



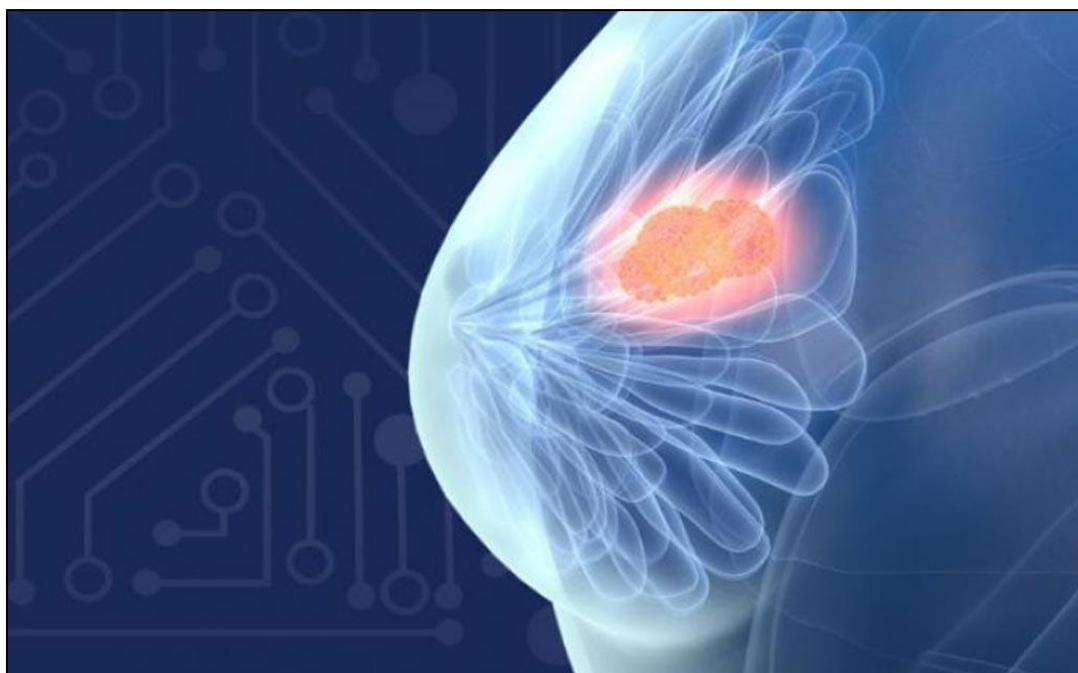
ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ

ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ

ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ & ΗΛΕΚΤΡΟΝΙΚΩΝ ΜΗΧΑΝΙΚΩΝ

Διπλωματική Εργασία

Ανάλυση βιοϊατρικών εικόνων για την ανίχνευση καρκίνου του μαστού
με χρήση αλγορίθμων μηχανικής μάθησης



Φοιτήτρια: Δατσώλη Κωνσταντίνα

ΑΜ: 262017042

Επιβλέπουσα Καθηγήτρια

Μαρία Ραγκούση

Καθηγήτρια

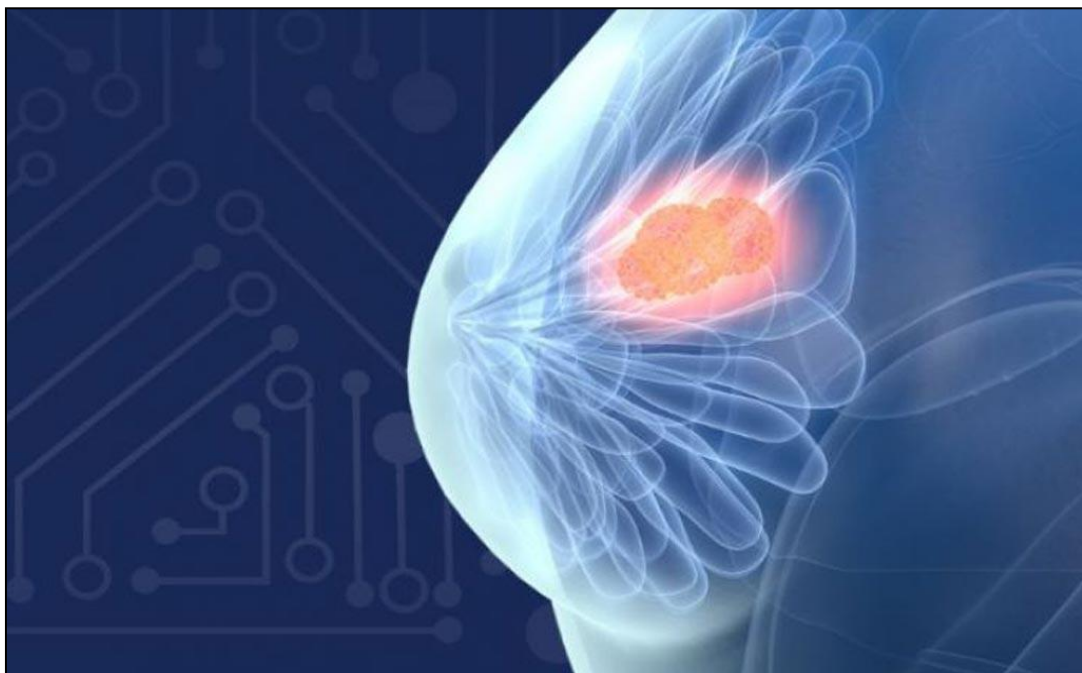
ΑΘΗΝΑ-ΑΙΓΑΛΕΩ, ΙΟΥΝΙΟΣ 2024



UNIVERSITY OF WEST ATTICA
FACULTY OF ENGINEERING
DEPARTMENT OF ELECTRICAL & ELECTRONICS ENGINEERING

Diploma Thesis

Biomedical image analysis for breast cancer diagnosis using machine learning algorithms



Student: Datsoli Konstantina

Registration Number: 262017042

Supervisor

Maria Rangoussi

Professor

ATHENS-EGALEO, JUNE 2024

Η Διπλωματική Εργασία έγινε αποδεκτή και βαθμολογήθηκε από την εξής τριμελή επιτροπή:

Μαρία Ραγκούση, Καθηγήτρια	Νικόλαος-Αλέξανδρος Τάτλας Καθηγητής	Ηλίας Ζώης, Αν. Καθηγητής
(Υπογραφή)	(Υπογραφή)	(Υπογραφή)

Copyright © Με επιφύλαξη παντός δικαιώματος. All rights reserved.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ και **Δατσώλη Κωνσταντίνα, Ιούνιος, 2024**

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

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον/την συγγραφέα του και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις θέσεις του επιβλέποντος, της επιτροπής εξέτασης ή τις επίσημες θέσεις του Τμήματος και του Ιδρύματος.

ΔΗΛΩΣΗ ΣΥΓΓΡΑΦΕΑ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ

Η κάτωθι υπογεγραμμένη **Δατσώλη Κωνσταντίνα** του **Αποστόλου**, με αριθμό μητρώου **262017042** φοιτήτρια του Πανεπιστημίου Δυτικής Αττικής της Σχολής ΜΗΧΑΝΙΚΩΝ του Τμήματος ΗΛΕΚΤΡΟΛΟΓΩΝ ΚΑΙ ΗΛΕΚΤΡΟΝΙΚΩΝ ΜΗΧΑΝΙΚΩΝ,

δηλώνω υπεύθυνα ότι:

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

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

Η Δηλούσα



Δατσώλη Κωνσταντίνα

ΕΥΧΑΡΙΣΤΙΕΣ

Με την ολοκλήρωση αυτής της διπλωματικής εργασίας, θα ήθελα να εκφράσω τις ειλικρινείς μου ευχαριστίες προς όλους τους ανθρώπους που με στήριξαν και με βοήθησαν κατά τη διάρκεια αυτής της ακαδημαϊκής πορείας.

Κατ' αρχάς, θέλω να ευχαριστήσω θερμά την καθηγήτριά μου, κ. Μαρία Ραγκούση, για την αφοσίωσή της και την καθοδήγησή της κατά τη διάρκεια της έρευνάς μου. Οι συμβουλές της και η συνεχής στήριξή της με βοήθησαν να ολοκληρώσω αυτήν την εργασία με επιτυχία και της είμαι ευγνώμων όχι μόνο για την ακαδημαϊκή, αλλά και για την ανθρώπινή της προσφορά.

Επίσης, θέλω να εκφράσω την ευγνωμοσύνη μου προς όλους τους καθηγητές του Πανεπιστημίου Δυτικής Αττικής για τις απαραίτητες βάσεις και τις γνώσεις που μου προσέφεραν κατά τη διάρκεια της φοίτησής μου. Χάρη σε αυτούς, κατάφερα να φτάσω στο σημείο που βρίσκομαι σήμερα και να ολοκληρώσω τις σπουδές μου με επιτυχία.

Δεν μπορώ να μην ευχαριστήσω ακόμα θερμά όλους τους φίλους και συμφοιτητές μου που με στήριξαν από το πρώτο έτος των σπουδών μου έως και την ολοκλήρωση αυτής της διπλωματικής εργασίας. Η συμπαράστασή τους, οι συζητήσεις μας και η αλληλοβοήθεια αποτέλεσαν τεράστιο κίνητρο για εμένα. Χωρίς την αμοιβαία υποστήριξη και τη φιλία μας, δεν θα είχα προχωρήσει τόσο μακριά. Ευχαριστώ τον καθένα από εσάς προσωπικά.

Τέλος, αφιερώνω αυτήν την εργασία στην οικογένειά μου που με την αγάπη και την συνεχή υποστήριξή τους με ενέπνευσαν και με ενθάρρυναν να φτάσω ως εδώ. Λίγο περισσότερο όμως, θα ήθελα να την αφιερώσω στην αγαπημένη μου μητέρα Παναγιώτα, που με την τεράστια δύναμη και ψυχική αντοχή της νίκησε τον καρκίνο του μαστού, αλλά και στον θείο και νονό μου Στέλιο, που αγωνίζεται για τρίτη φορά με τον καρκίνο και είμαι σίγουρη πως θα βγει και πάλι νικητής.

Περίληψη

Η παρούσα διπλωματική εργασία ασχολείται με την επεξεργασία βιοϊατρικών εικόνων, ειδικότερα υπερηχογραφήματων μαστού, και την ανάλυσή τους μέσω της εκπαίδευσης νευρωνικών δικτύων βαθιάς μάθησης για την αναγνώριση παθολογίας και την κατηγοριοποίησή τους σε εικόνες υγιείς ή παθολογικές. Η επεξεργασία και ανάλυση πραγματοποιείται στο περιβάλλον MATLAB®.

Αρχικά, παρουσιάζονται διάφορες μεθόδους επεξεργασίας εικόνων, οι οποίες αξιολογούνται για την ικανότητά τους να απλοποιούν τα δεδομένα και να αναδεικνύουν κρίσιμα χαρακτηριστικά τους, διευκολύνοντας κατ' αυτόν τον τρόπο την ορθή αναγνώριση των εικόνων από τα νευρωνικά δίκτυα.

Στη συνέχεια, εξετάζονται διεξοδικά διάφοροι τύποι νευρωνικών δικτύων, όπως τα νευρωνικά δίκτυα πρόσθιας τροφοδότησης (FNN) και τα συνελκτικά νευρωνικά δίκτυα (CNN), με σκοπό την εκτίμηση της απόδοσής τους στην ανάλυση των υπερηχογραφήματων μαστού. Μέσω πειραματικών διαδικασιών, διερευνώνται οι επιπτώσεις διαφόρων παραμετροποιήσεων των δικτύων και αξιολογούνται τα αποτελέσματα της εκπαίδευσής τους, παρέχοντας έτσι μια ολοκληρωμένη επισκόπηση της αποτελεσματικότητας των προτεινόμενων τεχνικών στην ιατρική διάγνωση στο πεδίο των μαστογραφιών.

Λέξεις – κλειδιά

Βιοϊατρική εικόνα, Ψηφιακή επεξεργασία εικόνας, Νευρωνικά δίκτυα, Βαθιά μάθηση, CNN, FNN, Καρκίνος του μαστού, Υπέρηχος μαστού

Abstract

This thesis focuses on the digital processing of biomedical images, specifically breast ultrasound images, and their analysis through the training of deep neural networks (DNN) aimed at the recognition of breast cancer cases and at the classification of images into healthy or cancerous classes. Processing and analysis are performed in MATLAB ®.

Initially, various image processing methods are presented and discussed. These are evaluated as to their ability to simplify data and to extract critical and informative features, thus facilitating the accurate recognition by neural networks.

Subsequently, different types of neural networks, such as Feedforward Neural Networks (FNN) and Convolutional Neural Networks (CNN), are examined in detail to assess their performance in the analysis of breast ultrasound images. Through experimental procedures, the effects of various network parameterizations are investigated and the results of the training of the respective NNs are evaluated, thereby providing a comprehensive overview of the effectiveness of the proposed techniques in medical diagnosis.

Keywords

Biomedical image, Digital image processing, Neural networks, Deep learning, CNN, FNN, Breast cancer, Breast ultrasound

Περιεχόμενα

Κατάλογος Πινάκων	10
Κατάλογος Εικόνων.....	11
Αλφαβητικό Ευρετήριο	15
ΕΙΣΑΓΩΓΗ	17
Αντικείμενο της διπλωματικής εργασίας.....	17
Σκοπός και στόχοι	17
Μεθοδολογία.....	17
Καινοτομία.....	17
Δομή.....	18
ΚΕΦΑΛΑΙΟ 1^ο: Επιστημονική και Ιστορική Ανάλυση του Καρκίνου του Μαστού.....	19
1.1 Βασικές Έννοιες και Χαρακτηριστικά.....	19
1.1.1 Ο καρκίνος του μαστού	20
1.1.2 Αιτίες του καρκίνου του μαστού	20
1.1.3 Συμπτώματα του καρκίνου του μαστού	22
1.2 Μέθοδοι Διάγνωσης και Απεικόνισης	24
1.3 Ιστορική Αναδρομή.....	25
1.3.1 Αρχαία Ελλάδα και Αίγυπτος.....	25
1.3.2 Θεωρίες και Θεραπευτικές Προσεγγίσεις (17 ^{ος} – 20 ^{ος} αιώνας)	25
1.3.3 Οι σημερινές προσεγγίσεις για την θεραπεία του καρκίνου του μαστού.....	26
ΚΕΦΑΛΑΙΟ 2^ο: Επεξεργασία ψηφιακών βιοϊατρικών εικόνων.....	29
2.1 Ορισμός Εικόνας και Διαστάσεις.....	29
2.1.1 Δυαδική Εικόνα	30
2.1.2 Μονοχρωματική Εικόνα.....	30
2.1.3 Έγχρωμη Εικόνα	31
2.2 Βελτίωση Ψηφιακής Εικόνας στο Πεδίο της Συχνότητας	33
2.2.1 Ο Μετασχηματισμός Fourier.....	34
2.2.2 Φίλτρα Συχνοτήτων.....	36
2.2.2.1 Χαμηλοπερατό (Low Pass) Φίλτρο	36
2.2.2.2 Υψηλοπερατό (High Pass) Φίλτρο	37
2.2.2.3 Γκαουσιανό (Gaussian) Φίλτρο	37
2.3 Επεξεργασία Εικόνων στην Βιοϊατρική	38
2.3.1 Ποιότητα της Εικόνας	39
2.3.2 Μέθοδοι Επεξεργασίας Εικόνων	40
ΚΕΦΑΛΑΙΟ 3^ο: Νευρωνικά δίκτυα βαθιάς μάθησης.....	43
3.1 Τεχνητά Νευρωνικά Δίκτυα (Artificial Neural Networks – ANNs)	43
3.1.1 Συνάρτηση Ενεργοποίησης Νευρώνα	44
3.1.2 Δομή Νευρωνικών Δικτύων	45
3.2 Νευρωνικά Δίκτυα Βαθιάς Μάθησης (Deep Neural Networks – DNNs).....	45
3.2.1 Κατηγορίες Deep Neural Networks	46
3.2.2 Αξιολόγηση της Απόδοσης μέσω Confusion Matrix	47
3.3 Νευρωνικά Δίκτυα Πρόσθιας Τροφοδότησης (Feedforward Neural Networks – FNNs)	49
3.3.1 Δίκτυα FNN ενός κρυφού επιπέδου	49
3.3.2 Δίκτυα FNN πολλαπλών κρυφών επιπέδων.....	50
3.4 Συνελκτικά Νευρωνικά Δίκτυα (Convolutional Neural Networks – CNNs).....	50

3.4.1 Συνέλιξη	53
3.4.2 Παράμετροι ενός CNN	53
ΚΕΦΑΛΑΙΟ 4ο: Πειραματικό μέρος - σχεδίαση έρευνας και data set	55
4.1 Σύνολο Δεδομένων.....	56
4.2 Δημιουργία και Σύγκριση Μασκών	57
4.2.1 Μέθοδος κατωφλίου και CLAHE	57
4.2.2 Μέθοδος κατωφλίου Otsu	59
4.3 Προετοιμασία και Προ-επεξεργασία Δεδομένων.....	60
4.3.1 Δημιουργία dataset με χρήση μετασχηματισμού FFT-2D και συχνοτικού φίλτρου	65
4.3.2 Δημιουργία dataset με άθροιση των αρχικών και των φιλτραρισμένων εικόνων	70
4.3.3 Επίπεδα Νευρωνικών Δικτύων.....	71
4.3.4 Παράμετροι Εκπαίδευσης (Training Options)	72
4.3.5 Εκπαίδευση και Αξιολόγηση του Δικτύου.....	74
ΚΕΦΑΛΑΙΟ 5ο: Ανάλυση με Feedforward Neural Network	76
5.1 Feedforward Neural Network (FNN)	76
5.2 Αρχιτεκτονική του FNN και Ρυθμίσεις Εκπαίδευσης	76
5.3 Εικόνες χωρίς επεξεργασία.....	78
5.3.1 Διάγραμμα Ακρίβειας.....	78
5.3.2 Διάγραμμα Απώλειας	78
5.3.3 Μετρήσεις Απόδοσης.....	79
5.3.4 Confusion Matrix.....	80
5.4 Εικόνες με FFT-2D, συχνοτικό φίλτρο και IFFT-2D.....	81
5.4.1 Διάγραμμα Ακρίβειας.....	81
5.4.2 Διάγραμμα Απώλειας	81
5.4.3 Μετρήσεις Απόδοσης.....	82
5.4.4 Confusion Matrix.....	83
5.5 Εικόνες από άθροιση των αρχικών και των φιλτραρισμένων.....	84
5.5.1 Διάγραμμα Ακρίβειας.....	84
5.5.2 Διάγραμμα Απώλειας	84
5.5.3 Μετρήσεις Απόδοσης.....	85
5.5.4 Confusion Matrix.....	86
ΚΕΦΑΛΑΙΟ 6ο: Ανάλυση με Convolutional Neural Network	87
6.1 Convolutional Neural Network (CNN).....	87
6.2 CNN – 1^η Αρχιτεκτονική	87
6.2.1 Εικόνες χωρίς επεξεργασία	89
6.2.1.1 Διάγραμμα Ακρίβειας	89
6.2.1.2 Διάγραμμα Απώλειας.....	90
6.2.1.3 Μετρήσεις απόδοσης	90
6.2.1.4 Confusion Matrix	91
6.2.2 Εικόνες με FFT-2D, συχνοτικό φίλτρο και IFFT-2D.....	92
6.2.2.1 Διάγραμμα Ακρίβειας	92
6.2.2.2 Διάγραμμα Απώλειας.....	92
6.2.2.3 Μετρήσεις απόδοσης	94
6.2.2.4 Confusion Matrix	95
6.2.3 Εικόνες από άθροιση των αρχικών και των φιλτραρισμένων εικόνων	96
6.2.3.1 Διάγραμμα Ακρίβειας	96
6.2.3.2 Διάγραμμα Απώλειας.....	96

6.2.3.3 Μετρήσεις Απόδοσης	97
6.2.3.4 Confusion Matrix	98
6.3 CNN – 2^η Αρχιτεκτονική	99
6.3.1 Εικόνες χωρίς επεξεργασία	100
6.3.1.1 Διάγραμμα Ακρίβειας	101
6.3.1.2 Διάγραμμα Απώλειας.....	101
6.3.1.3 Μετρήσεις απόδοσης	101
6.3.1.4 Confusion Matrix	102
6.3.2 Εικόνες με FFT-2D, συχνοτικό φίλτρο και IFFT-2D.....	104
6.3.2.1 Διάγραμμα Ακρίβειας	104
6.3.2.2 Διάγραμμα Απώλειας.....	104
6.3.2.3 Μετρήσεις Απόδοσης	105
6.3.2.4 Confusion Matrix	106
6.3.3 Εικόνες από άθροιση των αρχικών και των φιλτραρισμένων	107
6.3.3.1 Διάγραμμα Ακρίβειας	107
6.3.3.2 Διάγραμμα Απώλειας.....	107
6.3.3.3 Μετρήσεις Απόδοσης	108
6.3.3.4 Confusion Matrix	109
ΚΕΦΑΛΑΙΟ 7^ο: Συμπεράσματα και Περαιτέρω Έρευνα	110
7.1 Επίδραση της Ανισοροπίας των Δεδομένων	110
7.2 Αποτελεσματικότητα των Αλλαγών στην Αρχιτεκτονική.....	111
7.3 Συγκριτική Ανάλυση Δικτύων 2 Κλάσεων	111
7.4 Περαιτέρω Έρευνα.....	113
Βιβλιογραφία – Αναφορές – Διαδικτυακές Πηγές	114
Παράρτημα Α – Προ-επεξεργασία των δεδομένων.....	117
Κώδικας Δημιουργίας Φιλτραρισμένων Εικόνων	117
Εντολές για οπτικοποίηση των αποτελεσμάτων.....	117
Κώδικας πρόσθεσης εικόνων	118
Συνάρτηση shuffleImageDatstore	118
Δημιουργία και Σύγκριση Μασκών	119
Παράρτημα Β – Εκπαίδευση Νευρωνικών Δικτύων	120
Κώδικας εκπαίδευσης του FNN.....	120
Κώδικας εκπαίδευσης του CNN #1	122
Κώδικας εκπαίδευσης του CNN #2	123

Κατάλογος Πινάκων

Πίνακας 3.1 Πραγματικές Τιμές και Προβλέψεις ενός δικτύου.....	48
Πίνακας 4.1 Παρουσίαση των πειραμάτων που θα πραγματοποιηθούν ανά κεφάλαιο	55
Πίνακας 7.1 Αποτελέσματα Απόδοσης όλων των Νευρωνικών Δικτύων που εξετάστηκαν	110

Πίνακας 7.2 Αποτελέσματα Απόδοσης των Νευρωνικών Δικτύων με χρήση 2 κλάσεων (2-class classification, Normal v/s Malignant).....	112
--	-----

Κατάλογος Εικόνων

Εικόνα 1.1 Διαφορές φυσιολογικών και καρκινικών κυττάρων [Πηγή: https://www.verywellhealth.com/cancer-cells-vs-normal-cells-2248794	19
Εικόνα 1.2 Γονίδια BRCA1 και BRCA2 [Πηγή: https://www.cancer.gov/news-events/cancer-currents-blog/2017/brca-mutation-cancer-risk]	21
Εικόνα 1.3 Πιθανότητες ανάπτυξης καρκίνου του μαστού σε άτομα υπό θεραπεία ορμονικής υποκατάστασης (HRT) με (α) μόνο οιστρογόνα (ανοιχτόχρωμη γραμμή) (β) οιστρογόνα και προγεστερόνη, ως προς χρόνια χρήσης της θεραπείας (σκουρόχρωμη γραμμή) [Πηγή: https://www.auntminnie.com/clinical-news/womens-imaging/breast/article/15627161/some-types-of-hrt-more-risky-for-breast-cancer].....	22
Εικόνα 1.4 Συμπτώματα Καρκίνου του Μαστού [Πηγή: https://post.healthline.com/wp-content/uploads/2021/09/1535013-What-Are-the-Breast-Cancer-Warning-Signs-1296x1688-Body.jpg].....	23
Εικόνα 1.5 Απεικόνιση μαστού με: (α) Μαστογραφία, (β) Υπέρηχο, (γ) Μαγνητική Τομογραφία [Πηγή: https://breastcan.weebly.com/imaging.html].....	24
Εικόνα 1.6 Ιπποκράτης (460 π.Χ. – 365 π.Χ) [Πηγή: https://www.britannica.com/biography/Hippocrates]	25
Εικόνα 1.7 Παράδειγμα ενός συνελκτικού νευρωνικού δικτύου (CNN) με πολλαπλά επίπεδα [Πηγή: Breast tumor localization and segmentation using machine learning techniques: Overview of datasets, findings, and methods].....	27
Εικόνα 1.8 Παράδειγμα ενός συνελκτικού νευρωνικού δικτύου (CNN) που δέχεται εικόνες μαστογραφίας στην είσοδο [Πηγή: An Enhanced LightGBM-Based Breast Cancer Detection Technique Using Mammography Images, Abdul Rahaman Wahab Sait, and Ramprasad Nagaraj].....	27
Εικόνα 2.1 Μεγέθυνση σε εικονοστοιχεία (pixels) μιας εικόνας [Πηγή: http://barc.wi.mit.edu/graphics/scanning/resolution.php]	29
Εικόνα 2.2 Δυαδική (binary) εικόνα [Πηγή: https://www.robotix.in/tutorial/imageprocessing/imagetypes/]	30
Εικόνα 2.3 Μονοχρωματική (grayscale) εικόνα [Πηγή: https://www.robotix.in/tutorial/imageprocessing/imagetypes/]	31
Εικόνα 2.4 Βάθος (k) και αποχρώσεις του γκρι (L) [Πηγή: https://snadisplays.com/blog/bit-depth-and-why-it-matters/)].....	31
Εικόνα 2.5 Έγχρωμη (color) εικόνα [Πηγή: https://www.robotix.in/tutorial/imageprocessing/imagetypes/]	32

Εικόνα 2.6 Μεταβολή της ευκρίνειας μιας έγχρωμης εικόνας διατηρώντας σταθερό το βάθος χρώματος [Πηγή: https://www.travelplannersinternational.com/travel-agency-blog/useful-tips/the-confusion-about-image-resolution/].....	32
Εικόνα 2.7 Μεταβολή του βάθους χρώματος μιας εικόνας διατηρώντας σταθερές τις διαστάσεις της [Πηγή: https://en.wikipedia.org/wiki/Indexed_color].....	33
Εικόνα 2.8 Τεχνικές Βελτίωσης Εικόνας [Πηγή: Fundamentals of Digital Image Processing, Anil K. Jain]	33
Εικόνα 2.9 (α) Αρχική εικόνα (β) Χάρτης κατανομής συχνότητας μετά τον μετασχηματισμό Fourier [Πηγή: https://medium.com/ntust-aivc/opencv-fourier-transform-d9811aaac2d5].....	35
Εικόνα 2.10 Απεικόνιση του άπειρου περιοδικού φάσματος μετά την εφαρμογή μετασχηματισμού Fourier στην αρχική εικόνα [Πηγή: https://medium.com/ntust-aivc/opencv-fourier-transform-d9811aaac2d5#7656].....	35
Εικόνα 2.11 Εφαρμογή low pass φίλτρου στο πεδίο της συχνότητας, πολλαπλασιάζοντας με 0 (μαύρο) τις υψηλές συχνότητες και διατηρώντας τις χαμηλές στο κέντρο του συχνοτικού πεδίου που πολλαπλασιάζονται με 1 (άσπρο) [Πηγή: https://medium.com/ntust-aivc/opencv-fourier-transform-d9811aaac2d5#7656].....	36
Εικόνα 2.12 Αποτέλεσμα low pass φίλτρου στο πεδίο του χώρου [Πηγή: https://medium.com/ntust-aivc/opencv-fourier-transform-d9811aaac2d5#7656]	36
Εικόνα 2.13 Εφαρμογή high pass φίλτρου στο πεδίο της συχνότητας, πολλαπλασιάζοντας με 0 (μαύρο) τις χαμηλές συχνότητες στο κέντρο του συχνοτικού πεδίου και διατηρώντας τις υψηλές συχνότητες που πολλαπλασιάζονται με 1 (άσπρο) [Πηγή: https://medium.com/ntust-aivc/opencv-fourier-transform-d9811aaac2d5#7656].....	37
Εικόνα 2.14 Αποτέλεσμα high pass φίλτρου στο πεδίο του χώρου [Πηγή: https://medium.com/ntust-aivc/opencv-fourier-transform-d9811aaac2d5#7656]	37
Εικόνα 2.15 Αναπαράσταση ενός (α) απλού high pass φίλτρου (β) Γκαουσιανού high pass φίλτρου [Πηγή: Digital Image Processing Using MATLAB, Second Edition, Rafael C. Gonzalez, Richard E. Woods, Steven L. Eddins]	38
Εικόνα 2.16 Εφαρμογή ενός Γκαουσιανού low pass φίλτρου [Πηγή: Digital Image Processing Using MATLAB, Second Edition, Rafael C. Gonzalez, Richard E. Woods, Steven L. Eddins]	38
Εικόνα 2.17 Ενίσχυση της αντίθεσης μιας εικόνας από μαστογραφία (α) αρχική εικόνα (β) εικόνα με ενισχυμένη αντίθεση [Πηγή: Morphological image processing for quantitative shape analysis of biomedical structures: effective contrast enhancement, Yoshitaka Kimori].....	40
Εικόνα 2.18 Τμηματοποίηση εικόνας ανατομίας σώματος για τον διαχωρισμό οργάνων: συκώτι (κόκκινο), στομάχι (πράσινο), σπλήνα (κίτρινο) [Πηγή: https://www.comp.nus.edu.sg/~leowwk/model-seg/]	42
Εικόνα 3.1 Βιολογικός Νευρώνας [Πηγή: https://commons.wikimedia.org/wiki/File:Neuron_el.svg].....	43
Εικόνα 3.2 Τεχνητός Νευρώνας [Πηγή: https://eclass.upatras.gr/modules/document/file.php/CIV1756/8-Artificial%20Neural%20Networks%20%28ANN%29.pdf]	43

Εικόνα 3.3 Συναρτήσεις ενεργοποίησης νευρώνων: (α) threshold (β) γραμμική (γ) σιγμοειδής (δ) ReLU [Πηγή: https://www.naukri.com/code360/library/activation-functions]	44
Εικόνα 3.4 Ένα νευρωνικό δίκτυο με τέσσερις εισόδους και ένα κρυφό επίπεδο με τρεις κρυμμένους νευρώνες [Πηγή: https://otexts.com/fppgr/nnetar.html].....	45
Εικόνα 3.5 Σύγκριση ενός (α) ANN και (β) Deep Neural Network [Πηγή: https://pub.mdpi-res.com/energies/energies-13-04782/article_deploy/html/images/energies-13-04782-g002.png?1600070307]	46
Εικόνα 3.6 Παράδειγμα δομής ενός GAN [Πηγή: https://sthalles.github.io/intro-to-gans]	47
Εικόνα 3.7 Confusion Matrix σύμφωνα με τα δεδομένα του Πίνακα 3.1	49
Εικόνα 3.8 FNN ενός κρυφού επιπέδου νευρώνων [Πηγή: https://stackoverflow.com/questions/47499700/units-of-neural-network-layer-are-independent]	49
Εικόνα 3.9 FNN με πολλαπλά (πλήρως συνδεδεμένα) κρυφά επίπεδα νευρώνων [Πηγή: https://www.i2tutorials.com/hidden-layers-in-neural-networks/]	50
Εικόνα 3.10 Αρχιτεκτονική CNN του νευρωνικού δικτύου VGG-16 [Πηγή: https://learnopencv.com/understanding-convolutional-neural-networks-cnn/].....	51
Εικόνα 3.11 Οπτικοποίηση των σημείων ενεργοποίησης ενός καναλιού από ένα πρώιμο συνελκτικό επίπεδο ενός CNN στο MATLAB. Το συγκεκριμένο κανάλι ενεργοποιείται στα περιγράμματα [Πηγή: https://www.mathworks.com/help/deeplearning/ug/visualize-activations-of-a-convolutional-neural-network.html].....	52
Εικόνα 3.12 Οπτικοποίηση των σημείων ενεργοποίησης ενός καναλιού από ένα πρώιμο συνελκτικό επίπεδο ενός CNN στο MATLAB. Το συγκεκριμένο κανάλι ενεργοποιείται στα ερυθρά pixels. [Πηγή: https://www.mathworks.com/help/deeplearning/ug/visualize-activations-of-a-convolutional-neural-network.html].....	52
Εικόνα 3.13 Οπτικοποίηση των σημείων ενεργοποίησης ενός καναλιού από ένα πιο βαθύ επίπεδο ενός CNN στο MATLAB.. Το κανάλι ενεργοποιείται σε πιο περίπλοκα χαρακτηριστικά, όπως τα μάτια [Πηγή: https://www.mathworks.com/help/deeplearning/ug/visualize-activations-of-a-convolutional-neural-network.html].....	52
Εικόνα 3.14 Ενδεικτικό διάγραμμα της διαδικασίας της συνέλιξης [Πηγή: https://www.researchgate.net/figure/Principle-diagram-of-convolution-process_fig1_339918931].....	53
Εικόνα 3.15 Παραδείγματα διαφορετικών ειδών Padding (πάνω σειρά) και πώς αυτά φαίνονται σε θολή εικόνα (κάτω σειρά) [Πηγή: Computer Vision: Algorithms and Applications, 2nd Edition, Richard Szeliski]	54
Εικόνα 4.1 Παραδείγματα εικόνων του dataset (α) καλοήθεια, (β) κακοήθεια, (γ) φυσιολογικός ιστός..	56
Εικόνα 4.2 Οι μάσκες που αντιστοιχούν στα παραδείγματα της Εικόνας 4.1 (α) καλοήθεια, (β) κακοήθεια, (γ) φυσιολογικός ιστός.....	56
Εικόνα 4.3 Αποτελέσματα και σύγκριση εικόνων και масκών με εφαρμογή κατωφλίου και CLAHE.....	58
Εικόνα 4.4 Αποτελέσματα και σύγκριση εικόνων και масκών με εφαρμογή κατωφλίου Otsu και CLAHE	59

Εικόνα 4.5 Απαιτούμενα πρόσθετα εργαλεία στο MATLAB για την υλοποίηση των πειραμάτων	60
Εικόνα 4.6 Περιεχόμενα του normalDatastore στο MATLAB	61
Εικόνα 4.7 Περιεχόμενα του υποσυνόλου training για την κλάση ‘normal’ στο MATLAB	62
Εικόνα 4.8 Περιεχόμενα του trainNormal.Files στο MATLAB	62
Εικόνα 4.9 Περιεχόμενα του trainData στο MATLAB	63
Εικόνα 4.10 Ετικέτες του trainData (α) πριν το ανακάτεμα και (β) μετά το ανακάτεμα	64
Εικόνα 4.11 (α) Οπτικοποίηση των συχνοτήτων μετά τον μετασχηματισμό Fourier (β) Γκαουσιανό χαμηλοπερατό φίλτρο (γ) Φιλτραρισμένη εικόνα στο πεδίο της συχνότητας	66
Εικόνα 4.12 Η εικόνα benign (1).png (α) στην αρχική μορφή (β) κατόπιν εφαρμογής του γκαουσιανού χαμηλοπερατού φίλτρου	67
Εικόνα 4.13 Μεγέθυνση στο σημείο ενδιαφέροντος της εικόνας benign (1).png (α) στην αρχική μορφή (β) κατόπιν εφαρμογής του γκαουσιανού χαμηλοπερατού φίλτρου	68
Εικόνα 4.14 Οι τιμές γκρι της αρχικής εικόνας benign (1).png περίπου στο σημείο ενδιαφέροντος	68
Εικόνα 4.15 Οι τιμές γκρι της φιλτραρισμένης εικόνας benign (1).png περίπου στο σημείο ενδιαφέροντος	69
Εικόνα 4.16 Η εικόνα benign (1).png (α) στην αρχική μορφή (β) κατόπιν εφαρμογής του γκαουσιανού χαμηλοπερατού φίλτρου (γ) κατόπιν πρόσθεσης των (α) και (β)	70
Εικόνα 4.17 Οι τιμές γκρι της εικόνας benign (1).png που έχει προκύψει από την άθροιση της αρχικής και της φιλτραρισμένης, περίπου στο σημείο ενδιαφέροντος	70
Εικόνα 5.1 Στοιχεία γραφημάτων εκπαίδευσης	77
Εικόνα 5.2 Διαγράμματα Ακρίβειας και Απώλειας της εκπαίδευσης και επικύρωσης του δικτύου FNN με τις αρχικές εικόνες	78
Εικόνα 5.3 Απόδοση δικτύου FNN κατόπιν εκπαίδευσης με τις αρχικές εικόνες	79
Εικόνα 5.4 Confusion Matrix του FNN κατόπιν εκπαίδευσης με τις αρχικές εικόνες	80
Εικόνα 5.5 Διαγράμματα Ακρίβειας και Απώλειας της εκπαίδευσης και επικύρωσης του δικτύου FNN με τις φιλτραρισμένες εικόνες	81
Εικόνα 5.6 Απόδοση δικτύου FNN κατόπιν εκπαίδευσης με τις φιλτραρισμένες εικόνες	82
Εικόνα 5.7 Confusion Matrix του FNN κατόπιν εκπαίδευσης με τις φιλτραρισμένες εικόνες	83
Εικόνα 5.8 Διαγράμματα Ακρίβειας και Απώλειας της εκπαίδευσης και επικύρωσης του δικτύου FNN με τις αθροισμένες εικόνες	84
Εικόνα 5.9 Απόδοση δικτύου FNN κατόπιν εκπαίδευσης με τις αθροισμένες εικόνες	85
Εικόνα 5.10 Confusion Matrix του FNN κατόπιν εκπαίδευσης με τις αθροισμένες εικόνες	86
Εικόνα 6.1 Διαγράμματα Ακρίβειας και Απώλειας της εκπαίδευσης και επικύρωσης του δικτύου CNN #1 με τις αρχικές εικόνες	89
Εικόνα 6.2 Απόδοση δικτύου CNN #1 κατόπιν εκπαίδευσης με τις αρχικές εικόνες	90

Εικόνα 6.3 Confusion Matrix του CNN #1 κατόπιν εκπαίδευσης με τις αρχικές εικόνες.....	91
Εικόνα 6.4 Διαγράμματα Ακρίβειας και Απώλειας της εκπαίδευσης και επικύρωσης του δικτύου CNN #1 με τις φιλτραρισμένες εικόνες	92
Εικόνα 6.5 Απόδοση δικτύου CNN #1 κατόπιν εκπαίδευσης με τις φιλτραρισμένες εικόνες	94
Εικόνα 6.6 Confusion Matrix του CNN #1 κατόπιν εκπαίδευσης με τις φιλτραρισμένες εικόνες.....	95
Εικόνα 6.7 Διαγράμματα Ακρίβειας και Απώλειας της εκπαίδευσης και επικύρωσης του δικτύου CNN #1 με τις αθροισμένες εικόνες	96
Εικόνα 6.8 Απόδοση δικτύου CNN #1 κατόπιν εκπαίδευσης με τις αθροισμένες εικόνες	97
Εικόνα 6.9 Confusion Matrix του CNN #1 κατόπιν εκπαίδευσης με τις αθροισμένες εικόνες.....	98
Εικόνα 6.10 Διαγράμματα Ακρίβειας και Απώλειας της εκπαίδευσης και επικύρωσης του δικτύου CNN #2 με τις αρχικές εικόνες	101
Εικόνα 6.11 Απόδοση δικτύου CNN #2 κατόπιν εκπαίδευσης με τις αρχικές εικόνες	102
Εικόνα 6.12 Confusion Matrix του CNN #2 κατόπιν εκπαίδευσης με τις αρχικές εικόνες.....	103
Εικόνα 6.13 Διαγράμματα Ακρίβειας και Απώλειας της εκπαίδευσης και επικύρωσης του δικτύου CNN #2 με τις φιλτραρισμένες εικόνες	104
Εικόνα 6.14 Απόδοση δικτύου CNN #2 κατόπιν εκπαίδευσης με τις φιλτραρισμένες εικόνες	105
Εικόνα 6.15 Confusion Matrix του CNN #2 κατόπιν εκπαίδευσης με τις φιλτραρισμένες εικόνες.....	106
Εικόνα 6.16 Διαγράμματα Ακρίβειας και Απώλειας της εκπαίδευσης και επικύρωσης του δικτύου CNN #2 με τις αθροισμένες εικόνες	107
Εικόνα 6.17 Απόδοση δικτύου CNN #2 κατόπιν εκπαίδευσης με τις αθροισμένες εικόνες	108
Εικόνα 6.18 Confusion Matrix του CNN #2 κατόπιν εκπαίδευσης με τις αθροισμένες εικόνες.....	109

Αλφαβητικό Ευρετήριο

ML: Machine Learning
ANN: Artificial Neural Network
DNN: Deep Neural Network
FNN: Feedforward Neural Network
CNN: Convolutional Neural Network
RNN: Recurrent Neural Network
GAN: Generative Adversarial Network
FFT2: Fast Fourier Transform 2D
HRT: Hormonal Replacement Therapy
MRI: Magnetic Resonance Imaging
MRT: Magnetic Resonance Tomography
PET: Positron Emission Tomography
CAD: Computer-Aided Diagnostics

ΕΙΣΑΓΩΓΗ

Η παρούσα διπλωματική εργασία επικεντρώνεται στην επεξεργασία και ανάλυση βιοϊατρικών εικόνων, με ιδιαίτερη εστίαση στα υπερηχογραφήματα μαστού, και στη χρήση αυτών για την εκπαίδευση νευρωνικών δικτύων βαθιάς μάθησης. Το πεδίο αυτό παρουσιάζει ιδιαίτερο ενδιαφέρον και σημασία λόγω της ανάγκης για ακριβή και γρήγορη διάγνωση του καρκίνου του μαστού, μιας ασθένειας που πλήττει εκατομμύρια γυναίκες παγκοσμίως.

Αντικείμενο της διπλωματικής εργασίας

Η εργασία αποσκοπεί στην ανάπτυξη και αξιολόγηση τεχνικών επεξεργασίας υπερηχογραφημάτων μαστού, με στόχο τη βελτίωση της ακρίβειας στην αναγνώριση και κατηγοριοποίηση των εικόνων μέσω νευρωνικών δικτύων. Η έρευνα αυτή είναι σημαντική, καθώς μπορεί να συμβάλει στη βελτίωση των διαγνωστικών μεθόδων και στην έγκαιρη ανίχνευση του καρκίνου του μαστού.

Σκοπός και στόχοι

Ο σκοπός της εργασίας είναι να αξιολογήσει την απόδοση διαφόρων τύπων νευρωνικών δικτύων, όπως τα νευρωνικά δίκτυα πρόσθιας τροφοδότησης (FNN) και τα συνελκτικά νευρωνικά δίκτυα (CNN), στην ανάλυση υπερηχογραφημάτων μαστού. Στόχοι της εργασίας είναι:

- Η ανάπτυξη και εφαρμογή τεχνικών επεξεργασίας εικόνων για την εξαγωγή κρίσιμων χαρακτηριστικών.
- Η εκπαίδευση και αξιολόγηση νευρωνικών δικτύων για την αναγνώριση και κατηγοριοποίηση των υπερηχογραφημάτων (μαστογραφιών).
- Η διερεύνηση της επίδρασης διαφόρων παραμέτρων στην απόδοση των νευρωνικών δικτύων ως κατηγοριοποιητών εικόνων μαστογραφιών.

Μεθοδολογία

Η μεθοδολογία που ακολουθήθηκε περιλαμβάνει τα εξής βήματα:

- Συλλογή και προ-επεξεργασία των υπερηχογραφημάτων μαστού.
- Εφαρμογή τεχνικών επεξεργασίας εικόνας για την ανάδειξη κρίσιμων χαρακτηριστικών.
- Εκπαίδευση και αξιολόγηση διαφόρων τύπων νευρωνικών δικτύων χρησιμοποιώντας το λογισμικό MATLAB®.
- Ανάλυση των αποτελεσμάτων και αξιολόγηση της απόδοσης των δικτύων.

Καινοτομία

Η καινοτομία της παρούσας διπλωματικής εργασίας έγκειται στην εφαρμογή σύγχρονων τεχνικών βαθιάς μάθησης για την ανάλυση βιοϊατρικών εικόνων και στην ανάπτυξη μιας ολοκληρωμένης μεθοδολογίας για την αναγνώριση και κατηγοριοποίηση υπερηχογραφημάτων μαστού. Η εργασία επιδιώκει να συμβάλει στη βελτίωση των διαγνωστικών διαδικασιών, προσφέροντας πιο αξιόπιστες και ακριβείς λύσεις στην ιατρική διάγνωση του καρκίνου.

Δομή

Η παρούσα διπλωματική εργασία είναι δομημένη σε επτά κύρια κεφάλαια, καθένα από τα οποία συμβάλλει στην ολοκληρωμένη παρουσίαση και ανάλυση του θέματος. Αναλυτικά, η δομή της εργασίας είναι η εξής:

ΚΕΦΑΛΑΙΟ 1^ο: Παρέχεται μια εισαγωγή στις βασικές έννοιες και το ιστορικό πλαίσιο του καρκίνου και ειδικότερα του καρκίνου του μαστού. Αναλύονται οι αιτίες και τα συμπτώματα του καρκίνου του μαστού, καθώς και οι μέθοδοι διάγνωσης και απεικόνισης. Επιπλέον, γίνεται ιστορική αναδρομή από την αρχαία Ελλάδα και Αίγυπτο έως τις σύγχρονες θεραπευτικές προσεγγίσεις.

ΚΕΦΑΛΑΙΟ 2^ο: Επικεντρώνεται στην επεξεργασία ψηφιακών βιοϊατρικών εικόνων. Περιγράφονται οι διαστάσεις των εικόνων, από τις δυαδικές και μονοχρωματικές έως τις έγχρωμες εικόνες. Εξετάζονται οι τεχνικές βελτίωσης εικόνας στο πεδίο της συχνότητας, συμπεριλαμβανομένου του μετασχηματισμού Fourier και των φίλτρων συχνότητων, καθώς και οι μέθοδοι επεξεργασίας εικόνων που χρησιμοποιούνται στη βιοϊατρική.

ΚΕΦΑΛΑΙΟ 3^ο: Αναλύεται η δομή και η λειτουργία των τεχνητών νευρωνικών δικτύων, με έμφαση στα νευρωνικά δίκτυα βαθιάς μάθησης (DNNs). Παρουσιάζονται οι κατηγορίες των DNNs, η αξιολόγηση της απόδοσής τους μέσω confusion matrix, και εξετάζονται αναλυτικά τα συνελκτικά νευρωνικά δίκτυα (CNNs) με ιδιαίτερη ανάλυση στις παραμέτρους και τη λειτουργία τους.

ΚΕΦΑΛΑΙΟ 4^ο: Περιγράφεται η πειραματική μεθοδολογία και το σύνολο δεδομένων που χρησιμοποιήθηκαν στην έρευνα. Περιλαμβάνονται η δημιουργία και σύγκριση μασκών, η προετοιμασία και προ-επεξεργασία των δεδομένων, καθώς και οι παράμετροι εκπαίδευσης των νευρωνικών δικτύων.

ΚΕΦΑΛΑΙΟ 5^ο: Επικεντρώνεται στην ανάλυση των δεδομένων με τη χρήση Feedforward Neural Networks (FNNs). Παρουσιάζονται η αρχιτεκτονική και οι ρυθμίσεις εκπαίδευσης του FNN, καθώς και τα αποτελέσματα των δοκιμών με διαφορετικές παραμέτρους και μεθόδους επεξεργασίας εικόνων.

ΚΕΦΑΛΑΙΟ 6^ο: Εξετάζεται η ανάλυση των δεδομένων με τη χρήση Convolutional Neural Networks (CNNs). Αναλύονται δύο διαφορετικές αρχιτεκτονικές CNN, οι οποίες αξιολογούνται με διαφορετικές μεθόδους επεξεργασίας εικόνων και παρουσιάζονται τα αποτελέσματα της απόδοσής τους.

ΚΕΦΑΛΑΙΟ 7^ο: Συνοψίζονται τα ευρήματα της έρευνας και αξιολογείται η αποτελεσματικότητα των αλλαγών στην αρχιτεκτονική των νευρωνικών δικτύων. Τέλος, προτείνονται κατευθύνσεις για περαιτέρω έρευνα.

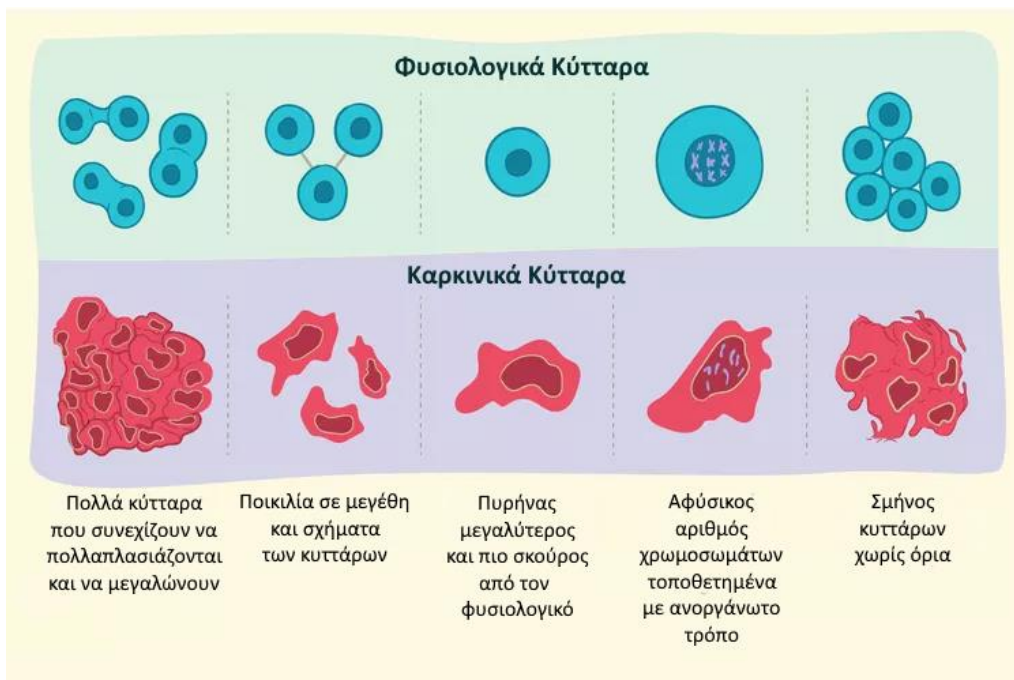
ΚΕΦΑΛΑΙΟ 1^ο: Επιστημονική και Ιστορική Ανάλυση του Καρκίνου του Μαστού

Τα τελευταία χρόνια, οι θάνατοι από καρκίνο αυξάνονται ολοένα και περισσότερο σε όλον τον κόσμο. Συγκριτικά με άλλους τύπους καρκίνου ο καρκίνος του μαστού θεωρείται η δεύτερη πιο κοινή αιτία θανάτου, ιδιαίτερα στις γυναίκες. Σύμφωνα με τον Παγκόσμιο Οργανισμό Υγείας, πάνω από 2 εκατομμύρια νέα περιστατικά καρκίνου του μαστού εντοπίστηκαν το 2020, καθιστώντας τον τη νόσο με τη μεγαλύτερη διάγνωση. Κάθε χρόνο, 626.700 γυναίκες πεθαίνουν από ασθένειες που σχετίζονται με τον καρκίνο. Ο καρκίνος του μαστού μπορεί να είναι θανατηφόρος εάν δεν εντοπιστεί και αντιμετωπιστεί εγκαίρως.

Η διάγνωση σε πρώιμο στάδιο σε συνδυασμό με μια αποτελεσματική θεραπεία, αυξάνει σε μεγάλο βαθμό την πιθανότητα επιβίωσης. Ο ασθενής έχει 85% πιθανότητα να βιώσει πλήρη ύφεση εάν ο καρκίνος ανακαλυφθεί πριν μεγαλώσει σε μέγεθος 10 mm. Προκειμένου να αναγνωριστεί άμεσα ο καρκίνος του μαστού, συνιστώνται μέθοδοι απεικόνισης όπως η μαστογραφία, η μαγνητική τομογραφία και ο υπέρηχος.

1.1 Βασικές Έννοιες και Χαρακτηριστικά

Ο καρκίνος είναι το αποτέλεσμα της ανεξέλεγκτης διαίρεσης των κυττάρων με γρήγορους ρυθμούς που οδηγεί σε σχηματισμό όγκου ή μάζας. Τα κύτταρα αυτά μπορεί να εξαπλωθούν μέσω του λεμφικού συστήματος ή της κυκλοφορίας του αίματος, επιδρώντας έτσι σε διάφορα σημεία του σώματος. Σε αυτήν την περίπτωση, ο όγκος θεωρείται κακοήθης (καρκινικός).



