



**ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΑΤΤΙΚΗΣ  
ΣΧΟΛΗ ΜΗΧΑΝΙΚΩΝ  
ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ**

**Χρήση Τεχνικών Μηχανικής Μάθησης για την  
Ανάλυση Ιατρικών Εικόνων με Στόχο την  
Υποβοήθηση στη Διάγνωση της Λοίμωξης  
COVID-19**

**ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ**

**ΨΑΡΑΥΤΗΣ-ΣΟΥΡΑΝΗΣ ΣΥΜΕΩΝ**

**Επιβλέποντες**

**Χρήστος Τρούσσας, Επίκουρος Καθηγητής  
Αθανάσιος Βουλόδημος, Επίκουρος Καθηγητής**

**Αθήνα, Οκτώβριος 2023**



**UNIVERSITY OF WEST ATTICA**  
**SCHOOL OF ENGINEERING**  
**DEPARTMENT OF INFORMATICS AND COMPUTER ENGINEERING**

# **Use of Machine Learning Techniques to Analyze Medical Images with the Aim of Assisting in the Diagnosis of COVID-19 Infection**

**DIPLOMA THESIS**

**PSARAFTIS-SOURANIS SYMEON**

**Supervisors**

**Christos Troussas, Assistant Professor**

**Athanasiros Voulodimos, Assistant Professor**

**Athens, October 2023**

Εγκρίθηκε από την τριμελή επιτροπή:

Χρήστος Τρούσσας Επίκουρος Καθηγητής	Αθανάσιος Βουλόδημος Επίκουρος Καθηγητής	Παναγιώτα Τσελέντη ΕΔΙΠ

## **ΔΗΛΩΣΗ ΣΥΓΓΡΑΦΕΑ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ**

Ο κάτωθι υπογεγραμμένος **Ψαραύτης – Σουράνης Συμεών** του **Ιωάννη**, με αριθμό μητρώου **711141049** φοιτητής του Πανεπιστημίου Δυτικής Αττικής της Σχολής **Μηχανικών** του Τμήματος **Μηχανικών Πληροφορικής και Υπολογιστών**, δηλώνω υπεύθυνα ότι:

«Είμαι συγγραφέας αυτής της διπλωματικής εργασίας και ότι κάθε βοήθεια την οποία είχα για την προετοιμασία της είναι πλήρως αναγνωρισμένη και αναφέρεται στην εργασία. Επίσης, οι όποιες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών ή λέξεων, είτε ακριβώς είτε παραφρασμένες, αναφέρονται στο σύνολό τους, με πλήρη αναφορά στους συγγραφείς, τον εκδοτικό οίκο ή το περιοδικό, συμπεριλαμβανομένων και των πηγών που ενδεχομένως χρησιμοποιήθηκαν από το διαδίκτυο. Επίσης, βεβαιώνω ότι αυτή η εργασία έχει συγγραφεί από μένα αποκλειστικά και αποτελεί προϊόν πνευματικής ιδιοκτησίας τόσο δικής μου, όσο και του Ιδρύματος.

Παράβαση της ανωτέρω ακαδημαϊκής μου ευθύνης αποτελεί ουσιώδη λόγο για την ανάκληση του πτυχίου μου».

Ο Δηλών

Ψαραύτης-Σουράνης Συμεών

*Σ. Ψαραύτης*

## **Ευχαριστίες**

Θα ήθελα να ευχαριστήσω τους επιβλέποντες καθηγητές μου, κύριο Χρήστο Τρούσσα και κύριο Αθανάσιο Βουλόδημο, για την πολύτιμη βοήθεια που μου παρείχαν για την επιτυχή εκπόνηση της παρούσας διπλωματικής εργασίας.

Επίσης, θα ήθελα να ευχαριστήσω την οικογένειά μου για τη στήριξη και συμπαράσταση που μου προσέφερε καθ' όλη τη διάρκεια των σπουδών μου.

## **Περίληψη**

Η ραγδαία εξέλιξη της τεχνολογίας έχει ως αποτέλεσμα το επιστημονικό πεδίο που ονομάζεται επιστήμη των υπολογιστών να βρίσκει εφαρμογές σε ολοένα και περισσότερους τομείς. Ένας από αυτούς του τομείς αποτελεί και η επιστήμη της ιατρικής. Η πρόοδος που έχει σημειωθεί στην επιστήμη των υπολογιστών, ιδιαίτερα στο υποπεδίο που ονομάζεται μηχανική μάθηση, καθιστά δυνατή την υποστήριξη μιας πληθώρας ιατρικών καθηκόντων μέσω τεχνικών της. Παραδείγματος χάρη, οι τεχνικές μηχανικής μάθησης μπορούν να αξιοποιηθούν για την ανίχνευση πιθανής παρουσίας κάποιας ασθένειας ή όγκου σε ιατρικές εικόνες, όπως ακτινογραφίες (X-rays), αξονικές τομογραφίες (CT scans) και μαγνητικές τομογραφίες (MRI scans).

Λαμβάνοντας υπόψη τη δύσκολη περίοδο την οποία διανύουμε εξαιτίας της πανδημίας COVID-19, καθώς και τη σημασία της ακρίβειας στην ανίχνευση της νόσου, μπορούμε ξεκάθαρα να διαπιστώσουμε ότι η βοήθεια που μπορεί να προσφέρει η μηχανική μάθηση στην ιατρική επιστήμη, είναι πιο πολύτιμη από ποτέ.

Στην παρούσα διπλωματική εργασία, πραγματευόμαστε την ανάλυση ιατρικών εικόνων για την υποβοήθηση στη διάγνωση COVID-19 κάνοντας χρήση τεχνικών μηχανικής μάθησης. Πιο συγκεκριμένα, εξετάζουμε την αποτελεσματικότητα τεχνικών βαθειάς μάθησης, και ειδικότερα συνελικτικών νευρωνικών δικτύων στην ανίχνευση της νόσου COVID-19, βασιζόμενοι σε πραγματικά δεδομένα ιατρικών εικόνων που έχουν συλλεχθεί παγκοσμίως.

## **Λέξεις - Κλειδιά**

Όραση Υπολογιστών, Μηχανική Μάθηση, Βαθιά Μάθηση, Μεταφορά Μάθησης, Συνελικτικά Νευρωνικά Δίκτυα, COVID-19, Ταξινόμηση Εικόνων, Σημασιολογική Κατάτμηση Εικόνων, Ιατρική Απεικόνιση, Αξονική Τομογραφία

## **Abstract**

The rapid development of technology has resulted in the scientific field called computer science finding applications in more and more areas. One of these fields is medicine. The progress made in computer science, especially in the subfield called machine learning, has made it possible for a variety of medical tasks to be assisted by its various techniques. For example, machine learning techniques can be utilized for the detection of the potential presence of a disease or tumor on medical images or scans, such as X-rays, CT scans and MRI scans.

Considering the difficult period we are going through due to the COVID-19 pandemic, as well as the importance of accuracy in the detection of the disease, we can clearly see that the assistance that machine learning can provide to the medical science is more valuable than ever.

In this diploma thesis we are dealing with the analysis of medical images for COVID-19 diagnosis support using machine learning techniques. More precisely, in this thesis we are examining the effectiveness of deep learning techniques, and particularly convolutional neural networks in detecting the COVID-19 disease, based on real-world medical image data that have been collected worldwide.

## **Keywords**

Computer Vision, Machine Learning, Deep Learning, Transfer Learning, Convolutional Neural Networks, COVID-19, Image Classification, Image Segmentation, Medical Imaging, Computed Tomography

## **Περιεχόμενα**

1	Εισαγωγή .....	18
1.1	Η σημασία της χρήσης τεχνικών μηχανικής μάθησης στην διαδικασία ανίχνευσης παθήσεων μέσω ιατρικών εικόνων .....	19
1.2	Περιγραφή του αντικειμένου και του στόχου της διπλωματικής εργασίας.....	20
1.3	Δομή Διπλωματικής Εργασίας.....	21
2	Θεωρητικό Υπόβαθρο .....	23
2.1	Αξονική Τομογραφία (Computed Tomography) .....	23
2.1.1	Παραθυροποίηση Εικόνας Αξονικής Τομογραφίας (CT Image Windowing) .	25
2.2	Όραση Υπολογιστών (Computer Vision) .....	26
2.2.1	Ταξινόμηση Εικόνων (Image Classification) .....	27
2.2.2	Σημασιολογική Κατάτμηση Εικόνων (Semantic Image Segmentation).....	27
2.3	Μηχανική Μάθηση (Machine Learning).....	28
2.3.1	Επιβλεπόμενη Μάθηση (Supervised Learning) .....	28
2.3.2	Μη Επιβλεπόμενη Μάθηση (Unsupervised Learning).....	29
2.3.3	Ενισχυτική Μάθηση (Reinforcement Learning) .....	29
2.3.4	Ημι-επιβλεπόμενη Μάθηση (Semi-supervised Learning) .....	30
2.4	Διαδικασία Εκπαίδευσης Αλγορίθμων Μηχανικής Μάθησης .....	30
2.4.1	Υποπροσαρμογή (Underfitting) .....	33
2.4.2	Υπερπροσαρμογή (Overfitting) .....	33
2.5	Τεχνικές Διαχωρισμού Δεδομένων.....	33
2.5.1	Τεχνική Holdout.....	34
2.5.2	Τεχνική Διασταυρούμενης Επικύρωσης K-Πτυχών (K-Fold Cross Validation)	
	34	
2.6	Βαθιά Μάθηση (Deep Learning).....	35
2.6.1	Βιολογικά Νευρωνικά Δίκτυα (Biological Neural Networks) .....	36
2.6.2	Τεχνητός Νευρώνας .....	38
2.6.3	Τεχνητά Νευρωνικά Δίκτυα (Artificial Neural Networks).....	39
2.7	Συνάρτηση Ενεργοποίησης (Activation Function).....	41
2.7.1	Σιγμοειδής Συνάρτηση Ενεργοποίησης (Sigmoid Activation Function).....	41
2.7.2	Συνάρτηση Ενεργοποίησης Softmax (Softmax Activation Function).....	41
2.7.3	Ανορθωμένη Γραμμική Μονάδα (ReLU).....	42
2.8	Συνάρτηση Απώλειας/Κόστους (Loss/Cost Function).....	42
2.8.1	Συνάρτηση Cross Entropy .....	43
2.9	Αλγόριθμοι Βελτιστοποίησης (Optimization Algorithms).....	43

2.9.1	Αλγόριθμος Βελτιστοποίησης Κατάβασης Κλίσης (Gradient Descent Optimization Algorithm) .....	44
2.9.2	Αλγόριθμος Βελτιστοποίησης Adaptive Moment Estimation (Adam) .....	46
2.10	Αλγόριθμος Οπισθοδιάδοσης Backward Propagation (Backpropagation) Algorithm	
	46	
2.10.1	Συνελικτικά Νευρωνικά Δίκτυα (Convolutional Neural Networks) .....	47
2.10.2	Επίπεδο Συνέλιξης (Convolutional Layer) .....	48
2.10.3	Επίπεδο Συγκέντρωσης (Pooling Layer) .....	50
2.10.4	Πλήρως Συνδεδεμένο Επίπεδο (Fully Connected Layer).....	51
2.11	Τεχνικές Κανονικοποίησης (Regularization Techniques) .....	51
2.11.1	Προσαύξηση Δεδομένων (Data Augmentation) .....	51
2.11.2	Γεωμετρικοί Μετασχηματισμοί .....	52
2.11.3	Πρόωρος Τερματισμός (Early Stopping).....	53
2.11.4	Επίπεδο Dropout .....	53
2.11.5	Επίπεδο Batch Normalization .....	54
2.12	Μεταφορά Μάθησης (Transfer Learning) .....	55
3	Περιγραφή Μεθοδολογίας Πειράματος Ταξινόμησης Εικόνων.....	57
3.1	Επιλογή Συνόλου Δεδομένων Πειράματος Ταξινόμησης Εικόνων (Image Classification) .....	58
3.1.1	Σύνολο Δεδομένων COVID-CT-MD .....	59
3.1.2	Σύνολο Δεδομένων CC-CCII.....	61
3.2	Προεπεξεργασία Δεδομένων .....	63
3.2.1	Κανονικοποίηση Αντίθεσης (Contrast) και Φωτεινότητας (Brightness) Τομών Αξονικών Τομογραφιών .....	64
3.2.2	Δημιουργία παραλλαγής του συνόλου δεδομένων όπου διατηρείται μόνο το πνευμονικό παρέγχυμα στις τομές αξονικών τομογραφιών.....	65
3.2.3	Μετατροπή εικόνων από μορφότυπο DICOM σε μορφότυπο TIFF.....	66
3.2.4	Αφαίρεση των τομών στις οποίες δεν διακρίνεται το πνευμονικό παρέγχυμα ή που δεν εμφανίζουν αλλοιώσεις.....	66
3.2.5	Αλλαγή Διαστάσεων Εικόνας.....	68
3.2.6	Κανονικοποίηση Τιμών Εικονοστοιχείων στο εύρος 0-1 .....	68
3.2.7	One-Hot Encoding .....	68
3.2.8	Διαχωρισμός σε σύνολα Εκπαίδευσης/Επικύρωσης/Ελέγχου .....	69
3.2.9	Προσαύξηση Δεδομένων Εκπαίδευσης (Data Augmentation) .....	71
3.3	Αρχιτεκτονικές Τεχνητών Νευρωνικών Δικτύων .....	71
3.3.1	Αρχιτεκτονική Inception .....	72

3.3.2	Αρχιτεκτονική ResNet.....	73
3.3.3	Αρχιτεκτονική DenseNet.....	74
3.3.4	Αρχιτεκτονική Xception .....	75
3.4	Αξιολόγηση Μοντέλων.....	76
3.4.1	Εφαρμογή Διασταυρούμενης Επικύρωσης 5-Πτυχών .....	77
3.4.2	Καμπύλες Εκπαίδευσης και Επικύρωσης.....	77
3.4.3	Πίνακας Σύγχυσης (Confusion Matrix).....	77
3.4.4	Ακρίβεια (Accuracy) .....	79
3.4.5	Απώλεια (Loss) .....	79
3.4.6	Αξιοπιστία (Precision).....	79
3.4.7	Ανάκληση (Recall).....	80
3.4.8	Συνδυασμός Ακρίβειας-Ανάκλησης (F1 Score) .....	80
3.4.9	Gradient-weighted Class Activation Mapping (Grad-CAM).....	80
3.5	Σχετικές Εργασίες .....	81
3.6	Πειραματική Διαδικασία.....	83
3.6.1	Δυαδική Ταξινόμηση των Τομών στις Κατηγορίες Covid-19 και Normal .....	86
3.6.2	Εφαρμογή της Μεθόδου Grad-CAM σε Ταξινομημένες Τομές στις Κατηγορίες COVID-19 και Normal.....	105
3.6.3	Σύγκριση Αποτελεσμάτων για την Περίπτωση Ταξινόμησης στις Κατηγορίες Covid-19 και Normal.....	109
3.6.4	Δυαδική Ταξινόμηση των Τομών Αξονικών Τομογραφιών στις Κατηγορίες Covid-19 και Other Pneumonia.....	110
3.6.5	Σύγκριση Αποτελεσμάτων για την Περίπτωση Ταξινόμησης στις Κατηγορίες Covid-19 και Other Pneumonia.....	118
4	Περιγραφή Μεθοδολογίας Πειράματος Σημασιολογικής Κατάτμησης Εικόνων (Semantic Image Segmentation) .....	120
4.1	Επιλογή Συνόλου Δεδομένων Πειράματος Σημασιολογικής Κατάτμησης Εικόνων (Semantic Image Segmentation).....	121
4.1.1	Σύνολο Δεδομένων COVID-19-CT-Seg .....	121
4.2	Προεπεξεργασία Δεδομένων .....	123
4.2.1	Κανονικοποίηση Αντίθεσης (Contrast) και Φωτεινότητας (Brightness) Τομών Αξονικών Τομογραφιών .....	123
4.2.2	Δημιουργία Παραλλαγής του Συνόλου Δεδομένων όπου Διατηρείται μόνο το Πνευμονικό Παρέγχυμα στις Τομές των Αξονικών Τομογραφιών .....	124
4.2.3	Μετατροπή Εικόνων από Μορφότυπο DICOM σε Μορφότυπο TIFF .....	125
4.2.4	Αφαίρεση των Τομών που δεν Εμφανίζουν Άλλοιώσεις .....	125

4.2.5	Αλλαγή Διαστάσεων Εικόνων .....	125
4.2.6	Κανονικοποίηση Τιμών Εικονοστοιχείων στο Εύρος Τιμών 0-1 .....	126
4.2.7	Διαχωρισμός Δεδομένων σε Σύνολα Εκπαίδευσης/Επικύρωσης/Ελέγχου ...	126
4.2.8	Προσαύξηση Δεδομένων (Data Augmentation) .....	126
4.3	Αρχιτεκτονικές Τεχνητών Νευρωνικών Δικτύων .....	127
4.3.1	Αρχιτεκτονική U-Net .....	127
4.3.2	Αρχιτεκτονική U-Net με κωδικοποιητή την αρχιτεκτονική VGG16 .....	129
4.3.3	Αρχιτεκτονική U-Net με κωδικοποιητή την αρχιτεκτονική DenseNet121....	129
4.4	Αξιολόγηση Μοντέλων.....	130
4.4.1	Εφαρμογή Διασταυρούμενης Επικύρωσης 5-Πτυχών .....	130
4.4.2	Καμπύλες Εκπαίδευσης και Επικύρωσης.....	130
4.4.3	Αξιοπιστία (Precision).....	130
4.4.4	Ανάκληση (Recall).....	131
4.4.5	Dice Similarity Coefficient (DSC).....	131
4.4.6	Intersection over Union (IoU).....	131
4.5	Σχετικές Εργασίες .....	132
4.6	Πειραματική Διαδικασία.....	133
4.6.1	Σημασιολογική Κατάτμηση Αλλοιώσεων Πνευμονίας από COVID-19 .....	135
4.6.2	Εφαρμογή της σημασιολογικής κατάτμησης αλλοιώσεων σε τομές των συνόλων επικύρωσης και ελέγχου.....	148
4.6.3	Σύγκριση αποτελεσμάτων των τεσσάρων παραλλαγών.....	152
5	Επίλογος .....	155

## **Ευρετήριο Εικόνων**

Εικόνα 2.1: Αξονικός τομογράφος.....	23
Εικόνα 2.2: Τομή αξονικής τομογραφίας .....	24
Εικόνα 2.3: Περιπτώσεις υποπροσαρμογής και υπερπροσαρμογής σε πρόβλημα ταξινόμησης (classification) και παλινδρόμησης (regression).....	32
Εικόνα 2.4: Μέθοδος holdout .....	34
Εικόνα 2.5: Μέθοδος διασταυρούμενης επικύρωσης k-πτυχών με k=5.....	35
Εικόνα 2.6: Δομή ενός βιολογικού νευρώνα .....	36
Εικόνα 2.7: Το Perceptron του Rosenblatt .....	40
Εικόνα 2.8: Perceptron πολλών επιπέδων .....	40
Εικόνα 2.9: Σχηματική αναπαράσταση της εκτέλεσης του αλγορίθμου gradient descent .....	45
Εικόνα 2.10: Βασική δομή ενός συνελικτικού νευρωνικού δίκτυου .....	47
Εικόνα 2.11: Παράδειγμα λειτουργίας επιπέδου συνέλιξης .....	48
Εικόνα 2.12: Παράδειγμα εφαρμογής του βήματος ολίσθησης (stride) .....	49
Εικόνα 2.13: Παράδειγμα εφαρμογής της τεχνικής padding (zero-padding) .....	49
Εικόνα 2.14: Παράδειγμα λειτουργίας των δύο τύπων επιπέδου συγκέντρωσης .....	50
Εικόνα 2.15: Σημείο πρόωρου τερματισμού της εκπαίδευσης πριν εμφανισθεί υπερπροσαρμογή .....	53
Εικόνα 2.16: Παράδειγμα εφαρμογής dropout σε ένα τεχνητό νευρωνικό δίκτυο .....	54
Εικόνα 2.17: Εφαρμογή της μεθόδου μεταφοράς μάθησης.....	56
Εικόνα 3.1: Διαδικασία εκπαίδευσης και αξιολόγησης μοντέλου ως προς τη ταξινόμηση τομών με βάση αλλοιώσεις από πνευμονία COVID-19, όπου χρησιμοποιείται μεταφορά μάθησης.....	57
Εικόνα 3.2: Ραβδόγραμμα κατανομής των ασθενών, στους οποίους ανήκουν οι τομές αξονικών τομογραφιών, στις κλάσεις του συνόλου δεδομένων COVID-CT-MD.....	60
Εικόνα 3.3: Ραβδόγραμμα κατανομής των τομών αξονικών τομογραφιών στις κλάσεις του συνόλου δεδομένων COVID-CT-MD.....	60
Εικόνα 3.4: Υποσύνολο από τις τομές αξονικών τομογραφιών του συνόλου δεδομένων COVID-CT-MD.....	61
Εικόνα 3.5: Ραβδόγραμμα κατανομής των ασθενών, στους οποίους ανήκουν οι τομές αξονικών τομογραφιών, στις κλάσεις του συνόλου δεδομένων CC-CCII .....	62
Εικόνα 3.6: Ραβδόγραμμα κατανομής των τομών αξονικών τομογραφιών στις κλάσεις του συνόλου δεδομένων CC-CCII .....	62
Εικόνα 3.7: Υποσύνολο από τις τομές αξονικών τομογραφιών του συνόλου δεδομένων CC-CCII.....	63
Εικόνα 3.8: Τομή αξονικής τομογραφίας του συνόλου COVID-CT-MD πριν και μετά την κανονικοποίηση αντίθεσης και φωτεινότητας .....	64
Εικόνα 3.9: Τομή αξονικής τομογραφίας του συνόλου COVID-CT-MD πριν και μετά την κατάτμηση του πνευμονικού παρεγχύματος.....	65
Εικόνα 3.10: Ραβδόγραμμα κατανομής των τομών αξονικών τομογραφιών στις κλάσεις του συνόλου δεδομένων COVID-CT-MD έπειτα από την αφαίρεση των τομών .....	67
Εικόνα 3.11: Ραβδόγραμμα κατανομής των τομών αξονικών τομογραφιών στις κλάσεις του συνόλου δεδομένων CC-CCII έπειτα από την αφαίρεση των τομών.....	67
Εικόνα 3.12: Παράδειγμα one-hot encoding .....	69

Εικόνα 3.13: Εφαρμογή προσαύξησης δεδομένων σε τομή αξονικής τομογραφίας του συνόλου COVID-CT-MD .....	71
Εικόνα 3.14: Ένα inception module της αρχιτεκτονικής InceptionV3 .....	73
Εικόνα 3.15: Ένα Residual μπλοκ .....	74
Εικόνα 3.16: Ένα dense μπλοκ 5 επιπέδων στο οποίο πραγματοποιείται επαναχρησιμοποίηση χαρακτηριστικών.....	75
Εικόνα 3.17: Η δομή ενός extreme inception module.....	76
Εικόνα 3.18: Πίνακας σύγχυσης για ένα πρόβλημα διαδικής ταξινόμησης .....	78
Εικόνα 3.19: Παράδειγμα εφαρμογής της μεθόδου Grad-CAM σε τομή που εμφανίζει πνευμονία COVID-19 .....	81
Εικόνα 3.20: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Τομή: Συνενωμένοι πίνακες σύγχυσης διαδικασίας ταξινόμησης των δεδομένων επικύρωσης στις κατηγορίες Covid-19 και Normal .....	87
Εικόνα 3.21: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Τομή: Καμπύλες εκπαίδευσης και επικύρωσης των μοντέλων όσον αφορά την ταξινόμηση στις κατηγορίες Covid-19 και Normal .....	89
Εικόνα 3.22: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Τομή: Συνενωμένοι πίνακες σύγχυσης διαδικασίας ταξινόμησης των δεδομένων ελέγχου στις κατηγορίες Covid-19 και Normal .....	91
Εικόνα 3.23: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Ασθενή: Συνενωμένοι πίνακες σύγχυσης διαδικασίας ταξινόμησης των δεδομένων επικύρωσης στις κατηγορίες Covid-19 και Normal .....	93
Εικόνα 3.24: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Ασθενή: Καμπύλες εκπαίδευσης και επικύρωσης των μοντέλων όσον αφορά την ταξινόμηση στις κατηγορίες Covid-19 και Normal .....	95
Εικόνα 3.25: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Ασθενή: Συνενωμένοι πίνακες σύγχυσης διαδικασίας ταξινόμησης των δεδομένων ελέγχου στις κατηγορίες Covid-19 και Normal .....	97
Εικόνα 3.26: Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Τομή: Συνενωμένοι πίνακες σύγχυσης διαδικασίας ταξινόμησης των δεδομένων επικύρωσης στις κατηγορίες Covid-19 και Normal .....	99
Εικόνα 3.27: Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Τομή: Καμπύλες εκπαίδευσης και επικύρωσης των μοντέλων όσον αφορά την ταξινόμηση στις κατηγορίες Covid-19 και Normal .....	101
Εικόνα 3.28: Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Ασθενή: Συνενωμένοι πίνακες σύγχυσης διαδικασίας ταξινόμησης των δεδομένων επικύρωσης στις κατηγορίες Covid-19 και Normal .....	103
Εικόνα 3.29: Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Ασθενή: Καμπύλες εκπαίδευσης και επικύρωσης των μοντέλων όσον αφορά την ταξινόμηση στις κατηγορίες Covid-19 και Normal .....	104
Εικόνα 3.30: Τομές συνόλου επικύρωσης χωρίς κατάτμηση πνευμονικού παρεγχύματος ορθά κατηγοριοποιημένες ως Covid-19.....	105
Εικόνα 3.31: Τομές συνόλου επικύρωσης με κατάτμηση πνευμονικού παρεγχύματος ορθά κατηγοριοποιημένες ως Covid-19.....	106
Εικόνα 3.32: Τομές συνόλου επικύρωσης χωρίς κατάτμηση πνευμονικού παρεγχύματος ορθά κατηγοριοποιημένες ως Normal .....	106

Εικόνα 3.33: Τομές συνόλου επικύρωσης με κατάτμηση πνευμονικού παρεγχύματος ορθά κατηγοριοποιημένες ως Normal .....	107
Εικόνα 3.34: Τομές συνόλου ελέγχου χωρίς κατάτμηση πνευμονικού παρεγχύματος ορθά κατηγοριοποιημένες ως Covid-19.....	108
Εικόνα 3.35: Τομές συνόλου ελέγχου χωρίς κατάτμηση πνευμονικού παρεγχύματος ορθά κατηγοριοποιημένες ως Normal .....	108
Εικόνα 3.36: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Τομή: Συνενωμένοι πίνακες σύγχυσης διαδικασίας ταξινόμησης των δεδομένων επικύρωσης στις κατηγορίες Covid-19 και Other Pneumonia .....	112
Εικόνα 3.37: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Τομή: Συνενωμένοι πίνακες σύγχυσης διαδικασίας ταξινόμησης των δεδομένων ελέγχου στις κατηγορίες Covid-19 και Other Pneumonia .....	112
Εικόνα 3.38: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Ασθενή: Συνενωμένοι πίνακες σύγχυσης διαδικασίας ταξινόμησης των δεδομένων επικύρωσης στις κατηγορίες Covid-19 και Other Pneumonia.....	114
Εικόνα 3.39: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Ασθενή: Συνενωμένοι πίνακες σύγχυσης διαδικασίας ταξινόμησης των δεδομένων ελέγχου στις κατηγορίες Covid-19 και Other Pneumonia .....	115
Εικόνα 3.40: Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Τομή: Συνενωμένοι πίνακες σύγχυσης διαδικασίας ταξινόμησης των δεδομένων επικύρωσης στις κατηγορίες Covid-19 και Other Pneumonia .....	116
Εικόνα 3.41: Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Ασθενή: Συνενωμένοι πίνακες σύγχυσης διαδικασίας ταξινόμησης των δεδομένων επικύρωσης στις κατηγορίες Covid-19 και Other Pneumonia.....	118
Εικόνα 4.1: Διαδικασία εκπαίδευσης και αξιολόγησης μοντέλου ως προς τη σημασιολογική κατάτμηση αλλοιώσεων από πνευμονία COVID-19 σε τομές, όπου αξιοποιείται η τεχνική της μεταφοράς μάθησης .....	120
Εικόνα 4.2: Ραβδόγραμμα υποσυνόλων συνόλου δεδομένων COVID-19-CT-Seg.....	122
Εικόνα 4.3: Υποσύνολο από τις τομές αξονικών τομογραφιών και τις αντίστοιχες «μάσκες» του συνόλου δεδομένων COVID-19-CT-Seg .....	123
Εικόνα 4.4: Τομή αξονικής τομογραφίας πριν και μετά την κανονικοποίηση αντίθεσης και φωτεινότητας .....	124
Εικόνα 4.5: Τομή αξονικής τομογραφίας πριν και μετά την κατάτμηση πνευμονικού παρεγχύματος.....	124
Εικόνα 4.6: Ραβδόγραμμα υποσυνόλων του συνόλου δεδομένων COVID-19-CT-Seg έπειτα από την αφαίρεση των τομών που δεν εμφανίζουν αλλοιώσεις .....	125
Εικόνα 4.7: Εφαρμογή προσαύξησης δεδομένων σε τομή αξονικής τομογραφίας του συνόλου COVID-19-CT-Seg .....	127
Εικόνα 4.8: Διάγραμμα Αρχιτεκτονικής U-Net .....	128
Εικόνα 4.9: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Τομή: Καμπύλες εκπαίδευσης και επικύρωσης των μοντέλων όσον αφορά τη σημασιολογική κατάτμηση αλλοιώσεων .....	137
Εικόνα 4.10: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Ασθενή: Καμπύλες εκπαίδευσης και επικύρωσης των μοντέλων όσον αφορά τη σημασιολογική κατάτμηση αλλοιώσεων.....	140

Εικόνα 4.11: Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Τομή: Καμπύλες εκπαίδευσης και επικύρωσης των μοντέλων όσον αφορά τη σημασιολογική κατάτμηση αλλοιώσεων .....	143
Εικόνα 4.12: Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Ασθενή: Καμπύλες εκπαίδευσης και επικύρωσης των μοντέλων όσον αφορά τη σημασιολογική κατάτμηση αλλοιώσεων.....	147
Εικόνα 4.13: Εφαρμογή σημασιολογικής κατάτμησης αλλοιώσεων σε τομές του συνόλου επικύρωσης όπου δεν έχει εφαρμοσθεί κατάτμηση πνευμονικού παρεγχύματος .....	149
Εικόνα 4.14: Εφαρμογή σημασιολογικής κατάτμησης αλλοιώσεων σε τομές του συνόλου ελέγχου όπου δεν έχει εφαρμοσθεί κατάτμηση πνευμονικού παρεγχύματος .....	150
Εικόνα 4.15: Εφαρμογή σημασιολογικής κατάτμησης αλλοιώσεων σε τομές του συνόλου επικύρωσης όπου έχει εφαρμοσθεί κατάτμηση του πνευμονικού παρεγχύματος .....	151
Εικόνα 4.16: Εφαρμογή σημασιολογικής κατάτμησης αλλοιώσεων σε τομές του συνόλου ελέγχου όπου έχει εφαρμοσθεί κατάτμηση του πνευμονικού παρεγχύματος .....	152

## **Ευρετήριο Πινάκων**

Πίνακας 2.1: Τιμές Hounsfield που αφορούν συνήθεις ιστούς .....	25
Πίνακας 2.2: Ρυθμίσεις παραθύρου που χρησιμοποιούνται στην απεικόνιση εικόνων αξονικών τομογραφιών .....	26
Πίνακας 2.3: Γεωμετρικοί μετασχηματισμοί εικόνων .....	52
Πίνακας 3.1: Αρχιτεκτονικές που χρησιμοποιήθηκαν για το πρόβλημα ταξινόμησης και οι παράμετροί τους.....	84
Πίνακας 3.2: Υπερπαράμετροι διαδικασίας εκπαίδευσης για την ταξινόμηση των τομών στις κατηγορίες Covid-19 και Normal .....	84
Πίνακας 3.3: Υπερπαράμετροι διαδικασίας εκπαίδευσης για την ταξινόμηση των τομών στις κατηγορίες Covid-19 και Other Pneumonia .....	85
Πίνακας 3.4: Παραλλαγές του συνόλου COVID-CT-MD .....	85
Πίνακας 3.5: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Τομή: Μετρικές αξιολόγησης διαδικασίας ταξινόμησης των δεδομένων επικύρωσης στις κατηγορίες Covid-19 και Normal .....	86
Πίνακας 3.6: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Τομή: Μετρικές αξιολόγησης της ταξινόμησης των δεδομένων ελέγχου στις κατηγορίες Covid-19 και Normal .....	90
Πίνακας 3.7: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Ασθενή: Μετρικές αξιολόγησης της ταξινόμησης των δεδομένων επικύρωσης στις κατηγορίες Covid-19 και Normal .....	92
Πίνακας 3.8: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Ασθενή: Μετρικές αξιολόγησης της ταξινόμησης των δεδομένων ελέγχου στις κατηγορίες Covid-19 και Normal .....	96
Πίνακας 3.9: Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Τομή: Μετρικές αξιολόγησης της ταξινόμησης των δεδομένων επικύρωσης στις κατηγορίες Covid-19 και Normal .....	98
Πίνακας 3.10: Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Ασθενή: Μετρικές αξιολόγησης της ταξινόμησης των δεδομένων επικύρωσης στις κατηγορίες Covid-19 και Normal .....	102
Πίνακας 3.11: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Τομή: Μετρικές αξιολόγησης της ταξινόμησης των δεδομένων επικύρωσης στις κατηγορίες Covid-19 και Other Pneumonia .....	111
Πίνακας 3.12: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Τομή: Μετρικές αξιολόγησης της ταξινόμησης των δεδομένων ελέγχου στις κατηγορίες Covid-19 και Other Pneumonia .....	111
Πίνακας 3.13: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Ασθενή: Μετρικές αξιολόγησης της ταξινόμησης των δεδομένων επικύρωσης στις κατηγορίες Covid-19 και Other Pneumonia .....	113
Πίνακας 3.14: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Ασθενή: Μετρικές αξιολόγησης της ταξινόμησης των δεδομένων ελέγχου στις κατηγορίες Covid-19 και Other Pneumonia .....	114
Πίνακας 3.15: Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Τομή: Μετρικές αξιολόγησης της ταξινόμησης των δεδομένων επικύρωσης στις Κατηγορίες Covid-19 και Other Pneumonia.....	116

Πίνακας 3.16: Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Ασθενή: Μετρικές αξιολόγησης της ταξινόμησης των δεδομένων επικύρωσης στις κατηγορίες Covid-19 και Other Pneumonia .....	117
Πίνακας 4.1: Αρχιτεκτονικές που χρησιμοποιήθηκαν στο πρόβλημα σημασιολογικής κατάτμησης αλλοιώσεων και οι παράμετροί τους .....	133
Πίνακας 4.2: Υπερπαράμετροι διαδικασίας εκπαίδευσης για τη σημασιολογική κατάτμηση αλλοιώσεων .....	134
Πίνακας 4.3: Παραλλαγές του συνόλου δεδομένων COVID-19-CT-Seg .....	134
Πίνακας 4.4: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Τομή: Μετρικές αξιολόγησης της εφαρμογής σημασιολογικής κατάτμησης αλλοιώσεων στα δεδομένα επικύρωσης .....	135
Πίνακας 4.5: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Τομή: Μετρικές αξιολόγησης της εφαρμογής σημασιολογικής κατάτμησης αλλοιώσεων στα δεδομένα ελέγχου.....	137
Πίνακας 4.6: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Ασθενή: Μετρικές αξιολόγησης της εφαρμογής σημασιολογικής κατάτμησης αλλοιώσεων στα δεδομένα επικύρωσης.....	139
Πίνακας 4.7: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Ασθενή: Μετρικές αξιολόγησης της εφαρμογής σημασιολογικής κατάτμησης αλλοιώσεων στα δεδομένα ελέγχου .....	141
Πίνακας 4.8: Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Τομή: Μετρικές αξιολόγησης της εφαρμογής σημασιολογικής κατάτμησης αλλοιώσεων στα δεδομένα επικύρωσης .....	142
Πίνακας 4.9: Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Τομή: Μετρικές αξιολόγησης της εφαρμογής σημασιολογικής κατάτμησης αλλοιώσεων στα δεδομένα ελέγχου.....	144
Πίνακας 4.10: Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Ασθενή: Μετρικές αξιολόγησης της εφαρμογής σημασιολογικής κατάτμησης αλλοιώσεων στα δεδομένα επικύρωσης.....	145
Πίνακας 4.11: Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Ασθενή: Μετρικές αξιολόγησης της εφαρμογής σημασιολογικής κατάτμησης αλλοιώσεων στα δεδομένα ελέγχου .....	147

## 1 Εισαγωγή

Στις αρχές του έτους 2020 τα μέσα μαζικής ενημέρωσης σε όλο τον κόσμο έκαναν λόγο για μία άγνωστη μέχρι τότε ιδιαίτερα μεταδοτική ασθένεια η οποία εμφανίστηκε για πρώτη φορά στα τέλη του έτους 2019 στη πόλη Γουχάν, πρωτεύουσα της επαρχίας Χουμπέι στην Κεντρική Κίνα, και η οποία πρόσβαλε το αναπνευστικό σύστημα προκαλώντας ιογενή πνευμονία. Στις 11 Φεβρουαρίου του 2020 ο Παγκόσμιος Οργανισμός Υγείας ονόμασε τη νέα αυτή ασθένεια “COVID-19” και τον ίο που την προκαλεί “SARS-CoV-2” [1]. Σύμφωνα με την επιστημονική κοινότητα, η νέα αυτή ασθένεια αποτελεί μια εξαιρετικά μεταδοτική και παθογόνο ιογενή λοίμωξη που προσβάλλει το αναπνευστικό σύστημα και προκαλείται από τον προαναφερθέντα ίο [2]. Τα συνήθη συμπτώματα που εμφανίζουν οι ασθενείς που προσβάλλονται από τον συγκεκριμένο ίο είναι πυρετός, βήχας, πονόλαιμος, δύσπνοια, κόπωση, δυσφορία, ανοσμία και αγευσία [3]. Στις 11 Μαρτίου του 2020 ο Παγκόσμιος Οργανισμός Υγείας χαρακτήρισε την έξαρση της COVID-19 ως πανδημία λόγω των ανησυχητικών επιπέδων εξάπλωσης και σοβαρότητας του ιού [4]. Σήμερα, η πανδημία της COVID-19 έχει προσβάλλει μια ιδιαίτερα μεγάλη μερίδα του πληθυσμού και παρότι έχουν ανακαλυφθεί εμβόλια που προστατεύουν αποτελεσματικά από αυτόν, συνεχίζει να αποτελεί σοβαρή απειλή για την ανθρωπότητα.

Σύμφωνα με τον Παγκόσμιο Οργανισμό Υγείας, από την πρώτη εμφάνιση του ιού SARS-CoV-2 και κατά την εκπόνηση της παρούσας εργασίας έχουν επιβεβαιωθεί 770.778.396 κρούσματα COVID-19 από τα οποία τα 6.958.499 κατέληξαν σε θάνατο [5]. Ο τόσο υψηλός αριθμός κρουσμάτων οφείλεται τόσο στο γεγονός ότι ο ιός εμφανίζει υψηλή μεταδοτικότητα, όσο και στο γεγονός ότι πολλά άτομα που προσβάλλονται από τον ίο SARS-CoV-2 είναι ασυμπτωματικά ή εμφανίζουν τα πρώτα συμπτώματα έως και 14 ημέρες μετά την προσβολή τους από αυτόν [3].

Η έγκαιρη ανίχνευση και διάγνωση της λοίμωξης COVID-19 αποτελεί βασική προϋπόθεση για τον περιορισμό της εξάπλωσής του ιού [6]. Τα διαγνωστικά τεστ αποτελούν τον πιο συνήθη τρόπο διάγνωσης του ιού SARS-CoV-2. Οι τρεις διαθέσιμοι τύποι τεστ είναι τα τεστ ανάστροφης μεταγραφής αλυσιδωτής αντίδρασης πολυμεράσης (RT-PCR test), τα τεστ ταχείας ανίχνευσης αντιγόνου (Rapid antigen test) καθώς και τα αυτοδιαγνωστικά τεστ (Self test) [7]. Από τους τρεις αυτούς τύπους, τα τεστ ανάστροφης μεταγραφής αλυσιδωτής αντίδρασης πολυμεράσης (RT-PCR test) αποτελούν τα πιο αξιόπιστα. Ωστόσο, το γεγονός ότι είναι χρονοβόρα αποτελεί σημαντικό μειονέκτημα, ειδικά όσον αφορά το συγκεκριμένο ίο όπου η έγκαιρη διάγνωση αποτελεί παράγοντα κλειδί για την επιτυχή αντιμετώπισή του [6].

Η μελέτη ιατρικών εικόνων όπως ακτινογραφίες θώρακος και αξονικές τομογραφίες αποτελεί μια μέθοδο που μπορεί να θέσει την υποψία για τη διάγνωση της νόσου COVID-19 μέσω της ανίχνευσης παθολογικών ευρημάτων που συνάδουν με λοίμωξη του κατωτέρου αναπνευστικού συστήματος (πνευμονία) [8]. Οι ιατρικές

εικόνες χρησιμοποιούνται ευρέως από τους ειδικούς ως διαγνωστικά μέσα για τη σχετιζόμενη με τον ιό SARS-CoV-2 πνευμονία, με τις αξονικές τομογραφίες να εμφανίζουν πολύ υψηλότερη διαγνωστική ακρίβεια καθώς μπορούν να αναγνωρίσουν αρχόμενες βλάβες στο πνευμονικό παρέγχυμα που δεν μπορούν να διακριθούν στην απλή ακτινογραφία θώρακος [9].

## 1.1 Η σημασία της χρήσης τεχνικών μηχανικής μάθησης στην διαδικασία ανίχνευσης παθήσεων μέσω ιατρικών εικόνων

Η μελέτη ιατρικών εικόνων όπως ακτινογραφιών, αξονικών τομογραφιών και μαγνητικών τομογραφιών για την έγκαιρη ανίχνευση, διάγνωση και θεραπεία τυχόν παθήσεων αποτελεί μια μέθοδο η οποία ακολουθείται από τους ειδικούς εδώ και δεκαετίες. Μέσω της επισκόπησης ιατρικών εικόνων οι ειδικοί έχουν τη δυνατότητα να αποφανθούν για την κατάσταση της υγείας του εκάστοτε ασθενή και στη περίπτωση ανακάλυψης κάποιου ανησυχητικού ευρήματος να πραγματοποιήσουν την σωστή διάγνωση και να ακολουθήσουν την κατάλληλη θεραπεία. Ωστόσο, παρότι η διαδικασία επισκόπησης ιατρικών εικόνων από ιατρικό προσωπικό οδηγεί στις περισσότερες περιπτώσεις στη σωστή και έγκαιρη ανίχνευση, διάγνωση και θεραπεία, το 40%-54% των αμελειών που οφείλονται σε ιατρούς ακτινολόγους συνδέεται με λάθη κατά τη διαδικασία ερμηνείας ιατρικών εικόνων [10]. Πιο συγκεκριμένα, η αδυναμία του ιατρού να εντοπίσει την περιοχή στην οποία εμφανίζεται κάποια αλλοίωση ή να αναγνωρίσει σε ποια πάθηση οφείλεται η συγκεκριμένη αλλοίωση, η συσσωρευμένη κούραση που ενδεχομένως έχει αποκομίσει ο ιατρός λόγω των απαιτήσεων του επαγγέλματος, καθώς και η ενδεχόμενη απειρία του ιατρού που έχει αναλάβει τη μελέτη της εκάστοτε ιατρικής εικόνας αποτελούν ορισμένους από τους λόγους που οι ειδικοί οδηγούνται σε λάθος συμπεράσματα [10].

Από τα προαναφερθέντα, γίνεται αντιληπτό ότι η μελέτη ιατρικών εικόνων για την ανίχνευση τυχόν ευρημάτων μέσω τεχνικών μηχανικής μάθησης δύναται να αποτελέσει ένα πολύτιμο εργαλείο για την έγκαιρη και σωστή ανίχνευση και κατ' επέκταση διάγνωση ενδεχόμενων παθήσεων. Πιο συγκεκριμένα, το γεγονός ότι ένα σύστημα το οποίο έχει εκπαιδευτεί μέσω τεχνικών μηχανικής μάθησης είναι ικανό να ερμηνεύσει έναν πολύ μεγάλο αριθμό ιατρικών εικόνων σε μικρό χρονικό διάστημα, καθώς και το γεγονός ότι εάν το σύστημα έχει εκπαιδευτεί κατάλληλα δύναται να πραγματοποιήσει την διαδικασία επισκόπησης ιατρικών εικόνων καλύτερα ακόμη και από το ίδιο το ιατρικό προσωπικό, καθιστά τη συγκεκριμένη μέθοδο ιδανική, ειδικά τη συγκεκριμένη περίοδο όπου λόγω της πανδημίας και του μεγάλου αριθμού κρουσμάτων COVID-19 που διαγιγνώσκονται καθημερινά, το βάρος που κουβαλούν στις πλάτες τους οι ειδικοί που ασχολούνται με την μελέτη ιατρικών εικόνων είναι ιδιαίτερα μεγάλο [11].

## 1.2 Περιγραφή του αντικειμένου και του στόχου της διπλωματικής εργασίας

Η χρήση τεχνικών μηχανικής μάθησης για την ανάλυση ιατρικών εικόνων με στόχο την υποβοήθηση στη διάγνωση πνευμονίας από COVID-19, αποτελεί το αντικείμενο το οποίο πραγματεύεται η παρούσα διπλωματική εργασία. Τα τελευταία χρόνια έχει υπάρξει αυξημένο ενδιαφέρον όσον αφορά την εφαρμογή μεθόδων μηχανικής μάθησης στο τομέα της ιατρικής και συγκεκριμένα στη διαδικασία μελέτης ιατρικών εικόνων, με αποτέλεσμα να έχει δημοσιευθεί μια πληθώρα από επιστημονικές εργασίες με αξιόλογα αποτελέσματα. Όπως ήταν αναμενόμενο, με την εμφάνιση του ιού SARS-CoV-2 και την έλευση της πανδημίας δημιουργήθηκε μεγάλο ενδιαφέρον από την επιστημονική κοινότητα σχετικά με την μελέτη της δυνατότητας εφαρμογής των μεθόδων μηχανικής μάθησης στη διαδικασία επισκόπησης ιατρικών εικόνων με στόχο την ανίχνευση ύπαρξης της νόσου COVID-19. Αυτό είχε ως αποτέλεσμα να δημοσιευθούν τόσο ερευνητικές εργασίες όσο και σύνολα δεδομένων για τον σκοπό αυτό. Η επικαιρότητα καθώς και το μεγάλο ενδιαφέρον που παρουσιάζει το συγκεκριμένο θέμα αποτέλεσαν το κίνητρο για την ενασχόληση μας με αυτό.

Το ζήτημα της ανάλυσης ιατρικών εικόνων για την υποβοήθηση στη διάγνωση της λοίμωξης COVID-19 μπορεί να κατηγοριοποιηθεί είτε ως ένα πρόβλημα ταξινόμησης εικόνων (image classification problem) όπου πραγματοποιείται εκπαίδευση μοντέλων μηχανικής μάθησης ούτως ώστε να μπορούν να ταξινομούν τομές αξονικών τομογραφιών με βάση την ύπαρξη ή μη πνευμονίας προερχόμενης από τον ιό SARS-CoV-2, είτε ως ένα πρόβλημα σημασιολογικής κατάτμησης εικόνων (semantic image segmentation problem) όπου εκπαιδεύονται μοντέλα μηχανικής μάθησης ούτως ώστε να μπορούν να πραγματοποιούν κατάτμηση των περιοχών εκείνων των τομών αξονικών τομογραφιών που εμφανίζουν βλάβες προερχόμενες από πνευμονία COVID-19.

Όσον αφορά τη παρούσα διπλωματική εργασία, αυτό που επιδιώκεται είναι η διερεύνηση του κατά πόσο μοντέλα βαθειάς μάθησης και συγκεκριμένα συνελικτικά νευρωνικά δίκτυα (CNNs) είναι ικανά έπειτα από εκπαίδευση να αποτελέσουν λύση για τους δύο τύπους προβλημάτων που προαναφέρθηκαν και στους οποίους μπορεί να κατηγοριοποιηθεί το θέμα που πραγματεύεται η συγκεκριμένη εργασία.

Για την αποδοτικότερη εκπαίδευση των νευρωνικών δικτύων που υλοποιούνται για τις ανάγκες της παρούσας εργασίας, επιστρατεύεται μια τεχνική η οποία ονομάζεται μεταφορά μάθησης (transfer learning), όπου στην ουσία πραγματοποιείται εκμετάλλευση της γνώσης η οποία έχει αποκτηθεί από την εφαρμογή μοντέλων βαθειάς μάθησης στην κατηγοριοποίηση ενός συνόλου δεδομένων που ονομάζεται ImageNet [12] και το οποίο αποτελείται από έναν μεγάλο όγκο εικόνων.

Απότερος στόχος είναι η συγκεκριμένη εργασία να παρέχει μία ολοκληρωμένη παρουσίαση της δυνατότητας εφαρμογής μεθόδων βαθειάς μάθησης στον τομέα της ιατρικής και συγκεκριμένα στην ανίχνευση αλλοιώσεων προερχόμενων από τον ιό SARS-CoV2.

### 1.3 Διπλωματικής Εργασίας

Η παρούσα διπλωματική εργασία διαρθρώνεται με τον ακόλουθο τρόπο:

Στο Κεφάλαιο 2 παρουσιάζεται το θεωρητικό υπόβαθρο που είναι απαραίτητο για την κατανόηση των εννοιών που πραγματεύεται η συγκεκριμένη εργασία καθώς και της μεθοδολογίας και της πειραματικής διαδικασίας που περιγράφονται στα επόμενα κεφάλαια. Πιο αναλυτικά, αρχικά πραγματοποιείται μια περιγραφή των τεχνικών ιατρικής απεικόνισης δίνοντας έμφαση στις αξονικές τομογραφίες θώρακος. Στη συνέχεια, γίνεται μια αναφορά στο πεδίο της επιστήμης των υπολογιστών που ονομάζεται μηχανική μάθηση (machine learning) και αναλύεται η διαδικασία με την οποία εκπαιδεύονται οι αλγόριθμοι που συγκαταλέγονται στο συγκεκριμένο πεδίο. Έπειτα, παρουσιάζεται το πεδίο της όρασης υπολογιστών (computer vision), καθώς και οι δύο κατηγορίες του συγκεκριμένου πεδίου στις οποίες υπάγεται το ζήτημα που πραγματεύεται η παρούσα εργασία, δηλαδή η κατηγορία της ταξινόμησης εικόνων (image classification) και η κατηγορία της σημασιολογικής κατάτμησης εικόνων (semantic image segmentation). Ακόμη, αναλύεται το πεδίο της μηχανικής μάθησης που ονομάζεται βαθιά μάθηση (deep learning) και περιγράφονται οι βασικές δομές του πεδίου που ονομάζονται τεχνητά νευρωνικά δίκτυα (ANN), δίνοντας ιδιαίτερη έμφαση σε βασικές πτυχές του τρόπου λειτουργίας τους, όπως τους τεχνητούς νευρώνες (artificial neurons), τη συνάρτηση ενεργοποίησης (activation function), τη συνάρτηση απώλειας/κόστους (loss/cost function), τον αλγόριθμο βελτιστοποίησης (optimization algorithm), καθώς και τον αλγόριθμο οπισθοδιάδοσης (backpropagation algorithm). Επιπρόσθετα, παρουσιάζεται η κατηγορία των συνελικτικών νευρωνικών δικτύων (CNN), στην οποία συγκαταλέγονται οι αρχιτεκτονικές που υλοποιήθηκαν, εκπαιδεύθηκαν και αξιολογήθηκαν στο πειραματικό μέρος της παρούσας διπλωματικής. Τέλος, αναλύονται οι διάφορες τεχνικές κανονικοποίησης (regularization techniques) που χρησιμοποιούνται κατά τη διεξαγωγή του πειραματικού μέρους της εργασίας, καθώς και η τεχνική της μεταφοράς μάθησης (transfer learning).

Στο Κεφάλαιο 3 περιγράφεται η μεθοδολογία που ακολουθήθηκε και παρουσιάζονται και σχολιάζονται τα αποτελέσματα του πρώτου πειράματος που εκπονήθηκε και το οποίο πραγματεύεται τη ταξινόμηση τομών αξονικών τομογραφιών, ενώ στο Κεφάλαιο 4 περιγράφεται η μεθοδολογία που ακολουθήθηκε και παρουσιάζονται και σχολιάζονται τα αποτελέσματα του δεύτερου πειράματος που εκπονήθηκε και το οποίο πραγματεύεται τη σημασιολογική κατάτμηση αλλοιώσεων

σε τομές αξονικών τομογραφιών. Πιο αναλυτικά, όσον αφορά το κομμάτι της μεθοδολογίας, αρχικά παρουσιάζονται τα σύνολα δεδομένων τα οποία αξιοποιήθηκαν για την πραγματοποίηση των πειραμάτων. Στη συνέχεια, αναλύονται οι τεχνικές προεπεξεργασίας που εφαρμόστηκαν στα δεδομένα των συγκεκριμένων συνόλων πριν αυτά τροφοδοτηθούν στα μοντέλα βαθειάς μάθησης. Έπειτα, περιγράφονται τα μοντέλα τεχνητών νευρωνικών δικτύων που επιλέχθηκαν να χρησιμοποιηθούν στη συγκεκριμένη εργασία. Επίσης, παρουσιάζεται μία σύνοψη των διαφόρων μετρικών αξιολόγησης που επιλέχθηκαν να αξιοποιηθούν ούτως ώστε να πραγματοποιηθεί η αξιολόγηση της αποδοτικότητας των μοντέλων βαθειάς μάθησης. Όσον αφορά το πειραματικό μέρος, αναφέρονται και σχολιάζονται τα αποτελέσματα της μεθοδολογίας που ακολουθήθηκε. Τέλος, γίνεται σύγκριση και σχολιασμός της αποδοτικότητας των μοντέλων βαθειάς μάθησης που αξιοποιήθηκαν.

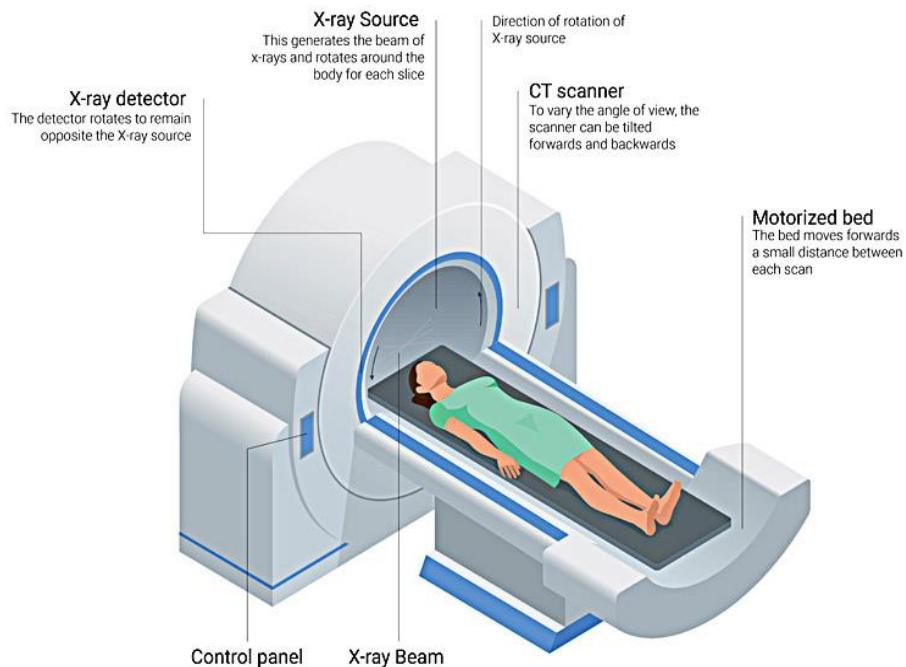
Στο Κεφάλαιο 5 παρουσιάζονται τα συμπεράσματα που αποκομίσθηκαν έπειτα από τη μελέτη που πραγματοποιήθηκε στα πλαίσια της συγκεκριμένης εργασίας. Τέλος, αναφέρονται τυχόν τροποποιήσεις και προσθήκες, που θα μπορούσαν να συμπεριληφθούν μελλοντικά ούτως ώστε να επεκταθεί περεταίρω η έρευνα πάνω στο ζήτημα που διερευνάται στα πλαίσια εκπόνησης της διπλωματικής εργασίας.

## 2 Θεωρητικό Υπόβαθρο

### 2.1 Αξονική Τομογραφία (Computed Tomography)

Η αξονική τομογραφία αποτελεί μέθοδο ιατρικής απεικόνισης η οποία χρησιμοποιείται για την υποβοήθηση στη διάγνωση ιατρικών παθήσεων ασθενών [13]. Στην Εικόνα 2.1 παρουσιάζονται τα μέρη ενός αξονικού τομογράφου. Η λήψη μιας αξονικής τομογραφίας, πραγματοποιείται με τον ασθενή τοποθετημένο στο κρεβάτι του αξονικού τομογράφου [14]. Ο αξονικός τομογράφος διαθέτει έναν σαρωτή ο οποίος αποτελείται από μία γεννήτρια ακτινών X και ένα σωλήνα που κατά τη διάρκεια λήψης της αξονικής τομογραφίας περιστρέφεται γύρω από τον ασθενή, εκπέμποντας μια σειρά από ακτίνες X οι οποίες τον διαπερνούν από διαφορετικές γωνίες [13], [14]. Κάποια από τα φωτόνια των ακτινών απορροφώνται από το σώμα του ασθενή, ενώ κάποια άλλα εξέρχονται και καταγράφονται από ανιχνευτές που βρίσκονται στην αντίθετη πλευρά του σαρωτή [13], [14].

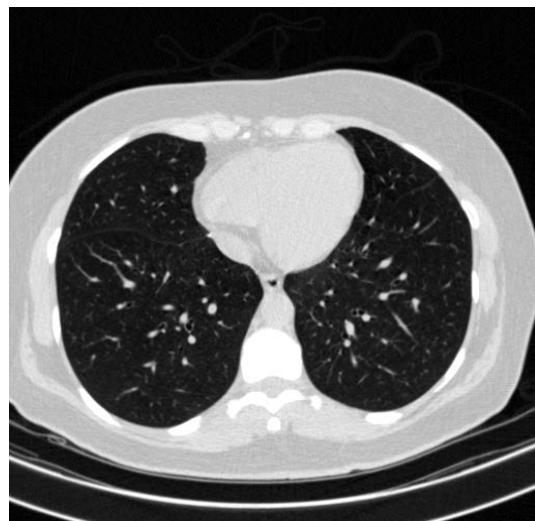
#### Computerized Tomography (CT) Scan



Εικόνα 2.1: Αξονικός τομογράφος [15]

Οι ανιχνευτές υπολογίζουν τις τιμές εξασθένισης των φωτονίων και τις μετατρέπουν σε ψηφιακά δεδομένα που αποστέλλονται σε έναν υπολογιστή [14]. Ο

υπολογιστής επεξεργάζεται τα δεδομένα και δημιουργεί δυσδιάστατες εικόνες που αναπαριστούν μια περιοχή του σώματος του ασθενούς και ονομάζονται τομές αξονικής τομογραφίας (CT slices) [13]. Στην Εικόνα 2.2, παρουσιάζεται μία τομή αξονικής τομογραφίας. Συνεχόμενες τομές μιας αξονικής τομογραφίας μπορούν να συνδυαστούν μεταξύ τους δημιουργώντας μια τρισδιάστατη απεικόνιση μιας δομής στο εσωτερικό του σώματος ενός ασθενούς [13].



Εικόνα 2.2: Τομή αξονικής τομογραφίας [16]

Οι αξονικές τομογραφίες είναι ιδιαίτερα χρήσιμες για τη λήψη λεπτομερών εικόνων οστών, μαλακών ιστών και αιμοφόρων αγγείων. Παρέχουν πιο λεπτομερείς πληροφορίες από τις συμβατικές ακτινογραφίες και μπορούν να ανιχνεύσουν ένα ευρύ φάσμα παθήσεων, όπως όγκους, εσωτερικούς τραυματισμούς, λοιμώξεις και ανωμαλίες στα όργανα.

Οι αξονικές τομογραφίες διαδραματίζουν σημαντικό ρόλο στην αξιολόγηση και τη διάγνωση πνευμονικών παθήσεων. Πιο αναλυτικά, παρέχουν μια λεπτομερή εικόνα των πνευμόνων, επιτρέποντας στους επαγγελματίες υγείας να αξιολογήσουν τη δομή του πνεύμονα και να ανιχνεύσουν τυχόν ανωμαλίες και αλλοιώσεις του πνευμονικού παρεγχύματος.

Χρησιμοποιούνται για την αξιολόγηση και διάγνωση μιας πληθώρας πνευμονικών παθήσεων, συμπεριλαμβανομένων του καρκίνου του πνεύμονα, πνευμονικών λοιμώξεων, διάμεσων πνευμονοπαθειών, πνευμονικής εμβολής, καθώς και της χρόνιας αποφρακτικής πνευμονοπάθειας (ΧΑΠ).

### 2.1.1 Παραθυροποίηση Εικόνας Αξονικής Τομογραφίας (CT Image Windowing)

Τα εικονοστοιχεία μιας εικόνας αξονικής τομογραφίας αντιστοιχίζονται με αριθμητικές τιμές που αποκαλούνται Hounsfield Units (HU) και είναι εκφρασμένες σε μία κλίμακα που ονομάζεται κλίμακα Hounsfield [17]. Η κλίμακα αυτή αποτελεί μία μέθοδο ποσοτικοποίησης της εξασθένισης των ακτινών X καθώς αυτές διέρχονται από μία δομή στο εσωτερικό του σώματος. Οι τιμές της κλίμακας Hounsfield αντιπροσωπεύουν τη ραδιοπυκνότητα των ιστών και άλλων δομών που βρίσκονται εντός του σώματος και αναπαριστούν τις αποχρώσεις του γκρι στις εικόνες αξονικών τομογραφιών [13], [17].

Στην κλίμακα Hounsfield η τιμή 0 αντιστοιχεί στην εξασθένιση του νερού και η τιμή -1000 στην εξασθένιση του αέρα [17]. Οι αρνητικές τιμές αποδίδονται σε ιστούς που έχουν χαμηλότερη εξασθένιση ακτινών X από το νερό, όπως δομές γεμάτες με αέρα, ενώ θετικές τιμές αποδίδονται σε ιστούς με υψηλότερη εξασθένιση ακτινών X από το νερό, όπως τα οστά [17]. Στον Πίνακα 2.1 αναφέρονται ορισμένες γενικές τιμές Hounsfield που αφορούν συνήθεις ιστούς.

Πίνακας 2.1: Τιμές Hounsfield που αφορούν συνήθεις ιστούς [18]

Τύπος Ιστού	Hounsfield Units
Αέρας (Air)	-1000
Πνεύμονας (Lung)	-700 με -500
Μαλακός Ιστός (Soft Tissue)	-300 με -100
Λίπος (Fat)	-100 με -50
Νερό (Water)	0
Εγκεφαλονωτιαίο Υγρό (CSF)	+15
Αίμα (Blood)	+30 με +45
Μυς (Muscle)	+10 με +40
Ήπαρ (Liver)	+40 με +60
Οστό (Bone)	+700 με +3000

Μεταβάλλοντας κανείς τις τιμές Hounsfield μιας εικόνας αξονικής τομογραφίας μπορεί να ρυθμίσει την αντίθεση (contrast) και τη φωτεινότητα (brightness) της συγκεκριμένης εικόνας [13], [14]. Η ρύθμιση των τιμών Hounsfield μπορεί να πραγματοποιηθεί μέσω της μεταβολής δύο παραμέτρων που ονομάζονται εύρος παραθύρου (window width) και επίπεδο παραθύρου (window level) [13], [14]. Η λειτουργία των δύο αυτών παραμέτρων είναι η εξής:

Το εύρος παραθύρου ορίζει το εύρος των τιμών Hounsfield, που εμφανίζονται ως διακριτές αποχρώσεις του γκρι [13], [14]. Το εύρος παραθύρου ρυθμίζει την αντίθεση μιας εικόνας αξονικής τομογραφίας [13], [14].

Το επίπεδο παραθύρου αντιπροσωπεύει την τιμή Hounsfield στο κέντρο του εύρους παραθύρου, δηλαδή την κεντρική τιμή στο εύρος των τιμών Hounsfield [13], [14]. Το επίπεδο παραθύρου ρυθμίζει τη φωτεινότητα μιας εικόνας αξονικής τομογραφίας [13], [14].

Η διαδικασία μεταβολής του εύρους παραθύρου και του επιπέδου παραθύρου ονομάζεται παραθυροποίηση (windowing). Η μεταβολή του εύρους των τόνων του γκρι στις εικόνες αξονικών τομογραφιών επιτρέπει την οπτικοποίηση συγκεκριμένων ιστών ή δομών, που μας ενδιαφέρουν, με μεγαλύτερη σαφήνεια [14].

Στον Πίνακα 2.2 παρουσιάζονται συχνές ρυθμίσεις παραθύρου που χρησιμοποιούνται στην απεικόνιση εικόνων αξονικών τομογραφιών.

**Πίνακας 2.2: Ρυθμίσεις παραθύρου που χρησιμοποιούνται στην απεικόνιση εικόνων αξονικών τομογραφιών [18]**

Εξέταση	Ανατομία	Εύρος Παραθύρου	Επίπεδο Παραθύρου
Θώρακος (Chest)	Πνεύμονας (Lung)	1500	-600
	Μεσοθωράκιο (Mediastinum)	350	50
Κοιλίας (Abdomen)	Μαλακός Ιστός (Soft Tissue)	400	50
	Ηπαρ (Liver)	150	30
Σπονδυλικής Στήλης (Spine)	Μαλακός Ιστός (Soft Tissue)	250	50
	Οστό (Bone)	1800	400

## 2.2 Όραση Υπολογιστών (Computer Vision)

Η όραση υπολογιστών αποτελεί πεδίο της τεχνητής νοημοσύνης, το οποίο παρέχει μία πληθώρα αλγορίθμων και τεχνικών μέσω των οποίων υπολογιστικά συστήματα αποκτούν την ικανότητα να αντιλαμβάνονται και να κατανοούν τον κόσμο μέσω οπτικών δεδομένων (εικόνες, βίντεο) [19]. Αξιοποιώντας την αντίληψη και την

κατανόηση του κόσμου που έχουν αποκτήσει, τα υπολογιστικά συστήματα μπορούν να λαμβάνουν κατάλληλες αποφάσεις [19].

Στόχος του συγκεκριμένου τομέα είναι να δώσει τη δυνατότητα στους ηλεκτρονικούς υπολογιστές να μπορούν να εξάγουν πληροφορίες από οπτικά δεδομένα με τρόπο παρόμοιο με αυτό που οι άνθρωποι αντιλαμβάνονται και κατανοούν τα οπτικά ερεθίσματα που λαμβάνουν από το περιβάλλον τους [19].

Οι τεχνικές όρασης υπολογιστών εφαρμόζονται σε μία πληθώρα προβλημάτων όπως είναι η ταξινόμηση εικόνων (image classification), η ανίχνευση και ο εντοπισμός αντικειμένων (object detection and localization), η αυτοματοποιημένη εισαγωγή λεζάντας σε εικόνες (automatic image captioning), η μεταφορά στυλ (style transfer), καθώς και η σημασιολογική κατάτμηση εικόνων (semantic image segmentation) [19]. Οι δύο κατηγορίες προβλημάτων στις οποίες υπάγεται το ζήτημα της ανάλυσης ιατρικών εικόνων για την υποβοήθηση στη διάγνωση της COVID-19, που πραγματεύεται η παρούσα διπλωματική εργασία, είναι η ταξινόμηση εικόνων και η σημασιολογική κατάτμηση εικόνων.

### 2.2.1 Ταξινόμηση Εικόνων (Image Classification)

Η ταξινόμηση εικόνων αποτελεί μια κατηγορία προβλημάτων για την επίλυση των οποίων μπορούν να εφαρμοσθούν τεχνικές όρασης υπολογιστών. Στο συγκεκριμένο τύπο προβλημάτων, πραγματοποιείται αντιστοίχιση εικόνων σε κατηγορίες με βάση το περιεχόμενό τους, χρησιμοποιώντας έναν αλγόριθμο όρασης υπολογιστών [19].

Η πιο διαδεδομένη τεχνική ταξινόμησης εικόνων είναι τα τεχνητά νευρωνικά δίκτυα (ANNs) και πιο συγκεκριμένα τα συνελικτικά νευρωνικά δίκτυα (CNNs) [19].

Το ζήτημα της ανάλυσης ιατρικών εικόνων για την υποβοήθηση στη διάγνωση πνευμονίας από COVID-19 μπορεί να θεωρηθεί ως ένα πρόβλημα ταξινόμησης ιατρικών εικόνων. Όσον αφορά την παρούσα διπλωματική, εκπαιδεύονται συνελικτικά νευρωνικά δίκτυα, λαμβάνοντας ως είσοδο τομές αξονικών τομογραφιών, ούτως ώστε να είναι σε θέση να πραγματοποιούν ταξινόμησή τομών σε κατηγορίες με βάση την ύπαρξη ή μη αλλοιώσεων από πνευμονία COVID-19.

### 2.2.2 Σημασιολογική Κατάτμηση Εικόνων (Semantic Image Segmentation)

Η σημασιολογική κατάτμηση εικόνων αποτελεί μια ακόμη κατηγορία προβλημάτων στα οποία δύναται να εφαρμοσθούν τεχνικές όρασης υπολογιστών. Στη συγκεκριμένη κατηγορία πραγματοποιείται απομόνωση των περιοχών μιας εικόνας

μέσω της αντιστοίχισης των εικονοστοιχείων της σε κατηγορίες σύμφωνα με χαρακτηριστικά που εμφανίζει [20].

