



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ
ΥΠΟΛΟΓΙΣΤΩΝ

Τομέας Τεχνολογίας Πληροφορικής και Υπολογιστών

Τεχνικές Βαθιάς Μηχανικής Μάθησης για
Ανάλυση Ιατρικών Εικόνων και Διάγνωση Covid-19

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΤΟΥ

Ιωάννη Βλάχου

Επιβλέπων: Κόλλιας Στέφανος
Καθηγητής Ε.Μ.Π.

Αθήνα, Οκτώβριος 2020



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ
ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ
ΤΟΜΕΑΣ ΤΕΧΝΟΛΟΓΙΑΣ ΠΛΗΡΟΦΟΡΙΚΗΣ
ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

Τεχνικές Βαθιάς Μηχανικής Μάθησης για
Ανάλυση Ιατρικών Εικόνων και Διάγνωση Covid-19

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΤΟΥ

Ιωάννη Βλάχου

Επιβλέπων: Κόλλιας Στέφανος
Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 19η Οκτωβρίου 2020 .

.....
Κόλλιας Στέφανος
Καθηγητής Ε.Μ.Π.

.....
Ανδρέας-Γεώργιος
Σταφυλοπάτης
Καθηγητής Ε.Μ.Π.

.....
Γιώργος Στάμου
Αναπληρωτής Καθηγητής
Ε.Μ.Π.

Αθήνα, Οκτώβριος 2020.

.....
Βλάχος Ιωάννης

Διπλωματούχος Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

© Βλάχος Ιωάννης, 2020 Με επιφύλαξη παντός δικαιώματος. All rights reserved. Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η εκτύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.

Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.

Περίληψη

Εν έτει 2020, η ασθένεια Covid-19, η οποία έλαβε διαστάσεις πανδημίας, αποτέλεσε πρωτόγνωρη απειλή για την παγκόσμια υγεία μεταβάλλοντας ραγδαία την καθημερινή μας ζωή.

Από νωρίς διαπιστώθηκε η ανάγκη άμεσης ανίχνευσης και διάγνωσης των ασθενών και φορέων του ιού. Σε αυτή την προσπάθεια η τεχνητή νοημοσύνη αποτελεί καθοριστικό παράγοντα στη διαμόρφωση αξιόπιστων προβλέψεων, οι οποίες δύνανται να συντελέσουν στη λήψη ιατρικών αποφάσεων μειώνοντας σημαντικά τον χρόνο διάγνωσης.

Στο πλαίσιο της διπλωματικής εργασίας, μέσα από την ανάλυση ιατρικών εικόνων αξονικών τομογραφιών θώρακος, σχεδιάζεται αξιόπιστο CNN για την διάγνωση της ασθένειας Covid-19 σε πιθανά κρούσματα. Αρχικά εκπαιδεύονται δίκτυα σε σύνολο δεδομένων ξένων ατόμων, ενώ τα αποτελέσματα προσαρμόζονται και επεκτείνονται σε σειρά ελληνικών δεδομένων.

Για την σαφέστερη κατανόηση και ανάλυση των αποτελεσμάτων, εξάγουμε αναπαραστάσεις για κάθε εικόνα με τις οποίες εφαρμόζουμε ομαδοποίηση στο σύνολο των δεδομένων. Η διαδικασία αυτή μας δείχνει τον τρόπο με τον οποίο κατανέμονται τα δεδομένα στον χώρο του προβλήματος και επιτρέπει την περαιτέρω βελτιστοποίηση των αποτελεσμάτων. Δοκιμάζονται επίσης έτοιμα μοντέλα κατάτμησης ιατρικών εικόνων κρουσμάτων Covid-19, με σκοπό να μελετηθεί η επίδρασή τους στην εκπαίδευση συγκεκριμένων δικτύων.

Τέλος, χρησιμοποιούμε αναδρομικά νευρωνικά δίκτυα για την εξαγωγή προβλέψεων σε ακολουθίες εικόνων και 3D αξονικών τομογραφιών.

Λέξεις Κλειδιά

Κορονοϊός (Covid-19), Βαθιά Μάθηση (Deep Learning), Βαθιά Συνελικτικά Δίκτυα (CNNs), Αναδρομικά Νευρωνικά Δίκτυα (RNNs), Αξονικές Τομογραφίες (CTs), προβλέψεις, ομαδοποίηση (clustering)

Abstract

In the year 2020, the Covid-19 disease, having reached pandemic proportions, constitutes an unprecedented threat to global health, rapidly changing our daily lives.

The need for immediate detection and diagnosis of patients and carriers of the virus was identified early on. In this endeavor, artificial intelligence is a key factor in making reliable predictions, which can contribute to medical decisions, significantly reducing the time of diagnosis.

In the context of this thesis, reliable CNNs are designed for the diagnosis of the disease in possible Covid-19 cases through the analysis of medical images of chest CT scans. Initially, networks are trained in a data set of foreign people, while the results are adapted and extended to a series of Greek data.

For a clearer understanding and analysis of the results, for each image we export representations with which we apply clustering to the data set. This process shows how the data is distributed in the problem space and allows further optimization of the results. Pretrained models of medical image segmentation of Covid-19 cases are also being tested in order to study their impact on the training of specific networks.

Finally we use Recurrent Neural Networks (RNNs) in order to extract predictions on series of images and 3D CT scans.

Key Words

Covid-19, Deep Learning, Convolutional Neural Networks (CNNs), Recurrent Neural Networks (RNNs), CT scans, predictions, clustering

Ευχαριστίες

Θα ήθελα να ευχαριστήσω καταρχήν τον καθηγητή Στέφανο Κόλλια για την εμπιστοσύνη που μου έδειξε στην έρευνα του εν λόγω θέματος, καθώς και για το γεγονός ότι με ενέταξε στην ομαδική προσπάθεια του εργαστηρίου για την εξαγωγή χρήσιμων συμπερασμάτων στο πλαίσιο της έρευνας.

Ευχαριστώ επίσης τον Δημήτρη Κόλλια για την καθοδήγησή του καθόλη τη διάρκεια εκπόνησης της διπλωματικής εργασίας και για τις χρήσιμες παρατηρήσεις του.

Τέλος, ένα μεγάλο ευχαριστώ στην οικογένειά μου για την υποστήριξη που μου έδωσε αυτά τα χρόνια των σπουδών μου.

Περιεχόμενα

Περίληψη	4
Abstract	5
Ευχαριστίες	6
Κατάλογος πινάκων	9
Κατάλογος σχημάτων	11
1 Εισαγωγή	13
1.1 Covid-19	13
1.2 Διάρθρωση Διπλωματικής	14
2 Θεωρία Μηχανικής Μάθησης	15
2.1 Εισαγωγή	15
2.2 Επιβλεπόμενη Μάθηση	15
2.3 Μη Επιβλεπόμενη Μάθηση	16
2.3.1 Συσταδοποίηση (Clustering)	16
2.3.2 Αλγόριθμος K-Means	16
2.4 Deep Neural Networks(DNN)	18
2.4.1 Ορισμός	18
2.4.2 CNN	19
2.4.3 RNN	22
2.4.4 CNN-RNN	25
2.5 Αρχιτεκτονικές Δικτύων	25
2.5.1 Efficient-Net	25
2.5.2 DenseNet	26
2.5.3 ResNet	27
2.6 Μεταφορά Μάθησης (Transfer Learning)	27
3 Σύνολα δεδομένων	29
3.1 Σύνολα Δεδομένων (Datasets)	29
3.1.1 Βασικό dataset Chest CTs	29

3.1.2	Ελληνικό Σύνολο Δεδομένων Chest CTs	31
3.2	Μετρικές Επίδοσης	33
4	Συγγενείς Εργασίες	35
4.1	Βαθιά μάθηση σε ιατρικά δεδομένα	35
4.1.1	Clustering	35
4.1.2	Πρόβλεψη σε ακολουθία δεδομένων	36
4.2	Εκπαίδευση CNN σε σύνολο δεδομένων από Chest CTs	36
5	Υλοποίηση και Πειράματα	38
5.1	Επαύξηση Δεδομένων - Μετασχηματισμοί	38
5.1.1	Εισαγωγή	38
5.1.2	Αποτελέσματα των μετασχηματισμών	38
5.1.3	Συσταδοποίηση (Clustering)	48
5.2	Προβλέψεις σε CT series Ελλήνων ασθενών και μη	51
5.2.1	fine tuning	51
5.2.2	From scratch implementation	53
5.3	Κατάτμηση (segmentation) εικόνων αρχικού Dataset	54
5.4	Επαύξηση στα ελληνικά δεδομένα	59
5.4.1	Δομή Συνόλων Δεδομένων	59
5.4.2	Αρχιτεκτονική μοντέλων πρόβλεψης	60
5.4.3	Αποτελέσματα σε σειρές από CTs	61
6	Επίλογος	67
6.1	Αποτελέσματα και Συμπεράσματα	67
6.2	Μελλοντικές Επεκτάσεις	68
	Βιβλιογραφία	70

Κατάλογος πινάκων

3.1	Κατανομή test συνόλου	30
3.2	Κατανομή validation συνόλου	30
4.1	Προγενέστερα Αποτελέσματα σε ResNet, DenseNet	37
5.1	Μετρικές χωρίς εφαρμογή μετασχηματισμών	39
5.2	Μετρικές με προσθήκη του Ιταλικού Συνόλου Δεδομένων	39
5.3	Μετρικές με Ιταλικό dataset και 2 επιπλέον layers	39
5.4	Αρχικό dataset με 2 επιπλέον layers	40
5.5	Μετρικές με περιστροφή εύρους 5 μοιρών	40
5.6	Μετρικές με περιστροφή εύρους 10 μοιρών	41
5.7	Μετρικές με περιστροφή εύρους 180 μοιρών	41
5.8	Μετρικές με οριζόντια αντιστροφή	41
5.9	Μετρικές με κάθετη αντιστροφή	42
5.10	Μετρικές χωρίς μετασχηματισμό και προσθήκη 1 layer	42
5.11	Μετρικές με προσθήκη Ιταλικού dataset και 1 layer	43
5.12	Μετρικές με περιστροφή 5 μοιρών, κάθετη αντιστροφή, προσθήκη Ιταλικού dataset και 1 επιπλέον layer	43
5.13	Μετρικές με περιστροφή 5 μοιρών, κάθετη και οριζόντια αντιστροφή, προσθήκη Ιταλικού dataset και 1 επιπλέον layer	44
5.14	Μετρικές με κάθετη αντιστροφή, προσθήκη Ιταλικού dataset και 1 επιπλέον layer	44
5.15	Μετρικές με κάθετη και οριζόντια αντιστροφή, προσθήκη Ιταλικού dataset και 1 επιπλέον layer	44
5.16	Μετρικές με κάθετη αντιστροφή, περιστροφή 5 μοιρών και 1 επιπλέον layer	46
5.17	Μετρικές με οριζόντια και κάθετη αντιστροφή και 1 επιπλέον layer	46
5.18	Μετρικές με οριζόντια και κάθετη αντιστροφή	47
5.19	Μοντέλο με την προσθήκη 24 εικόνων covid	48
5.20	Fine-tuning, end-to-end εκπαίδευση, ψήφος "at least one"	52
5.21	Fine-tuning, εκπαίδευση ενός layer, ψήφος "at least one"	52
5.22	Fine-tuning, εκπαίδευση end-to-end, ψήφος πλειοψηφίας	52
5.23	Fine-tuning, εκπαίδευση ενός layer, ψήφος πλειοψηφίας	53
5.24	From scratch, εκπαίδευση ενός layer, ψήφος "τουλάχιστον ένας"	53

5.25	Αποτελέσματα πρώτου segmentation	55
5.26	Αποτελέσματα 2ου segmentation	57
5.27	Αποτελέσματα 3D-Slicer με χρήση της ανίχνευσης μόλυνσης	57
5.28	Αποτελέσματα 3D-Slicer με ανάδειξη περιοχών πνεύμονα	57
5.29	Αποτελέσματα 3ου segmentation με επιλογή κεντρικής τομής	58
5.30	Αποτελέσματα 3ου segmentation με επιλογή slice τομών 1, 32, 64	59
5.31	Στατιστικά μήκους ακολουθίας των 200 Non-Covid ατόμων	59
5.32	Αποτελέσματα για τα dataset 1,2,3 στο test set	62
5.33	Αποτελέσματα για το dataset 4 στο test set	63

Κατάλογος σχημάτων

2.1	Διαδοχικές επαναλήψεις του K-Means αλγορίθμου, [31]	18
2.2	Παράδειγμα CNN για την κατηγοριοποίηση γραπτών ψηφίων, [30]	20
2.3	Πράξη συνέλιξης σε μήτρα εισόδου, [17]	20
2.4	Max-pooling επίπεδο με 2x2 πυρήνα και stride ίσο με 2, [28]	21
2.5	Fully-Connected Layer, [27]	21
2.6	RNN δομή, [32]	22
2.7	LSTM μηχανισμός, [26]	23
2.8	GRU μηχανισμός, [10]	24
2.9	Παράδειγμα δομής CNN-RNN, [12]	25
2.10	Επίδοση των παραλλαγών Efficient-Net, [19]	26
2.11	Δομή της αρχιτεκτονικής Efficient-Net, [19]	26
2.12	Μεταφορά γνώσης από το task A στο B	27
3.1	Τομογραφία ασθενή Covid-19, [34]	29
3.2	Τομογραφία υγιούς ανθρώπου, [34]	30
3.3	Παράδειγμα 1 από Ιταλικό dataset	31
3.4	Παράδειγμα 2 από Ιταλικό dataset	31
3.5	Δείγματα Covid από Zenodo	32
5.1	Εκπαίδευση με τις 110 Ιταλικές εικόνες και χωρίς	45
5.2	Εκπαίδευση με και χωρίς το επιπλέον επίπεδο	45
5.3	Εκπαίδευση βέλτιστου δικτύου σε σχέση με το αρχικό	46
5.4	Ακρίβεια πρόβλεψης κατά τη διάρκεια εκπαίδευσης του καλύτερου μοντέλου.	47
5.5	Συσταδοποίηση στο αρχικό σύνολο δεδομένων	48
5.6	Κέντρα Covid για 10 clusters	49
5.7	Κέντρα Non-Covid για 10 clusters	49
5.8	Συσταδοποίηση στο test set μετά την προσθήκη των 24 αναπαραστάσεων	50
5.9	Συσταδοποίηση στις υπόλοιπες Ιταλικές εικόνες μετά την προσθήκη των 24 αναπαραστάσεων	50
5.10	Κέντρα για 5 clusters	54
5.11	Παραδείγματα Inf-Net	55
5.12	Covid δείγμα και η επίδραση του 3DSlicer	56
5.13	Non-Covid δείγμα και η επίδραση του 3DSlicer	56

5.14 Τομές εξόδου Lung-Segmentation-3D	58
5.15 Εκπαίδευση του CNN-RNN με τα training series για τη 2η μέθοδο dataset	61
5.16 Clustering για το CNN-RNN δίκτυο	64
5.17 Slices εκτός ακτίνας των κέντρων στα οποία ανήκουν	65

Κεφάλαιο 1

Εισαγωγή

1.1 Covid-19

Η ασθένεια κορονοϊού 2019 (Coronavirus disease 2019, COVID-19), είναι μία μολυσματική ασθένεια που προκαλείται από τον κορονοϊό SARS-CoV-2. Ο ιός και η ασθένεια που προκαλεί πρωτοεντοπίστηκε στην πόλη Ουχάν της Κίνας στα τέλη του 2019. Πρόκειται για μία σημαντικά μεταδοτική ασθένεια με ραγδαία εξάπλωση σε όλη την υφήλιο.

Δεν υπάρχουν αποτελεσματικά θεραπευτικά σκευάσματα, με τις προσπάθειες να περιορίζονται στη διαχείριση των συμπτωμάτων και σε υποστηρικτικά μέτρα με οξυγόνο και ενδοφλέβια υγρά, όπου χρειάζεται, ενώ παρακολουθείται η λειτουργία ζωτικών οργάνων.

Εμβόλιο για τη νόσο δεν υπάρχει μέχρι σήμερα ή όσα υπάρχουν βρίσκονται σε πειραματικό στάδιο. Για πρόληψη από την ασθένεια συνιστάται καλή υγιεινή των χεριών. Επιπλέον σε περίοδο επιδημίας συνιστάται η αποφυγή του συγχρωτισμού ή του συνωστισμού, ο περιορισμός των κοινωνικών επαφών και η τήρηση απόστασης από τους άλλους. Η κοινή χειρουργική μάσκα προφυλάσσει σε κάποιο βαθμό τους άλλους από το άτομο το οποίο την φέρει, αλλά οι μάσκες είναι περισσότερο απαραίτητες για τους ασθενείς και εκείνους που τους φροντίζουν.

Διάγνωση με ιατρικές εικόνες είναι δυνατή από εξειδικευμένο ιατρικό προσωπικό. Σε αυτή την προσπάθεια η χρήση αξονικών τομογραφιών θώρακος μπορεί να φανερώσει συγκεκριμένα μοτίβα αναγνωρίσιμα από πνευμονολόγους ή ιατρούς του αντικειμένου. Ακτινογραφίες θώρακος μπορούν επίσης να χρησιμοποιηθούν ως τρόπος διάγνωσης και ανίχνευσης της ασθένειας αν και με λιγότερη ακρίβεια.

Είναι φανερό από τα παραπάνω η ανάγκη επιστράτευσης νέων τεχνολογιών και ενεργοποίησης της παγκόσμιας ερευνητικής κοινότητας για την καταπολέμηση του ιού. Η τεχνητή νοημοσύνη από την πλευρά της, δύναται να σχεδιάσει αξιόπιστα συστήματα για την ανίχνευση και πρόβλεψη του Covid-19, μειώνοντας ταυτόχρονα σημαντικά τον χρόνο διάγνωσης. Σε αυτή την προσπάθεια, εξετάζουμε σε αυτή τη διπλωματική εργασία την σχεδίαση συστημάτων πρόβλεψης με την χρήση βαθιών νευρωνικών δικτύων και πλήθος ιατρικών εικόνων.

1.2 Διάρθρωση Διπλωματικής

Η διπλωματική εργασία αναλύεται σε 6 κεφάλαια. Στο πρώτο και τρέχον κεφάλαιο εμπεριέχονται εισαγωγικές πληροφορίες σχετικά με την ασθένεια Covid-19 και την οργάνωση της εργασίας.

Στο δεύτερο κεφάλαιο συγκαταλέγεται η απαραίτητη θεωρία και οι βασικές γνώσεις για την κατανόηση των μεθόδων μηχανικής μάθησης που χρησιμοποιήθηκαν όπως επίσης και τους βασικούς αλγόριθμους πάνω στους οποίους βασίστηκε το σύνολο των πειραμάτων.

Το τρίτο κεφάλαιο περιλαμβάνει την περιγραφή των συνόλων που χρησιμοποιήθηκαν (ελληνικά και ξένα) για την υλοποίηση και εκπαίδευση των μοντέλων πρόβλεψης.

Στο τέταρτο κεφάλαιο γίνεται αναφορά σε συγγενικές εργασίες με γνώμονα τις οποίες έγινε η ανάλυση των ιατρικών δεδομένων, ούτως ώστε να χρησιμοποιηθούν οι βέλτιστες μέθοδοι για την εξαγωγή αξιόπιστων αποτελεσμάτων και να ληφθούν υπόψη πρότερες ακαδημαϊκές έρευνες.

Το πέμπτο κεφάλαιο περιλαμβάνει τα πειράματα που πραγματοποιήθηκαν πάνω στα δεδομένα.

Το έκτο και τελευταίο κεφάλαιο συμπυκνώνει τα αποτελέσματα της προηγούμενης ενότητας, διατυπώνει τα βασικά συμπεράσματα και εκφράζει τροφή για μελλοντική εργασία πάνω στο μείζον ζήτημα.

Κεφάλαιο 2

Θεωρία Μηχανικής Μάθησης

2.1 Εισαγωγή

Η Μηχανική Μάθηση αποτελεί τομέα της Τεχνητής Νοημοσύνης και αποσκοπεί στην εκπαίδευση μηχανών με στόχο την αυτοματοποίηση έργων. Αφορά τον σχεδιασμό αλγορίθμων και την σταδιακή εκμάθηση μηχανών μέσα από την επεξεργασία συνόλων δεδομένων. Τα δεδομένα αυτά έχουν μια σειρά από τιμές που διαμορφώνουν τα χαρακτηριστικά (features) του κάθε στοιχείου.

Η μηχανική μάθηση διακρίνεται σε τρεις βασικές κατηγορίες.

