

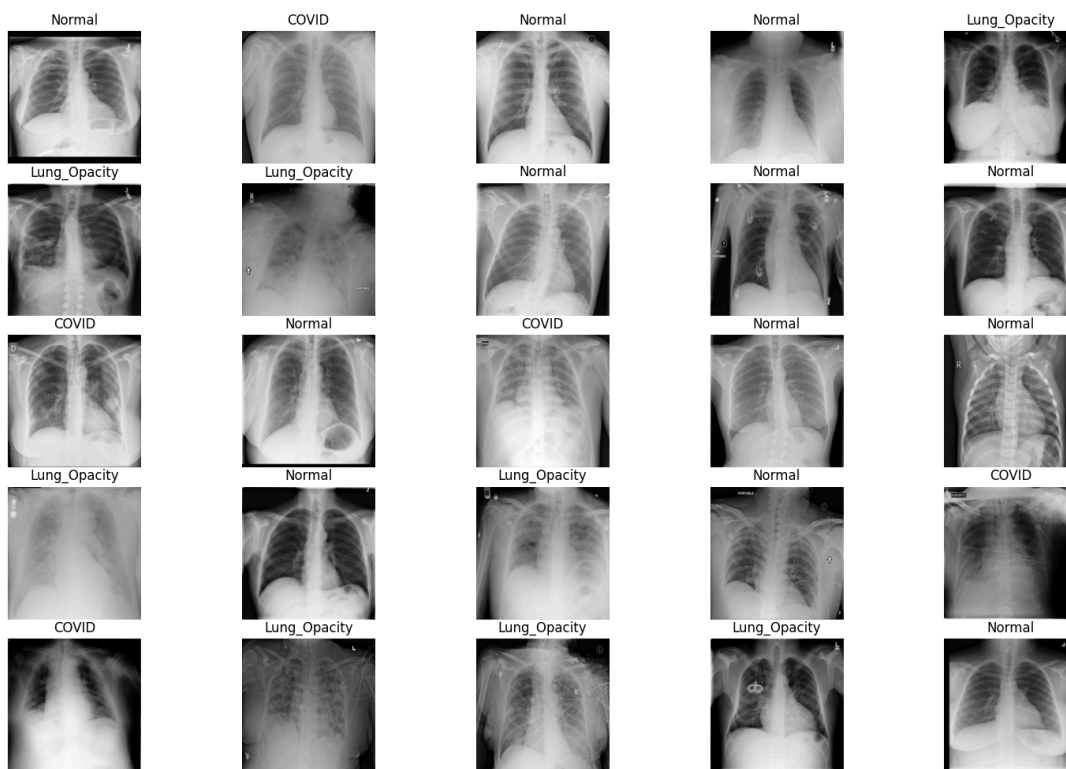
2η Εργασία στο μάθημα “Μηχανική Μάθηση και Εφαρμογές”

Κωνσταντίνος Χαφής, It22120

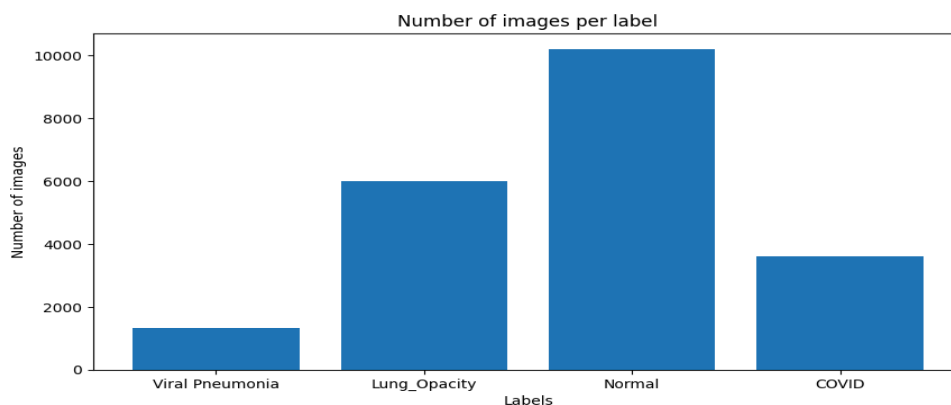
Ιανουάριος 2025

Ζητούμενο 1ο

Dataset: Παρατηρούμε πως υπάρχει μια άνιση κατανομή των εικόνων, και συγκεκριμένα οι ακτινογραφίες ασθενών χωρίς καμία ασθένεια. Αυτό είναι πολύ πιθανόν να επηρεάσει την αποδοτικότητα των μοντέλων για τη σωστή κατηγοριοποίηση των εικόνων. Ωστόσο, δεν θα είναι τόσο μεγάλο πρόβλημα, για ένα μοντέλο που έχει ένα ικανοποιητικό βάθος και που χρησιμοποιεί καλές μεθοδολογίες για την εκπαίδευση του (πχ resnet).