Τα συνελικτικά νευρωνικά δίκτυα αποτελούν την πιο δημοφιλή επιλογή για εφαρμογή σε προβλήματα σημασιολογικής κατάτμησης εικόνων.

Το ζήτημα της ανάλυσης ιατρικών εικόνων για την υποβοήθηση στη διάγνωση πνευμονίας από COVID-19 μπορεί να θεωρηθεί ως ένα πρόβλημα σημασιολογικής κατάτμησης ιατρικών εικόνων. Όσον αφορά την παρούσα διπλωματική, συνελικτικά νευρωνικά δίκτυα τροφοδοτούνται με τομές αξονικών τομογραφιών και εικόνες στις οποίες έχουν επισημανθεί από ειδικούς οι περιοχές των τομών που εμφανίζουν αλλοιώσεις. Τα δίκτυα στη συνέχεια εκπαιδεύονται, χρησιμοποιώντας τα δεδομένα εισόδου, ούτως ώστε να είναι σε θέση να πραγματοποιούν αντιστοίχιση των εικονοστοιχείων των εικόνων σε κατηγορίες με βάση την ύπαρξη ή μη αλλοιώσεων από πνευμονία COVID-19.

## 2.3 Μηχανική Μάθηση (Machine Learning)

Η μηχανική μάθηση σύμφωνα με τον Arthur Samuel ο οποίος έδωσε έναν πρώτο ορισμό της το 1959, αποτελεί το «ερευνητικό πεδίο που επιτρέπει στους ηλεκτρονικούς υπολογιστές να μαθαίνουν χωρίς να έχουν προγραμματισθεί ρητά» [21]. Ένας ακόμη ορισμός δόθηκε το 1997 από τον ερευνητή Tom Mitchell, ο οποίος αναφέρει πως «Ένα πρόγραμμα υπολογιστή λέγεται ότι μαθαίνει από εμπειρία E ως προς μια κατηγορία εργασιών T και μια μετρική επίδοσης P, εάν η επίδοση του σε εργασίες της κατηγορίας T, όπως αποτιμάται από τη μετρική P, βελτιώνεται με την εμπειρία E» [22].

Το πεδίο της μηχανικής μάθησης αποτελείται από ένα σύνολο αλγορίθμων, οι οποίοι παρέχουν τη δυνατότητα σε υπολογιστικά συστήματα να εκπαιδεύονται και να αποκτούν εμπειρία, μέσω της οποίας αυξάνουν την αποδοτικότητά τους στην επίλυση ενός προβλήματος [23]. Ένα βασικό πλεονέκτημα του πεδίου της μηχανικής μάθησης αποτελεί το γεγονός ότι μέσω αυτής καθίσταται δυνατή η επίλυση προβλημάτων, που είναι πολύ δύσκολο να επιτευχθεί μέσω συμβατικών μεθόδων προγραμματισμού [23]. Οι τέσσερις βασικές κατηγορίες μηχανικής μάθησης είναι η επιβλεπόμενη μάθηση, η μη επιβλεπόμενη μάθηση, η ημι-επιβλεπόμενη μάθηση, και η ενισχυτική μάθηση.

### 2.3.1 Επιβλεπόμενη Μάθηση (Supervised Learning)

Επιβλεπόμενη μάθηση ονομάζεται η κατηγορία μηχανικής μάθησης όπου ένας αλγόριθμος εκπαιδεύεται πάνω σε ένα σύνολο δεδομένων που ονομάζεται σύνολο

εκπαίδευσης (training set) και το οποίο αποτελείται από αντιστοιχίσεις εισόδων με εξόδους, οι οποίες είναι γνωστές εκ των προτέρων [24], [25].

Στόχος της επιβλεπόμενης μάθησης είναι ο αλγόριθμος μηχανικής μάθησης, μέσω της εκπαίδευσης, να είναι σε θέση να υλοποιήσει μία μαθηματική συνάρτηση που να αντιστοιχίζει με τον καλύτερο δυνατό τρόπο τις εισόδους του συνόλου εκπαίδευσης με τις επιθυμητές εξόδους [24], [25]. Απότερος σκοπός, η συνάρτηση να μπορεί να εφαρμοστεί και σε ένα άλλο σύνολο που ονομάζεται σύνολο ελέγχου (test set) και το οποίο αποτελείται από νέες εισόδους των οποίων οι αντιστοιχίσεις με εξόδους δεν είναι γνωστές [24], [25]. Η συγκεκριμένη κατηγορία έχει λάβει αυτή την ονομασία διότι οι αλγόριθμοι που ανήκουν σε αυτή απαιτούν εξωτερική επίβλεψη [25].

Η συγκεκριμένη κατηγορία μηχανικής μάθησης χρησιμοποιείται κυρίως σε προβλήματα ταξινόμησης (classification) και παλινδρόμησης (regression) [24].

### 2.3.2 Μη Επιβλεπόμενη Μάθηση (Unsupervised Learning)

Μη επιβλεπόμενη μάθηση ονομάζεται η κατηγορία μηχανικής μάθησης στην οποία ένας αλγόριθμος εκπαιδεύεται πάνω σε δεδομένα εισόδου των οποίων οι επιθυμητές έξοδοι δεν είναι γνωστές [24].

Στόχος της μη επιβλεπόμενης μάθησης είναι ο αλγόριθμος να καταφέρει να ανακαλύψει άγνωστα μέχρι τότε μοτίβα δεδομένων, με απότερο σκοπό να κατανοήσει καλύτερα τα δεδομένα και να εξάγει πληροφορίες από αυτά [24]. Η συγκεκριμένη κατηγορία έχει λάβει αυτήν την ονομασία διότι οι αλγόριθμοι που ανήκουν σε αυτή δεν λαμβάνουν εξωτερική επίβλεψη [24], [25].

Η συγκεκριμένη κατηγορία μηχανικής μάθησης χρησιμοποιείται κυρίως σε προβλήματα ομαδοποίησης (clustering) και συσχέτισης (association) [24].

### 2.3.3 Ενισχυτική Μάθηση (Reinforcement Learning)

Ενισχυτική μάθηση ονομάζεται η κατηγορία μηχανικής μάθησης η οποία πραγματοποιεί την εκπαίδευση ενός ευφυούς συστήματος, μέσω της επιβράβευσης της επιθυμητής συμπεριφοράς και αντίστοιχα μέσω της επιβολής κυρώσεων στη περίπτωση της μη επιθυμητής συμπεριφοράς [24].

Πιο αναλυτικά, το σύστημα παρατηρεί και αλληλεπιδρά με το περιβάλλον του, λαμβάνοντας ως είσοδο μια ένδειξη της τρέχουσας κατάστασης του περιβάλλοντος του [26]. Με βάση τη τρέχουσα κατάσταση και το στόχο που ευελπιστεί να επιτύχει, επιλέγει πως θα δράσει, παράγοντας κάποιο αποτέλεσμα [26]. Ο τρόπος δράσης του επηρεάζει τη κατάσταση του περιβάλλοντός του και στη συνέχεια το σύστημα

παρατηρώντας την επίδραση στο περιβάλλον, επιλέγει εάν θα επιβραβεύσει ή θα επιβάλλει κυρώσεις στη συγκεκριμένη δράση [26]. Με τη πάροδο της εκπαίδευσης του, το σύστημα μαθαίνει να αναγνωρίζει και να επιλέγει τις δράσεις εκείνες οι οποίες επιφέρουν τη μέγιστη επιβράβευση [26].

### 2.3.4 Ημι-επιβλεπόμενη Μάθηση (Semi-supervised Learning)

Η ημι-επιβλεπόμενη μάθηση αποτελεί έναν συνδυασμό της επιβλεπόμενης και της μη επιβλεπόμενης μάθησης, [25], [27]. Πιο συγκεκριμένα, στην ημι-επιβλεπόμενη μάθηση ένας αλγόριθμος εκπαιδεύεται πάνω σε ένα σύνολο δεδομένων, το οποίο αποτελείται τόσο από αντιστοιχίσεις εισόδων με εξόδους, που είναι γνωστές εκ των προτέρων, καθώς και από ένα μεγάλο αριθμό δεδομένων εισόδου των οποίων οι αντιστοιχίσεις με επιθυμητές εξόδους δεν είναι γνωστές [24], [27]. Το πλεονέκτημα που προσφέρει η συγκεκριμένη κατηγορία μηχανικής μάθησης είναι το γεγονός ότι μπορεί να αποτελέσει λύση στις ακόλουθες περιπτώσεις:

Όταν διαθέτουμε δεδομένα εισόδου των οποίων οι επιθυμητές έξοδοι δεν είναι γνωστές. Ο αριθμός των αντιστοιχίσεων εισόδων με γνωστές εξόδους του συνόλου δεδομένων δεν είναι επαρκής ούτως ώστε να εκπαιδευθεί ένας αλγόριθμος αξιοποιώντας τη μέθοδο της επιβλεπόμενης μάθησης. Η απόκτηση επιπλέον δεδομένων εισόδου των οποίων οι επιθυμητές έξοδοι είναι γνωστές εκ των προτέρων δεν είναι εφικτή [25].

## 2.4 Διαδικασία Εκπαίδευσης Αλγορίθμων Μηχανικής Μάθησης

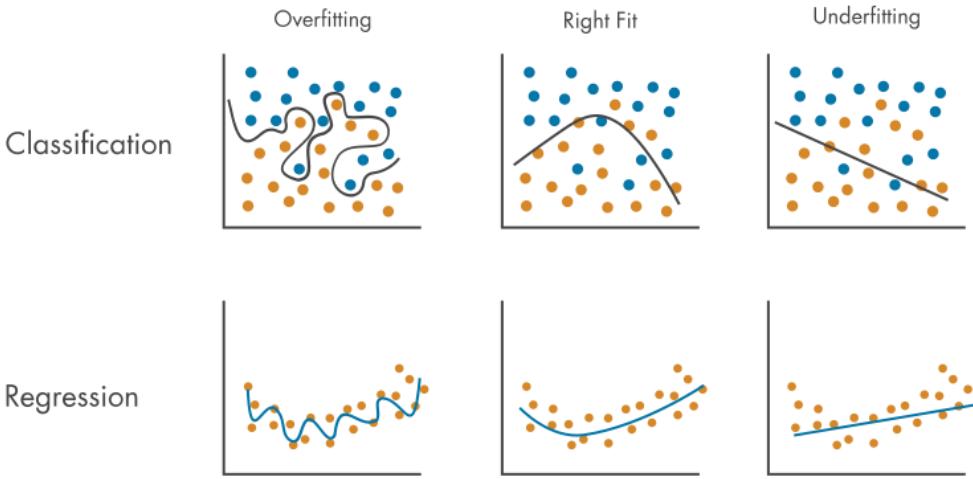
Μια διαδικασία μηχανικής μάθησης στοχεύει στη δημιουργία υπολογιστικών συστημάτων τα οποία στη συνέχεια με τη βοήθεια τεχνικών της εκπαίδευσης πραγματοποιώντας προβλέψεις πάνω σε ένα σύνολο δεδομένων που τους παρέχεται ως είσοδος και που ονομάζεται σύνολο εκπαίδευσης (training set). Η εκπαίδευση αποτελεί εξέχουσα διαδικασία ενός αλγορίθμου μηχανικής μάθησης. Κατά τη διάρκειά της, πραγματοποιείται υπολογισμός μιας μετρικής που ονομάζεται σφάλμα εκπαίδευσης (training error) [23]. Το σφάλμα εκπαίδευσης δείχνει πόσο λάθος κάνει ένα μοντέλο μηχανικής μάθησης ως προς τις προβλέψεις του και εκφράζει την απόσταση ανάμεσα στην προβλεπόμενη και στην αναμενόμενη έξοδο, όταν του δίνεται ως είσοδος το σύνολο εκπαίδευσης.

Εκτός από το σύνολο εκπαίδευσης, απαραίτητη είναι και η ύπαρξη του συνόλου επικύρωσης (validation set). Με αυτό ρυθμίζονται οι παράμετροι του εκάστοτε μοντέλου μηχανικής μάθησης (tuning) λαμβάνοντας υπόψη την επίδοσή του στο συγκεκριμένο σύνολο δεδομένων [28], [29]. Επίσης, το σύνολο επικύρωσης παρέχει

μια αξιολόγηση της επίδοσης του μοντέλου. Όπως και στο σύνολο εκπαίδευσης, έτσι και στο σύνολο επικύρωσης υπολογίζεται μια μετρική που ονομάζεται σφάλμα επικύρωσης (validation error) [23]. Αντίστοιχα με το σφάλμα εκπαίδευσης, το σφάλμα επικύρωσης εκφράζει την απόσταση ανάμεσα στην προβλεπόμενη έξοδο του μοντέλου μηχανικής μάθησης και στη πραγματική έξοδο του, όταν του δίνεται ως είσοδος το σύνολο επικύρωσης. Απώτερος σκοπός ενός μοντέλου μηχανικής μάθησης είναι η όσο το δυνατόν μεγαλύτερη μείωση των σφαλμάτων εκπαίδευσης και επικύρωσης [23].

Στόχος των ερευνητών που ασχολούνται με την υλοποίηση μοντέλων μηχανικής μάθησης, είναι να δημιουργήσουν μοντέλα τα οποία διαθέτουν ικανότητα γενίκευσης σε νέα δεδομένα. Ο όρος γενίκευση αναφέρεται στην ικανότητα ενός μοντέλου μηχανικής μάθησης να εκπαιδεύεται σε τέτοιο βαθμό ούτως ώστε στη συνέχεια να είναι σε θέση να πραγματοποιεί τη λειτουργία για την οποία έχει εκπαιδευθεί με αποδοτικό τρόπο σε νέα και άγνωστα σε αυτό δεδομένα.

Οι παράγοντες που καθορίζουν εάν ένα μοντέλο μηχανικής μάθησης έχει εκπαιδευθεί σωστά είναι η τιμή του σφάλματος εκπαίδευσης έπειτα από κάθε κύκλο εκπαίδευσης ή εποχή (epoch) και η απόσταση ανάμεσα στο σφάλμα εκπαίδευσης και στο σφάλμα επικύρωσης [23]. Όσο πιο μικρό είναι το σφάλμα εκπαίδευσης και όσο μικρότερη είναι η απόσταση ανάμεσα στο σφάλμα εκπαίδευσης και επικύρωσης, τόσο καλύτερη είναι η εκπαίδευση του μοντέλου. Εάν ένα μοντέλο αδυνατεί να μειώσει το σφάλμα εκπαίδευσης, που σημαίνει ότι αδυνατεί να αναγνωρίσει τις σχέσεις και τα μοτίβα που υπάρχουν μεταξύ των δεδομένων εκπαίδευσης, το μοντέλο λέγεται ότι υποπροσαρμόζεται στα δεδομένα [23]. Εάν η απόσταση ανάμεσα στο σφάλμα εκπαίδευσης και στο σφάλμα επικύρωσης είναι πολύ μεγάλη, αυτό σημαίνει ότι το μοντέλο έχει εκπαιδευθεί σε τέτοιο βαθμό στα δεδομένα εκπαίδευσης που αδυνατεί να γενικεύσει σε νέα, άγνωστα σε αυτό δεδομένα [23]. Σε αυτή την περίπτωση λέγεται ότι το μοντέλο υπερπροσαρμόζεται στα δεδομένα [23]. Οι δύο αυτές καταστάσεις, ονομάζονται υποπροσαρμογή (underfitting) και υπερπροσαρμογή (overfitting) αντίστοιχα. Στην Εικόνα 2.3, παρουσιάζονται περιπτώσεις υποπροσαρμογής και υπερπροσαρμογής σε ένα πρόβλημα ταξινόμησης και σε ένα πρόβλημα παλινδρόμησης.



**Εικόνα 2.3: Περιπτώσεις υποπροσαρμογής και υπερπροσαρμογής σε πρόβλημα ταξινόμησης (classification) και παλινδρόμησης (regression) [30]**

Το γεγονός ότι το σύνολο επικύρωσης χρησιμοποιείται για τη ρύθμιση των παραμέτρων των μοντέλων το καθιστά ακατάλληλο για την αξιολόγηση της ικανότητας γενίκευσης ενός μοντέλου. Αυτό, γιατί κατά τη διάρκεια της ρύθμισης των παραμέτρων το μοντέλο ενδεχομένως να έχει μάθει ακούσια μοτίβα και αναπαραστάσεις που εμφανίζονται αποκλειστικά στο σύνολο επικύρωσης. Επίσης, τα σύνολα εκπαίδευσης και επικύρωσης συνήθως προκύπτουν από το διαχωρισμό ενός ενιαίου συνόλου δεδομένων με αποτέλεσμα αυτά να εμφανίζουν κοινά χαρακτηριστικά μεταξύ τους [29]. Η αξιολόγηση ενός μοντέλου στα δεδομένα επικύρωσης, παρέχει μια εικόνα της ικανότητας γενίκευσης του μοντέλου σε νέα δεδομένα τα οποία ωστόσο είναι παρόμοια με τα δεδομένα του συνόλου εκπαίδευσης [29]. Αποτελεί λοιπόν καλή πρακτική να χρησιμοποιείται ένα τρίτο σύνολο δεδομένων το οποίο προέρχεται από διαφορετικό αρχικό σύνολο και το οποίο δεν έχει χρησιμοποιηθεί κατά τη διάρκεια της εκπαίδευτικής διαδικασίας. Το συγκεκριμένο σύνολο ονομάζεται σύνολο ελέγχου (test set) και χρησιμοποιείται για την τελική αξιολόγηση της ικανότητας γενίκευσης του μοντέλου. Όπως και στα σύνολα εκπαίδευσης και επικύρωσης, έτσι και στο σύνολο ελέγχου υπολογίζεται μια μετρική που ονομάζεται σφάλμα ελέγχου (test error) [23]. Το σφάλμα στο σύνολο δεδομένων ελέγχου εκφράζει την απόσταση ανάμεσα στην προβλεπόμενη έξοδο του μοντέλου μηχανικής μάθησης και στην πραγματική έξοδό του, όταν του δίνεται ως είσοδος το σύνολο ελέγχου.

Η γενίκευση αποτελεί πολυπόθητο χαρακτηριστικό για ένα μοντέλο μηχανικής μάθησης αφού μοντέλα τα οποία έχουν επιτύχει να γενικεύουν σε ικανοποιητικό βαθμό συνήθως κρίνονται κατάλληλα ούτως ώστε να εφαρμοσθούν σε πραγματικές συνθήκες.

#### 2.4.1 Υποπροσαρμογή (Underfitting)

Η υποπροσαρμογή αποτελεί την κατάσταση στην οποία βρίσκεται ένα εκπαιδευμένο μοντέλο μηχανικής μάθησης το οποίο αδυνατεί να μειώσει το σφάλμα εκπαίδευσης [23]. Βασική αιτία αυτής αποτελεί το γεγονός ότι το συγκεκριμένο μοντέλο ενδεχομένως να μην είναι αρκετά σύνθετο ούτως ώστε να εξάγει τα χαρακτηριστικά που θα το βιοθήσουν να ανακαλύψει τις σχέσεις και τα μοτίβα που εμφανίζονται μεταξύ των δεδομένων [19]. Ένα τέτοιο μοντέλο παρουσιάζει ανεπαρκή επίδοση τόσο στα δεδομένα εκπαίδευσης, όσο και σε νέα, άγνωστα σε αυτό δεδομένα.

#### 2.4.2 Υπερπροσαρμογή (Overfitting)

Υπερπροσαρμογή ονομάζουμε τη κατάσταση όπου το σφάλμα εκπαίδευσης και επικύρωσης ενός μοντέλου παρουσιάζουν μεγάλη απόσταση μεταξύ τους [23]. Στη συγκεκριμένη περίπτωση το μοντέλο μηχανικής μάθησης έχει εκπαιδευθεί σε υπερβολικό βαθμό με αποτέλεσμα να έχει «αποστηθίσει» τα δεδομένα εκπαίδευσης [19]. Ένα μοντέλο που εμφανίζει υπερπροσαρμογή, αδυνατεί να αναγνωρίσει τις σχέσεις και τα μοτίβα που εμφανίζονται μεταξύ των δεδομένων εκπαίδευσης, ούτως ώστε να πραγματοποιήσει προβλέψεις. Αντιθέτως, βασίζεται σε λανθασμένες σχέσεις και μοτίβα [19]. Βασική αιτία που ένα μοντέλο έχει περιέλθει σε κατάσταση υπερπροσαρμογής αποτελεί ενδεχομένως το γεγονός ότι το μοντέλο είναι πολύπλοκο για την επίλυση του προβλήματος για το οποίο έχει επιστρατευθεί [19]. Επίσης, πιθανόν τα δεδομένα να μην είναι αρκετά ή αντιπροσωπευτικά για την επίλυση του προβλήματος. Ένα μοντέλο που εμφανίζει υπερπροσαρμογή αποδίδει υπερβολικά καλά στα δεδομένα στα οποία εκπαιδεύεται και είναι ανεπαρκές όσον αφορά νέα και άγνωστα σε αυτό δεδομένα [19]. Ένας τρόπος εντοπισμού της ύπαρξης υπερπροσαρμογής αποτελεί η χρήση μιας τεχνικής η οποία ονομάζεται διασταυρούμενη επικύρωση k-πτυχών (k-fold cross validation).

### 2.5 Τεχνικές Διαχωρισμού Δεδομένων

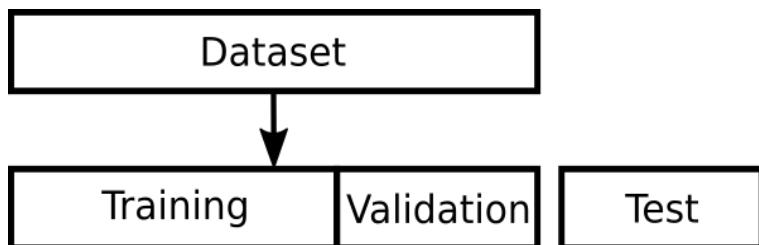
Ο διαχωρισμός των δεδομένων σε υποσύνολα εκπαίδευσης, επικύρωσης και ελέγχου είναι ιδιαίτερα σημαντικός γιατί:

Επιτρέπει να πραγματοποιηθεί ρύθμιση (tuning) μιας μεθόδου μηχανικής μάθησης. Αξιολόγηση της αποδοτικότητάς της. Ανίχνευση τυχόν υπερπροσαρμογής της μεθόδου στα δεδομένα εκπαίδευσης, και κατ' επέκταση διερεύνηση της ικανότητας γενίκευσής της σε νέα και άγνωστα δεδομένα.

Υπάρχουν διάφορες τεχνικές διαχωρισμού των δεδομένων. Η βασικότερη και απλούστερη μέθοδος είναι η holdout, ενώ η πιο ευρέως χρησιμοποιούμενη είναι η διασταυρούμενης επικύρωσης k-πτυχών [31].

### 2.5.1 Τεχνική Holdout

Η τεχνική holdout αποτελεί τη βασικότερη και απλούστερη μέθοδο διαχωρισμού των δεδομένων που τροφοδοτούνται ως είσοδος σε έναν αλγόριθμο μηχανικής μάθησης [28]. Στη συγκεκριμένη μέθοδο πραγματοποιείται αρχικά ανακάτεμα με τυχαίο τρόπο του συνόλου δεδομένων, το οποίο στη συνέχεια διαχωρίζεται σε δύο υποσύνολα, το σύνολο εκπαίδευσης και το σύνολο επικύρωσης [28]. Όσον αφορά το σύνολο ελέγχου (test set) αποτελεί καλή πρακτική να επιλέγεται από ένα ξεχωριστό σύνολο δεδομένων [28]. Επειδή ο διαχωρισμός των δεδομένων πραγματοποιείται με τυχαίο τρόπο, η μέθοδος holdout παρουσιάζει το μειονέκτημα ότι σε περιπτώσεις όπου το σύνολο δεδομένων είναι μικρό υπάρχει μεγάλη πιθανότητα το σύνολο επικύρωσης να μην είναι αντιπροσωπευτικό των δεδομένων [28]. Η διαδικασία διαχωρισμού δεδομένων από τη μέθοδο holdout παρουσιάζεται στην Εικόνα 2.4.



Εικόνα 2.4: Μέθοδος holdout

### 2.5.2 Τεχνική Διασταυρούμενης Επικύρωσης K-Πτυχών (K-Fold Cross Validation)

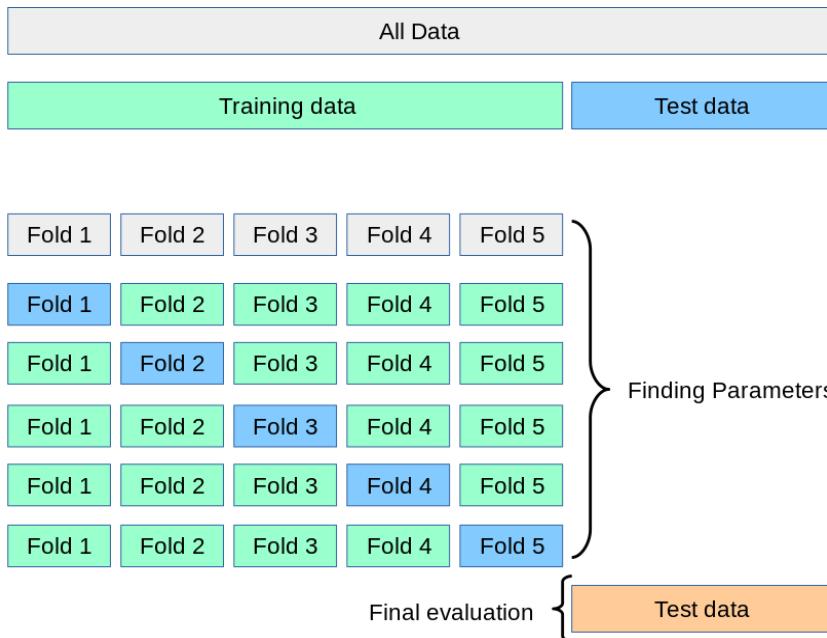
Η μέθοδος διασταυρούμενης επικύρωσης k-πτυχών αποτελεί μια βελτιωμένη εκδοχή της μεθόδου holdout. Πιο συγκεκριμένα στην τεχνική διασταυρούμενης επικύρωσης k-πτυχών, τα δεδομένα διαχωρίζονται με τυχαίο τρόπο σε  $k$  μη επικαλυπτόμενα υποσύνολα [31]. Στη συνέχεια για κάθε ένα από τα  $k$  υποσύνολα, πραγματοποιείται εκπαίδευση του μοντέλου πάνω στα  $k-1$  υποσύνολα του συνόλου δεδομένων και επικύρωσή του στο  $k$ -οστό υποσύνολο [31].

Η προαναφερθείσα διαδικασία επαναλαμβάνεται μέχρις ότου κάθε ένα από τα  $k$  υποσύνολα να έχει αποτελέσει υποσύνολο επικύρωσης [31]. Τέλος για τον

υπολογισμό των μετρικών αξιολόγησης της συνολικής διαδικασίας, υπολογίζεται η μέση τιμή των μετρικών αξιολόγησης κάθε μιας από τις  $k$  επιμέρους διαδικασίες [31].

Η συγκεκριμένη μέθοδος σε αντίθεση με τη μέθοδο holdout διασφαλίζει ότι οποιαδήποτε παρατήρηση από το αρχικό σύνολο δεδομένων έχει την ευκαιρία να εμφανιστεί τόσο στο σύνολο εκπαίδευσης όσο και στο σύνολο επικύρωσης. Επιπλέον, η συγκεκριμένη μέθοδος προσφέρει μια καλύτερη εικόνα της επίδοσης και της ικανότητας γενίκευσης της εκάστοτε τεχνικής μηχανικής μάθησης [31].

Ένα μειονέκτημα της συγκεκριμένης μεθόδου αποτελεί το γεγονός ότι ο αριθμός των φορών που θα πρέπει να πραγματοποιηθεί η διαδικασία εκπαίδευσης αυξάνεται κατά ένα παράγοντα  $k$ , πράγμα που σημαίνει ότι η διαδικασία εκπαίδευσης θα χρειαστεί  $k$  φορές περισσότερο χρόνο για να πραγματοποιηθεί [32]. Η διαδικασία διαχωρισμού δεδομένων μέσω της διασταυρούμενης επικύρωσης  $k$ -πτυχών με  $k=5$  φαίνεται στην Εικόνα 2.5.



Εικόνα 2.5: Μέθοδος διασταυρούμενης επικύρωσης  $k$ -πτυχών με  $k=5$  [33]

## 2.6 Βαθιά Μάθηση (Deep Learning)

Η βαθιά μάθηση σύμφωνα με τους LeCun κ.α. αποτελεί υποπεδίο της μηχανικής μάθησης, το οποίο «επιτρέπει σε υπολογιστικά μοντέλα που αποτελούνται από πολλαπλά επίπεδα επεξεργασίας να μάθουν αναπαραστάσεις δεδομένων με πολλαπλά επίπεδα αφαιρετικότητας» [34].

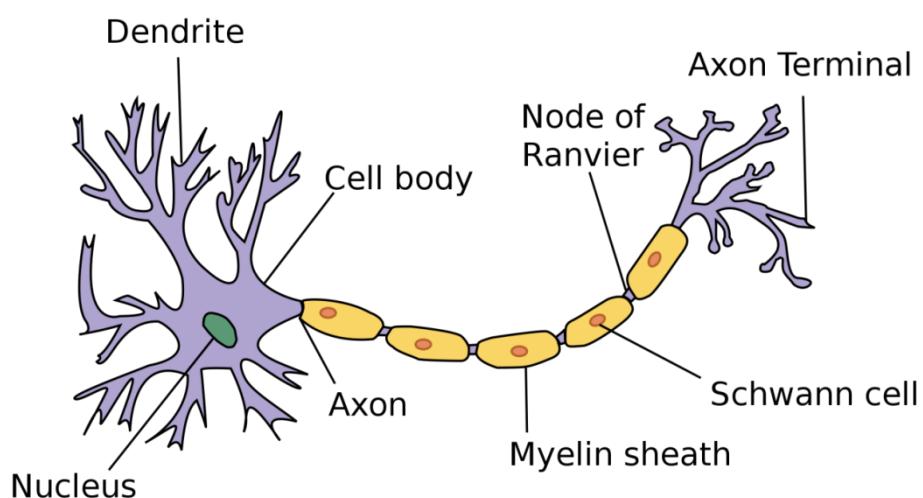
Ο όρος «βαθιά» στην ονομασία του συγκεκριμένου πεδίου έχει να κάνει με το γεγονός ότι τα μοντέλα που κατατάσσονται στο πεδίο αυτό αποτελούν κατά κύριο λόγο πολύπλοκες δομές που αποτελούνται από ένα μεγάλο αριθμό επιπέδων [34].

Τα τεχνητά νευρωνικά δίκτυα αποτελούν τη βασική δομή του συγκεκριμένου πεδίου [23]. Πρόκειται για πολύπλοκα μοντέλα τα οποία μαθαίνουν να αναγνωρίζουν αυτοματοποιημένα μοτίβα και αναπαραστάσεις δεδομένων [34].

Η βαθιά μάθηση έχει εφαρμοσθεί επιτυχώς σε μια πληθώρα από επιστημονικά πεδία όπως είναι η όραση υπολογιστών (computer vision), η επεξεργασία ομιλίας και ήχου (speech and audio processing), η επεξεργασία φυσικής γλώσσας (natural language processing), η ρομποτική (robotics), καθώς και η βιοπληροφορική (bioinformatics) [23].

### 2.6.1 Βιολογικά Νευρωνικά Δίκτυα (Biological Neural Networks)

Τα βιολογικά νευρωνικά δίκτυα αποτελούνται από νευρικά κύτταρα που ονομάζονται νευρώνες (neurons) [35] και τα οποία διασυνδέονται μεταξύ τους σχηματίζοντας ένα δίκτυο, το οποίο ονομάζεται νευρωνικό δίκτυο. Βασική λειτουργία των νευρώνων είναι να παραλαμβάνουν ερεθίσματα και να άγουν τη νευρική ώση [35]. Στην Εικόνα 2.6, παρουσιάζεται η δομή ενός βιολογικού νευρώνα. Παρατηρώντας τη συγκεκριμένη εικόνα μπορεί κανείς να διαπιστώσει πως ένας βιολογικός νευρώνας αποτελείται από τέσσερα βασικά δομικά στοιχεία: το σώμα του νευρικού κυττάρου (cell body), τον άξονα (axon), τους δενδρίτες (dendrites), καθώς και τις συνάψεις (synapses).



Εικόνα 2.6: Δομή ενός βιολογικού νευρώνα [36]

Το κυτταρικό σώμα (cell body) ενός βιολογικού νευρώνα αποτελείται από τη κυτταροπλασματική μάζα εντός της οποίας εντοπίζεται ο πυρήνας (nucleus), ενώ εξωτερικά περιβάλλεται από την κυτταροπλασματική μεμβράνη [35]. Από την επιφάνεια του κυτταρικού σώματος εκφύονται αποφυάδες οι οποίες ονομάζονται νευρίτες (neurites), και οι οποίες διακρίνονται σε δενδρίτες και έναν άξονα [35]. Το κυτταρικό σώμα αποτελεί τον χώρο όπου πραγματοποιείται το μεγαλύτερο μέρος του νευρωνικού «υπολογισμού».

Οι δενδρίτες χρησιμοποιούνται για τη διασύνδεση των νευρικών κυττάρων μεταξύ τους δημιουργώντας έτσι ένα δίκτυο νευρώνων [35]. Βασική λειτουργία των δενδριτών είναι να παραλαμβάνουν πληροφορίες και να τις άγουν προς το κυτταρικό σώμα [35].

Ο άξονας αποτελεί τη μακρύτερη αποφυάδα του κυτταρικού σώματος και η λειτουργία του είναι να απάγει τις ώσεις από το κυτταρικό σώμα [35]. Ο άξονας συνήθως δεν διακλαδίζεται κοντά στο κυτταρικό σώμα, παρουσιάζει όμως παράπλευρους κλάδους κατά την πορεία του [35]. Λίγο πριν από το περιφερικό άκρο του διακλαδίζεται πολύ [35]. Τα τελικά άκρα των τελικών κλάδων παρουσιάζονται διογκωμένα και ονομάζονται τελικά κομβία [35].

Το νευρικό σύστημα αποτελείται από έναν τεράστιο αριθμό νευρώνων οι οποίοι συνδέονται μεταξύ τους με τέτοιο τρόπο, ούτως ώστε να σχηματίζονται λειτουργικές οδοί μεταφοράς των ώσεων [35]. Η θέση όπου δύο νευρώνες πλησιάζουν αρκετά και τελείται λειτουργική διανευρωνική επικοινωνία μεταξύ τους αποκαλείται σύναψη [35]. Η επικοινωνία σε μια σύναψη, δηλαδή η μεταβίβαση της νευρικής ώσης υπό φυσιολογικές συνθήκες, γίνεται μόνο κατά μία κατεύθυνση [35]. Υπάρχουν διάφοροι τύποι συνάψεων [35]. Ο συχνότερος τύπος είναι μεταξύ του άξονα ενός νευρώνα και του δενδρίτη ή μεταξύ του άξονα ενός νευρώνα και του κυτταρικού σώματος ενός δευτέρου νευρώνα [35]. Κατά τα άκρα του τελικού δενδρίτη του άξονα, εγγύς της σύναψης, κάθε κλαδίσκος εμφανίζει ένα τελικό κομβίο ή μπορεί να εμφανίζει σειρά κομβίων, κάθε ένα από τα οποία κάνει συναπτική επαφή [35]. Σε άλλους τύπους συνάψεων, ο άξονας συνάπτεται με το αρχικό τμήμα άλλου άξονα, ή μπορεί να γίνονται συνάψεις των τελικών κομβίων από αρκετούς διαφορετικούς νευρώνες [35]. Ο τρόπος κατάληξης ενός άξονα ποικίλει στα διάφορα μέρη του νευρικού συστήματος [35]. Ένας άξονας μπορεί να συνάπτεται με ένα μόνο νευρώνα, ή μπορεί να συνάπτεται με πολλούς νευρώνες [35]. Αντίστοιχα, ένας νευρώνας μπορεί να έχει συνάψεις με τους άξονες πολλών διαφορετικών νευρώνων [35].

Ο τρόπος διασύνδεσης μεταξύ των βιολογικών νευρώνων αποτελεί τη βάση στην οποία έχει στηριχθεί η λειτουργία των τεχνητών νευρωνικών δικτύων.

## 2.6.2 Τεχνητός Νευρώνας

Ο τεχνητός νευρώνας αποτελεί το κύριο δομικό στοιχείο των τεχνητών νευρωνικών δικτύων [37]. Οι τεχνητοί νευρώνες αποτελούν μονάδες επεξεργασίας πληροφοριών σε υπολογιστικά μοντέλα. Η λειτουργία ενός τεχνητού νευρώνα είναι η ακόλουθη:

Ένας τεχνητός νευρώνας συνδέεται με ένα πλήθος συνάψεων κάθε μία από τις οποίες χαρακτηρίζεται από ένα βάρος (weight) [37]. Ο τεχνητός νευρώνας λαμβάνει σήματα εισόδου από τις συνάψεις σταθμισμένα από τα βάρη των συνάψεων, τα οποία στη συνέχεια αθροίζει μεταξύ τους μέσω ενός αθροιστή (addition) [37]. Τέλος, ένας εξωτερικός παράγοντας του δικτύου που ονομάζεται πόλωση (bias) προστίθεται στο σταθμισμένο αθροισμα των εισόδων [37], σύμφωνα με την Εξίσωση (2.1):

$$z = \sum x_i * w_i + b \quad (2.1)$$

όπου:

$x_i$  = το σήμα εισόδου που προέρχεται από τη σύναψη  $i$

$w_i$  = το βάρος της σύναψης  $i$

$b$  = η πόλωση

Το σταθμισμένο αθροισμα των εισόδων που παράγεται από τον αθροιστή καθώς και η πόλωση, διέρχονται στη συνέχεια μέσω μιας συνάρτησης που ονομάζεται συνάρτηση ενεργοποίησης (activation function) και η οποία εισάγει μη-γραμμικότητα στην έξοδο του νευρώνα [19]. Η συνάρτηση ενεργοποίησης λαμβάνει το σταθμισμένο αθροισμα των εισόδων και ενεργοποιεί τον νευρώνα εάν αυτό ξεπεράσει ένα συγκεκριμένο κατώφλι ενεργοποίησης [19], [37].

Η ισχύς των τεχνητών νευρώνων βρίσκεται στην ικανότητά τους να μαθαίνουν και να προσαρμόζονται μέσω της διαδικασίας της εκπαίδευσης. Κατά τη διάρκεια της εκπαίδευσης, τα βάρη και οι εφαρμοζόμενες πολώσεις των νευρώνων προσαρμόζονται σύμφωνα με τον αλγόριθμο μάθησης που ακολουθείται και την επιθυμητή έξοδο. Αυτό δίνει τη δυνατότητα στα τεχνητά νευρωνικά δίκτυα να μαθαίνουν μέσω παραδειγμάτων και να πραγματοποιούν προβλέψεις βασιζόμενα σε μοτίβα που παρουσιάζονται μεταξύ των δεδομένων.

### 2.6.3 Τεχνητά Νευρωνικά Δίκτυα (Artificial Neural Networks)

Τα τεχνητά νευρωνικά δίκτυα αποτελούν μαθηματικά μοντέλα τα οποία είναι εμπνευσμένα από τη δομή και τον τρόπο λειτουργίας των βιολογικών νευρωνικών δικτύων. Οι νευρώνες ενός τεχνητού νευρωνικού δικτύου οργανώνονται σε επίπεδα. Οι τρεις βασικοί τύποι επιπέδων είναι οι ακόλουθοι:

Το επίπεδο εισόδου (input layer), το οποίο αποτελεί το αρχικό επίπεδο του τεχνητού νευρωνικού δικτύου και το οποίο λαμβάνει ως είσοδο τα αρχικά δεδομένα ή αλλιώς χαρακτηριστικά του προβλήματος προς επίλυση [19]. Κάθε νευρώνας του επιπέδου εισόδου αντιπροσωπεύει ένα χαρακτηριστικό ή μια μεταβλητή εισόδου. Ο αριθμός των νευρώνων του επιπέδου εισόδου εξαρτάται από τις διαστάσεις των δεδομένων εισόδου. Το συγκεκριμένο επίπεδο μεταδίδει τα δεδομένα που λαμβάνει ως είσοδο στα επόμενα επίπεδα.

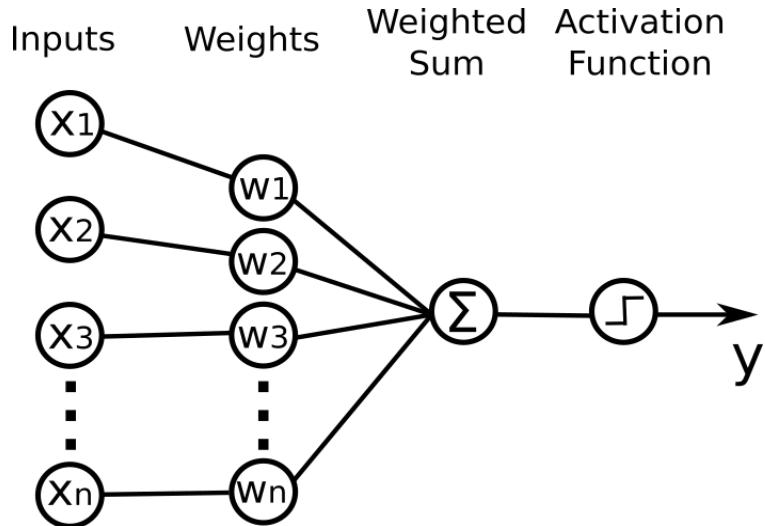
Το κρυφό επίπεδο (hidden layer), το οποίο αποτελεί κάθε επίπεδο που βρίσκεται μεταξύ του επιπέδου εισόδου και του επιπέδου εξόδου. Ένα νευρωνικό δίκτυο μπορεί να περιλαμβάνει περισσότερα από ένα κρυφά επίπεδα [37]. Στα συγκεκριμένα επίπεδα πραγματοποιείται η επεξεργασία των δεδομένων που λαμβάνονται ως είσοδος από το επίπεδο εισόδου και η εξαγωγή χαρακτηριστικών που θα βοηθήσουν το νευρωνικό δίκτυο στην εκμάθηση σύνθετων μοτίβων που παρουσιάζονται στα δεδομένα [19]. Οι νευρώνες ενός κρυφού επιπέδου λαμβάνουν δεδομένα από τους νευρώνες του προηγούμενου επιπέδου και εφαρμόζουν υπολογισμούς πάνω σε αυτά με στόχο να εξάγουν χαρακτηριστικά. Όσον αφορά τα χαρακτηριστικά που εξάγονται από τα κρυφά επίπεδα, τα αρχικά κρυφά επίπεδα ενός τεχνητού νευρωνικού δικτύου εξάγουν χαρακτηριστικά χαμηλού επιπέδου, ενώ τα τελευταία επίπεδα εξάγουν χαρακτηριστικά υψηλού επιπέδου [19]. Έπειτα από την επεξεργασία των δεδομένων, ένα κρυφό επίπεδο μεταδίδει τα δεδομένα του στους νευρώνες του επόμενου επιπέδου.

Το επίπεδο εξόδου (output layer), το οποίο αποτελεί το τελικό επίπεδο του τεχνητού νευρωνικού δικτύου και το οποίο πραγματοποιεί την πρόβλεψη λαμβάνοντας υπόψη τους υπολογισμούς που πραγματοποιήθηκαν και τα χαρακτηριστικά που εξήγησαν από τους νευρώνες των προηγούμενων επιπέδων [19]. Ο αριθμός των νευρώνων του επιπέδου εξόδου εξαρτάται από τη φύση του προβλήματος που καλείται να επιλύσει το νευρωνικό δίκτυο.

Όσον αφορά τις αρχιτεκτονικές τεχνητών νευρωνικών δικτύων, υπάρχουν δύο βασικές κατηγορίες αρχιτεκτονικών:

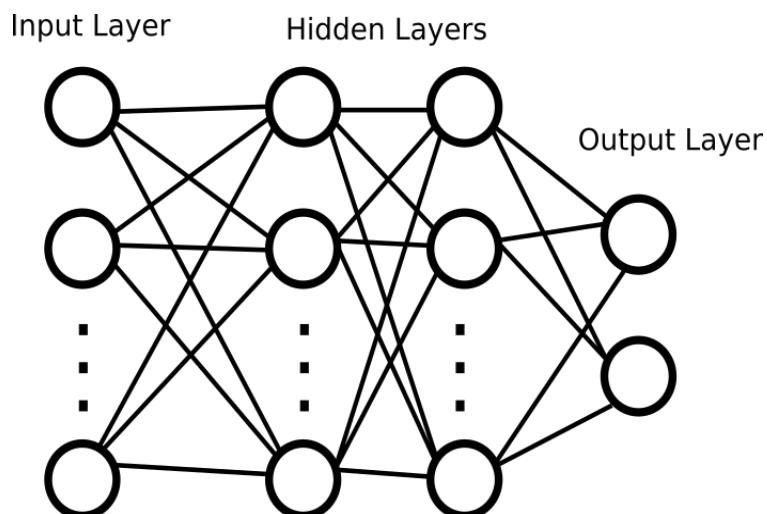
Το perceptron, που αποτελεί τον απλούστερο τύπο τεχνητού νευρωνικού δικτύου, ο οποίος υλοποιήθηκε το 1958 από τον Frank Rosenblatt [38]. Περιλαμβάνει ένα μόνο νευρώνα, ο οποίος δέχεται σήματα εισόδου από τις συνάψεις με τις οποίες διασυνδέεται και μέσω του αθροιστή υπολογίζει έναν γραμμικό συνδυασμό τους, εφαρμόζοντας μια εξωτερικά εφαρμοζόμενη πόλωση [37]. Η έξοδος του αθροιστή

διέρχεται μέσω μιας συνάρτησης ενεργοποίησης, η οποία στην ουσία αποτελεί μια συνάρτηση κατωφλίου που καθορίζει εάν θα πυροδοτηθεί ο νευρώνας ή όχι [37]. Στην Εικόνα 2.7, παρουσιάζεται η αρχιτεκτονική του perceptron του Rosenblatt.



Εικόνα 2.7: Το Perceptron του Rosenblatt

Το perceptron πολλών επιπέδων, που αποτελείται από επίπεδα κάθε ένα από τα οποία περιλαμβάνει περισσότερους από έναν νευρώνες [19], [37]. Ένα perceptron πολλών επιπέδων περιλαμβάνει ένα επίπεδο εισόδου, ένα η περισσότερα κρυφά επίπεδα, καθώς και ένα επίπεδο εξόδου [37]. Μεταξύ των επιπέδων υπάρχουν συνάψεις σταθμισμένων βαρών που διασυνδέουν τους νευρώνες των επιπέδων του δικτύου μεταξύ τους [19]. Στην Εικόνα 2.8, παρουσιάζεται η αρχιτεκτονική ενός perceptron πολλών επιπέδων.



Εικόνα 2.8: Perceptron πολλών επιπέδων

## 2.7 Συνάρτηση Ενεργοποίησης (Activation Function)

Η συνάρτηση ενεργοποίησης αποτελεί μία μαθηματική συνάρτηση μέσω της οποίας καθορίζεται εάν ένας συγκεκριμένος νευρώνας θα ενεργοποιηθεί ή όχι [19]. Η συνάρτηση ενεργοποίησης τοποθετείται στην έξοδο κάθε νευρώνα ενός επιπέδου του νευρωνικού δικτύου και μετασχηματίζει το γραμμικό συνδυασμό του σταθμισμένου αθροίσματος που λαμβάνει ως είσοδο σε μη-γραμμικό μοντέλο που στη συνέχεια μεταδίδεται στο επόμενο επίπεδο [19]. Με λίγα λόγια, η συνάρτηση ενεργοποίησης εισάγει μη-γραμμικότητα στο δίκτυο [19]. Μέσω της εισαγωγής μη-γραμμικότητας στο δίκτυο, η συνάρτηση ενεργοποίησης προσφέρει στο μοντέλο τη δυνατότητα να ανιχνεύει σύνθετες σχέσεις ανάμεσα στις εισόδους και στις εξόδους. Χωρίς τη χρήση συναρτήσεων ενεργοποίησης το τεχνητό νευρωνικό δίκτυο θα αποτελούσε μια γραμμική συνάρτηση και η προσθήκη πολλαπλών επιπέδων σε αυτό δεν θα παρείχε καμία επιπλέον ισχύ μοντελοποίησης [19].

### 2.7.1 Σιγμοειδής Συνάρτηση Ενεργοποίησης (Sigmoid Activation Function)

Η σιγμοειδής συνάρτηση ενεργοποίησης, η οποία ονομάζεται και λογιστική συνάρτηση (logistic function), αποτελεί τη πιο ευρέως χρησιμοποιούμενη συνάρτηση ενεργοποίησης [19]. Είναι μια μαθηματική συνάρτηση η οποία αντιστοιχίζει οποιοδήποτε πραγματικό αριθμό σε μία τιμή στο εύρος 0 έως 1 [39]. Η σιγμοειδής συνάρτηση χρησιμοποιείται σε τεχνητά νευρωνικά δίκτυα που έχουν κατασκευαστεί ούτως ώστε να επιλύουν δυαδικά προβλήματα [19]. Ο τύπος της σιγμοειδούς συνάρτησης δίνεται στην Εξίσωση (2.2):

$$f(x) = \frac{1}{1 + e^{-x}} \quad (2.2)$$

### 2.7.2 Συνάρτηση Ενεργοποίησης Softmax (Softmax Activation Function)

Η συνάρτηση ενεργοποίησης softmax αποτελεί γενίκευση της σιγμοειδούς συνάρτησης [19]. Είναι μια μαθηματική συνάρτηση που λαμβάνει ως είσοδο ένα διάνυσμα πραγματικών αριθμών και παράγει μία κατανομή πιθανοτήτων [40]. Πιο αναλυτικά, η συγκεκριμένη συνάρτηση υπολογίζει την εκθετική τιμή κάθε στοιχείου εισόδου και στη συνέχεια διαιρεί κάθε εκθετική τιμή με το άθροισμα όλων των εκθετικών τιμών. Αυτό εξασφαλίζει πως οι τιμές του αποτελέσματος θα βρίσκονται στο εύρος 0 έως 1 και θα έχουν άθροισμα 1 [40]. Η συνάρτηση softmax χρησιμοποιείται σε τεχνητά νευρωνικά δίκτυα που έχουν κατασκευαστεί ούτως ώστε

να επιλύουν πολυταξικά προβλήματα [19]. Ο τύπος της συνάρτησης softmax δίνεται στην Εξίσωση (2.3):

$$f(x_i) = \frac{e^{x_i}}{\sum e^{x_j}} \quad (2.3)$$

### 2.7.3 Ανορθωμένη Γραμμική Μονάδα (ReLU)

Η ανορθωμένη γραμμική μονάδα (ReLU) αποτελεί μια μη γραμμική μαθηματική συνάρτηση που χρησιμοποιείται ως συνάρτηση ενεργοποίησης στα τεχνητά νευρωνικά δίκτυα [39]. Η συγκεκριμένη συνάρτηση χρησιμοποιείται στα κρυφά επίπεδα του νευρωνικού δικτύου σε συνδυασμό με κάποια άλλη συνάρτηση ενεργοποίησης που χρησιμοποιείται στα επίπεδα εξόδου του νευρωνικού δικτύου [40]. Είναι γρηγορότερη και πιο αποδοτική από άλλες συναρτήσεις ενεργοποίησης, ενώ έχει καλύτερη ικανότητα γενίκευσης [40]. Αυτό συμβαίνει διότι σε αντίθεση με άλλες συναρτήσεις ενεργοποίησης, όταν χρησιμοποιείται η συνάρτηση ReLU δεν ενεργοποιούνται όλοι οι νευρώνες, αλλά αντίθετα ορισμένοι από αυτούς [39]. Ο τύπος της συνάρτησης ReLU δίνεται στην Εξίσωση (2.4).

$$f(x) = \max(0, x) = \begin{cases} 0, & x < 0 \\ x, & x \geq 0 \end{cases} \quad (2.4)$$

Παρατηρώντας τον τύπο της ανορθωμένης γραμμικής μονάδας (ReLU) μπορεί εύκολα να διαπιστωθεί πως αυτή επιστρέφει την τιμή 0 για κάθε αρνητική είσοδο και την τιμή x για κάθε θετική είσοδο [40].

## 2.8 Συνάρτηση Απώλειας/Κόστους (Loss/Cost Function)

Η συνάρτηση απώλειας ή συνάρτηση κόστους αποτελεί μία μαθηματική συνάρτηση η οποία χρησιμοποιείται για την αξιολόγηση της αποδοτικότητας ενός νευρωνικού δικτύου. Για την επίτευξη του προαναφερθέντος, η συνάρτηση απώλειας υπολογίζει την απόσταση ανάμεσα στη προβλεπόμενη έξοδο ενός νευρωνικού δικτύου και στη αναμενόμενη έξοδο όταν αυτό εκπαιδεύεται με συγκεκριμένα δεδομένα εισόδου. Η συνάρτηση απώλειας διαδραματίζει κρίσιμο ρόλο στην εκπαίδευση των νευρωνικών δικτύων αφού μέσω αυτής μπορεί να ελεγχθεί η επίδοση ενός νευρωνικού δικτύου για ένα συγκεκριμένο πρόβλημα [19]. Στόχος της

διαδικασίας εκπαίδευσης είναι η ελαχιστοποίηση όσο το δυνατόν περισσότερο της συνάρτησης απώλειας [19]. Όσο χαμηλότερη είναι η συνάρτηση απώλειας τόσο μεγαλύτερη είναι η ικανότητα ενός δικτύου να πραγματοποιεί ορθές προβλέψεις [19]. Η πιο ευρέως χρησιμοποιούμενη συνάρτηση απώλειας είναι η συνάρτηση cross entropy.

### 2.8.1 Συνάρτηση Cross Entropy

Η συνάρτηση cross entropy αποτελεί μια συνάρτηση απώλειας η οποία εκφράζει τη διαφορά ανάμεσα σε δύο κατανομές πιθανοτήτων [19]. Πιο αναλυτικά, η συγκεκριμένη συνάρτηση υπολογίζει την απόσταση μεταξύ της προβλεπόμενης κατανομής πιθανοτήτων και της πραγματικής κατανομής πιθανοτήτων. Για παράδειγμα, έστω ότι έχουμε ένα πρόβλημα ταξινόμησης εικόνων σε  $c$  κατηγορίες. Το νευρωνικό δίκτυο παράγει μια προβλεπόμενη πιθανότητα, τυπικά μεταξύ του εύρους 0 και 1, η οποία αντιπροσωπεύει τη πιθανότητα η είσοδος να ανήκει σε μία από τις  $c$  κατηγορίες. Ας ονομάσουμε τη προβλεπόμενη πιθανότητα μια είσοδος να ανήκει στην κατηγορία  $i$  ως  $p_i$  και την πραγματική πιθανότητα μια είσοδος να ανήκει στην κατηγορία  $i$  ως  $y_i$ , όπου  $y_i$  ισούται με το 1 εάν η είσοδος ανήκει στη συγκεκριμένη κατηγορία  $i$  και  $y_i$  ισούται με το 0 εάν ανήκει σε κάποια από τις άλλες κατηγορίες. Ο τύπος της συνάρτησης cross entropy δίνεται στην Εξίσωση (2.5):

$$L = - \sum_{i=1}^c y_i * \log(p_i) \quad (2.5)$$

Εάν ο αριθμός των κατηγοριών  $c$  ισούται με 2, τότε έχουμε να κάνουμε με δυαδικό (binary) πρόβλημα και η συνάρτηση απώλειας ονομάζεται binary cross entropy, ενώ εάν ο αριθμός των κατηγοριών  $c$  είναι μεγαλύτερος του 2, τότε έχουμε να κάνουμε με πολυταξικό (multiclass) πρόβλημα και η συνάρτηση απώλειας ονομάζεται categorical cross entropy.

## 2.9 Αλγόριθμοι Βελτιστοποίησης (Optimization Algorithms)

Οι αλγόριθμοι βελτιστοποίησης διαδραματίζουν σημαντικό ρόλο στη διαδικασία εκπαίδευσης των τεχνητών νευρωνικών δικτύων. Αυτοί οι αλγόριθμοι είναι υπεύθυνοι για την προσαρμογή των παραμέτρων ενός νευρωνικού δικτύου ούτως ώστε να ελαχιστοποιηθεί η συνάρτηση απώλειας [19]. Στόχος είναι να ανακαλυφθούν οι βέλτιστες τιμές βαρών των συνάψεων και πολώσεων των νευρώνων

που επιτρέπουν σε ένα δίκτυο να πραγματοποιεί ορθές προβλέψεις [19]. Υπάρχουν διάφοροι αλγόριθμοι βελτιστοποίησης, ο σημαντικότερος από τους οποίους είναι ο αλγόριθμος βελτιστοποίησης κατάβασης κλίσης (gradient descent).

### 2.9.1 Αλγόριθμος Βελτιστοποίησης Κατάβασης Κλίσης (Gradient Descent Optimization Algorithm)

Ο αλγόριθμος βελτιστοποίησης κατάβασης κλίσης (gradient descent) αποτελεί έναν αλγόριθμο βελτιστοποίησης ο οποίος χρησιμοποιείται συχνά στο πεδίο της μηχανικής μάθησης για την ελαχιστοποίηση της συνάρτησης απώλειας [19]. Είναι ιδιαίτερα χρήσιμος όταν έχουμε να κάνουμε με προβλήματα μεγάλης κλίμακας όπου είναι υπολογιστικά ακριβό ή ακατόρθωτο να υπολογισθεί το ακριβές σημείο που εμφανίζει την ελάχιστη απώλεια [19]. Ο αλγόριθμος gradient descent ενημερώνει επαναληπτικά τα βάρη και τις πολώσεις ενός μοντέλου με στόχο να ανακαλύψει το σημείο της συνάρτησης απώλειας στο οποίο η απώλεια είναι η ελάχιστη [19]. Η λειτουργία του αλγορίθμου gradient descent αποτελείται από τα εξής βήματα:

- 1) Αρχικοποίηση των τιμών των παραμέτρων που αντιστοιχούν στα βάρη και στις πολώσεις, οι οποίες ορίζονται τυχαία ή με κάποιες προκαθορισμένες τιμές.
- 2) Υπολογισμός, σε κάθε επανάληψη εκτέλεσης του αλγορίθμου η οποία ονομάζεται εποχή (epoch), της κλίσης της συνάρτησης απώλειας ως προς τις παραμέτρους που αντιστοιχούν στα βάρη και στις πολώσεις [19]. Ο υπολογισμός της κλίσης πραγματοποιείται μέσω του υπολογισμού της παραγώγου της απώλειας [19]. Η υπολογιζόμενη κλίση αντιπροσωπεύει τη κατεύθυνση της πιο απότομης ανόδου, υποδεικνύοντας πως οι διάφορες μεταβολές στις παραμέτρους επηρεάζουν τη τιμή της συνάρτησης απώλειας [19].
- 3) Υπολογισμός του ρυθμού μάθησης. Ο ρυθμός μάθησης αποτελεί υπερπαράμετρο ενός νευρωνικού δικτύου, η οποία καθορίζει το μέγεθος βήματος που θα πραγματοποιήσει το εκάστοτε μοντέλο κατά την διαδικασία προσέγγισης της ελάχιστης τιμής απώλειας, δηλαδή το μέγεθος της ενημέρωσης των τιμών των παραμέτρων που αντιστοιχούν στα βάρη και στις πολώσεις σε κάθε επανάληψη [19]. Ένας μικρότερος ρυθμός μάθησης έχει ως αποτέλεσμα η εκπαίδευση να είναι πιο αργή, αλλά εξασφαλίζει ότι θα βρεθεί το σημείο στο οποίο η απώλεια είναι ελάχιστο [19]. Από την άλλη, ένας μεγαλύτερος ρυθμός μάθησης μπορεί να επιταχύνει τη διαδικασία εκπαίδευσης, ωστόσο ενδέχεται να προκαλέσει υπέρβαση ή απόκλιση από το σημείο στο οποίο εμφανίζεται ελαχιστοποίηση της απώλειας [19].
- 4) Επανάληψη της διαδικασίας μέχρις ότου να ικανοποιηθεί κάποιο κριτήριο διακοπής. Συχνά χρησιμοποιούμενα κριτήρια διακοπής

αποτελούν το πέρας ενός συγκεκριμένου αριθμού εποχών, η μείωση της συνάρτησης απώλειας κάτω από ένα συγκεκριμένο όριο, ή ο τερματισμός της διαδικασίας λόγω ασήμαντων μεταβολών στην τιμή της συνάρτησης απώλειας μεταξύ των επαναλήψεων.

Η μεταβολή του βάρους σε κάθε βήμα δίνεται στην Εξίσωση (2.6):

$$\Delta_{W_i} = -\alpha \frac{dE}{dW_i} \quad (2.6)$$

όπου:

$W_i$  = το βάρος

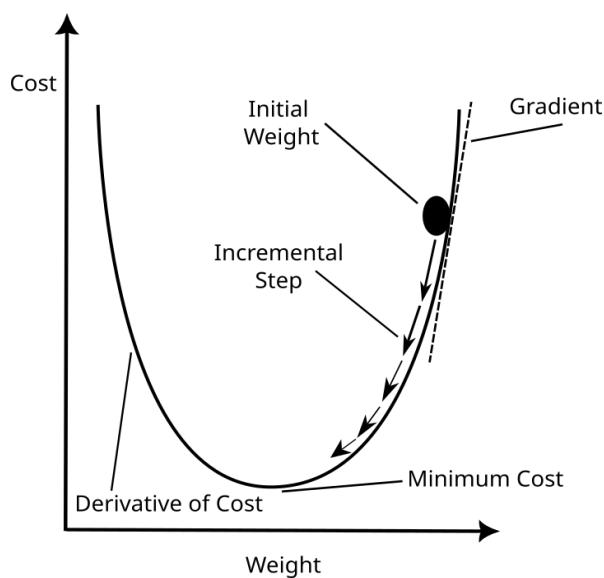
$\alpha$  = ο ρυθμός μάθησης

$E$  = η απώλεια

Ο υπολογισμός του επόμενου βήματος δίνεται στην Εξίσωση (2.7):

$$W_{\text{επόμενου\_βήματος}} = W_{\text{τωρινό}} + \Delta_W \quad (2.7)$$

Στην Εικόνα 2.9, παρουσιάζεται μία σχηματική αναπαράσταση της εκτέλεσης του αλγορίθμου gradient descent.



Εικόνα 2.9: Σχηματική αναπαράσταση της εκτέλεσης του αλγορίθμου gradient descent

### 2.9.2 Αλγόριθμος Βελτιστοποίησης Adaptive Moment Estimation (Adam)

Ο αλγόριθμος Adam (Adaptive Moment Estimation) είναι ένας πολύ δημοφιλής αλγόριθμος βελτιστοποίησης στον τομέα της βαθιάς μάθησης και αποτελεί επέκταση του αλγορίθμου βελτιστοποίησης stochastic gradient descent (SGD). Ο συγκεκριμένος αλγόριθμος, σε αντίθεση με άλλους αλγορίθμους όπως ο stochastic gradient descent (SGD) που χρησιμοποιούν έναν σταθερό ρυθμό μάθησης, προσαρμόζει το ρυθμό μάθησης για κάθε παράμετρο του δικτύου ξεχωριστά [41]. Ο αλγόριθμος Adam συνδυάζει τα πλεονεκτήματα δύο άλλων αλγορίθμων βελτιστοποίησης, του AdaGrad και του RMSprop ούτως ώστε να επιτύχει ενημέρωση παραμέτρων με αποδοτικό και προσαρμοστικό τρόπο [41]. Πιο συγκεκριμένα, προσαρμόζει τα βάρη βασιζόμενος στη 1<sup>η</sup> και στη 2<sup>η</sup> στιγμή των κλίσεων [41]. Τα πλεονεκτήματα του αλγορίθμου Adam είναι ότι αυξάνει την επίδοση του νευρωνικού δικτύου και το βοηθάει να οδηγηθεί γρηγορότερα σε σύγκλιση [41].

## 2.10 Αλγόριθμος Οπισθοδιάδοσης Backward Propagation (Backpropagation) Algorithm

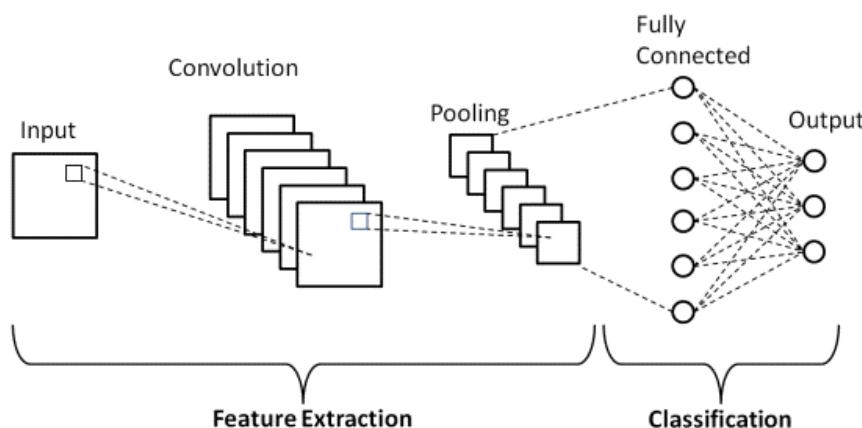
Ο αλγόριθμος οπισθοδιάδοσης (backward propagation ή backpropagation) αποτελεί έναν αλγόριθμο ο οποίος προτάθηκε το 1986 από τους Rummelhart κ.α. [42] και χρησιμοποιείται κατά τη φάση εκπαίδευσης ενός perceptron πολλών επιπέδων, ούτως ώστε να ενημερώσει τα βάρη σύμφωνα με την απώλεια μεταξύ της προβλεπόμενης εξόδου και της αναμενόμενης εξόδου. Στόχος είναι να προσαρμοσθούν τα βάρη με τέτοιο τρόπο ούτως ώστε να ελαχιστοποιείται η απώλεια και να βελτιώνεται η επίδοση του δικτύου. Ο αλγόριθμος αποτελείται από τα εξής στάδια, τα οποία επαναλαμβάνονται μέχρις ότου ολοκληρωθεί η εκπαίδευτική διαδικασία:

- 1) Εκτέλεση της διαδικασίας εμπρόσθιας διάδοσης (forward pass), όπου τα δεδομένα εισόδου τροφοδοτούνται στο επίπεδο εισόδου και διαμέσου των κρυφών επιπέδων φτάνουν στο επίπεδο εξόδου όπου πραγματοποιείται η εκάστοτε πρόβλεψη. Κατά τη διαδικασία μετάδοσης, σε κάθε επίπεδο υπολογίζεται το σταθμισμένο άθροισμα των εισόδων του, εφαρμόζεται μια εξωτερική πόλωση και στη συνέχεια το αποτέλεσμα διέρχεται μέσω της συνάρτησης ενεργοποίησης [19]. Η έξοδος του τελευταίου επιπέδου, δηλαδή του επιπέδου εξόδου, αντιπροσωπεύει την πρόβλεψη του δικτύου για τα δεδομένα εισόδου.
- 2) Υπολογισμός της απώλειας, δηλαδή της απόστασης ανάμεσα στην προβλεπόμενη έξοδο του δικτύου και στην αναμενόμενη έξοδο [19]. Ο υπολογισμός της απώλειας πραγματοποιείται μέσω μιας συνάρτησης απώλειας όπως είναι η συνάρτηση cross entropy [19].

- 3) Εκτέλεση της διαδικασίας οπισθοδιάδοσης (backpropagation), όπου ξεκινώντας από το τελευταίο προς το πρώτο επίπεδο υπολογίζεται η κλίση σε κάθε νευρώνα του δικτύου, χρησιμοποιώντας έναν αλγόριθμο βελτιστοποίησης όπως είναι ο gradient descent [19]. Σε κάθε νευρώνα πραγματοποιείται ενημέρωση των βαρών και των πολώσεων με βάση τις κλίσεις που υπολογίσθηκαν και το ρυθμό μάθησης [19].
- 4) Επανάληψη της διαδικασίας μέχρις ότου να ικανοποιηθεί κάποιο κριτήριο διακοπής.

### 2.10.1 Συνελικτικά Νευρωνικά Δίκτυα (Convolutional Neural Networks)

Τα συνελικτικά νευρωνικά δίκτυα αποτελούν έναν τύπο τεχνητών νευρωνικών δικτύων που έχει σχεδιαστεί ειδικά για την επεξεργασία δεδομένων όπως είναι οι εικόνες. Χρησιμοποιούνται ευρέως σε προβλήματα όρασης υπολογιστών, όπως είναι η ταξινόμηση εικόνων και η σημασιολογική κατάτμηση εικόνων. Ο όρος «συνελικτικά» στην ονομασία του συγκεκριμένου τύπου νευρωνικών δικτύων έχει να κάνει με το γεγονός ότι τα συγκεκριμένα δίκτυα αξιοποιούν την πράξη της συνέλιξης στα επίπεδά τους [23]. Η κύρια ιδέα πίσω από τα συνελικτικά νευρωνικά δίκτυα είναι η εξαγωγή χαρακτηριστικών από τα δεδομένα που τους τροφοδοτούνται ως είσοδος [19]. Το δίκτυο αποτελείται από πολλαπλά επίπεδα, συμπεριλαμβανομένων των συνελικτικών επιπέδων (convolutional layers), των επιπέδων ενεργοποίησης (activation layers), των επιπέδων συγκέντρωσης (pooling layers), και των πλήρως συνδεδεμένων επιπέδων (fully connected layers), η δομή και η λειτουργία των οποίων παρουσιάζονται στις επόμενες παραγράφους. Στην Εικόνα 2.10 παρουσιάζεται η βασική δομή ενός συνελικτικού νευρωνικού δικτύου.

