



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ

ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

Τμηματοποίηση εικόνων μαγνητικής τομογραφία με σκοπό τον εντοπισμό όγκου με χρήση Generative Adversarial Networks.

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

ΠΑΡΑΛΙΚΑ ΗΛΙΑ

Επιβλέπων : Γεώργιος Ματσόπουλος
Καθηγητής Ε.Μ.Π.

Συνεπιβλέπουσα : Ουρανία Πετροπούλου
ΕΔΙΠ Ε.Μ.Π.

Αθήνα, Φεβρουάριος 2025

Η σελίδα αυτή είναι σκόπιμα λευκή.



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ
ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

**Τμηματοποίηση εικόνων μαγνητικής τομογραφία με σκοπό
τον εντοπισμό όγκου με χρήση Generative Adversarial
Networks.**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

ΠΑΡΑΛΙΚΑ ΗΛΙΑ

Επιβλέπων : Γεώργιος Ματσόπουλος
Καθηγητής Ε.Μ.Π.

Συνεπιβλέπουσα : Ουρανία Πετροπούλου
ΕΔΙΠ Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 17^η Φεβρουαρίου 2025.

(Υπογραφή)

.....
Γ. Ματσόπουλος
Καθηγητής Ε.Μ.Π.

(Υπογραφή)

.....
Α. Παναγόπουλος
Καθηγητής Ε.Μ.Π.

(Υπογραφή)

.....
Π. Τσανάκας
Καθηγητής Ε.Μ.Π.

Αθήνα, Φεβρουάριος 2025

(Υπογραφή)

.....

ΠΑΡΑΛΙΚΑΣ ΗΛΙΑΣ

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright © Παραλίκας Ηλίας, 2025.

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

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

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

Περίληψη

Η μοντέρνα ιατρική έχει καταφέρει να αντιμετωπίσει έως και να εξαλείψει μεγάλο ποσοστό ασθενειών από την ιστορία της ανθρωπότητας. Ωστόσο, ο καρκίνος αποτελεί μία από τους μεγαλύτερους αντιπάλους της. Ειδικά στις ανεπτυγμένες χώρες, συνιστά μία από τις κυριότερες αιτίες θανάτου. Επιπλέον, η περίπλοκη φύση του, οι μη κατανοητές αιτίες καθώς και τα συμπτώματά του τον καθιστούν μεγάλη πρόκληση. Πολύ μεγάλα χρηματικά ποσά δαπανούνται για την ανάπτυξη φαρμάκων και θεραπειών. Ταυτόχρονα, την τελευταία δεκαετία η ανθρωπότητα έχει σημειώσει κολοσσιαία άλματα στον τομέα της τεχνίτης νοημοσύνης. Η πρόοδος αυτή, αναπόφευκτα θα αποτελέσει αναπόσπαστο σύμμαχο στην πόλεμο της υγείας. Τροχοπέδη στην εφαρμογή της είναι το υψηλό κόστος απόκτησης δεδομένων συγκριτικά με τις υπόλοιπες εφαρμογές, καθώς και κανονισμούς προστασίας προσωπικών δεδομένων. Η παρούσα εργασία επιδιώκει να αντιμετωπίσει το παραπάνω πρόβλημα, με την χρήση παραγωγικών ανταγωνιστικών δικτύων. Τα δίκτυα αυτά παράγουν αληθοφανή δεδομένα, τα οποία μπορούν να χρησιμοποιηθούν για την λύση του προβλήματος έλλειψης δεδομένων.

Λέξεις κλειδιά: Όγκος, Τεχνίτη νοημοσύνη, Τεχνητά νευρωνικά δίκτυα, Παραγωγικά ανταγωνιστικά δίκτυα,

Abstract

Modern medicine has managed to treat and even eliminate a large percentage of diseases from humanities history. However, cancer remains one of its greatest adversaries. Especially in developed countries, it is one of the leading causes of death. Furthermore, its complex nature, the incomprehensible causes, and its symptoms make it a significant challenge. Huge sums of money are spent on the development of drugs and treatments. At the same time, in the last decade, humanity has made monumental strides in the field of artificial intelligence. This progress will inevitably become an indispensable ally in the battle for health. A major obstacle to its implementation is the high cost of acquiring data compared to other applications, as well as personal data protection regulations. This paper aims to address the aforementioned problem through the use of Generative Adversarial Networks. These networks generate realistic data, which can be used to solve the problem of data scarcity.

Keywords: Tumor, Artificial Intelligence, Artificial Neural Networks, Generative Adversarial Networks.

Η σελίδα αυτή είναι σκόπιμα λευκή.

Ευχαριστίες

Θα ήθελα να εκφράσω τις ευχαριστίες μου στον επιβλέποντα καθηγητή, κ. Γεώργιο Ματσόπουλο, καθώς και στη συνεπιβλέπουσα καθηγήτρια Δρ. Ουρανία Πετροπούλου (Ε.ΔΙ.Π. Ε.Μ.Π.). Επίσης, ευχαριστώ θερμά τους καθηγητές κ. Αθανάσιο Παναγόπουλο και κ. Παναγιώτη Τσανάκα για την πολύτιμη συμμετοχή τους στην εξεταστική επιτροπή.

Τέλος, θα ήθελα να ευχαριστήσω την οικογένειά και τους φίλους μου.

Αθήνα, Φεβρουάριος 2025
Παραλίκας Ηλίας

Πίνακας περιεχομένων

Ευρετήριο εικόνων	11
Ευρετήριο Πινάκων.....	13
Εισαγωγή	14
Κεφάλαιο 1: Καρκίνος	16
1.1 Γενικά.....	16
1.1 Καρκίνος του Εγκεφάλου.....	16
1.2 Αιτίες του καρκίνου.....	17
1.3 Συμπτώματα	17
1.4 Διάγνωση.....	19
1.5 Θεραπεία.....	19
Κεφάλαιο 2: Μαγνητική τομογραφία.....	22
2.1 Γενικά.....	22
2.1 Πυρηνικός Μαγνητικός Συντονισμός.....	22
2.2 RF Συντονισμός.....	25
2.3 Χωρικός εντοπισμός του σήματος.....	28
2.4 Ανακατασκευή εικόνας	28
2.4.1 K-χώρος.....	28
2.4.2 Compressed Sensing.....	30
2.5 Μαγνητική Τομογραφία και Διάγνωση.....	31
2.5.1 BraTS 2020.....	32
Κεφάλαιο 3: Νευρωνικά Δίκτυα	33
3.1 Γενικά.....	33
3.2 Βιολογικοί νευρώνες	33
3.3 Διαισθητική ανάλυση	34
3.2.1 Μοντελοποίηση προβλήματος.....	35
3.2.2 Backpropagation.....	37
3.3 Φορμαλιστική ανάλυση.....	40
3.3.1 Μοντέλο νευρώνα.....	40
3.3.2 Perceptron Πολλών Επιπέδων	41
3.3.3 Gradient Decent.....	43
3.3.4 Backpropagation.....	44
Κεφάλαιο 4: Συνελκτικά Νευρωνικά Δίκτυα.....	48
4.1 Γενικά.....	48
4.2 Συνέλιξη	48
4.3 Συνέλιξη Στα Νευρωνικά Δίκτυα	49

4.4 Αρχιτεκτονική Unet και Τμηματοποίηση εικόνας	51
Κεφάλαιο 5: Παραγωγικά Ανταγωνιστικά Δίκτυα.....	54
5.1 Γενικά.....	54
5.2 Διασθητική Ανάλυση	54
5.3 Wasserstein GAN	58
5.3.1 Gradient Penalty	59
5.4 Pro-Gan	59
Κεφάλαιο 6: Υλοποίηση & Αποτελέσματα.....	61
6.1 Γενικά.....	61
6.2 Σχετική Βιβλιογραφία	62
6.3 Τμηματοποίηση εικόνας.....	63
6.4 Παραγωγή δεδομένων	65
6.3.1 DCGAN.....	65
6.3.2 Skip-Connections Generator.....	66
6.3.3 Συνδυασμός μεθόδων	67
6.4.4 Pro-GAN.....	68
6.5 Χρήση δεδομένων στην εκπαίδευση	69
6.5.1 Ερμηνεία αποτελεσμάτων	73
Συμπεράσματα-Μελλοντικές Προσεγγίσεις.....	74
Βιβλιογραφία.....	75

Ευρετήριο εικόνων

Εικόνα 1: Ακτινοθεραπευτικό μηχάνημα η κεφαλή του οποίου εκπέμπει τις ακτίνες X. Η κεφαλή καθώς και ολόκληρος ο βραχίονας έχουν την δυνατότητα περιστροφής, για προσέγγιση του όγκου από διαφορετικές γωνίες κρούσης. Ταυτόχρονα, το κρεβάτι στο οποίο τοποθετείται ο ασθενής, έχει την δυνατότητα κίνησης στους τρεις άξονες, για μεγαλύτερη ακρίβεια και αυτονομία [30].	20
Εικόνα 2: Μεταλλικές ράβδοι, με δυνατότητα μετακίνησης στον κατακόρυφο άξονα. Ο μηχανισμός αυτός τοποθετείται στην έξοδο της κεφαλής που εκπέμπει τις ακτίνες X. Σκοπός τους είναι να δώσουν στις ακτίνες το κατάλληλο σχήμα, έτσι ώστε να φτάσει η ακτινοβολία μόνο στις περιοχές στόχους και να περιοριστεί η ζημιά στα γειτονικά υγιή κύτταρα [8], [31], [32].	21
Εικόνα 3: Πυρήνας ατόμου. Στο σχήμα φαίνονται τα νετρόνια και τα πρωτόνια που απαρτίζουν τον πυρήνα. Κάθε ένα από αυτά έχει το δικό του spin. Το άθροισμα των επιμέρους spin αποτελεί το συνολικό spin [31].	23
Εικόνα 4: Εφαρμογή εξωτερικού μαγνητικού πεδίου σε πυρήνα. Το μαγνητικό πεδίο τείνει να αναγκάσει τον μαγνητικό πυρήνα να ευθυγραμιστεί με αυτό. Η στροφορμή του πυρήνα όμως, συνεχίζει να τον περιστρέφει, δημιουργώντας μία περιστροφή γύρω από τον άξονά του μαγνητικού πεδίου [38].	24
Εικόνα 5: Περιστροφική κίνηση των πυρήνων, με άξονα το εξωτερικό μαγνητικό πεδίο B. Το πλήθος των πυρήνων, με z συνηστώσα ομόρροπη στο B είναι ελαφρώς μεγαλύτερο από αυτό με αντίρροπη [33].	25
Εικόνα 6: Αριστερά: Η εικόνα που έχει ένας εξωτερικός παρατηρητής, ο οποίος βλέπει το μαγνητικό πεδίο, καθώς και την περιστροφή του πυρήνα. Δεξιά: Η εικόνα που έχει ένας παρατηρητής που περιστρέφεται με την ίδια συχνότητα με τον πυρήνα. Επιπλέον παρατηρεί μηδενικό B0 [39].	26
Εικόνα 7: Εφαρμογή RF παλμού στην συχνότητα Larmor. Το αποτέλεσμα είναι η συγκέντρωση του συνολικού μαγνητικού πεδίου στον άξονα x-y [40].	27
Εικόνα 8: Αντιστοίχιση ενός σημείου το κ-χώρου με την χωρική συχνότητα που του αντιστοιχεί στο πεδίο της τελικής εικόνας [44].	29
Εικόνα 9: Διαφορετικές περιοχές του κ-χώρου προσθέτουν στο τελικό αποτέλεσμα διαφορετικό είδος πληροφορίας Στην πρώτη στήλη φαίνεται ότι τα κεντρικά σημεία συμπεριλαμβάνουν την γενική δομή του ιστού, αλλά υστερούν στις λεπτομερείς. Αντίθετα, τα συνοριακά σημεία αποθανατίζουν την αντίθεση αλλά όχι την μάζα του ιστού. Ταυτόχρονα και οι δυο ελλείψεις εικόνες παρουσιάζουν μοτίβο κυματισμού, τα οποία επισημαίνονται με ένα κόκκινο βέλος [45].	30
Εικόνα 10: Διαδοχικές τομές από Μαγνητική Τομογραφία Εγκεφάλου.	31
Εικόνα 11: Αριστερά: Τομή μαγνητικής τομογραφίας Εγκεφάλου Δεξιά: Τμηματοποίηση εικόνας που αναδεικνύει τις προσβεβλημένες περιοχές, από ειδικούς νευροακτινολόγους.	31
Εικόνα 12: Δομή νευρώνα. Στο σχήμα φαίνονται τα βασικό δομικά χαρακτηριστικά του [49].	34
Εικόνα 13: Εικόνες ανάλυσης 2x2. Στο παρακάτω παράδειγμα, θα δομηθεί ένα νευρωνικό δίκτυο όπου θα αναλάβει να τις ξεχωρίζει μεταξύ τους.	35
Εικόνα 14: Νευρωνικό Δίκτυο τεσσάρων εισόδων και 2 εξόδων	37
Εικόνα 15: Νευρωνικό Δίκτυο τεσσάρων εισόδων και 2 εξόδων με τυχαία αρχικοποιημένες τιμές των βαρών	38
Εικόνα 16: Νευρωνικό Δίκτυο Πολλών επιπέδων [56]	41
Εικόνα 17: Εικόνες χειρόγραφων ψηφίων που ανήκουν στην βάση δεδομένων MNIST [59]	42

Εικόνα 18: Γραφική αναπαράσταση του αλγορίθμου gradient descend [62].....	44
Εικόνα 19: Νευρωνικό Δίκτυο πολλών επιπέδων. Οι τιμές που αναγράφονται επάνω στα βάρη και τους νεωρώνες είναι αυτές που χρησιμοποιούνται για τον υπολογισμό του κόστους.....	46
Εικόνα 20: Γραφική αναπαράσταση συνέλιξης δύο συναρτήσεων. Στα πέντε πρώτα γραφήματα φαίνεται η ολίσθηση της $h(t)$ επάνω στην $x(t)$, ενώ στην τελευταία γραφική παράσταση παρουσιάζεται το τελικό αποτέλεσμα	49
Εικόνα 21: Εφαρμογή συνελκτικού φίλτρου σε πίνακα.	50
Εικόνα 22: Σχηματική αναπαράσταση CNN. Σε κάθε επίπεδο του δικτύου φαίνονται οι διαστάσεις της εικόνας, τα κανάλια (channels) που έχει, οι διαστάσεις του φίλτρου. Γίνεται χρήση πολλών φίλτρων στην σειρά καθώς και παράλληλα. Τέλος, τοποθετείται ένα πλήρως συνδεδεμένο νευρωνικό δίκτυο, όπως παρουσιάστηκε στο προηγούμενο κεφάλαιο [68].	50
Εικόνα 23: Παράδειγμα τμηματοποίησης εικόνας. Σε κάθε pixel της εισόδου έχει αποδοθεί ένας χαρακτηρισμός από το μοντέλο. Στην πρώτη περίπτωση υπάρχουν δύο επιλογές, είτε άλογο είτε όχι, ενώ στην δεύτερη το μοντέλο μπορεί να επιλέξει από μία πικοιλία κλάσεων, όπως αμάξι, άνθρωπος, δένδρο [69].	51
Εικόνα 24: Αρχιτεκτονική Unet.....	52
Εικόνα 25: Σχηματική αναπαράσταση skip connection [70].....	52
Εικόνα 26: Σχηματική αναπαράσταση GAN. Η είσοδος το generator αποτελείται από ένα τυχαίο διάνυσμα, noise και παράγει το generated sample. Στην συνέχεια ο discriminator ανατεθεί την τιμή στα πραγματικά και συνθετικά δείγματα. Η τιμή αυτή χρησιμοποιείται ως συνάρτηση κόστους για την εκπαίδευση τους [73].	56
Εικόνα 27: Αρχιτεκτονική DCGAN	57
Εικόνα 28: Συνθετικά δεδομένα παραγόμενα από GAN, εκπαιδευμένα στην βάση δεδομένων MNIST [71]	57
Εικόνα 29: Σύγκριση διαχωριστικής ευθείας GAN με WGAN. Είναι φανερό ότι στην περίπτωση την GAN, σε ακραίες τιμές, η παράγωγος της συνάρτησης μηδενίζεται (vanishing gradient) και ο generator παύει να εκπαιδεύεται [75].....	59
Εικόνα 30: Σχηματική αναπαράσταση Pro-GAN. Στα αριστερά φαίνεται το δίκτυο το οποίο παράγει τις εικόνες με χαμηλή ανάλυση 4x4. Δεξιά του βρίσκεται το επόμενο βήμα που παράγει 8x8 εικόνες, το πρώτο επίπεδο του οποίου είναι το αρχικό δίκτυο. Η διαδικασία αυτή επαναλαμβάνεται μέχρι να επιτευχθεί η επιθυμητή ανάλυση [77].....	60
Εικόνα 31: Παράδειγμα τομής που χρησιμοποιήθηκε ως δεδομένο εκπαίδευσης. Επάνω φαίνονται τα 4 modalities, ενώ κάτω φαίνονται οι τέσσερις διαφορετικές κατηγοριοποιήσεις.	62
Εικόνα 32: Γραφική αναπαράσταση της διαδικασίας εκπαίδευσης. Αριστερά: Dice loss στο test dataset. Δεξιά: Dice loss στο validation dataset.	64
Εικόνα 33: Γραφική αναπαράσταση της διαδικασίας εκπαίδευσης με χρήση τεχνικών προσαύξησης δεδομένων. Αριστερά: Dice loss στο test dataset. Δεξιά: Dice loss στο validation dataset.	65
Εικόνα 34: Αποτελέσματα της αρχιτεκτονικής DCGAN	66
Εικόνα 35: Αποτελέσματα της αρχιτεκτονικής Skip-Connection GAN.....	66
Εικόνα 36: Ενδιάμεσο στάδιο παραγωγής δεδομένων, όπου είναι φανερή η παρουσία artifacts.....	67
Εικόνα 37: Αποτελέσματα της αρχιτεκτονικής πολλαπλών discriminators	68
Εικόνα 38: Αποτελέσματα αρχιτεκτονικής Pro-GAN	69
Εικόνα 39: Παράδειγμα μη πετυχημένου δείγματος με την χρήση Pro-GAN.....	70

Εικόνα 40: Γραφική αναπαράσταση της διαδικασίας εκπαίδευσης με χρήση προ εκπαιδευμένου μοντέλου στα συνθετικά δεδομένα. Η συμβολή των συνθετικών δεδομένων είναι ξεκάθαρη από το γεγονός ότι η συνάρτηση κόστους ξεκινάει από τιμές, χαμηλότερες από το 1, υποδεικνύοντας ότι το μοντέλο έχει μάθει χρήσιμες πληροφορίες.....	71
Εικόνα 41: Training και Validation Score κατά την διάρκεια της εκπαίδευσης.	71
Εικόνα 42: Σύγκριση αποτελεσμάτων με ή χωρίς εφαρμογή προ εκπαίδευσης σε συνθετικά δεδομένα.	72
Εικόνα 43: Πίνακας σύγκρισης της απόδοσης του καλύτερου μοντέλου.....	72

Ευρετήριο Πινάκων

Πίνακας 1: Μετρικές precision, recall και F1-Score για κάθε μία από τις κλάσεις.	73
--	----

Εισαγωγή

Ο καρκίνος αποτελεί μια από τις μεγαλύτερες προκλήσεις στον τομέα της ιατρικής αλλά και της μοντέρνας βιολογικής έρευνας [1], με σημαντικές κοινωνικές και οικονομικές επιπτώσεις [2]. Ο καρκίνος του εγκεφάλου αποτελεί μία από τις πιο πολύπλοκες και προκλητικές μορφές του, τόσο στη διάγνωση όσο και στην θεραπεία του [3]. Η πολυπλοκότητα του εγκεφάλου ως όργανο, σε συνδυασμό με την δυσκολία πρόσβασης σε αυτόν, καθιστά την έρευνα και την ανάπτυξη νέων θεραπευτικών προσεγγίσεων ιδιαίτερα απαιτητική [4]. Η ακριβής ανάλυση των δεδομένων από μαγνητικές τομογραφίες είναι κρίσιμη για την διάγνωση αλλά και την έγκαιρη και αποτελεσματική παρέμβαση [5]. Η μελέτη του καρκίνου του εγκεφάλου συνοδεύεται από πολλαπλές προκλήσεις, ιδιαίτερα όσον αφορά τη συλλογή και ανάλυση δεδομένων [4], [6]. Τα δεδομένα που αφορούν περιπτώσεις ασθενών είναι συχνά περιορισμένα λόγω του σπάνιου χαρακτήρα ορισμένων τύπων όγκων, καθώς και λόγω ηθικών και νομικών περιορισμών στη χρήση ιατρικών πληροφοριών [7]. Παράλληλα, η ανομοιογένεια των όγκων και οι διαφορετικές κλινικές εκδηλώσεις καθιστούν δύσκολη την εξαγωγή γενικεύσιμων συμπερασμάτων από τις υπάρχουσες βάσεις δεδομένων [5].

Τα προαναφερόμενα δεδομένα όμως, είναι απαραίτητα για την εφαρμογή αλγορίθμων μηχανικής μάθησης και τεχνίτης νοημοσύνης [8]. Οι αλγόριθμοι αυτοί έχουν αποφέρει εξαιρετικά αποτελέσματα σε πολλούς τομείς. Μοντέλα όπως το πλέον πασίγνωστο chat gpt [9], σε πολύ σύντομο χρονικό διάστημα έχουν εισχωρήσει στην καθημερινότητα. Στην εξαιρετική απόδοση των μοντέλων σαν και το προαναφερόμενο, σημαντική ήταν η συμβολή της πληθώρας δεδομένων που υπήρχαν στην διάθεσή κατά την διάρκεια της εκπαίδευσης [10]. Στις εφαρμογές ιατρικής απεικόνισης, όπως η τμηματοποίηση εικόνας, ο εντοπισμός δηλαδή όγκου σε μαγνητική τομογραφία, για τους λόγους που αναφέρθηκαν, δεν συναντάται αυτή η πληθώρα δεδομένων.

Η παρούσα διπλωματική εργασία έχει σκοπό να δώσει στον αναγνώστη μία ολοκληρωμένη εικόνα της εφαρμογής των νευρωνικών δικτύων στην αντιμετώπιση του παραπάνω προβλήματος. Δίνονται ορισμοί καθώς και διαισθητικές εξηγήσεις για όλες τις απαραίτητες έννοιες, έτσι ώστε να γίνει κατανοητή σε βάθος η ιατρική απεικόνιση με την χρήση νευρωνικών δικτύων. Επιπλέον εξετάζεται η εφαρμογή παραγωγικών ανταγωνιστικών δικτύων για την αντιμετώπιση του προβλήματος έλλειψης δεδομένων, το οποίο συνοδεύει τις ιατρικές εφαρμογές.

Γίνεται υλοποίηση και αξιολόγηση διαφορετικών αρχιτεκτονικών και μεθόδων εκπαίδευσης μοντέλων παραγωγής δεδομένων, με απώτερο σκοπό την βελτίωση της απόδοσης μοντέλων τμηματισμού μαγνητικής τομογραφίας.

Κεφάλαιο 1: Καρκίνος

1.1 Γενικά

Ο όρος καρκίνος αναφέρεται σε μία ομάδα νόσων που χαρακτηρίζονται από τον ανεξέλεγκτο διαχωρισμό των κυττάρων, λόγω μετάλλαξης στο γενετικό τους υλικό. Τα κύτταρα αυτά συσσωρεύονται και δημιουργούν έναν «όγκο». Η παρουσία του όγκου αυτού καθαυτού δεν αποτελεί απειλή για τον οργανισμό, ο λεγόμενος καλοήθεις [11], εκτός από την σπάνια περίπτωση, όπου ο όγκος πιέζει τα γειτονικά του όργανα. Το πρόβλημα που δημιουργείται, είναι όταν τα καρκινικά κύτταρα εξαπλώνονται, δηλαδή καταλαμβάνουν και καταστρέφουν υγιή κύτταρα [12]. Η εξάπλωση μπορεί να γίνει είτε στα κοντινά κύτταρα, είτε μέσω την κυκλοφορίας του αίματος και το λεμφικό σύστημα [13], σε ολόκληρο το σώμα [14]. Τα κακοήθη κύτταρα έχουν την ικανότητα να διαιρούνται χωρίς περιορισμούς [15]. Αντί να υπακούν στους μηχανισμούς που ρυθμίζουν τον κύκλο της κυτταρικής διαίρεσης, συνεχίζουν να διαιρούνται ακόμα και όταν δεν υπάρχει ανάγκη για αύξηση του αριθμού των κυττάρων. Ο τρόπος με τον οποίο τα καρκινικά κύτταρα βλάπτουν τον ασθενή διαφέρει, ανάλογα το σημείο που έχει προσβληθεί, καθώς και το στάδιο του καρκίνου. Μπορεί είτε ο όγκος να προκαλεί κάποιο μηχανικό εμπόδιο που παρεμβάλλει στην λειτουργία του οργάνου, όπως γίνεται στον καρκίνο του πεπτικού συστήματος [16], είτε να προσβάλλει κατευθείαν τα κύτταρα του οργάνου, καταστρέφοντάς τα, όπως στην περίπτωση του καρκίνου του πνεύμονα [17]. Ο καρκίνος αποτελεί την μεγαλύτερη αιτία θανάτου στις ανεπτυγμένες χώρες [1]. Η πολυπλοκότητά του, οι πολυάριθμες μορφές και υποπεριπτώσεις τις ασθένειας αυτής έχουν φέρει την εύρεση θεραπείας σε πολύ υψηλή προτεραιότητα, με την Ευρωπαϊκή Ένωση να διαθέτει 4 δισεκατομμύρια ευρώ για την πρόληψη και καταπολέμησή του [2].

1.1 Καρκίνος του Εγκεφάλου

Όπως προαναφέρθηκε, ο τρόπος που μπορεί να βλάψει ένας όγκος τον ασθενή είναι είτε μέσω πίεσης είτε μέσω επέκτασης και προσβολής στα γειτονικά κύτταρα. Υπάρχουν δύο βασικές κατηγορίες, τους πρωτοπαθείς όγκους, που ξεκινούν στον εγκέφαλο, και τους μεταστατικούς, που προέρχονται από καρκίνους άλλων οργάνων και εξαπλώνονται [14]. Όπως είναι φυσικό, όσο πιο ευαίσθητη είναι η περιοχή στην οποία εμφανίζεται, τόσο πιο σοβαρή η εμφάνισή του. Ο καρκίνος του εγκεφάλου αποτελεί

μεγάλη πρόκλησή, καθώς κάθε περιοχή του εγκεφάλου είναι υπεύθυνη για μία η παραπάνω λειτουργία ζωτικής σημασίας του ανθρώπου [3]. Για αυτόν τον λόγο, η χειρουργική αφαίρεση καθίσταται εξαιρετικά δύσκολη διαδικασία [4], σε σύγκριση με της μαστεκτομής, στην οποία αφαιρείται ολόκληρο το στήθος, μαζί με τον όγκο [18]. Τα συμπτώματα του καρκίνου του εγκεφάλου ποικίλλουν ανάλογα με το μέγεθος, την τοποθεσία και την ταχύτητα ανάπτυξης του όγκου. Καθώς προσβάλλεται το νευρικό σύστημα, οποιαδήποτε λειτουργία του εγκεφάλου, άρα και του σώματος μπορεί να επηρεαστεί [6].

