξόδου μετά την ολοκλήρωση όλων των επιμέρους υπολογισμών. Αυτό μας επιτρέπει να δούμε ότι η λειτουργία της συνέλιξης μπορεί να επιταχυνθεί χρησιμοποιώντας μια προσέγγιση παράλληλου προγραμματισμού. [47], [48]



Εικόνα 39. Παρουσίαση των Nvidia GPUs [48]

Πειραματικό Μέρος

7. Δεδομένα και Προεπεξεργασία

7.1 Μορφές Ιατρικών Δεδομένων

Ένας τύπος αρχείων εικόνων όπως τα jpeg, png, tiff είναι και τα αρχεία ιατρικών δεδομένων Dicom και Nifti. Η διαφορά τους έγκειται στην ανάγκη χρήσης κάποιων επιπλέον πληροφοριών ώστε να καταλαβαίνουμε τις παραμέτρους της εικόνας που βλέπουμε. Αυτές οι πληροφορίες ονομάζονται μετα-δεδομένα. Αποθηκεύονται μαζί με το αρχείο της εικόνας και απλώς καλούμε κάποια tags. Για παράδειγμα κάποια χρήσιμα tags είναι αυτά για τον υπολογισμό της περιοχής κάποιας δομής του σώματος, όπως κάποιος όγκος.

Αλλά ακόμα και οι μορφές ιατρικών δεδομένων παρουσιάζουν κάποιες διαφοροποιήσεις μεταξύ τους. Ένα αρχείο dicom αποτελείται από πολλές εικόνες (τομές) οι οποίες συνολικά συνθέτουν μια τρισδιάστατη εικόνα. Τα αρχεία μας υπέστησαν προεπεξεργασία και μετατράπηκαν από dicom σε tiff για την διαδικασία της 2D segmentation.

Βιβλιογραφικά η ανάπτυξη μοντέλων για 2D segmentation υπερισχύουν του 3D segmentation, όπως αναφέρουν οι Niyasa et al. οι υπολογιστικές απαιτήσεις και ο χρόνος εκπαίδευσης περιορίζουν τέτοιες προσεγγίσεις.^[54]

7.2 Dataset

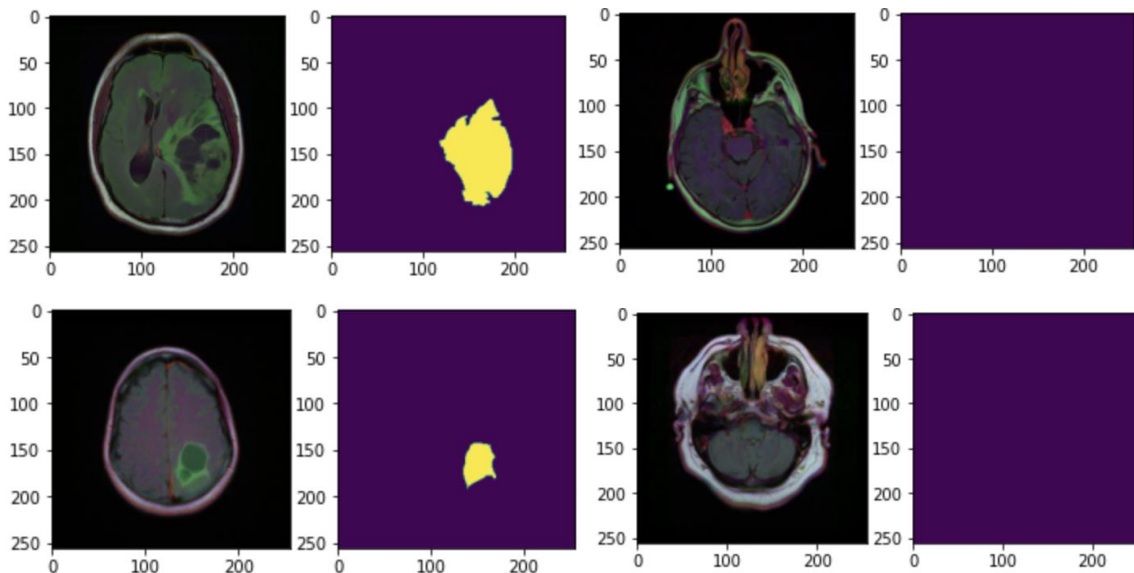
Οι εικόνες που χρησιμοποιούνται στην παρούσα διπλωματική εργασία αποτελούν σύνολο δεδομένων του The Cancer Imaging Archive (TCIA) με τίτλο LGG Segmentation Dataset (TCIA-LGG). Τα δεδομένα αποτελούν συλλογή από χαμηλού βαθμού γλοιώματα με το πρωτόκολλο της T2 – FLAIR (Fluid Attenuated Inversion Recovery) μαγνητικής τομογραφίας. Εκτός από τις εικόνες, παρέχονται και οι μάσκες των τμηματοποιημένων όγκων μεταξύ των 110 ασθενών. Οι ασθενείς χωρίζονται ως εξής βάσει των ιδρυμάτων από τα οποία συλλέχθηκαν : i) Thomas Jefferson University (16 ασθενείς), ii) Henry Ford Hospital (45 ασθενείς), iii) UNC (1 ασθενής), iv) Case Western (14 ασθενείς), v) Case Western – St. Joseph’s (34 ασθενείς). Οι εικόνες είναι διαθέσιμες σε tiff format, το οποίο υποστηρίζει γραφικά υψηλής ποιότητας για την εξαγωγή χρήσιμων patterns. Κάθε ασθενής είχε 20-88 τομές, το πλήθος των τομών ποικίλλει από modality σε modality, λαμβάνοντας υπόψιν τα διαφορετικά settings των μαγνητικών τομογράφων των νοσοκομείων που διεξήχθησαν οι εξετάσεις. Επίσης, οι εικόνες

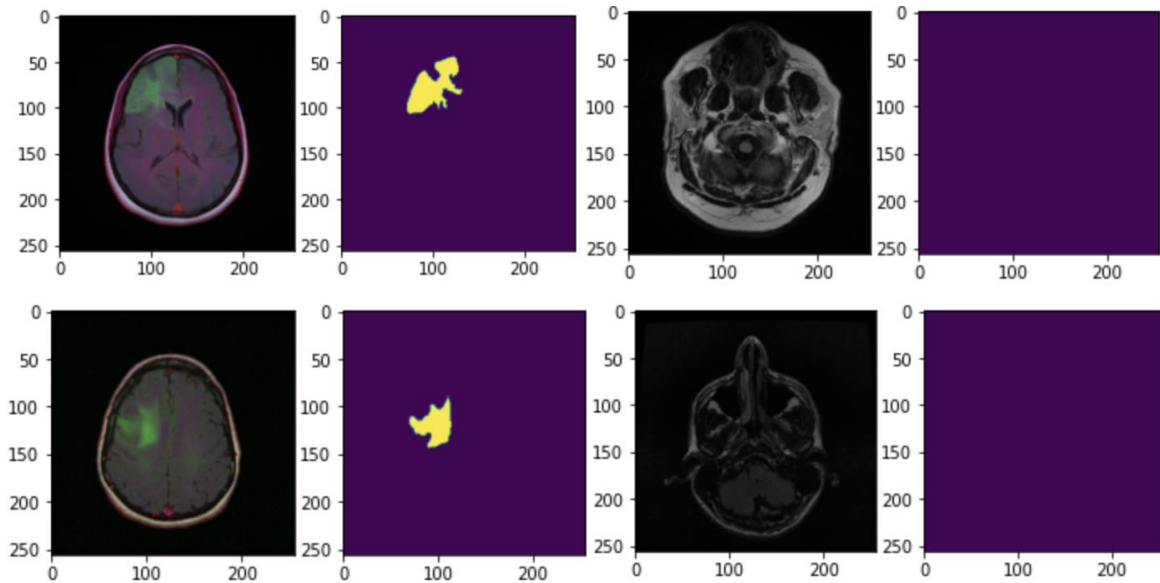
έχουν υποστεί προεπεξεργασία χρώματος, συνεπώς είναι έγχρωμες (RGB), δηλαδή έχουν 3 χρωματικά κανάλια. Αυτό είναι μια καλή στρατηγική όταν υπάρχουν εικόνες οι οποίες λήφθηκαν από διαφορετικούς μαγνήτες οπότε δημιουργείται ένα είδος τυποποίησης σε όλες τις εικόνες.^[75]

Οι μάσκες είναι δυαδικές εικόνες, αποτελούνται από ένα κανάλι, και χρησιμοποιούνται για την ανάδειξη της ανωμαλίας από την FLAIR ακολουθία όπως συνηθίζεται σε περιπτώσεις όγκων στον εγκέφαλο καθώς η αντίθεση μεταξύ όγκου και background είναι υψηλή.^[71]

Κατά την προεπεξεργασία οι εικόνες υπέστησαν αλλαγή μεγέθους με τελικές διαστάσεις 256x256 και οι μάσκες προστέθηκαν στο path των εικόνων. Το σύνολο των 3.929 εικόνων αποτελείται από 2556 εικόνες με όγκους και 1373 εικόνες χωρίς όγκους. Από το σύνολο χωρίστηκαν σε training set (70%), test set (10%) και validation set (20%).

Κατανοώντας τα δεδομένα τα οποία αξιοποιούμε για να εφαρμόσουμε τους αλγορίθμους μηχανικής μάθησης, έχουμε μια ευρύτερη εικόνα για τα αποτελέσματα που περιμένουμε. Συγκεκριμένα, dataset μας εμφανίζει μεγάλη ετερογένεια, οι εικόνες παρουσιάζουν αρκετό θόρυβο και πολλές φορές οι κακοήθειες δεν εμφανίζονται ξεκάθαρα ή χρησιμοποιούνται χρώματα. Έχοντας τα παραπάνω υπόψιν θα παρουσιάσουμε παρακάτω πως δίνουμε λύσεις και βελτιώνουμε τα μοντέλα μας. Ακολουθούν εικόνες μαζί με τις μάσκες τους, αριστερά εικόνες με όγκους ενώ δεξιά απεικονίζονται τομές χωρίς όγκους. (βλέπε Εικόνα 40)^[51]

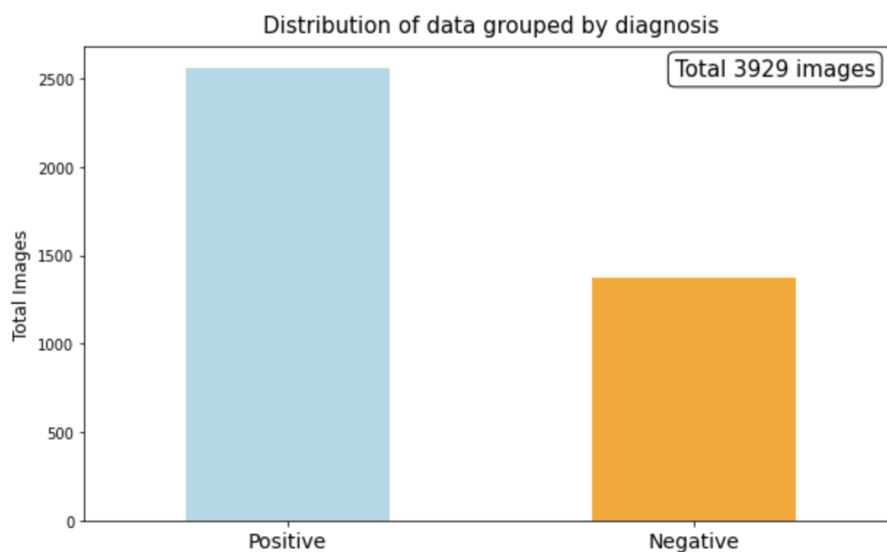




Εικόνα 40. Brain MR Images του Dataset

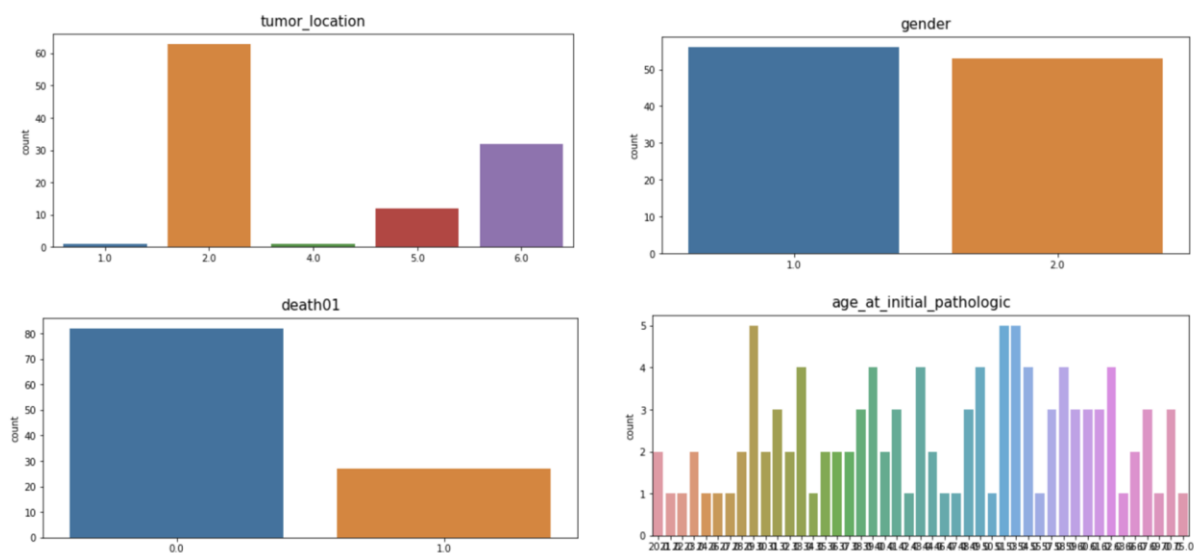
Οι παραπάνω εικόνες είναι low grade gliomas, παρακάτω θα αναλύσουμε κάποια από τα metadata ώστε να λάβουμε πληροφορία για τα γενομικά χαρακτηριστικά των ασθενών. Αυτά αποτελούν πληροφορία η οποία βρίσκεται κρυμμένη στις αλληλουχία των γονιδίων του DNA. Χρησιμοποιώντας υπολογιστικές και στατιστικές μεθόδους προσπαθούμε να αποκωδικοποιήσουμε την πληροφορία.

Όσο αφορά τα χαμηλού βαθμού γλοιώματα, παραθέτουμε την Εικόνα 41 σχετικά με την αναλογία των θετικών και αρνητικών εικόνων, αναλύσεις σχετικά με την γενομικά χαρακτηριστικά, την θέση του όγκου στον εγκέφαλο, την ηλικία εμφάνισης, το φύλο και το survival. (βλέπε Πίνακα 7.1 και Εικόνα 42) ^[51]



Εικόνα 41. Countplot για την κατανομή των κλάσεων

Παρατηρούμε ότι πρόκειται για unbalanced dataset, όπως αναφέρουμε σε αρκετά σημεία.



Εικόνα 42. Περιγραφή των δεδομένων σχετικά με την θέση της παθολογίας, το φύλο, την εξέλιξη της νόσου και την ηλικία εμφάνισης της παθολογία.

