



ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

Βαθιά Μάθηση με εφαρμογές στην Ιατρική

Αντώνιος Αθανασίου

A.M: 44518

Επιβλέπον Καθηγητής

Γρηγόριος Νικολάου

ΣΧΟΛΗ ΤΕΧΝΟΛΟΓΙΚΩΝ ΕΦΑΡΜΟΓΩΝ

ΤΜΗΜΑ ΑΥΤΟΜΑΤΙΣΜΟΥ

ΑΙΓΑΛΕΩ ΣΕΠΤΕΜΒΡΙΟΣ 2018

ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ (1^η ή 2^η σελίδα)

ΔΗΛΩΣΗ ΣΥΓΓΡΑΦΕΑ ΠΤΥΧΙΑΚΗΣ ΕΡΓΑΣΙΑΣ

Ο/Η κάτωθι υπογεγραμμένος/η Αθανασίου Αντώνιος, του Ιωάννη, φοιτητής του Τμήματος Μηχανικών Βιομηχανικής Σχεδίασης και Παραγωγής του Πανεπιστημίου Δυτικής Αττικής, πριν αναλάβω την εκπόνηση της Πτυχιακής Εργασίας μου, δηλώνω ότι ενημερώθηκα για τα παρακάτω:

«Η Πτυχιακή Εργασία (Π.Ε) αποτελεί προϊόν πνευματικής ιδιοκτησίας τόσο του συγγραφέα, όσο και του Ιδρύματος και θα πρέπει να έχει μοναδικό χαρακτήρα και πρωτότυπο περιεχόμενο.

Απαγορεύεται αυστηρά οποιοδήποτε κομμάτι κειμένου της να εμφανίζεται αυτούσιο ή μεταφρασμένο από κάποια άλλη δημοσιευμένη πηγή. Κάθε τέτοια πράξη αποτελεί προϊόν λογοκλοπής και εγείρει θέμα Ηθικής Τάξης για τα πνευματικά δικαιώματα του άλλου συγγραφέα. Αποκλειστικός υπεύθυνος είναι ο συγγραφέας της Π.Ε, ο οποίος φέρει και την ευθύνη των συνεπειών, ποινικών και άλλων, αυτής της πράξης.

Πέραν των όποιων ποινικών ευθυνών του συγγραφέα, σε περίπτωση που το Ίδρυμα του έχει απονείμει Πτυχίο, αυτό ανακαλείται με απόφαση της Συνέλευσης του Τμήματος. Η Συνέλευση του Τμήματος με νέα απόφασή της, μετά από αίτηση του ενδιαφερόμενου, του αναθέτει εκ νέου την εκπόνηση Π.Ε με άλλο θέμα και διαφορετικό επιβλέποντα καθηγητή. Η εκπόνηση της εν λόγω Π.Ε πρέπει να ολοκληρωθεί εντός τουλάχιστον ενός ημερολογιακού δμήνου από την ημερομηνία ανάθεσής της.

Ο Δηλών



Ημερομηνία

09/11/2018

Περίληψη

Η βαθιά μάθηση αποτελεί μέρος μιας ευρύτερης οικογένειας της μηχανικής μάθησης και βρίσκεται στην αιχμή της τεχνολογίας. Είναι μία τεχνική μηχανικής μάθησης η οποία διδάσκει στους υπολογιστές να πράττουν αυτό που στους ανθρώπους έρχεται φυσικά: μάθηση μέσω παραδειγμάτων. Η βαθιά μάθηση είναι η βασική τεχνολογία πίσω από τα αυτόνομα οχήματα και αποτελεί το κλειδί του φωνητικού ελέγχου σε καταναλωτικές συσκευές.

Η συγκεκριμένη εργασία θα εμβαθύνει στις εφαρμογές της στην Ιατρική, πιο συγκεκριμένα στη δημιουργία ενός μοντέλου εκπαίδευσης το οποίο θα μπορεί να κατηγοριοποιεί αξονικές τομογραφίες θώρακα σύμφωνα με την ύπαρξη ή μη καρκίνου του πνεύμονα στον ασθενή.

Στην εργασία χρησιμοποιήθηκε ή υψηλού επιπέδου γλώσσα προγραμματισμού Python, η Tensor Flow, μια βιβλιοθήκη λογισμικού ανοιχτού κώδικα για προγραμματισμό ροής δεδομένων καθώς και μία πληθώρα βιβλιοθηκών που κατέστησαν δυνατή τη χρήση και επεξεργασία των ιατρικών δεδομένων. Τα ιατρικά δεδομένα που χρησιμοποιήθηκαν πάρθηκαν από τον διαγωνισμό Data Science Bowl 2017 του Kaggle.

Abstract

Deep learning is part of a broader family of machine learning and is at the cutting edge of technology. Deep learning is a machine learning technique that teaches computers to do what comes naturally to humans: learn by example. Deep learning is a key technology behind driverless cars and the key to voice control in consumer devices.

This particular work will delve into its applications in Medicine, specifically in the creation of a training model that can classify computed tomography scans of the thorax according to whether or not the patient has lung cancer.

In this dissertation, Python, a high-level programming language, Tensor Flow, an open-source software library for dataflow programming, as well as a plethora of libraries were utilized to make possible the use and process of the medical data. The medical data used was taken from the Kaggle Data Science Bowl 2017 competition.

Περιεχόμενα

Κατάλογος Εικόνων	7
Κατάλογος Πινάκων.....	9
Ευχαριστίες.....	10
Εισαγωγή.....	11
Κεφάλαιο 1 ^ο Θεωρητικό Μέρος	12
1.1 Βαθιά μάθηση (Deep Learning)	12
1.1.1 Το ξεκίνημα (1940)	12
1.1.2 Η διάψευση του perceptron (1969).....	13
1.1.3 Η εμφάνιση της οπισθοδιάδοσης (1986).....	14
1.1.4 Μετονομασία τους σε «Βαθεία Μάθηση» (2006).....	15
1.1.5 Η επανάσταση (2012).....	15
1.2 Νευρωνικό δίκτυο (Neural Network).....	18
1.2.1 Ορισμός και λειτουργία.....	18
1.2.2 Συναρτήσεις ενεργοποίησης	20
1.2.3 Ιδιότητες	21
1.2.4 Εφαρμογές.....	22
1.2.4 Εκπαίδευση Νευρωνικών Δικτύων.....	23
Κεφάλαιο 2 ^ο Μέσα Υλοποίησης Πτυχιακής	26
2.1 Python	26
2.1.1 Jupyter Notebook.....	28
2.1.3 Απαραίτητες βιβλιοθήκες	30
2.2 TensorFlow	33
2.2.1 Το ξεκίνημα.....	34
2.2.2 Η δημιουργία του TensorFlow	35

2.2.3 Μονάδα επεξεργασίας Tensor	35
2.2.4 Εφαρμογές.....	36
2.2.5 TFLearn	37
2.2.6 Παράδειγμα Dogs vs Cats.....	39
Κεφάλαιο 3 ^ο Πρακτικό Μέρος	49
3.1 Εισαγωγή.....	49
3.2 Συναφή Έργα	50
3.3 Ιατρικά Δεδομένα.....	52
3.4 Σχεδιασμός και Υλοποίηση	52
3.4.1 Προεπεξεργασία των δεδομένων	53
3.4.2 Σχεδιασμός Νευρωνικού Δικτύου	58
3.5 Αποτελέσματα.....	62
3.6 Συμπεράσματα	66
Βιβλιογραφία	68

Κατάλογος Εικόνων

Εικόνα 1.1 - Ιστορία της Βαθιάς Μάθησης.....	12
Εικόνα 1.2 - Χρήση GPUs στο Imagenet και δείκτης σφάλματος των νικητών.....	177
Εικόνα 2.1 - Πλατφόρμα πληροφοριών TensorFlow.....	344
Εικόνα 2.2 - Διάγραμμα μπλοκ TPU.....	366
Εικόνα 2.3 - - Η πρώτη μονάδα επεξεργασίας Tensor (TPU) της Google σε μια πλακέτα τυπωμένου κυκλώματος (αριστερά). TPU που χρησιμοποιείται σε κέντρο δεδομένων της Google (δεξιά).....	388
Εικόνα 2.4 - Δείγμα του συνόλου δεδομένων.....	399
Εικόνα 2.5 - CNN αρχιτεκτονική του παραδείγματος Dogs vs Cats	422
Εικόνα 2.6 - Οπτικός έλεγχος του δικτύου σε μη επισημασμένα δεδομένα	433
Εικόνα 2.7 - ROC για CNN του παραδείγματος Dogs vs Cats	444
Εικόνα 2.8 - Αρχιτεκτονική μοντέλου AlexNet	466
Εικόνα 2.9 - CNN αρχιτεκτονική του εφαρμοσμένου μοντέλου AlexNet	477
Εικόνα 2.10 - Καμπύλη Ευστοχίας (Accuracy)	488
Εικόνα 2.11 - Καμπύλη Σφάλματος (Loss)	488
Εικόνα 3.1 - Ιστογράμματα των τιμών των εικονοστοιχείων σε μονάδα Hounsfield	544
Εικόνα 3.2 - Δεδομένα πριν την προεπεξεργασία.....	544
Εικόνα 3.3 - Δείγμα 3D εικόνας ασθενούς με τιμές εικονοστοιχείων μεγαλύτερες από 400 HU αποκαλύπτει το τμήμα των οστών	555

Εικόνα 3.4 - Το βρογχικό δέντρο για δείγμα ασθενούς	555
Εικόνα 3.5 - Αρχική απεικόνιση χωρίς αέρα για δείγμα ασθενούς	566
Εικόνα 3.6 - Τελική απεικόνιση συμπεριλαμβανομένων των βρογχιδιών για δείγμα ασθενούς	566
Εικόνα 3.7 - Δεδομένα μετά την προεπεξεργασία	588
Εικόνα 3.8 - CNN Αρχιτεκτονική	611
Εικόνα 3.9 - ROC για 3D CNN	655
Εικόνα 3.10 - Αρχιτεκτονική U-Net	677

Κατάλογος Πινάκων

Πίνακας 2.1 - CNN αρχιτεκτονική 6 επιπέδων (dropout με 0.8 πιθανότητα μετά από κάθε conv επίπεδο κατά την εκπαίδευση, Adam βελτιστοποιητής με δείκτη εκπαίδευσης 0.001).....	411
Πίνακας 2.2 - Πίνακας Σύγκυσης για CNN του παραδείγματος Dogs vs Cats.....	444
Πίνακας 2.3 - Έκθεση κατηγοριοποίησης για CNN του παραδείγματος Dogs vs Cats	444
Πίνακας 3.1 - Κλίμακα Hounsfield	533
Πίνακας 3.2 - CNN αρχιτεκτονική 4 επιπέδων (dropout με 0.8 πιθανότητα μετά από κάθε conv επίπεδο κατά την εκπαίδευση, Adam βελτιστοποιητής με δείκτη εκπαίδευσης 0.001).....	599
Πίνακας 3.3 - Πίνακας Σύγκυσης για 3D CNN.....	633
Πίνακας 3.4 - Έκθεση κατηγοριοποίησης για 3D CNN	644

Ευχαριστίες

Αρχικά θα ήθελα να ευχαριστήσω το ΑΕΙ Πειραιά ΤΤ, στο οποίο κατά τη διάρκεια φοίτησής μου απέκτησα πολύτιμες γνώσεις τόσο για την επιτυχή ολοκλήρωση της παρούσας πτυχιακής εργασίας όσο και για την μελλοντική μου σταδιοδρομία. Επίσης, θα ήθελα να ευχαριστήσω θερμά τον επιβλέποντα καθηγητή μου, κύριο Γρηγόριο Νικολάου, για την παρότρυνσή του να αναλάβω την παρούσα εργασία καθώς και για την βοήθεια και καθοδήγηση που μου πρόσφερε καθ' όλη την διάρκεια διεκπεραίωσής της. Τέλος, θα ήθελα να ευχαριστήσω την οικογένειά μου για την υποστήριξη όπως επίσης και για την εμπιστοσύνη που πρόσφεραν κατά την διάρκεια των σπουδών μου.

Εισαγωγή

Η βαθιά μάθηση, τα τελευταία πέντε χρόνια περίπου, έχει περάσει από ένα κάπως εξειδικευμένο πεδίο αποτελούμενο από απομονωμένες ομάδες ερευνητών, στο να αποτελεί ένα από τα κυρίαρχα ρεύματα του κλάδου. Η ταχεία άνοδος και η προφανής κυριαρχία της βαθιάς μάθησης σε σχέση με τις παραδοσιακές μεθόδους μηχανικής μάθησης σε μια ποικιλία καθηκόντων ήταν εκπληκτική και μάλιστα δύσκολο να εξηγηθεί.

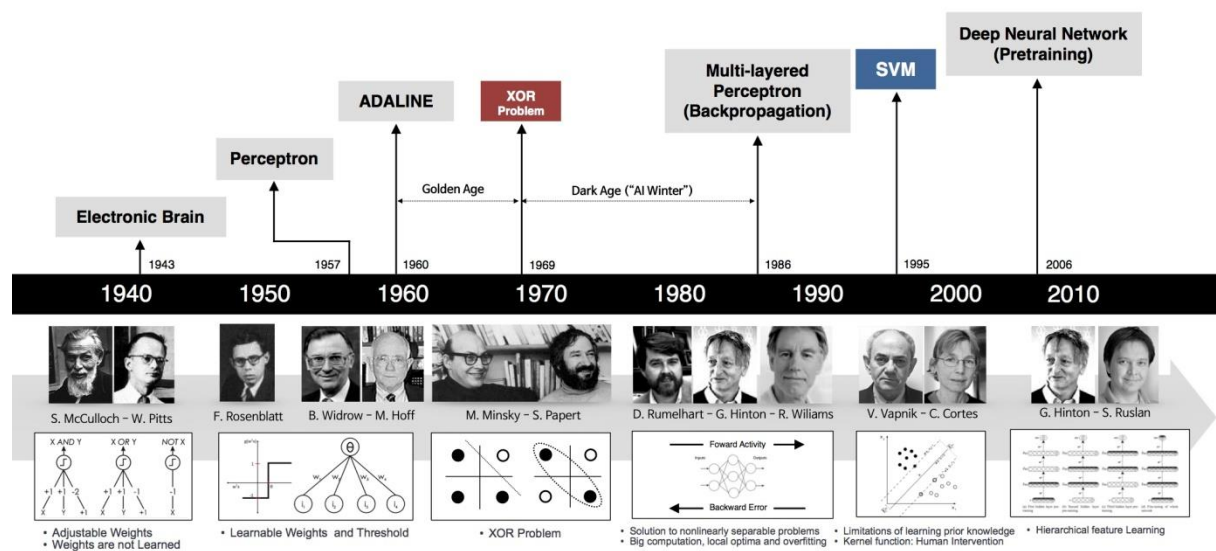
Κατά συνέπεια, το ενδιαφέρον για τη βαθιά εκμάθηση έχει ανέβει στα ύψη, με συνεχή κάλυψη σε δημοφιλή μέσα μαζικής ενημέρωσης όπως τα Science και Nature.

Τι σημαίνει αυτή η δημοσιότητα; Γιατί συμβαίνει αυτό τώρα; Σε αρχικό στάδιο, αυτό που είδαμε είναι ο συνεργατικός συνδυασμός των ιδεών από τη βελτιστοποίηση, την παραδοσιακή μηχανική μάθηση, την τεχνολογία λογισμικού, το σχεδιασμό ολοκληρωμένων κυκλωμάτων και την αξιοπιστία των ιδίων των ιδεών από μια ειδική ομάδα ερευνητών, μεταδιδακτορικών φοιτητών και προπτυχιακών φοιτητών. Αν και οι βασικές ιδέες πίσω από τη βαθιά εκμάθηση έχουν τεθεί σε ισχύ εδώ και δεκαετίες, δεν ήταν έως ότου ο όγκος των δεδομένων έγινε αρκετά μεγάλος και οι συμβατικοί υπολογιστές αρκετά γρήγοροι ώστε να αποκαλυφθεί η πραγματική δύναμή τους. Έχοντας πει αυτά, αξίζει να περάσουμε από την ιστορία των νευρωνικών δικτύων και της βαθιάς μάθησης για να δούμε πώς φτάσαμε εδώ.

Κεφάλαιο 1^ο Θεωρητικό Μέρος

1.1 Βαθιά μάθηση (Deep Learning)

1.1.1 Το ξεκίνημα (1940)



Εικόνα 1.1 - Ιστορία της Βαθιάς Μάθησης

Η ιδέα της δημιουργίας μιας «σκεπτόμενης μηχανής» είναι τουλάχιστον τόσο παλιά όσο και η σύγχρονη πληροφορική, αν όχι και παλαιότερη. Ο Alan Turing στο άρθρο του «Computing Machinery and Intelligence»[1] έθεσε διάφορα κριτήρια για να κρίνει αν μια μηχανή θα μπορούσε να ειπωθεί ως έξυπνη, η οποία από τότε έγινε γνωστή ως η «δοκιμή Turing».

Η αρχική εργασία στην εκμάθηση μηχανών ενημερώθηκε σε μεγάλο βαθμό από τις τρέχουσες θεωρίες για την λειτουργία του εγκεφάλου. Οι πρωτοπόροι της σκηνης ήταν ο Walter Pitts και ο Warren McCulloch. Είχαν αναπτύξει μια τεχνική γνωστή ως "thresholded logic unit" και σχεδιάστηκε για να μιμείται τον τρόπο με τον οποίο ένας νευρώνας θεωρούταν ότι λειτουργεί (που θα αποτελέσει ένα επαναλαμβανόμενο θέμα).

Επιστρέφοντας στο θέμα μας, δεν υπάρχει παρά ο αλγόριθμος "perceptron" του Frank Rosenblatt όπου βλέπουμε τον πρώτο πραγματικό πρόδρομο των σύγχρονων νευρωνικών δικτύων. Για την εποχή του ήταν αρκετά εντυπωσιακός. Ήρθε με μια διαδικασία μάθησης που θα μπορούσε να συγκλίνει αποδεδειγμένα στη σωστή λύση και θα μπορούσε να αναγνωρίσει γράμματα και αριθμούς. Ο Rosenblatt ήταν τόσο σίγουρος ότι το perceptron θα οδηγούσε στην πραγματική τεχνητή νοημοσύνη, ώστε το 1959 να σχολιάσει:

«[Το perceptron] είναι το έμβρυο ενός ηλεκτρονικού υπολογιστή που [το Ναυτικό] αναμένει ότι θα μπορέσει να περπατήσει, να μιλήσει, να δει, να γράψει, να αναπαραχθεί και να συνειδητοποιήσει την ύπαρξή του.» (Rosenblatt, 1959)

1.1.2 Η διάψευση του perceptron (1969)

Το perceptron του Rosenblatt άρχισε να συγκεντρώνει αρκετή προσοχή και ένα άτομο ξεκίνησε να λαμβάνει γνώση. Ο Μάρβιν Μίνσκι, ο οποίος συχνά θεωρείται ως ένας από τους πατέρες της τεχνητής νοημοσύνης (AI), άρχισε να αισθάνεται ότι κάτι πηγαίνει στραβά με το perceptron του Rosenblatt.

Μαζί με τη Seymour Papert, ο Minsky έγραψε ένα βιβλίο με τίτλο «Perceptrons»[2] που ουσιαστικά διέψευσε πλήρως το perceptron, βάζοντας τέλος στην εμβρυϊκή ιδέα ενός νευρωνικού δικτύου. Έδειξαν ότι το perceptron ήταν ανίκανο να μάθει την απλή αποκλειστική λειτουργία (XOR). Ακόμη χειρότερα, απέδειξαν ότι ήταν θεωρητικά αδύνατο να μάθει μια τέτοια λειτουργία, ανεξάρτητα από το πόσο χρόνο το αφήσουμε να εκπαιδευτεί. Αυτό δεν αποτελεί έκπληξη για εμάς, καθώς το μοντέλο που υπονοείται από το perceptron είναι γραμμικό και η λειτουργία XOR είναι μη γραμμική, αλλά κατά την περίοδο, αυτό ήταν αρκετό για να τερματίσει όλη την έρευνα στα νευρικά δίκτυα.

