

Διάγνωση Πνευμονίας

Εμμανουήλ Μορφιαδάκης 21110
Κυριάκος Ναβροζίδης 21113

Περιγραφή του προβλήματος

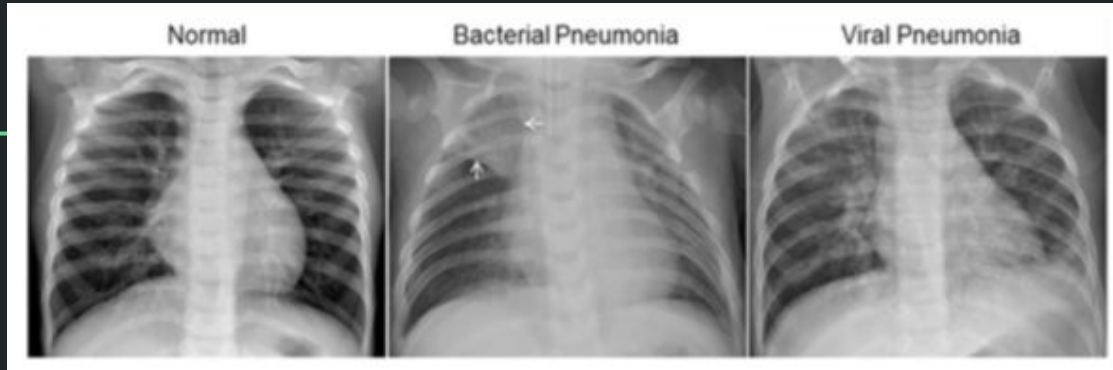
Σήμερα ο κόσμος μαστίζεται από διάφορα άγχη,όπως εύρεση εργασίας,επίδοση στη δουλειά του,να φτιάξει το σπιτικό του, εξεταστικές κλπ.Υπάρχουν όμως και κάποιοι άνθρωποι που έχουν κάποιους πόνους στο στήθος και φοβούνται μην έχουν προβλήματα με τον πνεύμονα συν τα παραπάνω άγχη.Ο ακτινολόγος μέχρι να κάνει την διάγνωση θέλει κάποιο χρόνο και μπορεί ο ασθενής να αγχωθεί πολύ με την αναμονή.

Γι αυτό σκοπεύουμε να εκπαιδεύσουμε ένα μοντέλο που να μαθαίνει αν ο ασθενής έχει πνευμονία και να δίνει τα αποτελέσματα σε πολύ λιγότερο χρόνο από ότι ο ακτινολόγος.

Dataset

To dataset έχει 5.856 cases

- viral pneumonia: 25.5%
- bacterial pneumonia: 47.5%
- Normal: 27%
- Size > 900*900 (Διαφορετικών μεγεθών)
- 3 Κατηγορίες



Ως έξοδο το σύστημα ταυτοποιεί αν κάποιος έχει πνευμονία 0 αν έχει viral , 1 αν έχει Bacterial , και 2 αν δεν έχει πνευμονία.

Preprocessing

- Μετατροπή κάθε φωτογραφίας σε array
- Μετατροπή των τιμών σε float32 και τις τιμές από [0-255] σε [0,1]
- Data augmentation on training dataset (5.856 => 15200)
- Color (RGB, GRAYSCALE)

Διαστάσεις εικόνας

- 224X224
- 200X200
- 100X100

Μεθοδολογία

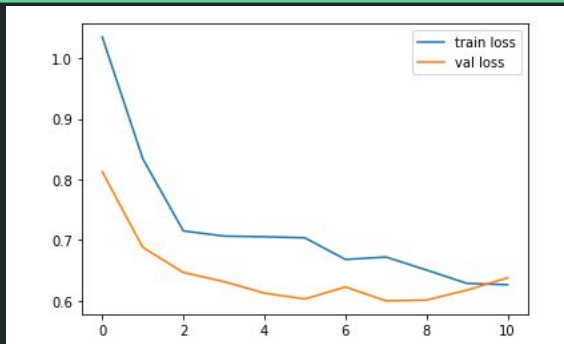
Η διαδικασία που ακολουθήσαμε περιλαμβάνει διερεύνηση της λύσης μέσα από διάφορες τεχνικές και εφαρμογές που διδαχθήκαμε σε όλη τη διάρκεια του εξαμήνου μέσα στο μάθημα. Προσπαθήσαμε να μην εντάξουμε οποιαδήποτε τεχνική έξω από τα πλαίσια αυτά.

Model Architectures Problems

- Adam optimizer με 3 διαφορετικά Learning Rates για το καθένα
τιμές $l=[0.01,0.001,0.0001]$
- Αυξομειώσεις του βάθους των μοντέλων (2-7) convolution layers, 3-11 artificial layers
- 20 - 100 εποχές
- Τεχνική dropout (0.1 , 0.2, 0.25)
- Callback Early Stopping με (patience 3-20, cross validation)
- L2 Regularization με τιμές α από $[10^{-2}, 10^{-7}]$
- padding (valid, same(επηρεαστήκαμε από την VGG16))
- cross validation με 4 διαχωρισμούς
- Batch (50-200)

Μοντελο 1

- 3 συνελκτικά επίπεδα (16,32,64)
- Max pooling(2,2) μετά από κάθε συνελκτικό
- Fully Connected + Relu
- 1 Dense layer("softmax")
- Callbacks (patience=3, mode=min)
- batch_size = 32
- epochs = 50
- Finish epochs = 14
- Accuracy = 45%
- Training score= 69%



