



ΠΑΝΕΠΙΣΤΗΜΙΟ
ΠΑΤΡΩΝ
UNIVERSITY OF PATRAS

Πολυτεχνική Σχολή

Τμήμα Μηχανικών Η/Υ & Πληροφορικής

Διπλωματική Εργασία

Ανάλυση εικόνας με στόχο τον εντοπισμό καρκίνου με
τεχνικές μηχανικής και εις βάθος μάθησης

Αντώνιος Ματζάκος-Χωριανόπουλος

A.M. 235831

Επιβλέπων

Καθηγητής Ιωάννης Χατζηλυγερούδης

Συνεπιβλέπων

Μετα-διδαστορικός Ισίδωρος Περίκος

Πάτρα, 2020

© Copyright συγγραφέας Αντώνιος Ματζάκος Χωριανόπουλος, 2020

© Copyright θέματος Ιωάννης Χατζυληγερούδης & Ισίδωρος Περίκος

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

Η έγκριση της διπλωματικής εργασίας από το Τμήμα Μηχανικών Ηλεκτρονικών Υπολογιστών & Πληροφορικής του Πανεπιστημίου Πατρών δεν υποδηλώνει απαραίτητως και αποδοχή των απόψεων του συγγραφέα εκ μέρους του Τμήματος.

ΕΥΧΑΡΙΣΤΙΕΣ

Αρχικά θα ήθελα να ευχαριστήσω τον καθηγητή μου κ. Ιωάννη Χατζηλυγερούδη για την ευκαιρία που μου έδωσε να ασχοληθώ με έναν τόσο ενδιαφέρον τομέα δίνοντας μου, εν μέσω του θέματος διπλωματικής μου εργασίας το ερέθισμα να ερευνήσω πτυχές της επιστήμης των υπολογιστών αλλά και της ζωής με τις οποίες δεν είχα συνδιαλεχθεί νωρίτερα. Επίσης θα ήθελα να ευχαριστήσω τους κ. Ισίδωρο Περίκο και κ.Ιωάννη Δαραμούσκα για τη συμβολή και καθοδήγηση τους καθ' όλη τη διάρκεια εκπόνησης της διπλωματικής μου. Τέλος θα ήθελα να ευχαριστήσω τον αδερφό μου κ. Ιωάννη Ματζάκο – Χωριανόπουλο για την καθοδήγηση του και συνεχή υποστήριξή του με κάθε τρόπο, καθώς και την οικογένειά μου.

Πρόλογος

Η παρούσα διπλωματική εκπονήθηκε το Ακαδημαϊκό Έτος 2019-2020 στα πλαίσια του προπτυχιακού προγράμματος σπουδών του τμήματος Μηχανικών Ηλεκτρονικών Υπολογιστών και Πληροφορικής του Πανεπιστημίου Πατρών. Η διπλωματική εργασία αυτή αποτελεί το μέσο για την λήψη του ενσωματωμένου μεταπτυχιακού τίτλου σπουδών. Η εργασία πραγματοποιήθηκε υπό την επίβλεψη του κ. Ιωάννη Χατζηλυγερούδη, καθηγητή του τμήματος Μηχανικών Ηλεκτρονικών Υπολογιστών και Πληροφορικής του Πανεπιστημίου Πατρών και συνεπιβλέποντα τον μετα-διδακτορικό του κ. Ισίδωρο Περίκο. Αντικείμενο της εργασίας αποτελεί η έρευνα και η δημιουργία ενός μοντέλου τεχνητής νοημοσύνης, το οποίο μέσω μεθόδων και τεχνικών Machine Learning και Deep Learning (Μηχανική και Εις Βάθος μάθηση) θα πραγματοποιεί ανάλυση εικόνας με στόχο να κάνει επιτυχή διάγνωση του καρκίνου του μαστού.

Περίληψη

Επηρεάζοντας περίπου το 10% των γυναικών παγκοσμίως, ο καρκίνος του μαστού έχει αποδειχθεί ότι είναι ένας από τους πιο συγκλονιστικούς και συνηθισμένους καρκίνους στις γυναίκες σήμερα. Αν και υπάρχουν διαθέσιμες θεραπείες για αυτόν τον καρκίνο, το πραγματικό ζήτημα που αντιμετωπίζουμε είναι όταν δεν μπορούμε να εντοπίσουμε σωστά την ασθένεια στα αρχικά της στάδια, ένα πρόβλημα που αντιμετωπίζουν συνεχώς τόσο οι χώρες του πρώτου κόσμου όσο και οι χώρες του τρίτου κόσμου. Η βαθιά μάθηση, ενώ ευδοκimeί στην υπολογιστική όραση, έχει αποδειχθεί ότι παίζει ζωτικό ρόλο στην πρόβλεψη τέτοιων ασθενειών. Μέθοδοι ταξινόμησης και εξαγωγής δεδομένων έχουν αποδειχθεί έως τώρα αξιόπιστες και αποτελεσματικές. Ωστόσο, πρόσφατες μελέτες έχουν δείξει, μέσω πολλαπλών πειραμάτων και δοκιμών, ότι οι τεχνικές βαθιάς μάθησης έχουν επιτύχει τα καλύτερα αποτελέσματα. Σε αυτή τη μελέτη, θα δούμε τη χρήση διαφορετικών βαθιών αρχιτεκτονικών συνελικτικών νευρικών δικτύων στην ιατρική απεικόνιση και θα τα συγκρίνουμε με ένα depth-wise συνελικτικό νευρωνικό δίκτυο σε πολλά σύνολα δεδομένων από εικόνες καρκίνου του μαστού. Οι συγκρίσεις απόδοσης γίνονται μεταξύ MobileNet, VGG16 και AlexNet σε υπερήχους μαστού και εικόνες ιστοπαθολογίας μαστού. Πρωταρχικός μας στόχος είναι να δείξουμε την αποτελεσματικότητα και την ακρίβεια τέτοιων συστημάτων στην ταξινόμηση διαφορετικών ειδών καρκίνου του μαστού. Το σύνολο δεδομένων υπερήχων περιέχει 250 εικόνες καλοήθων και κακοήθων περιπτώσεων, το δεύτερο σύνολο δεδομένων αποτελείται από 3268 εικόνες από μαστογραφίες με ιστοπαθολογία-Invasive Ductal Carcinoma (IDC) και το τρίτο αποτελείται από 10483 εικόνες επίσης με ιστοπαθολογία-IDC. Πειραματικά στοιχεία δείχνουν ότι τα προτεινόμενα μοντέλα βαθιάς μάθησης μπορούν να ταξινομήσουν αποτελεσματικά τους υπερήχους και τις ιστοπαθολογικές εικόνες. Η απόδοση του πιο αποτελεσματικού μοντέλου Convolutional Neural Network έφτασε το 96,82% ακρίβεια στους υπερήχους, το 88,23% στο δεύτερο σύνολο δεδομένων και το 91,04% στο τρίτο. Ως εκ τούτου,

κατέληξα με βάση τα αποτελέσματά μου, ότι οι αρχές της βαθιάς μάθησης στη διάγνωση του καρκίνου του μαστού είναι πράγματι πολύ ελπιδοφόρες και ο κόσμος μπορεί να ωφεληθεί αναμφίβολα από την εφαρμογή τέτοιων συστημάτων. Συνεπώς, φτάνω στο συμπέρασμα, βάση αποτελεσμάτων, ότι οι τεχνικές βαθιάς μάθησης στη διάγνωση του καρκίνου του μαστού είναι πολλά υποσχόμενες.

Λέξεις Κλειδιά : Μηχανική Μάθηση, Εις Βάθος Μάθηση, Νευρωνικά Δίκτυα, Συνελικτικά νευρωνικά δίκτυα, Καρκίνος του μαστού, Ανάλυση Εικόνας, Ιατρικές Εικόνες

Abstract

Affecting roughly about 10% of women worldwide, Breast Cancer has been proven to be one of the most shocking and common cancers among women today. Although, there are available treatments for this cancer, the real issue we face is when we cannot properly detect the disease in its early stages, a problem which both first world and third world countries encounter constantly. Deep Learning while thriving at computer vision field has been shown to play a vital role in predicting such diseases. Methods of classifying and extracting data have so far proved to be reliable and effective. However, recent studies have shown, through multiple experiments and trials, that deep learning techniques have achieved the best results. In this study, we will be seeing the use of different depth-wise convolutional neural networks in medical imaging and compare them with a simple convolutional neural net on multiple datasets of breast cancer images. Performance comparisons are made between MobileNet, VGG16 and AlexNet on breast ultrasounds and breast histopathology images. Our primary goal is to demonstrate the efficiency and accuracy of such systems in classifying different kind of breast cancer images. The ultrasounds dataset is containing 250 images of benign and malignant cases, the second (3268 images) and third (10483 images) datasets contain images of breast histology with cases of Invasive Ductal Carcinoma (IDC) and cancer free tissue. Experimental evidence suggests that the proposed deep learning models may effectively classify ultrasounds and histopathology images. The performance of the most efficient Convolutional Neural Network model reached 96.82% accuracy on ultrasounds, 88.23% on the second dataset and 91.04% on the third. Therefore, we concluded based on our results, that deep-learning tenets in the diagnosis of breast cancer are indeed very promising and the world can benefit undoubtedly from the implementation of such systems.

Keywords : Machine Learning, Deep Learning, Artificial Neural Nets, Convolutional Neural Nets, Breast Cancer, Image Analysis, Medical Imaging

Περιεχόμενα

<i>Πρόλογος</i>	5
<i>Περίληψη</i>	6
<i>Abstract</i>	8
<i>Περιεχόμενα</i>	i
<i>1</i>	11
<i>Εισαγωγή</i>	11
1.1 Κίνητρα	11
1.2 Ορισμός Προβλήματος	15
1.3 Στόχοι και Μεθοδολογία	17
1.4 Επιτομή Διπλωματικής.....	21
<i>2</i>	22
<i>Συναφείς Μελέτες & Γνωστικό Υπόβαθρο</i>	22
2.1 Ανάπτυξη των Classifiers.....	22
2.1.1 Naive Bayes Classifier	23
2.1.2 Support Vector Machine	24

2.2	Machine Learning	26
2.2.1	Ορισμός.....	26
2.2.2	Classification & Regression in Machine Learning	27
2.3	Deep Learning	30
2.3.1	Ορισμός & ανάπτυξη του Deep Learning.....	30
2.3.2	Artificial Neural Networks	32
2.3.3	Deep Learning & GPU Performance	35
2.4	Deep Learning σε Ιατρικές Εικόνες.....	37
2.4.1	Μηχανική Μάθηση για Διάγνωση του καρκίνου του μαστού	39
3.	42
	Μεθοδολογία & Υλοποίηση.....	42
3.1	Προκαταρκτική Μελέτη Δεδομένων.....	42
3.2	Προετοιμασία δεδομένων	45
3.2.1	Διαμόρφωση Συνόλου Δεδομένων	45
3.2.2	Τύποι Δεδομένων στη Μηχανική Μάθηση	46
3.3	Προεπεξεργασία Δεδομένων	48
3.3.1	Training Validation & Testing Set	48
3.3.2	Image Data Augmentation	51
3.3.3	Image Resize.....	53
3.3.4	Image Normalization	53

3.4	Αρχιτεκτονική των Νευρωνικών Δικτύων	54
3.4.1	Convolutional Neural Nets	55
3.4.2	CNN Layers	55
3.4.3	AlexNet Αρχιτεκτονική	59
3.4.4	MobileNet Αρχιτεκτονική (Lightweight Model)	64
3.4.5	Vgg16Net Αρχιτεκτονική	72
3.5	Μεθοδολογία.....	77
3.5.1	Προετοιμασία δεδομένων	77
3.5.2	Προεπεξεργασία δεδομένων.....	77
3.5.3	Deep Structured Learning Theory.....	78
3.5.4	Fine-Tuning.....	80
3.5.5	Υλοποίηση και Στρατηγική Εκπαίδευσης.....	81
3.5.6	System Specifications & Python Frameworks.....	82
4.....	84
Αποτελέσματα & Ανάλυση		84
4.1	Αποτελέσματα VGG16.....	88
4.1.1	Training & Testing Accuracy	88
4.1.2	Training & Testing Loss VGG16.....	89
4.1.3	Confusion Matrix & Classification Report.....	90
4.2	Αποτελέσματα MobileNet.....	92

4.2.1	Training & Testing Accuracy MobileNet.....	92
4.2.2	Training & Testing Loss MobileNet.....	93
4.2.3	Confusion Matrix & Classification Report.....	94
4.3	Αποτελέσματα AlexNet.....	95
4.3.1	Training & Testing Accuracy AlexNet.....	95
4.3.2	Training & Testing Loss AlexNet	96
4.3.3	Confusion Matrix & Classification Report.....	97
4.4	Πειράματα για Invasive Ductal Carcinoma (IDC).....	97
4.4.1	Vgg16 in Histopathology Images	99
4.4.2	MobileNet in Histopathology Images.....	103
4.4.3	AlexNet in Histopathology Images	107
5.....	111
<i>Συμπεράσματα & Μελλοντικές Υποθέσεις.....</i>	<i>.....</i>	<i>111</i>
<i>References.....</i>	<i>.....</i>	<i>113</i>
<i>Σύντομο Βιογραφικό Συγγραφέα.....</i>	<i>.....</i>	<i>119</i>

Λίστα Εικόνων

Εικόνα 3.1 Υπέρηχος Μαστού

Εικόνα 3.2 Data Augmentation in Breast Cancer Ultrasound

Εικόνα 3.3 Data Preparation

Εικόνα 3.4 CNN Βασική Αρχιτεκτονική

Εικόνα 3.5 Convolutional Layer

Εικόνα 3.6 Max Pooling Layer

Εικόνα 3.7 AlexNet Αρχιτεκτονική

Εικόνα 3.8 AlexNet Summary part1

Εικόνα 3.9 AlexNet Summary part2

Εικόνα 3.10 MobileNet Αρχιτεκτονική [22]

Εικόνα 3.11 MobileNet Layers [22]

Εικόνα 3.12 MobileNet Model Summary part1

Εικόνα 3.13 MobileNet Model Summary part2

Εικόνα 3.14 MobileNet Model Summary part3

Εικόνα 3.15 MobileNet Model Summary part4

Εικόνα 3.16 MobileNet Model Summary part5

Εικόνα 3.17 VGG16 Αρχιτεκτονική

Εικόνα 3.18 VGG16 Model Summary part 1 (Model Keras Functional API)

Εικόνα 3.19 VGG16 Model Summary part 2 (Model Keras Functional API)

Εικόνα 3.20 VGG16 Model Summary part 1 (Sequential Model)

Εικόνα 3.21 VGG16 Model Summary part 2 (Sequential Model)

Εικόνα 4.1 Vgg16 Model Accuracy/Epoch

Εικόνα 4.2 Vgg16 Model Loss/Epoch

Εικόνα 4.3 MobileNet Model Accuracy/Epoch

Εικόνα 4.4 MobileNet Model Loss/Epoch

Εικόνα 4.5 AlexNet Model Accuracy/Epoch

Εικόνα 4.6 AlexNet Model Loss/Epoch

Εικόνα 4.7 Invasive Ductal Carcinoma (IDC)

Λίστα Σχημάτων

Σχήμα 1.1: Computer Aided Diagnosis Schematic

Σχήμα 1.2: Training & Testing Neural Net CNN

Σχήμα 2.1: Naïve Bayes Classification

Σχήμα 2.2: Support Vector Machine

Σχήμα 2.3: Machine Learning Classification

Σχήμα 2.4: Deep Neural Network

Σχήμα 3.1: Dataset Split Ratio

Λίστα Πινάκων

Πίνακας 3.1 AlexNet

Πίνακας 4.1 Confusion Matrix VGG16

Πίνακας 4.2 Classification Report VGG16

Πίνακας 4.3 Confusion Matrix MobileNet

Πίνακας 4.4 Classification Report MobileNet

Πίνακας 4.5 Confusion Matrix AlexNet

Πίνακας 4.6 Classification Report AlexNet

Πίνακας 4.7 Classification Report 1st Dataset Vgg16

Πίνακας 4.8 Confusion Matrix Vgg16 1st Dataset

Πίνακας 4.9 Classification Report 2nd Dataset Vgg16

Πίνακας 4.10 Confusion Matrix Vgg16 2nd Dataset

Πίνακας 4.11 Classification Report 1st Dataset MobileNet

Πίνακας 4.12 Confusion Matrix MobileNet 1st Dataset

Πίνακας 4.13 Classification Report 2nd Dataset MobileNet

Πίνακας 4.14 Confusion Matrix MobileNet 2nd Dataset

Πίνακας 4.15 Classification Report 1st Dataset MobileNet

Πίνακας 4.16 AlexNet Confusion Matrix 1st Dataset

Πίνακας 4.17 AlexNet Classification Report 2nd Dataset

1

Εισαγωγή

1.1 Κίνητρα

Σύμφωνα με αναφορές του Παγκόσμιου Οργανισμού Υγείας (World Health Organization) , ο καρκίνος του μαστού είναι ο πιο κοινός καρκίνος μεταξύ των γυναικών. Αυτή η παθολογία είναι η πρώτη σημαντική αιτία του θανάτου στις πάσχουσες γυναίκες (Παγκόσμιος Οργανισμός Υγείας). Η έγκαιρη ανίχνευση του καρκίνου του μαστού απαιτείται για την αποτελεσματική διάγνωση και θεραπεία.

Ο καρκίνος του μαστού (Breast Cancer) είναι ο συχνότερος καρκίνος στις γυναίκες και επηρεάζει περίπου το 10% όλων των γυναικών σε ορισμένα στάδια της ζωής τους. Στη σύγχρονη εποχή, το ποσοστό συνεχίζει να αυξάνεται και τα δεδομένα δείχνουν ότι το ποσοστό επιβίωσης είναι

88% μετά από πέντε χρόνια από τη διάγνωση και 80% μετά από 10 χρόνια από τη διάγνωση. Η πρόωρη πρόγνωση του καρκίνου του μαστού μέχρι σήμερα έχει κάνει συνταρακτική βελτίωση, έχοντας μειώσει το ποσοστό θνησιμότητας που οφείλεται στον καρκίνο του μαστού κατά 39%, αρχής γενομένης από το 1989. Λόγω της ποικίλης φύσης των συμπτωμάτων του καρκίνου του μαστού, οι ασθενείς υποβάλλονται συχνά σε πολλαπλές δοκιμές, σε μαστογραφία, υπερηχογράφημα και βιοψία, για να ελέγξουν την πιθανότητα διάγνωσης του καρκίνου του μαστού. Η φύση του προβλήματος είναι τέτοια που καθιστά τις εικόνες έναν από τους πιο αποτελεσματικούς τρόπους για ανάλυση και διάγνωση των όγκων.

Οι μαστογραφίες και οι υπέρηχοι του μαστού είναι ο τρόπος καταγραφής των εικόνων και κατά συνέπεια συλλογής των στοιχείων και δεδομένων που είναι απαραίτητα για να απευθυνθούμε εύστοχα στο παρόν ζήτημα. Γίνονται πολύ μεγάλες προσπάθειες για τη δημιουργία και βελτίωση ήδη υπάρχοντων CAD (Computer Aided Diagnosis) συστημάτων τα οποία θα βοηθήσουν τους επιστήμονες στον τομέα της ιατρικής να αμβλύνουν το πρόβλημα, που σχετίζεται με τον καρκίνο, με την έγκαιρη και εύστοχη αναγνώρισή του. Η ταχύτητα και η ενίσχυση ακρίβειας πρόβλεψης αποτελούν τους δύο πιο σημαντικούς παράγοντες στους τομείς που σχετίζονται με τα θέματα υγείας.

Επιπλέον, μία διαφορετική προσέγγιση του ζητήματος είναι ότι η πληθώρα δειγμάτων που θα χρησιμοποιούμε κάνει την προσέγγιση αρκετά πιο επιτυχημένη, καθώς έχοντας πολλά δείγματα τα συστήματα θα καλύπτουν το μεγαλύτερο εύρος περιπτώσεων και τα ποσοστά αστοχίας θα τείνουν στο ελάχιστο δυνατό. Ένας μεγάλος αριθμός υπερήχων πραγματοποιείται κάθε μέρα, γεγονός που καθιστά δύσκολο το έργο της ανάλυσης της εικόνας, επειδή ένας ακτινολόγος δεν μπορεί να αναλύσει εκατοντάδες εικόνες με την ίδια ακρίβεια και σε ελάχιστο χρόνο. Επομένως, η ανάπτυξη συστημάτων διάγνωσης με τη βοήθεια υπολογιστή (CAD), τα οποία μπορούν να βοηθήσουν το ιατρικό προσωπικό με την έγκαιρη ανίχνευση του καρκίνου, αποτελούν μια κρίσιμη εναλλακτική λύση.

Τα συστήματα ανίχνευσης με υπολογιστή Computer Aided Diagnosis (CADx), που ονομάζονται επίσης Computer Assisted Detection (CADE), είναι συστήματα που βοηθούν τους γιατρούς στην ερμηνεία των ιατρικών εικόνων. Οι τεχνικές απεικόνισης στις ακτίνες X, τη μαγνητική τομογραφία και τη διάγνωση υπερήχων παρέχουν πολλές πληροφορίες που ο ακτινολόγος ή άλλος ιατρικός επαγγελματίας πρέπει να αναλύσει και να αξιολογήσει συνολικά σε σύντομο χρονικό διάστημα. Τα συστήματα CAD επεξεργάζονται ψηφιακές εικόνες για τυπικές εμφανίσεις και υπογραμμίζουν εμφανή τμήματα, όπως πιθανές ασθένειες, προκειμένου να παράσχουν στοιχεία για την υποστήριξη μιας απόφασης του επαγγελματία. Το CAD έχει επίσης πιθανές μελλοντικές εφαρμογές στην ψηφιακή παθολογία με την εμφάνιση αλγορίθμων ολικής διαφάνειας και μηχανικής μάθησης. Μέχρι στιγμής η εφαρμογή της έχει περιορισθεί στην ποσοτικοποίηση ανοσοχρώσης, αλλά είναι επίσης αντικείμενο έρευνας για το πρότυπο H & E χρώσης. Το CAD είναι μια διεπιστημονική τεχνολογία που συνδυάζει στοιχεία της τεχνητής νοημοσύνης και της όρασης του υπολογιστή με την επεξεργασία εικόνων ακτινολογικής και παθολογικής. Μια τυπική εφαρμογή είναι η ανίχνευση ενός όγκου. Για παράδειγμα, ορισμένα νοσοκομεία χρησιμοποιούν CAD για να υποστηρίξουν προληπτικούς ιατρικούς ελέγχους διάγνωση καρκίνου του μαστού, ανίχνευση πολυπόδων στο κόλον και καρκίνο πνεύμονα.

Κατόπιν, η βαθιά εκμάθηση είναι ένας νέος τομέας της μηχανικής μάθησης και της τεχνητής νοημοσύνης. Τα τελευταία χρόνια, η βαθιά εκμάθηση έχει προσελκύσει την προσοχή σε διάφορους ερευνητικούς τομείς, όπως η υπολογιστική όραση, κατηγοριοποίηση εικόνων καθώς και η ανάλυση μεγάλου όγκου δεδομένων. Αυτή η μέθοδος πέτυχε κρίσιμα αποτελέσματα σε πολλές προκλήσεις, όπως το ImageNet Large Scale Visual Recognition Competition (ILSVRC).

Η τεχνητή νοημοσύνη (Artificial Intelligence) κερδίζει μεγάλη προσοχή για την εξαιρετική της απόδοση σε συστήματα αναγνώρισης εικόνας και εφαρμόζεται ολοένα και περισσότερο στο υπερηχογράφημα του μαστού. Κάθε ευφύες σύστημα μπορεί να διεξάγει μια ποσοτική αξιολόγηση

αναγνωρίζοντας αυτόματα τις πληροφορίες απεικόνισης και να κάνει πιο ακριβή και αναπαραγωγική απεικόνιση. Ως εκ τούτου, η χρήση της τεχνητής νοημοσύνης στην ανίχνευση και τον εντοπισμό του καρκίνου του μαστού έχει μεγάλη σημασία, η οποία δεν μπορεί μόνο να εξοικονομήσει χρόνο για ακτινολόγους, όπως αναφέρθηκε προηγουμένως, αλλά και να καλύψει την έλλειψη εμπειρίας και δεξιοτήτων σε μερικούς αρχάριους επιστήμονες.

Επιπλέον, όσον αφορά στα υπερηχογραφήματα, είναι ένα εργαλείο απεικόνισης πρώτης γραμμής για τον χαρακτηρισμό του προβλήματος του μαστού για τη μεγάλη διαθεσιμότητά του, την αποδοτικότητα του κόστους, τις αποδεδειγμένες διαγνωστικές επιδόσεις και τις μη επεμβατικές δυνατότητες σε πραγματικό χρόνο. Εκτός από το υπερηχογράφημα τύπου Β, νέες τεχνικές όπως το Doppler χρώματος, το φασματικό Doppler, το υπερηχογράφημα με έντονη αντίθεσή και η ελαστογραφία μπορούν επίσης να βοηθήσουν τους γιατρούς να αποκτήσουν πιο ακριβείς και εύστοχες πληροφορίες. Ωστόσο, ένα μειονέκτημα το οποίο θα μπορούσαμε να σημειώσουμε σχετικά με τα υπερηχογραφήματα είναι η ύπαρξη εξάρτησης από τον χειριστή του μηχανήματος, δεδομένο το οποίο προσθέτει πιθανότητα λάθους, βασισμένη στον ανθρώπινο παράγοντα

Η Τεχνητή Νοημοσύνη εφαρμόζεται όλο και περισσότερο σε υπερήχους και αποδείχθηκε ένα ισχυρό εργαλείο για την παροχή αξιόπιστης διάγνωσης με μεγαλύτερη ακρίβεια και αποτελεσματικότητα και μείωση του φόρτου εργασίας του επιστημονικού προσωπικού. Είναι κατά προσέγγιση χωρισμένο σε πρώιμη μηχανική μάθηση (Machine Learning) που ελέγχεται από αλγόριθμους χειροκίνητης εισόδου δεδομένων και βαθιά μάθηση (Deep Learning) κατά την οποία υπάρχει μεγάλο εύρος τρόπων εκμάθησης. Δεν υπάρχουν ακόμη κατευθυντήριες γραμμές για την υποβολή προτάσεων για την εφαρμογή αυτών των ευφών συστημάτων σε υπερηχογραφήματα στην κλινική πρακτική και χρειάζονται περισσότερες μελέτες για να διερευνηθούν πιο προηγμένες μέθοδοι και να αποδειχθεί η χρησιμότητά τους. Στο εγγύς μέλλον, πιστεύω ότι η τεχνητή νοημοσύνη στο υπερηχογράφημα του μαστού δεν θα μπορεί μόνο να διακρίνει τις καλοήθεις και κακοήθεις μαστικές

μάζες, αλλά και να ταξινομεί περαιτέρω συγκεκριμένες καλοήθεις ασθένειες, όπως μολυσματική μαστική μάζα και νεοπλασία.

1.2 Ορισμός Προβλήματος

Η ιατρική είναι μια πολύπλοκη και βαρυσήμαντη επιστήμη στην οποία βασίζονται τα θέματα υγείας που προκύπτουν για την εύστοχη και έγκαιρη επίλυσή τους. Καθώς, ακόμα ο κόσμος αντιμετωπίζει πολλαπλά προβλήματα τα οποία με τη σειρά τους ανεβάζουν το επίπεδο ανάγκης και κατά συνέπεια δυσκολίας συνεχώς, αναπτύχθηκε η απαίτηση να συνδυαστούν οι επιδόσεις των υπολογιστικών συστημάτων με την επιστήμη της ιατρικής έτσι ώστε να ανοίξουν κι άλλες πόρτες όσον αφορά στους τρόπους αντιμετώπισης των ζητημάτων που προκύπτουν διαρκώς.

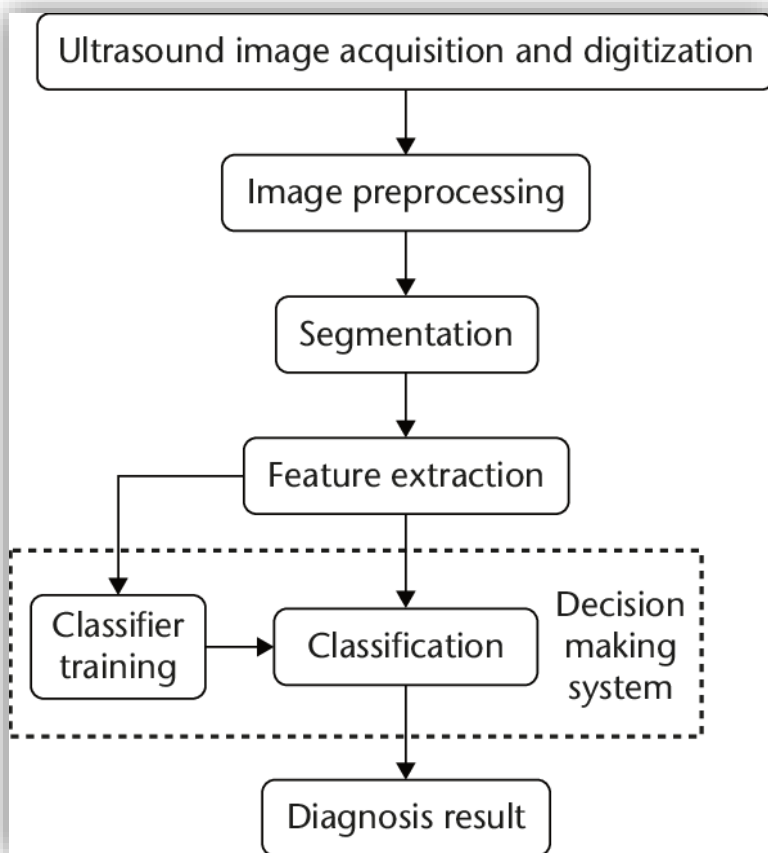
Η εξέταση των υπερηχογραφήμάτων του ασθενούς από παθολόγο είναι μια συμβατική μέθοδος για τη διάγνωση πολλών ασθενειών, ειδικά για καρκίνο, όπως ο καρκίνος του μαστού. Ωστόσο, η ανασκόπηση των δειγμάτων είναι επίπονη και χρονοβόρα, γεγονός που μπορεί να καθυστερήσει τη λήψη αποφάσεων. Η ανασκόπηση των διαφανειών παθολογίας είναι ένα πολύ περίπλοκο έργο. Μερικές φορές η συμφωνία στη διάγνωση ορισμένων μορφών καρκίνου του μαστού μπορεί να είναι τόσο χαμηλή που δεν μπορεί να υπάρξει ένα αρκετά εύστοχο συμπέρασμα. Η δυσκολία διάγνωσης ασθενειών από παθολόγους είναι αναπόφευκτη επειδή οι παθολόγοι πρέπει να αναθεωρήσουν όλες τις διαφάνειες ανά ασθενή και υπάρχει έντονο στοιχείο του ανθρώπινου παράγοντα.

Κατά αυτό τον τρόπο προκύπτουν τα υπολογιστικά συστήματα με χρήση ιατρικής υποβοήθειας. Αυτά τα συστήματα (CAD) επί της ουσίας είναι αλγόριθμοι οι οποίοι έχουν έναν ευφυή

τρόπο λειτουργίας ο οποίος αποκλείοντας το σφάλμα που οφείλεται στον ανθρώπινο παράγοντα μπορούν να επιφέρουν συνταρακτικά αποτελέσματα σε ένα πολύ εύλογο χρονικό διάστημα.

Πιο συγκεκριμένα τα υπολογιστικά συστήματα ανίχνευσης κάνουν χρήση τεχνολογιών τεχνητής νοημοσύνης κατά την οποία μαθαίνουν να αναγνωρίζουν και να βγάζουν συμπεράσματα από δεδομένα εισόδου τα οποία είναι απαιτούμενα ιατρικά στοιχεία.

Οπότε, το εν λόγω πρόβλημα είναι η δημιουργία ενός συστήματος τεχνητής νοημοσύνης το οποίο θα παίρνει ως δεδομένα εισόδου εικόνες από υπερηχογραφήματα με σκοπό να τα ανάγει σε ένα πρόβλημα κατηγοριοποίησης στις κλάσεις καλοηθών ή κακοηθών όγκων. Άρα το πρόβλημα ανάγεται σε ένα πρόβλημα classification μεταξύ δύο κατηγοριών και δεδομένα εικόνες.



Σχήμα 1 Computer Aided Diagnosis System

1.3 Στόχοι και Μεθοδολογία

Η πρώτη τεχνητή νοημοσύνη αναφέρεται κυρίως στην παραδοσιακή μηχανική μάθηση (machine learning). Επιλύει προβλήματα με δύο βήματα: ανίχνευση αντικειμένων και αναγνώριση αντικειμένων. Πρώτον, το μηχάνημα χρησιμοποιεί έναν αλγόριθμο ανίχνευσης πλαισίου οριοθέτησης για να σαρώσει ολόκληρη την εικόνα για να βρει την πιθανή περιοχή του αντικειμένου. Δεύτερον, ο αλγόριθμος αναγνώρισης αντικειμένου αναγνωρίζει το αντικείμενο με βάση το προηγούμενο βήμα.