Δείγμα 25 ακτινογραφιών [1]



Πλήθος ακτινογραφιών ανά κλάση (ασθένεια) [2]

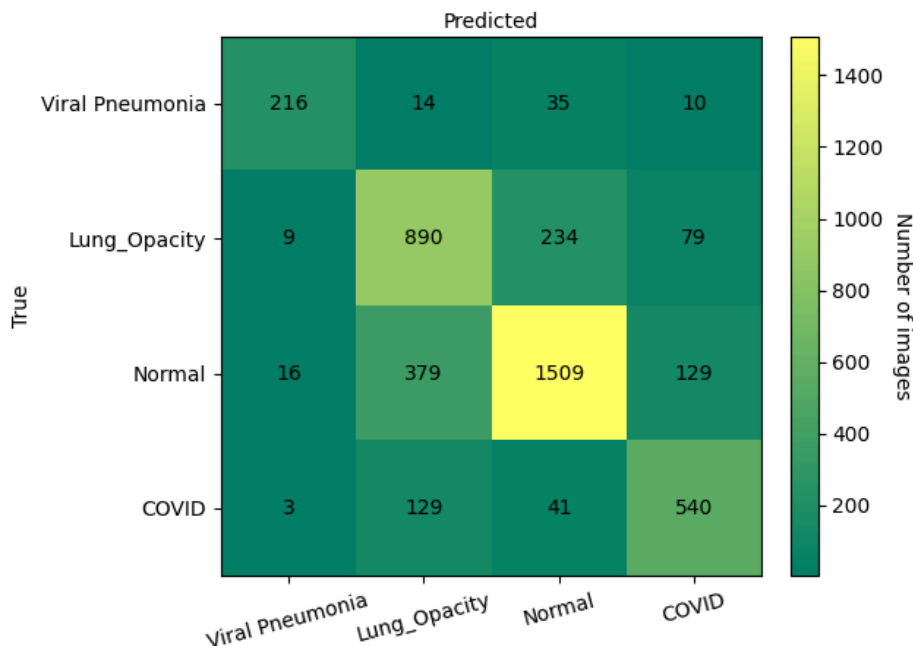
Ζητούμενο 2ο

CNN1: Για το πρώτο δίκτυο βλέπουμε πως γενικά οι επιδόσεις του μοντέλου δεν είναι ικανοποιητικές καθώς υπάρχει μια σημαντική διαφορά στην ευστοχία στα δεδομένα εκπαίδευσης και στα δεδομένα επικύρωσης και ελέγχου.

Αυτό μας οδηγεί στο συμπέρασμα πως το μοντέλο έχει μικρή πολυπλοκότητα, οπότε αδυνατεί να μάθει, να γενίκευση και να βρει τα απαραίτητα μοτίβα στα δεδομένα εκπαίδευσης και για αυτό η τελική ευστοχία είναι αφενός χειρότερη στα δεδομένα επικύρωσης και ελέγχου ([logs](#)).

```
===== Epoch: 20 =====  
Train Loss: 0.04,      Validation Loss: 3.37  
Train Accuracy: 98.51%,      Validation Accuracy: 73.37%  
  
Accuracy: 73.58%, Avg. Loss: 3.11
```

Σε αυτό το συμπέρασμα καταλήγουμε κοιτώντας και τον πίνακα σύγχυσης που δείχνει την δυνατότητα του μοντέλου να προβλέψει σε έναν σχετικά καλό βαθμό έναν υγιή άνθρωπο και έναν με πνευμονία, ωστόσο για τις άλλες δύο ασθένειες η ευστοχία είναι ιδιαίτερα χαμηλή.



Πίνακας Σύγχυσης Απλού CNN [\[3\]](#)

Επομένως, με τα προβλήματα του underfitting, του bias που έχει δημιουργηθεί λόγω της άνισης κατανομής των δεδομένων και η μη ικανοποιητική ευστοχία, αποτελούν συνδυασμός που καθιστούν το μοντέλο αδύνατο να χρησιμοποιηθεί για οποιαδήποτε χρήση στο ιατρικό χώρο, λόγω της αναξιοπιστίας του.

