



ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΜΕ ΕΦΑΡΜΟΓΕΣ
ΣΤΗ ΒΙΟΙΑΤΡΙΚΗ

**Μέθοδοι μηχανικής μάθησης για την υποβοήθηση της ιατρικής
διάγνωσης του SARS-CoV-2 με δεδομένα αξονικής τομογραφίας
θώρακα**

Βικτώρια Φίλη

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ
Υπεύθυνος
Μιχαήλ Σαβελώνας
Επίκουρος Καθηγητής

Λαμία, 2022



**ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΙΑΣ
ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΜΕ ΕΦΑΡΜΟΓΕΣ ΣΤΗ
ΒΙΟΙΑΤΡΙΚΗ**

**Μέθοδοι μηχανικής μάθησης για την υποβοήθηση της ιατρικής
διάγνωσης του SARS-CoV-2 με δεδομένα αξονικής τομογραφίας
θώρακα**

Βικτώρια Φίλη

**ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ
Επιβλέπων
Μιχαήλ Σαβελώνας
Επίκουρος Καθηγητής**

Λαμία, 2022

Με ατομική μου ευθύνη και γνωρίζοντας τις κυρώσεις ⁽¹⁾, που προβλέπονται από της διατάξεις της παρ. 6 του άρθρου 22 του Ν. 1599/1986, δηλώνω ότι:

1. Δεν παραθέτω κομμάτια βιβλίων ή άρθρων ή εργασιών άλλων αυτολεξεί **χωρίς να τα περικλείω σε εισαγωγικά** και χωρίς να αναφέρω το συγγραφέα, τη χρονολογία, τη σελίδα. Η αυτολεξεί παράθεση χωρίς εισαγωγικά χωρίς αναφορά στην πηγή, είναι λογοκλοπή. Πέραν της αυτολεξεί παράθεσης, λογοκλοπή θεωρείται και η παράφραση εδαφίων από έργα άλλων, συμπεριλαμβανομένων και έργων συμφοιτητών μου, καθώς και η παράθεση στοιχείων που άλλοι συνέλεξαν ή επεξεργάστηκαν, χωρίς αναφορά στην πηγή. Αναφέρω πάντοτε με πληρότητα την πηγή κάτω από τον πίνακα ή σχέδιο, όπως στα παραθέματα.
2. Δέχομαι ότι η αυτολεξεί **παράθεση χωρίς εισαγωγικά**, ακόμα κι αν συνοδεύεται από αναφορά στην πηγή σε κάποιο άλλο σημείο του κειμένου ή στο τέλος του, είναι αντιγραφή. Η αναφορά στην πηγή στο τέλος π.χ. μιας παραγράφου ή μιας σελίδας, δεν δικαιολογεί συρραφή εδαφίων έργου άλλου συγγραφέα, έστω και παραφρασμένων, και παρουσιάσή τους ως δική μου εργασία.
3. Δέχομαι ότι υπάρχει επίσης περιορισμός στο μέγεθος και στη συχνότητα των παραθεμάτων που μπορώ να εντάξω στην εργασία μου εντός εισαγωγικών. Κάθε μεγάλο παράθεμα (π.χ. σε πίνακα ή πλαίσιο, κλπ), προϋποθέτει ειδικές ρυθμίσεις, και όταν δημοσιεύεται προϋποθέτει την άδεια του συγγραφέα ή του εκδότη. Το ίδιο και οι πίνακες και τα σχέδια.
4. Δέχομαι όλες τις συνέπειες σε περίπτωση λογοκλοπής ή αντιγραφής.

Ημερομηνία: 11/10/2022

Η Δηλούσα

Βικτώρια Φίλη

(1) «Όποιος εν γνώσει του δηλώνει ψευδή γεγονότα ή αρνείται ή αποκρύπτει τα αληθινά με έγγραφη υπεύθυνη δήλωση του άρθρου 8 παρ. 4 Ν. 1599/1986 τιμωρείται με φυλάκιση τουλάχιστον τριών μηνών. Εάν ο υπαίτιος αυτών των πράξεων σκόπευε να προσπορίσει στον εαυτόν του ή σε άλλον περιουσιακό όφελος βλάπτοντας τρίτον ή σκόπευε να βλάψει άλλον, τιμωρείται με κάθειρξη μέχρι 10 ετών.

**Μέθοδοι μηχανικής μάθησης για την υποβοήθηση της ιατρικής
διάγνωσης του SARS-CoV-2 με δεδομένα αξονικής τομογραφίας
θώρακα**

Βικτώρια Φίλη

Τριμελής Επιτροπή:

Μιχαήλ Σαβελώνας, Επίκουρος Καθηγητής (επιβλέπων)

Κωνσταντίνος Δελήμπασης, Αναπληρωτής Καθηγητής

Χαράλαμπος Καρανίκας, Επίκουρος Καθηγητής

Στον πατέρα μου

Ευχαριστίες

Σε αυτό το σημείο, θα επιθυμούσα να ευχαριστήσω όλους τους ανθρώπους που με στήριξαν και με βοήθησαν να φέρω εις πέρας την πτυχιακή μου εργασία. Αρχικά, θα ήθελα να ευχαριστήσω τον επιβλέποντα καθηγητή, κ. Μιχαήλ Σαβελώνα, επίκουρο καθηγητή του τμήματος Πληροφορικής με Εφαρμογές στη Βιοϊατρική του Πανεπιστημίου Θεσσαλίας, για τη στήριξη και την καθοδήγηση του στην υλοποίηση της παρούσας εργασίας. Επιπλέον, θα ήθελα να ευχαριστήσω τη μητέρα μου και τους φίλους μου που ήταν δίπλα μου και με στήριξαν κατά τη διάρκεια των σπουδών μου. Τέλος, θα ήθελα να πω ένα μεγάλο ευχαριστώ και να αφιερώσω την πτυχιακή μου εργασία στον εκλιπόντα πατέρα μου καθώς χάρις εκείνον μπόρεσα και κατάφερα να πραγματοποιήσω τα όνειρά μου και να είμαι ο άνθρωπος που είμαι σήμερα.

Βικτώρια Φίλη,

Λαμία 2022

Περίληψη

Ο κορονοϊός (COVID-19) έχει χαρακτηριστεί ως μία από τις μεγαλύτερες πανδημίες παγκοσμίως προκαλώντας σοβαρές αλλοιώσεις στο αναπνευστικό σύστημα του ανθρώπου. Όπως είναι φυσικό, αναπτύχθηκε πλήθος υπολογιστικών μεθόδων υποβοήθησης της ιατρικής διάγνωσης (Computer-aided Diagnosis – CAD). Η πλειοψηφία των μεθόδων αυτών καταλήγουν στην ταξινόμηση κάποιας εξεταζόμενης εικόνας ως θετικής ή αρνητικής ως προς τον COVID-19. Ένα μειονέκτημα αυτών των προσεγγίσεων είναι ότι συνήθως λειτουργούν ως «μαύρα κουτιά» τα οποία δεν κερδίζουν εύκολα την εμπιστοσύνη της ιατρικής κοινότητας. Στα πλαίσια της παρούσας εργασίας, αναπτύχθηκε μία πρωτότυπη μέθοδος ανάκτησης εικόνων με βάση το περιεχόμενο (Content-based Image Retrieval – CBIR) σε αξονικές τομογραφίες θώρακα (Computed Tomography – CT). Η προτεινόμενη μέθοδος, όπως κάθε μέθοδος CBIR, επιστρέφει πέρα από ένα αποτέλεσμα διάγνωσης θετικό/αρνητικό και μια λίστα από εικόνες όμοιες με την υπό εξέταση εικόνα, διευκολύνοντας έτσι τον χρήστη/ιατρό κατά την ερμηνεία τους αποτελέσματος. Η ποσοτικοποίηση της ομοιότητας συμπεριλαμβάνει τη στάθμιση της σημασίας της κάθε περιοχής της εικόνας ως προς τον COVID-19, όπως αυτή αποδίδεται από τη μέθοδο Χαρτογράφησης Ενεργοποίησης Κλάσης με Στάθμιση Βαθμίδας (Gradient-activated Class Activation Mapping – Grad-CAM). Με αυτόν τον τρόπο τα αποτελέσματα της ανάκτησης είναι πιο εύστοχα στον COVID-19, σε σύγκριση με την απλή ανάκτηση χωρίς στάθμιση.

Σε πρώτο στάδιο, επιχειρήθηκε η χρήση μεθόδων μηχανικής μάθησης για την ιατρική υποβοήθηση της διάγνωσης του κορονοϊού με αξονικές τομογραφίες θώρακα. Πιο συγκεκριμένα, επιδιώχθηκε η χρήση συστημάτων βαθιάς μάθησης για τη δυαδική ταξινόμηση εικόνων και την ανίχνευση εκείνων των χαρακτηριστικών που καθιστούν διακριτή την ασθένεια. Πραγματοποιήθηκαν διάφορες δοκιμές με γνωστές αρχιτεκτονικές συνελκτικών νευρωνικών δικτύων. Τελικά, η αρχιτεκτονική που επιλέχθηκε για τη δυαδική ταξινόμηση των αξονικών τομογραφιών θώρακα ήταν το DenseNet201. Η αξιολόγηση της ταξινόμησης πραγματοποιήθηκε με τα κλασσικά μέτρα και με τη χρήση της μεθόδου Grad-CAM, η οποία σταθμίζει και οπτικοποιεί τις περιοχές της εικόνας ανάλογα με την επιρροή τους στο αποτέλεσμα της ταξινόμησης. Τα αποτελέσματα έδειξαν ότι το νευρωνικό δίκτυο λειτουργεί αρκετά καλά επιτυγχάνοντας ποσοστό ολική ορθότητα περίπου 96%.

Στο δεύτερο στάδιο της εργασίας, αναπτύχθηκε ένα πρωτότυπο σύστημα CBIR με σκοπό την εύρεση όμοιων αξονικών τομογραφιών θώρακα μέσα σε μία μεγάλη βάση δεδομένων. Η σύγκριση της ομοιότητας των εικόνων πραγματοποιήθηκε με τον κλασσικό τρόπο υπολογισμού της απόστασης των διανυσμάτων των χαρακτηριστικών και με δύο ακόμη παραλλαγές αυτού. Στη μία περίπτωση χρησιμοποιήθηκαν, ως συντελεστής βαρύτητας, τα βάρη του Grad-CAM, ενώ στην άλλη περίπτωση συνδυάστηκαν οι δύο προαναφερθέντες με την προσθήκη και μίας υπερπαραμέτρου λ . Η αξιολόγηση της απόδοσης των παραλλαγών της μεθόδου έγινε με μετρικές αξιολόγησης συστημάτων CBIR, καθώς και με την εφαρμογή του Grad-CAM. Το σύστημα ανάκτησης εικόνων βάσει περιεχομένου κατόρθωσε ποσοστό μέσου όρου ακριβειών (MAP) ίσο με 96%.

Λέξεις κλειδιά: κορονοϊός, αξονική τομογραφία θώρακα, συνελκτικά νευρωνικά δίκτυα, DenseNet201, Grad-CAM, CBIR

Abstract

Coronavirus disease (COVID-19) has been characterized as one of the biggest pandemics worldwide, causing serious changes in the human respiratory system. Naturally, many computer-aided diagnosis (CAD) methods have been developed. The majority of these methods result in classifying an image under consideration as positive or negative for COVID-19. A disadvantage of these approaches is that they usually act as "black boxes", which do not easily gain the trust of medical community. In the context of this work, a novel method for Content-based Image Retrieval (CBIR) on computed tomography (CT) images is proposed. As in the case with CBIR methods, the proposed method, returns a list of images similar to the image under consideration, instead of a positive/negative diagnosis. Therefore, the proposed method facilitates the user/physician when interpreting the results. Quantification of similarity involves weighting the importance of each image region to COVID-19, as assigned by Gradient-activated Class Activation Mapping (Grad-CAM). In this way, the retrieval results are more targeted to COVID-19, when compared to simple unweighted retrieval.

As a first stage, the use of machine learning methods was attempted to aid the diagnosis of coronavirus with chest CT scans. More specifically, deep learning systems were used for the binary classification of images and the extraction of those features that make the disease distinct. Several tests were performed with known convolutional neural network architectures. Finally, the architecture chosen for binary classification of chest CT scans was DenseNet201. Classification evaluation was performed with the classical measures and using the Grad-CAM method, which weights and visualizes image regions according to their impact on the classification result. The results showed that the neural network works quite well achieving an accuracy rate of around 96%.

As a second stage, a prototype CBIR system was developed to find similar chest CT scans within a large database. The quantification of image similarity was carried out with the classical way of calculating the distance of feature vectors, with two more variants. In the first variant, the weights of Grad-CAM were used as a weighting factor, whereas in the second variant the two previous distances were combined using a weighting hyperparameter λ . The evaluation of all variants was performed with

CBIR system evaluation metrics, as well as with the application of Grad-CAM. The content-based image retrieval system achieved a mean accuracy rate (MAP) of 96%.

Keywords: COVID-19, computed tomography, convolutional neural networks, DenseNet201, Grad-CAM, CBIR

Περιεχόμενα

1.	Εισαγωγή	17
2.	Βιβλιογραφική Ανασκόπηση	19
3.	Θεωρητικό Υπόβαθρο	22
3.1	Συνελκτικά Νευρωνικά Δίκτυα	22
3.2	Αρχιτεκτονική DenseNet201	28
3.3	Μεταφορά μάθησης (Transfer Learning)	30
3.4	Χαρτογράφηση Ενεργοποίησης Κλάσης με Στάθμιση Βαθμίδας (Grad-guided CAM – Grad CAM)	31
3.5	Ανάκτηση Εικόνας Βάσει Περιεχομένου (Content Based Image Retrieval - CBIR).....	33
3.5.1	Μετρική Ομοιότητας.....	34
4.	Μεθοδολογία.....	34
4.1	Φόρτωση και Προεπεξεργασία Δεδομένων (Data Loading and Preprocessing)	34
4.2	Καθορισμός του Νευρωνικού Δικτύου	36
4.3	Εμπλουτισμός Δεδομένων / Data Augmentation.....	38
4.4	Εκπαίδευση του Νευρωνικού Δικτύου.....	38
4.5	Κατασκευή συστήματος ανάκτησης εικόνων βάσει περιεχομένου (Content Based Image Retrieval).....	39
4.6	Οπτικοποίηση με τη μέθοδο Grad-CAM	40
5.	Αποτελέσματα και Δοκιμές	41
5.1	Μετρικές Αξιολόγησης	42
5.1.1	Αξιολόγηση Νευρωνικού Δικτύου.....	42
5.1.2	Αξιολόγηση του συστήματος ανάκτησης βάσει περιεχομένου	44
5.2	Πειράματα και δοκιμές.....	46
5.3	Αξιολόγηση της επιλεγθείσας αρχιτεκτονικής	48
5.3.1	Αξιολόγηση ταξινόμησης	48

5.3.2 Εφαρμογή του συστήματος CBIR.....	51
5.3.4 Ειδικές περιπτώσεις χρήσης του συστήματος ανάκτησης εικόνων ανά επερώτηση.....	61
6. Συμπεράσματα και Μελλοντικές προσεγγίσεις	66

Ευρετήριο Εικόνων

Εικόνα 1 - Βασική δομή Τεχνητού Νευρωνικού Δικτύου.....	23
Εικόνα 2 - Παράδειγμα δομής Συνελκτικού Νευρωνικού Δικτύου για ταξινόμηση εικόνων	23
Εικόνα 3 - Διαδικασία Συνελκτικού Επιπέδου	24
Εικόνα 4 - Μέθοδοι Υποδειματοληψίας.....	25
Εικόνα 5 - Η δομή του νευρωνικού δικτύου πριν και μετά την εφαρμογή του επιπέδου απόσυρσης	27
Εικόνα 6 - Αρχιτεκτονική DenseNet.....	29
Εικόνα 7 - Δομή επιπέδου μετάβασης.....	29
Εικόνα 8 - Εφαρμογή της μεταφοράς μάθησης [19].....	31
Εικόνα 9 - Παράδειγμα εφαρμογής του Grad-CAM [6]	32
Εικόνα 10 - Διαδικασία ανάκτησης εικόνων βάσει περιεχομένου [50].....	33
Εικόνα 11 - Δείγμα του συνόλου εκπαίδευσης και του συνόλου δοκιμής.....	36
Εικόνα 12 - Πίνακας συνάφειας της ταξινόμησης.....	49
Εικόνα 13 - Καμπύλη ROC της ταξινόμησης.....	50
Εικόνα 14 - Γραφική παράσταση του σφάλματος εκπαίδευσης και δοκιμής.....	51
Εικόνα 15 - Γραφική παράσταση της ακρίβειας εκπαίδευσης και δοκιμής	51
Εικόνα 16 - Απλή εφαρμογή CBIR με εικόνα επερώτησης την Covid(31)	52
Εικόνα 17 - Εφαρμογή Grad-CAM στην απλή περίπτωση CBIR με εικόνα επερώτησης την Covid(31).....	52
Εικόνα 18 - Παραλλαγή CBIR/Grad-CAM με εικόνα επερώτησης την Covid(31)	53
Εικόνα 19 - Εφαρμογή Grad-CAM στην παραλλαγή CBIR/Grad-CAM με εικόνα επερώτησης την Covid(31).....	53
Εικόνα 20 - Παραλλαγή CBIR/hyperparameter με εικόνα επερώτησης την Covid(31)	53
Εικόνα 21 - Εφαρμογή Grad-CAM στην παραλλαγή CBIR/hyperparameter με εικόνα επερώτησης την Covid(31).....	53
Εικόνα 22 - Απλή εφαρμογή CBIR με εικόνα επερώτησης την Covid(65)	54
Εικόνα 23 - Εφαρμογή Grad-CAM στην απλή περίπτωση CBIR με εικόνα επερώτησης την Covid(65).....	54
Εικόνα 24 - Παραλλαγή CBIR/Grad-CAM με εικόνα επερώτησης την Covid(65)	54

Εικόνα 25 - Εφαρμογή Grad-CAM στην παραλλαγή CBIR/Grad-CAM με εικόνα επερώτησης την Covid(65)	54
Εικόνα 26 - Παραλλαγή CBIR/hyperparameter με εικόνα επερώτησης την Covid(65)	55
Εικόνα 27 - Εφαρμογή Grad-CAM στην παραλλαγή CBIR/hyperparameter με εικόνα επερώτησης την Covid(65)	55
Εικόνα 28 - Απλή εφαρμογή CBIR με εικόνα επερώτησης την Covid(1168)	55
Εικόνα 29 - Εφαρμογή Grad-CAM στην απλή περίπτωση CBIR με εικόνα επερώτησης την Covid(1168)	55
Εικόνα 30 - Παραλλαγή CBIR/Grad-CAM με εικόνα επερώτησης την Covid(1168)	56
Εικόνα 31 - Εφαρμογή Grad-CAM στην παραλλαγή CBIR/Grad-CAM με εικόνα επερώτησης την Covid(1168)	56
Εικόνα 32 - Παραλλαγή CBIR/hyperparameter με εικόνα επερώτησης την Covid(1168)	56
Εικόνα 33 - Εφαρμογή Grad-CAM στην παραλλαγή CBIR/hyperparameter με εικόνα επερώτησης την Covid(1168)	56
Εικόνα 34 - Απλή εφαρμογή CBIR με εικόνα επερώτησης την Non-Covid(37)	57
Εικόνα 35 - Εφαρμογή Grad-CAM στην απλή περίπτωση CBIR με εικόνα επερώτησης την Non-Covid(37)	57
Εικόνα 36 - Παραλλαγή CBIR/Grad-CAM με εικόνα επερώτησης την Non-Covid(37)	57
Εικόνα 37 - Εφαρμογή Grad-CAM στην παραλλαγή CBIR/Grad-CAM με εικόνα επερώτησης την Non-Covid(37)	57
Εικόνα 38 - Παραλλαγή CBIR/hyperparameter με εικόνα επερώτησης την Non-Covid(37)	58
Εικόνα 39 - Εφαρμογή Grad-CAM στην παραλλαγή CBIR/hyperparameter με εικόνα επερώτησης την Non-Covid(37)	58
Εικόνα 40 - Απλή εφαρμογή CBIR με εικόνα επερώτησης την Non-Covid(409)	58
Εικόνα 41 - Εφαρμογή Grad-CAM στην απλή περίπτωση CBIR με εικόνα επερώτησης την Non-Covid(409)	58
Εικόνα 42 - Παραλλαγή CBIR/Grad-CAM με εικόνα επερώτησης την Non-Covid(409)	59

Εικόνα 43 - Εφαρμογή Grad-CAM στην παραλλαγή CBIR/Grad-CAM με εικόνα επερώτησης την Non-Covid(409)	59
Εικόνα 44 - Παραλλαγή CBIR/hyperparameter με εικόνα επερώτησης την Non-Covid(409)	59
Εικόνα 45 - Εφαρμογή Grad-CAM στην παραλλαγή CBIR/hyperparameter με εικόνα επερώτησης την Non-Covid(409)	59
Εικόνα 46 - Απλή εφαρμογή CBIR με εικόνα επερώτησης την Non-Covid(1198)	60
Εικόνα 47 - Εφαρμογή Grad-CAM στην απλή περίπτωση CBIR με εικόνα επερώτησης την Non-Covid(1198)	60
Εικόνα 48 - Παραλλαγή CBIR/Grad-CAM με εικόνα επερώτησης την Non-Covid(1198)	60
Εικόνα 49 - Εφαρμογή Grad-CAM στην παραλλαγή CBIR/Grad-CAM με εικόνα επερώτησης την Non-Covid(1198)	60
Εικόνα 50 - Παραλλαγή CBIR/hyperparameter με εικόνα επερώτησης την Non-Covid(1198)	61
Εικόνα 51 - Εφαρμογή Grad-CAM στην παραλλαγή CBIR/hyperparameter με εικόνα επερώτησης την Non-Covid(1198)	61

1. Εισαγωγή

Στα τέλη Δεκεμβρίου του 2019, καταγράφηκε μία πληθώρα πνευμονολογικών κρουσμάτων άγνωστης αιτίας [1]. Έρευνες έδειξαν ότι οι ασθενείς είχαν προσβληθεί από έναν νέο ιό ο οποίος ονομάστηκε κορονοϊός ή SARS-CoV-2 (Severe Acute Respiratory Syndrome Coronavirus 2). Η μετάδοσή του ξεκίνησε από την πόλη Ουχάν της Κίνας όπου και πραγματοποιήθηκαν οι πρώτες έρευνες σχετικά με την προέλευση του. Ο κορονοϊός αποτελεί 7^ο μέλος της οικογένειας των RNA ιών που μπορούν να προσβάλουν τον άνθρωπο και παρουσιάζει 70% ομοιότητα ως προς την αλληλουχία με έναν παλαιότερο ιό, τον SARS-CoV [2]. Η ραγδαία εξάπλωσή του παγκοσμίως, οδήγησε τον Παγκόσμιο Οργανισμό Υγείας (WHO) να ανακηρύξει στις 11 Μαρτίου του 2020 την ύπαρξη μίας νέας πανδημίας [3].

Ο κορονοϊός αποτελεί μία μεταδοτική και επικίνδυνη ασθένεια η οποία μεταδίδεται από άνθρωπο σε άνθρωπο μέσω των σταγονιδίων που εξέρχονται από τη στοματική ή ρινική κοιλότητα. Τα πιο συχνά συμπτώματα του είναι ο ξηρός βήχας, ο πυρετός, η δυσκολία στην αναπνοή και η διάρροια [1]. Το ποσοστό των ανθρώπων που κινδυνεύει περισσότερο από την ασθένεια είναι οι ηλικιωμένοι και άτομα που ανήκουν σε ευπαθείς ομάδες. Από την απαρχή της πανδημίας σημειώθηκε μεγάλο ποσοστό θανάτων και ασθενών με σοβαρά αναπνευστικά προβλήματα. Ήταν ιδιαίτερα σημαντικό, λοιπόν, να υπάρξει περιορισμός της νόσου έως ότου βρεθούν τρόποι εξάλειψής της. Αυτός ήταν και ο λόγος, που ο Παγκόσμιος Οργανισμός Υγείας ανακήρυξε μέτρα προστασίας κατά του κορονοϊού. Το κυριότερο εξ αυτών αφορούσε τη έγκαιρη διάγνωση και απομόνωση των παθόντων [3].