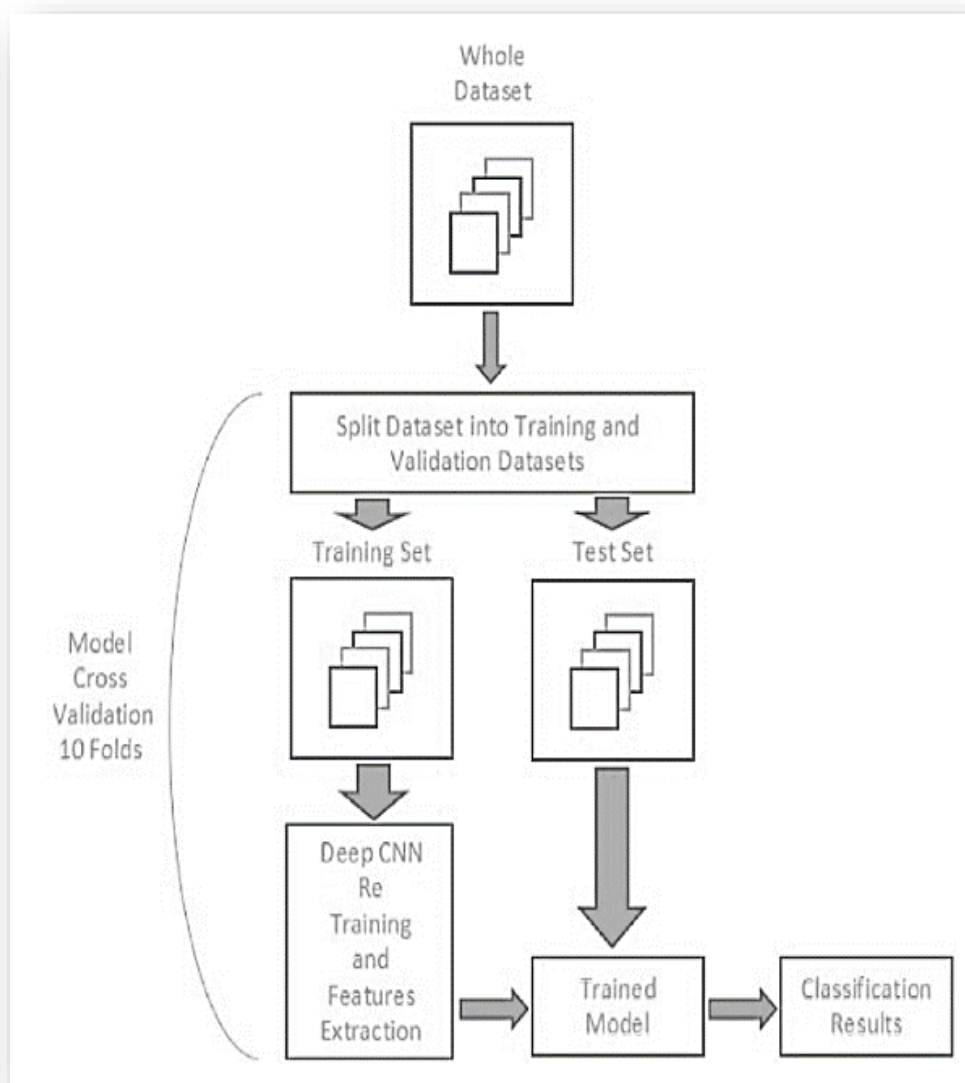
Στην συγκεκριμένη εργασία ο τρόπος που θα προσεγγίσουμε το πρόβλημα είναι μέσω διαχείρισης νευρωνικών δικτύων. Τα νευρωνικά δίκτυα (neural nets) αποτελούν μία σχετικά νέα περιοχή στις φυσικές επιστήμες, καθ' όσον έχουν γίνει γνωστά και έχουν αναπτυχθεί μόνο κατά τα τελευταία σαράντα περίπου χρόνια. Εν τούτοις, η περιοχή αυτή έχει δει μία μεγάλη άνθηση, κρίνοντας από την μεγάλη ανάπτυξη που έχει παρατηρηθεί, από τον αριθμό των επιστημόνων που ασχολούνται με αυτά τα θέματα, και βέβαια από τα πολύ σημαντικά επιτεύγματα, που έχουν συμβάλει στο να γίνουν γνωστά σε ένα ευρύτερο κύκλο. Αποτελούν επομένως ένα θέμα με μεγάλο ενδιαφέρον στις τεχνολογικές επιστήμες. Το κύριο χαρακτηριστικό τους είναι ότι οι πρώτες αρχές και λειτουργίες τους βασίζονται στο νευρικό σύστημα των ζώντων οργανισμών (και φυσικά του ανθρώπου), αλλά η μελέτη και η χρήση τους έχει προχωρήσει πολύ πέρα από τους βιολογικούς οργανισμούς, και σήμερα τα νευρωνικά δίκτυα χρησιμοποιούνται για να λύσουν κάθε είδους προβλήματα με ηλεκτρονικό υπολογιστή. Η φιλοσοφία τους όμως είναι διαφορετική από τον τρόπο με τον οποίο δουλεύουν οι κλασικοί υπολογιστές. Η λειτουργία τους προσπαθεί να συνδυάσει τον τρόπο σκέψης του ανθρώπινου εγκεφάλου με τον αφηρημένο μαθηματικό τρόπο σκέψης. Έχουν, λοιπόν, κατά κάποιο τρόπο μοντελοποιηθεί οι λειτουργίες των νευρώνων και κατά τα συνέπεια και ο τρόπος εκμάθησης δεδομένων

από εξωτερικό περιβάλλον. Ανάλογα και με το είδος της πληροφορίας έχουν δημιουργηθεί και διαφορετικού προσανατολισμού νευρωνικά δίκτυα. Στο deep learning, ένα παράδειγμα είναι ότι σε περίπτωση που η πληροφορία είναι εικόνα τότε καταλληλότερα είναι τα συνελικτικά νευρωνικά δίκτυα – Convolutional Neural Nets (CNN), ενώ σε μια άλλη περίπτωση που η πληροφορία είναι, για παράδειγμα, ηχητικό σήμα ή textual δεδομένα (ακολουθία γραμμάτων συμβόλων κειμένου), καταλληλότερα είναι τα επαναλαμβανόμενα νευρωνικά δίκτυα Recurrent Neural Nets (RNN)

Αρχικά ένας βασικός υποστόχος είναι η συγκέντρωση απαραίτητης πληροφορίας (data collection), έτσι ώστε να υποστεί την απαραίτητη επεξεργασία με σκοπό να δοθεί σαν είσοδο στο νευρωνικό δίκτυο και από αυτή τη συλλογή εικόνων να εκπαιδευτεί και κατά συνέπεια να κατανοήσει τις διαφορές μεταξύ τους.

Σε αυτήν τη μελέτη στοχεύω στο classification μαστογραφιών από υπερηχογραφήματα και εικόνων με ιστοπαθολογία με τη χρήση συνελικτικών νευρικών δικτύων. Εφαρμόζω τρεις διαφορετικές αρχιτεκτονικές CNN και συγκρίνω τα αποτελέσματα με βάση τις μετริกές, accuracy, precision sensitivity και f1-score. Η δομή της μεθόδου μου ξεκινά με τη συλλογή δεδομένων, ακολουθείται από την οργάνωση δεδομένων σε ομάδες training, testing και validation. Επιπλέον, το επόμενο και πιο κρίσιμο βήμα είναι η προεπεξεργασία των δεδομένων δεδομένων. Σε αυτό το βήμα ως προεπεξεργασία κάνω Image normalization and pixel scaling, data augmentation, and data resizing πριν από την τροφοδοσία των δεδομένων στα νευρωνικά δίκτυα. Στη συνέχεια φόρτωσα τα νευρωνικά δίκτυα VGG16, MobileNet και AlexNet και μετά από μια διαδικασία fine-tuning με τρόπο που τα προσαρμόσα ώστε να εξυπηρετούν τον συγκεκριμένο στόχο μου, ξεκίνησα την πειραματική φάση. Επί της ουσίας, σε αυτήν τη μελέτη, προσπαθώ να δώσω την οπτική μου για το πώς θα μπορούσαμε να αντιμετωπίσουμε στο εγγύς μέλλον ένα σοβαρό πρόβλημα όπως ο καρκίνος του μαστού ειδικά σε περιπτώσεις όπου υπάρχει περιορισμός στους υπολογιστικούς πόρους και πώς να συνεισφέρουμε σε αυτό το τεράστιο πεδίο έρευνας. Αυτό το δείχνω χρησιμοποιώντας στα πειράματά μου το δίκτυο

MobileNet το οποίο είναι ένα νευρωνικό δίκτυο διαφορετικής αρχιτεκτονικής (Depth-wise αρχιτεκτονική) η οποία είναι low computational cost και πιο γρήγορη σε σχέση με τα υπόλοιπα βαθιά συνελικτικά δίκτυα όπως το VGG16 και το AlexNet. Γενικότερα, παρουσιάζω και αξιολογώ τις επιδόσεις διαφόρων αρχιτεκτονικών CNN σε πολλά σύνολα δεδομένων διαφορετικών περιπτώσεων εξέτασης καρκίνου του μαστού. Συγκεκριμένα, σύγκρινα την απόδοση του depth-wise και low computational cost συνελικτικού νευρικού δικτύου με κάποια βαθιά συνελικτικά δίκτυα και ταυτόχρονα αξιολόγησα τη συνολική απόκρισή τους σε αυτά τα διαφορετικά ιατρικά σύνολα δεδομένων. Το πρώτο σύνολο δεδομένων είναι ένα σύνολο δεδομένων για τον καρκίνο του μαστού υπερήχων όπου εκπαιδεύσα όλα τα δίκτυα στο να μπορούν να διαφοροποιούν καλοήθεις όγκους από κακοήθεις και το καλύτερο αποτέλεσμα, που έλαβα ήταν ακρίβεια 96,82%. Το δεύτερο σύνολο δεδομένων ήταν περιοχές ενδιαφέροντος από εικόνες ιστολογίας με περιπτώσεις διεισδυτικού καρκίνου του πόρου (Invasive Ductal Carcinoma), ο πιο κοινός τύπος καρκίνου του μαστού και το καλύτερο αποτέλεσμα που λάβαμε ήταν 88,23%. Το τρίτο σύνολο δεδομένων ήταν και πάλι ιστολογικές εικόνες διεισδυτικού καρκίνου του πόρου με μεγαλύτερου όγκου δεδομένα, όπου το καλύτερο αποτέλεσμα ήταν ακρίβεια 91,04%.



Σχήμα 1.2 Training & Testing Neural Net CNN

1.4 Επιτομή Διπλωματικής

Ο καρκίνος του μαστού όπως και κάθε άλλο πρόβλημα αντίστοιχου βεληνεκούς χρήζει αντιμετώπισης και είναι πολύ σημαντικό να γίνονται διαρκώς μελέτες πάνω στους διαφορετικούς τρόπους προσέγγισής του. Το Machine Learning και γενικότερα η τεχνητή νοημοσύνη είναι ένας τομέας υπό συνεχή και ραγδαία εξέλιξη τα τελευταία χρόνια, ο οποίος αποτελεί πολύ σημαντικό βήμα στην επιτυχή αντιμετώπιση πολλαπλών ζητημάτων της ανθρωπότητας αλλά και μέσον βελτίωσης της καθημερινότητας των ανθρώπων μέσω ποικίλων εφαρμογών. Κατά αυτό τον τρόπο χρησιμοποιούνται οι πολλαπλές εφαρμογές και μελέτες ως πάτημα για την τεχνολογική εξέλιξη της ανθρωπότητας.

Η εν λόγω διατριβή δείχνει, επί της ουσίας, τις πολλά υποσχόμενες προοπτικές της τεχνητής νοημοσύνης και της βαθιάς μάθησης, μέσω ενός έργου classification υπερηχογραφήματων βασισμένο στον πολυδιάστατο χαρακτήρα των νευρωνικών δικτύων. Κατόπιν, αναδεικνύει τον εξέχον τρόπο με τον οποίο αναλύουν, τα διαφορετικών ειδών νευρωνικά δίκτυα και μόνο με τη βοήθεια μιας προεπεξεργασίας των εικόνων. Συγκρίνοντας τις επιδόσεις τους πάνω σε αυτήν την κατηγοριοποίηση υπερηχογραφήματων κατανοούμε ότι οι διαφορετικές αρχιτεκτονικές των διαφόρων συνελικτικών νευρωνικών επιφέρουν διαφορετικά αποτελέσματα, καθώς και το μέγεθος του δικτύου νευρώνων είναι ένα σημαντικό στοιχείο όσον αφορά στην αντιληπτική του ικανότητα.

2

Συναφείς Μελέτες & Γνωστικό Υπόβαθρο

2.1 Ανάπτυξη των Classifiers

Το πρόβλημα της κατηγοριοποίησης (Classification Problem) είναι ένα σημαντικό στοιχείο στον τομέα της βαθιάς μάθησης. Στόχος είναι να μπορεί να γίνει η, όσον το δυνατόν πιο έγκυρη τοποθέτηση δειγμάτων στις ανάλογες κατηγορίες. Το πρόβλημα κατηγοριοποίησης ονομάζεται επίσης εποπτευόμενη κατηγοριοποίηση (Supervised Classification), αφού όλα τα δείγματα έχουν επισημανθεί και όλες οι κατηγορίες είναι προκαθορισμένες. Ο κατηγοριοποιητής (classifier) είναι μία από τις εφαρμογές αναγνώρισης προτύπων. Ο πιο ευρέως χρησιμοποιούμενος κατηγοριοποιητής τεχνητής νοημοσύνης είναι το φίλτρο ηλεκτρονικού ταχυδρομείου ανεπιθύμητης αλληλογραφίας, το οποίο κατηγοριοποιεί κάθε μήνυμα ηλεκτρονικού ταχυδρομείου σε "κανονικό" ή "ανεπιθύμητο". Σε γενικές γραμμές, κάθε περίπτωση στο πρόβλημα κατηγοριοποίησης θα μετατραπεί σε ένα διάνυσμα που

μπορεί να αναλυθεί από τον υπολογιστή, ο οποίος συνήθως ονομάζεται "features". Ένα feature μπορεί να είναι μια απαρίθμηση ή ένας αριθμός.

2.1.1 Naive Bayes Classifier

Ο Naive Bayes classifier προτάθηκε το 1950. Πρόκειται για μια ομάδα απλών classifiers που προέρχονται από το Θεώρημα Bayes, υποθέτοντας ότι όλα τα χαρακτηριστικά των δειγμάτων είναι έντονα ανεξάρτητα. Από τότε που δημοσιεύτηκε έχει ευρέως μελετηθεί και χρησιμοποιηθεί. Τα πράγματα αποδείχτηκαν ότι τα πήγαν καλά για κατηγοριοποίηση κειμένου, με τον αριθμό εμφάνισης λέξεων ως χαρακτηριστικών. Μπορεί να κάνει την προαναφερθείσα εργασία κατηγοριοποίησης ηλεκτρονικού ταχυδρομείου σε σχετικά χαμηλό υπολογιστικό ποσό σε σύγκριση με τους πιο πρόσφατους αλγορίθμους ενώ παράλληλα επιτυγχάνει αποδεκτή ακρίβεια [1]. Με κατάλληλη προεπεξεργασία, εξακολουθεί να είναι ανταγωνιστική.

$$P(c | x) = \frac{P(x | c)P(c)}{P(x)}$$

Labels in the diagram:
- Likelihood (points to $P(x | c)$)
- Class Prior Probability (points to $P(c)$)
- Posterior Probability (points to $P(c | x)$)
- Predictor Prior Probability (points to $P(x)$)

$$P(c | X) = P(x_1 | c) \times P(x_2 | c) \times \dots \times P(x_n | c) \times P(c)$$

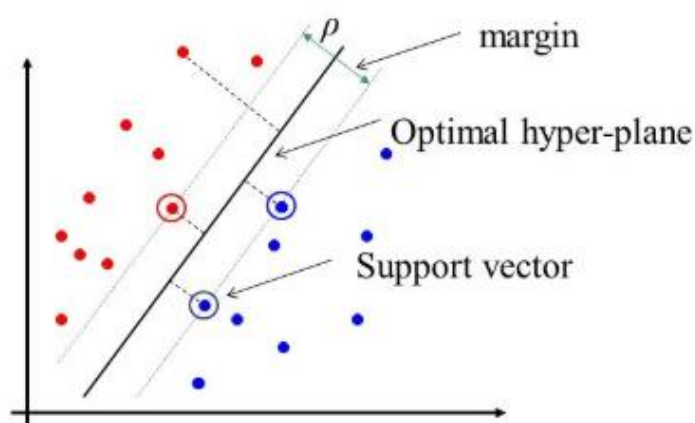
Σχήμα 2.1 Naïve Bayes Classifier

Ο Naive Bayes classifier μπορεί να έχει διαφορετικές υποθέσεις για την υποκείμενη διανομή χαρακτηριστικών. Για συνεχείς μεταβλητές, μπορούμε να υποθέσουμε ότι βρίσκονται κάτω από την κλασική Gaussian κατανομή. Για τα δεδομένα κειμένου, η τυπική παραδοχή είναι πολυωνυμική κατανομή, όπου λαμβάνεται υπόψη ο αριθμός εμφάνισης μιας λέξης. Μια απλοποιημένη έκδοση είναι η διανομή Bernoulli, η οποία εξετάζει μόνο αν εμφανίζεται ή όχι μια λέξη. Ο classifier Naive Bayes είναι πολύ πιο εκτεταμένος από άλλους αλγόριθμους. Ο αριθμός των παραμέτρων που χρειάζεται να μάθει είναι γραμμικός σε αριθμό χαρακτηριστικών, επομένως η πολυπλοκότητα του χρόνου εκπαίδευσης είναι επίσης γραμμική. Επιπλέον, η διαδικασία κατάρτισης έχει μια πολύ στενή έκφραση. Για πρόβλημα κατάταξης μέσω ηλεκτρονικού ταχυδρομείου, ο αριθμός των παραμέτρων είναι απλώς ο αριθμός των μοναδικών λέξεων σε όλα τα μηνύματα ηλεκτρονικού ταχυδρομείου. Αυτό αποφεύγεται η δαπανηρή γραμμική προσέγγιση που χρησιμοποιούν πολλοί άλλοι κατηγοριοποιητές.

2.1.2 Support Vector Machine

Η μεταγενέστερη Support Vector Machine (SVM) εισήχθη από τον Vladimir Naumovich Vapnik και τον Alexey Yakovlevich Chervonenkis [2]. Με δεδομένο ένα εκπαιδευτικό σύνολο, κάθε δείγμα αντιπροσωπεύεται από ένα σημείο στον Ευκλείδειο χώρο. Στη μηχανική μάθηση, τα μοντέλα είναι υπό εποπτεία μαθησιακά μοντέλα με συναφείς αλγόριθμους εκμάθησης που αναλύουν δεδομένα που χρησιμοποιούνται για ανάλυση κατηγοριοποίησης και παλινδρόμησης. Λαμβάνοντας υπόψη ένα σύνολο κατάρτισης ως παραδείγματα, κάθε σήμανση ότι ανήκουν σε μία ή την άλλη από τις δύο κατηγορίες, ένας αλγόριθμος εκπαίδευσης SVM χτίζει ένα μοντέλο που εικχωρεί νέα παραδείγματα στη μία ή στην άλλη κατηγορία. Ένα μοντέλο SVM είναι μια αναπαράσταση των παραδειγμάτων ως σημεία στο διάστημα, χαρτογραφημένα έτσι ώστε τα παραδείγματα των ξεχωριστών κατηγοριών χωρίζονται από ένα σαφές κενό που είναι όσο το δυνατόν ευρύτερο. Στη συνέχεια, νέα παραδείγματα

χαρτογραφούνται στον ίδιο χώρο και προβλέπεται να ανήκουν σε μια κατηγορία που βασίζεται στην πλευρά του χάσματος επί του οποίου πέφτουν. Εκτός από την εκτέλεση γραμμικής ταξινόμησης, τα SVM μπορούν να εκτελούν αποτελεσματικά μια μη γραμμική κατηγοριοποίηση χρησιμοποιώντας αυτό που ονομάζεται κόλπο του πυρήνα, χαρτογραφώντας τις εισόδους τους σε χώρους μεγάλης διάστασης. Όταν τα δεδομένα δεν έχουν επισημανθεί, η εποπτευόμενη μάθηση δεν είναι δυνατή και απαιτείται μια μη εποπτευόμενη μαθησιακή προσέγγιση, η οποία προσπαθεί να βρει φυσική ομαδοποίηση των δεδομένων σε ομάδες και στη συνέχεια να χαρτογραφήσει νέα δεδομένα σε αυτές τις σχηματισμένες ομάδες. Ο αλγόριθμος SVM [2], που δημιουργήθηκε από τον Hava Siegelmann και τον Vladimir Vapnik, εφαρμόζει τα στατιστικά στοιχεία των διανυσμάτων υποστήριξης που αναπτύσσονται στον αλγόριθμο για την κατηγοριοποίηση μη επισημασμένων δεδομένων και είναι ένας από τους πιο ευρέως χρησιμοποιούμενους αλγόριθμους ομαδοποίησης σε βιομηχανικές εφαρμογές. Το SVM είναι συνήθως ένας γραμμικός classifier. Ωστόσο, με μερικά κόλπα που ονομάζονται "κόλπο πυρήνα", το SVM μπορεί επίσης να κάνει μη γραμμική κατηγοριοποίηση. Η βασική ιδέα είναι, με τη χαρτογράφηση του αρχικού χώρου δείγματος σε ένα υψηλότερο χώρο διαστάσεων, το αρχικό μη γραμμικό διαχωριζόμενο σύνολο να μπορεί να γίνει διαχωρίσιμο.



Σχήμα 2.2 Support Vector Machine

2.2 Machine Learning

2.2.1 Ορισμός

Η μηχανική μάθηση είναι η επιστημονική μελέτη των αλγορίθμων και στατιστικών μοντέλων που τα συστήματα υπολογιστών χρησιμοποιούν για να εκτελέσουν μια συγκεκριμένη εργασία χωρίς τη χρήση ρητών οδηγιών. Θεωρείται τομέας ο οποίος ανήκει στον κλάδο της τεχνητής νοημοσύνης. Οι αλγόριθμοι μηχανικής μάθησης δημιουργούν ένα μαθηματικό μοντέλο που βασίζεται σε δεδομένα δείγματος, γνωστό ως " training data ", προκειμένου να προβεί σε προβλέψεις ή αποφάσεις χωρίς να έχει προγραμματιστεί ρητά για την εκτέλεση της εργασίας. Οι αλγόριθμοι μηχανικής μάθησης χρησιμοποιούνται σε μια ευρεία ποικιλία εφαρμογών, όπως το φιλτράρισμα ηλεκτρονικού ταχυδρομείου και την υπολογιστική όραση, όπου είναι δύσκολο ή ανέφικτο να αναπτυχθεί ένας συμβατικός αλγόριθμος για την αποτελεσματική εφαρμογή αυτής της εργασίας.

Η μηχανική μάθηση σχετίζεται στενά με τις στατιστικές υπολογιστικής, η οποία επικεντρώνεται στην πραγματοποίηση προβλέψεων με τη χρήση ηλεκτρονικών υπολογιστών. Η μελέτη της μαθηματικής βελτιστοποίησης παρέχει μεθόδους, θεωρία και τομείς εφαρμογής στον τομέα της μηχανικής μάθησης. Η εξόρυξη δεδομένων είναι ένα πεδίο σπουδών στο πλαίσιο της μηχανικής μάθησης και επικεντρώνεται στην διερευνητική ανάλυση δεδομένων μέσω της μάθησης χωρίς επίβλεψη. Στην εφαρμογή του σε διάφορα επιχειρηματικά προβλήματα, η μηχανική μάθηση αναφέρεται επίσης ως προγνωστική ανάλυση. Το 1959, ο Άρθουρ Σάμιουελ ορίζει τη μηχανική μάθηση ως "Πεδίο μελέτης που δίνει στους υπολογιστές την ικανότητα να μαθαίνουν, χωρίς να έχουν ρητά προγραμματιστεί".[5] Η μηχανική μάθηση διερευνά τη μελέτη και την κατασκευή αλγορίθμων που μπορούν να μαθαίνουν από τα δεδομένα και να κάνουν προβλέψεις σχετικά με αυτά. Τέτοιοι αλγόριθμοι λειτουργούν κατασκευάζοντας μοντέλα από πειραματικά δεδομένα, προκειμένου να

κάνουν προβλέψεις βασιζόμενες στα δεδομένα ή να εξάγουν αποφάσεις που εκφράζονται ως το αποτέλεσμα.

Η μηχανική μάθηση είναι στενά συνδεδεμένη και συχνά συγχέεται με υπολογιστική στατιστική, ένας κλάδος, που επίσης επικεντρώνεται στην πρόβλεψη μέσω της χρήσης των υπολογιστών. Έχει ισχυρούς δεσμούς με την μαθηματική βελτιστοποίηση, η οποία παρέχει μεθόδους, τη θεωρία και τομείς εφαρμογής. Η Μηχανική μάθηση εφαρμόζεται σε μια σειρά από υπολογιστικές εργασίες, όπου τόσο ο σχεδιασμός όσο και ο ρητός προγραμματισμός των αλγορίθμων είναι ανέφικτος. Παραδείγματα εφαρμογών αποτελούν τα φίλτρα spam (spam filtering), η οπτική αναγνώριση χαρακτήρων (OCR), οι μηχανές αναζήτησης και η υπολογιστική όραση. Η Μηχανική μάθηση μερικές φορές συγχέεται με την εξόρυξη δεδομένων, όπου η τελευταία επικεντρώνεται περισσότερο στην εξερευνητική ανάλυση των δεδομένων, γνωστή και ως μη επιτηρούμενη μάθηση.

Στο πεδίο της ανάλυσης δεδομένων, η μηχανική μάθηση είναι μια μέθοδος που χρησιμοποιείται για την επινόηση πολύπλοκων μοντέλων και αλγορίθμων που οδηγούν στην πρόβλεψη. Τα αναλυτικά μοντέλα επιτρέπουν στους ερευνητές, τους επιστήμονες δεδομένων, τους μηχανικούς και τους αναλυτές να παράγουν αξιόπιστες αποφάσεις και αποτελέσματα και να αναδείξουν αλληλοσυσχετίσεις μέσω της μάθησης από ιστορικές σχέσεις και τάσεις στα δεδομένα.[3],[4]

2.2.2 Classification & Regression in Machine Learning

Οι επιστήμονες χρησιμοποιούν πολλούς διαφορετικούς τύπους αλγορίθμων μηχανικής μάθησης για να ανακαλύψουν μοτίβα σε μεγάλα δεδομένα που οδηγούν σε ενέργειες που μπορούν να ασκηθούν. Σε υψηλό επίπεδο, αυτοί οι διαφορετικοί αλγόριθμοι μπορούν να ταξινομηθούν σε δύο ομάδες βάσει του τρόπου με τον οποίο «μαθαίνουν» τα δεδομένα για να κάνουν προβλέψεις: μάθηση υπό επίβλεψη και χωρίς επίβλεψη.

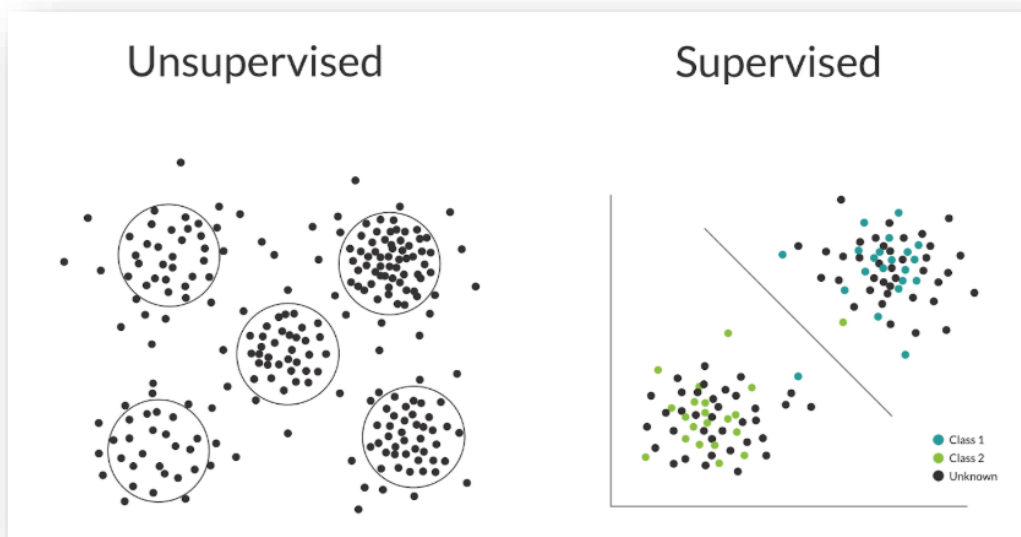
Supervised Learning

Η πλειονότητα της πρακτικής μηχανικής μάθησης χρησιμοποιεί supervised learning. Η εποπτευόμενη μάθηση είναι όπου έχετε μεταβλητές εισόδου (x) και μεταβλητή εξόδου (Y) και χρησιμοποιείται ένας αλγόριθμος για να μάθουμε τη λειτουργία χαρτογράφησης από την είσοδο στην έξοδο $Y = f(X)$. Ο στόχος είναι να προσεγγίσουμε τη λειτουργία χαρτογράφησης τόσο καλά ώστε όταν έχουμε νέα δεδομένα εισόδου (x) που μπορούμε να προβλέψουμε τις μεταβλητές εξόδου (Y) για αυτά τα δεδομένα. Οι τεχνικές των εποπτευόμενων αλγορίθμων μάθησης μηχανών περιλαμβάνουν γραμμική και λογική παλινδρόμηση, ταξινόμηση πολλαπλών κλάσεων, δέντρα αποφάσεων και μηχανήματα φορέα υποστήριξης. Το supervised learning ή αλλιώς εποπτευόμενη μάθηση απαιτεί τα δεδομένα που χρησιμοποιούνται για την εκπαίδευση του αλγορίθμου να έχουν ήδη επισημανθεί με σωστές απαντήσεις. Για παράδειγμα, ένας αλγόριθμος ταξινόμησης θα μάθει να αναγνωρίζει τα ζώα αφού εκπαιδεύεται σε ένα σύνολο δεδομένων εικόνων που φέρουν σωστή επισήμανση με τα είδη του ζώου και κάποια χαρακτηριστικά ταυτοποίησης. Τα εποπτευόμενα προβλήματα μάθησης μπορούν να ομαδοποιηθούν περαιτέρω σε προβλήματα παλινδρόμησης και ταξινόμησης. Και τα δύο προβλήματα έχουν ως στόχο την κατασκευή ενός συνοπτικού μοντέλου που μπορεί να προβλέψει την αξία του εξαρτώμενου χαρακτηριστικού από τις μεταβλητές χαρακτηριστικών. Η διαφορά μεταξύ των δύο κατηγοριών είναι το γεγονός ότι το εξαρτώμενο χαρακτηριστικό είναι αριθμητικό για παλινδρόμηση και κατηγορηματικό για ταξινόμηση.

Unsupervised Learning

Το unsupervised learning είναι η εκπαίδευση συστήματος που χρησιμοποιεί πληροφορίες που δεν είναι ούτε κατηγοριοποιημένες ούτε επισημασμένες και επιτρέπουν στον αλγόριθμο να ενεργεί σε αυτές τις πληροφορίες χωρίς καθοδήγηση. Σε αυτήν την περίπτωση μάθησης ο σκοπός του συστήματος

είναι να συγκεντρώνει τις ασαφείς πληροφορίες σύμφωνα με ομοιότητες, μοτίβα και διαφορές χωρίς προηγούμενη εκπαίδευση των δεδομένων. Οι unsupervised ή αλλιώς μη εποπτευόμενοι αλγόριθμοι μάθησης μάς επιτρέπουν να εκτελούμε πιο πολύπλοκα καθήκοντα επεξεργασίας σε σύγκριση με την εποπτευόμενη μάθηση. Αν και η μη εποπτευόμενη μάθηση μπορεί να είναι πιο απρόβλεπτη σε σύγκριση με άλλες φυσικές μεθόδους μάθησης. Δύο από τις κύριες μεθόδους που χρησιμοποιούνται στη μάθηση χωρίς επίβλεψη αποτελούν κύρια ανάλυση συστατικών στοιχείων και συμπλέγματος. Η ανάλυση συμπλέγματος χρησιμοποιείται σε μάθηση χωρίς επίβλεψη για την ομαδοποίηση ή την ταξινόμηση σειρών δεδομένων με κοινά χαρακτηριστικά για την παρεμβολή αλγοριθμικών σχέσεων. Η ανάλυση συμπλέγματος είναι ένας κλάδος της μηχανικής μάθησης που ομαδοποιεί τα δεδομένα που δεν έχουν επισημανθεί, ταξινομηθεί ή κατηγοριοποιηθεί. Αντί να αντιδράσει σε ανατροφοδότηση, η ανάλυση συμπλέγματος εντοπίζει τα κοινά στοιχεία στα δεδομένα και αντιδρά με βάση την παρουσία ή την απουσία τέτοιων κοινών στοιχείων σε κάθε νέο κομμάτι δεδομένων. Αυτή η προσέγγιση βοηθά στην ανίχνευση ανώμαλων σημείων δεδομένων που δεν ταιριάζουν σε καμία ομάδα.[6]



Σχήμα 2.3 Machine Learning Classification

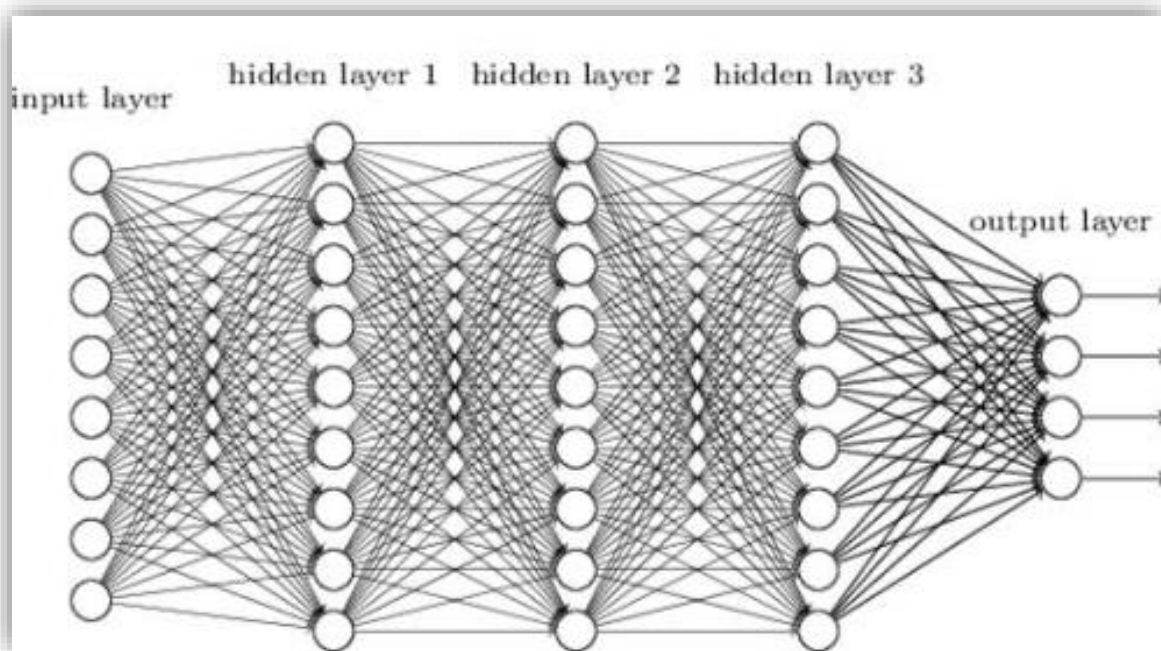
2.3 Deep Learning

2.3.1 Ορισμός & ανάπτυξη του Deep Learning

Η βαθιά μάθηση (επίσης γνωστή ως βαθιά δομημένη μάθηση ή διαφορικός προγραμματισμός) είναι μέρος μιας ευρύτερης οικογένειας μεθόδων μηχανικής μάθησης που βασίζονται σε τεχνητά νευρωνικά δίκτυα με εκπροσώπηση μάθησης. Η μάθηση μπορεί να εποπτεύεται (Supervised Learning), να εποπτεύεται μερικώς (Semi-Supervised Learning) ή να μην εποπτεύεται (Unsupervised). Η deep learning αρχιτεκτονικές ή αλλιώς αρχιτεκτονικές βαθιάς μάθησης όπως τα βαθιά νευρωνικά δίκτυα (Deep Neural Nets), τα επαναλαμβανόμενα νευρωνικά δίκτυα (RNN) και τα συνελικτικά νευρωνικά δίκτυα (CNN) έχουν εφαρμοστεί σε τομείς όπως ο υπολογιστής, η αναγνώριση ομιλίας, η επεξεργασία φυσικής γλώσσας, η αναγνώριση ήχου, το φιλτράρισμα κοινωνικών δικτύων, η μηχανική μετάφραση, η βιοπληροφορική, την ανάλυση ιατρικών εικόνων, τα προγράμματα επιθεώρησης υλικών και επιτραπέζιων παιχνιδιών, όπου παρήγαγαν συγκρίσιμα αποτελέσματα και, σε ορισμένες περιπτώσεις, υπερέβησαν τις επιδόσεις των εμπειρογνομόνων του ανθρώπου, [7], [8], [9].

