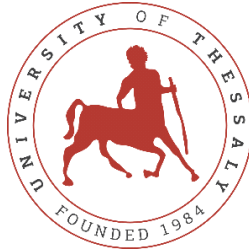


*Τίτλος: Προσέγγιση βαθιάς μάθησης για διάγνωση εικόνων αιμάτωσης
μυοκαρδίου και ανάπτυξη λογισμικού για αυτόματη κατηγοριοποίηση τους.*



Εκπαιδευτικό Ίδρυμα: Πανεπιστήμιο Θεσσαλίας

Τμήμα: Μηχανική Λογισμικού για Διαδικτυακές και Φορητές Εφαρμογές



Όνομα μεταπτυχιακού φοιτητή: Άννα Φελέκη

Επιβλέπουσα Καθηγήτρια: Δρ. Ελπινίκη Παπαγεωργίου

Κριτικός αναγνώστης: Δρ. Γεώργιος Κακαρόντζας

Λάρισα, 2021

Ευχαριστίες

Σε αυτό το σημείο θέλω να ευχαριστήσω θερμά την επιβλέπουσα καθηγήτρια μου κα. Ελπινίκη Παπαγεωργίου, για την αμέριστη καθοδήγηση και στήριξη της καθ' όλη τη διάρκεια της πτυχιακής, με τις πολύτιμες συμβουλές και τη συνεχή βοήθεια της, καθώς και για το τόσο ενδιαφέρον θέμα που επέλεξε για την εφαρμογή των αλγορίθμων βαθιάς μάθησης σε ιατρικές εικόνες και που διάλεξε εμένα να το αναπτύξω.

Επίσης, θα ήθελα να ευχαριστήσω τον κ. Παπανδριανό Νικόλαο, Πυρηνικό Ιατρό, Επίκουρο Καθηγητή του Τμήματος Συστημάτων Ενέργειας του Πανεπιστημίου Θεσσαλίας, για τη συλλογή και διάθεση των εικόνων, για την πολύτιμη βοήθειά του στην προεπεξεργασία τους και στη διάγνωσή τους, καθώς και για την αξιολόγηση των αποτελεσμάτων που προέκυψαν.

Τέλος, ευχαριστώ την οικογένεια μου που ήταν πάντα δίπλα μου για την απόκτηση του πτυχίου, αλλά και γενικά στη ζωή μου.

Περίληψη

Η παρούσα πτυχιακή αφορά την έρευνα γύρω από την εύρεση αλγορίθμου για την αυτόματη διάγνωση της στεφανιαίας νόσου, μέσα από εικόνες αιμάτωσης μυοκαρδίου, καθώς και την βιβλιογραφική ανασκόπηση της υπάρχουσας βιβλιογραφίας. Η έρευνα περιλαμβάνει μεθόδους τεχνητής νοημοσύνης και συγκεκριμένα τη μέθοδο βαθιάς μάθησης που βασίζεται στα συνελκτικά νευρωνικά δίκτυα. Τα δεδομένα, που χρησιμοποιούνται είναι ιατρικές εικόνες, οι οποίες απεικονίζουν την αιμάτωση του μυοκαρδίου, με την αριστερή κοιλιακή χώρα της καρδιάς, σε δυο ξεχωριστές αναπαραστάσεις, με σκοπό την ολοκληρωμένη μελέτη και την εύρεση του βέλτιστου αλγορίθμου βαθιάς μάθησης για τη διάγνωσή τους. Ο προτεινόμενος αλγόριθμος συνελκτικών νευρωνικών δικτύων παραμετροποιείται ανάλογα με το σετ δεδομένων, έτσι ώστε να δίνει επιθυμητό αποτέλεσμα κάθε φορά. Επιπλέον, παρουσιάζεται ένας εύκολος τρόπος προεπεξεργασίας των εικόνων, για να αφαιρεθεί ο θόρυβος, που είναι αναπόφευκτος στα αρχικά δεδομένα. Τέλος, προτείνεται ένα διαδικτυακό λογισμικό για την κατηγοριοποίηση εικόνων αιμάτωσης μυοκαρδίου, το οποίο παρέχει εύχρηστο και φιλικό περιβάλλον για τον ιατρό, με σκοπό την υποβοήθηση στη λήψη απόφασης της στεφανιαίας νόσου στον τομέα της πυρηνικής ιατρικής.

3.1.3	Ανάλυση όρθιων και πρόσθιων spect δεδομένων με μεθόδους βαθιάς μάθησης για την πρόβλεψη στεφανιαίας νόσου.....	49-51
3.1.4	Μέθοδοι βαθιάς μάθησης για την πρόβλεψη αποφρακτικής νόσου με δεδομένα spect.....	52-54
3.1.5	Μοντέλα κατηγοριοποίησης για δεδομένα Spect.....	54-56
3.1.6	Αυτόματη κατηγοριοποίηση πολικών χαρτών με μεθόδους βαθιάς μάθησης και αύξηση δεδομένων	57-59
3.1.7	Τεχνητή νοημοσύνη, μηχανική (βαθιά) μάθηση και ραδιονομική-γεωμική: ορισμού και εφαρμογές πυρηνικής ιατρικής	60

Κεφ. 4 Παρουσίαση του προβλήματος με δεδομένα

4.1	Ιατρικές φωτογραφίες.....	61-64
4.2	Προ επεξεργασία δεδομένων.....	65
4.2.1	Προεπεξεργασία SPECT MPI εικόνες.....	66-67
4.2.2	Προεπεξεργασία πολικών χαρτών.....	68-70
4.3	Ανάλυση λογισμικού και τον τρόπο υλοποίησης.....	71-75

Κεφ. 5: Αποτελέσματα εφαρμογής αλγορίθμων βαθιάς μάθησης στην διάγνωση της αιμάτωσης μυοκαρδίου

5.1	Περιγραφή της μεθοδολογίας της παρούσας έρευνας.....	76-77
5.2	Περιγραφή του μοντέλου.....	77-84
5.3	Παρουσίαση μετρικών απόδοσης μοντέλου.....	84-87
5.4	Αποτελέσματα εφαρμογής αλγορίθμου.....	88-106
5.5	Αποτελέσματα πολικών χαρτών.....	107-109
5.6	Μέθοδος cross-validation.....	109-113

5.7 Συγκριτικά τρεξίματα.....	113-116
5.8 Transfer Learning-RGB.....	116-117
6, Σύνοψη-Συμπεράσματα.....	118-119
7.Αναφορές.....	120-128

Πίνακας Συντομογραφιών

MPI	Myocardial Perfusion Imaging	Απεικόνιση Αιμάτωσης Μυοκαρδίου
SPECT	Single-photon emission computed tomography	Υπολογιστική τομογραφία εκπομπής ενός φωτονίου
CNN	Convolutional Neural Networks	Συνελκτικά νευρωνικά δίκτυα
PET	Positron Emission Tomography	Τομογραφία εκπομπής ποζιτρονίων
CAD	Coronary artery disease	Στεφανιαία νόσο
TPD	Total perfusion deficit	Ολικό έλλειμμα αιμάτωσης
CT	Computed-tomography	Υπολογιστική τομογραφία
AC	Attenuation-corrected	
NAC	Non-corrected	
PM	Polar maps	Πολικοί χάρτες
AUC	Area under curve	«Περιοχή κάτω από την καμπύλη»
ROC	Receiver Operating Characteristic curve	Χαρακτηριστική καμπύλη λειτουργίας
TPR	True positive rate	Πραγματικό θετικό ποσοστό
FPR	False positive rate	Ψευδώς θετικό ποσοστό
GAN	Generative Adversarial Networks	Γεννητικά δίκτυα αντιπαράθεσης
SVM	Support Vector Machines	Μηχανή φορέας υποστήριξης
KNN	K Nearest Neighbour	K-πλησιέστετος γείτονας
LAD	Left Anterior Descending	Αριστερά πρόσθια φθίνουσα
LCX	Left circumflex artery	Αριστερή περιφερειακή αρτηρία
RCA	Right Coronary Artery	Δεξιά στεφανιαία αρτηρία

Πίνακας 1

Κεφ 1.Εισαγωγή

1.1 Περιγραφή του προβλήματος ταξινόμησης της ιατρικής εικόνας

Τις τελευταίες δεκαετίες, η ιατρική απεικόνιση έχει βρει σημαντικές εφαρμογές στην ανίχνευση, τη διάγνωση και τη θεραπεία ασθενειών. Πιο συγκεκριμένα αποτελεί εργαλείο για περιοχές όπως η υπολογιστική τομογραφία (Computer Tomography), ο μαγνητικός συντονισμός (Magnetic resonance imaging), η τομογραφία εκπομπής ποζιτρονίων (positron emission tomography), η μαστογραφία, οι υπέρηχοι, οι ακτίνες X κλπ. [1]



Εικόνα 1 [2]

Στον τομέα της ιατρικής, η ερμηνεία της εικόνας έχει πραγματοποιηθεί ως επί το πλείστον από εμπειρογνώμονες στο χώρο της ιατρικής απεικόνισης, όπως ακτινολόγους και γιατρούς σχετικών ειδικοτήτων. Ωστόσο, λόγω του μεγάλου όγκου δεδομένων και της εμπειρίας που απαιτείται, οι ερευνητές και οι γιατροί άρχισαν πρόσφατα να επωφελούνται από παρεμβάσεις του υπολογιστή. Είναι εμφανής η πρόοδος στην υπολογιστική ανάλυση της ιατρικής εικόνας και οι τεχνικές μηχανικής μάθησης συνεισφέρουν ουσιαστικά στην γρήγορη και ακριβή διάγνωση στα σύγχρονα ιατρικά κέντρα.

Η εφαρμογή της μηχανικής μάθησης για ανάλυση δεδομένων και η ουσιαστική εξαγωγή/αναπαράσταση χαρακτηριστικών βρίσκεται στο επίκεντρο της επιτυχίας, καθώς έχουν υλοποιηθεί πάρα πολλές μέθοδοι. Στο παρελθόν, τα χαρακτηριστικά των δεδομένων ήταν δύσκολο να αναλυθούν επαρκώς και να χρησιμοποιηθούν από κάποιον που είναι μη ειδικός με

σκοπό την υποβοήθηση στη διάγνωση και λήψη αποφάσεων. Πρόσφατα, η τεχνητή νοημοσύνη και ειδικότερα η βαθιά μάθηση [3], ως κατηγορία της μηχανικής μάθησης, έχει βοηθήσει, ώστε τέτοια εμπόδια να μην υπάρχουν πλέον. Αντί της εξαγωγής χαρακτηριστικών με χειροκίνητο τρόπο, η βαθιά μάθηση απαιτεί μόνο ένα σύνολο δεδομένων με μικρή σχετικά προ-επεξεργασία, εάν είναι απαραίτητο, για να ανακαλύψει τις αναπαραστάσεις/μοτίβα των εικόνων. [4] [5]

Έτσι, τώρα το βάρος της εξαγωγής χαρακτηριστικών έχει μετατοπιστεί από την πλευρά του ανθρώπου στην πλευρά του υπολογιστή, επιτρέποντας έτσι σε μη ειδικούς με τη βοήθεια της μηχανικής μάθησης να χρησιμοποιούν αποτελεσματικά τα ιατρικά δεδομένα για τις δικές τους έρευνες και εφαρμογές.

Η επιτυχία της βαθιάς μάθησης προκύπτει κυρίως από τους ακόλουθους παράγοντες:

- εξελίξεις κεντρικών μονάδων επεξεργασίας υψηλής τεχνολογίας (Central Processing Unit) και μονάδων επεξεργασίας γραφικών (Graphics Processing Unit).
- διαθεσιμότητα τεράστιου όγκου δεδομένων
- εξελίξεις αλγορίθμων μηχανικής μάθησης [6] [7] [8].

Χάρη στο γεγονός ότι γίνεται εκμάθηση ιεραρχικών αναπαραστάσεων χαρακτηριστικών αποκλειστικά από δεδομένα, η βαθιά μάθηση έχει επιτύχει πρωτοποριακή απόδοση σε ποικιλία εφαρμογών τεχνητής νοημοσύνης [9] [10]. Ιδιαίτερα, οι μεγάλες βελτιώσεις στους αλγορίθμους που παρέχει ο υπολογιστής ενέπνευσαν την ύπαρξη διάφορων τεχνικών, όπως τμηματοποίηση εικόνας (segmentation) [11] [12], διάγνωση και πρόγνωση μέσω υπολογιστή, ανίχνευση βλαβών και ανάλυση μικροσκοπικής απεικόνισης. [13] [14] [15]

Οι μέθοδοι βαθιάς μάθησης είναι πολύ αποτελεσματικές όταν ο αριθμός των διαθέσιμων δειγμάτων είναι μεγάλος κατά τη διάρκεια της εκπαίδευσης. Για παράδειγμα, στο ImageNet Large Scale Visual Recognition Challenge (ILSVRC), που είναι μια γνωστή βάση δεδομένων εικόνων, παρασχέθηκαν περισσότερες από 1 εκατομμύριο κατηγοριοποιημένες εικόνες. [16]

Ωστόσο, δεν υπάρχει σύνολο δεδομένων απεικόνισης ιατρικής εικόνας μεγάλης κλίμακας συγκρίσιμο με το ImageNet, καθώς η απόκτηση δεδομένων είναι δύσκολη [17] και για αυτό το λόγο έχουμε πολύ περιορισμένο αριθμό εικόνων, π.χ. λιγότερες από 1.000 εικόνες. Για το λόγο, ότι έχουμε περιορισμένο αριθμό δειγμάτων εκπαίδευσης για τη δημιουργία μοντέλων βαθιάς μάθησης, οι ερευνητικές ομάδες έχουν επινοήσει διάφορες στρατηγικές, όπως

(i) να λαμβάνουν τμήματα εικόνας είτε από 2D είτε 3D ως είσοδο [14] [18] [19] , αντί για το πλήρες μέγεθος εικόνας, για να μειωθεί η κοινοτυπία δεδομένων, επομένως ο αριθμός των παραμέτρων του μοντέλου.

(ii) να επεκτείνουν το σύνολο δεδομένων τους δημιουργώντας τεχνητά δείγματα μέσω μετασχηματισμού (δηλαδή, αύξηση δεδομένων) και στη συνέχεια να εκπαιδεύσουν το δίκτυό τους από το μηδέν με το αυξημένο σύνολο. [18] [19] [20] [21]

(iii) να χρησιμοποιούν βαθιά μοντέλα που έχουν εκπαιδευτεί πάνω από έναν τεράστιο αριθμό φυσικών εικόνων και στη συνέχεια να εκπαιδεύσουν τον τελικό ταξινομητή (μεταφορά εκπαίδευσης). [22] [17]

(iv) να γίνει προετοιμασία στις παραμέτρους του μοντέλου με εκείνες των προ-εκπαιδευμένων μοντέλων από μη ιατρικές ή φυσικές ακόμη εικόνες και, στη συνέχεια, να προσαρμοστούν οι παράμετροι του δικτύου με τα δείγματα που σχετίζονται με την ήδη υπάρχουσα μελέτη [23] [24] [25] [26].

Τα μοντέλα συνελκτικών νευρωνικών δικτύων (Convolutional Neural Networks), που είναι ειδικά για δεδομένα εικόνων και εκπαιδεύονται σε παγκόσμια βάση δεδομένων, όπως το ImageNet [27] χρησιμεύουν ουσιαστικά στη βελτίωση πολλών προβλημάτων εντοπισμού αντικειμένων και τμηματοποίησης εικόνας. Είναι ιδιαίτερα χρήσιμο να χρησιμοποιούμε σύνολα άλλων δεδομένων για την κατηγοριοποίηση ιατρικής εικόνας. [22]

1.2 Συνελικτικά Νευρωνικά Δίκτυα στην επεξεργασία ιατρικής εικόνας

Η υλοποίηση της βαθιάς μάθησης στην ιατρική απεικόνιση πραγματοποιείται κυρίως από συνελικτικά νευρωνικά δίκτυα - Convolutional Neural Networks (CNNs) [28] [16], έναν σχετικά νέο και ισχυρό τρόπο εκμάθησης χρήσιμων αναπαραστάσεων εικόνων/φωτογραφιών. Πριν από την εφαρμογή τους, αυτά τα χαρακτηριστικά συνήθως έπρεπε να δημιουργηθούν από λιγότερο ισχυρά μοντέλα μηχανικής μάθησης ή ακόμα και χειροκίνητα.

Με την εισαγωγή των συνελικτικών δικτύων, μπορεί να γίνει εκμάθηση των χαρακτηριστικών απευθείας από τα δεδομένα, καθώς περιλαμβάνουν ορισμένες λειτουργίες στη δομή τους που τα καθιστούν ισχυρά μοντέλα βαθιάς μάθησης για ανάλυση εικόνας [29] [30] [31].

Τα τυπικά συνελικτικά νευρωνικά δίκτυα έχουν παρόμοια δομή με τα τεχνητά νευρωνικά δίκτυα (Artificial Neural Networks) και αποτελούνται από ένα ή περισσότερα φίλτρα (δηλαδή, συνελικτικά επίπεδα), ακολουθούμενα από συγκεντρωτικά στρώματα για την εξαγωγή χαρακτηριστικών για εργασίες ταξινόμησης [32].

Τα τελευταία χρόνια, τα CNNs έχουν αποκτήσει ευρεία αναγνώριση στον τομέα της ανάλυσης ιατρικής εικόνας [33] [34] [35]. Η εφαρμογή τους στην ιατρική απεικόνιση κυμαίνεται από απλή ακτινογραφία, και μικροσκοπικές εικόνες σε κλινικές εφαρμογές, για δερματολογικό έλεγχο και οπτική αναγνώριση, [36] [37] [38] [39] για προβλήματα όπως αυτόματη ανίχνευση φυματίωσης σε ακτινογραφίες θώρακα καθώς [40] και σε κατάτμηση όγκου εγκεφάλου. Επιπλέον, έχουν δείξει εξαιρετική απόδοση στην ακτινολογία και τη μοριακή απεικόνιση [38].

1.3 Εξαγωγή δεδομένων και απεικόνιση αιμάτωσης μυοκαρδίου

Το πρόβλημα, που μελετάται στα πλαίσια της παρούσας πτυχιακής, είναι η διάγνωση της αιμάτωσης του μυοκαρδίου [29] [34] [41], μέσω της απεικόνισης. Η απεικόνιση της αιμάτωσης του μυοκαρδίου (myocardial perfusion imaging or scanning (MPI or MPS) είναι μια διαδικασία ιατρικής απεικόνισης στην πυρηνική ιατρική, η οποία απεικονίζει τη λειτουργία του μυ της καρδιάς [42].

Για να γίνει η εξέταση του ασθενούς, πρέπει να γίνει εισαγωγή ενδοφλέβιου υγρού, έτσι ώστε να εντοπιστούν μέσα από την κάμερα απεικόνισης, οι περιοχές που θα διαπεράσει, και να φανεί αν κάποια αρτηρία είναι φραγμένη.

Συνήθως χρειάζονται δύο είδη εικόνων σε μορφή στρες (stress), και σε μορφή χαλάρωσης (rest). Είναι σημαντικό να σημειωθεί ότι για να ελεγχθεί η αιμάτωση του μυοκαρδίου, υπολογίζεται η αριστερή περιοχή της κοιλίας (left ventricular ejection fraction (LVEF)). Ο έλεγχος αυτός γίνεται σε συνδυασμό με ένα καρδιακό τεστ stress-rest, από όπου προκύπτουν οι δυο μορφές εικόνων αιμάτωσης του μυοκαρδίου. Τα δεδομένα για την φάση στρες παράγονται με την πρόκληση ελεγχόμενης τοπικής ισχαιμίας, δηλαδή μεταβλητή και ελεγχόμενη αιμάτωση, είτε με το ενδοφλέβιο υγρό, είτε με ελεγχόμενη από τον ιατρό γυμναστική. Όσον αφορά για την φάση χαλάρωσης γίνεται μια απλή δειγματοληψία με τον ασθενή καθημένο. Η απεικόνιση με την κάμερα υπολογιστικής τομογραφίας εκπομπής μονών φωτονίων έγχυσης SPECT (Single Photon Emission Computed Tomography) καταφέρνει να ολοκληρώνει τη διαδικασία σε λιγότερο από δέκα λεπτά. Το βασικό πλεονέκτημα είναι ότι μπορεί να εντοπίσει ακόμη και περιοχές με μικρές ανωμαλίες λόγω εμφράγματος, καθώς και την απόφραξη αιμοφόρων αγγείων (Jahng et al., 2014) [43] [42].

Οι εικόνες αποθηκεύονται ως αρχεία βίντεο και αναλύονται σε έναν ειδικό σταθμό εργασίας. Η πλειονότητα των κλινικών σαρώσεων αναλύεται ποιοτικά συγκρίνοντας οπτικά τις εικόνες σε κατάσταση στρες και τις εικόνες σε κατάσταση χαλάρωσης παράλληλα. Σε μια κανονική σάρωση μπορεί να θεωρηθεί ότι το μυοκάρδιο μετατρέπεται από μαύρο σε μεσαίο γκριζο ομοιόμορφα. Σε μια μη φυσιολογική σάρωση, μια περιοχή του μυοκαρδίου θα γίνει γκρι με πιο αργούς ρυθμούς από τον περιβάλλοντα ιστό, καθώς το αίμα (και ως εκ τούτου το γαδολίνιο, που κυκλοφορεί μέσω του αίματος) εισέρχεται πιο αργά λόγω της στένωσης της στεφανιαίας αρτηρίας. Αυτό ονομάζεται

ελάττωμα διάχυσης και συνήθως αντιπροσωπεύει ισχαιμία του μυοκαρδίου. Μπορεί να παρατηρηθεί τόσο στις υπόλοιπες σαρώσεις όσο και στις ανιχνεύσεις σε κατάσταση στρες, οπότε ονομάζεται ταιριασμένο ελάττωμα διάχυσης και πιθανότατα οφείλεται σε μια περιοχή ή ουλή από προηγούμενο έμφραγμα του μυοκαρδίου. Εάν παρατηρηθεί μόνο στη σάρωση στρες, ονομάζεται ισχαιμία [44] [45]. Η διαδικασία αυτή έχει την ικανότητα να ελέγχει πολλές παθήσεις της καρδιάς, όπως τη στεφανιαία νόσο (coronary artery disease (CAD)), υπερτροφική καρδιομυοπάθεια και ανωμαλίες στις κινήσεις της καρδιάς, αλλά και ακόμη παθήσεις σε διαφορετικά όργανα, όπως την ανίχνευση εγκεφαλικού [46].

Οι καρδιαγγειακές παθήσεις, ιδιαίτερα η στεφανιαία νόσος (CAD), είναι ένα από τα πιο κοινά προβλήματα υγείας. Η υπολογιστική τομογραφία εκπομπής μονών φωτονίων έγχυσης SPECT (Single Photon Emission Computed Tomography) απεικόνισης της αιμάτωσης του μυοκαρδίου παρέχει τρισδιάστατες πληροφορίες σχετικά με την κατανομή μιας ραδιενεργού ένωσης εντός της καρδιάς και είναι η πλέον προτιμώμενη, οικονομικά αποδοτική μέθοδος για τον εντοπισμό και τον προσδιορισμό της σοβαρότητας της στεφανιαίας νόσου. Αυτή η μέθοδος απεικόνισης επιτρέπει τον κατάλληλο σχεδιασμό θεραπείας. [47] [48] Στο SPECT MPI, το λογισμικό ανάλυσης ποσοτικοποιεί συνήθως την ασθένεια ως τη διαφορά μεταξύ των μεγεθών του στρες και του ήρεμης κατάστασης τα οποία λαμβάνονται από ομάδες ασθενών, όπου έχουν ήδη αναγνωριστεί ως φυσιολογικά. Η ερμηνεία των εικόνων SPECT πραγματοποιείται συνήθως από έμπειρους πυρηνικούς ιατρούς οπτικά. [49] [39] [50] Η ερμηνεία αυτών των εικόνων είναι χρονοβόρα εργασία και εξαρτάται από την εμπειρία του εκάστοτε γιατρού και ως εκ τούτου υπόκειται σε σημαντική διακύμανση μεταξύ των ειδικών ιατρών [51] [52].

Η αυτοματοποιημένη ανάλυση με τη βοήθεια εργαλείων διάγνωσης μέσω υπολογιστή μπορεί να βελτιώσει την αντικειμενικότητα τυποποιώντας την ερμηνεία. Βελτιώνει επίσης την ακρίβεια της διάγνωσης και μειώνει το χρόνο διάγνωσης και το κόστος υγειονομικής περίθαλψης. [53] Μπορεί να χρησιμοποιηθεί ως δεύτερη γνώμη για να βοηθήσει στην ανάλυση πολύπλοκων περιπτώσεων και να εκπαιδεύσει επίσης νέους ιατρούς.

Με την αυτοματοποίηση της διαδικασίας εξαγωγής συμπερασμάτων και υποβοήθηση στη λήψη αποφάσεων μπορούν να επωφεληθούν οι λιγότερο έμπειροι ιατροί [29].

1.4 Στόχος της πτυχιακής εργασίας

Ο στόχος της παρούσας πτυχιακής εργασίας είναι η διάγνωση της στεφανιαίας νόσου μέσω της αιμάτωσης του μυοκαρδίου στην πυρηνική ιατρική, με την χρήση αυτοματοποιημένου αλγορίθμου βασισμένου στα συνελκτικά νευρωνικά δίκτυα. Η υπολογιστική τομογραφία εκπομπής μονών φωτονίων (SPECT) που χρησιμοποιείται για την απεικόνιση της αιμάτωσης του μυοκαρδίου (MPI) είναι μια ευρέως χρησιμοποιούμενη τεχνική για τη διάγνωση της στεφανιαίας νόσου (CAD), με περισσότερες από 7 εκατομμύρια σαρώσεις ετησίως μόνο στις Ηνωμένες Πολιτείες. Ο ποσοτικός προσδιορισμός της σχετικής αιμάτωσης του μυοκαρδίου χρησιμοποιώντας κανονικές βάσεις δεδομένων επιτρέπει τον καθορισμό του ολικού ελλείμματος διάχυσης (total perfusion deficit (TPD)) σε στρες και rest φωτογραφίες. Η ποσοτικοποίηση TPD έχει αποδειχθεί ότι είναι ισοδύναμη με την οπτική ερμηνεία/διάγνωση έμπειρων πυρηνικών ιατρών για την ανίχνευση αποφρακτικού CAD, ενώ παρέχει σχεδόν αυτόματη εξαγωγή συμπεράσματος.

Αρχικά, μελετήθηκε η υπάρχουσα βιβλιογραφία τόσο ως προς το κομμάτι που αφορά τη βαθιά μάθηση, αλλά και ως προς την ανάλυση εικόνας, καθώς η τεχνητή νοημοσύνη παρέχει αλγόριθμους ως προς τη διαδικασία αυτή. Η εστίαση της έρευνας είναι η κατηγοριοποίηση του προβλήματος, μέσω μεθόδων μηχανικής μάθησης και ιδιαίτερα βαθιάς μάθησης. Τα δεδομένα χορηγήθηκαν από Πυρηνικό Ιατρό, ο οποίος επέλεξε τις εικόνες για την υλοποίηση της έρευνας και αφορούν δυο διαφορετικές αναπαραστάσεις.

Αρχικά διερευνήθηκαν τα μοντέλα των συνελκτικών νευρωνικών δικτύων τα οποία παραμετροποιήθηκαν κατάλληλα για την καλύτερη απόδοση και ακρίβεια της διάγνωσης, χρησιμοποιώντας μετρικές όπως ακρίβεια και το σφάλμα. Επίσης εξάχθηκε ROC καμπύλη, η οποία δείχνει την αντιστάθμιση μεταξύ της τιμής sensitivity (ή TPR) και της specificity (1-FPR). Οι ταξινομητές που δίνουν καμπύλες πιο κοντά στην επάνω αριστερή γωνία υποδηλώνουν καλύτερη απόδοση. Όσο πλησιάζει η καμπύλη στη διαγώνιο 45 μοιρών του χώρου ROC, τόσο λιγότερο ακριβής είναι η δοκιμή[54]. Επιπλέον έχουμε και την AUC, η οποία δείχνει πόσο το μοντέλο είναι ικανό να διακρίνει μεταξύ κλάσεων. Όσο υψηλότερη είναι η AUC, τόσο καλύτερο είναι το μοντέλο στην πρόβλεψη εξόδων στα δεδομένα, όπως στη διάκριση μεταξύ ασθενών με νόσο και χωρίς ασθένεια [55].

Επίσης, δημιουργήθηκε λογισμικό, το οποίο βασίζεται πάνω στον προτεινόμενο αλγόριθμο κατηγοριοποίησης, και παρέχει διεπαφή, που θα υποβοηθήσει στη διάγνωση αν υπάρχει ισχαιμία, ή έμφραγμα ή είναι φυσιολογική, ανάλογα με την εικόνα SPECT που θα παρέχεται από τον πυρηνικό ιατρό.

Τόσο η αλγοριθμική προσέγγιση για την κατηγοριοποίηση, όσο και το λογισμικό που έχει αναπτυχθεί είναι νέα αποδοτικά εργαλεία και μπορούν να βοηθήσουν ουσιαστικά στη λήψη απόφασης του πυρηνικού ιατρού.

Κεφ. 2 Αλγόριθμοι Βαθιάς μάθησης

2.1 Τεχνητή Νοημοσύνη

Η τεχνητή νοημοσύνη είναι γεγονός ότι έχει φέρει επανάσταση στην κλινική έρευνα. Από την πρόσφατη εξέλιξη ισχυρών αλγορίθμων μηχανικής μάθησης και την αυξανόμενη υπολογιστική ισχύ, η μελέτη πολύ μεγάλων συνόλων δεδομένων κατέδειξε ότι η κλινική ταξινόμηση και η πρόβλεψη μπορούν να βελτιστοποιηθούν με την εξερεύνηση πολύ-διάστατων δεδομένων. Όσον αφορά το πρόβλημα, που ερευνάται στην εργασία αυτή, ισχύει ότι κατά την αξιολόγηση, έχει βελτιστοποιηθεί η αναγνώριση των ανωμαλιών της αιμάτωσης πέρα από τα κλασσικά μέτρα, καθώς έχει σημειωθεί βελτίωση της πρόβλεψης ανεπιθύμητων ενεργειών, όπως καρδιαγγειακού επεισοδίου. Οι αλγόριθμοι, γενικά χαρακτηρίζονται από την ικανότητά τους να βελτιώνουν σταδιακά την απόδοσή τους μέσω της έκθεσης σε μεγάλες ποσότητες δεδομένων, όπως για παράδειγμα στην εκπαίδευση. Υπάρχουν πολλά παραδείγματα ισχυρών αλγορίθμων, όπως μηχανήματα υποστήριξης φορέα [56], ενισχυμένα σύνολα [57] και random forest [58]. Αλλά το πιο αξιοσημείωτο πρόσφατο παράδειγμα είναι τα συνελκτικά νευρωνικά δίκτυα για την ανάλυση δεδομένων από εικόνες.

Με την χρήση της μηχανικής μάθησης, η πυρηνική καρδιολογία έχει αρχίσει να προβάλλει τεράστιο όφελος από βελτιωμένες δυνατότητες στην επεξεργασία της εικόνας και την ταξινόμηση μεγάλων δεδομένων από διάφορες πηγές. Η βαθιά μάθηση έχει αναμφισβήτητα τα πιο αξιοσημείωτα αποτελέσματα στο πλαίσιο της βελτίωσης της διάγνωσης και της πρόγνωσης στην πυρηνική καρδιολογία. Οπότε συμπερασματικά, η τεχνητή νοημοσύνη είναι ένα αλγοριθμικό σύστημα του υπολογιστή που δείχνει έξυπνη συμπεριφορά παρόμοια με τον άνθρωπο στην αντιμετώπιση προβλημάτων, στην εκτέλεση των εργασιών και στην προσαρμογή στη δημιουργία λύσεων [59].

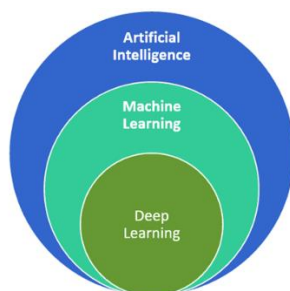
Οι επιστήμονες εργάζονται στις θεωρίες της τεχνητής νοημοσύνης από τον Γκόντελ τη δεκαετία του 1930 και από τον Τούρινγκ το 1950, ο οποίος ρώτησε για πρώτη φορά αν οι μηχανές θα μπορούσαν να σκεφτούν. Η τεχνητή νοημοσύνη μπορεί να οργανώσει και να ταξινομήσει δεδομένα γρήγορα και αποτελεσματικά, καθώς και να εξάγει δεδομένα για μοτίβα και στη συνέχεια να χρησιμοποιήσει τα δεδομένα αυτά για να κάνει προβλέψεις. Όλες αυτές οι λειτουργίες

μπορούν να γίνουν πλέον σε ταχύτητες πολύ μεγαλύτερες από τις ανθρώπινες δυνατότητες. [60] [61] [62].

Η τεχνητή νοημοσύνη απαιτεί επίσης δοκιμές και συνεπώς απαιτούνται κλινικές δοκιμές για να διασφαλιστεί ότι κάνει αυτό που σκοπεύει. Καθώς εισάγονται περισσότερα δεδομένα σε ένα σύστημα, η έξοδος θα ποικίλλει. Μία από τις ανησυχίες που εκφράζονται από ορισμένους είναι ότι η τεχνητή νοημοσύνη θα γίνει πιο έξυπνη από τον άνθρωπο – ένας όρος που ονομάζεται «υπερευφυΐα». Αυτό μπορεί να οδηγήσει σε επιταχυνόμενη τεχνολογική πρόοδο που ξεπερνά τον ανθρώπινο έλεγχο με στόχους που μπορεί να μην ταιριάζουν με τους ισχύοντες κοινωνικούς κανόνες.

Η γνώση των αποτελεσμάτων τα τελευταία χρόνια σχετικά με την ποιότητα στην υγειονομική περίθαλψη έχει δείξει ότι τόσο τα μηχανήματα όσο και οι άνθρωποι μπορούν να κάνουν λάθη, οπότε πρέπει να υπάρχουν συστήματα για την πρόληψη, τον εντοπισμό και τη διόρθωση αυτών. Τώρα είναι η ώρα να διασφαλίσουμε ότι τα δεδομένα εισόδου είναι έγκυρα και ότι διαθέτουμε υψηλή ακρίβεια στα στοιχεία αυτά και να καθαρίσουμε τυχόν βάσεις που υπάρχουν, αναπτύσσοντας εθνικές βάσεις δεδομένων για βελτιωμένη κλινική χρησιμότητα. [60] [61] [62]

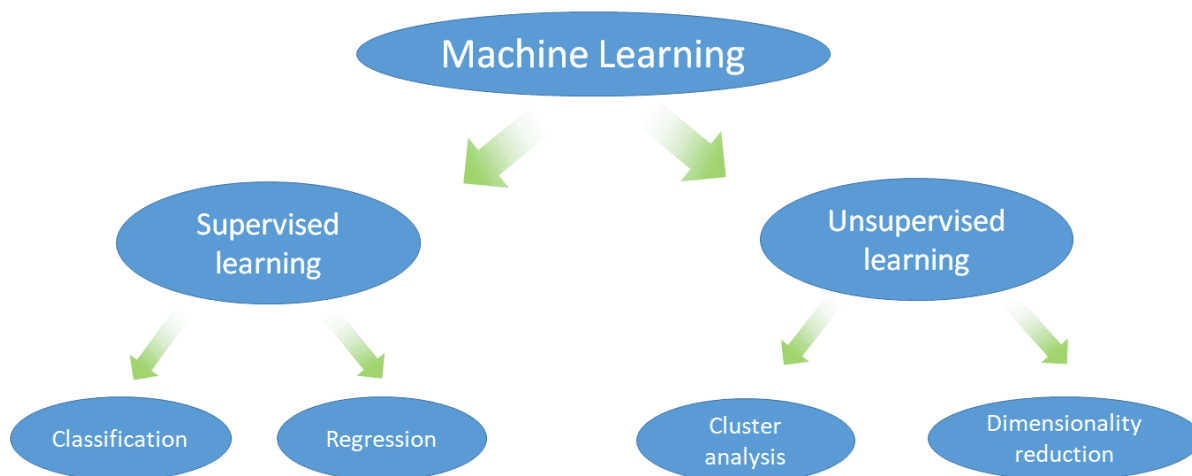
Η τεχνητή νοημοσύνη περιλαμβάνει όπως φαίνεται στην εικόνα 2, τη μηχανική μάθηση, η οποία έχει ως υποσύνολο την βαθιά μάθηση.



Εικόνα 2

2.2 Μηχανική μάθηση

Η μηχανική μάθηση είναι η ικανότητα ενός αλγορίθμου/προγράμματος να μαθαίνει από την εμπειρία και να προσαρμόζεται για την καλύτερη εξαγωγή αποτελέσματος, χωρίς να απαιτείται ρητός προγραμματισμός σε μια επαναληπτική διαδικασία. Η μάθηση μπορεί να είναι με επίβλεψη (supervised), δηλαδή να γνωρίζουμε την έξοδο-στόχο, που θέλουμε να πετύχουμε, ή αλλιώς χωρίς επίβλεψη (unsupervised), όπου δημιουργούνται ομάδες (clusters), από κοινά χαρακτηριστικά, καθώς δεν υπάρχει έξοδος, όπως φαίνεται στην εικόνα 3. Επιπλέον υπάρχει και μια συγκεκριμένη κατηγορία, η οποία είναι ημί-επίβλεψη, όπου συνδυάζονται τα δεδομένα με έξοδο και αυτά με χωρίς. Ένα ενδιάμεσο βήμα είναι η ημι-εποπτευόμενη μάθηση. Τα δεδομένα χωρίς ετικέτα χρησιμοποιούνται για τον προσδιορισμό της γενικής δομής των δεδομένων και τα δεδομένα με ετικέτα χρησιμοποιούνται για να μάθουν τη διαφορά ανάμεσα σε διαφορετικούς εξόδους.



Εικόνα 3

Η τεχνητή νοημοσύνη καλύπτει πολλές διαφορετικές μεθόδους και απαιτεί γνώση των μαθηματικών, της άλγεβρας, των στατιστικών και της πληροφορικής.

[63]

Η μηχανική μάθηση έχει δοκιμαστεί σε διάφορους τομείς της πυρηνικής ιατρικής, συμπεριλαμβανομένης της πυρηνικής καρδιολογίας, της ογκολογίας και της νευρολογίας. Υπήρξε επιτυχία με την αυτόματη ανίχνευση άκρων και τον καθορισμό όγκου και επιπλέον την αυτόματη ανίχνευση της ανατομίας και των πνευμονικών οζιδίων στο PET / CT. Η βελτιωμένη διάγνωση, η σταδιοποίηση και η αξιολόγηση της απόκρισης στη θεραπεία χρησιμοποιώντας PET / CT είχε

θετικά αποτελέσματα και η ανάλυση υφής βοήθησε στη διαφοροποίηση των περιπτώσεων καλοήθων/κακοήθων νοσών. [60]

2.3 Προεπεξεργασία και είσοδος δεδομένων στο σύστημα

Όσον αφορά το κομμάτι για την προεπεξεργασία των δεδομένων πριν εισέλθουν στον αλγόριθμο προηγούνται τα παρακάτω βήματα:

- Ανάγνωση δεδομένων (εικόνες). Σε αυτό το βήμα, αποθηκεύουμε τη τοποθεσία στον υπολογιστή που βρίσκονται τα δεδομένα σε μια μεταβλητή και στη συνέχεια φορτώνουμε τον φάκελο που περιέχουν τις εικόνες σε πίνακες.
- Αλλαγή μεγέθους εικόνας. Συνήθως τα δεδομένα ποικίλουν σε μέγεθος, και για αυτό πρέπει να καθοριστεί ένα καθολικό μέγεθος για όλες τις εικόνες που τροφοδοτούνται στους αλγόριθμους τεχνητής νοημοσύνης.
- Ανακάτεμα δεδομένων (shuffle), έτσι ώστε να είναι όσο πιο τυχαία γίνεται ο διαχωρισμός των δεδομένων
- Αφαίρεση θορύβου (Denoise). Μία από τις μεθόδους που χρησιμοποιείται είναι το Gaussian blur (επίσης γνωστό ως Gaussian smoothing) και είναι το αποτέλεσμα του θολώματος μιας εικόνας. Το οπτικό αποτέλεσμα αυτής της τεχνικής θαμπώματος είναι ένα ομαλό θόλωμα που μοιάζει με αυτό της προβολής της εικόνας μέσα από μια ημιδιαφανή οθόνη. Η εξομάλυνση Gauss χρησιμοποιείται με σκοπό την ενίσχυση των δομών εικόνας σε διαφορετικές κλίμακες.
- Τμηματοποίηση και μορφολογία. Τμηματοποιείται η εικόνα, διαχωρίζοντας το φόντο από αντικείμενα προσκηνίου, με σκοπό την περισσότερη και καλύτερη αφαίρεση θορύβου [64] [65].

Τα δύο τελευταία βήματα μπορούν να παραλειφθούν, αν χρησιμοποιείται ως αλγόριθμος τα συνελκτικά νευρωνικά δίκτυα, επειδή είναι ικανά από μόνα τους να προ-επεξεργάζονται τα δεδομένα για αυτό το κομμάτι.

Αφού φορτωθούν τα δεδομένα στον αλγόριθμο χωρίζονται σε τρία σετ δεδομένων: σετ εκπαίδευσης (training dataset), σετ ελέγχου (testing dataset), σετ επαλήθευσης (validation dataset). Όπως δείχνει και η εικόνα 4, το σετ εκπαίδευσης καταλαμβάνει το μεγαλύτερο ποσοστό (συνήθως το 75% του συνολικού σετ δεδομένων).