1. **Επιβλεπόμενη Μάθηση:** Το σύνολο δεδομένων διαθέτει για κάθε στοιχείο του μια τιμή (label) που δείχνει σε ποια κατηγορία ανήκει. Έτσι έχουμε ένα σύνολο εκπαίδευσης με βάση το οποίο μπορούμε να κάνουμε κατηγοριοποίηση των δεδομένων (classification) ή να προβλέψουμε μελλοντικές τιμές (regression).
2. **Μη Επιβλεπόμενη Μάθηση:** Τα δεδομένα που διατίθενται δεν διαθέτουν labels. Στόχος μας είναι η εξαγωγή συμπερασμάτων και εύρεση της δομής των δεδομένων εισόδου με βάση τα χαρακτηριστικά τους.
3. **Ενισχυτική Μάθηση:** Οι αλγόριθμοι μάθησης αλληλεπιδρούν με αντικείμενα σε έναν χώρο δράσης. Καθώς επιτυγχάνεται η περιήγηση στον χώρο του προβλήματος, ο αλγόριθμος εκπαιδεύεται δίνοντας επιβραβεύσεις σε αντίστοιχες ενέργειες και καλείται να δράσει σωστά δίνοντας το καλύτερο δυνατό κέρδος.

Στην παρούσα εργασία θα μας απασχολήσει περισσότερο η Επιβλεπόμενη Μάθηση καθώς και μέθοδοι μη Επιβλεπόμενης Μάθησης με σκοπό την κατηγοριοποίηση συνόλου δεδομένων.

2.2 Επιβλεπόμενη Μάθηση

Σε αυτό το είδος Μηχανικής Μάθησης τα δεδομένα εισόδου X αποτελούνται από N χαρακτηριστικές τιμές το καθένα, ενώ διαθέτουν και επιθυμητές τιμές εξόδου (la-

bels) Y . Στόχος είναι η εύρεση μιας συνάρτησης f που να απεικονίζει τα δεδομένα εισόδου στις επιθυμητές τιμές εξόδου.

$$f : X \rightarrow Y, f(X) = Y$$

Το σύνολο δεδομένων που θα χρησιμοποιηθεί για αυτή την διαδικασία είναι το σύνολο εκπαίδευσης (training set). Αν μας δοθούν δεδομένα εκπαίδευσης με άγνωστο label (test set), μπορούμε να εφαρμόσουμε στην είσοδο X την συνάρτηση f που έχει εκτιμηθεί μέσω ενός αλγορίθμου μηχανικής μάθησης και να λάβουμε μία εκτίμηση για την προβλεπόμενη τιμή Y .

2.3 Μη Επιβλεπόμενη Μάθηση

Σε αυτό το είδος Μηχανικής Μάθησης δε μας είναι γνωστές οι επιθυμητές τιμές εξόδου. Οι αλγόριθμοι που εφαρμόζονται διακρίνουν τα δεδομένα σε διαφορετικές κατηγορίες με βάση τα χαρακτηριστικά τους σχηματίζοντας συστάδες δεδομένων.

2.3.1 Συσταδοποίηση (Clustering)

Σημαντικό κομμάτι σε αυτή τη διαδικασία αποτελεί η ομαδοποίηση των δεδομένων εκπαίδευσης. Κάθε ομάδα διαθέτει παραδείγματα παρόμοιων χαρακτηριστικών, ενώ κάποιои αλγόριθμοι βασίζονται στην εύρεση των κατανομών που ακολουθούν τα δεδομένα αυτά στον χώρο των χαρακτηριστικών.

2.3.2 Αλγόριθμος K-Means

Ο αλγόριθμος k-means αποτελεί έναν αλγόριθμο συσταδοποίησης σε έναν πολυδιάστατο χώρο. Έστω λοιπόν αναπαραστάσεις x_1, x_2, \dots, x_N από N παρατηρήσεις σε έναν D -διάστατο Ευκλείδειο χώρο. Στόχος είναι η ομαδοποίηση σε K συστάδες όπου K ακέραιος που δίνεται. Θεωρούμε K κέντρα $\mu_k, k=1,2,\dots, K$. Ο αλγόριθμος βρίσκει μια ανάθεση κάθε σημείου σε μια συστάδα, όπως επίσης και ένα σύνολο κέντρων τέτοιο ώστε το άθροισμα των τετραγώνων των αποστάσεων κάθε σημείου στο πλησιέστερο κέντρο του μ_k να είναι ελάχιστο.

Σε αυτό το σημείο εισάγουμε για κάθε σημείο x_n ένα σύνολο από δυαδικές μεταβλητές-δείκτες $r_{nk} \in 0, 1$ για $k=1,2,\dots,K$ το οποίο περιγράφει σε ποιο cluster k έχει ανατεθεί η συγκεκριμένη μεταβλητή έτσι ώστε $r_{nk} = 1$ και $r_{nj} = 0$ για $j \neq k$.

Ορίζουμε επίσης την χαρακτηριστική συνάρτηση κόστους η οποία περιγράφει το άθροισμα των τετραγώνων των αποστάσεων κάθε σημείου από το κέντρο στο οποίο αντιστοιχεί.

$$J = \sum_{n=1}^N \sum_{k=1}^K r_{nk} \|x_n - \mu_k\|^2$$

Στόχος είναι να βρούμε τις τιμές r_{nk} και μ_k έτσι ώστε να ελαχιστοποιείται το κριτήριο J . Ακολουθούμε στη συνέχεια μια ειδική περίπτωση του Expectation - Maximization

αλγορίθμου ελαχιστοποιώντας το κριτήριο κόστους διαδοχικά ως προς τα r_{nk} και ύστερα ως προς τα μ_k . Συνεχίζεται αυτή η διπλή διαδικασία μέχρι να έχουμε σύγκλιση του αλγορίθμου.

Για την πρώτη ελαχιστοποίηση του κριτηρίου ως προς r_{nk} αρκεί να παρατηρήσουμε ότι αν αντιστοιχίσουμε σε κάθε σημείο το κέντρο από το οποίο έχει ελάχιστη απόσταση τότε το δοθέν άθροισμα ελαχιστοποιείται. Δηλαδή έχουμε:

$$r_{nk} = \begin{cases} 1 & \text{αν } k = \operatorname{argmin} \|x_n - \mu_j\|^2, \\ 0 & \text{αλλιώς} \end{cases}$$

Για την ελαχιστοποίηση ως προς μ_k παίρνουμε το κριτήριο της πρώτης παραγώγου και βρίσκουμε:

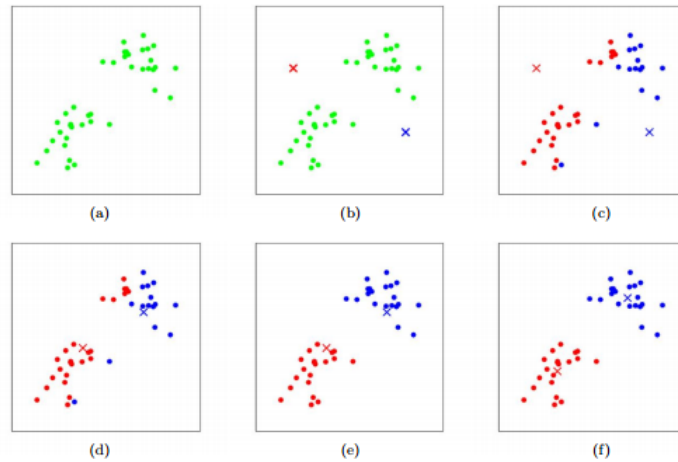
$$2 \sum_{n=1}^N r_{nk} (x_n - \mu_k) = 0$$

Αν το λύσουμε ως προς το ζητούμενο βρίσκουμε: $\mu_k = \frac{\sum_n r_{nk} x_n}{\sum_n r_{nk}}$

Ο παρονομαστής είναι το σύνολο των σημείων που ανατίθενται στο κέντρο μ_k . Έτσι η παραπάνω τιμή είναι ο μέσος όρος των σημείων που έχουν αντιστοιχηθεί στο κέντρο αυτό. Για αυτό το λόγο η παραπάνω διαδικασία ονομάζεται αλγόριθμος K-means.

Συνολικά ο αλγόριθμος φαίνεται στα παρακάτω βήματα:

- Διάλεξε αρχικές τιμές $\mu_k(0)$, $k=1,2,\dots,K$
- Επανάλαβε:
 - for** $i=1$ μέχρι N **do**
 - Βρες το κοντινότερο κέντρο του x_i , έστω μ_j
 - Θέσε κέντρο(i) = j
 - end**
 - for** $j=1$ μέχρι K **do**
 - μ_j = μέσος όρος των $x_i \in X$ με κέντρο(i)= j
 - end**
- Μέχρι να μην υπάρχουν αλλαγές ανάμεσα σε 2 διαδοχικές επαναλήψεις.



Σχήμα 2.1: Διαδοχικές επαναλήψεις του K-Means αλγορίθμου, [31]

2.4 Deep Neural Networks(DNN)

2.4.1 Ορισμός

Τα Βαθιά Νευρωνικά Δίκτυα εκτελούν υπολογισμούς κατά μήκος πολλαπλών επιπέδων νευρώνων. Η μέθοδος αυτή έχει σημαντικές εφαρμογές στην αναγνώριση εικόνων και χρησιμοποιείται όλο και περισσότερο στις σύγχρονες επιστημονικές μεθόδους.

Συνάρτηση Ενεργοποίησης

Η συνάρτηση ενεργοποίησης ή activation function, χρησιμοποιείται στο τέλος των επιπέδων ενός δικτύου στην έξοδο των νευρώνων. Θέτει τις εξόδους σε συγκεκριμένα εύρη τιμών ώστε να είναι διαχειρίσιμες ενώ λειτουργούν αποδοτικά κρατώντας την χρήσιμη πληροφορία για την εκπαίδευση ενός μοντέλου. Χαρακτηριστικές περιπτώσεις τέτοιων είναι οι:

- **Σιγμοειδής:** $Sigm(x) = \frac{1}{1+e^{-x}}$
- **Υπερβολική Εφαπτομένη:** $\tanh x = 2sigmoid(2x) - 1 = 1 - \frac{2}{1+e^{2x}}$
- **Relu:** $Relu(x) = \max(0, x)$

Συνάρτηση Κόστους

Η συνάρτηση αυτή μας λέει πόσο απέχει η έξοδος από το επιθυμητό αποτέλεσμα και αποτελεί μέτρο σύγκλισης των αλγορίθμων εκπαίδευσης που εφαρμόζονται στο δίκτυο. Η παράγωγος αυτής της μεταβλητής είναι αυτό που ανατροφοδοτείται στο

δίκτυο ως πληροφορία. Τα πιο συχνά παραδείγματα που χρησιμοποιούνται στην πράξη είναι:

- **Mean Square Error(MSE)**: Ορίζεται ως το άθροισμα των τετραγωνικών σφαλμάτων των εξόδων από τις επιθυμητές τιμές
$$C(x, y) = \frac{1}{n} \sum \|x_i - y_i\|^2$$
- **Binary Cross Entropy loss**: Πρόκειται για τη συνάρτηση κόστους που χρησιμοποιείται κατά κόρον στα πειράματα της εργασίας. Ορίζεται ως:

$$H_p(q) = -\frac{1}{n} \sum_{i=1}^N y_i \log(p(y_i)) + (1 - y_i) \log(1 - p(y_i))$$

όπου $p(y)$ είναι οι πιθανότητες να επιλεγεί η έξοδος y .

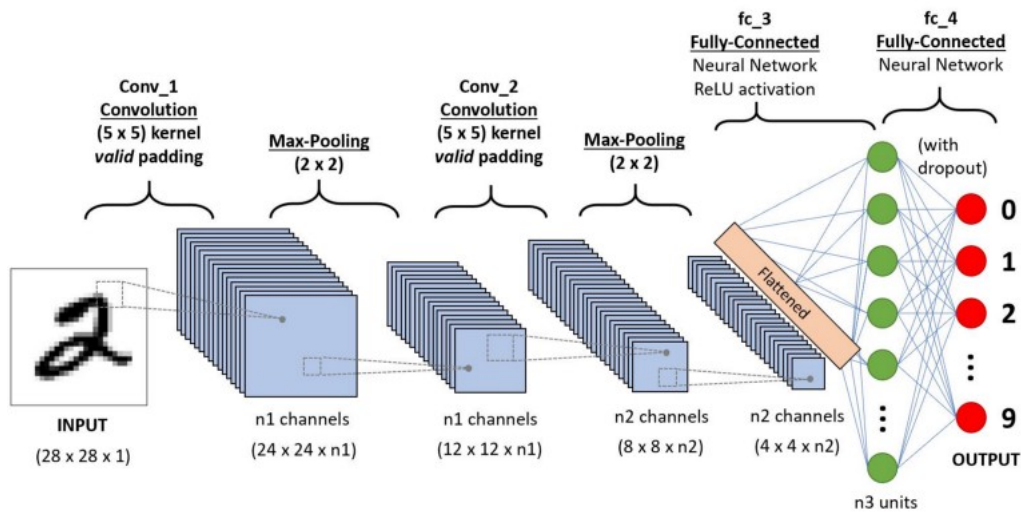
Optimization Αλγόριθμος

Οι αλγόριθμοι βελτιστοποίησης διαχειρίζονται τα βάρη του δικτύου και υπολογίζουν νέα σε κάθε επανάληψη με βάση την παράγωγο της συνάρτησης κόστους. Βασίζονται κυρίως στον αλγόριθμο κατάβασης κλίσης (gradient descent algorithm) και στον back-propagation (οπισθοδιάδοσης). Μερικά συχνά παραδείγματα τέτοιων αλγορίθμων είναι ο SGD, και ο Adam, που έχει επιλεγεί για την πλειοψηφία των πειραμάτων.

- **Stochastic Gradient Descent**: Είναι ένας επαναληπτικός αλγόριθμος που υπολογίζει συνεχώς νέα σημεία στην συνάρτηση κόστους, με σκοπό να φτάσει στο σημείο ελαχιστοποίησής της. Η επιλογή των σημείων γίνεται σύμφωνα με τον ρυθμό εκπαίδευσης και περιλαμβάνει τυχαιότητα που περιορίζει σημαντικά την πολυπλοκότητα εκτέλεσης.
- **Adam**: Ο Adam είναι ένας αλγόριθμος βελτιστοποίησης που αντικαθιστά το στοχαστικό gradient descent, κατάλληλος για την εκπαίδευση βαθιών νευρωνικών δικτύων. Συνδυάζει δύο άλλους αλγορίθμους: τον προσαρμοστικό gradient descent (AdaGrad) και την ρίζα μέσης τετραγωνικής επέκτασης (Root Mean Square Propagation). Χάρη σε αυτούς διαχειρίζεται αραιές παραγωγούς και θορυβώδη δεδομένα. Με τις προεπιλεγμένες παραμέτρους του πετυχαίνει καλά αποτελέσματα σε πληθώρα προβλημάτων.

2.4.2 CNN

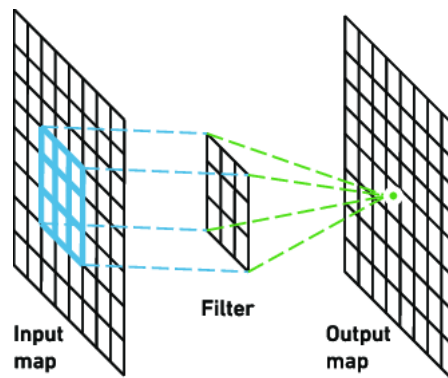
Στην βαθιά μάθηση (deep learning), τα CNN (Convolutional Neural Networks) είναι μια κλάση από DNNs που χρησιμοποιείται κατά κόρον στην ανάλυση εικόνων. Αποτελείται από μία σειρά από layers η οποία σταδιακά μειώνει τις διαστάσεις των δεδομένων εισόδου μέσω συνέλιξης με συγκεκριμένους πυρήνες και δημιουργεί αναπαραστάσεις με αποδοτικό τρόπο. Τα διάφορα είδη των CNNs αναλύονται παρακάτω και περιέχουν συγκεκριμένες δομές και πράξεις στα επιμέρους layers.



Σχήμα 2.2: Παράδειγμα CNN για την κατηγοριοποίηση γραπτών ψηφίων, [30]

Convolutional

Στην περίπτωση ενός συνελικτικού δικτύου, η είσοδος έχει διαστάσεις $N \times$ διαστάσεις εικόνες, όπου N το σύνολο των εικόνων σε ένα batch και η εικόνα είναι τρισδιάστατη έχοντας ύψος, πλάτος και βάθος. Στην έξοδο του συνελικτικού στρώματος δημιουργούνται features επιθυμητών διαστάσεων μετά από διδιάστατη συνέλιξη με συγκεκριμένο πυρήνα μικρότερων διαστάσεων.



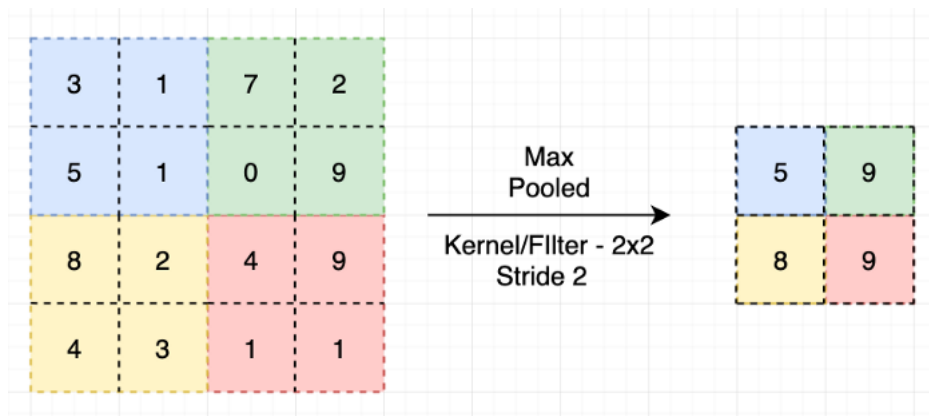
Σχήμα 2.3: Πράξη συνέλιξης σε μήτρα εισόδου, [17]

Pooling

Τα συγκεκριμένα layers μειώνουν τις διαστάσεις της εισόδου στο στρώμα εκτελώντας μια πράξη σε κελιά μικρού μεγέθους και δίνοντας στην έξοδο έναν μοναδικό

νευρώνα για κάθε κελί. Οι πράξεις που μπορεί να χρησιμοποιηθούν σε μια συστάδα από νευρώνες του προηγούμενου στρώματος είναι:

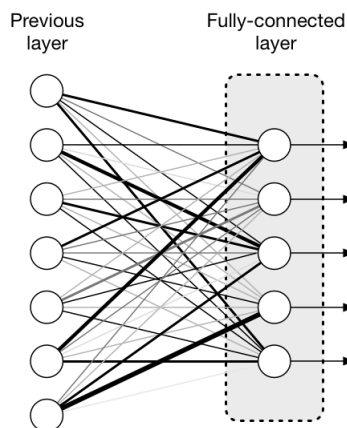
- Max pooling: χρησιμοποιεί την μέγιστη τιμή των τιμών που υπάρχουν στο κελί.
- Average pooling: υπολογίζει τη μέση τιμή των εξόδων των νευρώνων του προηγούμενου στρώματος που περιέχονται στο κελί.



Σχήμα 2.4: Max-pooling επίπεδο με 2x2 πυρήνα και stride ίσο με 2, [28]

Fully-Connected

Τα πλήρως συνδεδεμένα στρώματα στρώματα συνδέουν κάθε νευρώνα του ενός επιπέδου με το επόμενο επίπεδο. Έχουν παρόμοια λειτουργία με τα multi-layer perceptron (MLP). Ο πίνακας που δίνεται στην είσοδο αφού γίνει flat ώστε να έχει μία διάσταση περνάει από ένα πλήρως διασυνδεδεμένο επίπεδο για να κατηγοριοποιηθούν οι εικόνες.



Σχήμα 2.5: Fully-Connected Layer, [27]

2.4.3 RNN

Τα Αναδρομικά Νευρωνικά Δίκτυα ή RNN (Recursive Neural Networks) είναι σχεδιασμένα για να ανταποκρίνονται σε ακολουθιακά προβλήματα όπου πρέπει να γίνει μια πρόβλεψη σε σειρά εισόδων.

