



Πανεπιστήμιο Μακεδονίας

Τμήμα Εφαρμοσμένης Πληροφορικής

Απαλλακτική εργασία στο μάθημα Νευρωνικά Δίκτυα

MRI Classification Problem

Αλεξανδρίδης Λάζαρος (ics22080)
Βαρνάς Δημήτριος (ics22082)

Ιούνιος 2025

Περιεχόμενα

1. Εισαγωγή.....	6
2. Σχετική Βιβλιογραφία.....	7
3. Προτεινόμενη Μεθοδολογία.....	8
3.1. Προεπεξεργασία Δεδομένων.....	8
3.2. Προτεινόμενες Αρχιτεκτονικές.....	9
4. Πειραματικά Αποτελέσματα.....	12
4.1. Επεξήγηση του Dataset.....	12
4.2. Βελτιστοποίηση Υπερπαραμέτρων.....	13
4.3. Εκπαίδευση των μοντέλων.....	14
4.3.1. ResNet-50.....	14
4.3.2. EfficientNet-B0.....	17
4.3.3. DenseNet121.....	19
4.4. Αποτελέσματα.....	21
5. Συμπεράσματα.....	24
6. Βιβλιογραφία.....	25

Εικόνες

Εικόνα 1 - Διάγραμμα Ροής (Flowchart) της λειτουργίας των αναπτυσσόμενων μοντέλων.....	8
Εικόνα 2 - Αρχική και τροποποιημένη εικόνα MRI.....	9
Εικόνα 3 - Αρχιτεκτονική ResNet-50.....	10
Εικόνα 4 - Αρχιτεκτονική EfficientNet-B0.....	10
Εικόνα 5 - Αρχιτεκτονική DenseNet121.....	11
Εικόνα 6 - Stratified k-fold Cross Validation.....	13

Διαγράμματα

Διάγραμμα 1 - Πληθυσμός Dataset πριν και μετά την περικοπή.....	12
Διάγραμμα 2 - Loss Curve για την αρχιτεκτονική ResNet-18.....	15
Διάγραμμα 3 - Confusion Matrix για το μοντέλο ResNet-50.....	16
Διάγραμμα 4 - Loss και Accuracy Curves για το μοντέλο ResNet-50.....	16
Διάγραμμα 5 - Loss Curve για το μοντέλο EfficientNet-B0 πριν την προσθήκη Regularizers.....	17
Διάγραμμα 6 - Confusion Matrix για το μοντέλο EfficientNet-B0.....	18
Διάγραμμα 7 - Loss και Accuracy Curves για το μοντέλο EfficientNet-B0.....	18
Διάγραμμα 8 - Confusion Matrix για το μοντέλο DenseNet121.....	19
Διάγραμμα 9 - Loss και Accuracy Curves για το μοντέλο DenseNet121.....	20
Διάγραμμα 10 - Αποδόσεις των μοντέλων για τις μετρικές Accuracy, Precision και Recall.....	21
Διάγραμμα 11 - Απόδοση του κάθε μοντέλου σε όλες τις περιπτώσεις για την μετρική F1 Score.....	22

Πίνακες

Πίνακας 1 - Διακύμανση τιμών του Learning Rate.....	14
Πίνακας 2 - Μαθηματικοί τύποι των μετρικών Accuracy, Precision, Recall και F1 Score.....	21
Πίνακας 3 - Αποδόσεις των μοντέλων για τις μετρικές Accuracy, Precision και Recall.....	21

1. Εισαγωγή

Η εργασία αυτή στοχεύει στην δημιουργία ενός Νευρωνικού Δικτύου το οποίο λαμβάνει εικόνες από μαγνητικές τομογραφίες (Magnetic Resonance Imaging - MRI) και αναγνωρίζει εάν ο ασθενής πάσχει είτε από αιμορραγία είτε από ισχαιμία, ή εάν δεν πάσχει από κάποια πάθηση, δηλαδή είναι υγιής. Η διάγνωση μίας μαγνητικής τομογραφίας σε γρήγορο χρόνο είναι ιδιαίτερα σημαντική, καθώς δίνει την δυνατότητα στο ιατρικό προσωπικό να αναγνωρίσει την φύση του προβλήματος και να ελαττώσει τις συνέπειες της τυχόν πάθησης [1]. Επομένως, απαιτείται η ανάπτυξη ενός μηχανισμού ο οποίος έχει την ικανότητα να ανιχνεύει τις παραπάνω περιπτώσεις με ακρίβεια και ταχύτητα.

Για την επίλυση του παραπάνω προβλήματος αξιοποιείται η μέθοδος του classification, σε συνδυασμο με object detection. Συγκεκριμένα, τα δεδομένα χωρίζονται σε τρεις κλάσεις: Αιμορραγία, Ισχαιμία και Υγιής. Παράλληλα, η μέθοδος του object detection εξασφαλίζει την γρήγορη ανάγνωση κάποιας από τις παθήσεις, αναγνωρίζοντας τα σημεία στα οποία αυτή παρουσιάζεται.

Το dataset¹ που χρησιμοποιήθηκε είναι δημόσια διαθέσιμο και προσφέρει ένα σύνολο εικόνων από μαγνητικές τομογραφίες. Ωστόσο, κάθε φωτογραφία που υπάγεται στην κλάση Αιμορραγία ή στην κλάση Ισχαιμία συμπεριλαμβάνεται τρεις φορές: Αρχικά ως εικόνα σε μορφή DICOM² [2], έπειτα συνοδευόμενη από μια μάσκα overlay, η οποία προσδιορίζει με κόκκινο χρώμα το σημείο στο οποίο βρίσκεται το αιμάτωμα, και τέλος στην αρχική της μορφή. Το φαινόμενο αυτό δεν παρατηρείται για τις περιπτώσεις όπου η φωτογραφία ανήκει στην κλάση Υγιής. Για την εκπαίδευση των μοντέλων χρησιμοποιήθηκαν αποκλειστικά οι εικόνες της τρίτης περίπτωσης, δηλαδή οι τομογραφίες στην αρχική τους μορφή.

Ο σκοπός της παρούσας έρευνας είναι η ανάπτυξη και η σύγκριση ορισμένων μοντέλων βασισμένα σε ήδη υπάρχουσες αρχιτεκτονικές τα οποία κάνουν αποκλειστικά πρόβλεψη της κλάσης στην οποία κατατάσσεται οποιοδήποτε δείγμα, χωρίς να επισημαίνει το σημείο στο οποίο βρίσκεται (εάν υπάρχει) το αιμάτωμα.

¹ <https://www.kaggle.com/datasets/orvile/inme-veri-seti-stroke-dataset/data>

² Digital Imaging and Communications in Medicine | Αποτελεί το διεθνές πρότυπο για ιατρικές εικόνες και σχετικές πληροφορίες.

2. Σχετική Βιβλιογραφία

Το συγκεκριμένο πρόβλημα είναι ευρέως διαδεδομένο σε αυτόν τον χώρο, επομένως είναι προφανές ότι είναι εξίσου διαδεδομένη και η βιβλιογραφία που σχετίζεται με αυτό. Υπάρχουν πολλές έρευνες οι οποίες έχουν εξερευνήσει διάφορες αρχιτεκτονικές νευρωνικών δικτύων, χρησιμοποιώντας ξεχωριστές στρατηγικές για την προεπεξεργασία και την τροποποίηση των δεδομένων. Σκοπός αυτής της ενότητας είναι να επισημάνει ορισμένες βασικές μεθοδολογίες και να οριοθετήσει το σημείο αναφοράς για τα αναμενόμενα αποτελέσματα.

Το άρθρο [3] εφαρμόζει, μεταξύ άλλων, τις τεχνικές EfficientNet B0 και ResNet-18 για την κατάταξη των εικόνων στις κλάσεις Γλοιώμα, Μηνιγγίωμα, Όγκος Υπόφυσης, και Απουσία Όγκου. Το dataset περιείχε συνολικά 7023 εικόνες, τις οποίες οι ερευνητές αρχικά άλλαξαν μέγεθος και στην συνέχεια περιέστριψαν. Ως συνάρτηση ενεργοποίησης χρησιμοποιήθηκε η ReLU³, ενώ τα μοντέλα πέτυχαν 89% και 86% accuracy αντίστοιχα.

Η επόμενη έρευνα [4] προτείνει την χρήση του μοντέλου ResNet50 σε συνδυασμό με το Grad-CAM⁴ [5]. Το accuracy του μοντέλου αυτού ανέρχεται στο 98.52%, με τις μετρικές precision και recall να ξεπερνούν και αυτές το 98%. Η εφαρμογή του Grad-CAM παρέχει με λεπτομέρεια τα σημεία στα οποία επικεντρώνεται το μοντέλο για να κάνει τις προβλέψεις του.