1.2 Αιτίες του καρκίνου

Οι αιτίες του καρκίνου είναι πολύπλοκες και συχνά αποτελούν συνδυασμό διαφόρων παραγόντων. Ο καρκίνος οφείλεται στις αλλαγές στο γενετικό των κυττάρων, που προκαλεί την ανεξέλεγκτη διαίρεση και αναπαραγωγή του. Οι αιτίες την μετάλλαξης αυτής είναι ποικίλες αν και οι ακριβείς μηχανισμοί δεν είναι πλήρως κατανοητοί. Οι γενετικοί παράγοντες παίζουν σημαντικό ρόλο, καθώς ορισμένες μεταλλάξεις στο DNA μπορούν να κληρονομηθούν από τους γονείς. Μερικές μορφές καρκίνου είναι είτε αποκλειστικά, είτε σε μεγάλο βαθμό συνδεδεμένες με γενετικές προδιαθέσεις, όπως για παράδειγμα ο καρκίνος του παγκρέατος [19]. Ανάμεσα σε άλλα, μεταλλάξεις στα ογκοκατασταλτικά γονίδια, όπως είναι τα *BRCA1* και *BRCA2* [20] έχουν συνδεθεί με μεγάλη προδιάθεση, καθώς μειώνουν την δυνατότητα καταστολής της ασθένειας [21]. Ο τρόπος ζωής επίσης διαδραματίζει καθοριστικό ρόλο στην ανάπτυξη καρκίνου. Είναι πλέον γνωστό ότι συνήθειες όπως το κάπνισμα [22], η κατανάλωση αλκοόλ, η παχυσαρκία και η έλλειψη σωματικής άσκησης και η διατροφή είναι από τους κύριους παράγοντες, η χρόνια φλεγμονή, το στρες, και το ανοσοποιητικό σύστημα παίζουν επίσης ρόλο στην ανάπτυξη της νόσου [23]. Εκτιμάται ότι στις ανεπτυγμένες χώρες, οι μισές περιπτώσεις καρκίνου θα μπορούσαν να αποφευχθούν, με προσαρμογή του τρόπου ζωής [24]. Για παράδειγμα, στην περίπτωση του καρκίνου του μαστού, η μη εξαγωγή του γάλακτος μετά την γέννα, αυξάνει τις πιθανότητες εμφάνισής του [24]. Ένας ακόμα σημαντικός παράγοντας είναι η ηλικία [25]. Τέλος, η έκθεση σε ραδιενέργεια, αυξάνει την πιθανότητα κάθε μετάλλαξης του DNA, συνεπώς και την πιθανότητα εμφάνισης καρκίνου [20].

1.3 Συμπτώματα

Ο καρκίνος μπορεί να εκδηλώνεται με ποικίλα συμπτώματα, ανάλογα με το όργανο που προσβάλλει και το στάδιο της νόσου [6]. Μερικά από τα πιο συχνά γενικά συμπτώματα περιλαμβάνουν [26] :

- Ανεξήγητη απώλεια βάρους
- Επίμονο πυρετό ή νυχτερινές εφιδρώσεις
- Έντονη κόπωση που δεν υποχωρεί με την ξεκούραση
- Απώλεια όρεξης
- Αλλαγές στο δέρμα όπως κιτρίνισμα, σκουρόχρωμες κηλίδες ή αλλαγές σε υπάρχοντες σπίλους
- Ανεξήγητη αιμορραγία ή μώλωπες
- Επίμονος βήχας ή βραχνάδα
- Δυσκολία στην κατάποση
- Αλλαγές στις συνήθειες του εντέρου ή της ουροδόχου κύστης.
- Χρόνιος πόνος

Συγκεκριμένα, ο καρκίνος του εγκεφάλου μπορεί να εκδηλωθεί με διάφορα συμπτώματα, τα οποία συχνά εξαρτώνται από το μέγεθος και τη θέση του όγκου. Τα πιο συνηθισμένα συμπτώματα περιλαμβάνουν [12]:

- Επίμονους και έντονους πονοκεφάλους, ειδικά το πρωί ή κατά τη διάρκεια της νύχτας,
- Ναυτία και εμετό χωρίς προφανή αιτία
- Προβλήματα όρασης όπως θολή ή διπλή όραση, δυσκολία στην ισορροπία και το συντονισμό των κινήσεων, επιληπτικές κρίσεις (ιδιαίτερα σε άτομα που δεν είχαν ποτέ πριν),
- Αλλαγές στη συμπεριφορά ή την προσωπικότητα,
- Δυσκολία στην ομιλία ή την κατανόηση
- Προοδευτική αδυναμία ή παράλυση σε ένα μέρος του σώματος

Ταυτόχρονα, πέρα από τις σωματικές επιπτώσεις, υπάρχουν και πολύπλευρες κοινωνικές επιπτώσεις που επηρεάζουν τόσο τους ασθενείς όσο και το ευρύτερο κοινωνικό περιβάλλον τους [6]. Οι ασθενείς συχνά αντιμετωπίζουν σημαντικές αλλαγές στην καθημερινότητά τους, συμπεριλαμβανομένης της απώλειας της εργασιακής τους ικανότητας και της κοινωνικής τους ζωής. Η ασθένεια μπορεί να οδηγήσει σε κοινωνική απομόνωση, καθώς οι θεραπείες και οι σωματικές αλλαγές περιορίζουν τη συμμετοχή σε κοινωνικές δραστηριότητες [4].

1.4 Διάγνωση

Η ανάλυση των συμπτωμάτων του καρκίνου είναι μια περίπλοκη διαδικασία που απαιτεί ιδιαίτερη προσοχή και εξειδικευμένη ιατρική γνώση [27]. Τα συμπτώματα συχνά μπορεί να είναι παραπλανητικά, καθώς πολλά από αυτά μοιάζουν με συμπτώματα άλλων, λιγότερο σοβαρών παθήσεων. Για παράδειγμα, η κόπωση και η απώλεια βάρους θα μπορούσαν να οφείλονται σε άγχος ή διατροφικές διαταραχές, ενώ ο πονοκέφαλος θα μπορούσε να είναι απλά ένας πονοκέφαλος τάσης. Ωστόσο, το κρίσιμο στοιχείο είναι η επιμονή και η διάρκεια των συμπτωμάτων - όταν διαρκούν για μεγάλο χρονικό διάστημα ή επιδεινώνονται σταδιακά, αυτό αποτελεί σημαντική ένδειξη που χρήζει περαιτέρω διερεύνησης. Επιπλέον, ο συνδυασμός πολλαπλών συμπτωμάτων ταυτόχρονα μπορεί να είναι πιο ενδεικτικός σοβαρής κατάστασης σε σύγκριση με μεμονωμένα συμπτώματα. Είναι επίσης σημαντικό να ληφθεί υπόψη το οικογενειακό ιστορικό και οι παράγοντες κινδύνου του κάθε ατόμου κατά την αξιολόγηση των συμπτωμάτων. Η έγκαιρη αναγνώριση και σωστή ερμηνεία των συμπτωμάτων μπορεί να οδηγήσει σε πρόωμη διάγνωση, που είναι καθοριστική για την επιτυχή αντιμετώπιση της νόσου. [6]. Εάν υπάρχει υποψία για εμφάνιση καρκίνου, ο ασθενής θα υποβληθεί σε μία σειρά από νευρολογικές εξετάσεις [27], ενώ κυριότερη μέθοδος διάγνωσης είναι η Μαγνητική τομογραφία [4], [12], [27]. Η Μαγνητική συνήθως προτιμάται της αξονικής τομογραφίας, καθώς δεν επιβαρύνει τον οργανισμό με επιπλέον ακτινοβολία. Έπειτα ακολουθεί η διαδικασία της τμηματοποίησης εικόνας, κατά την οποία ένας γιατρός, συνήθως ακτινολόγος αναλαμβάνει να εντοπίσει τις περιοχές οι οποίες είναι προσβεβλημένες. Όπως είναι φυσικό, η διαδικασία αυτή είναι εξαιρετικά χρονοβόρα [7] ενώ γίνονται λάθη τα οποία οφείλονται σε συνθήκες όπως η εμπειρία του γιατρού καθώς η ψυχολογική και σωματική του κατάσταση την στιγμή της διάγνωσης [5], τα οποία θα μπορούσαν να αποφευχθούν εάν μειωνόταν ο ανθρώπινος παράγοντας από την διαδικασία.

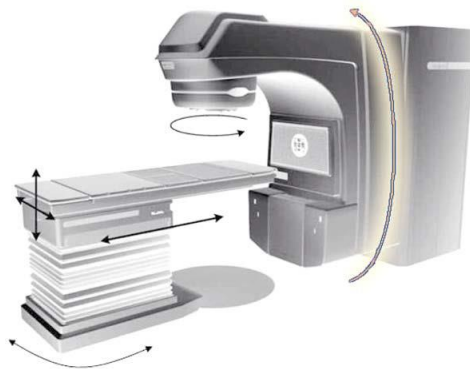
1.5 Θεραπεία

Καθώς ο καρκίνος αποτελεί μία πολύ περίπλοκη ασθένεια, δεν υπάρχει μοναδικός τρόπος θεραπείας, ενώ κάθε πρόβλημα προσεγγίζεται διαφορετικά, αναλόγως τις ιδιομορφίες του, τα μέσα που βρίσκονται στην διάθεση των γιατρών αλλά και την γνώμη της ασθενούς [28]. Οι κύριες τακτικές θεραπείας του καρκίνου είναι οι εξής.

- Αφαιρετική επέμβαση
- Ακτινοθεραπεία
- Χημειοθεραπεία

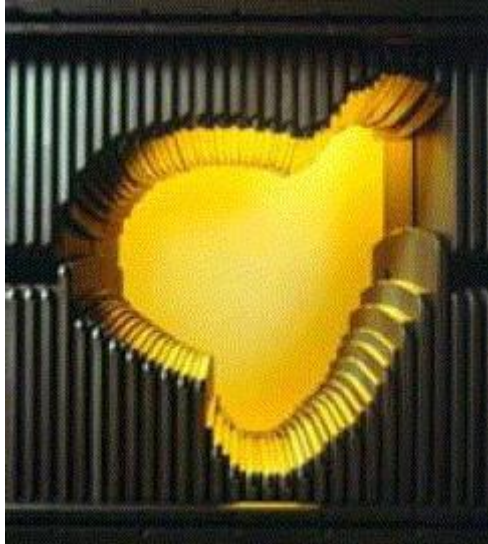
Η πιο συνηθισμένη προσέγγιση είναι ο συνδυασμός των παραπάνω μεθόδων, με το χειρουργείο να συμπεριλαμβάνεται πολύ συχνά [18], [29], όπου αφαιρείτε ο όγκος καθώς και μικρή περιοχή γύρω του [18].

Στην περίπτωση της ακτινοθεραπείας, ακτίνες X κατευθύνονται προς τον όγκο του ασθενή, με σκοπό να τον καταστρέψουν. Όπως είναι αναμενόμενο, οι ακτίνες υψηλής ενέργεια βλάπτουν και τα γειτονικά, υγιή κύτταρα. Για αυτόν τον λόγο, οι ακτίνες χορηγούνται στον ασθενή από περιστρεφόμενη πηγή, με σκοπό την ομοιόμορφη κατανομή της ακτινοβολίας στον όγκο, καθώς και την αποφυγή της επιβάρυνσης των ιστών και οργάνων που περιβάλλουν τα κύτταρα στόχο.



Εικόνα 1: Ακτινοθεραπευτικό μηχάνημα η κεφαλή του οποίου εκπέμπει τις ακτίνες X. Η κεφαλή καθώς και ολόκληρος ο βραχίονας έχουν την δυνατότητα περιστροφής, για προσέγγιση του όγκου από διαφορετικές γωνίες κρούσης. Ταυτόχρονα, το κρεβάτι στο οποίο τοποθετείται ο ασθενής, έχει την δυνατότητα κίνησης στους τρεις άξονες, για μεγαλύτερη ακρίβεια και αυτονομία [30].

Επιπλέον, η ακτίνα περνά μέσα από ρυθμιζόμενες ράβδους από υλικό, με σκοπό η ακτίνα να λάβει το σχήμα του όγκου, περειαίρω περιορίζοντας την διαρροή στα γειτονικά κύτταρα.



Εικόνα 2: Μεταλλικές ράβδοι, με δυνατότητα μετακίνησης στον κατακόρυφο άξονα. Ο μηχανισμός αυτός τοποθετείται στην έξοδο της κεφαλής που εκπέμπει τις ακτίνες Χ. Σκοπός τους είναι να δώσουν στις ακτίνες το κατάλληλο σχήμα, έτσι ώστε να φτάσει η ακτινοβολία μόνο στις περιοχές στόχους και να περιοριστεί η ζημιά στα γειτονικά υγιή κύτταρα [8], [31], [32].

Τέλος, η χημειοθεραπεία, χρησιμοποιεί χημικές ουσίες για την καταπολέμηση των καρκινικών κυττάρων. Στόχος της είναι να εμποδίσει την ανάπτυξη και την εξάπλωση αυτών των κυττάρων, συχνά επηρεάζοντας την ικανότητά τους να διαιρούνται.

Όπως είναι φανερό, η ακριβής οριοθέτηση του όγκου είναι εξαιρετικά σημαντική. Είναι απαραίτητο να συμπεριλάβουμε ολόκληρο τον όγκο, καθώς εάν απομείνουν κύτταρα υπάρχει η πιθανότητα να πολλαπλασιαστούν και να επανεμφανιστεί ο καρκίνος. Ταυτόχρονα, είναι απαραίτητο να μην αναδειχθούν υγιή κύτταρα ως καρκινικά, καθώς θα βλάψουμε αχρείαστα τον ασθενή. Στον διαχωρισμό αυτόν είναι απαραίτητη μία απεικονιστική μέθοδος, η μαγνητική τομογραφία.

Κεφάλαιο 2: Μαγνητική τομογραφία

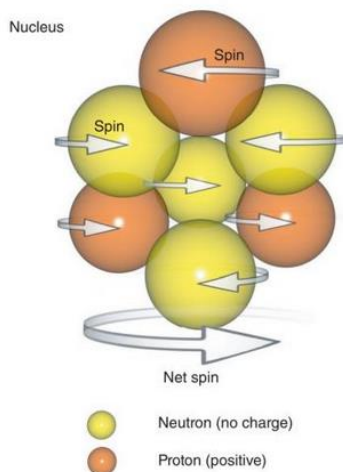
2.1 Γενικά

Όπως προαναφέρθηκε, μία από τις μεθόδους για την διάγνωση, εντοπισμό και την αξιολόγηση της θεραπείας είναι η μαγνητική τομογραφία (MT). Εφευρέθηκε το 1973 από τον Paul Lauterbur και η χρήση της εγκρίθηκε το 1985 [33]. Από τότε και μετά έχει αποτελέσει αποτελεί μία από τις σημαντικότερες εφευρέσεις στην απεικονιστική ιατρική, ίσως η σημαντικότερη μετά την ακτινογραφία [34]. Έχει την δυνατότητα να καταγράφει τρισδιάστατες εικόνες διατομής στο εσωτερικό του σώματος. Αυτό που την ξεχωρίζει είναι ότι αποτελεί μη επεμβατική μέθοδο [35], δηλαδή δεν χρειάζεται κάποια τομή για να εισέλθει κάποιου είδους κάμερα ή αισθητήρας. Επιπλέον, καθώς γίνεται χρήση μη ιονίζουσας ηλεκτρομαγνητικής ακτινοβολίας [36], γεγονός που συνεπάγεται ότι δεν συμπεριλαμβάνει καθόλου ραδιενέργεια, σε αντίθεση με την ακτινογραφία. Η μόνη πιθανή παρεμβολή της μαγνητικής τομογραφίας είναι στην περίπτωση χορήγησης σκιαγραφικού φαρμάκου. Αυτό το φάρμακο συμβάλει στην αντίθεση των ιστών του τελικού προϊόντος, κάνοντας πιο εύκολο τον διαχωρισμό τους. Μερικά σκιαγραφικά φάρμακα, όπως τα παραμαγνητικά υλικά [33], έχουν σχετικά μεγάλη οξύτητα. Παρόλα αυτά, η χορήγησή τους δεν είναι απαραίτητη, άρα υπάρχει η επιλογή μηδενικής παρέμβασης στο σώμα. Ένα επιπλέον πλεονέκτημα της μαγνητικής τομογραφίας, σε σχέση με παρόμοιες απεικονιστικές μεθόδους, όπως η αξονική τομογραφία, είναι η ευελιξία της, καθώς έχει πολλές παραμέτρους που μπορούν να προσαρμοστούν σε κάθε συγκεκριμένο πρόβλημα [37]. Η μαγνητική τομογραφία συνήθως εκμεταλλεύεται το φαινόμενο του πυρηνικού μαγνητικού συντονισμού [35].

2.1 Πυρηνικός Μαγνητικός Συντονισμός

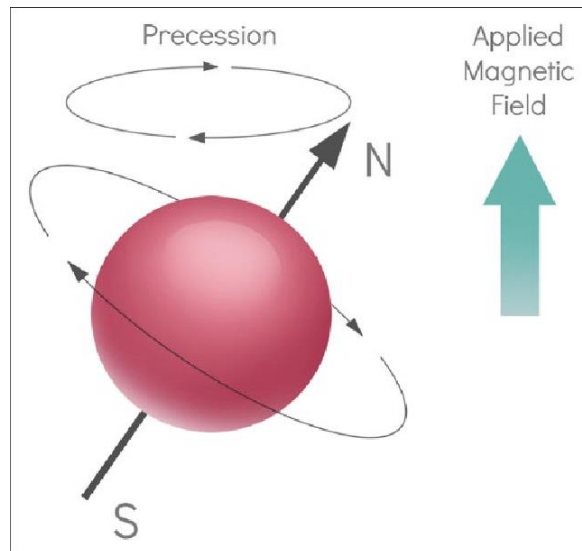
Η καλύτερη ερμηνεία που έχει δημιουργηθεί από την ανθρωπότητα για την εξήγηση των φυσικών φαινομένων, είναι η κβαντομηχανική. Έτσι και το φαινόμενο του πυρηνικού μαγνητικού συντονισμού, έχει τις ρίζες του σε κβαντομηχανικά φαινόμενα [36]. Παρόλα αυτά, για την κατανόηση της μαγνητικής τομογραφίας, μπορούν να γίνουν μερικές κλασικές παραδοχές, οι οποίες εάν και όχι απολύτως ορθές, προσεγγίζουν με ικανοποιητική ακρίβεια την πραγματικότητα. Αρχικά τα πρωτόνια και τα νετρόνια θεωρούνται θετικά φορτισμένα και ουδέτερα σωματίδια, αντίστοιχα, τα

οποία διαθέτουν στροφορμή, γνωστό και ως spin. Στην περίπτωση του πρωτονίου, ο συνδυασμός της στροφορμής με το φορτίο που διαθέτει, δημιουργεί ένα μαγνητικό δίπολο [35]. Οι πυρήνες όλων των ατόμων απαρτίζονται από πρωτόνια και νετρόνια. Αυτό έχει ως αποτέλεσμα, το σύνολο του πυρήνα να συμπεριφέρεται ως μία μάζα, με συνολικό φορτίο, spin και μαγνητικό πεδίο, το άθροισμα των χαρακτηριστικών των επιμέρους στοιχείων του [36].



Εικόνα 3: Πυρήνας ατόμου. Στο σχήμα φαίνονται τα νετρόνια και τα πρωτόνια που απαρτίζουν τον πυρήνα. Κάθε ένα από αυτά έχει το δικό του spin. Το άθροισμα των επιμέρους spin αποτελεί το συνολικό spin [31]

Όπως όλοι οι μαγνήτες, έτσι και οι πυρήνες, τη στιγμή που θα βρεθούν μέσα σε ένα μαγνητικό πεδίο B , έχουν την τάση να ευθυγραμμιστούν, ή να βρεθούν αντίρροπα με αυτό, καθώς αυτές αποτελούν τις καταστάσεις ισορροπίας. Ταυτόχρονα, η στροφορμή του πυρήνα τον περιστρέφει γύρω από τον άξονά του. Όταν ο άξονας περιστροφής τείνει να μεταβληθεί, η παρουσία περιστροφικής ταχύτητας, αθροίζεται με την επιτάχυνση του εξωτερικού μαγνητικού πεδίου και έχει ως αποτέλεσμα κίνηση σε διαφορετική κατεύθυνση από αυτή που υποδεικνύει το πεδίο. Αντί λοιπόν να ευθυγραμμιστούνε πλήρως, θα βρεθούν σε μία μόνιμη κατάσταση μετάπτωσης [36], όπου το μαγνητικό πεδίο θα προσπαθεί να τα ευθυγραμμίσει, ενώ η στροφορμή θα τα περιστρέφει. Αυτό έχει ως αποτέλεσμα, την κυκλική κίνηση του πυρήνα, άρα και του διανύσματος του μαγνητικού πεδίου του, γύρω από τον άξονα του εξωτερικά επιβαλλόμενου μαγνητικού πεδίου [37].



Εικόνα 4: Εφαρμογή εξωτερικού μαγνητικού πεδίου σε πυρήνα. Το μαγνητικό πεδίο τείνει να αναγκάσει τον μαγνητικό πυρήνα να ευθυγραμιστεί με αυτό. Η στροφορμή του πυρήνα όμως, συνεχίζει να τον περιστρέφει, δημιουργώντας μία περιστροφή γύρω από τον άξονά του μαγνητικού πεδίου [38].

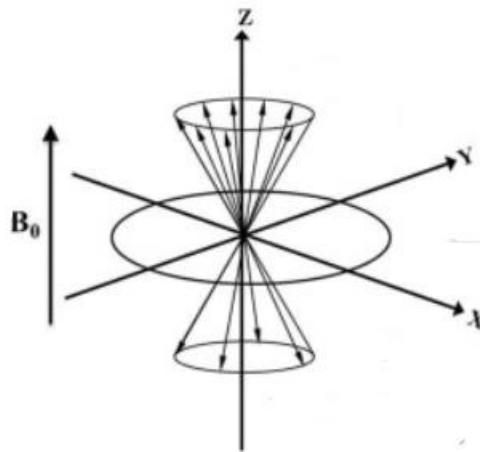
Η συχνότητα περιστροφής υπολογίζεται μέσω της εξίσωσης Larmor:

$$\omega_0 = \gamma * B_0 \quad [37]$$

Όπου

- γ ο γυρομαγνητικός λόγος, μοναδικός για κάθε πυρήνα
- B_0 ένταση του μαγνητικού πεδίου
- ω_0 Η συχνότητα περιστροφής

Κάθε πυρήνας μπορεί να περιστρέφεται μόνο παράλληλα ή αντι-παράλληλα με το μαγνητικό πεδίο. Η φάση, παρόλα αυτά, είναι αυθαίρετη, άρα μπορεί να πάρει όλες τις δυνατές τιμές [36].



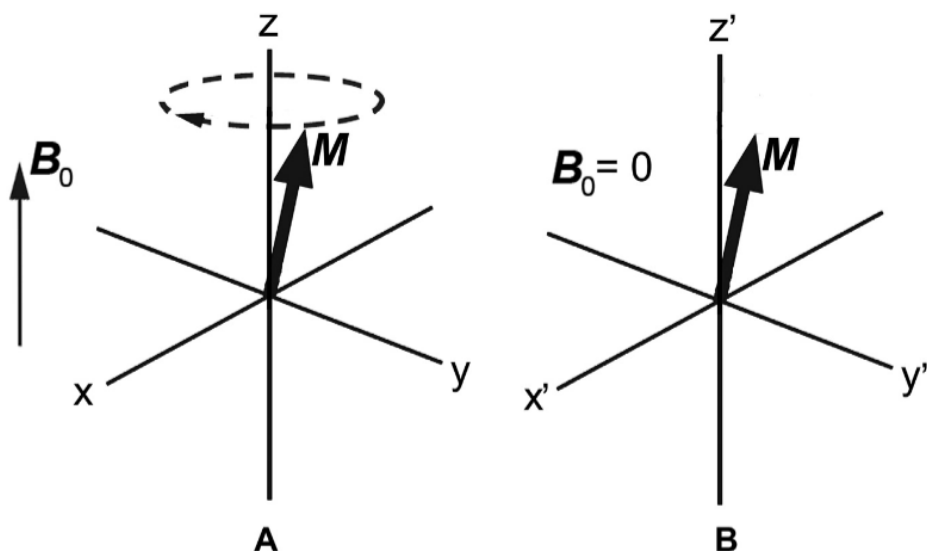
Εικόνα 5: Περιστροφική κίνηση των πυρήνων, με άξονα το εξωτερικό μαγνητικό πεδίο B . Το πλήθος των πυρήνων, με z συνιστώσα ομόρροπη στο B είναι ελαφρώς μεγαλύτερο από αυτό με αντίρροπη [33].

Η τυχαιότητα της φάσης των πυρήνων μεταξύ τους έχει ως αποτέλεσμα, πρακτικά, να εξαφανίζεται η x και y συνιστώσα του μαγνητικού πεδίου, καθώς για κάθε πυρήνα, θα υπάρχει και ένας που έχει με διαφορά φάσης 180° , άρα αντίθετη x, y συνιστώσα. Αυτό δεν συμβαίνει στον z άξονα, καθώς οι πυρήνες που είναι ευθυγραμμισμένοι με το πεδίο είναι ελαφρώς περισσότεροι από αυτούς που είναι στην αντίθετη κατεύθυνση [37]. Συνεπώς, η συνολική μαγνήτιση που αποτελεί το μετρήσιμο μέγεθος, είναι ένα στάσιμο διάνυσμα, παράλληλο με τον άξονα του B_0 , δηλαδή τον z άξονα. Για να εξασφαλίζουμε ότι θα ισχύει η στατιστική υπόθεση που κάναμε παραπάνω, καθώς και το ότι θα λάβουμε σήμα επαρκής έντασης, διαλέγουμε να επικεντρωθούμε στους πυρήνες υδρογόνο, που αποτελούνται από ένα μόνο πρωτόνιο, καθώς υπάρχουν άφθονοι στο ανθρώπινο [37]. Το υδρογόνο έχει γυρομαγνητικό λόγο $42.58 \text{ MegaHertz/Tesla}$ [35]. Για τις τιμές του B_0 από 1.5 και 3 Tesla [39], που συναντάται στην πράξη, η συχνότητα Larmor είναι στην περιοχή των ραδιοκυμάτων(RF).

2.2 RF Συντονισμός

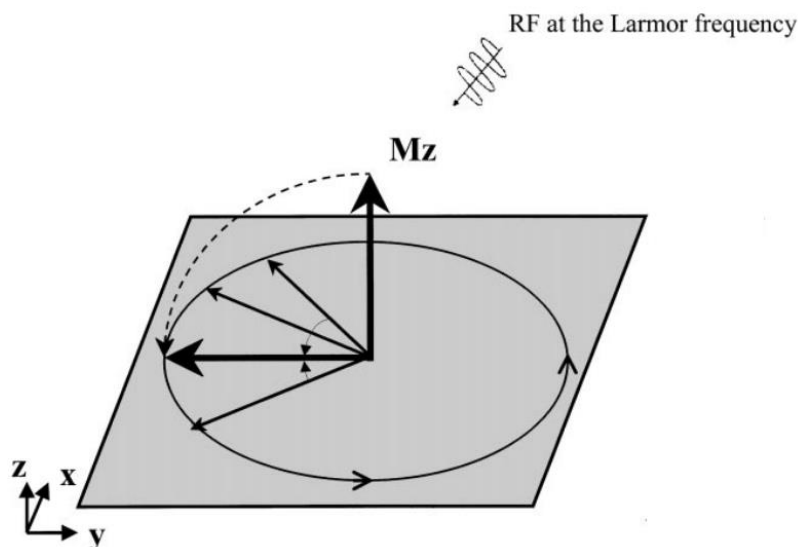
Η ιδιοσυχνότητα σε ένα σύστημα ταλαντώνεται είναι η τιμή στην οποία, όταν το σύστημα ταλαντωθεί, έχει το μέγιστο δυνατό πλάτος. Παρομοίως και εδώ, ένας πυρήνας έχει την δυνατότητα να αλληλεπιδράσει με ένα εξωτερικό μαγνητικό πεδίο B_1 , εάν και μόνο αυτό περιστρέφεται με την συχνότητα Larmor του [35], [36], [37].

Για να γίνει πιο εύκολα κατανοητό το παρακάτω φαινόμενο, θα γίνει αλλαγή του σημείου αναφοράς, από στατικό, σε ένα το οποίο περιστρέφεται με συχνότητα ω .



Εικόνα 6: Αριστερά: Η εικόνα που έχει ένας εξωτερικός παρατηρητής, ο οποίος βλέπει το μαγνητικό πεδίο, καθώς και την περιστροφή του πυρήνα. Δεξιά: Η εικόνα που έχει ένας παρατηρητής που περιστρέφεται με την ίδια συχνότητα με τον πυρήνα. Επιπλέον παρατηρεί μηδενικό B_0 [39].