1.1.3 Η εμφάνιση της οπισθοδιάδοσης (1986)

Ένας άνθρωπος με το όνομα Geoff Hinton, ολοκλήρωσε τις διδακτορικές του σπουδές στα νευρωνικά δίκτυα το 1978 και το 1986 μετακόμισε στο Ινστιτούτο για τη γνωστική επιστήμη στο UC San Diego. Μαζί με τον David Rumelhart και τον Ronald Williams, ο Hinton δημοσίευσε ένα άρθρο με τίτλο "Learning representations by back-propagating errors".[3] Σε αυτή την εργασία έδειξαν ότι τα νευρωνικά δίκτυα με πολλά κρυμμένα στρώματα θα μπορούσαν να εκπαιδεύονται αποτελεσματικά μέσω σχετικά απλής διαδικασίας. Αυτό θα επέτρεπε στα νευρωνικά δίκτυα να ξεπεράσουν την αδυναμία του perceptron επειδή τα πρόσθετα στρώματα έδιναν στα δίκτυα τη δυνατότητα να μάθουν μη γραμμικές λειτουργίες. Περίπου την ίδια χρονική στιγμή αποδείχθηκε ότι τέτοια δίκτυα είχαν την ικανότητα να μαθαίνουν οποιαδήποτε λειτουργία, αποτέλεσμα γνωστό ως «Universal approximation theorem». Και κάπως έτσι, τα νευρωνικά δίκτυα ήταν και πάλι στο προσκήνιο.

Ο αλγόριθμος λειτουργεί με τη λήψη του παραγώγου της συνάρτησης απώλειας του δικτύου και την οπισθοδιάδοση των σφαλμάτων για την ενημέρωση των παραμέτρων στα χαμηλότερα στρώματα.

Αυτός ο αλγόριθμος οδηγεί σε κάποιες πρώιμες επιτυχίες, κυρίως στην εκπαίδευση των συνελκτικών νευρωνικών δικτύων (γνωστών και ως CNN) για την αναγνώριση χειρόγραφων ψηφίων, ένα έργο το οποίο διεξήχθη με επικεφαλή τον Yann Le Cun στο AT&T Bell Labs.[4]

Δυστυχώς, τα νευρωνικά δίκτυα βρίσκονταν σε άλλη μια παύση. Η προσέγγιση δεν κλιμακώθηκε σε μεγαλύτερα προβλήματα και από τη δεκαετία του '90 η διανυσματική μηχανή υποστήριξης (SVM) έγινε η μέθοδος επιλογής και τα νευρωνικά δίκτυα αφέθηκαν στην άκρη. Εκ των υστέρων, τα νευρωνικά δίκτυα ήταν απλά μια ιδέα πριν από την εποχή τους, θα χρειαστούν σχεδόν δύο δεκαετίες πριν τα δεδομένα και η υπολογιστική ισχύς να είναι τέτοια που τα νευρωνικά δίκτυα θα μπορέσουν εντέλει να φτάσουν στην τελική τους μορφή.

1.1.4 Μετονομασία τους σε «Βαθιά Μάθηση» (2006)

Γύρω στο 2006, ο Hinton δήλωσε για άλλη μια φορά ότι ήξερε πώς λειτουργεί ο εγκέφαλος και εισήγαγε την ιδέα της μη εποπτικής προπαίδευσης και των deep belief network. Η ιδέα ήταν να εκπαιδεύσουμε ένα απλό δύστρομα μη εποπτευόμενου μοντέλου όπως μια περιορισμένη μηχανή Boltzmann, να παγώσουμε όλες τις παραμέτρους, να τοποθετήσουμε ένα νέο στρώμα στην κορυφή και να εκπαιδεύσουμε μόνο τις παραμέτρους για το νέο στρώμα. Θα συνεχίζαμε να προσθέτουμε και να εκπαιδεύουμε στρώματα με αυτόν τον «άπληστο αλγόριθμο» μέχρι να έχουμε ένα βαθύ δίκτυο και στη συνέχεια θα χρησιμοποιούσαμε το αποτέλεσμα αυτής της διαδικασίας για την προετοιμασία των παραμέτρων ενός παραδοσιακού νευρωνικού δικτύου. Χρησιμοποιώντας αυτή τη στρατηγική, οι άνθρωποι ήταν σε θέση να εκπαιδεύσουν δίκτυα που ήταν βαθύτερα από τις προηγούμενες προσπάθειες, προκαλώντας τον επαναπροσδιορισμό των «νευρωνικών δικτύων» σε «βαθιά μάθηση».

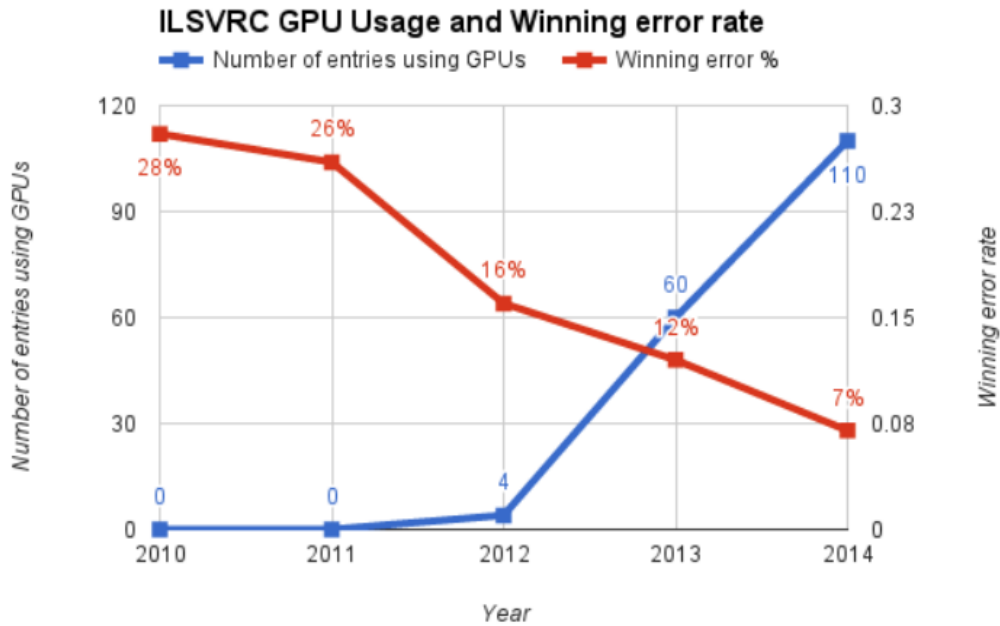
1.1.5 Η επανάσταση (2012)

Με το ανανεωμένο ενδιαφέρον που προκάλεσε η μη εποπτική προπαίδευση, όλο και περισσότερα επιστημονικά περιοδικά με θέμα τα νευρωνικά δίκτυα άρχισαν να διαρρέουν από πολλά ερευνητικά κέντρα. Στη συνέχεια ήρθε μια σχετική ανακάλυψη χρησιμοποιώντας βαθιά δίκτυα για αναγνώριση ομιλίας όπου, για πρώτη φορά, ένα μοντέλο νευρωνικών δικτύων θεωρήθηκε ως η τελευταία λέξη της τεχνολογίας.

Ωστόσο, καθώς τα νευρωνικά δίκτυα άρχισαν να ξεπερνούν τις παραδοσιακές μεθόδους, αυτό έγινε ολοένα και λιγότερο αληθές. Τα πράγματα θα φτάσουν τελικά στο προσκήνιο το 2012 με την Large Scale Visual Recognition Challenge (LSVRC). Το 2009 δημιουργήθηκε μια μεγάλη βάση δεδομένων γνωστή ως Imagenet[5] που περιείχε εκατομμύρια ετικετές εικόνων και δημοσιεύθηκε από την

ομάδα της Fei-Fei Li στο Stanford. Αυτή η βάση δεδομένων συνδυάστηκε με το ετήσιο LSVRC, όπου οι διαγωνιζόμενοι θα κατασκεύαζαν μοντέλα computer vision, θα υπέβαλαν τις προβλέψεις τους και θα έλαβαν βαθμολογία με βάση το πόσο ακριβείς ήταν. Στα πρώτα δύο χρόνια του διαγωνισμού, τα κορυφαία μοντέλα είχαν ποσοστά σφάλματος 28% και 26%. Το 2012, ο Alex Krizhevsky, ο Ilya Sutskever και ο Geoff Hinton εισήγαγαν μια πρόταση που θα μείωνε το ποσοστό σφάλματος στο 16%. Το μοντέλο συνδυάζει αρκετά κρίσιμα στοιχεία που θα μπορούσαν να γίνουν μία βάση σε μελλοντικά μοντέλα βαθιάς μάθησης.

Πιθανώς το πιο σημαντικό κομμάτι ήταν η χρήση των graphics processing units (GPUs) για την εκπαίδευση του μοντέλου. Οι μονάδες GPU είναι ουσιαστικά παράλληλοι υπολογιστές κινητής υποδιαστολής με εκατοντάδες έως χιλιάδες πυρήνες. Η επιτάχυνση που προσέφεραν οι GPUs σήμαινε ότι θα μπορούσαν να εκπαιδεύσουν μεγαλύτερα μοντέλα, τα οποία οδήγησαν σε χαμηλότερα ποσοστά σφάλματος. Εισηγάγαν επίσης μια μέθοδο για να μειώσουν την υπερφόρτωση που είναι γνωστή ως dropout και χρησιμοποίησαν την rectified linear activation unit (ReLU), και τα δύο είναι πλέον βασικά μέρη στη σύγχρονη βαθιά μάθηση. Το δίκτυο άρχισε να γίνεται γνωστό ως "Alexnet"[6] και το έγγραφο που το περιγράφει αναφέρθηκε σχεδόν 25.000 φορές από τη δημοσίευσή του στο NIPS το 2012. Παρακάτω υπάρχει ένα γράφημα που δείχνει το ποσοστό σφάλματος στην Imagenet με την πάροδο του χρόνου, σημειώστε τη δραματική πτώση το 2012 και την τεράστια άνοδο στις ομάδες που χρησιμοποιούν GPUs:



Εικόνα 1.2 - Χρήση GPUs στο Imagenet και δείκτης σφάλματος των νικητών

Πολλές σημαντικές καινοτομίες θα έρθουν μετά από αυτό το αποτέλεσμα, αλλά νομίζω ότι είναι ασφαλές να πούμε ότι αυτή είναι η στιγμή που είδαμε τις πραγματικές δυνατότητες της βαθιάς μάθησης. Πολλοί είναι πεπεισμένοι ότι βρισκόμαστε στη μέση μίας μεγάλης επανάστασης στον χώρο της τεχνητής νοημοσύνης και δισεκατομμύρια δαπανώνται από εταιρείες στην έρευνα και ανάπτυξη αυτής της τεχνολογίας. Είναι δύσκολο να ξέρει κανείς τι επιφυλάσσει το μέλλον και σε τι βαθμό η βαθιά μάθηση θα παίξει ρόλο, αλλά προς το παρόν, φαίνεται σαν να βιώσαμε μια αλλαγή στην ιδεολογική δομή της μηχανικής μάθησης που είναι εδώ για να μείνει.

1.2 Νευρωνικό δίκτυο (Neural Network)

1.2.1 Ορισμός και λειτουργία

Το νευρωνικό δίκτυο είναι ένα δίκτυο από απλούς υπολογιστικούς κόμβους (νευρώνες, νευρώνια), διασυνδεδεμένους μεταξύ τους. Είναι εμπνευσμένο από το Κεντρικό Νευρικό Σύστημα (ΚΝΣ), το οποίο προσπαθεί να προσομοιώσει.

Οι νευρώνες είναι τα δομικά στοιχεία του δικτύου. Κάθε τέτοιος κόμβος δέχεται ένα σύνολο αριθμητικών εισόδων από διαφορετικές πηγές (είτε από άλλους νευρώνες, είτε από το περιβάλλον), επιτελεί έναν υπολογισμό με βάση αυτές τις εισόδους και παράγει μία έξοδο. Η εν λόγω έξοδος είτε κατευθύνεται στο περιβάλλον, είτε τροφοδοτείται ως είσοδος σε άλλους νευρώνες του δικτύου. Υπάρχουν τρεις τύποι νευρώνων: οι νευρώνες εισόδου, οι νευρώνες εξόδου και οι υπολογιστικοί νευρώνες ή κρυμμένοι νευρώνες. Οι νευρώνες εισόδου δεν επιτελούν κανέναν υπολογισμό, μεσολαβούν απλώς ανάμεσα στις περιβαλλοντικές εισόδους του δικτύου και στους υπολογιστικούς νευρώνες. Οι νευρώνες εξόδου διοχετεύουν στο περιβάλλον τις τελικές αριθμητικές εξόδους του δικτύου. Οι υπολογιστικοί νευρώνες πολλαπλασιάζουν κάθε είσοδό τους με το αντίστοιχο συναπτικό βάρος και υπολογίζουν το ολικό άθροισμα των γινομένων. Το άθροισμα αυτό τροφοδοτείται ως όρισμα στη συνάρτηση ενεργοποίησης, την οποία υλοποιεί εσωτερικά κάθε κόμβος. Η τιμή που λαμβάνει η συνάρτηση για το εν λόγω όρισμα είναι και η έξοδος του νευρώνα για τις τρέχουσες εισόδους και βάρη.

Εάν X_{ki} είναι η i -οστή είσοδος του k νευρώνα, W_{ki} το i -οστό συναπτικό βάρος του k νευρώνα και $\Phi(\cdot)$ η συνάρτηση ενεργοποίησης του νευρωνικού δικτύου, τότε η έξοδος Y_k του k νευρώνα δίνεται από την εξίσωση:

$$Y_k = \Phi\left(\sum_{i=0}^N X_{ki} W_{ki}\right) \quad (1.1)$$

Στον k -οστό νευρώνα υπάρχει ένα συναπτικό βάρος W_{k0} με ιδιαίτερη σημασία, το οποίο καλείται πόλωση ή κατώφλι (bias, threshold). Η τιμή της εισόδου του είναι πάντα η μονάδα $X_{k0} = 1$. Εάν το συνολικό άθροισμα από τις υπόλοιπες εισόδους του νευρώνα είναι μεγαλύτερο από την τιμή αυτή, τότε ο νευρώνας ενεργοποιείται. Εάν

είναι μικρότερο, τότε ο νευρώνας παραμένει ανενεργός. Η ιδέα προέκυψε από τα βιολογικά νευρικά κύτταρα.

Όπως είναι φανερό, οι αριθμοί οι οποίοι συναποτελούν το διάνυσμα εισόδου (κάθε στοιχείο του διανύσματος τροφοδοτείται κατά τη λειτουργία του δικτύου σε έναν νευρώνα εισόδου), αλλά και οι αριθμοί οι οποίοι συναποτελούν το διάνυσμα εξόδου (κάθε στοιχείο του οποίου εμφανίζεται, μετά το πέρας του ολικού υπολογισμού, σε έναν νευρώνα εξόδου), περιγράφουν χαρακτηριστικά του προς επίλυση προβλήματος. Συνήθως αυτό που μας ενδιαφέρει είναι το δίκτυο να απεικονίζει με ορθό τρόπο διανύσματα εισόδου σε κατάλληλα διανύσματα εξόδου, το πρόβλημα δηλαδή είναι η υλοποίηση μίας συνάρτησης πολλαπλών μεταβλητών, κατά κανόνα περίπλοκης και με άγνωστο ακριβή τύπο. Τέτοιες απεικονίσεις έχουν εφαρμογή σε ποικιλία τομέων της επιστήμης και της τεχνολογίας, αφού λειτουργούν ως αριθμητικά μοντέλα για πολλά διαφορετικά ζητήματα. Το ίδιο δίκτυο μπορεί να υλοποιήσει άπειρες διαφορετικές απεικονίσεις, μία για κάθε διαφορετική επιλογή συνόλου συναπτικών βαρών.

Το κύριο χαρακτηριστικό των νευρωνικών δικτύων είναι η εγγενής ικανότητα μάθησης. Ως μάθηση μπορεί να οριστεί η σταδιακή βελτίωση της ικανότητας του δικτύου να επιλύει κάποιο πρόβλημα (π.χ. η σταδιακή προσέγγιση μίας συνάρτησης). Η μάθηση επιτυγχάνεται μέσω της εκπαίδευσης, μίας επαναληπτικής διαδικασίας σταδιακής προσαρμογής των παραμέτρων του δικτύου (συνήθως των βαρών και της πόλωσής του) σε τιμές κατάλληλες ώστε να επιλύεται με επαρκή επιτυχία το προς εξέταση πρόβλημα. Αφού ένα δίκτυο εκπαιδευτεί, οι παράμετροί του συνήθως «παγώνουν» στις κατάλληλες τιμές και από εκεί κι έπειτα είναι σε λειτουργική κατάσταση. Το ζητούμενο είναι το λειτουργικό δίκτυο να χαρακτηρίζεται από μία ικανότητα γενίκευσης: αυτό σημαίνει πως δίνει ορθές εξόδους για εισόδους καινοφανείς και διαφορετικές από αυτές με τις οποίες εκπαιδεύτηκε

1.2.2 Συναρτήσεις ενεργοποίησης

Η συνάρτηση ενεργοποίησης μπορεί να είναι βηματική (step transfer function), γραμμική (linear transfer function), μη γραμμική (non-linear transfer function), στοχαστική (stochastic transfer function).

Βηματική συνάρτηση ενεργοποίησης

Η βηματική συνάρτηση ενεργοποίησης μπορεί να είναι:

$$\Phi(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x \leq 0 \end{cases} \quad (1.2)$$

ή οποιαδήποτε άλλη βηματική συνάρτηση.

Η βηματική συνάρτηση δεν θεωρείται χρήσιμη ως συνάρτηση ενεργοποίησης στα ΤΝΔ (Τεχνητά Νευρωνικά Δίκτυα), καθώς σύμφωνα με τον απειροστικό λογισμό έχει το βασικό μειονέκτημα ότι η παράγωγός της απειρίζεται. Έτσι προέκυψε η ανάγκη συναρτήσεων ενεργοποίησης που η γραφική τους παράσταση τους να μοιάζει με τη βηματική, αλλά ταυτόχρονα να είναι συνεχείς και παραγωγίσιμες σε όλο το πεδίο ορισμού τους. Τέτοια συνάρτηση είναι και η σιγμοειδής.

Γραμμική συνάρτηση ενεργοποίησης

Η γραμμική συνάρτηση ενεργοποίησης μπορεί να είναι:

$$\Phi(x) = x \quad (1.3)$$

ή οποιαδήποτε άλλη γραμμική συνάρτηση.

Μη γραμμική συνάρτηση ενεργοποίησης

Η μη γραμμική συνάρτηση ενεργοποίησης που χρησιμοποιείται συνήθως στα νευρωνικά δίκτυα καλείται σιγμοειδής συνάρτηση. Οι τυπικές σιγμοειδείς είναι δύο:

Λογιστική σιγμοειδής:

$$\Phi(x) = \frac{1}{1+e^x} \quad (1.4)$$

Υπερβολική εφαπτομένη:

$$\Phi(x) = \tanh x \quad (1.5)$$

1.2.3 Ιδιότητες

Τα τελευταία χρόνια έχει υπάρξει μία έκρηξη ενδιαφέροντος για τα νευρωνικά δίκτυα καθώς εφαρμόζονται με μεγάλη επιτυχία σε ένα ασυνήθιστα μεγάλο φάσμα τομέων της επιστήμης και της τεχνολογίας, όπως τα χρηματοοικονομικά, η ιατρική, η επιστήμη μηχανικού, η γεωλογία, η φυσική, η ρομποτική, η επεξεργασία σήματος κτλ. Στην πραγματικότητα, τα νευρωνικά δίκτυα εισάγονται οπουδήποτε τίθεται θέμα πρόβλεψης, ταξινόμησης ή ελέγχου. Η σαρωτική αυτή επιτυχία, μπορεί να αποδοθεί σε δύο βασικά στοιχεία: την ισχύ και την ευχρηστία.

Ισχύς: Τα νευρωνικά δίκτυα είναι πολύ εξελιγμένες τεχνικές μη γραμμικής μοντελοποίησης, ικανές να μοντελοποιήσουν εξαιρετικά πολύπλοκες λειτουργίες. Η γραμμική μοντελοποίηση υπήρξε ευρέως διαδεδομένη για πολύ καιρό, δεδομένου ότι στα γραμμικά μοντέλα εφαρμόζονται πολύ γνωστές στρατηγικές βελτιστοποίησης. Στις συνήθεις, όμως, περιπτώσεις όπου η γραμμική προσέγγιση δεν ήταν έγκυρη, τα μοντέλα αυτά αποτύγχαναν αναλόγως. Τα νευρωνικά δίκτυα βέβαια, αν και επιτρέπουν τη μη γραμμικότητα μέσω χρήσης μη γραμμικών συναρτήσεων ενεργοποίησης, μεταθέτουν με τη σειρά τους το πρόβλημα στο ζήτημα της διάστασης (του πλήθους των διαφορετικών εισόδων και εξόδων), το οποίο αποτελεί αγκάθι στις προσπάθειες μοντελοποίησης μη γραμμικών συναρτήσεων με μεγάλο αριθμό μεταβλητών.