Σε αντίθεση με τα feed-forward νευρωνικά δίκτυα, οι καταστάσεις εισόδων είναι εξαρτημένες μεταξύ τους λόγω της αναδρομικής φύσης του δικτύου. Η κατάσταση περιγράφεται από την εξής σχέση:

$$h_t = f(h_{t-1}, x_t)$$

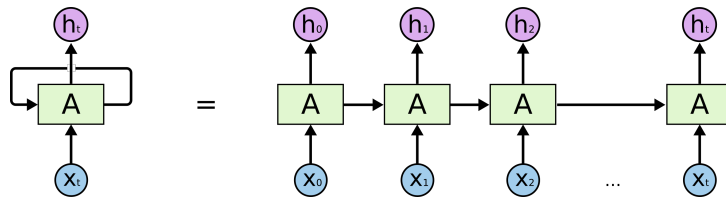
Για συνάρτηση ενεργοποίησης την υπερβολική εφαπτομένη έχουμε την εξής αναδρομική σχέση:

$$h_t = \tanh(W_{hh}h_{t-1}, W_{xh}x_t)$$

όπου W είναι τα βάρη, h το κρυφό επίπεδο, W_{hh} τα βάρη στην προηγούμενη κρυφή κατάσταση, W_{xh} τα βάρη στην τωρινή κατάσταση εισόδου και \tanh η συνάρτηση ενεργοποίησης που εφαρμόζει μη γραμμικότητα στις εξόδους. Η έξοδος θα είναι:

$$y_t = W_{hy}h_t$$

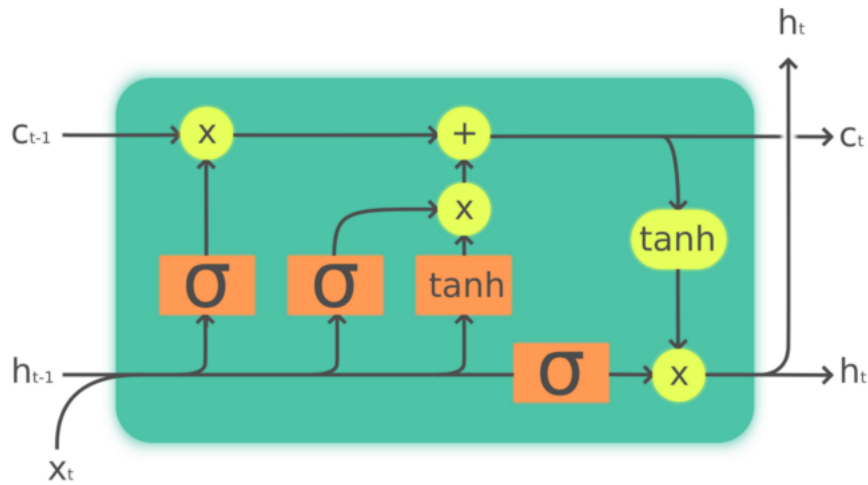
όπου W_{hy} είναι τα βάρη της κατάστασης εξόδου.



Σχήμα 2.6: RNN δομή, [32]

LSTM(Long Short Term Memory)

Ίσως η πιο αποτελεσματική δομή RNN είναι το LSTM το οποίο έχει αποδειχθεί στην πράξη ιδιαίτερα αποτελεσματικό.



Σχήμα 2.7: LSTM μηχανισμός, [26]

Όπως φαίνεται και στην εικόνα ο μηχανισμός αυτός αποτελείται από επιμέρους πύλες, οι οποίες επιτελούν συγκεκριμένες λειτουργίες και συνθέτουν τις εξισώσεις του LSTM.

- **Forget Gate:** Αρχικά τα δεδομένα πολλαπλασιάζονται με τα αντίστοιχα βάρη του μηχανισμού και περνάνε μέσα από μια σιγμοειδή συνάρτηση σύμφωνα με την εξίσωση

$$f_t = \sigma(W^{(f)}[h_{t-1}, x_t] + b_f)$$
- **Input Gate:** Στη συνέχεια, αποφασίζεται ποιες τιμές θα αποθηκευτούν στην τωρινή κατάσταση του κελιού. Η έξοδος του input gate και η ενδιάμεση κατάσταση που προκύπτει φαίνονται στις εξισώσεις

$$i_t = \sigma(W^{(i)}[h_{t-1}, x_t] + b_i)$$

$$\tilde{C}_t = \tanh(W^{(C)}[h_{t-1}, x_t] + b_C)$$

- **Cell State:** Από τις παραπάνω τιμές προκύπτει η νέα κατάσταση του κελιού:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t$$
- **Output Gate:** Τέλος προκύπτει η έξοδος του κελιού ως μία φιλτραρισμένη μορφή της τρέχουσας κατάστασης, συναρτήσει των βαρών του μηχανισμού.

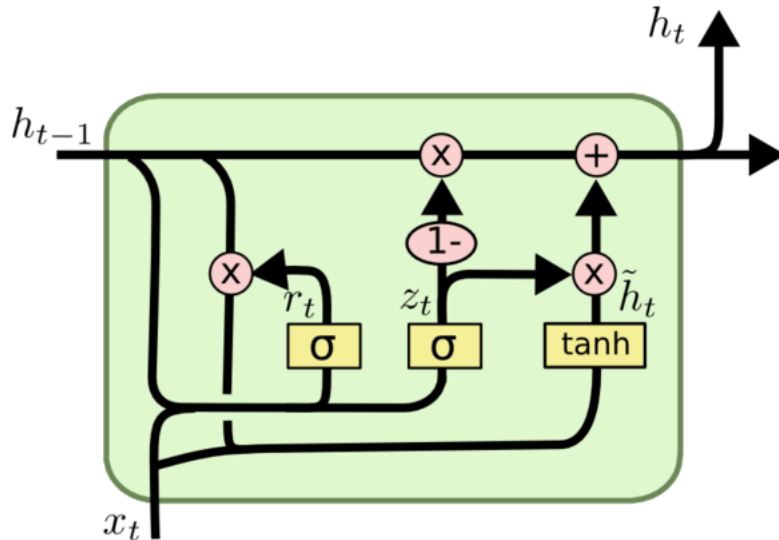
$$o_t = \sigma(W^{(o)}[h_{t-1}, x_t] + b_o)$$

$$h_t = o_t * \tanh C_t$$

GRU (Gated Recurrent Unit)

Στην εργασία αυτή γίνεται χρήση των GRU layers, τα οποία είναι μηχανισμοί των RNN δικτύων. Πρόκειται ουσιαστικά για τη νέα γενιά των RNN δικτύων και έχει

πολλές ομοιότητες με το lstm. Έχει λιγότερες παραμέτρους από το lstm αφού δεν έχει πύλη εξόδου, έχει μόνο reset και update πύλες, μπορεί να θεωρηθεί όμως το ίδιο, αν όχι πιο αποτελεσματικό από το τελευταίο.



Σχήμα 2.8: GRU μηχανισμός, [10]

Ομοίως με το LSTM, το GRU έχει τα εξής κομμάτια που το συνθέτουν:

- **Update Gate:** Η πύλη update για κάθε χρονική στιγμή t υπολογίζεται από τον τύπο

$$z_t = \sigma(W^{(z)}x_t + U^{(z)}h_{t-1})$$
 Έτσι λοιπόν η είσοδος πολλαπλασιάζεται με τα βάρη W ενώ το update γίνεται συνάρτηση του προηγούμενου cell state h_{t-1} . Η πύλη αυτή βοηθάει το μοντέλο να μάθει πόση από την πρότερη πληροφορία χρειάζεται να μεταβιβαστεί σε μελλοντικές καταστάσεις.
- **Reset Gate:** Η πύλη Reset λειτουργεί με βάση την παρακάτω εξίσωση

$$r_t = \sigma(W^{(r)}x_t + U^{(r)}h_{t-1})$$
 Η διαφορά με την παραπάνω πύλη έγκειται στα βάρη που χρησιμοποιούνται για το x και το h_{t-1} .

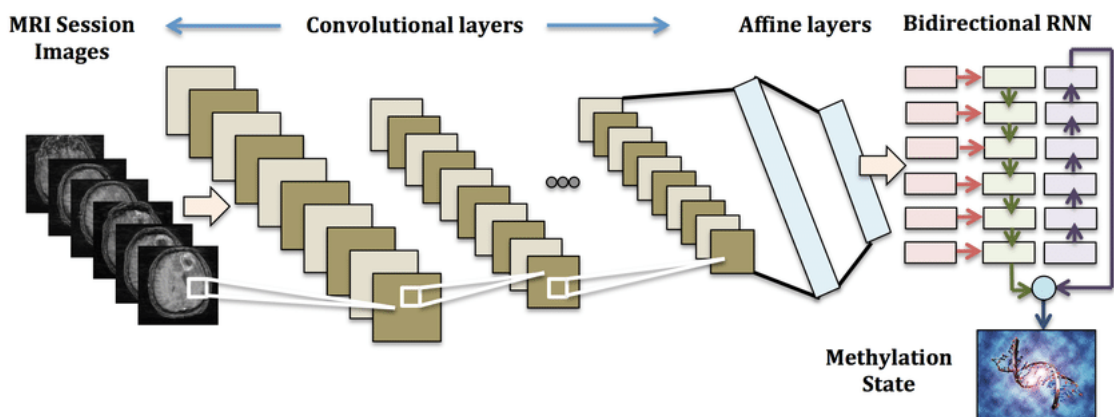
Το \tilde{h}_t είναι ενδιάμεση κατάσταση που προκύπτει μέσα στην δομή του GRU. Από αυτήν και την προηγούμενη κατάσταση κελιού προκύπτει η νέα κατάσταση. Συνολικά, οι 2 εξισώσεις καταστάσεων φαίνονται στη συνέχεια.

$$\tilde{h}_t = \tanh(W^{(h)}x_t + U^{(h)}h_{t-1} * r_t)$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t$$

2.4.4 CNN-RNN

Οι παραπάνω δομές CNN - RNN είναι δυνατό να συνδυαστούν ώστε να κάνουμε αποδοτικές προβλέψεις. Συγκεκριμένα, το CNN δίνει αντιπροσωπευτικά χαρακτηριστικά για κάθε εικόνα εισόδου. Τα δεδομένα αυτά τροφοδοτούνται στη συνέχεια στο RNN ή LSTM όπου γίνονται προβλέψεις για μια σειρά εικόνων που δίνονται στην είσοδο. Έτσι μπορεί να γίνει πρόβλεψη για ένα series αξονικής τομογραφίας ενός συγκεκριμένου ατόμου, πιθανού ασθενή.

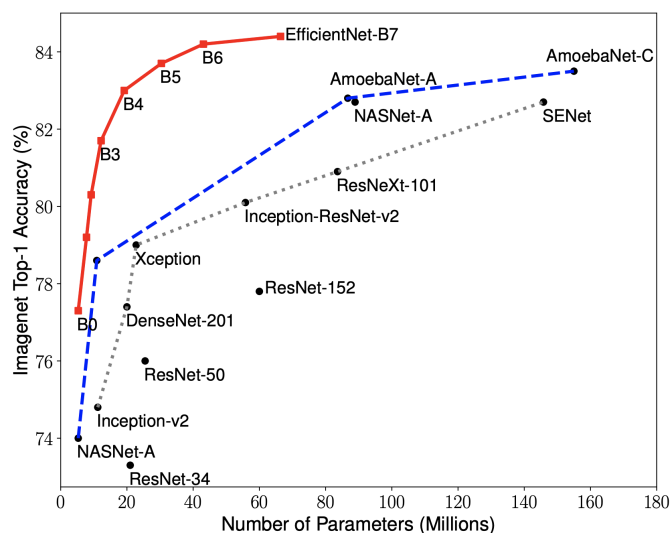


Σχήμα 2.9: Παράδειγμα δομής CNN-RNN, [12]

2.5 Αρχιτεκτονικές Δικτύων

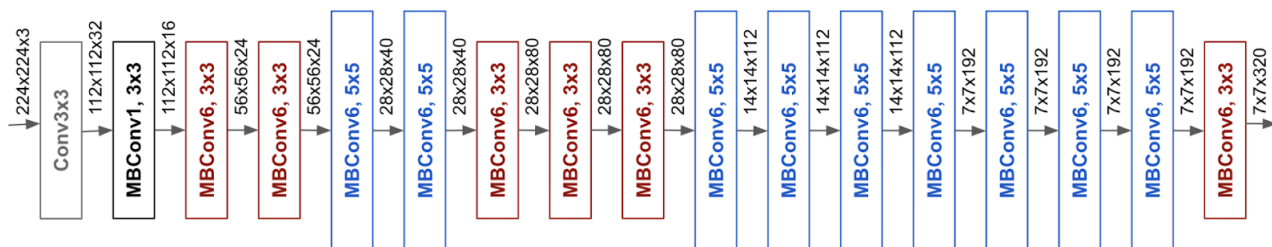
2.5.1 Efficient-Net

Το EfficientNet εισήχθη στην επιστημονική κοινότητα από τους Tan και Le το 2019 και είναι από τα πιο αποτελεσματικά μοντέλα στο imagenet καθώς και σε κλασικά προβλήματα κατηγοριοποίησης εικόνων.



Σχήμα 2.10: Επίδοση των παραλλαγών Efficient-Net, [19]

Το EfficientNet περιλαμβάνει στην λογική του έναν ευριστικό τρόπο για να κλιμακώνει το μοντέλο. Παρέχει έτσι μια οικογένεια από μοντέλα (B0 μέχρι B7) που αποτελεί έναν καλό συνδυασμό αποδοτικότητας και ακρίβειας σε διάφορες κλίμακες. Το βασικό μοντέλο B0 επιτρέπει να ξεπερνά άλλα παρόμοια μοντέλα αποφεύγοντας grid-search των υπερπαραμέτρων.



Σχήμα 2.11: Δομή της αρχιτεκτονικής Efficient-Net, [19]

Αξίζει επίσης να αναφερθούν δύο ακόμα δομές οι οποίες χρησιμοποιήθηκαν σε συγγενείς με την παρούσα διπλωματική εργασίες.

2.5.2 DenseNet

Το DenseNet [8] αποτέλεσε μια σημαντική δομή στην προσπάθεια δημιουργίας όλο και μεγαλύτερων βαθιών νευρωνικών δικτύων. Ξεπέρασε τα εκατοντάδες επίπεδα σε αριθμό, έχοντας όμως πιο στενή δομή σε σχέση με άλλα γνωστά δίκτυα. Αυτό συντέλεσε στην μείωση των υπερπαραμέτρων προς εκμάθηση, πετυχαίνοντας ταυτόχρονα εξαιρετικές επιδόσεις σε γνωστά datasets όπως το ImageNet.

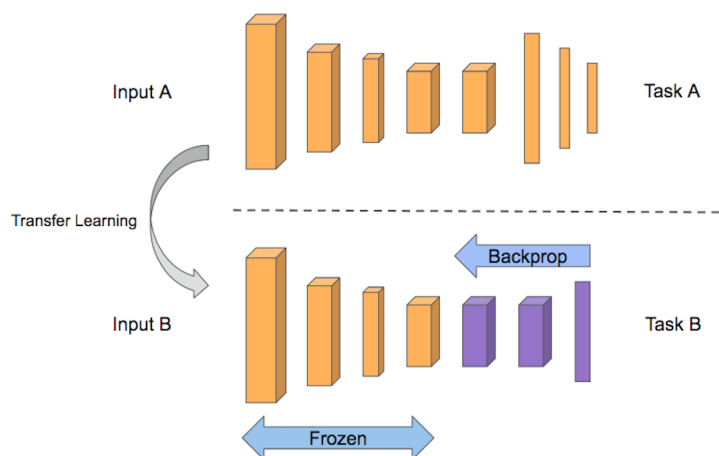
2.5.3 ResNet

Το ResNet [6] είναι μια άλλη γνωστή δομή προεκπαιδευμένου δικτύου. Επιλύει αποδοτικά το πρόβλημα της εκμηδένισης κλίσης (vanishing gradient) το οποίο προκύπτει σε βαθιά νευρωνικά δίκτυα στα οποία λόγω του μεγάλου βάθους χάνονται οι παράγωγοι καθώς γίνονται οι συνεχείς υπολογισμοί του στο backpropagation.

2.6 Μεταφορά Μάθησης (Transfer Learning)

Η ιδέα είναι εμπνευσμένη από την ικανότητα του ανθρώπου να προσαρμόζεται σε νέες προκλήσεις χρησιμοποιώντας την εμπειρία του σε άλλα παρόμοιου τύπου έργα. Όπως ακριβώς ο ανθρώπινος εγκέφαλος δεν μαθαίνει τις εκάστοτε δεξιότητές του από την αρχή, έτσι και η μεταφορά μάθησης εφαρμόζεται πλέον με επιτυχία στην τεχνητή νοημοσύνη.

Η μεταφορά μάθησης (transfer learning) είναι μία μέθοδος που συναντάται συχνά στην βαθιά μάθηση και αξιοποιεί το γεγονός ότι κάποιες συγγενείς εφαρμογές έχουν παρόμοιους στόχους σε προβλήματα μηχανικής μάθησης. Συγκεκριμένα, ένα μοντέλο μπορεί να μεταφέρει τη γνώση που έχει αποκτήσει από ένα άλλο πρόβλημα και ένα σύνολο εικόνων στα βάρη του δικτύου του. Στη συνέχεια, γίνεται προσαρμογή του μοντέλου αυτού στις απαιτήσεις του δεδομένου προβλήματος σε αντιπαράθεση με την ιδέα να εκπαιδευτεί ένα δίκτυο αναγκαστικά από την αρχή στα δεδομένα που αφορούν μια πιο συγκεκριμένη εφαρμογή. Έτσι αποφεύγεται η αποτυχία μάθησης με το πρόβλημα υπερεκπαίδευσης ενός πολύπλοκου μοντέλου CNN με λίγα δεδομένα.



Σχήμα 2.12: Μεταφορά γνώσης από το task A στο B

Η γνώση αυτή είναι ουσιαστικά βάρη προεκπαιδευμένου δικτύου ή αναπαραστάσεις

συγκεκριμένων επιπέδων, ενώ ως μέθοδος μπορεί να εφαρμοστεί σε προβλήματα για τα οποία διαθέτουμε λιγότερα δεδομένα. Η μέθοδος που ακολουθείται για την επιλογή αυτών των επιπέδων του δικτύου ποικίλει, καθώς σε κάθε στρώμα ενός βαθιού νευρωνικού δικτύου, το μοντέλο μαθαίνει διαφορετικά χαρακτηριστικά των δεδομένων εισόδου που είναι εικόνες στην περίπτωσή μας. Έτσι, συνήθως χρησιμοποιούνται προεκπαιδευμένα δίκτυα χωρίς το τελικό επίπεδο, ή γίνεται συγκεκριμένη επιλογή των επιπέδων που θα συνεχίσουν να εκπαιδεύονται στο νέο πρόβλημα κατηγοριοποίησης δεδομένων.

ImageNet

Το ImageNet είναι ένα σύνολο δεδομένων βασισμένο σε συγκεκριμένη ιεραρχία λέξεων κάθε κόμβος της οποίας περιλαμβάνει χιλιάδες εικόνες. Αυτή η βάση δεδομένων και η ποικιλία των εικόνων της βρίσκει μεγάλη εφαρμογή στη μηχανική μάθηση και συγκεκριμένα στην προεκπαίδευση μοντέλων με σκοπό την καλύτερη επίδοσή τους σε εφαρμογές κατηγοριοποίησης εικόνων. Η συλλογή αυτή συγκεντρώθηκε μετά από σημαντική προσπάθεια του πανεπιστημίου του Stanford, ενώ υπάρχει και αντίστοιχο challenge κατά το οποίο το μοντέλο προς εξέταση καλείται να αναγνωρίσει την κατηγορία στην οποία ανήκει το εικονιζόμενο αντικείμενο.

Κεφάλαιο 3

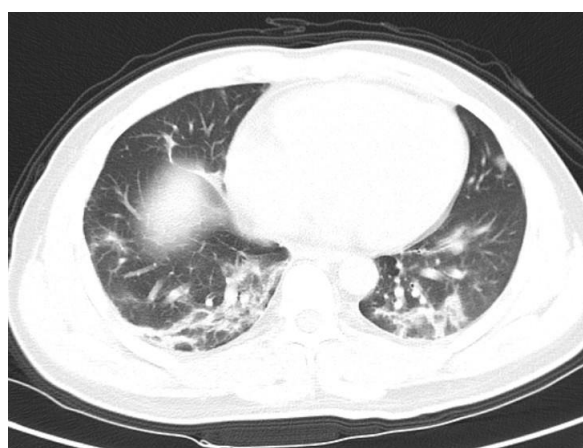
Σύνολα δεδομένων

3.1 Σύνολα Δεδομένων (Datasets)

Κατά την περίοδο έξαρσης του ιού Covid-19, ολόκληρη η επιστημονική κοινότητα κινητοποιήθηκε με σκοπό την απόκτηση ιατρικών δεδομένων σχετικά με τον κορονοϊό. Σε αυτή την προσπάθεια, συγκεντρώθηκαν τομογραφίες θώρακος ασθενών και μη από όλη την Ελλάδα, ενώ χρησιμοποιήθηκε και ένα dataset για την εκπαίδευση μοντέλων και την εξαγωγή προεκπαιδευμένων δικτύων.