Πίνακας 7.1 Δεδομένα σχετικά με το φύλο και την ηλικία των περιστατικών

Characteristics	Patients(N=110)
Age(years)	
Median	47
Range	20-75
Gender	
Female	56
Male	32
Not Available	1

Έπειτα τα ιστοπαθολογικά δεδομένα που αφορούν την ομαδοποίηση των περιστατικών με βάση την μεθυλίωση του DNA χωρίζονται σε Methylation , RNA Sequence, miRNA, Reverse Phase Protein lysate Array (RPPA), CN και Oncogenic signature (βλέπε Πίνακα 7.2). Παρόλο που είναι βέβαιη η συσχέτιση με μεταλλάξεις στα γονίδια IDH, TP53 και ATRX των χρωμοσωμικών βραχιόνων 1p/19q, δεν είναι επαρκής η ιστολογική κατηγορία ενώ επικρατεί βεβαιότητα για την εξέλιξη των χαμηλών βαθμού γλοιομάτων σε γλοιοβλάστωματα (GBM).

Καθώς παρουσιάζουν αρκετά μεταβλητή κλινική συμπεριφορά και οδηγούν τους ιατρούς σε διχογνωμίες όσο αφορά την διάγνωση. Τέλος, οδηγούμαστε με σιγουριά, στην διαπίστωση ότι ιστολογικοί υπότυποι, διαφέρουν μεταξύ τους ως προς την πορεία της νόσου και την συνολική επιβίωση. ^{[51], [53]}

Πίνακας 7.2 Ανάλυση των ιστολογικών υποτύπων στο LGG Dataset

Clusters	Num of Classes	Description
IDH mutation & 1p/19q co-deletion	3	mutation/no-mutation, Wild type
RNASeq	4	R1-R4
DNA methylation	5	M1-M5
DNA copy number	3	C1-C3
microRNA expression	4	mi1-mi4
Cluster of clusters	3	coc1-coc3

Τέλος παραθέτουμε τον πίνακα 7.3 όπου αναφέρουμε συγκεντρωτικά τους υποτύπους των περιστατικών που περιλαμβάνουν το dataset, χωρίς να εστιάσουμε στην τμηματοποίηση ανά κατηγορία.

Πίνακας 7.3 Υπότυποι και βαθμοί κακοήθειας

Characteristics	Cases (N=110)
Histologic type and grade	
Astrocytoma	
Grade II	8
Grade III	25
Oligoastrocytoma	
Grade II	14
Grade III	21
Oligodendroglioma	
Grade II	29
Grade III	18
Not Available	1
IDH-1p/19q subtype	
IDH mutation, 1p/19q co-deletion	26
IDH mutation, no 1p/19q co-deletion	56
IDH wild type	25
Not available	3

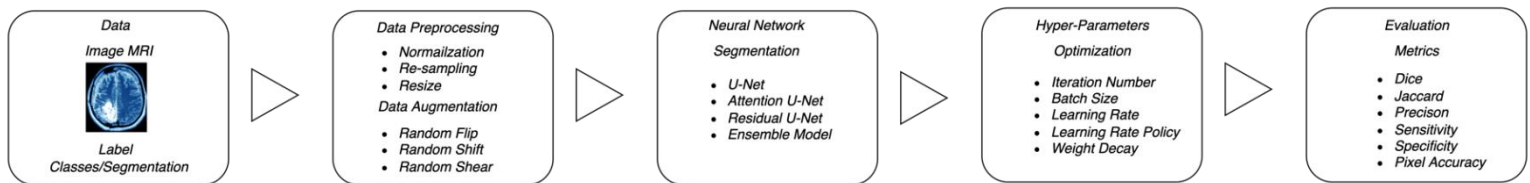
Ο πίνακας που ακολουθεί περιγράφει τον τρόπο που χωρίσαμε το dataset μας, βάσει της επιλογής των νευρωνικών δικτύων. Βιβλιογραφικά παρουσιάζονται ποικίλοι τρόποι όπως αυτό των Ghosh et al [55] που χωρίζουν το dataset σε training set με 3143, σε validation και test set με 393 εικόνες αντίστοιχα. Συνήθως η στρατηγική αποτελεί πεδίο πειραματισμού παρόλο που υπάρχει ο γενικός κανόνας που μιλάει για training set με το 70% του συνολικού dataset.

Πίνακας 7.4 Στρατηγική χωρισμού του Dataset

	Training Set	Validation Set	Test Set
No. of images	2750	384	795

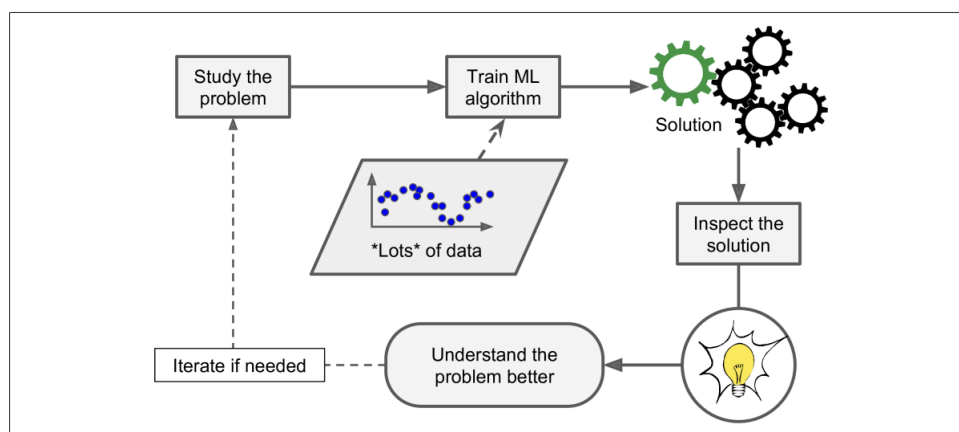
Διάγραμμα Ροής

Ακολουθεί μια σύνοψη της διαδικασίας που ακολουθήσαμε από την λήψη και προεπεξεργασία των δεδομένων μέχρι την ανάπτυξη των μοντέλων και την αξιολόγηση τους. Το παρακάτω διάγραμμα παρουσιάζει όλα τα σημαντικά βήματα της μεθοδολογία που σχεδιάσαμε και αναπτύξαμε σε περιβάλλον της γλώσσας προγραμματισμού Python.



Σε όλο το μήκος της διαδρομής μας στην ανάπτυξη και την διερεύνησή των αλγορίθμων υπήρχαν προβλήματα τα οποία χρειάστηκαν χειροκίνητες λύσεις επιφέροντας αλυσιδωτές αλλαγές σε αρκετούς κανόνες των αλγορίθμων μας. Επίσης πολύπλοκα προβλήματα για οποία δεν υπήρχαν βέλτιστες λύσεις, αντιμετωπίστηκαν με παραδοσιακές τεχνικές Μηχανικής Μάθησης.

Τέλος, η λήψη αποφάσεων γίνεται βάσει αποτελεσμάτων σε μεγάλη κλίμακα δεδομένων ώστε να υπάρχει μια ευρεία εικόνα. Παρακάτω η εικόνα συνοψίζει την στρατηγική με την οποία αναπτύξαμε το pipeline μας και επιλύσαμε πολλά από τα προβλήματα. Η εικόνα αποτελεί state-of-art διάγραμμα ροής για κάθε πρόβλημα στο πεδίο της Μηχανικής Μάθησης.



Εικόνα 43. Προσέγγιση στρατηγικής Μηχανικής Μάθησης ^[66]

7.3 Προεπεξεργασία

Κάθε τμήμα του διαγράμματος ροής υλοποιήθηκε χρησιμοποιώντας γνωστές βιβλιοθήκες και frameworks της Python. Συγκεκριμένα χρησιμοποιήσαμε βιβλιοθήκες όπως η Scikit-Learn, Sklearn, Numpy, Pandas, Matplotlib, OpenCV. Όλες οι βιβλιοθήκες χρησιμοποιήθηκαν πάνω στα Machine learning frameworks Tensorflow, Pytorch.

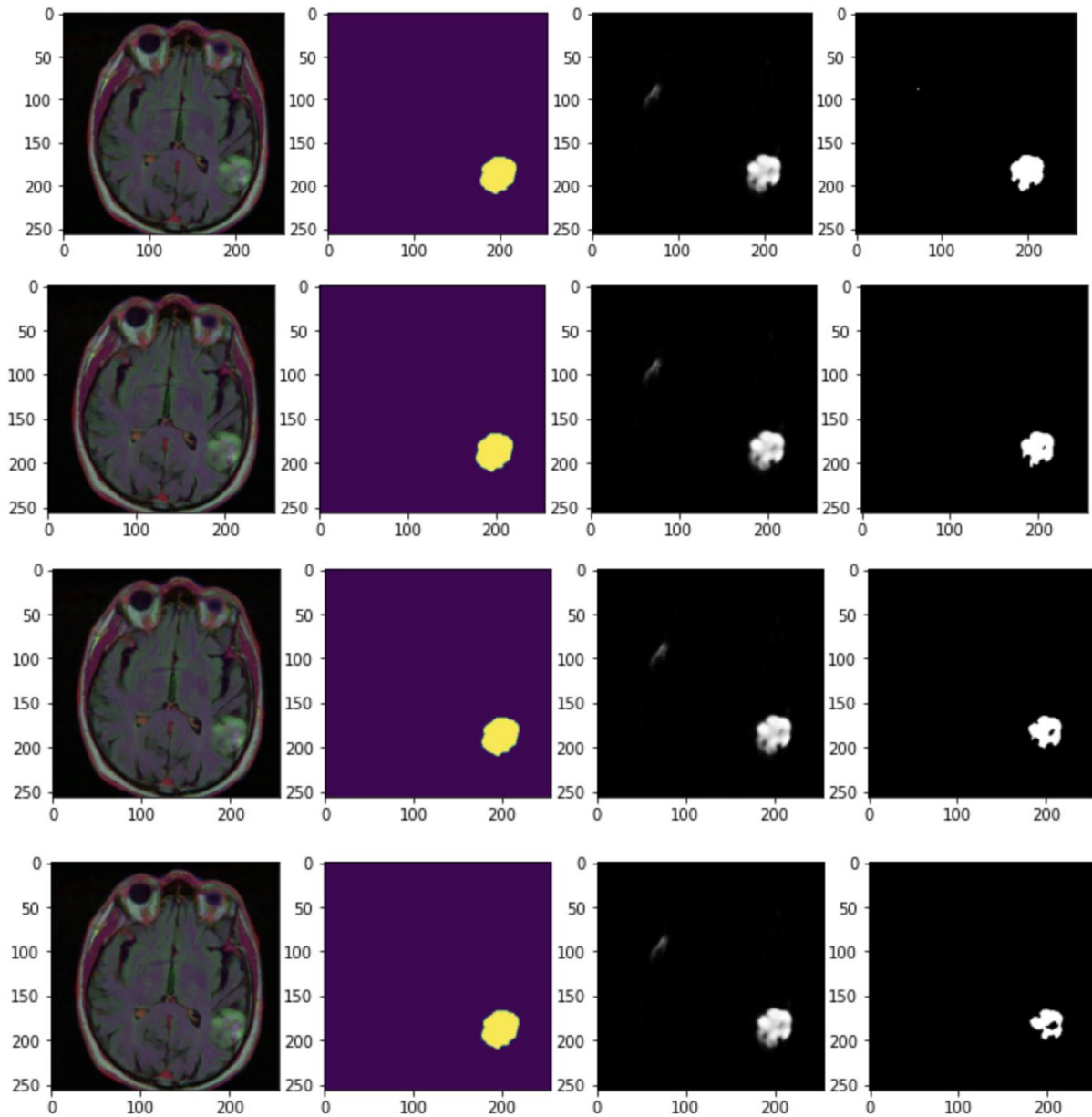
Μετά την επεξεργασία και την μεταφόρτωση των δεδομένων στους φάκελους training, test και validation, μετασχηματίσαμε τις διαστάσεις των εικόνων μας ώστε να μπορέσουν να προωθηθούν στην αρχιτεκτονική του μοντέλου μας. Ακολουθούμε κοινή στρατηγική για όλα τα μοντέλα μας και μετασχηματίσαμε τις εικόνες μας σε διαστάσεις 224x224. Το πλεονέκτημα που μας δίνεται αφορά το μέγεθος των εικόνων, το οποίο μειώνεται δραστικά και συνεπώς μειώνεται ο χρόνος εκπαίδευσης. Επίσης, ελέγχοντας τις μάσκες παρατηρήθηκαν μικρές διακυμάνσεις οπότε τις απαλείψαμε χρησιμοποιώντας φίλτρα για δυαδικές εικόνες. Τέλος, τεχνικές data augmentation χρησιμοποιήθηκαν ώστε να αυξήσουν τη μεταβλητότητα του dataset. Ενδεικτικά χρησιμοποιήσαμε μεθόδους shifting, shearing και flipping καθώς παρέχονται έτοιμες από τα πακέτα της python.

Κατωφλίωση Προβλέψεων

Ένα σημαντικό βήμα το οποίο αποτελεί μέρος της στρατηγικής μας μετά την διαδικασία της εκμάθησης από τα μοντέλα είναι η κατωφλίωση των εικόνων των όγκων ώστε να διατηρηθούν τα pixels που ο αλγόριθμος εκτιμά υψηλή πιθανότητα για την ύπαρξη της κακοήθειας. Να σημειωθεί ότι η εργασία της τμηματοποίησης αποτελεί μια μεγάλης κλίμακας ταξινόμηση των pixels σε κατηγορίες. Έτσι λοιπόν ακολουθούμε ένα post-processing stage στο οποίο σαρώνουμε τις μάσκες ώστε να δημιουργήσουμε δυαδικές εικόνες που απεικονίζουν βάσει της πιθανότητας που προκύπτει από την πρόβλεψη. Η μάσκα αποτελεί ένα πίνακα που τα pixels φέρουν τιμές 0 και 1. Όσο κινούμαστε προς το 1 οι αποχρώσεις των pixels είναι άσπρες ενώ στο 0 επικρατεί το μαύρο. Μετά την τροφοδότηση της εικόνας στο μοντέλο, ο παραγόμενος πίνακας (array) περιέχει τιμές σε εύρος 0-1 οι οποίες αντιπροσωπεύουν την πιθανότητα ένα pixel να υπάγεται σε μία κατηγορία. Αυτό το πίνακα μετατρέπουμε σε δυαδικό χρησιμοποιώντας την κατωφλίωση.

Η κατωφλίωση γίνεται μέσω μιας τιμής την οποία επιλέγουμε πειραματικά λαμβάνοντας υπόψιν τα μέτρα αξιολόγησης. Μετά από πλήθος πειραμάτων θέσαμε την τιμή 0.45, θεωρώντας ότι τα μοντέλα μας έχουν μια απόκλιση όσο απομακρύνονται από τον πυρήνα της κακοήθειας λόγω της μεταβλητότητας των εικόνων του dataset. Αυτό διαπιστώνεται εύκολα

από τις εικόνες που θα ακολουθήσουν στην ανάπτυξη της παρούσας εργασίας. Παρακάτω παρατίθενται οι πίνακες με τα μετρικά που λήφθηκαν χρησιμοποιώντας αυθαίρετες τιμές κατωφλίωσης. Η επιλογή της τιμής κατωφλίωσης εξαρτάται από την επίδοση του συντελεστή Dice και του πίνακα αλήθειας όπως παρουσιάζεται παρακάτω (βλέπε Εικόνα 44 και Πίνακα 7.5).



