



ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ Η/Υ ΚΑΙ ΠΛΗΡΟΦΟΡΙΚΗΣ  
ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ  
ΥΠΟΛΟΓΙΣΤΩΝ  
ΤΜΗΜΑ ΜΗΧΑΝΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ & ΑΕΡΟΝΑΠΥΓΗΩΝ

ΑΥΤΟΜΑΤΗ ΔΙΑΓΝΩΣΗ ΝΕΟΠΛΑΣΙΩΝ  
ΣΕ ΛΟΞΕΣ ΚΑΙ ΠΛΑΓΙΕΣ ΠΡΟΒΟΛΕΣ  
ΜΑΣΤΟΓΡΑΦΙΑΣ

ΣΥΣΤΗΜΑΤΑ ΥΠΟΒΟΗΘΗΣΗΣ ΙΑΤΡΙΚΗΣ ΔΙΑΓΝΩΣΗΣ

Παναγιώτης Ηλίας  
Α.Μ.:1019362

ΕΠΙΒΛΕΠΩΝ: Καθηγητής Ευάγγελος Δερματάς

ΠΑΤΡΑ 2020

## Πρόλογος

Η παρούσα διπλωματική εκπονήθηκε στο πλαίσιο της του διατμηματικού μεταπτυχιακού προγράμματος (ΣΜΗΝ) Συστήματα Επεξεργασίας Πληροφορίας και Μηχανική Νοημοσύνη, οπού αποτέλεσε το 1/3 της συνολικής βαθμολογίας του Πτυχίου.

Αρχικά, θα ήθελα να ευχαριστήσω όσους συνέβαλλαν άμεσα η έμμεσα στην ολοκλήρωση της διπλωματικής εργασίας ιδιαίτερα:

Τον Αναπληρωτή Καθηγητή κ.ε. Δερματά, που είναι υπεύθυνος του τομέα Τηλεπικοινωνιών και Τεχνολογία Πληροφορίας του Τμήματος Ηλεκτρολόγων Μηχανικών και Τεχνολογίας Υπολογιστών, που μου έδωσε την ευκαιρία και την δυνατότητα να ασχοληθώ με ένα τόσο ενδιαφέρον και επίκαιρο θέμα.

Την Διδακτορικό κ. Ιωάννα .Χριστογιάννη του Τμήματος Ηλεκτρολόγων Μηχανικών και Τεχνολογίας Υπολογιστών,, μέλος της τριμελούς εξεταστικής επιτροπής, που μου με βοήθησε πολύ στο θέμα, παρέχοντάς μου την κατάλληλη βιβλιογραφία και τις σωστές κατευθύνσεις.

Τον Διδακτορικό κ. Αθανάσιο Κούτρα του Τμήματος Ηλεκτρολόγων Μηχανικών και Τεχνολογίας Υπολογιστών, για της χρήσιμες συμβουλές και για την πολύτιμη βοήθεια του στην υλοποίηση της παρούσας διπλωματικής.

Τον καθηγητή κ. Κωσταντίνο Μπερμπερίδη που είναι αντιπρόεδρος στο Τμήμα Μηχανικών Η/Υ και Πληροφορικής, επικεφαλής του τομέα επεξεργασίας σήματος και επικοινωνίας, για τις χρήσιμες παρατηρήσεις του στην εργασία.

Την οικογένεια μου για την βοήθεια και την στήριξη σε όλη την διάρκεια της μεταπτυχιακής διπλωματικής εργασίας.

## Περίληψη

Στη συγκεκριμένη μεταπτυχιακή διπλωματική εργασία παρουσιάζεται ένα ολοκληρωμένο υπολογιστικό σύστημα υποβοήθησης (Computer-Aided System) σχετικά με την διάγνωση του καρκίνου του μαστού, βασισμένο στα νευρωνικά δίκτυα και την εκπαίδευση με εποπτεία. Ως κύρια βάση δεδομένων για την παραπάνω διαδικασία του έλαβε χώρα, αξιοποιήθηκε η DDSM – CBIS, μίας ευρέως γνωστή και έγκυρη πηγή στοιχείων. Αρχικά, πραγματοποιήθηκε διαχωρισμός, αξιοποιώντας τις διαφορετικές προβολές μαστογραφίας, για το εκάστοτε στήθος του αντίστοιχου ασθενή, των δύο βασικών κατηγοριών καρκινικών όγκων, δηλαδή, καλοήγη και κακοήγη. Για την παραπάνω διαδικασία, τέθηκαν σε εφαρμογή μία ποικιλία αρχιτεκτονικών συνελκτικών νευρωνικών δικτύων, με σκοπό την αξιολόγηση και την κατασκευή του βέλτιστου δυνατού κατηγοροποιητή. Μετά την επιτυχημένη δημιουργία της παραπάνω διάταξης, έλαβε χώρα η δημιουργία ενός συστήματος με σκοπό την παροχή μίας ευρύτερης διάγνωσης για το άτομο που είχε λάβει υποβληθεί στην εξέταση, βασιζόμενη πάνω στην διάγνωση της εκάστοτε προβολής μαστογραφίας. Για την συγκεκριμένη διεργασία, αξιοποιήθηκαν shallow νευρωνικά δίκτυα.

## Abstract

This postgraduate thesis presents a comprehensive system that acts as Computer Aided System (C.A.D) in order to provide an assistive prediction, regarding breast cancer diagnose procedure. This solution was rendered possible by using deep learning, image based supervised classification. As the main dataset for training, validation and testing procedures, DDSM-CBIS, a widely used, database, with size over 152 GB, was deployed. First, we managed to identify whether mass cases were malignant or benign, by training our system with both lateral and oblique mammographic projections. The main training architecture that we deployed was convolutional neural networks for image classification. Finally, we developed a system which evaluates the predictions for each mammographic view and provides a general estimation regarding the malignance of the cancer mass. To do that, the previous estimations are reviewed by shallow neural networks which provide the final estimation for the respective patient.

# Πίνακας Περιεχομένων

Πρόλογος.....	2
Περίληψη .....	3
Abstract.....	4
Εισαγωγή .....	9
Οργάνωση και Πρότυπα Σημεία της Διπλωματικής .....	11
Πίνακας Εικόνων .....	13
<b>Κεφάλαιο 1<sup>ο</sup></b> .....	<b>16</b>
Γενικά για τον Καρκίνο .....	16
1.1 Αίτιες καρκίνου.....	19
1.2 Τύποι καρκίνου .....	19
1.3 Καρκίνος του μαστού: Αιτιολογία και παράγοντες κινδύνου .....	21
1.4 Τύποι καρκίνου του μαστού .....	24
1.5 Διαγνωστικές τεχνικές για τον καρκίνο του μαστού .....	25
1.6 Μέθοδοι αντιμετώπισης του καρκίνου .....	31
1.7 Στατιστικά για τον καρκίνο του μαστού .....	33
Βιβλιογραφία 1 <sup>ου</sup> Κεφαλαίου.....	34
<b>Κεφάλαιο 2ο.....</b>	<b>38</b>
Απεικόνιση του Καρκίνου του Μαστού - Μαστογραφία .....	38
2.1 Μαστογραφία.....	38
2.2 Βασική αρχή της μαστογραφίας .....	38
2.3 Εξοπλισμός μαστογράφου .....	39
2.4 Λυχνία ακτίνων Χ, φίλτρα και φάσμα .....	40
2.5 Αντιδιαχυτικό διάφραγμα .....	42
2.6 Ανιχνευτές.....	42
2.7 Σύστημα συμπίεσης μαστού .....	43
2.8 Σύστημα αυτόματης έκθεσης (AEC) .....	44
2.9 Ψηφιακή Μαστογραφία .....	44
2.10 Ψηφιακή Τομοσύνθεση μαστού.....	47
2.11 Μαστογραφία: Κίνδυνοι και εσφαλμένες αντιλήψεις .....	49
Βιβλιογραφία 2 <sup>ου</sup> Κεφαλαίου.....	50
<b>Κεφάλαιο 3ο.....</b>	<b>53</b>
Ψηφιακή Επεξεργασία και Ανάλυση της Μαστογραφίας .....	53
3.1 Ψηφιακή Επεξεργασία μαστογραφίας .....	53
3.2 Τεχνικές βελτίωσης και αποκατάστασης μαστογραφίας .....	54

3.3 Το ιστόγραμμα μιας εικόνας.....	55
3.4 Επεξεργασία Ιστογράμματος .....	55
3.5 Τροποποίηση ιστογράμματος .....	56
3.6 Ισοστάθμιση Ιστογράμματος .....	58
3.7 Τεχνικές κατάτμησης της μαστογραφίας.....	59
3.8 Τεχνικές κατάτμησης με συνένωση, αύξηση και διαίρεση περιοχών .....	60
3.9 Ψηφιακή ανάλυση μαστογραφίας.....	61
3.10 Μορφολογικοί και γεωμετρικοί παράμετροι .....	63
Βιβλιογραφία 3 <sup>ου</sup> Κεφαλαίου .....	65
<b>Κεφάλαιο 4ο.....</b>	<b>67</b>
Νευρωνικά Δίκτυα .....	67
4.1 Ορισμός Νευρωνικού Δικτύου .....	68
4.2 Εκπαίδευση ενός Νευρωνικού δικτύου .....	69
4.3 Εφαρμογή των ANN .....	70
4.4 Πλεονεκτήματα Νευρωνικών Δικτύων.....	71
4.5 Κατηγορίες Νευρωνικών δικτύων .....	72
4.6 Πολυεπίπεδα Νευρωνικά δίκτυα Perceptron .....	72
4.7 Συνελκτικά νευρωνικά δίκτυα .....	78
4.8 Νευρωνικά δίκτυα ακτινικών συναρτήσεων.....	81
Βιβλιογραφία 4 <sup>ου</sup> Κεφαλαίου.....	86
<b>Κεφάλαιο 5<sup>ο</sup> .....</b>	<b>90</b>
Βαθιά Νευρωνικά Δίκτυα .....	90
5.1 Θεμελιώδεις έννοιες.....	92
5.2 Εκπαίδευση νευρωνικού δικτύου.....	94
5.3 Δομές DNN .....	95
5.4 Προβλήματα με Deep Neural Networks .....	96
5.5 Εφαρμογές βαθιάς μάθησης.....	96
Βιβλιογραφία 5 <sup>ου</sup> Κεφαλαίου.....	98
<b>Κεφάλαιο 6<sup>ο</sup> .....</b>	<b>103</b>
Σχετική θεωρία και έρευνες .....	103
6.1 Νευρωνικά Δίκτυα .....	103
6.1.4 AlexNet .....	111
6.2 Σύγκριση μεθόδων εκπαίδευσης.....	112
6.3 Transfer Learning.....	112
6.4 ImageNet.....	113

6.5 Δείκτες Αξιολόγησης .....	114
6.6 Άλλες Έννοιες .....	117
6.7 Έρευνες .....	118
<b>Κεφάλαιο 7ο.....</b>	<b>125</b>
Κατασκευή SingleView Συστήματος .....	125
7.1 DDSM .....	125
7.2 CBIS-DDSM.....	126
7.3 Διαδικασίες Προεπεξεργασίας.....	127
7.4 Εκπαίδευση SingleView Συστήματος.....	139
Βιβλιογραφία 7ου Κεφαλαίου .....	140
Κώδικας      141	
<b>Κεφάλαιο 8ο .....</b>	<b>148</b>
Κατασκευή Multiview-Fusion Συστήματος .....	150
8.1 Προεπεξεργασία.....	150
8.2 Εκπαίδευση Fusion Συστήματος.....	155
Βιβλιογραφία 8 <sup>ου</sup> Κεφαλαίου.....	157
Κώδικας      158	
<b>Κεφάλαιο 9ο .....</b>	<b>163</b>
Συζήτηση-Βελτιώσεις .....	163
<b>Επίλογος.....</b>	<b>164</b>

Page Intentionally Left Blank



## Εισαγωγή

Ο καρκίνος του μαστού αποτελεί την δεύτερη, περισσότερο συχνή αιτία θανάτου, για τον γυναικείο πληθυσμό. Η συγκεκριμένη νόσος, προσβάλλει ένα σημαντικό μέρος της συγκεκριμένης ομάδας ενώ οι πιθανότητες να αναπτυχθεί σε ένα άτομο, που φέρει τα παραπάνω χαρακτηριστικά, κυμαίνονται περίπου στο 10%. Η πιθανότητα, αυτή αυξάνεται σε συνάρτηση με την ηλικία καθώς και από άλλους παράγοντες. Η κυρίαρχη αίτια πρόκλησης του καρκίνου δεν είναι ακόμα σαφής, επομένως δεν υπάρχουν, ακόμη, ξεκάθαρες οδηγίες για την ασφαλής προφύλαξη εναντίων της συγκεκριμένης νόσου. Η έγκαιρη διάγνωση, στη σύγχρονη εποχή, παρουσιάζεται ως η κυριότερη μέθοδος αντιμετώπισης, με σημαντικά ποσοστά επιτυχία-ίαςης, προλαμβάνοντας την μετάσταση του καρκίνου σε άλλα μέρη του σώματος. Η έγκαιρη διάγνωση επιτυγχάνεται κυρίως με την Κλινική Εξέταση από εξειδικευμένο ιατρό και με την διαγνωστική μέθοδο της μαστογραφίας ενώ πολλές φορές συνδυάζεται και με άλλες διαγνωστικές τεχνικές,

Η μαστογραφία, είναι η καταγραφή των ακτινών Χ που διέρχονται από τον μαστό, είναι η πλέον αποτελεσματική, χαμηλού κόστους και υψηλής ευαισθησίας τεχνική για την ανίχνευση μικρών νεοπλασιών, οδηγώντας καταλυτικά στην μείωση των θανάτων κατά 30%. Επίσης, αποτελεί την πιο ασφαλή εξέταση για την ανίχνευση των βλαβών και της αρχιτεκτονικής τους. Αυτό δικαιολογείται από το γεγονός ότι η μαστογραφία πραγματοποιείται με χαμηλή ακτινοβολία καθιστώντας την εντελώς ακίνδυνη, για τον ασθενή.

Τα παθολογικά ευρήματα του καρκίνου του μαστού σχετίζονται με την καρκινική μάζα και την επίδραση που ασκεί ο καρκίνος στους γειτονικούς ιστού και παρουσιάζονται συνήθως ως:

- Οζώδης σκίαση με σαφή όρια.
- Οζώδης σκίαση με απότομα όρια σε σχήμα αστεριού.
- Οζώδης σκίαση με ακανόνιστο σχήμα και λιγότερα ευδιάκριτα όρια.
- Μικροαποτιτανώσεις, με ή χωρίς σκίαση.

Γενικά μπορούμε να τα ταξινομήσουμε τον καρκίνο σε καλοήθη και κακοήθη με αρκετές ενδιάμεσες περιπτώσεις. Από τα παραπάνω είδη καρκίνου τα πιο συχνά είναι και επικινδυνά είναι η οζώδης σκίαση με απότομα όρια σε σχήμα αστεριού, η οζώδης σκίαση με ευδιάκριτα όρια, και οι μικροαποτιτανώσεις με ή χωρίς σκίαση αποτελούν ευρήματα που αποτελούν ένδειξη για τον καρκίνο του μαστού.

Η διάγνωση του καρκίνου του μαστού πραγματοποιείται από ειδικούς ακτινολόγους, οι οποίοι εξετάζουν την δομή των ακτινογραφιών για ίχνη μη φυσιολογικών περιοχών, όμως η πολύπλοκη δομή και ποικιλομορφία τους αποτελεί εμποδίζουν πολλές φορές την πρόωπη διάγνωση τους εξαιτίας του μικρού μεγέθους, όπως επίσης γιατί οι ενδείξεις του καρκίνου στα πρώτα στάδια είναι ελάχιστα. Παράλληλα, η λανθασμένη διάγνωση μπορεί επίσης να οφείλεται σε ανθρωπίνους παράγοντες, όπως η λανθασμένη

εκτίμηση δομικών χαρακτηριστικών των ιστών , η ανθρώπινη αβλεψία η και ακόμη ανθρώπινη εξάντληση λόγω μεγάλου φόρτου εργασίας.

Η έγκαιρη ανίχνευση μέσω της μαστογραφίας αυξάνει το πλήθος των δυνατών θεραπειών του καρκίνου του μαστού και, όπως είναι λογικό, το ποσοστό επιβίωσης [C. H. Lee, “Screening mammography: proven benefit, continued controversy,” *Radiol. Clinics North Am.* 40, 395–407 (2002).]. Ωστόσο, η τεχνική της μαστογραφίας διαθέτει ψεγάδια. Η ανίχνευση ύποπτων ανωμαλιών είναι μια επαναλαμβανόμενη και κουραστική πνευματική εργασία. Συγκεκριμένα έρευνες έχουν δείξει ότι, οι ακτινολογικές διαγνώσεις, αποτυγχάνουν να εντοπίσουν το 10% έως το 30% των καρκίνων [C. H. Lee, “Screening mammography: proven benefit, continued controversy,” *Radiol. Clinics North Am.* 40, 395–407 (2002).]. Περίπου τα δύο τρίτα αυτών των ψευδώς αρνητικών αποτελεσμάτων είναι λόγω χαμένων βλαβών που είναι εμφανείς αναδρομικά [M. L. Giger, “Computer-aided diagnosis in radiology,” *Acad. Radiol.* 9, 1–3 (2002).]. Ακόμη, ένα υψηλό ποσοστό βιοψιών πραγματοποιούνται σε καλοήθεις αλλοιώσεις. Είναι ευνόητο, ότι είναι αρκετά σημαντικό να προταθούν και να εφαρμοστούν λύσεις με σκοπό την αποφυγή των παραπάνω φαινομένων.

Τα συστήματα ανίχνευσης με τη βοήθεια υπολογιστή (CAD) έχουν αναπτυχθεί για να βοηθήσει τους ακτινολόγους στην ανίχνευση βλαβών στη μαστογραφία που μπορεί να υποδηλώνουν την παρουσία καρκίνου του μαστού. Αυτά τα συστήματα λειτουργούν μόνο ως επιπλέον αναγνώστης-σύμβουλος και η τελική απόφαση καθορίζεται από τον ακτινολόγο. Πρόσφατες μελέτες έχουν επίσης δείξει ότι τα συστήματα ανίχνευσης, όταν χρησιμοποιούνται ως βοήθημα, έχουν βελτιώσει την ακρίβεια διαγνώσεων που παρέχονται από τους ακτινολόγους για την ανίχνευση καρκίνου του μαστού [C. J. Vyborny, M. L. Giger, and R. M. Nishikawa, “Compueraided detection and diagnosis of breast cancer,” *Radiol. Clin. North Am.* 38, 725–740 (2000).]. Συστήματα διάγνωσης με τη βοήθεια υπολογιστή (CDX) για βοήθεια στην απόφαση μεταξύ παρακολούθησης και βιοψίας βρίσκεται ακόμη σε εξέλιξη.

## Οργάνωση και Πρότυπα Σημεία της Διπλωματικής

Η παρούσα διπλωματική εργασία αποτελείται από 7(8) κεφάλαια, τα οποία έχουν την εξής δομή:

Το **Κεφάλαιο 1** αποτελεί μια εισαγωγή στην πάθηση του του καρκίνου του μαστού όπου γίνεται αναφορά στις αίτιες, στην φυσιολογία του μαστού, στα είδη νεοπλασιών καθώς επίσης και στην βιολογική εξέλιξη του. Δίνεται έμφαση στους ενδογενείς και στους εξωγενείς παράγοντες που προκαλούν την εμφάνιση της νόσου. Έπειτα, παρουσιάζονται οι τεχνικές απεικόνισης του καρκίνου του μαστού όπου γίνεται αντιληπτό ότι η μαστογραφία αποτελεί την κυρίαρχη τεχνική για την διάγνωση, αλλά έχει καταστεί σαφές ότι καμία τεχνική δεν υπερτερεί έναντι της άλλης, όμως συμπληρώνει η μια την άλλη . Παράλληλα γίνεται περιγραφή των μεθόδων αντιμετώπισης του καρκίνου του μαστού και παρουσιάζονται τα στατιστικά του Παγκοσμίου Οργανισμού Υγείας για τον αριθμό των γυναικών που εμφανίζουν την νόσο και παθαίνουν(Π.Ο.Υ).

Το **Κεφάλαιο 2** έχει αφιερωθεί εξολοκλήρου στην διαδικασία της μαστογραφίας. Παρουσιάζεται η βασική αρχή της μαστογραφίας καθώς και ο ορισμός της. Επίσης, γίνεται μια εκτενής περιγραφή του εξοπλισμού του αυτογράφου, όπου παρουσιάζονται οι λυχνίες ακτίνων Χ, τα φίλτρα και γίνεται επεξήγηση του αντίστοιχου φάσματος. Επιπρόσθετα πραγματοποιείται αναφορά στους ανιχνευτές αναλογικούς , ψηφιακούς και αναλύονται τα πλεονεκτήματα των ψηφιακών που οδήγησαν στην αντίσταση των αναλογικών. Παράλληλα γίνεται μια σύντομη αναφορά στο αυτόματο σύστημα έκθεσης και στην ψηφιακή τομοσύνθεση, ενώ περιγράφεται η ψηφιακή μαστογραφία αλλά και οι διαφορές που έχει με την αναλογική. Τέλος, καταγράφηκαν οι κίνδυνοι και λανθασμένες αντιλήψεις που συνοδεύουν την μαστογραφία.

Στο **Κεφάλαιο 3** αναλύονται οι μέθοδοι που χρησιμοποιούνται για την ψηφιακή επεξεργασία ενός μαστογραφήματος και οι μέθοδοι που υπάρχουν για την εξαγωγή χαρακτηριστικών της μαστογραφίας από φυσιολογική και μη. Τέλος γίνεται περιγραφή της στατιστικής, δομικής και φασματικής μεθόδου.

Στο **Κεφάλαιο 4** έχει περιγράψει εκτενώς η μέθοδος μηχανικής μάθησης και συγκεκριμένα των νευρωνικών δικτύων. Παρουσιάζεται ο ορισμός ,οι εφαρμογές, τα πλεονεκτήματα των νευρωνικών δικτύων αλλά και η εκπαίδευση με επιτήρηση και χωρίς. Έπειτα γίνεται μια εκτενής περιγραφή των κατηγοριών των νευρωνικών δικτύων που είναι τα συνελκτικά, τα πολυεπίπεδα και τα δίκτυα ακτίνων συναρτήσεων. Για την κάθε κατηγορία αναλύονται τα χαρακτηριστικά τους, οι εφαρμογές τους, τα πλεονεκτήματα, οι αρχιτεκτονικές τους καθώς και το μαθηματικό τους υπόβαθρο.

Το **Κεφάλαιο 5** αποτελεί μια περιγραφή της μεθόδου μηχανικής μάθησης που είναι τα βαθιά νευρωνικά δίκτυα. Αρχικά γίνεται μια πλήρης περιγραφή του τρόπου που προέκυψαν καθώς και της πνοής που έδωσαν στο τομέα της μηχανικής μάθησης , εξαιτίας των ιδιαίτερων χαρακτηριστικών που διαθέτουν. Έτσι , αφού αναλύθηκαν οι θεμελιώδεις έννοιες που το απαρτίζουν καθώς και του τρόπου εκπαίδευσης που απαιτεί συγκεκριμένες προϋποθέσεις τόσο ως προς τα δεδομένα όσο και στις δυνατότητες του συστήματος. Έπειτα ,έγινε μια αναφορά στις δομές που ποικίλουν καθώς και στα

προβλήματα που παρουσιάζουν. Τέλος, παρουσιάστηκαν τα πλεονεκτήματα της βαθιάς μάθησης

Στο **Κεφάλαιο 6** γίνεται μία εισαγωγή τόσο στις αρχιτεκτονικές των νευρωνικών δικτύων που αξιοποιήθηκαν τόσο σε προηγούμενες έρευνες όσο και στην συγκεκριμένη ερευνητική/μεταπτυχιακή εργασία. Τα συγκεκριμένα, αποτελούν τα περισσότερα διάσημα CNNs για την κατηγοριοποίηση dicom αρχείων, είτε σχετικών με μαγνητικές εγκεφάλου ή με διαδικασίες μαστογραφίας, όπως στην συγκεκριμένη περίπτωση. Επιπλέον, παρουσιάζεται μία, συνοπτική, αναφορά σχετικά με τις έννοιες του Training from Scratch και του Transfer Learning καθώς και της βάσης δεδομένων ImageNet που είναι, ιδιαίτερα, διάσημη για την προαναφερόμενη τακτική.

Στο **Κεφάλαιο 7** πραγματοποιείται αναφορά στην βάση δεδομένων που χρησιμοποιήθηκε και αναλύονται όλα τα βήματα που έλαβαν χώρα προκειμένου να κατασκευαστεί ο single-view classifier.

Στο **Κεφάλαιο 8** αναλύονται όλα τα βήματα που οδήγησαν στην κατασκευή του τελικού fusion συστήματος που λαμβάνει υπόψιν όλες τις προβλέψεις, για έναν ασθενή, από τους εκάστοτε single-view κατηγοροποιητές και εξαγεί τα κατάλληλα συμπεράσματα.

Στο **Κεφάλαιο 9** παρουσιάζονται συνοπτικά το σύνολο των πιθανών βελτιώσεων που είναι δυνατόν να λάβουν χώρα με σκοπό, μελλοντικές ερευνητικές ομάδες να παράγουν καλύτερα αποτελέσματα.

Οι κώδικες, στα κεφάλαια 7 και 8, που βρίσκονται στο τέλος, μετά την εκάστοτε ενότητα της βιβλιογραφίας, έχουν τοποθετηθεί με σειρά αναφοράς.

## Πίνακας Εικόνων

Εικόνα 1: Συνολικά ετήσιος αριθμός θανάτων από καρκίνο σε όλα τα φύλα και ηλικίες, καταναμημένα ανά μορφή καρκίνου, για το 2016 [12].	17
Εικόνα 2: Συνολικά ετήσιος αριθμός ατόμων που πάσχουν από καρκίνο όλων των τύπων σε όλα τα φύλα και ηλικίες, για το 2016 [12].	18
Εικόνα 3: Συνολικά ετήσιος αριθμός ατόμων που πάσχουν από καρκίνο όλων των τύπων σε όλα τα φύλα και ηλικίες, για το 2016. [12].	18
Εικόνα 4: Καλοήθης όγκος.	20
Εικόνα 5: Σχηματική αναπαράσταση κακοήθους όγκου με ανεξέλεγκτο πολλαπλασιασμό των κυττάρων	21
Εικόνα 6: Ψηλαφιστή[42]	26
Εικόνα 7: Αξονικός τομογράφος.	28
Εικόνα 8: Εξοπλισμός μαστογραφίας.	39
Εικόνα 9: Παραδείγματα φάσματος μαστογραφίας ακτινών X.	41
Εικόνα 10: Αντιδιαχυτικό διάφραγμα ακτινών X.	42
Εικόνα 11: Σχέση μαστογραφικού φιλμ- γενικού ακτινογραφικού φιλμ.	43
Εικόνα 12: Μεταφορά αντίθεσης σε ψηφιακή μαστογραφία.	43
Εικόνα 13: Σύστημα συμπίεσης του μαστού στη μαστογραφία.	44
Εικόνα 14: Αναλογική και ψηφιακή εικόνα μαστού.	45
Εικόνα 15: Αριστερά: Ψηφιακή απεικόνιση Δεξιά: Απεικόνιση σε film	46
Εικόνα 16: Τυπικές γωνιακές προβολές κατά την ψηφιακή μαστογραφία.	48
Εικόνα 17: Ψηφιακή τομοσύνθεση μίας προβολής και πολλών προβολών.	48
Εικόνα 18: α) Αρχικό ιστόγραμμα, β) Ανάπτυξη ιστογράμματος.	56
Εικόνα 19: α) Αρχικό Ιστόγραμμα, β) Συμπιεσμένο Ιστόγραμμα.	57
Εικόνα 20 : α) αρχικό ιστόγραμμα β) Ολισθημένο ιστόγραμμα[10]	58
Εικόνα 21: α) Αρχικό ιστόγραμμα β) ισοσταθμισμένο ιστόγραμμα[11]	59
Εικόνα 22 : Απλό Νευρωνικό Δίκτυο	68
Εικόνα 23: Νευρωνικό δίκτυο εμπρόσθιας διάδοσης.	72
Εικόνα 24: Περιοχές απόφασης σε πολυεπίπεδο perceptron.	73
Εικόνα 25: Σιγμοειδής συνάρτηση	73
Εικόνα 26: Σιγμοειδής Μονοπολική.	74
Εικόνα 27: Πολυεπίπεδος Perceptron.	74
Εικόνα 28: Παράδειγμα δίκτυο εμπρόσθιας διάδοσης.	74
Εικόνα 29 Συνηθισμένοι τύποι δικτύων: α) επαναλαμβανόμενα δίκτυο; β) ένα στρωματοποιημένο δίκτυο · γ) ένα δίκτυο με συνδέσεις μεταξύ μονάδων εισόδου και εξόδου · δ) μια τροφή δίκτυο προώθησης πλήρως συνδεδεμένο στα δίκτυα προώθησης τροφοδοσίας, οι μονάδες είναι συνήθως διευθετημένες σε επίπεδα (επίπεδα) όπως στο Σχήμα 7.β αλλά και άλλες τοπολογίες μπορεί να χρησιμοποιηθεί. Το σχήμα 7.ε δείχνει έναν τύπο δικτύου που είναι χρήσιμο σε ορισμένες εφαρμογές στις οποίες οι απευθείας συνδέσεις μεταξύ των μονάδων εισόδου και εξόδου χρησιμοποιούνται. Εικόνα 7.δ δείχνει ένα δίκτυο με 3 μονάδες που είναι πλήρως συνδεδεμένο, δηλαδή επιτρέπεται σε όλες τις διασυνδέσεις περιορισμός ροής προς τα εμπρός.	75
Εικόνα 30: Η γενική αρχιτεκτονική ενός RBFN[14].	82
Εικόνα 31: Νευρωνικό δίκτυο ενός επιπέδου.	91
Εικόνα 32: Βαθιά Νευρωνικά δίκτυα	92
Εικόνα 33: Αρχιτεκτονική VGG16.	103

Εικόνα 34:Architecture Map του VGG16(Detailed) .....	104
Εικόνα 35:Residual Blocks .....	105
Εικόνα 36:Identity Mapping .....	106
Εικόνα 37: ResNet50 Architecture .....	107
Εικόνα 38: Inception Module .....	109
Εικόνα 39: Αρχιτεκτονική Googlenet.....	110
Εικόνα 40: AlexNet Architecture .....	111
Εικόνα 41: Δείκτες Μετρήσεων-Είδος Προβλήματος.....	114
Εικόνα 42: TP-TN-FP-FN .....	115
Εικόνα 43: Υπολογισμός ROC-AUC .....	116
Εικόνα 44: Accuracies των διαφορετικών μοντέλων (pretained και from scratch)...	119
Εικόνα 45: Validation Loss σε σύγκριση με τις εποχές.....	120
Εικόνα 46: Fine tuning στην INbreast, pretrained σε Imagenet/CBIS .....	120
Εικόνα 47: From Scratch Training Evaluation .....	121
Εικόνα 48: Pre-trained CNNs .....	122
Εικόνα 49: Pretrained/From Scratch ResNet50 .....	123
Εικόνα 50: Fine Tuning ResNet50.....	123
Εικόνα 51: Mass_case_train_set.....	128
Εικόνα 52: Calc_case_train_set .....	128
Εικόνα 53: Τροποποιημένο σετ δεδομένων mass για test .....	132
Εικόνα 54: Cropped Dicom εικόνα που προβάλλεται στο MATLAB Περιβάλλον..	133
Εικόνα 55: LayersToReplace.m.....	142
Εικόνα 56: Διαχωρισμός normal/abnormal περιπτώσεων .....	149
Εικόνα 57: Mass Detection/Calcification .....	149
Εικόνα 58: Malignant/Benign Detection .....	149
Εικόνα 59: Πίνακες Ασθενών .....	154

Page Intentionally Left Blank

# Κεφάλαιο 1<sup>ο</sup>

## Γενικά για τον Καρκίνο

Ο καρκίνος αποτελεί μια θανατηφόρο ασθένεια η οποία προκαλείται από την ανεξέλεγκτη αύξηση κυττάρων και τον σχηματισμό από περίσσεια αυτών συσσωματωμάτων, δηλαδή όγκων [1].

Το ανθρώπινο σώμα περιέχει εκατομμύρια κύτταρα τα οποία αναπτύσσονται, διαιρούνται-πολλαπλασιάζονται και πεθαίνουν με συμβατικό τρόπο. Σε ορισμένες περιπτώσεις η μη σωστή λειτουργία του ανθρώπινου οργανισμού, οδηγεί σε ανεξέλεγκτο πολλαπλασιασμό των κυττάρων, τα οποία με τη σειρά του συνδυάζονται και οδηγούν σε σχηματισμό μαζικών ιστών (όγκοι) [2]. Η παθογένεια αυτή οδηγεί στην εμφάνιση καρκίνου. Αντίθετα, τα κύτταρα που είναι φυσιολογικά/υγιή ελέγχουν τον τρόπο με τον οποίο αναπτύσσονται και σε περίπτωση που αλλοιωθούν έχουν την δυνατότητα να καταστρέφονται από μόνα τους.

Ο καρκίνος είναι μια κοινή ασθένεια που εξαπλώνεται σε όλο το ανθρώπινο σώμα είτε μέσω του αίματος είτε μέσω των λεμφαγγείων (μετάσταση) [3]. Για παράδειγμα, η λευχαιμία τροποποιεί τα κύτταρα του αίματος και επηρεάζει την ωρίμανση ή όχι αυτών [4]. Ωστόσο, μερικοί από τους όγκους δεν εξαπλώνονται σε ολόκληρο το σώμα, αλλά αναπτύσσονται ανεξέλεγκτα σαν καλοήθεις όγκοι [5].

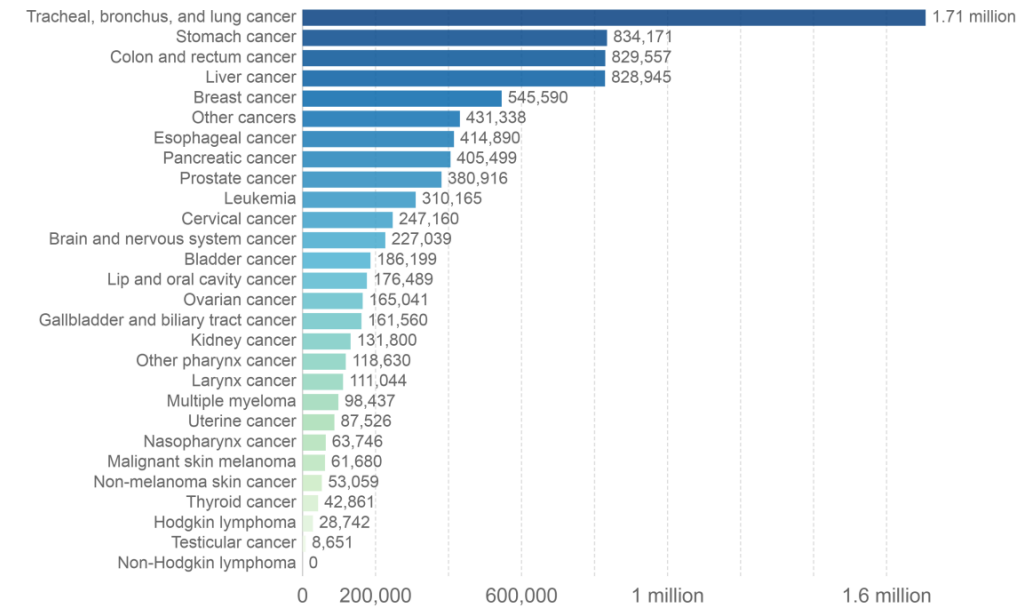
Πολλοί είναι εκείνοι οι λόγοι που μπορεί να οδηγήσουν στην εμφάνιση του καρκίνου. Για παράδειγμα, μεταλλάξεις σε συγκεκριμένα γονίδια καθώς και η διατροφή διαδραματίζουν ζωτικό ρόλο στη θνησιμότητα του καρκίνου [6]-[8]. Η έκθεση σε αλδεΐδες, φορμαλδεΐδη και άλλους οργανικούς διαλύτες σχετίζονται επίσης με τον υψηλό κίνδυνο εμφάνισης καρκίνου (λέμφωμα) [9]. Ωστόσο, δεν είναι δυνατόν να εντοπιστεί η συγκεκριμένη αιτία για την εμφάνιση του καρκίνου. Η ανάπτυξη των καρκινικών κυττάρων ελέγχονται σε μεγάλο βαθμό από το εξωκυττάριο περιβάλλον [10]. Παρόλα αυτά, υπάρχουν και εξωτερικοί παράγοντες που αυξάνουν τις πιθανότητες εμφάνισης καρκίνου, όπως το κάπνισμα, το αλκοόλ, η κακή διατροφή, η παχυσαρκία, η έκθεση στην υπεριώδη ακτινοβολία, η έλλειψη σωματικής δραστηριότητας και σχεδόν το 5% των διαφόρων μορφών καρκίνου προκαλείται λόγω κληρονομικών παραγόντων [1].

Καθώς ο καρκίνος κατατάσσεται στην δεύτερη θέση ως αίτιο θνησιμότητας παγκοσμίως μετά τις καρδιαγγειακές παθήσεις [11], [2], αποτελεί μια από τις πιο ενδιαφέρουσες προκλήσεις της μοντέρνας ιατρικής. Συγκεκριμένα κάθε 6 θάνατοι στον κόσμο οφείλονται στον καρκίνο. Το 2016, υπολογίστηκε ότι 8.9 εκατομμύρια άτομα πέθαναν από διάφορες μορφές καρκίνου. Το Ινστιτούτο Μετρήσεων και Αξιολόγησης της Υγείας (IMAY) έθεσε σχετικά μικρά περιθώρια σφάλματος γύρω από αυτόν τον συνολικό αριθμό. Στη παρακάτω εικόνα, παρουσιάζεται ένα διάγραμμα εκτίμησης ποσοστού θανάτων παγκοσμίως, για το 2016 ανάλογα με τους διάφορους τύπου καρκίνου.



## Cancer deaths by type, World, 2016

Total annual number of deaths from cancers across all ages and both sexes, broken down by cancer type.



Source: IHME, Global Burden of Disease (GBD)

CC BY-SA

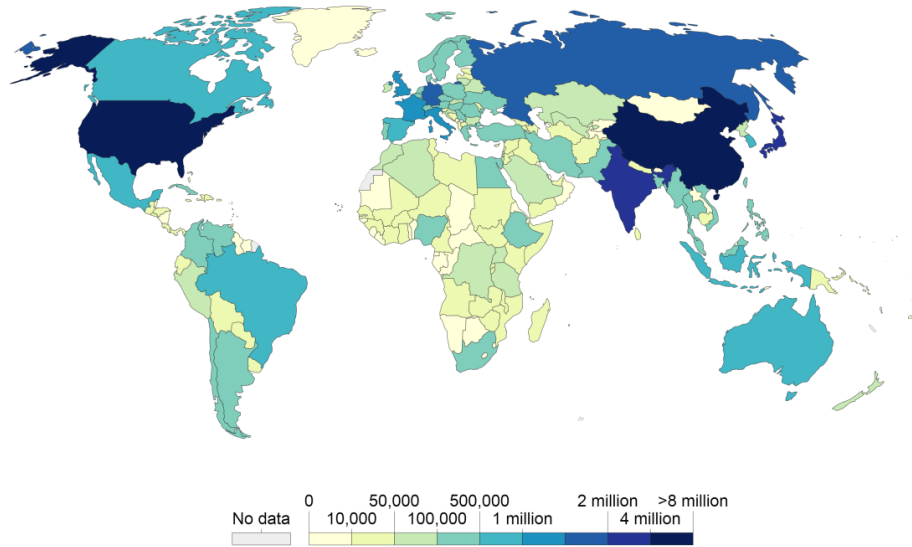
Εικόνα 1: Συνολικά ετήσιος αριθμός θανάτων από καρκίνο σε όλα τα φύλα και ηλικίες, κατανεμημένα ανά μορφή καρκίνου, για το 2016 [12].

Κάθε χρόνο διαγιγνώσκεται η ύπαρξη καρκίνου σε έναν στους διακόσιους πενήντα άνδρες και σε μία στις τριακόσιες γυναίκες [11],[2]. Η επίπτωση του καρκίνου παρουσιάζει αύξηση στην Ευρώπη. Σήμερα ο καρκίνος διαγιγνώσκεται σε περισσότερους από 1.2 εκατομμύρια ανθρώπους το χρόνο εντός των χωρών της Ευρωπαϊκής Ένωσης [13], [3].

Στις ανεπτυγμένες χώρες, οι πιο συχνά εμφανιζόμενες μορφές καρκίνου είναι αυτές του πνεύμονα, μαστού, του δέρματος, του εντέρου και τέλος του προστάτη. Στην Εικόνα 2, παρουσιάζεται ένα διάγραμμα συνολικού αριθμού ατόμων που έπασχαν από οποιοδήποτε τύπο καρκίνου παγκοσμίως για το 2016.

## Number of people with cancer, 2016

Total number of people suffering from any type of cancer at a given time. This is measured across both sexes, and all ages.

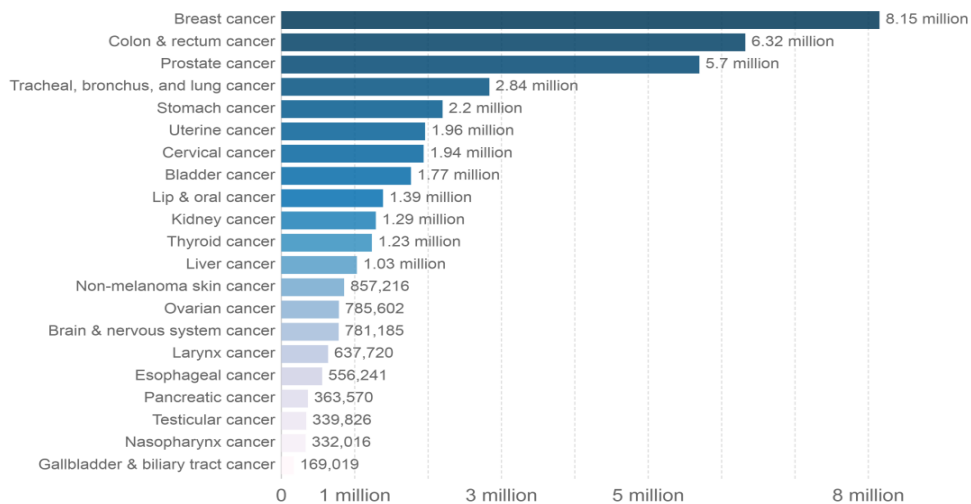


Εικόνα 2: Συνολικά ετήσιος αριθμός ατόμων που πάσχουν από καρκίνο όλων των τύπων σε όλα τα φύλα και ηλικίες, για το 2016 [12].

Πιο συγκεκριμένα στην Εικόνα 3, παρουσιάζονται περιστατικά ατόμων που διαγνώστηκαν και πάσχουν από καρκίνο, διαφόρων μορφών, όπως προέκυψε από στατιστικές μελέτες σε όλο τον κόσμο για όλα τα φύλα και ηλικίες, για το 2016. Όπως φαίνεται, ο καρκίνος του μαστού είναι και αυτός που με μόλις 8.5 εκατομμύρια ατόμων.

## Number of people with cancer by type, World, 2016

Total number of people suffering from cancer at any given time, differentiated by cancer type. This is measured across both sexes and all ages.



Εικόνα 3: Συνολικά ετήσιος αριθμός ατόμων που πάσχουν από καρκίνο όλων των τύπων σε όλα τα φύλα και ηλικίες, για το 2016. [12]

### 1.1 Αίτιες καρκίνου

Στη διερεύνηση των αιτιών του καρκίνου, δύο κύριες οδοί μπορούν να ακολουθηθούν: Η πρώτη αφορά στο γενετικό επίπεδο και η άλλη στις αιτιολογικές συσχετίσεις. Οι αιτίες καρκίνου που περιγράφονται μέχρι σήμερα είναι αρκετές, και οι μηχανισμοί της δημιουργίας του είναι αρκετά πολύπλοκοι [14]. Οι αιτίες αυτές μπορεί να είναι χημικές (χημικά καρκινογόνες ουσίες, ιοντίζουσες και υπεριώδεις ακτινοβολίες), ιογενείς (ιοί και ογκογονίδια), και άλλες, όπως η κληρονομικότητα, τα τραύματα, και οι προκαρκινικές κακώσεις.

Η βιολογική μονάδα του καρκίνου είναι το καρκινικό κύτταρο, το οποίο είναι παρόμοιο με ένα φυσιολογικό κύτταρο από πολλές απόψεις. Οι θεμελιώδεις ιδιότητες όμως του καρκινικού κυττάρου, είναι αυτές που το διακρίνουν από τα φυσιολογικά αντίστοιχά του. Δύο βασικά χαρακτηριστικά του είναι ο μη ελεγχόμενος πολλαπλασιασμός και η ικανότητα να κάνει μεταστάσεις [15]. Χαρακτηριστικό είναι το γεγονός ότι μεταξύ των καρκινικών κυττάρων παρατηρούνται διαφορές στη λειτουργία ή/και τη δομή, όταν προέρχονται από καρκίνους διαφορετικής προέλευσης, πράγμα που δεν παρατηρείται ανάμεσα στα φυσιολογικά κύτταρα [16]. Παρόλο, όμως, που ο καρκίνος είναι μια διαταραχή των κυττάρων, συνήθως κατά τη διάγνωση, εμφανίζεται σαν ένας ορατός όγκος που είναι το τελικό αποτέλεσμα μιας ολόκληρης σειράς γενετικών αλλαγών στα κύτταρα που προάγουν την εμφάνιση νέων χαρακτηριστικών στο ανθρώπινο σώμα, που μπορεί να έχει πάρει χρόνια για να αναπτυχθεί [15].

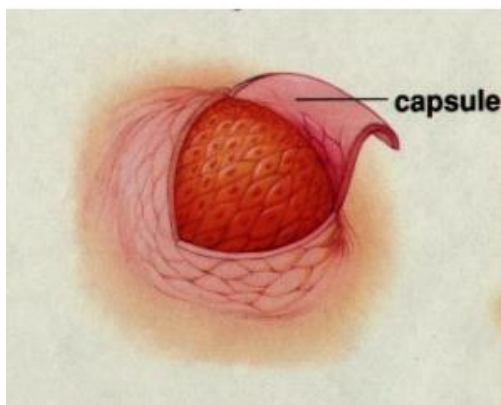
Ο καρκίνος, λοιπόν, είναι μια νόσος που προκαλείται όταν τα φυσιολογικά κύτταρα διαταραχθούν από κάποιον αιτιολογικό παράγοντα και τότε οδηγούνται σε μη ελεγχόμενη λειτουργία και ανάπτυξη. Αυτή η μη ελεγχόμενη ανάπτυξη προκαλεί το σχηματισμό του όγκου. Αν δεν αντιμετωπισθεί, ο όγκος μπορεί να προκαλέσει προβλήματα με την εισβολή του σε φυσιολογικούς γειτονικούς ιστούς ή ασκώντας πίεση σε άλλες γειτονικές δομές του σώματος, όπως επίσης και να αποβεί θανατηφόρος για τη ζωή του ασθενούς.

### 1.2 Τύποι καρκίνου

Υπάρχουν πάνω από διακόσιοι διαφορετικοί τύποι καρκίνου, καθώς μπορεί να υπάρξουν στο σώμα πάνω από διακόσιοι διαφοροποιημένοι τύποι κυττάρων [17],[18]. Οι διαφορετικοί ιστοί του σώματος αποτελούνται από διαφορετικούς τύπους κυττάρων τα οποία εκτελούν διαφορετικές λειτουργίες όμως όλα είναι παρόμοια όσο αφορά τη δομή τους. Φέρουν πυρήνα, μέσα στον οποίο υπάρχουν τα γονίδια, τα οποία ελέγχουν τις λειτουργίες του κυττάρου, και δίνουν οδηγίες για την αναπαραγωγή αλλά και τον θάνατο τους. Όταν διαταραχθεί αυτή η φυσιολογική διαδικασία τότε το κύτταρο αρχίζει να διαιρείται χωρίς κανένα έλεγχο, έτσι ώστε τα κύτταρα αυξάνονται σε αριθμό. Η συσσώρευση τους στην περιοχή οδηγεί στον σχηματισμό όγκου [13],[15],[18]. Οι όγκοι μπορεί να είναι καλοήθεις ή κακοήθεις.

Οι **καλοήθεις όγκοι** σχηματίζονται από πολλαπλασιαζόμενα κύτταρα εξαιτίας κάποιου ερεθισμού, φαινόμενο που ονομάζεται καλοήθης νεοπλασία. Ωστόσο, όταν ο ερεθισμός αυτός πάψει να υφίσταται, η νεοπλασία διακόπτεται, ο όγκος δεν

αναπτύσσεται περαιτέρω, δεν καταστρέφει περιβάλλοντες υγιείς ιστούς (μολονότι μπορεί να ασκήσει βλαπτική επίδραση πάνω τους) και δεν προκαλεί το σχηματισμό νέων όγκων σε άλλα σημεία του σώματος. Παράλληλα, όταν ένας καλοήθης όγκος αφαιρείται από την περιοχή του οργανισμού στην οποία αναπτύχθηκε, δεν παρατηρείται εκ νέου ανάπτυξή του. Οι καλοήθεις όγκοι σπάνια προκαλούν σοβαρά προβλήματα σε έναν οργανισμό, και συνήθως αντιμετωπίζονται εύκολα. Προβλήματα μπορούν να προκαλέσουν μόνο όταν γίνουν αρκετά μεγάλοι, όταν πιέζουν άλλα όργανα, όταν αλλοιωθεί η μορφή τους, και αν εκκρίνουν ορμόνες που επηρεάζουν τις λειτουργίες του σώματος.



Εικόνα 4: Καλοήθης όγκος.

Οι **κακοήθεις όγκοι** είναι γνωστοί και ως νεοπλάσματα. Σχηματίζονται με τρόπο παρόμοιο με τους καλοήθεις όγκους, αλλά λειτουργούν εντελώς διαφορετικά. Τα πολλαπλασιαζόμενα κύτταρα φέρουν μια παθογένεια που προκαλεί τη συνεχή και άναρχη αύξησή τους. Η κακοήθης νεοπλασία - η διαδικασία του άναρχου πολλαπλασιασμού των κυττάρων, που καταλήγει σε σχηματισμό νέων ιστών - εξακολουθεί να δρα, ακόμη και όταν ο αρχικός όγκος αφαιρεθεί από το σώμα, σχηματίζοντας στη θέση του έναν καινούργιο. Ταυτόχρονα, επιδρά καταστροφικά στους γεινιάζοντες ιστούς, στους οποίους εισβάλλει (διαδικασία που ονομάζεται διήθηση) και έχει τη δυνατότητα σχηματισμού νέων όγκων σε άλλα σημεία του σώματος, όταν καρκινικά κύτταρα, αφού αποσπαστούν από τον πρωτογενή όγκο (το σημείο του σώματος όπου ξεκινάει η ανάπτυξη ενός καρκίνου), μετακινούνται μέσω του αίματος ή της λέμφου στο υπόλοιπο σώμα. Η διαδικασία αυτή, όπως προαναφέρθηκε, ονομάζεται μετάσταση (μετεγκατάσταση, δηλαδή, των καρκινικών κυττάρων). Οι κακοήθεις όγκοι, οι οποίοι χαρακτηρίζονται, ανάλογα με το είδος τους, ως καρκινώματα ή σαρκώματα, είναι εν δυνάμει θανατηφόροι, ειδικά αν δε διαγνωστούν έγκαιρα.

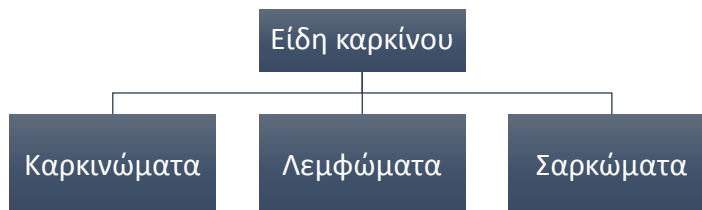


Εικόνα 5: Σχηματική αναπαράσταση κακοήθους όγκου με ανεξέλεγκτο πολλαπλασιασμό των κυττάρων

### Βασικοί τύποι καρκίνου

Ανάλογα με την δράση των διαφοροποιημένων κύτταρων που προσβάλλουν τους ιστούς, αναπτύσσονται διάφοροι τύποι καρκίνου. Κάθε μορφή καρκίνου γι' αυτόν το λόγο παρουσιάζει και διαφορετικά συμπτώματα σε κάθε ασθενή, καθώς μπορεί να εμφανίζεται και σε διαφορετικό σημείο του σώματος κάθε ασθενούς. Επίσης η ταχύτητα ανάπτυξης των κυττάρων παίζει καθοριστικό ρόλο στην επίδραση που θα έχει στο σώμα του ασθενούς, αφού διαφορετικές θα είναι και οι ουσίες που θα εκκρίνονται στο αίμα. Λόγω λοιπόν της διαφορετικής τάσης ανάπτυξης των καρκινικών κυττάρων, η ανταπόκριση του κάθε ασθενούς στις μεθόδους αντιμετώπισης του θα είναι εντελώς διαφορετική [15],[18].

Καθώς λοιπόν οι λειτουργία και ανάπτυξη των καρκινικών κυττάρων μπορεί να διαφέρει, είναι δύσκολη οποιαδήποτε κατηγοριοποίηση. Ωστόσο, μια βασική διάκριση αυτών είναι η ακόλουθη:

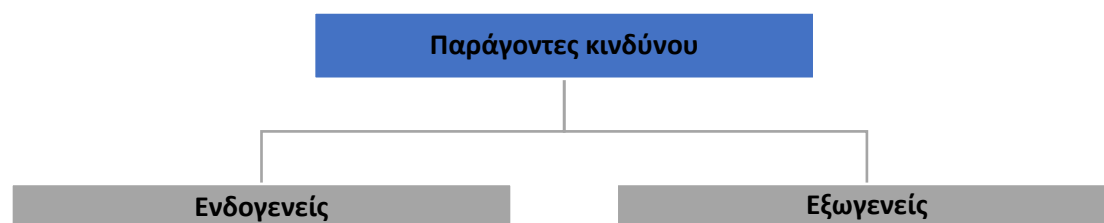


Στην συνέχεια παρουσιάζονται συνοπτικά οι τύποι καρκίνου και ορισμένα παραδείγματα αυτών καθώς και οι αιτίες και τα συμπτώματά τους.

### 1.3 Καρκίνος του μαστού: Αιτιολογία και παράγοντες κινδύνου

Όπως αναφέρθηκε και προηγουμένως ο καρκίνος του μαστού αποτελεί τον πιο κοινό τύπο καρκίνου για τις γυναίκες και τη δεύτερη αιτία θανάτου. Συγκεκριμένα μια στις 8 γυναίκες πάσχουν από καρκίνο του μαστού όπου απαιτείται για την αντιμετώπιση του πλήρης αφαίρεση, χημειοθεραπεία, ακτινοθεραπεία καθώς και ορμονική θεραπεία τις περισσότερες φορές [14].

Τα αίτια που προκαλούν την εμφάνιση καρκίνου του μαστού δεν έχουν διευκρινιστεί πλήρως. Ωστόσο μια βασική διάκριση παραγόντων που ευθύνονται για την εμφάνιση καρκίνου του μαστού, είναι σε ενδογενείς και εξωγενείς παράγοντες κινδύνου.



### 1.3.1 Ενδογενείς παράγοντες

Αν και ο καρκίνος του μαστού αυξάνεται με την ηλικία, η ο ρυθμός αύξησης του είναι βραδύτερος μετά την εμμηνόπαυση. Αυτό αντανακλά το γεγονός ότι ο καρκίνος του μαστού εξαρτάται από τις ορμόνες της νόσου. Τρεις χρονικές περίοδοι στη ζωή μιας γυναίκας: ηλικία στην εμμηνόρροια, ηλικία στην πρώτη πλήρη κύηση, και ηλικία κατά την εμμηνόπαυση. Γυναίκες που βιώνουν εμμηνόρροια στην ηλικία των 16 ετών έχουν μόνο 50 - 60% του κινδύνου για εμφάνιση καρκίνου του μαστού. Ομοίως, εμμηνόπαυση που συμβαίνουν 10 χρόνια πριν τη μέση ηλικία της εμμηνόπαυσης (52 ετών), φυσικό ή χειρουργικά προκαλείται, μειώνει τον κίνδυνο διάρκειας ζωής κατά περίπου 35%. Οι γυναίκες που έχουν πρώτη πλήρη εγκυμοσύνη κατά ηλικία των 18 ετών έχουν 30-40% χαμηλότερο κίνδυνο εμφάνισης καρκίνου του μαστού. Έτσι, η διάρκεια της εμμήνου ρύσεως ιδιαίτερα το διάστημα που εμφανίζεται πριν από την πρώτη πλήρους εγκυμοσύνη αποτελεί ουσιαστική συνιστώσα για τον συνολικό κίνδυνο του καρκίνου του μαστού. Αυτοί οι τρεις παράγοντες μπορούν να ληφθούν υπόψη για το 70-80% της διακύμανσης της συχνότητας του καρκίνου του μαστού σε διάφορες χώρες.

Επίσης η εθνικότητα παίζει επίσης καθοριστικό ρόλο. Ο πληθυσμός λευκών γυναικών έχει υψηλότερη συχνότητα εμφάνισης καρκίνου του μαστού, ενώ ακολουθούν οι γυναίκες αφρικανικής καταγωγής, οι Αμερικανίδες και οι γυναίκες ισπανικής καταγωγής [14].

### 1.3.2 Εξωγενείς παράγοντες

Κατανόηση του δυνητικού ρόλου των εξωγενών ορμονών στον καρκίνο του μαστού είναι εξαιρετικής σημασίας γιατί τα εκατομμύρια των Αμερικάνων γυναικών τακτικά χρήση από του στόματος αντισυλληπτικών και μετα-εμμηνοπαυσιακής ορμόνης θεραπεία (HRT). Η πιο αξιόπιστη μετά από αναλύσεις της από στόματος χρήση αντισυλληπτικών προτείνουν ότι οι παράγοντες αυτοί προκαλούν μόνο ένα μικρό αυξημένο κίνδυνο του καρκίνου του μαστού. Η HRT έχει μεγάλο αντίκτυπο στον κίνδυνο του καρκίνου του μαστού. Μελέτες έδειξαν ότι η HRT σε συνδυασμό με οιστρογόνα συν προγεστίνη για 6 έως 7 χρόνια σχεδόν, διπλασίασε τον κίνδυνο εμφάνισης του.

Επίσης καθοριστικό ρόλο παίζει και ο δείκτης μάζας σώματος (ΔΜΣ). Ο υψηλός ΔΜΣ έχει προστατευτικό αποτέλεσμα σε προ εμμηνοπαυσιακές περιόδους. Υπάρχει

ελάχιστος ή σχεδόν καθόλου συσχετισμός μεταξύ υψηλού ΔΜΣ και του καρκίνου μετά την εμμηνόπαυση. Ο υψηλότερος ΔΜΣ έχει συσχετιστεί με θνησιμότητα, πιθανώς λόγω της δυσκολίας στη διάγνωση των μικρών όγκων σε μεγαλύτερα στήθη.

Τέλος, υπάρχει συσχέτιση μεταξύ της κατανάλωσης αλκοόλ και του κινδύνου: για κάθε αύξηση των 10 g / ημέρα υπάρχει 9% αύξηση του κινδύνου εμφάνισης καρκίνου του μαστού [14].

Γενικά Παράγοντες που ευνοούν την ανάπτυξη του είναι [19]:

- **Το φύλο.** Η αναλογία του καρκίνου του μαστού στη γυναίκα σε σχέση με τον άνδρα είναι 100:1
- **Η κληρονομικότητα.** Διπλάσια ή τριπλάσια πιθανότητα σε σχέση με το γενικό πληθυσμό έχουν οι γυναίκες που η μητέρα τους ή η αδερφή τους εμφάνισαν καρκίνο του μαστού. Ο κίνδυνος επίσης μεγαλώνει εάν οι συγγενείς του πρώτου βαθμού (μητέρα- αδερφή) έχουν προσβληθεί σε μικρή ηλικία ή αν ο καρκίνος ήταν και στους δύο μαστούς (αμφοτερόπλευρα). Έτσι, ο κίνδυνος γίνεται 5 φορές μεγαλύτερος στην αμφοτερόπλευρη εντόπιση και εννεαπλάσιος στην περίπτωση που η νόσος έχει εμφανιστεί αμφοτερόπλευρα πριν από την εμμηνόπαυση. Εκτιμάται ότι το 5% των γυναικών με καρκίνο μαστού έχουν ταυτοποιημένη μετάλλαξη πρωτογενούς αλληλουχίας στο γονίδιο BRCA 1, εντοπισμένη στο χρωμόσωμα 17q21. Κατά συνέπεια, οι συγγενείς που φέρουν μεταλλάξεις του BRCA 1 έχουν υψηλό κίνδυνο (μέχρι και 85%) για καρκίνο μαστού στη διάρκεια της ζωής τους, με 50% πιθανότητα να εμφανιστεί πριν από την ηλικία των 50 ετών.
- **Η τεκνοποίηση.** Γυναίκες που ποτέ δεν γέννησαν ή που τεκνοποίησαν σε μεγάλη ηλικία (π.χ. μετά τα 30 έτη) έχουν τετραπλάσια πιθανότητα να αναπτύξουν καρκίνο του μαστού. Αντίθετα, όσο μεγαλύτερος είναι ο αριθμός των κυήσεων μίας γυναίκας τόσο μειώνεται ο κίνδυνος. Επιπλέον, η πρώτη κύηση σε μικρή ηλικία (γύρω στα 20 έτη) θεωρείται προστατευτικός παράγων.
- **Πρώιμη εμμηναρχή/ όψιμη εμμηνόπαυση.** Έμμηνος ρύση που αρχίζει σε μικρή ηλικία (πριν από τα 11 έτη) και εμμηνόπαυση σε μεγάλη ηλικία (μετά τα 50 έτη) αυξάνουν τον κίνδυνο για εμφάνιση καρκίνου του μαστού. Αντίθετα η πρώιμη ή η τεχνητή εμμηνόπαυση αποτελεί προστατευτικό παράγοντα.
- **Το ατομικό ιστορικό.** Ο κίνδυνος να παρουσιαστεί καρκίνος στον άλλο μαστό μετά από κάποιο χρονικό διάστημα από την εμφάνιση του πρώτου (δευτερο-πρωτοπαθής όγκος του μαστού) είναι 5 φορές μεγαλύτερος από τον κίνδυνο που διατρέχει ο γενικός πληθυσμός.
- **Η ακτινοβολία.** Η έκθεση στην ιοντίζουσα ακτινοβολία αυξάνει τον κίνδυνο για καρκίνο του μαστού. Εδώ επισημαίνεται ότι η ακτινοβολία που δέχεται η γυναίκα από την μαστογραφία είναι άνευ σημασίας λόγω της πολύ χαμηλής δόσης.
- **Οι διατροφικοί παράγοντες.** Η δίαιτα υψηλής περιεκτικότητας σε ζωικές πρωτεΐνες ή ζωικά λίπη, η υπερβολική κατανάλωση οινοπνεύματος και η παχυσαρκία αποτελούν παράγοντες αυξημένου κινδύνου για την εκδήλωση της



νόσου. Ως προστασία συνιστάται τροφή πλούσια σε φυτικές ίνες, φρούτα και λαχανικά.