3.1.1 Βασικό dataset Chest CTs

Αρχικά, η εργασία βασίστηκε στο dataset CT εικόνων [34]. Το σύνολο αυτό περιλαμβάνει 349 εικόνες από ένα σύνολο 216 ασθενών Covid. Περιλαμβάνει επίσης 463 εικόνες δείγματα από διάφορες πηγές για το σύνολο των υγιών ατόμων που αριθμούνται σε 55. Είναι open-source και διαθέτει και metadata.



Σχήμα 3.1: Τομογραφία ασθενή Covid-19, [34]

Οι εικόνες αυτές έχουν διαφορετικά μεγέθη. Η ελάχιστη, μέση και μέγιστη τιμή του ύψους των διαστάσεων είναι 153, 491, 1853. Για το πλάτος τα αντίστοιχα μεγέθη είναι 124, 383, 1485.

Από τους ασθενείς Covid, οι αρσενικοί είναι περισσότεροι από τους θηλυκούς με αντιστοιχία 86 προς 51 σε αριθμό.

Για τις περιπτώσεις που είναι υγιείς, οι εικόνες συγκεντρώθηκαν από άλλα σύνολα δεδομένων όπως το Luna, MedPix, PMC, Radiopaedia. Συνολικά συγκεντρώθηκαν 463 εικόνες από 55 υγιή άτομα για το σύνολο εκπαίδευσης (training set) όσον αφορά τα υγιή δείγματα.

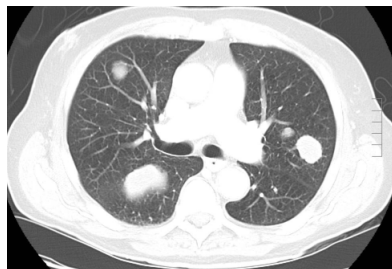
Συνολικά, μέρος αυτών χρησιμοποιήθηκαν για το σύνολο επαλήθευσης (validation set), όπως επίσης και για το σύνολο δοκιμών (test set). Τα στατιστικά για το κάθε σύνολο φαίνονται στους παρακάτω πίνακες.

	άτομα	εικόνες
covid	4	173
non-covid	20	168

Πίνακας 3.1: Κατανομή test συνόλου

	άτομα	εικόνες
covid	4	39
non-covid	88	64

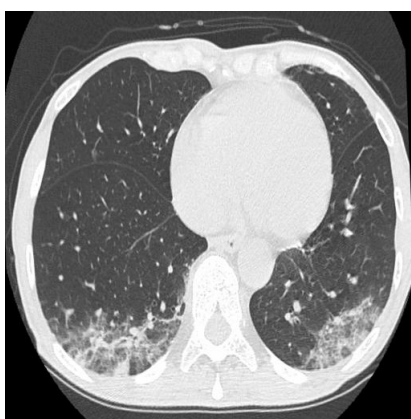
Πίνακας 3.2: Κατανομή validation συνόλου



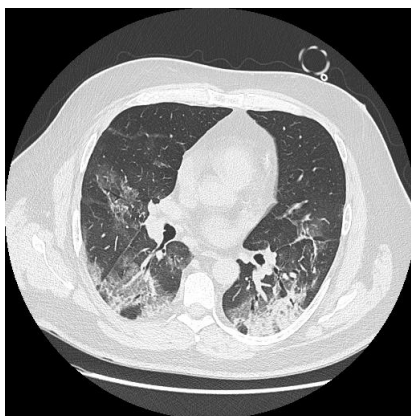
Σχήμα 3.2: Τομογραφία υγιούς ανθρώπου, [34]

Εικόνες Ιταλών ασθενών

Ως συμπληρωματικό υλικό στο προηγούμενο σύνολο δεδομένων, ελήφθη υπόψη και αυτή η σειρά 110 εικόνων [29]. Αρχικός στόχος του συνόλου αυτού είναι για χρήση κατάτμησης (segmentation), μια και οι εικόνες έχουν καταταμηθεί σε περιοχές ενδιαφέροντος από εξειδικευμένους ιατρούς. Αποτελείται έτσι από 100 εικόνες για σκοπούς εκπαίδευσης μοντέλων και επιπλέον 10 εικόνες για testing σκοπούς. Οι εικόνες αυτές είναι όλες από παραπάνω από 40 ασθενείς Covid-19 προερχόμενες από την Ιταλική Ραδιολογική Υπηρεσία (SIRM) και ήταν από τα πρώτα προσβάσιμα σύνολα δεδομένων που εμφανίστηκαν.



Σχήμα 3.3: Παράδειγμα 1 από Ιταλικό dataset



Σχήμα 3.4: Παράδειγμα 2 από Ιταλικό dataset

3.1.2 Ελληνικό Σύνολο Δεδομένων Chest CTs

Χάρη στην ΕΔΥΤΕ (GRNet) έγινε δυνατή η μελέτη τομογραφιών θώρακος ατόμων, ασθενών και μη ασθενών, από πολλά νοσοκομεία της χώρας. Συγκεκριμένα, μέσω

ενιαίου διασυνδεδεμένου συστήματος, έγινε δυνατή η πρόσβαση σε πλήθος παλαιών τομογραφιών αλλά και νέων τομογραφιών θώρακος ασθενών Covid, και μη, που επέτρεψε την σημαντική συνεισφορά στην γενικότερη προσπάθεια που γίνεται για διεύρυνση των συνόλων δεδομένων παγκοσμίως, αλλά και για την δημιουργία αξιόπιστων συστημάτων ανίχνευσης του Covid-19 εντός και εκτός χώρας.

Αναλυτικότερα, η ελληνική βάση δεδομένων αποτελείται από τα παρακάτω μέρη. Οι πηγές κάθε τμήματος του συνόλου αναφέρονται ως αριθμημένα νοσοκομεία.

- 6 πρόσφατες σειρές τομογραφιών Covid και 4 πρόσφατες σειρές Non Covid. Από τις Covid σειρές, 4 είναι από το νοσοκομείο 1 και 2 από το νοσοκομείο 2.
- 100 σειρές Non Covid από το νοσοκομείο 1
- 92 σειρές Non Covid από το νοσοκομείο 2
- 100 σειρές Non Covid από το νοσοκομείο 3
- 100 σειρές Non Covid από το νοσοκομείο 4
- 100 σειρές Non Covid από το νοσοκομείο 5

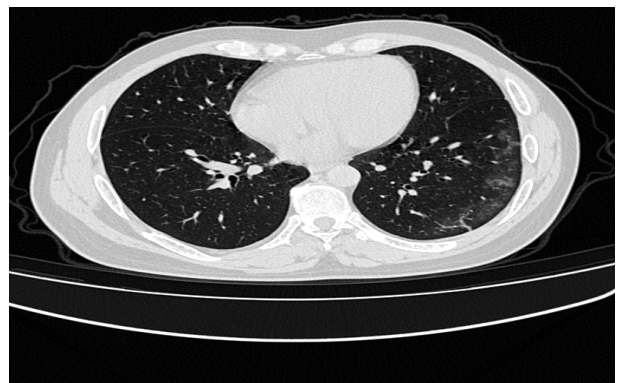
Προσθήκη επιπλέον περιπτώσεων Covid

Προκειμένου να δημιουργηθεί ένα ισορροπημένο dataset από Covid και Non Covid cases είναι απαραίτητο να γίνει επαύξηση (augmentation) των παραπάνω 6 περιπτώσεων Covid που έχουν συλλεχθεί από ελληνικά Νοσοκομεία.

Για το λόγο αυτό στην παρούσα εργασία γίνεται εν μέρει και χρήση του dataset [22] από το Zenodo, το οποίο αποτελείται από 20 ολοκληρωμένες Covid σειρές (series) ξένων ασθενών.



(a) Κεντρική τομή (slice) ασθενή 1



(b) Κεντρική τομή (slice) ασθενή 2

Σχήμα 3.5: Δείγματα Covid από Zenodo

3.2 Μετρικές Επίδοσης

Στα παραπάνω σύνολα δεδομένων υπάρχουν 2 κατηγορίες εικόνων, αυτές που αντιστοιχούν σε ασθενείς οι οποίοι είναι θετικοί (positive) στον Covid-19 σε αυτούς που είναι αρνητικοί (negative) στον ιό.

Προβλέψεις μπορούν να γίνουν σε μια συγκεκριμένη εικόνα, είτε σε μια σειρά εικόνων στην περίπτωση μιας συνολικής αξονικής τομογραφίας θώρακος ενός ατόμου. Σε κάθε περίπτωση, η επίδοση κάθε μοντέλου που επεξεργάζεται τα δεδομένα φαίνεται σε συγκεκριμένες μετρικές τις οποίες αναλύουμε στη συνέχεια. Η κατανόηση αυτών είναι απαραίτητη για την ερμηνεία των αποτελεσμάτων.

Με βάση την πρόβλεψη που έχει γίνει σε ένα αντικείμενο (subject - σειρά εικόνων ή μια εικόνα), δηλαδή αν είναι θετική ή αρνητική (Positive-Negative) και την αρχική κατηγορία στην οποία ανήκει κάθε δεδομένο, δηλαδή αν έγινε σωστή πρόβλεψη ή όχι (True-False), το σύνολο χωρίζεται σε 4 κατηγορίες που είναι οι εξής:

- **True Positive** : Οι ασθενείς Covid-19 που διαγνώστηκαν σωστά ότι όντως έχουν την ασθένεια.
- **False Positive** : Πρόκειται για περιπτώσεις που διαγνώστηκαν με Covid-19 αλλά στην πραγματικότητα δεν είναι ασθενείς.
- **True Negative** : Είναι δεδομένα στα οποία έγινε σωστά η πρόβλεψη ότι δεν υπάρχει Covid-19
- **False Negative** : Σε αυτή την κατηγορία ανήκουν οι ασθενείς που είναι θετικοί στον Covid-19, αλλά διαγνώστηκαν λανθασμένα ως αρνητικοί

	Predicted Negative	Predicted Positive
Non - Covid - 19	TN	FP
Covid - 19	FN	TP

Με βάση αυτές τις απλές μεταβλητές ορίζονται οι παρακάτω μετρικές επίδοσης.

- **Accuracy** : Είναι ποσοστό των σωστών προβλέψεων από τις συνολικές.

$$accuracy = \frac{TP+TN}{TP+FP+TN+FN}$$
- **Precision** : Είναι το ποσοστό των ασθενών που προβλέφθηκαν ως Covid στους οποίους έγινε σωστή πρόβλεψη $precision = \frac{TP}{TP+FP}$
- **Recall** : Είναι το ποσοστό των πραγματικών ασθενών Covid στους οποίους έγινε σωστή πρόβλεψη $recall = \frac{TP}{TP+FN}$
- **F1 score** : Η μετρική αυτή συνδυάζει τις μετρικές recall, precision και αποτελεί καλό μέτρο για να εκτιμηθεί η ισορροπία ανάμεσα στις δύο μεταβλητές

$$F1score = \frac{2*precision*recall}{precision+recall}$$

Φυσικά κάθε μετρική μπορεί να εφαρμοστεί σε κάθε κάθε κατηγορία ξεχωριστά, ιδιαίτερα αν πρόκειται για πρόβλημα πολλαπλών κλάσεων. Σε τέτοιες περιπτώσεις συναντώνται οι παρακάτω ορισμοί των μετρικών, οι οποίοι χρησιμεύουν στο να λάβουμε ενδεικτικές τιμές των μετρικών για το σύνολο των κλάσεων και δεδομένων.

- **micro** : Οι μετρικές υπολογίζονται στο σύνολο των TP, TN, FN, FP κάθε κατηγορίας και αθροίζονται συνολικά για τον τελικό υπολογισμό.
- **macro** : Οι μετρικές υπολογίζονται για κάθε κλάση ξεχωριστά και ύστερα συνυπολογίζονται σε έναν μη σταθμισμένο μέσο.
- **weighted** : Πρόκειται για την ίδια διαδικασία με την macro μέθοδο αλλά αυτή τη φορά ο μέσος είναι σταθμισμένος ανάλογα με το πόσα στοιχεία έχει κάθε κλάση. Λαμβάνεται έτσι υπόψη η ανισορροπία στο πλήθος των στοιχείων κάθε κλάσης του συνόλου δεδομένων.

Κεφάλαιο 4

Συγγενείς Εργασίες

Το κεφάλαιο αυτό συγκεντρώνει σχετικές εργασίες πάνω στις οποίες βασίστηκε η συγγραφή του εν λόγω έργου. Μελετήσαμε έτσι και επεκτείναμε τις εργασίες των [14], [11],[16], [24], [18], [33] πάνω σε παρόμοιες έρευνες ιατρικής θεματολογίας και των [13],[35],[7],[2],[3],[4],[5],[23],[15],[24],[9] για ανάπτυξη αξιόπιστων DNN, αλλά και μοντέλων πρόβλεψης DNN-RNN.

4.1 Βαθιά μάθηση σε ιατρικά δεδομένα

4.1.1 Clustering

Μία βασική μέθοδος που χρησιμοποιείται στην εργασία είναι η εξαγωγή χαρακτηριστικών αναπαραστάσεων του συνόλου δεδομένων μέσα από ένα DNN, όπως έχει γίνει σε παρόμοια έρευνα σε ιατρικές εικόνες για την ανίχνευση Parkinson σε εγκεφαλικά MRIs [35], [11], [18], [33], αλλά και γενικότερα σε νευροεκφυλιστικές ασθένειες όπως το Alzheimer [14].

Συγκεκριμένα ένα DNN είναι ικανό να κάνει προβλέψεις για την είσοδο (input) που του δίνεται. Ωστόσο αυτό λειτουργεί χωρίς να έχουμε ιδέα για την διαδικασία με την οποία γίνεται ο διαχωρισμός. Με σκοπό την καλύτερη κατανόηση της κατανομής που έχουν τα δεδομένα μας στον χώρο του προβλήματος, εξάγουμε πληροφορία για κάθε εικόνα που εισάγεται ως είσοδος στο DNN αφαιρώντας τις τιμές ενός συγκεκριμένου layer στα τελικά στρώματα του δικτύου.

Έτσι λαμβάνουμε αξιόπιστες αναπαραστάσεις μίας διάστασης για κάθε εικόνα, με βάση τις οποίες μπορούμε να εφαρμόσουμε τον αλγόριθμο K-means στο σύνολο των δεδομένων εκπαίδευσης και να πάρουμε αντιπροσωπευτικά κέντρα στα οποία συγκεντρώνονται οι εικόνες. Αυτό μας δίνει μια γενικότερη εικόνα για το πώς ομαδοποιούνται οι εικόνες και σε ποιες ομάδες covid ή non-covid εικόνων είναι πιο κοντά.

Μάλιστα η συγκεκριμένη μέθοδος επιτρέπει, μέσω της κατάλληλης επιλογής του πλήθους των κέντρων κατά την εκτέλεση του αλγορίθμου K-means, να πετύχουμε ακόμα καλύτερα αποτελέσματα, βελτιστοποιώντας τις μετρικές επίδοσης όταν εφαρ-

μόζουμε Nearest-Neighbour στο σύνολο εικόνων test set.

4.1.2 Πρόβλεψη σε ακολουθία δεδομένων

Στα εγκεφαλικά MRI, τα δεδομένα είναι τρισδιάστατα και προκύπτουν έτσι ακολουθίες εικόνων για κάθε πιθανό ασθενή. Χρησιμοποιούνται έτσι δομές αναδρομικών δικτύων (lstm-gru), όπως στο [23] τα οποία συνδυαζόμενα με CNN για την εξαγωγή χαρακτηριστικών των δεδομένων εξάγουν αξιόπιστες προβλέψεις για το σύνολο της εξέτασης ενός δείγματος [14].

Τέτοιες μέθοδοι μπορούν να εφαρμοστούν σε ποικιλία προβλημάτων που αφορούν υγειονομική φροντίδα όπως έχει γίνει για τις ασθένειες Parkinson και Alzheimer [16], [14] και χρησιμοποιούνται στην παρούσα εργασία για την ανάλυση αξονικών τομογραφιών και την ανίχνευση Covid-19.

4.2 Εκπαίδευση CNN σε σύνολο δεδομένων από Chest CTs

Με βάση το dataset των 349 εικόνων Covid που περιγράφηκε αναλυτικά στο προηγούμενο κεφάλαιο, είναι σημαντικό να αναφερθούν τα αρχικά πειράματα που έγιναν από τους δημιουργούς του, τα οποία αποτέλεσαν σημαντική βάση για την έναρξη της παρούσας έρευνας όπως επίσης και μέτρο σύγκρισης για τα αποτελέσματα που εξήχθησαν από τα μοντέλα της εργασίας.

Ως προεπεξεργασία των δεδομένων, οι εικόνες τροποποιήθηκαν όλες σε διαστάσεις 480x480. Για το σύνολο εκπαίδευσης γίνεται με πιθανότητα 0.5 αποκοπή των εξωτερικών μερών της εικόνας, γίνεται επίσης οριζόντια αντιστροφή, τυχαία προσαρμογή φωτεινότητας και αντίθεσης με πιθανότητα 0.2.

Όσον αφορά την αρχιτεκτονική, τα βάρη βελτιστοποιούνται με τον αλγόριθμο Adam, με αρχικό ρυθμό εκμάθησης (learning rate) ίσο με 0.0001 και μέγεθος του batch ίσο με 16. Η αρχιτεκτονική έγινε σε pytorch.

Για τα αποτελέσματα πάνω στα test δεδομένα, χρησιμοποιήθηκαν δύο προεκπαιδευμένες δομές, το DenseNet και το ResNet. Συγκεκριμένα για αυτά έχουμε:

	Accuracy	F1 score	AUC
DenseNet	79.5	76	90.1
ResNet	77.4	74.6	86.4

Πίνακας 4.1: Προγενέστερα Αποτελέσματα σε ResNet, DenseNet

Κεφάλαιο 5

Υλοποίηση και Πειράματα

5.1 Επαύξηση Δεδομένων - Μετασχηματισμοί

5.1.1 Εισαγωγή

Βασισμένοι στον κώδικα [21] και στο σύνολο δεδομένων [34] εκπαιδεύεται ποικιλία μοντέλων με σκοπό την εξαγωγή στατιστικών και την σύγκριση της απόδοσης αυτών.

Για τον λόγο αυτό, δοκιμάζονται διαφόρων ειδών μετασχηματισμοί ως προεπεξεργασία των δεδομένων και μέθοδος επαύξησής τους. Με βάση τις παρακάτω παρατηρήσεις, συμπεραίνουμε ποιος συνδυασμός μετασχηματισμών βοηθάει το δίκτυο στο να μάθει καλύτερα το σύνολο δεδομένων και να δώσει καλά αποτελέσματα στα άγνωστα δεδομένα του test set. Στα δεδομένα αυτά, έχουμε συνολικά 425 εικόνες εκπαίδευσης, 118 εικόνες για το σύνολο επαλήθευσης και 203 εικόνες προς εξέταση.

Όσον αφορά την αρχιτεκτονική, χρησιμοποιήθηκε ως βάση το προεκπαιδευμένο μοντέλο EfficientNet-b0, τα βάρη βελτιστοποιούνται με τον αλγόριθμο Adam, με αρχικό ρυθμό εκμάθησης (learning rate) ίσο με 0.0001 και μέγεθος του batch ίσο με 20. Η αρχιτεκτονική έγινε σε pytorch.

5.1.2 Αποτελέσματα των μετασχηματισμών

Παρακάτω λοιπόν φαίνονται αναλυτικά, οι σταδιακοί μετασχηματισμοί που εφαρμόστηκαν στα πειράματα και τα αντίστοιχα αποτελέσματα που είχαν αυτοί στο test set των 203 εικόνων.