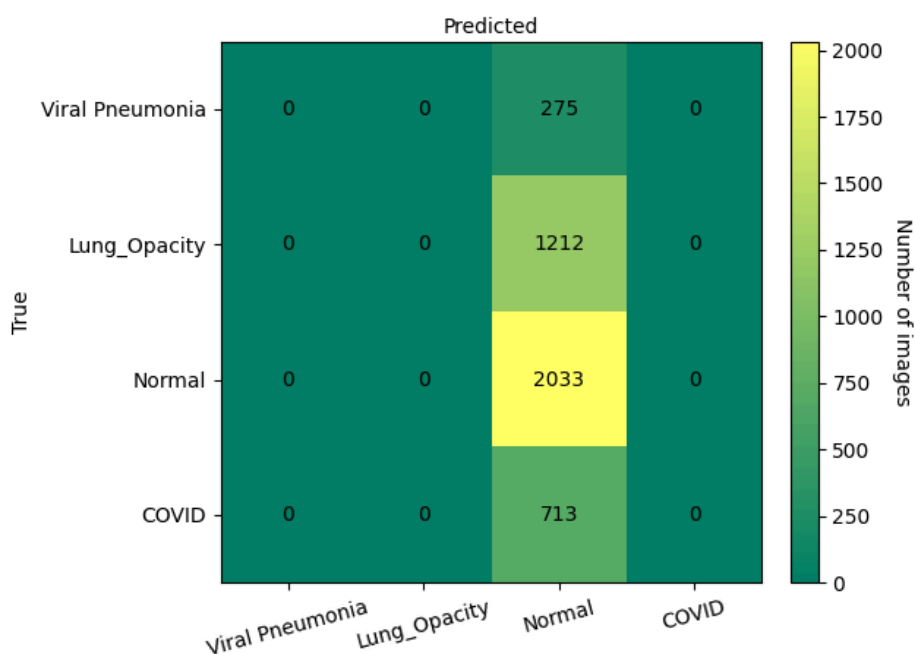
Ζητούμενο 3ο

CNN2: Τώρα για το δεύτερο δίκτυο, συναντάμε το ανάποδο από αυτό που είδαμε παραπάνω ([logs](#)). Δηλαδή το δίκτυο μας είναι αρκετά βαθύ και αδυνατεί να εκπαιδευτεί λόγω της πολυπλοκότητας του, με αποτέλεσμα να 'μαθαίνει' πάνω σε τυχόν φασαρία ή να εκπαιδεύεται πάνω σε λεπτομέρειες που δεν αποτελούν καλοί παράγοντες για την κατηγοριοποίηση των ακτινογραφιών.

```
===== Epoch: 20 =====  
Train Loss: 1.19,      Validation Loss: 1.18  
Train Accuracy: 47.90%,      Validation Accuracy: 48.02%  
  
Accuracy: 47.41%, Avg. Loss: 1.19
```

Ακόμη, το μοντέλο υποφέρει από το πρόβλημα των vanishing gradients καθώς το training loss παραμένει το ίδιο μετά από μόλις μια μόνο εποχή εκπαίδευσης.

Τα δύο αυτά προβλήματα είναι ακόμη πιο φανερά παρατηρώντας τον πίνακα σύγχυσης που δείχνει πως το μοντέλο μπορεί να κάνει πρόβλεψη μόνο για τις ακτινογραφίες με υγιείς ασθενείς. Και είναι έτσι αδύνατο να μάθει τις ιδιαιτερότητες των ακτινογραφιών με κάποια ασθένεια.