Με αυτόν τον τρόπο, το B_0 μπορεί να θεωρηθεί μηδενικό, καθώς είναι χρονοσταθερό, ενώ το διάνυσμα που εκφράζει την μαγνήτιση του κάθε πυρήνα, M , σταματάει να περιστρέφεται. Όπως προηγουμένως, η συνολική μαγνήτιση βρίσκεται επάνω στον άξονα z . Σε αυτό το στατικό σύστημα αναφοράς, γίνεται εφαρμογή ενός επιπλέον μαγνητικού πεδίου, B_1 [35]. Το διάνυσμα αυτό δεν είναι σταθερό, όπως το B_0 , αλλά περιστρέφεται με την συχνότητα Larmor του υδρογόνου. Συνεπώς, στο νέο σύστημα αναφοράς φαίνεται ακίνητο, όπως και οι πυρήνες. Είναι κάθετο στο αρχικό μαγνητικό πεδίο το B_0 βρίσκεται δηλαδή στο επίπεδο x - y [37]. Η εφαρμογή του έχει ως αποτέλεσμα την ευθυγράμμιση των x - y συνιστωσών των μαγνητικών πεδίων των πυρήνων με αυτό. Η συχνότητά του βρίσκεται στην συχνότητα των ραδιοκυμάτων και για αυτό συχνά αναγράφεται ως RF (radio frequency) παλμός [40].



Εικόνα 7: Εφαρμογή RF παλμού στην συχνότητα Larmor. Το αποτέλεσμα είναι η συγκέντρωση του συνολικού μαγνητικού πεδίου στον άξονα x-y [40]

Ο ευθύγραμμος των πυρήνων στο x-y επίπεδο έχει ως αποτέλεσμα την εμφάνιση ενός εγκάρσιου μαγνητικού πεδίου, το οποίο δεν παρατηρούνταν προηγουμένως, λόγω της αταξίας που επικρατούσε στο επίπεδο αυτό [35]. Όπως κάθε αλληλεπίδραση μαγνητικού πεδίου και δίπολου, το μαγνητικό πεδίο επιβάλλει στα μαγνητικά δίπολα να βρεθούν στο επίπεδό του, συγκεκριμένα στο x-y επίπεδο [37]. Αυτή η κατάσταση διαρκεί μόνο όσο παρέχουμε το πεδίο B_1 . Όταν σταματήσει να εφαρμόζεται το πεδίο αυτό, το πρωτόνια, δηλαδή οι πυρήνες του υδρογόνου, θα επιστρέψουν στην κατάσταση ισορροπίας που περιεγράφηκε προηγουμένως. Αυτή η διαδικασία δεν γίνεται στιγμιαία, με αποτέλεσμα το μέγεθος που παρατηρείται, δηλαδή η συνολική μαγνήτιση, να παρουσιάζει μία συνεχή συμπεριφορά. Κατά την διάρκεια αυτής της μετάβασης, μειώνεται η x-y συνιστώσα και αυξάνεται η z, μέχρι να φτάσει στην αρχική κατάσταση, όπου η πρώτη είναι μηδενική και η δεύτερη μέγιστη. Οι δύο αυτές μεταβάσεις καθορίζουν τους χρόνους διαμήκης χαλάρωσης T1 και εγκάρσιας χαλάρωσης T2 του ιστού [37]. Κάθε είδος ιστού έχει διαφορετικούς χρόνους χαλάρωσης, οι οποίοι καθορίζονται από τον αριθμό των πυρήνων υδρογόνου καθώς και την αλληλεπίδραση με το περιβάλλον τους [41].

2.3 Χωρικός εντοπισμός του σήματος

Η διαδικασία που προαναφέρθηκε, παρέχει πληροφορία σχετικά με το είδος του ιστού. Παρόλα αυτά, παραμένει ένα πολύ σημαντικό πρόβλημα. Δεν υπάρχει καμία πληροφορία για την τοποθεσία από την οποία προέρχεται το σήμα. Καθώς το μαγνητικό πεδίο υπάρχει παντού στον χώρο, ανά πάσα στιγμή, θα ληφθεί σήμα από κάθε σημείο του σώματος, καθιστώντας αδύνατο να εξαχθεί κάποια χρήσιμη πληροφορία. Για να αντιμετωπιστεί το παραπάνω πρόβλημα, αντί για την εφαρμογή ενός μαγνητικού πεδίου σε όλο το σώμα του ασθενή, γίνεται χρήση βαθμοτών, B_0 [39], [42], δηλαδή διαφορετικών σε μέγεθος μαγνητικών πεδίων. Αυτό έχει ως αποτέλεσμα, σε κάθε σημείο του χώρου, η συχνότητα Larmor είναι διαφορετική για τα πρωτόνια, λόγω του διαφορετικού B_0 . Κατ' αυτόν τον τρόπο, οι RF παλμοί, μπορούν να απορριφθούν μόνο από ένα συγκεκριμένο σημείο, το σημείο το οποίο έχει την συγκεκριμένη ιδιοσυχνότητα. Έτσι, όταν σταλεί ο παλμός και γίνει αισθητό το σήμα της χαλάρωσης, μπορεί να αποφανθεί το σημείο του χώρου από το οποίο προήλθε [41]. Το σημείο αυτό είναι το έχει ένταση μαγνητικού πεδίου ίση με την συχνότητα του RF παλμού, προς τον γυρομαγνητικό λόγο του υδρογόνου. Με την χρήση πολλών διαφορετικών παλμών, συλλέγεται πληροφορία για κάθε ένα από τα σημεία του σώματος. Οι χρόνοι χαλάρωσης, για κάθε ένα από τα σημεία συλλέγονται και απαρτίζουν τον κ-χώρο

2.4 Ανακατασκευή εικόνας

Έως τώρα έχει γίνει περιγραφή για την διαδικασία μέτρησης μαγνητικών πεδίων, από το σώμα του ασθενή. Για να μπορέσουν αυτές οι μετρήσεις να ερμηνευτούν, πρέπει πρώτα να μεταφραστούν σε μία μορφή, αναγνώσιμη από το ιατρικό προσωπικό, από μετρήσεις μεταβολών στην ένταση του μαγνητικού πεδίου, στην εικόνα της μαγνητικής τομογραφίας όπως πολύ πιθανό ο αναγνώστης να έχει συναντήσει. Η διαδικασία αυτή ονομάζεται ανακατασκευή εικόνας.

2.4.1 Κ-χώρος

Ο κ-χώρος είναι μία συλλογή από μιγαδικούς αριθμούς, στους οποίους το πραγματικό μέρος εκφράζει ένταση και το φανταστικό φάση. Κάθε σημείο αυτού του χώρου εξάγεται με τις εξής συναρτήσεις:

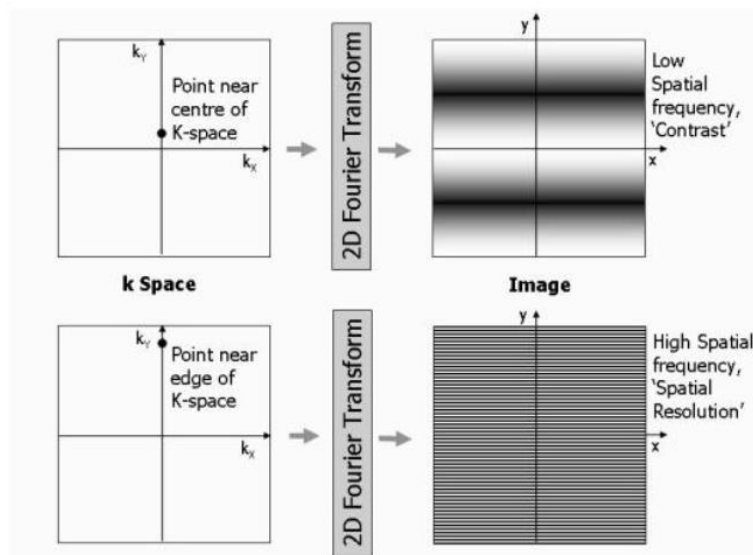
$$k_x(t') = \int_0^{t'} \gamma B_{1x}(t) dt$$

$$k_y(t') = \int_0^{t'} \gamma B_{1y}(t) dt$$

όπου

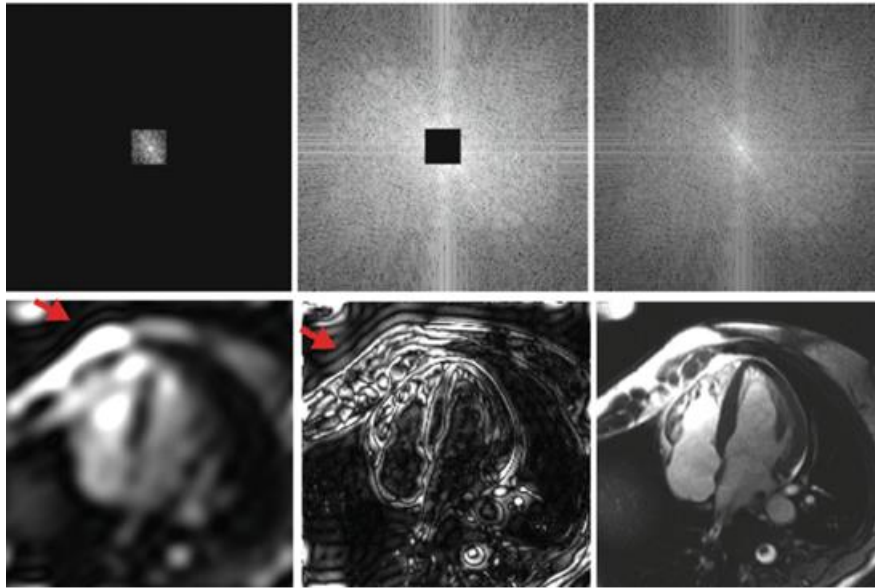
- B_1 η ένταση του RF παλμού
- t' ο χρόνος συλλογής των δεδομένων.

Η ισχύς και η διάρκεια των παλμών RF που εφαρμόζονται στους βιολογικούς ιστούς καθορίζουν την απόσταση και την κατεύθυνση του σημείου στον κ-χώρου από το κέντρο του [43]. Κάθε σημείο του κ-χώρου αντιστοιχεί σε μία χωρική συνιστώσα εναλλασσόμενης έντασης και διαφορετικής συχνότητας.



Εικόνα 8: Αντιστοίχιση ενός σημείου το κ-χώρου με την χωρική συχνότητα που του αντιστοιχεί στο πεδίο της τελικής εικόνας [44].

Λαμβάνοντας πολλά τέτοια σημεία και αθροίζοντας τα, είναι δυνατή η ανακατασκευή μίας εικόνας. Στο κέντρο του κ-χώρου βρίσκεται η πληροφορία για την μεγαλύτερη μάζα του ιστού, ενώ στις εξωτερικές συχνότητες, κωδικοποιούνται οι ακμές, δηλαδή το contrast μεταξύ δύο αντικειμένων που συνορεύουν.



Εικόνα 9: Διαφορετικές περιοχές του κ-χώρου προσθέτουν στο τελικό αποτέλεσμα διαφορετικό είδος πληροφορίας. Στην πρώτη στήλη φαίνεται ότι τα κεντρικά σημεία συμπεριλαμβάνουν την γενική δομή του ιστού, αλλά υστερούν στις λεπτομερείς. Αντίθετα, τα συνοριακά σημεία αποθανατίζουν την αντίθεση αλλά όχι την μάζα του ιστού. Ταυτόχρονα και οι δυο ελλείψεις εικόνες παρουσιάζουν μοτίβο κυματισμού, τα οποία επισημαίνονται με ένα κόκκινο βέλος [45].

Όπως φαίνεται παραπάνω, όσο περισσότερα σημεία του κ-χώρου συμπεριληφθούν, τόσο καλύτερη ανάλυση εικόνας επιτυγχάνεται. Δυστυχώς, για κάθε στήλη του χώρο, θα χρειαστεί να γίνει εκπομπή ενός παλμού, γεγονός το οποίο καθιστά την διαδικασία πολύ χρονοβόρα [43].

2.4.2 Compressed Sensing

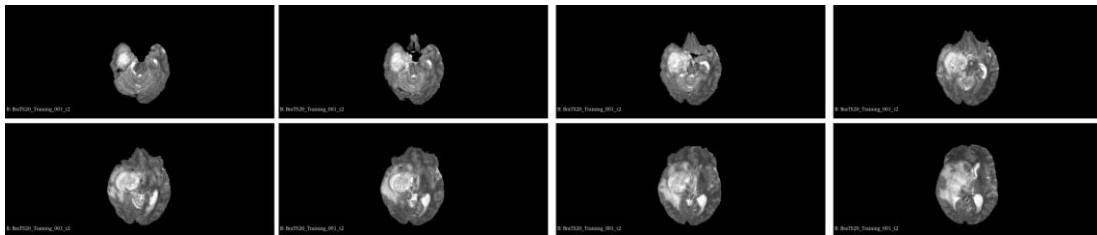
Η πιο γνωστή από τις διαδικασίες ανακατασκευής ονομάζεται Compressed sensing. Βασίζεται στην υπόθεση ότι οι εικόνες που συλλέγονται δεν είναι μία ακολουθία από τυχαία pixel, ή, στην περίπτωσή μας, τυχαίες συνιστώσες του κ-χώρου, αλλά υπάρχει ένα είδος δομής και τάξης. Από μαθηματικής άποψης συνεπάγεται ότι συλλέγεται πλεονάζουσα πληροφορία. Επιλέγοντας να μετρηθούν συγκεκριμένα σημεία του κ-χώρου, δημιουργείται ένας αραιός πίνακας y , ένας πίνακας δηλαδή που σε ένα, συνήθως μεγάλο ποσοστό των θέσεων του έχει μηδενική τιμή. Από αυτές τις μετρήσεις αυτές, μπορεί να παραχθεί μία εικόνα, προφανώς χαμηλότερης ποιότητας από αυτή την επιθυμητή. Στην συνέχεια, εφαρμόζεται ένας αλγόριθμος αφαίρεσης θορύβου [45]. Αυτό έχει ως αποτέλεσμα, πέρα από τον θόρυβο, να χάνεται και κάποια χρήσιμη πληροφορία. Για να την επαναφορά της, γίνεται σύγκριση του κ-χώρου συχνοτήτων που παράχθηκε, με τον κ-χώρο που μετρήθηκε.

Προφανώς, οι μετρήσεις που έγιναν, αν και περιορισμένες στο πλήθος τους, αντιπροσωπεύουν την πραγματικότητα. Μπορεί λοιπόν να επαναφερθεί την

πληροφορία που χάθηκε μέσω του αλγόριθμο αφαίρεσης θορύβου, χωρίς όμως να επιβαρυνθεί ο κ-χώρος με τον θόρυβο της αρχικής μέτρησης. Αυτή η διαδικασία μπορεί να γίνει επαναληπτικά, μέχρι να επιτευχθεί το επιθυμητό αποτέλεσμα. Το πρόβλημα αυτής της λύσης είναι ότι ο αλγόριθμος, καθώς υλοποιεί μετασχηματισμούς από και προς τον κ-χώρο, τον αλγόριθμο αφαίρεσης θορύβου και συγκρίσεις των αποτελεσμάτων τους κατ' επανάληψη, αποτελεί χρονοβόρα διαδικασία.

2.5 Μαγνητική Τομογραφία και Διάγνωση

Όπως αναφέρθηκε στο Κεφάλαιο 1, για την διάγνωση ασθενειών όπως ο καρκίνος του εγκεφάλου, συχνά γίνεται χρήση μαγνητικής τομογραφίας. Η διαδικασία αυτή συμπεριλαμβάνει την λήψη μαγνητικής τομογραφίας από έναν ασθενή. Καθώς η αναπαράσταση τρισδιάστατων όγκων είναι δύσκολη, σε μία δισδιάστατη οθόνη, συνήθως η μαγνητικές τομογραφίες απεικονίζονται σαν διαδοχικές τομές όπως φαίνεται στη παρακάτω εικόνα.



Εικόνα 10: Διαδοχικές τομές από Μαγνητική Τομογραφία Εγκεφάλου.

Έπειτα ο γιατρός καλείτε να επιβλέψει την εικόνα και, εάν θεωρήσει ότι κάποιες από τις περιοχές που βλέπει περιέχουν καρκινικά κύτταρα, να τα αναδείξει, κάνοντας χρήση εξειδικευμένου λογισμικού όπως το 3D-Slicer [46].



Εικόνα 11: Αριστερά: Τομή μαγνητικής τομογραφίας Εγκεφάλου Δεξιά: Τμηματοποίηση εικόνας που αναδεικνύει τις προσβεβλημένες περιοχές, από ειδικούς νευροακτινολόγους.

Όπως είναι κατανοητό, η διαδικασία αυτή καταλαμβάνει πολύ από τον χρόνο του γιατρού, που πρέπει να διαλέξει κάθε ένα από τα Pixel που θεωρεί προσβεβλημένα, σε κάθε τομή του ασθενή. Οι παραπάνω εικόνες ανήκουν στην ανοιχτή βάση δεδομένων BraTS 2020 [32].

2.5.1 BraTS 2020

Τα δεδομένα αυτά παρέχονται από το πανεπιστήμιο της Πενσυλβάνιας και συμπεριλαμβάνει δεδομένα από 369 ασθενείς. Οι εικόνες, έχουν διαστάσεις 240x240x155. Για κάθε ασθενή, περιέχονται 4 διαφορετικοί τρόποι απεικόνισης (modalities).

- T1
- T2
- T1 contrast Enhanced
- T2 Fluid Attenuated Inversion Recovery

Όπου οι T1 και T2 περιεγράφηκαν παραπάνω, η T1 contrast Enhanced, αποτελεί μία επεξεργασία της T1, όπου μετά την λήψη, έχει τονωθεί η αντίθεση μεταξύ των χρωμάτων

Σε κάθε σημείο του έχει ανατεθεί ένας από τους εξής χαρακτηρισμούς.

- Φυσιολογικά κύτταρα
- Ενισχυμένο τμήμα του όγκου
- Περιφερικό οίδημα
- Νεκρωτικός πυρήνας

Οι τέσσερις αυτή χαρακτηρισμοί φαίνονται με διαφορετικό χρώμα, όπως φαίνεται στην **Εικόνα 11**. Όπως είναι εμφανές, η διαδικασία χαρακτηρισμού κάθε σημεία κάθε τομής από έναν ή περισσότερους [7] ιατρούς αποτελεί μία χρονοβόρα και ακριβή διαδικασία, στην οποία ο ρόλος της τεχνολογίας, της τεχνίτης νοημοσύνης και των νευρωνικών δικτύων είναι κομβικός.

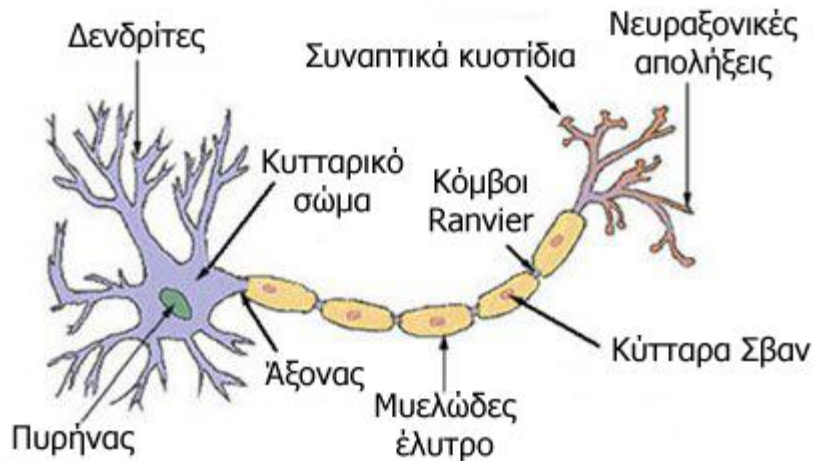
Κεφάλαιο 3: Νευρωνικά Δίκτυα

3.1 Γενικά

Ο κόσμος στον οποίο ζούμε, μπορεί πολλές φορές να μοιάζει χαοτικός και απρόβλεπτος, σχεδόν τυχαίος. Παρόλα αυτά, στον δειγματικό χώρο, όλων των ενδεχομένων που μπορούν να συμβούν, αυτά που παρατηρούνται είναι ένα ελάχιστο υποσύνολο. Στην φύση, όση ποικιλομορφία και να παρατηρείται, στην πραγματικότητα, υπακούσετε μία πολύ αυστηρή δομή. Εάν ένας υπολογιστή να παράγει τυχαίες εικόνες, η πιθανότητα να δημιουργήσει κάτι που να θυμίζει κάτι που συναντάται στην φύση είναι ελάχιστη και ρεαλιστικά μηδενική. Αυτός ο προβληματισμός είναι γνωστός ως το θεώρημα των άπειρων μαϊμούδων [47]. Όλα τα παραπάνω προδίδουν ότι στην φύση υπάρχει σε ένα πολύ μεγάλο βαθμό ένα μοτίβο, μία τάξη. Εάν το μοτίβο εξαχθεί από μία διαδικασία, τότε μπορεί να χρησιμοποιηθεί για να πραγματοποιηθούν προβλέψεις για το μέλλον. Αυτό το μοτίβο προσπαθούν να εξάγουν οι τεχνικές μηχανικής μάθησης. Η εργασία αυτή ασχολείται με μία υποκατηγορία του ευρύ αυτού αντικειμένου, τα νευρωνικά δίκτυα.

3.2 Βιολογικοί νευρώνες

Η έμπνευση για αυτή την τεχνική ήρθε από τον ίδιο τον ανθρώπινο εγκέφαλο. Ο εγκέφαλος, όπως και ολόκληρο το νευρικό μας σύστημα έχει ως δομική μονάδα το νευρικό κύτταρο, ή αλλιώς νευρώνα. Όπως κάθε άλλο κύτταρο, διαθέτει το κυτταρικό σώμα που περιέχει τις βασικές δομές, όπως τα μιτοχόνδρια, λυσοσώματα και φυσικά, τον πυρήνα. Η διαφοροποίηση του με τα υπόλοιπα κύτταρα, είναι ότι διαθέτει έναν ακόμα άξονα, τον νευράξονα, καθώς και αποφυάδες, δηλαδή ινοειδή δομές που εξέρχονται από το κύτταρο, τους δένδριτες και τις νευρικές απολήξεις [48].



Εικόνα 12: Δομή νευρώνα. Στο σχήμα φαίνονται τα βασικά δομικά χαρακτηριστικά του [49].

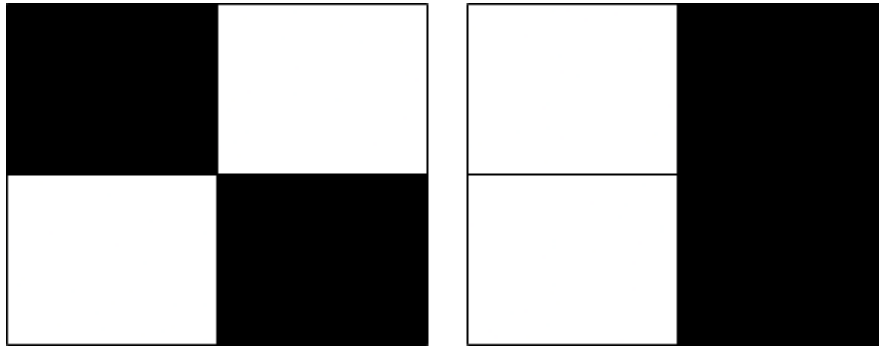
Η λειτουργία τους είναι να μεταφέρουν σήματα από ένα σημείο του σώματος σε ένα άλλο. Οι νευρικές απολήξεις το ενός νευρώνα συνδέονται με τους δενδρίτες του επόμενου. Όταν φτάσει η νευρική ώση, δηλαδή η διαφορά δυναμικού στο πρώτο νευρώνα σε ένα ικανοποιητικό κατώφλι, ελευθερώνονται νευροδιαβιβαστές. Οι νευροδιαβιβαστές είναι χημικές ουσίες, οι οποίες σηματοδοτούν το επόμενο νεύρο να παράξει αυτό την δικιά του νευρική ώση, η οποία μεταφέρεται μέσω του άξονα στις νευρικές απολήξεις αυτού του νεύρου και να επαναληφθεί αυτή η διαδικασία μέχρι να φτάσει το σήμα στον στόχο του [50]. Όταν μία σύναψη ενεργοποιηθεί, γίνεται ελάχιστα πιο ισχυρή [51]. Αυτό έχει ως αποτέλεσμα την δημιουργία συγκεκριμένων μονοπατιών στον εγκέφαλό μας, που στην ουσία είναι η διαδικασία της μάθησης από την επανάληψη.

3.3 Διαισθητική ανάλυση

Οι βασικές έννοιες που δομούν τα νευρωνικά δίκτυα είναι εκπληκτικά απλές. Παρόλα αυτά, ίσως λόγω της νεαρής ηλικίας του αντικειμένου, ίσως λόγω των εξαιρετικών αποτελεσμάτων του, υπάρχει η λανθασμένη αντίληψη ότι είναι απαραίτητη γνώση περίπλοκων μαθηματικών, κάτι το οποίο δεν ισχύει. Πολλές φορές παρουσιάζονται δυσνόητη μαθηματικοί τύποι και η κατανόηση του πως προέκυψαν ή σε τι χρησιμεύει κάθε σύμβολο και πράξη που παρουσιάζεται παραμένει μυστήρια. Παρακάτω θα παρουσιαστεί ένας συλλογισμός, ξεκινώντας από ένα εξαιρετικά απλό παράδειγμα, μέσα από του οποίου θα δομηθούν και θα εξηγηθούν όλες οι δομικές έννοιες, απαραίτητες για την κατανόηση τους.

3.2.1 Μοντελοποίηση προβλήματος

Το πρόβλημα που θα λυθεί έχει ως εξής. Υπάρχουν δύο ειδών σχήματα, για τα οποία χρειάζεται να αναπτυχθεί ένας αλγόριθμος διαχωρισμού τους. Τα σχήματα αυτά φαίνονται στην παρακάτω εικόνα.



Εικόνα 13: Εικόνες ανάλυσης 2x2. Στο παρακάτω παράδειγμα, θα δομηθεί ένα νευρωνικό δίκτυο όπου θα αναλάβει να τις ξεχωρίζει μεταξύ τους.

Πρώτο βήμα για να λυθεί αλγοριθμικά ένα τέτοια πρόβλημα, το οποίο δίνεται περιγραφικά, είναι να διατυπωθεί σε νούμερα. Για τον σκοπό αυτό, ορίζονται οι 2x2 πίνακες που αντιπροσωπεύουν τις εικόνες

$$X_a = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, X_b = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$$

Με 1 να συμβολίζει το χρώμα μαύρο και 0 το άσπρο. Ορίζονται επίσης τα διανύσματα

$$Y_a = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, Y_b = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Με αυτόν τον τρόπο, ένα περιγραφικό πρόβλημα, που παρατέθηκε με λόγια, έχει μετατραπεί σε ένα μαθηματικό πρόβλημα. Χρειάζεται να βρεθεί μία διαδικασία, μία συνάρτηση $f: \mathbb{R}^{2 \times 2} \rightarrow \mathbb{R}^{2 \times 1}$

$$f(X_a) = Y_a, f(X_b) = Y_b$$

Λόγω διαστάσεων των πινάκων και μόνο, ένας μαθηματικός θα μπορούσε να προτείνει, η συνάρτηση f να έχει την μορφή πολλαπλασιασμού με πίνακα διαστάσεων 2x1, W .

$$f(X) = X * W, W \in \mathbb{R}^{2 \times 1}$$

Σε αυτό το σημείο θα οριστεί ένας φορμαλισμός που θα διατηρηθεί καθόλη την της παρούσας εργασίας.

- Με κεφαλαία γράμματα X, Y, W γράφονται οι πίνακες ως σύνολο

- Εάν οι πίνακες αυτό έχουν δείκτης γράμμα, X_a , αναφέρονται σε ένα συγκεκριμένο δείγμα, για παράδειγμα την εικόνα a .
- Με πεζά γράμματα και δύο δείκτες, x_{ij} συμβολίζεται το στοιχείο του πίνακα, X , που βρίσκεται στην γραμμή i και στήλη j . Όταν γίνεται αναφορά σε διανύσματα, συναντάται μόνο ένας δείκτης.

Ο φορμαλισμός αυτός θα επεκταθεί στην συνέχεια, ανάλογα με τις ανάγκες διατύπωσης περίπλοκων.