Εικόνα 44. Πειραματικά δεδομένα αριστερά η τομή της μαγνητικής, δίπλα η μάσκα της τομής, δεξιότερα η πρόβλεψη του μοντέλου του UNET και δεξιότερα η επιπτώσεις της κατωφλίωσης στις εικόνες. Μπορεί κάποιος να παρατηρήσει την απουσία εικονοστοιχείων στον πυρήνα του όγκου στην εικόνα της μαγνητικής τομογραφίας

Πίνακας 7.5 Ανάλυση των μετρικών βάσει της προσαρμογής της τιμής κατωφλίωσης

Εικόνες	Threshold	TP/TN	FP/FN	Dice Coefficient
1	0.45	1065/64133	62/276	0.86
2	0.6	947/64150	45/394	0.81
3	0.7	855/64158	37/486	0.77
4	0.8	705/64169	26/636	0.68

7.4 Εκπαίδευση Μοντέλων

Παρακάτω θα αναπτύξουμε την μεθοδολογία για την εκπαίδευση των τριών νευρωνικών δικτύων. Σε αυτό το σημείο παραθέτουμε ένα πίνακα με τα trainable parameters επισημαίνοντας το μέγεθος του attention UNET. Το μέγεθος του είναι μεγάλο σχετικά με τα άλλα δύο, παρόλα αυτά δεν πετυχαίνει την μεγαλύτερη απόδοση κοιτώντας παρακάτω τα μετρικά του λαμβάνοντας υπόψιν το μέγεθος και τις υπολογιστικές του απαιτήσεις.

Πίνακας 7.6 Model trainable parameters

Model	Parameters
UNET	1,179,409
Attention UNET	32,094,725
Residual UNET	1,211,473

7.4.1 Η εκπαίδευση του UNET

Το πρώτο νευρωνικό δίκτυο που εκπαιδεύσαμε είναι το UNET, το οποίο αποτελεί state-of-art για segmentation tasks. Στο medical imaging αποδίδει ικανοποιητικά σε διαφορετικά modalities και σε διαφορετικούς τύπους ιστών. Εμπειρικά, ο εγκέφαλος αποτελεί δημοφιλή εφαρμογή segmentation και λόγω των ποικίλων datasets που υπάρχουν διαθέσιμα, μερικά από τα πιο γνωστά challenges είναι το MICCAI, BRATS και το Decathlon. ^[55]

Η αρχιτεκτονική αποτελείται από το μέρος της συστολής, το μέρος της διαστολής και το bottleneck μέρος. Αξίζει να αναφέρουμε ότι η διατήρηση της χωρικής διακριτικής ικανότητας εξασφαλίζεται χάρις στα skip connections που προωθούν τα feature maps στο μέρος της διαστολής (βλέπε Εικόνα 45).

Αποτελείται από 4 blocks από διαδοχικά convolution, batch normalization και maxpooling. Αρχικοποιείται με kernel μεγέθους 3x3 και σε κάθε μπλοκ τα φίλτρα από τα οποία προκύπτουν

τα feature maps διπλασιάζονται, το αρχικό μέγεθος είναι 16x16 και φτάνει μέχρι 256x256. Η χρήση kernel 3x3 εγγυάται την σταδιακή μείωση των διαστάσεων της εικόνας ώστε να μην χαθεί χρήσιμη πληροφορία σχετικά με την παθολογία του ασθενούς.

Η συνάρτηση ενεργοποίησης είναι η γνωστή ReLU ενώ το max pooling έχει kernel 2x2 και χρησιμοποιείται μόνο στο μέρος της συστολής. Ορίζουμε την παράμετρο του dropout ως 0.1 ώστε να επηρεάσει το generalizability του μοντέλου απενεργοποιώντας διαφορετικούς νευρώνες με το πέρας των εποχών.

Η αρχικοποίηση των βαρών του νευρωνικού δικτύου ακολούθησε την μέθοδο Xavier initialization, δηλαδή στάθμιση γύρω από το 0 και διακύμανση γύρω από μια αυστηρή τιμή. Η συνάρτηση απώλειας είναι η διασταυρούμενης εντροπίας και αποτελεί κύρια επιλογή για ταξινόμηση και τμηματοποίηση και ορίζουμε μετρικά τους accuracy, precision και τους συντελεστές Dice και IoU. Οι γραφικές δηλαδή θα απεικονίζουν τα scores των μετρικών με το πέρας των εποχών.

Στην εκπαίδευση χρησιμοποιούμε Adam optimizer ώστε να ενημερώνει τα βάρη κατά την οπισθοδρόμηση χωρίς weight decay policy. Επίσης χρησιμοποιήσαμε τεχνικές data augmentation, περιστροφή με μικρό εύρος γωνίας, μετατόπιση στον δυσδιάστατο επίπεδο, zoom και shearing (θόλωμα λόγω περιστροφής) σε μικρό βαθμό.

Επιπλέον μέτρα για την αντιμετώπιση του overfitting πήραμε εφαρμόζοντας τεχνικές early stopping και ReduceLROnPlateau οι οποίες στοχεύουν στον τερματισμό της εκπαίδευσης όταν το loss δεν μειώνεται χρησιμοποιώντας μια παράμετρο patience η οποία υποδηλώνει τον αριθμό των εποχών που το loss μένει σταθερό και η μείωση του ρυθμού εκμάθησης ώστε να η βαθμωτή σύγκλιση να προχωρήσει καθώς φαίνεται να είναι εγκλωβισμένη σε κάποιο τοπικό ελάχιστο στην προσπάθεια εύρεσης του ολικού ελάχιστου, χρησιμοποιείται μια παράμετρος patience που θα ρυθμίσει μετά από πόσες εποχές χωρίς αύξηση των μετρικών θα επέλθει μείωση του ρυθμού εκμάθησης. Στην περίπτωση μας ορίζουμε patience=10 για το early stopping και patience, 5, για το ReduceLROnPlateau (βλέπε Πίνακα 7.7).

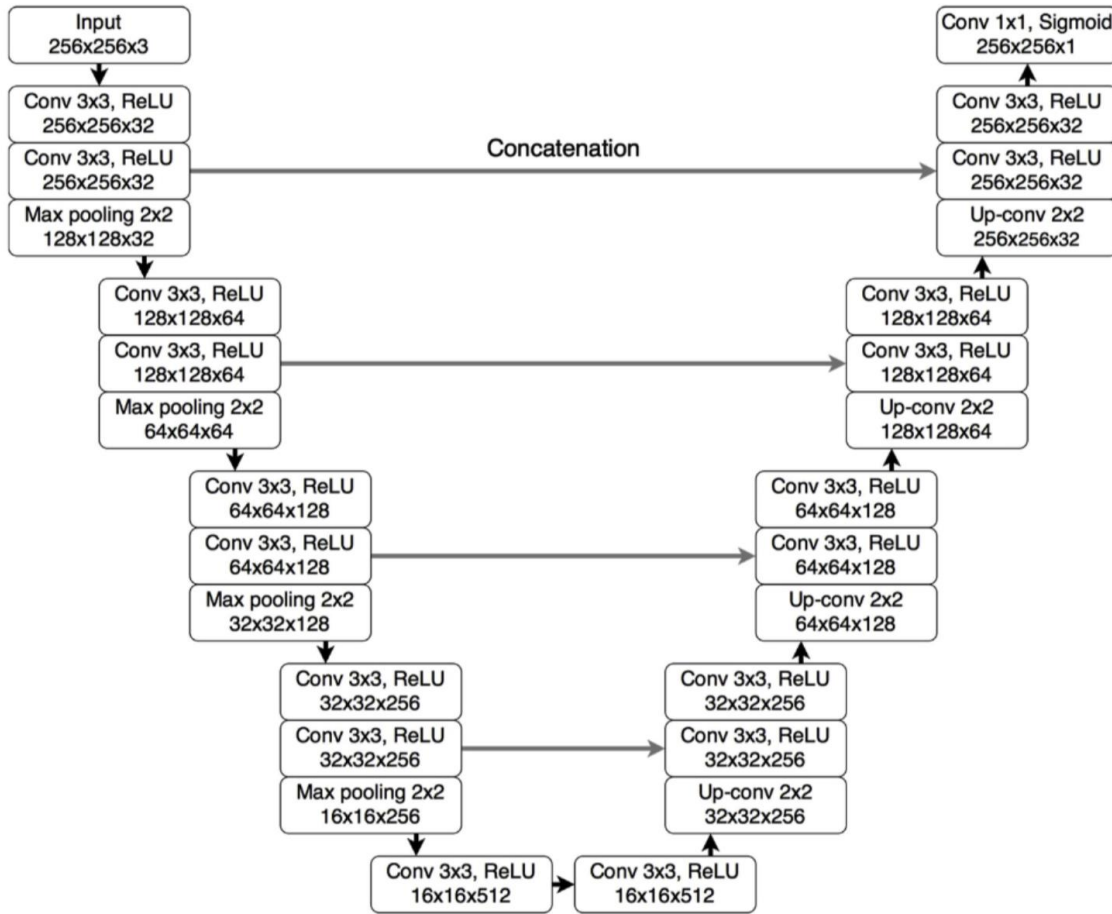
Λογικό θα μπορούσαμε να πούμε καθώς αν χρησιμοποιούσαμε τις τιμές αντίστροφα το μοντέλο μας θα σταματούσε την εκπαίδευση σχετικά γρήγορα και θα δεν θα είχε κάποια σημαντική απόδοση (underfitting). Στην παράμετρο του το ReduceLROnPlateau ορίζουμε εμπειρικά ένα ελάχιστο ρυθμό εκμάθησης ώστε να μην παδιγεύσουμε την βαθμωτή σύγκλιση σε κάποια τοπικό ελάχιστο.

Έπειτα ορίζουμε τον αριθμό των εποχών στις 35 και τα βήματα που θα γίνουν σε κάθε εποχή διαιρώντας τον αριθμό των εικόνων με το batch size.

Συμπερασματικά, η εκπαίδευση διήρκησε 32 εποχές και κλήθηκε η μέθοδος early stopping καθώς όπως θα δούμε και παρακάτω το loss δεν μειώθηκε περαιτέρω. Η τεχνική ReduceLRonPlateau ρύθμισε το ρυθμό εκμάθησης μέχρι την ελάχιστη τιμή.

(βλέπε Πίνακα 7.7)

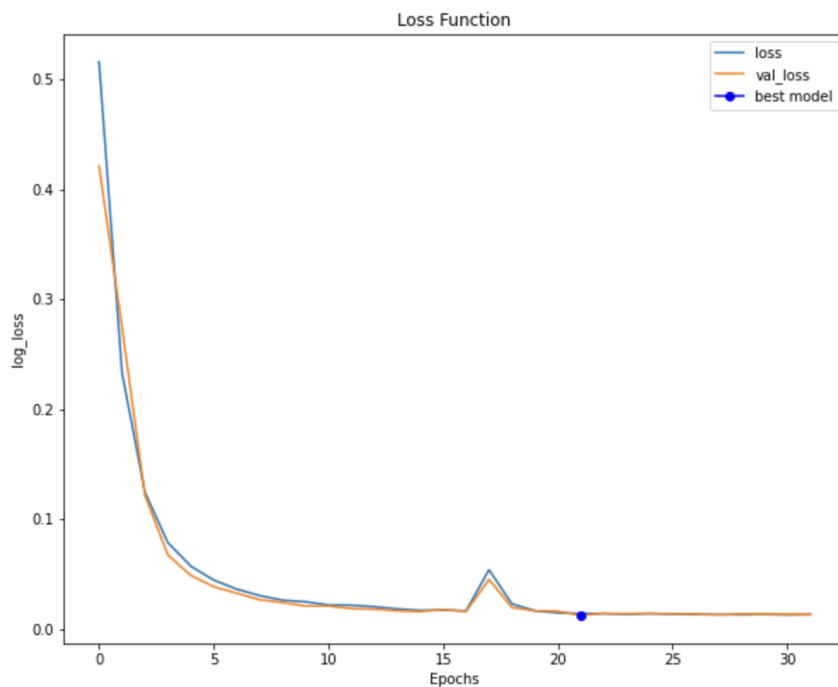
Ακολουθούν τα διαγράμματα για τα μετρικά που ορίσαμε τόσο για training όσο και για validation (βλέπε Εικόνες 46-47) και κάποιες εικόνες με προβλέψεις (βλέπε Εικόνα 48). Όλες οι τιμές που αναφέρονται παραπάνω είναι αποτελέσματα επαναλαμβανομένων περαμάτων.



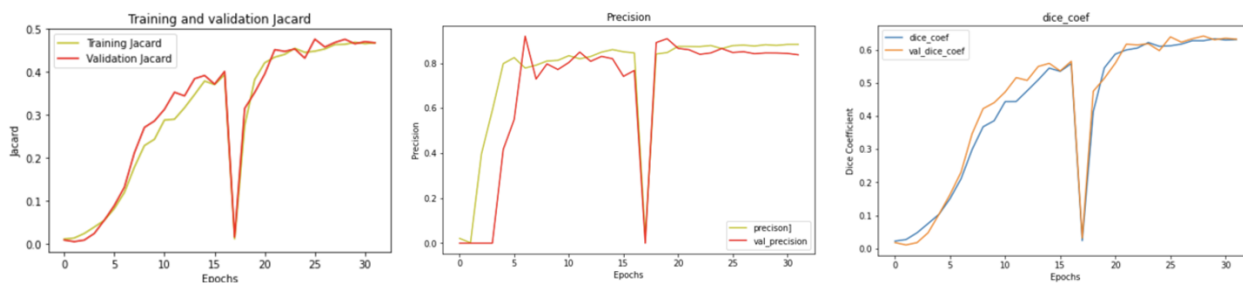
Εικόνα 45. Λεπτομερής Αρχιτεκτονική του UNET. ^[51]

Πίνακας 7.7 Λεπτομερής παρουσίαση των υπερπαραμέτρων που χρησιμοποιήθηκαν στο πείραμα