Ευχρηστία: Τα νευρωνικά δίκτυα εκπαιδεύονται με παραδείγματα. Ο χρήστης συγκεντρώνει αντιπροσωπευτικά δεδομένα και στη συνέχεια, καθώς τα τροφοδοτεί συστηματικά στο δίκτυο μέσω των κατάλληλων αλγορίθμων εκπαίδευσης, το δίκτυο «αντιλαμβάνεται» αυτομάτως τη δομή των δεδομένων και η «γνώση» αυτή εκφράζεται ως κατάλληλες επιλογές συναπτικών βαρών. Επομένως το τελικό αποτέλεσμα της εκπαίδευσης με ένα συγκεκριμένο σύνολο παραδειγμάτων είναι ο προσδιορισμός των κατάλληλων βαρών του δικτύου. Ο χρήστης χρειάζεται να έχει

κάποιες ουσιώδεις γνώσεις σχετικά με τον τρόπο επιλογής και προετοιμασίας των δεδομένων, τον τρόπο εκλογής του κατάλληλου νευρωνικού δικτύου και στο πως θα ερμηνευτούν τα αποτελέσματα. Παρά ταύτα, το επίπεδο των γνώσεων του χρήστη που απαιτούνται για μια επιτυχημένη εφαρμογή των νευρωνικών δικτύων, είναι πολύ χαμηλότερο συγκριτικά με κάποια περίπτωση που θα χρησιμοποιούνταν ορισμένες πιο παραδοσιακές, μη γραμμικές στατιστικές μέθοδοι.

1.2.4 Εφαρμογές

Τα νευρωνικά δίκτυα είναι εφαρμόσιμα σχεδόν σε κάθε κατάσταση στην οποία ισχύει μια σχέση μεταξύ μεταβλητών πρόβλεψης (ανεξάρτητες, εισροές) και προβλεπόμενες μεταβλητές (εξαρτημένες, εκροές), ακόμα και όταν αυτή η σχέση είναι πολύ περίπλοκη για να αποδοθεί με τους συνηθισμένους όρους της «συσχέτισης» ή των «διαφόρων ομάδων». Ενδεικτικά αντιπροσωπευτικά παραδείγματα προβλημάτων στα οποία η ανάλυση των νευρωνικών δικτύων έχει εφαρμοστεί με επιτυχία είναι τα εξής:

Ιατρική διάγνωση: Ένα ευρύ φάσμα ιατρικά συσχετιζόμενων ενδείξεων, όπως ο συνδυασμός της καρδιακής συχνότητας, τα επίπεδα των διαφόρων ουσιών στο αίμα, ο ρυθμός της αναπνοής μπορούν να παρακολουθηθούν. Η εκδήλωση μιας συγκεκριμένης ιατρικής κατάστασης, γίνεται να συσχετιστεί με ένα πολύπλοκο συνδυασμό μεταβολών σε ένα υποσύνολο μεταβλητών που παρακολουθούνται. Τα νευρωνικά δίκτυα έχουν χρησιμοποιηθεί για την αναγνώριση αυτού του προτύπου πρόβλεψης, ώστε να χορηγηθεί η κατάλληλη θεραπεία.

Χρηματιστηριακές προβλέψεις: Οι διακυμάνσεις των τιμών των μετοχών και των χρηματιστηριακών δεικτών είναι ακόμα ένα παράδειγμα ενός πολύπλοκου, πολυδιάστατου, αλλά και σε ορισμένες περιπτώσεις εν μέρει ντετερμινιστικού φαινομένου. Τα νευρωνικά δίκτυα χρησιμοποιούνται από πολλούς τεχνικούς αναλυτές, ώστε να κάνουν προβλέψεις σχετικά με τις τιμές των μετοχών, βασιζόμενοι σε ένα μεγάλο αριθμό παραγόντων, όπως δηλαδή, τις προηγούμενες επιδόσεις άλλων αποθεμάτων και διαφόρων οικονομικών δεικτών.

Πιστωτική ανάθεση: Μια ποικιλία από κομμάτια πληροφοριών, τα οποία είναι συνήθως γνωστά για ένα απαιτούμενο δάνειο. Για παράδειγμα, η ηλικία του αιτούντος, η εκπαίδευση, το επάγγελμα και πολλά άλλα στοιχεία που μπορεί να είναι διαθέσιμα. Μετά την εκπαίδευση ενός νευρωνικού δικτύου σε ιστορικά δεδομένα η ανάλυση μπορεί να εκτοπίσει τα πιο κατάλληλα και σχετικά χαρακτηριστικά και να τα χρησιμοποιήσει για την ταξινόμηση των αιτούντων ως χαμηλού ή υψηλού κινδύνου.

Παρακολούθηση της κατάστασης των μηχανημάτων: Τα νευρωνικά δίκτυα μπορούν να συμβάλλουν στη μείωση του κόστους με την εξασφάλιση της πρόσθετης εμπειρογνωμοσύνης για τον προγραμματισμό προληπτικής συντήρησης των μηχανημάτων. Ένα νευρωνικό δίκτυο, λοιπόν, μπορεί να εκπαιδευτεί με τέτοιο τρόπο, ώστε να διακρίνει από τους ήχους τους οποίους παράγει μια μηχανή είτε αν εκτελεί κανονικά τις λειτουργίες της, είτε βρίσκεται στα πρόθυρα εμφάνισης οποιασδήποτε δυσλειτουργίας. Μετά από αυτήν την περίοδο εκπαιδευτικής κατάρτισης, η εμπειρία του ίδιου δικτύου είναι δυνατό να χρησιμοποιηθεί με σκοπό την προειδοποίηση ενός τεχνικού για κάποια επικείμενη βλάβη προτού συμβεί και ενδεχομένως προκαλέσει πολυδάπανες και απρόβλεπτες χρονικές καθυστερήσεις.

Συστήματα διαχείρισης κινητήρα: Τα νευρωνικά δίκτυα έχουν χρησιμοποιηθεί για την ανάλυση των εισροών που δέχονται οι αισθητήρες ενός κινητήρα. Το νευρωνικό δίκτυο ελέγχει μια ποικιλία παραμέτρων με τις οποίες λειτουργεί ο κινητήρας, προκειμένου να επιτευχθεί ένας συγκεκριμένος στόχος. Για παράδειγμα, το δίκτυο αυτό επιχειρεί την ελαχιστοποίηση της κατανάλωσης των καυσίμων.

1.2.4 Εκπαίδευση Νευρωνικών Δικτύων

Μια από τις πιο βασικές ιδιότητες των Νευρωνικών Δικτύων είναι η ικανότητά τους για εκπαίδευση. Η εκπαίδευση αυτή επιτυγχάνεται μέσω της ανταλλαγής τιμών και βαρών, που αποσκοπεί στη βαθμιαία σύλληψη της πληροφορίας η οποία στη συνέχεια θα είναι διαθέσιμη προς ανάκτηση. Υπάρχουν, βέβαια, πολλοί αλγόριθμοι που η εφαρμογή τους έχει στόχο την προσαρμογή των

τιμών των βαρών ενός Τεχνητού Νευρωνικού Δικτύου. Όλες οι μέθοδοι μάθησης μπορούν να καταταχτούν σε δύο κατηγορίες : τη μάθηση με επίβλεψη (supervised learning) και τη μάθηση χωρίς επίβλεψη (unsupervised learning).

Μάθηση με επίβλεψη: Η μάθηση αυτή είναι μια διαδικασία η οποία συνδυάζει έναν εξωτερικό εκπαιδευτή και τη συνολική ή γενικευμένη πληροφορία. Κάποιες από τις μεθόδους οι οποίες συγκαταλέγονται σε αυτή την κατηγορία είναι η μάθηση με διόρθωση σφάλματος, η στοχαστική μάθηση. Παραδείγματα τα οποία αντιπροσωπεύουν την μάθηση με επίβλεψη συμπεριλαμβάνουν αποφάσεις για το πότε θα πρέπει να σταματήσει η διαδικασία εκπαίδευσης, αποφάσεις αναφορικά με τη συχνότητα παρουσίασης στο δίκτυο τα πρότυπα εκπαίδευσης και η παρουσίαση προόδου του δικτύου. Η μάθηση με επίβλεψη χωρίζεται σε δύο ακόμα κατηγορίες: στη δομική (structural) και στην προσωρινή (temporal) εκμάθηση. Οι αλγόριθμοι οι οποίοι βρίσκονται στην πρώτη κατηγορία, χρησιμοποιούνται για την εύρεση της βέλτιστης σχέσης μεταξύ εισόδων και εξόδων για κάθε ξεχωριστό ζευγάρι προτύπων. Παραδείγματα της δομικής εκμάθησης αποτελούν η αναγνώριση και η κατηγοριοποίηση προτύπων, ενώ παραδείγματα της προσωρινής εκμάθησης η πρόβλεψη και ο έλεγχος.

Μάθηση χωρίς επίβλεψη: Οι αλγόριθμοι της εν λόγω μάθησης αναφέρονται ως αυτό-οργανώμενοι (self-organized) και είναι διαδικασίες οι οποίες δεν απαιτούν να είναι παρών ένας «εξωτερικός» δάσκαλος ή επιβλέπων. Βασίζονται, μάλιστα, μόνο σε τοπική πληροφορία καθ' όλη τη διάρκεια της εκπαίδευσης του Τεχνητού Νευρωνικού Δικτύου. Οι συγκεκριμένοι αλγόριθμοι οργανώνουν τα δεδομένα και ανακαλύπτουν τις σημαντικές συλλογικές ιδιότητες. Για παράδειγμα, αλγόριθμοι εκπαίδευσης χωρίς επίβλεψη είναι ο αλγόριθμος Hebbian, ο διαφορικός αλγόριθμος Hebbian και ο Min-Max αλγόριθμος.

Κατά κύριο λόγο οι περισσότερες διαδικασίες εκπαίδευσης είναι off line. Όταν χρησιμοποιείται όλο το δείγμα προτύπων για την τροποποίηση των τιμών των βαρών, πριν την τελική χρήση του δικτύου ως εφαρμογή, τότε ονομάζεται off line εκπαίδευση. Οι αλγόριθμοι εκπαίδευσης off line έχουν την απαίτηση να βρίσκονται στην εκπαίδευση του δικτύου παρόντα όλα τα πρότυπα. Το γεγονός αυτό αποκλείει

την πιθανότητα εισαγωγής νέων πληροφοριών μέσω νέων προτύπων. Βέβαια, υπάρχουν και Τεχνητά Νευρωνικά Δίκτυα τα οποία δεν αποκλείουν την εισαγωγή νέας πληροφορίας, μετά την τελική τους μοντελοποίηση. Αν παρουσιαστεί ανάγκη εισαγωγής νέου προτύπου στο δίκτυο, μπορεί να γίνει απευθείας χωρίς τον κίνδυνο να χαθεί κανένα μέρος της αρχικής πληροφορίας. Το πλεονέκτημα των δικτύων που χρησιμοποιούν off line διαδικασίες εκπαίδευσης επικεντρώνεται κυρίως στη δυνατότητα να δίνουν καλύτερες λύσεις σε δύσκολα προβλήματα

Κεφάλαιο 2^ο Μέσα Υλοποίησης Πτυχιακής

2.1 Python

Η Python είναι μια υψηλού επιπέδου γλώσσα προγραμματισμού η οποία δημιουργήθηκε από τον Ολλανδό Γκίντο βαν Ρόσσομ (Guido van Rossum) το 1990. Ο κύριος στόχος της είναι η αναγνωσιμότητα του κώδικά της και η ευκολία χρήσης της. Το συντακτικό της επιτρέπει στους προγραμματιστές να εκφράσουν έννοιες σε λιγότερες γραμμές κώδικα απ' ό τι θα ήταν δυνατόν σε γλώσσες όπως η C++ ή η Java. Διακρίνεται λόγω του ότι έχει πολλές βιβλιοθήκες που διευκολύνουν ιδιαίτερα αρκετές συνηθισμένες εργασίες και για την ταχύτητα εκμάθησής της.

Οι διερμηνευτές της Python είναι διαθέσιμοι για εγκατάσταση σε πολλά λειτουργικά συστήματα, επιτρέποντας στην Python την εκτέλεση κώδικα σε ευρεία γκάμα συστημάτων. Χρησιμοποιώντας εργαλεία τρίτων, όπως το Py2exe ή το Pyinstaller, ο κώδικας της Python μπορεί να πακεταριστεί σε αυτόνομα εκτελέσιμα προγράμματα για μερικά από τα πιο δημοφιλή λειτουργικά συστήματα, επιτρέποντας τη διανομή του βασισμένου σε Python λογισμικού για χρήση σε αυτά τα περιβάλλοντα χωρίς να απαιτείται εγκατάσταση του διερμηνευτή της Python.

Η Python αναπτύσσεται ως ανοιχτό λογισμικό (open source) και η διαχείρισή της γίνεται από τον μη κερδοσκοπικό οργανισμό Python Software Foundation. Ο κώδικας διανέμεται με την άδεια Python Software Foundation License η οποία είναι συμβατή με την GPL. Το όνομα της γλώσσας προέρχεται από την ομάδα άγγλων κωμικών Μόντυ Πάιθον.

Αρχικά, η Python ήταν γλώσσα σεναρίων που χρησιμοποιούνταν στο λειτουργικό σύστημα Amoeba, ικανή και για κλήσεις συστήματος.

Η Python 2.0 κυκλοφόρησε στις 16 Οκτωβρίου του 2000. Στις 3 Δεκεμβρίου 2008 κυκλοφόρησε η έκδοση 3.0 (γνωστή και ως py3k ή python 3000). Πολλά από τα καινούργια χαρακτηριστικά αυτής της έκδοσης έχουν μεταφερθεί στις εκδόσεις 2.6 και 2.7 που είναι προς τα πίσω συμβατές.

Η Python 3.0 είναι ιστορικά η πρώτη γλώσσα προγραμματισμού που σπάει την προς τα πίσω συμβατότητα με προηγούμενες εκδόσεις ώστε να διορθωθούν κάποια λάθη που υπήρχαν σε προγενέστερες εκδόσεις και να καταστεί ακόμα πιο σαφής ο απλός τρόπος με τον οποίο μπορούν να γίνουν κάποια πράγματα.

Για τη συγγραφή προγραμμάτων είναι απαραίτητος ένας κειμενογράφος ή ακόμα καλύτερα ένα ολοκληρωμένο περιβάλλον ανάπτυξης (Integrated Development Environment - IDE), το οποίο είναι ένα ειδικό λογισμικό για την ανάπτυξη εφαρμογών. Η Python έρχεται μαζί με ένα εύχρηστο περιβάλλον ανάπτυξης με την ονομασία IDLE. Τα αρχικά του έρχονται από τις λέξεις Interactive DeveLopment Environment και είναι γραμμένο σε Python από τον Guido van Rossum. Χρησιμοποιεί τη βιβλιοθήκη γραφικών Tkinter, οπότε μπορεί να εκτελεσθεί σε περιβάλλον Linux, Windows και Mac OS X. Το IDLE μας δίνει τη δυνατότητα να χρησιμοποιήσουμε διαδραστικά τον διερμηνευτή της γλώσσας, να γράψουμε και να επεξεργαστούμε προγράμματα, να τα αποθηκεύσουμε σε αρχεία, να τα εκτελέσουμε, να κάνουμε αποσφαλμάτωση.

Η γλώσσα χρησιμοποιεί μεταγλωττιστή (compiler) για την δημιουργία του εκτελέσιμου κώδικα και σχετίζεται με τις γλώσσες προγραμματισμού Tcl, Perl, Scheme, Java και Ruby, καθώς και με την ABC η οποία υπήρξε η αρχική πηγή έμπνευσης για τη δημιουργία της.

Ένα ιδιαίτερο χαρακτηριστικό της γλώσσας είναι η χρήση κενών διαστημάτων (whitespace) για τον διαχωρισμό των συντακτικών δομών που προγράμματος, σε αντίθεση με την πρακτική σε άλλες γλώσσες όπου για τον ίδιο σκοπό χρησιμοποιούνται ειδικά σύμβολα (πχ αγκύλες). Αυτό, σε συνδυασμό με το ότι χρησιμοποιεί πλήρεις αγγλικές λέξεις στη θέση συμβόλων, καθιστούν τον κώδικα της Python ευανάγνωστο από όσους έχουν βασική γνώση των αγγλικών.

2.1.1 Jupyter Notebook

Το Jupyter Notebook (Παλαιότερα IPython Notebooks) είναι ένα διαδικτυακό υπολογιστικό περιβάλλον για τη δημιουργία εγγράφων Jupyter notebooks. Ο όρος "σημειωματάριο" στην καθομιλουμένη, μπορεί να κάνει αναφορές σε πολλές διαφορετικές οντότητες, κυρίως την εφαρμογή Ιστού Jupyter, το web server Jupyter rython ή τη μορφή εγγράφου Jupyter ανάλογα με το πλαίσιο. Ένα έγγραφο Jupyter Notebook είναι ένα έγγραφο JSON, το οποίο περιέχει μια λίστα κατανεμημένων κελιών εισόδου / εξόδου τα οποία μπορούν να περιέχουν κώδικα, κείμενο (με Markdown), μαθηματικά, σχέδια και εμπλουτισμένες διαφημίσεις, συνήθως τελειώνοντας με το ".ipynb" ως επέκταση.

Το έγγραφο Jupyter notebooks μπορεί να μετατραπεί σε μια σειρά ανοιχτών τυποποιημένων μορφών εξόδου (HTML, παρουσίαση διαφανειών, LaTeX, PDF, ReStructuredText, Markdown, Python) μέσω του 'Download As' στην διεπαφή ιστού μέσω της βιβλιοθήκης nbconvert ή στην γραμμή εντολών μέσω της εντολής 'jupyter nbconvert'.

Για να απλοποιήσει την απεικόνιση των εγγράφων σημειωματάρων Jupyter στον ιστό, η βιβλιοθήκη nbconvert παρέχεται ως υπηρεσία μέσω του NbViewer, η οποία μπορεί να πάρει μια διεύθυνση URL από οποιοδήποτε, διαθέσιμο στο κοινό, έγγραφο φορητού υπολογιστή, να το μετατρέψει σε αρχείο HTML και να το εμφανίσει στον χρήστη.

Το Jupyter Notebook παρέχει ένα REPL με βάση το πρόγραμμα περιήγησης το οποίο είναι δομημένο με μια σειρά από δημοφιλείς βιβλιοθήκες ανοιχτού κώδικα:

- IPython
- ØMQ
- Tornado (διακομιστής ιστού)
- jQuery
- Bootstrap (front-end πλαίσιο)
- MathJax

Το Jupyter Notebook μπορεί να συνδεθεί σε πολλούς πυρήνες, (από προεπιλογή το Jupyter Notebook συνοδεύεται από τον πυρήνα IPython) για να επιτρέψει τον προγραμματισμό σε πολλές γλώσσες. Από την έκδοση 2.3 (Οκτώβριος 2014), υπάρχουν διαθέσιμοι 49 πυρήνες συμβατοί με Jupyter για πολλές γλώσσες προγραμματισμού, συμπεριλαμβανομένων των Python, R, Julia και Haskell.

Η διασύνδεση του Notebook προστέθηκε στο IPython στην έκδοση 0.12 (Δεκέμβριος 2011), μετονομασμένη σε Jupyter Notebook το 2015 (IPython 4.0 - Jupyter 1.0). Το Jupyter Notebook είναι παρόμοιο με το περιβάλλον εργασίας άλλων προγραμμάτων όπως το Maple, το Mathematica και το SageMath, ένα υπολογιστικό στυλ διεπαφής που δημιουργήθηκε με το Mathematica στη δεκαετία του 1980 . Σύμφωνα με το The Atlantic[7], το ενδιαφέρον του Jupyter ξεπέρασε τη δημοτικότητα της διασύνδεσης του Notebook Mathematica στις αρχές του 2018.

Ο πυρήνας Jupyter είναι ένα πρόγραμμα που είναι υπεύθυνο για τον χειρισμό διαφόρων τύπων αιτημάτων (εκτέλεση κώδικα, ολοκλήρωση κώδικα, επιθεώρηση) και παροχή απάντησης. Οι πυρήνες μιλούν με τα υπόλοιπα στοιχεία του Jupyter χρησιμοποιώντας το ZeroMQ μέσω του δικτύου και έτσι μπορούν να είναι στην ίδια ή σε απομακρυσμένες μηχανές. Σε αντίθεση με πολλές άλλες φορητές διεπαφές, στο Jupyter, οι πυρήνες δεν γνωρίζουν ότι είναι συνδεδεμένοι με ένα συγκεκριμένο έγγραφο και μπορούν να συνδεθούν από πολλούς πελάτες ταυτόχρονα. Συνήθως οι πυρήνες υλοποιούνται και επιτρέπουν την εκτέλεση μιας μόνο γλώσσας με μερικές εξαιρέσεις.