Η βαθιά μάθηση είναι ένα υποσύνολο της μηχανικής μάθησης. Πρόκειται για μια οικογένεια αλγορίθμων μάθησης χαρακτηριστικών στον τομέα της μηχανικής μάθησης. Οι τιμές παρατήρησης μπορούν να εκπροσωπούνται με διάφορους τρόπους, όπως ένας φορέας που περιέχει τις τιμές RGB κάθε εικονοστοιχείου, ή πιο αφηρημένα μια σειρά ακμών και περιοχών [10]. Προσπαθεί να κάνει έναν εξαιρετικά αφηρημένο υπολογισμό δεδομένων με πολλαπλά στρώματα διεργασίας που μπορεί να περιέχουν περίπλοκη δομή ή μη γραμμική χαρτογράφηση. Γενικά, είναι μια πιο αφηρημένη μέθοδος μάθησης, καθώς δεν είναι συγκεκριμένη για οποιαδήποτε εργασία. Υπάρχουν ήδη πολλά ευρέως χρησιμοποιούμενα πλαίσια ευρείας μάθησης, όπως το βαθύ νευρωνικό δίκτυο, το συνελικτικό νευρωνικό δίκτυο και το αναδρομικό νευρωνικό δίκτυο. Η βαθιά εκμάθηση έχει χρησιμοποιηθεί ευρέως σε εφαρμογές, συμπεριλαμβανομένης της ορατότητας του υπολογιστή, της επεξεργασίας

φυσικής γλώσσας και της βιοπληροφορικής, και επιτυγχάνει τα υπέρτατα αποτελέσματα. Το 1989, ο Yann LeCun [11] πρότεινε τη λειτουργία βαθιάς εκμάθησης. Μέσα από αυτό θα μπορούσε να τρέξει, το κόστος υπολογισμού ήταν τόσο μεγάλο που η εκπαίδευση χρειάστηκε περίπου τρεις ημέρες. Επομένως, η πρώτη προσπάθεια βαθιάς εκμάθησης απέτυχε στην πραγματική εφαρμογή. Η τάση του ΑΙ στη συνέχεια μετατοπίστηκε σε Υποστηρίξη Vector Machine. Ωστόσο, το 1992, ο Schmidhuber [12] πρότεινε έναν αποτελεσματικό αλγόριθμο για την εκπαίδευση νευρωνικών δικτύων. Αυτός ο αλγόριθμος αντιμετωπίζει κάθε στρώμα στο δίκτυο ως μη επιτηρούμενο και, στη συνέχεια, συντονίζει τις παραμέτρους του με αλγόριθμο επιτήρησης οπίσθιας διάδοσης. Στο πείραμα, αποδείχθηκε ότι αυτή η μέθοδος κατάρτισης μπορεί πράγματι να βελτιώσει την ταχύτητα της αμαξοστοιχίας υπό εποπτεία. Το πλεονέκτημα της βαθιάς μάθησης είναι ότι χρησιμοποιεί αποτελεσματική μη επιτηρούμενη ή ημι-εποπτευόμενη μάθηση χαρακτηριστικών γνώσεων και εξαγωγή στρωματοποιημένων χαρακτηριστικών αντί για εξαγωγή χαρακτηριστικών από τον άνθρωπο. Ο στόχος της μάθησης χαρακτηριστικών είναι η αναζήτηση καλύτερης αντιπροσώπευσης των δεδομένων και η δημιουργία καλύτερου μοντέλου για να μάθουν αυτές τις αναπαραστάσεις από μεγάλης κλίμακας μη ετικέτα σύνολο δεδομένων. Η αντιπροσώπευση είναι σαν την ανάπτυξη πραγματικού νευρικού δικτύου και βασίζεται στην κατανόηση του τρόπου με τον οποίο οι πληροφορίες επεξεργάζονται και μεταδίδονται σε νευρωνικά συστήματα [12].

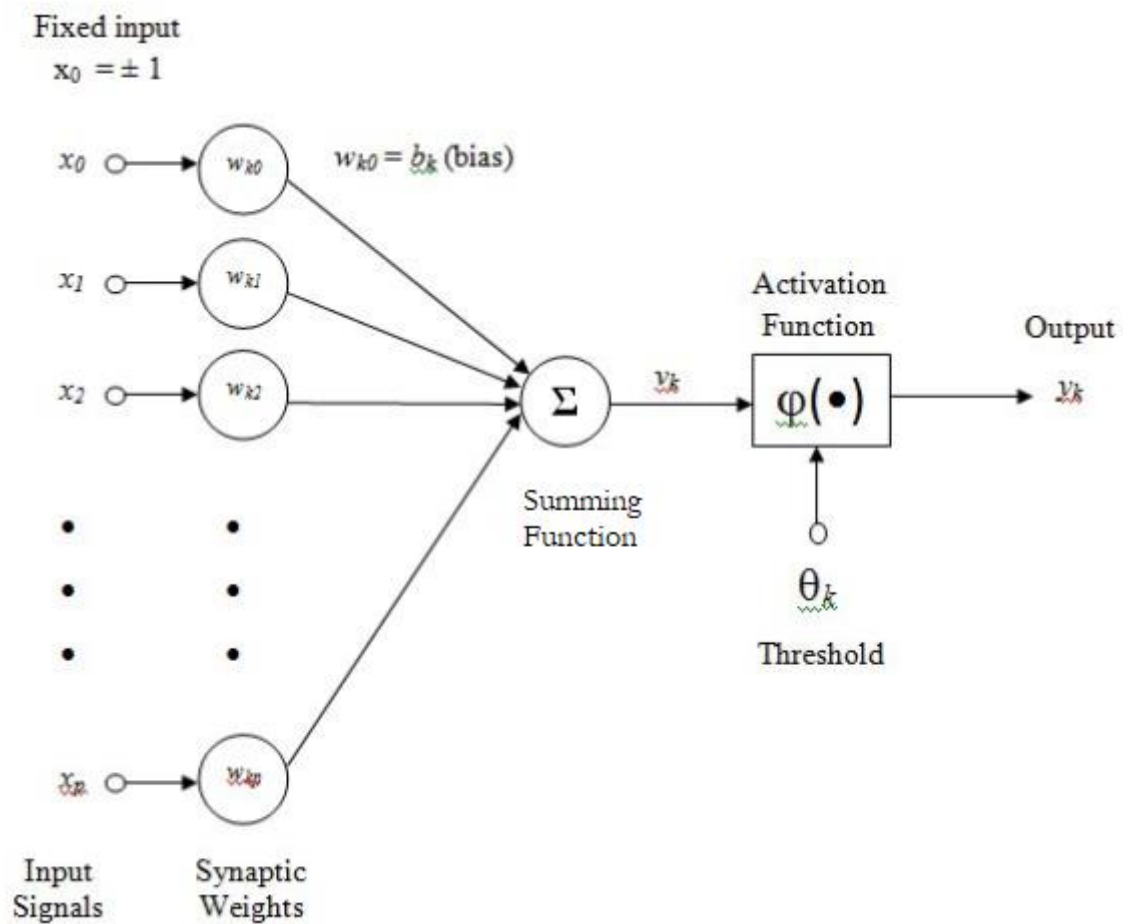


Σχήμα 2.4 Deep Neural Network

2.3.2 Artificial Neural Networks

Τα τεχνητά νευρωνικά δίκτυα ή Artificial Neural Networks (ANN) είναι υπολογιστικά συστήματα ασαφώς εμπνευσμένα από τα βιολογικά νευρωνικά δίκτυα που αποτελούν εγκέφαλο των ζώων. [13] Αυτά τα συστήματα "μαθαίνουν" να εκτελούν εργασίες, εξετάζοντας παραδείγματα, γενικά χωρίς να προγραμματίζονται με κανόνες συγκεκριμένους στόχους. Για παράδειγμα, στην αναγνώριση εικόνων μπορεί να μάθουν να εντοπίζουν εικόνες που περιέχουν γάτες αναλύοντας παραδείγματα εικόνων που έχουν επισημανθεί με το χέρι ως "γάτα" ή "μη γάτα" και χρησιμοποιώντας τα αποτελέσματα για τον εντοπισμό των γατών σε άλλες εικόνες. Το κάνουν αυτό χωρίς προηγούμενη γνώση των γατών, για παράδειγμα, ότι έχουν γούνα, ουρές, μουστάκια και πρόσωπα που μοιάζουν με γάτες. Αντ' αυτού, παράγουν αυτόματα τα αναγνωριστικά χαρακτηριστικά από τα παραδείγματα που επεξεργάζονται. Ένα ANN βασίζεται σε μια συλλογή από συνδεδεμένες μονάδες ή κόμβους που

ονομάζονται τεχνητοί νευρώνες, οι οποίοι χαλαρά μοντελοποιούν τους νευρώνες σε έναν βιολογικό εγκέφαλο. Κάθε σύνδεση, όπως οι συνάψεις σε έναν βιολογικό εγκέφαλο, μπορεί να μεταδώσει ένα σήμα σε άλλους νευρώνες. Ένας τεχνητός νευρώνας που λαμβάνει ένα σήμα στη συνέχεια επεξεργάζεται αυτό και μπορεί να σηματοδοτήσει νευρώνες που συνδέονται με αυτό. Στις υλοποιήσεις του ANN, το "σήμα" σε μια σύνδεση είναι ένας πραγματικός αριθμός και η έξοδος κάθε νευρώνα υπολογίζεται από κάποια μη γραμμική συνάρτηση του αθροίσματος των εισόδων του. Οι συνδέσεις ονομάζονται άκρα. Οι νευρώνες και τα άκρα έχουν συνήθως ένα βάρος που προσαρμόζεται ως έσοδα της μάθησης. Το βάρος αυξάνει ή μειώνει τη δύναμη του σήματος σε μια σύνδεση. Οι νευρώνες μπορεί να έχουν ένα κατώφλι τέτοιο ώστε ένα σήμα να αποστέλλεται μόνο αν το συνολικό σήμα διασχίζει αυτό το όριο. Τυπικά, οι νευρώνες συσσωματώνονται σε στρώματα. Τα διαφορετικά στρώματα μπορούν να εκτελούν διαφορετικούς μετασχηματισμούς στις εισόδους τους. Τα σήματα μετακινούνται από το πρώτο στρώμα (το στρώμα εισόδου) στο τελευταίο στρώμα (το στρώμα εξόδου), πιθανώς μετά από πολλαπλές διαδρομές. Ο αρχικός στόχος της προσέγγισης ANN ήταν να λυθούν τα προβλήματα με τον ίδιο τρόπο που θα είχε και ο ανθρώπινος εγκέφαλος. Αλλά με την πάροδο του χρόνου, η προσοχή μεταφέρθηκε στην εκτέλεση συγκεκριμένων καθηκόντων, οδηγώντας σε αποκλίσεις από τη βιολογία. Τα ANN έχουν χρησιμοποιηθεί σε ποικίλα καθήκοντα, όπως η ηλεκτρονική όραση, η αναγνώριση ομιλίας, η μηχανική μετάφραση, το φιλτράρισμα κοινωνικών δικτύων, τα παιχνίδια και τα βιντεοπαιχνίδια, η ιατρική διάγνωση και ακόμη και σε δραστηριότητες που παραδοσιακά θεωρούνται αποκλειστικά για τον άνθρωπο, όπως ζωγραφική. [13],[14].

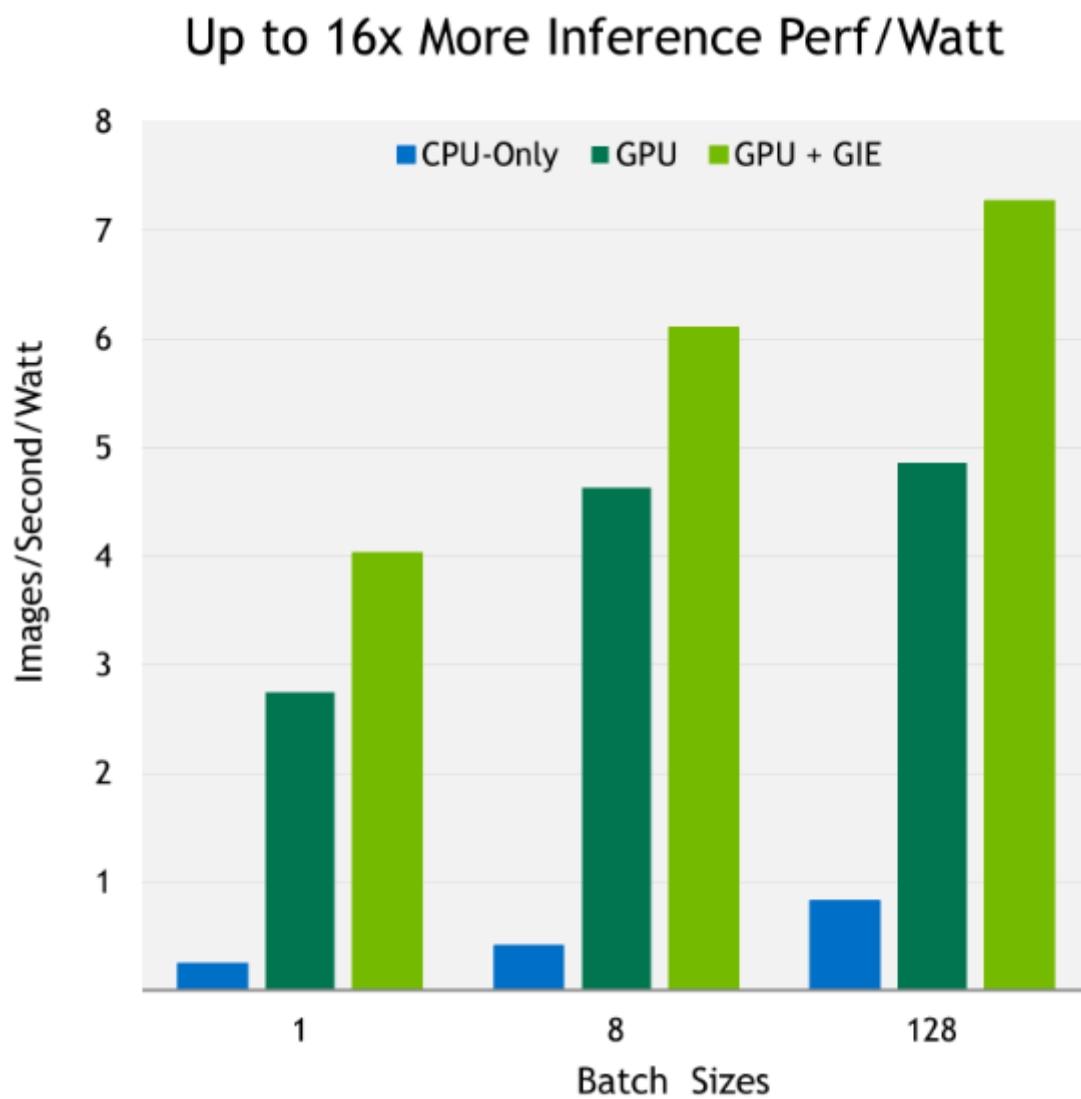


Σχήμα 2.5 Neural Network

2.3.3 Deep Learning & GPU Performance

Η βάση της βαθιάς μάθησης είναι η νόηση ότι οι παρατηρήσεις πηγάζουν από το συνδυασμό και την αλληλεπίδραση διαφορετικών παραγόντων. Περαιτέρω, αυτή η αλληλεπίδραση συμβαίνει και σε πολλαπλά στρώματα πράγμα που συνεπάγεται την συνδυαστική σκέψη επί της παρατηρούμενης τιμής. Πολλαπλά στρώματα σημαίνουν αυξημένος βαθμός αφαιρετικής ικανότητας, καθώς και τέτοιο επίπεδο αφαιρετικής ικανότητας υποκινεί τη δυνατότητα εκμάθησης εννοιών υψηλότερου επιπέδου, μέσω της εκμάθησης εννοιών χαμηλότερου επιπέδου.

Κατά αυτόν τον τρόπο είναι κατανοητό πως όσο μεγαλύτερο το βάθος τόσο μεγαλύτερη η ανάγκη για ενίσχυση των υπολογιστικών πόρων. Η ανάπτυξη ισχυρότερου υλικού και η αύξηση των διαθέσιμων δεδομένων κατέστησαν δυνατή την πραγματοποίηση βαθύτερων νευρωνικών δικτύων. Το 2009, η Nvidia μπήκε στην περιοχή της βαθιάς μάθησης και άρχισε να προωθεί τις επιδόσεις των GPU της εταιρείας. Επιβεβαιώθηκε ότι η συμμετοχή της GPU μπορεί να αυξήσει την ταχύτητα εκπαίδευσης κατά περισσότερο από 100 φορές. Δεδομένου ότι η GPU είναι αρκετά κατάλληλη για υπολογισμό μητρώου/διανύσματος σε αλγόριθμο βαθιάς μάθησης, μια GPU μπορεί να μειώσει τον απαιτούμενο χρόνο από εβδομάδες σε ημέρες. Από την εμφάνιση της βαθιάς μάθησης, έχει γίνει ένα από τα πιο προηγμένα συστήματα σε διάφορους τομείς, ειδικά στον τομέα της υπολογιστικής όρασης και της ταχύτητας αναγνώρισης. Σε τυπικά σύνολα δεδομένων επαλήθευσης τα πειράματα έδειξαν ότι η βαθιά μάθηση μπορεί να βελτιώσει την ακρίβεια αναγνώρισης σε πολύ μεγάλα ποσοστά. [15]



Σχήμα 2.5 GPU over CPU performance

2.4 Deep Learning σε Ιατρικές Εικόνες

Όσον αφορά στην επεξεργασία ιατρικής εικόνας, η βαθιά εκμάθηση γίνεται όλο και πιο ελκυστική. Η πρόσφατη εξέλιξη της βαθιάς μάθησης έχει σημειώσει ένα μεγάλο άλμα. Σε γενικές γραμμές, η έρευνα σχετικά με τη βαθιά μάθηση για ιατρικές εικόνες επικεντρώνεται κυρίως σε τέσσερις πτυχές: ανίχνευση δομών, τμηματοποίηση, επισήμανση και λεζάντες, και ανίχνευση ή διάγνωση με υπολογιστή. Η ανίχνευση της δομής είναι ένα από τα πιο σημαντικά βήματα στη διαδικασία ιατρικής απεικόνισης. Οι παθολόγοι γενικά ολοκληρώνουν αυτό το έργο αναγνωρίζοντας κάποιο ανατομικό χαρακτηριστικό στην εικόνα. Αν και η επιτυχία της βαθιάς μάθησης σε αυτόν τον τομέα εξαρτάται κυρίως από το πόσα ανατομικά χαρακτηριστικά του αλγόριθμου μπορούν να εξάγουν. Η πρόσφατη τάση δείχνει ότι η βαθιά μάθηση είναι αρκετά ώριμη για να λύσει τα προβλήματα του πραγματικού κόσμου.

Η τμηματοποίηση ή segmentation είναι η διαδικασία διαίρεσης μιας ψηφιακής εικόνας σε πολλά δευτερεύοντα σήματα. [16] Ένα τμήμα είναι ένα σύνολο εικονοστοιχείων και επομένως ονομάζεται επίσης υπέρ pixel. Ο στόχος της τμηματοποίησης της εικόνας είναι να απλοποιηθεί ή να μεταβληθεί η αναπαράσταση της εικόνας, έτσι ώστε να γίνει πιο εύκολη η κατανόησή της ή η ανάλυση της. Η τμηματοποίηση χρησιμοποιείται συνήθως για τον εντοπισμό αντικειμένων ή ακμών στην εικόνα. Πιο συγκεκριμένα, η τμηματοποίηση είναι μια διαδικασία για την ετικέτα κάθε εικονοστοιχείου στην εικόνα, πράγμα που κάνει εικονοστοιχεία με την ίδια ετικέτα να έχουν παρόμοιο οπτικό χαρακτηριστικό, όπως χρώμα, φωτεινότητα ή υφή. . Moeskops et al. [17] σχεδίασε CNN πολλαπλών κλιμάκων για ακριβή κατάρτηση ιστών, χρησιμοποιώντας πολλαπλά μεγέθη patch και πολλαπλά μεγέθη πυρήνα συνέλιξης για να αποκτήσει πληροφορίες πολλαπλών κλιμάκων για κάθε εικονοστοιχείο και απέκτησε ακρίβεια από 82% έως 91%.

Η επισήμανση και η λεζάντα (labeling & captioning of images) είναι ο πιο διαδεδομένος τρόπος για να περιγράφουν τα περιεχόμενα μιας εικόνας. Πρόκειται για το κλασικό πρόβλημα κατηγοριοποίησης, στο οποίο προαναφέρθηκα, στον τομέα των ιατρικών εικόνων. Καταβάλλονται συνεχείς προσπάθειες για την εξασφάλιση της επισήμανσης των ασθενειών για συγκεκριμένες ασθένειες. Εμπνευσμένο από τα νευρωνικά δίκτυα για τις κανονικές εικόνες, κάποια έρευνα [18] [19] εισήγαγε το RNN μαζί με την τελευταία πρόοδο στην υπολογιστική όραση με τις ακτινογραφίες θωρακικού υποτίτλου σε συγκεκριμένα πλαίσια. Οι συντάκτες χρησιμοποίησαν λεζάντες εικόνες σε δημόσιο διαθέσιμο σύνολο δεδομένων για την κατάρτιση του CNN. Για να αποφευχθούν μεγάλα σφάλματα, εφαρμόστηκαν πολλές τεχνικές ομαλοποίησης. Στη συνέχεια, το νευρωνικό αυτό δίκτυο χρησιμοποιήθηκε για να περιγράψει την κατάσταση της ανιχνευθίσας ασθένειας. Η ανίχνευση ή η διάγνωση με υποβοήθηση υπολογιστή περιλαμβάνει εύρεση ή εντοπισμό ανωμαλιών και ύποπτων περιοχών και, στη συνέχεια, προειδοποίηση για κλινικούς ιατρούς. Ο κύριος στόχος της ανίχνευσης με ηλεκτρονικό υπολογιστή είναι η αύξηση του ποσοστού ανίχνευσης μολυσμένης περιοχής και η μείωση του ψευδώς αρνητικού λόγω του σφάλματος του παρατηρητή. Παρόλο που θεωρείται μια ώριμη περιοχή στις ιατρικές εικόνες, η βαθιά εκμάθηση βελτίωσε τις επιδόσεις σε πολλές εφαρμογές και επέτρεψε κάποια σχεδίαση που ήταν αδύνατη στο παρελθόν. Παραδοσιακά, η ανίχνευση υπολογιστών απαιτεί μια προεπεξεργασμένη υποψήφια περιοχή και εργατικό δυναμικό για την εξαγωγή χαρακτηριστικών όπως για παράδειγμα το σχήμα ή τα στατιστικά στοιχεία στην περιοχή. Τότε μετά τα χαρακτηριστικά μπορούν να τροφοδοτούν τον classifier. Ωστόσο, το πλεονέκτημα της μάθησης χαρακτηριστικών είναι ο πυρήνας των νέων εξελίξεων. Η βαθιά εκμάθηση μπορεί να μάθει ανεξάρτητα τα ιεραρχικά χαρακτηριστικά από το σύνολο δεδομένων αντί για εξαρτώμενα χειροποίητα χαρακτηριστικά ειδικά στοχοθετημένα για συγκεκριμένο τομέα γνώσης. Σύντομα αποδείχθηκε η πιο εξελιγμένη τεχνολογία. [20], [21]. Εξέτασαν το πρόβλημα υπερφόρτωσης τραβώντας τυχαία ή κουνώντας την αρχική εικόνα και έπειτα τροφοδοτούν τις υπό εικόνες στο μοντέλο. Λόγω της ευημερίας στην έρευνα, όλο και περισσότερες εμπορικές προσπάθειες διεξάγονται πρόσφατα. Οι

αρχάριοι που εισέρχονται στην ιατρική περιοχή ΑΙ αυξάνονται. Από το 2012 έως το 2016, οι επενδύσεις σε ιατρικό κλινικό σύστημα αυξάνονται από 20 περιπτώσεις ετησίως σε 70 περιπτώσεις ετησίως. Περισσότεροι από 100 μεγάλες εταιρείες προσπαθούν να εφαρμόσουν βαθιά μάθηση για να μειώσουν το χρόνο παροχής βοήθειας στον ασθενή και για αυτόματη διάγνωση ασθενειών με ιατρικές εικόνες. Ο IBM Watson Group υποστηρίζει μια έρευνα για την ανίχνευση ασθενών με καρκίνο με μια προσιτή διαδικασία. Προσπαθούν να κάνουν τη βαθιά εκμάθηση κατάλληλη για παραγωγή.

2.4.1 Μηχανική Μάθηση για Διάγνωση του καρκίνου του μαστού

Στην συγκεκριμένη εργασία χρησιμοποιούνται μοντέλα για να προβλέψουν την επανεμφάνιση του καρκίνου του μαστού. Γίνεται σύγκριση τριών μοντέλων.

Data mining αλγόριθμοι:

1. **Decision Trees**
2. **Support Vector Machine**
3. **Artificial Neural Networks**

Σε αυτή την εργασία, χρησιμοποιήθηκε **DT**, **SVM** και **ANN** μηχανική μάθηση αλγορίθμων για την πρόβλεψη της επανεμφάνισης του καρκίνου του μαστού για να βρεθεί ποια μέθοδος εκτελείται καλύτερα.

DT

Κάθε κόμβος δέντρου είναι είτε κόμβος φύλλων είτε κόμβος απόφασης. Όλοι οι κόμβοι αποφάσεων έχουν διαχωρισμούς, ελέγχοντας τις τιμές ορισμένων λειτουργιών των χαρακτηριστικών δεδομένων. Κάθε κλάδος από τον κόμβο απόφασης αντιστοιχεί σε διαφορετικό αποτέλεσμα της δοκιμής. Κάθε κόμβος φύλλων έχει μια ετικέτα κατηγορίας προσαρμοσμένη σε αυτό. Το λογισμικό Weka εφαρμόστηκε για να αναλύσει τα δεδομένα με το C4.5.

SVM

Ο support vector machine (SVM) είναι μια αναδυόμενη ισχυρή μηχανή μάθησης για την κατηγοριοποίηση περιπτώσεων. Το SVM έχει χρησιμοποιηθεί σε μια σειρά από προβλήματα και έχουν ήδη επιτύχει την αναγνώριση των προτύπων στη βιοπληροφορική, στη διάγνωση του καρκίνου και άλλα.

ANN

Το multilayer perceptron έχει τη δυνατότητα να κάνει map ένα σύνολο input δεδομένων σε ένα σύνολο output δεδομένων.

Αποτελέσματα της εν λόγω εργασίας

	DT	SVM	ANN
Accuracy	0.936	0.957	0.947
Sensitivity	0.958	0.971	0.956
Specificity	0.907	0.945	0.928

Υπάρχουν πολλές διαφορετικές τεχνικές data mining που μπορούν να χρησιμοποιηθούν για την πρόβλεψη του καρκίνου του μαστού και για την επανεμφάνισή του.

Στη συγκεκριμένη μελέτη αναλύθηκαν τρεις βασικοί αλγόριθμοι ταξινόμησης για την σύγκριση των αποτελεσμάτων τους.

Τα αποτελέσματα έδειξαν ότι ο Support Vector Machine είχε τις καλύτερες επιδόσεις με δεύτερο το νευρωνικό δίκτυο και τρίτο το decision tree. [22]

3

Μεθοδολογία & Υλοποίηση

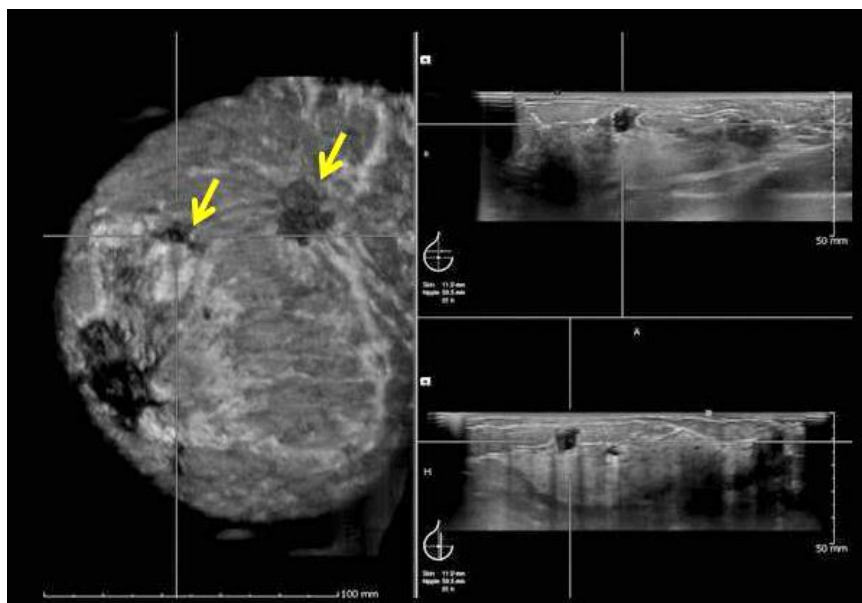
3.1 Προκαταρκτική Μελέτη Δεδομένων

Σε αυτή την ενότητα θα συζητήσω για το ιατρικό υπόβαθρο που μελετήσαμε για αυτό το έργο. Τα θέματα που καλύπτονται περιλαμβάνουν την παθοφυσιολογία και την τρέχουσα μέθοδο διάγνωσης του καρκίνου του μαστού.

Υπέρηχοι Μαστού

Η απεικόνιση με υπερήχους του στήθους χρησιμοποιεί ηχητικά κύματα για την παραγωγή εικόνων των εσωτερικών δομών του μαστού. Χρησιμοποιείται κυρίως για να βοηθήσει στη διάγνωση θρόμβων μαστού ή άλλες ανωμαλίες που μπορεί να έχει εντοπίσει ένας γιατρός κατά τη διάρκεια μιας φυσικής εξέτασης, μαστογραφίας ή μαγνητικής τομογραφίας στήθους. Ο υπέρηχος είναι ασφαλής, μη επεμβατικός και δεν χρησιμοποιεί ακτινοβολία. Ο υπέρηχος παράγει εικόνες από το εσωτερικό του

σώματος χρησιμοποιώντας ηχητικά κύματα. Η απεικόνιση με υπερήχους ονομάζεται επίσης σάρωση υπερήχων ή υπερηχογράφημα. Χρησιμοποιεί ένα μικρό ανιχνευτή που ονομάζεται μετατροπέας και πηγήμα τοποθετημένο απευθείας στο δέρμα. Τα ηχητικά κύματα υψηλής συχνότητας μετακινούνται από τον καθετήρα μέσω του πηκτώματος μέσα στο σώμα. Ο καθετήρας συλλέγει τους ήχους που αναπήδησαν πίσω. Ένας υπολογιστής χρησιμοποιεί αυτά τα ηχητικά κύματα για να δημιουργήσει μια εικόνα. Οι εξετάσεις με υπερήχους δεν χρησιμοποιούν ακτινοβολία (όπως χρησιμοποιείται στις ακτίνες X). Επειδή οι εικόνες συλλαμβάνονται σε πραγματικό χρόνο, μπορούν να δείξουν τη δομή και την κίνηση των εσωτερικών οργάνων του σώματος. Μπορούν επίσης να δείξουν αίμα που ρέει μέσα από αιμοφόρα αγγεία. Η απεικόνιση με υπερήχους είναι μια μη επεμβατική ιατρική εξέταση που βοηθάει τους ιατρούς να διαγνώσουν και να θεραπεύσουν τις ιατρικές παθήσεις. Οι υπέρηχοι Doppler είναι μια ειδική τεχνική υπερήχων που αξιολογεί την κίνηση των υλικών στο σώμα. Επιτρέπει στον γιατρό να βλέπει και να αξιολογεί τη ροή του αίματος μέσω αρτηριών και φλεβών στο σώμα. Η απεικόνιση με υπερήχους του μαστού παράγει μια εικόνα των εσωτερικών δομών του μαστού. Κατά τη διάρκεια μιας υπερηχογραφικής εξέτασης στο μαστό, ο sonographer ή ο γιατρός που εκτελεί τη δοκιμασία μπορεί να χρησιμοποιήσει τεχνικές Doppler για να αξιολογήσει τη ροή του αίματος ή την έλλειψη ροής σε οποιαδήποτε μάζα του μαστού. Σε ορισμένες περιπτώσεις, αυτό μπορεί να παρέχει πρόσθετες πληροφορίες σχετικά με την αιτία της μάζας.



Εικόνα 3.1 Υπέρηχος Μαστού

Παθοφυσιολογία του καρκίνου του Μαστού

Έχω διερευνήσει την παθοφυσιολογική εξήγηση του καρκίνου του μαστού. Αυτό με βοήθησε να κατανοήσω τα χαρακτηριστικά των εικόνων και επίσης με βοήθησε να αναπτύξω μια πιο καλή εικόνα για τις προοπτικές του έργου που υλοποιώ. Ο καρκίνος είναι ανοσολογική αποτυχία του ανοσοποιητικού συστήματος. Το ανοσοποιητικό σύστημα συνήθως αναζητά καρκινικά κύτταρα και κύτταρα με κατεστραμμένο DNA και τα καταστρέφει. Ο καρκίνος του μαστού μπορεί να είναι αποτέλεσμα της αποτυχίας μιας τέτοιας αποτελεσματικής ανοσολογικής άμυνας και παρακολούθησης. Ο καρκίνος περιλαμβάνει στρωματικά κύτταρα και επιθηλιακά κύτταρα. Αυτά είναι διάφορα συστήματα σηματοδότησης αυξητικών παραγόντων και άλλων μεσολαβητών που αλληλεπιδρούν μεταξύ στρωματικών κυττάρων και επιθηλιακών κυττάρων. Η διακοπή αυτών μπορεί επίσης να οδηγήσει σε καρκίνο του μαστού.

3.2 Προετοιμασία δεδομένων

3.2.1 Διαμόρφωση Συνόλου Δεδομένων

Όταν μιλάμε για δεδομένα, συνήθως σκέφτομαι μερικά μεγάλα σύνολα δεδομένων με τεράστιο αριθμό σειρών και στηλών. Ενώ αυτό είναι ένα πιθανό σενάριο, δεν είναι πάντοτε η περίπτωση - τα δεδομένα θα μπορούσαν να είναι σε τόσες πολλές μορφές: Δομημένοι πίνακες, εικόνες, αρχεία ήχου, βίντεο και πολλά άλλα. Οι μηχανές δεν κατανοούν τα ελεύθερα κείμενα, τα δεδομένα εικόνας ή βίντεο όπως είναι, κατανοούν τα 1 και 0. Έτσι πιθανότατα δεν θα είναι αρκετά καλό αν βάλουμε σε ένα slideshow όλες τις εικόνες μας και περιμένουμε το μοντέλο μηχανικής εκμάθησης μας να εκπαιδευτεί μόνο με αυτό. Σε οποιαδήποτε διαδικασία Μηχανικής Μάθησης, η Προεπεξεργασία Δεδομένων είναι το βήμα με το οποίο μετατρέπονται τα δεδομένα, ή αλλιώς κωδικοποιούνται με κατάλληλο τρόπο έτσι ώστε να τα φέρουμε σε μια κατάσταση που τώρα η μηχανή μπορεί εύκολα να τα αναλύσει. Με άλλα λόγια, τα χαρακτηριστικά των δεδομένων μπορούν πλέον να ερμηνευθούν εύκολα από τον αλγόριθμο. Η προεπεξεργασία δεδομένων είναι ένα σημαντικό και απαραίτητο βήμα στη διαδικασία του έργου αυτού. Αυτό είναι ένα θέμα για το οποίο δεν διατίθενται πολλοί εκπαιδευτικοί πόροι στο διαδίκτυο. Χρειάστηκε πολύς χρόνος έρευνας και όμως όλες οι διαθέσιμες πληροφορίες στο διαδίκτυο που βρήκα δεν απαντούν στις πιο ουσιώδεις ερωτήσεις μας που είναι οι εξής: Γιατί πρέπει να το κάνουμε; Τι πραγματικά κάνει; Και, τέλος, πώς μπορούμε να το κάνουμε οι ίδιοι; Ως υπό-πεδίο επεξεργασίας ψηφιακού σήματος, η ψηφιακή επεξεργασία εικόνας έχει πολλά πλεονεκτήματα σε σχέση με την επεξεργασία αναλογικών εικόνων. Επιτρέπει μια ευρύτερη κλίμακα αλγορίθμων που πρέπει να εφαρμοστούν στα δεδομένα εισόδου. Ο σκοπός της ψηφιακής επεξεργασίας εικόνας είναι να βελτιώσει τα δεδομένα εικόνας (χαρακτηριστικά) καταργώντας τις ανεπιθύμητες στρεβλώσεις ή/και την ενίσχυση ορισμένων σημαντικών χαρακτηριστικών εικόνας έτσι ώστε ο υπολογιστής μας να μπορεί να επωφεληθεί από αυτά τα βελτιωμένα δεδομένα για να υλοποιήσει το απαραίτητο δοθέν έργο.[23]

3.2.2 Τύποι Δεδομένων στη Μηχανική Μάθηση

Ένα σύνολο δεδομένων μπορεί να θεωρηθεί ως μια συλλογή αντικειμένων δεδομένων, τα οποία συχνά καλούνται επίσης ως εγγραφές, σημεία, φορείς, μοτίβα, συμβάντα, περιπτώσεις, δείγματα, παρατηρήσεις ή οντότητες. Τα αντικείμενα δεδομένων περιγράφονται από έναν αριθμό χαρακτηριστικών που καταγράφουν τα βασικά χαρακτηριστικά ενός αντικειμένου, όπως η μάζα ενός φυσικού αντικειμένου ή ο χρόνος κατά τον οποίο συνέβη ένα γεγονός. Επί της ουσίας, ένα χαρακτηριστικό είναι μια μεμονωμένη μετρήσιμη ιδιότητα ή χαρακτηριστικό ενός φαινομένου που παρατηρείται.

