



ΠΑΝΕΠΙΣΤΗΜΙΟ ΜΑΚΕΔΟΝΙΑΣ  
ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ  
ΤΜΗΜΑ ΕΦΑΡΜΟΣΜΕΝΗΣ ΠΛΗΡΟΦΟΡΙΚΗΣ

ΠΡΟΒΛΕΨΗ COVID-19

Ομαδική Εργασία

Μέλη ομάδας

Εμμανουηλίδης Αθανάσιος – Μισαηλίδης Σάββας  
ΑΜ: ics21190 – ics21166

Θεσσαλονίκη, 11 Απριλίου 2024



## ΠΡΟΒΛΕΨΗ COVID

Αθανάσιος Εμμανουηλίδης - Σάββας Μισαηλίδης

Απαλλακτική Εργασία

υποβαλλόμενη για την εκπλήρωση των απαιτήσεων του

Μαθήματος: ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ  
του Τμήματος ΕΦΑΡΜΟΣΜΕΝΗΣ ΠΛΗΡΟΦΟΡΙΚΗΣ

Επιβλέπων Καθηγητής  
Ευτύχιος Πρωτοπαπαδάκης

## Περίληψη

Η παρούσα μελέτη επικεντρώνεται στην ανάλυση εικόνων με απώτερο σκοπό την πρόβλεψη COVID-19 σε ακτινογραφίες X-RAY. Πιο συγκεκριμένα, τα δεδομένα επεξεργασίας, αποτελούνται από ακτινογραφίες θώρακα, κατατασσόμενες σε κατηγορίες όπως, ασθενείς που είχαν COVID, πνευμονία, θολότητα του πνεύμονα και υγιής. Έπειτα, με την εκτεταμένη χρήση των Νευρωνικών Δικτύων, αναγνωρίζονται και αξιολογούνται αποδοτικά οι διάφορες αρχιτεκτονικές που εφαρμόστηκαν, με στόχο τη δημιουργία ενός μοντέλου που να μπορεί να προβλέπει, αν ο ασθενής ανήκει σε μία από τις παραπάνω κατηγορίες με επίκεντρο την διάγνωση COVID-19. Τέλος, με την αξιοποίηση κατάλληλων τεχνικών, έγινε η σωστή αξιολόγηση, καθώς και η σωστή μετατροπή των αρχιτεκτονικών, ώστε τα αποτελέσματα να είναι έγκυρα και να έχουν μεγάλα ποσοστά επιτυχίας.

Η ικανότητα πρόβλεψης ενός ασθενή, ο οποίος πάσχει από Covid-19 είναι ένα ζήτημα που απασχολεί τη ζωή των ανθρώπων ειδικά τα τελευταία χρόνια. Τα αποτελέσματα βασίζονται στη σύγκριση των ακτινογραφιών θώρακα μεταξύ υγιών και ασθενών. Στην παρούσα αναφορά γίνεται σύγκριση των αρχιτεκτονικών ώστε να εκτιμηθεί η απόδοση κάθε αρχιτεκτονικής αλλά και η τελική επιλογή της πιο κατάλληλης.

**Λέξεις Κλειδιά:** Deep Learning, Machine Learning, Neural Networks, Data Analysis, COVID-19 Imaging, Model Evaluation, Image Processing, Convolutional Neural Networks, Dataset Characteristics, Model Training, Train test, Test set

## **Abstract**

In the present study, image analysis is performed with the aim of predicting COVID-19 in XRAY radiographs. In the data there are images of lungs from patients who had covid, pneumonia, lung opacities or if they were healthy. It deepens the use of neural networks in the recognition and evaluation of the efficiency of various architectures, with the aim of creating a model that can perceive whether the patient belongs to one of the above categories with a focus on Covid-19. By using appropriate techniques, the correct evaluation was done, as well as the correct conversion of the architectures, so that the results are valid and have high success rates.

The ability to predict a patient suffering from Covid-19 is an issue that concerns people's lives especially in recent years. The results are based on the comparison of healthy and non-patient images. In this report, the architectures are compared in order to make an assessment of the performance of each architecture and the final choice of the most suitable one.

**Keywords:** Deep Learning, Machine Learning, Neural Networks, Data Analysis, COVID-19 Imaging, Model Evaluation, Image Processing, Convolutional Neural Networks, Dataset Characteristics, Model Training, Train test, Test set

# Περιεχόμενα

1 Εισαγωγή	1
1.1 Σχετική έρευνα	1
1.2 Συνεισφορά	4
2 ΑΡΧΙΤΕΚΤΟΝΙΚΕΣ ΚΑΙ ΜΕΘΟΔΟΙ	4
Τεχνικές Βαθιάς Μάθησης	6
2.1.1 Συνελικτικά νευρωνικά δίκτυα (CNN)	6
Προηγμένες Αρχιτεκτονικές: VGG16 και ResNet50	6
3 ΑΠΟΤΕΛΕΣΜΑΤΑ	7
3.1 Περιγραφή Dataset	7
3.2 Εφαρμογή και περιορισμοί των μοντέλων ταξινόμησης	8
3.3 Πειραματικά Αποτελέσματα	9
3.4 Οπτικοποίηση αποτελεσμάτων	13
3.5 Τελικά Συμπεράσματα	16
Βιβλιογραφία	19

## Κατάλογος Εικόνων

Εικόνα 1: Συγκριτική ανάλυση απώλειας εκπαίδευσης επικύρωσης και Ακρίβειας στις 30	
Εποχές για το CNN.....	11
Εικόνα 2: Συγκριτική ανάλυση απώλειας εκπαίδευσης επικύρωσης και Ακρίβειας στις 30	
Εποχές για το VGG16 .....	12
Εικόνα 3: Συγκριτική ανάλυση απώλειας εκπαίδευσης επικύρωσης και Ακρίβειας στις 30	
Εποχές για το ResNet50 .....	12

## Κατάλογος Πινάκων

Πίνακας 1: Μέσοι όροι αποτελεσμάτων Train.....	15
Πίνακας 2: Μέσοι όροι αποτελεσμάτων Test .....	15
Πίνακας 3: Μέσοι όροι αποτελεσμάτων Validation .....	15



# 1 Εισαγωγή

Υπάρχει μεγάλη εξέλιξη στα νευρωνικά δίκτυα ειδικά τα τελευταία χρόνια και αυτό έχει ως αποτέλεσμα να επηρεάζεται σημαντικά ο τομέας της ιατρικής. Χάρη στην απεικόνιση δεδομένων παρέχεται η δυνατότητα δημιουργίας μοντέλων για τη διάγνωση παθήσεων. Παράδειγμα αποτελεί και η εργασία αυτή που έχει ως στόχο τη διάγνωση παθήσεων των πνευμόνων. Η ικανότητα των νευρωνικών δικτύων συμπεριλαμβανομένων των Convolutional Neural Networks (CNN), του ResNet50 και του VGG16 έχει αποτελέσει κβαντικό παράγοντα στην έγκαιρη ανίχνευση ασθενειών οδηγώντας στη φροντίδα των ασθενών.

Μελέτες έχουν δείξει ότι ο ρόλος των νευρωνικών δικτύων ενισχύει στην ανίχνευση καρκίνου στον πνεύμονα, κάτι που μας επιδεικνύει με υψηλή ακρίβεια στον εντοπισμό κακοήθων όγκων μέσω από ιατρικές εικόνες [1]. Επίσης στα Συνελκτικα Νευρωνικά Δίκτυα μέσω της πολλαπλής ανάλυσης τους, έχουν αποδείξει ότι μπορούν να κάνουν κατάτμηση όζων πνεύμονα, κάτι το οποίο βοηθάει στην έγκαιρη διάγνωση του καρκίνου του πνεύμονα σε περιβάλλοντα υπολογιστικού νέφους [2]. Μια νέα προοπτική για τη λήψη κλινικών αποφάσεων και τη διαχείριση μιας νόσου προσφέρει η εφαρμογή τεχνικών νευρωνικών δικτύων [3].

Το έργο αυτό αξιοποιεί αυτές τις πρόσφατες εξελίξεις και συγκρίνει την απόδοση των αρχιτεκτονικών CNN, ResNet50 και VGG16 σε ένα σύνολο δεδομένων εικόνων για ακτίνες X-RAY. Οι ανακαλύψεις στοχεύουν στη βελτίωση των εφαρμογών των νευρωνικών δικτύων στην ιατρική απεικόνιση δίνοντας μια βάση για μελλοντική έρευνα και πρακτική ανάπτυξη στη διάγνωση παθήσεων του πνεύμονα

## 1.1 Σχετική έρευνα

Σε αυτήν την ενότητα θα γίνει εξέταση σε έρευνες που έχουν επικεντρωθεί στην εφαρμογή νευρωνικών δικτύων και τεχνικών μηχανικής μάθησης για τη διάγνωση ασθενειών από εικόνες ακτινών X.