- **Τα οιστρογόνα και τα αντισυλληπτικά.** Η μακροχρόνια λήψη οιστρογόνων θεωρείται παράγοντας αυξημένου κινδύνου για καρκίνου του μαστού. Όσον αφορά στα αντισυλληπτικά δισκία, δεν έχει αποδειχτεί σημαντική σχέση μεταξύ λήψης αυτών και του καρκίνου του μαστού. Εκείνο που πρέπει να τονιστεί είναι ότι οι μετεμμηνοπαυσιακές γυναίκες που χρησιμοποιούν οιστρογόνα πρέπει να υπόκεινται σε τακτική παρακολούθηση των μαστών.
- **Η κυστική μαστοπάθεια.** Θεωρείται από τους πλέον ισχυρούς προδιαθεσικούς παράγοντες ανάπτυξης καρκίνου του μαστού. Η "επιθηλίωση", δηλαδή η αύξηση των επιθηλιακών στοιχείων χωρίς υπερπλασία, αυξάνει τον κίνδυνο 1,9 φορές, ενώ η παρουσία υπερπλασίας αυξάνει τον κίνδυνο εμφάνισης της νόσου κατά 5 φορές περίπου. Η παρουσία ατυπίας συνοδευόμενη από θετικό οικογενειακό ιστορικό αυξάνει τον κίνδυνο κατά 11 φορές. Τέλος, η παρουσία και ασβεστώσεων μαζί με την επιθηλίωση κάνει τον κίνδυνο ακόμα μεγαλύτερο για την δημιουργία καρκίνου του μαστού.
- **Το κάπνισμα.** Αν και το κάπνισμα δεν έχει αποδειχτεί ότι προκαλεί καρκίνο στο μαστό, υπάρχουν μερικές μελέτες που συσχετίζουν το κάπνισμα με αύξηση της συχνότητας του καρκίνου του μαστού.

#### 1.4 Τύποι καρκίνου του μαστού

Υπάρχουν πολλά είδη καρκίνων του μαστού. Εξαρτάται πώς τους κατατάσσουμε.

Μια διάκριση με βάση την καλοήθεια ή κακοήθεια των μορφών καρκίνου, όπως περιεγράφηκε πιο πάνω, και συγκεκριμένα του καρκίνου του μαστού, παρουσιάζεται ακολούθως.

<i><b>Καλοήθεις παθήσεις</b></i>	<i><b>Κακοήθεις παθήσεις</b></i>
• Ινοαδένωμα	• Επί τόπου Πορογενές Καρκίνωμα (Ductal carcinoma in situ [DCIS])
• Κύστη	• Πορογενές διηθητικό καρκίνωμα (Invasive Ductal Carcinoma [IDC])
• Λίπωμα	• Σωληνώδες Καρκίνωμα του μαστού [Tubular carcinoma]
• Φυλλοειδής όγκος	• Μυελώδες Καρκίνωμα στο στήθος
• Θήλωμα	• Βλεννοπαράγωγο Μαστικό Καρκίνωμα
• Φλεγμονή-Απόστημα	• Θηλώδες (Θηλοειδές) Καρκίνωμα στο μαστό (Papillary Carcinoma of the Breast)
• Ινοκυστική μαστοπάθεια	• Διάτρητο Καρκίνωμα του Μαστού [Cribriform Carcinoma of the Breast]



Ανάλογα με το πόσο βαθιά έχει προχωρήσει ένας καρκίνος μέσα στο μαστό, χωρίζεται σε **in situ** και σε **διηθητικό**. Ανάλογα με τα κύτταρα από τα οποία προέρχεται, χωρίζεται σε **πυρογενή** και σε **λοβιακό**. Ανάλογα με το πόσο κακοήγη είναι τα κύτταρά του, χωρίζεται σε **βαθμό κακοήθειας 1** (χαμηλό), **2** (μέτριο) ή **3** (υψηλό). Ανάλογα με το πόσο διαφοροποιημένα είναι τα κύτταρά του χωρίζεται σε **χαμηλής, μέσης και υψηλής διαφοροποίησης**. Επίσης χωρίζονται σε δύο ομάδες ανάλογα με το αν τα κύτταρα έχουν **θετικούς ορμονικούς υποδοχείς** οιστρογόνων και προγεστερόνης, ή αν έχουν **αρνητικούς ορμονικούς υποδοχείς**. Χωρίζονται επίσης σε δύο ομάδες ανάλογα με το αν έχουν υποδοχείς **HER2** ή όχι δηλαδή **HER-2 θετικοί** και **HER-2 αρνητικοί** [20]

### 1.5 Διαγνωστικές τεχνικές για τον καρκίνο του μαστού

Οι ακτινογραφίες παράγονται με ακτινοβολία με ακτίνες X ενός μέρους του ανθρώπινου σώματος, του μαστού στην περίπτωση μας, και την ακτινογραφία της ακτινοβολίας που δεν απορροφήθηκε από αυτό. Η ακτινογραφία είναι η πλέον διαδεδομένη μέθοδος ιατρικής απεικόνισης.

Μια μαστογραφία είναι μια ακτινογραφία του μαστού. Σε παρακάτω κεφάλαιο θα γίνει εκτενής περιγραφή της διαγνωστικής τεχνικής. Ενώ οι διαγνωστικές μαστογραφίες χορηγούνται συνήθως για την ανίχνευση του καρκίνου του μαστού σε γυναίκες που δεν έχουν εμφανή συμπτώματα. Χρησιμοποιούνται διαγνωστικά μαστογραφία δείγματα μετά από ύποπτα αποτελέσματα σε μαστογραφία προ συμπτωματικού ελέγχου ή μετά από μερικά σημάδια καρκίνου του μαστού προειδοποιούν τον γιατρό να ελέγξει τον ιστό.

Τέτοια σημεία μπορεί να περιλαμβάνουν:

- Πόνος στο στήθος
- Απορρόφηση θηλών
- Σκλήρυνση του δέρματος στο στήθος
- Αλλαγές στο μέγεθος ή το σχήμα του μαστού

Μια διαγνωστική μαστογραφία μπορεί να βοηθήσει να καθοριστεί εάν αυτά τα συμπτώματα είναι ενδεικτικά της παρουσίας του καρκίνου. Σε σύγκριση με τα μαστογραφία διαγνωστικά, οι διαγνωστικές μαστογραφίες παρέχουν μια πιο λεπτομερή ακτινογραφία του μαστού χρησιμοποιώντας εξειδικευμένες τεχνικές. Επίσης χρησιμοποιούνται σε ειδικές περιστάσεις, όπως για ασθενείς με εμφυτεύματα στήθους.

Ακόμη, σε γυναίκες νεαρής ηλικίας δεν συσταίνεται η εξέταση της μαστογραφίας για τις παρακάτω αιτίες:

- Σε νεαρή ηλικία (20-30) ο μαστός είναι πυκνός και επομένως η εξέταση έχει χαμηλή διαγνωστική αξία.

- Όταν είναι εφικτό να αποφευχθεί η ακτινοβολία πρέπει να αποφεύγεται λόγω της ακτινοβολίας.
- Στην νεαρή ηλικία γίνεται συνήθως μόνο αν υπάρχει υποψία για κακοήθεια.

Ο γιατρός αν συνταγογραφήσει μια διαγνωστική μαστογραφία, αυτό σημαίνει ότι θα διαρκέσει περισσότερο από μια κανονική μαστογραφία διαλογής, επειδή λαμβάνονται περισσότερες ακτινογραφίες, παρέχοντας όψεις του μαστού από πολλαπλά πλεονεκτικά σημεία. Ο ακτινολόγος που διενεργεί την εξέταση μπορεί επίσης να κάνει μεγέθυνση σε μια συγκεκριμένη περιοχή του μαστού όπου υπάρχει υποψία ανωμαλίας. Αυτό θα δώσει στον γιατρό μια καλύτερη εικόνα του ιστού για να επιτύχει μια ακριβή διάγνωση.

Η ικανότητα μιας μαστογραφίας να ανιχνεύει καρκίνο του μαστού μπορεί να εξαρτάται από το μέγεθος του όγκου, την πυκνότητα του ιστού του μαστού και την ικανότητα του ακτινολόγου να χορηγεί και να διαβάζει τη μαστογραφία. Η μαστογραφία είναι λιγότερο πιθανό να αποκαλύψει όγκους του μαστού σε γυναίκες ηλικίας κάτω των 50 ετών από ό,τι σε γυναίκες μεγαλύτερης ηλικίας. Αυτό μπορεί να οφείλεται στο γεγονός ότι οι νεότερες γυναίκες έχουν πυκνότερο ιστό μαστού που εμφανίζεται λευκό σε μαστογραφία, έτσι ένας όγκος εμφανίζεται λευκός σε μια μαστογραφία με αποτέλεσμα να είναι δύσκολο να εντοπιστεί.

#### 1.5.1 Ψηλάφηση

Γενικά οι γυναίκες ανεξάρτητου ηλικίας πρέπει μια φορά τον μήνα, και συγκεκριμένα κατά το τέλος της εμμήνου ρήσεως, όπου συνήθως οι μαστοί είναι μαλακοί και λιγότερο ευαίσθητοι, να ψηλαφούν τους μαστούς τους. Επιπλέον, μετά την εμμηνοπάυση ή σε γυναίκες που έχουν κάνει αφαίρεση μήτρας, η εξέταση πρέπει να υλοποιείται την ίδια μέρα κάθε μήνα. Παράλληλα μια φορά το εξάμηνο ή τον χρόνο η ψηλάφηση πρέπει να γίνεται από ειδικό γιατρό.



Εικόνα 6: Ψηλαφιστή[42]

#### 1.5.2 Υπερηχογράφημα

Οι μηχανικές δονήσεις διαδίδονται μέσα από στερεά, υγρά και αέρια καθώς η δόνηση μεταφέρεται από το ένα σωματίδιο της ύλης στο διπλανό του. Τα σωματίδια της ύλης ταλαντώνονται και η κίνησή τους είναι κανονική τότε μετρίεται σε κύκλους ανά δευτερόλεπτο ( $\text{Hz} = \text{sec}^{-1}$ ). Όταν, η συχνότητα της δόνησης είναι μεταξύ 10 και 20000 Hz, ο ήχος είναι αντληπτός από το ανθρώπινο αυτί. Οι δονήσεις με συχνότητα πάνω από 20000 Hz οι «ήχοι» δεν ακούγονται και ονομάζονται υπέρηχοι (ultrasound ή

ultrasonic).Ακόμη , οι υπόηχοι αφορούν χαμηλές συχνότητες κάτω από 16.000 κύκλους ανά δευτερόλεπτο και είναι αρκετά επιβλαβείς.[37]

Η παραγωγή και η ανίχνευση των υπερήχων βασίζεται στο πιεζοηλεκτρικό φαινόμενο, κατά το οποίο μετά από συμπίεση ή εφελκυσμό ενός υλικού, παρατηρούνται ετερόνυμα ηλεκτρικά φορτία στα άκρα του υλικού και αντιστρόφως, η εφαρμογή μιας ηλεκτρικής τάσης (διαφορά δυναμικού) θα έχει ως αποτέλεσμα τη συμπίεση ή τον εφελκυσμό του υλικού. Τα υλικά με αυτές τις ιδιότητες ονομάζονται πιεζοηλεκτρικά (π.χ. οι κρύσταλλοι χαλαζία, διάφορα κεραμικά υλικά κτλ.). Η παραγωγή των υπερήχων γίνεται με την εφαρμογή του κατάλληλου εναλλασσόμενου ηλεκτρικού δυναμικού στα άκρα του υλικού, που προκαλεί περιοδική μετακίνηση της επιφάνειας του υλικού ταλάντωση των εφαιπτόμενων μορίων αέρα και την παραγωγή υπερήχων. Για την ανίχνευση των υπερήχων χρησιμοποιείται η ίδια πηγή, όπου το κύμα των υπερήχων προσπίπτει επάνω στην πηγή και προκαλεί παλμική κίνηση στην επιφάνειά του, δημιουργώντας διαφορά δυναμικού στα άκρα του, η οποία και καταγράφεται, ή σκεδάζονται όταν συναντά εμπόδια ή ασυνέχειες, ανάλογα με τη συχνότητά τους. Αυτή η ιδιότητα, χρησιμοποιείται για να εντοπιστούμε τυχόν ατέλειες στο εσωτερικό ενός σώματος χωρίς να το τεμαχίσουμε (ΜΚΕ).

Οι πιεζοηλεκτρικοί κρύσταλλοι που χρησιμοποιούνται στους διαγνωστικούς ηχολογείς έχουν μορφολογία κυκλικού δίσκου. Όταν γίνεται μία εξωτερική τάση στον κρύσταλλο, τότε δονείται και παράγοντας κύματα συγκεκριμένης συχνότητας. Παράλυτα, οι υπέρηχοι που χρησιμοποιούνται για διαγνωστικούς σκοπούς στην Ιατρική δεν είναι ένα μόνο ηχητικό κύμα αλλά μια δέσμη ηχητικών κυμάτων.[37]

Το υπερηχογράφημα χρησιμοποιεί υπερηχητικά κύματα για την εξέταση του μαστού. Η εξέταση είναι εντελώς ανώδυνη και δεν εκπέμπεται ακτινοβολία, και είναι κατάλληλη για την εξέταση πεύκων μαστών συνήθως νέων ασθενών.

Τέλος, για τη μελέτη οργάνων που σχετίζονται με το βάθος π.χ. στην κοιλιά, χρησιμοποιούνται υπέρηχοι με μικρότερες συχνότητες, ενώ για τη μελέτη επιφανειακών δομών π.χ. μαστοί, άρχεις ή μαλακά μόρια (πχ διογκώσεις δέρματος) χρησιμοποιούνται υπέρηχοι με υψηλότερες συχνότητες.[38]

### 1.5.3 Αξονική Τομογραφία

Μια αξονική τομογραφία (σάρωση CT ή CAT) χρησιμοποιεί υπολογιστή και **περιστρεφόμενες μηχανές ακτίνων X**, για τη δημιουργία δια τομεακών εικόνων του σώματος. Κατά τη διάρκεια της εξέτασης ο αξονικός τομογράφος διαιρεί το σώμα σε εγκάρσιες τομές τις οποίες και απεικονίζει σε ψηφιακή μορφή. Οι τομές αυτές είναι πάχος ενός χιλιοστού. Τα αποτελέσματα μπορούν να κρατηθούν σε ψηφιακό αρχείο (CD) και να εκτυπωθούν σε μορφή φιλμ.

Ο γιατρός για να προβεί στην εξέταση, παρέχει μια ειδική χρωστική ουσία που ονομάζεται σκιαγραφικό υλικό, έτσι ώστε οι εσωτερικές δομές να εμφανιστούν με μεγαλύτερη ακρίβεια στις εικόνες ακτίνων X. Το σκιαγραφικό υλικό αποκλείει τις ακτίνες X και εμφανίζεται λευκό στις εικόνες, με την δυνατότητα επισήμανσης **τα**

**έντερα, τα αιμοφόρα αγγεία ή άλλες δομές** στην εξεταζόμενη περιοχή. Ανάλογα με το μέρος του σώματος που εξετάζεται, ίσως χρειαστεί ένα υγρό που περιέχει το σκιαγραφικό. Εναλλακτικά, το σκιαγραφικό μπορεί να χρειαστεί να εγχυθεί σε φλέβα του βραχίονά του εξεταζόμενου.



Εικόνα 7: Αξονικός τομογράφος.

#### 1.5.4 Εφαρμογές Αξονικής Τομογραφίας

Μια αξονική τομογραφία έχει εφαρμογές, αλλά είναι ιδιαίτερα κατάλληλη για τη διάγνωση ασθενειών και την αξιολόγηση των τραυματισμών. Η τεχνική απεικόνισης μπορεί να βοηθήσει τον γιατρό σας:

- Στη διάγνωση λοιμώξεων, μυϊκών διαταραχών και καταγμάτων οστών.
- Στον εντοπισμό τη θέση των μαζών και των όγκων (συμπεριλαμβανομένου του καρκίνου).
- Να μελετήσει τα αιμοφόρα αγγεία και άλλες εσωτερικές δομές.
- Να αξιολογήσει την έκταση των εσωτερικών τραυματισμών και την εσωτερική αιμορραγία.
- Στις διαδικασίες καθοδήγησης, όπως χειρουργικές επεμβάσεις και βιοψίες.

να παρακολουθεί την αποτελεσματικότητα των θεραπειών για ορισμένες ιατρικές καταστάσεις, συμπεριλαμβανομένου του καρκίνου και των καρδιακών παθήσεων. Η εξέταση είναι ελάχιστα επεμβατική και μπορεί να διεξαχθεί γρήγορα[39].

#### 1.5.5 Δόση ακτινοβολίας

Η ακτινοβολία που προσλαμβάνει ο εξεταζόμενος είναι μικρότερη από μια τομογραφία και μεγαλύτερη από τις ακτινογραφίες. Ακόμη όταν γίνεται λήψη μιας τομής από το ίδιο το σημείο τότε έχουμε μεγαλύτερη ακτινοβολία από μια απλή ακτινογραφία. Συνήθως αυτή η μέθοδος γίνεται με ξεχωριστές περιπτώσεις για περαιτέρω διερεύνηση.[40]

#### 1.5.6 Μαγνητική τομογραφία(MRI)

Η απεικόνιση του μαστού με μαγνητική τομογραφία (MRI) αποτελεί χρήσιμο διαγνωστικό εργαλείο καθώς προσφέρεται ιδιαίτερα για την εξέταση των μαστών με πυκνό αδενικό ιστό. Το MRI μπορεί να συμβάλει αποτελεσματικά στη συλλογή πληροφοριών για ανωμαλίες του μαστού οι οποίες υποδεικνύονται είτε μαστογραφία, είτε μετά από φυσική εξέταση ή με άλλες απεικονιστικές μεθόδους. Επίσης, αξιολόγηση είναι η συμβολή του στη σταδιοποίηση του καρκίνου του μαστού, στην αξιολόγηση θεραπευτικών δυνατοτήτων καθώς και για τη μεταθεραπευτική παρακολούθηση του ασθενούς.

Το MRI αποτελεί σημαντική εφαρμογή στην αξιολόγηση μαστού με ενθέματα καθώς και της ρήξης αυτών, την απεικόνιση ανεστραμμένων θηλών, τη διερεύνηση της πρωτοπαθούς εστίας καρκίνου σε ασθενείς με διηθημένους μασχαλιαίους λεμφαδένες όπως και την ανάδειξη υπολειπόμενης νόσου μετά ογκεκτομή. Η συμβολή του MRI στην επιλογή της χειρουργικής επέμβασης (ογκεκτομή ή μαστεκτομή) και η αξιολόγηση περιστατικών για υποτροπή μετά από ογκεκτομή, θεωρείται επίσης σημαντική.[40]

Οι περιορισμοί του μαγνητικού συντονισμού στην απεικόνιση του μαστού έχουν να κάνουν με τη δυσχέρεια στη χρησιμοποίηση της μεθόδου ως εργαλείο για «screening» καρκίνου του μαστού. Περιορισμός στη διαθεσιμότητα των μηχανημάτων, υψηλό κόστος της εξέτασης, και αυξημένη μη ειδικότητα (αδυναμία διάκρισης μεταξύ κακοήθους και μη κακοήθους βλάβης) έχουν θέσει εμπόδια στη διάδοσή της χρήσης της μεθόδου. Ακόμη, η μέθοδος δεν απεικονίζει την παρουσία αποτιτανώσεων, οπότε και είναι πιθανή η μη έγκαιρη διάγνωση πρώιμων μαστογραφία αναδεικνυόμενων καρκίνων. Στους περιορισμούς της μεθόδου συμπεριλαμβάνονται η κλειστοφοβική συμπεριφορά κάποιων εξεταζόμενων και η απαραίτητη χρήση παραμαγνητικών σκιαγραφικών ουσιών. [40]

#### 1.5.7 Βιοψία

Η βιοψία μαστού είναι μια δοκιμή που αφαιρεί ιστό ή μερικές φορές υγρό από την ύποπτη περιοχή. Τα αναιρεθέντα κύτταρα εξετάζονται υπό μικροσκόπιο και εξετάζονται περαιτέρω για να ελεγχθεί η παρουσία καρκίνου του μαστού. Μια βιοψία είναι η μόνη διαγνωστική διαδικασία που μπορεί σίγουρα να καθορίσει εάν η ύποπτη περιοχή είναι καρκινική.

Τα καλά νέα είναι ότι το 80% των γυναικών που έχουν βιοψία μαστού δεν έχουν καρκίνο του μαστού.

Υπάρχουν τρεις τύποι βιοψιών:

- Λεπτή βελόνα
- Βιοψία πυρήνα-βελόνας
- Χειρουργική βιοψία

Υπάρχουν διάφοροι παράγοντες που βοηθούν τον γιατρό να αποφασίσει ποιος τύπος βιοψίας θα συστήσει. Αυτές περιλαμβάνουν την εμφάνιση, το μέγεθος και την τοποθεσία της ύποπτης περιοχής στο στήθος.

Μόλις ολοκληρωθεί η βιοψία, ένας ειδικώς εκπαιδευμένος γιατρός που ονομάζεται παθολόγος εξετάζει τα δείγματα ιστών ή υγρών υπό μικροσκόπιο, αναζητώντας μη φυσιολογικά ή καρκινικά κύτταρα. Η αναφορά παθολογίας, η οποία μπορεί να διαρκέσει μία ή δύο εβδομάδες για να ολοκληρωθεί, αποστέλλεται στον γιατρό του ασθενούς. Δείχνει αν η ύποπτη περιοχή είναι καρκινική και παρέχει μια πλήρη εικόνα της κατάστασής.[41]

Αν δεν βρεθούν καρκινικά κύτταρα, η αναφορά θα δείξει ότι τα κύτταρα στο χονδρόκοκκο είναι καλοήγη, δηλαδή μη καρκινικά. Ωστόσο, ενδέχεται να χρειαστεί κάποιο είδος παρακολούθησης ή θεραπείας.[41]

Εάν εντοπιστούν καρκινικά κύτταρα, η αναφορά θα παρέχει περισσότερες πληροφορίες για να καθορίσει τα επόμενα βήματα.

Η αναφορά για δείγμα βιοψίας πυρήνα-βελόνας θα περιλαμβάνει τον τύπο του όγκου και τον ρυθμό ανάπτυξης ή τον βαθμό του όγκου. Εάν εντοπιστεί καρκίνος, ο παθολόγος συνήθως θα εκτελέσει εργαστηριακές εξετάσεις για να εξετάσει τα κύτταρα για υποδοχείς οιστρογόνου ή προγεστερόνης.[41]

Στην περίπτωση μιας χειρουργικής βιοψίας, τα αποτελέσματα αποκαλύπτουν δεδομένα σχετικά με τον τύπο, την ποιότητα και την κατάσταση του υποδοχέα του όγκου, καθώς και την απόσταση μεταξύ του περιβάλλοντος φυσιολογικού ιστού και του εκτομηθέντος όγκου. Το περιθώριο, όπως αναφέρθηκε προηγουμένως, δείχνει αν η περιοχή είναι απαλλαγμένη από καρκινικά κύτταρα.[41]

#### 1.5.7 Πυρηνική Ιατρική

Η Πυρηνική Ιατρική είναι μία Ιατρική ειδικότητα, η οποία χρησιμοποιεί «ανοικτές πηγές ακτινοβολίας» για διαγνωστικούς και θεραπευτικούς σκοπούς. Με τον όρο «ανοικτές πηγές ακτινοβολίας» εννοούμε ραδιοϊσότοπα ή ουσίες επισημασμένες με ραδιοϊσότοπα που χορηγούνται κατά κύριο λόγο ενδοφλέβια ή από το στόμα προκειμένου να λάβουμε πληροφορίες για τη λειτουργία και μεταβολισμό οργάνων του ανθρώπινου σώματος ή για θεραπεία ορισμένων παθήσεων.

Η λειτουργική απεικόνιση της Πυρηνικής Ιατρικής είναι και η ειδοποιός διάφορα των εξετάσεων της από τις εξετάσεις με ακτίνες Χ, υπερήχους ή άλλες ακτινολογικές διαδικασίες που καθορίζουν την παρουσία της ασθένειας βασιζόμενες στην ανατομική απεικόνιση και όχι στη διαταραχή μια λειτουργίας.

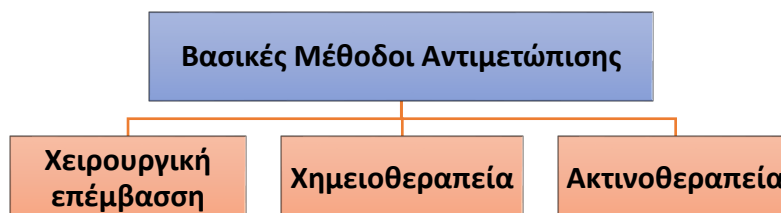
Ο τρόπος εξέτασης προσφέρει χρήσιμες πληροφορίες για την σωστή και έγκαιρη διάγνωση του ιατρικού προβλήματος. Οι εξετάσεις συνήθως διέπονται από τον κανόνα όφελος προς κίνδυνο. Το κέρδος από μια έγκαιρη και ακριβή διάγνωση αντισταθμίζει με τον παραπάνω κίνδυνο.



Για την λήψη των εικόνων χρησιμοποιείται μια ειδική κάμερα κοντά στο εξεταζόμενο μέρος του σώματος. Παράλληλα, θα γίνει λήψη μιας σειράς εικόνων οι οποίες εμφανίζονται στην κάμερα και στον ηλεκτρονικό υπολογιστή για παρκετάρω μελέτη. Η κάμερα δεν παράγει ραδιενέργεια, απλά λαμβάνει σήματα, τα σήματα από το φάρμακο. Συνήθως οι εξετάσεις πυρηνικής ιατρικής είναι το σπινθηρογράφημα οστών, νεφρών και καρδιά

## 1.6 Μέθοδοι αντιμετώπισης του καρκίνου

Σήμερα, η κλινική αντιμετώπιση του καρκίνου στηρίζεται στη χειρουργική, στην ακτινοθεραπεία, στη χορήγηση χημικοθεραπευτικών φαρμάκων (χημειοθεραπεία), καθώς και σε νεότερες θεραπείες, όπως η ορμονοθεραπεία και οι στοχευμένες θεραπείες. Με τις μεθόδους αυτές επιτυγχάνεται στην πλειονότητα των περιπτώσεων απλά μια προσωρινή υποχώρηση της νόσου, ενώ σε λίγες παρατηρείται ίαση ή παρατεταμένη ύφεση [1].



Συγκεκριμένα:

- Η **χειρουργική επέμβαση** αποτελεί, προφανώς, την παλιότερη αλλά και πιο κοινή θεραπευτική προσέγγιση κατά του καρκίνου. Η διάνοιξη του σώματος του ασθενούς με σκοπό την αφαίρεση του όγκου αποτέλεσε ουσιαστικά την πρώτη λύση κατά της νόσου, ιδιαίτερα σε περιπτώσεις που ο καρκίνος δεν έχει εξαπλωθεί σε άλλα όργανα (μετάσταση). Η εξέλιξη της τεχνολογίας τις τελευταίες δεκαετίες έχει εξοπλίσει τους επιστήμονες με εργαλεία διάγνωσης και χειρουργικής επέμβασης που καθιστούν πολύ πιο αποτελεσματικές τις προσπάθειές τους και επιδρά σε μικρό ποσοστό βλαπτικά στους περιβάλλοντες τον όγκο υγιείς ιστούς [13], [29], [30]
- Η **χρήση ακτινοβολίας (ακτινοθεραπεία)** κατά του καρκίνου αποσκοπεί στην καταστροφή των καρκινικών κυττάρων και στη συρρίκνωση των καρκινικών όγκων. Ενέργεια σε μορφή ακτίνων κατευθύνεται μέσω κυμάτων ή σωματιδίων στον όγκο των καρκινικών κυττάρων που επιδρούν βλαπτικά στα γονίδια και τη μοριακή υπόσταση των κυττάρων με αποτέλεσμα την αδυναμία περαιτέρω ανάπτυξης και διαίρεσης των καρκινικών κυττάρων. Η ακτινοθεραπεία είναι η πρώτη μη χειρουργική θεραπευτική προσέγγιση κατά του καρκίνου, και εμπλούτισε σημαντικά το οπλοστάσιο της ιατρικής στις μάχες της κατά της νόσου.
- Η **χημειοθεραπεία** περιλαμβάνει τη χρήση φαρμάκων που αποσκοπούν στην καταστροφή των καρκινικών κυττάρων. Αποτελεί συνήθη πρακτική για την αντιμετώπιση των περισσότερων ειδών καρκίνου, και χρησιμοποιείται ποικιλοτρόπως, είτε σε συνδυασμό με ακτινοβολία και χειρουργική επέμβαση

(λήψη πριν ή μετά την εγχείρηση), είτε ως μονοθεραπεία. Η χημειοθεραπεία έκανε την εμφάνιση της στη δεκαετία του 1950 και έχει έκτοτε αποδειχτεί εξαιρετικά αποτελεσματική σε αρκετές περιπτώσεις. Παγκόσμια, διενεργούνται συνεχείς έρευνες γύρω από υπάρχοντες ή και νέες φαρμακευτικές ουσίες, που στόχο έχουν την πληρέστερη αντιμετώπιση διαφόρων ειδών καρκίνου, με όσο το δυνατόν λιγότερες επιβλαβείς επιδράσεις στα υγιή κύτταρα [29], [31], [32], [33].

- Η **ορμονοθεραπεία** είναι μια αντικαρκινική θεραπεία που εμποδίζει την παραγωγή συγκεκριμένων ορμονών στον οργανισμό, δεσμεύοντας τους ορμονικούς υποδοχείς στην επιφάνεια των καρκινικών κυττάρων ή υποκαθιστώντας τις φυσικές ορμόνες με χημικώς όμοιές τους, τις οποίες όμως δεν μπορούν να χρησιμοποιήσουν τα καρκινικά κύτταρα, δυσχεραίνοντας έτσι την ανάπτυξη τους
- Οι **στοχευμένες θεραπείες** χρησιμοποιούν τις φυσιολογικές ουσίες του σώματος ή φάρμακα κατασκευασμένα από φυσιολογικά συστατικά του σώματος. Χαρακτηρίζονται ως «έξυπνες βόμβες», επειδή κατευθύνονται σε ειδικούς στόχους στα καρκινικά κύτταρα, αφήνοντας ανεπηρέαστα τα φυσιολογικά κύτταρα του οργανισμού. Κατ' αυτόν τον τρόπο η θεραπεία γίνεται περισσότερο συγκεκριμένη, πιο ειδική, λιγότερο τοξική και μπορεί να βοηθήσει στον έλεγχο των παρενεργειών των άλλων αντικαρκινικών θεραπειών, όπως της χημειοθεραπείας. Αποτελεί ένα σχετικά νέο τρόπο αντικαρκινικής θεραπείας. Οι θεραπείες αυτές καλούνται επίσης «βιολογικοί τροποποιητές της ανταπόκρισης», «ανοσοθεραπευτικοί παράγοντες», καθώς και «βιολογικοί παράγοντες», γιατί διεγείρουν τη φυσιολογική ανταπόκριση του σώματος στην αντιμετώπιση του καρκίνου. Είναι διαθέσιμες σήμερα πολλές μορφές βιολογικών θεραπευτικών παραγόντων [5]. Ανάμεσα σε αυτές περιλαμβάνονται τα μονοκλωνικά αντισώματα (όπως είναι η τρανστουζουμάβη και η ριτουξιμάβη), τα αντικαρκινικά εμβόλια, οι αυξητικοί παράγοντες των κυττάρων του αίματος, ακόμα οι αναστολείς των αυξητικών παραγόντων, οι αντιαγγειογενετικές ουσίες, η ιντερφερόνη, η ιντερλευκίνη II, και τα φάρμακα γονιδιακής θεραπείας.
- Η **ανοσοθεραπεία** αποτελεί τύπο θεραπείας που ενεργοποιεί και κατευθύνει το ανοσοποιητικό σύστημα του οργανισμού να καταπολεμήσει την ανάπτυξη των καρκινικών κυττάρων. Μπορεί να γίνει μόνη της ή σε συνδυασμό με άλλες θεραπείες για τον καρκίνο. Σε σχετικές έρευνες αποδείχθηκε ότι πλεονεκτεί σε αρκετά σημεία, επειδή το ανοσοποιητικό σύστημα διαδραματίζει ούτως ή άλλως ουσιαστικό ρόλο προστατεύοντας τον οργανισμό εναντίον του καρκίνου. Επιπλέον, το ανοσοποιητικό σύστημα συμβάλλει στην καταπολέμηση του καρκίνου όταν αυτός έχει ήδη αναπτυχθεί.

Η πρόοδος στον τομέα της μοριακής βιολογίας και η κατανόηση της καρκινογένεσης του καρκίνου οδήγησαν στην ανάπτυξη διαφόρων τεχνικών γονιδιακής θεραπείας. Με τον όρο γονιδιακή θεραπεία εννοούμε την εισαγωγή στον οργανισμό «κομματιών» γενετικού υλικού με διάφορους φορείς (π.χ. ιούς). Διακρίνεται σε διορθωτική,



κυτταροτοξική και ανοσορυθμιστική. Η διορθωτική γονιδιακή θεραπεία έχει ως σκοπό την αποκατάσταση της λειτουργίας των διαταραγμένων ογκοκατασταλτικών γονιδίων ή/και την αναστολή της ενεργοποίησης των ογκογονιδίων που προάγουν την αύξηση του όγκου. Παρόλο που σε μοντέλα ζώων φαίνεται να έχει θετικά αποτελέσματα, δε συμβαίνει το ίδιο στον άνθρωπο, γιατί είναι πολύ δύσκολο να «διορθωθεί» κάθε ένα καρκινικό κύτταρο. Η κυτταροτοξική γονιδιακή θεραπεία στηρίζεται κατά κύριο λόγο στην εισαγωγή στον οργανισμό γονιδίων που επάγουν τον κυτταρικό θάνατο των καρκινικών κυττάρων, ενώ η ανοσορυθμιστική γονιδιακή θεραπεία στοχεύει στην ενεργοποίηση του ανοσοποιητικού συστήματος, ώστε ο ίδιος ο οργανισμός να καταπολεμήσει τα καρκινικά κύτταρα.

Παρόλο που υπάρχουν ενθαρρυντικά αποτελέσματα σε πειραματικές μελέτες, η γονιδιακή θεραπεία βρίσκεται ακόμα σε αρχικό στάδιο εξαιτίας της δυσκολίας δημιουργίας φορέων του γενετικού υλικού μέσα στον ανθρώπινο οργανισμό [17], [31], [32]-[36].

### 1.7 Στατιστικά για τον καρκίνο του μαστού

**Καρκίνος του μαστού:** Κάθε χρόνο στη χώρα μας εντοπίζονται περίπου **6.000 νέα περιστατικά!**

#### **Ο καρκίνος του μαστού Παγκοσμίως:**

Σύμφωνα με τον Παγκόσμιο Οργανισμό Υγείας, περισσότερες από **1.600.000** γυναίκες το χρόνο διαγιγνώσκονται με καρκίνο του μαστού.

#### **Καρκίνος του μαστού στην Ευρώπη:**

- Στην Ευρώπη, περισσότερες από 450.000 γυναίκες εμφανίζουν καρκίνο στο μαστό. Περίπου **140.000** χάνουν τη ζωή τους λόγω της ασθένειας.

#### **Καρκίνος μαστού στις ΗΠΑ:**

- Στις ΗΠΑ το 2004 οι νέες περιπτώσεις καρκίνου του μαστού στις γυναίκες ήταν **217.000** περίπου. Σήμερα, στη χώρα αυτή ζουν περισσότερες από δύο εκατομμύρια γυναίκες με καρκίνο του μαστού.

#### **Ο καρκίνος του μαστού στην Ελλάδα:**

- Στην Ελλάδα αναφέρονται **6.000** περίπου νέες περιπτώσεις το χρόνο, ενώ υπολογίζεται ότι **1 στις 8** γυναίκες παγκοσμίως θα παρουσιάσει καρκίνο μαστού σε κάποια φάση της ζωής της.

## Βιβλιογραφία 1<sup>ου</sup> Κεφαλαίου

- [1]. V. Pudata, Subrahmanyam V., Jhansi K., (2011), 'A Short Note on Cancer', Journal of Carcinogenesis & Mutagene
- [2]. <http://coloncancer.about.com/od/glossaries/g/Cancer.htm>
- [3]. <http://en.wikipedia.org/wiki/Cancer>
- [4]. Nair LS, Mahesh S, Smitha LS, Sujathan K, Remani P, (2011), 'Expression of Canavalia gladiata lectin in leukemic cells', J Cancer Sci Ther 3: 088-091
- [5]. Bali A, Singh MP, Padmavathi, Khorate M, Ahmed J, (2010), 'Malignant Fibrous Histiocytoma - An Unusual Transformation from Benign to Malignant', J Cancer Sci Ther 2: 053-057
- [6]. Naga Deepthi CH, VVL Pavan Kumar A, Rameshbabu, Indirapriyadarshini U, (2011), 'Role of Tumor Suppressor Protein p53 in Apoptosis and Cancer Therapy'. J Cancer Sci Ther S17:001
- [7]. Turconi G, (2011), 'Healthy Aging: Nutritional Intervention to Improve and Extend Quality of Life among Older People'. J Nutr Food Sci 1:e101
- [8]. Liu CY, Chang LC, Yang SW (2011), 'Metastatic Hepatocellular Carcinoma to the Nasal Cavity: A Case Report and Review of the Literature', J Cancer Sci Ther 3: 081-083
- [9]. Kikawa KD, Noah T, Ahwah SM, Pardini RS, (2011), 'Docosahexaenoic Acid (DHA) Induces P53-Dependent Growth Inhibition in Transformed Colon and Lung Cell Lines Expressing Wildtype P53', J Cancer Sci Ther 3: 182-185
- [10]. Nagy M.A. (2011), 'HIF-1 is the Commander of Gateways to Cancer', J Cancer Sci Ther 3: 035-040
- [11]. Kardamakis D., (2004), 'Πανεπιστημιακές παραδόσεις: Ακτινοβιολογίας, Ακτινοπροστασίας και Ακτινοθεραπείας'
- [12]. <https://ourworldindata.org/cancer#global-perspective-on-cancer-prevalence>
- [13]. Παντελάκος Σ. Παναγιώτης, (2005) 'Διασφάλιση ποιότητας στην Ακτινοθεραπεία.' 2<sup>ο</sup> Διεταρικό Αντικαρκινικό Συνέδριο 364-375
- [14]. S. Libson, M. Lippman, (2014), 'A review of clinical aspects of breast cancer' International Review of Psychiatry,; 26(1): 4-15
- [15]. Harrison's Principles of Internal Medicine. 14th edition 282-287
- [16]. Landis SH, Murray T, Bolden S, et al. (1998), 'Cancer statistics, 1999'. CA: A Cancer Journal for Clinicians 48 6-29
- [17]. Zangemeister-Wittke U, (2005) 'Antibodies for targeted cancer therapy -- technical aspects and clinical perspectives', Pathobiology.; 72(6):279-86
- [18]. Chan AT, Manson JE, Albert CM, Chae CU et al., (2006), 'Non-steroidal anti-inflammatory drugs, acetaminophen, and the risk of cardiovascular events', Circulation 113(12) 1578-87
- [19]. Emery IF, Battelli C, Cai L, Hayes DM (2011) Assessment of DNA Repair Mechanisms to Determine the Susceptibility of Non-small Cell Lung Cancer to Alkylating Agents. J Cancer Sci Ther 3: 092-095

- [20]. Evans S, Moieni M, Subramanian S, Tsao JCI, Sternlieb B, et al. (2011) “Now I see a brighter day”: expectations and perceived benefits of an Iyengar yoga intervention for young patients with rheumatoid arthritis. *J Yoga Phys Therapy* 1:101
- [21]. Oncology Centre: Καρκίνος του μαστού: αιτιολογία και παράγοντες κινδύνου’ Διαθέσιμο στο: <http://oncologikoathens.gr/gr/interestingissues/132362531>
- [22]. ‘Είδη Καρκίνου Μαστού (Στήθους) & Βαθμοί Διαφοροποίησης’, [fyssas.gr](http://fyssas.gr)
- [23]. ‘Πορογενής & Λοβιακός Διηθητικός Καρκίνος’, [fyssas.gr](http://fyssas.gr)
- [24]. ‘In Situ Καρκίνος Μαστού (Στήθους): Θεραπεία – Τι Είναι’;, [fyssas.gr](http://fyssas.gr)
- [25]. ‘Σάρκωμα Μαστού’, [fyssas.gr](http://fyssas.gr) .
- [26]. ‘Νόσος Paget Του Μαστού’, [fyssas.gr](http://fyssas.gr)
- [27]. ‘Φλεγμονώδης Καρκίνος Μαστού’, [fyssas.gr](http://fyssas.gr) .
- [28]. Kirkova J, Davis MP, Walsh D, Tiernan E, O’Leary N, LeGrand SB, Lagman RL, Russell KM, (2006), Cancer symptom assessment instruments: a systematic review. *J Clin Oncol*. Mar 20;24(9):1459-73
- [29]. Βώρος Δ. (2005), Απώτερες επιπτώσεις ογκολογικών εγχειρήσεων. 2ο Διαιτητικό Αντικαρκινικό Συνέδριο (2005) 269-272
- [30]. Harvey A. Richard, Champe C. Pamela. *Pharmacology*. 2nd edition 425-453
- [31]. Morris MJ, Tong WP, Gordon-Cardo C, et al. (2002), Phase I trial of BCL-2 antisense oligonucleotide (G3139) administered by continuous intravenous infusion in patients with advanced cancer. *Clinical Cancer Research* 8 (2002) 679-683
- [32]. Σανιδάς Η, Φίλης Δ. Διαχείριση του ογκολογικού ασθενή σε περιβάλλον ποιότητας Μονάδα χειρουργικής ογκολογίας. 2ο Διαιτητικό Αντικαρκινικό Συνέδριο (2005), 299-301
- [33]. Thalasila A, Poplin E, Shih J, et al. A phase I trial of weekly Paclitaxel 13-cis retinoic acid, and interferon alpha in patients with prostate cancer and other advanced malignancies. *Cancer Chemotherapy and Pharmacology* 52 (2003) 119-124
- [34]. Καλόφωνος ΧΠ, Κούτρας ΑΚ. Ο ρόλος των μονοκλωνικών αντισωμάτων. 2ο Διαιτητικό Αντικαρκινικό Συνέδριο (2005) 68-73
- [35]. Τσαβαρής ΝΒ. Βιολογικές θεραπείες-Γαστρεντερολογικό Σύστημα. 2ο Διαιτητικό Αντικαρκινικό Συνέδριο (2005) 65-68
- [36]. Χριστακοπούλου Ι. Θεραπεία με Ραδιοεπισημασμένα βιομόρια. 2ο Διαιτητικό Αντικαρκινικό Συνέδριο (2005) 120-124
- [37]. <http://users.teiath.gr>
- [38]. <https://echohealth.gr>
- [39]. <https://www.iprolipsi.gr>
- [40]. <http://www.cancer-society.gr>
- [41]. <https://www.nationalbreastcancer.org>
- [42]. <http://www.Breastcure.com>

[43]. <http://www.Medinova.gr>

Page Intentionally Left Blank

## Κεφάλαιο 2ο

### Απεικόνιση του Καρκίνου του Μαστού - Μαστογραφία

Η πιο συνήθης και κύρια εξέταση απεικόνισης του μαστού είναι η μαστογραφία, η οποία έχει αποδειχθεί ότι οδηγεί στη δραστική μείωση της θνησιμότητας από τον καρκίνο του μαστού. Για το συγκεκριμένο είδος απεικόνισης η ευαισθησία, βάσει πληθυσμού, κυμαίνεται από 75% έως 80%, μειώνεται όμως σε γυναίκες οι οποίες εμφανίζουν υψηλή επικινδυνότητα, με πυκνά στήθη, στο 50%. Ενισχυτικά της μαστογραφίας, στη διάγνωση της νόσου, λειτουργούν το υπερηχογράφημα του μαστού και η αντίθεσης- ενίσχυσης μαγνητική τομογραφία (MRI). Κάθε απεικονιστική μέθοδος δεν υπερτερεί της άλλης, ούτε την αντικαθιστά. Η μία συμπληρώνει την άλλη, καθεμία έχει τις ενδείξεις της, για το λόγο αυτό πρέπει να γίνονται με μια λογική σειρά ανάλογα με τα ευρήματα σε κάθε περίπτωση. Τα αποτελέσματα από αυτές τις εξετάσεις συγκεντρώνονται και αξιολογούνται από κατάλληλα εκπαιδευμένο ακτινοδιαγνωστικά. Σήμερα οι προσπάθειες επικεντρώνονται στην ανάπτυξη νέων τεχνικών απεικόνισης (αξονική τομογραφία (CT), τομογραφία εκπομπής ποζιτρονίων (PET), τομογράφος PET/CT, απεικόνιση μαστού με ραδιονουκλίδια) και στη βελτίωση των ήδη υπάρχοντων τεχνικών (ψηφιακή τομοσύνθεση του μαστού, μαστογραφία με ενισχυμένη αντίθεση και υπερηχογράφημα με πλαστογραφία ή μικροφυσαλλίδες), με σκοπό τη διάγνωση του καρκίνου του μαστού σε πρώιμο στάδιο.[1]

#### 2.1 Μαστογραφία

Η μαστογραφία είναι ένας ειδικός τύπος απεικόνισης του μαστού που χρησιμοποιεί ακτίνες X χαμηλής δόσης προκειμένου να ανιχνευθούν πιθανές παθολογικές αλλοιώσεις του μαστού, έγκαιρα πριν την παρουσίαση σχετικών συμπτωμάτων ή ψηλάφηση των όγκων. Τα φωτόνια που χρησιμοποιούνται είναι σχετικά μικρών ενεργειών (20 με 30 KeV) για να απεικονίσουν πιο έντονα τις αλλαγές στους ιστούς και το λίπος.[2]

#### 2.2 Βασική αρχή της μαστογραφίας

Όπως προαναφέρθηκε η μαστογραφία είναι μια ακτινολογική εξέταση από την οποία λαμβάνονται πληροφορίες σχετικά με το περιεχόμενο του μαστού. Για να επιτευχθεί απεικόνιση υψηλής αντίθεσης και ευκρίνειας χρησιμοποιούνται ειδικά συστήματα λυχνιών που παράγουν ακτινοβολία ακτίνων X χαμηλής ενέργειας (20 έως 30 KeV). Ανάλογα με το πάχος και τη σύσταση του μαστού η υψηλή τάση που εφαρμόζεται στη λυχνία είναι μεταξύ των τιμών 25 και 35 kVp.[3]

Η αντίθεση δύο ιστών με διαφορετικούς γραμμικούς συντελεστές και διαφορετικό πάχος δίνεται από την παρακάτω σχέση :

$$C = \mu^{2\Delta x_2} - \mu^{1\Delta x_1}$$

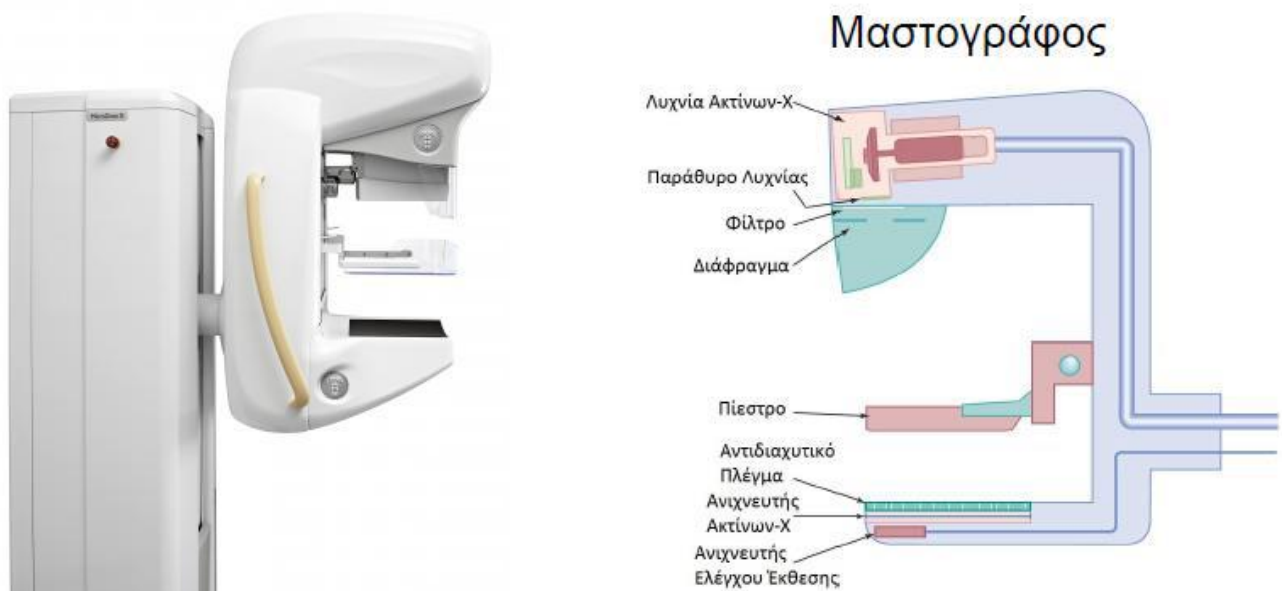
όπου  $\mu_1$  και  $\mu_2$  είναι οι γραμμικοί συντελεστές εξασθένισης και  $\Delta x_1, \Delta x_2$  είναι τα αντίστοιχα πάχη των ιστών.

Όπως είναι γνωστό στις υψηλές ενέργειες η διαφορά μεταξύ των γραμμικών συντελεστών εξασθένισης μειώνεται, ενώ στις χαμηλές αυξάνεται. Συνεπώς είναι απαραίτητο η χρήση χαμηλών kV για την καλύτερη επεξεργασία της πληροφορίας στην ακτινογραφική εικόνα.

Στις χαμηλές ενέργειες ο βασικός μηχανισμός αλληλεπίδρασης της ακτινοβολίας με τους ιστούς είναι το φωτοηλεκτρικό φαινόμενο. Με το φωτοηλεκτρικό φαινόμενο οι διαφορές στον ατομικό αριθμό μεταξύ δυο διαφορετικών ιστών ενισχύονται με αποτέλεσμα την καλύτερη κατανόηση της εικόνας. Με τη χρήση υψηλών kV θα ήταν αδύνατο να κατανοηθούν οι διαφορές στην εικόνα ανάμεσα σε έναν υγιή και καρκινικό ιστό.[3]

### 2.3 Εξοπλισμός μαστογράφου

Το σύστημα μαστογραφίας αποτελείται από το σώμα μέσα στην οποία βρίσκεται η γεννήτρια υψηλής τάσης. Τα μέρη τα οποία αποτελούν το σώμα είναι η τράπεζα χειρισμού καθώς και ένα προστατευτικό φύλλο από μολυβδύαλο για την προστασία του χειριστή. Τα κύρια μέρη που αφορούν την απεικόνιση του συστήματος είναι η ακτινολογική λυχνία, καθώς και το σύστημα λήψης εικόνας όπου την αποτελούν η βάση στήριξης, όπου εκεί τοποθετείται ο εξεταζόμενος μαστός, το αντιδιαχυτικό διάφραγμα και η θέση της κασέτας. Αυτά τα μέρη συνδέονται με το σώμα μέσω του περιστρεφόμενου βραχίονα. Οπότε ο προσανατολισμός του βραχίονα καθορίζεται από την επιθυμητή εξέταση που θέλει ο χειριστής να επιτευχθεί.



Εικόνα 8: Εξοπλισμός μαστογραφίας.

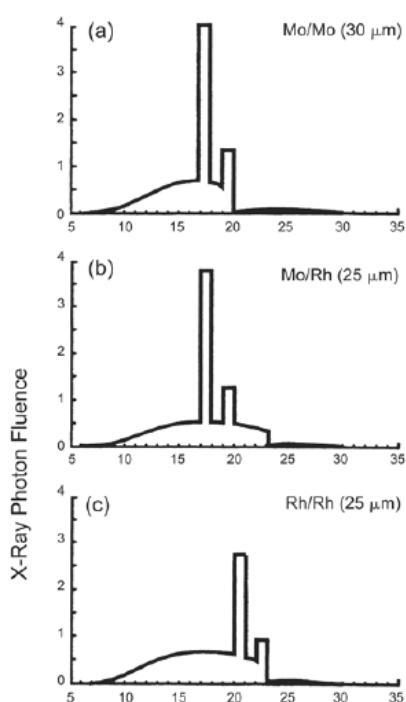
## 2.4 Λυχνία ακτινών Χ, φίλτρα και φάσμα

Στο σύστημα μαστογραφίας η ειδική λυχνία ακτινών Χ παράγει φωτόνια ακτινών Χ χαμηλής ενέργειας στην οποία η υψηλή τάσης της κυμαίνεται από 25-30 KeV. Η άνοδος είναι περιστρεφόμενη και η εστία της δημιουργείται από μολυβδαίνιο (Mo),

το οποίο αποσκοπεί στην παραγωγή χαμηλής ενέργειας φωτονίων. **Το φάσμα της ακτινοβολίας στις λυχνίες αυτές είναι σύνθετο**, αποτελούμενο από την ακτινοβολία πεδήσεως αλλά και την χαρακτηριστική ακτινοβολία του μολυβδαίνιου με ενέργειες 17,5 και 19,6. Στην έξοδο της λυχνίας τοποθετείται φίλτρο πάχους 30μm από μολυβδαίνιο, ώστε να αποκόπτει τις χαμηλές και υψηλές ενέργειες με αποτέλεσμα να παρέχει ενέργειες έως από 15 έως 20 keV. Οι εστίες στους μαστογράφους έχουν διαστάσεις μήκους 0,1-0,3 mm. Στους σύγχρονους μαστογράφους γίνεται χρήση διπλών εστιών ώστε να επιλέγεται κατάλληλα το μέγεθος του εστιακού σημείου ανάλογα με τον τύπο του μαστού. Όσο προσεγγίζει το ιδανικό μέγεθος η εστία τόσο καλύτερη είναι η διακριτική ικανότητα της εικόνας (αποφυγή παρασκιάς).



Η ύπαρξη των φίλτρων είναι απαραίτητη ώστε να αποκόπτονται τα χαμηλής ενέργειας φωτόνια που αυξάνουν τη δόση στον ασθενή χωρίς κανένα διαγνωστικό αποτέλεσμα. Εκτός από μολυβδύαλο που αναφέρθηκε, γίνεται χρήση και φίλτρου από Ρόδιο (Rh) το οποίο παρέχει μεγαλύτερης ενέργειας φωτόνια έως 23 KeV που απεικονίζουν αποτελεσματικότερα τους πυκνούς μαστούς. Οι συνδυασμοί υλικού εστίας-φίλτρου συμβολίζονται με Mo/Mo ή Mo/Rh και επιλέγεται ανάλογα με τη δομή του μαστού. Στην Εικόνα 10 παρουσιάζεται το φάσμα των ακτίνων X, προερχόμενο από άνοδο μολυβδαίνιου με (Mo/Mo) , από άνοδο (Mo/Rh) και (Rh/Rh), αντίστοιχα. Ο συνδυασμός Rh/Rh παρέχει πιο διεισδυτική ακτινοβολία με αποτέλεσμα να είναι πιο αποδοτική για την απεικόνιση πυκνών μαστών σε

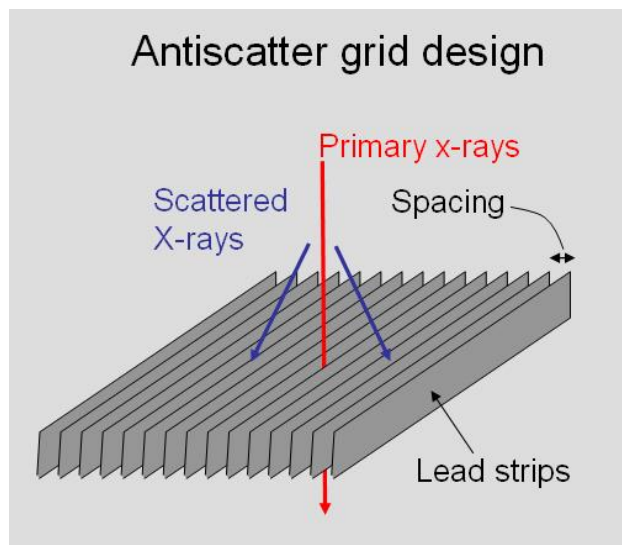


Εικόνα 9: Παραδείγματα φάσματος μαστογραφίας ακτίνων X.

συνδυασμό με μικρότερη απορροφούμενη δόση. Συνεπώς η επιλογή του υλικού εξαρτάται από την πυκνότητα και το πάχος του μαστού που θα απεικονιστεί.[4]

## 2.5 Αντιδιαχυτικό διάφραγμα

Για την μείωση της σκεδαζόμενης ακτινοβολίας η οποία φτάνει στο φιλμ γίνεται τοποθέτηση αντιδιαχυτικού διαφράγματος ανάμεσα στο μαστό και στην ενισχυτική πινακίδα-φιλμ. Στη μαστογραφία χρησιμοποιούνται παλινδρομικά αντιδιαχυτικά διαφράγματα με λόγο 4:1 ή 5:1. Το αντιδιαχυτικό διάφραγμα αυξάνει τη δόση ενισχύοντας όμως την ποιότητα εικόνας με δυνατότητα ανίχνευσης μικρών αποτυπωμάτων. Η δυνατότητα αποφυγής του αντιδιαχυτικού διαφράγματος συμβαίνει μόνο σε μικρούς μαστούς χωρίς σημαντική επίπτωση στην ποιότητα εικόνας.[3,4]



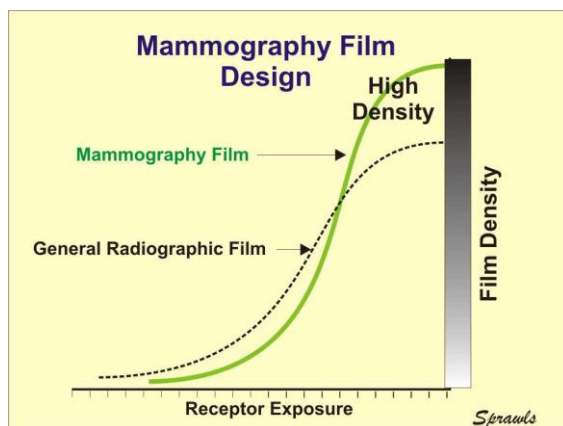
Εικόνα 10: Αντιδιαχυτικό διάφραγμα ακτίνων X.

## 2.6 Ανιχνευτές

### 2.6.1 Συστήματα ενισχυτικών πινακίδων- φιλμ

Ο συνδυασμός φιλμ και ενισχυτικών πινακίδων αποτελεί βασικό μέρος του εξοπλισμού της μαστογραφίας, καθώς επικαθορίζει την ευκρινέστερη απεικόνιση του μαστικού ιστού. Το ακτινογραφικό φιλμ αποτελεί το πιο διαδεδομένο μέσο καταγραφής διαγνωστικών εικόνων, με βασικό του στοιχείο μια φωτοευαίσθητη ουσία που ονομάζεται φωτογραφικό γαλάκτωμα (emulsion). Τα ακτινογραφικά φιλμ διακρίνονται σε φιλμ μονού ή διπλού φωτογραφικού γαλακτώματος.[5] Το φιλμ που χρησιμοποιείται για την καταγραφή των μαστογραφιών έχει κόκκους μικρών διαστάσεων (λεπτόκοκκο) και είναι μονού φωτογραφικού γαλακτώματος. [6] Οι ενισχυτικές πινακίδες έχουν σκοπό τη μετατροπή των ακτίνων X σε άλλης μορφής ακτινοβολία, στην οποία είναι περισσότερο ευαίσθητο το φιλμ. Η μετατροπή αυτή βασίζεται στο φαινόμενο του φθορισμού, το οποίο συνίσταται στην εκπομπή φωτονίων από ορισμένες ουσίες, όταν αυτές διεγείρονται με ιοντίζουσες ακτινοβολίες (φωτόνια X,  $\gamma$ , ηλεκτρόνια κ.λπ.) Τα υλικά που παρουσιάζουν φθορισμό ονομάζονται και σπινθηριστές (scintillators).[5] Φιλμ μονής επίστρωσης, τοποθετούνται σε κασέτες με μια ενισχυτική πινακίδα ( $\text{CaWO}_4$  ή σπανίων γαιών). Οι κασέτες είναι ειδικής

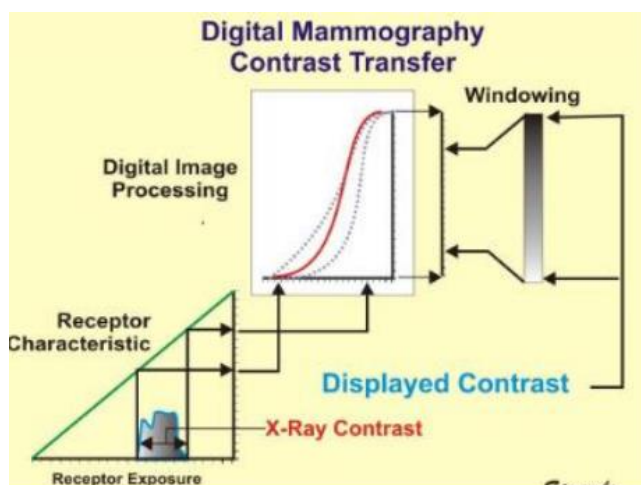
κατασκευής, διαστάσεων 18 x 24 ή 24 x 30 εξασφαλίζοντας την καλύτερη δυνατή επαφή ενισχυτικής πινακίδας - φιλμ.