Η διάγνωση πραγματοποιείται κατά κύριο λόγο με δύο τρόπους: την αλυσιδωτή αντίδραση πολυμεράσης (PCR) και την ταχεία δοκιμή αντιγόνου (Rapid Test) [4]. Αυτές οι μέθοδοι διάγνωσης, δυστυχώς, έχουν συγκεκριμένα όρια ως προς την αξιοπιστία τους καθώς μπορούν να δώσουν λανθασμένες προβλέψεις. Επιπλέον, η μέθοδος PCR είναι χρονοβόρα, έχει ένα κόστος, απαιτεί εξειδικευμένο προσωπικό και το αποτέλεσμά της εξαρτάται από τη φάση στην οποία βρίσκεται ο ιός. Οι παραπάνω περιορισμοί καθιστούν τις δύο αυτές μεθόδους προσωρινές αλλά όχι μόνιμες λύσεις. Το γεγονός αυτό οδήγησε πολλούς ερευνητές στην αναζήτηση νέων μεθόδων διάγνωσης που θα δίνουν περισσότερες πληροφορίες για τη λοίμωξη.

Στα πλαίσια της παρούσας εργασίας, μελετήθηκε ένας διαφορετικός τρόπος διάγνωσης του κορονοϊού. Πιο συγκεκριμένα, χρησιμοποιήθηκαν μέθοδοι μηχανικής μάθησης για την υποβοήθηση της διάγνωσης του COVID-19 με αξονικές τομογραφίες θώρακα. Η αξονική τομογραφία (CT-Scan) αποτελεί μία μορφή εξέτασης με ακτίνες-X και παρουσιάζει ευαισθησία στην ανάδειξη αλλοιώσεων στο αρχικό στάδιο του κορονοϊού [2]. Μεταξύ άλλων τεχνικών, τα συνελκτικά νευρωνικά δίκτυα (Convolutional Neural Network - CNN), τα οποία είναι νευρωνικά δίκτυα εμπρόσθιας τροφοδότησης, έχουν αποδειχθεί πολύ αποδοτικά ως προς την εύρεση χαρακτηριστικών. Βασικός στόχος της εργασίας αποτελεί η ταξινόμηση των εικόνων σε «Covid» και «Non-Covid» με τη χρήση ενός προεκπαιδευμένου μοντέλου. Σε δεύτερο στάδιο, επιδιώχθηκε η οπτικοποίηση των σημαντικών περιοχών του πνεύμονα με τη μέθοδο Grad-CAM η οποία παράγει έναν χάρτη με τα σημαντικότερα χαρακτηριστικά, όπως αυτά έχουν αξιολογηθεί από το νευρωνικό δίκτυο [6]. Στόχος, είναι η εύρεση τυχόν αλλοιώσεων που ενδέχεται να υπάρχουν στον πνεύμονα.

Από την απαρχή της πανδημίας οι ερευνητές έστρεψαν το ενδιαφέρον τους στον συνδυασμό συστημάτων ταξινόμησης με μεθόδους οπτικοποίησης, όπως το Grad-CAM. Στα πλαίσια της παρούσας ερευνητικής ανάλυσης, δοκιμάστηκε η ενσωμάτωση και μίας άλλης τεχνικής, της ανάκτησης εικόνων βάσει περιεχομένου (Content Based Image Retrieval – CBIR) [7]. Αρχικά, εφαρμόστηκε η απλή περίπτωση CBIR όπου ως μετρική απόστασης επιλέχθηκε η L_1 νόρμα. Στη συνέχεια, προτάθηκαν δύο παραλλαγές της μεθόδου CBIR οι οποίες αναφέρθηκαν περιληπτικά και παραπάνω. Η καινοτομία, λοιπόν, που προτείνεται αφορά την κατασκευή ενός συστήματος που θα ανακτά εικόνες βάσει των κοινών χαρακτηριστικών τους με μεγαλύτερη ευστοχία και ολική ορθότητα.

Η παρούσα πτυχιακή εργασία περιλαμβάνει επτά βασικές ενότητες. Η δεύτερη ενότητα αποτελεί μία σύντομη βιβλιογραφική ανασκόπηση των σημαντικότερων συναφών ερευνητικών έργων, ενώ στην τρίτη ενότητα παρουσιάζεται το θεωρητικό υπόβαθρο της εργασίας. Στην τέταρτη ενότητα, περιγράφεται αναλυτικά η προτεινόμενη μεθοδολογία, ενώ τα αποτελέσματά της παρουσιάζονται στην πέμπτη ενότητα. Τέλος, οι ενότητες 6 και 7 αναφέρονται στα συμπεράσματα και στις μελλοντικές προοπτικές αυτής της έρευνας.

2. Βιβλιογραφική Ανασκόπηση

Τα τελευταία χρόνια διεξήχθησαν διάφορες έρευνες που αφορούν τη χρήση μεθόδων μηχανικής μάθησης στη διάγνωση της πανδημίας του COVID-19. Παρακάτω παρέχεται μία σύντομη βιβλιογραφική ανασκόπηση των σημαντικότερων ερευνητικών έργων που προσεγγίζουν με παρόμοιο τρόπο το πρόβλημα της ταξινόμησης αξονικών τομογραφιών θώρακα.

Αρχικά, οι Butt et al. [8] επιδίωξαν την καθιέρωση ενός συστήματος διάκρισης της πνευμονίας που οφείλεται στον κορονοϊό από τις ιογενείς πνευμονίες της γρίπης-Α καθώς και από υγιείς περιπτώσεις, με αξονικές τομογραφίες θώρακα χρησιμοποιώντας το νευρωνικό δίκτυο ResNet18. Πολύ σύντομα προτάθηκαν μέθοδοι παρακολούθησης που βασίζονται στη μεταφορά μάθησης (transfer learning), εκ των οποίων οι περισσότερες χρησιμοποιούν δημοφιλείς υπάρχουσες αρχιτεκτονικές, όπως το VGG, το ResNet και το DenseNet. [9] Τον Απρίλιο του 2020, οι He et al. [10] κατασκεύασαν ένα δικό τους σύνολο δεδομένων με πολλές αξονικές τομογραφίες θώρακα για να αντιμετωπίσουν το πρόβλημα της υπερπροσαρμογής του δικτύου, ενώ για το στάδιο της εκπαίδευσης χρησιμοποίησαν τη μεταφορά μάθησης. Η υπερπροσαρμογή αποτελεί το φαινόμενο κατά το οποίο το δίκτυο επιτυγχάνει καλύτερη ταξινόμηση στα δεδομένα δοκιμής από ότι στα δεδομένα εκπαίδευσης. Ο περιορισμός του είναι αναγκαίος ώστε να επιτευχθεί η καλύτερη δυνατή γενίκευση του προβλήματος. [11] Οι Zheng et al. [12] σχεδίασαν ένα τρισδιάστατο βαθύ CNN (DeCoVNet) για την ανίχνευση του COVID-19 χρησιμοποιώντας φέτες αξονικής τομογραφίας θώρακα. Οι Wang, Liu, και Dou [13] ανακατασκεύασαν τη βασική δομή του μοντέλου COVID-Net με σκοπό τη βελτίωση της αρχιτεκτονικής και της εκμάθησης του νευρωνικού δικτύου. Οι δοκιμές που έγιναν ανέδειξαν την προτεινόμενη μέθοδο ως την καλύτερη σε σύγκριση με άλλες, πιο παλιές, μεθοδολογίες, επιτυγχάνοντας 90.83 ± 0.93 ολική ορθότητα. Η επιτυχία του δικτύου τους φάνηκε ιδιαίτερα με τη χρήση της μεθόδου Grad-CAM, ενώ μέσω της έρευνας τους έδειξαν ότι για κάθε σύνολο δεδομένων είναι απαραίτητη η αναπροσαρμογή της βασικής δομής οποιοδήποτε δικτύου. Την ίδια χρονιά, οι Walvekar και Shinde [14], κατασκεύασαν ένα σύστημα το οποίο βασίζεται στο μοντέλο Resnet-50 και θα μπορούσε να διαχωρίσει μία εικόνα ασθενή με κορονοϊό από εικόνες παθόντων άλλων δημοφιλών ασθενειών. Η επιτυχία του έφτασε το ποσοστό 96.23%. Μία σύγκριση που έγινε από τους Seum. et al. [15] ανάμεσα σε

γνωστές αρχιτεκτονικές ανέδειξε ως καλύτερα τα ResNet18 και DenseNet201. Η καινοτομία που προτάθηκε εδώ ήταν η ένταξη της τμηματοποίησης των εικόνων μέσω του U-Net αμέσως μετά την προεπεξεργασία του και προτού εισαχθούν στο νευρωνικό ώστε να αξιολογηθεί η απόδοση του κάθε μοντέλου. Το 2021, οι X. Gao, Y. Qian, και A. Gao [16], στα πλαίσια του διαγωνισμού MIA-COV19D, που διοργανώθηκε στα πλαίσια του International Conference on Computer Vision (ICCV), πρότειναν ένα σύστημα που βασίζεται σε μεθόδους μετασχηματισμού όρασης (vision transformer – ViT) οι οποίες τα τελευταία χρόνια έχουν αποδειχθεί να λειτουργούν καλύτερα από τα κλασσικά CNN καθώς χρησιμοποιούν λιγότερους υπολογιστικούς πόρους. Την ίδια λογική χρήσης λιγότερων υπολογιστικών πόρων ακολούθησε και ο Teli [17] με την κατασκευή του μοντέλου TeliNet. Το μοντέλο TeliNet αποδείχθηκε συγκριτικά με το κλασσικό VGG16 πιο αποδοτικό ως προς τη δυαδική ταξινόμηση των εικόνων με τη χρήση πολύ λιγότερων παραμέτρων. Οι Shah, Keniya, Shridharani, Punjabi, Shah, και Mehendale [18] πραγματοποίησαν μία σύγκριση ανάμεσα σε κλασσικά μοντέλα ταξινόμησης εκ των οποίων εκείνο που κατόρθωσε το μεγαλύτερο ποσοστό ολική ορθότητα (93%) ήταν το VGG19. Το VGG19 γενικά αποτελεί ένα μεγάλο δίκτυο το οποίο εμπεριέχει πολλές παραμέτρους. Γεννήθηκε, λοιπόν, η ανάγκη χρήσης πιο ελαφριών δικτύων ώστε να υπάρξει κέρδος ως προς την ταχύτητα και το υπολογιστικό κόστος. Έτσι, οι Song et al. [19] πρότειναν ένα νέο μοντέλο με όνομα DRENet η αρχιτεκτονική του οποίου βασίζεται στο προεκπαιδευμένο συνελκτικό δίκτυο Resnet50. Η οπτικοποίηση των σημαντικότερων περιοχών της εικόνας έγινε με τη χρήση της μεθόδου Grad-CAM και τελικά το σύστημα έδωσε ποσοστό επιτυχίας 93% το οποίο απεδείχθη, ύστερα από σύγκριση, μεγαλύτερα από άλλα κλασσικά προεκπαιδευμένα μοντέλα. Ένα από τα πιο αποδοτικά μοντέλα που προτάθηκε εκείνη τη χρονιά από τους Jaiswal, Gianchandani, Singh, Kumar, Kaur [20] αφορούσε τη χρήση της τεχνικής βαθιάς μεταφοράς μάθησης (Deep Transfer Learning - DTL) σε συνδυασμό με το δίκτυο DenseNet201 με σκοπό τον διαχωρισμό παθόντων και μη από τον κορονοϊό. Η προτεινόμενη μεθοδολογία άγγιξε το 97% ξεπερνώντας κλασσικά μοντέλα με μεγάλη απόδοση. Σε αυτό το σημείο πρέπει να σημειωθεί ότι τόσο τα ResNets όσο και τα DenseNets χαρακτηρίζονται από λίγες σχετικά παραμέτρους. Επιπρόσθετα, οι Hasan, Bao, Shawon, Huang [21] κινούμενοι στην ίδια κατηγορία δικτύων (DenseNet121) κατόρθωσαν την ταξινόμηση ενός μεγάλου συνόλου δεδομένων με ποσοστό ολική

ορθότητα 93% μόνο με τη χρήση 50 εποχών. Οι Halder και Datta [22] περιέγραψαν την ανάπτυξη ενός πλαισίου DL που περιλαμβάνει προεκπαιδευμένα μοντέλα (DenseNet201, VGG16, ResNet50V2 και MobileNet) ως τη βάση του, γνωστό ως KarNet. Μεταξύ των τεσσάρων μοντέλων του KarNet, το DenseNet201 επέδειξε εξαιρετική διαγνωστική ικανότητα σε μη επαυξημένα και σε επαυξημένα σύνολα δεδομένων, αντίστοιχα. Μεγάλη έμφαση στη δοκιμή παραμέτρων έδωσαν οι Mishra, Singh, και Joshi [23] οι οποίοι μετά από πολλούς και ποικίλους πειραματισμούς πέτυχαν εξαιρετικά καλή απόδοση στην περίπτωση της δυαδικής ταξινόμησης (COVID-19 έναντι κανονικής) με μέση ολική ορθότητα μεγαλύτερη από 99% τόσο στο VGG16 όσο και στο ResNet50. Αντιθέτως, στην ταξινόμηση πολλαπλών κλάσεων (COVID-19 έναντι φυσιολογικού έναντι πνευμονίας), το προτεινόμενο μοντέλο κατόρθωσε μέση ολική ορθότητα ταξινόμησης 86,74% και 88,52% χρησιμοποιώντας αρχιτεκτονικές VGG16 και ResNet50, αντίστοιχα. Αυτή η εργασία προσθέτει στην υπάρχουσα βιβλιογραφία ένα μοντέλο που βασίζεται σε μία απλή αρχιτεκτονική και στη μεταφορά μάθησης σε δύο στάδια για την ανίχνευση περιπτώσεων COVID-19 με βάση αξονικές τομογραφίες, ώστε να δημιουργηθεί μία ανταγωνιστική απόδοση συγκριτικά με τα κλασσικά μοντέλα. [24] Πιο πρόσφατες έρευνες φαίνεται να κεντρίζουν το ενδιαφέρον τους στη σύγκριση αρχιτεκτονικών για την ταξινόμηση των εικόνων, όπως αυτή που προτάθηκε από τους Chouat, Echtioui, Khemakhem, Zouch, Ghorbel, και Hamida [25], τον Ιανουάριο του 2022, η οποία ανέδειξε ως καλύτερο μοντέλο ταξινόμησης το VGG19. Οι Ashraf, Malik, και Khan [26] ασχολήθηκαν και εκείνοι με τη σύγκριση των μοντέλων ταξινόμησης με σκοπό την επίτευξη της καλύτερης διάγνωσης του κορονοϊού. Χρησιμοποιώντας το VGG19 οι Namani, Akkapeddi, και Bantu [27] πέτυχαν καλύτερη απόδοση (93%) η οποία φαίνεται να οφείλεται στη διαφορετική δομή του δικτύου και στην επιλογή συγκεκριμένων παραμέτρων. Έρχονται, λοιπόν, στο προσκήνιο και άλλοι παράγοντες που πρέπει να λαμβάνονται υπόψη πέρα από την επιλογή του προεκπαιδευμένου μοντέλου. Η έρευνα των Sassi, Ouarda, και Amar [28] πρότεινε ένα σύστημα ανάκτησης εικόνων βάσει περιεχομένου (CBIR) χρησιμοποιώντας μια νέα διαμόρφωση ενός συνελκτικού νευρωνικού δικτύου για την ανίχνευση του COVID-19. Τέλος, το άρθρο του Li [29] αποδεικνύει ότι ακόμη και τον Ιούλιο του 2022 τα μοντέλα που βρίσκονται στο επίκεντρο για τη διάγνωση της νόσου είναι τα VGG, ResNet και DenseNet με το τελευταίο να κερδίζει ως προς την απόδοση (95%).

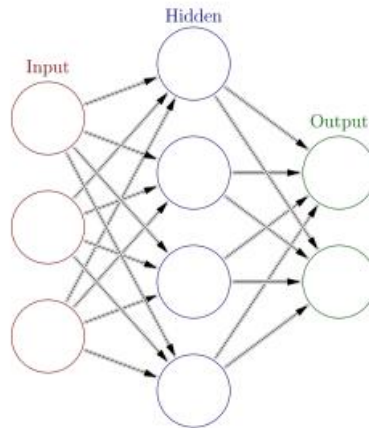
Παρατηρείται, λοιπόν, μεγάλο ενδιαφέρον από τον επιστημονικό κλάδο για τη ασθένεια του κορονοϊού καθώς έχουν προταθεί ποικίλες μεθοδολογίες ταξινόμησης και επεξεργασίας δεδομένων, όπου όλες αποσκοπούν στην αποδοτικότερη και ταχύτερη διάγνωση.

3. Θεωρητικό Υπόβαθρο

3.1 Συνελικτικά Νευρωνικά Δίκτυα

Τα Συνελικτικά Νευρωνικά Δίκτυα (Convolutional Neural Networks – CNNs) ανήκουν στον τομέα της Μηχανικής Μάθησης και η δομή τους βασίζεται στα Τεχνητά Νευρωνικά Δίκτυα (Artificial Neural Networks – ANN). Η καινοτομία τους συγκριτικά με άλλα μοντέλα ταξινόμησης έγκειται στο ότι λαμβάνουν υπόψιν τους τη χωρική πληροφορία [30]. Η απαρχή τους ορίζεται το 1980 με τη χρήση της αρχιτεκτονικής Neocognitron για την επίλυση προβλημάτων αναγνώρισης προτύπων. Το πρώτο, όμως, ολοκληρωμένο συνελικτικό νευρωνικό δίκτυο που εφαρμόστηκε για την ταξινόμηση εικόνων έκανε την εμφάνισή του το 2012 κερδίζοντας το διαγωνισμό ILSVRC (ImageNet Large-Scale Visual Recognition Challenge). Το AlexNet κατασκευάστηκε από τους Alex Krizhevsky, Sutskever, και Hinton και κατόρθωσε την ταξινόμηση μίας μεγάλης βάσης δεδομένων (ImageNet) σε 1000 κλάσεις [31]. Αποτελεί, λοιπόν, την πρώτη επιτυχή προσπάθεια για το σχεδιασμό ενός Βαθέως Νευρωνικού Δικτύου και την πηγή έμπνευσης όλων των υπόλοιπων αρχιτεκτονικών.

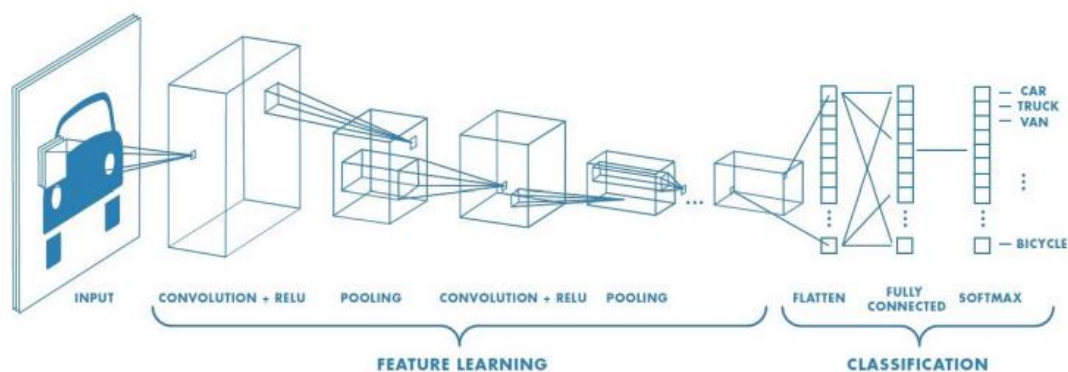
Τα τεχνητά νευρωνικά δίκτυα αποτελούν ευφυή υπολογιστικά συστήματα και είναι εμπνευσμένα από τα βιολογικά νευρωνικά δίκτυα του ανθρώπινου εγκεφάλου [32]. Πιο συγκεκριμένα, η δομή τους αποτελείται από συνδεδεμένους κόμβους κατανομημένους σε στρώματα οι οποίοι ονομάζονται νευρώνες, ενώ οι συνδέσεις που σχηματίζονται ανάμεσα τους χαρακτηρίζονται ως ακμές. Κάθε νευρώνας δέχεται ένα πλήθος εισόδων, εκτελεί έναν υπολογισμό και παράγει μία έξοδο [33]. Επιπλέον, οι νευρώνες κάθε κρυφού επιπέδου είναι πλήρως συνδεδεμένοι με τους νευρώνες του επόμενου και του προηγούμενου επιπέδου [34]. Γενικά, η επιλογή του αριθμού των κρυφών επιπέδων και των κόμβων ανά στρώμα βασίζεται στην εμπειρική προσέγγιση του προβλήματος. Ωστόσο, η πιο απλή δομή ενός τεχνητού νευρωνικού δικτύου χαρακτηρίζεται από ένα στρώμα εισόδου (Input Layer), ένα κρυφό στρώμα (Hidden Layer) και ένα στρώμα εξόδου (Output Layer), όπως φαίνεται και στην Εικόνα 1 [35].



Εικόνα 1 - Βασική δομή Τεχνητού Νευρωνικού Δικτύου

(Πηγή: https://en.wikipedia.org/wiki/Artificial_neural_network)

Γενικά, τα συνελκτικά νευρωνικά δίκτυα παρουσιάζουν σύνθετη αρχιτεκτονική η οποία δομείται από πολλά και διαφορετικά στρώματα. (Εικόνα 2) Ειδικότερα, αποτελούνται από ένα ή περισσότερα στρώματα συνέλιξης τα οποία συνοδεύονται από ένα στρώμα υποδειγματοληψίας. Το μοτίβο αυτό επαναλαμβάνεται συνήθως πολλές φορές με αποτέλεσμα να αυξάνεται η πολυπλοκότητα του δικτύου. Το τελευταίο στρώμα υποδειγματοληψίας μετατρέπεται σε ένα μονοδιάστατο διάνυσμα το οποίο ονομάζεται στρώμα ισοπέδωσης (flatten). Έπειτα, ακολουθεί ένα πλήρως συνδεδεμένο στρώμα (fully connected layer). Τέλος, εφαρμόζεται μία συνάρτηση ενεργοποίησης και προκύπτουν οι έξοδοι του δικτύου οι οποίες αναλογούν στις κλάσεις διαχωρισμού των δεδομένων [36].