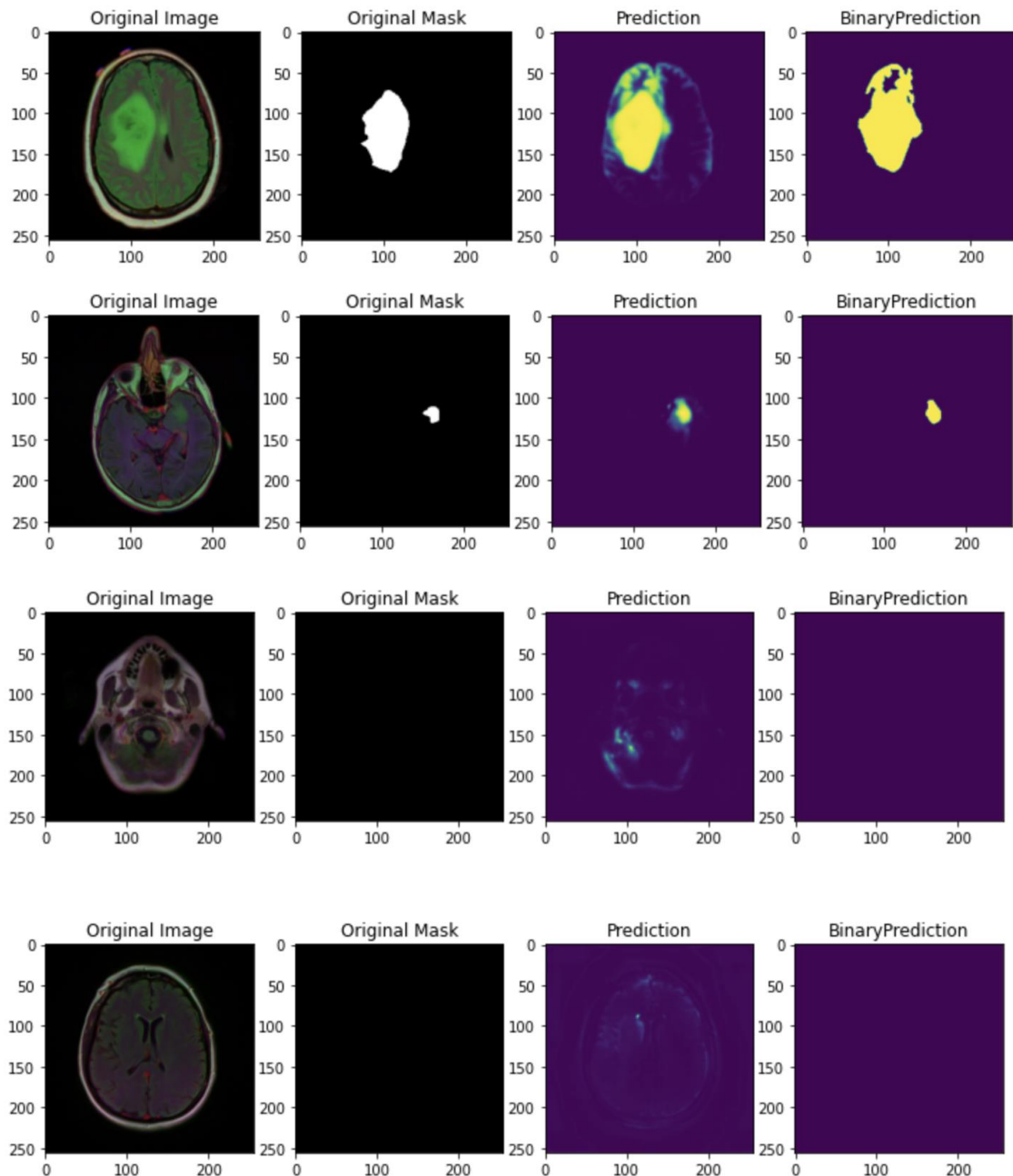
Batch Size	Optimizer	Loss Function	Learning rate	Epochs	Early Stopping	Reduce LRonPlateau
32	Adam	BinaryCrossEntropy	$1e^{-5}$	35	Patience : 10	Patience : 5



Εικόνα 46. Γράφημα της συνάρτησης απώλειας κατά το training

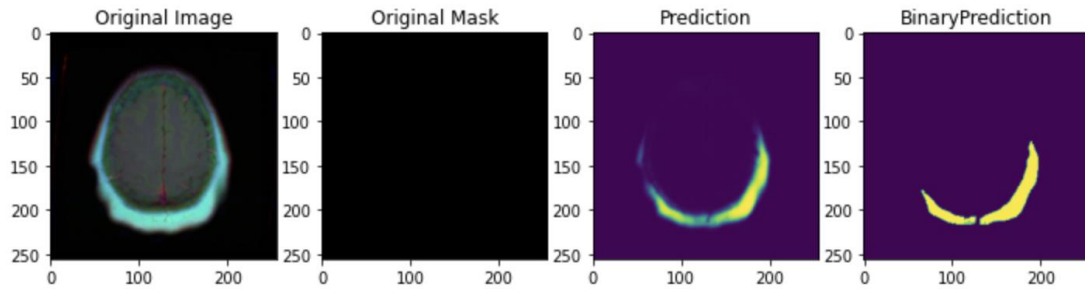


Εικόνα 47. Διαγράμματα των μετρικών Jaccard, Precision και Dice



Εικόνα 48. Προβλέψεις στο validation set

Η χρήση RGB εικόνων επηρεάζει τις προβλέψεις αφού σε αρκετά μεγάλο εύρος της εικόνας προκύπτουν misclassified pixels. Τέτοιες περιπτώσεις δεν εμφανίζονται συχνά παρόλα αυτά δείχνουν να μειώνουν την γενικά ικανοποιητική απόδοση του UNET και των παραλλαγών του όπως θα δούμε και στην συνέχεια. Επίσης, πρέπει να επισημανθεί ότι η προεπεξεργασία αποτελεί μέρος του dataset το οποίο διατίθεται στο Kaggle.



Εικόνα 49. Παράδειγμα κακής πρόβλεψης που οφείλεται στην προεπεξεργασία με έγχρωμα κανάλια, στην εικόνα δεν υπάρχει κάποια παθολογία.

7.4.2 Η εκπαίδευση του Attention UNET

Το attention UNET αποτελεί το δεύτερο μοντέλο που υλοποιήσαμε και είναι ένα αρκετά μεγάλο δίκτυο οπότε πριν το εκπαιδεύσουμε ήμασταν βέβαιοι ότι θα ξεπεράσει στα μετρικά, το UNET.

Η διαφοροποίηση έγκειται στο attention gate που τοποθετείται στο μέρος της διαστολής. Σε κάθε skip connection αντιστοιχεί ένα attention gate και ακολουθεί μια διαδικασία μεταξύ των δύο εισόδων που δέχεται. Η πρώτη είσοδος είναι η εικόνα που έρχεται από το μέρος της συστολής και η δεύτερη έρχεται από το αμέσως προηγούμενο επίπεδο στο μέρος της διαστολής όπου η αναπαράσταση των χαρακτηριστικών είναι αρκετά καλύτερη λόγω της προέλευσης της από βαθιά επίπεδα του δικτύου. Η συνάθροιση των δύο εισόδων αποσκοπεί την ανάδειξη των βαρών που είναι ευθυγραμμισμένα και άρα θα γίνουν μεγαλύτερα ενώ το αντίθετο θα συμβεί για τα μη-ευθυγραμμισμένα. Έτσι αναδεικνύουμε την χρήσιμη πληροφορία και υποβαθμίζουμε την ανεπιθύμητη (βλέπε Εικόνα 50).

Η αρχιτεκτονική αποτελείται από το μέρος της συστολής, το μέρος της διαστολής και το bottleneck μέρος όπως ακριβώς στο UNET. Αποτελείται από 4 blocks από διαδοχικά convolution με kernel μεγέθους 3x3, batch normalization και max pooling με kernel 2x2. Σε κάθε μπλοκ τα φίλτρα από τα οποία προκύπτουν τα feature maps διπλασιάζονται, το αρχικό μέγεθος είναι 64x64 και φτάνει μέχρι 512x512. Τα πρώτα trainings έγινα με kernel 16x16 ως αρχικά feature maps αλλά στην πορεία συμπεράναμε ότι τα 64x64 βελτιώνουν την διαδικασία της μάθησης της κατανομής του dataset από το νευρωνικό δίκτυο. Η συνάρτηση ενεργοποίησης είναι η γνωστή ReLU ενώ το max pooling έχει kernel 2x2 και χρησιμοποιείται μόνο στο μέρος της συστολής. Ορίζουμε την παράμετρο του dropout ως 0.1 σε κάθε μπλοκ καθώς η αύξηση του δεν μας οδήγησε σε κάποια σημαντική αύξηση των μετρικών.

Όσο αφορά την αρχικοποίηση των βαρών του νευρωνικού δικτύου ακολουθούμε την μέθοδο Xavier initialization, δηλαδή στάθμιση γύρω από το 0 και διακύμανση γύρω από μια αυστηρή τιμή.

Η συνάρτηση απώλειας είναι αυτή της διασταυρούμενης εντροπίας και αποτελεί σε κάθε μοντέλο μας επιλογή για το task της τμηματοποίησης. Ορίζουμε μετρικά όπως και στο UNET, τα accuracy, precision και τους συντελεστές Dice και IoU. Τα μετρικά είναι ίδια καθώς θέλουμε να συγκρίνουμε τα μοντέλα γύρω από ένα κοινό επίπεδο αναφοράς.

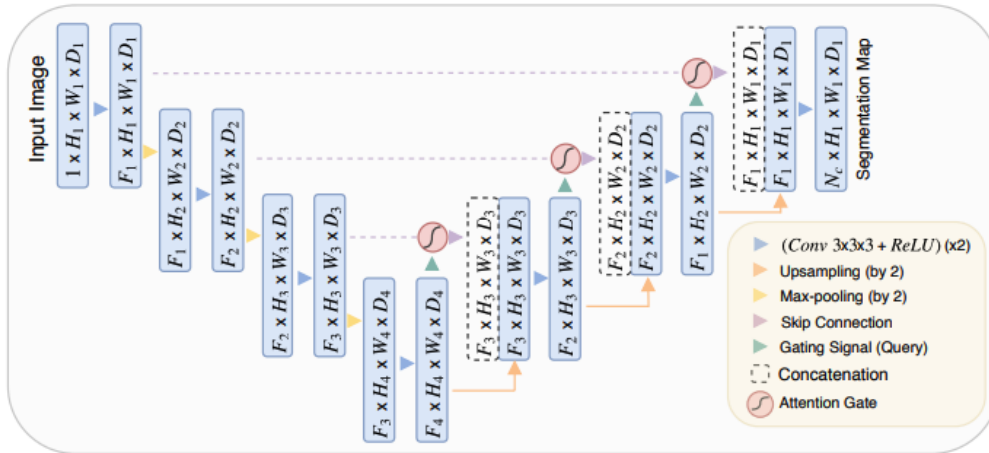
Στην εκπαίδευση χρησιμοποιούμε Adam optimizer ώστε να ενημερώνει τα βάρη κατά την οπισθοδρόμηση δίχως την χρήση weight decay policy. Επιλέγουμε κάποιες τεχνικές data augmentation, όπως η περιστροφή με μικρό εύρος γωνίας, η μετατόπιση στον δυσδιάστατο επίπεδο, και τα zoom και shearing (θόλωμα λόγω περιστροφής) σε μικρό βαθμό. Παρόλα αυτά δεν είδαμε να έχουν σημαντική συνεισφορά στην απόδοση του μοντέλου.

Εφαρμόσαμε τεχνικές early stopping και ReduceLROnPlateau οι οποίες στοχεύουν στον τερματισμό της εκπαίδευσης. Όταν το loss δεν μειώνεται ενεργοποιείται η παράμετρος patience η οποία υποδηλώνει τον αριθμό των εποχών που το loss μένει σταθερό και η μείωση του ρυθμού εκμάθησης ώστε να η βαθμωτή σύγκλιση να προχωρήσει καθώς φαίνεται να είναι εγκλωβισμένη σε κάποιο τοπικό ελάχιστο στην προσπάθεια εύρεσης του ολικού ελάχιστου. Η παράμετρος patience που ρυθμίζει την συχνότητα μείωσης του ρυθμού εκμάθησης.

Στην περίπτωση μας ορίζουμε patience 10 για το early stopping και patience 5 για το ReduceLROnPlateau (βλέπε Πίνακα 7.8). Η λογική προκύπτει από το γεγονός ότι χρησιμοποιούσαμε τις τιμές αντίστροφα το μοντέλο μας θα σταματούσε την εκπαίδευση σχετικά γρήγορα και θα δεν θα είχε κάποια σημαντική απόδοση (underfitting). Όσο αφορά την υπερπαράμετρο ReduceLROnPlateau ορίζουμε εμπειρικά ένα ελάχιστο ρυθμό εκμάθησης ώστε να μην παδιγεύσουμε την βαθμωτή σύγκλιση σε κάποιο τοπικό ελάχιστο. Έπειτα ορίζουμε τον αριθμό των εποχών στις 100 και τα βήματα της εποχής ως η διαίρεση του αριθμού των εικόνων με το batch size.

Συμπερασματικά, η εκπαίδευση διήρκεσε 50 εποχές και κλήθηκε η μέθοδος early stopping καθώς όπως θα δούμε και παρακάτω το loss δεν μειώθηκε περαιτέρω (βλέπε Εικόνα 51). Η τεχνική ReduceLROnPlateau ρύθμισε το ρυθμό εκμάθησης μέχρι την ελάχιστη τιμή.

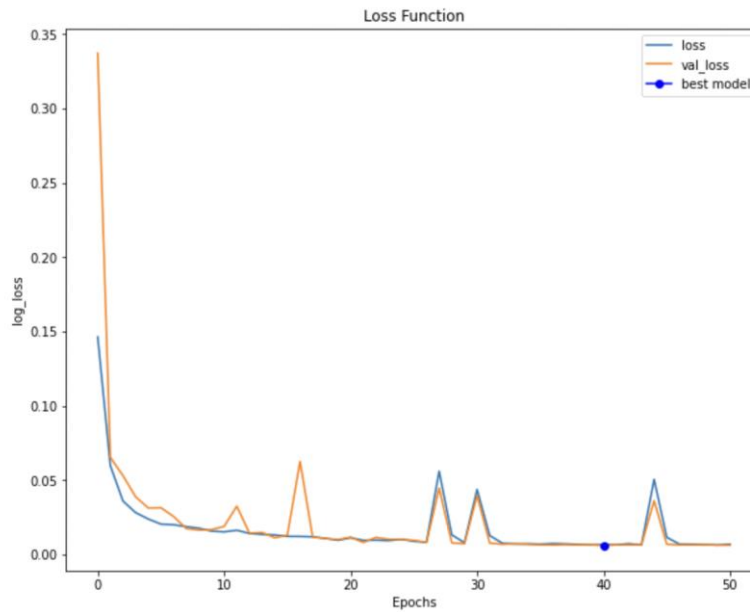
Ακολουθούν τα διαγράμματα για τα μετρικά που ορίσαμε τόσο για training όσο και για validation και κάποιες προβλέψεις (βλέπε Εικόνες 52-53). Όλες οι τιμές που αναφέρονται παραπάνω είναι αποτελέσματα επαναλαμβανομένων περαμάτων.



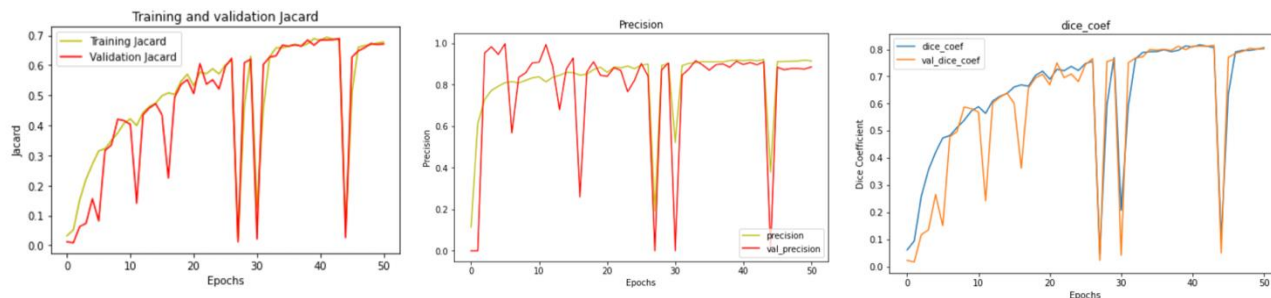
Εικόνα 50. Σχηματική Αναπαράσταση Attention U-NET [29]

Πίνακας 7.8 Λεπτομερής παρουσίαση των υπερπαραμέτρων του πειράματος

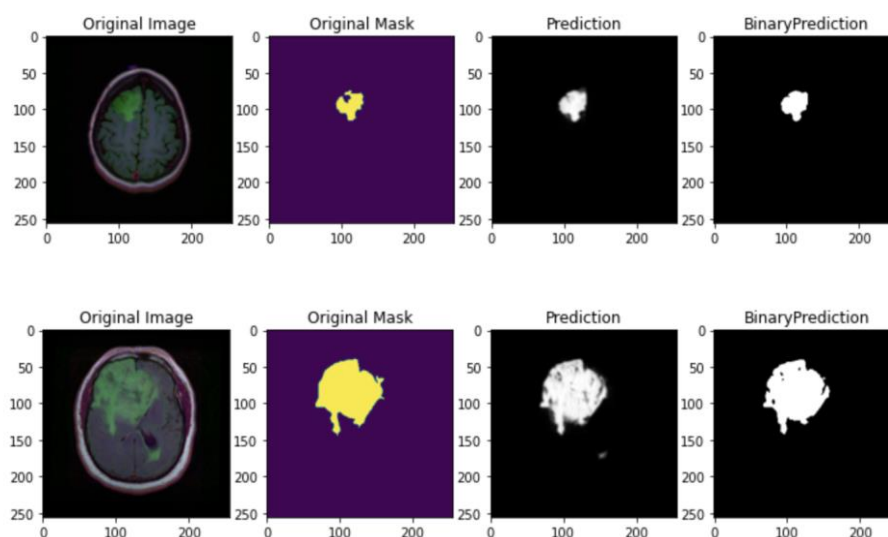
Batch Size	Optimizer	Loss Function	Learning rate	Epochs	Early Stopping	Reduce LROnPlateau
32	Adam	Binary Cross Entropy	$1e^{-5}$	35	Patience: 10	Patience: 5



Εικόνα 51. Γράφημα της συνάρτησης απώλειας κατά το training



Εικόνα 52. Διαγράμματα των μετρικών Jaccard, Precision και Dice



Εικόνα 53. Προβλέψεις με τιμή συντελεστή Dice μεγαλύτερη του 0.8

7.4.3 Η εκπαίδευση του ResUNET

Τρίτο υλοποιήσαμε ένα παρόμοιο δίκτυο με το UNET με την διαφορά ότι χρησιμοποιεί residual units, μια πολύ γνωστή τεχνική που χρησιμοποιούν deep neural networks όπως το ResNet για να αποφύγουν overfitting και να διατηρήσουν την ροή πληροφορίας στα βαθύτερα επίπεδα.