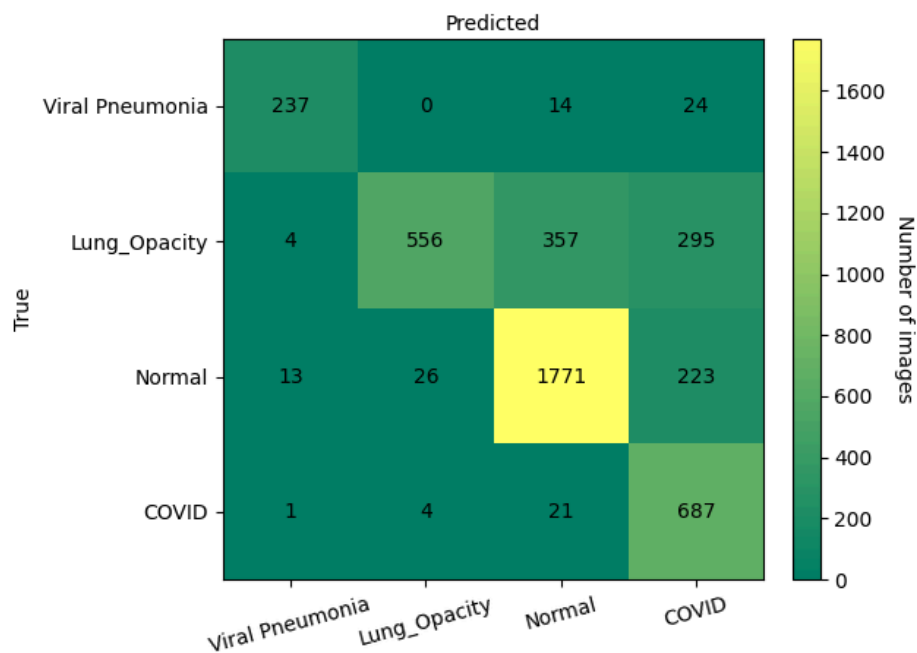


Πίνακας Σύγχυσης Βαθιά CNN [\[4\]](#)

Ζητούμενο 4ο

ResNet50: Για αυτά τα μοντέλα, η καλύτερη επίδοση τους είναι ξεκάθαρη από την 1η κιόλας εποχή ([logs](#)). Αυτό έχει να κάνει άμεσα με τη τεχνική του Residual Learning που βοηθά στην αντιμετώπιση του προβλήματος exploding/vanishing gradient.

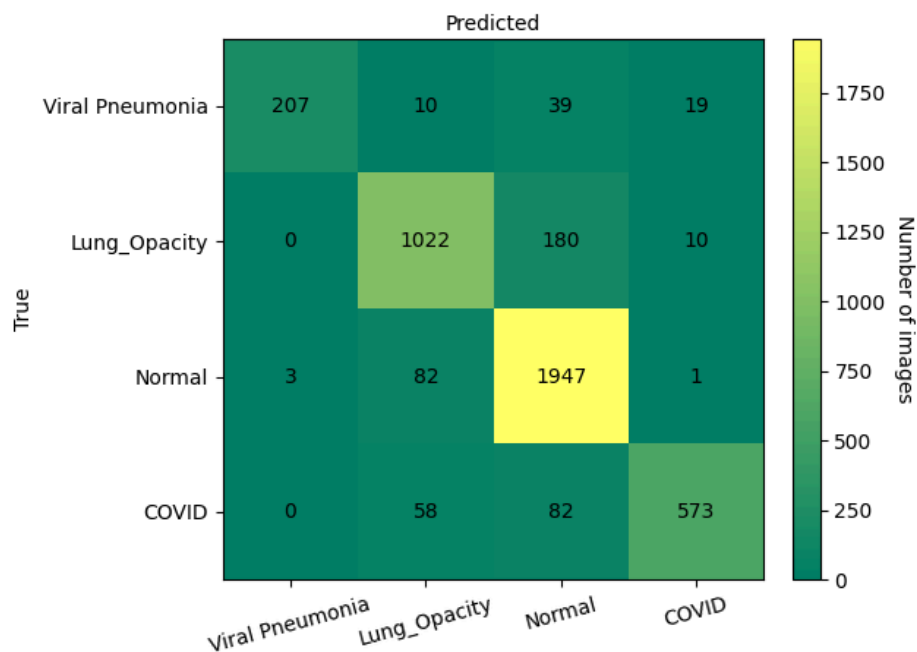
```
===== Epoch: 5 =====  
Train Loss: 0.35,      Validation Loss: 0.67  
Train Accuracy: 86.72%,      Validation Accuracy: 76.00%  
  
Accuracy: 75.82%, Avg. Loss: 0.67
```



Πίνακας Σύγχυσης Resnet [\[5\]](#)

Το **προεκπαιδευμένο** μοντέλο είναι ακόμη καλύτερο από την απλή εκδοχή του ResNet50 και αυτό φαίνεται άμεσα από την ευστοχία και τον πίνακα συσχέτισης ([logs](#)).

```
===== Epoch: 5 =====  
Train Loss: 0.11,      Validation Loss: 0.34  
Train Accuracy: 95.99%,      Validation Accuracy: 87.92%  
  
Accuracy: 87.43%, Avg. Loss: 0.35
```