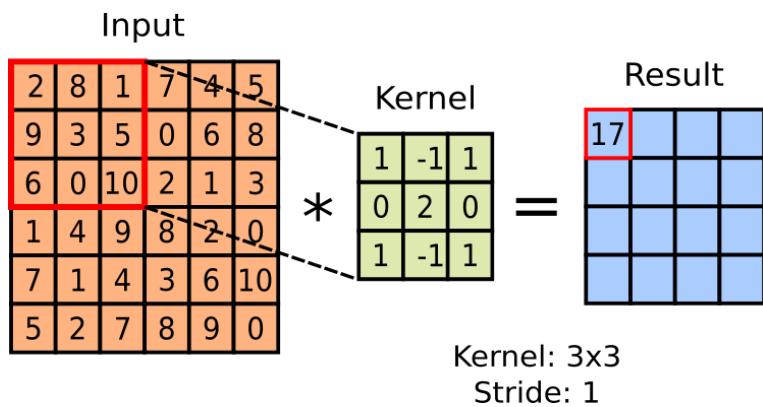


Εικόνα 2.10: Βασική δομή ενός συνελικτικού νευρωνικού δικτύου [43]

## 2.10.2 Επίπεδο Συνέλιξης (Convolutional Layer)

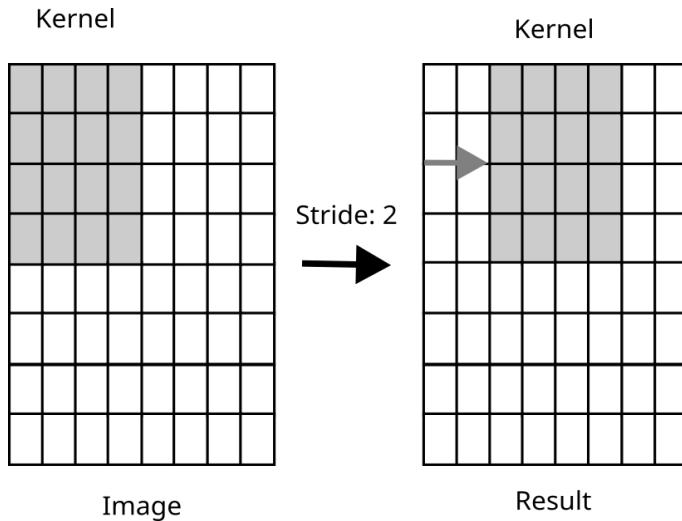
Αποτελεί το βασικό δομικό στοιχείο ενός συνελικτικού νευρωνικού δικτύου [19]. Στο συγκεκριμένο επίπεδο πραγματοποιείται εξαγωγή χαρακτηριστικών μέσω της εφαρμογής της πράξης της συνέλιξης (convolution) στις εικόνες που λαμβάνονται ως είσοδος [19].

Η συνέλιξη αποτελεί μια πράξη η οποία πραγματοποιείται μεταξύ των εικόνων που λαμβάνονται ως είσοδος και ενός δυσδιάστατου πίνακα που ονομάζεται πυρήνας (kernel) ή φίλτρο (filter) [19]. Πιο συγκεκριμένα, κατά τη πράξη της συνέλιξης πραγματοποιείται ολίσθηση του πυρήνα σε όλο το πλάτος και το ύψος της εισόδου, όπου σε κάθε βήμα κάθε στοιχείο του πυρήνα πολλαπλασιάζεται με το αντίστοιχο στοιχείο της εικόνας εισόδου και στη συνέχεια τα στοιχεία προστίθενται μεταξύ τους παράγοντας έναν δισδιάστατο πίνακα με νέα στοιχεία που ονομάζεται αντιστοίχιση χαρακτηριστικών (activation map) [19]. Στην Εικόνα 2.11 παρουσιάζεται ένα παράδειγμα της λειτουργίας του επιπέδου συνέλιξης.



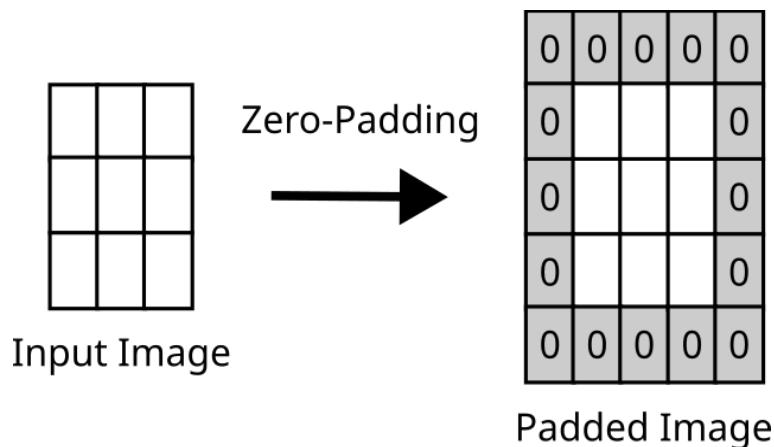
Εικόνα 2.11: Παράδειγμα λειτουργίας επιπέδου συνέλιξης

Το βήμα ολίσθησης (stride) του πυρήνα, δηλαδή ο αριθμός των στοιχείων κατά τα οποία ολισθαίνει κάθε φορά ο πυρήνας, καθορίζεται από την αντίστοιχη υπερπαράμετρο του επιπέδου συνέλιξης [19]. Στην Εικόνα 2.12 παρουσιάζεται ένα παράδειγμα εφαρμογής του βήματος ολίσθησης.



Εικόνα 2.12: Παράδειγμα εφαρμογής του βήματος ολίσθησης (stride)

Μία σημαντική τεχνική που χρησιμοποιείται σε ένα επίπεδο συνέλιξης είναι η padding, η οποία καθορίζεται από την αντίστοιχη υπερπαράμετρο του επιπέδου συνέλιξης [19]. Στην ουσία προστίθενται τιμές στα άκρα της εικόνας, διατηρώντας το χωρικό της μέγεθος σταθερό, ούτως ώστε το πλάτος και το ύψος των εικόνων εισόδου και εξόδου να είναι ίδια [19]. Στην Εικόνα 2.13 παρουσιάζεται ένα παράδειγμα εφαρμογής της τεχνικής padding, όπου προστίθενται μηδενικές τιμές στα άκρα της εικόνας. Η συγκεκριμένη τεχνική padding ονομάζεται zero-padding.



Εικόνα 2.13: Παράδειγμα εφαρμογής της τεχνικής padding (zero-padding)

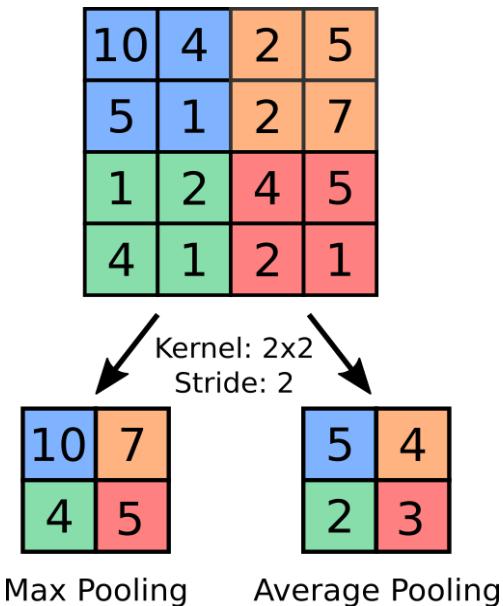
### 2.10.3 Επίπεδο Συγκέντρωσης (Pooling Layer)

Όσο περισσότερα συνελικτικά επίπεδα προστίθενται σε μία αρχιτεκτονική συνελικτικού νευρωνικού δικτύου, τόσο περισσότερες είναι οι παράμετροι που πρέπει να ρυθμιστούν από το δίκτυο, αυξάνοντας τη χρονική και χωρική πολυπλοκότητα του δικτύου [19]. Λόγη σε αυτό το ζήτημα αποτελεί η προσθήκη επιπέδων συγκέντρωσης στο δίκτυο [19]. Ένα επίπεδο συγκέντρωσης πραγματοποιεί υποδειγματοληψία (subsampling) μειώνοντας σημαντικά την υπολογιστική πολυπλοκότητα του δικτύου, αφού μειώνει τον αριθμό των χαρακτηριστικών που περνούν στο επόμενο επίπεδο και κατ' επέκταση τον αριθμό των παραμέτρων που πρέπει να ρυθμιστούν [19]. Τα πιο συνηθισμένα επίπεδα συγκέντρωσης είναι το επίπεδο συγκέντρωσης μεγίστου (max pooling layer) και το επίπεδο συγκέντρωσης μέσου (average pooling layer), τα οποία περιγράφονται παρακάτω:

Το επίπεδο συγκέντρωσης τοπικού μεγίστου (max pooling layer) πραγματοποιεί ολίσθηση ενός φίλτρου σε όλο το πλάτος και το ύψος της εικόνας εισόδου. Ωστόσο, δεν πραγματοποιεί την πράξη της συνέλιξης, αλλά αντιθέτως επιλέγει τη μέγιστη τιμή, η οποία θα σταλεί στο επόμενο επίπεδο, αγνοώντας τις υπόλοιπες τιμές [19].

Το επίπεδο συγκέντρωσης τοπικού μέσου (average pooling layer) λειτουργεί όπως και το επίπεδο συγκέντρωσης τοπικού μεγίστου, με τη διαφορά ότι επιλέγει το μέσο όρο των τιμών εισόδου, ο οποίος θα σταλεί στο επόμενο επίπεδο [19].

Στην Εικόνα 2.14 παρουσιάζεται ένα παράδειγμα της διαδικασίας συγκέντρωσης (pooling) για τους δύο τύπους επιπέδων συγκέντρωσης.



Εικόνα 2.14: Παράδειγμα λειτουργίας των δύο τύπων επιπέδου συγκέντρωσης

#### 2.10.4 Πλήρως Συνδεδεμένο Επίπεδο (Fully Connected Layer)

Τα χαρακτηριστικά των εικόνων που εξάγονται από τα προηγούμενα επίπεδα του δικτύου μετατρέπονται σε ένα μονοδιάστατο διάνυσμα (flattened). Αυτό το διάνυσμα στη συνέχεια διέρχεται από ένα ή περισσότερα πλήρως συνδεδεμένα επίπεδα. Τα συγκεκριμένα επίπεδα πραγματοποιούν τη πρόβλεψη του δικτύου λαμβάνοντας υπόψη τις εικόνες που δίνονται ως είσοδος, καθώς και τα χαρακτηριστικά που έχουν εξαχθεί από τα προηγούμενα επίπεδα [19]. Ονομάζονται «πλήρως συνδεδεμένα» επειδή στα συγκεκριμένα επίπεδα κάθε νευρώνας ενός επιπέδου συνδέεται με όλους τους νευρώνες του επόμενου επιπέδου. Το τελευταίο επίπεδο ή αλλιώς επίπεδο εξόδου ενός συνελικτικού νευρωνικού δικτύου είναι ένα πλήρως συνδεδεμένο επίπεδο το οποίο περιλαμβάνει έναν ή περισσότερους νευρώνες, ανάλογα με το πρόβλημα που καλείται το νευρωνικό δίκτυο να επιλύσει.

### 2.11 Τεχνικές Κανονικοποίησης (Regularization Techniques)

Όπως έχει γίνει αντιληπτό, η υπερπροσαρμογή αποτελεί μείζον ζήτημα στα πεδία της μηχανικής και βαθιάς μάθησης. Υπάρχουν διάφορες τεχνικές κανονικοποίησης μέσω των οποίων δύναται να επιτευχθεί μείωση της υπερπροσαρμογής ενός μοντέλου βαθιάς μάθησης. Οι τεχνικές κανονικοποίησης που επιστρατεύονται στην παρούσα διπλωματική εργασία είναι η τεχνική της προσαύξησης δεδομένων (data augmentation), του πρόωρου τερματισμού (early stopping), καθώς και η προσθήκη επιπέδων dropout και επιπέδων batch normalization στο εκάστοτε τεχνητό νευρωνικό δίκτυο.

#### 2.11.1 Προσαύξηση Δεδομένων (Data Augmentation)

Η προσαύξηση δεδομένων αποτελεί μια τεχνική η οποία χρησιμοποιείται ευρέως στα πεδία της μηχανικής και της βαθιείας μάθησης για την αντιμετώπιση του φαινομένου της υπερπροσαρμογής. Επιτυγχάνεται μέσω της προσθήκης νέων δεδομένων στο ήδη υπάρχον σύνολο εκπαίδευσης, αυξάνοντας με αυτό τον τρόπο το μέγεθος και τη ποικιλομορφία του συνόλου και βελτιώνοντας τη ποιότητα του [44]. Τα νέα δεδομένα μπορούν να δημιουργηθούν με δύο τρόπους, είτε εφαρμόζοντας μετασχηματισμούς στα ήδη υπάρχοντα δεδομένα δημιουργώντας ελαφρώς παραλλαγμένα αντίγραφά τους, είτε δημιουργώντας συνθετικά δεδομένα από τα ήδη υπάρχοντα [44]. Η εφαρμογή των τεχνικών προσαύξησης μπορεί να πραγματοποιηθεί μεμονωμένα ή συνδυαστικά καθώς και ντετερμινιστικά ή με τυχαίο τρόπο στα δεδομένα εκπαίδευσης. Ωστόσο, αξίζει να σημειωθεί πως ο συνδυασμός των μετασχηματισμών πρέπει να πραγματοποιείται με προσοχή, διότι η υπερβολική

εφαρμογή μετασχηματισμών ενδεχομένως να δημιουργήσει εικόνες με χαρακτηριστικά που δεν εμφανίζονται στο σύνολο δεδομένων, επιφέροντας τα αντίθετα αποτελέσματα και οδηγώντας σε όξυνση του φαινομένου της υπερπροσαρμογής [44]. Το ποιοι μετασχηματισμοί είναι κατάλληλο να χρησιμοποιηθούν εξαρτάται από τη φύση του προβλήματος που ο εκάστοτε αλγόριθμος καλείται να επιλύσει καθώς και από τα χαρακτηριστικά των εικόνων του συνόλου δεδομένων.

Υπάρχει μια πληθώρα από τεχνικές προσαύξησης δεδομένων. Ορισμένες από αυτές είναι οι γεωμετρικοί μετασχηματισμοί (geometric transformations), οι μετασχηματισμοί χώρου χρωμάτων (color space transformations), τα φίλτρα πυρήνα (kernel filters), η ανάμειξη εικόνων (mixing images), η τυχαία αφαίρεση (random erasing), η προσαύξηση δεδομένων μέσω νευρωνικής μεταφοράς στυλ (neural style transfer data augmentation), καθώς και η προσαύξηση βασιζόμενη σε γενετικά ανταγωνιστικά δίκτυα (GAN-based data augmentation) [44]. Στα πλαίσια της πειραματικής διαδικασίας της παρούσας διπλωματικής εργασίας αξιοποιήθηκαν οι γεωμετρικοί μετασχηματισμοί.

### 2.11.2 Γεωμετρικοί Μετασχηματισμοί

Οι γεωμετρικοί μετασχηματισμοί (geometric transformations) αποτελούν έναν πολύ δημοφιλή τύπο μετασχηματισμών λόγω της ευκολίας με την οποία μπορούν να εφαρμοσθούν [44]. Στον Πίνακα 2.3 παρουσιάζονται ορισμένοι από τους πιο δημοφιλείς γεωμετρικούς μετασχηματισμούς.

**Πίνακας 2.3: Γεωμετρικοί μετασχηματισμοί εικόνων**

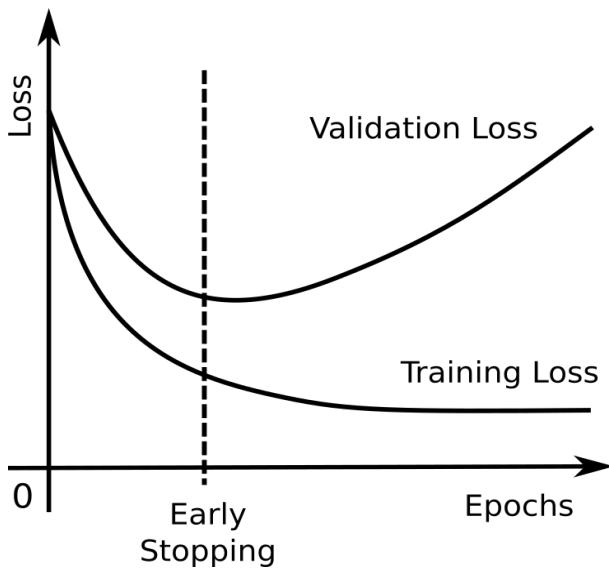
Μετασχηματισμός	Αποτέλεσμα
Αντιστροφή (Flipping)	Πραγματοποιεί αντιστροφή του περιεχομένου μιας εικόνας κατά μήκος του οριζόντιου ή κατακόρυφου άξονα
Περιστροφή (Rotation)	Πραγματοποιεί αλλαγή του προσανατολισμού μιας εικόνας μέσω της περιστροφής του περιεχομένου της προς τα αριστερά ή δεξιά
Μετατόπιση (Translation)	Πραγματοποιεί μετακίνηση του περιεχομένου της εικόνας οριζοντίως ή καθέτως
Περικοπή (Cropping)	Πραγματοποιεί αφαίρεση ενός μέρους του περιεχομένου της εικόνας ενώ ή υπόλοιπη εικόνα διατηρείται ως έχει

### 2.11.3 Πρόωρος Τερματισμός (Early Stopping)

Ο πρόωρος τερματισμός αποτελεί μέθοδο που χρησιμοποιείται συχνά στο πεδίο της μηχανικής μάθησης και συγκεκριμένα κατά τη διαδικασία εκπαίδευσης ενός νευρωνικού δικτύου ούτως ώστε να καθορίσει ποια είναι η κατάλληλη στιγμή για να διακοπεί η εκπαιδευτική διαδικασία [19]. Στόχος της εφαρμογής ενός αλγορίθμου πρόωρου τερματισμού είναι να τερματιστεί η διαδικασία πριν παρουσιασθεί το φαινόμενο της υπερπροσαρμογής του μοντέλου στα δεδομένα εκπαίδευσης [19].

Αυτό επιτυγχάνεται μέσω της παρακολούθησης της τιμής της συνάρτησης απώλειας. Εάν σε μια δεδομένη φάση της εκπαιδευτικής διαδικασίας η τιμή της συνάρτηση απώλειας γίνει μεγαλύτερη από τη τιμή που η συνάρτηση είχε στη προγενέστερη φάση, πραγματοποιείται διακοπή της εκπαιδευτικής διαδικασίας [19].

Στην Εικόνα 2.15 παρουσιάζεται το γράφημα ενός μοντέλου στο οποίο εφαρμόζεται η μέθοδος του πρόωρου τερματισμού πριν προλάβει να εμφανιστεί το φαινόμενο της υπερπροσαρμογής.



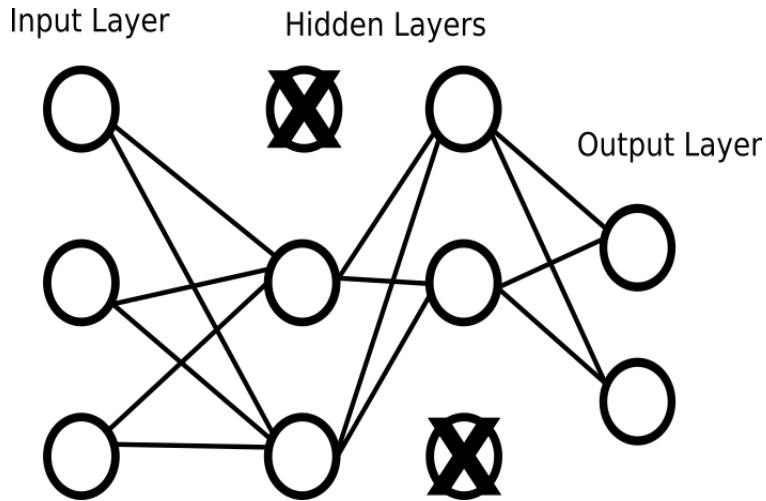
Εικόνα 2.15: Σημείο πρόωρου τερματισμού της εκπαίδευσης πριν εμφανισθεί υπερπροσαρμογή.

### 2.11.4 Επίπεδο Dropout

Κατά τη διαδικασία εκπαίδευσης ενός τεχνητού νευρωνικού δικτύου οι νευρώνες του δημιουργούν εξαρτήσεις μεταξύ τους [19], [45]. Αυτό σημαίνει ότι κάποιοι νευρώνες εξαρτώνται από χαρακτηριστικά που λαμβάνουν από άλλους νευρώνες ούτως ώστε να αναγνωρίσουν μοτίβα και σχέσεις που εμφανίζονται μεταξύ των δεδομένων και να εξάγουν επιτυχώς χαρακτηριστικά από αυτά [19]. Οι νευρώνες που

δημιουργούν εξαρτήσεις μεταξύ τους ενδεχομένως να αποδίδουν καλά στο σύνολο εκπαίδευσης, ώστόσο περιορίζουν την ικανότητα γενίκευσης του δικτύου σε νέα, άγνωστα δεδομένα.

Η εισαγωγή dropout επιπέδων στο δίκτυο αποτελεί λύση για το συγκεκριμένο θέμα. Ένα dropout επίπεδο εξαιρεί προσωρινά ένα τυχαίο ποσοστό των νευρώνων καθώς και των εισερχόμενων και εξερχόμενων συνδέσεών τους από ένα επίπεδο ενός νευρωνικού δικτύου [19], [45]. Το ποσοστό των νευρώνων που απενεργοποιεί ένα επίπεδο dropout αποτελεί υπερπαράμετρο που καθορίζεται πριν την έναρξη της διαδικασίας εκπαίδευσης [19]. Οι νευρώνες που απενεργοποιούνται μέσω του επιπέδου dropout, αγνοούνται από το δίκτυο [19]. Με την απενεργοποίηση ορισμένων από τους νευρώνες του δικτύου, οι νευρώνες που εξαρτώνται από την έξοδο των νευρώνων που απενεργοποιήθηκαν, αναγκάζονται να πραγματοποιούν την εξαγωγή χαρακτηριστικών από τα δεδομένα χωρίς να εξαρτώνται από τα χαρακτηριστικά που λαμβάνουν από άλλους νευρώνες, αφού ανά πάσα στιγμή κάποιος νευρώνας μπορεί να τεθεί εκτός λειτουργίας [19]. Στην Εικόνα 2.16 παρουσιάζεται το αποτέλεσμα της εφαρμογής επιπέδων dropout σε κρυφά επίπεδα ενός τεχνητού νευρωνικού δικτύου.



Εικόνα 2.16: Παράδειγμα εφαρμογής dropout σε ένα τεχνητό νευρωνικό δίκτυο

### 2.11.5 Επίπεδο Batch Normalization

Η διαδικασία εκπαίδευσης ενός νευρωνικού δικτύου επηρεάζεται από το γεγονός ότι η κατανομή των εισόδων του κάθε επιπέδου μεταβάλλεται κατά τη διάρκεια της εκπαίδευσης, καθώς μεταβάλλονται οι παράμετροι των προηγούμενων επιπέδων [46]. Αυτή η μεταβολή έχει ως αποτέλεσμα το κάθε επίπεδο να χρειάζεται να προσαρμόζεται συνεχώς στη νέα κατανομή [46]. Αυτό το φαινόμενο ονομάζεται internal covariate shift και προκαλεί καθυστέρηση στην εκπαίδευση διότι απαιτεί

χαμηλότερους ρυθμούς μάθησης και προσεκτική αρχικοποίηση των παραμέτρων [46].

To batch normalization προτάθηκε ως λύση του συγκεκριμένου φαινομένου [46]. Πιο συγκεκριμένα, αποτελεί επίπεδο που προστίθεται στα τεχνητά νευρωνικά δίκτυα και το οποίο πραγματοποιεί κανονικοποίηση (normalization) των εισόδων των επιπέδων ενός δικτύου [46]. Για να κανονικοποιήσει τα δεδομένα, ένα επίπεδο batch normalization αρχικά υπολογίζει τη μέση τιμή (mean) και τη διακύμανση (variance) των δεδομένων και στη συνέχεια πραγματοποιεί κανονικοποίηση των δεδομένων χρησιμοποιώντας τη μέση τιμή και τη διακύμανση που υπολόγισε [46]. Τέλος εφαρμόζει μετατόπιση (shift) και κλιμάκωση (scale) στα δεδομένα [46].

Η προσθήκη του συγκεκριμένου επιπέδου σε ένα νευρωνικό δίκτυο, επιτρέπει τη χρήση πολύ υψηλότερου ρυθμού μάθησης και κάνει την αρχικοποίηση των παραμέτρων λιγότερο σημαντική [46]. Η χρήση επιπέδων batch normalization σε ορισμένες περιπτώσεις βελτιώνει την ικανότητα γενίκευσης ενός μοντέλου, εξαλειφούντας την ανάγκη χρήσης επιπέδων dropout [46].

## 2.12 Μεταφορά Μάθησης (Transfer Learning)

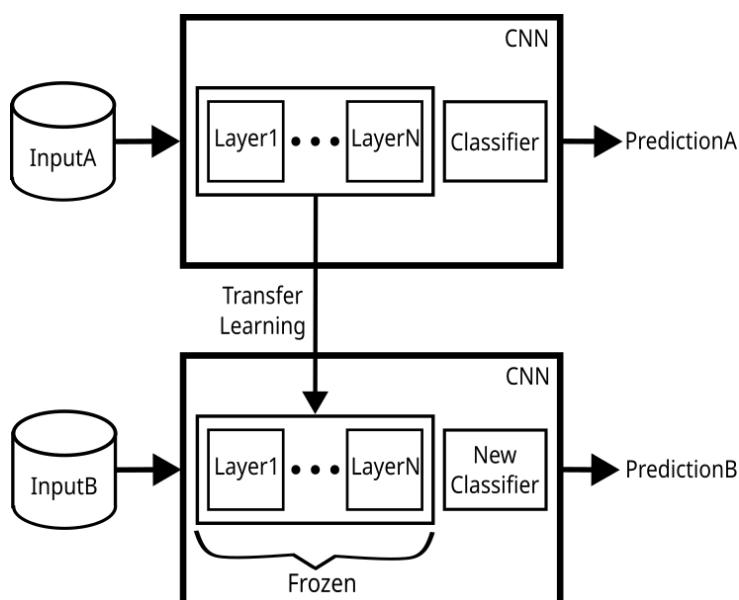
Στο πεδίο της μηχανικής μάθησης και ιδιαίτερα της βαθιάς μάθησης αποτελεί συχνό φαινόμενο τα δεδομένα που είναι κατάλληλα για να χρησιμοποιηθούν για την επίλυση ενός προβλήματος να μην επαρκούν [19]. Η συλλογή ενός ικανοποιητικού αριθμού από δεδομένα αποτελεί στις περισσότερες περιπτώσεις μια δύσκολη, κοστοβόρα και χρονοβόρα διαδικασία, ιδιαίτερα σε προβλήματα που υπάγονται στο πεδίο της ιατρικής όπως είναι το πρόβλημα που πραγματεύεται η παρούσα διπλωματική εργασία [19]. Κάτι τέτοιο έχει ως αποτέλεσμα οι παραδοσιακές προσεγγίσεις εκπαίδευσης μοντέλων βαθειάς μάθησης όπου ένα μοντέλο εκπαιδεύεται από το μηδέν για να επιλύσει ένα συγκεκριμένο πρόβλημα να μην μπορούν να εφαρμοσθούν, διότι απαιτούν έναν πολύ μεγάλο όγκο δεδομένων [19]. Επιπλέον, ακόμη και εάν είναι διαθέσιμος ένας μεγάλος όγκος δεδομένων, η εκπαίδευση ενός τεχνητού νευρωνικού δικτύου από το μηδέν αποτελεί μία υπολογιστικά ακριβή διαδικασία [19].

Η τεχνική της μεταφοράς μάθησης στοχεύει στο να δώσει λύση στο συγκεκριμένο πρόβλημα. Μία εφαρμογή της μεθόδου της μεταφοράς μάθησης φαίνεται στην Εικόνα 2.17. Η μεταφορά μάθησης αποτελεί τεχνική όπου συστήματα βαθιάς μάθησης, για την επίλυση ενός προβλήματος για το οποίο διαθέτουμε μικρό αριθμό δεδομένων, χρησιμοποιούν τη γνώση που έχει αποκομισθεί από την επίλυση ενός άλλου παραπλήσιου προβλήματος για το οποίο διατίθεται ένας μεγάλος όγκος δεδομένων [19]. Η αξιοποίηση αυτής της τεχνικής έχει ως αποτέλεσμα να παράγονται συστήματα τα οποία είναι πιο αξιόπιστα και αποδίδουν καλύτερα [19]. Υπάρχουν τρεις προσεγγίσεις μεταφοράς μάθησης:

Στην πρώτη προσέγγιση το προεκπαιδευμένο μοντέλο χρησιμοποιείται ως ταξινομητής [19]. Αυτό σημαίνει ότι δεν πραγματοποιείται εκ νέου εκπαίδευση του προεκπαιδευμένου μοντέλου [19]. Το προεκπαιδευμένο μοντέλο χρησιμοποιείται απευθείας για την ταξινόμηση των δεδομένων του προβλήματος για το οποίο έχει επιστρατευθεί [19]. Η συγκεκριμένη προσέγγιση χρησιμοποιείται σε περιπτώσεις όπου το νέο πρόβλημα είναι όμοιο με το πρόβλημα στο οποίο έχει προεκπαιδευθεί το μοντέλο [19].

Στη δεύτερη προσέγγιση το προεκπαιδευμένο μοντέλο χρησιμοποιείται ως εξαγωγέας χαρακτηριστικών (feature extractor) [19]. Αυτό σημαίνει ότι τα τελευταία επίπεδα του μοντέλου, που αποτελούν τον ταξινομητή (classifier), αφαιρούνται και προστίθενται νέα επίπεδα για τη προσαρμογή του μοντέλου στο νέο πρόβλημα που καλείται να επιλύσει [19]. Τα βάρη του εξαγωγέα χαρακτηριστικών του προεκπαιδευμένου μοντέλου «παγώνουν», δηλαδή δεν ενημερώνονται κατά τη διαδικασία της εκπαίδευσης και μόνο τα βάρη των νέων επιπέδων του ταξινομητή ενημερώνονται [19]. Αυτή η προσέγγιση είναι κατάλληλη σε περιπτώσεις όπου το νέο πρόβλημα είναι παρόμοιο με το αρχικό πρόβλημα [19].

Στην τρίτη προσέγγιση τα επίπεδα του ταξινομητή αντικαθίστανται με νέα και εκπαιδεύονται μαζί με τα επίπεδα του εξαγωγέα χαρακτηριστικών στα δεδομένα του νέου προβλήματος [19]. Το εάν θα εκπαιδευθούν όλα τα επίπεδα του εξαγωγέα χαρακτηριστικών ή όχι εξαρτάται από το πρόβλημα που καλείται να επιλύσει το δίκτυο καθώς και από τα διαθέσιμα δεδομένα. Τα πρώτα επίπεδα ενός μοντέλου βαθιειάς μάθησης εξάγουν χαμηλού επιπέδου χαρακτηριστικά τα οποία είναι κοινά σε ένα μεγάλο μέρος προβλημάτων οπότε ενδεχομένως να μη χρειάζεται να εκπαιδευθούν ξανά [19]. Η προαναφερθείσα προσέγγιση είναι κατάλληλη σε περιπτώσεις όπου το νέο πρόβλημα διαφέρει από το αρχικό [19].



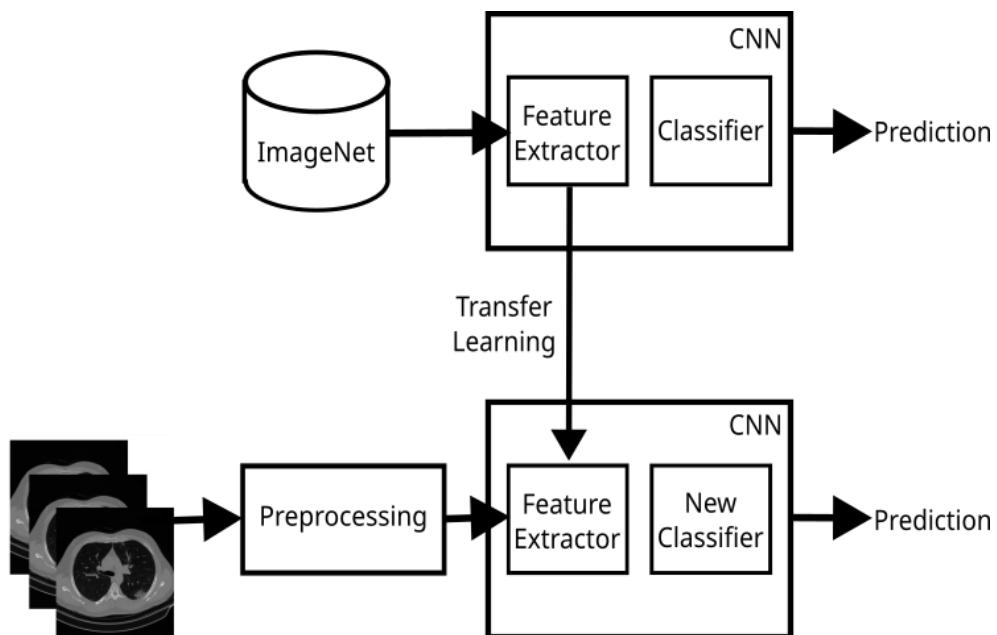
Εικόνα 2.17: Εφαρμογή της μεθόδου μεταφοράς μάθησης

### 3 Περιγραφή Μεθοδολογίας Πειράματος Ταξινόμησης Εικόνων

Στο πρώτο εκ των δύο πειραμάτων, που διεξάγονται στα πλαίσια εκπόνησης της παρούσας διπλωματικής εργασίας, στόχος είναι να μελετηθούν και να αξιολογηθούν μέθοδοι βαθειάς μάθησης. Συγκεκριμένα, αρχιτεκτονικές τεχνητών νευρωνικών δικτύων αξιολογούνται ως προς την ικανότητά τους, έπειτα από κατάλληλη εκπαίδευση, να ταξινομούν τομές αξονικών τομογραφιών. Απότερος σκοπός είναι τα εκπαιδευμένα μοντέλα να μπορούν να διακρίνουν τομές αξονικών τομογραφιών που εμφανίζουν αλλοιώσεις οφειλόμενες σε πνευμονία προερχόμενη από τον ιό SARS-CoV-2.

Η πειραματική διαδικασία που ακολουθείται και παρουσιάζεται στην Εικόνα 3.1, αποτελείται από τα ακόλουθα βήματα:

- 1) Επιλογή κατάλληλων συνόλων δεδομένων, σύμφωνα με τις ανάγκες του πειράματος ταξινόμησης των τομών αξονικών τομογραφιών.
- 2) Επεξεργασία των δεδομένων, πριν αυτά τροφοδοτηθούν στο εκάστοτε μοντέλο τεχνητών νευρωνικών δικτύων.
- 3) Επιλογή κατάλληλων μοντέλων, εκπαίδευση και επικύρωση τους χρησιμοποιώντας σύνολα δεδομένων και ακολουθώντας τη μέθοδο της μεταφοράς μάθησης.
- 4) Αξιολόγηση, έπειτα από το πέρας της εκπαίδευτικής διαδικασίας, της ικανότητας των μοντέλων να ταξινομούν τομές αξονικών τομογραφιών.  
(Η αξιολόγηση των μοντέλων πραγματοποιείται με τη βοήθεια κατάλληλων μετρικών αξιολόγησης).



Εικόνα 3.1: Διαδικασία εκπαίδευσης και αξιολόγησης μοντέλου ως προς τη ταξινόμηση τομών με βάση αλλοιώσεις από πνευμονία COVID-19, όπου χρησιμοποιείται μεταφορά μάθησης

### **3.1 Επιλογή Συνόλου Δεδομένων Πειράματος Ταξινόμησης Εικόνων (Image Classification)**

Ένας από τους σημαντικότερους παράγοντες που συμβάλουν στην επιτυχή εκπαίδευση ενός μοντέλου μηχανικής μάθησης αποτελεί η επιλογή κατάλληλων δεδομένων. Στα προβλήματα επιβλεπόμενης μάθησης, όπως είναι αυτά που καλούμαστε να μελετήσουμε στη παρούσα διπλωματική εργασία, μας ενδιαφέρει τα δεδομένα να είναι όσο το δυνατόν περισσότερα, όσο το δυνατόν πιο ποικίλα, να έχουν την υψηλότερη δυνατή ποιότητα, καθώς επίσης και να είναι έγκυρα, ορθά κατηγοριοποιημένα, και ισοκατανεμημένα στις κλάσεις του συνόλου δεδομένων. Όσο περισσότερα από τα προαναφερθέντα κριτήρια πληρούν τα δεδομένα που χρησιμοποιούμε, τόσο καλύτερη πρόκειται να είναι η εκπαίδευση του εκάστοτε μοντέλου που έχουμε επιλέξει να αξιοποιήσουμε και κατ' επέκταση η γενίκευσή του σε νέα δεδομένα. Η εύρεση δεδομένων που να πληρούν όλα τα παραπάνω κριτήρια αποτελεί δύσκολη υπόθεση στο πεδίο της μηχανικής μάθησης. Αυτό γίνεται ακόμη περισσότερο αντιληπτό σε νέα προβλήματα όπως είναι αυτό που μελετάται στη παρούσα εργασία. Πιο συγκεκριμένα, τα σύνολα δεδομένων που είναι συναφή με το πρόβλημα της ανάλυσης ιατρικών εικόνων για την υποβοήθηση στη διάγνωση COVID-19, τα οποία διατίθενται μέχρι στιγμής ελεύθερα στο διαδίκτυο είναι λίγα σε αριθμό. Επιπλέον, κάθε αξονική τομογραφία αποτελείται από ένα μεγάλο αριθμό από συνεχόμενες τομές που παρουσιάζουν ελάχιστες διαφορές μεταξύ τους, όσον αφορά τα ευρήματα που εμφανίζονται εντός του πνευμονικού παρεγχύματος. Αυτό σημαίνει ότι σε κάθε ασθενή αντιστοιχεί ένας μεγάλος αριθμός από εικόνες με ελάχιστες διαφορές μεταξύ τους. Επιπρόσθετα, τα δεδομένα σε αρκετά σύνολα δεδομένων δεν είναι ορθά κατηγοριοποιημένα, η ποιότητα ορισμένων συνόλων δεδομένων δεν είναι αρκετά ικανοποιητική, ενώ ένας σημαντικός αριθμός από σύνολα δεδομένων αποτελείται από μη ισοκατανεμημένα δεδομένα. Οι προαναφερθείσες παρατηρήσεις αποτελούν ορισμένες από τις σημαντικότερες προκλήσεις που παρουσιάζει η μελέτη του θέματος της παρούσας διπλωματικής εργασίας.

Για την επίλυσή του προβλήματος, που αναφέρεται στο παρόν κεφάλαιο, επιλέχθηκαν σύνολα δεδομένων τα οποία αποτελούνται από αξονικές τομογραφίες. Άλλες επιλογές θα ήταν να χρησιμοποιηθούν ως σύνολα δεδομένων ακτινογραφίες θώρακος, είτε τόσο ακτινογραφίες θώρακος όσο και αξονικές τομογραφίες.

Οι λόγοι που έγινε η επιλογή να χρησιμοποιηθούν μόνο σύνολα δεδομένων που αποτελούνται από αξονικές τομογραφίες είναι οι ακόλουθοι:

Στην πλειοψηφία τους τα σύνολα δεδομένων που διατίθενται δωρεάν στο διαδίκτυο και αποτελούνται από ακτινογραφίες θώρακος έχουν προκύψει από τη συνένωση άλλων συνόλων δεδομένων τα οποία έχουν προέλθει από διαφορετικές πηγές χωρίς να έχει δοθεί προσοχή για τυχόν πανομοιότυπες εικόνες. Αυτό έχει ως αποτέλεσμα αφενός στα περισσότερα από τα προαναφερθέντα σύνολα να εμφανίζονται οι ίδιες εικόνες περισσότερες από μία φορές και αφετέρου να υπάρχουν

μεγάλες διαφορές τόσο ως προς τη ποιότητα όσο και ως προς τον τρόπο λήψης των δεδομένων από τα ιατρικά μηχανήματα.

Από την άλλη, τα περισσότερα σύνολα δεδομένων που αποτελούνται από αξονικές τομογραφίες είναι πολύ καλύτερα ποιοτικά και η συλλογή τους έχει πραγματοποιηθεί με πιο προσεκτικό τρόπο.

Επιπρόσθετα, σε ορισμένα πειράματα που πραγματοποιήθηκαν στα πρώτα στάδια εκπόνησης της παρούσας διπλωματικής εργασίας, παρατηρήθηκε ότι τα μοντέλα βαθιάς μάθησης που εκπαιδεύθηκαν πάνω σε σύνολα που αποτελούνταν από ακτινογραφίες θώρακος, κατηγοριοποιούσαν τις ιατρικές εικόνες των συγκεκριμένων συνόλων δεδομένων βασιζόμενα σε διαφορές που είχαν να κάνουν με τη ποιότητα τους και τον τρόπο λήψης τους από τα ιατρικά μηχανήματα και όχι σύμφωνα με τα ευρήματα που εμφανίζονταν εντός του πνευμονικού παρεγχύματος. Αυτό είχε ως αποτέλεσμα η αρχική ιδέα να πραγματοποιηθούν πειράματα κάνοντας χρήση τόσο συνόλων δεδομένων που αποτελούνταν από αξονικές τομογραφίες όσο και συνόλων που αποτελούνταν από ακτινογραφίες θώρακος να εγκαταλειφθεί γρήγορα.

Για τη διεξαγωγή του πειράματος του παρόντος κεφαλαίου επιλέχθηκε να αξιοποιηθούν τα σύνολα δεδομένων COVID-CT-MD και China Consortium of Chest CT Image Investigation (CC-CCII) Dataset [47].

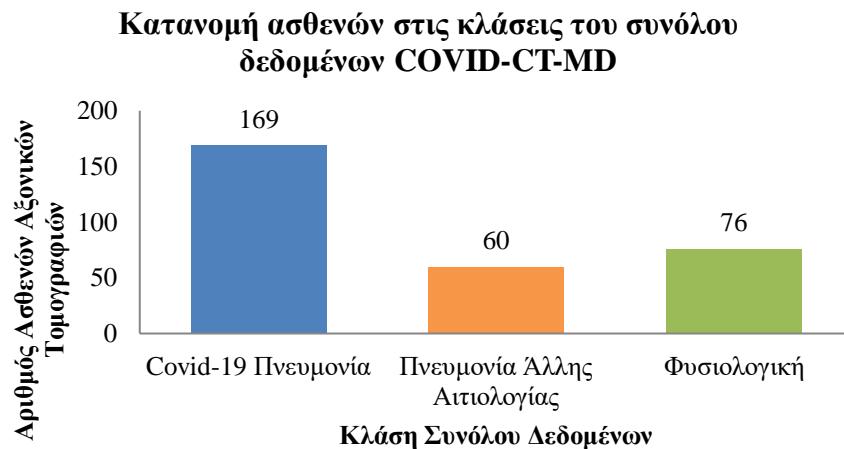
Πιο συγκεκριμένα, το σύνολο δεδομένων COVID-CT-MD αξιοποιήθηκε στα πλαίσια του πειράματος τόσο στη φάση της εκπαίδευσης (training phase) όσο και στη φάση της επικύρωσης (validation phase) των μοντέλων.

Το σύνολο δεδομένων CC-CCII αξιοποιήθηκε ως εξωτερικό σύνολο αξιολόγησης στα πλαίσια του πειράματος του συγκεκριμένου κεφαλαίου, για την αξιολόγηση της ικανότητας γενίκευσης των μοντέλων που εκπαιδεύθηκαν.

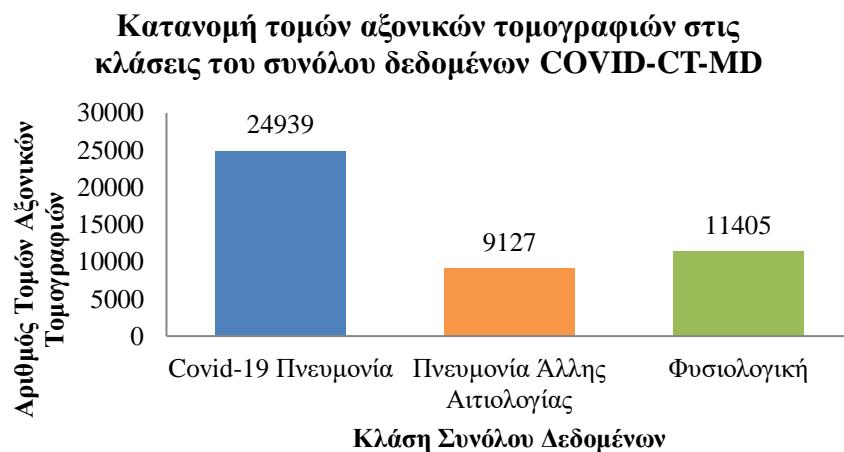
### 3.1.1 Σύνολο Δεδομένων COVID-CT-MD

Ένα από τα σύνολα δεδομένων, που αξιοποιήθηκαν για τις ανάγκες του παρόντος πειράματος είναι το σύνολο δεδομένων COVID-CT-MD. Το συγκεκριμένο σύνολο δεδομένων αποτελείται από 45.471 τομές αξονικών τομογραφιών που έχουν συλλεχθεί από 305 ασθενείς στο Απεικονιστικό Κέντρο Babak της Τεχεράνης [16]. Από τις 45.471 τομές αξονικών τομογραφιών, οι 24.939 αντιστοιχούν σε 169 ασθενείς και εμφανίζουν αλλοιώσεις οφειλόμενες σε πνευμονία που προέρχεται από τον ιό SARS-CoV-2, οι 9.127 αντιστοιχούν σε 60 ασθενείς και εμφανίζουν αλλοιώσεις οφειλόμενες σε πνευμονία που δεν προέρχεται από τον ιό SARS-CoV-2, ενώ οι 11.405 αντιστοιχούν σε 76 ασθενείς και δεν εμφανίζουν αλλοιώσεις [16].

Στην Εικόνα 3.2 παρουσιάζεται ένα ραβδόγραμμα της κατανομής των ασθενών στις κλάσεις του συνόλου COVID-CT-MD και στην Εικόνα 3.3 παρουσιάζεται ένα ραβδόγραμμα της κατανομής των τομών στις κλάσεις του ίδιου συνόλου.



**Εικόνα 3.2:** Ραβδόγραμμα κατανομής των ασθενών, στους οποίους ανήκουν οι τομές αξονικών τομογραφιών, στις κλάσεις του συνόλου δεδομένων COVID-CT-MD

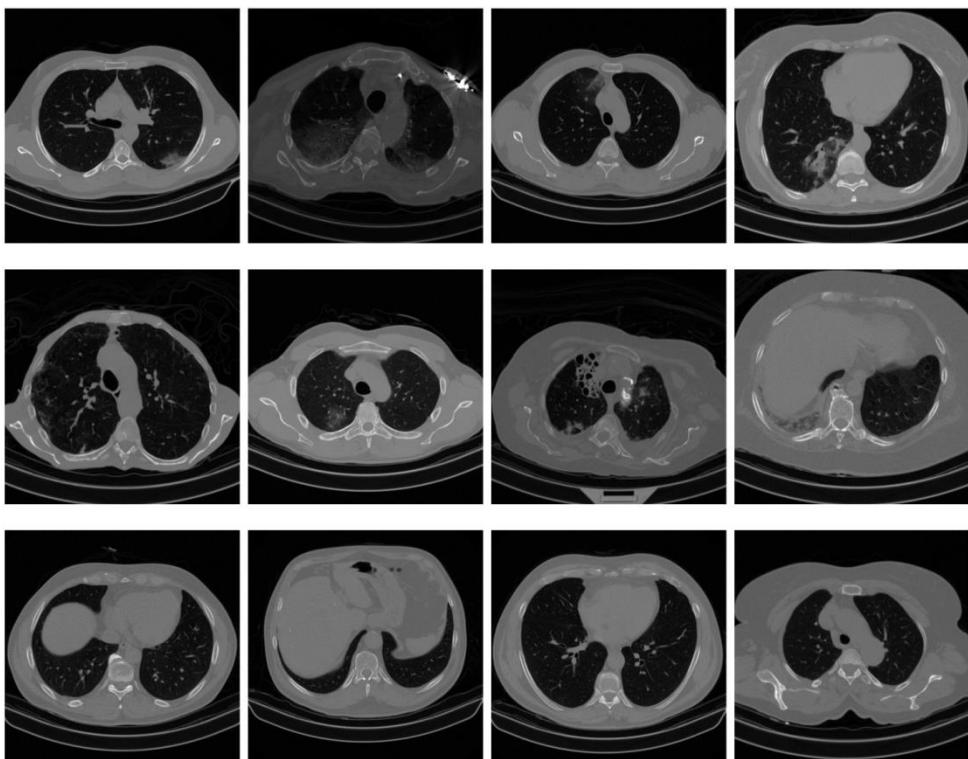


**Εικόνα 3.3:** Ραβδόγραμμα κατανομής των τομών αξονικών τομογραφιών στις κλάσεις του συνόλου δεδομένων COVID-CT-MD

Στα πλαίσια του πειράματος που περιγράφεται στο παρόν κεφάλαιο, επιλέχθηκε να διερευνηθεί η αποδοτικότητα μοντέλων βαθιάς μάθησης σε δύο περιπτώσεις δυαδικής ταξινόμησης. Στην ταξινόμηση τομών που εμφανίζουν πνευμονία προερχόμενη από τον ιό SARS-CoV-2 από τομές που δεν εμφανίζουν πνευμονία και στην ταξινόμηση τομών που εμφανίζουν πνευμονία προερχόμενη από τον ιό SARS-CoV-2 από τομές που εμφανίζουν πνευμονία άλλης αιτιολογίας.

Επομένως, για τις ανάγκες του συγκεκριμένου πειράματος επιλέχθηκαν να αξιοποιηθούν από το σύνολο δεδομένων COVID-CT-MD αξονικές τομογραφίες ασθενών με πνευμονία από COVID-19, αξονικές τομογραφίες ασθενών με πνευμονία άλλης αιτιολογίας, καθώς και φυσιολογικές αξονικές τομογραφίες. Οι τομές του συνόλου δεδομένων που αντιστοιχούν σε κάθε ασθενή στο συγκεκριμένο σύνολο είναι αποθηκευμένες σε μορφότυπο Digital Imaging and Communications in Medicine (DICOM).

Στην Εικόνα 3.4 παρουσιάζεται ένα υποσύνολο από τις τομές αξονικών τομογραφιών του συνόλου δεδομένων COVID-CT-MD.



Εικόνα 3.4: Υποσύνολο από τις τομές αξονικών τομογραφιών του συνόλου δεδομένων COVID-CT-MD

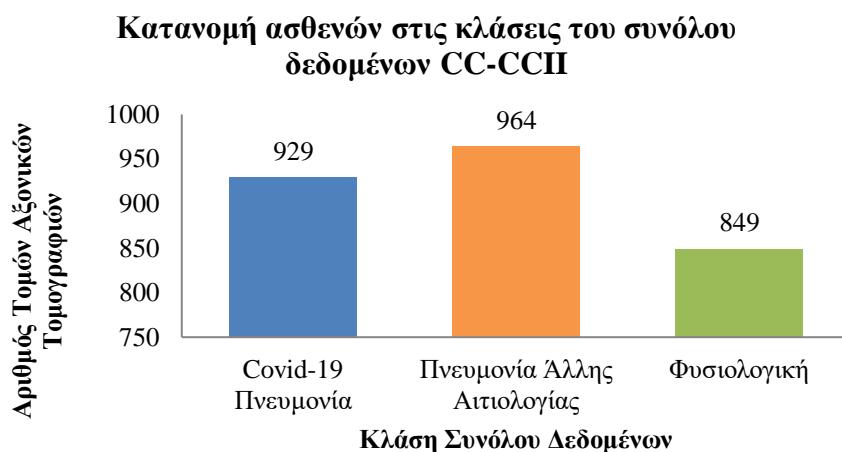
### 3.1.2 Σύνολο Δεδομένων CC-CCII

Ένα δεύτερο σύνολο δεδομένων που αξιοποιήθηκε για τις ανάγκες του πειράματος του παρόντος κεφαλαίου αποτελεί το σύνολο δεδομένων CC-CCII.

Το συγκεκριμένο σύνολο δεδομένων αποτελείται από 411.529 τομές αξονικών τομογραφιών που έχουν προέλθει από 2.742 ασθενείς [47]. Από τις 411.529 τομές αξονικών τομογραφιών, οι 156.071 αντιστοιχούν σε 929 ασθενείς και εμφανίζουν αλλοιώσεις οφειλόμενες σε πνευμονία που προέρχεται από τον ιό SARS-CoV-2, οι

159.702 αντιστοιχούν σε 964 ασθενείς και εμφανίζουν αλλοιώσεις οφειλόμενες σε πνευμονία που δεν προέρχεται από τον ιό SARS-CoV-2, ενώ οι 95.756 αντιστοιχούν σε 849 ασθενείς και δεν εμφανίζουν αλλοιώσεις [47].

Στην Εικόνα 3.5 παρουσιάζεται ένα ραβδόγραμμα της κατανομής των ασθενών στις κλάσεις του συνόλου CC-CCII και στην Εικόνα 3.6 παρουσιάζεται ένα ραβδόγραμμα της κατανομής των τομών στις κλάσεις του ίδιου συνόλου.



**Εικόνα 3.5:** Ραβδόγραμμα κατανομής των ασθενών, στους οποίους ανήκουν οι τομές αξονικών τομογραφιών, στις κλάσεις του συνόλου δεδομένων CC-CCII

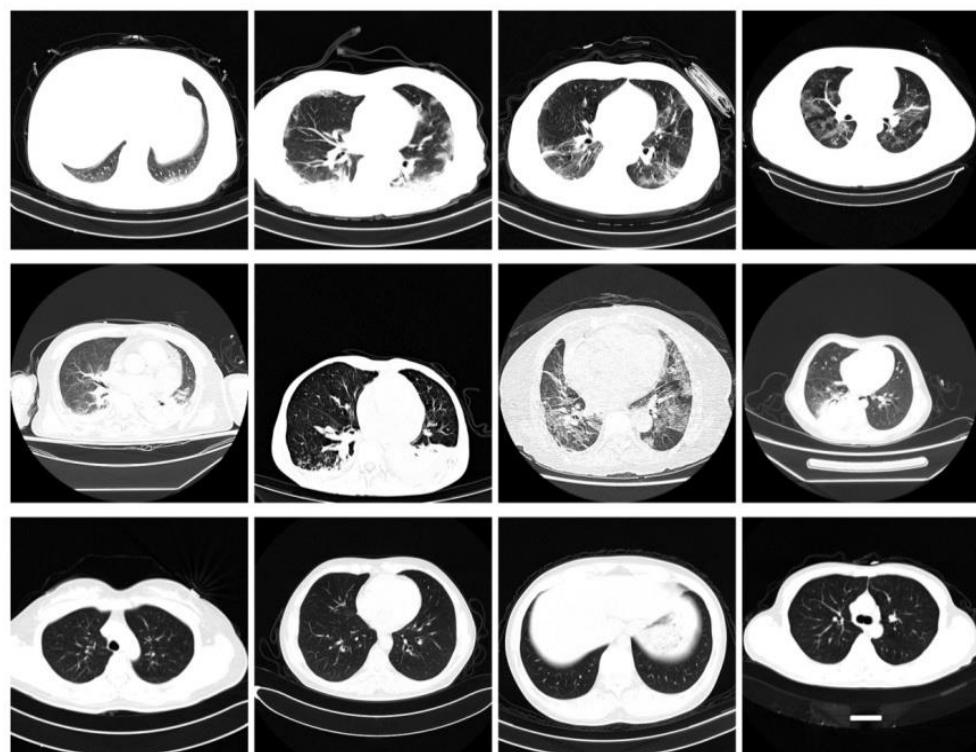


**Εικόνα 3.6:** Ραβδόγραμμα κατανομής των τομών αξονικών τομογραφιών στις κλάσεις του συνόλου δεδομένων CC-CCII

Για τις ανάγκες του πειράματος αξιοποιούνται από το σύνολο δεδομένων CC-CCII αξονικές τομογραφίες με πνευμονία από SARS-CoV-2, αξονικές τομογραφίες με

πνευμονία άλλης αιτιολογίας, καθώς και φυσιολογικές αξονικές τομογραφίες. Οι τομές του συνόλου δεδομένων που αντιστοιχούν σε κάθε ασθενή στο συγκεκριμένο σύνολο, είναι αποθηκευμένες σε μορφότυπο PNG, JPG και BMP.

Στην Εικόνα 3.7 παρουσιάζεται ένα υποσύνολο από τις τομές του συνόλου δεδομένων CC-CCII.



Εικόνα 3.7: Υποσύνολο από τις τομές αξονικών τομογραφιών του συνόλου δεδομένων CC-CCII

## 3.2 Προεπεξεργασία Δεδομένων

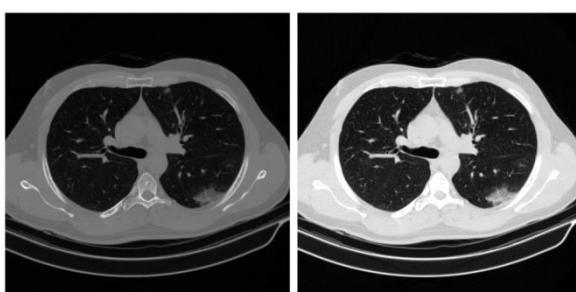
Με τον όρο προεπεξεργασία δεδομένων εννοούμε όλες εκείνες τις τεχνικές που εφαρμόζονται στα ακατέργαστα δεδομένα ενός συνόλου πριν αυτά διοχετευθούν ως είσοδος σε έναν αλγόριθμο μηχανικής μάθησης. Σκοπός της προεπεξεργασίας είναι η βελτίωση της ποιότητας των δεδομένων, η μείωση του χρόνου εκπαίδευσης, καθώς και η παραγωγή καλύτερων αποτελεσμάτων.

Για τις ανάγκες του πειράματος του παρόντος κεφαλαίου επιστρατεύεται ένα σύνολο από τεχνικές προεπεξεργασίας δεδομένων οι οποίες περιγράφονται στις επόμενες παραγράφους.

### 3.2.1 Κανονικοποίηση Αντίθεσης (Contrast) και Φωτεινότητας (Brightness) Τομών Αξονικών Τομογραφιών

Παρατηρώντας κανείς τα σύνολα δεδομένων που χρησιμοποιούνται στο παρόν κεφάλαιο θα διαπιστώσει ότι παρουσιάζουν μεγάλες διαφορές ως προς την αντίθεση (contrast) και τη φωτεινότητα (brightness). Οι διαφορές αυτές δυσχεραίνουν την ανίχνευση των χαρακτηριστικών που είναι απαραίτητα για την ορθή σημασιολογική κατάτμηση των εικόνων από τα μοντέλα βαθειάς μάθησης. Μια λύση στο πρόβλημα αυτό αποτελεί η κανονικοποίηση των εικόνων ως προς την αντίθεση και τη φωτεινότητα, υιοθετώντας συγκεκριμένες τιμές αντίθεσης και φωτεινότητας για όλες τις εικόνες. Αυτό μπορεί να υλοποιηθεί εφαρμόζοντας το παράθυρο πνεύμονα, μια τεχνική που περιγράφεται στο Κεφάλαιο 2 της παρούσας εργασίας, δηλαδή ρυθμίζοντας την τιμή των παραμέτρων μήκος παραθύρου (window width) και επίπεδο παραθύρου (window level) στα 1400 Hounsfield Units (HU) και -500 Hounsfield Units (HU) αντίστοιχα. Οι τιμές αυτές επιλέγονται ώστε το εσωτερικό του πνευμονικού παρεγχύματος να είναι καλύτερα ευδιάκριτο. Δυστυχώς, αυτό το στάδιο προεπεξεργασίας δύναται να εφαρμοσθεί μόνο στο σύνολο δεδομένων COVID-CT-MD, διότι οι τομές αξονικών τομογραφιών του συνόλου δεδομένων CC-CCII είναι αποθηκευμένες σε επεκτάσεις PNG, JPG, BMP και για την εφαρμογή του παραθύρου πνεύμονα είναι αναγκαίο οι εικόνες να είναι αποθηκευμένες σε μορφότυπο DICOM. Για να πραγματοποιηθεί η προαναφερθείσα μετατροπή, αρχικά ελέγχεται ότι οι τομές της εκάστοτε αξονικής τομογραφίας είναι εκφρασμένες σε μονάδες της κλίμακας Hounsfield. Αφού οι εικόνες αξονικών τομογραφιών του συνόλου COVID-CT-MD έχουν μορφότυπο DICOM, ο έλεγχος αυτός μπορεί να πραγματοποιεί παρατηρώντας τις τιμές των πεδίων Rescale\_Slope και Rescale\_Intercept της επικεφαλίδας του κάθε αρχείου DICOM. Έπειτα, σειρά έχει η μετατροπή του παραθύρου πνεύμονα στο κατάλληλο εύρος ούτως ώστε να είναι ευδιάκριτα τα επιθυμητά στοιχεία. Για τη μετατροπή αυτή, χρησιμοποιούνται οι τιμές των πεδίων Window\_Center και Window\_Width της επικεφαλίδας του κάθε αρχείου DICOM.

Στην Εικόνα 3.8 φαίνεται το αποτέλεσμα της μετατροπής σε μία από τις τομές του συνόλου COVID-CT-MD.



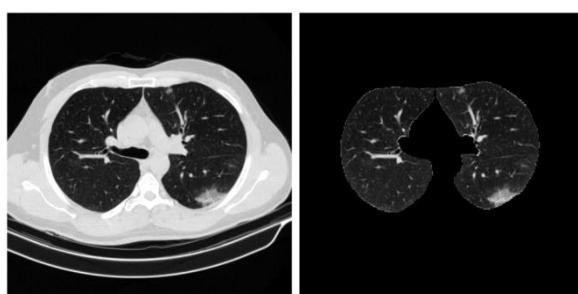
Εικόνα 3.8: Τομή αξονικής τομογραφίας του συνόλου COVID-CT-MD πριν και μετά την κανονικοποίηση αντίθεσης και φωτεινότητας

### **3.2.2 Δημιουργία παραλλαγής του συνόλου δεδομένων όπου διατηρείται μόνο το πνευμονικό παρέγχυμα στις τομές αξονικών τομογραφιών**

Η περιοχή των τομών των αξονικών τομογραφιών που ελέγχεται για την ύπαρξη λοίμωξης από τον ιό SARS-CoV-2 είναι το πνευμονικό παρέγχυμα. Μια πρακτική που δύναται να συμβάλλει στην αύξηση της αποδοτικότητας των μοντέλων βαθιάς μάθησης που αξιοποιούνται στην παρούσα εργασία αποτελεί η εφαρμογή τεχνικών κατάτμησης των τομών ούτως ώστε να διατηρείται στην εκάστοτε τομή μόνο το πνευμονικό παρέγχυμα. Για το σκοπό αυτό αξιοποιείται το πακέτο που έχει δημιουργηθεί από τους Hofmanninger κ.α. [48] και το οποίο πραγματοποιεί αυτοματοποιημένη κατάτμηση του πνευμονικού παρεγχύματος των τομών αξονικών τομογραφιών. Η μέθοδος που περιλαμβάνει το συγκεκριμένο πακέτο λειτουργεί σε δεδομένα αποθηκευμένα σε μορφότυπο DICOM (σύνολο COVID-CT-MD). Επομένως, είναι αδύνατη η εφαρμογή της συγκεκριμένης μεθόδου στις τομές του συνόλου δεδομένων CC-CCII.

Έπειτα από την εφαρμογή του συγκεκριμένου βήματος προεπεξεργασίας διαθέτουμε δύο παραλλαγές του συνόλου δεδομένων COVID-CT-MD που θα χρησιμοποιηθούν και οι δύο κατά τη διαδικασία εκπαίδευσης των μοντέλων. Η μία από τις δύο παραλλαγές αποτελείται από το σύνολο δεδομένων COVID-CT-MD πριν την εφαρμογή του προαναφερθέντος βήματος προεπεξεργασίας, ενώ η άλλη αποτελείται από τομές στις οποίες απεικονίζεται μόνο το εσωτερικό του πνευμονικού παρεγχύματος.

Στην Εικόνα 3.9 φαίνεται το αποτέλεσμα της κατάτμησης του πνευμονικού παρεγχύματος σε μία από τις τομές του συνόλου COVID-CT-MD.



**Εικόνα 3.9: Τομή αξονικής τομογραφίας του συνόλου COVID-CT-MD πριν και μετά την κατάτμηση του πνευμονικού παρεγχύματος**

### **3.2.3 Μετατροπή εικόνων από μορφότυπο DICOM σε μορφότυπο TIFF**

Όπως έχει ήδη αναφερθεί, κάθε αξονική τομογραφία ενός ασθενούς αποτελείται από πολλές τομές οι οποίες όσον αφορά το σύνολο COVID-CT-MD είναι οργανωμένες σε μορφότυπο DICOM. Για να τροφοδοτηθούν οι τομές στο εκάστοτε νευρωνικό δίκτυο προς εκπαίδευση θα πρέπει πρώτα να πραγματοποιηθεί εξαγωγή των τομών από τα αρχεία DICOM και αποθήκευσή τους σε αρχεία με επεκτάσεις που αντιστοιχούν σε αρχεία εικόνας. Επιλέχθηκε οι τομές των αξονικών τομογραφιών να αποθηκευθούν σε αρχεία μορφότυπου TIFF.

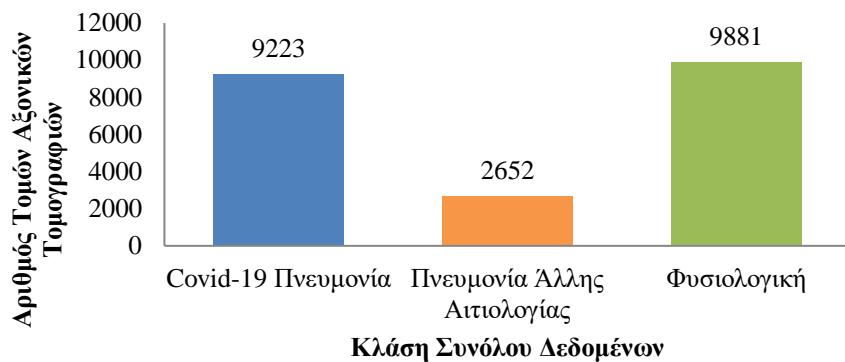
### **3.2.4 Αφαίρεση των τομών στις οποίες δεν διακρίνεται το πνευμονικό παρέγχυμα ή που δεν εμφανίζουν αλλοιώσεις**

Όσον αφορά τις τομές των αξονικών τομογραφιών του συνόλου δεδομένων COVID-CT-MD, στις πρώτες και στις τελευταίες τομές των αξονικών τομογραφιών δεν διακρίνεται το πνευμονικό παρέγχυμα του ασθενούς. Οι συγκεκριμένες τομές δεν έχουν καμία αξία για το πείραμα του παρόντος κεφαλαίου και γι' αυτό το λόγο αφαιρούνται από το σύνολο δεδομένων.

Επίσης, παρατηρώντας κανείς τις τομές αξονικών τομογραφιών ασθενών με πνευμονία από SARS-CoV-2, καθώς και ασθενών με πνευμονία άλλης αιτιολογίας, θα διαπιστώσει πως δεν εμφανίζουν όλες οι τομές αλλοιώσεις. Η διατήρηση των συγκεκριμένων τομών εντός του συνόλου δεδομένων προκαλεί σύγχυση στα εκπαιδεύμενα μοντέλα οδηγώντας τα σε λάθος συμπεράσματα και δυσχεραίνοντας με αυτό το τρόπο τη διαδικασία της εκπαίδευσης. Για να αποφευχθεί κάτι τέτοιο, πραγματοποιείται αφαίρεση των συγκεκριμένων τομών από το σύνολο δεδομένων πριν αυτό διοχετευθεί στο εκάστοτε μοντέλο βαθιάς μάθησης.

Πραγματοπιώντας τις παραπάνω μεταβολές στο σύνολο δεδομένων COVID-CT-MD, οι τομές που αφορούν ασθενείς που έχουν πνευμονία από ιό SARS-CoV2 μειώνονται στις 9.223 τομές, οι τομές αυτών με πνευμονία άλλης αιτιολογίας μειώνονται στις 2.652 τομές, και οι τομές φυσιολογικών αξονικών τομογραφιών μειώνονται στις 9.881. Η κατανομή των τομών έπειτα από τις προαναφερθείσες μεταβολές παρουσιάζεται στην Εικόνα 3.10.