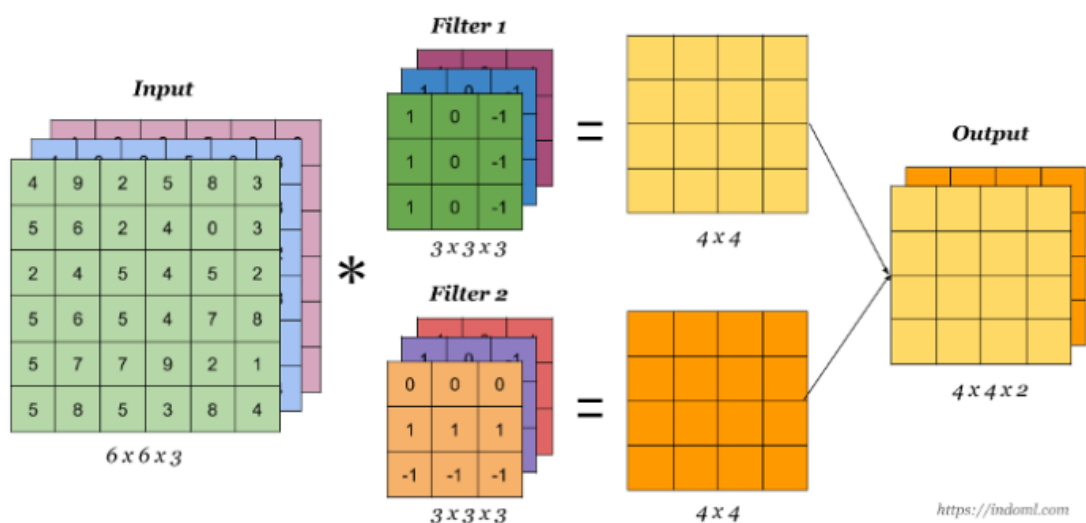


Εικόνα 2 - Παράδειγμα δομής Συνελκτικού Νευρωνικού Δικτύου για ταξινόμηση εικόνων

(Πηγή: <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>)

Συνελικτικό Στρώμα (Convolutional Layer)

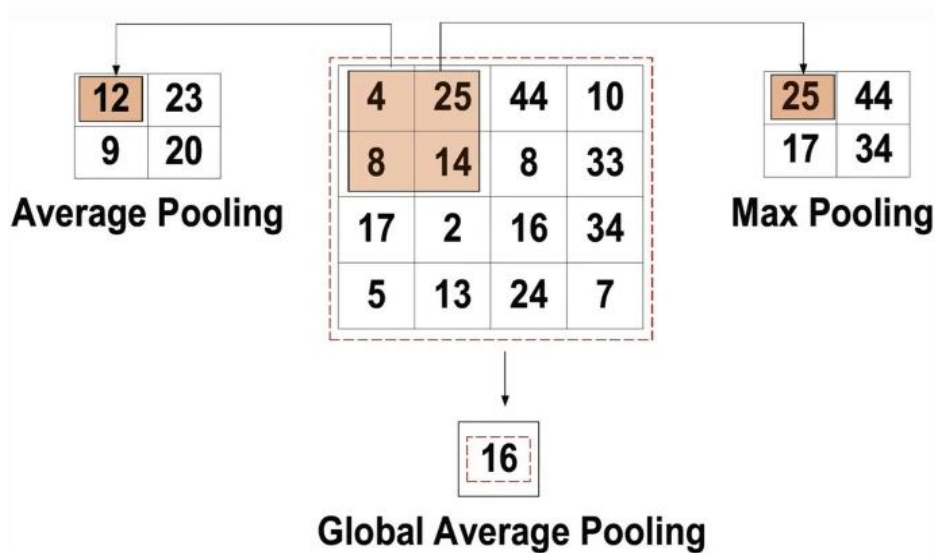
Το συνελικτικό στρώμα αποτελεί το πιο βασικό δομικό στοιχείο ενός συνελικτικού νευρωνικού δικτύου. Στην ουσία είναι μία μαθηματική πράξη η οποία πραγματοποιείται ώστε να συγχωνευτεί πληροφορία από δύο διαφορετικές πηγές (Εικόνα 3). Κάθε συνελικτικό στρώμα περιλαμβάνει φίλτρα των οποίων οι τιμές καθορίζονται κατά το στάδιο της εκπαίδευσης. Τα φίλτρα συνελίσσονται με την εικόνα εισόδου ώστε να παραχθούν οι χάρτες χαρακτηριστικών οι οποίοι περιέχουν τις θέσεις ενεργοποίησης των φίλτρων. Συγκεκριμένα, το φίλτρο ολισθαίνει πάνω στην εικόνα εισόδου και πραγματοποιείται η πράξη της συνέλιξης ανάμεσα στο φίλτρο και σε ένα παράθυρο της εικόνας εισόδου το οποίο οφείλει να έχει το ίδιο μέγεθος με το φίλτρο ώστε να είναι δυνατός ο πολλαπλασιασμός στοιχείο προς στοιχείο. Η περιοχή όπου λαμβάνει χώρα η πράξη της συνέλιξης χαρακτηρίζεται ως οπτικό πεδίο νευρώνων (receptive field) [37]. Οι διαστάσεις των φίλτρων πρέπει να είναι μικρότερες από αυτές της εικόνας εισόδου. Ο χάρτης χαρακτηριστικών που θα προκύψει από τη συνέλιξη μιας εικόνας εισόδου $M \times N \times 3$ με ένα φίλτρο $m \times n$ θα είναι διαστάσεων $(M-m+1) \times (N-n+1) \times 3$ με $m, n < M, N$. Στην πραγματικότητα, μία εικόνα αναπαρίσταται από μία τρισδιάστατη μήτρα με διαστάσεις ύψους M , πλάτους N και βάθους 3, όπου το βάθος αντιστοιχεί στον αριθμό των χρωματικών καναλιών RGB [38]. Από την άλλη πλευρά, το βάθος ενός συνελικτικού επιπέδου ισούται με το πλήθος των φίλτρων που εφαρμόζονται. Είναι σημαντικό να σημειωθεί ότι οι χάρτες χαρακτηριστικών που προκύπτουν συρρικνώνονται σε κάθε στρώμα.



Εικόνα 3 - Διαδικασία Συνελικτικού Επιπέδου

Στρώμα Υποδειματοληψίας (Pooling Layer)

Το στρώμα υποδειματοληψίας συναντάται συνήθως μετά τα συνελκτικά και εφαρμόζεται για τη μείωση της χωρικής διάστασης στους χάρτες χαρακτηριστικών. Συγκεκριμένα, μικραίνει το ύψος και το πλάτος ενός χάρτη χαρακτηριστικών χωρίς όμως να επηρεάζεται το βάθος του. Σε σύγκριση με την πράξη της συνέλιξης η υποδειματοληψία δεν περιέχει παραμέτρους. Το μόνο που είναι αναγκαίο να καθοριστεί είναι το βήμα και το μέγεθος του παραθύρου το οποίο ολισθαίνει πάνω στην είσοδο [36]. Υπάρχουν δύο βασικές κατηγορίες υποδειματοληψίας οι οποίες είναι: η υποδειματοληψία μεγίστου (max pooling) και η υποδειματοληψία μέσου όρου (average pooling). Στην πρώτη κατηγορία επιλέγεται η μέγιστη τιμή για ένα παράθυρο της εικόνας, ενώ στη δεύτερη κατηγορία η τελική τιμή αποτελεί το μέσο όρο των τιμών του κάθε παραθύρου (Εικόνα 4). Η μείωση της χωρικής διάστασης είναι απαραίτητη και αποσκοπεί στην ελαχιστοποίηση του πλήθους των παραμέτρων και κατά συνέπεια στον περιορισμό της πολυπλοκότητας του δικτύου. Επιπρόσθετα, η υποδειματοληψία αποτελεί και μία μέθοδο αντιμετώπισης του φαινομένου της υπερεκπαίδευσης [39].



Εικόνα 4 - Μέθοδοι Υποδειματοληψίας

Πλήρες Συνδεδεμένο Στρώμα (Fully Connected Layer)

Τα πλήρως συνδεδεμένα στρώματα αποτελούν το τελικό στάδιο για την ολοκλήρωση της δομής ενός συνελκτικού νευρωνικού δικτύου και τοποθετούνται συνήθως μετά από τελευταίο συνελκτικό στρώμα. Πιο συγκεκριμένα, σε ένα πλήρως

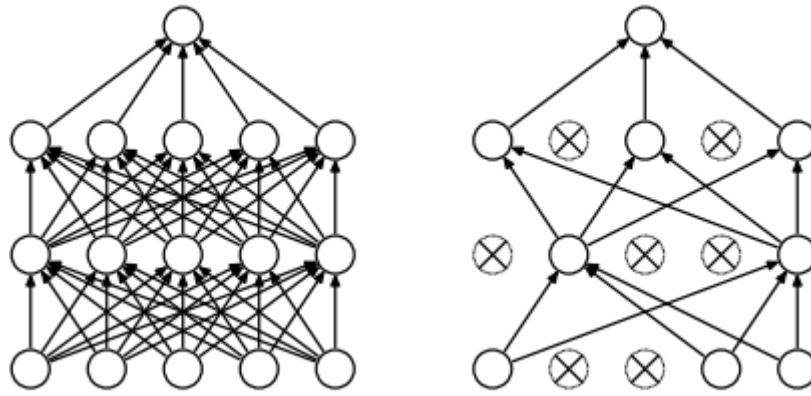
συνδεδεμένο στρώμα όλοι οι νευρώνες συνδέονται με τους νευρώνες του προηγούμενου και του επόμενου επιπέδου. Το τελευταίο πλήρως συνδεδεμένο αποτελείται από ένα πλήθος νευρώνων ισάξιο του αριθμού των κλάσεων της ταξινόμησης. Σκοπός του αποτελεί η εξαγωγή ενός μονοδιάστατου διανύσματος χαρακτηριστικών [36].

Συνάρτηση Ενεργοποίησης (Activation Function)

Η συνάρτηση ενεργοποίησης αποτελεί μία μαθηματική αναπαράσταση η οποία καθορίζει την έξοδο ενός νευρώνα. Στην ουσία, ρυθμίζει κάθε φορά εάν ένας νευρώνας πρέπει να ενεργοποιηθεί ή να απενεργοποιηθεί εισάγοντας τη μη γραμμικότητα στην έξοδο του νευρώνα, δηλαδή μετατρέπει ένα πρόβλημα από γραμμικό σε μη γραμμικό. Στην περίπτωση που δεν εισαχθεί η μη γραμμικότητα, το δίκτυο μπορεί να μάθει μόνο γραμμικές συναρτήσεις διαχωρισμού. Επιπλέον, χρησιμοποιείται και σε περιπτώσεις όπου απαιτείται κανονικοποίηση της εξόδου του νευρώνα. Υπάρχουν διάφορες συναρτήσεις ενεργοποίησης. Σε προβλήματα ταξινόμησης χρησιμοποιούνται κατά κύριο λόγο: η συνάρτηση Ανόρθωσης (ReLU), η Σιγμοειδής συνάρτηση (Sigmoid) [36].

Στρώμα Απόσυρσης (Dropout)

Το στρώμα απόσυρσης δεν αποτελεί βασικό δομικό στοιχείο ενός συνελκτικού νευρωνικού δικτύου και εφαρμόζεται συνήθως πριν από το πλήρως συνδεδεμένο στρώμα. Πρόκειται για μία τεχνική που χρησιμοποιείται για την αποφυγή του προβλήματος της υπερπροσαρμογής του δικτύου στα δεδομένα εκπαίδευσης. Πιο συγκεκριμένα, θέτει τυχαία, σε κάθε επανάληψη του αλγόριθμου εκπαίδευσης, ένα ποσοστό των Ενεργοποιήσεων (Activations) του προηγούμενου επιπέδου ίσο με το 0. Το ποσοστό απόσυρσης επιλέγεται από τον εκάστοτε χρήστη, δηλαδή αποτελεί μία υπερπαραμέτρο του δικτύου. Οι αποσυρόμενοι νευρώνες δε συμμετέχουν στον τρέχοντα κύκλο εκπαίδευσης και έτσι οι ενεργοί νευρώνες αναγκάζονται να εντοπίσουν νέα χαρακτηριστικά. Το στρώμα απόσυρσης εφαρμόζεται μόνο κατά το στάδιο της εκπαίδευσης του δικτύου [40].



(α) Κανονική μορφή ενός νευρωνικού δικτύου (β) Μετά την εφαρμογή του dropout
 Εικόνα 5 - Η δομή του νευρωνικού δικτύου πριν και μετά την εφαρμογή του επιπέδου απόσυρσης

Κανονικοποίηση Δέσμης (Batch Normalization)

Η κανονικοποίηση δέσμης αποσκοπεί στον περιορισμό της εσωτερικής μεταβολής συνδιακύμανσης (internal covariance shift) η οποία καθυστερεί την εκπαίδευση του δικτύου επιβάλλοντας τη χρήση μικρών ρυθμών εκπαίδευσης (learning rate). Η εσωτερική μεταβολή συνδιακύμανσης παρατηρείται κατά το στάδιο εκπαίδευσης του δικτύου, καθώς οι παράμετροι των στρώματων μεταβάλλονται, μεταβάλλεται και η κατανομή των εισόδων στο επόμενο στρώμα. Οι έξοδοι, λοιπόν, κάθε στρώματος κανονικοποιούνται ώστε να έχουν μηδενική μέση τιμή και μοναδιαία διακύμανση. Η κανονικοποίηση δεν εφαρμόζεται σε όλο το σύνολο δεδομένων αλλά σε μικρά πακέτα. Εν κατακλείδι, η κανονικοποίηση δέσμης επιτρέπει τη χρήση μεγαλύτερου ρυθμού μάθησης, ενώ ταυτόχρονα βελτιώνει τη βελτιστοποίηση [41].

Συνοπτικά, τα συνελκτικά νευρωνικά δίκτυα δέχονται ως είσοδο εικόνες (έγχρωμες ή άχρωμες) διαφορετικού μεγέθους. Στη συνέχεια, διέρχονται από διάφορα στρώματα και εξάγονται σταδιακά όλο και πιο σύνθετα χαρακτηριστικά. Η εξαγωγή των χαρακτηριστικών στηρίζεται στην προσαρμογή των διαφόρων φίλτρων των συνελκτικών επιπέδων. Στα χαμηλότερα στρώματα εξάγονται χαρακτηριστικά πρώτης τάξης (ακμές, σημεία), ενώ σε υψηλότερα στρώματα εξάγονται πιο σύνθετες δομές (ψηφία, γράμματα, πρόσωπα). Τα χαρακτηριστικά που προκύπτουν παράγουν τελικά τους χάρτες χαρακτηριστικών (feature maps) [42].

Πλέον, τα βαθιά νευρωνικά δίκτυα είναι πολύ χρήσιμα εργαλεία και παρουσιάζουν ποικίλες εφαρμογές σε διάφορους τομείς. Οι πιο σημαντικές από αυτές αφορούν την ταξινόμηση εικόνων (Image Classification), την αναγνώριση αντικειμένων (Object Detection) και ψηφίων (Digit Recognition) τα οποία αποτελούν

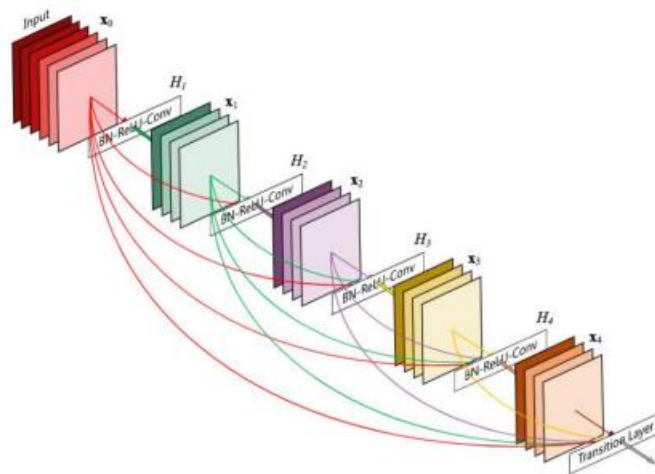
πεδία της Υπολογιστικής Όρασης. Επιπλέον, χρησιμοποιούνται ιδιαίτερα στην επίλυση πιο σύνθετων προβλημάτων όπως είναι η αναγνώριση της νοηματικής γλώσσας (Sign Language Recognition) [43].

3.2 Αρχιτεκτονική DenseNet201

Στις κλασσικές αρχιτεκτονικές συνελκτικών νευρωνικών δικτύων, η πληροφορία διέρχεται από το ένα στρώμα στο άλλο. Με την πάροδο των χρόνων, όμως, υπήρξε ανάγκη αύξησης του βάθους του δικτύου το οποίο οδήγησε στην εμφάνιση του προβλήματος των Εξαφανιζόμενων Κλίσεων (Vanishing Gradient) [44]. Επιπλέον, παρατηρήθηκε ότι η προσθήκη νέων στρωμάτων οδήγησε στην ταυτόχρονη αύξηση του πλήθους παραμέτρων και της υπολογιστικής πολυπλοκότητας του δικτύου [45].

Τη λύση στα προαναφερθέντα προβλήματα έδωσαν οι He et al. [45] προτείνοντας το δίκτυο ResNet. Τα Residual Networks συμβάλουν στην εκπαίδευση πολύ βαθέων νευρωνικών δικτύων τα οποία αποτελούνται από 150+ στρώματα και παρουσιάζουν για πρώτη φορά την έννοια της παράβλεψης σύνδεσης (skip connection). Το DenseNet αποτελεί προέκταση αυτού με αποτέλεσμα τα δύο δίκτυα να παρουσιάζουν πολλές ομοιότητες. Συγκεκριμένα, χαρακτηρίζεται ως ένα βαθύ συνελκτικό νευρωνικό δίκτυο, ο σχεδιασμός του οποίου αποσκοπεί στην εξάλειψη της αλλοίωσης των χαρακτηριστικών της εικόνας. Η αρχιτεκτονική του αποτελείται από πολλά στρώματα κάθε ένα από τα οποία περιέχει πολλές πυκνές στοίβες (Dense Blocks), όπως φαίνεται και στην Εικόνα 5. Μία πυκνή στοίβα συνδέεται με την επόμενη της μέσω των στρωμάτων μετάβασης (Transition Layers) και συντελείται από τα επιμέρους τρία στρώματα: [46]

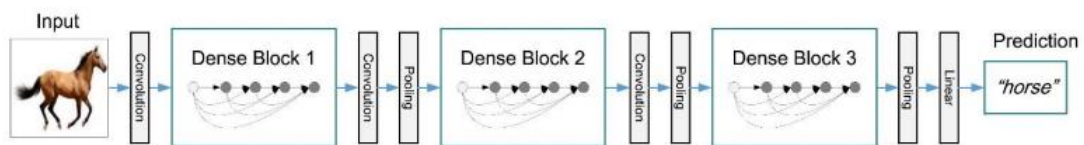
1. Κανονικοποίηση Δέσμης δεδομένων εισόδου (Batch Normalization)
2. Συνάρτηση Ενεργοποίησης ReLU
3. Συνελκτικό στρώμα με φίλτρο μεγέθους 3x3



Εικόνα 6 - Αρχιτεκτονική DenseNet

Το στρώμα μετάβασης περιέχει και αυτό συγκεκριμένα στρώματα τα οποία είναι:

1. Κανονικοποίηση Δέσμης δεδομένων εισόδου (Batch Normalization)
2. Συνελκτικό στρώμα με 1x1 πυρήνα
3. Υποδειγματοληψία μέσου όρου (Average Pooling)



Εικόνα 7 - Δομή επιπέδου μετάβασης

Στη συνέχεια, από κάθε στρώμα εξάγεται ένα πλήθος χαρτών χαρακτηριστικών το οποίο δίνεται ως είσοδος στο επόμενο στρώμα, ενώ το τελευταίο στρώμα θα λάβει ως είσοδο το άθροισμα των χαρτών χαρακτηριστικών όλων των προηγούμενων στρωμάτων. Η συνένωση αυτή θα έχει ως απόρροια τη δημιουργία ενός ανύσματος (tensor). Κάθε στρώμα, λοιπόν, δέχεται μία πληθώρα γνώσεων από τα προηγούμενα στρώματα με αποτέλεσμα το δίκτυο να είναι μικρότερο και κατά επέκταση να καταλαμβάνει λιγότερη μνήμη. Συνεπώς, ο αριθμός των επιπλέον χαρακτηριστικών κάθε στρώματος ορίζει ένα νέο μέγεθος που ονομάζεται ρυθμός ανάπτυξης (growth rate). Επομένως, η βασική καινοτομία που προτείνει το DenseNet είναι η συνένωση (concatenation) των χαρτών χαρακτηριστικών [\[46\]](#).

Η πυκνή συνδεσιμότητα του μοντέλου παρέχει το πλεονέκτημα της διάδοσης του σφάλματος μεταξύ των στρώματων με αποτέλεσμα να αντιμετωπίζεται το φαινόμενο της εξαφανιζόμενης κλίσης της παραγώγου. Στα πλεονεκτήματα συγκαταλέγεται και η επαναχρησιμοποίηση χαρακτηριστικών προερχόμενων από διαφορετικά στρώματα τα οποία σχηματίζουν πιο εμπλουτισμένα μοτίβα. Τα δύο αυτά πλεονεκτήματα σε συνδυασμό με το μικρό υπολογιστικό κόστος καθιστούν αυτό το μοντέλο μία από τις πιο εύστοχες επιλογές για την επίλυση σύνθετων προβλημάτων.

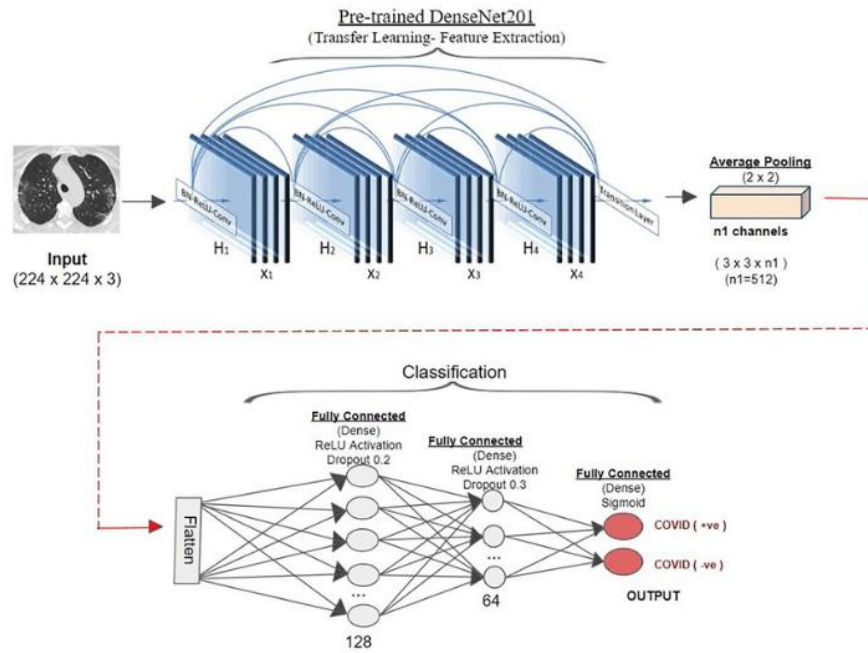
DenseNet201

Το DenseNet201 αποτελείται από 201 στρώματα και περιλαμβάνει συνολικά 20 εκατομμύρια παραμέτρους. Συγκριτικά με άλλα πιο μικρά δίκτυα, όπως το VGG16, περιέχει λίγες παραμέτρους. Αυτό οφείλεται στις πυκνές συνδέσεις που σχηματίζονται ανάμεσα στα στρώματα του δικτύου καθιστώντας έτσι εφικτή τη χρήση λιγότερων παραμέτρων [\[47\]](#).

3.3 Μεταφορά μάθησης (Transfer Learning)

Τρωτό σημείο των συνελκτικών νευρωνικών δικτύων αποτελεί η ανάπτυξη της ικανότητας γενίκευσης που απαιτεί ένα μεγάλο σύνολο δεδομένων. Εύκολα γίνεται σαφές ότι είναι μία αρκετά ακριβή και χρονοβόρα διαδικασία. Τη λύση σε αυτό το πρόβλημα έρχεται να δώσει η μεταφορά μάθησης που έχει ως βασική αρχή την επαναχρησιμοποίηση ήδη εξαχθέντων πόρων από δίκτυα που έχουν προεκπαιδευτεί. Προκειμένου, να αποφευχθεί η επανεκπαίδευση τους εξ' ολοκλήρου, ένα κομμάτι της αρχιτεκτονικής του παραμένει το ίδιο και αφαιρείται μόνο το τελευταίο στρώμα εξόδου. Πιο συγκεκριμένα, η μεταφορά μάθησης έγκειται στη μεταβίβαση των βαρών ενός προεκπαιδευμένου δικτύου. Στη συνέχεια, μπορούν να προστεθούν και άλλα στρώματα ανάλογα με τις ανάγκες του προβλήματος.