Εικόνα 11: Σχέση μαστογραφικού φιλμ- γενικού ακτινογραφικού φιλμ.

### 2.6.2 Ψηφιακός ανιχνευτής

Οι ψηφιακοί ανιχνευτές βασίζονται στη δημιουργία ψηφιακού σήματος, καθώς τα φωτόνια της δέσμης ανιχνεύονται απευθείας από φωτοευαίσθητους ημιαγωγούς (συνήθως από Σελήνιο). Στην πορεία η εικόνα αξιολογείται σε ψηφιακές οθόνες.[6]



Εικόνα 12: Μεταφορά αντίθεσης σε ψηφιακή μαστογραφία..

### 2.7 Σύστημα συμπίεσης μαστού

Η αντίθεση θέματος εξαρτάται από τη διαφορά  $\mu_2 - \mu_1$ , στην οποία η αύξηση της διαφοράς των γραμμικών συντελεστών εξασθένησης  $\mu_2 - \mu_1$  αυξάνει στα χαμηλά kV. Αυτό είναι επιθυμητό μόνο αν θεωρηθεί ότι η διαφορά των παχών  $x_2 - x_1$  των αντίστοιχων γειτονικών ιστών είναι μηδενική, δηλαδή  $x_2 = x_1$ . Συνεπώς στη μαστογραφία για να επιτευχθεί το γεγονός αυτό, ο μαστός τοποθετείται σε μια διάταξη στην οποία συμπίεζεται, ώστε το πάχος των ιστών να παραμένει σταθερό σε όλη την διάρκεια της ακτινοβολήσής του. Επίσης κατά την ακινητοποίηση του μαστού εξασφαλίζεται η αποφυγή της ασάφειας λόγω κίνησης. Με την πίεση του μαστού το πάχος μειώνεται στα **4-6 cm**. Το αποτέλεσμα αυτής της διαδικασίας είναι ο μαστός να

απορροφά λιγότερη δόση και να μειώνει την σκεδαζόμενη ακτινοβολία η οποία υποβιβάζει την ευκρίνεια της εικόνας.[3,4]



Εικόνα 13: Σύστημα συμπίεσης του μαστού στη μαστογραφία.

## 2.8 Σύστημα αυτόματης έκθεσης (AEC)

Οι σύγχρονοι μαστογράφοι διαθέτουν σύστημα αυτόματου ελέγχου έκθεσης. Οι ανιχνευτές AEC τοποθετούνται στο πίσω μέρος της κασέτας. Οι ανιχνευτές αυτοί τερματίζουν την έκθεση των ακτίνων X όταν ο αριθμός των φωτονίων που προσπίπτει στο φιλμ είναι ικανός ώστε να προκαλέσει αμαύρωση του. Ένα πρόβλημα που παρατηρείται στους AEC έχει σχέση με την διεισδυτικότητα της δέσμης. Αν ακτινοβοληθεί ένας παχύς μαστός τα φωτόνια χαμηλής ενέργειας θα εξασθενούν στο μαστό, με αποτέλεσμα τον πρόωρο τερματισμό της έκθεσης, νωρίτερα από μια τυπική ταινία.

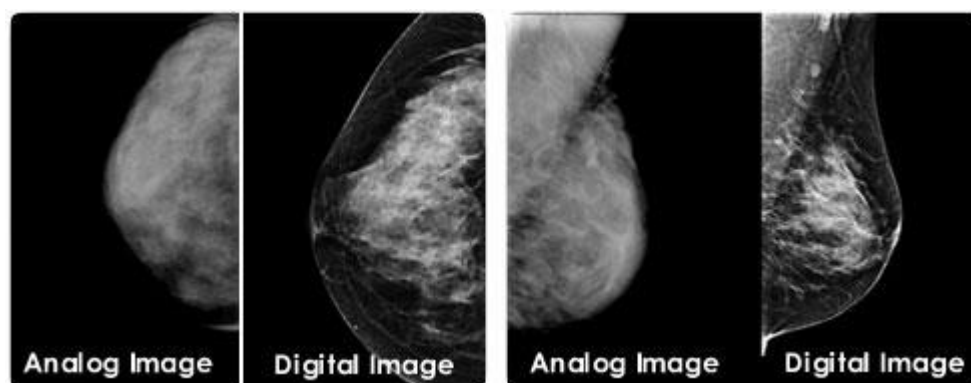
## 2.9 Ψηφιακή Μαστογραφία

Οι μελέτες και οι βελτιώσεις στα συστήματα απεικόνισης του μαστού, τις τελευταίες δεκαετίες, βασιζόμενες στην αλματώδη εξέλιξη της τεχνολογίας, οδήγησαν ερευνητές και ακτινοδιαγνώστες στην ψηφιακή μαστογραφία (**full-field digital mammography**). Η βασική διαφορά της με την κλασική μαστογραφία έγκειται στον διαφορετικό τρόπο ανίχνευσης των εξερχόμενων φωτονίων από το μαστό. Στην κλασική χρησιμοποιείται φιλμ ενώ στην ψηφιακή ημιαγωγοί. Διακρίνονται δύο κατηγορίες ψηφιακής μαστογραφίας, η έμμεση (CR) και η άμεση (DR).

Αν και τα τελευταία χρόνια η τεχνολογία φθορίζουσας οθόνης- φιλμ έχει κάνει σημαντική πρόοδο στην ανίχνευση και την παρακολούθηση ύποπτων αλλοιώσεων του μαστού, υπόκειται σε περιορισμούς όσον αφορά την ανίχνευση συγκεκριμένων αλλοιώσεων του μαλακού ιστού, ιδιαίτερα σε περιπτώσεις που η παρουσία πυκνού αδενικού ιστού είναι έντονη.[7]

Η ψηφιακή μαστογραφία διαχωρίζει τις διαδικασίες απόκτησης εικόνας και απεικόνισης, επιτρέποντας με τον τρόπο αυτό την δυνατότητα βελτίωσης και των δύο.

Με την ψηφιακή μαστογραφία επίσης, μέσω της επεξεργασίας ψηφιακών δεδομένων κατά την απεικόνιση, καθίσταται δυνατή η διαχείριση του βαθμού της αντίθεσης σε μια εικόνα, με τελικό αποτέλεσμα την αύξηση της αντίθεσης σε περιοχές με πυκνό ιστό του μαστού.[8,9]



Εικόνα 14: Αναλογική και ψηφιακή εικόνα μαστού.

#### 2.9.1 Διαφορές της ψηφιακής από την αναλογική μαστογραφία

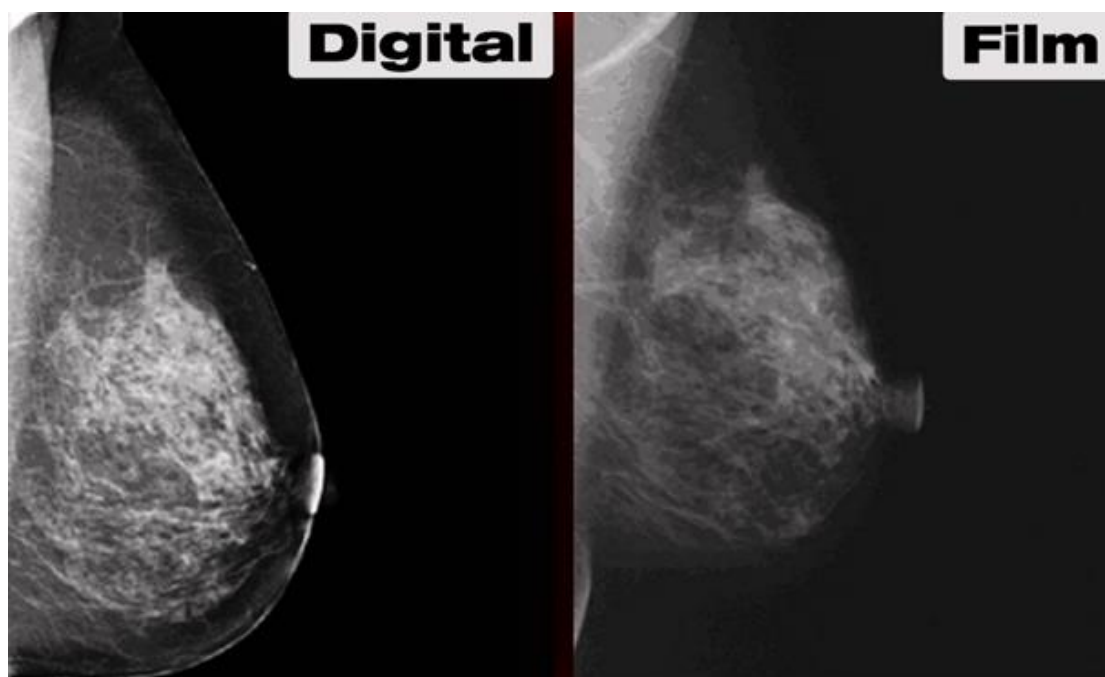
Η βασική διαφορά της ψηφιακής μαστογραφίας με την κλασσική μαστογραφία εντοπίζεται στην αντικατάσταση της φθορίζουσας οθόνης και του φιλμ, που χρησιμοποιούνται στην κλασσική μαστογραφία, με έναν ψηφιακό ανιχνευτή. Ο ψηφιακός ανιχνευτής καταγράφει τις ακτίνες X που ανιχνεύει ως ηλεκτρικά σήματα, τα οποία στη συνέχεια μετατρέπονται σε ψηφιακά δεδομένα. [10]

Με την έλευση της ψηφιακής μαστογραφίας, γεννήθηκε ταυτόχρονα και η ελπίδα για ανίχνευση περιπτώσεων καρκίνου του μαστού οι οποίοι στην κλασσική μαστογραφία δεν ήταν ορατοί λόγω ύπαρξης πυκνού φυσιολογικού ιστού. Σε σύγκριση με το κλασσικό ακτινογραφικό φιλμ, ο ψηφιακός ανιχνευτής έχει καλύτερη διακριτική ικανότητα αντίθεσης και μεγαλύτερη ελευθερία σε επιλογές και περιορισμούς (wider latitude).[11] Εντούτοις, ένας ψηφιακός ανιχνευτής δεν έχει τόσο υψηλή χωρική διακριτική ικανότητα όπως το φιλμ, για αντικείμενα υψηλής αντίθεσης. Η χωρική διακριτική ικανότητα και η διακριτική ικανότητα αντίθεσης είναι δύο χαρακτηριστικά που παίζουν εξίσου σημαντικό ρόλο στον χαρακτηρισμό μιας περίπτωσης ως καρκίνο του μαστού. Έτσι, είναι δύσκολο να επιτευχθεί μια ισορροπία μεταξύ των δύο αυτών χαρακτηριστικών, προκειμένου μια εκ των δύο τεχνικών μαστογραφίας να υπερέχει της άλλης [10].

Επιπλέον, η κλινική απόδοση της ψηφιακής μαστογραφίας εξαρτάται από την απόδοση του ακτινολόγου κατά την ερμηνεία των εικόνων καθώς και από την προσωπική του εμπειρία και ικανότητα κατά τη φάση επεξεργασίας των αρχικών εικόνων.

Συμπερασματικά, καμία τεχνική δεν υπερέχει της άλλης στο σύνολο των περιπτώσεων. Δηλαδή, μπορεί η ψηφιακή μαστογραφία να έχει καλύτερη διακριτική ικανότητα αντίθεσης και να προσφέρει περισσότερες επιλογές επεξεργασίας των εικόνων σε σχέση με την κλασσική μαστογραφία, αλλά κάποια χαρακτηριστικά της κλασσικής

μαστογραφίας, όπως για παράδειγμα η χωρική διακριτική ικανότητα του φιλμ, καθιστούν και τις δύο τεχνικές αναγκαίες. [10]



Εικόνα 15: Αριστερά: Ψηφιακή απεικόνιση Δεξιά: Απεικόνιση σε film

### 2.9.2 Τυπικές προβολές Μαστογραφίας

Οι εικόνες από τη μαστογραφία είναι συνήθως από δύο έως τέσσερις γωνίες (ή απόψεις)

Υπάρχει η "Cranial-Caudal" (CC) είναι μια λήψη από ψηλά. ενώ μια πλάγια ή γωνιακή όψη «mediolateral-πλάγια», ή (MLO). Η «μέσο-πλευρική όψη», (ML) είναι μια όψη από το κέντρο του θώρακα προς τα έξω, ενώ η «πλευρική-μεσαία» (ή LM) είναι μια όψη από την εξωτερική πλευρά του μαστού, προς τη μέση του στήθους.

Συνήθως, οι ακτινολόγοι προτιμούν την προβολή MLO ή «από την πλευρά-υπό γωνία», την όψη σε προβολή 90 μοιρών. Αυτό συμβαίνει επειδή η προβολή MLO επιτρέπει την απεικόνιση περισσότερου από το στήθος στο άνω-εξωτερικό τεταρτημόριο, καθώς και στην μασχάλη.

Με την προβολή από πάνω προς τα κάτω ή Cranial Caudal (CC), απεικονίζεται ολόκληρο το στήθος. Ο λιπώδης ιστός που βρίσκεται πλησιέστερα στον μυ του μαστού πρέπει να εμφανίζεται ως μια σκοτεινή λωρίδα στην ακτινογραφία. Επίσης, η προβολή CC τείνει επίσης να απεικονίζει σαφώς τη θηλή.[17]

Η προβολή Mediolateral (ή ML) είναι πολύ σημαντική επειδή η πλευρική πλευρά του μαστού είναι πιθανώς το πιο κοινό μέρος για την εμφάνιση παθολογικών αλλαγών. Η θέα από το κέντρο του στήθους, προς τα έξω προς τα πλάγια, (ML) δίνει την καλύτερη θέα της πλευρικής πλευράς του μαστού. Σε αυτήν την άποψη ο μυς του θώρακα

(θωρακικός) εμφανίζεται στη μαστογραφία ως στενή λωρίδα φωτός στο ήμισυ περίπου της εικόνας. Ακόμη, η απεικόνιση της θηλής είναι επίσης σαφής στο προφίλ.

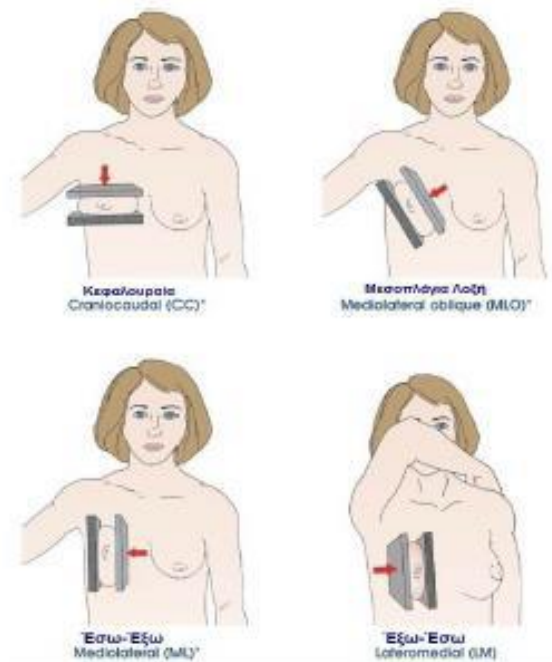
Τέλος, με την προβολή Latero-Medial (LM) το στήθος ακτινογραφείται από την πλευρά προς τη μέση και αυτό δίνει την καλύτερη άποψη της μεσαίας πλευράς του μέσου του στήθους.[17]

## 2.10 Ψηφιακή Τομοσύνθεση μαστού

Η ψηφιακή τομοσύνθεση είναι μια μέθοδος ανακατασκευής τομογραφικών εικόνων περιορισμένου τόξου (limited angle reconstruction of tomographic images), οι οποίες παράγονται σε μεταβλητά ύψη, στη βάση μιας σειράς από γωνιακές προβολές, (Εικόνα 9).[12] Ιδιαίτερα στην απεικόνιση του μαστού, οι μελέτες οι οποίες συμπεριλαμβάνουν και κλινικές περιπτώσεις, έχουν αποδείξει ότι η ψηφιακή τομοσύνθεση μπορεί να παρέχει πολύ ανώτερη ποιότητα εικόνας σε σύγκριση με την αντίστοιχη των κλασσικών μαστογραφιών εικόνων. Στη μέθοδο αυτή, ο απεικονιζόμενος όγκος (volume) ανακατασκευάζεται από δισδιάστατες προβολές έτσι ώστε να παρέχει τρισδιάστατη πληροφορία σχετικά με τη δομή της ανατομίας του εξεταζόμενου. Οι βασικότεροι αλγόριθμοι που χρησιμοποιούνται για την ανακατασκευή, είναι η οπισθοπροβολή και η μέθοδος μετατόπισης και πρόσθεσης. Ωστόσο, οι αλγόριθμοι αυτοί οδηγούν στην εμφάνιση σημαντικής ασάφειας στην εικόνα, η οποία προκαλείται από άλλες δομές που δεν βρίσκονται στο επίπεδο ενδιαφέροντος. Το γεγονός αυτό οδηγεί σε κακή ανιχνευσιμότητα των αντικειμένων που βρίσκονται μέσα στο επίπεδο που εστιάζουμε.[13]



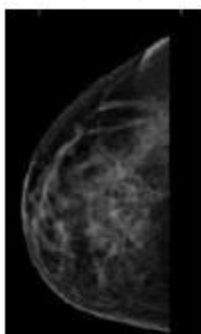
## Τυπικές προβολές



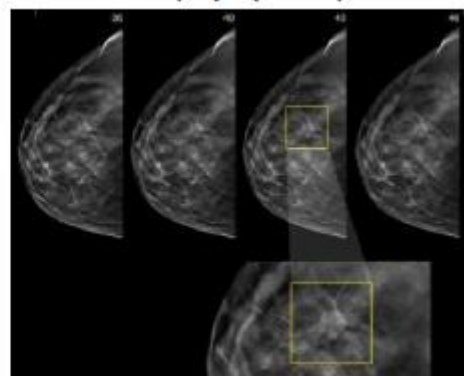
Εικόνα 16: Τυπικές γωνιακές προβολές κατά την ψηφιακή μαστογραφία.

Η ψηφιακή τομοσύνθεση μαστού αποτελεί μια τεχνική απεικόνισης του γυναικείου μαστού η οποία προσφέρεται είτε σαν ένα ξεχωριστό σύστημα, είτε ως επιλογή στα ήδη υπάρχοντα δισδιάστατα συστήματα μαστογραφίας.[14] Κατά τη διάρκεια της απόκτησης των εικόνων, η ακτινογραφική λυχνία κινείται διαγράφοντας μια μικρή γωνία γύρω από τον μαστό και με τον τρόπο αυτό λαμβάνεται μια σειρά εικόνων. Η κίνηση της λυχνίας μπορεί να είναι συνεχής είτε σε μορφή step-and shoot.

Μια προβολή



Πολλές προβολές



Εικόνα 17: Ψηφιακή τομοσύνθεση μίας προβολής και πολλών προβολών.



### 2.11 Μαστογραφία: Κίνδυνοι και εσφαλμένες αντιλήψεις

Η μαστογραφία κατηγορείται σήμερα υπερβολικά για την ακτινοβολία της και ως εκ τούτου αποτρέπει μέρος των γυναικών στη διενέργειά της. Όμως η ακτινοβολία είναι μέγεθος απόλυτα μετρήσιμο και οι μετρήσεις δεν επιβεβαιώνουν τις κατηγορίες. Το πρόβλημα της μαστογραφίας δεν είναι η ακτινοβολία αλλά η μεγάλη πιθανότητα σφάλματος στους πυκνούς μαστούς. Η μαστογραφία παραμένει πάντα μια πολύτιμη εξέταση. [15,16]

Αξίζει να σημειωθεί ότι μετρώντας τα επίπεδα ιονίζουσας ακτινοβολίας σε μονάδες mGy κατά τη διάρκεια διενέργειας ψηφιακής μαστογραφίας, η ακτινοβολία δεν ξεπερνά τα 0.8 mGy, τη στιγμή που μια απλή ακτινογραφία θώρακος μετράται σε 3 mGy, ενώ σε μία αξονική τομογραφία κυμαίνεται από 10- 40 mGy (ανάλογα με το είδος της εξέτασης). Η θανατηφόρος δόση ιονίζουσας ακτινοβολίας υπολογίζεται στα 8.000 mGy.

Λαμβάνοντας υπόψιν, αφενός τα αποτελέσματα στατιστικών ερευνών, την έλλειψη βιβλιογραφίας περί άμεσης σύνδεσης της εμφάνισης καρκίνου του μαστού λόγω ακτινοβολίας στη διάρκεια μαστογραφίας και αφετέρου το ανεκτίμητο κέρδος της άμεσης διάγνωσης ενός πρώιμου καρκίνου, οδηγούμαστε στο συμπέρασμα της χρήσης της μεθόδου αυτής ως χρήσιμο μέρος μιας συνολικότερης και πιο ορθής διαδικασίας αντιμετώπισης της νόσου.

## Βιβλιογραφία 2<sup>ου</sup> Κεφαλαίου

- [1]. M. Jochelson, 'Advanced imaging techniques for the detection of breast cancer', Am Soc Clin Oncol Educ Book, σσ 65–69, 2012.
- [2]. R. S. of N. A. (RSNA) and A. C. of Radiology (ACR), 'Mammography (Mammogram)'. Διαθέσιμο στο: <https://www.radiologyinfo.org/en/info.cfm?pg=mammo>
- [3]. Ε.Γεωργίου,Ε.Γιακουμάκης,Π.Δημητρίου,Π.Καραίσκος,Σ.Κόττου,Α.Λουίζη, κ.α (2014)“Ιατρική Φυσική” “Διαγνωστικές και θεραπευτικές εφαρμογές των ακτινοβολιών”.
- [4]. Κ.Ψαρρακός, Α. Μολυβδά-Αθανασουπούλου, Α. Γ. Ψαρρακού, Α. Σιούτας, (2012), 'Ιατρική Φυσική', Θεσσαλονίκη, University Studio Press.
- [5]. Μέθοδοι καταγραφής της ακτινοβολίας X και γ στα συστήματα ιατρικής απεικόνισης. Εισηγητής: Ι. Κανδαράκης, καθηγητής ΤΕΙ Αθήνας, Τμήμα Τεχνολογίας Ιατρικών Οργάνων.
- [6]. Ευάγγελος Παντελής, Επ. Καθ. Ιατρικής Φυσικής, Ιατρική Σχολή Αθηνών, Ιατρική Φυσική – Διαγνωστικές και θεραπευτικές εφαρμογές ακτινοβολιών, Κεφάλαιο 8.
- [7]. J. M. Lewin, R. E. Hendrick, C. J. D’Orsi, P. K. Isaacs, L. J. Moss, A. Karellas, G. A. Sisney, C. C. Kuni, and Gary R. Cutter, et al. (2001), “Comparison of Full-Field Digital Mammography with Screen-Film Mammography for Cancer Detection: Results of 4,945 Paired Examinations”, Radiology, Vol. 218, Issue 3, 873-880.
- [8]. E.D. Pisano, M.J. Yaffe, B.M. Hemminger et al. (2000), “Current status of full-field digital mammography”, Acad Radiol, Vol. 7, Number 4, 266-80.
- [9]. E.D. Pisano, M.J. Yaffe (2005), “Digital mammography Radiology”, Vol. 234, Issue 2, 353-362.
- [10]. J. M. Lewin, C. J. D’Orsi, R. E. Hendrick, L. J. Moss, P. K. Isaacs, A. Karellas and G. R. Cutter et al. (2002), “Clinical Comparison of Full-Field Digital Mammography and Screen-Film Mammography for Detection of Breast Cancer”, American Journal of Roentgenology, Vol. 179, Issue 3, 671-677.
- [11]. S.A. Feig, M.J. Yaffe (1995), “Digital mammography, computer- aided diagnosis, and telemammography”, Radiol Clin North Am, Vol. 33, Issue 6, 1205–1230.
- [12]. Τσαβαρής ΝΒ. Βιολογικές θεραπείες-Γαστρεντερολογικό Σύστημα. 2ο Διαιτητικό Αντικαρκινικό Συνέδριο (2005) 65-68.
- [13]. K. Bliznakova, Z. Bliznakov, I. Buliev (2012), “Comparison of algorithms for out-of-plane artifacts removal in digital tomosynthesis reconstructions”, Computer Methods and Programs in Biomedicine, Vol. 107, Issue 1, 75-83.
- [14]. ]‘Καρκίνος του μαστού: Στατιστικά στοιχεία’, Άλμα Ζωής - Πανελλήνιος Σύλλογος Γυναικών με Καρκίνο Μαστού.
- [15]. E. C. Lin, ‘Radiation Risk from Medical Imaging’, Mayo Clin Proc, τ. 85, τχ. 12, σσ 1142–1146, Δεκεμβρίου 2010.

[16]. 'Mammograms: Benefits and Risks', Breastcancer.org. Διαθέσιμο στο: [https://www.breastcancer.org/symptoms/testing/types/mammograms/benefits\\_risks](https://www.breastcancer.org/symptoms/testing/types/mammograms/benefits_risks).

[17]. [www.breast-cancer.ca](http://www.breast-cancer.ca)

Page Intentionally Left Blank

## Κεφάλαιο 3ο

### Ψηφιακή Επεξεργασία και Ανάλυση της Μαστογραφίας

Η όραση είναι η πιο δυναμική μας αίσθηση και μας δίνει ένα μεγάλο πλήθος πληροφοριών για το περιβάλλον. Ο τεράστιος όγκος δεδομένων και η ανάγκη για την επεξεργασία τους οδήγησε επιστήμονες και τεχνικούς στην αναζήτηση μέθοδων ψηφιακής αποθήκευσης και επεξεργασίας με την χρήση ηλεκτρονικών υπολογιστών προσπάθεια αυτή οδήγησε σε ένα καινούργιο κλάδο πληροφορικής που ονομάζεται ψηφιακή επεξεργασία και ανάλυση εικόνας.

Η ψηφιακή ανάλυση εικόνας καταπιάνεται με την εγγραφή και αναγνώριση του περιεχομένου της εικόνας. Συνήθως η διαδικασία είναι συμβολική. Αρά, η είσοδος είναι ψηφιακή εικόνα και η έξοδος είναι συμβολική περιγραφή. Ουσιαστικά προσπαθεί να μιμηθεί την ανθρώπινη όραση και εξαιτίας αυτού συχνά γίνεται χρήση του ορού τεχνητή όραση (computer vision). Αλλά η τεχνητή όραση, Παρόλου που έχει εξελιχτεί και χρησιμοποιείται σε πληθώρα εφαρμογών απέχει πολύ από το να φτάσει στην ανθρώπινη όραση. Τα μαθηματικά μοντέλα που χρησιμοποιούνται έχουν περιορισμούς και είναι δύσκολη η προσομοίωση σε ηλεκτρονικό υπολογιστή.

Η ψηφιακή επεξεργασία εικόνας καταπιάνεται με επεξεργασία και την ψηφιακή καταγραφή εικόνων. Αρά, τόσο η είσοδος όσο και η έξοδος είναι ψηφιακές εικόνες. Το προς αντικείμενο επεξεργασία μπορεί να είναι το φιλτράρισμα θορύβου καταγραφής η μετάδοσης, η βελτίωση της ποιότητας, η συμπίεση του όγκου των πληροφοριών η και η αποθήκευση εικόνας και η ψηφιακή μετάδοσης της.

#### 3.1 Ψηφιακή Επεξεργασία μαστογραφίας

Η δημιουργία ενός συστήματος εντοπισμού νεοπλασιών είναι άρρηκτα συνδεδεμένο με την ψηφιακή επεξεργασία του μαστογραφήματος. Ο λόγος ύπαρξης είναι ότι χρειάζεται η αξιόπιστη απεικόνιση της ανατομίας του στήθους. οι τεχνικές για την επεξεργασία της μαστογραφίας κατηγοριοποιούνται σε δυο ομάδες:

Τεχνικές βελτίωσης και αποκατάστασης της μαστογραφίας. Υλοποιείται βελτίωση της εικόνας δίνοντας έμφαση τονίζοντας χαρακτηριστικά όπως περίγραμμα, αντίθεση η εξομαλύνοντας κάποια χαρακτηριστικά. Αυτές οι τεχνικές συνήθως χρησιμοποιούνται κυρίως στο στάδιο της προ επεξεργασίας της ψηφιακής μαστογραφίας σε συστήματα ανίχνευσης. Ακόμη, χρησιμοποιούνται για την βελτίωση της εικόνας έτσι ώστε να υπάρξει καλύτερη ερμηνεία της εξέτασης από τον ακτινολόγο

Τεχνικές κατάτμησης της μαστογραφίας, για την εξαγωγή της περιοχής ενδιαφέροντος (ROI). Αυτές αποτελούν το κύριο στάδιο επεξεργασίας καθώς εκτιμούν τις περιοχές ενδιαφέροντος της μαστογραφίας.

### 3.2 Τεχνικές βελτίωσης και αποκατάστασης μαστογραφίας

Οι τεχνικές αποκατάστασης της μαστογραφίας χρησιμοποιείται σε για την αφαίρεση θαμπώματος και θορύβου που συναντάται στη καταγραφή. Αυτές οι μέθοδοι έχουν την δυνατότητα να αφαιρέσουν τον θόρυβο με αποτέλεσμα να αυξάνεται η ποιότητας της εικόνας προσφέροντας καλύτερη αντίληψη στους ακτινολόγους. Οι μέθοδοι που συνήθως χρησιμοποιούνται είναι φίλτρα εξομάλυνσης, ανάλυση και επεξεργασία ιστογράμματος με την μέθοδο της ισοστάθμισης , με χρήση μεσαίου ζωνοδιαβατού καθώς και ψηφιακού φίλτρου

Τα συστήματα ανίχνευσης υπόπτων περιοχών που έχουν μικροαποτιτανώσεις προκειμένου να ελαττώσουν την ποσότητα της ασυσχέτιστης πληροφορίας στους ιστού χρησιμοποιούνται wavelet και ζωνοδιαβατά φίλτρα.

Επιπλέον υπάρχουν πολλές τεχνικές που βελτιώνουν την ποιότητα της εικόνας τονίζοντας χαρακτηριστικά όπως αντίθεση με χρήση υπεραπενών φίλτρων. Με την χρήση των φίλτρων δίνεται έμφαση στα χαρακτηριστικά υψηλής συχνότητας στην μαστογραφία. Για τον σκοπό αυτό χρησιμοποιούνται επίσης τεχνικές ανίχνευσης ακμών, που βρίσκουν εφαρμογή στην εύρεση περιγραμμάτων.

#### 3.2.1 Γραμμικά Φίλτρα Εξομάλυνσης

Τα φίλτρα χρησιμοποιούνται για την μείωση του θορύβου και για το θάμπωμα. Το θάμπωμα προκύπτει σε στάδια επεξεργασίας όπως η κάλυψη ασυνεπειών σε καμπύλες ή γραμμές

#### 3.2.2 Κατωδιαβατά φίλτρα χώρου

Η διαδικασία κατά την οποία γίνεται χρήση της μάσκας χώρου για την επεξεργασία εικόνας ονομάζεται φιλτράρισμα διαστήματος. Τα φίλτρα εξομάλυνσης είναι κατωδιαβατά φίλτρα, και ονομάζονται έτσι επειδή αποκόπτουν τις υψηλές συχνότητες αφήνοντας ανεπηρέαστες τις χαμηλές συχνότητες. Οι συνιστώσες των υψηλών συχνοτήτων αποτελούν χαρακτηριστικά των ακμών και των λεπτομερειών και έχουν ως συνέπεια το θάμπωμα όταν εφαρμόζεται ένα κατωδιαβατό φίλτρο.

Επιπλέον, τα γραμμικά φίλτρα ανεξάρτητα από το είδος, η εικόνα της εξόδου προκύπτει από την άθροιση γινομένων μεταξύ των συντελεστών της μάσκας με αντίστοιχα στοιχεία, στην καθορισμένη περιοχή της εικόνας.

#### 3.2.3 Φίλτρο Μέσης τιμής(mean filter)

Ένα από τα πλέον γραμμικά φίλτρα είναι το φίλτρο μέσης τιμής. Χρησιμοποιείται για την εξομάλυνση εικόνων και την μείωση του θορύβου σε αυτές. Η λειτουργία φίλτρου μέσης τιμής συνίσταται στην αντικαταστατή της φωτεινότητας σε κάθε εικονοστοιχείο, με την μέση φωτεινότητα σε μια γειτονιά του. Αυτό έχει ως αποτέλεσμα τη μείωση της μεταβλητότητας τοπικά σε κάθε εικονοστοιχείων και συνεπώς στην θαμπώσει της εικόνας.[1]

#### 3.2.4 Μεσαίο φίλτρο(median filter)

Το φιλτράρισμα με ένα φίλτρο μαίας τιμής (median filter) είναι μια μη γραμμική τσέχικη που ανέπτυξε ο Tukey για την ελάττωση του θορύβου στις ψηφιακές εικόνες. Η τιμή median ενός συνόλου  $A$  είναι ίση με την μεσσία τιμή του συνόλου. Σε ένα σύνολο το οποίο έχει περιττό αριθμό στοιχείων , γίνεται η ταξινόμηση κατά αύξουσα

τιμή, οπότε το median του συνόλου ισούται με το μεσαίο στοιχείο αν υπάρχει περιττός αριθμός στοιχείων ή με την μέση τιμή των δυο γειτονικών στοιχείων στην περίπτωση του αρτίου αριθμού στοιχείων

Αυτή η μέθοδος εξομάλυνσης προκαλεί θάμπωμα στις λεπτομέρειες και τις ακμές της εικόνας. Στην περίπτωση που μας αφορά ο θόρυβος, παραπάνω από το θόρυβο τότε κάνουμε χρήση του μεσαίου φίλτρου. Το φίλτρο αυτό αντικαθιστά τις τιμές του γκρι με την μεσαία τιμή αυτού του εικονοστοιχείων. Η τεχνική είναι ιδιαίτερα αποτελεσματική όταν ο θόρυβος εμφανίζεται σαν αιχμές στο διάγραμμα του και το χαρακτηριστικό που μας αφορά είναι η οξύτητα των ακμών. Τέλος το φίλτρο πρέπει να ανέρθει ότι είναι μη γραμμικό.[2]

### 3.2.5 Ανωδιαβατό φίλτρο.

Τα φίλτρα εξομάλυνσης είναι ακόμη τα ανωδιαβατά φίλτρα, και ονομάζονται έτσι επειδή αφήνουν ανεπηρέαστες τις υψηλές συχνότητες και αποκόπτουν τις χαμηλές συχνότητες. Οι συνιστώσες των υψηλών συχνοτήτων αποτελούν χαρακτηριστικό των ακμών και των λεπτομερειών και έχουν ως συνέπεια το θάμπωμα όταν εφαρμόζεται ένα κατωδιαβατό φίλτρο. Η ανάδειξη και ο τονισμός της εικόνα μπορεί να γίνει με την χρήση υψιπερατών φίλτρων. Το υψιπερατό φιλτράρισμα μπορεί να γίνει με την χρήση μάσκας που αποτελείται από ένα μείγμα αρνητικών και θετικών συντελεστών.[7]

## 3.3 Το ιστόγραμμα μιας εικόνας

Το ιστόγραμμα μιας εικόνας είναι εύκολο να προσδιορισθεί. Εκφράζει την κατανομή των αποχρώσεων του γκρι στην εικόνα και στις περισσότερες περιπτώσεις καθορίζει απολυτά την εικόνα. Ένα ιστόγραμμα είναι ένα γράφημα που στο οριζόντιο άξονα έχει τις φωτεινότητες από 0-255 και στον κατακόρυφο άξονα το πλήθος των εικονοστοιχείων που έχουν κάθε φωτεινότητα. Στην περίπτωση που η μέγιστη τιμή του ιστογράμματος αντιστοιχεί με το ένα, τότε το κανονικοποιημένο ιστόγραμμα που προκύπτει μας δίνει ουσιαστικά την κατανομή πυκνότητας πιθανότητας των επίπεδων του γκρι στην εικόνα.

## 3.4 Επεξεργασία Ιστογράμματος

Η φωτεινότητα  $f(i,j)$  στην εικόνα είναι μια τυχαία μεταβλητή που περιγράφεται με την βοήθεια της συνάρτησης πυκνότητας πιθανότητας  $p_f(f)$ . Η συνάρτηση πυκνότητας πιθανότητας δεν δίνεται έτοιμη αλλά πρέπει να υπολογισθεί με τη βοήθεια της στατιστικής πυκνότητας πιθανότητα που ονομάζεται ιστόγραμμα. Το μέγεθος αυτό, παρέχει πληροφορία για το συνολικό περιεχόμενο της εικόνας. Μια ψηφιακή εικόνα έχει  $M$  διακριτά επίπεδα του γκρι και ο αριθμός των εικονοστοιχείων είναι  $k=0,1 \dots M-1$ , οπότε το ιστόγραμμα ορίζεται από την παρακάτω σχέση:

$$P_f(f_k) = \frac{nk}{n}, k=0,1 \dots M-1$$

Οπού  $n$  είναι ο συνολικός αριθμός των εικονοστοιχείων.

Το ιστόγραμμα αποτελεί σημαντική πηγή πληροφορίας για την φύση της εικόνας. Μια εικόνα που έχει φωτεινά χαρακτηριστικά οι τιμές είναι συγκεντρωμένες στις υψηλές τιμές φωτεινότητας, ενώ μια εικόνα που έχει σκοτεινά χαρακτηριστικά οι τιμές είναι συγκεντρωμένες στις χαμηλές.[7]

Επιπλέον, για παράδειγμα ένα απλωμένο ιστόγραμμα σημαίνει ότι η εικόνα είναι πλούσια σε αντιθέσεις, ένα στενό ιστόγραμμα σημαίνει μια εικόνα χωρίς ιδιαίτερες αντιθέσεις. Ακόμη, ένα ιστόγραμμα με δυο μεγάλες συγκεντρώσεις σημαίνει ότι η εικόνα έχει δυο αντικείμενα με διαφορετικές τιμές φωτεινότητας.

### 3.5 Τροποποίηση ιστογράμματος

Το ιστόγραμμα μπορεί να συμπιεστεί, να αποσυμπιεστεί καθώς και ολισθήσει με την χρήση μιας προσαρμοστικής συνάρτησης. Η συρρίκνωση και η εξάπλωση περιλαμβάνονται στην κλιμάκωση του ιστογράμματος.

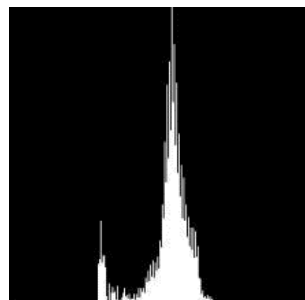
#### 3.5.1 Ανάπτυξη ιστογράμματος(stretching)

Η συνάρτηση που προσαρμόζεται προκειμένου να επιτευχθεί το άνοιγμα του ιστογράμματος είναι:

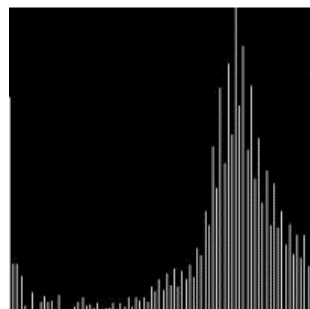
$$Stretch(I(r, c)) = \left[ \frac{I(r, c) - I(r, c)_{MIN}}{I(r, c)_{MAX} - I(r, c)_{MIN}} \right] [MAX - MIN] + MIN$$

Οπού τα  $r, c$  αντιστοιχούν στην γραμμή και την στήλη,  $I(r, c)_{MAX}$  είναι η μέγιστη τιμή της που συναντάμε στην εικόνα,  $I(r, c)_{MIN}$  είναι η μικρότερη τιμή που υπάρχει,  $MIN$  και  $MAX$  συμβολίζουν την ελάχιστη και την μέγιστη τιμή της φωτεινότητας οπού είναι από 0 μέχρι 255.

Ο μετασχηματισμός αυτός θα ανοίξει το ιστόγραμμα και θα απλωθεί σε όλες τις κλίμακες φωτεινότητας με αποτέλεσμα να αυξηθεί η αντίθεση της εικόνας



α)



β)

Εικόνα 18: α) Αρχικό ιστόγραμμα, β) Ανάπτυξη ιστογράμματος.



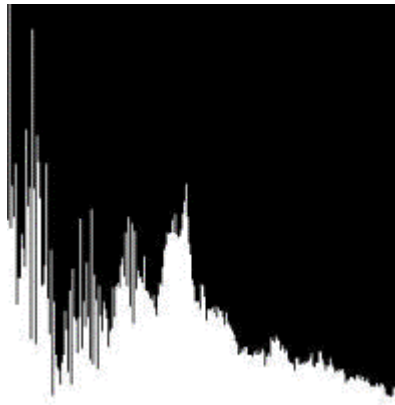
### 3.5.2 Συμπίεση Ιστογράμματος

Η συμπίεση του ιστογράμματος λειτουργεί με τον αντίθετο τρόπο με αποτέλεσμα να μειωθεί η φωτεινότητα αφού η εικόνα συμπιέζεται στις τιμές φωτεινότητας του γκρι. Η προσαρμοστική συνάρτηση δίνεται από την παρακάτω σχέση:

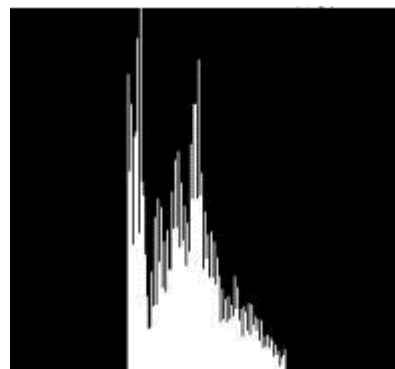
$$Shrink(I(r, c)) = \left[ \frac{ShrinkMAX - ShrinkMIN}{I(r, c)MAX - I(r, c)MIN} \right] [I(r, c) - I(r, c)MIN] + ShrinkMIN$$

Οπού: ShrinkMAX και ShrinkMIN το μέγιστο και ελάχιστο οροί που Έχουμε σκοπό να συμπίεσουμε το ιστόγραμμα.

Παράδειγμα



α)



β)

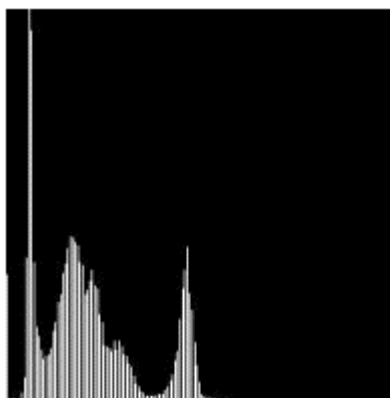
Εικόνα 19: α) Αρχικό Ιστόγραμμα, β) Συμπιεσμένο Ιστόγραμμα.

### 3.5.3 Ολίσθηση ιστογράμματος

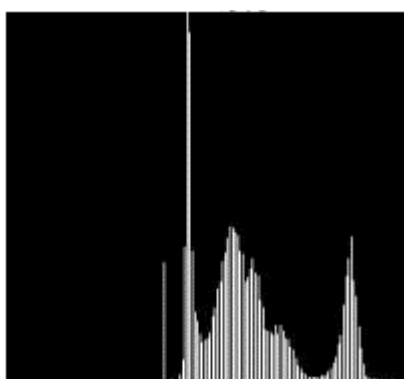
Η ολίσθηση ιστογράμματος έχει σκοπό να γίνει η εικόνα είτε πιο φωτεινή είτε πιο φωτεινή, αλλά με διατήρηση των σχετικών επίπεδων του γκρι. Αυτό μπορεί να επιτευχθεί με την απλή αφαίρεση ή πρόσθεση συγκεκριμένου αριθμού σε όλες τις φωτεινότητες των εικονοστοιχείων. Η προσαρμοστική συνάρτηση δίνεται από την παρακάτω εξίσωση:

$$Sl(I(r, c)) = I(r, c) + OFFSET$$

Οπού OFFSET είναι η ποσότητα όπου θα ολισθήσουμε το ιστόγραμμα.



α)



β)

Εικόνα 20 :α)αρχικό ιστόγραμμα β) Ολισθημένο ιστόγραμμα[10]

### 3.6 Ισοστάθμιση Ιστογράμματος

Η ισοστάθμιση αποτελεί μια μέθοδο βελτίωσης της ποιότητας της εικόνας που είναι χαμηλή. Η τεχνική της εξισορρόπησης ιστογράμματος μετασχηματίζει τις γκρι φωτεινότητες μιας εικόνας έτσι ώστε αυτές να κατανέμονται ισόμορφα σε όλη την κλίμακα φωτεινότητων. Η εικόνα που προκύπτει με τον τρόπο αυτό είναι αυξημένης αντίθεσης σε σχέση με την αρχική.[RIWI96]

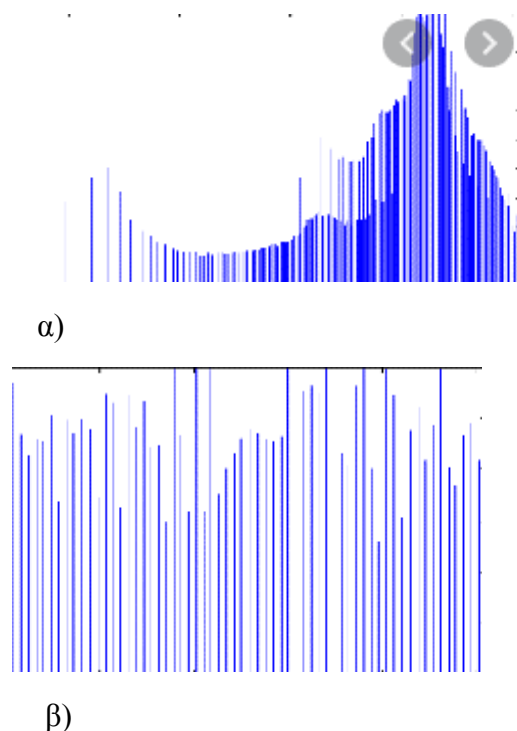
Η ισοστάθμισή ιστογράμματος επιτυγχάνεται με τα τέσσερα παρακάτω στάδια:

- Υπολογίζονται τα αθροίσματα της αρχικής εικόνας.
- Κανονικοποίηση των τιμών που προκύπτουν από το πρώτο βήμα και διαίρεση των αθροισμάτων με τον αριθμό των συνολικών εικονοστοιχείων
- Πολλαπλασιασμός των τιμών που προκύπτουν από το δεύτερο στάδιο με το την μέγιστη τιμή φωτεινότητας και έπεται στρογγυλοποίηση των αποτελεσμάτων
- Στο τέταρτο στάδιο γίνεται αντιστοίχιση μια προς μια των αρχικών τιμών των επίπεδων φωτεινότητας που προέκυψαν από το τρίτο στάδιο.[3]

Επιπλέον, επισημαίνεται ότι η ισοστάθμιση δεν θα δώσει ένα τελείως επίπεδο ιστόγραμμα αλλά θα το κάνει όσο πιο επίπεδο γίνεται. Το τελείως επίπεδο ιστόγραμμα θα χρειαστεί να χωρίσουμε τα εικονοστοιχεία σε περισσότερα από ένα επίπεδα

φωτεινότητας το οποίο θα δημιουργήσει πρόβλημα λόγω της πολυπλοκότητας που προκύπτει.

Παράδειγμα:



Εικόνα 21: α) Αρχικό ιστόγραμμα β) ισοσταθμισμένο ιστόγραμμα[11]

### 3.7 Τεχνικές κατάτμησης της μαστογραφίας

Η μορφή που έχει ένα αντικείμενο μπορεί να περιγράψει με την βοήθεια του περιγράμματος ή με την περιοχή που έχει καταλάβει. Η περιγραφή αντικειμένου που στηρίζεται στην περιοχή που έχει καταλάβει χρειάζεται κατάτμηση εικόνας σε ομοιογενείς περιοχές. Συνεπώς, η κατάτμηση και η ανίχνευση ακμών είναι απαραίτητη στην ανάλυση της εικόνας. Οι περιοχές της εικόνας θα πρέπει να έχουν παρόμοια χαρακτηριστικά όπως η ένταση, τα οποία διαφέρουν σε κάθε περιοχή της εικόνας. Ακόμη, διανύσματα σχηματίζονται από τα χαρακτηριστικά τα οποία χρησιμοποιούνται για να επιτευχθεί η διάκριση από μια περιοχή με μια άλλη. Οι παράμετροι στην διαδικασία χωρισμού της εικόνας χρησιμοποιούνται από κανόνες για να γίνει σωστά.

Οι μέθοδοι που υπάρχουν για την κατάτμηση των ψηφιακών μαστογραφιών συνηθώς είναι οι καθολικές που στηρίζονται στις πληροφορίες που λαμβάνονται καθολικά όπως για παράδειγμα από την χρήση ενός ιστογράμματος. Επίσης χρησιμοποιούνται οι τεχνικές συνένωσης αύξησης και διαίρεσης βασίζονται σε στην γεωμετρική ανισότητα και στις εννοείς της ομοιογένειας. Οι περιοχές αν είναι όμοιες ή παροιμίες μεταξύ τους τότε μπορούν να ενωθούν. Στην εικόνα μια περιοχή που δεν είναι ομοιογενής μπορεί να χωριστεί σε μικρότερες υποπεριοχές καθώς και μια περιοχή μπορεί να αυξηθεί προσθέτοντας περιοχές παραμένοντας όμοια.[4]

### 3.7.1 Κατάτμηση εικόνας με κατώφλι

Το απλό πρόβλημα κατάτμησης είναι εκείνο όπου στην εικόνα υπάρχει ένα αντικείμενο με ομοιογενή ένταση και ένα φόντο με διαφορετικό επίπεδο φωτεινότητας, και χωρίζεται η εικόνα σε δυο περιοχές με κατωφλίωση:

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

Όπου  $f(x, y)$ : η φωτεινότητα του εικονοστοιχείων με συντεταγμένες  $x, y$  αντίστοιχα

Το κατώφλι επιλέγεται χρησιμοποιώντας το ιστόγραμμα της εικόνας. Αν η εικόνα περιέχει ένα αντικείμενο και φόντο με παρόμοια ένταση τότε το ιστόγραμμα θα έχει δυο λοβούς. Το κατώφλι συνήθως επιλέγεται το τοπικό ελάχιστο του ιστογράμματος που εντοπίζεται αναμεσα σε δυο λοβούς. Παράλληλα, συνίσταται η εξομάλυνση του ιστογράμματος για να εξαλειφθεί όσο το δυνατόν περισσότερο ο θόρυβος ειδικά αν το ιστόγραμμα θα είναι θορυβώδες με αποτέλεσμα να είναι δύσκολο να υπολογιστεί το τοπικό μέγιστο. [KIT186]

Επίσης, σε κάποιες εφαρμογές η ένταση του αντικειμένου δεν μεταβάλλεται γρηγορά. Σε αυτήν την περίπτωση μπορεί να εφαρμοσθεί κατώφλι που να μεταβάλλεται στον χώρο. Η εικόνα χωρίζεται σε τετράγωνα και αν το τοπικό ιστόγραμμα δεν είναι δυο λοβών, το κατώφλι θα υπολογισθεί με την βοήθεια των κατωφλίων στις γειτονικές περιοχές της εικόνας. Όταν βρεθεί το τοπικό κατώφλι εφαρμόζεται η παραπάνω σχέση (Σχέση 3.). Τέλος, η μέθοδος της κατωφλίωσης εξασφαλίζει την ομοιογένεια των περιοχών αλλά όχι την συνεκτικότητά τους, εξαιτίας του γεγονότος ότι η τεχνική εφαρμόζεται στο επίπεδο των εικονοστοιχείων δίχως την πληροφορία γειτονικών περιοχών. [KOHL81]

### 3.8 Τεχνικές κατάτμησης με συνένωση, αύξηση και διαίρεση περιοχών

Κομβικό ρολό στην κατάτμηση εικόνας έχει η γεωμετρική αμεσότητα. Τα εικονοστοιχεία που βρίσκονται στην ίδια γειτόνευα συνήθως έχουν κοινές στατιστικές ιδιότητες και βρίσκονται στην ίδια περιοχή της εικόνας. Επομένως, οι αλγόριθμοι κατάτμησης θα πρέπει να λαβαίνουν υπόψη την ομογένεια και την γεωμετρική αμεσότητα για να παράγουν συνεκτικές περιοχές εικόνας.

Στην κατάτμηση εικόνας, μια απλή μέθοδος είναι να χρησιμοποιήσουμε αντιπροσωπευτικά εικονοστοιχεία (σπόρους) προκειμένου να αυξηθούν μέχρι να καλυφθεί όλη η εικόνα. Για κάθε περιοχή επιλέγεται ένας σπόρος για κάθε περιοχή ξεχωριστά. Επιπλέον, προκειμένου να υλοποιηθεί η αύξηση των περιοχών θα πρέπει να φτιαχτεί ένα μηχανισμός αύξησης των περιοχών και ένα μηχανισμός που θα ελέγχει την ομοιογένεια μετά από κάθε βήμα. Ο μηχανισμός αύξησης είναι σχετικά απλός καθώς σε κάθε στάδιο και για κάθε περιοχή γίνεται έλεγχος αν υπάρχουν εικονοστοιχεία στην πιο κοντινή γειτόνευα του κάθε εικονοστοιχείων που δεν έχει καταταχθεί. Αν γίνει η προσθήκη εικονοστοιχείων και δεν αλλάξει η ομογένεια τότε μπορεί να γίνει η προσθήκη. Συγκεκριμένα, είναι εφικτό να γίνεται έλεγχος αν το υποψήφιο εικονοστοιχείων έχει ένταση  $g$  κοντά στην μέση τιμή  $m$  της έντασης της περιοχής, δηλαδή το  $|g - m|$  πρέπει να είναι μικρότερο από ένα κατώφλι. [9]

Τώρα, σχετικά με την διαδικασία συνενώνεις, δυο προσκείμενες περιοχές με παρόμοιες στατιστικές ιδιότητες μπορούν να συνενωθούν. Ο αριθμητικός μέσος  $m$  και η τυπική απόκλιση  $\sigma$  της φωτεινότητας των εικονοστοιχείων μιας περιοχής μπορούν να χρησιμοποιηθούν για να κριθεί αν αυτή η περιοχή μπορεί να συνενωθεί με μια άλλη. Ειδικά, αν οι αριθμητικοί μεσοί είναι  $m_1, m_2$  δυο περιοχών είναι κοντά ο ένας στον άλλον:

$$|m_1 - m_2| < k\sigma, i = 1, 2$$

οι περιοχές επιτρέπεται να συνενωθούν συνένωση περιοχών μπορεί να ενσωματωθεί στον μηχανισμό ανάπτυξης με στόχο την δημιουργία ενός αλγορίθμου κατάτμησης με συνδυασμό συνένωσης και αύξησης.

Η τεχνική της διαίρεσης είναι αντίθετη από εκείνη της συνένωσης περιοχών, ξεκινώντας από την υπόθεση ότι όλη η εικόνα είναι ομοιογενής. Αν η υπόθεση δεν υφίσταται τότε η εικόνα χωρίζεται σε τέσσερις περιοχές. Η διαδικασία αυτή επαναλαμβάνεται συνέχεια μέχρι να επιτευχθεί ομοιογενείς περιοχές στην εικόνα. Οι μέθοδοι διαρκέσεις δημιουργούν ένα σημαντικό πρόβλημα καθώς μπορεί να φτιαχτούν περιοχές που είναι προσκείμενες και ομοιογενείς αλλά να μην συνενώνεται, με αποτέλεσμα να φτιάχνεται ο αλγόριθμος που λέγεται αλγόριθμος διαίρεσης και συνένωσης.[9]

Ο αλγόριθμος αυτός είναι επαναληπτικός, οπότε σε κάθε επανάληψη διαιρεί και συνενώνει περιοχές. Αν μια περιοχή είναι ανομοιογενής τότε αρχικά χωρίζεται σε τέσσερις υποπεριοχές, έπειτα αν δυο προσκείμενες υποπεριοχές είναι όμοιες συνενώνονται. Ο αλγόριθμος σταματά όταν δεν είναι δυνατόν να γίνει άλλη διαίρεσή. Ο συγκεκριμένος αλγόριθμος παράγει πιο συμπαγείς περιοχές συγκριτικά με τον αλγόριθμο διαίρεσης. Τέλος, μια εκδοχή του αλγορίθμου διαίρεσης και συνένωσης είναι ο αλγόριθμος διαίρεσης και να αναβληθεί μέχρις ότου να μην είναι δυνατόν η διαίρεση και οι προσκείμενες περιοχές συνενώνονται στο τελευταίο βήμα της συνένωσης.

### 3.9 Ψηφιακή ανάλυση μαστογραφίας

Η ψηφιακή ανάλυση εικόνας έχει κύριο σκοπό την περιγραφή και αναγνώριση του περιεχομένου που είναι συνήθως συμβολικό και πολύ σημαντικό όταν έχουν προσδιοριστεί οι ύποπτες περιοχές με την βοήθεια της ψηφιακής επεξεργασίας εικόνας. Κομβικό ρολό έχει στην ανάλυση, στην διαδικασία ανίχνευσης και στην διάγνωσης υπόπτων περιοχών είναι η εξαγωγή κατάλληλων παραμέτρων για να είναι εφικτή η διάκριση μεταξύ υγιής και μη υγιή φυσιολογική περιοχή. Για την καλύτερη περιγραφή μια περίπτωσης καρκίνου ομαδοποιούνται σε δυο κατηγορίες:

- Παράμετροι βασιζόμενοι στην περιγραφή μιας υφής
- Γεωμετρικοί και μορφολογικοί παράμετροι.

Περιγραφή μιας εικόνας

Η υφή είναι ένα φαινόμενο που είναι διαδεδομένο, εύκολο να αναγνωριστεί από τον άνθρωπο αλλά δύσκολο να καθοριστεί. Η κλίμακα με την οποία παρατηρούμε μια εικόνα καθορίζει το αν απέχουμε κίονα υφής η όχι. Για παράδειγμα, ένα φύλλο που καταλαμβάνει το μεγαλύτερο μέρος μιας εικόνας είναι ένα αντικείμενο, αλλά το

φύλλωμα ενός δέντρου αν παρατηρηθεί από μακριά δημιουργεί μια υφή. Επίσης παρατηρώντας την εικόνα μιας σκέπης με κεραμιδιά από μεγάλη απόσταση παρατηρείται η υφή που δημιουργούν τα κεραμιδιά αλλά τα σχήματα των κεραμιδιών δεν διακρίνονται. όταν η ίδια σκηνή παρατηρηθεί από κοντά, έτσι ώστε μόνο μερικά κεραμιδιά να είναι στο πεδίο όρασης, αρχίζει να διακρίνεται το σχήμα πλέξης το οποίο προκύπτει από την θέση των λεπτομερειών σχεδίων τα οποία στοιχειοθετούν κάθε κεραμιδί.

Ο ορισμός της υφής δεν είναι πλήρως καθορισμένος, όμως ένας πλήρης ορισμός θεωρείται ότι γενικά είναι το μέγεθος που μετρά την ομαλότητα και την κανονικότητα μιας εικόνας. [7] Διάφοροι συγγραφείς έχουν προσπαθήσει να καθορίσουν ποιητικά είναι η υφή. Οι μέθοδοι για την περιγραφή της διακρίνονται σε τρεις μεγάλες κατηγορίες, στατιστικές, φασματικές και δομικές.[8]

#### Στατιστικοί Μέθοδοι

Οι στατιστικοί μέθοδοι είναι εύχρηστοι και αρκετά απλοί μέθοδοι, καθώς χρησιμοποιούνται κατά κύριο λόγο οι ροπές φωτεινότητας του ιστογράμματος  $p_f(f)$  του ιστογράμματος. Θεωρούμε ότι  $f_k, k=1, \dots, N$  είναι τα διαφορά επίπεδα του γκρι της εικόνας. Οι ροπές που μας νοιάζουν περισσότερο είναι οι πρώτες 3 για την εξαγωγή παραμέτρων:

Μέση Τιμή	$\mu = \frac{\sum F_i}{N}$
Διασπορά	$\sigma^2 = \sum (f_k - \mu)^2 p_f(f_k)$
Κλίση	$\mu_3 = \frac{1}{\sigma^3}$

Η μέση τιμή είναι το μετρό της μέσης φωτεινότητας της περιοχής ενώ η διασπορά αποτελεί το μετρό της αντίθεσης της περιοχής[ Η κλίση δίνει ένα μετρό της συμμετρίας της φωτεινότητας της περιοχής και η κύρτωση αποτελεί ένα μετρό της ουράς του ιστογράμματος. Το μεγάλο πλεονέκτημα αυτών των περιγραφών υφής είναι η απλότητα σε ότι αφορά την ικανότητα υπολογισμού. Το κύριο μειονέκτημα είναι ότι αδυνατούν να εκφράσουν χωρικά χαρακτηριστικά της υφής

Επιπλέον, χρησιμοποιώντας τα ιστογράμματα των διάφορων επιπέδων του γκρι μπορεί να περιγράψει η χωρική πληροφορία. Υποθέτουμε ότι έστω  $d=(d_1, d_2)$  το διάνυσμα μετατόπισης μεταξύ δυο εικονοστοιχείων εικόνας και  $z(d)$  η διαφορά επιπέδου γκρι σε απόσταση  $d$ .

$$z(d)=|f(k,l)-f(k+d_1,l+d_2)|$$

ορίζουμε το  $p_z(z,d)$  το ιστόγραμμα των διάφορων επιπέδων του γκρι για την απόσταση  $d$ . Ένα ιστόγραμμα παρέχει χρήσιμες πληροφορίες για την χωρική οργάνωση ρέων εντάσεων της εικόνας. Αν μια περιοχή της εικόνας έχει τραχιά υφή, το ιστόγραμμα  $p_z(z,d)$  τείνει να έχει τιμή  $z=0$  για μικρές αποστάσεις  $d$ . Παράλληλα, αν η υφή είναι της περιοχής είναι λεπτή  $p_z(z,d)$  έχει την τάση να εξαπλώνεται, ακόμα και για μικρά  $d$  που ξεπερνούν το μέγεθος του κόκκου. Επίσης από το ιστόγραμμα των διάφορων επιπέδων του γκρι μπορούν παραλύουν παράμετροι της υφής όπως η μέση τιμή, μεταβλητότητα, η αντίθεση και η εντροπία:

Μέση τιμή:  $\mu_d = \sum g_k p_g(g_k, d)$

Οπού οι μικρές μέσες τιμές συμβολίζουν τραχιά υφή που έχει μέγεθος κόκκου ίσο η μεγαλύτερο από το μετρό του διανύσματος μετατόπισης  $d$ . αν εκφραστεί σε πολικές συντεταγμένες η μέση τιμή μπορεί να παρέχει πληροφορίες για την υφή της εικόνας.