Εικόνα 1.1 Διαφορές φυσιολογικών και καρκινικών κυττάρων

Οι πυρήνες των κακοηθών ιστών είναι συνήθως σημαντικά μεγαλύτεροι από εκείνους των φυσιολογικών ιστών, κάτι το οποίο μπορεί να επιφέρει θανατηφόρες επιπτώσεις σε μεταγενέστερα στάδια της νόσου.

Οι καλοήθειες περιπτώσεις από την άλλη πλευρά θεωρούνται μη καρκινικές και μη απειλητικές για τη ζωή του ατόμου. Ωστόσο, σε ορισμένες περιπτώσεις υπάρχει η πιθανότητα να εξελιχθούν σε καρκίνο.

1.1.1 Ο καρκίνος του μαστού

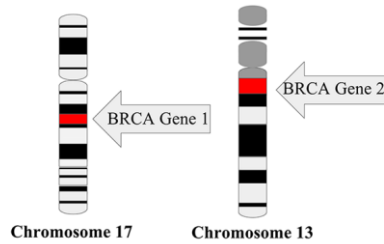
Ο καρκίνος του μαστού αποτελεί την πιο συχνή μορφή καρκίνου και την δεύτερη πιο κοινή αιτία θανάτου στις γυναίκες παγκοσμίως. Η έγκαιρη διάγνωση σε πρώιμο στάδιο σε συνδυασμό με μια αποτελεσματική θεραπεία μπορεί να αυξήσει σε μεγάλο βαθμό την πιθανότητα επιβίωσης. Συνήθως το 70 – 80% των περιπτώσεων είναι ιάσιμο για ασθενείς που διαγιγνώσκονται σε μη μεταστατικό στάδιο. Οι τεχνικές απεικόνισης μαστού, μπορούν να εντοπίσουν την θέση των καρκινικών κυττάρων με μεγάλη ακρίβεια, όπως επίσης και να παρέχουν διάφορα είδη σημαντικών πληροφοριών συμπεριλαμβανομένων της δομής, των λειτουργιών και της μορφολογίας.

1.1.2 Αιτίες του καρκίνου του μαστού

Αν και οι ακριβείς αιτίες του καρκίνου του μαστού δεν είναι απόλυτες, υπάρχουν κάποιοι ενδεικτικοί παράγοντες που αυξάνουν τις πιθανότητες εμφάνισης καρκίνου. Μερικοί από αυτούς μπορούν να αλλάξουν, όπως για παράδειγμα παράγοντες που αφορούν τον τρόπο ζωής, ενώ άλλοι όχι. Οι παράγοντες που ενδεχομένως επηρεάζουν την εμφάνιση του καρκίνου χωρίζονται σε ελέγξιμους και μη ελέγξιμους.

Μη ελέγξιμοι παράγοντες:

- **Ηλικία:** Ο κίνδυνος καρκίνου του μαστού είναι ανάλογος της ηλικίας. Το μεγαλύτερο ποσοστό των περιπτώσεων, συγκεκριμένα ένα 80%, αφορά γυναίκες άνω των 50 ετών και σε γυναίκες που είχαν εμμηνόπαυση. Γι' αυτό είναι απαραίτητο οι γυναίκες μεταξύ 50 και 70 ετών να ελέγχονται τακτικά. Είναι επίσης πιθανό και οι γυναίκες άνω των 70 ετών να εμφανίσουν καρκίνο του μαστού.
- **Γένος:** Ο καρκίνος του μαστού μπορεί σπάνια να εμφανιστεί ακόμα και στους άνδρες. Οι γυναίκες ωστόσο έχουν 100 φορές περισσότερες πιθανότητες να εμφανίσουν καρκίνο του μαστού σε σχέση με τους άνδρες.
- **Πυκνότητα στήθους:** Οι γυναίκες με πυκνό στήθος έχουν υψηλότερο κίνδυνο ανάπτυξης καρκίνου σε σύγκριση με γυναίκες με λιγότερο πυκνό στήθος. Αυτό συμβαίνει διότι υπάρχουν περισσότερα κύτταρα που μπορούν να εξελιχθούν σε καρκινικά. Η υψηλή πυκνότητα του μαστού καθιστά επίσης πιο δύσκολη την διάγνωση όγκων χρησιμοποιώντας μαστογραφία.
- **Οικογενειακό ιστορικό:** Παρόλο που οι περισσότεροι από τους καρκίνους του μαστού δεν έχουν σχέση με την κληρονομικότητα, υπάρχουν γονίδια τα οποία καθορίζουν την πιθανότητα εμφάνισης καρκίνου του μαστού. Τα γονίδια είναι τα BRCA1 και BRCA2 μπορούν να αυξήσουν τον κίνδυνο ανάπτυξης τόσο του καρκίνου του μαστού όσο και του καρκίνου των ωοθηκών. Είναι επίσης πιθανό αυτά τα γονίδια να περάσουν από έναν γονέα στο παιδί. Έτσι, συμπεραίνουμε ότι οι γυναίκες που έχουν στενή συγγένεια με πρόσωπο που είχε εμφανίσει καρκίνο του μαστού ή καρκίνο των ωοθηκών έχουν υψηλότερο κίνδυνο.



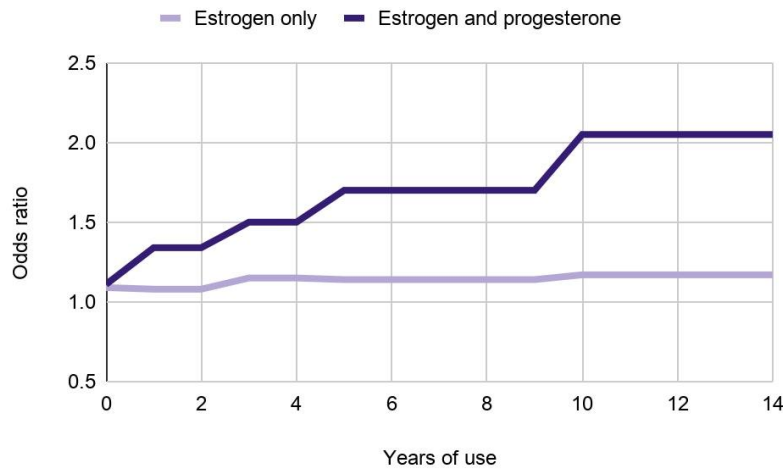
Εικόνα 1.2 Γονίδια BRCA1 και BRCA2

- **Προηγούμενο ιστορικό με καρκίνο του μαστού ή καλοήθεις όγκους:** Γυναίκες που εμφάνισαν καρκίνο στον ένα μαστό είναι πιθανό να εμφανίσουν αργότερα καρκίνο είτε στον άλλο μαστό είτε στον ίδιο. Επιπλέον, αν και ένας καλοήθης όγκος δεν υποδηλώνει καρκίνο, είναι πιθανό μια καλοήθεια να προηγείται ενός καρκίνου.

Ελέγξιμοι παράγοντες:

- **Παχυσαρκία:** Τα παχύσαρκα ή υπέρβαρα άτομα έχουν περισσότερες πιθανότητες να εμφανίσουν καρκίνο του μαστού. Υπάρχει μία θεωρία ότι οι γυναίκες αυτές μπορεί να έχουν υψηλότερα επίπεδα οιστρογόνων στο αίμα. Επιπλέον, μετά την εμμηνόπαυση, οι γυναίκες που είναι υπέρβαρες ή παχύσαρκες παράγουν περισσότερα οιστρογόνα.
- **Έκθεση σε ακτινοβολία:** Μια μικρή αύξηση στον κίνδυνο εμφάνισης του καρκίνου του μαστού μπορεί να προκύψει από απεικονιστικές εξετάσεις όπως η αξονική τομογραφία και οι ακτινογραφίες. Ο κίνδυνος καρκίνου του μαστού είναι υψηλότερος σε γυναίκες που έχουν υποβληθεί στο παρελθόν σε ακτινοθεραπεία για καρκίνο.
- **Αλκοόλ:** Ο καρκίνος του μαστού είναι πιο συχνός σε γυναίκες που καταναλώνουν μεγάλες ποσότητες αλκοόλ. Για κάθε 200 γυναίκες που καταναλώνουν συστηματικά δύο αλκοολούχα ποτά κάθε μέρα, τρεις περισσότερες γυναίκες από τον ίδιο αριθμό γυναικών που δεν πίνουν καθόλου διαγιγνώσκονται με καρκίνο του μαστού.
- **Θεραπεία ορμονικής υποκατάστασης (HRT – Hormonal Replacement Therapy):** Γυναίκες που βρίσκονται υπό θεραπεία ορμονικής υποκατάστασης πρέπει να παίρνουν χάπια τα οποία περιέχουν προγεστερόνη, οιστρογόνα ή και τα δύο. Ο κίνδυνος καρκίνου του μαστού είναι κάπως αυξημένος σε όσους βρίσκονται σε θεραπεία ορμονικής υποκατάστασης. Η HRT μόνο με οιστρογόνα και η συνδυαστική HRT μπορούν να αυξήσουν τον κίνδυνο καρκίνου του μαστού, ωστόσο ο κίνδυνος της συνδυαστικής HRT είναι οριακά υψηλότερος.

Odds of developing breast cancer based on recent HRT use



Εικόνα 1.3 Πιθανότητες ανάπτυξης καρκίνου του μαστού σε άτομα υπό θεραπεία ορμονικής υποκατάστασης (HRT) με (α) μόνο οιστρογόνα (ανοιχτόχρωμη γραμμή) (β) οιστρογόνα και προγεστερόνη, ως προς χρόνια χρήση της θεραπείας (σκουρόχρωμη γραμμή)

Για κάθε 1.000 γυναίκες που λαμβάνουν συνδυαστική θεραπεία ορμονικής υποκατάστασης για μια δεκαετία, υπάρχουν περίπου 19 περισσότερες περιπτώσεις καρκίνου του μαστού. Μόλις διακοπεί η HRT, ο κίνδυνος μειώνεται.

- **Έκθεση σε διαιθυλοστυλβεστερόλη (DES):** Οι γυναίκες λάμβαναν αυτό το φάρμακο μεταξύ της δεκαετίας του 1940 και της δεκαετίας του 1960 κυρίως. Μετά την ηλικία των 40, υπήρχε υψηλότερος κίνδυνος καρκίνου του μαστού για τις γυναίκες που χρησιμοποιούσαν διαιθυλοστυλβεστερόλη (DES) για να αποφύγουν τις αποβολές.

Τα εμφυτεύματα στήθους (εγκεκριμένα από ρυθμιστικούς φορείς), η χρήση συρμάτινων στηθόδεσμων, η χρήση αντιδρωτικών και η έκθεση σε φυτοφάρμακα *δεν* αυξάνουν τον κίνδυνο καρκίνου του μαστού.

1.1.3 Συμπτώματα του καρκίνου του μαστού

Ένα πολύ σημαντικό βήμα για την πρόληψη αλλά και για την αντιμετώπιση του καρκίνου του μαστού είναι να μπορεί κανείς να αναγνωρίζει τις πιθανές ενδείξεις. Αν και τα συμπτώματα μπορεί να ποικίλουν, η γνώση των πιο κοινών συμπτωμάτων είναι απαραίτητη για την έγκαιρη ανίχνευση του προβλήματος όπως επίσης και την έγκαιρη αντιμετώπιση. Πιο αναλυτικά μερικά από τα συμπτώματα είναι:

- **Όγκοι μέσα στο στήθος:** Μπορεί να βρίσκονται είτε πάνω στο δέρμα του στήθους είτε στις δομές κάτω από το δέρμα, όπως το θωρακικό τοίχωμα. Ένα τέτοιο ογκίδιο είναι αρκετά συχνά ενδεικτικό κακοήθους όγκου.
- **Όγκοι κοντά στις μασχάλες, την κλείδα, κλπ.:** Αυτό είναι ένδειξη διογκωμένων λεμφαδένων. Αν και υπάρχουν πολλοί λόγοι για τους οποίους μπορεί ένας λεμφαδένας να έχει διογκωθεί, συνιστάται άμεσος έλεγχος για τυχόν ανωμαλίες ή ασθένειες του μαστού.
- **Ιστός που είναι πυκνότερος από το δέρμα του μαστού:** Μπορεί να υποδεικνύει παθολογικές αλλαγές στον ιστό του μαστού. Αυτό μπορεί να απαιτεί περαιτέρω ιατρική αξιολόγηση για την εξακρίβωση των αιτιών αυτής της αυξημένης πυκνότητας.

- **Αλλαγές στο σχήμα ή την εμφάνιση των θηλών:** Η εμφάνιση της θηλής συχνά αλλοιώνεται και ισοπεδώνεται από έναν όγκο που βρίσκεται κάτω από αυτήν. Για παράδειγμα, η θέση της θηλής μπορεί να αλλάξει καθώς σηκώνονται και τα δύο χέρια.
- **Εξανθήματα ή πληγές γύρω από τη θηλή:** Είναι ανησυχητικά συμπτώματα που απαιτούν ιατρική προσοχή, καθώς μπορεί να υποδεικνύουν προβλήματα στον ιστό του μαστού ή δερματικές παθήσεις.
- **Εκκρίσεις από τη μία ή και τις δύο θηλές:** Η έκκριση μπορεί να είναι διαυγής, κίτρινη, πράσινη ή ακόμα να περιέχει και αίμα.
- **Πόνος στο στήθος ή στις μασχάλες που δεν σχετίζεται με την έμμηνο ρύση:** Αν και οι όγκοι συνήθως δεν πονάνε, μπορεί να προκαλέσουν μια αίσθηση τσιμπήματος.
- **Πόνος στα οστά, εύκολα κατάγματα, δερματικά έλκη, απώλεια βάρους, κόπωση, σπασμοί ή επιληπτικές κρίσεις κλπ.:** Αυτά είναι συνήθως συμπτώματα σε προχωρημένα στάδια του καρκίνου, γι' αυτό χρειάζεται άμεση ιατρική προσοχή καθώς μπορεί να υποδεικνύουν μεταστάσεις και σοβαρές επιπλοκές.



Εικόνα 1.4 Συμπτώματα Καρκίνου του Μαστού

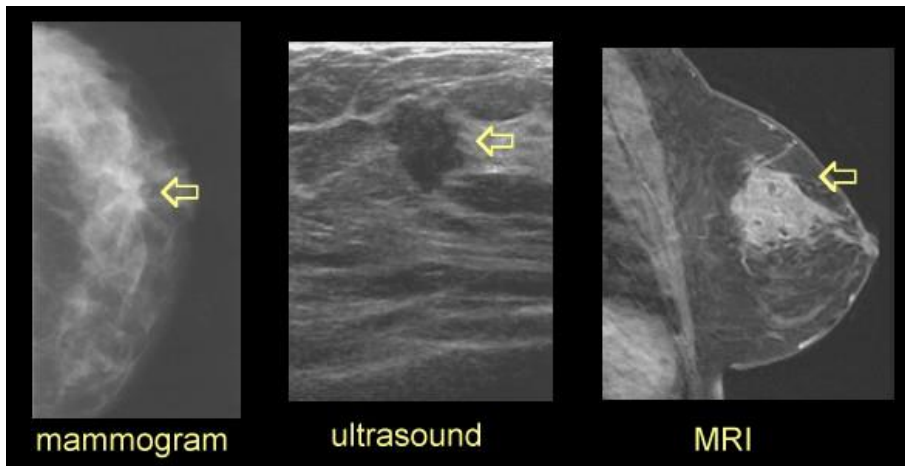
Σε κάθε περίπτωση, ένας γιατρός μπορεί να αξιολογήσει τα συμπτώματα, να ελέγξει τον μαστό ή τους μαστούς που έχουν προσβληθεί και, εάν απαιτείται περαιτέρω έρευνα, να την συστήσει. Προκειμένου να αποκλειστεί η μόλυνση ή άλλοι πιθανοί λόγοι, μπορεί να συσταθεί από τον γιατρό μαστογραφία, υπερηχογράφημα, άλλες απεικονιστικές εξετάσεις ή εξετάσεις αίματος.

1.2 Μέθοδοι Διάγνωσης και Απεικόνισης

Για την ανίχνευση του καρκίνου του μαστού όπως επίσης και την ακριβή θέση των ύποπτων όγκων, η χρήση τεχνικών απεικόνισης είναι απαραίτητη. Η απεικόνιση, εκτός από τις σημαντικές πληροφορίες που μπορεί να παρέχει, είναι ένα πολύ βασικό μέρος της διαδικασίας της διάγνωσης και αποτελεί το κύριο μέρος των κλινικών πρωτοκόλλων. Επιπλέον, με μια ποικιλία εξετάσεων και ελέγχων είναι δυνατόν να ανιχνευθεί και η εξάπλωση του καρκίνου.

Οι πιο συνηθισμένες εξετάσεις περιλαμβάνουν την αξονική τομογραφία, μαγνητική τομογραφία, τομογραφία εκπομπής ποζιτρονίων (PET scan – Positron Emission Tomography), σάρωση οστών και υπερηχογράφημα.

Συγκεκριμένα για την απεικόνιση του μαστού, χρησιμοποιείται η **μαστογραφία**, το **υπερηχογράφημα** και η **μαγνητική τομογραφία** μαστού (MRT – Magnetic Resonance Tomography ή MRI – Magnetic Resonance Imaging).



Εικόνα 1.5 Απεικόνιση μαστού με: (α) Μαστογραφία, (β) Υπέρηχο, (γ) Μαγνητική Τομογραφία

Η ιατρική απεικόνιση λοιπόν, αποτελεί καθιερωμένο κομμάτι της αντιμετώπισης του καρκίνου του μαστού και χρησιμοποιείται σε όλες τις φάσεις του ελέγχου του καρκίνου, από την αναγνώριση και τον εντοπισμό των όγκων μέχρι και την παρακολούθηση της θεραπείας και της μετεγχειρητικής πορείας. Ωστόσο, υπάρχουν κάποια προβλήματα που προκύπτουν κατά την διάγνωση με τη χρήση αυτών των μεθόδων απεικόνισης.

Μερικά παραδείγματα μπορεί να είναι η λανθασμένη ή η ανακριβής ερμηνεία των αποτελεσμάτων λόγω κούρασης του ειδικού, η ανακριβής απεικόνιση στην μαστογραφία με αποτέλεσμα να προκύψουν ψευδώς θετικά ή αρνητικά αποτελέσματα ή η χαμηλή ευαισθησία λόγω παρεμφερών πυκνοτήτων ιστού, δηλαδή οι μαστογραφίες να μην ανιχνεύουν καρκίνους σε περιοχές όπου ο ιστός του μαστού είναι πυκνός και έτσι να μην μπορεί να γίνει σωστά η διάκριση των καρκινικών περιοχών από τον φυσιολογικό ιστό, οδηγώντας σε πιθανά ψευδώς αρνητικά αποτελέσματα, όταν στην πραγματικότητα υπάρχει καρκίνος.

Επομένως, η ερμηνεία των εικόνων αποτελεί μια χρονοβόρα και απαιτητική εργασία που εξαρτάται κυρίως από τον χειριστή.

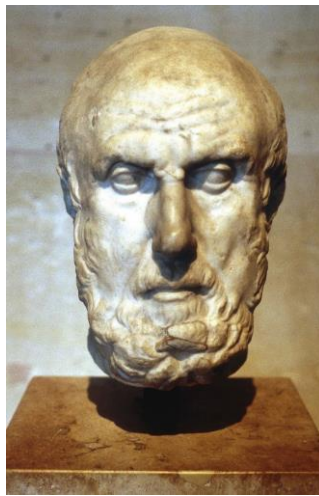
1.3 Ιστορική Αναδρομή

Ήδη από την αρχαιότητα, ο καρκίνος του μαστού είναι ένα γνωστό ιατρικό πρόβλημα και αναφέρεται σε σχεδόν κάθε περίοδο της καταγεγραμμένης ιστορίας. Σε αντίθεση με άλλες μορφές (εσωτερικού) καρκίνου, οι όγκοι στον μαστό εκδηλώνονταν ως ορατοί όγκοι σε μεταγενέστερα στάδια.

1.3.1 Αρχαία Ελλάδα και Αίγυπτος

Οι αρχαίοι Αιγύπτιοι ήταν οι πρώτοι που κατέγραψαν τη νόσο πριν από περισσότερα από 3.500 χρόνια, παρέχοντας σχετικά ακριβείς περιγραφές στους παπύρους *Edwin Smith* και *George Ebers*. Μια από τις περιγραφές αναφέρεται σε «ανίατους όγκους στο στήθος που δεν έχουν θεραπεία».

Καρκινικοί όγκοι έχουν εντοπιστεί και σε αιγυπτιακές μούμιες, όπως, για παράδειγμα, εκτεταμένος καρκίνος των ωοθηκών, με κοιλιακές μεταστάσεις, που εντοπίστηκε σε μια μούμια γυναίκας.



Εικόνα 1.6 Ιπποκράτης (460 π.Χ. – 365 π.Χ.)

Το 460 π.Χ. – 365 π.Χ., ο *Ιπποκράτης*, ο πατέρας της Ιατρικής, έδωσε την ονομασία «καρκίνος», λόγω του σχήματος των όγκων που τείνουν να μοιάζουν με την εικόνα του κάβουρα (καρκίνου).

Έπειτα, το 131 μ.Χ., ο *Γαληνός* διέκρινε τον καρκίνο από την σκληρότητα, την ακινησία και τους πόνους που προκαλούσε και έτσι παρομοίασε και αυτός τον καρκίνο με τον κάβουρα, λόγω της ομοιότητας των αποφύσεων του όγκου του μαστού με τα πόδια ενός καβουριού.

1.3.2 Θεωρίες και Θεραπευτικές Προσεγγίσεις (17^{ος} – 20^{ος} αιώνας)

Άλλες θεωρίες κυρίως από τον 17^ο και 18^ο αιώνα περιλάμβαναν τον *Giovanni Morgagni* που κατηγορούσε το πηγμένο γάλα, τον *Johanes de Gorter* που κατηγορούσε τις πυώδεις φλεγμονές στο στήθος, τον *Claude-Nicolas Le Cat* από τη Ρουέν που κατηγορούσε τις καταθλιπτικές ψυχικές διαταραχές, ο *Lorenz Heister* που κατηγορούσε την άτεκνη ζωή, και άλλους που κατηγορούσαν την καθιστική ζωή.

Η ριζική μαστεκτομή ήταν η βάση της θεραπείας για τις πρώτες τέσσερις δεκαετίες του 20^{ου} αιώνα. Αν και ήταν μια μέθοδος που βοήθησε τις γυναίκες να επιβιώσουν περισσότερο, ειδικά αν εκτελούνταν νωρίς, πολλές γυναίκες δεν την επέλεξαν αφού τις άφηνε παραμορφωμένες.

Το 1895, ο Σκωτσέζος χειρουργός *George Beatson* ανακάλυψε ότι η αφαίρεση των ωοθηκών από μια ασθενή του, συρρίκνωσε τον όγκο του μαστού της. Αυτή η μείωση του όγκου μετά την αφαίρεση των ωοθηκών οφειλόταν στο γεγονός ότι τα οιστρογόνα από τις ωοθήκες συνέβαλλαν στην ανάπτυξη του όγκου και έτσι, με την αφαίρεσή τους, το μέγεθος του όγκου παρουσίαζε μείωση. Ως αποτέλεσμα, πολλοί χειρουργοί άρχισαν να αφαιρούν και τις δύο ωοθήκες και να πραγματοποιούν ριζική μαστεκτομή όταν είχαν να αντιμετωπίσουν περιπτώσεις καρκίνου του μαστού.

1.3.3 Οι σημερινές προσεγγίσεις για την θεραπεία του καρκίνου του μαστού

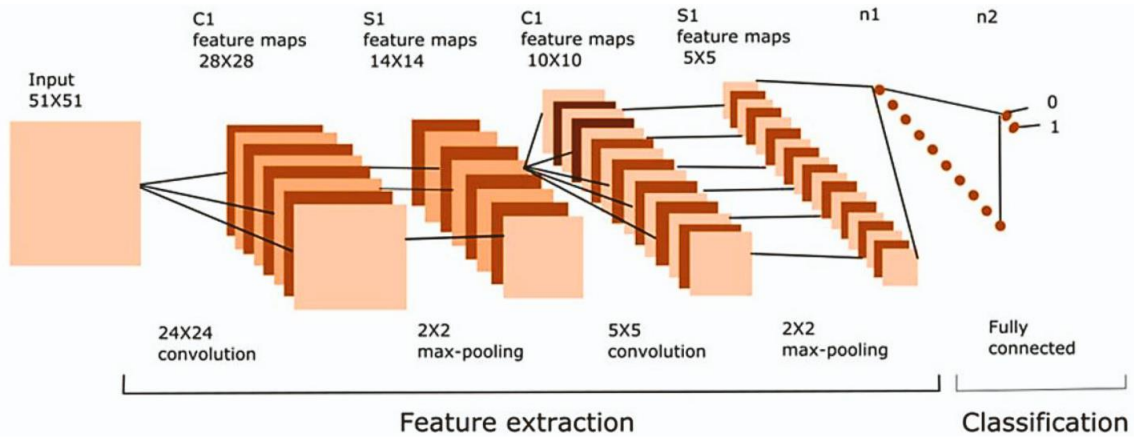
Καθώς οι γιατροί σήμερα αποκτούν περισσότερες γνώσεις για τον καρκίνο του μαστού, η θεραπεία που παρέχουν στους ασθενείς γίνεται όλο και πιο εξατομικευμένη. Επί του παρόντος, ο καρκίνος αναγνωρίζεται ως ασθένεια με διάφορους υπο-τύπους που επηρεάζουν το σώμα με ποικίλους τρόπους και μοτίβα. Πιο εξειδικευμένες επιλογές θεραπείας έχουν ήδη ξεκινήσει να γίνονται διαθέσιμες, παρέχοντας τη δυνατότητα εντοπισμού συγκεκριμένων γονιδίων.

Η αποτελεσματική διάγνωση του καρκίνου του μαστού σήμερα εξαρτάται κυρίως από τις μεθόδους μηχανικής μάθησης (Machine Learning – ML) και τη διάγνωση με τη βοήθεια υπολογιστή (Computer-Aided Diagnostics – CAD). Αυτά τα βασικά εργαλεία χρησιμοποιούν εξελιγμένους αλγόριθμους για την αξιολόγηση των βιοϊατρικών εικόνων και βοηθούν τους γιατρούς να εντοπίσουν και να κατηγοριοποιήσουν τους καρκίνους.

Ειδικότερα, τα συστήματα CAD βοηθούν τόσο τους ακτινολόγους όσο και άλλους γιατρούς να αξιολογούν τις ιατρικές εικόνες χρησιμοποιώντας αλγόριθμους μηχανικής μάθησης. Αυτά τα συστήματα έχουν τη δυνατότητα να εντοπίζουν αυτόματα και να υποδεικνύουν τυχόν «προβληματικές» περιοχές στις εικόνες, επισημαίνοντας συγκεκριμένα σημεία που θα μπορούσαν να χρειαστούν περισσότερη εξέταση. Η CAD παρέχει στους εργαζόμενους στον τομέα της υγειονομικής περίθαλψης μια «δεύτερη γνώμη» βοηθώντας στον εντοπισμό ανωμαλιών που μπορεί να είναι δύσκολο να παρατηρηθούν με γυμνό μάτι.

Η ακριβής αναγνώριση των καρκίνων του μαστού στις ιατρικές φωτογραφίες καθίσταται δυνατή με προσεγγίσεις τμηματοποίησης (segmentation) που βασίζονται στη μηχανική μάθηση. Αυτές οι μέθοδοι βοηθούν στον εντοπισμό της ακριβούς θέσης των κακοήθων κυττάρων στον ιστό του μαστού, οριοθετώντας με ακρίβεια τους όγκους. Ο σχεδιασμός της θεραπείας απαιτεί αυτό το επίπεδο τμηματοποίησης και εντοπισμού, καθώς παρέχει συγκεκριμένες πληροφορίες και λεπτομέρειες σχετικά με τη θέση και το μέγεθος των όγκων, επιτρέποντας πιο στοχευμένες θεραπείες και παρεμβάσεις.

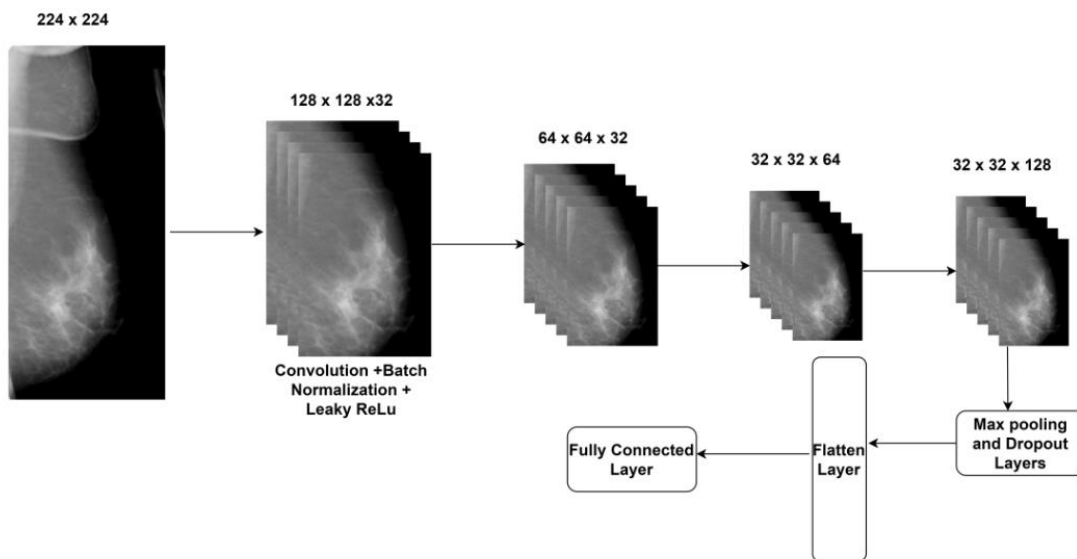
Για την επεξεργασία και ανάλυση μαζικών συνόλων δεδομένων ιατρικών εικόνων, (όπως μαγνητική τομογραφία, ακτινογραφία και μαστογραφία) χρησιμοποιούνται μέθοδοι μηχανικής μάθησης, όπως μοντέλα βαθιάς μάθησης. Αυτοί οι αλγόριθμοι είναι σε θέση να αναγνωρίζουν μοτίβα και χαρακτηριστικά στις φωτογραφίες που μπορούν να υποδεικνύουν την ύπαρξη όγκων ή κακοήθων κυττάρων. Αυτές οι μέθοδοι μπορούν να βοηθήσουν στην έγκαιρη αναγνώριση του καρκίνου του μαστού μαθαίνοντας να διακρίνουν μεταξύ φυσιολογικού και καρκινικού ιστού μέσω της χρήσης μοντέλων μηχανικής μάθησης που έχουν εκπαιδευτεί σε ήδη χαρακτηρισμένα σύνολα δεδομένων.



Εικόνα 1.7 Παράδειγμα ενός συνελκτικού νευρωνικού δικτύου (CNN) με πολλαπλά επίπεδα

Η ενσωμάτωση συστημάτων μηχανικής μάθησης και διάγνωσης CAD έχει αποδειχθεί ότι βελτιώνει σημαντικά την ακρίβεια στη διάγνωση του καρκίνου του μαστού. Αυτές οι τεχνολογίες επιτρέπουν στα συστήματα να είναι πιο ευαίσθητα και να παρέχουν πιο αξιόπιστα αποτελέσματα, βοηθώντας έτσι στην ερμηνεία των ιατρικών εικόνων. Επιπλέον, με την αυτοματοποίηση συγκεκριμένων πτυχών της ανάλυσης εικόνων, τα συστήματα μηχανικής μάθησης και CAD αυξάνουν την αποδοτικότητα, επιτρέποντας στους επαγγελματίες του κλάδου υγείας να επικεντρωθούν περισσότερο στη λήψη κλινικών αποφάσεων και τη φροντίδα των ασθενών.

Η χρήση τεχνικών μηχανικής μάθησης είναι ζωτικής σημασίας για την έγκαιρη διάγνωση του καρκίνου του μαστού. Από την κατάτμηση των όγκων μέχρι την ανίχνευση και κατηγοριοποίηση βιοϊατρικών εικόνων, οι αλγόριθμοι μηχανικής μάθησης βοηθούν στον ακριβή εντοπισμό πιθανών περιοχών καρκινικών κυττάρων. Για παράδειγμα, η εφαρμογή των επαναλαμβανόμενων νευρωνικών δικτύων (RNN – Recurrent Neural Networks) και των προσεγγίσεων βαθιάς μάθησης επιτρέπει την ανίχνευση του καρκίνου σε πραγματικό χρόνο, βοηθώντας στην έγκαιρη διάγνωση και την αποτελεσματικότερη αντιμετώπιση της νόσου.



Εικόνα 1.8 Παράδειγμα ενός συνελκτικού νευρωνικού δικτύου (CNN) που δέχεται εικόνες μαστογραφίας στην είσοδο

Τα τελευταία χρόνια η βαθιά μάθηση, και πιο συγκεκριμένα η εφαρμογή συνελκτικών νευρωνικών δικτύων (CNN – Convolutional Neural Networks), έχει δείξει επιτυχημένα αποτελέσματα κατηγοριοποίησης. Ορισμένα πρόσφατα ανεπτυγμένα συστήματα CAD είχαν καλά αποτελέσματα αξιοποιώντας δημοφιλείς αλγόριθμους βαθιάς μάθησης. Η σημαντική πρόοδος και τα εξαιρετικά αποτελέσματα που έχουν σημειώσει αυτές οι μέθοδοι στους τομείς της μηχανικής όρασης (computer vision) και της επεξεργασίας εικόνας έχουν κινητοποιήσει πολλούς ερευνητές να εφαρμόσουν τη βαθιά μάθηση για την κατηγοριοποίηση βιοϊατρικών εικόνων.

Τα συνελκτικά νευρωνικά δίκτυα (CNN), που είναι ένας από τους δημοφιλέστερους τύπους βαθιάς μάθησης, έχουν αποδειχθεί αξιόπιστα στον εντοπισμό και στην κατηγοριοποίηση οπτικών χαρακτηριστικών. Ωστόσο, για να δημιουργηθούν αποτελεσματικά βαθιά CNN, απαιτούνται μεγάλα κατηγοριοποιημένα σύνολα δεδομένων, τα οποία στον τομέα της ιατρικής, ιδίως όσον αφορά τον καρκίνο του μαστού, σπανίζουν. Επιπλέον, η πολυπλοκότητα της ανατομίας του μαστού, τα δευτερεύοντα χαρακτηριστικά (όπως περιοχές ανωμαλίας, δομή των ιστών, κ.α.) και η ανθρώπινη κόπωση μπορεί να επηρεάσουν αρνητικά την ακρίβεια της διαδικασίας διάγνωσης.

Συνεπώς, παρουσιάζεται η ανάγκη για χρήση μεθόδων προ-επεξεργασίας των ψηφιακών εικόνων με σκοπό τη βελτίωση της ποιότητάς τους και την ανάδειξη ορισμένων σημαντικών χαρακτηριστικών. Οι μέθοδοι αυτές περιλαμβάνουν διάφορες τεχνικές βελτίωσης και ανάλυσης εικόνων, όπως η αφαίρεση θορύβου, η ενίσχυση της αντίθεσης και η απομόνωση σημαντικών λεπτομερειών. Η σωστή προ-επεξεργασία συμβάλλει όχι μόνο στην ορθότερη ανάλυση, αλλά και στη μείωση του χρόνου που απαιτείται για την επεξεργασία των δεδομένων, προετοιμάζοντας κατάλληλα τις εικόνες για την εφαρμογή αλγορίθμων μηχανικής μάθησης και διασφαλίζοντας έτσι την εξαγωγή πιο αξιόπιστων διαγνωστικών αποτελεσμάτων και την επίτευξη μεγαλύτερης ακρίβειας στα συστήματα κατηγοριοποίησης.

ΚΕΦΑΛΑΙΟ 2^ο: Επεξεργασία ψηφιακών βιοϊατρικών εικόνων

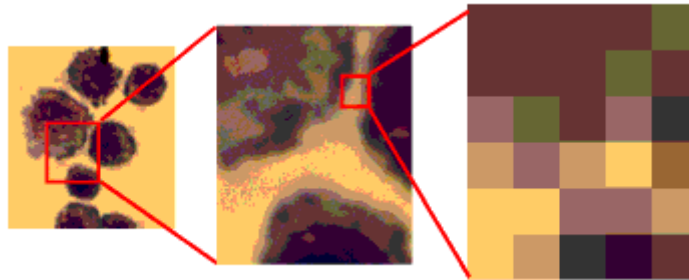
Στην εποχή της ψηφιοποίησης και της ταχείας ανάπτυξης της τεχνολογίας, οι εικόνες αποτελούν κυρίαρχο μέσο για τη μεταφορά πληροφοριών και την επικοινωνία. Από τα μέσα κοινωνικής δικτύωσης έως και τις ιατρικές εφαρμογές, οι ψηφιακές εικόνες κατακλύζουν τον ψηφιακό χώρο, δημιουργώντας την ανάγκη για αποτελεσματική διαχείριση και ανάλυση.

Η επεξεργασία των ψηφιακών εικόνων αποτελεί ένα βασικό στάδιο στη διαχείριση και στην ανάλυσή τους. Αυτή η διαδικασία περιλαμβάνει μια σειρά από τεχνικές που εφαρμόζονται στις εικόνες, με στόχο τη βελτίωση της ποιότητάς τους ή την απόκτηση σημαντικών πληροφοριών από αυτές. Ανάμεσα σε αυτές είναι η εφαρμογή φίλτρων, η μείωση του θορύβου, η βελτίωση της ευκρίνειας, και άλλες τεχνικές.

Πλέον, η ανάγκη για γρήγορη και αποτελεσματική ανάλυση εικόνων έχει αυξηθεί παράλληλα με την αύξηση των εφαρμογών που χρησιμοποιούν τεχνικές αναγνώρισης προτύπων, όπως η ραγδαία εξέλιξη της τεχνητής νοημοσύνης ακόμα και σε απλές καθημερινές δραστηριότητες. Αυτό οδηγεί σε αυξημένη ανάγκη για βέλτιστη και γρήγορη επεξεργασία των εικόνων και για την επιτυχή ανάλυση και ερμηνεία τους.

2.1 Ορισμός Εικόνας και Διαστάσεις

Μια ψηφιακή εικόνα μπορεί να οριστεί ως μια *διδιάστατη αναπαράσταση του χώρου*, η οποία περιλαμβάνει πληροφορία σχετικά με τα χαρακτηριστικά ενός αντικειμένου, σκηνής ή φαινομένου. Από την κατανόηση της δομής και των διαστάσεων της εικόνας, αναπτύσσεται η ικανότητα αναγνώρισης προτύπων, εξαγωγής πληροφοριών και η χρήση διάφορων τεχνικών επεξεργασίας για την ανάλυση ή τη βελτίωση της εικόνας. Συνεπώς, η κατανόηση του ορισμού της είναι ουσιαστική για την ανάπτυξη αλγορίθμων επεξεργασίας εικόνων και την ανάλυση των δεδομένων που περιέχονται σε αυτές.



Εικόνα 2.1 Μεγέθυνση σε εικονοστοιχεία (pixels) μιας εικόνας

Οι τρεις κύριες μορφές που μπορεί να έχει μια ψηφιακή εικόνα είναι:

1. **δωαδική (binary)**, που αποτελείται μόνο από άσπρα και μαύρα εικονοστοιχεία (pixels),
2. **μονοχρωματική (grayscale)**, που περιέχει μόνο αποχρώσεις του γκρι,
3. **έγχρωμη (color)**, που περιλαμβάνει πολλαπλά κανάλια χρωμάτων.

Οι διάφορες μορφές των ψηφιακών εικόνων παρέχουν διαφορετικά επίπεδα πληροφορίας και δυνατοτήτων επεξεργασίας, εξυπηρετώντας έτσι διαφορετικές ανάγκες και εφαρμογές.

2.1.1 Δυαδική Εικόνα

Μια εικόνα σε δυαδική μορφή αναπαρίσταται μέσω ενός πίνακα διαστάσεων $M \times N$, όπου κάθε στοιχείο του αντιπροσωπεύει ένα pixel και μπορεί να πάρει μόνο δύο τιμές: 1 (αληθής) ή 0 (ψευδής). Συνήθως, το μαύρο αντιστοιχεί στην τιμή "0", ενώ το άσπρο αντιστοιχεί στην τιμή "1".



Εικόνα 2.2 Δυαδική (binary) εικόνα

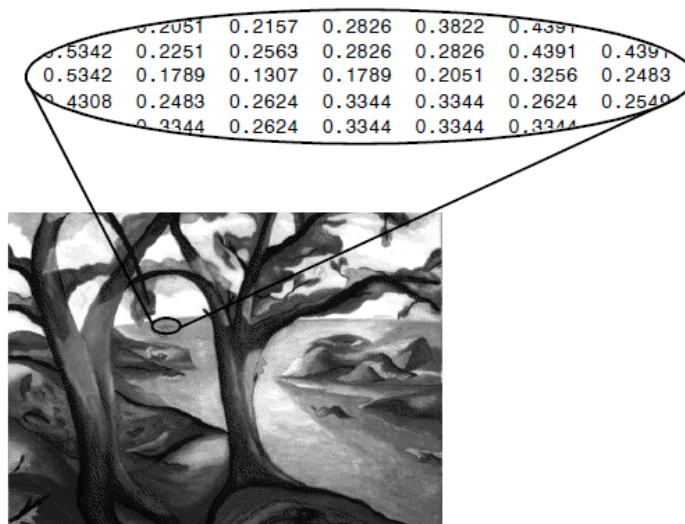
Αυτή η μορφή εικόνας είναι η πιο απλή και για αυτόν τον λόγο απαιτεί λιγότερη μνήμη και μικρότερο υπολογιστικό κόστος όταν πρόκειται για την επεξεργασία της. Παρά την απλότητά της, μπορεί να μεταφέρει σημαντικές πληροφορίες, όπως το εμβαδόν και τη θέση αντικειμένων, καθώς και τα χαρακτηριστικά τους. Οι δυαδικές εικόνες εξυπηρετούν πολλές εφαρμογές της Ψηφιακής Επεξεργασίας Εικόνας, όπως η αναγνώριση χαρακτήρων, η αναγνώριση υπογραφής και η αναγνώριση αποτυπωμάτων.

2.1.2 Μονοχρωματική Εικόνα

Η μονοχρωματική ψηφιακή εικόνα περιγράφεται ως ένας πίνακας διαστάσεων $M \times N$, με κάθε στοιχείο του πίνακα να αντιστοιχεί σε ένα pixel. Οι τιμές αυτών των στοιχείων αντιστοιχούν στη φωτεινότητα του αντίστοιχου σημείου της εικόνας.

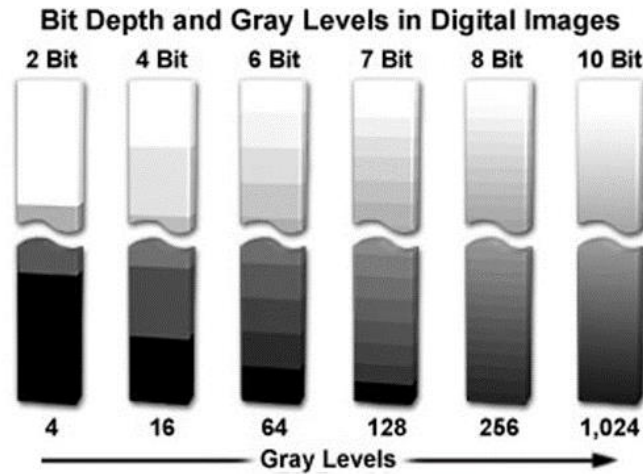
$$I(i, j) \text{ με } i = 1, \dots, M \text{ και } j = 1, \dots, N$$

Η τιμή κάθε pixel, συμβολίζεται ως $I(i, j)$, όπου i και j είναι οι συντεταγμένες του pixel στον πίνακα, με $i = 1, \dots, M$ και $j = 1, \dots, N$.



Εικόνα 2.3 Μονοχρωματική (grayscale) εικόνα

Οι τιμές των pixel κυμαίνονται μεταξύ 0 και $L - 1$, όπου το L είναι συνήθως μια δύναμη του 2, δηλαδή $L = 2^k$, και το k είναι συνήθως 8, που αντιστοιχεί σε 256 αποχρώσεις του γκρι.



Εικόνα 2.4 Βάθος (k) και αποχρώσεις του γκρι (L)

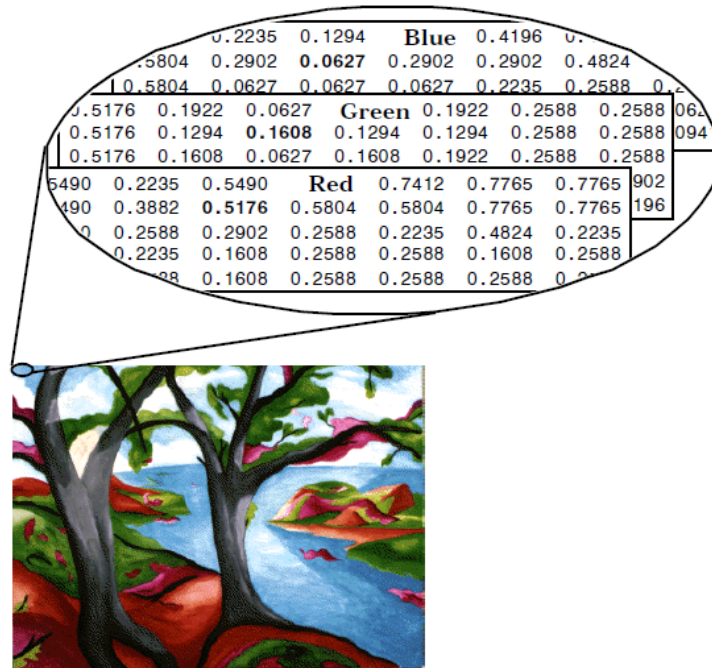
2.1.3 Έγχρωμη Εικόνα

Η έγχρωμη ψηφιακή εικόνα χρησιμοποιείται ευρέως για την αναπαράσταση της πραγματικότητας. Σε μια έγχρωμη εικόνα το κάθε pixel αναπαρίσταται ως ένας συνδυασμός τριών μονοχρωματικών συνιστωσών. Δηλαδή μια έγχρωμη εικόνα είναι τρεις μονοχρωματικές, όπου η τιμή της φωτεινότητας σε καθεμία από αυτές δημιουργεί το τελικό χρώμα του pixel της έγχρωμης εικόνας.

$$I_c(i, j) \text{ με } i = 1, \dots, M \text{ και } j = 1, \dots, N$$

Όπου για κάθε $c = 1, 2, 3$ οι τιμές των pixel κυμαίνονται μεταξύ 0 και $L - 1$. Τελικά το χρώμα του pixel έχει προκύψει συνδυάζοντας τις τρεις χρωματικές συνιστώσες:

$$Color(i, j) = [I_1(i, j), I_2(i, j), I_3(i, j)]$$



Εικόνα 2.5 Έγχρωμη (color) εικόνα

Το χρωματικό σύστημα RGB αποτελεί ένα συνηθισμένο παράδειγμα όπου το κάθε χρώμα αποτελείται από τη συνδυαστική παρουσία των τριών βασικών χρωμάτων: κόκκινο (**Red**), πράσινο (**Green**) και μπλε (**Blue**).



Εικόνα 2.6 Μεταβολή της ευκρίνειας μιας έγχρωμης εικόνας διατηρώντας σταθερό το βάθος χρώματος

Ως *ευκρίνεια* μιας εικόνας ορίζεται η ικανότητα να διακρίνονται σαφώς οι λεπτομέρειές της και εξαρτάται από το πλήθος των pixels που την απαρτίζουν (ανάλυση) όπως επίσης και από το πλήθος των αποχρώσεών της.

Το πλήθος των αποχρώσεων (L) σε μια εικόνα εξαρτάται από το *βάθος χρώματος* k και υπολογίζεται ως 2^k . Συνολικά, τα bits που χρειάζονται για την εικόνα είναι $M \times N \times k$. Όσο μεγαλύτερο είναι το βάθος χρώματος τόσο περισσότερες διαφορετικές αποχρώσεις μπορούν να αναπαρασταθούν. Μια έγχρωμη εικόνα χρειάζεται τα τριπλάσια bits συγκριτικά με μια μονοχρωματική εικόνα.

Αν σε μια εικόνα μειωθεί η ανάλυση διατηρώντας σταθερό το βάθος χρώματος, τότε ως αποτέλεσμα θα έχουμε μια εικόνα «σκακιέρα», όπου τα pixels θα διατηρούνται σε ομοιόμορφα χρωματικά ζώνες (Εικόνα 2.6). Αντίστοιχα, αν σε μια εικόνα μειωθεί το βάθος χρώματος διατηρώντας σταθερές τις διαστάσεις της, τότε εμφανίζονται και πάλι ομοιόμορφες περιοχές που εξαρτώνται από το βάθος χρώματος και την ομοιότητα των αποχρώσεων με τις γειτονικές τους (Εικόνα 2.7).

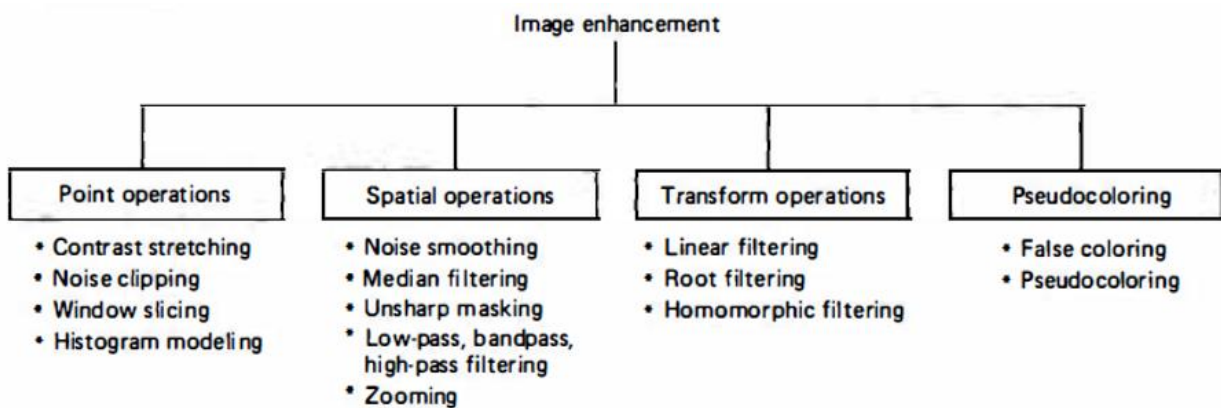


Εικόνα 2.7 Μεταβολή του βάθους χρώματος μιας εικόνας διατηρώντας σταθερές τις διαστάσεις της

2.2 Βελτίωση Ψηφιακής Εικόνας στο Πεδίο της Συχνότητας

Ο όρος «βελτίωση της ψηφιακής εικόνας» αναφέρεται στην ενίσχυση των χαρακτηριστικών της, όπως η αντίθεση, τα όρια ή οι άκρες, προκειμένου να γίνουν πιο ευδιάκριτα και κατανοητά κατά την προβολή και την ανάλυσή τους. Αν και αυτή η διαδικασία δεν αυξάνει την πληροφορία που περιέχεται στην εικόνα, αυξάνει το εύρος των χαρακτηριστικών που είναι ανιχνεύσιμα και μπορούν να αναλυθούν.

Μερικές από τις τεχνικές που περιλαμβάνει η διαδικασία βελτίωσης της εικόνας είναι η ρύθμιση του επιπέδου γκρι και της αντίθεσης, η μείωση του θορύβου, η ενίσχυση των ακμών, το φιλτράρισμα, η μεγέθυνση αλλά και η εφαρμογή ψευδοχρωμάτων.



Εικόνα 2.8 Τεχνικές Βελτίωσης Εικόνας

Επειδή ωστόσο πολλές από αυτές τις τεχνικές βασίζονται σε εμπειρικές μεθόδους και συχνά απαιτούνται δοκιμές και διαδραστικές διαδικασίες προκειμένου να επιτευχθεί ένα επιθυμητό αποτέλεσμα, είναι δύσκολο να καθοριστεί ένα συγκεκριμένο κριτήριο για την αποτελεσματικότητα της βελτίωσης. Κάθε εφαρμογή έχει διαφορετικές απαιτήσεις και στόχους, και επομένως, η επιτυχία μιας τεχνικής βελτίωσης εικόνας πρέπει να κρίνεται με βάση το αν επιτυγχάνει τα συγκεκριμένα αποτελέσματα που χρειάζονται για εκείνη την εφαρμογή. Για παράδειγμα, σε ιατρικές εικόνες η βελτίωση μπορεί να στοχεύει στην καλύτερη ανίχνευση λεπτομερειών ή ανωμαλιών ενώ στην αστρονομία μπορεί να χρειάζεται να μειωθεί ο θόρυβος για να φανούν τα πιο ασθενή ουράνια σώματα.