Στην περίπτωση της κατηγοριοποίησης εικόνων τα μοντέλα ταξινόμησης αναγνωρίζουν αδρά χαρακτηριστικά και σχήματα στα κατώτερα στρώματα και πιο σύνθετα μοτίβα στα ανώτερα. Ο συνδυασμός των μοντέλων ταξινόμησης με την τεχνική της μεταφοράς μάθησης, παρείχε κέρδος σε χρόνο και σε υπολογιστικούς πόρους. (Εικόνα 8) [\[48\]](#).



Εικόνα 8 - Εφαρμογή της μεταφοράς μάθησης [19]

3.4 Χαρτογράφηση Ενεργοποίησης Κλάσης με Στάθμιση Βαθμίδας (Grad-guided CAM – Grad CAM)

Το νευρωνικό δίκτυο ταξινομεί τις εικόνες βάσει ενός διανύσματος χαρακτηριστικών. Η οπτικοποίηση εστιάζει στις περιοχές από τις οποίες εξάγονται αυτά τα χαρακτηριστικά με αποτέλεσμα να παρουσιάζει διττό ρόλο ως προς το πρόβλημα. Συγκεκριμένα, παρέχει σημαντικές πληροφορίες που σχετίζονται με τη λειτουργία του νευρωνικού, ενώ ταυτόχρονα μπορεί να αποτελέσει ένα ισχυρό μέσο υποβοήθησης στον ιατρικό κλάδο, καθώς επισημαίνει τις περιοχές στις οποίες εντοπίζεται η ασθένεια.

Στην παρούσα εργασία, χρησιμοποιήθηκε η μέθοδος «Χαρτογράφησης Ενεργοποίησης Κλάσης με Στάθμιση Βαθμίδας» (Grad-guided CAM – Grad CAM) [6]. Πιο συγκεκριμένα, η μέθοδος χρησιμοποιεί τα χαρακτηριστικά που εξάγονται από κάποιο από τα συνελκτικά στρώματα, συνήθως χρησιμοποιείται το τελευταίο στρώμα, τα οποία συντελούν στην παραγωγή ενός χάρτη ενεργοποίησης. Ο χάρτης ενεργοποίησης αποσκοπεί στον εντοπισμό και στην οπτικοποίηση των σημαντικότερων περιοχών της εικόνας. Παρακάτω παρουσιάζεται η διαδικασία έως την τελική οπτικοποίηση. Αρχικά, υπολογίζεται η κλίση της πρόβλεψης y^c της κλάσης c ως προς τους χάρτες χαρακτηριστικών A^k κάποιου συνελκτικού επιπέδου $\left(\frac{\partial y^c}{\partial A^k}\right)$,

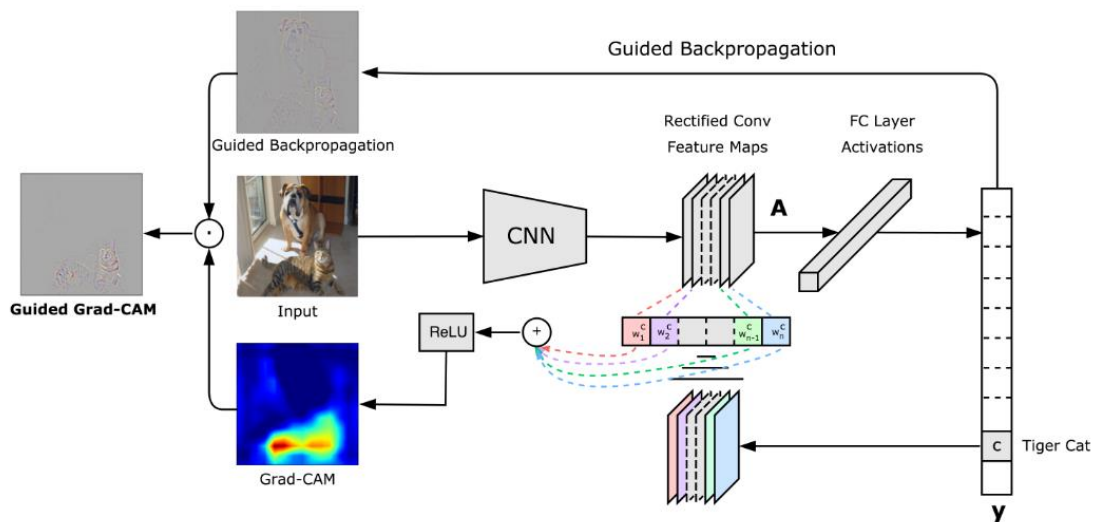
προτού εφαρμοστεί η συνάρτηση ενεργοποίησης softmax. Έπειτα εφαρμόζεται η τεχνική «Global Average Pooling» των βαθμίδων ώστε να ληφθούν τα απαιτούμενα βάρη για κάθε χάρτη χαρακτηριστικών. Το βάρος περιγράφει τη σημασία του k χάρτη χαρακτηριστικών για την κλάση c [6]. Ο υπολογισμός των βαρών περιγράφεται από τη σχέση (1).

$$a_k^c = \frac{1}{Z} \sum_i \sum_j \frac{\partial y^c}{\partial A_{ij}^k} \quad (1)$$

Τέλος, γίνεται υπολογισμός του χάρτη ενεργοποίησης (heatmap). Κάθε χάρτης χαρακτηριστικών πολλαπλασιάζεται με το αντίστοιχο βάρος του. Στη συνέχεια, εφαρμόζεται η συνάρτηση ανόρθωσης (ReLU) η οποία αποσκοπεί στον εντοπισμό των χαρακτηριστικών με θετική επιρροή στην κλάση ενδιαφέροντος. Τα εικονοστοιχεία που σχετίζονται με την επιθυμητή κλάση είναι εκείνα που η αύξηση της έντασης τους προκαλεί αύξηση και στην y^c . Η παραπάνω διαδικασία αναπαρίσταται από τη σχέση (2) [6].

$$L_{Grad-CAM}^c = ReLU(\sum_k a_k^c A^k) \quad (2)$$

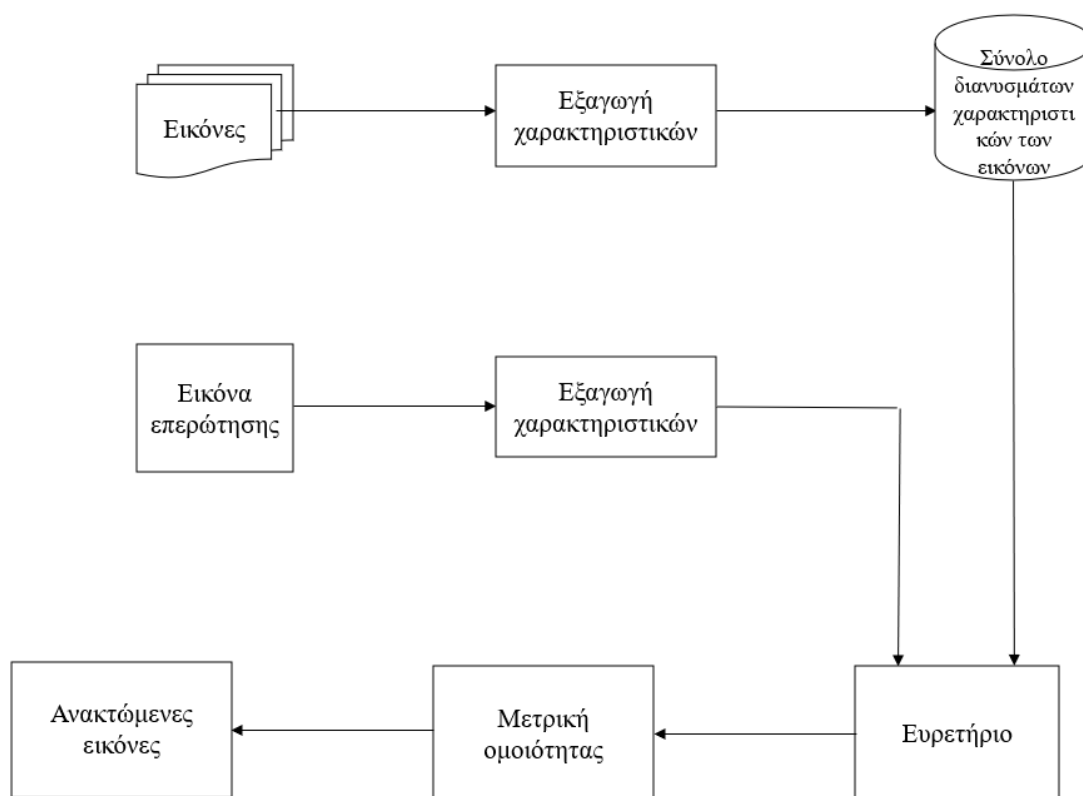
Η βασική λειτουργία της μεθόδου φαίνεται στην Εικόνα 9



Εικόνα 9 - Παράδειγμα εφαρμογής του Grad-CAM [6]

3.5 Ανάκτηση Εικόνας Βάσει Περιεχομένου (Content Based Image Retrieval - CBIR)

Η ανάκτηση εικόνας βάσει περιεχομένου αποσκοπεί στην εύρεση όμοιων εικόνων σε μία μεγάλη βάση δεδομένων αξιοποιώντας συγκεκριμένες υπολογιστικές τεχνικές [49]. Πιο συγκεκριμένα, εντοπίζει τις εικόνες βάσει του οπτικού περιεχομένου το οποίο περιλαμβάνει την υφή, το χρώμα, το σχήμα (χαρακτηριστικά χαμηλού επιπέδου) και οποιαδήποτε άλλη πληροφορία μπορεί να εξαχθεί από την εικόνα [50]. Τα συστήματα ανάκτησης εικόνων βάσει περιεχομένου μπορούν να συμβάλλουν στην ορθή διάγνωση και θεραπεία μιας ασθένειας, καθώς και στην εκπαίδευση ενός ιατρού. Η αρχιτεκτονική ενός τέτοιου συστήματος περιγράφεται συνοπτικά από το παρακάτω σχήμα (Εικόνα 10) [49].



Εικόνα 10 - Διαδικασία ανάκτησης εικόνων βάσει περιεχομένου [50]

Αρχικά, ο χρήστης δίνει ως είσοδο στο σύστημα μία εικόνα επερώτησης (query image) από την οποία εξάγονται ορισμένα οπτικά χαρακτηριστικά (feature extraction). Τα εξαγόμενα χαρακτηριστικά σχηματίζουν ένα διάνυσμα χαρακτηριστικών (feature vector). Παράλληλα, πραγματοποιείται εξαγωγή χαρακτηριστικών από όλες τις εικόνες που συντελούν τη βάση δεδομένων τα οποία

αποθηκεύονται σε ένα ευρετήριο. Στη συνέχεια, αναζητούνται όμοιες εικόνες, δηλαδή εικόνες της βάσης δεδομένων που παρουσιάζουν όμοια χαρακτηριστικά με αυτά της εικόνας επερώτησης. Το σύστημα εφαρμόζει στο στάδιο της σύγκρισης ένα μέτρο ομοιότητας και επιστρέφει τις εικόνες (retrieved images) που εντοπίζονται πιο κοντά στην εικόνα επερώτησης [50].

3.5.1 Μετρική Ομοιότητας

Η έξοδος ενός τέτοιου συστήματος αποτελεί στην ουσία μία λίστα εικόνων η οποία είναι ταξινομημένη βάσει του βαθμού ομοιότητας με την εικόνα επερώτησης. Ειδικότερα, η μέτρηση της ομοιότητας δύο εικόνων πραγματοποιείται με τη χρήση μετρικών ομοιότητας. Η διαφορά υπολογίζεται μεταξύ του διανύσματος χαρακτηριστικών του ερωτήματος και των διανυσμάτων χαρακτηριστικών της βάσης δεδομένων χρησιμοποιώντας μετρικές απόστασης. Η μικρή διαφορά μεταξύ δύο διανυσμάτων χαρακτηριστικών δείχνει τη μεγάλη ομοιότητα και τη μικρή απόσταση [51].

Γενικά, η έννοια της ομοιότητας είναι συμπληρωματική με την έννοια της απόστασης [51]. Οι αποστάσεις που χρησιμοποιούνται κατά κόρον σε CBIR συστήματα είναι οι εξής: η απόσταση μπλοκ πόλεων, η ευκλείδεια απόσταση και το άθροισμα της απόλυτης διαφοράς [52] [53].

$$\text{Απόσταση μπλοκ πόλεων ή } L_1 \text{ νόρμα: } d_1 (H, K) = \sum_m |h_i - k_i| \quad (3)$$

$$\text{Ευκλείδεια απόσταση ή } L_2 \text{ νόρμα: } d_2 (H, K) = \sqrt{(\sum_m (h_i - k_i)^2)} \quad (4)$$

$$\text{Άθροισμα απόλυτης διαφοράς ή } L_{inf} \text{ νόρμα: } d_{inf} (H, K) = \max (|h_i - k_i|) \quad (5)$$

4. Μεθοδολογία

Η ενότητα αυτή παρουσιάζει αναλυτικά τα βήματα δημιουργίας των συστημάτων ταξινόμησης και ανάκτησης εικόνων.

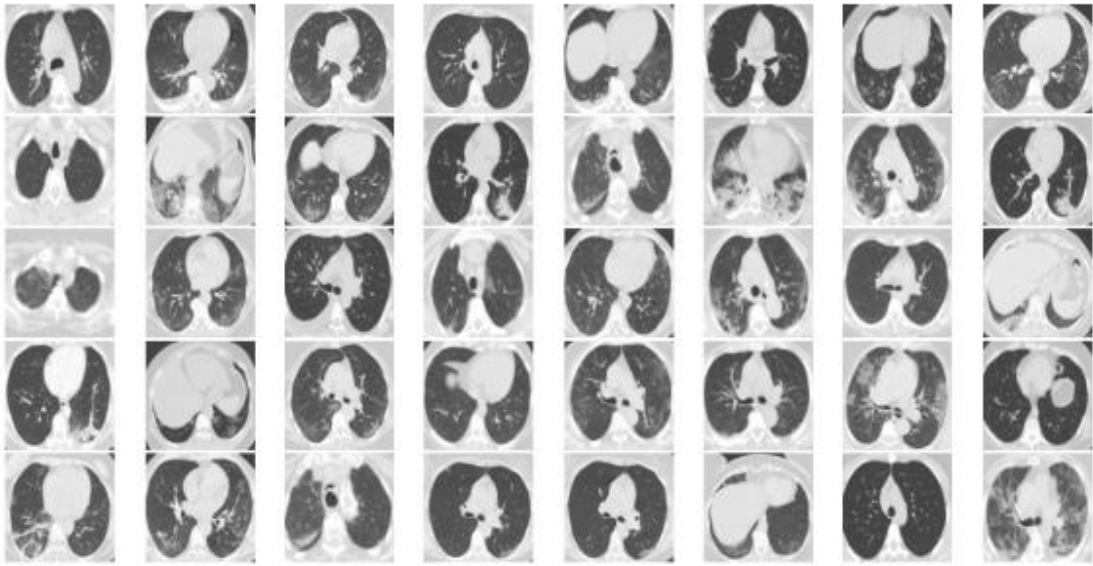
4.1 Φόρτωση και Προεπεξεργασία Δεδομένων (Data Loading and Preprocessing)

Το πρώτο βήμα αφορά την προεπεξεργασία των συνόλων δεδομένων. Παρατηρείται ότι οι εικόνες παρουσιάζουν μεγάλο μέγεθος γεγονός που καθιστά αδύνατη τη φόρτωση τους στο νευρωνικό δίκτυο. Για να μπορέσουν να δοθούν ως είσοδος, μετασχηματίζονται και λαμβάνουν μέγεθος 224×224. Επιπλέον, τα

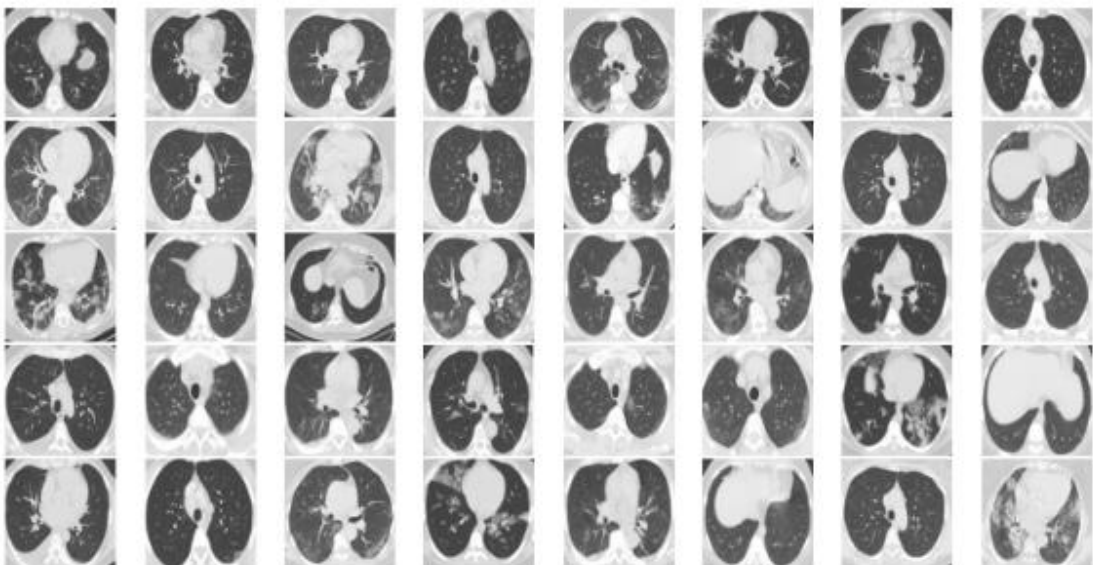
εικονοστοιχεία (pixels) των εικόνων κυμαίνονταν στο διάστημα $[0, 255]$. Ήταν απαραίτητο, λοιπόν, κάθε εικονοστοιχείο της εικόνας να κανονικοποιηθεί στο διάστημα $[0, 1]$.

Σε δεύτερο στάδιο, οι Covid και οι Non-Covid εικόνες χωρίζονται σε σύνολο εκπαίδευσης (train set) και σε σύνολο δοκιμής (test set) με αναλογία 80:20 του αρχικού συνόλου δεδομένων, αντίστοιχα. Το σύνολο εκπαίδευσης των Covid εικόνων και των Non-Covid εικόνων ενώνεται για να σχηματίσει ένα κοινό σύνολο εκπαίδευσης. Η ίδια διαδικασία ακολουθείται και για το σύνολο ελέγχου. Οι εικόνες χρησιμοποιήθηκαν ως είσοδοι στο συνελκτικό νευρωνικό δίκτυο DenseNet201 με διάσταση $224 \times 224 \times 3$, όπου ο αριθμός τρία δηλώνει τα κανάλια του χρωματικού χώρου RGB. Ένα μικρό δείγμα εικόνων που ανήκουν στο σύνολο δεδομένων SARS-CoV-2 και έχουν χαρακτηριστεί ως Covid φαίνεται στην Εικόνα 11.

X_train



X_test



Εικόνα 11 - Δείγμα του συνόλου εκπαίδευσης και του συνόλου δοκιμής

Τέλος, ελέγχεται εάν οι κλάσεις που εμπεριέχονται στο σύνολο δεδομένων εκπαίδευσης και δοκιμής είναι ισορροπημένες. Η διαδικασία αυτή πραγματοποιείται μετρώντας το πλήθος των παρατηρήσεων ανά κλάση.

4.2 Καθορισμός του Νευρωνικού Δικτύου

Αρχικά, εφαρμόζεται η τεχνική της μεταφοράς μάθησης, σύμφωνα με την οποία το προεκπαιδευμένο δίκτυο DenseNet201 φορτώνεται μαζί με τα βάρη του τα οποία

έχουν προκύψει από την ταξινόμηση του συνόλου δεδομένων ImageNet. Έπειτα, αφαιρείται το τελευταίο πλήρως συνδεδεμένο στρώμα που δίνει την έξοδο του νευρωνικού, δηλαδή την πρόγνωση. Ακολουθεί η προσθήκη πέντε νέων επιπέδων ώστε η εκπαίδευση να γίνει με βάση το παρόν σύνολο δεδομένων και την εξής σειρά:

1. **Καθολική Υποδειγματοληψία Μέσου Όρου (Global Average Pooling):** η οποία μειώνει το μέγεθος των δεδομένων, το πλήθος των παραμέτρων, συμβάλλοντας στην ελάττωση του υπολογιστικού κόστους.
2. **Στρώμα Ισοπέδωσης (Flatten Layer):** το οποίο αναδομεί τα δεδομένα ξεδιπλώνοντας τους χάρτες γνωρισμάτων εξόδου ενός συνελκτικού δικτύου και μετατρέποντας τα σε ένα μονοδιάστατο πίνακα.
3. **Πλήρως Συνδεδεμένο Στρώμα (Fully Connected Layer):** το οποίο περιέχει 128 νευρώνες και με συνάρτηση ενεργοποίησης, την Ανορθωμένη Γραμμική Μονάδα - ReLU (Rectified Linear Unit).
4. **Τεχνική Dropout:** η οποία επιτυγχάνει τυχαία ανάθεση, σε κάθε επανάληψη του αλγόριθμου εκπαίδευσης. Το 20% των ενεργοποιήσεων (Activations) του προηγούμενου επιπέδου γίνεται ίσο με 0. Η τεχνική αυτή χρησιμοποιείται προκειμένου να αποφευχθεί η υπερπροσαρμογή του δικτύου στα δεδομένα εκπαίδευσης [40].
5. **Πλήρως Συνδεδεμένο Στρώμα (Fully Connected Layer):** το οποίο δομείται από 64 νευρώνες και με συνάρτηση ενεργοποίησης τη 'ReLU'.
6. **Τεχνική Dropout:** το 30% των ενεργοποιήσεων του προηγούμενου επιπέδου έγινε ίσο με 0.
7. **Πλήρως Συνδεδεμένο Στρώμα (Dense):** το οποίο αποτελεί το στρώμα εξόδου του συστήματος. Ως συνάρτηση ενεργοποίησης επιλέγεται η σιγμοειδής συνάρτηση (sigmoid) η οποία ακολουθεί κατανομή μεταξύ διαστήματος $[0,1]$.

Στη συνέχεια, ορίζονται οι πληροφορίες αναφορικά με τον «Αλγόριθμο Προσαρμογής» (Fitting Algorithm). Ως «Συνάρτηση Σφάλματος» (loss function) που ελαχιστοποιείται επιλέγεται η κατηγορηματική διασταυρωμένη εντροπία (Categorical

Cross-entropy) [54], ενώ ως βελτιστοποιητής ο «Αλγόριθμος Προσαρμοστικής Εκτίμησης Βαρών» (Adaptive Moment Estimation - Adam) [55]. Επιπλέον, δίνεται η δυνατότητα ορισμού μίας λίστας από επιπλέον μετρικές οι οποίες θα υπολογίζονται κατά τη διάρκεια εκπαίδευσης. Στη συγκεκριμένη περίπτωση, ως επιπλέον μετρική επιλέγεται η ολική ορθότητα.

