



**ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ**

**ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ  
ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ**

**« Ταξινόμηση ιστολογικών εικόνων οστεοσαρκώματος με τη  
χρήση μηχανικής μάθησης »**

**Εξαμηνιαία Εργασία**

**στο μάθημα «Εργαστήριο Βιοϊατρικής Τεχνολογίας»**

**των φοιτητών**

**Αβραμίδης Σταύρος , Α.Μ.: 03117811  
Ηλιόπουλος Ανδρέας , Α.Μ.: 03120815  
Κανδρέλη Παρασκευή Μαρίνα, Α.Μ.: 03120047  
Κανταρτζής Βασίλειος, Α.Μ.: 03119935**

**Διδάσκοντες: Κ. Νικήτα, Δ. Κουτσούρης, Ο. Πετροπούλου**

**Αθήνα, Ιούνιος 2023**



## Περίληψη

Η παρούσα εργασία παρουσιάζει μια προσέγγιση για την ταξινόμηση ιστολογικών εικόνων οστεοσαρκώματος σε τρεις διαφορετικές κατηγορίες: Μη βιώσιμος όγκος, βιώσιμος όγκος και μη όγκος. Το οστεοσάρκωμα, ένας πρωτογενής κακοήθης όγκος των οστών, αποτελεί σημαντική πρόκληση στην ακριβή ταξινόμηση λόγω των ποικίλων και πολύπλοκων μορφολογικών χαρακτηριστικών του. Για την αντιμετώπιση αυτού του προβλήματος, αξιοποιούμε τη δύναμη των Συνελκτικών Νευρωνικών Δικτύων (CNN) και συγκεκριμένα χρησιμοποιούμε την αρχιτεκτονική EfficientNet-B7. Η μελέτη ξεκίνησε με την λήψη του dataset που αποτελείται από μια συλλογή ιστολογικών H&E απεικονίσεων οστεοσαρκώματος που λαμβάνονται από Πανεπιστήμιο του Ντάλας (UT Southwestern/UT Dallas). Το σύνολο δεδομένων επιμελείται σχολαστικά, εξασφαλίζοντας μια ισορροπημένη εκπροσώπηση των διαφόρων κατηγοριών ενδιαφέροντος. Εφαρμόζονται τεχνικές προεπεξεργασίας, η αλλαγή μεγέθους και η επαύξηση, για να βελτιωθεί η ποιότητα και η ποικιλομορφία του συνόλου δεδομένων. Στη συνέχεια, το προ-εκπαιδευμένο μοντέλο EfficientNet-B7 τελειοποιείται στο σύνολο δεδομένων για το οστεοσάρκωμα, επιτρέποντάς του να μάθει σημαντικά χαρακτηριστικά και πρότυπα που σχετίζονται με διαφορετικά χαρακτηριστικά όγκων. Η προτεινόμενη προσέγγιση αξιολογείται με τη χρήση αυστηρών πειραματικών μεθοδολογιών, συμπεριλαμβανομένης της πίνακα σύγχυσης και μετρικών επιδόσεων όπως η ακρίβεια, η ακρίβεια, η ευαισθησία και το F1-score. Τα αποτελέσματα καταδεικνύουν την ανώτερη απόδοση της μεθόδου μας στην ακριβή κατηγοριοποίηση εικόνων οστεοσαρκώματος σε κλάσεις Μη Όγκου, Μη-Βιώσιμου Όγκου και Βιώσιμου Όγκου. Συνολικά, η παρούσα έρευνα παρουσιάζει μια αποτελεσματική προσέγγιση για την ταξινόμηση εικόνων οστεοσαρκώματος με τη χρήση του EfficientNet-B7, με τη δυνατότητα βελτίωσης της έγκαιρης ανίχνευσης και της ακριβούς ταξινόμησης, οδηγώντας τελικά σε βελτιωμένα αποτελέσματα για τους ασθενείς στον τομέα της διάγνωσης και της θεραπείας του οστεοσαρκώματος.

## Λέξεις Κλειδιά

Osteosarcoma, histopathology, machine learning, deep learning, convolutional neural networks, transfer learning



# Πίνακας Περιεχομένων

<b>1</b>	<b>Εισαγωγή.....</b>	<b>8</b>
1.1	Οστεοσάρκωμα .....	8
1.1.1	Επιδημιολογία .....	8
1.1.2	Τοποθεσίες όγκου.....	9
1.1.3	Παθολογία.....	9
1.1.4	Ιστολογία και ιστολογική απεικόνιση .....	9
1.2	Σχετική έρευνα .....	14
1.3	Πρακτικό ενδιαφέρον .....	16
1.3.1	Ανάγκη για αυτοματοποίηση.....	16
1.3.2	Απόπειρες αυτοματοποίησης .....	17
1.4	Προκλήσεις .....	17
<b>2</b>	<b>Υλικό και Μέθοδοι.....</b>	<b>19</b>
2.1	Νευρωνικά Δίκτυα.....	19
2.1.1	Τεχνητά Νευρωνικά Δίκτυα.....	19
2.1.2	Συνελικτικά Νευρωνικά Δίκτυα - CNN.....	20
2.1.3	Αρχιτεκτονική Συνελικτικών Νευρωνικών Δικτύων.....	21
2.1.4	Convolutional Layer.....	22
2.1.5	Pooling Layer.....	23
2.1.6	Fully Connected Layer.....	23
2.1.7	Transfer Learning.....	23
2.1.8	EfficientNet CNN.....	24
2.2	Σύνολο δεδομένων προς ανάλυση .....	26
2.3	Πειραματική διάταξη.....	27
<b>3</b>	<b>Αποτελέσματα.....</b>	<b>30</b>

3.1	Οπτικοποίηση Προόδου Εκπαίδευσης και Αξιολόγηση Απόδοσης .....	30
3.1.1	<i>Loss and Accuracy per Epoch</i> .....	30
3.1.2	<i>Test Set Analysis</i> .....	31
3.2	Δείκτες αξιολόγησης .....	32
3.2.1	<i>Ανάλυση Πίνακα Σύγχυσης</i> .....	34
<b>4</b>	<b>Συμπεράσματα - Επίλογος</b> .....	<b>36</b>
4.1	Συμπεράσματα της Εργασίας .....	36
4.2	Μελλοντικές Κατευθύνσεις Έρευνας.....	37
<b>5</b>	<b>Βιβλιογραφία</b> .....	<b>40</b>



# 1

## Εισαγωγή

### 1.1 Οστεοσάρκωμα

Ο όρος οστεοσάρκωμα (ή οστεογενές σάρκωμα) αναφέρεται σε νεοπλάσματα που, ιστολογικά, παρουσιάζουν παραγωγή οστών και οστεοειδών με κακοήγη μεσεγγυματικά κύτταρα και αποτελεί τον συχνότερο κακοήγη όγκο των οστών[1]. Η επιθετική του φύση και συστηματική τάση για πρώιμες μεταστάσεις καθιστούν την έρευνα για τρόπους γρήγορης και ακριβούς ταξινόμησης απεικονίσεων του όγκου πολύτιμη για την ιατρική κοινότητα. Η δυσκολία στην έγκυρη ταυτοποίηση και ταξινόμησή του έγκειται στη μεγάλη ιστολογική διαφοροποίηση που παρουσιάζεται τόσο μεταξύ υποκατηγοριών του, αλλά ακόμα και μεταξύ ανατομιών του ίδιου ασθενούς. Η διαδικασία ταξινόμησης και ανάλυσης των ιστών, εκτός από το ότι είναι ενεργοβόρα για τους παθολόγους που την διεξάγουν, πολλές φορές προβαίνει ανακριβής, καθώς συχνά υπάρχουν διαφορετικά συμπεράσματα μεταξύ παθολόγων για την ίδια απεικόνιση[2]. Σε αυτή την εργασία, θα αποπειραθούμε να αυτοματοποιήσουμε τη διαδικασία και να αυξήσουμε τα ποσοστά ακρίβειας κάνοντας χρήση Συνελικτικών Νευρωνικών Δικτύων (CNN) για την ταξινόμηση απεικονίσεων σε βιώσιμο όγκο (Viable), μη-βιώσιμο (Non-Viable) και μη-καρκινικό ιστό (Non-Tumor)[3].

#### 1.1.1 Επιδημιολογία

Το οστεοσάρκωμα αποτελεί ασθένεια που εμφανίζεται κυρίως σε εφήβους και παιδιά, αλλά υπάρχει και εμφάνιση περιστατικών σε άτομα ηλικίας άνω των 60[4]. Συγκεκριμένα, αποτελεί τον τρίτο συχνότερο καρκίνο στις εφηβικές και προεφηβικές ηλικίες -με τα λεμφώματα και τους εγκεφαλικούς όγκους να είναι οι μόνοι συχνότεροι-, ενώ σε ηλικίες κάτω των 5 ετών είναι πολύ σπάνιος[1]. Το οστεοσάρκωμα στις εφηβικές ηλικίες χαρακτηρίζεται ως πρωτογενές, εμφανίζεται συχνότερα σε άρρενες με ποσοστό 60% και συνδέεται με την αυξημένη δραστηριότητα στα κέντρα ανάπτυξης του οστού[5]. Το δευτερογενές οστεοσάρκωμα, που εμφανίζεται σε αρκετά



μεγαλύτερες ηλικίες, συνήθως συνοδεύεται από προϋπάρχουσες ασθένειες όπως η Paget's disease και άλλες οστικές βλάβες.

Επιδημιολογικά, η ηλικία του ασθενούς συνδέεται με την πιθανότητα επιβίωσης, με τους νεαρότερους ασθενείς να έχουν την υψηλότερη. Καθώς στο παρελθόν η θεραπεία συνίσταντο μόνο από ακρωτηριασμό, το ποσοστό πενταετούς επιβίωσης ήταν 20%. Ωστόσο, με την εισαγωγή της ακτινοθεραπείας και την συνεχή καινοτομία στις μεθόδους θεραπείας, το ποσοστό αυξήθηκε σταθερά κατά 1.3% ετησίως, φτάνοντας σήμερα περίπου το 68%, ανεξαρτήτως φύλου[6].

### **1.1.2 Τοποθεσίες όγκου**

Το πρωτογενές οστεοσάρκωμα, από τη μία πλευρά, εμφανίζεται κατά κύριο λόγο σε μακρά οστά των άκρων, στις μεταφύσεις δίπλα στις αναπτυξιακές πλάκες, με πιθανότητα εμφάνισης έως και 60%. Αυτή η περιγραφή περιλαμβάνει δομές όπως: άπω μηριαίο οστό, πρόσθια κνήμη και πρόσθιο βραχιόνιο. Το δευτερογενές οστεοσάρκωμα, από την άλλη, εμφανίζεται συνήθως στο γνάθο, την πύελο και τη σπονδυλική στήλη<sup>1</sup>. Λιγότερο συχνές περιοχές αποτελούν τα μικρά οστά των χεριών και των ποδιών ή ακόμα και μαλακά μόρια (εξωσκελετικό οστεοσάρκωμα).