2.2.1 Ο Μετασχηματισμός Fourier

Ο μετασχηματισμός Fourier είναι εξαιρετικά σημαντικός στην ψηφιακή επεξεργασία εικόνας, καθώς χρησιμοποιείται για την ενίσχυση, την ανάλυση, την αποκατάσταση και τη συμπίεση εικόνων. Ειδικά, ο 2D μετασχηματισμός Fourier εφαρμόζεται σε δισδιάστατες εικόνες για να μετασχηματίσει τα δεδομένα από το πεδίο του χώρου στο πεδίο της συχνότητας, επιτρέποντας την ανάλυση των συχνοτικών περιεχομένων της εικόνας.

Ο 2-D μετασχηματισμός Fourier μιας διδιάστατης εικόνας $f(m, n)$ είναι:

$$F(\omega_1, \omega_2) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} f(m, n) e^{-j\omega_1 m} e^{-j\omega_2 n}$$

Όπου

- $f(m, n)$ είναι η αρχική εικόνα στο πεδίο του χώρου, με m και n να είναι οι συντεταγμένες των pixel.
- $F(\omega_1, \omega_2)$ είναι η εικόνα στον χώρο της συχνότητας μετά την εφαρμογή του μετασχηματισμού Fourier, με ω_1 και ω_2 να είναι οι κυκλικές συχνότητες.
- Ο όρος $e^{-j\omega_1 m}$ είναι η εκθετική συνάρτηση που συνδέει τη συχνότητα ω_1 με τη συντεταγμένη m .
- Ο όρος $e^{-j\omega_2 n}$ είναι η εκθετική συνάρτηση που συνδέει τη συχνότητα ω_2 με τη συντεταγμένη n .

Δηλαδή, η αρχική εικόνα μετατρέπεται σε μια αναπαράσταση συχνοτικών συνιστωσών, που δείχνει ποιες συχνότητες υπάρχουν στην εικόνα και πόσο ισχυρές είναι. Αυτή η πληροφορία είναι εξαιρετικά χρήσιμη για διάφορες εφαρμογές, όπως:

1. **Ενίσχυση Χαρακτηριστικών Εικόνας:** Εντοπισμός και ενίσχυση συγκεκριμένων συχνοτήτων για βελτίωση της ευκρίνειας.
2. **Αποκατάσταση Εικόνας:** Αφαίρεση θορύβου που εμφανίζεται σε συγκεκριμένες συχνότητες.
3. **Συμπίεση Εικόνας:** Αποθήκευση μόνο των σημαντικών συχνοτήτων για μείωση του όγκου δεδομένων.

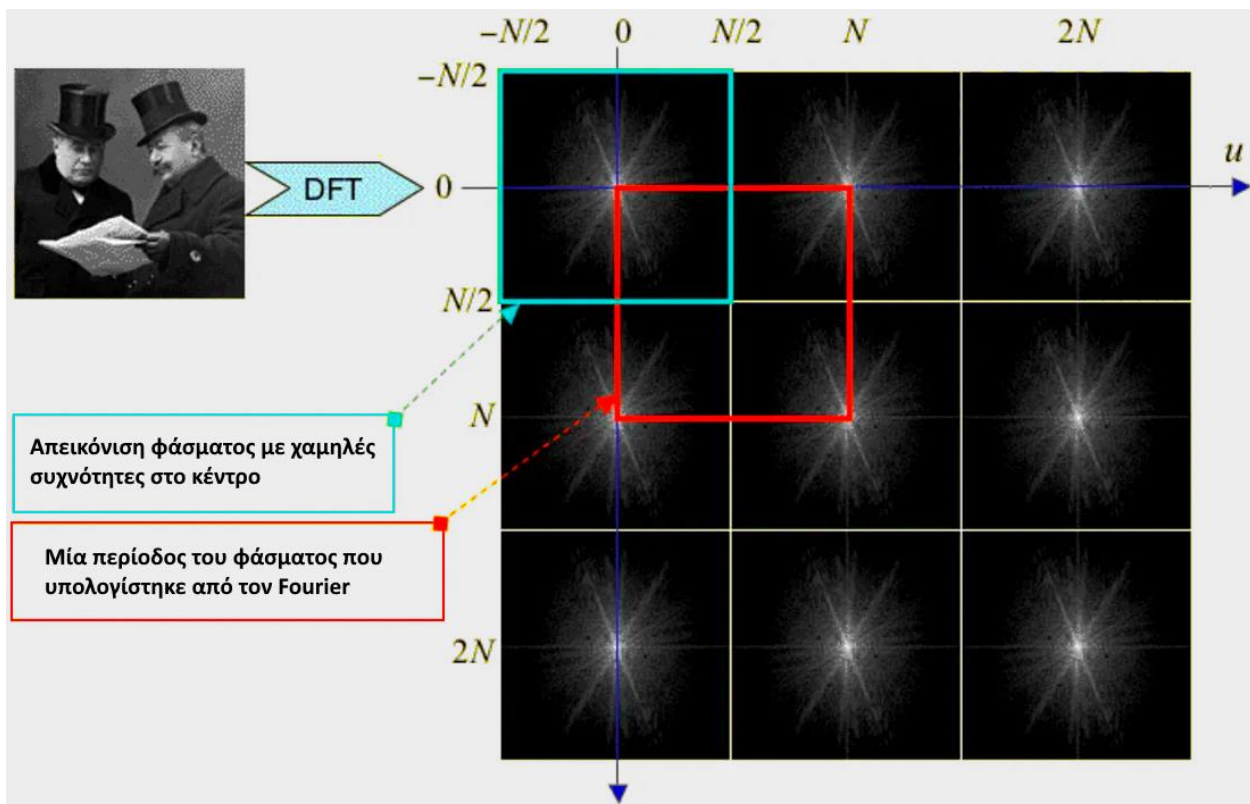


Εικόνα 2.9 (α) Αρχική εικόνα (β) Χάρτης κατανομής συχνότητας μετά τον μετασχηματισμό Fourier.

Για την επαναφορά της εικόνας στο πεδίο του χώρου, χρησιμοποιείται ο αντίστροφος μετασχηματισμός Fourier σύμφωνα με την εξίσωση:

$$f(m, n) = \frac{1}{2\pi} \int_{\omega_1=-\pi}^{\pi} \int_{\omega_2=-\pi}^{\pi} F(\omega_1, \omega_2) e^{j\omega_1 m} e^{j\omega_2 n}$$

Το φάσμα Fourier στο πεδίο της συχνότητας θεωρείται ότι είναι περιοδικό και επαναλαμβάνεται προς όλες τις κατευθύνσεις επ’ άπειρον (Εικόνα 2.10). Συνήθως προτιμάται να γίνεται μετακίνησή του, τέτοια ώστε η πληροφορία που βρίσκεται στις χαμηλές συχνότητες να βρίσκεται στο κέντρο (Εικόνα 2.9).



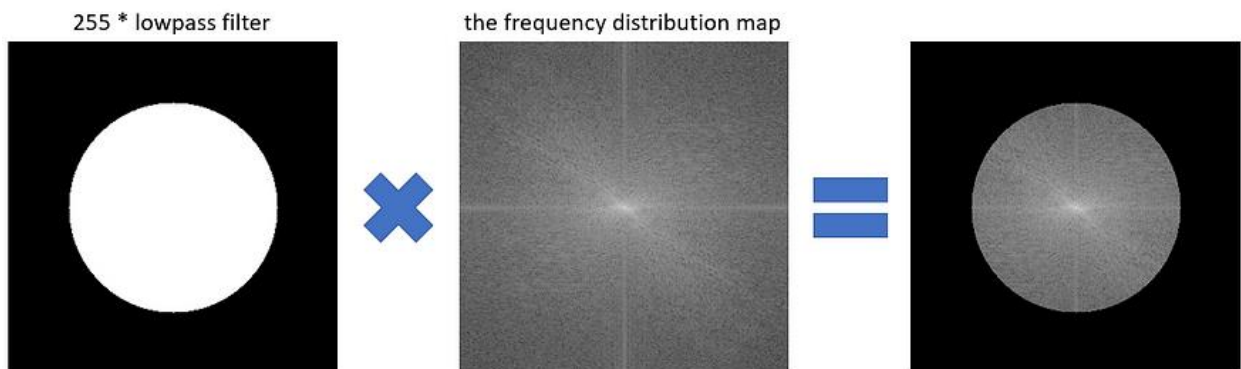
Εικόνα 2.10 Απεικόνιση του άπειρου περιοδικού φάσματος μετά την εφαρμογή μετασχηματισμού Fourier στην αρχική εικόνα.

2.2.2 Φίλτρα Συχνότητας

Η χρήση φίλτρων στο πεδίο της συχνότητας έχει τον ρόλο επιλογής ή απόρριψης ορισμένων συχνοτήτων και κατά συνέπεια κάποιων χαρακτηριστικών της αρχικής εικόνας. Αν εφαρμοστεί φίλτρο σε συγκεκριμένες συχνότητες, το τελικό αποτέλεσμα που θα προκύψει στο πεδίο του χώρου επηρεάζεται άμεσα από το ποιες συχνότητες θα διατηρηθούν και ποιες όχι. Αυτό μπορεί να οδηγήσει σε εικόνες με επισημασμένες δομές ή εξομαλυσμένα χαρακτηριστικά, ανάλογα με τον τρόπο που προσαρμόζεται το φίλτρο. Έτσι, η χρήση φίλτρων στο πεδίο της συχνότητας επιτρέπει τη διαμόρφωση της τελικής εικόνας σύμφωνα με τις απαιτήσεις και τους στόχους της επεξεργασίας εικόνας.

2.2.2.1 Χαμηλοπερατό (Low Pass) Φίλτρο

Φιλτράροντας τις υψηλές συχνότητες και κρατώντας μόνο τις χαμηλές, η εικόνα γίνεται πιο λεία και εξομαλύνεται. Αυτό μπορεί να έχει ως αποτέλεσμα τη μείωση κάποιων λεπτομερειών στην εικόνα και την εμφάνιση πιο ομαλών μεταβολών μεταξύ περιοχών της εικόνας με υψηλή αντίθεση, αλλά είναι ιδιαίτερα χρήσιμο φίλτρο για τη μείωση του θορύβου. Το φίλτρο αυτό ονομάζεται *low pass* ή *χαμηλοπερατό*, αφού επιτρέπει την διέλευση των χαμηλών συχνοτήτων και αποκόπτει τις υψηλές.



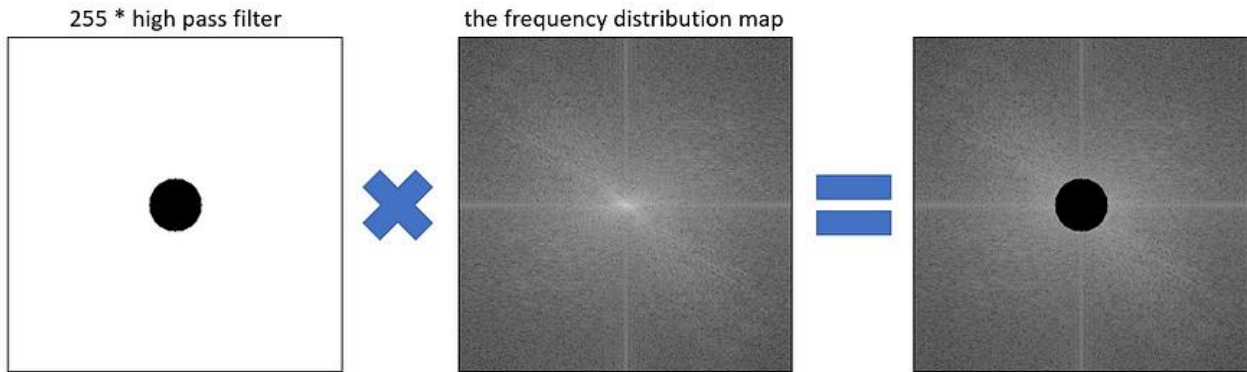
Εικόνα 2.11 Εφαρμογή low pass φίλτρου στο πεδίο της συχνότητας, πολλαπλασιάζοντας με 0 (μαύρο) τις υψηλές συχνότητες και διατηρώντας τις χαμηλές στο κέντρο του συχνοτικού πεδίου που πολλαπλασιάζονται με 1 (άσπρο).



Εικόνα 2.12 Αποτέλεσμα low pass φίλτρου στο πεδίο του χώρου.

2.2.2.2 Υψηλοπερατό (High Pass) Φίλτρο

Φιλτράροντας τις χαμηλές συχνότητες, ενισχύονται οι λεπτομέρειες και οι «αιχμηρές» αλλαγές στην εικόνα. Αυτό μπορεί να οδηγήσει σε πιο επισημασμένες δομές και στον τονισμό λεπτομερειών όπως είναι τα περιγράμματα μιας εικόνας, κάτι που μπορεί να είναι χρήσιμο σε εφαρμογές όπως η ανίχνευση αντικειμένων ή η αναγνώριση προτύπων. Το φίλτρο αυτό ονομάζεται *high pass* ή *υψηλοπερατό*, αφού επιτρέπει την διέλευση των υψηλών συχνοτήτων και αποκόπτει τις χαμηλές.



Εικόνα 2.13 Εφαρμογή high pass φίλτρου στο πεδίο της συχνότητας, πολλαπλασιάζοντας με 0 (μαύρο) τις χαμηλές συχνότητες στο κέντρο του συχνωτικού πεδίου και διατηρώντας τις υψηλές συχνότητες που πολλαπλασιάζονται με 1 (άσπρο).



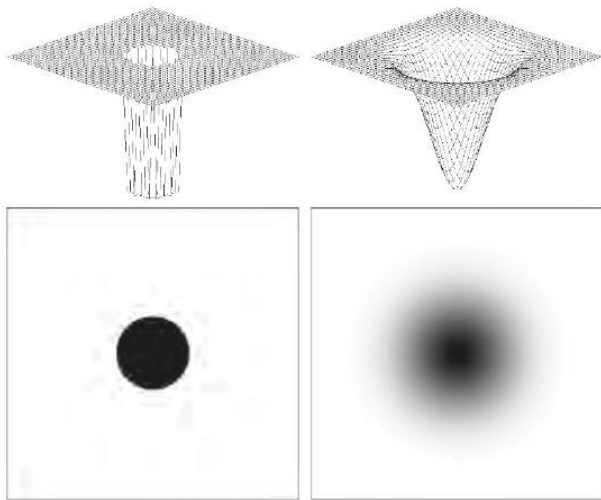
Εικόνα 2.14 Αποτέλεσμα high pass φίλτρου στο πεδίο του χώρου.

2.2.2.3 Γκαουσιανό (Gaussian) Φίλτρο

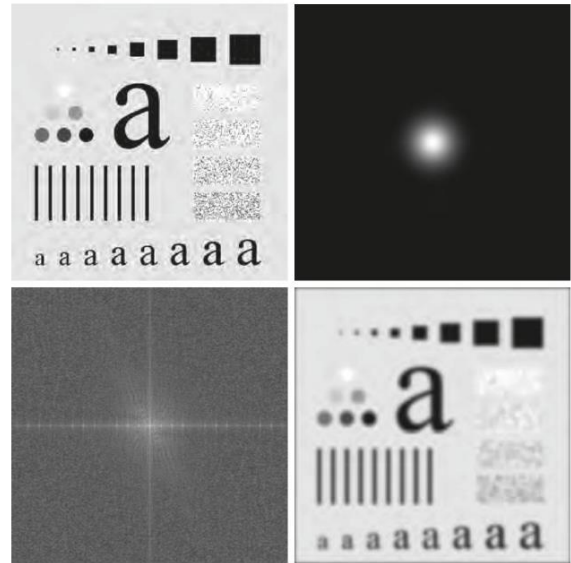
Το Γκαουσιανό φίλτρο είναι ένα είδος χαμηλοπερατού ή υψηλοπερατού φίλτρου. Η βασική διαφορά με τα δύο προηγούμενα είναι ο τρόπος που γίνεται η εξομάλυνση των συχνοτήτων στην εικόνα.

$$H(u, v) = e^{-D(u,v)^2/(2\sigma^2)}$$

Όπου $D(u, v)$ είναι η Ευκλείδεια απόσταση στον συχνωτικό χώρο μεταξύ του κέντρου της συχνότητας και του σημείου (u, v) , και σ είναι η τυπική απόκλιση, η παράμετρος δηλαδή που ελέγχει το μέγεθος του πυρήνα. Το μέγεθος του πυρήνα είναι ανάλογο του σ .



Εικόνα 2.15 Αναπαράσταση ενός (α) απλού high pass φίλτρου (β) Γκαουσιανού high pass φίλτρου



Εικόνα 2.16 Εφαρμογή ενός Γκαουσιανού low pass φίλτρου

Το Γκαουσιανό φίλτρο χρησιμοποιεί μια γκαουσιανή συνάρτηση για να κατασκευάσει έναν πυρήνα, ο οποίος είναι πυκνότερος στο κέντρο και αραιότερος στις άκρες, προσδίδοντας έτσι ένα συγκεκριμένο βάρος σε κάθε περιοχή γύρω από το σημείο της κεντρικής συχνότητας. Αυτό έχει ως αποτέλεσμα την εξομάλυνση της εικόνας, διατηρώντας παράλληλα την αιχμηρότητα των ακμών και των λεπτομερειών.

Το υψηλοπερατό γκαουσιανό φίλτρο προκύπτει από το αντίστροφο του χαμηλοπερατού γκαουσιανού φίλτρου. Μπορεί δηλαδή να υπολογιστεί ως $1 - H(u, v)$, όπου $H(u, v)$ είναι η τιμή του χαμηλοπερατού φίλτρου.

2.3 Επεξεργασία Εικόνων στην Βιοϊατρική

Η επεξεργασία των ψηφιακών εικόνων, ειδικά μετά τη μετάβαση από αναλογικές σε ψηφιακές εικόνες, αποτελεί ένα απαραίτητο εργαλείο σε πολλούς τομείς προσφέροντας τη δυνατότητα βελτίωσης, ανάλυσης και αποτελεσματικότερης αποθήκευσης.

Οι βιοϊατρικές εικόνες, όπως οι ακτινογραφίες, οι μαγνητικές τομογραφίες και οι υπέρηχοι, μπορούν να παρέχουν ουσιαστικές πληροφορίες για την ανάλυση και την αξιολόγηση της κατάστασης του ανθρώπινου σώματος, στις οποίες οι ειδικοί μπορούν να έχουν άμεση πρόσβαση ώστε να τις επεξεργαστούν και να τις αξιολογήσουν. Ορισμένες τεχνικές επεξεργασίας εικόνας επιτρέπουν την ενίσχυση της ποιότητας, την εξαγωγή σημαντικών χαρακτηριστικών και πληροφοριών από αυτές όπως επίσης και τη διευκόλυνση της διαδικασίας της ανάλυσης από τους ιατρούς. Μέθοδοι όπως η απομάκρυνση θορύβου, η ενίσχυση της αντίθεσης, των ακμών, η εφαρμογή φίλτρων βοηθούν στην ακριβέστερη απεικόνιση και στην βελτίωση της εικόνας γενικότερα διευκολύνοντας παράλληλα και την ανάλυση των βιοϊατρικών δεδομένων. Ως αποτέλεσμα, οι ιατροί έχουν την δυνατότητα να παρακολουθούν την πορεία μιας ασθένειας, να εκτιμούν τα αποτελέσματα, να λαμβάνουν άμεσες αποφάσεις και να παρέχουν σωστή καθοδήγηση όσον αφορά την θεραπεία των ασθενών.

2.3.1 Ποιότητα της Εικόνας

Η ποιότητα της ιατρικής εικόνας είναι καθοριστική για την ακριβή διάγνωση και αξιολόγηση των παθήσεων. Για αυτόν τον λόγο, η επεξεργασία αυτών των εικόνων θα πρέπει να στοχεύει στη βελτίωση της ποιότητας στο μέγιστο δυνατό βαθμό. Μεγαλύτερη ακρίβεια και ευκρίνεια επιτρέπουν όχι μόνο στους ιατρούς να εντοπίζουν ακόμη και τις πιο μικρές λεπτομέρειες, οι οποίες μπορεί να είναι κρίσιμες, αλλά και στην πιο αποτελεσματική επεξεργασία τους από υπολογιστές και αλγορίθμους μηχανικής μάθησης.

Η υψηλή ποιότητα εικόνας μπορεί να επιτρέψει σε τέτοιους αλγορίθμους να ανιχνεύουν πρότυπα και ανωμαλίες με πολύ μεγαλύτερη ακρίβεια, αξιοποιώντας όλη την διαθέσιμη πληροφορία που παρέχει η εικόνα. Αυτό όχι μόνο ενισχύει τη διαγνωστική ικανότητα, αλλά μειώνει επίσης τον χρόνο που απαιτείται για την ερμηνεία των αποτελεσμάτων, επιταχύνοντας τη διαδικασία λήψης αποφάσεων και διευκολύνοντας στην παρακολούθηση της προόδου της θεραπείας και της ανταπόκρισης του ασθενούς.

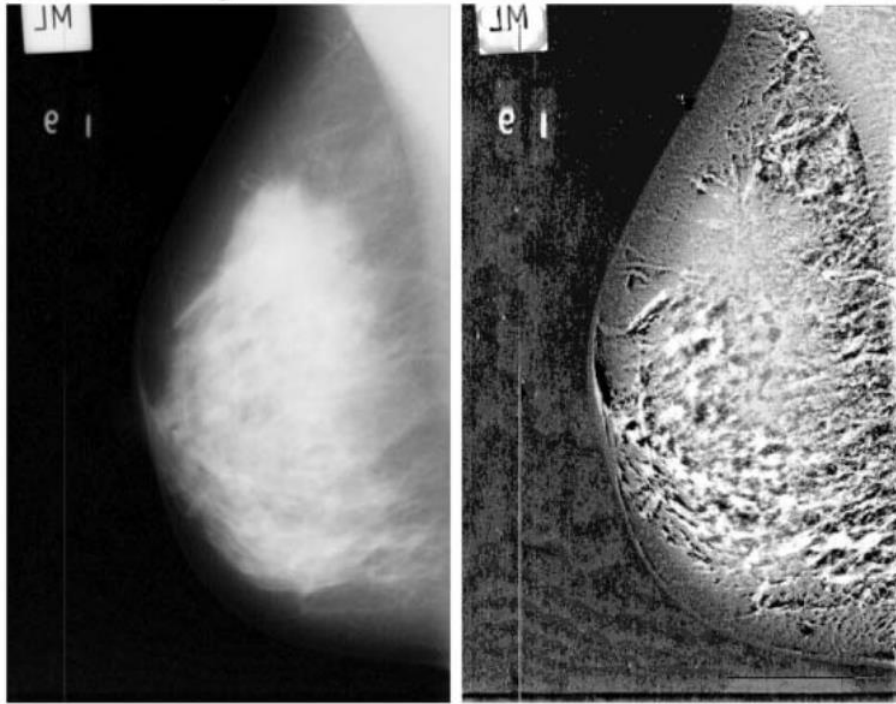
Τα βασικά χαρακτηριστικά που επηρεάζουν την ποιότητα της εικόνας, όπως επίσης και τη διαγνωστική της αξία είναι τα εξής:

- **Χωρική Ανάλυση:** Η χωρική ανάλυση είναι μια ένδειξη της ευκρίνειας της εικόνας και κατά πόσο καλά μπορούν να διαχωρίζονται «γειτονικά» σημεία της εικόνας, τα οποία δηλαδή βρίσκονται κοντά μεταξύ τους. Στις ψηφιακές εικόνες, η χωρική ανάλυση έχει να κάνει με τον αριθμό των pixel ανά μονάδα επιφάνειας.
- **Ανάλυση Αντίθεσης:** Η ανάλυση της αντίθεσης αφορά την ικανότητα διάκρισης μικρών διαφορών στην ένταση σε διάφορες περιοχές της εικόνας. Ένα παράδειγμα θα ήταν οι διαφορές απορρόφησης ακτινών X και πόσο καλά διακρίνεται η διαφορά σε μια τέτοια εικόνα. Στις ψηφιακές εικόνες, η ανάλυση της αντίθεσης σχετίζεται με τον αριθμό των bits ανά pixel.

Χρονική Ανάλυση: Η χρονική ανάλυση έχει να κάνει με τον χρόνο που απαιτείται για τη δημιουργία μιας εικόνας. Μια διαδικασία απεικόνισης που μπορεί να παράγει εικόνες ταυτόχρονα με τη φυσική διαδικασία που απεικονίζει, θεωρείται πραγματικού χρόνου. Για παράδειγμα, είναι δυνατόν να παραχθούν καθαρές εικόνες της καρδιάς που χτυπά, δεδομένου ότι ο ρυθμός των εικόνων που παράγονται είναι τουλάχιστον 30 εικόνες ανά δευτερόλεπτο.

Αυτά τα χαρακτηριστικά καθιστούν δυνατή την αντικειμενική αξιολόγηση των εικόνων που λαμβάνονται χρησιμοποιώντας τεχνικές ψηφιακής απεικόνισης. Η ενίσχυση της αντίθεσης για παράδειγμα (Εικόνα 2.8) αποκαλύπτει και αναδεικνύει διάφορα δομικά χαρακτηριστικά που δύσκολα ανιχνεύονται από το ανθρώπινο μάτι. Συνεπώς, επιλέγοντας την κατάλληλη μέθοδο επεξεργασίας αναλόγως το πρόβλημα, είναι εφικτό να ανιχνευθούν διάφορες λεπτομέρειες που είναι απαραίτητες για την αποτελεσματικότερη ανάλυση των εικόνων και έτσι όλη η πληροφορία που βρίσκεται στην εικόνα να είναι εκμεταλλεύσιμη.

Η επιλογή της καλύτερης προσέγγισης επηρεάζεται επίσης και από άλλους παράγοντες, συμπεριλαμβανομένου του κόστους και του πόσο παρεμβατική είναι η θεραπεία. Μια ιδανική προσέγγιση θα πρέπει να παρέχει εικόνες υψηλής ποιότητας και να είναι απλή, ανώδυνη και μη επεμβατική.



Εικόνα 2.17 Ενίσχυση της αντίθεσης μιας εικόνας από μαστογραφία (α) αρχική εικόνα (β) εικόνα με ενισχυμένη αντίθεση

2.3.2 Μέθοδοι Επεξεργασίας Εικόνων

Οι τεχνικές επεξεργασίας εικόνων μεταχειρίζονται την εικόνα ως ένα διδιάστατο σήμα και εφαρμόζουν μετασχηματισμούς, φίλτρα, και άλλες μαθηματικές λειτουργίες για τη βελτίωση της εικόνας ή την εξαγωγή πληροφοριών από αυτήν.

Οι κυριότερες χρήσεις της επεξεργασίας εικόνας στον χώρο της βιοϊατρικής είναι:

- **η βελτίωση της εικόνας**, με σκοπό την καλύτερη απεικόνιση των λεπτομερειών εκείνων που διαφορετικά θα ήταν δυσδιάκριτες,
- **ο διαγνωστικός έλεγχος**, ο οποίος αυτοματοποιημένα ή ημι-αυτοματοποιημένα χρησιμοποιείται για τον έλεγχο μεγάλων όγκων εικόνων με σκοπό την εύρεση ανωμαλιών ή σημείων που απαιτούν περαιτέρω εξέταση,
- **η ποσοτική ανάλυση**, που περιλαμβάνει την μέτρηση των ποσοτικών παραμέτρων που προκύπτουν από μια εικόνα, όπως η πυκνότητα των ιστών, το μέγεθος ενός όγκου, ή άλλες μετρήσεις.

Οι μέθοδοι που συνήθως χρησιμοποιούνται στην επεξεργασία των εικόνων, μπορούν να χωριστούν σε τρεις κύριες κατηγορίες:

1. **Χαμηλού Επιπέδου:** Η εικόνα αναπαρίσταται με τον πιο απλό τρόπο, χρησιμοποιώντας τα pixel της με τον ελάχιστο δυνατό υπολογιστικό φόρτο. Αυτό επιτρέπει την εστίαση των υπολογιστικών πόρων σε πιο προηγμένα στάδια ανάλυσης, όπως η τμηματοποίηση και η σύγκριση με άλλες εικόνες.

2. **Υψηλού Επιπέδου:** Για κάθε εικόνα αντιστοιχίζεται και μία ετικέτα βάσει του περιεχομένου της, κάτι το οποίο απαιτεί μεγαλύτερη υπολογιστική ισχύ στο στάδιο της αναπαράστασης. Χρησιμοποιούνται επίσης αυτόματες μέθοδοι τμηματοποίησης για την αναγνώριση των περιοχών ενδιαφέροντος και τεχνικές εκμάθησης για την κατηγοριοποίησή τους.
3. **Μεσαίου Επιπέδου:** Η εικόνα μετατρέπεται σε ένα σύνολο χαρακτηριστικών που περιγράφουν τα σημαντικά της στοιχεία. Αυτά τα χαρακτηριστικά μπορεί να είναι η φωτεινότητα, το χρώμα, η δομή ή άλλα χαρακτηριστικά που βοηθούν στην ανάλυσή της. η αναπαράσταση της εικόνας βρίσκεται σε μεσαίο επίπεδο, μεταβαίνοντας από τα pixels σε σημαντικά χαρακτηριστικά. Έτσι, αντί να αποθηκεύονται όλες οι πληροφορίες της εικόνας σε μορφή pixels, χρησιμοποιούνται συγκεντρωμένες περιγραφές των σημαντικών χαρακτηριστικών της εικόνας. Αυτές οι περιγραφές, που ονομάζονται διανύσματα χαρακτηριστικών, αντιπροσωπεύουν το φάσμα του περιεχομένου της εικόνας με λιγότερες πληροφορίες. Η ανάλυση της εικόνας γίνεται σύμφωνα με αυτήν την αναπαράσταση.

Συνήθως χρησιμοποιούνται μία ή περισσότερες από τις ακόλουθες θεμελιώδεις μεθόδους επεξεργασίας εικόνας:

1. **Γενική Επεξεργασία (Global Processing):** Περιλαμβάνει υπολογισμούς που εφαρμόζονται σε ολόκληρη την εικόνα, χωρίς να εστιάζουν σε συγκεκριμένες περιοχές. Ο κύριος σκοπός αυτής της προσέγγισης είναι να βελτιώσει την εικόνα για καλύτερη οπτική αντίληψη από τον άνθρωπο ή για να προετοιμάσει την εικόνα για περαιτέρω ανάλυση από τον υπολογιστή. Για παράδειγμα, στην περίπτωση των ακτινογραφιών, επειδή το ανθρώπινο μάτι δεν μπορεί να διακρίνει όλες τις αποχρώσεις του γκρι που παράγονται από τον σαρωτή, οι τιμές των pixel ρυθμίζονται με τέτοιο τρόπο ώστε οι ιατροί να μπορούν να διακρίνουν καλύτερα τις λεπτομέρειες.
2. **Βελτίωση Εικόνας (Image Enhancement):** Σκοπός της είναι να κάνει την εικόνα πιο ευδιάκριτη μέσω της ενίσχυσης της αντίθεσης, της μείωσης του θορύβου, και άλλες τεχνικές που έχουν ήδη αναφερθεί. Ένα παράδειγμα αυτής της μεθόδου είναι η αφαίρεση μιας θολής εικόνας από την αρχική ώστε να αυξηθεί η τοπική αντίθεση και να ενισχυθούν οι λεπτομέρειες.
3. **Δημιουργία/Οπτικοποίηση Εικόνας (Image Rendering/Visualization):** Περιλαμβάνει τη δημιουργία οπτικών αναπαραστάσεων από τις αρχικές εικόνες για να βοηθήσει στην κατανόηση ή την ανάλυσή τους. Αυτό μπορεί να περιλαμβάνει την τρισδιάστατη απεικόνιση ή τη δημιουργία εικόνων από δεδομένα που δεν είναι απευθείας ορατά. Αυτή η τεχνική είναι ιδιαίτερα χρήσιμη στην ανάλυση και οπτικοποίηση τρισδιάστατων δεδομένων όπως εικόνες από μαγνητική τομογραφία, υπερήχους και PET.
4. **Κβαντισμός Εικόνας (Image Quantitation):** Η μέθοδος αυτή εξάγει αριθμητικές παραμέτρους από εικόνες ή συγκεκριμένες περιοχές τους, οι οποίες μπορεί να είναι χρήσιμες από μόνες τους, όπως το μέγεθος της καρδιάς ή του εμβρύου, ή να χρησιμοποιούνται για ταξινόμηση αντικειμένων. Αυτή η διαδικασία περιλαμβάνει την επεξεργασία ολόκληρης της εικόνας και την τμηματοποίηση για να δοθούν αριθμητικές τιμές σε περιοχές ενδιαφέροντος, χρησιμοποιώντας χαρακτηριστικά που προέρχονται από χρώμα, υφή και σχήμα.
5. **Τμηματοποίηση Εικόνας (Image Segmentation):** Είναι η διαδικασία διαίρεσης μιας εικόνας σε διάφορα τμήματα ή περιοχές για την ευκολότερη ανάλυση. Συνήθως χρησιμοποιείται για την απομόνωση σημαντικών δομών, όπως όργανα ή μέρη αυτών, ή ακόμα και προβληματικούς ιστούς.

Η τμηματοποίηση γίνεται κατά ένα μεγάλο ποσοστό χειροκίνητα, ειδικά όταν υπάρχουν δομές με πολύπλοκα σύνορα, ώστε να επιλεγθούν προσεκτικά τις περιοχές που θεωρούνται σημαντικές για ανάλυση.



Εικόνα 2.18 Τμηματοποίηση εικόνας ανατομίας σώματος για τον διαχωρισμό οργάνων: συκώτι (κόκκινο), στομάχι (πράσινο), σπλήνα (κίτρινο)

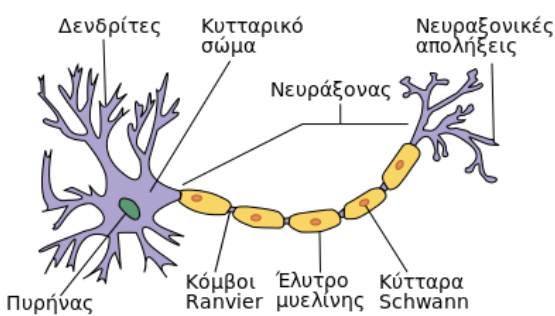
6. **Αντιστοίχιση Εικόνας (Image Registration):** Αφορά τη σύνδεση πολλαπλών εικόνων ώστε να συνδυαστούν σε μία ενιαία αναπαράσταση. Αυτό μπορεί να περιλαμβάνει τον συνδυασμό εικόνων από διαφορετικές χρονικές στιγμές ή διαφορετικές απεικονιστικές μεθόδους. Ένα παράδειγμα είναι η αντιστοίχιση διαφορετικών εικόνων του εγκεφάλου από διαφορετικές γωνίες και χρονικές στιγμές.
7. **Ερμηνεία Εικόνας (Image Interpretation and Computer Reasoning):** Είναι η χρήση τεχνικών για την ανάλυση και την εξαγωγή πληροφοριών από τις εικόνες. Αυτό μπορεί να περιλαμβάνει την ταξινόμηση των εικόνων σε διαφορετικές κατηγορίες, βάση χαρακτηριστικών όπως το περιεχόμενο ή άλλες σημαντικές παραμέτρους. Η διαδικασία της ανάλυσης των εικόνων συνήθως γίνεται αυτοματοποιημένα με τη χρήση υπολογιστών και ειδικών λογισμικών που εφαρμόζουν αλγόριθμους μηχανικής μάθησης για την εξαγωγή συμπερασμάτων από αυτές.

ΚΕΦΑΛΑΙΟ 3^ο: Νευρωνικά δίκτυα βαθιάς μάθησης

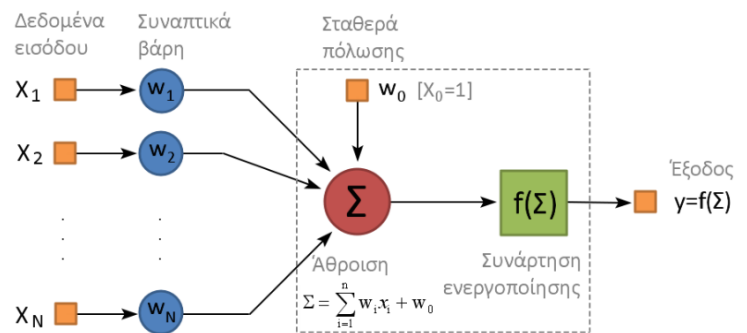
3.1 Τεχνητά Νευρωνικά Δίκτυα (Artificial Neural Networks – ANNs)

Εμπνευσμένα από τον τρόπο λειτουργίας του ανθρώπινου εγκεφάλου, τα *Τεχνητά Νευρωνικά Δίκτυα* (Artificial Neural Networks – ANNs) είναι ένα από τα πιο ισχυρά και διαδεδομένα εργαλεία στον τομέα της μηχανικής μάθησης. Η μηχανική μάθηση είναι ένας κλάδος της τεχνητής νοημοσύνης που ασχολείται με την ανάπτυξη αλγορίθμων και μοντέλων που επιτρέπουν στους υπολογιστές να μαθαίνουν από ένα σύνολο δεδομένων και έπειτα να λαμβάνουν αποφάσεις, να πραγματοποιούν προβλέψεις ή να εξάγουν μοτίβα, χωρίς να είναι ρητά προγραμματισμένα για κάθε συγκεκριμένη εργασία.

Τα ANN χρησιμοποιούνται ευρέως σε εφαρμογές όπως η αναγνώριση εικόνας, η επεξεργασία φυσικής γλώσσας, η πρόβλεψη τάσεων, η ιατρική διάγνωση, επιτρέποντας την αυτοματοποίηση και την ενίσχυση διαφόρων διαδικασιών. Αναπτύχθηκαν ως γενικευμένα μοντέλα που βασίζονται στα μαθηματικά των βιολογικών νευρικών συστημάτων και αποτελούν ισχυρά μαθηματικά εργαλεία που έχουν τη δυνατότητα να «εκπαιδεύονται» με σκοπό να κατανοήσουν τη συμπεριφορά ενός συστήματος, χωρίς να απαιτείται προγενέστερη γνώση του συστήματος, παρά μόνο με βάση τα δεδομένα εισόδου και εξόδου.



Εικόνα 3.1 Βιολογικός Νευρώνας



Εικόνα 3.2 Τεχνητός Νευρώνας

Ένας *βιολογικός νευρώνας* (Εικόνα 3.1) αποτελείται από το κυτταρικό σώμα, τους δενδρίτες και τον νευράξονα. Οι δενδρίτες λαμβάνουν σήματα από άλλους νευρώνες και τα μεταφέρουν στο σώμα του κυττάρου. Το σώμα του κυττάρου επεξεργάζεται τα σήματα και αν η συνολική είσοδος υπερβεί ένα ορισμένο κατώφλι, το σήμα μεταδίδεται μέσω του άξονα σε άλλους νευρώνες. Αυτή η διαδικασία επιτρέπει στον εγκέφαλο να εκτελεί πολύπλοκες λειτουργίες, όπως η αποθήκευση αναμνήσεων και η λήψη αποφάσεων, μέσα από ένα δίκτυο εκατομμυρίων συνδεδεμένων νευρώνων.

Ένας *τεχνητός νευρώνας* (Εικόνα 3.2), γνωστός και ως τεχνητή νευρωνική μονάδα, λειτουργεί με τρόπο παρόμοιο με έναν βιολογικό νευρώνα. Λαμβάνει εισόδους από πολλές άλλες νευρωνικές μονάδες, κάθε μία από τις οποίες έχει έναν συντελεστή βάρους που καθορίζει τη σημασία της συγκεκριμένης εισόδου. Ο τεχνητός νευρώνας υπολογίζει το συνολικό άθροισμα των βαρών των εισόδων και αν αυτό το άθροισμα υπερβεί ένα καθορισμένο κατώφλι, ενεργοποιείται και μεταδίδει το σήμα στην έξοδο. Με αυτόν τον τρόπο, τα τεχνητά νευρωνικά δίκτυα μπορούν να μαθαίνουν και να προσαρμόζονται, επιλύοντας προβλήματα

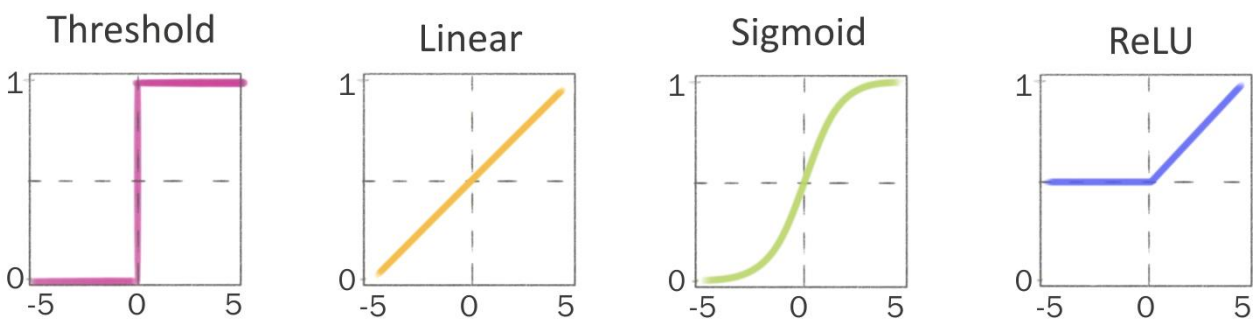
μέσω της προσαρμογής των βαρών των συνδέσεων μεταξύ των νευρώνων, παρόμοια με τον τρόπο που ο εγκέφαλος μαθαίνει και προσαρμόζεται μέσω της πολυπλοκότητας των συνάψεων.

Ο τεχνητός νευρώνας αποτελείται από τρία βασικά στοιχεία:

- **Συνάψεις:** Κάθε σύναψη έχει ένα βάρος που πολλαπλασιάζεται με το εισερχόμενο σήμα, ενισχύοντάς το ή αποδυναμώνοντάς το.
- **Αθροιστής:** Συνδυάζει τα εισερχόμενα σήματα με γραμμικό τρόπο.
- **Συνάρτηση ενεργοποίησης:** Περιορίζει το πλάτος του σήματος εξόδου σε συγκεκριμένες τιμές, γι' αυτό λέγεται και συνάρτηση σύνθλιψης.

3.1.1 Συνάρτηση Ενεργοποίησης Νευρώνα

Οι πιο συνηθισμένες συναρτήσεις ενεργοποίησης είναι οι εξής:



Εικόνα 3.3 Συναρτήσεις ενεργοποίησης νευρώνων: (α) threshold (β) γραμμική (γ) σιγμοειδής (δ) ReLU

Συνάρτηση κατωφλίου (threshold)

Η συνάρτηση κατωφλίου (Εικόνα 3.3α) είναι μια απλή συνάρτηση που ενεργοποιεί το νευρώνα μόνο εάν η είσοδος υπερβαίνει ένα συγκεκριμένο κατώφλι, συνήθως 0.

$$f(x) = \begin{cases} 0 & \text{για } x < 0 \\ 1 & \text{για } x \geq 0 \end{cases}$$

Γραμμική συνάρτηση

Η γραμμική συνάρτηση ενεργοποίησης (Εικόνα 3.3β) είναι μια συνάρτηση που πολλαπλασιάζει το σήμα εισόδου με μια σταθερά K και λειτουργεί ως γραμμικός ενισχυτής.

$$f(x) = K \cdot x$$

Σιγμοειδής συνάρτηση

Η σιγμοειδής συνάρτηση (Εικόνα 3.3γ) είναι μια ομαλή και συνεχής συνάρτηση η οποία συμπιέζει την είσοδο σε ένα εύρος μεταξύ 0 και 1, προσομοιώνοντας την κατανομή πιθανοτήτων.

$$f(x) = \frac{1}{1 + e^{-ax}}$$

Η καμπυλότητα της σιγμοειδούς συνάρτησης αλλάζει καθώς το a μεγαλώνει ή μικραίνει.

Όταν το x είναι πολύ μεγάλο (δηλαδή $x \rightarrow \infty$), η τιμή της σιγμοειδούς προσεγγίζει το 1.

Όταν το x είναι πολύ μικρό (δηλαδή $x \rightarrow -\infty$), η τιμή της προσεγγίζει το 0.

Όσο μεγαλύτερο είναι το a , τόσο πιο απότομη γίνεται η καμπύλη, προσεγγίζοντας τη συμπεριφορά της συνάρτησης κατωφλίου.

Όσο μικρότερο είναι το a , τόσο πιο ομαλή γίνεται η καμπύλη, προσεγγίζοντας τη συμπεριφορά της γραμμικής συνάρτησης.

Συνάρτηση Rectified Linear Unit (ReLU)

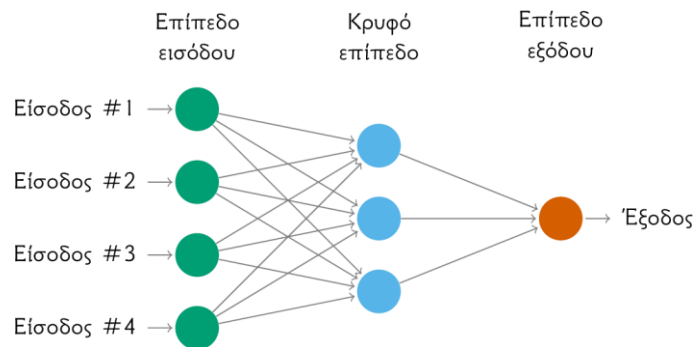
Η συνάρτηση ReLU (Εικόνα 3.3δ) είναι γραμμική κατά τμήματα, αλλά μη γραμμική στο πεδίο ορισμού της. Είναι αρκετά δημοφιλής συνάρτηση σε βαθιά νευρωνικά δίκτυα λόγω της ευκολίας της στον υπολογισμό και της αποτελεσματικότητάς της στη βελτιστοποίηση των αλγορίθμων μηχανικής μάθησης.

$$f(x) = \begin{cases} 0 & \text{για } x < 0 \\ K \cdot x & \text{για } x \geq 0 \end{cases}$$

3.1.2 Δομή Νευρωνικών Δικτύων

Για να δημιουργηθεί ένα νευρωνικό δίκτυο με δυνατότητα μάθησης και αποθήκευσης γνώσης, απαιτείται η σύνδεση πολλών νευρώνων. Οι τρεις παράγοντες που καθορίζουν την δομή του δικτύου είναι:

1. το είδος των νευρώνων,
2. η τοπολογία τους
3. και οι συνδέσεις μεταξύ τους.

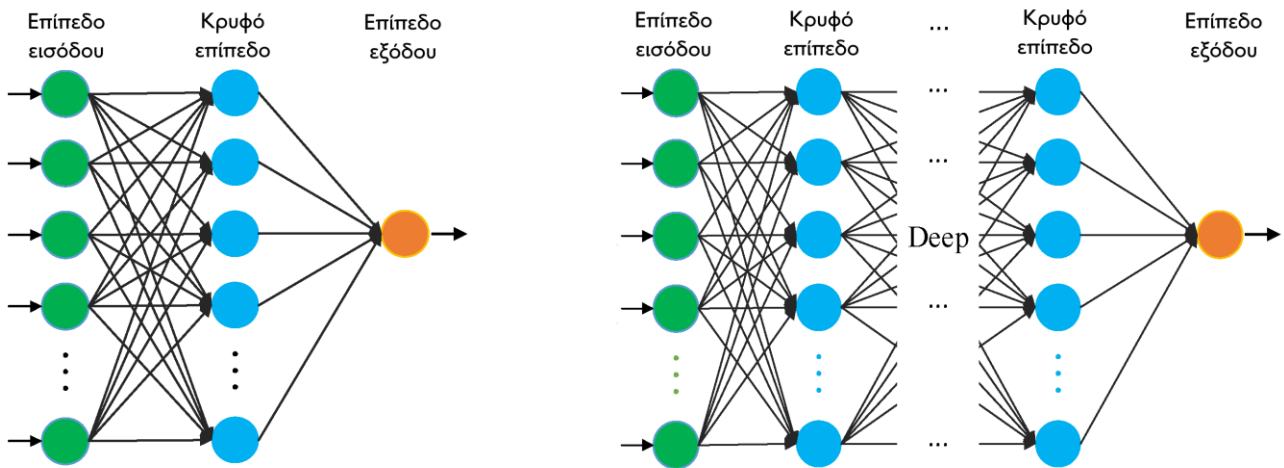


Εικόνα 3.4 Ένα νευρωνικό δίκτυο με τέσσερις εισόδους και ένα κρυφό επίπεδο με τρεις κρυμμένους νευρώνες

Η αρχιτεκτονική των νευρωνικών δικτύων συνδέεται άμεσα και με τον τρόπο που αποθηκεύουν την γνώση από το περιβάλλον και λειτουργούν. Υπάρχουν πολλές αρχιτεκτονικές νευρωνικών δικτύων, καθεμία με πλεονεκτήματα και μειονεκτήματα, που εφαρμόζονται σε διάφορα προβλήματα, από πιο εξειδικευμένα έως και ευρύτερης χρήσης.

3.2 Νευρωνικά Δίκτυα Βαθιάς Μάθησης (Deep Neural Networks – DNNs)

Η βαθιά μάθηση είναι ένα υποσύνολο της μηχανικής μάθησης, που βασίζεται στη χρήση πολυεπίπεδων νευρωνικών δικτύων, γνωστά ως *Νευρωνικά Δίκτυα Βαθιάς Μάθησης (Deep Neural Networks – DNNs)*. Αυτά τα δίκτυα σχεδιάζονται με σκοπό να προσομοιάσουν την πολύπλοκη σκέψη και λήψη αποφάσεων που εκτελεί ο ανθρώπινος εγκέφαλος, όπως να αναγνωρίζουν σύνθετα μοτίβα σε δεδομένα και να εξάγουν χρήσιμες πληροφορίες, χρησιμοποιώντας σύνθετα μαθηματικά μοντέλα. Η βαθιά μάθηση είναι υπεύθυνη για το μεγαλύτερο μέρος της τεχνητής νοημοσύνης που συναντάται σήμερα, καθώς εξελίσσεται σε ένα ευρύ φάσμα εφαρμογών, από την αναγνώριση προτύπων έως την αυτόνομη οδήγηση.



Εικόνα 3.5 Σύγκριση ενός (α) ANN και (β) Deep Neural Network

Στην πιο απλή μορφή, ως DNN μπορεί να θεωρηθεί ένα νευρωνικό δίκτυο με τουλάχιστον δύο κρυφά επίπεδα νευρώνων (Εικόνα 3.5). Πρακτικά, τα περισσότερα βαθιά νευρωνικά δίκτυα έχουν πολύ περισσότερα επίπεδα. Αυτά τα δίκτυα εκπαιδεύονται σε μεγάλο όγκο δεδομένων για να αναγνωρίζουν και να ταξινομούν φαινόμενα, να αναγνωρίζουν πρότυπα και σχέσεις, να αξιολογούν πιθανότητες και να κάνουν προβλέψεις και αποφάσεις. Η χρήση πολλαπλών επιπέδων σε ένα τέτοιο νευρωνικό δίκτυο βοηθάει στη βελτίωση και την ακρίβεια των αποτελεσμάτων.

3.2.1 Κατηγορίες Deep Neural Networks

Τα Νευρωνικά Δίκτυα Βαθιάς Μάθησης (Deep Neural Networks – DNNs) έχουν αναπτυχθεί σε μια πληθώρα δομών και αρχιτεκτονικών, προσαρμοσμένων για διάφορες εφαρμογές και τύπους δεδομένων. Κάθε ένα από αυτά τα δίκτυα είναι ειδικά σχεδιασμένο για να αξιοποιεί διαφορετικά χαρακτηριστικά των δεδομένων και να τα επεξεργάζεται με το βέλτιστο τρόπο για την επίτευξη των καλύτερων αποτελεσμάτων.

Συνήθως, κάθε είδος DNN εξειδικεύεται σε διαφορετικού είδους προβλήματα, προσφέροντας δυνατότητες και πλεονεκτήματα ανάλογα με την εφαρμογή. Με την κατάλληλη προσαρμογή της δομής τους, τα DNNs μπορούν να επιτύχουν υψηλή απόδοση και ακρίβεια σε ένα ευρύ φάσμα εφαρμογών.

Μερικές από τις πιο ευρέως γνωστές και χρησιμοποιούμενες κατηγορίες των Deep Neural Networks (DNNs) είναι:

1. Feedforward Neural Networks (FNNs):

- Είναι η πιο απλή μορφή DNN, όπου τα δεδομένα κινούνται μόνο προς τα εμπρός από την είσοδο προς την έξοδο.
- Κάθε νευρώνας σε ένα επίπεδο συνδέεται με τους νευρώνες του επόμενου επιπέδου.
- Χρησιμοποιούνται κυρίως για απλές ταξινομήσεις και προβλέψεις.