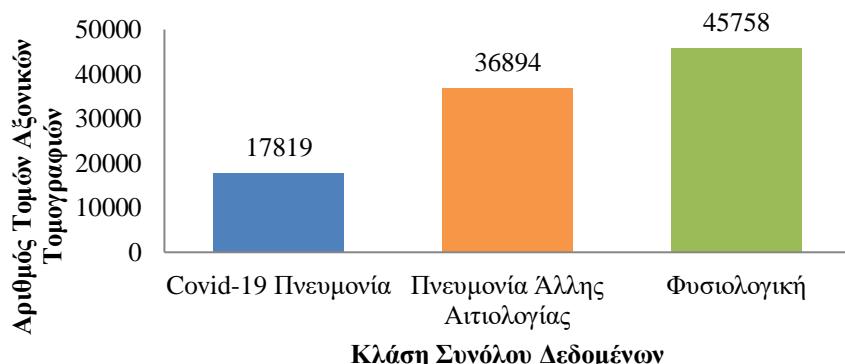
### Κατανομή τομών αξονικών τομογραφιών στις κλάσεις του συνόλου δεδομένων COVID-CT-MD



**Εικόνα 3.10:** Ραβδόγραμμα κατανομής των τομών αξονικών τομογραφιών στις κλάσεις του συνόλου δεδομένων COVID-CT-MD έπειτα από την αφαίρεση των τομών

Με τις παραπάνω μεταβολές στο σύνολο δεδομένων CC-CCII, οι τομές που αφορούν ασθενείς που έχουν πνευμονία από το SARS-CoV2 μειώνονται στις 17.819 τομές, οι τομές αυτών με πνευμονία άλλης αιτιολογίας μειώνονται στις 36.894 τομές, και οι τομές φυσιολογικών αξονικών τομογραφιών μειώνονται στις 45.758, όπως φαίνεται στην Εικόνα 3.11.

### Κατανομή τομών αξονικών τομογραφιών στις κλάσεις του συνόλου δεδομένων CC-CCII



**Εικόνα 3.11:** Ραβδόγραμμα κατανομής των τομών αξονικών τομογραφιών στις κλάσεις του συνόλου δεδομένων CC-CCII έπειτα από την αφαίρεση των τομών

### 3.2.5 Αλλαγή Διαστάσεων Εικόνας

Ένας σημαντικός παράγοντας που πρέπει να καθοριστεί πριν την έναρξη της διαδικασίας της εκπαίδευσης αποτελεί η επιλογή των διαστάσεων των εικόνων που πρόκειται να τροφοδοτηθούν στα μοντέλα βαθειάς μάθησης. Όσο μεγαλύτερες είναι οι διαστάσεις των εικόνων που τροφοδοτούνται, τόσο περισσότερο χώρο καταλαμβάνουν στη μνήμη και επιπλέον τόσο περισσότερος χρόνος και υπολογιστική ισχύς απαιτείται ούτως ώστε να πραγματοποιηθεί η διαδικασία της εκπαίδευσης. Οι διαστάσεις των εικόνων θα πρέπει να επιλεγούν με τέτοιο τρόπο ούτως ώστε να διασφαλίζεται ότι μπορεί να τοποθετηθεί ένας ικανός αριθμός από αυτές στη μνήμη, καθώς και ότι ο χρόνος που απαιτείται για την εκπαίδευση βρίσκεται εντός λογικών πλαισίων. Για το σκοπό αυτό επιλέγεται οι διαστάσεις όλων των εικόνων να μετατραπούν στις διαστάσεις 256x256 πριν αυτές τροφοδοτηθούν στο εκάστοτε τεχνητό νευρωνικό δίκτυο.

### 3.2.6 Κανονικοποίηση Τιμών Εικονοστοιχείων στο εύρος 0-1

Όσον αφορά τις εικόνες των συνόλων δεδομένων που χρησιμοποιούνται για τις ανάγκες του πειράματος του συγκεκριμένου κεφαλαίου, αυτές αποτελούνται από εικονοστοιχεία τα οποία λαμβάνουν ως τιμές ακέραιους αριθμούς που βρίσκονται εντός του εύρους 0 έως 255. Η τροφοδότηση εικόνων με εικονοστοιχεία που αποτελούνται από υψηλές αριθμητικές τιμές στα μοντέλα βαθειάς μάθησης οδηγεί τα μοντέλα στην ανάγκη να πραγματοποιούν πιο σύνθετους υπολογισμούς. Μια καλή πρακτική αποτελεί το να κανονικοποιηθούν οι τιμές των εικονοστοιχείων της κάθε εικόνας μέσω της μετατροπής τους από το εύρος τιμών 0 έως 255, στο εύρος τιμών 0 έως 1. Αυτό μπορεί να πραγματοποιηθεί μέσω της διαίρεσης κάθε εικονοστοιχείου με τον ακέραιο αριθμό 255. Το μικρότερο εύρος τιμών έχει ως αποτέλεσμα γρηγορότερη σύγκλιση του μοντέλου, αυξημένη επίδοσή του, καθώς και ευκολότερους και ταχύτερους υπολογισμούς.

### 3.2.7 One-Hot Encoding

Η τεχνική one-hot encoding αποτελεί συχνά χρησιμοποιούμενη τεχνική στο πεδίο της μηχανικής μάθησης. Σε ένα πρόβλημα ταξινόμησης όπως είναι αυτό που παρουσιάζεται στο παρόν κεφάλαιο, τα δεδομένα του συνόλου δεδομένων είναι κατηγορικά, δηλαδή είναι δεδομένα που προέρχονται από μεταβλητές, οι τιμές των οποίων εκφράζουν κλάσεις ή κατηγορίες [49]. Με τη τεχνική one-hot encoding τα κατηγορικά δεδομένα αναπαρίστανται ως διανύσματα, δηλαδή σε μορφή που μπορεί να χρησιμοποιηθεί αποτελεσματικά από αλγόριθμους μηχανικής μάθησης και κατ'

επέκταση μοντέλα τεχνητών νευρωνικών δικτύων [49]. Ένα παράδειγμα εφαρμογής της τεχνικής one-hot encoding φαίνεται στην Εικόνα 3.12.

Class	
Covid-19	
Normal	

One-Hot Encoding →

	Covid-19	Normal
Covid-19	1	0
Normal	0	1

**Εικόνα 3.12: Παράδειγμα one-hot encoding**

Από την Εικόνα 3.12 γίνεται αντιληπτό πως στο συγκεκριμένο παράδειγμα όπου έχει εφαρμοσθεί one-hot encoding στα δεδομένα, η κλάση Covid-19 στην οποία συμπεριλαμβάνονται οι τομές που εμφανίζουν πνευμονία οφειλόμενη στον ιό SARS-CoV-2 αναπαρίσταται με την τιμή “10”, ενώ η κλάση Normal στην οποία συμπεριλαμβάνονται οι τομές φυσιολογικών αξονικών τομογραφιών αναπαρίσταται με την τιμή “01”.

### 3.2.8 Διαχωρισμός σε σύνολα Εκπαίδευσης/Επικύρωσης/Ελέγχου

Για τις ανάγκες του πειράματος ταξινόμησης τομών αξονικών τομογραφιών με βάση την ύπαρξη πνευμονίας COVID-19, τα δεδομένα διαχωρίζονται σε τρία υποσύνολα δεδομένων:

Στο σύνολο εκπαίδευσης, το οποίο αποτελείται από τα δεδομένα εκείνα τα οποία πρόκειται να χρησιμοποιηθούν για την εκπαίδευση του εκάστοτε μοντέλου βαθιάς μάθησης. Για την κατασκευή του συνόλου εκπαίδευσης αξιοποιείται το σύνολο COVID-CT-MD.

Στο σύνολο επικύρωσης, το οποίο αποτελείται από τα δεδομένα εκείνα τα οποία πρόκειται να χρησιμοποιηθούν για την αξιολόγηση της επίδοσης του εκάστοτε μοντέλου βαθιάς μάθησης έπειτα από κάθε κύκλο εκπαίδευσης ή εποχή, και την προσαρμογή των παραμέτρων του ανάλογα με την επίδοση που επιτυγχάνει. Για την κατασκευή του συνόλου επικύρωσης αξιοποιείται το σύνολο COVID-CT-MD.

Στο σύνολο ελέγχου, το οποίο αποτελείται από άγνωστα δεδομένα για το μοντέλο βαθιάς μάθησης, δηλαδή δεδομένα τα οποία δεν έχουν χρησιμοποιηθεί ούτε στη φάση εκπαίδευσης, αλλά ούτε και στη φάση επικύρωσης του μοντέλου και πρόκειται να χρησιμοποιηθούν για την αξιολόγηση της ικανότητας γενίκευσης του. Για την

κατασκευή του συνόλου ελέγχου αξιοποιείται ένα υποσύνολο 2000 τομών για κάθε κατηγορία από το σύνολο CC-CCII.

Για την κατασκευή των συνόλων εκπαίδευσης και επικύρωσης επιλέγονται από το σύνολο COVID-CT-MD οι τομές του συνόλου έπειτα από την αφαίρεση των τομών στις οποίες δεν διακρίνεται το πνευμονικό παρέγχυμα ή που δεν εμφανίζουν αλλοιώσεις. Δηλαδή επιλέγονται 9223 τομές από την κατηγορία που αφορά τον τομέα που εμφανίζουν COVID-19 πνευμονία, 2652 τομές από την κατηγορία που αφορά τον τομέα που εμφανίζουν πνευμονία άλλης αιτιολογίας και 9881 τομές από την κατηγορία που αφορά τον τομέα που δεν εμφανίζουν αλλοιώσεις.

Παρατηρώντας την κατανομή των τομών στις κλάσεις του συνόλου δεδομένων, μπορεί να διαπιστωθεί πως οι τομές αξονικών τομογραφιών που ανήκουν στην κατηγορία που αφορά τον τομέα που εμφανίζουν πνευμονία άλλης αιτιολογίας είναι πολύ λιγότερες από τις τομές που ανήκουν στις άλλες δύο κατηγορίες. Για το λόγο αυτό, όσον αφορά το πρόβλημα ταξινόμησης των τομών σε τομές που εμφανίζουν COVID-19 πνευμονία και σε τομές που εμφανίζουν πνευμονία άλλης αιτιολογίας, εφαρμόσθηκε υποδειγματοληψία (undersampling) της κατηγορίας που αντιστοιχεί σε τομές ασθενών που εμφανίζουν COVID-19 πνευμονία.

Κάτι άλλο που θα πρέπει να επισημανθεί σε αυτό το σημείο είναι ότι στα πλαίσια του πειράματος πραγματοποιείται διαχωρισμός των δεδομένων σε σύνολα εκπαίδευσης και επικύρωσης με τους ακόλουθους δύο τρόπους:

Σε επίπεδο τομής (slice level), όπου πραγματοποιείται τυχαίος διαχωρισμός των τομών αξονικών τομογραφιών στα σύνολα εκπαίδευσης και επικύρωσης χωρίς κάποιο κριτήριο.

Σε επίπεδο ασθενή (patient level), όπου πραγματοποιείται τυχαίος διαχωρισμός των τομών αξονικών τομογραφιών στα σύνολα εκπαίδευσης και επικύρωσης με βάση τον ασθενή στον οποίο ανήκουν.

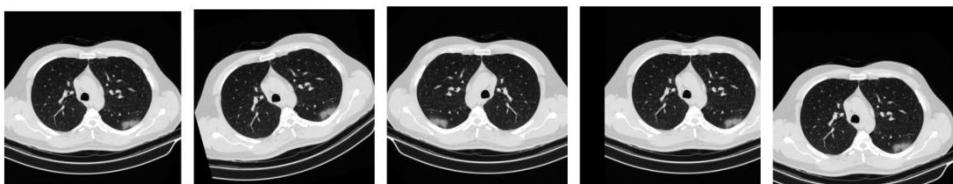
Επομένως, η διεξαγωγή του πειράματος πραγματοποιείται αξιοποιώντας δύο μεθόδους διαχωρισμού των δεδομένων. Αυτή η απόφαση λήφθηκε διότι εάν κάποιος παρατηρήσει το σύνολο δεδομένων COVID-CT-MD, θα διαπιστώσει ότι οι συνεχόμενες τομές μιας αξονικής τομογραφίας ενός ασθενή έχουν ελάχιστες διαφορές μεταξύ τους όσον αφορά το πνευμονικό παρέγχυμα. Κάτι τέτοιο έχει ως αποτέλεσμα εάν ο διαχωρισμός γίνει σε επίπεδο τομής (slice level), συνεχόμενες τομές του ίδιου ασθενή που διαφέρουν ελάχιστα μεταξύ τους να εμφανίζονται τόσο στο σύνολο εκπαίδευσης (training set) όσο και στο σύνολο επικύρωσης (validation set) [50], [51]. Το συγκεκριμένο φαινόμενο ονομάζεται διαρροή δεδομένων (data leakage) και οδηγεί σε ψευδώς υπερβολικά υψηλή επίδοση ενός μοντέλου και κατ' επέκταση μείωση της αξιοπιστίας των αποτελεσμάτων της διαδικασίας εκπαίδευσης [50].

### 3.2.9 Προσαύξηση Δεδομένων Εκπαίδευσης (Data Augmentation)

Ένα από τα σημαντικότερα προβλήματα, αν όχι το σημαντικότερο που παρουσιάζεται κατά την εκπαίδευση ενός μοντέλου μηχανικής μάθησης, αποτελεί το φαινόμενο της υπερπροσαρμογής (overfitting), το οποίο αναλύεται διεξοδικά στο Κεφάλαιο 2. Ένας τρόπος περιορισμού του προαναφερθέντος φαινομένου αποτελεί η εφαρμογή τεχνικών προσαύξησης δεδομένων (data augmentation) στο υποσύνολο εκπαίδευσης του συνόλου δεδομένων.

Στο πείραμα του παρόντος κεφαλαίου, η τεχνική προσαύξησης των δεδομένων πραγματοποιείται “on the fly”, δηλαδή κάθε φορά που πρέπει να τροφοδοτηθεί το εκάστοτε τεχνητό νευρωνικό δίκτυο με δεδομένα, πραγματοποιείται επιτόπου η προσαύξηση των δεδομένων μέσω της εφαρμογής τυχαίων προκαθορισμένων μετασχηματισμών.

Για τις ανάγκες του προβλήματος της ταξινόμησης εικόνων εφαρμόσθηκαν συνδυασμοί περιστροφής εικόνας στο εύρος 0 έως 15 μοίρες, οριζόντιας αντιστροφής εικόνας, οριζόντιας και κατακόρυφης μετακίνησης εικόνας στο εύρος 0 έως 15%. Στην Εικόνα 3.13 παρουσιάζεται η εφαρμογή προσαύξησης δεδομένων μέσω γεωμετρικών μετασχηματισμών σε τομή του συνόλου δεδομένων COVID-CT-MD.



Εικόνα 3.13: Εφαρμογή προσαύξησης δεδομένων σε τομή αξονικής τομογραφίας του συνόλου COVID-CT-MD

## 3.3 Αρχιτεκτονικές Τεχνητών Νευρωνικών Δικτύων

Στη συγκεκριμένη υποενότητα παρουσιάζονται οι αρχιτεκτονικές τεχνητών νευρωνικών δικτύων που επιστρατεύθηκαν ούτως ώστε να εκπαιδευτούν στην ταξινόμηση τομών αξονικών τομογραφιών. Όλα τα μοντέλα που χρησιμοποιήθηκαν αποτελούν προεκπαιδευμένα μοντέλα συνελικτικών νευρωνικών δικτύων, δηλαδή μοντέλα που έχουν εκπαιδευθεί προηγουμένως πάνω στο σύνολο δεδομένων ImageNet.

Οι αρχιτεκτονικές που αξιοποιήθηκαν είναι οι InceptionV3 [52], ResNet50 [53], DenseNet121 [54] και Xception [55]. Επιπλέον, κατά την εκπαίδευση των συγκεκριμένων αρχιτεκτονικών εφαρμόσθηκε η μέθοδος της μεταφοράς μάθησης, η

οποία αναλύεται στο Κεφάλαιο 2. Για την υλοποίηση των συγκεκριμένων αρχιτεκτονικών αξιοποιήθηκε το πακέτο Keras [56] που είναι γραμμένο σε γλώσσα προγραμματισμού Python.

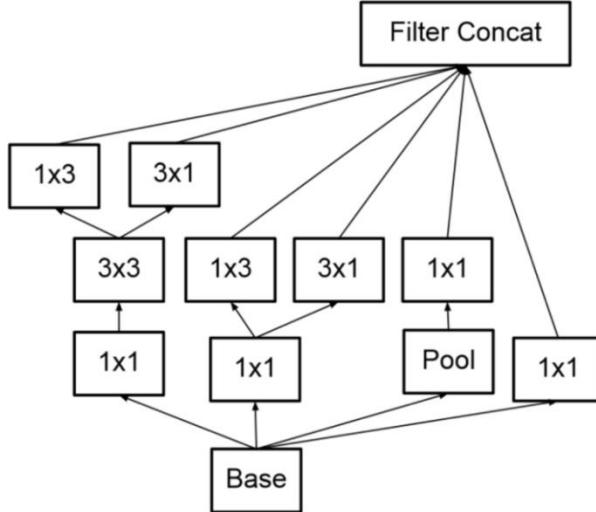
### 3.3.1 Αρχιτεκτονική Inception

Η αρχιτεκτονική Inception αποτελεί μια αρχιτεκτονική συνελικτικού νευρωνικού δικτύου η οποία προτάθηκε από τους Szegedy κ.α. το έτος 2014 [57]. Η συγκεκριμένη αρχιτεκτονική προσφέρει το πλεονέκτημα ότι έχει χαμηλότερο υπολογιστικό κόστος συγκριτικά με άλλες [52]. Η πρώτη έκδοση ονομαζόταν Inception ή GoogleNet [57]. Η έκδοση που αξιοποιείται στη συγκεκριμένη εργασία είναι η έκδοση InceptionV3 [52].

Το βασικό δομικό στοιχείο της InceptionV3 αρχιτεκτονικής είναι το inception module [52], [57]. Ένα inception module αποτελείται από παράλληλα επίπεδα συνέλιξης (convolutional layers) με πυρήνες (kernels) διαφορετικών διαστάσεων, τα οποία συλλαμβάνουν χαρακτηριστικά των δεδομένων σε διαφορετικές κλίμακες, και από παράλληλα επίπεδα συγκέντρωσης (pooling layers) [52]. Σε ένα inception module η πράξη της συνέλιξης είναι υπολογιστικά ακριβή. Για να αντιμετωπισθεί αυτό, τοποθετούνται επίπεδα συνέλιξης με μέγεθος πυρήνα  $1 \times 1$  πριν από κάθε επίπεδο συνέλιξης, μειώνοντας τις διαστάσεις του δικτύου και αυξάνοντας την ταχύτητα των υπολογισμών [52], [57]. Το δίκτυο εκτός από inception modules περιλαμβάνει και ταξινομητές που ονομάζονται auxiliary classifiers και οι οποίοι τοποθετούνται σε ενδιάμεσα επίπεδα του δικτύου [52]. Οι auxiliary classifiers βελτιώνουν την ικανότητα σύγκλισης των μοντέλων αντιμετωπίζοντας το πρόβλημα της εξαφανιζόμενης κλίσης (vanishing gradient) [52]. Η δομή ενός inception module του μοντέλου InceptionV3 φαίνεται στην Εικόνα 3.14.

Όπως έχει αναφερθεί στο Κεφάλαιο 2, κατά τη διαδικασία εκτέλεσης του αλγορίθμου οπισθοδιάδοσης (backpropagation algorithm) ενός τεχνητού νευρωνικού δικτύου, πραγματοποιείται ενημέρωση των βαρών των επιπέδων του δικτύου ανάλογα με την τιμή της κλίσης (gradient), έπειτα από κάθε κύκλο εκπαίδευσης του μοντέλου [58]. Ωστόσο, ανάλογα με τη συνάρτηση ενεργοποίησης και την αρχιτεκτονική νευρωνικού δικτύου που χρησιμοποιείται, η τιμή της κλίσης γίνεται υπερβολικά μικρή κατά τη διαδικασία οπισθοδιάδοσης [58]. Αυτό οδηγεί σε αργή ενημέρωση των βαρών των αρχικών επιπέδων του μοντέλου, με αποτέλεσμα η διαδικασία της εκπαίδευσης να είναι αργή και σε ορισμένες περιπτώσεις να σταματάει εντελώς [58]. Αυτό αποτελεί το πρόβλημα της εξαφανιζόμενης κλίσης, εξαιτίας του οποίου, το μοντέλο αδυνατεί να ανιχνεύσει τις σχέσεις και τα μοτίβα που εμφανίζονται μεταξύ των δεδομένων που λαμβάνει ως είσοδο και άρα παρουσιάζει χαμηλή επίδοση. Το πρόβλημα της εξαφανιζόμενης κλίσης περιορίζει τον αριθμό των

επιπέδων που μπορεί να έχει ένα δίκτυο, και εμποδίζει τη δημιουργία πολύ βαθύτερων αρχιτεκτονικών [58].

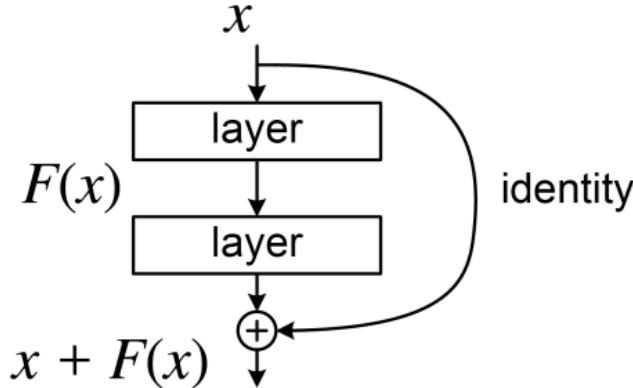


Εικόνα 3.14: Ένα inception module της αρχιτεκτονικής InceptionV3 [52]

### 3.3.2 Αρχιτεκτονική ResNet

Η αρχιτεκτονική ResNet αποτελεί μια αρχιτεκτονική συνελικτικού νευρωνικού δικτύου που προτάθηκε από τους He κ.α. το έτος 2015 [53]. Στόχος της ερευνητικής ομάδας ήταν να προτείνει ένα μοντέλο βαθιάς μάθησης που να αντιμετωπίζει το πρόβλημα της εξαφανιζόμενης κλίσης, το οποίο εμφανίζεται σε νευρωνικά δίκτυα όπως τα μοντέλα της οικογένειας VGG [59] που αποτελούνται από μεγάλο αριθμό επιπέδων [53].

Η ερευνητική ομάδα που υλοποίησε το ResNet αντιμετώπισε το πρόβλημα της εξαφανιζόμενης κλίσης εισάγοντας συνδέσεις στο δίκτυο, οι οποίες ονομάζονται residual συνδέσεις [53]. Στην ουσία μια residual σύνδεση προσπερνά ένα ή περισσότερα επίπεδα και αθροίζει την είσοδο ενός επιπέδου με την έξοδο ενός μεταγενέστερου επιπέδου [53]. Με αυτό τον τρόπο δίνεται η δυνατότητα στις κλίσεις να ρέουν διαμέσου των επιπέδων του δικτύου χωρίς να παρουσιάζεται το φαινόμενο της εξαφανιζόμενης κλίσης [53]. Το πλεονέκτημα που προσφέρουν οι residual συνδέσεις είναι ότι επιτρέπουν την προσθήκη πολλών επιπέδων σε ένα δίκτυο. Οι residual συνδέσεις ενσωματώνονται μέσα σε μπλοκ τα οποία ονομάζονται residual μπλοκ [53]. Η δομή ενός residual μπλοκ φαίνεται στην Εικόνα 3.15.

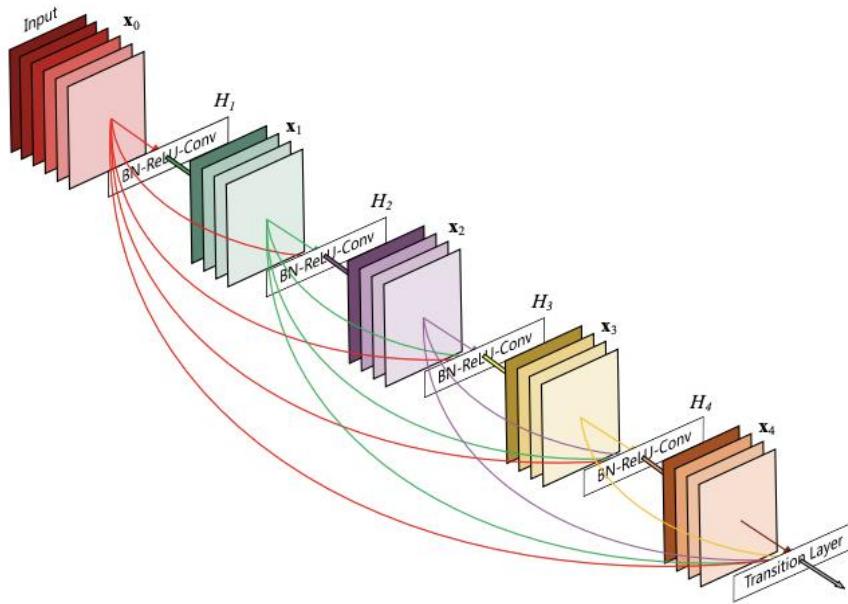


Εικόνα 3.15: Ένα Residual μπλοκ [53]

Από τις διάφορες παραλλαγές της αρχιτεκτονικής ResNet, επιλέχθηκε να χρησιμοποιηθεί η αρχιτεκτονική ResNet50, η οποία αποτελείται από 50 επίπεδα [53].

### 3.3.3 Αρχιτεκτονική DenseNet

Η αρχιτεκτονική DenseNet αποτελεί μια αρχιτεκτονική συνελικτικού νευρωνικού δικτύου που προτάθηκε από τους Huang κ.α. το έτος 2017 και η οποία αντιμετωπίζει αποτελεσματικά το πρόβλημα της εξαφανιζόμενης κλίσης (vanishing gradient problem) [54]. Η αρχιτεκτονική των μοντέλων της οικογένειας DenseNet ακολουθεί ένα απλό μοτίβο συνδεσμότητας [54]. Πιο συγκεκριμένα, τα επίπεδα του δικτύου συνδέονται απευθείας μεταξύ τους πράγμα που διευκολύνει τη ροή πληροφοριών μεταξύ των επιπέδων του δικτύου [54]. Επίσης, πραγματοποιείται επαναχρησιμοποίηση των χαρακτηριστικών, αφού κάθε επίπεδο λαμβάνει επιπλέον εισόδους από όλα τα προηγούμενα επίπεδα του δικτύου και στέλνει τις εξόδους του σε όλα τα επόμενα επίπεδα [54]. Κάτι τέτοιο αυξάνει την ποικιλομορφία των εισόδων και βελτιώνει την αποδοτικότητα του μοντέλου [54]. Τα επίπεδα της αρχιτεκτονικής DenseNet οργανώνονται σε μπλοκ τα οποία ονομάζονται dense μπλοκ [54]. Ανάμεσα στα dense μπλοκ τοποθετούνται μπλοκ που ονομάζονται transition και τα οποία πραγματοποιούν μείωση των διαστάσεων των δεδομένων [54]. Στην Εικόνα 3.16 παρουσιάζεται η δομή ενός dense μπλοκ 5 επιπέδων στο οποίο πραγματοποιείται επαναχρησιμοποίηση χαρακτηριστικών.



**Εικόνα 3.16:** Ένα dense μπλοκ 5 επιπέδων στο οποίο πραγματοποιείται επαναχρησιμοποίηση χαρακτηριστικών [54]

Σε αντίθεση με την αρχιτεκτονικές της οικογένειας ResNet όπου ο συνδυασμός των χαρακτηριστικών πραγματοποιείται με άθροιση (summation) πριν τροφοδοτηθούν στο επόμενο επίπεδο, στις αρχιτεκτονικές της οικογένειας DenseNet πραγματοποιείται συνδυασμός μέσω συνένωσης (concatenation) τους [54].

Πλεονεκτήματα της οικογένειας αρχιτεκτονικών DenseNet είναι ότι αποφεύγεται η εκμάθηση περιττών χαρακτηριστικών από το μοντέλο και βελτιώνεται η ροή πληροφοριών και κλίσεων σε όλο το δίκτυο, η οποία διευκολύνει τη διαδικασία της εκπαίδευσης [54].

Όσον αφορά την αρχιτεκτονική του μοντέλου DenseNet121 που αξιοποιείται στη παρούσα εργασία, αυτή αποτελείται από 121 επίπεδα [54].

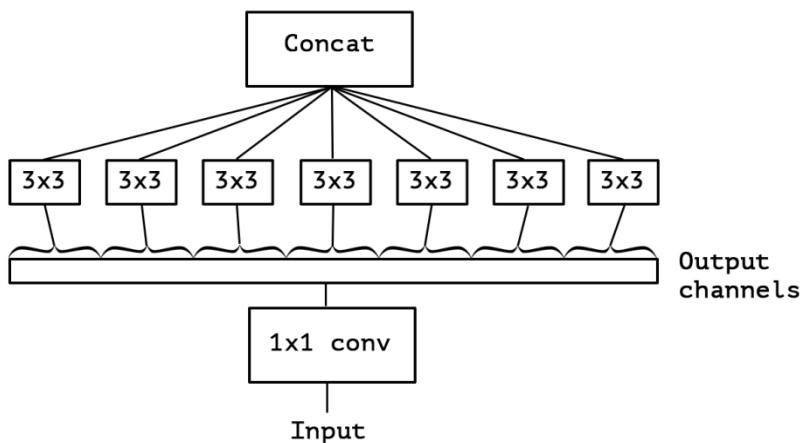
### 3.3.4 Αρχιτεκτονική Xception

Η αρχιτεκτονική Xception αποτελεί μια αρχιτεκτονική συνελικτικού νευρωνικού δικτύου η οποία προτάθηκε από τον Chollet το έτος 2017 και η οποία είναι εμπνευσμένη από τα inception modules της αρχιτεκτονικής Inception [55]. Η καινοτομία που εισάγει το μοντέλο Xception είναι η χρήση διαχωρίσιμων σε βάθος συνελίξεων (depthwise separable convolutions) [55]. Στην ουσία το μοντέλο Xception αντικαθιστά τα inception modules της αρχιτεκτονικής Inception με modules

τα οποία αποτελούνται από διαχωρίσιμες σε βάθος συνέλιξεις και τα οποία ονομάζονται extreme inception modules [55].

Ένα επίπεδο σε βάθος διαχωρίσιμης συνέλιξης (depthwise separable convolution layer) εφαρμόζει δύο τεχνικές. Μία συνέλιξη σε βάθος (depthwise convolution) και μία σημειακή συνέλιξη (pointwise convolution) [55]. Η πρώτη πραγματοποιεί χωρική συνέλιξη ανεξάρτητα σε κάθε κανάλι της εισόδου [55]. Η δεύτερη πραγματοποιεί εφαρμογή μίας συνέλιξης διαστάσεων 1x1 στο αποτέλεσμα της εξόδου της συνέλιξης σε βάθος και στην ουσία συνδυάζει την πληροφορία των διαφόρων καναλιών και παράγει τα χαρακτηριστικά εξόδου [55].

Όσον αφορά την αρχιτεκτονική του μοντέλου Xception, αυτή περιλαμβάνει 36 επίπεδα συνέλιξης [55]. Εκτός από το πρώτο και το τελευταίο, τα υπόλοιπα επίπεδα συνέλιξης είναι οργανωμένα σε 14 modules τα οποία περιλαμβάνουν residual συνδέσεις [55]. Με λίγα λόγια, το μοντέλο Xception αποτελείται από ένα σύνολο από επίπεδα διαχωρίσιμων σε βάθος συνδέσεων με residual συνδέσεις [55]. Η δομή ενός extreme inception module του μοντέλου InceptionV3 φαίνεται στην Εικόνα 3.17.



Εικόνα 3.17: Η δομή ενός extreme inception module [55]

### 3.4 Αξιολόγηση Μοντέλων

Στη συγκεκριμένη υποενότητα παρουσιάζονται οι μέθοδοι που χρησιμοποιούνται για την αξιολόγηση του πειράματος που περιγράφεται στο παρόν κεφάλαιο. Οι μέθοδοι αξιολόγησης που αξιοποιούνται είναι η μέθοδος της διασταυρούμενης επικύρωσης 5-πτυχών (5-fold cross validation), οι καμπύλες εκπαίδευσης και επικύρωσης, ο πίνακας σύγχυσης (confusion matrix), καθώς και οι μετρικές Ακρίβεια (Accuracy), Απώλεια (Loss), Αξιοπιστία (Precision), Ανάκληση (Recall) και ο Συνδυασμός Ακρίβειας-Ανάκλησης (F1 Score).

### **3.4.1 Εφαρμογή Διασταυρούμενης Επικύρωσης 5-Πτυχών**

Για την αξιολόγηση των μοντέλων ως προς την ικανότητα τους να ταξινομούν τομές αξονικών τομογραφιών, καθώς και για την ανίχνευση του φαινομένου της υπερπροσαρμογής, εφαρμόζεται η τεχνική διασταυρούμενης επικύρωσης k-πτυχών (k-fold cross validation). Για την τεχνική αυτή, που αναλύεται διεξοδικά στο Κεφάλαιο 2, επιλέχθηκε η τιμή  $k=5$ , δηλαδή το αρχικό σύνολο να διαχωριστεί σε 5 υποσύνολα.

### **3.4.2 Καμπύλες Εκπαίδευσης και Επικύρωσης**

Για την αξιολόγηση της επίδοσης των μοντέλων κατά την εκπαιδευτική διαδικασία, τον προσδιορισμό ύπαρξης ή μη υποπροσαρμογής και υπερπροσαρμογής τους και κατ' επέκταση για τη διερεύνηση της ικανότητας γενίκευσης τους, χρησιμοποιούνται γραφικές παραστάσεις που παρουσιάζουν τις καμπύλες της διαδικασίας εκπαίδευσης και επικύρωσης.

### **3.4.3 Πίνακας Σύγχυσης (Confusion Matrix)**

Ο πίνακας σύγχυσης (confusion matrix) αποτελεί μια πολύ δημοφιλή μέθοδο αξιολόγησης της επίδοσης ενός αλγορίθμου ταξινόμησης. Μπορεί να κατασκευαστεί τόσο σε περιπτώσεις όπου το πρόβλημα ταξινόμησης είναι δυαδικό, όσο και σε περιπτώσεις όπου έχουμε να κάνουμε με πολυταξική ταξινόμηση [60]. Αποτελεί έναν δυσδιάστατο πίνακα ο οποίος παρέχει μια αναπαράσταση της επίδοσης του εκάστοτε αλγορίθμου ταξινόμησης. Ο συγκεκριμένος πίνακας στην ουσία δείχνει πόσα δείγματα του συνόλου δεδομένων ταξινομήθηκαν σωστά, και πόσα δείγματα ταξινομήθηκαν λάθος από τον αλγόριθμο ταξινόμησης. Στην Εικόνα 3.18 παρουσιάζεται η δομή ενός πίνακα σύγχυσης για δυαδικό πρόβλημα ταξινόμησης. Κάθε γραμμή του πίνακα αναπαριστά το σύνολο των δειγμάτων του συνόλου δεδομένων τα οποία αντιστοιχούν στις πραγματικές κατηγορίες, ενώ κάθε στήλη του πίνακα αναπαριστά το σύνολο των δειγμάτων του συνόλου δεδομένων τα οποία αντιστοιχούν στις προβλεπόμενες κατηγορίες.

		Predicted Class	
		1	0
True Class	1	True Positives	False Negatives
	0	False Positives	True Negatives

Εικόνα 3.18: Πίνακας σύγχυσης για ένα πρόβλημα δυαδικής ταξινόμησης

Για την καλύτερη κατανόηση του τι αναπαριστά ένας πίνακας σύγχυσης είναι αναγκαίο να δοθούν κάποιοι ορισμοί:

Τα Αληθώς Θετικά (True Positives) αποτελούν τον αριθμό των δειγμάτων που ανήκουν στη θετική κατηγορία και ταξινομούνται ορθά σε αυτή από τον αλγόριθμο ταξινόμησης [60], [61].

Τα Αληθώς Αρνητικά (True Negatives) αποτελούν τον αριθμό των δειγμάτων που ανήκουν στην αρνητική κατηγορία και ταξινομούνται ορθά σε αυτή [60], [61].

Τα Ψευδώς Θετικά (False Positives) αποτελούν τον αριθμό των δειγμάτων που ανήκουν στην αρνητική κατηγορία και ταξινομούνται λανθασμένα στη θετική κατηγορία από τον αλγόριθμο ταξινόμησης [60], [61].

Τα Ψευδώς Αρνητικά (False Negatives) αποτελούν τον αριθμό των δειγμάτων που ανήκουν στην θετική κατηγορία και ταξινομούνται λανθασμένα στην αρνητική κατηγορία [60], [61].

Στο πείραμα του παρόντος κεφαλαίου οι πίνακες σύγχυσης δημιουργούνται από τη συνένωση (concatenation) των πινάκων που παράγονται κατά τη διάρκεια εκτέλεσης της μεθόδου διασταυρούμενης επικύρωσης 5-πτυχών.

### 3.4.4 Ακρίβεια (Accuracy)

Η Ακρίβεια (Accuracy) αποτελεί μία από τις συνηθέστερες μετρικές για την αξιολόγηση ενός αλγορίθμου ταξινόμησης [60]. Στόχος ενός αλγορίθμου ταξινόμησης είναι η βελτιστοποίηση της συγκεκριμένης μετρικής. Η μετρική Ακρίβεια (Accuracy) ορίζεται ως ο λόγος μεταξύ των ορθά ταξινομημένων δειγμάτων προς το συνολικό αριθμό δειγμάτων και δίνεται στην Εξίσωση (3.1):

$$\text{Ακρίβεια (Accuracy)} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (3.1)$$

### 3.4.5 Απώλεια (Loss)

Η Απώλεια (Loss) αποτελεί μετρική αξιολόγησης ενός αλγορίθμου ταξινόμησης. Στόχος ενός αλγορίθμου ταξινόμησης είναι η όσο το δυνατόν μεγαλύτερη μείωση της συγκεκριμένης μετρικής [23]. Η μετρική Απώλεια για το πρόβλημα δυαδικής ταξινόμησης, που εξετάζεται στο παρόν κεφάλαιο, δίνεται από τη συνάρτηση binary cross entropy. Ο τύπος της συνάρτησης binary cross entropy δίνεται στην Εξίσωση (3.2):

$$L = -\frac{1}{N} \sum_{i=1}^N y_i * \log(p_i) + (1 - y_i) * \log(1 - p_i) \quad (3.2)$$

### 3.4.6 Αξιοπιστία (Precision)

Η μετρική Αξιοπιστία (Precision) ορίζεται ως ο λόγος των δειγμάτων που ορθά έχουν ταξινομηθεί ως θετικά, προς το σύνολο των δειγμάτων που έχουν ταξινομηθεί ως θετικά, από τον αλγόριθμο [60]. Η μετρική Αξιοπιστία δίνεται στην Εξίσωση (3.3):

$$\text{Αξιοπιστία (Precision)} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (3.3)$$

### 3.4.7 Ανάκληση (Recall)

Η μετρική Ανάκληση (Recall) ή αλλιώς Ευαισθησία (Sensitivity) ορίζεται ως ο λόγος των δειγμάτων που ορθά έχουν ταξινομηθεί ως θετικά από τον αλγόριθμο, προς το συνολικό αριθμό των θετικών δειγμάτων του συνόλου [60]. Η μετρική Ανάκληση δίνεται στην Εξίσωση (3.4):

$$\text{Ανάκληση (Recall)} = \frac{TP}{TP + FN} \quad (3.4)$$

### 3.4.8 Συνδυασμός Ακρίβειας-Ανάκλησης (F1 Score)

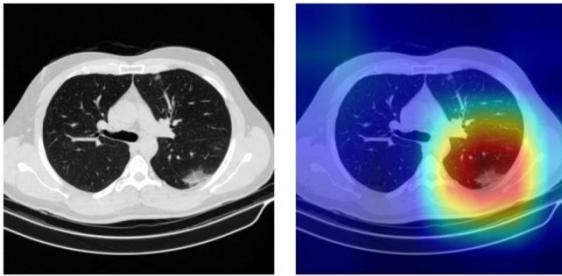
Ο Συνδυασμός Ακρίβειας-Ανάκλησης ή F1 Score αποτελεί μια μετρική αξιολόγησης που αναπαριστά τον αρμονικό μέσο μεταξύ των μετρικών αξιολόγησης Αξιοπιστία (Precision) και Ανάκληση (Recall) [60]. Η μετρική F1 Score δίνεται στην Εξίσωση (3.5):

$$F1\ Score = 2 * \frac{Precision * Recall}{Precision + Recall} = \frac{2 * TP}{2 * TP + FP + FN} \quad (3.5)$$

### 3.4.9 Gradient-weighted Class Activation Mapping (Grad-CAM)

Η έξοδος ενός τεχνητού νευρωνικού δικτύου είναι ακατόρθωτο να προσδιοριστεί ακόμη και από τον ίδιο το δημιουργό του [62]. Για το λόγο αυτό, τα τεχνητά νευρωνικά δίκτυα συχνά αποκαλούνται «μαύρα κουτιά», διότι είναι αδύνατο να γνωρίζει κάποιος τις διεργασίες που πραγματοποιούνται εντός τους [62].

Μία λύση στο συγκεκριμένο ζήτημα αποτελεί η τεχνική Grad-CAM που χρησιμοποιείται στο πεδίο της βαθιάς μάθησης ούτως ώστε να εξηγήσει το λόγο που ένα τεχνητό νευρωνικό δίκτυο έκανε μια συγκεκριμένη πρόβλεψη για μια εικόνα [63]. Για να επιτευχθεί αυτό, η τεχνική Grad-CAM υλοποιεί έναν θερμικό χάρτη (heatmap) στον οποίο εμφανίζονται οι σημαντικές περιοχές της εικόνας [63]. Ο θερμικός χάρτης τοποθετείται πάνω από την αρχική εικόνα, ούτως ώστε να οπτικοποιήσει τις περιοχές στις οποίες επικεντρώνεται το εκάστοτε μοντέλο όταν πραγματοποιεί μια πρόβλεψη [63]. Ένα παράδειγμα εφαρμογής της τεχνικής Grad-CAM φαίνεται στην Εικόνα 3.19.



**Εικόνα 3.19: Παράδειγμα εφαρμογής της μεθόδου Grad-CAM σε τομή που εμφανίζει πνευμονία COVID-19**

### 3.5 Σχετικές Εργασίες

Οι ερευνητικές ομάδες που έχουν ασχοληθεί με την ταξινόμηση τομών αξονικών τομογραφιών χρησιμοποιούν προεκπαιδευμένα μοντέλα βαθιάς μάθησης, τα οποία εκπαιδεύουν εκ νέου σε δυαδικά ή και πολυταξικά σύνολα δεδομένων. Για ευκολία, θα αναφερόμαστε στην κατηγορία που αφορά τομές που εμφανίζουν COVID-19 πνευμονία με την ονομασία “Covid-19”, πνευμονία άλλης αιτιολογίας με την ονομασία “Other Pneumonia” και στις τομές χωρίς αλλοιώσεις με την ονομασία “Normal”.

Η ομάδα των Alshazly κ.α. [64] εκπαίδευσε και αξιολόγησε προεκπαιδευμένες παραλλαγές των μοντέλων ResNet και DenseNet καθώς και τα προεκπαιδευμένα μοντέλα InceptionV3, Xception, SqueezeNet [65] και ShuffleNet [66] ως προς την ικανότητά τους να πραγματοποιούν ταξινόμηση των τομών των συνόλων SARS-CoV-2 CT-scan [67] και COVID19-CT [68] στις κατηγορίες Covid-19 και Non Covid19. Η τελευταία κατηγορία περιέχει τόσο φυσιολογικές όσο και τομές που εμφανίζουν πνευμονία άλλης αιτιολογίας.

Οι Ulutas κ.α. [69] υλοποίησαν ένα μοντέλο το οποίο ονόμασαν COVIDxNet-CT και εκπαίδευσαν αυτό καθώς και προεκπαιδευμένα μοντέλα MobileNetv2 [70] και ResNet50 ούτως ώστε να αξιολογήσουν την ικανότητα τους να ταξινομούν τα δεδομένα του συνόλου COVID-CT-MD σε τρεις κατηγορίες (Covid-19, Other Pneumonia, Normal).

Οι ερευνητές Gunraj κ.α. [71], [72] κατασκεύασαν συνελικτικά νευρωνικά δίκτυα τα οποία δοκίμασαν στο σύνολο COVIDx CT. Πιο συγκεκριμένα, στην εργασία [71] εκπαίδευσαν ένα συνελικτικό νευρωνικό δίκτυο το οποίο ονόμασαν COVID-Net CT, αρχικά στο σύνολο ImageNet και στη συνέχεια στο σύνολο COVIDx CT, ούτως ώστε να αξιολογήσουν την ικανότητά του να πραγματοποιεί ταξινόμηση σε τρεις κατηγορίες (Covid-19, Other Pneumonia, Normal). Στην εργασία [72] επέκτειναν το συγκεκριμένο δίκτυο, το οποίο ονόμασαν COVID-Net CT-2 και το εκπαίδευσαν σε δύο νέα σύνολα δεδομένων τα COVIDx CT-2A και COVIDx CT-2B.

Οι Chaudhary κ.α. [73] στα πλαίσια του διαγωνισμού IEEE ICASSP 2021 Signal Processing Grand Challenge (SPGC) πρότειναν μία μεθοδολογία δύο σταδίων για την ταξινόμηση τομών αξονικών τομογραφιών του συνόλου δεδομένων του διαγωνισμού σε τρεις κατηγορίες (Covid-19, Other Pneumonia, Normal). Οι ερευνητές χρησιμοποίησαν μόνο τις ενδιάμεσες τομές κάθε ασθενή και στο πρώτο στάδιο χρησιμοποίησαν δύο προεκπαιδευμένα DenseNet121 δίκτυα ώστε να ταξινομήσουν τις τομές των αξονικών που ανήκουν στις κλάσεις COVID-19 και Other Pneumonia σε τομές που εμφανίζουν αλλοιώσεις και σε τομές που δεν εμφανίζουν αλλοιώσεις. Στο δεύτερο στάδιο εκπαίδευσαν ένα προεκπαιδευμένο EfficientNetB6 [74] δίκτυο ούτως ώστε να πραγματοποιεί ταξινόμηση σε τρεις κατηγορίες (Normal, Covid-19, Other Pneumonia). Στις δύο τελευταίες κατηγορίες, χρησιμοποιήθηκαν μόνο οι τομές που εμφάνιζαν αλλοιώσεις. Η εκπαίδευση πραγματοποιήθηκε τόσο με διαχωρισμό δεδομένων ανά τομή όσο και με διαχωρισμό δεδομένων ανά ασθενή.

Στην εργασία τους οι Rahimzadeh [75] κ.α. πρότειναν ένα μοντέλο που συνδυάζει το μοντέλο ResNet50V2 με ένα Feature Pyramid Network (FPN) [76], και εκπαίδευσαν το συγκεκριμένο δίκτυο καθώς και ένα μοντέλο Xception και ένα μοντέλο ResNet50V2, σε ένα δικό τους σύνολο δεδομένων που ονομάζεται COVID-CTset [75], ούτως ώστε τα δίκτυα να πραγματοποιούν ταξινόμηση τομών σε δύο κατηγορίες (Covid-19, Normal).

Η ομάδα των Attallah κ.α. [77] αξιοποίησε προεκπαιδευμένες παραλλαγές του μοντέλου ResNet για τη ταξινόμηση τομών σε δύο κατηγορίες (Covid-19, Normal) χρησιμοποιώντας ένα ιδιωτικό σύνολο που ονομάζεται OMNIAHCOV και ένα δημόσιο σύνολο που ονομάζεται SARS-CoV-2 CT-scan. Ενδιαφέρον παρουσιάζει το γεγονός ότι η ομάδα χρησιμοποίησε διακριτό μετασχηματισμό κυματιδίων (DWT) για την προεπεξεργασία των τομών.

Οι Vaidyanathan κ.α. [78] διερεύνησαν την ικανότητα ενός προεκπαιδευμένου InceptionV3 μοντέλου να ταξινομεί δεδομένα τριών ιδιωτικών συνόλων σε τρεις κατηγορίες (Covid-19, Other Pneumonia, Normal) και πραγματοποίησαν εξωτερική αξιολόγηση του μοντέλου στο σύνολο COVID-CT-MD.

Η ομάδα των Ko κ.α. [79] επέλεξε να αξιολογήσει προεκπαιδευμένα μοντέλα VGG16, ResNet50, InceptionV3 και Xception ως προς την ικανότητα ταξινόμησης των δεδομένων δύο ιδιωτικών και ενός δημοσίου συνόλου σε τρεις κατηγορίες (COVID-19, Other Pneumonia, Normal) και επιπλέον πραγματοποίησε εξωτερική αξιολόγηση του μοντέλου σε τομές αξονικών τομογραφιών χαμηλής ποιότητας προερχόμενες από άλλες δημοσιευμένες εργασίες.

Οι ερευνητές Nguyen κ.α. [80] εκπαίδευσαν και αξιολόγησαν ένα συνελικτικό νευρωνικό δίκτυο, χρησιμοποιώντας ένα ιδιωτικό και τρία δημόσια σύνολα δεδομένων, ως προς την ικανότητα ταξινόμησης των δεδομένων σε δύο κατηγορίες (Covid-19, Normal). Το ένα εκ των δημοσίων συνόλων ήταν το COVID-CT-MD.

Η ομάδα των Fallahpoor κ.α. [81] αξιολόγησε την ικανότητα προεκπαιδευμένων παραλλαγών των μοντέλων ResNet και DenseNet να ταξινομούν τομές στις οποίες έχει εφαρμοσθεί κατάτμηση πνευμονικού παρεγχύματος. Οι τομές προήλθαν από δύο δημόσια και δύο ιδιωτικά σύνολα δεδομένων και ανήκουν σε δύο κατηγορίες (Covid-19, Normal). Το ένα από τα δύο δημόσια σύνολα ήταν το COVID-CT-MD.

Τα μοντέλα των τεσσάρων τελευταίων ομάδων που αναφέρθηκαν επιτυγχάνουν υψηλή επίδοση τόσο στο εσωτερικό όσο και στο εξωτερικό σύνολο αξιολόγησης.

### 3.6 Πειραματική Διαδικασία

Στη συγκεκριμένη υποενότητα, πραγματοποιείται αξιολόγηση της επίδοσης των μοντέλων που εκπαιδεύθηκαν ως προς την ικανότητα τους να ταξινομούν τομές αξονικών τομογραφιών, με στόχο τη διάκριση των τομών που εμφανίζουν αλλοιώσεις που προέρχονται από τον ιό SARS-CoV-2. Πιο συγκεκριμένα, αρχικά γίνεται μια αναφορά στα σύνολα δεδομένων και στις αρχιτεκτονικές νευρωνικών δικτύων που επιστρατεύονται. Στη συνέχεια, αναφέρονται οι υπερπαράμετροι που επιλέχθηκαν για τη διαδικασία εκπαίδευσης των μοντέλων. Έπειτα, παρουσιάζονται και σχολιάζονται τα αποτελέσματα των διαδικασιών εκπαίδευσης και αξιολόγησης των μοντέλων.

Για την εκπαίδευση των μοντέλων του πειράματος και την αξιολόγηση της ικανότητας τους να ταξινομούν ιατρικές εικόνες ούτως ώστε να διακρίνουν τις αλλοιώσεις που οφείλονται σε πνευμονία COVID-19, χρησιμοποιήθηκε το σύνολο COVID-CT-MD. Το σύνολο CC-CCII αξιοποιήθηκε ως ένα εξωτερικό σύνολο αξιολόγησης των μοντέλων.

Στα πλαίσια του παρόντος πειράματος, για την κατηγορία που αφορά τομές που εμφανίζουν COVID-19 πνευμονία θα χρησιμοποιούμε την ονομασία “Covid-19”, για την κατηγορία που αφορά τομές που εμφανίζουν πνευμονία άλλης αιτιολογίας την ονομασία “Other Pneumonia” και για τις τομές που δεν εμφανίζουν αλλοιώσεις την ονομασία “Normal”.

Όσον αφορά το πειραματικό μέρος, για τη διεξαγωγή των αξιοποιήθηκαν τα μοντέλα InceptionV3, ResNet50, Xception και DenseNet121, τα οποία επιστρατεύουν την τεχνική της μεταφοράς μάθησης για την εκπαίδευσή τους. Οι παράμετροι των αρχιτεκτονικών παρουσιάζονται στον Πίνακα 3.1.

**Πίνακας 3.1: Αρχιτεκτονικές που χρησιμοποιήθηκαν για το πρόβλημα ταξινόμησης και οι παράμετροί τους**

Αρχιτεκτονική	Συνολικές Παράμετροι	Εκπαιδεύσιμες Παράμετροι	Μη Εκπαιδεύσιμες Παράμετροι
InceptionV3	21.806.882	21.772.450	34.432
ResNet50	23.591.810	23.538.690	53.120
Xception	20.865.578	20.811.050	54.528
DenseNet121	7.039.554	6.955.906	83.648

Η εκπαίδευση και επικύρωση των αρχιτεκτονικών που υλοποιήθηκαν πραγματοποιήθηκε μέσω της πλατφόρμας Kaggle [82] αξιοποιώντας μια Nvidia Tesla P100 GPU χωρητικότητας 16 GB.

Τα μοντέλα που κατασκευάσθηκαν έλαβαν ως είσοδο εικόνες διαστάσεων 256x256 και εκπαιδεύθηκαν για 20 εποχές (epochs) με ρυθμό μάθησης 0,000001 όσον αφορά το πρόβλημα δυαδικής ταξινόμησης στις κατηγορίες Covid-19 και Normal και με ρυθμό μάθησης 0,000001 όσον αφορά το πρόβλημα δυαδικής ταξινόμησης στις κατηγορίες Covid-19 και Other Pneumonia. Ακόμη, κατά τη διάρκεια της εκπαιδευτικής διαδικασίας, κάθε φορά που παρουσιάζονταν βελτίωση των τιμών των μετρικών επικύρωσης, πραγματοποιούνταν αποθήκευση των βαρών του εκπαιδευόμενου μοντέλου.

Επειδή η μνήμη δεν ήταν αρκετή ούτως ώστε να δεχτεί όλες τις εικόνες του συνόλου δεδομένων σε κάθε εποχή, τα δεδομένα διαχωρίζονταν σε παρτίδες (batches) και τροφοδοτούνταν σταδιακά στη μνήμη. Το μέγεθος παρτίδας (batch size) που επιλέχθηκε ήταν 32. Ως βελτιστοποιητής (optimizer) επιλέχθηκε ο Adam. Οι υπερπαράμετροι που επιλέχθηκαν συνοψίζονται στους Πίνακες 3.2 και 3.3.

**Πίνακας 3.2: Υπερπαράμετροι διαδικασίας εκπαίδευσης για την ταξινόμηση των τομών στις κατηγορίες Covid-19 και Normal**

Υπερπαράμετρος	Τιμή
Εποχές (Epochs)	20
Διαστάσεις Εικόνων	256x256
Ρυθμός Μάθησης (Learning Rate)	0,000001
Μέγεθος Παρτίδας (Batch Size)	32
Βελτιστοποιητής (Optimizer)	Adam

**Πίνακας 3.3: Υπερπαράμετροι διαδικασίας εκπαίδευσης για την ταξινόμηση των τομών στις κατηγορίες Covid-19 και Other Pneumonia**

Υπερπαράμετρος	Τιμή
Εποχές (Epochs)	20
Διαστάσεις Εικόνων	256x256
Ρυθμός Μάθησης (Learning Rate)	0,00001
Μέγεθος Παρτίδας (Batch Size)	32
Βελτιστοποιητής (Optimizer)	Adam

Η εκπαίδευση των μοντέλων πραγματοποιήθηκε σε τέσσερεις παραλλαγές του συνόλου δεδομένων COVID-CT-MD. Στην Παραλλαγή 1 δεν πραγματοποιείται κατάτμηση του πνευμονικού παρεγχύματος και τα δεδομένα διαχωρίζονται ανά τομή (slice), το ίδιο και στην Παραλλαγή 2, ωστόσο τα δεδομένα διαχωρίζονται ανά ασθενή (patient). Στην Παραλλαγή 3 πραγματοποιείται κατάτμηση του πνευμονικού παρεγχύματος και τα δεδομένα διαχωρίζονται ανά τομή (slice), το ίδιο και στην Παραλλαγή 4, ωστόσο εδώ τα δεδομένα διαχωρίζονται ανά ασθενή (patient). Οι διάφορες παραλλαγές συνοψίζονται στον Πίνακα 3.4.

**Πίνακας 3.4: Παραλλαγές του συνόλου COVID-CT-MD**

Παραλλαγή	Κατάτμηση Πνευμονικού Παρεγχύματος	Διαχωρισμός ανά Ασθενή
1	OXI	OXI
2	OXI	NAI
3	NAI	OXI
4	NAI	NAI

### 3.6.1 Δυαδική Ταξινόμηση των Τομών στις Κατηγορίες Covid-19 και Normal

#### 3.6.1.1 Παραλλαγή 1: Μη Κατάτμηση Πνευμονικού Παρεγχύματος – Διαχωρισμός Δεδομένων ανά Τομή

Στον Πίνακα 3.5 παρουσιάζονται οι μετρικές αξιολόγησης της ικανότητας των μοντέλων να πραγματοποιούν, έπειτα από εκπαίδευση, ταξινόμηση των τομών του συνόλου δεδομένων επικύρωσης της Παραλλαγής 1 στις κατηγορίες Covid-19 και Normal.

**Πίνακας 3.5: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Τομή: Μετρικές αξιολόγησης διαδικασίας ταξινόμησης των δεδομένων επικύρωσης στις κατηγορίες Covid-19 και Normal**

Μοντέλο	Μέση Ακρίβεια	Μέση Απώλεια	Μέση Αξιοπιστία	Μέση Ανάκληση	Μέση F1 Score
InceptionV3	0,9647 ± 0,0028	0,0966 ± 0,0046	0,9865 ± 0,0037	0,9398 ± 0,0051	0,9626 ± 0,0031
ResNet50	0,9818 ± 0,0029	0,0507 ± 0,0057	0,9935 ± 0,0019	0,9686 ± 0,0044	0,9809 ± 0,0031
DenseNet121	0,9768 ± 0,0017	0,0637 ± 0,0077	0,9933 ± 0,0020	0,9585 ± 0,0027	0,9756 ± 0,0018
Xception	0,9583 ± 0,0035	0,1103 ± 0,0078	0,9775 ± 0,0012	0,9352 ± 0,0079	0,9558 ± 0,0039

Παρατηρώντας τις μετρικές αξιολόγησης στον Πίνακα 3.5 διαπιστώνουμε πως όλα τα μοντέλα επιτυγχάνουν πολύ υψηλή μέση τιμή για όλες τις μετρικές.

Τις υψηλότερες τιμές όσον αφορά τη Μέση Ακρίβεια και Μέση F1 Score και τη χαμηλότερη τιμή όσον αφορά τη Μέση Απώλεια, επιτυγχάνουν τα μοντέλα ResNet50 και DenseNet121. Πιο συγκεκριμένα, το μοντέλο ResNet50 επιτυγχάνει τιμές 0,9818, 0,9809 και 0,0507 όσον αφορά τη Μέση Ακρίβεια, Μέση F1 Score και Μέση Απώλεια αντίστοιχα. Το μοντέλο DenseNet121 επιτυγχάνει τιμές 0,9768, 0,9756 και 0,0637 για τις ίδιες μετρικές.

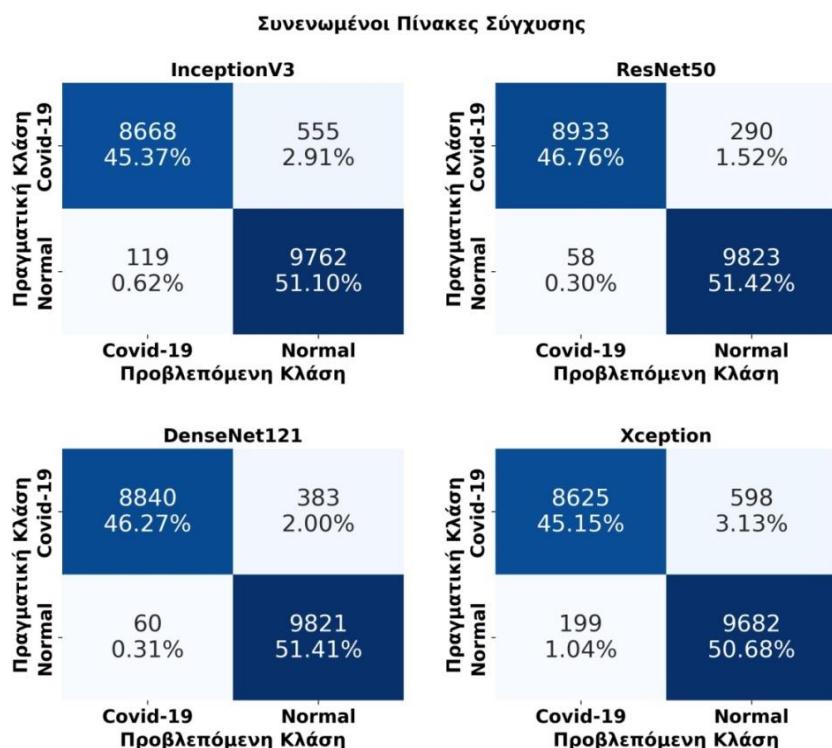
Από την άλλη, το μοντέλο που εμφανίζει τις χαμηλότερες τιμές όσον αφορά την Μέση Ακρίβεια και Μέση F1 Score και την υψηλότερη τιμή όσον αφορά τη Μέση Απώλεια, είναι το μοντέλο Xception. Πιο συγκεκριμένα, το μοντέλο Xception επιτυγχάνει τιμές 0,9583, 0,9558 και 0,1103 όσον αφορά τη Μέση Ακρίβεια, Μέση F1 Score και Μέση Απώλεια αντίστοιχα. Το μοντέλο InceptionV3 επιτυγχάνει τιμές 0,9647, 0,9626 και 0,0966 για τις ίδιες μετρικές.

Όσον αφορά τη Μέση Αξιοπιστία, τα μοντέλα InceptionV3, ResNet50, DenseNet121 και Xception παρουσιάζουν τιμές 0,9865, 0,9935, 0,9933 και 0,9775

αντίστοιχα. Η υψηλή μέση τιμή Αξιοπιστίας των μοντέλων δείχνει πως είναι ικανά να πραγματοποιούν με μεγάλη ακρίβεια ταξινόμηση των τομών που εμφανίζουν αλλοιώσεις από πνευμονία COVID-19.

Όσον αφορά τη Μέση Ανάκληση, τα μοντέλα InceptionV3, ResNet50, DenseNet121 και Xception παρουσιάζουν μέση τιμή 0,9398 0,9686 0,9585 και 0,9352 αντίστοιχα. Η υψηλή μέση τιμή Ανάκλησης των μοντέλων δείχνει πως είναι ικανά να πραγματοποιούν ταξινόμηση ενός μεγάλου πλήθους τομών που παρουσιάζουν αλλοιώσεις από πνευμονία COVID-19.

Στην Εικόνα 3.20 παρουσιάζονται οι πίνακες σύγχυσης των μοντέλων έπειτα από την ταξινόμηση των τομών του συνόλου δεδομένων επικύρωσης της Παραλλαγής 1 στις κατηγορίες Covid-19 και Normal.



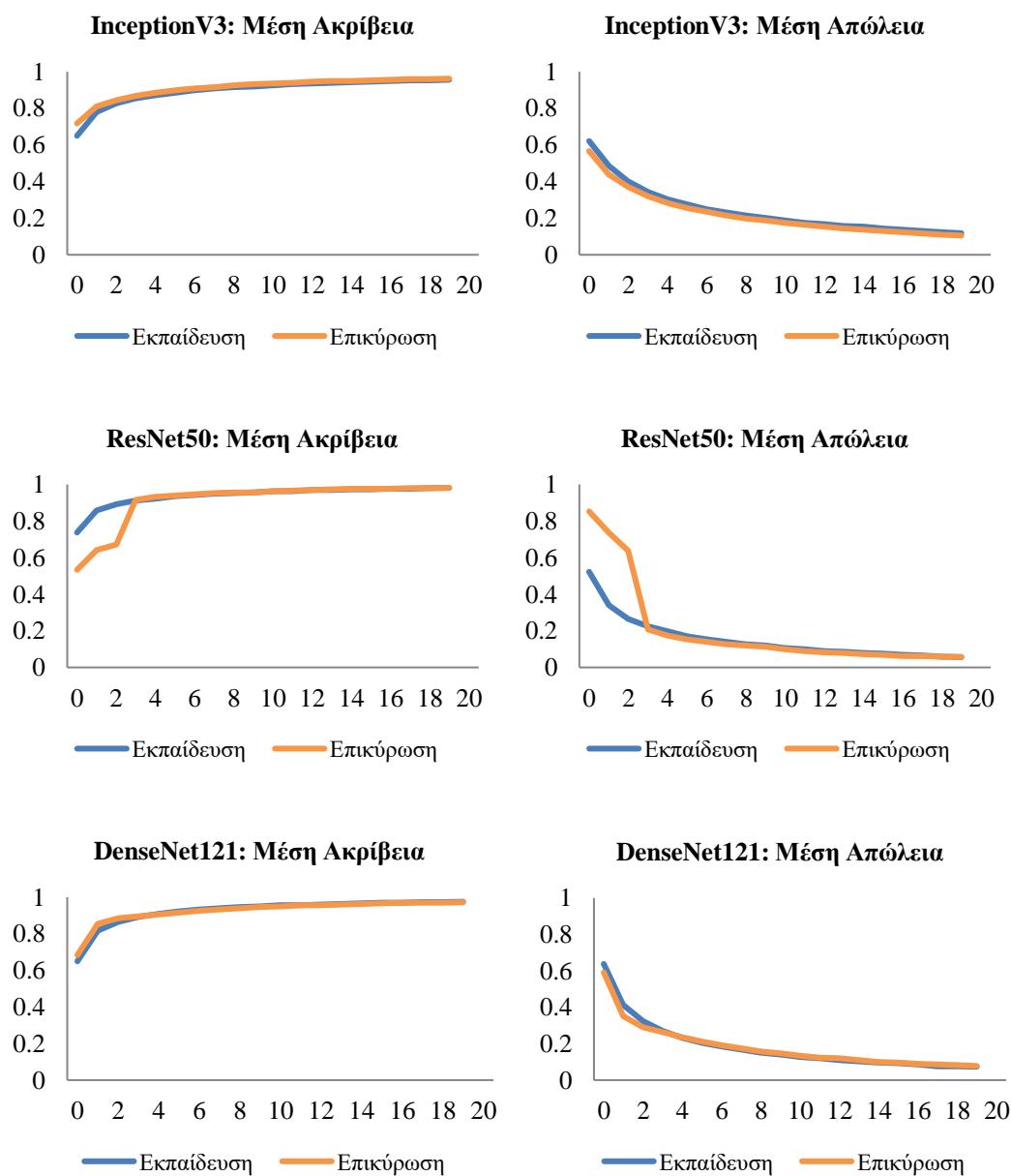
**Εικόνα 3.20: Μη Κατάτμηση Πνευμονικού Παρεγγύματος - Διαχωρισμός Δεδομένων ανά Τομή:**  
Συνενωμένοι πίνακες σύγχυσης διαδικασίας ταξινόμησης των δεδομένων επικύρωσης στις κατηγορίες Covid-19 και Normal

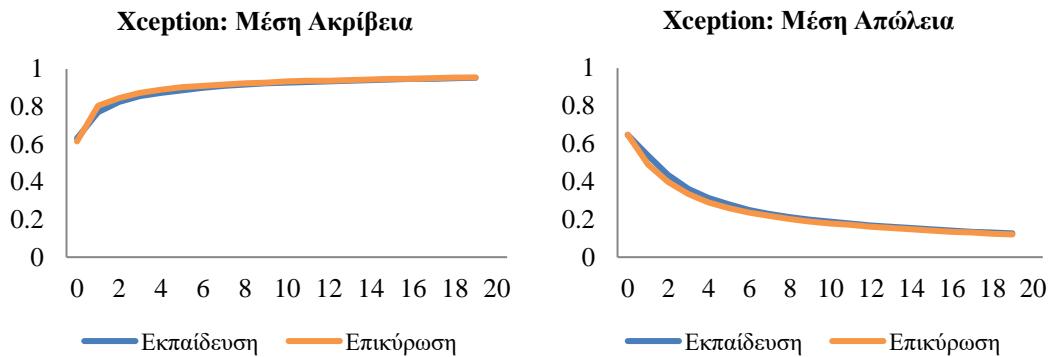
Παρατηρώντας τους πίνακες σύγχυσης στην Εικόνα 3.20, διαπιστώνουμε πως όλα τα μοντέλα που εκπαιδεύθηκαν στο σύνολο εκπαίδευσης πραγματοποίησαν την ταξινόμηση των τομών αξονικών τομογραφιών του συνόλου επικύρωσης στις κατηγορίες Covid-19 και Normal του προβλήματος σε πολύ ικανοποιητικό βαθμό. Τα μοντέλα ResNet50 και DenseNet121 είναι αυτά που ταξινομούν ορθά τις

περισσότερες εικόνες, όπως φαίνεται και από τις μετρικές αξιολόγησης στον Πίνακα 3.5.

Ακόμη, μπορεί να παρατηρηθεί πως τα Ψευδώς Αρνητικά αποτελέσματα είναι περισσότερα από τα Ψευδώς Θετικά. Αυτό σημαίνει πως τα μοντέλα δυσκολεύονται περισσότερο στο να ταξινομήσουν ορθά τομές που εμφανίζουν αλλοιώσεις που οφείλονται σε πνευμονία COVID-19.

Στην Εικόνα 3.21 παρουσιάζονται οι καμπύλες εκπαίδευσης και επικύρωσης των μοντέλων όσον αφορά την ταξινόμηση των δεδομένων της Παραλλαγής 1 στις κατηγορίες Covid-19 και Normal.





**Εικόνα 3.21: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Τομή: Καμπύλες εκπαίδευσης και επικύρωσης των μοντέλων όσον αφορά την ταξινόμηση στις κατηγορίες Covid-19 και Normal**

Παρατηρώντας τις καμπύλες στην Εικόνα 3.21 διαπιστώνουμε πως όλα τα μοντέλα που εκπαιδεύθηκαν και αξιολογήθηκαν αντίστοιχα στα σύνολα εκπαίδευσης και επικύρωσης της Παραλλαγής 1, συγκλίνουν μετά το πέρας της εκπαιδευτικής διαδικασίας.