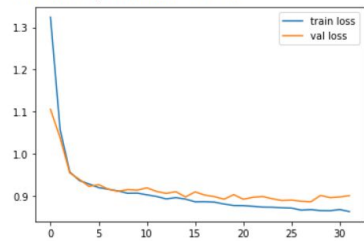
Μοντελο 2

- 4 συνελικτικά επίπεδα (40 νευρώνες)
 - Max pooling(2,2) μετά από κάθε συνελικτικό
 - Fully Connected + Relu
 - 3 Dense layers("RELU" 30 νευρώνες το καθένα)
 - Dropout(0.2) μετά από κάθε RELU
 - 1 Dense layer("softmax")
 - L2=0.01(Για Dense Layers)
 - Callbacks (patience=4, mode=max)
 - batch_size = 60
 - epochs = 200
 - Finish epochs 24
 - Accuracy= 76%
 - Training score= 79%
-

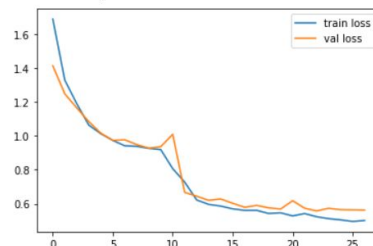
Μοντελο 3

- 6 συνελκτικα επίπεδα (32,64,128,128,256,256)
- Max pooling(2,2) μετά από αλλαγή στον αριθμό των νευρώνων
- Fully Connected + Relu
- BatchNormalization για κάθε convolutional layer
- 2 Dense layers("RELU" 40,30 νευρώνες)
- Dropout(0.2) μετά από κάθε RELU
- 1 Dense layer("softmax")
- $L2=10^{-7}$ (Για Dense Layers)
- Callbacks (patience=20, mode=max,keep_best_weights)
- batch_size = 50
- epochs = 100
- Finish epochs 62
- Accuracy = 80.7%
- Training score= 83.4%

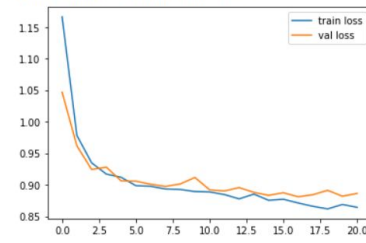
Test loss: 0.548294723033905
Test accuracy: 0.7738907933235168



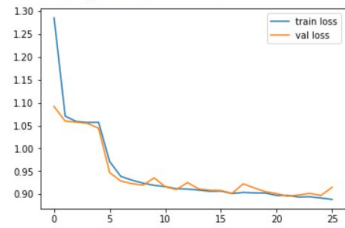
Test loss: 0.5551278591156006
Test accuracy: 0.782423198223114



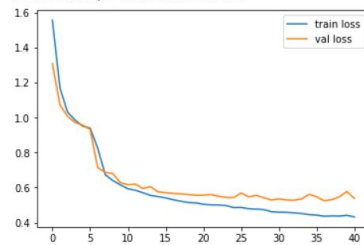
Test loss: 0.5240959525108337
Test accuracy: 0.7849829196929932



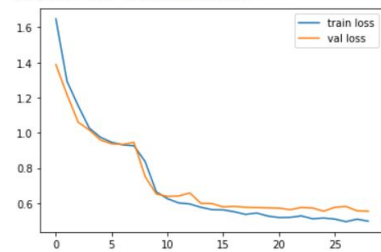
Test loss: 0.626929983978271
Test accuracy: 0.7508532404899597



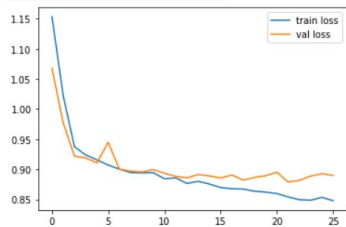
Test loss: 0.5372496843338013
Test accuracy: 0.7790102362632751



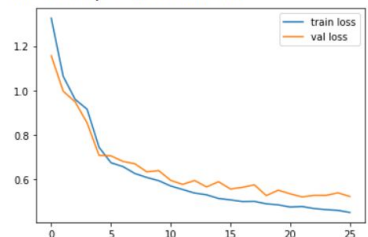
Test loss: 0.5589168071746826
Test accuracy: 0.782423198223114



Test loss: 0.5458223019732666
Test accuracy: 0.7773037552833557

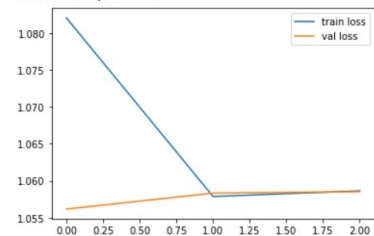


Test loss: 0.526958703994751
Test accuracy: 0.7662116289138794



non trainable params: 0

Test loss: 1.061693787574768
Test accuracy: 0.4744027256965637



Προβληματισμοί

- Μέγεθος εικόνας
- Μη επαρκής GPU

Διαπιστώσεις

- Καλύτερα αποτελέσματα μετά από Data Augmentation
-
- Μεγάλο patience και keep best weights
 - Απότομη αύξηση του validation score δεν είναι απαραίτητα θετικό
 - Αργή αύξηση του training score => καλύτερο test score

Ευχαριστούμε

Any questions?