Στην συνέχεια παρουσιάζεται ένα άρθρο στο οποίο χρησιμοποιήθηκε η αρχιτεκτονική EfficientNet-B0 [6]. Παρά το γεγονός ότι η έρευνα εστίασε στο accuracy του μοντέλου, η οποία ανέρχεται στο 99%, έγινε και εκτενής αναφορά στην σημασία που έχει η σωστή εφαρμογή του μοντέλου σε κλινικά περιβάλλοντα, προτείνοντας την χρήση των Explainable AI [7] σε μελλοντικές υλοποιήσεις.

Τέλος, το ακόλουθο άρθρο επικεντρώνεται στην αύξηση της απόδοσης του μοντέλου με την τροποποίηση των δεδομένων και την αξιοποίηση της αρχιτεκτονικής DenseNet121 [8]. Καθώς το πρόβλημα ήταν δυαδικής ταξινόμησης, χρησιμοποιήθηκε η Σιγμοειδής⁵ συνάρτηση ενεργοποίησης. Οι ερευνητές έλεγξαν την απόδοση του μοντέλου πριν και μετά την τροποποίηση των δεδομένων, με το accuracy να είναι 90.63% στην πρώτη περίπτωση και 94.83% στην δεύτερη. Είναι εμφανές, λοιπόν, η αποτελεσματικότητα της τεχνικής αυτής πάνω στο συγκεκριμένο dataset.

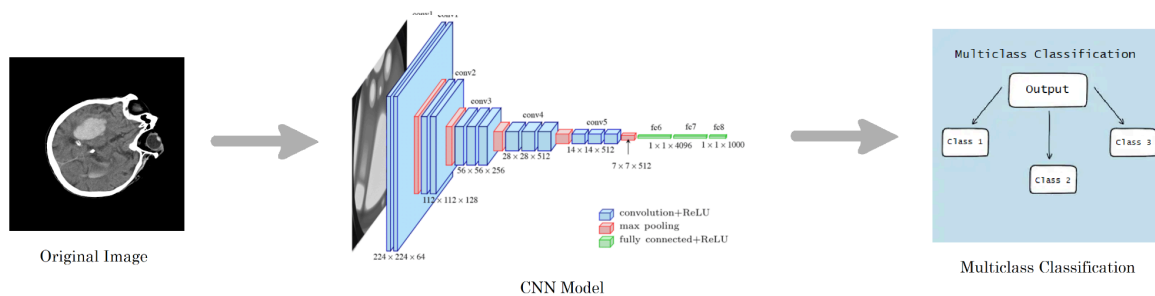
³ Rectified Linear Unit | $\text{ReLU}(z) = \max\{0, z\}$

⁴ Gradient-weighted Class Activation Mapping | Δημιουργία heatmap για την εύρεση σημαντικών περιοχών στην εικόνα.

⁵ $g(z) = 1/[1+e^{(-z)}]$

3. Προτεινόμενη Μεθοδολογία

Όπως αναφέρθηκε και στην εισαγωγή, το πρόβλημα μας αναφέρεται στην αναγνώριση αιμορραγίας ή ισχαιμίας, βάσει μιας μαγνητικής τομογραφίας. Ωστόσο, για να μπορέσουν τα μοντέλα μας να υλοποιήσουν αυτή την εφαρμογή, είναι πρώτα απαραίτητη η προεπεξεργασία των δεδομένων, με σκοπό να βελτιωθεί η ποιότητα αλλά και η ποικιλία τους. Κατά αυτόν τον τρόπο, τα μοντέλα μας μπορούν να εκπαιδευτούν καλύτερα και να γενικευτούν σε διαφορετικά, παρόμοια ωστόσο δεδομένα. Στην εικόνα 1 απεικονίζεται το διάγραμμα ροής (flowchart) που θα ακολουθήσουμε στην ανάπτυξη των μοντέλων.

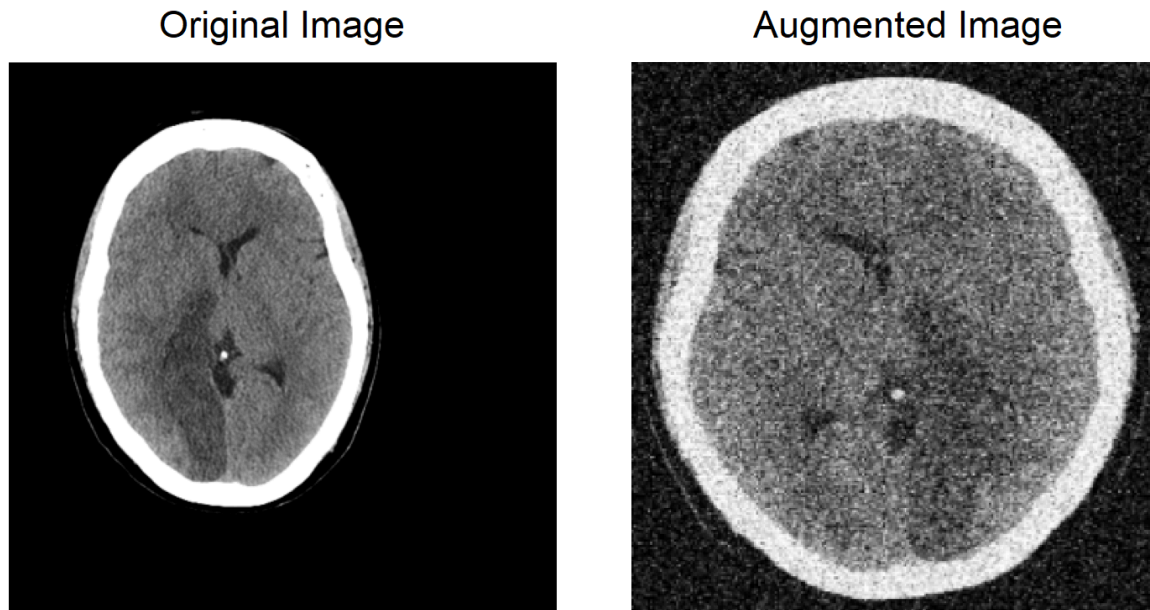


Εικόνα 1 - Διάγραμμα Ροής (Flowchart) της λειτουργίας των αναπτυσσόμενων μοντέλων

3.1. Προεπεξεργασία Δεδομένων

Για όλα τα μοντέλα τα οποία αναπτύξαμε πάνω στην έρευνα αυτή, ακολουθήσαμε την ίδια φιλοσοφία όσον αφορά την προεπεξεργασία των δεδομένων. Το πρώτο βήμα της διαδικασίας αυτής ήταν η αλλαγή του μεγέθους των εικόνων σε διαστάσεις 224 x 224, καθώς αποτελούν τις προεπιλεγμένες διαστάσεις που δέχονται τα μοντέλα αυτά ως είσοδο. Είναι σημαντικό να αναφέρουμε πως, ενώ η διαδικασία αυτή γίνεται με συμπίεση της εικόνας, ταυτόχρονα κόβουμε τις άκρες των εικόνων που αποτελούν περιττή πληροφορία, βελτιώνοντας έτσι την ανάλυση των εικόνων.

Έπειτα, τροποποιήσαμε τα δεδομένα εφαρμόζοντας διάφορες τεχνικές, για την ορθότερη εκπαίδευση των μοντέλων. Συγκεκριμένα, ανατρέψαμε οριζόντια το 50% των εικόνων, αυξομειώσαμε τυχαία την φωτεινότητα και την αντίθεση του 20% των εικόνων και γυρίσαμε το 30% των εικόνων με κατώφλι 15 μοιρών προς οποιαδήποτε κατεύθυνση. Ακόμα, προσθέσαμε Gaussian Noise [9] στο 20% των εικόνων και εφαρμόσαμε τυχαία κλιμάκωση 10% στο ίδιο ποσοστό. Σε αυτό το σημείο πρέπει να σημειωθεί πως οποιαδήποτε εικόνα μπορεί να τροποποιηθεί ταυτόχρονα με πολλούς τρόπους, γεγονός που επιβεβαιώνεται και από το άθροισμα των παραπάνω ποσοστών. Ένα παράδειγμα μίας τέτοιας τροποποίησης φαίνεται στην Εικόνα 2.



Εικόνα 2 - Αρχική και τροποποιημένη εικόνα MRI

Τέλος, έγινε κανονικοποίηση των τιμών των pixel στην κλίμακα $[0, 1]$ με σκοπό να υπάρχει μια σταθερή κλίμακα εισόδου και να γίνει πιο εύκολη η γενίκευση των μοντέλων. Παράλληλα, έγινε και φόρτωση των εικόνων σε grayscale⁶. Κατά αυτόν τον τρόπο, η είσοδος αλλάζει από τρία κανάλια χρώματος (RGB) σε ένα, μειώνοντας σημαντικά τον όγκο των δεδομένων εισόδου. Η τεχνική αυτή είναι ιδανική για επεξεργασία εικόνων μαγνητικής τομογραφίας.