Σε αντίθεση με τα παραδοσιακά νευρωνικά δίκτυα κάθε επίπεδο προωθεί την έξοδο του στον επόμενο επίπεδο. Όμως, σε ένα δίκτυο με residual blocks ταυτόχρονα με την προώθηση της εισόδου κατά μήκος του επιπέδου, η είσοδος παραλείπει το επίπεδα και στην έξοδο αθροίζεται με την είσοδο. Το skip connection συνεισφέρει στην διατήρηση της πληροφορίας κατά μήκος του νευρωνικού δικτύου διασφαλίζοντας την απουσία φαινομένων όπως vanishing/exploding gradients (βλέπε Εικόνα 54).

Η αρχιτεκτονική αποτελείται από το μέρος της συστολής, το μέρος της διαστολής και το bottleneck μέρος όπως ακριβώς στο UNET. Αποτελείται από 4 blocks από διαδοχικά

convolution με kernel μεγέθους 3x3, batch normalization και max pooling με kernel 2x2. Σε κάθε μπλοκ τα φίλτρα από τα οποία προκύπτουν τα feature maps διπλασιάζονται, το αρχικό μέγεθος είναι 16x16 και φτάνει μέχρι 256x256. Τα πρώτα trainings έγιναν με kernel 64x64 ως αρχικά feature maps αλλά στην πορεία συμπεράναμε ότι τα kernels 16x16 συνεισφέρουν σημαντικότερα στην εκμάθηση της κατανομής του dataset. Η συνάρτηση ενεργοποίησης είναι η γνωστή ReLU ενώ το max pooling έχει kernel 2x2 και χρησιμοποιείται μόνο στο μέρος της συστολής. Δεν χρησιμοποιούμε το dropout καθώς η χρήση του στα πρώτα πειράματα δεν μας οδήγησε σε κάποια σημαντική αύξηση των μετρικών. Όσο αφορά την αρχικοποίηση των βάρων του νευρωνικού δικτύου ακολουθούμε την μέθοδο Xavier initialization.

Η συνάρτηση απώλειας είναι αυτή της διασταυρούμενης εντροπίας και αποτελεί σε κάθε μοντέλο μας επιλογή για το task της τμηματοποίησης. Ορίζουμε μετρικά όπως και στο UNET, τα accuracy, precision και τους συντελεστές Dice και IoU. Τα μετρικά είναι ίδια με τα προηγούμενα μοντέλα που αναπτύξαμε ώστε να είμαστε σε θέση να κάνουμε συγκρίσεις σχετικά με την απόδοσή τους.

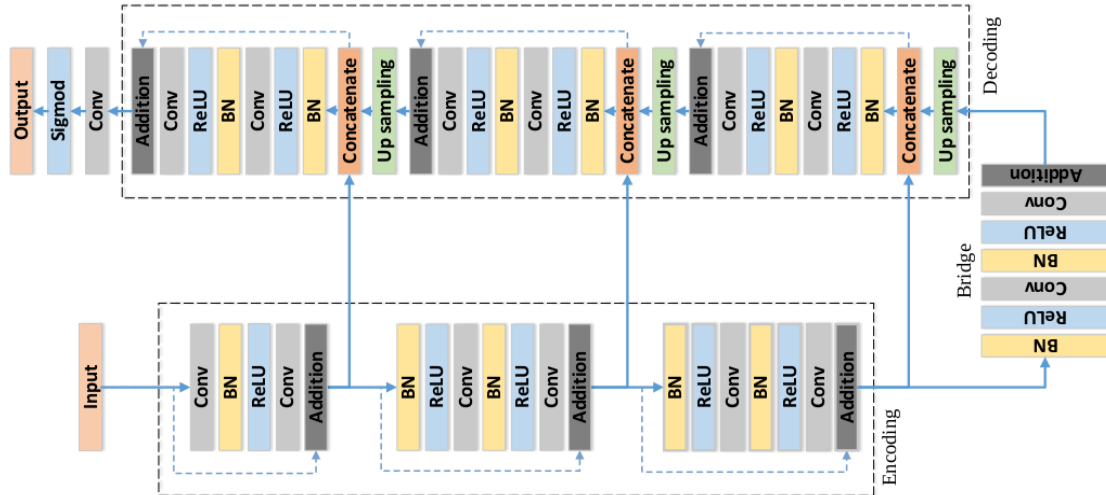
Στην εκπαίδευση χρησιμοποιούμε Adaptive Moment Estimation Optimizer (Adam) ώστε να ενημερώνει τα βάρη κατά την οπισθοδρόμηση δίχως την χρήση weight decay policy. Επιλέγουμε παρόμοιες τεχνικές data augmentation με τα προηγούμενα μοντέλα, τη περιστροφή με μικρό εύρος γωνίας, τη μετατόπιση στον δυσδιάστατο επίπεδο, zoom και shearing (θόλωμα λόγω περιστροφής) σε μικρό βαθμό. Γενική διαπίστωση αποτελεί η αμελητέα συνεισφορά των τεχνικών data augmentation στην απόδοση των μοντέλων μας.

Η τεχνική early stopping και ReduceLRonPlateau οι οποίες στοχεύουν στον τερματισμό της εκπαίδευσης εφαρμόστηκαν. Η παράμετρος patience που ρυθμίζει την συχνότητα μείωσης του ρυθμού εκμάθησης ορίστηκε σε 10 εποχές για το early stopping και 5 εποχές για το ReduceLRonPlateau. Η λογική της στρατηγικής μας στοχεύει στην αποφυγή του underfitting (βλέπε Πίνακα 7.9).

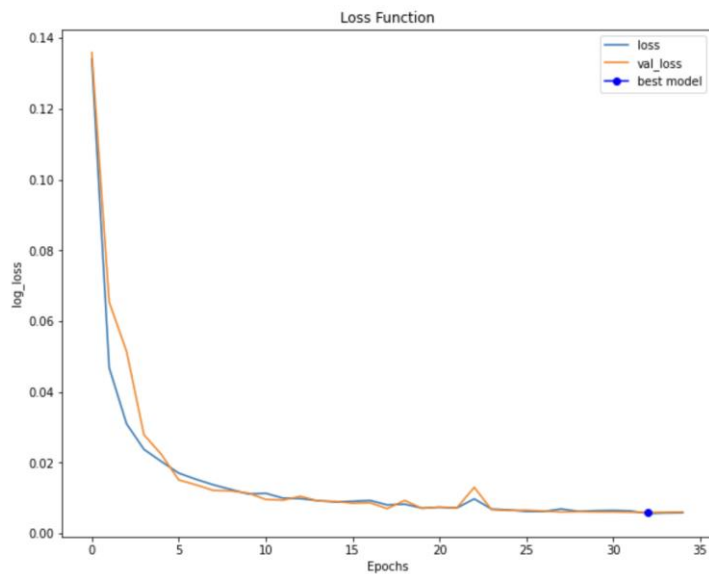
Συμπερασματικά, η εκπαίδευση διήρκεσε 50 εποχές και κλήθηκε η μέθοδος early stopping καθώς όπως θα δούμε και παρακάτω το loss δεν μειώθηκε περαιτέρω. Η τεχνική ReduceLRonPlateau ρύθμισε το ρυθμό εκμάθησης μέχρι την ελάχιστη τιμή. Ακολουθούν τα διαγράμματα για τα μετρικά που ορίσαμε τόσο για training όσο και για validation (βλέπε Εικόνες 55-56) και κάποιες εικόνες με προβλέψεις (βλέπε Εικόνα 57). Όλες οι τιμές που αναφέρονται παραπάνω είναι αποτελέσματα επαναλαμβανόμενων περαμάτων.

Πίνακας 7.9 Λεπτομερής παρουσίαση των υπερπαραμέτρων που χρησιμοποιήθηκαν στο πείραμα

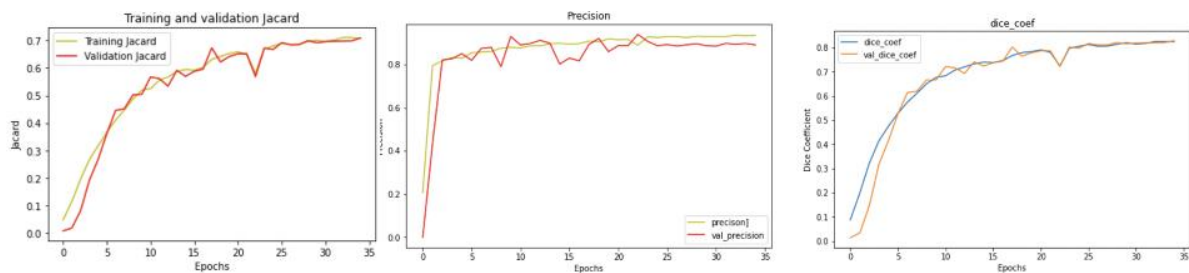
Batch Size	Optimizer	Loss Function	Learning rate	Epochs	Early Stopping	Reduce LROnPlateau
32	Adam	Binary Cross Entropy	$1e^{-5}$	50	Patience : 10	Patience : 5



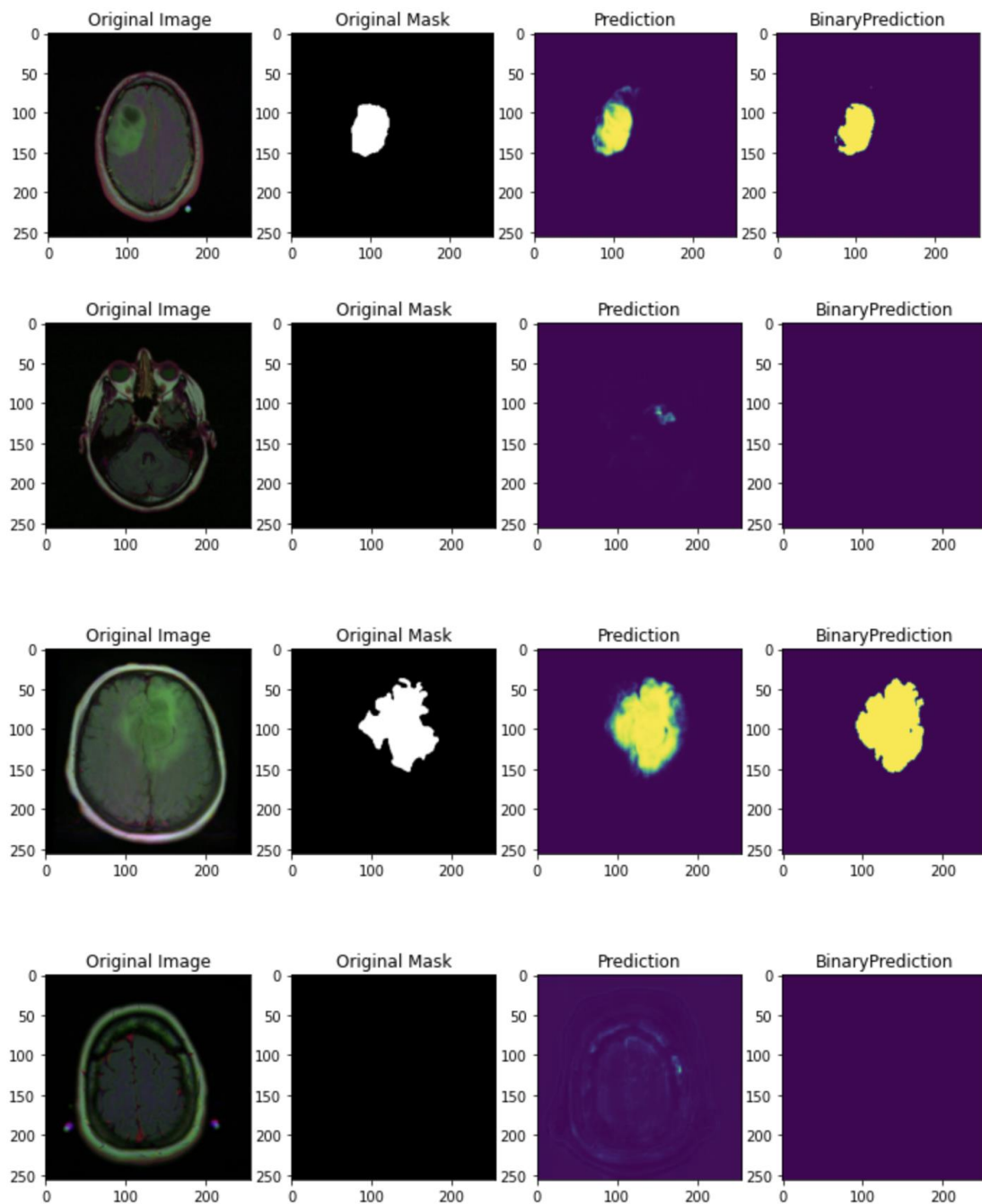
Εικόνα 54. Αναπαράσταση της αρχιτεκτονικής του ResUnet. Με γαλάζια διακεκομμένα τα res blocks. [32], [33], [34]



Εικόνα 55. Γράφημα της συνάρτησης απώλειας κατά το training



Εικόνα 56. Διαγράμματα Jaccard , Precision, Dice



Εικόνα 57. Προβλέψεις στο validation set για το Attention UNET

8. Αποτελέσματα

Θα μπορούσαμε να κατατάξουμε τις προβλέψεις μας σε τρεις μεγάλες κατηγορίες: α) τις προβλέψεις όπου το μεγαλύτερο ποσοστό των pixels ταξινομούνται σωστά, β) αυτές που υπάρχει ένα ποσοστό γύρω στο 10-15% σε FP/FN και τις προβλέψεις που γίνεται εσφαλμένη πρόβλεψη αναγνωρίζοντας όγκο, ο οποίος δεν υπάρχει (~1% των προβλέψεων).