Εικόνα 4

Το πρώτο βήμα για τα δεδομένα, ύστερα από την προ-επεξεργασία είναι το σετ εκπαίδευσης. Το σετ εκπαίδευσης περιέχει τα δεδομένα εισόδου, δηλαδή στη δική μας περίπτωση τις φωτογραφίες σε μορφή πίνακα με πίξελ, με αντίστοιχα την έξοδο, της κάθε οντότητας. Αφού τα δεδομένα διασχίσουν από το μοντέλο, ο αλγόριθμος εξάγει μοτίβα για να είναι σε θέση να κάνει εύκολη κατηγοριοποίηση και στο μέλλον. Σε αυτό το στάδιο ακολουθείται η backpropagation τεχνική, δηλαδή γίνεται προσαρμογή των τιμών, το οποίο γίνεται για τον εξής λόγο: αν μετά τη διάσχιση των δεδομένων από τον αλγόριθμο, η προτεινόμενη έξοδος δεν συμβαδίζει με την πραγματική έξοδο, προσαρμόζονται από την αρχή οι παράμετροι και τα βάρη, τα οποία αρχικά είχαν οριστεί τυχαία, ώστε να φτάσουν όσο πιο κοντά στην κανονική τιμή.[66]

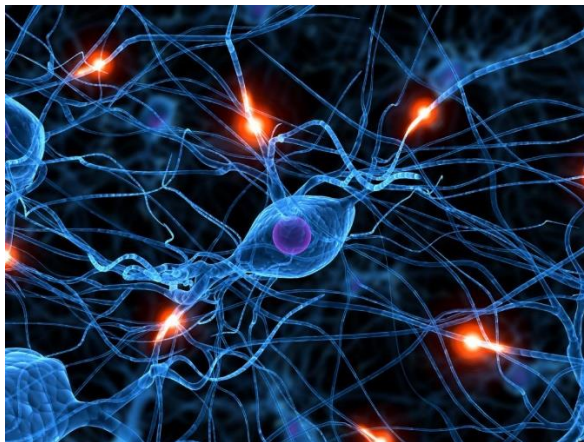
Το σετ επαλήθευσης είναι ένα κομμάτι από το σετ εκπαίδευσης, το οποίο δεν περνάει μέσα από το μοντέλο και η χρησιμότητα του είναι να προσφέρει εκτίμηση στην αλλαγή των παραμέτρων, με σκοπό να αποτρέψει την υπερπροσαρμογή (overfitting). Η υπερπροσαρμογή είναι όταν ένα μοντέλο σε συνδυασμό με τις τιμές των παραμέτρων έχει μάθει να εξάγει μοτίβα μόνο από ένα συγκεκριμένο σετ δεδομένων, το οποίο είναι μη επιθυμητό, καθώς είναι πολύ πιθανόν με καινούργιες φωτογραφίες να μην έχει καλή απόδοση. Τα δεδομένα που βρίσκονται μέσα στο σετ επαλήθευσης δίνονται στον αλγόριθμο, με σκοπό να τα κατηγοριοποιήσει, αφού έχει εκπαιδευτεί για να δούμε την απόδοση του, με ποσοστιαίες τιμές. [67]

Τέλος, το σετ ελέγχου, χρησιμοποιείται με σκοπό την τελική εκτίμηση του μοντέλου, χωρίς να προσφέρει βελτιστοποίηση στις παραμέτρους. Υπάρχει αξιολόγηση σε αυτό το βήμα, καθώς γίνεται κατηγοριοποίηση δεδομένων από τον αλγόριθμο, που δεν είχε επεξεργαστεί στο παρελθόν. Είναι το βήμα, όπου επιβλέπουμε πως θα συμπεριφερθεί ο αλγόριθμος σε πραγματικά δεδομένα, που δεν έχουν έξοδο. [68] [69] [70]

2.4 Νευρωνικά Δίκτυα

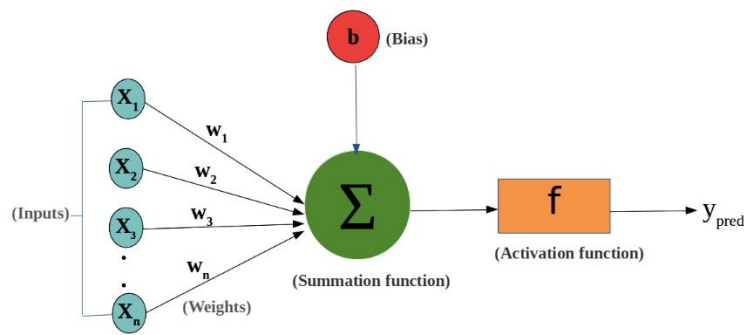
Τα νευρωνικά δίκτυα είναι μια από τις μεθόδους της τεχνητής νοημοσύνης, τα οποία έχουν επηρεαστεί από τα βιολογικά νευρωνικά δίκτυα, που βρίσκονται στον εγκέφαλο των ζωντανών οργανισμών.

Παρακάτω βλέπουμε πως είναι ένας εγκεφαλικός νευρώνας.



Εικόνα 5

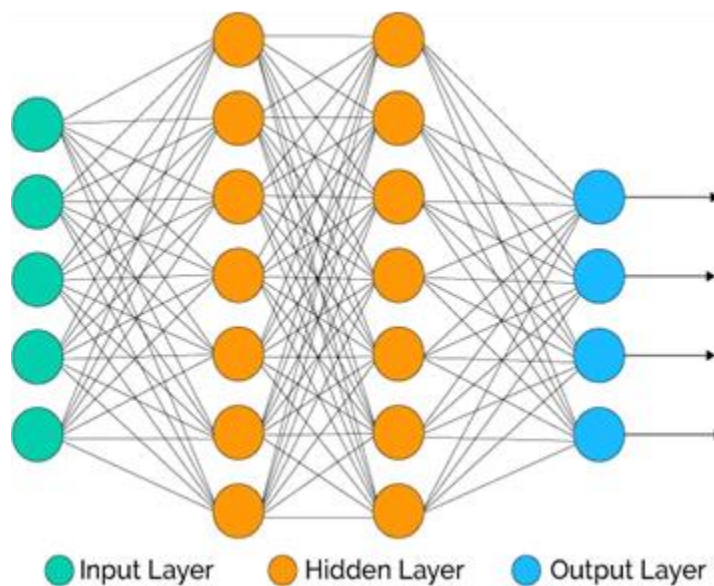
Είναι μια συλλογή από συνδεδεμένους κόμβους, όπως φαίνεται στην εικόνα 5, οι οποίοι αποκαλούνται τεχνητοί κόμβοι και αναπαριστούν τους νευρώνες στο βιολογικό εγκέφαλο. Κάθε σύνδεση, όπως οι συνάψεις μεταδίδουν ένα σήμα στους υπόλοιπους νευρώνες. Το σήμα που εισέρχεται είναι μια αριθμητική τιμή, η οποία προκύπτει ύστερα από τον υπολογισμό μιας μη γραμμικής συνάρτησης, της συνέλιξης από τα αθροίσματα των εισόδων. Οι συνδέσεις λέγονται κορυφές. Οι νευρώνες και οι κορυφές έχουν μια μεταβλητή το βάρος, το οποίο αρχικά του δίνεται μια τιμή, αλλά στη συνέχεια προσαρμόζεται καθώς προχωράει η μάθηση, ανάλογα με την σημαντικότητα του νευρώνα, που έχει στην εξαγωγή αποτελέσματος.



Εικόνα 5

Τα νευρωνικά δίκτυα αποτελούνται από επίπεδα, όπου τα διαφορετικά επίπεδα κάνουν διαφορετικές μετατροπές στις εισόδους. [67] [71] [3] [72]

Ένα παράδειγμα ενός νευρωνικού δικτύου παρουσιάζεται παρακάτω στην εικόνα 6. Η συγκεκριμένη αναπαράσταση αποτελείται από ένα στρώμα εισόδου με 5 κόμβους, δύο κρυφά επίπεδα με 7 κόμβους το καθένα και ένα στρώμα εξόδου με 4 κόμβους, δηλαδή 16 πιθανούς εξόδους, καθώς υπολογίζεται ως 2^x , όπου x οι κόμβοι.



Εικόνα 6

Τα επίπεδα εισόδου παίρνουν ως είσοδο στη δική μας περίπτωση ένα πίνακα από εικονοστοιχεία, τα οποία αναπαριστούν την ιατρική φωτογραφία. Παρόλα αυτά μπορούν να αποθηκεύσουν αρκετές μεταβλητές.

Η είσοδος σε κάθε διαδοχικό στρώμα, όσον αφορά για τα κρυφά επίπεδα και για το επίπεδο εξόδου είναι η έξοδος από το προηγούμενο στρώμα, έτσι ώστε να αποκτούνται πολλαπλά επίπεδα αναπαράστασης που αντιστοιχούν σε διαφορετικά επίπεδα αφαίρεσης. [73]

Γενικά ισχύει, ότι η λειτουργία του κάθε επιπέδου στα νευρωνικά δίκτυα είναι να μεταμορφώνει τα δεδομένα της εισόδου σε μια πιο αφαιρετική αναπαράσταση, έτσι ώστε η τελική μορφή να είναι όσο πιο γενική γίνεται, με σκοπό την ανίχνευση των μοτίβων για τις μελλοντικές μη κατηγοριοποιημένες φωτογραφίες, στις οποίες θα πρέπει να αντιστοιχηθεί μια έξοδος. Όσον αφορά για την έξοδο, εξαρτάται από το πρόβλημα που έχουμε: αν είναι δυαδικό, δηλαδή αν η ιατρική κατάσταση ενός ασθενή είναι παθολογική ή όχι έχουμε έναν κόμβο, ο οποίος θα έχει τιμή 0 ή 1, ανάλογα την έξοδο. Αν όμως το πρόβλημα έχει τρεις πιθανές εξόδους θα έχουμε 2 κόμβους εξόδου. Το σύστημα αυτό βασίζεται στο δυαδικό. [74] [28] [5]

2.5 Βαθιά Μάθηση

Η βαθιά μάθηση είναι ένα κομμάτι της μηχανικής μάθησης, η οποία βασίζεται στα τεχνητά νευρωνικά δίκτυα. Οι αρχιτεκτονικές που περιλαμβάνονται είναι τα βαθιά νευρωνικά δίκτυα, τα αναδρομικά νευρωνικά δίκτυα, καθώς και τα συνελκτικά νευρωνικά δίκτυα, που θα μελετήσουμε μιας και μπορούν να εφαρμοστούν σε φωτογραφίες, αναγνώριση φωνής, επεξεργασία φυσικής γλώσσας, αναγνώριση σε βίντεο, ανάλυση ιατρικών φωτογραφιών, όπως και σε πολλούς άλλους τομείς και έχουν αποφέρει εξαιρετικά αποτελέσματα.

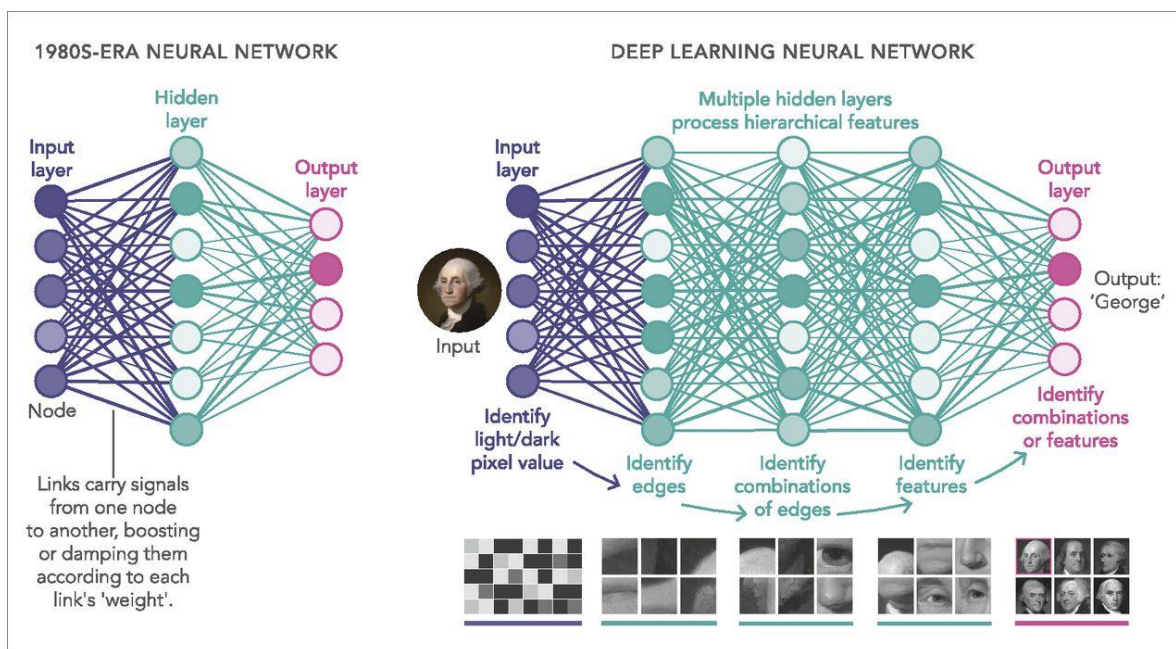
Ο όρος βαθύ προέρχεται από το γεγονός της χρήσης πολλών επιπέδων στο δίκτυο, καθώς έχουμε την είσοδο (input layer), κρυφά επίπεδα (hidden layer) και επίπεδα εξόδου (output layer). Ανάλογα την πολυπλοκότητα του προβλήματος, προσαρμόζουμε τα επίπεδα που θα περιέχει το δίκτυο, δηλαδή πόσοι κόμβοι θα περιέχονται στο κάθε επίπεδο, αλλά το πιο σημαντικό είναι πόσα επίπεδα θα είναι τα κρυφά. [72] [3] [5] [75]

2.6.Βαθιά Νευρωνικά Δίκτυα

Ένα βαθύ νευρωνικό δίκτυο είναι ένα τεχνητό νευρωνικό δίκτυο με πολλαπλά επίπεδα μεταξύ των επιπέδων εισόδου και εξόδου. Υπάρχουν διάφοροι τύποι νευρωνικών δικτύων, αλλά αποτελούνται πάντα από τα ίδια συστατικά: νευρώνες, συνάψεις, βάρη, bias και λειτουργίες ενεργοποίησης. Αυτά τα συστατικά λειτουργούν παρόμοια με τον ανθρώπινο εγκέφαλο και μπορούν να εκπαιδευτούν όπως οποιοσδήποτε άλλος αλγόριθμος μηχανικής μάθησης.

Συγκεκριμένα, ένα βαθύ νευρωνικό δίκτυο δημιουργεί έναν χάρτη εικονικών νευρώνων και εκχωρεί τυχαίες αριθμητικές τιμές, ή αλλιώς λέγονται "βάρη", στις συνδέσεις μεταξύ τους. Τα βάρη και οι εισοδοί πολλαπλασιάζονται και επιστρέφουν μια έξοδο ανάλογη του προβλήματος. Εάν το δίκτυο δεν αναγνωρίζει με ακρίβεια ένα συγκεκριμένο μοτίβο, ένας αλγόριθμος θα προσαρμόσει τα βάρη. Με αυτόν τον τρόπο ο αλγόριθμος μπορεί να κάνει ορισμένες παραμέτρους να συμμετέχουν ή όχι, έως ότου καθορίσει τον σωστό μαθηματικό χειρισμό για την πλήρη επεξεργασία των δεδομένων. Σε αντίθεση με τα βάρη έχουμε και τη μεταβλητή bias, η οποία είναι μια σταθερά που «επιδρά» στον νευρώνα και συμμετέχει στον υπολογισμό της τιμής εξόδου του ανάλογου νευρώνα. [3] [4] [76] [74]

Παρακάτω στην εικόνα 7 φαίνεται η διαφορά, ανάμεσα σε δύο δίκτυα μιας και το δεξί αποτελεί βαθύ δίκτυο, καθώς έχει περισσότερα επίπεδα, σε σχέση με το αριστερό, το οποίο είναι πιο απλό, με ένα μόνο κρυφό στρώμα.



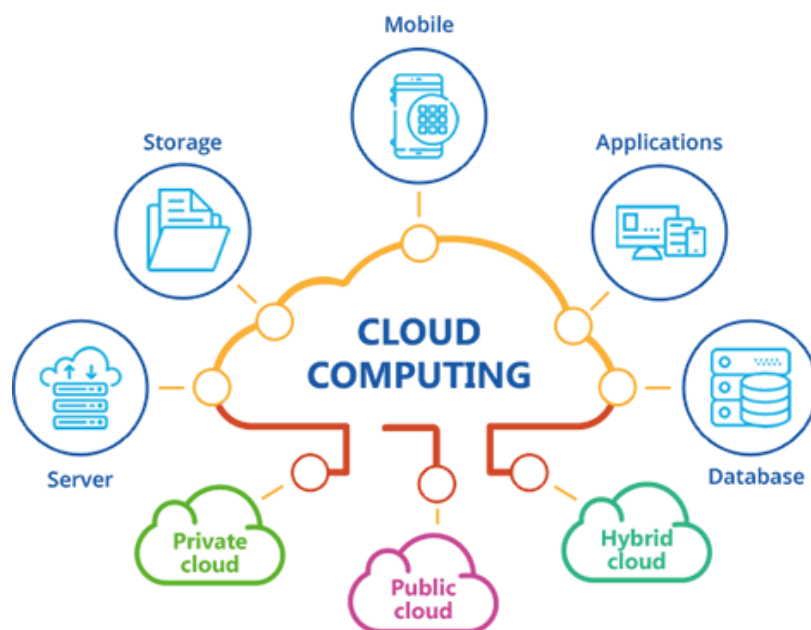
Εικόνα 7

Τα βαθιά νευρωνικά δίκτυα έχουν μεγάλο αριθμό υπερπαραμέτρων και η διερεύνηση του χώρου παραμέτρων για τον εντοπισμό βέλτιστων συνήθως δεν είναι εφικτή λόγω περιορισμών στους υπολογιστικούς πόρους και το χρόνο. Μέθοδοι όπως ο υπολογισμός της διαβάθμισης σε πολλά παραδείγματα ταυτόχρονα, όπου ονομάζεται batch processing, δηλαδή πόσα στοιχεία να διασχίζουν το νευρωνικό δίκτυο σε κάθε πέρασμα, μπορούν να βοηθήσουν στην επιτάχυνση του υπολογισμού. Οι μεγάλες δυνατότητες επεξεργασίας των GPU επέτρεψαν σημαντικές βελτιώσεις στην ταχύτητα εκπαίδευσης. [73] [74]

Η πολυεπίπεδη, ιεραρχική δομή του βαθιού CNN, του δίνει τη δυνατότητα να εξαγάγει χαρακτηριστικά χαμηλού, μεσαίου και υψηλού επιπέδου. Τα χαρακτηριστικά υψηλού επιπέδου (πιο αφηρημένα χαρακτηριστικά) είναι ένας συνδυασμός λειτουργιών χαμηλότερου και μεσαίου επιπέδου. Οι βαθιές αρχιτεκτονικές έχουν συχνά πλεονέκτημα έναντι των ρηχών αρχιτεκτονικών όταν αντιμετωπίζουν πολύπλοκα μαθησιακά προβλήματα. Η στοίβαξη πολλαπλών μονάδων επεξεργασίας ανά στρώμα παρέχει τη δυνατότητα εκμάθησης σύνθετων αναπαραστάσεων. Κατά συνέπεια, σε εργασίες αναγνώρισης που αποτελούνται από εκατοντάδες κατηγορίες εικόνων, τα βαθιά συνελκτικά νευρωνικά δίκτυα έχουν δείξει σημαντική βελτίωση της απόδοσης σε σχέση με τα συμβατικά μοντέλα.

Τα βαθιά συνελκτικά νευρωνικά δίκτυα βασίζονται σε εποπτευόμενους μηχανισμούς μάθησης και, ως εκ τούτου, απαιτείται η διαθεσιμότητα μεγάλων και σχολιασμένων δεδομένων για τη σωστή εκμάθησή του. Η επιλογή υπερπαραμέτρων επηρεάζει ιδιαίτερα την απόδοση του. Μια μικρή αλλαγή στις τιμές των υπερπαραμέτρων μπορεί να επηρεάσει τη συνολική απόδοση. Αυτός είναι ο λόγος για τον οποίο η προσεκτική επιλογή τους είναι ένα σημαντικό ζήτημα σχεδιασμού που πρέπει να αντιμετωπιστεί με κάποια κατάλληλη στρατηγική βελτιστοποίησης. Η αποτελεσματική εκπαίδευση απαιτεί ισχυρούς πόρους υλικού όπως GPU. Ο συνδυασμός πολλαπλών και διαφορετικών αρχιτεκτονικών μπορεί να βοηθήσει το μοντέλο στη βελτίωση της γενίκευσης και της ευρωστίας σε διάφορες κατηγορίες εικόνων.

Η μαθησιακή ικανότητα του συνελκτικού νευρωνικού δικτύου γενικά ενισχύεται αυξάνοντας το μέγεθος του δικτύου και μπορεί να γίνει σε εύλογο χρονικό διάστημα με τη βοήθεια της τρέχουσας προηγμένης τεχνολογίας υλικού όπως κάποιος υπερυπολογιστής. Ωστόσο, η κατάρτιση αρχιτεκτονικών βαθιάς και υψηλής χωρητικότητας εξακολουθεί να αποτελεί σημαντική επιβάρυνση για τη χρήση μνήμης και τους υπολογιστικούς πόρους. Κατά συνέπεια, εξακολουθούμε να απαιτούμε πολλές βελτιώσεις στην τεχνολογία υλικού που μπορούν να επιταχύνουν την έρευνα. Το κύριο μέλημα είναι η δυνατότητα εκτέλεσης χρόνου εκτέλεσης. Στο μέλλον, αναμένεται ότι θα αξιοποιηθεί το δυναμικό των πλατφόρμων που βασίζονται στη τεχνολογία cloud για την ανάπτυξη υπολογιστικών εντάσεων εφαρμογών. Τα βαθιά δίκτυα αποτελούν κρίσιμη πρόκληση για την εφαρμογή και την εκτέλεση τους σε συσκευές περιορισμένης χρήσης πόρων. Το cloud computing όχι μόνο επιτρέπει τη διαχείριση τεράστιων όγκων δεδομένων, αλλά αξιοποιεί επίσης το όφελος της υψηλής υπολογιστικής απόδοσης με αμελητέο κόστος. Παγκόσμιες κορυφαίες εταιρείες όπως η Amazon, η Microsoft, η Google και η IBM προσφέρουν τις δημόσιες εγκαταστάσεις υπολογιστικού νέφους σε υψηλή επεκτασιμότητα, ταχύτητα και ευελιξία για να εκπαιδεύσουν αρχιτεκτονικές που θέλουν πόρους. Επιπλέον, το περιβάλλον cloud καθιστά εύκολη τη διαμόρφωση βιβλιοθηκών τόσο για ερευνητές όσο και για νέους επαγγελματίες. [77] [78] [79]



Εικόνα 8 [80]

Από την εφαρμογή οπτικής και απεικόνισης μέσω υπολογιστών τα συνελκτικά νευρωνικά δίκτυα θεωρήθηκαν πραγματική ανακάλυψη κατά τα έτη 2011 και 2012. Ύστερα από τότε οι μέθοδοι βαθιάς μάθησης, ειδικά τα συνελκτικά νευρωνικά δίκτυα, αλλά και άλλοι τύποι δικτύων (π.χ. επαναλαμβανόμενα νευρωνικά δίκτυα και γενετικά δίκτυα αντιπαλότητας) έχουν αξιοποιηθεί για την αντιμετώπιση υφισταμένων προκλήσεων σε μια σειρά από εργασίες ιατρικής απεικόνισης, Στα περισσότερα από αυτά τα καθήκοντα, είχαν συχνά επιτυχία καθιστώντας το νέο πρότυπο και σημείο αναφοράς στον ερευνητικό τομέα. [75] [81]

2.7 Συνελικτικά νευρωνικά δίκτυα

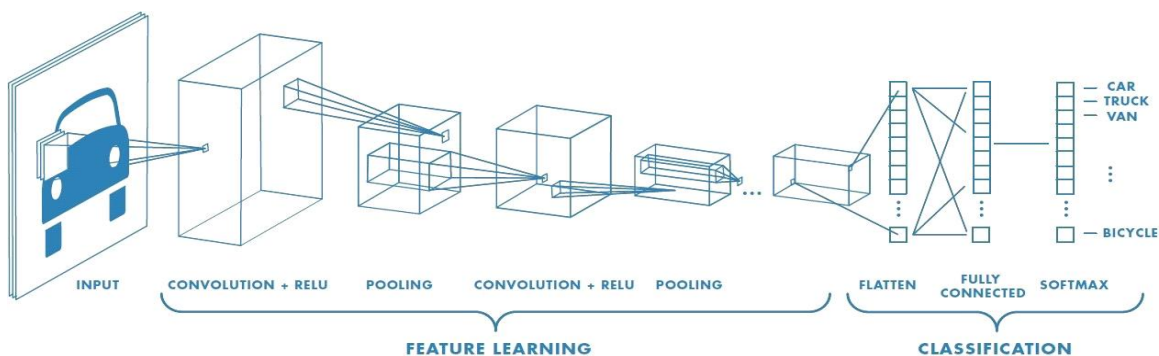
Τα συνελικτικά νευρωνικά δίκτυα είναι μια κλάση των βαθιά νευρωνικών δικτύων, τα οποία χρησιμοποιούνται κυρίως στην ανάλυση φωτογραφιών. Η ονομασία τους προέρχεται από το γεγονός ότι περιέχουν μια μαθηματική πράξη την συνέλιξη, η οποία είναι μια γραμμική πράξη.

Η αρχιτεκτονική τους είναι η κλασική των δικτύων, με είσοδο, έξοδο και πολλά κρυμμένα στρώματα. Τα κρυμμένα στρώματα έχουν ως συνάρτηση ενεργοποίησης, συνήθως την RELU.

Η πράξη της συνέλιξης είναι η απλή εφαρμογή ενός φίλτρου σε μια είσοδο που οδηγεί σε ενεργοποίηση κόμβων. Η επαναλαμβανόμενη εφαρμογή του ίδιου φίλτρου σε μια είσοδο οδηγεί σε έναν χάρτη ενεργοποιήσεων που ονομάζεται χάρτης δυνατοτήτων, υποδεικνύοντας τις θέσεις και την ισχύ ενός ανιχνευμένου χαρακτηριστικού σε μια είσοδο, όπως μια εικόνα. [82]

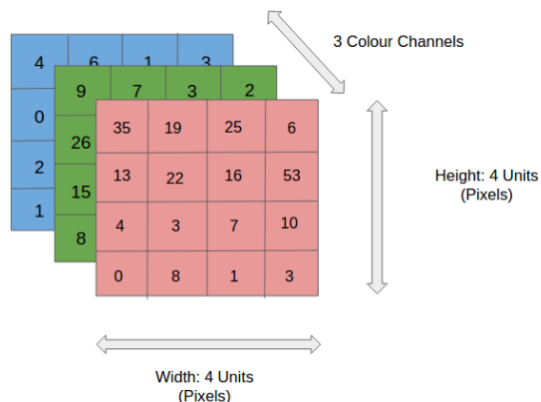
Παράμετροι συνελικτικών δικτύων

Παρακάτω στην εικόνα 9 βλέπουμε μια αναπαράσταση της διαδικασίας που συμβαίνει στα συνελικτικά νευρωνικά δίκτυα.



Εικόνα 9

Στη συνέχεια ακολουθούν λεπτομέρειες για τις παραμέτρους των συνελικτικών νευρωνικών δικτύων:



Εικόνα 10

4x4x3 RGB Image

Αρχικά στην εικόνα 10 βλέπουμε πως βλέπει ένα δίκτυο μια εικόνα. Σε ένα συνελικτικό δίκτυο, η είσοδος είναι με το εξής σχήμα (αριθμός εικόνων) x (ύψος εικόνας) x (πλάτος εικόνας) x (κανάλια εισόδου). Έστω ότι έχουμε μια εικόνα RGB που έχει διαχωριστεί από τα τρία χρωματικά επίπεδα - Κόκκινο, Πράσινο και Μπλε.

Ο ρόλος του συνελικτικού δικτύου είναι να εξάγει τις εικόνες σε μια μορφή που είναι πιο εύκολη στην επεξεργασία, χωρίς να χαθούν χαρακτηριστικά που είναι κρίσιμα για τη λήψη μιας σωστής πρόβλεψης. Αυτό είναι σημαντικό όταν πρόκειται να σχεδιάσουμε μια αρχιτεκτονική που δεν είναι μόνο καλή στην εκμάθηση χαρακτηριστικών αλλά και είναι επεκτάσιμη σε τεράστια σύνολα δεδομένων. [83]

Αριθμός πίξελ

Καταρχάς επιλέγουμε τιμή για τον αριθμό πίξελ, που θα μετατρέψει ο αλγόριθμος όλες τις φωτογραφίες.

Αριθμός batch

Όπως έχει αναφερθεί ο όρος batch είναι πόσα δεδομένα θα διαπερνάνε τον αλγόριθμο κάθε φορά.

Epochs

Ο αριθμός των φορών που θα διαπεράσουν το νευρωνικό δίκτυο.

Παρακάτω στην εικόνα 11 βλέπουμε πως είναι προγραμματιστικά ένα συνελκτικό νευρωνικό δίκτυο. Αρχικά έχουμε το επίπεδο συνέλιξης με αριθμό πυρήνων/φίλτρων, με συνάρτηση ενεργοποίησης και το μέγεθος της είσοδος της εικόνας.



```

model = models.Sequential()
model.add(layers.Conv2D(32, (3, 3), activation='relu', input_shape=(32, 32, 3)))
model.add(layers.MaxPooling2D((2, 2)))
model.add(layers.Conv2D(64, (3, 3), activation='relu'))
model.add(layers.MaxPooling2D((2, 2)))
model.add(layers.Conv2D(64, (3, 3), activation='relu'))
  
```

Annotations in the image:

- Filter**: points to the first `(3, 3)` in the first `Conv2D` layer.
- Kernel size**: points to the second `(3, 3)` in the first `Conv2D` layer.
- Pixel**: points to the `3` in the `input_shape` parameter.
- 3: RGB, 1: Gray**: points to the `3` in the `input_shape` parameter.
- Pooling shape**: points to the `(2, 2)` in the first `MaxPooling2D` layer.

Εικόνα 11 [84]

Αυτές οι διαδικασίες (convolutional neural network – pooling – dropout) επαναλαμβάνονται όσες φορές θεωρείται αρκετό για την καλύτερη απόδοση του δικτύου. Στη συνέχεια ακολουθεί το βήμα-βήμα τι γίνεται σε κάθε φάση.

Convolutional layers

Αυτό το επίπεδο δημιουργεί έναν πυρήνα συνέλιξης συνδυαστικά με την τιμή kernel size και με την είσοδο του επιπέδου για να παράγει την έξοδο. Όταν χρησιμοποιείται αυτό το επίπεδο ως το πρώτο επίπεδο σε ένα μοντέλο, πρέπει να καταχωρίσουμε τη μεταβλητή `input_shape`, όπου είναι ο αριθμός των πίζελ.

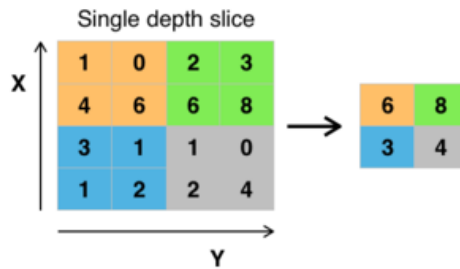
Τα στρώματα αυτά μειώνουν τις διαστάσεις των δεδομένων.

Αριθμός φίλτρων

Δεδομένου ότι το μέγεθος των χαρακτηριστικών μειώνεται με το βάθος, τα επίπεδα κοντά στο επίπεδο εισόδου τείνουν να έχουν λιγότερα φίλτρα ενώ τα υψηλότερα επίπεδα μπορούν να έχουν περισσότερα.

Η μέθοδος που ακολουθείται είναι τοπική ή ολική συνέλιξη, η οποία εξαρτάται από το pooling layer.

Pooling layer



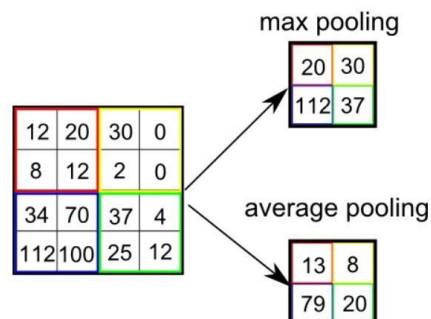
Εικόνα 12

Max pooling with a 2x2 filter and stride = 2

Η διαδικασία, που απεικονίζεται στην εικόνα 12 είναι μια μορφή μη γραμμικής συνάρτησης για τη μείωση διαστάσεων, της υπολογιστικής ισχύος, όπως επίσης για τη μείωση του αριθμού των παραμέτρων και ως εκ τούτου για τον έλεγχο της υπερβολικής τοποθέτησης. Διαχωρίζει την εικόνα εισόδου σε ένα σύνολο μη επικαλυπτόμενων ορθογωνίων και για κάθε τέτοια υποπεριοχή, εξάγει το μέγιστο ή μέσο όρο.

Είναι σύνηθες να γίνεται εισαγωγή ενός συγκεντρωτικού στρώματος μεταξύ διαδοχικών συνελκτικών στρωμάτων (το καθένα συνήθως ακολουθείται από ένα στρώμα ReLU) σε μια αρχιτεκτονική CNN.

Υπάρχουν δύο τύποι Pooling, οι οποίοι φαίνονται και οι δύο στην εικόνα 13: Max Pooling and Average Pooling. Το Max Pooling επιστρέφει τη μέγιστη τιμή από το τμήμα της εικόνας που καλύπτεται από τον πυρήνα. Από την άλλη πλευρά, το μέσο Pooling επιστρέφει το μέσο όρο όλων των τιμών από το τμήμα της εικόνας που καλύπτεται από τον πυρήνα.



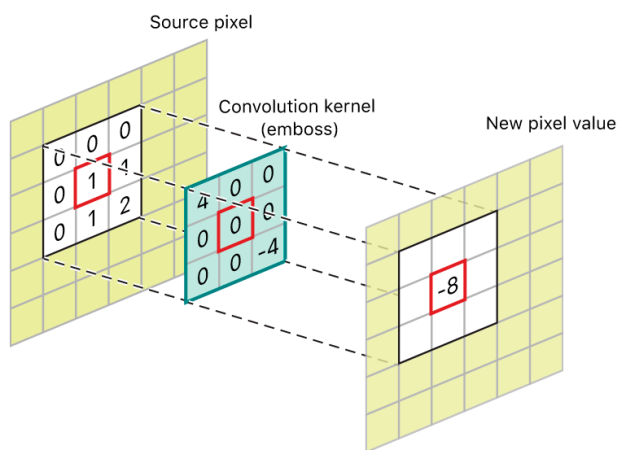
Εικόνα 13

[85],[86],[83]

Pooling shape

Οι τυπικές τιμές είναι 2×2 και απεικονίζουν το παράθυρο επεξεργασίας, στο οποίο θα εφαρμοστεί η συνάρτηση Pooling. Πολύ μεγάλοι όγκοι εισόδου μπορεί να απαιτούν συγκέντρωση 4×4 στα κάτω επίπεδα. Ωστόσο, η επιλογή μεγαλύτερων σχημάτων θα μειώσει δραματικά τη διάσταση του σήματος και μπορεί να οδηγήσει σε υπερβολική απώλεια πληροφοριών.

Ένα σχήμα που συνδυάζει την συνέλιξη με τη μέθοδο pooling είναι το παρακάτω στην εικόνα 14.



Εικόνα 14

Συνάρτηση ενεργοποίησης

Η συνάρτηση ενεργοποίησης, που εφαρμόζεται συνήθως είναι RELU. Αφαιρεί αποτελεσματικά τις αρνητικές τιμές από έναν χάρτη ενεργοποίησης θέτοντάς τις στο μηδέν και δεν επηρεάζει τα δεκτικά πεδία του στρώματος της συνέλιξης.

Άλλες συναρτήσεις χρησιμοποιούνται επίσης για την αύξηση της μη γραμμικότητας, για παράδειγμα είναι η σιγμοειδής συνάρτηση, αλλά προτιμάται στα τελικά στάδια του δικτύου για την εξαγωγή αποτελέσματος. Το ReLU προτιμάται συχνά από άλλες λειτουργίες [87]

Dropout

Επειδή ένα πλήρως συνδεδεμένο στρώμα καταλαμβάνει τις περισσότερες από τις παραμέτρους, είναι επιρρεπές σε υπερβολική τοποθέτηση. Οπότε η μέθοδος dropout είναι πολύ σημαντική για αυτό το σκοπό. Σε κάθε στάδιο εκπαίδευσης, οι μεμονωμένοι κόμβοι είτε παραλείπονται από το

δίκτυο, είτε διατηρούνται, έτσι ώστε να μείνει ένα μειωμένο δίκτυο. Οι εισερχόμενες και εξερχόμενες άκρες σε έναν αποκλεισμένο κόμβο αφαιρούνται επίσης. Μόνο το μειωμένο δίκτυο εκπαιδεύεται στα δεδομένα σε αυτό το στάδιο. Αποφεύγοντας την εκπαίδευση όλων των κόμβων σε όλα τα δεδομένα εκπαίδευσης, μειώνεται αισθητά η υπερβολική τοποθέτηση. Η μέθοδος αυτή βελτιώνει επίσης σημαντικά την ταχύτητα της εκπαίδευσης. Αυτό καθιστά τον συνδυασμό μοντέλου πρακτικό, ακόμη και για βαθιά νευρωνικά δίκτυα. Η τεχνική μειώνει τις αλληλεπιδράσεις κόμβων, οδηγώντας τους να μάθουν πιο ισχυρές λειτουργίες που γενικεύονται καλύτερα σε νέα δεδομένα.

[84]

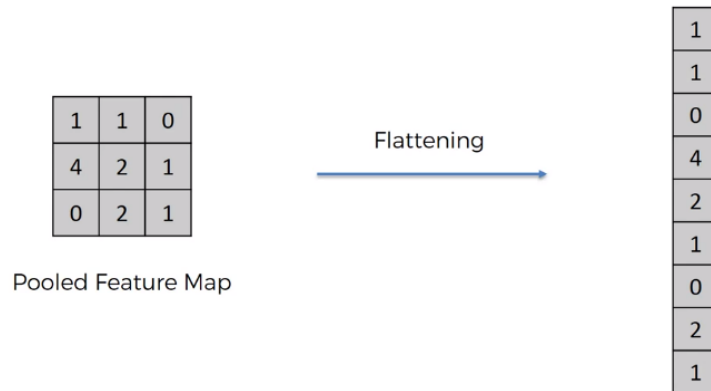
Στη συνέχεια έχουμε το παρακάτω κομμάτι κώδικα, στην εικόνα 15 και απεικονίζει την συνάρτηση Flatten και 2 πλήρως συνδεδεμένα δίκτυα.

```
model.add(layers.Flatten())  
model.add(layers.Dense(64, activation='relu'))  
model.add(layers.Dense(10))
```

Εικόνα 15

Flatten

Η μέθοδος Flatten μετατρέπει τα δεδομένα σε έναν μονοδιάστατο πίνακα για την εισαγωγή τους στο επόμενο επίπεδο. Εξισορροπούν την έξοδο των συνελκτικών επιπέδων για να δημιουργηθεί ένας απλός διανυσματικός πίνακας. Και συνδέεται με το τελικό μοντέλο ταξινόμησης, το οποίο ονομάζεται πλήρως συνδεδεμένο επίπεδο, όπως φαίνεται στην εικόνα 16. [88]



Εικόνα 16

Fully connected layer (Dense)

Αφού παίρνουν σαν είσοδο τον μονοδιάστατο πίνακα τα πλήρως συνδεδεμένα στρώματα συνδέουν κάθε νευρώνα σε ένα στρώμα με κάθε νευρώνα σε ένα άλλο στρώμα. Είναι το ίδιο με ένα απλό νευρωνικό δίκτυο πολλαπλών επιπέδων. Ο μονοδιάστατος πίνακας, που έχει δημιουργηθεί περνά από ένα πλήρως συνδεδεμένο επίπεδο για να ταξινομήσει τις εικόνες. Οι νευρώνες σε ένα πλήρως συνδεδεμένο στρώμα έχουν συνδέσεις με όλες τις ενεργοποιήσεις στο προηγούμενο στρώμα, όπως φαίνεται στα κανονικά (μη συνελκτικά) τεχνητά νευρικά δίκτυα.

Στο βήμα του compile ορίζεται το loss layer, ο optimizer και οι μετρικές, όπως αναπαρίσταται στην εικόνα 17.

```
model.compile(optimizer='adam',
              loss=tf.keras.losses.SparseCategoricalCrossentropy(from_logits=True),
              metrics=['accuracy'])

history = model.fit(train_images, train_labels, epochs=10,
                   validation_data=(test_images, test_labels))
```

Εικόνα 17

Loss layer

Το "επίπεδο απώλειας" καθορίζει πώς η εκπαίδευση θα υπολογίσει την απόκλιση μεταξύ της προβλεπόμενης (εξόδου) και της πραγματικής ετικέτας και είναι συνήθως το τελικό στρώμα ενός νευρωνικού δικτύου. Μπορούν να χρησιμοποιηθούν διάφορες λειτουργίες απώλειας.

Στην τελευταία γραμμή προγραμματιστικά της αρχιτεκτονικής του συνελκτικού νευρωνικού δικτύου ορίζεται η έξοδος με ένα πλήρως συνδεδεμένο επίπεδο, το οποίο περιλαμβάνει τον αριθμό

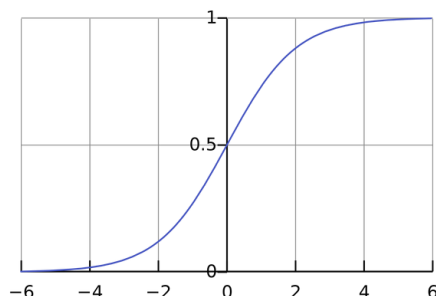
των συνολικών εξόδων (δηλαδή αν είναι δυαδικό `num_classes = 2`) και τη συνάρτηση ενεργοποίησης που στο τελικό επίπεδο χρησιμοποιείται η `sigmoid` ή `softmax`. Αλγοριθμικά απεικονίζεται στην εικόνα 18.

```
model.add(tf.keras.layers.Dense(num_classes, activation='softmax'))
```

Εικόνα 18

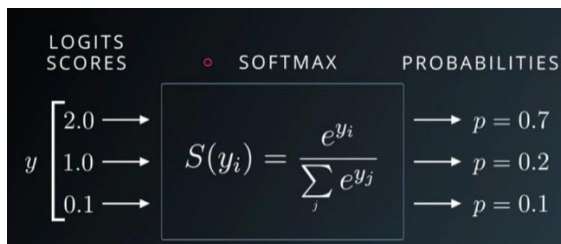
Συνάρτηση ενεργοποίησης

Η συνάρτηση ενεργοποίησης Sigmoid, όπου φαίνεται στην εικόνα 19 είναι μια μαθηματική συνάρτηση, η οποία μετατρέπει την οποιαδήποτε είσοδο σε τιμή μεταξύ του 0 και του 1, καθώς για μικρές τιμές εξόδου, το `sigmoid` επιστρέφει μια τιμή κοντά στο μηδέν και για μεγάλες τιμές το αποτέλεσμα της συνάρτησης πλησιάζει το 1. Για αυτό είναι κατάλληλη για δυαδικά προβλήματα.