### **1.1.3 Παθολογία**

Εξ ορισμού, το οστεοσάρκωμα είναι ένας κακοήθης όγκος του οποίου τα κύτταρα παράγουν οστά ή οστεοειδή. Ο όγκος αρχικά εξαπλώνεται εντός του μυελού, έπειτα διεισδύει στον οστικό φλοιό και, αφού τον καταστρέφει εν μέρει, επεκτείνεται κάτω από το περίοστεο, ενίοτε ακόμα και σε μαλακά μόρια[7].

Το οστεοσάρκωμα χωρίζεται, σύμφωνα με την ιστολογική ταξινόμηση οστικών όγκων του Παγκόσμιου Οργανισμού Υγείας, σε κεντρικό ή ενδομυελικό (central/ intramedullary), που αποτελεί το 80% των περιπτώσεων, επιφανειακό (surface), που αποτελεί το 10-15% και άλλες σπανίζουσες κατηγορίες, όπως το εξωσκελετικό (extraskeletal) οστεοσάρκωμα που συνιστά <1% των περιπτώσεων[6]. Οι σημαντικότερες υποκατηγορίες κεντρικού οστεοσάρκωματος είναι οι εξής: Συμβατικό (conventional), Τηλαγγειεκτατικό (telangiectatic), Μικροκυτταρικό (small-cell) και Χαμηλού βαθμού (low-grade). Οι σημαντικότερες δε κατηγορίες επιφανειακού είναι Παραοστεϊκό (Parosteal), Περιοστεϊκό (Periosteal) και Ενδοφλοιϊκό (Intracortical). Περισσότερες κατηγορίες μπορούν να φανούν στον πίνακα 1.

### **1.1.4 Ιστολογία και ιστολογική απεικόνιση**

Καθώς η πλειονότητα της βιβλιογραφίας, συμπεριλαμβανομένου και του συνόλου δεδομένων μας, μελετά ιστολογικές απεικονίσεις χρωματισμένες με αιματοξυλίνη και ηωσίνη (H&E) -ουσίες οι οποίες βάφουν τους πυρήνες μπλε και τους ιστούς του υποβάθρου ροζ [1], θα ασχοληθούμε μόνο με H&E εικόνες.

---

<sup>1</sup> <https://www.pathologyoutlines.com/topic/boneosteosarcomageneral.html>

**Table 1**  
**Osteosarcoma Types**

---

Central
High-grade
Conventional
Telangiectatic
Small cell
Epithelioid
Osteoblastoma-like
Chondroblastoma-like
Fibrohistiocytic
Giant cell-rich
Low-grade
Low-grade central
Fibrous dysplasia-like
Desmoplastic fibroma-like
Surface
Low-grade
Parosteal
Intermediate-grade
Periosteal
High-grade
Dedifferentiated parosteal
High-grade surface
Intracortical
Gnathic
Extraskeletal
High-grade
Low-grade

---

**Πίνακας 1: Υποκατηγορίες οστεοσαρκώματος [8]**

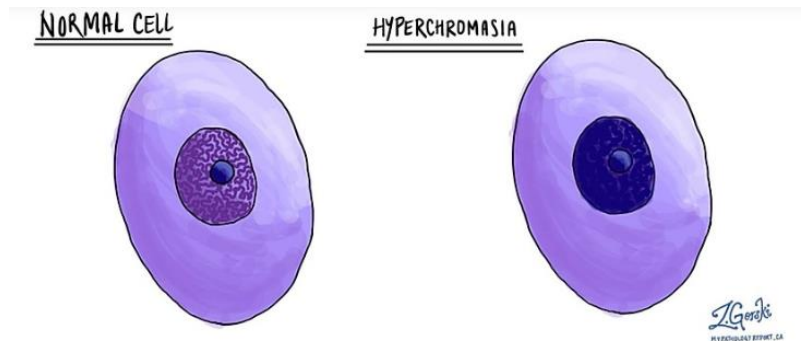
#### 1.1.4.1 Υποκατηγορίες οστεοσαρκώματος

Αν και χρησιμοποιούνται διάφορες διαγνωστικές μέθοδοι για την αναζήτηση του οστεοσαρκώματος, η ιστολογική απεικόνιση είναι απαραίτητη για την διάγνωσή του και την ταξινόμηση των κυττάρων. Λόγω του μεγάλου πλήθους υποκατηγοριών, αλλά και της δυσκολίας σαφούς διάγνωσης λόγω αυτής της τάσης για εκτενή υποκατηγοριοποίηση, η λεπτομερής ταξινόμηση συνήθως έχει σημασία μόνο σε περιπτώσεις όπου ένας συγκεκριμένος υπότυπος συστηματικά ανταποκρίνεται στη θεραπεία με διαφορετικό τρόπο. Τα ιστολογικά χαρακτηριστικά των κυριότερων υποκατηγοριών είναι [8]:

- Συμβατικό οστεοσάρκωμα:

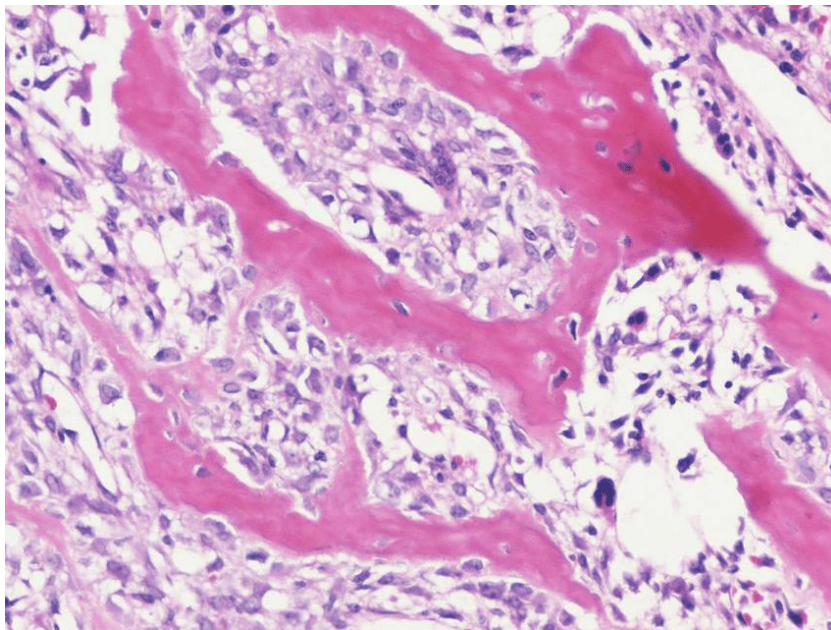
Αποτελεί την πιο εύκολα διαγνώσιμη υποκατηγορία οστεοβλαστώματος. Τα καρκινικά κύτταρα έχουν πληθώρα διαφορετικών σχημάτων, από ατρακτοειδή έως και πολυεδρικά, και οι πυρήνες τους είναι πλειόμορφοι (παρουσιάζουν πολλές διαφοροποιήσεις σε μέγεθος, σχήμα ή/και χρώμα)

και υπερχρωματικοί (μετά από χρώση με H&E, οι πυρήνες φαίνονται πιο σκούροι από το κανονικό σε εξέταση υπό το μικροσκόπιο, βλ. εικόνα 1).



**Εικόνα 1:** Παράδειγμα υπερχρωματικού πυρήνα<sup>2</sup>

Προϋπόθεση για τη διάγνωση όγκου ως συμβατικό οστεοσάρκωμα είναι η παραγωγή οστού ή οστεοειδούς από κάποια από τα καρκινικά κύτταρα του όγκου.



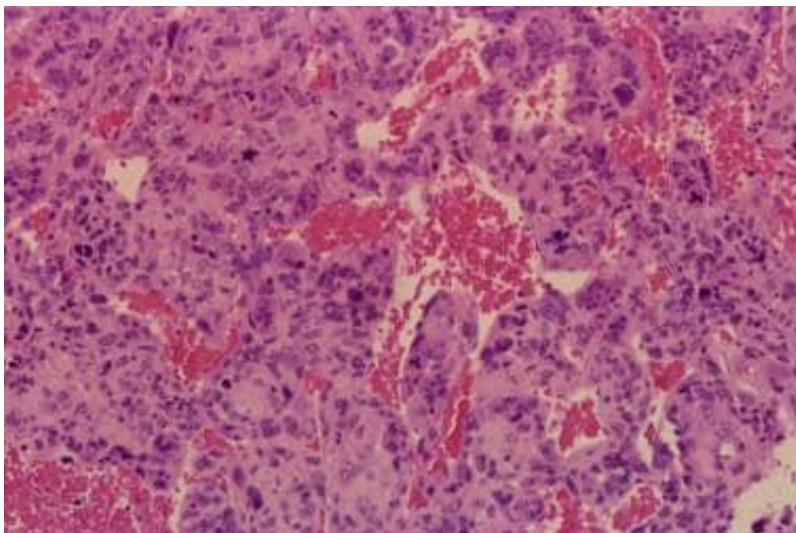
**Εικόνα 2:** Ιστολογική απεικόνιση συμβατικού οστεοσαρκώματος [9]

---

<sup>2</sup> <https://www.mypathologyreport.ca/pathology-dictionary/definition-hyperchromatic/>

- Τηλαγγειεκτατικό οστεοσάρκωμα:

Η ιστολογική απεικόνιση του τηλαγγειεκτατικού οστεοσαρκώματος (TOS) χαρακτηρίζεται από διασταλμένες κοιλότητες γεμάτες με αίμα και κυτταρικά κύτταρα υψηλού βαθμού (στα μεσοδιποκυττάρια διαφράγματα και στο περιφερικό χείλος). Το γεγονός αυτό συχνά προκαλεί σύγχυση μεταξύ του TOS και των ανευρυσματικών κυστών οστού, που έχουν παρόμοια απεικόνιση, οδηγώντας ενίοτε σε λανθασμένη διάγνωση[6].