4.3 Εμπλουτισμός Δεδομένων / Data Augmentation

Σε προβλήματα με λίγα δεδομένα, όπως στις ιατρικές απεικονίσεις, είναι πολύ εύκολο το μοντέλο να μάθει να ταξινομεί καλύτερα από όσο απαιτείται ζημιώνοντας την ικανότητα γενίκευσής του σε άγνωστα δεδομένα. Το πρόβλημα αυτό καλείται υπερπροσαρμογή, όπως αναφέρθηκε και παραπάνω, και μία τεχνική αντιμετώπισης του ο εμπλουτισμός του συνόλου εκπαίδευσης με δεδομένα που έχουν υποστεί ελαφριές παραμορφώσεις, έχοντας ως σκοπό το δίκτυο να μη βλέπει τις ίδιες εικόνες σε πολλές εποχές (epochs) [56]. Στο Keras αυτό επιτυγχάνεται με τη χρήση της «ImageDataGenerator» η οποία δέχεται μία ομάδα παρατηρήσεων (batch), εφαρμόζει τυχαίες γεωμετρικές παραμορφώσεις και επιστρέφει μία νέα ομάδα παρατηρήσεων για να εκπαιδευτεί το δίκτυο. Είναι σημαντικό να τονιστεί ότι με τον εμπλουτισμό δεδομένων δεν αυξάνεται το μέγεθος του ήδη υπάρχοντος συνόλου δεδομένων.

4.4 Εκπαίδευση του Νευρωνικού Δικτύου

Με το πέρας των παραπάνω βημάτων, έγινε η έναρξη της εκπαίδευσης του νευρωνικού δικτύου. Η εκπαίδευση του δικτύου, πραγματοποιείται με τη χρήση της μεθόδου «fit» του μοντέλου. Ως πρώτη παράμετρος της μεθόδου αυτής, εισάγονται τα δεδομένα εκπαίδευσης, όπου η μεταβλητή «X_train» αναφέρεται στα δεδομένα που μπορούν να χρησιμοποιηθούν για την πρόβλεψη, ενώ η μεταβλητή «y_train» αναπαριστά την πραγματική τιμή εξόδου για κάθε είσοδο.

Οι παράμετροι «epochs» και «batch size» παίζουν σημαντικό ρόλο στην εκπαίδευση του δικτύου. Ειδικότερα, η παράμετρος batch size αποτελεί το μέγεθος κάθε δέσμης δειγμάτων και δηλώνει ότι σε κάθε επανάληψη, ο αλγόριθμος βελτιστοποίησης επιλέγει τυχαία μια ομάδα δειγμάτων πλήθους ίσου με την τιμή της παραμέτρου, για τον υπολογισμό της βαθμίδας (gradient) [56]. Σε αυτή την περίπτωση το μέγεθος της δέσμης δειγμάτων ορίζεται στην τιμή 32. Από την άλλη πλευρά, η παράμετρος «epochs» ορίζεται στην τιμή 500 και στην ουσία δηλώνει το

πλήθος των εποχών, όπου μία εποχή αντιστοιχεί σε μία επανάληψη της επεξεργασίας όλων των δειγμάτων εκπαίδευσης από το νευρωνικό δίκτυο [56].

Η παράμετρος «validation_data», αφορά την εισαγωγή δεδομένων επαλήθευσης, δηλαδή δειγμάτων, που δεν χρησιμοποιούνται στην εκπαίδευση του δικτύου, αλλά επιτρέπουν την παρακολούθηση της προόδου του με το πέρασμα των εποχών. Η μέθοδος «fit» επιστρέφει τελικά ένα αντικείμενο στο οποίο έχουν αποθηκευτεί οι τιμές της συνάρτησης σφάλματος και οι τιμές των επιπλέον μετρικών για τα δεδομένα εκπαίδευσης και δοκιμής για κάθε εποχή. Στην συγκεκριμένη υλοποίηση, ως επιπλέον μετρική έχει καθοριστεί η ολική ορθότητα. Οι αποθηκευμένες αυτές τιμές φέρουν τη δομή ενός λεξικού (dictionary) με όνομα history και κλειδιά (keys) τα αντίστοιχα ονόματα των μετρικών.

Τέλος, το εκπαιδευμένο νευρωνικό δίκτυο αποθηκεύεται σε ένα αρχείο με κατάληξη .h5 ώστε να υπάρχει δυνατότητα επαναχρησιμοποίησης του. Η συνάρτηση «save_weights» που παρέχεται από το μοντέλο δίνει τη δυνατότητα αποθήκευσης μόνο των βαρών του νευρωνικού σε ένα αρχείο με κατάληξη .hdf5.

4.5 Κατασκευή συστήματος ανάκτησης εικόνων βάσει περιεχομένου (Content Based Image Retrieval)

Η σύγκριση της ομοιότητας των εικόνων πραγματοποιήθηκε με τον κλασσικό τρόπο υπολογισμού της απόστασης των διανυσμάτων των χαρακτηριστικών και με δύο ακόμη παραλλαγές αυτού. Η μετρική απόστασης που επιλέχθηκε ήταν η L_1 νόρμα. Αρχικά, εφαρμόστηκε το κλασσικό μοντέλο CBIR. Τα διανύσματα χαρακτηριστικών προέρχονται από το τελευταίο στρώμα του νευρωνικού. Σε αυτή την απλή περίπτωση, όπως και στις δύο άλλες, δίνεται μία εικόνα επερώτησης (query image) και ανακτώνται οι τέσσερις πιο όμοιες με αυτήν. Οι δοκιμές που έγιναν έδειξαν ότι ορισμένες φορές η χρήση μίας Covid εικόνας ως επερώτηση επέστρεφε μία λίστα με Covid και Non-Covid εικόνες, κάτι που φάνηκε να ισχύει και στην περίπτωση των Non-Covid εικόνων. Αυτή ήταν και η αδυναμία που παρουσιάζει η συγκεκριμένη τεχνική η οποία φαίνεται να αποτυγχάνει αρκετές φορές στην ορθή αναζήτηση.

Προκειμένου, να περιοριστεί η πιθανότητα αστοχίας της μεθόδου και να παραχθεί μία βελτιωμένη λίστα ανακτημένων εικόνων-προτείνεται η αξιοποίηση της στάθμισης

της σημασίας των περιοχών κάθε εικόνας, ως προς τον χαρακτηρισμό Covid ή Non-Covid, με βάση τη μέθοδο Grad-CAM.

Στην πρώτη παραλλαγή που προτείνεται χρησιμοποιούνται τα βάρη a_k^c του Grad-CAM ως συντελεστής βαρύτητας για την επιλογή των πιο όμοιων εικόνων. Στόχος είναι τα χαρακτηριστικά που προέρχονται από το Grad-CAM να καθοδηγούν την ανάκτηση. Με αυτόν τον τρόπο το κριτήριο ομοιότητας διαφοροποιεί τη βαρύτητα που αποδίδει σε διάφορες περιοχές της εικόνας με βάση τη σημασία τους.

Σε μια δεύτερη παραλλαγή που προτείνεται, ακολουθείται μια υβριδική προσέγγιση που συνδυάζει το κλασσικό μοντέλο και την πρώτη παραλλαγή. Συγκεκριμένα, χρησιμοποιούνται οι αποστάσεις που προκύπτουν από τα δύο CBIR και εξάγεται ένας γραμμικός συνδυασμός με κάποιο συντελεστή βαρύτητας (υπερπαραμέτρος) της μορφής $d_1 + \lambda \times d_2$, όπου το λ έλαβε τιμές στο εύρος 0-1. Ύστερα από διάφορες δοκιμές, ως καλύτερη τιμή για το λ αναδείχθηκε το 0.5.

Τέλος, πρέπει να σημειωθεί ότι και στις τρεις περιπτώσεις ύστερα από την ανάκτηση εικόνων εφαρμόστηκε η μέθοδος Grad-CAM ώστε να διαπιστωθεί εάν ο χρωματισμός των σημαντικών περιοχών της εικόνας ήταν πιο εύστοχος και συνδεδεμένος με πιθανές γνωστές αλλοιώσεις που προκαλεί ο κορονοϊός. Τα αποτελέσματα θα αναλυθούν στην Ενότητα 5.

4.6 Οπτικοποίηση με τη μέθοδο Grad-CAM

Εκτός από μέθοδος στάθμισης, η μέθοδος Grad-CAM εφαρμόζεται κατά την υλοποίηση της εργασίας και ως μέσο οπτικοποίησης, που είναι και ο κατεξοχήν στόχος της μεθόδου όταν προτείνεται. Πιο συγκεκριμένα, χρησιμοποιείται για τον χρωματισμό των σημείων της εικόνας βάσει χαρτών ενεργοποίησης. Η κλίμακα των χρωμάτων που επιλέχθηκε περιλαμβάνει το κόκκινο, το κίτρινο, το πράσινο και το μπλε χρώμα. Το κόκκινο χρώμα υποδηλώνει τις πιο σημαντικές περιοχές της εικόνας, ενώ το μπλε χρώμα επισημαίνει τις λιγότερο σημαντικές. Υπάρχει, λοιπόν, μία κλίμακα χρωμάτων που αντιστοιχεί στην κλίμακα σημασίας. Αυτή η πληροφορία επιδιώκεται να ενσωματωθεί στη μέθοδο ανάκτησης.

Στην περίπτωση των Covid εικόνων ελέγχεται, επίσης, η πιθανότητα οι επισημασμένες περιοχές να υποδηλώνουν γνωστές αλλοιώσεις του πνεύμονα. Οι πιο συχνές βλάβες που εντοπίζονται μέσω της αξονικής τομογραφίας θώρακα είναι οι

διηθήσεις δίκην θολής υάλου ή/και οζώδεις διηθήσεις στην περιφέρεια των πνευμόνων (Ground-glass opacities - GGO) καθώς και η πάχυνση του μεσολοβιακού διαφράγματος (interlobular septal thickening) [57][58]. Αυτά τα ευρήματα είναι συνήθως αμφίπλευρα και πολυλοβιακά, ως επί το πλείστον κατανεμημένα στις υπουπεζωκοτικές/περιφερικές και οπίσθιες περιοχές των πνευμόνων, αλλά περιστασιακά μπορεί να δείχνουν βρογχοαγγειακή κατανομή ή συνδυασμό και των δύο [58].

5. Αποτελέσματα και Δοκιμές

Η υλοποίηση της εργασίας πραγματοποιήθηκε σε γλώσσα Python με χρήση του Keras, το οποίο αποτελεί μία ανοιχτού κώδικα Διεπαφή Προγραμματισμού Εφαρμογών (Application Programming Interface - API) υψηλού επιπέδου, και της βιβλιοθήκης Tensorflow ως υπολογιστικό backend [59]. Προκειμένου, να ικανοποιηθούν οι απαιτητικοί υπολογισμοί του προβλήματος το νευρωνικό δίκτυο εκπαιδεύτηκε με χρήση GPU. Επιλέχθηκαν δύο σύνολα δεδομένων τα οποία αποτελούν τα μεγαλύτερα δημόσια διαδεδομένα.

Το σύνολο δεδομένων SARS-CoV-2 αποτελείται από 2482 αξονικές τομογραφίες προερχόμενες από 120 ασθενείς, εκ των οποίων οι 1252 αξονικές τομογραφίες συλλέχθηκαν από 60 ασθενείς που μολύνθηκαν από τον ιό (άνδρες: 32, γυναίκες: 28) και 1230 αξονικές τομογραφίες θώρακα 60 μη μολυσμένων ασθενών (άνδρες: 30, γυναίκες: 30). Τα δεδομένα συλλέχθηκαν από νοσοκομεία του Σάο Πάολο της Βραζιλίας. Οι εικόνες δεν έχουν κάποιο συγκεκριμένο πρότυπο ως προς το μέγεθος τους (οι διαστάσεις της μικρότερης εικόνας είναι 104×153, ενώ οι μεγαλύτερες εικόνες είναι 484×416). Η βάση δεδομένων είναι ελεύθερα διαθέσιμη στον εξής σύνδεσμο <https://www.kaggle.com/plameneduardo/sarscov2-ctscan-dataset> [60].

Το σύνολο δεδομένων COVID-CT συλλέχθηκε από 216 ασθενείς και είναι ελεύθερα διαδεδομένο στον εξής σύνδεσμο <https://github.com/UCSD-AI4H/COVID-CT>. Πιο συγκεκριμένα, το σύνολο δεδομένων αποτελείται από 812 αξονικές τομογραφίες θώρακα εκ των οποίων οι 349 αναφέρονται σε θετικά διαγνωσμένους ασθενείς με κορονοϊό και οι υπόλοιπες 463 σε αξονικές τομογραφίες θώρακα αρνητικά διαγνωσμένων ασθενών. Οι εικόνες των παθόντων από κορονοϊό πάρθηκαν από επιστημονικά άρθρα που κατατέθηκαν στα αποθετήρια medRxiv και bioRxiv, από τις 19 Ιανουαρίου έως τις 25 Μαρτίου καθώς και από δωρεές που έγιναν σε νοσοκομεία (<http://medicalsegmentation.com/covid19/>). Τα

μεταδεδομένα εξήχθησαν με μη αυτόματο τρόπο και περιείχαν τις εξής πληροφορίες: ηλικία ασθενούς, φύλο, τοποθεσία, ιατρικό ιστορικό, χρόνος σάρωσης, σοβαρότητα της μόλυνσης και ιατρική αναφορά. Το συγκεκριμένο σύνολο δεδομένων έχει επικυρωθεί ως προς την ορθότητα του από έναν ραδιολόγο ο οποίος ασχολείται με τη διάγνωση και τη θεραπεία του COVID-19 από τα πρώιμα στάδια της πανδημίας. Συγκριτικά με το προηγούμενο σύνολο δεδομένων, έχουν οριστεί πρότυπα για το μέγεθος και την αντίθεση των εικόνων. Είναι σημαντικό να τονιστεί ότι ορισμένες εικόνες περιέχουν πληροφορίες κειμένου που μπορεί να επηρεάσουν την πρόβλεψη του μοντέλου [61].

5.1 Μετρικές Αξιολόγησης

5.1.1 Αξιολόγηση Νευρωνικού Δικτύου

Η αξιολόγηση της συμπεριφοράς και της απόδοσης του μοντέλων που κατασκευάστηκαν, καθώς και η εξακρίβωση ότι δεν υπάρχει υπερπροσαρμογή πραγματοποιήθηκε με τις κλασσικές μετρικές αξιολόγησης. Οι μετρικές αυτές αναφέρονται στην ολική ορθότητα (6), στην καμπύλη ROC και στον Πίνακα Συνάφειας.

Ολική ορθότητα (Accuracy)

Η ολική ορθότητα ορίζεται ως το ποσοστό του συνολικού αριθμού προβλέψεων που είναι ορθές [62].

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (6)$$

όπου:

TP (True positive): αληθώς θετικών αποτελεσμάτων

TN (True negative): αληθώς αρνητικών αποτελεσμάτων

FP (False positive): ψευδώς θετικών αποτελεσμάτων

FN (False negative): ψευδώς αρνητικών αποτελεσμάτων

Η τιμή της ολικής ορθότητας, δηλώνει ότι το νευρωνικό δίκτυο πρόβλεψε σωστά ένα ποσοστό εικόνων του συνόλου δεδομένων δοκιμής. Στο παρόν πείραμα η ολική ορθότητα έφτασε στην καλύτερη περίπτωση το ποσοστό του 96%. Η τιμή αυτή

μπορεί να θεωρηθεί αξιόπιστη καθώς το σύνολο δεδομένων δοκιμής είναι ισορροπημένο.

Μέτρο F

Ωστόσο, η ολική ορθότητα δεν εξετάζει τον τρόπο διανομής των δεδομένων. Το μέτρο F (F-measure), λοιπόν, χρησιμοποιείται για να χειριστεί το πρόβλημα διανομής με ακρίβεια. Είναι χρήσιμο όταν το σύνολο δεδομένων περιέχει κλάσεις ανισορροπίας. Προκειμένου, να γίνει αντιληπτή η μαθηματική έκφραση υπολογισμού του μέτρου F πρέπει να δοθεί ο ορισμός των επιμέρους εννοιών [63].

Ακρίβεια (Precision): Η ακρίβεια χρησιμοποιείται για την αξιολόγηση της ακρίβειας των ταξινομητών. Η χαμηλή τιμή του υποδεικνύει ότι ο ταξινομητής πάσχει από μεγάλο αριθμό ψευδώς θετικών αποτελεσμάτων. Η μαθηματική διατύπωση του δίνεται παρακάτω:

$$Precision = \frac{TP}{TP+FP} \quad (7)$$

Ανάκληση (Recall): Η ανάκληση που ονομάζεται επίσης και ευαισθησία χρησιμοποιείται ως μέτρο για τον υπολογισμό της πληρότητας ενός ταξινομητή. Η χαμηλότερη τιμή της υποδεικνύει ότι ο ταξινομητής υποφέρει από μεγάλες τιμές ψευδώς θετικών αποτελεσμάτων.

$$Recall = \frac{TP}{TP+FN} \quad (8)$$

Επομένως, το μέτρο F δίνεται από τον παρακάτω τύπο:

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (9)$$

Πίνακας Συνάφειας (Confusion Matrix)

Η κατανόηση των προαναφερθέντων εννοιών μπορεί να γίνει με τη συμβολή του πίνακα συνάφειας ο οποίος περιέχει πληροφορίες σχετικά με τα πραγματικά και προβλέψιμα αποτελέσματα, δοσμένα από έναν ταξινομητή.

Πίνακας 1 - Πίνακας συνάφειας

	Ορθή πρόβλεψη (Predicted positive)	Λανθασμένη πρόβλεψη (Predicted negative)
Πραγματικά ορθή (Actual positive)	TP	FP
Πραγματικά λανθασμένη (Actual negative)	FN	TN

Καμπύλη ROC

Το γράφημα λειτουργικών χαρακτηριστικών δείκτη ή καμπύλη ROC αποτελεί μια τεχνική οπτικοποίησης ταξινομητών που είναι βασισμένη στην απόδοσή τους. Πιο συγκεκριμένα, χρησιμοποιείται συνήθως σε συστήματα λήψης ιατρικών αποφάσεων και εξαρτάται από δύο παράγοντες. Στην ουσία, αποτελεί μία γραφική παράσταση δύο διαστάσεων όπου στους άξονες του έχει το ποσοστό των αληθώς θετικών αποτελεσμάτων (TPR) που αποτελεί την ευαισθησία (sensitivity) και το ποσοστό των ψευδών θετικών αποτελεσμάτων (FPR), που αποτελεί την ειδικότητα (specificity). Κάθε TPR και FPR δημιουργούν στην ουσία ένα σημείο της καμπύλης το οποίο αντιστοιχεί σε ένα συγκεκριμένο κατώφλι (cutoff), παράγοντας έτσι το τελικό σχήμα. Όσο πιο κοντά βρίσκεται η γραφική παράσταση της καμπύλης στα άνω αριστερά όρια (ευαισθησία = 1, ειδικότητα = 1) τόσο πιο ακριβής θεωρείται η διάγνωση. Αντιθέτως, όταν βρίσκεται πιο κοντά στη διαγώνιο τόσο λιγότερο ακριβής είναι. Το εμβαδόν της περιοχής κάτω από την καμπύλη ονομάζεται **AUC (Area Under the Curve)** και αποτελεί ένα μέτρο διαχωρισμού των δύο κατηγοριών που συμμετέχουν στην ταξινόμηση σε όλα τα πιθανά κατώφλια. Το εύρος στο οποίο κυμαίνεται η AUC είναι από 0-1. Όσο η τιμή της AUC πλησιάζει το 1 τόσο πιο αξιόπιστο είναι το μοντέλο, ενώ όταν πλησιάζει στο 0.5 σημαίνει ότι το μοντέλο παρουσιάζει δυσκολία ως προς τη διάκριση των θετικών και των αρνητικών αποτελεσμάτων και ότι πιθανότατα ο διαχωρισμός ήταν τυχαίος [64].

5.1.2 Αξιολόγηση του συστήματος ανάκτησης βάσει περιεχομένου

Δύο από τα πιο δημοφιλή μέτρα για την αξιολόγηση της απόδοσης ενός συστήματος ανάκτησης εικόνων βάσει περιεχομένου είναι ο μέσος όρος ακρίβειας

και ο κοντινότερος γείτονας. Η πρώτη μετρική προϋποθέτει τον υπολογισμό της ακρίβειας και της ανάκλησης. Ωστόσο, στην περίπτωση της ανάκτησης πληροφοριών οι δύο αυτές μετρικές παρουσιάζουν διαφορετική σημασία και υπολογίζονται με διαφορετικό τρόπο από αυτόν που δόθηκε στην αξιολόγηση της ταξινόμησης [65].

Ακρίβεια (Precision)

Η ακρίβεια ορίζεται ως το σύνολο των σχετικά με το ερώτημα ανακτημένων εικόνων προς το σύνολο των ανακτημένων εικόνων.

$$Precision = \frac{\text{Number of relevant images retrieved}}{\text{Number of retrieved images}} \quad (10)$$

Ανάκληση (Recall)

Η ανάκληση είναι το σύνολο των σχετικών εικόνων με το ερώτημα που ανακτώνται με επιτυχία προς το σύνολο των σχετικών εικόνων.

$$Recall = \frac{\text{Number of relevant images retrieved}}{\text{Number of relevant images}} \quad (11)$$

Μέσος Όρος Ακρίβειας (Mean Average Precision – MAP)

Ο μέσος όρος ακρίβειας αποτελεί το άθροισμα των μέσων ακριβειών (Average Precision – AP) διαιρούμενο με το σύνολο των εικόνων επερώτησης [67]. Δίνεται μία εικόνα επερώτησης και παράγεται μία λίστα με τις πιο συσχετισμένες εικόνες βάσει της μετρικής της απόστασης. Το μέγεθος της λίστας επιλέγεται από τον κάθε χρήστη ανάλογα με τις απαιτήσεις του προβλήματος. Υπολογίζοντας την ακρίβεια και την ανάκληση για κάθε στοιχείο της λίστας μπορεί να προκύψει η συνάρτηση $p(r)$ και να βρεθεί η μέση τιμή της $p(r)$ στο διάστημα $r = [0,1]$ η οποία χαρακτηρίζεται ως μέση ακρίβεια και δίνεται από τον εξής τύπο [66]:

$$AP = \frac{\sum_{k=1}^n (P(k) \times r(k))}{\text{Number of relevant images}} \quad (12)$$

Ο τύπος που δίνει τελικά το μέσο όρο ακριβειών είναι:

$$MAP = \frac{\sum_{q=1}^Q AveP(k)}{Q} \quad (13)$$

όπου Q είναι το σύνολο των εικόνων επερώτησης και $AveP(k)$ η μέση ακρίβεια για μία εικόνα επερώτησης q .