- Κατηγορηματικά : Χαρακτηριστικά των οποίων οι τιμές λαμβάνονται από ένα καθορισμένο σύνολο τιμών. Για παράδειγμα, ημέρες σε μια εβδομάδα: {Δευτέρα, Τρίτη, Τετάρτη, Πέμπτη, Παρασκευή, Σάββατο, Κυριακή} είναι μια κατηγορία επειδή η αξία της λαμβάνεται πάντοτε από αυτό το σετ. Ένα άλλο παράδειγμα θα μπορούσε να είναι το σύνολο Boolean: {True, False}
- Αριθμητικά : Χαρακτηριστικά των οποίων οι τιμές είναι συνεχείς ή με ακέραια αξία. Αντιπροσωπεύονται από αριθμούς και κατέχουν τις περισσότερες από τις ιδιότητες των αριθμών. Για παράδειγμα, ο αριθμός των βημάτων που περπατάμε μέσα σε μια μέρα ή η ταχύτητα με την οποία οδηγεί κάποιος το αυτοκίνητό του.

Τα βήματα της προετοιμασίας δεδομένων είναι αυτά που ακολουθούνται για να κωδικοποιήσουμε κατάλληλα τα δεδομένα ώστε να μπορούν να αναλυθούν σωστά από τον υπολογιστή. Εννοείται πως δεν ισχύουν όλα τα βήματα για κάθε πρόβλημα και εξαρτάται σε μεγάλο βαθμό από τα δεδομένα με τα οποία εργαζόμαστε, οπότε ίσως χρειαστεί μόνο λίγα βήματα με το σύνολο δεδομένων. Αυτά τα βήματα γενικά είναι : Αξιολόγηση ποιότητας δεδομένων, Συγκέντρωση χαρακτηριστικών, Δειγματοληψία χαρακτηριστικών, Μείωση διατομής, Κωδικοποίηση χαρακτηριστικών.

Αξιολόγηση Ποιότητας Δεδομένων

Αυτό το βήμα γίνεται, λόγω του ότι τα δεδομένα λαμβάνονται συχνά από πολλές πηγές, οι οποίες κατά κανόνα δεν είναι πολύ αξιόπιστες και ότι σε διαφορετικές μορφές, περισσότερο από το ήμισυ του χρόνου μας καταναλώνεται στην αντιμετώπιση προβλημάτων ποιότητας δεδομένων κατά την επεξεργασία ενός προβλήματος μηχανικής μάθησης. Είναι απλώς μη ρεαλιστικό να περιμένουμε ότι τα δεδομένα θα είναι τέλεια. Μπορεί να υπάρχουν προβλήματα λόγω ανθρώπινου σφάλματος, περιορισμών των συσκευών μέτρησης ή ελαττωμάτων στη διαδικασία συλλογής δεδομένων.

Συγκέντρωση Χαρακτηριστικών

Οι συσσωρευμένες λειτουργίες εκτελούνται έτσι ώστε να λαμβάνουν τις συγκεντρωτικές τιμές για να βάλουν τα δεδομένα σε μια καλύτερη προοπτική. Σκεφτείτε τα δεδομένα συναλλαγών, ας υποθέσουμε ότι έχουμε καθημερινές συναλλαγές ενός προϊόντος από την καταγραφή των ημερήσιων πωλήσεων αυτού του προϊόντος σε διάφορες τοποθεσίες καταστημάτων κατά τη διάρκεια του έτους. Η συγκέντρωση των συναλλαγών σε μηνιαίες ή ετήσιες συναλλαγές ενός καταστήματος θα μας βοηθήσει να μειώσουμε τις εκατοντάδες ή ενδεχομένως χιλιάδες συναλλαγές που συμβαίνουν καθημερινά σε ένα συγκεκριμένο κατάστημα, μειώνοντας έτσι τον αριθμό των αντικείμενων δεδομένων. Αυτό έχει ως αποτέλεσμα τη μείωση της κατανάλωσης μνήμης και του χρόνου επεξεργασίας. μας παρέχει μια άποψη υψηλού επιπέδου των δεδομένων καθώς η συμπεριφορά των ομάδων είναι πιο σταθερή από τα μεμονωμένα αντικείμενα δεδομένων.

Δειγματοληψία Χαρακτηριστικών

Η δειγματοληψία είναι μια πολύ κοινή μέθοδος για την επιλογή ενός υποσυνόλου του συνόλου δεδομένων που αναλύουμε. Στις περισσότερες περιπτώσεις, η εργασία με το πλήρες σύνολο δεδομένων μπορεί να αποδειχθεί υπερβολικά δαπανηρή λαμβάνοντας υπόψη τους περιορισμούς μνήμης και χρόνου. Η χρήση ενός αλγόριθμου δειγματοληψίας μπορεί να μας βοηθήσει να μειώσουμε το μέγεθος

του συνόλου δεδομένων σε ένα σημείο όπου μπορούμε να χρησιμοποιήσουμε έναν καλύτερο, αλλά πιο ακριβό αλγόριθμο εκμάθησης μηχανών. Η βασική αρχή εδώ είναι ότι η δειγματοληψία πρέπει να γίνεται με τέτοιο τρόπο ώστε το παραγόμενο δείγμα να έχει περίπου τις ίδιες ιδιότητες με το αρχικό σύνολο δεδομένων, πράγμα που σημαίνει ότι το δείγμα είναι αντιπροσωπευτικό. Αυτό συνεπάγεται την επιλογή του σωστού μεγέθους δείγματος και της στρατηγικής δειγματοληψίας. Η απλή τυχαία δειγματοληψία υπαγορεύει ότι υπάρχει ίση πιθανότητα επιλογής οποιασδήποτε συγκεκριμένης οντότητας.

3.3 Προεπεξεργασία Δεδομένων

3.3.1 Training Validation & Testing Set

Στη διαδικασία προετοιμασίας των δεδομένων μας, προτού δώσουμε το dataset σαν είσοδο στους αλγόριθμους μας, συνιστάται πάντα να διαιρέσουμε το σύνολο δεδομένων σε 2 ή μερικές φορές σε 3 μέρη. Οι αλγόριθμοι μηχανικής μάθησης ή οποιοσδήποτε άλλος αλγόριθμος σχετικά με αυτό το ζήτημα, πρέπει πρώτα να εκπαιδευτεί στην διαθέσιμη κατανομή δεδομένων και στη συνέχεια να επικυρωθεί και να δοκιμαστεί, προτού να μπορέσει να χρησιμοποιηθεί για την αντιμετώπιση πραγματικών δεδομένων.

Training Data

Αυτό είναι το μέρος στο οποίο οι αλγόριθμοι μηχανικής μάθησης εκπαιδεύονται πραγματικά για την κατασκευή ενός μοντέλου το οποίο θα κατανοεί δεδομένα πραγματικού κόσμου. Το μοντέλο προσπαθεί να μάθει το σύνολο δεδομένων και τα διάφορα χαρακτηριστικά του και τις περιπλοκές του, γεγονός που εγείρει και το ζήτημα του overfitting και underfitting.

Validation Data

Αυτό είναι το μέρος του συνόλου δεδομένων που χρησιμοποιείται για την επικύρωση των διαφόρων μοντέλων μας. Με απλά λόγια, χρησιμοποιούμε δεδομένα επικύρωσης για να επιλέξουμε και να βελτιώσουμε τις υπέρ-παραμέτρους των μοντέλων μας. Το μοντέλο δεν μαθαίνει validation set αλλά το χρησιμοποιεί για να φτάσει σε μια καλύτερη κατάσταση υπέρ-παραμέτρων.

Testing Data

Αυτό το τμήμα του συνόλου δεδομένων χρησιμοποιείται για να ελέγξει την υπόθεση του μοντέλου μας. Αφήνεται ανέγγιχτο και αόρατο μέχρις ότου αποφασιστούν το μοντέλο και οι υπέρ-παραμέτρους και μόνο μετά από αυτό το μοντέλο εφαρμόζεται στα δεδομένα των δοκιμών για να πάρει ένα ακριβές μέτρο για το πώς θα εκτελέσει όταν αναπτύσσεται σε δεδομένα πραγματικού κόσμου.

Split Ratio

Η αναλογία κατά την οποία χωρίζουμε τα δεδομένα μας, εξαρτάται σε μεγάλο βαθμό από τον τύπο του μοντέλου που χτίζουμε και από το ίδιο το σύνολο δεδομένων. Εάν το σύνολο δεδομένων και το μοντέλο μας είναι τέτοιο που απαιτείται μεγάλη κατάρτιση, τότε χρησιμοποιούμε ένα μεγαλύτερο κομμάτι των δεδομένων μόνο για εκπαιδευτικούς σκοπούς (στο μεγαλύτερο μέρος των περιπτώσεων). Για παράδειγμα, η κατάρτιση σε δεδομένα κειμένου, δεδομένα εικόνας ή δεδομένα βίντεο συνήθως περιλαμβάνει χιλιάδες χαρακτηριστικά. Εάν το μοντέλο έχει πολλές υπέρ-παραμέτρους που μπορούν να ρυθμιστούν, τότε συνιστάται η διατήρηση υψηλότερου ποσοστού δεδομένων για το validation set. Τα μοντέλα με μικρότερο αριθμό υπέρ-παραμέτρων είναι εύκολο να συντονιστούν και να ενημερωθούν και έτσι μπορούμε να διατηρηθεί ένα μικρότερο validation set. Όπως και πολλά άλλα πράγματα στη μηχανική μάθηση, ο λόγος διαχωρισμού εξαρτάται σε μεγάλο βαθμό από το πρόβλημα που προσπαθούμε να λύσουμε και πρέπει να αποφασίσουμε αφού λάβουμε υπόψη όλες τις διάφορες λεπτομέρειες σχετικά με το μοντέλο και το σύνολο δεδομένων.



Σχήμα 3.1 Dataset Split Ratio

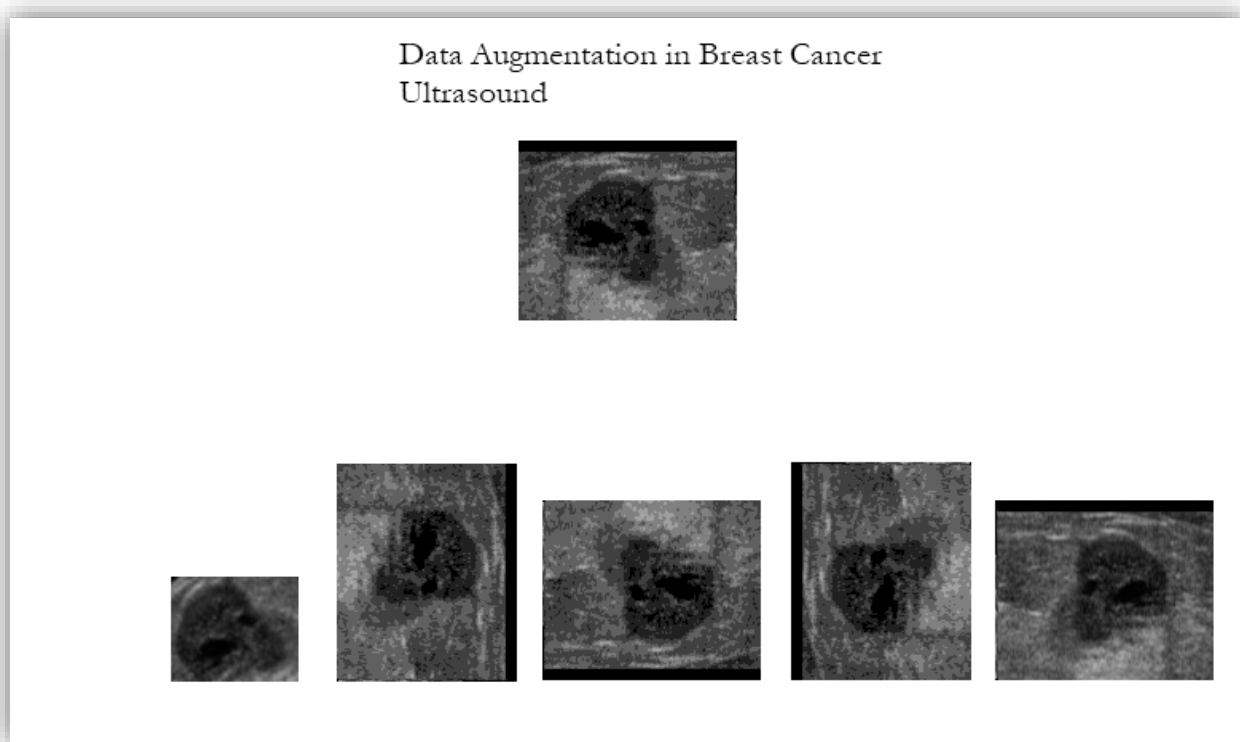
Πιο συγκεκριμένα στο δικό μου έργο χρησιμοποιώ ένα dataset υπερηχογραφημάτων του μαστού από το Kaggle. Στο συγκεκριμένο dataset υπάρχουν εικόνες χωρισμένες σε benign και malignant καταστάσεις, δηλαδή εικόνες με καλοήθεις και κακοήθεις όγκους αντίστοιχα. Οπότε ο στόχος είναι να χωρίσω τα δείγματα μου, βάση όλων όσων ανέφερα από πάνω σε training data, validation data και testing data.

Για να επιτελέσω το εν λόγω έργο, τα δεδομένα μου τα χώρισα σε αναλογία 70% training set, 15% validation set και 15% testing set από ένα σύνολο 250 υπερηχογραφημάτων. Όσον αφορά στην προεπεξεργασία των εικόνων αυτών, αν και οι μέθοδοι συλλογής εικόνων συχνά ελέγχονται αυστηρά για το σύνολο δεδομένων, τα αρχικά δεδομένα εξακολουθούν να έχουν διαφορετικές ιδιότητες όπως φωτεινότητα, αντίθεση και κορεσμό. Η ανάλυση δεδομένων που δεν έχουν κανονικοποιηθεί προσεκτικά μπορεί να παραγάγει παραπλανητικά αποτελέσματα. Έτσι, πρέπει να διασφαλιστεί η αναπαράσταση και η ποιότητα των δεδομένων εισόδου πριν από την εκπαίδευση. Για αυτόν τον λόγο ακολούθησα κάποια βήματα για την προεπεξεργασία των δειγμάτων/εικόνων μου με σκοπό την βελτίωση της εκπαίδευσης και κατά συνέπεια ευστοχίας των αποτελεσμάτων μου.

3.3.2 Image Data Augmentation

Η απόδοση των νευρωνικών δικτύων βαθιάς μάθησης συχνά βελτιώνεται με την ποσότητα των διαθέσιμων δεδομένων. Η αύξηση δεδομένων/data augmentation είναι μια τεχνική για την τεχνητή δημιουργία νέων δεδομένων εκπαίδευσης από τα υπάρχοντα δεδομένα εκπαίδευσης. Αυτό γίνεται με την εφαρμογή τεχνικών συγκεκριμένων τομέων σε παραδείγματα από τα δεδομένα εκπαίδευσης που δημιουργούν νέα και διαφορετικά παραδείγματα εκπαίδευσης. Η αύξηση των δεδομένων εικόνας είναι ίσως ο πιο γνωστός τύπος αύξησης δεδομένων και περιλαμβάνει τη δημιουργία μετασχηματισμένων εκδόσεων εικόνων στο σύνολο δεδομένων κατάρτισης που ανήκουν στην ίδια κατηγορία με την αρχική εικόνα. Οι μετασχηματισμοί περιλαμβάνουν μια σειρά λειτουργιών από το πεδίο χειρισμού εικόνας, όπως μετατοπίσεις, ανατροπές, ζουμ και πολλά άλλα. Η πρόθεση είναι να επεκταθεί το σύνολο δεδομένων κατάρτισης με νέα, εύλογα παραδείγματα. Αυτό σημαίνει, παραλλαγές των εικόνων εκπαιδευτικής σειράς που είναι πιθανό να παρατηρηθούν από το μοντέλο. Για παράδειγμα, μια οριζόντια αναστροφή μιας εικόνας μιας γάτας μπορεί να έχει νόημα, επειδή η φωτογραφία θα μπορούσε να έχει ληφθεί από αριστερά ή δεξιά. Μια κάθετη αναστροφή της φωτογραφίας μιας γάτας δεν έχει νόημα και πιθανόν να μην είναι κατάλληλη δεδομένου ότι το μοντέλο είναι πολύ απίθανο να δει μια φωτογραφία μιας ανάποδης γάτας. Ως εκ τούτου, είναι σαφές ότι η επιλογή των τεχνικών αύξησης των ειδικών δεδομένων που χρησιμοποιούνται για ένα σύνολο δεδομένων κατάρτισης πρέπει να επιλεγεί προσεκτικά και στο πλαίσιο του συνόλου δεδομένων κατάρτισης και της γνώσης του τομέα προβλημάτων. Ουσιαστικά δηλαδή η τεχνική αύξησης δειγμάτων προσπαθεί να ωθήσει έναν αλγόριθμο βαθιάς μάθησης να αντιληφθεί το δεδομένο με περισσότερους από έναν τρόπους, έτσι ώστε να αντλήσει ακόμα περισσότερη πληροφορία ίσως και κατά πολύ μεγαλύτερη σε σχέση με την προηγούμενη του κατάσταση. Αυτή η τεχνική data augmentation βοηθάει κυρίως σε περιπτώσεις όπου η συλλογή περισσότερων δεδομένων δεν είναι πάντα εφικτή. Οι σύγχρονοι αλγόριθμοι βαθιάς μάθησης, όπως το συνελικτικό νευρωνικό δίκτυο ή CNN, μπορούν να μάθουν χαρακτηριστικά που

είναι αμετάβλητα στη θέση τους στην εικόνα. Εντούτοις, η αύξηση μπορεί να βοηθήσει περαιτέρω σε αυτή τη μεταβλητή αμετάβλητη προσέγγιση της μάθησης και μπορεί να βοηθήσει το μοντέλο στην εκμάθηση χαρακτηριστικών που είναι επίσης αμετάβλητα για μετασχηματισμούς όπως από αριστερά προς δεξιά στην κορυφή προς τα κάτω παραγγελία, επίπεδα φωτός σε φωτογραφίες και πολλά άλλα. Η αύξηση των δεδομένων εικόνας/image data augmentation συνήθως εφαρμόζεται μόνο στο σύνολο δεδομένων κατάρτισης/training set και όχι στο σύνολο δεδομένων επικύρωσης/validation set ή δοκιμής/testing set. Αυτό είναι διαφορετικό από την προετοιμασία δεδομένων, όπως αλλαγή μεγέθους εικόνας και κλιμάκωση εικονοστοιχείων. πρέπει να εκτελούνται με συνέπεια σε όλα τα σύνολα δεδομένων που αλληλεπιδρούν με το μοντέλο.



Εικόνα 3.2 Data Augmentation in Breast Cancer Ultrasound

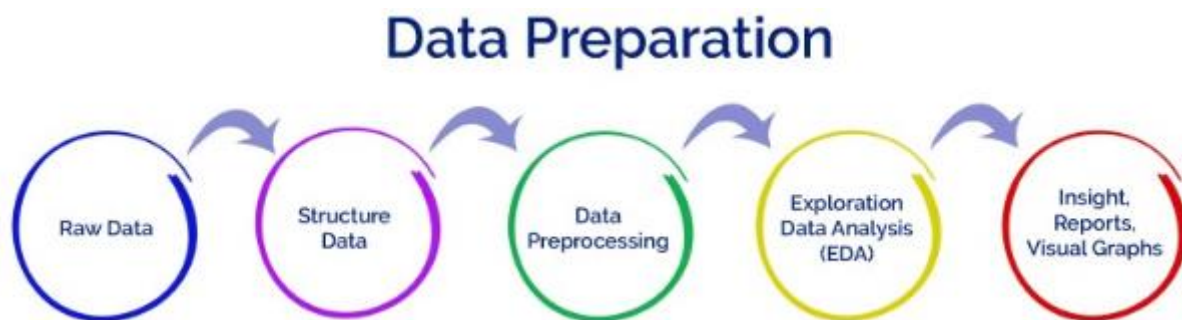
3.3.3 Image Resize

Η τεχνική image resize είναι μια τεχνική κατά την οποία αλλάζουμε τις διαστάσεις των δεδομένων/εικόνων μας, με τρόπο ώστε να εξυπηρετούν το συγκεκριμένο σκοπό του έργου που καλούμαστε να υλοποιήσουμε. Γιατί όμως ακριβώς αλλάζουμε το μέγεθος της εικόνας μας κατά τη φάση προεπεξεργασίας; Ορισμένες εικόνες που έχουμε στο δείγμα μας και τροφοδοτούνται στον αλγόριθμο βαθιάς μάθησης που χρησιμοποιώ ποικίλλουν σε μέγεθος, επομένως, θα πρέπει να δημιουργήσουμε ένα μέγεθος βάσης για όλες τις εικόνες που τροφοδοτούνται στους αλγορίθμους. Για αυτόν τον λόγο κυρίως χρειαζόμαστε αυτό το βήμα επεξεργασίας.

3.3.4 Image Normalization

Οι εικόνες αποτελούνται από μήτρες τιμών εικονοστοιχείων. Οι ασπρόμαυρες εικόνες είναι μονές μήτρες των εικονοστοιχείων, ενώ οι έγχρωμες εικόνες έχουν ξεχωριστή σειρά τιμών εικονοστοιχείων για κάθε κανάλι χρώματος, όπως το κόκκινο, το πράσινο και το μπλε. Οι τιμές pixel είναι συχνά μη προσημασμένοι ακέραιοι αριθμοί μεταξύ 0 και 255. Αν και αυτές οι τιμές εικονοστοιχείων μπορούν να παρουσιαστούν απευθείας σε μοντέλα νευρωνικών δικτύων στην ακατέργαστη μορφή τους, αυτό μπορεί να οδηγήσει σε προκλήσεις κατά τη μοντελοποίηση, όπως στην πιο αργή από την αναμενόμενη εκπαίδευση του μοντέλου. Αντί αυτού, μπορεί να υπάρξει μεγάλο πλεονέκτημα στην προετοιμασία των τιμών εικονοστοιχείων εικόνας πριν από τη μοντελοποίηση, όπως απλά κλιμάκωση τιμών εικονοστοιχείων στην περιοχή 0-1 έως κεντραρίσματος και ακόμη και τυποποίηση των τιμών.

Εφάρμοσα λοιπόν αυτές τις τεχνικές image preprocessing στο σύνολο δεδομένων μου, καθώς ήθελα να επιτύχω επιθυμητά αποτελέσματα και τα νευρωνικά μου δίκτυα να μπορούν με μεγάλη ευστοχία να καταλαβαίνουν σε ποια κατηγορία ανήκουν οι εικόνες που παίρνουν ως είσοδο.



Εικόνα 3.3 Data Preparation

3.4 Αρχιτεκτονική των Νευρωνικών Δικτύων

Η αρχιτεκτονική των νευρωνικών δικτύων είναι ένας πολύ σημαντικός παράγοντας στην πολυπλοκότητα της αντίληψης του δικτύου νευρώνων. Υπάρχουν πολλών ειδών αρχιτεκτονικές και κάθε μία από αυτές έχει διαφορετική δυναμική και κατά συνέπεια προσφέρει διαφορετική ευστοχία σε στην αναγνώριση του δοθέντος συνόλου δεδομένων. Τα τεχνητά νευρωνικά δίκτυα είναι ένα από τα κύρια εργαλεία που χρησιμοποιούνται στη μηχανική μάθηση. Όπως υποδηλώνει το «νευρωνικά» μέρος του ονόματός τους, είναι συστήματα εμπνευσμένα από εγκεφάλους, τα οποία προορίζονται να αναπαράγουν τον τρόπο με τον οποίο μαθαίνουν οι άνθρωποι. Τα νευρωνικά δίκτυα αποτελούνται από στρώματα εισόδου και εξόδου, καθώς και (στις περισσότερες περιπτώσεις) από ένα κρυφό στρώμα που αποτελείται από μονάδες που μετατρέπουν την είσοδο σε κάτι που μπορεί να χρησιμοποιήσει το στρώμα εξόδου. Είναι εξαιρετικά εργαλεία για την εύρεση σχεδίων που είναι πάρα πολύ περίπλοκα ή μεγάλου όγκου για έναν ανθρώπινο προγραμματιστή να εξαγάγει και να διδάξει τη μηχανή να αναγνωρίσει. [24]

Στο εν λόγω έργο, χρησιμοποιώ διαφόρων ειδών αρχιτεκτονικές νευρωνικών δικτύων, καθώς και μια αυτοσχέδια με λίγα layers (μοντέλο τύπου sequential) και συγκρίνω τα αποτελέσματά τους

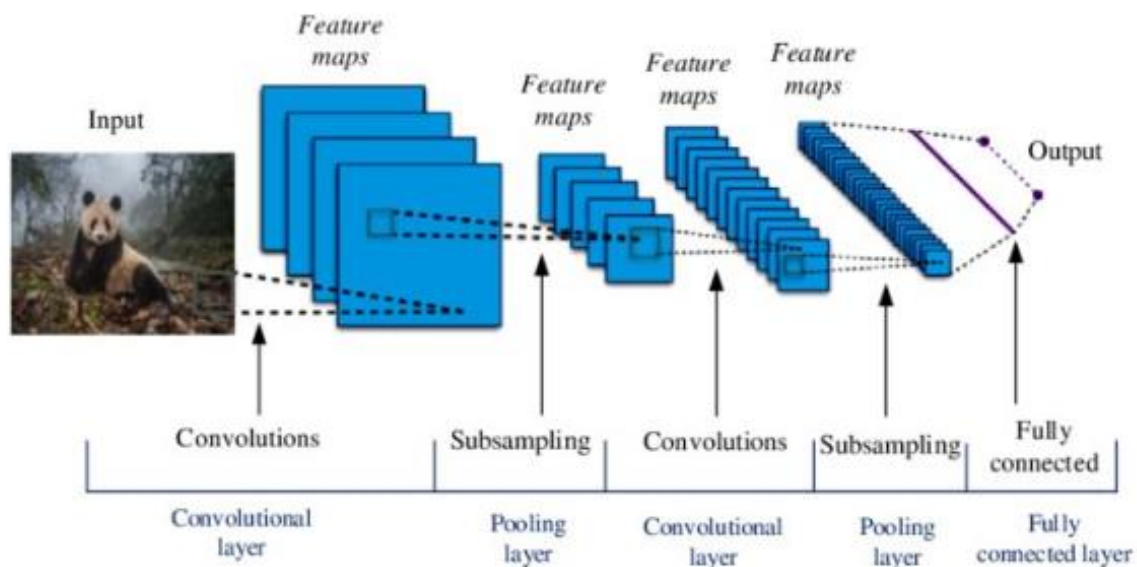
(με το ίδιο image preprocessing σε κάθε περίπτωση) και οι διαφορές είναι εμφανείς βάση δυναμικής μεγέθους και, προφανώς, πολυπλοκότητας. Θα εξηγήσω, λοιπόν, αναλυτικά την κάθε αρχιτεκτονική που χρησιμοποιώ από την πιο απλή μέχρι και την πιο αποδοτική.

3.4.1 Convolutional Neural Nets

Τα Convolutional Neural Networks (CNN) είναι παντού. Είναι αναμφισβήτητα ο πιο δημοφιλής τύπος νευρωνικών δικτύων βαθιάς μάθησης. Η πρόσφατη αύξηση του ενδιαφέροντος για τη βαθιά μάθηση οφείλεται στην τεράστια δημοτικότητα και αποτελεσματικότητα των συνεδρίων. Το ενδιαφέρον για τα CNN ξεκίνησε με το AlexNet το 2012 και έκτοτε έχει αυξηθεί εκθετικά. Σε μόλις τρία χρόνια, οι ερευνητές προχώρησαν από 8 στρώματα AlexNet σε 152 επίπεδα ResNet. Τα CNN είναι τώρα τα μοντέλα για κάθε πρόβλημα που σχετίζεται με την εικόνα. Όσον αφορά την ακρίβεια, ξεπερνούν κάθε άλλο τύπο αρχιτεκτονικής νευρωνικών δικτύων σε ανταγωνισμό. Εφαρμόζονται επίσης με επιτυχία σε συστήματα σύστασης, επεξεργασία φυσικής γλώσσας και άλλα. Το κύριο πλεονέκτημα των CNN σε σύγκριση με τους προκατόχους του είναι ότι ανιχνεύει αυτόματα τα σημαντικά χαρακτηριστικά χωρίς καμία ανθρώπινη επίβλεψη. Για παράδειγμα, με δεδομένα πολλές εικόνες και, μαθαίνει ξεχωριστά χαρακτηριστικά για κάθε κλάση από μόνο του. Τα CNN είναι επίσης υπολογιστικά αποτελεσματικά. Χρησιμοποιούν ειδικές λειτουργίες συνέλιξης και ομαδοποίησης και εκτελούν κοινή χρήση παραμέτρων. Αυτό επιτρέπει στα μοντέλα CNN να λειτουργούν σε οποιαδήποτε συσκευή, καθιστώντας τα παγκοσμίως ελκυστικά. Αντιμετωπίζουμε ένα πολύ ισχυρό και αποδοτικό μοντέλο που εκτελεί αυτόματη εξαγωγή χαρακτηριστικών για την επίτευξη υπεράνθρωπης ακρίβειας.

3.4.2 CNN Layers

Όλα τα μοντέλα CNN ακολουθούν παρόμοια αρχιτεκτονική, όπως φαίνεται στο παρακάτω σχήμα.



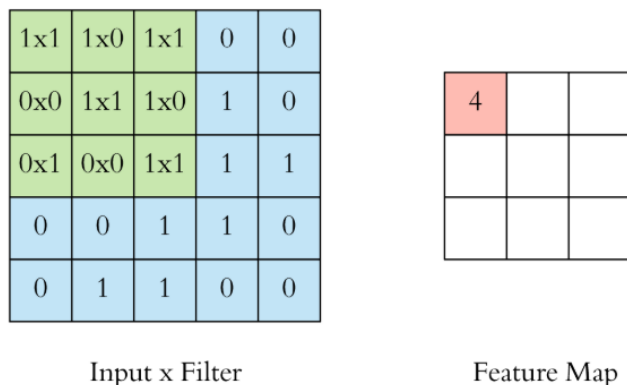
Εικόνα 3.4 CNN Βασική Αρχιτεκτονική

Υπάρχει μια εικόνα εισαγωγής με την οποία εργαζόμαστε. Πραγματοποιούμε μια σειρά συνελίξεων(convolution) και pooling, ακολουθούμενες από έναν αριθμό fully connected επιπέδων. Εάν πραγματοποιούμε ταξινόμηση πολλαπλών κλάσεων, η έξοδος είναι SoftMax. Τώρα θα βρούμε κάθε συστατικό.

- **Convolution**

Το κύριο δομικό στοιχείο του CNN είναι το convolution επίπεδο. Το Convolution είναι μια μαθηματική πράξη για τη συγχώνευση δύο συνόλων πληροφοριών. Στην περίπτωσή μας, η συνέλιξη εφαρμόζεται στα δεδομένα εισόδου χρησιμοποιώντας ένα φίλτρο συνελύσεων για την παραγωγή ενός χάρτη χαρακτηριστικών. Υπάρχουν πολλοί όροι, οπότε ας τους οπτικοποιήσουμε έναν προς έναν. Στην αριστερή πλευρά βρίσκεται η είσοδος στο επίπεδο της συνέλιξης, για παράδειγμα η εικόνα εισόδου. Στα δεξιά βρίσκεται το φίλτρο συνέλιξης, που ονομάζεται επίσης πυρήνας. Αυτό ονομάζεται συνέλιξη 3x3 λόγω του σχήματος του φίλτρου. Πραγματοποιούμε τη διαδικασία συνέλιξης σύροντας αυτό το

φίλτρο πάνω από την είσοδο. Σε κάθε τοποθεσία, κάνουμε πολλαπλασιασμό μήτρας και αθροίζουμε το αποτέλεσμα. Αυτό το άθροισμα πηγαίνει στο χάρτη χαρακτηριστικών. Η πράσινη περιοχή όπου λαμβάνει χώρα η διαδικασία συνέλιξης ονομάζεται receptive field. Λόγω του μεγέθους του φίλτρου, το receptive field είναι επίσης 3x3.



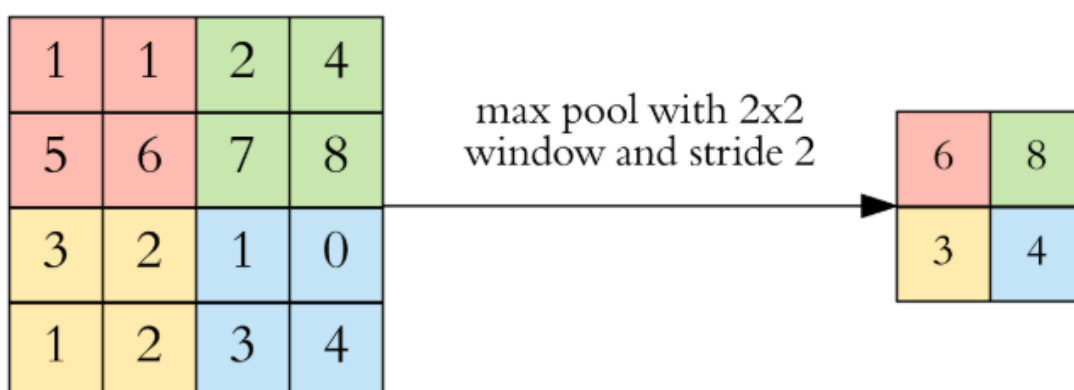
Εικόνα 3.5 Convolutional Layer

Εδώ το φίλτρο βρίσκεται επάνω αριστερά, η έξοδος της λειτουργίας συνέλιξης "4" εμφανίζεται στον χάρτη χαρακτηριστικών που προκύπτει. Στη συνέχεια το φίλτρο μετακινείται ένα βήμα δεξιά και εκτελείται η ίδια λειτουργία, προσθέτοντας αυτό το αποτέλεσμα και στο feature map. Στην πραγματικότητα αυτές οι συνέλιξεις εκτελούνται σε 3D. Στην πραγματικότητα, μια εικόνα αντιπροσωπεύεται ως 3D μήτρα με διαστάσεις ύψους, πλάτους και βάθους, όπου το βάθος αντιστοιχεί σε κανάλια χρώματος (RGB). Ένα convolutional layer έχει συγκεκριμένο ύψος και πλάτος, όπως 3x3 ή 5x5, και από τη σχεδίαση καλύπτει ολόκληρο το βάθος της εισόδου του, οπότε πρέπει να είναι και 3D.

- Pooling

Μετά από μια διαδικασία συνέλιξης συνήθως γίνεται το pooling για να μειώσουμε τη διάσταση. Αυτό μας επιτρέπει να μειώσουμε τον αριθμό των παραμέτρων, οι οποίες συντομεύουν τόσο τον χρόνο του training όσο και προλαμβάνουν για το overfitting. Γίνεται ουσιαστικά ομαδοποίηση επιπέδων

δειγματοληψίας κάθε χάρτη χαρακτηριστικών ανεξάρτητα, μειώνοντας το ύψος και το πλάτος, διατηρώντας το βάθος ανέπαφο. Ο πιο συνηθισμένος τύπος pooling layer είναι το max pooling που παίρνει μόνο τη μέγιστη τιμή στο παράθυρο συγκέντρωσης. Σε αντίθεση με τη λειτουργία της συνέλιξης, η συγκέντρωση δεν έχει παραμέτρους. Μετακινείται ένα παράθυρο πάνω από την είσοδό του και παίρνει απλά τη μέγιστη τιμή στο παράθυρο. Παρόμοια με μια συνέλιξη, καθορίζουμε το μέγεθος και το βήμα παραθύρου.