Στους πίνακες 8.1-8.3 παρουσιάζονται τα αποτελέσματα για τα μετρικά Jaccard, συντελεστής Dice, Accuracy, Sensitivity, Specificity και Precision. Στην πρώτη γραμμή είναι η μέση τιμή και στην δεύτερη γραμμή η τυπική απόκλιση. Σημαντική παρατήρηση είναι η τάση αύξησης την μέσης τιμής με την ταυτόχρονη μείωση της τυπικής απόκλισης στο Ensemble Model.

Πίνακας 8.1 Αποτελέσματα UNET

	Jaccard	Dice	Accuracy	Precision	Specificity	Sensitivity
Mean	0.68	0.78	0.99	0.79	1.0	0.79
STD	0.23	0.2	0.01	0.25	0.01	0.24

Πίνακας 8.2 Αποτελέσματα Attention UNET

	Jaccard	Dice	Accuracy	Precision	Specificity	Sensitivity
Mean	0.73	0.83	0.99	0.86	1.0	0.84
STD	0.19	0.16	0.01	0.16	0.01	0.18

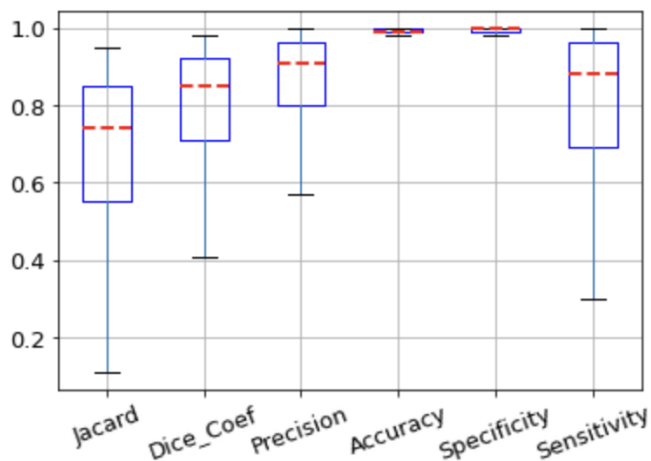
Πίνακας 8.3 Αποτελέσματα ResUnet

	Jaccard	Dice	Accuracy	Precision	Specificity	Sensitivity
Mean	0.72	0.82	0.99	0.85	1.0	0.84
STD	0.18	0.15	0.01	0.15	0.01	0.18

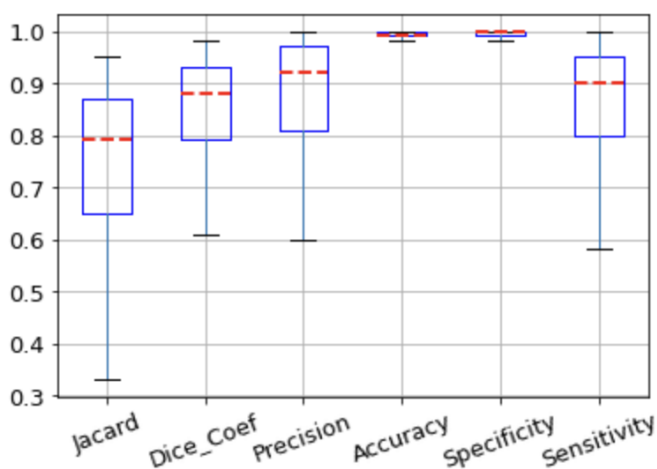
Πίνακας 8.4 Αποτελέσματα Ensemble Model

	Jaccard	Dice	Accuracy	Precision	Specificity	Sensitivity
Mean	0.74	0.84	0.99	0.88	1.0	0.85
STD	0.18	0.15	0.01	0.13	0.01	0.17

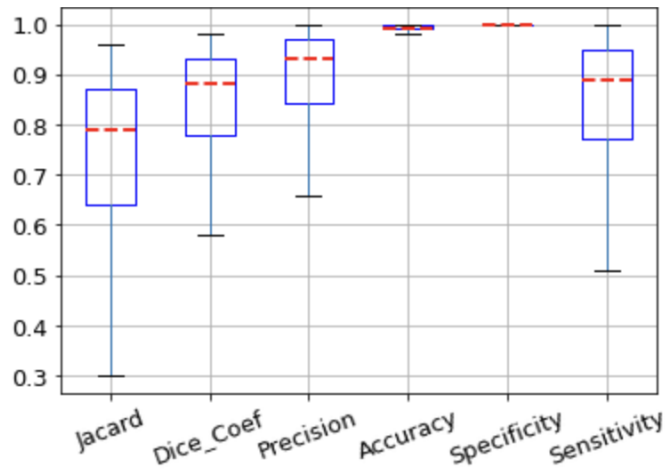
Στις εικόνες από 58-61 παρουσιάζονται τα θηκογράμματα των μετρικών διαφόρων δεικτών αξιολόγησης για τα τρία μοντέλα που αναπτύξαμε στο validation set. Η κόκκινη διακεκομμένη γραμμή δηλώνει την ενδιάμεση τιμή των αποτελεσμάτων ενώ το μπλε κουτί δηλώνει το εύρος των τιμών μεταξύ δύο τυπικών αποκλίσεων.



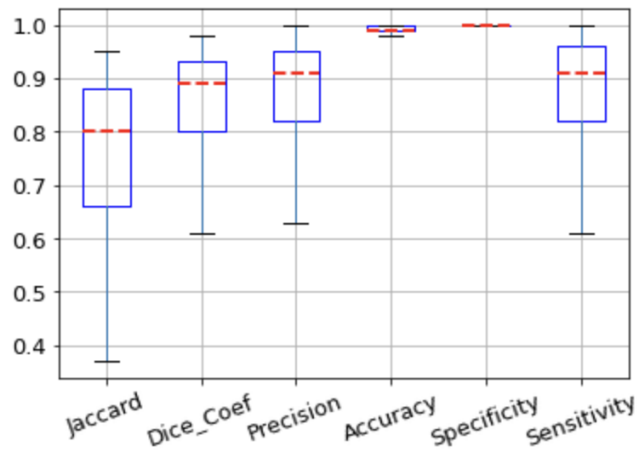
Εικόνα 58. Boxplot UNET



Εικόνα 59. Boxplot Attention UNET



Εικόνα 60. Boxplot ResUNET



Εικόνα 61. Boxplot Ensemble Model

Παρακάτω ακολουθεί η Εικόνα 62 (A-P) από το validation set όπου παρουσιάζονται οι προβλέψεις τόσο από τα μεμονωμένα μοντέλα όσο και από το πλειοψηφικό. Στον Πίνακα 8.5: παρατίθενται τα scores του Dice για την καλύτερη κατανόηση των προβλέψεων. Παρατηρείται μεγάλη ετερογένεια στο σχήμα, το μέγεθος, την υφή και την συμμετρία των όγκων. Παρόλα αυτά τα μοντέλα είναι ικανά να προσεγγίσουν την κακοήθεια σε ικανοποιητικά επίπεδα με ευαισθησία στις παρυφές όπως στις περιπτώσεις των b, d, i και e.

Τμηματοποίηση όγκων εγκεφάλου σε εικόνες μαγνητικής τομογραφίας μέσω αλγορίθμων βαθιάς μάθησης

	Original Image	Ground Truth Mask	UNET	AttentionUNET	ResUNET	Ensemble Model
A						
B						
C						
D						
E						
F						
G						
H						

	Original Image	Ground Truth Mask	UNET	AttentionUNET	ResUNET	Ensemble Model
I						
J						
K						
L						
M						
N						
O						
P						

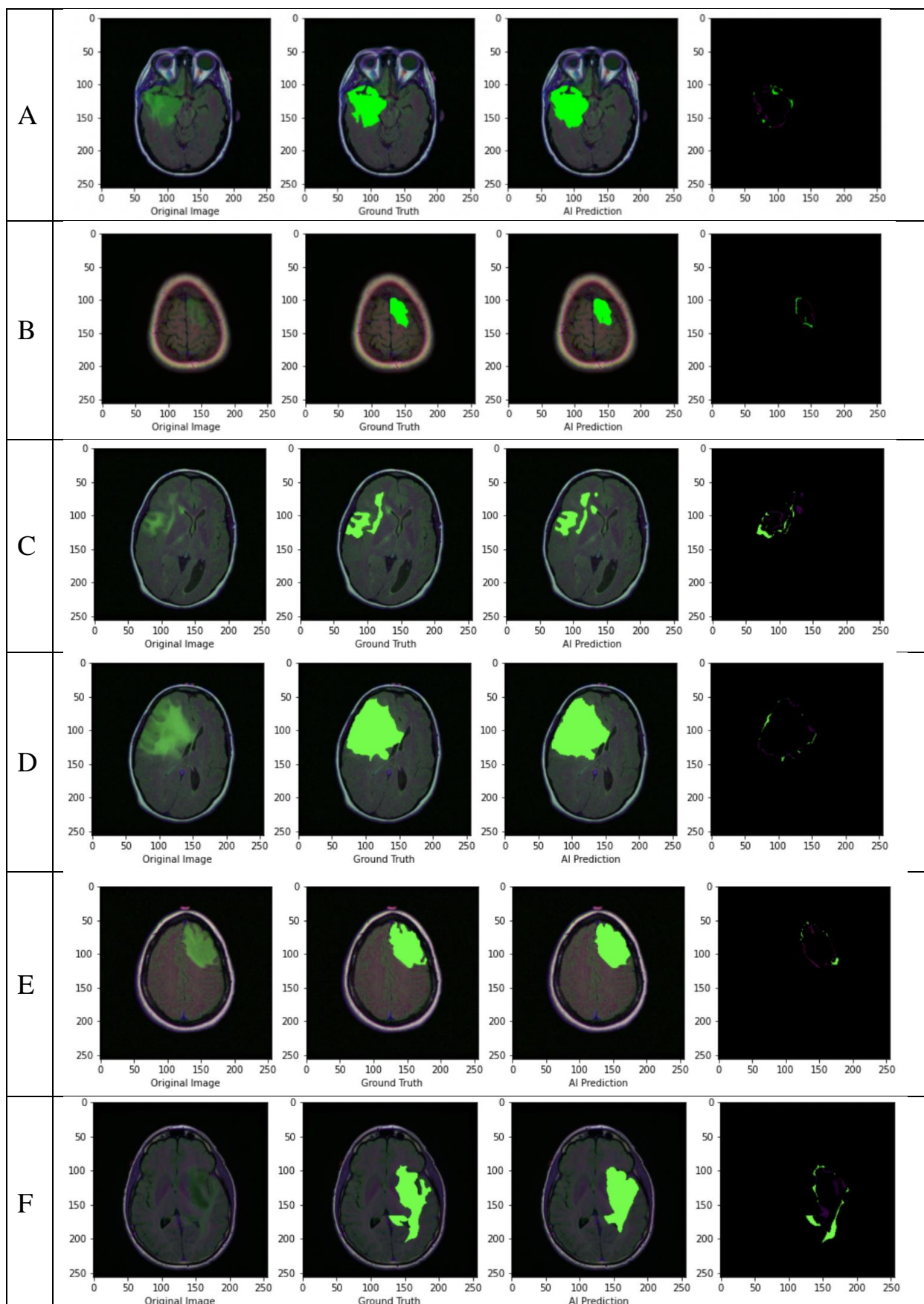
Εικόνα 62. (A-P) Παραδείγματα σε 16 εικόνες με τα αντίστοιχα αποτελέσματα τμηματοποίησης για κάθε ένα από τα δοκιμαζόμενα δίκτυα

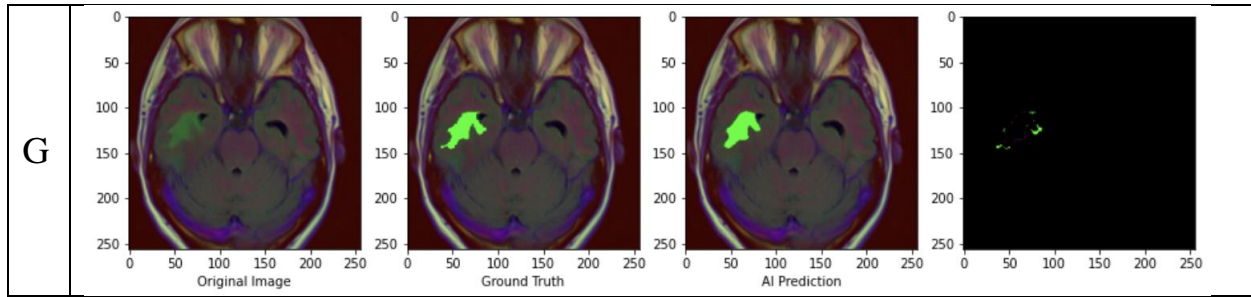
Πίνακας 8.5: Τα αποτελέσματα του συντελεστή Dice για κάθε μια από τις εικόνες του παραδείγματος (Εικόνα 62 A-P)

	Dice		
	UNET	Attention UNET	Res UNET
A	0.9	0.92	0.92
B	0.72	0.77	0.72
C	0.96	0.97	0.96
D	0.75	0.82	0.82
E	0.9	0.89	0.88
F	0.82	0.9	0.6
G	0.94	0.94	0.95
H	0.85	0.84	0.84
I	0.88	0.89	0.87
J	0.9	0.91	0.95
K	0.87	0.9	0.86
L	0.89	0.88	0.71
M	0.94	0.96	0.95
N	0.97	0.96	0.97
O	0.68	0.78	0.82
P	0	0	0

Ακολουθεί η Εικόνα 63 (A-G) με τις προβλέψεις του ensemble model για να διαπιστωθεί η προσέγγιση της βλάβης. Αριστερά η κανονική εικόνα, δεξιότερα η μάσκα, δεξιότερα η πρόβλεψη του μοντέλου μας και τέλος δεξιά η διαφορά της πρόβλεψης μας από την αληθινή. Στις παρακάτω εικόνες θα φανεί η προβλεπτική αξία του μοντέλου μας καθώς και οι δυνατότητες του ensembling σε μικρά datasets.