Από προεπιλογή, ο Jupyter συμπεριλαμβάνει το IPython ως προεπιλεγμένο πυρήνα. Υπάρχουν πυρήνες διαφόρων ποιοτήτων και χαρακτηριστικών για πολλές γλώσσες.

Το JupyterHub είναι ένας διακομιστής πολλαπλών χρηστών για τα Jupyter Notebooks. Έχει σχεδιαστεί για να υποστηρίζει πολλούς χρήστες μέσω της αναπαραγωγής, της διαχείρισης και την λειτουργία του ως διακομιστή μεσολάβησης πολλών μοναδικών διακομιστών Jupyter Notebook.

Το JupyterLab είναι η διεπαφή χρήστη επόμενης γενιάς για το Project Jupyter. Προσφέρει όλα τα γνωστά δομικά στοιχεία του κλασικού Jupyter Notebook (notebook, τερματικό, πρόγραμμα επεξεργασίας κειμένου, πρόγραμμα περιήγησης αρχείων, πλούσιες εξόδους κ.λπ.) σε ένα ευέλικτο και ισχυρό περιβάλλον εργασίας χρήστη.

Το Jupyter Notebook έχει γίνει μία δημοφιλής διεπαφή χρήστη για το Cloud computing και οι σημαντικότεροι προμηθευτές cloud έχουν υιοθετήσει τα εργαλεία Jupyter Notebook ή παράγωγα εργαλεία ως διεπαφή frontend για χρήστες cloud.

2.1.3 Απαραίτητες βιβλιοθήκες

Καθ' όλη την πορεία της εργασίας, διαφορές βιβλιοθήκες, κατέστησαν απαραίτητες για την επεξεργασία των δεδομένων.

Οι σημαντικότερες αυτών αναφέρονται παρακάτω:

Το ερευνητικό σχέδιο Pandas: Στον προγραμματισμό υπολογιστών, το pandas είναι μια βιβλιοθήκη λογισμικού γραμμένη για τη γλώσσα προγραμματισμού Python για χειρισμό και ανάλυση δεδομένων. Συγκεκριμένα, προσφέρει δομές δεδομένων και λειτουργίες για τον χειρισμό αριθμητικών πινάκων και χρονοσειρών. Είναι ελεύθερο λογισμικό που εκδίδεται βάσει της άδειας BSD με τρεις ρήτρες. Το όνομα προέρχεται από τον όρο "δεδομένα πίνακα" (panel data), έναν οικονομετρικό όρο για σύνολα δεδομένων που περιλαμβάνουν παρατηρήσεις για πολλαπλές χρονικές περιόδους για τα ίδια άτομα.

Λειτουργίες βιβλιοθήκης:

- Το αντικείμενο DataFrame για χειρισμό δεδομένων με ενσωματωμένη ευρετηρίαση.
- Εργαλεία για ανάγνωση και εγγραφή δεδομένων μεταξύ δομών δεδομένων εντός μνήμης και διαφορετικών μορφών αρχείων.

- Ευθυγράμμιση δεδομένων και ολοκληρωμένος χειρισμός των δεδομένων που λείπουν.
- Επαναδιαμόρφωση και περιστροφή των συνόλων δεδομένων.
- Ετικετοποίηση με βάση το τεμαχισμό, φανταχτερή ευρετηρίαση και υποσύνολο μεγάλων συνόλων δεδομένων.
- Εισαγωγή και διαγραφή στήλης δομής δεδομένων.
- Ομαδοποίηση ανά μηχανή, επιτρέποντας τις λειτουργίες split-apply-combine σε σύνολα δεδομένων.
- Συγχώνευση και ένωση συνόλου δεδομένων.
- Ιεραρχικός άξονας ευρετηρίασης για εργασία με δεδομένα μεγάλης διαστάσεως σε δομή δεδομένων χαμηλότερων διαστάσεων.
- Λειτουργικότητα χρονοσειρών: Γενική χρονική περίοδος και μετατροπή συχνότητας, μετακίνηση στατιστικών παραθύρων, γραμμικές παλινδρομήσεις κινούμενων παραθύρων, μετατόπιση ημερομηνιών και υστέρηση.

Η βιβλιοθήκη είναι εξαιρετικά βελτιστοποιημένη για απόδοση, με κρίσιμες διαδρομές κώδικα γραμμένες σε Cython ή C.

Βιβλιοθήκη SciPy: Η βιβλιοθήκη SciPy είναι ένα από τα βασικά πακέτα που αποτελούν τη στοίβα SciPy. Παρέχει πολλές φιλικές προς το χρήστη και αποτελεσματικές αριθμητικές ρουτίνες, όπως ρουτίνες για αριθμητική ολοκλήρωση και βελτιστοποίηση.

Βιβλιοθήκη σχεδίασης Matplotlib: Το Matplotlib είναι μια βιβλιοθήκη σχεδίασης για τη γλώσσα προγραμματισμού Python και την αριθμητική μαθηματική επέκταση NumPy. Παρέχει ένα αντικειμενοστρεφές API για την ενσωμάτωση διαγραμμάτων σε εφαρμογές που χρησιμοποιούν γενικές GUI εργαλειομηχανές όπως Tkinter, wxPython, Qt ή GTK+. Υπάρχει επίσης μια διαδικαστική διασύνδεση "pylab"

βασισμένη σε μηχανές καταστάσεων (όπως το OpenGL), σχεδιασμένη να μοιάζει πολύ με εκείνο του MATLAB, αν και αποθαρρύνεται η χρήση του. Το SciPy κάνει χρήση του matplotlib. Το Pyplot είναι μια μονάδα matplotlib που παρέχει μια διεπαφή τύπου MATLAB. Το Matplotlib έχει σχεδιαστεί έτσι ώστε να είναι τόσο χρησιμοποιήσιμο όσο το MATLAB, με τη δυνατότητα να χρησιμοποιεί την Python και το πλεονέκτημα του να είναι ελεύθερο και ανοιχτό λογισμικό.

Το πακέτο NumPy: Το NumPy είναι το θεμελιώδες πακέτο για πληροφορική ερευνά με την Python. Εκτός από τις προφανείς επιστημονικές του χρήσεις, το NumPy μπορεί επίσης να χρησιμοποιηθεί ως αποτελεσματικό πολυδιάστατο δοχείο γενικών δεδομένων. Μπορούν να οριστούν αυθαίρετοι τύποι δεδομένων. Αυτό επιτρέπει στην NumPy να ενσωματώνεται απρόσκοπτα και γρήγορα με μια μεγάλη ποικιλία βάσεων δεδομένων. Το NumPy είναι εγκεκριμένο βάσει της άδειας BSD, επιτρέποντας την επαναχρησιμοποίηση με λίγους περιορισμούς.

Περιέχει μεταξύ άλλων:

- ένα ισχυρό αντικείμενο N-διάστατης συστοιχίας
- εξελιγμένες λειτουργίες (εκπομπής)
- εργαλεία για την ενσωμάτωση κώδικα C / C ++ και Fortran
- χρήσιμη γραμμική άλγεβρα, μετασχηματισμό Fourier και δυνατότητες τυχαίου αριθμού

Το πακέτο Pydicom: Το Pydicom είναι ένα πακέτο Python για εργασία με αρχεία DICOM όπως ιατρικές εικόνες, αναφορές και αντικείμενα ακτινοθεραπείας. Επίσης, διευκολύνει την ανάγνωση αυτών των σύνθετων αρχείων σε φυσικές, για την Python, δομές για εύκολο χειρισμό. Τροποποιημένα σύνολα δεδομένων μπορούν να συνταχθούν ξανά σε αρχεία μορφής DICOM.

2.2 TensorFlow

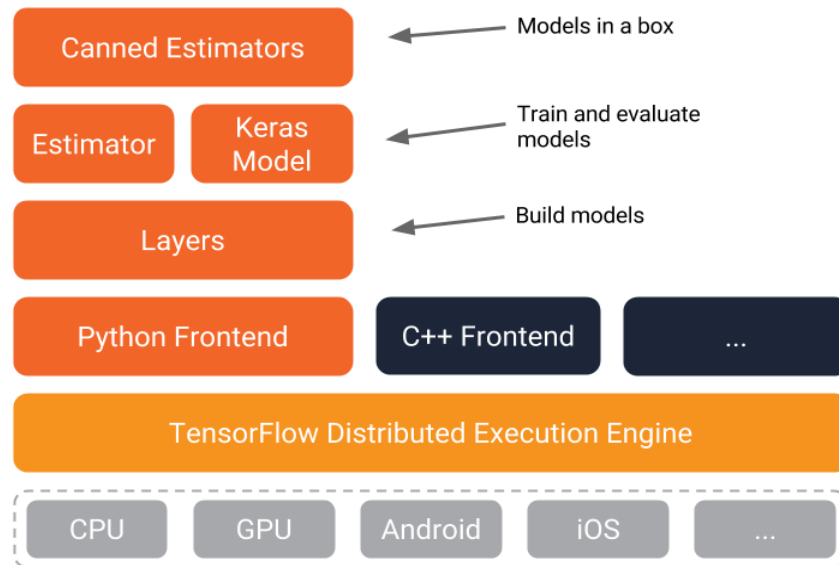
Το TensorFlow είναι μια βιβλιοθήκη λογισμικού ανοιχτού κώδικα για προγραμματισμό ροής δεδομένων σε μια σειρά εργασιών. Είναι μια συμβολική βιβλιοθήκη μαθηματικών και χρησιμοποιείται επίσης για εφαρμογές μηχανικής μάθησης όπως νευρωνικά δίκτυα. Χρησιμοποιείται τόσο για την έρευνα όσο και για την ανάπτυξη στην Google[8] για μηχανική μάθηση και έρευνα για τα βαθιά νευρωνικά δίκτυα, αντικαθιστώντας συχνά τον κλειστού κώδικα προκάτοχό, DistBelief. Το σύστημα είναι αρκετά γενικό ώστε να μπορεί να εφαρμοστεί σε ευρύ φάσμα άλλων τομέων. Έφθασε την έκδοση 1.0 τον Φεβρουάριο του 2017 και συνέχισε την ταχεία ανάπτυξη, με 21.000 συνεισφορές έως τώρα, πολλές από εξωτερικούς συνεργάτες. Απελευθερώθηκε υπό την άδεια ανοικτού κώδικα Apache 2.0 στις 9 Νοεμβρίου 2015.[9]

Το TensorFlow είναι cross-platform. Τρέχει σχεδόν στα πάντα: σε μονάδες GPU και CPU, συμπεριλαμβανομένων των κινητών και ενσωματωμένων πλατφορμών και ακόμη μονάδες επεξεργασίας tensor (TPU), οι οποίες είναι εξειδικευμένο υλικό για την εκτέλεση μαθηματικών με βάση τους τανυστές. Δεν διατίθενται ακόμη ευρέως, αλλά πρόσφατα ξεκίνησαν ένα πρόγραμμα alpha.

Η κατανεμημένη μηχανή εκτέλεσης TensorFlow αφαιρεί τις πολλές υποστηριζόμενες συσκευές και παρέχει έναν υψηλής απόδοσης πυρήνα που εφαρμόζεται στην C++ για την πλατφόρμα TensorFlow.

Πάνω από αυτό κάθονται τα μετωπιαία άκρα των Python και C++. Το Layer API παρέχει μια απλούστερη διεπαφή για τα χρησιμοποιούμενα στρώματα σε μοντέλα βαθιάς μάθησης. Επάνω σε αυτό κάθονται τα API υψηλότερου επιπέδου, συμπεριλαμβανομένου του Keras και του API Estimator, το οποίο κάνει ευκολότερη την κατάρτιση και την αξιολόγηση των κατανεμημένων μοντέλων.

Και τέλος, μια σειρά από κοινά χρησιμοποιούμενα μοντέλα είναι έτοιμα για χρήση, με περισσότερα να έρθουν στο μέλλον.



Εικόνα 2.1 - Πλατφόρμα πληροφοριών TensorFlow

2.2.1 Το ξεκίνημα

Ξεκινώντας το 2011, η ομάδα του Google Brain δημιούργησε το DistBelief ως ιδιόκτητο σύστημα εκμάθησης μηχανών που βασίζεται σε νευρωνικά δίκτυα εκπαίδευσης βαθιάς μάθησης. Η χρήση του αυξήθηκε ταχύτατα σε ποικίλες εταιρείες της πολυεθνικής Alphabet τόσο σε ερευνητικές όσο και σε εμπορικές εφαρμογές. Η Google ανέθεσε σε πολλούς επιστήμονες πληροφορικής, συμπεριλαμβανομένου του Jeff Dean, να απλοποιήσουν και να επαναπροσδιορίσουν τον βασικό κώδικα του DistBelief σε μια ταχύτερη και πιο στιβαρή βιβλιοθήκη εφαρμογών, δημιουργώντας έτσι το TensorFlow.[10] Το 2009, η ομάδα, με επικεφαλής τον Geoffrey Hinton, είχε εφαρμόσει γενική οπισθοδιάδοση και άλλες βελτιώσεις που επέτρεψαν τη δημιουργία νευρωνικών δικτύων με σημαντικά υψηλότερη ακρίβεια, για παράδειγμα μείωση κατά 25% των σφαλμάτων στην αναγνώριση ομιλίας.[11]

2.2.2 Η δημιουργία του TensorFlow

Το TensorFlow είναι το σύστημα δεύτερης γενιάς του Google Brain. Η έκδοση 1.0.0 κυκλοφόρησε στις 11 Φεβρουαρίου 2017. Ενώ η εφαρμογή αναφοράς λειτουργεί σε μεμονωμένες συσκευές, το TensorFlow μπορεί να λειτουργεί σε πολλαπλές CPU και GPU (με προαιρετικές επεκτάσεις CUDA και SYCL για μονάδες επεξεργασίας γραφικών σε υπολογιστές γενικής χρήσης). Το TensorFlow διατίθεται σε πλατφόρμες υπολογιστών Linux, macOS, Windows και κινητής τηλεφωνίας 64-bit, συμπεριλαμβανομένων των Android και iOS.

Οι υπολογισμοί TensorFlow εκφράζονται ως γραφικά σταδιακής ροής δεδομένων. Το όνομα TensorFlow προέρχεται από τις λειτουργίες που τέτοιου είδους νευρωνικά δίκτυα εκτελούν σε πολυδιάστατες συστοιχίες δεδομένων. Αυτές οι συστοιχίες αναφέρονται ως "τανυστές". Τον Ιούνιο του 2016, ο Dean δήλωσε ότι 1.500 αποθετήρια στο GitHub ανέφεραν το TensorFlow, εκ των οποίων μόνο 5 ήταν από την Google.

2.2.3 Μονάδα επεξεργασίας Tensor

Τον Μάιο του 2016, η Google ανακοίνωσε τη μονάδα επεξεργασίας Tensor (TPU), ένα ASIC που κατασκευάστηκε ειδικά για μηχανική μάθηση και προσαρμόστηκε για το TensorFlow. Το TPU είναι ένας προγραμματιζόμενος επιταχυντής AI που έχει σχεδιαστεί για να παρέχει υψηλή διεκπεραιωτή ικανότητα σε χαμηλής ακρίβειας αριθμητική (π.χ. 8-bit) και να είναι προσανατολισμένη προς τη χρήση ή τη λειτουργία μοντέλων αντί να τα εκπαιδεύει. Η Google ανήγγειλε ότι χρησιμοποιούσαν TPU μέσα στα κέντρα δεδομένων τους για περισσότερο από ένα χρόνο με αποτέλεσμα να έχουν μία πολύ πιο βελτιωμένη απόδοση ανά watt για μηχανική μάθηση.[12]

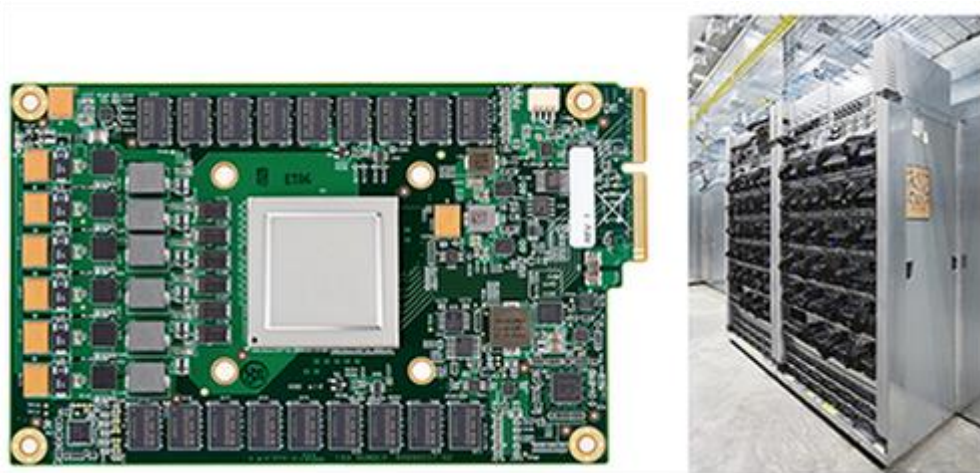
2.2.5 TFLearn

Το TFLearn είναι μια δομοστοιχειωτή και διαφανής βιβλιοθήκη βαθιάς μάθησης που είναι χτισμένη στην κορυφή του TensorFlow. Σχεδιάστηκε για να παρέχει API υψηλότερου επιπέδου στο TensorFlow, προκειμένου να διευκολύνει και να επιταχύνει τις πειραματικές διαδικασίες, παραμένοντας όμως πλήρως διαφανές και συμβατό με αυτό.

Τα χαρακτηριστικά του TFLearn περιλαμβάνουν:

- Εύκολο στη χρήση και κατανόηση API υψηλού επιπέδου για την υλοποίηση βαθιών νευρωνικών δικτύων, με διδακτικό υλικό και παραδείγματα.
- Γρήγορα πρωτότυπα μοντέλα μέσω των εξαιρετικά δομοστοιχειωτών ενσωματωμένων επιπέδων νευρωνικού δικτύου, ρυθμιστές, βελτιστοποιητές, μετρήσεις κ.α.
- Πλήρης διαφάνεια έναντι του TensorFlow. Όλες οι λειτουργίες είναι χτισμένες πάνω στους τανυστήρες και μπορούν να χρησιμοποιηθούν ανεξάρτητα από το TFLearn.
- Ισχυρές βοηθητικές λειτουργίες για την εκπαίδευση οποιουδήποτε γραφήματος TensorFlow, με υποστήριξη πολλαπλών εισόδων, εξόδων και βελτιστοποιητών.
- Εύκολη και όμορφη απεικόνιση γραφημάτων, με λεπτομέρειες για βάρη, κλίσεις, ενεργοποιήσεις και πολλά άλλα
- Εύκολη τοποθέτηση συσκευών για χρήση πολλαπλών CPU / GPU.