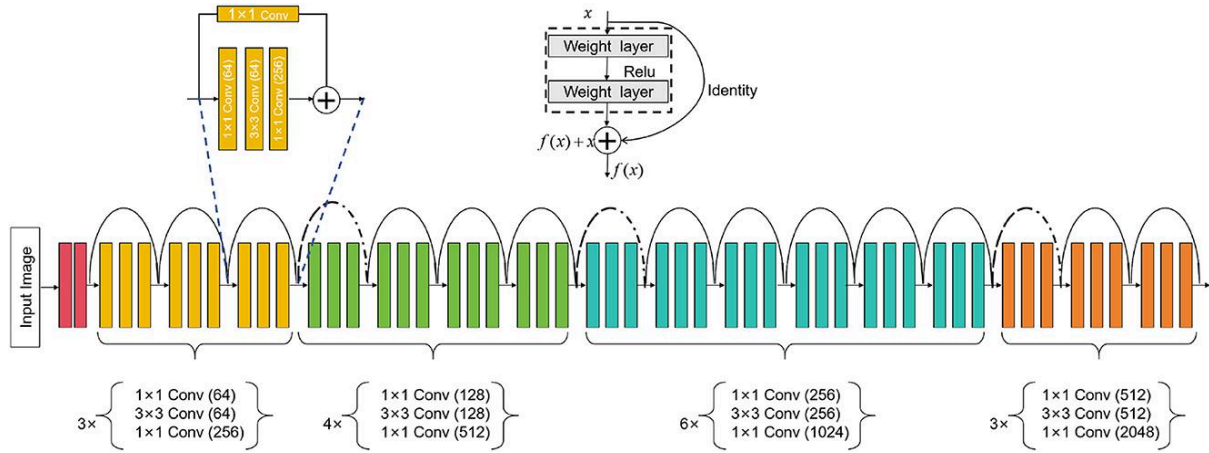
3.2. Προτεινόμενες Αρχιτεκτονικές

Για την υλοποίηση του Νευρωνικού μας Δικτύου χρησιμοποιήσαμε τρεις αρχιτεκτονικές, οι οποίες σχετίζονται με την ταξινόμηση εικόνων (image classification) και ανίχνευση αντικειμένων (object detection). Συγκεκριμένα, αναπτύξαμε και εκπαιδεύσαμε τα μοντέλα μας πάνω στις αρχιτεκτονικές ResNet-50, EfficientNet-B0 και DenseNet121.

ResNet-50 - Το μοντέλο ResNet-50 αποτελεί μία αρχιτεκτονική Συνελικτικού Νευρωνικού Δικτύου [10] (Convolutional Neural Network - CNN) η οποία χρησιμοποιεί υπολειπόμενη μάθηση για να μπορεί να προσπεράσει ένα ή περισσότερα επίπεδα. Αυτή η τεχνική επιτρέπει στο μοντέλο να μάθει υπολειπόμενες χαρτογραφήσεις, διευκολύνοντας έτσι την εκπαίδευση ενός Βαθύ Νευρωνικού Δικτύου. Το ResNet-50 αποτελείται από 50 επίπεδα, συμπεριλαμβανόμενων και των συνελικτικών επιπέδων, της κανονικοποίησης παρτίδας (batch normalization) και

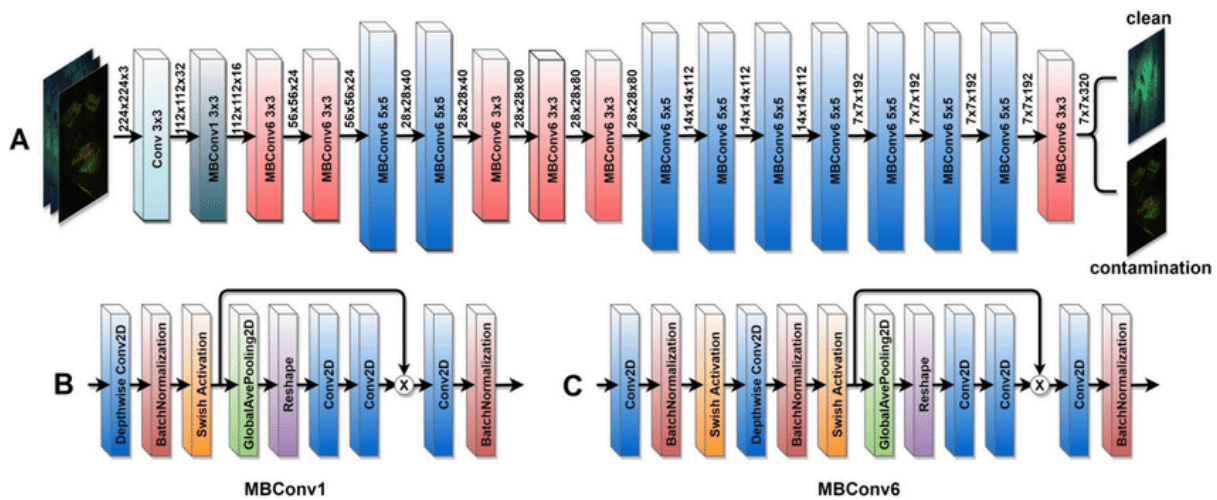
⁶ Εικόνες όπου αντί για 3 κανάλια χρώματος (RGB) έχουμε 1 κανάλι χρώματος, όπου κάθε τιμή εκφράζει την ένταση της φωτεινότητας στο συγκεκριμένο pixel, με την τιμή 0 να εκφράζει το μαύρο και την τιμή 255 το άσπρο.

συναρτήσεις ενεργοποίησης ReLU [11]. Στην Εικόνα 3 φαίνεται η αρχιτεκτονική του μοντέλου ResNet-50.



Εικόνα 3 - Αρχιτεκτονική ResNet-50

EfficientNet-Bo - Στην Εικόνα 4 φαίνεται η αρχιτεκτονική του μοντέλου EfficientNet-Bo. Το μοντέλο αυτό είναι μια αρχιτεκτονική Συνελκτικού Νευρωνικού Δικτύου η οποία κλιμακώνει τις διαστάσεις του ύψους, του πλάτους και της ανάλυσης της εικόνας χρησιμοποιώντας μια συλλογή από σταθερούς συντελεστές. Το μοντέλο αυτό αξιοποιεί έναν συνδυασμό δομικών στοιχείων τα οποία στοχεύουν στην εύρεση χαρακτηριστικών με έναν αποτελεσματικό τρόπο. Συγκεκριμένα, αποτελείται από MBConv Blocks [12], επίπεδα Squeeze and Excitation (SE) [13], και Swish συναρτήσεις ενεργοποίησης⁷ [14].



Εικόνα 4 - Αρχιτεκτονική EfficientNet-Bo

⁷ $f(z) = z * \text{sigmoid}(z)$

DenseNet121 - Το μοντέλο αυτό είναι μία αρχιτεκτονική Συνελικτικού Νευρωνικού Δικτύου η οποία έχει μια πυκνή δομή μπλοκ, με κάθε επίπεδο να συνδέεται με όλα τα υπόλοιπα επίπεδα με έναν ανατροφοδοτικό τρόπο. Επίσης, χρησιμοποιεί επίπεδα Bottleneck για να ελαχιστοποιήσει τις παραμέτρους χωρίς να μειώνει τα χαρακτηριστικά που μαθαίνει το Νευρωνικό Δίκτυο. Το μοντέλο αυτό περιέχει 121 επίπεδα τα οποία συντελούν στην ισορροπημένη σχέση μεταξύ υπολογιστικής ισχύος και απόδοσης [15]. Η αρχιτεκτονική του μοντέλου DenseNet121 φαίνεται στην παρακάτω Εικόνα 5.