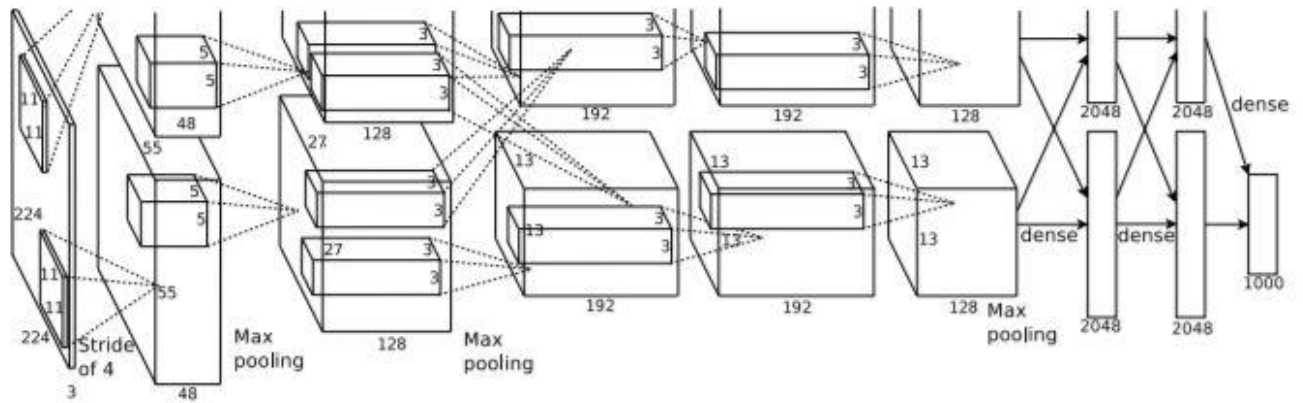


Εικόνα 3.6 Max Pooling Layer

- **Fully Connected Layers**

Μετά τα στρώματα συνέλιξης/convolution και συγκέντρωσης/pooling προσθέτουμε μερικά πλήρως συνδεδεμένα επίπεδα για να ολοκληρώσουμε την αρχιτεκτονική του CNN. Η έξοδος τόσο των συνέλιξεων όσο και των συγκεντρωτικών επιπέδων είναι τρισδιάστατοι χώροι, αλλά ένα πλήρως συνδεδεμένο επίπεδο αναμένει ένα 1D διάνυσμα αριθμών. Έτσι ισοπεδώνουμε την έξοδο του τελικού στρώματος συγκέντρωσης σε ένα διάνυσμα και αυτό το διάνυσμα γίνεται η είσοδος στο πλήρως συνδεδεμένο στρώμα. Η ισοπέδωση αυτή, ουσιαστικά απλώς τακτοποιεί τον τρισδιάστατο όγκο αριθμών σε ένα διάνυσμα 1D.

3.4.3 AlexNet Αρχιτεκτονική



Εικόνα 3.7 AlexNet Αρχιτεκτονική

Η αρχιτεκτονική AlexNet αποτελείται από πέντε συνελικτικά στρώματα, μερικά από τα οποία ακολουθούνται από μέγιστα συγκεντρωτικά στρώματα και έπειτα από τρία πλήρως συνδεδεμένα στρώματα και τελικά από έναν SoftMax classifier.

AlexNet Layers

Πρώτο επίπεδο:

Η είσοδος για το AlexNet είναι μια εικόνα RGB 227x227x3 η οποία περνάει μέσα από το πρώτο στρώμα περιελίξεων με 96 χάρτες χαρακτηριστικών ή φίλτρα μεγέθους 11×11 και ένα βήμα 4. Οι διαστάσεις της εικόνας αλλάζουν σε 55x55x96. Στη συνέχεια, το AlexNet εφαρμόζει max pooling με μέγεθος φίλτρου 3×3 και ένα βήμα δύο. Οι προκύπτουσες διαστάσεις της εικόνας θα μειωθούν σε 27x27x96.

Δεύτερο επίπεδο:

Στη συνέχεια, υπάρχει ένα δεύτερο στρώμα περιστροφής με 256 χάρτες χαρακτηριστικών που έχουν μέγεθος 5×5 και ένα βήμα 1. Στη συνέχεια, υπάρχει και πάλι ένα μέγιστο στρώμα

συγκέντρωσης με μέγεθος φίλτρου 3×3 και ένα βήμα 2. Αυτό το στρώμα είναι το ίδιο με το δεύτερο στρώμα εκτός από τους 256 χάρτες χαρακτηριστικών έτσι ώστε η απόδοση να μειωθεί στα $13 \times 13 \times 256$.

Τρίτο, τέταρτο και πέμπτο επίπεδο:

Το τρίτο, το τέταρτο και το πέμπτο στρώμα είναι συνελικτικά στρώματα με μέγεθος φίλτρου 3×3 και ένα βήμα ενός. Οι πρώτοι δύο χρησιμοποιούσαν 384 χάρτες χαρακτηριστικών όπου το τρίτο χρησιμοποίησε 256 φίλτρα. Τα τρία στρώματα συνέλιξης ακολουθούνται από ένα μέγιστο στρώμα συγκέντρωσης με μέγεθος φίλτρου 3×3 , ένα βήμα 2 και έχουν 256 χάρτες χαρακτηριστικών.

Έκτο Επίπεδο:

Η έξοδος του συνελικτικού στρώματος ισοπεδώνεται μέσω ενός πλήρως συνδεδεμένου στρώματος με 9216 χάρτες χαρακτηριστικών κάθε μεγέθους 1×1 .

Έβδομο και Όγδοο Επίπεδο:

Ακολουθούν δύο πλήρως συνδεδεμένα επίπεδα με 4096 μονάδες.

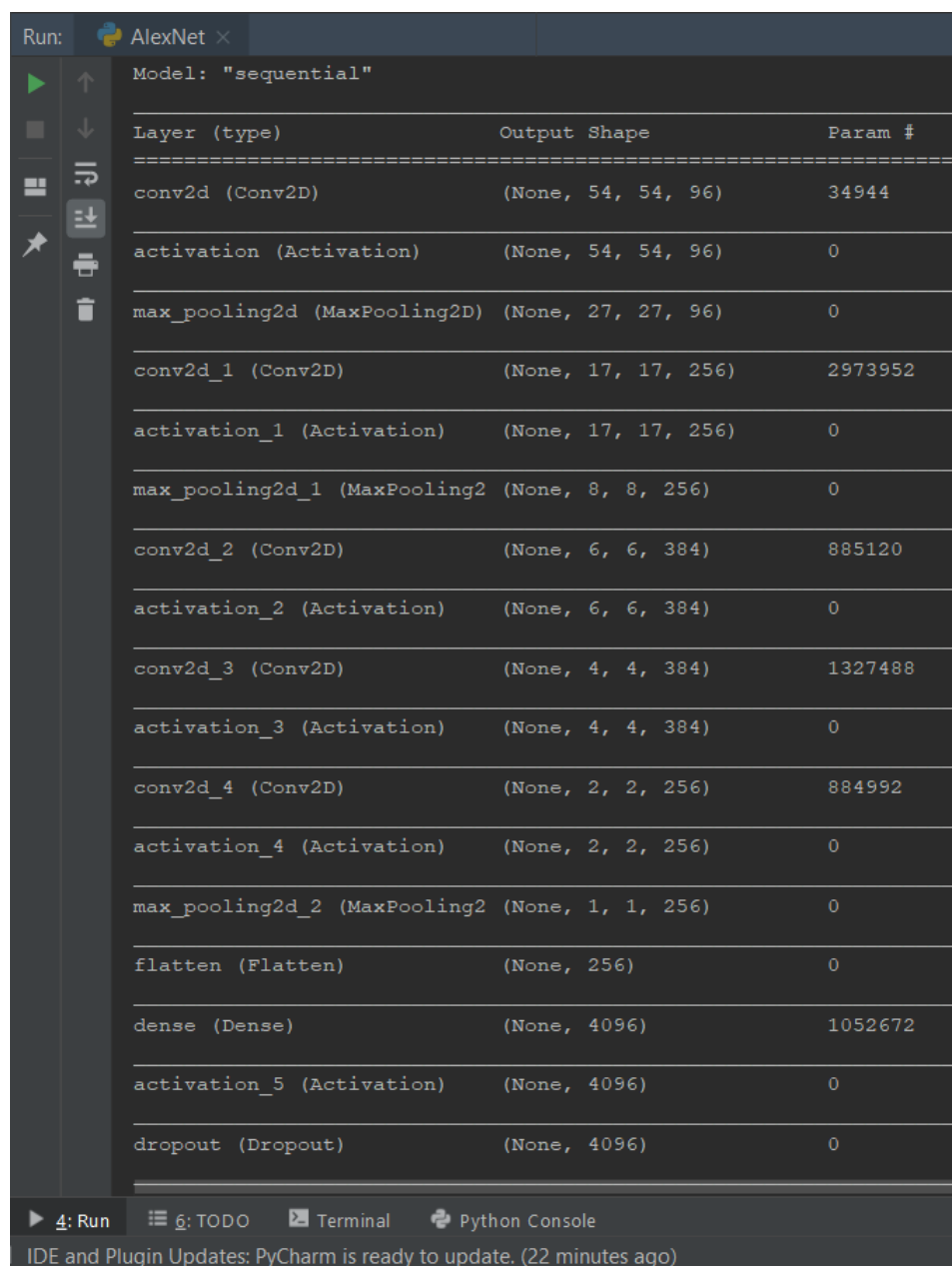
Επίπεδο εξόδου:

Τέλος, υπάρχει ένα επίπεδο εξόδου softmax \hat{y} με 2 πιθανές τιμές.

Layer		Feature Map	Size	Kernel Size	Stride	Activation
Input	Image	1	227x227x3	-	-	-
1	Convolution	96	55 x 55 x 96	11x11	4	relu
	Max Pooling	96	27 x 27 x 96	3x3	2	relu
2	Convolution	256	27 x 27 x 256	5x5	1	relu
	Max Pooling	256	13 x 13 x 256	3x3	2	relu
3	Convolution	384	13 x 13 x 384	3x3	1	relu
4	Convolution	384	13 x 13 x 384	3x3	1	relu
5	Convolution	256	13 x 13 x 256	3x3	1	relu
	Max Pooling	256	6 x 6 x 256	3x3	2	relu
6	FC	-	9216	-	-	relu
7	FC	-	4096	-	-	relu
8	FC	-	4096	-	-	relu
Output	FC	-	1000	-	-	Softmax

Πίνακας 3.1 AlexNet

Σε αυτό το σημείο φαίνεται το model summary του AlexNet CNN όπως εμφανίζεται στον IDE που χρησιμοποιώ PyCharm.



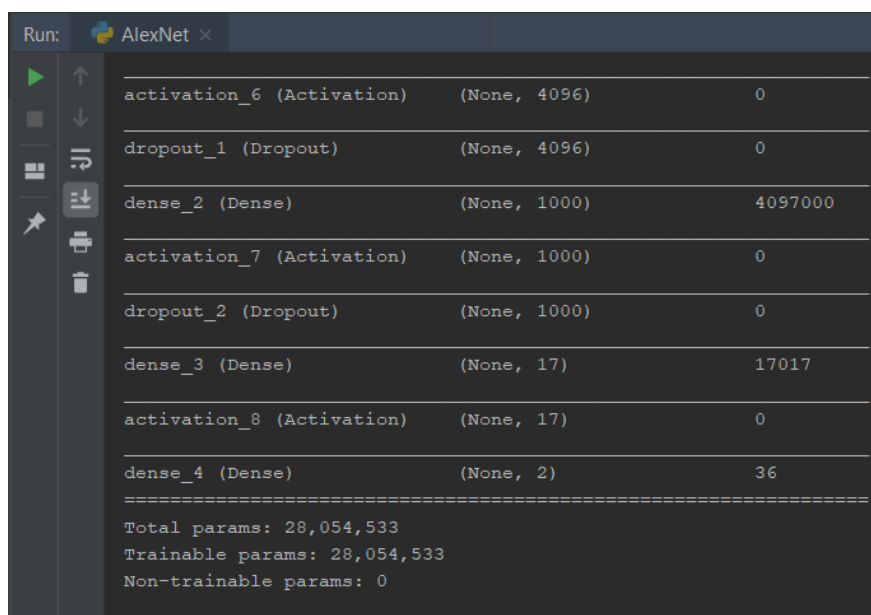
The screenshot shows the PyCharm Run console with the AlexNet model summary. The console output is as follows:

```
Run: AlexNet x
Model: "sequential"

Layer (type)                Output Shape                Param #
=====
conv2d (Conv2D)              (None, 54, 54, 96)          34944
activation (Activation)       (None, 54, 54, 96)          0
max_pooling2d (MaxPooling2D) (None, 27, 27, 96)          0
conv2d_1 (Conv2D)             (None, 17, 17, 256)         2973952
activation_1 (Activation)     (None, 17, 17, 256)         0
max_pooling2d_1 (MaxPooling2 (None, 8, 8, 256)          0
conv2d_2 (Conv2D)             (None, 6, 6, 384)           885120
activation_2 (Activation)     (None, 6, 6, 384)           0
conv2d_3 (Conv2D)             (None, 4, 4, 384)           1327488
activation_3 (Activation)     (None, 4, 4, 384)           0
conv2d_4 (Conv2D)             (None, 2, 2, 256)           884992
activation_4 (Activation)     (None, 2, 2, 256)           0
max_pooling2d_2 (MaxPooling2 (None, 1, 1, 256)          0
flatten (Flatten)             (None, 256)                  0
dense (Dense)                 (None, 4096)                 1052672
activation_5 (Activation)     (None, 4096)                 0
dropout (Dropout)             (None, 4096)                 0
```

At the bottom of the console, there are tabs for '4: Run', '6: TODO', 'Terminal', and 'Python Console'. A status bar at the very bottom indicates 'IDE and Plugin Updates: PyCharm is ready to update. (22 minutes ago)'.

Εικόνα 3.8 AlexNet Summary part1



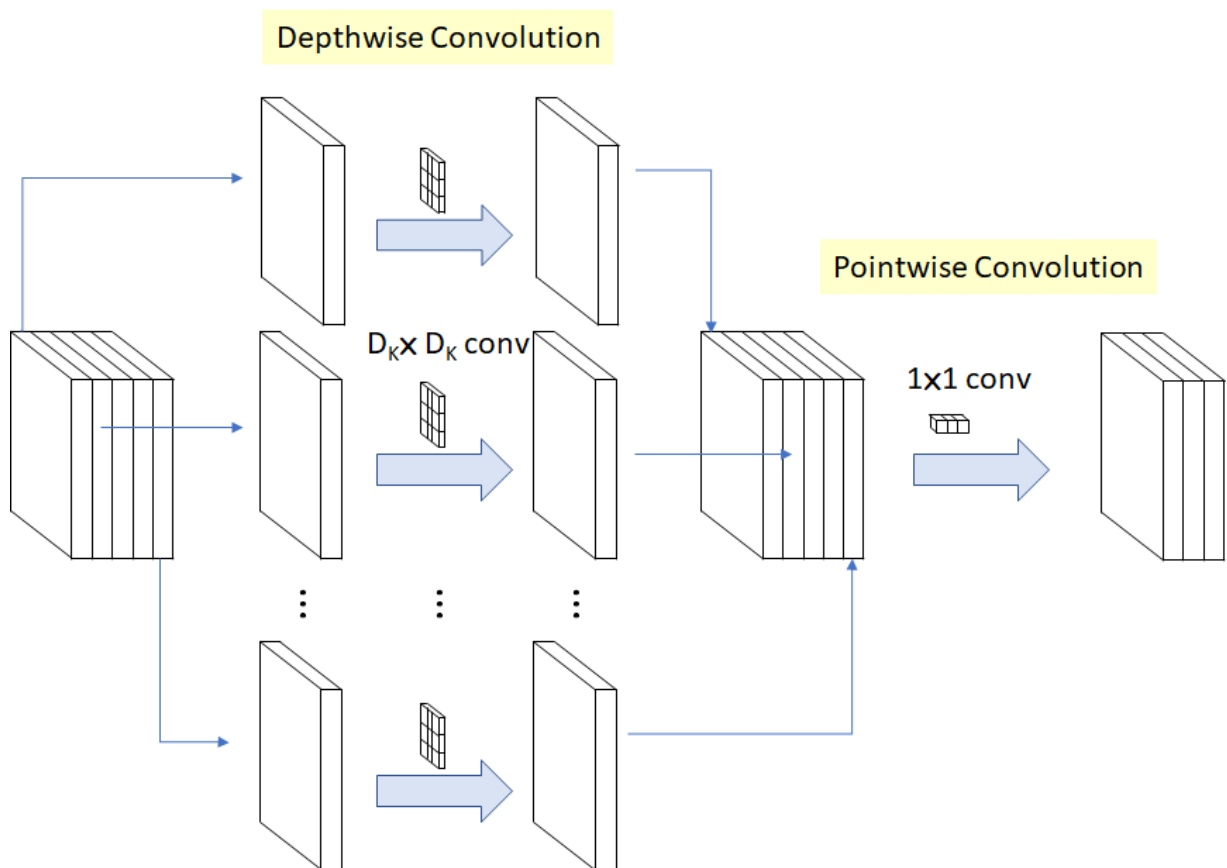
Layer	Output Shape	Param Count
activation_6 (Activation)	(None, 4096)	0
dropout_1 (Dropout)	(None, 4096)	0
dense_2 (Dense)	(None, 1000)	4097000
activation_7 (Activation)	(None, 1000)	0
dropout_2 (Dropout)	(None, 1000)	0
dense_3 (Dense)	(None, 17)	17017
activation_8 (Activation)	(None, 17)	0
dense_4 (Dense)	(None, 2)	36

=====
Total params: 28,054,533
Trainable params: 28,054,533
Non-trainable params: 0

Εικόνα 3.9 AlexNet Summary part2

Έδωσα ως input το us-dataset των υπερηχογραφήμάτων το οποίο με το preprocessing που του έκανα όπως προαναφέρθηκα στις προηγούμενες ενότητες του κεφαλαίου έδωσε κάποια αρκετά καλά αποτελέσματα δεδομένου της απλότητας από πλευράς αρχιτεκτονικής και μεγέθους (περίπου 60% accuracy στο testing).

3.4.4 MobileNet Αρχιτεκτονική (Lightweight Model)



Εικόνα 3.10 MobileNet Architecture [28]

Το MobileNet είναι μια αρχιτεκτονική που είναι πιο κατάλληλη για φορητές και ενσωματωμένες εφαρμογές υπολογιστικής όρασης όπου υπάρχει έλλειψη υπολογιστικής ισχύος. Αυτή η αρχιτεκτονική προτάθηκε από την Google. Το δίκτυο αυτό είναι ένα αρκετά ελαφρύ δίκτυο νευρώνων το οποίο όμως λόγω της έξυπνης αρχιτεκτονικής του υπάρχει πολύ καλή απόδοση.[28]

Table 1. MobileNet Body Architecture

Type / Stride	Filter Shape	Input Size
Conv / s2	$3 \times 3 \times 3 \times 32$	$224 \times 224 \times 3$
Conv dw / s1	$3 \times 3 \times 32$ dw	$112 \times 112 \times 32$
Conv / s1	$1 \times 1 \times 32 \times 64$	$112 \times 112 \times 32$
Conv dw / s2	$3 \times 3 \times 64$ dw	$112 \times 112 \times 64$
Conv / s1	$1 \times 1 \times 64 \times 128$	$56 \times 56 \times 64$
Conv dw / s1	$3 \times 3 \times 128$ dw	$56 \times 56 \times 128$
Conv / s1	$1 \times 1 \times 128 \times 128$	$56 \times 56 \times 128$
Conv dw / s2	$3 \times 3 \times 128$ dw	$56 \times 56 \times 128$
Conv / s1	$1 \times 1 \times 128 \times 256$	$28 \times 28 \times 128$
Conv dw / s1	$3 \times 3 \times 256$ dw	$28 \times 28 \times 256$
Conv / s1	$1 \times 1 \times 256 \times 256$	$28 \times 28 \times 256$
Conv dw / s2	$3 \times 3 \times 256$ dw	$28 \times 28 \times 256$
Conv / s1	$1 \times 1 \times 256 \times 512$	$14 \times 14 \times 256$
$5 \times$	Conv dw / s1 $3 \times 3 \times 512$ dw	$14 \times 14 \times 512$
	Conv / s1 $1 \times 1 \times 512 \times 512$	$14 \times 14 \times 512$
Conv dw / s2	$3 \times 3 \times 512$ dw	$14 \times 14 \times 512$
Conv / s1	$1 \times 1 \times 512 \times 1024$	$7 \times 7 \times 512$
Conv dw / s2	$3 \times 3 \times 1024$ dw	$7 \times 7 \times 1024$
Conv / s1	$1 \times 1 \times 1024 \times 1024$	$7 \times 7 \times 1024$
Avg Pool / s1	Pool 7×7	$7 \times 7 \times 1024$
FC / s1	1024×1000	$1 \times 1 \times 1024$
Softmax / s1	Classifier	$1 \times 1 \times 1000$

Εικόνα 3.11 MobileNet Layers [28]

Το MobileNet βασίζονται σε μια απλοποιημένη αρχιτεκτονική (streamed line architecture) που χρησιμοποιεί διαχωρίσιμες συνελιξής βάθους για τη δημιουργία ελαφρών νευρωνικών δικτύων. [28]

Run: MobileNetTransferLearning ×

Model: "model"

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	[None, 224, 224, 3]	0
conv1_pad (ZeroPadding2D)	(None, 225, 225, 3)	0
conv1 (Conv2D)	(None, 112, 112, 32)	864
conv1_bn (BatchNormalization)	(None, 112, 112, 32)	128
conv1_relu (ReLU)	(None, 112, 112, 32)	0
conv_dw_1 (DepthwiseConv2D)	(None, 112, 112, 32)	288
conv_dw_1_bn (BatchNormaliza	(None, 112, 112, 32)	128
conv_dw_1_relu (ReLU)	(None, 112, 112, 32)	0
conv_pw_1 (Conv2D)	(None, 112, 112, 64)	2048
conv_pw_1_bn (BatchNormaliza	(None, 112, 112, 64)	256
conv_pw_1_relu (ReLU)	(None, 112, 112, 64)	0
conv_pad_2 (ZeroPadding2D)	(None, 113, 113, 64)	0
conv_dw_2 (DepthwiseConv2D)	(None, 56, 56, 64)	576
conv_dw_2_bn (BatchNormaliza	(None, 56, 56, 64)	256
conv_dw_2_relu (ReLU)	(None, 56, 56, 64)	0
conv_pw_2 (Conv2D)	(None, 56, 56, 128)	8192
conv_pw_2_bn (BatchNormaliza	(None, 56, 56, 128)	512

4: Run 6: TODO Terminal Python Console

IDE and Plugin Updates: PyCharm is ready to update. (13 minutes ago)

Εικόνα 3.12 MobileNet Model Summary part1

Run: MobileNetTransferLearning x

Εικόνα 3.13 MobileNet Model Summary part2

Run:	Python	MobileNetTransferLearning	x		
▶	↑	conv_pw_5_relu (ReLU)	(None, 28, 28, 256)	0	
■	↓	conv_pad_6 (ZeroPadding2D)	(None, 29, 29, 256)	0	
≡	↺	conv_dw_6 (DepthwiseConv2D)	(None, 14, 14, 256)	2304	
✦	↻	conv_dw_6_bn (BatchNormaliza	(None, 14, 14, 256)	1024	
	🗑	conv_dw_6_relu (ReLU)	(None, 14, 14, 256)	0	
		conv_pw_6 (Conv2D)	(None, 14, 14, 512)	131072	
		conv_pw_6_bn (BatchNormaliza	(None, 14, 14, 512)	2048	
		conv_pw_6_relu (ReLU)	(None, 14, 14, 512)	0	
		conv_dw_7 (DepthwiseConv2D)	(None, 14, 14, 512)	4608	
		conv_dw_7_bn (BatchNormaliza	(None, 14, 14, 512)	2048	
		conv_dw_7_relu (ReLU)	(None, 14, 14, 512)	0	
		conv_pw_7 (Conv2D)	(None, 14, 14, 512)	262144	
		conv_pw_7_bn (BatchNormaliza	(None, 14, 14, 512)	2048	
		conv_pw_7_relu (ReLU)	(None, 14, 14, 512)	0	
		conv_dw_8 (DepthwiseConv2D)	(None, 14, 14, 512)	4608	
		conv_dw_8_bn (BatchNormaliza	(None, 14, 14, 512)	2048	
		conv_dw_8_relu (ReLU)	(None, 14, 14, 512)	0	
		conv_pw_8 (Conv2D)	(None, 14, 14, 512)	262144	
		conv_pw_8_bn (BatchNormaliza	(None, 14, 14, 512)	2048	
▶ 4: Run	☰ 6: TODO	📄 Terminal	🔗 Python Console	IDE and Plugin Updates: PyCharm is ready to update. (16 minutes ago)	

Εικόνα 3.14 MobileNet Model Summary part3

Run:	MobileNetTransferLearning ×
▶	conv_pw_8_relu (ReLU) (None, 14, 14, 512) 0
■	conv_dw_9 (DepthwiseConv2D) (None, 14, 14, 512) 4608
⌵	conv_dw_9_bn (BatchNormaliza (None, 14, 14, 512) 2048
⌵	conv_dw_9_relu (ReLU) (None, 14, 14, 512) 0
⌵	conv_pw_9 (Conv2D) (None, 14, 14, 512) 262144
⌵	conv_pw_9_bn (BatchNormaliza (None, 14, 14, 512) 2048
⌵	conv_pw_9_relu (ReLU) (None, 14, 14, 512) 0
⌵	conv_dw_10 (DepthwiseConv2D) (None, 14, 14, 512) 4608
⌵	conv_dw_10_bn (BatchNormaliz (None, 14, 14, 512) 2048
⌵	conv_dw_10_relu (ReLU) (None, 14, 14, 512) 0
⌵	conv_pw_10 (Conv2D) (None, 14, 14, 512) 262144
⌵	conv_pw_10_bn (BatchNormaliz (None, 14, 14, 512) 2048
⌵	conv_pw_10_relu (ReLU) (None, 14, 14, 512) 0
⌵	conv_dw_11 (DepthwiseConv2D) (None, 14, 14, 512) 4608
⌵	conv_dw_11_bn (BatchNormaliz (None, 14, 14, 512) 2048
⌵	conv_dw_11_relu (ReLU) (None, 14, 14, 512) 0
⌵	conv_pw_11 (Conv2D) (None, 14, 14, 512) 262144
⌵	conv_pw_11_bn (BatchNormaliz (None, 14, 14, 512) 2048
⌵	conv_pw_11_relu (ReLU) (None, 14, 14, 512) 0
▶ 4: Run	6: TODO Terminal Python Console
IDE and Plugin Updates: PyCharm is ready to update. (18 minutes ago)	

Εικόνα 3.15 MobileNet Model Summary part4

Run: MobileNetTransferLearning ×

▶

⬆

■

⬇

≡

⌵

⬇

🔍

🖨

🗑

conv_pad_12	(ZeroPadding2D)	(None, 15, 15, 512)	0
conv_dw_12	(DepthwiseConv2D)	(None, 7, 7, 512)	4608
conv_dw_12_bn	(BatchNormaliz	(None, 7, 7, 512)	2048
conv_dw_12_relu	(ReLU)	(None, 7, 7, 512)	0
conv_pw_12	(Conv2D)	(None, 7, 7, 1024)	524288
conv_pw_12_bn	(BatchNormaliz	(None, 7, 7, 1024)	4096
conv_pw_12_relu	(ReLU)	(None, 7, 7, 1024)	0
conv_dw_13	(DepthwiseConv2D)	(None, 7, 7, 1024)	9216
conv_dw_13_bn	(BatchNormaliz	(None, 7, 7, 1024)	4096
conv_dw_13_relu	(ReLU)	(None, 7, 7, 1024)	0
conv_pw_13	(Conv2D)	(None, 7, 7, 1024)	1048576
conv_pw_13_bn	(BatchNormaliz	(None, 7, 7, 1024)	4096
conv_pw_13_relu	(ReLU)	(None, 7, 7, 1024)	0
global_average_pooling2d	(Gl	(None, 1024)	0
dense	(Dense)	(None, 2)	2050
=====			
Total params: 3,230,914			
Trainable params: 3,209,026			
Non-trainable params: 21,888			

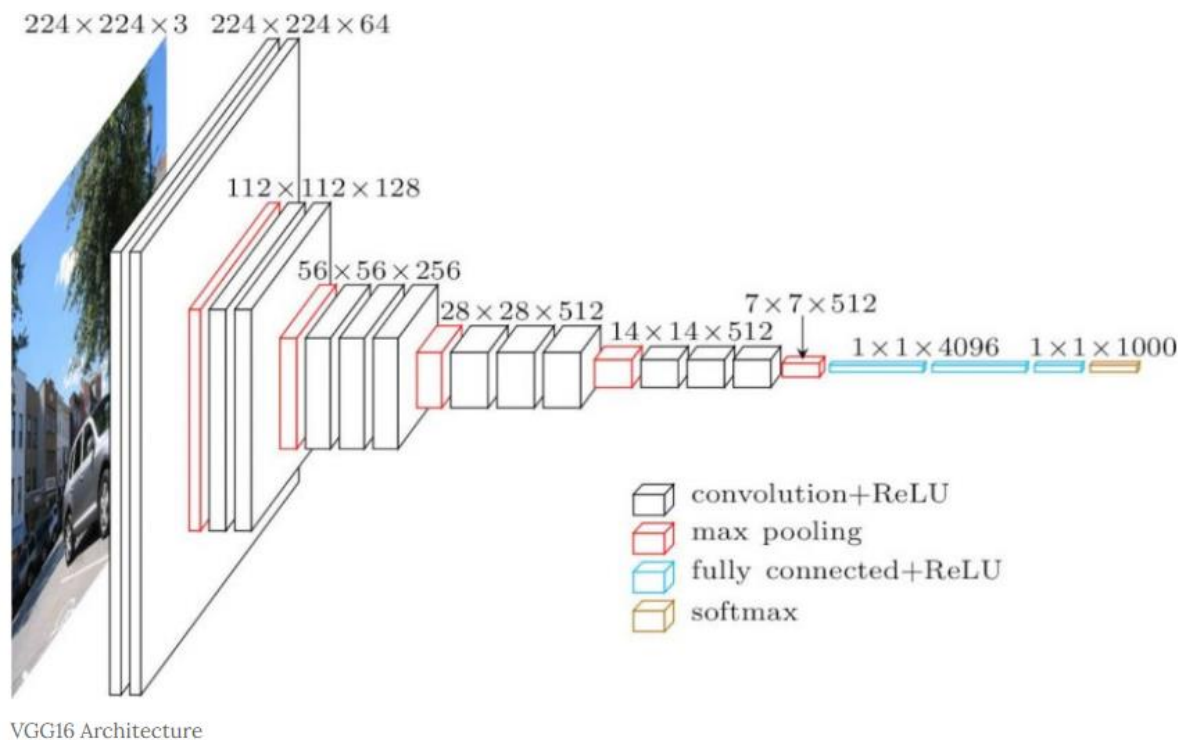
Εικόνα 3.16 MobileNet Model Summary part5

Αυτή η αρχιτεκτονική χρησιμοποιεί σε βάθος διαχωρίσιμες συνελιξείς που μειώνουν, όπως ανέφερα σημαντικά τον αριθμό των παραμέτρων σε σύγκριση με το δίκτυο με κανονικές συνελιξείς με το ίδιο βάθος στα δίκτυα.

Οι εικόνες που παραθέτω από πάνω δείχνουν το μοντέλο MobileNet το οποίο είναι και τύπου Model Keras Functional API και έχει γίνει ένα fine tune με τρόπο ώστε να τρέξει όπως θέλω για το δικό μου binary classification πρόβλημα. Επειδή το MobileNet δίκτυο, όταν το φόρτωσα ήταν εκπαιδευμένο να κατηγοριοποιεί εικόνες από το ImageNet σε 1000 κατηγορίες, χρειάστηκε να του αφαιρέσω τα τελευταία 5 layers μαζί με και το dense layer και κατόπιν να του προσθέσω ένα dense layer το οποίο δρομολογεί το δίκτυο στο να αντιμετωπίζει ένα πρόβλημα binary classification.

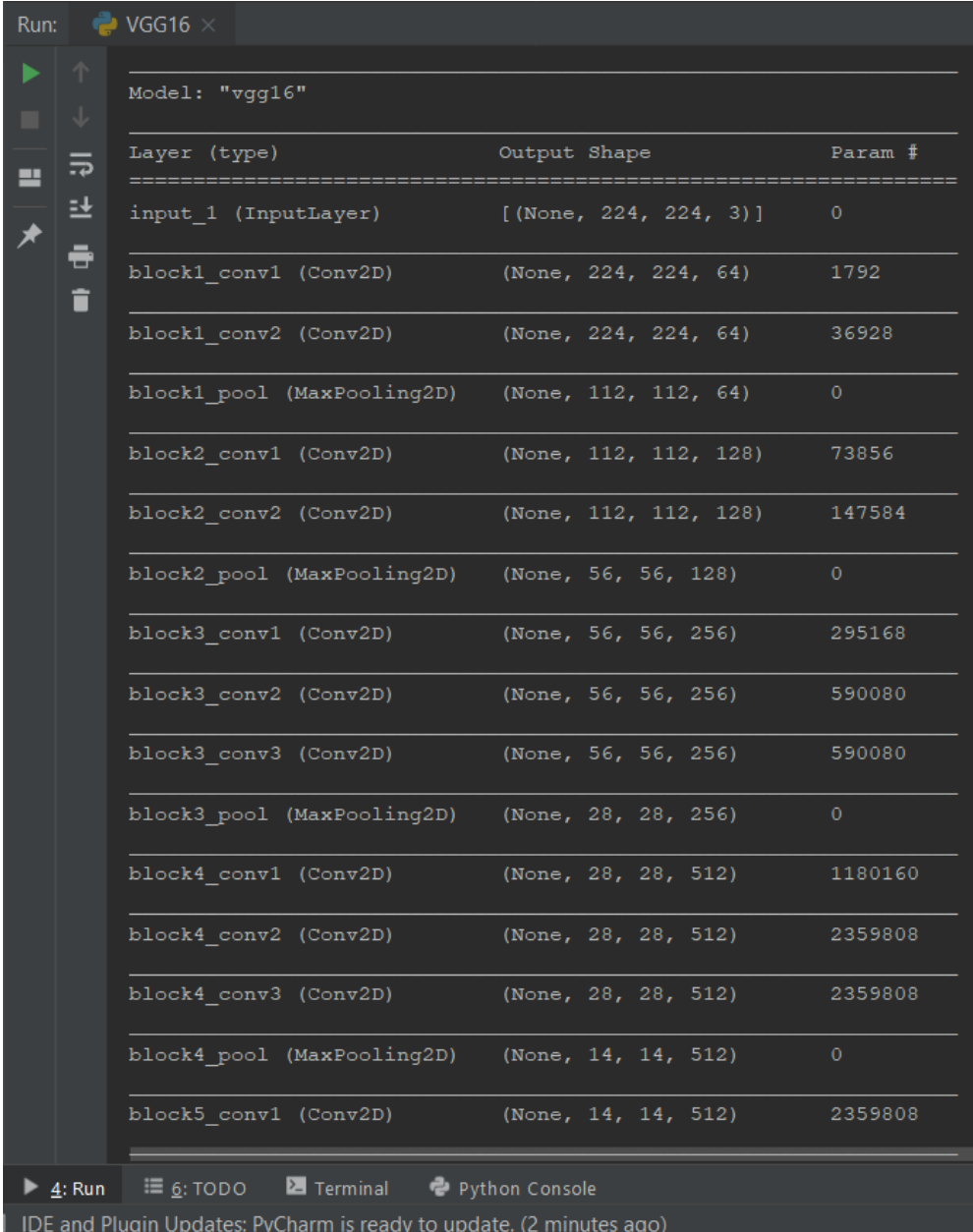
Τα αποτελέσματα που έλαβα, τα οποία θα αναλυθούν περαιτέρω στο επόμενο κεφάλαιο, από αυτή την αρχιτεκτονική ήταν πολύ ικανοποιητικά δεδομένου του χρόνου υλοποίησης και μεγέθους του δικτύου (87,3%).