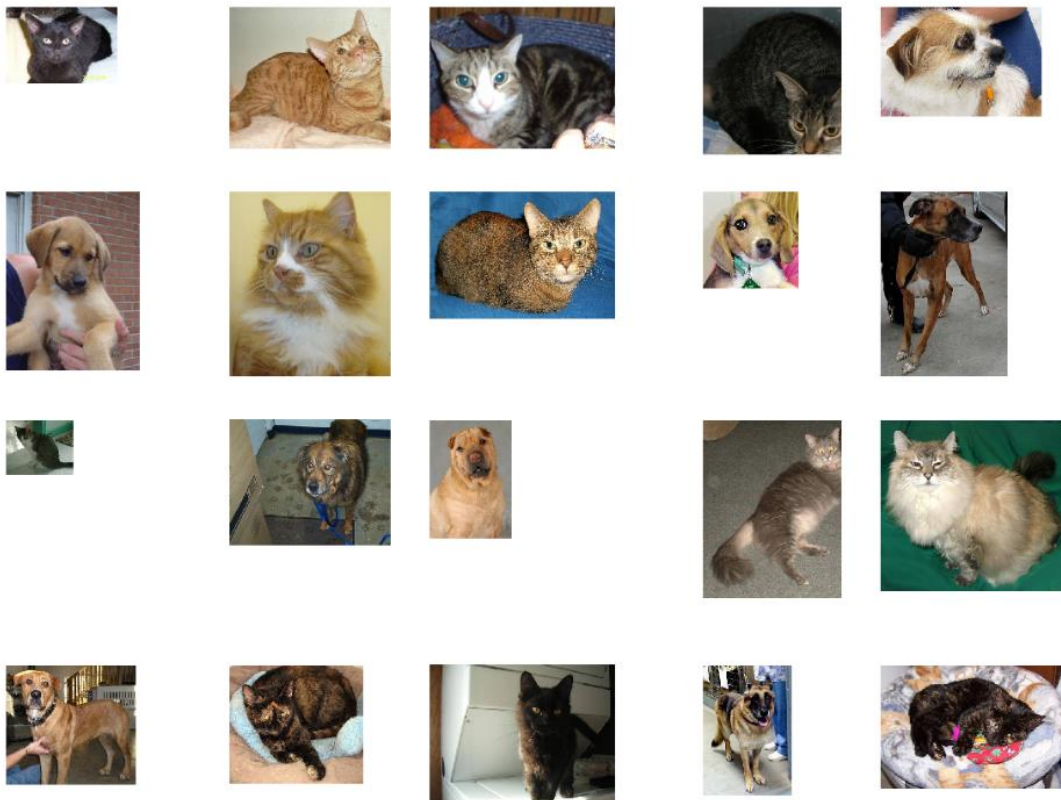
Το API υψηλού επιπέδου υποστηρίζει επί του παρόντος τα περισσότερα από τα πρόσφατα μοντέλα βαθιάς μάθησης, όπως Convolutions, LSTM, BiRNN, BatchNorm, PReLU, Residual networks και Generative networks. Στο μέλλον, το TFLearn προορίζεται επίσης να παραμείνει επίκαιρο με τις τελευταίες τεχνικές βαθιάς μάθησης.



Εικόνα 2.3 - - Η πρώτη μονάδα επεξεργασίας Tensor (TPU) της Google σε μια πλακέτα τυπωμένου κυκλώματος (αριστερά). TPU που χρησιμοποιείται σε κέντρο δεδομένων της Google (δεξιά)

2.2.6 Παράδειγμα Dogs vs Cats

Δεδομένα: Τα δεδομένα που χρησιμοποιήσαμε πάρθηκαν από τον διαγωνισμό «Dogs vs. Cats Redux: Kernels Edition»[15]. Το σύνολο των δεδομένων αποτελείται από 25000 φωτογραφίες, σε μορφή .jpg, που απεικονίζουν είτε σκύλους είτε γάτες. Κάθε φωτογραφία περιέχει έναν αύξοντα αριθμό καθώς και την ετικέτα 'dog' για την περίπτωση που είναι σκύλος και 'cat' για αν είναι γάτα.



Εικόνα 2.4 - Δείγμα του συνόλου δεδομένων

Μοντέλο: Στα δεδομένα έγινε αλλαγή μεγέθους σε 150x150 και στην συνέχεια τροφοδοτήθηκαν σε συνελκτικό νευρωνικό δίκτυο αποτελούμενο από έξι επίπεδα συνέλιξης με συνάρτηση ενεργοποίηση ReLU και την απόσυρση (dropout) κατά τη διάρκεια της εκπαίδευσης. Επίσης, χρησιμοποιεί σταθμισμένη απώλεια διασταυρούμενης εντροπίας για συνάρτηση ενεργοποίησης softmax και του βελτιστοποιητή Adam.

Κώδικας περιγραφής του μοντέλου:

```
>> import tflearn
>> from tflearn.layers.conv import conv_2d, max_pool_2d
>> from tflearn.layers.core import input_data, dropout,
fully_connected
>> from tflearn.layers.estimator import regression

>> convnet = input_data(shape=[None, IMG_SIZE, IMG_SIZE, 1],
name='input')

>> convnet = conv_2d(convnet, 32, 5, activation='relu')
>> convnet = max_pool_2d(convnet, 5)

>> convnet = conv_2d(convnet, 64, 5, activation='relu')
>> convnet = max_pool_2d(convnet, 5)

>> convnet = conv_2d(convnet, 32, 5, activation='relu')
>> convnet = max_pool_2d(convnet, 5)

>> convnet = conv_2d(convnet, 64, 5, activation='relu')
>> convnet = max_pool_2d(convnet, 5)

>> convnet = conv_2d(convnet, 32, 5, activation='relu')
>> convnet = max_pool_2d(convnet, 5)

>> convnet = conv_2d(convnet, 64, 5, activation='relu')
>> convnet = max_pool_2d(convnet, 5)

>> convnet = fully_connected(convnet, 1024, activation='relu')
>> convnet = dropout(convnet, 0.8)
```



```
>> convnet = fully_connected(convnet, 2, activation='softmax')
>> convnet = regression(convnet, optimizer='adam',
learning_rate=LR, loss='categorical_crossentropy', name='targets')
```

Πίνακας 2.1 - CNN αρχιτεκτονική 6 επιπέδων (dropout με 0.8 πιθανότητα μετά από κάθε conv επίπεδο κατά την εκπαίδευση, Adam βελτιστοποιητής με δείκτη εκπαίδευσης 0.001)

Layer	Parameters	Activation	Output
Input			150x150x1
Conv1	5x5	ReLU	150x150x3
Max Pool	5x5, stride 5		30x30x32
Conv2	5x5	ReLU	30x30x64
Max Pool	5x5, stride 5		6x6x64
Conv3	5x5	ReLU	6x6x32
Max Pool	5x5, stride 5		2x2x32
Conv4	5x5	ReLU	2x2x64
Max Pool	5x5, stride 5		1x1x64
Conv5	5x5	ReLU	1x1x32
Max Pool	5x5, stride 5		1x1x32
Conv6	5x5	ReLU	1x1x64
Max Pool	5x5, stride 5		1x1x64

Αποτελέσματα: Τα αποτελέσματα ήταν εξαιρετικά λαμβάνοντας υπόψη πως δεν χρησιμοποιήθηκαν εξωτερικά δεδομένα και προ-εκπαιδευμένα μοντέλα. Συγκεκριμένα, είχαμε ποσοστό επιτυχίας 83.37% με ακρίβεια 84% και μέσω των καμπυλών λειτουργικού χαρακτηριστικού δέκτη (ROC) βλέπουμε πως το εμβαδόν της περιοχής κάτω από την καμπύλη (AUC) είναι 0.842. Η πλήρη σημασία των παραπάνω χαρακτηριστικών θα αναλυθεί στο επόμενο κεφάλαιο.

Τέλος, αξίζει να σημειωθεί πως σε αντίθεση με τα ιατρικά δεδομένα του επόμενου κεφαλαίου, τα δεδομένα του παραδείγματος είναι δύο διαστάσεων και όχι τριών.



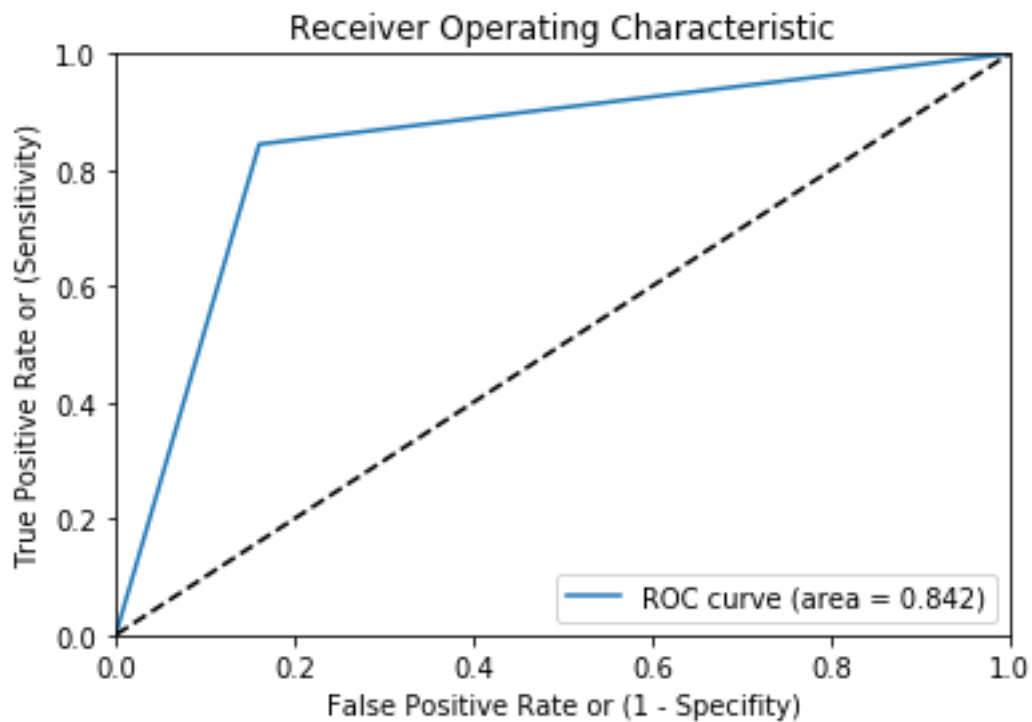
Εικόνα 2.6 - Οπτικός έλεγχος του δικτύου σε μη επισημασμένα δεδομένα

Πίνακας 2.2 - Πίνακας Σύγκρισης για CNN του παραδείγματος Dogs vs Cats

	Predicted	
	Normal	Abnormal
Actual		
Normal	0.839	0.161
Abnormal	0.156	0.844

Πίνακας 2.3 - Έκθεση κατηγοριοποίησης για CNN του παραδείγματος Dogs vs Cats

Classification report:				
	Precision	Recall	F1-Score	Support
0	0.84	0.84	0.84	12262
1	0.84	0.84	0.84	12238
Avg/Total	0.84	0.84	0.84	24500



Εικόνα 2.7 - ROC για CNN του παραδείγματος Dogs vs Cats

Εφαρμογή προπαιδευμένου μοντέλου AlexNet:

Το AlexNet είναι ένα συνελικτικό νευρικό δίκτυο, που εφευρέθηκε από τους Alex Krizhevsky, Ilya Sutskever και Geoffrey Hinton. Το AlexNet είχε μεγάλο αντίκτυπο στον τομέα της μηχανικής μάθησης, συγκεκριμένα στην εφαρμογή της βαθιάς μάθησης στη μηχανική όραση.

Το AlexNet συμμετείχε στον διαγωνισμό 'ImageNet Large Scale Visual Recognition Challenge' το 2012. Το δίκτυο πέτυχε Top-5 ποσοστό σφάλματος 15,3%. Το επόμενο καλύτερο αποτέλεσμα ήταν χαμηλότερο κατά 10,8%. Το κύριο σημείο του άρθρου ήταν ότι το βάθος του μοντέλου ήταν απαραίτητο για την υψηλή απόδοση του, η οποία ήταν υπολογιστικά δαπανηρή, αλλά έγινε εφικτή λόγω της χρήσης GPU κατά τη διάρκεια της εκπαίδευσης.

Επιλύει το πρόβλημα της ταξινόμησης εικόνων όπου, η είσοδος είναι μια εικόνα η οποία, μπορεί να ανήκει σε μια από τις 1000 διαφορετικές κατηγορίες (π.χ. γάτες, σκυλιά κλπ.) και η έξοδος είναι ένα διάνυσμα 1000 αριθμών. Το v -οστό στοιχείο του φορέα εξόδου ερμηνεύεται ως η πιθανότητα ότι η εικόνα εισόδου ανήκει στην v -τάξη. Επομένως, το άθροισμα όλων των στοιχείων του φορέα εξόδου είναι 1.

Η είσοδος στο AlexNet είναι μια εικόνα RGB μεγέθους 256×256 . Αυτό σημαίνει ότι όλες οι εικόνες στο σετ εκπαίδευσης και όλες οι εικόνες δοκιμής πρέπει να έχουν μέγεθος 256×256 .

Το AlexNet ήταν πολύ μεγαλύτερο από τα προηγούμενα CNN που χρησιμοποιήθηκαν για τα καθήκοντα μηχανικής όρασης. Έχει 60 εκατομμύρια παραμέτρους και 650.000 νευρώνες και χρειάστηκε πέντε με έξι ημέρες για να εκπαιδευτεί, χρησιμοποιώντας δύο GPU GTX 580 3GB.

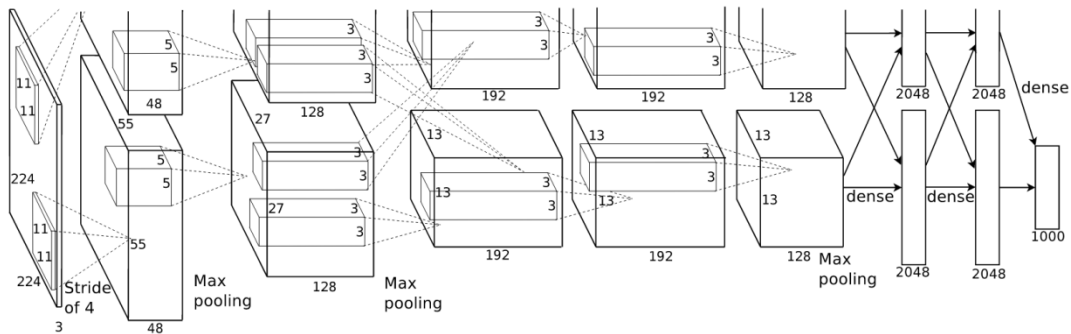
Το AlexNet αποτελείται από 5 επίπεδα συνέλιξης και 3 πλήρως συνδεδεμένα επίπεδα.

Πολλαπλοί συνελικτικοί πυρήνες (γνωστά επίσης ως φίλτρα) εξάγουν ενδιαφέροντα χαρακτηριστικά από μια εικόνα. Σε ένα ενιαίο στρώμα συνέλιξης, υπάρχουν

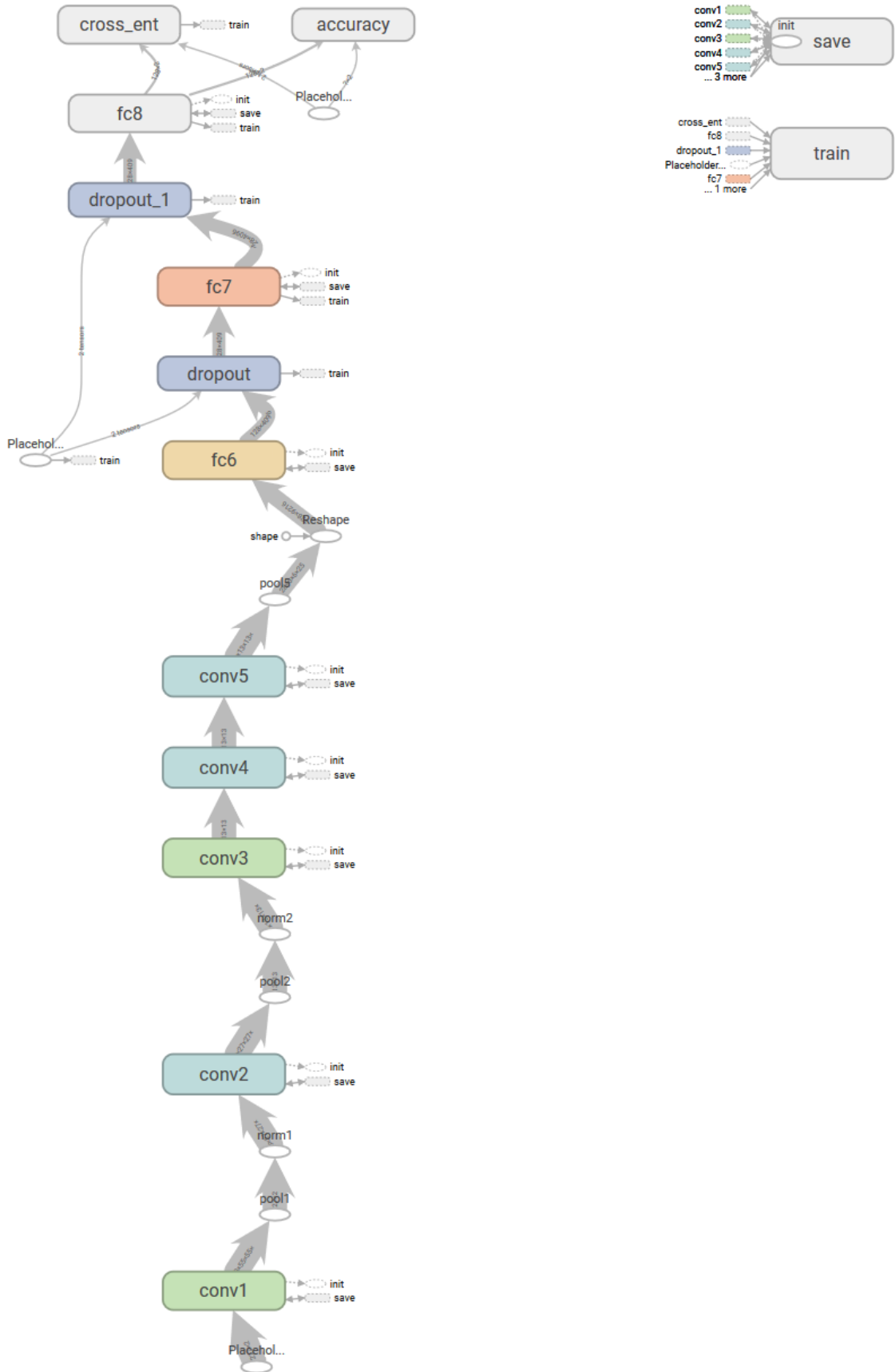
συνήθως πολλοί πυρήνες του ίδιου μεγέθους. Για παράδειγμα, το πρώτο επίπεδο Conv του AlexNet περιέχει 96 πυρήνες μεγέθους 11x11x3. Σημειώση ότι το πλάτος και το ύψος του πυρήνα είναι συνήθως ίδια και το βάθος είναι το ίδιο με τον αριθμό των καναλιών.

Τα πρώτα δύο στρώματα συνέλιξης ακολουθούνται από τα στρώματα Overlapping Max Pooling. Το τρίτο, τέταρτο και πέμπτο συνελκτικό στρώμα συνδέονται άμεσα. Το πέμπτο στρώμα συνέλιξης ακολουθείται από ένα στρώμα Overlapping Max Pooling, η έξοδος του οποίου, εισέρχεται σε μια σειρά δυο πλήρως συνδεδεμένων στρωμάτων. Το δεύτερο πλήρως συνδεδεμένο στρώμα τροφοδοτεί έναν ταξινομητή softmax με 1000 ετικέτες διαφορετικών κατηγοριών.

Η συνάρτηση ενεργοποίηση ReLU εφαρμόζεται μετά από όλα τα στρώματα συνέλιξης και πλήρως συνδεδεμένα στρώματα. Η ReLU του πρώτου και του δεύτερου στρώματος συνέλιξης ακολουθείται από ένα βήμα τοπικής εξομάλυνσης πριν από τη συγκέντρωση. Αλλά οι ερευνητές αργότερα δεν βρήκαν την εξομάλυνση πολύ χρήσιμη.