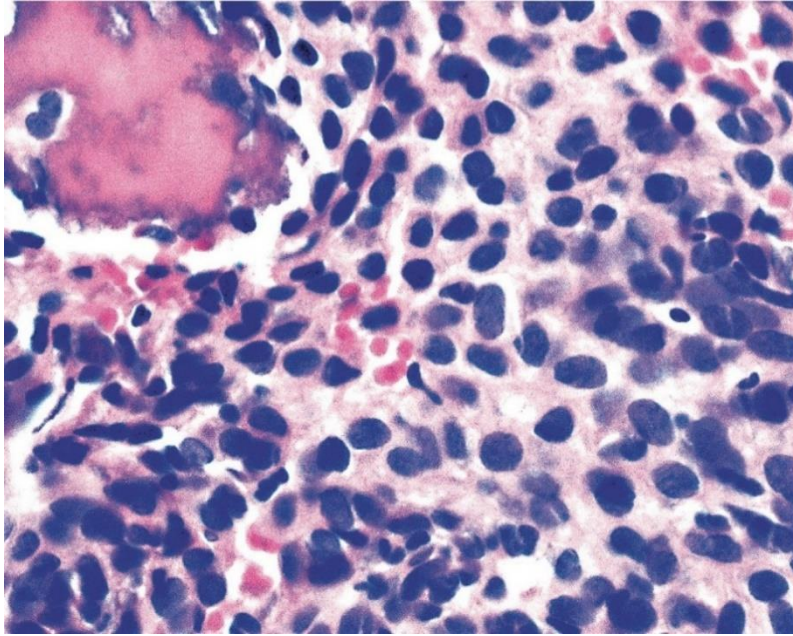
*Εικόνα 3: Ιστολογική απεικόνιση τηλαγγειεκτατικού οστεοσαρκώματος<sup>3</sup>*

- Μικροκυτταρικό οστεοσάρκωμα:

Ιστολογικά, το μικροκυτταρικό οστεοσάρκωμα παρουσιάζει κύτταρα μικρά σε μέγεθος, με στρογγυλούς υπερχρωματικούς πυρήνες που έχουν μικρές διαφοροποιήσεις στην εμφάνιση, σε αντίθεση με το συμβατικό οστεοσάρκωμα. Η εμφάνιση αυτή είναι παρόμοια με του Ewing's sarcoma, το οποίο όμως δεν έχει την παραγωγή οστού ή οστεοειδών ως χαρακτηριστικό γνώρισμα, διευκολύνοντας τον διαχωρισμό των δύο κατηγοριών.[8]

---

<sup>3</sup> <https://emedicine.medscape.com/article/1257219-workup>



*Εικόνα 4: Ιστολογική απεικόνιση μικροκυτταρικού οστεοσαρκώματος<sup>4</sup>*

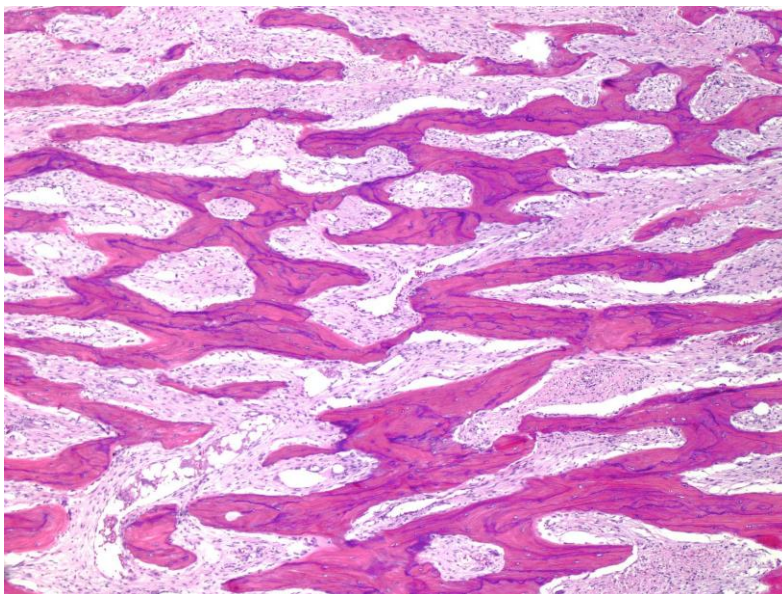
- Παραοστεϊκό οστεοσάρκωμα:

Η ιστολογική απεικόνιση του παραοστεϊκού OS εμφανίζει ταινίες οστικών δοκίδων με υψηλό βαθμό παράλληλου προσανατολισμού, παρόμοιο με αυτό που μπορεί να παρατηρηθεί σε μια περιοστεϊκή αντίδραση νέου οστού (παραγωγή νέου οστού ως αντίδραση σε τραυματισμό ή άλλα ερεθίσματα).

---

<sup>4</sup> <https://www.pathologyoutlines.com/wick/osteosarcoma%20small%20cell%20type%20micro02.jpg>





*Εικόνα 5: Ιστολογική απεικόνιση παραοστεϊκού οστεοσαρκώματος<sup>5</sup>*

## **1.2 Σχετική έρευνα**

Υπάρχει εκτενής σχετική βιβλιογραφία ερευνητών που ασχολήθηκαν με την ταξινόμηση ιστολογικών απεικονίσεων οστεοσαρκώματος με διαφορετικές προσεγγίσεις. Συγκεκριμένα, οι Z. Ling et al. (2022) ανέπτυξαν έναν αλγόριθμο Βαθιάς Μηχανικής Μάθησης που βασίζεται στην Βελτιστοποίηση του Συνόλου Δεδομένων (Data Set Optimization) μέσω ενός αλγορίθμου Mean-Teacher, στην Προ-επεξεργασία (Pretreatment) των εικόνων μέσω δυαδοποίησής τους ενώ τέλος επικεντρώνεται στην Ανάλυση και Τμηματοποίηση Εικόνας (Image Analysis and Segmentation) με χρήση των μεθόδων SepUNet (Deep Separable U-Shaped Network) και CRF (Conditional RandomField). [10] Τέλος, ο αλγόριθμος υπολογίζει τους δείκτες αξιολόγησης (Precision, Recall, F1 Score, IOU, DSC), οι οποίοι λαμβάνουν υψηλότερες τιμές στο προτεινόμενο από το άρθρο μοντέλο CNN συγκριτικά με άλλα γνωστά μοντέλα CNN (βλ. Πίνακα 2)

---

<sup>5</sup> <https://www.pathologyoutlines.com/topic/boneparostealosteo.html>

TABLE 3: Comparison of performance of different methods in MRI images of patients with osteosarcoma.

Model	Pr	Re	F1	IOU	DSC	Params	FLOPs
FCN-16s	0.922	0.882	0.900	0.824	0.859	134.3M	190.35G
FCN-8s	0.941	0.873	0.901	0.830	0.876	134.3M	190.08G
PSPNet	0.856	0.888	0.872	0.772	0.870	49.07M	101.55G
MSFCN [3]	0.881	0.936	0.906	0.841	0.874	20.38M	1524.34G
MSRN [4]	0.893	0.945	0.918	0.853	0.887	14.27M	1461.23G
FPN	0.914	0.924	0.919	0.852	0.888	88.63M	141.45G
UNet	0.922	0.924	0.923	0.867	0.892	17.26M	160.16G
Our (SepUNet)	0.927	0.932	0.930	0.869	0.896	20.32M	199.26G
Our (SepUNet+CRF)	0.939	0.926	0.932	0.874	0.901	20.32M	200.07G
Our (SepUNet+CRF +Prop)	0.937	0.938	0.937	0.883	0.914	20.32M	200.07G

**Πίνακας 2:** Αποτελέσματα σύγκρισης μοντέλων του άρθρου των Z. Ling et al. (2022) [10]

Οι Arunachalam et al. (2017) από την άλλη, δημιούργησαν έναν αυτοματοποιημένο αλγόριθμο με χρήση CNN [11], ο οποίος πραγματοποιεί σε αρχικό στάδιο την ψηφιακή επεξεργασία των MRI εικόνων, την μετατροπή τους σε WSIs και την SVS ανάλυσή τους. Έπειτα επιχειρεί τον συνδυασμό των μεθόδων K-Means Clustering, Otsu Thresholding και Flood-Fill για να επιτύχει την σωστή τμηματοποίηση και οριοθέτηση των περιοχών ενδιαφέροντος. Το μειονέκτημα της παραπάνω διαδικασίας είναι πως λόγω της πληθώρας των τεχνικών που χρησιμοποιήθηκαν υπάρχει σημαντική επιβάρυνση των υπολογιστικών πόρων αλλά και του χρόνου που απαιτείται για να ολοκληρωθεί η διαδικασία επιτυχώς. Οι Gawade et al. (2023) διάλεξαν οι εικόνες να φορτώνονται από το DataSet και έπειτα γίνεται επεξεργασία τους από έναν αυτοματοποιημένο αλγόριθμο Βαθιάς Μηχανικής Μάθησης [12] μέσω των ξεχωριστών τεχνικών μοντέλων δικτύων εικόνας VGG16, VGG19, DenseNet201 και ResNet101 τα οποία και συγκρίνει. Υπολογίζονται οι δείκτες αξιολόγησης (Accuracy, F1, Precision, Recall, AUC and V-score) για να αξιολογηθεί το τελικό μοντέλο που θα χρησιμοποιηθεί βάση της αποδοτικότητάς του στην ταξινόμηση των εικόνων. Εκ των τεσσάρων μοντέλων CNN που συγκρίθηκαν το άρθρο αναφέρει ως πιο αποδοτικό το ResNet101 με τις τιμές Accuracy, F1 Score, Recall, AUC, V-Score να είναι 0.93, 0.89, 0.89, 0.89, 0.95, 2.72 αντίστοιχα.

Σύγκριση διαφορετικών μοντέλων συνελκτικών δικτύων έκαναν επίσης οι Vezakis et al. (2023). Ειδικότερα, ανέπτυξαν μία συγκριτική μεθοδολογική προσέγγιση για την αξιολόγηση διαφόρων δικτύων βαθιάς μάθησης για τη διάγνωση και την ταξινόμηση των όγκων με χρήση των μοντέλων MobileNetV2 και EfficientNetB0, τα οποία κρίθηκαν πως είναι τα πιο αποδοτικά μετά από διαδικασία σύγκρισης[13]. Με την χρήση των μοντέλων αυτών, οι συντάκτες πετυχαίνουν αποτελεσματικότερη ταξινόμηση των μη βιώσιμων όγκων, των βιώσιμων όγκων και των περιοχών νέκρωσης στις νέες υποβαθμισμένης ανάλυσης εικόνες, ενώ ύστερα από επανεκπαίδευση του MobileNetV2 επιτυγχάνεται συνολική ακρίβεια (accuracy) 0.91 και μέση ανάκληση (mean recall) 0.95, 0.93, 0.83 για μη-βιώσιμους όγκους, βιώσιμους όγκους και περιοχές νέκρωσης, αντίστοιχα.