Κοντινότερος Γείτονας (Nearest Neighbor)

Ο κοντινότερος γείτονας είναι το ποσοστό των πλησιέστερων αντιστοιχιών που έχει βρεθεί ότι ανήκουν στην ίδια κλάση με την εικόνα επερώτησης. Η μετρική αυτή αποτελεί μία ένδειξη για το πόσο καλά θα απέδιδε ένας ταξινομητής πλησιέστερου γείτονα. Για κάθε εικόνα επερώτησης εντοπίζεται ο πρώτος πλησιέστερος γείτονας ($k=1$) και ελέγχεται εάν η κλάση στην οποία ανήκει ταυτίζεται με αυτή της εικόνας επερώτησης [68].

$$Nearest\ neighbor = \frac{\sum nearest\ neighbors}{Number\ of\ total\ images} \quad (14)$$

5.2 Πειράματα και δοκιμές

Στο στάδιο της ταξινόμησης, πραγματοποιήθηκαν αρκετά διαφορετικά πειράματα εκ των οποίων τα πιο σημαντικά φαίνονται στον παρακάτω πίνακα. Συγκεκριμένα, χρησιμοποιήθηκε διαφορετικό σύνολο εποχών εύρους 100-500 σε κάθε περίπτωση. Επιπλέον, δοκιμάστηκαν διάφορες αρχιτεκτονικές όσον αφορά τα επιπλέον στρώματα που προστέθηκαν στα προεκπαιδευμένα δίκτυα. Σε αυτό το σημείο είναι σημαντικό να αναφερθεί ότι στόχος ήταν η χρήση λίγων επιπλέον στρωμάτων ώστε να επιτευχθεί ένα αρκετά καλό ποσοστό ολικής ορθότητας και ταυτόχρονα να περιοριστεί η πολυπλοκότητα του δικτύου. Σε όλες τις περιπτώσεις εφαρμόστηκε Data Augmentation καθώς έχει αποδειχθεί ότι επηρεάζει θετικά το αποτέλεσμα της ταξινόμησης. Δεν πραγματοποιήθηκε κάποια αλλαγή στο batch size το οποίο παρέμεινε στην τιμή 32. Μία υποσημείωση που πρέπει να γίνει εδώ είναι ότι ανάμεσα στην περίπτωση (6) και (8) υπάρχει μόνο ένα επιπλέον Dropout στρώμα, η προσθήκη του οποίου φαίνεται να αυξάνει το ποσοστό ολικής ορθότητας κατά 1%. Τα τρία μοντέλα ResNet50, VGG19, DenseNet201 αξιολογήθηκαν βάσει της ολικής ορθότητας ταξινόμησης και το σφάλμα εκπαίδευσης.

Πίνακας 2 - Πειράματα και αποτελέσματα στο στάδιο της ταξινόμησης

Μεθοδολογία	Σύνολο Δεδομένων	Ολική ορθότητα ταξινόμησης	Σφάλμα εκπαίδευσης
1. Pretrained ResNet50+Flatten layer, Dropout layer (0.5), Dense layer (softmax)	COVID-CT dataset	73%	1.3853

2. Pretrained ResNet50+Flatten layer, Dropout layer (0.5), Dense layer (softmax)	SARS-CoV-2 dataset	73%	1.1140
3. Pretrained VGG19 +Flatten layer, Dropout layer (0.5), Dense layer (softmax)	COVID-CT dataset	77%	0.7437
4. Pretrained VGG19 +Flatten layer, Dropout layer (0.5), Dense layer (softmax)	SARS-CoV-2 dataset	93%	0.2849
5. Pretrained DenseNet201 +2D GlobalAveragePooling, Flatten layer, Dense layer (128, Relu), Dropout layer (0.3), Dense layer (64, Relu), Dense layer (softmax activation)	COVID-CT dataset	82%	0.6633
6. Pretrained DenseNet201 +2D GlobalAveragePooling, Flatten layer, Dense layer (128, Relu), Dropout layer (0.3), Dense layer (64, Relu), Dense layer (softmax activation)	SARS-CoV-2 dataset	95%	0.1307
7. Pretrained DenseNet201 +2D GlobalAveragePooling, Flatten layer, Dense layer (128, Relu), Dropout layer (0.2), Dense layer (64, Relu), Dropout layer	COVID-CT dataset	82%	0.6929

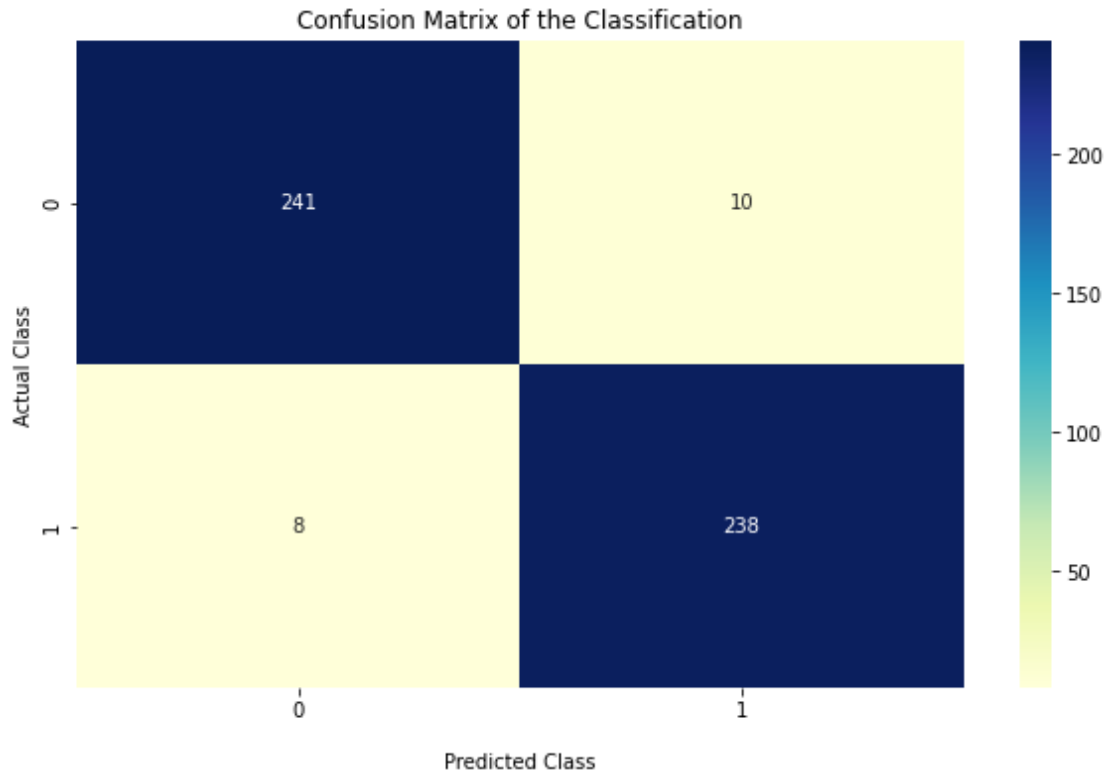
(0.3), Dense layer (sigmoid activation)/			
8. Pretrained DenseNet201 +2D GlobalAveragePooling, Flatten layer, Dense layer (128, Relu), Dropout layer (0.2), Dense layer (64, Relu), Dropout layer (0.3), Dense layer (sigmoid activation)	SARS-CoV-2 dataset	96%	0.0977

5.3 Αξιολόγηση της επιλεγθείσας αρχιτεκτονικής

Το μοντέλο που επιλέχθηκε τελικά για την κατηγοριοποίηση των συνόλων δεδομένων ήταν το DenseNet201 το οποίο χρησιμοποιήθηκε έπειτα για την κατασκευή του συστήματος CBIR και για την εφαρμογή του Grad-CAM. Πιο συγκεκριμένα, τα αποτελέσματα που παρουσιάζονται παρακάτω αφορούν μόνο την περίπτωση 8 του πίνακα 2 η οποία αναδείχθηκε ως η καλύτερη και ως προς την αρχιτεκτονική και ως προς το σύνολο δεδομένων.

5.3.1 Αξιολόγηση ταξινόμησης

Πίνακας συνάφειας

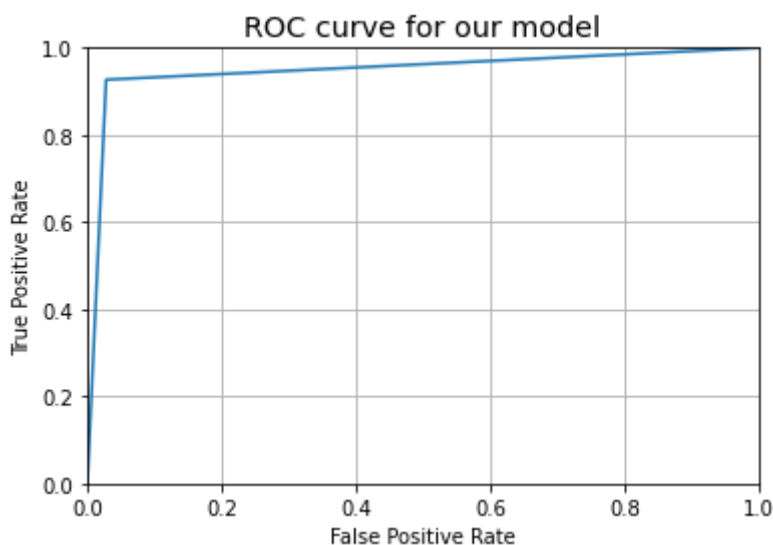


Εικόνα 12 - Πίνακας συνάφειας της ταξινόμησης

Πίνακας 3 - Πίνακας αποτελεσμάτων της ταξινόμησης

	Ακρίβεια	Ανάκληση	F-μέτρο	Σύνολο εικόνων
0	0.97	0.96	0.96	251
1	0.96	0.97	0.96	246
Ολική ορθότητα	-	-	0.96	497
Μέσος όρος μετρικών για κάθε κλάση (Macro avg)	0.96	0.96	0.96	497
Σταθμισμένος μέσος όρος (Weighted avg)	0.96	0.96	0.96	497

Καμπύλη ROC



Εικόνα 13 - Καμπύλη ROC της ταξινόμησης

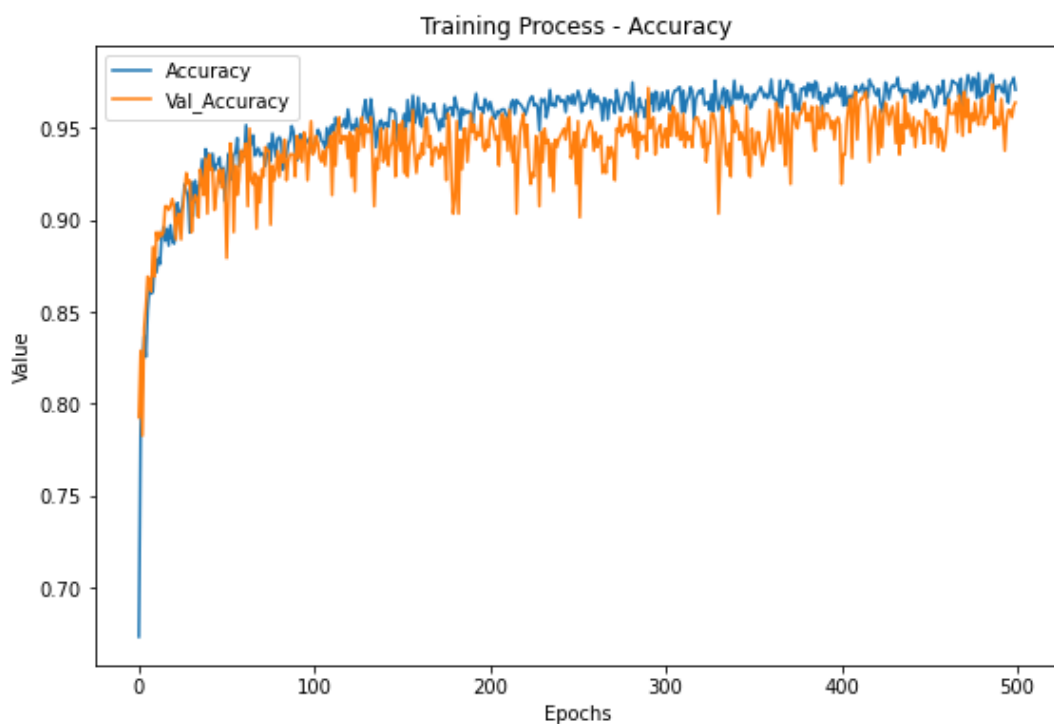
Η καμπύλη φαίνεται να τείνει προς την τιμή 1 γεγονός που καθιστά την ταξινόμηση ακριβή.

Γραφικές παραστάσεις

Στο παρόν πείραμα η ολική ορθότητα έφτασε το 96% και η τιμή αυτή είναι αξιόπιστη, καθώς οι κλάσεις στο σύνολο δεδομένων εκπαίδευσης και δοκιμής είναι ισορροπημένες. Γίνεται εύκολα αντιληπτό ότι η συνάρτηση σφάλματος ξεκινάει από μία τιμή και ελαττώνεται κατά τη διάρκεια της εκπαίδευσης, ενώ η ολική ορθότητα ξεκινάει από μία χαμηλή τιμή και με την πάροδο του χρόνου αυξάνεται. Επιπλέον, η καμπύλη της συνάρτησης σφάλματος τείνει να συγκλίνει με αυτή της συνάρτησης σφάλματος του συνόλου δοκιμής. Το ίδιο φαινόμενο παρατηρείται και στο γράφημα της ολικής ορθότητας. Επιπρόσθετα, η καμπύλη και στις δύο περιπτώσεις δεν παρουσιάζει στασιμότητα το οποίο υποδηλώνει ότι έχει αποφευχθεί το φαινόμενο της υπερπροσαρμογής. Τα αποτελέσματα των δύο γραφημάτων είναι πολύ σημαντικά για την αξιολόγηση της εκπαίδευσης η οποία φαίνεται να έχει πραγματοποιηθεί με επιτυχία σε αρκετά μεγάλο ποσοστό.



Εικόνα 14 - Γραφική παράσταση του σφάλματος εκπαίδευσης και δοκιμής



Εικόνα 15 - Γραφική παράσταση της ακρίβειας εκπαίδευσης και δοκιμής

5.3.2 Εφαρμογή του συστήματος CBIR

Το σύστημα ανάκτησης εικόνων εφαρμόστηκε για ένα υποσύνολο του συνόλου δεδομένων το οποίο περιλαμβάνει 251 αξονικές τομογραφίες θώρακα παθόντων και

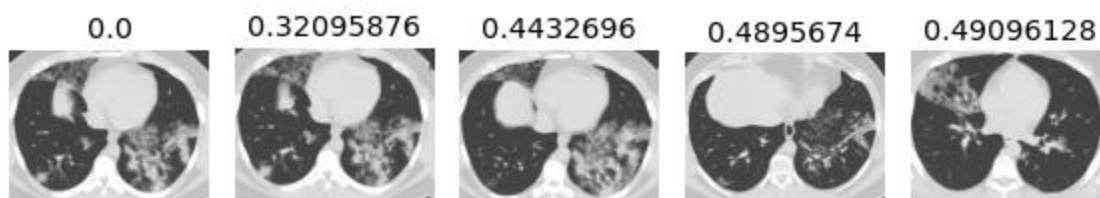
249 μη παθόντων. Εξ' αυτών παρουσιάζονται μόνο έξι παραδείγματα τα οποία επιτυγχάνουν για τις τέσσερις πρώτες ανακτώμενες εικόνες 100% ακρίβεια. Είναι απαραίτητο να δοθούν ορισμένες διευκρινίσεις που θα συνδράμουν στην κατανόηση των αποτελεσμάτων που ακολουθούν:

1. Ο αριθμός δίπλα από την ετικέτα Covid/Non-Covid υποδηλώνει τη θέση της εικόνας μέσα στο σύνολο δεδομένων.
2. Σε κάθε παράδειγμα παρουσιάζονται η λίστα των ανακτώμενων εικόνων και η οπτικοποίηση τους με τη μέθοδο Grad-CAM αφότου έχει υπολογιστεί η απόσταση των διανυσμάτων των χαρακτηριστικών με τον κλασσικό τρόπο και με τις δύο παραλλαγές του.
3. Η πρώτη εικόνα που εμφανίζεται σε κάθε περίπτωση είναι η εικόνα επερώτησης.
4. Πάνω από κάθε εικόνα εμφανίζεται και η απόσταση του διανύσματος χαρακτηριστικών της από αυτού της εικόνας επερώτησης.

1. Covid (31)

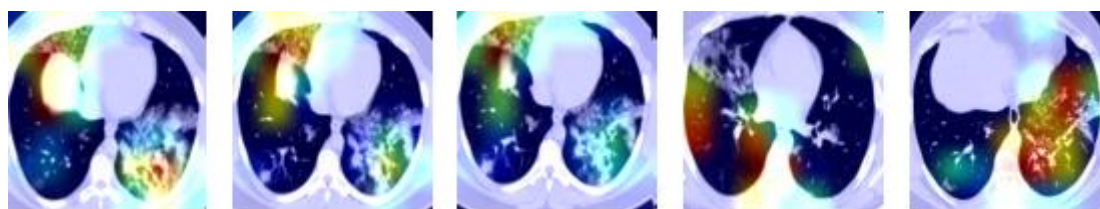
Η εικόνα επερώτησης για αυτό το παράδειγμα είναι η Covid (31).

Κλασσικό CBIR



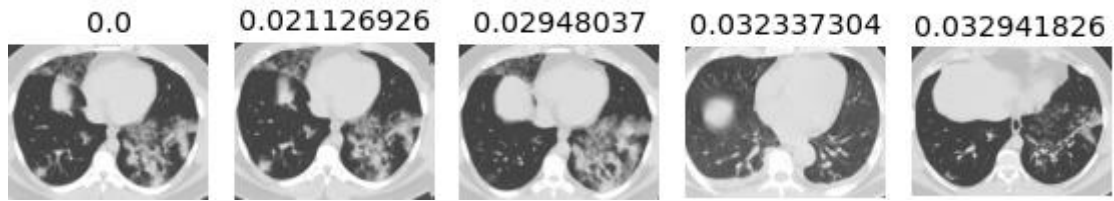
Εικόνα 16 - Απλή εφαρμογή CBIR με εικόνα επερώτησης την Covid(31)

Οπτικοποίηση των ανακτώμενων εικόνων με Grad-CAM



Εικόνα 17 - Εφαρμογή Grad-CAM στην απλή περίπτωση CBIR με εικόνα επερώτησης την Covid(31)

1^η παραλλαγή (CBIR / Grad-CAM)



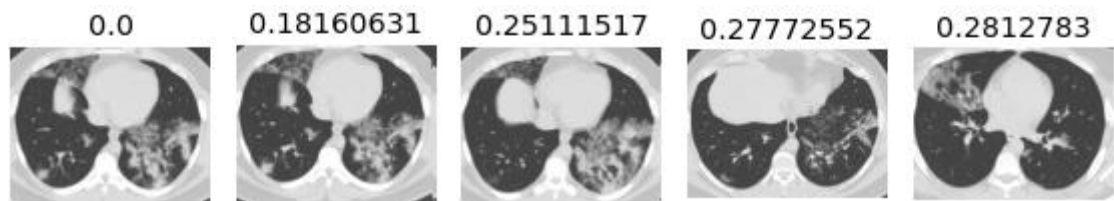
Εικόνα 18 - Παραλλαγή CBIR/Grad-CAM με εικόνα επερότησης την Covid(31)

Οπτικοποίηση των ανακτώμενων εικόνων με Grad-CAM



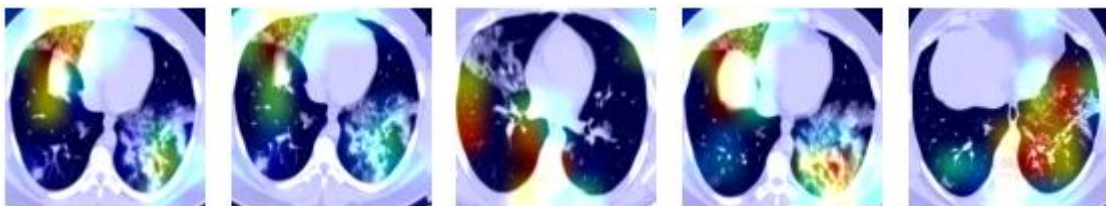
Εικόνα 19 - Εφαρμογή Grad-CAM στην παραλλαγή CBIR/Grad-CAM με εικόνα επερότησης την Covid(31)

2^η παραλλαγή (CBIR with hyperparameter)



Εικόνα 20 - Παραλλαγή CBIR/υπερparameter με εικόνα επερότησης την Covid(31)

Οπτικοποίηση των ανακτώμενων εικόνων με Grad-CAM

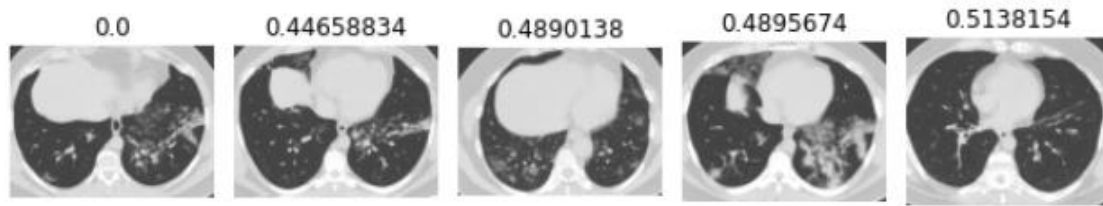


Εικόνα 21 - Εφαρμογή Grad-CAM στην παραλλαγή CBIR/υπερparameter με εικόνα επερότησης την Covid(31)

2. Covid (65)

Η εικόνα επερώτησης σε αυτό το παράδειγμα είναι η Covid (65).