Για να βρεθεί ο πίνακας που ικανοποιεί το δείγμα a , λύνουμε την παρακάτω εξίσωση

$$f(X_a) = Y_a \leftrightarrow X_a * W = Y_a \leftrightarrow \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} * \begin{bmatrix} w_0 \\ w_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \leftrightarrow$$

$$w_0 = 1, w_1 = 0$$

Ο συνδυασμός αυτός όμως, παράγει λανθασμένα αποτελέσματα για το δεύτερο δείγμα

$$f(X_b) = X_b * W = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} * \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \neq Y_b$$

Οι διαστάσεις των πινάκων που επιλέχθηκαν περιορίζουν τον υπολογισμό του πρώτου στοιχείου εξόδου, y_0 να είναι συνάρτηση μόνο των στοιχείων εισόδου x_{00}, x_{01} , και αντίστοιχα για y_1 . Για να εξαλειφθεί ο παραπάνω περιορισμός, εκφράζουμε τους πίνακες εισόδου σε μονοδιάστατη μορφή, με

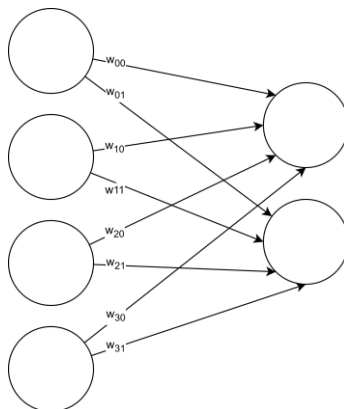
$$X_a = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, X_b = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

Πλέον η συνάρτηση μας αλλάζει διαστάσεις $f: \mathbb{R}^{4 \times 1} \rightarrow \mathbb{R}^{2 \times 1}$. Για χάρην ευκολίας, οι πίνακες θα με τους ανάστροφους τους θα χρησιμοποιούνται ως συνώνυμοι, δηλαδή θα στην θέση του X_a θα γίνεται χρήση του ανάστροφου του X_a^T , χωρίς να αναφέρεται ρητά. $X_a \leftrightarrow X_a^T$. Ένας αναγνώστης έστω και λίγο εξοικειωμένος με πράξεις πινάκων είναι ικανός να αναγνωρίσει ότι το αποτέλεσμα μένει αναλλοίωτο.

Ο πίνακας W έχει πλέον την μορφή

$$W = \begin{bmatrix} w_{00} & w_{01} \\ w_{10} & w_{11} \\ w_{20} & w_{21} \\ w_{30} & w_{31} \end{bmatrix}$$

Ο παραπάνω πίνακας, οπτικοποιείται ως εξής. Τα τέσσερα στοιχεία εισόδου πολλαπλασιάζονται με τα αντίστοιχα βάρη του πίνακα W και παράγουν το διάνυσμα εξόδου Y .



Εικόνα 14: Νευρωνικό Δίκτυο τεσσάρων εισόδων και 2 εξόδων

Για τα δείγματα a και b

$$f(X_a) = X_a * W = [1 \ 0 \ 0 \ 1] * \begin{bmatrix} w_{00} & w_{01} \\ w_{10} & w_{11} \\ w_{20} & w_{21} \\ w_{30} & w_{31} \end{bmatrix} = \begin{bmatrix} w_{00} + w_{30} \\ w_{01} + w_{31} \end{bmatrix}$$

$$f(X_b) = X_b * W = [0 \ 1 \ 0 \ 1] * \begin{bmatrix} w_{00} & w_{01} \\ w_{10} & w_{11} \\ w_{20} & w_{21} \\ w_{30} & w_{31} \end{bmatrix} = \begin{bmatrix} w_{10} + w_{30} \\ w_{11} + w_{31} \end{bmatrix}$$

Για να ικανοποιηθούν οι συνθήκες, πρέπει

$$\begin{bmatrix} w_{00} + w_{30} \\ w_{01} + w_{31} \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

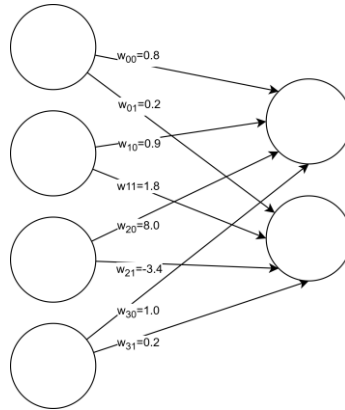
$$\begin{bmatrix} w_{10} + w_{30} \\ w_{11} + w_{31} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Είναι ένα σύστημα με τέσσερις (4) εξισώσεις και έξι (6) μεταβλητές, $w_{00}, w_{30}, w_{01}, w_{31}, w_{10}, w_{11}$, οπότε οι λύσεις είναι άπειρες και πολύ εύκολο να βρεθούν.

3.2.2 Backpropagation

Η παραπάνω μέθοδος, ανάγει ένα οποιοδήποτε πρόβλημα, το οποίο μπορεί να εκφραστεί με την μορφή πινάκων εισόδου εξόδου, σε ένα γραμμικό σύστημα εξισώσεων. Παρόλα αυτά, όταν η διαστάσεις αλλά και το πλήθος των δειγμάτων

εισόδου αυξηθούν, η εύρεση των ιδανικών βαρών γίνεται αδύνατη. Για τον λόγο αυτό, εφαρμόζεται ο ευριστικός αλγόριθμος backpropagation [52]. Αρχικά οι τιμές των βαρών αρχικοποιούνται με τυχαίες τιμές.



Εικόνα 15: Νευρωνικό Δίκτυο τεσσάρων εισόδων και 2 εξόδων με τυχαία αρχικοποιημένες τιμές των βαρών

Προφανώς το αποτέλεσμα του νευρωνικού δικτύου είναι τυχαίο. Για το δείγμα a του παραδείγματός μας

$$f(X_a) = X_a * W = [1 \ 0 \ 0 \ 1] * \begin{bmatrix} w_{00} & w_{01} \\ w_{10} & w_{11} \\ w_{20} & w_{21} \\ w_{30} & w_{31} \end{bmatrix} = \begin{bmatrix} w_{00} + w_{30} \\ w_{01} + w_{31} \end{bmatrix} = \begin{bmatrix} 1.8 \\ 0.4 \end{bmatrix}$$

Με βάση την παραπάνω, τυχαία έξοδο, μπορεί να υπολογιστεί το πόσο κοντά ήταν η απάντηση του δικτύου, για το συγκεκριμένο δείγμα, κάνοντας χρήση μίας συνάρτησης κόστους. Στο παράδειγμά θα χρησιμοποιηθεί μία από τις συχνότερες συναρτήσεις, mean squared error [53].

$$MSE(y, label) = \frac{1}{2} * \sum (y - label)^2$$

- y η έξοδος του δικτύου μας
- $label$ η επιθυμητή έξοδος

Για την είσοδο X_a

$$Loss = \frac{1}{2} * \sum (y - label)^2 = \frac{1}{2} * \sum \left(\begin{bmatrix} 1.8 \\ 0.4 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right)^2 = \frac{1}{2} * \sum \left(\begin{bmatrix} 0.8 \\ 0.4 \end{bmatrix} \right)^2 = \frac{1}{2} * (0.8^2 + 0.4^2) = 0.4$$

Εάν η συνάρτηση κόστους μηδενιστεί, σημαίνει ότι το κάθε δείγμα έχει την επιθυμητή είσοδο. Αφού λοιπόν το πρόβλημα έχει εκφραστεί πλήρως με μαθηματικούς όρους, μπορεί να υπολογιστεί η παράγωγος της συνάρτησης κόστους ως προς κάθε βάρος.

$$\frac{dLoss}{dw_{00}} = \frac{d \frac{1}{2} \sum (y - label)^2}{dw_{00}} = \frac{1}{2} * \frac{d \sum (x^T * w - label)^2}{dw_{00}}$$

Όπου

$$\begin{aligned} \sum (x^T * w - label)^2 &= \sum ([1 \ 0 \ 0 \ 1] * \begin{bmatrix} w_{00} & w_{01} \\ w_{10} & w_{11} \\ w_{20} & w_{21} \\ w_{30} & w_{31} \end{bmatrix} - label)^2 = \\ \sum \left(\begin{bmatrix} w_{00} + w_{30} \\ w_{01} + w_{31} \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right)^2 &= (w_{00} + (w_{30} - 1))^2 + (w_{01} + w_{31})^2 = \\ &w_{00}^2 + 2 * w_{00} * (w_{30} - 1) + (w_{30} - 1)^2 + (w_{01} + w_{31})^2 \end{aligned}$$

Άρα

$$\begin{aligned} \frac{dLoss}{dw_{00}} &= \frac{1}{2} * \frac{d}{dw_{00}} [w_{00}^2 + 2 * w_{00} * (w_{30} - 1) + (w_{01} + w_{31})^2] = \\ &\frac{1}{2} * (2 * w_{00} + 2 * (w_{30} - 1)) \end{aligned}$$

$$\frac{dLoss}{dw_{00}} = 0.8$$

Από τους παραπάνω υπολογισμούς, μπορεί να αποφανθεί ότι η αύξηση του βάρους w_{00} θα επιφέρει την αύξηση της συνάρτησης κόστους. Για να ελαχιστοποιηθεί το κόστος, θα χρειαστεί το βάρος να μειωθεί.

Εάν υπολογιστεί η έξοδος είναι

$$\begin{aligned} f(X_a) &= [1 \ 0 \ 0 \ 1] * \begin{bmatrix} w_{00} - \frac{dLoss}{dw_{00}} & w_{01} \\ w_{10} & w_{11} \\ w_{20} & w_{21} \\ w_{30} & w_{31} \end{bmatrix} = \\ \begin{bmatrix} w_{00} - \frac{dLoss}{dw_{00}} + w_{30} \\ w_{01} + w_{31} \end{bmatrix} &= \begin{bmatrix} 0.8 - 0.8 + 1 \\ 0.2 + 0.2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0.4 \end{bmatrix} \end{aligned}$$

Πράγματι, η έξοδος είναι πιο κοντά στο label, ενώ και το κόστος, όπως ήταν αναμενόμενο, μειώθηκε.

$$Loss = \frac{1}{2} * \sum (y - label)^2 = \frac{1}{2} * \sum ([0.4] - [1])^2 = 0.08$$

Η διαδικασία αυτή επαναλαμβάνεται για όλα τα βάρη, για όλα τα δείγματα.

3.3 Φορμαλιστική ανάλυση

Η παραπάνω ανάλυση, εάν και διαισθητικά ορθή, δεν αποτελεί τον συστηματικό τρόπο που τα νευρωνικά δίκτυα απαντώνται στην βιβλιογραφία. Για λόγους συμβατότητας και συνέπειας, ακολουθεί μία αυστηρότερη περιγραφή, η οποία όμως συμπεριλαμβάνει τις ίδιες ακριβώς έννοιες.

3.3.1 Μοντέλο νευρώνα

Ο τεχνητός νευρώνας, ή αλλιώς perceptron, δεν είναι τίποτα παραπάνω από ένα άθροισμα παραμέτρων και την εφαρμογή μίας συνάρτησης ενεργοποίησης. Ο κάθε νευρώνας δέχεται ως είσοδο την ενεργοποίηση των προηγούμενων, πολλαπλασιασμένη με ένα ξεχωριστό για τον καθέναν βάρος. Στην συνέχεια, προστίθεται μία σταθερή τιμή, το bias. Τέλος εφαρμόζεται μία συνάρτηση ενεργοποίησης. Η έξοδος αυτή μπορεί να τροφοδοτηθεί στους επόμενους νευρώνες, ή να αποτελέσει μίας από τις εξόδους του μοντέλου.

$$a_j = \sigma(\sum w_{ij} * x_i + b_j)[54]$$

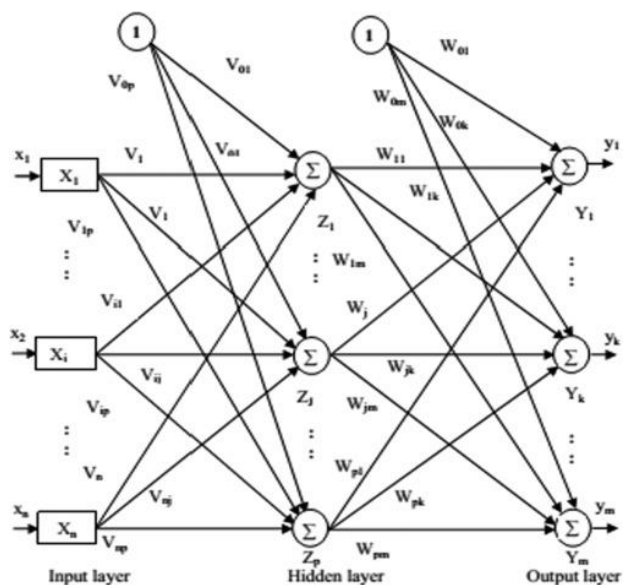
Όπου

- a_j είναι η έξοδος του νευρώνα j
- x_i η τιμή του νευρώνα i, που ανήκει στο προηγούμενο επίπεδο
- w_{ij} , το βάρος που ενώνει τον νευρώνα i του προηγούμενου επιπέδου, με τον νευρώνα j του επόμενου
- b_j το bias του νευρώνα j
- σ , η συνάρτηση ενεργοποίησης.

Η συνάρτηση ενεργοποίησης σ επιλέγεται συνήθως να είναι μία μη γραμμική συνάρτηση. Τα περισσότερα προβλήματα που απαντώνται στην φύση και χρειάζονται την χρήση νευρωνικών δικτύων είναι μη γραμμικά [55]. Τέλος το Bias μπορεί να εκφραστεί ως ένα επιπλέον στοιχείο στους νευρώνες εισόδου, με σταθερή τιμή 1 και βάρος την τιμή b

3.3.2 Perceptron Πολλών Επιπέδων

Για να λυθούν προβλήματα τα οποία, πέρα από μη γραμμικά, έχουν και συνδυαστική σχέση μεταξύ τους, θα χρειαστεί μία πιο περίπλοκη δομή. Η δομή αυτή είναι τα νευρωνικά δίκτυα πολλών επιπέδων. Όπως δηλώνει το όνομα, δεν πρόκειται για ένα δίκτυο, που συμπεριλαμβάνει πολλούς νευρώνες. Ο κάθε ένας ξεχωριστά λειτουργεί όπως περιγράψαμε προηγουμένως.



Εικόνα 16: Νευρωνικό Δίκτυο Πολλών επιπέδων [56]

Στο σύνολο η διαδικασία μπορεί να περιγράψει ως ένας πολλαπλασιασμός πινάκων [57], όπου η είσοδος είναι ένα διάνυσμα $1 \times n$, που πολλαπλασιάζεται με τον πίνακα βαρών $n \times m$ και παράγει την έξοδο των επόμενων νευρώνων $1 \times m$. Τα δεδομένα, αφού έχουν εκφραστεί σε μορφή διανύσματος, εισέρχονται στο δίκτυο από τα αριστερά, ενώ τα αποτελέσματα παράγονται στα δεξιά. Μία από τις πολύ γνωστές και φιλικές προς νέους αναγνώστες εφαρμογές είναι η αναγνώριση χειρόγραφων ψηφίων, βασιζόμενοι στην βάση δεδομένων MNIST [58]. Η βάση αυτή περιέχει συνολικά 70000 φωτογραφίες με διαστάσεις 28×28 .



Εικόνα 17: Εικόνες χειρόγραφων ψηφίων που ανήκουν στην βάση δεδομένων MNIST [59]

Κάθε φωτογραφία έχει ανάλυση 28x28 και συνοδεύεται από τον αριθμό που είναι γραμμένος σε αυτήν. Για να αναγνωριστεί ποιο ψηφίο είναι γραμμένο σε κάθε φωτογραφία, θα δημιουργηθεί ένα νευρωνικό δίκτυο με $28 \times 27 = 784$ νευρώνες εισόδου, δηλαδή ένας πίνακας 1×785 (μαζί με το bias), που ο κάθε ένας θα έχει την τιμή του αντίστοιχου pixel (και 1 το bias). Στην συνέχεια, θα γίνει επιλογή, από τον προγραμματιστή ένας ή παραπάνω ακέραιοι αριθμοί, οι οποίοι θα περιγράφουν τον αριθμό των νευρώνων στα ενδιάμεσα επίπεδα. Για απλότητα, θα γίνει χρήση ενός μόνο κρυμμένου επιπέδου, με 1000 νέρωνες. Το επίπεδο αυτό εκφράζεται ως ένας πίνακας 785×1000 , που όταν πολλαπλασιαστεί με την είσοδο, θα παράξει ένα διάνυσμα 1×1000 . Το τελευταίο επίπεδο, το οποίο παράγει το αποτέλεσμα, έχει 10 νευρώνες, έναν για κάθε πιθανό ψηφίο που μπορεί να αναπαρίσταται στην εικόνα. Ο πίνακας βαρών που το συνδέει με το προηγούμενο επίπεδο έχει διαστάσει 1000×10 . Στο τέλος αυτής της διαδικασίας θα προκύψει ένα διάνυσμα με 10 τιμές. Οι τιμές αυτές εκφράζουν το πόσο σίγουρο είναι το μοντέλο μας για το ποιος αριθμός απεικονίζεται στην εικόνα. Ως τελική απάντηση του μοντέλου, θα θεωρηθεί η θέση στην οποία το διάνυσμα εξόδου λαμβάνει την μέγιστή του τιμή. Εάν δηλαδή το μοντέλο δώσει ως έξοδο το διάνυσμα

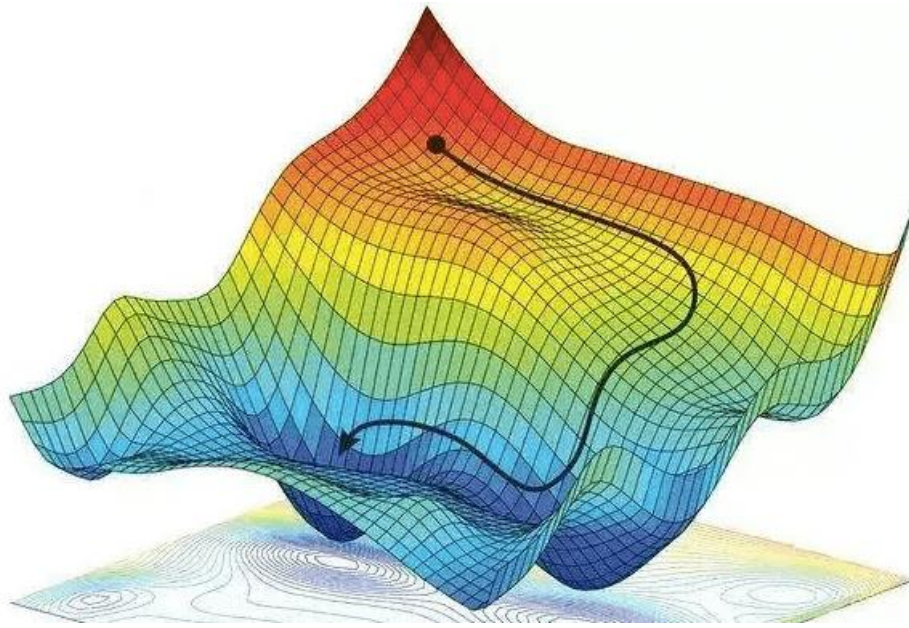
$$y = [2 \ 2.5 \ 31 \ 2 \ 46 \ -0.3 \ 3 \ 91 \ 0.2 \ 2]$$

Η τελική απάντηση του μοντέλου θα είναι ο αριθμός 8 καθώς η μέγιστη τιμή του διανύσματος βρίσκεται στην όγδοη θέση, αρχίζοντας την αρίθμηση από το μηδέν. Εάν η θέση της μέγιστης τιμής συμπίπτει με την πραγματική τιμή, τότε το μοντέλο έχει απαντήσει σωστά για το συγκεκριμένο δείγμα. Μία εικόνα δεν μπορεί να συμβολίζει δύο αριθμούς ταυτόχρονα. Έτσι, η έξοδος ενός σωστού μοντέλου, πέρα από το γεγονός ότι η θέση της μέγιστης τιμής του θα πρέπει να συμπίπτει με την πραγματική, θα πρέπει να έχει μηδενικές τιμές σε όλες τις άλλες. Όσο πιο μακριά βρίσκεται η έξοδος από την παραπάνω περιγραφή από αυτές τις τιμές, τόσο μεγαλύτερο σφάλμα έχουμε. Για να

επιτευχθεί το αποτέλεσμα αυτό, χρειάζεται οι πίνακες των βαρών του νευρωνικού δικτύου να έχουν τις κατάλληλες τιμές. . Η Προφανώς οι ιδανικές τιμές δεν είναι γνωστές εκ των προτέρων αλλά είναι το ζητούμενο της όλης διαδικασίας και προσεγγίζονται με την αλγόριθμο backpropagation.

3.3.3 Gradient Decent

Εάν θεωρηθεί μία συνάρτηση που δέχεται ως παραμέτρους τα βάρη του μοντέλου μας και παράγει ως έξοδο το συνολικό σφάλμα στα δεδομένα μας, είναι εύκολο να αντιληφθεί κανείς ότι εάν βρεθεί το ελάχιστο της παραπάνω συνάρτησης, δηλαδή τις συντεταγμένες όπου η παράγωγος μηδενίζεται, το ιδανικό μοντέλο θα έχει βρεθεί [60]. Δυστυχώς, καθώς τα βάρη είναι πίνακες πολύ μεγάλων διαστάσεων, η υπολογιστική εύρεση του σημείου αυτού είναι σχεδόν αδύνατη. Για αυτόν τον λόγο, εφαρμόζεται η μέθοδος της κλίσης κατάβασης, ή αλλιώς Gradient Decent [61]. Ο αλγόριθμος έχει ως εξής. Αρχικοποιούνται τυχαία οι πίνακες βαρών του μοντέλου. Υπολογίζεται η κλίση της συνάρτησης στο συγκεκριμένο σημείο και πραγματοποιείται ένα μικρό βήμα προς την κατεύθυνση που το ελαχιστοποιεί. Το πόσο μικρό θα είναι αυτό το βήμα ορίζεται από μία υπερπάρμετρο, που καθορίζεται από τον χρήστη, το learning rate. Εάν επαναληφθεί η διαδικασία αυτή αρκετές φορές, η παραπάνω διαδικασία θα βρεθεί σε ένα ελάχιστο. Υπάρχει πάντα ο κίνδυνος το ελάχιστο αυτό να μην είναι ολικό, αλλά τοπικό. Σε αυτήν την περίπτωση, όλες οι κατευθύνσεις αυξάνουν το σφάλμα, οπότε ο αλγόριθμος προτιμά να παραμείνει σε αυτό το σημείο.



Εικόνα 18: Γραφική αναπαράσταση του αλγορίθμου gradient descend [62]

3.3.4 Backpropagation

Ο αλγόριθμος εύρεσης της κατάλληλης κατεύθυνσης μεταβολής είναι ο αλγόριθμος backpropagation. Χρησιμοποιεί το σφάλμα που προαναφέρθηκε, για να υπολογίσει το ποια βάρη πρέπει να μεταβληθούν και πόσο, έτσι ώστε να φτάσουμε στο ελάχιστο της συνάρτησης. Για τους νευρώνες εξόδου, το σφάλμα υπολογίζεται πολύ απλά, από την διαφορά της πρόβλεψής μας με το πραγματικό αποτέλεσμα, κάνοντας χρήση κάποιας μετρικής, όπως το mean squared error που έγινε χρήση στο κεφάλαιο της διαισθητικής ανάλυσης. Στην συνέχεια, για να μετακινηθούν οι προβλέψεις προς την επιθυμητή κατεύθυνση, θα έπρεπε να μεταβληθούν τα βάρη που συνδέουν τους νευρώνες εξόδου με το προηγούμενο επίπεδο, δεδομένου τις ενεργοποιήσεις τους. Επιπλέον υπολογίζεται, πως θα έπρεπε να μεταβληθούν οι τιμές των νευρώνων του προηγούμενου επιπέδου, με δεδομένα τα τωρινά βάρη

Για να διατηρηθεί η συνοχή με την πηγή [63], από όπου αντλήθηκαν οι παρακάτω εξισώσεις θα ορίσουμε τους συμβολισμούς ως εξής.

- a_j^l Η ενεργοποίηση του νευρώνα j στο επίπεδο l
- w_{jk}^l Το βάρος που ενώνει τον νευρώνα j του επιπέδου l με τον νευρώνα k του προηγούμενου επιπέδου
- b_j^l Το bias του νευρώνα j του επιπέδου l
- σ Η συνάρτηση ενεργοποίησης

- C Το κόστος

Σκοπός είναι ο υπολογισμός τις μερικές παραγώγους της συνάρτησης κόστους ως προς τα βάρη, $\frac{\partial C}{\partial w_{jk}^l}$ και τα biases $\frac{\partial C}{\partial b_j^l}$ καθώς είναι οι μόνοι μεταβλητοί παράγοντες.

Εάν θεωρήθει ως συνάρτηση κόστους το mean squared error

$$C_n = \frac{1}{2} \sum_j (y_j - a_j^l)^2$$

Η συνάρτηση κόστους για όλα τα δεδομένα, αποτελείται από το άθροισμά τους

$$C = \frac{1}{2} \sum_n C_n$$

Οι όροι w_{jk}^l και b_j^l δεν εμφανίζονται απευθείας στην παραπάνω εξίσωση. Παρόλα αυτά, επηρεάζουν έμμεσα το κόστος καθώς γίνεται η χρήση τους για τον υπολογισμό του του a_j^l , το οποίο δίνεται από την εξίσωση

$$a_j^l = \sigma \left(\sum_k a_k^{l-1} * w_{jk}^l + b_j^l \right)$$

Επιπλέον, για λόγο συντόμευσης συνηθίζεται να ορίζεται η ποσότητα

$$z_j^l = \sum_k a_k^{l-1} * w_{jk}^l + b_j^l$$

Η παραπάνω έκφραση συναντάται συχνά και σε μορφή πινάκων

$$a^l = \sigma(a^{l-1} * w^l + b^l)$$

$$z^l = a^{l-1} * w^l + b^l$$

$$a^l = \sigma(z^l)$$

Με τους παραπάνω ορισμούς και τον κανόνα της αλυσίδας, μπορεί να υπολογιστεί η παράγωγος για κάθε βάρος που συνδέει το προηγούμενο με το τελευταίο επίπεδο ως εξής.

$$\frac{\partial C}{\partial w_{jk}^l} = \frac{\partial C}{\partial a_j^l} \frac{\partial a_j^l}{\partial z_j^l} \frac{\partial z_j^l}{\partial w_{jk}^l}$$

Και παρομοίως για το bias

$$\frac{\partial C}{\partial b_j^l} = \frac{\partial C}{\partial a_j^l} \frac{\partial a_j^l}{\partial z_j^l} \frac{\partial z_j^l}{\partial b_j^l}$$

Όπου κάθε όρος υπολογίζεται

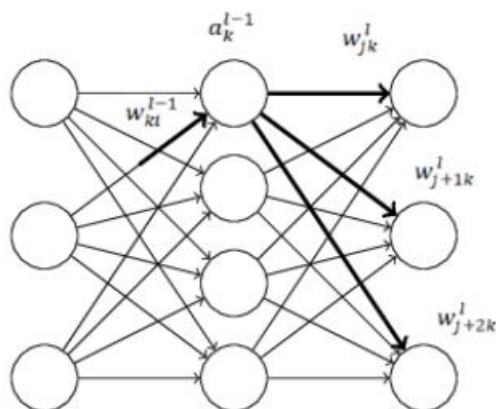
$$\frac{\partial C}{\partial a_j^l} = y_j - a_j^l$$

$$\frac{\partial a_j^l}{\partial z_j^l} = \sigma'(z_j^l)$$

$$\frac{\partial z_j^l}{\partial w_{jk}^l} = a_k^{l-1}$$

$$\frac{\partial z_j^l}{\partial b_j^l} = 1$$

Για την εύρεση των παραγώγων προηγούμενων επιπέδων, αρκεί να ληφθεί υπόψιν ο τρόπος με τον οποίο η κάθε παράμετρος επηρεάζει το κόστος. Το βάρος w_{ki}^{l-1} του προηγούμενου επιπέδου συνεισφέρει στον υπολογισμό της ενεργοποίησης του νευρώνα a_k^{l-1} . Πολύ προσοχή πρέπει να δοθεί στο γεγονός ότι η ενεργοποίηση του νευρώνα a_k^{l-1} συμβάλει στον υπολογισμό όλων των νευρώνων του επόμενου επιπέδου. Συνεπώς πρέπει να υπολογιστεί να άθροισμα όλων των συνεισφορών.