Εικόνα 19

Η λειτουργία ενεργοποίησης Softmax είναι μια συνάρτηση ενεργοποίησης που λαμβάνει διανύσματα πραγματικών αριθμών ως εισόδους και τα ομαλοποιεί σε κατανομή πιθανότητας ανάλογη με τα εκθετικά των αριθμών εισόδου. Μετά την εφαρμογή του Softmax, κάθε στοιχείο θα είναι στην περιοχή από 0 έως 1 και τα όλες οι τιμές των στοιχείων εξόδων θα έχουν άθροισμα. Φαίνεται αρκετά ξεκάθαρα στην εικόνα 20.



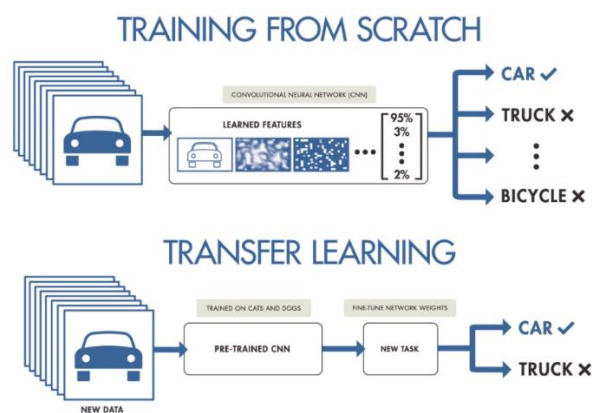
Εικόνα 20

Γενικά, το Softmax χρησιμοποιείται για πολλαπλή ταξινόμηση, ενώ το Sigmoid χρησιμοποιείται για δυαδική ταξινόμηση.[89] [90] [85] [91]

Η αρχιτεκτονική παίζει σημαντικό ρόλο, καθώς είναι υπεύθυνη για τη μείωση του αριθμού των παραμέτρων, που θα συμμετέχουν και στην επαναχρησιμοποίηση των βαρών και για αυτό το λόγο με τη σωστή δομή το δίκτυο μπορεί να εκπαιδευτεί για να κατανοήσει καλύτερα την πολυπλοκότητα της εικόνας. [83]

2.8 Μεταφορά εκμάθησης

Επίσης είναι σημαντικό να αναλυθεί η έννοια της μεταφοράς γνώσης (transfer learning), καθώς εφαρμόστηκε στα δικά μας δεδομένα, με σκοπό την εύρεση του βέλτιστου αλγορίθμου. Όπως βλέπουμε στην εικόνα 21 το πρώτο διάγραμμα απεικονίζει την διαδικασία κατηγοριοποίησης με δίκτυο, που δημιουργήθηκε και προσαρμόστηκε για το σετ το ίδιο και κάνει σωστή πρόβλεψη. Ενώ στο κάτω διάγραμμα βλέπουμε το ήδη εκπαιδευμένο δίκτυο σαν ένα «μαύρο κουτί», με προσαρμοσμένα βάρη, όπου γίνεται επιτυχώς η πρόβλεψη της εξόδους



Εικόνα 21 [92]

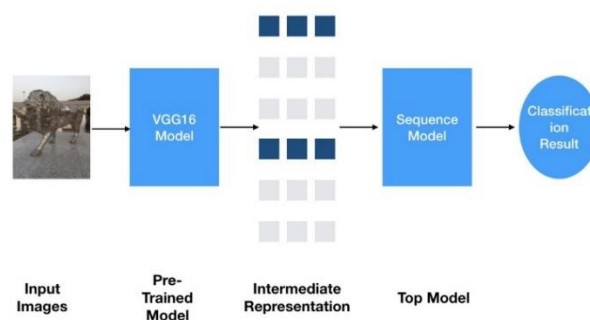
Η μεταφορά εκμάθησης είναι μια μέθοδος στη μηχανική μάθηση που επικεντρώνεται στην αποθήκευση των γνώσεων που αποκτώνται κατά την επίλυση ενός προβλήματος και την εφαρμογή των γνώσεων σε ένα διαφορετικό αλλά σχετικό πρόβλημα. Για παράδειγμα, οι γνώσεις που αποκτήθηκαν κατά την εκμάθηση της αναγνώρισης ενός αντικειμένου θα μπορούσαν να

εφαρμοστούν κατά την προσπάθεια αναγνώρισης κάποιου διαφορετικού αντικειμένου σε ένα άλλο πρόβλημα. Από πρακτική άποψη, η επαναχρησιμοποίηση ή η μεταφορά πληροφοριών από προηγούμενα αναγνωρισμένα καθήκοντα για την εκμάθηση νέων μοτίβων έχει τη δυνατότητα να βελτιώσει σημαντικά την αποτελεσματικότητα.

Το 1976 οι Stevo Bozinovski και Ante Fulgosi δημοσίευσαν ένα άρθρο που αφορούσε ρητά τη μάθηση μεταφοράς σε εκπαίδευση νευρωνικών δικτύων. Το έγγραφο δίνει ένα μαθηματικό και γεωμετρικό μοντέλο της μάθησης μεταφοράς. Το 1981 δόθηκε μια έκθεση σχετικά με την εφαρμογή της μάθησης μεταφοράς στην εκπαίδευση ενός νευρικού δικτύου σε ένα σύνολο δεδομένων εικόνων που αντιπροσωπεύουν γράμματα τερματικών υπολογιστών. Το 2020 ανακαλύφθηκε ότι, λόγω των παρόμοιων φυσικών τους φύσεων, είναι δυνατή η εκμάθηση μεταφοράς μεταξύ των ηλεκτρομυογραφικών (EMG) σημάτων από τους μύες κατά την ταξινόμηση των συμπεριφορών των ηλεκτροεγκεφαλογραφικών (EEG) εγκεφαλικών κυμάτων. Τα πειράματα σημείωσαν ότι η ακρίβεια των νευρωνικών δικτύων και των συνελκτικών νευρικών δικτύων βελτιώθηκε μέσω της εκμάθησης μεταφοράς. Επιπλέον, ο τελικός χρήστης ενός προ-εκπαιδευμένου μοντέλου μπορεί να αλλάξει τη δομή των πλήρως συνδεδεμένων επιπέδων για να επιτύχει μια ανώτερη απόδοση και να το προσαρμόσει στο δικό του πρόβλημα. [93] [94]

Ένα από τα προ-εκπαιδευμένα δίκτυα είναι το VGG-16, το οποίο δέχεται σαν είσοδο τα δεδομένα και σύμφωνα με τα ήδη προσαρμοσμένα βάρη, που έχει μπορεί να εξάγει αξιόπιστα αποτελέσματα, όπως φαίνεται στο παράδειγμα της εικόνας 22.

Transfer Learning



Εικόνα 22 [95]

Κεφ. 3 Βιβλιογραφική ανασκόπηση

3.1 Ταξινόμηση ιατρικών εικόνων στην πυρηνική καρδιολογία

Στο παρόν κεφάλαιο παρατίθεται η βιβλιογραφική ανασκόπηση των συνελκτικών δικτύων, στον τομέα της πυρηνικής ιατρικής με δεδομένα εικόνες SPECT απεικόνισης αιμάτωσης μυοκαρδίου. Είναι ξεκάθαρο ότι πλέον στον τομέα αυτό η εφαρμογή αλγορίθμων μηχανικής μάθησης έχει πρόσφατα επικεντρωθεί στη διάγνωση καρδιαγγειακών προβλημάτων. Η πυρηνική καρδιολογία αντιπροσωπεύει μια συγκεκριμένη θέση για την τεχνητή νοημοσύνη καθώς ασχολείται με πολύπλοκες εικόνες, που αποκτήθηκαν με το SPECT και το PET, δηλαδή υπολογιστική τομογραφία εκπομπών μονών φωτονίων [SPECT] και την τομογραφία εκπομπής ποζιτρονίων [PET]. [96]

Καταγράφηκαν 6 εργασίες που είναι σχετικές με το πρόβλημα της διάγνωσης της καρδιακής νόσου και παραθέτονται παρακάτω. Γενικά στις παρακάτω ερευνητικές μελέτες, οι συγγραφείς χρησιμοποιούν τεχνητά νευρωνικά δίκτυα για τη διάγνωση της καρδιαγγειακής νόσου χρησιμοποιώντας πολικούς χάρτες και διαφορετικές μεταβλητές εισόδου, συμπεριλαμβανομένων κλινικών και ποσοτικών δεδομένων. Υπάρχουν επίσης αρκετές μελέτες που διαγιγνώσκουν τη στεφανιαία νόσο από το SPECT MPI χρησιμοποιώντας έναν ενισχυμένο αλγόριθμο μηχανικής εκμάθησης συνόλου, το φορέα υποστήριξης και τη βαθιά μάθηση. Οι περισσότερες από αυτές τις μελέτες απαιτούν μεγάλο αριθμό εικόνων για εκπαίδευση.

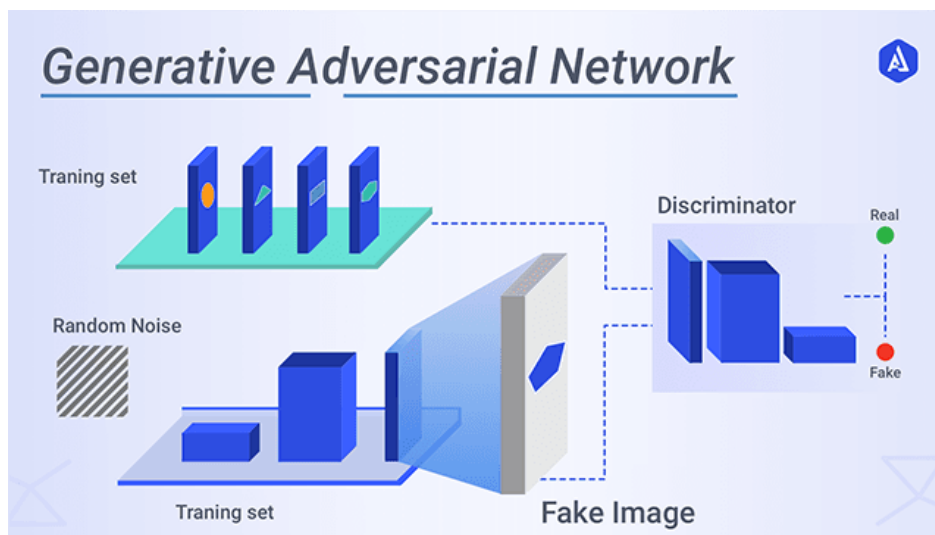
Τα δεδομένα εικόνων στην πυρηνική καρδιολογία απεικονίζουν την κατανομή μιας ποσότητας ραδιοφαρμάκου εντός της καρδιάς και είναι η πλέον προτιμώμενη, καθώς αποτελεί μια οικονομικά αποδοτική μέθοδος για τον εντοπισμό και τον προσδιορισμό της σοβαρότητας του καρδιαγγειακής νόσου πριν από την εμφάνιση της στεφανιαίας αγγειογραφίας. Αυτή η μέθοδος απεικόνισης μειώνει τον αριθμό των περιττών αγγειογραφιών και επιτρέπει τον κατάλληλο σχεδιασμό θεραπείας. Για να δημιουργηθεί ο πολικός χάρτης αιμάτωσης μυοκαρδίου, η περιοχή που φαίνεται η έλλειψη διάχυσης σκιαγραφείται με μαύρα σημεία, τα οποία είναι κάτω από το όριο που ορίστηκε για να θεωρηθεί φυσιολογικό. [97] [98]. Γενικά οι εικόνες αιμάτωσης μυοκαρδίου αποτελούν συνήθως την πρόσθια κατερχόμενη αρτηρία (LAD), την αριστερή

περίμετρο της αρτηρίας (LCx) ή τη δεξιά στεφανιαία αρτηρία (RCA), καθώς έχει θεωρηθεί σημαντική και χρησιμοποιείται ως το πρότυπο για την στεφανιαία νόσο της καρδιάς.

Ακολουθεί αναλυτική περιγραφή των εργασιών:

3.1.1 Εφαρμογή μηχανικής μάθησης στην αξιολόγηση μυοκαρδιακής ισχαιμίας μέσω της πυρηνικής ιατρικής

Αρχικά έχουμε την εργασία [96], η οποία ασχολείται με την αναφορά της θεωρίας γύρω από την τεχνητή νοημοσύνη, την βαθιά μάθηση και την αναφορά διάφορων μεθόδων, οι οποίες έχουν αναπτυχθεί από ερευνητές. Πιο συγκεκριμένα αναφέρει μια προηγμένη μέθοδο, η οποία παράγει φωτογραφίας και βασίζεται πάνω στα Generative Adversarial Networks (GANs), όπου φαίνεται η αναπαράσταση τους στην εικόνα 23. Τα GAN προέρχονται από την ιδέα να υπάρχουν δύο δίκτυα που ανταγωνίζονται μεταξύ τους. Ο στόχος ενός δικτύου (D) είναι η διάκριση μεταξύ πραγματικών και ψεύτικων εικόνων, όπως σε έναν τυπικό ταξινομητή. Ο στόχος του δεύτερου δικτύου (G) είναι η δημιουργία εικόνων, που φαίνονται όσο το δυνατόν πιο κοντά στις πραγματικές εικόνες.



Εικόνα 23 [99]

Και τα δύο δίκτυα ανταγωνίζονται το ένα το άλλο για να γίνουν καλύτερα στην εργασία τους έως ότου το G μπορέσει να δημιουργήσει νέες εικόνες με χαρακτηριστικά ενός συγκεκριμένου πληθυσμού. Αυτό ανοίγει ένα ευρύ φάσμα δυνατοτήτων, επειδή στην ιατρική απεικόνιση μπορεί να είναι δύσκολο να ληφθούν μεγάλα σύνολα δεδομένων εικόνων με συγκεκριμένα

χαρακτηριστικά λόγω του κόστους, αλλά και της συγκατάθεσης του ασθενούς. Κατά συνέπεια, αυτή η τεχνική έχει εφαρμοστεί σε εικόνες MR [100], στους οζώδεις πνεύμονες [98] και στις βλάβες του δέρματος [98], μεταξύ άλλων. Έχουν επίσης διερευνηθεί για τη μετατροπή εικόνων από τη μία εικόνα σε άλλη (π.χ. MR σε CT). [98] Χρησιμοποιήθηκαν επίσης για να βελτιώσουν την ανασυγκρότηση της εικόνας μέσω απενεργοποίησης σε τεχνικές όπως CT [102] και PET [103].

Η δημοσίευση επίσης αναφέρει την ομάδα της Pioneer, όπου χρησιμοποιήθηκαν εικόνες SPECT απεικόνισης αιμάτωσης μυοκαρδίου (MPI) για να τροφοδοτήσουν αλγόριθμους προκειμένου να βελτιωθεί η ακρίβεια της ανίχνευσης CAD σε ένα αυτοματοποιημένο σύστημα. Ο Arsanjani και οι συνεργάτες του εξήγαγαν ποσοτικές και λειτουργικές μεταβλητές από εικόνες MPI και τις συνδύασαν μέσω ενός αλγορίθμου μηχανής υποστήριξης για τη βελτίωση της διάγνωσης της στεφανιαίας στένωσης. [104]

Η απόδοση της μηχανικής μάθησης συγκρίθηκε με την κλασσική ποσοτική ανάλυση. Είναι αξιοσημείωτο ότι ο αλγόριθμος μηχανικής μάθησης SVM (Support Vector Machine) ξεπέρασε όλες τις ποσοτικές αναλύσεις, καθώς και την οπτική ερμηνεία δύο αναγνωστών. Η φυσική συνέχεια ήταν η ενσωμάτωση πρόσθετων δεδομένων για μια πιο εκλεπτυσμένη ανάλυση.

Αυτή ήταν η πρώτη μελέτη που συνδύασε με επιτυχία πολλά χαρακτηριστικά για να βελτιώσει τη διαγνωστική ακρίβεια του SPECT MPI. Μια διαδοχική μελέτη συγκέντρωσε κλινικά δεδομένα που εκτείνονται από την ηλικία και το φύλο έως την πιθανότητα απόφραξης μυοκαρδίου χρησιμοποιώντας τα κριτήρια Diamond-Forrester.

Baseline demographics by Diamond-Forrester classification

	Diamond-Forrester Classification		p-value
	Atypical/Non-cardiac (n=163)	Typical (n=46)	
Age - years	45 ± 9.9	46 ± 8.3	0.55
Female	103 (63.2)	32 (69.6)	0.43
Diabetes mellitus	17 (10.4)	5 (10.9)	0.93
Hypertension	60 (36.8)	19 (41.3)	0.58
Stroke	1 (0.6)	3 (6.5)	0.03
Peripheral vascular disease	0	1 (2.2)	0.22
Hyperlipidemia	23 (14.1)	15 (32.6)	0.004
Tobacco use	40 (24.5)	10 (21.7)	0.69
Family history of CAD	76 (46.6)	27 (58.7)	0.15
GERD	33 (20.2)	15 (32.6)	0.08
Depression	31 (19)	10 (21.7)	0.68
Anxiety	39 (23.9)	12 (26.1)	0.76
Chronic pain	17 (10.4)	6 (13)	0.62
Aspirin	9 (5.5)	5 (10.9)	0.20
Statin	8 (4.9)	6 (13)	0.05
Beta blocker	9 (5.5)	4 (8.7)	0.43

All data expressed as no. (%) or mean ± standard deviation.

Εικόνα 24 [105]

και τα συνδυάζει με ποσοτικές μεταβλητές διάχυσης μέσω ενός αλγορίθμου μηχανικής μάθησης, που ονομάζεται LogitBoost και φαίνονται στην εικόνα 24. [104]

Το πιο σημαντικό είναι να αναφερθεί ότι υπάρχει βελτιωμένη ακρίβεια της διάγνωσης στεφανιαίας νόσου σε σύγκριση με την εκδοχή της μεθόδου της μηχανικής μάθησης που δεν χρησιμοποίησε τα κλινικά δεδομένα. Αυτό ήταν ένα κρίσιμο εύρημα, καθώς οι μη ποσοτικές, κλινικές μεταβλητές είχαν σημαντική σημασία.

Ωστόσο, μια περαιτέρω ανακάλυψη επιτεύχθηκε λίγα χρόνια αργότερα, όταν μια μελέτη ανέπτυξε ένα μοντέλο βαθιάς μαθήσεως ικανό να εντοπίσει ανωμαλίες απευθείας από τις εικόνες SPECT και να μιμείται την ερμηνεία των εμπειρογνομόνων. [106] Πιο συγκεκριμένα, χρησιμοποιήθηκαν

εικόνες stress και rest και δημιουργήθηκε ένα λογισμικό που θα μπορούσε να ανιχνεύσει περισσότερες από 5000 υποψήφιες περιοχές, οι οποίες αξιολογήθηκαν από πυρηνικούς ιατρούς ως μη φυσιολογικές ή φυσιολογικές. Αυτές οι τελικές αξιολογήσεις των υποψηφίων περιοχών χρησιμοποιήθηκαν ως είσοδος για την εκπαίδευση του νευρωνικού δικτύου, μαζί με άλλες κλινικές μεταβλητές. Το μοντέλο έχει ως έξοδο μια τιμή από 0 έως 1 για να εκτιμήσει την πιθανότητα ελαττωμάτων ή ισχαιμίας σε εικόνες SPECT και εξήγαγε υψηλή αποτελέσματα, όχι μόνο όσον αφορά την ανάλυση ROC, αλλά και τις κρίσεις των εμπειρογνομόνων.

Επίσης, ο Bentacur και οι συνεργάτες του χρησιμοποίησαν ακατέργαστους και ποσοτικούς πολικούς χάρτες από εικόνες SPECT MPI για να τροφοδοτήσουν και να εκπαιδεύσουν ένα μοντέλο Deep Learning. [97] Σε ένα πρώτο βήμα, ένα τέτοιο μοντέλο εξήγαγε πολλά χαρακτηριστικά από τις εικόνες που στη συνέχεια μεταφέρθηκαν σε ένα δεύτερο βήμα, αποτελούμενο από τρία πλήρως συνδεδεμένα επίπεδα. Κάθε στρώμα έδωσε μια έξοδο, η οποία αντιστοιχούσε σε βαθμολογία διάγνωσης της αποφρακτικής νόσου. Η μελέτη ανέφερε τη βελτίωση της διάγνωσης CAD σε σύγκριση με την τυπική ποσοτική ανάλυση.

Εκτός αυτού, ενσωμάτωσαν ένα μεγάλο αριθμό από δεδομένα SPECT, κλινικά και δεδομένα stress για να αξιολογήσουν την προγνωστική αξία του αλγορίθμου βαθιάς μάθησης τους [97]. Όπως αναμενόταν, ο αλγόριθμος που συνδύασε όλα τα δεδομένα είχε ως αποτέλεσμα καλύτερη απόδοση, σε σχέση με τις κλασσικές μεθόδους, αλλά και τη μέθοδο μηχανικής μάθησης με δεδομένα απεικόνισης μόνο, τονίζοντας για άλλη μια φορά τη σημασία της ενσωμάτωσης κλινικών μεταβλητών.

Επίσης, δοκιμάστηκαν διαφορετικοί αλγόριθμοι μηχανικής μάθησης για την πρόβλεψη της καρδιακής θνησιμότητας και συγκρίθηκαν με τις παραδοσιακές στατιστικές προσεγγίσεις. [107]. Όχι μόνο όλοι οι αλγόριθμοι ξεπέρασαν τους παραδοσιακούς, αλλά και ένας ειδικός αλγόριθμος μηχανικής μάθησης ονομαζόμενος LASSO που εκπαιδεύτηκε να συρρικνώνει τις μεταβλητές εισόδου και ήταν ακόμα καλύτερος, από ένα απλό λογισμικό μοντέλο με 100 μεταβλητές. Αυτό επισημαίνει τις δυνατότητες στις μεθόδους της μηχανικής μάθησης στην ακριβή επιλογή των πιο σημαντικών μεταβλητών.

Ένα βήμα παραπέρα επιτεύχθηκε από τον Juarez-Orozco και τους συναδέλφους του.[108] Χρησιμοποιήθηκε ένα μοντέλο LogitBoost εκπαιδευμένο με κλινικές τιμές για να προβλεφθεί ισχαιμία του μυοκαρδίου όπως εκτιμάται από τις εικόνες PET. Συγκρίθηκε ανάλυσή τους απευθείας με μια παραδοσιακή προσέγγιση λογιστικής παλινδρόμησης. Ο αλγόριθμος μηχανικής μάθησης ξεπέρασε τα αποτελέσματα σε σύγκριση με τη λογιστική παλινδρόμηση. Αυτό πιθανότατα θα είναι επωφέλές στην κλινική πρακτική για μια πιο ακριβή αξιολόγηση ασθενών με γνωστή ή υποψία ισχαιμίας μέσω αυτοματοποίησης της ανίχνευσης ισχαιμίας.

Οι κλινικοί γιατροί συχνά αναμένουν την επιτυχία του καλύτερου προγνωστικού αποτελέσματος, με σκοπό τη μείωση περιττών διαδικασιών, καθώς και την αποφυγή των μη επιθυμητών αποτελεσμάτων. Συνεπώς, οι επιστήμονες έχουν αποδεχτεί ότι υπάρχει ένα πιθανό όφελος από την εφαρμογή συστημάτων υποστήριξης τεχνητής νοημοσύνης για μεταβλητές και απεικονίσεις ενσωμάτωση δεδομένων. Στην εικόνα 24, φαίνεται ότι οι συμβατικές μέθοδοι, όπως SVM, KNN είναι αξιόπιστα για δομημένα αριθμητικά δεδομένα, αλλά όταν έχουμε φωτογραφίες ως δεδομένα, τότε είναι κατάλληλη η βαθιά μάθηση, καθώς είναι πιο περίπλοκα δεδομένα.

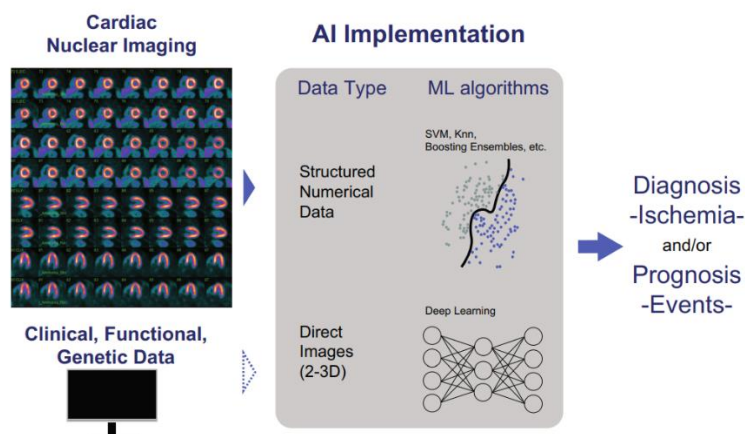


Fig. 3 Workflow of nuclear cardiology for data or image processing and complementary variable integration

Εικόνα 24 [96]

3.1.2 Κατηγοριοποίηση πολικών χαρτών με δεδομένα απεικόνισης καρδιακής αιμάτωσης με γραφικά συνελκτικά νευρωνικά δίκτυα

Στην επόμενη εργασία [109] παρουσιάζεται η μέθοδος των συνελκτικών νευρωνικών δικτύων ύστερα από την εφαρμογή αρκετών ερευνητών. Παρόλα αυτά η παρούσα δημοσίευση παρουσιάζει ως υλοποίηση για την αυτόματη ταξινόμηση των πολικών χαρτών τα graph convolutional neural networks.

Ακολουθούν διάφορες προσεγγίσεις που πρότειναν ερευνητές. Ένας εξ αυτών χρησιμοποίησε ένα δίκτυο πολλών επιπέδων με 16×16 κόμβους εισόδου, 100 κρυμμένους κόμβους και οκτώ κόμβους εξόδου για να ταξινομήσει τους ψηφιοποιημένους και δειγματοληπτικούς πολικούς χάρτες σε φυσιολογικούς ή παθολογικούς. Αυτή η προσέγγιση έχει αναφερθεί ότι η επιτευχθείσα απόδοση αναγνώρισης ήταν καλύτερη από την οπτική εκτίμηση ενός ακτινολόγου, αλλά χειρότερη από έναν έμπειρο ακτινολόγο. Σε μία διαφορετική έρευνα χρησιμοποιήθηκε παρόμοια αρχιτεκτονική δικτύου με 45 κόμβους εισόδου, 15 κρυμμένους κόμβους και έναν κόμβο εξόδου που υποδεικνύει την παρουσία της αποφρακτική νόσου. Επίσης παρόμοια προσέγγιση με τους προηγούμενους ήταν η χρήση τεχνικού νευρωνικού δικτύου με κόμβους εισόδου 15×40 , 5 κρυμμένους κόμβους και έναν κόμβο εξόδου. Όλες αυτές οι μέθοδοι έχουν κοινά ότι χρησιμοποιούν απευθείας εκδόσεις των αρχικών πολικών χαρτών ως είσοδο σε ένα νευρωνικό δίκτυο 3 επιπέδων με ένα κρυφό στρώμα. Ως εκ τούτου, ονομάζουμε αυτές τις προσεγγίσεις άμεσες. Ένας ερευνητής επέκτεινε αυτήν την ιδέα και πρότεινε μια μέθοδο, όπου χρησιμοποίησε ακόμη περισσότερες εισόδους, δηλαδή upright (όρθια θέση) / supine (ανάσκελα σε κάθισμα) TPD, αλλαγές stress / rest, παροδική ισχαιμική διαστολή, ηλικία, φύλο και πιθανότητα CAD μετά το ηλεκτροκαρδιογράφημα, ως είσοδο σε έναν ενισχυμένο ταξινομητή συνόλου. Η διαγνωστική ακρίβεια του ταξινομητή στην τελευταία μελέτη έχει αποδειχθεί ότι είναι ισοδύναμη ή ελαφρώς καλύτερη από αυτήν που επιτεύχθηκε από δύο συμμετέχοντες εμπειρογνώμονες, αντίστοιχα. Επίσης, προτάθηκε μια υβριδική μέθοδος βασισμένη σε βαθιά συνελκτικά νευρωνικά δίκτυα, η οποία λαμβάνει υπόψη τόσο τους μη επεξεργασμένους όσο και τους ποσοτικούς (με βάση την TPD) πολικούς χάρτες για την πρόβλεψη της απόφραξης στεφανιαίας αορτής. αποφρακτικής στένωσης.

Παρόλα αυτά, σε αυτή τη μελέτη, αξιολογούνται οι δυνατότητες του Graph Convolutional Neural Networks (GCNNs) για την εργασία ταξινόμησης πολικών χαρτών που λαμβάνονται από MPI, σε φυσιολογικό και μη φυσιολογικό περιστατικό. Καθώς αυτοί οι πολικοί χάρτες είναι εικόνες με τιμές έντασης που είναι διατεταγμένες σε ένα πολικό πλέγμα, τα GCNN παρέχουν έναν φυσικό τρόπο εφαρμογής συνελκτικών νευρωνικών δικτύων σε πολικούς χάρτες χωρίς την ανάγκη επαναληπτικής δειγματοληψίας τους, σε αντίθεση με τα προηγούμενα έργα που χρησιμοποιούν τα νευρωνικά δίκτυα, τα οποία απαιτούν την επαναληπτική δειγματοληψία.

Στην εικόνα 26, έχουμε τρία ξεχωριστά σχήματα

Α) Πολικός χάρτης που προέρχεται από MPI.

β) Γράφημα που κωδικοποιεί τη χωρική σχέση των τιμών έντασης του πολικού χάρτη που απεικονίζεται στο (α).

γ) Ανατομική διαίρεση που δείχνει που αντιστοιχούν τα διαφορετικά τμήματα του μυοκαρδίου.

Προσαρμοσμένο από τους Cerqueira et al.

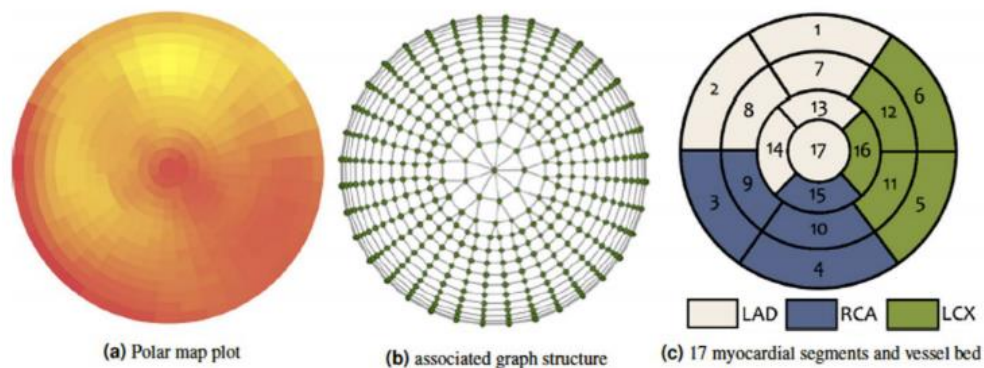


Figure 1. (a) Polar map derived from MPI. (b) Graph encoding the spatial relationship of the intensity values of the polar map depicted in (a). (c) Anatomical division showing to which major vessel bed each myocardial segment is assigned to. Adapted from Cerqueira et al.²³.

Εικόνα 26

Επίσης, μια διαφορετική απεικόνιση των πολικών χαρτών είναι στην εικόνα 27, όπου έχουμε θερμικούς χάρτες που υποδεικνύουν τη συχνότητα των υποδεδειγμένων παθολογικών τμημάτων και τη συχνότητα των τμημάτων που ανιχνεύονται ως παθολογικά. Η γραμμή χρωμάτων υποδεικνύει τη συχνότητα των ανιχνεύσεων σε μια συγκεκριμένη περιοχή:

- Υψηλή συχνότητα - κόκκινο,

- χαμηλή συχνότητα - κίτρινο



Figure 3. Heatmaps indicating the frequency of annotated pathological segments (left) and the frequency of segments being detected as pathological. Color bar indicates frequency of detections in a particular region: high frequency - red, low frequency - yellow.

Εικόνα 27

Οι πολικοί χάρτες που προέρχονται από MPI μπορούν να τροφοδοτηθούν απευθείας σε συνελκτικά νευρωνικά δίκτυα, τα οποία βασίζονται σε αυτές τις γενικευμένες λειτουργίες συνέλιξης. Μια άλλη σημαντική διαφορά στην παρούσα έρευνα είναι το γεγονός ότι η προτεινόμενη μέθοδος λαμβάνει μόνο τους πολικούς χάρτες ως είσοδο χωρίς πρόσθετα δεδομένα. Ο κύριος στόχος αυτής της μελέτης είναι να διερευνήσει εάν, και σε ποιο βαθμό, η γνώση ενός έμπειρου στον τομέα αυτό, που αξιολογεί αυτούς τους πολικούς χάρτες μπορεί να αποκτηθεί από τέτοια δίκτυα.

Για την ταξινόμηση CAD αξιολογήθηκαν ένα δίκτυο βασισμένο σε πολώνυμα Chebyshev, ένα βασισμένο σε πολώνυμα Cayley, ένα πλήρως συνδεδεμένο δίκτυο (FCN), καθώς και ένα συνελκτικό νευρωνικό δίκτυο ως βασικές μεθόδους. Πρώτα, αξιολογήθηκαν τα μοντέλα με την ικανότητά τους να ταξινομήσουν έναν δεδομένο πολικό χάρτη σε μη φυσιολογικό. Προκειμένου να αξιοποιηθεί πλήρως το σύνολο δεδομένων και να επικυρωθεί η απόδοση των μοντέλων, πραγματοποιήθηκε 4fold cross validation, δηλαδή το αρχικό σετ δεδομένων χωρίστηκε σε 4 σετ μοντέλων και εκπαιδεύτηκε το μοντέλο χρησιμοποιώντας τα 3 σετ δεδομένων και στη συνέχεια αξιολογήθηκε με το υπολειπόμενο 1. Ύστερα, από την έρευνα, η χρήση του GCNN με Chebyshev filters έδειξε σημαντικές βελτιώσεις και για αυτό το λόγο συγκρίθηκε με ένα πλήρως συνδεδεμένο

νευρωνικό δίκτυο και ένα νευρωνικό δίκτυο με λίγους κόμβους. Ενώ θα μπορούσαν να χρησιμοποιηθούν πιο περίπλοκες δομές δικτύου, θα διέτρεχαν μεγαλύτερο κίνδυνο overfitting (υπερ-τοποθέτησης). Δημιουργήθηκαν δύο πολικοί χάρτες για κάθε ασθενή και έγινε η κατηγοριοποίηση, κατά την ερμηνεία του συνολικού αποτελέσματος του ασθενούς. Κρίθηκε σωστό να γίνει εκπαίδευση για δύο εργασίες: ταξινόμηση rest, δηλαδή να αξιολογηθεί εάν κατάσταση rest είναι μη φυσιολογική ή όχι και ταξινόμηση stress, δηλαδή να εκτιμηθεί εάν η κατάσταση stress είναι ανώμαλη ή όχι. Ως βήμα προεπεξεργασίας, κάθε πολικός χάρτης κανονικοποιήθηκε με την υψηλότερη τιμή διάχυσης και δεν πραγματοποιείται η αύξηση δεδομένων (Image data augmentation).

Οι μέθοδοι που υλοποιήθηκαν στην έρευνα αυτή, αναλύονται παρακάτω.

- 1D Πλήρως συνδεδεμένο νευρωνικό δίκτυο. Πρώτον, εξετάστηκε η συμβατική προσέγγιση. Μετατράπηκε ο αρχικός πολικός χάρτης σε μονοδιάστατο πίνακα και χρησιμοποιήθηκε ως είσοδο στο πλήρως συνδεδεμένο νευρωνικό δίκτυο 3 επιπέδων με 460 κρυμμένους κόμβους και δύο κόμβους εξόδου. Αυτή η προσέγγιση χρησιμοποιεί τις πλήρεις πληροφορίες του πολικού χάρτη χωρίς υπο-δειγματοληψία ή ποσοτικοποίηση των τιμών έντασης. Μια λειτουργία ενεργοποίησης Rectified Linear Unit (ReLU) χρησιμοποιείται για τους νευρώνες στο κρυφό στρώμα και μια λειτουργία ενεργοποίησης softmax χρησιμοποιείται για το επίπεδο εξόδου. Η categorical cross entropy (κατηγοριοποιημένη εγκάρσια εντροπία) χρησιμοποιείται ως συνάρτηση απώλειας.
- 2D Συνελικτικό νευρωνικό δίκτυο. Σε αυτό το μοντέλο, δημιουργήθηκε ένα συνελικτικό δίκτυο αναδιαμορφώνοντας τον πολικό χάρτη σε ένα δισδιάστατο πίνακα μεγέθους 23x20 pixel. Η αρχιτεκτονική στο CNN απασχολεί μόνο 2 στρώσεις συνθετικών φίλτρων με μέγεθος πυρήνα 5x5 στο πρώτο στρώμα και 3x3 στο δεύτερο. Κάθε στρώμα ακολουθήθηκε από μια λειτουργία ενεργοποίησης ReLU, καθώς και ένα στρώμα μέγιστης συγκέντρωσης με παράθυρο 2x2 pixel. Το τελευταίο στρώμα δικτύου είναι ένα πλήρως συνδεδεμένο επίπεδο με softmax λειτουργία ενεργοποίησης. Η categorical cross entropy (κατηγοριοποιημένη εγκάρσια εντροπία) χρησιμοποιείται ως συνάρτηση απώλειας.

Στην εικόνα 28 έχουμε αναπαράσταση μιας αρχιτεκτονικής Graph Convolutional Neural Network. Η είσοδος του δικτύου είναι τα δεδομένα του πολικού χάρτη, τα οποία αντιστοιχούνται σε ένα γράφημα που κωδικοποιεί τις γειτονικές σχέσεις, μεταξύ των επιμέρους σημείων δεδομένων. Χρησιμοποιούνται δύο στρώματα περιστροφής, τα οποία πλέκονται το καθένα, ακολουθούμενο από στρώματα συγκέντρωσης γραφήματος. Κατά την εκπαίδευση με τα πολώνυμα Chebyshev χρησιμοποιούμε φίλτρα τάξης 16 στο πρώτο στρώμα και 32 στο δεύτερο. Αντίστοιχα με τα πολώνυμα Cayley της τάξης 4 και 6 χρησιμοποιούνται στο πρώτο και δεύτερο στρώμα αντίστοιχα. Και τα δύο μοντέλα χρησιμοποιούν μέγεθος συγκέντρωσης 4 στο πρώτο στρώμα και 2 στο δεύτερο.

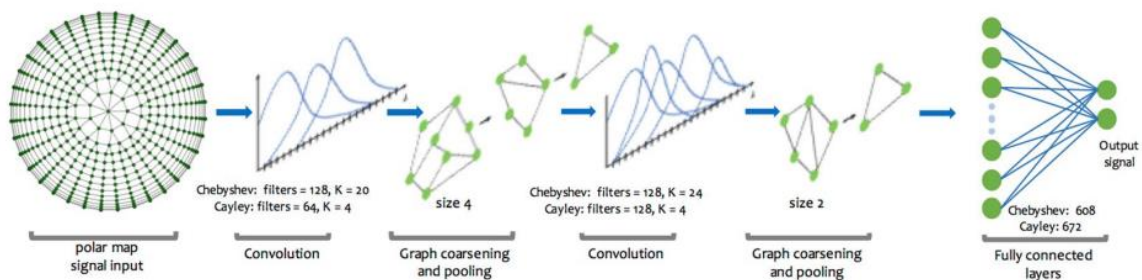


Figure 5. Graph Convolutional Neural Network Architecture. The network's input is the the polar map data defined on a graph that encodes the neighborhood relationships between the individual data points. We use two layers of convolution filter each followed by graph coarsening and pooling layers. When training with Chebyshev polynomials we use filters of order 16 in the first layer and 32 in the second. Cayley polynomials of order 4 and 6 are used in the first and second layer respectively. Both models employ a pooling size of 4 in the first layer and 2 in the second.

Εικόνα 28

- Συνελικτικό νευρωνικό δίκτυο που βασίζεται σε γραφήματα χρησιμοποιώντας πολώνυμα chebyshev. Και στις δύο προηγούμενες υλοποιήσεις δεν λαμβάνεται υπόψη η χωρική διάταξη των δεδομένων. Αντιμετωπίζεται αυτό το πρόβλημα αναπαριστώντας τον πολικό χάρτη ως μη κατευθυνόμενο γράφημα. Η συγκέντρωση γίνεται με τη συμπίκνωση του

γραφήματος χρησιμοποιώντας τον αλγόριθμο Graclus21. Η αρχιτεκτονική GCCN, που απεικονίζεται

αποτελείται από 2 συνελκτικά στρώματα, το καθένα ακολουθούμενο από ένα στρώμα συγκέντρωσης και μια λειτουργία ενεργοποίησης ReLU. Το δίκτυο εκπαιδεύεται με μια λειτουργία απώλειας εντροπίας.

Τα καλύτερα αποτελέσματα εξήγαγε το GCNN μοντέλο χρησιμοποιώντας Chebyshev φίλτρα με την υψηλότερη απόδοση για rest και stress φωτογραφίες, με τιμή 89.3%.

3.1.3 Ανάλυση όρθιων και πρόσθιων spect δεδομένων με μεθόδους βαθιάς μάθησης για την πρόβλεψη στεφανιαίας νόσου

Σε αυτήν την εργασία [110], ο στόχος ήταν η εφαρμογή μεθόδων βαθιάς μάθησης σε upright και supine εικόνες που αποκτήθηκαν για τον ίδιο ασθενή σε μια κάμερα υψηλής απόδοσης.

Οι γιατροί έπρεπε να εξετάσουν αυτές τις 2 εικόνες, που ήταν σε μορφή πολικών χαρτών για την τελική ερμηνεία πριν περάσουν στο βήμα της υλοποίησης αλγορίθμου. Τα δεδομένα που χρησιμοποιήθηκαν στην έρευνα είναι δείγματα πολικών χαρτών που προέρχονται από τις ακατέργαστες εικόνες, που δείχνουν τις κατανομές ραδιοανιχνευτών κανονικοποιημένες στις μέγιστες μετρήσεις και χρησιμοποιήθηκαν μόνο εικόνες stress και όχι rest. . Κατά τη διαδικασία του αλγορίθμου, χρησιμοποιήθηκε 4-fold cross validation. Οι προβλέψεις για κάθε σετ συγχωνεύτηκαν για να υπάρξει μια συνολική εκτίμηση της απόδοσης [111]. Οπότε η πρόβλεψη ήταν πλήρως γενικευμένη, καθώς το νευρωνικό δίκτυο εκπαιδεύτηκε με δεδομένα που δεν είχε προσπελάσει προηγουμένως[112].

Οι εικόνες SPECT ποσοτικοποιήθηκαν με φυσιολογικά όρια όσον αφορά τη διάχυση. Οι πολικοί χάρτες καθορίζουν τις περιοχές της υπερ-σύντηξης. Το φύλο των ασθενών συμπεριλήφθηκε για να ληφθούν υπόψη οι διαφορές μεταξύ ανδρών και γυναικών. Το μοντέλο εκτιμά την αποφρακτική πιθανότητα της στεφανιαίας νόσου CAD χωρίς τη χρήση προκαθορισμένων στεφανιαίων περιοχών ή οποιαδήποτε άλλη είσοδο στο μοντέλο.