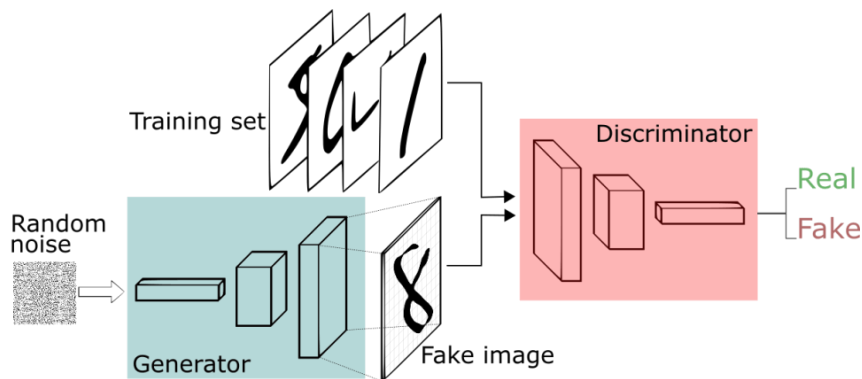
2. Recurrent Neural Networks (RNNs):

- Είναι τα δίκτυα που περιέχουν τουλάχιστον έναν βρόγχο ανατροφοδότησης.
- Η έξοδος τουλάχιστον ενός από τους νευρώνες, ανατροφοδοτείται είτε στον ίδιο νευρώνα είτε σε άλλους ως είσοδος.

- Κατάλληλα για μοντελοποίηση δυναμικών συστημάτων και χρονοσειρών.

3. Convolutional Neural Networks (CNNs):

- Τα δίκτυα αυτά δέχονται είτε εικόνες είτε διδιάστατους πίνακες στην είσοδο.
- Χρησιμοποιούν φίλτρα συνέλιξης για την ανίχνευση χαρακτηριστικών σε διαφορετικές περιοχές της εικόνας.
- Είναι εξειδικευμένα στην επεξεργασία εικόνων και την αναγνώριση προτύπων σε αυτές και χρησιμοποιούνται για την κατηγοριοποίησή τους, όπως επίσης και σε εφαρμογές μηχανικής όρασης (computer vision).



Εικόνα 3.6 Παράδειγμα δομής ενός GAN

4. Generative Adversarial Networks (GANs):

- Αποτελούνται από δύο δίκτυα, ένα δίκτυο – γεννήτρια (generator), που προσπαθεί να δημιουργήσει νέα δεδομένα, και ένα δίκτυο – ανιχνευτή (discriminator), που προσπαθεί να διακρίνει μεταξύ των πραγματικών και των παραγόμενων δεδομένων.
- Τα δύο δίκτυα εκπαιδεύονται ανταγωνιστικά. Καθώς το δίκτυο – γεννήτρια μαθαίνει να παράγει νέα δεδομένα, το δίκτυο – ανιχνευτής εκπαιδεύεται για να γίνεται όλο και πιο αποτελεσματικό στη διάκριση των δημιουργιών της γεννήτριας.
- Χρησιμοποιούνται ευρέως για τη δημιουργία νέων δεδομένων που προσομοιώνουν τα πραγματικά δεδομένα, όπως ρεαλιστικές εικόνες, ήχοι, βίντεο, μουσική και άλλες μορφές πολυμέσων.

3.2.2 Αξιολόγηση της Απόδοσης μέσω Confusion Matrix

Στην περίπτωση νευρωνικών δικτύων που εξειδικεύονται στην αναγνώριση προτύπων, όπως τα CNNs, είναι σημαντικό να υπάρχει ένας δείκτης ακρίβειας ώστε να μπορεί να αξιολογηθεί η απόδοση του νευρωνικού δικτύου. Ένα πολύ χρήσιμο εργαλείο στην αξιολόγηση αυτή είναι ο πίνακας σύγχυσης (confusion matrix).

Ο πίνακας σύγχυσης έχει τετραγωνική δομή και οι διαστάσεις του είναι $N \times N$, όπου N είναι ο αριθμός των κατηγοριών. Κάθε γραμμή του πίνακα αντιπροσωπεύει τις προβλέψεις του μοντέλου, ενώ κάθε στήλη τις πραγματικές κατηγορίες στις οποίες ανήκουν τα δεδομένα. Κάθε κελί του πίνακα δείχνει τις σωστές ή τις λάθος προβλέψεις.

Έστω ένας πίνακας σύγκρισης 2×2 , δηλαδή $N = 2$ κλάσεις δεδομένων. Για λόγους ευκολίας και κατανόησης, οι δύο κλάσεις του παραδείγματος είναι “0” και “1”. Στον Πίνακα 3.1 φαίνεται ένα παράδειγμα 10 περιπτώσεων προβλέψεων συγκριτικά με τις πραγματικές τους τιμές.

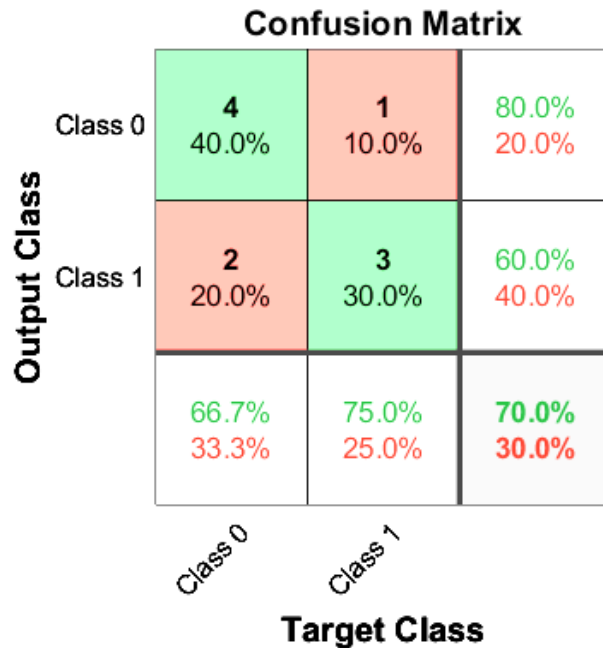
Πίνακας 3.1 Πραγματικές Τιμές και Προβλέψεις ενός δικτύου

Αριθμός Δεδομένου	Πραγματικές Τιμές	Προβλέψεις
1	0	0
2	1	1
3	1	0
4	1	1
5	0	1
6	0	0
7	1	1
8	0	0
9	0	1
10	0	0

Από τα δεδομένα του Πίνακα 3.1 προκύπτει ότι:

- Στην **Κλάση 0** ανήκουν 6/10 περιπτώσεις, ενώ προβλέφθηκαν σωστά 4/10 = **40%**.
 - Οι συνολικές σωστές προβλέψεις (4) ως προς τον αριθμό των πραγματικών δεδομένων (6) είναι: $4/6 = 66,7\%$
 - Οι συνολικές λάθος προβλέψεις (2) ως προς τον αριθμό των πραγματικών δεδομένων (6) είναι: $2/6 = 33,3\%$
 - Οι συνολικές σωστές προβλέψεις (4) ως προς τον συνολικό αριθμό προβλέψεων (5) που έγιναν για αυτήν την κλάση είναι: $4/5 = 80\%$
 - Οι συνολικές λάθος προβλέψεις (1) ως προς τον συνολικό αριθμό προβλέψεων (5) που έγιναν για αυτήν την κλάση είναι: $1/5 = 20\%$
- Στην **Κλάση 1** ανήκουν 4/10 περιπτώσεις, ενώ προβλέφθηκαν σωστά 3/10 = **30%**.
 - Οι συνολικές σωστές προβλέψεις (3) ως προς τον αριθμό των πραγματικών δεδομένων (4) είναι: $3/4 = 75\%$
 - Οι συνολικές λάθος προβλέψεις (1) ως προς τον αριθμό των πραγματικών δεδομένων (4) είναι: $1/4 = 25\%$
 - Οι συνολικές σωστές προβλέψεις (3) ως προς τον συνολικό αριθμό προβλέψεων (5) που έγιναν για αυτήν την κλάση είναι: $3/5 = 60\%$
 - Οι συνολικές λάθος προβλέψεις (2) ως προς τον συνολικό αριθμό προβλέψεων (5) που έγιναν για αυτήν την κλάση είναι: $2/5 = 40\%$

Εν τέλει, το μοντέλο έχει συνολικό ποσοστό επιτυχίας $40\% + 30\% = 70\%$. Ο πίνακας σύγχυσης του παραδείγματος αυτού κατασκευάζεται σύμφωνα με αυτούς τους υπολογισμούς (Εικόνα 3.7).



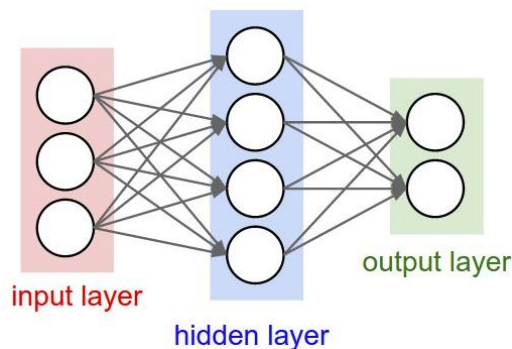
Εικόνα 3.7 Confusion Matrix σύμφωνα με τα δεδομένα του Πίνακα 3.1

3.3 Νευρωνικά Δίκτυα Πρόσθιας Τροφοδότησης (Feedforward Neural Networks – FNNs)

Τα *Νευρωνικά Δίκτυα Πρόσθιας Τροφοδότησης (Feedforward Neural Networks – FNNs)* αποτελούν μια από τις πιο βασικές αρχιτεκτονικές στον τομέα της τεχνητής νοημοσύνης και της μηχανικής μάθησης. Χαρακτηρίζονται από την κατεύθυνση ροής των δεδομένων, η οποία είναι πάντα προς τα εμπρός, από την είσοδο προς την έξοδο, χωρίς ανατροφοδοτήσεις. Τα δίκτυα αυτά χρησιμοποιούνται ευρέως λόγω της απλότητάς τους σε προβλήματα προσέγγισης συναρτήσεων (μοντελοποίηση συστημάτων) όπως επίσης και σε προβλήματα ταξινόμησης προτύπων. Με την προσθήκη περισσότερων κρυφών επιπέδων νευρώνων, αυξάνεται η πολυπλοκότητά τους και συνεπώς η ικανότητά τους να προσεγγίσουν πολύπλοκες συναρτήσεις και να πραγματοποιήσουν πιο ακριβείς προβλέψεις.

3.3.1 Δίκτυα FNN ενός κρυφού επιπέδου

Το πιο απλό δίκτυο πρόσθιας τροφοδότησης (FNN) είναι το δίκτυο που περιέχει μόνο ένα κρυφό επίπεδο νευρώνων (Εικόνα 3.8).



Εικόνα 3.8 FNN ενός κρυφού επιπέδου νευρώνων

Το επίπεδο εισόδου δεν θεωρείται επίπεδο νευρώνων, διότι δεν εκτελεί κανέναν υπολογισμό. Η λειτουργία του είναι να μεταφέρει τις μεταβλητές εισόδου προς τους νευρώνες του επόμενου επιπέδου.

Σε αυτήν την βασική δομή, οι εισοδοί του δικτύου μεταφέρονται στο μοναδικό κρυφό επίπεδο, το οποίο εκτελεί τους υπολογισμούς και μαθαίνει τις απαραίτητες σχέσεις και μοτίβα από τα δεδομένα.

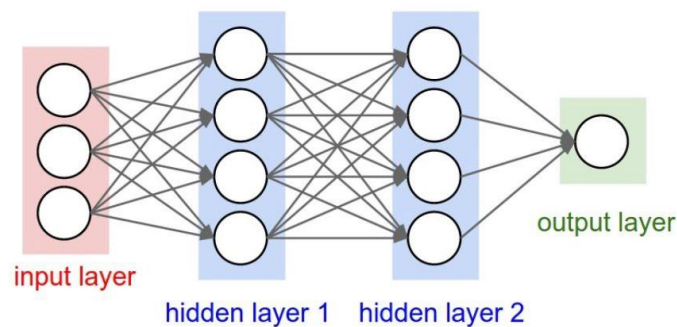
Το δίκτυο με ένα μόνο κρυφό επίπεδο νευρώνων είναι συχνά αρκετό για να λύσει απλά προβλήματα, ενώ παράλληλα αποτελεί τη βάση για την κατανόηση πιο σύνθετων αρχιτεκτονικών. Ωστόσο, καθώς τα περισσότερα προβλήματα είναι συνήθως πιο σύνθετα, προκειμένου το δίκτυο να αποκτήσει την ικανότητα να μαθαίνει και να προσαρμόζεται σε πιο πολύπλοκα μοτίβα και δεδομένα είναι σχεδόν αναπόφευκτη η χρήση περισσότερων κρυφών επιπέδων.

3.3.2 Δίκτυα FNN πολλαπλών κρυφών επιπέδων

Τα δίκτυα πρόσθιας τροφοδότησης πολλών επιπέδων αποτελούνται από πολλά κρυφά επίπεδα νευρώνων, ανάμεσα στο επίπεδο της εισόδου και στο επίπεδο της εξόδου. Τα κρυφά επίπεδα δεν είναι απευθείας ορατά από την είσοδο ή την έξοδο του δικτύου και είναι υπεύθυνα για την επεξεργασία των δεδομένων.

Σε *πλήρως συνδεδεμένα δίκτυα*, κάθε νευρώνας ενός επιπέδου συνδέεται με κάθε νευρώνα του επόμενου επιπέδου (Εικόνα 3.9). Αυτή η πολυπλοκότητα αυξάνει την ικανότητα του δικτύου να μαθαίνει και να προσαρμόζεται σε σύνθετα πρότυπα και δεδομένα.

Σε *μερικώς συνδεδεμένα δίκτυα*, οι νευρώνες ενός επιπέδου δεν συνδέονται με όλους τους νευρώνες του επόμενου επιπέδου, αλλά μόνο με μερικούς από αυτούς. Αυτή η προσέγγιση μειώνει την υπολογιστική ισχύ που απαιτείται όπως επίσης και τον κίνδυνο υπερπροσαρμογής (overfitting), επιτρέποντας έτσι στο δίκτυο να επικεντρώνεται στα πιο σημαντικά χαρακτηριστικά, και βελτιώνοντας τη γενίκευση και την απόδοση σε πραγματικά δεδομένα.



Εικόνα 3.9 FNN με πολλαπλά (πλήρως συνδεδεμένα) κρυφά επίπεδα νευρώνων

Με την προσθήκη περισσότερων κρυφών επιπέδων, το δίκτυο μπορεί να αναγνωρίζει και να μαθαίνει πιο πολύπλοκες σχέσεις και πρότυπα στα δεδομένα, γεγονός που το καθιστά ιδιαίτερα αποτελεσματικό σε εφαρμογές όπως η προσέγγιση συναρτήσεων και η ταξινόμηση δεδομένων.

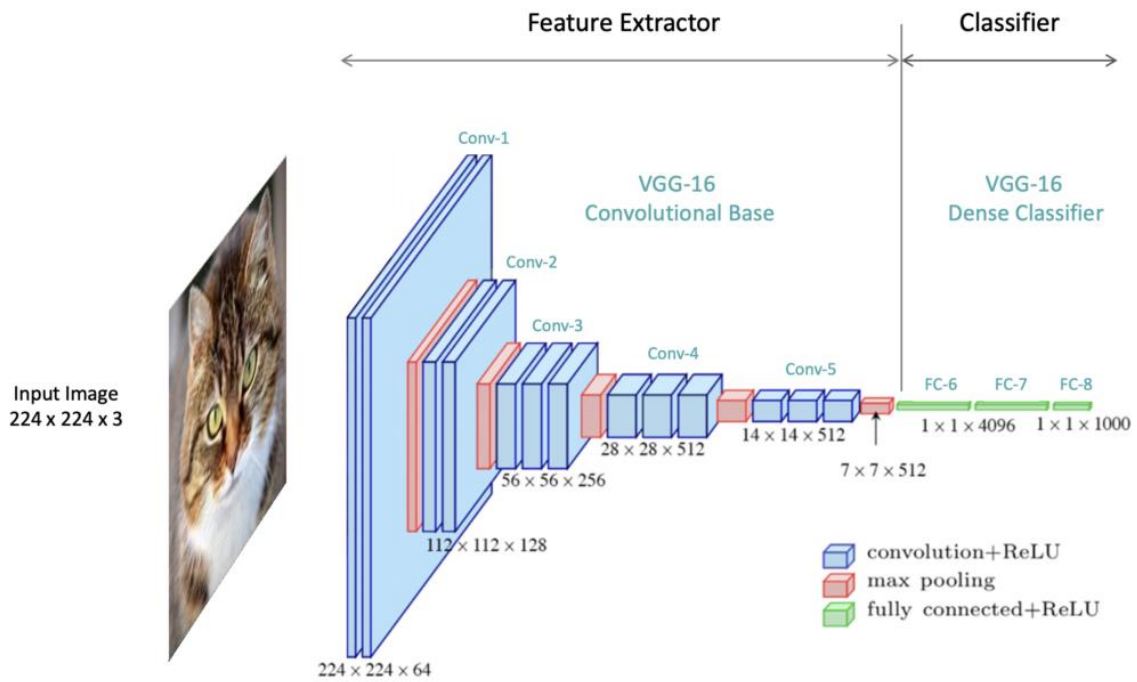
3.4 Συνελκτικά Νευρωνικά Δίκτυα (Convolutional Neural Networks – CNNs)

Τα *Συνελκτικά Νευρωνικά Δίκτυα* (Convolutional Neural Networks – CNNs) αποτελούν μια ειδική κατηγορία των νευρωνικών δικτύων βαθιάς μάθησης που έχει σχεδιαστεί για να αντιμετωπίζει αποκλειστικά προβλήματα επεξεργασίας εικόνας και αναγνώρισης προτύπων. Η ανάπτυξη των CNNs

βασίζεται στην δομή του ανθρώπινου οπτικού συστήματος, επιτρέποντας την αυτόματη εξαγωγή και αναγνώριση συνθέτων χαρακτηριστικών από εικόνες.

Αποτελούνται από τρία κύρια είδη επιπέδων:

- 1. Συνελικτικά Επίπεδα:** Τα συνελικτικά επίπεδα εφαρμόζουν συνελίξεις στην εικόνα που λαμβάνουν στην είσοδό τους, χρησιμοποιώντας φίλτρα τα οποία ανιχνεύουν χαρακτηριστικά όπως ακμές και χρώματα σε μικρές περιοχές της εικόνας. Το πρώτο επίπεδο ενός CNN είναι συνελικτικό επίπεδο.
- 2. Επίπεδα Συγκέντρωσης (Pooling):** Τα επίπεδα συγκέντρωσης μειώνουν την χωρική διάσταση των χαρακτηριστικών χωρίς να αλλοιώνουν σημαντικά την πληροφορία. Αυτό το επιτυγχάνουν είτε επιλέγοντας το μέγιστο στοιχείο (max pooling) είτε το μέσο στοιχείο (average pooling) σε μικρές περιοχές των χαρακτηριστικών που έχουν δημιουργηθεί από το συνελικτικό επίπεδο που προηγήθηκε.
- 3. Πλήρως Συνδεδεμένα Επίπεδα:** Τα πλήρως συνδεδεμένα επίπεδα είναι τα τελευταία επίπεδα ενός CNN. Χρησιμοποιούνται για την τελική ταξινόμηση ή πρόβλεψη του δικτύου.



Εικόνα 3.10 Αρχιτεκτονική CNN του νευρωνικού δικτύου VGG-16

Με κάθε προσθήκη τέτοιων επιπέδων, το νευρωνικό δίκτυο αυξάνει την πολυπλοκότητά του και μαθαίνει να αναγνωρίζει μεγαλύτερα τμήματα της αρχικής εικόνας. Τα πρώτα επίπεδα ενός CNN επικεντρώνονται κυρίως σε βασικά χαρακτηριστικά, όπως περιγράμματα ή χρώματα (Εικόνα 3.11 και Εικόνα 3.12) ενώ τα πιο βαθιά επίπεδα μπορούν να αναγνωρίζουν πιο περίπλοκα στοιχεία όπως σχήματα ή σημαντικά χαρακτηριστικά (Εικόνα 3.13).

Έτσι το δίκτυο μαθαίνει σταδιακά να εστιάζει σε πιο σύνθετη και σημαντική πληροφορία πάνω στην εικόνα, αναλογικά με τον αριθμό και την δομή των επιπέδων του, και να κατανοεί στοιχεία όπως συγκεκριμένα αντικείμενα ή περίπλοκα μοτίβα. Καθώς το κάθε νέο επίπεδο προσθέτει μεγαλύτερη εξειδίκευση στην ανάλυση των δεδομένων εισόδου, το CNN μπορεί τελικά να αποκτήσει τη δυνατότητα να αντιληφθεί και να διακρίνει ακόμα και τις πιο λεπτομερείς δομές και σχέσεις στις εικόνες.



Εικόνα 3.11 Οπτικοποίηση των σημείων ενεργοποίησης ενός καναλιού από ένα πρώιμο συνελκτικό επίπεδο ενός CNN στο MATLAB. Το συγκεκριμένο κανάλι ενεργοποιείται στα περιγράμματα.



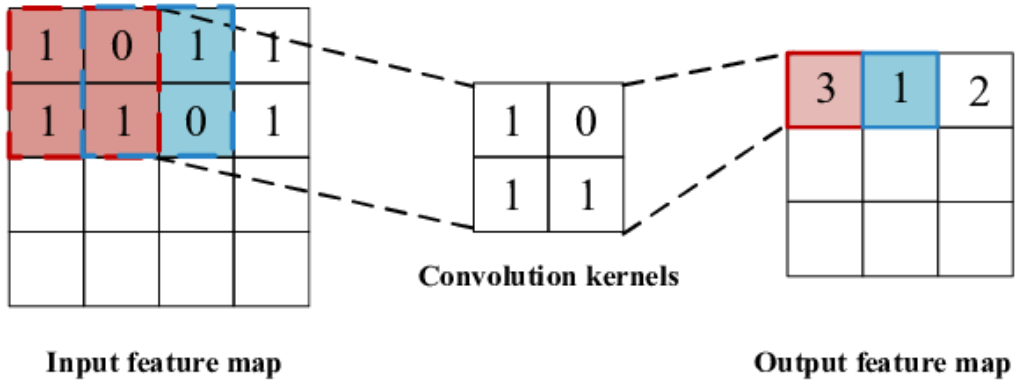
Εικόνα 3.12 Οπτικοποίηση των σημείων ενεργοποίησης ενός καναλιού από ένα πρώιμο συνελκτικό επίπεδο ενός CNN στο MATLAB. Το συγκεκριμένο κανάλι ενεργοποιείται στα ερυθρά pixels.



Εικόνα 3.13 Οπτικοποίηση των σημείων ενεργοποίησης ενός καναλιού από ένα πιο βαθύ επίπεδο ενός CNN στο MATLAB.. Το κανάλι ενεργοποιείται σε πιο περίπλοκα χαρακτηριστικά, όπως τα μάτια.

3.4.1 Συνέλιξη

Στο πλαίσιο της επεξεργασίας εικόνας και των συνελκτικών νευρωνικών δικτύων (CNN), η συνέλιξη είναι μια μαθηματική πράξη που χρησιμοποιείται για να συνδυάσει δύο σύνολα πληροφοριών. Περιλαμβάνει την ολίσθηση ενός φίλτρου, γνωστού και ως πυρήνα, πάνω από την εικόνα εισόδου, την εκτέλεση πολλαπλασιασμών ανά στοιχείο (element-wise) και τέλος το άθροισμα των αποτελεσμάτων αυτών για την παραγωγή ενός χάρτη χαρακτηριστικών.

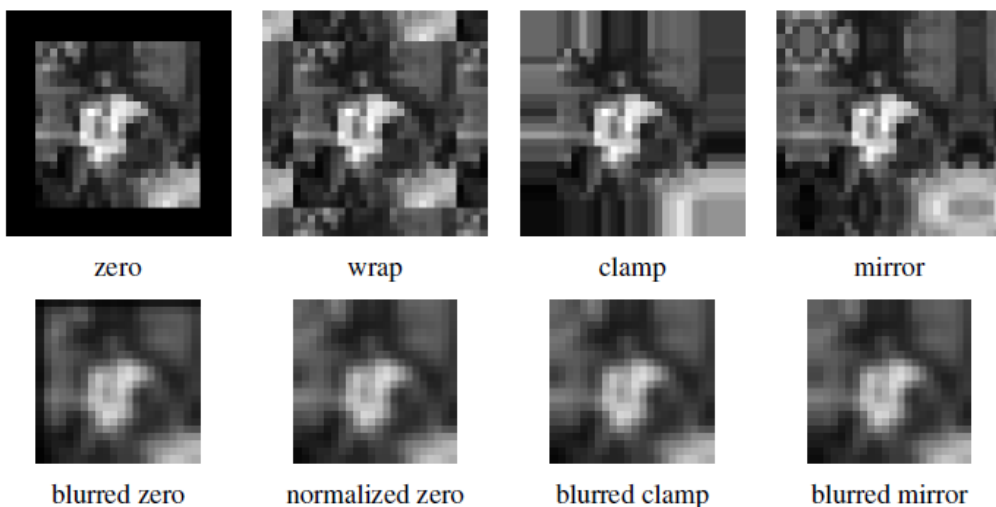


Εικόνα 3.14 Ενδεικτικό διάγραμμα της διαδικασίας της συνέλιξης

Τα βήματα που ακολουθούνται στη συνέλιξη είναι τα εξής:

1. **Φίλτρο (Πυρήνας):** Επιλογή μιας μικρής μήτρας – φίλτρου (συνήθως 3×3, 5×5 ή 7×7), που θα ολισθαίνει πάνω από την εικόνα εισόδου.
2. **Ολίσθηση φίλτρου:** Μετακίνηση του φίλτρου πάνω στην εικόνα, ένα βήμα τη φορά (το μέγεθος του βήματος ονομάζεται ‘stride’).
3. **Πολλαπλασιασμός βάσει στοιχείων:** Σε κάθε θέση που τοποθετείται το φίλτρο, γίνεται πολλαπλασιασμός των στοιχείων της μήτρας με τα αντίστοιχα στοιχεία της εικόνας.
4. **Άθροισμα γινομένων:** Το άθροισμα των γινομένων αυτών των πολλαπλασιασμών δίνει μια νέα τιμή η οποία τοποθετείται στην αντίστοιχη θέση στον χάρτη χαρακτηριστικών εξόδου.
5. **Επανάληψη:** Η διαδικασία επαναλαμβάνεται για όλες τις θέσεις του φίλτρου στην εικόνα εισόδου.

3.4.2 Παράμετροι ενός CNN



Εικόνα 3.15 Παραδείγματα διαφορετικών ειδών Padding (πάνω σειρά) και πώς αυτά φαίνονται σε θολή εικόνα (κάτω σειρά)

Μερικές από τις παραμέτρους που μπορούν να ρυθμιστούν στα συνελκτικά επίπεδα των CNN περιλαμβάνουν:

- **Αριθμός φίλτρων:** Αυτός ο αριθμός καθορίζει πόσα φίλτρα θα εφαρμοστούν σε κάθε στάδιο συνέλιξης και κατά συνέπεια πόσα διαφορετικά χαρακτηριστικά θα μπορέσει να ανιχνεύσει το CNN σε κάθε επίπεδο συνέλιξης. Αυξάνοντας τον αριθμό των φίλτρων, το δίκτυο μπορεί να μάθει περισσότερα και πιο σύνθετα χαρακτηριστικά, αλλά αυτό αυξάνει και την υπολογιστική δύναμη που απαιτείται.
- **Μέγεθος φίλτρου:** Καθορίζει τις διαστάσεις του φίλτρου (π.χ. 3×3) που χρησιμοποιείται για τις συνελίξεις. Μικρότερα φίλτρα μπορούν να ανιχνεύσουν μικρότερες λεπτομέρειες, ενώ μεγαλύτερα φίλτρα είναι καλύτερα για την αναγνώριση μεγαλύτερων χαρακτηριστικών.
- **Πλάτος παρεμβολής (Padding):** Αυτό καθορίζει το μέγεθος των «περιθωρίων» που προστίθενται γύρω από την εικόνα πριν από την εκτέλεση των συνελίξεων, όπως π.χ. συμπλήρωση με μηδενικά ή αναπαραγωγή των pixel που βρίσκονται στις άκρες (Εικόνα 3.15), ώστε να μην συρρικνώνεται η εικόνα μετά από κάθε συνέλιξη.
- **Βήμα (Stride):** Το προεπιλεγμένο βήμα για τη συνέλιξη είναι 1 pixel, αλλά μπορεί επίσης να εκτελεστεί η συνέλιξη σε κάθε n στήλη και γραμμή. Για παράδειγμα, το πρώτο επίπεδο συνέλιξης στο AlexNet χρησιμοποιεί ένα βήμα 4. Ένα μικρό βήμα αυξάνει την ακρίβεια, αλλά αυξάνει και τον χρόνο εκτέλεσης και ολοκλήρωσης των συνελίξεων, ενώ ένα μεγαλύτερο βήμα μπορεί να μειώσει την ακρίβεια, αλλά να μειώσει και τον υπολογιστικό φόρτο.

Η επιλογή αυτών των παραμέτρων εξαρτάται σε μεγάλο βαθμό από την εφαρμογή ή το πρόβλημα που το CNN καλείται να λύσει. Για παράδειγμα, η αναγνώριση αντικειμένων σε εικόνες θα απαιτούσε μεγαλύτερα μεγέθη φίλτρων ώστε να αναγνωρίζονται μεγαλύτερα τμήματα της εικόνας, ενώ η αναγνώριση προσώπων θα χρησιμοποιούσε μικρότερα φίλτρα και μικρότερα βήματα, καθώς τα χαρακτηριστικά του προσώπου που πρέπει να αναγνωριστούν είναι συνήθως πιο μικρά και λεπτομερή.

Οι περιορισμοί σε υπολογιστική ισχύ, μνήμη και χρόνο επηρεάζουν επίσης την επιλογή των παραμέτρων. Μπορεί να είναι αναγκαίο να περιοριστεί ο αριθμός των φίλτρων ή το μέγεθος τους για να εξασφαλιστεί η εκτέλεση του μοντέλου σε ρεαλιστικό χρόνο.

Επίσης, η δομή του μοντέλου παίζει εξίσου σημαντικό ρόλο στην επιλογή των παραμέτρων. Ένα πιο επιφανειακό μοντέλο χρειάζεται διαφορετική προσαρμογή στις παραμέτρους για να επιτευχθεί η καλύτερη δυνατή απόδοση, συγκριτικά με ένα πιο βαθύ μοντέλο.

ΚΕΦΑΛΑΙΟ 4^ο: Πειραματικό μέρος - σχεδίαση έρευνας και data set

Στο κεφάλαιο αυτό, θα παρουσιαστούν λεπτομερώς οι μεθοδολογίες και τα πειράματα που ακολουθήθηκαν για την υλοποίηση της ερευνητικής διαδικασίας. Θα αναλυθούν τα βήματα που αφορούν στην επεξεργασία των δεδομένων στις απεικονιστικές μεθόδους που εφαρμόστηκαν, όπως επίσης και στην κατηγοριοποίηση των εικόνων με χρήση νευρωνικών δικτύων CNN και FNN. Παράλληλα, θα αναλυθούν και τα αποτελέσματα που προέκυψαν, ώστε να προσφέρουν μια ολοκληρωμένη και εμπειριστατωμένη εικόνα του πειράματος.

Πιο συγκεκριμένα, θα εξεταστεί η απόδοση τριών διαφορετικών αρχιτεκτονικών νευρωνικών δικτύων. Αυτές περιλαμβάνουν:

1. ένα απλό **FNN** (Feedforward Neural Network) καθώς και
2. ένα **CNN** (Convolutional Neural Network) – Αρχιτεκτονική #1, και
3. ένα **CNN** (Convolutional Neural Network) – Αρχιτεκτονική #2.

Και στις 2 ανωτέρω περιπτώσεις, θα γίνει δοκιμή στην εκπαίδευση των δικτύων:

1. χρησιμοποιώντας τις αρχικές εικόνες χωρίς καμία επεξεργασία,
2. μετά από επεξεργασία τους με τη χρήση του μετασχηματισμού **FFT-2D** (**Fast Fourier Transform 2D**), εφαρμογή ενός **χαμηλοπερατού γκαουσιανού φίλτρου** στο πεδίο της συχνότητας, και **αντίστροφο FFT-2D**,
3. χρησιμοποιώντας τις **εικόνες-αθροίσματα των εικόνων** των δύο παραπάνω περιπτώσεων.

Μέσα από αυτήν την αναζήτηση, στόχος είναι να βρεθεί η πιο αποτελεσματική προσέγγιση που θα οδηγήσει στην επιτυχή κατηγοριοποίηση των εικόνων υπερήχων μαστού, αποτελώντας έτσι μια σημαντική πρόοδο στην παρούσα έρευνα.

Για λόγους διευκόλυνσης του αναγνώστη, στον παρακάτω πίνακα 3×3 παρουσιάζονται τα πειράματα που θα ακολουθήσουν στα επόμενα κεφάλαια, οργανωμένα σύμφωνα με την ανωτέρω κατηγοριοποίηση. Πρόκειται για 3 διαφορετικά νευρωνικά δίκτυα, το καθένα από τα οποία δοκιμάζεται και αξιολογείται με 3 διαφορετικά επεξεργασμένα σύνολα εικόνων στην κάθε περίπτωση. Θα εξεταστούν 1 FNN και 2 CNN (διαφορετικής αρχιτεκτονικής), ώστε να συγκριθούν τα αποτελέσματα όλων των πειραμάτων και να διαπιστωθεί ποιος συνδυασμός έχει την μεγαλύτερη αξιοπιστία.

Πίνακας 4.1 Παρουσίαση των πειραμάτων που θα πραγματοποιηθούν ανά κεφάλαιο

	Εικόνες χωρίς επεξεργασία	Εικόνες συχνοτικά φιλτραρισμένες (FFT-2D, χαμηλοπερατό γκαουσιανό φίλτρο, IFFT-2D)	Εικόνες-αθροίσματα αρχικών και φιλτραρισμένων εικόνων
FNN	Κεφάλαιο 5.3	Κεφάλαιο 5.4	Κεφάλαιο 5.5
CNN (Αρχιτεκτονική #1)	Κεφάλαιο 6.2.1	Κεφάλαιο 6.2.2	Κεφάλαιο 6.2.3

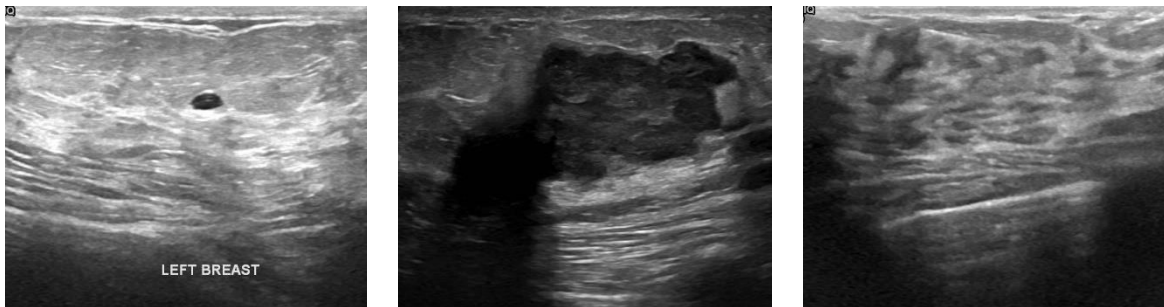
CNN (Αρχιτεκτονική #2)	Κεφάλαιο 6.3.1	Κεφάλαιο 6.3.2	Κεφάλαιο 6.3.3
---------------------------	----------------	----------------	----------------

4.1 Σύνολο Δεδομένων

Το σύνολο δεδομένων (dataset) που χρησιμοποιήθηκε στην παρούσα διπλωματική εργασία αποτελείται από 780 εικόνες υπερήχου που έχουν ήδη κατηγοριοποιηθεί σε τρεις βασικές κατηγορίες. Κάθε εικόνα ανήκει σε μία από τις ακόλουθες κατηγορίες:

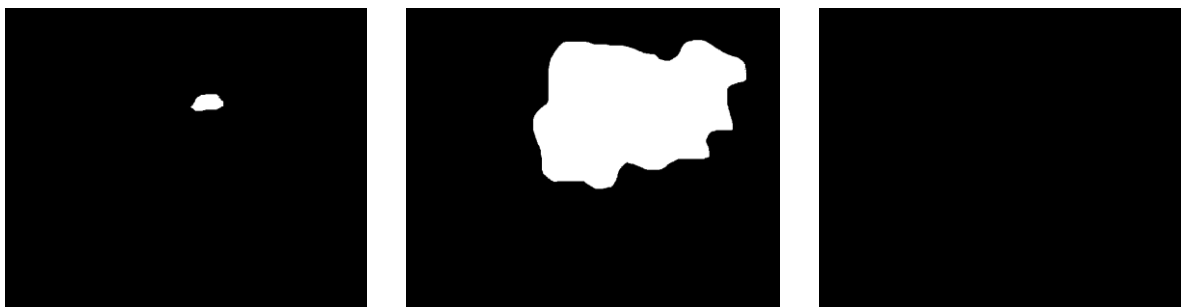
- benign (καλοήθεις)
- malignant (κακοήθεις)
- normal (φυσιολογικές)

Η ταξινόμηση αυτή βασίζεται σε κριτήρια που έχουν οριστεί από ειδικούς στον τομέα της ιατρικής. Κάθε κατηγορία περιέχει εικόνες από υπερήχους μαστών με διαφορετικά χαρακτηριστικά, προσφέροντας έτσι μια ευρεία γκάμα παραδειγμάτων για την αξιολόγηση και τον εκπαιδευτικό σκοπό του μοντέλου κατηγοριοποίησης, ώστε να εκτιμηθεί η απόδοση και η αξιοπιστία του σε διάφορα κλινικά σενάρια.



Εικόνα 4.1 Παραδείγματα εικόνων του dataset
(α) καλοήθεια, (β) κακοήθεια, (γ) φυσιολογικός ιστός

Το dataset περιλαμβάνει συνολικά **437 εικόνες** που αντιστοιχούν σε καλοήθεις περιπτώσεις, **210 εικόνες** που αντιπροσωπεύουν κακοήθειες και **133 εικόνες** φυσιολογικών κυττάρων. Κάθε εικόνα συνοδεύεται από την αντίστοιχη μάσκα, η οποία παρέχει πληροφορίες σχετικά με τις περιοχές του ενδιαφέροντος στην εικόνα. Αυτές οι μάσκες λειτουργούν ως επιπρόσθετες πληροφορίες για την εστίαση σε συγκεκριμένες περιοχές της εικόνας που είναι σημαντικές για την ανίχνευση του καρκίνου και την ορθή ταξινόμηση των εικόνων.



Εικόνα 4.2 Οι μάσκες που αντιστοιχούν στα παραδείγματα της Εικόνας 4.1
(α) καλοήθεια, (β) κακοήθεια, (γ) φυσιολογικός ιστός

Για την αξιολόγηση της απόδοσης του μοντέλου και την αποφυγή overfitting, το dataset θα διαιρεθεί σε τρία υποσύνολα: το **υποσύνολο εκπαίδευσης** (training), το **υποσύνολο επικύρωσης** (validation) και το **υποσύνολο ελέγχου** (testing).

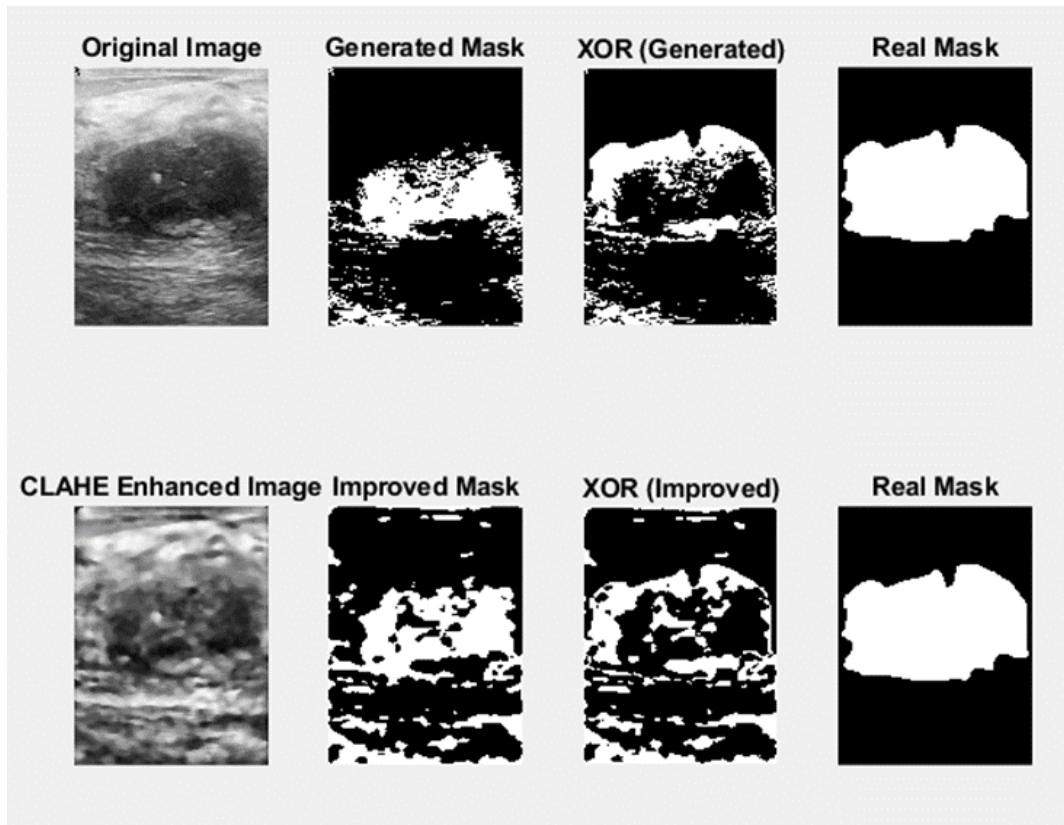
Το υποσύνολο εκπαίδευσης θα χρησιμοποιηθεί για την εκπαίδευση του μοντέλου, ενώ το υποσύνολο επικύρωσης θα χρησιμοποιηθεί για την αξιολόγηση της απόδοσης του μοντέλου κατά τη διάρκεια της εκπαίδευσης και την επιλογή των βέλτιστων υπερπαραμέτρων. Τέλος, το υποσύνολο ελέγχου θα χρησιμοποιηθεί για την αξιολόγηση της τελικής απόδοσης του μοντέλου σε ανεξάρτητα δεδομένα που δεν έχουν χρησιμοποιηθεί κατά την εκπαίδευση. Ακολουθώντας αυτή τη διαδικασία διαχωρισμού του dataset, εξασφαλίζεται η μεγαλύτερη αποτελεσματικότητα της εκπαίδευσης του μοντέλου ώστε να μάθει να γενικεύει σε νέα άγνωστα δεδομένα.

4.2 Δημιουργία και Σύγκριση Μασκών

Αρχικά, αναπτύχθηκε ένας κώδικας στο περιβάλλον MATLAB προκειμένου να δημιουργηθεί μια μάσκα για κάθε μία από τις εικόνες υπερήχου και να γίνει σύγκρισή της με την πραγματική μάσκα. Στόχος ήταν να μπορεί να δημιουργηθεί η αντίστοιχη μάσκα σε κάθε περίπτωση ώστε τα δεδομένα που θα χρησιμοποιηθούν στον έλεγχο (μιας και αυτά θα ήταν νέα δεδομένα για το νευρωνικό δίκτυο) να μπορούν να εισαχθούν αυτόνομα χωρίς να συνοδεύονται από την ήδη υπάρχουσα μάσκα. Θα πρέπει δηλαδή το ίδιο το νευρωνικό δίκτυο να μπορεί να απομονώσει τις περιοχές ενδιαφέροντος, αφού ρεαλιστικά ο σκοπός του πειράματος είναι να μην απαιτείται η σχεδίαση των масκών από κάποιον ειδικό. Για τις δοκιμές που ακολουθούν, τυχαία επιλέχθηκε να εξεταστεί η πρώτη εικόνα από την κατηγορία με τους κακοήθεις υπερήχους.

4.2.1 Μέθοδος κατωφλίου και CLAHE

Εξετάστηκαν δύο διαφορετικοί τρόποι δημιουργίας των масκών: Αρχικά, με τη χρήση ενός κατωφλίου (threshold) χωρίς περαιτέρω επεξεργασία της αρχικής εικόνας, και έπειτα, εφαρμόζοντας ένα φίλτρο median για τη μείωση του θορύβου μαζί με την μέθοδο CLAHE (Contrast Limited Adaptive Histogram Equalization) για τη βελτίωση της αντίθεσης.



Εικόνα 4.3 Αποτελέσματα και σύγκριση εικόνων και масκών με εφαρμογή κατωφλίου και CLAHE

Στην διαδικασία σύγκρισης της πραγματικής μάσκας με την παραγόμενη (generated) μάσκα, χρησιμοποιήθηκε ο λογικός τελεστής XOR (exclusive OR). Αυτός ο τελεστής συγκρίνει δύο δυαδικές τιμές και επιστρέφει μια νέα τιμή που δηλώνει αν οι δύο τιμές είναι διαφορετικές ή όχι.

Πίνακας 4.2 Πίνακας Αληθείας του λογικού τελεστή XOR

A	B	$A \oplus B$
0	0	0
0	1	1
1	0	1
1	1	0

Πρακτικά, ο λογικός τελεστής XOR εφαρμόζεται pixel προς pixel στις δύο εικόνες μάσκας. Όταν τα αντίστοιχα pixel στις δύο μάσκες έχουν την ίδια τιμή (ίσο με 0 ή ίσο με 1), τότε το αντίστοιχο pixel στην παραγόμενη μάσκα που παράγει ο XOR τελεστής είναι 0. Αντίθετα, αν τα αντίστοιχα pixel έχουν διαφορετική τιμή, τότε το αντίστοιχο pixel στην παραγόμενη μάσκα είναι 1. Αυτή η διαδικασία επιτρέπει την αξιολόγηση της ομοιότητας ή της διαφοράς μεταξύ των δύο масκών σε επίπεδο pixel και την απεικόνιση της διαφοράς τους σε μια νέα εικόνα, όπου τα κοινά σημεία τους θα είναι μαύρα και τα σημεία στα οποία διαφέρουν άσπρα.

Συγκρίνοντας την πραγματική μάσκα με τις δύο που έχουν προκύψει (Generated Mask και Improved Mask) προκύπτουν τα εξής αποτελέσματα:

Similarity Score (Generated Mask): 78.10 %

Threshold (Generated Mask): 65

Similarity Score (Improved Mask): 72.95 %

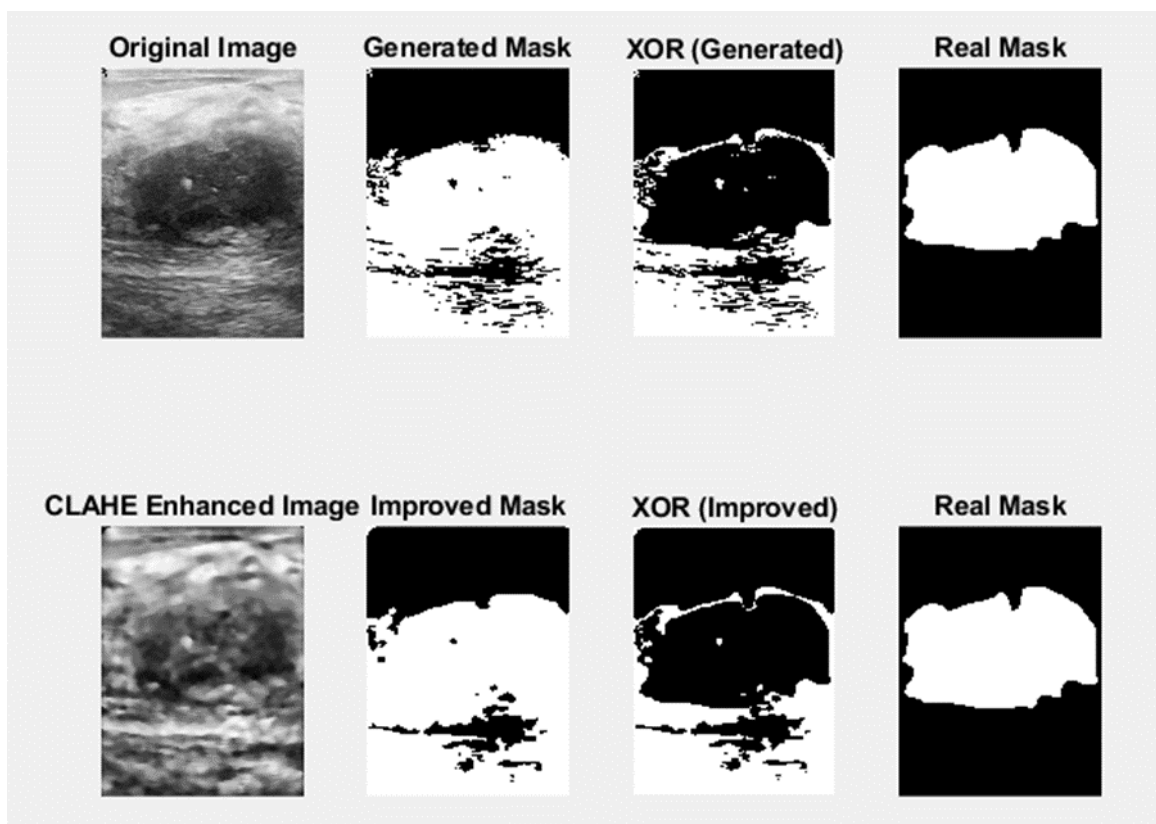
Threshold (Improved Mask): 90

Σε κάθε περίπτωση η επιλογή του Threshold έχει γίνει έπειτα από πειραματισμό με διαφορετικές τιμές. Στο συγκεκριμένο παράδειγμα η τιμή 65 έδωσε την μεγαλύτερη ομοιότητα στην περίπτωση της παραγόμενης μάσκας, ενώ η τιμή 90 έδωσε αντίστοιχα την βέλτιστη δυνατή ομοιότητα στην περίπτωση της μάσκας από την εικόνα που υπέστη επεξεργασία.

Παρατηρήθηκε ότι δεν υπήρξε σημαντική βελτίωση στα αποτελέσματα μετά την εφαρμογή των πρόσθετων επεξεργασιών (CLAHE και φίλτρου median). Δοκιμάζοντας και διαφορετικές εικόνες του dataset, παρατηρήθηκε επίσης ότι για κάθε εικόνα απαιτείται μια διαφορετική βέλτιστη τιμή κατωφλίου, γεγονός που επέβαλε πολλαπλές χειροκίνητες δοκιμές για την εύρεση των βέλτιστων παραμέτρων για κάθε εικόνα ξεχωριστά.

4.2.2 Μέθοδος κατωφλίου Otsu

Μελετώντας διάφορες μεθόδους για την εύρεση της βέλτιστης τιμής threshold για τη δημιουργία των μασκών, δοκιμάστηκε η μέθοδος κατωφλίου Otsu. Στόχος ήταν να διερευνηθεί η δυνατότητα αυτόματης επιλογής της τιμής threshold και η σύγκρισή της με τη χειροκίνητη επιλογή. Αφού ενσωματώθηκε το αντίστοιχο σκέλος κώδικα στον ίδιο αλγόριθμο, ώστε να επιλέγεται αυτόματα η τιμή κατωφλίου, έγινε εκ νέου δοκιμή στην ίδια εικόνα ώστε τα αποτελέσματα να είναι συγκρίσιμα.



Εικόνα 4.4 Αποτελέσματα και σύγκριση εικόνων και μασκών με εφαρμογή κατωφλίου Otsu και CLAHE

Similarity Score (Generated Mask): 66.45 %

Threshold (Generated Mask) : 5.137255e-01

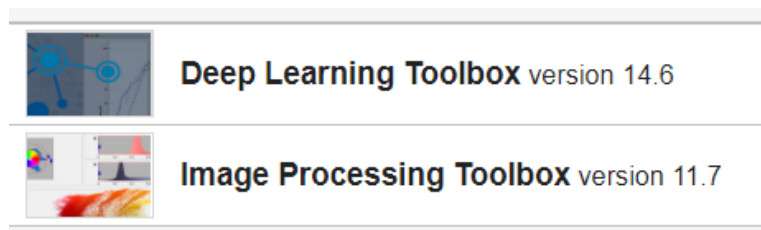
Similarity Score (Improved Mask) : 65.35 %

Threshold (Improved Mask) : 5.176471e-01

Παρ' όλη την αρχική υπόθεση, τα αποτελέσματα από τη μέθοδο Otsu δεν ήταν ικανοποιητικά. Παρατηρήθηκε ότι η αυτόματη διαδικασία επιλογής του threshold δεν προσφέρει καλύτερα ή όμοια αποτελέσματα με την χειροκίνητη επιλογή. Αντιθέτως, φαίνεται ότι για την ίδια εικόνα η τιμή κατωφλίου είναι πολύ διαφορετική, και η παραγόμενη μάσκα έχει μικρότερο ποσοστό ομοιότητας με την πραγματική μάσκα και στις δύο περιπτώσεις, συγκριτικά με προηγούμενως. Έτσι, αφού τα αποτελέσματα δεν ήταν ικανοποιητικά, οι μάσκες δεν θα χρησιμοποιηθούν καθόλου στην τελική μεθοδολογία.