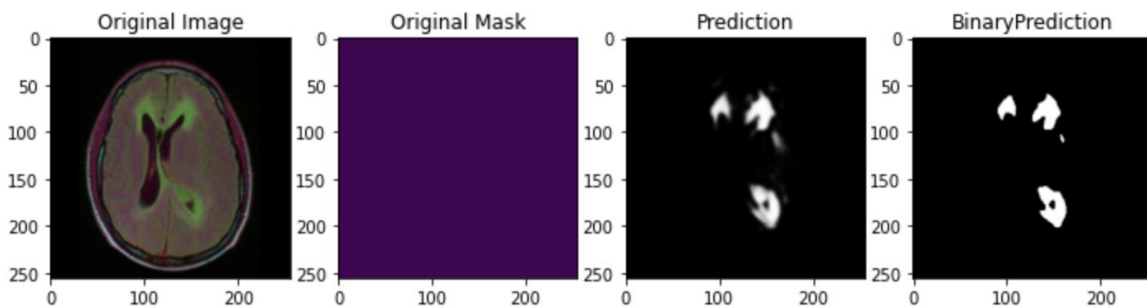
Τμηματοποίηση όγκων εγκεφάλου σε εικόνες μαγνητικής τομογραφίας μέσω αλγορίθμων βαθιάς μάθησης





Εικόνα 63. (A-G) Αριστερά η τομή, στη επόμενη έχουμε υπερθέσει την μάσκα για σαφή προβολή, δεξιά απεικονίζεται η πρόβλεψη του EDL μοντέλου μας και τέλος η διαφορά μεταξύ προβλεπόμενης και αληθινής ώστε να κατανοηθεί η απόδοση του μοντέλου (a-g)

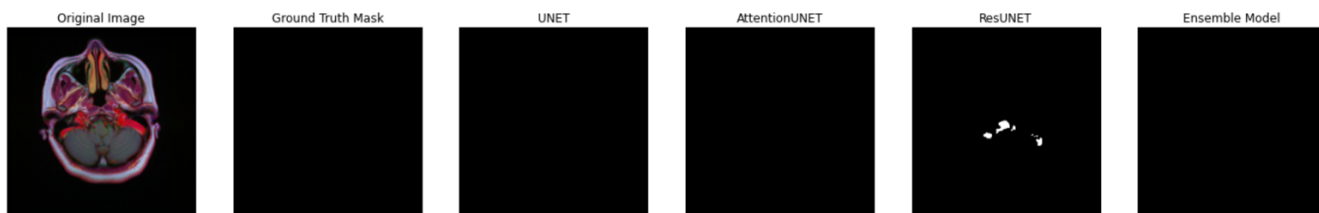
Στην Εικόνα 64 παρακάτω παρατηρούμε ένα παράδειγμα όπου η εικόνα φαίνεται να περιέχει πολλαπλές βλάβες, που όμως είναι αστοχία της προεπεξεργασίας που έχει υποστεί το dataset καθώς η μάσκα δεν περιέχει κάποια παθολογία. Αυτό οδηγεί το μοντέλο σε False Positives. Τέτοια παραδείγματα μας υποβαθμίζουν την απόδοση του μοντέλου και προκαλούν την απόκλιση που φαίνεται στο θηκόγραμμα και στον πίνακα της μέσης τιμής του Dice. (βλέπε Πίνακες 8.1-8.4)



Εικόνα 64. Παράδειγμα κακής πρόβλεψης λόγω αστοχίας του dataset

Ενδεικτικά αποτελέσματα για τα μετρικά στο validation set. Παρατηρούμε ότι το specificity και το accuracy είναι μέγιστα λόγω της φύσης του unbalanced dataset. Δηλαδή τα TNs δυσανάλογα με τα TPs. Φυσικά αναφερόμαστε στην επικράτηση των True Negatives.

Τέλος, ένα από τα πολλά παραδείγματα που αποδεικνύει την χρησιμότητα του ensembling στην βαθιά μάθηση απεικονίζεται παρακάτω (Εικόνα 65) όπου η τεχνική majority vote είναι πολύ πιο ουσιαστική από την οπτική που προσφέρει κάθε μοντέλο ξεχωριστά. Στην συγκεκριμένη περίπτωση το ένα μοντέλο έχει αποτύχει να ανιχνεύσει την κενή μάσκα αλλά δεν επηρεάζει το τελικό αποτέλεσμα.



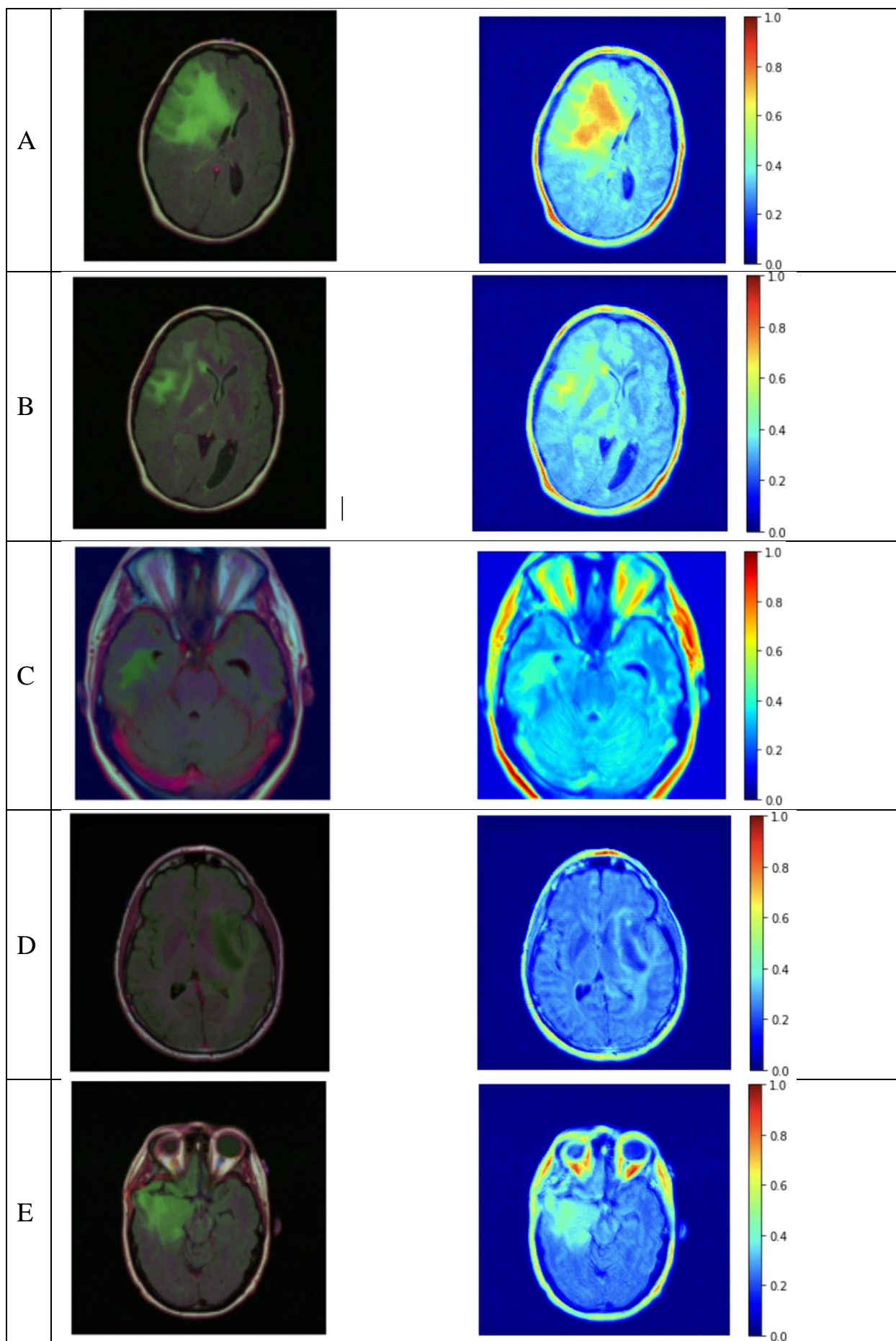
Εικόνα 65. Η επίδραση του ensembling : Αρχική εικόνα όπου δεν είχε δοθεί μάσκα παθολογίας (normal) και τα αντίστοιχα αποτελέσματα των δικτύων.

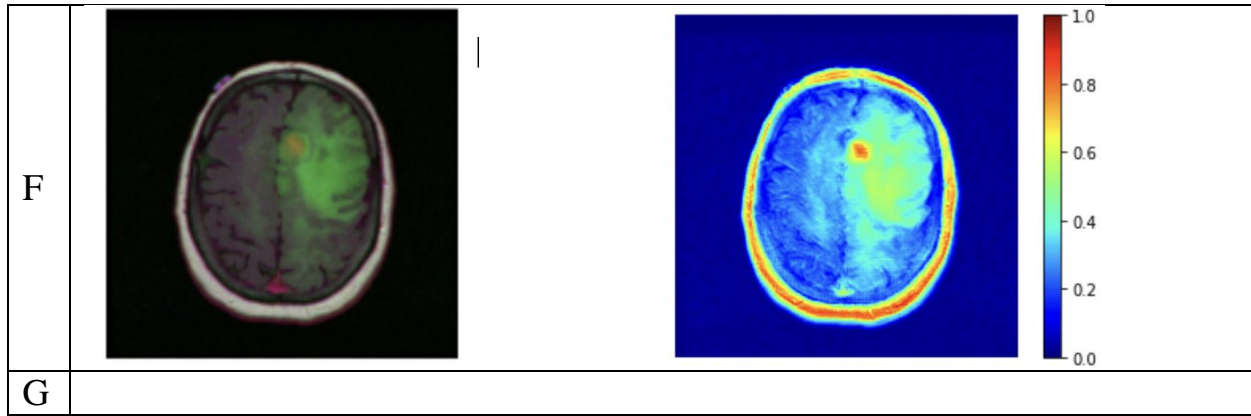
8.1 Κατανόηση των αποφάσεων στο Deep Learning

Για να γίνουν οι αποφάσεις του μοντέλου κατανοητές, χρησιμοποιήσαμε ένα visualization tool, το Grad-CAM (Gradient-weighted Class Activation Map) για να απεικονίσουμε τον τρόπο με τον οποίο γίνονται οι προβλέψεις όπως προτάθηκε από τους Selvaraju et al. [75] Τροφοδοτούμε το προεκπαιδευμένο δίκτυο μας με τις εικόνες του dataset για το validation. Χρησιμοποιώντας τα gradients του τελευταίου συνελκτικού επίπεδο κάθε δικτύου μας, δημιουργείται το class activation map που επισημαίνει τις περιοχές με σχετική πληροφορία με τις προβλέψεις των μοντέλων μας μέσω της συνάρτησης softmax (εξίσωση 1.3). Έπειτα εφαρμόζεται στο heatmap ανερχόμενη δειγματοληψία μέσω διγραμμικής (bilinear) παρεμβολής ώστε να συμπίπτει με το μέγεθος της αρχικής εικόνας.

Υπάρχουν και άλλες δημοφιλείς τεχνικές όπως local approximation (LIME), αλλά στην παρούσα διπλωματική εστίασαμε μόνο σε visualization τεχνικές.

Στην Εικόνα 66 παρακάτω παρατίθενται οι χάρτες πυκνότητας πιθανότητας (heatmaps) οι οποίοι μας παρουσιάζουν τα scores του τελευταίου επιπέδου στο νευρωνικό δίκτυο ώστε να κατανοήσουμε τις προβλέψεις των νευρωνικών δικτύων μας. [75], [76]





Εικόνα 66. (A-F) Αριστερά η εικόνα και δεξιά ο χάρτης πυκνότητας πιθανότητας (Grad-CAM)

9. Συμπεράσματα-Συζήτηση

Το πλειοψηφικό μοντέλο μας εντοπίζει με μεγάλη ακρίβεια τους όγκους παρόλο την ετερογένεια που παρουσιάζουν. Έχει ικανοποιητική απόδοση τμηματοποίησης των όγκων, αντίστοιχα με την βιβλιογραφία και η ικανότητα γενίκευσης σύμφωνα με τα αποτελέσματα σε δεδομένα ελέγχου (validation set), λαμβάνοντας υπόψη το μικρό μέγεθος και την ανομοιογένεια του dataset.

Μπορούμε να παρατηρήσουμε ότι οι προβλέψεις έχουν ένα bias λόγω της χρήσης του max pooling επιπέδου το οποίο δίνει μία ομαλή τμηματοποίηση των όγκων με μικρή ευαισθησία στις δομές που κυριαρχούν οι απότομες καμπύλες. Συνεπώς οι παραγόμενες μάσκες είναι πιο λείες σε σχέση με αυτές που σχεδίασαν οι ιατροί χειροκίνητα. Επίσης, η προεπεξεργασία που έγινε στις εικόνες του dataset πριν δημοσιευτεί στο Kaggle οδηγεί τις προβλέψεις μας σε κάποια λάθη λόγω των RGB τόνων σε σημεία που δεν έχουν σχέση με την διαγνωστική πληροφορία.

Η παρούσα διπλωματική εξετάζει μία εφαρμογή που στηρίζεται στην χρήση αλγορίθμων βαθιάς μάθησης για την ανίχνευση όγκων από εικόνες μαγνητικής τομογραφίας εγκεφάλου. Παρόλο που η απόδοση των μοντέλων μηχανικής μάθησης είναι σε υψηλό επίπεδο όπως αναφέρεται στην βιβλιογραφία σε πληθώρα εφαρμογών της ιατρικής απεικόνισης, η ανίχνευση παθολογιών δεν έχει δοκιμαστεί ακόμα σε κλινικό περιβάλλον που είναι και ο σκοπός όλων των προσπαθειών. Επίσης, περιορισμοί όπως αυτοί των μικρών datasets και άλλων κλινικών παραμέτρων που υφίστανται στο κλινικό περιβάλλον όπως εμφυτεύσιμες συσκευές (ενδοκρανιακές), artifacts στην απεικόνιση, ακόμα και το ιστορικό κάποιας χειρουργικής επέμβασης δεν παρέχονται. Συνεπώς η απόδοση των μοντέλων μας στην βαθιά μάθηση είναι σε ένα εξιδανικευμένο περιβάλλον, μακριά από την κλινική ρουτίνα.

Άρα λύσεις όπως η τυποποίηση των σαρωτών των διάφορων εταιριών σε ένα κοινό πρότυπο (π.χ CT, MRI), τα μεγάλα datasets και η συνεχής εκπαίδευση των μοντέλων μέσα στο κλινικό περιβάλλον(MLOps) θα ενισχύσει σημαντικά την ευρωστία των μοντέλων μας για όλες τις υπολογιστικά υποβοηθούμενες εφαρμογές στην ιατρική.

Βιβλιογραφία

- [1] *Simon Haykin (2010), Νευρωνικά Δίκτυα και Μηχανική Μάθηση, Εκδόσεις Παπασωτηρίου.*
- [2] *Gonzalez C. Rafael., Woods. E. Richard. (2018), Ψηφιακή Επεξεργασία Εικόνας, Εκδόσεις ΤΖΙΟΛΑ.*
- [3] *Deisenroth Marc Peter, Faisal A. Aldo, Ong C. S. (2020), Mathematics for Machine Learning, Cambridge University Press*
- [4] *IBM Course (2020), Gradient Descent, <https://www.ibm.com/cloud/learn/gradient-descent>*
- [5] *Long J., Shelhamer E., Darrell T. (2015), Fully Convolutional Networks for Semantic Segmentation, IEEE Conf. on Computer Vision and Pattern Recognition.*
- [6] *Rotha R. H., Odaa H., Zhou X., Shimizu N., Yanga Y., Hayashia Y., Odaa M., Fujiwarac M., Misawad K., Mori K. (2018), An application of cascaded 3D fully convolutional networks for medical image segmentation, Computerized medical imaging and graphics: the official journal of the Computerized Medical Imaging Society.*
- [7] *Parmar P.(2018), Detection and Segmentation through ConvNets, Towards Data Science. <https://towardsdatascience.com/detection-and-segmentation-through-convnets-47aa42de27ea>*
- [8] *Nickolas Papanikolaou, Joao Santinha (2018), An Introduction to Radiomics: Capturing Tumour Biology in Space and Time.*
- [9] *Patrick Henry Winston (2021), Lecture of Artificial Intelligence, Department of Electrical Engineering and Computer Science, Course Number :6.034, MIT OpenCourseWare,.*
- [10] *Daniel S. Kermany, Michael Goldbaum, Wenjia Cai M. Anthony Lewis, Huimin Xia, Kang Zhang (2018), Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning.*
- [11] *Seth Weidman (2019), Deep Learning from Scratch, O'REILLY Press*
- [12] *Tuan D. Pham (2021), Classification of COVID-19 chest X-rays with deep learning: new models or fine tuning, Health Inf Sci Syst. 2021.*
- [13] *Miroslav Kubat (2017), An Introduction to Machine Learning, Second Edition, Springer Press.*

[14] Olaf Ronneberger, Philipp Fischer, Thomas Brox (2015), *U-Net: Convolutional Networks for Biomedical Image Segmentation* <https://lmb.informatik.uni-freiburg.de/people/ronneber/u-net/>

[15] Heet Sankesara (2019), *UNet-Introducing Symmetry in Segmentation*
<https://towardsdatascience.com/u-net-b229b32b4a71>

[16] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio(2014), *Generative Adversarial Networks* , Arxiv, 1406.2661.