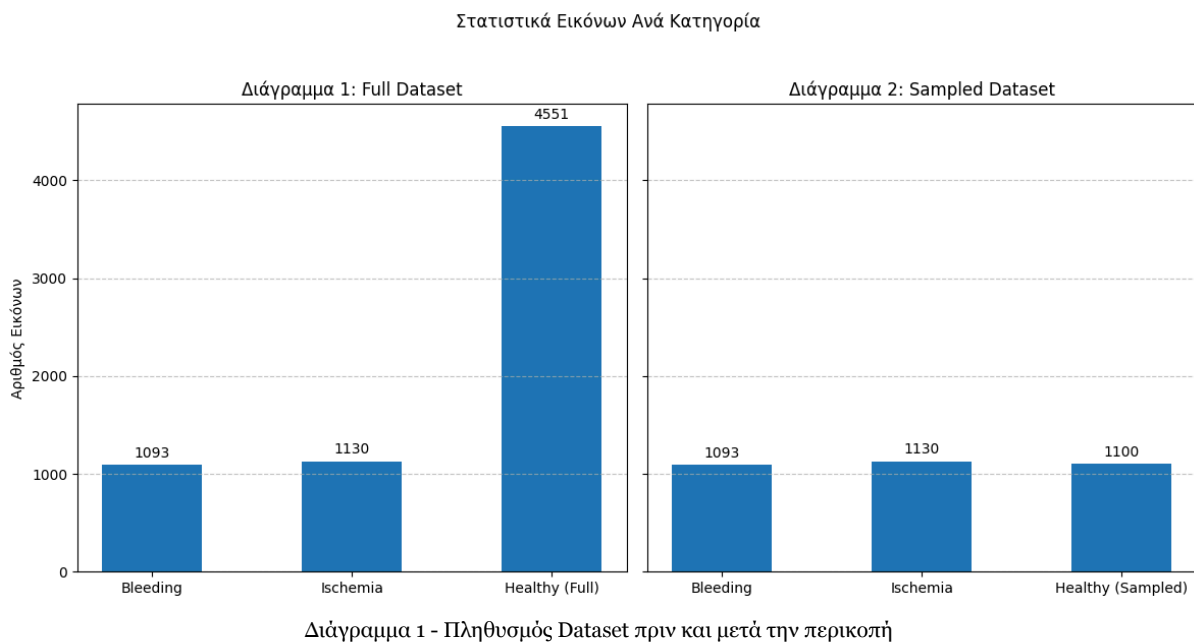
Εικόνα 5 - Αρχιτεκτονική DenseNet121

4. Πειραματικά Αποτελέσματα

Στην ενότητα αυτή θα αναλύσουμε την απόδοση των μοντέλων που υλοποιήσαμε για το πρόβλημά μας, παρουσιάζοντας διάφορα γραφήματα και συγκρίνοντάς τα με βάση διάφορες μετρικές.

4.1. Επεξήγηση του Dataset

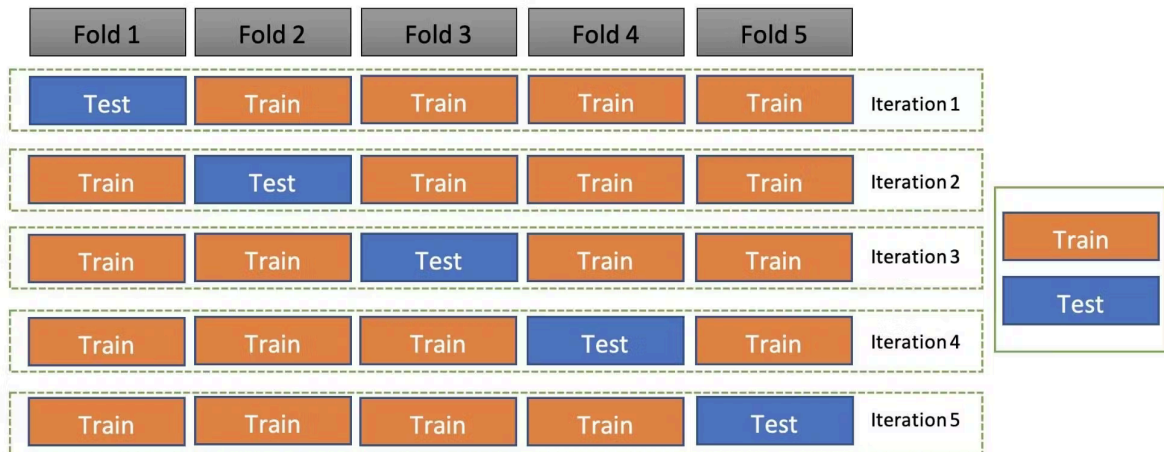
Το dataset που χρησιμοποιήσαμε περιέχει συνολικά 6774 εικόνες (χρησιμοποιώντας αποκλειστικά τις αρχικές εικόνες, όπως αναφέρθηκε στην εισαγωγή). Από το σύνολο των εικόνων αυτών, 1093 αποτελούν περιπτώσεις αιμορραγίας, 1130 περιπτώσεις ισχαιμίας, ενώ οι υπόλοιπες 4551 εικόνες αφορούν υγιείς εγκεφάλους. Είναι φανερό το γεγονός ότι υπάρχει μεγάλη ανισορροπία μεταξύ των υγιών ασθενών με αυτούς που πάσχουν από κάποια πάθηση. Για τον λόγο αυτό και για την ομαλή εκπαίδευση των μοντέλων, περιορίσαμε τις εικόνες της υγιούς κλάσης που αυτά δέχονται ώστε να συμβαδίζουν με τις υπόλοιπες κλάσεις. Ορίσαμε το κατώφλι στις 1100 φωτογραφίες, με την επιλογή τους να γίνεται με τυχαίο τρόπο. Το νέο σύνολο των εικόνων ανέρχεται στις 3323 (Διάγραμμα 1).



Πλέον, μπορούμε να κάνουμε τον διαχωρισμό των δεδομένων σε Train, Validation και Test sets. Για να επιτευχθεί αυτός ο διαχωρισμός, αξιοποιήσαμε την τεχνική Stratified k-fold Cross Validation [16] (Εικόνα 6). Συγκεκριμένα, χρησιμοποιήσαμε 5 folds ($k = 5$), όπου το κάθε ένα από αυτά αντιστοιχεί στο 20% του συνόλου των δεδομένων. Επομένως, για κάθε περίπτωση το Train set, το οποίο μεταφράζεται σε ένα fold από τα πέντε, αποτελείται από το 20% των εικόνων, με το υπόλοιπο 80% να χωρίζεται σε Train και Validation sets. Από αυτό το ποσοστό, το

85% αφορά το Train set, ενώ το 15% το Validation set. Δηλαδή, το dataset συνολικά χωρίζεται ως εξής:

- Train set: 68%
- Validation Set: 12%
- Test Set: 20%



Εικόνα 6 - Stratified k-fold Cross Validation

4.2. Βελτιστοποίηση Υπερπαραμέτρων

Η βελτιστοποίηση των υπερπαραμέτρων στην παρούσα έρευνα έγινε με εμπειρικό τρόπο, καθώς δεν βρέθηκε κάποια τεχνική αυτοματοποίησης της διαδικασίας αυτής πάνω στο πρόβλημά μας. Οι αρχιτεκτονικές που χρησιμοποιήσαμε έχουν τις εξής υπερπαραμέτρους: Learning Rate, Batch Size, Dropout Rate, Patience και Number of Workers. Ωστόσο, στην αρχιτεκτονική ResNet-50 αξιοποιήσαμε και την υπερπαραμέτρο pseudo-RGB, διότι δεν δέχεται δεδομένα σε grayscale. Έτσι, σε αυτό το μοντέλο πραγματοποιούμε κανονικοποίηση περνώντας τρεις φορές την ίδια τιμή ως παράμετρο.

Ακόμα, στο μοντέλο EfficientNet-B0, προσθέσαμε δύο regularizers, ώστε να αποτρέψουμε το overfit του μοντέλου. Αναλυτικότερα, η τεχνική Weight Decay (L2)⁸ αποσκοπεί στην μείωση των βαρών όσο εκπαιδεύεται το μοντέλο, ώστε να μην υπάρχουν αυστηρά σημεία απόφασης. Παράλληλα, η τεχνική Lasso (L1)⁹ παροτρύνει το μοντέλο να αγνοήσει τις λιγότερο σημαντικές πληροφορίες, με σκοπό την μείωση του θορύβου.

Συγκεκριμένα για το Learning Rate, αφότου γίνει η αρχικοποίηση των τιμών, τα μοντέλα χρησιμοποιούν τον Adam¹⁰, ο οποίος αποτελεί έναν optimizer που

⁸ $\text{loss} += \lambda * \sum ||w||^2$

⁹ $\text{loss} += \lambda * \sum |w|$

¹⁰ Adaptive Moment Estimation

τροποποιεί τις τιμές αυτές αξιοποιώντας το Momentum και το RMSProp¹¹ [17]. Κατά αυτόν τον τρόπο εξασφαλίζεται η βελτίωση των μοντέλων κατά την διάρκεια της εκπαίδευσης, ενώ παράλληλα ελαχιστοποιεί την ανάγκη για fine-tuning των παραμέτρων. Στον παρακάτω πίνακα φαίνεται η διακύμανση των τιμών του Learning Rate για κάθε μοντέλο.