4.3 Προετοιμασία και Προ-επεξεργασία Δεδομένων

Πριν την εκπαίδευση του νευρωνικού δικτύου θα χρειαστεί να γίνει κάποια προετοιμασία των δεδομένων. Η προετοιμασία αυτή περιλαμβάνει σε πρώτο στάδιο την φόρτωση, την προ-επεξεργασία και την διαχείριση τους από το MATLAB προκειμένου να μπορούν να χρησιμοποιηθούν από το δίκτυο.



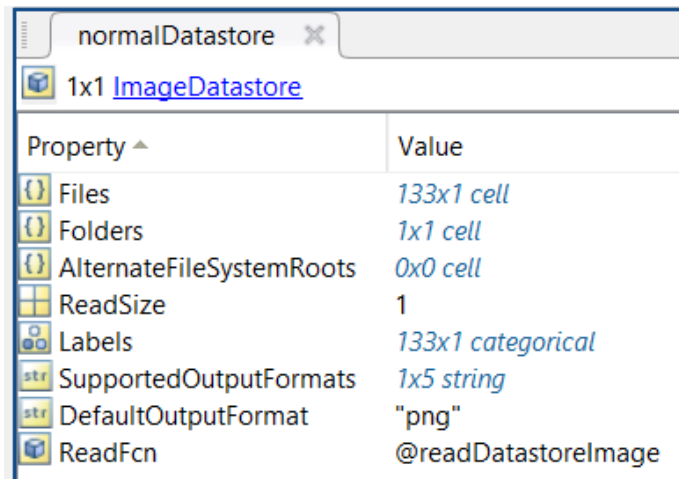
Εικόνα 4.5 Απαιτούμενα πρόσθετα εργαλεία στο MATLAB για την υλοποίηση των πειραμάτων

```
% Path to dataset
datasetPath = './dataset';

% Create datastores for each class
normalDatastore = imageDatastore(fullfile(datasetPath, 'normal'), 'LabelSource',
'foldernames', 'IncludeSubfolders', true);
benignDatastore = imageDatastore(fullfile(datasetPath, 'benign'), 'LabelSource',
'foldernames', 'IncludeSubfolders', true);
malignantDatastore = imageDatastore(fullfile(datasetPath, 'malignant'), 'LabelSource',
'foldernames', 'IncludeSubfolders', true);
```

Για τη φόρτωση και την προ-επεξεργασία των δεδομένων, χρησιμοποιείται η συνάρτηση `imageDatastore` του MATLAB. Αυτή η συνάρτηση επιτρέπει την φόρτωση ενός συνόλου εικόνων από έναν καθορισμένο φάκελο στον υπολογιστή. Έπειτα, με χρήση της συνάρτησης `fullfile` ορίζεται η πλήρης διαδρομή του φακέλου που περιέχει τις εικόνες. Οι παράμετροι `LabelSource`, `foldernames` και `IncludeSubfolders` δηλώνουν ότι ως ετικέτες κλάσεων στα δεδομένα θα χρησιμοποιηθούν τα ονόματα των υποφακέλων, δηλαδή, σε αυτό το πείραμα οι τρεις κλάσεις θα έχουν τα ονόματα **benign**, **malignant** και **normal** όπως ακριβώς ονομάζονται και οι αντίστοιχοι υποφάκελοι που περιέχουν τις αντίστοιχες εικόνες.

Οι εικόνες έχουν αποθηκευτεί σε τρία διαφορετικά `datastores`. Το `datastore` είναι μια δομή δεδομένων που χρησιμοποιείται στο MATLAB για την διαχείριση μεγάλων συνόλων δεδομένων όπως εικόνες, ήχοι, βίντεο, και άλλα πολυμέσα. Έτσι είναι πιο εύκολη και αποτελεσματική η φόρτωση, η επεξεργασία και η πρόσβαση σε αυτά.



Property	Value
Files	133x1 cell
Folders	1x1 cell
AlternateFileSystemRoots	0x0 cell
ReadSize	1
Labels	133x1 categorical
SupportedOutputFormats	1x5 string
DefaultOutputFormat	"png"
ReadFcn	@readDatastoreImage

Εικόνα 4.6 Περιεχόμενα του normalDatastore στο MATLAB

Όπως φαίνεται στην Εικόνα 4.6, το κάθε datastore φέρει ορισμένα χαρακτηριστικά που χρειάζονται για την πρόσβαση και την επεξεργασία τους από το MATLAB. Για παράδειγμα, το στοιχείο Labels που θα χρειαστεί αργότερα είναι ένας πίνακας διαστάσεων 133×1 (όπου 133 είναι το σύνολο των εικόνων normal μιας και απεικονίζεται το normalDatastore), ο οποίος γράφει σε κάθε κελί την λέξη “normal”. Το κείμενο αυτό που υπάρχει στα κελιά θα χρησιμοποιηθεί αργότερα ως ετικέτα της κάθε εικόνας.

```
% Shuffle image order in each datastore  
rng(1); % Set a random seed for reproducibility  
normalDatastore = shuffleImageDatastore(normalDatastore);  
benignDatastore = shuffleImageDatastore(benignDatastore);  
malignantDatastore = shuffleImageDatastore(malignantDatastore);
```

Η συνάρτηση shuffleImageDatastore δημιουργήθηκε ώστε να ανακατεύει τυχαία τα αρχεία με τις ετικέτες τους. Αρχικά, τα αποθηκεύει, έπειτα δημιουργεί έναν τυχαίο πίνακα αναδιάταξης με χρήση της randperm, και στη συνέχεια ανακατεύει τα αρχεία και τις αντίστοιχες ετικέτες βάσει αυτού του πίνακα.

Τέλος, δημιουργεί και επιστρέφει σαν έξοδο ένα νέο datastore με τα ανακατεμένα πλέον αρχεία και ετικέτες, εξασφαλίζοντας έτσι ότι το μοντέλο δεν θα μάθει τυχαίες σχέσεις που προκύπτουν από τη σειρά των δεδομένων. Η συνάρτηση αυτή θα είναι ακόμα πιο χρήσιμη παρακάτω όπου θα συμπεριλαμβάνονται ετικέτες και δεδομένα από όλες τις κλάσεις.

```
% Split the data into training, validation, and testing sets  
[trainNormal, valNormal, testNormal] = splitEachLabel(normalDatastore, 0.7, 0.15, 0.15);  
[trainBenign, valBenign, testBenign] = splitEachLabel(benignDatastore, 0.7, 0.15, 0.15);  
[trainMalignant, valMalignant, testMalignant] = splitEachLabel(malignantDatastore, 0.7, 0.15, 0.15);
```

Αφού έχουν ήδη φορτωθεί τα δεδομένα και έχουν κατηγοριοποιηθεί σε κλάσεις, θα χρειαστεί να διαιρεθούν σε τρία υποσύνολα (training, validation, testing). Χρησιμοποιώντας την συνάρτηση splitEachLabel τα ήδη υπάρχοντα datastores χωρίζονται σε συγκεκριμένα ποσοστά:

- **70%** των εικόνων για την εκπαίδευση (training),
- **15%** για την επικύρωση (validation) και
- **15%** για δοκιμή (testing).

Property	Value
Files	93x1 cell
Folders	0x1 cell
AlternateFileSystemRoots	0x0 cell
ReadSize	1
Labels	93x1 categorical
SupportedOutputFormats	1x5 string
DefaultOutputFormat	"png"
ReadFcn	@readDatastoreImage

Εικόνα 4.7 Περιεχόμενα του υποσυνόλου training για την κλάση ‘normal’ στο MATLAB

Όπως φαίνεται στην Εικόνα 4.7, το υποσύνολο της εκπαίδευσης των δεδομένων της κλάσης ‘normal’, περιέχει πράγματι το 70% των δεδομένων αυτής της κλάσης ($133 \cdot 0.7 = 93,1 \cong 93$), αφού το στοιχείο ‘Files’ περιέχει έναν πίνακα 93×1 .

	1
1	D:\Users\... \Dataset_BUSI_with_GT_no_masks\normal\normal (100).png
2	D:\Users\... \Dataset_BUSI_with_GT_no_masks\normal\normal (68).png
3	D:\Users\... \Dataset_BUSI_with_GT_no_masks\normal\normal (75).png
4	D:\Users\... \Dataset_BUSI_with_GT_no_masks\normal\normal (133).png
5	D:\Users\... \Dataset_BUSI_with_GT_no_masks\normal\normal (24).png
6	D:\Users\... \Dataset_BUSI_with_GT_no_masks\normal\normal (88).png
7	D:\Users\... \Dataset_BUSI_with_GT_no_masks\normal\normal (89).png
8	D:\Users\... \Dataset_BUSI_with_GT_no_masks\normal\normal (111).png
9	D:\Users\... \Dataset_BUSI_with_GT_no_masks\normal\normal (9).png
10	D:\Users\... \Dataset_BUSI_with_GT_no_masks\normal\normal (123).png
11	D:\Users\... \Dataset_BUSI_with_GT_no_masks\normal\normal (37).png
12	D:\Users\... \Dataset_BUSI_with_GT_no_masks\normal\normal (29).png
13	D:\Users\... \Dataset_BUSI_with_GT_no_masks\normal\normal (82).png
14	D:\Users\... \Dataset_BUSI_with_GT_no_masks\normal\normal (122).png
15	D:\Users\... \Dataset_BUSI_with_GT_no_masks\normal\normal (103).png
16	D:\Users\... \Dataset_BUSI_with_GT_no_masks\normal\normal (126).png
17	D:\Users\... \Dataset_BUSI_with_GT_no_masks\normal\normal (33).png

Εικόνα 4.8 Περιεχόμενα του trainNormal.Files στο MATLAB

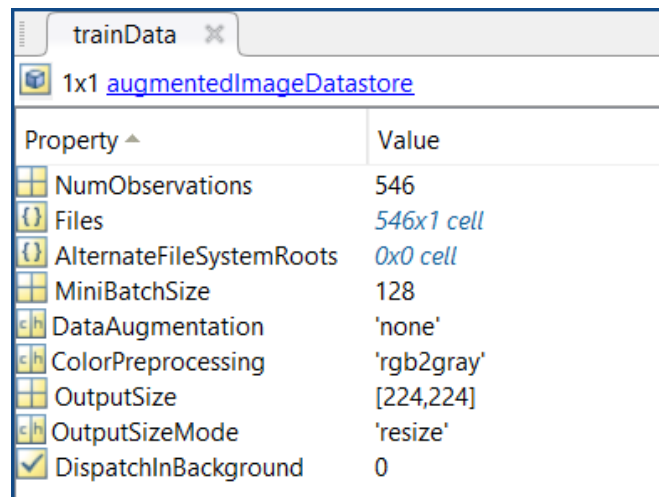
Παρατηρείται επίσης στην Εικόνα 4.8 ότι τα 93 αρχεία που θα χρησιμοποιηθούν στην εκπαίδευση για την κλάση ‘normal’ είναι πράγματι ανακατεμένα όπως ήταν αναμενόμενο μετά την χρήση της συνάρτησης shuffleImageDatastore. Τα αρχεία έχουν τυχαία σειρά και δεν είναι τα πρώτα 93 του φακέλου (δηλαδή **δεν** είναι normal (1).png, normal (2).png, ..., normal (93).png με την σειρά που ήταν στον φάκελο).

Επομένως, έχουν δημιουργηθεί 3 υποσύνολα σε κάθε μία από τις 3 κλάσεις. Όπως ήδη αναφέρθηκε, για τις εικόνες της κλάσης normal, το 70% των εικόνων βρίσκεται πλέον σε ένα datastore που ονομάζεται trainNormal και θα χρησιμοποιηθεί στην εκπαίδευση, το 15% σε ένα δεύτερο datastore με όνομα valNormal που θα χρησιμοποιηθεί κατά την διάρκεια της εκπαίδευσης για την επικύρωση και τέλος το

υπόλοιπο 15% σε ένα τρίτο datastore που ονομάζεται `testNormal` και θα χρησιμοποιηθεί στην δοκιμή του νευρωνικού δικτύου. Αντίστοιχα και για τις άλλες δύο κλάσεις έχουν προκύψει ομοίως τρία διαφορετικά datastores.

% Combine the datastores for all classes

```
trainData = imageDatastore(cat(1, trainNormal.Files, trainBenign.Files,  
trainMalignant.Files), 'LabelSource', 'foldernames');  
valData = imageDatastore(cat(1, valNormal.Files, valBenign.Files, valMalignant.Files),  
'LabelSource', 'foldernames');  
testData = imageDatastore(cat(1, testNormal.Files, testBenign.Files, testMalignant.Files),  
'LabelSource', 'foldernames');
```



Property	Value
NumObservations	546
Files	546x1 cell
AlternateFileSystemRoots	0x0 cell
MiniBatchSize	128
DataAugmentation	'none'
ColorPreprocessing	'rgb2gray'
OutputSize	[224,224]
OutputSizeMode	'resize'
DispatchInBackground	0

Εικόνα 4.9 Περιεχόμενα του `trainData` στο MATLAB

Αφού λοιπόν έχουν ήδη χωριστεί, θα πρέπει να γίνει η συγχώνευση όλων των δεδομένων training μεταξύ τους, όλων των δεδομένων validation και φυσικά όλων των δεδομένων testing. Αυτό επιτυγχάνεται με την συνάρτηση `cat` του MATLAB η οποία θα ενώσει τα `trainMalignant`, `trainBenign` και `trainNormal` σε ένα ενιαίο datastore που πλέον θα λέγεται `trainData` (Εικόνα 4.9) και θα περιλαμβάνει όλα τα δεδομένα από κάθε κλάση που θα χρησιμοποιηθούν στην εκπαίδευση του νευρωνικού δικτύου. Ομοίως δημιουργούνται και τα `valData` και `testData` για την επικύρωση και την δοκιμή αντίστοιχα.

% Shuffle image order for augmented/combined datastores

```
trainData = shuffleImageDatastore(trainData);  
valData = shuffleImageDatastore(valData);  
testData = shuffleImageDatastore(testData);
```

	1	2	3
1	normal		
2	normal		
3	normal		
4	normal		
5	normal		
6	normal		
7	normal		
8	normal		
9	normal		
10	normal		
11	normal		
12	normal		
13	normal		
14	normal		
15	normal		
16	normal		
17	normal		
18	normal		

	1	2	3
1	malignant		
2	benign		
3	malignant		
4	benign		
5	normal		
6	benign		
7	benign		
8	benign		
9	benign		
10	malignant		
11	benign		
12	benign		
13	benign		
14	benign		
15	benign		
16	benign		
17	benign		
18	benign		

Εικόνα 4.10 Ετικέτες του trainData (α) πριν το ανακάτεμα και (β) μετά το ανακάτεμα

Τέλος, τα δεδομένα και οι αντίστοιχες ετικέτες τους ανακατεύονται με τη χρήση της συνάρτησης `shuffleImageDatastore` που χρησιμοποιήθηκε και προηγουμένως, για να εξασφαλιστεί η τυχαιότητα στη σειρά των δεδομένων κατά την εκπαίδευση του μοντέλου. Σε αντίθετη περίπτωση τα πρώτα δείγματα θα ήταν όλα από μία συγκεκριμένη κατηγορία, όπως για παράδειγμα τα `normal` στην περίπτωση του `trainData` που τυχαίνει να βρίσκονται πρώτα σε σειρά κατά την συγχώνευση (Εικόνα 4.10α), με αποτέλεσμα το μοντέλο να εκπαιδευτεί με έμφαση σε αυτήν την κατηγορία και να αγνοήσει τις άλλες. Αυτό θα οδηγούσε σε ένα μοντέλο που είναι υπερ-εκπαιδευμένο στα δείγματα της πρώτης κατηγορίας (`normal`) και δεν θα ήταν σε θέση να κατηγοριοποιήσει σωστά εικόνες από τις άλλες δύο κατηγορίες, (`benign` και `malignant`).

```
% Get the true labels from the testDatastore
trueLabels = testData.Labels;
```

Ολοκληρώνοντας την διαδικασία προεπεξεργασίας των δεδομένων, οι πραγματικές ετικέτες των δεδομένων `testing` (δηλαδή από το `testData`) αποθηκεύονται στην μεταβλητή `trueLabels`, ώστε μετά την εκπαίδευση του νευρωνικού δικτύου να είναι δυνατή η σύγκριση των προβλέψεων του μοντέλου με τις πραγματικές κλάσεις που ανήκουν τα δεδομένα αυτά για να αξιολογηθεί η απόδοσή του.

```
% Preprocess and resize the images
trainData = augmentedImageDatastore(inputSize(1:2), trainData, 'ColorPreprocessing',
'rgb2gray');
valData = augmentedImageDatastore(inputSize(1:2), valData, 'ColorPreprocessing',
'rgb2gray');
testData = augmentedImageDatastore(inputSize(1:2), testData, 'ColorPreprocessing',
'rgb2gray');
```

Με την συνάρτηση `augmentedImageDatastore` γίνονται οι τελικές απαραίτητες μετατροπές στις εικόνες των ήδη υπάρχοντων `datastores` ώστε να αναδιαμορφωθούν σε ένα επιθυμητό μέγεθος. Συγκεκριμένα, η επιλογή `'ColorPreprocessing', 'rgb2gray'` δηλώνει ότι οι εικόνες θα μετατραπούν από έγχρωμες σε κλίμακα γκρι, δηλαδή σε μονόχρωμες εικόνες.

Επιπλέον, ορίζεται το επιθυμητό μέγεθος των εικόνων με την παράμετρο `inputSize(1:2)`, έτσι ώστε όλες οι εικόνες να έχουν τις ίδιες διαστάσεις πριν από την εισαγωγή τους στο νευρωνικό δίκτυο. Το `inputSize(1:2)` αναφέρεται στις πρώτες δύο διαστάσεις του μεγέθους των εικόνων, δηλαδή στο μήκος και στο πλάτος των εικόνων, αγνοώντας το βάθος ή τον αριθμό των καναλιών χρώματος.

Στο συγκεκριμένο πείραμα, το `inputSize` είναι είτε `[224, 224, 3]` είτε `[224, 224, 1]`, ανάλογα τις εικόνες μιας και μερικές από αυτές έχουν κάποια έγχρωμα στοιχεία, άρα το `inputSize(1:2)` θα είναι `[224, 224]`, εξασφαλίζοντας έτσι ότι όλες οι εικόνες θα έχουν τις ίδιες διαστάσεις πριν την εισαγωγή τους στο νευρωνικό δίκτυο. Η διαδικασία αυτή κατά την προ-επεξεργασία εξασφαλίζει την ομοιογένεια στα δεδομένα.

4.3.1 Δημιουργία dataset με χρήση μετασχηματισμού FFT-2D και συχνοτικού φίλτρου

Εφόσον είναι ήδη έτοιμο το βασικότερο μέρος του πειράματος, δηλαδή ο κυρίως κώδικας που θα γίνεται η φόρτωση των δεδομένων και η εκπαίδευση του νευρωνικού δικτύου, θα χρειαστεί για κάποια από τα πειράματα που θα ακολουθήσουν να γίνει επιπλέον επεξεργασία των εικόνων του dataset και αποθήκευσή τους εκ νέου σε διαφορετικό φάκελο, ώστε να δοθούν στο δίκτυο σε διαφορετική μορφή προκειμένου να γίνει σύγκριση της συμπεριφοράς και της αποτελεσματικότητάς του μοντέλου στην κατηγοριοποίησή τους. Συγκεκριμένα, θα εξαχθούν από τις αρχικές εικόνες οι συντελεστές του διδιάστατου μετασχηματισμού Fourier τους. Στη συνέχεια θα εφαρμοστεί χαμηλοπερατό 2D φίλτρο στο πεδίο των χωρικών συχνοτήτων, και το φιλτραρισμένο αποτέλεσμα θα επιστρέψει στο αρχικό πεδίο του σήματος με αντίστροφο μετασχηματισμό Fourier-2D.

Δημιουργήθηκε κώδικας στο MATLAB ο οποίος εκτελεί την επεξεργασία όλων των εικόνων από τον φάκελο εισόδου, εφαρμόζοντας ένα γκαουσιανό χαμηλοπερατό φίλτρο στο πεδίο της συχνότητας μέσω του μετασχηματισμού Fourier-2D, και στο τέλος αποθηκεύει τις φιλτραρισμένες εικόνες σε έναν νέο φάκελο εξόδου.

```
% Αρχικοποίηση φακέλων
inputFolder = './Dataset_BUSI_with_GT_no_masks\benign';
outputFolder = './fft2_lowgaussian\benign';

% Δημιουργία φακέλου εξόδου εάν δεν υπάρχει
if ~exist(outputFolder, 'dir')
    mkdir(outputFolder);
end

% Ανάγνωση όλων των αρχείων από τον φάκελο
imageFiles = dir(fullfile(inputFolder, '*.png'));
```

Αρχικά, ορίζονται οι φάκελοι εισόδου και εξόδου, και εάν ο φάκελος εξόδου δεν υπάρχει, δημιουργείται. Στη συνέχεια, διαβάζονται όλα τα αρχεία εικόνων (.png) από τον φάκελο εισόδου.

```
for i = 1:length(imageFiles)
    inputImagePath = fullfile(inputFolder, imageFiles(i).name);
    img = imread(inputImagePath);
    img = rgb2gray(img);

    % Μετασχηματισμός Fourier 2D
    F = fft2(double(img));

    % Μετακίνηση του μηδενικού συχνοτικού συντελεστή στο κέντρο
```

```
F_shifted = fftshift(F);
```

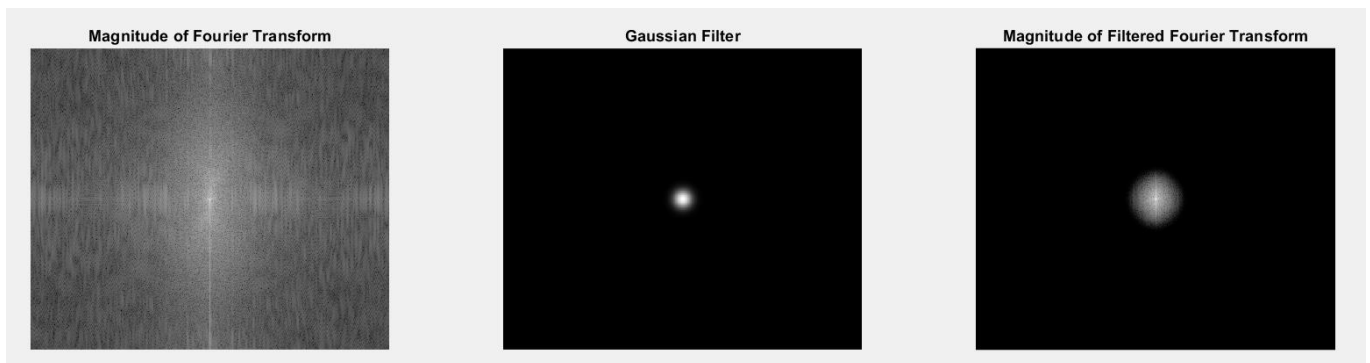
Κάθε εικόνα στον φάκελο διαβάζεται και μετατρέπεται σε εικόνα κλίμακας του γκρι. Εφαρμόζεται στη συνέχεια ο μετασχηματισμός Fourier-2D για να μεταβεί η εικόνα στο πεδίο της χωρικής συχνότητας, και ακολούθως ο μηδενικός συχνοτικός συντελεστής (σημείο 0,0) μετακινείται στο κέντρο του φάσματος με τη χρήση της συνάρτησης `fftshift`.

```
% Δημιουργία γκαουσιανού φίλτρου
[M, N] = size(F_shifted);
[X, Y] = meshgrid(-floor(N/2):floor((N-1)/2), -floor(M/2):floor((M-1)/2));
sigma = 10; % Παράμετρος σ του φίλτρου
gaussian_filter = exp(-(X.^2 + Y.^2) / (2*sigma^2)); %low pass

% Εφαρμογή του φίλτρου στο πεδίο της συχνότητας
F_filtered = F_shifted .* gaussian_filter;
```

Οι διαστάσεις της εικόνας `F_shifted` αποθηκεύονται στις μεταβλητές `M` (γραμμές) και `N` (στήλες) και στη συνέχεια δημιουργούνται δύο πίνακες `X` και `Y` που αντιστοιχούν στις διαστάσεις του πλέγματος συντεταγμένων που θα χρησιμοποιηθεί για την δημιουργία του φίλτρου.

- Τα `floor(N/2)` και `floor((N-1)/2)` εξασφαλίζουν ότι οι συντεταγμένες εκτείνονται από το αρνητικό μισό του εύρους της μήτρας μέχρι το θετικό μισό, χωρίς να υπερβαίνουν τα όρια.
- Για παράδειγμα, αν το `N` είναι 10, τότε το `floor(N/2)` είναι 5 και το `floor((N-1)/2)` είναι 4, οπότε οι συντεταγμένες καλύπτουν το εύρος από -5 έως 4.
- Έτσι εξασφαλίζεται ο σωστός αριθμός στοιχείων που είναι αναγκαία για το πλέγμα, δηλαδή 10 στοιχεία συνολικά στο συγκεκριμένο παράδειγμα (από -5 έως 4 υπάρχουν 10 διακριτά σημεία, ενώ από -5 έως 5 θα υπήρχαν 11, κάτι που δεν ταιριάζει με τις συμμετρικές διαστάσεις της μήτρας συχνοτήτων που χρειάζεται για το φίλτρο).



Εικόνα 4.11 (α) Οπτικοποίηση των συχνοτήτων μετά τον μετασχηματισμό Fourier (β) Γκαουσιανό χαμηλοπερατό φίλτρο (γ) Φιλτραρισμένη εικόνα στο πεδίο της συχνότητας

Δημιουργείται στη συνέχεια ένα γκαουσιανό φίλτρο χαμηλής διέλευσης σύμφωνα με τον τύπο $H(u, v) = e^{-D(u,v)^2/(2\sigma^2)}$ (Κεφάλαιο 2.2.2.3) και εφαρμόζεται στο πεδίο της συχνότητας (Εικόνα 4.11).

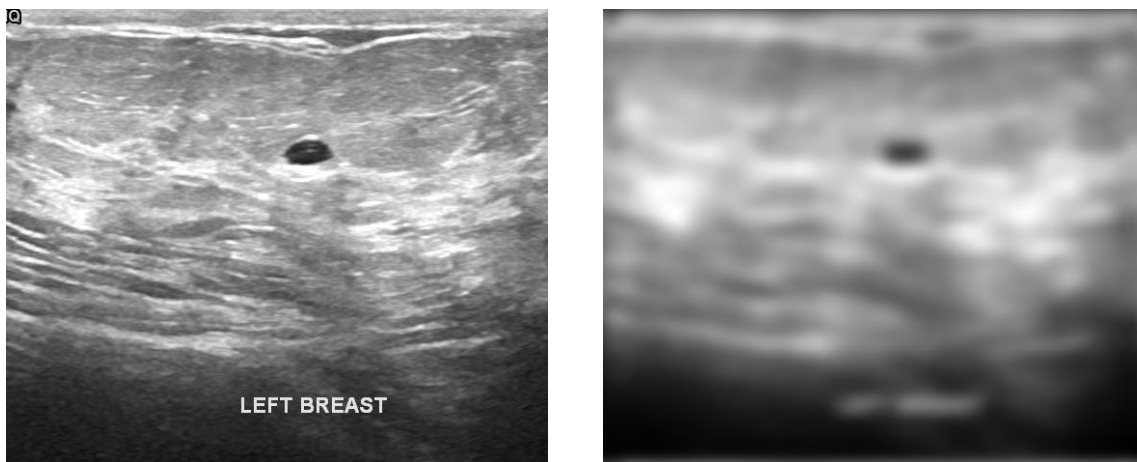
```
% Αντίστροφος μετασχηματισμός Fourier
F_inv_shifted = ifftshift(F_filtered);
img_filtered = ifft2(F_inv_shifted);

% Πραγματικό μέρος της εικόνας
img_filtered = real(img_filtered);
```

```
% Κανονικοποίηση της εικόνας (τιμές μεταξύ 0 και 255)
img_filtered = uint8(255 * mat2gray(img_filtered));

% Αποθήκευση της φιλτραρισμένης εικόνας
outputImagePath = fullfile(outputFolder, imageFiles(i).name);
imwrite(img_filtered, outputImagePath);
end
```

Έπειτα, εκτελείται ο αντίστροφος μετασχηματισμός Fourier για να επανέλθει η εικόνα στο χωρικό πεδίο και λαμβάνεται το πραγματικό μέρος της. Η εικόνα κανονικοποιείται ώστε οι τιμές να βρίσκονται στο εύρος 0 – 255 (που αναπαριστά την φωτεινότητα των pixels, ώστε να επανέλθει σε κλίμακα γκρι με το 0 να αναπαριστά το μαύρο ως ελάχιστη τιμή και το 255 το άσπρο ως μέγιστη τιμή), και τελικά αποθηκεύεται στον φάκελο εξόδου.



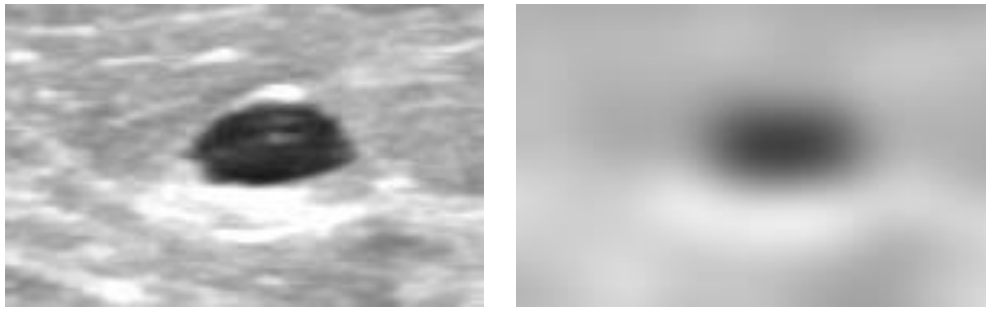
Εικόνα 4.12 Η εικόνα benign (1).png (α) στην αρχική μορφή (β) κατόπιν εφαρμογής του γκαουσιανού χαμηλοπερατού φίλτρου

Τελικά οι εικόνες που έχουν προκύψει έχουν πιο εξομαλυμένα χαρακτηριστικά, χωρίς απότομες αλλαγές στις γκρι αποχρώσεις, πράγμα το οποίο απαλείφει την παρουσία ακμών ή περιττών περιγραμμάτων έξω από τις περιοχές ενδιαφέροντος που μπορεί να μάθει το δίκτυο (Εικόνα 4.12).

Επειδή οι εικόνες υπερήχου συχνά περιέχουν πολύ θόρυβο, ο οποίος μπορεί να επηρεάσει αρνητικά την απόδοση των νευρωνικών δικτύων, η εφαρμογή του φίλτρου στοχεύει να βοηθήσει στη μείωση αυτού του θορύβου που περιέχεται στις εικόνες, διατηρώντας τα σημαντικά χαρακτηριστικά και επιτρέποντας έτσι στο οποιοδήποτε δίκτυο να επικεντρωθεί σε πιο καθαρά και αναγνωρίσιμα μοτίβα, χωρίς να «αποσπάται» από ασήμαντες λεπτομέρειες.

Οι θολωμένες εικόνες μέσω του γκαουσιανού χαμηλοπερατού φίλτρου περιέχουν κυρίως πληροφορίες χαμηλής συχνότητας, δηλαδή βασικές δομές, αποχρώσεις και γενικά σχήματα της εικόνας, τα οποία είναι τα πιο χρήσιμα χαρακτηριστικά για την αναγνώριση και κατηγοριοποίηση των όγκων στο παρόν πείραμα.

Απλοποιώντας λοιπόν τις εικόνες μειώνεται και η πολυπλοκότητα των δεδομένων, πράγμα το οποίο μπορεί να κάνει το πρόβλημα πιο διαχειρίσιμο για το νευρωνικό δίκτυο και να το βοηθήσει να μάθει τα σωστά χαρακτηριστικά γρηγορότερα και πιο αποτελεσματικά κατά την εκπαίδευση.



Εικόνα 4.13 Μεγέθυνση στο σημείο ενδιαφέροντος της εικόνας benign (1).png
(α) στην αρχική μορφή (β) κατόπιν εφαρμογής του γκαουσιανού χαμηλοπερατού φίλτρου

Ήδη όπως φαίνεται στην Εικόνα 4.14 και Εικόνα 4.15, δεν παρουσιάζονται απότομες αλλαγές στις αποχρώσεις, τόσο στις σκούρες αποχρώσεις μεταξύ τους, όσο και μεταξύ σκούρων και ανοιχτόχρωμων.

Επομένως, η φύση των δεδομένων φαίνεται να είναι πιο απλοποιημένη και τα σημαντικότερα στοιχεία να ξεχωρίζουν περισσότερο.

Ειδικά στην ανοιχτή γκρι περιοχή γύρω από τον όγκο (Εικόνα 4.13β) φαίνεται να υπάρχουν παρόμοιες αποχρώσεις του γκρι, με αποτέλεσμα να ξεχωρίζει περισσότερο η πιο απότομη μεταβολή απόχρωσης στο σκούρο σημείο του όγκου.

	303	304	305	306	307	308	309	310	311	312	313	314	315	316
136	195	199	208	204	193	179	168	161	157	153	153	155	161	170
137	199	187	184	164	150	140	126	119	117	112	114	115	114	112
138	181	169	151	130	113	102	90	86	87	85	85	86	85	82
139	125	113	104	91	82	78	73	71	71	69	68	69	69	65
140	73	66	63	60	60	63	64	63	63	63	70	74	77	75
141	52	52	52	54	56	58	60	60	61	65	74	81	88	92
142	47	50	54	61	65	67	70	68	67	71	74	81	88	93
143	46	49	49	57	59	62	69	67	63	66	69	72	74	75
144	41	43	49	52	47	46	55	58	56	61	59	59	58	56
145	41	42	42	42	41	44	47	47	51	60	61	67	71	69
146	42	45	42	43	43	47	53	56	62	73	87	99	107	104
147	45	49	50	53	54	58	65	69	77	89	101	116	127	123
148	51	56	62	65	66	69	73	75	79	88	98	108	116	113
149	58	62	65	69	69	70	71	68	67	70	79	80	81	82
150	62	62	62	65	65	64	65	61	56	57	59	54	52	55
151	64	61	60	61	59	59	62	61	56	56	56	52	50	50
152	67	62	59	59	55	55	61	62	59	59	55	56	53	49
153	63	64	64	63	61	59	57	56	56	55	53	51	47	45

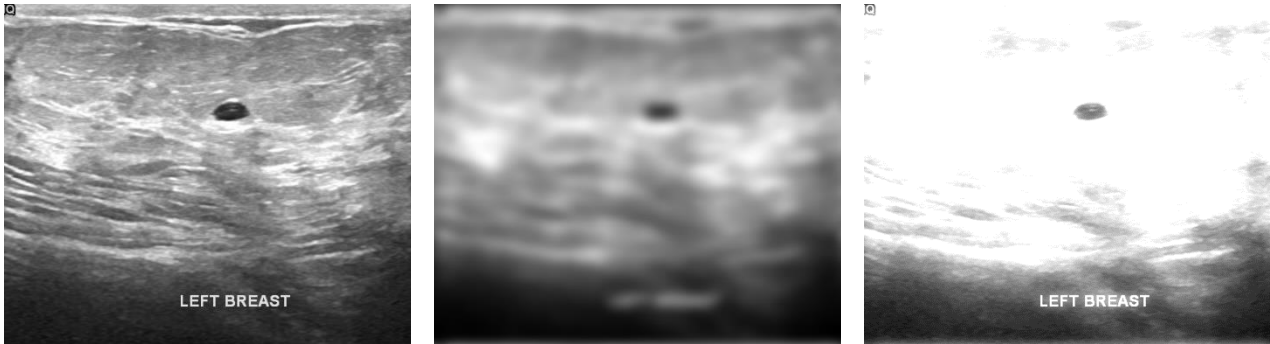
Εικόνα 4.14 Οι τιμές γκρι της αρχικής εικόνας benign (1).png περίπου στο σημείο ενδιαφέροντος

	303	304	305	306	307	308	309	310	311	312	313	314	315	316	
136	138	137	136	136	135	135	135	135	135	135	135	136	136	136	^
137	132	131	130	129	129	128	128	128	128	128	128	128	129	129	
138	125	124	123	122	122	121	121	121	121	121	121	121	121	122	
139	119	118	116	116	115	114	114	114	113	113	113	114	114	114	
140	113	111	110	109	108	107	107	106	106	106	106	106	106	107	
141	106	105	103	102	101	100	100	99	99	99	99	99	99	99	
142	101	99	97	96	95	94	93	93	92	92	92	92	92	93	
143	95	93	91	90	89	88	87	87	86	86	86	86	86	86	
144	90	88	86	85	83	82	82	81	81	80	80	80	80	81	
145	86	84	82	80	79	78	77	76	76	75	75	75	75	76	
146	83	80	78	76	75	74	73	72	72	71	71	71	71	72	
147	80	78	75	74	72	71	70	69	69	68	68	68	68	69	
148	78	76	74	72	70	69	68	67	67	66	66	66	66	67	
149	78	75	73	71	69	68	67	66	66	65	65	65	65	66	
150	78	76	73	71	70	68	67	66	66	65	65	65	66	66	
151	80	77	75	73	71	70	69	68	67	67	67	67	67	68	
152	83	80	78	75	74	72	71	70	70	70	69	70	70	70	
153	87	84	81	79	78	76	75	74	74	73	73	74	74	74	v

Εικόνα 4.15 Οι τιμές γκρι της φιλτραρισμένης εικόνας benign (1).png περίπου στο σημείο ενδιαφέροντος

4.3.2 Δημιουργία dataset με άθροιση των αρχικών και των φιλτραρισμένων εικόνων

Η επόμενη μέθοδος επεξεργασίας του υπάρχοντος dataset, είναι η άθροιση των αρχικών εικόνων με τις φιλτραρισμένες που δημιουργήθηκαν από την εφαρμογή του γκαουσιανού χαμηλοπερατού φίλτρου προηγουμένως (Κεφάλαιο 4.3.1). Στόχος αυτής της μεθόδου είναι η ανάδειξη των σκουρόχρωμων σημείων, καθώς με την πρόσθεση των εικόνων οι πιο ανοιχτόχρωμες περιοχές γύρω από αυτά θα γίνουν άσπρες ή θα τείνουν περισσότερο σε τιμές κοντά στο 255 (Εικόνα 4.16).



Εικόνα 4.16 Η εικόνα benign(1).png (α) στην αρχική μορφή (β) κατόπιν εφαρμογής του γκαουσιανού χαμηλοπερατού φίλτρου (γ) κατόπιν πρόσθεσης των (α) και (β)

	303	304	305	306	307	308	309	310	311	312	313	314	315	316
132	255	255	255	255	255	255	255	255	255	255	255	255	255	255
133	255	255	255	255	255	255	255	255	255	255	255	255	255	255
134	255	255	255	255	255	255	255	255	255	255	255	255	255	255
135	255	255	255	255	255	255	255	255	255	255	255	255	255	255
136	255	255	255	255	255	255	255	255	255	255	255	255	255	255
137	255	255	255	255	255	255	254	247	245	240	242	243	243	243
138	255	255	255	255	252	235	223	211	207	208	206	206	207	206
139	251	244	231	220	207	197	192	187	185	184	182	181	183	183
140	195	186	177	173	169	168	170	171	169	169	169	176	180	183
141	163	158	157	155	156	157	158	160	159	160	164	173	180	187
142	152	148	149	151	157	160	161	163	161	159	163	166	173	180
143	149	141	142	140	147	148	150	156	154	149	152	155	158	160
144	140	131	131	135	137	130	128	137	139	137	141	139	139	138
145	131	127	126	124	122	120	122	124	123	127	135	136	142	146
146	126	125	125	120	119	118	121	126	128	134	144	158	170	178
147	124	125	127	125	127	126	129	135	138	146	157	169	184	195
148	126	129	132	136	137	136	138	141	142	146	154	164	174	182
149	133	136	137	138	140	138	138	138	134	133	135	144	145	146
150	140	140	138	135	136	135	137	137	127	127	127	124	110	118

Εικόνα 4.17 Οι τιμές γκρι της εικόνας benign(1).png που έχει προκύψει από την άθροιση της αρχικής και της φιλτραρισμένης, περίπου στο σημείο ενδιαφέροντος

Ο κώδικας που δημιουργήθηκε διαβάζει όλες τις εικόνες από δύο φακέλους εισόδου, προσθέτει τις αντίστοιχες εικόνες μεταξύ τους και τέλος αποθηκεύει τις νέες εικόνες σε έναν νέο φάκελο εξόδου.

Από την ίδια την εικόνα που έχει προκύψει (Εικόνα 4.16γ) όπως επίσης και από τις τιμές των pixel σε αυτήν (Εικόνα 4.17), πράγματι επιβεβαιώνεται ότι οι νέες εικόνες έχουν περισσότερο άσπρο γύρω από τις σκουρόχρωμες περιοχές ενδιαφέροντος.

Η αύξηση της φωτεινότητας γύρω από τις ύποπτες περιοχές μπορεί να καταστήσει πιο εμφανή τα σύνορα των όγκων ή τις μικρές λεπτομέρειες της δομής τους, όπως είναι το σχήμα τους. Για παράδειγμα, στις περισσότερες καλοήθειες περιπτώσεις οι όγκοι είναι συνήθως πιο στρογγυλοί και έχουν πιο ομαλό σχήμα,

ενώ στις κακοήθεις περιπτώσεις οι όγκοι είναι συχνά ακαθόριστου σχήματος. Άρα με αυτήν την μέθοδο ενισχύονται αυτά τα χαρακτηριστικά που είναι σημαντικά για την διάκριση των περιπτώσεων.

4.3.3 Επίπεδα Νευρωνικών Δικτύων

Τα επίπεδα που χρησιμοποιήθηκαν κατά την παρούσα έρευνα στην δημιουργία των νευρωνικών δικτύων βαθιάς μάθησης που θα αναλυθούν σε επόμενα κεφάλαια περιλαμβάνουν:

1. Επίπεδο Εισόδου (Input Layer)

```
imageInputLayer([224 224 1])
```

Αυτό το επίπεδο ορίζει το μέγεθος των εικόνων εισόδου. Εδώ, οι εικόνες εισόδου έχουν διαστάσεις 224×224 pixels και είναι μονόχρωμες (1 κανάλι). Το επίπεδο εισόδου δεν πραγματοποιεί καμία επεξεργασία αλλά προετοιμάζει τα δεδομένα για τα επόμενα επίπεδα.

2. Συνελικτικά Επίπεδα (Convolution Layers)

```
convolution2dLayer(3, 32, 'Padding', 'same')
```

Αυτό το επίπεδο εκτελεί συνελίξεις (convolutions) στις εικόνες εισόδου με φίλτρα διαστάσεων 3×3. Ο αριθμός 32 υποδεικνύει ότι χρησιμοποιούνται 32 φίλτρα. Άρα στο συγκεκριμένο επίπεδο, υπάρχουν 32 φίλτρα 3×3 τα οποία «γλιστράνε» πάνω στην εικόνα και εκτελούν συνελίξεις σε κάθε θέση που τοποθετούνται.

Η παράμετρος 'Padding', 'same' διατηρεί το μέγεθος της εικόνας μετά τη συνέλιξη, προσθέτοντας μηδενικά περιμετρικά της εικόνας. Αν δεν καθοριστεί το Padding, τότε η συνέλιξη εκτελείται μόνο στα μέρη της εικόνας όπου το φίλτρο μπορεί να εφαρμοστεί πλήρως, με αποτέλεσμα να χάνεται η πληροφορία που βρίσκεται στις άκρες της εικόνας.

3. Επίπεδα Κανονικοποίησης (Batch Normalization Layers)

```
batchNormalizationLayer
```

Αυτό το επίπεδο κανονικοποιεί τις μικρές παρτίδες δεδομένων (mini-batch) προκειμένου να επιταχύνει την εκπαίδευση του νευρωνικού δικτύου. Χρησιμοποιείται μεταξύ του συνελικτικού επιπέδου convolution2dLayer και του μη γραμμικού επιπέδου reluLayer.

4. Επίπεδα ReLU (Rectified Linear Unit Layers)

```
reluLayer
```

Αυτό το επίπεδο εφαρμόζει τη συνάρτηση ενεργοποίησης Rectified Linear Unit (ReLU), που εισάγει μη-γραμμικότητα στο μοντέλο, επιτρέποντας την εκμάθηση πολύπλοκων χαρακτηριστικών.

5. Επίπεδα Υποδειγματοληψίας (Pooling Layers)

```
maxPooling2dLayer(2, 'Stride', 2)
```

Αυτό το επίπεδο εκτελεί μέγιστη υποδειγματοληψία (max pooling) σε περιοχές 2×2 της εικόνας, διαιρώντας την είσοδο σε ορθογώνιες περιοχές και στη συνέχεια υπολογίζοντας το μέγιστο κάθε περιοχής. Η μετατόπιση (stride) 2 σημαίνει ότι το παράθυρο υποδειγματοληψίας κινείται κατά 2 pixels κάθε φορά.

6. Επίπεδα Αποφυγής Υπερεκπαίδευσης (Dropout Layers)

```
dropoutLayer(0.1)
```

Το επίπεδο αυτό χρησιμοποιείται για την αποφυγή υπερεκπαίδευσης (overfitting) κατά την εκπαίδευση ενός νευρωνικού δικτύου. Η παράμετρος 0.1 στο συγκεκριμένο παράδειγμα καθορίζει το ποσοστό των νευρώνων που θα «απενεργοποιηθούν» τυχαία στην κάθε ενημέρωση του βάρους, δηλαδή το 10% των νευρώνων σε αυτό το επίπεδο δεν θα συμμετέχει στην εκάστοτε ενημέρωση, γεγονός που αποτρέπει το δίκτυο από το να εξαρτάται υπερβολικά από συγκεκριμένους νευρώνες και ενισχύει τη γενική του ικανότητα να μαθαίνει από τα δεδομένα εκπαίδευσης.

7. Πλήρως Συνδεδεμένα Επίπεδα (Fully Connected Layers)

`fullyConnectedLayer(256)`

Το επίπεδο αυτό συνδέεται πλήρως με όλες τις εξόδους του προηγούμενου επιπέδου και το συγκεκριμένο έχει 256 νευρώνες. Το τελευταίο πλήρως συνδεδεμένο επίπεδο σε όλα τα πειράματα έχει 3 νευρώνες `fullyConnectedLayer(3)` και χρησιμοποιείται για την κατηγοριοποίηση των δεδομένων εισόδου σε μία από τις τρεις κατηγορίες. Ο κάθε νευρώνας αντιστοιχεί και σε μία κατηγορία εξόδου (benign, malignant, normal).

8. Επίπεδο Softmax (Softmax Layer)

`softmaxLayer`

Το επίπεδο softmax εφαρμόζει τη συνάρτηση softmax στις εξόδους του προηγούμενου επιπέδου, μετατρέποντας τις εξόδους σε πιθανότητες με συνολικό άθροισμα το 1. Παρέχει δηλαδή την πιθανότητα της εξόδου να ανήκει σε κάθε μία από τις 3 κατηγορίες.

9. Επίπεδο Κατηγοριοποίησης (Classification Layer)

`classificationLayer`

Το επίπεδο κατηγοριοποίησης είναι το τελικό επίπεδο του δικτύου και ορίζει την κατηγορία με την υψηλότερη πιθανότητα ως την τελική πρόβλεψη του μοντέλου.

4.3.4 Παράμετροι Εκπαίδευσης (Training Options)

Η εκπαίδευση ενός νευρωνικού δικτύου επηρεάζεται, εκτός από την αρχιτεκτονική του, και από διάφορες άλλες παραμέτρους. Κατανοώντας και βελτιστοποιώντας αυτές τις παραμέτρους, μπορεί να βελτιωθεί σημαντικά η απόδοση του μοντέλου.

Οι κυριότερες παράμετροι εκπαίδευσης που χρησιμοποιήθηκαν και τροποποιήθηκαν κατά την διάρκεια του πειράματος είναι οι εξής:

'sgdm', ...

Ορίζει τον αλγόριθμο βελτιστοποίησης ως “Stochastic Gradient Descent with Momentum” (SGDM). Αυτός ο αλγόριθμος χρησιμοποιείται για να βελτιώσει τη σταθερότητα και την ταχύτητα της εκπαίδευσης εφαρμόζοντας την ορμή (momentum) για να αποφευχθούν τα τοπικά ελάχιστα και να επιταχυνθεί η ισορροπία.

'MaxEpochs', 10, ...

Ορίζει τον μέγιστο αριθμό εποχών (epochs) για την εκπαίδευση. Ως εποχή ορίζεται μία πλήρης διέλευση από το σύνολο δεδομένων εκπαίδευσης. Στο συγκεκριμένο παράδειγμα, το μοντέλο θα εκπαιδευτεί για 10 εποχές. Η ρύθμιση του αριθμού των εποχών είναι σημαντική για να διασφαλιστεί ότι το μοντέλο δεν υπερεκπαιδεύεται ή υπο-εκπαιδεύεται.

'MiniBatchSize', 32, ...

Καθορίζει το μέγεθος του mini-batch, δηλαδή τον αριθμό των δειγμάτων που χρησιμοποιούνται για κάθε ενημέρωση των βαρών του μοντέλου. Ένα mini-batch size των 32 σημαίνει ότι τα βάρη θα ενημερώνονται μετά από κάθε 32 δείγματα. Αυτό επηρεάζει την ταχύτητα και τη σταθερότητα της εκπαίδευσης, όπως επίσης και την απαιτούμενη μνήμη. Μεγαλύτερο MiniBatchSize σημαίνει λιγότερες επαναλήψεις (iterations) και αντίστροφα.

'Shuffle', 'every-epoch', ...

Ορίζει το ανακάτεμα των δεδομένων να γίνεται κάθε εποχή. Αυτό βοηθάει να αποφευχθεί η υπερεκπαίδευση σε συγκεκριμένες σειρές δεδομένων και διασφαλίζει ότι το μοντέλο δεν μαθαίνει μοτίβα που προκύπτουν από τη σειρά των δεδομένων.

'InitialLearnRate', 1e-05, ...

Ορίζει τον αρχικό ρυθμό μάθησης. Στο συγκεκριμένο παράδειγμα, ο ρυθμός μάθησης έχει οριστεί 1e-05, δηλαδή 10^{-5} . Ένας μικρός αρχικός ρυθμός μάθησης σημαίνει ότι οι ενημερώσεις των βαρών θα είναι μικρές, πράγμα που βοηθά στη σταθερότερη εκπαίδευση και αποτρέπει την υπερβολική προσαρμογή.

'ValidationData', valData, ...

Παρέχει τα δεδομένα επικύρωσης που θα χρησιμοποιούνται για την αξιολόγηση της απόδοσης του μοντέλου κατά τη διάρκεια της εκπαίδευσης. Τα δεδομένα επικύρωσης χρησιμοποιούνται για να παρακολουθείται η απόδοση του μοντέλου και να εντοπίζεται πιθανή υπερεκπαίδευση (overfitting).

'ValidationFrequency', 10, ...

Καθορίζει πόσο συχνά θα αξιολογείται η απόδοση του μοντέλου στα δεδομένα επικύρωσης. Το συγκεκριμένο μοντέλο του παραδείγματος θα αξιολογείται κάθε 10 επαναλήψεις (iterations) στα δεδομένα επικύρωσης.

'Plots', 'training-progress', ...

Ορίζει την εμφάνιση της προόδου εκπαίδευσης επιτρέποντας την παρακολούθησή της σε πραγματικό χρόνο μέσω διαγραμμάτων κατά την εκτέλεση του κώδικα.

'Verbose', true, ...

Καθορίζει αν θα εμφανίζονται λεπτομερείς πληροφορίες κατά τη διάρκεια της εκπαίδευσης. Ενεργοποιεί δηλαδή την εκτύπωση λεπτομερειών εκπαίδευσης στο command window, διευκολύνοντας την παρακολούθηση της προόδου.

'VerboseFrequency', 5

Καθορίζει πόσο συχνά θα εμφανίζονται οι λεπτομέρειες εκπαίδευσης. Οι λεπτομέρειες εκπαίδευσης του συγκεκριμένου παραδείγματος θα εμφανίζονται κάθε 5 επαναλήψεις (iterations).

Προαιρετικά μπορούν επίσης να χρησιμοποιηθούν οι παρακάτω παράμετροι για μη σταθερό ρυθμό μάθησης:

'LearnRateSchedule', 'piecewise', ...