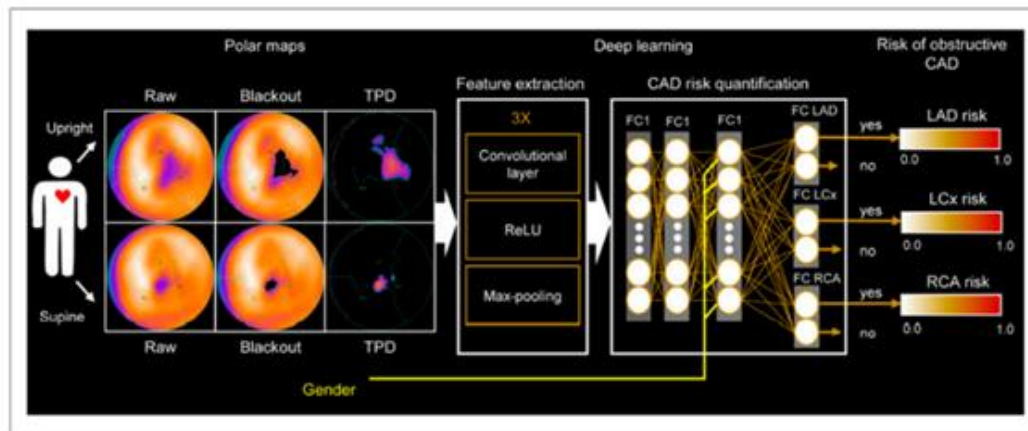


FIGURE 1. DL prediction of obstructive CAD from upright and supine MPI. A deep convolutional neural network trained from obstructive stenosis correlations by ICA was used to simultaneously estimate probability of obstructive CAD for LAD, LCx, and RCA territories from upright and supine polar MPI maps. Maximum probability was retained as probability of patient disease. FC = fully connected layer; Max-pooling = function that returns maximum value for image patch; ReLU = rectified linear unit.

Εικόνα 29

Στην εικόνα 29 βλέπουμε ότι έχουμε δεδομένα πολικών χαρτών, όπου είναι σε όρθια και πρόσθια όψη. Αν κάποιο σημείο βρεθεί ότι βρίσκεται κάτω από το όριο, ορίζεται ως παθολογικό και γίνεται blackout. Από εκεί και έπειτα εισάγονται στη διαδικασία εξαγωγής χαρακτηριστικών, μέσω του μοντέλου βαθιάς μάθησης και τέλος έχουμε σαν αποτέλεσμα την πιθανότητα των περιοχών LAD, LCX, RCA να είναι παθολογικές.

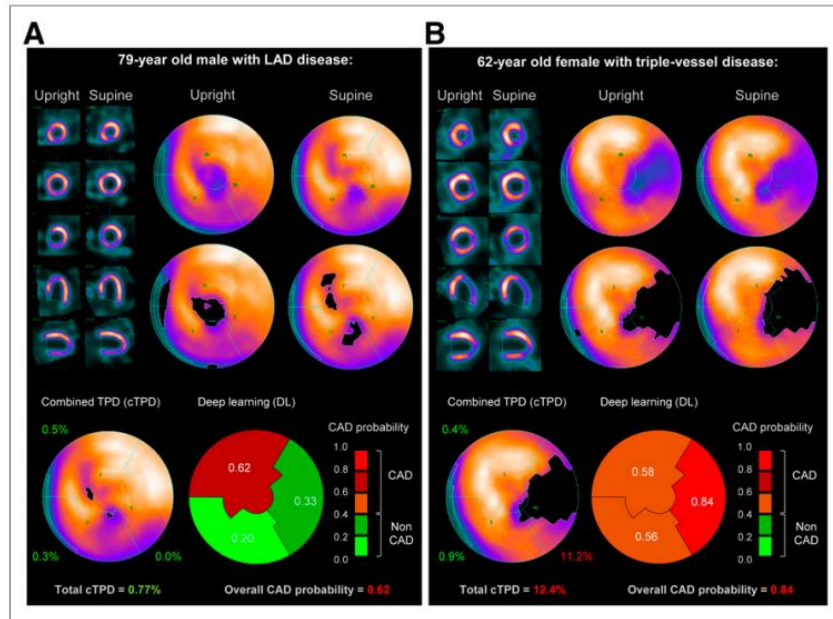


FIGURE 6. Prediction of obstructive CAD from upright and supine stress MPI. Short/long axis views, polar maps depicting normalized radiotracer count distribution and perfusion defects (top), and predictions by cTPD and DL (bottom) are shown for 2 patients with obstructive CAD. (A) In 79-y-old man (85% proximal LAD stenosis) quantified with normal cTPD (per-patient cTPD < 3% and per-vessel cTPD < 1%), DL correctly identified LAD disease. Patient had body mass index of 30 kg/m² and diabetes and underwent exercise stress MPI. (B) In 62-y-old woman (70% mid LAD stenosis, 95% proximal LCX stenosis, and 80% proximal RCA stenosis) with cTPD abnormal for 1 vessel only, DL correctly identified triple-vessel disease. Patient had body mass index of 25 kg/m², dyslipidemia, and family history of cardiac disease and underwent exercise stress MPI. BMI = body mass index.

Εικόνα 30

Η προτεινόμενη μέθοδος είναι αυτόνομη και χρησιμοποιεί μόνο δεδομένα εικόνας, συνδυάζοντας βέλτιστα την κανονική και «σκιασμένη» όψη, κάτι που είναι δύσκολο οπτικά, όπως φαίνεται στην εικόνα 30, όπου παρατηρούμε το «σκιασμένο» σημείο στους πολικούς χάρτες.

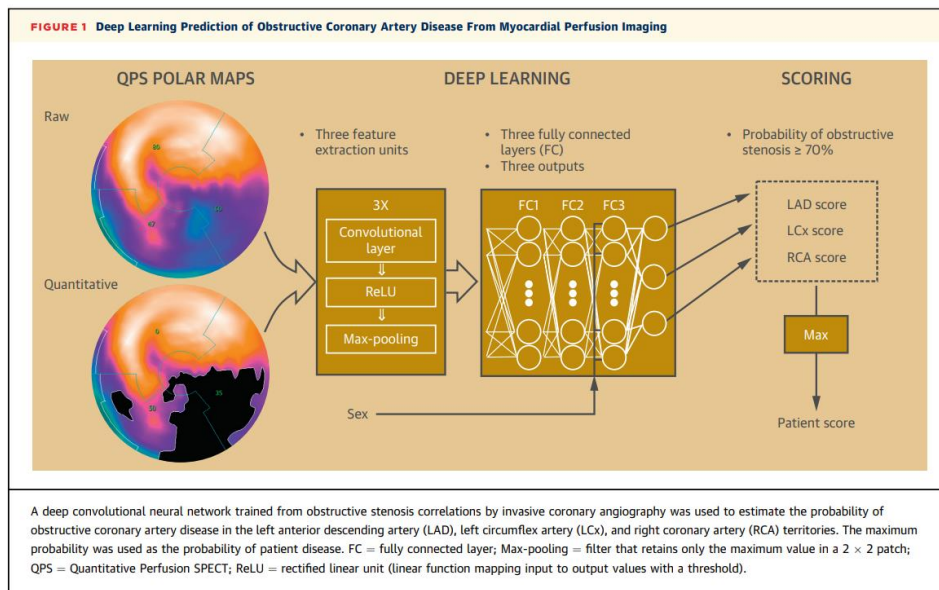
Σε μια νέα διαδικασία διασταυρούμενης επικύρωσης «one-center-out», ισοδύναμη με την εξωτερική επικύρωση, παρατηρήσαμε ότι η βαθιά μάθηση ξεπέρασε το cTPD στην πρόβλεψη της αποφρακτικής νόσου, καθώς παρέχει βελτιωμένη επεξεργασία των ίδιων πληροφοριών που χρησιμοποιούνται χωρίς πρόσθετες δοκιμές, ακτινοβολία ή κόστος. [113]

Τα αποτελέσματα που εξήγαγε είναι AUC: 0.81 και απόδοση 84.8%.

3.1.4 Μέθοδοι βαθιάς μάθησης για την πρόβλεψη αποφρακτικής νόσου με δεδομένα spect) (Deep Learning for Prediction of Obstructive Disease From Fast Myocardial Perfusion SPECT)

Η επόμενη μελέτη [97] αξιολόγησε την αυτόματη πρόβλεψη της αποφρακτικής νόσου, μέσω της απεικόνισης της αιμάτωσης του μυοκαρδίου (MPI) με μεθόδους βαθιάς μάθησης σε σύγκριση με τη μέθοδο του (TPD). Στην έρευνα συμπεριλήφθηκαν ασθενείς χωρίς προηγουμένως γνωστή στεφανιαία νόσο. Όσον αφορά για τα δεδομένα δυο μέθοδοι εξετάστηκαν με και χωρίς ποσοτικούς χάρτες (raw and quantitative), για να είναι ξεκάθαρη η απόδοση που θα εξάγουν. Το μοντέλο της παρούσας έρευνας αξιολογήθηκε σε διαδικασία 10 fold cross validation. Οι πολικοί χάρτες είναι δύο, για κάθε ασθενή και αποτελούνται από έναν πολικό χάρτη και ένα χάρτη TPD και είναι σε μέγεθος 64x64 πίξελ.

Το δίκτυο οργανώθηκε σε 2 στάδια. Το πρώτο στάδιο εξάγει χαρακτηριστικά από εικόνες. Το δίκτυο συνδέεται απευθείας με τα pixel του πολικού χάρτη μέσω ενός συνελκτικού στρώματος, ακολουθούμενο από φίλτρα μέγιστης συγκέντρωσης — φίλτρα που διατηρούν μόνο τη μέγιστη τιμή από μια είσοδο 2x2, για να αντισταθμίσουν τις μικρές αλλαγές και παραμορφώσεις εικόνας. Αυτές οι λειτουργίες επαναλαμβάνονται 3 φορές διαδοχικά για εξαγωγή χαρακτηριστικών εικόνας. Στο πρώτο πέρασμα ανιχνεύονται χαρακτηριστικά χαμηλού επιπέδου, ενώ τα 2 επόμενα περάσματα επιτρέπουν τον εντοπισμό λειτουργιών υψηλότερου επιπέδου (π.χ., περιγραφή περιφερειακών ελαττωμάτων διάχυσης). Στο δεύτερο στάδιο περιλαμβάνονται 3 πλήρως συνδεδεμένα στρώματα ακολουθούμενες από 3 παράλληλες εξόδους - 1 ανά στεφανιαία περιοχή. Τα πλήρως συνδεδεμένα στρώματα μετασχηματίζουν τα χαρακτηριστικά εικόνας και προσαρμόζουν τα βάρη για ενεργοποιήσεις νευρώνων κατά τη διάρκεια της εκπαίδευσης.



Εικόνα 31

Φαίνεται η διαδικασία που ακολουθήθηκε στην εικόνα 31, όπου το αποτέλεσμα, που ήταν θεμιτό να εξαχθεί ήταν η εκτίμηση της πιθανότητας αποφρακτικής στεφανιαίας νόσου στην αριστερή πρόσθια φθίνουσα αρτηρία (LAD), στην αριστερή περιφερειακή αρτηρία (LCx) και στη δεξιά στεφανιαία αρτηρία (RCA).

Η συμπερίληψη των ποσοτικών χαρτών διάχυσης βελτίωσε σημαντικά την απόδοση του μοντέλου βαθιάς μάθησης, σε σύγκριση με τη χρήση ακατέργαστων πολικών χαρτών μόνο, αλλά επίσης ήταν σημαντικά υψηλότερη από το TPD. Οπότε αποδείχθηκε, ότι η βαθιά μάθηση χρησιμοποιώντας έναν συνδυασμό ακατέργαστων και ποσοτικών πολικών χαρτών διάχυσης ξεπερνά το πρότυπο TPD στην πρόβλεψη της απόφραξης στεφανιαίας αορτής, που προκαλείται λόγω της αποφρακτικής νόσου. Παρατηρήθηκε επίσης ότι η προσθήκη ποσοτικών πολικών χαρτών σε ακατέργαστους χάρτες βελτιώνει την πρόβλεψη CAD.

Η απόδοση και η γενική εκτίμηση σφάλματος της πρόβλεψης βαθιάς μάθησης αξιολογήθηκε χρησιμοποιώντας τη μέθοδο cross validation με 10 fold για τη μείωση του σφάλματος πρόβλεψης, και αποτρέποντας την υπερβολική τοποθέτηση του μοντέλου. Έχει βρεθεί ότι η ικανότητα της βαθιάς μάθησης να ποσοτικοποιεί σχετικές μετρήσεις από τα ανεπεξέργαστα δεδομένα θα έχει

σημαντικό ρόλο στην καρδιαγγειακή ιατρική, και ως ένα μη εποπτευόμενο εργαλείο για την απεικόνιση της ηχοκαρδιογραφίας.

Τα αποτελέσματα είναι AUC: 0.8 και απόδοση 82.3%.

3.1.5 Μοντέλα κατηγοριοποίησης για δεδομένα Spect

Στην έρευνα [98] προτείνονται δύο διαφορετικά μοντέλα ταξινόμησης, ένα με βάση τη βαθιά μάθηση και ένα που βασίζεται στη μηχανική μάθηση. Επίσης υπάρχει και η αξιολόγηση των εμπειρογνώμων. Όλες οι επιλεγμένες εικόνες SPECT MPI ερμηνεύθηκαν και ταξινομήθηκαν ανεξάρτητα από τους ειδικούς αναγνώστες χρησιμοποιώντας δύο διαφορετικές ετικέτες εξόδου (1: φυσιολογική και 2: μη φυσιολογική) για περαιτέρω επεξεργασία.

Ο πρώτος τύπος μοντέλου χρησιμοποιεί τη μεταφορά μάθησης με προ-εκπαιδευμένα βαθιά νευρωνικά δίκτυα και το άλλο στο μοντέλο SVM με βαθιά και shallow χαρακτηριστικά, δηλαδή χωρίς πολλά επίπεδα εκμάθησης.

Στο μοντέλο που βασίζεται στη βαθιά μάθηση, η εκμάθηση μεταφοράς χρησιμοποιείται με τη χρήση διάφορων προ-εκπαιδευμένων βαριών νευρωνικών δικτύων, αντί να εκπαιδεύεται ένα νέο δίκτυο από το μηδέν, έτσι ώστε να επιτυγχάνεται υψηλότερη απόδοση ταξινόμησης, ακόμη και με μικρό αριθμό εικόνων για εκπαίδευση. Σε συνδυασμό, χρησιμοποιείται ένας ισχυρός ταξινομητής, η μηχανή φορέα υποστήριξης (SVM), με βαθιά και ρηχά χαρακτηριστικά που εξάγονται από διάφορα προ-εκπαιδευμένα βαθιά νευρωνικά δίκτυα για την εργασία ταξινόμησης.

Σε αυτήν την προσέγγιση, η εκμάθηση μεταφοράς χρησιμοποιείται με τη χρήση αρκετών προ-εκπαιδευμένων βαθιών νευρωνικών δικτύων αντί να εκπαιδεύεται ένα νέο δίκτυο με τυχαία αρχικοποιημένα βάρη από το μηδέν. Με αυτόν τον τρόπο, οι εκφρασμένες απεικονίσεις χαρακτηριστικών για ένα ευρύ φάσμα εικόνων μπορούν να μεταφερθούν γρήγορα για να ταξινομήσουν τις εικόνες SPECT χρησιμοποιώντας λιγότερες εικόνες στη διαδικασία της εκπαίδευσης [114]. Τέλος, το δίκτυο εκπαιδεύεται χρησιμοποιώντας τις εικόνες εκπαίδευσης και ελέγχου, έτσι ώστε οι εικόνες SPECT να μπορούν να ταξινομηθούν από το εκ νέου εκπαιδευμένο δίκτυο. Η πραγματική απόδοση του τελικού δικτύου αξιολογείται με εικόνες που δεν

χρησιμοποιούνται στη φάση της εκπαίδευσης. Επίσης, σε αυτό το βήμα, γίνεται πρόωρη διακοπή για να αποφευχθεί η υπερβολική τοποθέτηση.

Στο μοντέλο που βασίζεται σε βαθιά μάθηση, το χρησιμοποιούμενο δίκτυο δημιουργεί μια ιεραρχική αναπαράσταση εικόνων εισόδου. Τα βαθύτερα στρώματα περιέχουν χαρακτηριστικά υψηλότερου επιπέδου, τα οποία κατασκευάζονται χρησιμοποιώντας τα χαρακτηριστικά χαμηλότερου επιπέδου των προηγούμενων επιπέδων. Τα προηγούμενα στρώματα συνήθως εξάγουν λιγότερα, πιο ρηχά χαρακτηριστικά, έχουν υψηλότερη χωρική ανάλυση και έχουν μεγαλύτερο συνολικό αριθμό ενεργοποιήσεων.

Ως εκ τούτου, στη δεύτερη προσέγγιση, το SVM, που είναι ένας παραδοσιακός και ισχυρός ταξινομητής, χρησιμοποιείται με τα βαθιά και ρηχά χαρακτηριστικά που εξάγονται από διάφορα προ-εκπαιδευμένα βαθιά νευρωνικά δίκτυα.

Από την άλλη πλευρά, το μοντέλο βασισμένο στους ειδικούς, στοχεύει στη μετατροπή της γνώσης των ειδικών αναγνωστών σε κατάλληλες μεθόδους επεξεργασίας εικόνας χαμηλού και μεσαίου επιπέδου για την ταξινόμηση των εικόνων SPECT. Οι εικόνες που ελήφθησαν μετά το στρες SPECT MPI συγκρίθηκαν οπτικά με τις εικόνες που ελήφθησαν μετά το υπόλοιπο SPECT MPI για την εκτίμηση των ελαττωμάτων που προκλήθηκαν από τον ραδιοφαρμακευτικό παράγοντα ή την άσκηση. Εάν δεν παρατηρηθεί ελάττωμα διάχυσης στις εικόνες SPECT μετά την άσκηση στις υπόλοιπες εικόνες, αυτή η κατάσταση ορίζεται ως ισχαιμία. Από την άλλη πλευρά, ορίζεται ως έμφραγμα εάν παρατηρηθεί το ελάττωμα τόσο στις εικόνες stress όσο και στην rest κατάσταση.

Η μέγιστη ακρίβεια επιτυγχάνεται μέσω του ταξινομητή SVM με shallow χαρακτηριστικά. Σε αυτό το μοντέλο, δεδομένου ότι το μέγεθος του συνόλου δεδομένων είναι σχετικά μικρό, υπάρχουν μόνο λίγα δεδομένα για να γίνει εκμάθηση από τη ρύθμιση των βαθύτερων επιπέδων του δικτύου. Δεδομένου ότι οι εικόνες στο σύνολο δεδομένων είναι επίσης πολύ διαφορετικές από τις αρχικές εικόνες που χρησιμοποιούνται για την εκπαίδευση των προ-εκπαιδευμένων δικτύων, η χρήση γενικότερων χαρακτηριστικών παρέχει καλύτερη απόδοση ταξινόμησης. Επομένως, επιτυγχάνεται υψηλότερη ακρίβεια όταν χρησιμοποιούνται shallow χαρακτηριστικά και όχι βαθιά.

Για το μοντέλο που βασίζεται σε τεχνικές βαθιάς μάθησης, ο χρόνος επεξεργασίας είναι άμεσα ανάλογος με τον μέγιστο αριθμό εποχών, τον αριθμό των επιπέδων και τις συνδέσεις μεταξύ των επιπέδων. Επομένως, ο ελάχιστος χρόνος επεξεργασίας (0,09 δευτερόλεπτα) επιτεύχθηκε με το AlexNet, το οποίο έχει τα λιγότερα επίπεδα και συνδέσεις και τον μέγιστο αριθμό εποχών σε σύγκριση με τα άλλα δίκτυα. Από την άλλη πλευρά, ο χρόνος επεξεργασίας για το μοντέλο βασισμένο στη γνώση υπολογίστηκε ως 4,03 δευτερόλεπτα.

Λαμβάνοντας υπόψη τους περιορισμούς των προηγούμενων έργων, τα προτεινόμενα μοντέλα είναι ανεξάρτητα από το φύλο και δεν απαιτούν πολικούς χάρτες ή μια κανονικοποιημένη βάση δεδομένων ελέγχου, η οποία είναι συγκεκριμένη για διαφορετικούς πληθυσμούς. Ένα άλλο πλεονέκτημα των προτεινόμενων μοντέλων είναι ότι λαμβάνουν μόνο τις εικόνες SPECT ως είσοδο, δεν απαιτούν πρόσθετα δεδομένα και λειτουργούν απευθείας στις ίδιες τις εικόνες. Επίσης, οι γιατροί μπορούν να χρησιμοποιήσουν τα αποτελέσματα αυτών των συστημάτων ως δεύτερη γνώμη για τη βελτίωση της διαγνωστικής τους ακρίβειας.

Τα αποτελέσματα για την μέθοδο της βαθιάς μάθησης είναι accuracy 94%, sensitivity 88%, και specificity 100% και για το μοντέλο βασισμένο στη γνώση είναι απόδοση 93%, sensitivity 100%, specificity 86%.

3.1.6 Αυτόματη κατηγοριοποίηση πολικών χαρτών με μεθόδους βαθιάς μάθησης και αύξηση δεδομένων

Στην επόμενη έρευνα [115] έγινε διερεύνηση τεχνικής βαθιάς μάθησης, για τον αυτόματο χαρακτηρισμό πολικών χαρτών που προέρχονται από μελέτες απεικόνισης αιμάτωσης μυοκαρδίου (MPI) για τη διάγνωση της στεφανιαίας νόσου.

Στο σετ δεδομένων της παρούσας έρευνας περιλήφθηκαν πολικοί χάρτες σε μορφή AC attenuation-corrected (AC) and non-corrected (NAC). Καθώς το αρχικό σύνολο δεδομένων αυτής της μελέτης ήταν μικρό για να γίνει σωστή εκπαίδευση ενός μοντέλου βαθιάς μάθησης ακολουθήθηκαν δύο στρατηγικές. Το πρώτο ονομάζεται μεταφορά μάθησης. Για αυτό, χρησιμοποιήθηκε το συνελκτικό νευρωνικό δίκτυο VGG16. Η δεύτερη στρατηγική περιλαμβάνει την αύξηση δεδομένων, η οποία επιτυγχάνεται με την περιστροφή των πολικών χαρτών, για την επέκταση του σετ εκπαίδευσης. Αξιολογήθηκε το VGG16 με 10 fold cross validation. Τα αποτελέσματα συγκρίθηκαν με τη οπτική διαγνωστική απόδοση των ιατρών.

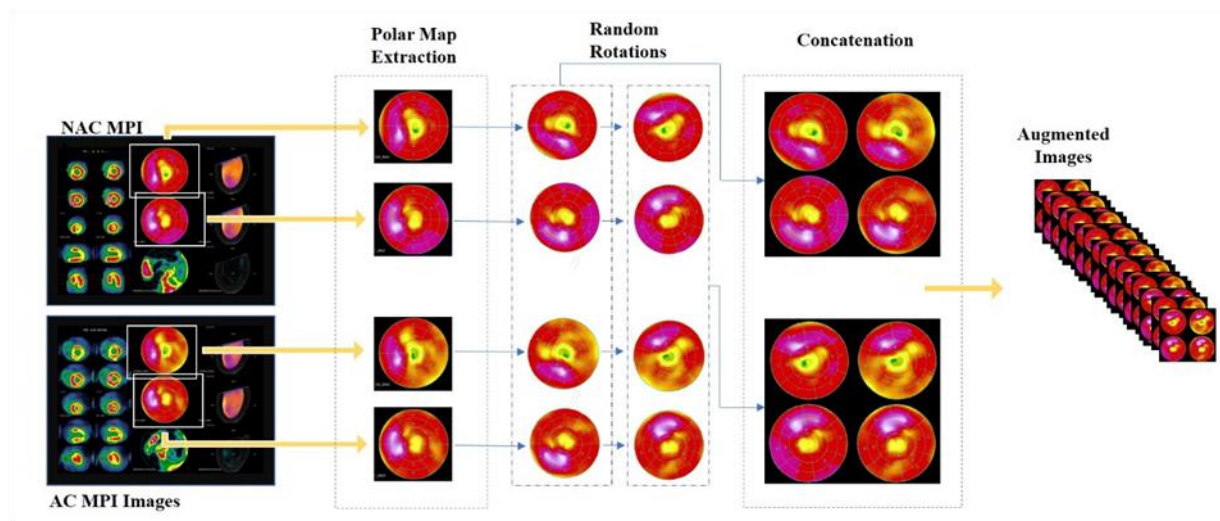
Το συνελκτικό νευρωνικό δίκτυο (VGG) [116] χρησιμοποιείται γενικά σε έρευνες. Για την εργασία ιατρικής ταξινόμησης χρησιμοποιείται η εκμάθηση μεταφοράς για να γίνει εκμετάλλευση των δυνατοτήτων των προκαθορισμένων δικτύων. Από τα παράγωγα της κύριας αρχιτεκτονικής του VGG, επιλέχθηκαν δύο αρχιτεκτονικές που ονομάζονται VGG16 και VGG19. Και οι δύο έχουν το όνομά τους από τον αριθμό των συνολικών εκπαιδευσιμων στρωμάτων, που είναι αντίστοιχα 16 και 19. Ο αριθμός των επιπέδων επηρεάζει το βάθος του δικτύου και τα βήματα εξαγωγής χαρακτηριστικών. Δεδομένου ότι υπάρχει μικρή διαφορά μεταξύ αυτών των δύο αρχιτεκτονικών, ήταν προτιμότερο να χρησιμοποιηθούν και τα δύο.

Η χρήση ενός αυτοματοποιημένου συστήματος ανάλυσης εικόνας και χαρακτηρισμού MPI, θα μπορούσε ενδεχομένως να απελευθερώσει τους γιατρούς από μακροχρόνιες διαδικασίες ή και ακόμη να λειτουργήσει ως βοηθητικό εργαλείο για εκπαιδευτικούς σκοπούς.

Οι εικόνες AC και NAC του στρες και της rest εξήχθησαν με πίξελ εισόδου 150x150, για να μειωθεί το υπολογιστικό κόστος. Οι τέσσερις πολικοί χάρτες που αντιστοιχούν σε κάθε περίπτωση συνδυάστηκαν σε μία εικόνα. Για την εκπαίδευση βαθιών νευρωνικών δικτύων, απαιτούνται μεγάλης κλίμακας σύνολα δεδομένων.

Δημιουργήθηκαν τρία διαφορετικά σύνολα δεδομένων, που ονομάστηκαν AC, NAC και ACNAC. Τα σύνολα δεδομένων AC και NAC περιείχαν δύο πολικούς χάρτες για κάθε περίπτωση ασθενούς (κατάσταση stress και rest), ενώ το σύνολο δεδομένων ACNAC περιείχε 4 πολικούς χάρτες συνδυασμένους σε μία εικόνα για κάθε περίπτωση ασθενούς. Για αύξηση δεδομένων, έγινε ελαφρά περιστροφή των πολικών χαρτών (κατά μέγιστο 45 μοίρες). Οι περιστρεφόμενοι πολικοί χάρτες στη συνέχεια ενώθηκαν, δημιουργώντας νέες εικόνες. Κατά την αύξηση δεδομένων, δημιουργήθηκαν 1000 νέες εικόνες και των δύο τάξεων για εκπαίδευση. Για την εξισορρόπηση του συνόλου δεδομένων, δημιουργήθηκαν περισσότερες εικόνες που ανήκουν στην κλάση υγιείς. Έτσι, η εκπαίδευση πραγματοποιήθηκε με ισορροπημένα δεδομένα, παρά το αρχικό μικρό πρόβλημα ανισορροπίας.

Το σχήμα παρακάτω (Εικόνα 32) παρουσιάζει τα βήματα προ-επεξεργασίας και αύξησης δεδομένων που σχετίζονται με το σύνολο δεδομένων ACNAC.



Εικόνα 32

Λόγω των περιορισμών κλίμακας συνόλων δεδομένων, πραγματοποιήθηκε 10 fold cross validation προκειμένου να επιτευχθεί σωστή εκπαίδευση και αξιόπιστη εξαγωγή απόδοσης και στο απλό CNN, εκτός από το VGG. Για να γίνει αξιολόγηση της στρατηγικής μάθησης μεταφοράς για τη συγκεκριμένη αρχιτεκτονική VGG, πραγματοποιήθηκε ένα άλλο πείραμα. Αρχικά, εκπαιδεύτηκε το VGG16 από το μηδέν, δηλαδή χωρίς να διατηρηθούν τα βάρη που έχει μάθει, το οποίο θεωρείται η συμβατική διαδικασία προπόνησης. Αυτό το μοντέλο αναφέρεται ως VGG16-

16. Δεύτερον, έγινε πείραμα με την αλλαγή του αριθμού και της διάταξης των συνελκτικών στρωμάτων που διατηρούμε εκπαιδύσιμα κάθε φορά. Αυτή η διαδικασία μπορεί να θεωρηθεί ως μέρος μιας διαδικασίας που ονομάζεται fine-tuning και περιλαμβάνει πειραματισμό με τις παραμέτρους του μοντέλου για την επίτευξη του καλύτερου δυνατού αποτελέσματος. Ξεκινώντας από την κορυφή του δικτύου, έγινε προσαρμογή σταδιακά στα μπλοκ συνέλιξης. Επισημαίνεται ότι η μεταφορά μάθησης είναι μια αποτελεσματική στρατηγική για την ταξινόμηση των συγκεκριμένων εικόνων, ικανών να αντιμετωπίζουν μικρά σύνολα δεδομένων.

Όσον αφορά για τα αποτελέσματα το VGG16 έφτασε απόδοση 74.53%, sensitivity 75.69% και specificity 71.18% όταν οι φωτογραφίες AC και NAC συγχωνεύθηκαν σε μια. Οι εμπειρογνώμονες εξήγαγαν απόδοση 75.00%, sensitivity 76.97% specificity 70.31%.

3.1.7 Τεχνητή νοημοσύνη, μηχανική (βαθιά) μάθηση και ραδιονομική-γεωμική: ορισμού και εφαρμογές πυρηνικής ιατρικής

Στην εργασία [73] παρουσιάζεται μια εργασία, όπου έχει δημιουργηθεί ένα δίκτυο, του οποίου ο στόχος είναι να υπολογίσει από τα δεδομένα ορισμένα χαρακτηριστικά, τα οποία στη συνέχεια τροφοδοτούνται σε έναν άλλο αλγόριθμο, ο οποίος τα χαρτογραφεί, με στόχο τη διαφοροποίηση μεταξύ δύο τύπων όγκων ασθενών με βάση το κλινικό αποτέλεσμα. Σύμφωνα με αυτό, είναι εύκολο να εννοηθεί, ότι μία από τις σημαντικότερες διαφορές μεταξύ των σύγχρονων τεχνικών και των «παλαιότερων» προσεγγίσεων μηχανικής μάθησης είναι ότι ο στόχος αυτών των δικτύων είναι να μάθουν συγκεκριμένα μοτίβα που σχετίζονται με μια δεδομένη εργασία (π.χ. τμηματοποίηση ή πρόβλεψη τελικού σημείου) από τα ίδια τα δεδομένα (π.χ. εικόνες) , αντί να βασίζονται σε γνώσεις των εμπειρογνομώνων. Η παρέμβαση του χρήστη, όπως για ανίχνευση και επιλογή αντικειμένων θα μπορούσε επομένως να απλοποιηθεί σημαντικά ή ακόμη και να θεωρηθεί περιττή. Μεγάλο μέρος απεικόνισης της τεχνητής νοημοσύνης περιλαμβάνει τη χρήση νευρωνικών δικτύων. Τα νευρωνικά δίκτυα αποδίδουν καλύτερα από τις κλασικές μεθόδους της μηχανικής μάθησης, αντιμετωπίζοντας καλύτερα μεγάλες ποσότητες δεδομένων. [117] [118] [73] [119] [5] [75]

Συμπερασματικά, είναι δυνατή η ταξινόμηση ενός μικρού συνόλου δεδομένων εικόνων MPI σε επίπεδο που ανταγωνίζεται τη διαγνωστική ικανότητα των ειδικών με μεθόδους βαθιάς μάθησης, χρησιμοποιώντας τεχνικές μάθησης μεταφοράς και αύξησης δεδομένων. Το προτεινόμενο μοντέλο μπορεί να αναπτυχθεί εύκολα και άμεσα κλινικά για την αξιολόγηση των εικόνων NAC και AC.

Κεφ. 4 Παρουσίαση προβλήματος με δεδομένα

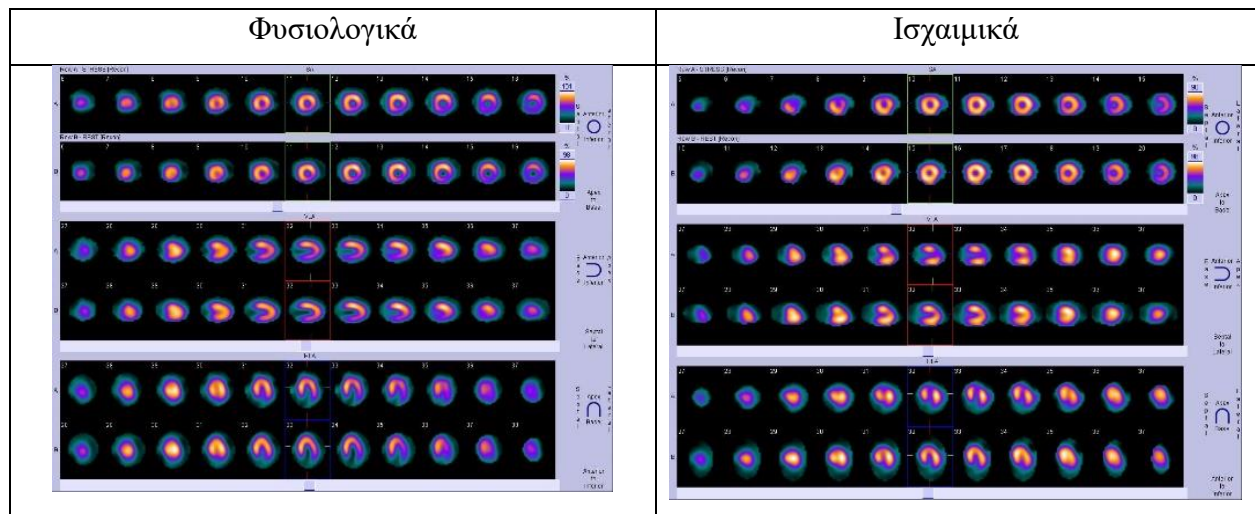
4.1 Ιατρικές φωτογραφίες

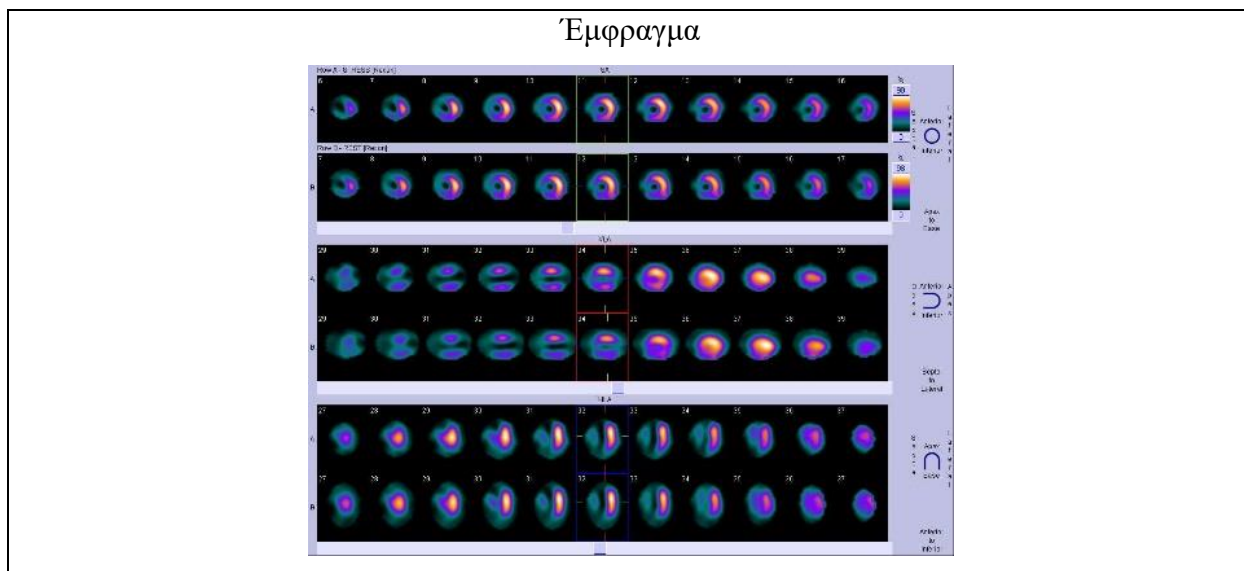
Σύμφωνα με την καθοδήγηση των υπαρχόντων δημοσιεύσεων ακολούθησε και η δικιά μας έρευνα. Προκειμένου να γίνει διερεύνηση της εφαρμογής των συνελκτικών νευρωνικών δικτύων στην απεικόνιση του ιατρικού προβλήματος απεικόνισης αιμάτωσης του μυοκαρδίου, χρησιμοποιήθηκαν εικόνες που προ-επεξεργάστηκαν και διαγνωστήκαν από τον Πυρηνικό Ιατρό, σε τρεις κατηγορίες. Πιο συγκεκριμένα στον πίνακα 2 φαίνονται οι κατηγορίες με τον αριθμό των δεδομένων αντίστοιχα:

Κατηγορία	Περιστατικά
Φυσιολογικά	262
Ισχαιμικά	251
Έμφραγμα	134

Πίνακας 2

Ακολουθεί παράδειγμα το πως είναι η κάθε φωτογραφία για κάθε μία περίπτωση (δες Πίνακα 3)





Πίνακας 3

Χάρη στη δυνατότητα του μεγάλου όγκου δεδομένων, επιτεύχθηκε η δημιουργία τεσσάρων διαφορετικών σετ δεδομένων. Παραθέτονται παρακάτω:

Όσον αφορά για τις 2 κλάσεις

Σετ α) Φυσιολογικά – Ισχαιμικά

Σετ β) Φυσιολογικά – Ισχαιμικά, Έμφραγμα

Σετ γ) Ισχαιμικά– Έμφραγμα

Για τις 3 κλάσεις

Σετ δ) Φυσιολογικά – Ισχαιμικά – Έμφραγμα

Τα αποτελέσματα του αλγορίθμου που προέκυψαν αναφέρονται σε επόμενο κεφάλαιο.

Είναι σημαντικό να σημειωθεί ότι είχαμε δεδομένα σε μορφή κατηγορίας SPECT (Single-photon emission computed tomography) , και διαφορετικά δεδομένα σε μορφή πολικού χάρτη, τα οποία δόθηκαν από το Τμήμα Πυρηνικής Ιατρικής του Πανεπιστημιακού Νοσοκομείου Πατρών και

χρησιμοποιήθηκαν για τη διάγνωση του μυοκαρδίου και συγκρίθηκαν με ήδη δημοσιεύσιμα αποτελέσματα από το ίδιο σετ δεδομένων.

Σε αυτήν την περίπτωση, χρησιμοποιώντας τους πολικούς χάρτες, έχουμε ένα σετ δεδομένων για ανάλυση που αποτελείται από δύο κατηγορίες περιστατικών, τα φυσιολογικά και τα παθολογικά. Στον πίνακα 4 φαίνεται ο αριθμός των δεδομένων για κάθε κλάση αντίστοιχα.

Κατηγορία	Περιστατικά
Φυσιολογικά	146
Παθολογικά	176

Πίνακας 4

Η μορφή τους είναι στις ακόλουθες μορφές

- AC
- NAC

όπως τις παρέδωσε ο ιατρός, όπως είχαμε δει σε μία από τις έρευνες που περιεγράφηκαν.

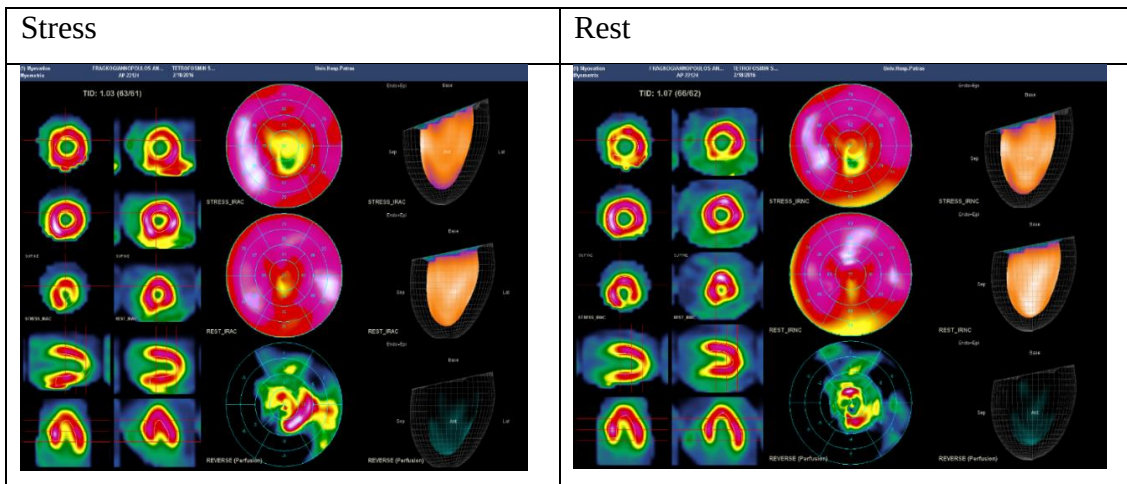
Παρέχονται σε δυο φάσεις ανάλυσης,

- Stress
- Rest

Η μορφή άγχους stress γίνεται τεχνητά με ενδοφλέβια ουσία ή με φυσικό τρόπο, δηλαδή με ελεγχόμενη γυμναστική. Όσον αφορά για την μορφή rest είναι απλά ξαπλωμένος ο ασθενής και γίνεται η δειγματοληψία.