Τέλος, το ερευνητικό άρθρο των Anisuzzaman et al. (2020) εξέτασε επίσης μοντέλα για την ανίχνευση των περιοχών βιώσιμου όγκου[14], μη βιώσιμου όγκου και περιοχών νέκρωσης, συγκεκριμένα τα Transfer Learning μοντέλα VGG-19 και Inception V3, τα οποία εκπαιδεύθηκαν σε εικόνες WSIs, έτσι ώστε να επιτευχθούν υψηλότερα ποσοστά ακρίβειας. Τα μοντέλα εφαρμόζονται σε διάφορα προβλήματα ταξινόμησης, συμπεριλαμβανομένων ταξινομητών δυαδικών και πολλαπλών κατηγοριών. Τα πειραματικά αποτελέσματα αυτού του άρθρου καταλήγουν στο ότι την υψηλότερη απόδοση έχει η ακρίβεια του VGG-19 μοντέλου συνελκτικών δικτύων, 0.96, μεταξύ όλων των δυαδικών και των πολλαπλών κλάσεων ταξινόμησης.

Εξετάζοντας τη βιβλιογραφία, παρατηρήσαμε λοιπόν ότι υπάρχει μία πληθώρα νευρωνικών δικτύων, αλλά και αρχιτεκτονικών συνελκτικών νευρωνικών δικτύων συγκεκριμένα, τα οποία έχουν εφαρμοστεί στην έρευνα για την ταξινόμηση απεικονίσεων οστεοσαρκώματος και από τα οποία μπορούμε να διαλέξουμε για την εργασία μας.

## ***1.3 Πρακτικό ενδιαφέρον***

### ***1.3.1 Ανάγκη για αυτοματοποίηση***

Επί του παρόντος, οι ερευνητές πρέπει να αξιολογούν χειροκίνητα τις ιστολογικές διαφάνειες οστεοσαρκώματος με μικροσκόπιο για να εκτιμήσουν τόσο την έκταση του όγκου όσο και τον βαθμό νέκρωσης του, και αναλόγως με τα συμπεράσματα της μελέτης προτείνεται κατάλληλο πλάνο αντιμετώπισης του όγκου. Μελέτη σχετικά με το νεφρικό κυτταρικό καρκίνωμα διαπίστωσε ότι υπήρχε μεγάλη διαφωνία μεταξύ των παθολόγων, για τα ίδια δείγματα δεδομένων, γεγονός που κατά κόρο οφείλεται στην προαναφερθείσα ετερογένεια των καρκινικών ιστών. Το γεγονός αυτό, σε συνδυασμό με τον μεγάλο βαθμό ετερογένειας μεταξύ καρκινικών ιστών, καθιστά σαφές ότι η χειροκίνητη ταξινόμηση καρκινικών (ή μη) ιστών αποτελεί μια διαδικασία που επιτάσσει εξαιρετικό φόρτο εργασίας και διέπεται από την υποκειμενικότητα του παρατηρητή [14],[15].

Όμως είναι επιθυμητή η ανάπτυξη μιας αυτόματης προσέγγισης για την ταξινόμηση των ιστολογικών διαφανειών οστεοσαρκώματος. Η σάρωση και η ψηφιοποίηση ολόκληρης της ιστολογικής διαφάνειας παρέχει τη δυνατότητα αυτοματοποίησης της διαδικασίας ανάλυσης. Η ψηφιοποίηση των διαφανειών με τον χρωματισμένο ιστό έχει ως αποτέλεσμα μια εικόνα υψηλής ανάλυσης. Αυτές οι ψηφιακές εικόνες ολόκληρων ιστολογικών διαφανειών (WSI) επιτρέπουν την επεξεργασία και ανάλυση εικόνας με τεχνικές που αξιοποιούν τις μορφολογικές και context based ενδείξεις που υπάρχουν στις WSI ως χαρακτηριστικά για την ταξινόμηση των ιστών[15].

Ωστόσο, υπάρχουν αρκετά εμπόδια στην πορεία προς ένα πλήρως αυτοματοποιημένο σύστημα. Οι WSIs είναι μεγάλες ψηφιακές εικόνες υψηλής ανάλυσης που λαμβάνονται από σαρωτές διαφανειών, συχνά μεγέθους αρκετών gigabytes και περιέχουν εκατομμύρια pixels. Η ανάλυση τόσο μεγάλων εικόνων παρουσιάζει προκλήσεις λόγω περιορισμών υλικού[13] κυτταρική μορφολογία έχει ως αποτέλεσμα τη μεταβλητότητα στον ίδιο τύπο κυττάρων και η ομοιότητα σε



διαφορετικές κυτταρικές δομές(π.χ. μεταξύ υγιούς δομής και μη υγιούς) μπορεί να καταστήσει την ταξινόμηση των διαφανειών των κυττάρων αρκετά δύσκολη. Ιδιαίτερα στο οστεοσάρκωμα, τόσο τα ογκολογικά κύτταρα όσο και ορισμένοι τύποι φυσιολογικών κυττάρων (πρόδρομα κύτταρα) χρωματίζονται με το ίδιο μπλε χρώμα, αλλά τα καρκινικά κύτταρα έχουν ακανόνιστο σχήμα, ενώ τα πρόδρομα κύτταρα είναι πιο στρογγυλά, στενά και κανονικά. Επιπλέον, κάθε τύπος όγκου διαφέρει σημαντικά από άλλους τύπους, γεγονός που καθιστά δύσκολη την εφαρμογή μιας πανάκειας μεθόδου που αναπτύχθηκε για έναν τύπο όγκου σε έναν άλλο τύπο όγκου. Κατ' αυτό τον τρόπο λοιπόν, γίνεται σαφές ότι η παροχή ενός μηχανισμού ταξινόμησης υψηλού επιπέδου όχι μόνο δε συνίσταται την αντικατάσταση των ερευνητών/ιατρών, αλλά απαιτεί την παρουσία και τη συνεργασία μαζί τους ώστε να επιτευχθεί μια αξιόπιστη ταξινόμηση[11], [15].

### **1.3.2 Απόπειρες αυτοματοποίησης**

Στο παρελθόν, οι πρώτες μελέτες μηχανικής μάθησης στην ιστοπαθολογία επικεντρώθηκαν σε απλές τεχνικές ανάλυσης εικόνας, όπως η κατωφλίωση και η ανίχνευση ακμών, για τον εντοπισμό συγκεκριμένων δομών. Καθώς η τεχνολογία εξελισσόταν, οι ερευνητές αναζήτησαν πιο εξελιγμένες μεθόδους, όπως η ανάλυση υφής και η αναγνώριση προτύπων, για να βελτιώσουν την ταξινόμηση των WSI[13].

Η μηχανική μάθηση έχει επιφέρει σημαντική πρόοδο στην ταξινόμηση εικόνων. Οι συμβατικοί αλγόριθμοι, όπως οι Μηχανές Διανυσμάτων Υποστήριξης (SVM) και οι Τυχαίες Δυνάμεις (RF), έχουν χρησιμοποιηθεί ευρέως στο παρελθόν, αξιοποιώντας χειροποίητα edited χαρακτηριστικά που εξάγονται από εικόνες. Ωστόσο, η απόδοσή τους περιορίζεται από την ποιότητα αυτών των χαρακτηριστικών, ιδίως σε μεγάλα σύνολα δεδομένων με υψηλή μεταβλητότητα[13].

Αντίθετα, τα μοντέλα βαθιάς μάθησης, όπως τα Συνελικτικά Νευρωνικά Δίκτυα (CNN) και ο Vision Transformer (ViT), έχουν επιτύχει εντυπωσιακά αποτελέσματα στην ταξινόμηση εικόνων, επιτυγχάνοντας συχνά αποτελέσματα συγκρίσιμα με τις ανθρώπινες επιδόσεις. Πρόσφατες αναφορές έχουν αναδείξει την υπεροχή της βαθιάς μάθησης έναντι της συμβατικής μηχανικής μάθησης, ιδίως στον τομέα της μικροσκοπίας[13].

## **1.4 Προκλήσεις**

Η ιστολογική ταξινόμηση ιστολογικών απεικονίσεων οστεοσαρκώματος παρουσιάζει διάφορες δυσκολίες. Αρχικά, μια ιστολογική απεικόνιση οστεοσαρκώματος με χρώση H&E παρουσιάζει επιπλέον δυσκολίες στην ταξινόμηση, καθώς η απεικόνιση μπορεί να περιέχει, πέρα από τα ενεργά καρκινικά κύτταρα, και άλλα είδη ιστών των ανατομικών δομών της περιοχής, όπως ινώδεις ιστοί, αιμοφόρα κύτταρα, ασβεστοποιημένα τμήματα οστών, μυελικά κύτταρα, λιποκύτταρα, οστεοβλάστες, οστεοκλάστες, χόνδρους, πρόδρομα κύτταρα, αυξητικές πλάκες και οστεοειδή, ενώ πιθανή είναι και η ύπαρξη μη βιώσιμων (νεκρωτικών) καρκινικών κυττάρων.[11] Η ύπαρξη αυτών των ιστών δημιουργεί ανομοιογένεια ως προς το χρώμα, το σχήμα, το μέγεθος, την

πυκνότητα και την υφή, ενώ χρωματική ανομοιογένεια μπορεί να υπάρχει και μεταξύ διαφορετικών εικόνων ενός συνόλου δεδομένων λόγω αλλαγών όπως ο φωτισμός. Όλα τα παραπάνω, σε συνδυασμό και με τον υψηλό βαθμό ανομοιογένειας του οστεοσαρκώματος συνεπάγονται αυξημένη περιπλοκότητα στα δεδομένα.

Επιπλέον, ένα ακόμα πρόβλημα συνιστά η περιορισμένη διαθεσιμότητα και ενίοτε ποιότητα των training data. Λόγω του χρονοβόρου και ενεργοβόρου έργου που απαιτείται, όσο για τη δημιουργία, τόσο και για τον σχολιασμό των εικόνων από κλινικούς εμπειρογνώμονες, υπάρχουν περιορισμένα δείγματα προς μελέτη. Επιπροσθέτως, τα δεδομένα που αποκτώνται μπορεί να είναι μη ισορροπημένα, δηλαδή να είναι πολυπληθέστερα σε μία κλάση από ότι σε άλλες, με αποτέλεσμα να είναι μεροληπτικά προς αυτή.