Ορίζει τον τρόπο με τον οποίο θα αλλάζει ο ρυθμός μάθησης κατά τη διάρκεια της εκπαίδευσης. Με την χρήση της παραμέτρου 'piecewise', ο ρυθμός μάθησης θα αλλάζει σε συγκεκριμένα σημεία κατά τη διάρκεια της εκπαίδευσης, όπως θα καθορίζεται από το 'LearnRateDropPeriod' και το 'LearnRateDropFactor'.

'LearnRateDropPeriod', 5, ...

Καθορίζει μετά από πόσες εποχές θα μειωθεί ο ρυθμός μάθησης. Σε αυτή την περίπτωση, ο ρυθμός μάθησης θα μειώνεται κάθε 5 εποχές, αν το μοντέλο δεν παρουσιάζει βελτίωση.

'LearnRateDropFactor', 0.1, ...

Καθορίζει τον συντελεστή μείωσης του ρυθμού μάθησης. Ο ρυθμός μάθησης θα μειώνεται κατά 90% (δηλαδή, θα πολλαπλασιάζεται με 0.1) κάθε 5 εποχές.

Όλες αυτές οι παράμετροι καθορίζουν τον τρόπο εκπαίδευσης ενός νευρωνικού δικτύου, επηρεάζοντας την ταχύτητα, τη σταθερότητα και την απόδοση της διαδικασίας εκπαίδευσης. Η επιλογή των τιμών για κάθε παράμετρο στα πειράματα που θα ακολουθήσουν έγινε μετά από δοκιμές με διάφορους συνδυασμούς, από τους οποίους διαπιστώθηκε ότι οι συγκεκριμένες τιμές που επιλέχθηκαν οδηγούν στα βέλτιστα αποτελέσματα για το κάθε νευρωνικό δίκτυο.

4.3.5 Εκπαίδευση και Αξιολόγηση του Δικτύου

Η εκπαίδευση του νευρωνικού δικτύου και η αξιολόγηση της απόδοσής του είναι τα πιο σημαντικά βήματα για την ανάπτυξη ενός επιτυχημένου μοντέλου. Το δίκτυο εκπαιδεύεται χρησιμοποιώντας τα προκαθορισμένα σύνολα δεδομένων εκπαίδευσης και τις παραμέτρους εκπαίδευσης που έχουν οριστεί προηγουμένως.

% Train the network

```
net = trainNetwork(trainData, layers, options);
```

Η συνάρτηση `trainNetwork` λαμβάνει τρία ορίσματα:

1. **trainData**: Το σύνολο δεδομένων εκπαίδευσης.
2. **layers**: Την αρχιτεκτονική του νευρωνικού δικτύου με τα επίπεδα που ορίστηκαν.
3. **options**: Τις παραμέτρους εκπαίδευσης που καθορίστηκαν.

Το εκπαιδευμένο δίκτυο αποθηκεύεται στην μεταβλητή `net` ώστε να μπορεί να χρησιμοποιηθεί μετά την ολοκλήρωση της εκπαίδευσής του.

% Evaluate the network on the test set

```
[predictedLabels, scores] = classify(net, testData);
```

Μετά την λήξη της εκπαίδευσης, το δίκτυο αξιολογείται στο σύνολο δεδομένων δοκιμής (`testData`). Η συνάρτηση `classify` χρησιμοποιείται για να προβλέψει τις ετικέτες των εικόνων στα δεδομένα δοκιμής χρησιμοποιώντας το δίκτυο `net`. Αυτή η συνάρτηση επιστρέφει τις προβλεπόμενες ετικέτες (`predictedLabels`) και τις αντίστοιχες βαθμολογίες (`scores`), που δείχνουν την εμπιστοσύνη του δικτύου για την κάθε πρόβλεψη.

% Calculate accuracy

```
accuracy = sum(predictedLabels == trueLabels) / numel(trueLabels)
```

Η ακρίβεια του μοντέλου υπολογίζεται συγκρίνοντας τις προβλεπόμενες ετικέτες με τις πραγματικές ετικέτες (`trueLabels`). Η ακρίβεια είναι ο λόγος των σωστών προβλέψεων προς το συνολικό αριθμό των δειγμάτων στο σύνολο δοκιμής και αποθηκεύεται στην μεταβλητή `accuracy` η οποία στο τέλος εμφανίζεται και στο `command window`.

% Plot Confusion Matrix

```
figure();
```

plotconfusion(trueLabels, predictedLabels)

Τέλος, η απόδοση του μοντέλου οπτικοποιείται μέσω ενός πίνακα σύγχυσης (confusion matrix). Ο πίνακας σύγχυσης παρουσιάζει τον αριθμό των σωστών και λανθασμένων προβλέψεων για κάθε κατηγορία. Αυτό βοηθά στην κατανόηση των επιδόσεων του μοντέλου και στον εντοπισμό των κατηγοριών όπου το μοντέλο μπορεί να χρειάζεται βελτίωση.

ΚΕΦΑΛΑΙΟ 5^ο: Ανάλυση με Feedforward Neural Network

5.1 Feedforward Neural Network (FNN)

Η πρώτη σειρά πειραμάτων θα γίνει χρησιμοποιώντας ένα δίκτυο πρόσθιας τροφοδότησης (FNN). Το δίκτυο αυτό περιλαμβάνει κυρίως πλήρως συνδεδεμένα επίπεδα νευρώνων και επίπεδα ReLU. Στόχος είναι να διαπιστωθεί αν τα αποτελέσματα που θα δώσει είναι ικανοποιητικά, μιας και απαιτεί λιγότερη υπολογιστική δύναμη και χρόνο κατά την εκπαίδευση, συγκριτικά με ένα CNN.

Τα πειράματα που θα ακολουθήσουν θα περιλαμβάνουν:

- 1) εκπαίδευση με τις αρχικές εικόνες **χωρίς επεξεργασία**,
- 2) εκπαίδευση με εικόνες που έχουν υποστεί μετασχηματισμό Fourier **FFT-2D**, εφαρμογή γκαουσιανού χαμηλοπερατού **φίλτρου**, και αντίστροφο μετασχηματισμό Fourier, **IFFT-2D**,
- 3) εκπαίδευση με εικόνες που έχουν προκύψει από την **πρόσθεση** των δύο παραπάνω.

5.2 Αρχιτεκτονική του FNN και Ρυθμίσεις Εκπαίδευσης

Η αρχιτεκτονική του δικτύου είναι αρκετά απλή και αποτελείται από πλήρως συνδεδεμένα επίπεδα με αριθμό νευρώνων, που αρχικά αυξάνεται και στη συνέχεια μειώνεται, και επίπεδα ReLU μετά από κάθε πλήρως συνδεδεμένο επίπεδο.

% FNN Architecture

```
layers = [  
    imageInputLayer([224 224 1])  
    fullyConnectedLayer(128)  
    reluLayer  
    fullyConnectedLayer(256)  
    reluLayer  
    fullyConnectedLayer(512)  
    reluLayer  
    fullyConnectedLayer(1024)  
    reluLayer  
    fullyConnectedLayer(512)  
    reluLayer  
    fullyConnectedLayer(256)  
    reluLayer  
    fullyConnectedLayer(128)  
    reluLayer  
    fullyConnectedLayer(64)  
    reluLayer  
    fullyConnectedLayer(32)  
    reluLayer  
    fullyConnectedLayer(3)  
    softmaxLayer  
    classificationLayer  
];
```

Ο λόγος που επιλέχθηκε η ανωτέρω αρχιτεκτονική για το συγκεκριμένο σύνολο εικόνων που εκτελείται το πείραμα, είναι διότι κατόπιν δοκιμών διαπιστώθηκε ότι:

- Αν τα επίπεδα περιλαμβάνουν μόνο σταδιακή *αύξηση* του αριθμού νευρώνων, το δίκτυο δεν αποδίδει καλά στην εκπαίδευση και συνεπώς δεν έχει καλή ακρίβεια στις προβλέψεις του.

- Αν τα επίπεδα περιλαμβάνουν μόνο σταδιακή μείωση του αριθμού νευρώνων, το δίκτυο μαθαίνει σε υπερβολικά γρήγορο βαθμό τα δεδομένα της εκπαίδευσης και δεν έχει καλή ακρίβεια στην κατηγοριοποίηση των υποσυνόλων validation και testing.

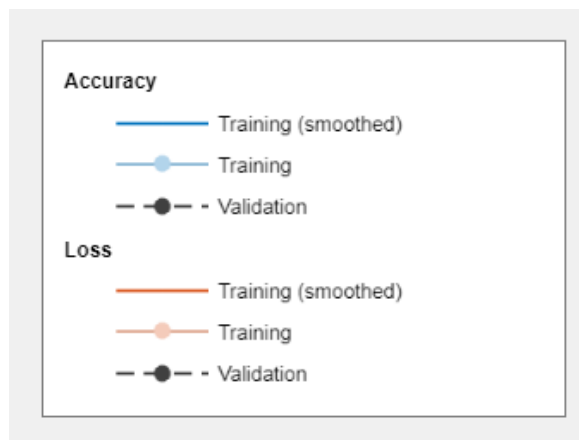
Οι ρυθμίσεις εκπαίδευσης του FNN είναι:

% Training options

```
options = trainingOptions('sgdm', ...  
    'MaxEpochs', 20, ...  
    'MiniBatchSize', 16, ...  
    'Shuffle','every-epoch', ...  
    'InitialLearnRate', 1e-05, ...  
    'ValidationData', valData, ...  
    'ValidationFrequency', 10, ...  
    'Plots', 'training-progress', ...  
    'Verbose', true, ...  
    'VerboseFrequency', 5);
```

Το δίκτυο εκπαιδεύεται για 20 εποχές συνολικά με σταθερό ρυθμό μάθησης $1e-05$ (δηλ. 10^{-5}) και Mini Batch Size 16. Η επιλογή αυτών των παραμέτρων έγινε κατόπιν δοκιμών, όπου διαπιστώθηκε ότι έδιναν το πιο ικανοποιητικό αποτέλεσμα.

Κατά την διάρκεια της εκπαίδευσης του νευρωνικού δικτύου, δημιουργείται σε πραγματικό χρόνο ένα διάγραμμα που απεικονίζει την πρόοδο της εκπαίδευσης (όπως αυτό της Εικόνας 5.2). Σε αυτό το διάγραμμα φαίνονται ορισμένα βασικά στοιχεία που θα χρειαστεί να αναλυθούν πριν την ερμηνεία των αποτελεσμάτων.



Εικόνα 5.1 Στοιχεία γραφημάτων εκπαίδευσης

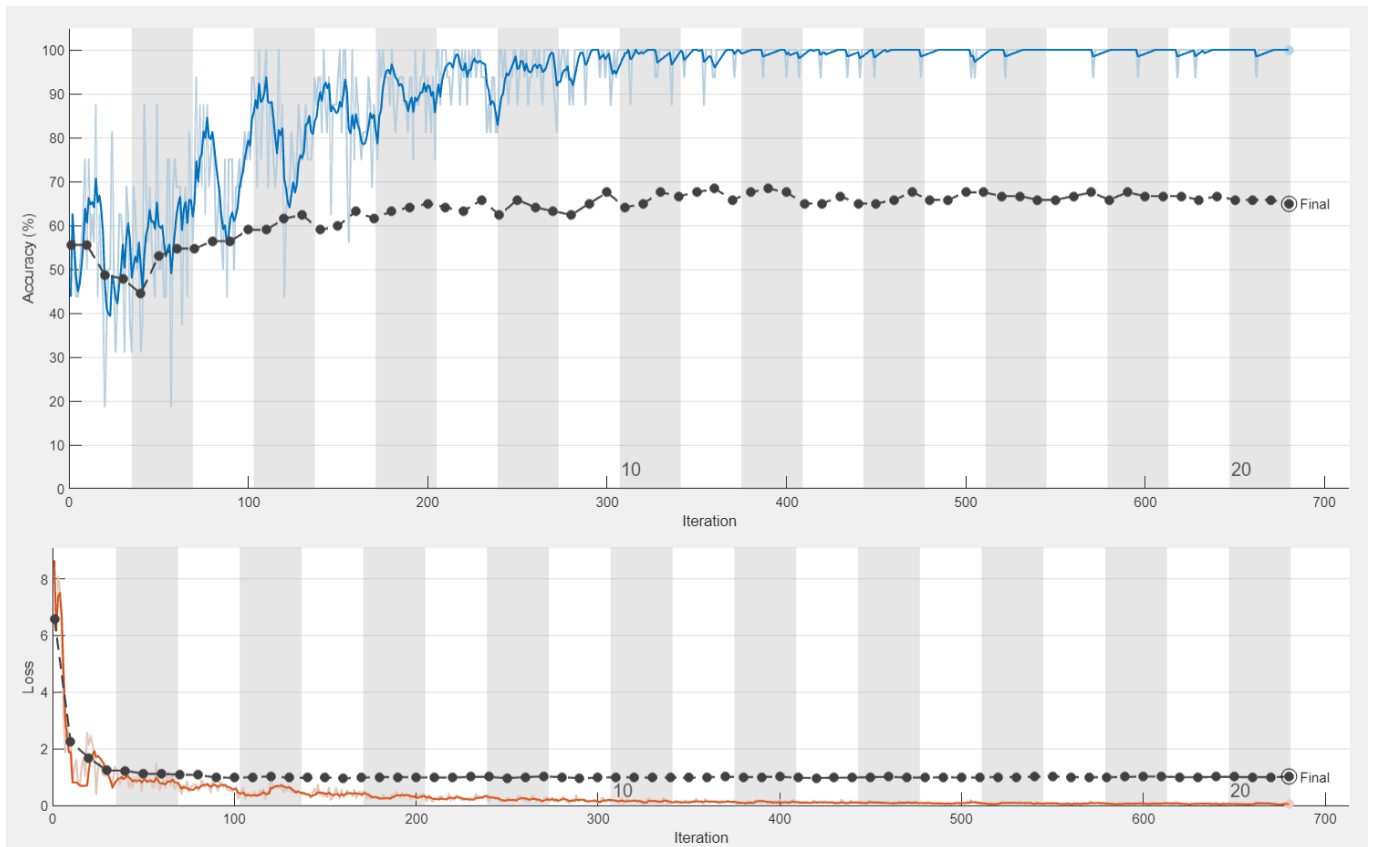
Όπως φαίνεται και στην Εικόνα 5.1, στα δύο γραφήματα της Ακρίβειας (Accuracy) και Απώλειας (Loss) που θα αναλυθούν με μεγαλύτερη λεπτομέρεια κατά την εκτέλεση των πειραμάτων, η μπλε γραμμή απεικονίζει την ακρίβεια της εκπαίδευσης (ομαλή), ενώ η ανοιχτή μπλε γραμμή τις πραγματικές τιμές της ακρίβειας της εκπαίδευσης. Τέλος, με μαύρη διακεκομμένη γραμμή απεικονίζεται η ακρίβεια της επικύρωσης (validation).

Αντίστοιχα, για το διάγραμμα της Απώλειας με πορτοκαλί γραμμή απεικονίζεται η απώλεια της εκπαίδευσης (ομαλή), με ανοιχτό πορτοκαλί οι πραγματικές τιμές της απώλειας της εκπαίδευσης και με μαύρη διακεκομμένη γραμμή η απώλεια της επικύρωσης (validation).

Στην Εικόνα 5.2, απεικονίζεται στο επάνω πεδίο το γράφημα της Ακρίβειας (ποσοστό %) ως προς τον αριθμό των επαναλήψεων, ενώ στο κάτω πεδίο το γράφημα της Απώλειας (σφάλμα πρόβλεψης) ως προς τον αριθμό των επαναλήψεων.

5.3 Εικόνες χωρίς επεξεργασία

Στο πρώτο πείραμα, δίνεται ως είσοδος στο FNN το dataset που περιέχει τις εικόνες στην αρχική τους μορφή. Δηλαδή η μόνη επεξεργασία που γίνεται στις εικόνες είναι τα απαραίτητα βήματα για την εισαγωγή τους στο δίκτυο, όπως αναλύθηκε στο Κεφάλαιο 4.3.



Εικόνα 5.2 Διαγράμματα Ακρίβειας και Απώλειας της εκπαίδευσης και επικύρωσης του δικτύου FNN με τις αρχικές εικόνες

5.3.1 Διάγραμμα Ακρίβειας

Η ακρίβεια της εκπαίδευσης παρουσιάζει μια σταθερή ανοδική τάση για τις πρώτες 8 εποχές της εκπαίδευσης, ενώ στη συνέχεια πλησιάζει και φτάνει το 100%. Αυτό από την μία πλευρά αυτό δείχνει ότι το μοντέλο μαθαίνει αρκετά καλά τα χαρακτηριστικά των δεδομένων εκπαίδευσης, αλλά από την άλλη πλευρά σημαίνει ότι υπερ-εκπαιδύεται σε αυτά (overfitting).

Η ακρίβεια επικύρωσης δεν παρουσιάζει την ίδια ανοδική τάση και σταθεροποιείται λίγο παρακάτω από το 70%, υποδηλώνοντας ότι το δεν βελτιώνεται ιδιαίτερα, κυρίως από την 8^η εποχή και έπειτα.

5.3.2 Διάγραμμα Απώλειας

Η απώλεια της εκπαίδευσης μειώνεται πολύ γρήγορα, φτάνοντας σε σημεία πολύ κοντά στο 0 ήδη από τις πρώτες εποχές.

Η απώλεια της επικύρωσης παρουσιάζει την ίδια απότομη μείωση και παραμένει σταθερή σε υψηλότερο επίπεδο (περίπου στο 1).

5.3.3 Μετρήσεις Απόδοσης

Results	
Validation accuracy:	64.96%
Training finished:	Max epochs completed
Training Time	
Start time:	19-Jun-2024 19:34:23
Elapsed time:	2 min 53 sec
Training Cycle	
Epoch:	20 of 20
Iteration:	680 of 680
Iterations per epoch:	34
Maximum iterations:	680
Validation	
Frequency:	10 iterations
Other Information	
Hardware resource:	Single CPU
Learning rate schedule:	Constant
Learning rate:	1e-05

Εικόνα 5.3 Απόδοση δικτύου FNN κατόπιν εκπαίδευσης με τις αρχικές εικόνες

Η τελική ακρίβεια επικύρωσης σταθεροποιείται στο **64,96%**. Η τιμή αυτή δεν είναι πολύ ικανοποιητική, καθώς δείχνει ότι το FNN δεν έχει τόσο καλή δυνατότητα να αναγνωρίσει και να γενικεύσει τα χαρακτηριστικά των εικόνων.

Η εκπαίδευσή του είναι αρκετά γρήγορη, καθώς για 680 επαναλήψεις στις 20 εποχές, χρειάστηκε περίπου 3 λεπτά. Το δίκτυο εκπαιδεύτηκε με σταθερό ρυθμό μάθησης 1e-05. Επομένως δεν απαιτεί ιδιαίτερη υπολογιστική ισχύ.

5.3.4 Confusion Matrix

		Target Class				
		benign	malignant	normal		
Output Class	benign	62 53.0%	9 7.7%	11 9.4%	75.6%	24.4%
	malignant	4 3.4%	20 17.1%	1 0.9%	80.0%	20.0%
	normal	0 0.0%	2 1.7%	8 6.8%	80.0%	20.0%
		93.9% 6.1%	64.5% 35.5%	40.0% 60.0%	76.9%	23.1%

Εικόνα 5.4 Confusion Matrix του FNN κατόπιν εκπαίδευσης με τις αρχικές εικόνες

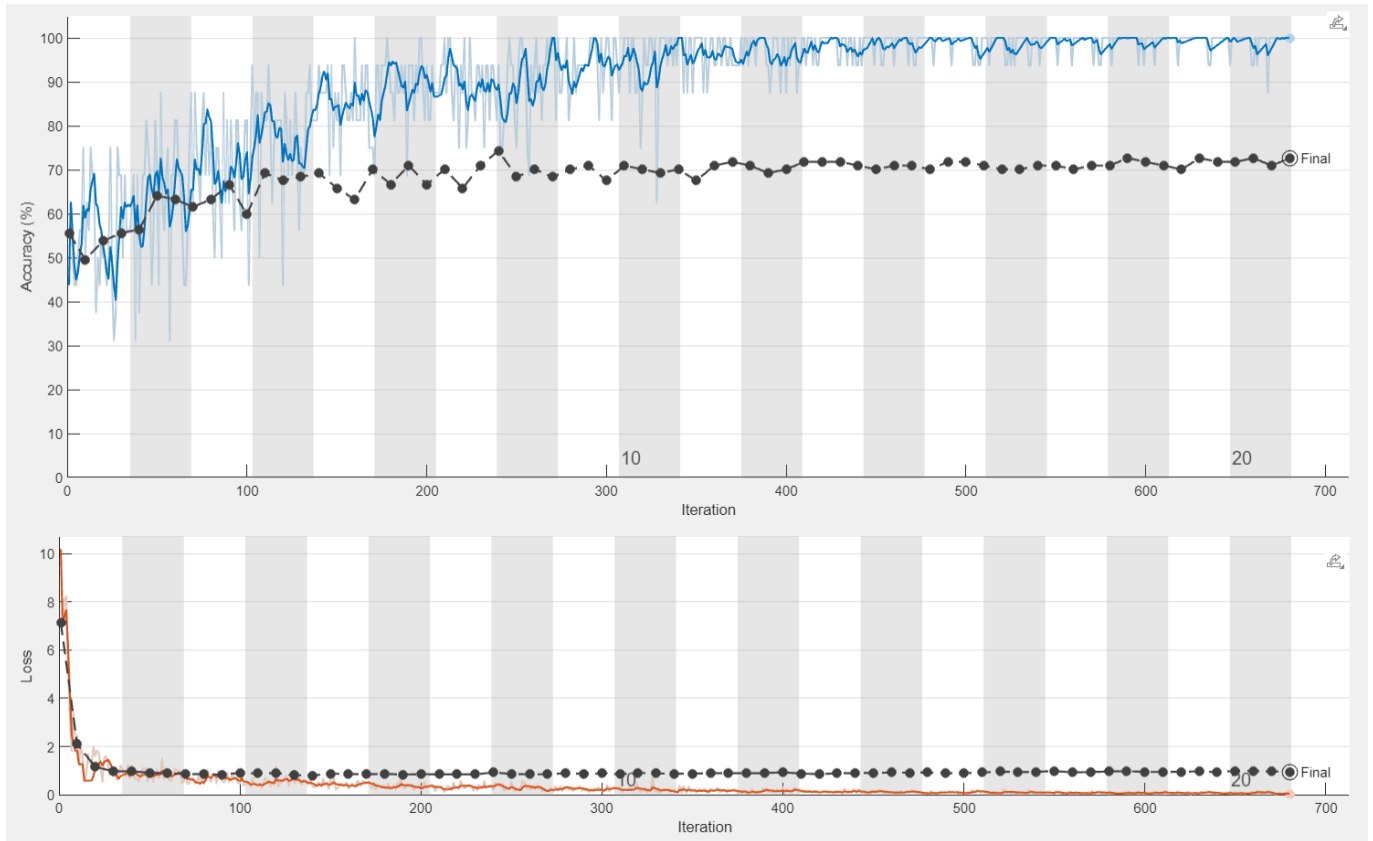
Το συνολικό ποσοστό επιτυχίας του FNN είναι **76,9%**, έχει δηλαδή μέτρια απόδοση.

- **Benign:** Το μοντέλο φαίνεται να αναγνωρίζει με πολύ μεγάλη ακρίβεια (93,9%) τα παραδείγματα της κατηγορίας benign. Ωστόσο έχουν γίνει και κάποιες λανθασμένες προβλέψεις σε αυτήν την κατηγορία, από τις οποίες οι περισσότερες (11) στην πραγματικότητα ανήκαν στην κατηγορία normal.
- **Malignant:** Η κατηγορία malignant παρουσιάζει μέτρια απόδοση, με το μοντέλο να αναγνωρίζει ένα σχετικά καλό ποσοστό κακοήθων δειγμάτων, αλλά να συγχίζεται σε κάποιες περιπτώσεις και λανθασμένα να τις κατηγοριοποιεί ως καλοήθεις όγκους. Από τις λάθος προβλέψεις που έγιναν, οι περισσότερες από αυτές (9) προβλέφθηκαν ως καλοήθειες, ενώ ένα μικρό μέρος (2) ως φυσιολογικές εικόνες.
- **Normal:** Η κατηγορία normal παρουσιάζει τη χαμηλότερη απόδοση, γεγονός που οφείλεται ενδεχομένως στα λιγότερα δεδομένα εκπαίδευσης για τις φυσιολογικές εικόνες. Ως αποτέλεσμα το μοντέλο δυσκολεύεται περισσότερο να αναγνωρίσει αυτές τις περιπτώσεις. Το ποσοστό λανθασμένων προβλέψεων (60%) για αυτήν την κατηγορία είναι μεγαλύτερο από το ποσοστό των σωστών (40%), με πολλές από αυτές να έχουν προβλεφθεί ότι ανήκουν στην κατηγορία benign.

Η συνολική απόδοση του FNN μέτρια. Το FNN δεν φαίνεται να έχει την ικανότητα να εκμεταλλευτεί μοτίβα και πληροφορίες από τις εικόνες ή να αναγνωρίσει την διαφορά μεταξύ των σημαντικότερων χαρακτηριστικών των εικόνων.

5.4 Εικόνες με FFT-2D, συχνοτικό φίλτρο και IFFT-2D

Αφού το FNN φαίνεται να παρουσιάζει δυσκολία στην αναγνώριση των στοιχείων που διαφοροποιούν τις 3 κατηγορίες μεταξύ τους, θα γίνει δοκιμή με τις εικόνες που έχουν φιλτραριστεί στο πεδίο της συχνότητας με γκαουσιανό χαμηλοπερατό φίλτρο. Στο δίκτυο δίνονται αυτή τη φορά οι εικόνες αυτές, οι οποίες έχουν επανέλθει στο πεδίο του χώρου μέσω αντίστροφου μετασχηματισμού Fourier, IFFT-2D.



Εικόνα 5.5 Διαγράμματα Ακρίβειας και Απώλειας της εκπαίδευσης και επικύρωσης του δικτύου FNN με τις φιλτραρισμένες εικόνες

5.4.1 Διάγραμμα Ακρίβειας

Η ακρίβεια της εκπαίδευσης φαίνεται να έχει παρόμοια μορφή με αυτή του προηγούμενου πειράματος (Εικόνα 5.2). Έχει και σε αυτήν την περίπτωση ανοδική τάση και δείχνει να μαθαίνει σε μεγάλο βαθμό τα δεδομένα της εκπαίδευσης.

Η ακρίβεια της επικύρωσης παρουσιάζει επίσης παρόμοια μορφή με την προηγούμενη. Ειδικότερα από την 9^η εποχή της εκπαίδευσης του δικτύου και έπειτα, δεν φαίνεται να βελτιώνεται ιδιαίτερα που σημαίνει ότι το μοντέλο πιθανόν να έχει μάθει λεπτομερή χαρακτηριστικά από τα δεδομένα της εκπαίδευσης, τα οποία δεν υπάρχουν στα δεδομένα επικύρωσης.

5.4.2 Διάγραμμα Απώλειας

Όπως και προηγουμένως, η απώλεια της εκπαίδευσης μειώνεται αρκετά απότομα και πλησιάζει το 0, που σημαίνει ότι το δίκτυο έχει μάθει πολύ καλά τα δεδομένα της εκπαίδευσης.

Η απώλεια της επικύρωσης εξακολουθεί να μένει σταθερή γύρω από το 1, χωρίς ιδιαίτερη μείωση από την αρχή της εκπαίδευσης.

5.4.3 Μετρήσεις Απόδοσης

Results	
Validation accuracy:	72.65%
Training finished:	Max epochs completed
Training Time	
Start time:	19-Jun-2024 19:30:01
Elapsed time:	2 min 4 sec
Training Cycle	
Epoch:	20 of 20
Iteration:	680 of 680
Iterations per epoch:	34
Maximum iterations:	680
Validation	
Frequency:	10 iterations
Other Information	
Hardware resource:	Single CPU
Learning rate schedule:	Constant
Learning rate:	1e-05

Εικόνα 5.6 Απόδοση δικτύου FNN κατόπιν εκπαίδευσης με τις φιλτραρισμένες εικόνες

Η ακρίβεια επικύρωσης τελικά σταθεροποιείται στο **72,65%**. Αν και είναι λίγο υψηλότερη από το προηγούμενο πείραμα (Εικόνα 5.5) εξακολουθεί να έχει περιθώριο βελτίωσης.

Το δίκτυο φαίνεται να χρειάστηκε 2 λεπτά για την ολοκλήρωση της εκπαίδευσης (680 επαναλήψεις, 20 εποχές, με ρυθμό μάθησης 1e-05). Το πείραμα αυτό χρειάστηκε ελαφρώς λιγότερη υπολογιστική ισχύ, αφού ολοκληρώθηκε ένα λεπτό νωρίτερα, που σημαίνει ότι η απλοποίηση των δεδομένων με την χρήση φίλτρου ίσως να βοήθησε στην πιο εύκολη αναγνώριση των σημαντικών χαρακτηριστικών.

5.4.4 Confusion Matrix

	benign	malignant	normal	
benign	55 47.0%	12 10.3%	8 6.8%	73.3% 26.7%
malignant	6 5.1%	19 16.2%	0 0.0%	76.0% 24.0%
normal	5 4.3%	0 0.0%	12 10.3%	70.6% 29.4%
	83.3% 16.7%	61.3% 38.7%	60.0% 40.0%	73.5% 26.5%
	benign	malignant	normal	

Εικόνα 5.7 Confusion Matrix του FNN κατόπιν εκπαίδευσης με τις φιλτραρισμένες εικόνες

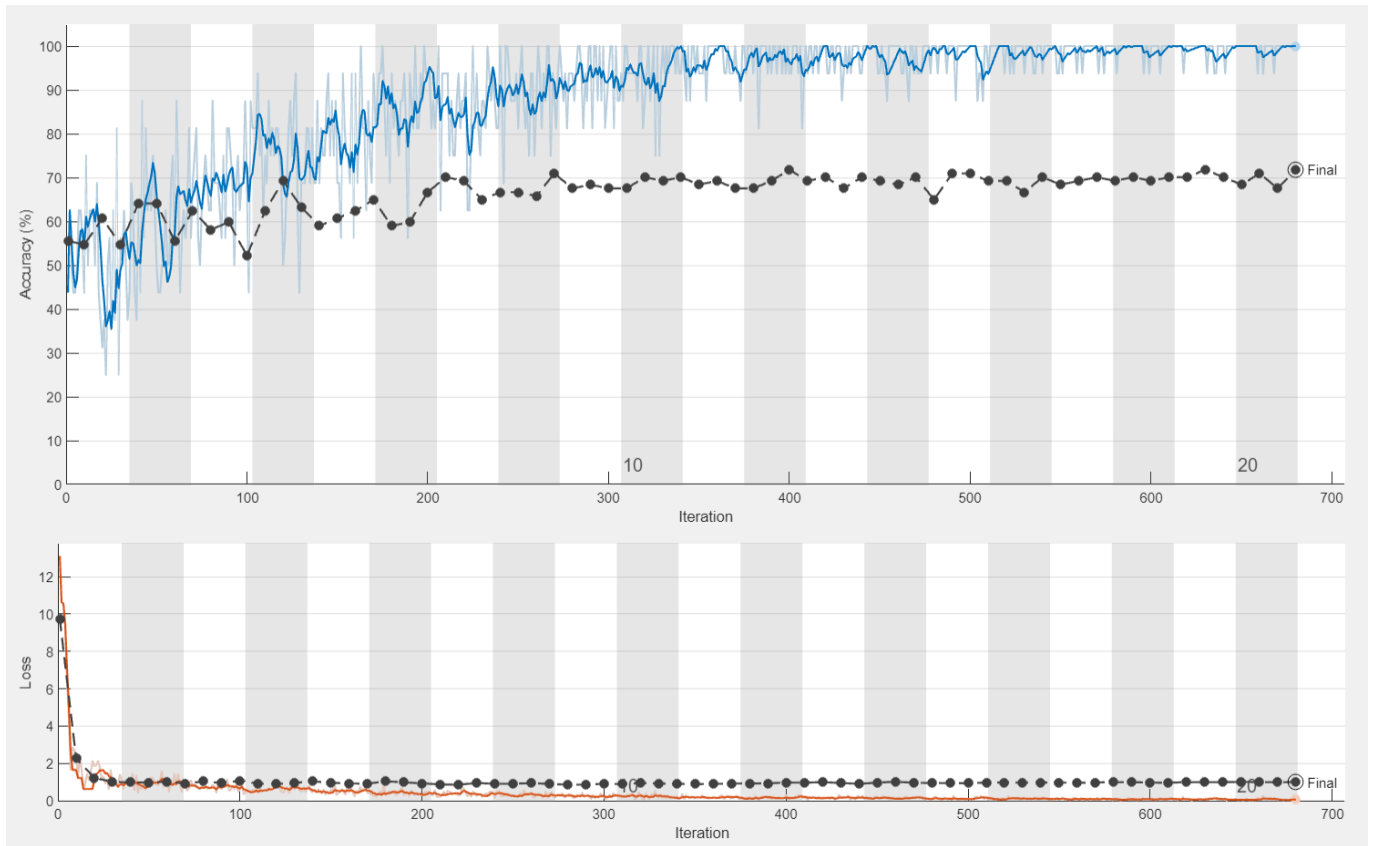
Το συνολικό ποσοστό επιτυχίας κατά την εκπαίδευση με τις φιλτραρισμένες εικόνες μειώθηκε στο **73,5%**.

- **Benign:** Η κλάση benign φαίνεται και πάλι να έχει την καλύτερη απόδοση, γεγονός που οφείλεται στα περισσότερα δείγματα από αυτήν. Για ακόμα μια φορά φαίνεται ότι κάποιες περιπτώσεις που ανήκουν σε άλλες κλάσεις, έχουν λανθασμένα προβλεφθεί ως καλοήθεις όγκοι. Αυτή τη φορά μάλιστα, περισσότερα δείγματα κακοηθειών προβλέφθηκαν ως καλοήθειες, ενώ τα δείγματα φυσιολογικών εικόνων που αναγνωρίστηκαν ως καλοήθεις όγκοι φαίνεται να μειώθηκαν από 11 (Κεφάλαιο 5.3.4) σε 8.
- **Malignant:** Το FNN φαίνεται να μπορεί να διακρίνει επιτυχώς τις διαφορές μεταξύ κακοηθών και φυσιολογικών περιπτώσεων, αφού καμία εικόνα της μίας κλάσης δεν προβλέφθηκε να ανήκει στην άλλη. Ωστόσο, φαίνεται να υπάρχει μεγαλύτερη δυσκολία στην διάκριση μεταξύ καλοήθειας και κακοήθειας συγκριτικά με προηγούμενως.
- **Normal:** Η κλάση normal έχει καλύτερη απόδοση. Περισσότερα από τα μισά δείγματα αυτής της κατηγορίας προβλέφθηκαν σωστά, ενώ οι λάθος προβλέψεις έγιναν όλες στην κλάση benign. Το δίκτυο μπορεί επιτυχώς να διακρίνει μεταξύ κακοήθειας και φυσιολογικού ιστού. Υπάρχει ωστόσο μια σχετικά έντονη τάση να μπερδεύονται κάποιες περιπτώσεις καλοήθειας ως φυσιολογικές και αντίστροφα.

Οι προβλέψεις του συγκεκριμένου δικτύου δεν είναι πολύ αξιόπιστες. Φαίνεται να παρουσιάζει έντονες προκλήσεις στη διάκριση μεταξύ καλοηθών και κακοήθων όγκων, καθώς και στη διάκριση μεταξύ φυσιολογικού ιστού και καλοήθειας. Παρά τις δυσκολίες αυτές, το δίκτυο καταφέρνει να διαχωρίζει αποτελεσματικά τους κακοήθεις από τους φυσιολογικούς ιστούς, γεγονός που είναι αρκετά σημαντικό στην ιατρική διάγνωση.

5.5 Εικόνες από άθροιση των αρχικών και των φιλτραρισμένων

Ως τελευταίο πείραμα σε αυτήν την αρχιτεκτονική, θα γίνει δοκιμή με τις εικόνες που δημιουργήθηκαν αθροίζοντας τις αρχικές με τις φιλτραρισμένες, όπως εξηγήθηκε αναλυτικά στο Κεφάλαιο 4.3.2.



Εικόνα 5.8 Διαγράμματα Ακρίβειας και Απώλειας της εκπαίδευσης και επικύρωσης του δικτύου FNN με τις αθροισμένες εικόνες

5.5.1 Διάγραμμα Ακρίβειας

Η ακρίβεια της εκπαίδευσης φαίνεται και πάλι να ανεβαίνει και να φτάνει μεταξύ 95% και 100% περίπου στα μισά της εκπαίδευσης του δικτύου.

Η ακρίβεια της επικύρωσης δείχνει να έχει σταθεροποιηθεί αρκετά νωρίς. Αυτό δείχνει καθαρά το φαινόμενο της υπερεκπαίδευσης (overfitting), όπου το δίκτυο έχει μάθει πολύ συγκεκριμένα χαρακτηριστικά των εικόνων που χρησιμοποιήθηκαν για την εκπαίδευση, τα οποία δεν παρουσιάζονται απαραίτητα στα δεδομένα επικύρωσης. Όπως και στα προηγούμενα πειράματα, επειδή το σύνολο των δεδομένων είναι μικρό, η ακρίβεια επικύρωσης εμφανίζει χαμηλότερο ποσοστό λόγω της μικρότερης ποικιλίας δεδομένων.

5.5.2 Διάγραμμα Απώλειας

Η απώλεια της εκπαίδευσης έχει αντίστοιχη μορφή με αυτή των δύο πειραμάτων που προηγήθηκαν. Πλησιάζει δηλαδή πολύ κοντά στο 0 και άρα τα δεδομένα της εκπαίδευσης είναι πλέον αρκετά γνώριμα για το νευρωνικό δίκτυο.

Η απώλεια της επικύρωσης δείχνει επίσης να έχει σταθερή πορεία γύρω από το 1. Η πολύ χαμηλή απώλεια εκπαίδευσης σε συνδυασμό με την υψηλότερη και σχετικά σταθερή απώλεια επικύρωσης είναι ενδείξεις υπερεκπαίδευσης (overfitting).

Συνεπώς, η αυξημένη φωτεινότητα και αντίθεση στις εικόνες μπορεί να δημιουργεί πολύ συγκεκριμένα χαρακτηριστικά που το μοντέλο μαθαίνει καλά, αλλά αυτά τα χαρακτηριστικά δεν είναι γενικεύσιμα.

5.5.3 Μετρήσεις Απόδοσης

Results	
Validation accuracy:	71.79%
Training finished:	Max epochs completed
Training Time	
Start time:	19-Jun-2024 19:25:08
Elapsed time:	2 min 24 sec
Training Cycle	
Epoch:	20 of 20
Iteration:	680 of 680
Iterations per epoch:	34
Maximum iterations:	680
Validation	
Frequency:	10 iterations
Other Information	
Hardware resource:	Single CPU
Learning rate schedule:	Constant
Learning rate:	1e-05

Εικόνα 5.9 Απόδοση δικτύου FNN κατόπιν εκπαίδευσης με τις αθροισμένες εικόνες

Η ακρίβεια επικύρωσης λαμβάνει τελική τιμή **71,79%** μετά το πέρας της εκπαίδευσης. Γενικά μπορεί να θεωρηθεί μέτρια προς καλή, αλλά αναμένεται να εξεταστούν και τα αποτελέσματα από το υποσύνολο της δοκιμής για την εξαγωγή συμπερασμάτων.

Ο χρόνος εκπαίδευσης είναι σχεδόν 2,5 λεπτά (680 επαναλήψεις, 20 εποχές, με ρυθμό μάθησης 1e-05). Επομένως το FNN εκπαιδύεται αρκετά γρήγορα και σε αυτό το dataset εικόνων.

5.5.4 Confusion Matrix

		Target Class				
		benign	malignant	normal		
Output Class	benign	59 50.4%	9 7.7%	8 6.8%	77.6%	22.4%
	malignant	2 1.7%	21 17.9%	1 0.9%	87.5%	12.5%
	normal	5 4.3%	1 0.9%	11 9.4%	64.7%	35.3%
		89.4%	67.7%	55.0%	77.8%	22.2%
		10.6%	32.3%	45.0%		

Εικόνα 5.10 Confusion Matrix του FNN κατόπιν εκπαίδευσης με τις αθροισμένες εικόνες

Το συνολικό ποσοστό επιτυχίας του μοντέλου αυξήθηκε στο **77,8%**, παρουσιάζοντας μια γενική βελτίωση σε σχέση με προηγούμενες προσπάθειες. Παρ' όλα αυτά, εξακολουθούν να υπάρχουν σημαντικά ποσοστά λανθασμένων κατηγοριοποιήσεων, που σημαίνει ότι υπάρχει η ανάγκη για περαιτέρω βελτιώσεις.

- **Benign:** Η κλάση benign φαίνεται να παρουσιάζει το μεγαλύτερο ποσοστό επιτυχίας καθώς το 89,4% των δειγμάτων της προβλέφθηκαν σωστά. Ωστόσο, όσον αφορά το ποσοστό των προβλέψεων, το δίκτυο φαίνεται να προβλέπει πολύ πιο συχνά τις εικόνες ως benign σε αντίθεση με τις άλλες δύο κατηγορίες. Συνολικά 17 εικόνες από άλλες κατηγορίες αναγνωρίστηκαν ψευδώς ως καλοήθεις όγκοι.
- **Malignant:** Οι προβλέψεις για την κλάση malignant είναι αρκετά καλές, με το 87,5 % αυτών να είναι σωστές. Από τα πραγματικά δείγματα κακοήθων όγκων, οι 9 κατηγοριοποιήθηκαν ως καλοήθεις ενώ μόνο 1 από αυτούς ως φυσιολογικός ιστός.
- **Normal:** Οι σωστές προβλέψεις για αυτήν την κατηγορία δεν φαίνεται να βελτιώθηκαν. Λόγω της μεγαλύτερης έκθεσης του δικτύου σε δείγματα της κλάσης benign, σχεδόν οι μισές εικόνες που δόθηκαν στο δίκτυο για κατηγοριοποίηση, κατατάχθηκαν λανθασμένα ως καλοήθεις όγκοι.

Παρά τη βελτίωση του συνολικού ποσοστού επιτυχίας, το FNN εξακολουθεί να αντιμετωπίζει δυσκολίες στη διάκριση μεταξύ καλοήθων και κακοήθων όγκων, καθώς και στη σωστή κατηγοριοποίηση φυσιολογικού ιστού. Η προσθήκη φωτεινότητας και αντίθεσης γύρω από τους όγκους μπορεί να ενισχύει την ανίχνευση καλοήθων όγκων, αλλά ταυτόχρονα φαίνεται να προκαλεί σύγχυση στην κατηγοριοποίηση φυσιολογικών και κακοήθων περιοχών.

ΚΕΦΑΛΑΙΟ 6^ο: Ανάλυση με Convolutional Neural Network

6.1 Convolutional Neural Network (CNN)

Εφόσον η απόδοση του FNN δεν ήταν πολύ ικανοποιητική, σε αυτό το κεφάλαιο θα επαναληφθούν τα πειράματα με χρήση δύο ξεχωριστών CNN, διαφορετικής αρχιτεκτονικής το κάθε ένα.

Θα εξεταστούν αρχικά τα διάφορα επίπεδα του δικτύου, οι παράμετροι εκπαίδευσης, καθώς και οι τεχνικές βελτιστοποίησης που εφαρμόζονται για την επίτευξη των βέλτιστων αποτελεσμάτων και στη συνέχεια θα ακολουθήσουν οι πειραματικές διαδικασίες, ώστε να προκύψουν συγκρίσιμα αποτελέσματα.

Θα ακολουθήσουν όπως και στο προηγούμενο κεφάλαιο τρεις διαφορετικές προσεγγίσεις κατά την εκπαίδευση των δύο αυτών CNN:

- 1) εκπαίδευση με τις αρχικές εικόνες **χωρίς επεξεργασία**,
- 2) εκπαίδευση με εικόνες που έχουν υποστεί μετασχηματισμό **FFT-2D**, εφαρμογή γκαουσιανού χαμηλοπερατού **φίλτρου**, και αντίστροφο μετασχηματισμό **IFFT-2D**,
- 3) εκπαίδευση με εικόνες-αθροίσματα, που έχουν προκύψει από την **πρόσθεση** των δύο παραπάνω.

Και στις δύο αρχιτεκτονικές CNN θα εφαρμοστούν βάρη στις τρεις κλάσεις, εισάγοντας την παράμετρο `ClassWeights` στο τελευταίο επίπεδο του κάθε νευρωνικού δικτύου, δηλαδή το επίπεδο της κατηγοριοποίησης.

Τα βάρη θα είναι αντιστρόφως ανάλογα με τον αριθμό των δειγμάτων σε κάθε κατηγορία, δηλαδή η κατηγορία `normal` για παράδειγμα, που έχει και τα λιγότερα δείγματα, θα έχει και το μεγαλύτερο βάρος.

Με αυτόν τον τρόπο, οι κατηγορίες με λιγότερα δείγματα θα έχουν μεγαλύτερη επίδραση κατά την εκπαίδευση του μοντέλου έτσι ώστε να ισορροπήσει η επίδρασή τους και να αντιμετωπιστεί η «προκατάληψη» του δικτύου προς την κατηγορία `benign`, την οποία τείνει να προβλέπει πολύ πιο συχνά λόγω της μεγαλύτερης «εμπειρίας» πάνω σε αυτή, εφόσον έχει εκπαιδευτεί περισσότερο σε παραδείγματα από αυτήν την κατηγορία.

6.2 CNN – 1^η Αρχιτεκτονική

Η πρώτη αρχιτεκτονική CNN, περιλαμβάνει 3 συνελκτικά επίπεδα σε συνδυασμό με επίπεδα κανονικοποίησης, επίπεδα ReLU και επίπεδα υποδειγματοληψίας, τα οποία παρουσιάστηκαν αναλυτικά στο Κεφάλαιο 4.3.3:

% CNN architecture

```
layers = [  
    imageInputLayer([224 224 1])  
  
    convolution2dLayer(3,32, 'Padding', 'same')  
    batchNormalizationLayer  
    reluLayer  
  
    maxPooling2dLayer(2, 'Stride', 2)  
  
    convolution2dLayer(3,64, 'Padding', 'same')  
    batchNormalizationLayer  
    reluLayer
```

```
maxPooling2dLayer(2, 'Stride', 2)

convolution2dLayer(3, 128, 'Padding', 'same')
batchNormalizationLayer
reluLayer

fullyConnectedLayer(3)
softmaxLayer
classificationLayer('Classes', categorical({'benign', 'malignant', 'normal'}), ...
                    'ClassWeights', [1, 2.08, 3.29])
];
```

Στο τελευταίο επίπεδο του δικτύου, έχουν εφαρμοστεί τα βάρη:

- 1 για την κλάση benign,
- 2,08 για την κλάση malignant και
- 3,29 για την κλάση normal.

Αφού υπάρχουν 437 δείγματα για την κατηγορία benign, 210 δείγματα για την κατηγορία malignant και 133 δείγματα για την κατηγορία normal, τα βάρη υπολογίζονται ως εξής:

1) Υπολογισμός Συνολικού αριθμού δειγμάτων:

$$\text{Συνολικά Δείγματα} = 437 + 210 + 133 = 780$$

2) Ποσοστό δειγμάτων σε κάθε κατηγορία:

$$\text{benign}(\%) = \frac{437}{780} \cong 0,5602 = 56,02\%$$

$$\text{malignant}(\%) = \frac{210}{780} \cong 0,2692 = 26,92\%$$

$$\text{normal}(\%) = \frac{133}{780} \cong 0,1705 = 17,05\%$$

3) Υπολογισμός βαρών ως τα αντιστρόφως ανάλογα ποσοστά:

$$\text{benign weight} = \frac{1}{0,5602} \cong 1,78$$

$$\text{malignant weight} = \frac{1}{0,2692} \cong 3,71$$

$$\text{normal weight} = \frac{1}{0,1705} \cong 5,86$$

4) Κανονικοποίηση των βαρών:

Για λόγους απλότητας και αποτελεσματικότητας, κανονικοποιούνται τα βάρη έτσι ώστε το βάρος της κατηγορίας με τον μεγαλύτερο αριθμό δειγμάτων (benign) να είναι 1, και υπόλοιπα βάρη να είναι αναλογικά.

$$\text{benign weight} = \frac{1,78}{1,78} = 1$$

$$\text{malignant weight} = \frac{3,71}{1,78} \cong 2,08$$

$$\text{normal weight} = \frac{5,86}{1,78} \cong 3,29$$

Οι ρυθμίσεις του δικτύου είναι παρόμοιες με αυτές που χρησιμοποιήθηκαν και στο FNN στο Κεφάλαιο 5.2

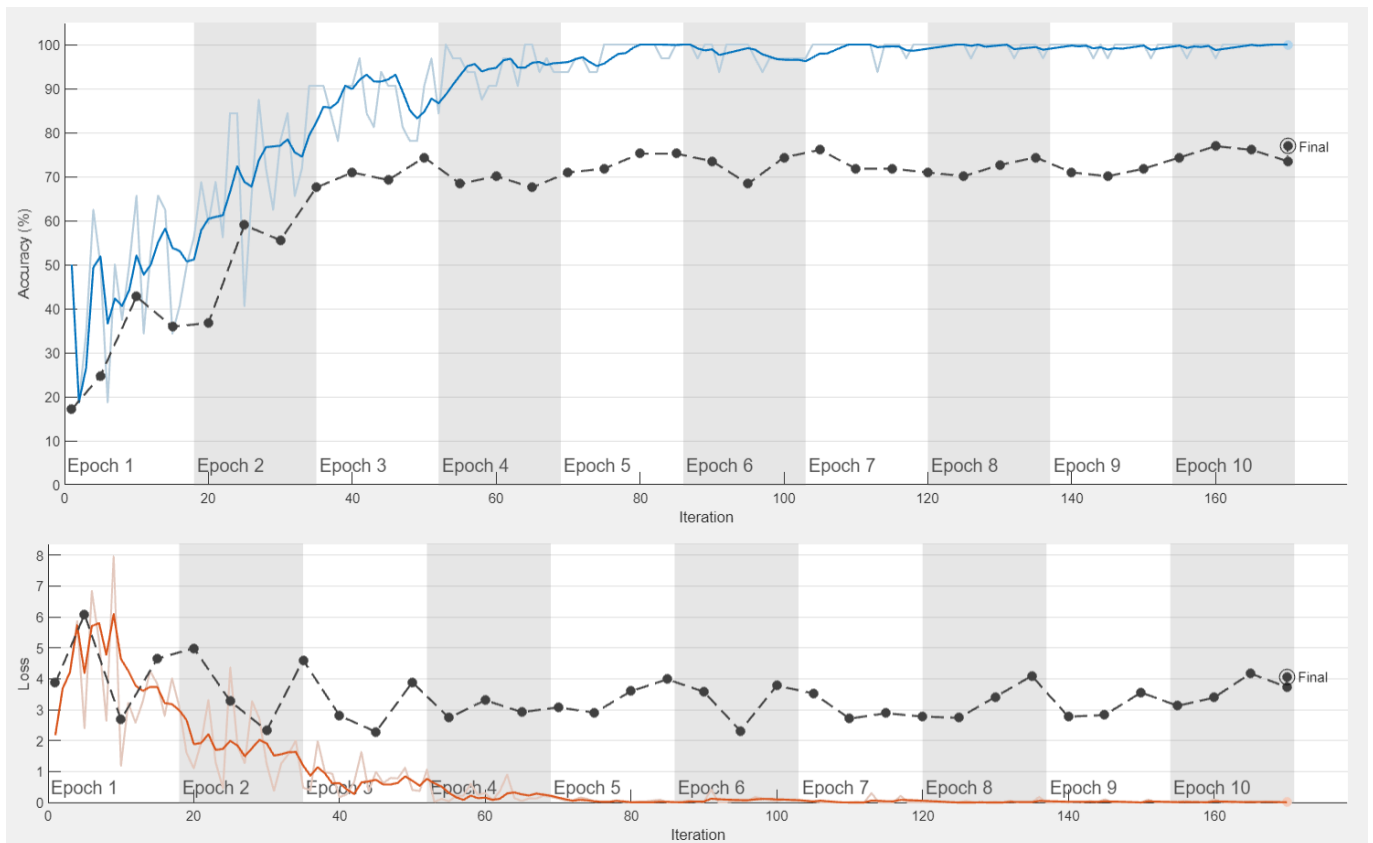
% Training options

```
options = trainingOptions('sgdm', ...  
    'MaxEpochs', 10, ...  
    'MiniBatchSize', 32, ...  
    'Shuffle','every-epoch', ...  
    'InitialLearnRate', 1e-04, ...  
    'ValidationData', valData, ...  
    'ValidationFrequency', 5, ...  
    'Plots', 'training-progress', ...  
    'Verbose', true, ...  
    'VerboseFrequency', 5);
```

Συγκεκριμένα, το δίκτυο εκπαιδεύεται για 10 εποχές αντί για 20, με σταθερό ρυθμό μάθησης 1e-04 και Mini-Batch Size ίσο με 32 αντί για 16.