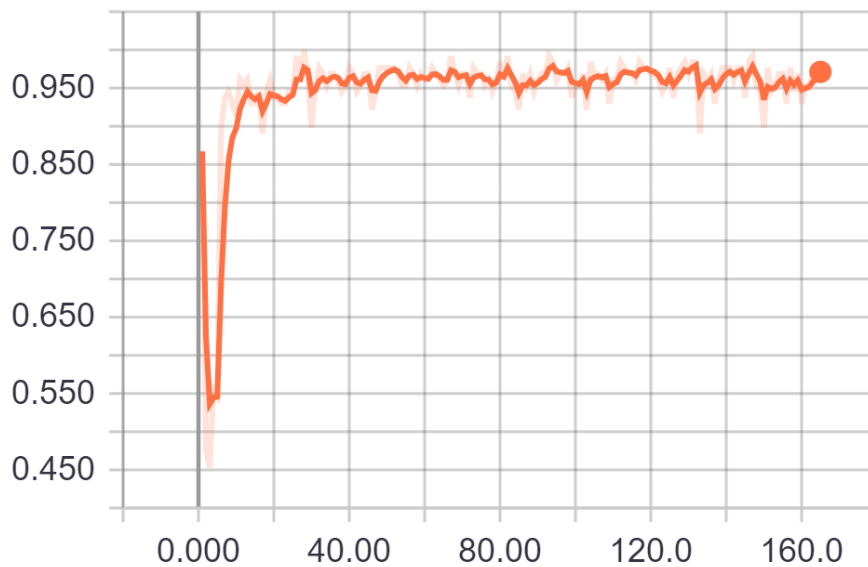


Εικόνα 2.8 - Αρχιτεκτονική μοντέλου AlexNet

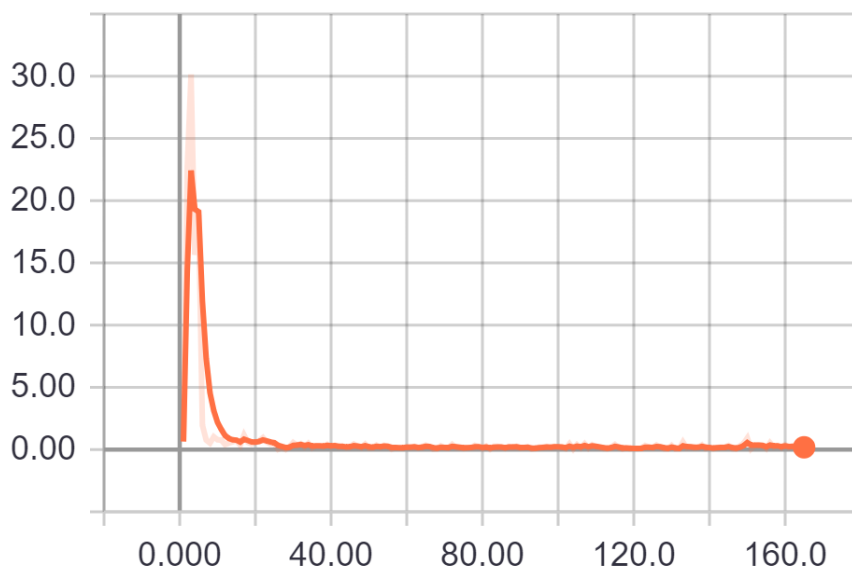


Εικόνα 2.9 - CNN αρχιτεκτονική του εφαρμοσμένου μοντέλου AlexNet

Αποτελέσματα: Τα αποτελέσματα της εφαρμογής προπαιδευμένου μοντέλου AlexNet είναι εξαιρετικά με ακρίβεια εκπαίδευσης 98.44% και σφάλμα 0.17. Στο δεδομένα επαλήθευσης είχαμε ποσοστό επιτυχίας 94.32%. Σύμφωνα με τα αποτελέσματα, το ποσοστό επιτυχίας στο κομμάτι της εκπαίδευσης είναι ελαφρώς μεγαλύτερο από αυτό της επαλήθευσης, που σημαίνει πως έχουμε μία μικρή Υπερπροσαρμογή (Overfitting), δηλαδή το μοντέλο έχει εξοικειωθεί σε μεγαλύτερο βαθμό με τα δεδομένα εκπαίδευσης και δεν αποδίδει το ίδιο περὰ αυτού. Τρόποι ελαχιστοποίησης της Υπερπροσαρμογής είναι η εκπαίδευση σε μεγαλύτερο εύρος δεδομένων, την μείωση των χαρακτηριστικών του δικτύου και την πρώιμη διακοπή της διαδικασίας εκπαίδευσης σε αριθμό επαναλήψεων όπου το σύστημα δεν έχει αρχίσει να γενικολογεί.



Εικόνα 2.10 - Καμπύλη Ευστοχίας (Accuracy)



Εικόνα 2.11 - Καμπύλη Σφάλματος (Loss)

Κεφάλαιο 3^ο Πρακτικό Μέρος

3.1 Εισαγωγή

Παγκοσμίως, ο καρκίνος του πνεύμονα είναι η πιο κοινή μορφή καρκίνου τόσο από άποψη συχνότητας όσο από άποψη θνησιμότητας. Το 2008, υπήρχαν 1,61 εκατομμύρια νέα κρούσματα, και 1,38 εκατομμύρια θάνατοι που οφείλονται στον καρκίνο του πνεύμονα. Τα υψηλότερα ποσοστά είναι σε Ευρώπη και τη Βόρεια Αμερική.[16] Το τμήμα του πληθυσμού που είναι πιθανότερο να αναπτύξει καρκίνο του πνεύμονα είναι οι άνθρωποι άνω των 50 ετών που έχουν ιστορικό καπνίσματος. Σε αντίθεση με το ποσοστό θνησιμότητας στους άνδρες, το οποίο άρχισε να μειώνεται πάνω από 20 χρόνια πριν, τα ποσοστά θνησιμότητας στις γυναίκες από καρκίνο των πνευμόνων έχουν αυξηθεί τις τελευταίες δεκαετίες, και μόλις πρόσφατα άρχισαν να σταθεροποιούνται. Στις ΗΠΑ, ο κίνδυνος ανάπτυξης καρκίνου του πνεύμονα είναι 8% στους άνδρες και 6% στις γυναίκες.[17] Η Ανατολική Ευρώπη έχει την υψηλότερη θνησιμότητα του καρκίνου των πνευμόνων μεταξύ των ανδρών, ενώ η βόρεια Ευρώπη και οι ΗΠΑ έχουν την υψηλότερη θνησιμότητα μεταξύ των γυναικών. Στις Ηνωμένες Πολιτείες, οι μαύροι άνδρες και γυναίκες έχουν την υψηλότερη συχνότητα.[18] Η συχνότητα του καρκίνου του πνεύμονα είναι σήμερα μικρότερη στις αναπτυσσόμενες χώρες.[19] Με την αύξηση του καπνίσματος στις αναπτυσσόμενες χώρες, η συχνότητα αναμένεται να αυξηθεί κατά τα επόμενα χρόνια, κυρίως στην Κίνα και την Ινδία.[20][21]

Η σταδιοποίηση του καρκίνου του πνεύμονα είναι η αξιολόγηση του βαθμού εξάπλωσης του καρκίνου από την αρχική του προέλευση. Είναι ένας από τους παράγοντες που επηρεάζουν την πρόγνωση και πιθανή θεραπεία του καρκίνου του πνεύμονα.[22] Τα στάδια 1 και 2 αναφέρονται σε καρκίνους που εντοπίζονται οι

πνεύμονες και τα μεταγενέστερα στάδια αναφέρονται σε καρκίνους που έχουν εξαπλωθεί και μετασταθεί σε άλλα όργανα. Οι τρέχουσες διαγνωστικές μέθοδοι περιλαμβάνουν βιοψίες και απεικόνισης, όπως η αξονική τομογραφία (CT). Πρόωρη ανίχνευση του καρκίνου του πνεύμονα (ανίχνευση κατά τα πρώιμα στάδια) βελτιώνει σημαντικά τις πιθανότητες επιβίωσης, αλλά είναι επίσης και πιο δύσκολος ο εντοπισμός νωρίς στα στάδια του καρκίνου του πνεύμονα, καθώς υπάρχουν λιγότερα συμπτώματα.

Σκοπός μας είναι η επίλυση ενός προβλήματος δυαδικής ταξινόμησης στην ανίχνευση της παρουσίας πρώιμου ή μη καρκίνου του πνεύμονα στις αξονικές τομογραφίες (CT) του θώρακα του ασθενή. Στόχος μας είναι η χρήση μεθόδων μηχανικής όρασης και βαθιάς μάθησης, συγκεκριμένα δισδιάστατων και τρισδιάστατων συνελκτικών νευρωνικών δικτύων, για την κατασκευή ενός ακριβούς ταξινομητή. Ένας ταξινομητής με υψηλή ακρίβεια στον εντοπισμό του καρκίνου του πνεύμονα θα μπορούσε να επιταχύνει και να μειώσει το κόστος της ανίχνευσης του καρκίνου, επιτρέποντας παράλληλα την διάδοση της έγκαιρης ανίχνευσής του καθώς και την βελτίωση των ποσοστών επιβίωσης των ασθενών. Πρόθεση μας είναι η δημιουργία ενός σύστημα υποβοηθούμενης διάγνωσης με χρήση υπολογιστή (CAD) το οποίο θα λαμβάνει ως είσοδο αξονικές τομογραφίες και θα εξάγει το αν ο ασθενής έχει ή όχι καρκίνο του πνεύμονα.

3.2 Συναφή Έργα

Η υποβοηθούμενη μέσω υπολογιστή διάγνωση (CAD) είναι σε θέση να βοηθήσει του γιατρούς στην κατανόηση των ιατρικών εικόνων, επιτρέποντας τη διάγνωση του καρκίνου με μεγαλύτερη ευαισθησία και ακρίβεια, το οποίο είναι κρίσιμο για του ασθενείς. Οι Demir και Yener[23] εξετάζουν την δομή των CAD, χωρίζοντας τα σε προεπεξεργασία, εξαγωγή χαρακτηριστικών, επιλογή και τέλος ταξινόμηση. Επιπλέον, τεκμηριώνουν τη χρήση της λογιστικής παλινδρόμησης (logistic regression), διαγράμματα αποφάσεων, κ- πλησιέστερου γείτονα και των νευρωνικών δικτύων σε υπάρχουσες προσεγγίσεις.

Ο Khalvati και η ομάδα του[24] χρησιμοποιούν ένα διάνυσμα υποστήριξης μηχανής (SVM) πάνω σε μαγνητικές τομογραφίες (MRI) για την ανίχνευση καρκίνου του προστάτη σε ασθενείς. Οι νικητές της πρόκλησης Camelyon16[25], για παράδειγμα, ανιχνεύουν τον καρκίνο του μαστού από εικόνες λεμφαδένων.

Τα βαθιά συνελκτικά νευρωνικά δίκτυα (CNN) αποδειχθήκαν ότι αποδίδουν καλά στην ταξινόμηση εικόνων[26][27], την ανίχνευση αντικειμένων[28] και σε άλλες οπτικές εργασίες. Έχουν βρει μεγάλες επιτυχίες σε εφαρμογές ιατρικής απεικόνισης[29] και είναι για παράδειγμα ικανές να ανιχνεύσουν μεταστάσεις του καρκίνου του δέρματος[30], επιτυγχάνοντας ουσιαστικά καλύτερη απόδοση στην ταχύτητα απόκρισης από ότι οι ανθρωπινοί παθολόγοι. Όλες αυτές οι μέθοδοι λειτουργούν σε δυσδιάστατες εικόνες, συνήθως σε μια διατομεακή εικόνα του μέρους του σώματος που έχει πληγεί.

Σε σύγκριση, οι αξονικές τομογραφίες (CT) είναι τρισδιάστατοι όγκοι και είναι συνήθως ανισότροποι. Επίσης, τα βαθιά δίκτυα έχουν δείξει ότι έχουν καλή απόδοση σε τρισδιάστατη τμηματοποίηση[31] και έχουν προσαρμοστεί με επιτυχία από δύο διαστάσεις σε τρεις διαστάσεις[32][33]. Ο Liao και η ομάδα του[34] παρουσίασαν ένα τρισδιάστατο πλαίσιο βαθιάς μάθησης για την αυτόματη κατάτμηση του προστάτη. Ο Kumar και η ομάδα του [35] εκτελούν χωριστά την κατάτμηση και την ταξινόμηση των οζιδίων του καρκίνου του πνεύμονα.

Η πρόκληση LUNA16[36] είχε δύο καθήκοντα: την ανίχνευση των πνευμονικών οζιδίων με την χρήση αξονικών τομογραφιών και τη μείωση του ψευδούς θετικού ρυθμού από την αναγνώριση αυτών των οζιδίων. Το πρώτο λύθηκε από την ομάδα της Moira[37] χρησιμοποιώντας το UNet[38] σε στοίβες των τριών διαδοχικών οριζόντιων κομματιών του πνεύμονα και το δεύτερο πραγματοποιήθηκε από την ομάδα της Dou[39] εφαρμόζοντας πολυκλαδικά (multi-contextual) τρισδιάστατα συνελκτικά νευρωνικά δίκτυα.

Οι δύο πρώτες ομάδες στο Kaggle Data Science Bowl 2017 έχουν δημοσιεύσει τις λύσεις τους στην πρόκληση. Και οι δύο ομάδες προχωρούν σε ένα ενδιάμεσο βήμα προτού δώσουν, στους ασθενείς, την πιθανότητα καρκίνου. Μετά την ταυτοποίηση των περιοχών με πιθανές ανωμαλίες (οζίδια), η ομάδα που

βρίσκεται στην δεύτερη θέση, του de Wit[40], χρησιμοποιεί 17 διαφορετικά τρισδιάστατα συνελκτικά νευρωνικά δίκτυα για να εξαγάγει ιατρικά συναφή χαρακτηριστικά για τα οζίδια. Στη συνέχεια, συγκεντρώνουν τις προβλέψεις αυτών των χαρακτηριστικών των οζιδίων σε έναν περιγραφέα επιπέδου ασθενούς. Η ομάδα που κατέκτησε την πρώτη θέση[41] ανιχνεύει τα οζίδια μέσω ενός τρισδιάστατου συνελκτικού νευρωνικού δικτύου, στη συνέχεια χρησιμοποιεί τις ανιχνεύσεις υψηλότερης εμπιστοσύνης καθώς και τη χειροκίνητη επισήμανση των οζιδίων για να προβλέψει τον καρκίνο μέσω ενός απλού ταξινομητή.

3.3 Ιατρικά Δεδομένα

Τα δεδομένα που χρησιμοποιήσαμε πάρθηκαν από τον διαγωνισμό «Kaggle Data Science Bowl 2017»[42]. Το σύνολο των δεδομένων αποτελείται από 2101 αξονικές τομογραφίες (CT) στη θωρακική κοιλότητα των ασθενών. Από τις 2101, οι 1595 εκδόθηκαν αρχικά στο πρώτο στάδιο του διαγωνισμού, με τις 1397 να ανήκουν στα δεδομένα εκπαίδευσης και 198 στα δεδομένα έλεγχου. Τα υπόλοιπα 506 εκδοθήκαν στο δεύτερο στάδιο του διαγωνισμού ως τα τελευταία δεδομένα έλεγχου.

Κάθε αξονική τομογραφία περιείχε την ετικέτα ‘with cancer’ αν ο ασθενής που συσχετιζόταν με αυτή είχε διαγνωστεί με καρκίνο εντός ενός χρόνου από την στιγμή της εξέτασης και σε διαφορετική περίπτωση είχε την ετικέτα ‘without cancer’. Δυστυχώς, η θέση ή το μέγεθος των καρκινικών οζιδίων δεν είχαν ετικέτα.

3.4 Σχεδιασμός και Υλοποίηση

Η αρχική μας προσέγγιση ήταν να τροφοδοτούμε τα 3D CNNs με τις ανεπεξέργαστες τρισδιάστατες αξονικές τομογραφίες αλλά τα αποτελέσματα ήταν ανεπαρκή κυρίως λόγω του μεγάλου όγκου των δεδομένων και της περιορισμένης

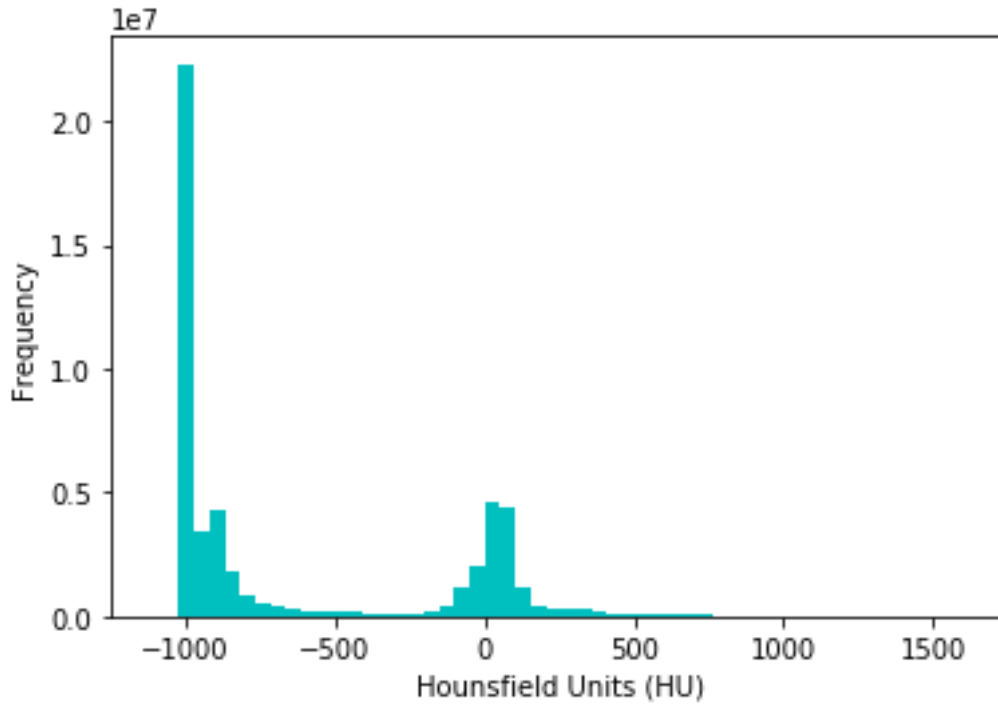
υπολογιστικής ισχύς του εξοπλισμού μας, έτσι χρειάστηκε να εφαρμόσουμε μηχανισμούς προεπεξεργασίας των αρχικών δεδομένων για τον καθαρισμό τους από διαφόρους θορύβους καθώς και την μετατροπή τους σε πιο προσιτές μορφές επεξεργασίας από τα 3D CNNs. Αρχικά, εφαρμόσαμε τεχνικές ομαλοποίησης, δειγματοληψίας, αλλαγής μεγέθους τους και τέλος την τοποθέτηση με κέντρο το μηδέν (zero-center) ώστε η μέση τιμή των δεδομένων, να είναι 0'.

3.4.1 Προεπεξεργασία των δεδομένων

Η μονάδα μέτρησης στις αξονικές τομογραφίες είναι το Hounsfield Unit (HU), το οποίο είναι ένα μέτρο ραδιοπυκνότητας. Οι αξονικοί τομογράφοι είναι προσεκτικά ρυθμισμένοι στην μέτρηση αυτής.

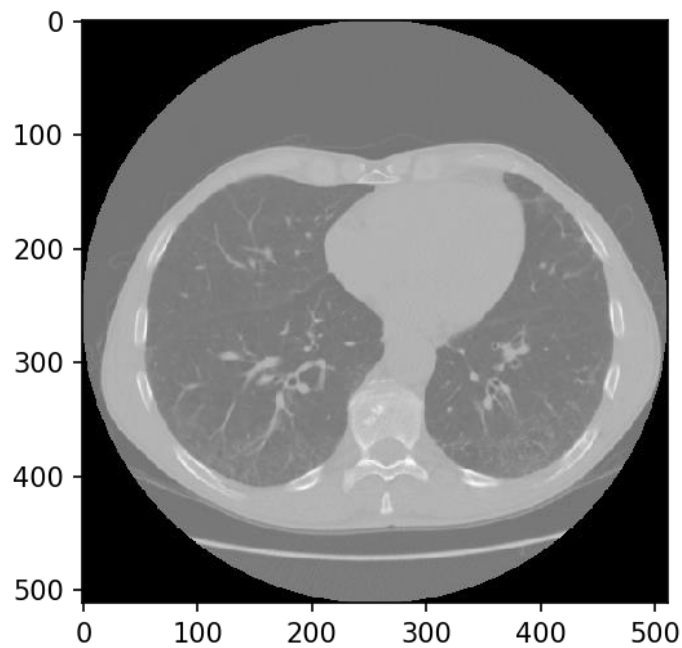
Πίνακας 3.1 - Κλίμακα Hounsfield

Ουσία	HU
Αέρας	-1000
Πνεύμονας	-500
Λίπος	-100 έως -50
Νερό	0
Εγκεφαλονωτιαίο υγρό (CSF)	15
Νεφρός	30
Αίμα	+30 έως +45
Μύες	+10 έως +40
Φαία ουσία	+37 έως +45
Λευκή ουσία	+20 έως +30
Ήπαρ	+40 έως +60
Μαλακός ιστός, Αντίθεση	+100 έως +300
Οστά	+700 (σπογγώδες οστό) έως +3000 (φλοιώδες οστό)



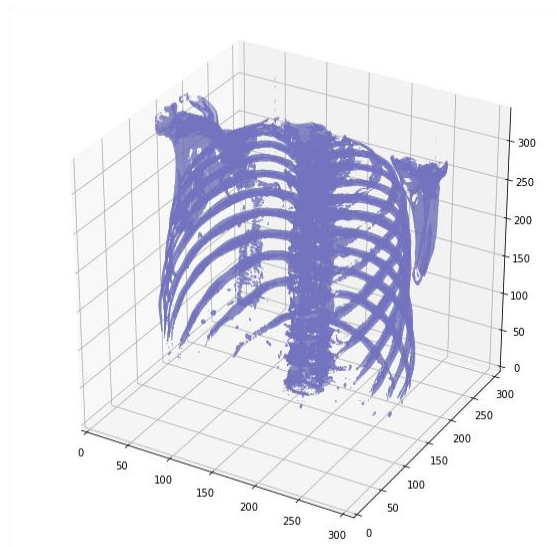
Εικόνα 3.1 - Ιστογράμματα των τιμών των εικονοστοιχείων σε μονάδα Hounsfield

Κάποιοι τομογράφοι έχουν κυλινδρικά όρια σάρωσης αλλά η τελική εικόνα είναι τετράγωνη. Τα εικονοστοιχεία που βρίσκονται εκτός των ορίων λαμβάνουν την προκαθορισμένη τιμή -2000. Το πρώτο βήμα είναι ο ορισμός αυτών σε μηδέν, το οποίο επί του παρόντος αντιστοιχεί σε αέρα.