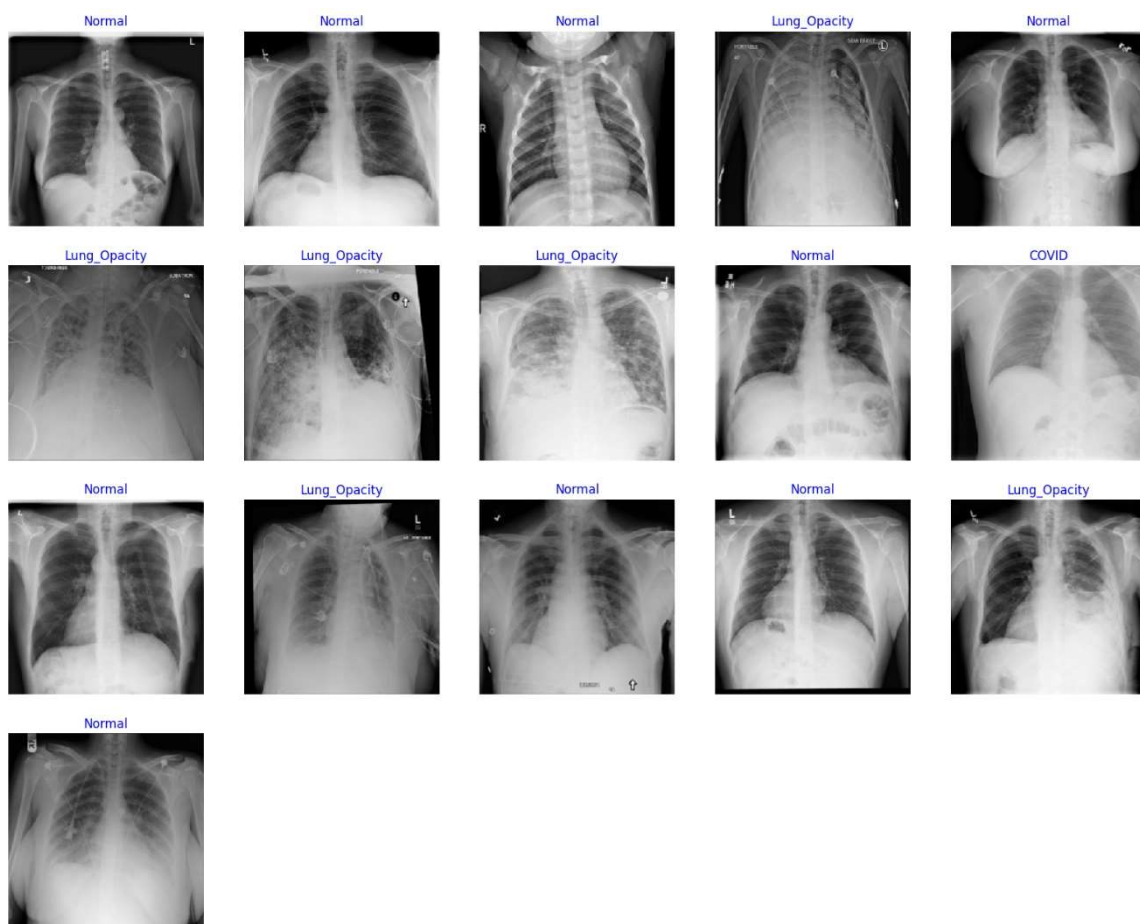
Μελέτες που έχουν διεξαχθεί από τους Wang και Dai έχουν φανερώσει ότι νευρωνικά δίκτυα πολλαπλής ανάλυσης είναι αποτελεσματικά στην τμηματοποίηση των οζιδίων του πνεύμονα. Παράλληλα αυτό δείχνει μια καινοτομική προσέγγιση η οποία ενισχύει την ακρίβεια των μοντέλων διάγνωσης βαθιάς μάθησης. Το έργο των Sheremet et al. εμφανίζει τις δυνατότητες των συνελκτικων νευρωνικών δικτύων στη διάγνωση

πνευμονικών παθήσεων. Αυτό ανοίγει τον δρόμο, κάνοντας ένα σημαντικό βήμα στην αυτοματοποίηση με στόχο τη βελτίωση της αξιοπιστίας και της ανάλυσης ιατρικής απεικόνισης.

Έχουν χρησιμοποιηθεί αρκετές μέθοδοι για να αντιμετωπιστεί η πρόκληση της αναγνώρισης της νόσου από ακτινογραφίες θώρακα. Οι Pandimeena et al. Διερεύνησαν άρθρο για τα νευρωνικά δίκτυα ώστε να ανιχνεύουν καρκίνο του πνεύμονα και να απεικονίζουν την προσαρμοστικότητα και την ακρίβεια σπονδυλωτής προσέγγισης [3]. Από την αντίθετη πλευρά υπάρχει και η μελέτη των Damayanti, Ananda και Nugraha η οποία εφάρμοσε το DenseNet μαζί με παραδοσιακές αρχιτεκτονικές του CNN για την ταξινόμηση του καρκίνου του πνεύμονα βρίσκοντας την ικανότητα να ξαναχρησιμοποιούνται χαρακτηριστικά του ώστε να βελτιωθεί η αποτελεσματικότητα και η ακρίβεια του μοντέλου [4]. Αυτές οι μεθοδολογίες τονίζουν τα υβριδικά και σύνθετα μοντέλα τα οποία έχουν ως στόχο να αξιοποιήσουν τα πλεονεκτήματα πολλαπλών αρχιτεκτονικών των νευρωνικών δικτύων κάτι το οποίο απαιτεί σημαντικούς υπολογιστικούς πόρους και πολύπλοκη προ επεξεργασία δεδομένων.

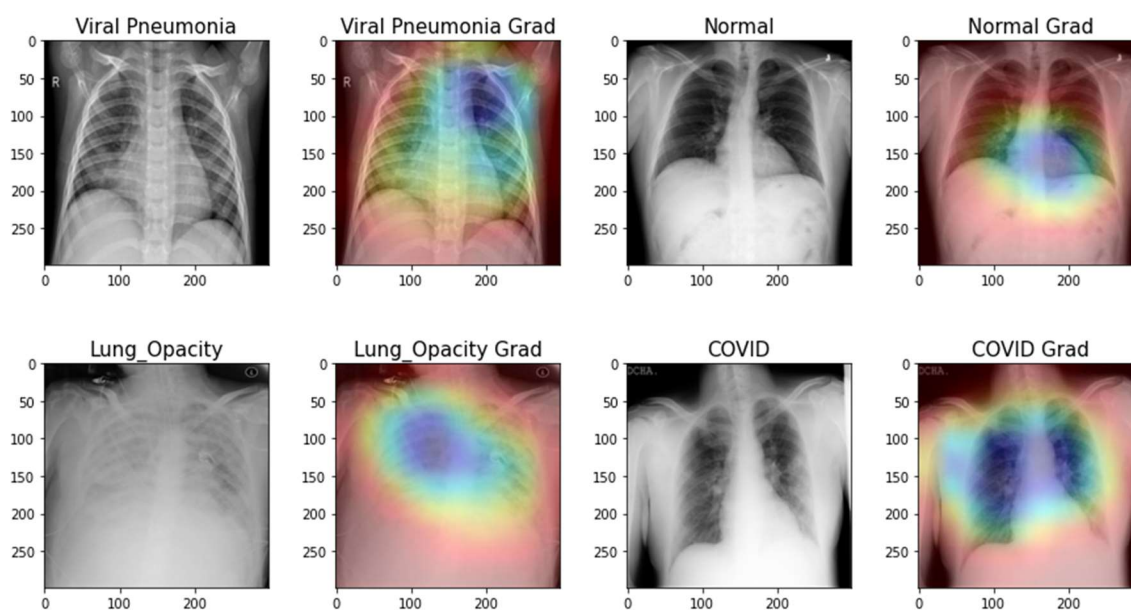
Σε όλες αυτές τις έρευνες υπήρξε ένα κενό το οποίο αποτελούσε τη δυσκολία προσαρμογής των μοντέλων που έχουν εκπαιδευτεί σε συγκεκριμένα σύνολα δεδομένων. Αυτό συμβαίνει διότι τα σύνολα δεδομένων μπορεί να διαφέρουν σημαντικά ως προς την ποιότητα της εικόνας τα δημογραφικά στοιχεία και τα πρότυπα σχολιασμού Sheshulin et al. (2021). Πολλές μελέτες όπως των Badr et al. (2022) βασίζονται σε δημόσια σύνολα δεδομένων τα οποία είναι διαθέσιμα για όλους αλλά ενδέχεται να μην περιλαμβάνουν διαφορετικούς ή ολοκληρωμένους σχολιασμούς κάτι το οποίο είναι απαραίτητο για την εκπαίδευση ισχυρών μοντέλων. Τέτοιοι περιορισμοί επηρεάζουν την ικανότητα των μοντέλων όσον αφορά τη γενίκευση την οποία αποδίδουν με συνέπεια σε διαφορετικά κλινικά περιβάλλοντα (Badr, Al-Otaibi, Alturki, & Abir, 2022).

Ένα εξίσου πολύ σημαντικό πρόβλημα αποτελεί η πολυπλοκότητα και το υπολογιστικό κόστος που χρειάστηκε για να γίνει η υλοποίηση του προβλήματος. Οι υπολογιστικές απαιτήσεις της εκπαίδευσης και της ανάπτυξης των συγκεκριμένων προηγμένων νευρωνικών δικτύων για την ανάλυση των εικόνων δεν είναι ασήμαντη. Στις έρευνες δεν είχαν επικεντρωθεί τόσο σε αυτό τον κλάδο αλλά περισσότερο στην ανάπτυξη αποδοτικών μοντέλων με ισχυρή υπολογιστική δύναμη κάτι το οποίο καθιστά τις εφαρμογές μη προσιτές στον πραγματικό κόσμο ειδικά σε περιπτώσεις όπου υπάρχει ρύθμιση για περιορισμένους πόρους.