3.4.5 Vgg16Net Αρχιτεκτονική



Εικόνα 3.17 VGG16 Αρχιτεκτονική

Το VGG16 είναι μια αρχιτεκτονική neural net (CNN) που χρησιμοποιήθηκε για να κερδίσει το διαγωνισμό ILSVR (Imagenet) το 2014. Θεωρείται ως μια από τις εξαιρετικές αρχιτεκτονικές μοντέλων υπολογιστικής όρασης μέχρι σήμερα. Το εξέχον στοιχείο όσον αφορά στο VGG16 είναι ότι, αντί να έχουν μεγάλο αριθμό υπέρ-παραμέτρων, επικεντρώθηκαν στο να έχουν στρώματα συνελίξεων 3x3 με stride 1 και πάντα χρησιμοποιούσαν το ίδιο στρώμα padding και maxpooling 2x2 με stride 2. Στο τέλος έχει 2 FC (Fully Connected Layers) ακολουθούμενα από ένα softmax για έξοδο. Ο αριθμός 16 στο VGG16 αναφέρεται στο ότι έχει 16 στρώματα που έχουν βάρη. Αυτό το δίκτυο είναι ένα αρκετά μεγάλο δίκτυο και έχει περίπου 138 εκατομμύρια (περίπου) παραμέτρους.



Run: VGG16 x

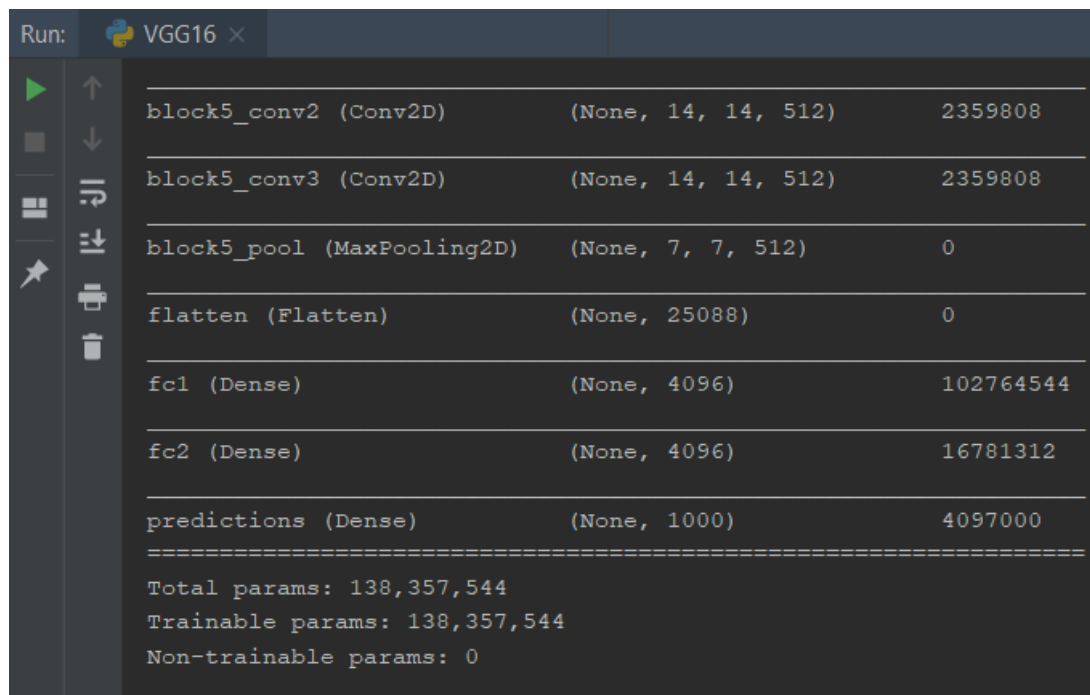
Model: "vgg16"

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	[(None, 224, 224, 3)]	0
block1_conv1 (Conv2D)	(None, 224, 224, 64)	1792
block1_conv2 (Conv2D)	(None, 224, 224, 64)	36928
block1_pool (MaxPooling2D)	(None, 112, 112, 64)	0
block2_conv1 (Conv2D)	(None, 112, 112, 128)	73856
block2_conv2 (Conv2D)	(None, 112, 112, 128)	147584
block2_pool (MaxPooling2D)	(None, 56, 56, 128)	0
block3_conv1 (Conv2D)	(None, 56, 56, 256)	295168
block3_conv2 (Conv2D)	(None, 56, 56, 256)	590080
block3_conv3 (Conv2D)	(None, 56, 56, 256)	590080
block3_pool (MaxPooling2D)	(None, 28, 28, 256)	0
block4_conv1 (Conv2D)	(None, 28, 28, 512)	1180160
block4_conv2 (Conv2D)	(None, 28, 28, 512)	2359808
block4_conv3 (Conv2D)	(None, 28, 28, 512)	2359808
block4_pool (MaxPooling2D)	(None, 14, 14, 512)	0
block5_conv1 (Conv2D)	(None, 14, 14, 512)	2359808

4: Run 6: TODO Terminal Python Console

IDE and Plugin Updates: PyCharm is ready to update. (2 minutes ago)

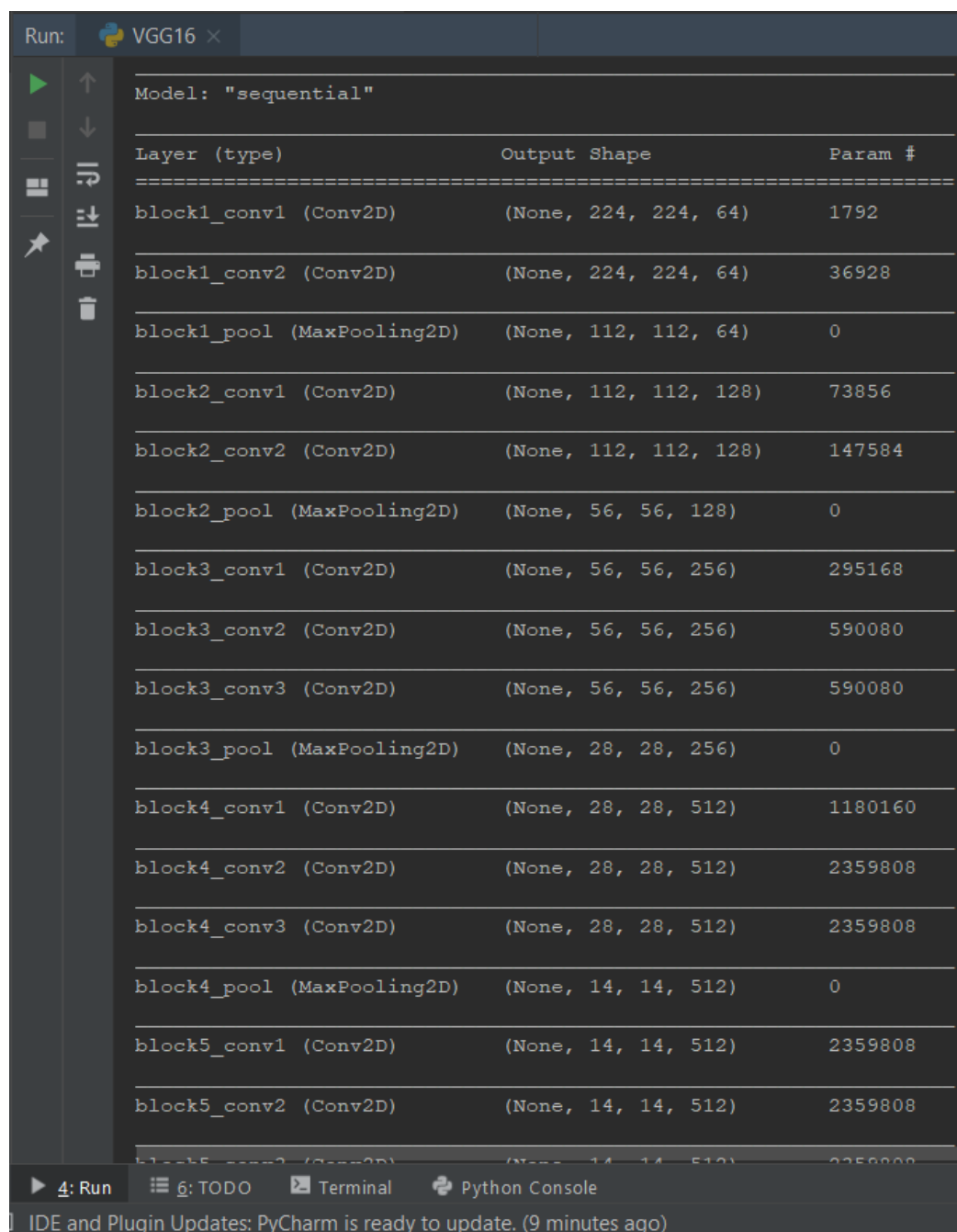
Εικόνα 3.18 VGG16 Model Summary part 1 (Model Keras Functional API)



Layer	Shape	Params
block5_conv2 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv3 (Conv2D)	(None, 14, 14, 512)	2359808
block5_pool (MaxPooling2D)	(None, 7, 7, 512)	0
flatten (Flatten)	(None, 25088)	0
fc1 (Dense)	(None, 4096)	102764544
fc2 (Dense)	(None, 4096)	16781312
predictions (Dense)	(None, 1000)	4097000
=====		
Total params: 138,357,544		
Trainable params: 138,357,544		
Non-trainable params: 0		

Εικόνα 3.19 VGG16 Model Summary part 2 (Model Keras Functional API)

Σε αυτό το σημείο οι εικόνες που δείχνουν τα layers του δικτύου VGG16 είναι τύπου Keras Functional API και το τελευταίο είναι ένα Dense Layer το οποίο κάνει classify 1000 κατηγορίες επειδή αυτό το δίκτυο είναι ήδη εκπαιδευμένο στο ImageNet. Για αυτόν το λόγο του κάνω ένα Fine Tuning αφαιρώντας του το τελευταίο Dense Layer και κατόπιν του προσθέτω ένα καινούργιο Dense Layer το οποίο κάνει classify 2 κατηγορίες Benign και Malignant.



Run: VGG16 x

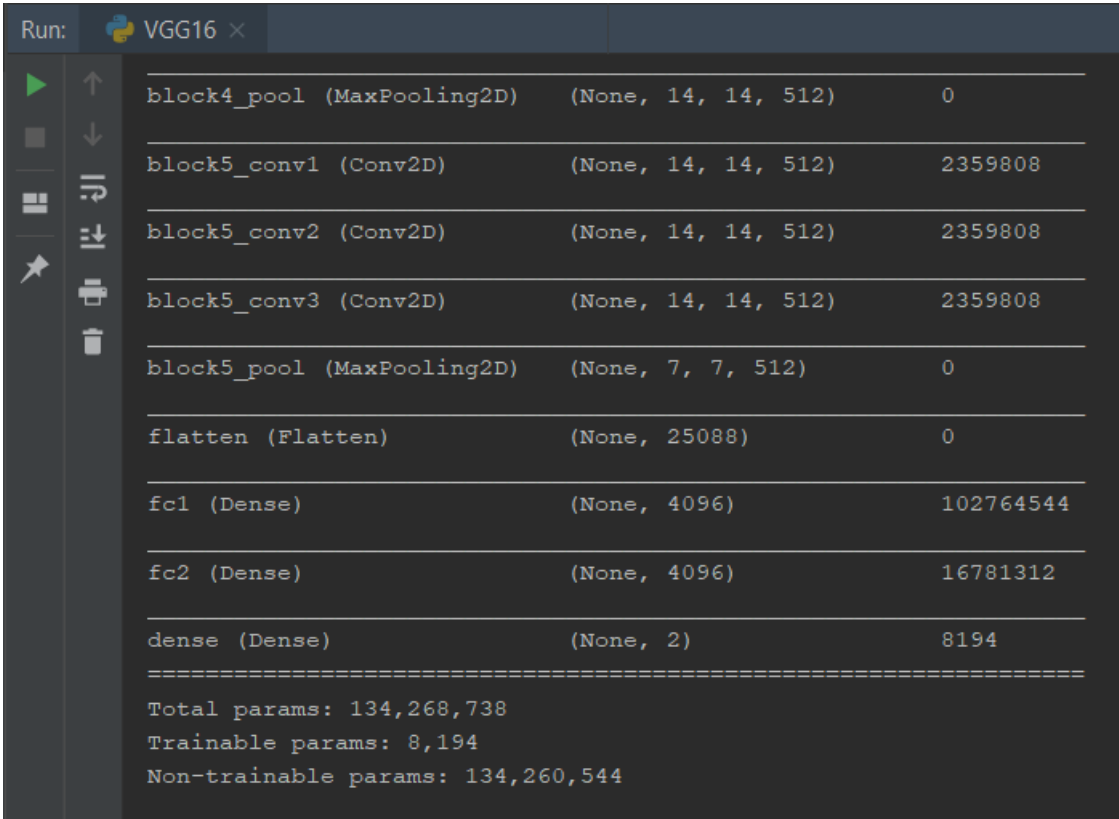
Model: "sequential"

Layer (type)	Output Shape	Param #
block1_conv1 (Conv2D)	(None, 224, 224, 64)	1792
block1_conv2 (Conv2D)	(None, 224, 224, 64)	36928
block1_pool (MaxPooling2D)	(None, 112, 112, 64)	0
block2_conv1 (Conv2D)	(None, 112, 112, 128)	73856
block2_conv2 (Conv2D)	(None, 112, 112, 128)	147584
block2_pool (MaxPooling2D)	(None, 56, 56, 128)	0
block3_conv1 (Conv2D)	(None, 56, 56, 256)	295168
block3_conv2 (Conv2D)	(None, 56, 56, 256)	590080
block3_conv3 (Conv2D)	(None, 56, 56, 256)	590080
block3_pool (MaxPooling2D)	(None, 28, 28, 256)	0
block4_conv1 (Conv2D)	(None, 28, 28, 512)	1180160
block4_conv2 (Conv2D)	(None, 28, 28, 512)	2359808
block4_conv3 (Conv2D)	(None, 28, 28, 512)	2359808
block4_pool (MaxPooling2D)	(None, 14, 14, 512)	0
block5_conv1 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv2 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv3 (Conv2D)	(None, 14, 14, 512)	2359808

4: Run 6: TODO Terminal Python Console

IDE and Plugin Updates: PyCharm is ready to update. (9 minutes ago)

Εικόνα 3.20 VGG16 Model Summary part 1 (Sequential Model)



block4_pool (MaxPooling2D)	(None, 14, 14, 512)	0
block5_conv1 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv2 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv3 (Conv2D)	(None, 14, 14, 512)	2359808
block5_pool (MaxPooling2D)	(None, 7, 7, 512)	0
flatten (Flatten)	(None, 25088)	0
fc1 (Dense)	(None, 4096)	102764544
fc2 (Dense)	(None, 4096)	16781312
dense (Dense)	(None, 2)	8194
=====		
Total params:		134,268,738
Trainable params:		8,194
Non-trainable params:		134,260,544

Εικόνα 3.21 VGG16 Model Summary part 2 (Sequential Model)

Σε αυτό το σημείο φαίνεται το Fine tuned VGG16. Για να φτάσω σε αυτό το σημείο όρισα ένα καινούργιο κενό μοντέλο τύπου sequential και ύστερα σε ένα loop πέρασα κάθε layer από το προηγούμενο μοντέλο στο κενό μοντέλο, εκτός όμως από το τελευταίο layer (predictions Dense), έτσι ώστε να προσθέσω μετά το καινούργιο Dense layer για να εξασφαλίσω την κατηγοριοποίηση σε 2 κατηγορίες και όχι σε 1000.

Τώρα στο εν λόγω CNN τα αποτελέσματα, στα οποία θα αναφερθώ εκτενώς στο επόμενο κεφάλαιο, που είδα ήταν τα καλύτερα σε σχέση με τα άλλα νευρωνικά που ανέφερα προηγουμένως, πράγμα το οποίο είναι λογικό καθώς το δίκτυο είναι και μεγάλο (528 MB) αλλά και βαρύ (97% accuracy στο testing).

3.5 Μεθοδολογία

3.5.1 Προετοιμασία δεδομένων

Η προετοιμασία δεδομένων είναι ένα κρίσιμο και αναπόφευκτο βήμα στη διαδικασία δημιουργίας ενός αξιόπιστου συστήματος. Σε αυτήν τη συγκεκριμένη μελέτη στην πρώτη φάση χρησιμοποίησα 250 εικόνες μαστού από υπερήχους και χώρισα τα δεδομένα σε ομάδες εκπαίδευσης και παρτίδες επικύρωσης με αναλογία εκπαίδευσης 75%, δοκιμές 20% και επικύρωση 5%. Οι κλάσεις μου χαρακτηρίστηκαν ως benign και malignant και το σύνολο δεδομένων ήταν ισορροπημένο. Επίσης, με τον ίδιο τρόπο ετοίμασα τα επόμενα δύο σύνολα δεδομένων ιστοπαθολογικών εικόνων με περιπτώσεις IDC (Invasive Ductal Carcinoma) όπου επισημαίνουμε τα δεδομένα ως Cancerous (Καρκινικά) και Non-Cancerous (Μη Καρκινικά) και για τα δύο σύνολα δεδομένων. Η πρώτη αποτελείται από 3628 εικόνες ιστού του μαστού και η δεύτερη αποτελείται από 10483 εικόνες και τα δύο σύνολα δεδομένων προέρχονται από το Kaggle και όπου διατίθενται στο κοινό για κάθε ερευνητή. Επίσης, όσον αφορά στις ιστοπαθολογικές εικόνες και τα δύο σύνολα δεδομένων ήταν ισορροπημένα.

3.5.2 Προεπεξεργασία δεδομένων

Όσο για την προεπεξεργασία αυτών των εικόνων, αυτό που έκανα ήταν, αρχικά ένα image normalization κάνοντας rescaling pixel values, στη συνέχεια, resizing αυτών των εικόνων και, τέλος, data augmentation σε πραγματικό χρόνο μέσω του ImageDataGenerator.flowfromdirectory class του keras functional api, ενώ παράγονται παρτίδες ελαστικών εικόνων για τροφοδοσία στα μοντέλα μας για εκπαίδευση. Πιο συγκεκριμένα Horizontal flip: True που σημαίνει ότι αναστρέφει τυχαία τις εισόδους οριζόντια. Vertical flip: True, που σημαίνει ότι αναστρέφει τυχαία τις εισόδους κάθετα. Rescale: 1./255 που αποτελεί παράγοντα επανασύνδεσης. Shear range-Εύρος διάτμησης: 0,2. Τώρα σχετικά με το νευρωνικό δίκτυο MobileNet, η Keras διαθέτει μια ενσωματωμένη λειτουργία

προεπεξεργασίας την οποία χρησιμοποίησα επίσης για να επεξεργαστώ τα δεδομένα μου ανάλογα. Αυτή η λειτουργία προεπεξεργασίας του MobileNet εφαρμόστηκε σε κάθε είσοδο πριν δοθεί στο νευρωνικό δίκτυο. Η λειτουργία εκτελέστηκε μετά το `resizing` και το `augmentation procedure` κάθε εικόνας εισόδου.

Data Augmentation

Το Data Augmentation είναι ένα πολύ σοβαρό βήμα στο στάδιο της προεπεξεργασίας του έργου μου λαμβάνοντας υπόψη το γεγονός ότι οι όγκοι δεν έχουν εγγενή προσανατολισμό και η ανίχνευσή τους είναι διαρκής για τον μετασχηματισμό, πράγμα που σημαίνει ότι αυτή η διαδικασία της επαύξησης δεδομένων δεν μπορεί να αλλοιώσει το αποτέλεσμα της διάγνωσης.

3.5.3 Deep Structured Learning Theory

Η βαθιά δομημένη μάθηση είναι μια υποκατηγορία αλγορίθμων μηχανικής μάθησης που χρησιμοποιεί πολλαπλά επίπεδα για να εξάγει προοδευτικά χαρακτηριστικά υψηλότερου επιπέδου από την ακατέργαστη είσοδο.

- **Input Layers:** αυτό το επίπεδο φορτώνει είσοδο και παράγει έξοδο που χρησιμοποιείται για την τροφοδοσία συνελικτικών στρωμάτων. Σε αυτή τη μελέτη, οι είσοδοι είναι ιατρικές εικόνες. Για το MobileNet και το VGG16 το μέγεθος των εικόνων έγινε σε 224x224x3 πριν δοθεί στο δίκτυο. Για το AlexNet οι εικόνες άλλαξαν το μέγεθος τους σε 227x227x3.

- **Convolutional layers:** ένα επίπεδο συνελίξης ενοποιεί την εικόνα εισόδου με ένα σύνολο από learnable filters που μπορούν να διδαχθούν, το καθένα παράγει έναν χάρτη χαρακτηριστικών στην εικόνα εξόδου. Στο MobileNet υπάρχουν πολλές depth-wise συνελίξεις που καθιστούν το υπολογιστικό κόστος πολύ χαμηλό και έτσι το δίκτυο εξαιρετικά αποτελεσματικό σε δύσκολες εργασίες. Στο VGG16 υπάρχουν δεκατρία συνελικτικά στρώματα και τα δεκτικά πεδία (πυρήνες) έχουν μέγεθος 3×3 . Το AlexNet έχει ένα πρώτο συνελικτικό στρώμα με μέγεθος φίλτρου 11x11, τα

δεικτικά πεδία (πυρήνες) έχουν μέγεθος 3×3 . Το Conv_1 αποτελείται από 11×11 φίλτρα, το Conv_2 χρησιμοποιεί φίλτρα 5×5 και το Conv_3 χρησιμοποιεί φίλτρα 3×3 , το Conv_4 χρησιμοποιεί φίλτρα 3×3 και το Conv_5 χρησιμοποιεί φίλτρα 3×3 .

- Pooling Layers: ο στόχος αυτών των επιπέδων είναι η υποδειγματοληψία της χωρικής διάστασης της εισόδου. Στο VGG16 υπάρχει ένα στρώμα συγκέντρωσης μετά από κάθε συνελικτικό στρώμα. Όλα είναι έτοιμα να χρησιμοποιήσουν ένα πεδίο υποδοχής 2×2 (χωρική έκταση) με ένα βήμα 2. Επίσης, όλα τα στρώματα συγκέντρωσης χρησιμοποιούν τη συνηθέστερη μέγιστη λειτουργία πάνω από το πεδίο υποδοχής. Στο AlexNet, το πέμπτο συνελικτικό επίπεδο ακολουθείται από ένα στρώμα Overlapping Max Pooling.

- Flatten Layer: Ένα επίπεδο επίπεδο καταρρέει τις χωρικές διαστάσεις της εισόδου στην διάσταση του καναλιού.

- Dense Layer: Το πυκνό στρώμα είναι το συνηθισμένο βαθύ συνδεδεμένο στρώμα νευρωνικού δικτύου και το πιο κοινό επίπεδο εξόδου. Σε αυτή τη μελέτη τα Vgg16, AlexNet και MobileNet έχουν ένα τελευταίο πυκνό επίπεδο για δυαδική ταξινόμηση.

Depth-Wise Convolutional Neural Network

MobileNet

Το MobileNet είναι μια αρχιτεκτονική που ταιριάζει καλύτερα σε φορητές και ενσωματωμένες εφαρμογές υπολογιστικής όρασης όπου υπάρχει έλλειψη υπολογιστικής ισχύος. Αυτή η αρχιτεκτονική προτάθηκε από την Google. Αυτό το δίκτυο είναι ένα αρκετά ελαφρύ δίκτυο νευρώνων, αλλά λόγω της έξυπνης αρχιτεκτονικής του, λειτουργεί πολύ αποδοτικά. Το MobileNet βασίζεται σε μια απλοποιημένη αρχιτεκτονική που χρησιμοποιεί ουσιαστικά depthwise separable convolutions που είναι μια μορφή παραγοντοποιημένων συνελίξεων που παραγοντοποιούν μια τυπική συνέλιξη σε μια depth-wise συνέλιξη και μια συνέλιξη 1×1 που ονομάζεται point-wise convolution. Για το

MobileNet, η συνεχής ανάλυση εφαρμόζει ένα μόνο φίλτρο σε κάθε κανάλι εισόδου. Είναι βασικά ένα πολύ μικρό μοντέλο χαμηλού λανθάνοντος χρόνου που μπορεί να ταιριάζει με τις απαιτήσεις σχεδίασης για εφαρμογές κινητής και ενσωματωμένης όρασης. Η βασική σχεδίαση MobileNet είναι μικρή και χαμηλή καθυστέρηση. Συνήθως μια συγκεκριμένη περίπτωση χρήσης ή εφαρμογή μπορεί να απαιτεί το μοντέλο να είναι μικρότερο και γρηγορότερο.

Depth-wise Separable Convolution

Οι depth-wise συνελίξεις είναι μια μορφή παραγοντοποιημένων συνελίξεων που παραγοντοποιούν μια τυπική συνέλιξη σε μια βαθιά συνεκτική συνέλιξη. Στο MobileNet, η βαθιά συνέλιξη εφαρμόζει ένα μόνο φίλτρο σε κάθε κανάλι εισόδου. Στη συνέχεια, η περιστροφική κατάκλιση εφαρμόζει μια συνέλιξη $1 \times 11 \times 1$ για να συνδυάσει τις εξόδους με τη συνεχή συνεχή, αυτή είναι που καθιστά τη διαχωριστική συνέλιξη σε βάθος πολύ χαμηλού κόστους υπολογιστική. Έτσι το MobileNet μπορεί εύκολα να χρησιμοποιηθεί για συστήματα χωρίς σημαντική υπολογιστική χωρητικότητα, όπως φορητοί υπολογιστές χωρίς GPU ή CUDA για παράλληλη επεξεργασία, ενσωματωμένα συστήματα, εφαρμογές για κινητά.

3.5.4 Fine-Tuning

Σε αυτήν τη μελέτη, βελτίωσα την αρχιτεκτονική των νευρωνικών δικτύων VGG16, AlexNet και MobileNet, προκειμένου να τη χρησιμοποιήσω στο σύνολο των ιατρικών εικόνων. Το Fine-tuning είναι ένα επιτακτικό βήμα στη διαδικασία, καθώς πρέπει να εκμεταλλευτώ, όσο γνωρίζω, τα προνόμια αυτής της εικεντρικής αρχιτεκτονικής. Αρχικά εισάγω τα δίκτυα και στη συνέχεια το συντονίζω διαγράφοντας ουσιαστικά τα τελευταία επίπεδα των αρχικών μοντέλων που κατασκευάστηκε για να ταξινομήσει δεδομένα 1000 κατηγοριών. Επιπλέον, προσθέτω ένα επίπεδο εξόδου (προβλέψεις) που είναι ένα dense επίπεδο με δύο εξόδους για benign και malignant, ή cancerous και non-cancerous.

3.5.5 Υλοποίηση και Στρατηγική Εκπαίδευσης

Η προτεινόμενη μέθοδος στοχεύει στη σύγκριση διαφορετικών δομών συνελικτικών νευρωνικών δικτύων σε διάφορα σύνολα δεδομένων για θέματα καρκίνου του μαστού διαφορετικών όγκων και μεγεθών. Αρχικά, πριν καταρτίσουμε τα μοντέλα μας, αρχίζουμε να τα προσαρμόζουμε το καθένα με σκοπό τη δυαδική ταξινόμηση [31]. MobileNet, VGG16 & AlexNet: Μετά τη φόρτωση του μοντέλου, αφαιρέσαμε τα τελευταία επίπεδα και στη συνέχεια προσθέσαμε ένα νέο πυκνό επίπεδο που ταξινομεί δύο κατηγορίες εικόνων μαζί με μια λειτουργία ενεργοποίησης Softmax. Όσον αφορά τη συνάρτηση Softmax, είναι επιτακτική ανάγκη να κατανοήσουμε ότι είναι ένα βασικό στοιχείο στις εργασίες ταξινόμησης βαθιάς μάθησης. Ουσιαστικά, είναι μια συνάρτηση ενεργοποίησης που μετατρέπει τους αριθμούς σε πιθανότητες που συνοψίζονται σε μία. Η συνάρτηση Softmax εξάγει ένα διάλυμα που καθορίζει τις κατανομές πιθανότητας μιας λίστας πιθανών αποτελεσμάτων.

$$Softmax(x)_i = \frac{e^{x_i}}{\sum_{j=1}^n e^{x_j}}$$

Training Strategy

Το πρωτόκολλο εκπαίδευσης που χρησιμοποιήσαμε εποπτεύεται καθαρά. Ως βελτιστοποιητής χρησιμοποιήσαμε το πρόγραμμα βελτιστοποίησης Adam με ρυθμό εκμάθησης 0,0001, το οποίο είναι ένα εργαλείο βελτιστοποίησης ειδικά εξοπλισμένο για την εκπαίδευση νευρωνικών δικτύων. Είναι ένας αλγόριθμος που βασικά αξιοποιεί τη δύναμη των προσαρμοστικών μεθόδων εκμάθησης για να βρει ατομικά ποσοστά εκμάθησης για κάθε παράμετρο. [15] Ως συνάρτηση απώλειας χρησιμοποιήσαμε κατηγορηματική εγκάρσια εντροπία, η οποία είναι μια λειτουργία ενεργοποίησης Softmax μαζί με μια απώλεια Cross-Entropy και ουσιαστικά την χρησιμοποιήσαμε για να αποδώσουμε μια πιθανότητα για τις τάξεις μας για κάθε ιατρική εικόνα. Τέλος, ως μετρήσεις χρησιμοποιήσαμε την Accuracy, Sensitivity, Specificity, F1-score.



$$CE = -\log\left(\frac{e^{S_p}}{\sum_j^C e^{S_j}}\right)$$

Softmax Activation Function & Cross-Entropy Loss

3.5.6 System Specifications & Python Frameworks

Για την υλοποίηση αυτού του εν λόγω έργου χρησιμοποίησα το υπολογιστικό σύστημα με τα χαρακτηριστικά που παραθέτω παρακάτω.

System Specifications

- **Processor** Intel® Core™ i76700HQ CPU @ 2.60GHz, 4 cores, 8 Logical Processors
- **OS** Microsoft Windows 10
- **Version** - 10.0.17134 Build 17134
- **System Type** x64-based PC
- **Installed Physical Memory (RAM)** 8.00 GB
- **Total Physical Memory** 7.89 GB
- **GPU** Nvidia GTX 950m
- **GPU Memory** 4GB
- **CUDA Version** 10.0.130

Τον κώδικα τον έγραψα σε Python (python Version 3.6) και το περιβάλλον το έστησα στο VirtualEnv της python.

Τον κώδικά μου τον τρέχω με Tensorflow Backend και χρησιμοποιώ βιβλιοθήκες και του Keras API καθώς και άλλες γνωστές όπως Scikitlearn, Numpy, Matplotlib.

Για την πιο εύκολη και γρήγορη υλοποίηση χρησιμοποιώ το σύστημα παράλληλης επεξεργασίας CUDA της κάρτας γραφικών, διότι το πρότζεκτ απαιτεί μεγάλη υπολογιστική ισχύ για να φορτώσει και να φέρει σε πέρας δύσκολη ανάλυση πολλών δεδομένων.

4

Αποτελέσματα & Ανάλυση

Στο προηγούμενο μέρος ανέλυσα την μεθοδολογία μου για την αναγνώριση των εικόνων σχετικά, πρώτον με το αν είναι οι όγκοι καλοήθεις ή κακοήθεις, δεύτερον τις αρχιτεκτονικές των νευρωνικών δικτύων τις οποίες χρησιμοποίησα για την ανάλυση των εικόνων, καθώς και την προεπεξεργασία την οποία έκανα, έτσι ώστε να λάβω τα καλύτερα δυνατά αποτελέσματα. Κατόπιν θα αναλύσω τα αποτελέσματα μου ως προς την ευστοχία της κάθε αρχιτεκτονικής αλλά και τη διαφορά που υπήρξε σε σχέση με την προεπεξεργασία που έκανα στα υπερηχογραφήματα.

Τα αποτελέσματα σε γενικές γραμμές κυμαίνονται από μέτρια ως εξαιρετικά, ανάλογα και με το μέγεθος αλλά και την πολυπλοκότητα των δικτύων. Στα παρακάτω υποκεφάλαια παραθέτω τα αντίστοιχα διαγράμματα που δείχνουν το επίπεδο κατάρτισης που φτάνει κάθε αρχιτεκτονική, τα επίπεδα ακρίβειας/accuracy που επιτυγχάνουν όσο περνάνε οι εποχές εκπαίδευσης/epochs, καθώς και το loss που έχει σε σχέση πάλι με τον χρόνο.

Η ακρίβεια/accuracy μπορεί να θεωρηθεί ως ο αριθμός σφαλμάτων που έκανα ουσιαστικά στα δεδομένα. Αυτό σημαίνει ότι η χαμηλή ακρίβεια και τεράστια απώλεια σημαίνει ότι έγιναν τεράστια σφάλματα σε πολλά δεδομένα, η χαμηλή ακρίβεια αλλά και χαμηλή απώλεια σημαίνει ότι έγιναν μικρά λάθη σε πολλά δεδομένα και μια μεγάλη ακρίβεια με χαμηλή απώλεια σημαίνει ότι έγιναν λίγα σφάλματα σε μερικά δεδομένα, δηλαδή η καλύτερη περίπτωση όλων.

Για τις αρχιτεκτονικές μου τώρα η διαπίστωση είναι :

1. CNN (απλό sequential model με 4 layers) : χαμηλή ακρίβεια, τεράστια απώλεια.
2. AlexNet (Deep CNN) : μέτρια ακρίβεια, μέτρια απώλεια.
3. MobileNet (Lightweight & efficient CNN) : μεγάλη ακρίβεια, πολύ μικρή απώλεια.
4. VGG16 (Very Deep and efficient net) : πολύ μεγάλη ακρίβεια, πολύ μικρή απώλεια.

Σε γενικές γραμμές, χρησιμοποίησα του tensorflow και του keras APIs built-in optimizers, loss functions και metrics όταν έκανα compile τα μοντέλα μου.

- **Optimizer** χρησιμοποίησα Adam Optimizer με $lr = 0.0001$ (learning rate)
- **Loss Function** χρησιμοποίησα Categorical_crossentropy
- **Metrics** χρησιμοποίησα Accuracy

Όσον αφορά στην εκπαίδευση των μοντέλων, για κάθε εκπαίδευση :

- Steps per epoch = 25
- Epochs = 200

Σε αυτό το σημείο θα αναλύσω τον τρόπο με τον οποίο εξηγούνται τα αποτελέσματα μου παρακάτω.

Δηλαδή τι πρεσβεύουν οι βασικές ενδείξεις και από ποιους υπολογισμούς προκύπτουν.

Confusion Matrix

Ο Confusion Matrix, είναι μια μέτρηση απόδοσης για πρόβλημα classification μηχανικής μάθησης όπου η παραγωγή μπορεί να είναι δύο ή περισσότερες τάξεις. Είναι ένας πίνακας με 4 διαφορετικούς συνδυασμούς προβλεπόμενων και πραγματικών τιμών.

		Actual Values	
		Positive (1)	Negative (0)
Predicted Values	Positive (1)	TP	FP
	Negative (0)	FN	TN

Πίνακας 4.1 Confusion Matrix

Είναι εξαιρετικά χρήσιμο για τη μέτρηση των Recall, Precision, Specificity, Accuracy[29].

- TP -True Positive
- FN-False Negative
- FP-False Positive
- TN-True Negative

Precision

Από όλες τις θετικές τάξεις που έχουμε προβλέψει σωστά, πόσες είναι πραγματικά θετικές.

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

Recall

Από όλες τις θετικές τάξεις, πόσο προβλέψαμε σωστά. Θα πρέπει να είναι όσο το δυνατόν υψηλότερο.