Κλασσικό CBIR



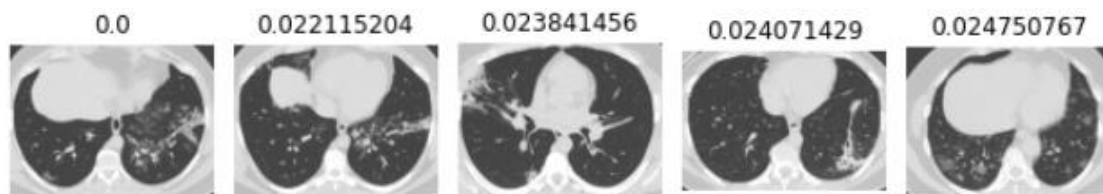
Εικόνα 22 - Απλή εφαρμογή CBIR με εικόνα επερώτησης την Covid(65)

Οπτικοποίηση των ανακτώμενων εικόνων με Grad-CAM



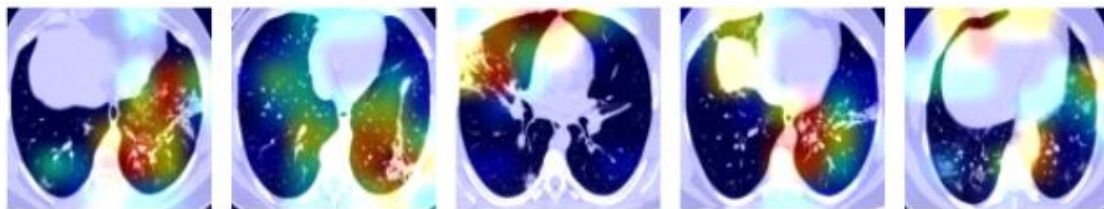
Εικόνα 23 - Εφαρμογή Grad-CAM στην απλή περίπτωση CBIR με εικόνα επερώτησης την Covid(65)

1^η παραλλαγή (CBIR/Grad-CAM)



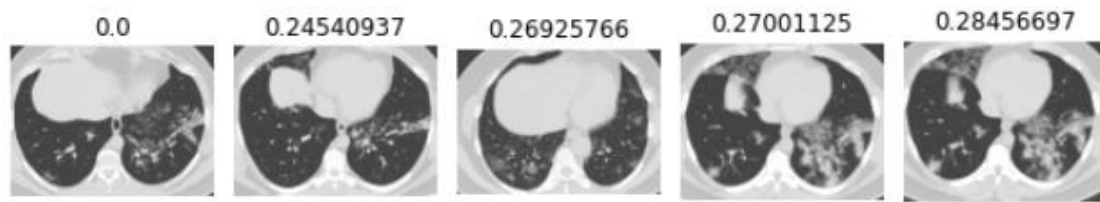
Εικόνα 24 - Παραλλαγή CBIR/Grad-CAM με εικόνα επερώτησης την Covid(65)

Οπτικοποίηση των ανακτώμενων εικόνων με Grad-CAM



Εικόνα 25 - Εφαρμογή Grad-CAM στην παραλλαγή CBIR/Grad-CAM με εικόνα επερώτησης την Covid(65)

2^η παραλλαγή (CBIR with hyperparameter)



Εικόνα 26 - Παραλλαγή CBIR/hyperparameter με εικόνα επερώτησης την Covid(65)

Οπτικοποίηση των ανακτώμενων εικόνων με Grad-CAM

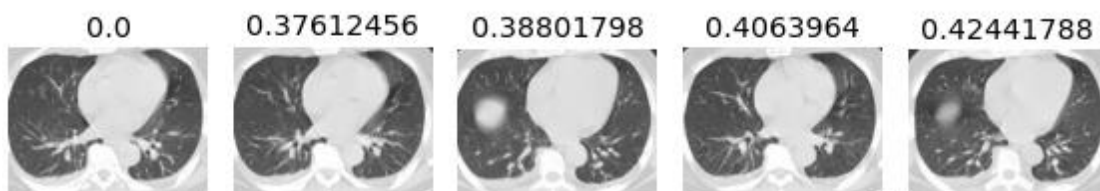


Εικόνα 27 - Εφαρμογή Grad-CAM στην παραλλαγή CBIR/hyperparameter με εικόνα επερώτησης την Covid(65)

3. Covid (1168)

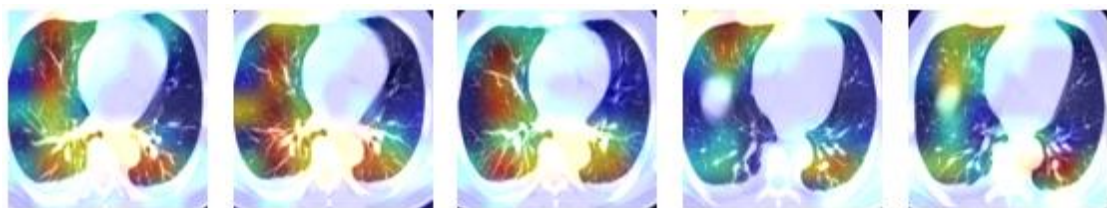
Η εικόνα επερώτησης σε αυτό το παράδειγμα είναι η Covid (1168)

Κλασικό CBIR



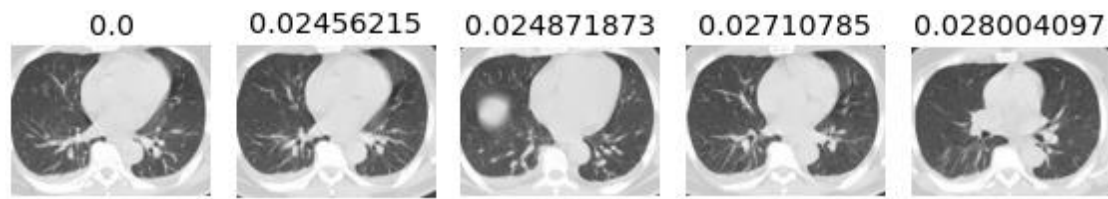
Εικόνα 28 - Απλή εφαρμογή CBIR με εικόνα επερώτησης την Covid(1168)

Οπτικοποίηση των ανακτώμενων εικόνων με τη μέθοδο Grad-CAM



Εικόνα 29 - Εφαρμογή Grad-CAM στην απλή περίπτωση CBIR με εικόνα επερώτησης την Covid(1168)

1^η παραλλαγή (CBIR/Grad-CAM)



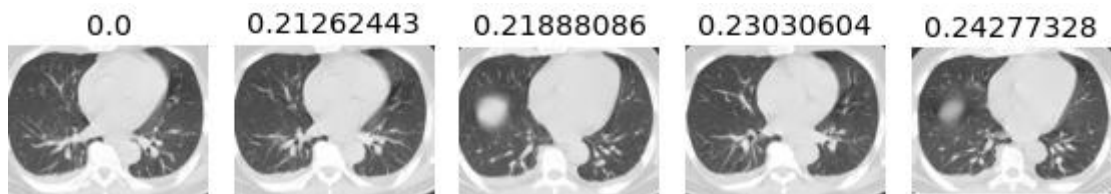
Εικόνα 30 - Παραλλαγή CBIR/Grad-CAM με εικόνα επερώτησης την Covid(1168)

Οπτικοποίηση των ανακτώμενων εικόνων με τη μέθοδο Grad-CAM



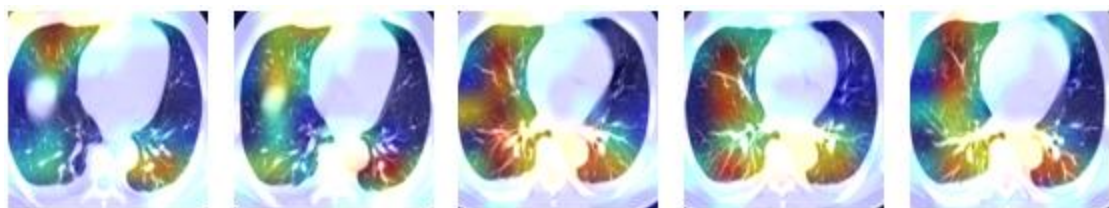
Εικόνα 31 - Εφαρμογή Grad-CAM στην παραλλαγή CBIR/Grad-CAM με εικόνα επερώτησης την Covid(1168)

2^η παραλλαγή (CBIR with hyperparameter)



Εικόνα 32 - Παραλλαγή CBIR/hyperparameter με εικόνα επερώτησης την Covid(1168)

Οπτικοποίηση των ανακτώμενων εικόνων με τη μέθοδο Grad-CAM

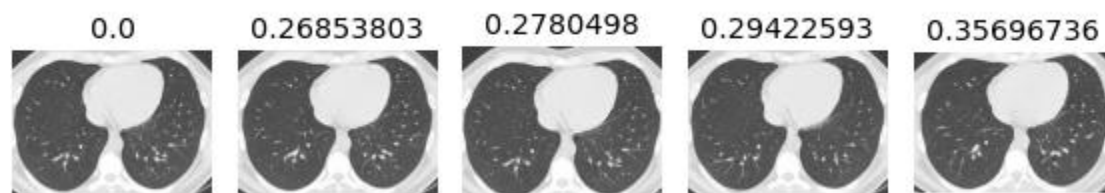


Εικόνα 33 - Εφαρμογή Grad-CAM στην παραλλαγή CBIR/hyperparameter με εικόνα επερώτησης την Covid(1168)

1. Non-Covid (37)

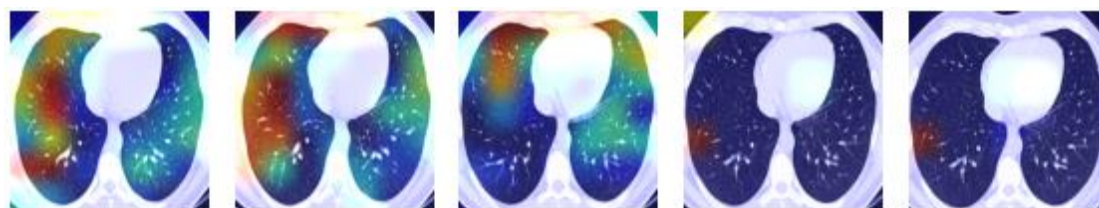
Η εικόνα επερώτησης που επιλέχθηκε είναι η Non-Covid (37).

Κλασικό CBIR



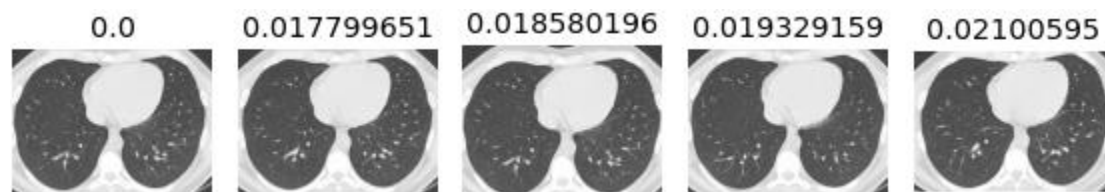
Εικόνα 34 - Απλή εφαρμογή CBIR με εικόνα επερώτησης την Non-Covid(37)

Οπτικοποίηση των ανακτώμενων εικόνων με τη μέθοδο Grad-CAM



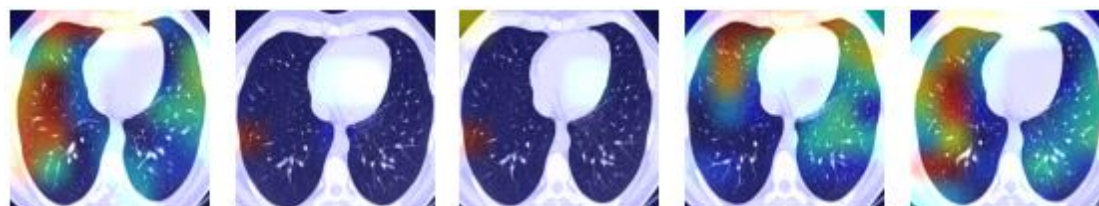
Εικόνα 35 - Εφαρμογή Grad-CAM στην απλή περίπτωση CBIR με εικόνα επερώτησης την Non-Covid(37)

1^η παραλλαγή (CBIR/Grad-CAM)



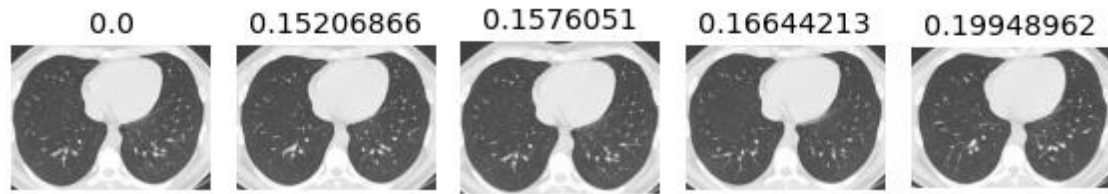
Εικόνα 36 - Παραλλαγή CBIR/Grad-CAM με εικόνα επερώτησης την Non-Covid(37)

Οπτικοποίηση των ανακτώμενων εικόνων με τη μέθοδο Grad-CAM



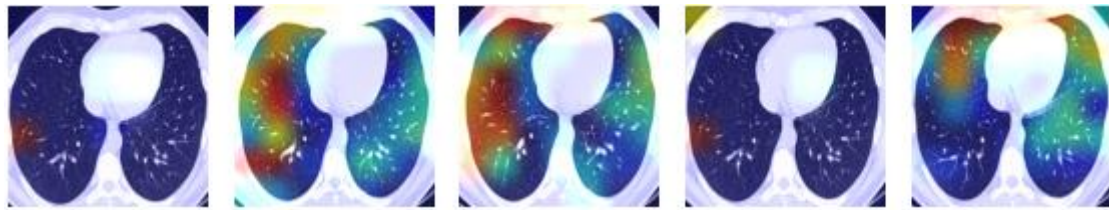
Εικόνα 37 - Εφαρμογή Grad-CAM στην παραλλαγή CBIR/Grad-CAM με εικόνα επερώτησης την Non-Covid(37)

2^η παραλλαγή (CBIR with hyperparameter)



Εικόνα 38 - Παραλλαγή CBIR/hyperparameter με εικόνα επερώτησης την Non-Covid(37)

Οπτικοποίηση των ανακτώμενων εικόνων με τη μέθοδο Grad-CAM



Εικόνα 39 - Εφαρμογή Grad-CAM στην παραλλαγή CBIR/hyperparameter με εικόνα επερώτησης την Non-Covid(37)

2. Non-Covid (409)

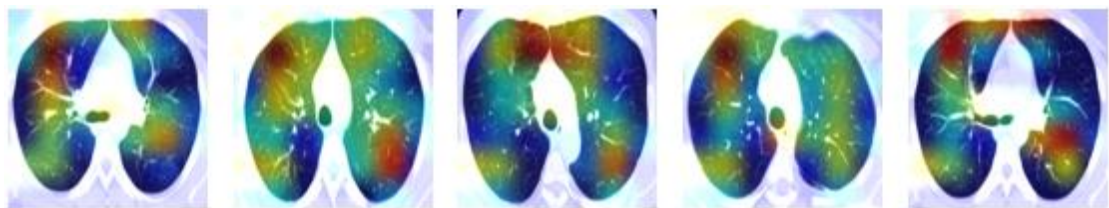
Η εικόνα επερώτησης αυτού του παραδείγματος είναι η Non-Covid (409).

Κλασσικό CBIR



Εικόνα 40 - Απλή εφαρμογή CBIR με εικόνα επερώτησης την Non-Covid(409)

Οπτικοποίηση των ανακτώμενων εικόνων με τη μέθοδο Grad-CAM



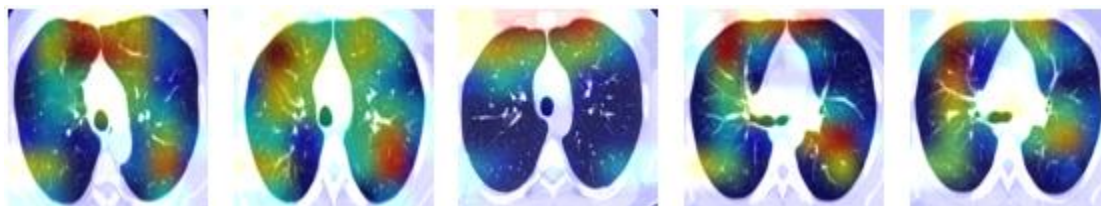
Εικόνα 41 - Εφαρμογή Grad-CAM στην απλή περίπτωση CBIR με εικόνα επερώτησης την Non-Covid(409)

1^η παραλλαγή (CBIR/Grad-CAM)



Εικόνα 42 - Παραλλαγή CBIR/Grad-CAM με εικόνα επερώτησης την Non-Covid(409)

Οπτικοποίηση των ανακτώμενων εικόνων με τη μέθοδο Grad-CAM



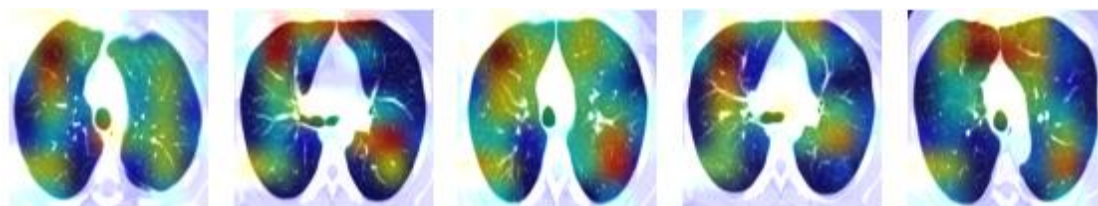
Εικόνα 43 - Εφαρμογή Grad-CAM στην παραλλαγή CBIR/Grad-CAM με εικόνα επερώτησης την Non-Covid(409)

2^η παραλλαγή (CBIR with hyperparameter)



Εικόνα 44 - Παραλλαγή CBIR/hyperparameter με εικόνα επερώτησης την Non-Covid(409)

Οπτικοποίηση των ανακτώμενων εικόνων με τη μέθοδο Grad-CAM

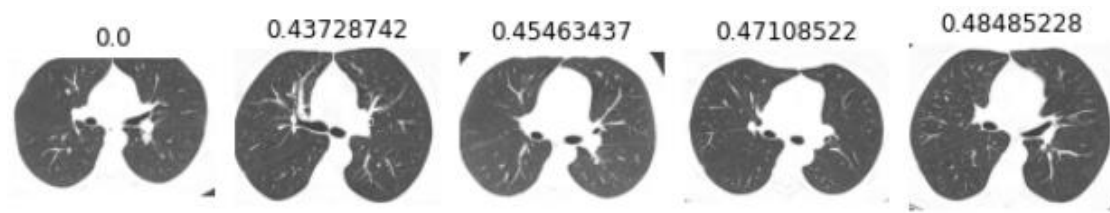


Εικόνα 45 - Εφαρμογή Grad-CAM στην παραλλαγή CBIR/hyperparameter με εικόνα επερώτησης την Non-Covid(409)

3. Non-Covid (1198)

Στο τελευταίο παράδειγμα επιλέχθηκε ως εικόνα επερώτησης η Non-Covid (1198).

Κλαστικό CBIR



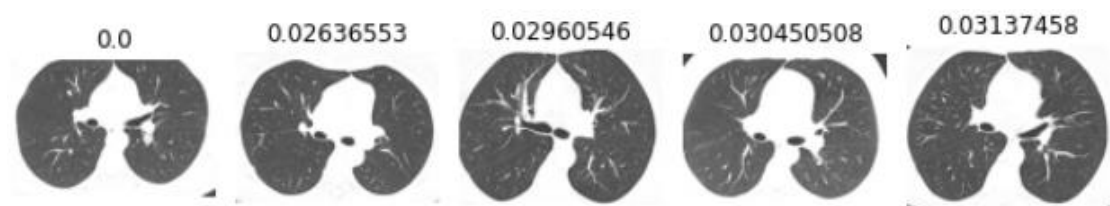
Εικόνα 46 - Απλή εφαρμογή CBIR με εικόνα επερώτησης την Non-Covid(1198)

Οπτικοποίηση των ανακτώμενων εικόνων με τη μέθοδο Grad-CAM



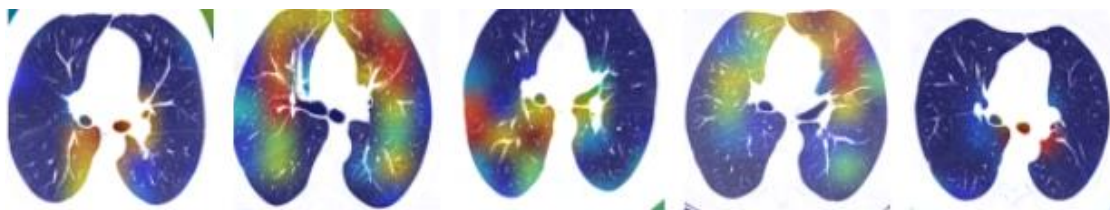
Εικόνα 47 - Εφαρμογή Grad-CAM στην απλή περίπτωση CBIR με εικόνα επερώτησης την Non-Covid(1198)

1^η παραλλαγή (CBIR/Grad-CAM)



Εικόνα 48 - Παραλλαγή CBIR/Grad-CAM με εικόνα επερώτησης την Non-Covid(1198)

Οπτικοποίηση των ανακτώμενων εικόνων με τη μέθοδο Grad-CAM



Εικόνα 49 - Εφαρμογή Grad-CAM στην παραλλαγή CBIR/Grad-CAM με εικόνα επερώτησης την Non-Covid(1198)

2^η παραλλαγή (CBIR with hyperparameter)



Εικόνα 50 - Παραλλαγή CBIR/hyperparameter με εικόνα επερώτησης την Non-Covid(1198)

Οπτικοποίηση των ανακτώμενων εικόνων με τη μέθοδο Grad-CAM



Εικόνα 51 - Εφαρμογή Grad-CAM στην παραλλαγή CBIR/hyperparameter με εικόνα επερώτησης την Non-Covid(1198)

Από τα παραπάνω αποτελέσματα προκύπτουν ορισμένες αξιόλογες παρατηρήσεις. Αρχικά, παρατηρείται ότι οι αποστάσεις που προκύπτουν από τη χρήση των δύο παραλλαγών είναι μικρότερες από αυτές του κλασσικού συστήματος ανάκτησης εικόνων. Ανάμεσα στην 1^η και στη 2^η παραλλαγή, η δεύτερη παράγει μικρότερες αποστάσεις από τη πρώτη. Τα πειράματα που έγιναν, έδειξαν, επίσης, ότι η 1^η παραλλαγή (CBIR/Grad-CAM) και η 2^η παραλλαγή (CBIR with hyperparameter) δίνουν καλύτερη βαθμονομημένη λίστα από το κλασσικό CBIR.

Επιπρόσθετα, στην περίπτωση των covid εικόνων φαίνεται ότι το Grad-CAM χρωματίζει με έντονο χρώμα περιοχές της εικόνας που εμφανίζουν λευκά στίγματα καθώς και αλλοιώσεις στην περιφέρεια των πνευμόνων. Αντίθετα, στις non-covid εικόνες ο χρωματισμός είναι πιο γενικός και όχι τόσο εύστοχος. Σε πολλές από τις περιπτώσεις non-covid οι εικόνες έχουν πολύ αμυδρό χρωματισμό.

5.3.4 Ειδικές περιπτώσεις χρήσης του συστήματος ανάκτησης εικόνων ανά επερώτηση

Τα παραδείγματα που παρουσιάστηκαν παραπάνω έχουν ποσοστό ακρίβειας 100%. Ωστόσο, υπάρχουν και περιπτώσεις όπου η ανάκτηση εικόνων αποτυγχάνει. Στον Πίνακα 3, λοιπόν, δίνονται ορισμένες εκ των περιπτώσεων που οι παραλλαγές

της μεθόδου CBIR επιτυγχάνουν καλύτερο ποσοστό ακρίβειας από την απλή περίπτωση.