Μεταβλητότητα σχέσης  $\sigma_d^2 = \sum (g_k - \mu)^2 p_g(g_k, d)$

Η μεταβλητότητα ορίζεται ως το μετρό διασποράς των διάφορων επίπεδων του γκρι σε μια συγκεκριμένη απόσταση  $d$ .

Αντίθεση  $c_d = \sum g_k^2 p_g(g_k, d)$

Η αντίθεση εκφράζει την ροπή αδράνειας του ιστογράμματος  $p_z(z, d)$  γύρω από το  $z=0$ , και ουσιαστικά μετρά την αντίθεση στα διαφορά επίπεδα του γκρι

Εντροπία  $H_d = -\sum p_g(g_k, d) \ln p_g(g_k, d)$

Η εντροπία μετά την ομογένεια του ιστογράμματος  $p_z(z, d)$ , και μεγιστοποιείται για τα ομοιόμορφα ιστογράμματα

Το ιστόγραμμα των διάφορων επίπεδων του γκρι έχουν μεγάλη χρησιμότητα καθώς διακρίνονται από την υπολογιστική απλότητα και για το γεγονός ότι παρέχουν πληροφορίες για την χωρική οργάνωση της υφής.

### Δομική μέθοδος

Η μέθοδος χαρακτηρίζεται από την χρήση προτύπων συνοδευόμενων από καθορισμένους κανόνες τοποθέτησης και στηρίζεται στην έννοια της επεξεργασίας της εικόνας. Η εύρεση τροπών για να οριστούν οι σχέσεις μεταξύ των εικόνων δίνουν την δυνατότητα να χαρακτηριστεί πλήρως η εικόνα, και πρακτικά γίνεται μοντελοποίηση της εικόνας για να πιο εύκολη η επεξεργασία της.

### Φασματική μέθοδος

Η μέθοδος έχει ως βάση την συνάρτηση αυτοσυσχέτιση μιας περιοχής η στην κατανομή ισχύος στην περιοχή του μετασχηματισμού Fourier με στόχο να γίνει η ανίχνευση περιοδικότητας της υφής. Επίσης, ελέγχονται οι διακεκριμένες κορυφές στο φάσμα, που δείχνουν την κατεύθυνση της υφής καθώς και η θέση τους στο πεδίο της συχνότητας, που υποδηλώνει στις χαμηλές συχνότητες τραχιά υφή και λεπτή υφή σε υψηλές συχνότητες.

### 3.10 Μορφολογικοί και γεωμετρικοί παράμετροι

Ανάλογα με το είδος του καρκίνου, υπάρχουν συγκεκριμένα χαρακτηριστικά που πολλές φορές την καθιστούν δια κρίσιμη. Κάποια από αυτά τα χαρακτηριστικά είναι το πάχος του δέρματος, η ομαλότητα, η πυκνότητα του μαστικού αδένος καθώς και το μέγεθος της περιοχής ενδιαφέροντος μαζί με κάποια χαρακτηριστικά. Τα χαρακτηριστικά που σχετίζονται με το σχήμα της περιοχής είναι η αντικανονικότητα, η ομαλότητα του σχήματος, το ποσό συμπαγής είναι, η διάμετρος, το εμβαδόν καθώς και χαρακτηριστικά των ακμών της.

Τα γεωμετρικά χαρακτηριστικά προκύπτουν άμεσα από την γεωμετρική μορφή αντικειμένων. Στη συνέχεια θα περιγράψουν ορισμένες από τα σημαντικότερα χαρακτηριστικά.

### Περίμετρος

Η περίμετρος ενός αντικειμένου ορίζεται ως το μήκος του εξωτερικού περιγράμματος ενός αντικειμένου. Σε μια ψηφιακή εικόνα το χαρακτηριστικό μπορεί να ισούται απλά με το πλήθος των εικονοστοιχείων της παραμέτρου

### Επιφάνεια

Σε μια δυαδική εικόνα, ως επιφάνεια ενός αντικειμένου καθορίζεται το πλήθος των εικονοστοιχείων από τα οποία αποτελείται ένα αντικείμενο.

### Μέγιστη και Ελάχιστη Ακτίνα

Το  $R_{max}$  και  $R_{min}$  ορίζεται ως η μέγιστη και η ελάχιστη απόσταση αντίστοιχα του κέντρου μάζας του αντικειμένου από την περιφέρεια του. Μερικές φορές ο λόγος  $R_{max}/R_{min}$  τους λαμβάνεται ως επιπλέον χαρακτηριστικό για την εκκεντρότητα ή τη μήκυνση του αντικειμένου.

### Οπές

Το πλήθος των οπών ενός αντικείμενο μπορεί να ληφθεί ως ένα ισχυρό χαρακτηριστικό.

Ο αριθμός Euler ορίζεται από την σχέση

$E = \text{πλήθος συνδεδεμένων περιοχών} - \text{πλήθος οπών}.$

### Καμπυλότητα

Η καμπυλότητα ορίζεται από την σχέση :

$$C = \frac{4\pi A}{P^2}$$

### Οπού

A: η επιφάνεια του αντικειμένου

P: η περίμετρος του αντικειμένου

Είναι φανερό ότι στην περίπτωση ενός κυκλικού αντικειμένου Έχουμε  $C=1$ , στην περίπτωση ενός τετράγωνου  $\frac{\pi}{4}$  και γενικά σε όλα τα αντικείμενα με επίμηκες σχήμα Έχουμε  $C<1$ .



### Βιβλιογραφία 3<sup>ου</sup> Κεφαλαίου

- [1]. Ritter, G.X. and Wilson, J.N., Handbook of Computer Vision Algorithms in Algebra, CRC Press, 1996.
- [2]. Tukey, J.W, Exploratory Data Analysis, Addison – Wesley, Reading MA, 1971.
- [3]. Klette, R. and Zamperoni , P., Handbook of Image Processing Operators, John Wiley & Sons, 1996.
- [4]. Haralic, R.M, “ Statistical and Stractural Approach to Texture,” Proc IEEE, 67, 5 pp. 786-804, 1979.
- [5]. Klitter , J., and Illingrowth, L., “Minimum error thresholding”, Pattern Recognition vol 19, pp. 41-47, 1986.
- [6]. Kohler, R., “A segmedation system based on threseholding”, Computer Graphics and Image Processing, CGIP-15, pp. 319-338, 1981.
- [7]. Gonzalez, R.G and Woods , R.E., Digital Image Proccessign, Addisno-Wesley, 1992.
- [8]. Hawkins, J.K, “Textural Prosperties for Pattern Recognition,”in Picture Processing and Psychopictorics, B.C Lipkin and A. Rosenfeld, Eds., Academic Press, New York, pp. 347-370, 1970.
- [9]. Brice, C.R. and Fenema, C.L., “Scene analysis using regions” Artificial Intelligence, Vol .1, pp. 205-226, 1970.
- [10] [www.uobabylon.edu.iq](http://www.uobabylon.edu.iq)
- [11][www.semanticscholar.org/paper/A-Comparative-Study-of-Histogram-Equalization-Based-Patel-Maravid](http://www.semanticscholar.org/paper/A-Comparative-Study-of-Histogram-Equalization-Based-Patel-Maravid)

Page Intentionally Left Blank

## Κεφάλαιο 4ο

### Νευρωνικά Δίκτυα

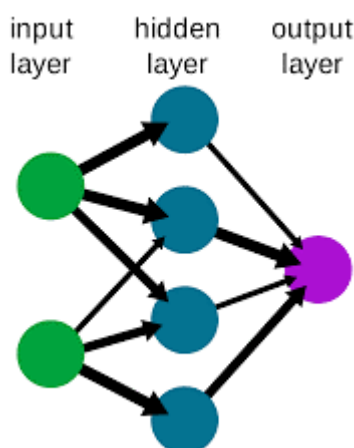
Η μελέτη του ανθρώπινου εγκεφάλου έχει διάρκεια χιλιάδων ετών. Με την έλευση της σύγχρονης ηλεκτρονικής, ήταν αναμενόμενο να γίνει προσπάθεια να αξιοποιηθεί. Το πρώτο βήμα προς την κατεύθυνση τεχνητά νευρωνικά δίκτυα ήρθαν το 1943 όταν ο Warren McCulloch, ένας νευροφυσιολόγος, μαζί με έναν νέος μαθηματικός, ο Walter Pitts, έγραψε ένα έγγραφο για το πώς νευρώνες μπορούν να δουλέψουν. Σχεδίασαν ένα απλό νευρωνικό δίκτυο με ηλεκτρικά κυκλώματα. Τα νευρωνικά δίκτυα, έχουν αξιοσημείωτο χαρακτηριστικό την ικανότητα να εξάγουν νόημα από περίπλοκα ή ασαφείς δεδομένα, μπορούν να χρησιμοποιηθούν για την εξαγωγή μοτίβων και για την ανίχνευση τάσεων που είναι πολύ περίπλοκες για να παρατηρηθούν είτε από ανθρώπους είτε από άλλους τεχνικές ηλεκτρονικών υπολογιστών. Ένα εκπαιδευμένο νευρωνικό δίκτυο μπορεί να είναι λειτουργεί ως «εμπειρογνώμονας» στην κατηγορία των πληροφοριών έχει δοθεί να αναλύσει.[1] Άλλα πλεονεκτήματα περιλαμβάνουν:

- Προσαρμοσμένη μάθηση: Δυνατότητα εκμάθησης σχετικά με την υλοποίηση καθηκόντων με βάση τα δεδομένα που δίνονται για εκπαίδευση ή αρχική εμπειρία.
- Αυτο-οργάνωση: Ένα ANN μπορεί να δημιουργήσει τη δική του την οργάνωση ή την εκπροσώπηση των πληροφοριών που διαθέτει κατά τη διάρκεια της εκμάθησης.
- Λειτουργία σε πραγματικό χρόνο: Οι υπολογισμοί των ANN μπορεί να είναι που εκτελούνται παράλληλα, καθώς και ειδικές συσκευές υλικού που σχεδιάζονται και κατασκευάζονται και εκμεταλλεύονται της δυνατότητας αυτής.
- Αντοχή σφάλματος μέσω περιττής κωδικοποίησης πληροφοριών: Μερική καταστροφή ενός δικτύου οδηγεί στην αντίστοιχη υποβάθμιση των επιδόσεων του. Ωστόσο, ορισμένες δυνατότητες δικτύου μπορεί να διατηρηθούν ακόμη και με μεγάλες ζημιές στο δίκτυο.

Τα νευρωνικά δίκτυα υιοθετούν μια διαφορετική προσέγγιση στο πρόβλημα λύση από εκείνη των συμβατικών υπολογιστών. Οι συμβατικοί υπολογιστές χρησιμοποιούν μια αλγοριθμική προσέγγιση δηλαδή τον υπολογιστή ακολουθεί ένα σύνολο οδηγιών για την επίλυση ενός προβλήματος. Εκτός από τα συγκεκριμένα βήματα που πρέπει να ακολουθήσει ο υπολογιστής είναι γνωστό ότι ο υπολογιστής δεν μπορεί να λύσει το πρόβλημα. Ότι περιορίζει την ικανότητα επίλυσης προβλημάτων των συμβατικών υπολογιστών σε προβλήματα που ήδη καταλαβαίνουμε και γνωρίζουμε πώς να λύσει. Τα νευρωνικά δίκτυα επεξεργάζονται πληροφορίες με παρόμοιο τρόπο όπως ο ανθρώπινος εγκέφαλος. Το δίκτυο αποτελείται από ένα μεγάλο αριθμό στοιχείων διασυνδεδεμένης επεξεργασίας (νευρώνες) που εργάζονται παράλληλα για την επίλυση ενός συγκεκριμένου προβλήματος. Τα νευρωνικά δίκτυα μαθαίνουν με παράδειγμα Τα παραδείγματα πρέπει να επιλεγθούν προσεκτικά, διαφορετικά χρήσιμος χρόνος χάνεται ή ακόμη και χειρότερα το δίκτυο ενδέχεται να λειτουργεί εσφαλμένα. Το μειονέκτημα είναι ότι επειδή το δίκτυο βρίσκει πώς να λύσει το πρόβλημα από μόνο του, η λειτουργία του μπορεί να είναι απρόβλεπτη. Από την άλλη πλευρά, οι συμβατικοί υπολογιστές ,χρησιμοποιούν μια γνωστική προσέγγιση για την επίλυση

προβλημάτων. Ο τρόπος με τον οποίο το πρόβλημα είναι να λυθεί πρέπει να είναι γνωστό και να αναφέρεται σε μικρές σαφείς οδηγίες. Αυτές οι οδηγίες είναι τότε μετατρέπονται σε ένα πρόγραμμα γλώσσας υψηλού επιπέδου και στη συνέχεια τον υπολογιστή που μπορεί να καταλάβει ο υπολογιστής[1-2].

Νευρωνικά δίκτυα και οι συμβατικοί αλγοριθμικοί υπολογιστές δεν ανταγωνίζονται αλλά αλληλοσυμπληρώνονται. Υπάρχουν εργασίες πιο κατάλληλες σε μια αλγοριθμική προσέγγιση όπως οι αριθμητικές πράξεις και εργασίες που είναι πιο κατάλληλες για νευρωνικά δίκτυα. Ακόμη περισσότερο, ένας μεγάλος αριθμός εργασιών, απαιτούν συστήματα που χρησιμοποιούν ένα συνδυασμός των δύο προσεγγίσεων (συνήθως συμβατικών ο υπολογιστής χρησιμοποιείται για την επίβλεψη του νευρικού δικτύου) προκειμένου να εκτελεστεί με τη μέγιστη απόδοση.



Εικόνα 22 : Απλό Νευρωνικό Δίκτυο

#### 4.1 Ορισμός Νευρωνικού Δικτύου

Τα τεχνητά νευρικά δίκτυα είναι σχετικά ακατέργαστα ηλεκτρονικά μοντέλα βασισμένα στη νευρική δομή του εγκεφάλου. Ο εγκέφαλος, βασικά μαθαίνει από την εμπειρία. Αυτό το μοντέλο εγκεφάλου, υπόσχεται επίσης μια λιγότερο τεχνική τρόπο για την ανάπτυξη λύσεων μηχανών. Η νέα αυτή προσέγγιση παρέχει επίσης μια πιο χαριτωμένη υποβάθμιση κατά τη διάρκεια υπερφόρτωσης του συστήματος από ό, τι οι πιο παραδοσιακοί αντίστοιχοι τρόποι. Αυτές οι βιολογικά εμπνευσμένες μέθοδοι υπολογιστικής, θεωρείται η επόμενη μεγάλη πρόοδος στη βιομηχανία. Ακόμα και οι απλοί ζωικοί εγκέφαλοι είναι ικανοί για λειτουργίες που είναι αδύνατο σήμερα για υπολογιστές. Οι υπολογιστές κάνουν τα πράγματα καλά, όπως τη διατήρηση βιβλίων ή την εκτέλεση πολύπλοκων μαθηματικών.

Αλλά οι υπολογιστές δυσκολεύονται να αναγνωρίσουν ακόμη και απλά μοτίβα πολύ λιγότερο γενικεύοντας αυτά τα πρότυπα του παρελθόντος σε δράσεις του μέλλοντος. Τώρα, η πρόοδος στο τομέα της βιολογίας, καθώς η έρευνα υπόσχεται μια αρχική κατανόηση του φυσικού μηχανισμού σκέψης. Αυτή η έρευνα δείχνει ότι ο εγκέφαλος, αποθηκεύει πληροφορίες ως πρότυπα. Μερικά από αυτά τα σχέδια είναι πολύ πολύπλοκα και μας επιτρέπουν την ικανότητα να αναγνωρίζουμε τα άτομα αντιμετωπίζει από πολλές διαφορετικές γωνίες. Αυτή η διαδικασία αποθήκευσης πληροφοριών ως πρότυπα, χρησιμοποιώντας αυτά τα πρότυπα, και στη συνέχεια η

επίλυση των προβλημάτων περιλαμβάνει ένα νέο πεδίο υπολογισμού. Αυτό το πεδίο, όπως προαναφέρθηκε, δεν χρησιμοποιεί παραδοσιακά εργαλεία προγραμματισμού, αλλά περιλαμβάνει τη δημιουργία μαζικά παράλληλων δικτύων και την κατάρτιση αυτών των δικτύων για την επίλυση συγκεκριμένων προβλημάτων. Αυτό το πεδίο χρησιμοποιεί επίσης λέξεις πολύ διαφορετικές από παραδοσιακά computing, λέξεις όπως συμπεριφέρονται, αντιδρούν, αυτοργανώνουν, μαθαίνουν, γενικεύουν και ξεχνούν.[2]

Όποτε γίνεται αναφορά σε ένα νευρικό δίκτυο, θα πρέπει να αναφερθεί πως το πιο δημοφιλές - Artificial Neural Network(ANN) [1], ANN είναι υπολογιστές των οποίων η αρχιτεκτονική είναι αυτή που διαμορφώθηκε με τον εγκέφαλο. Συνήθως αποτελούνται από εκατοντάδες απλών μονάδων επεξεργασίας που είναι ενωμένες σε ένα πολύπλοκο δίκτυο επικοινωνίας. Κάθε μονάδα ή κόμβος είναι απλοποιημένο μοντέλο πραγματικού νευρώνα που αποστέλλει ένα νέο σήμα ή πυρκαγιά εάν λάβει ένα επαρκώς ισχυρό σήμα εισόδου από τους άλλους κόμβους στους οποίους είναι συνδεδεμένος.

Παραδοσιακά, το νευρωνικό δίκτυο χρησιμοποιήθηκε ως δίκτυο ή ως κύκλωμα των βιολογικών νευρώνων, αλλά η σύγχρονη χρήση του όρου αναφέρεται συχνά στην ANN. Το ANN είναι μαθηματικό μοντέλο ή αλλιώς υπολογιστικό μοντέλο, ένα υπόδειγμα επεξεργασίας πληροφοριών δηλ. εμπνέεται από το βιολογικό νευρικό σύστημα,. Το ANN αποτελείται από διασυνδεδεμένους τεχνητούς νευρώνες που προγραμματίζονται σαν να μιμούνται τις ιδιότητες των βιολογικών νευρώνων m. Αυτοί οι νευρώνες που εργάζονται από κοινού για την επίλυση συγκεκριμένων προβλημάτων.

Τα ANN έχουν ρυθμιστεί για την επίλυση προβλημάτων τεχνητής νοημοσύνης χωρίς να δημιουργηθεί ένα μοντέλο πραγματικού βιολογικού συστήματος. Τα ANN χρησιμοποιούνται για την αναγνώριση ομιλίας, ανάλυση εικόνας, προσαρμοστική έλεγχο κλπ. Αυτές οι εφαρμογές γίνονται μέσω μιας διαδικασίας μάθησης, όπως η μάθηση στο βιολογικό σύστημα, που περιλαμβάνει η προσαρμογή μεταξύ των νευρώνων μέσω της συνοπτικής σύνδεση. [3]

## 4.2 Εκπαίδευση ενός Νευρωνικού δικτύου

Μόλις ένα δίκτυο έχει δομηθεί για μια συγκεκριμένη εφαρμογή, το δίκτυο αυτό είναι έτοιμο να εκπαιδευτεί. Για να ξεκινήσει η διαδικασία της εκπαίδευσης των αρχικών βαρών επιλέγονται τυχαία. Μετά την κατάρτιση ή τη μάθηση. Υπάρχουν δύο προσεγγίσεις εκπαίδευσης με επιτήρηση και χωρίς. Εποπτευόμενη εκπαίδευση περιλαμβάνει έναν μηχανισμό παροχής του δικτύου με το επιθυμητή έξοδο είτε με χειροκίνητη "ταξινόμηση" των δικτύων ή με την παροχή των επιθυμητών εξόδων με τις εισόδους. Η μη επιτηρούμενη εκπαίδευση είναι εκεί όπου το δίκτυο πρέπει να είναι να έχουν νόημα των εισροών χωρίς εξωτερική βοήθεια.

### 4.2.1 Εκπαίδευση με επιτήρηση

Στην επιτηρούμενη κατάρτιση, τόσο οι εισόδοι όσο και οι εξόδιο παρέχονται .Το δίκτυο στη συνέχεια επεξεργάζεται τις εισόδους και συγκρίνει τις προσκόπτουσες εξόδους του με τις επιθυμητές. Τα λάθη μεταδίδονται στη συνέχεια μέσω του συστήματος, προκαλώντας στο σύστημα να ρυθμίσει τα βάρη που ελέγχουν το δίκτυο. Αυτή η διαδικασία εμφανίζεται ξανά και ξανά ως βάρη. Το σύνολο δεδομένων που επιτρέπει την εκπαίδευση ονομάζεται "εκπαιδευτικό σετ". Κατά τη διάρκεια της

εκπαίδευσης δικτύου, το ίδιο σύνολο δεδομένων επεξεργάζεται πολλές φορές όπως τα βάρη σύνδεσης είναι εκλεπτυσμένα.. Ωστόσο, μερικά δίκτυα δεν μαθαίνουν ποτέ. Αυτό θα μπορούσε να συμβεί επειδή τα δεδομένα εισόδου δεν περιέχουν τις συγκεκριμένες πληροφορίες από τις οποίες προέρχεται η επιθυμητή έξοδος[2].

Τα δίκτυα επίσης δεν συγκλίνουν εάν δεν υπάρχουν αρκετά δεδομένα που επιτρέπουν την πλήρη μάθηση. Στην ιδανική περίπτωση, θα πρέπει να υπάρχει αρκετό δεδομένων, έτσι ώστε ένα μέρος των δεδομένων να μπορεί να συγκρατηθεί ως δοκιμή. Πολλά δίκτυα πολλαπλών κόμβων με πολλαπλούς κόμβους απομνημονεύοντας δεδομένα. Για να παρακολουθήσετε το δίκτυο για να διαπιστώσετε εάν το σύστημα απλά απομνημονεύει τα δεδομένα του ένα σετ που πρέπει να χρησιμοποιούνται για τη δοκιμή του συστήματος μετά την ολοκλήρωσή της εκπαίδευσης

#### 4.2.2 Μη επιτηρούμενη ή προσαρμοστική εκπαίδευση

Ο άλλος τύπος εκπαίδευσης καλείται μη επιτηρούμενη εκπαίδευση. Σε μη επιτηρούμενη εκπαίδευση, το δίκτυο παρέχεται με εισόδους αλλά όχι με τις επιθυμητές εξόδους. Το ίδιο το σύστημα πρέπει να αποφασίσει ποια χαρακτηριστικά θα χρησιμοποιήσει ομαδοποιώντας τα δεδομένα εισόδου. Αυτό συχνά αναφέρεται ως αυτοοργάνωση ή προσαρμογή. Επί του παρόντος, χωρίς επίβλεψη η μάθηση δεν είναι καλά κατανοητή. Αυτή η προσαρμογή στο περιβάλλον είναι η υπόσχεση που θα επιτρέψει την επιστήμη να μαθαίνουν συνεχώς από μόνα τους και αντιμετωπίζουν νέες καταστάσεις και νέα περιβάλλοντα. Η ζωή είναι που γεμίζουν με καταστάσεις όπου δεν υπάρχουν ακριβή σύνολα εκπαίδευσης. Ορισμένες από αυτές τις καταστάσεις περιλαμβάνουν στρατιωτική δράση όπως καινούργιες τεχνικές μάχης και νέα όπλα. Λόγω αυτής της απροσδόκητης πτυχής στη ζωή και στον άνθρωπο επιθυμία να είναι προετοιμασμένοι, εξακολουθούν να διεξάγονται έρευνες, στον τομέα αυτό. Ωστόσο, προς το παρόν, το τεράστιο το μεγαλύτερο μέρος της εργασίας του νευρικού δικτύου είναι σε συστήματα με επίβλεψη μάθηση. Η εποπτευόμενη μάθηση επιτυγχάνει αποτελέσματα[2].

### 4.3 Εφαρμογή των ANN

Οι διάφορες εφαρμογές του Τεχνητού Νευρικού σε πραγματικό χρόνο είναι οι εξής:

1. Προσέγγιση λειτουργιών ή ανάλυση παλινδρόμησης, συμπεριλαμβανομένης της πρόβλεψης χρονοσειρών και της μοντελοποίησης.
2. Έλεγχος κλήσεων - απάντηση σε εισερχόμενη κλήση (ηχείο-ON) με ένα κύμα του χεριού κατά την οδήγηση.
3. Ταξινόμηση, συμπεριλαμβανομένου του σχεδίου και της ακολουθίας αναγνώριση, ανίχνευση καινοτομίας και διαδοχική απόφαση κατασκευή.
4. Παραλείψτε κομμάτια ή ελέγξτε την ένταση του ήχου στο media player σας χρησιμοποιώντας απλές κινήσεις του χεριού - χαλαρώστε πίσω και χωρίς να χρειάζεται για να μεταβείτε στη συσκευή.
5. Επεξεργασία δεδομένων, συμπεριλαμβανομένου του φιλτραρίσματος, της ομαδοποίησης, του τυφλού διαχωρισμού σήματος και συμπίεση.

6. Μετακίνηση σε ιστοσελίδες, ή μέσα σε ένα eBook με απλή αριστερά και δεξιές χειρονομίες.[3]

7. Οι τομείς εφαρμογής των ANN περιλαμβάνουν το σύστημα αναγνώριση και έλεγχος (έλεγχος οχήματος, διαδικασία έλεγχος), παιχνίδι και λήψη αποφάσεων (τάβλι, σκάκι, αγώνες), αναγνώριση προτύπων (ραντάρ συστήματα αναγνώρισης προσώπων, αναγνώριση αντικειμένων κ.λπ.), αναγνώριση ακολουθίας (χειρονομία, ομιλία, χειρόγραφο κείμενο αναγνώριση), ιατρική διάγνωση, οικονομική εφαρμογές, εξόρυξη δεδομένων (ή ανακάλυψη γνώσεων στο βάσεις δεδομένων, "KDD").

8. Μια άλλη ενδιαφέρουσα περίπτωση χρήσης είναι όταν χρησιμοποιείτε το Smartphones ως κόμβος πολυμέσων, ο χρήστης μπορεί να συνδέσει τη συσκευή στην τηλεόραση και να παρακολουθήσετε περιεχόμενο από τη συσκευή, ενώ ελέγχοντας το περιεχόμενό του με τρόπο απαλλαγμένο από επαφή μακριά.[3]

#### 4.4 Πλεονεκτήματα Νευρωνικών Δικτύων

Κάποια χαρακτηριστικά πλεονεκτήματα νευρωνικών δικτύων:

- Προσαρμοσμένη μάθηση: Ικανότητα να μάθει πώς να κάνει τα καθήκοντα με βάση τα δεδομένα που δίνονται για εκπαίδευση ή αρχική εμπειρία.
- Αυτο-οργάνωση: Ένα ANN μπορεί να δημιουργήσει το δικό του την οργάνωση ή την εκπροσώπηση των πληροφοριών που διαθέτει κατά τη διάρκεια της εκμάθησης.
- Λειτουργία σε πραγματικό χρόνο: Οι υπολογισμοί των ANN μπορεί να είναι που εκτελούνται παράλληλα, και ειδικές συσκευές υλικού σχεδιάζονται και κατασκευάζονται πλεονέκτημα αυτής της δυνατότητας.
- Η αναγνώριση μοτίβων είναι μια ισχυρή τεχνική για αξιοποίηση των πληροφοριών στα δεδομένα και γενικεύοντας γι 'αυτό. Τα νευρικά δίκτυα μαθαίνουν να αναγνωρίζουν τα πρότυπα που υπάρχουν στο σύνολο δεδομένων.
- Το σύστημα αναπτύσσεται μέσω της μάθησης αντί του προγραμματισμού. Τα νευρωνικά δίκτυα διδάσκουν τον εαυτό τους πρότυπα στα δεδομένα που απελευθερώνουν τον δημιουργό του για περισσότερα ενδιαφέρουσα εργασία.
- Τα νευρωνικά δίκτυα είναι ευέλικτα σε μια αλλαγή στο περιβάλλον. Αν και τα νευρωνικά δίκτυα μπορεί να πάρουν κάποια στιγμή να μάθουν μια ξαφνική δραστητική αλλαγή είναι εξαιρετική στην προσαρμογή στις συνεχώς μεταβαλλόμενες πληροφορίες.
- Τα νευρωνικά δίκτυα μπορούν να δημιουργήσουν ενημερωτικά μοντέλα όταν συμβατικές προσεγγίσεις αποτυγχάνουν. Επειδή τα νευρωνικά δίκτυα μπορούν να χειριστούν πολύ σύνθετες αλληλεπιδράσεις μπορούν εύκολα να μοντελοποιήσουν δεδομένα τα οποία είναι πολύ δύσκολα μοντέλο με παραδοσιακές προσεγγίσεις, όπως στατιστικές ή λογικός προγραμματισμός.
- Η απόδοση των νευρωνικών δικτύων είναι τουλάχιστον τόσο καλή όσο κλασική στατιστική μοντελοποίηση και καλύτερα στα περισσότερα προβλήματα. Τα νευρωνικά δίκτυα δημιουργούν μοντέλα που είναι που αντικατοπτρίζει περισσότερο τη δομή των δεδομένων στο σημαντικά λιγότερο χρόνο.



## 4.5 Κατηγορίες Νευρωνικών δικτύων

Τα νευρωνικά δίκτυα που συνήθως χρησιμοποιούνται για την διάγνωση σε μαστογραφίες είναι:

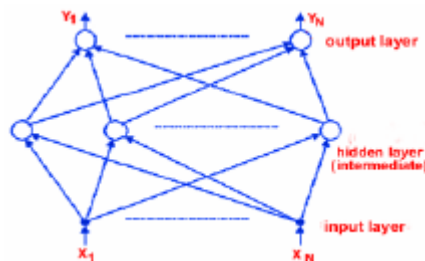
- Τα πολυεπίπεδα νευρωνικά δίκτυα perceptron
- Τα συνελκτικά νευρωνικά δίκτυα
- Τα νευρωνικά δίκτυα ακτινικών συναρτήσεων

Έπειτα θα αναλυθούν αναλυτικά αυτές οι κατηγορίες.

## 4.6 Πολυεπίπεδα Νευρωνικά δίκτυα Perceptron

### 4.6.1 Δομή νευρωνικού δικτύου Perceptron

Το πολυεπίπεδο perceptron είναι το πιο γνωστό και ο πιο συχνά χρησιμοποιούμενος τύπος νευρικού δικτύου. Συνήθως, τα σήματα μεταδίδονται εντός του δίκτυο προς μία κατεύθυνση: από είσοδο σε έξοδο. Αυτή η αρχιτεκτονική ονομάζεται feedforward(εμπρόσθιας διάδοσης). Ένα χαρακτηριστικό σχήμα φαίνεται στην εικόνα 2.



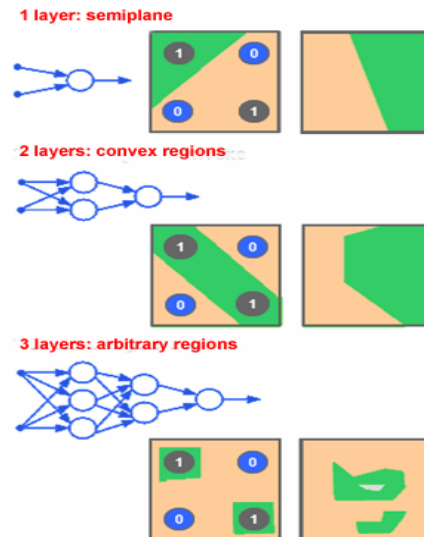
Εικόνα 23: Νευρωνικό δίκτυο εμπρόσθιας διάδοσης.

Επίπεδα που δεν είναι άμεσα συνδεδεμένα με το περιβάλλον ονομάζονται κρυφά. Υπάρχει μια διαμάχη σχετικά με το πρώτο επίπεδο (το επίπεδο εισόδου) θεωρείται αυτόνομο επίπεδο στο δίκτυο, αφού είναι μόνο.

Υπάρχουν επίσης δίκτυα τροφοδοσίας, τα οποία μπορούν να μεταδίδουν παλμούς στο και τις δύο κατευθύνσεις, λόγω των συνδέσεων αντίδρασης στο δίκτυο. Αυτοί οι τύποι δικτύων είναι πολύ ισχυροί και μπορεί να είναι εξαιρετικά περίπλοκοι. Αυτή η δυναμική, μεταβάλλει την κατάστασή τους όλη την ώρα, μέχρι το δίκτυο να φτάσει σε κατάσταση ισορροπίας και η αναζήτηση για μια νέα ισορροπία συμβαίνει με κάθε αλλαγή στην είσοδο. Η εισαγωγή πολλών επιπέδων καθορίζεται από την ανάγκη να αυξηθεί η πολυπλοκότητα του.

Όπως φαίνεται στη προηγούμενη παράγραφο, ένα perceptron με ένα μόνο στρώμα και μια η είσοδος δημιουργεί περιοχές απόφασης υπό τη μορφή ημίν επίπεδα. Προσθέτοντας ένα άλλο στρώμα, κάθε νευρώνας ενεργεί ως τυπικό perceptron για τις εξόδους του νευρώνες στο πρόσθιο στρώμα, έτσι η έξοδος του δίκτυο μπορεί να εκτιμήσει περιοχές αποφάσεων, που προκύπτει από τη διασταύρωση των ημί επιπέδων δημιουργούνται από τους νευρώνες. Με τη σειρά του, ένα τριών επιπέδων Το perceptron μπορεί να δημιουργήσει αυθαίρετους τομείς αποφάσεων (Εικόνα. 2).



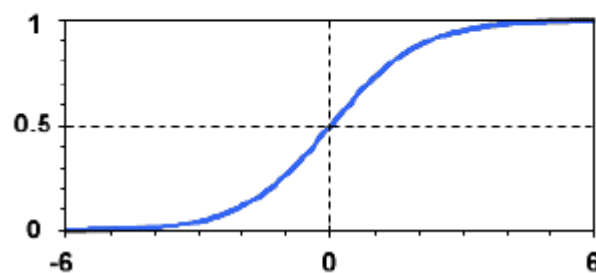


Εικόνα 24: Περιοχές απόφασης σε πολυεπίπεδο perceptron.

Όσον αφορά τη λειτουργία ενεργοποίησης τους νευρώνες, βρέθηκε ότι τα δίκτυα πολλαπλών επιπέδων δεν παρέχει αύξηση της υπολογιστικής ισχύος σε σύγκριση με δίκτυα με ένα μόνο επίπεδο. Αν οι λειτουργίες ενεργοποίησης είναι γραμμικές, επειδή μια γραμμική συνάρτηση των γραμμικών συναρτήσεων είναι επίσης μια γραμμική συνάρτηση.

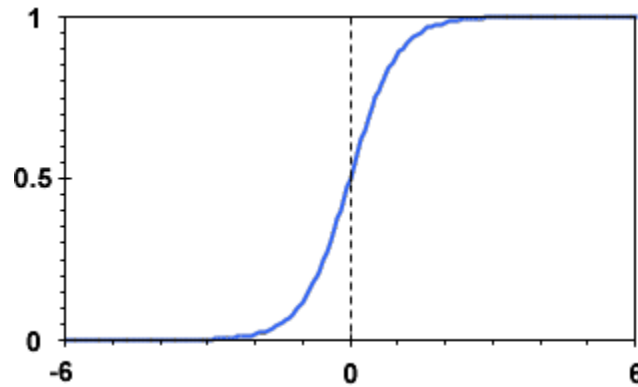
Η ισχύς του πολυεπίπεδου perceptron προκύπτει ακριβώς από μη γραμμικές λειτουργίες ενεργοποίησης. Σχεδόν οποιαδήποτε μη γραμμική συνάρτηση μπορεί να χρησιμοποιηθεί για αυτό σκοπό, εκτός από τις πολωνυμικές συναρτήσεις. Επί του παρόντος, οι λειτουργίες που χρησιμοποιούνται συνήθως είναι μονοπολικό (ή λογιστικό) σιγμοειδές, που φαίνεται στη Εικόνα 4:

$$f(s) = \frac{1}{1 + e^{-s}}.$$



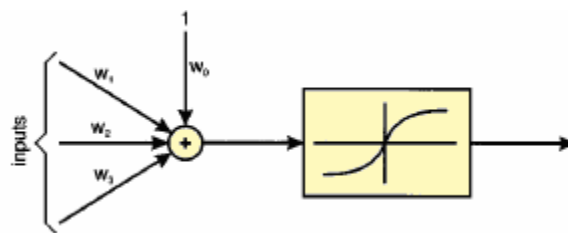
Εικόνα 25: Σιγμοειδής συνάρτηση

Θα πρέπει να σημειωθεί ότι οι σιγμοειδείς λειτουργίες ενεργούν περίπου γραμμικά για μικρές απόλυτες τιμές, και αναλαμβάνουν τον ρόλο του κατωφλίου για υψηλές απόλυτες τιμές. Έχει αποδειχθεί [4] ότι ένα δίκτυο με ένα κρυφό στρώμα είναι σε θέση να υπολογίσει κατά προσέγγιση οποιαδήποτε συνεχή λειτουργία.

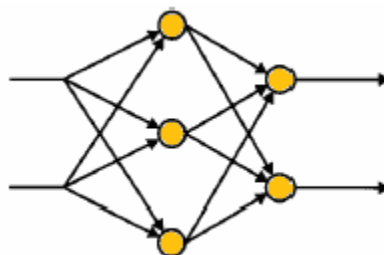


Εικόνα 26: Σιγμοειδής Μονοπολική.

Αυτό δικαιολογεί την ιδιότητα του πολυεπίπεδου perceptron να ενεργεί ως μέθοδος καθολικής προσέγγισης. Επίσης, εφαρμόζοντας το θεώρημα Stone-Weierstrass στο νευρωνικό δίκτυο, αποδείχθηκε ότι μπορούν να υπολογιστούν ορισμένες πολωνυμικές εκφράσεις: εάν υπάρχουν δύο δίκτυα που υπολογίζουν ακριβώς δύο συναρτήσεις  $f_1$ , δηλαδή το  $f_2$ , τότε υπάρχει ένα μεγαλύτερο δίκτυο που υπολογίζει ακριβώς μια πολωνυμική έκφραση των  $f_1$  και  $f_2$ . Το Multi Perceptron είναι το πιο γνωστό και πιο δημοφιλές τύπος νευρωνικών δικτύων είναι εκπαιδευμένες μονάδες του τύπου φαίνεται στο Εικόνα 6. Κάθε μία από αυτές τις μονάδες σχηματίζει α σταθμισμένο άθροισμα των εισόδων του στις οποίες προστίθενται α συνεχής. Αυτό το ποσό μεταφέρεται στη συνέχεια μέσω ενός μη γραμμικού συνάρτηση που συχνά ονομάζεται ενεργοποίηση λειτουργία. Οι περισσότερες μονάδες διασυνδέονται κατά τρόπο "feed forward", δηλαδή διασυνδέσεις που σχηματίζουν βρόχο όπως φαίνεται στο Σχ. 6.

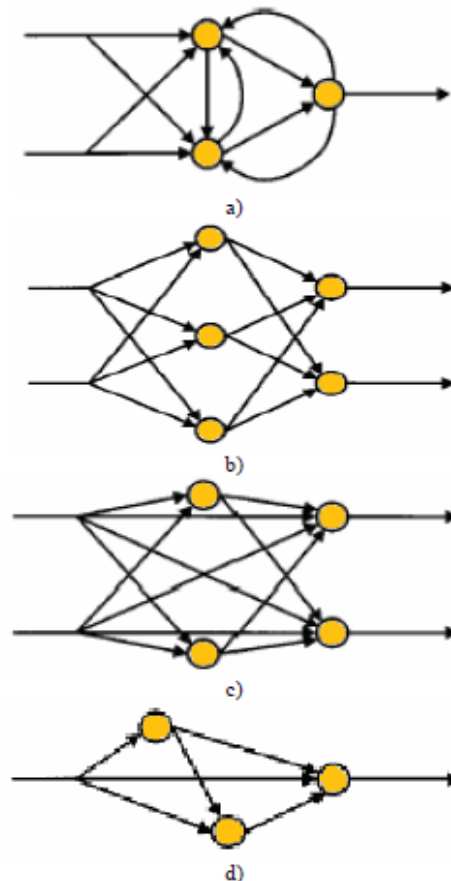


Εικόνα 27: Πολυεπίπεδος Perceptron.



Εικόνα 28: Παράδειγμα δίκτυο εμπρόσθιας διάδοσης.

Για ορισμένους τύπους εφαρμογών επαναλαμβανόμενα δίκτυα (π.χ. όχι "feed forward"), στα οποία ορισμένες διασυνδέσεις χρησιμοποιούν βρόχο διαμόρφωσης, Έχει παρατηρηθεί όπως παρατηρείται στην Εικόνα 7 ένα παράδειγμα δικτύου προώθησης τροφοδοσίας. Όπως αναφέρθηκε μονάδες διασύνδεσης αυτού του τύπου δικτύου κάνει ένα και όχι μορφή βρόχου, επομένως το δίκτυο ονομάζεται τροφοδοσία προς τα εμπρός. Δίκτυα στα οποία υπάρχει ένας ή περισσότεροι βρόχοι διασυνδέσεις όπως απεικονίζονται στο σχήμα της Εικόνας 8.α .



Εικόνα 29 Συνηθισμένοι τύποι δικτύων: α) επαναλαμβανόμενο δίκτυο; β) ένα στρωματοποιημένο δίκτυο · γ) ένα δίκτυο με συνδέσεις μεταξύ μονάδων εισόδου και εξόδου · δ) μια τροφή δίκτυο προώθησης πλήρως συνδεδεμένο στα δίκτυα προώθησης τροφοδοσίας, οι μονάδες είναι συνήθως διευθετημένες σε επίπεδα (επίπεδα) όπως στο Σχήμα 7.β αλλά και άλλες τοπολογίες μπορεί να χρησιμοποιηθεί. Το σχήμα 7.ε δείχνει έναν τύπο δικτύου που είναι χρήσιμο σε ορισμένες εφαρμογές στις οποίες οι απευθείας συνδέσεις μεταξύ των μονάδων εισόδου και εξόδου χρησιμοποιούνται. Εικόνα 7.δ δείχνει ένα δίκτυο με 3 μονάδες που είναι πλήρως συνδεδεμένο, δηλαδή επιτρέπεται σε όλες τις διασυνδέσεις περιορισμός ροής προς τα εμπρός.

#### 4.6.2 Ο αλγόριθμος backpropagation

Τα δίκτυα μάθησης επιτυγχάνονται συνήθως με ένα εποπτευόμενο τρόπο. Μπορεί να θεωρηθεί ότι είναι διαθέσιμο ένα μαθησιακό περιβάλλον που περιέχει και τα δύο μοντέλα μάθησης και μοντέλα επιθυμητής παραγωγής αντιστοιχεί στην είσοδο (αυτό είναι γνωστό ως "στόχος" μοντέλα "). Όπως θα δούμε, η μάθηση βασίζεται συνήθως σχετικά με την ελαχιστοποίηση των σφαλμάτων μέτρησης μεταξύ εξόδων δικτύου και επιθυμητών εξόδων. Αυτό συνεπάγεται, πίσω διάδοση μέσω ενός δικτύου παρόμοιου με αυτό που γίνεται η μάθηση. Για αυτόν τον λόγο η εκμάθηση αλγορίθμων ονομάζεται

οπίσθια διάδοση. Η μέθοδος ήταν η πρώτη προτάθηκε [5], αλλά εκείνη τη στιγμή ήταν ουσιαστικά αγνοήθηκε, επειδή υποτίθεται ότι υπολόγιζε τον μεγάλο όγκο.

Στη συνέχεια ανακαλύφθηκε ξανά από [6], αλλά μόνο στα μέσα της δεκαετίας του '80 ξεκίνησε ο Williams [18] ως ένα γενικά αποδεκτό εργαλείο για την εκπαίδευση του πολυεπίπεδο perceptron. Η ιδέα είναι να βρείτε το συνάρτηση ελάχιστου σφάλματος  $e(w)$  σε σχέση με τα βάρη συνδέσεων. Ο αλγόριθμος για ένα πολυστρωματικό στρώμα το perceptron με ένα κρυφό επίπεδο είναι το ακόλουθο [8]:

Βήμα 1: Αρχικοποίηση. Όλα τα βάρη δικτύου και τα όρια αρχικοποιούνται με τυχαίες τιμές, κατανέμεται ομοιόμορφα σε μικρό εύρος, για παράδειγμα  $\left(\frac{-2.4}{F_i}, \frac{2.4}{F_i}\right)$ , όπου το  $F_i$  είναι συνολικός αριθμός εισόδων του νευρώνα  $i$  [9]. Εάν αυτές οι τιμές είναι 0, οι κλίσεις που θα υπολογιστούν κατά τη διάρκεια της δοκιμής θα είναι επίσης 0 (εάν δεν υπάρχει άμεση σύνδεση μεταξύ της εισόδου και έξοδος) και το δίκτυο δεν θα μάθει. Περισσότερο υποδεικνύονται προσπάθειες εκπαίδευσης, με διαφορετικά αρχικά βάρη, για να βρεθεί η καλύτερη τιμή για τη συνάρτηση κόστους (ελάχιστο σφάλμα). Αντίθετα, εάν οι αρχικές τιμές είναι μεγάλο, τείνουν να κορεστούν αυτές τις μονάδες. Σε αυτήν την περίπτωση, η σιγμοειδής συνάρτηση είναι πολύ μικρή. Λειτουργεί ως πολλαπλασιαστικός παράγοντας κατά τη διάρκεια της μαθησιακής διαδικασίας και έτσι οι κορεσμένες μονάδες θα είναι σχεδόν μπλοκαρισμένες, η οποία καθιστά τη μάθηση πολύ αργή.

Βήμα 2: Μια νέα εποχή εκπαίδευσης, παρουσιάζει όλα τα παραδείγματα στο σετ προπόνησης. Στις περισσότερες περιπτώσεις, η εκπαίδευση του δικτύου περιλαμβάνει περισσότερα εκπαιδευτικές εποχές. Για να διατηρηθεί η μαθηματική αυστηρότητα, τα βάρη θα ρυθμίζονται μόνο μετά από όλα τα διανύσματα δοκιμής που θα εφαρμοστούν στο δίκτυο. Επομένως, οι κλίσεις των βαρών πρέπει να απομνημονεύονται και να προσαρμόζονται μετά από κάθε μοντέλο στο σετ προπόνησης, τα βάρη θα αλλάξουν μόνο μία φορά (υπάρχει μια "on-line" παραλλαγή, πιο απλό, στο οποίο ενημερώνονται τα βάρη άμεσα, σε αυτήν την περίπτωση, τη σειρά με την οποία τα διανύσματα του δικτύου που παρουσιάζονται μπορεί να έχει σημασία. Όλες οι βαθμίδες των βαρών και το τρέχον σφάλμα αρχικοποιούνται με 0 ( $\Delta w_{ij} = 0$  και  $E = 0$ )).

Βήμα 3: Η προώθηση του σήματος προς τα εμπρός

Ένα παράδειγμα από το σετ εκπαίδευσης εφαρμόζεται στις εισόδους, και οι έξοδοι των νευρώνων από το κρυφό στρώμα υπολογίζονται:

$$y_j(p) = f\left(\sum_{i=1}^n x_i(p) \cdot w_{ij} - \theta_j\right) \quad (3)$$

όπου  $n$  είναι ο αριθμός εισόδων για τον νευρώνα  $j$  από το κρυφό στρώμα, και  $f$  είναι η σιγμοειδής συνάρτηση.

Ακόμη, οι πραγματικές έξοδοι του δικτύου υπολογίζονται:

$$y_k(p) = f\left(\sum_{j=1}^m x_{jk}(p) \cdot w_{jk}(p) - \theta_k\right) \quad (4)$$

όπου  $m$  είναι ο αριθμός εισόδων για τον νευρώνα  $k$  από το επίπεδο εξόδου.

Το σφάλμα ενημερώνεται:

$$E = E + \frac{(e_k(p))^2}{2} \quad (5)$$

Βήμα 4: Η προς τα πίσω διάδοση των λαθών και τις προσαρμογές των βαρών.

Οι κλίσεις των σφαλμάτων για τους νευρώνες υπολογίζεται στο επίπεδο εξόδου:

$$\delta_k(p) = f' \cdot e_k(p) \quad (6)$$

όπου  $f'$  είναι η παράγωγη συνάρτηση για την ενεργοποίηση, και το σφάλμα  $e_k(p) = y_{d,k}(p) - y_k(p)$ .

Εάν χρησιμοποιούμε τη μονοπολική σιγμοειδής (εξίσωση 1), παράγεται είναι:

$$f'(x) = \frac{e^{-x}}{(1+e^{-x})^2} = f(x) \cdot (1-f(x)) \quad (7)$$

Εάν χρησιμοποιούμε το διπολικό σιγμοειδές (εξίσωση 2), τότε προκύπτει:

$$f'(x) = \frac{2a \cdot e^{-ax}}{(1+e^{-ax})^2} = \frac{a}{2} \cdot (1-f(x)) \cdot (1+f(x)) \quad (8)$$

Επιπλέον, θεωρείται ότι η συνάρτηση που χρησιμοποιείται είναι η μονοπολική σιγμοειδής. Στη συνέχεια, η εξίσωση (6) γίνεται:

$$\delta_k(p) = y_k(p) \cdot (1-y_k(p)) \cdot e_k(p) \quad (9)$$

Οι κλίσεις για τα βάρη μεταξύ των κρυφών και το επίπεδο εξόδου ενημερώνονται:

$$\Delta w_{jk}(p) = \Delta w_{jk}(p) + y_j(p) \cdot \delta_k(p) \quad (10)$$

Οι κλίσεις των σφαλμάτων για τους νευρώνες υπολογίζονται το κρυφό επίπεδο:

$$\delta_j(p) = y_j(p) \cdot (1-y_j(p)) \cdot \sum_{k=1}^l \delta_k(p) \cdot w_{jk}(p) \quad (11)$$

όπου  $l$  είναι ο αριθμός των εξόδων για το δίκτυο.

Οι διαβαθμίσεις των βαρών μεταξύ της εισόδου και το κρυφό επίπεδο ενημερώνονται:

$$\Delta w_{ij}(p) = \Delta w_{ij}(p) + x_i(p) \cdot \delta_j(p) \quad (12)$$

Βήμα 5: Μια νέα επανάληψη.

Εάν εξακολουθούν να υπάρχουν διανύσματα δοκιμής στην τρέχουσα εκπαίδευση, γίνεται μετάβαση στο βήμα 3. Εάν όχι, τα βάρη όλα οι συνδέσεις θα ενημερωθούν με βάση τις διαβαθμίσεις των βαρών:

$$w_{ij} = w_{ij} + \eta \cdot \Delta w_{ij}$$

όπου  $\eta$  είναι το ποσοστό εκμάθησης.

#### 4.6.3 Μεθοδος επιτάχυνσης της μάθησης

Η μέθοδος ορμής [7] προτείνει την προσθήκη ενός όρου για να προσαρμοστούν τα βάρη. Αυτός ο όρος είναι ανάλογος με τον τελευταίο τροποποίηση του βάρους, δηλαδή των τιμών με τις οποίες τα βάρη ρυθμίζονται αποθηκεύονται και απευθείας επηρεάζουν όλες τις περαιτέρω προσαρμογές:

$$\Delta w_{ij}(p) = \Delta w_{ij}(p) + \alpha \cdot \Delta w_{ij}(p-1)$$

Η προσθήκη ενός νέου όρου γίνεται μετά την ενημέρωση της κλίσεις για τα βάρη από τις εξισώσεις 10 και 12. Η μέθοδος του μεταβλητού ρυθμού εκμάθησης [19] είναι να χρησιμοποιήσετε ένα ατομικό ποσοστό εκμάθησης για κάθε βάρος και προσαρμογή αυτές οι παράμετροι σε κάθε επανάληψη, ανάλογα με το διαδοχικά σημάδια των κλίσεων [9]:

$$\eta_{ij}(p) = \begin{cases} u \cdot \eta_{ij}(p-1), \text{sign}(\Delta w_{ij}(p)) = \text{sign}(\Delta w_{ij}(p-1)) \\ d \cdot \eta_{ij}(p-1), \text{sign}(\Delta w_{ij}(p)) \neq \text{sign}(\Delta w_{ij}(p-1)) \end{cases}$$

Εάν κατά τη διάρκεια της προπόνησης το σφάλμα αρχίσει να αυξάνεται, αντί για μείωση, τα ποσοστά εκμάθησης επαναφέρονται σε αρχικές τιμές και στη συνέχεια η διαδικασία συνεχίζεται.

### 4.7 Συνελικτικά νευρωνικά δίκτυα

Τα Convolutional Neural Networks (CNN) είναι ανάλογα με τα παραδοσιακά ANN στο ότι αποτελούνται από νευρώνες που αυτο-βελτιστοποιούνται μέσω της μάθησης. Κάθε νευρώνας θα εξακολουθεί να λαμβάνει μια είσοδο και να εκτελεί μια λειτουργία (όπως μια βαθμίδα) προϊόν που ακολουθείται από μια μη γραμμική συνάρτηση) - η βάση των αμέτρητων ANN. Από τα διανύσματα ακατέργαστης εικόνας εισόδου στην τελική έξοδο της βαθμολογίας τάξης, το σύνολο των δικτύων θα εξακολουθεί να εκφράζει μια συνάρτηση αντίληψης σκορ (το βάρος). Το τελευταίο επίπεδο θα περιέχει συναρτήσεις απώλειας που σχετίζονται με τις κλάσεις και όλες εξακολουθούν να ισχύουν κανονικά όπως τα παραδοσιακά ANN.

Η μόνη αξιοσημείωτη διαφορά μεταξύ των CNN και των παραδοσιακών ANN είναι ότι τα CNN χρησιμοποιούνται κυρίως στον τομέα της αναγνώρισης προτύπων μέσα σε εικόνες. Αυτό επιτρέπει την κωδικοποίηση ειδικών χαρακτηριστικών εικόνας στην αρχιτεκτονική, κάνοντας το δίκτυο πιο κατάλληλο για εργασίες που εστιάζουν στην εικόνα - μειώνοντας παράλληλα τις παραμέτρους που απαιτείται για τη ρύθμιση του μοντέλου. Ένας από τους μεγαλύτερους περιορισμούς των παραδοσιακών μορφών του ANN είναι ότι τείνουν ανταγωνίζονται με την υπολογιστική πολυπλοκότητα που απαιτείται για τον υπολογισμό των δεδομένων εικόνας.[16]

#### 4.7.1 Overfitting

Σίγουρα θα ήταν εφικτό απλώς να αυξήσουμε τον αριθμό των κρυφών επιπέδων στο δίκτυό, και ίσως να αυξηθεί ο αριθμός των νευρώνων μέσα τους, όμως δεν συνίσταται. Αυτό οφείλεται σε δύο λόγους, ο ένας είναι το απλό πρόβλημα της μη απεριορίστης υπολογιστικής ισχύος και χρόνου για να γίνει η εκπαίδευση σε τεράστια ANN.

Ο δεύτερος λόγος είναι η διακοπή ή η μείωση των επιπτώσεων της overfitting). Είναι βασικά όταν ένα δίκτυο δεν είναι σε θέση να μάθει αποτελεσματικά λόγω ενός αριθμού αιτιών. Είναι μια σημαντική έννοια των περισσότερων, αν όχι όλων των αλγορίθμων μηχανικής μάθησης και είναι σημαντικό να λαμβάνεται κάθε προφύλαξη για τη μείωση των επιπτώσεων της. Εάν τα μοντέλα μας παρουσίαζαν σημάδια υπερβολικής τοποθέτησης, τότε μπορεί να δούμε μειωμένη δυνατότητα εντοπισμού γενικευμένων λειτουργιών όχι μόνο για το σύνολο δεδομένων εκπαίδευσης, αλλά επίσης τα σετ δοκιμών και προβλέψεων[17].

Αυτός είναι ο κύριος λόγος για τη μείωση της πολυπλοκότητας των ANN μας. οι λιγότεροι παράμετροι που απαιτούνται για την προπόνηση, τόσο λιγότερο πιθανό θα είναι υπερβολικό το δίκτυο – και φυσικά, να βελτιωθεί η προγνωστική απόδοση του μοντέλου.

#### 4.7.2 Αρχιτεκτονική CNN

Όπως σημειώθηκε νωρίτερα, τα CNN εστιάζουν κυρίως στη βάση ότι θα περιληφθεί η εισαγωγή εικόνων. Αυτό εστιάζει την αρχιτεκτονική που πρέπει να διαμορφωθεί έτσι ώστε να ταιριάζει καλύτερα την ανάγκη αντιμετώπισης του συγκεκριμένου τύπου δεδομένων.

Μία από τις βασικές διαφορές είναι ότι οι νευρώνες που στρώνονται μέσα στο CNN αποτελούνται από νευρώνες οργανωμένους σε τρεις διαστάσεις, τη χωρική διάσταση της εισόδου (ύψος και πλάτος) και το βάθος. Το βάθος δεν είναι ο συνολικός αριθμός επιπέδων εντός του ANN, αλλά στην τρίτη διάσταση του όγκου ενεργοποίησης. Σε αντίθεση με τα τυπικά ANNs, οι νευρώνες σε οποιοδήποτε δεδομένο στρώμα θα συνδεθεί μόνο σε μια μικρή περιοχή του στρώματος που προηγείται.

#### 4.7.3 Συνθετικό στρώμα

Όπως υποδηλώνει το όνομα, το συνελκτικό επίπεδο παίζει ζωτικό ρόλο στον τρόπο με τον οποίο τα CNNs λειτουργεί. Οι παράμετροι επιπέδων επικεντρώνονται στη χρήση μαθημάτων που μπορούν να διδαχθούν. Αυτοί οι πυρήνες είναι συνήθως μικροί σε χωρική διάσταση, αλλά απλώνονται κατά μήκος του ολόκληρου το βάθος της εισόδου. Όταν τα δεδομένα χτυπήσουν ένα συνελκτικό επίπεδο, το στρώμα εμπλέκει κάθε φίλτρο σε όλη τη χωρική διάσταση της εισόδου στο να δημιουργηθεί ένας χάρτη 2D ενεργοποίησης. Αυτοί οι χάρτες ενεργοποίησης μπορούν να απεικονιστούν, όπως φαίνεται στο παρακάτω σχήμα

Καθώς περνάμε μέσα από την είσοδο, το κλιμακωτό προϊόν υπολογίζεται για κάθε τιμή σε αυτόν τον πυρήνα. Από αυτό το δίκτυο θα μάθει πυρήνες που «πυροδοτούν» όταν βλέπουν ένα συγκεκριμένο χαρακτηριστικό σε μια δεδομένη χωρική θέση της εισόδου. Αυτά είναι κοινώς γνωστό ως ενεργοποιήσεις.

Κάθε πυρήνας θα έχει έναν αντίστοιχο χάρτη ενεργοποίησης, του οποίου θα στοιβάζεται κατά μήκος της διάστασης βάθους για να σχηματίσει τον πλήρη όγκο εξόδου από το συνελκτικό στρώμα.

Όπως αναφέραμε νωρίτερα, η εκπαίδευση ANNs σε εισόδους όπως εικόνες οδηγεί σε μοντέλα των οποίων είναι πολύ μεγάλα για αποτελεσματική εκπαίδευση. Αυτό έρχεται με πλήρως συνδεδεμένο τρόπο των τυπικών νευρώνων ANN, έτσι ώστε να



μετριαστούν από αυτό. Ο νευρώνας σε ένα συνελκτικό στρώμα συνδέεται μόνο με μια μικρή περιοχή της εισόδου. Η διάσταση αυτής της περιοχής αναφέρεται συνήθως ως δεκτική μέγεθος πεδίου του νευρώνα. Το μέγεθος της συνδεσιμότητας μέσω του το βάθος είναι σχεδόν πάντα ίσο με το βάθος της εισόδου.

Τα συνελκτικά στρώματα μπορούν επίσης να μειώσουν σημαντικά την πολυπλοκότητα του μοντέλου μέσω της βελτιστοποίησης της παραγωγής του. Αυτά βελτιστοποιούνται μέσω τριών υπερπαραμέτρων, το βάθος, το βήμα και τη ρύθμιση μηδενικής επένδυσης. Το βάθος του όγκου εξόδου που παράγεται από τα συνελκτικά στρώματα μπορούν να ρυθμιστούν χειροκίνητα τον αριθμό των νευρώνων εντός του στρώματος στην ίδια περιοχή της εισόδου. Αυτό μπορεί να φανεί με άλλες μορφές ANN, όπου όλοι οι νευρώνες στο κρυφό στρώμα συνδέονται άμεσα με κάθε έναν νευρώνα εκ των προτέρων. Η μείωση αυτού του υπερπαραμέτρου μπορεί να ελαχιστοποιήσει σημαντικά τον συνολικό αριθμό των νευρώνων του δικτύου, αλλά μπορεί επίσης να μειώσει σημαντικά τις δυνατότητες αναγνώρισης προτύπων του μοντέλου.[19]

Μπορούμε επίσης να καθορίσουμε το βήμα στο οποίο καθορίζουμε το βάθος γύρω από τη χωρική διάσταση της εισόδου προκειμένου να τοποθετηθεί το δεκτικό πεδίο. Για παράδειγμα εάν έπρεπε να κάνουμε ένα βήμα ως 1, τότε θα είχαμε ένα πολύ επικαλυπτόμενο δεκτικό πεδίο που παράγει εξαιρετικά μεγάλες ενεργοποιήσεις. Εναλλακτικά, ρυθμίζοντας το βήμα σε ένα μεγαλύτερο αριθμός θα μειώσει την επικάλυψη και θα παράγει έξοδο χαμηλότερων χωρικών διαστάσεων.