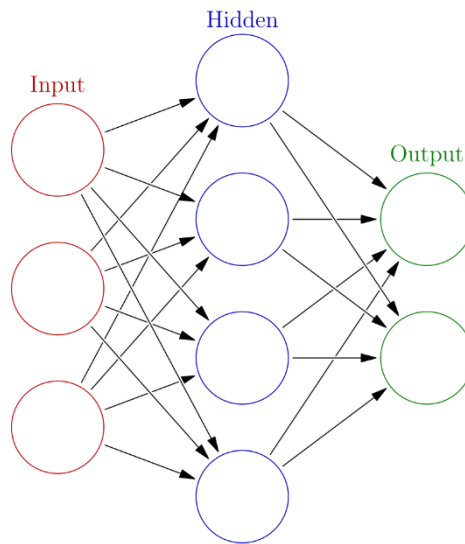
# 2

## *Υλικό και Μέθοδοι*

### *2.1 Νευρωνικά Δίκτυα*

#### *2.1.1 Τεχνητά Νευρωνικά Δίκτυα*

Τα Τεχνητά Νευρωνικά Δίκτυα (Artificial Neural Networks-ANN) είναι υπολογιστικά συστήματα εμπνευσμένα από τον τρόπο λειτουργίας του εγκεφάλου του ανθρώπου. Αποτελούνται από διασυνδεδεμένους κόμβους που ονομάζονται νευρώνες οι οποίοι συνεργάζονται για την επεξεργασία πληροφοριών και τη βελτιστοποίηση της τελικής εξόδου. Τα ANN έχουν συνήθως ένα επίπεδο εισόδου που λαμβάνει πολυδιάστατα διανύσματα, τα οποία στη συνέχεια περνούν από κρυφά επίπεδα. Αυτά τα κρυφά στρώματα λαμβάνουν αποφάσεις με βάση την είσοδο που λαμβάνουν και προσαρμόζουν τις εσωτερικές τους παραμέτρους για να βελτιώσουν την έξοδο, μια διαδικασία γνωστή ως μάθηση. Όταν ένα ANN έχει πολλαπλά κρυφά στρώματα, είναι γνωστό ως μοντέλο βαθιάς μάθησης ("deep learning") [16]. Η σχηματική αναπαράσταση της βασικής αρχιτεκτονικής των ANN φαίνεται στην **Error! Reference source not found..**



*Εικόνα 6: Αυτή η δομή αποτελεί τη βάση της αρχιτεκτονικής των ANN<sup>6</sup>*

### 2.1.2 Συνελκτικά Νευρωνικά Δίκτυα - CNN

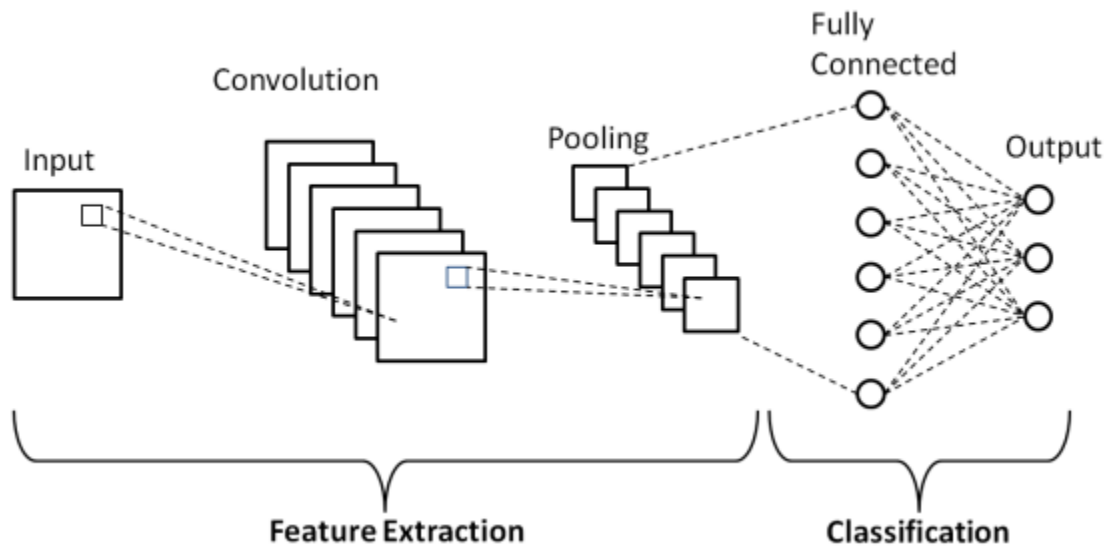
Τα CNN μοιράζονται ομοιότητες με τα παραδοσιακά ANN στο ότι αποτελούνται από νευρώνες που μαθαίνουν και βελτιστοποιούν τον εαυτό τους. Κάθε νευρώνας δέχεται είσοδο και εκτελεί λειτουργίες όπως τα κλιμακωτά γινόμενα που ακολουθούνται από μη γραμμικές συναρτήσεις, οι οποίες αποτελούν τη βάση πολλών ANN. Ξεκινώντας από ακατέργαστα διανύσματα εικόνας ως είσοδο και καταλήγοντας σε βαθμολογίες κλάσεων ως έξοδο, ολόκληρο το δίκτυο αναπαριστά μια ενιαία συνάρτηση αντιληπτικής βαθμολογίας (το βάρος). Το τελικό στρώμα ενός CNN περιέχει συναρτήσεις απωλειών που σχετίζονται με τις διάφορες κλάσεις και οι τεχνικές και οι στρατηγικές που αναπτύχθηκαν για τα παραδοσιακά ANN μπορούν ακόμη να εφαρμοστούν. Η βασική διάκριση μεταξύ των CNN και των παραδοσιακών ANN είναι ότι τα CNN χρησιμοποιούνται κυρίως για εργασίες αναγνώρισης προτύπων που αφορούν εικόνες. Επιτρέπουν στην αρχιτεκτονική να συλλάβει χαρακτηριστικά ειδικά για την εικόνα, όπως το χρώμα και η υφή, καθιστώντας το δίκτυο πιο κατάλληλο για εργασίες που εστιάζουν στην εικόνα, ενώ παράλληλα μειώνουν τον αριθμό των παραμέτρων που απαιτούνται για τη διαμόρφωση του μοντέλου [16].

<sup>6</sup> [https://en.wikipedia.org/wiki/Artificial\\_neural\\_network](https://en.wikipedia.org/wiki/Artificial_neural_network)

### 2.1.3 Αρχιτεκτονική Συνελκτικών Νευρωνικών Δικτύων

Τα CNN αποτελούνται από τρεις τύπους στρώματων: τα συνελκτικά στρώματα, τα στρώματα συγκέντρωσης (pooling) και τα πλήρως συνδεδεμένα στρώματα, τα οποία μαζί αποτελούν την αρχιτεκτονική του CNN[16]. Η αρχιτεκτονική μπορεί να απλοποιηθεί ως εξής:

1. Το στρώμα εισόδου περιέχει τις τιμές των pixel της εικόνας.
2. Στο στρώμα συνελίξεων, οι νευρώνες συνδέονται με τοπικές περιοχές της εισόδου. Η έξοδος προσδιορίζεται με τον υπολογισμό του κλιμακωτού γινομένου μεταξύ των βαρών και του όγκου εισόδου. Μια συνάρτηση ενεργοποίησης εφαρμόζεται στην έξοδο.
3. Το στρώμα pooling εκτελείται υποδειγματοληψία κατά μήκος των χωρικών διαστάσεων της εισόδου, μειώνοντας τον αριθμό των παραμέτρων.
4. Τα πλήρως συνδεδεμένα στρώματα, παρόμοια με τα τυπικά ANN, παράγουν βαθμολογίες κλάσεων για την ταξινόμηση.



**Εικόνα 7:** Αρχιτεκτονική CNN τεσσάρων σταδίων<sup>7</sup>

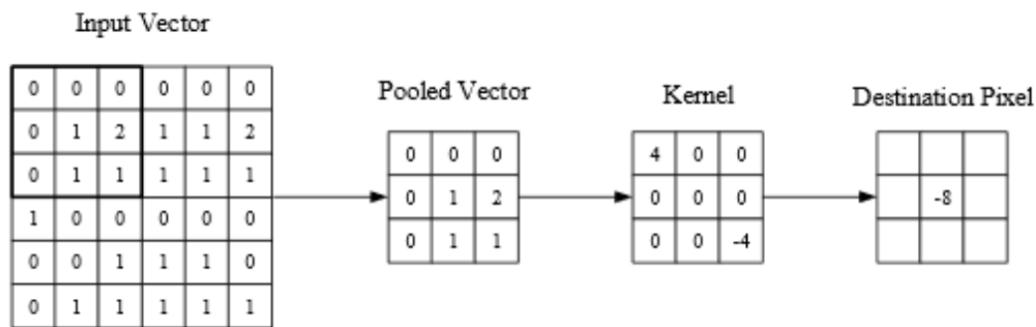
Εφαρμόζοντας τεχνικές συνελκτικής και υποδειγματοληψίας, τα CNN μετασχηματίζουν προοδευτικά το αρχικό στρώμα εισόδου για να παράγουν βαθμολογίες κλάσεων για εργασίες ταξινόμησης και παλινδρόμησης [16].

<sup>7</sup>

[https://www.researchgate.net/figure/Schematic-diagram-of-a-basic-convolutional-neural-network-CNN-architecture-26\\_fig1\\_336805909](https://www.researchgate.net/figure/Schematic-diagram-of-a-basic-convolutional-neural-network-CNN-architecture-26_fig1_336805909)

### 2.1.4 Convolutional Layer

Το στρώμα συνελκτικής ανάλυσης (convolutional layer) είναι ζωτικής σημασίας για τη λειτουργία των CNN. Οι βασικές του παράμετροι περιστρέφονται γύρω από τη χρήση μαθησιακών πυρήνων (kernels), οι οποίοι είναι μικροί σε μέγεθος αλλά καλύπτουν όλο το βάθος της εισόδου. Όταν τα δεδομένα περνούν από ένα επίπεδο συνελκτικής επεξεργασίας, κάθε φίλτρο συνελίσσεται στις χωρικές διαστάσεις της εισόδου, δημιουργώντας έναν χάρτη ενεργοποίησης 2D. Αυτοί οι χάρτες ενεργοποίησης μπορούν να οπτικοποιηθούν, όπως απεικονίζεται στην **Εικόνα 8**.



**Εικόνα 8:** Το συνελκτικό επίπεδο δείχνει το κεντρικό στοιχείο του πυρήνα τοποθετημένο πάνω στο διάνυσμα εισόδου. Οι τιμές εντός του πυρήνα υπολογίζονται και το κεντρικό στοιχείο αντικαθίσταται με ένα σταθμισμένο άθροισμα του εαυτού του και των γειτονικών pixel [16]

Κατά τη διάρκεια της επεξεργασίας, το στρώμα συνελκτικής επεξεργασίας υπολογίζει το γινόμενο τελείας μεταξύ κάθε τιμής στον πυρήνα και του αντίστοιχου τμήματος της εισόδου. Αυτό επιτρέπει στο δίκτυο να μάθει τους πυρήνες που ενεργοποιούνται όταν ανιχνεύονται συγκεκριμένα χαρακτηριστικά σε συγκεκριμένες χωρικές θέσεις εντός της εισόδου. Αυτά τα ενεργοποιημένα χαρακτηριστικά είναι συνήθως γνωστά ως ενεργοποιήσεις ή χάρτες χαρακτηριστικών [16].