Επιπλέον, το γεγονός ότι η απόσταση ανάμεσα στην καμπύλη εκπαίδευσης και στην καμπύλη επικύρωσης τόσο στη γραφική παράσταση Μέσης Ακρίβειας όσο και στη γραφική παράσταση Μέσης Απόλειας είναι μικρή και παραμένει σταθερή καθ' όλη τη διάρκεια της εκπαιδευτικής διαδικασίας για όλα τα μοντέλα, αποτελεί ένδειξη ότι τα μοντέλα δεν παρουσιάζουν υπερπροσαρμογή και ενδεχομένως έχουν μεγάλη ικανότητα γενίκευσης.

Στον Πίνακα 3.6 παρουσιάζονται οι μετρικές αξιολόγησης της ικανότητας των μοντέλων να πραγματοποιούν, έπειτα από εκπαίδευση, ταξινόμηση των τομών του συνόλου δεδομένων ελέγχου της Παραλλαγής 1 στις κατηγορίες Covid-19 και Normal.

**Πίνακας 3.6: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Τομή: Μετρικές αξιολόγησης της ταξινόμησης των δεδομένων ελέγχου στις κατηγορίες Covid-19 και Normal**

Μοντέλο	Μέση Ακρίβεια	Μέση Απώλεια	Μέση Αξιοπιστία	Μέση Ανάκληση	Μέση F1 Score
<b>InceptionV3</b>	0,9030 ± 0,0150	0,2522 ± 0,0376	0,9955 ± 0,0012	0,8096 ± 0,0309	0,8927 ± 0,0181
<b>ResNet50</b>	0,9527 ± 0,0140	0,1340 ± 0,0362	0,9811 ± 0,0048	0,9233 ± 0,0305	0,9511 ± 0,0153
<b>DenseNet121</b>	0,9752 ± 0,0053	0,0888 ± 0,0142	0,9719 ± 0,0103	0,9788 ± 0,0074	0,9753 ± 0,0052
<b>Xception</b>	0,9378 ± 0,0019	0,1632 ± 0,0085	0,9142 ± 0,0057	0,9663 ± 0,0072	0,9395 ± 0,0019

Παρατηρώντας τις μετρικές αξιολόγησης στον Πίνακα 3.6 διαπιστώνουμε πως όλα τα μοντέλα επιτυγχάνουν πολύ υψηλή μέση τιμή για όλες τις μετρικές.

Τις υψηλότερες τιμές όσον αφορά τη Μέση Ακρίβεια και Μέση F1 Score και τη χαμηλότερη τιμή όσον αφορά τη Μέση Απώλεια, επιτυγχάνει τα μοντέλο DenseNet121. Πιο συγκεκριμένα, το μοντέλο DenseNet121 επιτυγχάνει τιμές 0,9752, 0,9753 και 0,0888 όσον αφορά τη Μέση Ακρίβεια, Μέση F1 Score και Μέση Απώλεια αντίστοιχα.

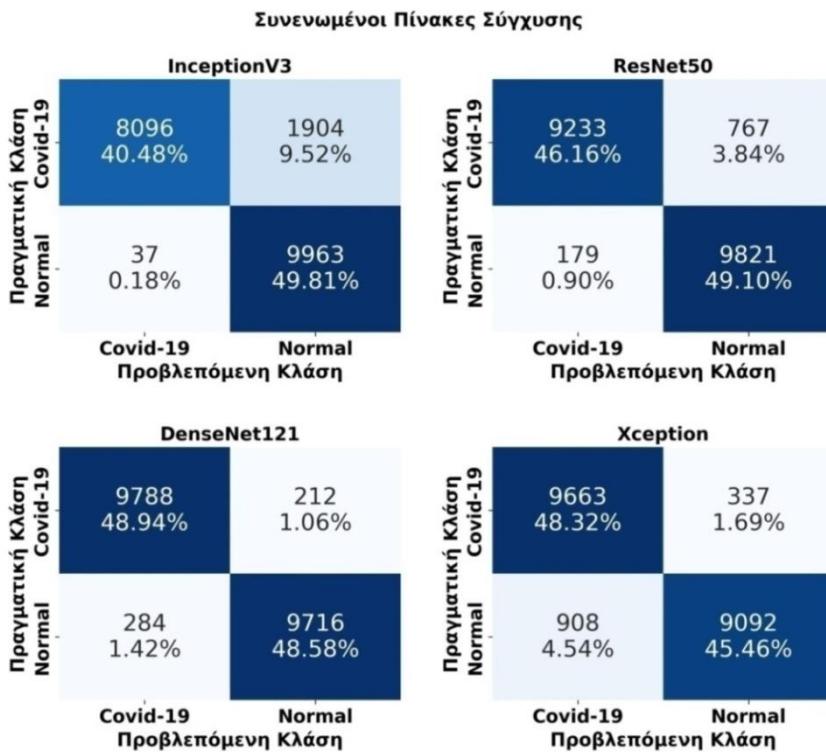
Από την άλλη, το μοντέλο που εμφανίζει τις χαμηλότερες τιμές όσον αφορά τη Μέση Ακρίβεια και Μέση F1 Score και την υψηλότερη τιμή όσον αφορά τη Μέση Απώλεια, είναι το μοντέλο InceptionV3. Πιο συγκεκριμένα, το μοντέλο InceptionV3 επιτυγχάνει τιμές 0,9030, 0,8927 και 0,2522 όσον αφορά τη Μέση Ακρίβεια, Μέση F1 Score και Μέση Απώλεια αντίστοιχα.

Το μοντέλο ResNet50 επιτυγχάνει τιμές 0,9527, 0,9511 και 0,1340 όσον αφορά τη Μέση Ακρίβεια, Μέση F1 Score και Μέση Απώλεια αντίστοιχα. Το μοντέλο Xception επιτυγχάνει τιμές 0,9378, 0,9395 και 0,1632 για τις ίδιες μετρικές.

Όσον αφορά τη Μέση Αξιοπιστία, τα μοντέλα InceptionV3, ResNet50, DenseNet121 και Xception παρουσιάζουν τιμές 0,9955, 0,9811, 0,9719 και 0,9142 αντίστοιχα. Η υψηλή μέση τιμή Αξιοπιστίας των μοντέλων δείχνει πως είναι ικανά να πραγματοποιούν με μεγάλη ακρίβεια ταξινόμηση των τομών με βάση την ύπαρξη ή μη αλλοιώσεων από πνευμονία COVID-19.

Όσον αφορά τη Μέση Ανάκληση, τα μοντέλα InceptionV3, ResNet50, DenseNet121 και Xception παρουσιάζουν τιμές 0,8096, 0,9233, 0,9788 και 0,9663 αντίστοιχα. Η υψηλή μέση τιμή Ανάκλησης των μοντέλων δείχνει πως είναι ικανά να πραγματοποιούν ταξινόμηση ενός μεγάλου πλήθους τομών που παρουσιάζουν αλλοιώσεις από πνευμονία COVID-19.

Στην Εικόνα 3.22 παρουσιάζονται οι πίνακες σύγχυσης των μοντέλων έπειτα από την ταξινόμηση των τομών του συνόλου δεδομένων ελέγχου της Παραλλαγής 1 στις κατηγορίες Covid-19 και Normal.



**Εικόνα 3.22: Μη Κατάτημη Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Τομή:**  
Συνενωμένοι πίνακες σύγχυσης διαδικασίας ταξινόμησης των δεδομένων ελέγχου στις κατηγορίες Covid-19 και Normal

Παρατηρώντας τους πίνακες σύγχυσης στην Εικόνα 3.22, διαπιστώνουμε πως όλα τα μοντέλα που εκπαιδεύθηκαν στο σύνολο εκπαίδευσης πραγματοποίησαν τη ταξινόμηση των τομών του συνόλου ελέγχου σε πολύ ικανοποιητικό βαθμό. Το μοντέλο DenseNet121 είναι αυτό που ταξινομεί ορθά τις περισσότερες εικόνες, όπως φαίνεται και από τις μετρικές αξιολόγησης στον Πίνακα 3.6.

Από το μεγάλο αριθμό των Ψευδώς Αρνητικών αποτελεσμάτων στους πίνακες σύγχυσης των μοντέλων InceptionV3 και ResNet50, μπορεί να διαπιστωθεί πως τα συγκεκριμένα μοντέλα δυσκολεύονται περισσότερο στο να κατηγοριοποιήσουν ορθά τις εικόνες που παρουσιάζουν αλλοιώσεις από πνευμονία COVID-19. Αξίζει να σημειωθεί, πως με το μοντέλο Xception συμβαίνει το ακριβώς αντίθετο, δηλαδή εμφανίζει μεγάλο αριθμό Ψευδώς Θετικών αποτελεσμάτων στον πίνακα σύγχυσης, πράγμα που σημαίνει πως δυσκολεύεται περισσότερο στο να ταξινομήσει ορθά τις εικόνες που δεν παρουσιάζουν αλλοιώσεις.

### 3.6.1.2 Παραλλαγή 2: Μη Κατάτμηση Πνευμονικού Παρεγχύματος – Διαχωρισμός Δεδομένων ανά Ασθενή

Στον Πίνακα 3.7 παρουσιάζονται οι μετρικές αξιολόγησης της ικανότητας των μοντέλων να πραγματοποιούν, έπειτα από εκπαίδευση, ταξινόμηση των τομών του συνόλου δεδομένων επικύρωσης της Παραλλαγής 2 στις κατηγορίες Covid-19 και Normal.

**Πίνακας 3.7: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Ασθενή:**  
Μετρικές αξιολόγησης της ταξινόμησης των δεδομένων επικύρωσης στις κατηγορίες Covid-19 και Normal

Μοντέλο	Μέση Ακρίβεια	Μέση Απώλεια	Μέση Αξιοπιστία	Μέση Ανάκληση	Μέση F1 Score
InceptionV3	0,9337 ± 0,0154	0,1852 ± 0,0454	0,9445 ± 0,0327	0,9174 ± 0,0257	0,9302 ± 0,0159
ResNet50	0,9535 ± 0,0155	0,1506 ± 0,0585	0,9636 ± 0,0352	0,9401 ± 0,0211	0,9512 ± 0,0159
DenseNet121	0,9590 ± 0,0144	0,1212 ± 0,0470	0,9719 ± 0,0307	0,9435 ± 0,0235	0,9570 ± 0,0142
Xception	0,9299 ± 0,0150	0,1842 ± 0,0357	0,9391 ± 0,0275	0,9146 ± 0,0196	0,9264 ± 0,0148

Παρατηρώντας τις μετρικές αξιολόγησης στον Πίνακα 3.7 διαπιστώνουμε πως όλα τα μοντέλα επιτυγχάνουν πολύ υψηλή μέση τιμή για όλες τις μετρικές.

Τις υψηλότερες τιμές όσον αφορά τη Μέση Ακρίβεια και Μέση F1 Score και τη χαμηλότερη τιμή όσον αφορά τη Μέση Απώλεια, επιτυγχάνουν τα μοντέλα ResNet50 και DenseNet121. Πιο συγκεκριμένα, το μοντέλο ResNet50 επιτυγχάνει τιμές 0,9535, 0,9512 και 0,1506 όσον αφορά τη Μέση Ακρίβεια, Μέση F1 Score και Μέση Απώλεια αντίστοιχα. Το μοντέλο DenseNet121 επιτυγχάνει τιμές 0,9590, 0,9570 και 0,1212 για τις ίδιες μετρικές.

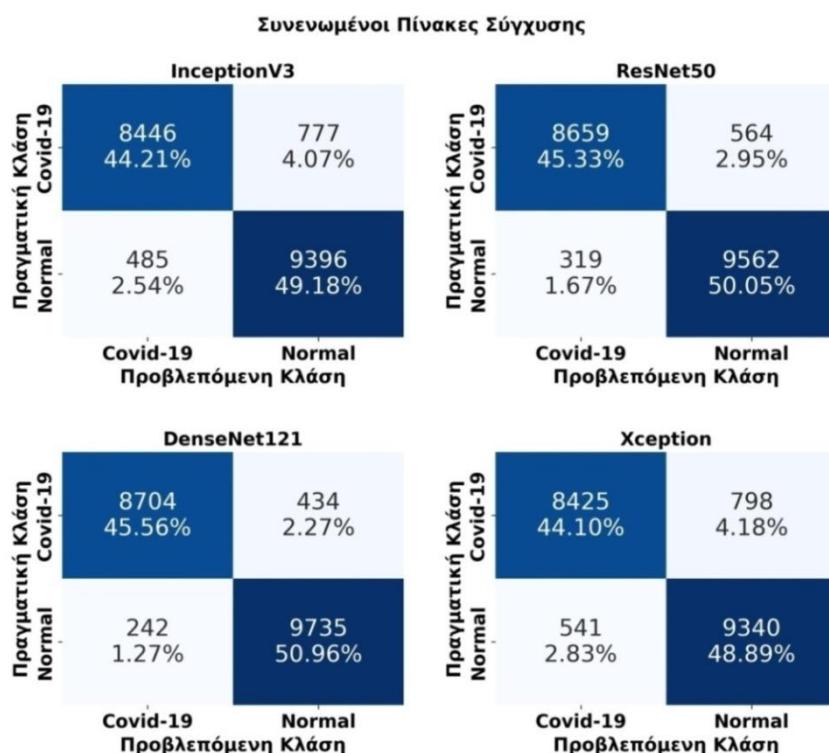
Από την άλλη, το μοντέλο που εμφανίζει τις χαμηλότερες τιμές όσον αφορά τις μετρικές αξιολόγησης Μέση Ακρίβεια και Μέση F1 Score και την υψηλότερη τιμή όσον αφορά τη Μέση Απώλεια, είναι το μοντέλο Xception. Πιο συγκεκριμένα, το μοντέλο Xception επιτυγχάνει τιμές 0,9299, 0,9264 και 0,1842 όσον αφορά τις μετρικές αξιολόγησης Μέση Ακρίβεια, Μέση F1 Score και Μέση Απώλεια αντίστοιχα. Το μοντέλο InceptionV3 επιτυγχάνει τιμές 0,9337, 0,9302 και 0,1852 για τις ίδιες μετρικές.

Όσον αφορά τη Μέση Αξιοπιστία, τα μοντέλα InceptionV3, ResNet50, DenseNet121 και Xception παρουσιάζουν τιμές 0,9445, 0,9636, 0,9719 και 0,9391 αντίστοιχα. Η υψηλή μέση τιμή Αξιοπιστίας των μοντέλων δείχνει πως είναι ικανά να

πραγματοποιούν με μεγάλη ακρίβεια ταξινόμηση των τομών με βάση την ύπαρξη ή μη αλλοιώσεων από πνευμονία COVID-19.

Όσον αφορά τη Μέση Ανάκληση, τα μοντέλα InceptionV3, ResNet50, DenseNet121 και Xception παρουσιάζουν τιμές 0,9174, 0,9401, 0,9435 και 0,9146 αντίστοιχα. Η υψηλή μέση τιμή Ανάκλησης των μοντέλων δείχνει πως είναι ικανά να πραγματοποιούν ταξινόμηση ενός μεγάλου πλήθους τομών που παρουσιάζουν αλλοιώσεις από πνευμονία COVID-19.

Στην Εικόνα 3.23 παρουσιάζονται οι πίνακες σύγχυσης των μοντέλων έπειτα από την ταξινόμηση των τομών του συνόλου δεδομένων επικύρωσης της Παραλλαγής 2 στις κατηγορίες Covid-19 και Normal.

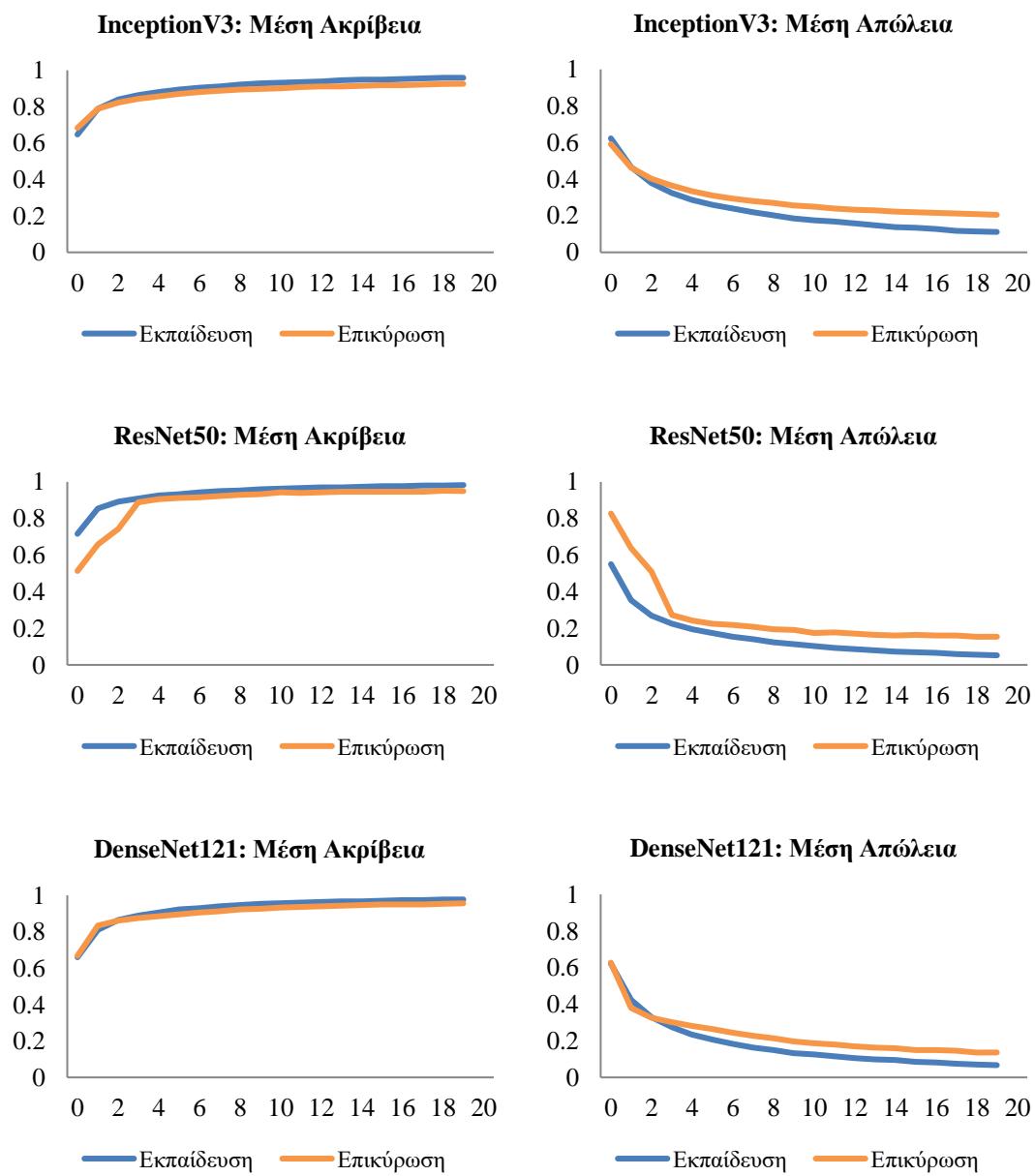


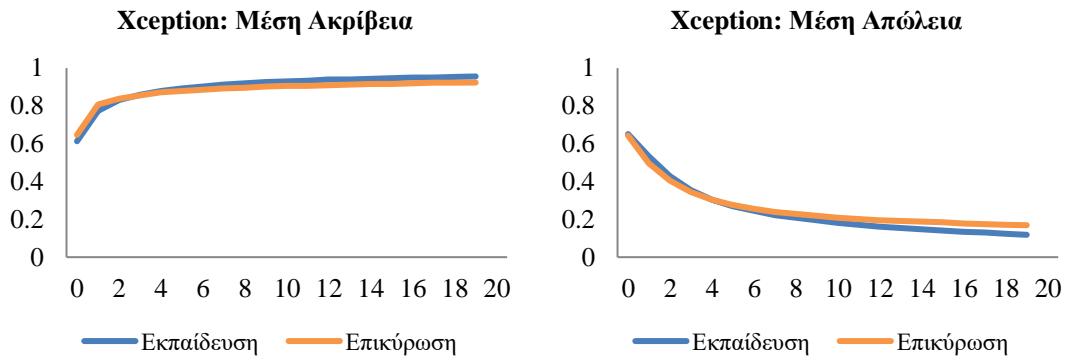
**Εικόνα 3.23: Μη Κατάτμηση Πνευμονικού Παρεγγύματος - Διαχωρισμός Δεδομένων ανά Ασθενή: Συνενωμένοι πίνακες σύγχυσης διαδικασίας ταξινόμησης των δεδομένων επικύρωσης στις κατηγορίες Covid-19 και Normal**

Παρατηρώντας τους πίνακες σύγχυσης στην Εικόνα 3.23, διαπιστώνουμε πως όλα τα μοντέλα που εκπαιδεύθηκαν στο σύνολο εκπαίδευσης πραγματοποίησαν την ταξινόμηση των τομών αξονικών τομογραφιών του συνόλου επικύρωσης σε πολύ ικανοποιητικό βαθμό. Τα μοντέλα ResNet50 και DenseNet121 είναι αυτά που ταξινομούν ορθά τις περισσότερες εικόνες, όπως φαίνεται και από τις μετρικές αξιολόγησης στον Πίνακα 3.7.

Ακόμη, μπορεί να παρατηρηθεί πως τα Ψευδώς Αρνητικά αποτελέσματα είναι περισσότερα από τα Ψευδώς Θετικά. Αυτό σημαίνει πως τα μοντέλα δυσκολεύονται περισσότερο στο να ταξινομήσουν ορθά τομές που εμφανίζουν αλλοιώσεις που οφείλονται σε πνευμονία COVID-19.

Στην Εικόνα 3.24 παρουσιάζονται οι καμπύλες εκπαίδευσης και επικύρωσης των μοντέλων όσον αφορά την ταξινόμηση των δεδομένων της Παραλλαγής 2 στις κατηγορίες Covid-19 και Normal.





**Εικόνα 3.24: Μη Κατάτμηση Πνευμονικού Παρεγγύματος - Διαχωρισμός Δεδομένων ανά Ασθενή: Καμπύλες εκπαίδευσης και επικύρωσης των μοντέλων όσον αφορά την ταξινόμηση στις κατηγορίες Covid-19 και Normal**

Παρατηρώντας τις καμπύλες εκπαίδευσης και επικύρωσης στην Εικόνα 3.24, διαπιστώνουμε πως όλα τα μοντέλα που εκπαιδεύθηκαν και αξιολογήθηκαν αντίστοιχα στα σύνολα εκπαίδευσης και επικύρωσης της Παραλλαγής 2 συγκλίνουν μετά το πέρας της εκπαιδευτικής διαδικασίας. Υπάρχει μία απόσταση ανάμεσα στην καμπύλη εκπαίδευσης και στην καμπύλη επικύρωσης, η οποία παραμένει σταθερή καθ' όλη τη διάρκεια της εκπαιδευτικής διαδικασίας για όλα τα μοντέλα. Η απόσταση αυτή είναι μεγαλύτερη για τα μοντέλα InceptionV3 και ResNet50 και μικρότερη για τα μοντέλα DenseNet121 και Xception. Αυτό αποτελεί ένδειξη πως τα μοντέλα DenseNet121 και Xception παρουσιάζουν μικρότερη υπερπροσαρμογή από τα μοντέλα InceptionV3 και ResNet50 και ενδεχομένως έχουν μεγαλύτερη ικανότητα γενίκευσης.

Στον Πίνακα 3.8 παρουσιάζονται οι μετρικές αξιολόγησης της ικανότητας των μοντέλων να πραγματοποιούν, έπειτα από εκπαίδευση, ταξινόμηση των τομών του συνόλου δεδομένων ελέγχου της Παραλλαγής 2 στις κατηγορίες Covid-19 και Normal.

**Πίνακας 3.8: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Ασθενή:**  
**Μετρικές αξιολόγησης της ταξινόμησης των δεδομένων ελέγχου στις κατηγορίες Covid-19 και Normal**

Μοντέλο	Μέση Ακρίβεια	Μέση Απώλεια	Μέση Αξιοπιστία	Μέση Ανάκληση	Μέση F1 Score
<b>InceptionV3</b>	0,8987 ± 0,0253	0,2674 ± 0,0587	0,9884 ± 0,0099	0,8073 ± 0,0593	0,8875 ± 0,0310
<b>ResNet50</b>	0,9475 ± 0,0121	0,1484 ± 0,0316	0,9730 ± 0,0157	0,9211 ± 0,0350	0,9458 ± 0,0138
<b>DenseNet121</b>	0,9700 ± 0,0037	0,0946 ± 0,0100	0,9743 ± 0,0170	0,9659 ± 0,0174	0,9698 ± 0,0036
<b>Xception</b>	0,9419 ± 0,0072	0,1562 ± 0,0158	0,9272 ± 0,0114	0,9591 ± 0,0100	0,9428 ± 0,0070

Παρατηρώντας τις μετρικές αξιολόγησης στον Πίνακα 3.8 διαπιστώνουμε πως όλα τα μοντέλα επιτυγχάνουν πολύ υψηλή μέση τιμή για όλες τις μετρικές.

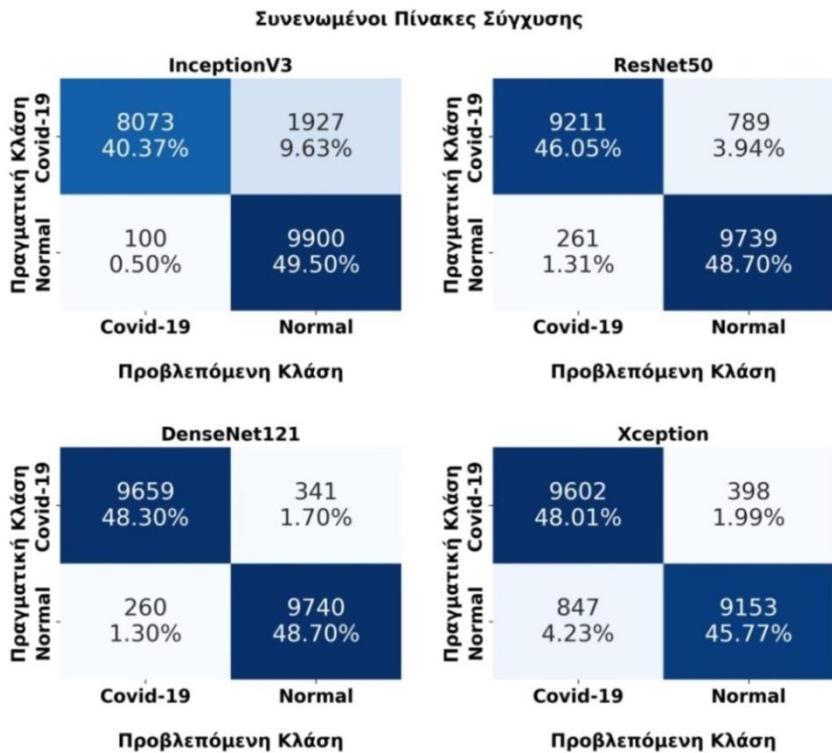
Το μοντέλο που επιτυγχάνει την υψηλότερη τιμή όσον αφορά τις μετρικές αξιολόγησης Μέση Ακρίβεια και Μέση F1 Score και τη χαμηλότερη τιμή όσον αφορά τη μετρική Μέση Απώλεια, είναι το μοντέλο DenseNet121. Πιο συγκεκριμένα, το μοντέλο DenseNet121 επιτυγχάνει τιμές 0,9700, 0,9698 και 0,0946 όσον αφορά τις μετρικές αξιολόγησης Μέση Ακρίβεια, Μέση F1 Score και Μέση Απώλεια αντίστοιχα.

Από την άλλη, το μοντέλο που εμφανίζει τη χαμηλότερη τιμή όσον αφορά τις μετρικές αξιολόγησης Μέση Ακρίβεια και Μέση F1 Score και την υψηλότερη τιμή όσον αφορά τη μετρική Μέση Απώλεια, είναι το μοντέλο InceptionV3. Πιο συγκεκριμένα, το μοντέλο InceptionV3 επιτυγχάνει τιμές 0,8987, 0,8875 και 0,2674 όσον αφορά τις μετρικές αξιολόγησης Μέση Ακρίβεια, Μέση F1 Score και Μέση Απώλεια αντίστοιχα. Το μοντέλο ResNet50 επιτυγχάνει τιμές 0,9475, 0,9458 και 0,1484 όσον αφορά τις μετρικές αξιολόγησης Μέση Ακρίβεια, Μέση F1 Score και Μέση Απώλεια αντίστοιχα. Το μοντέλο Xception επιτυγχάνει τιμές 0,9419, 0,9428 και 0,1562 για τις ίδιες μετρικές.

Όσον αφορά τη Μέση Αξιοπιστία, τα μοντέλα InceptionV3, ResNet50, DenseNet121 και Xception παρουσιάζουν τιμές 0,9884, 0,9730, 0,9743 και 0,9272 αντίστοιχα. Η υψηλή μέση τιμή Αξιοπιστίας των μοντέλων δείχνει πως είναι ικανά να πραγματοποιούν με μεγάλη ακρίβεια ταξινόμηση των τομών με βάση την ύπαρξη ή μη αλλοιώσεων από πνευμονία COVID-19.

Όσον αφορά τη Μέση Ανάκληση, τα μοντέλα InceptionV3, ResNet50, DenseNet121 και Xception παρουσιάζουν τιμές 0,8073, 0,9211, 0,9659 και 0,9591 αντίστοιχα. Η υψηλή μέση τιμή Ανάκλησης των μοντέλων δείχνει πως είναι ικανά να πραγματοποιούν ταξινόμηση ενός μεγάλου πλήθους τομών που παρουσιάζουν αλλοιώσεις από πνευμονία COVID-19.

Στην Εικόνα 3.25 παρουσιάζονται οι πίνακες σύγχυσης των μοντέλων έπειτα από την ταξινόμηση των τομών του συνόλου δεδομένων ελέγχου της Παραλλαγής 2 στις κατηγορίες Covid-19 και Normal.



**Εικόνα 3.25: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Ασθενή:**  
Συνενωμένοι πίνακες σύγχυσης διαδικασίας ταξινόμησης των δεδομένων ελέγχου στις κατηγορίες Covid-19 και Normal

Παρατηρώντας τους πίνακες σύγχυσης στην Εικόνα 3.25, διαπιστώνουμε πως όλα τα μοντέλα που εκπαιδεύθηκαν στο σύνολο εκπαίδευσης πραγματοποίησαν την ταξινόμηση των τομών αξονικών τομογραφιών του συνόλου ελέγχου σε πολύ ικανοποιητικό βαθμό. Το μοντέλο DenseNet121 είναι αυτό που ταξινομεί ορθά τις περισσότερες εικόνες, όπως φαίνεται και από τις μετρικές αξιολόγησης στον Πίνακα 3.8.

Από το μεγάλο αριθμό των Ψευδώς Αρνητικών αποτελεσμάτων στους πίνακες σύγχυσης των μοντέλων InceptionV3 και ResNet50, μπορεί να διαπιστωθεί πως τα συγκεκριμένα μοντέλα δυσκολεύονται περισσότερο στο να κατηγοριοποιήσουν ορθά τις εικόνες που παρουσιάζουν αλλοιώσεις από πνευμονία COVID-19. Αξίζει να σημειωθεί, πως με το μοντέλο Xception συμβαίνει το ακριβώς αντίθετο, δηλαδή εμφανίζει μεγάλο αριθμό Ψευδώς Θετικών αποτελεσμάτων στον πίνακα σύγχυσης, πράγμα που σημαίνει πως δυσκολεύεται περισσότερο στο να ταξινομήσει ορθά τις εικόνες που δεν παρουσιάζουν αλλοιώσεις.

### 3.6.1.3 Παραλλαγή 3: Κατάτμηση Πνευμονικού Παρεγχύματος – Διαχωρισμός Δεδομένων ανά Τομή

Στον Πίνακα 3.9 παρουσιάζονται οι μετρικές αξιολόγησης της ικανότητας των μοντέλων να πραγματοποιούν, έπειτα από εκπαίδευση, ταξινόμηση των τομών του συνόλου δεδομένων επικύρωσης της Παραλλαγής 3 στις κατηγορίες Covid-19 και Normal.

**Πίνακας 3.9: Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Τομή: Μετρικές αξιολόγησης της ταξινόμησης των δεδομένων επικύρωσης στις κατηγορίες Covid-19 και Normal**

Μοντέλο	Μέση Ακρίβεια	Μέση Απώλεια	Μέση Αξιοπιστία	Μέση Ανάκληση	Μέση F1 Score
<b>InceptionV3</b>	0,9736 ± 0,0032	0,0826 ± 0,0075	0,9908 ± 0,0026	0,9542 ± 0,0061	0,9722 ± 0,0034
<b>ResNet50</b>	0,9713 ± 0,0035	0,0817 ± 0,0040	0,9957 ± 0,0025	0,9447 ± 0,0052	0,9695 ± 0,0038
<b>DenseNet121</b>	0,9717 ± 0,0027	0,0817 ± 0,0040	0,9922 ± 0,0034	0,9488 ± 0,0058	0,9700 ± 0,0029
<b>Xception</b>	0,9620 ± 0,0039	0,1111 ± 0,0059	0,9872 ± 0,0052	0,9335 ± 0,0069	0,9596 ± 0,0042

Παρατηρώντας τις μετρικές αξιολόγησης στον Πίνακα 3.9 διαπιστώνουμε πως όλα τα μοντέλα επιτυγχάνουν πολύ υψηλή μέση τιμή για όλες τις μετρικές.

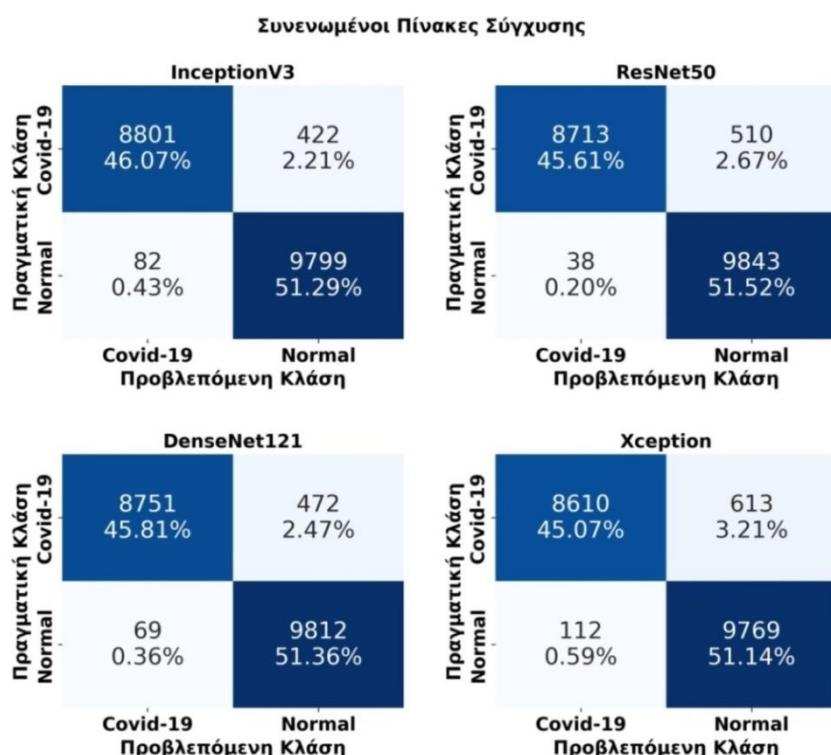
Τα μοντέλα που επιτυγχάνουν τις υψηλότερες τιμές όσον αφορά τις μετρικές αξιολόγησης Μέση Ακρίβεια και Μέση F1 Score και τη χαμηλότερη τιμή όσον αφορά τη μετρική Μέση Απώλεια, είναι τα InceptionV3, ResNet50, και DenseNet121. Πιο συγκεκριμένα, το μοντέλο InceptionV3 επιτυγχάνει τιμές 0,9736, 0,9722 και 0,0826 όσον αφορά τις μετρικές αξιολόγησης Μέση Ακρίβεια, Μέση F1 Score και Μέση Απώλεια αντίστοιχα. Το μοντέλο ResNet50 επιτυγχάνει τιμές 0,9713, 0,9695 και 0,0817 όσον αφορά τις μετρικές αξιολόγησης Μέση Ακρίβεια, Μέση F1 Score και Μέση Απώλεια αντίστοιχα. Το μοντέλο DenseNet121 επιτυγχάνει τιμές 0,9717, 0,9700 και 0,0817 όσον αφορά τις μετρικές αξιολόγησης Μέση Ακρίβεια, Μέση F1 Score και Μέση Απώλεια αντίστοιχα.

Από την άλλη, το μοντέλο που εμφανίζει τις χαμηλότερες τιμές όσον αφορά τις μετρικές αξιολόγησης Μέση Ακρίβεια και Μέση F1 Score και την υψηλότερη τιμή όσον αφορά τη μετρική Μέση Απώλεια, είναι το μοντέλο Xception. Πιο συγκεκριμένα, το μοντέλο Xception επιτυγχάνει τιμές 0,9620, 0,9596 και 0,1111 όσον αφορά τις μετρικές αξιολόγησης Μέση Ακρίβεια, Μέση F1 Score και Μέση Απώλεια αντίστοιχα.

Όσον αφορά τη Μέση Αξιοπιστία, τα μοντέλα InceptionV3, ResNet50, DenseNet121 και Xception παρουσιάζουν τιμές 0,9908 0,9957 0,9922 και 0,9872 αντίστοιχα. Η υψηλή μέση τιμή Αξιοπιστίας των μοντέλων δείχνει πως είναι ικανά να πραγματοποιούν με μεγάλη ακρίβεια ταξινόμηση των τομών με βάση την ύπαρξη ή μη αλλοιώσεων από πνευμονία COVID-19.

Όσον αφορά τη Μέση Ανάκληση, τα μοντέλα InceptionV3, ResNet50, DenseNet121 και Xception παρουσιάζουν τιμές 0,9542 0,9447 0,9488 και 0,9335 αντίστοιχα. Η υψηλή μέση τιμή Ανάκλησης των μοντέλων δείχνει πως είναι ικανά να πραγματοποιούν ταξινόμηση ενός μεγάλου πλήθους τομών που παρουσιάζουν αλλοιώσεις από πνευμονία COVID-19.

Στην Εικόνα 3.26 παρουσιάζονται οι πίνακες σύγχυσης των μοντέλων έπειτα από την ταξινόμηση των τομών του συνόλου δεδομένων επικύρωσης της Παραλλαγής 3 στις κατηγορίες Covid-19 και Normal.



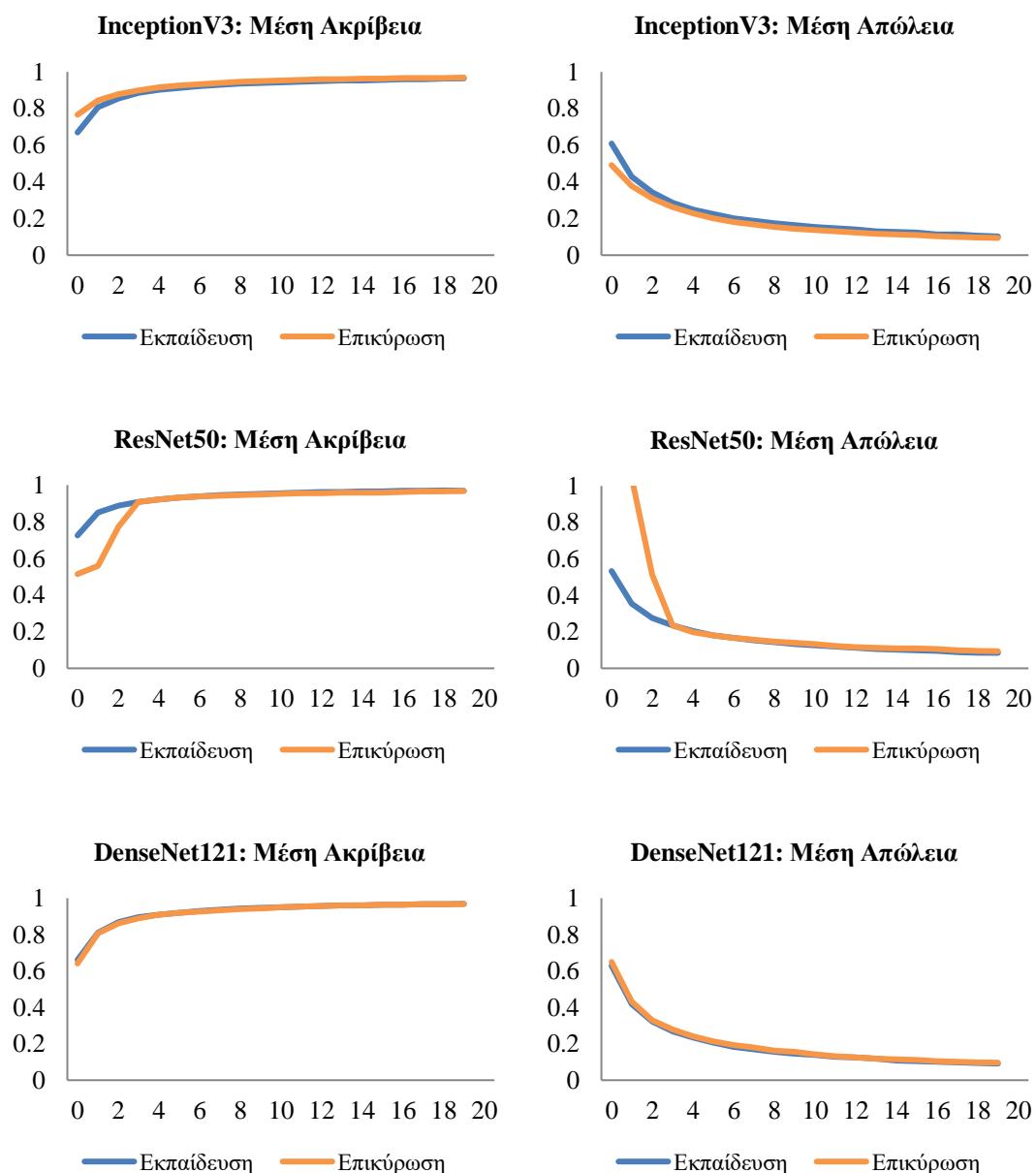
Εικόνα 3.26: Κατάτμηση Πνευμονικού Παρεγγύματος - Διαχωρισμός Δεδομένων ανά Τομή: Συνενωμένοι πίνακες σύγχυσης διαδικασίας ταξινόμησης των δεδομένων επικύρωσης στις κατηγορίες Covid-19 και Normal

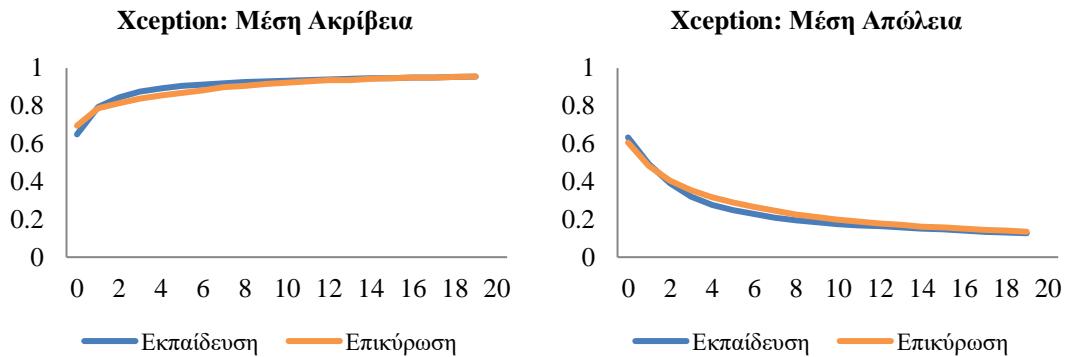
Παρατηρώντας τους πίνακες σύγχυσης στην Εικόνα 3.26, διαπιστώνουμε πως όλα τα μοντέλα που εκπαιδεύθηκαν στο σύνολο εκπαίδευσης πραγματοποίησαν την ταξινόμηση των τομών αξονικών τομογραφιών του συνόλου επικύρωσης σε πολύ ικανοποιητικό βαθμό. Τα μοντέλα InceptionV3, ResNet50 και DenseNet121 είναι

αυτά που ταξινομούν ορθά τις περισσότερες εικόνες, όπως φαίνεται και από τις μετρικές αξιολόγησης στον Πίνακα 3.9.

Ακόμη, μπορεί να παρατηρηθεί πως τα Ψευδώς Αρνητικά αποτελέσματα είναι περισσότερα από τα Ψευδώς Θετικά. Αυτό σημαίνει πως τα μοντέλα δυσκολεύονται περισσότερο στο να ταξινομήσουν ορθά τομές που εμφανίζουν αλλοιώσεις οφειλόμενες σε πνευμονία COVID-19.

Στην Εικόνα 3.27 παρουσιάζονται οι καμπύλες εκπαίδευσης και επικύρωσης των μοντέλων όσον αφορά την ταξινόμηση των δεδομένων της Παραλλαγής 3 στις κατηγορίες Covid-19 και Normal.





**Εικόνα 3.27: Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Τομή:** Καμπύλες εκπαίδευσης και επικύρωσης των μοντέλων όσον αφορά την ταξινόμηση στις κατηγορίες Covid-19 και Normal

Παρατηρώντας τις καμπύλες εκπαίδευσης και επικύρωσης στην Εικόνα 3.27, διαπιστώνουμε πως όλα τα μοντέλα που εκπαιδεύθηκαν και αξιολογήθηκαν αντίστοιχα στα σύνολα εκπαίδευσης και επικύρωσης της Παραλλαγής 3 συγκλίνουν μετά το πέρας της εκπαιδευτικής διαδικασίας. Επιπλέον, το γεγονός ότι η απόσταση ανάμεσα στη καμπύλη εκπαίδευσης και στη καμπύλη επικύρωσης τόσο στη γραφική παράσταση Μέσης Ακρίβειας όσο και στη γραφική παράσταση Μέσης Απώλειας είναι μικρή και παραμένει σταθερή καθ' όλη τη διάρκεια της εκπαιδευτικής διαδικασίας για όλα τα μοντέλα, αποτελεί ένδειξη ότι τα μοντέλα δεν παρουσιάζουν υπερπροσαρμογή και ενδεχομένως έχουν μεγάλη ικανότητα γενίκευσης.

### 3.6.1.4 Παραλλαγή 4: Κατάτμηση Πνευμονικού Παρεγχύματος – Διαχωρισμός Δεδομένων ανά Ασθενή

Στον Πίνακα 3.10 παρουσιάζονται οι μετρικές αξιολόγησης της ικανότητας των μοντέλων να πραγματοποιούν, έπειτα από εκπαίδευση, ταξινόμηση των τομών του συνόλου δεδομένων επικύρωσης της Παραλλαγής 4 στις κατηγορίες Covid-19 και Normal.

**Πίνακας 3.10: Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Ασθενή: Μετρικές αξιολόγησης της ταξινόμησης των δεδομένων επικύρωσης στις κατηγορίες Covid-19 και Normal**

Μοντέλο	Μέση Ακρίβεια	Μέση Απώλεια	Μέση Αξιοπιστία	Μέση Ανάκληση	Μέση F1 Score
InceptionV3	0,9640 ± 0,0142	0,1133 ± 0,0513	0,9863 ± 0,0065	0,9390 ± 0,0192	0,9621 ± 0,0131
ResNet50	0,9637 ± 0,0106	0,1130 ± 0,0416	0,9903 ± 0,0040	0,9343 ± 0,0159	0,9615 ± 0,0096
DenseNet121	0,9680 ± 0,0136	0,1034 ± 0,0494	0,9894 ± 0,0042	0,9444 ± 0,0222	0,9663 ± 0,0128
Xception	0,9540 ± 0,0114	0,1356 ± 0,0356	0,9775 ± 0,0075	0,9264 ± 0,0171	0,9512 ± 0,0106

Παρατηρώντας τις μετρικές αξιολόγησης στον Πίνακα 3.10 διαπιστώνουμε πως όλα τα μοντέλα επιτυγχάνουν πολύ υψηλή μέση τιμή για όλες τις μετρικές.

Τα μοντέλα που επιτυγχάνουν τις υψηλότερες τιμές όσον αφορά τις μετρικές αξιολόγησης Μέση Ακρίβεια και Μέση F1 Score και τη χαμηλότερη τιμή όσον αφορά τη μετρική Μέση Απώλεια, είναι τα InceptionV3, ResNet50, και DenseNet121. Πιο συγκεκριμένα, το μοντέλο InceptionV3 επιτυγχάνει τιμές 0,9640, 0,9621 και 0,1133 όσον αφορά τις μετρικές αξιολόγησης Μέση Ακρίβεια, Μέση F1 Score και Μέση Απώλεια αντίστοιχα. Το μοντέλο ResNet50 επιτυγχάνει τιμές 0,9637, 0,9615 και 0,1130 όσον αφορά τις μετρικές αξιολόγησης Μέση Ακρίβεια, Μέση F1 Score και Μέση Απώλεια αντίστοιχα. Το μοντέλο DenseNet121 επιτυγχάνει τιμές 0,9680, 0,9663 και 0,1034 όσον αφορά τις μετρικές αξιολόγησης Μέση Ακρίβεια, Μέση F1 Score και Μέση Απώλεια αντίστοιχα.

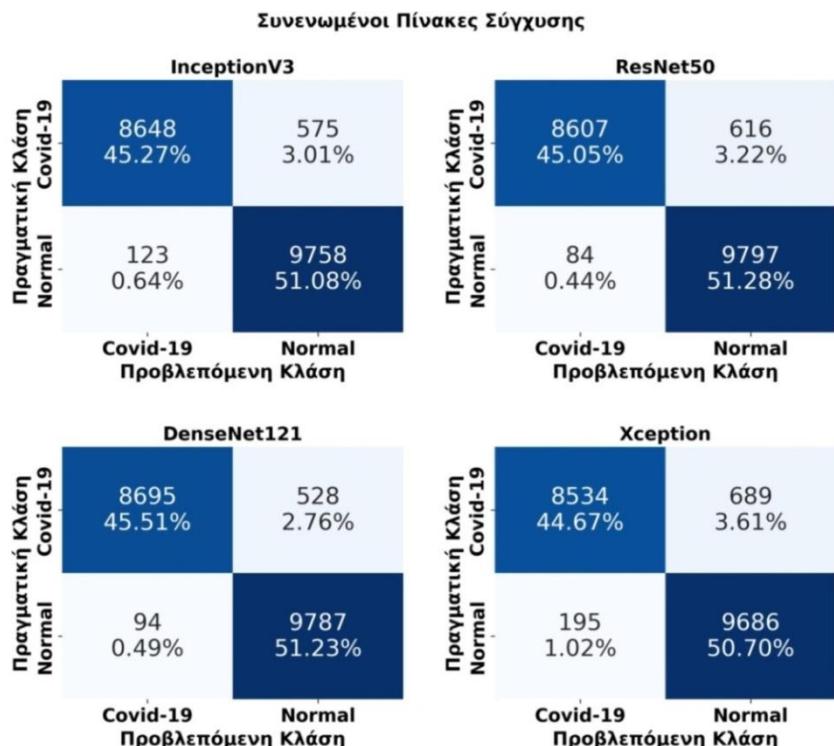
Από την άλλη, το μοντέλο που εμφανίζει τις χαμηλότερες τιμές όσον αφορά τις μετρικές αξιολόγησης Μέση Ακρίβεια και Μέση F1 Score και την υψηλότερη τιμή όσον αφορά τη μετρική Μέση Απώλεια, είναι το μοντέλο Xception. Πιο συγκεκριμένα, το μοντέλο Xception επιτυγχάνει τιμές 0,9540, 0,9512 και 0,1356 όσον αφορά τις μετρικές αξιολόγησης Μέση Ακρίβεια, Μέση F1 Score και Μέση Απώλεια αντίστοιχα.

Όσον αφορά τη Μέση Αξιοπιστία, τα μοντέλα InceptionV3, ResNet50, DenseNet121 και Xception παρουσιάζουν τιμές 0,9863, 0,9903, 0,9894 και 0,9775 αντίστοιχα. Η υψηλή μέση τιμή Αξιοπιστίας των μοντέλων δείχνει πως είναι ικανά να πραγματοποιούν με μεγάλη ακρίβεια ταξινόμηση των τομών με βάση την ύπαρξη ή μη αλλοιώσεων από πνευμονία COVID-19.

Όσον αφορά τη Μέση Ανάκληση, τα μοντέλα InceptionV3, ResNet50, DenseNet121 και Xception παρουσιάζουν τιμές 0,9390, 0,9343, 0,9444 και 0,9264 αντίστοιχα. Η υψηλή μέση τιμή Ανάκλησης των μοντέλων δείχνει πως είναι ικανά να

πραγματοποιούν ταξινόμηση ενός μεγάλου πλήθους τομών που παρουσιάζουν αλλοιώσεις από πνευμονία COVID-19.

Στην Εικόνα 3.28 παρουσιάζονται οι πίνακες σύγχυσης των μοντέλων έπειτα από την ταξινόμηση των τομών του συνόλου δεδομένων επικύρωσης της Παραλλαγής 4 στις κατηγορίες Covid-19 και Normal.

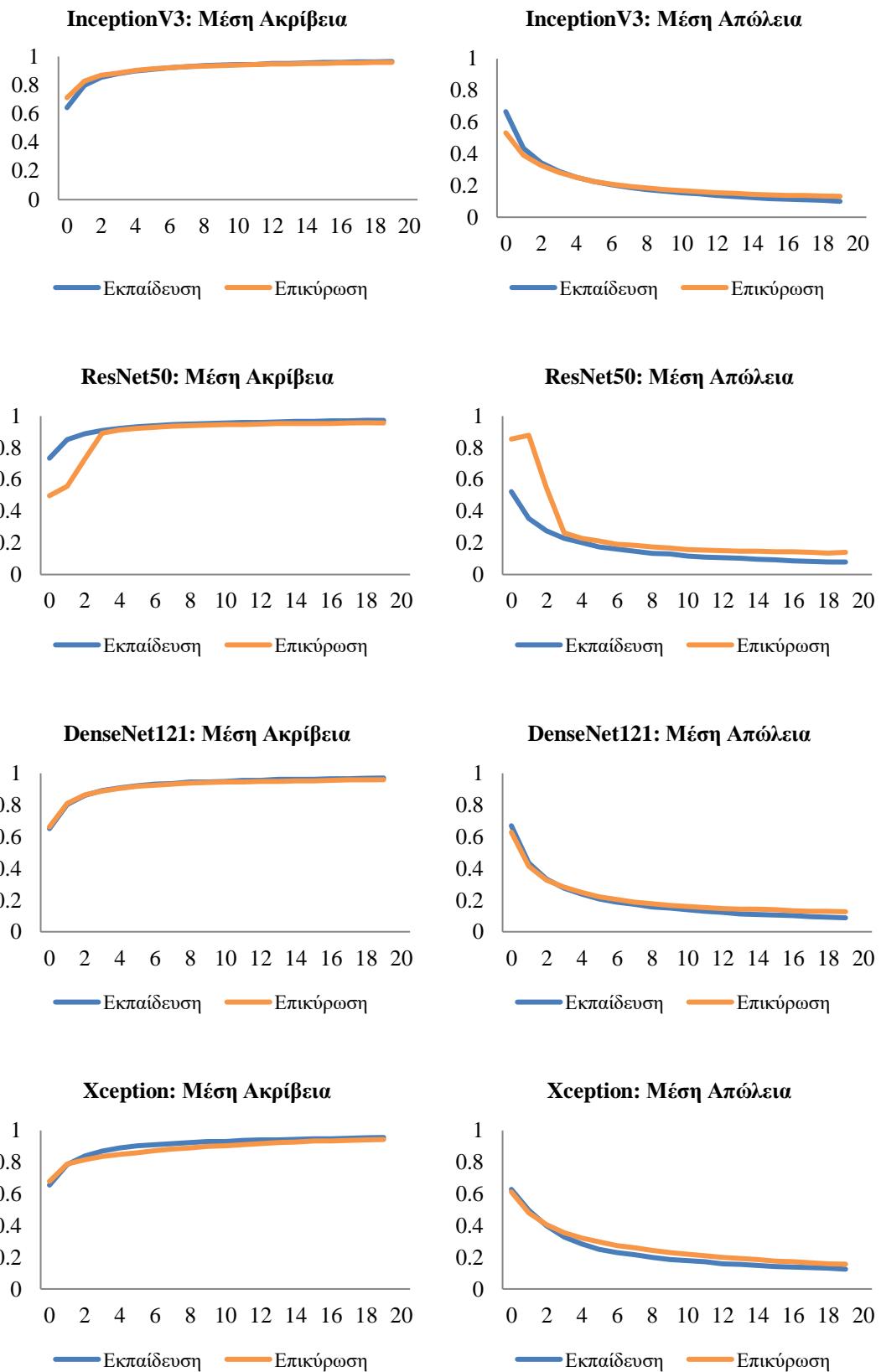


**Εικόνα 3.28: Κατάτημα Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Ασθενή: Συνενωμένοι πίνακες σύγχυσης διαδικασίας ταξινόμησης των δεδομένων επικύρωσης στις κατηγορίες Covid-19 και Normal**

Παρατηρώντας τους πίνακες σύγχυσης στην Εικόνα 3.28, διαπιστώνουμε πως όλα τα μοντέλα που εκπαιδεύθηκαν στο σύνολο εκπαίδευσης πραγματοποίησαν την ταξινόμηση των τομών του συνόλου επικύρωσης σε πολύ ικανοποιητικό βαθμό. Τα μοντέλα InceptionV3, ResNet50 και DenseNet121 είναι αυτά που ταξινομούν ορθά τις περισσότερες εικόνες, όπως φαίνεται και στα αποτελέσματα του Πίνακα 3.10.

Ακόμη, μπορεί να παρατηρηθεί πως τα Ψευδώς Αρνητικά αποτελέσματα είναι περισσότερα από τα Ψευδώς Θετικά, που σημαίνει πως τα μοντέλα δυσκολεύονται περισσότερο να ταξινομήσουν ορθά τομές που εμφανίζουν αλλοιώσεις οφειλόμενες σε πνευμονία COVID-19.

Στην Εικόνα 3.29 παρουσιάζονται οι καμπύλες εκπαίδευσης και επικύρωσης των μοντέλων όσον αφορά την ταξινόμηση των δεδομένων της Παραλλαγής 4 στις κατηγορίες Covid-19 και Normal.

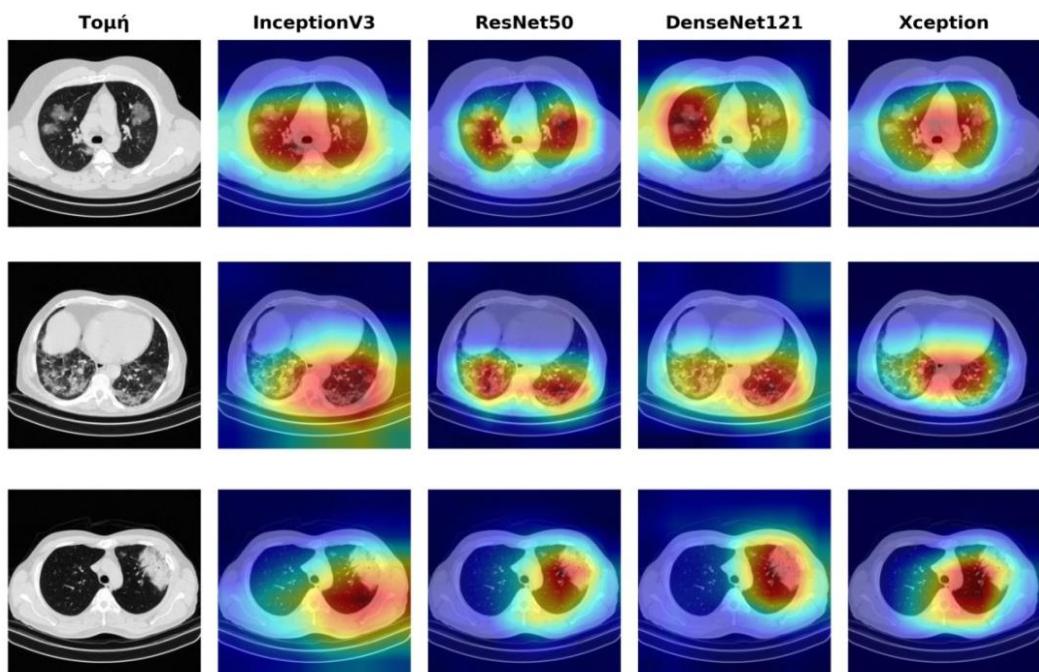


**Εικόνα 3.29: Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Ασθενή: Καμπύλες εκπαίδευσης και επικύρωσης των μοντέλων όσον αφορά την ταξινόμηση στις κατηγορίες Covid-19 και Normal**

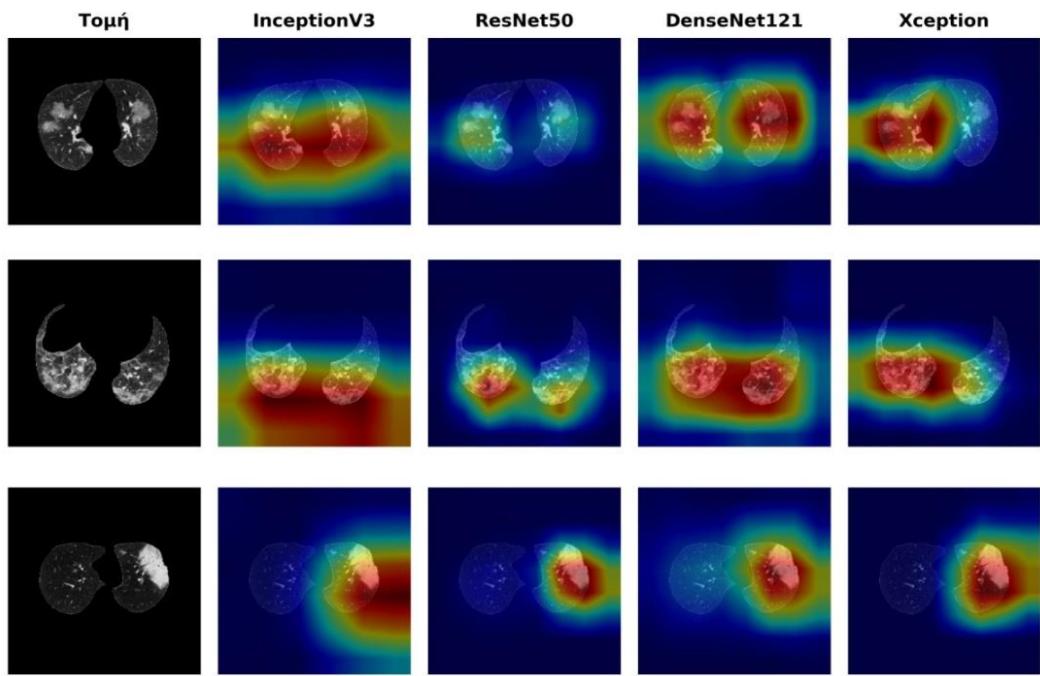
Παρατηρώντας τις καμπύλες εκπαίδευσης και επικύρωσης στην Εικόνα 3.29, διαπιστώνουμε πως όλα τα μοντέλα που εκπαιδεύθηκαν και αξιολογήθηκαν αντίστοιχα στα σύνολα εκπαίδευσης και επικύρωσης της Παραλλαγής 4 συγκλίνουν μετά το πέρας της εκπαιδευτικής διαδικασίας. Υπάρχει μία απόσταση ανάμεσα στην καμπύλη εκπαίδευσης και στην καμπύλη επικύρωσης, η οποία παραμένει σταθερή καθ' όλη τη διάρκεια της εκπαιδευτικής διαδικασίας για όλα τα μοντέλα. Η απόσταση αυτή είναι μεγαλύτερη για το μοντέλο ResNet50 και μικρότερη για τα μοντέλα InceptionV3, DenseNet121 και Xception. Αυτό αποτελεί ένδειξη πως τα μοντέλα InceptionV3, DenseNet121 και Xception παρουσιάζουν μικρότερη υπερπροσαρμογή από το μοντέλο ResNet50 και ενδεχομένως έχουν μεγαλύτερη ικανότητα γενίκευσης.

### 3.6.2 Εφαρμογή της Μεθόδου Grad-CAM σε Ταξινομημένες Τομές στις Κατηγορίες COVID-19 και Normal

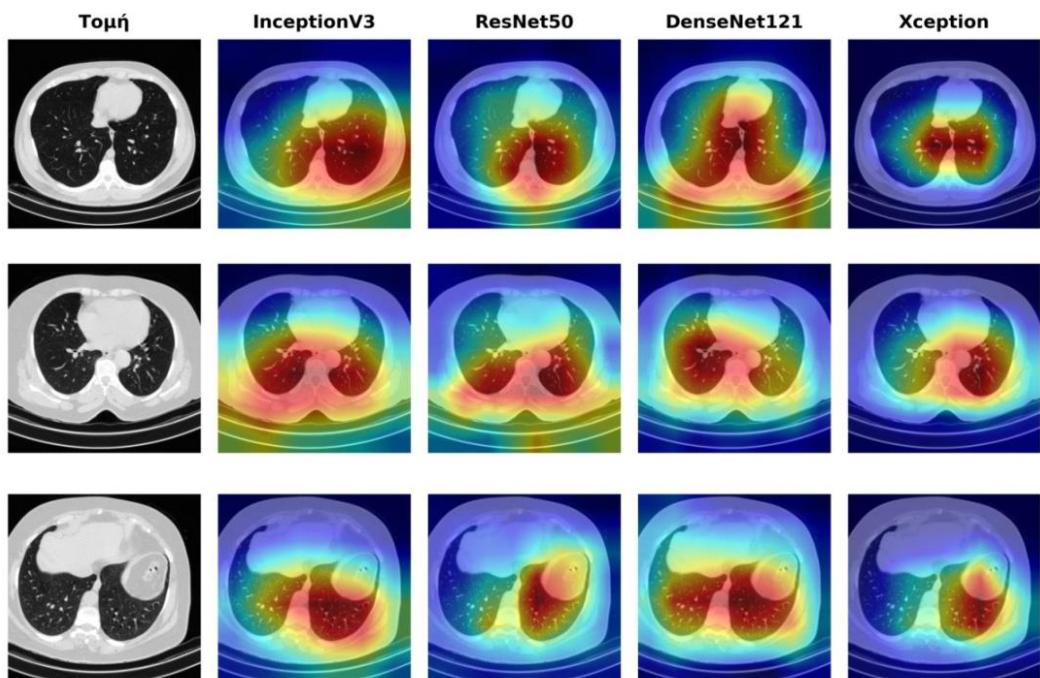
Στις Εικόνες 3.30 και 3.31 παρουσιάζεται η εφαρμογή της τεχνικής Grad-CAM σε τομές αξονικών τομογραφιών του συνόλου επικύρωσης οι οποίες ταξινομήθηκαν ορθά στην κατηγορία Covid-19 από τα μοντέλα βαθιάς μάθησης. Παρατηρώντας τις συγκεκριμένες εικόνες, διαπιστώνουμε πως όλα τα μοντέλα λαμβάνουν υπόψη τους στοιχεία που βρίσκονται εντός του πνευμονικού παρεγχύματος ούτως ώστε να πραγματοποιήσουν την ταξινόμηση των τομών. Παρατηρούμε επίσης πως οι περιοχές στις οποίες δίνουν βαρύτητα τα μοντέλα για να πραγματοποιήσουν την πρόβλεψη εμφανίζουν διαφορές από μοντέλο σε μοντέλο, με τα μοντέλα ResNet50 και DenseNet121 να εστιάζουν περισσότερο στα σημεία που εμφανίζουν αλλοιώσεις.