Αρχικά βλέπουμε τα αποτελέσματα για την αρχική δομή του δικτύου χωρίς να εφαρμόσουμε μετασχηματισμούς στις εικόνες.

	precision	recall	f1-score	support
covid	0.78	0.72	0.75	98
non-covid	0.76	0.81	0.78	105
macro	0.77	0.77	0.77	203
weighted	0.77	0.77	0.77	203
accuracy	0.77			203

Πίνακας 5.1: Μετρικές χωρίς εφαρμογή μετασχηματισμών

Στο σημείο αυτό δοκιμάστηκε να ενσωματωθούν οι 110 εικόνες του Ιταλικού συνόλου δεδομένων [29] στις Covid εικόνες του υπάρχοντος dataset. Τα αποτελέσματα είναι παρόμοια με αυτά της αρχικής δομής.

	precision	recall	f1-score	support
covid	0.81	0.67	0.74	98
non-covid	0.74	0.86	0.79	105
macro	0.78	0.77	0.77	203
weighted	0.77	0.77	0.77	203
accuracy	0.77			203

Πίνακας 5.2: Μετρικές με προσθήκη του Ιταλικού Συνόλου Δεδομένων

Μια σημαντική προσπάθεια βελτιστοποίησης των αποτελεσμάτων είναι η προσθήκη επιπλέον στρωμάτων-επιπέδων στα τελευταία στάδια της δομής. Προσθέτουμε έτσι 2 επιπλέον fully-connected layers που μειώνουν σταδιακά τις διαστάσεις του αριθμού των νευρώνων πρώτα από 1000 σε 32 και έπειτα από 32 σε 2 που είναι και ο αριθμός των κλάσεων. Παρακάτω φαίνονται τα αποτελέσματα με τη δομή για το επαυξημένο με τις 110 εικόνες σύνολο δεδομένων.

	precision	recall	f1-score	support
covid	0.84	0.66	0.74	98
non-covid	0.74	0.89	0.81	105
macro	0.79	0.77	0.77	203
weighted	0.79	0.78	0.78	203
accuracy	0.78			203

Πίνακας 5.3: Μετρικές με Ιταλικό dataset και 2 επιπλέον layers

Γενικά τα αποτελέσματα είναι παρόμοια με τα αρχικά. Το ίδιο παρατηρούμε και αν στη νέα αυτή δομή δοκιμάσουμε το σύνολο δεδομένων χωρίς τις Ιταλικές εικόνες.

	precision	recall	f1-score	support
covid	0.74	0.78	0.76	98
non-covid	0.78	0.74	0.76	105
macro	0.76	0.76	0.76	203
weighted	0.76	0.76	0.76	203
accuracy	0.76			203

Πίνακας 5.4: Αρχικό dataset με 2 επιπλέον layers

Η προσθήκη αυτών των layers δεν φαίνεται να ενισχύει την αποδοτικότητα του δικτύου.

Διατηρώντας έτσι την αρχική δομή δοκιμάζουμε ως μετασχηματισμό την περιστροφή των εικόνων κατά 5 μοίρες.

	precision	recall	f1-score	support
covid	0.89	0.78	0.83	98
non-covid	0.81	0.91	0.86	105
macro	0.85	0.84	0.85	203
weighted	0.85	0.85	0.85	203
accuracy	0.85			203

Πίνακας 5.5: Μετρικές με περιστροφή εύρους 5 μοιρών

Η ελαφριά αυτή κλίση έχει πολύ καλά αποτελέσματα στην ακρίβεια των προβλέψεων. Δοκιμάζεται έτσι στη συνέχεια για περισσότερες μοίρες γωνιακής μετατόπισης και συγκεκριμένα 10.

	precision	recall	f1-score	support
covid	0.82	0.69	0.75	98
non-covid	0.75	0.86	0.80	105
macro	0.78	0.78	0.78	203
weighted	0.78	0.78	0.78	203
accuracy	0.78			203

Πίνακας 5.6: Μετρικές με περιστροφή εύρους 10 μοιρών

Τα αποτελέσματα δεν είναι τόσο καλά όσο πριν. Ωστόσο δοκιμάζουμε και για μεγαλύτερο εύρος γωνιακής μετατόπισης, δηλαδή 180 μοίρες.

	precision	recall	f1-score	support
covid	0.86	0.72	0.78	98
non-covid	0.78	0.89	0.83	105
macro	0.82	0.81	0.81	203
weighted	0.81	0.81	0.81	203
accuracy	0.81			203

Πίνακας 5.7: Μετρικές με περιστροφή εύρους 180 μοιρών

Τα αποτελέσματα είναι βελτιωμένα αλλά όχι τόσο καλά όσο με το εύρος των 5 μοιρών.

Παρακάτω φαίνονται τα αποτελέσματα από την εφαρμογή τυχαίου μετασχηματισμού οριζόντιας αντιστροφής της εικόνας (συμμετρία ως προς τον κάθετο άξονα).

	precision	recall	f1-score	support
covid	0.88	0.59	0.71	98
non-covid	0.71	0.92	0.80	105
macro	0.79	0.76	0.75	203
weighted	0.79	0.76	0.76	203
accuracy	0.76			203

Πίνακας 5.8: Μετρικές με οριζόντια αντιστροφή

Το επόμενο πείραμα αφορά στην εφαρμογή τυχαίας κάθετης αντιστροφής (συμμετρία ως προς τον οριζόντιο άξονα) στις εικόνες ως μετασχηματισμό.

	precision	recall	f1-score	support
covid	0.84	0.78	0.81	98
non-covid	0.81	0.87	0.83	105
macro	0.82	0.82	0.82	203
weighted	0.82	0.82	0.82	203
accuracy	0.82			203

Πίνακας 5.9: Μετρικές με κάθετη αντιστροφή

Ως μετασχηματισμός φαίνεται ικανοποιητικός, καθώς ενισχύει την ακρίβεια των προβλέψεων στον 82 %.

Ως επόμενο κύκλο πειραμάτων, γίνεται προσπάθεια ενσωμάτωσης των επιπλέον 110 εικόνων στο dataset με σκοπό την βελτιστοποίηση των αποτελεσμάτων. Ως νέο επίπεδο στην προεκπαιδευμένη δομή του EfficientNet-b0, αλλάζουμε το τελευταίο πλήρως διασυνδεδεμένο layer του σε διαστάσεις 1280x32, προσθέτουμε ένα επίπεδο ReLU, ενώ προσθέτουμε και το τελικό στάδιο εξαγωγής των προβλέψεων ως ένα πλήρως διασυνδεδεμένο επίπεδο 32x2.

Ως σημείο αναφοράς είναι πλέον τα παρακάτω αποτελέσματα στα οποία έχουν προστεθεί τα νέα layers στο δίκτυο και η ακρίβεια έχει αυξηθεί κατά 2 % σε σχέση με το αρχικό δίκτυο χωρίς τη χρήση μετασχηματισμών.

	precision	recall	f1-score	support
covid	0.78	0.85	0.81	98
non-covid	0.85	0.78	0.81	105
macro	0.81	0.81	0.81	203
weighted	0.81	0.81	0.81	203
accuracy	0.81			203

Πίνακας 5.10: Μετρικές χωρίς μετασχηματισμό και προσθήκη 1 layer

Παρακάτω φαίνονται αποτελέσματα για την οριζόντια αντιστροφή των εικόνων με τα επιπλέον 110 δείγματα και την εφαρμογή της νέας δομής του δικτύου.

	precision	recall	f1-score	support
covid	0.82	0.71	0.77	98
non-covid	0.76	0.86	0.81	105
macro	0.79	0.79	0.79	203
weighted	0.79	0.79	0.79	203
accuracy	0.79			203

Πίνακας 5.11: Μετρικές με προσθήκη Ιταλικού dataset και 1 layer

Παραμένοντας στο πλαίσιο των επιπλέον 110 εικόνων και των νέων επιπέδων, δοκιμάζεται ο μετασχηματισμός των εικόνων με περιστροφή κατά 5 μοίρες και η κάθετη αντιστροφή τους.

	precision	recall	f1-score	support
covid	0.85	0.74	0.79	98
non-covid	0.79	0.88	0.83	105
macro	0.82	0.81	0.81	203
weighted	0.82	0.81	0.81	203
accuracy	0.81			203

Πίνακας 5.12: Μετρικές με περιστροφή 5 μοιρών, κάθετη αντιστροφή, προσθήκη Ιταλικού dataset και 1 επιπλέον layer

Τα αποτελέσματα ξεπερνούν το 80 % αλλά υπάρχει δυσκολία στην περαιτέρω βελτίωσή τους. Στους παραπάνω μετασχηματισμούς προσθέτουμε και την τυχαία οριζόντια αντιστροφή.

	precision	recall	f1-score	support
covid	0.92	0.66	0.77	98
non-covid	0.75	0.94	0.84	105
macro	0.83	0.80	0.80	203
weighted	0.83	0.81	0.80	203
accuracy	0.81			203

Πίνακας 5.13: Μετρικές με περιστροφή 5 μοιρών, κάθετη και οριζόντια αντιστροφή, προσθήκη Ιταλικού dataset και 1 επιπλέον layer

Προκειμένου να επιλεγθούν οι κατάλληλες πράξεις επί των εικόνων δοκιμάζουμε και τον μετασχηματισμό των εικόνων μόνο με τυχαία κάθετη αντιστροφή, όμως ενδεικτικά η ακρίβεια των προβλέψεων έπεσε στο 77 %.

	precision	recall	f1-score	support
covid	0.89	0.60	0.72	98
non-covid	0.72	0.93	0.81	105
macro	0.80	0.77	0.76	203
weighted	0.80	0.77	0.77	203
accuracy	0.77			203

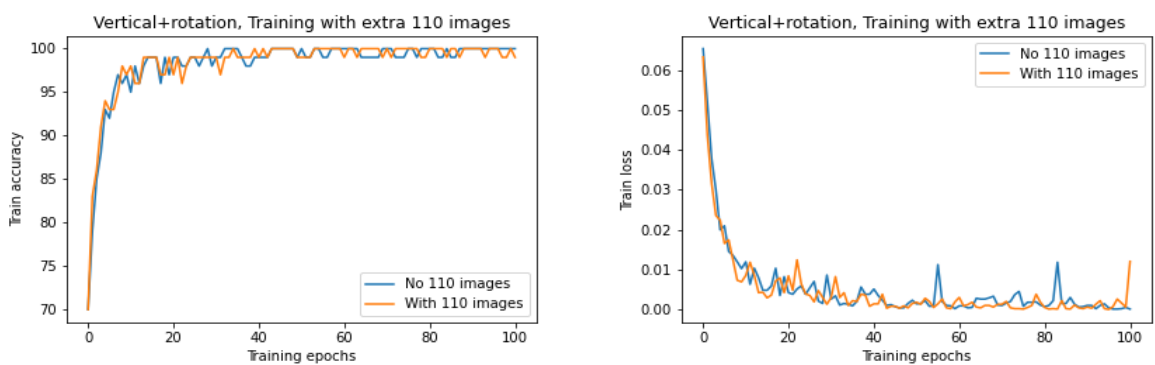
Πίνακας 5.14: Μετρικές με κάθετη αντιστροφή, προσθήκη Ιταλικού dataset και 1 επιπλέον layer

Στον παραπάνω μετασχηματισμό προσθέτουμε και την οριζόντια αντιστροφή, παρατηρώντας έτσι μια ελαφρά βελτίωση.

	precision	recall	f1-score	support
covid	0.85	0.71	0.78	98
non-covid	0.77	0.89	0.82	105
macro	0.81	0.80	0.80	203
weighted	0.81	0.80	0.80	203
accuracy	0.80			203

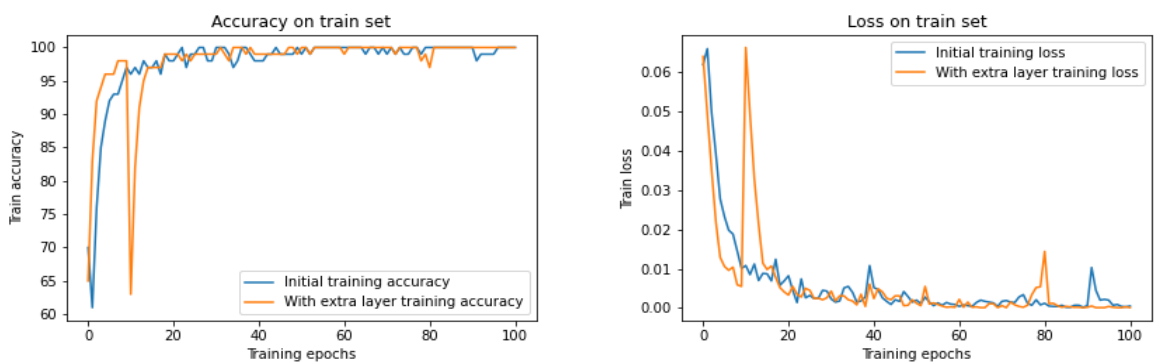
Πίνακας 5.15: Μετρικές με κάθετη και οριζόντια αντιστροφή, προσθήκη Ιταλικού dataset και 1 επιπλέον layer

Συμπερασματικά από τον νέο αυτό κύκλο πειραμάτων μπορούμε να πούμε ότι η προσθήκη των νέων στρωμάτων λειτουργεί θετικά στα αποτελέσματα, ενώ το Ιταλικό σύνολο δεδομένων επιδρά με ελαφρώς επιβαρυντικό τρόπο στην εκπαίδευση του δικτύου. Από τους μετασχηματισμούς ξεχωρίζουν οι οριζόντιες και κάθετες αντιστροφές όπως επίσης και η περιστροφή εύρους 5 μοιρών.



(a) Ακρίβεια train set με και χωρίς το Ιταλικό dataset (b) Κόστος train set με και χωρίς το Ιταλικό dataset

Σχήμα 5.1: Εκπαίδευση με τις 110 Ιταλικές εικόνες και χωρίς



(a) Ακρίβεια train set με και χωρίς το extra layer (b) Κόστος train set με και χωρίς το extra layer

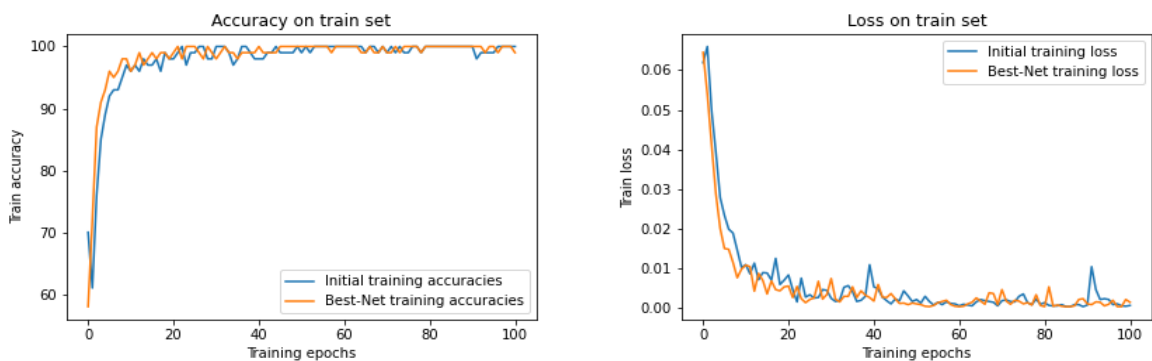
Σχήμα 5.2: Εκπαίδευση με και χωρίς το επιπλέον επίπεδο

Με βάση αυτά, κρατώντας την προσθήκη του νέου επιπέδου και το αρχικό dataset, εφαρμόζουμε συνδυασμό περιστροφής 5 μοιρών και κάθετης αντιστροφής.

	precision	recall	f1-score	support
covid	0.89	0.75	0.81	98
non-covid	0.79	0.91	0.85	105
accuracy	0.83			203

Πίνακας 5.16: Μετρικές με κάθετη αντιστροφή, περιστροφή 5 μοιρών και 1 επιπλέον layer

Τέλος, προκύπτει ότι το καλύτερο μοντέλο με πρότυπες επιδόσεις σε όλες τις μετρικές, είναι όταν εφαρμόζουμε τις 2 πιθανές αντιστροφές μαζί με την προσθήκη του τελικού επιπέδου. Το μοντέλο αυτό αποτελεί σημείο αναφοράς για την χρήση προεκπαιδευμένου μοντέλου σε μελλοντικές εφαρμογές της παρούσας εργασίας βασισμένο στο συγκεκριμένο σύνολο δεδομένων. Στη συνέχεια των πειραμάτων θα συναντάται ως **βασικό μοντέλο**.



(a) Ακρίβεια train set αρχικού και βέλτιστου δικτύου (b) Κόστος train set αρχικού και βέλτιστου δικτύου

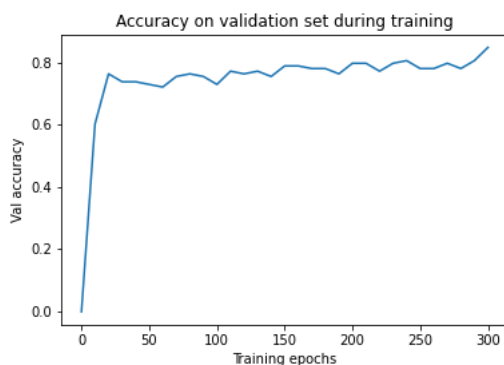
Σχήμα 5.3: Εκπαίδευση βέλτιστου δικτύου σε σχέση με το αρχικό

	precision	recall	f1-score	support
covid	0.88	0.79	0.83	98
non-covid	0.82	0.90	0.85	105
macro	0.87	0.78	0.83	203
accuracy	0.84			203

Πίνακας 5.17: Μετρικές με οριζόντια και κάθετη αντιστροφή και 1 επιπλέον layer

Τα αποτελέσματα αυτά του βασικού μας μοντέλου φαίνονται και στην εργασία

[24], όπου αναφέρεται το f1-score 0.842 (με 0.855 για την μη Covid κλάση και 0.828 για την Covid κλάση).



Σχήμα 5.4: Ακρίβεια πρόβλεψης κατά τη διάρκεια εκπαίδευσης του καλύτερου μοντέλου.

Συγκριτικά, φαίνεται η βελτίωση με πανομοιότυπο μοντέλο αλλά χωρίς την προσθήκη των νέων στρωμάτων.

	precision	recall	f1-score	support
covid	0.79	0.88	0.83	98
non-covid	0.85	0.76	0.80	105
macro	0.79	0.88	0.83	203
accuracy	0.82			203

Πίνακας 5.18: Μετρικές με οριζόντια και κάθετη αντιστροφή

Διαχείριση του Ιταλικού συνόλου δεδομένων

Παρά την δυσκολία βελτιστοποίησης αποτελεσμάτων με την προσθήκη του Ιταλικού συνόλου δεδομένων, επιχειρείται η ενσωμάτωσή τους στην εκπαίδευση των μοντέλων. Ως βάση χρησιμοποιείται το πιο αποδοτικό από τα παραπάνω μοντέλα όπως σημειώθηκε παραπάνω.

Συγκεκριμένα, συλλέγουμε τα δεδομένα τα οποία το μοντέλο μας προβλέπει σωστά. Αυτές είναι 24 covid εικόνες. Τα αποτελέσματα στο test set φαίνονται παρακάτω και είναι ελάχιστα χειρότερα από ό,τι είναι με το βασικό μοντέλο που προέκυψε στα προηγούμενα πειράματα.

	precision	recall	f1-score	support
covid	0.81	0.84	0.82	98
non-covid	0.82	0.79	0.80	105
macro	0.81	0.83	0.82	203
accuracy	0.81			203

Πίνακας 5.19: Μοντέλο με την προσθήκη 24 εικόνων covid

5.1.3 Συσταδοποίηση (Clustering)

Εισαγωγή

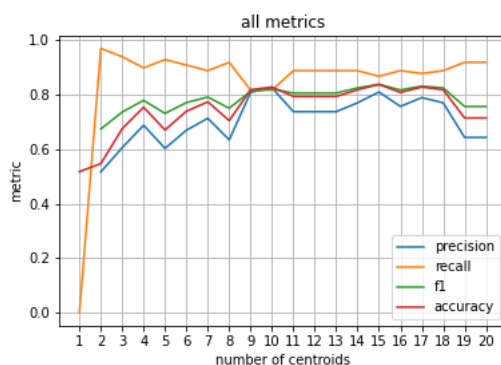
Ως συνέχεια στην επεξεργασία των δεδομένων, από το καλύτερο μοντέλο συλλέγουμε τις αναπαραστάσεις μεγέθους 32 για κάθε εικόνα από το στρώμα που προστέθηκε στην παραπάνω διαδικασία. Με βάση τις training αναπαραστάσεις εφαρμόζουμε τον αλγόριθμο K-Means για την εύρεση k κέντρων όπου το k παίρνει επιλεγμένες τιμές.