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

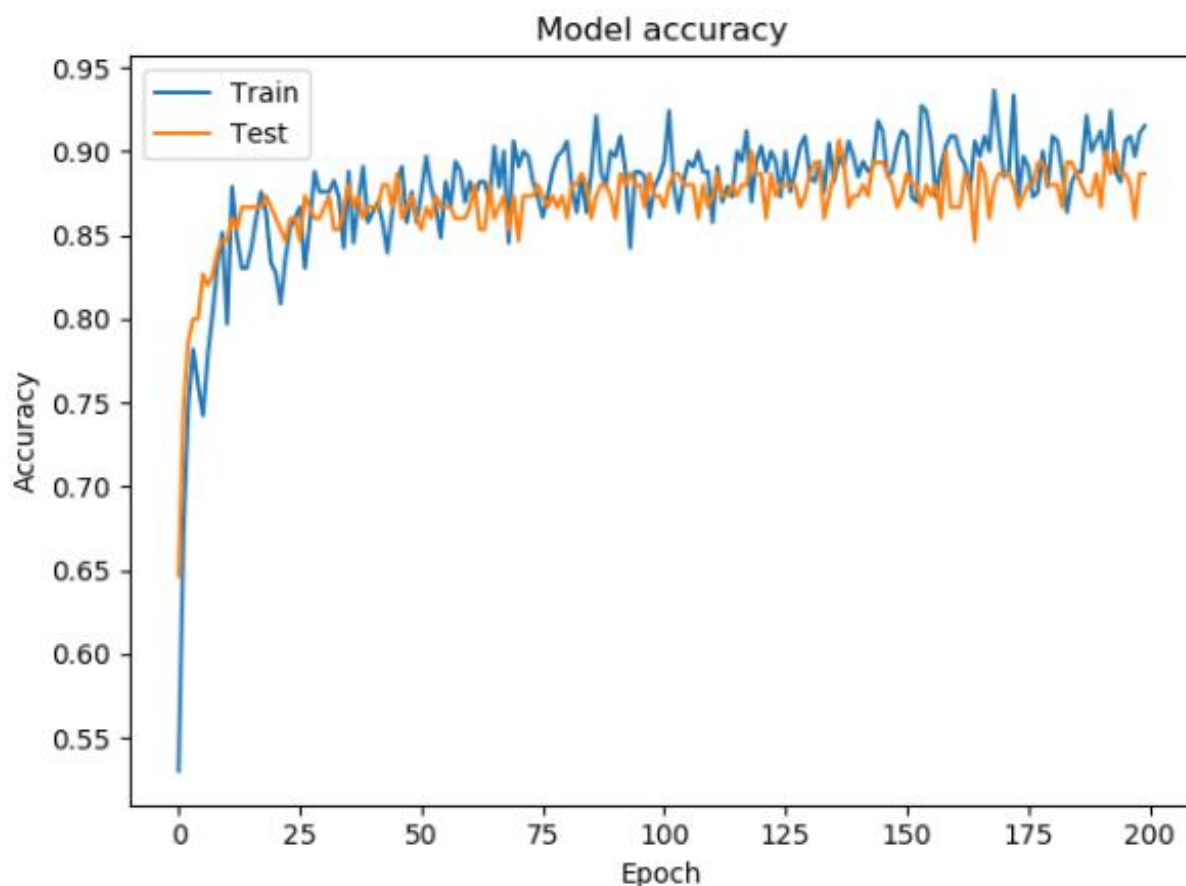
F-measure

Είναι δύσκολο να συγκρίνουμε δύο μοντέλα με χαμηλό precision και υψηλό recall ή αντίστροφα. Για να τα συγκρίνουμε, χρησιμοποιούμε το F-Score. Η βαθμολογία F βοηθά στη μέτρηση του recall και του precision ταυτόχρονα. Χρησιμοποιεί το Αρμονικό Μέσο στη θέση του Αριθμητικού Μέσου τιμωρώντας τις ακραίες τιμές περισσότερο.

$$F - \text{measure} = \frac{2 * \text{Recall} * \text{Precision}}{\text{Recall} + \text{Precision}}$$

4.1 Αποτελέσματα VGG16

4.1.1 Training & Testing Accuracy



Εικόνα 4.1 Vgg16 Model Accuracy/Epoch

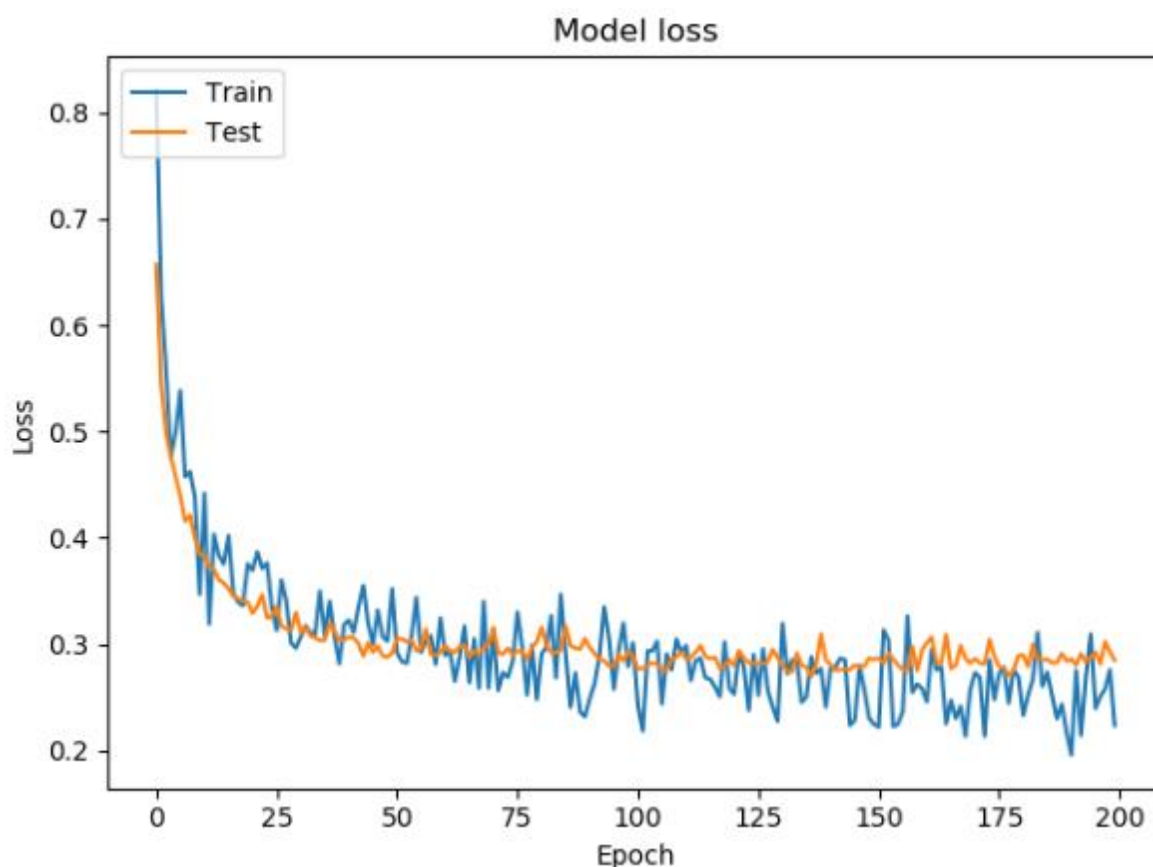
Σε αυτό το διάγραμμα φαίνεται Training Accuracy και το Testing Accuracy του VGG16, στο οποίο όπως φαίνεται σταδιακά ανεβαίνει και το training όπως και το testing και μετά από αρκετές εποχές βρίσκονται στο ίδιο επίπεδο σιγά σιγά. Αξίζει να σημειωθεί, ότι όταν φαίνεται το validation accuracy δηλαδή η ευστοχία του δικτύου κυμαίνεται στα ίδια ή κοντινά επίπεδα με το accuracy (training) τότε ξέρουμε ότι το δίκτυο επικυρώνει την εκμάθησή του σε ένα καλό ποσοστό επιτυχίας. Στην Εικόνα 4.1 φαίνεται η εξέλιξη της του training και του testing ως προς το accuracy του μοντέλου VGG16 για το σύνολο δεδομένων των υπερηχογραφήμάτων του καρκίνου του μαστού σε 200 εποχές εκπαίδευσης.

Με την εξέλιξη του χρόνου, η ακρίβεια του μοντέλου βελτιώνεται. Μπορούμε να παρατηρήσουμε ότι υπήρχαν εποχές όπου η ακρίβεια του μοντέλου είναι σταθερή και στη συνέχεια συνεχίζει να ανεβαίνει. Ωστόσο, υπάρχουν μερικές μικρές πτώσεις στην ακρίβεια σε ορισμένα στάδια της προπόνησης, αλλά ανακάμπτει καλά. Αυτές οι πτώσεις δεν είναι τόσο ακραίες σε σύγκριση με το MobileNet AlexNet.

```
234 4/4 - 9s - loss: 0.2116 - accuracy: 0.9432 - val_loss: 0.
    2226 - val_accuracy: 0.9750
235 Epoch 20/40
236 4/4 - 8s - loss: 0.2224 - accuracy: 0.9512 - val_loss: 0.
    2152 - val_accuracy: 0.9750
```

Εικόνα 4.2 Accuracy & Validation Accuracy

4.1.2 Training & Testing Loss VGG16



Εικόνα 4.2 Vgg16 Model Loss/Epoch

Αντίστοιχα, εδώ βλέπουμε το τι loss έχει το δίκτυο καθ' όλη τη διάρκεια της εκπαίδευσης καθώς και του testing. Εννοείται, πως όσο πιο καταρτισμένο γίνεται το δίκτυο νευρώνων στο να αντιλαμβάνεται τις διαφορές μεταξύ κάποιων κλάσεων, όπως στο συγκεκριμένο μεταξύ benign & malignant, τόσο λιγότερη πληροφορία θα πρέπει να χάνει με την πάροδο των εποχών εκπαίδευσης. Όπως φαίνεται στην Εικόνα 4.2 η πληροφορία που χάνεται κατά τη διάρκεια της εκπαίδευσης και του testing μειώνεται εκθετικά όσο περνάνε οι εποχές. Ειδικά παρατηρούμε στην φάση εκπαίδευσης (μπλε γραμμή) διαπιστώνεται αρκετά μεγάλη διακύμανση, ενώ στην φάση πειράματος (κίτρινη γραμμή) διαπιστώνεται μικρότερη διακύμανση. Αυτό συμβαίνει λόγω του μεγάλου όγκου πληροφορίας που διαχειρίζεται στην φάση εκπαίδευσης σε σχέση με τη φάση πειράματος, όπου τα δείγματα είναι πολύ λιγότερα.

4.1.3 Confusion Matrix & Classification Report

		Predicted	
		Benign	Malignant
Actual	Benign	50	0
	Malignant	4	72

Πίνακας 4.2 Confusion Matrix VGG16

Όσον αφορά στον Confusion Matrix της συγκεκριμένης αρχιτεκτονικής βλέπουμε ότι στις εικόνες με κακοήθεια δεν είχε κανένα false negative, ενώ στις εικόνες με καλοήθεια, είχε κάποιες τις οποίες έβγαλε λάθος αποτέλεσμα ως προς την ύπαρξη κακοήθειας (4 False Positive).

Ανάλυση εικόνας με στόχο τον εντοπισμό καρκίνου με τεχνικές μηχανικής και εις βάθος μάθησης.

Βάση υπολογισμών του συστήματος αλλά και όπως φαίνεται παρακάτω το συνολικό

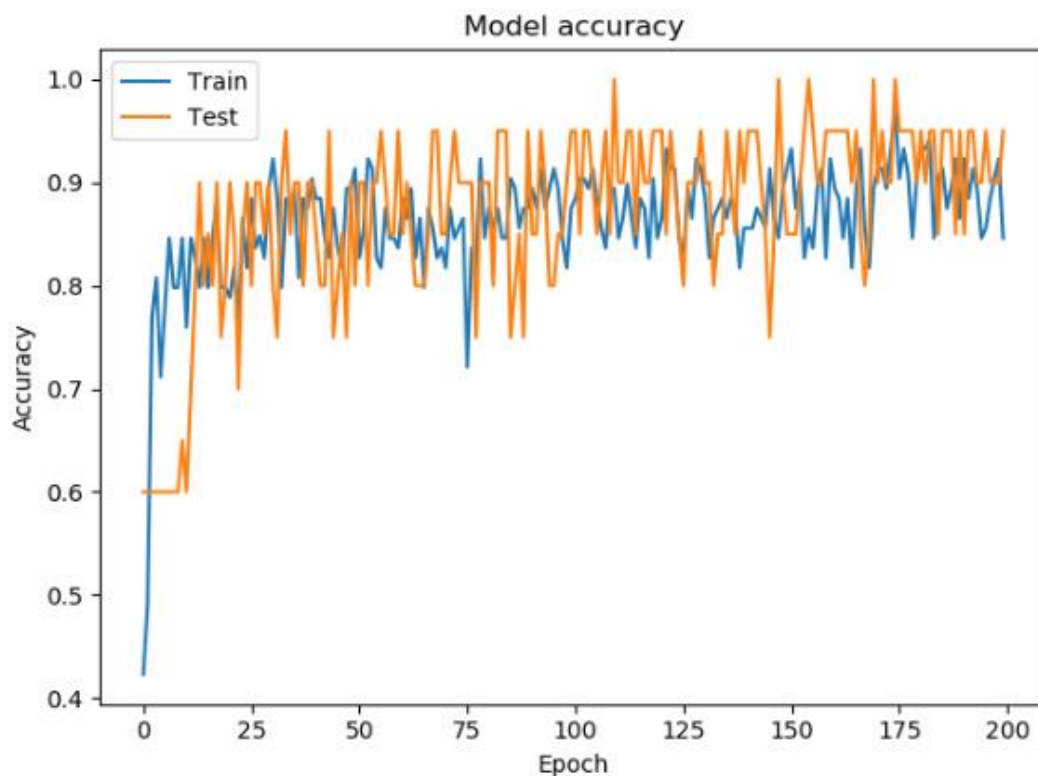
Accuracy = 0,96825397 .

VGG16	Breast Cancer Ultrasounds Dataset			
Case	<i>Precision</i>	<i>Sensitivity</i>	<i>F1-Score</i>	<i>Accuracy</i>
Benign	0.91	0.91	0.91	0.97
Malignant	0.82	0.82	0.82	

Πίνακας 4.3 Classification Report VGG16

4.2 Αποτελέσματα MobileNet

4.2.1 Training & Testing Accuracy MobileNet

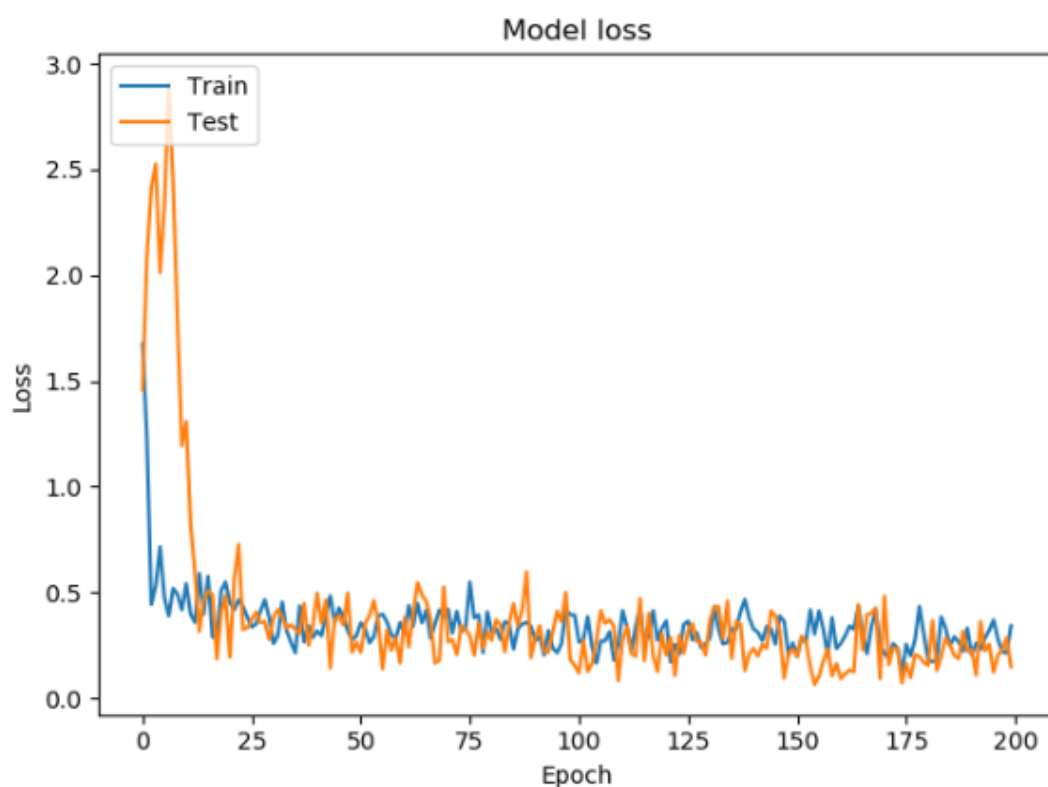


Εικόνα 4.3 MobileNet Model Accuracy/Epoch

Το εν λόγω διάγραμμα δείχνει αντίστοιχα την επίδοση του training και του testing του δικτύου MobileNet σε σχέση με της εποχές εκπαίδευσης. Όπως είναι εμφανές το MobileNet από την πολύ αρχή τα επίπεδα training είναι τέλεια, ενώ το testing του αναπτύσσεται σε σχέση με την πάροδο του χρόνου μέχρι που φτάνει σε πολύ καλά αποτελέσματα. Στην Εικόνα 4.3 φαίνεται η εξέλιξη του training και testing accuracy του μοντέλου MobileNet για σύνολο δεδομένων των υπερήχων του καρκίνου του μαστού σε 200 εποχές εκπαίδευσης. Με την πάροδο του χρόνου η ακρίβεια του μοντέλου βελτιώνεται, εν τω μεταξύ μπορούμε να εντοπίσουμε μερικές πτώσεις στην ακρίβεια ειδικότερα στην αρχή μεταξύ

των εποχών 15 -25 και μετά επανέρχεται σε υψηλά επίπεδα κατάρτισης. Ωστόσο, αυτές οι πτώσεις δεν είναι τόσο ακραίες σε σύγκριση με την απόδοση του AlexNet.

4.2.2 Training & Testing Loss MobileNet



Εικόνα 4.4 MobileNet Model Loss/Epoch

Επίσης το Loss του MobileNet και πάλι ως προς το training και testing δείχνει τα ανάλογα αποτελέσματα. Δηλαδή στο training από την αρχή εκπαιδεύεται πολύ γρήγορα και επίπεδα του loss πέφτουν χαμηλά από τις πρώτες ήδη εποχές εκπαίδευσης, ενώ στο testing πέφτουν λιγότερο απότομα.

4.2.3 Confusion Matrix & Classification Report

		Predicted	
		Benign	Malignant
Actual	Benign	50	0
	Malignant	16	60

Πίνακας 4.4 Confusion Matrix MobileNet

Όσον αφορά στον Confusion Matrix της συγκεκριμένης αρχιτεκτονικής βλέπουμε ότι στις εικόνες με κακοήθεια δεν είχε κανένα False negative, ενώ στις εικόνες με καλοήθεια, είχε κάποιες τις οποίες έβγαλε λάθος αποτέλεσμα ως προς την ύπαρξη κακοήθειας (16 False Positive).

Βάση υπολογισμών του συστήματος αλλά και όπως φαίνεται παρακάτω το συνολικό

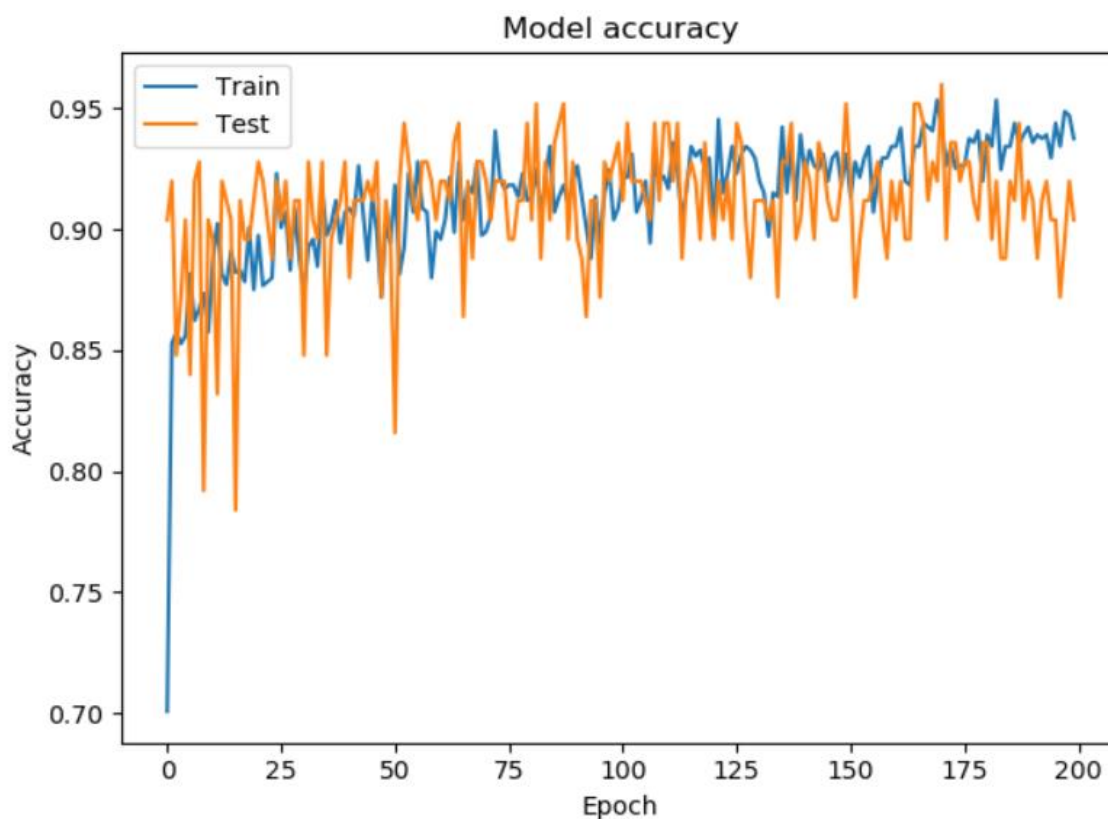
Accuracy = 0.87301587 .

Classification Report				
	precision	recall	f1-score	support
Benign	0.76	1.00	0.86	50
Malignant	1.00	0.79	0.88	76
accuracy			0.87	126
macro avg	0.88	0.89	0.87	126
weighted avg	0.90	0.87	0.87	126

Πίνακας 4.5 MobileNet Classification Report

4.3 Αποτελέσματα AlexNet

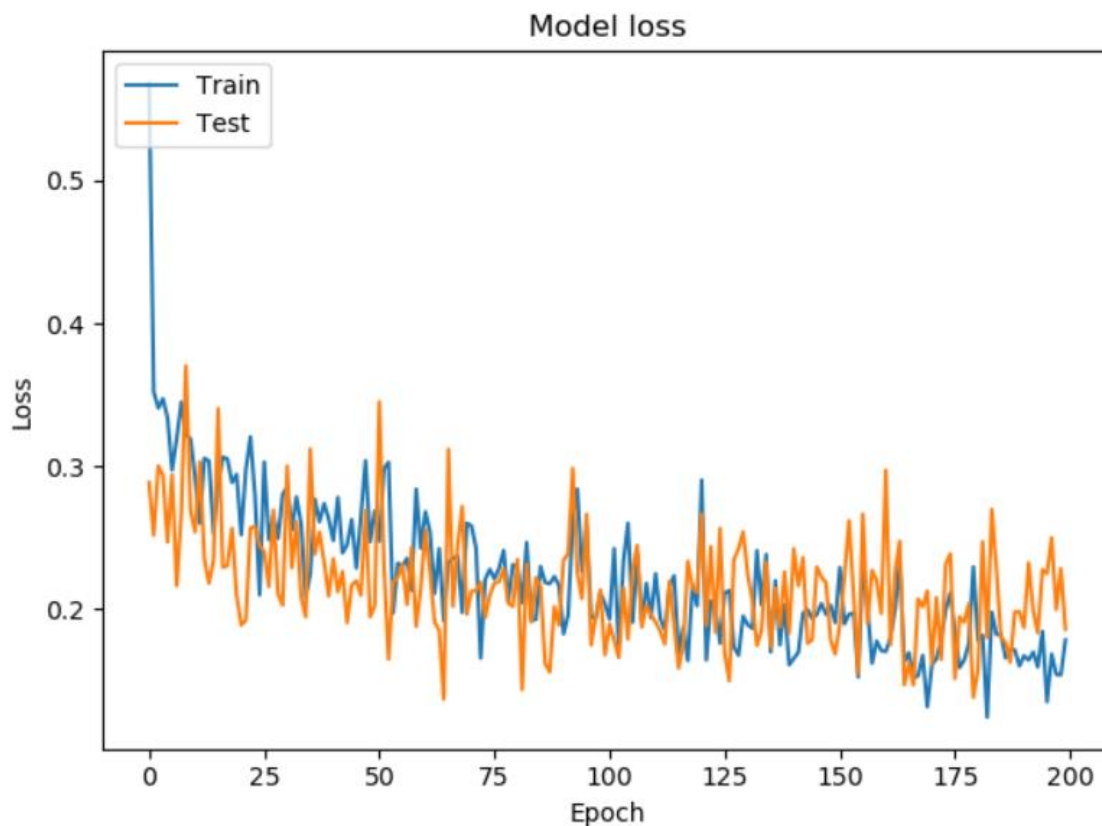
4.3.1 Training & Testing Accuracy AlexNet



Εικόνα 4.5 AlexNet Model Accuracy

Στην εικόνα 4.5 φαίνεται η εξέλιξη του training και του testing του μοντέλου AlexNet για το σύνολο δεδομένων από υπερηχογραφήματα μαστού σε γυναίκες με περιπτώσεις benign & malignant tumors σε 200 εποχές εκπαίδευσης. Με την εξέλιξη του χρόνου, η ακρίβεια του μοντέλου βελτιώνεται, εν τω μεταξύ μπορούμε να εντοπίσουμε κάποιες ξαφνικές πτώσεις στην ακρίβεια και αυτό συμβαίνει λόγω του ρυθμού εκμάθησης που πιθανότατα είναι 0,0001.

4.3.2 Training & Testing Loss AlexNet



Εικόνα 4.6 AlexNet Model Loss

Στην εικόνα 4.6 φαίνεται το Loss που έχει το μοντέλο κατά τη διάρκεια του training και του testing του μοντέλου AlexNet για το σύνολο δεδομένων των υπερηχογραφήματων του μαστού σε γυναίκες σε περιπτώσεις με benign & malignant tumors σε 200 εποχές εκπαίδευσης. Καθώς η περίοδος εκπαίδευσης εξελίσσεται, το δίκτυο χάνει όλο και λιγότερες πληροφορίες και ταυτόχρονα βελτιώνεται η ακρίβεια με τις εποχές εκπαίδευσης. Το ποσοστό εκμάθησης είναι 0,0001 και τα βήματα εκπαίδευσης ανά εποχή ορίζονται στα 25.

4.3.3 Confusion Matrix & Classification Report

Το AlexNet είναι από τα πρώτα βαθιά νευρωνικά δίκτυα το οποίο διακρίνεται από μια απλή σχετικά sequential αρχιτεκτονική και τα αποτελέσματά του δείχνουν πολύ μεγάλο επίπεδο κατάρτισης ανεξάρτητα από την απλότητα της δομής του και την ταχύτητα της εκπαίδευσής του. Πολύ εύκολα παρατηρείται σε αυτή τη φάση ότι ακόμα και με τις πιο απλές προσεγγίσεις και αρχιτεκτονικές συστημάτων μπορούν να προσεγγιστούν οι λύσεις σε διαφόρων ειδών δύσκολα classification προβλήματα.

Accuracy = 0.823257

AlexNet	Breast Cancer Ultrasounds Dataset			
Case	Precision	Sensitivity	F1-Score	Accuracy
Benign	0.75	0.84	0.79	0.82
Malignant	0.88	0.81	0.85	

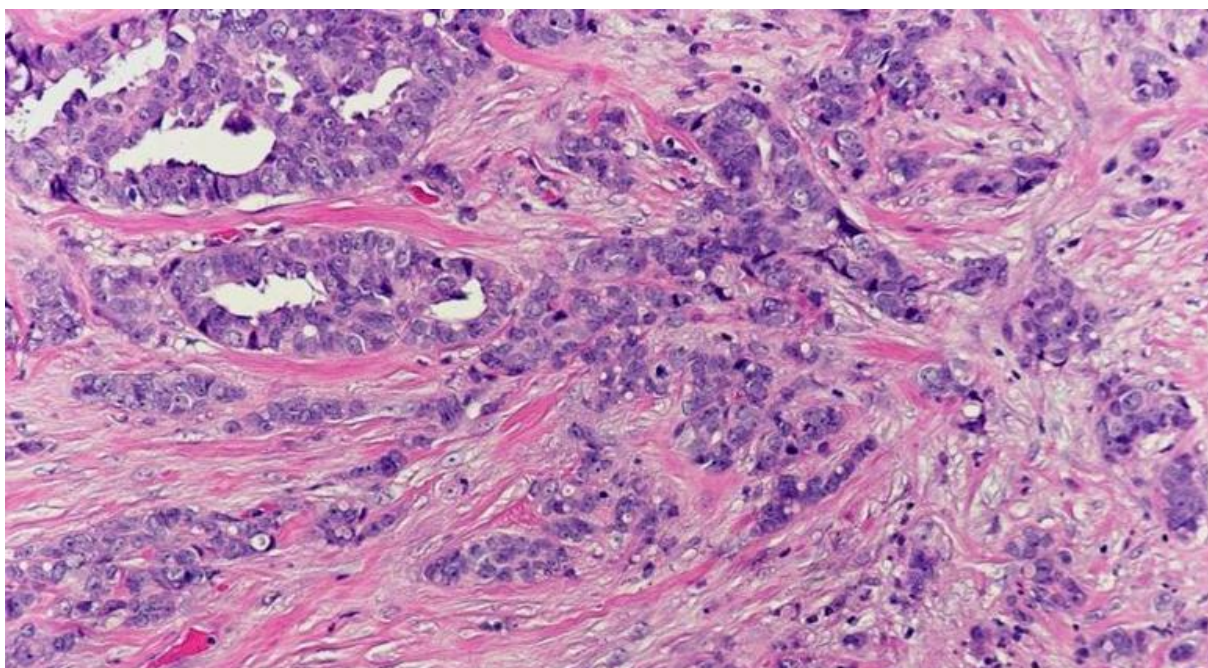
Πίνακας 4.6 AlexNet Classification Report

4.4 Πειράματα για Invasive Ductal Carcinoma (IDC)

Για να εξετάσω και περεταίρω τη δυναμική των αρχιτεκτονικών αυτών συνέλεξα επιπλέον πειραματικά δεδομένα τα οποία είναι και πιο δύσκολα ως προς το κομμάτι της εκπαίδευσης, αλλά και πιο πολλά σε όγκο. Πιο συγκεκριμένα αναφέρομαι σε δύο datasets τα οποία αποτελούνται από εικόνες που αφορούν σε μια υποπερίπτωση στου καρκίνου του μαστού (Invasive Ductal Carcinoma IDC), η οποία μάλιστα είναι και η πιο συχνή περίπτωση σε ανθρώπους που πάσχουν από μια τέτοια ασθένεια.

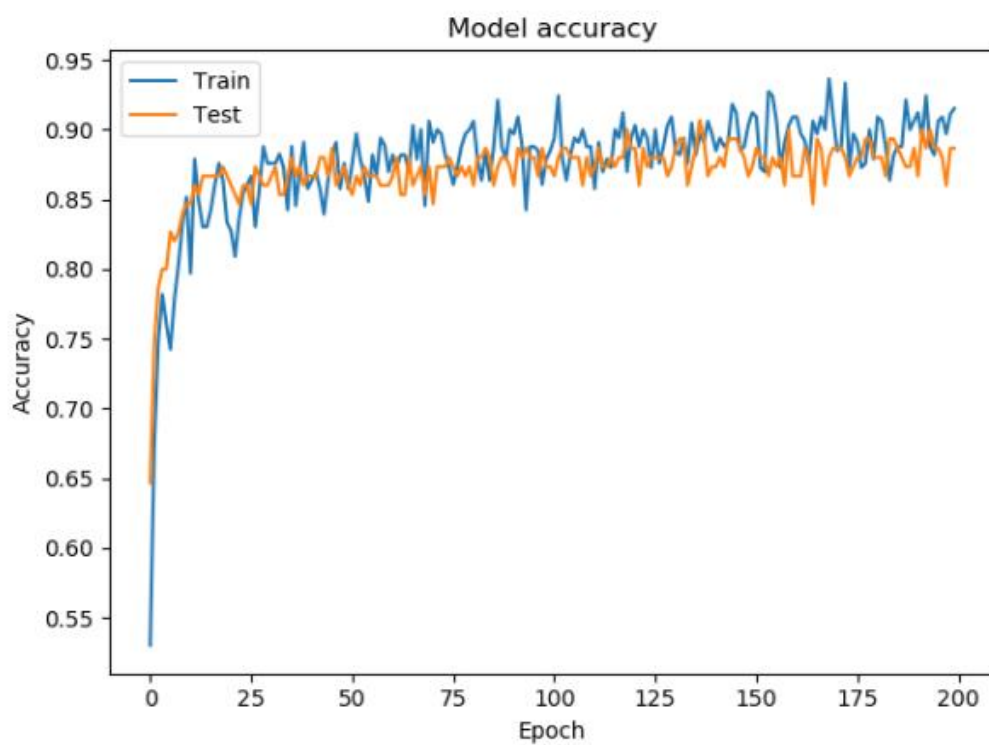
Όσον αφορά τώρα στην περιγραφή των νέων δεδομένων, το Invasive Ductal Carcinoma (IDC) είναι ο πιο κοινός τύπος καρκίνου από όλους τους καρκίνους του μαστού. Για να αναγνωριστεί ο βαθμός επιθετικότητας σε ένα δείγμα ολόκληρης εικόνας, οι παθολόγοι συνήθως εστιάζουν στις περιοχές που περιέχουν το IDC. Ως αποτέλεσμα, ένα από τα συνηθισμένα βήματα προεπεξεργασίας για την αυτόματη καταγραφή επιθετικότητας είναι να οριοθετηθούν οι ακριβείς περιοχές του IDC στο εσωτερικό μιας ολόκληρης διαφάνειας. Οπότε έχουμε δείγματα δύο κατηγοριών εκ των οποίων η μία είναι αυτή όπου ο ιστός δεν έχει δείγματα καρκίνου, ενώ στην άλλη υπάρχουν σημεία με ανεπτυγμένα καρκινικά κύτταρα.

Το ένα dataset με Breast Cancer Histopathological Images περιέχει 3628 εικόνες, ενώ στο δεύτερο dataset χρησιμοποιήθηκαν 10483 εικόνες. Για κάθε dataset καταγράφηκαν τα αποτελέσματα εκπαίδευσης αλλά και προβλέψεων για τα τρία νευρωνικά δίκτυα Vgg16, MobileNet και AlexNet. Στην πρώτη περίπτωση τα καλύτερα αποτελέσματα τα είχε το Vgg16 με accuracy 88% στις προβλέψεις και μετά το MobileNet με 84% accuracy, ενώ το AlexNet δεν μπόρεσε να ανταποκριθεί επάξια όντας μικρό και πιο απλή αρχιτεκτονική του οποίου οι προβλέψεις χαρακτηρίζονται και πάλι από μια τυχειότητα. Στη δεύτερη περίπτωση την καλύτερη επίδοση την είχε το MobileNet με απόδοση 84% και μετά το Vgg16 με accuracy 82%.