Hyperparameter	ResNet-50	EfficientNet-B0	DenseNet121
Learning Rate	0.000160 - 0.000001	0.000200 - 0.000002	0.000200 - 0.000001

Πίνακας 1 - Διακύμανση τιμών του Learning Rate

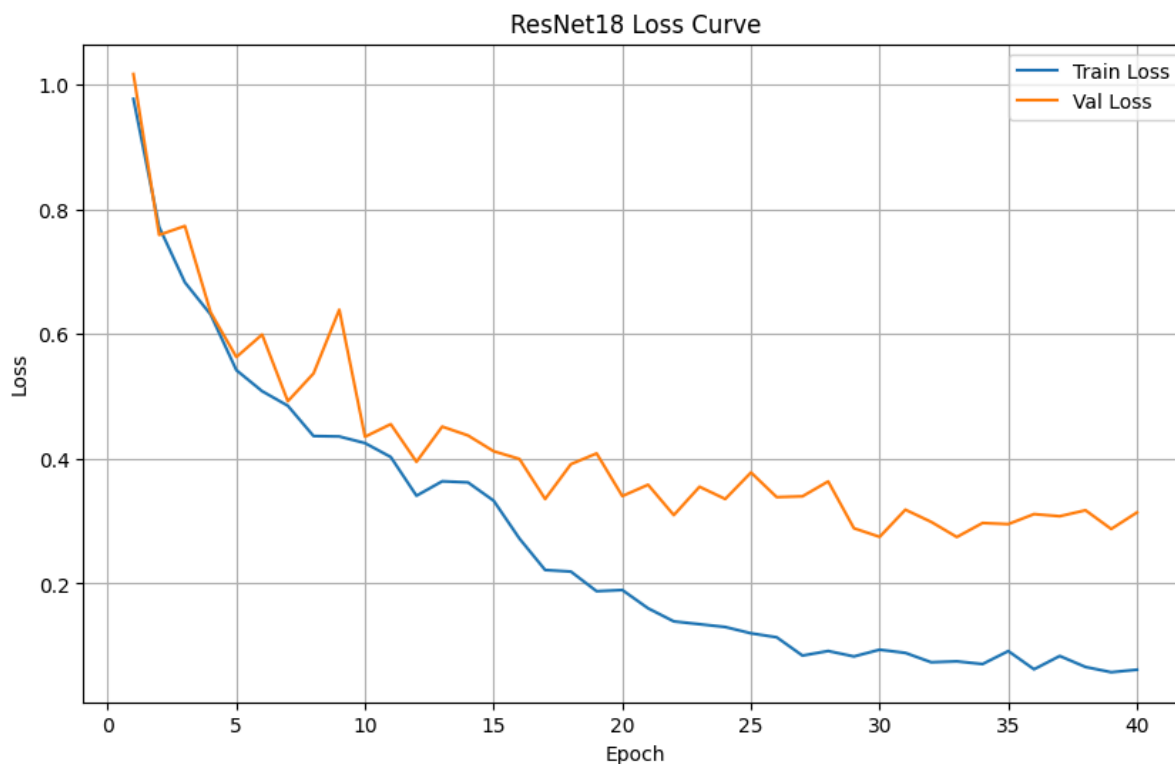
4.3. Εκπαίδευση των μοντέλων

Στην συνέχεια αναλύεται η διαδικασία της εκπαίδευσης των μοντέλων που υλοποιήσαμε χρησιμοποιώντας τις τεχνικές που αναφέρθηκαν παραπάνω. Η εκπαίδευση έγινε στην δωρεάν έκδοση του Google Collab, αξιοποιώντας την παροχή μιας GPU για την γρήγορη παραγωγή αποτελεσμάτων. Το backbone που αξιοποιείται σε όλα τα παρακάτω μοντέλα είναι το ImageNet.

4.3.1. ResNet-50

Αρχίζοντας με την αρχιτεκτονική ResNet, η πρώτη έκδοση που χρησιμοποιήσαμε ήταν η ResNet-18, η οποία αποτελεί μια απλούστερη εκδοχή της ResNet-50, καθώς διαθέτει μόλις 18 επίπεδα. Παρά τα ικανοποιητικά αποτελέσματα που παρείχε εκ πρώτης όψεως, σύντομα έγινε αντιληπτό πως το πρόβλημά που αντιμετωπίζουμε απαιτεί μια πιο σύνθετη αρχιτεκτονική για την ορθότερη λειτουργία του συστήματός. Στο Διάγραμμα 2 φαίνεται το Loss Curve της αρχιτεκτονικής ResNet-18.

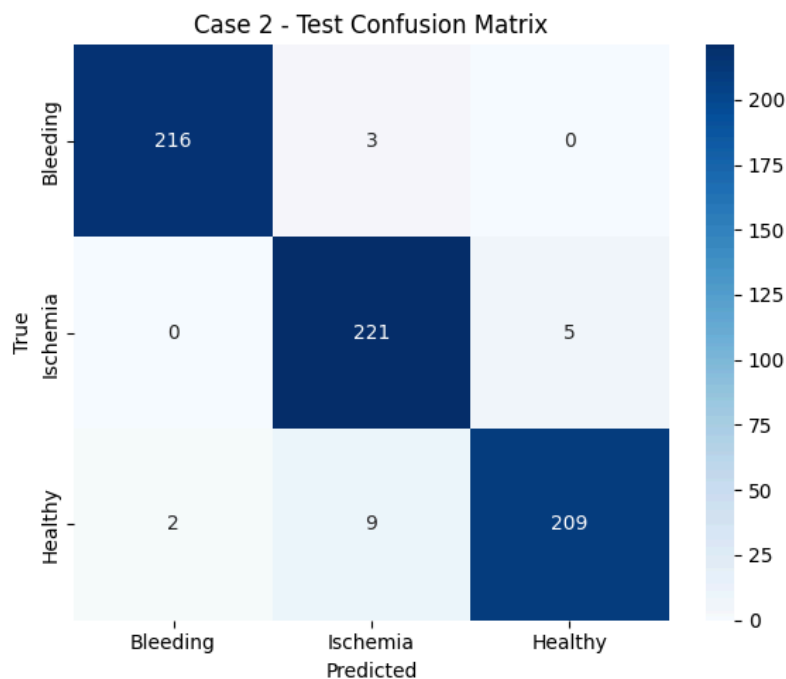
¹¹ Root Mean Square Propagation



Διάγραμμα 2 - Loss Curve για την αρχιτεκτονική ResNet-18

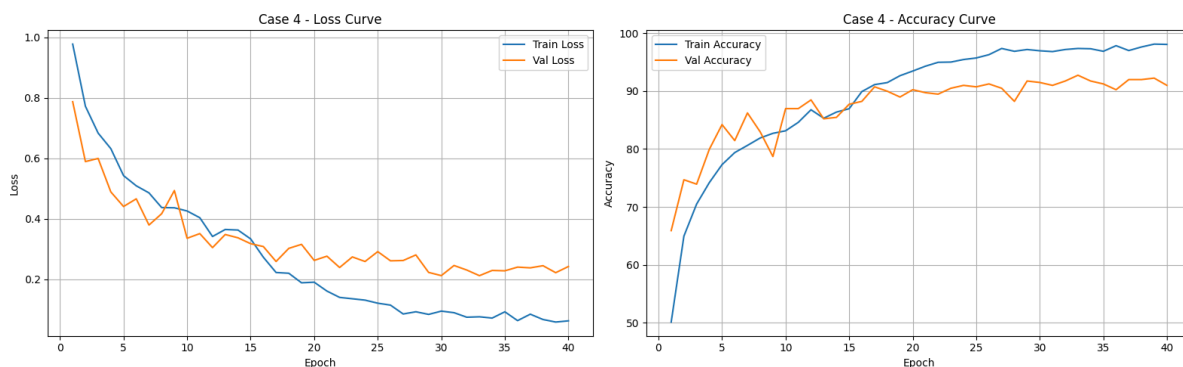
Όπως αναφέρθηκε στις προηγούμενες ενότητες, για την εκπαίδευση του ResNet-50 έγιναν οι απαραίτητες τροποποιήσεις στα δεδομένα, τόσο για να την αποφυγή του overfit, όσο και για την μεγαλύτερη ακρίβεια στα αποτελέσματα του μοντέλου. Η εκπαίδευση κάθε fold πραγματοποιήθηκε για 40 εποχές (epochs)¹², με το μοντέλο πάνω στο fold set 3 να αναδεικνύεται ως το καλύτερο. Στο Διάγραμμα 3 μπορούμε να διακρίνουμε το συγκεκριμένο confusion matrix, το οποίο είναι πλούσιο στην κύρια διαγώνιό του, γεγονός που επιβεβαιώνει την σωστή λειτουργία του μοντέλου.

¹² Μία πλήρης πέραση της εκπαίδευσης του μοντέλου πάνω στο dataset ή σε κάποιο fold αυτού.