Εικόνα 19: Νευρωνικό Δίκτυο πολλών επιπέδων. Οι τιμές που αναγράφονται επάνω στα βάρη και τους νευρώνες είναι αυτές που χρησιμοποιούνται για τον υπολογισμό του κόστους.

$$\frac{\partial C}{\partial w_{ki}^{l-1}} = \sum_j \frac{\partial C}{\partial a_j^l} \frac{\partial a_j^l}{\partial z_j^l} \frac{\partial z_j^l}{\partial a_k^{l-1}} \frac{\partial a_j^l}{\partial z_k^{l-1}} \frac{\partial z_j^l}{\partial w_{ki}^{l-1}}$$

Και αντίστοιχα για τα biases

$$\frac{\partial C}{\partial b_k^{l-1}} = \sum_j \frac{\partial C}{\partial a_j^l} \frac{\partial a_j^l}{\partial z_j^l} \frac{\partial z_j^l}{\partial a_k^{l-1}} \frac{\partial a_j^l}{\partial z_k^{l-1}} \frac{\partial z_j^l}{\partial b_k^{l-1}}$$

Οι νέοι όροι που προκύπτουν υπολογίζονται ως εξής

$$\frac{\partial z_j^l}{\partial a_k^{l-1}} = w_{jk}^l$$

$$\frac{\partial a_j^l}{\partial z_k^{l-1}} = \sigma'(z_k^{l-1})$$

$$\frac{\partial z_j^l}{\partial w_{ki}^{l-1}} = a_i^{l-2}$$

$$\frac{\partial z_j^l}{\partial b_k^{l-1}} = 1$$

Είναι προφανές ότι επαναλαμβάνοντας αυτή την μέθοδο είναι δυνατός ο υπολογισμός όλων των παραμέτρων. Όσο πιο εσωτερικά είναι η παράμετρος, τόσο πιο περίπλοκος γίνεται ο συμβολισμός. Παρόλα αυτά η μέθοδος παραμένει η ίδια και βασίζεται σε πολύ απλές αρχές ανάλυσης και γραμμικής άλγεβρας, τις οποίες όλοι οι πρωτοετής φοιτητές έχουν διδαχθεί.

Κεφάλαιο 4: Συνελικτικά Νευρωνικά Δίκτυα

4.1 Γενικά

Τα τεχνητά νευρωνικά δίκτυα αποφέρουν πολύ καλά αποτελέσματα στα αριθμητικά δεδομένα, αλλά δεν διαθέτουν έναν εγγενή μηχανισμό για την συσχέτιση των δεδομένων μεταξύ τους. Ένα τέτοιο παράδειγμα είναι η αναγνώριση εικόνων. Η τιμές δυο pixel που βρίσκονται σε διαφορετικές γωνίες της εικόνας δεν έχουν την ίδια σύνδεση με δυο γειτονικά pixel. Τα συνελικτικά νευρωνικά δίκτυα(CNN) από την άλλη, εκμεταλλεύονται αυτή την χωρική εξάρτηση [64] και αποδίδουν εκπληκτικά αποτελέσματα σε τομείς που συμπεριλαμβάνουν, αλλά δεν περιορίζονται στην, όραση υπολογιστών και επεξεργασία φυσικής γλώσσας [65].

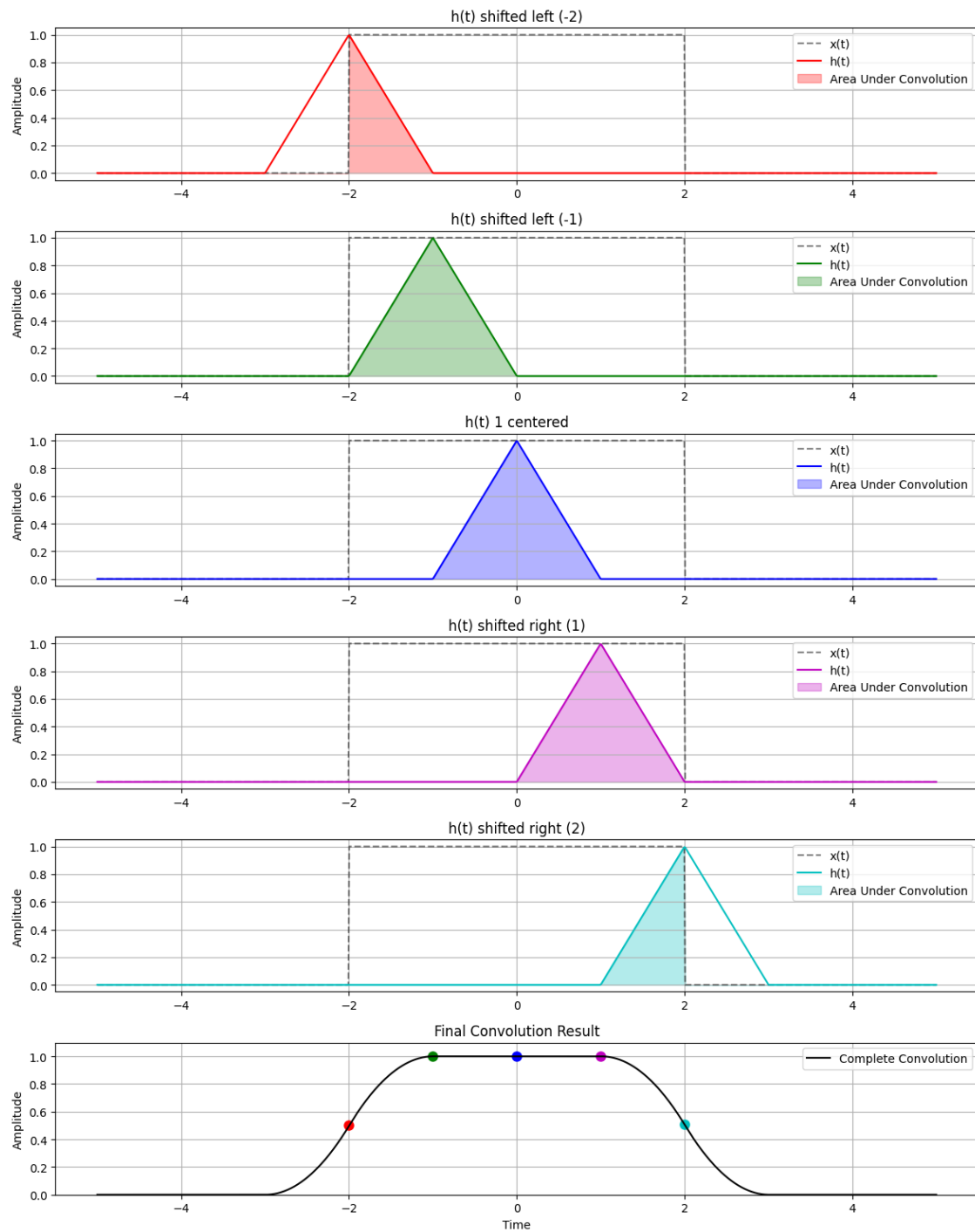
Τα CNN έχουν δύο κύρια είδη επιπέδων, τα συνελικτικά και τα συγκεντρωτικά.

4.2 Συνέλιξη

Η συνέλιξη είναι μία μαθηματική διαδικασία που εφαρμόζεται επάνω σε δύο συναρτήσεις, λίστες, πίνακες και η χρήση δεν περιορίζεται στα νευρωνικά δίκτυα. Συμβολίζεται με * και για δύο συνεχείς συναρτήσεις ορίζεται ως εξής.

$$y(t) = x(t) * h(t) = \int_{-\infty}^{\infty} x(\tau) h(t - \tau) d\tau [66]$$

Ο τύπος αυτός μοιάζει δυσνόητος αλλά, στην ουσία εκφράζει την ολίσθηση της συνάρτησης $h(t)$ επάνω στην $x(t)$, υπολογίζοντας σε κάθε βήμα το εμβαδό που βρίσκεται κάτω και από τις 2 συναρτήσεις.

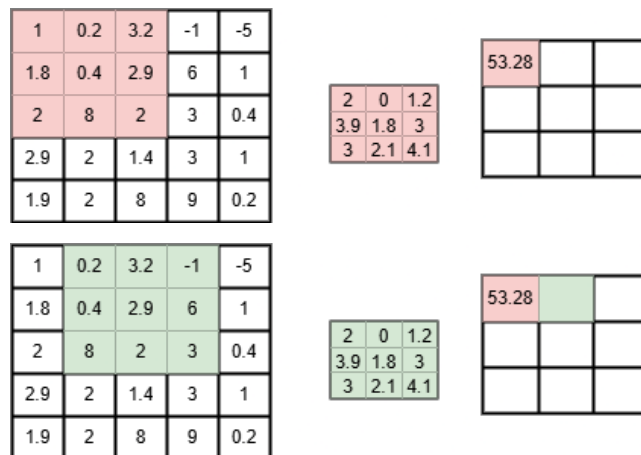


Εικόνα 20: Γραφική αναπαράσταση συνέλιξης δύο συναρτήσεων. Στα πέντε πρώτα γραφήματα φαίνεται η ολίσθηση της $h(t)$ επάνω στην $x(t)$, ενώ στην τελευταία γραφική παράσταση παρουσιάζεται το τελικό αποτέλεσμα

4.3 Συνέλιξη Στα Νευρωνικά Δίκτυα

Στα πλαίσια της όραση υπολογιστών με χρήση νευρωνικών δικτύων, η συνάρτηση $x(t)$, δηλαδή η είσοδος μας, είναι ένα πίνακας, όπου κάθε στοιχείο του αντιστοιχεί στην ένταση των pixel της εικόνας. Η $h(t)$ ονομάζεται φίλτρο[67]

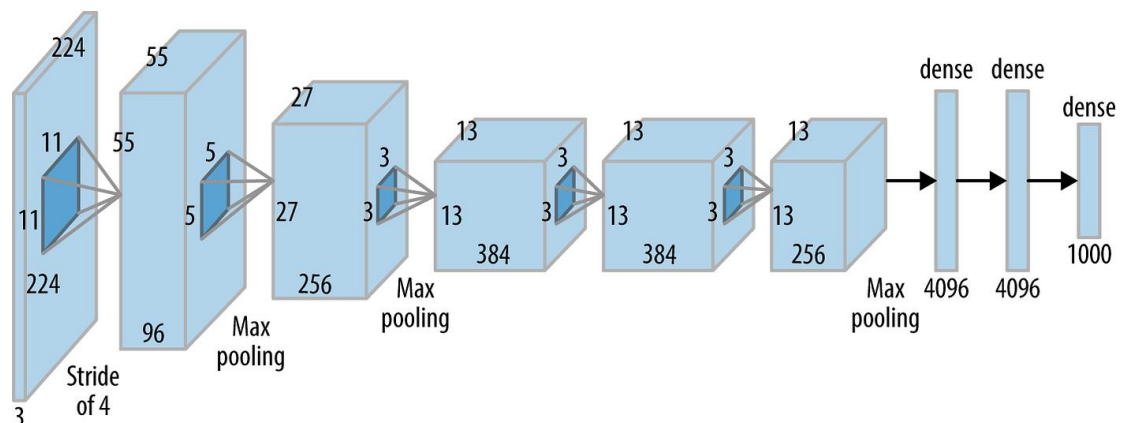
αποτελείτε επίσης από έναν πίνακα, συνήθως αρκετά μικρότερων διαστάσεων, με στοιχεία τα οποία ονομάζονται βάρη.



Εικόνα 21: Εφαρμογή συνελκτικού φίλτρου σε πίνακα.

Όπως φαίνεται στην παραπάνω εικόνα, όπως ακριβώς στην συνέλιξη δύο συναρτήσεων, το φίλτρο ολισθήσει επάνω στην εικόνα, πολλαπλασιάζοντας τα στοιχεία του με τα στοιχεία της εικόνας με τα οποία συμπίπτουν. Το άθροισμα των πολλαπλασιασμών αυτών, αποτελεί ένα στοιχείο της νέας μας εικόνας.

Ένα μοντέρνο CNN αποτελείται από πάρα πολλές συνελίξεις, είτε παράλληλα, δηλαδή εφαρμόζονται στην ίδια εικόνα και παράγουν διαφορετικό αποτέλεσμα, είτε σε σειρά, όπου η μία χρησιμοποιεί ως είσοδο των προηγούμενων.



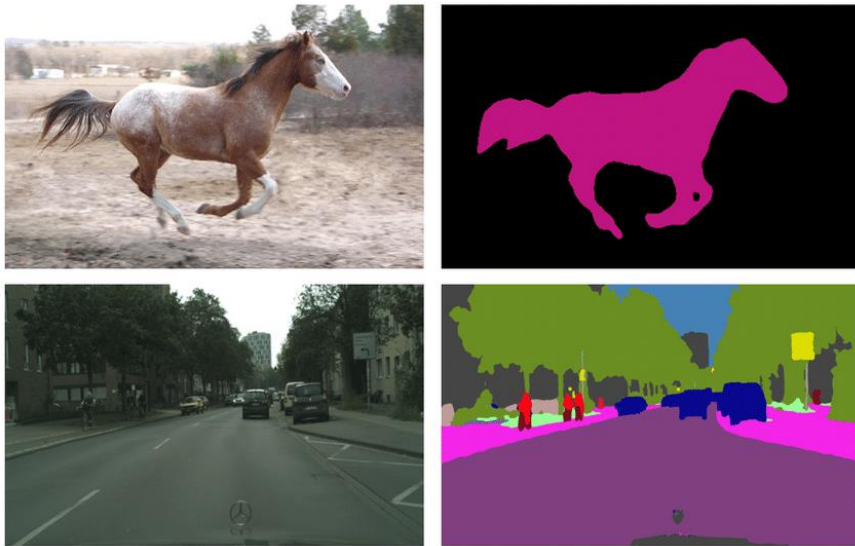
Εικόνα 22: Σχηματική αναπαράσταση CNN. Σε κάθε επίπεδο του δικτύου φαίνονται οι διαστάσεις της εικόνας, τα κανάλια (channels) που έχει, οι διαστάσεις του φίλτρου. Γίνεται χρήση πολλών φίλτρων στην σειρά καθώς και παράλληλα. Τέλος, τοποθετείται ένα πλήρως συνδεδεμένο νευρωνικό δίκτυο, όπως παρουσιάστηκε στο προηγούμενο κεφάλαιο [68].

Η δομή του CNN διαφέρει αρκετά από του νευρωνικού δικτύου,. Παρόλα αυτά, ο αλγόριθμος ο οποίος χρησιμοποιείται για την εκπαίδευση είναι ο αλγόριθμος backpropagation, όπως και σχεδόν σε όλους τους μοντέρνους τομείς της τεχνίτης

νοημοσύνης[52]. Όπως ακριβώς μπορεί, για ένα δείγμα να υπολογιστεί η παράγωγος του σφάλματος, ως προς ένα βάρος, έτσι και εδώ, υπολογίζεται για κάθε στοιχείο κάθε φίλτρου.

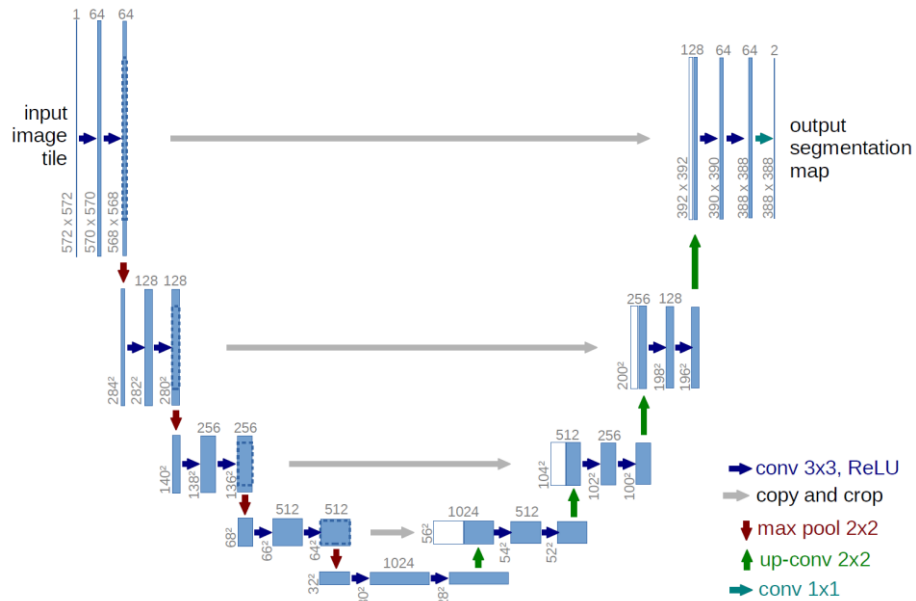
4.4 Αρχιτεκτονική Unet και Τμηματοποίηση εικόνας

Οι παραπάνω αρχιτεκτονικές είναι ικανές να διαχειριστούν εισόδους που αποτελούνται είτε από γραμμικά είτε από δεδομένα πινάκων, όπως για παράδειγμα μία εικόνα. Παρόλα αυτά, η έξοδος τους παραμένει είτε αριθμητική είτε κατηγορηματική. Στη περίπτωση του του καρκίνου, τα παραπάνω μοντέλα θα είχαν την δυνατότητα, με την μαγνητική τομογραφία του ασθενή να αποφανθούν εάν υπάρχει ή όγκος. Παρόλα αυτά, η ύπαρξη ή όχι όγκου, πολλές φορές δεν είναι η δυσκολότερη για το ιατρικό προσωπικό καθώς και η διαδικασία στην οποία γίνονται τα περισσότερα λάθη[5], [32]. Ένα ιδανικό μοντέλο θα μπορούσε να τοποθετήσει τον όγκο επάνω στην μαγνητική τομογραφία. Η έξοδος ενός τέτοιου μοντέλου, θα ήταν μία εικόνα ίδιων διαστάσεων με την είσοδο, όπου κάθε ρixel θα έχει κατηγοριοποιηθεί σε μία από τις διαθέσιμες κλάσεις.



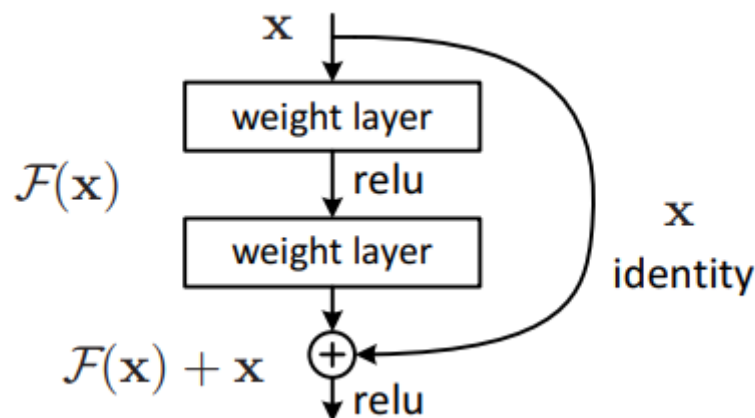
Εικόνα 23: Παράδειγμα τμηματοποίησης εικόνας. Σε κάθε ρixel της εισόδου έχει αποδοθεί ένας χαρακτηρισμός από το μοντέλο. Στην πρώτη περίπτωση υπάρχουν δύο επιλογές, είτε άλογο είτε όχι, ενώ στην δεύτερη το μοντέλο μπορεί να επιλέξει από μία πικοιλία κλάσεων, όπως αμάξι, άνθρωπος, δένδρο [69].

Για τον σκοπό της τμηματοποίησης εικόνας, μία από τις πιο διαδεδομένες αρχιτεκτονικές είναι γνωστή ως Unet [31]. Η αρχιτεκτονική αυτή κάνει χρήση συνελκτικών επιπέδων για την κωδικοποίηση και την αποκωδικοποίηση της εικόνας.



Εικόνα 24: Αρχιτεκτονική Unet.

Η κύρια διαφορά του παραπάνω μοντέλου, σε σχέση με ένα απλό CNN είναι η εκτεταμένη χρήση skip connections [70]. Η έξοδος ενός επιπέδου, όπως συνυθίζεται στα νευρωνικά δίκτυα, αποτελεί την είσοδο στο επόμενο επίπεδο. Έπειτα από κάποια επίπεδα, αθροίζεται με την εκάστοτε έξοδο περνώντας, κάνοντας δηλαδή skip κάποια από τα επίπεδα.



Εικόνα 25: Σχηματική αναπαράσταση skip connection [70]

Η αρχιτεκτονική αυτή επιλογή, επιτρέπει στα επόμενα επίπεδα να προσθέτουν πληροφορία, χωρίς να πρέπει να διατηρήσουν αυτή της εισόδου, καθώς αυτή θα προστεθεί στα επόμενα επίπεδα. Στο παράδειγμα της τμηματοποίησης εικόνας.

Κεφάλαιο 5: Παραγωγικά Ανταγωνιστικά Δίκτυα

5.1 Γενικά

Τα νευρωνικά δίκτυα, όπως έχουν περιγραφεί παραπάνω, ελαχιστοποιούν μία συνάρτηση κόστους, με βάση τα δεδομένα εκπαίδευσης. Έχουν την δυνατότητα να αναγνωρίσουν τα μοτίβο σε είδη υπάρχουσα δείγματα. Η δημιουργία νέων δεδομένων, αποτελούσε ανοιχτό πρόβλημα. Παρόλο που, υπήρχαν τεχνικές παραγωγής δεδομένων, η δημοσίευση που πρότεινε τα παραγωγικά ανταγωνιστικά δίκτυα, generative adversarial networks (GAN)[71], έφερε επανάσταση τον τομέα της τεχνητής νοημοσύνης και συγκεκριμένα των παραγωγικών μοντέλων.

5.2 Διαισθητική Ανάλυση

Ένας έμπειρος ζωγράφος μπορεί να δημιουργήσει αμέτρητα αντικείμενα του ίδιου είδους, όπως για παράδειγμα ένα δέντρο, τα οποία όχι μόνο δεν τα έχει δει πουθενά, αλλά κατά πάσα πιθανότητα δεν υπάρχουν πουθενά στον κόσμο. Μπορεί δηλαδή να συνθέσει μία μοναδική αλληλουχία από χρώματα, τα οποία όμως μοιράζονται τα κύρια χαρακτηριστικά με όλα τα υπόλοιπα αντικείμενα της ίδιας κατηγορίας. Το παραπάνω εγχείρημα, μοιάζει απίθανο για έναν αλγόριθμο, μία διαδικασία, η οποία εξ ορισμού μαθαίνει να μιμείται τα δεδομένα που του παρέχονται για εκπαίδευση. Επιπλέον, για να μπορέσει να εκπαιδευτεί ένα νευρωνικά δίκτυο, χρειάζεται μία συνεχή συνάρτηση κόστους, η οποία υποδεικνύει την κατεύθυνση μεταβολής των βαρών, με σκοπό την ελαχιστοποίησή της. Ένα δίκτυο το οποίο παράγει δεδομένα, όπως το παράδειγμα των δέντρων, με βάση τους ορισμούς που έχουν παρατεθεί παραπάνω, θα χρειαζόταν έναν ειδικό ζωγράφο, γεωπόνο ή κάποιον ειδικό να ποσοτικοποίησή το πόσο μακριά είναι το κάθε δείγμα, από ένα πραγματικό δέντρο. Όπως είναι κατανοητό, η διαδικασία αυτή θα απαιτούσε ανέφικτα πολύ χρόνο ανθρώπινης παρέμβασης.

Τα GAN, ξεπερνάνε το παραπάνω πρόβλημα, αφαιρώντας την ανάγκη για επίβλεψη, αντικαθιστώντας τον άνθρωπο από ένα άλλο νευρωνικό δίκτυο. Η εκπαίδευση του συχνά παρομοιάζεται με την διαδικασία πλαστογραφίας χρημάτων. Ο πλαστογράφος προσπαθεί να παράγει αληθοφανή, ενώ ένας αστυνόμος, discriminator,

προσπαθεί να τα ξεχωρίσει από τα πραγματικά χαρτονομίσματα. Κατά την εκκίνηση της εκπαίδευσης και οι δύο «παίκτες» έχουν μηδενική γνώση για το αντικείμενο. Όπως είναι φυσικό, ο παραγωγός, generator, παράγει δείγματα, τα οποία αποτελούνται αποκλειστικά από θόρυβο. Στην συνέχεια, ένα ζεύγος από πλαστά και αληθινά δείγματα παραπέθονται στον discriminator, για να αποφανθεί ποιο είναι πραγματικό και ποιο όχι. Ένας ιδανικός discriminator, θα διαχώριζε τα ψεύτικα, με τα πραγματικά. Καθώς το πρόβλημα έχει δυο αμοιβαίως αποκλειόμενα ενδεχόμενα, τους αναθέτουμε την κλάση 0 και 1 αντίστοιχα. Για αυτόν τον λόγο, η έξοδος του discriminator λαμβάνει συνεχείς τιμές σε αυτό το διάστημα, με την τιμή 0 να συμβολίζει ότι θεωρεί πως το δείγμα που εξέτασε είναι συνθετικό και 1 ότι είναι αληθινό. Σε αυτό το σημείο, γίνεται η μετατροπή του προβλήματος, από περιγραφικό, σε μαθηματικό. με βάση την παραπάνω περιγραφή η συνάρτηση κόστους του discriminator μπορεί διαισθητικά να οριστεί, ως εξής.

$$Loss_D = (1 - D(x_{real})) + D(x_{fake})$$

Όπου

- x_{real} είναι τα πραγματικά δεδομένα
- x_{fake} είναι τα συνθετικά, η ψεύτικα δεδομένα
- $D(x)$ είναι η έξοδος του discriminator, με είσοδο x

Όπως είναι φανερό, η παραπάνω συνάρτηση μηδενίζεται όταν ο discriminator δίνει την τιμή 1 στα πραγματικά δεδομένα, ενώ την τιμή μηδέν στα ψεύτικα, ξεχωρίζει δηλαδή με απόλυτη σιγουριά τις δύο κατηγορίες. Όσο διαρκεί η διαδικασία εκπαίδευσης, εφαρμόζεται ο αλγόριθμος backpropagation και ανανεώνονται τα βάρη, προς την κατεύθυνση ελαχιστοποίησής της.

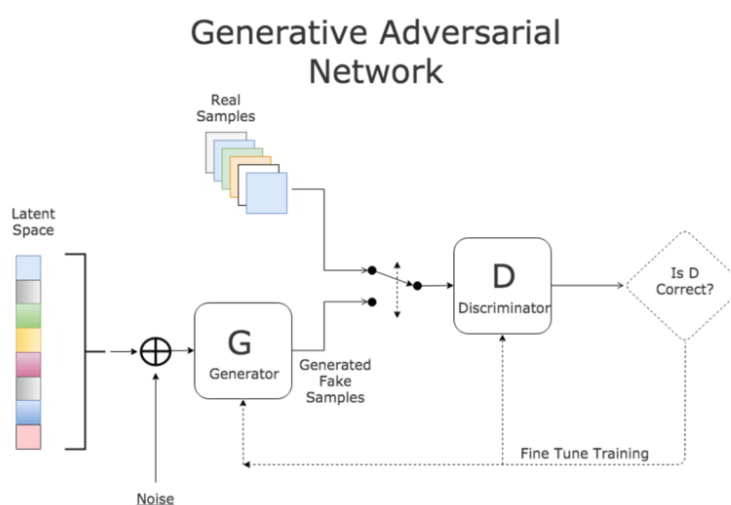
Για να παραχθούν όμως αληθοφανή δεδομένα, απαραίτητος είναι ο ορισμός συνάρτησης κόστους του generator. Ο σκοπός του είναι λοιπόν να ξεγελάσει τον discriminator, ώστε να κατατάξει το δεδομένα που έχει συνθέσει ως αληθινά. Ο δεύτερος κατά κάποιον τρόπο, αναλαμβάνει τον ρόλο του «ειδικού» που προαναφέρθηκε, να ποσοτικοποιεί την ομοιότητα των δύο αυτών ομάδων. Επόμενο είναι λοιπόν η συνάρτηση κόστους του να έχει την μορφή.

$$Loss_G = 1 - D(x_{fake})$$

Η παραπάνω συνάρτηση μειώνεται, όσο ο discriminator θεωρεί τα συνθετικά δεδομένα αληθοφανή. Σε τεχνικό επίπεδο, σημαντικό είναι να δοθεί έμφαση στο γεγονός ότι, για να εφαρμοστεί ο αλγόριθμος μάθησης στα βάρη του generator, πρέπει

να υπολογιστεί η παράγωγος ως προς τα βάρη του discriminator και να διαδοθεί μέσω του κανόνα της αλυσίδας σε κάθε βάρη του generator.