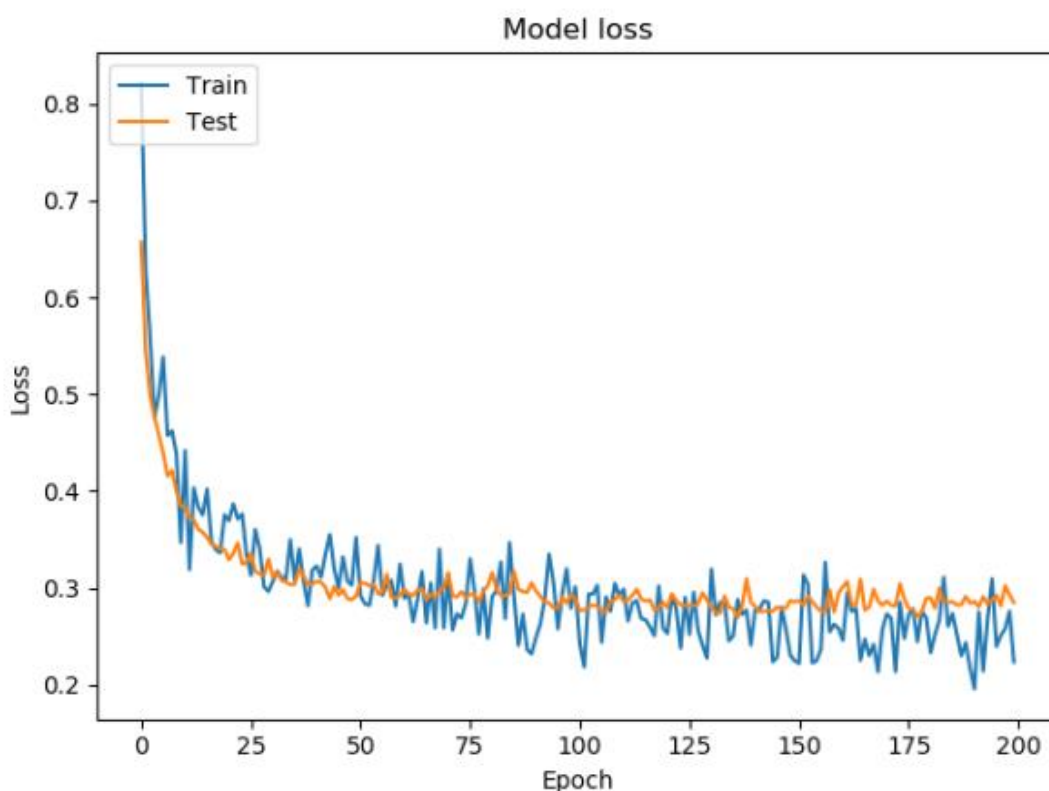
Εικόνα 4.7 Invasive Ductal Carcinoma (IDC)

4.4.1 Vgg16 in Histopathology Images



Εικόνα 4.8 VGG16 Training & Testing Accuracy in IDC Dataset

Στην Εικόνα 4.8 φαίνεται πως εξελίσσεται η διαδικασία εκπαίδευσης του μοντέλου VGG16 για το 2ο σύνολο δεδομένων του IDC καρκίνου σε 200 εποχές εκπαίδευσης. Με την εξέλιξη του χρόνου, η ακρίβεια του μοντέλου βελτιώνεται. Μπορούμε να παρατηρήσουμε ότι υπήρχαν εποχές όπου η ακρίβεια του μοντέλου είναι κάπως πιο σταθερή και στη συνέχεια συνεχίζει να ανεβαίνει. Ωστόσο, υπάρχουν μερικές μικρές πτώσεις στην ακρίβεια σε ορισμένα στάδια της εκπαίδευσης, αλλά ανακάμπτει καλά. Αυτές οι πτώσεις δεν είναι τόσο ακραίες σε σύγκριση με το MobileNet, AlexNet.



Εικόνα 4.10 VGG16 Training and Testing Loss in IDC Dataset

Στην εικόνα 4.10 φαίνεται το validation loss του μοντέλου VGG16 για το 2ο σύνολο δεδομένων του επεμβατικού καρκίνου του πόρου σε 200 εποχές εκπαίδευσης. Καθώς η περίοδος εκπαίδευσης εξελίσσεται, το δίκτυο χάνει όλο και λιγότερες πληροφορίες και ταυτόχρονα βελτιώνεται η ακρίβεια με τις εποχές εκπαίδευσης. Το ποσοστό εκμάθησης είναι 0,0001 και τα βήματα εκπαίδευσης ανά εποχή ορίζονται σε 25. Είναι προφανές ότι υπάρχουν αιχμές απώλειας πληροφοριών μεταξύ των εποχών 50-

75, οι οποίες ανακτήθηκαν κατά τη διάρκεια της εκπαίδευσης και η απώλεια συνεχίζει να μειώνεται σταδιακά λόγω της προεπεξεργασίας των εικόνων.

VGG16	1 st IDC (Invasive Ductal Carcinoma) Dataset			
Case	<i>Precision</i>	<i>Sensitivity</i>	<i>F1-Score</i>	<i>Accuracy</i>
Non-Cancerous	0.90	0.91	0.90	0.88
Cancerous	0.84	0.82	0.83	

Πίνακας 4.7 Classification Report 1st Dataset Vgg16

		Predicted	
		Cancer	No Cancer
Actual	Cancer	888	90
	No Cancer	87	409

Πίνακας 4.8 Confusion Matrix Vgg16 1st Dataset

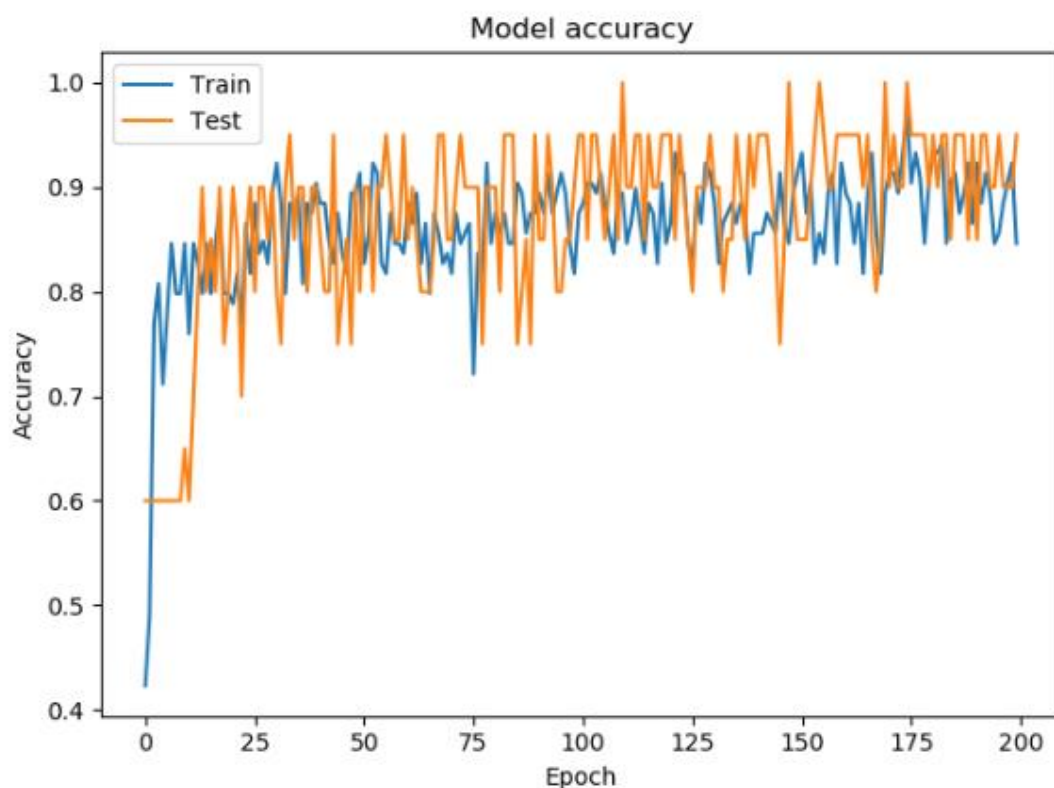
VGG16	2 nd IDC (Invasive Ductal Carcinoma) Dataset			
Case	<i>Precision</i>	<i>Sensitivity</i>	<i>F1-Score</i>	<i>Accuracy</i>
Non-Cancerous	0.95	0.83	0.89	0.86
Cancerous	0.74	0.92	0.82	

Πίνακας 4.9 Classification Report 2nd Dataset Vgg16

		Predicted	
		Cancer	No Cancer
Actual	Cancer	2814	352
	No Cancer	704	2050

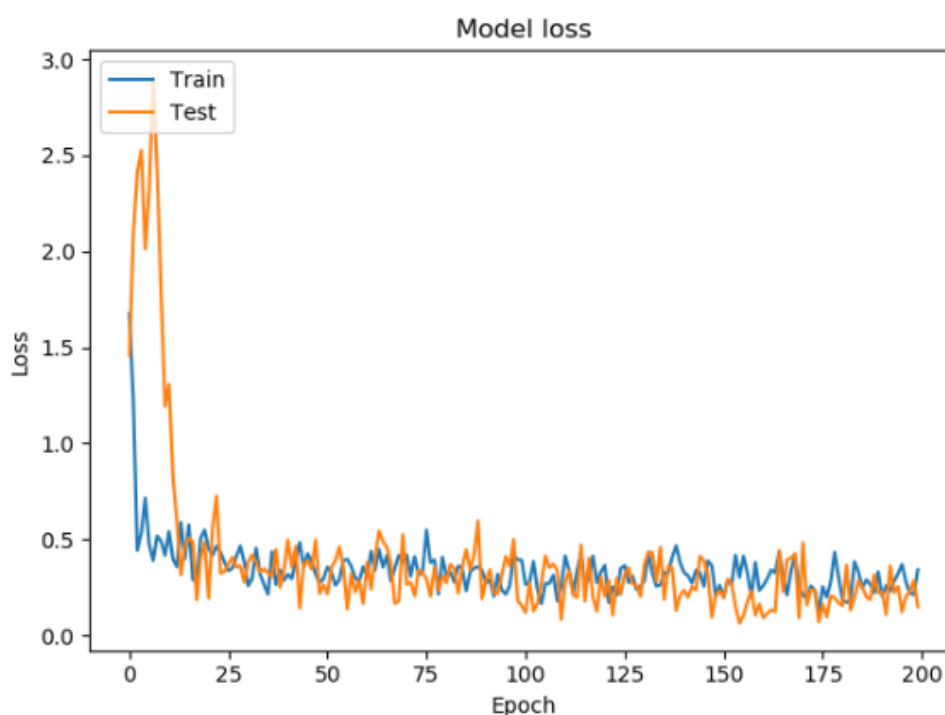
Πίνακας 4.10 Confusion Matrix Vgg16 2nd Dataset

4.4.2 MobileNet in Histopathology Images



Εικόνα 4.12 MobileNet Training & Testing Accuracy

Στην Εικόνα 4.12 φαίνεται η πρόοδος του training και του testing του μοντέλου MobileNet για 2ο σύνολο δεδομένων του IDC καρκίνου του μαστού σε 200 εποχές εκπαίδευσης. Με την πάροδο του χρόνου η ακρίβεια του μοντέλου βελτιώνεται, εν τω μεταξύ μπορούμε να εντοπίσουμε μερικές απότομες πτώσεις στην ακρίβεια. Ωστόσο, αυτές οι πτώσεις δεν είναι τόσο ακραίες σε σύγκριση με την απόδοση του AlexNet.



Εικόνα 4.13 MobileNet Model Loss Training & Testing

Στην Εικόνα 4.13 παρατηρούμε την εξέλιξη της πληροφορίας που χάνεται κατά την διάρκεια training και testing του μοντέλου MobileNet για το 2ο σύνολο δεδομένων του επεμβατικού καρκίνου του πόρου σε 200 εποχές εκπαίδευσης. Καθώς η περίοδος εκπαίδευσης εξελίσσεται, το δίκτυο χάνει όλο και λιγότερες πληροφορίες και ταυτόχρονα βελτιώνεται η ακρίβεια με τις εποχές εκπαίδευσης. Το ποσοστό εκμάθησης είναι 0,0001 και τα βήματα εκπαίδευσης ανά εποχή ορίζονται σε 25. Είναι προφανές ότι υπάρχει μια αύξηση στην απώλεια πληροφοριών μεταξύ των εποχών 0-25, η οποία καλύφθηκε κατά τη διάρκεια της προπόνησης και η απώλεια συνεχίζει να μειώνεται σταδιακά λόγω της προεπεξεργασία των εικόνων.

MobileNet	1 st IDC (Invasive Ductal Carcinoma) Dataset			
Case	<i>Precision</i>	<i>Sensitivity</i>	<i>F1-Score</i>	<i>Accuracy</i>
Non-Cancerous	1.00	0.76	0.86	0.84
Cancerous	0.68	1.00	0.81	

Πίνακας 4.11 Classification Report 1st Dataset MobileNet

		Predicted	
		Cancer	No Cancer
Actual	Cancer	741	237
	No Cancer	1	495

Πίνακας 4.12 Confusion Matrix MobileNet 1st Dataset

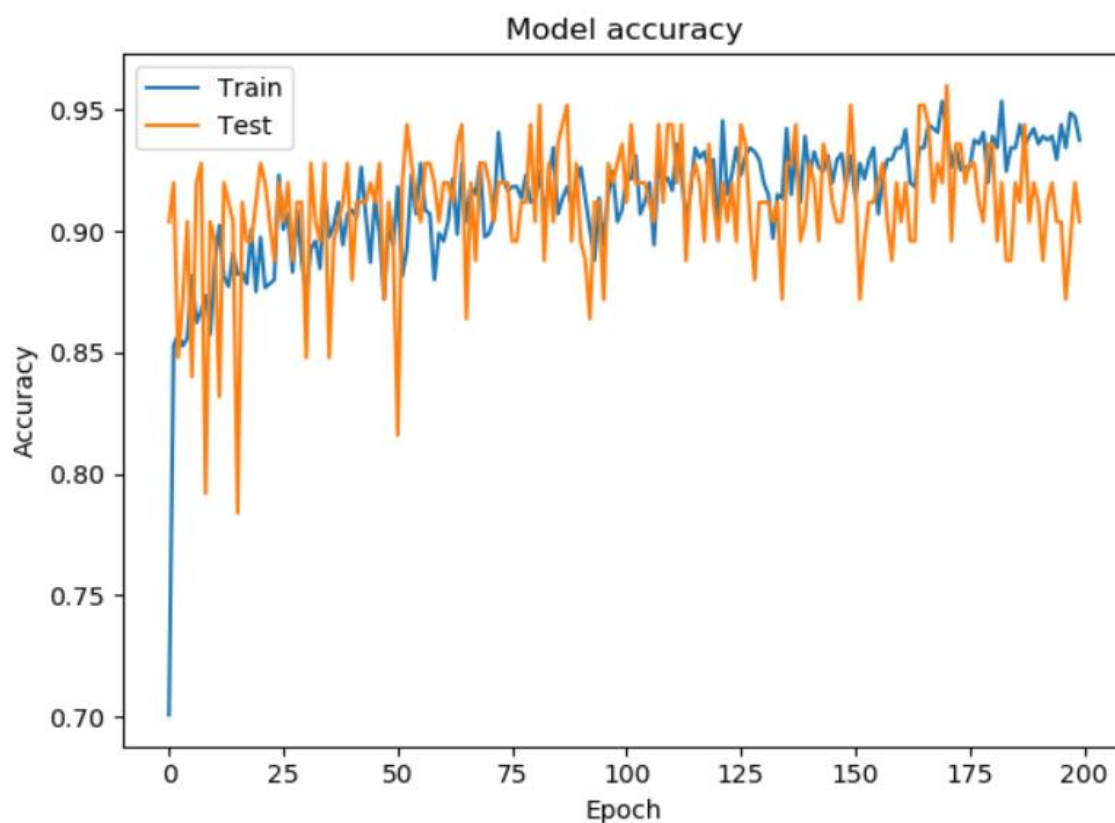
MobileNet	2 nd IDC (Invasive Ductal Carcinoma) Dataset			
Case	<i>Precision</i>	<i>Sensitivity</i>	<i>F1-Score</i>	<i>Accuracy</i>
Non-Cancerous	0.98	0.89	0.93	0.91
Cancerous	0.81	0.96	0.88	

Πίνακας 4.13 Classification Report 2nd Dataset MobileNet

		Predicted	
		Cancer	No Cancer
Actual	Cancer	2804	362
	No Cancer	590	2164

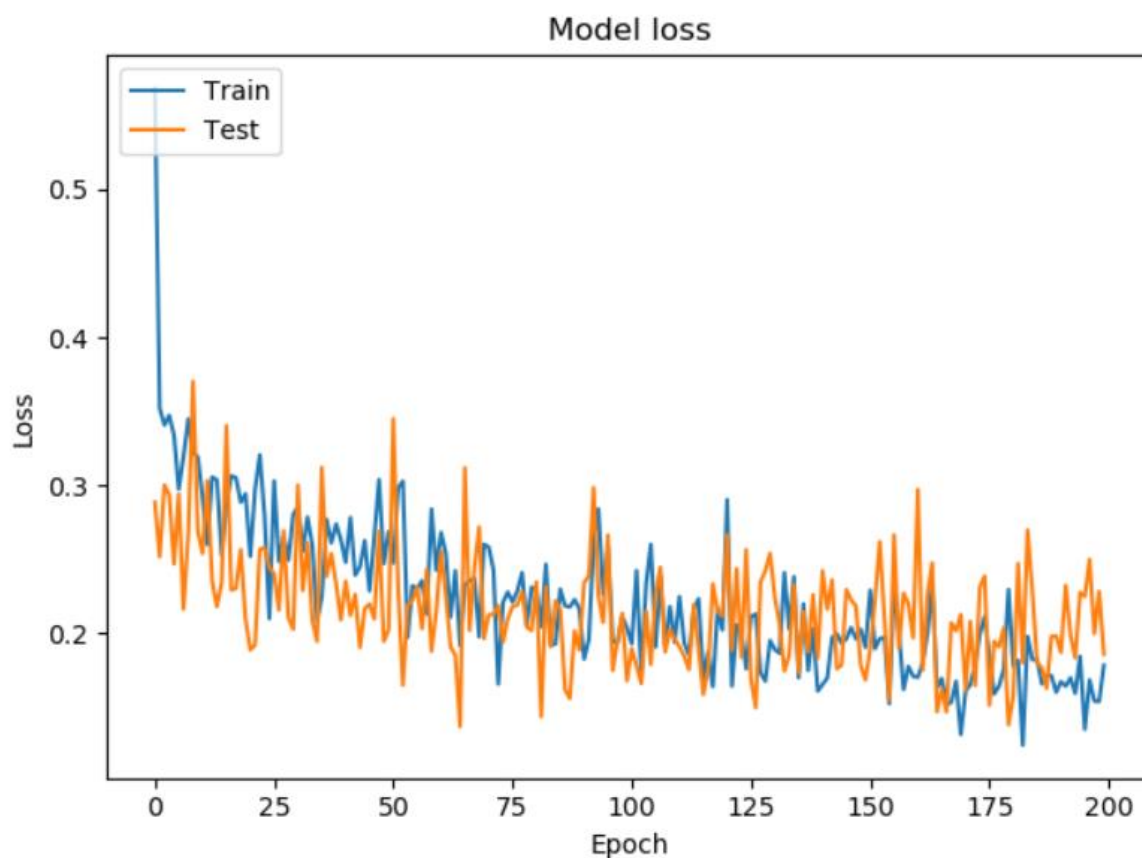
Πίνακας 4.14 Confusion Matrix MobileNet 2nd Dataset

4.4.3 AlexNet in Histopathology Images



Εικόνα 4.11 AlexNet Model Accuracy Training & Testing

Στην Εικόνα 4.11 παρατηρούμε την εξέλιξη της μοντέλου AlexNet για το 2ο σύνολο δεδομένων του IDC καρκίνου του μαστού σε 200 εποχές εκπαίδευσης. Με την εξέλιξη του χρόνου, η ακρίβεια του μοντέλου βελτιώνεται, εν τω μεταξύ μπορούμε να εντοπίσουμε κάποιες ξαφνικές πτώσεις στην ακρίβεια μεταξύ των εποχών 20-22 και 50-53 και αυτό συμβαίνει λόγω του ρυθμού εκμάθησης πιθανότατα ο οποίος είναι 0,0001.



Εικόνα 4.12 AlexNet Model Accuracy Training & Testing

Στην Εικόνα 4.11 παρουσιάζεται η πορεία του training και testing loss του μοντέλου AlexNet για το 2ο σύνολο δεδομένων του IDC καρκίνου σε 200 εποχές εκπαίδευσης. Καθώς η περίοδος εκπαίδευσης εξελίσσεται, το δίκτυο χάνει όλο και λιγότερες πληροφορίες και ταυτόχρονα βελτιώνεται η ακρίβεια με τις εποχές εκπαίδευσης. Το ποσοστό εκμάθησης είναι 0,0001 και τα βήματα εκπαίδευσης ανά εποχή ορίζονται στα 25.

AlexNet	1st IDC (Invasive Ductal Carcinoma) Dataset			
Case	<i>Precision</i>	<i>Sensitivity</i>	<i>F1-Score</i>	<i>Accuracy</i>
Non-Cancerous	0.90	0.83	0.86	0.82
Cancerous	0.70	0.83	0.76	

Πίνακας 4.15 Classification Report 1st Dataset MobileNet

		Predicted	
		Cancer	No Cancer
Actual	Cancer	978	0
	No Cancer	496	0

Πίνακας 4.16 AlexNet Confusion Matrix 1st Dataset

AlexNet	2nd IDC (Invasive Ductal Carcinoma) Dataset			
Case	<i>Precision</i>	<i>Sensitivity</i>	<i>F1-Score</i>	<i>Accuracy</i>
Non-Cancerous	0.95	0.86	0.90	0.88
Cancerous	0.77	0.91	0.83	

Πίνακας 4.17 AlexNet Classification Report 2nd Dataset

5

Συμπεράσματα & Μελλοντικές Υποθέσεις

Στον σύγχρονο κόσμο, εξακολουθούμε να αντιμετωπίζουμε συνεχώς κάποιες ανεπίλυτες περιπτώσεις ασθενειών που απειλούν την ύπαρξη ορισμένων ανθρώπων και ενώ έχουν γίνει πολλές προσπάθειες από την επιστημονική κοινότητα για τη βελτίωση αυτής της κατάστασης κάθε έρευνα και κάθε νέα μελέτη μπορεί να είναι πολύτιμη για τα επόμενα χρόνια. Σε αυτήν τη μελέτη, προσπαθώ να δώσω την οπτική μου για το πώς θα μπορούσαμε να αντιμετωπίσουμε στο εγγύς μέλλον ένα σοβαρό πρόβλημα όπως ο καρκίνος του μαστού ειδικά σε περιπτώσεις όπου υπάρχει περιορισμός στους υπολογιστικούς πόρους και πώς να συνεισφέρουμε σε αυτό το τεράστιο πεδίο έρευνας. Η απόδοση του Deep Learning στην αναγνώριση αντικειμένων και στην ταξινόμηση εικόνας έχει σημειώσει τεράστια πρόοδο τα τελευταία χρόνια. Πρόσφατα, πολλές μελέτες βασίστηκαν στα CNN για εξέταση ιατρικής απεικόνισης [20], όπως η ανίχνευση καρκίνου του μαστού [21], αναλύοντας ιατρικές εικόνες και έχουν

επιτύχει πολύ ενδιαφέροντα και πολλά υποσχόμενα αποτελέσματα [22]. Σε αυτήν τη μελέτη, παρουσιάζω και αξιολογώ τις επιδόσεις διαφόρων αρχιτεκτονικών CNN σε πολλά σύνολα δεδομένων διαφορετικών περιπτώσεων εξέτασης καρκίνου του μαστού. Συγκεκριμένα, σύγκρινα την απόδοση ενός depth-wise και low computational cost συνελικτικού νευρικού δικτύου με κάποια βαθιά συνελικτικά δίκτυα και ταυτόχρονα αξιολόγησα τη συνολική απόκρισή τους σε αυτά τα διαφορετικά ιατρικά σύνολα δεδομένων. Το πρώτο σύνολο δεδομένων είναι ένα σύνολο δεδομένων για τον καρκίνο του μαστού υπερήχων όπου εκπαιδευσα όλα τα δίκτυα στο να μπορούν να διαφοροποιούν καλοήθεις όγκους από κακοήθεις και το καλύτερο αποτέλεσμα, που έλαβα ήταν ακρίβεια 96,82%. Το δεύτερο σύνολο δεδομένων ήταν περιοχές ενδιαφέροντος από εικόνες ιστολογίας με περιπτώσεις διεισδυτικού καρκίνου του πόρου (Invasive Ductal Carcinoma), ο πιο κοινός τύπος καρκίνου του μαστού και το καλύτερο αποτέλεσμα που λάβαμε ήταν 88,23%. Το τρίτο σύνολο δεδομένων ήταν και πάλι ιστολογικές εικόνες διεισδυτικού καρκίνου του πόρου με μεγαλύτερου όγκου δεδομένα, όπου το καλύτερο αποτέλεσμα ήταν ακρίβεια 91,04%. Η ουσία είναι ότι υπάρχει τεράστιο δυναμικό για τα CNN να φτάσουν στο επίπεδο εκπαίδευσης για να αξιολογήσουν διεξοδικά τις περιπτώσεις καρκίνου του μαστού τόσο σε υπερήχους όσο και σε εικόνες ιστολογίας. Επιπλέον, αυτό που κατέληξα επίσης ήταν ότι όχι μόνο οι βαριές αρχιτεκτονικές μπορούν να επιτύχουν εξαιρετικά αποτελέσματα, αλλά και έξυπνη αρχιτεκτονική όπως το MobileNet που είναι πολύ ελαφρύ με πολύ μικρό κόστος στο accuracy και στο sensitivity. Ωστόσο, απέδειξα επίσης ότι είναι επιτακτική ανάγκη να αντιμετωπίσουμε αποτελεσματικά τέτοια προβλήματα classification και ότι κάθε λεπτομέρεια μπορεί να κάνει τεράστια διαφορά σε αυτόν τον τομέα της έρευνας. Τέλος, τα αποτελέσματα αυτής της σύγκρισης έδειξαν την πιθανή αποτελεσματικότητα που μπορεί να επιτευχθεί η διάγνωση με τη βοήθεια υπολογιστών, με χαμηλό επίπεδο υπολογιστικού κόστους, βάσει depth-wise convolutional neural networks. Τέτοια οικονομικά αποδοτικά συστήματα έχουν τη δυνατότητα να ενσωματωθούν σε περιβάλλοντα όπου υπάρχουν υπολογιστικοί περιορισμοί και πόροι.

References

- [1]. Rish, Irina. "An empirical study of the naive Bayes classifier." IJCAI 2001 workshop on empirical methods in artificial intelligence. Vol. 3. No. 22. 2001.
- [2]. Ben-Hur, A., Horn, D., Siegelmann, H. T., & Vapnik, V. (2001). Support vector clustering. *Journal of machine learning research*, 2(Dec), 125-137.
- [3]. Alpaydin, Ethem. *Introduction to machine learning*. MIT press, 2020.
- [4]. Bishop, C. M. (2006), *Pattern Recognition and Machine Learning*, Springer, ISBN 978-0-387-31073-2
- [5]. The definition "without being explicitly programmed" is often attributed to Arthur Samuel, who coined the term "machine learning" in 1959, but the phrase is not found verbatim in this publication, and may be a paraphrase that appeared later. Confer "Paraphrasing Arthur Samuel (1959), the question is: How can computers learn to solve problems without being explicitly programmed?" in Koza, John R.; Bennett, Forrest H.; Andre, David; Keane, Martin A. (1996). *Automated Design of Both the Topology and Sizing of Analog Electrical Circuits Using Genetic Programming. Artificial Intelligence in Design '96*. Springer, Dordrecht. pp. 151–170. doi:10.1007/978-94-009-0279-4_9.
- [6]. Roman, Victor (2019-04-21). "Unsupervised Machine Learning: Clustering Analysis". Medium. Retrieved 2019-10-01
- [7]. Bengio, Y.; Courville, A.; Vincent, P. (2013). "Representation Learning: A Review and New Perspectives". *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 35 (8): 1798–1828. arXiv:1206.5538. doi:10.1109/tpami.2013.50. PMID 23787338.

- [8]. Schmidhuber, J. (2015). "Deep Learning in Neural Networks: An Overview". *Neural Networks*. 61: 85–117. arXiv:1404.7828. doi:10.1016/j.neunet.2014.09.003. PMID 25462637.
- [9]. Ciresan, D.; Meier, U.; Schmidhuber, J. (2012). "Multi-column deep neural networks for image classification". 2012 IEEE Conference on Computer Vision and Pattern Recognition. pp. 3642–3649. arXiv:1202.2745. doi:10.1109/cvpr.2012.6248110. ISBN 978-1-4673-1228-8.
- [10]. Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.
- [11]. P. J. Werbos. *Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences*. PhD thesis, Harvard University, 1974.
- [12]. B. A. Olshausen and D. J. Field, "Emergence of simple-cell receptive field properties by learning a sparse code for natural images," *Nature*, vol. 381, no. 6583, pp. 607–609, 1996.
- [13]. Chen, Yung-Yao; Lin, Yu-Hsiu; Kung, Chia-Ching; Chung, Ming-Han; Yen, I.-Hsuan (January 2019). "Design and Implementation of Cloud Analytics-Assisted Smart Power Meters Considering Advanced Artificial Intelligence as Edge Analytics in Demand-Side Management for Smart Homes". *Sensors*. 19 (9): 2047. doi:10.3390/s19092047. PMC 6539684. PMID 31052502.
- [14]. Bethge, Matthias; Ecker, Alexander S.; Gatys, Leon A. (26 August 2015). "A Neural Algorithm of Artistic Style". arXiv:1508.06576 [cs.CV].
- [15]. S. Shi, Q. Wang, P. Xu and X. Chu, "Benchmarking State-of-the-Art Deep Learning Software Tools," 2016 7th International Conference on Cloud Computing and Big Data (CCBD), Macau, 2016, pp. 99-104.

- [16]. Shapiro, Linda G., and George C. Stockman. Computer vision. Prentice Hall, 2001.
- [17]. Moeskops P, Viergever MA, Mendrik AM, de Vries LS, Benders M, Išgum I. Automatic segmentation of MR brain images with a convolutional neural network. *IEEE Transactions on Medical Imaging*. 2016;35:1252–1261.
- [18]. Mao J, Xu W, Yang Y, Wang J, Huang Z, Yuille A. Explain images with multimodal recurrent neural networks. *ArXiv.org Web site*. [Accessed April 1, 2017]. <https://arxiv.org/abs/1410.1090>.
- [19]. Socher R, Karpathy A, Le QV, Manning CD, Ng AY. Grounded compositional semantics for finding and describing images with sentences. *Trans Assoc Comput Linguist*. 2014;2:207–218.
- [20]. Spanhol, F. A., Oliveira, L. S., Petitjean, C., & Heutte, L. (2016, July). Breast cancer histopathological image classification using convolutional neural networks. In 2016 international joint conference on neural networks (IJCNN) (pp. 2560-2567). IEEE.
- [21]. Abbass, H. A. (2002). An evolutionary artificial neural networks approach for breast cancer diagnosis. *Artificial intelligence in Medicine*, 25(3), 265-281.
- [22]. Ahmad, L. G., Eshlaghy, A. T., Poorebrahimi, A., Ebrahimi, M., & Razavi, A. R. (2013). Using three machine learning techniques for predicting breast cancer recurrence. *J Health Med Inform*, 4(124), 3.
- [23]. Rajakumari, S. Brintha, and C. Nalini. "An efficient data mining dataset preparation using aggregation in relational database." *Indian Journal of Science and Technology* 7 (2014): 44.
- [24]. Milletari, Fausto, et al. "Hough-CNN: deep learning for segmentation of deep brain regions in MRI and ultrasound." *Computer Vision and Image Understanding* 164 (2017): 92-102.

-
- [25]. Yan, L. C., Y. S. Bengio, and G. Hinton. "Deep learning." *nature* 521.7553 (2015): 436-444.
- [26]. Geras, K. J., Wolfson, S., Shen, Y., Wu, N., Kim, S., Kim, E., ... & Cho, K. (2017). High-resolution breast cancer screening with multi-view deep convolutional neural networks. *arXiv preprint arXiv:1703.07047*.
- [27]. Litjens, G., Sánchez, C. I., Timofeeva, N., Hermsen, M., Nagtegaal, I., Kovacs, I., ... & Van Der Laak, J. (2016). Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. *Scientific reports*, 6, 26286.
- [28]. Howard, Andrew G., et al. "Mobilenets: Efficient convolutional neural networks for mobile vision applications." *arXiv preprint arXiv:1704.04861* (2017).
- [29]. Visa, Sofia, et al. "Confusion Matrix-based Feature Selection." *MAICS* 710 (2011): 120-127.
- [30]. Ramos-Pollán, R., Guevara-López, M. A., Suárez-Ortega, C., Díaz-Herrero, G., Franco-Valiente, J. M., Rubio-Del-Solar, M., ... & Ramos, I. (2012). Discovering mammography-based machine learning classifiers for breast cancer diagnosis. *Journal of medical systems*, 36(4), 2259-2269.
- [31]. Rasti, R., Teshnehlal, M., & Phung, S. L. (2017). Breast cancer diagnosis in DCE-MRI using mixture ensemble of convolutional neural networks. *Pattern Recognition*, 72, 381-390.
- [32]. Becker, A. S., Marcon, M., Ghafoor, S., Wurnig, M. C., Frauenfelder, T., & Boss, A. (2017). Deep learning in mammography: diagnostic accuracy of a multipurpose image analysis software in the detection of breast cancer. *Investigative radiology*, 52(7), 434-440.
- [33]. Cireşan, D. C., Giusti, A., Gambardella, L. M., & Schmidhuber, J. (2013, September). Mitosis detection in breast cancer histology images with deep neural networks.

- In International Conference on Medical Image Computing and Computer-assisted Intervention (pp. 411-418). Springer, Berlin, Heidelberg.
- [34]. Wang, D., Khosla, A., Gargeya, R., Irshad, H., & Beck, A. H. (2016). Deep learning for identifying metastatic breast cancer. arXiv preprint arXiv:1606.05718.
- [35]. Bejnordi, B. E., Veta, M., Van Diest, P. J., Van Ginneken, B., Karssemeijer, N., Litjens, G., ... & Geessink, O. (2017). Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. *Jama*, 318(22), 2199-2210.
- [36]. Sun, W., Tseng, T. L. B., Zhang, J., & Qian, W. (2017). Enhancing deep convolutional neural network scheme for breast cancer diagnosis with unlabeled data. *Computerized Medical Imaging and Graphics*, 57, 4-9.
- [37]. Abdel-Zaher, A. M., & Eldeib, A. M. (2016). Breast cancer classification using deep belief networks. *Expert Systems with Applications*, 46, 139-144.
- [38]. Wang, Z., Li, M., Wang, H., Jiang, H., Yao, Y., Zhang, H., & Xin, J. (2019). Breast Cancer Detection Using Extreme Learning Machine Based on Feature Fusion with CNN Deep Features. *IEEE Access*.
- [39]. Cruz-Roa, A., Gilmore, H., Basavanthally, A., Feldman, M., Ganesan, S., Shih, N. N., ... & Madabhushi, A. (2017). Accurate and reproducible invasive breast cancer detection in whole-slide images: A Deep Learning approach for quantifying tumor extent. *Scientific reports*, 7, 46450.
- [40]. <http://dx.doi.org/10.17632/wmy84gzngw.1#file-fa206ab0-b2af-453a-b8d0-1e6e89230377>

Σύντομο Βιογραφικό Συγγραφέα

Ονομάζομαι Αντώνιος Ματζάκος Χωριανόπουλος και γεννήθηκα στην Αθήνα το 1995. Φοίτησα στην Ελληνογαλλική Σχολή «Ευγένιος Ντελακρουά» από την οποία αποφοίτησα το 2013. Σπουδάζω στο Τμήμα Μηχανικών Ηλεκτρονικών Υπολογιστών της Πολυτεχνικής Σχολής του Πανεπιστημίου Πατρών. Μιλώ 4 γλώσσες (μαζί με την μητρική) και πέρα από την ενασχόληση μου με το Πανεπιστήμιο έχω έντονη δραστηριότητα και στον τομέα της μουσικής.