Πίνακας Σύγχυσης Resnet50 Pretrained [6]

Αν και τα δύο μοντέλα είναι αρκετά πιο εύστοχα στην πρόβλεψη τους, ωστόσο δεν μπορούν παρά μόνο να αποτελέσουν σαν βοήθημα για ένα γιατρό ή ένα νοσοκομείο, εφόσον οποιαδήποτε διάγνωση ενός ασθενούς δεν μπορεί να είναι θέμα τυχαιότητας λόγω τις πιθανούς ανακρίβειας τους μοντέλου.

Ζητούμενο Bonus

CNN3-BasicBlock: Το custom μοντέλο με τη χρήση της τεχνικής του Residual Learning, αποδίδει πολύ καλύτερα αποτελέσματα στην ευστοχία. Ωστόσο για να το καταφέρω αυτό χρειάστηκαν μερικές δοκιμές, πειραματισμοί και έμπνευση από την δομή τους ResNet50.

Αν και αρχικά, η ευστοχία των μοντέλων μετά από μόλις λίγες εποχές ήταν ικανοποιητική, η ευστοχία στα δεδομένα επικύρωσης μεταβαλλόταν μεταξύ 20-60-80%, γεγονός που πιθανόν να καθιστά αναξιόπιστα τα μοντέλα για το αν όντως μπορούν να κάνουν καλή κατηγοριοποίηση ή όχι.

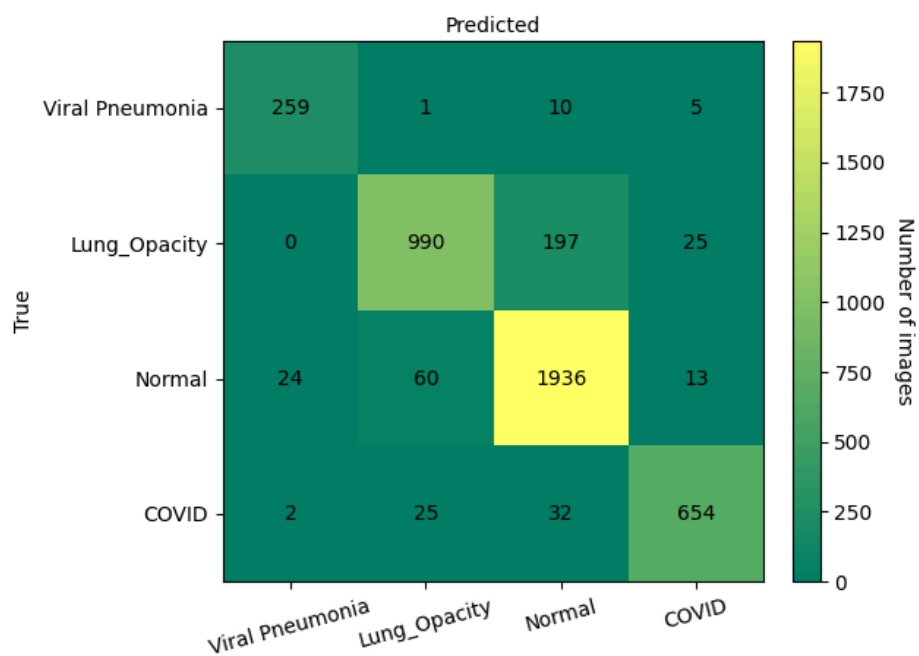
Ωστόσο, το τελευταίο μοντέλο στο οποίο κατέληξα, φαίνεται να μην έχει αυτό το πρόβλημα ([logs](#)) και με αποτέλεσμα να έχει σχεδόν την ίδια ευστοχία με το ResNet50-Pretrained, αν και βέβαια μετά από 10 εποχές εκπαίδευσης.

```

===== Epoch: 10 =====
Train Loss: 0.18,      Validation Loss: 0.31
Train Accuracy: 93.19%,      Validation Accuracy: 89.39%

Accuracy: 89.53%, Avg. Loss: 0.30

```



Πίνακας Σύγχυσης Βαθιά CNN με Residual Learning [\[7\]](#)

Αποθετήριο Κώδικα

<https://github.com/Kostas-Xafis/DeepLearning>