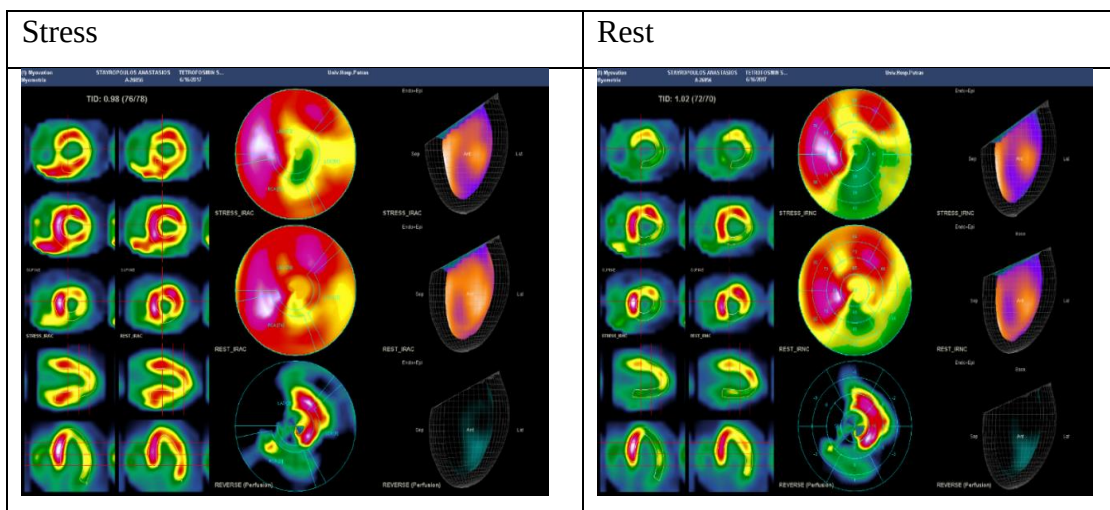
Τα φυσιολογικά περιστατικά έχουν την εξής απεικόνιση (δείτε Πίνακες 5):

Σετ ε) Φυσιολογικά PM(Polar Maps) – Παθολογικά PM (Polar Maps)



Πίνακας 5

Και τα παθολογικά είναι ως εξής στον πίνακα 6:



Πίνακας 6

4.2.Προεπεξεργασία δεδομένων (SPECT και πολικοί χάρτες)

Όσον αφορά για τη διαδικασία της προεπεξεργασίας δεδομένων, συνήθως δεν χρειάζεται να γίνει από την στιγμή που ακολουθεί η κατηγοριοποίηση μέσω συνελκτικών νευρωνικών δικτύων, καθώς είναι ικανά από μόνα τους να εξάγουν μοτίβα. Οπότε δοκιμάστηκε να γίνει κατηγοριοποίηση με αυτή την μορφή των δεδομένων και επειδή περιέχουν αρκετό «θόρυβο», δεν ήταν τα αποτελέσματα αξιόλογα. Ο λόγος που περιέχουν «θόρυβο» είναι λόγω του λογισμικού εξαγωγής δεδομένων, το οποίο παρεμβάλλει στη διαδικασία κατηγοριοποίησης του αλγορίθμου και δεν είναι σε θέση να αναγνωρίσει μοτίβα και να ξεχωρίσει τις φωτογραφίες και να τις ταξινομήσει σωστά.

Η επιθυμητή προεπεξεργασία που χρειαζόταν να γίνει ήταν να σβηστεί οτιδήποτε δεν αντιστοιχεί σε απεικόνιση καρδιάς. Ύστερα, από έρευνα βρέθηκε η τεχνική, η οποία ονομάζεται inpainting και ανήκει στη βιβλιοθήκη cv της γλώσσας προγραμματισμού της Python, το οποίο βασίζεται σε μια πολύ εύκολη μέθοδο:

Γίνεται αντικατάσταση των σημείων που δεν επιθυμούμε με τις τιμές των πίξελ των γειτονικών σημείων. Παραθέτονται δυο διαφορετικές εκδοχές του αλγορίθμου. Και οι δύο είναι προσβάσιμοι από την ίδια συνάρτηση cv.inpaint().

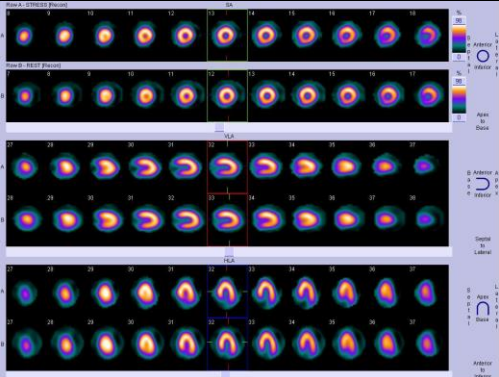
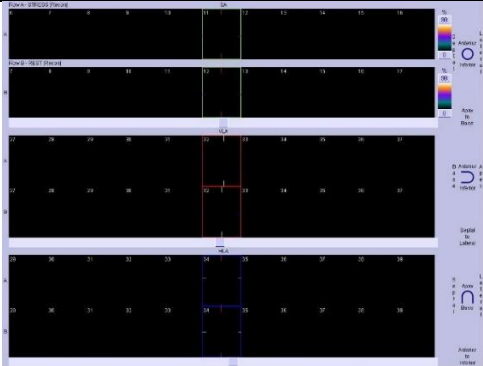
- cv.INPAINT_TELEA
- cv.INPAINT_NS

Αυτή η μέθοδος περιγράφεται στην εργασία [120].

Υπάρχουν δυο υλοποιήσεις αυτού του αλγορίθμου, οι οποίες στο δικό μας σετ δεδομένων είχαν την ίδια συμπεριφορά. Περιγραφικά, ο αλγόριθμος παίρνει σαν είσοδο δύο ιατρικές εικόνες, αυτή που θέλουμε να επεξεργαστούμε, και την φωτογραφία-μάσκα, όπως ονομάζεται, η οποία θα υποδείξει στον αλγόριθμο, ποια σημεία πρέπει να επικαλύψει, έτσι ώστε να αφαιρεθούν από την δικιά μας φωτογραφία.

4.3 Προεπεξεργασία SPECT MPI εικόνες

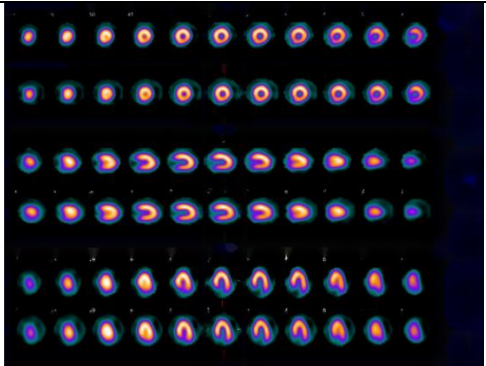
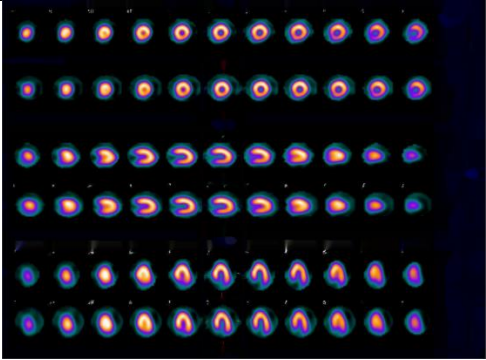
Παρακάτω, στον πίνακα 7 παρατίθεται ένα παράδειγμα, πως έγινε απαλοιφή θορύβου και για τις δύο μορφές απεικόνισης.

Spect Data	Mask
	

Πίνακας 7

Είναι εύκολο να κατανοηθεί ότι η μάσκα είναι σαν την αρχική φωτογραφία, αλλά μόνο με τα σημεία, που επιθυμούμε να σβηστούν.

Εφαρμόζοντας τον αλγόριθμο, που αναφέρθηκε παραπάνω έχουμε τις παραπάνω δύο εκδοχές, δηλαδή `cv.INPAINT_TELEA` και `cv.INPAINT_NS`, όπου και οι δυο απεικονίζονται στον πίνακα 8:

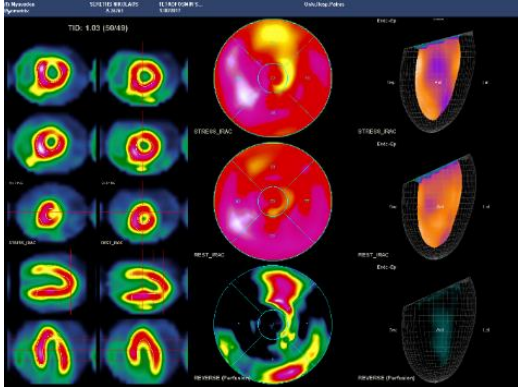
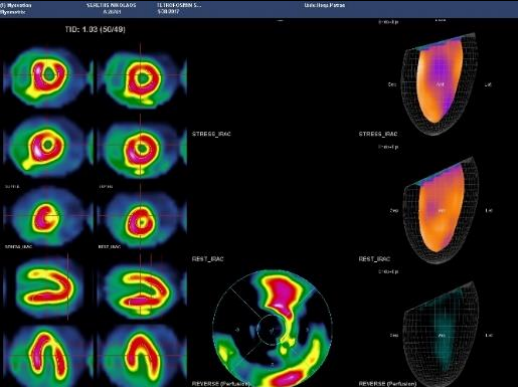
<code>cv.INPAINT_TELEA</code>	
<code>cv.INPAINT_NS</code>	

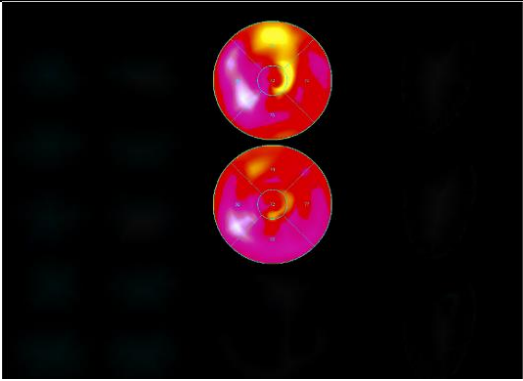
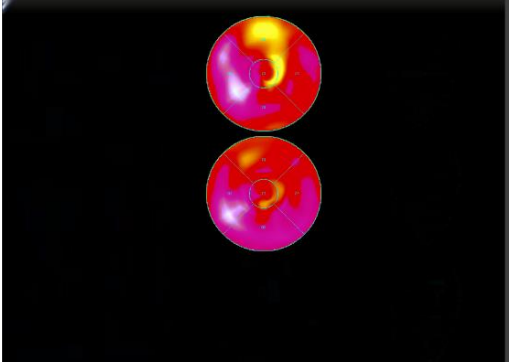
Πίνακας 8

Παρατηρούμε ότι τα αποτελέσματα είναι αρκετά ικανοποιητικά, καθώς και παρόμοια μεταξύ τους.

4.2.2 Προεπεξεργασία πολικών χαρτών

Στη δεύτερη απεικόνιση μυοκαρδίου ακολουθήθηκε η ίδια διαδικασία για την αφαίρεση θορύβου, με τα ανάλογα δεδομένα. Έχουμε τον πίνακα 9:

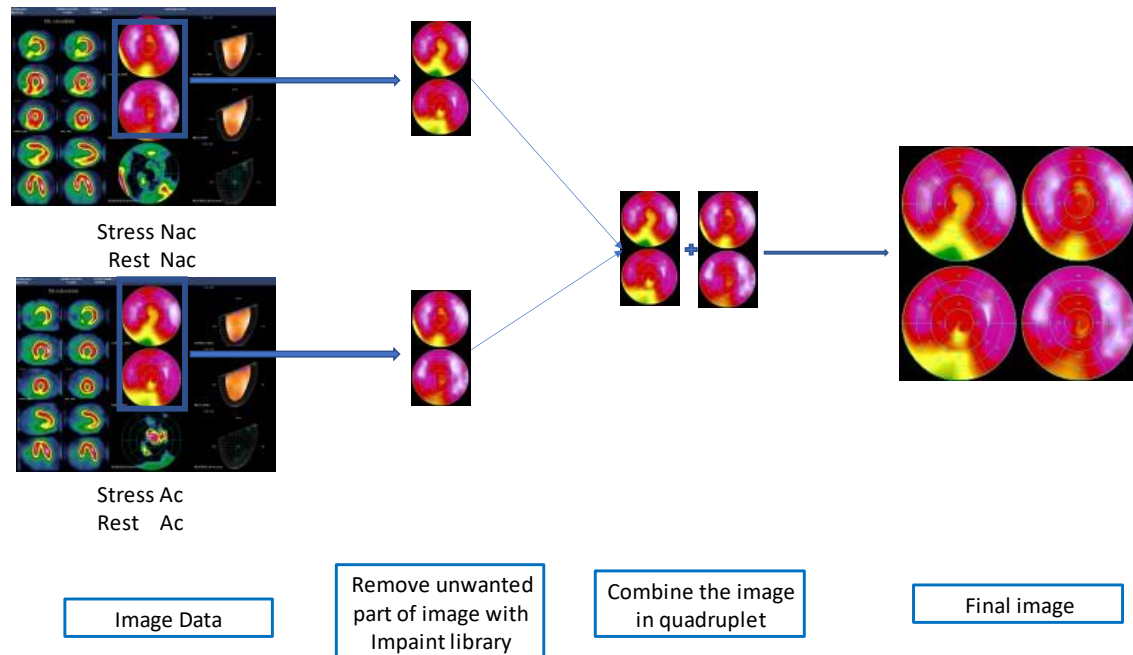
Αρχική φωτογραφία	
Εικόνα μάσκα	
Αποτελέσματα	

cv.INPAINT_TELEA	
cv.INPAINT_NS	

Πίνακας 9

Το μόνο διαφορετικό που έγινε στο σεν αυτό είναι ότι μετά την αφαίρεση θορύβου, οι δυο φωτογραφίες από κάθε κατηγορία ενώθηκαν σε μια τετράδα. Το αποτέλεσμα, που προκύπτει θεωρήθηκε ως μια οντότητα του σεν δεδομένων, της ανάλογης κατηγορίας. Η μέθοδος που ακολουθήθηκε είναι ότι σβήστηκε το μαύρο πλαίσιο και με έναν αλγόριθμο σε Python, συνενώθηκαν οι δυάδες σε τετράδα φωτογραφιών.

Το παρακάτω σχήμα στην εικόνα 33 απεικονίζει την μεθοδολογία αυτή.



Εικόνα 33

Κάθε μια τετράδα απεικονίζει μια οντότητα των δεδομένων. Ύστερα από αυτή τη διαδικασία τα δεδομένα, που έχουμε είναι αρκετά για να ξεκινήσει η αλγοριθμική διαδικασία.

4.3 Ανάλυση λογισμικού και τον τρόπο υλοποίησης

Σε περίπτωση που γίνει θεμιτό αυτός ο αλγόριθμος να υλοποιηθεί σε πραγματικό σενάριο, από ιατρικό προσωπικό θα πρέπει να περάσει το λογισμικό του νευρωνικού δικτύου σε κάποια εφαρμογή επιφάνειας εργασίας ή σε κάποια σελίδα στο ίντερνετ. Η λειτουργία της θα είναι να ανεβάζουμε μια φωτογραφία ή ένα σετ δεδομένων και να γίνεται η κατηγοριοποίηση στο back end και να εμφανίζεται η έξοδος που πρόβλεψε ο αλγόριθμος σε ένα περιβάλλον φιλικό για τον χρήστη. Θα υπάρχει βέβαια και η κρίση του ιατρού, οπότε ο αλγόριθμος θα ενεργεί ως βοηθητικό εργαλείο, στα πλαίσια υποστήριξης λήψης αποφάσεων. Ύστερα από έρευνα βρέθηκε ένα λογισμικό, το οποίο υλοποιεί κατηγοριοποιήσεις σε image classification. [121]

Το λογισμικό που βρέθηκε υλοποιείται με τη βιβλιοθήκη Flask, όπου είναι ένα micro web framework γραμμένο σε Python, και σε αυτό οφείλεται το γεγονός της ύπαρξης της σελίδας και της δυνατότητας να ενεργεί με το νευρωνικό δίκτυο στο back end. Είναι αρκετά εύχρηστο, καθώς δεν απαιτεί συγκεκριμένα εργαλεία ή βιβλιοθήκες. [122] [123]

Ο παρακάτω κώδικας στην εικόνα 34 δείχνει μια απλή εφαρμογή ιστού που εμφανίζει το "Hello World!":

```
from flask import Flask
app = Flask(__name__)

@app.route("/")
def hello():
    return "Hello World"

if __name__ == "__main__":
    app.run(debug=False)
```

Εικόνα 34

Οπότε η σελίδα αυτή προσαρμόστηκε με τα βάρη και το μοντέλο που έχουμε ήδη και είναι σε θέση να κατηγοριοποιεί για image classification 3-class.

Το μόνο που πρέπει να γίνει για να τρέξει η σελίδα είναι να πάμε στο φάκελο που υπάρχουν τα αρχεία και να τρέξουμε την εντολή, που φαίνεται στην εικόνα 35, μέσω του command prompt:

```
>python app.py
```

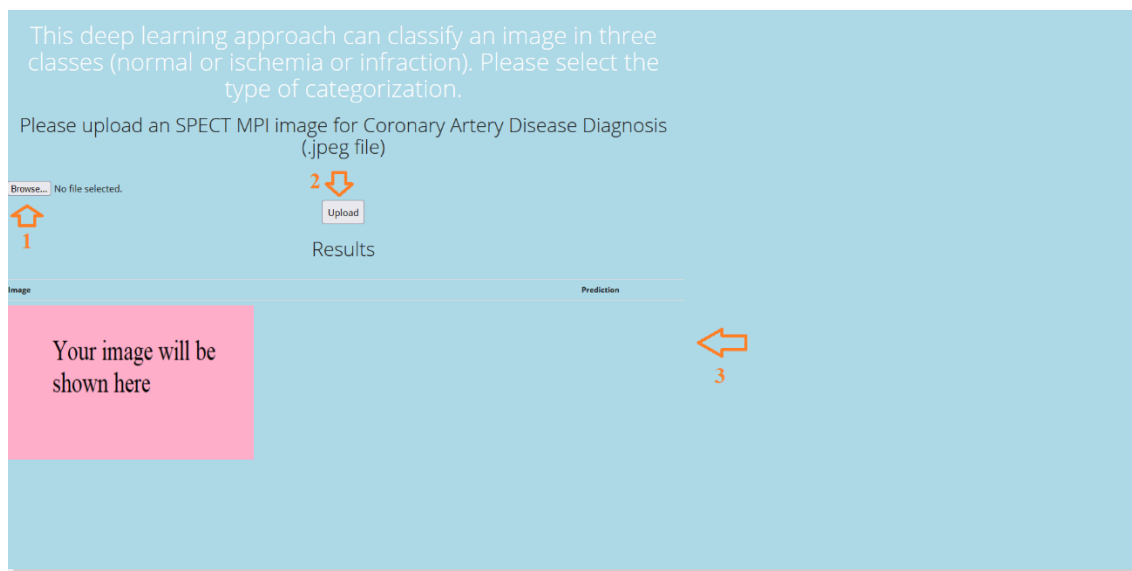
Εικόνα 35

Στη συνέχεια αν ανοίξουμε κάποιο περιηγητή και κατευθυνθούμε στη σελίδα **localhost:3000**

Θα δούμε το interface της σελίδας, το οποίο είναι το παρακάτω.

Ακολουθούν στιγμιότυπα με το πως λειτουργεί η σελίδα.

Στη συνέχεια η αρχική σελίδα που έχουμε φαίνεται στην εικόνα 36:



Εικόνα 36

Στην επιλογή 1 επιλέγουμε από τον υπολογιστή μας ποια φωτογραφία θέλουμε να προβλέψει ο αλγόριθμος. Αφού γίνει η επιλογή πατάμε την επιλογή 2 για να αποθηκευτεί στο Flask App. Στο 3 εμφανίζεται η πρόβλεψη του αλγορίθμου.

Αρχικά το παράδειγμα στην εικόνα 37 απεικονίζει μια φυσιολογική φωτογραφία, καθώς και τη διαδικασία πρόβλεψης.

This deep learning approach can classify an image in three classes (normal or ischemia or infraction). Please select the type of categorization.

Please upload an SPECT MPI image for Coronary Artery Disease Diagnosis (.jpeg file)

normal (522).jpg

Results

Image	Prediction
Your image will be shown here	

Εικόνα 37

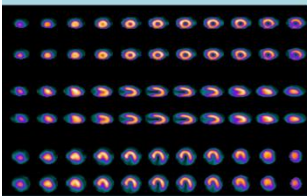
Παρατηρούμε ότι εμφανίστηκε το όνομα της φωτογραφίας στη σελίδα. Αφού πατήσουμε την επιλογή 2, εμφανίζεται η φωτογραφία στο ροζ πλαίσιο και στο πεδίο 3 εμφανίζεται η πρόβλεψη, όπως είναι στην εικόνα 38.

This deep learning approach can classify an image in three classes (normal or ischemia or infraction). Please select the type of categorization.

Please upload an SPECT MPI image for Coronary Artery Disease Diagnosis (.jpeg file)

No file selected.

Results

Image	Prediction
	Normal

Εικόνα 38

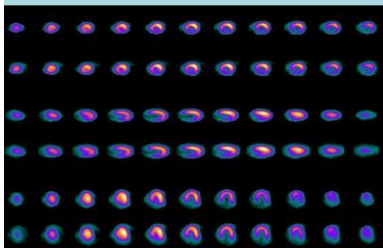
Ακολουθεί το ίδιο παράδειγμα αλλά για ισχαιμική φωτογραφία, στην εικόνα 39.

This deep learning approach can classify an image in three classes (normal or ischemia or infraction). Please select the type of categorization.

Please upload an SPECT MPI image for Coronary Artery Disease Diagnosis (.jpeg file)

No file selected.

Results

Image	Prediction
	Isxaimia

Εικόνα 39

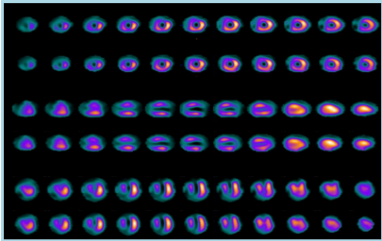
Και τέλος για φωτογραφία, που απεικονίζει έμφραγμα, στην εικόνα 40.

This deep learning approach can classify an image in three classes (normal or ischemia or infraction). Please select the type of categorization.

Please upload an SPECT MPI image for Coronary Artery Disease Diagnosis (.jpeg file)

No file selected.

Results

Image	Prediction
	Infraction

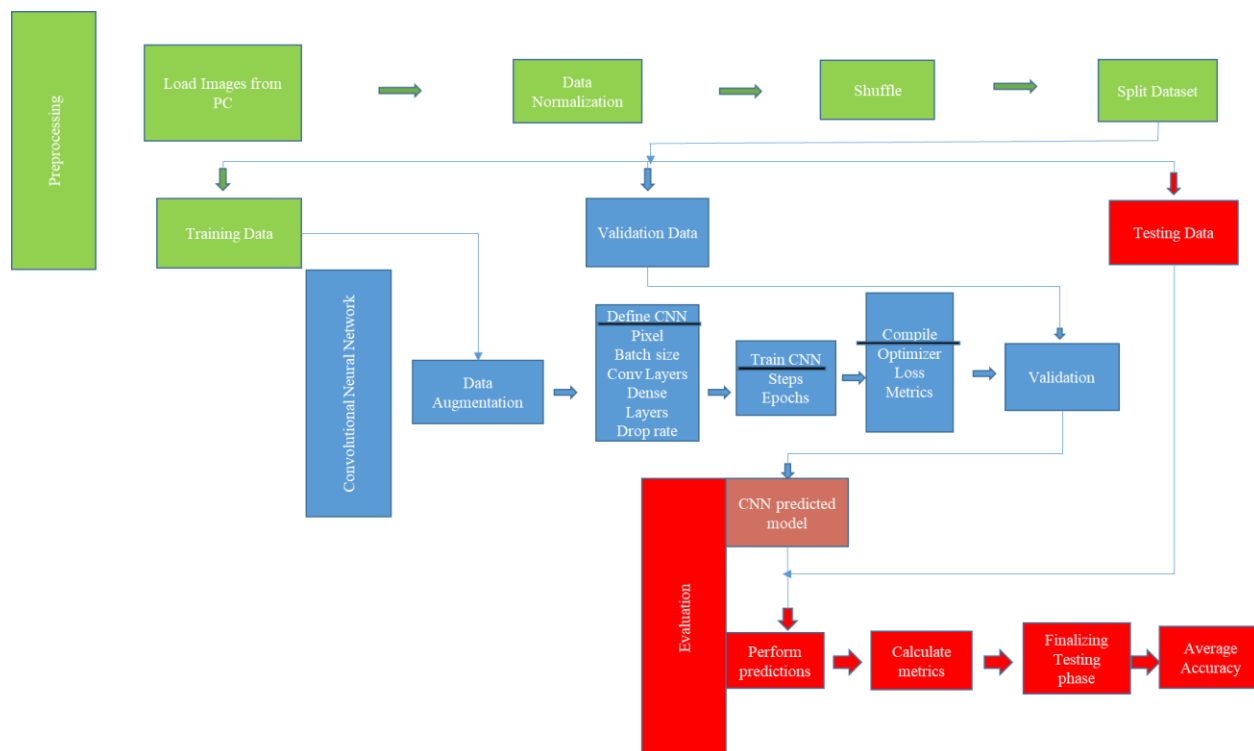
Εικόνα 40

Κεφ. 5: Αποτελέσματα εφαρμογής αλγορίθμων βαθιάς μάθησης στην διάγνωση της αιμάτωσης μυοκαρδίου

5.1 Περιγραφή της μεθοδολογίας της παρούσας έρευνας

Ακολουθεί το παρακάτω σχήμα το οποίο απεικονίζει όλη τη διαδικασία της μεθοδολογίας, σύμφωνα με όλα όσα ειπώθηκαν στα προηγούμενα κεφάλαια. Το παρακάτω σχήμα της εικόνας 41 δείχνει τα βήματα, που ακολουθήθηκαν για να εξαχθεί η απόδοση και το σφάλμα του αλγορίθμου στη συγκεκριμένη πτυχιακή.

Με πράσινο πλαίσιο είναι τα πρώτα στάδια, δηλαδή αρχικά είναι η φόρτωση των δεδομένων, στον αλγόριθμο, δηλαδή η μετατροπή τους από αρχεία εικόνων σε μορφή ικανή, ώστε να απεικονίζουν δεδομένα εισόδου για νευρωνικό δίκτυο. Στη συνέχεια ακολουθεί κανονικοποίηση των τιμών (προαιρετικό), «ανακάτεμα» δεδομένων, ώστε να είναι το αποτέλεσμα, όσο πιο τυχαίο γίνεται. Έπειτα έχουμε τον διαχωρισμό σε δεδομένα εκπαίδευσης (training data), δεδομένα ελέγχου (testing data) και δεδομένα εγκυρότητας (validation data). Μετά ακολουθούν τα μπλε πλαίσια, στα οποία έχουμε το εξής. Η αύξηση δεδομένων, κάποιες φορές κρίνεται απαραίτητη, για να είναι ο αριθμός των δεδομένων αρκετός για την σωστή λειτουργία του νευρωνικού δικτύου. Έπειτα πρέπει να καθοριστούν και οι τιμές των παραμέτρων του νευρωνικού δικτύου, δηλαδή, ο αριθμός των συνελκτικών δικτύων, το batch size, η τιμή των πίξελ, η τιμή των epochs, καθώς και αξιολογητή κ.ά. Τέλος με τα κόκκινα πλαίσια είναι η εξαγωγή αποτελεσμάτων και για αυτό το λόγο έχουμε τον υπολογισμό των μετρικών και του μέσου όρου.



Εικόνα 41

5.2 Περιγραφή του μοντέλου

Σε όλα τα μοντέλα έτρεξε αρχικά RGB συνελκτικό νευρωνικό δίκτυο το οποίο είχε προταθεί σε προηγούμενες εργασίες μας και περιγράφεται παρακάτω, και στη συνέχεια χρησιμοποιώντας τη μέθοδο μεταφορά εκμάθησης, εκτελέστηκαν τα δίκτυα VGG16, MobileNet, DenseNet. Δοκιμάστηκαν αρκετές παραμετροποιήσεις των μεταβλητών για να δούμε ποιο ταιριάζει καλύτερα στο ανάλογο σετ δεδομένων. Όσον αφορά για τις μεταβλητές του αλγορίθμου αξίζει να σημειωθεί ότι οι μεταβλητές που θέλαμε να προσαρμοστούνε είναι το πίξελ, το batch size, ο αριθμός των συνελκτικών επιπέδων, καθώς και οι κόμβοι που θα περιέχουν, όπως επίσης και τα πλήρως συνδεδεμένα δίκτυα και η μεταβλητή dropout rate = 0.2, έμεινε σταθερή, μιας και τα δεδομένα μας είναι ισορροπημένα.

Χρησιμοποιήθηκε ο ίδιος αλγόριθμος για όλες τις κατηγορίες, με μερικές τροποποιήσεις στις παραμέτρους για να εφαρμόζει ανάλογα την περίπτωση κάθε φορά και να δέχεται τις ανάλογες εξόδους. Παρακάτω ακολουθούν οι αρχιτεκτονικές που επιλέχθηκαν για όλα τα σετ δεδομένων, με έμφαση σε αυτές που εξήγαγαν τα καλύτερα αποτελέσματα.

Η ανάλυση των δικτύων γίνεται αρχικά για τα δεδομένα SPECT.

Πρώτα έχουμε το σετ α:

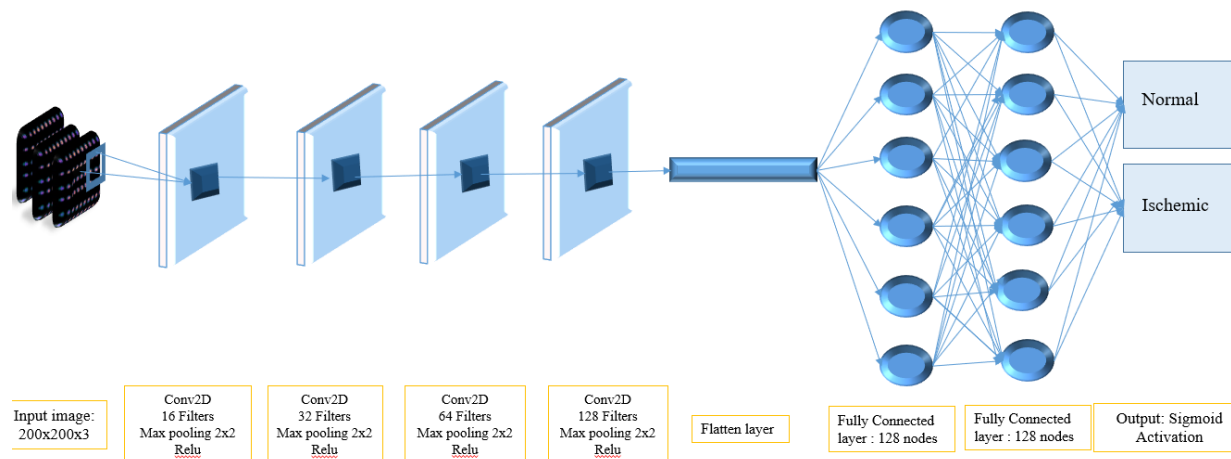
Σετ α) Φυσιολογικά – Ισχαιμία

Παρακάτω στον πίνακα 10 φαίνονται οι μεταβλητές που επιλέχθηκαν για το σετ α). Όπως αναφέρθηκε σε προηγούμενο κεφάλαιο οι μεταβλητές, οι οποίες καθορίζουν κατά κύριο λόγο την απόδοση του μοντέλου είναι ο αριθμός των πίζελς, το batch size, ο αριθμός των κόμβων και των επιπέδων των συνελικτικών δικτύων, και το αντίστοιχο για τα πλήρως συνδεδεμένα δίκτυα. Ύστερα, από αρκετή έρευνα καταλήξαμε στον παρακάτω συνδυασμό για το συγκεκριμένο σετ δεδομένων και αντίστοιχα ο κάθε πίνακας δηλώνει τους παραμέτρους για το ανάλογο σετ δεδομένων. Για τα πίζελ δοκιμάστηκαν 200x200, 250x250, 300x300, 350x350, για batch size 8-16-32-64, για convolutional layers 16-32-64, 16-32-64-128, 16-32-64-128-256 και για dense layers 32-32, 64-64, 128-128, 256-256. Τα αποτελέσματα, που εξήγαγε ο αλγόριθμος για κάθε σετ δεδομένων ακολουθεί σε επόμενο κεφάλαιο

RGB	Pixel	Batch size	Conv layers	Dense layers
	200x200	32	16-32-64-128	128-128

Πίνακας 10

Έχουμε το παρακάτω σχήμα, στην εικόνα 42, το οποίο απεικονίζει οπτικά το συνελκτικό δίκτυο, που μελετήθηκε για το σετ α) Φυσιολογικά – Ισχαιμία



Εικόνα 42

Αρχικά στην εικόνα 42 φαίνεται το σετ δεδομένων, το οποίο είναι η είσοδος μας στο δίκτυο. Στη συνέχεια έχουμε την ακολουθία Conv layers-Max Pooling-Dropout, το οποίο έγινε τέσσερις φορές, με μικρή τροποποίηση κάθε φορά στις παραμέτρους, για να είναι κλιμακωτή και πιο «γενική» η εξαγωγή των μοτίβων. Έπειτα έχουμε την συνάρτηση Flatten, η οποία όπως έχει αναφερθεί, μετατρέπει τον πίνακα πίξελ σε μονοδιάστατο πίνακα. Μετά έχουμε τα πλήρως συνδεδεμένα δίκτυα, τα οποία ακολουθούνται δύο φορές με παρόμοιες τιμές στις παραμέτρους και τέλος έχουμε την πρόβλεψη των δεδομένων, όπου μπορεί να είναι φυσιολογικά ή ισχαιμικά, λόγω του παρόντος σετ δεδομένων.

Επίσης στην εικόνα 43, το οποίο δείχνει ακριβώς την αρχιτεκτονική του συνελκτικού δικτύου. Κάθε μια γραμμή δείχνει τα βήματα του συνελκτικού νευρωνικού δικτύου, καθώς και την κάθε γραμμή αλγοριθμική, όπου συντάχθηκε. Στη δεξιά μεριά έχουμε τους παραμέτρους που επιλέχθηκαν για το κάθε επίπεδο, όπως περιγράφηκε παραπάνω, αλλά αυτή την φορά με ακριβή τιμή στις παραμέτρους.

Architecture: Pixel Size = (200, 200), batch_size = 32, dropout_rate = 0.2, epochs = 400	
Type of layers	Parameters
Conv2D	filters = 16, kernel_size = (3,3), activation = relu, input_shape = (250, 250, 3)
MaxPooling2D	Pool_size = (2,2)
Dropout	Drop_Rate = 0.2
Conv2D	filters = 32, kernel_size = (3,3), activation = relu
MaxPooling2D	Pool_size = (2,2)
Dropout	Drop_Rate = 0.2
Conv2D	filters = 64, kernel_size = (3,3), activation = relu
MaxPooling2D	Pool_size = (2,2)
Dropout	Drop_Rate = 0.2
Conv2D	filters = 128, kernel_size = (3,3), activation = relu
MaxPooling2D	Pool_size = (2,2)
Dropout	Drop_Rate = 0.2
Flatten	
Dropout	Drop_rate = 0.2
Dense	Nodes = 128, activation = relu
Dense	Nodes = 128, activation = relu
Dense	Node = 1, activation = sigmoid

Εικόνα 43

Τα δεδομένα μας εκτός από το μοντέλο RGB εκπαιδεύτηκαν και μέσω της διαδικασίας που ονομάζεται «μεταφορά εκμάθησης» (transfer learning). Τα δίκτυα, τα οποία επιλέχθηκαν είναι τα VGG-16, DenseNet-121, MobileNet. Στον πίνακα 11 φαίνονται οι αντίστοιχες τιμές στις παραμέτρους, αλλά στην περίπτωση της μεταφοράς εκμάθησης έχουμε διαφορετικές παραμέτρους. Έχουμε ξανά τον αριθμό πίξελ, το batch size, αλλά στην συνέχεια επειδή έχουμε το ήδη υλοποιημένο δίκτυο, επιλέγουμε πόσα επίπεδα δικά του θα συμπεριλάβουμε στο πρόβλημα κατηγοριοποίησης μας. Το VGG-16 έχει 16 επίπεδα, το DenseNet έχει 121 και το MobileNet έχει 28. Και η στήλη του πίνακα 11, με όνομα True layers υποδηλώνει πόσα επίπεδα του δικτύου συμπεριλάβαμε στην εκπαίδευση του μοντέλου. Τα dense layers είναι τα επίπεδα, που προσθέτουμε εμείς πριν τη εξαγωγή αποτελέσματος του δικτύου, τα οποία προστίθενται στο δίκτυο ύστερα από το τελευταίο επίπεδο του ήδη εκπαιδευμένου δικτύου.

Transfer Learning	Pixel	Batch size	True layers	Dense layers
VGG	200x200	16	True 13	GlobalAveragePooling2D 32 32
DenseNet	250x250	16	True 118	GlobalAveragePooling2D 16 16
MobileNet	300x300	8	True 26	GlobalAveragePooling2D 16 16

Πίνακας 11

Συνεχίζουμε για το επόμενο σετ δεδομένων.

Σετ β) Φυσιολογικά – Ισχαιμία, Έμφραγμα

Στον πίνακα 12 παρακάτω βλέπουμε ποιες παράμετροι δοκιμάστηκαν και τι τιμές κατέληξαν να έχουν μιας και εξήγαγαν τα καλύτερα αποτελέσματα, για την περίπτωση του RGB.

RGB	Pixel	Batch size	Conv layers	Dense layers
	200x200	16	16 32 64 128	GlobalAveragePooling2D 128 128

Πίνακας 12

Και στη συνέχεια στον πίνακα 13 βλέπουμε τις μεταβλητές για τη μέθοδο μεταφορά εκμάθησης με τα δίκτυα VGG-16, DenseNet, MobileNet.

Transfer Learning	Pixel	Batch size	True layers	Dense layers
VGG	250x250	16	True 13	GlobalAveragePooling2D 32 32
DenseNet	200x200	16	True 118	GlobalAveragePooling2D 32 32
MobileNet	250x250	8	True 26	GlobalAveragePooling2D 16 16

Πίνακας 13

Το επόμενο σετ δεδομένων είναι

Σετ γ) Ισχαιμικά, έμφραγμα

Φαίνονται και οι τιμές των παραμέτρων, που επιλέχθηκαν στον πίνακα 14, όσον αφορά για το RGB.

RGB	Pixel	Batch size	Conv layers	Dense layers
	250x250	8	16 32 64 128 256	64-64

Πίνακας 14

Και όσον αφορά για τη μέθοδο μεταφορά εκμάθησης στο σετ γ) έχουμε τις εξής τιμές στον πίνακα 15, για τα δίκτυα VGG, DenseNet και MobileNet.

Transfer Learning	Pixel	Batch size	True layers	Dense layers
VGG	250x250	16	True 13	GlobalAveragePooling2D 32 32
DenseNet	300x300	8	True 26	GlobalAveragePooling2D 16 16
MobileNet	300x300	8	True 26	GlobalAveragePooling2D 16 16

Πίνακας 15

Στη συνέχεια ακολουθεί το σετ, το οποίο είναι για 3 κλάσεις. Δεν έχει κάποια διαφορά, όσον αφορά για την επιλογή των παραμέτρων, συγκριτικά με το πρόβλημα των 2 κλάσεων. Για το RGB έχουμε τις τιμές των παραμέτρων στον πίνακα 16.

Σετ δ) Φυσιολογικά-Ισχαιμικά - έμφραγμα

RGB	Pixel	Batch size	Conv layers	Dense layers
	250x250	16	16 32 64 128	128-128

Πίνακας 16

Για τη μέθοδο μεταφοράς εκμάθησης έχουμε τις εξής μεταβλητές για τις 3 κλάσεις στον πίνακα 17, για τα δίκτυα VGG, DenseNet, MobileNet.