Η μηδενική επένδυση είναι η απλή διαδικασία πλήρωσης του περιγράμματος της εισόδου και είναι μια αποτελεσματική μέθοδος για να δοθεί περαιτέρω έλεγχος ως προς τη διάσταση του όγκου εξόδου.

Είναι σημαντικό να γίνει αντιληπτό ότι με τη χρήση αυτών των τεχνικών, θα αλλάξει η χωρική διάσταση της εξόδου των συνελκτικών στρωμάτων. Για τον υπολογισμό, μπορεί να χρησιμοποιηθεί ο ακόλουθος τύπος:

$$\frac{(V - R) + 2Z}{S + 1}$$

Οπού, το V αντιπροσωπεύει το μέγεθος της έντασης εισόδου (βάθος πλάτους ύψους), το R αντιπροσωπεύει το δεκτικό μέγεθος πεδίου, το Z είναι το ποσό του μηδενικού σερ γεμίσματος και του S που αναφέρεται στο βήμα.[19]

#### 4.7.4 Επίπεδο συγκέντρωσης

Τα συγκεντρωτικά στρώματα στοχεύουν στη σταδιακή μείωση της διασταλτικότητας της αναπαράστασης, και έτσι μειώνεται περαιτέρω ο αριθμός παραμέτρων και ο υπολογισμός πολυπλοκότητα του μοντέλου. Το στρώμα συγκέντρωσης λειτουργεί σε κάθε χάρτη ενεργοποίησης στην είσοδο και κλίμακες της διασταλτικότητά του χρησιμοποιώντας τη λειτουργία "MAX". Στα περισσότερα CNN, αυτά έρχονται στο εφαρμόζεται μορφή ομαδοποίησης στρώσεων με πυρήνες διαστάσεων 2 x 2 με ένα βήμα 2 κατά μήκος των χωρικών διαστάσεων της εισόδου. Αυτό κλιμακώνει ο χάρτης



ενεργοποίησης έως και 25% του αρχικού μεγέθους - διατηρώντας παράλληλα το βάθος όγκου στο κανονικό του μέγεθος.

Λόγω της καταστροφικής φύσης του στρώματος συγκέντρωσης, υπάρχουν μόνο δύο γενικά παρατηρήσιμες μέθοδοι μέγιστης συγκέντρωσης. Συνήθως, το βήμα και τα φίλτρα της συγκέντρωσης. Τα στρώματα ρυθμίζονται και τα δύο σε  $2 \times 2$ , τα οποία θα επιτρέψουν στο επίπεδο να επεκταθεί μέσω του σύνολο της χωρικής διάστασης της εισόδου. Επιπλέον αλληλεπικαλύπτεται.[18]

Ομαδοποίηση μπορεί να χρησιμοποιηθεί, όπου το βήμα ρυθμίζεται σε 2 με το μέγεθος του πυρήνα 3. Λόγω της καταστροφικής φύσης της συγκέντρωσης, έχει μέγεθος πυρήνα πάνω από συνήθως μειώνει σημαντικά την απόδοση του μοντέλου.

Είναι επίσης σημαντικό να κατανοήσουμε ότι πέρα από τη μέγιστη συγκέντρωση, οι αρχιτεκτονικές CNN μπορεί να περιέχουν γενική συγκέντρωση. Τα γενικά στρώματα συγκέντρωσης αποτελούνται από συγκέντρωση νευρώνες που είναι σε θέση να εκτελούν πολλές κοινές επεμβάσεις, συμπεριλαμβανομένων L1 / L2- κανονικοποίηση και μέσος όρος συγκέντρωσης.

#### *4.7.5 Πλήρως συνδεδεμένο στρώμα*

Το πλήρως συνδεδεμένο στρώμα περιέχει νευρώνες από τους οποίους συνδέονται άμεσα οι νευρώνες στα δύο γειτονικά στρώματα, χωρίς να συνδέονται με κανένα στρώμα μέσα τους. Αυτό είναι ανάλογο με τον τρόπο που οι νευρώνες διατάσσονται με παραδοσιακό τρόπο μορφές ANN. (Σχήμα 1)

### *4.8 Νευρωνικά δίκτυα ακτινικών συναρτήσεων*

#### *4.8.1 Λειτουργία Νευρωνικού δικτύου ακτινικών συναρτήσεων*

Τα δίκτυα λειτουργίας ακτινικής βάσης (RBFN) μπορεί σε κάποιο βαθμό να χαρακτηριστούν ως γραμμικά δίκτυα, αν και για τη λήψη καλών γενικεύσεων, μη γραμμική βελτιστοποίηση πρέπει να χρησιμοποιηθεί.[12]. Τα μη γραμμικά μοντέλα που χρησιμοποιούνται στις στατιστικές είναι πιθανό να εφαρμοστούν στις RBFN.

Τα RBFN μπορούν να θεωρηθούν τεχνητά νευρικά δίκτυα με ένα επίπεδο εισόδου, ένα κρυφό επίπεδο, ένα επίπεδο εξόδου και RBF στους κόμβους (οι κρυφές μονάδες) όπως φαίνεται στο παρακάτω σχήμα(Σχήμα ).Πρόκειται για τεχνητά νευρικά δίκτυα τύπου πολλαπλών στρώσεων feedforward. Τα RBFN συγκλίνουν γρήγορα στα βέλτιστα σημεία και έχουν πολύ υψηλές δυνατότητες τοποθέτησης. Αυτά τα γεγονότα κάνουν RBFNs πολύ κατάλληλα για έλεγχο του χάους[13].

#### *4.8.2 Προσδιορισμοί παραμέτρων*

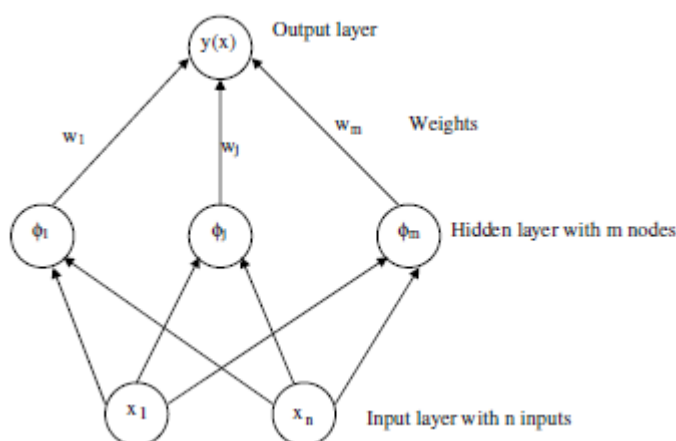
Ο σχεδιασμός ενός RBFN περιλαμβάνει την επιλογή ορισμένων παραμέτρων του RBFN, και συγκεκριμένα:

- Ο τύπος των συναρτήσεων βάσης ( $\phi$ ).
- Ο αριθμός των βασικών συναρτήσεων ( $M$ ).
- Τα κέντρα των λειτουργιών βάσης ( $\mu_i$ ).
- Τα πλάτη των βασικών συναρτήσεων ( $\sigma_i$ ).
- Τα βάρη των συναρτήσεων βάσης ( $w_j$ ).

Οι τύποι των συναρτήσεων βάσης και ο αριθμός τους καθορίζεται από την αρχή. Όλα τα άλλα καθορίζονται από την εκπαίδευση και ο προσδιορισμός πραγματοποιείται σε τέτοιο τρόπο ελαχιστοποίησης μιας κατάλληλης συνάρτησης κόστους.

Συνήθως, οι Gaussian ή, γενικότερα, άλλοι τύποι λειτουργούν με επιλέγεται η συμπαγής υποστήριξη. Ο αριθμός  $M$  των βασικών συναρτήσεων εξαρτάται από το μοντέλο επιλογή. Τα πλάτη των βασικών συναρτήσεων μπορούν να προσδιοριστούν έτσι ώστε κάθε βάση

Η συνάρτηση αποκτά τη δική της ατομική τιμή αντί για την παγκόσμια τιμή κοινή για όλους συναρτήσεις βάσης.



Εικόνα 30: Η γενική αρχιτεκτονική ενός RBFN[14].

#### 4.8.3 Πολυπλοκότητα μοντέλου

Τα τεχνητά δίκτυα νευρώνων, συμπεριλαμβανομένου του RBFN, δεν έχουν ως στόχο να βρουν την καλύτερη εφαρμογή για τα υπάρχοντα δεδομένα. Αντίθετα, ο σκοπός της εκπαίδευσης είναι να βρει μια γενική προσέγγιση, ένα μοντέλο πώς να βρείτε την καλύτερη εφαρμογή για δεδομένα εκτός του εκπαιδευτικού συνόλου. Αυτό το πρόβλημα του περιλαμβάνει μια αντιστάθμιση μεταξύ των δύο στοιχείων του σφάλματος γενίκευσης: κλίση και διακύμανση.

Η επιλογή ενός πολύ απλού μοντέλου θα οδηγήσει σε υψηλή μεροληψία, ενώ θα επιλέξει επίσης το σύνθετο μοντέλο μπορεί να οδηγήσει σε υψηλή διακύμανση και μπορεί έτσι να εξαρτάται πάρα πολύ από τη συγκεκριμένη επιλογή του σετ κατάρτισης. Ένα πολύ περίπλοκο μοντέλο μπορεί, ωστόσο, να έχει χαμηλό bias. Η εύρεση του καλύτερου συνδυασμού κλίσης και διακύμανσης είναι πρόβλημα εύρεσης του σωστού αριθμού παραμέτρων. Στην περίπτωση των RBFN, αυτό σημαίνει, μεταξύ άλλων, τον προσδιορισμό του βέλτιστου αριθμού κρυφών μονάδων. Ένας τρόπος για να γίνει αυτό θα ήταν να εκπαιδευτούν αρκετά δίκτυα το καθένα με διαφορετικούς αριθμούς κρυφών μονάδων.

Ένα πλεονέκτημα του RBFN είναι ότι αυτή η μέθοδος ωμής βίας δεν χρειάζεται να χρησιμοποιηθεί στον τοπικό χαρακτήρα των RBF και σε δύο στάδια εκπαίδευσης. Υπάρχουν κυρίως δύο τρόποι να το κάνουμε αυτό για RBFN.

Το πρώτο είναι να επιλέγει ένας αριθμός συναρτήσεων βάσης και μετά να προστεθεί ένας όρος κανονικοποίησης που είναι η συνάρτηση κόστους. Αυτό μπορεί να χρησιμοποιηθεί για να αυξηθεί π.χ. εξομάλυνση.

Μία εναλλακτική είναι να διακόψετε πρόωρα τη διαδικασία προπόνησης, ώστε να μην προσαρμόσετε το μοντέλο επίσης στενά στα δεδομένα εκπαίδευσης. Η δεύτερη μέθοδος αποτελείται είτε από την προσθήκη είτε τη διαγραφή κρυφών μονάδων κατά τη διάρκεια της εκπαιδευτικής διαδικασίας. Αυτό θα οδηγήσει σε λεγόμενους αλγόριθμους κλάδευσης ή ανάπτυξης.

#### 4.8.4 Εποπτευόμενη εκπαίδευση

Η ελαχιστοποίηση της συνάρτησης κόστους πραγματοποιείται με την επαναληπτική ενημέρωση των παραμέτρων μία φορά ανά δείγμα εκπαίδευσης. Επιλέγοντας τη συνάρτηση κόστους ως

$$E = \sum_n E_n \quad (18)$$

με

$$E_n = \frac{1}{2} \sum_k \{t_k^n - y_k(\bar{x}^n)\}^2 \quad (19)$$

όπου  $t_k^n$  είναι η τιμή-στόχος της μονάδας εξόδου  $k$  που προκύπτει από το φορέα εισόδου  $x^n$ .

Οι εξισώσεις ενημέρωσης μπορούν εύκολα να ληφθούν [15] ως

$$\begin{aligned} \Delta w_{ij} &= \eta_1 \{t_k^n - y_k(\bar{x}^n)\} \\ \Delta \bar{\mu}_j &= \eta_2 \phi_j(\bar{x}^n) \frac{\|\bar{x}^n - \bar{\mu}_j\|}{\sigma_j^2} \sum_k \{y_k(t_k^n) - \bar{x}^n\} w_{kj} \\ \Delta \sigma_j &= \eta_3 \phi_j(\bar{x}^n) \frac{\|\bar{x}^n - \bar{\mu}_j\|}{\sigma_j^3} \sum_k \{y_k(t_k^n) - \bar{x}^n\} w_{kj} \end{aligned} \quad (21)$$

όπου  $n_1, n_2$  και  $n_3$  είναι τα ποσοστά μάθησης.

#### 4.8.5 Εκπαίδευση δύο σταδίων

Και τα τρία αυτά σύνολα παραμέτρων μπορούν να ενημερώνονται ταυτόχρονα. Για μη στάσιμα περιβάλλοντα και σε απευθείας σύνδεση ρυθμίσεις, αυτή η διαδικασία μπορεί να είναι κατάλληλη. Ωστόσο, υπάρχουν καταστάσεις όπου η ενημέρωση της παραμέτρου μπορεί να διαχωριστεί σε μια διαδικασία δύο σταδίων.

Αυτό συμβαίνει π.χ. για στατική χαρτογράφηση. Η διαδικασία δύο σταδίων αποτελείται από:

- Προσδιορισμός των κέντρων των λειτουργιών βάσης ( $u_m$ ) και τα πλάτη του βασικές συναρτήσεις ( $j$ ).
- Στη συνέχεια, με τη χρήση των κέντρων των λειτουργιών βάσης ( $j$  m) και το πλάτος των συναρτήσεων βάσης ( $j$ ) που λαμβάνονται έτσι, τα βάρη της βάσης προσδιορίζονται οι συναρτήσεις ( $j$  w).

Οι πιο χρησιμοποιούμενοι αλγόριθμοι μάθησης ή κατάρτισης για τον προσδιορισμό των κεντρικών θέσεων είναι ή μέθοδος K-clustering, για τον προσδιορισμό του πλάτους του K-πλησιέστερου γείτονα μέθοδος και για τον προσδιορισμό του βάρους η μέθοδος κλίσης-καθόδου.[16]

Αποδεικνύεται ότι αυτή η διαδικασία δύο σταδίων οδηγεί σε μικρή απώλεια ποιότητας σε σύγκριση με προσδιορίζοντας ταυτόχρονα και τα τρία σύνολα παραμέτρων. Αντίθετα, μπορεί ακόμη και δώστε καλύτερες λύσεις, ειδικά για πεπερασμένα δεδομένα εκπαίδευσης και υπολογιστικούς πόρους.

Το πρώτο στάδιο της μάθησης είναι χωρίς επίβλεψη, δεδομένου ότι χρησιμοποιεί μόνο τις τιμές εισαγωγής  $\{x_n\}$ .

Στο επόμενο στάδιο της εποπτευόμενης μάθησης, δηλαδή της εκπαίδευσης όπου χρησιμοποιούνται οι πληροφορίες στόχου, μπορεί να εκτελεστεί.

Με τις βασικές παραμέτρους των RBF που καθορίζονται, αποκτάμε ένα δίκτυο ενός επιπέδου (ένα RBFN). Αυτό το δίκτυο έχει γραμμικές μονάδες εξόδου.

#### 4.8.6 Μη επιτηρούμενη εκπαίδευση

Η βασική ιδέα του RBFN είναι να εντοπίσει τα κέντρα σε περιοχές με υψηλά δεδομένα συγκέντρωση και να έχουν τα πλάτη που σχετίζονται με τη διάδοση των δεδομένων στο αντίστοιχες θέσεις και τις αποστάσεις από τις κοντινές συγκεντρώσεις δεδομένων. Αυτό μπορεί να ληφθούν με διάφορες μεθόδους:

- Επιλέγοντας τυχαία υποσύνολα.
- Χρήση αλγορίθμων ομαδοποίησης.
- Προσδιορισμός του πλάτους.

##### Επιλέγοντας τυχαία υποσύνολα

Από το σύνολο δεδομένων εκπαίδευσης, επιλέξτε τυχαία ένα υποσύνολο σημείων δεδομένων. Στη συνέχεια εντοπίζεται το Κέντρο RBF ( $j$  m) εκεί. Αυτό δίνει ένα σημείο εκκίνησης από το οποίο μπορεί κανείς να συντονιστεί περαιτέρω τις τοποθεσίες του κέντρου.

##### Χρήση αλγορίθμων ομαδοποίησης

Υπάρχουν διάφορες τεχνικές ομαδοποίησης που μπορούν να χρησιμοποιηθούν για τον εντοπισμό των κέντρων RBF(ιμ). Ο αλγόριθμος ομαδοποίησης K-σημαίνει περιγράφεται εδώ. Σημεία δεδομένων διαμερισμάτων  $N \times n$  σε συστάδες  $K$  ή υποσύνολα  $j$  S. Στη συνέχεια, ελαχιστοποιήστε το άθροισμα της συνάρτησης των τετραγώνων

$$J = \sum_{j=1}^K \sum_{x \in S_j} \|x^* - \bar{\mu}_j\|^2 \quad (23)$$

Οπού το δίνεται από

$$\bar{\mu}_j = \frac{1}{N_j} \sum_{x \in S_j} x^* \quad (24)$$

είναι ο γεωμετρικός μέσος όρος των σημείων δεδομένων στο υποσύνολο  $S_j$ .

#### Προσδιορισμός του πλάτους

Τα πλάτη των RBF μπορούν είτε να επιλεγούν ίσα για όλες τις μονάδες είτε διαφορετικά για το καθένα. Εάν επιλεγεί ένα κοινό, ή καθολικό, πλάτος, μπορεί να οριστεί ως πολλαπλάσιο του η μέση απόσταση μεταξύ των κέντρων βάσης. Η επιλογή καθορίζει τον βαθμό εξομάλυνση, με μικρό πλάτος που οδηγεί σε μικρότερες λειτουργίες εξομάλυνσης.

Διαφορετικά τα τοπικά πλάτη μπορούν να ληφθούν καθορίζοντας, για κάθε μονάδα, τον μέσο όρο απόσταση από τους  $L$  πλησιέστερους γείτονες και στη συνέχεια ρυθμίζοντας το πλάτος ίσο με το πολλαπλάσιο αυτό σημαίνει. Συνήθως επιλέγεται τιμή 1,5 έως 2 για αυτό το πολλαπλό.

#### Διαμόρφωση μήτρας

Ελαχιστοποιώντας τη συνάρτηση κόστους (Σχέση. 18), λαμβάνουμε τα βάρη ως (Σχέση. 10)

$$W^T = \Phi^+ T \quad (25)$$

όπου  $(T)_a = t_a^*$  και  $(\Phi)_a = \phi_j(x^*)$  Το F είναι η μήτρα σχεδιασμού και

$$\Phi^+ = (\Phi^T \Phi)^{-1} \Phi^T \quad (26)$$

δηλώνει το ψευδο-αντίστροφο του  $\Phi$ , στο οποίο

$$A = \Phi^T \Phi \quad (27)$$

ονομάζεται πίνακας διακύμανσης.

#### 4.8.7 Συγκρίσεις μεθόδων εκπαίδευσης

Στη μη εποπτευόμενη εκπαίδευση τα δεδομένα εξόδου παραμελούνται. Αυτό οδηγεί σε υπο-βελτιστοποίηση των παραμέτρων RBFN, τα κέντρα των λειτουργιών βάσης ( $j$  m) και τα πλάτη του βασικές συναρτήσεις ( $j$ ). Η εποπτευόμενη εκπαίδευση, όπου τα δεδομένα εξόδου δεν παραμελούνται, μπορεί να οδηγούν σε βέλτιστες τιμές των κέντρων και των πλάτους.

- Ωστόσο, η εποπτευόμενη εκπαίδευση έχει ορισμένα μειονεκτήματα
- Οι μη γραμμικές τεχνικές βελτιστοποίησης που χρησιμοποιούνται στην εποπτευόμενη εκπαίδευση είναι συχνά υπολογιστικό ακριβό.
- Η εποπτευόμενη εκπαίδευση μπορεί να οδηγήσει σε μη εντοπισμένους RBF.
- Με την εποπτευόμενη εκπαίδευση, το πλεονέκτημα της μάθησης δύο σταδίων του RBFN η πιθανότητα έχει χαθεί.

#### 4.8.8 Εφαρμογές RBFN

Τα RBFN χρησιμοποιούνται για πολλές διαφορετικές εφαρμογές. Χαρακτηριστικές εφαρμογές είναι: πρόβλεψη χρονοσειρών, ομιλία ταξινόμηση προτύπων, επεξεργασία εικόνας, ιατρική διάγνωση, μη γραμμικό σύστημα ταυτοποίησης, προσαρμοστική εξίσωση στα συστήματα επικοινωνίας, μη γραμμικό χαρακτηριστικό εξαγωγή, αναγνώριση χειρονομίας, αναγνώριση προσώπου, ταξινόμηση ΗΚΓ, χαρτογράφηση χειρονομιών σε ομιλία, αναγνώριση αντικειμένων, ανίχνευση ναρκοπεδίων και σχεδιασμό προσαρμοστικών ισοσταθμιστών είναι μερικά παραδείγματα εφαρμογών RBFN.

Αυτές οι εφαρμογές χρησιμοποιούν τις δυνατότητες προσέγγισης των RBFN, βάσει την τοπική δομή και αποτελεσματικοί αλγόριθμοι εκπαίδευσης. Ωστόσο, είναι συχνά δυνατό να χρησιμοποιηθεί ένα, MLP (Multi-Layered Perceptron).

Ανάλογα με το πρόβλημα, τα MLP μπορεί μερικές φορές να έχουν καλύτερη απόδοση, αλλά υπάρχουν πολλά παραδείγματα για τα οποία τα RBFN αποδίδουν καλύτερα ή τουλάχιστον αποκτούν παρόμοια επίπεδα απόδοσης. Το πλεονέκτημα ενός RBF μπορεί συχνά να είναι ότι είναι πιο γρήγορο να προπονηθείτε ενώ το μειονέκτημά του είναι οι υψηλότερες απαιτήσεις μνήμης.

## Βιβλιογραφία 4<sup>ου</sup> Κεφαλαίου

[1] Bradshaw, J.A., Carden, K.J., Riordan, D., 1991. Ecological —Applications Using a Novel Expert System Shell. Comp. Appl. Biosci. 7, 79–83.

- [2] Lippmann, R.P., 1987. An introduction to computing with neural nets. IEEE Accoust. Speech Signal Process. Mag., April: 4-22.
- [3] N. Murata, S. Yoshizawa, and S. Amari, —Learning curves, model selection and complexity of neural networks, in Advances in Neural Information Processing Systems 5, S. Jose Hanson, J. D. Cowan, and C. Lee Giles, ed. San Mateo, CA: Morgan Kaufmann, 1993, pp. 607-614
- [4] Cybenko, G. Approximation by superpositions of a sigmoidal function, Math. Control, Signal Syst. 2, pp.303-314, 1989.
- [5] Bryson, A.E., Ho, Y.C. Applied Optimal Control, Blaisdell, New York, 1969.7
- [6] Silva, F.M., Almeida, L.B. Acceleration techniques for the backpropagation algorithm in L.B. Almeida, C.J. Wellekens (eds.), Neural Networks, Springer, Berlin, pp.110–19, 1990.
- [7] Principe, J.C., Euliano, N.R., Lefebvre, W.C. Neural and Adaptive Systems. Fundamentals Through Simulations, John Wiley & Sons, Inc, 2000.
- [8] Negnevitsky, M. Artificial Intelligence: A Guide to Intelligent Systems, Addison Wesley, England, 2002.
- [9] Haykin, S. Neural Networks: A Comprehensive Foundation, Maxmillan, IEEE Press, 1994.
- [10] Rumelhart, D.E., Hinton, G.E., Williams, R.J. Learning representations by backpropagating errors, Nature 323, pp.533-536, 1986.
- [11] Popescu M.C., Hybrid neural network for prediction of process parameters in injection moulding, Annals of University of Petroşani, Electrical Engineering, Universities Publishing House, Petroşani, Vol. 9, pp.312-319, 2007.
- [12] Qun Zhao and Zheng Bao, “Radar Target Recognition Using a Radial Basis Function Neural Network”, Neural Networks, Vol. 9, No. 4, pp. 709-720, 1996.
- [13] T.J. Sejnowski and C.R. Rosenberg, ”Parallel networks that learn to pronounce English text”, Complex Systems, 1, pp. 145-168, 1987. 41
- [14] T. Stubbings and H. Hutter, “Classification of analytical images with radial basis function networks and forward selection”, Chemometrics and Intelligent Systems, No. 49, pp. 163-172, 1999.
- [15] B. Walczak and D.L. Massart, ”Local modelling with radial basis function networks”, Chemometrics and Intelligent Laboratory Systems, 50, pp. 179-198, 2000.
- [16]. Cireşan, D.C., Meier, U., Masci, J., Maria Gambardella, L., Schmidhuber, J.: Flexible, high performance convolutional neural networks for image classification. In: IJCAI Proceedings-International Joint Conference on Artificial Intelligence. vol. 22, p. 1237 (2011)

- [17]. Hinton, G.E., Srivastava, N., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.R.: Improving neural networks by preventing co-adaptation of feature detectors. arXiv preprint arXiv:1207.0580 (2012)
- [18]. Zeiler, M.D., Fergus, R.: Visualizing and understanding convolutional networks. In: Computer Vision–ECCV 2014, pp. 818–833. Springer (2014)
- [19]. Tivive, F.H.C., Bouzerdoun, A.: A new class of convolutional neural networks (siconnets) and their application of face detection. In: Neural Networks, 2003. Proceedings of the International Joint Conference on. vol. 3, pp. 2157–2162. IEEE (2003)



Page Intentionally Left Blank

## Κεφάλαιο 5<sup>ο</sup>

### Βαθιά Νευρωνικά Δίκτυα

Η βαθιά μάθηση (βαθιά δομημένη μάθηση, ιεραρχική μάθηση ή βαθιά μηχανική μάθηση) είναι ένας κλάδος της μηχανικής μάθησης με βάση ένα σύνολο αλγορίθμων που επιχειρούν σε επίπεδα επεξεργασίας με σύνθετες δομές ή αλλιώς αποτελείται από πολλαπλούς μη γραμμικούς μετασχηματισμούς [1-6]. Η βαθιά μάθηση είναι μέρος μιας ευρύτερης οικογένειας των μεθόδων μηχανικής μάθησης που βασίζονται σε μαθησιακές αναπαραστάσεις δεδομένων. Μια παρατήρηση (π.χ. Εικόνα) μπορεί να αναπαρασταθεί με πολλούς τρόπους, όπως ένα διάνυσμα τιμών έντασης ανά εικονοστοιχείο, ή περισσότερο αφηρημένος τρόπος ως ένα σύνολο άκρων, περιοχών ιδιαίτερου σχήματος κ.λπ. Ορισμένες παραστάσεις διευκολύνουν για την εκπαίδευση (π.χ. αναγνώριση προσώπου ή αναγνώριση έκφρασης προσώπου [7]). Οι δυνατότητες της βαθιάς μάθησης αντικαθιστά χειροποίητα χαρακτηριστικά με αποτελεσματικούς αλγόριθμους για μη εποπτευόμενη ή ημι-εποπτευόμενη εκμάθηση χαρακτηριστικών και εξαγωγή χαρακτηριστικών ιεραρχίας [8, 9].

Μελέτες σε αυτόν τον τομέα προσπαθούν να κάνουν καλύτερες αναπαραστάσεις και να δημιουργήσουν μοντέλα για να τα μάθουν αναπαραστάσεις από δεδομένα μεγάλης κλίμακας χωρίς σήμανση. Μερικές από τις παραστάσεις εμπνέονται από τις εξελίξεις στη νευροεπιστήμη και βασίζονται στην ερμηνεία της επεξεργασίας πληροφοριών και στα πρότυπα επικοινωνίας σε ένα νευρικό σύστημα, όπως η νευρική κωδικοποίηση που επιχειρεί να ορίσει τη σχέση μεταξύ διαφόρων ερεθισμάτων και συναφών νευρωνικών αποκρίσεων στον εγκέφαλο [10].

Διάφορες αρχιτεκτονικές βαθιάς μάθησης, όπως τα βαθιά νευρωνικά δίκτυα, τα συνεχή βαθιά νευρωνικά δίκτυα, βαθιά δίκτυα πεποιθήσεων και επαναλαμβανόμενα νευρωνικά δίκτυα έχουν εφαρμοστεί σε πεδία όπως όραση υπολογιστή, αυτόματη αναγνώριση ομιλίας, επεξεργασία φυσικής γλώσσας, αναγνώριση ήχου και βιοπληροφορική όπου έχουν αποδειχθεί ότι παράγουν αποτελέσματα τελευταίας τεχνολογίας σε διάφορες εργασίες.

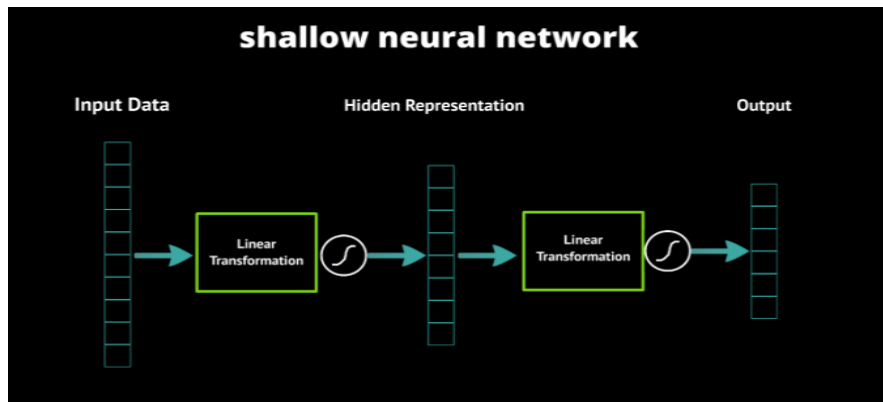
Εναλλακτικά, η βαθιά μάθηση έχει χαρακτηριστεί ως λέξη-κλειδί, ή μετονομασία νευρωνικά δίκτυα [11, 12]. Η βαθιά μάθηση θα μπορούσε να χαρακτηριστεί ως μια κατηγορία αλγορίθμων μηχανικής μάθησης που χρησιμοποιούν πολλά επίπεδα μη γραμμικών μονάδων επεξεργασίας για εξαγωγή χαρακτηριστικών. Κάθε διαδοχικό επίπεδο χρησιμοποιεί την έξοδο από το προηγούμενο επίπεδο ως είσοδο.

Οι αλγόριθμοι μπορεί να επιβλέπονται ή να μην ελέγχονται και οι εφαρμογές περιλαμβάνουν ανάλυση προτύπων (χωρίς επίβλεψη) και ταξινόμηση (υπό επίβλεψη). Βασίζονται στην (χωρίς επίβλεψη) εκμάθηση πολλαπλών επιπέδων χαρακτηριστικών ή αναπαραστάσεων του δεδομένα. Τα χαρακτηριστικά υψηλότερου επιπέδου προέρχονται από χαρακτηριστικά χαμηλότερου επιπέδου για να σχηματίσουν μια ιεραρχική αναπαράσταση. Είναι μέρος του ευρύτερου πεδίου μηχανικής μάθησης των μαθησιακών αναπαραστάσεων δεδομένων.

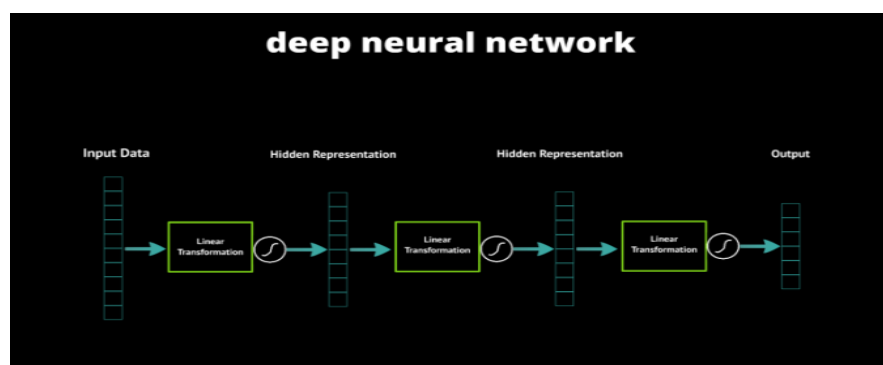
Αυτοί οι χαρακτηρισμοί έχουν κοινά (1) πολλαπλά επίπεδα μη γραμμικών μονάδων επεξεργασίας και (2) την εποπτευόμενη ή χωρίς επίβλεψη εκμάθηση των αναπαραστάσεων χαρακτηριστικών σε κάθε επίπεδο, με τα επίπεδα σχηματίζοντας μια ιεραρχία από χαρακτηριστικά χαμηλού επιπέδου έως υψηλού επιπέδου [1]. Η σύνθεση ενός στρώματος από μη γραμμικές μονάδες επεξεργασίας που χρησιμοποιούνται σε έναν αλγόριθμο βαθιάς μάθησης εξαρτάται από το πρόβλημα που πρέπει να λυθεί.

Τα επίπεδα που έχουν χρησιμοποιηθεί στη βαθιά μάθηση περιλαμβάνουν κρυμμένα στρώματα ενός τεχνητού νευρικού δικτύου και σύνολα περίπλοκων προτεινόμενων τύπων [2]. Μπορούν επίσης να περιλαμβάνουν λανθάνουσες μεταβλητές οργανωμένο επίπεδο σε βαθιά γενετικά μοντέλα, όπως οι κόμβοι στα Deep Belief Networks [13] και Βαθιά μηχανήματα Boltzmann [14, 15].

Τα νευρικά δίκτυα λαμβάνουν τις αναπαραστάσεις τους από τη χρήση επιπέδων μάθησης. Οι πρωτεύοντες εγκέφαλοι κάνουν ένα παρόμοιο πράγμα στον οπτικό φλοιό, οπότε η ελπίδα ήταν ότι η χρήση περισσότερων στρωμάτων σε ένα νευρικό δίκτυο θα μπορούσε αφήσει να μάθει καλύτερα μοντέλα. Ωστόσο, μελέτες έχουν δείξει ότι οι εσωτερικές αναπαραστάσεις μεταξύ αυτών των μοντέλων δεν μπόρεσαν να λειτουργήσουν, αλλά ωστόσο, τα επιτυχημένα μοντέλα συνειδητοποιήθηκαν ότι θα κατασκευάσουν ένα ρηχό δίκτυο, ένα με ένα μόνο επίπεδο αναπαράστασης δεδομένων [16]. Μαθαίνοντας σε βάθος νευρωνικό δίκτυο, ένα με περισσότερα από ένα επίπεδα αναπαράστασης δεδομένων, απλά δεν λειτουργήσε. Στην πραγματικότητα, η βαθιά μάθηση υπήρχε για όσο διάστημα υπήρχαν νευρωνικά δίκτυα, όπως απεικονίζεται στα σχήματα στις Εικόνες 1 και 2.



Εικόνα 31: Νευρωνικό δίκτυο ενός επιπέδου



Οι αλγόριθμοι βαθιάς μάθησης έρχονται σε αντίθεση με τους αλγόριθμους ρηχής μάθησης συγκριτικά με τον αριθμό παραμετροποιημένων μετασχηματισμών που συναντά ένα σήμα καθώς μεταδίδεται από το επίπεδο εισόδου στην έξοδο επίπεδο, όπου ένας παραμετροποιημένος μετασχηματισμός είναι μια μονάδα επεξεργασίας που έχει εκπαιδευμένες παραμέτρους, όπως βάρη και κατώτατα όρια [4]. Μια αλυσίδα μετασχηματισμών από είσοδο σε έξοδο είναι μια πίστωση διαδρομή ανάθεσης (CAP). Οι CAP περιγράφουν δυνητικά αιτιώδεις συνδέσεις μεταξύ εισόδου και εξόδου και μπορεί να διαφέρουν σε μήκος. Για ένα νευρικό δίκτυο τροφοδοσίας, το βάθος των CAPs, και επομένως το βάθος του δικτύου, είναι ο αριθμός των κρυφών επιπέδων συν ένα (το επίπεδο εξόδου είναι επίσης παράμετρος). Για επαναλαμβανόμενα νευρωνικά δίκτυα, στα οποία ένα σήμα μπορεί να διαδοθεί μέσω ενός επιπέδου περισσότερες από μία φορές, η CAP είναι δυνητικά απεριόριστη σε μήκος. Δεν υπάρχει καθολικά συμφωνημένο κατώφλι του βάθους που διαιρεί τη ρηχή μάθηση από τη βαθιά μάθηση, αλλά οι περισσότεροι ερευνητές στον τομέα συμφωνούν ότι η βαθιά μάθηση έχει πολλαπλά μη γραμμικά επίπεδα ( $CAP > 2$ ) και [4] θεωρεί ότι η  $CAP > 10$  χαρακτηρίζεται ως πολύ βαθιά μάθηση

### 5.1 Θεμελιώδεις έννοιες

Οι αλγόριθμοι βαθιάς μάθησης βασίζονται σε κατανεμημένες αναπαραστάσεις. Η υποκείμενη υπόθεση πίσω από τις κατανεμημένες αναπαραστάσεις είναι ότι τα παρατηρούμενα δεδομένα δημιουργούνται από τις αλληλεπιδράσεις των παραγόντων οργανωμένοι σε στρώματα. Η βαθιά μάθηση προσθέτει την υπόθεση ότι αντιστοιχούν αυτά τα επίπεδα παραγόντων επίπεδα άντλησης ή σύνθεσης. Μπορούν να χρησιμοποιηθούν ποικίλοι αριθμοί επιπέδων και μεγεθών επιπέδων παρέχουν διαφορετικές ποσότητες άντλησης [3]. Η βαθιά μάθηση εκμεταλλεύεται αυτήν την ιδέα ιεραρχικής επεξηγηματικοί παράγοντες όπου υψηλότερο επίπεδο, πιο αφηρημένες έννοιες μαθαίνονται από το κατώτερο επίπεδο αυτοί. Αυτές οι αρχιτεκτονικές κατασκευάζονται συχνά με μια άπληστη μέθοδο στρώματος προς στρώμα. Βαθιά μάθηση βοηθά στην αποδέσμευση αυτών των αφαιρέσεων και διαλέγει ποια χαρακτηριστικά είναι χρήσιμα για την εκμάθηση [3].

Για εποπτευόμενες μαθησιακές εργασίες, οι μέθοδοι βαθιάς μάθησης αποφεύγουν τη μηχανική χαρακτηριστικών, μεταφράζοντας τα δεδομένα σε συμπαγείς ενδιάμεσες αναπαραστάσεις που μοιάζουν με κύρια συστατικά στοιχεία και παράγονται σε στρώσεις δομών που αφαιρούν τον πλεονασμό στην αναπαράσταση [1]. Πολλοί αλγόριθμοι βαθιάς μάθησης εφαρμόζεται σε μη εποπτευόμενες μαθησιακές εργασίες. Αυτό είναι ένα σημαντικό όφελος επειδή τα δεδομένα χωρίς ετικέτα είναι συνήθως πιο άφθονα από τα επισημασμένα δεδομένα.

Ένα παράδειγμα βαθιάς δομής που μπορεί να εκπαιδευτεί σε ένα είναι ο μη εποπτευόμενος τρόπος είναι ένα δίκτυο βαθιάς πεποίθησης [3]. Τα βαθιά νευρωνικά δίκτυα γενικά ερμηνεύονται στους ορούς:

Θεώρημα καθολικής προσέγγισης [20-24] ή Πιθανοτική συναγωγή [1-4, 13, 25]. Το θεώρημα προσέγγισης της καθολικής αφορά την ικανότητα των νευρωνικών δικτύων

τροφοδοσίας σε ένα μόνο κρυφό στρώμα πεπερασμένου μεγέθους για προσέγγιση των συνεχών λειτουργιών [20-24]. Το 1989, η πρώτη απόδειξη δημοσιεύθηκε από το Cybenko [21] για τις λειτουργίες ενεργοποίησης του σιγμοειδούς και γενικεύτηκε σε feed-forward πολυεπίπεδη αρχιτεκτονική το 1991 από τον Hornik [22]. Η προσέγγιση θα μπορούσε να είναι μαθηματικά ως εξής:

Ας είναι μια συνεχής, μη-σταθερή, οριακή και μονοτονικά αυξανόμενη λειτουργία. Δηλώνει τον υπερ-κύβο μονάδας διαστάσεων  $[0,1]$ . Δηλώνεται ο χώρος των συνεχών λειτουργιών στο  $\text{Im } C(\text{Im})$ . Τότε, δεδομένης οποιασδήποτε συνάρτησης, υπάρχει ένας ακέραιος  $N$  και πραγματικές σταθερές, όπου  $i = 1, \dots, N$  έτσι ώστε να μπορεί να οριστεί:

$$F(x) = \sum_{i=1}^N v_i \varphi(W_i^T x + b_i)$$

είναι μια κατά προσέγγιση πραγματοποίηση της συνάρτησης  $f$  όπου το  $f$  είναι ανεξάρτητο από  $\varphi$  αυτό είναι

$$F(x) - f(x) < \varepsilon$$

για όλα  $x$ . Με άλλα λόγια, οι συναρτήσεις της φόρμας  $F(x)$  είναι πυκνές  $C(\text{Im})$

Το  $N$  αντιπροσωπεύει τον αριθμό των μονάδων ή του δείγματος, το  $W$  είναι η συμμετρική μήτρα αλληλεπίδρασης βάρους όρος, το  $b$  είναι ένας όρος μεροληψίας και το  $T$  είναι ο στόχος ή ο μεταθετικός φορέας που αντιπροσωπεύει τη μήτρα στήλης, όταν το  $\text{Im}$  αντικαθίσταται με οποιοδήποτε συμπαγές υποσύνολο  $\mathbb{R}^m$ . Η πιθανολογική ερμηνεία [25], η οποία οδήγησε η εισαγωγή της εγκατάλειψης ως κανονικοποιητή σε νευρωνικά δίκτυα [26] προέρχεται από το πεδίο του μηχανική εκμάθησης. Διαθέτει συμπεράσματα, καθώς και τις έννοιες βελτιστοποίησης της εκπαίδευσης και των δοκιμών, σχετίζονται με την τοποθέτηση και τη γενίκευση αντίστοιχα. Πιο συγκεκριμένα, η Πιθανοτική Ερμηνεία θεωρεί τη μη γραμμικότητα ενεργοποίησης ως αθροιστική συνάρτηση διανομής [25] όπως αναφέρεται στην παρακάτω εξίσωση.

$$P(H|E) = \frac{P(H|E)P(H)}{P(E)}$$

Η πιθανότητα ερμηνείας εισήχθη και διαδόθηκε από τον Geoff Hinton, Yoshua Bengio, Yann LeCun και Juergen Schmidhuber [27].

Το 2006, τρεις ξεχωριστές ομάδες ανέπτυξαν τρόπους για να ξεπεράσουν τις δυσκολίες που πολλοί στο κόσμο της μηχανικής μάθησης συναντήσαν κατά την προσπάθεια να εκπαιδεύσουν βαθιά νευρωνικά δίκτυα. Οι υπεύθυνοι αυτών των τριών ομάδων είναι οι πατέρες της εποχής της βαθιάς μάθησης. Αυτό δεν είναι καθόλου υπερβολικό. αυτά τα άτομα εγκαινιάστηκαν σε μια νέα εποχή. Η δουλειά τους έδωσε νέα ζωή σε νευρωνικά δίκτυα όταν πολλοί είχαν εγκαταλείψει τις υπηρεσίες κοινής ωφέλειας [28]. Λίγα χρόνια μετά, ο Geoff Hinton προσφέρεται για δουλειά από την Google, ο Yann H LeCun γίνεται διευθυντής της AI Research στο Facebook, ενώ η Yoshua Bengio αναλαμβάνει θέση ως ερευνητής για την Τεχνητή Νοημοσύνη στο Πανεπιστήμιο του Μόντρεαλ, χρηματοδοτούμενη εν μέρει από την εταιρεία βιντεοπαιχνιδιών Ubisoft [29]. Οι τροχιές τους δείχνουν ότι η δουλειά τους αποτέλεσε ένα σοβαρό εγχείρημα. Πριν από την ερευνά τους, τα πρώτα στρώματα σε ένα βαθύ δίκτυο απλά δεν μαθαίνουν χρήσιμες αναπαραστάσεις για τα δεδομένα. Σε πολλές περιπτώσεις δεν μαθαίνουν

καθόλου. Αντ' αυτού έμεναν κοντά στην τυχαία εκκίνηση λόγω της φύσης του αλγορίθμου εκπαίδευσης για νευρωνικά δίκτυα. Χρησιμοποιώντας διαφορετικές τεχνικές, καθεμία από αυτές τις τρεις ομάδες μπόρεσε να κάνει αυτά τα πρώτα στρώματα να μάθουν χρήσιμες αναπαραστάσεις, οι οποίες οδήγησαν σε πολύ πιο ισχυρά νευρωνικά δίκτυα.

## 5.2 Εκπαίδευση νευρωνικού δικτύου

Μία από τις θεωρητικές ιδιότητες των νευρωνικών δικτύων που έχει κρατήσει τους ερευνητές είναι ότι πρέπει να είναι διδακτικά. Είναι απλό να γίνει αντιληπτό ότι σε μικρή κλίμακα πώς μπορείτε να παράσχετε μια σειρά παραδειγμάτων εισόδων και αναμενόμενων εξόδων και να περάσετε από μια μηχανική διαδικασία πάρτε τα βάρη από αρχικές τυχαίες τιμές σε προοδευτικά καλύτερους αριθμούς που παράγουν περισσότερα ακριβείς προβλέψεις. Το πρόβλημα ήταν πάντα πώς να κάνετε το ίδιο πράγμα σε πολύ πιο περίπλοκα προβλήματα όπως η αναγνώριση ομιλίας ή ο υπολογιστής όραση με πολύ μεγαλύτερο αριθμό βαρών. Υπήρξε μια πραγματική ανακάλυψη το 2012 που δημοσιεύθηκε από μια εφημερίδα ImageNet [18] πυροδοτώντας την τρέχουσα αναγέννηση στα νευρικά δίκτυα.

Ο Χίντιν συγκέντρωσε μια σειρά από διαφορετικούς τρόπους επιτάχυνσης της μαθησιακής διαδικασίας, συμπεριλαμβανομένων των συνελκτικών δικτύων, της έξυπνης χρήσης GPU και ορισμένων νέων μαθηματικών κόλπων όπως ReLU and εγκατάλειψη, και έδειξε ότι σε λίγες εβδομάδες θα μπορούσαν να εκπαιδεύσουν ένα πολύ περίπλοκο δίκτυο σε ένα επίπεδο που θα μπορούσε να ξεπεράσει τις συμβατικές προσεγγίσεις για την όραση του υπολογιστή [1, 18]. Δεν πρόκειται για παρέκκλιση. παρόμοιες προσεγγίσεις που έχουν χρησιμοποιηθεί με μεγάλη επιτυχία στη φυσική επεξεργασία γλώσσας και αναγνώριση ομιλίας. Αυτή είναι η καρδιά της βαθιάς μάθησης – οι νέες τεχνικές που έχουν ανακαλυφθεί που μας επιτρέπουν να χτίζουμε και να εκπαιδεύουμε νευρωνικά δίκτυα για να χειριστούμε προηγούμενως άλυτα προβλήματα. Με την πλειονότητα της μηχανικής μάθησης, το δύσκολο μέρος είναι να εντοπίσει τα χαρακτηριστικά στα πρωτογενή δεδομένα εισόδου, για παράδειγμα SIFT ή SURF σε εικόνες [19].

Η βαθιά μάθηση αφαιρεί το χειροκίνητο βήμα, αντί να στηριχτεί στη διαδικασία προπόνησης για να ανακαλυφθούν τα πιο χρήσιμα μοτίβα σε ολόκληρα τα δείγματα εισόδου. Πρέπει ακόμη να κάνετε επιλογές σχετικά με την εσωτερική διάταξη των δικτύων πριν ξεκινήσει την εκπαίδευση, αλλά η αυτόματη ανακάλυψη χαρακτηριστικών καθιστά τη ζωή πολύ πιο εύκολη. Με άλλους τρόπους, επίσης, νευρωνικά τα δίκτυα είναι πιο γενικά από τις περισσότερες άλλες τεχνικές μηχανικής μάθησης. Οι ίδιες τεχνικές για την αρχιτεκτονική και την κατάρτιση δικτύων είναι χρήσιμες σε όλα τα είδη φυσικών δεδομένων, από ήχος σε σεισμικούς αισθητήρες ή φυσική γλώσσα. Καμία άλλη προσέγγιση δεν είναι σχεδόν τόσο ευέλικτη. Η Βαθιά μάθηση λειτουργεί πολύ καλά, και αν ασχοληθεί ποτέ με ακατάστατα δεδομένα από τον πραγματικό κόσμο, θα είναι βασικό στοιχείο στην εργαλειοθήκη τα επόμενα χρόνια



### 5.3 Δομές DNN

Υπάρχουν τεράστιοι αριθμοί παραλλαγών βαθιών αρχιτεκτονικών. Τα περισσότερα από αυτά είναι κλάδοι μερικών πρωτότυπων της γονικής αρχιτεκτονικής. Δεν είναι πάντα δυνατή η σύγκριση της απόδοσης πολλαπλών αρχιτεκτονικών όλες μαζί, επειδή δεν αξιολογούνται όλες βάσει των ίδιων συνόλων δεδομένων. Η Βαθιά μάθηση είναι ένα ταχέως αναπτυσσόμενο πεδίο, και νέες αρχιτεκτονικές, παραλλαγές ή αλγόριθμοι εμφανίζονται κάθε λίγες εβδομάδες. Κάποιες από τις βασικότερες θα αναλυθούν παρακάτω.

#### 5.3.1 Δίκτυα *Deep Belief*

Ένα δίκτυο βαθιάς πεποίθησης (DBN) είναι ένα πιθανό, δημιουργικό μοντέλο που αποτελείται από πολλαπλάσια στρώματα κρυφών μονάδων. Μπορεί να θεωρηθεί μια σύνθεση απλών μαθησιακών ενοτήτων που συνθέτουν κάθε στρώμα [42]. Ένα DBN μπορεί να χρησιμοποιηθεί για να προ-εκπαιδεύσει γενικά ένα DNN χρησιμοποιώντας το εκπαιδευμένο DBN βάρη ως τα αρχικά βάρη DNN. Παράλληλα, οι αλγόριθμοι διάδοσης ή άλλοι διακρίσεις μπορούν τότε να εφαρμοστεί για τη ρύθμιση αυτών των βαρών. Αυτό είναι ιδιαίτερα χρήσιμο όταν είναι περιορισμένα τα εκπαιδευτικά δεδομένα διαθέσιμα, επειδή τα κακώς αρχικοποιημένα βάρη μπορούν να παρεμποδίσουν σημαντικά τα μαθημένα μοντέλα να εκτελεστούν. Αυτά τα προ-εκπαιδευμένα βάρη βρίσκονται σε μια περιοχή του χώρου βάρους που είναι πιο κοντά στο βέλτιστα βάρη από ότι επιλέγονται τυχαία αρχικά βάρη. Αυτό επιτρέπει και τα δύο βελτιωμένα μοντέλα και ταχύτερη σύγκλιση της φάσης βελτίωσης [43].

Ένα DBN μπορεί να εκπαιδευτεί αποτελεσματικά με τρόπο χωρίς επίβλεψη, στρώμα προς στρώμα, όπου τα στρώματα είναι συνήθως κατασκευασμένα από μηχανές περιορισμένης πρόσβασης Boltzmann (RBM). Ένα RBM είναι ένα μη κατευθυνόμενο, μοντέλο βασισμένο σε ενέργεια με ένα "ορατό" επίπεδο εισόδου και ένα κρυφό στρώμα, με συνδέσεις μεταξύ των στρωμάτων αλλά όχι εντός των στρωμάτων.

#### 5.3.2 Συνελικτικά Νευρωνικά Δίκτυα (CNN)

Το CNN αποτελείται από ένα ή περισσότερα συνελικτικά στρώματα με πλήρως συνδεδεμένα στρώματα (ταιριάζουν με εκείνα των τυπικών τεχνητών νευρικών δικτύων) στην κορυφή. Χρησιμοποιεί επίσης δεμένα βάρη και ομαδοποίηση στρώματα. Αυτή η αρχιτεκτονική επιτρέπει στα CNN να επωφεληθούν από τη δισδιάστατη δομή δεδομένων εισόδου. Σε σύγκριση με άλλες βαθιές αρχιτεκτονικές, αρχίζουν να εμφανίζονται συνελικτικά νευρωνικά δίκτυα ανώτερα αποτελέσματα τόσο σε εφαρμογές εικόνας όσο και σε ομιλία [47]. Μπορούν επίσης να εκπαιδευτούν με στάνταρ οπίσθια διάδοση. Τα CNN είναι πιο εύκολο να εκπαιδευτούν από άλλα κανονικά, βαθιά, νευρωνικά δίκτυα τροφοδοσίας και έχουν πολύ λιγότερες παραμέτρους για εκτίμηση, καθιστώντας τις μια εξαιρετικά ελκυστική αρχιτεκτονική στη χρήση [48]. Παραδείγματα εφαρμογών στο Computer Vision περιλαμβάνουν το DeepDream [49].

#### 5.3.3 Συνδυαστικά Δίκτυα *Convolutional Belief*

Ένα πρόσφατο επίτευγμα στη βαθιά μάθηση είναι η χρήση των Convolutional Deep Belief Networks (CDBN). Τα CDBN έχουν δομή πολύ παρόμοια με τα συνελικτικά νευρικά δίκτυα και είναι εκπαιδευμένα παρόμοια με τα δίκτυα βαθιάς πεποίθησης. Επομένως, εκμεταλλεύονται τη δισδιάστατη δομή των εικόνων, όπως κάνουν τα CNN,

και χρησιμοποιήστε την προ-εκπαίδευση όπως τα δίκτυα βαθιάς πεποίθησης. Παρέχουν μια γενική δομή που μπορεί να χρησιμοποιηθεί σε πολλές εργασίες επεξεργασίας εικόνας και σήματος.

## 5.4 Προβλήματα με Deep Neural Networks

Όπως και με τα ANN, πολλά προβλήματα μπορούν να προκύψουν με τα DNNs εάν είναι αφελώς εκπαιδευμένα. Δύο κοινά ζητήματα είναι η ικανότητα υπολογισμού και ο αντίστοιχος απαιτούμενος χρόνος. Τα DNNs είναι επιρρεπή σε υπερφόρτωση λόγω των προστιθέμενων επιπέδων αφαίρεσης, που τους επιτρέπουν να μοντελοποιήσουν σπάνιες εξαρτήσεις στα δεδομένα εκπαίδευσης.

Οι μέθοδοι κανονικοποίησης όπως η μείωση του βάρους (κανονικοποίηση l2) ή η αραιότητα (1-κανονικοποίηση) μπορούν να εφαρμοστούν κατά τη διάρκεια της προπόνησης για να βοηθήσει στην καταπολέμηση της υπερβολικού overfitting [35]. Μια πιο πρόσφατη μέθοδος τακτοποίησης εφαρμόζεται στα DNNs και είναι η κανονικοποίηση της εγκατάλειψης. Κατά την εγκατάλειψη, ορισμένοι αριθμοί μονάδων παραλείπονται τυχαία από τα κρυμμένα στρώματα κατά τη διάρκεια της προπόνησης. Αυτό βοηθά στη διακοπή των σπάνιων εξαρτήσεων που μπορεί να εμφανιστούν τα δεδομένα εκπαίδευσης [36].

Η κυρίαρχη μέθοδος για την εκπαίδευση αυτών των δομών ήταν η εκπαίδευση διόρθωσης σφαλμάτων (τέτοια ως οπίσθια διάδοση με κλίση κατάβασης) λόγω της ευκολίας εφαρμογής του και της τάσης του, σε σύγκριση σε καλύτερα τοπικά βέλτιστα από άλλες μεθόδους εκπαίδευσης. Ωστόσο, αυτές οι μέθοδοι μπορεί να είναι υπολογιστικά ακριβές, ειδικά για DNN. Υπάρχουν πολλές παράμετροι εκπαίδευσης, όπως το μέγεθος (αριθμός επιπέδων και αριθμός μονάδων ανά στρώση), το ποσοστό μάθησης και τα αρχικά βάρη. Ουσιαστικά, σαρώνει τον χώρο παραμέτρων για βέλτιστες παραμέτρους που δεν είναι εφικτό λόγω του κόστους στο χρόνο και των υπολογιστικών πόρων. Διάφορα «κόλπα» όπως η χρήση mini-batching (υπολογισμός της κλίσης σε πολλά παραδείγματα εκπαίδευσης ταυτόχρονα παρά μεμονωμένα παραδείγματα) [37] έχει αποδειχθεί ότι επιταχύνει τον υπολογισμό. Η μεγάλη απόδοση επεξεργασίας του. Οι Μονάδες Επεξεργασίας Γραφικών (GPUs) έχουν δημιουργήσει σημαντικές επιταχύνσεις στην εκπαίδευση, λόγω του πίνακα και οι υπολογισμοί διανυσμάτων απαιτούνται κατάλληλοι για GPU [4]. Ριζοσπαστικές εναλλακτικές λύσεις αντί της αναπαράστασης όπως Extreme Learning Machines [38], "No-prop" network [39], επαναλαμβανόμενη εκπαίδευση δίκτυα χωρίς backtracking [40], και Weightlessneural δίκτυα [41] κερδίζουν την προσοχή

## 5.5 Εφαρμογές βαθιάς μάθησης

Υπήρξαν αρκετές μελέτες που αποδεικνύουν την αποτελεσματικότητα των μεθόδων βαθιάς μάθησης σε ένα ποικίλο τομέα εφαρμογών. Εκτός από το Μικτό Εθνικό Ινστιτούτο Προτύπων και Τεχνολογία (MNIST) πρόκληση γραφής [52], υπάρχουν εφαρμογές στην ανίχνευση προσώπου [53, 54], αναγνώριση και ανίχνευση ομιλίας [55], γενική αναγνώριση αντικειμένων [56], επεξεργασία φυσικής γλώσσας [57] και ρομποτική.

Είναι αλήθεια ότι η πραγματικότητα του πολλαπλασιασμού των δεδομένων και της αφθονίας των πολυτονικών αισθητηριακών πληροφοριών μια πρόκληση και ένα



επαναλαμβανόμενο θέμα σε πολλές στρατιωτικές και μη στρατιωτικές εφαρμογές, όπως εξελιγμένα συστήματα παρακολούθησης. Κατά συνέπεια, το ενδιαφέρον για τη βαθιά μηχανική μάθηση δεν ήταν περιορίζεται στην ακαδημαϊκή έρευνα. Πρόσφατα, ο Οργανισμός Προηγμένων Ερευνητικών Έργων Άμυνας (DARPA) ανακοίνωσε ένα ερευνητικό πρόγραμμα που επικεντρώνεται αποκλειστικά στη βαθιά μάθηση [6, 58]. Πολλά ιδιωτικοί οργανισμοί έχουν εστιάσει την προσοχή τους στην εμπορευματοποίηση τεχνολογιών βαθιάς μάθησης με εφαρμογές σε ευρείς τομείς.

Οι Lenz et al [59] παρουσίασαν πρόσφατα ένα σύστημα για την ανίχνευση ρομποτικών πιέσεων από τη χρήση δεδομένων RGB-D μια προσέγγιση βαθιάς μάθησης που έχει πολλά πλεονεκτήματα έναντι των τρεχουσών προηγμένων μεθόδων. Η δικιά τους προσέγγιση απέδειξε αρχικά ότι η χρήση βαθιάς μάθησης σας επιτρέπει να αποφύγετε τη χρήση μηχανικής χειρός χαρακτηριστικά, αλλά αντ' αυτού τα μάθε. Δεύτερον, τα αποτελέσματά τους έδειξαν ότι οι μέθοδοι βαθιάς μάθησης ξεπέρασε σημαντικά ακόμη και τα καλά σχεδιασμένα χαρακτηριστικά χειροποίητα από προηγούμενες εργασίες.

## Βιβλιογραφία 5<sup>ου</sup> Κεφαλαίου

- [1] D. Li and D. Yu, "Deep Learning: Methods and Applications," Foundations and Trends in Signal Processing, Now Publishers, 2014.
- [2] Y. Bengio, "Learning deep architectures for AI," Foundations and trends® in Machine Learning, vol. 2, pp. 1-127, 2009.
- [3] Y. Bengio, A. Courville, and P. Vincent, "Representation learning: A review and new perspectives," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 35, pp. 1798-1828, 2013.
- [4] J. Schmidhuber, "Deep learning in neural networks: An overview," Neural Networks, vol. 61, pp. 85-117, 2015.
- [5] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, pp. 436-444, 2015.
- [6] I. Arel, D. C. Rose, and T. P. Karnowski, "Deep machine learning-a new frontier in artificial intelligence research [research frontier]," Computational Intelligence Magazine, IEEE, vol. 5, pp. 13-18, 2010.
- [7] P. O. Glauner, "Deep Convolutional Neural Networks for Smile Recognition," arXiv preprint arXiv:1508.06535, 2015.
- [8] H. A. Song and S.-Y. Lee, "Hierarchical Representation Using NMF," in Neural Information Processing, 2013, pp. 466-473.
- [9] M. Xiong, J. Chen, Z. Wang, C. Liang, Q. Zheng, Z. Han, et al., "Deep Feature Representation via Multiple Stack Auto-Encoders," in Advances in Multimedia Information Processing--PCM 2015, ed: Springer, 2015, pp. 275-284.
- [10] B. A. Olshausen, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," Nature, vol. 381, pp. 607-609, 1996.
- [11] R. Collobert, "Deep learning for efficient discriminative parsing," in International Conference on Artificial Intelligence and Statistics, 2011.
- [12] L. Gomes, "Machine-learning maestro michael jordan on the delusions of big data and other huge engineering efforts," IEEE Spectrum, Oct, vol. 20, 2014.
- [13] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," Neural computation, vol. 18, pp. 1527-1554, 2006.
- [14] R. Salakhutdinov and G. Hinton, "An efficient learning procedure for deep Boltzmann machines," Neural computation, vol. 24, pp. 1967-2006, 2012.
- [15] D. H. Staelin and C. H. Staelin, "Models for Neural Spike Computation and Cognition," ed: CreateSpace.[Links], 2011.
- [16] N. Kriegeskorte, "Deep Neural Networks: A New Framework for Modeling Biological Vision and Brain Information Processing," Annual Review of Vision Science, vol. 1, pp. 417-446, 2015.