Εικόνα 1: Εμφάνιση Δείγματος Εικόνων από το πρώτο batch του train\_gen για οπτική επιβεβαίωση των δεδομένων που εισέρχονται στο νευρωνικό δίκτυο.

## GRAD-CAM COVID-19 Image Analysis



Εικόνα 2: Υπόδειξη περιοχών στους πνεύμονες που απεικονίζουν φλεγμονή, μόλυνση ή άλλες ανωμαλίες

## 1.2 Συνεισφορά

Στην εργασία αυτή παρουσιάζεται ένα πλαίσιο βαθιάς μάθησης που ανιχνεύει ανωμαλίες σε ακτίνες θώρακα και προβλέπει σε καινούριες εικόνες την πάθηση του ασθενή. Αντίθετα με τα παραπάνω προβλήματα η παρούσα εργασία προσπαθεί με λίγους υπολογιστικούς πόρους να εξάγει αποτελέσματα με ποσοστά ακρίβειας άνω του 90% ενώ παράλληλα χρησιμοποιεί ένα μεγάλο σύνολο δεδομένων που καλύπτει πολλές περιπτώσεις. Αυτό δημιουργεί ένα αποδοτικό μοντέλο το οποίο μπορεί ανεξαρτήτως της κατάστασης της εικόνας να κάνει μια αρκετά στοχευμένη και σωστή πρόβλεψη. Για την ταχύτερη εμφάνιση αποτελεσμάτων γίνεται χρήση T4 GPU σε περιβάλλον Google Colab, ενώ υιοθετούνται μοντέλα που δεν χρειάζονται πολλούς υπολογιστικούς πόρους όπως το ResNet-50, το CNN και το VGG16.

## 2 ΑΡΧΙΤΕΚΤΟΝΙΚΕΣ ΚΑΙ ΜΕΘΟΔΟΙ

Η ανίχνευση καθώς και η διαφοροποίηση μεταξύ των διαφορετικών πνευμονικών παθήσεων, συμπεριλαμβανομένων των φυσιολογικών καταστάσεων των πνευμόνων, του

COVID-19, του lung opacity και της πνευμονίας από εικόνες ακτινών Χ αποτελεί μια πρόκληση ταξινόμησης για πολλές κατηγορίες. Σε αυτή την εργασία μπορεί να γίνει μια ισχυρή ανάλυση των χαρακτηριστικών των εικόνων για κάθε κατάσταση, κάτι που θα περιγραφεί εκτεταμένα στις παρακάτω ενότητες. Η ανάλυση αυτή θα γίνει αξιοποιώντας τεχνικές βαθιάς μάθησης για την εξαγωγή και ταξινόμηση των χαρακτηριστικών. Κάθε εικόνα υποβάλλεται σε επεξεργασία ώστε να εντοπιστεί ένα διακριτό μοτίβο το οποίο προαναφέρθηκε σε κάποια πνευμονική κατάσταση. Με τη χρήση πλαισίων βαθιάς μάθησης υπάρχει δυνατότητα να εκτελούνται ταυτόχρονα η εξαγωγή και ταξινόμηση χαρακτηριστικών χρησιμοποιώντας ένα ολοκληρωμένο σύνολο δεδομένων.

Η έρευνα περιστρέφεται γύρω από την επιλογή κατάλληλης αρχιτεκτονικής βαθιάς εκμάθησης η οποία θα ταιριάζει καλύτερα στη συγκεκριμένη εργασία: συνελκτικά νευρωνικά δίκτυα (CNN) έναντι πιο εξελιγμένων αρχιτεκτονικών όπως το ResNet50 και το VGG16. Παραδοσιακά το CNN, είναι μια διαδικασία που περιλαμβάνει την εξαγωγή χαρακτηριστικών από τα τμήματα μιας εικόνας αυτό ονομάζεται "patches" και στη συνέχεια γίνεται η ταξινόμηση κάθε τμήματος μέσω ενός εκπαιδευμένου μοντέλου. Αυτή η διαδικασία απαιτεί αρχικά μια διαίρεση της κάθε εικόνας σε πολυάριθμα επικαλυπτόμενα patches με την ταξινόμηση CNN να προσδιορίζει αν υπάρχουν ειδικοί δείκτες ή όχι σε αυτά τα τμήματα. Με αυτόν τον τρόπο ξεπερνιούνται οι παραδοσιακές τεχνικές εξαγωγής χειροκίνητων χαρακτηριστικών ενώ προσφέρει παράλληλα αποτελεσματικότητα και ακρίβεια.

Από την αντίθετη μεριά υπάρχουν προηγμένες αρχιτεκτονικές όπως οι VGG16 και ResNet50 οι οποίες προσφέρουν την ανάλυση ολόκληρων εικόνων σε ένα μοναδικό πέρασμα χωρίς την ανάγκη τμηματοποίησης. Χάρη σε αυτές τις μεθόδους ενισχύεται η ικανότητα του μοντέλου να διακρίνει περίπλοκα μοτίβα τα οποία σχετίζονται με την κάθε περίπτωση των πνευμόνων. Αυτές οι αρχιτεκτονικές περιλαμβάνουν την εκπαίδευση σε ένα σύνολο δεδομένων το οποίο έχει σχολιαστεί τοποθετώντας ανάλογα την κατάσταση κάθε συγκεκριμένης εικόνας το αντίστοιχο label που την αντιπροσωπεύει.

Κύριος στόχος είναι οι προσεγγίσεις βαθιάς μάθησης που αφορούν την αποτελεσματικότητα στην ταξινόμηση των πνευμονοπαθειών από ακτινογραφίες θώρακα. Οι τεχνικές αυτές επιδεικνύουν μια υψηλή απόδοση χρησιμοποιώντας δυνατότητες για μια ολιστική κατανόηση των δεδομένων απεικόνισης. Υπάρχει ένα μεγάλο όφελος λόγω της χρήσης των συγκεκριμένων εξελιγμένων μοντέλων το οποίο

είναι η βελτίωση της ακρίβειας, η ικανότητα αποτελεσματικής επεξεργασίας μεγάλων συνόλων δεδομένων ενώ γίνεται παράλληλα και ο χειρισμός της ανισορροπίας τάξης στα δεδομένα εκπαίδευσης. Στις παρακάτω ενότητες θα γίνει εμβάθυνση στα αποτελέσματα καθώς και ανάλυση των συμπερασμάτων, υπογραμμίζοντας τα πλεονεκτήματα και τους περιορισμούς που αντιμετωπίστηκαν κατά την εκτέλεση του έργου.

## **Τεχνικές Βαθιάς Μάθησης**

Στην προσπάθεια δημιουργίας ενός μοντέλου το οποίο θα διαγνώσει με ακρίβεια για πνευμονικές παθήσεις από εικόνες ακτίνων X, χρησιμοποιήθηκαν πολλές τεχνικές βαθιάς μάθησης. Κάθε μία από αυτές τις αρχιτεκτονικές έχει επιλεγεί διότι έχει μοναδικά πλεονεκτήματα και δυνατότητες στην ανίχνευση χαρακτηριστικών και στην ταξινόμηση εικόνων. Παρακάτω ακολουθούν τα μοντέλα βαθιάς μάθησης τα οποία συντέλεσαν αναπόσπαστο ρόλο στη μελέτη:

### **2.1.1 Συνελικτικά νευρωνικά δίκτυα (CNN)**

Η αρχιτεκτονική αυτή είναι πολύ σημαντική για την προσέγγιση αυτού του προβλήματος. Χρησιμοποιεί στρώματα συνέλιξης (Convolutional Layers) για την αποτελεσματική επεξεργασία και ανάλυση των εικόνων. Τα χαρακτηριστικά αυτά εξάγονται από εικόνες που δεν έχουν δεχτεί επεξεργασία. Η διαδικασία αυτή γίνεται μέσω πολλαπλών στρωμάτων φίλτρων καθένα από τα οποία είναι υπεύθυνο για να εντοπίζει διαφορετικές πτυχές της εικόνας άκρες, υφές ή πιο περίπλοκα μοτίβα όπως τα σχήματα των ανωμαλιών των πνευμόνων (Sarvamangala & Kulkarni, 2021).

### **Προηγμένες Αρχιτεκτονικές: VGG16 και ResNet50**

Αυτές οι τεχνικές είναι διάσημες για τα διαδοχικά συνελικτικά στρώματά τους, το VGG16 υπερέχει στην εξαγωγή όλο και πιο αφηρημένο χαρακτηριστικό από τις εικόνες ακτίνων X. Έχει υψηλά αποτελέσματα κάτι που τον καθιστά απαραίτητο για θέματα που είναι ζωτικής σημασίας προσφέροντας λύση στην ταξινόμηση πολύπλοκων παθήσεων των πνευμόνων. Η αρχιτεκτονική αυτή είναι ομοιόμορφη, ενώ επιτρέπει τη συστηματική εξερεύνηση χαρακτηριστικών σε διάφορα βάθη (Bertels & Robben, 2022).