Διάγραμμα 3 - Confusion Matrix για το μοντέλο ResNet-50

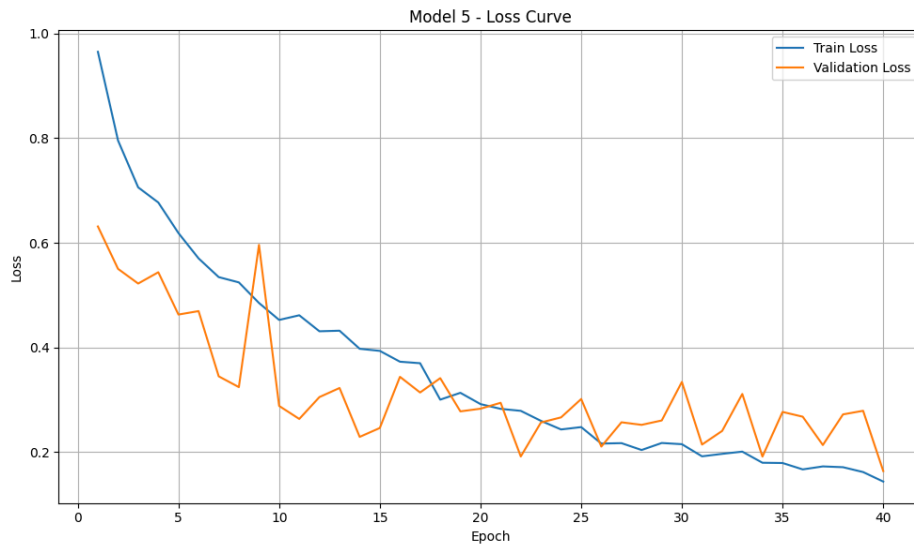
Στο ακόλουθο Διάγραμμα 4, φαίνεται το Loss Curve μαζί με το Accuracy που πέτυχε το μοντέλο αυτό στην ίδια περίπτωση. Μπορούμε να διακρίνουμε ότι η καμπύλη του Loss Curve έχει το επιθυμητό σχήμα, διατηρώντας μια ομαλή σύγκλιση προς το 0 κατά την διάρκεια της εκπαίδευσης. Παρομοίως, η καμπύλη του Accuracy συγκλίνει προς το 100 επιβεβαιώνοντας την σωστή λειτουργία του μοντέλου.



Διάγραμμα 4 - Loss και Accuracy Curves για το μοντέλο ResNet-50

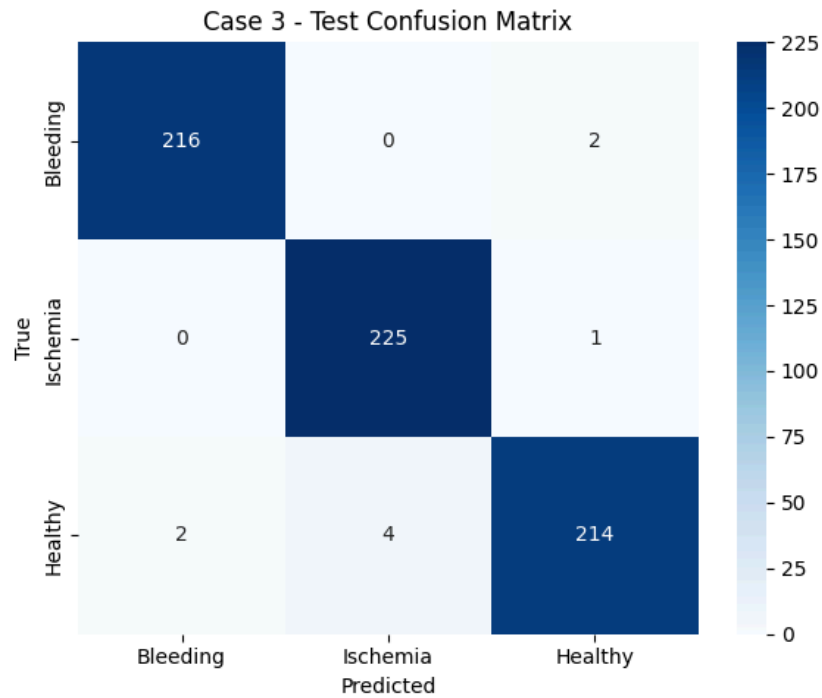
4.3.2. EfficientNet-B0

Για την εκπαίδευση του EfficientNet-B0 ακολουθήσαμε την ίδια φιλοσοφία με το ResNet-50, με την κύρια διαφορά να βρίσκεται στην προσθήκη των regularizers που αναφέρθηκαν παραπάνω. Αυτή η υλοποίηση έγινε διότι το EfficientNet-B0 δεν προσφέρει μία τόσο σύνθετη αρχιτεκτονική όσο το ResNet-50, επομένως ήταν απαραίτητη η χρήση των L1 και L2 ώστε να μην οδηγηθεί το μοντέλο σε overfit. Στο παρακάτω διάγραμμα 5 παρατηρούμε το Loss Curve που παρουσίασε το μοντέλο πριν την παραπάνω τροποποίηση.



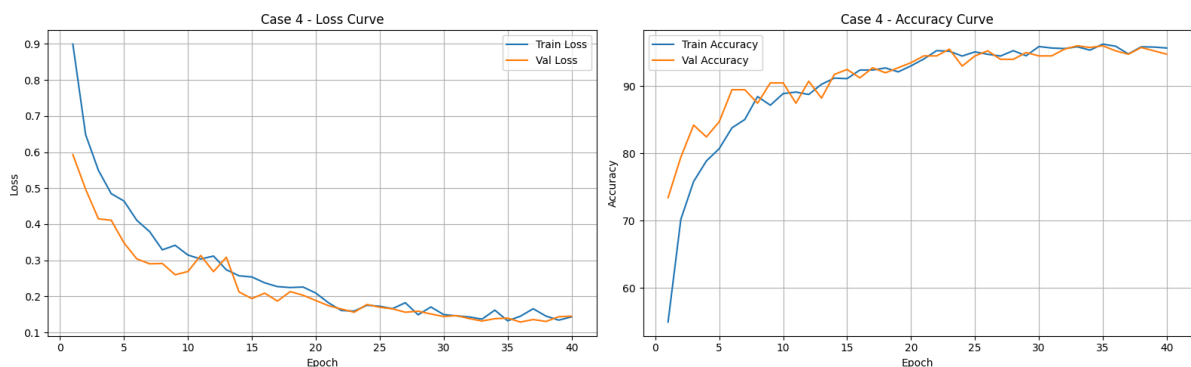
Διάγραμμα 5 - Loss Curve για το μοντέλο EfficientNet-B0 πριν την προσθήκη Regularizers

Όπως και στο προηγούμενο μοντέλο, έτσι και σε αυτό η εκπαίδευση διήρκησε για 40 εποχές, με την τέταρτη περίπτωση διάταξης των fold να αναδεικνύεται ως η καλύτερη εκδοχή. Στο Διάγραμμα 6 φαίνεται το confusion matrix για την συγκεκριμένη περίπτωση, το οποίο έχει σημειώσει άρτια αποτελέσματα, με την κύρια διαγώνιο να είναι αρκετά έντονη.



Διάγραμμα 6 - Confusion Matrix για το μοντέλο EfficientNet-Bo

Στο Διάγραμμα 7 γίνεται μελέτη μιας διαφορετικής περίπτωσης, της περίπτωσης 5. Το συγκεκριμένο μοντέλο παρουσίασε εξαιρετικά θετικές επιδόσεις, με την καμπύλη του Train Loss να είναι σχεδόν ταυτόσημη με αυτή του Validation Loss μετά τις 20 εποχές. Όπως είναι φυσικό, την ίδια μοίρα ακολούθησαν και οι καμπύλες του Accuracy, έχοντας παρόμοια πορεία από ακόμα νωρίτερες εποχές.