- [17] S. Bengio, L. Deng, H. Larochelle, H. Lee, and R. Salakhutdinov, "Guest Editors' Introduction: Special Section on Learning Deep Architectures," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 35, pp. 1795-1797, 2013.
- [18] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097-1105.
- [19] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: an efficient alternative to SIFT or SURF," in *Computer Vision (ICCV), 2011 IEEE International Conference on*, 2011, pp. 2564- 2571.
- [20] B. C. Csáji, "Approximation with artificial neural networks," *Faculty of Sciences, Eötvös Loránd University, Hungary*, vol. 24, 2001.
- [21] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Mathematics of control, signals and systems*, vol. 2, pp. 303-314, 1989.
- [22] K. Hornik, "Approximation capabilities of multilayer feedforward networks," *Neural networks*, vol. 4, pp. 251-257, 1991.
- [23] S. Haykin and R. Lippmann, "Neural Networks, A Comprehensive Foundation," *International Journal of Neural Systems*, vol. 5, pp. 363-364, 1994.
- [24] M. H. Hassoun, *Fundamentals of artificial neural networks*: MIT press, 1995.
- [25] K. P. Murphy, *Machine learning: a probabilistic perspective*: MIT press, 2012.
- [26] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov, "Improving neural networks by preventing co-adaptation of feature detectors," *arXiv preprint arXiv:1207.0580*, 2012.
- [27] Y. Bengio, "Deep learning of representations: Looking forward," in *Statistical language and speech processing*, ed: Springer, 2013, pp. 1-37.
- [28] A. Testolin, I. Stoianov, M. De Filippo De Grazia, and M. Zorzi, "Deep unsupervised learning on a desktop PC: a primer for cognitive scientists," *Front. Psychol*, vol. 4, p. 10.3389, 2013.
- [29] A. Weapons, "an Open Letter from AI & Robotics Researchers, 2015," ed.
- [30] C. Szegedy, A. Toshev, and D. Erhan, "Deep neural networks for object detection," in *Advances in Neural Information Processing Systems*, 2013, pp. 2553-2561.
- [31] S. Kombrink, T. Mikolov, M. Karafiát, and L. Burget, "Recurrent Neural Network Based Language Modeling in Meeting Recognition," in *INTERSPEECH*, 2011, pp. 2877-2880.
- [32] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, pp. 2278-2324, 1998.

- [33] T. N. Sainath, A.-r. Mohamed, B. Kingsbury, and B. Ramabhadran, "Deep convolutional neural networks for LVCSR," in *Acoustics, Speech and Signal Processing (ICASSP)*, 2013 IEEE International Conference on, 2013, pp. 8614-8618.
- [34] G. Hinton, L. Deng, D. Yu, G. E. Dahl, A.-r. Mohamed, N. Jaitly, et al., "Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups," *Signal Processing Magazine, IEEE*, vol. 29, pp. 82-97, 2012.
- [35] Y. Bengio, N. Boulanger-Lewandowski, and R. Pascanu, "Advances in optimizing recurrent networks," in *Acoustics, Speech and Signal Processing (ICASSP)*, 2013 IEEE International Conference on, 2013, pp. 8624-8628.
- [36] G. E. Dahl, T. N. Sainath, and G. E. Hinton, "Improving deep neural networks for LVCSR using rectified linear units and dropout," in *Acoustics, Speech and Signal Processing (ICASSP)*, 2013 IEEE International Conference on, 2013, pp. 8609-8613.
- [37] G. Hinton, "A practical guide to training restricted Boltzmann machines," *Momentum*, vol. 9, p. 926, 2010.
- [38] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70, pp. 489-501, 2006.
- [39] B. Widrow, A. Greenblatt, Y. Kim, and D. Park, "The no-prop algorithm: A new learning algorithm for multilayer neural networks," *Neural Networks*, vol. 37, pp. 182-188, 2013.
- [40] Y. Ollivier and G. Charpiat, "Training recurrent networks online without backtracking," *arXiv preprint arXiv:1507.07680*, 2015.
- [41] I. Aleksander, M. De Gregorio, F. M. G. França, P. M. V. Lima, and H. Morton, "A brief introduction to Weightless Neural Systems," in *ESANN*, 2009.
- [42] G. E. Hinton, "Deep belief networks," *Scholarpedia*, vol. 4, p. 5947, 2009.
- [43] H. Larochelle, D. Erhan, A. Courville, J. Bergstra, and Y. Bengio, "An empirical evaluation of deep architectures on problems with many factors of variation," in *Proceedings of the 24<sup>th</sup> international conference on Machine learning*, 2007, pp. 473-480.
- [44] G. E. Hinton, "Training products of experts by minimizing contrastive divergence," *Neural computation*, vol. 14, pp. 1771-1800, 2002.
- [45] A. Fischer and C. Igel, "Training restricted Boltzmann machines: an introduction," *Pattern Recognition*, vol. 47, pp. 25-39, 2014.
- [46] G. Barnich, H. A. Gonzalez, A. Maloney, and B. Oblak, "One loop partition function of threedimensional flat gravity," *Journal of High Energy Physics*, vol. 2015, pp. 1-8, 2015.
- [47] T. Sercu, C. Puhersch, B. Kingsbury, and Y. LeCun, "Very deep multilingual convolutional neural networks for LVCSR," *arXiv preprint arXiv:1509.08967*, 2015.

- [48] C. Yan, F. Coenen, and B. Zhang, "Driving posture recognition by convolutional neural networks," *IET Computer Vision*, 2015.
- [49] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, et al., "Going deeper with convolutions," *arXiv preprint arXiv:1409.4842*, 2014.
- [50] A. Krizhevsky and G. Hinton, "Convolutional deep belief networks on cifar-10," *Unpublished manuscript*, 2010.
- [51] H. Lee, R. Grosse, R. Ranganath, and A. Y. Ng, "Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations," in *Proceedings of the 26<sup>th</sup> Annual International Conference on Machine Learning*, 2009, pp. 609-616.
- [52] Y. LeCun, "The MNIST database of handwritten digits.{Online})," ed, 2007.
- [53] B. Kwolek, "Face detection using convolutional neural networks and Gabor filters," in *Artificial Neural Networks: Biological Inspirations–ICANN 2005*, ed: Springer, 2005, pp. 551-556.
- [54] M. Osadchy, Y. L. Cun, and M. L. Miller, "Synergistic face detection and pose estimation with energy-based models," *The Journal of Machine Learning Research*, vol. 8, pp. 1197-1215, 2007.
- [55] S. Sukittanon, A. C. Surendran, J. C. Platt, and C. J. Burges, "Convolutional networks for speech detection," in *Interspeech*, 2004.
- [56] F. J. Huang and Y. LeCun, "Large-scale learning with svm and convolutional for generic object categorization," in *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, 2006, pp. 284-291.
- [57] H. Lee, P. Pham, Y. Largman, and A. Y. Ng, "Unsupervised feature learning for audio classification using convolutional deep belief networks," in *Advances in neural information processing systems*, 2009, pp. 1096-1104.
- [58] T. Khorshed, "Research Problem Definition Part."
- [59] I. Lenz, H. Lee, and A. Saxena, "Deep learning for detecting robotic grasps," *The International Journal of Robotics Research*, vol. 34, pp. 705-724, 2015.

Page Intentionally Left Blank

## Κεφάλαιο 6<sup>ο</sup>

### Σχετική θεωρία και έρευνες

Καταρχάς, πρέπει να γίνει μία εισαγωγή στις αρχιτεκτονικές των νευρωνικών δικτύων που αξιοποιήθηκαν τόσο σε προηγούμενες έρευνες όσο και στην συγκεκριμένη ερευνητική/μεταπτυχιακή εργασία. Τα συγκεκριμένα, αποτελούν τα περισσότερα διάσημα CNNs για την κατηγοριοποίηση dicom αρχείων, είτε σχετικών με μαγνητικές εγκεφάλου ή, όπως εδώ, με διαδικασίες μαστογραφίας,.

Επιπλέον, παρουσιάζεται μία, συνοπτική, αναφορά σχετικά με τις έννοιες του Training from Scratch και του Transfer Learning καθώς και της βάσης δεδομένων ImageNet που είναι, ιδιαίτερα, διάσημη για την προαναφερόμενη τακτική.

Τέλος, αναφέρονται οι βασικές έρευνες, πάνω στις οποίες στηρίχθηκε το implementation στην συγκεκριμένη εργασία.

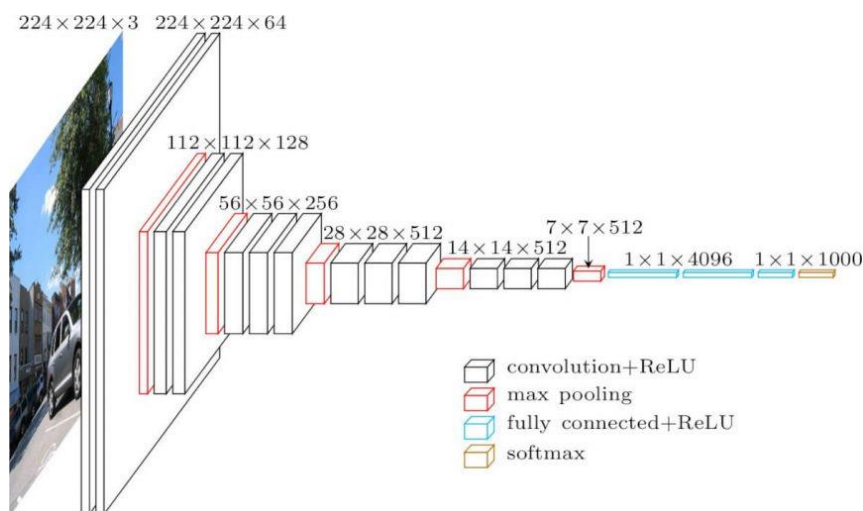
### 6.1 Νευρωνικά Δίκτυα

Λόγω της πληθώρας των δικτύων, μελετώνται αυτά που έχουν χρησιμοποιηθεί στην ίδια βάση δεδομένων με αυτή της συγκεκριμένης εργασίας, δηλαδή της DDSM-CBIS. Αυτά είναι τα VGG16, ResNet50, Inception v3 και το AlexNet.

#### 6.1.1 VGG16

Το VGG16 αποτελεί ένα CNN, το οποίο παρουσιάστηκε το 2014 (Simonyan & Zisserman, 2014). Σε μία από τις πιο γνωστές βάσεις δεδομένων μεγάλου μεγέθους, την ImageNet, το ίδιο πετυχαίνει ακρίβεια 92.7%. Η ίδια θα αναφερθεί εκτενέστερα παρακάτω.

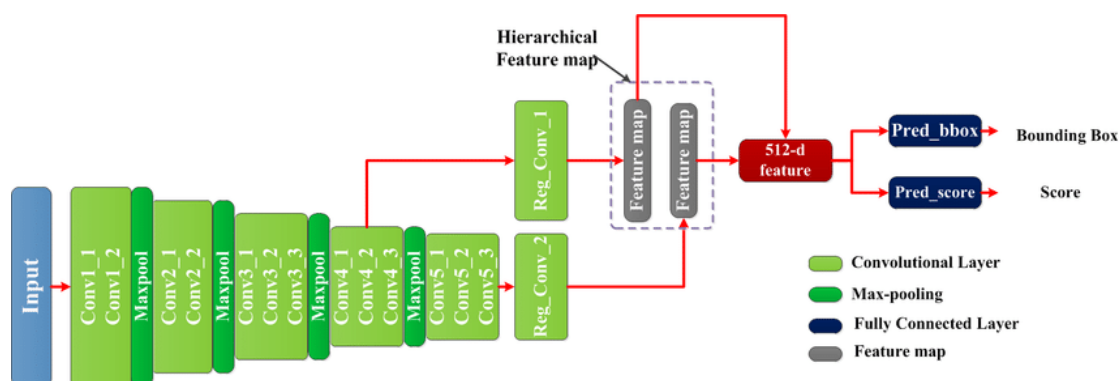
Η αρχιτεκτονική του VGG16, μπορεί να φανεί παρακάτω (Zhang, Zou, He, & Sun, 2015):



Εικόνα 33: Αρχιτεκτονική VGG16

Παρατηρείται ότι η είσοδος, απαιτείται να είναι RGB εικόνα μεγέθους 224X224. Η εικόνα που παρέχεται ως input, αρχικά, διέρχεται μέσω σειρά convolutional επιπέδων, με φίλτρα που φτάνουν μέχρι και διαστάσεις 3X3 (καθώς αυτή είναι η μικρότερη δυνατή που μπορεί να χρησιμοποιηθεί για την εύρεση των εννοιών πάνω-κάτω-αριστερά-δεξιά σε μία εικόνα). Ειδικότερα, τα πρώτα δύο layers έχουν 64 κανάλια φίλτρων μεγέθους 3X3 αξιοποιώντας τις ίδιες τιμές padding.

Στη συνέχεια, εφαρμόζεται ένα max-pooling επίπεδο με stride (2,2) και άλλα δύο convolution layers με φίλτρα μεγέθους 3X3 και συνολικού αριθμού 256, όπως φαίνεται παρακάτω:



Εικόνα 34: Architecture Map του VGG16 (Detailed)

Από την παραπάνω εικόνα, παρατηρείται στη συνέχεια η εφαρμογή 2 convolution επιπέδων, πάλι με μέγεθος φίλτρων 3X3 και μεγέθους 256. Έπειτα, εφαρμόζονται 2 σετ 3 convolution επιπέδων και ενός max-pool layer. Έκαστο επίπεδο διαθέτει 512 φίλτρα, διαστάσεων 3X3 με τις ίδιες τιμές padding.

Στη συνέχεια, ακολουθείται παρόμοια διαδικασία με την εικόνα να διέρχεται ξανά από δύο convolution και ενός max-pooling layer των οποίων τα φίλτρα έχουν μέγεθος 3X3 σε αντίθεση με τις διαστάσεις 11X11 που παρατηρούνται στο AlexNet.

Εδώ, είναι σημαντικό να υπογραμμιστεί η χρήση pixel 1X1 με σκοπό την διαχείριση των αριθμού των καναλιών εισόδου, ενώ σε κάθε convolution layer υπάρχει padding 1 pixel με σκοπό να αποφευχθούν τα χωρικά χαρακτηριστικά (spatial features) της εκάστοτε εικόνας.

Τέλος, το αποτέλεσμα της παραπάνω διάταξης διέρχεται από έναν χάρτη χαρακτηριστικών με μέγεθος (7,7,512), με σκοπό την παραγωγή ενός διανύσματος (feature vector) 1 γραμμής και 25088 στηλών. Έτσι, τοποθετούνται, έπειτα, 3 πλήρως συνδεδεμένα (fully connected) layers, το καθένα από τα οποία, παράγει (ξεκινώντας από τις τελευταίες τιμές) ένα αντίστοιχο διάνυσμα μεγέθους 1x4096. Το τελευταίο από τα 3 επίπεδα, όμως, παράγει μία έξοδο 1000 καναλιών για τον αντίστοιχο αριθμό κατηγοριών της ILSVRC (θα αναφερθεί παρακάτω), ενώ στη συνέχεια, διέρχεται από ένα SoftMax επίπεδο για να επιτευχθεί κανονικοποίηση τιμών του αντίστοιχου διανύσματος.

Όλα τα κρυφά επίπεδα χρησιμοποιούν ως συνάρτηση ενεργοποίησης την ReLU, λόγω της αποδοτικότητάς της στις υπολογιστικές διαδικασίες καθώς και της ικανότητάς να περιορίζει το φαινόμενο των vanishing gradients που περιορίζει την αλλαγή των τιμών



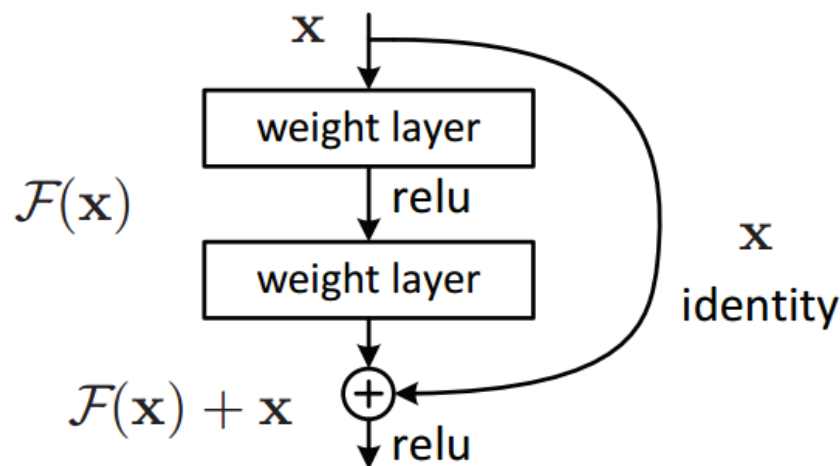
των βαρών (weights), γεγονός που μπορεί να οδηγήσει ακόμη και στον τερματισμό της εκπαίδευσης

### 6.1.2 ResNet50

Το όνομα του συγκεκριμένου νευρωνικού δικτύου αποτελεί συντομογραφία της λέξης Residual Network (Res Net), υποδηλώνοντας ουσιαστικά ότι η συγκεκριμένη αρχιτεκτονική υποστηρίζει ένα είδος εκπαίδευσης στον τομέα της μηχανικής μάθησης που ονομάζεται Residual Learning.

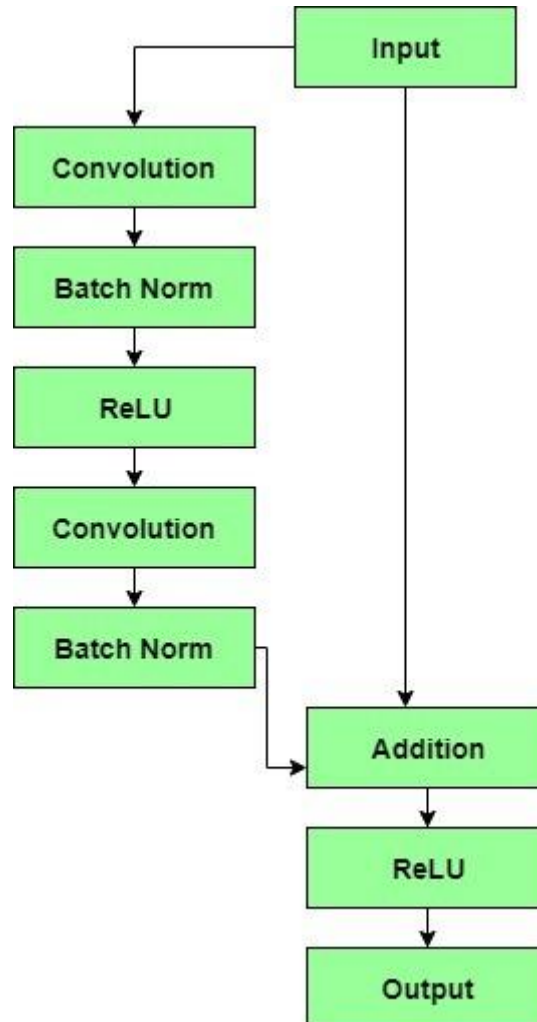
Συγκεκριμένα, με την είσοδο των βαθιών (deep-D) convolutional νευρωνικών δικτύων (D-CNNs) στον τομέα της μηχανικής μάθησης, ήταν έκδηλη η συνεισφορά του μεγαλύτερου αριθμού κρυφών επιπέδων στην ικανότητα πρόβλεψης μίας αρχιτεκτονικής νευρωνικού δικτύου (Krizhevsky, Sutskever, & Hinton, 2012) όπου ένα συνελκτικό (convolutional) επίπεδο ακολουθείται από μία συνάρτηση ενεργοποίησης (activation function) και μία max-pooling διεργασία. Το γεγονός ότι τέτοια είδη επιπέδων είναι σε θέση να αποκρυπτογραφούν περισσότερο πολύπλοκα χαρακτηριστικά, καθιστά την χρήση όλο και περισσότερων “βαθιών” νευρωνικών δικτύων, μία φαινομενικά ορθή πρακτική γενικής χρήσης.

Παρόλα αυτά, αποδείχθηκε ότι η συγκεκριμένη τακτική παρουσιάζει ανώτατα όρια σχετικά με την αποτελεσματικότητά της στον τομέα της εκπαίδευσης νευρωνικών δικτύων (He, Zhang, Ren, & Sun, 2015). Συγκεκριμένα, η παραπάνω μελέτη έδειξε ότι η τιμή λάθους ενός συνελκτικού νευρωνικού δικτύου βάθους 20 σε σύγκριση με ενός με 56 και με την χρήση των σωστών παραμέτρων, είναι μικρότερη. Για την επίλυση του παραπάνω προβλήματος, οι ερευνητές παρουσίασαν ένα νέο layer, το Residual Block, όπως φαίνεται παρακάτω:



Εικόνα 35: Residual Blocks

Το πρόβλημα που αναφέρεται παραπάνω σχετίζεται κυρίως με το vanishing gradient, όπως αναφέρεται και στο VGG16. Η αντιμετώπιση της μικρής τιμής των παραγώγων κατά την διάρκεια του backpropagation, εντοπίζεται στην εισαγωγή identity mappings. Έτσι, προστίθεται το error της εισόδου στο αποτέλεσμα της εκάστοτε συνάρτησης του συνελκτικού δικτύου, όπως φαίνεται παρακάτω:

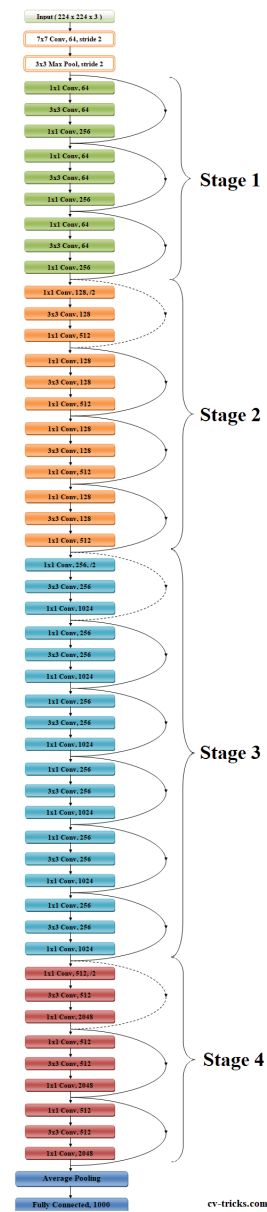


Εικόνα 36: Identity Mapping

Το ResNet50 αποτελεί ένα νευρωνικό δίκτυο που αξιοποιεί την τεχνική εκπαίδευσης Residual Learning και έχει βάθος επιπέδων 50. Η αρχιτεκτονική του, είναι δυνατόν να διαχωριστεί σε 4 βασικά στάδια.

Η εικόνα που παρέχεται ως είσοδος, πρέπει να είναι RGB μεγέθους 224X224. Ως αρχική διεργασία, το input διέρχεται από συνελκτικά και max-pooling στάδια μεγέθους φίλτρων 7X7 και 3X3 αντίστοιχα. Έτσι, ξεκινάει το πρώτο από τα τέσσερα στάδια που περιλαμβάνει 3 Residual Blocks καθένα από τα οποία περιέχει 3 επίπεδα. Τα μεγέθη των kernels, εδώ είναι 64X64 (1<sup>ο</sup> επίπεδο), 64X64 (2<sup>ο</sup> επίπεδο) και 128X128 (3<sup>ο</sup> επίπεδο). Το stride που χρησιμοποιείται σε τέτοιου είδους block είναι μεγέθους 2X2. Για μεγάλα βάθη όπως το συγκεκριμένο νευρωνικό δίκτυο αξιοποιούνται αντίστοιχα ήδη Residual συναρτήσεων με τρία stacked συνελκτικά επίπεδα μεγέθους 1X1, 3X3 και 1X1. Τέλος, υπάρχει και ένα Average Pooling Layer που είναι fully

connected και διαθέτει 1000 νευρώνες (έξοδοι της βάσης δεδομένων ImageNet που αναφέρεται παρακάτω εκτενέστερα).



Εικόνα 37: ResNet50 Architecture

### 6.1.3 Inception v3

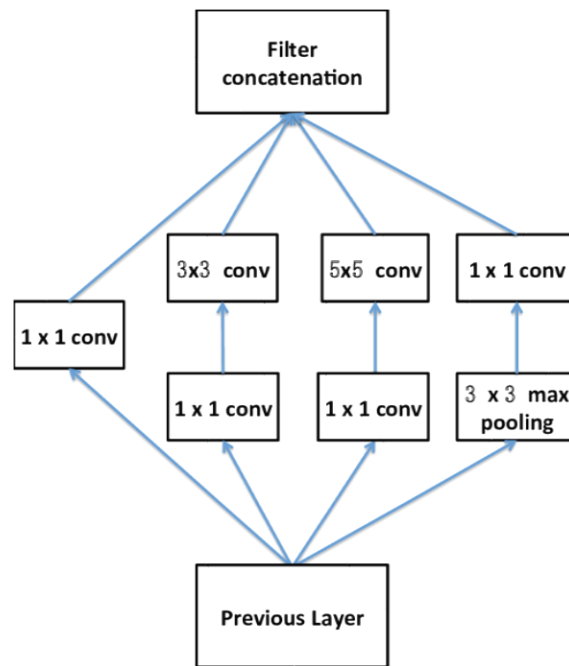
Το συγκεκριμένο είδος νευρωνικού δικτύου, κατασκευάστηκε από την Google και αποτελεί την τρίτη έκδοση μίας σειράς αρχιτεκτονικών Συνελικτικών Νευρωνικών Δικτύων μεγάλου βάθους (DCNN) που έπεται των αντίστοιχων προηγούμενων (Inception v1, Inception v2).

Το πρότυπο στο οποίο στηρίζονται οι παραπάνω εκδόσεις, αποτέλεσε μία εναλλακτική τακτική συγκριτικά με την πρακτική της συνεχόμενης αύξησης του βάθους ενός συνελικτικού νευρωνικού δικτύου με σκοπό την επίτευξη βελτιωμένων ικανοτήτων πρόβλεψης. Εδώ, θα αναλυθούν οι πρώτες τρεις εκδόσεις, αν και υπάρχουν τουλάχιστον άλλες δύο ακόμη αρχιτεκτονικές όπως το Inception v4 και το Inception-ResNet.

Συγκεκριμένα, η πρώτη έκδοση, ξεκίνησε με την θεωρία ότι σε μία εικόνα μπορεί να υπάρχει σημαντική ποικιλομορφία μεταξύ του background και του κεντρικού περιεχομένου-ενδιαφέροντος που είναι επιθυμητό να αναγνωριστεί (Szegedy, et al., 2014). Το εύρος αυτό καθιστά απαραίτητη την επιλογή του κατάλληλου μεγέθους kernel με σκοπό να απομονωθεί η πληροφορία ενώ είναι έκδηλο ότι οι διαστάσεις θα ποικίλλουν σε σημαντικό βαθμό.

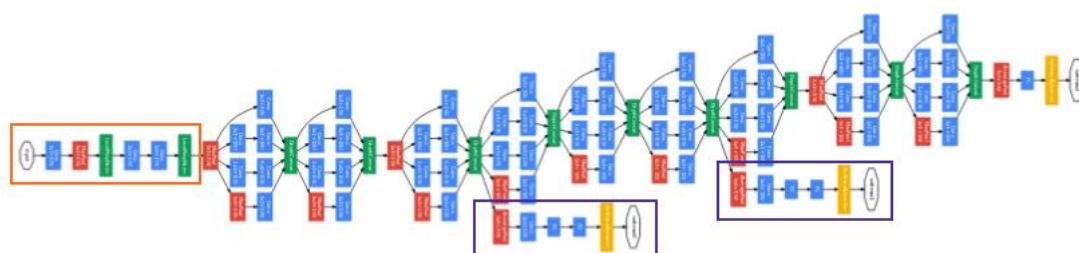
Ως λύση, λοιπόν, παρουσιάστηκε η χρήση πολλαπλών μεγεθών στο ίδιο επίπεδο, με αποτέλεσμα την αύξηση του πλάτους και όχι του βάθους του νευρωνικού δικτύου. Αναλυτικά, προτάθηκε ως “naïve”, η λύση της εφαρμογής συνέλιξης για 3 διαφορετικά φίλτρα (1X1, 3X3 και 5X5) ενώ, ταυτόχρονα, εφαρμόζονται max-pooling τεχνικές με τα αποτελέσματα να προωθούνται στο επόμενο module.

Στη συνέχεια, όμως, περιορίστηκαν τα κανάλια εισόδου με την πρόσθεση μίας επιπλέον 1X1 συνέλιξης πριν τις αντίστοιχες με μεγέθη 3X3 και 5X5, η οποία προηγείται το max-pooling layer, όπως φαίνεται παρακάτω.



Εικόνα 38: Inception Module

Με την αξιοποίηση της παραπάνω πρακτικής, κατασκευάστηκε το Inception v1 ή αλλιώς Google Net, η αρχιτεκτονική του οποίου παρουσιάζεται παρακάτω:



Εικόνα 39: Αρχιτεκτονική Googlenet

Το συγκεκριμένο είχε συνολικό βάθος 27 και εξαιτίας αυτού είναι επιρρεπής στο vanishing gradient problem. Για τον λόγο αυτό, προστέθηκαν δύο auxiliary κατηγοροποιητές (τα δύο μπλε κιβώτια στην εικόνα 7) με σκοπό να αποτρέψουν, με την χρήση SoftMax συναρτήσεων στις εξόδους, την εμφάνιση ιδιαίτερος μικρών τιμών βαρών στο δίκτυο.

Στην ίδια έρευνα (Szegedy, και συν., 2014) παρουσιάστηκαν οι άλλες δύο εκδόσεις που αναφέρθηκαν παραπάνω (Inception v2 και Inception v3), στις οποίες προτείνεται ένα σύνολο βελτιώσεων με σκοπό τόσο την αύξηση της ακρίβειας πρόβλεψης του δικτύου όσο και την μείωση του υπολογιστικού κόστους. Επειδή, στην συγκεκριμένη εργασία, μελετάται το μοντέλο Inception v3, θα παραληφθεί η αναφορά στην δεύτερη έκδοση.

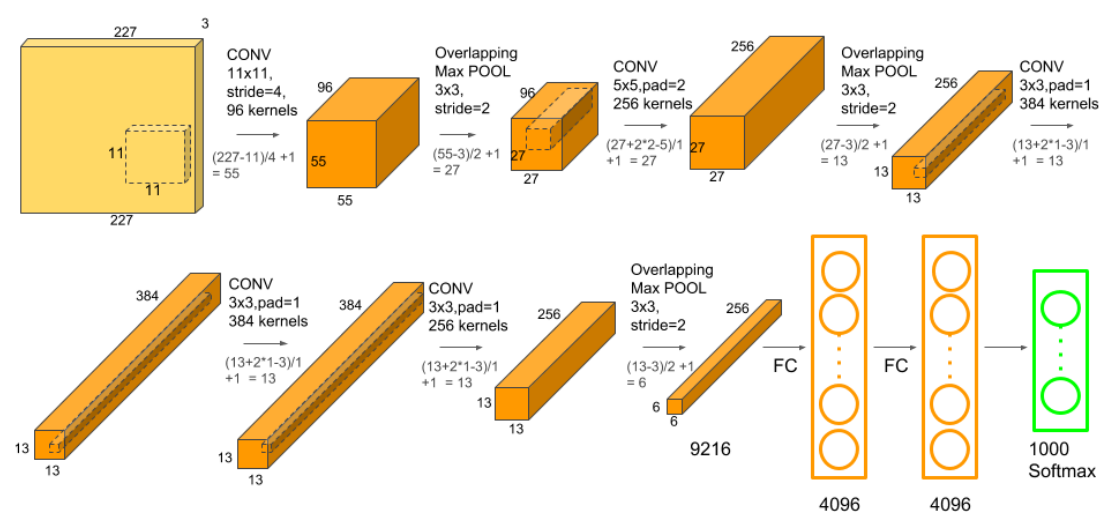
Σύμφωνα, με τους ερευνητές, οι auxiliary classifiers που αναφέρθηκαν πριν, παρουσίαζαν ελάχιστη συνεισφορά όσο η διαδικασία εκπαίδευσης έφτανε στο τελικό της στάδιο και οι τιμές της ακρίβειας είχαν σταθεροποιηθεί ενώ ταυτόχρονα, σύμφωνα με τους ίδιους, υπήρχε πληθώρα δυνατοτήτων βελτίωσης των αποτελεσμάτων της δεύτερης έκδοσης χωρίς να είναι απαραίτητες δραστικές αλλαγές.

Έτσι, με την αξιοποίηση για την optimization διαδικασία του RMSProp, την χρήση συνελίξεων με διαστάσεις 7X7, την τοποθέτηση Batch Norms στους Auxiliary κατηγοροποιητές και την εισαγωγή μίας πληθώρας τεχνικών με σκοπό την μείωση της υπερβολικής “αυτοπεποίθησης” του νευρωνικού δικτύου στην ικανότητα πρόβλεψής του (Label Smoothing) παράχθηκε το μοντέλο Inception Net v3.

### 6.1.4 AlexNet

Το συγκεκριμένο νευρωνικό δίκτυο αποτέλεσε τον νικητή στον διαγωνισμό ImageNet Large Scale Visual Recognition Challenge το 2012 (περισσότερες πληροφορίες παρακάτω). Η εικόνα που παρέχεται ως είσοδος, πρέπει να είναι RGB μεγέθους 256X256. Στην αντίστοιχη έρευνα (Krizhevsky, Sutskever, & Hinton, 2012), αναφέρεται ως ορθή είσοδος οι διάσταση 224 αλλά έχει αποδειχθεί ότι περισσότερο νόημα έχει το μέγεθος 227X227.

Το ίδιο, όταν συστήθηκε στο ευρύ κοινό το 2012, αποτελούσε ένα σύστημα μεγαλύτερου βάθους σε σύγκριση με τα μέχρι τότε νευρωνικά δίκτυα, γεγονός που μπορεί να φανεί και από την αρχιτεκτονική του στο παρακάτω σχήμα:



Εικόνα 40: AlexNet Architecture

Συγκεκριμένα, αποτελείται από 5 συνελκτικά και 3 πλήρως συνδεδεμένα layers. Τα φίλτρα (kernels) που αξιοποιούνται στο πρώτο convolutional επίπεδο είναι 96 στον αριθμό και μεγέθους 11X11X3. (τρία κανάλια εισόδου λόγω RGB εικόνας). Τα πρώτα δύο συνελκτικά επίπεδα οδηγούν τις εξόδους στα αντίστοιχα επικαλυπτόμενα max-pooling σε αντίθεση με το τρίτο και τέταρτο. Για το τελευταίο (δηλαδή το 5<sup>ο</sup>), η έξοδος οδηγείται, πάλι σε ένα max-pooling επίπεδο, του οποίου το αποτέλεσμα διαδραματίζει τον ρόλο της εισόδου σε μία σειρά από δύο πλήρως συνδεδεμένα επίπεδα, όπως μπορεί να φανεί στην εικόνα 8.

Τέλος, το δεύτερο, από αυτά τα επίπεδα, τροφοδοτεί έναν SoftMax κατηγοροποιητή με 1000 διαφορετικά είδη τάξεων (ImageNet).

Ως συνάρτηση ενεργοποίησης, χρησιμοποιείται η ReLU, η οποία εφαρμόζεται τόσο στο τέλος κάθε συνελκτικού αλλά και πλήρως συνδεδεμένου επιπέδου.

## 6.2 Σύγκριση μεθόδων εκπαίδευσης

Στην σύγχρονη βιβλιογραφία σχετικά με την εκπαίδευση νευρωνικών δικτύων, παρουσιάζονται δύο βασικές μέθοδοι. Αυτές είναι, πρώτον η εκπαίδευση ενός συστήματος εξολοκλήρου από την αρχή (training from scratch) και δεύτερον η χρήση ενός, ήδη εκπαιδευμένου, δικτύου (transfer learning) με την διόρθωση ορισμένων παραμέτρων του (hyperparameters).

Η μέθοδος του transfer learning είναι ιδιαίτερα δημοφιλής λόγω της φύσης της. Συγκεκριμένα, είναι γνωστό ότι ένας σημαντικός παράγοντας που συμβάλλει στην επιτυχία ενός νευρωνικού δικτύου είναι ο μεγάλος όγκος δεδομένων ως είσοδο. Στην συγκεκριμένη πρακτική, εφόσον τα δεδομένα είναι καλής ποιότητας, δεν υπάρχει όριο και ισχύει η φράση όσοι περισσότεροι τόσο καλύτερα.

Παρόλα αυτά, όμως, δεν είναι δυνατόν, σε αρκετές περιπτώσεις να είναι διαθέσιμη μεγάλη ποσότητα δεδομένων, ενώ άλλες φορές το πρόβλημα επεκτείνεται και στο hardware που χρησιμοποιείται (CPU/GPU) στο εκάστοτε εργαστήριο, θέτοντας σημαντικούς περιορισμούς στο υπολογιστικό κόστος/χρόνο.

## 6.3 Transfer Learning

Πανάκεια στο παραπάνω τροχοπέδη, είναι δυνατόν να αποτελέσει η τακτική του transfer learning για τα συνελκτικά νευρωνικά δίκτυα μεγάλου βάθους (DCNNs) (Cadène, Thome, & Cord, 2016). Συγκεκριμένα, αξιοποιώντας μία μεγάλη βάση δεδομένων (εδώ την ImageNet που αναλύεται παρακάτω), είναι δυνατόν να χρησιμοποιηθούν νευρωνικά δίκτυα με ήδη ορισμένο, σε βέλτιστο βαθμό, ένα σημαντικό ποσοστό των βαρών, σε αντίθεση με την αρχικοποίησή τους με τυχαίες τιμές και την εύρεση από την αρχή όλων των μεγεθών τους. Με αυτόν τον τρόπο, υλοποιώντας ορισμένες ακόμη τεχνικές όπως το “πάγωμα” (freeze) των επιπέδων εισόδου (αφού εκεί θα τοποθετηθούν ως είσοδοι οι καινούργιες βάσεις δεδομένων) και την διόρθωση ορισμένων παραμέτρων στα συνελκτικά επίπεδα της εξόδου (αριθμός κλάσεων που κατηγοριοποιούνται κτλ.) είναι δυνατόν, ο ερευνητής να παράγει σε σύντομο χρονικό διάστημα ένα σύστημα που έχει, ιδιαίτερα μεγάλες, πιθανότητες να παρουσιάσει υψηλή ικανότητα πρόβλεψης.

Παρόλα αυτά, όμως, υπάρχουν και ορισμένα μειονεκτήματα στην παραπάνω διαδικασία. Αρχικά, ένα από τα περισσότερο σημαντικά, μπορούν να εντοπιστούν στο είδος της εκπαίδευσης που έχει προηγηθεί στο νευρωνικό δίκτυο.

Αναλυτικά, το έκαστο pre-trained σύστημα, εφόσον αναφερόμαστε στην βάση δεδομένων ImageNet, έχει εκπαιδευτεί σε ένα τεράστιο σύνολο φυσικών (nature) RGB φωτογραφιών. Το γεγονός αυτό, θέτει ορισμένους προβληματισμούς, όταν αναφερόμαστε σε μη φυσικές φωτογραφίες όπως πχ μία εικόνα μαστογραφίας αλλά και στις grayscale εικόνες, σημαντικό ποσοστό των οποίων αποτελούν τα dicom αρχεία.

Επιπλέον, πρέπει να τονιστεί ότι, αρκετά συχνά, η χρήση τέτοιων συστημάτων θα μπορούσε να αποφευχθεί με την from scratch εκπαίδευση περισσότερο απλοϊκών αρχιτεκτονικών ή/και λιγότερων κύκλων. Ως αποτέλεσμα, τα pre-trained αυτά συστήματα εμφανίζονται, συνήθως, με πολύ μεγαλύτερο μέγεθος από τα, πιθανά πιο απλά, γεγονόσ που σε περιπτώσεις ιδιαίτερα περιορισμένων αρχιτεκτονικών (web



services) μπορεί να δημιουργήσει σημαντικό πρόβλημα στην υλοποίηση της εκάστοτε υπηρεσίας.

#### 6.4 ImageNet

Η ίδια, αποτελεί μία τεράστια συλλογή από φωτογραφίες που έχουν γίνει annotated από ανθρώπους και ο σκοπός χρήσης της αφορά, κυρίως, την ανάπτυξη του τομέα έρευνας της μηχανικής όρασης. Ειδικότερα, αποτελείται από περισσότερες από 14 εκατομμύρια εικόνες, οι οποίες έχουν διαχωριστεί σε λίγο παραπάνω από 21 χιλιάδες κατηγορίες ενώ 1 εκατομμύριο από τις φωτογραφίες έχουν συνοδεύονται από τα προαναφερόμενα annotation boxes.

##### *ImageNet Large Scale Visual Recognition Challenge*

Όπως έχει ήδη αναφερθεί, ένας από τους καταλυτικότερους διαγωνισμούς, ο οποίος επέτρεψε την άνθιση των νευρωνικών δικτύων μεγάλου βάθους ήταν ο ImageNet Large Scale Visual Recognition Challenge (ILSVRC). Πραγματοποιείται κάθε χρόνο, δέχεται συμμετοχές παγκοσμίως και έχει ως βάση ανάπτυξης την βάση δεδομένων ImageNet.

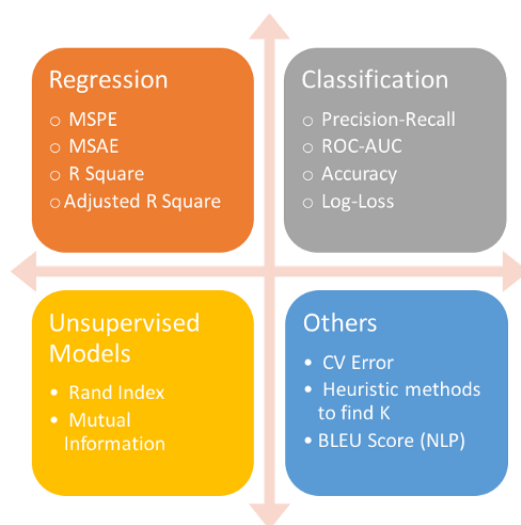
Ο συγκεκριμένος διαγωνισμός, αξιολογεί αλγορίθμους για διεργασίες που σχετίζονται με εντοπισμό αντικειμένων και κατηγοριοποίηση εικόνων σε μεγάλη κλίμακα.

## 6.5 Δείκτες Αξιολόγησης

Πριν την αναφορά στις μελέτες που έχουν γίνει, είναι σημαντικό να αναφερθούν, συνοπτικά οι δείκτες των μετρήσεων που αξιοποιούνται, συνήθως, σε τέτοιου είδους προβλήματα κατηγοριοποίησης.

Η παρακολούθηση τους, αποτελεί, συχνά, την πιο ουσιαστική, μετά - πειραματική διαδικασία και είναι σε θέση να παρέχει αξιόπιστες πληροφορίες για τα παραγόμενα αποτελέσματα.

Εδώ, πρέπει να τονιστεί, ότι ανάλογα την φύση του προβλήματος, συνήθως, επιλέγεται και ένας αντίστοιχος, όπως φαίνεται στην παρακάτω εικόνα:



Εικόνα 41: Δείκτες Μετρήσεων-Είδος Προβλήματος

### Βασικές έννοιες

Οι δείκτες αυτοί σχετίζονται, στην πλειονότητά τους, με συγκεκριμένες έννοιες που συναντώνται, αμιγώς στα προβλήματα κατηγοριοποίησης.

#### Αληθώς Θετικό (True Positive)

Αφορά την αναγνώριση μίας κατάστασης από το σύστημα, όταν εκείνη, συμβαίνει στην πραγματικότητα.

#### Αληθές Αρνητικό (True Negative)

Αφορά την αναγνώριση της απουσίας μίας κατάστασης από το σύστημα, όταν εκείνη, δεν συμβαίνει στην πραγματικότητα.

#### Ψευδές Θετικό (False Positive)

Αφορά την αναγνώριση μίας κατάστασης από το σύστημα, ενώ εκείνη δεν συμβαίνει στην πραγματικότητα.

#### Ψευδές Αρνητικό (False Negative)

Αφορά την αναγνώριση της απουσίας μίας κατάστασης από το σύστημα, ενώ εκείνη συμβαίνει στην πραγματικότητα.

	Predicted: NO	Predicted: YES
Actual: NO	TN = ??	FP = ??
Actual: YES	FN = ??	TP = ??

Εικόνα 42: TP-TN-FP-FN

### Recall

Με την χρήση του δείκτη Recall ή αλλιώς Ευαισθησίας, που συμβολίζεται και ως True Positive Result (TPR), απεικονίζεται ο αριθμός των δεδομένων που έχουν κατηγοριοποιηθεί ως Θετικό προς τον συνολικό αριθμό των Αληθές Θετικών στοιχείων.

### Precision

Η ακρίβεια ή αλλιώς Precision, αφορά τον αριθμό των αντικειμένων που έχουν αναγνωριστεί ως Θετικό προς το σύνολο των, αναγνωρισμένων Θετικό, δειγμάτων.

Σύμφωνα με την θεωρία αλλά και εμπειρικά, οι παρακάτω συνδυασμοί των δύο προηγούμενα αναφερθέντων, δεικτών (Recall και Precision), μπορούν να εξάγουν τα εξής αποτελέσματα για την ποιότητα του εκπαιδευμένου αλγορίθμου:

- Υψηλές τιμές Recall και Precision : Το μοντέλο που έχει χρησιμοποιηθεί είναι κατάλληλο για γενίκευση των εκτιμήσεων του και έχει αξιόπιστες θετικές δυνατότητες στην επίλυση του ζητήματος για το οποίο έχει επιλεγεί.
- Χαμηλή τιμή Recall και υψηλό Precision : Το μοντέλο δυσκολεύεται αρκετά να αναγνωρίσει την κλάση αλλά η επιτυχής του προσπάθεια χαίρει σημαντικής αξιοπιστίας.
- Υψηλή τιμή Recall και χαμηλό Precision: Οι προβλέψεις του μοντέλου δεν είναι ιδιαίτερα αξιόπιστες αλλά οι κατηγορίες προς πρόβλεψη, αναγνωρίζονται με ιδιαίτερη ευκολία.
- Χαμηλές τιμές Recall και Precision : Το μοντέλο που έχει χρησιμοποιηθεί είναι, μη ικανοποιητικά, αποτελεσματικό για την επιθυμητή γενίκευση ή/και χρίζει περισσότερου ελέγχου - προ-επεξεργασίας προκειμένου να είναι σε θέση να ανταποκριθεί στις απαιτήσεις του προβλήματος.

### Specificity

Η ακρίβεια - ιδιαιτερότητα ή Specificity, σχετίζεται με τον αριθμό των αντικειμένων που έχουν αναγνωριστεί ως Αρνητικό προς το σύνολο των αρνητικών δειγμάτων.

### Τύπος I Λάθους

Ο συγκεκριμένος δείκτης συμβολίζεται ως FPR (False Positive Rate) και αφορά το ποσό των δειγμάτων που αναγνωρίζονται, λανθασμένα ως θετικά, προς το σύνολο των αρνητικών.

### Τύπος II Λάθους

Η διαφορά του συγκεκριμένου δείκτη, σε σύγκριση με τον I, εντοπίζεται στο γεγονός ότι ο συγκεκριμένος αναγνωρίζει τα δείγματα που έχουν αναγνωριστεί ως αρνητικά, προς το σύνολο των αληθών θετικών. Αναλυτικά, αφορά το ποσό των δειγμάτων που αναγνωρίζονται, λανθασμένα ως θετικά, προς το σύνολο των αρνητικών.

### F1 Score

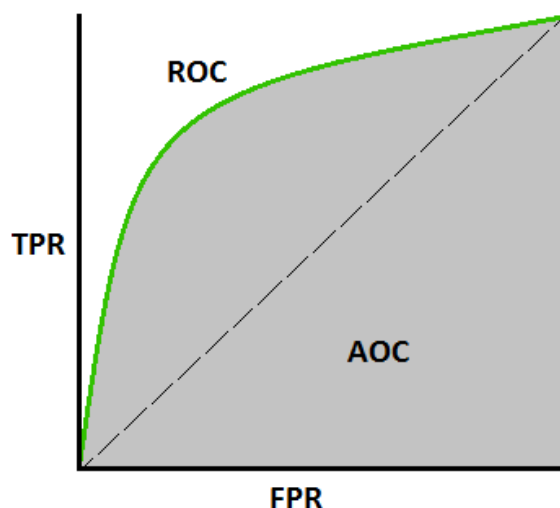
Ο συγκεκριμένος αφορά τον αρμονικό μέσο όρο των δεικτών Precision και Recall.

### Accuracy

Η ακρίβεια (Accuracy), ως περισσότερο απλοϊκός δείκτης, υπολογίζει το ποσοστό των συνολικών δειγμάτων που έχουν ταξινομηθεί με σωστό τρόπο.

### ROC – AUC

Το σκορ στους συγκεκριμένους δείκτες αφορά την πιθανότητα, σε περίπτωση τυχαία επιλογής μίας θετικής και μίας αρνητικής κατάστασης, η πρώτη να ξεπερνάει την δεύτερη. Γεωμετρικά, επεξηγείται από τον υπολογισμό της περιοχής που βρίσκεται κάτω από την καμπύλη που σχηματίζεται, τοποθετώντας στο καρτεσιανό επίπεδο, στον άξονα των x τον δείκτη FPR ή αλλιώς Τύπο I Λάθους και στον y, τον TPR ή αλλιώς, Recall.



Εικόνα 43: Υπολογισμός ROC-AUC

Η τιμή τους κυμαίνεται μεταξύ στο εύρος  $[0,1]$  και τυπικά για ένα ισορροπημένο σετ δεδομένων και για έναν τυχαίο κατηγοροποιητή, αυτή θα προσεγγίζει το 0.5 .

### Λογαριθμικές Απώλειες

Για δυαδική κατηγοριοποίηση, ο συγκεκριμένος δείκτης που ονομάζεται, επίσημα, Log Loss, παρέχει την πληροφορία σχετικά με την εμπιστοσύνη, σε μορφή πιθανότητας, στις προβλέψεις.

Για παράδειγμα, έστω ότι υπάρχει διαθέσιμη, μία πρόβλεψη με πιθανότητα 0.5 για όλες τις παρατηρήσεις. Η τιμή της Log Loss σύμφωνα με την αντίστοιχη εξίσωση, θα είναι 0.69, επιτρέποντας στον ερευνητή, να θεσπίσει ένα όριο για την ποιότητα των μετρήσεων του (στην συγκεκριμένη περίπτωση ότι οποιαδήποτε μέτρηση παρέχει λογαριθμικές απώλειες άνω της τιμής 0.6 θεωρείται μη αποδεκτή).

### Μίκρο-Μάκρο Μέσοι Όροι

Οι μικρο - μέθοδοι αφορούν το άθροισμα των ανεξαρτήτων αληθών, ψευδών θετικών και των ψευδών αρνητικών του συστήματος για διαφορετικά σεντ τα οποία αξιοποιεί για την παραγωγή των αντιστοιχών στατιστικών μεθόδων. Προτιμώνται κατά την ύπαρξη μη - ισορροπημένων συνόλου δεδομένων.

Στις μάκρο - μεθόδους, ευρίσκεται το μέσο των δεικτών Precision και Recall του συστήματος για διαφορετικά σεντ.

## 6.6 Άλλες Έννοιες

Τέλος, παρουσιάζονται ορισμένοι όροι που χρησιμοποιήθηκαν, συχνά, στην περιγραφή των αλγορίθμων παραπάνω και θεωρήθηκε απαραίτητη, η εκτεταμένη ανάλυσή τους.

### Imbalanced Dataset

Η ύπαρξη μη - ισορροπημένων συνόλου δεδομένων εκπαίδευσης σε ένα πρόβλημα Επιβλεπόμενης Μάθησης με Κατηγοριοποίηση Πολλαπλών Κλάσεων αποτελεί ένα ζήτημα καίριας σημασίας, η παράβλεψη του οποίου, οδηγεί συνήθως, σε χαμηλές αποδόσεις των παρεχόμενων προβλέψεων.

Ένα κλασσικό παράδειγμα αποτελεί το πρόβλημα εντοπισμού απάτης σε χρηματοπιστωτικές συναλλαγές καθώς το ποσοστό τους σε σύγκριση με τις έγκυρες περιπτώσεις είναι σημαντικά μικρότερο.

Στις περιπτώσεις αυτές, προτείνεται ο πειραματισμός με πολλές παραμέτρους όπως :

- Η αλλαγή του αλγορίθμου μάθησης που χρησιμοποιείται
- Η αλλαγή του δείκτη αξιολόγησης που εφαρμόζεται
- Επαναδειγματοληψία (Resampling), με την προσθήκη αντιγράφων της κατηγορίας με την μικρότερη συχνότητα παρουσίας στο σεντ, ειδικά σε περιπτώσεις όπου δεν είναι δυνατή η εύρεση μεγαλύτερου αριθμού δεδομένων
- Σε αντίθεση με την προηγούμενη τεχνική, η λήψη μικρότερου αριθμού δειγμάτων από τις " κυρίαρχες " κλάσεις στο σύνολο των δεδομένων.
- Η δημιουργία συνθετικών δειγμάτων με μεθόδους όπως η Synthetic Minority Oversampling Technique (S.M.O.T.E) στην τάξη με τον μικρότερο αριθμό στοιχείων.

## 6.7 Έρευνες

Το πρώτο κομμάτι της διπλωματικής, αυτής, εργασίας, αφορά την επιβεβαίωση των ισχυρισμών των περισσότερων διάσημων πρακτικών που έχουν χρησιμοποιηθεί με σκοπό την κατηγοριοποίηση των εικόνων μαστογραφίας, μέσω της εκπαίδευσης στα αρχεία της βάσης δεδομένων DDSM-CBIS.

Η πρώτη από τις συγκριτικές μελέτες, που θα αναφερθούν, επικεντρώθηκε στην αναγνώριση των μαζών στις αντίστοιχες μαστογραφίες (Richa, Oliver, Xavier, Moi, & Robert, 2019). Η συγκεκριμένη, εκτός της βάσης δεδομένων DDSM-CBIS, ανέλυσε και το σετ δεδομένων INbreast, ένα σύνολο μαστογραφιών, επίσης σε dicom μορφή που όμως διαφέρει από το πρώτο καθώς ανήκει στην κατηγορία Full Field Digital Mammograms (FFDM) σε αντίθεση με την DDSM-CBIS που ανήκει στην δεύτερη, δηλαδή στα Scanned/Screened Film Mammograms (SFM).

Αρχικά, αυτή η συγκριτική έρευνα χρησιμοποίησε τρεις από τις προαναφερόμενες, διάσημες, αρχιτεκτονικές συνελκτικών δικτύων και συγκεκριμένα, τις VGG16, ResNet50 και Inception v3.

Όπως είναι γνωστό, τα παραπάνω δίκτυα, απαιτούν ως είσοδο RGB εικόνες διαστάσεων 224x224. Παρόλα αυτά, οι μαστογραφίες της DDSM-CBIS, όπως και οι περισσότερες εικόνες ιατρικού τύπου, διαθέτουν μόνο ένα κανάλι (gray level) αντί για τα απαιτούμενα τρία (red, green, blue). Ως λύση, πραγματοποιήθηκε επαναληπτική χρήση του ενός αυτού καναλιού και για τα τρία ενώ, ακόμη, εφαρμόστηκε zero-mean κανονικοποίηση των τιμών έντασης (intensity) της κάθε φωτογραφίας (θα αναλυθεί παρακάτω), ενώ σχετικά με το μέγεθος, πραγματοποιήθηκε rescaling για κάθε dicom εικόνα μαστογραφίας στις απαιτούμενες διαστάσεις (224X224).

Όσον αφορά τον σχεδιασμό της διαδικασίας της εκπαίδευσης, έγινε ένας τυπικός διαχωρισμός στο σύνολο της εκπαίδευσης, με σκοπό την δημιουργία ενός συνόλου για την επιβεβαίωση της καλής λειτουργίας εκάστου classifier (validation phase) σε κάθε εποχή (epoch), μία τιμή που αφορά τον αριθμό των επαναλήψεων, κατά τις οποίες ο αντίστοιχος αλγόριθμος επεξεργάζεται ολόκληρο το σετ δεδομένων.

Ορισμένα, ιδιαιτέρως σημαντικά, βήματα στον τομέα της προ επεξεργασίας τα οποία ακολουθήθηκαν, εδώ, αφορούν την αύξηση του αριθμού των δεδομένων προς εκπαίδευση, με την μέθοδο data augmentation. Η συγκεκριμένη σχετίζεται με την παραγωγή περισσότερων δειγμάτων από το ήδη υπάρχον σύνολο εκτελώντας ορισμένους μετασχηματισμούς όπως αλλαγή μεγέθους (resize) ή/ και περιστροφή(rotation). Στην συγκεκριμένη έρευνα, αξιοποιήθηκε το horizontal flipping, η περιστροφή μέχρι και 30 μοίρες και η μετατροπή των διαστάσεων σε ένα εύρος [0.75 1.25] της αρχικής τιμής (rescaling).

Ως μηχανισμός βελτιστοποίησης, για όλες τις αρχιτεκτονικές, χρησιμοποιήθηκε ο Adam ενώ το μέγεθος του batch ήταν 128. Επίσης, με σκοπό τον περιορισμό της απώλειας validation, έχει ενεργοποιηθεί και η δυνατότητα του Early Stopping με κατώτερο όριο τις 10 εποχές.

Ακόμη, είναι σημαντικό να αναφερθεί ότι για κάθε δίκτυο, χρησιμοποιήθηκε τόσο η τεχνική training from scratch όσο και το transfer learning, επιλέγοντας έναν υψηλότερο

ρυθμό μάθησης (learning rate) στην πρώτη περίπτωση, λόγω του επιβαρυνμένου υπολογιστικού κόστους/χρόνου που, εκ φύσεως, παρουσιάζει.

Ως δείκτες αξιολόγησης, χρησιμοποιήθηκαν οι μεταβλητές Accuracy, ROC και η FROC καμπύλη (δεν αναλύεται περαιτέρω, καθώς αφορά μόνο την ικανότητα πρόβλεψης για το INbreast σετ δεδομένων).

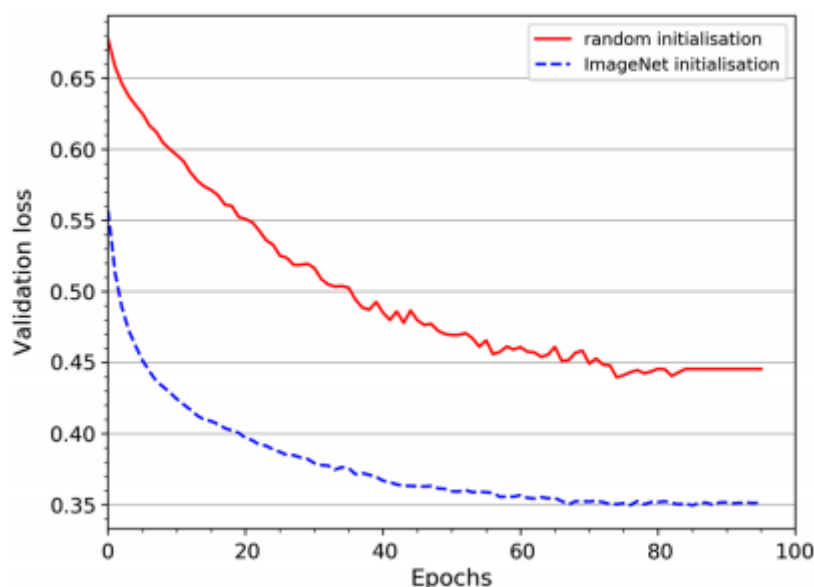
Πραγματοποιήθηκαν τρεις διαφορετικές πειραματικές διαδικασίες οι οποίες αναλύονται παρακάτω.

Αρχικά, αξιοποιούνται οι προαναφερόμενες αρχιτεκτονικές νευρωνικών δικτύων με την μέθοδο του transfer learning, χρησιμοποιώντας τα pretrained μοντέλα της που εκπαιδεύτηκαν στην βάση δεδομένων ImageNet. Η τακτική αυτή, συγκρίνεται με την from scratch training για την διάγνωση και ονομασία μαζών στις εικόνες μαστογραφίας της DDSM-CBIS.

Με βάση τις παρακάτω εικόνες, είναι έκδηλο ότι τα pretrained μοντέλα παρουσίασαν βελτιωμένες επιδόσεις τόσο στο Accuracy (βλ. εικόνα 12) όσο και στην πρόοδο του Validation Loss (βλ. εικόνα 13).