Η αρχιτεκτονική ResNet50 χρησιμοποιεί residual learning, αντιμετωπίζοντας προβλήματα διαβαθμίσεων μέσω skip connections τα οποία επιτρέπουν στα gradients να ρέουν μέσω μιας σύντομης διαδρομής. με αυτό τον τρόπο η ικανότητα του μοντέλου ενισχύεται στο να μαθαίνει τα δεδομένα εκπαίδευσης χωρίς να μικραίνει η τιμή της

απόδοσης καθιστώντας το κατάλληλο για λεπτομερή ανάλυση πνευμονικής νόσου (Pan, 2022).

Κάθε αρχιτεκτονική έχει το δικό της ρόλο στην επεξεργασία και στην ανάλυση πνευμονικής νόσου από εικόνες. Το VGG16 προσφέρει βαθιά, ενδεδειγμένη εξαγωγή χαρακτηριστικών, το ResNet50 είναι πολλοί εύχρηστο στο χειρισμό βαθύτερων αρχιτεκτονικών δικτύων χωρίς απώλεια αποτελεσματικότητας κάτι που είναι ζωτικής σημασίας για τον εντοπισμό προβλημάτων σε ιατρικές εικόνες.

### 3 ΑΠΟΤΕΛΕΣΜΑΤΑ

Όλα τα μοντέλα έχουν γραφτεί στη γλώσσα προγραμματισμού python χρησιμοποιώντας Keras και TensorFlow libraries. Τα αποτελέσματα που θα αναλυθούν παρακάτω. Έχουν ολοκληρωθεί στο Google Colab και για την εκπαίδευση έγινε χρήση T4 GPU. Στο συγκεκριμένο πρόβλημα δημιουργήθηκαν 3 διαφορετικά περιβάλλοντα ένα για κάθε αρχιτεκτονική. Αυτό έγινε για λόγους περιορισμένης μνήμης.

#### 3.1 Περιγραφή Dataset

Το σύνολο δεδομένων επιλέχθηκε μέσω του Kaggle μίας σελίδας που παρέχει ένα εύρος από δημόσιες διαθέσιμες βάσεις δεδομένων. Η βάση αυτή που περιέχει το σύνολο των ακτινογραφιών αρχικά συγκεντρώθηκε και κοινοποιήθηκε από ερευνητές με στόχο την διευκόλυνση, για την ταχύτερη αυτοματοποίηση διάγνωσης του COVID-19 από αυτές τις εικόνες. Οι εικόνες αυτές είναι ειδικά σχεδιασμένες ώστε να βοηθήσουν στην σωστή ανίχνευση και μελέτη του COVID-19, της πνευμονίας (υικικής και βακτηριακής) και των φυσιολογικών περιπτώσεων.

Το σύνολο των δεδομένων περιλαμβάνει συνολικά 21.165 εικόνες ακτινογραφιών θώρακα. Οι εικόνες αυτές χωρίζονται σε τέσσερις κατηγορίες: COVID-19, πνευμονία, πνευμονική αδιαφάνεια (λοίμωξη των πνευμόνων χωρίς COVID) και φυσιολογικές καταστάσεις. Κάθε εικόνα έχει διάσταση  $256 \times 256$  pixel, το μέγεθος αυτό αλλάζει από τις αρχικές καταστάσεις για ομοιομορφία και για επιτάχυνση της υπολογιστικής επεξεργασίας. Οι εικόνες αυτές έχουν τμηματοποιηθεί, και επαγγελματίες γιατροί τις έχουν επαληθεύσει διασφαλίζοντας την κλινική τους ακρίβεια και συνάφεια.

Πιο συγκεκριμένα στο σύνολο των δεδομένων περιλαμβάνονται 3616 θετικά κρούσματα COVID-19 μαζί με 10.192 φυσιολογικές, 6012 πνευμονική αδιαφάνεια (λοίμωξη των πνευμόνων χωρίς COVID) και 1345 εικόνες ιογενούς πνευμονίας. Για να γίνει ισχυρή η εκπαίδευση και η επικύρωση των μοντέλων βαθιάς μάθησης το 85% από αυτές τις εικόνες διατίθεται για σκοπούς εκπαίδευσης ενώ το υπόλοιπο 15% έχει ως στόχο τη δοκιμή της απόδοσης των μοντέλων. Παράλληλα το 10% των εικόνων που προορίζονται για εκπαίδευση χρησιμοποιείται κατά τη φάση επικύρωσης για τον συντονισμό των μοντέλων και την αποφυγή υπερπροσαρμογής, ενώ το 90% για την πραγματική φάση της εκπαίδευσης.

### **3.2 Εφαρμογή και περιορισμοί των μοντέλων ταξινόμησης**

Η υλοποίηση των μοντέλων ταξινόμησής περιλαμβάνει την αξιοποίηση του TensorFlow και Keras για τη χρήση αρχιτεκτονικών CNN, ResNet50 και VGG16. Για να γίνει όμως μια τέτοια υλοποίηση υπήρξαν βασικά βήματα τα οποία αποτέλεσαν σημαντικό ρόλο καθώς ήταν αναγκαία η επεξεργασία των δεδομένων. Αρχικά, έπρεπε να γίνει αυτοματοποιημένη επαύξηση των δεδομένων, καθώς και κανονικοποίηση σε μέρη, ώστε να βελτιωθεί η ικανότητα των μοντέλων να μπορούν να γενικεύουν από τα δεδομένα εκπαίδευσης σε δεδομένα που δεν γνωρίζουν.

Κατά την υλοποίηση της εργασίας υπήρξαν κάποιοι περιορισμοί οι οποίοι έπρεπε να αντιμετωπιστούν, ένας εκ των οποίων ήταν η ανισορροπία δεδομένων. Ακόμα και μετά την εφαρμογή υπό δειγματοληψίας υπήρξε ανισορροπία μεταξύ των κατηγοριών ασθενειών στα ιατρικά σύνολα δεδομένων. Ως αποτέλεσμα η ανισορροπία αυτή επηρέασε σημαντικά την εκπαίδευση του μοντέλου. Για να αντιμετωπιστεί αυτό το πρόβλημα έγινε χρήση συνάρτησης `train_test_split` από το `sklearn.model_selection` με την παράμετρο `stratify` βοηθά να διασφαλιστεί ότι τα σύνολα δεδομένων εκπαίδευσης, επικύρωσης και δοκιμής περιέχουν μια αναλογική αναπαράσταση κάθε κλάσης. Με αυτόν τον τρόπο διατηρούμε την ισορροπία μεταξύ των συνόλων δεδομένων κάτι που είναι ζωτικής σημασίας για τα μοντέλα ώστε να υπάρχει μεγαλύτερη ακρίβεια και καλύτερες τιμές οι οποίες δείχνουν μια καλή απόδοση σε όλες τις κατηγορίες.

Ένα εξίσου σημαντικό πρόβλημα το οποίο έπρεπε να αντιμετωπιστεί ήταν οι ανακρίβειες στους σχολιασμούς (manual annotations) οι οποίες μπορούν να οδηγήσουν σε λάθος δεδομένα εκπαίδευσης, ο αλγόριθμος μάθησης μπορεί παραπλανημένα να οδηγηθεί σε λάθος συμπεράσματα το οποίο είναι επικίνδυνο στις διαφορετικές



πνευμονικές παθήσεις. Με την χρήση του ImageDataGenerator διευκολύνεται η εκπαίδευση των μοντέλων γίνεται πιο ισχυρή εφαρμόζοντας μετασχηματισμούς όπως η οριζόντια ανατροπή (horizontal flipping). Με αυτήν την επαύξηση το μοντέλο γίνεται λιγότερο ευαίσθητο σε μικρές ανακρίβειες στους σχολιασμούς. Αυτό βοηθάει στη γενίκευση έναντι πιθανών σφαλμάτων.

Η εκπαίδευση σε βαθιά νευρωνικά δίκτυα απαιτεί μνήμη και επεξεργαστική ισχύ, αυτό μπορεί να αποτελέσει σημαντικό περιοριστικό μέτρο ειδικά όταν χρησιμοποιούνται ιατρικές εικόνες υψηλής ανάλυσης. Για αυτόν τον λόγο με τη χρήση της επιλογής `include_top=False`, επιτρέπεται η μόχλευση της εκμάθησης μεταφοράς ώστε να μειωθεί το υπολογιστικό φόρτο. Αυτό έγινε καθώς τα βάρη από ένα μοντέλο προ εκπαιδευμένο σε παρόμοια εργασία βοηθάει στην προετοιμασία του μοντέλου και μειώνει σημαντικά το χρόνο ενώ χρειάζεται μόνο η προσαρμογή των προηγούμενων χαρακτηριστικών στη συγκεκριμένη εργασία της πνευμονικής νόσου ταξινόμησης. Με αυτήν την προσέγγιση δεν επιτυγχάνεται μόνον η διαδικασία της εκπαίδευσης αλλά παράλληλα ενισχύεται η απόδοση του μοντέλου χρησιμοποιώντας λιγότερους υπολογιστικούς πόρους.