Transfer Learning	Pixel	Batch size	True layers	Dense layers
VGG	200x200	16	True 14	GlobalAveragePooling2D 32 32
DenseNet	250x250	32	True 26	GlobalAveragePooling2D 32 32
MobileNet	200x200	16	True 26	GlobalAveragePooling2D 16 16

Πίνακας 17

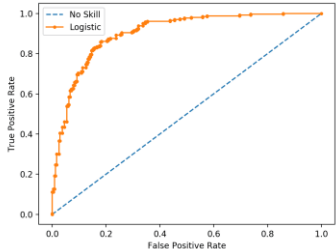
5.3 Παρουσίαση μετρικών απόδοσης μοντέλου

Στη διαδικασία της έρευνας των σωστών παραμέτρων χρησιμοποιήθηκαν οι μετρικές απόδοσής για να γίνει η σύγκριση με ποιες παραμέτρους θα καταλήξουμε. Οι μετρικές απόδοσης που χρησιμοποιήθηκαν είναι οι εξής και φαίνονται στον πίνακα 18:

Validation accuracy	Υπολογίζεται ως ο αριθμός των σωστών προβλέψεων διαιρώντας με το συνολικό αριθμό δεδομένων του τεστ επαλήθευσης (validation dataset). [124]
Validation loss	Αποτελεί το σφάλμα ανάμεσα στο πόσο επιτυχές είναι το validation dataset. [124]

Testing accuracy	Το αποτέλεσμα της διαίρεσης των σωστών προβλέψεων με το συνολικά δεδομένα του τεστ ελέγχου (testing dataset). [124]									
Test loss	Επιδεικνύει το σφάλμα του μοντέλου στο testing dataset. [124]									
Πίνακας σύγχυσης (Confusion matrix)	Είναι μια μέτρηση απόδοσης για προβλήματα ταξινόμησης μηχανικής μάθησης όπου μπορεί να αφορά δύο ή περισσότερες τάξεις. Είναι ένας πίνακας με 4 διαφορετικούς συνδυασμούς προβλεπόμενων και πραγματικών τιμών. <table><tr><td></td><td>Class 1 Predicted</td><td>Class 2 Predicted</td></tr><tr><td>Class 1 Actual</td><td>TP</td><td>FN</td></tr><tr><td>Class 2 Actual</td><td>FP</td><td>TN</td></tr></table> [124]		Class 1 Predicted	Class 2 Predicted	Class 1 Actual	TP	FN	Class 2 Actual	FP	TN
	Class 1 Predicted	Class 2 Predicted								
Class 1 Actual	TP	FN								
Class 2 Actual	FP	TN								
TP	Ο αριθμός των δεδομένων που προβλέφθηκαν ως θετικό και είναι σωστό. [124]									
TN	Ο αριθμός των δεδομένων που προβλέφθηκαν ως αρνητικό και είναι σωστό. [124]									
FP	Ο αριθμός των δεδομένων που προβλέφθηκαν ως θετικό και είναι λάθος. [124]									
FN	Ο αριθμός των δεδομένων που προβλέφθηκαν ως αρνητικό και είναι λάθος. [124]									

Precision	$Precision = \frac{TP}{TP + FP}$ Από τον αριθμό των δεδομένων που προβλέφθηκαν ως θετικά, πόσα είναι σωστά [125]
Accuracy	$Accuracy = \frac{TrueNegatives + TruePositive}{TruePositive + FalsePositive + TrueNegative + FalseNegative}$ Ο αριθμός των δεδομένων που προβλέφθηκαν σωστά. [125]
Loss	Υποδηλώνει πόσο άσχημα ή όχι συμπεριφέρεται ένα μοντέλο μετά από κάθε επανάληψη του κατηγοριοποιητή. [125]
Recall	$Recall = \frac{TP}{TP + FN}$ Υπολογίζεται ως η αναλογία μεταξύ του αριθμού των θετικών δειγμάτων που ταξινομούνται σωστά ως θετικά προς τον συνολικό αριθμό θετικών δειγμάτων. Είναι η ικανότητα του μοντέλου να ανιχνεύει θετικά δείγματα. Όσο υψηλότερη είναι τόσο περισσότερα θετικά δείγματα εντοπίστηκαν. [125]
F1-Score	$F - measure = \frac{2 * Recall * Precision}{Recall + Precision}$ Η βαθμολογία F βοηθά στη μέτρηση της precision και της υψηλής recall ταυτόχρονα. Χρησιμοποιεί το αρμονικό μέσο αντί του αριθμητικού μέσου έτσι ώστε οι ακραίες τιμές να μηδενιστούν. [125]
Sensitivity	(True Positive rate) μετρά την αναλογία των θετικών που έχουν ταυτοποιηθεί σωστά [120]
Specificity	(True Negative rate) μετρά την αναλογία των αρνητικών που προσδιορίζονται σωστά [120]

AUC	Είναι το μέτρο της ικανότητας ενός ταξινομητή να διακρίνει μεταξύ τάξεων και χρησιμοποιείται ως σύνοψη της καμπύλης ROC. Όσο υψηλότερη είναι η AUC, τόσο καλύτερη είναι η απόδοση του μοντέλου στη διάκριση μεταξύ θετικών και αρνητικών εξόδων [122]
ROC Curve	 Είναι η γραφική αναπαράσταση της ικανότητας του μοντέλου να διαχωρίζει σωστά τα δεδομένα. Δημιουργείται σχεδιάζοντας τις τιμές TPR (true positive rate) και FPR (false positive rate). Η μεταβλητή TPR ονομάζεται αλλιώς sensitivity ή recall και η FPR ονομάζεται ως «1-TPR». [122]
Plots	Η απεικόνιση της απόδοσης και του σφάλματος [125]

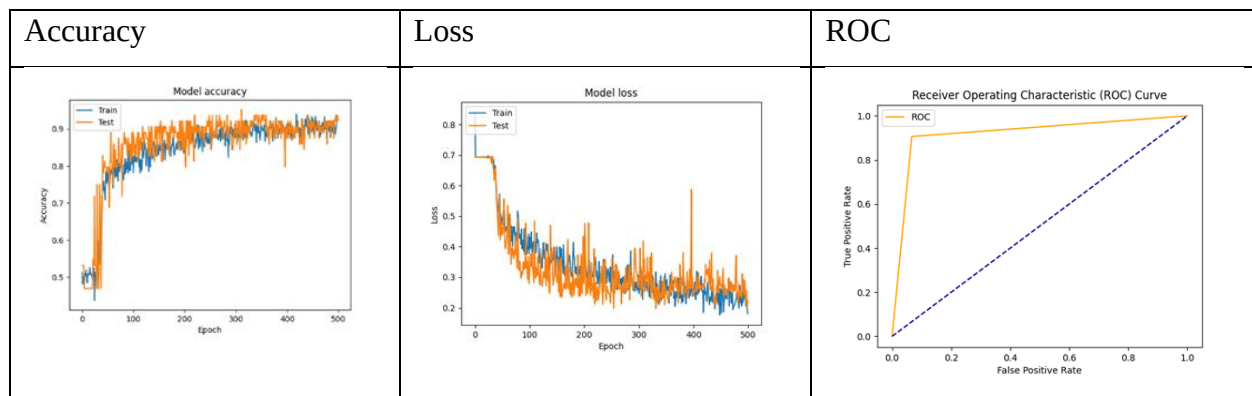
Πίνακας 18

5.4 Αποτελέσματα εφαρμογής αλγορίθμου

Στη συνέχεια ακολουθούν τα αποτελέσματα, που εξήγαγε ο κάθε αλγόριθμος, ανάλογα με το σετ δεδομένων που χρησιμοποιείται κάθε φορά.

Σετ Α) Φυσιολογικά – Ισχαιμία

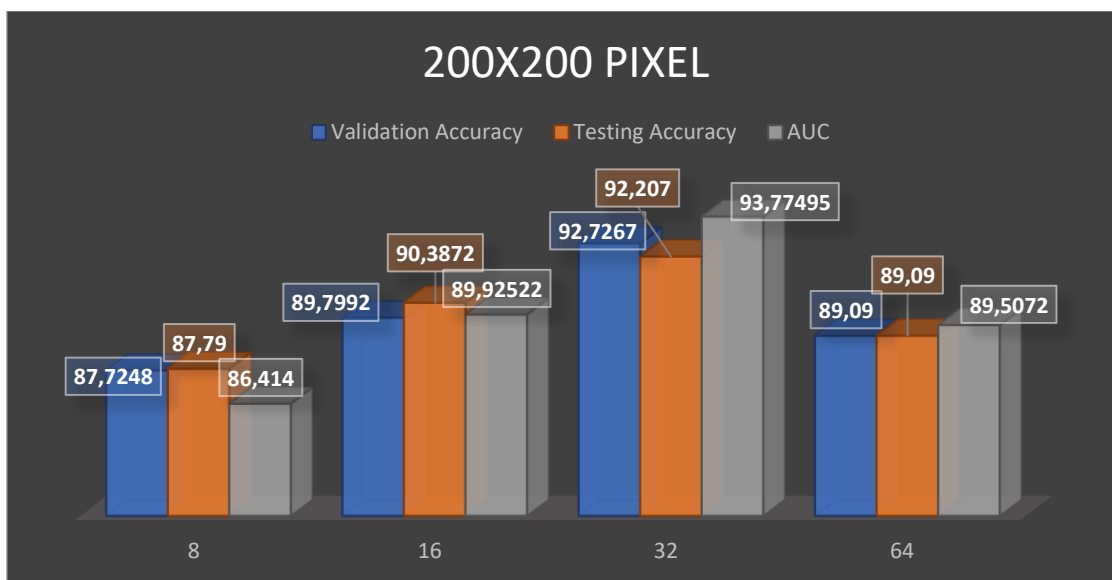
Στο σετ Α, τα καλύτερα αποτελέσματα τα οποία προέκυψαν με την αρχιτεκτονική Pixel:200x200, batch size:32, conv layers 16-32-64-128 και dense layers:128-128, μετά από 400 επαναλήψεις, απεικονίζονται στα παρακάτω διαγράμματα, και είναι τα εξής (Πίνακας 19): Στον πίνακα απεικονίζεται η απόδοση αρχικά, στη δεύτερη στήλη το σφάλμα και τέλος η καμπύλη ROC.



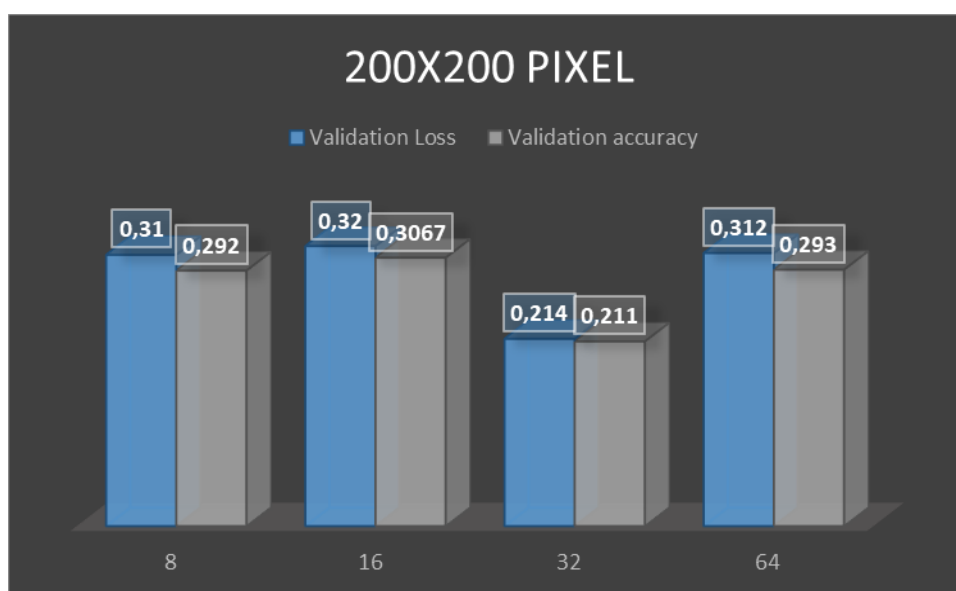
Πίνακας 19

Γνωρίζουμε ότι η καλύτερη αρχιτεκτονική ήταν με 32 batch size, όμως για να καταλήξουμε σε αυτή την αρχιτεκτονική έγιναν συγκρίσεις με διάφορα batch sizes, όπως 8-16-32-64, για να φανεί ποια αρχιτεκτονική έχει την καλύτερη απόδοση.

Το πρώτο σχήμα στην Εικόνα 44 δείχνει την απόδοση συγκριτικά με τα διαφορετικά batch sizes και το δεύτερο στην εικόνα 44 δείχνει το σφάλμα συγκριτικά αντίστοιχα.

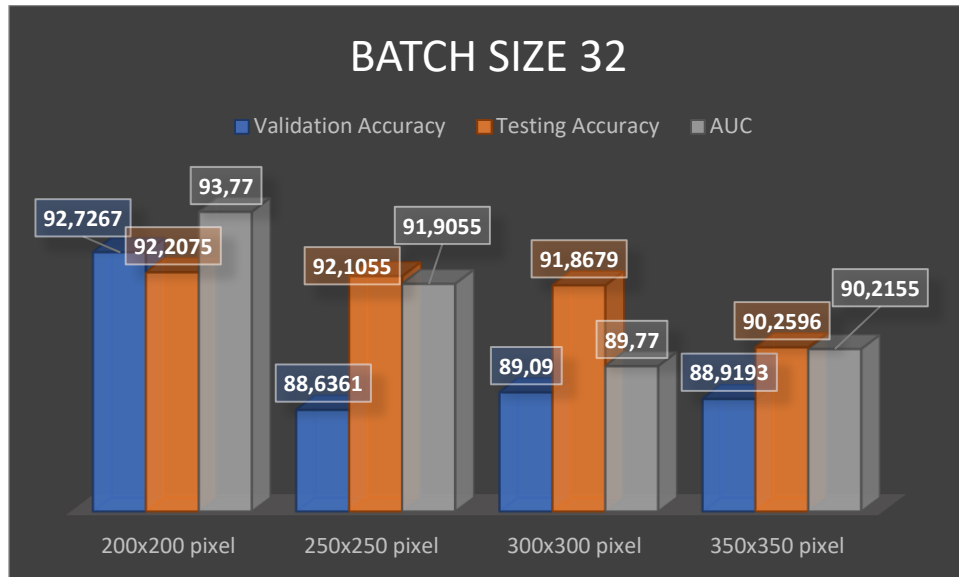


Εικόνα 44

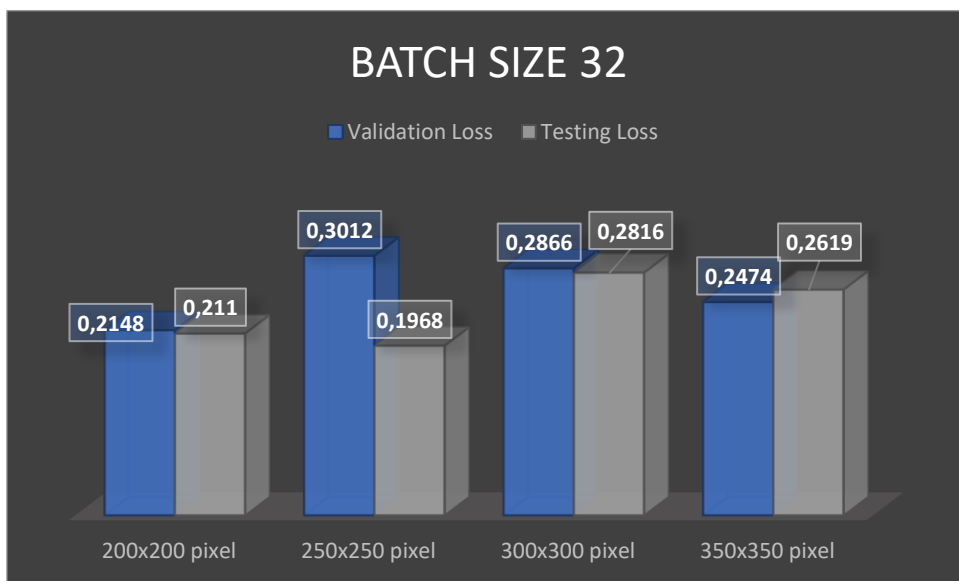


Εικόνα 45

Αφού έγινε ξεκάθαρο ότι με 32 batch size έχουμε τα καλύτερα αποτελέσματα, σειρά έχουν τα πίζελ και να δούμε πως τροποποιείται η απόδοση και το σφάλμα, με τη διαφοροποίηση της συγκεκριμένης τιμής. Οι παραμετροποιήσεις έγιναν με τιμές όπως 200x200, 250x250, 300x300 και 350x250. Παρατηρούμε ότι με 200x200 πίζελ, έχουμε τα καλύτερα αποτελέσματα, σύμφωνα με την εικόνα 46 και 47, που απεικονίζουν την απόδοση και το σφάλμα αντίστοιχα.



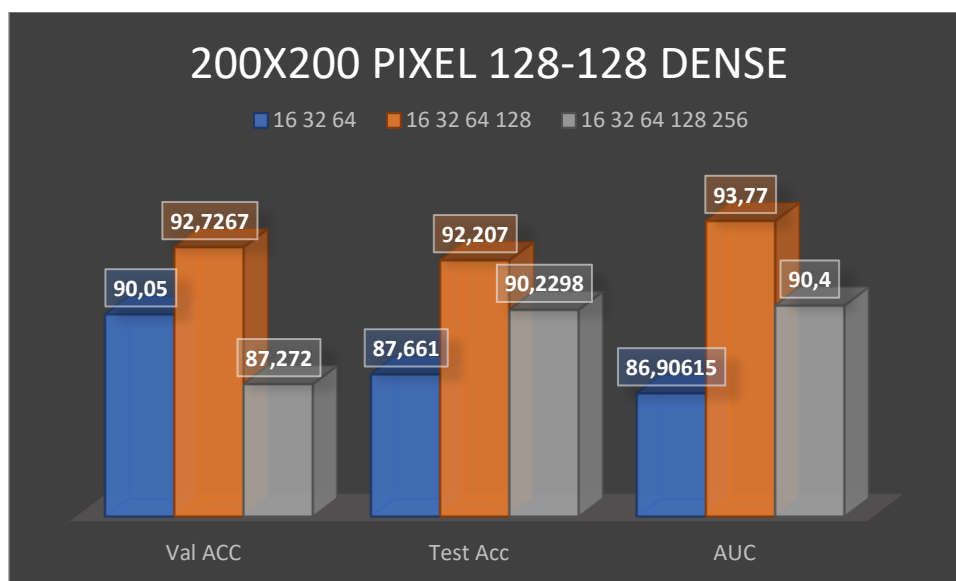
Εικόνα 46



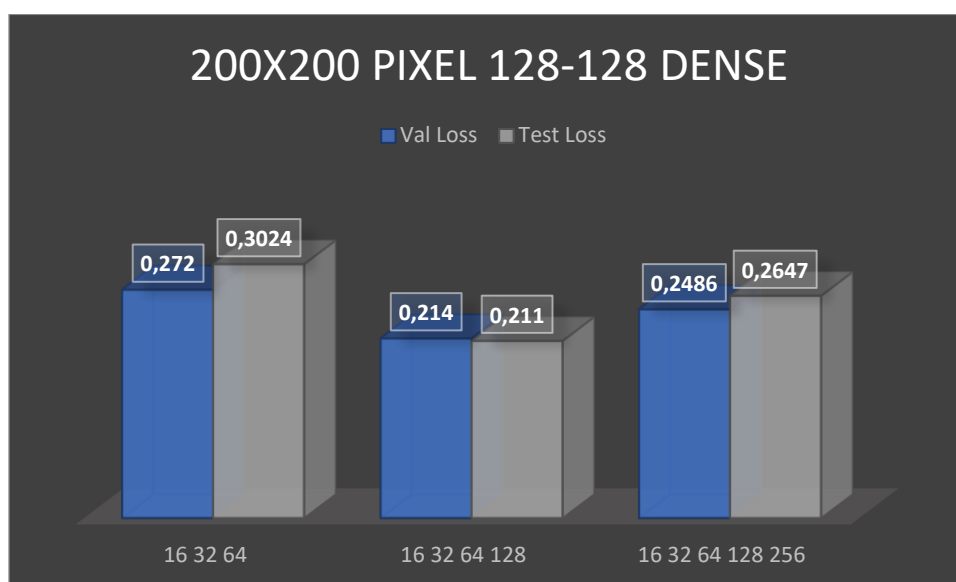
Εικόνα 47

Στη συνέχεια έχουμε την επιλογή για τα συνελκτικά επίπεδα, μιας και καθορίσαμε τα πίξελ και το batch size. Ακολουθεί αλλαγή στον αριθμό συνελκτικών δικτύων, καθώς και στους εκάστοτε κόμβους, για να επαληθεύσουμε συγκριτικά με τους συνδυασμούς 16-32-64, 16-32-64-128, 16-

32-64-128-256, ποιο είναι το καλύτερο, κρατώντας σταθερό τον αριθμό των πλήρως συνδεδεμένα δικτύων με τιμή 128-128, δύο επίπεδα, δηλαδή. Στην εικόνα 50 βλέπουμε την απόδοση και στην εικόνα 51 το σφάλμα, για τους διάφορους συνδυασμούς, και ότι έχουμε τα καλύτερα αποτελέσματα για τον συνδυασμό 16-32-64-128.

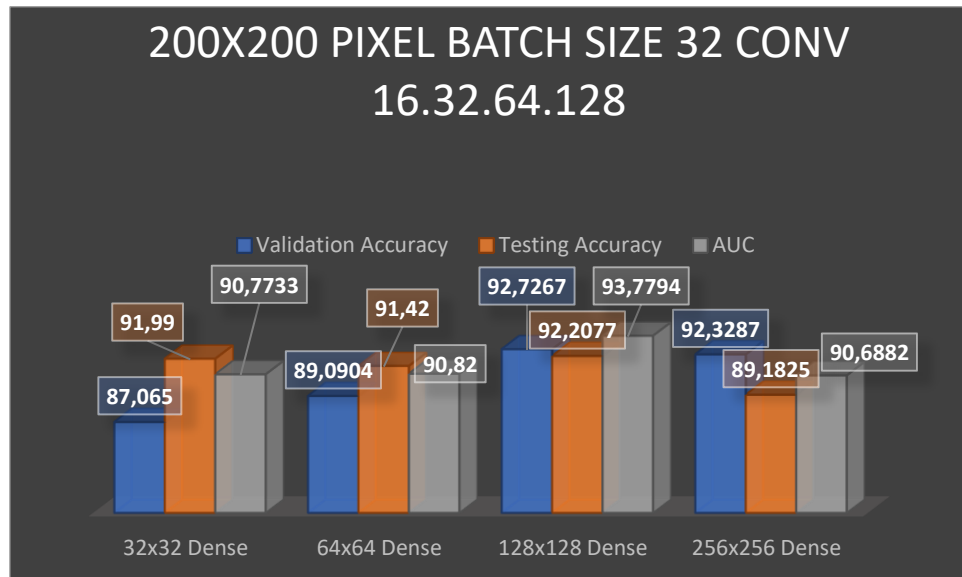


Εικόνα 50

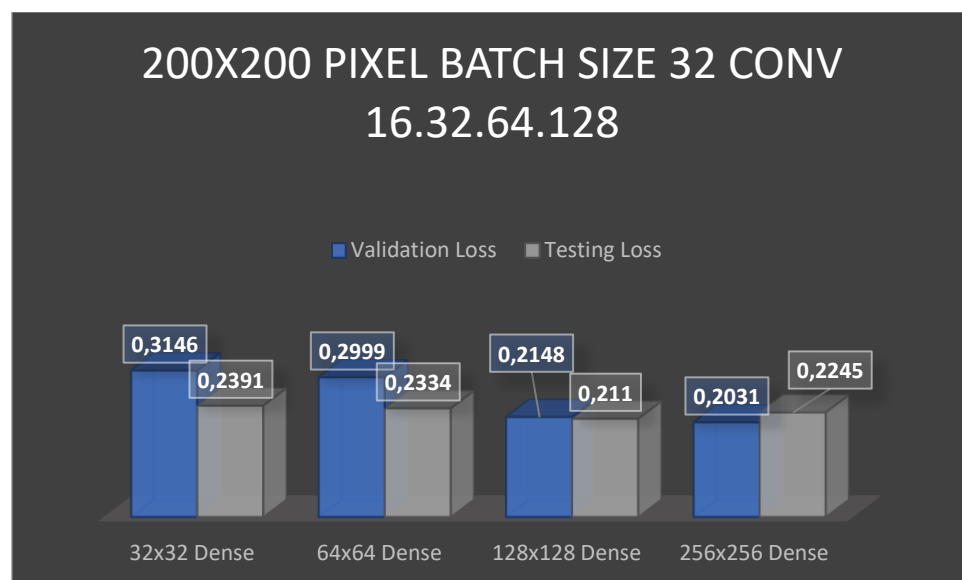


Εικόνα 51

Τέλος, αφού καταλήξαμε στο batch size και στο πίξελ και στα συνελικτικά δίκτυα ακολουθεί η επιλογή των πλήρως συνδεδεμένων δικτύων. Τα παρακάτω διαγράμματα δημιουργούνται όταν αλλάζουμε τα πλήρως συνδεδεμένα δίκτυα και κρατώντας σταθερές όλες τις άλλες μεταβλητές. Απεικονίζεται η απόδοση στην εικόνα 48 και το σφάλμα στην εικόνα 49. Βλέπουμε ότι με 128-128 είναι πιο καλά τα αποτελέσματα.

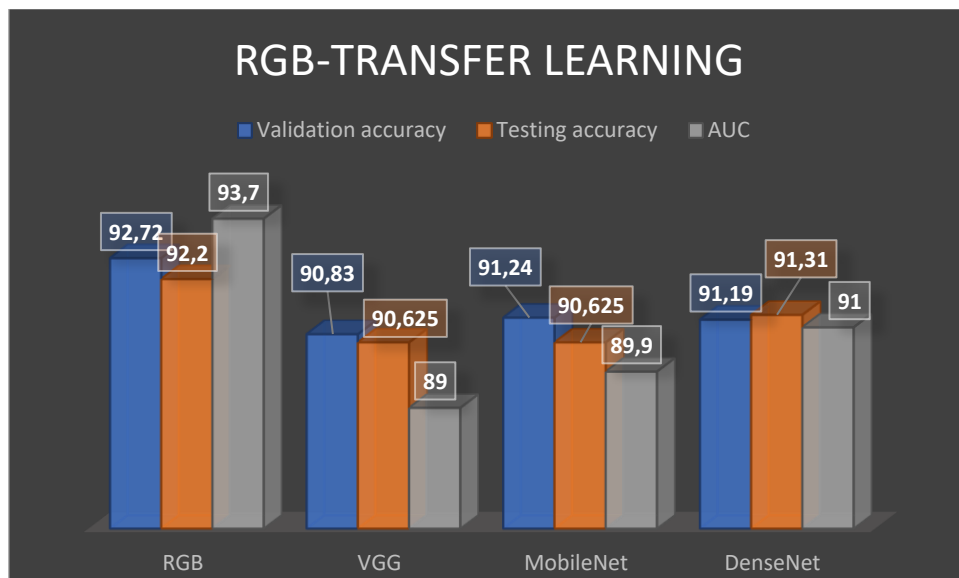


Εικόνα 48

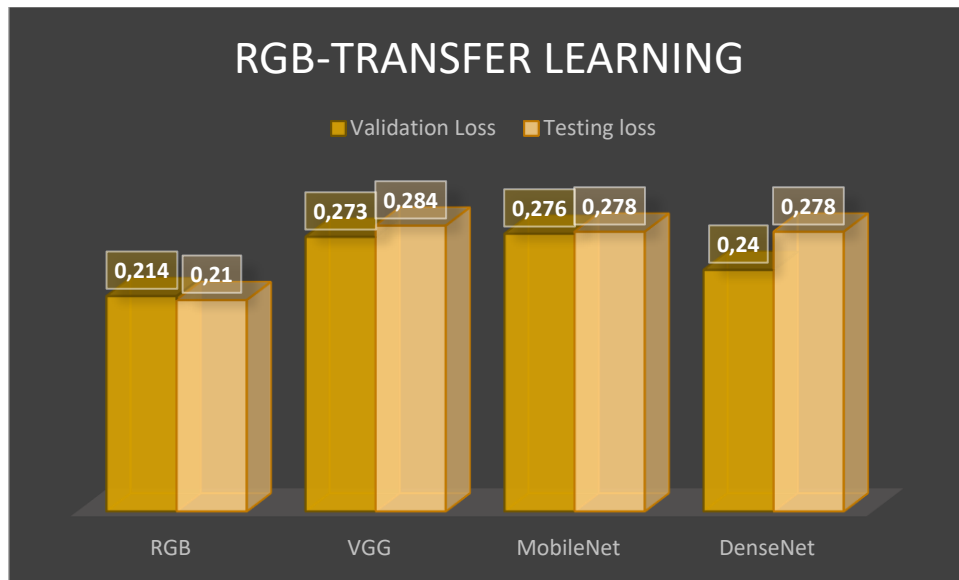


Εικόνα 49

Επίσης εκτός από την απλή τεχνική RGB υλοποιήθηκε και η μέθοδος transfer learning, όπου αναλύθηκε η μεθοδολογία της παραπάνω. Επιλέχθηκαν τα δίκτυα VGG16, DenseNet και MobileNet και συγκρίνεται η απόδοση, όπως φαίνεται στην εικόνα 52 και το σφάλμα στην εικόνα 53 συγκριτικά με το RGB, όπου αναλύθηκε ήδη, για να αποφανθούν τα καλύτερα αποτελέσματα. Οπότε έχουμε την καλύτερη αρχιτεκτονική για το RGB, και για κάθε ένα από τα δίκτυα ξεχωριστά. Το RGB φαίνεται ότι λειτουργεί καλύτερα, όσον αφορά για το Validation accuracy, αλλά για testing accuracy και auc το Mobilenet έχει καλύτερα αποτελέσματα. Και όσον αφορά για το σφάλμα το RGB έχει πάει καλύτερα.



Εικόνα 52



Εικόνα 53

Όσον αφορά για το σετ β) φυσιολογικά – ισχαιμία, έμφραγμα, ακολουθήθηκε η διαδικασία που προηγήθηκε παραπάνω, έγινε δηλαδή έρευνα στις παραμέτρους για να εξαχθεί το καλύτερο αποτέλεσμα και τα αποτελέσματα παραθέτονται σε μορφή πίνακα (βλέπε Πίνακα 20). Η καλύτερη αρχιτεκτονική είναι Pixel: 200x200, Batch size 16, Conv 16.32.64.128 Dense 128-128,

Αλλάζοντας το batch size, παρατηρούμε ότι το 16 απέδωσε καλύτερα.

Pixel: 200x200, Batch size , Conv 16.32.64.128 Dense 128-128,					
Batch size	Validation accuracy	Validation loss	Testing Accuracy	Testing Loss	AUC
8	89,8127	0,2577	92,2445	0,2295	0,9305
16	93,1567	0,1857	93,6224	0,2034	0,94
32	91,2688	0,2416	91,6284	0,2552	0,9249
64	88,635	0,2416	86,2244	0,3309	0,8620

Πίνακας 20

Κρατώντας σταθερό το batch size και αλλάζοντας τα πιξελ, παρατηρούμε ότι συνολικά στις μετρικές απόδοσης έχουμε καλύτερα αποτελέσματα με 200x200 πιξελ, όπως φαίνεται στον πίνακα 21, δηλαδή το validation loss είναι το χαμηλότερο συγκριτικά με τους υπόλοιπους συνδυασμούς pixels. Καθώς και το testing accuracy είναι αισθητά καλύτερο, όπως και το testing loss και το

AUC. Όσον αφορά για το validation accuracy παρατηρούμε ότι ο συνδυασμός 250x250, έχει αποδώσει λίγο καλύτερα, αλλά συνολικά το 200x200, έχει εξάγει καλύτερα αποτελέσματα.

Pixel, Batch size=16, Conv 16.32.64.128 Dense 128-128,					
Pixel	Validation accuracy	Validation loss	Testing Accuracy	Testing Loss	AUC
200x200	93,1567	0,1857	93,6224	0,2034	0,94
250x250	93,3069	0,2207	88,3672	0,2639	0,893
300x300	90,1574	0,2352	91,8364	0,2307	0,9281
350x350	88,7218	0,2737	92,0405	0,2189	0,9253

Πίνακας 21

Παραμετροποιώντας τα συνελκτικά επίπεδα, παρατηρούμε ότι με τέσσερα συνελκτικά δίκτυα και με τιμές 16-32-64-128 έχουμε τα καλύτερα αποτελέσματα σε όλες τις μετρικές, όπως φαίνεται στον πίνακα 22. Για να είναι ξεκάθαρη ότι αυτή η αρχιτεκτονική έχει φέρει την υψηλότερη απόδοση και το χαμηλότερο σφάλμα έγιναν τρεξίματα με διαφορετικό αριθμό συνελκτικών δικτύων και διαφορετική τιμή στους κόμβους για να είναι το δίκτυο, όσο γίνεται πιο αξιόπιστο για τα δεδομένα αυτά, και καταλήξαμε στον συνδυασμό 16-32-64-128.

Pixel: 200x200, Batch size=16, Conv Dense layers:128-128					
Dense layers	Validation accuracy	Validation loss	Testing Accuracy	Testing Loss	AUC
16 32 64	92,9521	0,224	89,7958	0,2560	0,9101
16 32 64 128	93,1567	0,1857	93,6224	0,2034	0,94
16 32 64 128 256	91,269	0,2613	90,8149	0,25133	0,9165

Πίνακας 22

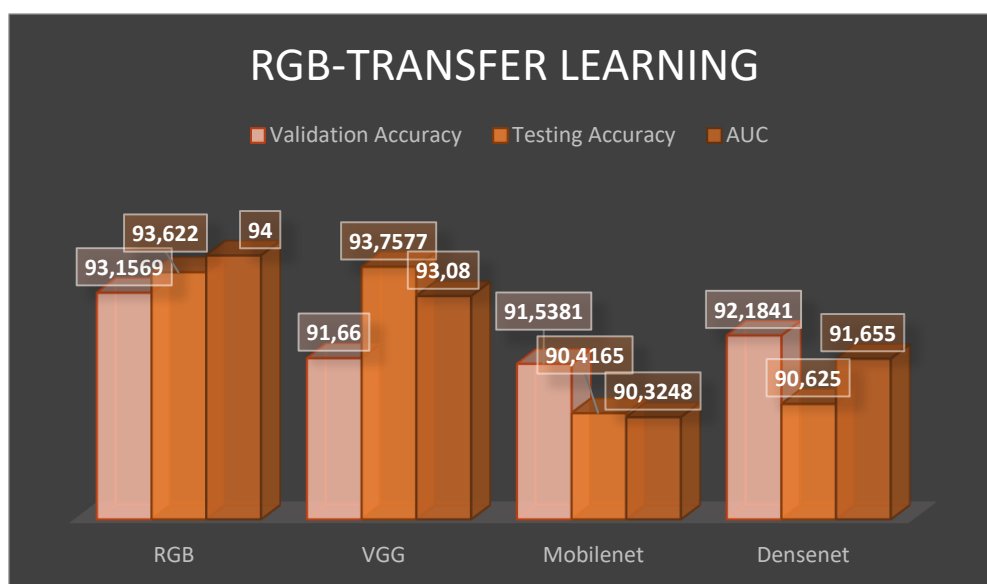
Αλλάζοντας τα πλήρως συνδεδεμένα επίπεδα, που αποτελούνται από συνδυασμό layers 32-32, έχουμε την καλύτερη validation accuracy συγκριτικά με την επιλογή 32-32, 64-64 και 256-256,

όπως φαίνεται στον πίνακα 23. Το ίδιο ισχύει και για το validation loss, testing accuracy και AUC. Εξάιρεση αποτελεί το testing loss, όπου η επιλογή για dense 256-256 έχει φέρει καλύτερα αποτελέσματα, αλλά δεν θα ήταν σωστό να επιλεγεί ως προτεινόμενη γιατί οι υπόλοιπες μετρικές δεν είναι συγκρίσιμες με το 128-128.

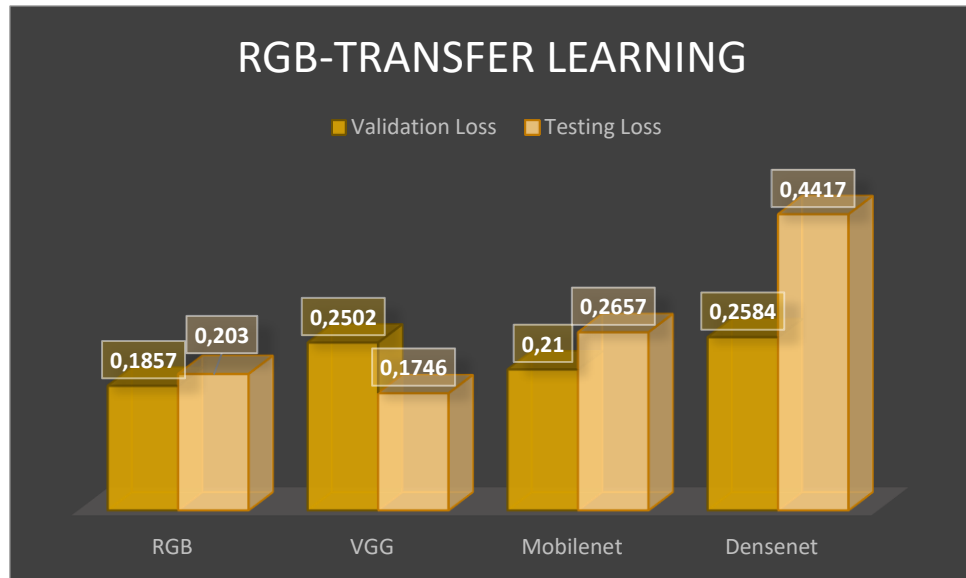
Pixel: 200x200, Batch size 16, Epochs:400, Dropout rate 0.2, Conv 16.32.64.128 Dense					
Batch size	Validation accuracy	Validation loss	Testing Accuracy	Testing Loss	AUC
32 32	90,3595	0,2622	91,0203	0,2704	0,9145
64 64	89,8166	0,241	91,0203	0,2559	0,9142
128 128	93,1567	0,1857	93,6224	0,2034	0,94
256 256	89,5448	0,2317	93,3673	0,1684	0,9386

Πίνακας 23

Όπως για το προηγούμενο σετ δεδομένων, που εφαρμόστηκε η μέθοδος transfer learning (μεταφορά γνώσης), έτσι εφαρμόστηκε και εδώ. Οπότε στα επόμενα διαγράμματα βλέπουμε ότι υλοποιήθηκαν τα δίκτυα VGG-16, MobileNet, DenseNet με σκοπό την αναπαράσταση των αποτελεσμάτων τους, αλλά και τι ποσοστό απόδοσης και σφάλματος έχουν συγκριτικά με το RGB. Στην εικόνα 54 βλέπουμε την απόδοση και στην εικόνα 55 το σφάλμα. Είναι κατανοητό ότι το RGB έχει εξάγει τα καλύτερα αποτελέσματα συνολικά, σύμφωνα με τα υπόλοιπα δίκτυα, αν και το VGG έχει εξάγει ικανοποιητικά αποτελέσματα.



Εικόνα 54



Εικόνα 55

Για το σετ δεδομένων γ) ισχαιμία-έμφραγμα η καλύτερη αρχιτεκτονική είναι Pixel: 250x250, Batch size 8, Conv 16.32.64.128.256 Dense 64-64. Για να αποφανθεί ότι αυτή είναι η καλύτερη αρχιτεκτονική έγινε έρευνα με διάφορα batch sizes (8,16,32,64) για να δούμε τι αποτελέσματα εξάγονται με αυτά στον Πίνακα 22 . Παρατηρούμε ότι για κάθε μετρική απόδοση (εκτός AUC) υπάρχει μια τιμή που είναι καλύτερη από το προτεινόμενο batch size. Παρόλα αυτά και οι τέσσερις μετρικές απόδοσης αποδίδουν συνολικά καλύτερα με το batch size 8.

Pixel: 250x250, Batch size, Conv 16.32.64.128.256 Dense 64-64,					
Batch size	Validation accuracy	Validation loss	Testing Accuracy	Testing Loss	AUC
8	91,285	0,3131	92,2405	0,278	0,914
16	83,596	0,3943	88,3620	0,32	0,865
32	87,120	0,287	87,929	0,304	0,876
64	85,984	0,37889	87,9	0,306	0,8486

Πίνακας 24

Αφού είδαμε ότι τα καλύτερα αποτελέσματα είναι με batch size 8, συνεχίζουμε την έρευνα για τα πιξελ. Οι συνδυασμοί που εξετάστηκαν είναι 200x200, 250x250 και 300x200, 250x250 και 300x300. Η προτεινόμενη αρχιτεκτονική περιέχει τον συνδυασμό 250x250, καθώς η απόδοση και το σφάλμα, που έχουν εξαχθεί είναι καλύτερα.

Pixel , Batch size=8, Conv 16.32.64.128-256 Dense 64-64,					
Pixel	Validation accuracy	Validation loss	Testing Accuracy	Testing Loss	AUC
200x200	87,120	0,287	87,929	0,304	0,876
250x250	91,285	0,313	92,24	0,278	0,914
300x300	84,09	0,331	86,219	0,38	0,856

Πίνακας 25

Μιας και επιλέχθηκαν και τα πιξελ, σειρά έχει η ανάλυση των συνελκτικών δικτύων. Έγινε έρευνα για 3 επίπεδα (16-32-64) για 4 επίπεδα (16-32-64-128) και για 5 επίπεδα (16-32-64-128-256). Ύστερα από τις ανάλυση των αποτελεσμάτων επιλέχθηκε η ακολουθία 16-32-64-128-256 να είναι η προτεινόμενη.

Pixel: 250x250, Batch size=8, Conv Dense layers:64-64					
Dense layers	Validation accuracy	Validation loss	Testing Accuracy	Testing Loss	AUC
16 32 64	87,12	0,3655	85,3442	0,389	0,849
16 32 64 128	87,12	0,28775	87,929	0,304175	0,87
16 32 64 128 256	91,285	0,3131	92,2405	0,278	0,914

Πίνακας 26

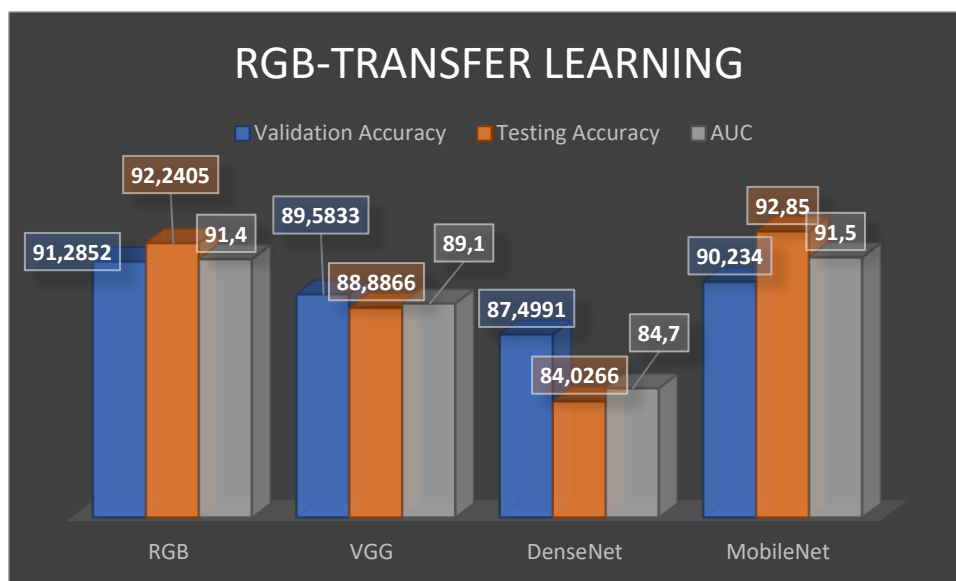
Τέλος έχουμε και τα πλήρως συνδεδεμένα δίκτυα, όπου έγιναν τρεξίματα για τους συνδυασμούς 32-32, 64-64, 128-128, 256-256. Όπως φαίνεται στον πίνακα 25, τα αποτελέσματα έχουν πιο καλές τιμές για την αρχιτεκτονική με δίκτυα 64-64, με αρκετή διαφορά συγκριτικά με τα υπόλοιπα.