Στη συνέχεια βρίσκουμε σε ποια κατηγορία (covid ή non-covid) ανήκει το κάθε cluster, ανάλογα με το ποια κατηγορία έχει την πλειοψηφία των εικόνων που ανήκουν στο cluster σύμφωνα με τον αλγόριθμο K-Means.

Τέλος, κάνουμε πρόβλεψη για κάθε εικόνα το label του κοντινότερου σε αυτήν κέντρου. Με βάση τις νέες προβλέψεις εξετάζουμε πάλι την συμπεριφορά των μετρικών.

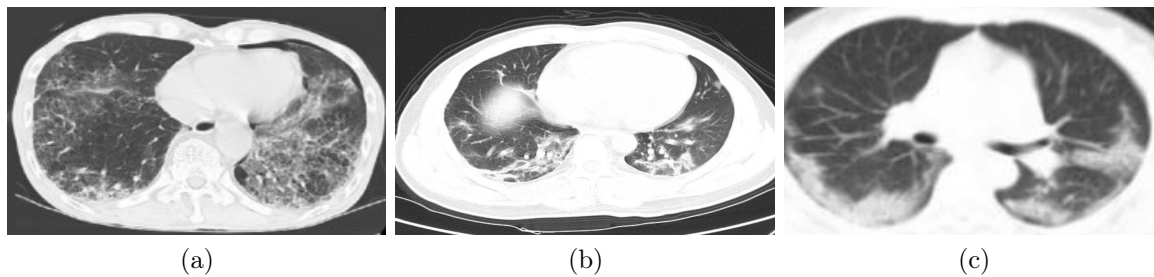
Αρχικό clustering

Παρακάτω φαίνονται οι μετρικές για πλήθος κέντρων από 1 έως 20. Παρατηρούμε ότι τα καλύτερα αποτελέσματα τα λαμβάνουμε για k=10.

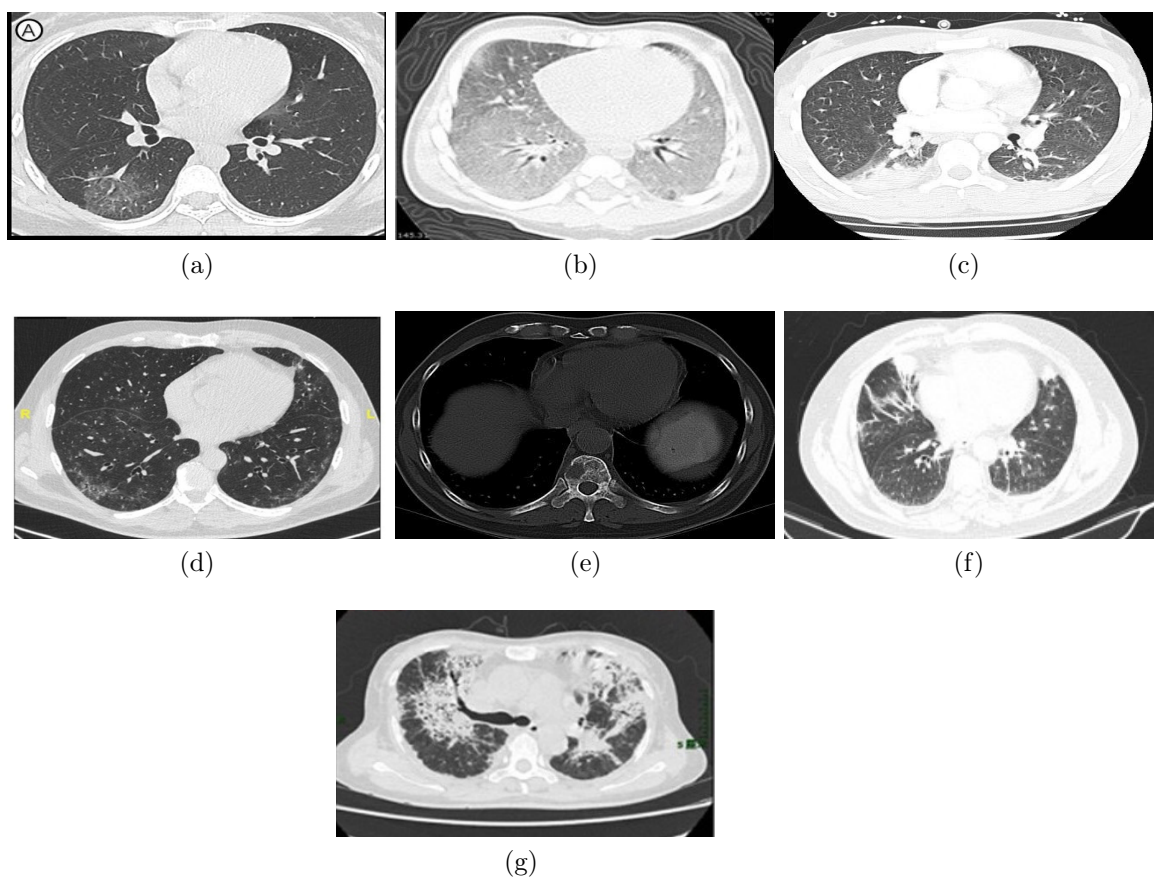


Σχήμα 5.5: Συσταδοποίηση στο αρχικό σύνολο δεδομένων

Οι επιδόσεις είναι ελαφρώς χειρότερες από το αντίστοιχο CNN ωστόσο αποκτούμε πλήρη εικόνα του πώς σχηματίζονται τα clusters.



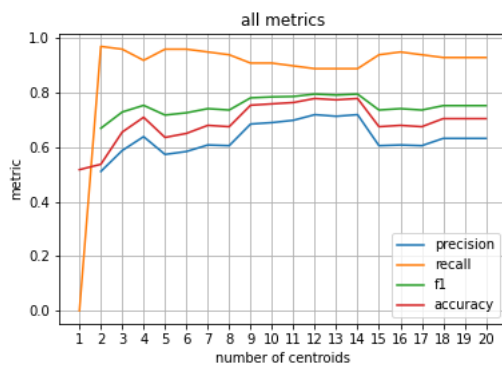
Σχήμα 5.6: Κέντρα Covid για 10 clusters



Σχήμα 5.7: Κέντρα Non-Covid για 10 clusters

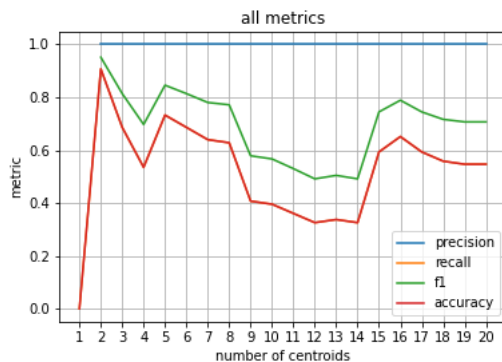
Νέο clustering

Προσθέτουμε στα δεδομένα εκπαίδευσης τις 24 εικόνες που προέβλεψε σωστά το CNN. Πάλι στην καλύτερη περίπτωση τα αποτελέσματα κυμαίνονται σε ακρίβεια 80%.



Σχήμα 5.8: Συσταδοποίηση στο test set μετά την προσθήκη των 24 αναπαραστάσεων

Το πλεονέκτημα της μεθόδου συσταδοποίησης φαίνεται ωστόσο αν εφαρμοστεί στις υπόλοιπες 86 εικόνες του Ιταλικού dataset. Για $k=2$, φτάνουμε σε ακρίβεια μεγαλύτερη του 80%, ενώ το αντίστοιχο CNN που είχε εκπαιδευτεί με τις 24 σωστά προβλεπόμενες εικόνες Covid, είχε ποσοστό ακρίβειας μόλις 32%.



Σχήμα 5.9: Συσταδοποίηση στις υπόλοιπες Ιταλικές εικόνες μετά την προσθήκη των 24 αναπαραστάσεων

5.2 Προβλέψεις σε CT series Ελλήνων ασθενών και μη

Στο κομμάτι αυτών των πειραμάτων, εργαζόμαστε με τα πιο πρόσφατα διαθέσιμα ελληνικά δεδομένα και συγκεκριμένα με τους 6 Covid ασθενείς και τα 4 δείγματα που διαγνώστηκαν ως υγιείς. Στα υγιή δείγματα προσθέτουμε και 2 υγιείς σειρές εικόνων από το σύνολο δεδομένων του νοσοκομείου 1. Χωρίζουμε στη συνέχεια το σύνολο των δειγμάτων σε 8 δείγματα εκπαίδευσης (4 Covid και 4 Non-Covid) και 4 δείγματα για το σύνολο test (2 Covid και 2 Non-Covid).

Data cleaning

Προκειμένου το σύνολο δεδομένων να αποτελείται από καθαρές εικόνες του κεντρικών τομών των πνευμόνων όπου το μοτίβο του κορονοϊού μπορεί να αναγνωρισθεί από το δίκτυο, γίνεται μια εκκαθάριση των σειρών κάθε δείγματος. Συγκεκριμένα, για κάθε σειρά εικόνων σε μια τομογραφία, απορρίπτονται τομές στην αρχή και το τέλος της σειράς, ώστε να κρατηθούν οι κεντρικές εικόνες. Συνολικά τα δεδομένα εκπαίδευσης διαθέτουν έτσι 533 εικόνες Covid και 531 εικόνες Non-Covid. Αντίστοιχα το test set διαθέτει 45 εικόνες Covid και 39 εικόνες Non-Covid, ωστόσο οι προβλέψεις εξετάζονται ανά άτομο.

Για την εκπαίδευση των δικτύων χρησιμοποιούμε είτε το καλύτερο pretrained μοντέλο που έχουμε και το εκπαιδεύουμε με το σύνολο των εικόνων που διαθέτουμε, είτε την αρχική δομή με το προεκπαιδευμένο EfficientNet-b0.

Τα στατιστικά αποτελέσματα που θα εμφανίζονται παρακάτω θα είναι με βάση τους 4 ασθενείς.

5.2.1 fine tuning

Στη μέθοδο fine tuning, αξιοποιούμε το γεγονός ότι το δίκτυο έχει εκπαιδευτεί σε παρόμοιου τύπου δεδομένα. Έτσι στις παρακάτω εφαρμογές, ξεκινώντας από το έτοιμο δίκτυο εκπαιδεύουμε εν συνεχεία στα νέα δεδομένα για λίγες μόνο εποχές.

Ψήφος "τουλάχιστον ένας"

Σε αυτό το είδος πρόβλεψης, βρίσκουμε αρχικά όλες τις προβλέψεις που υπάρχουν σε κάθε εικόνα της σειράς της εκάστοτε τομογραφίας θώρακος. Στη συνέχεια, αν υπάρχει έστω και μία θετική πρόβλεψη, προβλέπουμε covid. Σε αντίθετη περίπτωση προβλέπουμε non-covid.

Δοκιμάζουμε σε αυτό το πλαίσιο να κάνουμε fine tuning εκπαιδύοντας το σύνολο του προεκπαιδευμένου μοντέλου για λίγες μόνο εποχές. Τα αποτελέσματα φαίνονται παρακάτω.

	Σύνολο	Σωστές προβλέψεις	Λάθος προβλέψεις	Ακρίβεια(%)
covid	2	2	0	100
non-covid	2	0	2	0
Συνολικά				50

Πίνακας 5.20: Fine-tuning, end-to-end εκπαίδευση, ψήφος "at least one"

Ως δεύτερη επιλογή εκπαιδύουμε μόνο το layer μεγέθους 32 νευρώνων που έχει προστεθεί στα τελευταία layers του δικτύου. Αυτή η ρύθμιση δεν προβλέπει σωστά τον 1 ασθενή covid.

	Σύνολο	Σωστές προβλέψεις	Λάθος προβλέψεις	Ακρίβεια(%)
covid	2	1	1	50
non-covid	2	2	0	100
Συνολικά				75

Πίνακας 5.21: Fine-tuning, εκπαίδευση ενός layer, ψήφος "at least one"

Ψήφος πλειοψηφίας

Ένας άλλος τρόπος πρόβλεψης σε 3D τομογραφία ενός μεμονωμένου ατόμου είναι να εφαρμόσουμε ψήφο πλειοψηφίας στις προβλέψεις των τομών. Έτσι, γίνονται οι προβλέψεις σε κάθε τομή της σειράς και ανάλογα με το αν υπάρχουν περισσότερες προβλέψεις Covid ή Non Covid, γίνεται και η αντίστοιχη πρόβλεψη για τον πιθανό ασθενή συνολικά.

Λαμβάνοντας ξανά τις δύο περιπτώσεις, παρακάτω φαίνονται τα αποτελέσματα στα 4 δείγματα του test set, για εξολοκλήρου εκπαίδευση του δικτύου ή μόνο του επιπρόσθετου επιπέδου μεγέθους 32 αντίστοιχα.

	Σύνολο	Σωστές προβλέψεις	Λάθος προβλέψεις	Ακρίβεια(%)
covid	2	1	1	50
non-covid	2	1	1	50
Συνολικά				50

Πίνακας 5.22: Fine-tuning, εκπαίδευση end-to-end, ψήφος πλειοψηφίας

	Σύνολο	Σωστές προβλέψεις	Λάθος προβλέψεις	Ακρίβεια(%)
covid	2	1	1	50
non-covid	2	2	0	100
Συνολικά				75

Πίνακας 5.23: Fine-tuning, εκπαίδευση ενός layer, ψήφος πλειοψηφίας

5.2.2 From scratch implementation

Αυτή τη φορά εκπαιδεύουμε ολόκληρο το δίκτυο εξαρχής με τα νέα δεδομένα μας.

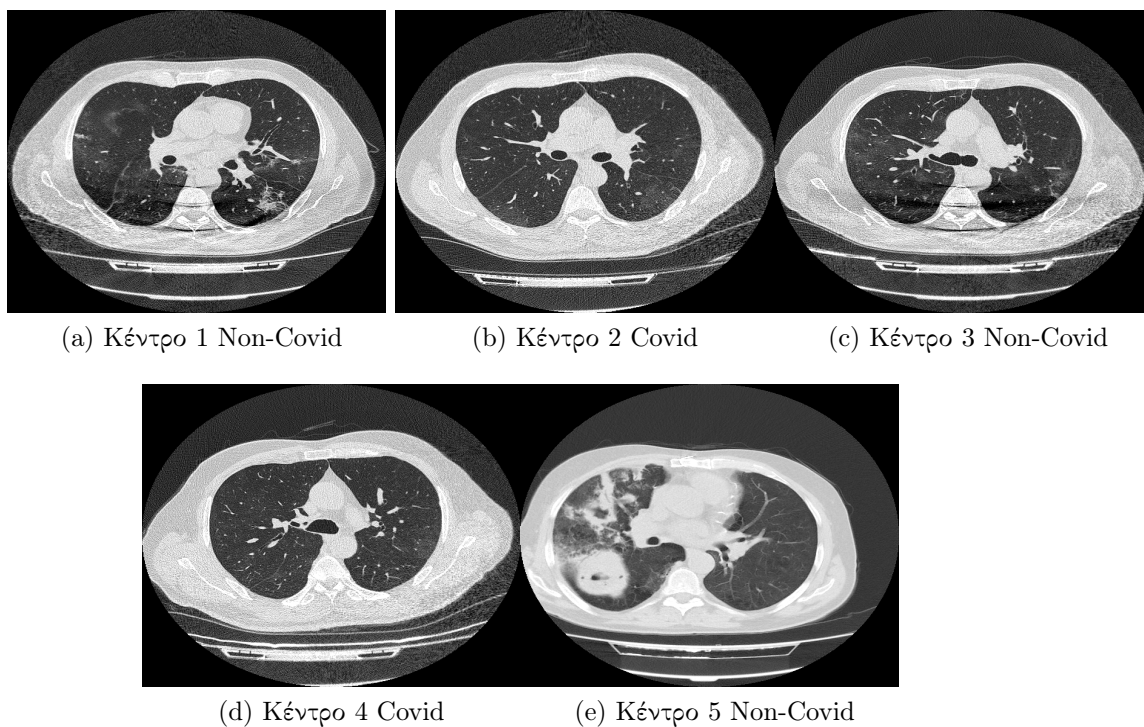
Ψήφος "τουλάχιστον ένας"

Το δίκτυο φαίνεται να προσαρμόζεται καλά σε αυτά, καθώς μετά από 60 εποχές εκπαίδευσης, η μέθοδος "τουλάχιστον 1" δίνει 100% ακρίβεια στους 4 test πιθανούς ασθενείς.

	Σύνολο	Σωστές προβλέψεις	Λάθος προβλέψεις	Ακρίβεια(%)
covid	2	2	0	100
non-covid	2	2	0	100
Συνολικά				100

Πίνακας 5.24: From scratch, εκπαίδευση ενός layer, ψήφος "τουλάχιστον ένας"

Παρακάτω φαίνονται και τα clusters που προκύπτουν αν επιλέξουμε το καλύτερο k για το οποίο βρίσκει όλες τις προβλέψεις των τεσσάρων ατόμων σωστά.



Σχήμα 5.10: Κέντρα για 5 clusters

5.3 Κατάτμηση (segmentation) εικόνων αρχικού Dataset

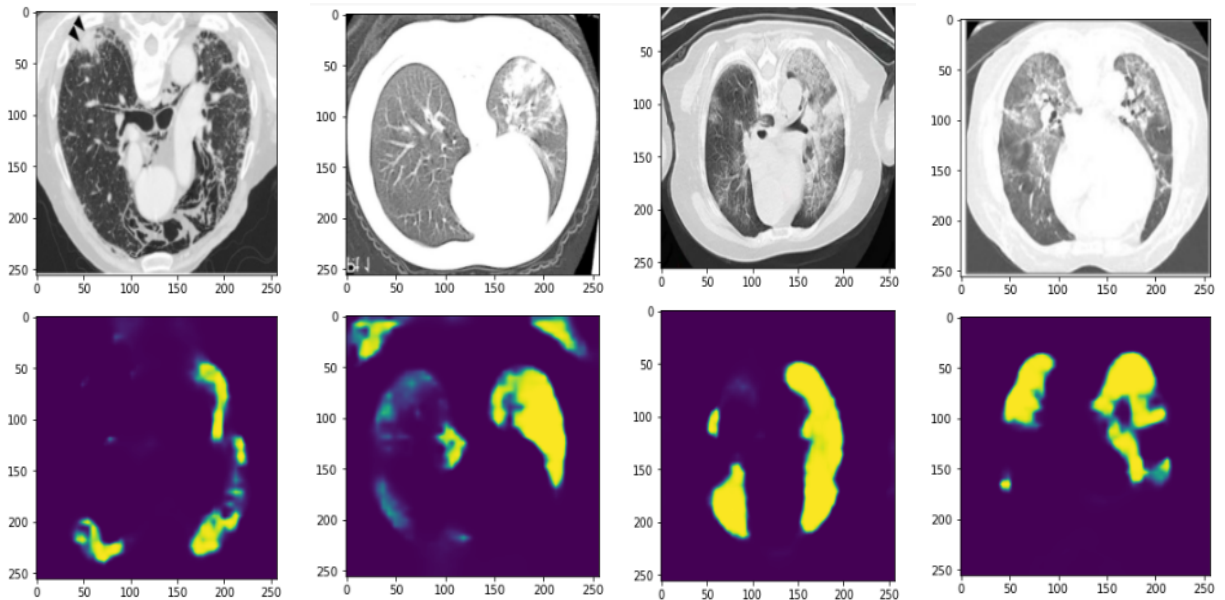
Στην ενότητα αυτή βλέπουμε πώς επιδρά σε εικόνες του αρχικού dataset η εφαρμογή segmentation χρησιμοποιώντας έτοιμα προεκπαιδευμένα μοντέλα σε covid εικόνες με τέτοια λειτουργία.

Εξετάζουμε επίσης τα αποτελέσματα που προκύπτουν από την χρήση αυτών για την εκπαίδευση του bestnet δικτύου και αν αξίζει να γίνει χρήση των συγκεκριμένων μοντέλων για εξαγωγή βελτιωμένων αποτελεσμάτων.

1ο Segmentation Δίκτυο

Πρόκειται για το δίκτυο Inf-Net [20] εκπαιδευμένο με μεθόδους επιβλεπόμενης μάθησης πάνω σε εικόνες στις οποίες έχει γίνει κατάτμηση των σημαντικών περιοχών από εξειδικευμένο επιστημονικό προσωπικό.