Η παραπάνω περιγραφή δεν εξασφαλίζει την ποικιλομορφία, ενώ ταυτόχρονα είναι ευάλωτη στο ενδεχόμενο ο generator να παράγει επανειλημμένα το ίδιο δείγμα, να έχει εντοπίσει δηλαδή ένα τοπικό ελάχιστο στον discriminator, το φαινόμενο του mode collapse [72]. Για να αποφευχθεί, ο generator δέχεται ως είσοδο ένα διάνυσμα από τυχαίους αριθμούς, τους οποίους καλείται να μετατρέψει στην εικόνα στόχο. Με αυτόν τον τρόπο εξαναγκάζεται να παράγει διαφορετικά δείγματα καθώς κάθε φορά η είσοδος του είναι διαφορετική.



Εικόνα 26: Σχηματική αναπαράσταση GAN. Η είσοδος το generator αποτελείται από ένα τυχαίο διάνυσμα, noise και παράγει το generated sample. Στην συνέχεια ο discriminator ανατεθεί την τιμή στα πραγματικά και συνθετικά δείγματα. Η τιμή αυτή χρησιμοποιείται ως συνάρτηση κόστους για την εκπαίδευση τους [73].

Σε αυτό το σημεία χρειάζεται να αναφερθεί ότι οι συναρτήσεις κόστους δεν είναι γραμμικές ως προς την έξοδο του discriminator, αλλά λογαριθμικές. Ο λόγος είναι σταθερότητα στην εκπαίδευση, όπως ακριβώς στην χρήση του cross entropy[74] Οι νέες συναρτήσεις κόστους είναι οι εξής.

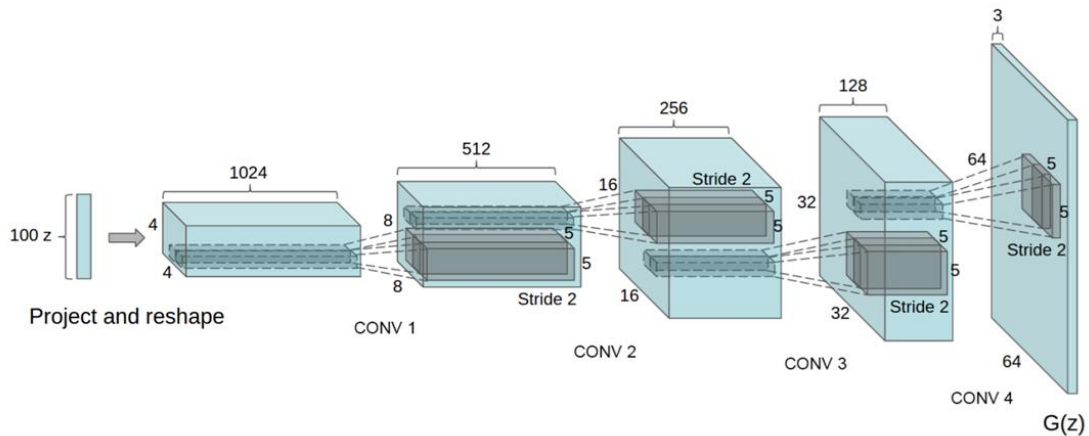
$$Loss_D = \log(D(G(z))) + \log(1 - D(x_{real}))$$

$$Loss_G = \log(1 - D(G(z)))$$

Όπου

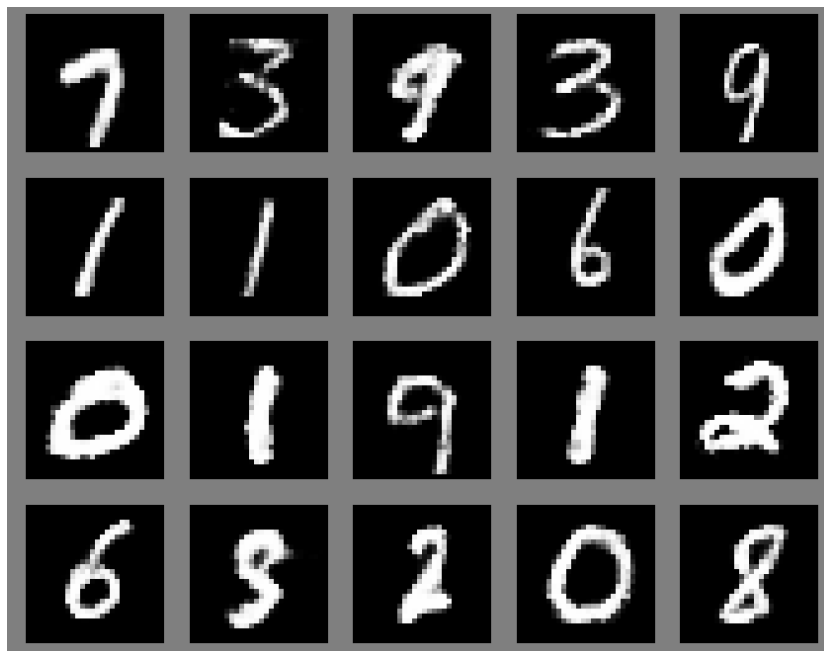
- z διάνυσμα τυχαίων τιμών, η είσοδος στον generator
- $G(z)$, η έξοδος του, συνώνυμο με x_{fake} στην προηγούμενη περιγραφή

Οι συναρτήσεις που παρατέθηκαν στην αρχή, λόγω μονοτονίας του λογαρίθμου έχουν ακριβώς την ίδια κατεύθυνση με τις παραπάνω. Για την παραγωγή εικόνων, όπως και στα συνηθισμένα νευρωνικά δίκτυα γίνεται χρήση συνελκτικών νευρωνικών δικτύων στον discriminator και χρήση de-convolution στον generator.



Εικόνα 27: Αρχιτεκτονική DCGAN

Η διαδικασία του de-convolution είναι η αντίθετη από αυτή της συνέλιξης, με κάθε επίπεδο να αυξάνει τις διαστάσεις της εικόνας. Με αυτόν τον τρόπο, από ένα μονοδιάστατο διάνυσμα, παράγεται μία εικόνα, με όσα channels και διαστάσεις χρειαστεί.



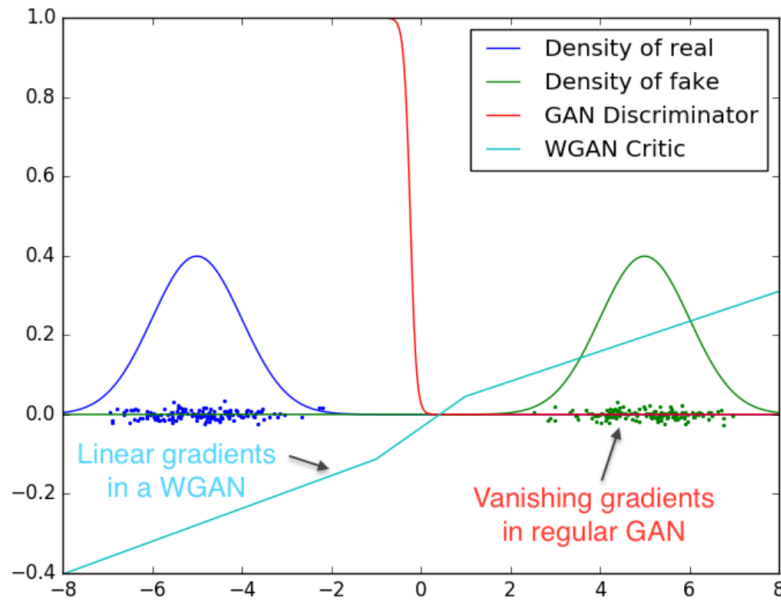
Εικόνα 28: Συνθετικά δεδομένα παραγόμενα από GAN, εκπαιδευμένα στην βάση δεδομένων MNIST [71]

5.3 Wasserstein GAN

Τα GAN αποτελούν μία εξαιρετική ιδέα. Παρόλα αυτά, η εφαρμογή τους στην πραγματικότητα έχει αποδειχθεί αρκετά δύσκολη. Σε κάθε διαδικασία εκπαίδευσης νευρωνικών δικτύων, ο προγραμματιστής καλείται να βρει, μέσα από διαδικασία trial and error, την σωστή αρχιτεκτονική και υπερπαραμέτρους που αποφέρουν τα καλύτερα αποτελέσματα. Στην περίπτωση των GAN, η διαδικασία αυτή γίνεται εκθετικά πιο περίπλοκη καθώς πρέπει οι παραπάνω επιλογές να ληφθούν για δύο δίκτυα, αντί για ένα.

Επιπλέον στην αρχή της διαδικασίας, όπου ο generator παράγει απλό θόρυβο, είναι φυσικό ο discriminator μετά από πολύ λίγες επαναλήψεις να μάθει πολύ εύκολα την διαφορά μεταξύ συνθετικών και πραγματικών δεδομένων και να έχει loss αρκετά κοντά στο μηδέν. Αντίστοιχα, το loss του generator τείνει στο ένα. Οι ακραίες αυτές λόγω της φύσης του λογαρίθμου που χρησιμοποιείται σαν συνάρτηση κόστους, έχουν ως αποτέλεσμα το gradient που παρέχεται στον generator να είναι κοντά στο μηδέν, άρα να γίνονται πολύ μικρά έως και μηδενικά βήματα εκπαίδευσης.

Για τον λόγο αυτό, αναπτύχθηκαν τα Wasserstein Gan (WGAN)[75] τα οποία, αφαιρούν τον περιορισμό ότι η έξοδος του discriminator (αναφέρεται ως critic στην περίπτωση των WGAN) πρέπει να ανήκει στο διάστημα $[0,1]$ και μπορεί να λαμβάνει οποιαδήποτε πραγματική τιμή. Επιπλέον, αφαιρέθηκε ο λογάριθμος από την συνάρτηση κόστους, για πιο ομαλή παράγωγο.



Εικόνα 29: Σύγκριση διαχωριστικής ευθείας GAN με WGAN. Είναι φανερό ότι στην περίπτωση της GAN, σε ακραίες τιμές, η παράγωγος της συνάρτησης μηδενίζεται (vanishing gradient) και ο generator πάει να εκπαιδεύεται [75].

5.3.1 Gradient Penalty

Καθώς αφαιρείται η ο περιορισμός στις τιμές που μπορεί να λάβει η έξοδος του critic, υπάρχει ο κίνδυνος να παρατηρηθούν υπερβολικές τιμές οι οποίες μπορούν να αποσταθεροποιήσουν την διαδικασία εκπαίδευσης [75]. Για να αποφευχθεί η ανεξέλεγκτη αύξηση των τιμών, οι διαγραφείς εφαρμόζουν περιορισμό στις τιμές που μπορούν να λάβουν τα βάρη του critic, στο εύρος τιμών $[-0.01, 0.01]$. Όπως παραδέχονται και οι ίδιοι, η τεχνική αυτή δεν είναι βέλτιστη, καθώς χάνεται είναι πολύ επεμβατική στην διαδικασία εκπαίδευσης.

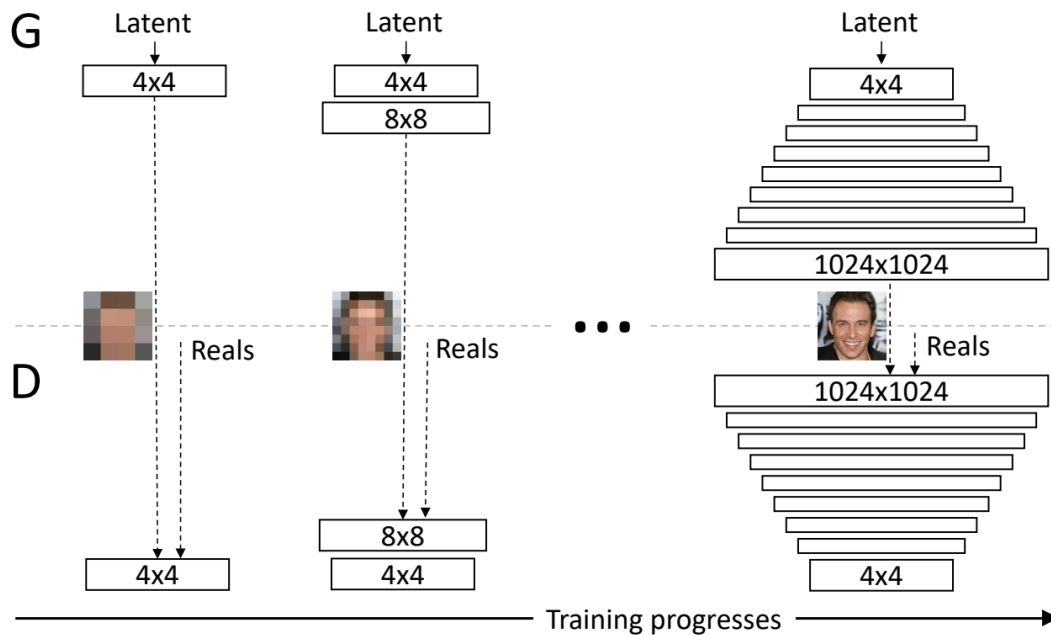
Για τον λόγο αυτό, σε πολύ μικρό χρονικό διάστημα μετά την δημοσίευση των WGAN, αναπτύχθηκε η τεχνική της ποινή παραγώγου (gradient penalty) [76]. Η διαδικασία αυτή, αντί να επιβάλλει τον περιορισμό των βαρών, προσθέτει στην συνάρτηση κόστους έναν επιπλέον όρο, ανάλογο του τετραγώνου της παραγώγου του critic, με είσοδο έναν συνδυασμό πραγματικών και συνθετικών δεδομένων. Με τον τρόπο αυτό, αντί να επιβάλλεται ο περιορισμός τιμών, απλά ενθαρρύνεται μέσω της συνάρτησης κόστους.

5.4 Pro-Gan

Η τελευταία αρχιτεκτονική που εξετάστηκε ήταν αυτή του Progressive -Gan (Pro-GAN) [77]. Η ιδέα σε αυτή την αρχιτεκτονική είναι αρκετά απλή, αλλά

συνοδεύεται από μία εξαιρετικά προσεγμένη υλοποίηση. Η ιδέα βασίζεται στην παρατήρηση ότι, ενώ τα GAN είχα σχετική επιτυχία στην παραγωγή εικόνων μικρών διαστάσεων, δυσκολευόντουσαν όταν η ανάλυση αυξανόταν όπως σε τιμές 1024×1024 , αναλύσεις οι οποίες μοιάζουν πολύ μεγάλες, σε σχέση με τις εφαρμογές τεχνητής νοημοσύνης (το Brats έχει διασπαστείς μόλις 240×240 ανά slice), δεν είναι τόσο μεγάλη σε σχέση με αυτές που συναντώνται στην καθημερινότητα.

Με αφορμή την παρατήρηση αυτή, οι συγγραφείς του Pro-Gan, αποφάσισαν να εκπαιδεύσουν ένα δίκτυο που παράγει εικόνες χαμηλής ποιότητας και στην συνέχεια, να γίνει χρήση του μοντέλου αυτού ως βάση, για την παραγωγή εικόνων διπλάσιας ανάλυσης. Συγκεκριμένα, εκπαιδεύτηκε ένα δίκτυο με μόλις 3 συνελκτικά επίπεδα για την παραγωγή εικόνων 4×4 . Στην συνέχεια, προστέθηκαν 3 επιπλέον επίπεδα, για την μετατροπή της εικόνας αυτής σε εικόνες ανάλυσης 8×8 .



Εικόνα 30: Σχηματική αναπαράσταση Pro-GAN. Στα αριστερά φαίνεται το δίκτυο το οποίο παράγει τις εικόνες με χαμηλή ανάλυση 4×4 . Δεξιά του βρίσκεται το επόμενο βήμα που παράγει 8×8 εικόνες, το πρώτο επίπεδο του οποίου είναι το αρχικό δίκτυο. Η διαδικασία αυτή επαναλαμβάνεται μέχρι να επιτευχθεί η επιθυμητή ανάλυση [77]

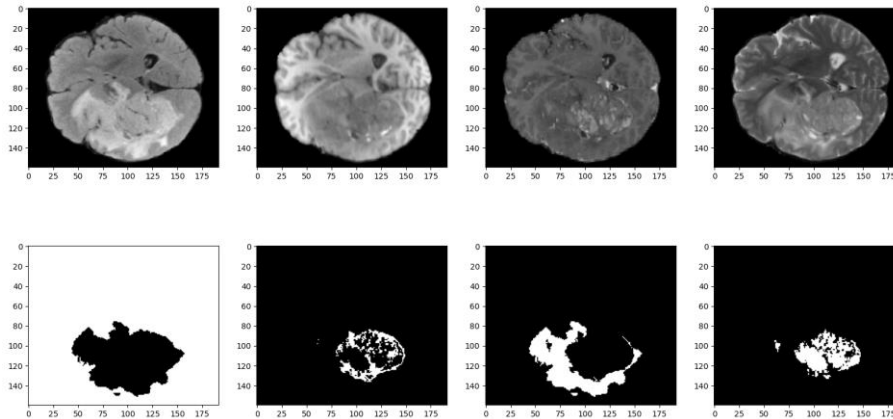
Η αρχιτεκτονική Pro-GAN, κάνει χρήση των τεχνικών που αναφέρθηκαν παραπάνω, όπως Wasserstein Loss και Gradient Penalty, για την εκπαίδευση των Νευρωνικών δικτύων.

Κεφάλαιο 6: Υλοποίηση & Αποτελέσματα

6.1 Γενικά

Στα πλαίσια της παρούσας διπλωματικής εργασίας, έχουν διεξαχθεί μία σειρά από πειράματα, με σκοπό την παραγωγή επιπλέον δεδομένων εκπαίδευσης, με σκοπό την βελτίωση της ακρίβειας. Η εκπαίδευση νευρωνικών δικτύων και συγκεκριμένα GAN, απαιτεί πολύ ακριβό και δυσεύρετο εξοπλισμό, κάρτες επεξεργασίας γραφικών GPU, ενώ δεν συναντώνται συχνά στην βιβλιογραφία[78]. Όσο μεγαλύτερο το μοντέλο, τόσο περισσότερη μνήμη και επεξεργαστική ισχύ είναι απαραίτητη. Καθώς δεν υπήρχε πρόσβαση σε τόσο ακριβό εξοπλισμό, έτσι ώστε να διαχειριστεί το υπολογιστικό βάρος των τρισδιάστατων εικόνων, τα πειράματα διεξάχθηκαν σε μία τομή του x-y άξονα για κάθε ασθενή. Η επιλογή αυτή προσθέτει μία επιπλέον δυσκολία στο πρόβλημα, καθώς δεν είναι δεδομένο ότι σε κάθε τομή θα υπάρχουν προσβεβλημένα κύτταρα. Παρόλα αυτά, η επιλογή αυτή θεωρήθηκε ρεαλιστική, καθώς στην πραγματική ζωή, μεγάλο ποσοστό των εξεταζόμενων, είναι πιθανό να μην πάσχουν από καρκίνο. Ένα μοντέλο το οποίο έχει εκπαιδευτεί μόνο σε ασθενής, είναι αρκετά πιθανό να παράγει πολλά δείγματα ως false-positive, δηλαδή να θεωρεί ότι υπάρχει καρκίνος σε υγιή κύτταρα.

Επιπλέον, καθώς στην πολυπήφια των ασθενών, παρατηρούνταν αρκετός χώρος δίχως χρήσιμη πληροφορία, έγινε περικοπή των δεδομένων από διαστάσεις 224x224 σε 160x192. Η επιλογή αυτή έγινε καθώς, λόγω των συνελίξεων του Unet και κυρίως των downsample layers, οι διαστάσεις της εικόνα πρέπει να είναι πολλαπλάσιες του 16. Οι αριθμοί αυτοί είναι οι μικρότεροι ανά διάσταση, όπου κόβεται μόνο background image, και πληρούν τον παραπάνω περιορισμό.



Εικόνα 31: Παράδειγμα τομής που χρησιμοποιήθηκε ως δεδομένο εκπαίδευσης. Επάνω φαίνονται τα 4 modalities, ενώ κάτω φαίνονται οι τέσσερις διαφορετικές κατηγοριοποιήσεις.

Η υλοποίηση του κώδικα βασίστηκε στις υλοποιήσεις που βρίσκονται στο παρακάτω github [79]. Η κύρια διαφοροποίηση με τις παραπάνω υλοποιήσεις είναι πως, στην έξοδο των generator, τα μισά channels της εξόδου του generator έχουν στο τελικό στάδιο συνάρτηση ενεργοποίησης $\text{Tanh}()$, τα οποία αντιπροσωπεύουν τα τέσσερα modalities που περιεγράφηκαν παραπάνω. Ταυτόχρονα, τα τέσσερα επόμενα channels αποτελούν την μάσκα και οι τιμές τους είναι αμοιβαίως αποκλεισόμενα. Για αυτόν τον λόγο, στα επόμενα τέσσερα channels τοποθετήθηκε συνάρτηση ενεργοποίησης softmax [80], η οποία δίνει περισσότερο βάρος σε μία από τις επιλογές που τις παρέχονται.

Για επιβεβαίωση των αποτελεσμάτων έγινε διαχωρισμός του dataset σε 3 τμήματα. Τα δεδομένα εκπαίδευσης (train), επιβεβαίωσης (validation) και δομικής (test) [78].

6.2 Σχετική Βιβλιογραφία

Η χρήση νευρωνικών δικτύων για την τμηματοποίηση ιατρικής εικόνας είναι ένα θέμα που πληθώρα βιβλιογραφίας [7], [8], [31], [32]. Όσον αφορά την χρήση GAN, έχει γίνει εφαρμογή της αρχιτεκτονικής CycleGan [81] για την απευθείας τμηματοποίηση εικόνας [82], χωρίς την διαδικασία παραγωγής δεδομένων. Επιπλέον, αρκετά συχνά απαντάται η χρήση GAN, όχι για την παραγωγή εντελώς νέων δεδομένων, αλλά ως ένα βήμα επαύξησης δεδομένων (data augmentation) [83], [84], [85] [86]. Οι αρχιτεκτονικές αυτές ανήκουν στην οικογένεια των Conditional GAN [87]. Τα δίκτυα αυτά δεν δέχονται ως είσοδο τυχαίο θόρυβο, όπως περιεγράφηκε προηγουμένως, αλλά μία πληροφορία από τα δεδομένα μας, είτε την εικόνα [84], [86], είτε την μάσκα [85], από τα υπάρχουσα δεδομένα. Ως έξοδο έχουν το

συμπληρωματικό της εισόδου, το οποίο ιδανικά έχει διατηρήσει τα κομβικά χαρακτηριστικά, αλλά είναι αρκετά διαφορετικό έτσι ώστε να μπορεί να θεωρηθεί ένα επιπλέον δείγμα.

Δημοσιεύσεις οι οποίες παράγουν συνθετικά πλήρως δεδομένα, συχνά περιορίζονται στην χρήση τους για απλούστερα προβλήματα, όπως binary classification για την ύπαρξη όγκου[88], ή πρόβλεψη όγκου μεγέθυνσης όγκου [89]. Και τα δύο αποτελούν πολύ σημαντικά προβλήματα στον πραγματικό κόσμο, αλλά από υπολογιστικής άποψης, έχουν λιγότερη πολυπλοκότητα. Η παραγωγή πλήρως συνθετικών δεδομένων μαζί με την μάσκα, για εφαρμογή τμηματοποίησης εικόνας συναντάται πιο σπάνια[78], [90], [91]. Επιπλέον στα [90], [91] γίνεται πρόβλεψη σε ιστούς, όπως το δέρμα, με αρκετά πιο απλές δομές, από τον εγκέφαλο στον οποίο μπορεί να εμφανιστεί ένας όγκος.

Στην βιβλιογραφία στην οποία έγινε επισκόπηση, η πιο κοντινή προσέγγιση στην παρούσα εργασία αποτελεί η δημοσίευση [78], όπου γίνεται προσπάθεια για παραγωγή πλήρως συνθετικών δεδομένων, κάνοντας χρήση των Brats 2020 καθώς και Brats 2021 datasets, με σύνολο 1500 ασθενείς, σε αντίθεση με τα 369 τα οποία χρησιμοποιήθηκαν. Το συμπέρασμα ήταν πως η χρήση συνθετικών δεδομένων, αντί για αληθινών μειώνει, σε μικρό βαθμό την απόδοση των μοντέλων [78]. Συμπέρασμα τους είναι ότι τα συνθετικά δεδομένα μπορούν να αντικαταστήσουν τα πραγματικά, εάν υπάρχει περιορισμός στην πρόσβαση των δεύτερων.

6.3 Τμηματοποίηση εικόνας

Για τον σκοπό της τμηματοποίησης εικόνας έγινε χρήση ενός μοντέλου Unet με και χωρίς Ως συνάρτηση κόστους, όπως συνηθίζεται σε περιπτώσεις τμηματοποίησης εικόνας [92], όπως η πλειοψηφία της μάσκας αποτελείται από background class, ήταν η dice loss [93].

$$DiceLoss = 1 - 2 * \frac{\sum y * label + c}{\sum y + label + c}$$

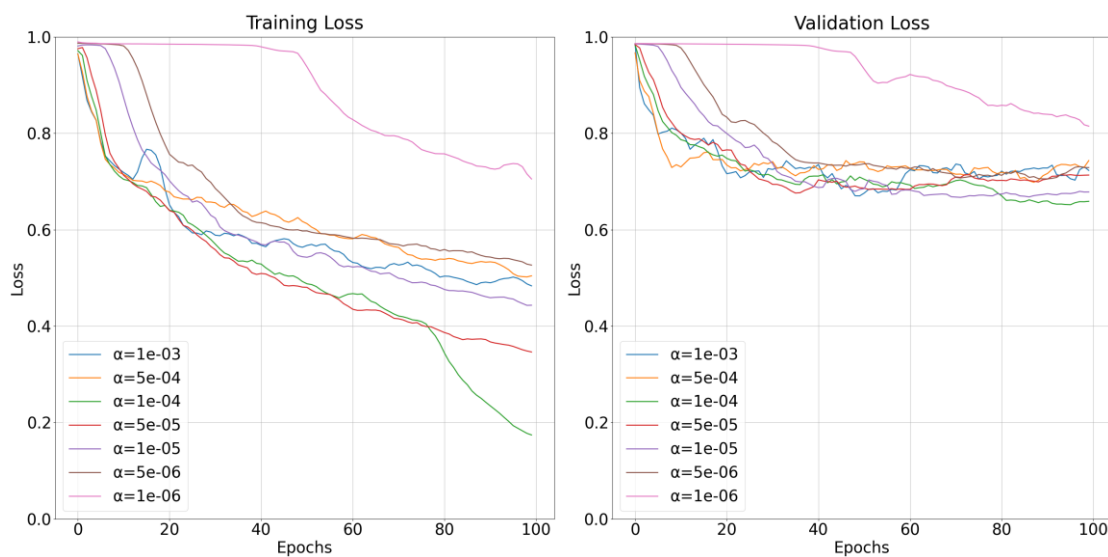
Όπου

- y η έξοδος του μοντέλου σε μορφή εικόνας
- $label$ η επιθυμητή έξοδος, σε μορφή εικόνας
- c μικρός αριθμός για να αποφευχθεί διαίρεση με το 0 (σε περίπτωση που $y = label = 0$).