Model	Pretrained	Time per epoch (s)	Testing accuracy ( $\mu \pm \sigma$ )	Az
VGG16	No	518	82.39% $\pm$ 0.52	0.90 $\pm$ 0.01
	Yes	465	83.69% $\pm$ 0.24	0.92 $\pm$ 0.01
Resnet50	No	483	82.30% $\pm$ 0.70	0.91 $\pm$ 0.01
	Yes	438	83.69% $\pm$ 0.15	0.92 $\pm$ 0.01
InceptionV3	No	338	82.10% $\pm$ 0.58	0.90 $\pm$ 0.01
	Yes	310	84.16% $\pm$ 0.19	0.93 $\pm$ 0.01

Εικόνα 44: Accuracies των διαφορετικών μοντέλων (pretrained και from scratch)



Εικόνα 45: Validation Loss σε σύγκριση με τις εποχές

Σύμφωνα με τα παραπάνω, παρατηρήθηκε ότι το μοντέλο Inception v3 παρουσίασε τις καλύτερες αποδόσεις και για τον λόγο αυτό, οι ερευνητές, προχώρησαν στην εκτέλεση των υπόλοιπων πειραμάτων αποκλειστικά με αυτό. Παρόλα, αυτά, όπως αναφέρεται και παρακάτω, χρήσιμα συμπεράσματα μπορούν να εξαχθούν και από την χρήση των υπόλοιπων, υποψήφιων μοντέλων καθώς οι μικρές διαφορές στις τιμές του Accuracy μπορεί να αποτελούν συνέπεια μη-βέλτιστου tuning των παραμέτρων της κάθε αρχιτεκτονικής.

Στην δεύτερη πειραματική διαδικασία, οι ερευνητές ασχολήθηκαν με το fine-tuning, μία διαδικασία κατά την οποία πραγματοποιούνται ορισμένες αλλαγές σε συγκεκριμένα layers μίας αρχιτεκτονικής νευρωνικού δικτύου με σκοπό την καλύτερη προσαρμογή στο δεδομένο πρόβλημα, όπως η τροποποίηση του SoftMax layer, η χρήση μικρότερης τιμής ρυθμού μάθησης (learning rate) και το “πάγωμα” (freeze) των αρχικών layers που είναι υπεύθυνα για την εύρεση χαρακτηριστικών όπως καμπύλες και γωνίες, features που σχετίζονται, σημαντικά με το είδος και τις κατηγορίες των εικόνων που έχει γίνει η αρχική εκπαίδευση του δικτύου.

Με το fine-tuning στο σετ δεδομένων INbreast, έχοντας ως σημείο αναφοράς, πλέον, την βάση δεδομένων DDSM-CBIS, το μοντέλο Inception v3 παρουσίασε, ακόμη, καλύτερη δυνατότητα εκτίμησης συγκριτικά με το pre-trained στην ImageNet, όπως φαίνεται παρακάτω:

Model	Pretrained weight	Fine-tuning on INbreast	Testing accuracy ( $\mu \pm \sigma$ )	Training cascade
InceptionV3	ImageNet	Yes	85.29% $\pm$ 4.29	ImageNet $\rightarrow$ INbreast
InceptionV3	CBIS-DDSM	Yes	88.86% $\pm$ 2.96	ImageNet $\rightarrow$ CBIS-DDSM $\rightarrow$ INbreast

Εικόνα 46: Fine tuning στην INbreast, pretrained σε Imagenet/CBIS

Στο τελευταίο πείραμα, αξιοποιήθηκε η αρχιτεκτονική νευρωνικού δικτύου που, μέχρι τώρα, παρουσίαζε τα βέλτιστα αποτελέσματα, δηλαδή το μοντέλο Inception v3 που



έχει εκπαιδευτεί στην CBIS-DDSM και έχει πραγματοποιηθεί fine-tuning στην INbreast, με σκοπό την τελική κατηγοριοποίηση-εντοπισμό μαζών στη βάση δεδομένων INbreast.

Για την βέλτιστη αξιολόγηση αποτελεσμάτων, οι όγκοι χωρίστηκαν σε τρεις κατηγορίες, με βάση το μέγεθος τους. Έτσι, στους μικρούς ( $<1 \text{ cm}^2$ ) παρουσίασαν  $\text{TPR}=0.89$  με  $0.5 \text{ FPI}$ . Στους όγκους μεσαίου μεγέθους ( $<4 \text{ cm}^2$ ) καθώς και στους μεγάλους, οι τιμές κυμαίνονται στα  $0.97$  για  $0.5 \text{ FPI}$ , φανερώνοντας, συμπερασματικά, μία λίγο χειρότερη ικανότητα πρόβλεψης για τους όγκους μικρού μεγέθους.

Στην επόμενη συγκριτική έρευνα (Tsochatzidis, Costaridou, & Pratikakis, 2019) που υλοποιήθηκε στον ίδιο τομέα και βάση δεδομένων, αξιοποιήθηκαν ξανά το δίκτυο AlexNet, ενώ εκτός από το VGG16, δοκιμάστηκε και η αρχιτεκτονική του VGG19. Όσον αφορά, τα ResNet, χρησιμοποιήθηκαν τα ResNet50, ResNet101, ResNet152. Τέλος, για τα μοντέλα Inception, έλαβε χώρα η εκπαίδευση των πρώτων δύο εκδόσεων (GoogLeNet/Inception v1 και Inception v2).

Κατά την πειραματική διαδικασία, αξιολογήθηκε η ικανότητα πρόβλεψης των παραπάνω συστημάτων, χρησιμοποιώντας πρακτικές training from scratch και fine-tuning στις βάσεις δεδομένων DDSM-400 και CBIS-DDSM.

Ως βασικά βήματα προ επεξεργασίας, ειδικότερα στην εκπαίδευση from scratch, αξιοποιείται η αρχικοποίηση Glorot ενώ για το fine-tuning αξιοποιήθηκαν, τα pre-trained δίκτυα από την βάση δεδομένων ImageNet, πραγματοποιώντας τις τυπικές τροποποιήσεις στα τελικά SoftMax και αρχικά layers, ενώ για το optimization, αξιοποιήθηκε ο αλγόριθμος Adam. Ακολουθήθηκαν, ακόμη, τεχνικές data augmentation, όπως είχαν αναφερθεί παραπάνω (horizontal flipping etc.).

Όσον αφορά τους ρυθμούς εκπαίδευσης (learning rate), κυμαίνονται στο εύρος μεταξύ  $[10^{-4}, 10^{-7}]$ . Το μέγεθος του κάθε batch που αξιοποιήθηκε σε κάθε αρχιτεκτονική ποικίλλει μεταξύ 10 και 32 εικόνων.

Εδώ, παρατηρείται, ότι εφαρμόστηκε και η τεχνική του early stopping για τον περιορισμό του Validation Loss με όριο τις 15 εποχές. Σχετικά με το cross-validation, αξιοποιήθηκε 10-fold.

Τα αποτελέσματα της έρευνας, για την from scratch εκπαίδευση, φαίνονται παρακάτω:

CNN	DDSM-400		CBIS-DDSM	
	AUC	ACC	AUC	ACC
AlexNet	0.657	0.610	0.716	0.656
VGG-16	0.621	0.590	0.702	0.580
VGG-19	0.644	0.588	0.707	0.581
ResNet-50	0.595	0.548	0.637	0.627
ResNet-101	0.637	0.588	0.641	0.662
ResNet-152	0.596	0.543	0.609	0.647
GoogLeNet	0.580	0.569	0.590	0.598
Inception-BN (v2)	0.652	0.590	0.577	0.654

Εικόνα 47: From Scratch Training Evaluation

Αντίστοιχα, για την διαδικασία του fine-tuning, παρουσιάζεται η απόδοση των pre-trained νευρωνικών δικτύων παρακάτω:

CNN	DDSM-400		CBIS-DDSM	
	AUC	ACC	AUC	ACC
AlexNet	0.805	0.733	0.802	0.753
VGG-16	0.844	0.748	0.781	0.716
VGG-19	0.835	0.738	0.783	0.736
ResNet-50	0.856	0.743	0.804	0.749
ResNet-101	0.859	0.785	0.791	0.753
ResNet-152	0.786	0.630	0.793	0.755
GoogLeNet	0.830	0.758	0.767	0.720
Inception-BN (v2)	0.850	0.780	0.774	0.747

Εικόνα 48: Pre-trained CNNs

Εδώ, λοιπόν, παρατηρούνται οι βέλτιστες αποδόσεις για τα μοντέλα ResNet50, ResNet101 και το AlexNet.

Σε άλλη έρευνα, μελετάται, αποκλειστικά, η ικανότητα πρόβλεψης της αρχιτεκτονικής του συνελκτικού νευρωνικού δικτύου ResNet50 (Agarwal, Diaz, Martí, & Llado, 2018). Εδώ, συγκρίνεται για άλλη μία φορά, οι αποδόσεις μεταξύ ενός νευρωνικού δικτύου που διέρχεται από την διαδικασία του from scratch training με μία pre-trained αρχιτεκτονική στην ImageNet βάση δεδομένων. Η εκτιμητική τους ικανότητα επικεντρώνεται στα σύνολα DDSM-CBIS και INbreast.

Αναλυτικά, χρησιμοποιώντας ως πρακτική προ επεξεργασίας την μέθοδο Otsu, διαφοροποιείται, αρχικά, η περιοχή του στήθους από το background της εικόνας της μαστογραφίας. Επιπλέον, θεωρώντας γνωστό, πλέον, ότι οι είσοδοι που απαιτούνται έχουν την μορφή RGB εικόνας διαστάσεων 224X224, πραγματοποιείται το κατάλληλο rescaling καθώς και replicating του gray καναλιού και για τις τρεις εισόδους. Όσον αφορά τον αλγόριθμο βελτιστοποίησης και εδώ αξιοποιείται ο Adam με μέγεθος batch 128 ενώ το early stopping περιορίζεται στις 10 εποχές.

Όπως παρατηρείται παρακάτω, με την χρήση Evaluation Metrics όπως το AUC και το Accuracy στο Testing Dataset, το pre-trained νευρωνικό δίκτυο, παρουσιάζει βελτιωμένη ικανότητα πρόβλεψης:

Dataset	Weights	Testing accuracy on CBIS-DDSM	AUC
CBIS-DDSM	random	0.75	0.88
	ImageNet	0.84	0.92

Εικόνα 49: Pretrained/From Scratch ResNet50

Ακολουθώντας παρόμοια λογική με τις πειραματικές διαδικασίες της (Richa, Oliver, Xavier, Moi, & Robert, 2019), αξιολογήθηκε η ικανότητα πρόβλεψης των παραπάνω δύο πρακτικών εκπαίδευσης του ResNet50 στη βάση δεδομένων INbreast, με την χρήση του fine-tuning σε αυτή, σε μηδενικό (καμία αλλαγή), μικρό (αλλαγή των Fully-connected SoftMax επιπέδων και freezing των συνελκτικών) καθώς και ολική ενημέρωση των βαρών (weights) όλων των επιπέδων του νευρωνικού δικτύου, με την τελευταία τακτική να αποδεικνύεται η καλύτερη, όπως φαίνεται παρακάτω:

Training	Testing accuracy on INbreast	AUC	Total positive #231		Total neg #10,663	
			TP	FN	TN	FP
CBIS-DDSM	0.81	0.81	149	82	8,719	1,944
FC layer	0.86	0.93	186	45	9,233	1,430
All layer	0.92	0.98	214	17	9,820	843

Εικόνα 50: Fine Tuning ResNet50

Page Intentionally Left Blank

# Κεφάλαιο 7ο

## Κατασκευή SingleView Συστήματος

### 7.1 DDSM

Η βάση δεδομένων CBIS-DDSM αποτέλεσε, την κύρια πηγή εκπαίδευσης των νευρωνικών δικτύων, που έλαβε χώρα στην μεταπτυχιακή εργασία. Η **Curated Breast Imaging Subset of DDSM** (CBIS-DDSM) αποτελεί, εξορισμού ένα υποσύνολο της **Digital Database for Screening Mammography** (DDSM).

Η τελευταία, αναπτύχθηκε με βασικό προσανατολισμό την διευκόλυνση της ανάπτυξης αλγορίθμων και υπολογιστικών συστημάτων για την υποστήριξη και ενίσχυση της πιθανότητας επιτυχημένων διαγνώσεων μαστογραφικών προβολών, από τους ειδικούς. Ως δευτερεύων, αναφέρεται και η πιθανή χρήση του συγκεκριμένου σετ ως μέσο εκπαίδευσης διδακτικού και ιατρικού προσωπικού. Περιλαμβάνει 2500 διαφορετικές περιπτώσεις ασθενών, με έκαστη να φέρει 2 εικόνες για κάθε μαστό σε συνδυασμό με ορισμένες πληροφορίες για το άτομο, όπως ηλικία, αξιολόγηση τυχών παρατηρούμενων ανωμαλιών κ.α. Σε αυτές περιλαμβάνονται περιπτώσεις καλοηθών, κακοηθών όγκων αλλά και υγιείς, όλες διαγνωσμένες από ειδικούς με επιβεβαιωμένες πληροφορίες για την εκάστοτε παθολογία.

Σε συντριπτικό βαθμό, οι περιπτώσεις προέρχονται από τα προγράμματα μαστογραφίας του Γενικού Νοσοκομείου της Μασσαχουσέτης (MGH) και της σχολής Ιατρικής του Πανεπιστημίου του Wake Forest (WFUSM). Με βάση φυλετικά κριτήρια, παρουσιάζονται τα παρακάτω στατιστικά του πληθυσμού των ασθενών:

	MGH	WFUSM
Asian	2.06	0.2
Black	4.12	20.4
Spanish	6.55	1.8
American Indian	0.00	0.1
Other	0.75	0.1
Unknown	30.34	0.3
White	56.18	77.0

Αξίζει να αναφερθεί εδώ, ότι το παραπάνω πρότζεκτ αποτέλεσε κύριο μέλημα του Γενικού Νοσοκομείου της Μασσαχουσέτης, του Πανεπιστημίου της Νότιας Φλόριντα και των Εθνικών Εργαστηρίων της Σάντια (Αλμπουκέρκη, Νέο Μεξικό). Τέλος, η υλοποίηση του έγινε υπό την αιγίδα και χρηματοδότηση του ερευνητικού προγράμματος των ερευνητικών τμημάτων ιατρικής του Αμερικάνικου Στρατού για τον καρκίνο του μαστού.

## 7.2 CBIS-DDSM

Η CBIS-DDSM αποτελεί μία ενημερωμένη και τυποποιημένη έκδοση της DDSM. Το υποσύνολο των περιπτώσεων, προέκυψε μέσω της επιμέλειας ενός αντίστοιχου ειδικού. Η ενημέρωση των δεδομένων, μπορεί, ακόμη, να εντοπιστεί στον επανακαθορισμό των περιοχών ενδιαφέροντος (Region Of Interest – ROI) και των πληροφοριών παθολογίας που συνοδεύουν τους ασθενείς.

Συγκεκριμένα, στην βάση δεδομένων CBIS-DDSM παρατηρούνται ενημερωμένες και βελτιωμένες εκδόσεις εικόνων Region Of Interest (ROI) μέσω του κατάλληλου που έχει πραγματοποιηθεί από τους ίδιους τους ερευνητές, γεγονός που τονίζουν και οι ίδιοι καθώς η ποιότητα των αντίστοιχων ROI στην DDSM ήταν σαφώς μικρότερη, σε σχέση με την υψηλή ποιότητα των προαναφερόμενων. (Song, 2009)

Συγκεκριμένα, οι ενέργειες που πραγματοποιήθηκαν είναι οι εξής:

1. Αφαίρεση περιπτώσεων με αμφίβολες μάζες: Στο review που πραγματοποιήσε ο ειδικός, εντοπίστηκαν 254 προβολές όπου οι όγκοι δεν ήταν ξεκάθαρα διακριτοί, με αποτέλεσμα να απουσιάζουν από την ενημερωμένη βάση δεδομένων
2. Αποσυμπίεση εικόνων: Στην αρχική βάση δεδομένων, οι εικόνες φέραν την μορφή Lossless JPEG (LJPEG) που, πλέον, θεωρείται obsolete. Αναπτύσσοντας κώδικα σε python, οι ερευνητές ήταν σε θέση να μετατρέψουν την συγκεκριμένη μορφή πρώτα σε TIFF και στη συνέχεια σε dicom (Digital Imaging And Communications in Medicine) αρχεία, μία τυποποιημένη μορφή για την μετάδοση και διαχείριση ιατρικών δεδομένων. Με την συγκεκριμένη διαδικασία επιτυγχάνεται μηδενική απώλεια πληροφοριών κατά τους μετασχηματισμούς.
3. Προεπεξεργασία: Όπως έχει ήδη προαναφερθεί, εκτός από τις επιλεγμένες, αρχικές, προβολές μαστογραφίας, μέσω της ενημέρωσης των ROI μασκών παρέχονται τόσο οι ίδιες, όσο και τμήματα της αρχικής εικόνας με μικρότερο μέγεθος, που αφορούν την κυρίως περιοχή ενδιαφέροντος, ως προς την αναγνώριση του είδους του καρκίνου. Εδώ, πρέπει να τονιστεί ότι καινούργιοι αλγόριθμοι, για τις περιπτώσεις των μαζών, αξιοποιήθηκαν με σκοπό την βελτίωση των περιοχών που λαμβάνοντα υπόψιν, σε σχέση με τα annotations της DDSM.
4. Τυποποιημένος διαχωρισμός Σετ: Το σύνολο των περιπτώσεων, χωρίστηκε και παρέχεται με τον κλασσικό διαχωρισμό 80/20, ξεχωριστά για τις περιπτώσεις των δυσκαμνίων και των μαζών.

Το σύνολο των ασθενών, στο νέο σύνολο δεδομένων ανέρχεται στους 1566 , με 6775 μελέτες περιπτώσεων, 10239 προβολών μαστογραφίας και συνολικού μεγέθους 163.6 GB.

Ο κύριος λόγος που πραγματοποιήθηκε η συγκεκριμένη επανέκδοση της βάσης δεδομένων, αφορά την ικανότητα επαλήθευσης/σύγκρισης των αποτελεσμάτων των εκάστοτε αλγορίθμων και υπολογιστικών συστημάτων που αναπτύσσονται από τους ερευνητές στα αντίστοιχα εκπαιδευτικά ιδρύματα. Συγκεκριμένα, λόγω της απουσίας ύπαρξης σημαντικού αριθμού τυποποιημένων, γενικευμένων τεστ για το εκάστοτε σύστημα, δεν ήταν δυνατόν στην κοινότητα που ασχολείται με την ανάλυση

μαστογραφιών, να μπορεί να τίθεται σε γενικό έλεγχο, παρά σε απομονωμένη αξιολόγηση σε ιδιωτικές βάσεις δεδομένων.

Σε συνδυασμό με τα παραπάνω, ακόμη και στα σύνολα στοιχείων που ήταν διαθέσιμα στο ευρύ κοινό όπως οι βάσεις DDSM, η Mammographic Imaging Analysis Society (MIAS) και το πρότζεκτ Image Retrieval in Medical Applications (IRMA), υπήρχαν προβλήματα. Αναλυτικά, ειδικότερα για την DDSM, λόγω του ότι δημιουργήθηκε το 1997, υπήρχαν σημαντικές αμφιβολίες για το κατά πόσο οι μη-τυποποιημένες μορφές που είχαν αποθηκευτεί οι εικόνες αλλά και τα ROIs που είχαν καταγραφεί, ήταν σε θέση να παρέχουν, στο σύνολό τους, αξιόπιστες πληροφορίες, εισάγοντας αυτές, ως έχουν σε νέα προγράμματα και αλγορίθμους. Ως αποτέλεσμα, η πλειονότητα των ερευνητών προχωρούσε στην χρήση ιδιωτικών τεχνικών και αλγορίθμων ώστε να μπορούν να προβούν σε διορθώσεις, αφαιρώντας, ουσιαστικά οποιαδήποτε δυνατό γενικευμένο έλεγχο.

### 7.3 Διαδικασίες Προεπεξεργασίας

Η προεπεξεργασία, στην συγκεκριμένη διπλωματική εργασία, είναι δυνατόν να χωριστεί σε δύο βασικές κατηγορίες, την μία να αφορά τον κατάλληλο μετασχηματισμό της αρχικής δομής της βάσης δεδομένων και την άλλη τις διαδικασίες που πρέπει να λάβουν χώρα ώστε οι εικόνες να είναι σε θέση να τεθούν ως είσοδοι στις αρχιτεκτονικές νευρωνικών δικτύων.

#### 7.3.1 Μετασχηματισμός Πινάκων

Με την ολοκλήρωση της λήψης δεδομένων, στα χέρια του φοιτητή υπήρχε μία λίστα φακέλων, με διαφορετικές ονομασίες που αφορούσαν, τον ασθενή, τον μαστό που μελετάται, και αρχεία για την κάθε προβολή, συμπεριλαμβανομένων της ολόκληρης μαστογραφικής προβολής (image file), της ROI μάσκας (ROI Mask) και μίας εικόνας που είχε προκύψει από την περικοπή της αρχικής (cropped image file), λαμβάνοντας υπόψιν την περιοχή ενδιαφέροντος.

Σε συνδυασμό με τα παραπάνω, στην ιστοσελίδα όπου φιλοξενείται το παραπάνω σετ δεδομένων, υπήρχαν και άλλα 4 αρχεία της μορφής csv, τα οποία διαδραματίζουν επεξηγηματικό ρόλο στο σετ δεδομένων. Συγκεκριμένα:

- Calc\_case\_description\_test\_set: Περιγραφή του σετ δεδομένων που θα χρησιμοποιηθεί για την αξιολόγηση του νευρωνικού δικτύου, ως προς την δυσκαμψία στήθους.
- Calc\_case\_description\_train\_set: Περιγραφή του σετ δεδομένων που διαδραματίζει τον ρόλο του πακέτου δεδομένων που χρησιμοποιείται για την εκπαίδευση του αντίστοιχου συστήματος, για την περίπτωση της δυσκαμψίας στήθους.
- Mass\_case\_description\_test\_set: Αντίστοιχα, εδώ για την εύρεση μαζών.
- Mass\_case\_description\_train\_set: Ομοίως, για την προαναφερόμενη, απαραίτητη εκπαίδευση που πρέπει να πραγματοποιηθεί.



Να σημειωθεί εδώ, ότι για τον διαχωρισμό των δεδομένων σε σετ εκπαίδευσης/τεστ, σύμφωνα με την ιστοσελίδα, πραγματοποιήθηκε ένας κλασσικός διαχωρισμός 80-20.

Σε κάθε αρχείο περιλαμβάνεται, με διαφορετικές, προφανώς, τιμές η ίδια μορφή πίνακα που παρουσιάζεται παρακάτω:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	
1	patient_id	breast_de	left or right	image vie	abnormal	abnormal	mass shap	mass mar	assessment	pathology	subtlety	image file	cropped i	ROI mask	file path	
2	P_00001	3	LEFT	CC		1 mass	IRREGULA	SPICULATI		4	MALIGNA		4	Mass-Trail	Mass-Trail	Mass-
3	P_00001	3	LEFT	MLO		1 mass	IRREGULA	SPICULATI		4	MALIGNA		4	Mass-Trail	Mass-Trail	Mass-
4	P_00004	3	LEFT	CC		1 mass	ARCHITEC	ILL_DEFIN		4	BENIGN		3	Mass-Trail	Mass-Trail	Mass-
5	P_00004	3	LEFT	MLO		1 mass	ARCHITEC	ILL_DEFIN		4	BENIGN		3	Mass-Trail	Mass-Trail	Mass-
6	P_00004	3	RIGHT	MLO		1 mass	OVAL	CIRCUMSC		4	BENIGN		5	Mass-Trail	Mass-Trail	Mass-
7	P_00009	3	RIGHT	CC		1 mass	OVAL	ILL_DEFIN		4	MALIGNA		4	Mass-Trail	Mass-Trail	Mass-
8	P_00009	3	RIGHT	MLO		1 mass	OVAL	ILL_DEFIN		4	MALIGNA		4	Mass-Trail	Mass-Trail	Mass-
9	P_00015	3	LEFT	MLO		1 mass	IRREGULA	ILL_DEFIN		4	MALIGNA		4	Mass-Trail	Mass-Trail	Mass-
10	P_00018	2	RIGHT	CC		1 mass	OVAL	CIRCUMSC		4	BENIGN		5	Mass-Trail	Mass-Trail	Mass-
11	P_00018	2	RIGHT	MLO		1 mass	OVAL	CIRCUMSC		4	BENIGN		5	Mass-Trail	Mass-Trail	Mass-
12	P_00021	1	LEFT	CC		1 mass	OVAL	CIRCUMSC		3	BENIGN		5	Mass-Trail	Mass-Trail	Mass-
13	P_00021	1	LEFT	MLO		1 mass	OVAL	CIRCUMSC		3	BENIGN		5	Mass-Trail	Mass-Trail	Mass-
14	P_00021	1	RIGHT	CC		1 mass	OVAL	CIRCUMSC		3	BENIGN		5	Mass-Trail	Mass-Trail	Mass-
15	P_00021	1	RIGHT	MLO		1 mass	OVAL	CIRCUMSC		3	BENIGN		5	Mass-Trail	Mass-Trail	Mass-
16	P_00023	1	RIGHT	CC		1 mass	IRREGULA	SPICULATI		5	MALIGNA		4	Mass-Trail	Mass-Trail	Mass-
17	P_00023	1	RIGHT	MLO		1 mass	IRREGULA	SPICULATI		5	MALIGNA		4	Mass-Trail	Mass-Trail	Mass-
18	P_00026	2	LEFT	CC		1 mass	OVAL	OBSCURE		3	BENIGN		3	Mass-Trail	Mass-Trail	Mass-
19	P_00026	2	LEFT	MLO		1 mass	OVAL	OBSCURE		3	BENIGN		3	Mass-Trail	Mass-Trail	Mass-
20	P_00027	2	RIGHT	CC		1 mass	OVAL	ILL_DEFIN		4	BENIGN		3	Mass-Trail	Mass-Trail	Mass-
21	P_00027	2	RIGHT	MLO		1 mass	OVAL	CIRCUMSC		4	BENIGN		4	Mass-Trail	Mass-Trail	Mass-
22	P_00034	3	RIGHT	CC		1 mass	IRREGULA	SPICULATI		5	MALIGNA		4	Mass-Trail	Mass-Trail	Mass-
23	P_00034	3	RIGHT	MLO		1 mass	IRREGULA	SPICULATI		5	MALIGNA		4	Mass-Trail	Mass-Trail	Mass-
24	P_00039	2	RIGHT	CC		1 mass	IRREGULA	SPICULATI		5	MALIGNA		5	Mass-Trail	Mass-Trail	Mass-
25	P_00039	2	RIGHT	MLO		1 mass	IRREGULA	SPICULATI		5	MALIGNA		5	Mass-Trail	Mass-Trail	Mass-
26	P_00041	1	LEFT	CC		1 mass	LYMPH_NI	CIRCUMSC		2	BENIGN_V		5	Mass-Trail	Mass-Trail	Mass-
27	P_00041	1	LEFT	MLO		1 mass	LOBULATE	CIRCUMSC		2	BENIGN_V		5	Mass-Trail	Mass-Trail	Mass-
28	P_00044	2	RIGHT	CC		1 mass	OVAL	CIRCUMSC		3	BENIGN		5	Mass-Trail	Mass-Trail	Mass-
29	P_00044	2	RIGHT	CC		2 mass	OVAL	CIRCUMSC		3	BENIGN		5	Mass-Trail	Mass-Trail	Mass-
30	P_00044	2	RIGHT	CC		3 mass	OVAL	CIRCUMSC		3	BENIGN		5	Mass-Trail	Mass-Trail	Mass-
31	P_00044	2	RIGHT	CC		4 mass	LOBULATE	CIRCUMSC		3	BENIGN		5	Mass-Trail	Mass-Trail	Mass-
32	P_00044	2	RIGHT	MLO		1 mass	LOBULATE	CIRCUMSC		3	BENIGN		5	Mass-Trail	Mass-Trail	Mass-
33	P_00044	2	RIGHT	MLO		2 mass	OVAL	CIRCUMSC		3	BENIGN		5	Mass-Trail	Mass-Trail	Mass-
34	P_00045	4	LEFT	CC		1 mass	FOCAL_AS	ILL_DEFIN		4	MALIGNA		3	Mass-Trail	Mass-Trail	Mass-
35	P_00045	4	LEFT	MLO		1 mass	FOCAL_AS	ILL_DEFIN		4	MALIGNA		3	Mass-Trail	Mass-Trail	Mass-
36	P_00046	3	RIGHT	MLO		1 mass	IRREGULA	OBSCURE		3	MALIGNA		1	Mass-Trail	Mass-Trail	Mass-
37	P_00051	2	LEFT	CC		1 mass	IRREGULA	SPICULATI		5	MALIGNA		5	Mass-Trail	Mass-Trail	Mass-
38	P_00051	2	LEFT	MLO		1 mass	IRREGULA	SPICULATI		5	MALIGNA		5	Mass-Trail	Mass-Trail	Mass-

Εικόνα 51: Mass\_case\_train\_set

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	
	patient_id	breast_de	left or right	image view	abnormal	abnormal	calc type	calc distrib	assessment	pathology	subtlety	image file	cropped image	ROI mask	file path
	P_00005	3	RIGHT	CC	1	calcificati	AMORPHC CLUSTER		3	MALIGNA		3	Calc-Train Calc-	Calc-Training_P_00005_RIG	
	P_00005	3	RIGHT	MLO	1	calcificati	AMORPHC CLUSTER		3	MALIGNA		3	Calc-Train Calc-	Calc-Training_P_00005_RIG	
	P_00007	4	LEFT	CC	1	calcificati	PLEOMOR LINEAR		4	BENIGN		4	Calc-Train Calc-	Calc-Training_P_00007_LEF	
	P_00007	4	LEFT	MLO	1	calcificati	PLEOMOR LINEAR		4	BENIGN		4	Calc-Train Calc-	Calc-Training_P_00007_LEF	
	P_00008	1	LEFT	CC	1	calcificati	N/A	REGIONAL	2	BENIGN_V		3	Calc-Train Calc-	Calc-Training_P_00008_LEF	
	P_00008	1	LEFT	CC	2	calcificati	N/A	REGIONAL	2	BENIGN_V		3	Calc-Train Calc-	Calc-Training_P_00008_LEF	
	P_00008	1	LEFT	CC	3	calcificati	N/A	REGIONAL	2	BENIGN_V		3	Calc-Train Calc-	Calc-Training_P_00008_LEF	
	P_00008	1	LEFT	MLO	1	calcificati	N/A	REGIONAL	2	BENIGN_V		3	Calc-Train Calc-	Calc-Training_P_00008_LEF	
	P_00008	1	LEFT	MLO	2	calcificati	N/A	REGIONAL	2	BENIGN_V		3	Calc-Train Calc-	Calc-Training_P_00008_LEF	
	P_00008	1	LEFT	MLO	3	calcificati	N/A	REGIONAL	2	BENIGN_V		3	Calc-Train Calc-	Calc-Training_P_00008_LEF	
	P_00008	1	RIGHT	CC	1	calcificati	N/A	REGIONAL	2	BENIGN_V		3	Calc-Train Calc-	Calc-Training_P_00008_RIG	
	P_00008	1	RIGHT	CC	2	calcificati	N/A	REGIONAL	2	BENIGN_V		3	Calc-Train Calc-	Calc-Training_P_00008_RIG	
	P_00008	1	RIGHT	CC	3	calcificati	N/A	REGIONAL	2	BENIGN_V		3	Calc-Train Calc-	Calc-Training_P_00008_RIG	
	P_00008	1	RIGHT	CC	4	calcificati	N/A	REGIONAL	2	BENIGN_V		3	Calc-Train Calc-	Calc-Training_P_00008_RIG	
	P_00008	1	RIGHT	CC	5	calcificati	N/A	REGIONAL	2	BENIGN_V		3	Calc-Train Calc-	Calc-Training_P_00008_RIG	
	P_00008	1	RIGHT	MLO	1	calcificati	N/A	REGIONAL	2	BENIGN_V		3	Calc-Train Calc-	Calc-Training_P_00008_RIG	
	P_00008	1	RIGHT	MLO	2	calcificati	N/A	REGIONAL	2	BENIGN_V		3	Calc-Train Calc-	Calc-Training_P_00008_RIG	
	P_00008	1	RIGHT	MLO	3	calcificati	N/A	REGIONAL	2	BENIGN_V		3	Calc-Train Calc-	Calc-Training_P_00008_RIG	
	P_00008	1	RIGHT	MLO	4	calcificati	N/A	REGIONAL	2	BENIGN_V		3	Calc-Train Calc-	Calc-Training_P_00008_RIG	
	P_00008	1	RIGHT	MLO	5	calcificati	N/A	REGIONAL	2	BENIGN_V		3	Calc-Train Calc-	Calc-Training_P_00008_RIG	
	P_00010	3	LEFT	CC	1	calcificati	ROUND_A DIFFUSELY		2	BENIGN_V		4	Calc-Train Calc-	Calc-Training_P_00010_LEF	
	P_00010	3	LEFT	MLO	1	calcificati	ROUND_A DIFFUSELY		2	BENIGN_V		4	Calc-Train Calc-	Calc-Training_P_00010_LEF	
	P_00011	3	LEFT	CC	1	calcificati	PLEOMOR CLUSTER		4	BENIGN		1	Calc-Train Calc-	Calc-Training_P_00011_LEF	
	P_00011	3	LEFT	MLO	1	calcificati	PLEOMOR CLUSTER		4	BENIGN		1	Calc-Train Calc-	Calc-Training_P_00011_LEF	
	P_00012	2	LEFT	CC	1	calcificati	PLEOMOR CLUSTER		4	MALIGNA		4	Calc-Train Calc-	Calc-Training_P_00012_LEF	
	P_00012	2	LEFT	MLO	1	calcificati	PLEOMOR CLUSTER		4	MALIGNA		3	Calc-Train Calc-	Calc-Training_P_00012_LEF	
	P_00013	4	RIGHT	MLO	1	calcificati	PLEOMOR SEGMENT		4	BENIGN		4	Calc-Train Calc-	Calc-Training_P_00013_RIG	
	P_00014	4	LEFT	CC	1	calcificati	PLEOMOR CLUSTER		4	MALIGNA		4	Calc-Train Calc-	Calc-Training_P_00014_LEF	
	P_00014	4	LEFT	MLO	1	calcificati	PLEOMOR CLUSTER		4	MALIGNA		4	Calc-Train Calc-	Calc-Training_P_00014_LEF	
	P_00016	4	LEFT	CC	2	calcificati	PLEOMOR CLUSTER		5	MALIGNA		5	Calc-Train Calc-	Calc-Training_P_00016_LEF	
	P_00016	4	LEFT	MLO	2	calcificati	PLEOMOR CLUSTER		5	MALIGNA		5	Calc-Train Calc-	Calc-Training_P_00016_LEF	
	P_00019	4	RIGHT	MLO	1	calcificati	PLEOMOR CLUSTER		4	BENIGN		2	Calc-Train Calc-	Calc-Training_P_00019_RIG	
	P_00019	4	RIGHT	CC	1	calcificati	PLEOMOR CLUSTER		4	BENIGN		2	Calc-Train Calc-	Calc-Training_P_00019_RIG	
	P_00020	2	LEFT	CC	1	calcificati	PLEOMOR CLUSTER		4	MALIGNA		4	Calc-Train Calc-	Calc-Training_P_00020_LEF	
	P_00020	2	LEFT	MLO	1	calcificati	PLEOMOR CLUSTER		4	MALIGNA		4	Calc-Train Calc-	Calc-Training_P_00020_LEF	
	P_00022	3	LEFT	CC	1	calcificati	PLEOMOR CLUSTER		4	BENIGN		3	Calc-Train Calc-	Calc-Training_P_00022_LEF	
	P_00022	3	LEFT	MLO	1	calcificati	AMORPHC CLUSTER		4	BENIGN		3	Calc-Train Calc-	Calc-Training_P_00022_LEF	

Εικόνα 52: Calc\_case\_train\_set



Παρατηρείται, λοιπόν, ότι εκτός από, προφανώς, την 5<sup>η</sup>, 6<sup>η</sup>, 7<sup>η</sup> και 8<sup>η</sup> στήλη που σχετίζεται με την αντίστοιχη περίπτωση κατηγοριοποίησης που μελετάται, τα υπόλοιπα πεδία ανήκουν στην ίδια κατηγορία. Αυτές οι ταξινομήσεις, αναλύονται παρακάτω:

- Patient\_id: Ένας αριθμός που έχει δοθεί σε ένα συμμετέχοντα προκειμένου να αναγνωρίζεται από το σύστημα αλλά και να προστατευθούν ευαίσθητα προσωπικά δεδομένα.
- Breast density: Πυκνότητα των ιστών του μαστού στην συγκεκριμένη μαστογραφία.
- Left or Right: Ποιος από τους δύο μαστούς μελετάται
- Image View: Αν η μαστογραφία ανήκει στην cc ή την mlo προβολή
- Assessment: Η αξιολόγηση της σοβαρότητας, από ειδικευμένο ογκολόγο, σύμφωνα με τις πηγές.
- Pathology: Η κατηγοριοποίησή της περίπτωσης σε καλοήγη, κακοήγη κτλ.
- Image file path: Η διεύθυνση στον σκληρό δίσκο του υπολογιστή, όπου βρίσκεται ολόκληρη η εικόνα της μαστογραφίας
- Cropped image file path: Σε αρκετά συστήματα υποβοήθησης προβλέψεων (CAD), τοποθετούνται ως ανώτερο όριο, συγκεκριμένες διαστάσεις ενώ ταυτόχρονα, μελετούν μόνο την περιοχή ενδιαφέροντος (ROI) και όχι ολόκληρη την συγκεκριμένη προβολή της μαστογραφίας. Για αυτόν τον λόγο, η ενημερωμένη έκδοση της βάσης δεδομένων DDSM, έχει προνοήσει για την παραχώρηση των αντίστοιχων δεδομένων, σε μικρότερα μεγέθη που απεικονίζουν, αμιγώς, τις ROI, με την αντίστοιχη abnormality να βρίσκεται πάντα στο κέντρο του τετραγώνου που παρέχεται ως εικόνα που έχει δεχθεί περικοπή. Εδώ, παρέχεται η αντίστοιχη θέση στην μνήμη του υπολογιστή
- ROI mask file path: Η διεύθυνση στον σκληρό δίσκο του υπολογιστή, όπου βρίσκεται η μάσκα/φίλτρο που χρησιμοποιήθηκε για την παραγωγή της παραπάνω ROI εικόνας. Παρόλα αυτά για λόγους οικονομίας και όπως παρατηρείται από την βιβλιογραφία, δεν αξιοποιείται η συγκεκριμένη μεταβλητή.

Είναι προφανές, ότι δεν είναι δυνατόν με την συγκεκριμένη μορφή που παρέχεται, αρχικά, το συγκεκριμένο αρχείο που λαμβάνεται από το διαδίκτυο, να γνωρίζει τις επακριβείς θέσεις των δεδομένων, καθώς τα paths φέρουν dummy τιμές.

Όμως, η εκπαίδευση με επιτήρηση, εκτός των εικόνων, χρειάζεται τα αντίστοιχα δεδομένα για τις κατηγορίες στις οποίες θέλουμε να εκτεθεί με σκοπό να μπορεί στη συνέχεια να αναγνωρίσει καινούργιες προβολές. Έτσι, τα παραπάνω αρχεία φαντάζουν, ιδιαίτερος βολικός τρόπος λύσης, που φέρουν το μειονέκτημα, ότι δεν δείχνουν στα paths των εικόνων. Ως εναλλακτική λύση, θα μπορούσε να προταθεί η αντιγραφή όλης της βάσης δεδομένων σε έναν άλλον φάκελο με προκαθορισμένους υποφάκελους.

Έτσι, λοιπόν, είναι κατανοητό ότι το πρώτο μέλημα του development έπρεπε να επικεντρωθεί στον μετασχηματισμό αυτών των paths, με σκοπό, πλέον, να είναι δυνατόν να διαβαστούν τα dicom αρχεία στο MATLAB περιβάλλον.

Στο παρακάτω κομμάτι κώδικα, αφού μπαίνει,hardcoded, το path του φακέλου που περιλαμβάνει το σύνολο της βάσης δεδομένων, λαμβάνεται το σύνολο των υποφακέλων με τις αντίστοιχες υποκατηγορίες που περιλαμβάνουν ενώ δημιουργούνται ορισμένες μεταβλητές που αποθηκεύουν ορισμένα μεγέθη, όπως το σύνολο των εικόνων.

```
dataPath='C:/Users/.../CBIS-DDSM';
topLevelFolder = uigetdir(dataPath);
if topLevelFolder == 0
    return;
end
% Get list of all subfolders.
allSubFolders = genpath(topLevelFolder);
% Parse into a cell array.
remain = allSubFolders;
listOfFolderNames = {};
while true
    [singleSubFolder, remain] = strtok(remain, ';');
    if isempty(singleSubFolder)
        break;
    end
    listOfFolderNames = [listOfFolderNames singleSubFolder];
end
numberOfFolders = length(listOfFolderNames);
mass_test='mass_case_description_test_set.csv';
opts = detectImportOptions(mass_test,'NumHeaderLines',0);
A = readtable(mass_test,opts);
numberOfImages=size(A,1);
```

Παραπάνω, πραγματοποιείται κατάλληλες αλλαγές, όπως το parsing σε cell arrays με σκοπό να είναι δυνατόν να γίνει η αντιστοιχία μεταξύ των επιθυμητών μεταβλητών και των μορφών που διαβάζονται τα δεδομένα στην MATLAB.

Στη συνέχεια, εκτελώντας τρία for-loop για την κάθε κατηγορία (imageFile, ROIMask, croppedImageFile) είμαστε σε θέση, μετασχηματίζοντας τα strings που υπάρχουν στα paths και εκμεταλλευόμενοι patterns που υπήρχαν στους πίνακες όπως, πχ το γεγονός ότι για κάθε προβολή, το πρώτο (πριν την πρώτη \) και το τελευταίο string (μετά την τελευταία \) έδειχνε όντως στην πραγματική εικόνα, να δημιουργήσουμε τα κατάλληλα paths που να αντιστοιχούν στις σωστές εικόνες.

```
for i=1:numberOfImages
    str=A.imageFilePath(i);
    initialPath=strcat(extractBefore(str,'/'),'\\');
    initialPath=strrep(initialPath,'/','\\');
    for k = 1 : numberOfFolders
        thisFolder = listOfFolderNames{k};
        filePattern = sprintf('%s/*1-1.dcm', thisFolder);
        baseFileNames = dir(filePattern);
        numberOfImageFiles = length(baseFileNames);
        if numberOfImageFiles >= 1
            % Go through all those image files.
            for f = 1 : numberOfImageFiles
                fullFileName = fullfile(thisFolder,
baseFileNames(f).name);
                index = strfind(fullFileName, initialPath);
                foundIt = ~isempty(index);
                if(foundIt==1)
                    A.imageFilePath(i)=cellstr(fullFileName);
                end
            end
        end
    end
end
end
```

Ένα πρόβλημα που αξίζει να σημειωθεί, είναι ότι στην τελευταία ενημερωμένη έκδοση, οι θέσεις των μασκών (ROI Masks) ήταν μπερδεμένες με τις εικόνες που είχαν υποστεί περικοπή από τις αρχικές. Με manual έλεγχο παρατηρήθηκε αρκετά συχνά το συγκεκριμένο φαινόμενο, ενώ δεν φάνηκε ότι οι θέσεις των αρχικών εικόνων να παρουσιάζουν τέτοιες ανωμαλίες. Για τον λόγο αυτό, στον παρακάτω κώδικα, για τις υπόλοιπες δύο κατηγορίες έπρεπε να πραγματοποιηθεί ένας επιπλέον έλεγχος και σε περίπτωση εύρεσης λάθους, να γίνουν οι κατάλληλες αλλαγές. Συγκεκριμένα, όπως φαίνεται παρακάτω, θέτουμε ένα όριο για το ύψος της εικόνας, γύρω στο 1000,

```
for i=1:height(Test_Dataset.A)
    outputImage =
dicominfo(string(Test_Dataset.A.croppedImageFilePath(i)));
    if outputImage.Height>1000
        temp=string(Test_Dataset.A.ROIMaskFilePath(i));

Test_Dataset.A.ROIMaskFilePath(i)=Test_Dataset.A.croppedImageFileP
ath(i);
        Test_Dataset.A.croppedImageFilePath(i)=cellstr(temp);
    end
end
```

αποκλείαμε όλες τις cropped περιπτώσεις ενώ δεν υπήρχε κίνδυνος να συναντήσουμε κάποια μάσκα που να μην ικανοποιούσε την παραπάνω συνθήκη.

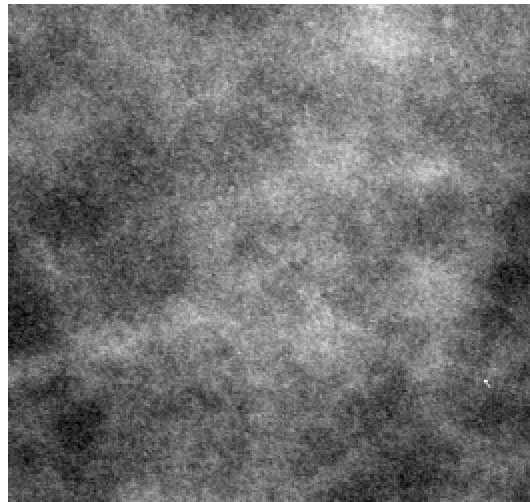
Με το πέρας της παραπάνω προ επεξεργασίας, είναι δυνατόν να διαπιστωθεί η επιτυχής αποθήκευση των ορθών διευθύνσεων για την κάθε εικόνα.

12	13
imageNamePath	croppedImageFilePath
5' C:\Users\	DDSM\Mass-Test_P_00016_LEFT_CC_1\10-04-2016-DDSM
5' C:\Users\	DDSM\Mass-Test_P_00016_LEFT_MLO_1\10-04-2016-DD
4' C:\Users\	DDSM\Mass-Test_P_00017_LEFT_CC_1\10-04-2016-DDSM
4' C:\Users\	DDSM\Mass-Test_P_00017_LEFT_MLO_1\10-04-2016-DD
2' C:\Users\	DDSM\Mass-Test_P_00032_RIGHT_CC_1\10-04-2016-DD
2' C:\Users\	DDSM\Mass-Test_P_00032_RIGHT_MLO_1\10-04-2016-DD
5' C:\Users\	DDSM\Mass-Test_P_00037_RIGHT_CC_1\10-04-2016-DD
5' C:\Users\	DDSM\Mass-Test_P_00037_RIGHT_MLO_1\10-04-2016-DD
4' C:\Users\	DDSM\Mass-Test_P_00056_LEFT_MLO_1\10-04-2016-DD
3' C:\Users\	DDSM\Mass-Test_P_00066_LEFT_CC_1\10-04-2016-DDSM
3' C:\Users\	DDSM\Mass-Test_P_00066_LEFT_MLO_1\10-04-2016-DD
4' C:\Users\	DDSM\Mass-Test_P_00099_LEFT_MLO_1\10-04-2016-DD
5' C:\Users\	DDSM\Mass-Test_P_00114_LEFT_MLO_1\10-04-2016-DD
5' C:\Users\	DDSM\Mass-Test_P_00116_RIGHT_CC_1\10-04-2016-DD
5' C:\Users\	DDSM\Mass-Test_P_00116_RIGHT_CC_2\10-04-2016-DD
5' C:\Users\	DDSM\Mass-Test_P_00116_RIGHT_MLO_1\10-04-2016-DD
5' C:\Users\	DDSM\Mass-Test_P_00116_RIGHT_MLO_2\10-04-2016-DD
5' C:\Users\	DDSM\Mass-Test_P_00118_RIGHT_CC_1\10-04-2016-DD
5' C:\Users\	DDSM\Mass-Test_P_00118_RIGHT_MLO_1\10-04-2016-DD
4' C:\Users\	DDSM\Mass-Test_P_00124_RIGHT_CC_1\10-04-2016-DD
4' C:\Users\	DDSM\Mass-Test_P_00126_RIGHT_CC_1\10-04-2016-DD
4' C:\Users\	DDSM\Mass-Test_P_00126_RIGHT_MLO_1\10-04-2016-DD
5' C:\Users\	DDSM\Mass-Test_P_00131_LEFT_CC_1\10-04-2016-DDSM
5' C:\Users\	DDSM\Mass-Test_P_00131_LEFT_MLO_1\10-04-2016-DD
4' C:\Users\	DDSM\Mass-Test_P_00145_LEFT_CC_1\10-04-2016-DDSM
4' C:\Users\	DDSM\Mass-Test_P_00145_LEFT_MLO_1\10-04-2016-DD
5' C:\Users\	DDSM\Mass-Test_P_00147_RIGHT_CC_1\10-04-2016-DD
5' C:\Users\	DDSM\Mass-Test_P_00147_RIGHT_MLO_1\10-04-2016-DD
4' C:\Users\	DDSM\Mass-Test_P_00156_RIGHT_MLO_1\10-04-2016-DD
5' C:\Users\	DDSM\Mass-Test_P_00158_RIGHT_MLO_1\10-04-2016-DD
3' C:\Users\	DDSM\Mass-Test_P_00159_RIGHT_MLO_1\10-04-2016-DD
1' C:\Users\	DDSM\Mass-Test_P_00171_RIGHT_CC_1\10-04-2016-DD
1' C:\Users\	DDSM\Mass-Test_P_00171_RIGHT_MLO_1\10-04-2016-DD
5' C:\Users\	DDSM\Mass-Test_P_00173_LEFT_CC_1\10-04-2016-DDSM
5' C:\Users\	DDSM\Mass-Test_P_00173_LEFT_MLO_1\10-04-2016-DD
5' C:\Users\	DDSM\Mass-Test_P_00173_RIGHT_CC_1\10-04-2016-DD
5' C:\Users\	DDSM\Mass-Test_P_00173_RIGHT_CC_2\10-04-2016-DD
5' C:\Users\	DDSM\Mass-Test_P_00173_RIGHT_MLO_1\10-04-2016-DD

Εικόνα 53: Τροποποιημένο σετ δεδομένων mass για test

### 7.3.2 Προετοιμασία για εκπαίδευση

Με την επιτυχημένη ανακατασκευή του σετ δεδομένων, πλέον ήταν δυνατόν να διαβαστούν dicom εικόνες και να προβληθούν στο περιβάλλον MATLAB. Το επόμενο ζήτημα αφορούσε τον τρόπο με τον οποίο θα πραγματοποιούνταν η εκπαίδευση των νευρωνικών δικτύων.



```
Format: 'DICOM'  
FormatVersion: 3  
Width: 265  
Height: 249  
BitDepth: 16  
ColorType: 'grayscale'
```

Εικόνα 54: Cropped Dicom εικόνα που προβάλλεται στο MATLAB Περιβάλλον

Είναι δυνατόν να κατασκευαστούν από την αρχή τόσο οι αλγόριθμοι για το backpropagation όσο και οι αρχιτεκτονικές των προτεινόμενων, σύμφωνα με τις έρευνες που αναφέρθηκαν παραπάνω, νευρωνικών δικτύων αλλά η συγκεκριμένη μέθοδος εκτός από χρονοβόρα ελλοχεύει κινδύνους ως προς την μη ορθή υλοποίησή της. Αντίθετα, από την στιγμή που ο τρόπος οργάνωσης των δικτύων είναι fixed, μία καλύτερη λύση εντοπίζεται στην χρήση του ToolBox της Matlab με την ονομασία Deep Learning Toolbox και την εκάστοτε λήψη της αρχιτεκτονικής του κάθε νευρωνικού δικτύου.

Έτσι, όπως φαίνεται παρακάτω, η φόρτωση μία αρχιτεκτονικής νευρωνικού δικτύου, ο διαχωρισμός των δεδομένων, η ονομασία των κατηγοριών και ορισμένες ακόμη σημαντικές αλλαγές διευκολύνονται σημαντικά.

```
[imdsTrain,imdsValidation] =  
splitEachLabel(imds,0.7,'randomized');  
numTrainImages = numel(imdsTrain.Labels);  
net=googlenet;  
inputSize = net.Layers(1).InputSize;
```

Μία πολύ βασική εντολή που συνεισφέρει στην δημιουργία βάσεων δεδομένων στην MATLAB με χαρακτηριστικά που διευκολύνουν ιδιαίτερα την εκπαίδευση νευρωνικών

δικτύων μέσω του Deep Learning Toolbox, είναι η imageDatastore, η οποία, όμως, απαιτεί την ύπαρξη φακέλων με τα αντίστοιχα labels και τις εικόνες, σωστά ταξινομημένες σε αυτές. Αυτή η δομή, δυστυχώς, δεν ταυτίζεται με την αρχιτεκτονική που έχουν δομηθεί οι φάκελοι στην δική μας περίπτωση, γεγονός που μας οδήγησε στην κατασκευή κώδικα ώστε να ικανοποιηθεί και αυτή η προϋπόθεση.

```
Train_Dataset=load('Train_Dataset.mat');  
DataLen=height(Train_Dataset.A);  
for i=1:DataLen  
    filePath=string(Train_Dataset.A.imageFilePath(i));  
    fileName=string(Train_Dataset.A.leftOrRightBreast(i));  
  
    fileMask=string(Train_Dataset.A.ROIMaskFilePath(i));  
    if ~exist(fileName, 'dir')  
        mkdir(fileName)  
    end  
end
```

Το δεύτερο πρόβλημα που αντιμετωπίστηκε, αφορούσε το γεγονός ότι τα dicom αρχεία δεν μπορούσαν να διαβαστούν με τον επιθυμητό τρόπο από την συγκεκριμένη εντολή και, εν γένει από το δίκτυο. Έτσι, θα έπρεπε να γίνει μετασχηματισμός σε μία δεκτή μορφή, αποφεύγοντας, όμως σημαντικές απώλειες στην ποιότητα της εικόνας. Έτσι, αποφασίστηκε ο μετασχηματισμός σε tiff αρχεία, που σύμφωνα με την βιβλιογραφία, ικανοποιούσαν τις παραπάνω προϋποθέσεις σε συνδυασμό με το γεγονός ότι η TIFF

```
imwrite(inputImage, fileName+'/'+ 'imageTest'+string(i)+'.tif');
```

έκδοση, χρησιμοποιήθηκε από τους ίδιους τους ερευνητές, πριν μετασχηματιστεί σε dicom, όπως προαναφέρθηκε.

### 7.3.3 Επιβεβαίωση Καλής Λειτουργίας

Πλέον, είναι δυνατόν να εισαχθούν απευθείας, οι εικόνες στα νευρωνικά δίκτυα και να γίνει η κατάλληλη κατηγοροποίησή τους. Αρχικά, για να διαπιστωθεί η καλή λειτουργία του συστήματος, συντάχθηκε κώδικας ώστε να εκπαιδευτεί ένα νευρωνικό δίκτυο πάνω στα training data της βάσης δεδομένων.

Λαμβάνοντας υπόψιν ότι χρησιμοποιούσαμε pretrained νευρωνικά δίκτυα, έπρεπε να συντάξουμε κώδικα και να πραγματοποιήσουμε συγκεκριμένες απαραίτητες αλλαγές ώστε να δουλέψει η CNN εκπαίδευση.

Αρχικά, πρέπει να αναφερθεί ότι έπρεπε να γίνει resizing των raw εικόνων με σκοπό να μπορούν να διαβαστούν από το δίκτυο. Το μέγεθος λαμβάνει είτε τις τιμές 224X224 ή 227X227 (για το AlexNet).

```
outputImage=imresize(inputImage,[224,224]);
```

Ακόμη, λοιπόν, έχοντας ήδη αναφέρει, ότι οι παραπάνω εικόνες αποτελούν Grayscale και όχι RGB, έπρεπε να πραγματοποιήσουμε την συγκεκριμένη αλλαγή. Σύμφωνα με την θεωρία και τις έρευνες που αναφέρθηκαν, αυτό μπορούσε να πραγματοποιηθεί κάνοντας replicate την grayscale εικόνα και στα τρία κανάλια του RGB όπως φαίνεται παρακάτω.

```
If numberOfColorChannels==1  
  
    inputImage=cat(3,inputImage,inputImage,inputImage);  
  
end
```

Έπειτα, ένα απαραίτητο βήμα αποτελούσε η αλλαγή των εξόδων που περίμενε το νευρωνικό σύστημα, μίας και η βάση ImageNet περιλάμβανε 1000 κατηγορίες σε αντίθεση με την δική μας περίπτωση που περιλάμβανε μόνο 2. Η συγκεκριμένη λειτουργία πραγματοποιείται στον κώδικα layerToReplace.m που μπορεί να εντοπιστεί στο τέλος του κεφαλαίου.

Επιπλέον, το γεγονός ότι τα δίκτυα ήταν pretrained, δεν αποκλείει την αναγκαιότητα μηδενισμού-reset κάποιων βαρών προκειμένου να ληφθούν υπόψιν τα καινούργια features των εικόνων που εισάγονται στο πρόβλημα. Ως περισσότερο διάσημη τακτική, συναντάται το ‘πάγωμα’ των αρχικών layers.

```
layers(1:10) = freezeWeights(layers(1:10));
```

Το συγκεκριμένο implementation μπορεί να βρεθεί στο αντίστοιχο snippet κώδικα που φαίνεται παρακάτω.

```
function layers = freezeWeights(layers)  
  
    for ii = 1:size(layers,1)  
        props = properties(layers(ii));  
        for p = 1:numel(props)  
            propName = props{p};  
            if ~isempty(regexpi(propName, 'LearnRateFactor$',  
'once'))  
                layers(ii).(propName) = 0;  
            end  
        end  
    end
```

Τέλος, απαραίτητο ήταν και ένα ακόμα script σχετικό με την δημιουργία συνδέσεων του εκάστοτε νευρωνικού δικτύου, που μπορεί να βρεθεί στο τέλος, του κεφαλαίου με το όνομα createLgraphUsingConnections.m.

Έτσι, θεωρώντας ως εισόδους τις εικόνες, έως έχουν, σε συνδυασμό με τις παραπάνω συναρτήσεις, κατασκευάστηκε κώδικας για να γίνει η εκπαίδευση ενός νευρωνικού δικτύου. Το πλήρες snippet μπορεί, κάποιος, να το βρεί στο τέλος του κεφαλαίου με το όνομα Classification.m.

Τα βασικά του μέρη, αποτελούνται, αρχικά, από τον διαχωρισμό του σε training και validation set, με τον κλασσικό κανόνα του 70:30

```
[imdsTrain,imdsValidation] =  
splitEachLabel(imds,0.7,'randomized');
```

Στην επιλογή της αντίστοιχης αρχιτεκτονικής pretrained νευρωνικού δικτύου (εδώ χρησιμοποιείται το google net):

```
net=googlenet;  
inputSize = net.Layers(1).InputSize;
```

Αυτό που πρέπει να σημειωθεί, είναι ότι εξαιτίας της χαρακτηριστικής ανάγκης των νευρωνικών δικτύων για μεγάλο όγκο δεδομένων, ένας περιορισμός που, αναφέρεται συχνά, αντιμετωπίστηκε με την χρήση της τεχνικής του augmentation τόσο μεγέθους (scaling στο 0.75 μέχρι το 1.25 της τιμής) όσο και φοράς (περιστροφή -30 και 30).

Έπειτα, πραγματοποιούνται όλες οι προαναφερόμενες ενέργειες που αφορούν το 'πάγωμα' των αρχικών layers, ενώ εισάγεται η πρώτη ενέργεια προεπεξεργασίας της ίδιας της εικόνας, που αφορά την περιστροφή της σε ένα εύρος μεταξύ -30 και 30 μοιρών

```
pixelRange = [-30 30];
```

Όσο και του μεγέθους της:

```
scaleRange = [0.9 1.1];
```

Αυτή η augmentation τεχνική συναντάται αρκετά συχνά, στην μηχανική μάθηση και αποτελεί ένα καλό τρόπο να αυξηθεί ο αριθμός των ήδη υπάρχων δεδομένων, προς εκπαίδευση.

Τέλος, γίνεται η επιλογή τόσο της optimization μεθόδου, του ρυθμού μάθησης (learning rate) και του μεγέθους των batch που θα χρησιμοποιούνται κατά την διάρκεια της εκπαίδευσης.

```
miniBatchSize = 10;  
optimizers="sgdm";  
learningrate=[1e-4,1e-7];
```

Παρόλα αυτά, όπως είναι λογικό, η τροφοδότηση της συνολικής προβολής της εκάστοτε μαστογραφίας του εκάστοτε στήθους για την αντίστοιχη, χωρίς ιδιαίτερη προεπεξεργασία, μειώνει σημαντικά την ικανότητα πρόβλεψης των εκπαιδευμένων αρχιτεκτονικών. Αυτό οφείλεται στο γεγονός, ότι το κομμάτι της πληροφορίας δεν διαχέεται σε όλη την εικόνα αλλά, εντοπίζεται σε συγκεκριμένες περιοχές της. Επιπλέον, πρέπει να τονιστεί ότι τα αρχεία ήταν σε greyscale μορφή και όχι RGB, γεγονός που συμβάλλει στο να μην μπορούν να διαβαστούν με βέλτιστο τρόπο από ένα νευρωνικό δίκτυο.

Επιπλέον, όπως και στις προηγούμενες μελέτες, εφαρμόστηκε early stopping στην περίπτωση όπου κατά την αύξηση των εποχών, το validation error παρουσίαζε αύξηση ή το accuracy έμενε αμετάβλητο για αρκετή ώρα.

Παρακάτω παρουσιάζεται η εκπαίδευση που πραγματοποιήθηκε για το σύνολο των νευρωνικών δικτύων. Επειδή, παρατηρήθηκαν παρόμοια αποτελέσματα, όπως ήταν αναμενόμενο από τις μελέτες, η εκπαίδευση πραγματοποιήθηκε σε 4 δίκτυα, το AlexNet, το Inception v1(Googlenet), το ResNet50 και ResNet101. Κατά την διάρκειά τους, δοκιμάστηκαν 3 διαφορετικοί αλγόριθμοι βελτιστοποίησης (adam, RMSprop και η Stochastic Gradient Descent with Momentum ή αλλιώς sgdm) και 4 διαφορετικοί



ρυθμοί μάθησης (από  $10^{-4}$  έως  $10^{-7}$ ), ενώ μέγεθος batchsize επιλέχθηκε ο αριθμός 10, σύμφωνα και με τις προηγούμενες μελέτες.