Παραδείγματα εικόνων του βασικού μας dataset μαζί με την επίδραση του Inf-Net φαίνονται στο σχήμα.



Σχήμα 5.11: Παραδείγματα Inf-Net

Τα αποτελέσματα που προκύπτουν αν εκπαιδεύσουμε εξαρχής το καλύτερό μας δίκτυο με τις κατατημένες εικόνες, είναι τα παρακάτω:

	precision	recall	f1-score	support
covid	0.61	0.63	0.62	98
non-covid	0.64	0.62	0.63	105
macro	0.63	0.63	0.63	203
weighted	0.63	0.63	0.63	203
accuracy	0.63			203

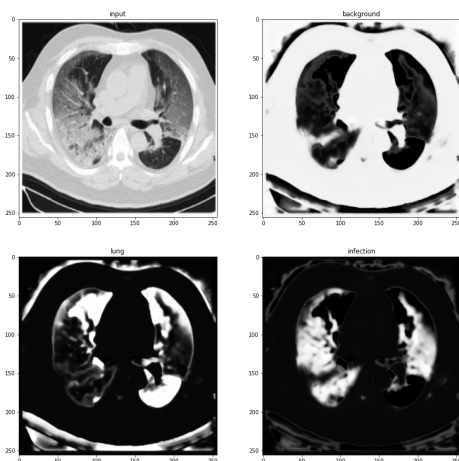
Πίνακας 5.25: Αποτελέσματα πρώτου segmentation

2ο Segmentation Δίκτυο

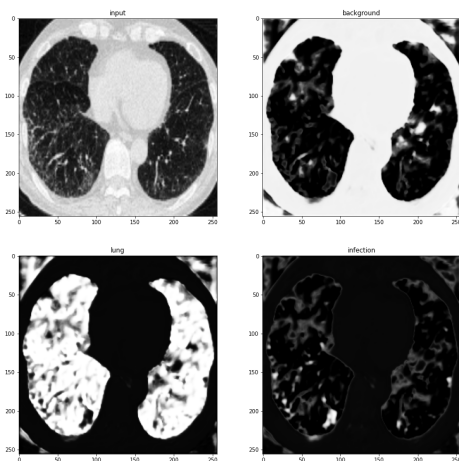
Πρόκειται για τον 3DSlicer [1] ο οποίος επιστρέφει μια εικόνα με 3 κανάλια, τα οποία αντιστοιχούν αντίστοιχα στο φόντο, στον πνεύμονα και την μόλυνση. Παραδείγματα της επίδρασης πάνω σε εικόνες της κάθε κλάσης για το σύνολο δεδομένων που δουλεύουμε φαίνονται στο σχήμα.

Πάνω αριστερά είναι η αρχική εικόνα. Δεξιά της είναι η πρόβλεψη του φόντου της

εικόνας. Στη δεύτερη σειρά είναι από αριστερά προς τα δεξιά, η ανίχνευση των πνευμόνων και η ανίχνευση των μολύνσεων.



Σχήμα 5.12: Covid δείγμα και η επίδραση του 3DSlicer



Σχήμα 5.13: Non-Covid δείγμα και η επίδραση του 3DSlicer

Εδώ το segmentation έχει καλύτερα αποτελέσματα από το Inf-Net:

	precision	recall	f1-score	support
covid	0.71	0.60	0.65	98
non-covid	0.68	0.77	0.72	105
macro	0.69	0.69	0.69	203
weighted	0.69	0.69	0.69	203
accuracy	0.69			203

Πίνακας 5.26: Αποτελέσματα 2ου segmentation

Ενδεικτικά, οι εικόνες ανάδειξης των περιοχών μόλυνσης του 3D-Slicer φαίνεται να είναι καλές καθώς αν τροφοδοτήσουμε τα 3 κανάλια των εικόνων εισόδου του δικού μας μοντέλου με αυτές, η ακρίβεια αυξάνεται στο 73 %.

	precision	recall	f1-score	support
covid	0.74	0.68	0.71	98
non-covid	0.73	0.78	0.75	105
macro	0.74	0.73	0.73	203
weighted	0.73	0.73	0.73	203
accuracy	0.73			203

Πίνακας 5.27: Αποτελέσματα 3D-Slicer με χρήση της ανίχνευσης μόλυνσης

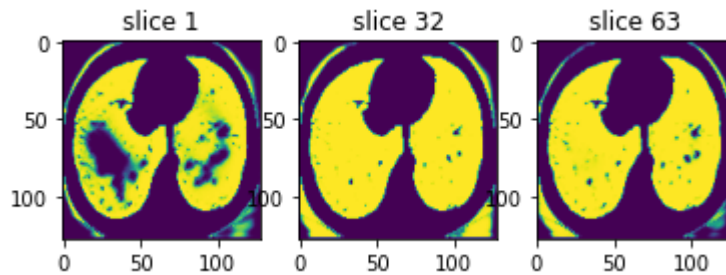
Αντίστοιχα, αν επιλέξουμε μόνο τις εικόνες ανάδειξης των περιοχών που αφορούν τον πνεύμονα παίρνουμε παρόμοια αποτελέσματα με πριν και συγκεκριμένα ακρίβεια της τάξης του 70%.

	precision	recall	f1-score	support
covid	0.64	0.91	0.75	98
non-covid	0.86	0.51	0.64	105
macro	0.75	0.71	0.70	203
weighted	0.75	0.70	0.69	203
accuracy	0.70			203

Πίνακας 5.28: Αποτελέσματα 3D-Slicer με ανάδειξη περιοχών πνεύμονα

3ο Segmentation Δίκτυο

Βασισμένο στη γνωστή δομή U-Net, το Lung-Segmentation-3D [25] εκτελεί κατάτμηση σε ακολουθία από scans μήκους 64. Προκειμένου να εξάγουμε αποτελέσματα για τις μεμονωμένες εικόνες, αναπαράγουμε 64 φορές την εικόνα εισόδου. Από την έξοδο επιλέγουμε ένα slice από τα κεντρικά. Τα αποτελέσματα μαζί με παράδειγμα εξόδου του δικτύου είναι παρακάτω:



Σχήμα 5.14: Τομές εξόδου Lung-Segmentation-3D

	precision	recall	f1-score	support
covid	0.64	0.91	0.75	98
non-covid	0.86	0.51	0.64	105
macro	0.75	0.71	0.70	203
weighted	0.75	0.70	0.69	203
accuracy	0.70			203

Πίνακας 5.29: Αποτελέσματα 3ου segmentation με επιλογή κεντρικής τομής

Ωστόσο εάν τροφοδοτήσουμε το δίκτυο με εικόνες 3 καναλιών στις οποίες επισυνάπτουμε τα slices 1, 32, 64, τότε έχουμε βελτίωση στην απόδοση του δικτύου που φτάνει ακρίβεια 74 %. Αυτό οφείλεται στο ότι δίνουμε παραπάνω πληροφορία στο δίκτυο, η οποία απορρέει από το εκπαιδευμένο σε series 3D δίκτυο κατάτμησης.

	precision	recall	f1-score	support
covid	0.72	0.74	0.73	98
non-covid	0.75	0.73	0.74	105
macro	0.74	0.74	0.74	203
weighted	0.74	0.74	0.74	203
accuracy	0.74			203

Πίνακας 5.30: Αποτελέσματα 3ου segmentation με επιλογή slice τομών 1, 32, 64

5.4 Επαύξηση στα ελληνικά δεδομένα

Σκοπός αυτού του συνόλου πειραμάτων είναι η δημιουργία πολλών ασθενών Covid που θα έχουν ως βάση τους 6 Έλληνες Covid ασθενείς που διαθέτουμε. Παρακάτω αναπτύσσονται 4 διαφορετικές μέθοδοι για την δημιουργία ενός τέτοιου dataset.

Προκειμένου να εκπαιδύσουμε ένα RNN με σειρές από τομές ασθενών και μη, είναι απαραίτητο οι αναπαραστάσεις που προκύπτουν από κάθε άτομο και το δείγμα εισόδου του, να είναι του ίδιου μεγέθους. Για αυτό το λόγο, στο σύνολο δεδομένων που δημιουργείται, δημιουργούμε δείγματα με ίδιο αριθμό από τομές, και συγκεκριμένα 55 σε πλήθος.

5.4.1 Δομή Συνόλων Δεδομένων

Όσον αφορά τα non-covid δείγματα, επιλέγουμε 100 3D τομογραφίες από το νοσοκομείο 1 και 100 3D τομογραφίες από το νοσοκομείο 2. Ο αριθμός των slices για κάθε έναν από τα υγιή δείγματα. Συγκεκριμένα έχουμε τα εξής στατιστικά για το μήκος των ακολουθιών:

μέση τιμή	222.615
διάμεσος	132.5
ελάχιστη τιμή	49
μέγιστη τιμή	589

Πίνακας 5.31: Στατιστικά μήκους ακολουθίας των 200 Non-Covid ατόμων

Οι παραπάνω τιμές μας οδηγούν στην απόφαση να επιλέξουμε ως κοινό μήκος ακολουθίας το 55 που είναι κοντά στην ελάχιστη τιμή. Δημιουργούμε έτσι για κάθε

ακολουθία μια αναπαράσταση 55 εικόνων, είτε υποδειγματοληπτώντας αν το μήκος της ακολουθίας είναι αρχικά μεγαλύτερο, ή επαναλαμβάνοντας τις τελευταίες εικόνες μέχρι να γίνουν 55 αν το μήκος της ακολουθίας είναι αρχικά μικρότερο. Χωρίζουμε έπειτα αυτό το σύνολο των 200 σειρών σε 160 σειρές για εκπαίδευση και 40 σειρές για testing.

Όσον αφορά τις covid εικόνες, επιλέγουμε τις 4 ακολουθίες covid με τις λιγότερες εικόνες και αναπαράγουμε σειρές ώστε να προκύψουν πολλαπλά series μεγέθους 55. Παρακάτω περιγράφονται 4 μέθοδοι με τις οποίες πετυχαίνουμε αυτόν τον σκοπό:

1. **Επανάληψη τελευταίων scans:** Για τις 4 ακολουθίες covid, επαναλαμβάνουμε τις τελευταίες εικόνες των scans τους, μέχρι να συμπληρωθεί το όριο των 55 που έχουμε βάλει. Προκύπτουν έτσι 4 συμπληρωμένες ακολουθίες οι οποίες για να συνεκπαιδευτούν με τις 160 non-covid ακολουθίες χρειάζεται είτε να ισορροπήσουμε το dataset, είτε να χρησιμοποιηθεί κατάλληλο cross entropy loss που να λαμβάνει υπόψη την δεδομένη ισορροπία των κλάσεων.
2. **Προσθήκη εικόνων από ελληνικές ακολουθίες:** Συμπληρώνουμε τις υπόλοιπες τομογραφίες από τις ελληνικές ακολουθίες covid ασθενών που απομένουν. Διαλέγουμε κάθε φορά διαφορετικά κομμάτια αυτών, ώστε να προκύψει ποικιλία νέων δειγμάτων. Δημιουργούνται έτσι 196 δείγματα.
3. **Προσθήκη εικόνων από το αρχικό Dataset:** Τα scans αυτά προστίθενται ενδιάμεσα, στο κεντρικό κομμάτι των ακολουθιών μια και οι εικόνες αυτές αφορούν εγκάρσιες τομές σε κεντρικό σημείο των πνευμόνων. Χρησιμοποιώντας διαφορετικά υποσύνολα των 349 αυτών εικόνων προκύπτουν 210 νέα δείγματα.
4. **Προσθήκη εικόνων από τα ξένα covid series:** Ομοίως με πριν δημιουργούμε 210 νέα δείγματα για εκπαίδευση, αλλά χρησιμοποιώντας εικόνες από τις ακολουθίες ξένων τομογραφιών covid ασθενών.

Για τα test δείγματα της κλάσης covid παίρνουμε τους 2 ασθενείς covid που απομένουν. Για να δημιουργήσουμε αναπαραστάσεις μήκους 55, υποδειγματοληπτούμε λαμβάνοντας πολλαπλές ακολουθίες για κάθε αξονική ενός ατόμου. Για τον έναν προκύπτουν 5 ακολουθίες ενώ για τον άλλο 51, όλες μήκους 55. Σε αυτές μπορούμε να εφαρμόσουμε τις μεθόδους "majority voting" ή "at least one" προκειμένου να κάνουμε πρόβλεψη για κάθε έναν από αυτούς όπως έχει επεξηγηθεί σε προηγούμενα πειράματα.

5.4.2 Αρχιτεκτονική μοντέλων πρόβλεψης

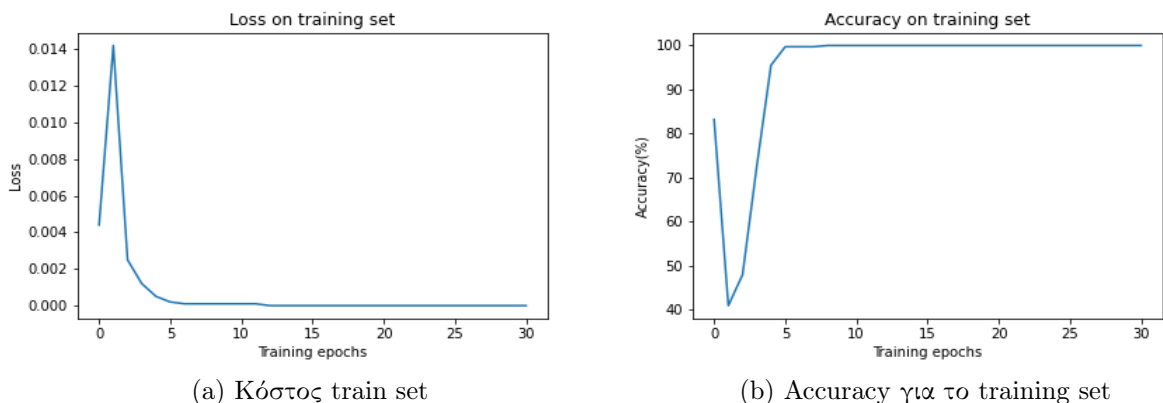
Η δομή που χρησιμοποιείται είναι το καλύτερο pretrained μοντέλο που προέκυψε από τους αρχικούς μετασχηματισμούς (transformations) αλλά αφαιρώντας τα τελευταία layers. Στη θέση τους μπαίνουν ένα gru layer με 128 units το οποίο τροφοδοτεί

ένα fully-connected layer μεγέθους 32. Ακολουθεί το τελικό layer της αρχιτεκτονικής.

Για την εκπαίδευση αυτού του δικτύου χρησιμοποιείται ο αλγόριθμος βελτιστοποίησης Adam με ρυθμό εκμάθησης ίσο με 0.001. Για την επανεκπαίδευση της δομής αυτής μπορούμε να παγώσουμε κάποια layer, και συγκεκριμένα τα layer που αφορούν το CNN κομμάτι της δομής. Στη συνέχεια γίνεται εκπαίδευση end-to-end, δηλαδή στο σύνολο της αρχιτεκτονικής του δικτύου.

5.4.3 Αποτελέσματα σε σειρές από CTs

Κατά την εκπαίδευση του δικτύου, η ακρίβεια (accuracy) φτάνει γρήγορα το 100 % ενώ και το κόστος εκπαίδευσης (training loss) πρακτικά μηδενίζεται μετά από λίγες μόνο εποχές, και συγκεκριμένα λιγότερες από 10.



Σχήμα 5.15: Εκπαίδευση του CNN-RNN με τα training series για τη 2η μέθοδο dataset

Ακολουθούνται δύο μέθοδοι για την εκπαίδευση των δικτύων. Η μία είναι η εκπαίδευση του δικτύου παγώνοντας τα επίπεδα του προεκπαιδευμένου CNN (frozen) και αφήνοντας τα τελευταία επίπεδα ελεύθερα στην ανανέωση των βαρών τους κατά τη διάρκεια της εκπαίδευσης. Τα τελευταία αυτά επίπεδα περιλαμβάνουν το GRU layer των 128 units, το fully-connected layer που δίνει αναπαραστάσεις μεγέθους 32 για κάθε δεδομένο εισόδου που τροφοδοτείται στο δίκτυο, όπως επίσης και το Relu layer που προηγείται αυτού, καθώς και το επίπεδο εξόδου (output layer). Η άλλη μέθοδος είναι να εκπαιδεύσουμε όλο το δίκτυο (end-to-end), αναμένοντας καλύτερα αποτελέσματα λόγω της πλήρους προσαρμογής στο δίκτυο.

Παγωμένο CNN

Επιλέγουμε batch size ίσο με 8 τομογραφίες μεγέθους 55. Στις 3 τελευταίες περιπτώσεις τα sets είναι σχετικά ισορροπημένα. Στην πρώτη περίπτωση ωστόσο έχουμε μόνο 4 series. Η εκπαίδευση σε αυτή την περίπτωση γίνεται με 2 τρόπους:

- weighted binary Cross-Entropy loss: Τα βάρη ενημερώνονται στη συνάρτηση κόστους λαμβάνοντας υπόψη την ανισοροπία του δικτύου.
- ισοροπία μέσω επανάληψης: Ισορροπούμε κάθε batch εκπαίδευσης βάζοντας επαναλήψεις των 4 ασθενών ώστε να έχουμε ίδιο αριθμό δειγμάτων κάθε κλάσης. Έτσι στην περίπτωση μας που έχουμε 8 δείγματα ανά batch, μπορούμε να βάζουμε τους 4 ασθενείς του training set συν άλλους 4 από τα non-covid δείγματα.

Για τις περιπτώσεις 1, 2, 3 οι προβλέψεις είναι όλες non-covid για το σύνολο του test set. Αυτό σημαίνει 100% για την κλάση Non-Covid αλλά και λάθος πρόβλεψη για τους 2 ασθενείς covid του test set. Η πρόβλεψη είναι λάθος είτε αυτή γίνει με τη μέθοδο "at least one", είτε αυτή γίνει με "majority voting", αφού και στις 2 περιπτώσεις δεν έχουμε κάποια πρόβλεψη Covid σε κάποιο από τα series που εκπροσωπούν τον κάθε ασθενή. Πιο αναλυτικά τα αποτελέσματα φαίνονται παρακάτω:

Προβλέψεις στα υγιή δείγματα

Non-Covid	total	True	False	accuracy(%)
series	40	40	0	100

Προβλέψεις στους 2 Covid ασθενείς

Covid	total	True	False	majority	at-least-one
patient-1	5	0	5	Non-Covid	Non-Covid
patient-2	51	0	51	Non-Covid	Non-Covid
accuracy				0	0

Πίνακας 5.32: Αποτελέσματα για τα dataset 1,2,3 στο test set

Ωστόσο στην περίπτωση 4, αν εκπαιδύσουμε το δίκτυο για 10 εποχές παρατηρείται σωστή πρόβλεψη για τον έναν ασθενή Covid, καθώς 4 από τις 5 προβλέψεις για τα series του είναι Covid. Αυτό δίνει το επιθυμητό αποτέλεσμα για τον έναν ασθενή είτε με τη μέθοδο "majority voting" είτε με την "at least one", όπως φαίνεται παρακάτω.

Προβλέψεις στα υγιή δείγματα				
Non-Covid	total	True	False	accuracy(%)
series	40	40	0	100

Προβλέψεις στους 2 Covid ασθενείς					
Covid	total	True	False	majority	at-least-one
patient-1	5	4	1	Covid	Covid
patient-2	51	0	51	Non-Covid	Non-Covid
accuracy				50	50

Πίνακας 5.33: Αποτελέσματα για το dataset 4 στο test set

End-to-end εκπαίδευση

Λόγω της αύξησης των υπερπαραμέτρων μικραίνουμε το batch size σε μέγεθος 6 ασθενών. Τα αποτελέσματα δεν διαφοροποιούνται στην end-to-end μέθοδο καθώς οι προβλέψεις είναι όλες Non-Covid για τις 4 μεθόδους επαύξησης των δεδομένων. Άρα σύμφωνα με τα παραπάνω έχουμε πάλι 100 % accuracy στα 40 υγιή δείγματα, ωστόσο οι 2 ασθενείς Covid προβλέπονται ως υγιείς, γεγονός που δεν μας επιτρέπει να έχουμε σωστή διάγνωση.

Φυσικά στις παραπάνω μεθόδους τα 55 scans που αφορούν την τομογραφία ενός ατόμου δεν μπαίνουν ανακατεμένα στο δίκτυο αλλά με τη σειρά που έχουν δημιουργηθεί στην πραγματικότητα ώστε να αποτελούν μια 3D αξονική τομογραφία. Οι επιπλέον εικόνες προστίθενται με τον τρόπο που περιγράφηκε στην περιγραφή των συνόλων.