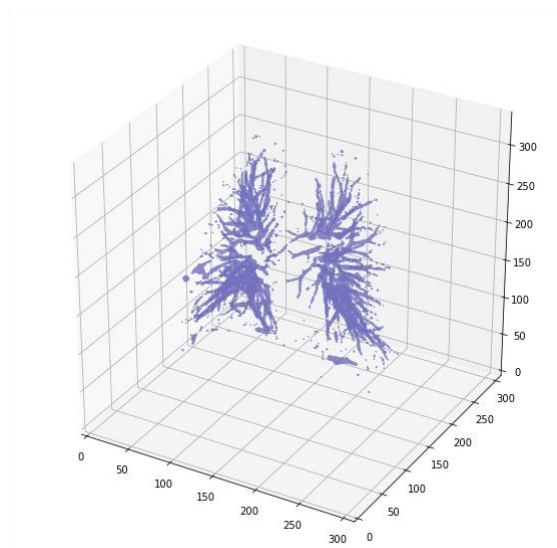


Εικόνα 3.2 - Δεδομένα πριν την προεπεξεργασία

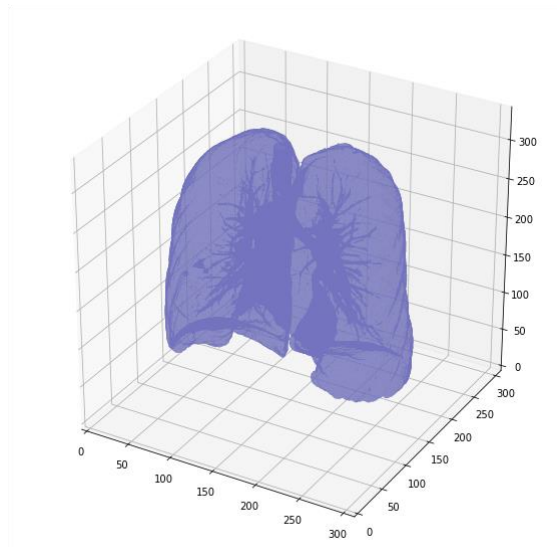
Μια σάρωση μπορεί να έχει απόσταση μεταξύ των εικονοστοιχείων [2,5, 0,5, 0,5], πράγμα που σημαίνει ότι η απόσταση μεταξύ των οριζόντιων κομματιών είναι 2,5 χιλιοστά. Για διαφορετική σάρωση, μπορεί να είναι [1,5, 0,725, 0,725], καθιστώντας το προβληματικό για την αυτόματη ανάλυση (π.χ. στην χρήση των CNNs).



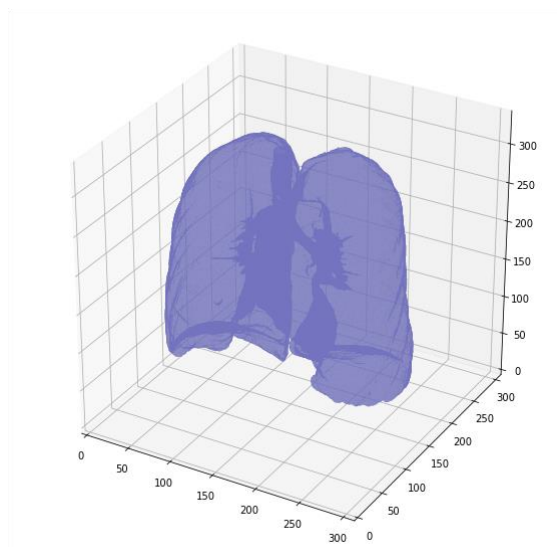
Εικόνα 3.3 - Δείγμα 3D εικόνας ασθενούς με τιμές εικονοστοιχείων μεγαλύτερες από 400 HU αποκαλύπτει το τμήμα των οστών



Εικόνα 3.4 - Το βρογχικό δέντρο για δείγμα ασθενούς



Εικόνα 3.5 - Αρχική απεικόνιση χωρίς αέρα για δείγμα ασθενούς



Εικόνα 3.6 - Τελική απεικόνιση συμπεριλαμβανομένων των βρογχιδίων για δείγμα ασθενούς

Μια κοινή μέθοδος αντιμετώπισης του προβλήματος αυτού είναι η επαναδειγματοληψία του πλήρους συνόλου των δεδομένων σε μία ορισμένη ιστροπική ανάλυση. Εάν επιλέξουμε την επαναδειγματοληψία σε εικονοστοιχεία εύρους [1mm, 1mm, 1mm] μπορούμε να χρησιμοποιήσουμε τα 3D CNNs χωρίς να ανησυχούμε για την αμεταβλητότητα του πάχους των οριζοντίων κομματιών κατά την εκμάθηση.

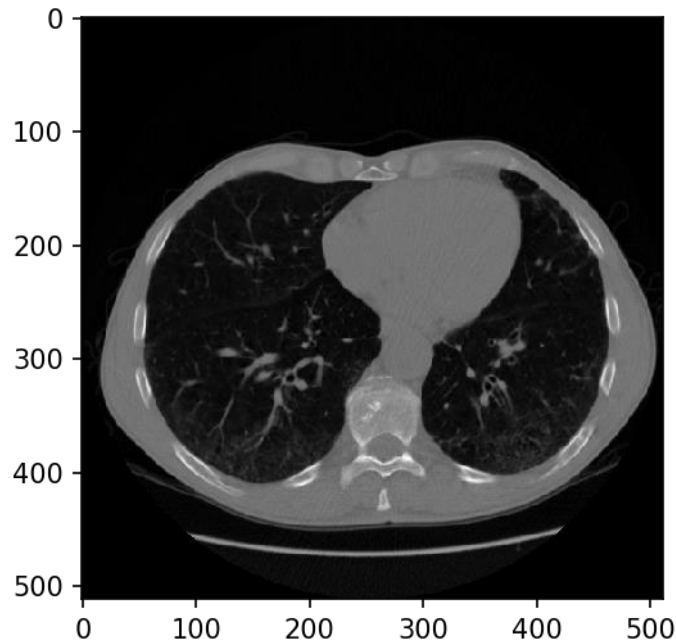
Ενώ αυτό μπορεί να μοιάζει με ένα πολύ απλό βήμα, έχει αρκετές ακραίες περιπτώσεις λόγω της στρογγυλοποίησης. Επίσης, είναι χρονοβόρα διαδικασία.

Οι τιμές μας, επί του παρόντος, κυμαίνονται από -1024 έως περίπου 2000. Οτιδήποτε πάνω από 400 δεν μας ενδιαφέρει, καθώς είναι απλά οστά με διαφορετική ραδιοπυκνότητα. Ένα κοινά χρησιμοποιούμενο εύρος κατώτατων ορίων για την ομαλοποίηση είναι μεταξύ -1000 και 400.

Συν τοις άλλοις, τοποθετούμε με, τα δεδομένα μας, κέντρο το μηδέν (zero-center) ώστε η μέση τιμή, των δεδομένων, να είναι '0'. Για να το κάνουμε αυτό απλά αφαιρούμε την μέση τιμή των εικονοστοιχείων από όλα τα εικονοστοιχεία.

Για να προσδιορίσουμε τη μέση τιμή απλά υπολογίζουμε τον μέσο όρο όλων των εικόνων σε ολόκληρο το σύνολο δεδομένων, στην περίπτωση μας είναι περίπου 0,25.

Ως τελικό βήμα της προεπεξεργασίας, αλλάζουμε το μέγεθος κάθε εικόνας των δεδομένων μας από 512x512 σε 50x50 και για κάθε ασθενή λαμβάνουμε προκαθορισμένο αριθμό εικόνων, στην περίπτωσή μας 50, με τελικό μέγεθος ταυστή 50x50x50, και όχι το συνολικό αριθμό εικόνων ανά ασθενή επειδή η τιμή αυτή είναι μεταβλητή και για την χρήση των δεδομένων σε 3D CNNs θέλουμε οι διαστάσεις κάθε ταυστή να είναι ίδιες και λόγω της περιορισμένης υπολογιστικής ισχύς του εξοπλισμού να έχουν μικρό μέγεθος.



Εικόνα 3.7 - Δεδομένα μετά την προεπεξεργασία

3.4.2 Σχεδιασμός Νευρωνικού Δικτύου

Ένα τρισδιάστατο συνελκτικό νευρωνικό δίκτυο χρησιμοποιείται ως γραμμικός ταξινομητής. Χρησιμοποιεί σταθμισμένη απώλεια διασταυρούμενης εντροπίας για συνάρτηση ενεργοποίησης softmax (το βάρος για μια ετικέτα είναι το αντίστροφο της συχνότητας της ετικέτας στο σύνολο εκπαίδευσης) και του βελτιστοποιητή Adam, τα CNNs χρησιμοποιούν την συνάρτηση ενεργοποίηση ReLU και την απόσυρση (dropout) μετά από κάθε στρώμα συνέλιξης κατά τη διάρκεια της εκπαίδευσης. Το δίκτυο έχει συρρικνωθεί για να αποφευχθεί η υπερφόρτωση παραμέτρων λόγω του σχετικά μικρού αρχείου δεδομένων από το Kaggle. Η αρχιτεκτονική 3D CNN περιγράφεται λεπτομερώς στον Πίνακα 3.2.

Το συνελκτικό νευρωνικό δίκτυο αποτελείται από έναν πεπερασμένο αριθμό συνελκτικών στρωμάτων, ακολουθούμενων από ένα ή και περισσότερων πλήρως συνδεδεμένων (Fully Connected) στρώματα και τέλος ένα στρώμα εξόδου. Ένα παράδειγμα αυτής της αρχιτεκτονικής απεικονίζεται στην Εικόνα 3.8.

Πίνακας 3.2 - CNN αρχιτεκτονική 4 επιπέδων (dropout με 0.8 πιθανότητα μετά από κάθε conv επίπεδο κατά την εκπαίδευση, Adam βελτιστοποιητής με δείκτη εκπαίδευσης 0.001)

Layer	Parameters	Activation	Output
Input			50x50x50x1
Conv1	3x3x3	ReLu	17x17x17x3
Max Pool	2x2, stride 2		9x9x9x32
Conv2	3x3x3	ReLu	3x3x3x64
Max Pool	2x2, stride 2		2x2x2x64
Conv3	3x3x3	ReLu	1x1x1x32
Max Pool	2x2, stride 2		1x1x1x32
Conv4	3x3x3	ReLu	1x1x1x64
Max Pool	2x2, stride 2		1x1x1x64

Κώδικας περιγραφής του μοντέλου:

```
import tflearn

from tflearn.layers.conv import conv_3d, max_pool_3d

from tflearn.layers.core import input_data, dropout, fully_connected

from tflearn.layers.estimator import regression

convnet = input_data(shape=[None, IMG_SIZE_PX, IMG_SIZE_PX,
SLICE_COUNT, 1], name='input')

convnet = conv_3d(convnet, 32, 3, 3, activation='relu')

convnet = max_pool_3d(convnet, 2, 2)

convnet = conv_3d(convnet, 64, 3, 3, activation='relu')

convnet = max_pool_3d(convnet, 2, 2)
```

```
convnet = conv_3d(convnet, 32, 3, 3, activation='relu')
```

```
convnet = max_pool_3d(convnet, 2, 2)
```

```
convnet = conv_3d(convnet, 64, 3, 3, activation='relu')
```

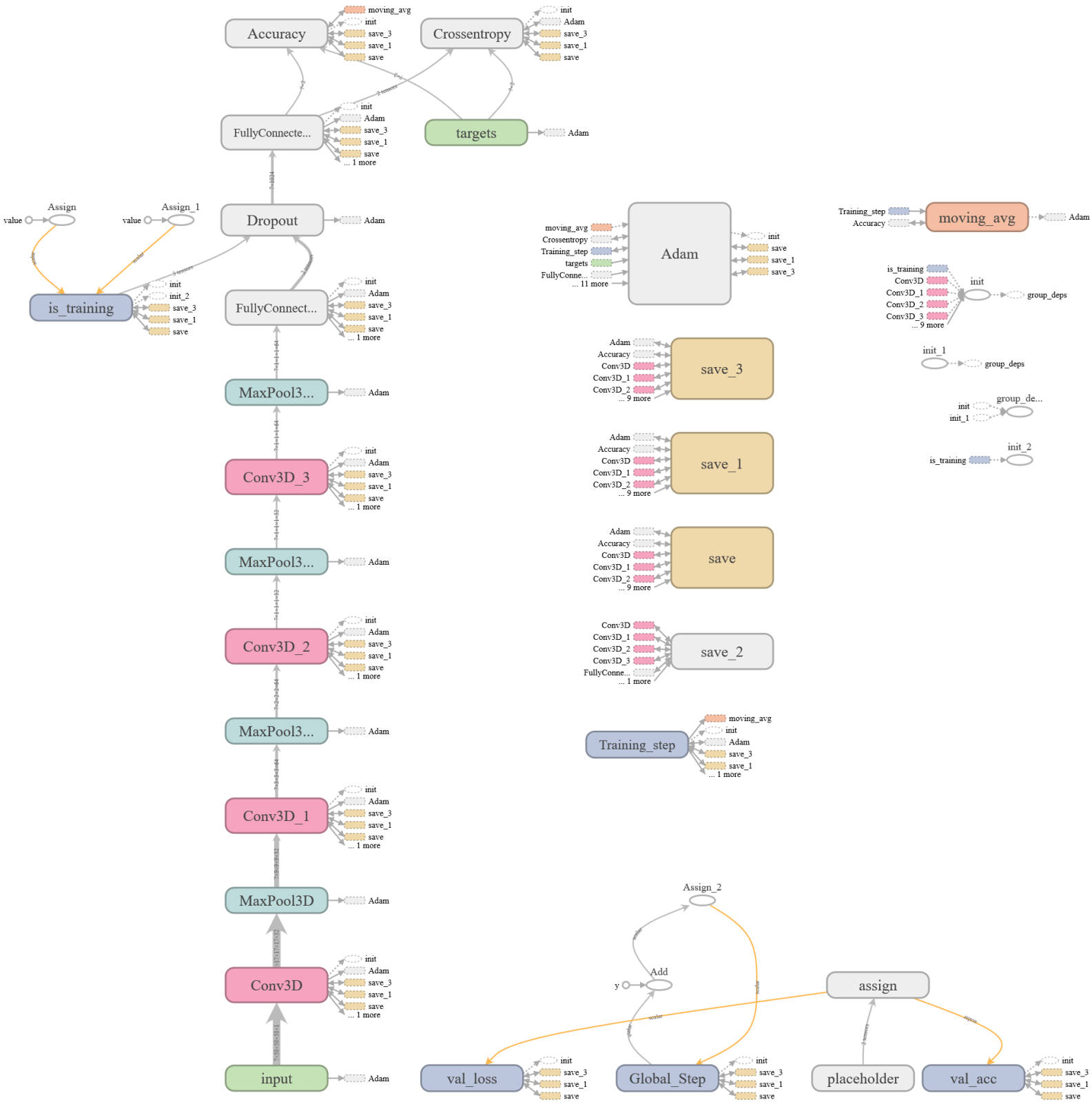
```
convnet = max_pool_3d(convnet, 2, 2)
```

```
convnet = fully_connected(convnet, 1024, activation='relu')
```

```
convnet = dropout(convnet, 0.8)
```

```
convnet = fully_connected(convnet, 2, activation='softmax')
```

```
convnet = regression(convnet, optimizer='adam', learning_rate=LR,  
loss='categorical_crossentropy', name='targets')
```



Εικόνα 3.8 - CNN Αρχιτεκτονική

3.5 Αποτελέσματα

Ο πίνακας σύγχυσης (Confusion Matrix) είναι ένα από τα πιο ευκολονόητα και εύστοχα κριτήρια που χρησιμοποιούνται για την εύρεση της ορθότητας και της ακρίβειας ενός μοντέλου. Χρησιμοποιείται για προβλήματα κατηγοριοποίησης όπου η έξοδος μπορεί να είναι δύο ή περισσότερων τύπων κλάσεων.

Όροι που σχετίζονται με τον πίνακα σύγχυσης:

Αληθή θετικά (True Positives): Αληθή θετικά είναι το ποσοστό των περιπτώσεων όπου η πραγματική τάξη του σημείου δεδομένων ήταν 1 (True) και η προβλεπόμενη είναι επίσης 1 (True), στην περίπτωση μας είναι 71.5%.

Αληθή αρνητικά (True Negatives): Αληθή αρνητικά είναι το ποσοστό των περιπτώσεων όπου η πραγματική τάξη του σημείου δεδομένων ήταν 0 (False) και η προβλεπόμενη είναι επίσης 0 (False), στην περίπτωση μας είναι 37.2%.

Ψευδές θετικά (False Positives): Ψευδές θετικά είναι το ποσοστό των περιπτώσεων όπου η πραγματική τάξη του σημείου δεδομένων ήταν 0 (False) και η προβλεπόμενη είναι 1 (True). Είναι Ψευδές επειδή το μοντέλο έχει προβλέψει λανθασμένα και Αληθή επειδή η τάξη που προβλεπόταν ήταν θετική (1), στην περίπτωση μας είναι 28.5%.

Ψευδές αρνητικά (False Negatives): Ψευδές αρνητικά είναι το ποσοστό των περιπτώσεων όπου η πραγματική τάξη του σημείου δεδομένων ήταν 1 (True) και η προβλεπόμενη είναι 0 (False). Είναι Αληθή επειδή το μοντέλο έχει προβλέψει λανθασμένα και Ψευδές επειδή η τάξη που προβλεπόταν ήταν αρνητική (0), στην περίπτωση μας είναι 62.8%.

Πίνακας 3.3 - Πίνακας Σύγκρισης για 3D CNN

	Predicted	
	Normal	Abnormal
Actual		
Normal	0.715	0.285
Abnormal	0.628	0.372

Μέσω των όρων του σχετίζονται με τον πίνακα σύγκρισης, μπορούμε να λάβουμε τις παρακάτω μετρήσεις:

Precision

Το Precision είναι ο αριθμός των True Positives διαιρούμενο με το αρθροισμα των True Positives και False Positives. Με άλλο τρόπο, είναι ο αριθμός των θετικών προβλέψεων διαιρούμενος με τον συνολικό αριθμό των θετικών τιμών της τάξης που προβλέπονται. Ονομάζεται επίσης Θετική προγνωστική αξία (Positive Predictive Value).

Το Precision μπορεί να θεωρηθεί και ως μέτρο της ακρίβειας ενός ταξινομητή. Ένα χαμηλο Precision μπορεί επίσης να υποδεικνύει έναν μεγάλο αριθμό False Positives.

Recall

Το Recall είναι ο αριθμός των True Positives διαιρούμενο με το αρθροισμα των True Positives και τον αριθμό των False Negatives. Με έναν άλλο τρόπο είναι ο αριθμός των θετικών προβλέψεων διαιρούμενος με τον αριθμό των θετικών τιμών της τάξης στα δεδομένα προς δοκιμή. Ονομάζεται επίσης Ευαισθησία ή True Positive Rate.

Το Recall μπορεί να θεωρηθεί και ως μέτρο της πληρότητας ενός ταξινομητή. Ένα χαμηλο Recall μπορεί να υποδεικνύει πολλά False Negatives.

F1-Score

Το F1-Score είναι $\eta\ 2 * ((\text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall}))$. Ονομάζεται επίσης F-Score ή F-Measure. Με άλλο τρόπο, το F1-Score δείχνει τη σχέση μεταξύ της ακρίβειας και της ανάκλησης.

Οι τιμές των παραπάνω μετρήσεων φαίνονται στο παρακάτω Πινάκα 4.