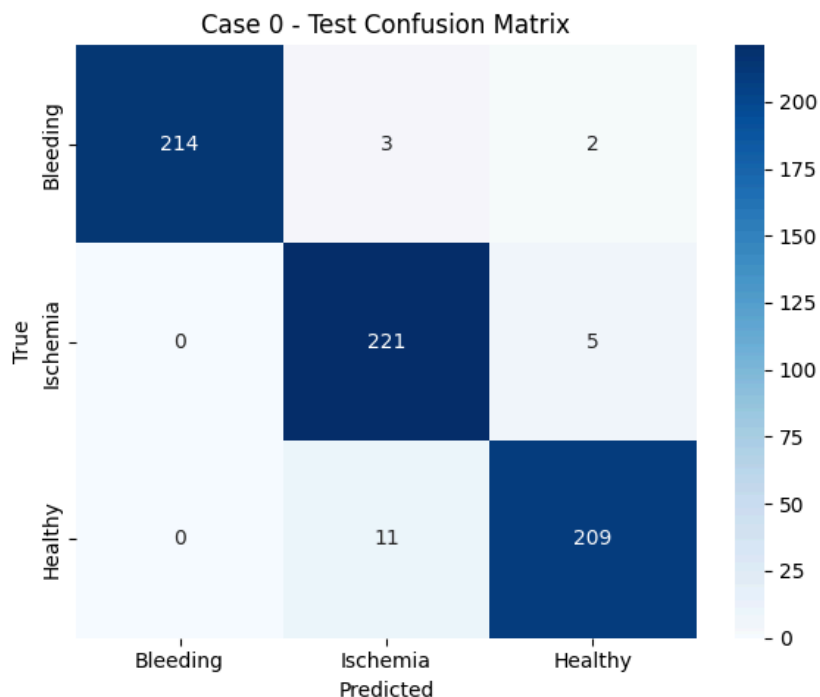


Διάγραμμα 7 - Loss και Accuracy Curves για το μοντέλο EfficientNet-Bo

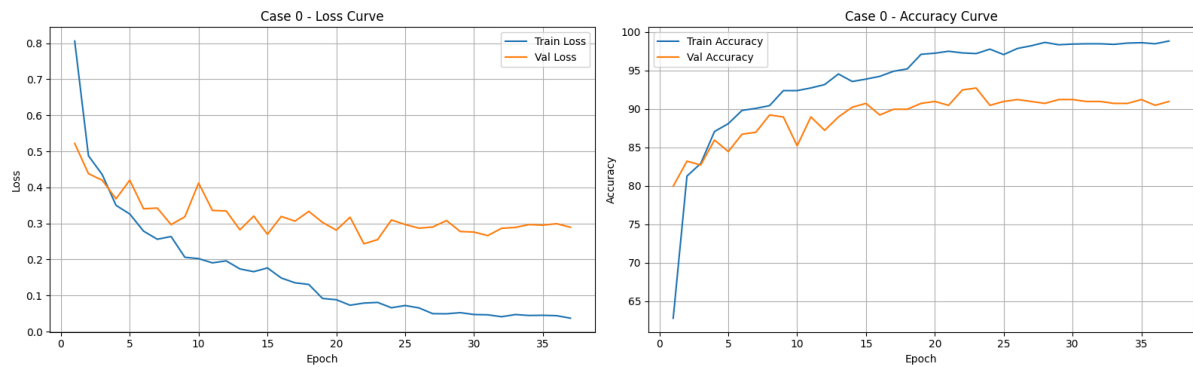
4.3.3. DenseNet121

Το τελευταίο μοντέλο που υλοποιήσαμε για την αντιμετώπιση του προβλήματος αξιοποιεί την αρχιτεκτονική DenseNet121. Η αρχιτεκτονική αυτή, όπως και η ResNet, δέχεται ως είσοδο εικόνες σε RGB μορφή. Ωστόσο, για την διαφοροποίηση των αποτελεσμάτων σε σχέση με το ResNet-50, αντί να χρησιμοποιήσουμε την τεχνική pseudo-RGB, απλώς αλλάξαμε το πρώτο επίπεδο της αρχιτεκτονικής ώστε να δέχεται μοναδική είσοδο. Επομένως, πλέον γίνονται δεκτές και εικόνες σε grayscale.

Η εκπαίδευση για άλλη μια φορά διήρκησε 40 εποχές. Όσον αφορά το Confusion Matrix του DenseNet121, τα αποτελέσματα ήταν παρόμοια με αυτά των προηγούμενων μοντέλων, με την κύρια διαγώνιο να είναι βαθύτατα εμπλουτισμένη, με ελάχιστες αστοχίες. Ωστόσο, δεν διαπιστώθηκε η ίδια επιτυχία στα Loss και Accuracy Curves για το μοντέλο αυτό, καθώς είναι εμφανές ότι υπάρχει μεγάλη απόκλιση μεταξύ της καμπύλης για τα Train δεδομένα με αυτά για το Validation. Επομένως, γίνεται αμέσως αντιληπτό ότι το DenseNet121 είναι αρκετά επιρρεπές σε overfit. Τα διαγράμματα 8 και 9 απεικονίζουν τις παραπάνω παρατηρήσεις.



Διάγραμμα 8 - Confusion Matrix για το μοντέλο DenseNet121



Διάγραμμα 9 - Loss και Accuracy Curves για το μοντέλο DenseNet121

Σε αυτό το σημείο πρέπει επίσης να σημειωθεί το γεγονός πως σε όλα τα μοντέλα αναπτύξαμε και έναν μηχανισμό πρόωρου τερματισμού (Early Stopping), ο οποίος τερματίζει την εκπαίδευση του μοντέλου σε περίπτωση που το τελευταίο δεν παρουσιάσει σημαντική βελτίωση για 5 συνεχόμενες εποχές ή εάν υπάρξει απροσδόκητη αύξηση στο Validation Loss. Για να έχει ισχύ ο παραπάνω μηχανισμός, απαραίτητη προϋπόθεση είναι να έχουν συμπληρωθεί τουλάχιστον 10 εποχές.

Τέλος, κατά την εκπαίδευση των μοντέλων παρατηρήσαμε πως υπήρχε μεγάλη σύγχυση στις κλάσεις Ισχαιμία και Υγής, συνεπώς μειώσαμε τα βάρη της κλάσης Αιμορραγία κατά 20% ώστε να εξισορροπήσουμε την εστίαση των μοντέλων μεταξύ των κλάσεων.

4.4. Αποτελέσματα

Εφόσον έχει ολοκληρωθεί η εκπαίδευση των μοντέλων, το επόμενο βήμα είναι να αναλύσουμε τις επιδόσεις τους με βάση ορισμένες μετρικές. Συγκεκριμένα, οι μετρικές τις οποίες θα αξιοποιήσουμε είναι το Accuracy, το Precision, το Recall και το F1 Score. Στον παρακάτω πίνακα 2 μπορούμε να δούμε τους τύπους των παραπάνω μετρικών.

Μετρική	Τύπος
Accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$
Precision	$\frac{TP}{TP + FP}$
Recall	$\frac{TP}{TP + FN}$
F1 Score	$\frac{2TP}{2TP + FP + FN}$

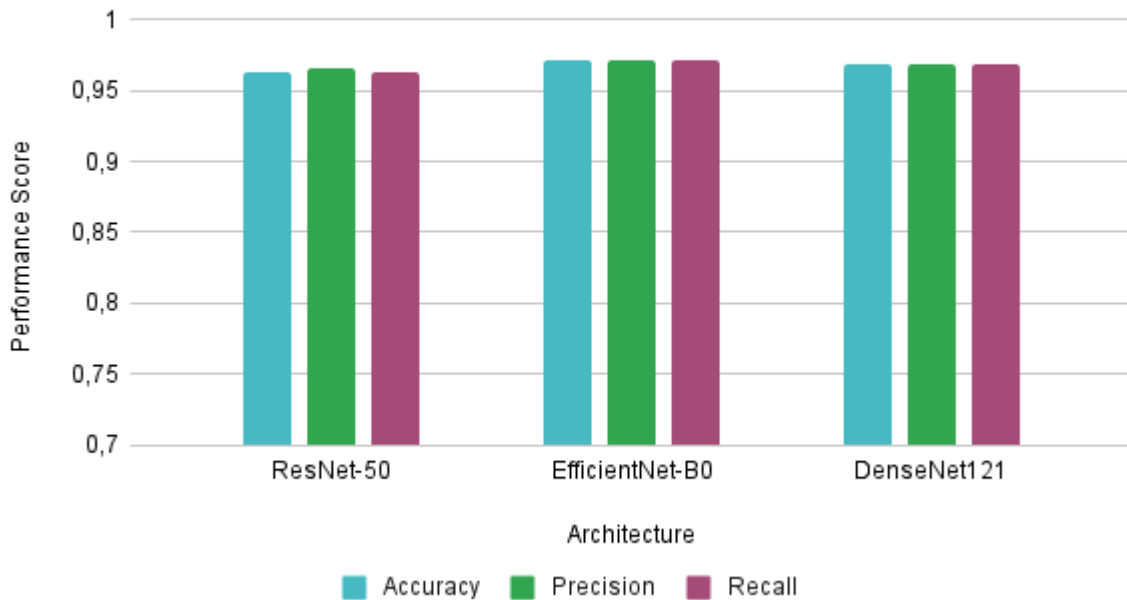
Πίνακας 2 - Μαθηματικοί τύποι των μετρικών Accuracy, Precision, Recall και F1 Score

Στον πίνακα 3 καταγράφονται οι τιμές του μέσου όρου όλων των περιπτώσεων κάθε μοντέλου για τις τρεις πρώτες μετρικές πάνω στο Test set. Είναι εμφανές ότι όλες οι αρχιτεκτονικές αποδίδουν ικανοποιητικά, με τις καλύτερες επιδόσεις να τις πετυχαίνει η αρχιτεκτονική EfficientNet-Bo. Στην συνέχεια ακολουθεί το αντίστοιχο διάγραμμα για την καλύτερη οπτικοποίηση των δεδομένων (Διάγραμμα 10).