Η υπερπροσαρμογή ήταν ένας παράγοντας ο οποίος λήφθηκε υπ' όψη κατά την υλοποίηση της εργασίας. Τα μοντέλα εκπαιδεύονται σε εξαιρετικά διαφορετικά και πολύπλοκα σύνολα δεδομένων, κάτι το οποίο διαφύλασσε πολλούς κινδύνους. Λύση σε αυτό αποτέλεσαν τα επιπέδα κανονικοποίησης δέσμης (Batch normalization layers) εντός της αρχιτεκτονικής του κάθε μοντέλου. Αυτό βοήθησε στη σταθεροποίηση της διαδικασίας μάθησης, και στη μείωση της εσωτερικής μεταβλητότητας των συσχετίσεων. Η μέθοδος αυτή εφαρμόζεται σε κάθε επίπεδο του δικτύου και έχει ως στόχο τη βελτίωση της σύγκλισης της εκπαίδευσης διατηρώντας τις τιμές εισόδου στο νευρωνικό επίπεδο κοντά στη βέλτιστη κατανομή. Αυτό μειώνει το φαινόμενο του "internal covariate shift" το οποίο προκαλεί διακυμάνσεις στην είσοδο των επιπέδων κατά τη διάρκεια της εκπαίδευσης κάτι που μπορεί να δυσκολέψει την εκμάθηση αλλά και να προκαλέσει αστάθεια στα βάρη του δικτύου.

### **3.3 Πειραματικά Αποτελέσματα**

Για την εξαγωγή των αποτελεσμάτων τα μοντέλα εκπαιδεύτηκαν σε ένα σύνολο δεδομένων καλά προετοιμασμένου το οποίο περιλάμβανε τις 4 κατηγορίες που έχουν προαναφερθεί. Το σύνολο των δεδομένων χωρίστηκε σε σύνολα train (80%), validation (10%), και test(10%). Οι μετρήσεις που χρησιμοποιήθηκαν ήταν το F1 Score, Precision,

Recall, Accuracy και Support. Ο λόγος επιλογής των συγκεκριμένων μετρήσεων έγινε ώστε να υπάρχει μια ολοκληρωμένη άποψη της αποτελεσματικότητας του μοντέλου, πρέπει να λαμβάνεται υπόψη τόσο η παρουσία όσο και η απουσία ενδείξεων ασθένειας. Το μοντέλο υπέδειξε υψηλά επίπεδα ακρίβειας σε όλα τα σύνολα, ενώ οι τιμές βρίσκονταν πάνω από το 80%. Οι μετρήσεις που προαναφέρθηκαν σημείωσαν ιδιαίτερα καλές τιμές κυρίως στο test set. Αυτοδικεί στο συμπέρασμα δημιουργίας ενός μοντέλου το οποίο μπορεί να κάνει με μεγάλη αποδοτικότητα σωστές προβλέψεις και έχοντας μια εικόνα να αναγνωρίζει σωστά σε ποια από τις 4 κατηγορίες βρίσκεται. Μετά από την ολοκλήρωση της εκπαίδευσης έγινε ο πίνακας σύγχυσης ώστε να απεικονιστεί η απόδοση του μοντέλου στις διαφορετικές κατηγορίες. Αυτή η απεικόνιση βοηθάει στον εντοπισμό οποιοδήποτε συγκεκριμένων κατηγοριών όπου το μοντέλο ενδέχεται να παρουσιάσει χαμηλές επιδόσεις ή να ταξινομήσει εσφαλμένα τις εικόνες. Έμφαση δόθηκε επίσης και στα τελικά αποτελέσματα του μοντέλου που αφορούσαν τις τιμές loss, accuracy, validation loss, validation accuracy. Για την καλύτερη ανάλυση των τιμών, έγινε η δημιουργία διαγραμμάτων, για να φανεί έντονα η απώλεια εκπαίδευσης και επικύρωσης.

Τα τελικά αποτελέσματα και οι παρατηρήσεις έγιναν σε ένα καλά προετοιμασμένο σύνολο δεδομένων, οι μετρήσεις επιλέχθηκαν για την αξιολόγηση της απόδοσης του μοντέλου. Τα γραφήματα που έγιναν είναι και για τα τρία μοντέλα CNN, VGG16, ResNet50 και παρουσιάζουν μια τάση βελτίωσης της ακρίβειας και μείωση της απώλειας κατά τη διάρκεια των εποχών. Για λόγους ασφάλειας της ακρίβειας των αποτελεσμάτων, η κάθε μέθοδος εξετάστηκε τρεις φορές και οι τιμές και τα συμπεράσματα που θα αναλυθούν, θα περιλαμβάνουν τα αποτελέσματα μόνων της φοράς που δεν αποτελείται ούτε από τις καλύτερες αλλά και όχι από τις χειρότερες τιμές αλλά της μέσης περίπτωσης. Οι τιμές αυτές δίνουν τη δυνατότητα να γίνει πρόβλεψη με ακρίβεια, καθώς και αποφυγή ψευδών θετικών ή αρνητικών αποτελεσμάτων, κρίσιμα για μια ιατρική διάγνωση.

Το μοντέλο του CNN φανερώνει ότι έχει σταθερή μείωση απώλειας εκπαίδευσης και επικύρωσης. Η καλύτερη εποχή είναι η 15 φανερώνοντας ένα σημείο βέλτιστης γενίκευσης πριν το μοντέλο κινδυνεύσει για υπερ προσαρμογή. Η ακρίβεια εκπαίδευσης και επικύρωσης άρχισε να αυξάνεται κατά την εποχή 28, γεγονός που δηλώνει καλή προσαρμογή στο σύνολο δεδομένων χωρίς να γίνεται υπερπροπόνηση (overtraining). (εικόνα 1)

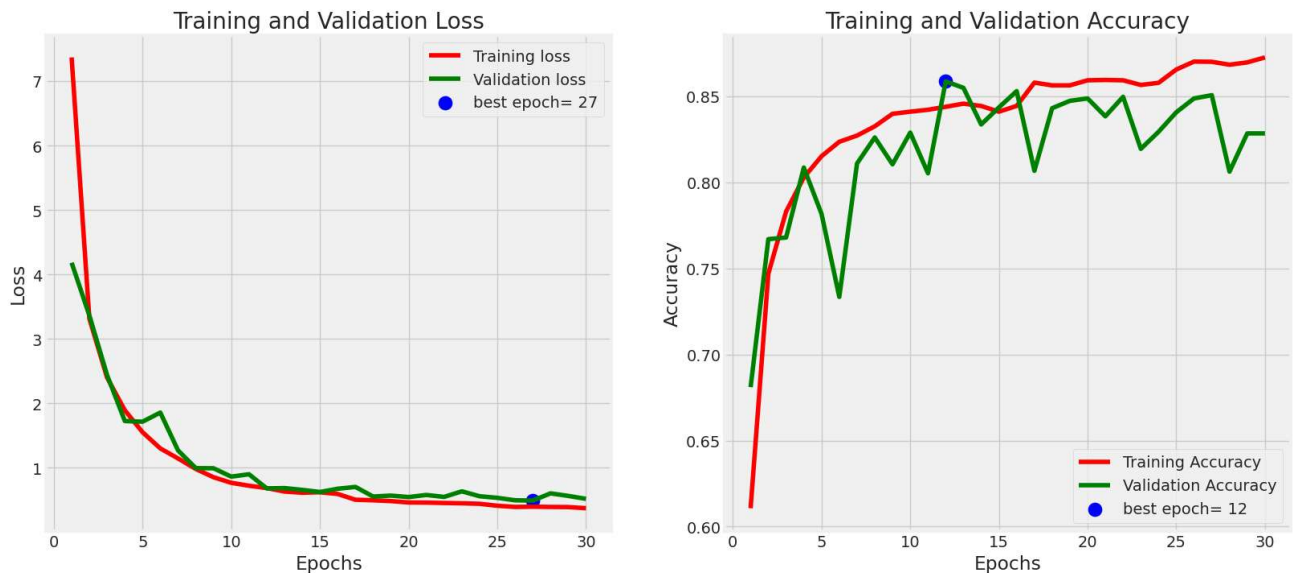


**Εικόνα 1: Συγκριτική ανάλυση απώλειας εκπαίδευσης επικύρωσης και Ακρίβειας στις 30 Εποχές για το CNN**

Στη συνέχεια η απόδοση του VGG16 φανερώνει απότομη πτώση απώλειας, ακολουθεί μια σταδιακή κάθοδο και έχει καλύτερη τιμή απώλειας επικύρωσης στην εποχή 27. Από τα αποτελέσματα αυτά προκύπτει ότι το μοντέλο έχει μάθει να λειτουργεί αποτελεσματικά παρόλο που χρειάζεται περισσότερες εποχές μέχρι να φτάσει στην πλήρη σύγκλιση, κάτι που πιθανώς οφείλεται στη σύνθετη αρχιτεκτονική του. Η ακρίβεια του VGG16 είχε μεγαλύτερες διακυμάνσεις σε σύγκριση με το CNN, συμπέρασμα που γίνεται έντονα αντιληπτό ιδιαίτερα σε μεταγενέστερες εποχές (εικόνα 2). Φανερώνεται μία πιθανή απόκλιση στην ικανότητα του μοντέλου να γενικεύει σε νέα δεδομένα. Αυτό είναι ένα κομβικό σημείο το οποίο επιζητά περαιτέρω βελτιστοποίηση και για αυτό τον λόγο το μοντέλο υποβλήθηκε σε δοκιμές πρόβλεψης με θετικά αποτελέσματα λόγω μη ύπαρξης λανθασμένων προβλέψεων.