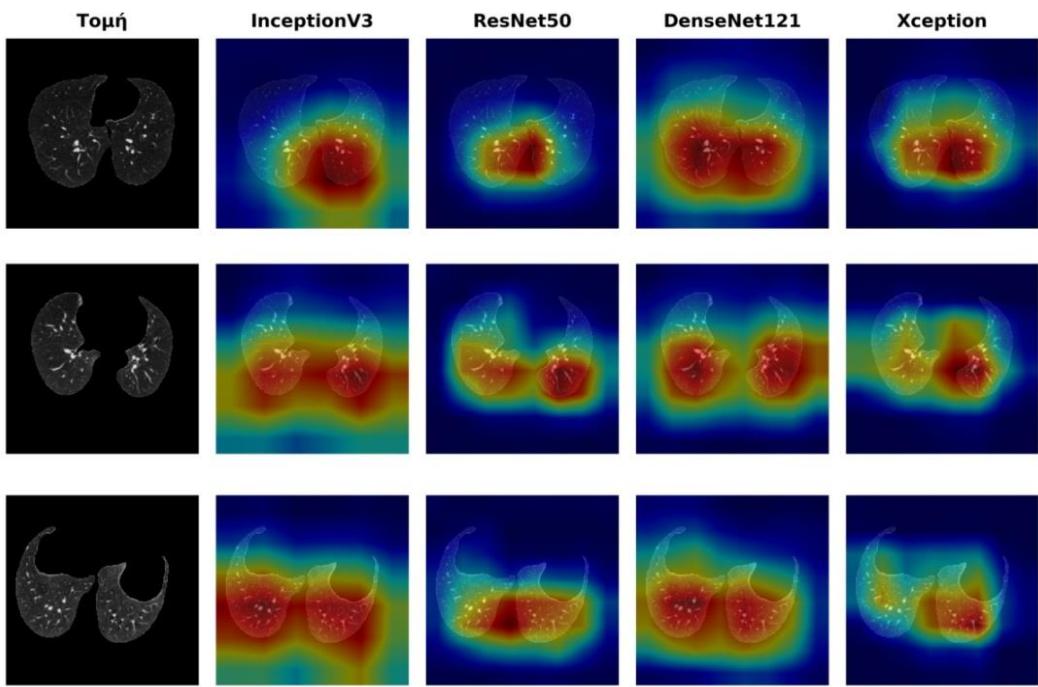
Εικόνα 3.30: Τομές συνόλου επικύρωσης χωρίς κατάτμηση πνευμονικού παρεγχύματος ορθά κατηγοριοποιημένες ως Covid-19



Εικόνα 3.31: Τομές συνόλου επικύρωσης με κατάτμηση πνευμονικού παρεγχύματος ορθά κατηγοριοποιημένες ως Covid-19



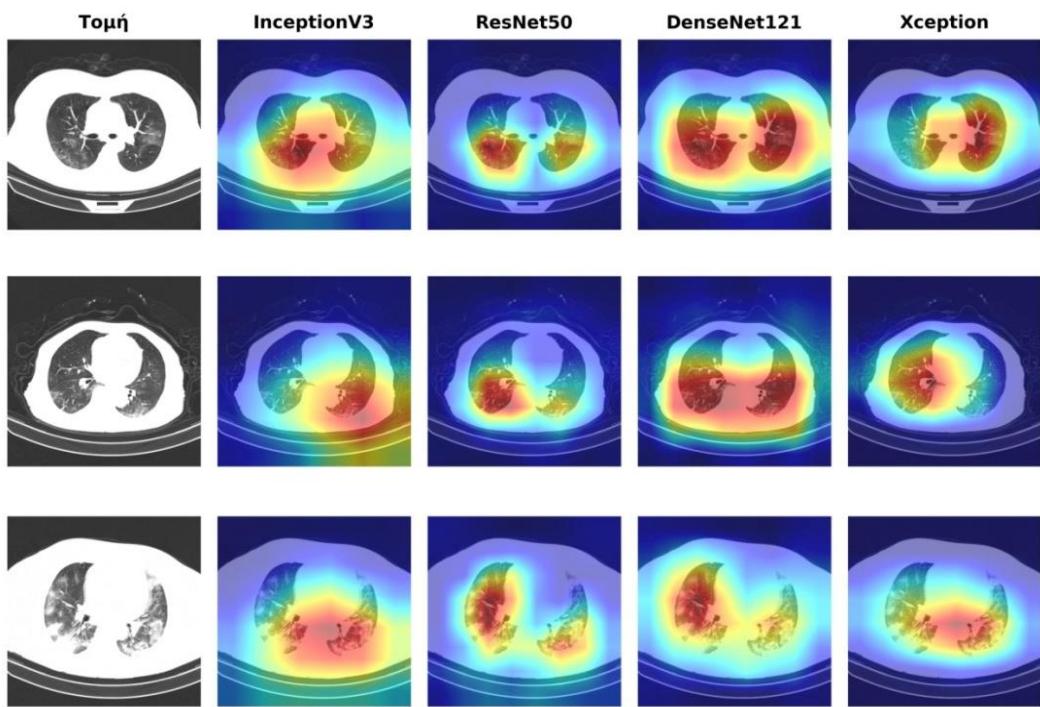
Εικόνα 3.32: Τομές συνόλου επικύρωσης χωρίς κατάτμηση πνευμονικού παρεγχύματος ορθά κατηγοριοποιημένες ως Normal



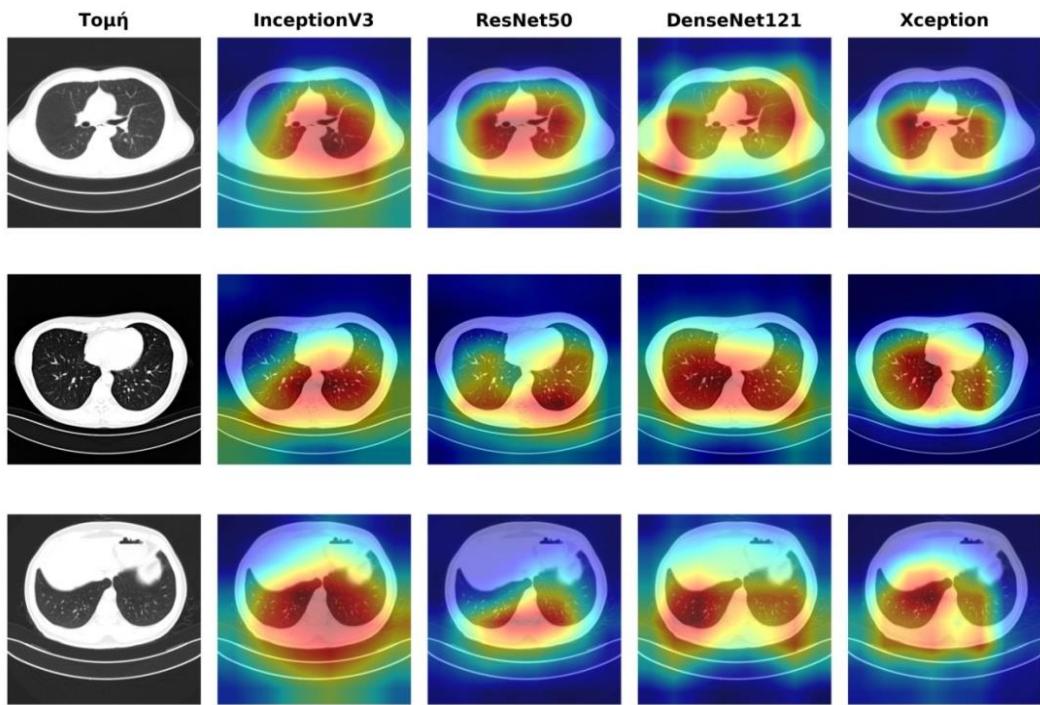
**Εικόνα 3.33:** Τομές συνόλου επικύρωσης με κατάτμηση πνευμονικού παρεγχύματος ορθά κατηγοριοποιημένες ως Normal

Στις Εικόνες 3.32 και 3.33 παρουσιάζεται η εφαρμογή της τεχνικής Grad-CAM σε τομές αξονικών τομογραφιών του συνόλου επικύρωσης οι οποίες ταξινομήθηκαν ορθά στην κατηγορία Normal από τα μοντέλα βαθιάς μάθησης. Επειδή, δεν εμφανίζονται αλλοιώσεις πνευμονίας COVID-19 στις συγκεκριμένες τομές, τα μοντέλα εστιάζουν σε μία μεγαλύτερη περιοχή της εικόνας ούτως ώστε να πραγματοποιήσουν την ταξινόμηση των τομών.

Στις Εικόνες 3.34 και 3.35 παρουσιάζεται η εφαρμογή της τεχνικής Grad-CAM σε τομές αξονικών τομογραφιών του συνόλου ελέγχου οι οποίες ταξινομήθηκαν ορθά στην κατηγορία Covid-19 και Normal αντίστοιχα, από τα μοντέλα βαθιάς μάθησης. Παρατηρώντας τις συγκεκριμένες εικόνες, διαπιστώνουμε πως όπως και στις Εικόνες 3.30 και 3.31 όλα τα μοντέλα λαμβάνουν υπόψη τους στοιχεία που βρίσκονται εντός του πνευμονικού παρεγχύματος ούτως ώστε να πραγματοποιήσουν την ταξινόμηση των τομών. Παρατηρούμε επίσης πως οι περιοχές στις οποίες δίνουν βαρύτητα τα μοντέλα για να πραγματοποιήσουν την πρόβλεψη εμφανίζουν διαφορές από μοντέλο σε μοντέλο, με τα μοντέλα ResNet50 και DenseNet121 στην Εικόνα 3.34 να εστιάζουν περισσότερο στα σημεία που εμφανίζουν αλλοιώσεις.



Εικόνα 3.34: Τομές συνόλου ελέγχου χωρίς κατάτμηση πνευμονικού παρεγχύματος ορθά κατηγοριοποιημένες ως Covid-19



Εικόνα 3.35: Τομές συνόλου ελέγχου χωρίς κατάτμηση πνευμονικού παρεγχύματος ορθά κατηγοριοποιημένες ως Normal

### **3.6.3 Σύγκριση Αποτελεσμάτων για την Περίπτωση Ταξινόμησης στις Κατηγορίες Covid-19 και Normal**

Από τα τέσσερα μοντέλα, που εκπαιδεύθηκαν για την ορθή ταξινόμηση τομών που εμφανίζουν αλλοιώσεις από COVID-19 πνευμονία και φυσιολογικών τομών, την υψηλότερη επίδοση, στην περίπτωση μη κατάτμησης πνευμονικού παρεγχύματος, εμφανίζουν τα μοντέλα ResNet50 και DenseNet121 στο σύνολο επικύρωσης και το μοντέλο DenseNet121 στο σύνολο ελέγχου. Αυτό σημαίνει πως τα μοντέλα ResNet50 και DenseNet121 πραγματοποιούν ορθότερα τη διαδικασία ταξινόμησης των τομών και πως το μοντέλο DenseNet121 γενικεύει καλύτερα σε σχέση με τα άλλα μοντέλα.

Στην περίπτωση όπου πραγματοποιείται κατάτμηση πνευμονικού παρεγχύματος, τα μοντέλα InceptionV3, ResNet50 και DenseNet121 είναι αυτά που εμφανίζουν την υψηλότερη επίδοση στο σύνολο επικύρωσης. Αυτό σημαίνει πως τα συγκεκριμένα μοντέλα παρουσιάζουν μεγαλύτερη ικανότητα ορθής ταξινόμησης των τομών αξονικών τομογραφιών και γενικεύουν καλύτερα συγκριτικά με το μοντέλο Xception. Επίσης, δείχνει πως στην περίπτωση κατάτμησης του πνευμονικού παρεγχύματος τα μοντέλα που επιτυγχάνουν καλύτερη επίδοση και μεγαλύτερη ικανότητα γενίκευσης είναι περισσότερα.

Παρατηρώντας τους πίνακες σύγχυσης των μοντέλων, μπορεί να διαπιστωθεί πως τα μοντέλα δυσκολεύονται περισσότερο στο να πραγματοποιήσουν ορθή ταξινόμηση των τομών αξονικών τομογραφιών που ανήκουν στην κατηγορία Covid-19. Αυτό οφείλεται στο γεγονός ότι ορισμένοι από τους ασθενείς εμφανίζουν ήπιας μορφής νόσηση από COVID-19 και επομένως οι αλλοιώσεις στις τομές των αξονικών τομογραφιών τους είναι λιγότερες σε αριθμό ή και λιγότερο έντονες. Λόγω αυτού, πιθανώς τα μοντέλα δυσκολεύονται να ταξινομήσουν ορθά τις τομές που προέρχονται από αυτούς τους ασθενείς. Ενδιαφέρον παρουσιάζει το γεγονός ότι στις περιπτώσεις εφαρμογής ταξινόμησης στα δεδομένα του συνόλου ελέγχου, το μοντέλο Xception παρουσιάζει μεγαλύτερη δυσκολία στο να ταξινομήσει ορθά τομές που ανήκουν στην κατηγορία Normal.

Συγκρίνοντας την επίδοση των μοντέλων στα δεδομένα επικύρωσης και στα δεδομένα ελέγχου, διαπιστώνεται πως οι τιμές των μετρικών αξιολόγησης στα πρώτα είναι εμφανώς μεγαλύτερες από τις αντίστοιχες στα δεύτερα. Αυτό είναι αναμενόμενο εάν σκεφτούμε ότι τα δεδομένα εκπαίδευσης και επικύρωσης προέρχονται από το ίδιο σύνολο, ενώ τα δεδομένα ελέγχου προέρχονται από ένα διαφορετικό σύνολο. Επομένως, είναι λογικό τα μοντέλα να μην έχουν τόσο καλή επίδοση σε ένα σύνολο δεδομένων το οποίο τους είναι εντελώς «άγνωστο». Επίσης, το σύνολο επικύρωσης χρησιμοποιείται κατά την εκπαίδευτική διαδικασία για τη ρύθμιση των παραμέτρων των δικτύων (tuning) πράγμα που ενδεχομένως να οδηγεί τα μοντέλα στο να «μαθαίνουν» χαρακτηριστικά του συνόλου επικύρωσης, παρότι δεν έχουν εκπαίδευθεί σε αυτό, και άρα να αποδίδουν καλύτερα.

Όσον αφορά τις καμπύλες εκπαίδευσης και επικύρωσης, παρατηρείται πως η απόσταση ανάμεσα στην καμπύλη εκπαίδευσης και επικύρωσης είναι μικρότερη στις περιπτώσεις όπου πραγματοποιείται διαχωρισμός ανά τομή. Η εξήγηση είναι ότι πιθανόν όταν πραγματοποιείται διαχωρισμός ανά τομή εμφανίζεται το φαινόμενο της διαρροής δεδομένων, αφού συνεχόμενες τομές του ίδιου ασθενή με ελάχιστες διαφορές μεταξύ τους παρουσιάζονται τόσο στο σύνολο εκπαίδευσης όσο και στο σύνολο επικύρωσης. Επίσης, στην περίπτωση διαχωρισμού ανά τομή, τα δεδομένα είναι περισσότερο ποικίλα με αποτέλεσμα τα μοντέλα να γενικεύουν καλύτερα όταν εκπαιδεύονται σε δεδομένα που διαχωρίζονται με αυτόν τον τρόπο.

Για τους ίδιους λόγους παρατηρούνται καλύτερες τιμές μετρικών αξιολόγησης στην περίπτωση που πραγματοποιείται διαχωρισμός ανά τομή σε σχέση με την περίπτωση όπου πραγματοποιείται διαχωρισμός ανά ασθενή.

Είναι σημαντικό να αναφερθεί, πως όταν πραγματοποιείται διαχωρισμός ανά ασθενή, παρατηρείται ότι τα μοντέλα στην περίπτωση κατάτμησης πνευμονικού παρεγχύματος παρουσιάζουν πολύ καλύτερη επίδοση συγκριτικά με τα αντίστοιχα μοντέλα στην περίπτωση μη κατάτμησης πνευμονικού παρεγχύματος.

Σε αντίθεση, όταν τα μοντέλα εκπαιδεύονται και αξιολογούνται σε δεδομένα με διαχωρισμό ανά τομή παρουσιάζουν καλύτερη επίδοση στη περίπτωση μη κατάτμησης του πνευμονικού παρεγχύματος. Σε αυτή την περίπτωση, τα μοντέλα ενδεχομένως βασίζονται και σε στοιχεία που βρίσκονται εκτός του πνευμονικού παρεγχύματος ούτως ώστε να πραγματοποιήσουν την ταξινόμηση. Η ύπαρξη περισσότερης πληροφορίας στην κάθε εικόνα πιθανόν να παίζει ρόλο στην αυξημένη επίδοση αυτών των μοντέλων.

Συγκρίνοντας τα αποτελέσματα της εφαρμογής της τεχνικής Grad-CAM σε ταξινομημένες τομές των συνόλων δεδομένων, διαπιστώνουμε πως σε όλες τις περιπτώσεις τα μοντέλα πραγματοποιούν την ταξινόμηση βασιζόμενα σε στοιχεία που βρίσκονται εντός του πνευμονικού παρεγχύματος.

### **3.6.4 Δυαδική Ταξινόμηση των Τομών Αξονικών Τομογραφιών στις Κατηγορίες Covid-19 και Other Pneumonia**

#### **3.6.4.1 Παραλλαγή 1: Μη Κατάτμηση Πνευμονικού Παρεγχύματος – Διαχωρισμός Δεδομένων ανά Τομή**

Στους Πίνακες 3.11 και 3.12 παρουσιάζονται αντίστοιχα οι μετρικές αξιολόγησης της ικανότητας των μοντέλων να πραγματοποιούν, έπειτα από εκπαίδευση, ταξινόμηση των τομών των συνόλων δεδομένων επικύρωσης και ελέγχου της Παραλλαγής 1 στις κατηγορίες Covid-19 και Other Pneumonia.

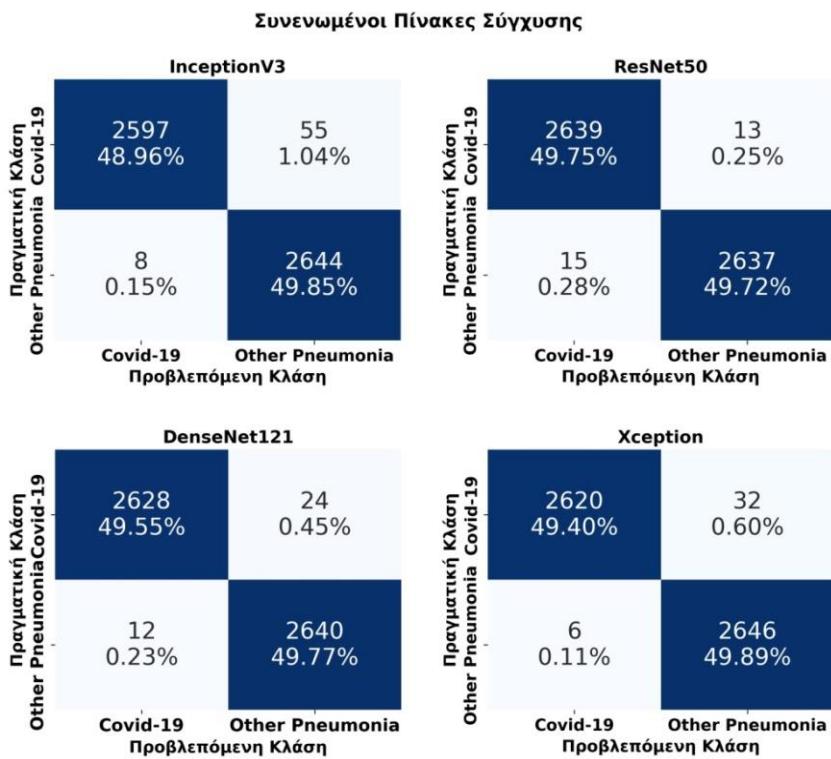
**Πίνακας 3.11: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Τομή: Μετρικές αξιολόγησης της ταξινόμησης των δεδομένων επικύρωσης στις κατηγορίες Covid-19 και Other Pneumonia**

Μοντέλο	Μέση Ακρίβεια	Μέση Απώλεια	Μέση Αξιοπιστία	Μέση Ανάκληση	Μέση F1 Score
<b>InceptionV3</b>	0,9881 ± 0,0030	0,0347 ± 0,0078	0,9969 ± 0,0011	0,9793 ± 0,0058	0,9880 ± 0,0030
<b>ResNet50</b>	0,9947 ± 0,0037	0,0175 ± 0,0073	0,9944 ± 0,0057	0,9951 ± 0,0032	0,9947 ± 0,0037
<b>DenseNet121</b>	0,9932 ± 0,0032	0,0194 ± 0,0070	0,9955 ± 0,0037	0,9909 ± 0,0043	0,9932 ± 0,0032
<b>Xception</b>	0,9928 ± 0,0034	0,0256 ± 0,0092	0,9977 ± 0,0021	0,9879 ± 0,0051	0,9928 ± 0,0034

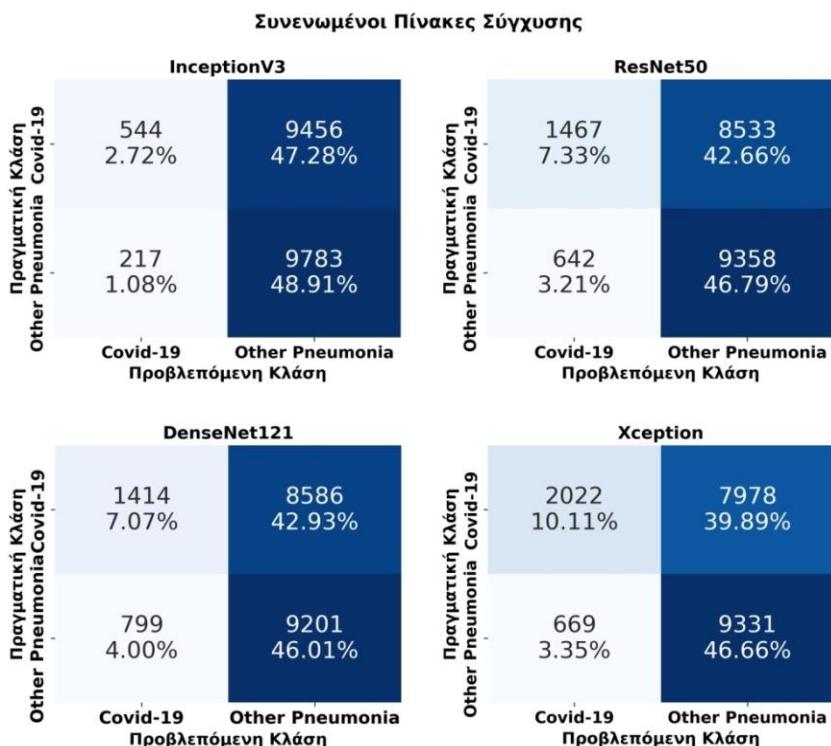
**Πίνακας 3.12: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Τομή: Μετρικές αξιολόγησης της ταξινόμησης των δεδομένων ελέγχου στις κατηγορίες Covid-19 και Other Pneumonia**

Μοντέλο	Μέση Ακρίβεια	Μέση Απώλεια	Μέση Αξιοπιστία	Μέση Ανάκληση	Μέση F1 Score
<b>InceptionV3</b>	0,5163 ± 0,0103	3,3259 ± 0,7260	0,6912 ± 0,1021	0,0544 ± 0,0255	0,1001 ± 0,0450
<b>ResNet50</b>	0,5413 ± 0,0334	2,8979 ± 1,1479	0,6992 ± 0,0976	0,1467 ± 0,0742	0,2359 ± 0,1050
<b>DenseNet121</b>	0,5308 ± 0,0020	2,4341 ± 0,4168	0,6283 ± 0,0605	0,1414 ± 0,0589	0,2272 ± 0,0824
<b>Xception</b>	0,5677 ± 0,0253	1,5477 ± 0,2581	0,7479 ± 0,0541	0,2022 ± 0,0534	0,3160 ± 0,0706

Στις Εικόνες 3.36 και 3.37 παρουσιάζονται αντίστοιχα οι πίνακες σύγχυσης των μοντέλων, έπειτα από την ταξινόμηση των τομών των συνόλων δεδομένων επικύρωσης και ελέγχου της Παραλλαγής 1 στις κατηγορίες Covid-19 και Other Pneumonia.



**Εικόνα 3.36: Μη Κατάτηση Πνευμονικού Παρεγγύματος - Διαχωρισμός Δεδομένων ανά Τομή:**  
Συνενωμένοι πίνακες σύγχυσης διαδικασίας ταξινόμησης των δεδομένων επικύρωσης στις κατηγορίες Covid-19 και Other Pneumonia



**Εικόνα 3.37: Μη Κατάτηση Πνευμονικού Παρεγγύματος - Διαχωρισμός Δεδομένων ανά Τομή:**  
Συνενωμένοι πίνακες σύγχυσης διαδικασίας ταξινόμησης των δεδομένων ελέγχου στις κατηγορίες Covid-19 και Other Pneumonia

Παρατηρώντας τις μετρικές αξιολόγησης των μοντέλων στα δεδομένα επικύρωσης στους Πίνακες 3.11 και 3.12 καθώς και τους αντίστοιχους πίνακες σύγχυσης στις Εικόνες 3.36 και 3.37, διαπιστώνουμε πως ενώ τα μοντέλα επιτυγχάνουν υψηλή επίδοση στο σύνολο επικύρωσης, εμφανίζουν πολύ χαμηλή επίδοση όταν πραγματοποιούν ταξινόμηση των δεδομένων του συνόλου ελέγχου. Αυτό σημαίνει ότι τα μοντέλα δυσκολεύονται να βρουν στοιχεία που διαχωρίζουν τις τομές που εμφανίζουν πνευμονία COVID-19 από τομές που εμφανίζουν πνευμονία άλλης αιτιολογίας και επομένως βασίζονται σε λανθασμένα στοιχεία ούτως ώστε να πραγματοποιήσουν την ταξινόμηση. Με λίγα λόγια τα μοντέλα εμφανίζουν υπερπροσαρμογή στα δεδομένα εκπαίδευσης.

### **3.6.4.2 Παραλλαγή 2: Μη Κατάτμηση Πνευμονικού Παρεγχύματος – Διαχωρισμός Δεδομένων ανά Τομή**

Στους Πίνακες 3.13 και 3.14 παρουσιάζονται αντίστοιχα οι μετρικές αξιολόγησης των μοντέλων ως προς την ικανότητά τους να πραγματοποιούν, έπειτα από εκπαίδευση, ταξινόμηση των τομών των συνόλων δεδομένων επικύρωσης και ελέγχου της Παραλλαγής 2 στις κατηγορίες Covid-19 και Other Pneumonia.

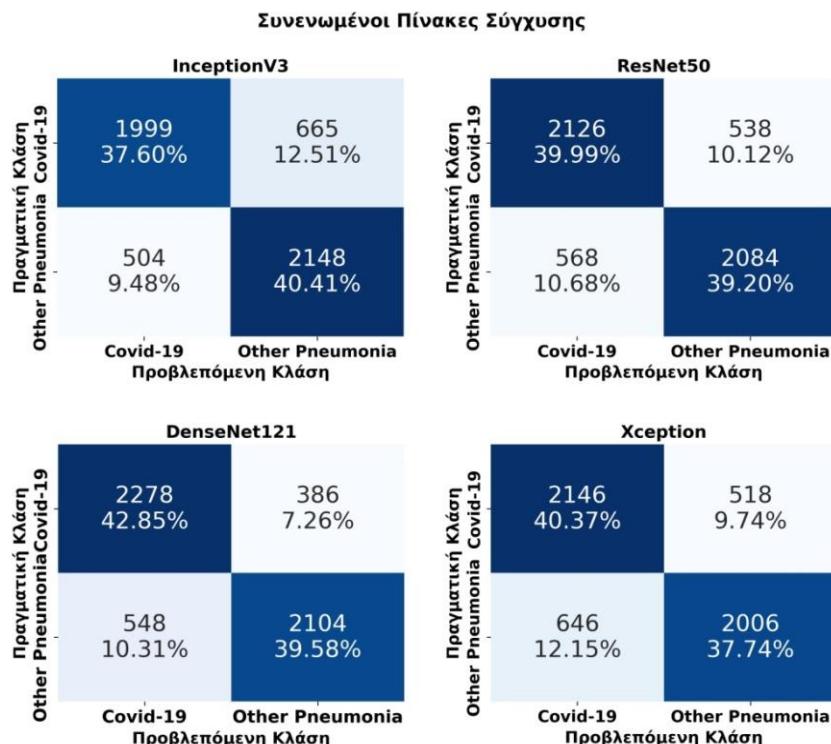
**Πίνακας 3.13: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Ασθενή:**  
Μετρικές αξιολόγησης της ταξινόμησης των δεδομένων επικύρωσης στις κατηγορίες Covid-19 και Other Pneumonia

Μοντέλο	Μέση Ακρίβεια	Μέση Απώλεια	Μέση Αξιοπιστία	Μέση Ανάκληση	Μέση F1 Score
<b>InceptionV3</b>	0,7844 ± 0,0502	0,4859 ± 0,0890	0,7976 ± 0,0533	0,7553 ± 0,1047	0,7741 ± 0,0746
<b>ResNet50</b>	0,7956 ± 0,0563	0,4424 ± 0,0646	0,7918 ± 0,0351	0,7960 ± 0,1828	0,7845 ± 0,1052
<b>DenseNet121</b>	0,8311 ± 0,0652	0,4163 ± 0,1187	0,8092 ± 0,0627	0,8584 ± 0,1245	0,8305 ± 0,0837
<b>Xception</b>	0,7867 ± 0,0598	0,4675 ± 0,0877	0,7688 ± 0,0467	0,8109 ± 0,1300	0,7868 ± 0,0804

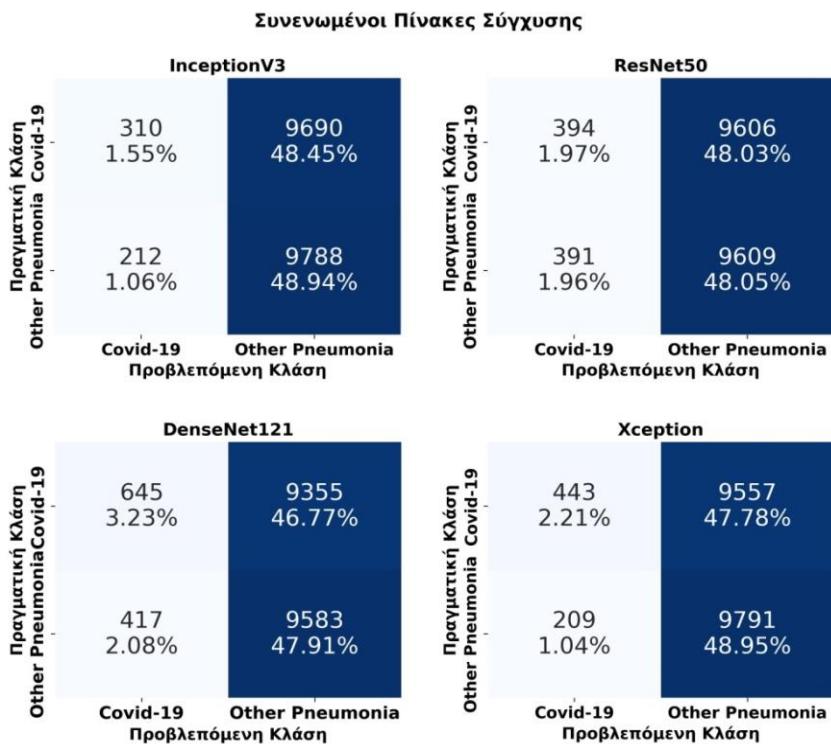
**Πίνακας 3.14: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Ασθενή:**  
Μετρικές αξιολόγησης της ταξινόμησης των δεδομένων ελέγχου στις κατηγορίες Covid-19 και Other Pneumonia

Μοντέλο	Μέση Ακρίβεια	Μέση Απώλεια	Μέση Αξιοπιστία	Μέση Ανάκληση	Μέση F1 Score
<b>InceptionV3</b>	0,5049 ± 0,0055	1,4442 ± 0,3535	0,5895 ± 0,0705	0,0310 ± 0,0131	0,0586 ± 0,0236
<b>ResNet50</b>	0,5002 ± 0,006	2,0504 ± 0,4423	0,4431 ± 0,1362	0,0394 ± 0,0352	0,0705 ± 0,0591
<b>DenseNet121</b>	0,5114 ± 0,0149	1,3952 ± 0,4283	0,5476 ± 0,1644	0,0645 ± 0,0482	0,1123 ± 0,0813
<b>Xception</b>	0,5117 ± 0,0090	1,0817 ± 0,1305	0,5476 ± 0,1644	0,0443 ± 0,0257	0,0820 ± 0,0447

Στις Εικόνες 3.38 και 3.39 παρουσιάζονται αντίστοιχα οι πίνακες σύγχυσης των μοντέλων έπειτα από την ταξινόμηση των τομών των συνόλων δεδομένων επικύρωσης και ελέγχου της Παραλλαγής 2 στις κατηγορίες Covid-19 και Other Pneumonia.



**Εικόνα 3.38: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Ασθενή:**  
Συνενωμένοι πίνακες σύγχυσης διαδικασίας ταξινόμησης των δεδομένων επικύρωσης στις κατηγορίες Covid-19 και Other Pneumonia



**Εικόνα 3.39: Μη Κατάτμηση Πνευμονικού Παρεγγύματος - Διαχωρισμός Δεδομένων ανά Ασθενή:**  
Συνενωμένοι πίνακες σύγχυσης διαδικασίας ταξινόμησης των δεδομένων ελέγχου στις κατηγορίες Covid-19 και Other Pneumonia

Παρατηρώντας τις μετρικές αξιολόγησης των μοντέλων στα δεδομένα επικύρωσης στους Πίνακες 3.13 και 3.14 καθώς και τους αντίστοιχους πίνακες σύγχυσης στις Εικόνες 3.38 και 3.39, διαπιστώνουμε πως ενώ τα μοντέλα επιτυγχάνουν πολύ καλή επίδοση στο σύνολο επικύρωσης, εμφανίζουν πολύ χαμηλή επίδοση όταν πραγματοποιούν ταξινόμηση των δεδομένων του συνόλου ελέγχου. Αυτό σημαίνει ότι τα μοντέλα δυσκολεύονται να βρουν στοιχεία που διαχωρίζουν τις τομές που εμφανίζουν πνευμονία COVID-19 από τομές που εμφανίζουν πνευμονία άλλης αιτιολογίας και επομένως βασίζονται σε λανθασμένα στοιχεία ούτως ώστε να πραγματοποιήσουν την ταξινόμηση. Με λίγα λόγια τα μοντέλα εμφανίζουν υπερπροσαρμογή στα δεδομένα εκπαίδευσης.

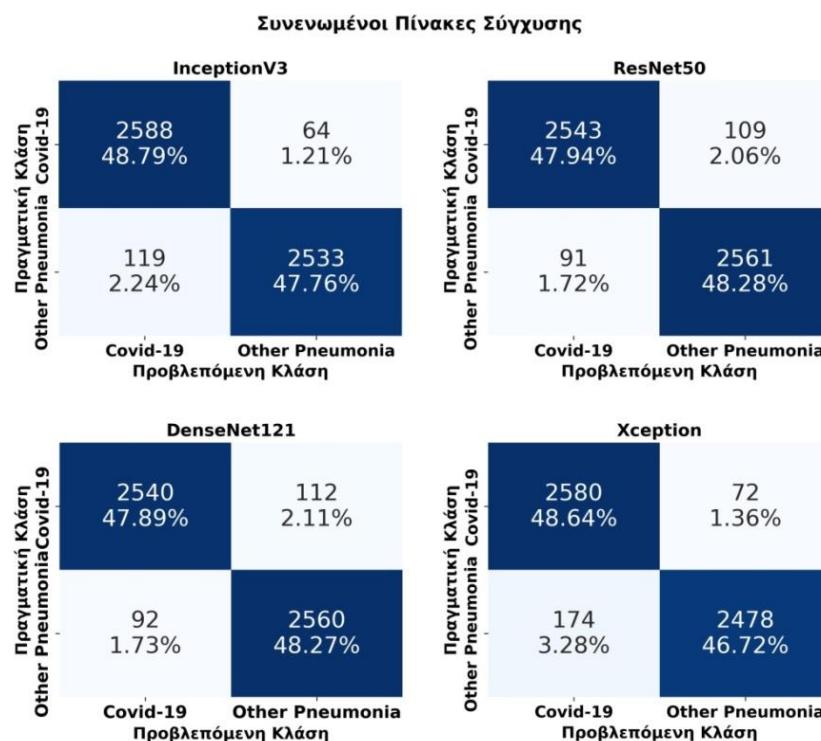
### 3.6.4.3 Παραλλαγή 3: Κατάτμηση Πνευμονικού Παρεγγύματος – Διαχωρισμός Δεδομένων ανά Τομή

Στον Πίνακα 3.15 παρουσιάζονται οι μετρικές αξιολόγησης των μοντέλων ως προς την ικανότητά τους να πραγματοποιούν, έπειτα από εκπαίδευση, ταξινόμηση των τομών του συνόλου δεδομένων επικύρωσης της Παραλλαγής 3 στις κατηγορίες Covid-19 και Other Pneumonia. Στην Εικόνα 3.40 παρουσιάζονται οι πίνακες σύγχυσης των μοντέλων έπειτα από την ταξινόμηση των τομών του συνόλου

δεδομένων επικύρωσης της Παραλλαγής 3 στις κατηγορίες Covid-19 και Other Pneumonia.

**Πίνακας 3.15: Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Τομή: Μετρικές αξιολόγησης της ταξινόμησης των δεδομένων επικύρωσης στις Κατηγορίες Covid-19 και Other Pneumonia**

Μοντέλο	Μέση Ακρίβεια	Μέση Απώλεια	Μέση Αξιοπιστία	Μέση Ανάκληση	Μέση F1 Score
InceptionV3	0,9655 ± 0,0063	0,0960 ± 0,0131	0,9561 ± 0,0111	0,9759 ± 0,0063	0,9659 ± 0,0061
ResNet50	0,9623 ± 0,0065	0,1005 ± 0,0174	0,9658 ± 0,0162	0,9589 ± 0,0155	0,9614 ± 0,0075
DenseNet121	0,9615 ± 0,0072	0,1061 ± 0,0174	0,9652 ± 0,0112	0,9578 ± 0,0167	0,9614 ± 0,0075
Xception	0,9536 ± 0,0080	0,1236 ± 0,0206	0,9373 ± 0,0206	0,9729 ± 0,0103	0,9546 ± 0,0073



**Εικόνα 3.40: Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Τομή: Συνενωμένοι πίνακες σύγχυσης διαδικασίας ταξινόμησης των δεδομένων επικύρωσης στις κατηγορίες Covid-19 και Other Pneumonia**

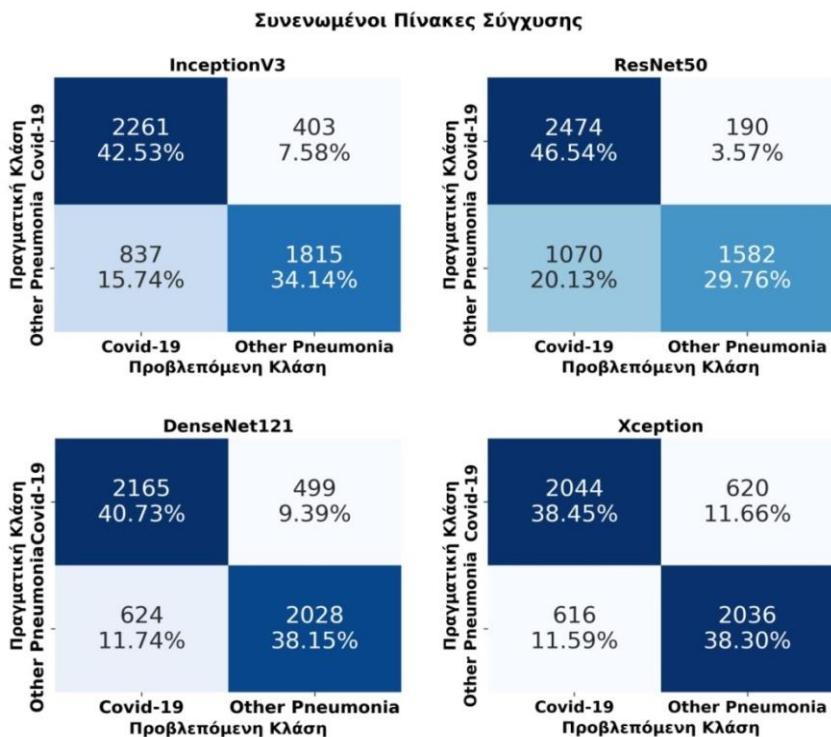
Παρατηρώντας τις μετρικές αξιολόγησης των μοντέλων στα δεδομένα επικύρωσης στον Πίνακα 3.15 καθώς και τους αντίστοιχους πίνακες σύγχυσης στην Εικόνα 3.40, διαπιστώνουμε πως τα μοντέλα επιτυγχάνουν υψηλή επίδοση στην περίπτωση διαχωρισμού ανά τομή. Επειδή δεν μπορούμε να πραγματοποιήσουμε κατάτμηση του πνευμονικού παρεγχύματος στα δεδομένα ελέγχου, δεν πραγματοποιήσαμε αξιολόγηση των μοντέλων στα συγκεκριμένα δεδομένα. Ωστόσο, πιστεύουμε ότι τα αποτελέσματα θα ήταν όμοια με των προηγούμενων περιπτώσεων.

### **3.6.4.4 Παραλλαγή 4: Κατάτμηση Πνευμονικού Παρεγχύματος – Διαχωρισμός Δεδομένων ανά Ασθενή**

Στον Πίνακα 3.16 παρουσιάζονται οι μετρικές αξιολόγησης των μοντέλων ως προς την ικανότητά τους να πραγματοποιούν, έπειτα από εκπαίδευση, ταξινόμηση των τομών του συνόλου δεδομένων επικύρωσης της Παραλλαγής 3 στις κατηγορίες Covid-19 και Other Pneumonia. Στην Εικόνα 3.41 παρουσιάζονται οι πίνακες σύγχυσης των μοντέλων έπειτα από την ταξινόμηση των τομών του συνόλου δεδομένων επικύρωσης της Παραλλαγής 3 στις κατηγορίες Covid-19 και Other Pneumonia.

**Πίνακας 3.16: Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Ασθενή: Μετρικές αξιολόγησης της ταξινόμησης των δεδομένων επικύρωσης στις κατηγορίες Covid-19 και Other Pneumonia**

Μοντέλο	Μέση Ακρίβεια	Μέση Απώλεια	Μέση Αξιοπιστία	Μέση Ανάκληση	Μέση F1 Score
<b>InceptionV3</b>	0,7662 ± 0,0403	0,5183 ± 0,0800	0,7278 ± 0,0499	0,8528 ± 0,1015	0,7824 ± 0,0564
<b>ResNet50</b>	0,7507 ± 0,1301	0,4940 ± 0,1298	0,7077 ± 0,1129	0,9321 ± 0,0531	0,7973 ± 0,0672
<b>DenseNet121</b>	0,7882 ± 0,0238	0,4678 ± 0,0662	0,7777 ± 0,0496	0,8190 ± 0,0942	0,7932 ± 0,0376
<b>Xception</b>	0,7653 ± 0,0287	0,5106 ± 0,0364	0,7706 ± 0,0490	0,7727 ± 0,0838	0,7673 ± 0,0232



**Εικόνα 3.41: Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Ασθενή: Συνενωμένοι πίνακες σύγχυσης διαδικασίας ταξινόμησης των δεδομένων επικύρωσης στις κατηγορίες Covid-19 και Other Pneumonia**

Παρατηρώντας τις μετρικές αξιολόγησης των μοντέλων στα δεδομένα επικύρωσης στον Πίνακα 3.16 καθώς και τον αντίστοιχο πίνακα σύγχυσης στην Εικόνα 3.41, διαπιστώνουμε πως τα μοντέλα επιτυγχάνουν πολύ καλή επίδοση στην περίπτωση διαχωρισμού ανά ασθενή. Επειδή δεν μπορούμε να πραγματοποιήσουμε κατάτμηση του πνευμονικού παρεγχύματος στα δεδομένα ελέγχου, δεν πραγματοποιήσαμε αξιολόγηση των μοντέλων στα συγκεκριμένα δεδομένα. Ωστόσο, πιστεύουμε ότι τα αποτελέσματα θα ήταν όμοια με των προηγούμενων περιπτώσεων.

### 3.6.5 Σύγκριση Αποτελεσμάτων για την Περίπτωση Ταξινόμησης στις Κατηγορίες Covid-19 και Other Pneumonia

Όσον αφορά τη δυαδική ταξινόμηση των τομών που εμφανίζουν COVID-19 πνευμονία και των τομών που εμφανίζουν πνευμονία άλλης αιτιολογίας, διαπιστώνουμε πως όταν τα μοντέλα αξιολογούνται στο σύνολο δεδομένων επικύρωσης επιτυγχάνουν υψηλά αποτελέσματα τόσο στην περίπτωση διαχωρισμού ανά τομή όσο και στην περίπτωση διαχωρισμού ανά ασθενή, ενώ όταν αξιολογούνται στο σύνολο ελέγχου επιτυγχάνουν πολύ χαμηλά αποτελέσματα.

Η υψηλή επίδοση στο σύνολο επικύρωσης ενδεχομένως να οφείλεται στο γεγονός ότι τα σύνολα εκπαίδευσης και επικύρωσης προέρχονται από το ίδιο αρχικό σύνολο και έχουν πολλές ομοιότητες μεταξύ τους. Αυτό έχει σαν αποτέλεσμα τα χαρακτηριστικά των εικόνων στο δεύτερο σύνολο να είναι «γνωστά» για τα μοντέλα που έχουν εκπαιδευθεί στο πρώτο σύνολο. Επίσης, το σύνολο επικύρωσης χρησιμοποιείται κατά την εκπαίδευτική διαδικασία για τη ρύθμιση των παραμέτρων των δικτύων (tuning) πράγμα που ενδεχομένως να οδηγεί τα μοντέλα στο να «μαθαίνουν» χαρακτηριστικά του παρότι δεν έχουν εκπαιδευθεί στο συγκεκριμένο σύνολο.

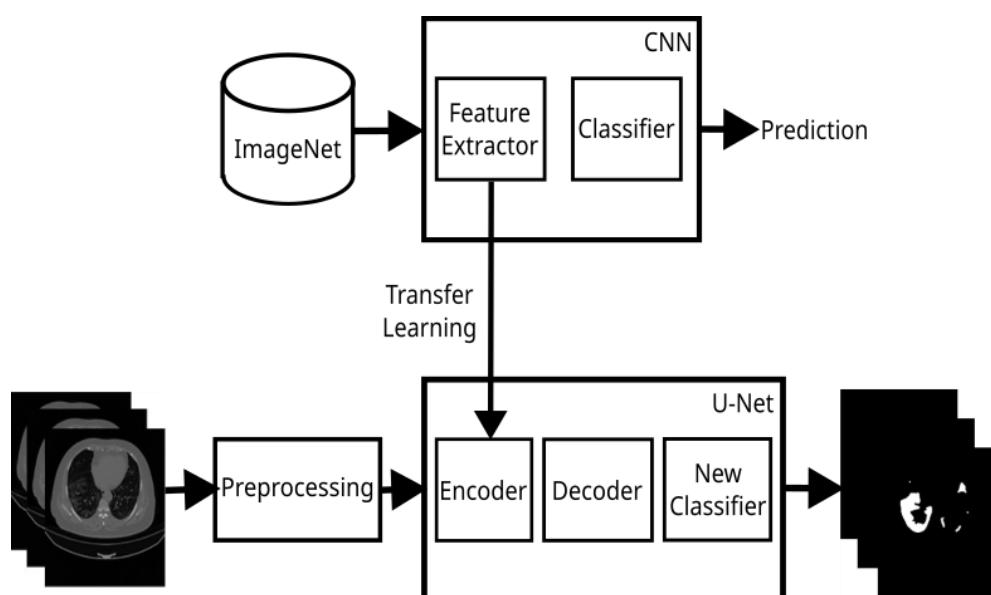
Αντίθετα, η χαμηλή επίδοση στην αξιολόγηση των δεδομένων στο σύνολο ελέγχου πιθανόν να οφείλεται στον περιορισμένο αριθμό δεδομένων του συνόλου εκπαίδευσης και στο γεγονός ότι η ταξινόμηση τομών αξονικών τομογραφιών με COVID-19 πνευμονία και πνευμονία άλλης αιτιολογίας είναι αρκετά απαιτητική.

## 4 Περιγραφή Μεθοδολογίας Πειράματος Σημασιολογικής Κατάτμησης Εικόνων (Semantic Image Segmentation)

Το δεύτερο εκ των δύο πειραμάτων έχει ως στόχο την εξέταση και αξιολόγηση αρχιτεκτονικών τεχνητών νευρωνικών δικτύων ως προς την ικανότητα τους να πραγματοποιούν σημασιολογική κατάτμηση αλλοιώσεων από πνευμονία COVID-19 σε τομές αξονικών τομογραφιών. Στόχος είναι, τα κατάλληλα εκπαιδευμένα μοντέλα, να μπορούν να διακρίνουν και να απομονώνουν τις περιοχές που εμφανίζουν αλλοιώσεις οφειλόμενες σε πνευμονία COVID-19.

Για τη διεξαγωγή του πειράματος του παρόντος κεφαλαίου ακολουθείται η ίδια διαδικασία με το πείραμα του Κεφαλαίου 3, η οποία παρουσιάζεται στην Εικόνα 4.1. Πιο συγκεκριμένα:

- 5) Επιλογή κατάλληλων συνόλων δεδομένων, σύμφωνα με τις ανάγκες του πειράματος σημασιολογικής κατάτμησης των τομών αξονικών τομογραφιών.
- 6) Επεξεργασία των δεδομένων, πριν αυτά τροφοδοτηθούν στο εκάστοτε μοντέλο τεχνητών νευρωνικών δικτύων
- 7) Επιλογή κατάλληλων μοντέλων, εκπαίδευση και επικύρωση τους χρησιμοποιώντας σύνολα δεδομένων και ακολουθώντας τη μέθοδο της μεταφοράς μάθησης
- 8) Αξιολόγηση, έπειτα από το πέρας της εκπαίδευτικής διαδικασίας, της ικανότητας των μοντέλων να ταξινομούν τομές αξονικών τομογραφιών. (Η αξιολόγηση των μοντέλων πραγματοποιείται με τη βοήθεια κατάλληλων μετρικών αξιολόγησης).



Εικόνα 4.1: Διαδικασία εκπαίδευσης και αξιολόγησης μοντέλου ως προς τη σημασιολογική κατάτμηση αλλοιώσεων από πνευμονία COVID-19 σε τομές, όπου αξιοποιείται η τεχνική της μεταφοράς μάθησης

## **4.1 Επιλογή Συνόλου Δεδομένων Πειράματος Σημασιολογικής Κατάτμησης Εικόνων (Semantic Image Segmentation)**

Όσον αφορά το πρόβλημα σημασιολογικής κατάτμησης αλλοιώσεων σε τομές αξονικών τομογραφιών που αναλύεται στο συγκεκριμένο κεφάλαιο, τα σύνολα δεδομένων που είναι διαθέσιμα στο διαδίκτυο αποτελούνται από αξονικές τομογραφίες.

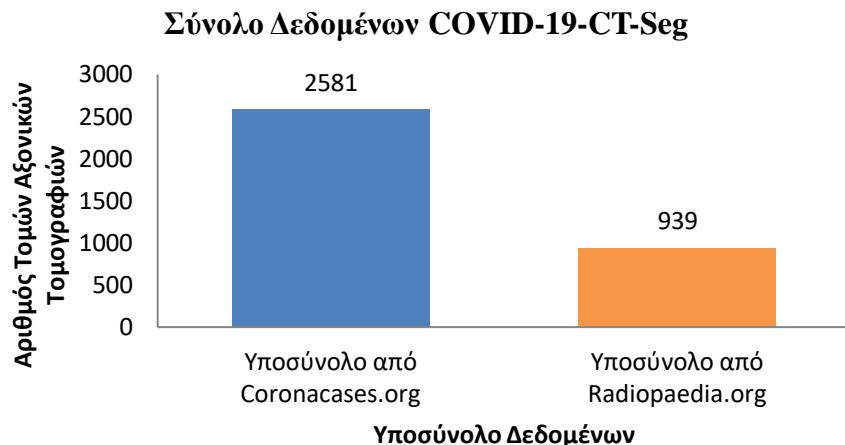
Για τη διεξαγωγή του πειράματος του παρόντος κεφαλαίου επιλέχθηκε να αξιοποιηθεί τα σύνολο δεδομένων COVID-19-CT-Seg [83].

Πιο συγκεκριμένα, το σύνολο δεδομένων COVID-19-CT-Seg αξιοποιήθηκε στα πλαίσια του πειράματος τόσο στη φάση της εκπαίδευσης (training phase) όσο και στη φάση της επικύρωσης (validation phase), καθώς και στη φάση εξωτερικής αξιολόγησης των μοντέλων. Πιο αναλυτικά, το υποσύνολο του συνόλου που συλλέχθηκε από το αποθετήριο Coronacases.org [84] αξιοποιήθηκε για την διαδικασία της εκπαίδευσης και επικύρωσης των μοντέλων βαθειάς μάθησης, ενώ το υποσύνολο του συνόλου που συλλέχθηκε από το αποθετήριο Radiopaedia.org [85] χρησιμοποιήθηκε ως εξωτερικό σύνολο αξιολόγησης για τη διερεύνηση της ικανότητας γενίκευσης των μοντέλων τεχνητών νευρωνικών δικτύων που αξιοποιήθηκαν στο παρόν πείραμα.

### **4.1.1 Σύνολο Δεδομένων COVID-19-CT-Seg**

Το σύνολο δεδομένων που αξιοποιήθηκε για τις ανάγκες του παρόντος πειράματος είναι το σύνολο δεδομένων COVID-19-CT-Seg.

Το συγκεκριμένο σύνολο δεδομένων αποτελείται από 3.520 τομές αξονικών τομογραφιών που έχουν συλλεχθεί από τα αποθετήρια Coronacases.org και Radiopaedia.org και οι οποίες ανήκουν σε 20 διαφορετικούς ασθενείς [83]. Πιο αναλυτικά, οι 2.581 τομές αξονικών τομογραφιών ανήκουν σε 10 από τους 20 ασθενείς και προέρχονται από το αποθετήριο Coronacases.org, ενώ οι υπόλοιπες 939 τομές αξονικών τομογραφιών ανήκουν στους υπόλοιπους 10 ασθενείς και προέρχονται από το αποθετήριο Radiopaedia.org. Στην Εικόνα 4.2 παρουσιάζεται ένα ραβδόγραμμα της κατανομής των ασθενών στις κλάσεις του συνόλου COVID-19-CT-Seg.

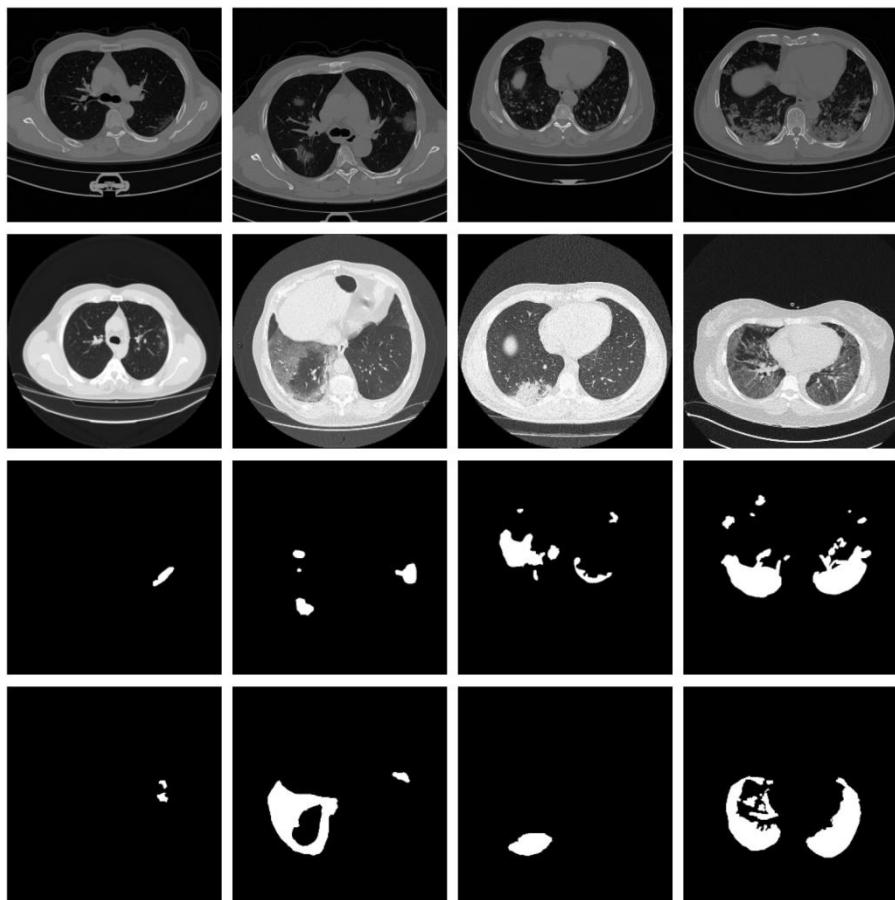


**Εικόνα 4.2: Ραβδόγραμμα υποσυνόλων συνόλου δεδομένων COVID-19-CT-Seg**

Όλες οι τομές αξονικών τομογραφιών εμφανίζουν αλλοιώσεις οφειλόμενες σε πνευμονία από το SARS-CoV-2 [83]. Οι δημιουργοί του συγκεκριμένου συνόλου δεδομένων έχουν συμπεριλάβει 3.520 τομές αξονικών τομογραφιών, καθώς επίσης και 3.520 «μάσκες» οι οποίες αποτελούν εικόνες στις οποίες έχει σκιαγραφηθεί το πνευμονικό παρέγχυμα αμφοτερόπλευρα. Επιπλέον, έχουν συμπεριλάβει άλλες 3.520 «μάσκες» που αποτελούν εικόνες στις οποίες έχουν σκιαγραφηθεί οι περιοχές όπου εμφανίζονται αλλοιώσεις από τον το SARS-CoV-2. Οι «μάσκες» που αντιστοιχούν στο πνευμονικό παρέγχυμα χρησιμοποιήθηκαν στην εργασία ούτως ώστε να πραγματοποιηθεί κατάτμηση των εικόνων και να διατηρηθεί μόνο το πνευμονικό παρέγχυμα της εκάστοτε τομής. Στη συνέχεια, οι κατατεμημένες τομές μαζί με τις αντίστοιχες «μάσκες» στις οποίες σκιαγραφούνται οι αλλοιώσεις, δίνονται ως είσοδος στο εκάστοτε μοντέλο, το οποίο εκπαιδεύεται ούτως ώστε να μπορεί να εντοπίσει και να πραγματοποιήσει τη σημασιολογική κατάτμηση των αλλοιώσεων σε εικόνες οι οποίες του είναι άγνωστες.

Η σκιαγράφηση για την κατασκευή των «μασκών» πραγματοποιήθηκε από δύο ιατρούς ακτινολόγους με 1-5 χρόνια εμπειρίας. Στη συνέχεια βελτιστοποιήθηκε από ιατρούς ακτινολόγους με 5-10 χρόνια εμπειρίας και τέλος επικυρώθηκε και βελτιστοποιήθηκε περαιτέρω από έναν ιατρό ακτινολόγο με πάνω από 10 χρόνια εμπειρίας στο τομέα της ακτινολογίας νοσημάτων του αναπνευστικού συστήματος [83].

Οι τομές που αντιστοιχούν σε κάθε ασθενή είναι αποθηκευμένες σε επέκταση Neuroimaging Informatics Technology Initiative (NiFTI). Στην Εικόνα 4.3 παρουσιάζεται ένα υποσύνολο του συνόλου δεδομένων COVID-19-CT-Seg και των αντίστοιχων «μασκών» όπου σκιαγραφούνται αλλοιώσεις από πνευμονία COVID-19.



**Εικόνα 4.3:** Υποσύνολο από τις τομές αξονικών τομογραφιών και τις αντίστοιχες «μάσκες» του συνόλου δεδομένων COVID-19-CT-Seg

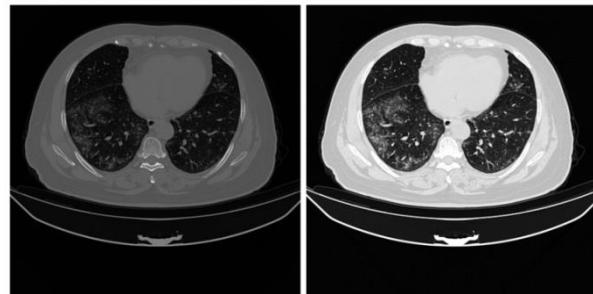
## 4.2 Προεπεξεργασία Δεδομένων

Τα σύνολα δεδομένων που έχουν επιλεγεί να χρησιμοποιηθούν, πριν τροφοδοτηθούν ως είσοδος σε ένα μοντέλο βαθιάς μάθησης, υπόκεινται σε τεχνικές προεπεξεργασίας, όπως αναφέρθηκε και στο Κεφάλαιο 3. Οι τεχνικές προεπεξεργασίας που εφαρμόζονται παρουσιάζονται στις επόμενες παραγράφους.

### 4.2.1 Κανονικοποίηση Αντίθεσης (Contrast) και Φωτεινότητας (Brightness) Τομών Αξονικών Τομογραφιών

Η κανονικοποίηση της αντίθεσης και φωτεινότητας των τομών αξονικών τομογραφιών που αξιοποιήθηκαν στο παρόν κεφάλαιο πραγματοποιείται όπως περιγράφεται στο Κεφάλαιο 3. Οι αξονικές τομογραφίες του συνόλου δεδομένων COVID-19-CT-Seg έχουν επέκταση Nifti. Αφού γίνει μετατροπή σε επέκταση DICOM, ελέγχεται εάν οι τομές είναι εκφρασμένες σε μονάδες της κλίμακας

Hounsfield και πραγματοποιείται μετατροπή του παραθύρου πνεύμονα στο κατάλληλο εύρος. Στην Εικόνα 4.4 παρουσιάζεται μία τομή αξονικής τομογραφίας πριν και μετά την εφαρμογή κανονικοποίησης αντίθεσης και φωτεινότητας.

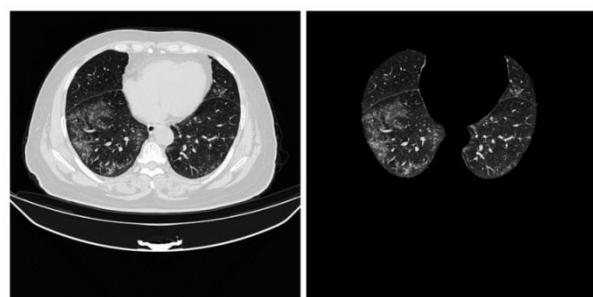


Εικόνα 4.4: Τομή αξονικής τομογραφίας πριν και μετά την κανονικοποίηση αντίθεσης και φωτεινότητας

#### 4.2.2 Δημιουργία Παραλλαγής του Συνόλου Δεδομένων όπου Διατηρείται μόνο το Πνευμονικό Παρέγχυμα στις Τομές των Αξονικών Τομογραφιών

Η διατήρηση μόνο του πνευμονικού παρεγχύματος στις τομές των αξονικών τομογραφιών πραγματοποιείται αξιοποιώντας τις «μάσκες» του συνόλου δεδομένων COVID-19-CT-Seg, στις οποίες σκιαγραφείται το εσωτερικό του πνευμονικού παρεγχύματος της εκάστοτε τομής.

Εν τέλει, δημιουργούνται δύο παραλλαγές των συνόλων δεδομένων που αξιοποιούνται στο παρόν πείραμα, όπως και στο Κεφάλαιο 3. Στην Εικόνα 4.5 παρουσιάζεται μία τομή αξονικής τομογραφίας πριν και μετά την κατάτμηση του πνευμονικού παρεγχύματος.



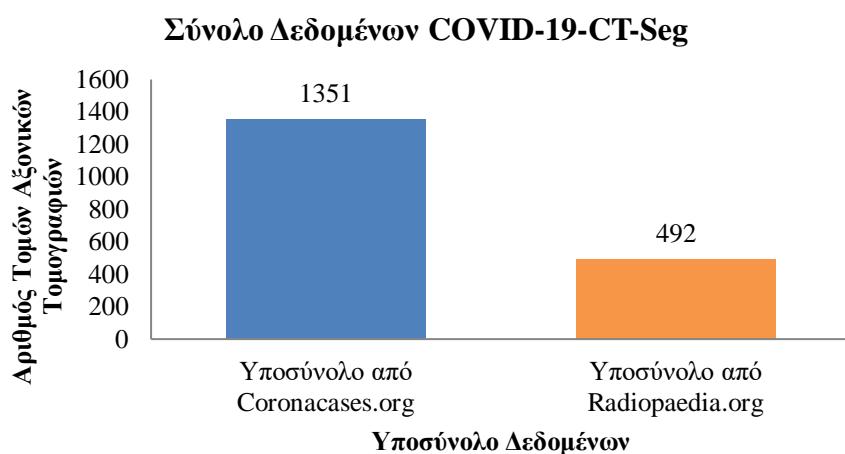
Εικόνα 4.5: Τομή αξονικής τομογραφίας πριν και μετά την κατάτμηση πνευμονικού παρεγχύματος

#### **4.2.3 Μετατροπή Εικόνων από Μορφότυπο DICOM σε Μορφότυπο TIFF**

Η μετατροπή των εικόνων από μορφότυπο DICOM σε μορφότυπο TIFF πραγματοποιείται όπως περιγράφεται στο Κεφάλαιο 3.

#### **4.2.4 Αφαίρεση των Τομών που δεν Εμφανίζουν Αλλοιώσεις**

Ομοίως με το πείραμα του Κεφαλαίου 3, πραγματοποιείται αφαίρεση των τομών που δεν εμφανίζουν αλλοιώσεις από το σύνολο COVID-19 CT-Seg. Επομένως, διατηρούνται στο σύνολο δεδομένων 1843 τομές, όπου οι 1351 τομές προέρχονται από το αποθετήριο Coronacases.org, ενώ οι υπόλοιπες 492 τομές προέρχονται από το αποθετήριο Radiopaedia.org. Η κατανομή των τομών στα δύο υποσύνολα έπειτα από την αφαίρεση των τομών που δεν εμφανίζουν αλλοιώσεις φαίνεται στην Εικόνα 4.6.



**Εικόνα 4.6: Ραβδόγραμμα υποσυνόλων των συνόλου δεδομένων COVID-19-CT-Seg έπειτα από την αφαίρεση των τομών που δεν εμφανίζουν αλλοιώσεις**

#### **4.2.5 Αλλαγή Διαστάσεων Εικόνων**

Η αλλαγή των διαστάσεων των εικόνων των συνόλων δεδομένων που αξιοποιούνται στο παρόν πείραμα πραγματοποιείται όπως περιγράφεται στο Κεφάλαιο 3.

#### **4.2.6 Κανονικοποίηση Τιμών Εικονοστοιχείων στο Εύρος Τιμών 0-1**

Η κανονικοποίηση των τιμών των εικονοστοιχείων των εικόνων των συνόλων δεδομένων που αξιοποιούνται στο παρόν πείραμα πραγματοποιείται όπως περιγράφεται στο Κεφάλαιο 3.

#### **4.2.7 Διαχωρισμός Δεδομένων σε Σύνολα Εκπαίδευσης/Επικύρωσης/Ελέγχου**

Για τις ανάγκες του πειράματος σημασιολογικής κατάτμησης αλλοιώσεων από COVID-19 πνευμονία σε τομές αξονικών τομογραφιών ακολουθείται η ίδια τακτική που ακολουθείται και στο Κεφάλαιο 3 όσον αφορά τον διαχωρισμό των δεδομένων. Τα δεδομένα επομένως διαχωρίζονται στα ακόλουθα τρία σύνολα:

Στα σύνολα εκπαίδευσης, επικύρωσης και ελέγχου, για την κατασκευή των οποίων αξιοποιούνται οι τομές του συνόλου COVID-19-CT-Seg έπειτα από την αφαίρεση των τομών που δεν εμφανίζουν αλλοιώσεις. Συγκεκριμένα, για την κατασκευή των συνόλων εκπαίδευσης και επικύρωσης αξιοποιούνται οι 1351 τομές του συνόλου COVID-19-CT-Seg που προέρχονται από το αποθετήριο Coronacases.org. Για την κατασκευή του συνόλου ελέγχου, αξιοποιούνται οι 492 τομές του συνόλου COVID-19-CT-Seg που προέρχονται από το αποθετήριο Radiopaedia.org.

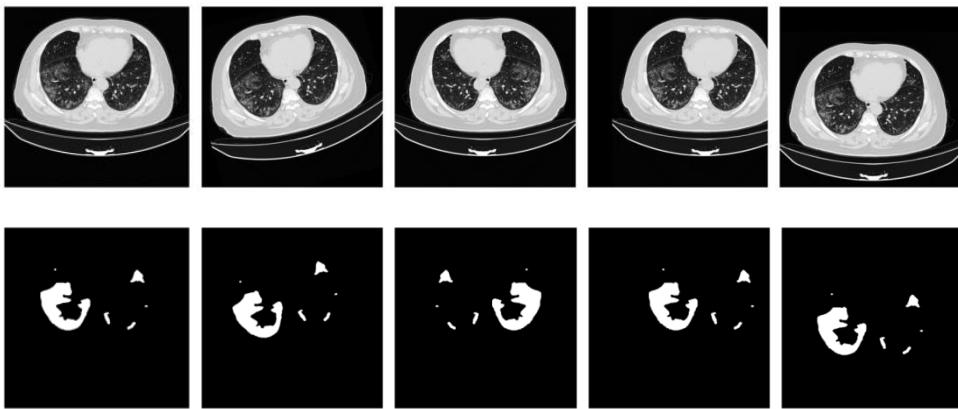
Επιπλέον, όπως και στο πείραμα του Κεφαλαίου 3, πραγματοποιείται διαχωρισμός των δεδομένων σε σύνολα εκπαίδευσης και επικύρωσης με τους ακόλουθους δύο τρόπους:

Σε επίπεδο τομής (slice level), όπου πραγματοποιείται τυχαίος διαχωρισμός των τομών αξονικών τομογραφιών στα σύνολα εκπαίδευσης και επικύρωσης χωρίς κάποιο κριτήριο. Σε επίπεδο ασθενή (patient level), όπου πραγματοποιείται τυχαίος διαχωρισμός των τομών αξονικών τομογραφιών στα σύνολα εκπαίδευσης και επικύρωσης με βάση τον ασθενή στον οποίο ανήκουν.

Επομένως και στο πείραμα του παρόντος κεφαλαίου η διεξαγωγή του πειράματος πραγματοποιείται πάνω σε δύο εκδοχές διαχωρισμού των δεδομένων του αρχικού συνόλου για τον ίδιο λόγο που αυτό συμβαίνει και στο πείραμα του Κεφαλαίου 3.

#### **4.2.8 Προσαύξηση Δεδομένων (Data Augmentation)**

Για την προσαύξηση των δεδομένων των συνόλων που αξιοποιούνται στο παρόν κεφάλαιο χρησιμοποιείται η ίδια τεχνική που αξιοποιείται και στο Κεφάλαιο 3. Στην Εικόνα 4.7 παρουσιάζεται η εφαρμογή προσαύξησης δεδομένων μέσω γεωμετρικών μετασχηματισμών σε τομή του συνόλου δεδομένων COVID-19-CT-Seg.