Κάθε πυρήνας στο συνελκτικό στρώμα παράγει το δικό του χάρτη ενεργοποίησης και αυτοί οι χάρτες στοιβάζονται κατά μήκος της διάστασης βάθους για να δημιουργήσουν τον πλήρη όγκο εξόδου του στρώματος [16].

Η εκπαίδευση ANN σε εισόδους εικόνας, όπως στην περίπτωση των CNN, αντιμετωπίζει συχνά προκλήσεις λόγω του μεγάλου μεγέθους του μοντέλου. Αυτό προκαλείται κυρίως από την πλήρως συνδεδεμένη φύση των τυπικών νευρώνων ANN. Για τον μετριασμό αυτού του προβλήματος, κάθε νευρώνας σε ένα στρώμα συνελκτικού τύπου συνδέεται μόνο με μια μικρή περιοχή του όγκου εισόδου, γνωστή ως μέγεθος του δεκτικού πεδίου του νευρώνα. Η συνδεσιμότητα βάθους συνήθως ρυθμίζεται ώστε να αντιστοιχεί στο βάθος της εισόδου [16].

Για παράδειγμα, εάν η είσοδος στο δίκτυο είναι μια εικόνα RGB  $64 \times 64 \times 3$  (με διαστάσεις  $64 \times 64$ ) και έχει επιλεγεί ένα μέγεθος δεκτικού πεδίου  $6 \times 6$ , κάθε νευρώνας μέσα στο στρώμα συνέλιξης

θα έχει συνολικά 108 βάρη ( $6 \times 6 \times 3$ , όπου το 3 αντιπροσωπεύει τη συνδεσιμότητα βάθους του όγκου). Ένας τυπικός νευρώνας σε άλλες μορφές ANN θα είχε συνήθως 12.288 βάρη[16].

### **2.1.5 Pooling Layer**

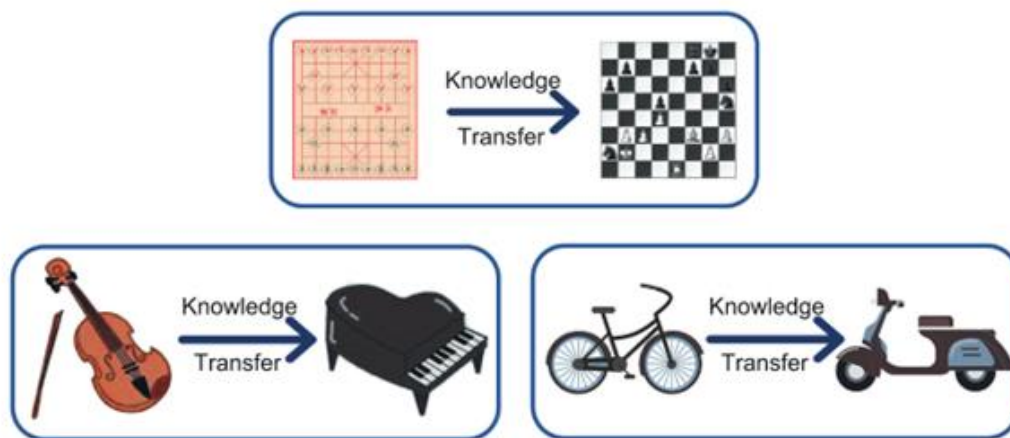
Τα στρώματα συγκέντρωσης χρησιμοποιούνται στα CNN για τη σταδιακή μείωση της διάστασης της αναπαράστασης, μειώνοντας έτσι τον αριθμό των παραμέτρων και την υπολογιστική πολυπλοκότητα του μοντέλου. Αυτά τα στρώματα λειτουργούν σε κάθε χάρτη ενεργοποίησης εντός της εισόδου και μειώνουν τη διάστασή του χρησιμοποιώντας τη συνάρτηση "MAX". Στα περισσότερα CNN, χρησιμοποιούνται συνήθως στρώματα max-pooling, τα οποία χρησιμοποιούν πυρήνες  $2 \times 2$  με βήμα 2 κατά μήκος των χωρικών διαστάσεων της εισόδου. Με τον τρόπο αυτό ο χάρτης ενεργοποίησης μειώνεται στο 25% του αρχικού του μεγέθους, ενώ ο όγκος βάθους διατηρείται στο τυπικό του μέγεθος. Η μέγιστη συγκέντρωση χρησιμοποιεί συνήθως δύο προσεγγίσεις. Τα στρώματα συγκέντρωσης ορίζονται συνήθως με ένα stride και φίλτρα  $2 \times 2$ , επιτρέποντάς τους να καλύπτουν ολόκληρη τη χωρική διάσταση της εισόδου. Επιπλέον, μπορεί να χρησιμοποιηθεί η αλληλεπικαλυπτόμενη συγκέντρωση, όπου το stride ορίζεται σε 2, με μέγεθος πυρήνα 3. Γενικά παρατηρείται ότι η χρήση μεγέθους πυρήνα μεγαλύτερου από 3 για τη συγκέντρωση μειώνει σημαντικά την απόδοση του μοντέλου λόγω της καταστροφικής του φύσης[16].

### **2.1.6 Fully Connected Layer**

Το πλήρως συνδεδεμένο στρώμα αποτελείται από νευρώνες που συνδέονται άμεσα με τους νευρώνες στα δύο γειτονικά στρώματα, παρόμοια με τη διάταξη των νευρώνων στις παραδοσιακές μορφές ANN. Σε αυτό το στρώμα, κάθε νευρώνας συνδέεται με όλους τους νευρώνες του προηγούμενου στρώματος και με όλους τους νευρώνες του επόμενου στρώματος. Αυτή η πλήρως συνδεδεμένη αρχιτεκτονική επιτρέπει την ολοκληρωμένη ανταλλαγή πληροφοριών μεταξύ των στρωμάτων, επιτρέποντας στο δίκτυο να μαθαίνει πολύπλοκα μοτίβα και σχέσεις. Το πλήρως συνδεδεμένο στρώμα διαδραματίζει κρίσιμο ρόλο στα τελικά στάδια του δικτύου, συνδυάζοντας τα χαρακτηριστικά που μαθαίνονται και κάνοντας προβλέψεις ή ταξινομήσεις με βάση την είσοδο[16].

### **2.1.7 Transfer Learning**

Το Transfer Learning είναι μια ελπιδοφόρα μεθοδολογία μηχανικής μάθησης που επικεντρώνεται στη μεταφορά γνώσης από έναν τομέα σε άλλους τομείς. Η έννοια της μεταφοράς προέρχεται αρχικά από



*Εικόνα 9: Παραδείγματα σχετικά με Transfer Learning [17]*

την εκπαιδευτική ψυχολογία, όπου πιστεύεται ότι η μεταφορά γίνεται μέσω της γενίκευσης μιας εμπειρίας. Αυτό σημαίνει ότι όταν κάποιος αποκτά γνώση σε έναν τομέα, μπορεί να την εφαρμόσει και σε άλλους τομείς, εκμεταλλευόμενος τις κοινές αρχές και γνώσεις μεταξύ τους.

Στην πράξη, αυτό μπορεί να συμβεί όταν κάποιος έχει μάθει ένα εργαλείο ή έναν τρόπο να λύνει ένα πρόβλημα σε έναν τομέα και τον εφαρμόζει επιτυχώς σε έναν άλλο τομέα που έχει κοινές αρχές ή παρόμοια δομή. Παραδείγματος χάρη ένα άτομο που έχει μάθει να παίζει βιολί μπορεί να μάθει πιάνο πιο γρήγορα από κάποιον που δεν έχει καμία μουσική εκπαίδευση. Καθώς και το βιολί και το πιάνο είναι μουσικά όργανα, υπάρχουν κοινές αρχές και τεχνικές που μπορούν να μεταφερθούν από το ένα όργανο στο άλλο. Φαίνονται και άλλα σχετικά παραδείγματα στην εικόνα 10. Με άλλα λόγια, το Transfer Learning στη μηχανική μάθηση αποσκοπεί στο να αξιοποιήσει τη γνώση από έναν σχετικό τομέα (ονομάζεται τομέας προέλευσης) για να βελτιώσει την απόδοση ή να μειώσει τον αριθμό των απαιτούμενων ετικετών που απαιτούνται σε έναν νέο τομέα (ονομάζεται τομέας στόχος). Ωστόσο, πρέπει να σημειωθεί ότι η μεταφερόμενη γνώση δεν είναι πάντα θετική για τις νέες εργασίες, καθώς μπορεί να υπάρχουν διαφορές στους τομείς που δυσχεραίνουν τη μεταφορά[17].

### **2.1.8 EfficientNet CNN**

Το EfficientNet είναι μια πρωτοποριακή οικογένεια CNN που έφερε επανάσταση στον τομέα της ταξινόμησης εικόνων. Ο πρωταρχικός στόχος του EfficientNet είναι η επίτευξη ισορροπίας μεταξύ ακρίβειας και αποδοτικότητας με προσεκτική διαχείριση του βάθους, του πλάτους και της ανάλυσης του δικτύου. Αυτή η εξισορρόπηση επιτυγχάνεται μέσω μιας μοναδικής μεθόδου κλιμάκωσης που ονομάζεται σύνθετος συντελεστής. Ο σύνθετος συντελεστής κλιμακώνει ομοιόμορφα τις διαστάσεις του δικτύου για τη δημιουργία μιας σειράς μοντέλων με διαφορετικά επίπεδα ακρίβειας και υπολογιστικής απόδοσης.