[17] Kawahara D., Nagata Y. (2021), *T1-weighted and T2-weighted MRI image synthesis with convolutional generative adversarial networks*, Department of Radiation Oncology, Institute of Biomedical and Health Sciences, Hiroshima University, Hiroshima, Japan
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8086713/>

[18] Platscher M., Zopes J., Federau C.(2021), *Image Translation for Medical Image Generation: Ischemic Stroke Lesion Segmentation*, Institute for Biomedical Engineering, University and ETH Zürich, Gloriastrasse 35, 8092 Zürich, Switzerland

[19] Jason Brownlee (2019), *Generative Adversarial Networks with Python Deep Learning Generative Models for Image Synthesis and Image Translation*, Edition: v1.5.

[20] Goodfellow I. (2016), *NIPS 2016 Tutorial: Generative Adversarial Networks*, Arxiv, 1701.00160.

[21] Goodfellow I., Bengio Y., Courville A. (2016), *DEEP LEARNING* , MIT Press Cambridge.
https://www.academia.edu/38223830/Adaptive_Computation_and_Machine_Learning_series_Deep_learning_The_MIT_Press_2016_pdf

[22] Hayflick S.J., Hartman M., Coryell J., Gitschier J., Rowley H. (2006), *Brain MRI in Neurodegeneration with Brain Iron Accumulation with and without PANK2 Mutations*, *AJNR Am J Neuroradiol* 27 (2006), 1230–1233.

[23] Naia Y.- H., Teo B. W., Tan N. L., O'Doherty S., Stephenson M. C., Thian Y.L., Chiong E., Reilhac A. (2021), *Comparison of metrics for the evaluation of medical segmentations using prostate MRI dataset*, *Computers in Biology and Medicine*, 134 (2021), 104497.

[24] Niranjana Kumar (2019), *Batch Normalization and Dropout in Neural Networks with Pytorch*, *Toward Data Science*

<https://towardsdatascience.com/batch-normalization-and-dropout-in-neural-networks-explained-with-pytorch-47d7a8459bcd>

[25] Niyas S., Pawan S. J., Anand Kumar M., Rajan J.,(2022) *Medical Image Segmentation with 3D Convolutional Neural Networks: A Survey*, Arxiv, 2108.08467.

[26] Eikon Minkoff, <https://s.mriquestions.com/who-invented-mri.html>

[27] Κωστόπουλος Σ. (2021), *Κατάτμηση Εικόνας και Μορφολογική επεξεργασία ιατρικών εικόνων με εφαρμογές, Εργαστήριο Επεξεργασίας Ιατρικού Σήματος & Εικόνας (ΕΙΣΕ), Τμήμα Μηχανικών Βιοϊατρικής, Πανεπιστήμιο Δυτικής Αττικής*

[28] Grand J. David, Woodfield A. Courtney, Mayo-Smith W. William(2012), *Practical Body MRI Protocols, Applications, and Image Interpretation*, Cambridge University Press.

[29] Oktay O., Schlemper J., Folgoc L. L., Lee M., Heinrich M., Misawa K., Mori K., McDonagh S., Hammerla N. Y., Kainz B., Glocker B., Rueckert D. (2018), *Attention U-Net: Learning Where to Look for the Pancreas*, Arxiv, 1804.03999.

[30] Bahdanau D., Cho K. H., Bengio Y. (2015), *Neural Machine Translation by Jointly Learning to Align and Translate*, Arxiv, 1409.0473.

[31] Graves A. (2020), *Lecture: Attention and Memory in Deep Learning*, DeepMind's deep learning videos 2020, University College of London.

[32] He K., Zhang X., Ren S., Sun J. (2016), *Identity Mappings in Deep Residual Networks*, Arxiv, 1603.05027.

[33] He K., Zhang X., Ren S., Sun J. (2015) *Deep Residual Learning for Image Recognition*, Arxiv, 1512.03385

[34] Zhang Z., Liu Q. (2017), *Road Extraction by Deep Residual U-Net*, *IEEE GEOSCIENCE AND REMOTE SENSING LETTERS*, Arxiv, 1711.10684.

[35] Sharif M. I., Li J. P., Amin J., Sharif A. (2021), *An improved framework for brain tumor analysis using MRI based on YOLOv2 and convolutional neural network*, SpringerLink, *Complex & Intelligent Systems*.

[36] Yang Q., Zhang H., Xia J., Zhang X.(2020), *Evaluation of magnetic resonance image segmentation in brain low-grade gliomas using support vector machine and convolutional neural network*, *Journal of Quantitative Imaging in Medicine and Surgery*.

[37] Imad Dabbura(2017), *Gradient Descent Algorithm and Its Variants*, *Towards Data Science*
<https://towardsdatascience.com/gradient-descent-algorithm-and-its-variants-10f652806a3>

[38] Pooley A.P. (2005), *AAPM/RSNA physics tutorial for residents: fundamental physics of MR imaging*. *Radiographics*. 2005 Jul-Aug;25(4):1087-99.

[39] Lipton L. M. (2008), *Totally Accessible MRI: A User's Guide to Principles, Technology, and Applications*, Springer 2008th Edition

[40] Καρατόπης Α. Κανδαράκης Ι. (2008), *Απεικόνιση Μαγνητικού Συντονισμού*, Πανεπιστημιακές Εκδόσεις "Αράκυνθος", Έκδοση 2008.

[41] *Article Conditions and Diseases (2022), Brain Anatomy and How the Brain Works*, Johns Hopkins Medicine.
<https://www.hopkinsmedicine.org/health/conditions-and-diseases/anatomy-of-the-brain>

[42] *Article Brain Basics (2022), Know Your Brain*, National Institute of Neurological Disorders and Stroke

[43] Gray H. (1993), *Gray's Anatomy*, Magpie Books Limited, Edition 1993.

[44] Εικόνα : <https://www.imaios.com/en/e-Courses/e-MRI/MRI-Sequences/inversion-recovery-stir-flair>

[45] Ekin Tiu (2019), *Metrics to Evaluate your Semantic Segmentation Model*, *Toward Data Science*.
<https://towardsdatascience.com/metrics-to-evaluate-your-semantic-segmentation-model-6bcb99639aa2>

[46] *Lecture 12, Convolutional Neural Networks*, Stanford University School of Engineering.
<https://www.youtube.com/watch?v=6wcs6szJWMY&list=PL3FW7Lu3i5JvHM8ljYj-zLfQRF3EO8sYv&index=12>

[47] Michelogiannakis G. (2016), *Lecture 23, Graphics Processing Units*, EECS, University of California at Berkeley.

[48] Εικόνα : <https://developer.nvidia.com/deep-learning-examples>

[49] Εικόνα: https://cdn-images-1.medium.com/max/800/1*vxxlTDnFOP5jQGud7uJrSg.png

[50] Zhou Z. H., Feng J.(2017), *Deep Forest: Towards An Alternative to Deep Neural Networks,*” *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence (IJCAI)*, pp. 3553–3559, 2017

[51] Buda M., Saha A., Mazurowskiab M. A. (2019), *Association of genomic subtypes of lower-grade gliomas with shape features automatically extracted by a deep learning algorithm*, *Computers in Biology and Medicine*, Elsevier.

[52] Chi-Feng Wang (2019), *The Vanishing Gradient Problem ,The problem, Its Cause , Its Significance and Its Solutions*
<https://towardsdatascience.com/the-vanishing-gradient-problem-69bf08b15484>

[53] Brat D.J., Verhaak R.G., Aldape K.D., Yung W.K., Salama S.R., Cooper L.A., Rheinbay E., Miller C.R., Vitucci M., Morozova O., Robertson A.G., Nounshmehr H. (2015) *Cancer Genome Atlas Research N, Comprehensive, Integrative Genomic Analysis of Diffuse Lower-Grade Gliomas*, *N Engl J Med*. 2015; 372:2481–2498.

[54] Niyasa S.,Pawana S. J., Anand Kumarb M., Rajana J. (2022), *Medical Image Segmentation using 3D Convolutional Neural Networks: A Review*, *ArXiv*, 2108.08467v2.

[55] Ghosh S., Chaki A., Santosh K. C. (2021), *Improved U-Net architecture with VGG-16 for brain tumor segmentation*, *Australasian College of Physical Scientists and Engineers in Medicine 2021*, 44(2021), 703–712

[56] Litjens G., Kooi T., Bejnordi B. E., Setio A. A., Ciompi F., Ghafoorian F., van Laak J. A.W.M., van Ginneken B., Sánchez C. I. (2017), *A survey on deep learning in medical image analysis*, *Diagnostic Image Analysis Group, Radboud University Medical Center, Nijmegen, Medical Imaging Analysis*, 42 (2017), 60-88.

[57] Kooi T., Litjens G., Ginneken B., Gubern-Mérida A., Sánchez C. I., Mann R., Heeten A., Karssemeijer N. (2017), *Large scale deep learning for computer aided detection of mammographic lesions*, *Diagnostic Image Analysis Group, Radboud University Medical Center, Nijmegen, Medical Image Analysis*, 35 (2017), 303-312.

[58] Charbonniera J. P., Rikxoort E. M., Setio A. A., Schaefer-Prokopc C. M., van Ginneken B., Ciompia F. (2017), *Improving airway segmentation in computed tomography using leak detection with convolutional*

networks, Diagnostic Image Analysis Group, Radboud University Medical Center, Nijmegen, Medical Imaging Analysis, 36 (2017), 52-60.

[59] *Ghafoorian M., Karssemeijer N., Heskes T., van Uder I. W. M., de Leeuw F. E., Marchiori E., van Ginneken B., Platel B.(2016) Non-uniform patch sampling with deep convolutional neural networks for white matter hyperintensity segmentation, 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), 2016, 1414-1417.*

[60] *van Grinsven M. J. J. P., van Ginneken B., Hoyng C. B., Theelen T., Sánchez C. I. (2016),Fast Convolutional Neural Network Training Using Selective Data Sampling: Application to Hemorrhage Detection, 2016 IEEE Transactions on Medical Imaging, 35, 5.*

[61] *Yu, L., Yang, X., Chen, H., Qin, J., Heng, P. A. (2017),Volumetric convnets with mixed residual connections for automated prostate segmentation from 3D MR images, 31st AAAI Conference on Artificial Intelligence, AAAI 2017 (pp. 66-72).*

[62] *Qi Dou, Hao Chen, Lequan Yu, Jing Qin, Pheng-Ann Heng (2016), Multilevel Contextual 3-D CNNs for False Positive Reduction in Pulmonary Nodule Detection, IEEE Trans Biomed Engineering, 64 (2017), 1558-1567.*

[63] *Tran P. V.(2017) A Fully Convolutional Neural Network for Cardiac Segmentationin Short-Axis MRI, Arxiv, 1604.00494.*

[64] *Michal Drozdal, Eugene Vorontsov, Gabriel Chartrand ,Samuel Kadoury, and Chris Pal, The Importance of Skip Connections in Biomedical Image Segmentation, Arxiv, 1608.04117.*

[65] *Xi Wei, Ming Gao, Ruiguo Yu, Zhiqiang Liu, Qing Gu, Xun Liu, Zhiming Zheng, Xiangqian Zheng, Jialin Zhu, Sheng Zhang(2020), Ensemble Deep Learning Model for Multicenter Classification of Thyroid Nodules on Ultrasound Images, Department of Diagnostic and Therapeutic Ultrasonography, Tianjin Medical University Cancer Institute and Hospital, Tianjin, P.R. China, Med Sci Monit. 2020 Jun 18;26:e926096.*

[66] *Géron A.(2019), Hands-on Machine Learning using Scikit-Learn, Keras & TensorFlow, O'Reilly 2nd Edition.*

[67] *Piepmeyer J.M and Christopher S. (1997), Low-grade gliomas: Introduction and overview Journal of Neuro-Oncology, 34 (1997), 1-3.*

[68] Karabacak M., Ozkara B. B., Mordag S., Bisdas S. (2020), *Deep learning for prediction of isocitrate dehydrogenase mutation in gliomas: a critical approach, systematic review and meta-analysis of the diagnostic test performance using a Bayesian approach*

[69] Wajeaha A., Appaji R. (2022), *Low Grade Gliomas*, StatPearls Publishing.

[70] Forst D. A., Nahed B.V., Loeffler J.S., Batchelor T.T.(2014) *Low-grade gliomas. Oncologist. 19(2014),403-13.*

[71] Essig M., Hawighorst H., Schoenberg S. O., Engenhart-Cabillic R., Fuss M., Debus J., Zuna I., Knopp M. V., van Kaick G. (2005), *Fast fluid-attenuated inversion-recovery (FLAIR) MRI in the assessment of intraaxial brain tumors, J Magn Reson Imaging, 1998 Jul-Aug, 8(1997):789-98.*

[72] Huang Z., Zhao Y., Liu Y., Song L. (2021), *GCAUNet: A group cross-channel attention residual UNet for slice based brain tumor segmentation, Biomedical Signal Processing and Control, 70(2021), 102958.*

[73] Schlemper J., Ozan Oktay, Schaap M., Heinrich M., Kainz B., Glocker B., Rueckert D. (2019), *Attention gated networks: Learning to leverage salient regions in medical images, Medical Image Analysis, 53(2019), Pages 197-207.*

[74] Isensee F., Kickingereder P., Wick W., Bendszus M., Maier-Hein K. H. (2018), *Brain Tumor Segmentation and Radiomics Survival Prediction: Contribution to the BRATS 2017 Challenge, Arxiv, 1802.10508.*

[75] Selvaraju R. R., Cogswell M., Das A., Vedantam R., Parikh D., Batra D. (2017). *Grad-CAM: Visual explanations from deep networks via gradient based localization. In Proceedings of the IEEE International Conference on Computer Vision, pages 618–626.*

[76] Ras G., Xie N., van Gerven M., Doran D. (2021), *Explainable Deep Learning: A Field Guide for the Uninitiated, Arxiv, 2004.14545.*

[77] Siddique N., Sidike P. , Elkin C., Devabhaktuni V. (2020) *U-Net and its variants for medical image segmentation: theory and applications, Arxiv, 2011.01118.*

[78] Dhall D., Kaur R., Juneja M. (2019), *Machine Learning: A Review of the Algorithms and Its Applications, Proceedings of ICRIC 2019 pp 47–63.*

[79] Singh J., Banerjee R.(2019), *A Study on Single and Multi-layer Perceptron Neural Network*, 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), pp. 35-40.

[80] Grover V. P., Tognarelli J. M., Crossey M. M., Cox I.J., Taylor-Robinson S. D., McPhail M. J. (2015), *Magnetic Resonance Imaging: Principles and Techniques: Lessons for Clinicians. J Clin Exp Hepatol. Sep, 5(2015), 246-55.*

[81] Mehan WA Jr, González RG, Buchbinder BR, Chen JW, Copen WA, et al. (2014) *Optimal Brain MRI Protocol for New Neurological Complaint. PLOS ONE 9(10): e110803.*

[82] PRINCE J. L., LINKS J. M.(2014), *Medical Imaging Signals and Systems*, Pearson Press.