**Εικόνα 4.7:** Εφαρμογή προσαυξησης δεδομένων σε τομή αξονικής τομογραφίας του συνόλου COVID-19-CT-Seg

### 4.3 Αρχιτεκτονικές Τεχνητών Νευρωνικών Δικτύων

Στη συγκεκριμένη υποενότητα παρουσιάζονται οι αρχιτεκτονικές τεχνητών νευρωνικών δικτύων που χρησιμοποιήθηκαν για την επίλυση του προβλήματος της σημασιολογικής κατάτμησης αλλοιώσεων από COVID-19 πνευμονία σε τομές αξονικών τομογραφιών. Όλα τα μοντέλα που εφαρμόσθηκαν αποτελούν παραλλαγές της βασικής αρχιτεκτονικής τεχνητών νευρωνικών δικτύων που χρησιμοποιείται για την επίλυση προβλημάτων κατάτμησης βιοιατρικών εικόνων και η οποία ονομάζεται U-Net [86]. Πιο συγκεκριμένα, χρησιμοποιούνται παραλλαγές της αρχιτεκτονικής U-Net με διαφορετικές «ραχοκοκαλιές» (backbones), δηλαδή αρχιτεκτονικές U-Net όπου το πρώτο μέρος της αρχιτεκτονικής, το οποίο ονομάζεται μονοπάτι συστολής (contractive path) και αποτελεί τον κωδικοποιητή (encoder) του δικτύου, έχει αντικατασταθεί με αρχιτεκτονικές προεκπαιδευμένων συνελικτικών νευρωνικών δικτύων, αξιοποιώντας έτσι την τεχνική που ονομάζεται μεταφορά μάθησης.

#### 4.3.1 Αρχιτεκτονική U-Net

Η αρχιτεκτονική U-Net παρουσιάσθηκε το 2015 από τους ερευνητές Ronneberger κ.α. και αποτελεί μία αρχιτεκτονική συνελικτικού νευρωνικού δικτύου που χρησιμοποιείται σε προβλήματα κατάτμησης βιοιατρικών εικόνων [86].

Το συγκεκριμένο δίκτυο αποτελείται από τρία κύρια μέρη, το μονοπάτι συστολής (contracting path), το επίπεδο συμπίεσης (bottleneck level) και το μονοπάτι επέκτασης (expansive path).

Το μονοπάτι συστολής, που αποτελεί τον κωδικοποιητή (encoder) του δικτύου, ακολουθεί την τυπική δομή ενός συνελικτικού νευρωνικού δικτύου και στην Εικόνα 4.7 απεικονίζεται στα αριστερά [86]. Σε κάθε μπλοκ του συγκεκριμένου μονοπατιού

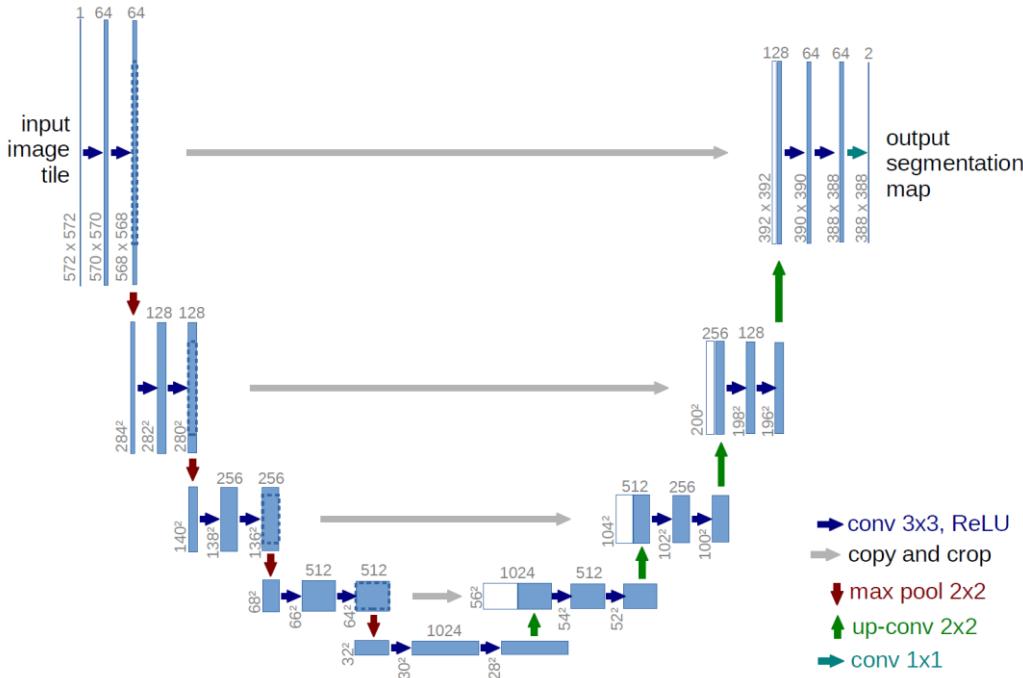
πραγματοποιείται μείωση των χωρικών διαστάσεων των εικόνων στο μισό και διπλασιασμός των καναλιών χαρακτηριστικών (feature channels) [85].

Το επίπεδο συμπίεσης (bottleneck level) συνδέει το μονοπάτι επέκτασης με το μονοπάτι συστολής [86].

Το μονοπάτι επέκτασης, που αποτελεί τον αποκωδικοποιητή (decoder) του δικτύου, απεικονίζεται στην Εικόνα 4.8 στα δεξιά [86]. Σε κάθε μπλοκ του συγκεκριμένου μονοπατιού πραγματοποιείται διπλασιασμός των χωρικών διαστάσεων των εικόνων και μείωση των καναλιών χαρακτηριστικών (feature channels) στο μισό [86].

Η έξοδος του εκάστοτε μπλοκ του μονοπατιού συστολής συννενώνεται (concatenated) με την είσοδο του αντίστοιχου μπλοκ του μονοπατιού επέκτασης. Αυτό επιτρέπει στο δίκτυο να διατηρήσει χαρακτηριστικά υψηλής ανάλυσης (high-resolution features) από το μονοπάτι συστολής [86]. Η πρώτη υλοποίηση του μοντέλου U-Net αποτελούνταν από 23 επίπεδα συνέλιξης [86].

Ο τρόπος με τον οποίο έχει υλοποιηθεί η αρχιτεκτονική του μοντέλου U-Net, του επιτρέπει να εκμεταλλευτεί τόσο υψηλού επιπέδου χαρακτηριστικά της εικόνας εισόδου όσο και χαμηλού επιπέδου χαρακτηριστικά αυτής, αυξάνοντας έτσι την ακρίβεια της κατάτμησης [86]. Ένα πλεονέκτημα του μοντέλου U-Net αποτελεί το γεγονός ότι διαθέτει μικρότερο αριθμό παραμέτρων, επιτυγχάνοντας έτσι χαμηλότερο χρόνο εκτέλεσης συγκριτικά με άλλες μεθόδους κατάτμησης [86].



Εικόνα 4.8: Διάγραμμα Αρχιτεκτονικής U-Net [86]

Στην παρούσα εργασία επιλέχθηκε να αξιολογηθεί η αποτελεσματικότητα του μοντέλου U-Net και των διαφόρων παραλλαγών του, που δημιουργούνται όταν αντικαταστήσουμε τον κωδικοποιητή του με κάποιο προεκπαιδευμένο συνελικτικό νευρωνικό δίκτυο. Η αντικατάσταση αυτή επιτρέπει την αξιοποίηση της τεχνικής που καλείται μεταφορά μάθησης. Ορισμένα χαρακτηριστικά χαμηλού επιπέδου όπως είναι οι ακμές, τα σχήματα και τα χρώματα είναι κοινά μεταξύ των εικόνων. Επομένως, η εμπειρία και η γνώση που έχουν αποκομίσει τα προεκπαιδευμένα συνελικτικά νευρωνικά δίκτυα από την προγενέστερη εκπαίδευση τους μπορούν να αξιοποιηθούν για την καλύτερη και γρηγορότερη εκπαίδευση του μοντέλου U-Net.

Τα προεκπαιδευμένα συνελικτικά νευρωνικά δίκτυα, που έχουν επιλεγεί να αποτελέσουν τον κωδικοποιητή του μοντέλου U-Net για το συγκεκριμένο πρόβλημα, είναι τα μοντέλα VGG16 και DenseNet121. Δημιουργούνται επομένως δύο παραλλαγές του μοντέλου U-Net, οι οποίες μαζί με το κλασικό μοντέλο U-Net, αξιολογούνται αν πραγματοποιούν ικανοποιητικά τη διαδικασία της κατάτμησης τομών έπειτα από εκπαίδευση. Για την υλοποίηση των παραλλαγών της αρχιτεκτονικής U-Net, αξιοποιήθηκε το πακέτο `segmentation_models` των Iakubovskii κ.α. [87].

#### 4.3.2 Αρχιτεκτονική U-Net με κωδικοποιητή την αρχιτεκτονική VGG16

Στη συγκεκριμένη αρχιτεκτονική ο κωδικοποιητής αντικαθίσταται με το μοντέλο VGG16. Το μοντέλο VGG16 αποτελείται από 21 επίπεδα τα οποία οργανώνονται σε πέντε μπλοκ επιπέδων [59]. Στη συνέχεια, το δίκτυο περιλαμβάνει το επίπεδο συμπίεσης, το οποίο συνδέει το μοντέλο VGG16 με το μονοπάτι συστολής. Όσον αφορά τον αποκωδικοποιητή, αυτός αποτελείται από πέντε μπλοκ επιπέδων. Οι έξοδοι των τεσσάρων τελευταίων μπλοκ του κωδικοποιητή συνενώνονται (concatenated) με τις εισόδους των τεσσάρων πρώτων μπλοκ του αποκωδικοποιητή.

#### 4.3.3 Αρχιτεκτονική U-Net με κωδικοποιητή την αρχιτεκτονική DenseNet121

Στη συγκεκριμένη αρχιτεκτονική ο κωδικοποιητής αντικαθίσταται με το μοντέλο DenseNet121. Το μοντέλο DenseNet121 περιλαμβάνει ένα επίπεδο συνέλιξης το οποίο ακολουθείται από ένα επίπεδο εξαγωγής τοπικού μεγίστου (max pooling layer) [54]. Στη συνέχεια η αρχιτεκτονική περιλαμβάνει 4 dense block, ανάμεσα στα οποία τοποθετούνται transition blocks [54]. Όσον αφορά τον αποκωδικοποιητή, αυτός αποτελείται από πέντε μπλοκ επιπέδων. Οι έξοδοι του επιπέδου συνέλιξης και των τριών πρώτων dense μπλοκ του κωδικοποιητή συνενώνονται (concatenated) με τις εισόδους των τεσσάρων πρώτων μπλοκ του αποκωδικοποιητή.

## 4.4 Αξιολόγηση Μοντέλων

Στη συγκεκριμένη υποενότητα παρουσιάζονται οι μέθοδοι που χρησιμοποιούνται για την αξιολόγηση του πειράματος που περιγράφεται στο παρόν κεφάλαιο. Οι μέθοδοι αξιολόγησης που αξιοποιούνται είναι η μέθοδος της διασταυρούμενης επικύρωσης 5-πτυχών (5-fold cross validation), οι καμπύλες εκπαίδευσης και επικύρωσης, καθώς και οι μετρικές Αξιοπιστία (Precision), Ανάκληση (Recall), Dice Similarity Coefficient (DSC) και Intersection over Union (IoU).

### 4.4.1 Εφαρμογή Διασταυρούμενης Επικύρωσης 5-Πτυχών

Για να ελέγξουμε την ικανότητα των μοντέλων να πραγματοποιούν σημασιολογική κατάτμηση των τομών σε περιοχές που εμφανίζουν αλλοιώσεις, καθώς και για την ανίχνευση του φαινομένου της υπερπροσαρμογής, εφαρμόζεται η τεχνική διασταυρούμενης επικύρωσης k-πτυχών (k-fold cross validation) με τιμή k=5 όπως και στο Κεφάλαιο 3.

### 4.4.2 Καμπύλες Εκπαίδευσης και Επικύρωσης

Για την αξιολόγηση της επίδοσης ενός μοντέλου κατά την εκπαιδευτική διαδικασία, τον προσδιορισμό ύπαρξης ή μη υποπροσαρμογής και υπερπροσαρμογής του μοντέλου και κατ' επέκταση για τη διερεύνηση της ικανότητας γενίκευσης του, χρησιμοποιούνται γραφικές παραστάσεις που παρουσιάζουν τις καμπύλες της διαδικασίας εκπαίδευσης και επικύρωσης όπως και στο Κεφάλαιο 3.

### 4.4.3 Αξιοπιστία (Precision)

Η μετρική Αξιοπιστία (Precision) όσον αφορά το πρόβλημα της σημασιολογικής κατάτμησης εικόνων, ορίζεται ως ο λόγος των εικονοστοιχείων που ορθά έχουν συμπεριληφθεί στην κατάτμηση της εικόνας, προς το σύνολο όλων των εικονοστοιχείων που έχουν συμπεριληφθεί στην κατάτμηση της εικόνας από τον αλγόριθμο. Η μετρική Αξιοπιστία δίνεται στην Εξίσωση (4.1):

$$\text{Αξιοπιστία (Precision)} = \frac{TP}{TP + FP} \quad (4.1)$$

#### 4.4.4 Ανάκληση (Recall)

Η μετρική Ανάκληση (Recall) ή αλλιώς Ευαισθησία (Sensitivity) όσον αφορά το πρόβλημα της σημασιολογικής κατάτμησης εικόνων, ορίζεται ως ο λόγος των εικονοστοιχείων που ορθά έχουν συμπεριληφθεί στην κατάτμηση της εικόνας προς το σύνολο όλων των εικονοστοιχείων που θα έπρεπε να έχουν συμπεριληφθεί στην κατάτμηση της εικόνας. Η μετρική Ανάκληση δίνεται στην Εξίσωση (4.2):

$$\text{Ανάκληση (Recall)} = \frac{TP}{TP + FN} \quad (4.2)$$

#### 4.4.5 Dice Similarity Coefficient (DSC)

Η μετρική Dice Similarity Coefficient ή Sørensen Dice ή F1 Score αποτελεί μία από τις δύο βασικότερες μετρικές αξιολόγησης ενός προβλήματος σημασιολογικής κατάτμησης εικόνων [88]. Η προαναφερθείσα μετρική αξιολόγησης αποτελεί τον αρμονικό μέσο μεταξύ των μετρικών αξιολόγησης Αξιοπιστία (Precision) και Ανάκληση (Recall) [88]. Όσον αφορά ένα πρόβλημα κατάτμησης, η συγκεκριμένη μετρική αξιολόγησης υπολογίζει τη χωρική αλληλοεπικάλυψη μεταξύ δύο περιοχών κατάτμησης [89]. Έστω ότι έχουμε δύο περιοχές κατάτμησης A και B, ο μαθηματικός τύπος της μετρικής Dice Similarity Coefficient δίνεται στην Εξίσωση (4.3):

$$DSC(A, B) = \frac{2 * |A \cap B|}{(A + B)} = \frac{2 * TP}{2 * TP + FP + FN} \quad (4.3)$$

#### 4.4.6 Intersection over Union (IoU)

Μία άλλη εξίσου σημαντική μετρική αξιολόγησης ενός προβλήματος σημασιολογικής κατάτμησης εικόνων, αποτελεί η μετρική Intersection over Union (IoU) ή αλλιώς Jaccard Index [90]. Χρησιμοποιείται και αυτή όπως και η Dice Similarity Coefficient για να υπολογίσει τη χωρική αλληλοεπικάλυψη μεταξύ δύο περιοχών κατάτμησης. Έστω ότι έχουμε δύο περιοχές κατάτμησης A και B, ο μαθηματικός τύπος της μετρικής Intersection over Union δίνεται στην Εξίσωση (4.4):

$$IoU(A, B) = \frac{|A \cap B|}{(A \cup B)} = \frac{TP}{TP + FP + FN} \quad (4.4)$$

## 4.5 Σχετικές Εργασίες

Μια από τις πρώτες εργασίες που δημοσιεύθηκαν και πραγματοποιούσαν εφαρμογή μοντέλων βαθειάς μάθησης για τη διερεύνηση της αποτελεσματικότητάς τους στη σημασιολογική κατάτμηση αλλοιώσεων σε τομές αξονικών τομογραφιών ήταν η εργασία των Ma κ.α. [83]. Οι συγκεκριμένοι ερευνητές πραγματοποίησαν εκπαίδευση και αξιολόγηση ενός 3D U-Net [91] χρησιμοποιώντας το σύνολο δεδομένων COVID-19-CT-Seg. Επιπλέον, πραγματοποίησαν εξωτερική αξιολόγησή του στο σύνολο MosMed [92]. Τα δεδομένα διαχωρίστηκαν ανά ασθενή και το 20% του συνόλου χρησιμοποιήθηκε για εκπαίδευση, ενώ το 80% για επικύρωση [83]. Για την καλύτερη αξιολόγηση της επίδοσης των μοντέλων, η ερευνητική ομάδα εφάρμοσε Διασταυρούμενη Επικύρωση 5-Πτυχών [83]. Άλλες εργασίες που αξιοποίησαν το σύνολο COVID-19-CT-Seg για να εκπαιδεύσουν 3D U-Nets είναι των Müller κ.α. [93], [94], που στην εργασία [93] πέτυχαν υψηλότερα αποτελέσματα από τους Ma κ.α., χρησιμοποιώντας το 80% του συνόλου για εκπαίδευση και το 20% για επικύρωση, ενώ στην εργασία [94] κατάφεραν να αυξήσουν ακόμη περισσότερο την επίδοση του 3D U-Net μέσω εφαρμογής προσαύξησης δεδομένων και κανονικοποίησης των τιμών HU των εικόνων. Οι Aswalthy κ.α. [95] υλοποίησαν ένα σύστημα το οποίο περιλαμβάνει δύο 3D U-Net αρχιτεκτονικές. Η πρώτη αρχιτεκτονική χρησιμοποιήθηκε για να πραγματοποιήσει κατάτμηση του πνευμονικού παρεγχύματος των τομών του συνόλου COVID-19-CT-Seg, ενώ η δεύτερη αρχιτεκτονική εκπαιδεύτηκε ούτως ώστε να πραγματοποιεί κατάτμηση αλλοιώσεων [95].

Ενδιαφέρον παρουσιάζει και η προσέγγιση των ερευνητών Fan κ.α. [96] που χρησιμοποιώντας 100 τομές αξονικών τομογραφιών που διαθέτουν αντίστοιχες μάσκες αλλοιώσεων και 1600 τομές αξονικών για τις οποίες δεν διατίθενται μάσκες, εκπαίδευσαν δύο παραλλαγές του ίδιου μοντέλου (Inf-Net και Semi-Inf-Net) ακολουθώντας μία προσέγγιση ημι-επιβλεπόμενης μάθησης.

Η ερευνητική ομάδα των Voulodimos κ.α. [97] εκπαίδευσε και αξιολόγησε ένα FCN-8 [98] και ένα U-Net μοντέλο και επιπλέον πρότεινε την εκπαίδευση του μοντέλου U-Net χρησιμοποιώντας πολύ μικρό αριθμό δεδομένων (few-shot learning).

Οι ομάδες των Saood κ.α. [99] και Polat κ.α. [100] αντιμετώπισαν το πρόβλημα της σημασιολογικής κατάτμησης αλλοιώσεων από πνευμονία COVID-19 ως ένα πολυταξικό πρόβλημα σημασιολογικής κατάτμησης. Οι μεν εκπαίδευσαν μία αρχιτεκτονική U-Net και μια αρχιτεκτονική SegNet [101] και οι δε μία αρχιτεκτονική DeepLabV3+ [102].

Οι Enshaei κ.α. [103] εκπαίδευσαν και αξιολόγησαν ένα μοντέλο που υλοποίησαν οι ίδιοι και το οποίο ονόμασαν COVID-Rate σε ένα ιδιωτικό και τρία δημόσια σύνολα δεδομένων [103]. Τα δημόσια σύνολα που χρησιμοποίησαν ήταν το υποσύνολο του συνόλου COVID-CT-Seg που προέρχεται από το αποθετήριο

Coronacases.org και τα σύνολα COVID-19 CT segmentation [104] και COVID-CT-MD. Το υποσύνολο δεδομένων του συνόλου COVID-CT-Seg που προέρχεται από το αποθετήριο Coronacases.org και το ιδιωτικό σύνολο δεδομένων χρησιμοποιήθηκαν για την εκπαίδευση και επικύρωση του μοντέλου, ενώ τα σύνολα COVID-19 CT segmentation και COVID-CT-MD χρησιμοποιήθηκαν ως σύνολα εξωτερικής αξιολόγησης [103]. Εφαρμόσθηκε κατάτμηση του πνευμονικού παρεγχύματος στις τομές και ο διαχωρισμός των δεδομένων πραγματοποιήθηκε ανά ασθενή και επίσης εφαρμόσθηκε διασταυρούμενη επικύρωση 10-πτυχών [103].

## 4.6 Πειραματική Διαδικασία

Στις παρακάτω παραγράφους πραγματοποιείται αξιολόγηση της επίδοσης των μοντέλων που εκπαιδεύθηκαν ως προς την ικανότητα σημασιολογικής κατάτμησης των τομών αξονικών τομογραφιών, για την απομόνωση αλλοιώσεων που προέρχονται από τον ιό SARS-CoV-2. Η δομή των παραγράφων είναι ίδια με του Κεφαλαίου 3.

Για την εκπαίδευση των μοντέλων του πειράματος και την αξιολόγηση της ικανότητας τους να πραγματοποιούν σημασιολογική κατάτμηση ιατρικών εικόνων ώστε να διακρίνουν τις αλλοιώσεις που οφείλονται σε πνευμονία COVID-19, χρησιμοποιήθηκε το υποσύνολο του συνόλου COVID-19-CT-Seg που προήλθε από το αποθετήριο Coronacases.org. Το υποσύνολο του συνόλου COVID-19-CT-Seg που προήλθε από το αποθετήριο Radiopaedia.org. αξιοποιήθηκε ως ένα εξωτερικό σύνολο αξιολόγησης των μοντέλων.

Στα μοντέλα που υλοποιήθηκαν για τις ανάγκες του παρόντος πειράματος συγκαταλέγεται η αρχιτεκτονική U-Net καθώς και παραλλαγές της, όπου ο κωδικοποιητής (encoder) του δικτύου αντικαθίσταται με μοντέλα συνελικτικών νευρωνικών δικτύων τα οποία έχουν προεκπαιδευθεί στο σύνολο δεδομένων ImageNet. Ως κωδικοποιητές χρησιμοποιούνται τα μοντέλα VGG-16 και DenseNet121. Οι παράμετροι της εκάστοτε αρχιτεκτονικής παρουσιάζονται στον Πίνακα 4.1.

**Πίνακας 4.1:** Αρχιτεκτονικές που χρησιμοποιήθηκαν στο πρόβλημα σημασιολογικής κατάτμησης αλλοιώσεων και οι παράμετροί τους

Αρχιτεκτονική	Συνολικές Παράμετροι	Εκπαιδεύσιμες Παράμετροι	Μη Εκπαιδεύσιμες Παράμετροι
U-Net	7.857.857	7.851.969	5.888
VGG16 + U-Net	23.752.279	23.748.247	4.032
DenseNet121 + U-Net	12.144.983	12.059.351	85.632

Στον Πίνακα 4.2 συνοψίζονται οι τιμές που επιλέχθηκαν για τις υπερπαραμέτρους των μοντέλων. Τα μοντέλα που κατασκευάστηκαν εκπαιδεύθηκαν λαμβάνοντας ως είσοδο εικόνες διαστάσεων 256x256, για 200 εποχές, με ρυθμό μάθησης (learning rate) 0.001. Κατά τη διάρκεια της εκπαίδευσης διαδικασίας, κάθε φορά που παρουσιάζονται βελτίωση των τιμών των μετρικών επικύρωσης, πραγματοποιούνται αποθήκευση των βαρών του εκπαίδευσης μοντέλου. Επειδή η μνήμη δεν ήταν αρκετή ούτως ώστε να δεχτεί όλες τις εικόνες του συνόλου δεδομένων σε κάθε εποχή, τα δεδομένα διαχωρίζονται σε παρτίδες (batches) μεγέθους 32 όπως και στο Κεφάλαιο 3. Ως βελτιστοποιητής (optimizer) επιλέχθηκε ο Adam.

**Πίνακας 4.2: Υπερπαράμετροι διαδικασίας εκπαίδευσης για τη σημασιολογική κατάτμηση αλλοιώσεων**

Υπερπαράμετρος	Τιμή
Εποχές (Epochs)	200
Διαστάσεις Εικόνων	256x256
Ρυθμός Μάθησης (Learning Rate)	0.001
Μέγεθος Παρτίδας (Batch Size)	32
Βελτιστοποιητής (Optimizer)	Adam

Η εκπαίδευση των μοντέλων πραγματοποιήθηκε, όπως και στο Κεφάλαιο 3, σε τέσσερεις παραλλαγές του συνόλου δεδομένων COVID-19-CT-Seg. Οι διάφορες παραλλαγές συνοψίζονται στον Πίνακα 4.3.

**Πίνακας 4.3: Παραλλαγές του συνόλου δεδομένων COVID-19-CT-Seg**

Παραλλαγή	Κατάτμηση Πνευμονικού Παρεγχύματος	Διαχωρισμός ανά Ασθενή
1	OXI	OXI
2	OXI	NAI
3	NAI	OXI
4	NAI	NAI

Στις επόμενες παραγράφους παρουσιάζονται τα αποτελέσματα των διαδικασιών εκπαίδευσης και αξιολόγησης των μοντέλων. Για ευκολία, όσον αφορά το κλασσικό μοντέλο U-Net, θα αναφερόμαστε σε αυτό χρησιμοποιώντας την ονομασία “U-Net”, όσον αφορά το μοντέλο U-Net με κωδικοποιητή το μοντέλο DenseNet121, θα αναφερόμαστε σε αυτό χρησιμοποιώντας την ονομασία “U-Net + DenseNet121”, ενώ όσον αφορά το μοντέλο U-Net με κωδικοποιητή το μοντέλο VGG16, θα αναφερόμαστε σε αυτό χρησιμοποιώντας την ονομασία “U-Net + VGG16”.

#### 4.6.1 Σημασιολογική Κατάτμηση Αλλοιώσεων Πνευμονίας από COVID-19

##### 4.6.1.1 Παραλλαγή 1: Μη Κατάτμηση Πνευμονικού Παρεγχύματος – Διαχωρισμός Δεδομένων ανά Τομή

Στον Πίνακα 4.4 παρουσιάζονται οι μετρικές αξιολόγησης της ικανότητας των μοντέλων, έπειτα από εκπαίδευση, να πραγματοποιούν σημασιολογική κατάτμηση αλλοιώσεων από πνευμονία COVID-19 στις τομές του συνόλου δεδομένων επικύρωσης της Παραλλαγής 1.

Πίνακας 4.4: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Τομή: Μετρικές αξιολόγησης της εφαρμογής σημασιολογικής κατάτμησης αλλοιώσεων στα δεδομένα επικύρωσης

Μοντέλο	Μέση Dice Similarity Coefficient	Μέση Intersection over Union	Μέση Αξιοπιστία	Μέση Ανάκληση
<b>U-Net</b>	0,8520 ± 0,0063	0,7422 ± 0,0096	0,8503 ± 0,0073	0,8538 ± 0,0127
<b>U-Net + VGG16</b>	0,8534 ± 0,0065	0,7444 ± 0,0098	0,8435 ± 0,0146	0,8580 ± 0,0132
<b>U-Net + DenseNet121</b>	0,8635 ± 0,0078	0,7599 ± 0,0123	0,8559 ± 0,0116	0,8714 ± 0,0124

Παρατηρώντας τις μετρικές αξιολόγησης στον Πίνακα 4.4, διαπιστώνουμε πως και τα τρία μοντέλα παρουσιάζουν μεγάλη ικανότητα σημασιολογικής κατάτμησης αλλοιώσεων σε τομές αξονικών τομογραφιών.

Πιο αναλυτικά, όλα τα μοντέλα επιτυγχάνουν παρόμοια επίδοση, ωστόσο την υψηλότερη μέση τιμή όσον αφορά τη Μέση DSC και Μέση IoU επιτυγχάνει το μοντέλο U-Net + DenseNet121. Ειδικότερα, το μοντέλο U-Net + DenseNet121 επιτυγχάνει τιμές 0,8635 και 0,7599 όσον αφορά τη Μέση DSC, και Μέση IoU αντίστοιχα.

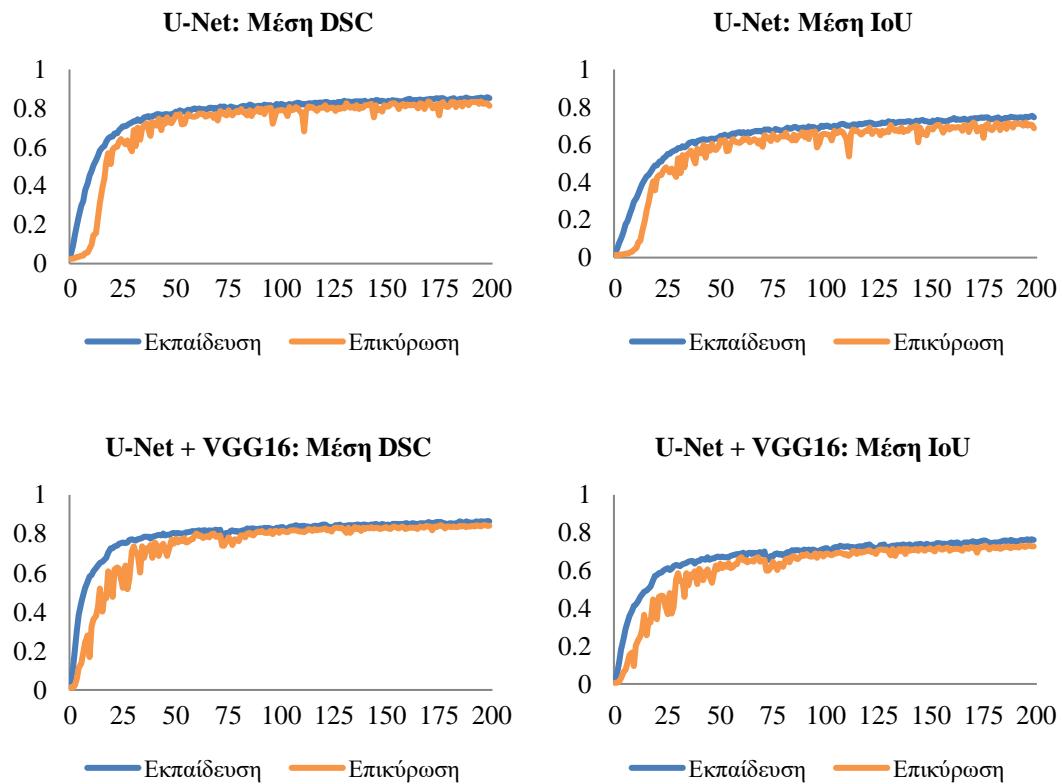
Τα άλλα δύο μοντέλα, δηλαδή το μοντέλο U-Net και το μοντέλο U-Net + VGG16 παρουσιάζουν παρόμοια επίδοση μεταξύ τους, με το μοντέλο U-Net + VGG16 να εμφανίζει ελαφρώς καλύτερα αποτελέσματα. Πιο συγκεκριμένα το μοντέλο U-Net + VGG16 επιτυγχάνει Μέση DSC και Μέση IoU 0,8534 και 0,7444 αντίστοιχα, ενώ το μοντέλο U-Net επιτυγχάνει Μέση DSC και Μέση IoU 0,8520 και 0,7422 αντίστοιχα.

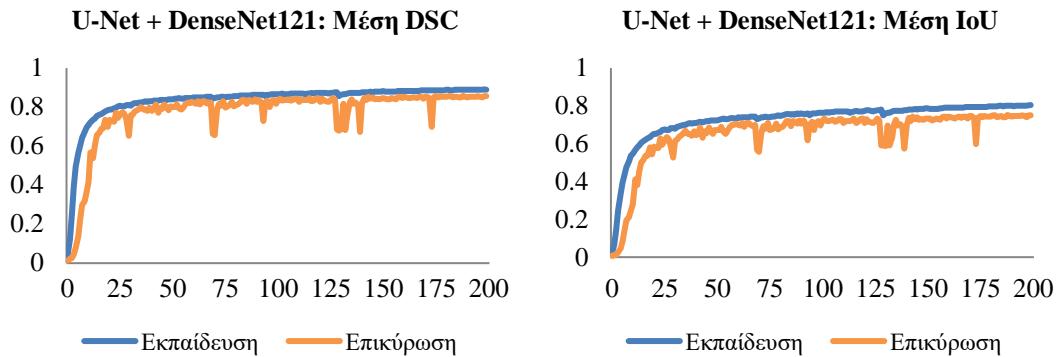
Όσον αφορά τη Μέση Αξιοπιστία, τα μοντέλα U-Net, U-Net + VGG16 και U-Net + DenseNet121 παρουσιάζουν τιμές 0,8503, 0,8435 και 0,8559 αντίστοιχα. Η πολύ υψηλή μέση τιμή Αξιοπιστίας των τριών μοντέλων δείχνει πως και τα τρία μοντέλα

είναι ικανά να πραγματοποιούν με πολύ μεγάλη ακρίβεια σημασιολογική κατάτμηση αλλοιώσεων σε μία τομή.

Όσον αφορά τη Μέση Ανάκληση, τα μοντέλα U-Net, U-Net + VGG16 και U-Net + DenseNet121 παρουσιάζουν τιμές 0,8538, 0,8580 και 0,8714 αντίστοιχα. Η πολύ υψηλή μέση τιμή Ανάκλησης των τριών μοντέλων δείχνει πως και τα τρία μοντέλα είναι ικανά να πραγματοποιούν σημασιολογική κατάτμηση ενός πολύ μεγάλου ποσοστού των αλλοιώσεων σε μία τομή.

Στην Εικόνα 4.9 παρουσιάζονται οι καμπύλες εκπαίδευσης και επικύρωσης των μοντέλων όσον αφορά τη σημασιολογική κατάτμηση αλλοιώσεων από πνευμονία COVID-19 στα δεδομένα της Παραλλαγής 1.





**Εικόνα 4.9: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Τομή: Καμπύλες εκπαίδευσης και επικύρωσης των μοντέλων όσον αφορά τη σημασιολογική κατάτμηση αλλοιώσεων**

Παρατηρώντας τις καμπύλες εκπαίδευσης και επικύρωσης στην Εικόνα 4.9, διαπιστώνουμε πως όλα τα μοντέλα που εκπαιδεύθηκαν και αξιολογήθηκαν στα σύνολα εκπαίδευσης και επικύρωσης αντίστοιχα, συγκλίνουν μετά το πέρας της εκπαιδευτικής διαδικασίας. Επιπλέον, το γεγονός ότι η απόσταση ανάμεσα στην καμπύλη εκπαίδευσης και στην καμπύλη επικύρωσης τόσο στη γραφική παράσταση Μέσης DSC όσο και στη γραφική παράσταση Μέσης IoU είναι μικρή και παραμένει σταθερή καθ' όλη τη διάρκεια της εκπαιδευτικής διαδικασίας για όλα τα μοντέλα, αποτελεί ένδειξη ότι τα μοντέλα δεν παρουσιάζουν υπερπροσαρμογή και ενδεχομένως έχουν μεγάλη ικανότητα γενίκευσης.

Στον Πίνακα 4.5 παρουσιάζονται οι μετρικές αξιολόγησης της ικανότητας των μοντέλων, έπειτα από εκπαίδευση, να πραγματοποιούν σημασιολογική κατάτμηση αλλοιώσεων από πνευμονία COVID-19 στις τομές του συνόλου δεδομένων ελέγχου της Παραλλαγής 1.

**Πίνακας 4.5: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Τομή: Μετρικές αξιολόγησης της εφαρμογής σημασιολογικής κατάτμησης αλλοιώσεων στα δεδομένα ελέγχου**

Μοντέλο	Μέση Dice Similarity Coefficient	Μέση Intersection over Union	Μέση Αξιοπιστία	Μέση Ανάκληση
<b>U-Net</b>	$0,6086 \pm 0,0390$	$0,4382 \pm 0,0393$	$0,7120 \pm 0,0378$	$0,5340 \pm 0,0570$
<b>U-Net + VGG16</b>	$0,6516 \pm 0,0338$	$0,4840 \pm 0,0369$	$0,8353 \pm 0,0485$	$0,5387 \pm 0,0607$
<b>U-Net + DenseNet121</b>	$0,6712 \pm 0,0322$	$0,5059 \pm 0,0358$	$0,8316 \pm 0,0417$	$0,5643 \pm 0,0424$

Παρατηρώντας τις μετρικές αξιολόγησης στον Πίνακα 4.5, διαπιστώνουμε πως και τα τρία μοντέλα παρουσιάζουν καλή ικανότητα σημασιολογικής κατάτμησης αλλοιώσεων σε τομές αξονικών τομογραφιών. Ωστόσο υπάρχει περιθώριο βελτίωσης.

Από τα τρία μοντέλα, το μοντέλο U-Net + DenseNet121 είναι αυτό που παρουσιάζει την καλύτερη επίδοση. Πιο συγκεκριμένα, το μοντέλο U-Net + DenseNet121 επιτυγχάνει τιμές 0,6712 και 0,5059 όσον αφορά τη Μέση DSC και Μέση IoU αντίστοιχα.

Όσον αφορά τα άλλα δύο μοντέλα, δηλαδή το μοντέλο U-Net και το μοντέλο U-Net + VGG16, το μοντέλο U-Net + VGG16 παρουσιάζει πολύ καλύτερη επίδοση από το μοντέλο U-Net. Αναλυτικότερα, το μοντέλο U-Net + VGG16 επιτυγχάνει τιμές 0,6516 και 0,4840 όσον αφορά τη Μέση DSC και Μέση IoU αντίστοιχα, ενώ το μοντέλο U-Net επιτυγχάνει τιμές 0,6086 και 0,4382 για τις ίδιες μετρικές.

Όσον αφορά τη Μέση Αξιοπιστία, τα μοντέλα U-Net, U-Net + VGG16 και U-Net + DenseNet121 παρουσιάζουν τιμές 0,7120, 0,8353 και 0,8316 αντίστοιχα. Η υψηλή μέση τιμή Αξιοπιστίας των τριών μοντέλων δείχνει πως και τα τρία μοντέλα είναι ικανά να πραγματοποιούν με μεγάλη ακρίβεια σημασιολογική κατάτμηση αλλοιώσεων σε μία τομή.

Όσον αφορά τη Μέση Ανάκληση, τα μοντέλα U-Net, U-Net + VGG16 και U-Net + DenseNet121 παρουσιάζουν τιμές 0,5340, 0,5387 και 0,5643 αντίστοιχα. Η μέτρια μέση τιμή Ανάκλησης των τριών μοντέλων δείχνει πως και τα τρία μοντέλα είναι ικανά να πραγματοποιούν σημασιολογική κατάτμηση ενός μέτριου ποσοστού των αλλοιώσεων σε μία τομή.

#### **4.6.1.2 Παραλλαγή 2: Μη Κατάτμηση Πνευμονικού Παρεγχύματος – Διαχωρισμός Δεδομένων ανά Ασθενή**

Στον Πίνακα 4.6 παρουσιάζονται οι μετρικές αξιολόγησης της ικανότητας των μοντέλων, έπειτα από εκπαίδευση, να πραγματοποιούν σημασιολογική κατάτμηση αλλοιώσεων από πνευμονία COVID-19 στις τομές του συνόλου δεδομένων επικύρωσης της Παραλλαγής 2.

**Πίνακας 4.6: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Ασθενή:**  
Μετρικές αξιολόγησης της εφαρμογής σημασιολογικής κατάτμησης αλλοιώσεων στα δεδομένα επικύρωσης

Μοντέλο	Μέση Dice Similarity Coefficient	Μέση Intersection over Union	Μέση Αξιοπιστία	Μέση Ανάκληση
<b>U-Net</b>	0,7477 ± 0,0615	0,6001 ± 0,0782	0,7554 ± 0,0741	0,7428 ± 0,0712
<b>U-Net + VGG16</b>	0,7618 ± 0,0619	0,6184 ± 0,0808	0,7545 ± 0,0909	0,7715 ± 0,0394
<b>U-Net + DenseNet121</b>	0,7813 ± 0,0546	0,6436 ± 0,0731	0,8083 ± 0,0618	0,7625 ± 0,0936

Παρατηρώντας τις μετρικές αξιολόγησης στον Πίνακα 4.6, διαπιστώνουμε πως και τα τρία μοντέλα παρουσιάζουν πολύ καλή ικανότητα σημασιολογικής κατάτμησης αλλοιώσεων σε τομές αξονικών τομογραφιών.

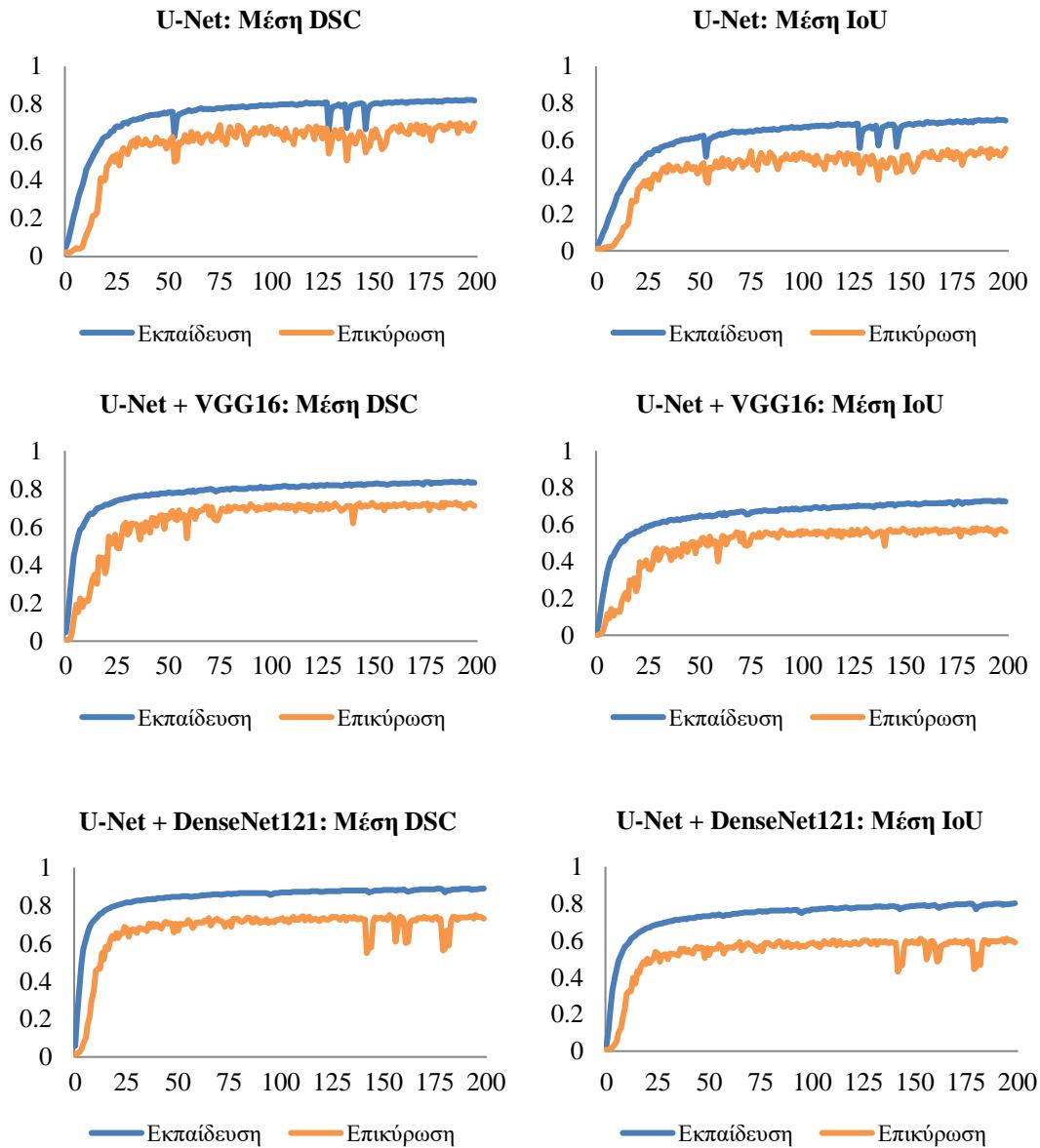
Όλα τα μοντέλα επιτυγχάνουν παρόμοια επίδοση, ωστόσο τις υψηλότερες τιμές όσον αφορά τη Μέση DSC και Μέση IoU επιτυγχάνει το μοντέλο U-Net + DenseNet121. Πιο αναλυτικά, το μοντέλο U-Net + DenseNet121 επιτυγχάνει τιμές 0,7813 και 0,6436 όσον αφορά τη Μέση DSC και Μέση IoU αντίστοιχα.

Τα άλλα δύο μοντέλα, δηλαδή το μοντέλο U-Net και το μοντέλο U-Net + VGG16 παρουσιάζουν χαμηλότερη επίδοση, με το μοντέλο U-Net + VGG16 να εμφανίζει καλύτερα αποτελέσματα. Πιο αναλυτικά, το μοντέλο U-Net + VGG16 επιτυγχάνει τιμές 0,7618 και 0,6184 όσον αφορά τη Μέση DSC και Μέση IoU αντίστοιχα, ενώ το μοντέλο U-Net επιτυγχάνει τιμές 0,7477 και 0,6001 για τις ίδιες μετρικές.

Όσον αφορά τη Μέση Αξιοπιστία, τα μοντέλα U-Net, U-Net + VGG16 και U-Net + DenseNet121 παρουσιάζουν τιμές 0,7554, 0,7545 και 0,8083 αντίστοιχα. Η υψηλή μέση τιμή Αξιοπιστίας των τριών μοντέλων δείχνει πως και τα τρία μοντέλα είναι ικανά να πραγματοποιούν με μεγάλη ακρίβεια σημασιολογική κατάτμηση αλλοιώσεων σε μία τομή.

Όσον αφορά τη Μέση Ανάκληση, τα μοντέλα U-Net, U-Net + VGG16 και U-Net + DenseNet121 παρουσιάζουν τιμές 0,7428, 0,7715 και 0,7625 αντίστοιχα. Η υψηλή μέση τιμή Ανάκλησης των τριών μοντέλων δείχνει πως και τα τρία μοντέλα είναι ικανά να πραγματοποιούν σημασιολογική κατάτμηση ενός μεγάλου ποσοστού των αλλοιώσεων σε μία τομή.

Στην Εικόνα 4.10 παρουσιάζονται οι καμπύλες εκπαίδευσης και επικύρωσης των μοντέλων όσον αφορά τη σημασιολογική κατάτμηση αλλοιώσεων από πνευμονία COVID-19 στα δεδομένα της Παραλλαγής 2.



**Εικόνα 4.10: Μη Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Ασθενή: Καμπύλες εκπαίδευσης και επικύρωσης των μοντέλων όσον αφορά τη σημασιολογική κατάτμηση αλλοιώσεων**

Παρατηρώντας τις καμπύλες εκπαίδευσης και επικύρωσης στην Εικόνα 4.10, διαπιστώνουμε πως όλα τα μοντέλα που εκπαιδεύθηκαν και αξιολογήθηκαν στα σύνολα εκπαίδευσης και επικύρωσης αντίστοιχα, συγκλίνουν μετά το πέρας της εκπαιδευτικής διαδικασίας. Ωστόσο, υπάρχει μία απόσταση ανάμεσα στην καμπύλη εκπαίδευσης και στην καμπύλη επικύρωσης τόσο στη γραφική παράσταση Μέσης DSC όσο και στη γραφική παράσταση Μέσης IoU, η οποία όμως παραμένει σταθερή καθ' όλη τη διάρκεια της εκπαιδευτικής διαδικασίας για όλα τα μοντέλα. Η απόσταση αυτή αποτελεί ένδειξη πως υπάρχει υπερπροσαρμογή των μοντέλων στα δεδομένα εκπαίδευσης και επίσης υποδηλώνει πως τα μοντέλα ενδεχομένως να μην διαθέτουν μεγάλη ικανότητα γενίκευσης.

Στον Πίνακα 4.7 παρουσιάζονται οι μετρικές αξιολόγησης της ικανότητας των μοντέλων, έπειτα από εκπαίδευση, να πραγματοποιούν σημασιολογική κατάτμηση αλλοιώσεων από πνευμονία COVID-19 στις τομές του συνόλου δεδομένων ελέγχου της Παραλλαγής 2.

**Πίνακας 4.7: Μη Κατάτμηση Πνευμονικού Παρεγγύματος - Διαχωρισμός Δεδομένων ανά Ασθενή:**  
Μετρικές αξιολόγησης της εφαρμογής σημασιολογικής κατάτμησης αλλοιώσεων στα δεδομένα ελέγχου

Μοντέλο	Μέση Dice Similarity Coefficient	Μέση Intersection over Union	Μέση Αξιοπιστία	Μέση Ανάκληση
<b>U-Net</b>	0,6270 ± 0,0235	0,4570 ± 0,0246	0,7190 ± 0,0883	0,5645 ± 0,0576
<b>U-Net + VGG16</b>	0,6600 ± 0,0290	0,4931 ± 0,0317	0,7947 ± 0,1003	0,5758 ± 0,0717
<b>U-Net + DenseNet121</b>	0,6615 ± 0,0347	0,4950 ± 0,0388	0,8039 ± 0,0750	0,5668 ± 0,0538

Παρατηρώντας τις μετρικές αξιολόγησης στον Πίνακα 4.7, διαπιστώνουμε πως και τα τρία μοντέλα παρουσιάζουν καλή ικανότητα σημασιολογικής κατάτμησης αλλοιώσεων σε τομές αξονικών τομογραφιών. Ωστόσο υπάρχει περιθώριο βελτίωσης.

Από τα τρία μοντέλα, τα μοντέλα U-Net + DenseNet121 και U-Net + VGG16 είναι αυτά που παρουσιάζουν τη καλύτερη επίδοση. Πιο συγκεκριμένα, το μοντέλο U-Net + DenseNet121 επιτυγχάνει τιμές 0,6615 και 0,4950 όσον αφορά τη Μέση DSC και Μέση IoU αντίστοιχα. Το μοντέλο U-Net + VGG16 επιτυγχάνει τιμές 0,6600 και 0,4931 όσον αφορά τη Μέση DSC και Μέση IoU αντίστοιχα, ενώ το μοντέλο U-Net επιτυγχάνει τιμές 0,6270 και 0,4570 για τις ίδιες μετρικές.

Οσον αφορά τη Μέση Αξιοπιστία, τα μοντέλα U-Net, U-Net + VGG16 και U-Net + DenseNet121 παρουσιάζουν τιμές 0,7190, 0,7947 και 0,8039 αντίστοιχα. Η υψηλή μέση τιμή Αξιοπιστίας των τριών μοντέλων δείχνει πως και τα τρία μοντέλα είναι ικανά να πραγματοποιούν με μεγάλη ακρίβεια σημασιολογική κατάτμηση αλλοιώσεων σε μία τομή.

Οσον αφορά τη Μέση Ανάκληση, τα μοντέλα U-Net, U-Net + VGG16 και U-Net + DenseNet121 παρουσιάζουν τιμές 0,5645, 0,5758 και 0,5668 αντίστοιχα. Η μέτρια μέση τιμή Ανάκλησης των τριών μοντέλων δείχνει πως και τα τρία μοντέλα είναι ικανά να πραγματοποιούν σημασιολογική κατάτμηση ενός μέτριου ποσοστού των αλλοιώσεων σε μία τομή.

#### 4.6.1.3 Παραλλαγή 3:Κατάτμηση Πνευμονικού Παρεγχύματος – Διαχωρισμός Δεδομένων ανά Τομή

Στον Πίνακα 4.8 παρουσιάζονται οι μετρικές αξιολόγησης της ικανότητας των μοντέλων, έπειτα από εκπαίδευση, να πραγματοποιούν σημασιολογική κατάτμηση αλλοιώσεων από πνευμονία COVID-19 στις τομές του συνόλου δεδομένων επικύρωσης της Παραλλαγής 3.

**Πίνακας 4.8: Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Τομή: Μετρικές αξιολόγησης της εφαρμογής σημασιολογικής κατάτμησης αλλοιώσεων στα δεδομένα επικύρωσης**

Μοντέλο	Μέση Dice Similarity Coefficient	Μέση Intersection over Union	Μέση Αξιοπιστία	Μέση Ανάκληση
<b>U-Net</b>	0,8423 ± 0,0131	0,7277 ± 0,0194	0,8312 ± 0,0267	0,8545 ± 0,0195
<b>U-Net + VGG16</b>	0,8573 ± 0,0072	0,7503 ± 0,0111	0,8489 ± 0,0120	0,8661 ± 0,0140
<b>U-Net + DenseNet121</b>	0,8644 ± 0,0091	0,7612 ± 0,0143	0,8597 ± 0,0112	0,8691 ± 0,0094

Παρατηρώντας τις μετρικές αξιολόγησης στον Πίνακα 4.8, διαπιστώνουμε πως και τα τρία μοντέλα παρουσιάζουν μεγάλη ικανότητα σημασιολογικής κατάτμησης αλλοιώσεων σε τομές αξονικών τομογραφιών.

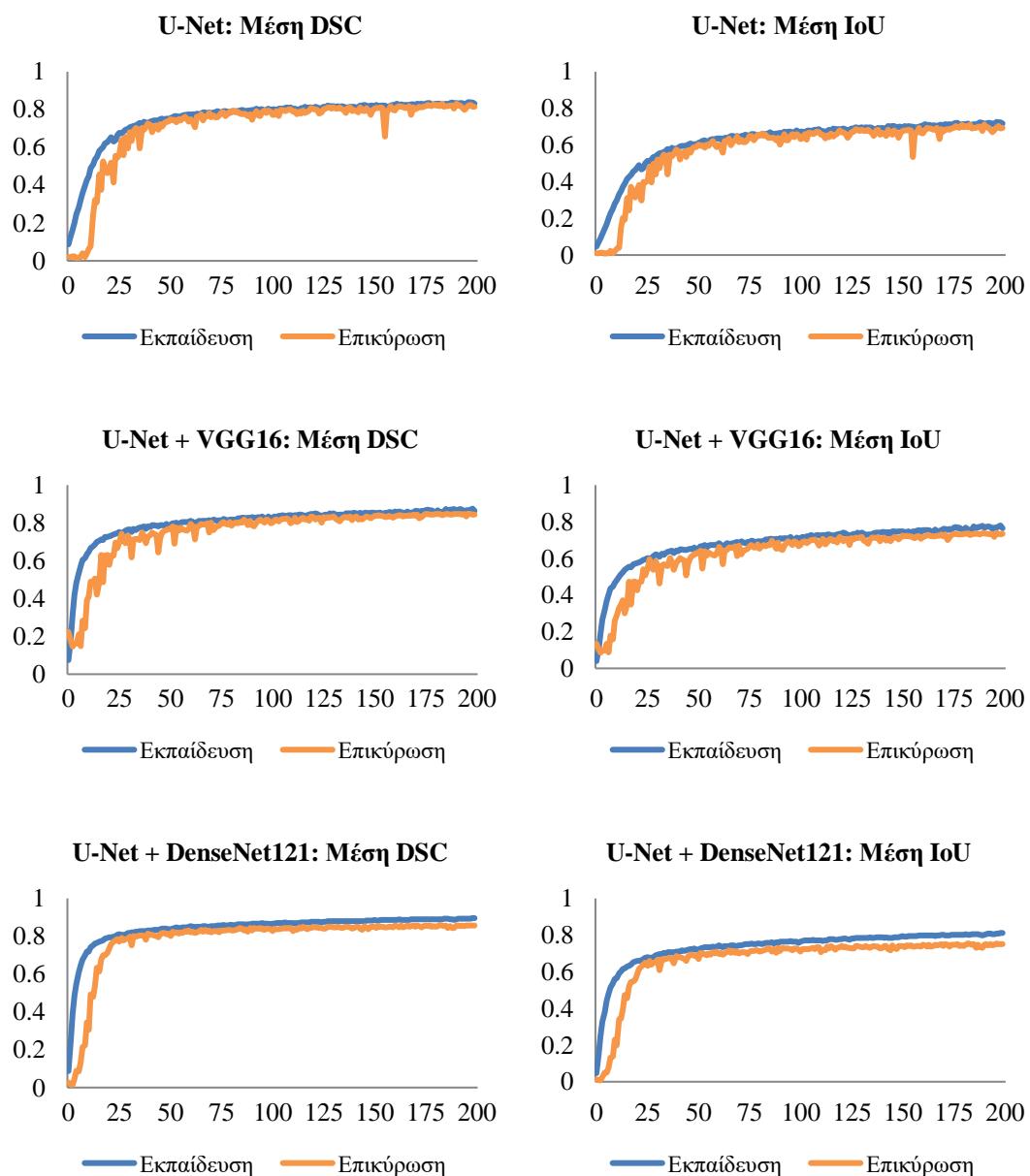
Πιο αναλυτικά, όλα τα μοντέλα επιτυγχάνουν παρόμοια επίδοση, ωστόσο την υψηλότερη μέση τιμή όσον αφορά τη Μέση DSC και Μέση IoU επιτυγχάνει το μοντέλο U-Net + DenseNet121. Πιο συγκεκριμένα, το μοντέλο U-Net + DenseNet121 επιτυγχάνει τιμές 0,8644 και 0,7612 όσον αφορά τη Μέση DSC, και Μέση IoU αντίστοιχα.

Τα άλλα δύο μοντέλα, δηλαδή το μοντέλο U-Net και το μοντέλο U-Net + VGG16 παρουσιάζουν χαμηλότερη επίδοση, με το μοντέλο U-Net + VGG16 να εμφανίζει καλύτερα αποτελέσματα. Ειδικότερα, το μοντέλο U-Net + VGG16 επιτυγχάνει τιμές 0,8573 και 0,7503 όσον αφορά τη Μέση DSC και Μέση IoU αντίστοιχα, ενώ το μοντέλο U-Net επιτυγχάνει τιμές 0,8423 και 0,7277 για τις ίδιες μετρικές.

Όσον αφορά τη Μέση Αξιοπιστία, τα μοντέλα U-Net, U-Net + VGG16 και U-Net + DenseNet121 παρουσιάζουν τιμές 0,8312, 0,8489 και 0,8597 αντίστοιχα. Η πολύ υψηλή μέση τιμή Αξιοπιστίας των τριών μοντέλων δείχνει πως και τα τρία μοντέλα είναι ικανά να πραγματοποιούν με πολύ μεγάλη ακρίβεια σημασιολογική κατάτμηση αλλοιώσεων σε μία τομή.

Όσον αφορά τη Μέση Ανάκληση, τα μοντέλα U-Net, U-Net + VGG16 και U-Net + DenseNet121 παρουσιάζουν τιμές 0,8545, 0,8661 και 0,8691 αντίστοιχα. Η πολύ υψηλή μέση τιμή Ανάκλησης των τριών μοντέλων δείχνει πως και τα τρία μοντέλα είναι ικανά να πραγματοποιούν σημασιολογική κατάτμηση ενός πολύ μεγάλου ποσοστού των αλλοιώσεων σε μία τομή.

Στην Εικόνα 4.11 παρουσιάζονται οι καμπύλες εκπαίδευσης και επικύρωσης των μοντέλων όσον αφορά τη σημασιολογική κατάτμηση αλλοιώσεων από πνευμονία COVID-19 στα δεδομένα της Παραλλαγής 3.



**Εικόνα 4.11: Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Τομή: Καμπύλες εκπαίδευσης και επικύρωσης των μοντέλων όσον αφορά τη σημασιολογική κατάτμηση αλλοιώσεων**

Παρατηρώντας τις καμπύλες εκπαίδευσης και επικύρωσης στην Εικόνα 4.11, διαπιστώνουμε πως όλα τα μοντέλα που εκπαιδεύθηκαν και αξιολογήθηκαν στα σύνολα εκπαίδευσης και επικύρωσης αντίστοιχα, συγκλίνουν μετά το πέρας της εκπαίδευτικής διαδικασίας. Επιπλέον, το γεγονός ότι η απόσταση ανάμεσα στη καμπύλη εκπαίδευσης και στη καμπύλη επικύρωσης τόσο στη γραφική παράσταση Μέσης DSC όσο και στη γραφική παράσταση Μέσης IoU είναι μικρή και παραμένει σταθερή καθ' όλη τη διάρκεια της εκπαίδευτικής διαδικασίας για όλα τα μοντέλα, αποτελεί ένδειξη ότι τα μοντέλα δεν παρουσιάζουν υπερπροσαρμογή και ενδεχομένως έχουν μεγάλη ικανότητα γενίκευσης.

Στον Πίνακα 4.9 παρουσιάζονται οι μετρικές αξιολόγησης της ικανότητας των μοντέλων, έπειτα από εκπαίδευση, να πραγματοποιούν σημασιολογική κατάτμηση αλλοιώσεων από πνευμονία COVID-19 στις τομές του συνόλου δεδομένων ελέγχου της Παραλλαγής 3.

**Πίνακας 4.9: Κατάτμηση Πνευμονικού Παρεγγύματος - Διαχωρισμός Δεδομένων ανά Τομή: Μετρικές αξιολόγησης της εφαρμογής σημασιολογικής κατάτμησης αλλοιώσεων στα δεδομένα ελέγχου**

Μοντέλο	Μέση Dice Similarity Coefficient	Μέση Intersection over Union	Μέση Αξιοπιστία	Μέση Ανάκληση
<b>U-Net</b>	0,6587 ± 0,0396	0,4921 ± 0,0435	0,8321 ± 0,0269	0,5487 ± 0,0631
<b>U-Net + VGG16</b>	0,7118 ± 0,0165	0,5528 ± 0,0200	0,8190 ± 0,0171	0,6305 ± 0,0325
<b>U-Net + DenseNet121</b>	0,7211 ± 0,0197	0,5641 ± 0,0238	0,8071 ± 0,0450	0,6566 ± 0,0584

Παρατηρώντας τις μετρικές αξιολόγησης στον Πίνακα 4.9, διαπιστώνουμε πως και τα τρία μοντέλα παρουσιάζουν μεγάλη ικανότητα σημασιολογικής κατάτμησης αλλοιώσεων σε τομές αξονικών τομογραφιών, ωστόσο υπάρχει περιθώριο βελτίωσης.

Από τα τρία μοντέλα, το μοντέλο U-Net + DenseNet121 είναι αυτά που παρουσιάζει τη καλύτερη επίδοση. Ειδικότερα, το μοντέλο U-Net + DenseNet121 επιτυγχάνει τιμές 0,7211 και 0,5641 όσον αφορά τη Μέση DSC και Μέση IoU αντίστοιχα.

Τα άλλα δύο μοντέλα, δηλαδή το μοντέλο U-Net και το μοντέλο U-Net + VGG16 παρουσιάζουν χαμηλότερη επίδοση, με το μοντέλο U-Net + VGG16 να εμφανίζει καλύτερα αποτελέσματα. Αναλυτικότερα, το μοντέλο U-Net + VGG16 επιτυγχάνει τιμές 0,7118 και 0,5528 όσον αφορά τη Μέση DSC και Μέση IoU αντίστοιχα, ενώ το μοντέλο U-Net επιτυγχάνει τιμές 0,6587 και 0,4921 για τις ίδιες μετρικές.

Όσον αφορά τη Μέση Αξιοπιστία, τα μοντέλα U-Net, U-Net + VGG16 και U-Net + DenseNet121 παρουσιάζουν τιμές 0,8321 0,8190 και 0,8071 αντίστοιχα. Η υψηλή μέση τιμή Αξιοπιστίας των τριών μοντέλων δείχνει πως και τα τρία μοντέλα είναι ικανά να πραγματοποιούν με μεγάλη ακρίβεια σημασιολογική κατάτμηση αλλοιώσεων σε μία τομή.

Όσον αφορά τη Μέση Ανάκληση, τα μοντέλα U-Net, U-Net + VGG16 και U-Net + DenseNet121 παρουσιάζουν τιμές 0,5487 0,6305 και 0,6566 αντίστοιχα. Η μέτρια μέση τιμή Ανάκλησης των τριών μοντέλων δείχνει πως και τα τρία μοντέλα είναι ικανά να πραγματοποιούν σημασιολογική κατάτμηση ενός μέτριου ποσοστού των αλλοιώσεων σε μία τομή.

#### **4.6.1.4 Παραλλαγή 4: Κατάτμηση Πνευμονικού Παρεγχύματος – Διαχωρισμός Δεδομένων ανά Ασθενή**

Στον Πίνακα 4.10 παρουσιάζονται οι μετρικές αξιολόγησης της ικανότητας των μοντέλων, έπειτα από εκπαίδευση, να πραγματοποιούν σημασιολογική κατάτμηση αλλοιώσεων από πνευμονία COVID-19 στις τομές του συνόλου δεδομένων επικύρωσης της Παραλλαγής 4.

**Πίνακας 4.10: Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Ασθενή: Μετρικές αξιολόγησης της εφαρμογής σημασιολογικής κατάτμησης αλλοιώσεων στα δεδομένα επικύρωσης**

Μοντέλο	Μέση Dice Similarity Coefficient	Μέση Intersection over Union	Μέση Αξιοπιστία	Μέση Ανάκληση
<b>U-Net</b>	0,7637 ± 0,0678	0,6216 ± 0,0883	0,7394 ± 0,0894	0,7912 ± 0,0442
<b>U-Net + VGG16</b>	0,7669 ± 0,0629	0,6253 ± 0,0830	0,7591 ± 0,0675	0,7759 ± 0,0673
<b>U-Net + DenseNet121</b>	0,7889 ± 0,0541	0,6540 ± 0,0738	0,8219 ± 0,0521	0,7603 ± 0,0706

Παρατηρώντας τις μετρικές αξιολόγησης στον Πίνακα 4.10, διαπιστώνουμε πως και τα τρία μοντέλα παρουσιάζουν πολύ καλή ικανότητα σημασιολογικής κατάτμησης αλλοιώσεων σε τομές αξονικών τομογραφιών.

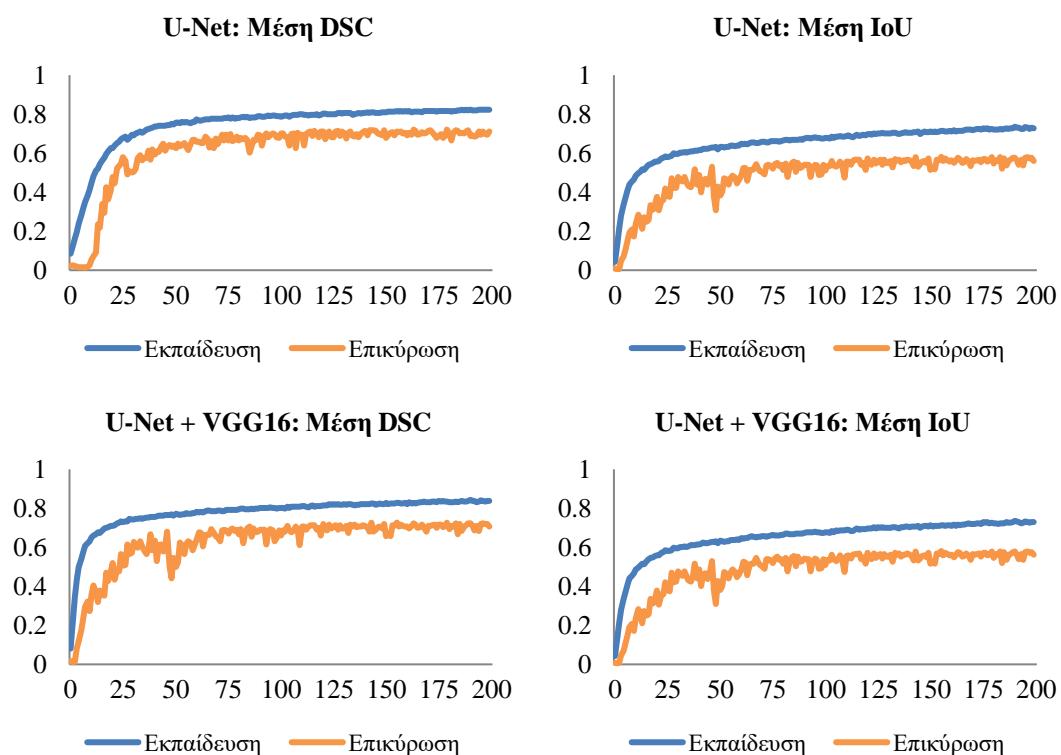
Όλα τα μοντέλα επιτυγχάνουν παρόμοια επίδοση, ωστόσο την υψηλότερη μέση τιμή όσον αφορά τη Μέση DSC και Μέση IoU επιτυγχάνει το μοντέλο U-Net + DenseNet121. Πιο αναλυτικά, το μοντέλο U-Net + DenseNet121 επιτυγχάνει τιμές 0,7889 και 0,6540 όσον αφορά τη Μέση DSC και Μέση IoU αντίστοιχα.