6.2.1 Εικόνες χωρίς επεξεργασία

Το πρώτο πείραμα περιλαμβάνει τις εικόνες του dataset όπως ακριβώς είναι, χωρίς καμία επιπλέον επεξεργασία εκτός από τα απαραίτητα βήματα για την εισαγωγή των εικόνων στο δίκτυο.



Εικόνα 6.1 Διαγράμματα Ακρίβειας και Απώλειας της εκπαίδευσης και επικύρωσης του δικτύου CNN #1 με τις αρχικές εικόνες

6.2.1.1 Διάγραμμα Ακρίβειας

Η ακρίβεια της εκπαίδευσης δείχνει να αυξάνεται κατά τη διάρκεια των πρώτων 3 – 4 εποχών, και έπειτα σταθεροποιείται σε πολύ υψηλό ποσοστό φτάνοντας τελικά στο 100%. Αυτό δείχνει ότι το μοντέλο μαθαίνει και βελτιώνει την απόδοσή του πολύ γρήγορα στα δεδομένα εκπαίδευσης, όμως πολύ πιθανόν να υπερεκπαιδεύεται σε αυτά.

Η ακρίβεια επικύρωσης δεν δείχνει το ίδιο επίπεδο βελτίωσης, υποδηλώνοντας και πάλι πιθανή υπερπροσαρμογή (overfitting), όπου το μοντέλο αποδίδει καλά στα δεδομένα της εκπαίδευσης, αλλά όχι το ίδιο καλά στα δεδομένα επικύρωσης, σε δεδομένα δηλαδή που το δίκτυο δεν έχει ήδη «δει». Ήδη από την 4^η εποχή περίπου, η ακρίβεια της επικύρωσης σταματάει να αυξάνεται και παρουσιάζει διακυμάνσεις μεταξύ του 70% – 80%.

6.2.1.2 Διάγραμμα Απώλειας

Η απώλεια της εκπαίδευσης μειώνεται γρήγορα και φτάνει στο 0 από την 5^η εποχή, πράγμα το οποίο σημαίνει ότι σε αυτό το σημείο το μοντέλο έχει εξαλείψει τα σφάλματα στο υποσύνολο της εκπαίδευσης.

Η απώλεια της επικύρωσης δείχνει μια μικρή αστάθεια στις τελευταίες εποχές. Αυτό επιβεβαιώνει την πιθανή υπερπροσαρμογή (overfitting), καθώς η απόδοση του μοντέλου στο υποσύνολο επικύρωσης δεν ακολουθεί την μειούμενη πορεία του υποσυνόλου της εκπαίδευσης.

6.2.1.3 Μετρήσεις απόδοσης

Results	
Validation accuracy:	76.92%
Training finished:	Max epochs completed
Training Time	
Start time:	19-Jun-2024 18:19:41
Elapsed time:	17 min 35 sec
Training Cycle	
Epoch:	10 of 10
Iteration:	170 of 170
Iterations per epoch:	17
Maximum iterations:	170
Validation	
Frequency:	5 iterations
Other Information	
Hardware resource:	Single CPU
Learning rate schedule:	Constant
Learning rate:	0.0001

Εικόνα 6.2 Απόδοση δικτύου CNN #1 κατόπιν εκπαίδευσης με τις αρχικές εικόνες

Η τελική ακρίβεια επικύρωσης (validation accuracy) λαμβάνει τελική τιμή ίση με **76,92%**. Οι μετρήσεις στα παραπάνω διαγράμματα που αφορούν την εκπαίδευση και την επικύρωση δείχνουν ότι υπάρχει

περιθώριο βελτίωσης, καθώς το χάσμα μεταξύ τους υποδηλώνει ότι το μοντέλο προσαρμόζεται περισσότερο από όσο θα έπρεπε στα δεδομένα της εκπαίδευσης, μαθαίνοντας πολύ καλά τα χαρακτηριστικά τους, τα οποία δεν μπορούν απαραίτητα να γενικευτούν στα δεδομένα επικύρωσης.

Το δίκτυο εκπαιδεύεται για 17 λεπτά σε διάρκεια 10 εποχών και 170 επαναλήψεων συνολικά. Απαιτείται λοιπόν πολύ περισσότερη ώρα και υπολογιστική δύναμη συγκριτικά με το FNN που εξετάστηκε στο προηγούμενο κεφάλαιο, όμως φαίνεται να αποδίδει λίγο καλύτερα.

6.2.1.4 Confusion Matrix

		Target Class				
		benign	malignant	normal		
Output Class	benign	60 51.3%	7 6.0%	9 7.7%	78.9%	21.1%
	malignant	5 4.3%	24 20.5%	0 0.0%	82.8%	17.2%
	normal	1 0.9%	0 0.0%	11 9.4%	91.7%	8.3%
		90.9%	77.4%	55.0%	81.2%	18.8%
		9.1%	22.6%	45.0%		

Εικόνα 6.3 Confusion Matrix του CNN #1 κατόπιν εκπαίδευσης με τις αρχικές εικόνες

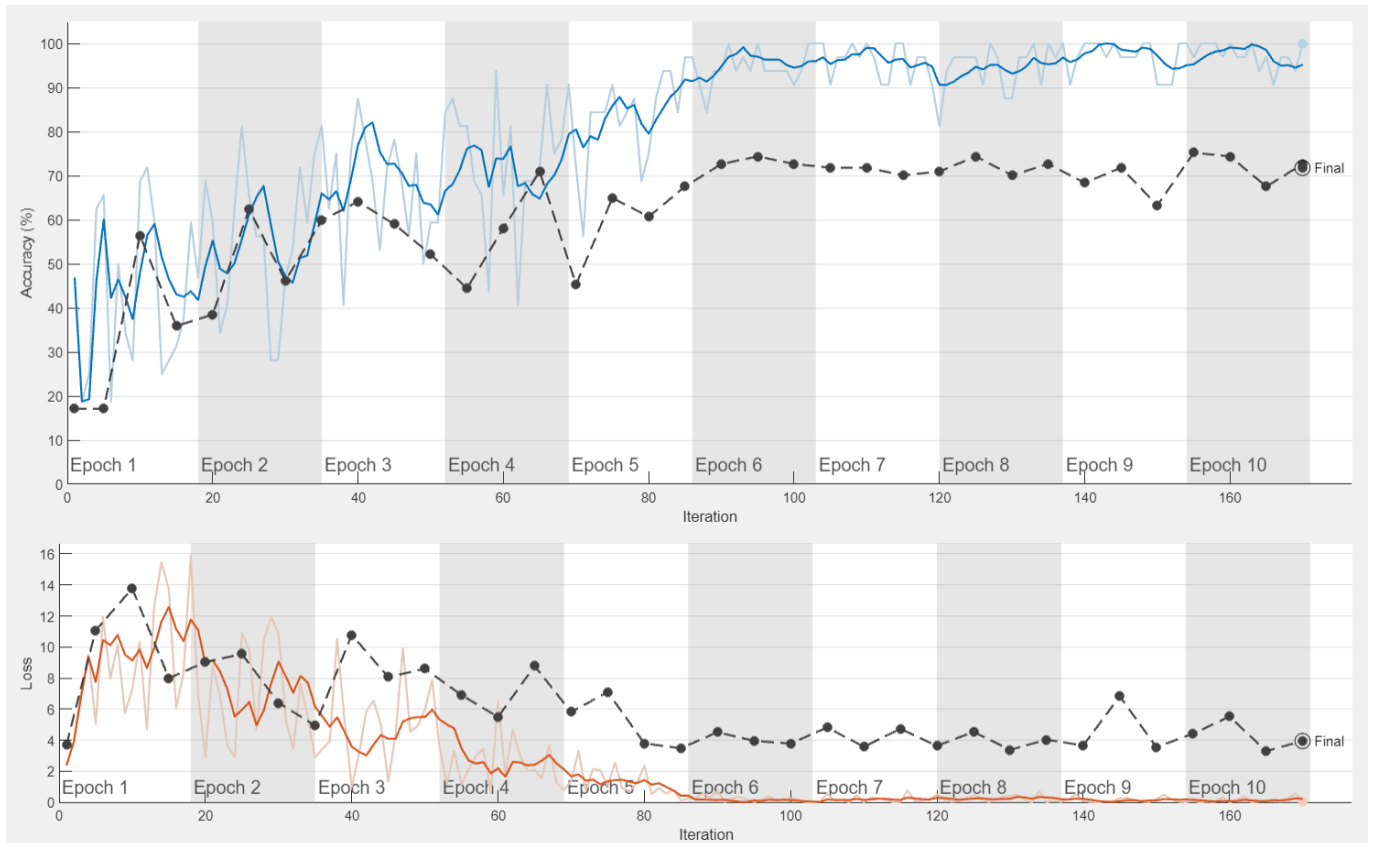
Από τον πίνακα σύγχυσης (confusion matrix) που έχει προκύψει κατόπιν κατηγοριοποίησης των δεδομένων δοκιμής (testing) και μετά την ολοκλήρωση της εκπαίδευσης του δικτύου, το συνολικό ποσοστό επιτυχίας του μοντέλου είναι **81,2%** ενώ το ποσοστό αποτυχίας είναι **18,8%**.

- **Benign:** Το δίκτυο, παρά τα βάρη των κλάσεων, εξακολουθεί να κατηγοριοποιεί λανθασμένα κάποιες εικόνες άλλων κλάσεων ως καλοήθειες. Από την άλλη πλευρά, τα πραγματικά παραδείγματα της κλάσης αυτής φαίνεται να αναγνωρίζονται με μεγάλη επιτυχία από το δίκτυο, όπου κατά 90,9% τα προβλέπει σωστά.
- **Malignant:** Η ακρίβεια των προβλέψεων είναι αρκετά καλή. Φαίνεται ότι το δίκτυο μπορεί να αναγνωρίσει πολύ καλά την διαφορά μεταξύ των malignant και normal εικόνων, καθώς κανένα δείγμα της κλάσης malignant δεν προβλέφθηκε ως normal και αντίστροφα.
- **Normal:** Οι προβλέψεις αυτής της κλάσης φαίνεται να είναι οι πιο αξιόπιστες, με ποσοστό επιτυχίας 91,7%. Από τις συνολικές προβλέψεις που έγιναν στην κατηγορία normal, μόνο 1 περίπτωση καλοήθους όγκου προβλέφθηκε λανθασμένα ως υγιής μαστός. Ωστόσο, ακόμη και με την χρήση βαρών, φαίνεται ότι σχεδόν τα μισά δείγματα της κατηγορίας normal κατηγοριοποιήθηκαν λανθασμένα ως καλοήθειες.

Γενικά το μοντέλο φαίνεται να έχει μια σχετικά καλή απόδοση, αλλά υπάρχει περιθώριο βελτίωσης. Παρά το γεγονός ότι είναι ένα σχετικά «ρηχό» δίκτυο, αφού δεν έχει πάρα πολλά συνεκλεκτικά επίπεδα και δεν μαθαίνει ίσως τόσο περίπλοκα χαρακτηριστικά, το ποσοστό επιτυχίας του φαίνεται να είναι καλύτερο από το αντίστοιχο του FNN στο πείραμα με το ίδιο σύνολο εικόνων (Κεφάλαιο 5.3.4).

6.2.2 Εικόνες με FFT-2D, συχνοτικό φίλτρο και IFFT-2D

Επόμενο πείραμα είναι η αξιολόγηση του CNN στην εκπαίδευσή του με εικόνες που έχουν υποστεί επεξεργασία στο πεδίο της συχνότητας και που έχουν φιλτραριστεί οι περιττές λεπτομέρειες και ο θόρυβος που ανήκουν στις χαμηλές συχνότητες.



Εικόνα 6.4 Διαγράμματα Ακρίβειας και Απώλειας της εκπαίδευσης και επικύρωσης του δικτύου CNN #1 με τις φιλτραρισμένες εικόνες

6.2.2.1 Διάγραμμα Ακρίβειας

Η ακρίβεια εκπαίδευσης αυξάνεται πιο αργά από αυτήν του προηγούμενου πειράματος (Εικόνα 6.1) και δείχνει να κυμαίνεται κοντά στο 90% – 100% στις τελευταίες εποχές της εκπαίδευσης.

Η ακρίβεια της επικύρωσης παρουσιάζει μια αστάθεια από την 8^η έως την 10^η εποχή της εκπαίδευσης, παρά το γεγονός ότι μεταξύ των εποχών 6 και 7 φαινόταν να έχει ισορροπήσει. Η τυχαία διαφοροποίηση και «ποικιλία» των δεδομένων που ανήκουν στο υποσύνολο της επικύρωσης μπορεί να οφείλεται στις αυξομειώσεις αυτές, ειδικά εφόσον το υποσύνολο αυτό είναι σχετικά μικρό.

6.2.2.2 Διάγραμμα Απώλειας

Η απώλεια της εκπαίδευσης είναι έντονη στις πρώτες εποχές, ωστόσο μετά την 6^η εποχή έχει ήδη μειωθεί αρκετά και πλησιάζει το 0, πράγμα το οποίο δείχνει ότι το δίκτυο τελικά έμαθε τα χαρακτηριστικά των δεδομένων.

Η απώλεια της επικύρωσης, αντίστοιχα με το διάγραμμα της ακρίβειας, παρουσιάζει απότομες αλλαγές στο τέλος της εκπαίδευσης, ειδικότερα στις εποχές 9 και 10.

6.2.2.3 Μετρήσεις απόδοσης

Results	
Validation accuracy:	71.79%
Training finished:	Max epochs completed
Training Time	
Start time:	19-Jun-2024 18:40:48
Elapsed time:	17 min 8 sec
Training Cycle	
Epoch:	10 of 10
Iteration:	170 of 170
Iterations per epoch:	17
Maximum iterations:	170
Validation	
Frequency:	5 iterations
Other Information	
Hardware resource:	Single CPU
Learning rate schedule:	Constant
Learning rate:	0.0001

Εικόνα 6.5 Απόδοση δικτύου CNN #1 κατόπιν εκπαίδευσης με τις φιλτραρισμένες εικόνες

Η τελική ακρίβεια επικύρωσης (validation accuracy) **71,79%** και είναι λίγο χαμηλότερη από αυτή του προηγούμενου πειράματος. Όμως λόγω του γεγονότος ότι παρουσιάζονται μικρές ταλαντώσεις γύρω από την τιμή σταθεροποίησης, δεν λαμβάνεται υπόψιν η διαφορά καθώς είναι πολύ μικρή.

Το CNN εκπαιδεύτηκε σε διάρκεια 17 λεπτών περίπου για 10 εποχές και 170 επαναλήψεις. Ο ρυθμός μάθησης ήταν σταθερός και ίσος με $1e-04$. Η μείωση των λεπτομερειών των μείωσε κατά μερικά δευτερόλεπτα τον χρόνο εκπαίδευσης του δικτύου. Παρατηρήθηκε όμοια πορεία εκπαίδευσης, λίγο πιο αργή ωστόσο και με μικρές διαφοροποιήσεις.

6.2.2.4 Confusion Matrix

		Target Class				
		benign	malignant	normal		
Output Class	benign	62 53.0%	8 6.8%	8 6.8%	79.5%	20.5%
	malignant	2 1.7%	21 17.9%	1 0.9%	87.5%	12.5%
	normal	2 1.7%	2 1.7%	11 9.4%	73.3%	26.7%
		93.9% 6.1%	67.7% 32.3%	55.0% 45.0%	80.3%	19.7%

Εικόνα 6.6 Confusion Matrix του CNN #1 κατόπιν εκπαίδευσης με τις φιλτραρισμένες εικόνες

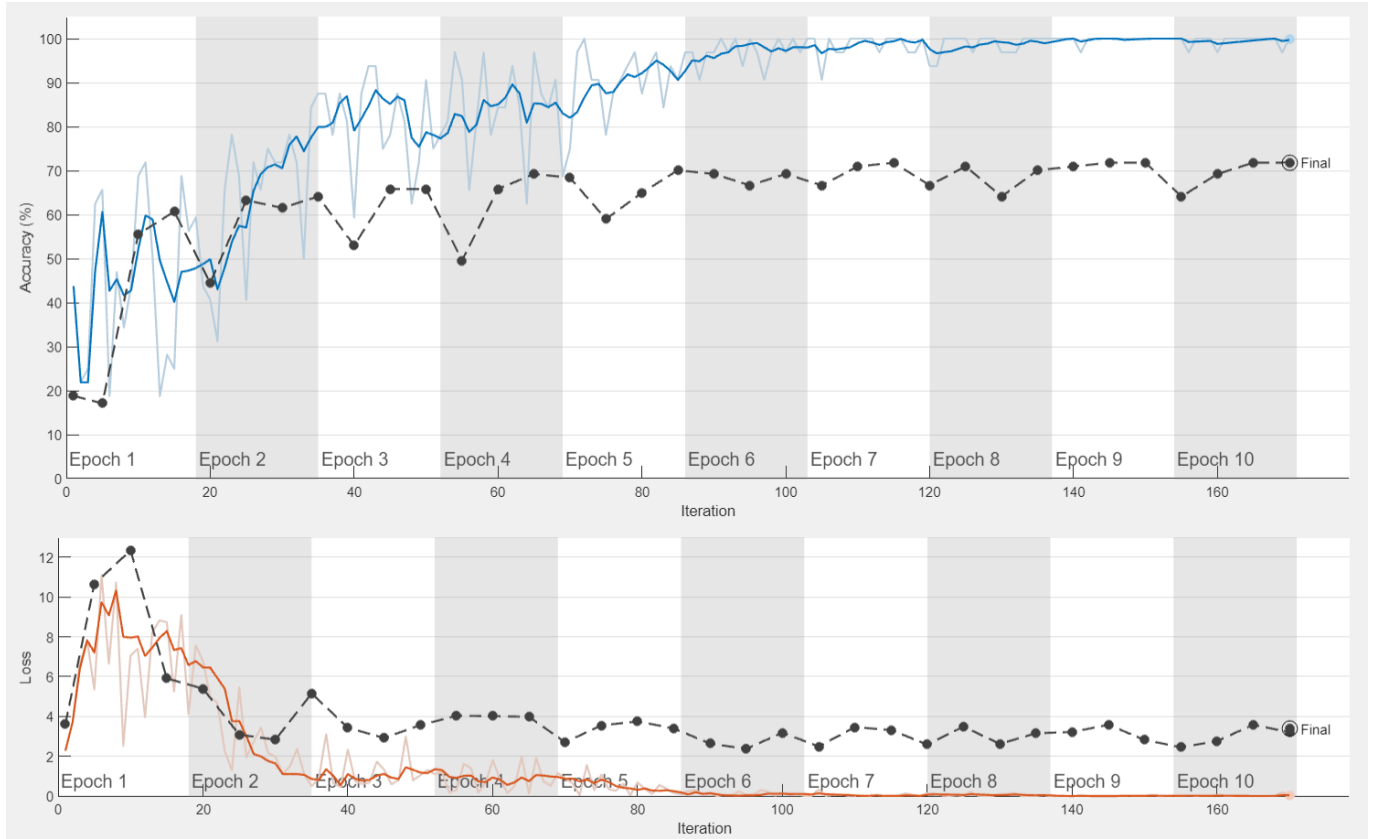
Το συνολικό ποσοστό επιτυχίας του δικτύου είναι **80,3%** ενώ το ποσοστό αποτυχίας **19,7%**.

- **Benign:** Το δίκτυο αποδίδει πολύ καλά στην αναγνώριση καλοήθων όγκων, αν και παρά την χρήση βαρών εξακολουθεί να έχει μια «προκατάληψη» προς αυτήν την κλάση και να την προβλέπει συχνότερα από τις υπόλοιπες. Το 93,9% των δειγμάτων αυτής της κατηγορίας ταξινομήθηκε σωστά, αλλά υπάρχουν κάποιες εικόνες κακοήθων όγκων και φυσιολογικών ιστών που κατηγοριοποιούνται λάθος ως καλοήθειες.
- **Malignant:** Το δίκτυο αποδίδει χειρότερα στην αναγνώριση κακοήθων όγκων, προβλέποντας 10 εικόνες αυτής της κλάσης λανθασμένα ως καλοήθειες (8) και φυσιολογικές (2). Μια θετική πτυχή είναι ότι η πρόβλεψη κακοήθους όγκου ως καλοήθους είναι προτιμότερη από την πρόβλεψη ως υγιούς ιστού κατά την διάγνωση, καθώς θα ζητηθούν ούτως ή άλλως περαιτέρω εξετάσεις σε μια τέτοια περίπτωση.
- **Normal:** Η πλειοψηφία των προβλέψεων για φυσιολογικούς ιστούς είναι σωστή, ωστόσο εξακολουθούν να υπάρχουν αρκετές περιπτώσεις οι οποίες ενώ ανήκουν σε αυτήν την κλάση προβλέπονται ως καλοήθειες.

Το μοντέλο γενικά παρουσιάζει καλή απόδοση στην αναγνώριση των παραδειγμάτων μεταξύ των κατηγοριών malignant και normal, αλλά εξακολουθεί να αποτυγχάνει στον σωστό διαχωρισμό των χαρακτηριστικών που διαφοροποιούν την κατηγορία benign με τις υπόλοιπες, αφού τείνει να την προβλέπει πολύ περισσότερο.

6.2.3 Εικόνες από άθροιση των αρχικών και των φιλτραρισμένων εικόνων

Τέλος, για την αξιολόγηση αυτής της αρχιτεκτονικής CNN, το δίκτυο θα εκπαιδευτεί στις εικόνες που έχουν προκύψει ως αθροίσματα των αρχικών και των φιλτραρισμένων, στις οποίες ήδη από την επεξεργασία που έχουν υποστεί ξεχωρίζουν περισσότερο τα σημεία ενδιαφέροντος. Αναμένεται να παρατηρηθεί βελτίωση της απόδοσης, όπως συνέβη και στην περίπτωση του FNN.



Εικόνα 6.7 Διαγράμματα Ακρίβειας και Απώλειας της εκπαίδευσης και επικύρωσης του δικτύου CNN #1 με τις αθροισμένες εικόνες

6.2.3.1 Διάγραμμα Ακρίβειας

Η ακρίβεια της εκπαίδευσης αυξάνεται, όπως και προηγουμένως, σταθερά μέχρι περίπου την 6^η εποχή και στη συνέχεια σταθεροποιείται στο 100%. Το δίκτυο εξακολουθεί να παρουσιάζει υπερεκπαίδευση (overfitting), αφού και αυτή τη φορά μαθαίνει πολύ καλά τα χαρακτηριστικά των εικόνων στις οποίες εκπαιδεύεται, που όμως δεν γενικεύονται σε νέα δεδομένα.

Η ακρίβεια επικύρωσης έχει κάπως πιο ασταθή συμπεριφορά γύρω από το σημείο ισορροπίας και ως αποτέλεσμα αυτής της αστάθειας η τελική τιμή είναι λίγο υψηλότερη από 70%.

6.2.3.2 Διάγραμμα Απώλειας

Η απώλεια κατά την εκπαίδευση μειώνεται πιο γρήγορα στις πρώτες εποχές, αφού από περίπου 10 φτάνει να μηδενίζεται στην 6^η εποχή.

Η απώλεια επικύρωσης έχει και σε αυτό το διάγραμμα πιο ασταθή μορφή, το οποίο ερμηνεύεται ως overfitting, αφού το δίκτυο δεν δείχνει να έχει το ίδιο καλή απόδοση σε δεδομένα που δεν έχει μάθει.

6.2.3.3 Μετρήσεις Απόδοσης

Results	
Validation accuracy:	71.79%
Training finished:	Max epochs completed
Training Time	
Start time:	19-Jun-2024 19:40:22
Elapsed time:	17 min 4 sec
Training Cycle	
Epoch:	10 of 10
Iteration:	170 of 170
Iterations per epoch:	17
Maximum iterations:	170
Validation	
Frequency:	5 iterations
Other Information	
Hardware resource:	Single CPU
Learning rate schedule:	Constant
Learning rate:	0.0001

Εικόνα 6.8 Απόδοση δικτύου CNN #1 κατόπιν εκπαίδευσης με τις αθροισμένες εικόνες

Όπως φάνηκε και από το Διάγραμμα της Ακρίβειας, πράγματι η τελική τιμή της ακρίβειας επικύρωσης (validation accuracy) είναι **71,79%**, δηλαδή λίγο παραπάνω από 70%.

Το δίκτυο εκπαιδεύτηκε για 17 λεπτά περίπου, με 10 εποχές και 170 επαναλήψεις κατά την διάρκεια της εκπαίδευσης. Ο ρυθμός μάθησης είναι $1e-04$ όπως και στα προηγούμενα πειράματα.

6.2.3.4 Confusion Matrix

		Target Class				
		benign	malignant	normal		
Output Class	benign	62 53.0%	9 7.7%	8 6.8%	78.5%	21.5%
	malignant	3 2.6%	22 18.8%	1 0.9%	84.6%	15.4%
	normal	1 0.9%	0 0.0%	11 9.4%	91.7%	8.3%
		93.9%	71.0%	55.0%	81.2%	18.8%
		6.1%	29.0%	45.0%		

Εικόνα 6.9 Confusion Matrix του CNN #1 κατόπιν εκπαίδευσης με τις αθροισμένες εικόνες

Από το confusion matrix του μοντέλου που εκπαιδεύτηκε με τις αθροισμένες εικόνες, το συνολικό ποσοστό επιτυχίας ανέβηκε στο **81,2%**. Ωστόσο δεν παρατηρείται μεγάλη διαφορά στις αποφάσεις του δικτύου ως προς την κατηγοριοποίηση των εικόνων στις κλάσεις.

- **Benign:** Είναι η κατηγορία με το χαμηλότερο ποσοστό σωστών προβλέψεων (78,5%), που όπως έχει ήδη αναφερθεί οφείλεται κυρίως στο γεγονός ότι οι περισσότερες εικόνες του dataset ανήκουν σε αυτήν την κλάση.
- **Malignant:** Το δίκτυο διατηρεί την αρκετά καλή απόδοσή του στις προβλέψεις για κακοήθεις όγκους, ενώ ένα μεγάλο ποσοστό αυτών (71%) αναγνωρίζεται σωστά. Κάποιες από τις περιπτώσεις κακοήθειας κατηγοριοποιούνται ως καλοήθειες, αλλά καμία ως φυσιολογικός ιστός. Έτσι, ακόμη και όταν ένας κακοήθης όγκος εκτιμάται λανθασμένα ως καλοήθης, θα ζητηθεί πάντα επιπλέον διερεύνηση για την επιβεβαίωση της διάγνωσης.
- **Normal:** Οι σωστές προβλέψεις για την κλάση normal, μετά την επιπλέον επεξεργασία των εικόνων που έγινε με σκοπό να επισημανθούν περισσότερο τα σημεία ενδιαφέροντος, δείχνουν σημαντική βελτίωση. Ενώ στο προηγούμενο πείραμα (Κεφάλαιο 6.2.2.4) είχαν προβλεφθεί λάθος 4 εικόνες στην κλάση normal, πλέον μόνο 1 είναι λάθος και αυτό ανεβάζει το ποσοστό των σωστών προβλέψεων στο 91,7%, πράγμα που οφείλεται και στο μεγαλύτερο βάρος της κλάσης. Οι διαφορές μεταξύ κακοήθων όγκων και φυσιολογικών ιστών ωστόσο είναι αρκετά ξεκάθαρες για το δίκτυο, καθώς κανένα δείγμα της κατηγορίας malignant δεν ταξινομήθηκε ως normal, ενώ μόνο 1 δείγμα της κατηγορίας normal ταξινομήθηκε ως malignant.

Σύμφωνα με αυτά τα αποτελέσματα της εκπαίδευσης του CNN με τις εικόνες που προέκυψαν από την πρόσθεση των αρχικών και των φιλτραρισμένων για την επισήμανση περιοχών ενδιαφέροντος, παρατηρείται μια πολύ μικρή βελτίωση στην συνολική απόδοση του δικτύου. Συνεπώς, για ακόμα

καλύτερα αποτελέσματα, η βελτίωση του CNN μπορεί να συμβάλει σε έναν βαθμό στην αύξηση της ακρίβειας, όμως η ανισορροπία των δεδομένων μπορεί να περιορίσει την επίτευξη σημαντικής βελτίωσης.

6.3 CNN – 2^η Αρχιτεκτονική

Η επόμενη αρχιτεκτονική CNN που χρησιμοποιήθηκε στα πλαίσια της έρευνας, αφορά ένα πιο βαθύ συνελκτικό νευρωνικό δίκτυο, με περισσότερα επίπεδα και μεγαλύτερη πολυπλοκότητα, ώστε να εξάγει πιο περίπλοκα χαρακτηριστικά από τις εικόνες, τα οποία ενδεχομένως να βοηθήσουν στην ορθότερη κατηγοριοποίηση αυτών.

% CNN architecture

```
layers = [  
    imageInputLayer([224 224 1])  
  
    convolution2dLayer(5,32, 'Padding', 'same')  
    batchNormalizationLayer  
    reluLayer  
  
    maxPooling2dLayer(2, 'Stride', 2)  
  
    convolution2dLayer(5,64, 'Padding', 'same')  
    batchNormalizationLayer  
    reluLayer  
  
    maxPooling2dLayer(2, 'Stride', 2)  
  
    convolution2dLayer(3,64, 'Padding', 'same')  
    batchNormalizationLayer  
    reluLayer  
  
    maxPooling2dLayer(2, 'Stride', 2)  
  
    convolution2dLayer(3,128, 'Padding', 'same')  
    batchNormalizationLayer  
    reluLayer  
  
    maxPooling2dLayer(2, 'Stride', 2)  
  
    convolution2dLayer(3,256, 'Padding', 'same')  
    batchNormalizationLayer  
    reluLayer  
    dropoutLayer(0.1)  
  
    maxPooling2dLayer(2, 'Stride', 2)  
  
    convolution2dLayer(3,512, 'Padding', 'same')  
    batchNormalizationLayer  
    reluLayer  
    dropoutLayer(0.1)  
  
    fullyConnectedLayer(512)  
    reluLayer  
    fullyConnectedLayer(256)  
    reluLayer  
    fullyConnectedLayer(64)  
    reluLayer  
    fullyConnectedLayer(3)  
    softmaxLayer
```

```
classificationLayer('Classes', categorical({'benign', 'malignant', 'normal'}), ...  
                    'ClassWeights', [1, 4.16, 6.58])  
];
```

Στην ανωτέρω αρχιτεκτονική CNN έχουν προστεθεί περισσότερα επίπεδα συνελίξεων, με διαφορετικό αριθμό φίλτρων και διαστάσεις. Επιπλέον, προκειμένου να αποφευχθεί η υπερ-εκπαίδευση του δικτύου, έχουν προστεθεί επίπεδα dropout, έτσι ώστε να μειωθεί η πιθανότητα υπερ-προσαρμογής στα δεδομένα της εκπαίδευσης. Πριν το επίπεδο κατηγοριοποίησης, έχουν προστεθεί περισσότερα πλήρως συνδεδεμένα επίπεδα. Τέλος, τα βάρη των κλάσεων malignant και normal έχουν διπλασιαστεί.

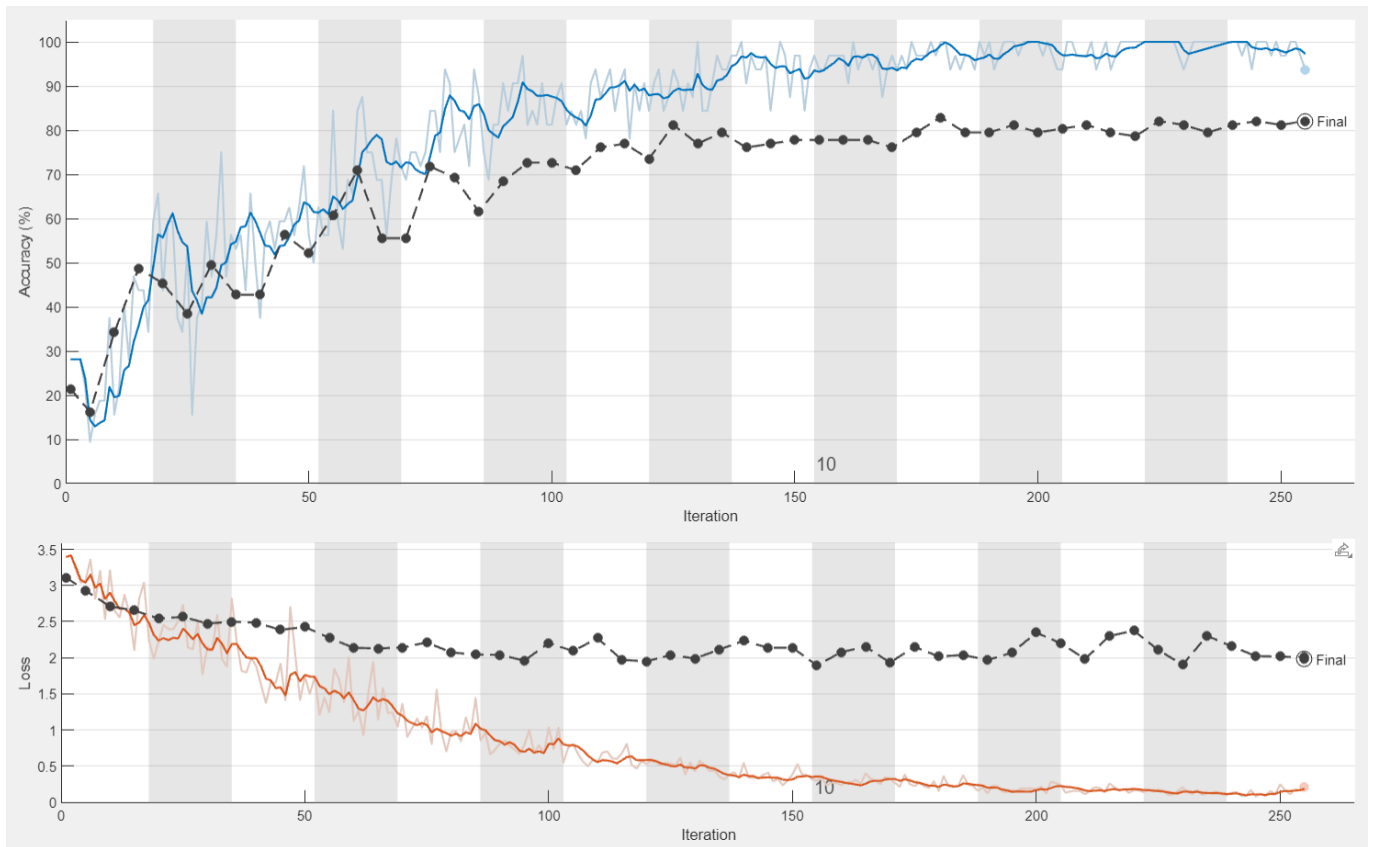
Οι ρυθμίσεις του δικτύου είναι:

```
% Training options  
options = trainingOptions('sgdm', ...  
    'MaxEpochs', 15, ...  
    'MiniBatchSize', 32, ...  
    'Shuffle', 'every-epoch', ...  
    'InitialLearnRate', 1e-04, ...  
    'ValidationData', valData, ...  
    'ValidationFrequency', 5, ...  
    'Plots', 'training-progress', ...  
    'Verbose', true, ...  
    'VerboseFrequency', 5);
```

Ο αριθμός των εποχών εκπαίδευσης του δικτύου αυξήθηκε σε 15, ενώ οι υπόλοιπες ρυθμίσεις παραμένουν ίδιες με αυτές του προηγούμενου CNN.

6.3.1 Εικόνες χωρίς επεξεργασία

Το πρώτο πείραμα με αυτήν την αρχιτεκτονική θα είναι, όπως και προηγουμένως, χρησιμοποιώντας τις εικόνες από το dataset όπως ακριβώς είναι χωρίς καμία επεξεργασία σε αυτές πέρα από τις απαραίτητες ενέργειες για την εισαγωγή τους ως δεδομένα στο δίκτυο.



Εικόνα 6.10 Διαγράμματα Ακρίβειας και Απώλειας της εκπαίδευσης και επικύρωσης του δικτύου CNN #2 με τις αρχικές εικόνες

6.3.1.1 Διάγραμμα Ακρίβειας

Η ακρίβεια της εκπαίδευσης αυξάνεται συνεχώς και σταθεροποιείται πολύ κοντά στο 100% στις τελευταίες 3 εποχές.

Η ακρίβεια της επικύρωσης φαίνεται να αυξάνεται με μεγαλύτερο βαθμό κατά τις πρώτες 8 εποχές, ενώ στη συνέχεια κυμαίνεται γύρω από το 70% και από την 12^η εποχή και έπειτα σταθεροποιείται κοντά στο 80%. Αυτό δείχνει ότι το δίκτυο βελτιώνεται συνεχώς, με μεγαλύτερο ρυθμό στην αρχή της εκπαίδευσης και με μικρότερο καθώς το δίκτυο έχει ήδη μάθει τα περισσότερα χαρακτηριστικά.

6.3.1.2 Διάγραμμα Απώλειας

Η απώλεια της εκπαίδευσης μειώνεται με αργό ρυθμό και πλησιάζει το 0, κάτι το οποίο δείχνει πως το δίκτυο μαθαίνει διάφορα χαρακτηριστικά με σταθερό ρυθμό, χωρίς να υπερεκπαιδεύεται σε κάποια από αυτά.

Η απώλεια της επικύρωσης μειώνεται πολύ πιο αργά και στη συνέχεια παρουσιάζει μια μικρή αστάθεια, γεγονός που οφείλεται σε χαρακτηριστικά τα οποία υπάρχουν στα δεδομένα της εκπαίδευσης αλλά πολύ πιθανόν να μην γενικεύονται το ίδιο καλά στα δεδομένα της επικύρωσης.

6.3.1.3 Μετρήσεις απόδοσης

Results	
Validation accuracy:	82.05%
Training finished:	Max epochs completed
Training Time	
Start time:	19-Jun-2024 15:36:33
Elapsed time:	30 min 3 sec
Training Cycle	
Epoch:	15 of 15
Iteration:	255 of 255
Iterations per epoch:	17
Maximum iterations:	255
Validation	
Frequency:	5 iterations
Other Information	
Hardware resource:	Single CPU
Learning rate schedule:	Constant
Learning rate:	0.0001

Εικόνα 6.11 Απόδοση δικτύου CNN #2 κατόπιν εκπαίδευσης με τις αρχικές εικόνες

Η τελική ακρίβεια επικύρωσης του δικτύου φτάνει στο **82,05%** που σημαίνει ότι η προσθήκη συνελκτικών επιπέδων βοήθησε το μοντέλο να αποδώσει λίγο καλύτερα. Το ποσοστό αυτό είναι αρκετά ικανοποιητικό, λαμβάνοντας υπόψιν τα λίγα δείγματα που περιέχονται στο dataset. Αναμένεται ότι τα αποτελέσματα από την αξιολόγηση στο confusion matrix θα είναι εξίσου ικανοποιητικά.

Το δίκτυο χρειάστηκε 30 λεπτά για να ολοκληρώσει την εκπαίδευση, διάρκειας 15 εποχών και 255 επαναλήψεων με ρυθμό μάθησης $1e-04$.

6.3.1.4 Confusion Matrix

Confusion Matrix

Output Class	benign	58 49.6%	7 6.0%	7 6.0%	80.6% 19.4%
	malignant	6 5.1%	23 19.7%	1 0.9%	76.7% 23.3%
	normal	2 1.7%	1 0.9%	12 10.3%	80.0% 20.0%
		87.9% 12.1%	74.2% 25.8%	60.0% 40.0%	79.5% 20.5%
		benign	malignant	normal	
		Target Class			

Εικόνα 6.12 Confusion Matrix του CNN #2 κατόπιν εκπαίδευσης με τις αρχικές εικόνες

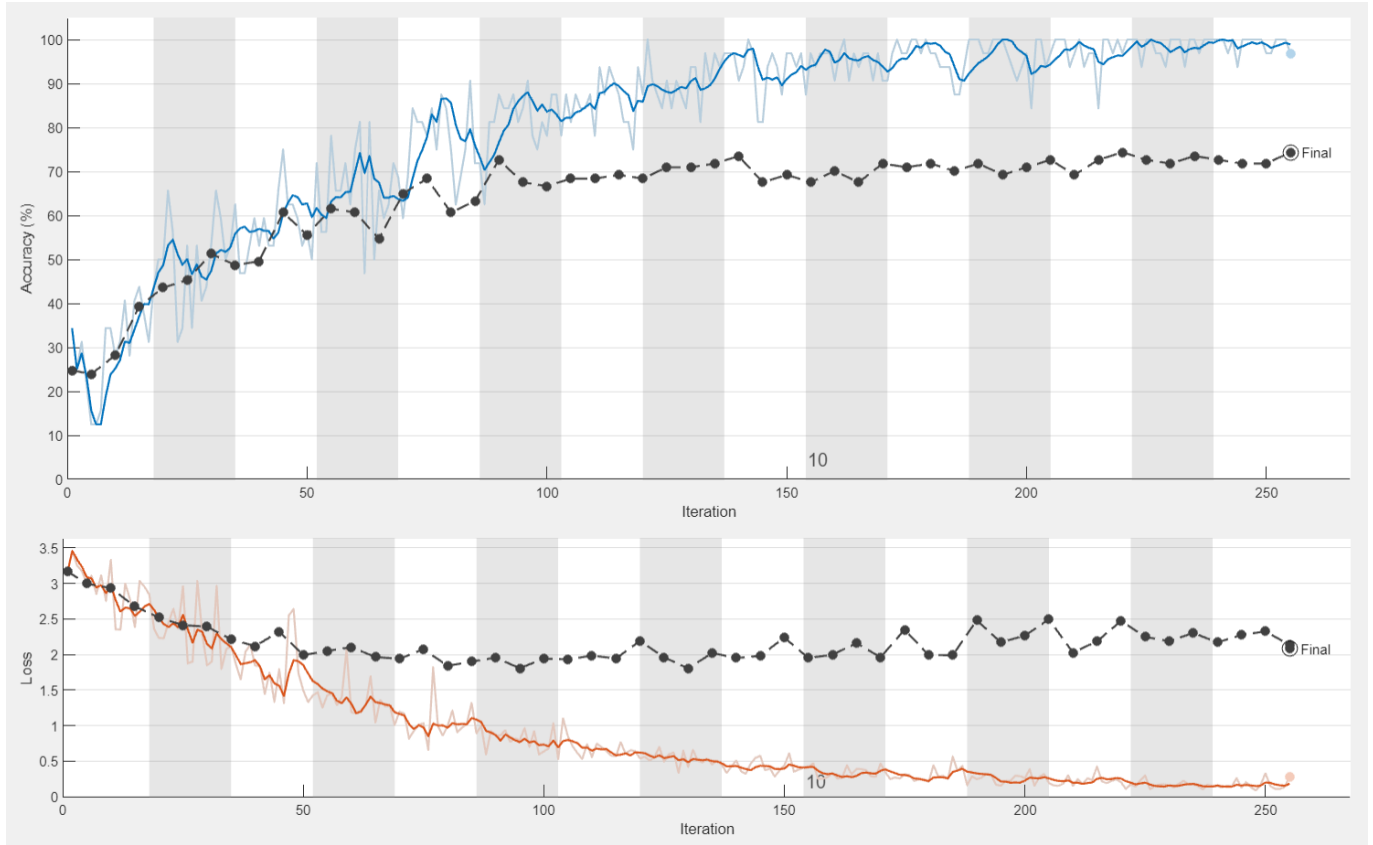
Το συνολικό ποσοστό επιτυχίας **79,5%** δείχνει σχετικά καλή γενική απόδοση του δικτύου, αν και υπάρχει περιθώριο βελτίωσης.

- **Benign:** Η κλάση benign, αν και έχει το μικρότερο βάρος, φαίνεται να διατηρεί την επιρροή της στις προβλέψεις του δικτύου. Λόγω των περισσότερων δειγμάτων αυτής της κατηγορίας είναι λογικό να παρουσιάζει μεγαλύτερο ποσοστό επιτυχίας ως προς την αξιολόγηση των δειγμάτων της (87,9%), καθώς μια λάθος πρόβλεψη έχει λιγότερο αρνητικό αντίκτυπο στο συνολικό ποσοστό επιτυχίας της.
- **Malignant:** Η απόδοση της κλάσης malignant είναι καλή. Το μεγαλύτερο ποσοστό των προβλέψεων για αυτή την κατηγορία ήταν σωστές, με μόνο 6 δείγματα καλοηθών όγκων να προβλέπονται λανθασμένα ως κακοήθειες, και μόλις 1 φυσιολογικός ιστός να κατατάσσεται λανθασμένα στην κλάση malignant.
- **Normal:** Η αναγνώριση των φυσιολογικών περιστατικών παρουσιάζει ικανοποιητικό ποσοστό σωστών προβλέψεων. Από τις 15 εικόνες που προβλέφθηκαν ως normal, οι 12 ήταν σωστές, υποδεικνύοντας ότι το δίκτυο αποδίδει αρκετά καλά και δεν προβλέπει φυσιολογικές περιπτώσεις πιο συχνά από όσο θα έπρεπε. Αυτό είναι ιδιαίτερα σημαντικό, καθώς η υπερβολική πρόβλεψη φυσιολογικών περιπτώσεων θα ήταν ένα κρίσιμο λάθος.

Η αύξηση των βαρών των κλάσεων malignant και benign αύξησε το συνολικό ποσοστό επιτυχίας του δικτύου, αφού οι σωστές προβλέψεις αυτών μετράνε πιο θετικά στην απόδοσή του. Παρά την χρήση των βαρών ωστόσο, οι περισσότερες λάθος προβλέψεις έχουν γίνει για ακόμα μια φορά στην κλάση benign, πιθανότατα λόγω των πολύ περισσότερων δεδομένων που υπάρχουν στο dataset.

6.3.2 Εικόνες με FFT-2D, συχνοτικό φίλτρο και IFFT-2D

Στο τρέχον πείραμα θα γίνει δοκιμή των φιλτραρισμένων εικόνων με την πιο βαθιά αρχιτεκτονική CNN ώστε να διαπιστωθεί αν τελικά η αφαίρεση των χαμηλών συχνοτήτων αυξάνει την απόδοση του δικτύου και την ικανότητά του να διακρίνει πιο αποτελεσματικά τα χαρακτηριστικά που διαφοροποιούν τις εικόνες.



Εικόνα 6.13 Διαγράμματα Ακρίβειας και Απώλειας της εκπαίδευσης και επικύρωσης του δικτύου CNN #2 με τις φιλτραρισμένες εικόνες

6.3.2.1 Διάγραμμα Ακρίβειας

Η ακρίβεια της εκπαίδευσης αυξάνεται με αργό και σταθερό ρυθμό. Το νευρωνικό δίκτυο δεν παρουσιάζει υπερπροσαρμογή (overfitting), αφού μόνο στις τελευταίες 2 εποχές φαίνεται η γραφική της ακρίβειας εκπαίδευσης να αγγίζει τιμές κοντά στο 100%.

Η ακρίβεια της επικύρωσης ακολουθεί την πορεία της ακρίβειας εκπαίδευσης μέχρι την 6^η εποχή και έπειτα αυξάνεται με μικρότερο ρυθμό, δημιουργώντας ένα χάσμα μεταξύ των δύο γραφημάτων. Η ζώνη σταθεροποίησης φαίνεται να βρίσκεται λίγο πιο πάνω από το 70% και δείχνει μια πολύ μικρή αύξηση, γεγονός που υποδεικνύει ότι ακόμα και στις τελευταίες εποχές, το δίκτυο εξακολουθεί να μαθαίνει κάποια επιπλέον πράγματα τα οποία συμβάλλουν στην καλύτερη απόδοσή του.

6.3.2.2 Διάγραμμα Απώλειας

Η απώλεια της εκπαίδευσης έχει ομαλή καθοδική πορεία. Στις τελευταίες εποχές τείνει να πλησιάσει το 0, που σημαίνει ότι το δίκτυο έχει μάθει αρκετά καλά τα δεδομένα της εκπαίδευσης και τα χαρακτηριστικά τους.

Η απώλεια της επικύρωσης μειώνεται πιο αργά και δεν φτάνει σε τόσο χαμηλό επίπεδο όσο η απώλεια της εκπαίδευσης. Μέχρι και την 7^η εποχή η μείωση είναι πιο ομαλή, ενώ από την 8^η εποχή και έπειτα παρουσιάζει μια μικρή αστάθεια μεταξύ 2 – 2,5.

6.3.2.3 Μετρήσεις Απόδοσης

Results	
Validation accuracy:	74.36%
Training finished:	Max epochs completed
Training Time	
Start time:	19-Jun-2024 16:10:04
Elapsed time:	30 min 25 sec
Training Cycle	
Epoch:	15 of 15
Iteration:	255 of 255
Iterations per epoch:	17
Maximum iterations:	255
Validation	
Frequency:	5 iterations
Other Information	
Hardware resource:	Single CPU
Learning rate schedule:	Constant
Learning rate:	0.0001

Εικόνα 6.14 Απόδοση δικτύου CNN #2 κατόπιν εκπαίδευσης με τις φιλτραρισμένες εικόνες

Η ακρίβεια της επικύρωσης, μετά την ολοκλήρωση της εκπαίδευσης του δικτύου, σταθεροποιείται στο **74,36%**. Αν και είναι λίγο χαμηλότερη από την τιμή του προηγούμενου πειράματος, η διαφορά τους είναι σχετικά μικρή και δεν αποτελεί κριτήριο για να αξιολογηθεί η γενική απόδοση του δικτύου.

Το δίκτυο εκπαιδεύτηκε σε διάρκεια 30 λεπτών, ομοίως με το προηγούμενο πείραμα, για 15 εποχές και 255 επαναλήψεις, και με σταθερό ρυθμό μάθησης $1e-04$.

6.3.2.4 Confusion Matrix

		Target Class				
		benign	malignant	normal		
Output Class	benign	60 51.3%	4 3.4%	9 7.7%	82.2%	17.8%
	malignant	3 2.6%	26 22.2%	0 0.0%	89.7%	10.3%
	normal	3 2.6%	1 0.9%	11 9.4%	73.3%	26.7%
		90.9%	83.9%	55.0%	82.9%	17.1%
		9.1%	16.1%	45.0%		

Εικόνα 6.15 Confusion Matrix του CNN #2 κατόπιν εκπαίδευσης με τις φιλτραρισμένες εικόνες

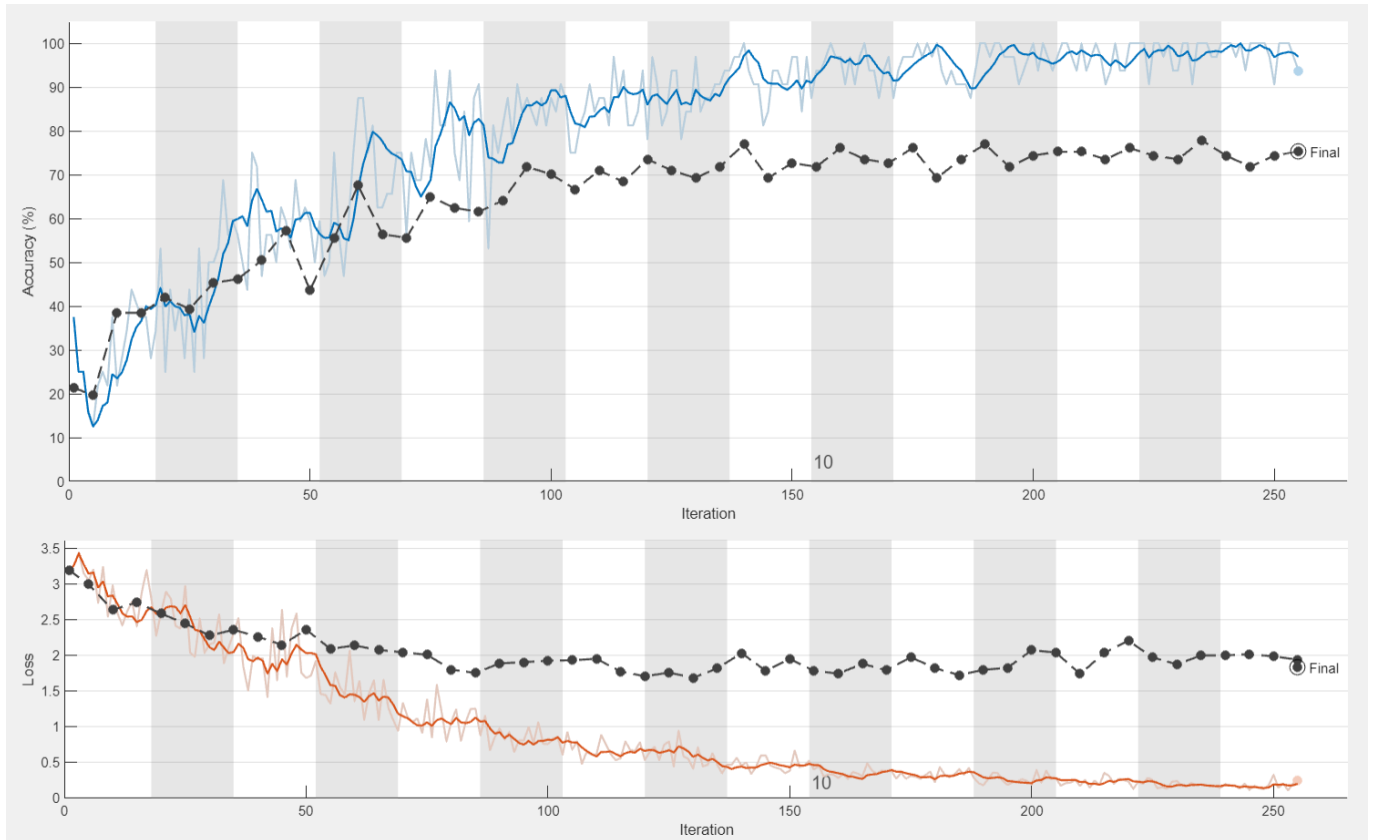
Το συνολικό ποσοστό επιτυχίας αυξήθηκε στο **82,9%**, δηλαδή περίπου 3% περισσότερο από το πείραμα με τις αρχικές εικόνες στην ίδια αρχιτεκτονική (Κεφάλαιο 6.3.1.4).