Πίνακας 4 - Ειδικές περιπτώσεις ανάκτησης εικόνων βάσει περιεχομένου

Ακρίβεια				
Εικόνα επερώτησης	Κλασσικό CBIR	1 ^η παραλλαγή (CBIR/Grad-CAM)	2 ^η παραλλαγή (CBIR with hyperparameter)	Καλύτερη περίπτωση
Covid(75)	60%	100%	60%	1 ^η παραλλαγή (CBIR/Grad-CAM)
Covid(85)	60%	80%	60%	1 ^η παραλλαγή (CBIR/Grad-CAM)
Covid(89)	80%	100%	100%	1 ^η παραλλαγή (CBIR/Grad-CAM), 2 ^η παραλλαγή (CBIR with hyperparameter)
Covid(117)	80%	60%	80%	Κλασσικό CBIR, 2 ^η παραλλαγή (CBIR with hyperparameter)
Covid(235)	80%	100%	100%	1 ^η παραλλαγή (CBIR/Grad-CAM), 2 ^η παραλλαγή (CBIR with hyperparameter)
Covid(302)	60%	60%	60%	Κλασσικό CBIR, 1 ^η παραλλαγή (CBIR/Grad-CAM), 2 ^η παραλλαγή (CBIR with hyperparameter)
Covid(323)	80%	100%	100%	1 ^η παραλλαγή (CBIR/Grad-CAM), 2 ^η παραλλαγή (CBIR with hyperparameter)
Covid(341)	80%	100%	80%	1 ^η παραλλαγή (CBIR/Grad-CAM)

Covid(357)	80%	100%	80%	1 ^η παραλλαγή (CBIR/Grad-CAM)
Covid(395)	40%	60%	60%	1 ^η παραλλαγή (CBIR/Grad-CAM), 2 ^η παραλλαγή (CBIR with hyperparameter)
Covid(635)	40%	20%	40%	Κλασσικό CBIR, 2 ^η παραλλαγή (CBIR with hyperparameter)
Covid(645)	60%	40%	60%	Κλασσικό CBIR, 2 ^η παραλλαγή (CBIR with hyperparameter)
Covid(653)	60%	60%	80%	2 ^η παραλλαγή (CBIR with hyperparameter)
Covid(692)	60%	40%	40%	Κλασσικό CBIR
Covid(803)	80%	20%	80%	Κλασσικό CBIR, 2 ^η παραλλαγή (CBIR with hyperparameter)
Covid(935)	80%	100%	80%	1 ^η παραλλαγή (CBIR/Grad-CAM)
Covid(1177)	80%	100%	80%	1 ^η παραλλαγή (CBIR/Grad-CAM)
Non-Covid (7)	80%	80%	80%	Κλασσικό CBIR, 1 ^η παραλλαγή (CBIR/Grad-CAM), 2 ^η παραλλαγή (CBIR with hyperparameter)
Non-Covid (18)	80%	100%	100%	1 ^η παραλλαγή (CBIR/Grad-CAM), 2 ^η παραλλαγή (CBIR with hyperparameter)
Non-Covid (109)	40%	80%	40%	1 ^η παραλλαγή (CBIR/Grad-CAM)
Non-Covid (163)	80%	80%	100%	2 ^η παραλλαγή (CBIR with hyperparameter)

				hyperparameter)
Non-Covid (171)	80%	100%	100%	1 ^η παραλλαγή (CBIR/Grad-CAM), 2 ^η παραλλαγή (CBIR with hyperparameter)
Non-Covid (309)	60%	40%	60%	Κλασσικό CBIR, 2 ^η παραλλαγή (CBIR with hyperparameter)
Non-Covid (333)	80%	100%	80%	1 ^η παραλλαγή (CBIR/Grad-CAM)
Non-Covid (343)	80%	60%	80%	Κλασσικό CBIR, 2 ^η παραλλαγή (CBIR with hyperparameter)
Non-Covid (382)	60%	80%	60%	1 ^η παραλλαγή (CBIR/Grad-CAM)
Non-Covid (496)	80%	40%	60%	Κλασσικό CBIR
Non-Covid (430)	80%	60%	80%	Κλασσικό CBIR, 2 ^η παραλλαγή (CBIR with hyperparameter)
Non-Covid(502)	60%	80%	80%	1 ^η παραλλαγή (CBIR/Grad-CAM), 2 ^η παραλλαγή (CBIR with hyperparameter)
Non-Covid(534)	60%	80%	60%	1 ^η παραλλαγή (CBIR/Grad-CAM)
Non-Covid(603)	40%	60%	40%	1 ^η παραλλαγή (CBIR/Grad-CAM)
Non-Covid(834)	60%	100%	80%	1 ^η παραλλαγή (CBIR/Grad-CAM)
Non-Covid(909)	80%	40%	80%	Κλασσικό CBIR, 2 ^η παραλλαγή (CBIR with hyperparameter)

Επομένως, προκύπτει το συμπέρασμα ότι πιθανόν οι παραλλαγές υπολογισμού της απόστασης σε ένα τέτοιο σύστημα μπορούν να δώσουν καλύτερη βαθμονομημένη λίστα και να περιορίσουν εν μέρει τις αστοχίες. Επιπλέον, όπως φαίνεται και στον πίνακα 4, σε πολλές περιπτώσεις που αποτυγχάνει η 1^η παραλλαγή, η 2^η παραλλαγή και το κλασσικό CBIR διατηρούν το ποσοστό της ακρίβειας υψηλό. Ωστόσο, υπάρχουν περιπτώσεις όπου το κλασσικό CBIR και η 2^η παραλλαγή αποτυγχάνουν, ενώ η 1^η παραλλαγή διατηρεί το ποσοστό της ακρίβειας στο 100%. Το σίγουρο είναι πάντως ότι στις περιπτώσεις αποτυχίας του κλασσικού CBIR μία από τις δύο παραλλαγές θα επιτυγχάνει πολύ καλό αποτέλεσμα ως προς την ανάκτηση των εικόνων. Η τελευταία στήλη του πίνακα 4 συμβάλλει στην κατανόηση των προαναφερθέντων σχολιασμών.

Πέρα από τον υπολογισμό της ακρίβειας χρειάζεται να υπολογιστούν πιο εξειδικευμένες μετρικές αξιολόγησης: ο μέσος όρος ακριβειών και ο κοντινότερος γείτονας. Οι μετρικές αυτές χρησιμοποιούνται ευρέως σε συστήματα ανάκτησης πληροφοριών και ανίχνευσης αντικειμένων.

Πίνακας 5 - Αποτελέσματα συστήματος ανάκτησης βάσει περιεχομένου

Μέθοδος υπολογισμού της απόστασης	Μέσος όρος ακριβειών (MAP)	Κοντινότερος γείτονας (Nearest Neighbor)
Κλασσικό CBIR	0.96	0.91
1 ^η παραλλαγή (CBIR/Grad-CAM)	0.94	0.81
2 ^η παραλλαγή (CBIR with hyperparameter)	0.96	0.92

Στον πίνακα 5 φαίνεται ότι στην απλή περίπτωση και στην ανάκτηση εικόνων με τη χρήση υπερπαραμέτρου επιτυγχάνονται καλύτερα αποτελέσματα σε όλες τις μετρικές. Αξίζει να σημειωθεί ότι η χρήση υπερπαραμέτρου βελτιστοποιεί κατά 1% τη μετρική του κοντινότερου γείτονα συγκριτικά με τις άλλες δύο.

6. Συμπεράσματα και Μελλοντικές προσεγγίσεις

Συμπερασματικά, η έρευνα που πραγματοποιήθηκε αποδεικνύει τη σκοπιμότητα της χρήσης της βαθιάς μάθησης στην υποβοήθηση των γιατρών να ανιχνεύσουν ασθενείς με κορονοϊό και να εντοπίσουν συγκεκριμένες αλλοιώσεις στον πνεύμονα του ανθρώπου μέσω των αξονικών τομογραφιών θώρακα. Η πρώιμη και ακριβής πρόγνωση μπορεί να περιορίσει σε μεγάλο βαθμό τον αριθμό των κρουσμάτων, γεγονός που καθιστά επιτακτική την ανάγκη σχεδιασμού συστημάτων υποβοήθησης της ιατρικής διάγνωσης.

Σε αυτή την εργασία προτείνεται, ένα σύνθετο σύστημα ταξινόμησης και ανάκτησης αξονικών τομογραφιών θώρακα. Το σύστημα βασίζεται στο προεκπαιδευμένο συνελκτικό νευρωνικό δίκτυο DenseNet201 και αξιοποιεί τα διανύσματα των χαρακτηριστικών που εξάγονται από κάθε εικόνα για την εύρεση όμοιων με αυτήν μέσα σε μία μεγάλη βάση δεδομένων βάση της μεθόδου CBIR. Η σύγκριση των εικόνων πραγματοποιείται με τον υπολογισμό της απόστασης μεταξύ των χαρακτηριστικών και στη συνέχεια παράγεται μία βαθμονομημένη λίστα. Ο υπολογισμός της απόστασης έγινε με τρεις διαφορετικούς τρόπους εκ των οποίων ως βέλτιστος αναδείχθηκε ο συνδυασμός των αποστάσεων με τη μεσολάβηση μίας υπερπαραμέτρου λ με τιμή 0.5.

Τα αποτελέσματα του συστήματος αξιολογήθηκαν με μία ποικιλία μετρικών και οπτικοποιήθηκαν με τη χρήση της μεθόδου Grad-CAM. Πέρα από την παρουσίαση του αποτελέσματος, η μέθοδος αυτή συντέλεσε και στην κατασκευή του συστήματος ανάκτησης εικόνων, καθώς τα βάρη της χρησιμοποιήθηκαν ως συντελεστής βαρύτητας στον υπολογισμό της απόστασης.

Οι δοκιμές που πραγματοποιήθηκαν έδωσαν τα εξής συμπεράσματα:

1. Το σύστημα ταξινόμησης δοκιμάστηκε σε όλο το σύνολο δεδομένων και πέτυχε ένα αξιόλογο ποσοστό ολικής ορθότητας 96%.
2. Η σύγκριση του κλασσικού τρόπου υπολογισμού της απόστασης και των δύο παραλλαγών του έδειξε ότι σε περιπτώσεις αποτυχίας εύρεσης μίας ορθής λίστας ανακτώμενων εικόνων οι δύο προτεινόμενοι τρόποι επιτυγχάνουν καλύτερο ποσοστό ακρίβειας.
3. Το σύστημα ανάκτησης εικόνων ελέγχθηκε σε ένα υποσύνολο του συνόλου δεδομένων για το οποίο κατόρθωσε ποσοστό μέσου όρου ακριβειών 96% στην κλασσική περίπτωση και με τη χρήση υπερπαραμέτρου.
4. Μεγάλο ενδιαφέρον προξένησε το γεγονός ότι η κλασσική εκδοχή CBIR έφτασε το 91% στη μετρική του κοντινότερου γείτονα, ενώ η παραλλαγή με χρήση της υπερπαραμέτρου κατόρθωσε να φτάσει στο 92%.

5. Τα heatmaps που εξάγει το Grad-CAM αναδεικνύουν στις covid εικόνες λευκά στίγματα και αλλοιώσεις στην περιφέρεια των πνευμόνων τα οποία πιθανόν να αντιστοιχούν στα GGOs καθώς και στην πάχυνση του μεσολοβιακού διαφράγματος.

Το σύνθετο σύστημα που δημιουργήθηκε προσφέρεται για τη διάγνωση ασθενών με κορονοϊό και την εύρεση εικόνων με κοινά χαρακτηριστικά μέσα σε μία μεγάλη βάση δεδομένων. Παρόλο που, το σύστημα δίνει αρκετά καλά αποτελέσματα ως προς το κομμάτι της ταξινόμησης, υπάρχουν περιθώρια βελτιστοποίησης στη μέθοδο ανάκτησης. Μία μελλοντική έρευνα θα μπορούσε να επικεντρωθεί στην αξιοποίηση υπερπαραμέτρων με σκοπό την επίτευξη μικρότερων αποστάσεων και κατά επέκταση καλύτερης βαθμονομημένης λίστας εικόνων.

Βιβλιογραφία

- [1] C. Huang, Y. Wang, X. Li, L. Ren, J. Zhao, Y. Hu, L. Zhang, G. Fan, J. Xu, X. Gu, Z. Cheng, T. Yu, J. Xia, Y. Wei, W. Wu, X. Xie, W. Yin, H. Li, M. Liu, Y. Xiao, H. Gao, L. Guo, J. Xie, G. Wang, R. Jiang, Z. Gao, Q. Jin, J. Wang, B. Cao, "Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China," *Lancet*, vol. 395, no. 10223, pp. 497–506, 2020.
- [2] T. Singhal, "A review of Coronavirus disease-2019 (COVID-19)," *Indian J. Pediatr.*, vol. 87, no. 4, pp. 281–286, 2020.
- [3] D. S. Hui, E. I Azhar, T. A. Madani, F. Ntoumi, R. Kock, O. Dar, G. Ippolito, T. D. Mchugh, Z. A. Memish, C. Drosten, A. Zumla, E. Petersen, "The continuing 2019-nCoV epidemic threat of novel Coronaviruses to global health–The latest 2019 novel Coronavirus outbreak in Wuhan China", *Int. J. Infect. Dis.*, vol. 91, pp. 264-266, 2020.
- [4] J. J. Deeks, J. Dinnes, Y. Takwoingi, C. Davenport, R. Spijker, S. Taylor-Phillips, A. Adriano, S. Beese, J. Dretzke, L. Ferrante di Ruffano, I. M Harris, M. J Price, S. Ditttrich, D. Emperor, L. Hooft, M. Mg Leeftang, A. Van den Bruel, Cochrane COVID-19 Diagnostic Test Accuracy Group, "Antibody tests for identification of current and past infection with SARS-CoV-2," *Cochrane Database Syst. Rev.*, vol. 6, no. 6, p. CD013652, 2020.
- [5] K. He, X. Zhang, S. Ren and J. Sun, "Deep residual learning for image recognition", *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 770-778, 2016.
- [6] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-cam: Visual explanations from deep networks via gradient-based localization," *International Journal of Computer Vision*, vol. 128, no. 2, pp. 336–359, 2019.
- [7] K. Ramanjaneyulu, K. V. Swamy, and C. H. S. Rao, "Novel CBIR System using CNN Architecture," in *2018 3rd International Conference on Inventive Computation Technologies (ICICT)*, 2018, pp. 379–383.
- [8] C. Butt, J. Gill, D. Chun, and B. A. Babu, "Deep learning system to screen coronavirus disease 2019 pneumonia," *Applied Intelligence*, Apr 2020.
- [9] L. Hall, D. Goldgof, R. Paul, and G. M. Goldgof, "Finding COVID-19 from chest X-rays using deep learning on a small dataset," 2020.
- [10] X. He, X. Yang, S. Zhang, J. Zhao, Y. Zhang, Eric Xing, P. Xie, "Sample-efficient deep learning for COVID-19 diagnosis based on CT scans," *bioRxiv*, 2020.
- [11] S. Salman and X. Liu, "Overfitting mechanism and avoidance in deep neural networks," *arXiv [cs.LG]*, 2019.

- [12] C. Zheng, X. Deng, Q. Fu, Q. Zhou, J. Feng, H. Ma, W. Liu, X. Wang, "Deep learning-based detection for COVID-19 from chest CT using weak label," *bioRxiv*, 2020.
- [13] Z. Wang, Q. Liu, and Q. Dou, "Contrastive cross-site learning with redesigned net for COVID-19 CT classification," *IEEE J. Biomed. Health Inform.*, vol. 24, no. 10, pp. 2806–2813, 2020.
- [14] S. Walvekar and D. S. Shinde, "Detection of COVID-19 from CT images using resnet50," *SSRN Electron. J.*, 2020.
- [15] A. Seum, A. H. Raj, S. Sakib, and T. Hossain, "A comparative study of CNN transfer learning classification algorithms with segmentation for COVID-19 detection from CT scan images," in *2020 11th International Conference on Electrical and Computer Engineering (ICECE)*, 2020, pp. 234–237.
- [16] X. Gao, Y. Qian, and A. Gao, "COVID-VIT: Classification of COVID-19 from CT chest images based on vision transformer models," *arXiv [eess.IV]*, 2021.
- [17] M. N. Teli, "TeliNet: Classifying CT scan images for COVID-19 diagnosis," in *2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, 2021.
- [18] V. Shah, R. Keniya, A. Shridharani, M. Punjabi, J. Shah, and N. Mehendale, "Diagnosis of COVID-19 using CT scan images and deep learning techniques," *Emerg. Radiol.*, vol. 28, no. 3, pp. 497–505, 2021.
- [19] Y. Song, S. Zheng, L. Li, X. Zhang, X. Zhang, Z. Huang, J. Chen, R. Wang, H. Zhao, Y. Chong, J. Shen, Y. Zha, and Y. Yang, "Deep learning enables accurate diagnosis of novel Coronavirus (COVID-19) with CT images," *IEEE/ACM Trans. Comput. Biol. Bioinform.*, vol. 18, no. 6, pp. 2775–2780, 2021.
- [20] A. Jaiswal, N. Gianchandani, D. Singh, V. Kumar, and M. Kaur, "Classification of the COVID-19 infected patients using DenseNet201 based deep transfer learning," *J. Biomol. Struct. Dyn.*, vol. 39, no. 15, pp. 5682–5689, 2021.
- [21] N. Hasan, Y. Bao, A. Shawon, and Y. Huang, "DenseNet convolutional neural networks application for predicting COVID-19 using CT image," *SN Comput Sci*, vol. 2, no. 5, p. 389, 2021.
- [22] A. Halder and B. Datta, "COVID-19 detection from lung CT-scan images using transfer learning approach," *Mach. Learn.: Sci. Technol.*, vol. 2, no. 4, p. 045013, 2021.
- [23] N. K. Mishra, P. Singh, and S. D. Joshi, "Automated detection of COVID-19 from CT scan using convolutional neural network," *Biocybern. Biomed. Eng.*, vol. 41, no. 2, pp. 572–588, 2021.
- [24] J. F. Hernández Santa Cruz, "An ensemble approach for multi-stage transfer learning models for COVID-19 detection from chest CT scans," *Intell Based Med*, vol. 5, no. 100027, p. 100027, 2021.

- [25] I. Chouat, A. Echioui, R. Khemakhem, W. Zouch, M. Ghorbel, and A. B. Hamida, "COVID-19 detection in CT and CXR images using deep learning models," *Biogerontology*, vol. 23, no. 1, pp. 65–84, 2022.
- [26] A. Ashraf, A. Malik, and Z. Khan, "POSTER: Diagnosis of COVID-19 through transfer learning techniques on CT scans: A comparison of deep learning models," *arXiv [eess.IV]*, 2022.
- [27] S. Namani, L. S. Akkapeddi, and S. Bantu, "Performance analysis of VGG-19 deep learning model for COVID-19 detection," in *2022 9th International Conference on Computing for Sustainable Global Development (INDIACom)*, 2022, pp. 781–787.
- [28] A. Sassi, W. Ouarda, and C. B. Amar, "Deep content information retrieval for COVID-19 detection from chromatic CT scans," *Arab. J. Sci. Eng.*, pp. 1–11, 2022.
- [29] Z. Li, "COVID-19 classification with CT scan and advanced deep learning technologies," in *2022 3rd International Conference on Computer Vision, Image and Deep Learning & International Conference on Computer Engineering and Applications (CVIDL & ICCEA)*, 2022, pp. 458–462.
- [30] S. Sakib, N. Ahmed, A. J. Kabir, and H. Ahmed, "An overview of Convolutional Neural Network: Its architecture and applications," *Preprints*, 2019.
- [31] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, 2017.
- [32] Z. R. Yang and Z. Yang, "Comprehensive Biomedical Physics," *Comprehensive Biomedical Physics. Karolinska Institute*, 2014.
- [33] X. Yao, "Evolving Artificial Neural Networks," *Bham.ac.uk*. [Online]. Available: https://www.cs.bham.ac.uk/~xin/papers/published_iproc_sep99.pdf. [Accessed: 19-Aug-2022].
- [34] M. F. Abbod, J. W. F. Catto, D. A. Linkens, and F. C. Hamdy, "Application of artificial intelligence to the management of urological cancer," *J. Urol.*, vol. 178, no. 4 Pt 1, pp. 1150–1156, 2007.
- [35] B.-M. Hsu, "Braille recognition for reducing asymmetric communication between the blind and non-blind," *Symmetry (Basel)*, vol. 12, no. 7, p. 1069, 2020.]
- [36] L. Alzubaidi *et al.*, "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *J. Big Data*, vol. 8, no. 1, p. 53, 2021.
- [37] A. Ajit, K. Acharya, and A. Samanta, "A review of convolutional neural networks," in *2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE)*, 2020, pp. 1–5.

- [38] J. Wu, "Introduction to Convolutional Neural Networks," *Edu.cn*, 2017. [Online]. Available: <https://cs.nju.edu.cn/wujx/paper/CNN.pdf>. [Accessed: 04-Aug-2022].
- [39] H. Gholamalinezhad and H. Khosravi, "Pooling methods in Deep Neural Networks, a review," *arXiv [cs.CV]*, 2020.
- [40] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *Jmlr.org*. [Online]. Available: <https://jmlr.org/papers/volume15/srivastava14a/srivastava14a.pdf>. [Accessed: 19-Aug-2022].
- [41] S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating deep network training by reducing internal covariate shift," *arXiv [cs.LG]*, 2015.
- [42] C. Du, Y. Wang, C. Wang, C. Shi, and B. Xiao, "Selective feature connection mechanism: Concatenating multi-layer CNN features with a feature selector," *Pattern Recognit. Lett.*, vol. 129, pp. 108–114, 2020.
- [43] A. Shamsaldin, P. Fattah, T. Rashid, and N. Al-Salihi, "A study of the convolutional neural networks applications," *UKH j. sci. eng.*, vol. 3, no. 2, pp. 31–40, 2019.
- [44] S. Basodi, C. Ji, H. Zhang, and Y. Pan, "Gradient amplification: An efficient way to train deep neural networks," *Big Data Min. Anal.*, vol. 3, no. 3, pp. 196–207, 2020.
- [45] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.
- [46] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.
- [47] F. D. Adhinata, D. P. Rakhmadani, M. Wibowo, and A. Jayadi, "A deep learning using DenseNet201 to detect masked or non-masked face," *JUITA J. Inform.*, vol. 9, no. 1, p. 115, 2021.
- [48] M. Shaha and M. Pawar, "Transfer Learning for Image Classification," in *2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, 2018, pp. 656–660.
- [49] S. Maji and S. Bose, "CBIR using features derived by Deep Learning," *arXiv [cs.IR]*, 2020.
- [50] M. M. Elmogy, H. El-bakry, and M. Al-khawlani, "Text-based, content-based, and semantic-based image retrievals: A survey," 2015.
- [51] F. Malik and B. Baharudin, "Analysis of distance metrics in content-based image retrieval using statistical quantized histogram texture features in the DCT domain," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 25, no. 2, pp. 207–218, 2013.

- [52] S. Kulkarni and Manu, “Content based Image Retrieval: A Literature Review,” in *2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS)*, 2022, pp. 1580–1587.
- [53] S. Sergyan, “Color histogram features based image classification in content-based image retrieval systems,” in *2008 6th International Symposium on Applied Machine Intelligence and Informatics*, 2008, pp. 221–224.
- [54] R. Gómez, Understanding Categorical Cross-Entropy Loss Binary Cross-Entropy Loss Softmax Loss Logistic Loss Focal Loss and all those confusing names, 2018, [online] Available: https://gombu.github.io/2018/05/23/cross_entropy_loss/.
- [55] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *arXiv [cs.LG]*, 2014.
- [56] J. Brownlee, “What is the Difference Between a Batch and an Epoch in a Neural Network,” *Machine Learning Mastery*, 2018.
- [57] W.-C. Dai., H.-W. Zhang, J. Yu, H.-J. Xu, H. Chen, S.-P. Luo, H. Zhang, L.-H. Liang, X.-L. Wu, Y. Lei, F. Lin, “CT imaging and differential diagnosis of COVID-19,” *Can. Assoc. Radiol. J.*, vol. 71, no. 2, pp. 195–200, 2020.
- [58] M. E. Laino, A. Ammirabile, L. Lofino, D. J. Lundon, A. Chiti, M. Francone and V. Savevski, “Prognostic findings for ICU admission in patients with COVID-19 pneumonia: baseline and follow-up chest CT and the added value of artificial intelligence,” *Emerg. Radiol.*, vol. 29, no. 2, pp. 243–262, 2022.
- [59] N. Ketkar, “Introduction to Keras,” in *Deep Learning with Python*, Berkeley, CA: Apress, 2017, pp. 97–111.
- [60] E. Soares, P. Angelov, S. Biaso, M. H. Froes, and D. K. Abe, *Sars-cov-2 ct-scan dataset: a large dataset of real patients ct scans for sars-cov-2 identification medRxiv*. 2020.
- [61] J. Close, Y. Zhao, X. Zhang, and P. X. He, *ct-dataset: a ct scan dataset about covid-19 (2020)*.
- [62] C. Liu, M. White, and G. Newell, “Measuring and comparing the accuracy of species distribution models with presence-absence data,” *Ecography (Cop.)*, vol. 34, no. 2, pp. 232–243, 2011.
- [63] M. Sokolova, N. Japkowicz, and S. Szpakowicz, “Beyond accuracy, F-score and ROC: A family of discriminant measures for performance evaluation,” in *Lecture Notes in Computer Science*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2006, pp. 1015–1021.
- [64] T. Fawcett, “An introduction to ROC analysis,” *Pattern Recognit. Lett.*, vol. 27, no. 8, pp. 861–874, 2006.

- [65] B. Syam and Y. Rao, “An effective similarity measure via genetic algorithm for content based image retrieval with extensive features,” *Int. Arab J. Inf. Technol*, vol. 10, pp. 143–151, 2013.
- [66] M. Zhu, “Recall, precision and average precision,” *Archive.org*, 2004. [Online]. Available: https://web.archive.org/web/20110504130953/http://sas.uwaterloo.ca/stats_navigation/techreports/04WorkingPapers/2004-09.pdf. [Accessed: 24-Sep-2022].
- [67] R. Padilla, S. L. Netto, and E. A. B. da Silva, “A survey on performance metrics for object-detection algorithms,” in *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*, 2020, pp. 237–242.
- [68] P. Shilane, P. Min, M. Kazhdan, and T. Funkhouser, “The princeton shape benchmark,” in *Proceedings Shape Modeling Applications, 2004*, 2004.