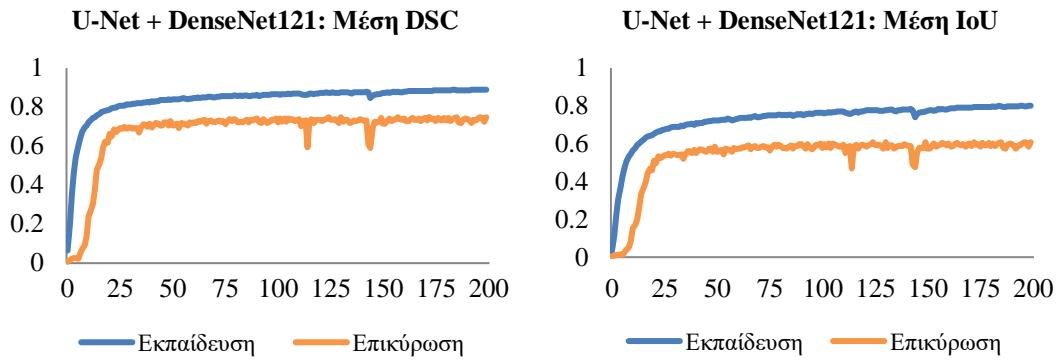
Τα άλλα δύο μοντέλα, δηλαδή το μοντέλο U-Net και το μοντέλο U-Net + VGG16 παρουσιάζουν χαμηλότερη επίδοση, με το μοντέλο U-Net + VGG16 να εμφανίζει ελαφρώς καλύτερα αποτελέσματα. Πιο συγκεκριμένα το μοντέλο U-Net + VGG16 επιτυγχάνει τιμές 0,7669 και 0,6253 όσον αφορά τη Μέση DSC και Μέση IoU αντίστοιχα, ενώ το μοντέλο U-Net επιτυγχάνει τιμές 0,7637 και 0,6216 για τις ίδιες μετρικές.

Όσον αφορά τη Μέση Αξιοπιστία, τα μοντέλα U-Net, U-Net + VGG16 και U-Net + DenseNet121 παρουσιάζουν τιμές 0,7394, 0,7591 και 0,8219 αντίστοιχα. Η υψηλή μέση τιμή Αξιοπιστίας των τριών μοντέλων δείχνει πως και τα τρία μοντέλα είναι ικανά να πραγματοποιούν με μεγάλη ακρίβεια σημασιολογική κατάτμηση αλλοιώσεων σε μία τομή.

Όσον αφορά τη Μέση Ανάκληση, τα μοντέλα U-Net, U-Net + VGG16 και U-Net + DenseNet121 παρουσιάζουν τιμές 0,7912, 0,7759 και 0,7603 αντίστοιχα. Η υψηλή μέση τιμή Ανάκλησης των τριών μοντέλων δείχνει πως και τα τρία μοντέλα είναι ικανά να ανιχνεύουν και να πραγματοποιούν σημασιολογική κατάτμηση ενός μεγάλου ποσοστού των αλλοιώσεων σε μία τομή.

Στην Εικόνα 4.12 παρουσιάζονται οι καμπύλες εκπαίδευσης και επικύρωσης των μοντέλων όσον αφορά τη σημασιολογική κατάτμηση αλλοιώσεων από πνευμονία COVID-19 στα δεδομένα της Παραλλαγής 4.





**Εικόνα 4.12: Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Ασθενή: Καμπύλες εκπαίδευσης και επικύρωσης των μοντέλων όσον αφορά τη σημασιολογική κατάτμηση αλλοιώσεων**

Παρατηρώντας τις καμπύλες εκπαίδευσης και επικύρωσης στην Εικόνα 4.12, διαπιστώνουμε πως όλα τα μοντέλα που εκπαιδεύθηκαν και αξιολογήθηκαν στα σύνολα εκπαίδευσης και επικύρωσης αντίστοιχα, συγκλίνουν μετά το πέρας της εκπαιδευτικής διαδικασίας. Ωστόσο, υπάρχει μία απόσταση ανάμεσα στη καμπύλη εκπαίδευσης και στη καμπύλη επικύρωσης τόσο στη γραφική παράσταση Μέσης DSC όσο και στη γραφική παράσταση Μέσης IoU, η οποία όμως παραμένει σταθερή καθ' όλη τη διάρκεια της εκπαιδευτικής διαδικασίας για όλα τα μοντέλα. Η απόσταση αυτή αποτελεί ένδειξη πως υπάρχει υπερπροσαρμογή των μοντέλων στα δεδομένα εκπαίδευσης και επίσης υποδηλώνει πως τα μοντέλα ενδεχομένως να μην διαθέτουν μεγάλη ικανότητα γενίκευσης.

Στον Πίνακα 4.11 παρουσιάζονται οι μετρικές αξιολόγησης της ικανότητας των μοντέλων, έπειτα από εκπαίδευση, να πραγματοποιούν σημασιολογική κατάτμηση αλλοιώσεων από πνευμονία COVID-19 στις τομές του συνόλου δεδομένων ελέγχου της Παραλλαγής 4.

**Πίνακας 4.11: Κατάτμηση Πνευμονικού Παρεγχύματος - Διαχωρισμός Δεδομένων ανά Ασθενή: Μετρικές αξιολόγησης της εφαρμογής σημασιολογικής κατάτμησης αλλοιώσεων στα δεδομένα ελέγχου**

Μοντέλο	Μέση Dice Similarity Coefficient	Μέση Intersection over Union	Μέση Αξιοπιστία	Μέση Ανάκληση
<b>U-Net</b>	0,6613 ± 0,0372	0,4940 ± 0,0404	0,7302 ± 0,1209	0,6392 ± 0,1431
<b>U-Net + VGG16</b>	0,7022 ± 0,0217	0,5414 ± 0,0258	0,7592 ± 0,0974	0,6654 ± 0,0639
<b>U-Net + DenseNet121</b>	0,7079 ± 0,0123	0,5480 ± 0,0146	0,8037 ± 0,0558	0,6369 ± 0,0466

Παρατηρώντας τις μετρικές αξιολόγησης στον Πίνακα 4.11, διαπιστώνουμε πως και τα τρία μοντέλα παρουσιάζουν καλή ικανότητα σημασιολογικής κατάτμησης αλλοιώσεων σε τομές αξονικών τομογραφιών, ωστόσο υπάρχει περιθώριο βελτίωσης.

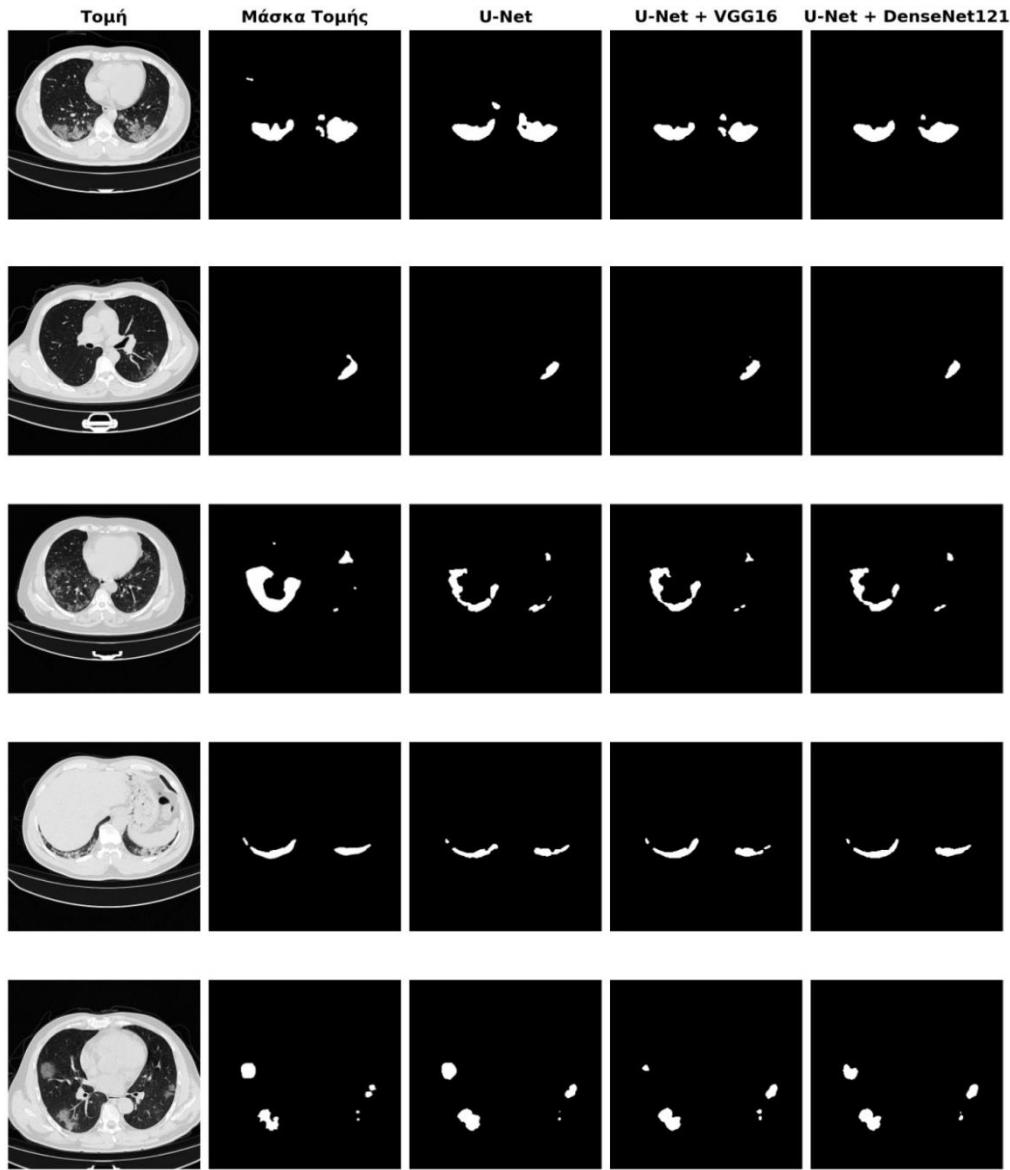
Από τα τρία μοντέλα, τα μοντέλα U-Net + DenseNet121 και U-Net + VGG16 είναι αυτά που παρουσιάζουν τη καλύτερη επίδοση. Αναλυτικότερα, το μοντέλο U-Net + DenseNet121 επιτυγχάνει τιμές 0,7079 και 0,5480 όσον αφορά τη Μέση DSC και Μέση IoU αντίστοιχα. Το μοντέλο U-Net + VGG16 επιτυγχάνει τιμές 0,7022 και 0,5414 όσον αφορά τη Μέση DSC και Μέση IoU αντίστοιχα, ενώ το μοντέλο U-Net επιτυγχάνει τιμές 0,6613 και 0,4940 για τις ίδιες μετρικές.

Όσον αφορά τη Μέση Αξιοπιστία, τα μοντέλα U-Net, U-Net + VGG16 και U-Net + DenseNet121 παρουσιάζουν τιμές 0,7302, 0,7592 και 0,8037 αντίστοιχα. Η υψηλή μέση τιμή Αξιοπιστίας των τριών μοντέλων δείχνει πως και τα τρία μοντέλα είναι ικανά να πραγματοποιούν με μεγάλη ακρίβεια σημασιολογική κατάτμηση αλλοιώσεων σε μία τομή.

Όσον αφορά τη Μέση Ανάκληση, τα μοντέλα U-Net, U-Net + VGG16 και U-Net + DenseNet121 παρουσιάζουν τιμές 0,6392, 0,6654 και 0,6369 αντίστοιχα. Η μέτρια μέση τιμή Ανάκλησης των τριών μοντέλων δείχνει πως και τα τρία μοντέλα είναι ικανά να πραγματοποιούν σημασιολογική κατάτμηση ενός μέτριου ποσοστού των αλλοιώσεων σε μία τομή.

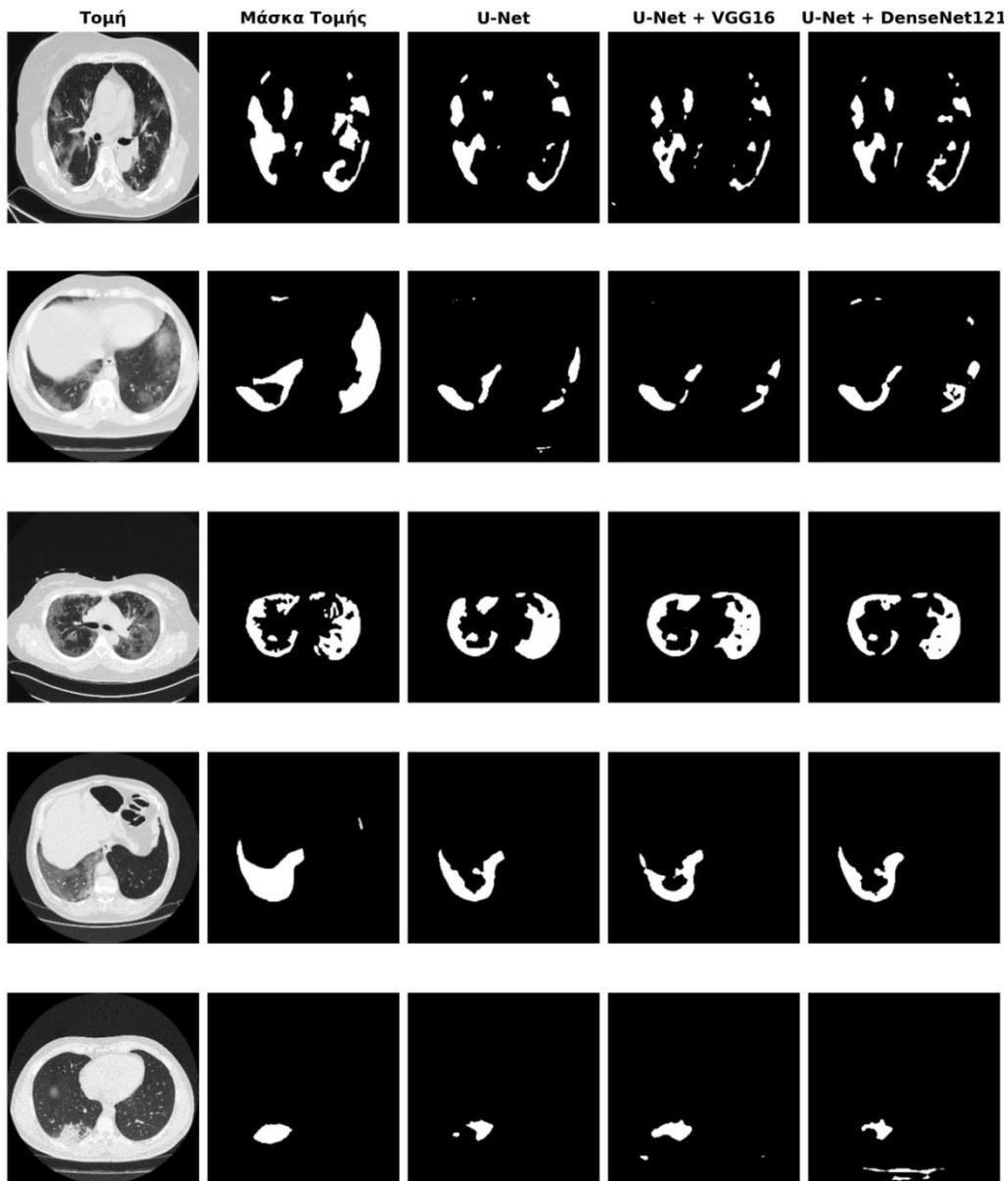
#### **4.6.2 Εφαρμογή της σημασιολογικής κατάτμησης αλλοιώσεων σε τομές των συνόλων επικύρωσης και ελέγχου**

Στην Εικόνα 4.13 παρουσιάζονται τα αποτελέσματα της εφαρμογής σημασιολογικής κατάτμησης αλλοιώσεων οφειλόμενων στον ιό SARS-CoV-2, από τα εκπαιδευμένα μοντέλα, σε ορισμένες τομές αξονικών τομογραφιών του συνόλου επικύρωσης, όπου δεν έχει εφαρμοσθεί κατάτμηση πνευμονικού παρεγχύματος. Παρατηρώντας τη συγκεκριμένη εικόνα, διαπιστώνουμε πως όλα τα μοντέλα ανιχνεύουν και απομονώνουν τις αλλοιώσεις των τομών αξονικών τομογραφιών σε πολύ ικανοποιητικό βαθμό.



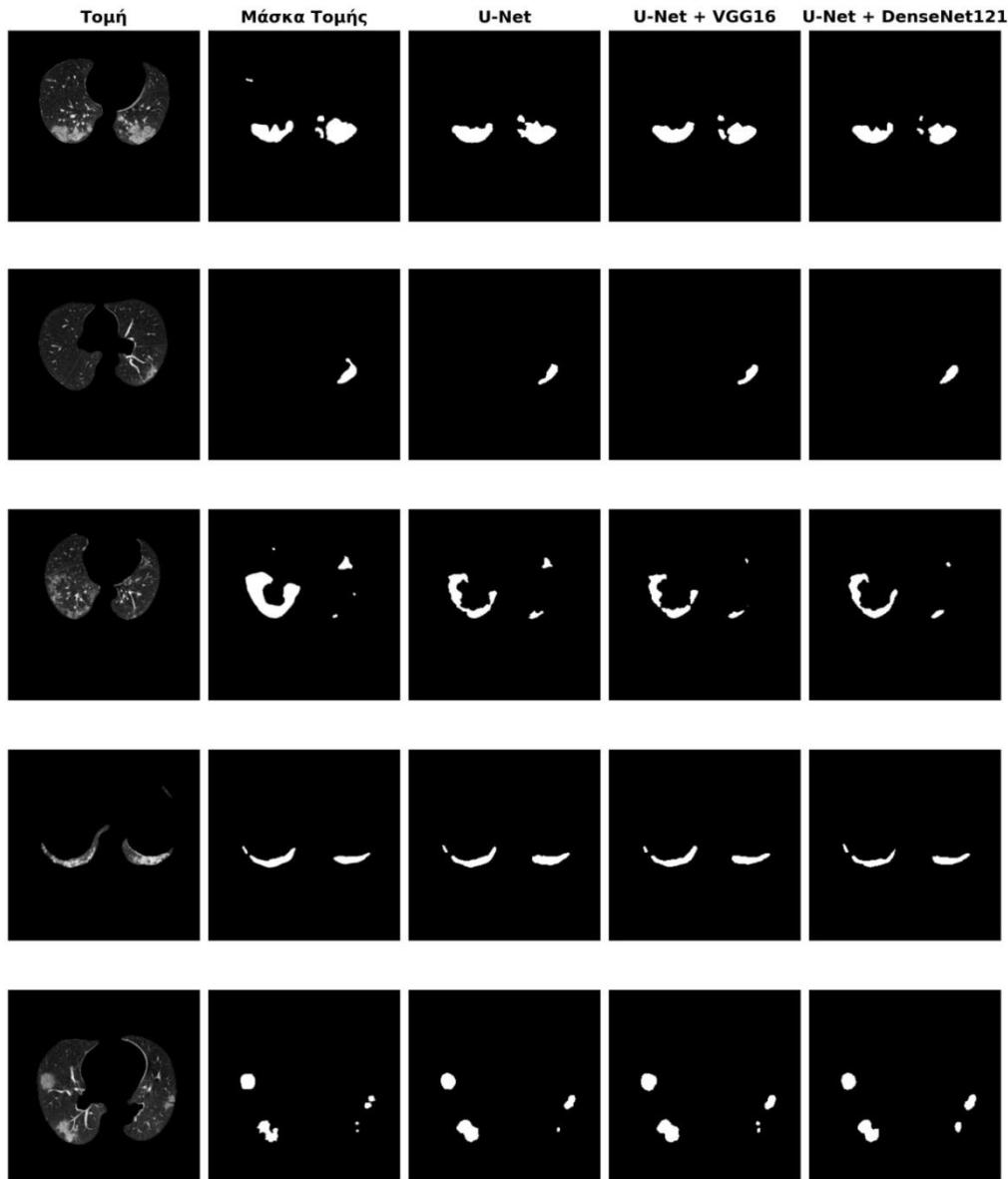
Εικόνα 4.13: Εφαρμογή σημασιολογικής κατάτμησης αλλοιώσεων σε τομές του συνόλου επικύρωσης όπου δεν έχει εφαρμοσθεί κατάτμηση πνευμονικού παρεγχύματος

Στην Εικόνα 4.14 παρουσιάζονται τα αποτελέσματα της εφαρμογής σημασιολογικής κατάτμησης αλλοιώσεων οφειλόμενων στον ιό SARS-CoV-2, από τα εκπαιδευμένα μοντέλα, σε ορισμένες τομές αξονικών τομογραφιών του συνόλου ελέγχου, όπου δεν έχει εφαρμοσθεί κατάτμηση πνευμονικού παρεγχύματος. Παρατηρώντας τη συγκεκριμένη εικόνα, διαπιστώνουμε πως τα μοντέλα ανιχνεύουν και απομονώνουν τις αλλοιώσεις των τομών στις περισσότερες περιπτώσεις σε ικανοποιητικό βαθμό. Ωστόσο, σε ορισμένες περιπτώσεις τα μοντέλα απομονώνουν και άσχετα στοιχεία τα οποία βρίσκονται εκτός του πνευμονικού παρεγχύματος.



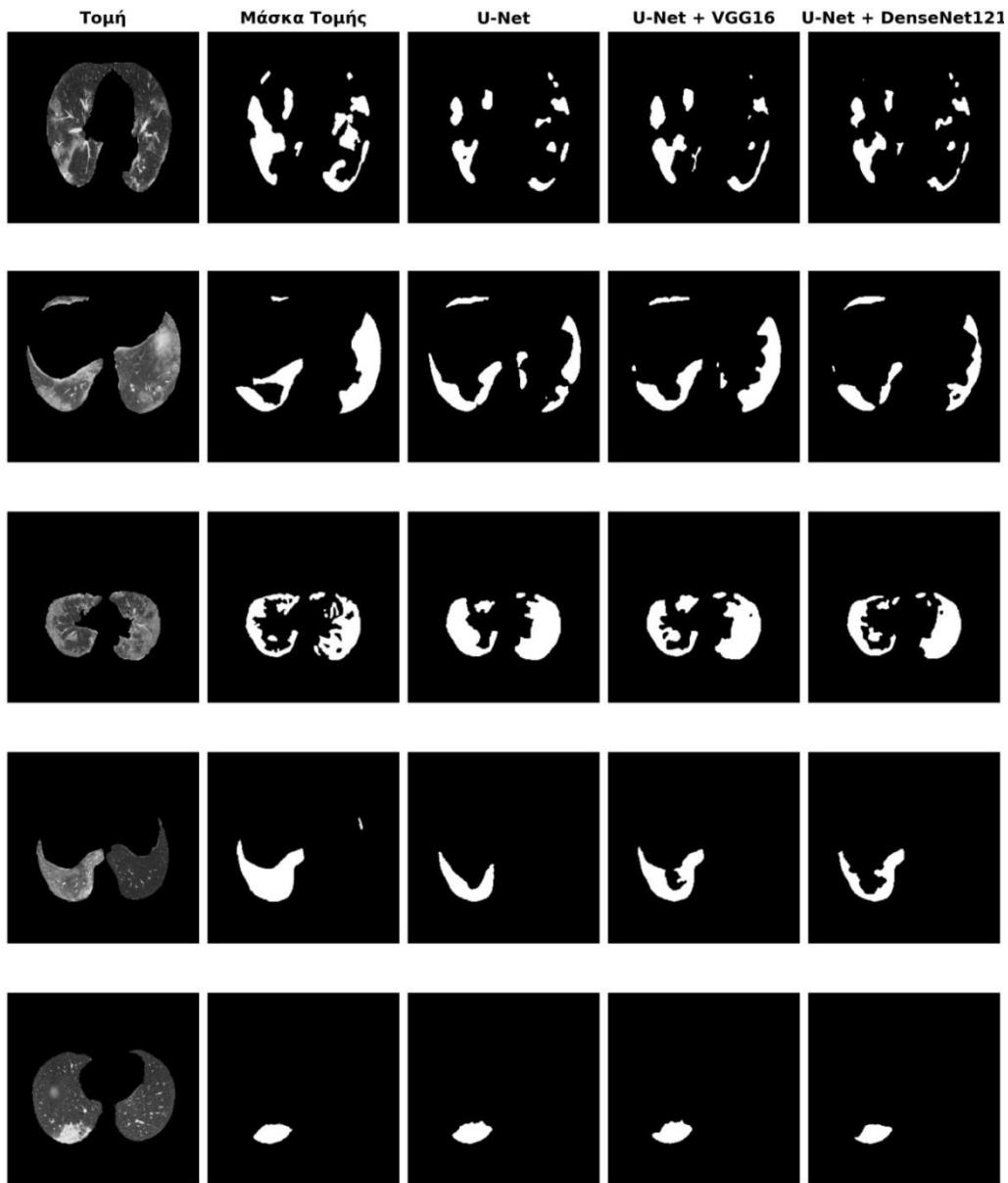
Εικόνα 4.14: Εφαρμογή σημασιολογικής κατάτμησης αλλοιώσεων σε τομές του συνόλου ελέγχου όπου δεν έχει εφαρμοσθεί κατάτμηση πνευμονικού παρεγχύματος

Στην Εικόνα 4.15 παρουσιάζονται τα αποτελέσματα της εφαρμογής σημασιολογικής κατάτμησης αλλοιώσεων οφειλόμενων στον ιό SARS-CoV-2, από τα εκπαιδευμένα μοντέλα, σε ορισμένες τομές αξονικών τομογραφιών του συνόλου επικύρωσης, όπου έχει εφαρμοσθεί κατάτμηση πνευμονικού παρεγχύματος. Παρατηρώντας τις συγκεκριμένες εικόνες, διαπιστώνουμε πως όλα τα μοντέλα ανιχνεύουν και απομονώνουν τις αλλοιώσεις των τομών σε πολύ ικανοποιητικό βαθμό.



**Εικόνα 4.15:** Εφαρμογή σημασιολογικής κατάτμησης αλλοιώσεων σε τομές του συνόλου επικύρωσης όπου έχει εφαρμοσθεί κατάτμηση του πνευμονικού παρεγχύματος

Στην Εικόνα 4.16 παρουσιάζονται τα αποτελέσματα της εφαρμογής σημασιολογικής κατάτμησης αλλοιώσεων οφειλόμενων στον ιό SARS-CoV-2, από τα εκπαιδευμένα μοντέλα, σε ορισμένες τομές αξονικών τομογραφιών του συνόλου ελέγχου, όπου έχει εφαρμοσθεί κατάτμηση πνευμονικού παρεγχύματος.  
Παρατηρώντας τις συγκεκριμένες εικόνες, διαπιστώνουμε πως όλα τα μοντέλα ανιχνεύουν και απομονώνουν τις αλλοιώσεις των τομών στις περισσότερες περιπτώσεις σε πολύ ικανοποιητικό βαθμό εφόσον έχουν αφαιρεθεί στοιχεία εκτός του πνευμονικού παρεγχύματος τα οποία αποπροσανατολίζουν τα μοντέλα βαθιάς μάθησης.



**Εικόνα 4.16:** Εφαρμογή σημασιολογικής κατάτμησης αλλοιώσεων σε τομές του συνόλου ελέγχου όπου έχει εφαρμοσθεί κατάτμηση του πνευμονικού παρεγχύματος

#### 4.6.3 Σύγκριση αποτελεσμάτων των τεσσάρων παραλλαγών

Παρατηρώντας την επίδοση των μοντέλων ως προς την εφαρμογή σημασιολογικής κατάτμησης αλλοιώσεων στα δεδομένα επικύρωσης και ελέγχου των τεσσάρων παραλλαγών του συνόλου δεδομένων, μπορούν να διαπιστωθούν τα εξής:

Στις περιπτώσεις όπου εφαρμόζεται η τεχνική της μεταφοράς μάθησης, δηλαδή όπου ο κωδικοποιητής του μοντέλου U-Net αντικαθίσταται από ένα προεκπαιδευμένο μοντέλο, το U-Net παρουσιάζει πολύ καλύτερη επίδοση.

Από τα τρία μοντέλα που εκπαιδεύθηκαν, το U-Net + DenseNet121 είναι αυτό που παρουσιάζει την καλύτερη επίδοση. Πιο συγκεκριμένα, επιτυγχάνει την υψηλότερη Μέση DSC και Μέση IoU τόσο στα σύνολα δεδομένων επικύρωσης όσο και στα ελέγχου και των τεστάρων παραλλαγών. Αυτό αποτελεί ένδειξη πως το U-Net + DenseNet121 πραγματοποιεί σημασιολογική κατάτμηση αλλοιώσεων στις τομές των συνόλων δεδομένων καλύτερα και επιδεικνύει μεγαλύτερη ικανότητα γενίκευσης από τα άλλα δύο μοντέλα.

Στις περιπτώσεις όπου πραγματοποιείται διαχωρισμός ανά τομή, το μοντέλο U-Net + VGG16 επιτυγχάνει σχεδόν ίδια επίδοση με το μοντέλο U-Net + DenseNet121, όταν αυτά αξιολογούνται στο εξωτερικό σύνολο αξιολόγησης.

Συγκρίνοντας την επίδοση των μοντέλων στα δεδομένα επικύρωσης με αυτή στα δεδομένα ελέγχου, διαπιστώνεται πως οι τιμές των μετρικών αξιολόγησης στα πρώτα είναι εμφανώς μεγαλύτερες από αυτές στα δεύτερα. Αυτό είναι αναμενόμενο διότι το σύνολο δεδομένων ελέγχου αποτελείται από δεδομένα που προέρχονται από ένα εξωτερικό σύνολο αξιολόγησης, δηλαδή ένα σύνολο το οποίο δεν έχει χρησιμοποιηθεί ούτε στη φάση εκπαίδευσης αλλά ούτε και στη φάση επικύρωσης της εκπαιδευτικής διαδικασίας. Επομένως, είναι λογικό τα μοντέλα να μην έχουν τόσο καλή επίδοση στο συγκεκριμένο σύνολο δεδομένων το οποίο τους είναι εντελώς «άγνωστο». Επίσης, το σύνολο επικύρωσης χρησιμοποιείται κατά την εκπαιδευτική διαδικασία για τη ρύθμιση των παραμέτρων των δικτύων (tuning) πράγμα που ενδεχομένως να οδηγεί τα μοντέλα στο να «μαθαίνουν» χαρακτηριστικά του παρότι δεν έχουν εκπαιδευθεί στο συγκεκριμένο σύνολο και άρα να αποδίδουν καλύτερα σε αυτό.

Τα μοντέλα παρουσιάζουν παρόμοια επίδοση στο σύνολο επικύρωσης είτε έχει πραγματοποιηθεί κατάτμηση πνευμονικού παρεγχύματος, είτε όχι. Ωστόσο στο σύνολο ελέγχου, παρατηρείται πως τα μοντέλα στην περίπτωση κατάτμησης πνευμονικού παρεγχύματος παρουσιάζουν πολύ καλύτερη επίδοση συγκριτικά με τα αντίστοιχα μοντέλα στην περίπτωση μη κατάτμησης πνευμονικού παρεγχύματος. Αυτό οφείλεται στο γεγονός ότι στην πρώτη περίπτωση δεν εμφανίζονται στοιχεία που βρίσκονται εκτός του πνευμονικού παρεγχύματος και τα οποία δυσκολεύουν την ορθή σημασιολογική κατάτμηση των τομών αξονικών τομογραφιών.

Όσον αφορά τις καμπύλες εκπαίδευσης και επικύρωσης, παρατηρείται πως η απόσταση ανάμεσά τους είναι μικρότερη στις περιπτώσεις όπου πραγματοποιείται διαχωρισμός ανά τομή σε σχέση με εκείνες όπου πραγματοποιείται διαχωρισμός ανά ασθενή. Αυτό μπορεί να εξηγηθεί εν μέρει από το φαινόμενο της διαρροής δεδομένων που αναφέρθηκε στο Κεφάλαιο 3. Επιπρόσθετα στην περίπτωση διαχωρισμού ανά τομή τα δεδομένα είναι περισσότερο ποικίλα με αποτέλεσμα τα μοντέλα να γενικεύουν καλύτερα, όταν εκπαιδεύονται σε δεδομένα που διαχωρίζονται με αυτόν τον τρόπο. Αντίθετα, στην περίπτωση διαχωρισμού ανά ασθενή, λόγω του μικρού αριθμού των ασθενών, τα δεδομένα του συνόλου εκπαίδευσης πιθανόν δεν είναι

αρκετά ώστε το εκάστοτε εκπαιδευόμενο μοντέλο να μάθει τα μοτίβα και τις σχέσεις μεταξύ τους.

Όσον αφορά τις δύο μεθόδους διαχωρισμού των δεδομένων, παρατηρείται πως οι τιμές των μετρικών αξιολόγησης στην περίπτωση όπου πραγματοποιείται διαχωρισμός ανά τομή είναι υψηλότερες σε σχέση με τη περίπτωση όπου πραγματοποιείται διαχωρισμός ανά ασθενή. Οι μεγαλύτερες τιμές των μετρικών αξιολόγησης στην πρώτη περίπτωση ενδεχομένως να οφείλονται στο φαινόμενο της διαρροής δεδομένων το οποίο οδηγεί σε ψευδώς υπερβολικά υψηλή επίδοση των μοντέλων. Επιπλέον, το γεγονός ότι τα δεδομένα είναι περισσότερο ποικίλα όταν διαχωρίζονται ανά τομή πιθανώς να οδηγούν τα μοντέλα στο να αποδίδουν καλύτερα στο σύνολο επικύρωσης.

Οι τιμές των μετρικών αξιολόγησης της διαδικασίας επικύρωσης πλησιάζουν περισσότερο τις αντίστοιχες τιμές των μετρικών της διαδικασίας ελέγχου στην περίπτωση όπου πραγματοποιείται διαχωρισμός ανά ασθενή. Αυτό δείχνει πως στις περιπτώσεις όπου πραγματοποιείται διαχωρισμός ανά ασθενή η αξιολόγηση στο σύνολο επικύρωσης παρέχει μια καλύτερη εικόνα της ικανότητας γενίκευσης των μοντέλων.

Συγκρίνοντας τις εικόνες που παράχθηκαν έπειτα από την εφαρμογή σημασιολογικής κατάτμησης αλλοιώσεων στις τομές των συνόλων επικύρωσης και ελέγχου, διαπιστώνουμε πως όσον αφορά το σύνολο επικύρωσης, τόσο στη περίπτωση κατάτμησης όσο και στη περίπτωση μη κατάτμησης πνευμονικού παρεγχύματος, τα μοντέλα απομονώνουν τις αλλοιώσεις με μεγάλη επιτυχία. Από την άλλη, στο σύνολο ελέγχου, φαίνεται πως στις περιπτώσεις όπου δεν έχει εφαρμοσθεί κατάτμηση πνευμονικού παρεγχύματος τα μοντέλα απομονώνουν και στοιχεία που δεν αποτελούν αλλοιώσεις.

## 5 Επίλογος

Στο συγκεκριμένο κεφάλαιο παρουσιάζονται τα τελικά συμπεράσματα από τη μεθοδολογία που ακολούθησε στην παρούσα διπλωματική εργασία, καθώς και οι επεκτάσεις που θα μπορούσαν να πραγματοποιηθούν μελλοντικά.

### 5.1 Τελικά Συμπεράσματα

Στόχος της παρούσας διπλωματικής εργασίας ήταν να διερευνηθεί η ικανότητα μεθόδων βαθιάς μάθησης να ανιχνεύουν πνευμονία οφειλόμενη στον ιό SARS-CoV-2 σε τομές αξονικών τομογραφιών, με απότερο στόχο την υποβοήθηση στη διάγνωση της λοίμωξης COVID-19. Στα πλαίσια εκπόνησής της, αξιολογήθηκε μία πληθώρα αρχιτεκτονικών τεχνητών νευρωνικών δικτύων ως προς την ικανότητά τους, αφού εκπαιδευθούν κατάλληλα, να είναι σε θέση να πραγματοποιήσουν ταξινόμηση και σημασιολογική κατάτμηση τομών αξονικών τομογραφιών. Από τις τιμές των μετρικών αξιολόγησης, καθώς και από τα αποτελέσματα εφαρμογής των μοντέλων σε δεδομένα των συνόλων, γίνεται αντιληπτό πως οι στόχοι της διπλωματικής εργασίας επιτεύχθηκαν και πως τα αποτελέσματα και των δύο πειραματικών διαδικασιών είναι πολύ ικανοποιητικά.

Από τα πειράματα που υλοποιήθηκαν προκύπτουν τα εξής συμπεράσματα:

Η επιλογή κατάλληλων δεδομένων αποτελεί μία από τις βασικότερες συνιστώσες επιτυχούς διεξαγωγής ενός πειράματος βαθιάς μάθησης. Όσον αφορά τη παρούσα διπλωματική εργασία, η επιλογή να χρησιμοποιηθούν τομές αξονικών τομογραφιών έναντι ακτινογραφιών θώρακος, ήταν ένας από τους λόγους που οδήγησε στην επιτυχή εκπαίδευση των μοντέλων.

Η χρήση μορφότυπου DICOM ή Nifti για την αποθήκευση των τομών των αξονικών τομογραφιών προσφέρει επιπλέον δυνατότητες προεπεξεργασίας των δεδομένων. Πιο συγκεκριμένα, επιτρέπει την κανονικοποίηση των τομών ως προς την αντίθεση και τη φωτεινότητα, μέσω της μεθόδου της παραθυροποίησης και επιπλέον βοηθάει στην κατάτμηση του πνευμονικού παρεγγύματος στις τομές.

Η εφαρμογή κατάλληλης προεπεξεργασίας στα δεδομένα αποτελεί κρίσιμο παράγοντα για την επιτυχή εκπαίδευση ενός μοντέλου βαθιάς μάθησης. Ειδικότερα, η κανονικοποίηση των τιμών αντίθεσης και φωτεινότητας στις τομές και η κατάτμηση του πνευμονικού παρεγγύματος οδήγησαν στο να επιτευχθούν υψηλές τιμές όσον αφορά τις μετρικές αξιολόγησης.

Η τεχνική της μεταφοράς μάθησης, οδηγεί τα μοντέλα στο να επιτύχουν πολύ καλύτερες επιδόσεις, ιδίως στις περιπτώσεις όπου το σύνολο δεδομένων είναι

περιορισμένο. Αυτό στη παρούσα εργασία γίνεται ιδιαίτερα αντιληπτό στο πείραμα σημασιολογικής κατάτμησης αλλοιώσεων, όπου το σύνολο δεδομένων που χρησιμοποιείται για την εκπαίδευση, επικύρωση και έλεγχο των μοντέλων αποτελείται από έναν περιορισμένο αριθμό τομών αξονικών τομογραφιών.

Η χρήση της τεχνικής προσαύξησης δεδομένων (data augmentation) μέσω εφαρμογής γεωμετρικών μετασχηματισμών στα ήδη υπάρχοντα δεδομένα, αποτελεί κατάλληλη μέθοδο για τον περιορισμό του φαινομένου της υπερπροσαρμογής στα δεδομένα εκπαίδευσης.

Ο τρόπος διαχωρισμού των δεδομένων είναι σημαντικός για τη σωστή διεξαγωγή των πειραμάτων και πρέπει να γίνει με ιδιαίτερη προσοχή διότι κάθε μέθοδος διαχωρισμού παρουσιάζει πλεονεκτήματα και μειονεκτήματα. Στο παρόν πόνημα, ο διαχωρισμός των δεδομένων ανά τομή δημιουργεί δεδομένα τα οποία είναι περισσότερο ποικίλα και άρα επιτρέπει στα μοντέλα να γενικεύουν καλύτερα. Ωστόσο, οδηγεί σε μοντέλα με ψευδώς υπερβολικά υψηλή επίδοση. Αυτό αποδίδεται, όπως έχει προαναφερθεί, στη διαρροή δεδομένων. Από την άλλη, όταν η εκπαίδευση και η επικύρωση γίνεται σε δεδομένα που έχουν διαχωριστεί ανά ασθενή, τα αποτελέσματα προσεγγίζουν περισσότερο την επίδοση των μοντέλων στα δεδομένα ελέγχου. Όμως, σε αυτή τη περίπτωση τα δεδομένα είναι λιγότερο ποικίλα, και οδηγούν σε μοντέλα που εμφανίζουν μεγαλύτερη υπερπροσαρμογή στα δεδομένα εκπαίδευσης.

Η χρήση της τεχνικής Grad-CAM για τις ανάγκες του προβλήματος ταξινόμησης, καθώς και η εμφάνιση ορισμένων από τις τομές στις οποίες πραγματοποιήσαμε σημασιολογική κατάτμηση των αλλοιώσεων, έπαιξε καθοριστικό ρόλο, αφού μας παρείχε τη δυνατότητα να αξιολογήσουμε εάν τα μοντέλα πραγματοποιούν σωστά τη διαδικασία για την οποία έχουν εκπαιδευθεί.

Σε αντίθεση με τη περίπτωση ταξινόμησης τομών στις κατηγορίες Covid-19 και Normal όπου οι μετρικές αξιολόγησης τόσο στο σύνολο επικύρωσης όσο και στο σύνολο ελέγχου είναι πολύ υψηλές, στη περίπτωση ταξινόμησης τομών στις κατηγορίες Covid-19 και Other Pneumonia, οι μετρικές αξιολόγησης είναι πολύ χαμηλές όταν τα μοντέλα αξιολογούνται στο σύνολο ελέγχου. Αυτό σημαίνει ότι τα μοντέλα δεν μπορούν να διακρίνουν την πνευμονία από COVID-19 από πνευμονίες άλλης αιτιολογίας και ενδεχομένως να οφείλεται στο ότι το σύνολο δεδομένων είναι μικρό καθώς και στο ότι η διάκριση της COVID-19 πνευμονίας από άλλες πνευμονίες είναι πιο απαιτητική.

## 5.2 Μελλοντικές Επεκτάσεις

Από τα αποτελέσματα της εκπαίδευσης και αξιολόγησης των μοντέλων ως προς την ικανότητά τους να επιλύουν τα προβλήματα που πραγματεύεται η παρούσα διπλωματική εργασία, προκύπτει το συμπέρασμα ότι η μεθοδολογία που ακολουθήθηκε οδήγησε στη δημιουργία μοντέλων τα οποία έχουν εκπαιδευθεί σωστά και πραγματοποιούν την λειτουργία για την οποία έχουν εκπαιδευθεί σε πολύ ικανοποιητικό βαθμό. Η μεθοδολογία της εργασίας θα μπορούσε να επεκταθεί μελλοντικά με τον ακόλουθο τρόπο:

Να συλλεχθούν περισσότερα δεδομένα. Αυτό, θα εξυπηρετούσε στο να πραγματοποιηθεί περαιτέρω εκπαίδευση των μοντέλων, τα δεδομένα να ποικίλουν ως προς την ηλικία, το φύλο και την εθνικότητα των ασθενών, καθώς επίσης και να έχουν ληφθεί από ποικιλία νοσηλευτικών ιδρυμάτων. Η συλλογή περισσοτέρων δεδομένων, θα εξυπηρετήσει ώστε κατά την εκπαίδευση των μοντέλων να υπάρχει ισοκατανομή τους στις κλάσεις των δεδομένων.

Οι τομές των αξονικών τομογραφιών που θα χρησιμοποιηθούν, να είναι αποθηκευμένες σε κατάλληλο μορφότυπο όπως DICOM ή Nifti. Κάτι τέτοιο είναι απαραίτητο για να εφαρμοσθεί κανονικοποίηση μέσω εφαρμογής παραθυροποίησης και κατάτμηση του πνευμονικού παρεγχύματος, όπως έχει αναφερθεί στο Κεφάλαιο 3.

Η ρύθμιση των υπερπαραμέτρων (hyperparameter tuning) να πραγματοποιηθεί αυτοματοποιημένα με τη βοήθεια μεθόδων όπως η Grid Search και η Random Search, που έχουν υψηλές απαιτήσεις σε υπολογιστικούς πόρους, και γι' αυτό δεν μπόρεσαν να εφαρμοσθούν στην παρούσα μελέτη.

Η χρήση περισσότερης μνήμης GPU θα επέτρεπε την εκπαίδευση με μέγεθος παρτίδας (batch size) μεγαλύτερο του 32. Τα πειράματα που πραγματοποιήθηκαν για το συγκεκριμένο ζήτημα έδειξαν ότι σε ορισμένες περιπτώσεις, μέγεθος παρτίδας μεγαλύτερο του 32 αύξανε την επίδοση των εκπαιδευόμενων μοντέλων.

Η εφαρμογή κανονικοποίησης (regularization), μέσω των μεθόδων Lasso Regression (L1 Regularization) και Ridge Regression (L2 Regularization), θα μπορούσε να ελαττώσει το φαινόμενο της υπερπροσαρμογής και κατ' επέκταση να βελτιώσει την ικανότητα γενίκευσης των μοντέλων.

## Βιβλιογραφία

- [1] “Naming the coronavirus disease (COVID-19) and the virus that causes it.” [https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-\(covid-2019\)-and-the-virus-that-causes-it](https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(covid-2019)-and-the-virus-that-causes-it) (accessed Jul. 21, 2023).
- [2] M. Adnan Shereen, S. Khan, A. Kazmi, N. Bashir, and R. Siddique, “COVID-19 infection: origin, transmission, and characteristics of human coronaviruses,” *Journal of Advanced Research*, vol. 24, pp. 91–98, Mar. 2020, doi: <https://doi.org/10.1016/j.jare.2020.03.005>.
- [3] T. Singhal, “A review of coronavirus disease-2019 (COVID-19),” *The Indian Journal of Pediatrics*, vol. 87, no. 4, pp. 281–286, Mar. 2020, doi: <https://doi.org/10.1007/s12098-020-03263-6>.
- [4] D. Cucinotta and M. Vanelli, “WHO Declares COVID-19 a Pandemic,” *Acta Bio-medica: Atenei Parmensis*, vol. 91, no. 1, pp. 157–160, Mar. 2020, doi: <https://doi.org/10.23750/abm.v91i1.9397>.
- [5] “WHO Coronavirus (COVID-19) dashboard,” *WHO Coronavirus (COVID-19) Dashboard With Vaccination Data*. <https://covid19.who.int/> (accessed Jul. 21, 2023).
- [6] Y.-J. Chen *et al.*, “Earlier diagnosis improves COVID-19 prognosis: a nationwide retrospective cohort analysis,” *Annals of Translational Medicine*, vol. 9, no. 11, pp. 941–941, Jun. 2021, doi: <https://doi.org/10.21037/atm-20-7210>.
- [7] B. Udugama *et al.*, “Diagnosing COVID-19: The Disease and Tools for Detection,” *ACS Nano*, vol. 14, no. 4, Mar. 2020, doi: <https://doi.org/10.1021/acsnano.0c02624>.
- [8] S. Machnicki *et al.*, “The Usefulness of Chest CT Imaging in Patients With Suspected or Diagnosed COVID-19: A Review of Literature,” *Chest*, vol. 160, no. 2, pp. 652–670, Aug. 2021, doi: <https://doi.org/10.1016/j.chest.2021.04.004>.
- [9] A. Borakati, A. Perera, J. Johnson, and T. Sood, “Diagnostic accuracy of X-ray versus CT in COVID-19: a propensity-matched database study,” *BMJ Open*, vol. 10, no. 11, p. e042946, Nov. 2020, doi: <https://doi.org/10.1136/bmjopen-2020-042946>.
- [10] A. Pinto, “Spectrum of diagnostic errors in radiology,” *World Journal of Radiology*, vol. 2, no. 10, p. 377, 2010, doi: <https://doi.org/10.4329/wjr.v2.i10.377>.
- [11] L. Wang *et al.*, “Artificial Intelligence for COVID-19: A Systematic Review,” *Frontiers in Medicine*, vol. 8, Sep. 2021, doi: <https://doi.org/10.3389/fmed.2021.704256>.
- [12] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, “ImageNet: A large-scale hierarchical image database,” *2009 IEEE Conference on Computer Vision and Pattern Recognition*, Miami, FL, USA, 2009, pp. 248-255, doi: <https://doi.org/10.1109/cvpr.2009.5206848>.

- [13] R. Davidson, “Physics and Principles of Computed Tomography,” in *Computed Tomography: A Primer for Radiographers*, S. Chau and C. M. Hayre, 1<sup>st</sup> edition, Boca Raton, FL, USA: CRC Press, 2022, doi: <https://doi.org/10.1201/9781003132554>.
- [14] E. Seeram, *Computed tomography: Physical principles, clinical applications, and quality control*, 4th ed. St. Louis, MO, USA: Elsevier, 2016.
- [15] “Computerized Tomography (CT) Scan | GI Update.” <https://www.giupdate.com/ct-scan> (accessed Jul. 21, 2023).
- [16] P. Afshar *et al.*, “COVID-CT-MD, COVID-19 computed tomography scan dataset applicable in machine learning and deep learning,” *Scientific Data*, vol. 8, no. 1, Apr. 2021, doi: <https://doi.org/10.1038/s41597-021-00900-3>.
- [17] L. E. Romans, *Computed tomography for technologists: a comprehensive text*. Philadelphia, PA, USA: Wolters Kluwer, 2011.
- [18] L. Hazel, “Introduction to CT Image Interpretation,” in *Computed Tomography: A Primer for Radiographers*, Shayne Chau and Christopher M. Hayre, 1<sup>st</sup> edition, Boca Raton, FL, USA: CRC Press, 2022, doi: <https://doi.org/10.1201/9781003132554>.
- [19] M. Elgendy, *Deep learning for vision systems*, Shelter Island, NY, USA: Manning Publications Co, 2020.
- [20] S. Minaee, Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz and D. Terzopoulos, “Image Segmentation Using Deep Learning: A Survey,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 7, pp. 3523-3542, Jul. 2022, doi: <https://doi.org/10.1109/TPAMI.2021.3059968>.
- [21] A. L. Samuel, “Some Studies in Machine Learning Using the Game of Checkers,” *IBM Journal of Research and Development*, vol. 3, no. 3, pp. 210-229, Jul. 1959, doi: <https://doi.org/10.1147/rd.33.0210>.
- [22] T. Mitchell, *Machine Learning*, NY, USA: McGraw Hill, 1997.
- [23] I. J. Goodfellow, Y. Bengio and A. Courville, *Deep Learning*, Cambridge, MA, USA: MIT Press, 2016.
- [24] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 4<sup>th</sup> ed. Upper Saddle River, NJ, USA: Prentice Hall, 2020.
- [25] B. Mahesh, “Machine Learning Algorithms – A Review,” *International Journal of Science and Research (IJSR)*, vol. 9, no. 1, pp. 381-386, Jan. 2020, doi: <http://dx.doi.org/10.21275/ART20203995>.
- [26] L. P. Kaelbling, M. L. Littman and A. W. Moore. “Reinforcement Learning: A Survey,” *Journal of Artificial Intelligence Research*, vol. 4, no. 1, pp. 237-285, May 1996, doi: <https://doi.org/10.1613/jair.301>.
- [27] O. Chapelle, B. Schölkopf and A. Zien, *Semi-Supervised Learning*, Cambridge, MA, USA: The MIT Press, 2006, doi: <https://10.7551/mitpress/9780262033589.001.0001>.
- [28] F. Chollet, *Deep Learning with Python*, NY, USA: Manning Publications, 2017, pp. 94-95.

- [29] G. Varoquaux and O. Colliot, “Evaluating Machine Learning Models and Their Diagnostic Value,” *Neuromethods*, vol. 197, pp. 601-630, Feb. 2012, doi: [https://doi.org/10.1007/978-1-0716-3195-9\\_20](https://doi.org/10.1007/978-1-0716-3195-9_20).
- [30] “Overfitting,” *MATLAB & Simulink*.  
<https://www.mathworks.com/discovery/overfitting.html> (accessed Jul. 21, 2023).
- [31] D. Berrar, “Cross-Validation,” *Encyclopedia of Bioinformatics and Computational Biology*, vol. 1, pp. 542-545, 2019, doi: <https://doi.org/10.1016/B978-0-12-809633-8.20349-X>.
- [32] C.M. Bishop, *Pattern Recognition and Machine Learning*, NY, USA: Springer, 2006.
- [33] “3.1. Cross-validation: evaluating estimator performance,” *Scikit-learn*.  
[https://scikit-learn.org/stable/modules/cross\\_validation.html](https://scikit-learn.org/stable/modules/cross_validation.html)
- [34] Y. LeCun, Y. Bengio and G. Hinton, “Deep Learning,” *Nature*, vol. 521, no. 7553, pp. 436-444, May 2015, doi: <https://doi.org/10.1038/nature14539>.
- [35] R. S. Snell, *Clinical Neuroanatomy*, 7th ed. Philadelphia, PA, USA: Lippincott Williams & Wilkins, 2010.
- [36] *File:Neuron.svg - Wikimedia Commons*. 2019. [Online]. Available: <https://commons.wikimedia.org/wiki/File:Neuron.svg#/media/File:Neuron.svg> (accessed Jul. 21, 2023).
- [37] S. Haykin, *Neural Networks and Learning Machines*, 3<sup>rd</sup> ed. Upper Saddle River, NJ, USA: Pearson Education, 2009.
- [38] F. Rosenblatt, “The perceptron: A probabilistic model for information storage and organization in the brain,” *Psychological Review*, vol. 65, no. 6, pp. 386-408, Jan. 1958, doi: <https://doi.org/10.1037/h0042519>.
- [39] S. Sharma, S. Sharma and A. Athaiya, “ACTIVATION FUNCTIONS IN NEURAL NETWORKS,” *International Journal of Engineering Applied Sciences and Technology*, vol. 4, no. 12, pp. 310-316, May 2020, doi: <https://doi.org/10.33564/ijeast.2020.v04i12.054>.
- [40] C. E. Nwankpa, W. Ijomah, A. Gachagan and S. Marshall, “Activation Functions: Comparison of Trends in Practice and Research for Deep Learning,” *2nd International Conference on Computational Sciences and Technology*, Jamshoro, Pakistan, 2021, pp. 124-133, doi: <https://doi.org/10.48550/arXiv.1811.03378>.
- [41] D. P. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization,” *arXiv (Cornell University)*, Dec. 2014, doi: <https://doi.org/10.48550/arxiv.1412.6980>.
- [42] D. E. Rumelhart, G. E. Hinton and R. J. Williams, “Learning representations by back-propagating errors,” *Nature*, vol. 323, no. 6088, pp. 533-536, Oct. 1986, doi: <https://doi.org/10.1038/323533a0>.
- [43] V. H. Phung and E. J. Rhee, “A High-Accuracy Model average ensemble of convolutional neural networks for classification of cloud image patches on small datasets,” *Applied Sciences*, vol. 9, no. 21, p. 4500, Oct. 2019, doi: <https://doi.org/10.3390/app9214500>.

- [44] C. Shorten and T. M. Khoshgoftaar, “A survey on Image Data Augmentation for Deep Learning,” *Journal of Big Data*, vol. 6, no. 1, Jul. 2019, doi: <https://doi.org/10.1186/s40537-019-0197-0>.
- [45] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever and R. Salakhutdinov, “Dropout: A Simple Way to Prevent Neural Networks from Overfitting,” *Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929-1958, Jun. 2014.
- [46] S. Ioffe and C. Szegedy, “Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift,” *arXiv (Cornell University)*, Feb. 2015, doi: <https://doi.org/10.48550/arXiv.1502.03167>.
- [47] K. Zhang *et al.*, “Clinically applicable AI system for accurate diagnosis, quantitative measurements, and prognosis of COVID-19 pneumonia using computed tomography,” *Cell*, vol. 181, no. 6, pp. 1423-1433.e11, Jun. 2020, doi: <https://doi.org/10.1016/j.cell.2020.04.045>.
- [48] J. Hofmanninger, F. Prayer, J. Pan, S. Röhrich, H. Prosch and G. Langs, “Automatic lung segmentation in routine imaging is primarily a data diversity problem, not a methodology problem,” *European Radiology Experimental*, vol. 4, no. 1, Aug. 2020, doi: <https://doi.org/10.1186/s41747-020-00173-2>.
- [49] J. T. Hancock and T. M. Khoshgoftaar, “Survey on categorical data for neural networks,” *Journal of Big Data*, vol. 7, no. 1, Apr. 2020, doi: <https://doi.org/10.1186/s40537-020-00305-w>.
- [50] E. Yagis *et al.*, “Effect of data leakage in brain MRI classification using 2D convolutional neural networks,” *Scientific Reports*, vol. 11, Nov. 2021, doi: <https://doi.org/10.1038/s41598-021-01681-w>.
- [51] J. Wen *et al.*, “Convolutional neural networks for classification of Alzheimer’s disease: Overview and reproducible evaluation,” *Medical Image Analysis*, vol. 63, p. 101694, Jul. 2020, doi: <https://doi.org/10.1016/j.media.2020.101694>.
- [52] C. Szegedy, V. Vanhoucke, S. Ioffe and J. Shlens and Z. Wojna, “Rethinking the Inception Architecture for Computer Vision,” *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016, pp. 2818-2826, doi: <https://doi.org/10.1109/CVPR.2016.308>.
- [53] K. He, X. Zhang, S. Ren and J. Sun, “Deep Residual Learning for Image Recognition,” *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016, pp. 770-778, doi: <https://doi.org/10.1109/CVPR.2016.90>.
- [54] G. Huang, Z. Liu, L. van der Maaten and K. Q. Weinberger, “Densely Connected Convolutional Networks,” *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, 2017, pp. 2261-2269, doi: <https://doi.org/10.1109/CVPR.2017.243>.
- [55] F. Chollet, “Xception: Deep Learning with Depthwise Separable Convolutions,” *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, 2017, pp. 1800-1807, doi: <https://doi.org/10.1109/CVPR.2017.195>.

- [56] F. Chollet *et al.*, “Home - Keras Documentation,” *Keras.io*, 2019. <https://keras.io>
- [57] C. Szegedy *et al.*, “Going Deeper with Convolutions,” *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, 2015, pp. 1-9, doi: <https://doi.org/10.1109/CVPR.2015.7298594>.
- [58] S. Basodi, C. Ji, H. Zhang and Y. Pan, “Gradient Amplification: An efficient way to train deep neural networks,” *Big Data Mining and Analytics*, vol. 3, no. 3, pp. 196-207, Sept. 2020, doi: <https://dl.acm.org/10.26599/BDMA.2020.9020004>.
- [59] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” *International Conference on Learning Representations 2015 (ICLR 2015)*, San Diego, CA, USA, 2014, doi: <https://doi.org/10.48550/arXiv.1409.1556>.
- [60] A. Tharwat, “Classification assessment methods,” *Applied Computing and Informatics*, vol. 17, no. 1, pp. 168–192, Jul. 2020, doi: <https://doi.org/10.1016/j.aci.2018.08.003>.
- [61] M. Sokolova and G. Lapalme, “A systematic analysis of performance measures for classification tasks,” *Information Processing & Management*, vol. 45, no. 4, pp. 427-437, Jul. 2009, doi: <https://doi.org/10.1016/j.ipm.2009.03.002>.
- [62] F. Xu, H. Uszkoreit, Y. Du, W. Fan, D. Zhao and J. Zhu, “Explainable AI: A Brief Survey on History, Research Areas, Approaches and Challenges,” *Natural Language Processing and Chinese Computing (NLPCC)*, Dunhuang, China, 2019, pp. 563-574, doi: [https://doi.org/10.1007/978-3-030-32236-6\\_51](https://doi.org/10.1007/978-3-030-32236-6_51).
- [63] 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,” *2017 IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, 2017, pp. 618-626, doi: <https://doi.org/10.1109/ICCV.2017.74>.
- [64] H. Alshazly, C. Linse, E. Barth and T. Martinetz, “Explainable COVID-19 Detection Using Chest CT Scans and Deep Learning,” *Sensors*, Jan 2021, vol. 21, no. 2, pp. 455, doi: <https://doi.org/10.3390/s21020455>.
- [65] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally and K. Keutzer, “SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size,” *International Conference on Learning Representations, ICLR 2017*, Toulon, France, 2017, doi: <https://doi.org/10.48550/arXiv.1602.07360>.
- [66] X. Zhang, X. Zhou, M. Lin and J. Sun, “ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices,” *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, 2018, pp. 6848-6856, doi: <https://doi.org/10.1109/CVPR.2018.00716>.
- [67] E. Soares, P. Angelov, S. Biaso, M. Higa 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, Apr. 2020,  
<https://doi.org/10.1101/2020.04.24.20078584>.
- [68] X. He *et al.*, “Sample-Efficient Deep Learning for COVID-19 Diagnosis Based on CT Scans,” *MedRxiv*, Apr. 2020, doi:  
<https://doi.org/10.1101/2020.04.13.20063941>.
- [69] H. Ulutas, M. E. Sahin and M. O. Karakus, “Application of a novel deep learning technique using CT images for COVID-19 diagnosis on embedded systems,” *Alexandria Engineering Journal*, vol. 74, pp. 345-358, Jul. 2023, doi: <https://doi.org/10.1016/j.aej.2023.05.036>.
- [70] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov and L.-C. Chen, “MobileNetV2: Inverted Residuals and Linear Bottlenecks,” *2018 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018, Salt Lake City, UT, USA, pp. 4510-4520, doi:  
<https://doi.org/10.48550/arXiv.1801.04381>.
- [71] H. Gunraj, L. Wang and A. Wong, “COVIDNet-CT: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases From Chest CT Images,” *Frontiers in Medicine*, vol. 7, Dec. 2020, doi:  
<https://doi.org/10.3389/fmed.2020.608525>.
- [72] H. Gunraj, A. Sabri, D. Koff and A. Wong, “COVID-Net CT-2: Enhanced Deep Neural Networks for Detection of COVID-19 From Chest CT Images Through Bigger, More Diverse Learning,” *Frontiers in Medicine*, vol. 8, Mar. 2022, doi: <https://doi.org/10.3389/fmed.2021.729287>.
- [73] S. Chaudhary, Sadbhawna, V. Jakhetiya, B. N Subudhi, U. Baid and S. C. Guntuku, “Detecting COVID-19 and Community Acquired Pneumonia using Chest CT scan images with Deep Learning,” *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Toronto, ON, Canada, 2021, pp. 8583-8587, doi:  
<https://doi.org/10.1109/ICASSP39728.2021.9414007>.
- [74] M. Tan and Q. V. Le, “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks,” *Proceedings of the 36th International Conference on Machine Learning, PMLR*, Long Beach, CA, USA, 2019, pp. 6105-6114, doi: <https://doi.org/10.48550/arXiv.1905.11946>.
- [75] M. Rahimzadeh, A. Attar and S. M. Sakhaei, “A fully automated deep learning-based network for detecting COVID-19 from a new and large lung CT scan dataset,” *Biomed Signal Process Control*, Jul. 2021, vol. 68, doi:  
<https://doi.org/10.1016/j.bspc.2021.102588>.
- [76] T.-Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan and S. Belongie, “Feature Pyramid Networks for Object Detection,” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, 2017, pp. 2117-2125, doi:  
<https://doi.org/10.48550/arXiv.1612.03144>.
- [77] O. Attalah and A. Samir, “A wavelet-based deep learning pipeline for efficient COVID-19 diagnosis via CT slices,” *Applied Soft Computing*, vol. 128, Oct. 2022, doi: <https://doi.org/10.1016/j.asoc.2022.109401>.

- [78] A. Vaidyanathan *et al.*, “An externally validated fully automated deep learning algorithm to classify COVID-19 and other pneumonias on chest CT,” *ERJ Open Research*, vol. 9, no. 5, doi: <https://doi.org/10.1183/23120541.00579-2021>.
- [79] H. Ko *et al.*, “COVID-19 Pneumonia Diagnosis Using a Simple 2D Deep Learning Framework With a Single Chest CT Image: Model Development and Validation,” *Journal of Medical Internet Research*, Jun. 2020, vol. 22, no. 6, doi: <https://doi.org/10.2196/19569>.
- [80] D. Nguyen *et al.*, “Deep Learning-Based COVID-19 Pneumonia Classification Using Chest CT Images: Model Generalizability,” *Frontiers in Artificial Intelligence*, vol. 4, Jun. 2021, doi: <https://doi.org/10.3389/frai.2021.694875>.
- [81] M. Fallahpoor, S. Chakraborty, M. T. Heshejin, H. Chegeni, M. J. Horry and B. Pradhan, “Generalizability assessment of COVID-19 3D CT data for deep learning-based disease detection,” *Computers in Biology and Medicine*, vol. 145, no. 18, doi: <https://doi.org/10.1016/j.combiomed.2022.105464>.
- [82] “Kaggle: your machine learning and data science community.” <https://www.kaggle.com/> (accessed Feb. 19, 2023).
- [83] J. Ma *et al.*, “Toward data-efficient learning: A benchmark for COVID-19 CT lung and infection segmentation,” *Medical Physics*, vol. 48, no. 3, pp. 1197–1210, Feb. 2021, doi: <https://doi.org/10.1002/mp.14676>.
- [84] O. A. Paiva, 2020, “CORONACASES.ORG – by RAI OSS.com.” *Coronacases*. <https://coronacases.org/> (accessed Feb. 19, 2023).
- [85] Y. Glick, “Playlist ‘COVID-19 pneumonia’ by Dr Yaïr Glick,” *Radiopaedia*. <https://radiopaedia.org/playlists/25887> (accessed Feb. 19, 2023).
- [86] O. Ronneberger, P. Fischer and T. Brox, “U-Net: Convolutional networks for biomedical image segmentation”, *Proc. 18th Int. Conf. Med. Image Comput. Comput.-Assist. Interv.*, pp. 234-241, 2015, doi: [https://doi.org/10.1007/978-3-319-24574-4\\_28](https://doi.org/10.1007/978-3-319-24574-4_28).
- [87] P. Iakubovskii, “GitHub - qubvel/segmentation\_models: Segmentation models with pretrained backbones. Keras and TensorFlow Keras.,” *GitHub*. [https://github.com/qubvel/segmentation\\_models](https://github.com/qubvel/segmentation_models)
- [88] A. A. Taha and A. Hanbury, “Metrics for evaluating 3D medical image segmentation: analysis, selection and tool,” *BMC Medical Imaging*, vol. 15, no. 1, Aug. 2015, doi: <https://doi.org/10.1186/s12880-015-0068-x>.
- [89] K. H. Zou *et al.*, “Statistical Validation of Image Segmentation Quality Based on a Spatial Overlap Index,” *Academic Radiology*, vol. 11, no. 2, pp. 178-189, Feb. 2004, doi: [https://doi.org/10.1016/s1076-6332\(03\)00671-8](https://doi.org/10.1016/s1076-6332(03)00671-8).
- [90] H. Rezatofighi, N. Tsoi, J. Y. Gwak, A. Sadeghian, I. Reid and S. Savarese, “Generalized Intersection over Union: A Metric and A Loss for Bounding Box Regression,” *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Long Beach, CA, USA, 2019, pp. 658-666, doi: <https://doi.org/10.1109/CVPR.2019.00075>.
- [91] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, O. Ronneberger, “3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation,”

- Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016, Athens, Greece, 2016, pp. 424-432, doi: [https://doi.org/10.1007/978-3-319-46723-8\\_49](https://doi.org/10.1007/978-3-319-46723-8_49).
- [92] S. P. Morozov *et al.*, “MosMedData: data set of 1110 chest CT scans performed during the COVID-19 epidemic,” *Digital Diagnostics*, vol. 1, no. 1, pp. 49-59, Dec. 2020, doi: <https://doi.org/10.17816/DD46826>
  - [93] D. Müller, I. S. Rey and F. Kramer, “Automated Chest CT Image Segmentation of COVID-19 Lung Infection based on 3D U-Net,” *arXiv*, Jun. 2020, <https://doi.org/10.48550/arXiv.1606.06650>.
  - [94] D. Müller, I. S. Rey and F. Kramer, “Robust chest CT image segmentation of COVID-19 lung infection based on limited data,” *Informatics in Medical Unlocked*, vol. 25, doi: <https://doi.org/10.1016/j.imu.2021.100681>.
  - [95] A. L. Aswathy and V. C. S. S., “Cascaded 3D UNet architecture for segmenting the COVID-19 infection from lung CT volume,” *Scientific Reports*, vol. 12, no. 1, pp. 3090, Feb. 2022, doi: <https://doi.org/10.1038/s41598-022-06931-z>.
  - [96] D.-P. Fan *et al.*, “Inf-Net: Automatic COVID-19 Lung Infection Segmentation From CT Images,” *IEEE Transactions on Medical Imaging*, vol. 39, no. 8, pp. 2626-2637, Aug. 2020, doi: <https://doi.org/10.1109/TMI.2020.2996645>.
  - [97] A. Voulodimos, E. Protopapadakis, I. Katsamenis, A. Doulamis and N. Doulamis, “A Few-Shot U-Net Deep Learning Model for COVID-19 Infected Area Segmentation in CT Images,” *Sensors*, vol. 21, no. 6, pp. 2215, doi: <https://doi.org/10.3390/s21062215>.
  - [98] J. Long, E. Shelhamer and T. Darrell, “Fully Convolutional Networks for Semantic Segmentation,” *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, 2015, pp. 3431-3440, doi: <https://doi.org/10.1109/CVPR.2015.7298965>.
  - [99] A. Saood and I. Hatem, “COVID-19 lung CT image segmentation using deep learning methods: U-Net versus SegNet,” *BMC Medical Imaging*, vol. 21, no. 19, Feb. 2021, doi: <https://doi.org/10.1186/s12880-020-00529-5>.
  - [100] H. Polat, “Multi-task semantic segmentation of CT images for COVID-19 infections using DeepLabV3+ based on dilated residual network,” *Physical Sciences in Engineering and Medicine*, vol. 45, no. 2, pp. 443-455, Jun. 2022, doi: <https://doi.org/10.1007/s13246-022-01110-w>.
  - [101] V. Badrinarayanan, A. Kendall and R. Cipolla, “SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 12, pp. 2481-2495, Dec. 2017, doi: <https://doi.org/10.1109/TPAMI.2016.2644615>.
  - [102] L. C. Chen, G. Papandreou, F. Schroff and H. Adam, “Rethinking Atrous Convolution for Semantic Image Segmentation”, *arXiv*, Jun. 2017, <https://doi.org/10.48550/arXiv.1706.05587>.
  - [103] N. Enshaei *et al.*, “COVID-Rate: An Automated Framework for Segmentation of COVID-19 Lesions from Chest CT Scans,” *Scientific Reports*, vol. 12, no. 1, Feb. 2022, doi: <https://doi.org/10.1038/s41598-022-06854-9>.

- [104] “COVID-19 - Medical segmentation,” *Medical Segmentation*, Jan. 05, 2021.  
<http://medicalsegmentation.com/covid19/> (accessed Sep. 04, 2023).