Τα παραπάνω οδήγησαν στα εξής αποτελέσματα:

Neural Network	Optimization Algorithms that were used	Learning Rate	Max Accuracy	Early Stopping was Applied?
AlexNet	Adam, Sgdm,RMSProp	1e-4,1e-5,1e-6,1e-7	~ 66%	Yes
GoogLeNet (Inception v1)	Adam, Sgdm,RMSProp	1e-4,1e-5,1e-6,1e-7	~ 68%	Yes
ResNet50	Adam, Sgdm,RMSProp	1e-4,1e-5,1e-6,1e-7	~ 65%	Yes
ResNet101	Adam, Sgdm,RMSProp	1e-4,1e-5,1e-6,1e-7	~ 62%	Yes

Αρχικά, λοιπόν, παρατηρείται, ιδιαίτερα χαμηλή δυνατότητα πρόβλεψης, γεγονός, όμως που είχε προβλεφθεί καθώς, δεν είχε πραγματοποιηθεί, πέραν του augmentation καμία περαιτέρω διαδικασία προεπεξεργασίας. Λαμβάνοντας υπόψιν, ότι η CBIS-DDSM αποτελούσε την μοναδική πηγή για το Training, Validation και Testing και δεδομένων των ερευνών που μελέτηθηκαν, προχωρήσαμε στην εφαρμογή/αξιολόγηση ορισμένων μεθόδων.

#### 7.3.4 Χρήση ROI Masks

Αρχικά, τα νευρωνικά δίκτυα είχαν εκπαιδευτεί με ολόκληρες τις εικόνες αποτέλεσμα να χαθεί η πληροφορία καθότι συμπεριζόταν μια μεγάλη εικόνα σε διαστάσεις που απαιτούσαν τα νευρωνικά δίκτυα και έτσι η ταξινόμηση είχε βγει περίπου σε όλες τις περιπτώσεις στο 68%. Η ορθή τακτική όμως ήταν να γίνει τμηματοποίηση της εικόνας και έπειτα να δοθεί προσοχή στην περιοχή ενδιαφέροντος που πρόκυπτε από κάθε εικόνα.

Έτσι, η προσοχή μας στρέφεται στις διαθέσιμες ROI μάσκες που παρέχονται στην βάση δεδομένων με σκοπό να αξιοποιηθούν οι πληροφορίες που διαθέτουν και να κατασκευαστούν εικόνες που θα φέρουν κυρίως τα απαιτούμενα δεδομένα για την αναγνώριση των αντίστοιχων χαρακτηριστικών.

Τα βήματα που γίνονται στο script CreateFolders.m είναι ότι πρώτα λαμβάνεται ολόκληρη η μαστογραφία, από την βάση δεδομένων CBIS-DDSM. Έπειτα, διαβάζεται η διαθέσιμη binary μάσκα για την αντίστοιχη εικόνα μαστογραφίας, από την βάση δεδομένων CBIS-DDSM. Έτσι, εφαρμόζεται η μάσκα πάνω στην εικόνα και λαμβάνεται το ROI από την μαστογραφία. Παράλληλα, χρησιμοποιείται ένα παράθυρο μεγέθους 224x224, λόγω του ότι στο GoogLeNet νευρωνικό δίκτυο, αυτό το μέγεθος έχουν οι απαιτούμενοι είσοδοι (πχ στο AlexNet είναι 227x227) με το οποίο γίνεται crop το συγκεκριμένο ROI. Τέλος, σε περίπτωση όπου το αρχικό μέγεθος του ROI είναι μικρότερο του 224x224, τότε πραγματοποιείται resizing στο αντίστοιχο μέγεθος.

Αναλυτικά, στο snippet CreateFolders.m (τέλος του κεφαλαίου), φαίνεται το διάβασμα τόσο της αρχικής dicom εικόνας, όσο και της μάσκας

```
inputImage = dicomread(filePath);  
Mask=dicomread(fileMask);  
  
okind=find(Mask>0);
```

Στη συνέχεια εντοπίζονται τα όρια της εικόνας που είναι δυνατόν να παραχθεί από την χρήση της binary μάσκας για την δημιουργία της ROI εικόνας:

```
[ii,jj]=ind2sub(size(Mask),okind);  
ymin=min(ii);  
ymax=max(ii);  
xmin=min(jj);  
xmax=max(jj);
```

Με τα συγκεκριμένα μεγέθη, είμαστε πλέον σε θέση να δημιουργήσουμε τις κατάλληλες βέλτιστες εισόδους για τα νευρωνικά μας συστήματα.

```
targetSize=[224 224];  
try  
    win1 = centerCropWindow2d(size(imCropped),targetSize);  
    croppedImage = imcrop(imCropped,win1);  
catch  
    croppedImage=imresize(imCropped,targetSize);  
end  
imwrite(croppedImage, fileName+'/'+ 'image'+string(i)+'.tif');
```

## 7.4 Εκπαίδευση SingleView Συστήματος

Καθώς, η συγκεκριμένη διαδικασία κυριώς λόγω του περιορισμένου μεγέθους δεδομένων, αποτέλεσε proof of concept, αξιοποιήθηκε ως validation, ένα μέρος του συνόλου που προορίζεται για την αξιολόγηση των CAD συστημάτων από την ίδια τη βάση.

Όπως και στις προηγούμενες εφαρμογές, παρατηρήθηκε early stopping στην περίπτωση όπου κατά την αύξηση των εποχών, το validation error παρουσίαζε αύξηση ή το accuracy έμενε αμετάβλητο για αρκετή ώρα.

Τα αποτελέσματα παρουσιάζονται παρακάτω, αντίστοιχες εικόνες από την εκπαίδευση αλλά και το εκάστοτε δίκτυο, είναι αποθηκευμένο σε μορφή .mat, με σκοπό περαιτέρω βελτίωση ή/και επαλήθευση.

Έτσι, λοιπόν, έχουμε τον παρακάτω πίνακα με τα βέλτιστα αποτελέσματα για την εκάστοτε αρχιτεκτονική:

Neural Network	Optimization Algorithm	Learning Rate	Accuracy	Early Stopping
AlexNet	Sgdm	1e-4	68.5%	Yes
GoogLeNet (Inception v1)	Adam	1e-4	74.94%	No
ResNet50	Adam	1e-4	72.53%	Yes

Η αρχιτεκτονική ResNet101 δεν επιλέχθηκε ξανά καθώς συνδύαζε υψηλό χρόνο εκπαίδευσης με μη εντυπωσιακά αποτελέσματα σε σύγκριση με τα υπόλοιπα δίκτυα.

Αυτό που πρέπει να σημειωθεί είναι ότι υπάρχει ξεκάθαρη βελτίωση στην ικανότητα εκτίμησης όλων των αρχιτεκτονικών των νευρωνικών δικτύων.

## Βιβλιογραφία 7ου Κεφαλαίου

Agarwal, R., Diaz, O., Martí, R., & Llado, X. (2018). Mass detection in mammograms using pre-trained deep learning models. Fourteenth International Workshop on Breast Imaging. Atlanta, Georgia, United States.

Cadène, R., Thome, N., & Cord, M. (2016). Master's Thesis : Deep Learning for Visual Recognition.

He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep Residual Learning for Image Recognition.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. Advances in neural information processing systems .

Richa, A., Oliver, D., Xavier, L., Moi, H., & Robert, M. (2019). Automatic mass detection in mammograms using deep convolution neural networks. Journal of Medical Imaging.

Simonyan, K., & Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., & Erhan, D. (2014). Going deeper with convolutions.

Tsochatzidis, L., Costaridou, L., & Pratikakis, I. (2019). Deep Learning for Breast Cancer Diagnosis from Mammograms - A Comparative Study. Journal of Imaging.

Zhang, X., Zou, J., He, K., & Sun, J. (2015). Accelerating Very Deep Convolutional Networks for Classification and Detection. Computer Vision and Pattern Recognition.

## Κώδικας

```
for i=1:numberOfImages
    str=A.croppedImageFilePath(i);
    initialPath=strcat(extractBefore(str,'/'),'\\');
    initialPath=strrep(initialPath,'/','\\');
    for k = 1 : numberOfFolders
        thisFolder = listOfFolderNames{k};
        filePattern = sprintf('%s/*1-1.dcm', thisFolder);
        baseFileNames = dir(filePattern);
        numberOfImageFiles = length(baseFileNames);
        if numberOfImageFiles >= 1
            % Go through all those image files.
            for f = 1 : numberOfImageFiles
                fullFileName = fullfile(thisFolder,
baseFileNames(f).name);
                index = strfind(fullFileName, initialPath);
                foundIt = ~isempty(index);
                if(foundIt==1)

A.croppedImageFilePath(i)=cellstr(fullFileName);
                    end
                end
            end
        end
    end

for i=1:numberOfImages
    str=A.ROIMaskFilePath(i);
    initialPath=strcat(extractBefore(str,'/'),'\\');
    initialPath=strrep(initialPath,'/','\\');
    for k = 1 : numberOfFolders
        thisFolder = listOfFolderNames{k};
        filePattern = sprintf('%s/*1-2.dcm', thisFolder);
        baseFileNames = dir(filePattern);
        numberOfImageFiles = length(baseFileNames);
        if numberOfImageFiles >= 1
            % Go through all those image files.
            for f = 1 : numberOfImageFiles
                fullFileName = fullfile(thisFolder,
baseFileNames(f).name);
                index = strfind(fullFileName, initialPath);
                foundIt = ~isempty(index);
                if(foundIt==1)
                    A.ROIMaskFilePath(i)=cellstr(fullFileName);
                end
            end
        end
    end
end
```

Figure 1: GroupData Τα άλλα δύο for loops

```

function [learnableLayer,classLayer] = layerToReplace(lgraph)

    if ~isa(lgraph,'nnet.cnn.LayerGraph')
        error('Argument must be a LayerGraph object.')
    end

    % Get source, destination, and layer names.
    src = string(lgraph.Connections.Source);
    dst = string(lgraph.Connections.Destination);
    layerNames = string({lgraph.Layers.Name}');

    % Find the classification layer. The layer graph must
    have a single
    % classification layer.
    isClassificationLayer = arrayfun(@(l) ...

        (isa(l,'nnet.cnn.layer.ClassificationOutputLayer')|isa(
        l,'nnet.layer.ClassificationLayer')), ...
        lgraph.Layers);

    if sum(isClassificationLayer) ~= 1
        error('Layer graph must have a single
        classification layer.')
    end
    classLayer = lgraph.Layers(isClassificationLayer);

    % Traverse the layer graph in reverse starting from the
    classification
    % layer. If the network branches, throw an error.
    currentLayerIdx = find(isClassificationLayer);
    while true

        if numel(currentLayerIdx) ~= 1
            error('Layer graph must have a single
            learnable layer preceding the classification layer.')
        end

        currentLayerType =
        class(lgraph.Layers(currentLayerIdx));
        isLearnableLayer = ismember(currentLayerType, ...

        ['nnet.cnn.layer.FullyConnectedLayer','nnet.cnn.layer.C
        onvolution2DLayer']);

        if isLearnableLayer
            learnableLayer =
            lgraph.Layers(currentLayerIdx);
            return
        end

        currentDstIdx = find(layerNames(currentLayerIdx)
        == dst);
        currentLayerIdx = find(src(currentDstIdx) ==
        layerNames);
    end

```

Εικόνα 55: *LayersToReplace.m*

```

function lgraph = createLgraphUsingConnections(layers,connections)

    lgraph = layerGraph();
    for i = 1:numel(layers)
        lgraph = addLayers(lgraph,layers(i));
    end

    for c = 1:size(connections,1)
        lgraph =
connectLayers(lgraph,connections.Source{c},connections.Destination
{c});
    end

end

```

*Figure 2: CreateLgraphUsingConnections*

```

imds=imageDatastore(pwd,'IncludeSubfolders',true,'LabelSource','fol-
ldernames');
[imdsTrain,imdsValidation] =
splitEachLabel(imds,0.7,'randomized');
numTrainImages = numel(imdsTrain.Labels);
net=googlenet;
inputSize = net.Layers(1).InputSize;
if isa(net,'SeriesNetwork')
    lgraph = layerGraph(net.Layers);
else
    lgraph = layerGraph(net);
end
[learnableLayer,classLayer] = layerToReplace(lgraph);
[learnableLayer,classLayer]
numClasses = numel(categories(imdsTrain.Labels));

if isa(learnableLayer,'nnet.cnn.layer.FullyConnectedLayer')
    newLearnableLayer = fullyConnectedLayer(numClasses, ...
        'Name','new_fc', ...
        'WeightLearnRateFactor',10, ...
        'BiasLearnRateFactor',10);

elseif isa(learnableLayer,'nnet.cnn.layer.Convolution2DLayer')
    newLearnableLayer = convolution2dLayer(1,numClasses, ...
        'Name','new_conv', ...
        'WeightLearnRateFactor',10, ...
        'BiasLearnRateFactor',10);
end

save(['googleNet_sgdm',num2str(learningrate(1)),'.mat'],'net');
lgraph =
replaceLayer(lgraph,learnableLayer.Name,newLearnableLayer);
newClassLayer = classificationLayer('Name','new_classoutput');
lgraph = replaceLayer(lgraph,classLayer.Name,newClassLayer);
figure('Units','normalized','Position',[0.3 0.3 0.4 0.4]);
plot(lgraph)
ylim([0,10])
layers = lgraph.Layers;
connections = lgraph.Connections;

layers(1:10) = freezeWeights(layers(1:10));
lgraph = createLgraphUsingConnections(layers,connections);
pixelRange = [-30 30];
scaleRange = [0.9 1.1];
imageAugmenter = imageDataAugmenter( ...
    'RandXReflection',true, ...
    'RandXTranslation',pixelRange, ...
    'RandYTranslation',pixelRange, ...
    'RandXScale',scaleRange, ...
    'RandYScale',scaleRange);
augimdsTrain = augmentedImageDatastore(inputSize(1:2),imdsTrain,
...
    'DataAugmentation',imageAugmenter);
augimdsValidation =
augmentedImageDatastore(inputSize(1:2),imdsValidation);
miniBatchSize = 10;
valFrequency = floor(numel(augimdsTrain.Files)/miniBatchSize);

```

Figure 3: Classification Part A



```

optimizers="sgdm";
learningrate=[1e-4,1e-7];
options = trainingOptions(optimizers, ...
    'MiniBatchSize',miniBatchSize, ...
    'MaxEpochs',100, ...
    'InitialLearnRate',learningrate(1), ...
    'Shuffle','every-epoch', ...
    'ValidationData',augimdsValidation, ...
    'ValidationFrequency',valFrequency, ...
    'Verbose',false, ...
    'Plots','training-progress');
net = trainNetwork(augimdsTrain,lgraph,options);
save(['googleNet_sgdm',num2str(learningrate(1)),'.mat'],'net');

```

*Figure 4: Classification Part B*

```

for i=1:DataLen

    filePath=string(Test_Dataset.A.imageFilePath(i));
    fileMask=string(Test_Dataset.A.ROIMaskFilePath(i));
    LeftRightfileName=string(Test_Dataset.A.leftOrRightBreast(i));
    CCMLofilename=string(Test_Dataset.A.imageView(i));
    Childtfilename=string(Test_Dataset.A.pathology(i));

    fileName=strcat(LeftRightfileName,CCMLofilename,'/',Childtfilename
);
    if ~exist(strcat(LeftRightfileName,CCMLofilename), 'dir')
        mkdir (strcat(LeftRightfileName,CCMLofilename));
        if ~exist(fileName,'dir')
            mkdir(fileName)
        end
    end
    if ~exist(fileName,'dir')
        mkdir(fileName)
    end

    inputImage = dicomread(filePath);
    Mask=dicomread(fileMask);

    okind=find(Mask>0);
    [ii,jj]=ind2sub(size(Mask),okind);
    ymin=min(ii);
    ymax=max(ii);
    xmin=min(jj);
    xmax=max(jj);
    imCropped=imcrop(inputImage,[xmin,ymin,xmax-xmin+1,ymax-
ymin+1]);
    [rows, columns, numberOfColorChannels] = size(imCropped);
    if numberOfColorChannels == 1
        imCropped = cat(3, imCropped, imCropped, imCropped);
    end
    targetSize=[224 224];
    try
        win1 = centerCropWindow2d(size(imCropped),targetSize);
        croppedImage = imcrop(imCropped,win1);
    catch
        croppedImage=imresize(imCropped,targetSize);
    end
    imwrite(croppedImage, fileName+'/'+ 'image'+string(i)+'.tif');
end

```

Figure 5: CreateFolders.m

Page Intentionally Left Blank

## Κεφάλαιο 8<sup>ο</sup>

Σύμφωνα, λοιπόν, με τις μελέτες αλλά και την επαλήθευση που πραγματοποιήθηκε μέσω των πειραμάτων της συγκεκριμένης ερευνητικής εργασίας, είναι δυνατόν, να εξαχθούν ασφαλή συμπεράσματα ως προς την ικανότητα πρόβλεψης των single-view CAD συστημάτων, όσον αφορά την διάγνωση του καρκίνου του μαστού.

Οι συγκεκριμένες αρχιτεκτονικές, ενώ παρουσιάζονται ικανά να διαγνώσουν ένα σημαντικό ποσοστό όγκων καθώς και να προβούν σε αναγνώριση του εκάστοτε είδους (καλοήθης-κακοήθης), απέχουν σημαντικά από το να αποτελούν, αξιόπιστα εργαλεία στην ακριβή διάγνωση όγκων από τους αντίστοιχους ειδικευμένους ιατρούς. Τα δίκτυα αυτά, ενώ μπορούν να δείξουν βελτιωμένες αποδόσεις με την χρήση βέλτιστων υπερ-παραμέτρων και με την συνεχόμενη εκπαίδευση σε πολλά και διαφορετικά, αξιόπιστα σετ δεδομένων, συνεχίζουν να αντιμετωπίζουν το σύνολο των προβολών (ipsilateral views) της μαστογραφίας ως ένα.

Αυτό το φαινόμενο, υπογραμμίζεται εντόνως από πληθώρα μελετών (Wolfson Stacey, 2018), (Jouirou, Walid, & Baazaoui, 2019), (Qian, Song, Lei, Sankar, & Eikman, 2007) ενώ ταυτόχρονα σε άλλες (Wang, et al., 2020), (T. Kooi, 2017), (P. Casti, 2016), (B. Kong, 2018) τονίζεται η σημασία της χρήσης πολλών διαφορετικών κατηγοριοποιητών (classifiers) για την εκάστοτε προβολή μαστογραφίας, καθώς είναι δυνατόν να καταγραφούν και, εν τέλει, να γίνουν αντικείμενο διάγνωσης, διαφορετικά χαρακτηριστικά ανάλογα με την γωνία του στήθους που χρησιμοποιείται. Βεβαίως, επειδή πραγματοποιείται αναφορά σε νευρωνικά δίκτυα, μία τεχνολογία που έχει γνωρίσει άνθιση την τελευταία δεκαετία, είναι αρκετά λογικό, ότι η συγκεκριμένη πρακτική δεν έχει τεθεί σε εφαρμογή στον απαιτούμενο βαθμό, ώστε να παρέχει στην επιστημονική κοινότητα περισσότερες μελέτες και εργαλεία με σκοπό την κατασκευή καλύτερων CAD συστημάτων που είναι σε θέση να προσφέρουν μεγαλύτερη και ουσιαστική βοήθεια στους ογκολόγους - ιατρούς που πραγματοποιούν διαγνωστικούς ελέγχους.

Η συγκεκριμένη ερευνητική εργασία, μετά την κατασκευή του singleview συστήματος, επιδιώκει να επαληθεύσει τα αποτελέσματα καθώς και να βελτιώσει τις προτεινόμενες αρχιτεκτονικές μίας από των πρότυπων καινοτόμων ερευνών στο τομέα της κατασκευής των Multiview CAD συστημάτων για την διάγνωση καρκινικών όγκων στην μαστογραφία (Raza, Alquhayz, & Khan, 2019).

Σύμφωνα με την ίδια, η μεμονωμένη αξιολόγηση της εκάστοτε προβολής (mlo,cc) για το αντίστοιχο στήθος του ασθενή (left, right), όπως και επίσης ο διαχωρισμός των μαστογραφιών που αφορούν τον εντοπισμό μαζών (mass-case) και του εντοπισμού δυσκαμψίας (calcification) με τον συνδυασμό της χρήσης αντίστοιχων πρακτικών προ επεξεργασίας, που έχουν αναλυθεί προηγουμένως όπως το data augmentation, το resizing και το rotating σε συγκεκριμένο εύρος τιμών μαζί με πρακτικές transfer learning, όπως εκείνες παρουσιάστηκαν στην προηγούμενη πειραματική διαδικασία είναι δυνατόν να παράγουν συστήματα με υψηλό δείκτη αξιοπιστίας, στην αναγνώριση των normal/abnormal περιπτώσεων όπως φαίνεται παρακάτω:

CNN Model	Augmentation	Testing Accuracy ( $\mu \pm \sigma$ )	Training Accuracy ( $\mu \pm \sigma$ )	Sensitivity	Specificity	AUC
VGG16	No	91.57% $\pm$ 1.40%	92.33% $\pm$ 2.10%	94.30%	88.11%	0.917
	Yes	92.13% $\pm$ 1.92%	95.89% $\pm$ 2.71%	94.66%	88.91%	0.918
VGG19	No	93.45% $\pm$ 1.56%	97.92% $\pm$ 1.88%	98.07%	88.13%	0.940
	Yes	93.54% $\pm$ 1.29%	96.13% $\pm$ 2.29%	97.25%	89.09%	0.932
InceptionV3	No	89.98% $\pm$ 2.02%	94.94% $\pm$ 1.75%	95.88%	83.53%	0.907
	Yes	90.92% $\pm$ 1.95%	98.05% $\pm$ 2.82%	96.28%	84.92%	0.906
ResNet50	No	90.92% $\pm$ 1.81%	96.84% $\pm$ 2.04%	92.65%	88.60%	0.908
	Yes	91.67% $\pm$ 1.72%	94.42% $\pm$ 1.96%	93.44%	89.30%	0.914
Proposed MVFF	No	92.70% $\pm$ 1.33%	97.05% $\pm$ 2.63%	96.56%	88.09%	0.931
Architecture	Yes	93.73% $\pm$ 1.61%	96.66% $\pm$ 2.23%	96.31%	90.47%	0.934

Εικόνα 56: Διαχωρισμός normal/abnormal περιπτώσεων

Παρόλα αυτά, επειδή το βασικό ενδιαφέρον της συγκεκριμένης εργασίας, αφορά την βάση δεδομένων CBIS-DDSM, η οποία περιλαμβάνει μόνο abnormal περιπτώσεις, η συγκεκριμένη κατηγοριοποίηση δεν ανήκει στην σφαίρα του αντικειμένου μελέτης.

Αντίθετα, όμως, αποτελεί βασικό εγχείρημα η επίτευξη της διάγνωσης των περιπτώσεων που σχετίζονται με δυσκαμψία και αυτών που αφορούν την διάγνωση υποψήφιων μαζών, που τα Multiview συστήματα της συγκεκριμένης έρευνας, παρουσίασαν και εδώ υψηλά ποσοστά:

CNN Model	Augmentation	Testing Accuracy ( $\mu \pm \sigma$ )	Training Accuracy ( $\mu \pm \sigma$ )	Sensitivity	Specificity	AUC
VGG16	No	89.48% $\pm$ 0.87%	93.26% $\pm$ 1.28%	90.16%	88.76%	0.897
	Yes	89.62% $\pm$ 0.68%	92.59% $\pm$ 1.72%	91.09%	88.14%	0.896
VGG19	No	86.26% $\pm$ 0.54%	90.43% $\pm$ 1.09%	89.74%	83.06%	0.864
	Yes	87.52% $\pm$ 0.94%	93.21% $\pm$ 1.56%	91.42%	84.00%	0.877
InceptionV3	No	86.82% $\pm$ 1.19%	88.12% $\pm$ 0.91%	86.67%	86.98%	0.868
	Yes	87.52% $\pm$ 1.36%	89.03% $\pm$ 1.23%	88.04%	86.96%	0.875
ResNet50	No	85.69% $\pm$ 0.73%	88.24% $\pm$ 1.68%	83.13%	89.03%	0.855
	Yes	85.55% $\pm$ 1.20%	90.90% $\pm$ 1.17%	82.28%	90.03%	0.862
Proposed MVFF	No	90.60% $\pm$ 0.92%	92.17% $\pm$ 1.35%	92.42%	88.80%	0.907
Architecture	Yes	92.29% $\pm$ 1.15%	95.68% $\pm$ 1.33%	93.37%	91.17%	0.923

Εικόνα 57: Mass Detection/Calcification

Τέλος, ακόμη μεγαλύτερης σημασίας εγχείρημα, αποτελεί η βελτιωμένη εκτίμηση/αξιολόγηση/διάγνωση των περιπτώσεων που περιλαμβάνονται στο συγκεκριμένο σετ δεδομένων, δηλαδή η επιτυχής ονομασία του εκάστοτε όγκου και η κατηγοριοποίησή του σε καλοήγη/κακοήγη:

CNN Model	Augmentation	Testing Accuracy ( $\mu \pm \sigma$ )	Training Accuracy ( $\mu \pm \sigma$ )	Sensitivity	Specificity	AUC
VGG16	No	73.47% $\pm$ 0.32%	80.51% $\pm$ 1.51%	76.42%	68.78%	0.721
	Yes	75.39% $\pm$ 0.98%	79.21% $\pm$ 1.72%	78.22%	70.98%	0.746
VGG19	No	69.11% $\pm$ 0.67%	75.27% $\pm$ 1.63%	71.93%	64.08%	0.671
	Yes	71.20% $\pm$ 0.55%	73.33% $\pm$ 1.34%	73.70%	66.83%	0.703
InceptionV3	No	68.24% $\pm$ 0.82%	76.12% $\pm$ 0.91%	71.43%	62.68%	0.663
	Yes	70.16% $\pm$ 1.12%	75.99% $\pm$ 1.23%	72.88%	65.38%	0.691
ResNet50	No	68.76% $\pm$ 0.25%	74.56% $\pm$ 1.68%	74.46%	61.29%	0.681
	Yes	69.98% $\pm$ 0.65%	74.08% $\pm$ 1.17%	75.46%	62.75%	0.691
Proposed MVFF	No	76.27% $\pm$ 0.89%	78.92% $\pm$ 1.35%	80.61%	70.37%	0.757
Architecture	Yes	77.66% $\pm$ 0.72%	80.56% $\pm$ 1.33%	81.82%	72.02%	0.769

Εικόνα 58: Malignant/Benign Detection

Εδώ, παρατηρούνται παρόμοιες τιμές, όπως ήταν αναμενόμενο, με αυτές που υλοποιήθηκαν στα παραπάνω πειράματα, με την προτεινόμενη Multiview αρχιτεκτονική να παρουσιάζει σαφώς βελτιωμένη περιθώρια διαγνωστικής ικανότητας.

## Κατασκευή Multiview-Fusion Συστήματος

Στο τελικό implementation της διπλωματικής εργασίας, χρησιμοποιήθηκε η προηγούμενη διαδικασία εκπαίδευσης για την κατασκευή των βέλτιστων single-view κατηγοροποιητών για κάθε στήθος, για κάθε προβολή, ως προς την αναγνώριση των μαζών και την διάκρισή τους σε benign και malignant. Στην συνέχεια, αφότου έγινε η κατάλληλη προεπεξεργασία, για την κατασκευή των απαραίτητων δεδομένων για κάθε ασθενή, στην προβλεπόμενη μορφή, εκτιμήθηκε η κάθε πρόβλεψη από τον αντίστοιχο κατηγοροποιητή της, γεγονός που οδήγησε στον σχηματισμό πινάκων με ανεξάρτητες εκτιμήσεις για τον κάθε ασθενή. Στη συνέχεια, με βάση τον πίνακα αυτόν, κατασκευάστηκαν shallow δίκτυα που λάμβαναν υπόψιν το σύνολο των εκτιμήσεων και προχωρούσαν στην τελική πρόβλεψη για την υγεία του ασθενή.

### 8.1 Προεπεξεργασία

Η διαδικασία της προ επεξεργασίας του συνόλου δεδομένων της βάσης CBIS-DDSM είναι όμοια με αυτή του transfer learning που πραγματοποιήθηκε στα προηγούμενα πειράματα, καθώς σύμφωνα με την βιβλιογραφία (Jouirou, Walid, & Baazaoui, 2019) (Agarwal, Diaz, Martí, & Llado, 2018) οι συγκεκριμένες πρακτικές αποτελούν γενικευμένη προτεινόμενη στρατηγική.

Ακολουθώντας αυτό το σκεπτικό, εκπαιδεύτηκε για κάθε προβολή, ένα σύνολο των προαναφερόμενων νευρωνικών δικτύων, ενώ σε ορισμένες περιπτώσεις, χάριν εξοικονομήσεως πόρων και υπολογιστικού κόστους, οι δομές που παρουσίαζαν, στις πρώτες κατηγορίες προβολών, μέτρια αποτελέσματα και απαιτούσαν σημαντικά περισσότερο χρόνο αποφεύχθηκαν, στις υπόλοιπες, να αξιοποιηθούν.

Το accuracy των εκάστοτε νευρωνικών δικτύων, σε όλες τις περιπτώσεις, ομοιάζουν αυτών, του προηγούμενου πειράματος, γεγονός που ήταν ιδιαίτερα λογικό καθώς, η φύση των εικόνων δεν έχει αλλάξει ιδιαίτερα.

#### 8.1.1 Δημιουργία βάσεων

Το πρώτο κομμάτι δεν είναι να γίνει ο διαχωρισμός των ασθενών, αλλά να δημιουργηθεί για την κάθε προβολή οι 4 διαφορετικές βάσεις και να εκπαιδύσουμε όλα τα νευρωνικά δίκτυα ξανά για την κάθε προβολή και να βρούμε το καλύτερο. Το script groupdata.m είναι ίδιο ακριβώς με το προηγούμενο που περιγράψαμε, και φτιάχνεται πάλι ένας πίνακας που έχει όλα τα δεδομένα, δεν αλλάζει τίποτα. Αυτό που αλλάζει είναι στο script create folders.m

Ουσιαστικά γίνεται load το table που δημιουργήθηκε. Κάθε φορά γίνεται διαχωρισμός περί τίνος πρόκειται Left ή Right Breast και μετά CC ή MLO View και δημιουργούνται οι αντιστοιχεί φάκελοι με τις περιοχές ενδιαφέροντος. Τέλος η προ επεξεργασία είναι ίδια με πριν.

Στην εκπαίδευση Νευρωνικών δικτύων για την κάθε προβολή, δεν αλλάζει τίποτα όπως φαίνεται παρακάτω. Το μόνο που αλλάζει είναι το path και δίνεται εντολή για κάθε μαστό με την κάθε προβολή ξεχωριστά. Τέλος, η εκπαίδευση γίνεται με τον ίδιο τρόπο που έγινε και προηγουμένως.

Μετά την ολοκλήρωση της επιμέρους εκπαίδευσης των εκάστοτε νευρωνικών δικτύων, επιλέχθηκαν τα βέλτιστα, με βάση τα υπάρχοντα δεδομένα, για την εκάστοτε προβολή.

#### CC View Left-Breast

Για την συγκεκριμένη προβολή επιλέχθηκε η αρχιτεκτονική Inception v1 (googlenet) με αλγόριθμο βελτιστοποίησης Adam και ρυθμό εκπαίδευσης (learning rate)  $10^{-4}$ . Το accuracy του συγκεκριμένου αλγορίθμου, πλησιάζει το 70%.

#### CC View Right-Breast

Για την συγκεκριμένη προβολή, το μοντέλο Inception v1 (googlenet) παρουσίασε καλύτερα αποτελέσματα (γύρω στο 72%) με αλγόριθμο βελτιστοποίησης Adam και ρυθμό εκπαίδευσης  $10^{-7}$ .

#### MLO View Left-Breast

Το μοντέλο Inception v1, παρουσίαζε, επίσης καλές ικανότητες πρόβλεψης σχετικά με την mlo προβολή του αριστερού στήθους, με τα αποτελέσματα να ομοιάζουν των προηγούμενων προβολών, επιτυγχάνοντας περίπου 70% ακρίβεια. Το learning rate έχει τιμή  $10^{-7}$  και ο αλγόριθμος βελτιστοποίησης είναι πάλι ο Adam.

#### MLO View Right-Breast

Εδώ, το Inception v1 (googlenet) δίκτυο παρουσίασε σε προαναφερόμενα επίπεδα προβλέψεις, με αλγόριθμο βελτιστοποίησης των RMSProp και learning rate το  $10^{-4}$ .

#### 8.1.2 Μετατροπή malignant/benign σε 0 1

Η ίδια τεχνική χρησιμοποιήθηκε και στην singleview εκπαίδευση και έχει σκοπό να αποφύγει την χρήση strings κατά την διάρκεια τόσο της deep όσο και της shallow κατηγοροποίησης. Η συγκεκριμένη αλλαγή δεν έχει καμία επίπτωση στο integrity των εκτιμήσεων που παρέχονται.

```
for i=1:height(Test_Dataset.A)
    if
        (contains(string(Test_Dataset.A.pathology(i)), 'BENIGN')==1)
            Test_Dataset.A.pathology(i)=num2cell(0);
        elseif
            (contains(string(Test_Dataset.A.pathology(i)), 'MALIGNANT')==1)
                Test_Dataset.A.pathology(i)=num2cell(1);
            end
    end
end
```

### 8.1.3 Κατασκευή Πινάκων Ασθενών

Στο script createTable.m, που μπορεί να φανεί στο τέλος του κεφαλαίου, κατασκευάζεται ο τελικός πίνακας που περιλαμβάνει όλους τους ασθενείς, τις εκτιμήσεις για την εκάστοτε προβολή του εκάστοτε μαστού τους και τέλος, μία στήλη με την τελική πρόβλεψη.

Αναλυτικά, δημιουργείται ένας πίνακας που περιλαμβάνει, 6 στήλες. Αυτές περιγραφικά είναι οι παρακάτω.

PatientID	LeftCC	LeftMLO	RightCC	RightMLO	Overall
<pre>headers={'PatientID','RIGHTCC','LEFTCC','RIGHTMLO','LEFTMLO','RIGHTCC_Mask','LEFTCC_Mask','RIGHTMLO_Mask','LEFTMLO_Mask','Diagnosis'}; load ('Train_Dataset.mat'); load ('Test_Dataset.mat');</pre>					

Βέβαια όπως φαίνεται παραπάνω, στο implementation προσθέσαμε και τις στήλες που αφορούν την μάσκα για την αντίστοιχη περίπτωση, ώστε να πραγματοποιείται εκείνη την στιγμή η εφαρμογή της και, εν τέλει η δημιουργία της ROI εικόνας.

Αρχικά, λοιπόν, για κάθε ασθενή, λαμβάνουμε το path ολόκληρης της εικόνας, της μάσκας και την περνάμε στην αντίστοιχη στήλη:

```
if (strcmp(Train_Dataset.A.leftOrRightBreast(i),"RIGHT")==1)

    if (strcmp(Train_Dataset.A.imageView(i),"CC")==1)

        T.RIGHTCC(lengthT)=Train_Dataset.A.imageFilePath(i);

        T.RIGHTCC_Mask(lengthT)=Train_Dataset.A.ROIMaskFilePath(i);
```

Το ίδιο πραγματοποιείται για όλες τις περιπτώσεις που υπάρχουν στην μνήμη. Σε περίπτωση που διαβάζεται καινούργιος ασθενής, τότε προστίθεται στον πίνακα, ενώ σε περίπτωση μη ύπαρξης δεδομένων για κάποιες προβολές, τοποθετείται κενή (null) τιμή. Ως τελική διάγνωση, για την εκπαίδευση, χρησιμοποιείται η κυριάρχων, αριθμητικά, τιμή που παρέχεται από τους κατηγοροποιητές, ενώ ταυτόχρονα συμπληρώνονται και οι κενές στήλες (labelEveryData.m)

```
if (cell2mat(Table.DataTable.T.RIGHTCC(i))~=0 &
    cell2mat(Table.DataTable.T.RIGHTCC(i))~=1)

    Table.DataTable.T.RIGHTCC(i)=Table.DataTable.T.Diagnosis(i);
```

Η τελευταία τεχνική χρησιμοποιήθηκε με σκοπό την δημιουργία ενός πρώιμου fusion συστήματος, με σκοπό στο μέλλον να αντικατασταθεί από βελτιωμένες.



Παρακάτω, παρατίθεται η μορφή του πίνακα

1 PatientID	2 RIGHTCC	3 LEFTCC	4 RIGHTMLO	5 LEFTMLO	6 Diagnosis
'P_00001'	[]	0	[]	1	1
'P_00004'	[]	[]	0	0	0
'P_00009'	[]	[]	1	[]	1
'P_00018'	[]	[]	1	[]	0
'P_00021'	0	[]	0	0	0
'P_00023'	[]	[]	0	[]	1
'P_00026'	[]	[]	[]	0	0
'P_00027'	[]	[]	1	[]	0
'P_00034'	[]	[]	1	[]	1
'P_00039'	[]	[]	1	[]	1
'P_00041'	[]	[]	[]	0	0
'P_00044'	0	[]	1	[]	0
'P_00045'	[]	[]	[]	1	1
'P_00051'	[]	[]	[]	1	1
'P_00055'	[]	[]	[]	[]	0
'P_00057'	[]	[]	1	[]	1
'P_00059'	[]	[]	[]	1	1
'P_00061'	[]	[]	0	[]	0
'P_00065'	[]	[]	[]	0	0
'P_00068'	[]	[]	1	[]	1
'P_00074'	0	[]	1	[]	1
'P_00076'	[]	[]	[]	1	0
'P_00079'	[]	[]	1	[]	1
'P_00080'	[]	[]	1	[]	1
'P_00081'	[]	[]	1	[]	0
'P_00086'	[]	[]	1	[]	1
'P_00090'	[]	[]	[]	0	0
'P_00092'	0	[]	1	1	1
'P_00094'	[]	[]	1	[]	0
'P_00095'	[]	[]	[]	0	1
'P_00096'	[]	[]	1	[]	0
'P_00106'	0	0	0	0	0
'P_00108'	[]	[]	[]	0	0
'P_00110'	0	[]	[]	1	1
'P_00119'	[]	[]	[]	0	0
'P_00120'	[]	[]	[]	1	0
'P_00128'	[]	[]	[]	[]	1
'P_00133'	[]	[]	[]	1	1

1 PatientID	2 RIGHTCC	3 LEFTCC	4 RIGHTMLO	5 LEFTMLO	6 Diagnosis
P_00001'	[]	P_00001\LEFTCC\image.tif	[]	P_00001\LEFTMLO\image.tif	'MALIGNANT'
P_00004'	[]	[]	P_00004\RIGHTMLO\image.tif	P_00004\LEFTMLO\image.tif	'BENIGN'
P_00009'	[]	[]	P_00009\RIGHTMLO\image.tif	[]	'MALIGNANT'
P_00015'	[]	[]	[]	[]	'MALIGNANT'
P_00018'	[]	[]	P_00018\RIGHTMLO\image.tif	[]	'BENIGN'
P_00021'	P_00021\RI...	[]	P_00021\RIGHTMLO\image.tif	P_00021\LEFTMLO\image.tif	'BENIGN'
P_00023'	[]	[]	P_00023\RIGHTMLO\image.tif	[]	'MALIGNANT'
P_00026'	[]	[]	[]	P_00026\LEFTMLO\image.tif	'BENIGN'
P_00027'	[]	[]	P_00027\RIGHTMLO\image.tif	[]	'BENIGN'
P_00034'	[]	[]	P_00034\RIGHTMLO\image.tif	[]	'MALIGNANT'
P_00039'	[]	[]	P_00039\RIGHTMLO\image.tif	[]	'MALIGNANT'
P_00041'	[]	[]	[]	P_00041\LEFTMLO\image.tif	'BENIGN_WITHOUT_CALLBACK'
P_00044'	P_00044\RI...	[]	P_00044\RIGHTMLO\image.tif	[]	'BENIGN'
P_00045'	[]	[]	[]	P_00045\LEFTMLO\image.tif	'MALIGNANT'
P_00046'	[]	[]	[]	[]	'MALIGNANT'
P_00051'	[]	[]	[]	P_00051\LEFTMLO\image.tif	'MALIGNANT'
P_00054'	[]	[]	[]	[]	'BENIGN_WITHOUT_CALLBACK'
P_00055'	[]	[]	[]	[]	'BENIGN'
P_00057'	[]	[]	P_00057\RIGHTMLO\image.tif	[]	'MALIGNANT'
P_00058'	[]	[]	[]	[]	'MALIGNANT'
P_00059'	[]	[]	[]	P_00059\LEFTMLO\image.tif	'MALIGNANT'
P_00061'	[]	[]	P_00061\RIGHTMLO\image.tif	[]	'BENIGN'
P_00064'	[]	[]	[]	[]	'BENIGN_WITHOUT_CALLBACK'
P_00065'	[]	[]	[]	P_00065\LEFTMLO\image.tif	'BENIGN'
P_00068'	[]	[]	P_00068\RIGHTMLO\image.tif	[]	'MALIGNANT'
P_00074'	P_00074\RI...	[]	P_00074\RIGHTMLO\image.tif	[]	'MALIGNANT'
P_00076'	[]	[]	[]	P_00076\LEFTMLO\image.tif	'BENIGN'
P_00079'	[]	[]	P_00079\RIGHTMLO\image.tif	[]	'MALIGNANT'
P_00080'	[]	[]	P_00080\RIGHTMLO\image.tif	[]	'MALIGNANT'
P_00081'	[]	[]	P_00081\RIGHTMLO\image.tif	[]	'BENIGN'
P_00086'	[]	[]	P_00086\RIGHTMLO\image.tif	[]	'MALIGNANT'
P_00090'	[]	[]	[]	P_00090\LEFTMLO\image.tif	'BENIGN'
P_00092'	P_00092\RI...	[]	P_00092\RIGHTMLO\image.tif	P_00092\LEFTMLO\image.tif	'MALIGNANT'
P_00094'	[]	[]	P_00094\RIGHTMLO\image.tif	[]	'BENIGN'
P_00095'	[]	[]	[]	P_00095\LEFTMLO\image.tif	'MALIGNANT'
P_00096'	[]	[]	P_00096\RIGHTMLO\image.tif	[]	'BENIGN'
P_00106'	P_00106\RI...	P_00106\LEFTCC\image.tif	P_00106\RIGHTMLO\image.tif	P_00106\LEFTMLO\image.tif	'BENIGN_WITHOUT_CALLBACK'
P_00107'	[]	[]	[]	[]	'BENIGN_WITHOUT_CALLBACK'

Εικόνα 59: Πίνακες Ασθενών

#### 8.1.4 Τελικά διανύσματα

Με σκοπό να διαβαστούν ορθά από τα shallow νευρωνικά δίκτυα, κατασκευάστηκαν στο script BinaryClassification.m τα αντίστοιχα διανύσματα που μπορούν να προκύψουν από τις στήλες του παραπάνω κατασκευασμένου πίνακα.

```
load('Labeled Data Everything1.mat');
LEFTCC=cell2mat(Table.DataTable.T.LEFTCC);
RIGHTCC=cell2mat(Table.DataTable.T.RIGHTCC);
LEFTMLO=cell2mat(Table.DataTable.T.LEFTMLO);
RIGHTMLO=cell2mat(Table.DataTable.T.RIGHTMLO);
LEFTPREDICTION= [LEFTCC LEFTMLO];
```

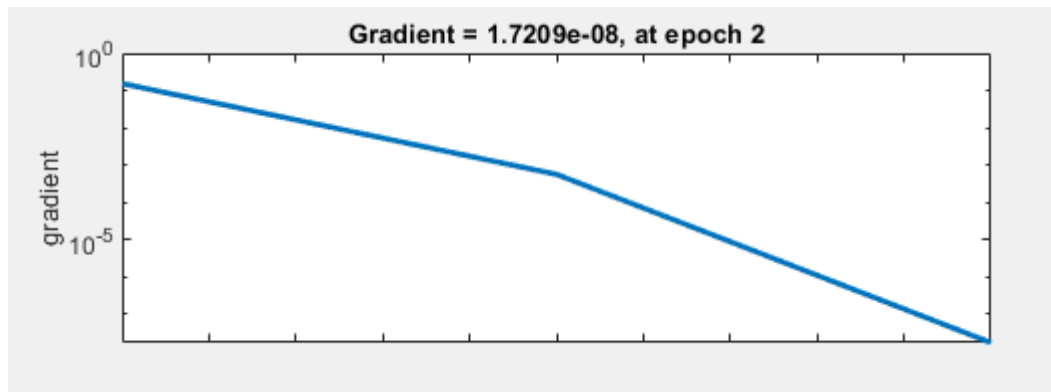
## 8.2 Εκπαίδευση Fusion Συστήματος

Στη συγκεκριμένη διπλωματική εργασία, λόγω της χρήσης μίας βάσης δεδομένων, αλλά και των μέτριων αποτελεσμάτων που παρέχονταν από τα νευρωνικά δίκτυα (οι καλύτερες προβλέψεις πλησίαζαν το 75%), ήταν σχεδόν βέβαιο, ότι ο επιμέρους περιορισμός των νευρωνικών δικτύων στις εκάστοτε προβολές και η προσπάθεια εξαγωγής ενός ενιαίου συστήματος, θα οδηγούσε σε χαμηλότερα αποτελέσματα. Για τον λόγο αυτό, κατασκευάστηκε ένα proof of concept ενιαίο σύστημα, θέτοντας ως στόχο, την, όσο δυνατόν, μικρότερη απώλεια accuracy στις προβλέψεις.

Στο Fusion σύστημα, πραγματοποιήθηκε η διόρθωση, μέσω της κατασκευής δύο shallow νευρωνικών δικτύων για την πρόβλεψη της εκάστοτε περίπτωσης και ενός ακόμη που λαμβάνει υπόψιν τους τις δύο αυτές εκτιμήσεις και παράγει το τελικό αποτέλεσμα. Η απόδοση κυμαίνεται γύρω στο 70%.

Για τα shallow νευρωνικά δίκτυα, επιλέχθηκαν 2 layers με 20 nodes το καθένα.

Τον τελικό παραγόμενο κώδικα, που περιγράφει το configuration του shallow δικτύου μπορεί κάποιος να το βρει στον FinalNN.m script.



		Training Confusion Matrix		
		0	1	
Output Class	0	192 55.2%	156 44.8%	55.2% 44.8%
	1	0 0.0%	0 0.0%	NaN% NaN%
		100% 0.0%	0.0% 100%	55.2% 44.8%
		0	1	
		Target Class		

		Validation Confusion Matrix		
		0	1	
Output Class	0	42 56.8%	32 43.2%	56.8% 43.2%
	1	0 0.0%	0 0.0%	NaN% NaN%
		100% 0.0%	0.0% 100%	56.8% 43.2%
		0	1	
		Target Class		

		Test Confusion Matrix		
		0	1	
Output Class	0	33 44.6%	41 55.4%	44.6% 55.4%
	1	0 0.0%	0 0.0%	NaN% NaN%
		100% 0.0%	0.0% 100%	44.6% 55.4%
		0	1	
		Target Class		

		All Confusion Matrix		
		0	1	
Output Class	0	267 53.8%	229 46.2%	53.8% 46.2%
	1	0 0.0%	0 0.0%	NaN% NaN%
		100% 0.0%	0.0% 100%	53.8% 46.2%
		0	1	
		Target Class		

## Βιβλιογραφία 8<sup>ου</sup> Κεφαλαίου

- Cadène, R., Thome, N., & Cord, M. (2016). Master's Thesis : Deep Learning for Visual Recognition.
- He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep Residual Learning for Image Recognition.
- Jouirou, A., Walid, B., & Baazaoui, A. (2019). Multi-View Information Fusion in Mammograms: A Comprehensive Overview. *Information Fusion*.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Advances in neural information processing systems*.
- Qian, W., Song, D., Lei, M., Sankar, R., & Eikman, E. (2007). Computer-aided Mass Detection Based on Ipsilateral Multiview Mammograms.
- Richa, A., Oliver, D., Xavier, L., Moi, H., & Robert, M. (2019). Automatic mass detection in mammograms using deep convolution neural networks. *Journal of Medical Imaging*.
- Simonyan, K., & Zisserman, A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition.
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., & Erhan, D. (2014). Going deeper with convolutions.
- Tsochatzidis, L., Costaridou, L., & Pratikakis, I. (2019). Deep Learning for Breast Cancer Diagnosis from Mammograms - A Comparative Study. *Journal of Imaging*.
- Wang, Y., Choi, E. J., Choi, Y., Zhang, H., Jin, G. Y., & Ko, S.-B. (2020). Breast Cancer Classification in Automated Breast Ultrasound Using Multiview Convolutional Neural Network With Transfer Learning. *Ultrasound Med. Biol.*
- Wolfson Stacey, Y. S. (2018). High-Resolution Breast Cancer Screening with Multi-View Deep Convolutional NN.
- Zhang, X., Zou, J., He, K., & Sun, J. (2015). Accelerating Very Deep Convolutional Networks for Classification and Detection. *Computer Vision and Pattern Recognition*.

## Κώδικας

```
headers={'PatientID','RIGHTCC','LEFTCC','RIGHTMLO','LEFTMLO','RIGHTCC_Mask','LEFTCC_Mask','RIGHTMLO_Mask','LEFTMLO_Mask','Diagnosis'};
load ('Train_Dataset.mat');
load ('Test_Dataset.mat');

patientIDBeingStudied=Train_Dataset.A.patient_id(1);
lengthT=1;

DataLen=height(Train_Dataset.A);
DataLen1=height(Test_Dataset.A);

T.PatientID(lengthT)=Train_Dataset.A.patient_id(1);
T.Diagnosis(lengthT)=Train_Dataset.A.pathology(1);

data = cell(DataLen+DataLen1,10);
T = cell2table(data);
T.Properties.VariableNames = headers;
for i=1:DataLen
    if
        (strcmp(Train_Dataset.A.patient_id(i),patientIDBeingStudied)==1)
            T.PatientID(lengthT)=Train_Dataset.A.patient_id(i);
            T.Diagnosis(lengthT)=Train_Dataset.A.pathology(i);
            if
                (strcmp(Train_Dataset.A.leftOrRightBreast(i),"RIGHT")==1)
                    if (strcmp(Train_Dataset.A.imageView(i),"CC")==1)

T.RIGHTCC(lengthT)=Train_Dataset.A.imageFilePath(i);

T.RIGHTCC_Mask(lengthT)=Train_Dataset.A.ROIMaskFilePath(i);
                    else

T.RIGHTMLO(lengthT)=Train_Dataset.A.imageFilePath(i);

T.RIGHTMLO_Mask(lengthT)=Train_Dataset.A.ROIMaskFilePath(i);
                    end
                else
                    if (strcmp(Train_Dataset.A.imageView(i),"CC")==1)

T.LEFTCC(lengthT)=Train_Dataset.A.imageFilePath(i);

T.LEFTCC_Mask(lengthT)=Train_Dataset.A.ROIMaskFilePath(i);
                    else

T.LEFTMLO(lengthT)=Train_Dataset.A.imageFilePath(i);

T.LEFTMLO_Mask(lengthT)=Train_Dataset.A.ROIMaskFilePath(i);
                    end
                end
            lengthT=lengthT+1;
        else
            patientIDBeingStudied=Train_Dataset.A.patient_id(i);
        end
    end
end

patientIDBeingStudied=Test_Dataset.A.patient_id(1);
```

6: Create Table - Μέρος 1

```

for i=1:DataLen1
    if
        (strcmp(Test_Dataset.A.patient_id(i),patientIDBeingStudied)==1)
            T.PatientID(lengthT)=Test_Dataset.A.patient_id(i);
            T.Diagnosis(lengthT)=Test_Dataset.A.pathology(i);
            if
                (strcmp(Test_Dataset.A.leftOrRightBreast(i),"RIGHT")==1)
                    if (strcmp(Test_Dataset.A.imageView(i),"CC")==1)

T.RIGHTCC(lengthT)=Test_Dataset.A.imageFilePath(i);

T.RIGHTCC_Mask(lengthT)=Test_Dataset.A.ROIMaskFilePath(i);
                    else

T.RIGHTMLO(lengthT)=Test_Dataset.A.imageFilePath(i);

T.RIGHTMLO_Mask(lengthT)=Test_Dataset.A.ROIMaskFilePath(i);
                    end
                else
                    if (strcmp(Test_Dataset.A.imageView(i),"CC")==1)
                        T.LEFTCC(lengthT)=Test_Dataset.A.imageFilePath(i);

T.LEFTCC_Mask(lengthT)=Test_Dataset.A.ROIMaskFilePath(i);
                    else

T.LEFTMLO(lengthT)=Test_Dataset.A.imageFilePath(i);

T.LEFTMLO_Mask(lengthT)=Test_Dataset.A.ROIMaskFilePath(i);
                    end
                end
                lengthT=lengthT+1;
            else
                patientIDBeingStudied=Test_Dataset.A.patient_id(i);
            end
        end
    end
end

```

8: Create Table - Μέρος 2

```

for i=1:rows
    if (cell2mat(Table.DataTable.T.LEFTCC(i))~=0 &
cell2mat(Table.DataTable.T.LEFTCC(i))~=1)

Table.DataTable.T.LEFTCC(i)=Table.DataTable.T.Diagnosis(i);
    end
    if (cell2mat(Table.DataTable.T.RIGHTCC(i))~=0 &
cell2mat(Table.DataTable.T.RIGHTCC(i))~=1)

Table.DataTable.T.RIGHTCC(i)=Table.DataTable.T.Diagnosis(i);
    end
    if (cell2mat(Table.DataTable.T.LEFTMLO(i))~=0 &
cell2mat(Table.DataTable.T.LEFTMLO(i))~=1)

Table.DataTable.T.LEFTMLO(i)=Table.DataTable.T.Diagnosis(i);
    end
    if (cell2mat(Table.DataTable.T.RIGHTMLO(i))~=0 &
cell2mat(Table.DataTable.T.RIGHTMLO(i))~=1)

Table.DataTable.T.RIGHTMLO(i)=Table.DataTable.T.Diagnosis(i);
    end
end

Table.DataTable.T(end,:)=[];

```

10: labelEveryData

```

load('Labeled_Data_Everything1.mat');
LEFTCC=cell2mat(Table.DataTable.T.LEFTCC);
RIGHTCC=cell2mat(Table.DataTable.T.RIGHTCC);
LEFTMLO=cell2mat(Table.DataTable.T.LEFTMLO);
RIGHTMLO=cell2mat(Table.DataTable.T.RIGHTMLO);
LEFTPREDICTION= [LEFTCC LEFTMLO];
RIGHTPREDICTION= [RIGHTCC RIGHTMLO];
LEFT_PRED=ceil(NN_L(LEFTPREDICTION));
RIGHT_PRED=ceil(NN_R(RIGHTPREDICTION));
PREDICTION=[LEFT_PRED RIGHT_PRED];
OUTPUT=cell2mat(Table.DataTable.T.Diagnosis);

```

Figure 9: BinaryClassification.m



```

function [y1] = final_NN()
%MYNEURALNETWORKFUNCTION neural network simulation function.
%
% Auto-generated by MATLAB, 19-Jul-2020 18:53:24.
%
% [y1] = myNeuralNetworkFunction() takes these arguments:
% and returns:
%   y = Qx1 matrix, output #1
% where Q is the number of samples.

%#ok<*RPMT0>

% ===== NEURAL NETWORK CONSTANTS =====

% Layer 1
b1 = [0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0;0];

% Layer 2
b2 = 0;
LW2_1 = [0.51195824186563088087 -0.95494565723802304902 -
0.45236862734041333178 0.69865709089092997708 -
0.85693348032997074259 -0.44683706787383203896
0.64201411084324466749 -0.48300762074555902981
0.33134958238541339925 0.54662865060079135482
0.5074299713627560271 0.39766781074904472559 -
0.37484369848081772192 0.59629054776404688543
0.67146544173684763468 0.42695039460450984192 -
0.75907018508905721443 -0.4698881607547106265 -
0.44836264059821773209 1.2009525956562658866];

% ===== SIMULATION =====

% Dimensions
Q = 0; % samples

% Layer 1
a1 = tansig_apply(repmat(b1,1,Q));

% Layer 2
a2 = logsig_apply(repmat(b2,1,Q) + LW2_1*a1);

% Output 1
y1 = a2;
y1 = y1';
end

% ===== MODULE FUNCTIONS =====

% Sigmoid Positive Transfer Function
function a = logsig_apply(n,~)
a = 1 ./ (1 + exp(-n));
end

% Sigmoid Symmetric Transfer Function
function a = tansig_apply(n,~)
a = 2 ./ (1 + exp(-2*n)) - 1;
end

```

*11: FinalNN*

Page Intentionally Left Blank

## Κεφάλαιο 9<sup>ο</sup>

### Συζήτηση-Βελτιώσεις

Η συγκεκριμένη ενότητα, αποτελεί μία χρήσιμη ευκαιρία, να αναλυθούν περαιτέρω βήματα, διορθωτικές κινήσεις και βελτιστοποιήσεις που μπορούν να εφαρμοστούν στο μέλλον με σκοπό να παρέχουν ακόμη καλύτερα αποτελέσματα.

Στο κομμάτι της προεπεξεργασίας, η μετατροπή από Grayscale σε RGB, η αλλαγή σε TIFF και η χρήση της παρεχόμενης ROI μάσκας με σκοπό την δημιουργία των αντίστοιχων εικόνων για εκπαίδευση, αποτελούν τεχνικές που ακολουθούνται κατά κόρον από πληθώρα μελετών.

Αυτό που θα μπορούσε να πραγματοποιηθεί, είναι η πλήρη μετατόπιση του ενδιαφέροντος και η προσήλωση σε χρήση νέων αλγορίθμων/τεχνικών με σκοπό τον εξαρχής εντοπισμό ύποπτων όγκων και αντίστοιχων ROIs.

Από εκεί και πέρα, η εκπαίδευση του SingleView συστήματος, από θέμα pipeline, ακολουθεί τα ίδια πρότυπα, με την λογική ότι εξαντλεί σε σημαντικό βαθμό την πληθώρα των βασικών παραμέτρων που μπορούν να αξιολογηθούν στα βαθιά νευρωνικά δίκτυα, χωρίς να προσανατολίζεται, σε σημαντικό βαθμό σε fine tuning τεχνικές, κυρίως λόγω της proof of concept φύσης του εγχειρήματος.

Παρόμοια λογική ισχύει και για το Fusion σύστημα όπου, ορισμένες τεχνικές επιλέχθηκαν έχοντας ως στόχο την παραγωγή proof of concept συστημάτων σε βάρος των βέλτιστων αποτελεσμάτων.

Τέλος, μία ακόμη ιδιαίτερα σημαντική βελτίωση, μπορεί να εντοπιστεί στον όγκο των δεδομένων που χρησιμοποιήθηκαν ο οποίος θεωρείται ως οριακά αποδεκτός. Η χρήση περισσότερων, γενικώς αποδεκτών βάσεων δεδομένων, τόσο screen filmed όσο και digital, μπορούν να βοηθήσουν τα μέγιστα στην κατασκευή αρχιτεκτονικών νευρωνικών δικτύων με σημαντικές δυνατότητες εξαγωγής εκτιμήσεων.

## Επίλογος

Ο καρκίνος του μαστού, αποτελεί, στην σύγχρονη εποχή, ένα, ιδιαιτέρως, σημαντικό ζήτημα που προσβάλλει ένα μεγάλο εύρος των γυναικών. Οι συχνές εξετάσεις και η ψηλάφηση από την ίδια την ασθενή, είναι δυνατόν να βοηθήσουν τον εντοπισμό του καρκίνου σε αρχικό στάδιο, γεγονός που προσφέρει σημαντικά περισσότερες πιθανότητες θεραπείας.

Για το πρώτο κομμάτι, είναι ευνόητο, ότι στην μαστογραφία, είναι πολύ σημαντικό να πραγματοποιηθούν, όσο το δυνατόν περισσότερο, ακριβής εκτιμήσεις για τον όγκο, το σχήμα και το είδος του καρκίνου. Σε αυτό το δύσκολο εγχείρημα, είναι κατανοητό ότι υπάρχει σημαντική πιθανότητα ανθρώπινου λάθους, λόγω κούρασης, ελλειπών εκπαιδευσης ή και δυσκολίας εντοπισμού του από το μάτι.

Σε αυτό, είναι δυνατόν να αξιοποιηθούν υπολογιστικά συστήματα και με την χρήση τεχνικών μηχανικής μάθησης και βαθιών νευρωνικών δικτύων, να κατασκευαστούν εκτιμητές που να μπορούν να λειτουργούν ως σύμβουλοι των παραπάνω ειδικών, παρέχοντας και οι ίδιοι μία πρόβλεψη για το είδος του όγκου.

Εδώ, σαν πιθανή βελτίωση, προτείνεται η κατασκευή όχι μόνο κατηγοροποιητών που θα είναι σε θέση να αναγνωρίσουν το είδος του όγκου σε μία προβολή της μαστογραφίας, αλλά και με βάση την συνολική εικόνα, να παρέχουν μία ολοκληρωμένη πρόβλεψη στον ιατρό που μελετά τον συγκεκριμένο ασθενή.

Τα συστήματα και τα pipelines που κατασκευάστηκαν στην συγκεκριμένη διπλωματική εργασία, αποτελούν προϊόν πολλαπλών μελετών πάνω στο παραπάνω θέμα και είχαν ως αποτέλεσμα την κατασκευή μίας πληθώρας κατηγοροποιητών μίας προβολής-εικόνας (SingleView) αλλά και ενός συνολικού συστήματος που λαμβάνει υπόψιν όλες τις προβλέψεις για την εκάστοτε προβολή (MultiView-Fusion). Το τελευταίο, αποτελεί κυρίως proof of concept implementation και επιδέχεται σημαντικές βελτιώσεις.

Η πρόοδος της τεχνολογίας σε συνδυασμό με κατάλληλη χρηματοδότηση και επιπρόσθετη έρευνα είναι σε θέση να προσφέρουν στο ανθρώπινο είδος καλύτερα και βελτιωμένα εργαλεία με σκοπό να προβούμε σε έγκυρες-πρόωρες προβλέψεις του καρκίνου του μαστού, αυξάνοντας σημαντικά τα ποσοστά επιβίωσης και της ποιότητας ζωής των ασθενών.