Μοντέλο	Accuracy	Precision	Recall
ResNet-50	0.962	0.966	0.962
EfficientNet-Bo	0.971	0.972	0.971
DenseNet121	0.968	0.969	0.968

Πίνακας 3 - Αποδόσεις των μοντέλων για τις μετρικές Accuracy, Precision και Recall

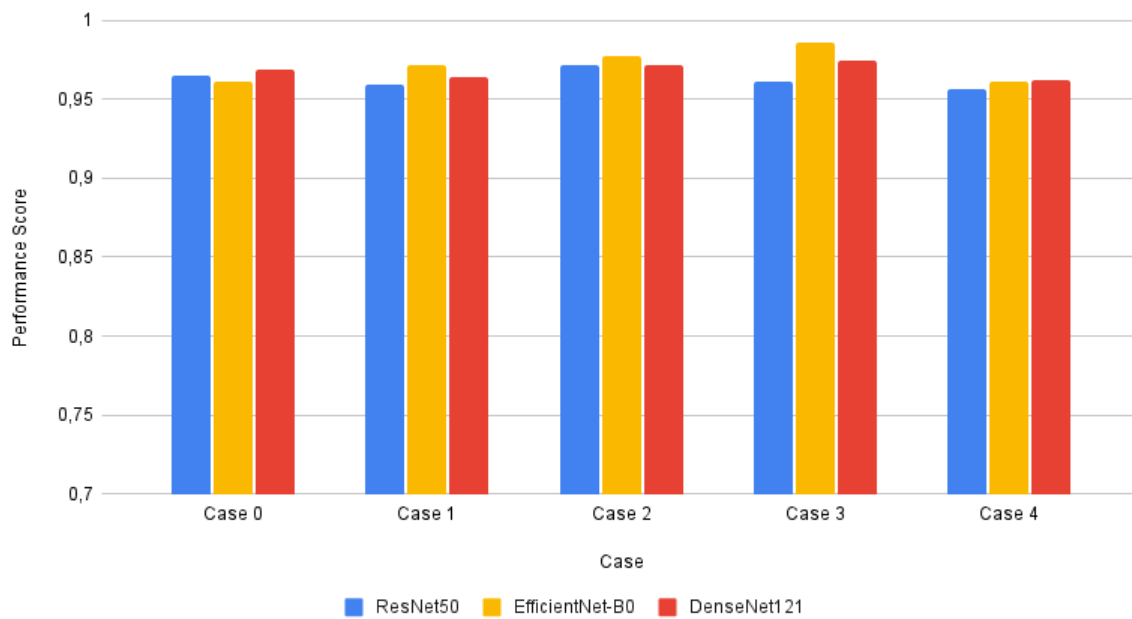
Metrics Score for each Architecture



Διάγραμμα 10 - Αποδόσεις των μοντέλων για τις μετρικές Accuracy, Precision και Recall

Βέβαια, σε τέτοιου είδους προβλήματα ταξινόμησης, η μετρική που έχει την μεγαλύτερη βαρύτητα δεν είναι άλλη από το F1 Score. Συνεπώς, ήταν συνετό να μελετήσουμε αναλυτικά τις επιδόσεις των μοντέλων μας στην μετρική αυτή. Στο Διάγραμμα 11 αποτυπώνονται οι τιμές του F1 Score κάθε αρχιτεκτονικής για κάθε fold set στο οποίο εξετάστηκε. Είναι εμφανές ότι για ακόμη μια φορά, το EfficientNet-B0 σημειώνει τις καλύτερες επιδόσεις, με τα υπόλοιπα μοντέλα να παρουσιάζουν εξίσου καλά αποτελέσματα.

Average F1 Score per Model for each case



Διάγραμμα 11 - Απόδοση του κάθε μοντέλου σε όλες τις περιπτώσεις για την μετρική F1 Score

5. Συμπεράσματα

Τελικά, είναι αντιληπτό ότι με βάση τις μετρικές που αναλύσαμε προηγουμένως, και συγκεκριμένα την μετρική F1 Score, το μοντέλο που αποδίδει καλύτερα πάνω στο συγκεκριμένο πρόβλημα είναι το EfficientNet-B0. Αυτό ίσως είναι αναμενόμενο, καθώς ο τρόπος με τον οποίο κλιμακώνεται ισορροπημένα το βάθος, το πλάτος, αλλά και η ανάλυση του μοντέλου, το καθιστά ικανό να εξάγει τα σωστά χαρακτηριστικά χωρίς υπερβολική αύξηση της πολυπλοκότητας. Αντίθετα, αν και τα μοντέλα ResNet-50 και DenseNet121 παρουσίασαν ικανοποιητικά αποτελέσματα, δεν μπόρεσαν να συναγωνιστούν με το EfficientNet-B0. Ειδικά στην περίπτωση του DenseNet121, το Διάγραμμα 9 καθιστά εμφανές το γεγονός ότι υπάρχει χώρος για βελτίωση όσον αφορά την εφαρμογή του μοντέλου σε πιο γενικευμένα δεδομένα.

Κλείνοντας, αξίζει να αναφερθεί πως η παρούσα έρευνα αποτελεί ένα δυνατό θεμέλιο για την επίλυση του προβλήματος αναγνώρισης εγκεφαλικών όγκων. Ωστόσο, είναι απαραίτητο να διεξαχθεί περαιτέρω μελέτη πάνω στο συγκεκριμένο θέμα ώστε να επιτευχθεί η βέλτιστη απόδοση των μοντέλων και να ελαχιστοποιηθεί η πιθανότητα λάθους κατά την διαδικασία αυτή.

6. Βιβλιογραφία

1. *IRM Sud-Ouest: Why Getting MRI Results Early is Important.*
<https://mobi.irmsudouest.com/why-getting-mri-results-early-is-important>
2. *The Medical Imaging Technology Association (MITA): About DICOM: Overview.*
<https://www.dicomstandard.org/about>
3. Xiaoyi L., Wang Z. (2024): *Deep Learning in Medical Image Classification from MRI-based Brain Tumor Images.*
<https://arxiv.org/abs/2408.00636>
4. Mohamed M. M., Mahesh T. R., Vinoth K. V., Guluwadi S. (2024): *Enhancing brain tumor detection in MRI images through explainable AI using Grad-CAM with Resnet 50.*
<https://bmcmedimaging.biomedcentral.com/articles/10.1186/s12880-024-01292-7>
5. Bhandari M. (2024): *Grad-CAM: A Beginner's Guide.*
<https://medium.com/@bmuskan007/grad-cam-a-beginners-guide-adf68e8of4bb>
6. Hossain M., Rahman A. M. (2023): *Brain Tumor Detection Using Deep Network EfficientNet-Bo.*
<https://www.researchgate.net/publication/371479936>
7. Bernardo V., Attoresì M., Lareo X., Velasco L. (2023): *Explainable Artificial Intelligence.*
https://www.edps.europa.eu/system/files/2023-11/23-11-16_techdispatch_xai_en.pdf
8. Zebari N., Marqas R., Salih S. M., Alkurdi A. A. H. (2023): *Enhancing Brain Tumor Classification with Data Augmentation and DenseNet121.*
<https://www.researchgate.net/publication/374746973>
9. *Geeks For Geeks (2024): Gaussian Noise.*
<https://www.geeksforgeeks.org/gaussian-noise/>
10. *IBM: What are convolutional neural networks?*
<https://www.ibm.com/think/topics/convolutional-neural-networks>

11. *Product Teacher: ResNet18 & ResNet50 in Computer Vision*
<https://www.productteacher.com/quick-product-tips/resnet18-and-resnet50>
12. Sandler M., Howard A., Zhu M., Zhmoginov A., Chen A. (2018): *MobileNetV2: Inverted Residuals and Linear Bottlenecks*
<https://arxiv.org/pdf/1801.04381v4>
13. Hu J., Shen L., Albanie S., Sun G., Wu E. (2018): *Squeeze-and-Excitation Networks*
<https://arxiv.org/pdf/1709.01507v4>
14. Tan M., Le Q. V. (2019): *EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks*.
<https://arxiv.org/pdf/1905.11946>
15. Das S. (2023): *Implementing DenseNet-121 in PyTorch: A Step-by-Step Guide*.
<https://medium.com/deepkapha-notes/implementing-densenet-121-in-pytorch-a-step-by-step-guide-c0c2625c2a60>
16. Lyashenko V., Jha A. (2025): *Cross-Validation in Machine Learning: How to Do It Right*.
<https://neptune.ai/blog/cross-validation-in-machine-learning-how-to-do-it-right>
17. Agarwal R. (2023): *Complete Guide to the Adam Optimization Algorithm*.
<https://builtin.com/machine-learning/adam-optimization>