Το ResNet50, είναι το τρίτο και τελευταίο μοντέλο το οποίο χαρακτηρίζεται σύνθετο λόγω της βαθιάς αρχιτεκτονικής του. Στο μοντέλο αυτό παρατηρείτε ομαλή κάθοδος της απώλειας, με την καλύτερη τιμή να βρίσκεται στην εποχή 23 . Μετά την εποχή 16 υπάρχει μια σχετική σταθερότητα στην ακρίβεια, φανερώνοντας ανθεκτικότητα στην υπερπροσαρμογή (εικόνα 3). Η πιθανότητα να γίνεται κάτι τέτοιο βρίσκεται στην αποτελεσματική χρήση των συνδέσεων παράκαμψης, οι οποίες μετριάζουν το πρόβλημα gradient (πρόβλημα της εξαφάνισης της κλίσης).

Από τον πίνακα σύγκρισης εξάγεται το συμπέρασμα ότι και τα 3 μοντέλα είχαν καλές αποδόσεις και των τεσσάρων καταστάσεων. Υπήρχαν καταστάσεις όπου η απόδοση δεν ήταν βέλτιστη και αυτές χρήζουν ανάλυσης, με πιθανότητα ανισορροπίας της κατηγορίας ή χαρακτηριστικών εντός των κατηγοριών, κάτι το οποίο είναι δύσκολο να βρεθεί.



**Εικόνα 2: Συγκριτική ανάλυση απώλειας εκπαίδευσης επικύρωσης και Ακρίβειας στις 30 Εποχές για το VGG16**



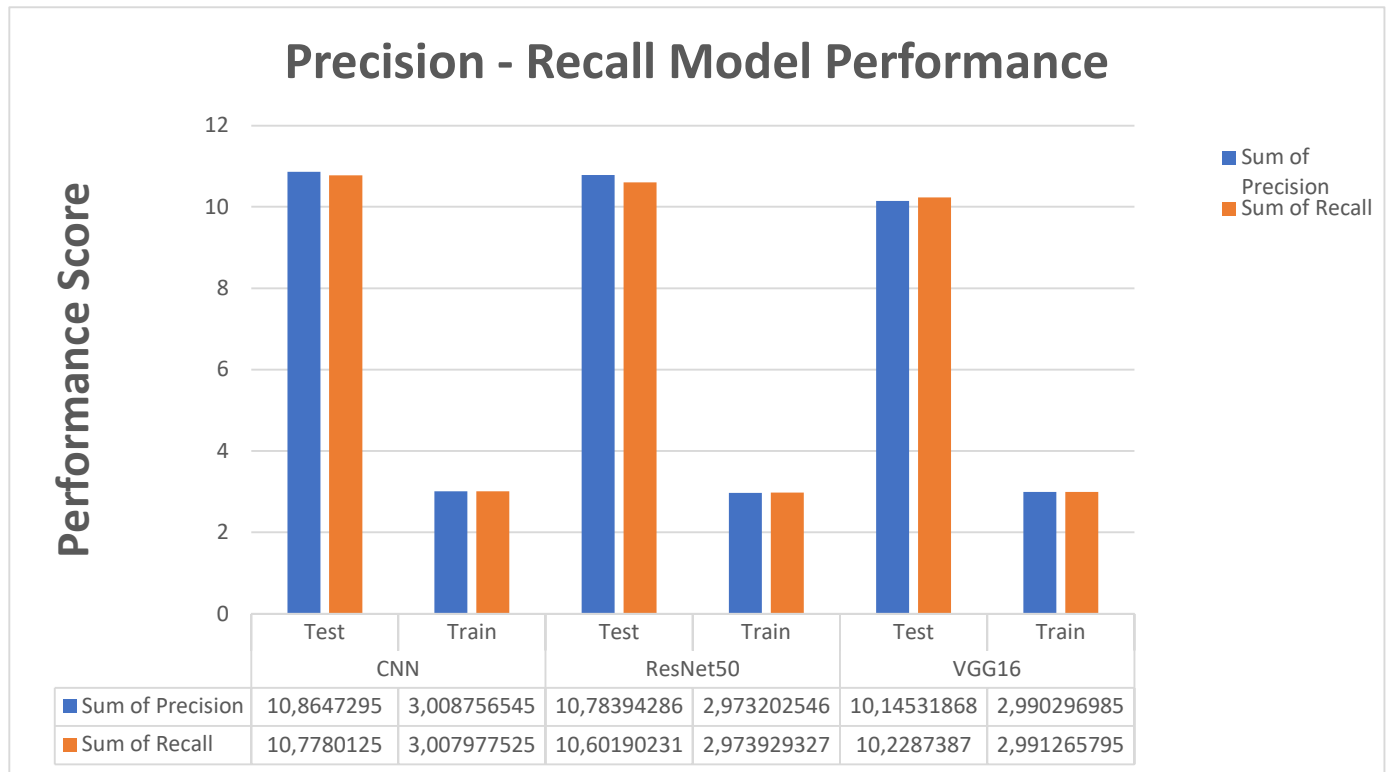
**Εικόνα 3: Συγκριτική ανάλυση απώλειας εκπαίδευσης επικύρωσης και Ακρίβειας στις 30 Εποχές για το ResNet50**

### 3.4 Οπτικοποίηση αποτελεσμάτων

Στο κεφάλαιο αυτό θα γίνει η ανάλυση και τα τελικά συμπεράσματα κάθε μεθόδου. Θα αναλυθούν οι τιμές του πίνακα σύγχυσης για κάθε μέθοδο, ο χρόνος που χρειάστηκε για την ολοκλήρωση της κάθε μεθόδου καθώς θα αναφερθούν και οι τελικές τιμές. Το συμπέρασμα της καλύτερης μεθόδου θα αποτυπωθεί στα τελικά συμπεράσματα, ενώ σε αυτό το μέρος την εργασίας θα γίνει μια γενική επισκόπηση και ένας πλήρης σχολιασμός.

Πριν ξεκινήσει η ανάλυση των αποτελεσμάτων των διαγραμμάτων θα γίνει η επισκόπηση των τιμών των ραβδογραμμάτων Accuracy – Precision – F1 Score – Recall. Λόγο περιορισμού της μνήμης αρχικά είναι απαραίτητο να γίνει αναφορά ότι τα βάρη που χρησιμοποιήθηκαν σε κάθε μέθοδο προήλθαν από ένα έτοιμο μοντέλο. Αυτό έγινε ώστε να υπάρχει έλεγχος των υπολογιστικών πόρων ώστε μία συσκευή με χρήση T4 GPU να μπορεί σε σύντομο χρονικό διάστημα να εξάγει θετικά αποτελέσματα. Για αυτό τον λόγο στα παρακάτω ραβδογράμματα οι τιμές του train set είναι αρκετά χαμηλές αλλά οι τιμές δεν είναι ανυσηχτικές εφόσον στο test set υπάρχουν καλά αποτελέσματα. Γίνεται αντιληπτό ότι τα μοντέλα στις τιμές αυτές δεν έχουν μεγάλες αποκλίσεις, γεγονός που τα καθιστά έγκυρα για σωστή πρόβλεψη πάθησης των πνευμόνων σε εικόνες X ray. Ο CNN έχει τις καλύτερες τιμές σε Precision – Recall (Ραβδόγραμμα 1), φανερώνοντας ότι το ποσοστό των θετικών προβλέψεων είναι πραγματικά σωστές και τα πραγματικά θετικά παραδείγματα μπορούν να ανιχνευτούν από το μοντέλο. Αν και ξεχώρισε μεταξύ των άλλων δύο η διαφορά της τιμής της μεθόδου δεν υπερέχει σε μεγάλο επίπεδο και ως αποτέλεσμα και ο ResNet50 και ο VGG16 λειτουργούν αποτελεσματικά σε νέα δεδομένα.

Παρόμοια μορφή φανερώνει και το ραβδόγραμμα Accuracy – F1 Score με καλύτερο το μοντέλο CNN το οποίο φέρει τις καλύτερες τιμές και στις δύο μετρικές. Αυτό συνεπάγεται ότι μπορεί να ταξινομήσει σωστά μεγάλο ποσοστό παραδειγμάτων και να παρέχει μια ισορροπημένη μετρική μεταξύ ακρίβειας και ανάκλησης. Οι τιμές αυτές είναι αποδεκτές στο test set, καθώς είναι πάνω από 85%, κάτι που σημαίνει ότι υπάρχει μεγάλη ακρίβεια και σωστή πρόβλεψη. Μια σημαντική παρατήρηση είναι ότι ο VGG16 έχει τις χειρότερες τιμές κάτι που τον καθιστά λιγότερο αποτελεσματικό σε σχέση με τις άλλες δύο μεθόδους (ραβδογράμματα 2).