Pixel: 250x250, Epochs:400, Dropout rate 0.2, Conv 16.32.64.128.256 Dense
--

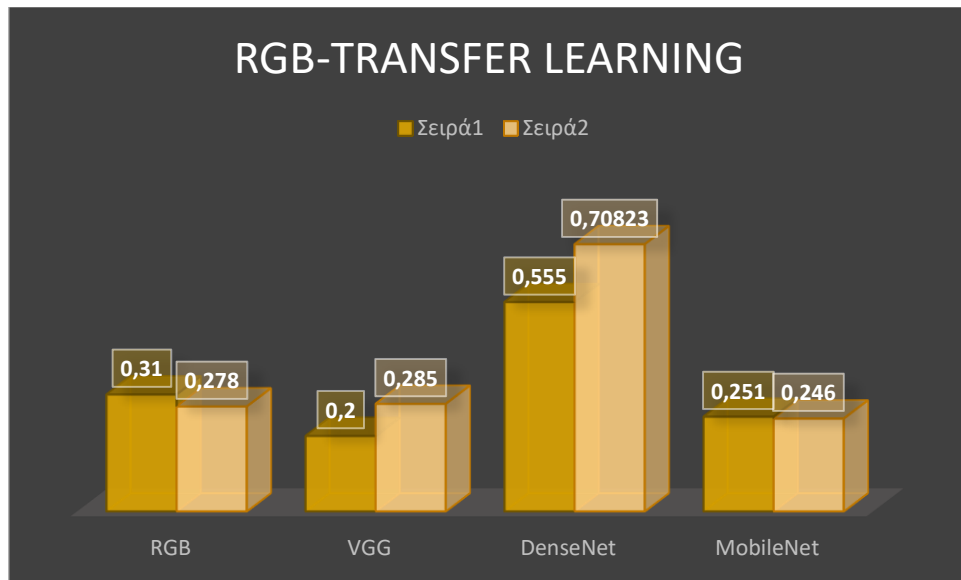
Batch size	Validation accuracy	Validation loss	Testing Accuracy	Testing Loss	AUC
32 32	84,847	0,395	86,2	0,329	0,862
64 64	91,28	0,3131	92,24	0,2787	0,914
128 128	87,1209	0,287	87,929	0,304	0,876
256 256	89,014	0,283	88,361	0,30	0,865

Πίνακας 27

Ομοίως και με τα προηγούμενα σετ δεδομένων έγινε υλοποίηση της μεθόδου transfer learning με τα δίκτυα VGG-16, MobileNet. Στη συνέχεια έγινε σύγκριση με το RGB και σύμφωνα με τις εικόνες 56 και 57 ακολουθούν τα παρακάτω διαγράμματα, τα οποία απεικονίζουν την απόδοση και το σφάλμα. Παρατηρούμε ότι για αυτό το σετ δεδομένων, το RGB απέδωσε καλύτερα, αν και το MobileNet έδωσε αρκετά καλά αποτελέσματα.



Εικόνα 56



Εικόνα 57

Εκτός από τα σετ δεδομένων με 2 κλάσεις, έχουμε και με 3 το εξής σετ δεδομένων:

Σετ δ) Φυσιολογικά – Ισχαιμικά – Έμφραγμα

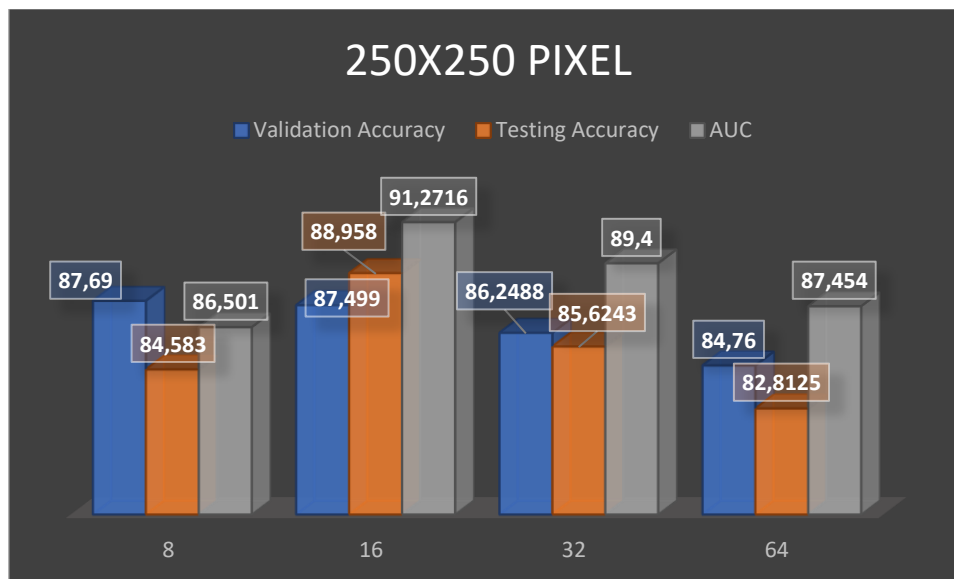
Για το best rgb τρέξαμε 700 epoch, καθώς είναι πιο λογικό να χρειάζεται περισσότερο χρόνο να γίνει πλήρης εκμάθηση των μοτίβων. Τα διαγράμματα που προέκυψαν φαίνονται στον πίνακα 28. Αυτή την φορά ο πίνακας έχει στην πρώτη στήλη την απόδοση, το σφάλμα και στη συνέχεια έχουμε ROC καμπύλη, για κάθε μία από τις κλάσεις, για φυσιολογικά, ισχαιμικά και έμφραγμα.

Accuracy	Loss	ROC		
		Φυσιολογικά	Ισχαιμικά	Έμφραγμα

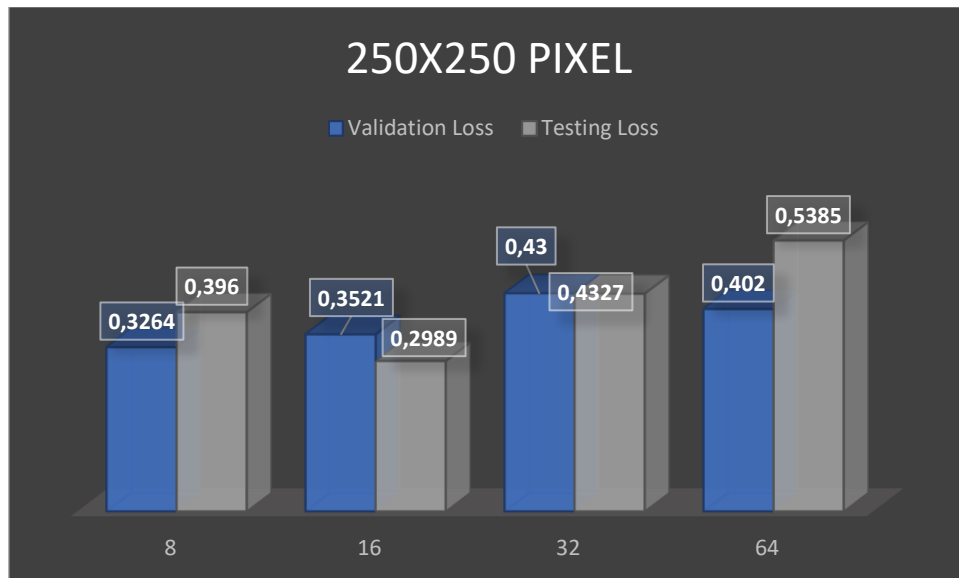
Πίνακας 28

Η καλύτερη αρχιτεκτονική, που επιλέχθηκε είναι Pixels:250x250, batch size:16, conv layers:16-32-64-128, και dense layers 128-128.

Για να καταλήξουμε όμως σε αυτή την αρχιτεκτονική έγινε έρευνα με διάφορες τιμές στις παραμέτρους, για να δούμε τι αποτελέσματα έχουν με διαφορετικές τιμές. Μιας και το καλύτερο batch size είναι το 16, βλέπουμε την validation accuracy, testing accuracy και auc τιμές έχουν για batch size 8, 32 και 64 στην εικόνα 58 και στην εικόνα 59 βλέπουμε το validation loss και testing loss. Παρατηρούμε ότι το batch size 16 φέρνει καλύτερα αποτελέσματα και στα δύο γραφήματα.

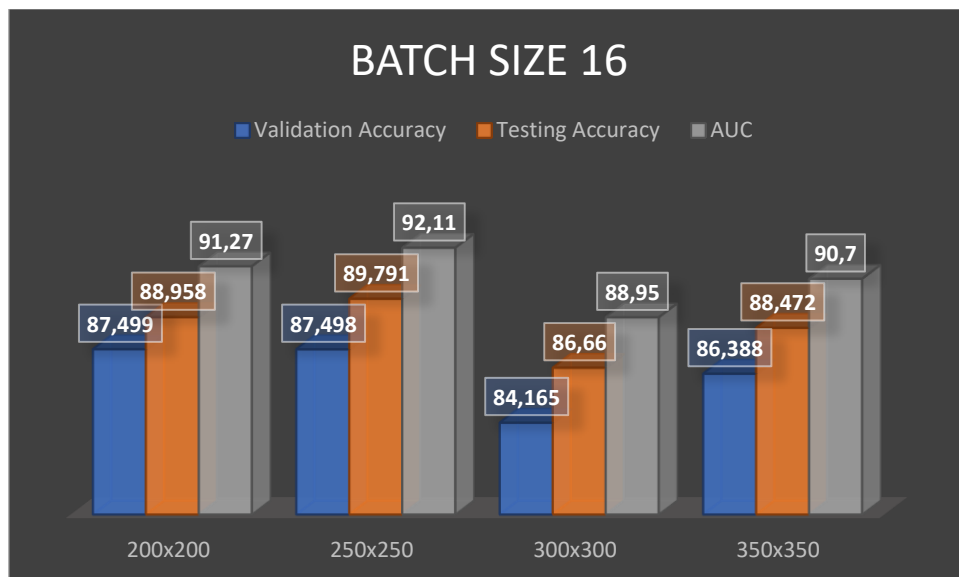


Εικόνα 58

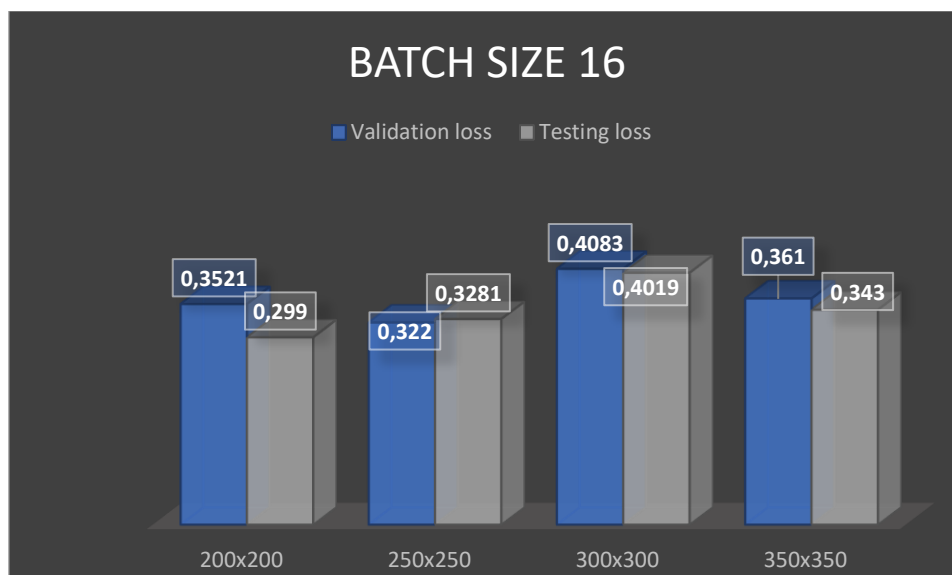


Εικόνα 59

Αφού επιβεβαιώθηκε ότι το batch size 16 είναι το κατάλληλο για αυτό το σετ δεδομένων εξετάστηκαν διάφοροι συνδυασμοί πίξελ για να δούμε ποιο εφαρμόζει καλύτερα. Είναι φανερό ότι ανάμεσα στο συνδυασμό 200x200, 250x250 και 300x300, το 250x250 είναι το καταλληλότερο, όπως φαίνεται στην εικόνα 60 που απεικονίζει την απόδοση και στην εικόνα 60 που απεικονίζει το σφάλμα 61 .

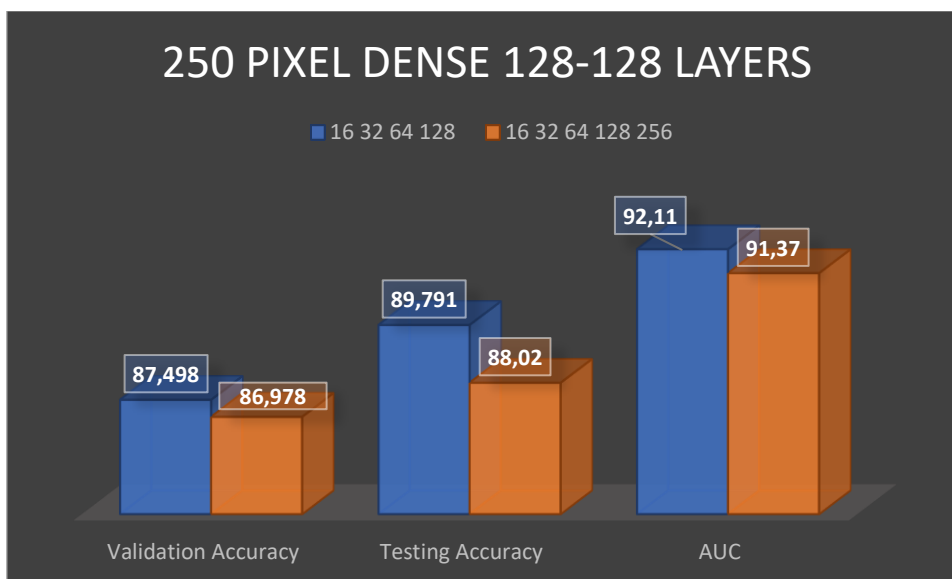


Εικόνα 60

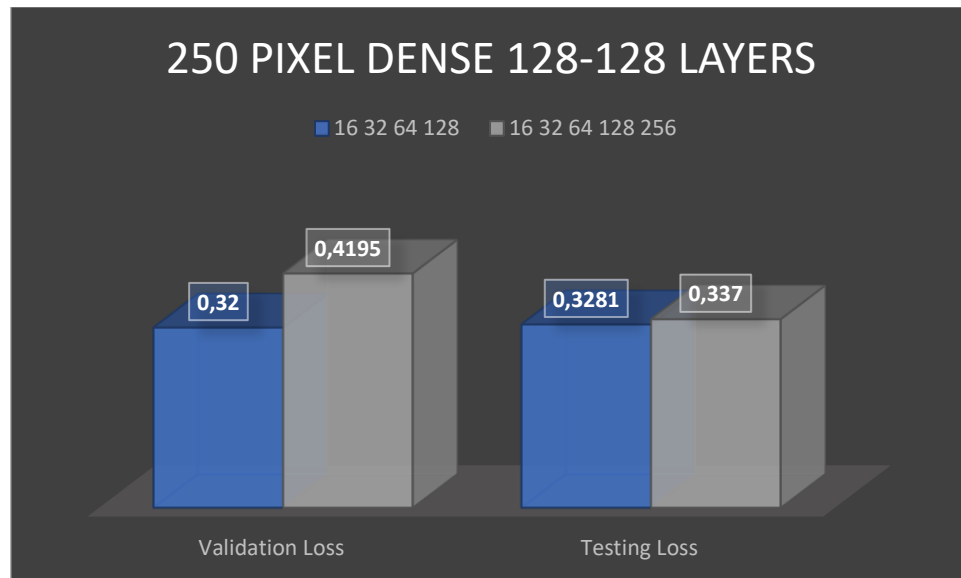


Εικόνα 61

Αφού αποφασίσαμε ότι το καλύτερο batch size είναι 16 για αυτό το σεντ δεδομένων, συνεχίζουμε με την αναζήτηση για τα συνελκτικα δίκτυα στις εικόνες 62 και 63 για την απόδοση και το σφάλμα αντίστοιχα. Η σύγκριση γίνεται για 16-32-64-128 και 16-32-64-128-256. Όπως φαίνεται και στα διαγράμματα, που απεικονίζουν τις μετρικές validation accuracy, testing accuracy και auc, τα καλύτερα αποτελέσματα έχουμε με τον συνδυασμό 16-32-64-128.

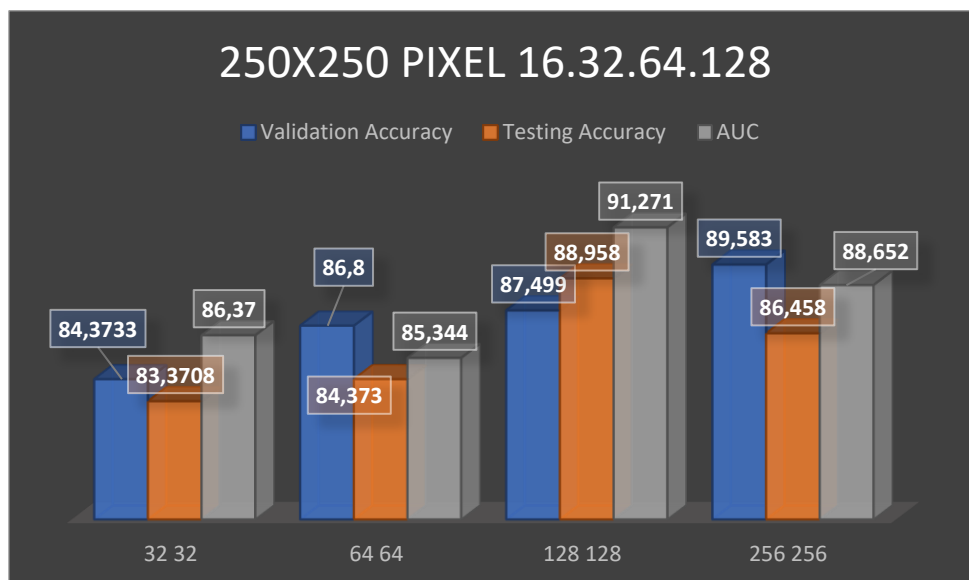


Εικόνα 62

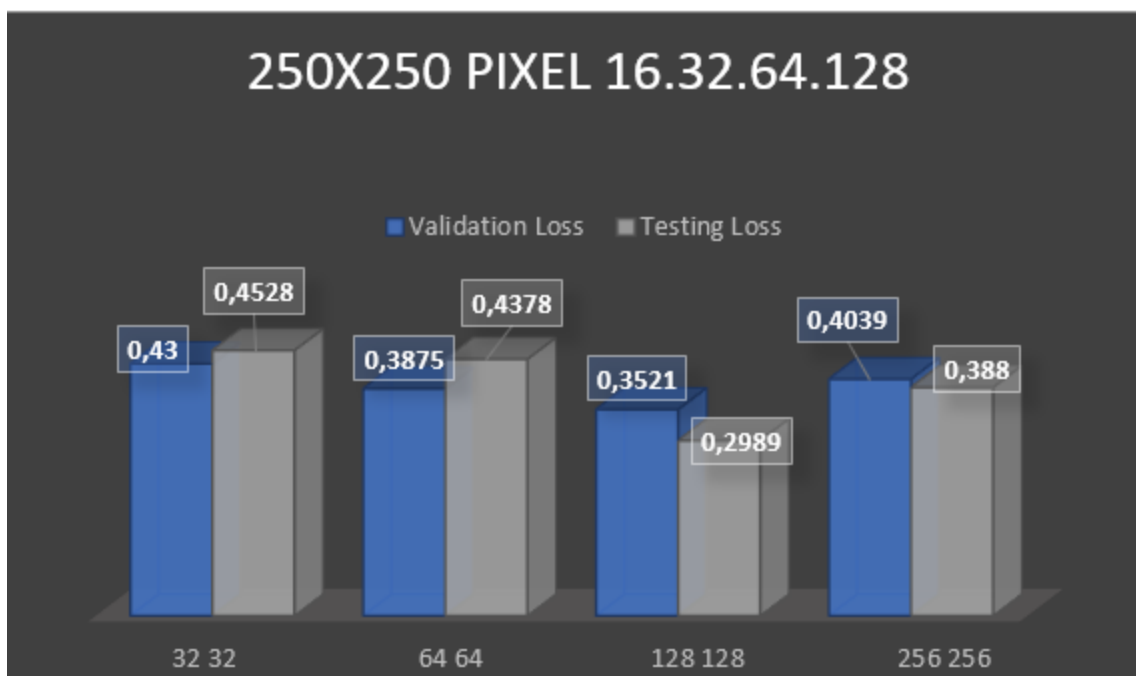


Εικόνα 63

Επόμενο και τελευταίο βήμα, για αυτό το μοντέλο είναι η αναζήτηση καλύτερου συνδυασμού για τα πλήρως συνδεδεμένα δίκτυα. Οι συνδυασμοί που εξετάστηκαν είναι 32-32, 64-64, 128-128, 256-256. Παρατηρούμε τα αποτελέσματα που παρήχθησαν στις εικόνες 64 και 65 για την απόδοση και το σφάλμα αντίστοιχα και για αυτό καταλήξαμε στον συνδυασμό 128-128.

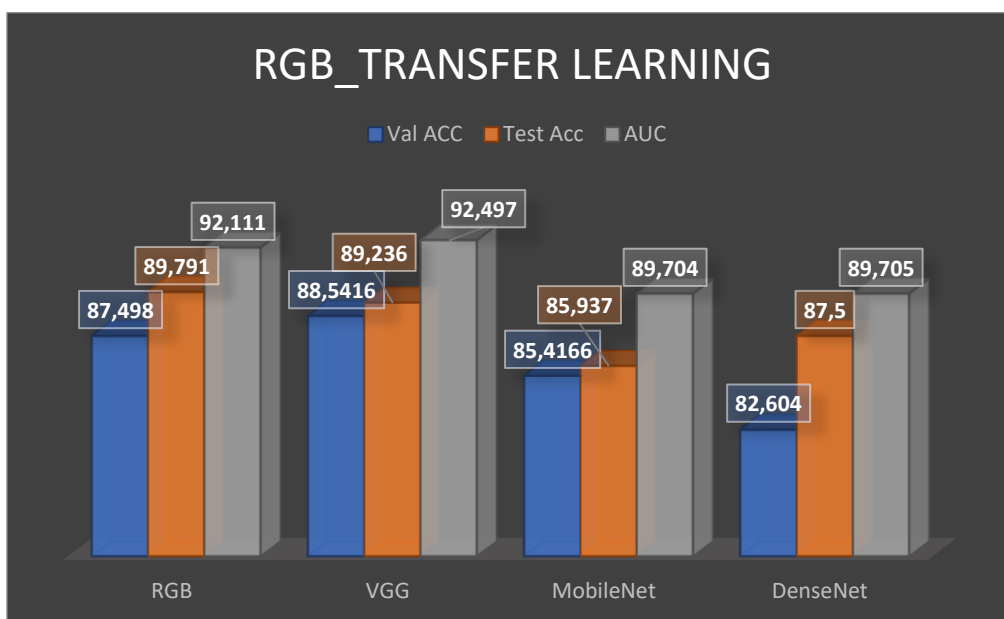


Εικόνα 64

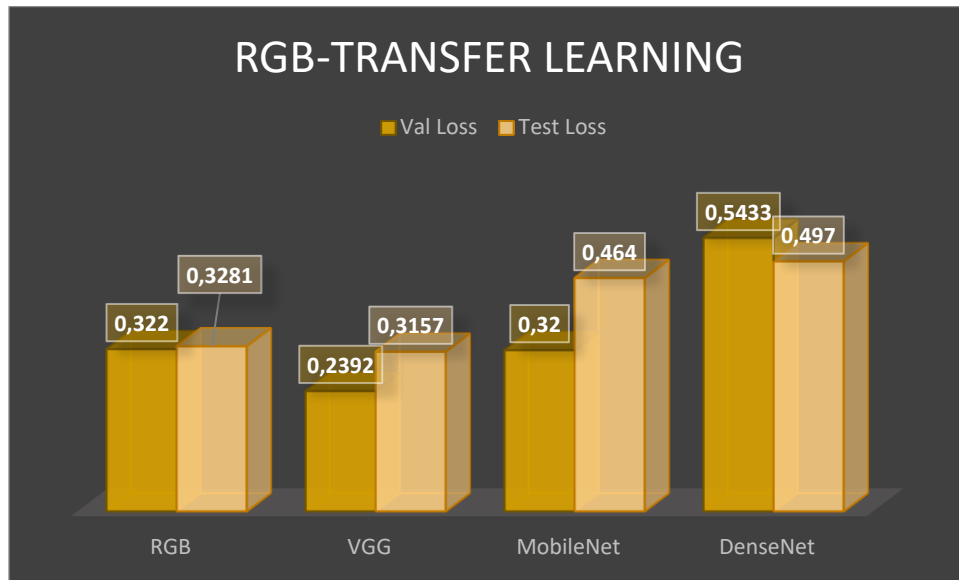


Εικόνα 65

Και τέλος δοκιμάστηκε η μέθοδος μεταφορά εκμάθησης (Transfer learning) με τα δίκτυα VGG-MobileNet, Densenet για να γίνει η σύγκριση με το RGB. Παρατηρούμε στις εικόνες 66 και 67 ότι το RGB και το VGG έχουν πολύ κοντινές τιμές, με το RGB να τα έχει πάει συνολικά καλύτερα.



Εικόνα 66



Εικόνα 67

5.5 Αποτελέσματα πολικών χαρτών

Σετ ε) Φυσιολογικά PM(Polar Maps) – Παθολογικά PM (Polar Maps)

Η ίδια μεθοδολογία που ακολουθήθηκε στα SPECT δεδομένα, θα ακολουθηθεί και στους πολικούς χάρτες.

Πρώτα παραθέτονται οι αρχιτεκτονικές, για το σετ δεδομένων φυσιολογικά-παθολογικά όσον αφορά τους πολικούς χάρτες.

Βλέπουμε την αρχιτεκτονική, που ακολουθήθηκε για το RGB, χωρισμένα με στήλες τα pixels, batch size, dropout, convolutional layers και dense layers.

	Pixels	Batch size	Dropout	Epochs	Conv Layers	Dense Layers
RGB	80x80	16	0.2	200	2-4-8	4-4

Πίνακας 29

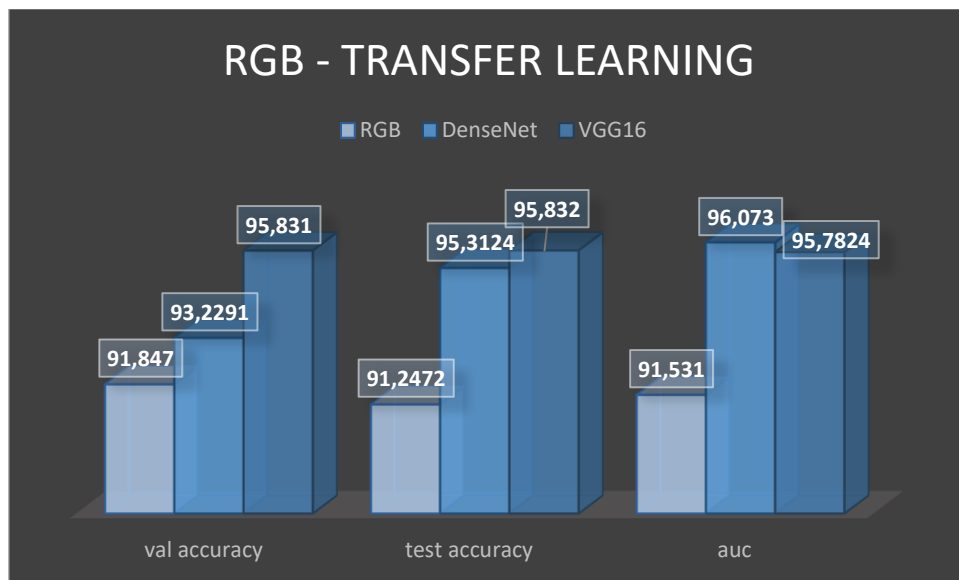
Για να καταλήξουμε ότι αυτές είναι οι καλύτερες τιμές για αυτό το σετ δεδομένων, ακολούθησε αρκετή έρευνα με διάφορες τιμές στις παραμέτρους. Είναι φανερό, ότι οι παράμετροι είναι ίδιοι, αλλά οι τιμές έχουν διαφορά, σε σχέση με τα SPECT δεδομένα, είναι δηλαδή πιο χαμηλά. Αυτό συμβαίνει όταν ένα σετ δεδομένων είναι εύκολο και με τις «φυσιολογικές» τιμές κάνει υπερ-εκπαίδευση (overfitting). Για αυτό και η τιμή για τα epochs, έχει την υπο-διπλάσια τιμή.

Αφού βρέθηκε η καλύτερη αρχιτεκτονική για το RGB, ακολούθησε ανάλυση της μεθόδου μεταφορά εκμάθησης. Μελετήθηκαν τα δίκτυα VGG-16 και DenseNet, όπως φαίνονται στον Πίνακα 28. Όσον αφορά για τις τιμές στα δίκτυα αυτά, είναι περίπου ίδια με τα δεδομένα SPECT, με εξαίρεση την τιμή των πλήρως συνδεδεμένων δικτύων, όπως φαίνεται στον πίνακα 30.

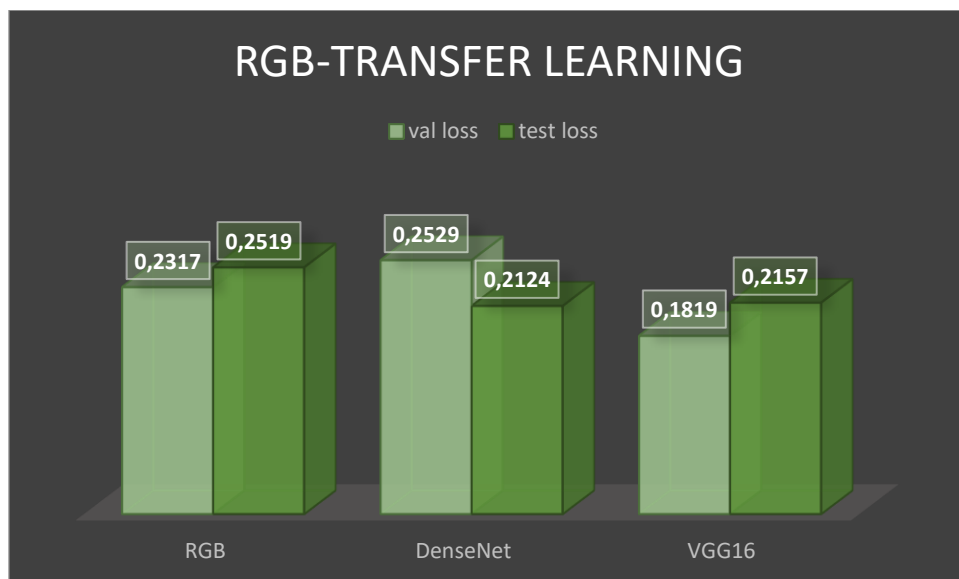
	Pixels	Batch size	Dropout	Epochs	Conv Layers	Dense Layers
VGG	200 200	16	0.2	200	True 13 Layers	Global Average Pooling 2D 16-16
DenseNet	200x200	8	0.2	200	True 119 Layers	Global Average Pooling 2D 4

Πίνακας 30

Αφού παρουσιάστηκαν οι αρχιτεκτονικές του RGB και του transfer learning ακολουθούν τα συγκριτικά αποτελέσματα μεταξύ τους στις εικόνες 68 και 69 . Παρατηρούμε ότι το VGG έχει τα καλύτερα αποτελέσματα, όσον αφορά την απόδοση και το σφάλμα.



Εικόνα 68



Εικόνα 69

5.6 Μέθοδος cross-validation

Επίσης μια μέθοδος που χρησιμοποιείται στα μοντέλα μηχανικής μάθησης είναι η cross-validation. Αποτελεί μια στατιστική μέθοδο που χρησιμοποιείται για την εκτίμηση της ικανότητας των μοντέλων μηχανικής μάθησης.

Αλγοριθμικά αναφέρεται ως k-Fold Cross-Validation, καθώς η διαδικασία έχει μια μόνο παράμετρο που ονομάζεται k που αναφέρεται στον αριθμό συστάδων στις οποίες πρέπει να διαιρεθεί ένα δεδομένο δείγμα δεδομένων. Όταν επιλέγεται μια συγκεκριμένη τιμή για το k, για παράδειγμα $k = 10$ γίνεται 10 Fold Cross-Validation. Οπότε το σετ δεδομένων χωρίζεται σε 10 ομάδες.

Η μέθοδος αυτή χρησιμοποιείται κατά κύριο λόγο στην εφαρμοσμένη μηχανική μάθηση για την εκτίμηση της ικανότητας ενός μοντέλου μηχανικής μάθησης σε δεδομένα που δεν έχει εκπαιδευτεί. Δηλαδή, γίνεται εφαρμογή της σε μοντέλο με περιορισμένο δείγμα δεδομένων για να εκτιμηθεί η απόδοση του μοντέλου με δεδομένα που δεν έχουν χρησιμοποιηθεί κατά την εκπαίδευση του μοντέλου.

Η γενική διαδικασία έχει ως εξής:

1. Τυχαία κατανομή δεδομένων.
2. Διαχώριση του σύνολο δεδομένων σε ομάδες k
3. Για κάθε μοναδική ομάδα:
 - 3.1 Ένα μέρος αποτελεί το σύνολο δεδομένων ελέγχου
 - 3.2 Το υπόλοιπο ως σύνολο δεδομένων εκπαίδευσης
 - 3.3 Τοποθέτηση ενός από τα k μοντέλα στο σετ προπόνησης και αξιολόγηση στο σετ δοκιμών
 - 3.4 Διατήρηση βαθμολογίας αξιολόγησης και απόρριψη μοντέλου
4. Σύνοψη αξιολόγησης του μοντέλου

Είναι σημαντικό ότι κάθε ένα δείγμα από το σετ δεδομένων, που ανατίθεται σε μια μεμονωμένη ομάδα, παραμένει σε αυτήν την ομάδα για τη διάρκεια της διαδικασίας. Αυτό σημαίνει ότι κάθε

δείγμα έχει την ευκαιρία να χρησιμοποιηθεί στο σετ ελέγχου 1 φορά και να χρησιμοποιηθεί για να εκπαιδεύσει το μοντέλο $k-1$ φορές.

Είναι σημαντικό να επιλεγεί προσεκτικά η τιμή k , σύμφωνα με το δείγμα δεδομένων, καθώς μια κακώς επιλεγμένη τιμή για το k μπορεί να οδηγήσει σε λανθασμένη εκτίμηση της ικανότητας του μοντέλου. Για αυτό το λόγο ακολουθούνται οι τρεις παρακάτω τακτικές για την επιλογή του k :

- Η τιμή για το k επιλέγεται έτσι ώστε κάθε ομάδα δειγμάτων από τα δεδομένα να είναι αρκετά μεγάλη ώστε να είναι στατιστικά αντιπροσωπευτική του ευρύτερου συνόλου δεδομένων.
- $k = 10$: Η τιμή για το k είναι σταθερή στο 10, μια τιμή που έχει βρεθεί μέσω πειραματισμού που οδηγεί γενικά σε μια εκτίμηση δεξιοτήτων μοντέλου με χαμηλή μεροληψία μια μέτρια διακύμανση.
- $k = n$: Η τιμή για το k είναι σταθερή στο n , όπου n είναι το μέγεθος του συνόλου δεδομένων για να δώσει σε κάθε δοκιμαστικό δείγμα την ευκαιρία να χρησιμοποιηθεί στο συγκρατημένο σύνολο δεδομένων. Αυτή η προσέγγιση ονομάζεται διασταυρούμενη επικύρωση

Ακολουθεί και παράδειγμα για να γίνει περισσότερο κατανοητή η εφαρμογή της.

Έχουμε ένα μοντέλο με 3 fold cross validation, οπότε δημιουργούνται 3 μοντέλα.

- **Model 1:** Εκπαίδευση στο Fold1 + Fold2, Έλεγχος στο Fold3
- **Model 2:** Εκπαίδευση στο Fold2 + Fold3, Έλεγχος στο Fold1
- **Model 3:** Εκπαίδευση στο Fold1 + Fold3, Έλεγχος στο Fold2

Τα μοντέλα στη συνέχεια απορρίπτονται αφού αξιολογηθούν καθώς έχουν εξυπηρετήσει τον σκοπό τους.

Οι τιμές της απόδοσης και του σφάλματος συλλέγονται για κάθε μοντέλο και συνοψίζονται για τον υπολογισμό του μέσου όρου [126].

Η μέθοδος k-fold cross validation δοκιμάστηκε και στο δικό μας σετ δεδομένων στο

Σετ α) Φυσιολογικά – Ισχαιμία

και προέκυψαν τα παρακάτω αποτελέσματα.

Για 5 folds, έχουμε τα αποτελέσματα στον πίνακα 31, με αρκετά καλά αποτελέσματα.

Folds = 5	<i>Accuracy</i>	<i>Loss</i>
<i>Fold 1</i>	93,5064	1,2704
<i>Fold 2</i>	96,1038	0,7948
<i>Fold 3</i>	93,5064	0,6371
<i>Fold 4</i>	94,8051	1,319
<i>Fold 5</i>	96,1038	0,4851
<i>Average for all folds</i>	94,8051 (+- 1,1615)	0,90133

Πίνακας 31

Δοκιμάστηκε και για k =10 και φαίνονται τα αποτελέσματα στον πίνακα 32, όπου έχουμε καλύτερη τιμή στο σφάλμα σε σύγκριση με k=5, αλλά χαμηλότερη απόδοση.

Folds = 10	<i>Accuracy</i>	<i>Loss</i>
<i>Fold 1</i>	92,2077	0,3028
<i>Fold 2</i>	93,5049	0,3717
<i>Fold 3</i>	94,8051	0,2431
<i>Fold 4</i>	94,8051	0,291
<i>Fold 5</i>	92,2077	0,2728
<i>Fold 6</i>	93,5064	0,2741
<i>Fold 7</i>	88,3116	0,3724
<i>Fold 8</i>	94,8051	0,2676
<i>Fold 9</i>	94,8051	0,4555
<i>Fold 10</i>	90,9090	0,4555
<i>Average for all folds</i>	92,9807 (+-2,2086)	0,3177

Πίνακας 32

Και τέλος για $k=15$ στον πίνακα 33, το σφάλμα έχει κακή τιμή, καθώς η τιμή για το k δεν είναι σωστή και αντιπροσωπευτική και είναι λογικό που δεν έχουν εξαχθεί καλά αποτελέσματα. Η απόδοση παρόλα αυτά έχει ικανοποιητική τιμή.

Folds = 15	<i>Accuracy</i>	<i>Loss</i>
<i>Fold 1</i>	93,5064	0,968
<i>Fold 2</i>	92,2077	0,8846
<i>Fold 3</i>	93,5064	1,3387
<i>Fold 4</i>	89,6103	0,7989
<i>Fold 5</i>	90,9090	0,6734
<i>Fold 6</i>	89,6103	1,0969
<i>Fold 7</i>	92,2077	0,8418
<i>Fold 8</i>	90,9090	1,3611
<i>Fold 9</i>	92,2077	1,7715
<i>Fold 10</i>	90,9090	1,76383
<i>Fold 11</i>	90,9090	0,463
<i>Fold 12</i>	93,5064	1,4318
<i>Fold 13</i>	89,6103	0,6925
<i>Fold 14</i>	92,2077	1,2409
<i>Fold 15</i>	88,3116	1,0730
Average for all folds	91,3419 (+-1,5478)	1,0933

Πίνακας 33

Είναι φανερό ότι συνδυαστικά με το μέσο όρο τιμών της απόδοσης και του σφάλματος, η καλύτερη τιμή για το k είναι το 10.

5.7 Συγκριτικά τρεξίματα

Θεωρήθηκε σωστό, να εξετάσουμε το μοντέλο συνελκτικού δικτύου που υλοποιήθηκε και εφαρμόστηκε σε αυτήν την εργασία (RGB), και σε διαφορετικό σετ δεδομένων, με ίδιου τύπου δεδομένα. Η εργασία [98] των συγγραφέων Selcan Kaplan Berkaya, Ilknur Ak Sivriköz, Serkan Gunal, παρουσιάζει τον σχεδιασμό ενός μοντέλου με παρόμοιο σετ δεδομένων.

Χρησιμοποιήθηκε η μέθοδος early stopping, για να γίνει αποφυγή του overfitting, από τους ερευνητές.

Παρακάτω στην εικόνα 70 έχουμε τα αποτελέσματα της αναφερόμενης έρευνας, τα οποία απεικονίζουν τα δίκτυα που αναπτύχθηκαν σύμφωνα με μέθοδο της μεταφοράς εκμάθησης (transfer learning). Για κάθε ξεχωριστό δίκτυο φαίνεται η απόδοση, η μετρική απόδοσης που ονομάζεται sensitivity και specificity, καθώς και ο χρόνος εκτέλεσης.

Models		Acc	Sen	Spe	Processing Time (sec)
DL-based	AlexNet	0.72/0.57/0.70	1.00 /0.96/0.96	0.43/0.17/0.43	2.09/0.18/0.09
	DenseNet-201	0.63/0.61/0.50	0.96/0.92/ 1.00	0.29/0.29/0.00	45.91/1.86/0.50
	GoogLeNet	0.72/0.59/0.65	1.00 /0.88/ 1.00	0.43/0.29/0.29	31.81/0.32/0.16
	NASNet-Large	0.57/0.61/0.50	1.00 /0.92/ 1.00	0.14/0.29/0.00	242.41/5.66/1.42
	ResNet-18	0.77/0.57/0.50	0.96/0.96/ 1.00	0.57/0.17/0.00	5.56/0.31/0.16
	VGG-16	0.86/0.68/0.92	1.00 /0.92/0.84	0.71/0.43/ 1.00	85.44/1.13/0.43
	VGG-19	0.84/0.79/ 0.94	0.96/ 1.00 /0.88	0.71/0.57/ 1.00	64.78/1.30/0.42

Εικόνα 70

Αφού ανακτήθηκαν τα δεδομένα από το σετ που παρέχονταν στο Kaggle, εφαρμόσαμε την προεπεξεργασία πρώτα με την μάσκα, με την μέθοδο inpaint(), για να αφαιρέσουμε τον θόρυβο και στη συνέχεια κάναμε μια μελέτη με τα δίκτυα RGB και VGG-16, για να εξάγουμε τα δικά μας αποτελέσματα.

Η αρχιτεκτονική του RGB φαίνεται στον πίνακα 34, και η αρχιτεκτονική του VGG-16 φαίνεται στον πίνακα 35 .

	Pixels	Batch size	Conv Layers	Dense Layers
--	--------	------------	-------------	--------------

RGB	150x150	8	16 32 64 128	128 128
-----	---------	---	--------------	---------

Πίνακας 34

	Pixels	Batch size	True Layers	Dense Layers
VGG-16	200x200	32	True 9	Global Average Pooling 2D 256-256

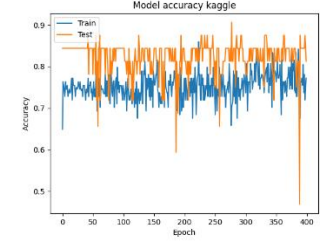
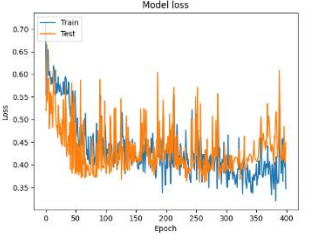
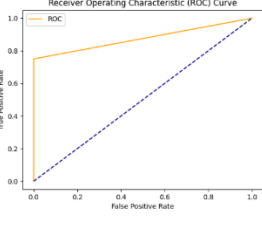
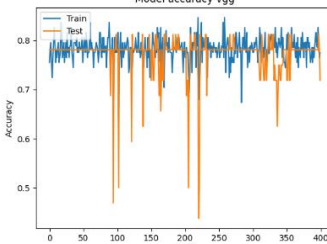
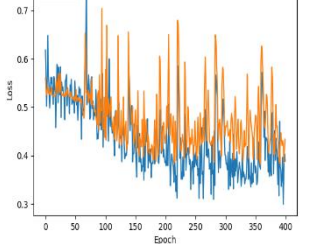
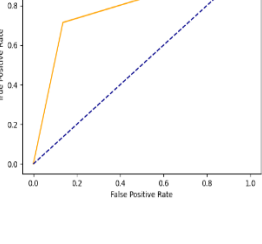
Πίνακας 35

Τα αποτελέσματα, που εξήγαγε το κάθε δίκτυο φαίνεται στον πίνακα 36. Παρατηρούμε ότι το RGB έχει καλύτερα αποτελέσματα.