- * Ο πολλαπλασιασμός στοιχεί με στοιχείο των εικόνων (element wise multiplication)

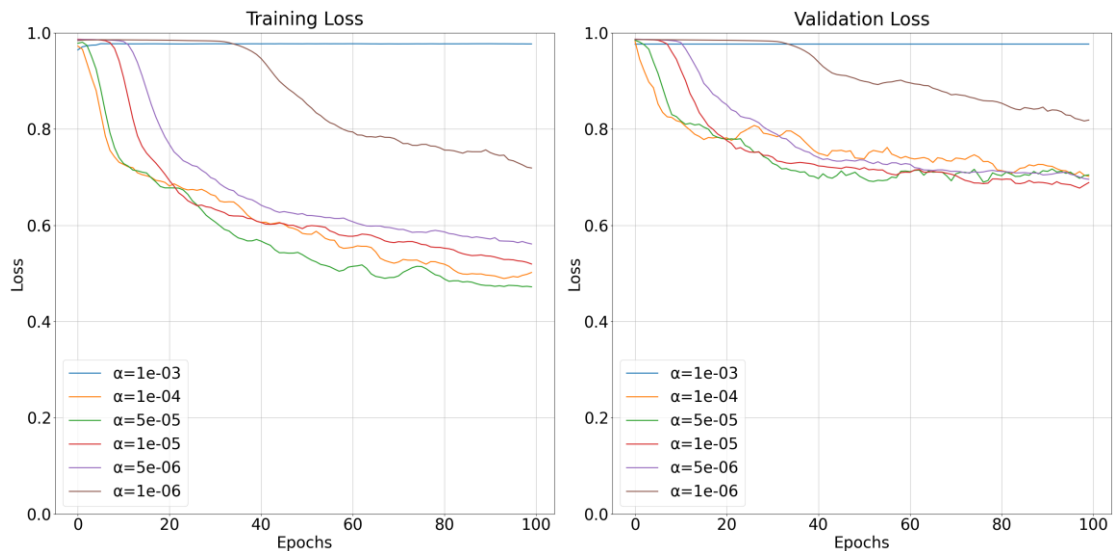
Με δεδομένο ότι τα στοιχεία των παραπάνω πινάκων λαμβάνουν τιμές στο διάστημα $[0,1]$, όταν $y = label$, $1*1 = 1$ για κάθε στοιχείο του πίνακα. Άρα $\sum y * label = \sum y$. Συνεπώς το κλάσμα θα λάβει την τιμή 0.5, η οποία όταν πολλαπλασιαστεί με 2 θα γίνει 1. Το κόστος λοιπόν θα μηδενιστεί, αφού $1-1=0$.

Παρακάτω φαίνονται γραφικές παραστάσεις εκπαίδευσης, για διάφορες τιμές της παραμέτρου learning rate, α .



Εικόνα 32: Γραφική αναπαράσταση της διαδικασίας εκπαίδευσης. Αριστερά: Dice loss στο test dataset. Δεξιά: Dice loss στο validation dataset.

Στο παραπάνω σχήμα φαίνεται ξεκάθαρα το φαινόμενο του overfitting[94], όπου το μοντέλο έχει «μάθει» πολύ καλά τα δεδομένα εκπαίδευσης. Η γνώση αυτή δεν γενικεύεται όμως σε δεδομένα που δεν έχει ξαναδεί. Όπως ένας μαθητής μπορεί να μάθει απέξω της ασκήσεις που υπάρχουν στο βιβλίο, αλλά δεν καταλαβαίνει τις έννοιες ενός φαινομένου, δεν μπορεί να λύσει τις ασκήσεις στο διαγώνισμα. Για να καταπολεμηθεί το παραπάνω πρόβλημα, συχνά γίνεται εφαρμογή τεχνικών επαύξησης δεδομένων, όπως περιστροφή, περικοπή και προσθήκη θορύβου [95]. Οι τεχνικές αυτές έχουν σκοπό να παραποιήσουν σε μικρό βαθμό τα δεδομένα εκπαίδευσης, χωρίς όμως να αλλοιωθούν τα δομικά χαρακτηριστικά της εικόνας.



Εικόνα 33: Γραφική αναπαράσταση της διαδικασίας εκπαίδευσης με χρήση τεχνικών προσαύξησης δεδομένων. Αριστερά: Dice loss στο test dataset. Δεξιά: Dice loss στο validation dataset.

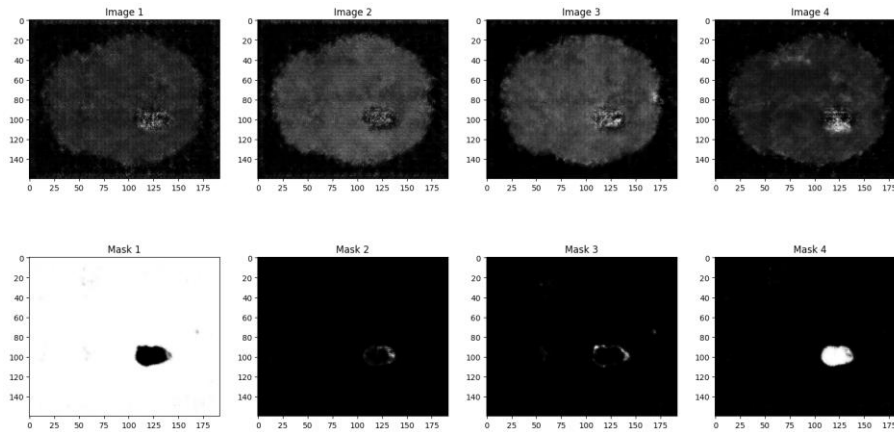
Παρατηρείται μικρότερο φαινόμενο overfitting, δηλαδή η απόδοση στο test και validation dataset έχουν μικρότερη διαφορά. Παρόλα αυτά, η τελική απόδοση χειροτερεύει, οπότε δεν θα γίνει χρήση τους στα μελλοντικά πειράματα της παρούσας εργασίας.

6.4 Παραγωγή δεδομένων

Για την παραγωγή δεδομένων, πραγματοποιήθηκαν δοκιμές με διάφορες αρχιτεκτονικές Generative Adversarial Networks (GAN), με στόχο τη δημιουργία ρεαλιστικών και ποικίλων συνόλων δεδομένων. Ο πειραματισμός επικεντρώθηκε στην στις διάφορες αρχιτεκτονικές GAN καθώς και στις κυριότερες υπερπαραμέτρους, όπως είναι το learning rate. Τα δεδομένα στα οποία είχε πρόσβαση ο discriminator, σε όλες τις περιπτώσεις εκπαίδευσης, ήταν μόνο τα δεδομένα εκπαίδευσης.

6.3.1 DCGAN

Η πρώτη δοκιμή έγινε με την χρήση του DCGAN. Όπως ήταν φυσικό, μία τόσο απλή αρχιτεκτονική δεν είναι ικανή να αποτυπώσει τις περίπλοκες αύλακες του εγκεφάλου.

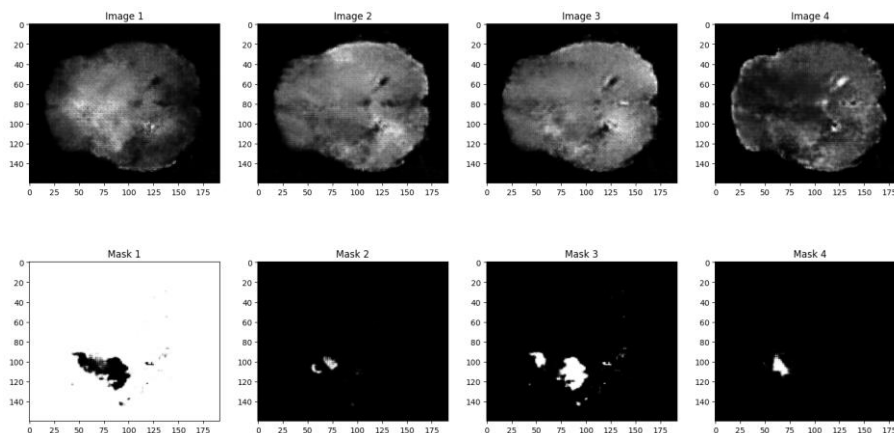


Εικόνα 34: Αποτελέσματα της αρχιτεκτονικής DCGAN

Όπως φαίνεται στην παραπάνω εικόνα, παρόλη την απλότητα της αρχιτεκτονικής, έχει αποτυπωθεί το γενικό σχήμα του εγκεφάλου. Ταυτόχρονα, στο σημείο που έχει τοποθετηθεί η μάσκα, υπάρχει διαφορά στις εικόνες από τα υπόλοιπα κύτταρα. Η διαφορά αυτή υποδηλώνει ότι έχει αποτυπωθεί στον generator ότι τα σημεία στα οποία βρίσκεται ο όγκος έχουν διαφορές από τα υπόλοιπα κύτταρα. Παρόλα αυτά, τα παραπάνω δεδομένα δεν είναι αρκετά ικανοποιητικά για να εκπαιδευτεί ένα μοντέλο.

6.3.2 Skip-Connections Generator

Όπως φάνηκε και στην περιγραφή του Unet, η εφαρμογή skip connections είναι εξαιρετικά χρήσιμη και βελτιώνει την απόδοση των νευρωνικών δικτύων, ειδικά σε ιατρικές εφαρμογές. Στην παρακάτω υλοποίηση, τα απλά de-convolutional layers, αντικαταστάθηκαν με convolution blocks τα οποία περιέχουν skip connections, όπως περιγράφεται παρακάτω.

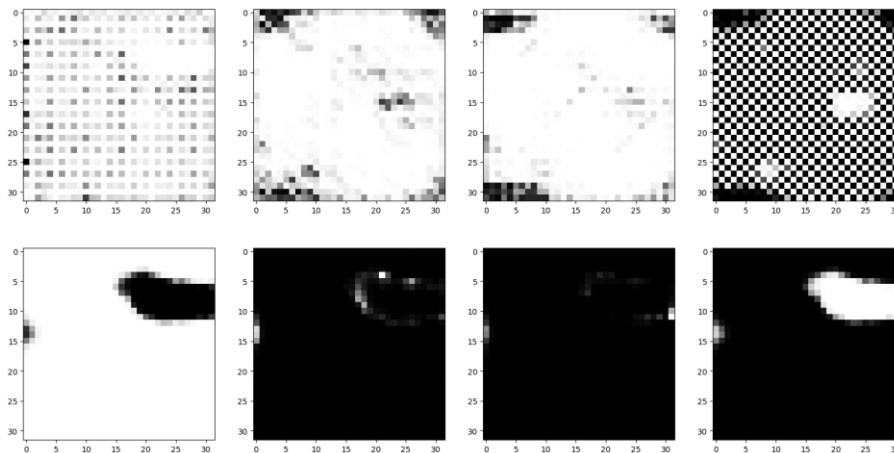


Εικόνα 35: Αποτελέσματα της αρχιτεκτονικής Skip-Connection GAN

Η βελτίωση των αποτελεσμάτων είναι ξεκάθαρη, με την αποτύπωση χαρακτηριστικών όπως οι εσωτερικές δομές. Παρόλα αυτά, τα αποτελέσματα ακόμα δεν είναι ικανοποιητικά.

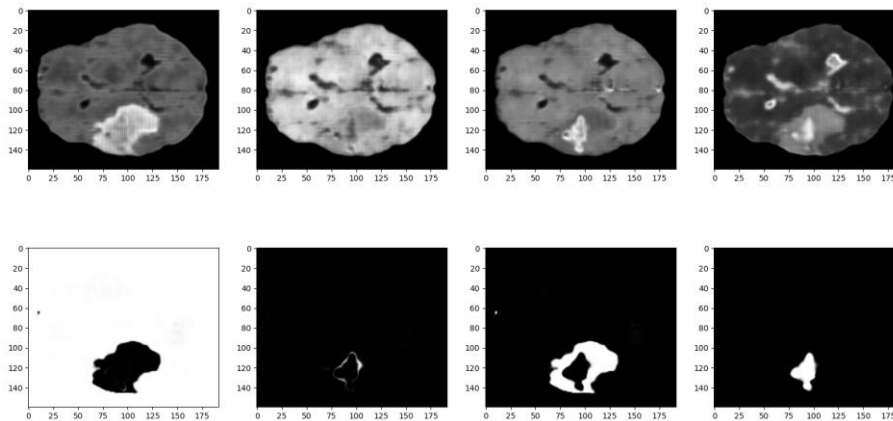
6.3.3 Συνδυασμός μεθόδων

Εάν γίνει εξέταση των παραπάνω εικόνων, ειδικά στην περίπτωση των skip-connections, θα γίνουν αντιληπτά κάποια μοτίβο, παρόμοια με μοτίβο σκακιέρας, όπως εμφανίζεται παρακάτω, τα λεγόμενα artifacts [96].



Εικόνα 36: Ενδιάμεσο στάδιο παραγωγής δεδομένων, όπου είναι φανερή η παρουσία artifacts.

Παρατηρήθηκε ότι η χρήση skip-connections βελτιώνει στην δημιουργία πιο περίπλοκων δομών από τον generator, αλλά υστερούσε σε χαρακτηριστικά όπως η υφή. Για αυτό τον λόγο, παρόμοιος με το [8], [85], έγινε χρήση πολλών discriminators. Σε αντίθεση με το προαναφερόμενο άρθρο, όλοι οι με διαφορετική δομή ο κάθε ένας είχε διαφορετική δομή, με τους 3 πρώτους να dcgan discriminators διαφορετικού βάθους, ενώ ο τελευταίος ο skip-connection discriminator. Η ιδέα είναι ότι εάν ο generator, από κάθε έναν discriminator, θα μαθαίνει διαφορετικά χαρακτηριστικά που περιέχονται στα δείγματα, από την υφή μέχρι τις περίπλοκες δομές. Σε κάθε εποχή, πραγματοποιούνται πέντε βήματα εκπαίδευσης για κάθε discriminator και ένα βήμα εκπαίδευσης του generator, με κάθε discriminator.

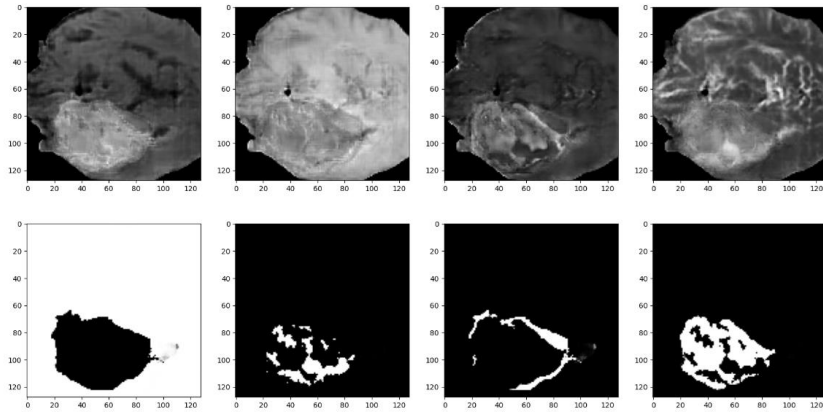


Εικόνα 37: Αποτελέσματα της αρχιτεκτονικής πολλαπλών discriminators

Τα αποτελέσματα που παρουσιάζονται είναι σαφώς τα καλύτερα που έχουν βρεθεί μέχρι στιγμής. Ο generator φαίνεται να έχει αντιληφθεί την συνολική δομή των δεδομένων, αλλά ακόμα υστερεί στην υφή, όπου συνεχίζει να εμφανίζει artifact.

6.4.4 Pro-GAN

Όπως αναφέρθηκε και προηγουμένως, η αρχιτεκτονική Pro-GAN όχι μόνο προσφέρει μία εξαιρετικά καλή διαισθητική ανάλυση για τους λόγους τους οποίους αποφέρει αποτελέσματα, αλλά έχει γίνει και χρήση της στην βιβλιογραφία[78], με αρκετά καλά αποτελέσματα. Η αρχιτεκτονική του Pro-GAN όμως, έχει πολύ συγκεκριμένους περιορισμούς ως προς τις διαστάσεις της εικόνας. Συγκεκριμένα Οι διαστάσεις κάθε επιπέδου είναι οι διπλάσιες του προηγούμενου. Για να διατηρηθεί το μοτίβο αυτό, έγινε επιπλέον περικοπή των εικόνων σε 128x128. Η εναλλακτική λύση θα ήταν η επένδυση (padding) έτσι ώστε να φτάσουν σε διαστάσεις 256x256 και έπειτα η περικοπή των συνθετικών σε 160x192. Η διαδικασία αυτή, εάν και πιο ορθή θα χρειαζόταν αρκετά περισσότερους πόρους, για μικρή βελτίωση καθώς, όπως φαίνεται παρακάτω, η χρήσιμη πληροφορία που αφαιρείται είναι ελάχιστη.



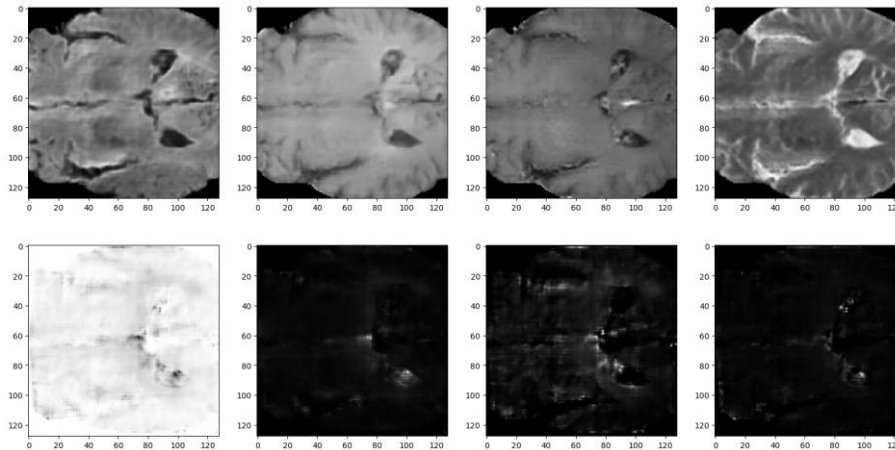
Εικόνα 38: Αποτελέσματα αρχιτεκτονικής Pro-GAN

Όπως φαίνεται παραπάνω, τα αποτελέσματα, τουλάχιστον σε ένα ανεκπαίδευτο μάτι, μοιάζουν αρκετά καλά, με τις αυλακώσεις του εγκεφάλου να αποτυπώνονται ικανοποιητικά. Επιπλέον, πιο σημαντικά για το πρόβλημα που η παρούσα εργασία προσπαθεί να λύσει, το σχήμα αλλά και οι διαφορετικές κατηγορίες του όγκου έχουν παρόμοια μορφή με αυτή που φαίνεται στην **Εικόνα 31**. Σε αντίθεση με τις προηγούμενες αρχιτεκτονικές, όπου ο όγκος είχε συχνά την μορφή ομόκεντρων σχημάτων, βλέπε **Εικόνα 37**, εμφανίζεται μία περιπλεγμένη δομή, με την μία κατηγορία όγκου να αναδύει μέσα από την άλλη, όπως στα πραγματικά δεδομένα.

6.5 Χρήση δεδομένων στην εκπαίδευση

Η οπτική ομοιότητα, για έναν ανειδίκευτο παρατηρητή, των παραπάνω δεδομένων δεν τα καθιστά αυτομάτως χρήσιμα. Για να επιτευχθεί ο σκοπός της παρούσας εργασίας, θα πρέπει η χρήση τους να συμβάλει στην βελτίωση του αποτελέσματος στην τμηματοποίηση εικόνας όπου είναι ο απώτερος σκοπός.

Αρχικά, έγινε παραγωγή συνθετικών 5000 δειγμάτων, με την χρήση της αρχιτεκτονικής Pro-GAN. Τα δείγματα αυτά δεν ήταν όλα ισάξια σε ποιότητα.

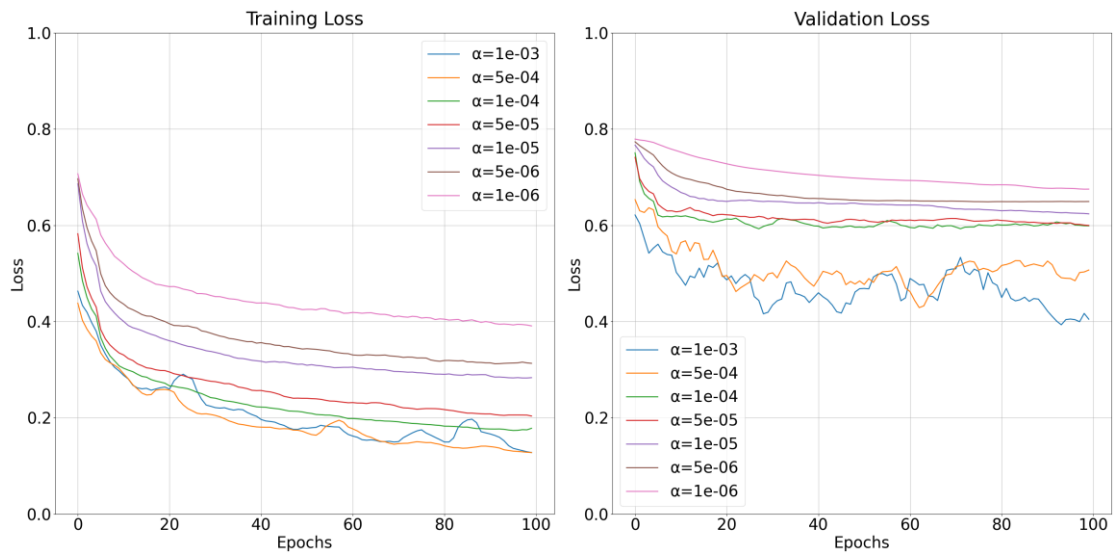


Εικόνα 39: Παράδειγμα μη πετυχημένου δείγματος με την χρήση Pro-GAN

Παραπάνω φαίνεται ένα τυχαίο παράδειγμα, όπου η μαγνητική τομογραφία έχει εξαιρετικά καλά αποτελέσματα, αλλά η μάσκα που την συνοδεύει είναι εξαιρετικά λανθασμένη. Εάν το δείγμα αυτό χρησιμοποιούταν για εκπαίδευση του Unet, τα αποτελέσματα θα ήταν πολύ χειρότερα. Για τον λόγο αυτό, παράλληλα με την διαδικασία παραγωγής δεδομένων από τον generator, γινόταν και η διαδικασία κρίσης από τον discriminator, ο οποίος ανέθετε μία τιμή, που αφορούσε το πόσο κοντά στην πραγματικότητα ήταν το κάθε δείγμα. Τα 1000 πιο ρεαλιστικά, με βάση τον discriminator δείγματα αποτελούσαν το συνθετικό dataset. Τα δεδομένα αυτά αξιοποιήθηκαν με τρεις διαφορετικούς τρόπους.

- Αποκλειστική χρήση συνθετικών δεδομένων.
- Συνδυασμός συνθετικών και πραγματικών δεδομένων.
- Χρήση συνθετικών δεδομένων για προ εκπαίδευση του μοντέλου και στην συνέχεια επανεκπαίδευση στα πραγματικά δεδομένα.

Όπως είναι φυσικό, η πρώτη μέθοδος απέτυχε, με αρκετά χειρότερα αποτελέσματα. Η δεύτερη μέθοδος, εάν και καλύτερο από την πρώτη, πάλι υστερούσε. Η μέθοδος που απέφερε τα καλύτερα αποτελέσματα ήταν η προ εκπαίδευση στα συνθετικά δεδομένα.

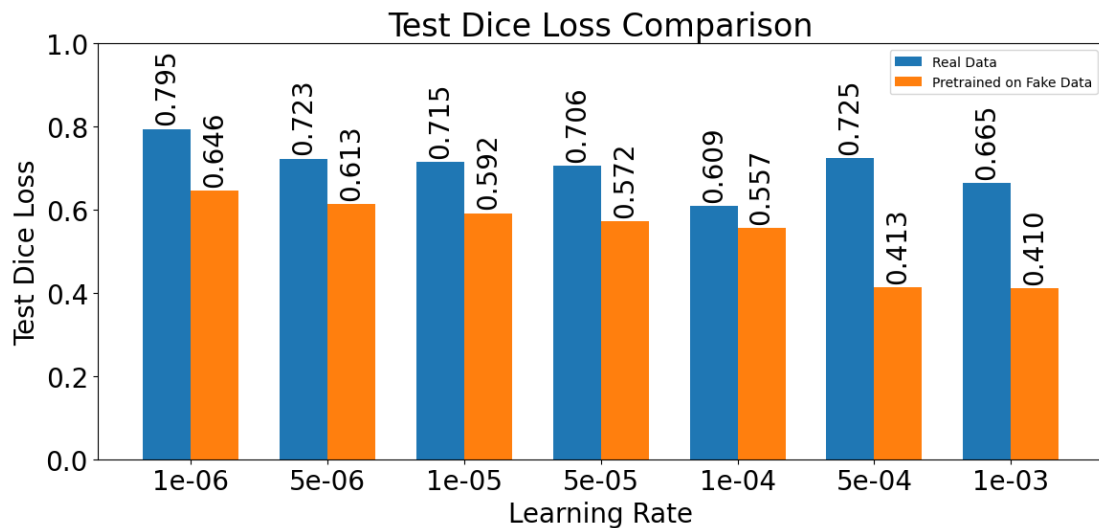


Εικόνα 40: Γραφική αναπαράσταση της διαδικασίας εκπαίδευσης με χρήση προ εκπαιδευμένου μοντέλου στα συνθετικά δεδομένα. Η συμβολή των συνθετικών δεδομένων είναι ξεκάθαρη από το γεγονός ότι η συνάρτηση κόστους ξεκινάει από τιμές, χαμηλότερες από το 1, υποδεικνύοντας ότι το μοντέλο έχει μάθει χρήσιμες πληροφορίες.



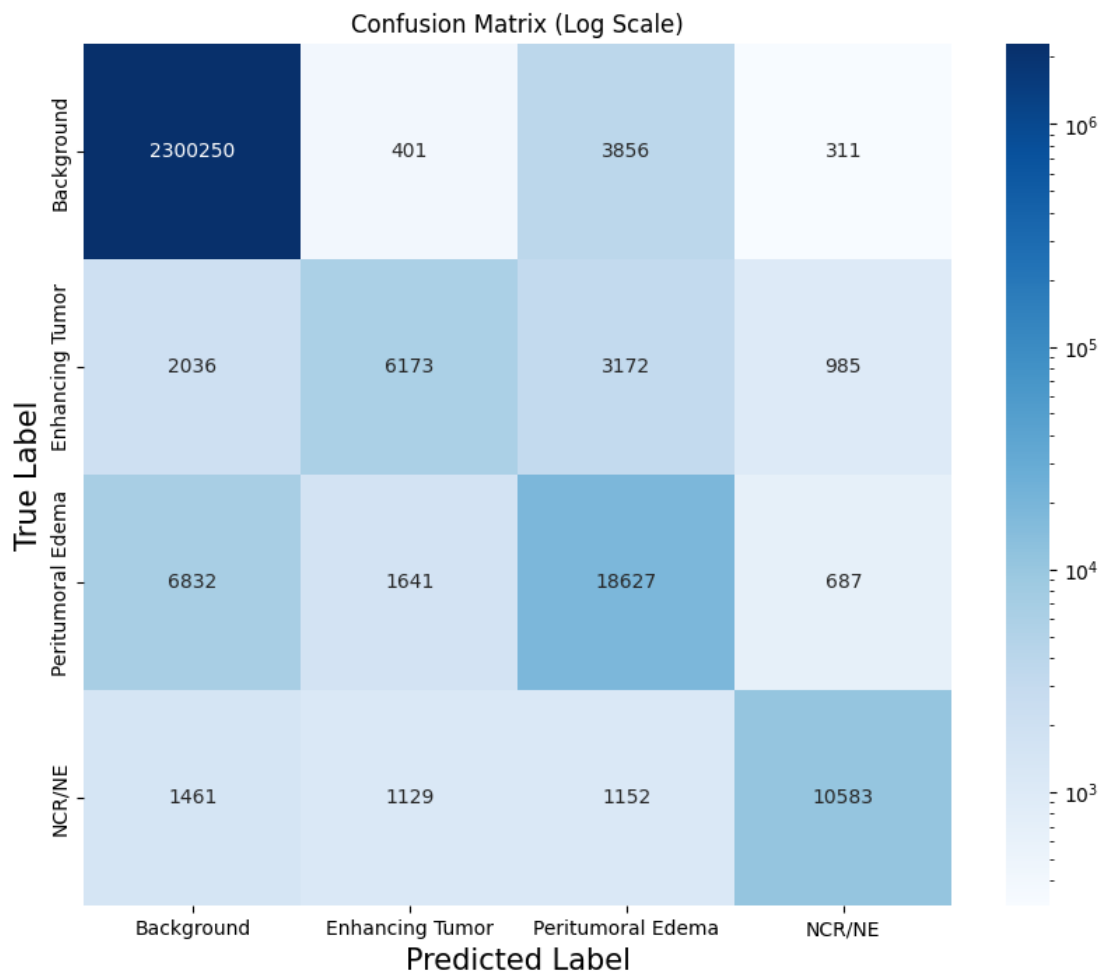
Εικόνα 41: Training και Validation Score κατά την διάρκεια της εκπαίδευσης.