Πίνακας 3.4 - Έκθεση κατηγοριοποίησης για 3D CNN

Classification report:				
	Precision	Recall	F1-Score	Support
0	0.76	0.71	0.74	890
1	0.3	0.35	0.32	307
Avg/Total	0.64	0.62	0.63	1197

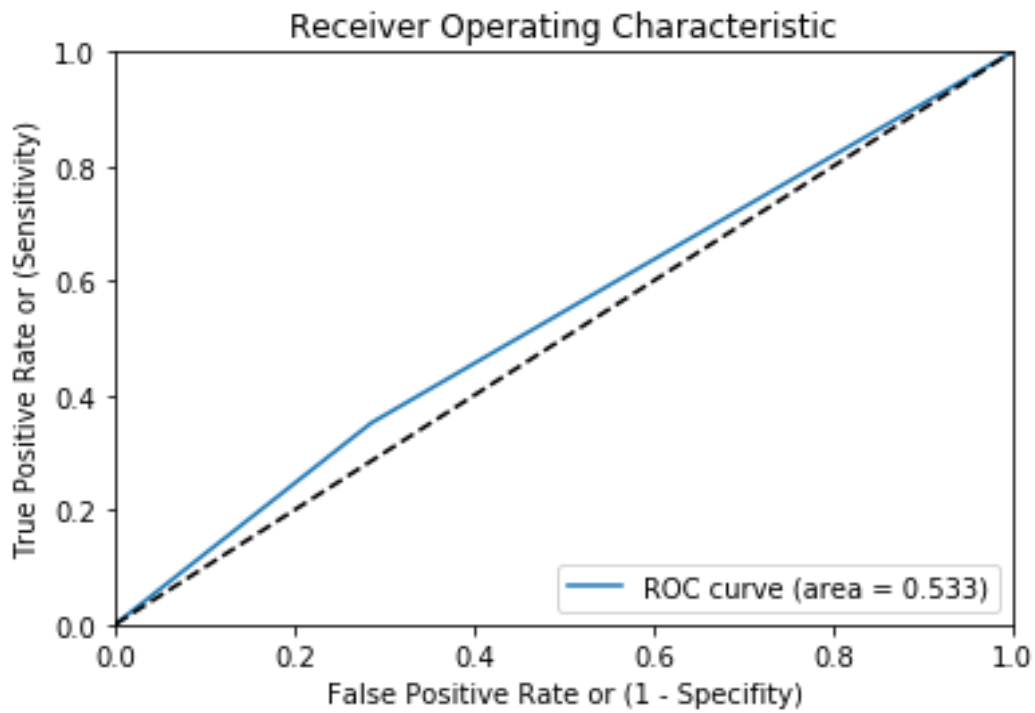
Αξιολόγηση της απόδοσης

Οι καμπύλες ROC μας δίνουν τη δυνατότητα να εκτιμήσουμε την απόδοση του κατηγοριοποιητή σε όλο το εύρος της λειτουργίας του. Το ευρύτερα χρησιμοποιούμενο μέτρο είναι η περιοχή κάτω από την καμπύλη (Area Under the Curve). Όπως μπορούμε να δούμε στην Εικόνα 13, η AUC για έναν ταξινομητή χωρίς ισχύ, ουσιαστικά τυχαία εικασία, είναι 0.5, επειδή η καμπύλη ακολουθεί τη διαγώνιο. Η AUC για τον τέλειο ταξινομητή, είναι 1.0. Οι περισσότεροι ταξινομητές έχουν τιμές AUC που πέφτουν κάπου μεταξύ αυτών των δύο τιμών.

Μια AUC μικρότερη από 0.5 μπορεί να υποδηλώνει ότι συμβαίνει κάτι ενδιαφέρον. Μια πολύ χαμηλή τιμή AUC μπορεί να υποδεικνύει ότι το πρόβλημα έχει δημιουργηθεί λανθασμένα, ο ταξινομητής βρίσκει μια σχέση στα δεδομένα που είναι, ουσιαστικά, το αντίθετο από αυτό που αναμένεται. Σε μια τέτοια περίπτωση, η επιθεώρηση ολόκληρης της καμπύλης ROC θα μπορούσε να δώσει κάποιες ενδείξεις σχετικά με το τι συμβαίνει: έχουν δοθεί λανθασμένες ετικέτες στα Αληθή και Ψευδές δεδομένα;

Σύγκριση κατηγοριοποιητών

Η AUC μπορεί να χρησιμοποιηθεί για να συγκρίνει την απόδοση δύο ή περισσότερων κατηγοριοποιητών. Μπορεί να επιλεγεί ένα ενιαίο όριο και να συγκριθεί η απόδοση των κατηγοριοποιητών σε αυτό το σημείο ή η συνολική απόδοση να μπορεί να συγκριθεί λαμβάνοντας υπόψη την AUC.



Εικόνα 3.9 - ROC για 3D CNN

3.6 Συμπεράσματα

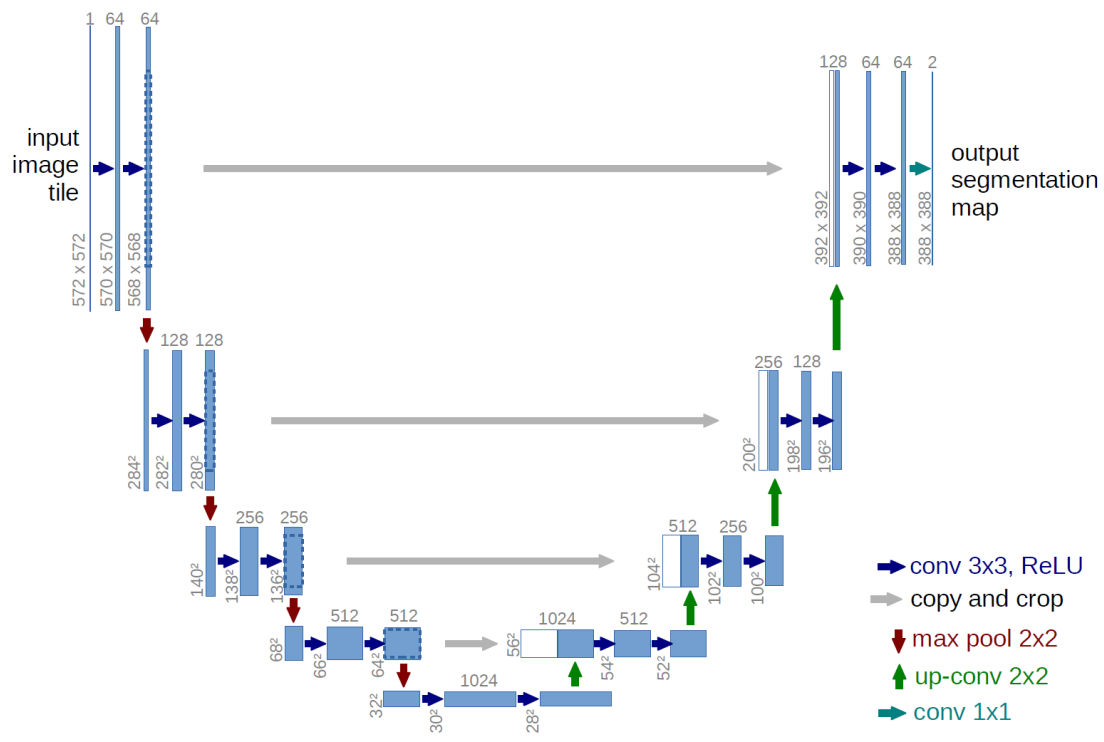
Στην παρούσα εργασία σχεδιάσαμε, αναπτύξαμε και εκπαιδεύσαμε ένα βαθύ τρισδιάστατο συνελκτικό νευρωνικό δίκτυο με σκοπό τον εντοπισμό και κατηγοριοποίηση των ασθενών με βάση την ύπαρξη ή όχι καρκίνου του πνεύμονα στις αξονικές τους τομογραφίες. Ο εντοπισμός του καρκίνου του πνεύμονα σε πλήρεις τρισδιάστατες αξονικές τομογραφίες είναι απαιτητική διεργασία.

Τα δεδομένα που χρησιμοποιήσαμε είχαν ατέλειες όπως την έλλειψη ετικετών σε περίπου 398 αξονικές τομογραφίες. Επίσης, σε όλο το εύρος των δεδομένων μας, η θέση ή το μέγεθος των καρκινικών οζιδίων δεν είχαν ετικέτα καθιστώντας την κατηγοριοποίηση τους ακόμα δυσκολότερη. Σε συνδυασμό με την χαμηλή επεξεργαστική ισχύ του εξοπλισμού μας τα αποτελέσματα ήταν μετρίου επιπέδου.

Συγκεκριμένα, είχαμε ποσοστό επιτυχίας 75.23% με ακρίβεια 64% και μέσω των καμπυλών λειτουργικού χαρακτηριστικού δέκτη (ROC) βλέπουμε πως το εμβαδόν της περιοχής κάτω από την καμπύλη (AUC) είναι 0.533. Στα δεδομένα μας έχουμε 890 ασθενείς χωρίς καρκίνο και 307 ασθενείς με καρκίνο όποτε ένα μοντέλο που θα πρόβλεπε σε όλες τις περιπτώσεις μη καρκίνο θα είχε $890/1197 = 74.35\%$ ποσοστό επιτυχίας ενώ αν πρόβλεπε καρκίνο $307/1197 = 25.64\%$ ποσοστό επιτυχίας. Επειδή όμως έχουμε ακρίβεια 64% δεν συμβαίνει κάτι τέτοιο στο μοντέλο μας. Για έναν ταξινομητή χωρίς ισχύ, ουσιαστικά τυχαία εικασία, το εμβαδόν της περιοχής κάτω από την καμπύλη (AUC) είναι 0.5, στην περίπτωσή μας είναι ελαφρώς καλύτερο με τιμή 0.533.

Υπάρχουν πολλοί τρόποι βελτιώσεις των αποτελεσμάτων. Αρχικά, εξοπλισμός με μεγαλύτερη επεξεργαστική ισχύ θα μπορούσε λάβει ως είσοδο τα δεδομένα στην αρχική τους μορφή χωρίς κάποια συμπίεση αρά και απώλεια πληροφορίας. Συν τοις άλλοις, δεδομένα στα οποία η θέση ή το μέγεθος των καρκινικών οζιδίων έχουν ετικέτα όπως τα δεδομένα του διαγωνισμού LUNA16[43] θα έδιναν την δυνατότητα, σε συνδυασμό με καλύτερο εξοπλισμό, χρήσης αρχιτεκτονικής δισδιάστατου συνελκτικού νευρωνικού δικτύου U-Net το οποίο

είναι διαδεδομένο για την κατάτμηση βιοϊατρικών εικόνων. Τέλος, η συμβολή εξειδικευμένων ιατρών θα βοηθούσε στην ευκολότερη και γρηγορότερη αναγνώριση και προσθήκη ετικετών στα καρκινικά οζίδια των αξονικών τομογραφιών.



Εικόνα 3.10 - Αρχιτεκτονική U-Net

Άξιο αναφοράς είναι το γεγονός πως στον διαγωνισμό του Kaggle η ομάδα του Julian de Wit, η οποία κατέκτησε την 2^η θέση, χρησιμοποίησε εργαλεία που επέτρεπαν επιτάχυνση υλικού για GPU της Nvidia, κάνοντας χρήση της GTX Titan X και όμως είχαν περιορισμούς στο μέγεθος και στον αριθμό των μοντέλων. Για αυτόν τον λόγο πειραματίστηκαν με υπηρεσίες on-demand Cloud Computing της Amazon (χρήση NVIDIA Tesla K80 GPU) και έτσι μπόρεσαν όχι μόνο να εκπαιδεύουν 5 μοντέλα ταυτόχρονα στον χρόνο του ενός αλλά να τα κάνουν και "βαθύτερα", καθιστώντας έτσι το κομμάτι της υπολογιστικής ισχύς εξίσου σημαντικό με την πληθώρα δεδομένων.

Βιβλιογραφία

- [1] Alan Mathison Turing 'Computing Machinery and Intelligence' *Mind*, New Series, Vol. 59, No. 236, pp. 433_460, 1950
- [2] Marvin L. Minsky, Seymour A. Papert 'Perceptrons: An Introduction to Computational Geometry' M.I.T. Press, 1969
- [3] David E. Rumelhart, Geoffrey E. Hinton, Ronald J. Williams 'Learning representations by back-propagating errors' *Nature*, Vol. 323, pp. 533_536, 1986
- [4] Ofer Matan, Jane Bromley, Christopher J. C. Burges, John S. Denker, Lawrence D. Jackel, Yann Le Cun, Eswin P. D. Pednault, Wilialm D. Satterfield, Charles E. Stenard, Timothy J. Thompson 'Reading Handwritten Digits: A Zip Code Recognition System' *AT&T Bell Laboratories*, 1992
- [5] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, Li Fei-Fei 'ImageNet: A Large-Scale Hierarchical Image Database' *CVPR09, 2009*
- [6] Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton 'ImageNet Classification with Deep Convolutional Neural Networks' *NIPS*, 2012
- [7] James Somers 'The Scientific Paper Is Obsolete' *The Atlantic*, 2018
- [8] Jeff Dean, Rajat Monga 'TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems' *TensorFlow.org*, 2015
- [9] Cade Metz 'Google Just Open Sourced TensorFlow, Its Artificial Intelligence Engine' *Wired*, 2015
- [10] Will Oremus 'What is TensorFlow and Why is Google So Excited About It' *Slate*, 2015
- [11] Jeff Ward-Baliley 'Google chairman: We're making real progress on artificial intelligence' *CSMonitor*, 2015

- [12] Norm Jouppi 'Google supercharges machine learning tasks with TPU custom chip' *Google Cloud Platform Blog*, 2016
- [13] Jeff Dean, Urs Hölzle 'Build and train machine learning models on our new Google Cloud TPUs' *Google*, 2017
- [14] John Barrus, Zak Stone 'Cloud TPU machine learning accelerators now available in beta' *Google Cloud Platform Blog*, 2018
- [15] Dogs vs. Cats Redux: Kernels Edition, *Kaggle*, 2016
<https://www.kaggle.com/c/dogs-vs-cats-redux-kernels-edition> προσπέλαση στις 15/3/2018
- [16] Jacques Ferlay, Hai-Rim Shin, Freddie Bray, David Forman, Colin Mathers, Donald Maxwell Parkin 'Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008' *International Journal of Cancer*, Vol. 127, pp. 2893_2917, 2010
- [17] Ahmedin Jemal, Ram C. Tiwari, Taylor Murray, Asma Ghafoor, Alicia Samuels, Elizabeth Ward, Eric J. Feuer, Michael J. Thun 'Cancer statistics, 2004' *CA Cancer J Clin*, Vol. 54, pp.8_29, 2008
- [18] 'SEER stat fact sheets: Lung and Bronchus. Surveillance Epidemiology and End Results' *National Cancer Institute*, 2010
- [19] 'Gender in lung cancer and smoking research' *World Health Organization*, 2004
- [20] Jing Zhang, Jia-Xian OU, Chun-Xue Bai 'Tobacco smoking in China: prevalence, disease burden, challenges and future strategies' *Respirology*, Vol. 16, pp.1165_1172
- [21] Behera, D, Balamugesh T 'Lung cancer in India' *Indian Journal of Chest Diseases and Allied Sciences* Vol. 46, pp.269_281
- [22] Dan Longo, Anthony Fauci, Dennis Kasper, Stephen Hauser, J. Jameson, Joseph Loscalzo '*Harrison's Principles of Internal Medicine*' McGraw Hill Professional, 2011
- [23] Cigdem Demir, Bülent Yener 'Automated cancer diagnosis based on histopathological images: a systematic survey' *Rensselaer Polytechnic Institute*, 2005

- [24] Farzad Khalvati, Alexander Wong, Masoom A. Haider 'Automated prostate cancer detection via comprehensive multi-parametric magnetic resonance imaging texture feature models' *BMC medical imaging*, 2015
- [25] 'Challenge on cancer metastases detection in lymph node' *Camelyon16*, 2016
<http://camelyon16.grand-challenge.org> προσπέλαση στις 19/6/2018
- [26] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun 'Deep residual learning for image recognition' *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp.770_778, 2016
- [27] Karen Simonyan, Andrew Zisserman 'Very Deep Convolutional Networks for Large-Scale Image Recognition' *CoRR abs/1409.1556*, 2014
- [28] Shaoqing Ren, Kaiming He, Ross Girshick, Jian Sun 'Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks' *NIPS*, 2015
- [29] Baris Kayalibay, Grady Jensen, Patrick van der Smagt 'CNN-based Segmentation of Medical Imaging Data' *CoRR abs/1701.03056*, 2017
- [30] Yun Liu, Krishna Gadepalli, Mohammad Norouzi, George E. Dahl, Timo Kohlberger, Aleksey Boyko, Subhashini Venugopalan, Aleksei Timofeev, Philip Q. Nelson, Greg S. Corrado, Jason D. Hipp, Lily Peng, Martin C. Stumpe 'Detecting Cancer Metastases on Gigapixel Pathology Images' *CoRR abs/1703.02442*, 2017
- [31] Fausto Milletari, Nassir Navab, Seyed-Ahmad Ahmadi 'V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation' *2016 Fourth International Conference on 3D Vision (3DV)*, 2016
- [32] Özgün Çiçek, Ahmed Abdulkadir, Soeren S. Lienkamp, Thomas Brox, Olaf Ronneberger '3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation' *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 424_432, 2016
- [33] Hao Chen, Qi Dou, Lequan Yu, Pheng-Ann Heng 'VoxResNet: Deep Voxelwise Residual Networks for Volumetric Brain Segmentation' *CoRR abs/1608.05895*, 2016

[34] Shu Liao, Yaozong Gao, Aytekin Oto, Dinggang Shen 'Representation Learning: A Unified Deep Learning Framework for Automatic Prostate MR Segmentation' *International Conference on Medical Image Computing and Computer Assisted Intervention*, pp. 254_261, 2013

[35] Devinder Kumar, Alexander Wong, David A. Clausi 'Lung Nodule Classification Using Deep Features in CT Images' *12th Conference on Computer and Robot Vision*, pp. 133_138, 2015

[36] Arnaud Arindra Adiyoso Setio, Alberto Traverso, Thomas de Bel, Moira S.N. Berens, Cas van den Bogaard, Piergiorgio Cerello, Hao Chen, Qi Dou, Maria Evelina Fantacci, Bram Geurts, Robbert van der Gugten, Pheng Ann Heng, Bart Jansen, Michael M.J. de Kaste, Valentin Kotov, Jack Yu-Hung Lin, Jeroen T.M.C. Manders, Alexander Sónora-Mengana, Juan Carlos García-Naranjo, Evgenia Papavasileiou, Mathias Prokop, Marco Saletta, Cornelia M Schaefer-Prokop, Ernst T. Scholten, Luuk Scholten, Miranda M. Snoeren, Ernesto Lopez Torres, Jef Vandemeulebroucke, Nicole Walasek, Guido C.A. Zuidhof, Bram van Ginneken, Colin Jacobs 'Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: The LUNA16 challenge' *Medical Image Analysis*, Vol. 42, pp. 1_13, 2017

[37] Moira Berens, Robbert van der Gugten, Michael de Kaste, Jeroen Manders, Guido Zuidhof 'ZNET - Lung Nodule Detection' 2016

[38] Olaf Ronneberger, Philipp Fischer, Thomas Brox 'U-Net: Convolutional Networks for Biomedical Image Segmentation' *International Conference on Medical Image Computing and Computer Assisted Intervention*, pp. 234-241, 2015

[39] Qi Dou, Hao Chen, Lequan Yu, Jing Qin, Pheng-Ann Heng 'Multilevel Contextual 3-D CNNs for False Positive Reduction in Pulmonary Nodule Detection' *IEEE Transactions on Biomedical Engineering*, vol. 64, No. 7, pp. 1558_1567, 2017

[40] Julian de Wit '2nd place solution for the 2017 national datascience bowl' 2017 <https://juliandewit.github.io/kaggle-ndsb2017/> προσπέλαση στις 9/5/2018

[41] grt123 'Solution of the grt123 Team' 2017

<https://github.com/lfz/DSB2017/blob/master/solution-grt123-team.pdf>

προσπέλαση στις 9/5/2018

[42] Data Science Bowl 2017, *Kaggle*, 2017

<https://www.kaggle.com/c/data-science-bowl-2017/> προσπέλαση στις 24/4/2018

[43] Lung Nodule Analysis 2016 <https://luna16.grand-challenge.org/> προσπέλαση στις 29/4/2018