**Ραβδόγραμμα 1: Συνολικές τιμές των μετρικών Precision και Recall για κάθε μοντέλο και set**

Αρχικά, είναι απαραίτητο να γίνει αναφορά στον χρόνο της ολοκλήρωσης του κάθε μοντέλου. Ο CNN αποτελεί μια απλή αλλά πρακτική μέθοδο η οποία μπορεί σε σύντομο χρόνο να παράγει ακριβή και σωστά αποτελέσματα. Κατά μέσο όρο χρειάστηκαν 39 λεπτά για την ολοκλήρωση των 30 εποχών. Οι μετρήσεις σταμάτησαν στην εποχή που θεωρείτε ότι το μοντέλο δεν μπορεί να βελτιωθεί παραπάνω, ενώ είχε ξεκινήσει και η υπερπροσαρμογή. Στον πίνακα χρόνων αναφέρονται και τα αποτελέσματα απώλειας και ακρίβειας τα οποία για το συγκεκριμένο μοντέλο υποδεικνύουν ότι έχει μάθει καλά το σύνολο δεδομένων εκπαίδευσης και έχει ισχυρή ικανότητα πρόβλεψης στο σετ εκπαίδευσης. Όσον αφορά την επικύρωση φανερώνει μεγαλύτερη απώλεια στην εκπαίδευση κάτι το οποίο είναι συνηθισμένο, καθώς το μοντέλο δοκιμάζεται σε αόρατα δεδομένα (πίνακας 3). Ωστόσο οι τιμές είναι σχετικά χαμηλές. Στο κομμάτι της ακρίβειας υπάρχει μια σταθερότητα, υπονοώντας ότι το μοντέλο γενικεύεται καλά αν και δεν αποδίδει τόσο έντονα στο σετ εκπαίδευσης (πίνακας 1). Όσον αφορά το τεστ σετ υπάρχει υψηλότερη τιμή σε σχέση με τα άλλα 2 σετ, το οποίο όμως είναι συνηθισμένο και υποδηλώνει ότι το μοντέλο μπορεί να προσαρμόζεται υπερβολικά στα δεδομένα εκπαίδευσης κάτι που ενδεχομένως να

προκαλέσει πρόβλημα στην απόδοση σε εντελώς νέα δεδομένα (πίνακας 2). Το ίδιο ισχύει και για την ακρίβεια η οποία είναι χαμηλότερη από τα άλλα 2 σετ αν και η τιμή της είναι αρκετά υψηλή με αποτέλεσμα το μοντέλο να είναι αρκετά ισχυρό να προβλέψει αόρατα δεδομένα.

Στη συνέχεια οι τιμές του VGG16 είχαμε μια καλή απόδοση με μέση ακρίβεια εκπαίδευσης η οποία υποδηλώνει σταθερή μάθηση στο σύνολο δεδομένων εκπαίδευσης (πίνακας 1). Η ολοκλήρωση των αποτελεσμάτων έγινε σε 42 λεπτά κατά μέσο όρο. Το μοντέλο φανερώνει ότι δυσκολεύεται να εκπαιδευτεί σε αόρατα δεδομένα με τη χαμηλή τιμή στην ακρίβεια επικύρωσης και δοκιμής (πίνακας 3). Αυτό ενδεχομένως να σημαίνει ότι υπάρχει υπερβολική προσαρμογή στα δεδομένα εκπαίδευσης και για αυτόν τον λόγο να μην μπορεί να γενικεύεται τόσο καλά σε νέα δεδομένα. (πίνακας 1)

Το μοντέλο ResNet δείχνει επίσης μια εξαιρετική απόδοση και μια μεγάλη τιμή στη μέση ακρίβεια της εκπαίδευσης (πίνακας 1). Οι αποδόσεις του μοντέλου είναι ισάξιες του χρόνου που χρειάστηκε για να ολοκληρωθεί. Ο μέσος χρόνος ολοκλήρωσης είναι 38 λεπτά κατατάσσοντας τον ως την γρηγορότερη μέθοδο μεταξύ των άλλων. Επίσης εξίσου μεγάλες τιμές διατήρησε και στην επικύρωση και στο τεστ κάτι το οποίο δείχνει την ισχυρή ικανότητα γενίκευσης του μοντέλου. (πίνακας 2) (πίνακας 3).

Model	Train Accuracy (%)	Train Loss	Time (minutes)
CNN	95.9	0.1166	39
VGG16	88.47	0.35885	42
ResNet	93.68	0.1556	38

**Πίνακας 1: Μέσοι όροι αποτελεσμάτων Train**

Model	Test Accuracy (%)	Test Loss	Time (minutes)
CNN	89.44	0.3858	39
VGG16	84.39	0.4985	42
ResNet	88.39	0.3521	38

**Πίνακας 2: Μέσοι όροι αποτελεσμάτων Test**

Model	Validation Accuracy (%)	Validation Loss	Time (minutes)
CNN	91.59	0.2649	39
VGG16	84.49	0.4985	42
ResNet	89.22	0.3427	38

**Πίνακας 3: Μέσοι όροι αποτελεσμάτων Validation**

### 3.5 Τελικά Συμπεράσματα

Στην μελέτη αυτή υπήρξε ένας στόχος ο οποίος ήταν να αναπτύξει και να αξιολογήσει μοντέλα βαθιάς μάθησης τα οποία να είναι ικανά να διαγνώσουν ασθένειες πνευμόνων σε εικόνες XRAY. Αρχικά έγινε η προετοιμασία του συνόλου των δεδομένων, έγινε η προετοιμασία των μοντέλων χρησιμοποιήθηκαν μετρήσεις οι οποίες επιδίωκαν να δημιουργήσουν μοντέλα ακριβή ελαχιστοποιώντας το ποσοστό λανθασμένης διάγνωσης.

Υπήρξε ουσιαστική δουλειά, κάτι που φανερώνουν τα αποτελέσματα των πειραμάτων για τα μοντέλα βαθιάς μάθησης, συγκεκριμένα τα CNN, VGG16 και ResNet50, κατάφεραν να πετύχουν μεγάλα επίπεδα απόδοσης στην προσπάθειά τους να ταξινομήσουν ιατρικές εικόνες. Τα μοντέλα έφεραν μοναδικά πλεονεκτήματα τα οποία διακρίνονται ξεχωριστά για κάθε μοντέλο.

Ο CNN ο οποίος είχε την υψηλότερη ακρίβεια εκπαίδευσης 95% δείχνει ότι είναι αποτελεσματικός στη μάθηση από το σύνολο δεδομένων. Ο χρόνος που χρειάστηκε για να τερματιστεί ήταν σχεδόν ίσος με τον ταχύτερο χρόνο κάτι που τον καθιστά εξαιρετικά κατάλληλο για έργα που απαιτούν γρήγορη ανάπτυξη και επανάληψη μοντέλων. Ένα πρόβλημα ήταν η υψηλή απώλεια στο σετ δοκιμής (0,38) Το οποίο υποδηλώνει πιθανή υπερπροσαρμογή στα δεδομένα.

Ο VGG16 Παρουσιάζει χαμηλότερα επίπεδα συνολικής ακρίβειας (84,39%) , έχει λεπτομερή ικανότητα εξαγωγής χαρακτηριστικών. Είναι ιδιαίτερα χρήσιμο για εντοπίσει μικρές ανωμαλίες σε εικόνες πνευμόνων όπου τα άλλα μοντέλα ενδέχεται να παραλείψουν. Έχει μεγάλο χρόνο εκπαίδευσης 42 λεπτά και μεγάλη τιμή απώλειας 0,49 αποτελώντας ενδεχόμενο πρόβλημα που απαιτεί προς βελτίωση για τη μείωση της μεταβλητότητας και τη βελτίωση της γενίκευσης.