Στο παραπάνω σχήμα φαίνεται ξεκάθαρα η επίδραση της προ εκπαίδευσης, όπου όλα τα μοντέλα ξεκινάνε με loss αρκετά μικρότερο του 1. Επιπλέον, για learning rate $1e^{-3}$ το validation loss, φτάνει το 0.4, ενώ το καλύτερο μοντέλο της συμβατικής προσέγγισης σε πέφτει κάτω από 0.6. Η τελική σύγκριση όμως 2 μοντέλων, η τεχνικών γίνεται στο test dataset. Παρακάτω φαίνεται η απόδοση όλων των μοντέλων που εκπαιδεύτηκαν, με και χωρίς την εφαρμογή προ εκπαίδευσης.



Εικόνα 42: Σύγκριση αποτελεσμάτων με ή χωρίς εφαρμογή προ εκπαίδευσης σε συνθετικά δεδομένα.

Όπως φαίνεται και παραπάνω, η μέθοδος που παρουσιάζεται στην παρούσα διπλωματική δεν αποφέρει μόνο το καλύτερο μοντέλο, με learning rate $1e^{-3}$, αλλά με σε κάθε δυνατό συνδυασμό.



Εικόνα 43: Πίνακας σύγκρισης της απόδοσης του καλύτερου μοντέλου.

Στην παραπάνω εικόνα φαίνεται ο πίνακας σύγκριση του μοντέλου. Τα στοιχεία της διάφωνου φανερώνουν τις σωστές προβλέψεις, ενώ τα υπόλοιπα τις λανθασμένες. Οι μεγαλύτεροι αριθμοί συναντώνται στην διαγώνιο, γεγονός που υποδεικνύει ότι το μοντέλο έχει σχετικά καλή απόδοση. Παρακάτω φαίνονται και οι μετρικές f1 score, recall και precision[97], μετρικές που χρησιμοποιούνται συχνά αντί για το accuracy, στις περιπτώσεις unbalanced dataset, όπως στην περίπτωσή μας, όπου η πλειοψηφία ανήκει στο background class.

	BACKGROUND	ENHANCING TUMOR	PERITUMORAL EDEMA	NCR/NE
PRECISION	0.996	0.661	0.695	0.842
RECALL	0.998	0.499	0.670	0.739
F1-SCORE	0.997	0.569	0.682	0.787

Πίνακας 1: Μετρικές precision, recall και F1-Score για κάθε μία από τις κλάσεις.

6.5.1 Ερμηνεία αποτελεσμάτων

Μια ερώτηση που είναι πιθανό να γεννηθεί σε έναν αναγνώστη είναι το γιατί η τεχνική αυτή να βελτιώνει το αποτέλεσμα. Τα συνθετικά δεδομένα, στην καλύτερη των περιπτώσεων, εάν έχει γίνει δηλαδή τέλεια εκπαίδευση του GAN, θα είναι όμοια με τα αληθινά. Γιατί κανείς να μπει σε όλη αυτή την διαδικασία, όπου θα υπάρχουν μικρές αστοχίες στην εκπαίδευση, για να συνθέσει δεδομένα χειρότερη ποιότητας. Η απάντηση βρίσκεται στην **Εικόνα 32**. Στην περίπτωση αυτή, είναι ξεκάθαρο το φαινόμενο του over fitting. Όταν τα δεδομένα είναι περιορισμένα σε πλήθος, το μοντέλο σχεδόν απομνημονεύει απέξω όλες εικόνες του παρέχονται για εκπαίδευση. Στο train dataset παρατηρείται πολύ καλή πρόοδος, αλλά αυτό δεν μεταφράζεται σε δεδομένα που δεν έχει ξαναδεί. Με την χρήση GAN, θυσιάζεται λίγη ποιότητα δεδομένων, ως προς την ποσότητα.

Συμπεράσματα-Μελλοντικές Προσεγγίσεις

Στην παρούσα εργασία έγινε μία εισαγωγή των βασικών εννοιών, απαραίτητες για την κατανόηση της διαδικασίας τμηματοποίησης ιατρικής εικόνας, συγκεκριμένα μαγνητικής τομογραφίας με την χρήση νευρωνικών δικτύων. Εφαρμόστηκαν τεχνικές για την βελτίωση της συμβατικής προσέγγισης, δηλαδή την χρήση της αρχιτεκτονικής Unet. Συγκεκριμένα, έγινε δοκιμή Generative adversarial networks για την δημιουργία συνθετικών δεδομένων, όσο πιο πιστά στα πραγματικά γινόταν. Τα δεδομένα αυτά στην συνέχεια αξιοποιήθηκαν για την προ εκπαίδευση του.

Τα αποτελέσματα ήταν αρκετά ικανοποιητικά καθώς η παραπάνω διαδικασία φαίνεται να αποφέρει καλύτερα αποτελέσματα όχι σε ένα, αλλά σε όλα τα πειράματα που έλαβαν χώρα. Το γεγονός ενθαρρύνει την περεταίρω εξερεύνηση της παραπάνω τεχνικής. Στην διαδικασία αυτή θα γινόταν δοκιμές παραπάνω αρχιτεκτονικών, τόσο για την τμηματοποίηση όσο και για την παραγωγή δεδομένων, διαφορετικές συναρτήσεις κόστους και τεχνικές προσαύξησης δεδομένων.

Συνοψίζοντας, η παραπάνω εργασία αποτελεί μία ισχυρή ένδειξη ότι η παραγωγή συνθετικών δεδομένων, τόσο στον τομέα της ιατρικής, όσο και σε άλλους τομείς όπου γίνεται χρήση αλγορίθμων μηχανικής μάθησης, έχει την δυνατότητα να συνεισφέρει στην βελτίωση των αποδόσεών τους.

Βιβλιογραφία

- [1] The Institute for Health Metrics and Evaluation (2020). *Global Burden of Disease (GBD)*. [online] www.healthdata.org. Available at: <https://www.healthdata.org/research-analysis/gbd>.
- [2] commission.europa.eu. (n.d.). *A cancer plan for Europe*. [online] Available at: https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/promoting-our-european-way-life/european-health-union/cancer-plan-europe_en.
- [3] Y. KOO *et al.*, “Brain cancer diagnosis and therapy with nanoplatforms☆,” *Adv Drug Deliv Rev*, vol. 58, no. 14, pp. 1556–1577, Dec. 2006, doi: 10.1016/j.addr.2006.09.012.
- [4] M. Lara-Velazquez *et al.*, “Advances in Brain Tumor Surgery for Glioblastoma in Adults,” *Brain Sci*, vol. 7, no. 12, p. 166, Dec. 2017, doi: 10.3390/brainsci7120166.
- [5] F. Yepes-Calderon and J. Gordon McComb, “Manual Segmentation Errors in Medical Imaging. Proposing a Reliable Gold Standard,” 2019, pp. 230–241. doi: 10.1007/978-3-030-32475-9_17.
- [6] J. Cahill, G. LoBiondo-Wood, N. Bergstrom, and T. Armstrong, “Brain Tumor Symptoms as Antecedents to Uncertainty: An Integrative Review,” *Journal of Nursing Scholarship*, vol. 44, no. 2, pp. 145–155, Jun. 2012, doi: 10.1111/j.1547-5069.2012.01445.x.
- [7] M. G. Linguraru *et al.*, “Tumor Burden Analysis on Computed Tomography by Automated Liver and Tumor Segmentation,” *IEEE Trans Med Imaging*, vol. 31, no. 10, pp. 1965–1976, Oct. 2012, doi: 10.1109/TMI.2012.2211887.
- [8] T. C. W. Mok and A. C. S. Chung, “Learning Data Augmentation for Brain Tumor Segmentation with Coarse-to-Fine Generative Adversarial Networks,” May 2018, doi: 10.1007/978-3-030-11723-8_7.
- [9] OpenAI *et al.*, “GPT-4 Technical Report,” Mar. 2023, [Online]. Available: <http://arxiv.org/abs/2303.08774>
- [10] L. Zhou, S. Pan, J. Wang, and A. V. Vasilakos, “Machine learning on big data: Opportunities and challenges,” *Neurocomputing*, vol. 237, pp. 350–361, May 2017, doi: 10.1016/j.neucom.2017.01.026.
- [11] J. Boutry *et al.*, “The evolution and ecology of benign tumors,” *Biochimica et Biophysica Acta (BBA) - Reviews on Cancer*, vol. 1877, no. 1, p. 188643, Jan. 2022, doi: 10.1016/J.BBCAN.2021.188643.
- [12] L. M. DeAngelis, “Brain Tumors,” *New England Journal of Medicine*, vol. 344, no. 2, pp. 114–123, Jan. 2001, doi: 10.1056/NEJM200101113440207.
- [13] M. Kanda *et al.*, “Pattern of Lymph Node Metastasis Spread in Pancreatic Cancer,” *Pancreas*, vol. 40, no. 6, pp. 951–955, Aug. 2011, doi: 10.1097/MPA.0b013e3182148342.
- [14] S. H. Jafari *et al.*, “Breast cancer diagnosis: Imaging techniques and biochemical markers,” *J Cell Physiol*, vol. 233, no. 7, pp. 5200–5213, Jul. 2018, doi: 10.1002/jcp.26379.
- [15] R. L. Siegel, K. D. Miller, H. E. Fuchs, and A. Jemal, “Cancer statistics, 2022,” *CA Cancer J Clin*, vol. 72, no. 1, pp. 7–33, Jan. 2022, doi: 10.3322/caac.21708.

- [16] S. Yang, T. Liu, Y. Cheng, Y. Bai, and G. Liang, "Immune cell infiltration as a biomarker for the diagnosis and prognosis of digestive system cancer," *Cancer Sci*, vol. 110, no. 12, pp. 3639–3649, Dec. 2019, doi: 10.1111/cas.14216.
- [17] J. D. Minna, J. A. Roth, and A. F. Gazdar, "Focus on lung cancer," *Cancer Cell*, vol. 1, no. 1, pp. 49–52, Feb. 2002, doi: 10.1016/S1535-6108(02)00027-2.
- [18] M. Akram, M. Iqbal, M. Daniyal, and A. U. Khan, "Awareness and current knowledge of breast cancer," *Biol Res*, vol. 50, no. 1, p. 33, Dec. 2017, doi: 10.1186/s40659-017-0140-9.
- [19] R. H. Hruban, G. M. Petersen, P. K. Ha, and S. E. Kern, "Genetics of Pancreatic Cancer," *Surg Oncol Clin N Am*, vol. 7, no. 1, pp. 1–23, Jan. 1998, doi: 10.1016/S1055-3207(18)30282-5.
- [20] K. N. Nathanson, R. Wooster, and B. L. Weber, "Breast cancer genetics: What we know and what we need," *Nat Med*, vol. 7, no. 5, pp. 552–556, May 2001, doi: 10.1038/87876.
- [21] A. R. Venkitaraman, "How do mutations affecting the breast cancer genes BRCA1 and BRCA2 cause cancer susceptibility?," *DNA Repair (Amst)*, vol. 81, p. 102668, Sep. 2019, doi: 10.1016/j.dnarep.2019.102668.
- [22] A. J. Sasco, M. B. Secretan, and K. Straif, "Tobacco smoking and cancer: a brief review of recent epidemiological evidence," *Lung Cancer*, vol. 45, pp. S3–S9, Aug. 2004, doi: 10.1016/j.lungcan.2004.07.998.
- [23] E. Weiderpass, "Lifestyle and Cancer Risk," *Journal of Preventive Medicine and Public Health*, vol. 43, no. 6, p. 459, 2010, doi: 10.3961/jpmph.2010.43.6.459.
- [24] V. A. Katzke, R. Kaaks, and T. Kühn, "Lifestyle and Cancer Risk," *The Cancer Journal*, vol. 21, no. 2, pp. 104–110, Mar. 2015, doi: 10.1097/PPO.000000000000101.
- [25] J. Campisi, J. K. Andersen, P. Kapahi, and S. Melov, "Cellular senescence: A link between cancer and age-related degenerative disease?," *Semin Cancer Biol*, vol. 21, no. 6, pp. 354–359, Dec. 2011, doi: 10.1016/J.SEMCANCER.2011.09.001.
- [26] Mayo Clinic Staff (2022). *Cancer - Symptoms and Causes*. [online] Mayo Clinic. Available at: <https://www.mayoclinic.org/diseases-conditions/cancer/symptoms-causes/syc-20370588>.
- [27] Mayo Clinic (2019). *Brain Tumor - Diagnosis and Treatment - Mayo Clinic*. [online] Mayo Clinic. Available at: <https://www.mayoclinic.org/diseases-conditions/brain-tumor/diagnosis-treatment/drc-20350088>.
- [28] L. Wang, "Early Diagnosis of Breast Cancer," *Sensors*, vol. 17, no. 7, p. 1572, Jul. 2017, doi: 10.3390/s17071572.
- [29] E. S. McDonald, A. S. Clark, J. Tchou, P. Zhang, and G. M. Freedman, "Clinical Diagnosis and Management of Breast Cancer," *Journal of Nuclear Medicine*, vol. 57, no. Supplement 1, pp. 9S-16S, Feb. 2016, doi: 10.2967/jnumed.115.157834.
- [30] D. M. Aleman, A. Kumar, R. K. Ahuja, H. E. Romeijn, and J. F. Dempsey, "Neighborhood search approaches to beam orientation optimization in intensity modulated radiation therapy treatment planning," *Journal of Global Optimization*, vol. 42, no. 4, pp. 587–607, Dec. 2008, doi: 10.1007/s10898-008-9286-x.

- [31] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," May 2015, [Online]. Available: <http://arxiv.org/abs/1505.04597>
- [32] www.med.upenn.edu. (n.d.). *Multimodal Brain Tumor Segmentation Challenge 2020: Data | CBICA | Perelman School of Medicine at the University of Pennsylvania*. [online] Available at: <https://www.med.upenn.edu/cbica/brats2020/data.html>.
- [33] G.-P. Yan, L. Robinson, and P. Hogg, "Magnetic resonance imaging contrast agents: Overview and perspectives," *Radiography*, vol. 13, pp. e5–e19, Dec. 2007, doi: 10.1016/j.radi.2006.07.005.
- [34] W. G. B. C. J. L. Ray Hashman Hashemi, *MRI: The Basics: The Basics* .
- [35] D. B. Plewes and W. Kucharczyk, "Physics of MRI: A primer," *Journal of Magnetic Resonance Imaging*, vol. 35, no. 5, pp. 1038–1054, May 2012, doi: 10.1002/jmri.23642.
- [36] Catherine Westbrook and John Talbot, *MRI in Practice*.
- [37] H. A. Sharma, "MRI physics–basic principles," *Acta Neuropsychiatr*, vol. 21, no. 4, pp. 200–201, Aug. 2009, doi: 10.1111/j.1601-5215.2009.00404.x.
- [38] *Precession: The alignment of nuclei with an applied magnetic field is not stagnant*. ResearchGate. Available at: https://www.researchgate.net/figure/Precession-The-alignment-of-nuclei-with-an-applied-magnetic-field-is-not-stagnant_fig1_323758452
- [39] T. Yousaf, G. Dervenoulas, and M. Politis, "Advances in MRI Methodology," *Int Rev Neurobiol*, vol. 141, pp. 31–76, Jan. 2018, doi: 10.1016/BS.IRN.2018.08.008.
- [40] R.-J. M. van Geuns *et al.*, "Basic principles of magnetic resonance imaging☆," *Prog Cardiovasc Dis*, vol. 42, no. 2, pp. 149–156, Sep. 1999, doi: 10.1016/S0033-0620(99)70014-9.
- [41] I. Young, "Nuclear magnetic resonance imaging," *Electronics and Power*, vol. 30, no. 3, p. 205, 1984, doi: 10.1049/ep.1984.0112.
- [42] V. Mlynárik, "Introduction to nuclear magnetic resonance," *Anal Biochem*, vol. 529, pp. 4–9, Jul. 2017, doi: 10.1016/J.AB.2016.05.006.
- [43] C. B. Paschal and H. D. Morris, "K-space in the clinic," *Journal of Magnetic Resonance Imaging*, vol. 19, no. 2, pp. 145–159, Feb. 2004, doi: 10.1002/jmri.10451.
- [44] Ridgway, J.P. Cardiovascular magnetic resonance physics for clinicians: part I. *J Cardiovasc Magn Reson* **12**, 71 (2010). <https://doi.org/10.1186/1532-429X-12-71>.
- [45] D. Moratal, A. Valles-Luch, L. Marti-Bonmati, and M. Brummer, "k-Space tutorial: an MRI educational tool for a better understanding of k-space," *Biomed Imaging Interv J*, vol. 4, no. 1, Jan. 2008, doi: 10.2349/bij.4.1.e15.
- [46] Slicer (2019). *3D Slicer*. [online] Slicer.org. Available at: <https://www.slicer.org/>.
- [47] C. R. S. Banerji, T. Mansour, and S. Severini, "A notion of graph likelihood and an infinite monkey theorem," *J Phys A Math Theor*, vol. 47, no. 3, p. 035101, Jan. 2014, doi: 10.1088/1751-8113/47/3/035101.
- [48] J. A. Anderson, *An introduction to neural networks*. 1995.
- [49] "wikipedia neuron."

- [50] cs.stanford.edu. (n.d.). *Neural Networks - Biology*. [online] Available at: <https://cs.stanford.edu/people/eroberts/courses/soco/projects/neural-networks/Biology/index.html>.
- [51] D. O. Hebb, *The Organization of Behavior: A Neuropsychological Theory*. 1949.
- [52] Beren Millidge, Tommaso Salvatori, Yuhang Song, Rafal Bogacz, and Thomas Lukasiewicz, "Predictive Coding: Towards a Future of Deep Learning beyond Backpropagation?," Feb. 2022.
- [53] P. Christoffersen and K. Jacobs, "The importance of the loss function in option valuation," *J financ econ*, vol. 72, no. 2, pp. 291–318, May 2004, doi: 10.1016/j.jfineco.2003.02.001.
- [54] Simon Haykin, *Neural Networks and Learning Machines*. 2009.
- [55] M. W. Gardner and S. R. Dorling, "Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences," *Atmos Environ*, vol. 32, no. 14–15, pp. 2627–2636, Aug. 1998, doi: 10.1016/S1352-2310(97)00447-0.
- [56] Abirami, S. and Chitra, P. (2020). *Multilayer Perceptron - an overview | ScienceDirect Topics*. [online] www.sciencedirect.com. Available at: <https://www.sciencedirect.com/topics/computer-science/multilayer-perceptron>.
- [57] C. Eckert, A. Subramaniyan, X. Wang, C. Augustine, R. Iyer, and R. Das, "Eidetic: An In-Memory Matrix Multiplication Accelerator for Neural Networks," *IEEE Transactions on Computers*, vol. 72, no. 6, pp. 1539–1553, Jun. 2023, doi: 10.1109/TC.2022.3214151.
- [58] LeCun, Y. (2009). *MNIST handwritten digit database, Yann LeCun, Corinna Cortes and Chris Burges*. [online] Lecun.com. Available at: <http://yann.lecun.com/exdb/mnist/>.
- [59] TensorFlow Datasets (2010). *mnist | TensorFlow Datasets*. [online] TensorFlow. Available at: <https://www.tensorflow.org/datasets/catalog/mnist>.
- [60] N. Ketkar, "Stochastic Gradient Descent," in *Deep Learning with Python*, Berkeley, CA: Apress, 2017, pp. 113–132. doi: 10.1007/978-1-4842-2766-4_8.
- [61] Nelson, D. (2020). *What is Gradient Descent?* [online] Unite.AI. Available at: <https://www.unite.ai/el/what-is-gradient-descent/>.
- [62] 产品经理的人工智能学习库. (n.d.). *Gradient Descent - Gradient descent - Product Manager's Artificial Intelligence Learning Library*. [online] Available at: <https://easyai.tech/en/ai-definition/gradient-descent/>.
- [63] M. Nielsen, "Neural Networks and Deep Learning." [Online]. Available: <http://neuralnetworksanddeeplearning.com>
- [64] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," in *2017 International Conference on Engineering and Technology (ICET)*, IEEE, Aug. 2017, pp. 1–6. doi: 10.1109/ICEngTechnol.2017.8308186.
- [65] Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, "A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects," *IEEE Trans Neural Netw Learn Syst*, vol. 33, no. 12, pp. 6999–7019, Dec. 2022, doi: 10.1109/TNNLS.2021.3084827.
- [66] Michael D. Adams, Ed., *Signals and Systems*, Edition 3.0.

- [67] N. Ketkar and J. Moolayil, "Convolutional Neural Networks," in *Deep Learning with Python*, Berkeley, CA: Apress, 2021, pp. 197–242. doi: 10.1007/978-1-4842-5364-9_6.
- [68] Towards Data Science (n.d.) 'The W3H of AlexNet, VGGNet, ResNet, and Inception', *Medium*. Available at: <https://medium.com/towards-data-science/the-w3h-of-alexnet-vggnet-resnet-and-inception-7baaecccc96>.
- [69] UpGrad (n.d.) 'Image Segmentation Techniques', *UpGrad Blog*. Available at: <https://www.upgrad.com/blog/image-segmentation-techniques/>.
- [70] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition." [Online]. Available: <http://image-net.org/challenges/LSVRC/2015/>
- [71] I. J. Goodfellow *et al.*, "Generative Adversarial Networks," Jun. 2014, [Online]. Available: <http://arxiv.org/abs/1406.2661>
- [72] *2018 IEEE High Performance Extreme Computing Conference (HPEC) : the Westin-Hotel Waltham-Boston, 70 Third Avenue, Waltham, Massachusetts USA, 25-27 September 2018*. IEEE, 2018.
- [73] Cuicaihao (2017) 'Sharing the opinion about Generative Adversarial Networks (GAN)', *Cuicaihao Blog*. Available at: <https://cuicaihao.com/2017/05/29/sharing-the-opinion-about-generative-adversarial-networks-gan/>.
- [74] Y. Ho and S. Wookey, "The Real-World-Weight Cross-Entropy Loss Function: Modeling the Costs of Mislabeling," *IEEE Access*, vol. 8, pp. 4806–4813, 2020, doi: 10.1109/ACCESS.2019.2962617.
- [75] M. Arjovsky, S. Chintala, and L. Bottou, "Wasserstein GAN," Jan. 2017, [Online]. Available: <http://arxiv.org/abs/1701.07875>
- [76] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. Courville, "Improved Training of Wasserstein GANs," Mar. 2017, [Online]. Available: <http://arxiv.org/abs/1704.00028>
- [77] T. Karras, T. Aila, S. Laine, and J. Lehtinen, "Progressive Growing of GANs for Improved Quality, Stability, and Variation," Oct. 2017, [Online]. Available: <http://arxiv.org/abs/1710.10196>
- [78] M. Usman Akbar, M. Larsson, I. Blystad, and A. Eklund, "Brain tumor segmentation using synthetic MR images - A comparison of GANs and diffusion models," *Sci Data*, vol. 11, no. 1, Dec. 2024, doi: 10.1038/s41597-024-03073-x.
- [79] Aladdin Persson (n.d.) *Machine Learning Collection*. Available at: <https://github.com/aladdinpersson/Machine-Learning-Collection>.
- [80] M. Wang, S. Lu, D. Zhu, J. Lin, and Z. Wang, "A High-Speed and Low-Complexity Architecture for Softmax Function in Deep Learning," in *2018 IEEE Asia Pacific Conference on Circuits and Systems (APCCAS)*, IEEE, Oct. 2018, pp. 223–226. doi: 10.1109/APCCAS.2018.8605654.
- [81] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks," Mar. 2017, [Online]. Available: <http://arxiv.org/abs/1703.10593>
- [82] V. Sandfort, K. Yan, P. J. Pickhardt, and R. M. Summers, "Data augmentation using generative adversarial networks (CycleGAN) to improve generalizability

- in CT segmentation tasks,” *Sci Rep*, vol. 9, no. 1, Dec. 2019, doi: 10.1038/s41598-019-52737-x.
- [83] C. Bowles *et al.*, “GAN Augmentation: Augmenting Training Data using Generative Adversarial Networks,” Oct. 2018, [Online]. Available: <http://arxiv.org/abs/1810.10863>
- [84] S. Dhivya, S. Mohanavalli, S. Karthika, S. Shivani, and R. Mageswari, “GAN based Data Augmentation for Enhanced Tumor Classification,” in *4th International Conference on Computer, Communication and Signal Processing, ICCSP 2020*, Institute of Electrical and Electronics Engineers Inc., Sep. 2020. doi: 10.1109/ICCSP49186.2020.9315189.
- [85] T. C. W. Mok and A. C. S. Chung, “Learning Data Augmentation for Brain Tumor Segmentation with Coarse-to-Fine Generative Adversarial Networks,” May 2018, doi: 10.1007/978-3-030-11723-8_7.
- [86] E. Carver, Z. Dai, E. Liang, J. Snyder, and N. Wen, “Improvement of Multiparametric MR Image Segmentation by Augmenting the Data with Generative Adversarial Networks for Glioma Patients.”
- [87] J. Gauthier, “Conditional generative adversarial nets for convolutional face generation.”
- [88] C. Han *et al.*, “Infinite Brain MR Images: PGGAN-based Data Augmentation for Tumor Detection,” Mar. 2019, [Online]. Available: <http://arxiv.org/abs/1903.12564>
- [89] A. Kamli, R. Saouli, H. Batatia, M. Ben Naceur, and I. Youkana, “Synthetic medical image generator for data augmentation and anonymisation based on generative adversarial network for glioblastoma tumors growth prediction,” *IET Image Process*, vol. 14, no. 16, Dec. 2020, doi: 10.1049/iet-ipr.2020.1141.
- [90] V. Thambawita *et al.*, “SinGAN-Seg: Synthetic training data generation for medical image segmentation,” *PLoS One*, vol. 17, no. 5 May, May 2022, doi: 10.1371/journal.pone.0267976.
- [91] J. Kugelman, D. Alonso-Caneiro, S. A. Read, S. J. Vincent, F. K. Chen, and M. J. Collins, “Dual image and mask synthesis with GANs for semantic segmentation in optical coherence tomography,” in *2020 Digital Image Computing: Techniques and Applications, DICTA 2020*, Institute of Electrical and Electronics Engineers Inc., Nov. 2020. doi: 10.1109/DICTA51227.2020.9363402.
- [92] S. Jadon, “A survey of loss functions for semantic segmentation,” in *2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, IEEE, Oct. 2020, pp. 1–7. doi: 10.1109/CIBCB48159.2020.9277638.
- [93] S. Jha, L. H. Son, R. Kumar, I. Priyadarshini, F. Smarandache, and H. V. Long, “Neutrosophic image segmentation with Dice Coefficients,” *Measurement*, vol. 134, pp. 762–772, Feb. 2019, doi: 10.1016/j.measurement.2018.11.006.
- [94] T. Dietterich, “Overfitting and Undercomputing in Machine Learning.”
- [95] A. Mumuni and F. Mumuni, “Data augmentation: A comprehensive survey of modern approaches,” *Array*, vol. 16, p. 100258, Dec. 2022, doi: 10.1016/j.array.2022.100258.
- [96] X. Zhang, S. Karaman, and S.-F. Chang, “Detecting and Simulating Artifacts in GAN Fake Images,” in *2019 IEEE International Workshop on Information*

- Forensics and Security (WIFS)*, IEEE, Dec. 2019, pp. 1–6. doi:
10.1109/WIFS47025.2019.9035107.
- [97] Labelf AI (n.d.) 'What is Accuracy, Precision, Recall, and F1 Score', *Labelf AI Blog*. Available at: <https://www.labelf.ai/blog/what-is-accuracy-precision-recall-and-f1-score>.