	Validation Accuracy	Validation Loss	Testing Accuracy	Testing Loss	AUC
RGB	84,09	0,3208	88,792	0,33632	0,866
VGG-16	82,724	0,32811	85,1553	0,3734	0,79512

Πίνακας 36

Στον πίνακα 37 φαίνονται τα διαγράμματα απόδοσης, σφάλματος και AUC για το RGB και για το VGG.

	Απόδοση	Σφάλμα	AUC
RGB			
VGG			

Πίνακας 37

Είναι πολύ σημαντικό προσόν για το δίκτυο, που υλοποιήσαμε να έχει τόσο καλά αποτελέσματα σε ένα σετ δεδομένων, στο οποίο δεν στηρίχτηκε η έρευνα μας. Αυτό σημαίνει ότι έχουμε “generalization” και ότι θα είναι πολύ εύκολη η κατηγοριοποίηση σε μελλοντικά σετ δεδομένων, το οποίο είναι το επιθυμητό για τέτοιου είδους αλγόριθμους, και πόσο μάλλον σε ιατρικά προβλήματα κατηγοριοποίησης.

5.8 Transfer Learning-RGB

Όπως εφαρμόζεται η μέθοδος που ονομάζεται μεταφορά γνώσης για τα προ-εκπαιδευμένα δίκτυα VGG-16, DenseNet-121, MobileNet, με την ίδια συμπεριφορά μπορεί να υλοποιηθεί η ίδια μέθοδος για τα απλά συνελκτικά μοντέλα. Συγκεκριμένα στην παρούσα έρευνα, αφού το δίκτυο ολοκλήρωσε τη διαδικασία της εκπαίδευσης και της αξιολόγησης είναι δυνατή η αποθήκευση των βαρών του μοντέλου, καθώς και το ίδιο το μοντέλο σε αρχεία στον υπολογιστή. Αυτά τα αρχεία μπορούμε στη συνέχεια να τα φορτώσουμε σε Python script και να γίνει κατηγοριοποίηση των δεδομένων χρησιμοποιώντας για μοντέλο, ότι έχει αποθηκευτεί στα αρχεία του υπολογιστή, από ήδη υπάρχουσα εκπαίδευση. Οι μετρικές και των 2 περιπτώσεων έχουν το ίδιο αποτέλεσμα, όπως

φαίνεται στον πίνακα 38. Τρέξαμε για το βήμα αυτό 3 φορές τον αλγόριθμο για να επικυρώσουμε τα αποτελέσματα του.

	Validation Accuracy	Validation Loss	Testing Accuracy	Testing Loss
1.CNN	81,818	0,4011	92,2077	0,2144
1. LoadedCNN	81,818	0,4011	92,2077	0,2144
2. CNN	89,7727	0,23453	85,714	0,2977
2. LoadedCNN	89,7727	0,23453	85,714	0,2977
3. CNN	93,1818	0,2092	87,0129	0,26714
3. LoadedCNN	93,1818	0,2092	87,0129	0,26714

Πίνακας 38

Παρατηρούμε ότι έχουμε τα ίδια αποτελέσματα και στις 3 περιπτώσεις. Ο κώδικας που αναπτύχθηκε είναι ο εξής:

```

# serialize model to JSON
model_json = model.to_json()
with open("model.json", "w") as json_file:
    json_file.write(model_json)
# serialize weights to HDF5
model.save_weights("model.h5")
print("Saved model to disk")

# load json and create model
json_file = open('model.json', 'r')
loaded_model_json = json_file.read()
json_file.close()
loaded_model = model_from_json(loaded_model_json)
# load weights into new model
loaded_model.load_weights("model.h5")
print("Loaded model from disk")

# evaluate loaded model on test data
loaded_model.compile(loss='binary_crossentropy', optimizer='adam', metrics=['accuracy'])
# score = loaded_model.evaluate(X_test, y_test, verbose=0)
[127]

```

6, Σύνοψη-Συμπεράσματα

Το συμπέρασμα της παραπάνω έρευνας είναι ότι πλέον τα συστήματα υποβοήθησης λήψης αποφάσεων στην ιατρική με μεθοδολογίες τεχνητής νοημοσύνης αποτελούν ένα σημαντικό και ίσως αναπόσπαστο εργαλείο στα χέρια των ιατρών, καθώς ο αριθμός των ιατρικών διαγνώσεων, που απαιτείται καθημερινά αυξάνεται ραγδαία. Οι προηγμένες μεθοδολογίες της μηχανικής μάθησης και οι νέες τεχνικές βαθιάς μάθησης είναι πλέον ικανές να ενσωματωθούν σε συστήματα επεξεργασίας ιατρικής εικόνας για διάγνωση και λήψη αποφάσεων. Για να έχουν σημαντικά και αξιόπιστα αποτελέσματα και να μπορούν να είναι ένα χρήσιμο εργαλείο στα χέρια των γιατρών, θα πρέπει τα δεδομένα που θα παρέχονται να είναι αρκετά για να εκπαιδεύσουν κατάλληλα τα μοντέλα διάγνωσης.

Σύμφωνα με την μελέτη, που έγινε στα πλαίσια αυτής της πτυχιακής, είναι κατανοητό ότι με την κατάλληλη διαχείριση των μεθόδων της βαθιάς μάθησης και την εντατική έρευνα, τα συνελκτικά νευρωνικά δίκτυα μπορούν να διαχειρίζονται τα δεδομένα και να λειτουργούν ως βοήθημα στη διαδικασία διάγνωσης στα ιατρικά προβλήματα. Είναι σε θέση να εξάγουν αξιόπιστα αποτελέσματα και να διαχειρίζονται μεγάλο όγκο δεδομένων.

Πρόσφατα, οι μεθοδολογίες της βαθιάς μάθησης έχουν προχωρήσει αρκετά και αυτό φαίνεται από το γεγονός ότι υπάρχει η μέθοδος μεταφοράς εκμάθησης (transfer learning). Έγινε φανερό, ότι η μέθοδος αυτή είναι σημαντικό βήμα, καθώς μπορούμε να υλοποιούμε αρκετά βαθιά νευρωνικά δίκτυα, τα οποία έχουν εκπαιδευτεί σε μεγάλο όγκο δεδομένων. Παρόλο, που χρειάζονται αρκετό χρόνο να εκπαιδευτούν παρέχουν κατηγοριοποίηση σε αρκετά προβλήματα και είναι τεράστιο προσόν.

Επίσης, μια σημαντική μεθοδολογία, που περιλαμβάνεται στις μεθόδους της τεχνητής νοημοσύνης είναι το cross validation, το οποίο μας βοηθάει να εξακριβώσουμε εάν ένα μοντέλο που έχει υλοποιηθεί είναι έγκυρο, σε δεδομένα, που είναι άγνωστα σε σχέση με το σετ εκπαίδευσης. Με την εφαρμογή του, είναι εύκολο να διαπιστωθεί αν το δίκτυο θα μπορεί να λειτουργεί ικανοποιητικά, όσο και «γενικό να είναι το νέο σετ δεδομένων και κυρίως ότι δεν θα γίνει υπερ-εκπαίδευση. [126]

Η στεφανιαία νόσος, που μελετήθηκε στην παρούσα έρευνα είναι ένα σημαντικό πρόβλημα στην ανθρώπινη υγεία. Το γεγονός ότι η διάγνωση της γίνεται μέσα από σπινθηρογραφικές εικόνες έδωσε το έναυσμα στην πληροφορική να παρέχει το λειτουργικό της και τους αλγόριθμους, με σκοπό να συντελέσει στο πιο ακριβές και ίσως πιο γρήγορο αποτέλεσμα για μεγάλα σετ δεδομένων. Είναι αξιοθαύμαστο, ότι στην έρευνα αυτή, όπως και σε τόσες άλλες, παρόμοιες με αυτές που αναφέρθηκαν, ένα απλό μοντέλο συνελκτικού δικτύου με τέσσερα μόλις συνελκτικά επίπεδα και με σχετικά μικρό σετ δεδομένων να φέρει αποτελέσματα ~90%, καλύτερα και από τη μέθοδο του transfer learning.

Είναι γνωστό, ότι τα νευρωνικά δίκτυα χρησιμοποιούνται όλο και περισσότερο για τη δημιουργία προγραμμάτων που μπορούν να προβλέψουν και να ταξινομήσουν διαφορετικά σετ δεδομένων. Και για αυτό το λόγο έχουν υλοποιηθεί νέα και χρήσιμα μοντέλα με αξιόπιστα αποτελέσματα,

όπως είδαμε και στη δική μας έρευνα. Ωστόσο, δεν γνωρίζουμε πραγματικά πώς λειτουργούν, για αυτό αρκετοί παρομοιάζουν τη λειτουργία τους σαν μαύρο κουτί, δηλαδή μια αναπαράσταση που αντιπροσωπεύει την άγνωστη εσωτερική μηχανική τους. Για αυτό το λόγο, διερευνώνται νέα μοντέλα, που ονομάζονται επεξηγήσιμα μοντέλα. [128] Είναι ιδιαίτερα σημαντικό να γίνει έρευνα στον τομέα μηχανισμού τους, καθώς η επεξήγηση είναι ένα από τα πιο έντονα θέματα όταν πρόκειται για την εφαρμογή της τεχνητής νοημοσύνης στον ιατρικό τομέα, μιας και το αποτέλεσμα αποτελεί τόσο κρίσιμο ζήτημα. Παρόλο που τα συστήματα που βασίζονται στην τεχνητή νοημοσύνη έχουν αποδειχθεί ότι υπερσχύουν των ανθρώπων σε ορισμένες αναλυτικές εργασίες, η έλλειψη επεξηγήσεων εξακολουθεί να προκαλεί κριτική, στο αν οι αλγόριθμοι της τεχνητής νοημοσύνης είναι άξιοι σύγκρισης [129].

7. Αναφορές

- [1] H. Brody, “Medical imaging,” *Nature*, vol. 502, no. 7473, p. S81, Oct. 2013, doi: 10.1038/502s81a.
- [2] “Artificial Intelligence in Healthcare | Stanford Online.” <https://online.stanford.edu/programs/artificial-intelligence-healthcare> (accessed Jul. 22, 2021).
- [3] J. Schmidhuber, “Deep Learning in Neural Networks: An Overview,” *Neural Netw.*, vol. 61, pp. 85–117, Jan. 2015, doi: 10.1016/j.neunet.2014.09.003.
- [4] Y. Bengio, “Learning Deep Architectures for AI,” *Foundations*, vol. 2, pp. 1–55, Jan. 2009, doi: 10.1561/22000000006.
- [5] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, Art. no. 7553, May 2015, doi: 10.1038/nature14539.
- [6] G. E. Hinton, “Reducing the Dimensionality of Data with Neural Networks,” *Science*, vol. 313, no. 5786, pp. 504–507, Jul. 2006, doi: 10.1126/science.1127647.

- [7] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, and P.-A. Manzagol, “Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion,” p. 38.
- [8] “Rectified linear units improve restricted boltzmann machines | Proceedings of the 27th International Conference on International Conference on Machine Learning.” <https://dl.acm.org/doi/10.5555/3104322.3104425> (accessed Mar. 27, 2021).
- [9] “unified architecture for natural language processing: Deep neural networks with multitask learning. Proceedings of International Conference on Machine Learning - Αναζήτηση Google.” <https://www.google.com/search?client=firefox-b-d&q=unified+architecture+for+natural+language+processing%3A+Deep+neural+networks+with+multitask+learning.+Proceedings+of+International+Conference+on+Machine+Learning+> (accessed Mar. 27, 2021).
- [10] I. Sutskever, J. Martens, and G. Hinton, “Generating Text with Recurrent Neural Networks,” p. 8.
- [11] W. Zhang *et al.*, “Deep convolutional neural networks for multi-modality iso-intense infant brain image segmentation,” *NeuroImage*, vol. 108, pp. 214–224, Mar. 2015, doi: 10.1016/j.neuroimage.2014.12.061.
- [12] J. Kleesiek *et al.*, “Deep MRI brain extraction: A 3D convolutional neural network for skull stripping,” *NeuroImage*, vol. 129, pp. 460–469, Apr. 2016, doi: 10.1016/j.neuroimage.2016.01.024.
- [13] G. Wu, M. Kim, Q. Wang, B. C. Munsell, and D. Shen, “Scalable High-Performance Image Registration Framework by Unsupervised Deep Feature Representations Learning,” *IEEE Trans. Biomed. Eng.*, vol. 63, no. 7, pp. 1505–1516, Jul. 2016, doi: 10.1109/TBME.2015.2496253.
- [14] H.-I. Suk, S.-W. Lee, and D. Shen, “Hierarchical Feature Representation and Multimodal Fusion with Deep Learning for AD/MCI Diagnosis,” *NeuroImage*, vol. 101, pp. 569–582, Nov. 2014, doi: 10.1016/j.neuroimage.2014.06.077.
- [15] H.-C. Shin, K. Roberts, L. Lu, D. Demner-Fushman, J. Yao, and R. M. Summers, “Learning to Read Chest X-Rays: Recurrent Neural Cascade Model for Automated Image Annotation,” *ArXiv160308486 Cs*, Mar. 2016, Accessed: Mar. 27, 2021. [Online]. Available: <http://arxiv.org/abs/1603.08486>
- [16] O. Russakovsky *et al.*, “ImageNet Large Scale Visual Recognition Challenge,” *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015, doi: 10.1007/s11263-015-0816-y.
- [17] H.-C. Shin *et al.*, “Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning,” *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1285–1298, May 2016, doi: 10.1109/TMI.2016.2528162.
- [18] J.-Z. Cheng *et al.*, “Computer-Aided Diagnosis with Deep Learning Architecture: Applications to Breast Lesions in US Images and Pulmonary Nodules in CT Scans,” *Sci. Rep.*, vol. 6, p. 24454, Apr. 2016, doi: 10.1038/srep24454.
- [19] H. R. Roth *et al.*, “Improving Computer-aided Detection using Convolutional Neural Networks and Random View Aggregation,” *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1170–1181, May 2016, doi: 10.1109/TMI.2015.2482920.
- [20] W. Shen, M. Zhou, F. Yang, C. Yang, and J. Tian, “Multi-scale Convolutional Neural Networks for Lung Nodule Classification,” in *Information Processing in Medical Imaging*, Cham, 2015, pp. 588–599. doi: 10.1007/978-3-319-19992-4_46.

- [21] A. A. A. Setio *et al.*, “Pulmonary Nodule Detection in CT Images: False Positive Reduction Using Multi-View Convolutional Networks,” *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1160–1169, May 2016, doi: 10.1109/TMI.2016.2536809.
- [22] F. Ciompi *et al.*, “Automatic classification of pulmonary peri-fissural nodules in computed tomography using an ensemble of 2D views and a convolutional neural network out-of-the-box,” *Med. Image Anal.*, vol. 26, no. 1, pp. 195–202, Dec. 2015, doi: 10.1016/j.media.2015.08.001.
- [23] A. Gupta, “Natural Image bases to represent Neuroimaging data,” p. 8.
- [24] T. Brosch and R. Tam, “Manifold Learning of Brain MRIs by Deep Learning,” in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013*, Berlin, Heidelberg, 2013, pp. 633–640. doi: 10.1007/978-3-642-40763-5_78.
- [25] Q. Dou *et al.*, “Automatic Detection of Cerebral Microbleeds From MR Images via 3D Convolutional Neural Networks,” *IEEE Trans. Med. Imaging*, vol. 35, no. 5, pp. 1182–1195, May 2016, doi: 10.1109/TMI.2016.2528129.
- [26] M. Khened, V. Kollerathu, and G. Krishnamurthi, “Fully Convolutional Multi-scale Residual DenseNets for Cardiac Segmentation and Automated Cardiac Diagnosis using Ensemble of Classifiers,” *Med. Image Anal.*, vol. 51, Jan. 2018, doi: 10.1016/j.media.2018.10.004.
- [27] “ImageNet,” *Wikipedia*. Aug. 19, 2021. Accessed: Aug. 26, 2021. [Online]. Available: <https://en.wikipedia.org/w/index.php?title=ImageNet&oldid=1039564413>
- [28] Y. Bengio, A. Courville, and P. Vincent, “Representation Learning: A Review and New Perspectives,” *ArXiv12065538 Cs*, Apr. 2014, Accessed: Mar. 28, 2021. [Online]. Available: <http://arxiv.org/abs/1206.5538>
- [29] A. Khan, A. Sohail, U. Zahoor, and A. S. Qureshi, “A Survey of the Recent Architectures of Deep Convolutional Neural Networks,” *Artif. Intell. Rev.*, vol. 53, no. 8, pp. 5455–5516, Dec. 2020, doi: 10.1007/s10462-020-09825-6.
- [30] D. Shen, G. Wu, and H.-I. Suk, “Deep Learning in Medical Image Analysis,” *Annu. Rev. Biomed. Eng.*, vol. 19, no. 1, pp. 221–248, Jun. 2017, doi: 10.1146/annurev-bioeng-071516-044442.
- [31] “Application of Deep Learning in Automated Analysis of Molecular Images in Cancer: A Survey.” <https://www.hindawi.com/journals/cmmi/2017/9512370/> (accessed Mar. 27, 2021).
- [32] K. O’Shea and R. Nash, “An Introduction to Convolutional Neural Networks,” *ArXiv151108458 Cs*, Dec. 2015, Accessed: Mar. 27, 2021. [Online]. Available: <http://arxiv.org/abs/1511.08458>
- [33] G. Litjens *et al.*, “A survey on deep learning in medical image analysis,” *Med. Image Anal.*, vol. 42, pp. 60–88, Dec. 2017, doi: 10.1016/j.media.2017.07.005.
- [34] M. Biswas *et al.*, “State-of-the-art review on deep learning in medical imaging,” *Front. Biosci. Landmark Ed.*, vol. 24, pp. 392–426, Jan. 2019.
- [35] D. Abdelhafiz, C. Yang, R. Ammar, and S. Nabavi, “Deep convolutional neural networks for mammography: advances, challenges and applications,” *BMC Bioinformatics*, vol. 20, no. 11, p. 281, Jun. 2019, doi: 10.1186/s12859-019-2823-4.
- [36] C. E. Kahn, “From Images to Actions: Opportunities for Artificial Intelligence in Radiology,” *Radiology*, vol. 285, no. 3, pp. 719–720, Dec. 2017, doi: 10.1148/radiol.2017171734.

- [37] K. J. Dreyer and J. R. Geis, "When Machines Think: Radiology's Next Frontier," *Radiology*, vol. 285, no. 3, pp. 713–718, Dec. 2017, doi: 10.1148/radiol.2017171183.
- [38] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural networks: an overview and application in radiology," *Insights Imaging*, vol. 9, no. 4, Art. no. 4, Aug. 2018, doi: 10.1007/s13244-018-0639-9.
- [39] H. Alaskar, A. Hussain, N. Al-Aseem, P. Liatsis, and D. Al-Jumeily, "Application of Convolutional Neural Networks for Automated Ulcer Detection in Wireless Capsule Endoscopy Images," *Sensors*, vol. 19, no. 6, Mar. 2019, doi: 10.3390/s19061265.
- [40] Z. Li, Y. Wang, J. Yu, Y. Guo, and W. Cao, "Deep Learning based Radiomics (DLR) and its usage in noninvasive IDH1 prediction for low grade glioma," *Sci. Rep.*, vol. 7, no. 1, Art. no. 1, Jul. 2017, doi: 10.1038/s41598-017-05848-2.
- [41] A. S. Lundervold and A. Lundervold, "An overview of deep learning in medical imaging focusing on MRI," *Z. Med. Phys.*, vol. 29, no. 2, pp. 102–127, May 2019, doi: 10.1016/j.zemedi.2018.11.002.
- [42] A. Elhendy, J. J. Bax, and D. Poldermans, "Dobutamine stress myocardial perfusion imaging in coronary artery disease," *J. Nucl. Med. Off. Publ. Soc. Nucl. Med.*, vol. 43, no. 12, pp. 1634–1646, Dec. 2002.
- [43] E. M. Prvulovich and J. B. Bomanji, "The role of nuclear medicine in clinical investigation," *BMJ*, vol. 316, no. 7138, pp. 1140–1146, Apr. 1998.
- [44] J. Rieber *et al.*, "Cardiac magnetic resonance perfusion imaging for the functional assessment of coronary artery disease: a comparison with coronary angiography and fractional flow reserve," *Eur. Heart J.*, vol. 27, no. 12, pp. 1465–1471, Jun. 2006, doi: 10.1093/eurheartj/ehl039.
- [45] N. M. Wilke, M. Jerosch-Herold, A. Zenovich, and A. E. Stillman, "Magnetic resonance first-pass myocardial perfusion imaging: clinical validation and future applications," *J. Magn. Reson. Imaging JMRI*, vol. 10, no. 5, pp. 676–685, Nov. 1999, doi: 10.1002/(sici)1522-2586(199911)10:5<676::aid-jmri10>3.0.co;2-l.
- [46] M. Koenig, E. Klotz, B. Luka, D. J. Venderink, J. F. Spittler, and L. Heuser, "Perfusion CT of the brain: diagnostic approach for early detection of ischemic stroke," *Radiology*, vol. 209, no. 1, pp. 85–93, Oct. 1998, doi: 10.1148/radiology.209.1.9769817.
- [47] A. Notghi and C. S. Low, "Myocardial perfusion scintigraphy: past, present and future," *Br. J. Radiol.*, vol. 84, no. Spec Iss 3, pp. S229–S236, Dec. 2011, doi: 10.1259/bjr/14625142.
- [48] "Nuclear Cardiology: Guidance on the Implementation of SPECT Myocardial Perfusion Imaging," Aug. 31, 2016. <https://www.iaea.org/publications/11076/nuclear-cardiology-guidance-on-the-implementation-of-spect-myocardial-perfusion-imaging> (accessed Mar. 27, 2021).
- [49] A. Khorsand, S. Graf, H. Sochor, E. Schuster, and G. Porenta, "Automated assessment of myocardial SPECT perfusion scintigraphy: a comparison of different approaches of case-based reasoning," *Artif. Intell. Med.*, vol. 40, no. 2, pp. 103–113, Jun. 2007, doi: 10.1016/j.artmed.2007.02.004.
- [50] R. S. Driessen *et al.*, "Automated SPECT analysis compared with expert visual scoring for the detection of FFR-defined coronary artery disease," *Eur. J. Nucl. Med. Mol. Imaging*, vol. 45, no. 7, pp. 1091–1100, 2018, doi: 10.1007/s00259-018-3951-1.

- [51] R. Arsanjani *et al.*, “Prediction of revascularization after myocardial perfusion SPECT by machine learning in a large population,” *J. Nucl. Cardiol.*, vol. 22, no. 5, pp. 877–884, Oct. 2015, doi: 10.1007/s12350-014-0027-x.
- [52] L. D *et al.*, “Scandinavian test of artificial neural network for classification of myocardial perfusion images,” *Clin. Physiol. Oxf. Engl.*, vol. 20, no. 4, pp. 253–261, Jul. 2000, doi: 10.1046/j.1365-2281.2000.00255.x.
- [53] “A survey of shaped-based registration and segmentation techniques for cardiac images | Computer Vision and Image Understanding.”
<https://dl.acm.org/doi/abs/10.1016/j.cviu.2012.11.017> (accessed Mar. 27, 2021).
- [54] “What is a ROC Curve and How to Interpret It,” *Displayr*, Jul. 05, 2018.
<https://www.displayr.com/what-is-a-roc-curve-how-to-interpret-it/> (accessed Aug. 26, 2021).
- [55] S. Narkhede, “Understanding AUC - ROC Curve,” *Medium*, Jun. 15, 2021.
<https://towardsdatascience.com/understanding-auc-roc-curve-68b2303cc9c5> (accessed Aug. 26, 2021).
- [56] “A tutorial on support vector regression | SpringerLink.”
<https://link.springer.com/article/10.1023/B:STCO.0000035301.49549.88> (accessed Apr. 25, 2021).
- [57] J. Friedman, T. Hastie, and R. Tibshirani, “Additive logistic regression: a statistical view of boosting (With discussion and a rejoinder by the authors),” *Ann. Stat.*, vol. 28, no. 2, pp. 337–407, Apr. 2000, doi: 10.1214/aos/1016218223.
- [58] “Random Forests | SpringerLink.”
<https://link.springer.com/article/10.1023/A:1010933404324> (accessed Apr. 25, 2021).
- [59] P. B. Soyer, “EU Proposes a Uniform Approach to the Regulation of Artificial Intelligence,” *International Maritime and Commercial Law*, Apr. 28, 2021.
<https://iistl.blog/2021/04/28/eu-proposes-a-uniform-approach-to-the-regulation-of-artificial-intelligence/> (accessed Jul. 22, 2021).
- [60] M. Hall, “Artificial intelligence and nuclear medicine,” *Nucl. Med. Commun.*, vol. 40, no. 1, pp. 1–2, Jan. 2019, doi: 10.1097/MNM.0000000000000937.
- [61] *Superintelligence: Paths, Dangers, Strategies*. Oxford, New York: Oxford University Press, 2014.
- [62] D. A. Clifton, K. E. Niehaus, P. Charlton, and G. W. Colopy, “Health Informatics via Machine Learning for the Clinical Management of Patients,” *Yearb. Med. Inform.*, vol. 10, no. 1, pp. 38–43, Aug. 2015, doi: 10.15265/IY-2015-014.
- [63] “Machine Learning (Supervised & unsupervised) - Diego Calvo.”
<https://www.diegocalvo.es/en/machine-learning-supervised-unsupervised/> (accessed Jul. 25, 2021).
- [64] J. Brownlee, “Best Practices for Preparing and Augmenting Image Data for CNNs,” *Machine Learning Mastery*, May 02, 2019. <https://machinelearningmastery.com/best-practices-for-preparing-and-augmenting-image-data-for-convolutional-neural-networks/> (accessed Apr. 04, 2021).
- [65] “Image Pre-processing. In this article, we are going to go... | by Prince Canuma | Towards Data Science.” <https://towardsdatascience.com/image-pre-processing-c1aec0be3edf> (accessed Apr. 04, 2021).
- [66] “Backpropagation - Wikipedia.” <https://en.wikipedia.org/wiki/Backpropagation> (accessed Sep. 01, 2021).

- [67] S. S. Haykin and S. S. Haykin, *Neural networks and learning machines*, 3rd ed. New York: Prentice Hall, 2009.
- [68] "Training, validation, and test sets," *Wikipedia*. Mar. 19, 2021. Accessed: Apr. 05, 2021. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Training,_validation,_and_test_sets&oldid=1013048903
- [69] "On Splitting Training and Validation Set: A Comparative Study of Cross-Validation, Bootstrap and Systematic Sampling for Estimating the Generalization Performance of Supervised Learning | SpringerLink." <https://link.springer.com/article/10.1007/s41664-018-0068-2> (accessed Apr. 05, 2021).
- [70] "What is the Difference Between Test and Validation Datasets?" <https://machinelearningmastery.com/difference-test-validation-datasets/> (accessed Apr. 05, 2021).
- [71] "A Logical Calculus of Ideas Immanent in Nervous Activity | BibSonomy." <https://www.bibsonomy.org/bibtex/13e8e0d06f376f3eb95af89d5a2f15957/schaul> (accessed Apr. 05, 2021).
- [72] "Artificial neural network," *Wikipedia*. Mar. 22, 2021. Accessed: Apr. 04, 2021. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Artificial_neural_network&oldid=1013668126
- [73] D. Visvikis, C. Cheze Le Rest, V. Jaouen, and M. Hatt, "Artificial intelligence, machine (deep) learning and radio(geno)mics: definitions and nuclear medicine imaging applications," *Eur. J. Nucl. Med. Mol. Imaging*, vol. 46, no. 13, pp. 2630–2637, Dec. 2019, doi: 10.1007/s00259-019-04373-w.
- [74] "Deep learning," *Wikipedia*. Mar. 14, 2021. Accessed: Apr. 04, 2021. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Deep_learning&oldid=1012107396
- [75] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, May 2017, doi: 10.1145/3065386.
- [76] "Deep neural networks for object detection | Proceedings of the 26th International Conference on Neural Information Processing Systems - Volume 2." <https://dl.acm.org/doi/10.5555/2999792.2999897> (accessed Apr. 03, 2021).
- [77] M. Stefanini, R. Lancellotti, L. Baraldi, and S. Calderara, "A Deep-learning-based approach to VM behavior Identification in Cloud Systems," Mar. 2021, pp. 308–315. Accessed: Mar. 29, 2021. [Online]. Available: <https://www.scitepress.org/Link.aspx?doi=10.5220/0007708403080315>
- [78] G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely Connected Convolutional Networks," *ArXiv160806993 Cs*, Jan. 2018, Accessed: Mar. 29, 2021. [Online]. Available: <http://arxiv.org/abs/1608.06993>
- [79] S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," in *International Conference on Machine Learning*, Jun. 2015, pp. 448–456. Accessed: Mar. 27, 2021. [Online]. Available: <http://proceedings.mlr.press/v37/ioffe15.html>
- [80] "Cloud Computing," *EXRION*. <https://www.exrion.com/cloud-computing.html> (accessed Jul. 25, 2021).

- [81] H. Chen, Q. Dou, X. Wang, J. Qin, and P.-A. Heng, "Mitosis detection in breast cancer histology images via deep cascaded networks," in *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*, Phoenix, Arizona, Feb. 2016, pp. 1160–1166.
- [82] J. Brownlee, "How Do Convolutional Layers Work in Deep Learning Neural Networks?," *Machine Learning Mastery*, Apr. 16, 2019.
<https://machinelearningmastery.com/convolutional-layers-for-deep-learning-neural-networks/> (accessed Apr. 10, 2021).
- [83] "A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way | by Sumit Saha | Towards Data Science." <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53> (accessed Apr. 06, 2021).
- [84] "Convolutional Neural Network (CNN) | TensorFlow Core," *TensorFlow*.
<https://www.tensorflow.org/tutorials/images/cnn> (accessed Apr. 10, 2021).
- [85] "Convolutional neural network," *Wikipedia*. Apr. 01, 2021. Accessed: Apr. 06, 2021.
[Online]. Available:
https://en.wikipedia.org/w/index.php?title=Convolutional_neural_network&oldid=1015376106
- [86] N. Kalchbrenner, E. Grefenstette, and P. Blunsom, "A Convolutional Neural Network for Modelling Sentences," *ArXiv14042188 Cs*, Apr. 2014, Accessed: Apr. 06, 2021. [Online]. Available: <http://arxiv.org/abs/1404.2188>
- [87] "Rectifier (neural networks)," *Wikipedia*. Jul. 22, 2021. Accessed: Aug. 10, 2021.
[Online]. Available:
[https://en.wikipedia.org/w/index.php?title=Rectifier_\(neural_networks\)&oldid=1034977791](https://en.wikipedia.org/w/index.php?title=Rectifier_(neural_networks)&oldid=1034977791)
- [88] "The Most Intuitive and Easiest Guide for Convolutional Neural Network | by Jiwon Jeong | Towards Data Science." <https://towardsdatascience.com/the-most-intuitive-and-easiest-guide-for-convolutional-neural-network-3607be47480> (accessed Apr. 07, 2021).
- [89] "The Differences between Sigmoid and Softmax Activation Functions | by Nikola Basta | Arteos AI | Medium." <https://medium.com/arteos-ai/the-differences-between-sigmoid-and-softmax-activation-function-12adee8cf322> (accessed Apr. 06, 2021).
- [90] V. Romanuke, "Appropriate Number and Allocation of ReLUs in Convolutional Neural Networks," *Res. Bull. Natl. Tech. Univ. Ukr. Kyiv Politech. Inst.*, no. 1, Art. no. 1, Mar. 2017, doi: 10.20535/1810-0546.2017.1.88156.
- [91] J. Zang *et al.*, "Attention-based Temporal Weighted Convolutional Neural Network for Action Recognition," *ArXiv180307179 Cs*, Mar. 2018, Accessed: Apr. 06, 2021. [Online]. Available: <http://arxiv.org/abs/1803.07179>
- [92] purnasai gudikandula, "Deep view on Transfer learning with Image classification Pytorch," *Medium*, Mar. 22, 2019. <https://purnasaigudikandula.medium.com/deep-view-on-transfer-learning-with-image-classification-pytorch-5cf963939575> (accessed Jul. 25, 2021).
- [93] L. Pratt, "Discriminability-Based Transfer between Neural Networks," 1992.
- [94] "Transfer learning," *Wikipedia*. Apr. 04, 2021. Accessed: Apr. 06, 2021. [Online]. Available:
https://en.wikipedia.org/w/index.php?title=Transfer_learning&oldid=1016005675
- [95] "Image classification: A comparison of DNN, CNN and Transfer Learning approach | by Lalit Pal | Analytics Vidhya | Medium." <https://medium.com/analytics-vidhya/image->

- classification-a-comparison-of-dnn-cnn-and-transfer-learning-approach-704535beca25 (accessed Jul. 25, 2021).
- [96] L. E. Juarez-Orozco, O. Martinez-Manzanera, A. E. Storti, and J. Knuuti, "Machine Learning in the Evaluation of Myocardial Ischemia Through Nuclear Cardiology," *Curr. Cardiovasc. Imaging Rep.*, vol. 12, no. 2, p. 5, Feb. 2019, doi: 10.1007/s12410-019-9480-x.
 - [97] J. Betancur *et al.*, "Deep Learning for Prediction of Obstructive Disease From Fast Myocardial Perfusion SPECT," *JACC Cardiovasc. Imaging*, vol. 11, no. 11, pp. 1654–1663, Nov. 2018, doi: 10.1016/j.jcmg.2018.01.020.
 - [98] S. Kaplan Berkaya, I. Ak Sivrikoz, and S. Gunal, "Classification models for SPECT myocardial perfusion imaging," *Comput. Biol. Med.*, vol. 123, p. 103893, Aug. 2020, doi: 10.1016/j.combiomed.2020.103893.
 - [99] "What are Generative Adversarial Networks - Akira AI," *Akira AI*, <https://www.akira.ai/glossary/generative-adversarial-networks/> (accessed Jul. 25, 2021).
 - [100] C. Han *et al.*, "GAN-based synthetic brain MR image generation," *IEEE 15th Int. Symp. Biomed. Imaging ISBI 2018*, pp. 734–738, 2018.
 - [101] D. Nie *et al.*, "Medical Image Synthesis with Deep Convolutional Adversarial Networks," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 12, pp. 2720–2730, Dec. 2018, doi: 10.1109/TBME.2018.2814538.
 - [102] "Generative Adversarial Networks for Noise Reduction in Low-Dose CT | IEEE Journals & Magazine | IEEE Xplore." <https://ieeexplore.ieee.org/document/7934380> (accessed Apr. 25, 2021).
 - [103] "(PDF) MedGAN: Medical Image Translation using GANs." https://www.researchgate.net/publication/325841281_MedGAN_Medical_Image_Translation_using_GANs (accessed Apr. 25, 2021).
 - [104] R. Arsanjani *et al.*, "Improved accuracy of myocardial perfusion SPECT for detection of coronary artery disease by machine learning in a large population," *J. Nucl. Cardiol. Off. Publ. Am. Soc. Nucl. Cardiol.*, vol. 20, no. 4, pp. 553–562, Aug. 2013, doi: 10.1007/s12350-013-9706-2.
 - [105] "Utility of the Diamond-Forrester classification in stratifying acute chest pain in an academic chest pain center." <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4874514/> (accessed Jul. 25, 2021).
 - [106] K. Nakajima *et al.*, "Diagnostic accuracy of an artificial neural network compared with statistical quantitation of myocardial perfusion images: a Japanese multicenter study," *Eur. J. Nucl. Med. Mol. Imaging*, vol. 44, no. 13, pp. 2280–2289, Dec. 2017, doi: 10.1007/s00259-017-3834-x.
 - [107] D. H. Alonso, M. N. Wernick, Y. Yang, G. Germano, D. S. Berman, and P. Slomka, "PREDICTION OF CARDIAC DEATH AFTER ADENOSINE MYOCARDIAL PERFUSION SPECT BASED ON MACHINE LEARNING," *J. Nucl. Cardiol. Off. Publ. Am. Soc. Nucl. Cardiol.*, vol. 26, no. 5, pp. 1746–1754, Oct. 2019, doi: 10.1007/s12350-018-1250-7.
 - [108] L. E. Juarez-Orozco, R. J. J. Knol, C. A. Sanchez-Catasus, O. Martinez-Manzanera, F. M. van der Zant, and J. Knuuti, "Machine learning in the integration of simple variables for identifying patients with myocardial ischemia," *J. Nucl. Cardiol. Off. Publ. Am. Soc. Nucl. Cardiol.*, vol. 27, no. 1, pp. 147–155, Feb. 2020, doi: 10.1007/s12350-018-1304-x.

- [109] N. Spier, S. Nekolla, C. Rupprecht, M. Mustafa, N. Navab, and M. Baust, “Classification of Polar Maps from Cardiac Perfusion Imaging with Graph-Convolutional Neural Networks,” *Sci. Rep.*, vol. 9, no. 1, p. 7569, Dec. 2019, doi: 10.1038/s41598-019-43951-8.
- [110] J. Betancur *et al.*, “Deep Learning Analysis of Upright-Supine High-Efficiency SPECT Myocardial Perfusion Imaging for Prediction of Obstructive Coronary Artery Disease: A Multicenter Study,” *J. Nucl. Med.*, vol. 60, no. 5, pp. 664–670, May 2019, doi: 10.2967/jnumed.118.213538.
- [111] E. W. Steyerberg and F. E. Harrell, “Prediction models need appropriate internal, internal-external, and external validation,” *J. Clin. Epidemiol.*, vol. 69, pp. 245–247, Jan. 2016, doi: 10.1016/j.jclinepi.2015.04.005.
- [112] S. E. Bleeker *et al.*, “External validation is necessary in prediction research: a clinical example,” *J. Clin. Epidemiol.*, vol. 56, no. 9, pp. 826–832, Sep. 2003, doi: 10.1016/s0895-4356(03)00207-5.
- [113] R. Arsanjani *et al.*, “Comparison of Fully Automated Computer Analysis and Visual Scoring for Detection of Coronary Artery Disease from Myocardial Perfusion SPECT in a Large Population,” *J. Nucl. Med. Off. Publ. Soc. Nucl. Med.*, vol. 54, no. 2, pp. 221–228, Feb. 2013, doi: 10.2967/jnumed.112.108969.
- [114] D. Sarkar, R. Bali, and T. Ghosh, *Hands-On Transfer Learning with Python: Implement advanced deep learning and neural network models using TensorFlow and Keras*. Packt Publishing, 2018.
- [115] I. D. Apostolopoulos, N. D. Papathanasiou, T. Spyridonidis, and D. J. Apostolopoulos, “Automatic characterization of myocardial perfusion imaging polar maps employing deep learning and data augmentation,” *Hell. J. Nucl. Med.*, vol. 23, no. 2, pp. 125–132, Aug. 2020, doi: 10.1967/s002449912101.
- [116] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” *ArXiv14091556 Cs*, Apr. 2015, Accessed: May 15, 2021. [Online]. Available: <http://arxiv.org/abs/1409.1556>
- [117] “Artificial intelligence,” *Wikipedia*. Mar. 27, 2021. Accessed: Apr. 05, 2021. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Artificial_intelligence&oldid=1014538985
- [118] “What is AI? Everything you need to know about Artificial Intelligence | ZDNet.” <https://www.zdnet.com/article/what-is-ai-everything-you-need-to-know-about-artificial-intelligence/> (accessed Apr. 05, 2021).
- [119] T. M. Press, “Foundations of Machine Learning | The MIT Press.” <https://mitpress.mit.edu/books/foundations-machine-learning> (accessed Apr. 05, 2021).
- [120] “OpenCV: Image Inpainting.” https://docs.opencv.org/3.4.0/df/d3d/tutorial_py_inpainting.html (accessed May 15, 2021).
- [121] “Create Flower Image Classification with Tensorflow, Keras and Flask | by Ferry Djaja | Medium.” <https://medium.com/@ferrygunawan/create-flower-image-classification-with-tensorflow-keras-and-flask-4fd6efd4c5ac> (accessed May 15, 2021).
- [122] “Flask (web framework),” *Wikipedia*. Jun. 01, 2021. Accessed: Jul. 25, 2021. [Online]. Available: [https://en.wikipedia.org/w/index.php?title=Flask_\(web_framework\)&oldid=1026379015](https://en.wikipedia.org/w/index.php?title=Flask_(web_framework)&oldid=1026379015)
- [123] “Flask Tutorial - Tutorialspoint.” <https://www.tutorialspoint.com/flask/index.htm> (accessed Jul. 25, 2021).

- [124] “Receiver operating characteristic - Wikipedia.”
https://en.wikipedia.org/wiki/Receiver_operating_characteristic (accessed Jul. 25, 2021).
- [125] “AUC-ROC Curve in Machine Learning Clearly Explained,” *Analytics Vidhya*, Jun. 15, 2020. <https://www.analyticsvidhya.com/blog/2020/06/auc-roc-curve-machine-learning/> (accessed Jul. 25, 2021).
- [126] J. Brownlee, “A Gentle Introduction to k-fold Cross-Validation,” *Machine Learning Mastery*, May 22, 2018. <https://machinelearningmastery.com/k-fold-cross-validation/> (accessed Jul. 25, 2021).
- [127] J. Brownlee, “How to Save and Load Your Keras Deep Learning Model,” *Machine Learning Mastery*, May 12, 2019. <https://machinelearningmastery.com/save-load-keras-deep-learning-models/> (accessed Jul. 25, 2021).
- [128] “An introduction to explainable AI, and why we need it,” *KDnuggets*.
<https://www.kdnuggets.com/an-introduction-to-explainable-ai-and-why-we-need-it.html/> (accessed Sep. 12, 2021).
- [129] J. Amann, A. Blasimme, E. Vayena, D. Frey, V. I. Madai, and the Precise4Q consortium, “Explainability for artificial intelligence in healthcare: a multidisciplinary perspective,” *BMC Med. Inform. Decis. Mak.*, vol. 20, no. 1, p. 310, Nov. 2020, doi: 10.1186/s12911-020-01332-6.