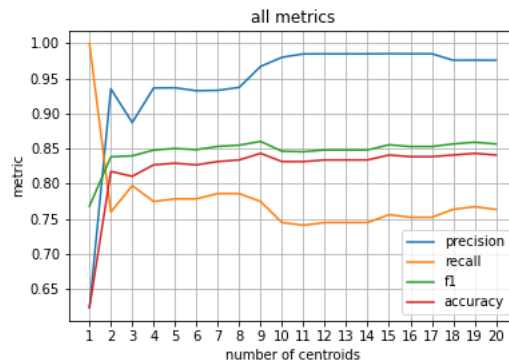
Σε αντίθεση με τα scans ανά ένα άτομο, τα δείγματα μπορούν να ανακατευτούν μεταξύ τους κατά την είσοδό τους στο δίκτυο. Στα παραπάνω πειράματα δοκιμάστηκαν ισορροπημένα batches τα οποία είχαν δείγματα και από τις 2 κλάσεις. Δοκιμάστηκε επίσης τα δείγματα να τροφοδοτούν το δίκτυο με τη σειρά, πρώτα δηλαδή τα δείγματα μίας κλάσης και στη συνέχεια της άλλης. Και οι δύο αυτές προσεγγίσεις έδωσαν ίδια αποτελέσματα.

Clustering

Επιλέγουμε αρχικά το δίκτυο το οποίο εκπαιδεύτηκε από τα δεδομένα εμπλουτισμένα με τα 20 ξένα covid series, δηλαδή η περίπτωση 4 της επαύξησης των δεδομένων. Από αυτό συλλέγουμε τις αναπαραστάσεις των δεδομένων όπως προκύπτουν από την έξοδο του 32-fully connected layer στα τελικά επίπεδα της αρχιτεκτονικής. Εφαρμόζοντας συσταδοποίηση στο πλήθος των δεδομένων, παρατηρούμε ότι συνεχίζει να κάνει τις ίδιες προβλέψεις για τους 2 ασθενείς Covid, δηλαδή να βρίσκει μόνο τον

έναν από τους δύο σωστά.

Συνενώνουμε έτσι τα training και test data και βλέπουμε για ποια k έχουμε την καλύτερη συμπεριφορά.



Σχήμα 5.16: Clustering για το CNN-RNN δίκτυο

Παρατηρούμε ότι η καλύτερη συμπεριφορά των μετρικών παρατηρείται για $k=9$. Φυσικά δεν είναι δυνατή η αποτελεσματική οπτικοποίηση των κέντρων, διότι τα δείγματα αυτή τη φορά είναι 3D αξονικές. Μπορούμε ωστόσο να παρατηρήσουμε τη συμπεριφορά των αξονικών των 2 ασθενών διατηρώντας την επιλογή $k=9$.

Για τον ασθενή με τα 5 subseries, η επιλογή $k=9$ έχει ως αποτέλεσμα να δείχνει όλα τα subseries ως Covid και να τον προβλέπει σωστά ως ασθενή.

Ο δεύτερος ασθενής με τα 51 subseries, έχει ακόμα ίδιες προβλέψεις με πριν, δηλαδή Non-Covid για όλα τα subseries. Για την περαιτέρω μελέτη του συγκεκριμένου ασθενή, βρίσκουμε αρχικά τις ακτίνες των κέντρων για $k=9$. Ως ακτίνα ορίζουμε τη μέγιστη απόσταση ενός δείγματος το οποίο ανήκει στο εν λόγω κέντρο από τα training δείγματα. Για τις αποστάσεις, συνεχίζουμε να τις υπολογίζουμε με χρήση της Ευκλείδειας απόστασης. Στη συνέχεια εξετάζουμε για όλα τα slices του ασθενή, αν βρίσκονται εντός της ακτίνας του κέντρου στο οποίο ανήκουν. Βρίσκουμε έτσι τα κέντρα στα οποία ανήκουν, τα οποία είναι όλα κέντρα non-Covid, υπολογίζουμε την απόστασή τους από αυτά και βλέπουμε αν αυτή η απόσταση ξεπερνάει την ακτίνα του cluster.

slice	slice distance from cluster	cluster radius	cluster number	is slice inside radius	prediction
27	2.292895	3.136405	0	True	Non-Covid
28	3.298008	3.136405	0	False	Non-Covid
29	3.390570	3.136405	0	False	Non-Covid
30	1.252148	3.826905	7	True	Non-Covid
31	2.195522	3.136405	0	True	Non-Covid
32	2.382035	3.136405	0	True	Non-Covid
33	2.602234	3.136405	0	True	Non-Covid
34	2.542834	3.136405	0	True	Non-Covid
35	3.241741	3.136405	0	False	Non-Covid
36	2.025687	3.826905	7	True	Non-Covid

Σχήμα 5.17: Slices εκτός ακτίνας των κέντρων στα οποία ανήκουν

Η διαδικασία αυτή όπως φαίνεται στην παραπάνω εικόνα, μας δίνει όντως τομές οι οποίες είναι εκτός του εύρους των κέντρων εκπαίδευσης. Αυτό μας δείχνει ότι το δείγμα του ασθενούς είναι πολύ διαφορετικό από τα δείγματα εκπαίδευσης και ότι η πρόβλεψη είναι δικαιολογημένα διαφορετική από το επιθυμητό αποτέλεσμα.

Μείωση μεγέθους συνόλου εκπαίδευσης

Η επαύξηση που υπέστησαν τα δεδομένα Covid μας βοήθησε να εκπαιδεύσουμε αποτελεσματικά το μοντέλο πρόβλεψης, όπως επίσης και να συμπεριλάβουμε πλήθος υγιών δειγμάτων που είχαμε στη διάθεσή μας κρατώντας το σύνολο δεδομένων ισορροπημένο. Ωστόσο, κατούσιαν τα δείγματα εκπαίδευσης Covid παρέμειναν 4, αριθμός αρκετά περιορισμένος. Σε αυτό το γεγονός οφείλεται και το ότι η πλειονότητα των παραπάνω μοντέλων είχε προβλέψεις Non-Covid.

Για τον λόγο αυτό δοκιμάστηκε να μειωθεί το πλήθος των δεδομένων εκπαίδευσης, ώστε να αξιοποιηθεί πιο αποτελεσματικά η πληροφορία που διαθέτουμε. Επιλέγουμε έτσι 50 δείγματα από τα 200 υγιή της βάσης δεδομένων. Δημιουργούμε ακόμα με τη μέθοδο επαύξησης δεδομένων 2, άλλα 50 δείγματα για την κλάση Covid.

Παρά την μείωση του όγκου, παρατηρήθηκε ίδια συμπεριφορά σε σύγκριση με τα δίκτυα στα οποία είχαμε 200 δείγματα ανά κλάση.

Συμπερασματικά μπορούμε να πούμε ότι ο λόγος που δεν παρατηρούνται καλά αποτελέσματα σε αυτές τις περιπτώσεις είναι η χρήση δεδομένων από διαφορετικά νοσοκομεία της χώρας. Η ανομοιογένεια των εικόνων που απαρτίζουν τα series, καθώς και

η σημαντική διαφορά στα μήκη των ακολουθιών είναι σημαντικοί παράγοντες που επηρεάζουν αρνητικά την έκβαση των αποτελεσμάτων. Πιο συγκεκριμένα:

- Τα 4 Covid series που χρησιμοποιήσαμε για training είναι από το νοσοκομείο 1, ενώ τα 2 που χρησιμοποιήθηκαν για testing είναι από το νοσοκομείο 2 και είχαν αρκετά μεγαλύτερο αριθμό από scans.
- Τα δείγματα Non-Covid είναι όπως έχει προαναφερθεί από τα νοσοκομεία 1 και 2 και έχουν κι αυτά μεγάλες διαφορές στο μήκος των ακολουθιών.

Κεφάλαιο 6

Επίλογος

6.1 Αποτελέσματα και Συμπεράσματα

Στην παρούσα εργασία εφαρμόσαμε τεχνολογίες βαθιάς μάθησης από την αιχμή της τεχνολογίας που έχουν παραχθεί τόσο στον διεθνή χώρο όσο και από μέλη του Εργαστηρίου Τεχνητής Νοημοσύνης και Συστημάτων Μάθησης του Ε.Μ. Πολυτεχνείου σε ένα μικρό σύνολο από Ελληνικά Δεδομένα που είχαμε διαθέσιμα. Οι τεχνολογίες αυτές θα χρησιμοποιηθούν στη συνέχεια για την ανάλυση πολύ μεγαλύτερων όγκων δεδομένων από Ελληνικά Δημόσια Νοσοκομεία, στο πλαίσιο και της συνεργασίας του Εθνικού Δικτύου Υποδομών Τεχνολογίας Έρευνας και του Υπουργείου Υγείας.

Συμπερασματικά, υπήρξε λεπτομερής ανάλυση των δεδομένων των συνόλων που παρουσιάστηκαν με σκοπό τη δημιουργία μοντέλων ανίχνευσης των ασθενών με κορονοϊό.

Αρχικά, εντοπίστηκαν οι κατάλληλοι μετασχηματισμοί ώστε να βελτιστοποιηθούν αποτελέσματα από προηγούμενες έρευνες πάνω στο βασικό σύνολο δεδομένων. Κρατώντας το καλύτερο μοντέλο που προέκυψε από τα πειράματα αυτά, προχωρήσαμε σε περαιτέρω δημιουργία προεκπαιδευμένων μοντέλων.

Ως συνέχεια αυτού, χρησιμοποιήθηκαν τεχνικές μη επιβλεπόμενης μάθησης για την ομαδοποίηση των δεδομένων και την εύρεση του πλήθους συστάδων με τα καλύτερα αποτελέσματα.

Στη συνέχεια, εξετάστηκε η μέθοδος κατάτμησης πάνω στο βασικό σύνολο δεδομένων. Βρέθηκαν έτσι ποια έτοιμα μοντέλα κατάτμησης λειτουργούν καλύτερα

οπτικά πάνω στα δεδομένα και ποια είχαν καλύτερη επίδοση αν τροφοδοτηθούν στα μοντέλα εντοπισμού ασθενών Covid-19.

Τέλος, έγινε ενδεδειγμένη έρευνα στα ελληνικά δεδομένα.

Από τη μία χρησιμοποιήθηκαν μέθοδοι εξαγωγής πρόβλεψης από τις μεμονωμένες προβλέψεις πάνω στο σύνολο των εικόνων ενός πιθανού ασθενή, μέσω της χρήσης των προηγούμενων νευρωνικών δικτύων.

Από την άλλη, αναπτύχθηκαν αναδρομικά νευρωνικά δίκτυα για την πρόβλεψη Covid-19 με είσοδο ακολουθίας εικόνων από το σύνολο των τομών της αξονικής τομογραφίας.

Συνοψίζοντας, τα προεκπαιδευμένα μοντέλα που δημιουργήθηκαν από το βασικό dataset οδήγησαν σε αποτελεσματικές μεθόδους πρόβλεψης των ασθενών, ενώ τελικά αναπτύχθηκαν CNN-RNN για την πρόβλεψη Covid-19 με είσοδο 3D αξονική τομογραφία.

6.2 Μελλοντικές Επεκτάσεις

Η ασθένεια Covid-19 έχει ισχυρή επίδραση στην παγκόσμια κοινότητα και η παρουσία της προμηνύεται να μείνει ενεργή για αρκετό καιρό ακόμα. Παρά την ενδεχόμενη εύρεση εμβολίων για την θωράκιση του ανοσοποιητικού συστήματος των πληθυσμών, η αμφισβητούμενη επίδοσή τους ενθαρρύνει περαιτέρω τις έρευνες πάνω στην ασθένεια.

Μελλοντική εργασία επομένως μπορεί να διεξαχθεί πάνω στα δεδομένα που αναφέρονται στην διπλωματική εργασία, όπως επίσης και διερεύνηση εναλλακτικών μεθόδων για τη δημιουργία αξιόπιστων μοντέλων.

Τα νέα δεδομένα μπορούν να εμπλουτίζονται συνεχώς με νέα δείγματα, χρήσιμα για την καλύτερη εκπαίδευση των νευρωνικών δικτύων που αναφέρονται στην εργασία. Επίσης, άλλα προεκπαιδευμένα μοντέλα ενδεχομένως να μπορούν να δοκιμαστούν ως βάση για την εκπαίδευση με τα δεδομένα του προβλήματος.

Φυσικά, περαιτέρω εργασία μπορεί να γίνει για την μέθοδο της ομαδοποίησης, όπως δοκιμή άλλων τύπου αποστάσεων κατά την εφαρμογή του K-Means.

Η κατάτμηση των εικόνων του συγκεκριμένου dataset παραμένει μία πρόκληση, όπως επίσης και η κατάτμηση των εικόνων του ελληνικού συνόλου δεδομένων.

Όσον αφορά τα μοντέλα πρόβλεψης σε series πιθανών ασθενών, αφήνεται το περιθώριο δοκιμών εναλλακτικών μηχανισμών αναδρομικών δικτύων με σκοπό τη διάγνωση 3D αξονικής τομογραφίας.

Βιβλιογραφία

- [1] Chen. J. & Frey E. C. “Medical Image Segmentation via Unsupervised Convolutional Neural Network. In Medical Imaging with Deep Learning”. In: (2020 January).
- [2] Yannis Avrithis, Nicolas Tsapatsoulis, and Stefanos Kollias. “Broadcast news parsing using visual cues: A robust face detection approach”. In: *2000 IEEE International Conference on Multimedia and Expo. ICME2000. Proceedings. Latest Advances in the Fast Changing World of Multimedia (Cat. No. 00TH8532)*. Vol. 3. IEEE. 2000, pp. 1469–1472.
- [3] Konstantinos Rapantzikos et al. “Bottom-up spatiotemporal visual attention model for video analysis”. In: *IET Image Processing 1.2* (2007), pp. 237–248.
- [4] Ilianna Kollia et al. “Semantic image analysis using a symbolic neural architecture”. In: *Image Analysis & Stereology 29.3* (2010), pp. 159–172.
- [5] Birte Glimm et al. “Lower and Upper Bounds for SPARQL Queries over OWL Ontologies.” In: *AAAI*. Citeseer. 2015, pp. 109–115.
- [6] Kaiming He et al. *Deep Residual Learning for Image Recognition*. 2015. arXiv: [1512.03385 \[cs.CV\]](#).
- [7] Dimitris Kollias et al. “Interweaving deep learning and semantic techniques for emotion analysis in human-machine interaction”. In: *2015 10th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP)*. IEEE. 2015, pp. 1–6.
- [8] Gao Huang et al. *Densely Connected Convolutional Networks*. 2016. arXiv: [1608.06993 \[cs.CV\]](#).
- [9] Dimitrios Kollias, Athanasios Tagaris, and Andreas Stafylopatis. “On line emotion detection using retrainable deep neural networks”. In: *2016 IEEE Symposium Series on Computational Intelligence (SSCI)*. IEEE. 2016, pp. 1–8.

- [10] Data-Blogger. “Python Deep Learning tutorial”. In: (2017). [Online; accessed September 14, 2020]. URL: <https://www.data-blogger.com/2017/08/27/gru-implementation-tensorflow/>.
- [11] Athanasios Tagaris, Dimitrios Kollias, and Andreas Stafylopatis. “Assessment of Parkinson’s disease based on deep neural networks”. In: *International Conference on Engineering Applications of Neural Networks*. Springer. 2017, pp. 391–403.
- [12] Lichy Han and Maulik Kamdar. “MRI to MGMT: predicting methylation status in glioblastoma patients using convolutional recurrent neural networks”. In: *Pacific Symposium on Biocomputing. Pacific Symposium on Biocomputing 23* (Jan. 2018), pp. 331–342.
- [13] Dimitrios Kollias and Stefanos Zafeiriou. “Training deep neural networks with different datasets in-the-wild: The emotion recognition paradigm”. In: *2018 International Joint Conference on Neural Networks (IJCNN)*. IEEE. 2018, pp. 1–8.
- [14] Dimitrios Kollias et al. “Deep neural architectures for prediction in healthcare”. In: *Complex & Intelligent Systems 4.2* (2018), pp. 119–131.
- [15] Dimitrios Kollias et al. “Photorealistic facial synthesis in the dimensional affect space”. In: *Proceedings of the European Conference on Computer Vision (ECCV)*. 2018, pp. 0–0.
- [16] Athanasios Tagaris et al. “Machine learning for neurodegenerative disorder diagnosis—survey of practices and launch of benchmark dataset”. In: *International Journal on Artificial Intelligence Tools 27.03* (2018), p. 1850011.
- [17] Hiromu Yakura et al. “Malware Analysis of Imaged Binary Samples by Convolutional Neural Network with Attention Mechanism”. In: Mar. 2018, pp. 127–134. DOI: [10.1145/3176258.3176335](https://doi.org/10.1145/3176258.3176335).
- [18] Ilianna Kollia, Andreas-Georgios Stafylopatis, and Stefanos Kollias. “Predicting Parkinson’s disease using latent information extracted from deep neural networks”. In: *2019 International Joint Conference on Neural Networks (IJCNN)*. IEEE. 2019, pp. 1–8.
- [19] Mingxing Tan and Quoc V. Le. *EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks*. 2019. arXiv: [1905.11946](https://arxiv.org/abs/1905.11946) [cs.LG].
- [20] Deng-Ping Fan et al. “Inf-Net: Automatic COVID-19 Lung Infection Segmentation from CT Images”. In: *IEEE TMI* (2020).

- [21] Xuehai He et al. “Sample-Efficient Deep Learning for COVID-19 Diagnosis Based on CT Scans”. In: *medrxiv* (2020).
- [22] Ma Jun et al. *COVID-19 CT Lung and Infection Segmentation Dataset*. Version Verson 1.0. Zenodo, Apr. 2020. DOI: [10.5281/zenodo.3757476](https://doi.org/10.5281/zenodo.3757476). URL: <https://doi.org/10.5281/zenodo.3757476>.
- [23] Dimitrios Kollias and Stefanos P Zafeiriou. “Exploiting multi-cnn features in cnn-rnn based dimensional emotion recognition on the omg in-the-wild dataset”. In: *IEEE Transactions on Affective Computing* (2020).
- [24] D Kollias et al. “Deep Transparent Prediction through Latent Representation Analysis”. In: *arXiv preprint arXiv:2009.07044* (2020).
- [25] “lung-segmentation-3d”. In: (2020). [Online; accessed September 14, 2020]. URL: <https://github.com/imlab-uip/lung-segmentation-3d>.
- [26] Medium. “Deep Learning Recurrent neural networks”. In: (2020). [Online; accessed September 3, 2020]. URL: <https://medium.com/deeplearningbrasil/deep-learning-recurrent-neural-networks-f9482a24d010>.
- [27] Medium. “Fully-Connected Layer with dynamic input shape”. In: (2020). [Online; accessed September 3, 2020]. URL: <https://medium.com/@tecokids.monastir/fully-connected-layer-with-dynamic-input-shape-70c869ae71af>.
- [28] Medium. “Pooling Layer”. In: (2020). [Online; accessed September 3, 2020]. URL: <https://medium.com/ai-in-plain-english/pooling-layer-beginner-to-intermediate-fa0dbdce80eb>.
- [29] MedSeg. *COVID-19 CT segmentation dataset*. [Online; accessed September 10, 2020]. 2020. URL: <http://medicalsegmentation.com/covid19/>.
- [30] Stanford.edu. “K Means”. In: (2020). [Online; accessed September 3, 2020]. URL: <https://stanford.edu/~cpiech/cs221/handouts/kmeans.html>.
- [31] Towards Data Science. “A comprehensive guide to convolutional neural networks”. In: (2020). [Online; accessed September 2, 2020]. URL: <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>.
- [32] Towards Data Science. “Introduction to recurrent neural network”. In: (2020). [Online; accessed September 3, 2020]. URL: <https://towardsdatascience.com/introduction-to-recurrent-neural-network-27202c3945f3>.

- [33] James Wingate et al. “Unified deep learning approach for prediction of Parkinson’s disease”. In: *IET Image Processing* 14.10 (2020), pp. 1980–1989.
- [34] Jinyu Zhao et al. “COVID-CT-Dataset: a CT scan dataset about COVID-19”. In: *arXiv preprint arXiv:2003.13865* (2020).
- [35] Dimitrios Kollias et al. “Adaptation and contextualization of deep neural network models”. In: *2017 IEEE symposium series on computational intelligence (SSCI)*. IEEE, pp. 1–8.