Το EfficientNet βασίζεται κυρίως σε μια προσέγγιση αναζήτησης νευρωνικής αρχιτεκτονικής (NAS). Αρχικά, αναπτύσσεται μια βασική αρχιτεκτονική δικτύου χρησιμοποιώντας ένα μικρότερο σύνολο δεδομένων. Αυτό το βασικό δίκτυο χρησιμεύει ως βάση για τα επόμενα πειράματα κλιμάκωσης. Στη συνέχεια, το δίκτυο επεκτείνεται συστηματικά με την προσθήκη περισσότερων στρωμάτων, την αύξηση του πλάτους κάθε στρώματος και την προσαρμογή της ανάλυσης των εικόνων εισόδου. Αυτή η διαδικασία επέκτασης καθοδηγείται από τον σύνθετο συντελεστή, ο οποίος καθορίζει πόσο θα κλιμακωθεί κάθε διάσταση. Το EfficientNet εισάγει έναν νέο κανόνα κλιμάκωσης που ονομάζεται "σύνθετη κλιμάκωση" για την επίτευξη βέλτιστων αποτελεσμάτων. Ο σύνθετος συντελεστής, που συμβολίζεται ως  $\phi$  ( $\phi$ ), ελέγχει ταυτόχρονα την κλιμάκωση βάθους, πλάτους και ανάλυσης. Οι συντελεστές για το βάθος, το πλάτος και την ανάλυση καθορίζονται από τις εξισώσεις:

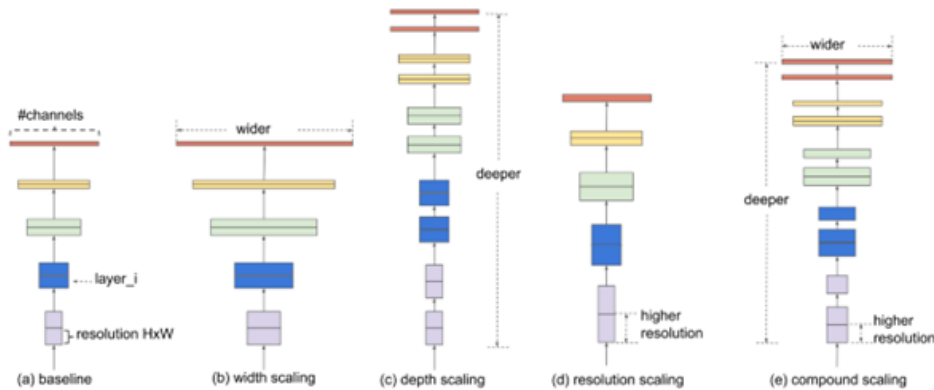
$$depth = \alpha \phi^{\phi} \quad (1)$$

$$width = \beta \phi^{\phi} \quad (2)$$

$$resolution = \gamma \phi^{\phi} \quad (3)$$

Εδώ, τα  $\alpha$ ,  $\beta$  και  $\gamma$  είναι σταθερές κλιμάκωσης που ορίζονται χειροκίνητα με βάση εμπειρικά αποτελέσματα. Μεταβάλλοντας την τιμή του  $\phi$ , το EfficientNet μπορεί να κλιμακώσει αποτελεσματικά το δίκτυο προς τα πάνω ή προς τα κάτω[18].

Η αρχιτεκτονική του EfficientNet αποτελείται κατά κύριο λόγο από επίπεδα συνελίξεων, ακολουθούμενα από ομαλοποίηση δέσμης και συναρτήσεις ενεργοποίησης. Χρησιμοποιεί μια παραλλαγή του δημοφιλούς συνελικτικού μπλοκ που ονομάζεται μπλοκ "MBConv" (Mobile Inverted Residual Bottleneck Convolutional block), το οποίο περιλαμβάνει μια διαχωρίσιμη σε βάθος συνελικτική συνέλιξη και έναν μηχανισμό συμπίεσης και διέγερσης (SE). Η κατά βάθος διαχωρίσιμη συνέλιξη μειώνει την υπολογιστική πολυπλοκότητα με το διαχωρισμό της συνέλιξης σε ξεχωριστές συνέλιξης κατά βάθος και κατά σημείο. Ο μηχανισμός SE ενισχύει τη δύναμη αναπαράστασης του δικτύου με προσαρμοστική στάθμιση των σημαντικών καναλιών[18].



**Εικόνα 10:** Σχηματική Απεικόνιση Αρχιτεκτονικής EfficientNet[18]

Το EfficientNet επιτυγχάνει κορυφαίες επιδόσεις σε διάφορες εργασίες ταξινόμησης εικόνων, ενώ είναι σημαντικά μικρότερο και ταχύτερο σε σύγκριση με προηγούμενες αρχιτεκτονικές CNN. Επιτυγχάνει ισορροπία μεταξύ ακρίβειας και αποδοτικότητας, καθιστώντας το κατάλληλο για περιβάλλοντα με περιορισμένους πόρους, όπως κινητές συσκευές και ενσωματωμένα συστήματα. Η επιτυχία του έγκειται στην προσεκτική ενορχήστρωση του βάθους, του πλάτους και της ανάλυσης του δικτύου μέσω της μεθόδου σύνθετης κλιμάκωσης, σε συνδυασμό με τη χρήση προηγμένων δομικών στοιχείων, όπως το μπλοκ MBConv το οποίο περιλαμβάνει μια διαχωρίσιμη κατά βάθος συνέλιξη και έναν μηχανισμό συμπίεσης και διέγερσης (SE). Η διαχωρίσιμη κατά βάθος συνέλιξη μειώνει την υπολογιστική πολυπλοκότητα με το διαχωρισμό της συνέλιξης σε ξεχωριστές συνέλιξης κατά βάθος και κατά σημείο. Ο μηχανισμός SE ενισχύει τη δύναμη αναπαράστασης του δικτύου με προσαρμοστική στάθμιση των σημαντικών καναλιών[18].

## 2.2 Σύνολο δεδομένων προς ανάλυση

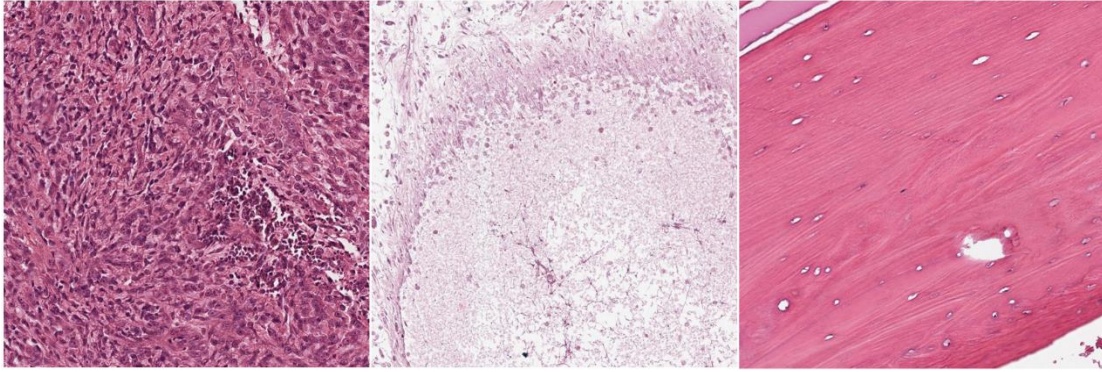
Το Dataset που χρησιμοποιήθηκε προέρχεται από τα ιατρικό κέντρο του Πανεπιστημίου του Ντάλας (UT Southwestern/UT Dallas) από την ιστοσελίδα The Cancer Imaging Archive (TCIA) Public Access<sup>8</sup>.

Σε σύνολο υπάρχουν στοιχεία από 50 ασθενείς, οι οποίοι όπου υποβλήθηκαν σε χειρουργική αφαίρεση, στο παιδιατρικό κέντρο (Children's Medical Center, Dallas) την περίοδο 1995 με 2015. Το οστό που αφαιρέθηκε κόπηκε σε κομμάτια, καθαρίστηκε από ασβεστώσεις, υποβλήθηκε σε χρώση H&E και μετατράπηκε σε αντικειμενοφόρες πλάκες. Οι διαφάνειες σαρώθηκαν σε απεικονίσεις ολόκληρης αντικειμενοφόρου πλάκας (WSI), 40 από τις οποίες επιλέχθηκαν χειροκίνητα από δύο παθολόγους με βάση την ετερογένεια του όγκου και τα χαρακτηριστικά ανταπόκρισης. Στην συνέχεια, από τα δείγματα αυτά, επιλέχθηκαν τυχαία 30 πλακίδια μεγέθους  $1024 \times 1024$  pixel, με αποτέλεσμα να προκύψουν 1200 πλακίδια. Εξ' αυτών, επιλέχθηκαν 1144 πλακίδια για ανάλυση, κατόπιν απόρριψης μη-ιστών, περιοχών με μελάνι και θολά δείγματα. Τα πλακίδια φέρουν έναν μόνο χαρακτηρισμό, καθώς κάθε εικόνα σχολιάστηκε από έναν μόνο παθολόγο, χειροκίνητα, με τις ακόλουθες κατανομές: 536 (47%) εικόνες χωρίς όγκο, 263 (23%) εικόνες με νεκρωτικό όγκο και 345 (30%) εικόνες με βιώσιμο όγκο.

Το σύνολο δεδομένων αυτό χωρίζει τις εικόνες, όπως προαναφέρθηκε, σε τρεις κατηγορίες [3]: 1. βιώσιμο όγκο (Viable-Tumor), 2. μη-βιώσιμο (Non-Viable-Tumor) και 3. μη-καρκινικό ιστό (Non-Tumor). Η πρώτη κατηγορία έχει ιστολογικές απεικονίσεις που αποτελούνται από υπερχρωματικούς πυρήνες πυκνά συσσωρευμένους. Η δεύτερη κατηγορία έχει απεικονίσεις με λιγότερο πυκνή χρώση και, σε περίπτωση που τα μη βιώσιμα κύτταρα έχουν υποστεί πηκτική νέκρωση, αποσυντεθειμένους πυρήνες. Η τρίτη κατηγορία εμπεριέχει ότι δεν εντάσσεται στις άλλες δύο, όπως π.χ. υγιή οστά, που απεικονίζονται με έντονο κόκκινο χρώμα. [11]

---

<sup>8</sup> <https://www.cancerimagingarchive.net/>



**Εικόνα 11:** Βιώσιμος όγκος (αριστερά), μη βιώσιμος (μέση) και μη καρκινικός ιστός (δεξιά) [3]

## 2.3 Πειραματική διάταξη

Μετά την απόκτηση του Dataset των εικόνων, τις μετασχηματίσαμε σε κατάλληλη δομή, φιλική προς το PyTorch. Πιο συγκεκριμένα, οι εικόνες είναι αρχικά κατανεμημένες σε 2 υπο-φακέλους που ανάλογα τον γιατρό που έκανε την διάγνωση. Μέσα σε αυτούς υπάρχουν σετ των περίπου 50 εικόνων συνοδευόμενες από ένα αρχείο csv για την περιγραφή τις κλάσης του/διάγνωση. Με ένα script σε python οι εικόνες μετασχηματίστηκαν, σε δομή 3 φακέλων, ανάλογα την κλάση την οποία ανήκουν ('Non-Tumor', 'Non-Viable-Tumor', 'Viable-Tumor').