- **Benign:** Η κλάση benign εξακολουθεί να έχει αρκετά καλή ακρίβεια όπως και στα προηγούμενα πειράματα. Λόγω και των μεγαλύτερων βαρών των άλλων κλάσεων, οι λάθος προβλέψεις της κλάσης benign έχουν μικρότερο αντίκτυπο στην συνολική απόδοση του δικτύου.
- **Malignant:** Το μοντέλο αναγνωρίζει την πλειοψηφία των κακοήθων εικόνων, επιδεικνύοντας καλή απόδοση σε αυτή την κατηγορία που είναι σχετικά κρίσιμη. Μόνο 4 κακοήθεις εικόνες ταξινομήθηκαν ως καλοήθεις, ενώ μόνο 1 ως φυσιολογική. Επιπλέον, οι προβλέψεις που γίνονται από το δίκτυο σε αυτήν την κατηγορία έχουν μεγάλο ποσοστό επιτυχίας (89,7%). Συνεπώς η πιθανότητα σοβαρών σφαλμάτων διάγνωσης είναι μικρή.
- **Normal:** Παρά την χρήση μεγαλύτερων βαρών, το δίκτυο εξακολουθεί να προβλέπει αρκετά από τα δείγματα της κλάσης normal ως καλοήθεις όγκους, Παράλληλα, 3 δείγματα της κατηγορίας benign και 1 δείγμα της κατηγορίας malignant προβλέφθηκαν λανθασμένα ως φυσιολογικές περιπτώσεις, γεγονός που υποδηλώνει την ανάγκη για περισσότερα δεδομένα εκπαίδευσης κυρίως από αυτήν την κατηγορία προκειμένου το δίκτυο να αποκτήσει την δυνατότητα να αναγνωρίζει τα χαρακτηριστικά που διαφοροποιούν αυτήν την κλάση από τις άλλες δύο.

Η συνολική απόδοση είναι θετική, με το μοντέλο να αναγνωρίζει επιτυχώς την πλειονότητα των εικόνων και να παρουσιάζει σχετικά χαμηλά ποσοστά λανθασμένων προβλέψεων, τα οποία κυρίως προέρχονται από λάθος προβλέψεις της κλάσης normal, η οποία έχει και τα λιγότερα δείγματα. Επομένως, ακόμα και με την εφαρμογή μεγαλύτερων βαρών και πιο σύνθετης αρχιτεκτονικής, εξακολουθεί να υπάρχει η ανάγκη για περισσότερα δεδομένα εκπαίδευσης προκειμένου να επιτευχθούν ακόμα καλύτερα αποτελέσματα και να βελτιωθεί η συνολική απόδοση του μοντέλου.

6.3.3 Εικόνες από άθροιση των αρχικών και των φιλτραρισμένων

Το τελευταίο πείραμα για την 2^η αρχιτεκτονική CNN, αφορά τις εικόνες που προέκυψαν αθροίζοντας τις αρχικές με τις φιλτραρισμένες. Οι εικόνες αυτές έχουν πιο απλά χαρακτηριστικά και οι περισσότεροι όγκοι έχουν απομονωθεί γύρω από άσπρες περιοχές. Αναμένεται να εξεταστεί η απόδοση της σύνθετης αυτής αρχιτεκτονικής και η ικανότητα του δικτύου να αναγνωρίσει επιτυχώς τα χαρακτηριστικά που διαφοροποιούν τις εικόνες μεταξύ τους.



Εικόνα 6.16 Διαγράμματα Ακρίβειας και Απώλειας της εκπαίδευσης και επικύρωσης του δικτύου CNN #2 με τις αθροισμένες εικόνες

6.3.3.1 Διάγραμμα Ακρίβειας

Η ακρίβεια της εκπαίδευσης αυξάνεται αρχικά με πιο γρήγορο ρυθμό έως την 6^η εποχή και έπειτα συνεχίζει να αυξάνεται με μικρότερο. Όπως και προηγουμένως, το δίκτυο δεν παρουσιάζει το φαινόμενο υπερπροσαρμογής (overfitting), που σημαίνει ότι η εκπαίδευσή του δείχνει να είναι αξιόπιστη.

Η ακρίβεια της επικύρωσης κινείται παράλληλα με αυτήν της εκπαίδευσης, σε λίγο χαμηλότερες τιμές. Πιθανόν κάποια από τα μοτίβα που μαθαίνει το δίκτυο από τα δεδομένα της εκπαίδευσης να μην μπορούν να γενικευτούν το ίδιο καλά στα δεδομένα της επικύρωσης.

6.3.3.2 Διάγραμμα Απώλειας

Η απώλεια κατά την εκπαίδευση δείχνει ομαλή μείωση η οποία σταθεροποιείται σε πολύ χαμηλές τιμές κοντά στο τέλος της εκπαίδευσης. Το δίκτυο μαθαίνει κατά τις επαναλήψεις που εκτελεί να ελαχιστοποιεί σταδιακά την απώλεια στα δεδομένα εκπαίδευσης.

Η απώλεια επικύρωσης από την άλλη, φαίνεται να είναι αρκετά υψηλότερη και δεν μειώνεται πολύ. Το μοντέλο δηλαδή δυσκολεύεται να αναγνωρίσει τα χαρακτηριστικά που μαθαίνει από τα δεδομένα της εκπαίδευσης στα δεδομένα επικύρωσης.

6.3.3.3 Μετρήσεις Απόδοσης

Results	
Validation accuracy:	75.21%
Training finished:	Max epochs completed
Training Time	
Start time:	19-Jun-2024 16:44:15
Elapsed time:	32 min 47 sec
Training Cycle	
Epoch:	15 of 15
Iteration:	255 of 255
Iterations per epoch:	17
Maximum iterations:	255
Validation	
Frequency:	5 iterations
Other Information	
Hardware resource:	Single CPU
Learning rate schedule:	Constant
Learning rate:	0.0001

Εικόνα 6.17 Απόδοση δικτύου CNN #2 κατόπιν εκπαίδευσης με τις αθροισμένες εικόνες

Η ακρίβεια επικύρωσης σταθεροποιείται τελικά στο **75,21%**, δηλαδή λίγο υψηλότερο ποσοστό συγκριτικά με το αμέσως προηγούμενο πείραμα με τις φιλτραρισμένες εικόνες.

Το δίκτυο εκπαιδεύεται σε λίγο περισσότερο χρόνο (32 λεπτά) για ίδιο αριθμό εποχών και επαναλήψεων (15 και 255 αντίστοιχα) με σταθερό ρυθμό μάθησης $1e-04$.

6.3.3.4 Confusion Matrix

		Target Class				
		benign	malignant	normal		
Output Class	benign	61 52.1%	7 6.0%	5 4.3%	83.6%	16.4%
	malignant	4 3.4%	23 19.7%	1 0.9%	82.1%	17.9%
	normal	1 0.9%	1 0.9%	14 12.0%	87.5%	12.5%
		92.4%	74.2%	70.0%	83.8%	16.2%
		7.6%	25.8%	30.0%		

Εικόνα 6.18 Confusion Matrix του CNN #2 κατόπιν εκπαίδευσης με τις αθροισμένες εικόνες

Το συνολικό ποσοστό επιτυχίας **83,8%** είναι αρκετά ικανοποιητικό και υποδηλώνει βελτίωση του δικτύου στην επιτυχή αναγνώριση των χαρακτηριστικών που διαφοροποιούν τις εικόνες μεταξύ τους.

- **Benign:** Οι προβλέψεις της κλάσης benign εξακολουθούν να περιέχουν τις περισσότερες λανθασμένες κατηγοριοποιήσεις (7 στην κλάση malignant και 5 στην κλάση normal, επομένως 12 συνολικά). Κάποιοι από τους καλοήθεις όγκους (4) κατηγοριοποιήθηκαν ψευδώς ως κακοήθειες, ενώ 1 μόνο από αυτούς προβλέφθηκε ως φυσιολογικός ιστός.
- **Malignant:** Οι λανθασμένες προβλέψεις της κλάσης malignant αυξήθηκαν λίγο σε σύγκριση με το προηγούμενο πείραμα (Κεφάλαιο 6.3.2.4), αφού 4 καλοήθεις όγκοι και 1 εικόνα από υγιή μαστό προβλέφθηκαν ως κακοήθειες. Η απόδοση της κλάσης είναι μέτρια προς καλή (74,2%), αφού τα περισσότερα παραδείγματα από αυτήν κατηγοριοποιήθηκαν σωστά.
- **Normal:** Η κλάση normal έχει τις πιο αξιόπιστες προβλέψεις, με μόνο 2 συνολικά λάθος προβλέψεις από τις άλλες δύο κλάσεις ως φυσιολογικές εικόνες. Επιπλέον, φαίνεται να μειώθηκαν και οι λάθος κατηγοριοποιήσεις των πραγματικών φυσιολογικών εικόνων, αφού μόνο 5 από αυτές προβλέφθηκαν ως καλοήθεις όγκοι και 1 ως κακοήθης.

Αυτά τα αποτελέσματα υποδεικνύουν ότι οι βελτιώσεις στην αρχιτεκτονική του δικτύου και η επιπλέον επεξεργασία των εικόνων έχουν θετική επίδραση στην εκπαίδευση του δικτύου, αλλά εξακολουθούν να χρειάζονται περισσότερα δεδομένα εκπαίδευσης για να επιτευχθεί ακόμα υψηλότερη ακρίβεια στις προβλέψεις.

ΚΕΦΑΛΑΙΟ 7^ο: Συμπεράσματα και Περαιτέρω Έρευνα

Τα αποτελέσματα που προέκυψαν από την αξιολόγηση όλων των πειραμάτων για τα τρία μοντέλα που εξετάστηκαν, συγκεντρωτικά είναι τα εξής:

Πίνακας 7.1 Αποτελέσματα Απόδοσης όλων των Νευρωνικών Δικτύων που εξετάστηκαν

	Εικόνες χωρίς επεξεργασία	Φιλτραρισμένες Εικόνες (FFT-2D, χαμηλοπερατό γκαουσιανό φίλτρο, IFFT-2D)	Εικόνες-αθροίσματα αρχικών και φιλτραρισμένων εικόνων
FNN	76,9%	73,5%	77,8%
CNN (Αρχιτεκτονική #1)	81,2%	80,3%	81,2%
CNN (Αρχιτεκτονική #2)	79,5%	82,9%	83,8%

Από τα παραπάνω αποτελέσματα προκύπτει ότι:

- Το **FNN** έδειξε την χαμηλότερη απόδοση σε σύγκριση με τα άλλα μοντέλα, κυρίως λόγω της περιορισμένης ικανότητάς του να εκμεταλλεύεται τις χωρικές πληροφορίες των εικόνων. Αυτό οφείλεται στο γεγονός ότι το FNN δεν έχει τη δυνατότητα να διαχειρίζεται την πολυπλοκότητα και τις λεπτομέρειες των εικόνων τόσο αποτελεσματικά όσο τα CNNs. Συνεπώς, η αξιοπιστία του FNN είναι χαμηλότερη, ειδικά για πιο περίπλοκα δεδομένα όπως οι ιατρικές εικόνες.
- Το «ρηχό» **CNN (Αρχιτεκτονική #1)** είχε καλύτερη απόδοση από το FNN, ιδιαίτερα στις αρχικές και στις αθροισμένες εικόνες, αλλά δεν παρουσίασε μεγάλη βελτίωση όταν χρησιμοποιήθηκαν φιλτραρισμένες εικόνες. Αυτό το μοντέλο κατάφερε να εκμεταλλευτεί καλύτερα τις χωρικές πληροφορίες από τις εικόνες, αλλά το περιορισμένο βάθος του, δεν του επέτρεψε να αναγνωρίσει πιο περίπλοκα πρότυπα.
- Το βαθύ **CNN (Αρχιτεκτονική #2)** ήταν το πιο αξιόπιστο μοντέλο και παρουσίασε την καλύτερη απόδοση, ειδικά όταν χρησιμοποιήθηκαν φιλτραρισμένες εικόνες ή εικόνες από άθροιση των αρχικών και των φιλτραρισμένων. Το μεγαλύτερο βάθος του μοντέλου του επέτρεψε να ανιχνεύσει και να μάθει πιο περίπλοκα χαρακτηριστικά από τις εικόνες, κάνοντάς το πιο αποτελεσματικό για αυτήν την εφαρμογή.

7.1 Επίδραση της Ανισορροπίας των Δεδομένων

Παρατηρήθηκε ότι η ανισορροπία των δεδομένων στις κλάσεις μπορεί να επηρεάσει σημαντικά την απόδοση του νευρωνικού δικτύου. Η ύπαρξη πολύ περισσότερων εικόνων στην κατηγορία benign (437 εικόνες) σε σύγκριση με τις κατηγορίες malignant (210 εικόνες) και normal (133 εικόνες) μπορεί να

οδηγήσει σε ένα μοντέλο που είναι πιο «ευαίσθητο» στο να προβλέπει την κατηγορία με τα περισσότερα δείγματα. Έτσι η ανισορροπία των δεδομένων έχει τα εξής αποτελέσματα:

- **Προκατάληψη προς την κατηγορία benign:** Το μοντέλο τείνει να προβλέπει πιο συχνά την κατηγορία με τα περισσότερα δεδομένα, καθώς έχει περισσότερη «εμπειρία» με αυτήν κατά την εκπαίδευση.
- **Μειωμένη ακρίβεια στις κατηγορίες malignant και normal:** Επειδή υπάρχουν λιγότερα δείγματα στις κατηγορίες malignant και normal, το μοντέλο δεν έχει αρκετές πληροφορίες για να μάθει να τις αναγνωρίζει αποτελεσματικά.

7.2 Αποτελεσματικότητα των Αλλαγών στην Αρχιτεκτονική

Το δίκτυο δυσκολεύεται να διαχωρίσει με ακρίβεια τις καλοήθειες από τις κακοήθειες και τις φυσιολογικές εικόνες, ακόμα και με αλλαγές στην αρχιτεκτονική του. Γενικότερα, παρατηρήθηκε ότι η μορφή των γραφημάτων δεν αλλάζει ιδιαίτερα. Η ακρίβεια και η απώλεια της επικύρωσης έχουν πάντα περίπου την ίδια απόσταση από την ακρίβεια και την απώλεια της εκπαίδευσης, υποδηλώνοντας ότι το δίκτυο ίσως χρειάζεται ένα μεγαλύτερο σύνολο δεδομένων για να αυξηθεί η απόδοσή του σημαντικά.

Επιπλέον παρατηρήσεις που έγιναν κατά την διεξαγωγή των πειραμάτων είναι ότι:

- Χωρίζοντας τα δεδομένα σε υποσύνολα 60% training, 20% validation, 20% testing, η ακρίβεια επικύρωσης δεν ξεπερνά το 60 – 70% σε κάθε περίπτωση, όπως επίσης και η τελική απόδοση του μοντέλου στο confusion matrix είναι πάντα μεταξύ 70 – 80%.
- Χωρίζοντας τα δεδομένα σε υποσύνολα 70% training, 15% validation, 15% testing, όπως έγινε στην παρούσα έρευνα, η ακρίβεια επικύρωσης αυξάνεται και κυμαίνεται μεταξύ 70 – 80% στις περισσότερες περιπτώσεις, ενώ τα αποτελέσματα στο confusion matrix φτάνουν συχνά σε ποσοστά κοντά στο 80 – 83%.

Η ανισορροπία των δεδομένων περιορίζει σημαντικά τη βελτίωση του δικτύου, καθώς όσες αλλαγές και βελτιστοποιήσεις και αν γίνουν στην αρχιτεκτονική του κάθε δικτύου και στις ρυθμίσεις εκπαίδευσής του, το ποσοστό επιτυχίας δεν μπορεί να ξεπεράσει ένα συγκεκριμένο όριο και αυξάνεται με πολύ μικρότερο ρυθμό σε σύγκριση με τις εφαρμοζόμενες αλλαγές. Αν και η χρήση βαρών στις κλάσεις είναι ένα σημαντικό βήμα για την βελτίωση της απόδοσης σε ανισο-κατανομημένα δεδομένα, δεν είναι πανάκεια. Ιδανικά λοιπόν, η εκπαίδευση ενός δικτύου θα χρειαστεί να πραγματοποιηθεί με μεγαλύτερο πλήθος εικόνων, έτσι ώστε να διαχωρίζονται επιτυχώς τα χαρακτηριστικά που διαφοροποιούν τις 3 κλάσεις μεταξύ τους και να φανεί κάποια ουσιαστική διαφορά στην απόδοση του νευρωνικού δικτύου.

7.3 Συγκριτική Ανάλυση Δικτύων 2 Κλάσεων

Κατά τη διάρκεια της έρευνας, πραγματοποιήθηκε ένα επιπλέον πείραμα με χρήση των ίδιων νευρωνικών δικτύων αλλά αυτή τη φορά για την ταξινόμηση δύο κλάσεων (2-class classification), χρησιμοποιώντας τις κλάσεις normal και malignant του ίδιου dataset. Οι αλγόριθμοι κατηγοριοποίησης εδώ καλούνται να επιλύσουν ένα «ευκολότερο» πρόβλημα, που επικεντρώνεται στις δύο «ακραίες» περιπτώσεις. Όπως ήταν αναμενόμενο, είχαν και καλύτερα αποτελέσματα σε σύγκριση με την ταξινόμηση τριών κλάσεων, γεγονός που μπορεί να αποδοθεί στη σαφή διαχωριστικότητα μεταξύ των δύο κλάσεων.

Στον Πίνακα 7.2 παρουσιάζονται συνοπτικά τα αποτελέσματα των confusion matrix κάθε ενός από τα πειράματα της δοκιμής αυτής (2-class classification), όπου χρησιμοποιήθηκαν μόνο οι κλάσεις normal και

malignant. Στα πειράματα αυτά χρησιμοποιήθηκαν ακριβώς τα ίδια νευρωνικά δίκτυα των πειραμάτων με τις 3 κλάσεις (3-class classification), με τη διαφορά ότι δεν υπήρχαν βάρη, καθώς αυτά είχαν υπολογιστεί αναλογικά με τα δείγματα της κλάσης benign, η οποία είχε πολύ περισσότερα δείγματα.

Πίνακας 7.2 Αποτελέσματα Απόδοσης των Νευρωνικών Δικτύων με χρήση 2 κλάσεων (2-class classification, Normal v/s Malignant)

	Εικόνες χωρίς επεξεργασία	Φιλτραρισμένες Εικόνες (FFT-2D, χαμηλοπερατό γκαουσιανό φίλτρο, IFFT-2D)	Εικόνες-αθροίσματα αρχικών και φιλτραρισμένων εικόνων
FNN	80,4%	86,3%	84,3%
CNN (Αρχιτεκτονική #1)	94,1%	84,3%	88,2%
CNN (Αρχιτεκτονική #2)	86,3%	90,2%	90,2%

Μία πρώτη παρατήρηση είναι ότι το CNN γενικά υπερτερεί εδώ του FNN. Επιπλέον, το CNN#1, με την απλούστερη αρχιτεκτονική, έδωσε πολύ ικανοποιητικά αποτελέσματα στην περίπτωση των εικόνων χωρίς επεξεργασία (94,1%). Ενδεχομένως η απλούστερη δομή του δικτύου επιτρέπει την αποτελεσματική αναγνώριση των βασικών χαρακτηριστικών στους υπερήχους μαστού, όπως περιγράμματα και διαφορές στις αποχρώσεις.

Στις περιπτώσεις των επεξεργασμένων εικόνων (είτε με φίλτρο, είτε με άθροιση) το CNN με την πιο βαθιά δομή (CNN#2) φαίνεται να υπερέχει των άλλων δύο. Η πιο σύνθετη αρχιτεκτονική του δικτύου μπορεί να επεξεργάζεται και να αντιλαμβάνεται καλύτερα τις επεξεργασμένες εικόνες, αξιοποιώντας πιο αποτελεσματικά την αφαιρεθείσα περιττή πληροφορία και τα ενισχυμένα χαρακτηριστικά.

Από τα αποτελέσματα αυτών των δοκιμών, διαπιστώνεται ότι η απλοποίηση του προβλήματος σε κατηγοριοποίηση δύο κλάσεων μπορεί να οδηγήσει σε πιο αξιόπιστες και ακριβείς προβλέψεις, ειδικά όταν οι διαφορές μεταξύ των δειγμάτων είναι έντονες, όπως σε αυτήν την περίπτωση όπου πρόκειται για περιπτώσεις υγιούς ιστού και κακοήθων όγκων.

Ωστόσο, είναι σημαντικό να σημειωθεί ότι η κατηγοριοποίηση σε τρεις κλάσεις (3-class classification), αν και ως πρόβλημα αποτελεί μεγαλύτερη πρόκληση για τους αλγορίθμους, ωστόσο παρέχει πιο λεπτομερή και ορθή διάγνωση στην περίπτωση του καρκίνου και είναι σαφώς πιο χρήσιμη ιατρικά και ρεαλιστική σε κλινικές περιπτώσεις όπου η διάκριση μεταξύ καλοήθειας και κακοήθειας είναι κρίσιμη για την επιλογή της κατάλληλης θεραπείας.

7.4 Περαιτέρω Έρευνα

Για την περαιτέρω βελτίωση της απόδοσης του μοντέλου και την αντιμετώπιση των προκλήσεων που αναδείχθηκαν κατά την εκπαίδευση με ανισο-κατανομημένα δεδομένα, προτείνονται οι εξής κατευθύνσεις για μελλοντική έρευνα:

- 1. Αύξηση του Συνόλου Δεδομένων:** Συλλογή περισσότερων δεδομένων, ιδίως για τις κατηγορίες malignant και normal, ώστε να μειωθεί η ανισορροπία και να βελτιωθεί η απόδοση του μοντέλου.
- 2. Χρήση Μεθόδων Data Augmentation:** Εφαρμογή τεχνικών data augmentation για την τεχνητή αύξηση του μεγέθους του συνόλου δεδομένων και την αντιμετώπιση της ανισορροπίας αλλά και της έλλειψης των δεδομένων εκπαίδευσης.
- 3. Βελτιστοποίηση Υπερπαραμέτρων:** Εφαρμογή τεχνικών βελτιστοποίησης υπερπαραμέτρων (όπως grid search και random search) για την ανακάλυψη των βέλτιστων τιμών για τις παραμέτρους του μοντέλου.
- 4. Εξερεύνηση Διαφορετικών Αρχιτεκτονικών:** Πειραματισμός με διάφορες αρχιτεκτονικές νευρωνικών δικτύων, όπως πιο σύνθετα CNNs ή ακόμα και υβριδικά μοντέλα, για τη βελτίωση της κατηγοριοποίησης.
- 5. Ενσωμάτωση Τεχνικών Βαθιάς Μάθησης:** Διερεύνηση της χρήσης τεχνικών βαθιάς μάθησης, όπως GANs (Generative Adversarial Networks), για τη δημιουργία συνθετικών εικόνων που μπορούν να χρησιμοποιηθούν για την εκπαίδευση του μοντέλου.
- 6. Εφαρμογή Μεθόδων Transfer Learning:** Χρήση προ-εκπαιδευμένων μοντέλων και τεχνικών transfer learning για τη βελτίωση της απόδοσης του μοντέλου σε περιορισμένα σύνολα δεδομένων.

Βιβλιογραφία – Αναφορές – Διαδικτυακές Πηγές

- [1] “Breast ultrasound Images Dataset,” *Kaggle*, Mar. 14, 2021. <https://www.kaggle.com/datasets/aryashah2k/breast-ultrasound-images-dataset>
- [2] R. Goldman, “History of breast cancer,” *Healthline*, Dec. 30, 2020. <https://www.healthline.com/health/history-of-breast-cancer>
- [3] J. H. M. Mfa Mph, “A comprehensive guide to breast cancer,” *Healthline*, Jan. 30, 2023. <https://www.healthline.com/health/breast-cancer#symptoms>
- [4] News-Medical, “History of breast cancer,” *News-Medical*, Feb. 26, 2019. <https://www.news-medical.net/health/History-of-Breast-Cancer.aspx>
- [5] News-Medical, “Breast cancer symptoms,” *News-Medical*, Jun. 23, 2023. <https://www.news-medical.net/health/Breast-Cancer-Symptoms.aspx>
- [6] News-Medical, “Breast cancer causes,” *News-Medical*, Jun. 22, 2023. <https://www.news-medical.net/health/Breast-Cancer-Causes.aspx>
- [7] News-Medical, “What is Breast Cancer?” *News-Medical*, Jun. 23, 2023. <https://www.news-medical.net/health/What-is-Breast-Cancer.aspx>
- [8] L. Eldridge MD, “Cancer Cells vs. Normal Cells: How Are They Different?” *Verywell Health*, Apr. 21, 2023. <https://www.verywellhealth.com/cancer-cells-vs-normal-cells-2248794>
- [9] S. Saxena, “Introduction to Softmax activation function for neural network,” *Analytics Vidhya*, Jun. 14, 2024. <https://www.analyticsvidhya.com/blog/2021/04/introduction-to-softmax-for-neural-network/>
- [10] “Create simple image classification network - MATLAB & Simulink.” <https://www.mathworks.com/help/deeplearning/gs/create-simple-deep-learning-classification-network.html>
- [11] “Softmax layer - MATLAB.” <https://www.mathworks.com/help/deeplearning/ref/nnet.cnn.layer.softmaxlayer.html>
- [12] “Classification output layer - MATLAB classificationLayer.” <https://www.mathworks.com/help/deeplearning/ref/classificationlayer.html>
- [13] “Max pooling layer - MATLAB.” <https://www.mathworks.com/help/deeplearning/ref/nnet.cnn.layer.maxpooling2dlayer.html>
- [14] “Batch normalization layer - MATLAB.” <https://www.mathworks.com/help/deeplearning/ref/nnet.cnn.layer.batchnormalizationlayer.html>
- [15] “High Pass filter fft2 on the coin image,” *MATLAB Answers - MATLAB Central*. <https://www.mathworks.com/matlabcentral/answers/429951-high-pass-filter-fft2-on-the-coin-image>
- [16] “How to recover image from IFFT after removing high frequency descri...” <https://www.mathworks.com/matlabcentral/answers/488662-how-to-recover-image-from-iff-after-removing-high-frequency-descriptors>

- [17] “Visualize activations of a convolutional neural Network - MATLAB & Simulink - MathWorks España.” <https://es.mathworks.com/help/deeplearning/ug/visualize-activations-of-a-convolutional-neural-network.html>
- [18] “Image types & data classes,” *Technology Robotix Society, IIT Kharagpur*. <https://www.robotix.in/tutorial/imageprocessing/imagetypes/>
- [19] “What is deep learning? | IBM.” <https://www.ibm.com/topics/deep-learning>
- [20] GeeksforGeeks, “Confusion Matrix in machine learning,” *GeeksforGeeks*, May 13, 2024. <https://www.geeksforgeeks.org/confusion-matrix-machine-learning/>
- [21] R. C. Gonzalez, *Digital Image Processing*, 2nd ed.
- [22] R. C. Gonzalez, R. E. Woods, and S. L. Eddins, *Digital image processing using MATLAB*. 2020.
- [23] A. K. Jain, *Fundamentals of digital image processing*. 1988. [Online]. Available: <http://ci.nii.ac.jp/ncid/BA05045393>
- [24] I. N. Έλληνας, *Ψηφιακή Επεξεργασία Εικόνας & Βίντεο: Από τη Θεωρία στην Πράξη*.
- [25] A. Αλεξανδρίδης and H. Ζώης, “UNIWA Open eClass | Υπολογιστική Νοημοσύνη.” <https://eclass.uniwa.gr/courses/EE140/>
- [26] Y. Kimori, “Morphological image processing for quantitative shape analysis of biomedical structures: effective contrast enhancement,” *Journal of Synchrotron Radiation*, vol. 20, no. 6, pp. 848–853, Sep. 2013, doi: 10.1107/s0909049513020761.
- [27] L. Yang *et al.*, “Rapid segmentation and diagnosis of breast tumor ultrasound images at the sonographer level using deep learning,” *Bioengineering*, vol. 10, no. 10, p. 1220, Oct. 2023, doi: 10.3390/bioengineering10101220.
- [28] A. R. W. Sait and R. Nagaraj, “An enhanced LightGBM-Based breast cancer detection technique using mammography images,” *Diagnostics*, vol. 14, no. 2, p. 227, Jan. 2024, doi: 10.3390/diagnostics14020227.
- [29] H. Zhang and Y. Qie, “Applying Deep Learning to Medical Imaging: A review,” *Applied Sciences*, vol. 13, no. 18, p. 10521, Sep. 2023, doi: 10.3390/app131810521.
- [30] D. L. Rubin, H. Greenspan, and J. F. Brinkley, “Biomedical Imaging Informatics,” in *Springer eBooks*, 2013, pp. 285–327. doi: 10.1007/978-1-4471-4474-8_9.
- [31] R. Szeliski, *Computer Vision: Algorithms and Applications*, 2nd ed. 2021.
- [32] M. Radak, H. Y. Lafta, and H. Fallahi, “Machine learning and deep learning techniques for breast cancer diagnosis and classification: a comprehensive review of medical imaging studies,” *Journal of Cancer Research and Clinical Oncology*, vol. 149, no. 12, pp. 10473–10491, Jun. 2023, doi: 10.1007/s00432-023-04956-z.
- [33] K. L. D. Santos and M. P. D. S. Silva, “Deep Cross-Training: An approach to improve deep neural network classification on mammographic images,” *Expert Systems With Applications*, vol. 238, p. 122142, Mar. 2024, doi: 10.1016/j.eswa.2023.122142.

- [34] V. Rathinam, R. Sasireka, and K. Valarmathi, “An Adaptive Fuzzy C-Means segmentation and deep learning model for efficient mammogram classification using VGG-Net,” *Biomedical Signal Processing and Control*, vol. 88, p. 105617, Feb. 2024, doi: 10.1016/j.bspc.2023.105617.
- [35] R. Jalloul, H. K. Chethan, and R. Alkhatib, “A Review of Machine Learning Techniques for the Classification and Detection of Breast Cancer from Medical Images,” *Diagnostics*, vol. 13, no. 14, p. 2460, Jul. 2023, doi: 10.3390/diagnostics13142460.
- [36] F.-F. Yao and Y. Zhang, “A review of quantitative diffusion-weighted MR imaging for breast cancer: Towards noninvasive biomarker,” *Clinical Imaging*, vol. 98, pp. 36–58, Jun. 2023, doi: 10.1016/j.clinimag.2023.03.016.
- [37] X. Zhao, J.-W. Bai, Q. Guo, K. Ren, and G.-J. Zhang, “Clinical applications of deep learning in breast MRI,” *Biochimica Et Biophysica Acta. Reviews on Cancer/Biochimica Et Biophysica Acta, CR. Reviews on Cancer*, vol. 1878, no. 2, p. 188864, Mar. 2023, doi: 10.1016/j.bbcan.2023.188864.
- [38] R. Ranjbarzadeh *et al.*, “Breast tumor localization and segmentation using machine learning techniques: Overview of datasets, findings, and methods,” *Computers in Biology and Medicine*, vol. 152, p. 106443, Jan. 2023, doi: 10.1016/j.combiomed.2022.106443.

Παράρτημα Α – Προ-επεξεργασία των δεδομένων

Κώδικας Δημιουργίας Φιλτραρισμένων Εικόνων

```
% Αρχικοποίηση φακέλων
inputFolder = './Dataset_BUSI_with_GT_no_masks\benign';
outputFolder = './fft2_lowgaussian\benign';

% Δημιουργία φακέλου εξόδου εάν δεν υπάρχει
if ~exist(outputFolder, 'dir')
    mkdir(outputFolder);
end

% Ανάγνωση όλων των αρχείων από τον φάκελο
imageFiles = dir(fullfile(inputFolder, '*.png'));

for i = 1:length(imageFiles)
    inputImagePath = fullfile(inputFolder, imageFiles(i).name);
    img = imread(inputImagePath);
    img = rgb2gray(img);

    % Μετασχηματισμός Fourier 2D
    F = fft2(double(img));

    % Μετακίνηση του μηδενικού συχνοτικού συντελεστή στο κέντρο
    F_shifted = fftshift(F);

    % Δημιουργία γκαουσιανού φίλτρου
    [M, N] = size(F_shifted);
    [X, Y] = meshgrid(-floor(N/2):floor((N-1)/2), -floor(M/2):floor((M-1)/2));
    sigma = 10; % Παράμετρος σ του φίλτρου
    gaussian_filter = exp(-(X.^2 + Y.^2) / (2*sigma^2)); %low pass

    % Εφαρμογή του φίλτρου στο πεδίο της συχνότητας
    F_filtered = F_shifted .* gaussian_filter;

    % Αντίστροφος μετασχηματισμός Fourier
    F_inv_shifted = ifftshift(F_filtered);
    img_filtered = ifft2(F_inv_shifted);

    % Πραγματικό μέρος της εικόνας
    img_filtered = real(img_filtered);

    % Κανονικοποίηση της εικόνας (τιμές μεταξύ 0 και 255)
    img_filtered = uint8(255 * mat2gray(img_filtered));

    % Αποθήκευση της φιλτραρισμένης εικόνας
    outputImagePath = fullfile(outputFolder, imageFiles(i).name);
    imwrite(img_filtered, outputImagePath);
end
```

Εντολές για οπτικοποίηση των αποτελεσμάτων

Χρησιμοποιείται στο command window, κατά την εκτέλεση με step, μέσα στον βρόγχο επανάληψης του παραπάνω κώδικα

```
% Οπτικοποίηση του magnitude του Fourier
```

```
magnitude = log(abs(F_shifted) + 1);  
% Οπτικοποίηση της φιλτραρισμένης εικόνας  
filtered_magnitude = log(abs(F_filtered) + 1);  
figure;  
  
subplot(1, 3, 1);  
imshow(magnitude, []);  
title('Magnitude of Fourier Transform');  
  
subplot(1, 3, 2);  
imshow(gaussian_filter, []);  
title('Gaussian Filter');  
  
subplot(1, 3, 3);  
imshow(filtered_magnitude, []);  
title('Magnitude of Filtered Fourier Transform');
```

Κώδικας πρόσθεσης εικόνων

```
% Αρχικοποίηση φακέλων  
inputFolder1 = './fft2_lowgaussian\benign';  
inputFolder2 = './Dataset_BUSI_with_GT_no_masks\benign';  
outputFolder = './added_images\benign';  
  
% Δημιουργία φακέλου εξόδου εάν δεν υπάρχει  
if ~exist(outputFolder, 'dir')  
    mkdir(outputFolder);  
end  
  
% Ανάγνωση όλων των αρχείων εικόνας από τους φακέλους εισόδου  
imageFiles1 = dir(fullfile(inputFolder1, '*.png'));  
imageFiles2 = dir(fullfile(inputFolder2, '*.png'));  
  
% Αριθμός αρχείων σε κάθε φάκελο  
numImages1 = numel(imageFiles1);  
numImages2 = numel(imageFiles2);  
  
% Επανάληψη για κάθε εικόνα στους δύο φακέλους  
for i = 1:numImages1  
    inputImagePath1 = fullfile(inputFolder1, imageFiles1(i).name);  
    inputImagePath2 = fullfile(inputFolder2, imageFiles2(i).name);  
    img1 = imread(inputImagePath1);  
    img2 = imread(inputImagePath2);  
    % img1 = rgb2gray(img1); % είναι ήδη  
    img2 = rgb2gray(img2);  
  
    % Πρόσθεση των δύο εικόνων  
    added_img = imadd(img1, img2);  
  
    % Αποθήκευση της νέας εικόνας  
    outputImagePath = fullfile(outputFolder, imageFiles1(i).name);  
    imwrite(added_img, outputImagePath);  
end
```

Συνάρτηση shuffleImageDatastore

```
function shuffledDatastore = shuffleImageDatastore(datastore)  
    % Get the number of files in the datastore  
    numFiles = numel(datastore.Files);
```

```
% Get the current files and labels
currentFiles = datastore.Files;
currentLabels = datastore.Labels;

% Generate a random permutation
perm = randperm(numFiles);

% Shuffle the files and labels using the permutation
shuffledFiles = currentFiles(perm);
shuffledLabels = currentLabels(perm);

% Create a new datastore with shuffled files and labels
shuffledDatastore = copy(datastore);
shuffledDatastore.Files = shuffledFiles;
shuffledDatastore.Labels = shuffledLabels;
end
```

Δημιουργία και Σύγκριση Μασκών

```
% Load image and mask
imageFilePath = './Dataset_BUSI_with_GT/malignant/malignant (1).png';
realMaskFilePath = './Dataset_BUSI_with_GT/malignant/malignant (1)_mask.png';

image = imread(imageFilePath);
grayImage = rgb2gray(image);
realMask = imread(realMaskFilePath);

% Generate mask
threshold = 65; % Otsu: threshold = graythresh(grayImage);
generatedMask = grayImage < threshold; % Otsu: generatedMask = ~imbinarize(grayImage,
threshold);

% Median Filter and CLAHE (Contrast Limited Adaptive Histogram Equalization) to enhance
contrast
filteredImage = medfilt2(grayImage, [15, 15]);
claheImage = adapthisteq(filteredImage);

% Generate improved mask
thresholdNew = 90; % Otsu: thresholdNew = graythresh(filteredImage);
improvedMask = claheImage < thresholdNew; % Otsu: improvedMask = ~imbinarize(filteredImage,
thresholdNew);

% Compare the 2 masks with the real mask (XOR)
resultXORGenerated = xor(generatedMask, realMask);
resultXORImproved = xor(improvedMask, realMask);

% Percentage of matching pixels
matchingPixelsGenerated = sum(resultXORGenerated(:) == 0);
totalPixelsGenerated = numel(generatedMask);
similarityScoreGenerated = (matchingPixelsGenerated / totalPixelsGenerated) * 100;

matchingPixelsImproved = sum(resultXORImproved(:) == 0);
totalPixelsImproved = numel(improvedMask);
similarityScoreImproved = (matchingPixelsImproved / totalPixelsImproved) * 100;

% Display image and mask comparisons
figure;
subplot(2, 4, 1), imshow(image), title('Original Image');
```

```
subplot(2, 4, 2), imshow(generatedMask), title('Generated Mask');
subplot(2, 4, 3), imshow(resultXORGenerated), title('XOR (Generated)');
subplot(2, 4, 4), imshow(realMask), title('Real Mask');
subplot(2, 4, 5), imshow(claheImage), title('CLAHE Enhanced Image');
subplot(2, 4, 6), imshow(improvedMask), title('Improved Mask');
subplot(2, 4, 7), imshow(resultXORImproved), title('XOR (Improved)');
subplot(2, 4, 8), imshow(realMask), title('Real Mask');

% Print similarity scores
fprintf('Similarity Score (Generated Mask): %.2f %%\n', similarityScoreGenerated);
fprintf('Threshold (Generated Mask): %d\n', threshold);
fprintf('Similarity Score (Improved Mask): %.2f %%\n', similarityScoreImproved);
fprintf('Threshold (Improved Mask): %d\n', thresholdNew);
```

Παράρτημα Β – Εκπαίδευση Νευρωνικών Δικτύων

Κώδικας εκπαίδευσης του FNN

```
% Path to dataset
datasetPath = './Dataset_BUSI_with_GT_no_masks';
% datasetPath = './fft2_lowgaussian';
% datasetPath = './added_images';

% Create datastores for each class
normalDatastore = imageDatastore(fullfile(datasetPath, 'normal'), 'LabelSource',
'foldernames', 'IncludeSubfolders', true);
benignDatastore = imageDatastore(fullfile(datasetPath, 'benign'), 'LabelSource',
'foldernames', 'IncludeSubfolders', true);
malignantDatastore = imageDatastore(fullfile(datasetPath, 'malignant'), 'LabelSource',
'foldernames', 'IncludeSubfolders', true);

% Shuffle image order in each datastore
rng(1); % Set a random seed for reproducibility
normalDatastore = shuffleImageDatastore(normalDatastore);
benignDatastore = shuffleImageDatastore(benignDatastore);
malignantDatastore = shuffleImageDatastore(malignantDatastore);

% Split the data into training, validation, and testing sets
[trainNormal, valNormal, testNormal] = splitEachLabel(normalDatastore, 0.7, 0.15, 0.15);
[trainBenign, valBenign, testBenign] = splitEachLabel(benignDatastore, 0.7, 0.15, 0.15);
[trainMalignant, valMalignant, testMalignant] = splitEachLabel(malignantDatastore, 0.7,
0.15, 0.15);

% Combine the datastores for all classes
trainData = imageDatastore(cat(1, trainNormal.Files, trainBenign.Files,
trainMalignant.Files), 'LabelSource', 'foldernames');
valData = imageDatastore(cat(1, valNormal.Files, valBenign.Files, valMalignant.Files),
'LabelSource', 'foldernames');
testData = imageDatastore(cat(1, testNormal.Files, testBenign.Files, testMalignant.Files),
'LabelSource', 'foldernames');

% Specify the desired input size
inputSize = [224, 224, 1];

% Shuffle image order for augmented/combined datastores
trainData = shuffleImageDatastore(trainData);
valData = shuffleImageDatastore(valData);
```

```
testData = shuffleImageDatastore(testData);

% Get the true labels from the testDatastore
trueLabels = testData.Labels;

% Preprocess and resize the images
trainData = augmentedImageDatastore(inputSize(1:2), trainData, 'ColorPreprocessing',
'rgb2gray');
valData = augmentedImageDatastore(inputSize(1:2), valData, 'ColorPreprocessing',
'rgb2gray');
testData = augmentedImageDatastore(inputSize(1:2), testData, 'ColorPreprocessing',
'rgb2gray');

% FNN Architecture
layers = [
    imageInputLayer([224 224 1])
    fullyConnectedLayer(128)
    reluLayer
    fullyConnectedLayer(256)
    reluLayer
    fullyConnectedLayer(512)
    reluLayer
    fullyConnectedLayer(1024)
    reluLayer
    fullyConnectedLayer(512)
    reluLayer
    fullyConnectedLayer(256)
    reluLayer
    fullyConnectedLayer(128)
    reluLayer
    fullyConnectedLayer(64)
    reluLayer
    fullyConnectedLayer(32)
    reluLayer
    fullyConnectedLayer(3)
    softmaxLayer
    classificationLayer
];

% Training options
options = trainingOptions('sgdm', ...
    'MaxEpochs', 20, ...
    'MiniBatchSize', 16, ...
    'Shuffle','every-epoch', ...
    'InitialLearnRate', 1e-05, ...
    'ValidationData', valData, ...
    'ValidationFrequency', 10, ...
    'Plots', 'training-progress', ...
    'Verbose', true, ...
    'VerboseFrequency', 5);

% Train the network
net = trainNetwork(trainData, layers, options);

% Evaluate the CNN on the test set
[predictedLabels, scores] = classify(net, testData);

% Calculate accuracy
accuracy = sum(predictedLabels == trueLabels) / numel(trueLabels)
```



```
% Plot Confusion Matrix
figure();
plotconfusion(trueLabels, predictedLabels)
```

Κώδικας εκπαίδευσης του CNN #1

```
% Path to dataset
datasetPath = './Dataset_BUSI_with_GT_no_masks';
% datasetPath = './fft2_lowgaussian';
% datasetPath = './added_images';

% Create datastores for each class
normalDatastore = imageDatastore(fullfile(datasetPath, 'normal'), 'LabelSource',
'foldernames', 'IncludeSubfolders', true);
benignDatastore = imageDatastore(fullfile(datasetPath, 'benign'), 'LabelSource',
'foldernames', 'IncludeSubfolders', true);
malignantDatastore = imageDatastore(fullfile(datasetPath, 'malignant'), 'LabelSource',
'foldernames', 'IncludeSubfolders', true);

% Shuffle image order in each datastore
rng(1); % Set a random seed for reproducibility
normalDatastore = shuffleImageDatastore(normalDatastore);
benignDatastore = shuffleImageDatastore(benignDatastore);
malignantDatastore = shuffleImageDatastore(malignantDatastore);

% Split the data into training, validation, and testing sets
[trainNormal, valNormal, testNormal] = splitEachLabel(normalDatastore, 0.7, 0.15, 0.15);
[trainBenign, valBenign, testBenign] = splitEachLabel(benignDatastore, 0.7, 0.15, 0.15);
[trainMalignant, valMalignant, testMalignant] = splitEachLabel(malignantDatastore, 0.7,
0.15, 0.15);

% Combine the datastores for all classes
trainData = imageDatastore(cat(1, trainNormal.Files, trainBenign.Files,
trainMalignant.Files), 'LabelSource', 'foldernames');
valData = imageDatastore(cat(1, valNormal.Files, valBenign.Files, valMalignant.Files),
'LabelSource', 'foldernames');
testData = imageDatastore(cat(1, testNormal.Files, testBenign.Files, testMalignant.Files),
'LabelSource', 'foldernames');

% Specify the desired input size
inputSize = [224, 224, 1];

% Shuffle image order for augmented/combined datastores
trainData = shuffleImageDatastore(trainData);
valData = shuffleImageDatastore(valData);
testData = shuffleImageDatastore(testData);

% Get the true labels from the testDatastore
trueLabels = testData.Labels;

% Preprocess and resize the images
trainData = augmentedImageDatastore(inputSize(1:2), trainData, 'ColorPreprocessing',
'rgb2gray');
valData = augmentedImageDatastore(inputSize(1:2), valData, 'ColorPreprocessing',
'rgb2gray');
testData = augmentedImageDatastore(inputSize(1:2), testData, 'ColorPreprocessing',
'rgb2gray');
```

```
% CNN architecture
layers = [
    imageInputLayer([224 224 1])

    convolution2dLayer(3,32, 'Padding', 'same')
    batchNormalizationLayer
    reluLayer

    maxPooling2dLayer(2, 'Stride', 2)

    convolution2dLayer(3,64, 'Padding', 'same')
    batchNormalizationLayer
    reluLayer

    maxPooling2dLayer(2, 'Stride', 2)

    convolution2dLayer(3,128, 'Padding', 'same')
    batchNormalizationLayer
    reluLayer

    fullyConnectedLayer(3)
    softmaxLayer
    classificationLayer('Classes', categorical({'benign', 'malignant', 'normal'}), ...
        'ClassWeights', [1, 2.08, 3.29])
];

% Training options
options = trainingOptions('sgdm', ...
    'MaxEpochs', 15, ...
    'MiniBatchSize', 32, ...
    'Shuffle', 'every-epoch', ...
    'InitialLearnRate', 1e-04, ...
    'ValidationData', valData, ...
    'ValidationFrequency', 5, ...
    'Plots', 'training-progress', ...
    'Verbose', true, ...
    'VerboseFrequency', 5);

% Train the CNN
net = trainNetwork(trainData, layers, options);

% Evaluate the CNN on the test set
[predictedLabels, scores] = classify(net, testData);

% Calculate accuracy
accuracy = sum(predictedLabels == trueLabels) / numel(trueLabels)

% Plot Confusion Matrix
figure();
plotconfusion(trueLabels, predictedLabels)
```

Κώδικας εκπαίδευσης του CNN #2

```
% Path to dataset
datasetPath = './Dataset_BUSI_with_GT_no_masks';
% datasetPath = './fft2_lowgaussian';
% datasetPath = './added_images';

% Create datastores for each class
```

```
normalDatastore = imageDatastore(fullfile(datasetPath, 'normal'), 'LabelSource',  
'foldernames', 'IncludeSubfolders', true);  
benignDatastore = imageDatastore(fullfile(datasetPath, 'benign'), 'LabelSource',  
'foldernames', 'IncludeSubfolders', true);  
malignantDatastore = imageDatastore(fullfile(datasetPath, 'malignant'), 'LabelSource',  
'foldernames', 'IncludeSubfolders', true);  
  
% Shuffle image order in each datastore  
rng(1); % Set a random seed for reproducibility  
normalDatastore = shuffleImageDatastore(normalDatastore);  
benignDatastore = shuffleImageDatastore(benignDatastore);  
malignantDatastore = shuffleImageDatastore(malignantDatastore);  
  
% Split the data into training, validation, and testing sets  
[trainNormal, valNormal, testNormal] = splitEachLabel(normalDatastore, 0.7, 0.15, 0.15);  
[trainBenign, valBenign, testBenign] = splitEachLabel(benignDatastore, 0.7, 0.15, 0.15);  
[trainMalignant, valMalignant, testMalignant] = splitEachLabel(malignantDatastore, 0.7,  
0.15, 0.15);  
  
% Combine the datastores for all classes  
trainData = imageDatastore(cat(1, trainNormal.Files, trainBenign.Files,  
trainMalignant.Files), 'LabelSource', 'foldernames');  
valData = imageDatastore(cat(1, valNormal.Files, valBenign.Files, valMalignant.Files),  
'LabelSource', 'foldernames');  
testData = imageDatastore(cat(1, testNormal.Files, testBenign.Files, testMalignant.Files),  
'LabelSource', 'foldernames');  
  
% Specify the desired input size  
inputSize = [224, 224, 1];  
  
% Shuffle image order for augmented/combined datastores  
trainData = shuffleImageDatastore(trainData);  
valData = shuffleImageDatastore(valData);  
testData = shuffleImageDatastore(testData);  
  
% Get the true labels from the testDatastore  
trueLabels = testData.Labels;  
  
% Preprocess and resize the images  
trainData = augmentedImageDatastore(inputSize(1:2), trainData, 'ColorPreprocessing',  
'rgb2gray');  
valData = augmentedImageDatastore(inputSize(1:2), valData, 'ColorPreprocessing',  
'rgb2gray');  
testData = augmentedImageDatastore(inputSize(1:2), testData, 'ColorPreprocessing',  
'rgb2gray');  
  
% CNN architecture  
layers = [  
    imageInputLayer([224 224 1])  
  
    convolution2dLayer(5,32, 'Padding', 'same')  
    batchNormalizationLayer  
    reluLayer  
  
    maxPooling2dLayer(2, 'Stride', 2)  
  
    convolution2dLayer(5,64, 'Padding', 'same')  
    batchNormalizationLayer  
    reluLayer
```

```
maxPooling2dLayer(2, 'Stride', 2)

convolution2dLayer(3, 64, 'Padding', 'same')
batchNormalizationLayer
reluLayer

maxPooling2dLayer(2, 'Stride', 2)

convolution2dLayer(3, 128, 'Padding', 'same')
batchNormalizationLayer
reluLayer

maxPooling2dLayer(2, 'Stride', 2)

convolution2dLayer(3, 256, 'Padding', 'same')
batchNormalizationLayer
reluLayer
dropoutLayer(0.1)

maxPooling2dLayer(2, 'Stride', 2)

convolution2dLayer(3, 512, 'Padding', 'same')
batchNormalizationLayer
reluLayer
dropoutLayer(0.1)

fullyConnectedLayer(512)
reluLayer
fullyConnectedLayer(256)
reluLayer
fullyConnectedLayer(64)
reluLayer
fullyConnectedLayer(3)
softmaxLayer
classificationLayer('Classes', categorical({'benign', 'malignant', 'normal'}), ...
                    'ClassWeights', [1, 4.16, 6.58])
];

% Training options
options = trainingOptions('sgdm', ...
    'MaxEpochs', 15, ...
    'MiniBatchSize', 32, ...
    'Shuffle', 'every-epoch', ...
    'InitialLearnRate', 1e-04, ...
    'ValidationData', valData, ...
    'ValidationFrequency', 5, ...
    'Plots', 'training-progress', ...
    'Verbose', true, ...
    'VerboseFrequency', 5);

% Train the CNN
net = trainNetwork(trainData, layers, options);

% Evaluate the CNN on the test set
[predictedLabels, scores] = classify(net, testData);

% Calculate accuracy
accuracy = sum(predictedLabels == trueLabels) / numel(trueLabels)

% Plot Confusion Matrix
```

```
figure();  
plotconfusion(trueLabels, predictedLabels)
```