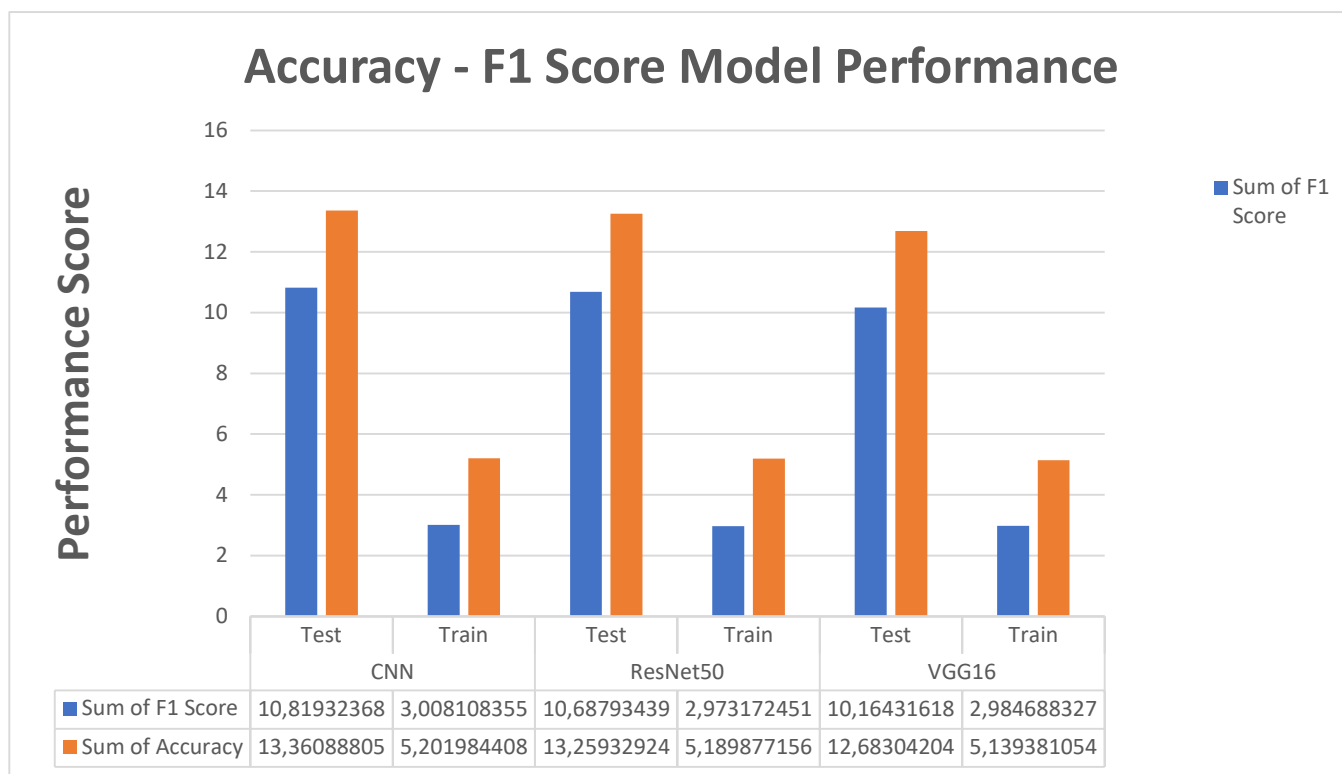
Ο ResNet50 είναι ισορροπημένος με την ακρίβεια στο 88,39% στα δεδομένα της δοκιμής έχει τη μικρότερη απώλεια δοκιμής, και αυτό προδιαθέτει ότι έχει ισχυρή δυνατότητα γενίκευσης. Μετριάζει τα αποτελέσματα της υπερπροσαρμογής, και στο σύνολο αποτελεί μια καλή επιλογή για κλινικές εφαρμογές όπου η ακρίβεια και η συνέπεια είναι πρωταρχικής σημασίας. Αποτελεί το μοντέλο με το μικρότερο χρόνο εκπαίδευσης κάτι το οποίο υπογραμμίζει την αποτελεσματικότητά του.

Σύμφωνα με τα παραπάνω ο ResNet50 Φαίνεται να είναι ο πιο πολλά υποσχόμενος και αναδεικνύεται ως το καλύτερο μοντέλο για τη διάγνωση παθήσεων των πνευμόνων σε εικόνες ακτίνων X. Η υψηλή του ακρίβεια, η χαμηλός χρόνος εξαγωγής

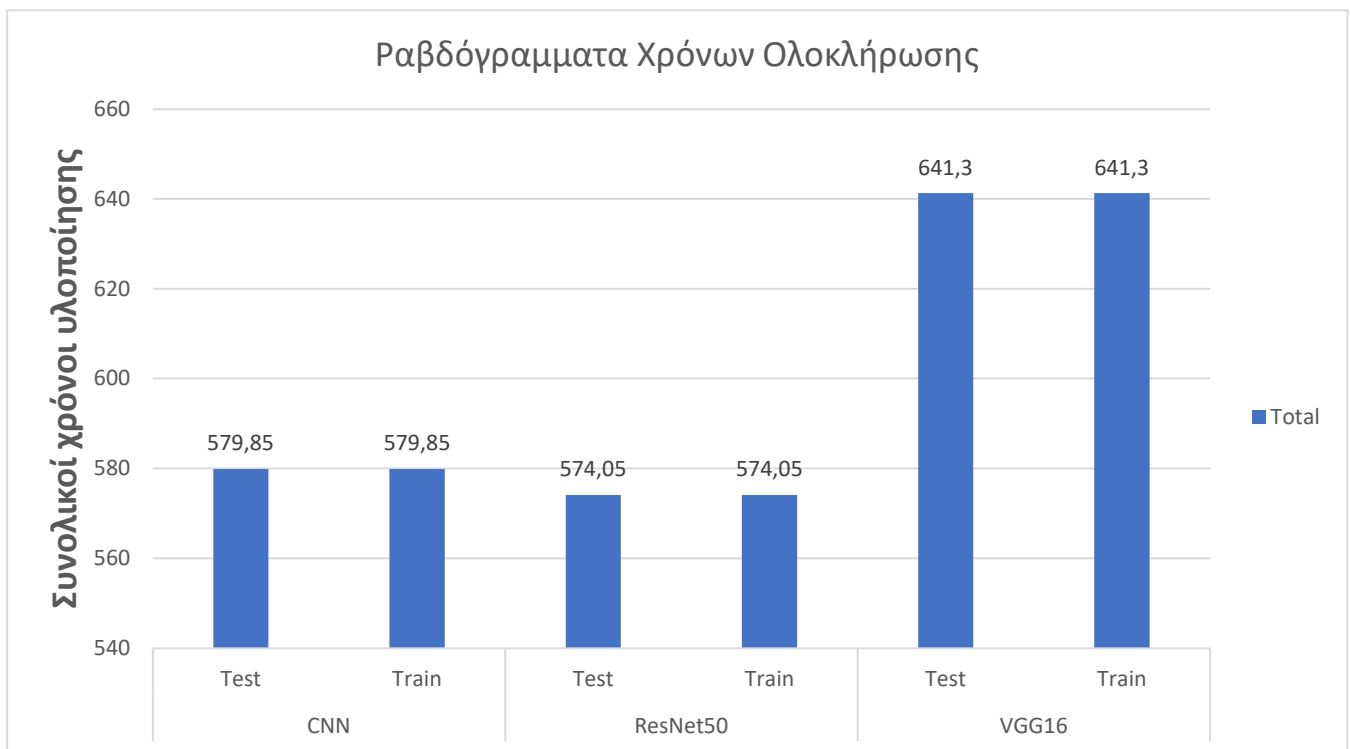


αποτελεσμάτων (ραβδόγραμμα 3), η ισορροπία στην απόδοση σε συνδυασμό με την αντοχή του στην υπερπροσαρμογή τον καθιστά καλύτερο μεταξύ των άλλων.

Στόχος για τη μελλοντική έρευνα θα είναι η βελτιστοποίηση αυτών των μοντέλων, και κυρίως η διερεύνηση τρόπων για τη βελτίωση της συνέπειας του VGG16 και τη μείωση της τάσης του CNN για υπερπροσαρμογή. Επίσης μια σημαντική δοκιμή θα είναι η ενσωμάτωση μεγαλύτερων και διαφορετικών συνόλων δεδομένων τα οποία θα συντελέσουν στη βελτίωση της ευρωστίας και της διαγνωστικής ακρίβειας των μοντέλων αυτών. Έτσι θα αντιμετωπίζονται ηθικά κριτήρια τα οποία θα είναι ζωτικής σημασίας και οι τεχνολογίες θα πλησιάζουν περισσότερο την εφαρμογή στον πραγματικό κόσμο.



**Ραβδόγραμμα 2: Συνολικές τιμές των μετρικών Accuracy και F1 Score για κάθε μοντέλο και set**



**Ραβδόγραμμα 3: Συνολικός χρόνος ολοκλήρωσης κάθε αρχιτεκτονικής**

## Βιβλιογραφία

### Σχετική έρευνα για το θεωρητικό υπόβαθρο:

- [1] R. Pandimeena, S. Bhuvaneshwari, T. A. Kumar, and P. M. Ram, "Detection of Lung Cancer Using Artificial Neural Networks," in 2022 International Conference on Data Science, Agents & Artificial Intelligence (ICDSAAI), 2022, pp. 1-5. doi: 10.1109/ICDSAAI55433.2022.10028794.
- [2] C. Wang and W. Dai, "Lung nodule segmentation via semi-residual multi-resolution neural networks," Open Life Sciences, vol. 18, 2023. doi: 10.1515/biol-2022-0727.
- [3] E. Prisciandaro, G. Sedda, A. Cara, C. Diotti, L. Spaggiari, and L. Bertolaccini, "Artificial Neural Networks in Lung Cancer Research: A Narrative Review," Journal of Clinical Medicine, vol. 12, 2023. doi: 10.3390/jcm12030880.
- 

### Dataset:

<https://www.kaggle.com/datasets/tawsifurrahman/covid19-radiography-database>

---

### Πηγές σχετικές για την παραγωγή κώδικα:

#### **TensorFlow:**

[https://www.tensorflow.org/guide/keras/transfer\\_learning](https://www.tensorflow.org/guide/keras/transfer_learning)

[https://www.tensorflow.org/api\\_docs/python/tf/keras/applications/ResNet50](https://www.tensorflow.org/api_docs/python/tf/keras/applications/ResNet50)

[https://www.tensorflow.org/api\\_docs/python/tf/keras/applications/VGG16](https://www.tensorflow.org/api_docs/python/tf/keras/applications/VGG16)

<https://www.tensorflow.org/tutorials/images/cnn>

---

#### **Kaggle:**

<https://www.kaggle.com/code/sana306/detection-of-covid-positive-cases-using-dl>

<https://www.kaggle.com/code/arunrk7/covid-19-detection-pytorch-tutorial>

<https://www.kaggle.com/code/ahmedtronic/covid-19-radiology-vgg19-f1-score-95>