Για την περεταίρω υλοποίηση της άσκησης χρησιμοποιήθηκε ένα Jupiter notebook στην πλατφόρμα Google Colab με χρήση της βιβλιοθήκης *PyTorch* σε σύστημα με μία Nvidia A100 GPU με 40GB vRAM. Χρησιμοποιήθηκαν επίσης οι βιβλιοθήκες *matplotlib* και *seaborn* για την δημιουργία γραφημάτων και *torchmetrics* για την αξιολόγηση του μοντέλου.

Όλα τα αρχεία κώδικα μπορούν να βρεθούν στο link: [github.com/purpl3F0x/biomed\\_project](https://github.com/purpl3F0x/biomed_project).

Το μοντέλο αρχιτεκτονικής που επιλέξαμε είναι το EfficientNet, καθώς είναι ιδιαίτερα δημοφιλές στη βιβλιογραφία και έχει παρουσιάσει εδραιωμένη επιτυχία. Ειδικότερα, διαλέξαμε το μοντέλο B7, που είναι και το μεγαλύτερο μοντέλο της οικογένειας των δικτύων αυτών, με πάνω από 66 εκατομμύρια νευρώνες. Επειδή η εφαρμογή μας είναι ιατρικού σκοπού, δεν θεωρήθηκε ότι υπάρχει περιορισμός ως προς την υπολογιστική δύναμη, αλλά δίνεται μεγαλύτερη έμφαση στην ακρίβεια των αποτελεσμάτων, έτσι η επιλογή του μεγαλύτερου B7 είναι συνετή (αν και υπογραμμίζεται ότι μεγαλύτερα μοντέλα, δεν ισοδυναμούν καθολικά με καλύτερα αποτελέσματα – η αξιολόγηση αυτή όμως είναι πέραν τα πλαίσια της εργασίας αυτής).

Στο πρόγραμμά μας, αρχικά εγκαταστάθηκαν και φορτώθηκαν οι βιβλιοθήκες. Στην συνέχεια έγινε λήψη των δεδομένων, φόρτωση και ανακάτεμά τους μέσω της συνάρτησης *torch.utils.data.random\_split*. Με αυτή, οι εικόνες χωρίστηκαν σε 3 σύνολα: training, validation και testing data σε ποσοστά 60%-25% και 15% αντίστοιχα. Οι εικόνες μετασχηματίστηκαν σε διαστάσεις  $600 \times 600$  pixel, για να είναι συνεπείς με το μέγεθος εισόδου του προ-εκπαιδευμένου μοντέλου που χρησιμοποιήθηκε. Επιπλέον, χρησιμοποιήθηκαν οι μέσες τιμές των καναλιών των χρωμάτων 0.485, 0.456 και 0.406 και οι τιμές τυπικής απόκλισης 0.229, 0.224 και 0.225 για τα κανάλια R, G και B, αντίστοιχα.

Ο διαχωρισμός του συνόλου στα τρία αυτά υποσύνολα είναι κοινή πρακτική στα νευρωνικά δίκτυα για την εκπαίδευση και αξιολόγηση ενός δικτύου.

Αναλυτικότερα:

- Training Set είναι το σύνολο δεδομένων που χρησιμοποιείται για την εκπαίδευση, ώστε να μάθει το μοντέλο τα κρυμμένα χαρακτηριστικά/πρότυπα των δεδομένων. Σε κάθε εποχή, τα ίδια δεδομένα εκπαίδευσης τροφοδοτούνται επανειλημμένα στην αρχιτεκτονική του νευρωνικού δικτύου με το μοντέλο να συνεχίζει να μαθαίνει τα χαρακτηριστικά των δεδομένων.
- Validation Set είναι ένα σύνολο δεδομένων, ξεχωριστό από το σύνολο εκπαίδευσης, το οποίο χρησιμοποιείται για την επικύρωση της απόδοσης του μοντέλου μας κατά τη διάρκεια της εκπαίδευσης. Η διαδικασία αυτής της επαλήθευσης παρέχει πληροφορίες που μας βοηθούν να προσαρμόσουμε ανάλογα τις υπερπαραμέτρους και τις διατάξεις του μοντέλου. Δηλαδή, το μοντέλο εκπαιδεύεται στο σύνολο εκπαίδευσης και, ταυτόχρονα, η αξιολόγηση του μοντέλου πραγματοποιείται στο σύνολο επικύρωσης μετά από κάθε εποχή. Ο διαχωρισμός των δεδομένων σε validation set, αποτρέπει το μοντέλο από το να υφίσταται υπερβολική προσαρμογή (overfitting), δηλαδή να γίνεται πολύ καλό στην ταξινόμηση των δειγμάτων του συνόλου εκπαίδευσης, αλλά να μην μπορεί να γενικεύσει και να κάνει ακριβείς ταξινομήσεις σε άγνωστα δεδομένα.
- Test Set είναι ένα ξεχωριστό σύνολο δεδομένων που χρησιμοποιείται για τη δοκιμή του μοντέλου μετά την ολοκλήρωση της εκπαίδευσης. Παρέχει μια αμερόληπτη τελική μετρική απόδοσης του μοντέλου όσον αφορά την ακρίβεια και τη δυνατότητά του να ταξινομήσει άγνωστα σε αυτό δεδομένα.

Με τα παραπάνω δεδομένα, πριν την εκπαίδευση του μοντέλου «παγώσαμε» όσα εσωτερικά στρώματα είναι εκπαιδεύσιμα, και ορίσαμε τον αριθμό των παραγόμενων κλάσεων στο τελευταίο στρώμα – το οποίο είναι προς εκπαίδευση- σε 3, μία κλάση για κάθε είδος ιστολογικής απεικόνισης ('Non-Tumor', 'Non-Viable-Tumor', 'Viable-Tumor').

Προχωρώντας στο στάδιο της εκπαίδευσης, μετά από αρκετές δοκιμές επιλέξαμε τον αριθμό των 15 εποχών. Αυτό διότι, σε μικρότερο αριθμό εποχών παρατηρούσαμε -όπως είναι αναμενόμενο- μειωμένη ακρίβεια, αλλά μειωμένη ακρίβεια παρατηρήσαμε και με την αύξηση των εποχών σε 20, λόγω του προαναφερθέντος φαινομένου του overfitting.

Τέλος, μετά το πέρας της διαδικασίας του transfer learning, αξιολογήσαμε το μοντέλο μας πάνω στο validation set, δηλαδή σε άγνωστες για το εκπαιδευμένο δίκτυο εικόνες. Για την ολοκληρωμένη και μεθοδική αξιολόγηση του δικτύου, υπολογίσαμε δείκτες αξιολόγησης και τον πίνακα σύγκρισης του μοντέλου με τις συναρτήσεις της βιβλιοθήκης *torchmetrics*.

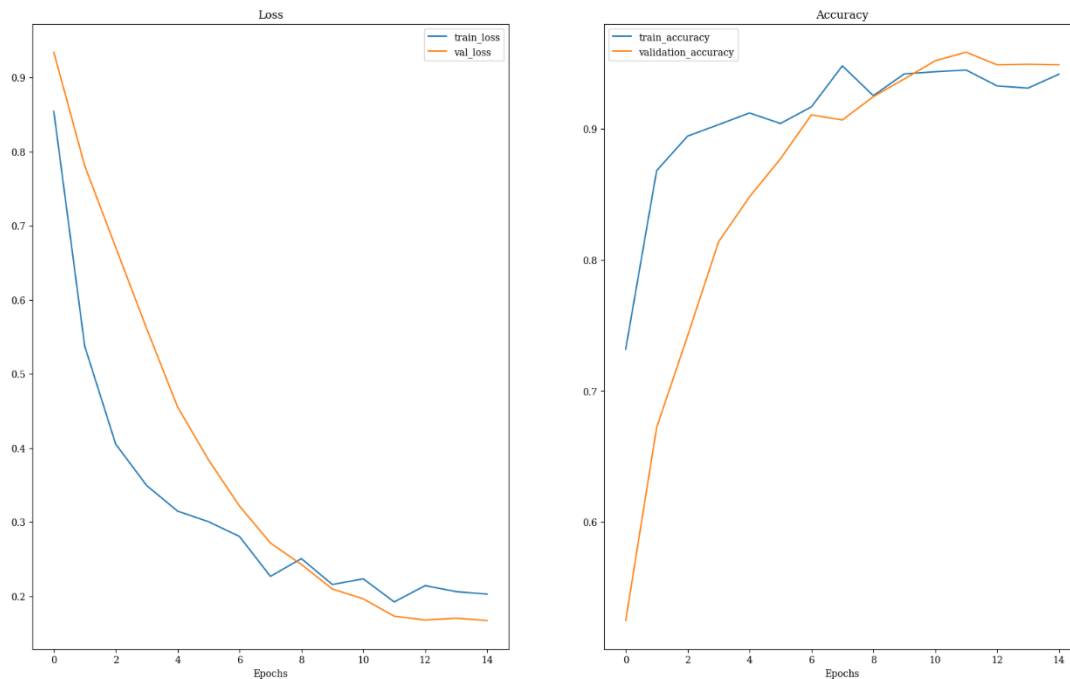
# 3

## Αποτελέσματα

### 3.1 Οπτικοποίηση Προόδου Εκπαίδευσης και Αξιολόγηση Απόδοσης

#### 3.1.1 Loss and Accuracy per Epoch

Κατά τη διάρκεια της εκπαίδευσης του μοντέλου υπολογίστηκε η απώλεια και η ακρίβεια αντίστοιχα για κάθε εποχή.



Εικόνα 12: Διαγράμματα Ακρίβειας(δεξιά) και Απώλειας(αριστερά)

Η εικόνα απεικονίζει την πρόοδο της εκπαίδευσης ενός μοντέλου μηχανικής μάθησης με δύο γραφικές αναπαραστάσεις. Στην αριστερή πλευρά, ένα γραμμικό γράφημα εμφανίζει την απώλεια ανά εποχή, δείχνοντας τη μεταβολή της συνάρτησης απώλειας του μοντέλου καθώς προχωρά η εκπαίδευση. Η απώλεια αντιπροσωπεύει την απόκλιση μεταξύ των προβλεπόμενων και των πραγματικών τιμών και το γράφημα παρέχει πληροφορίες σχετικά με το πόσο αποτελεσματικά το μοντέλο ελαχιστοποιεί αυτή την απόκλιση σε διαδοχικές εποχές. Μια φθίνουσα τάση υποδηλώνει ότι το μοντέλο μαθαίνει και βελτιώνει την απόδοσή του. Ο στόχος της εκπαίδευσης είναι η ελαχιστοποίηση αυτής της τιμής απώλειας. Παρατηρούμε ότι πράγματι με την αύξηση των εποχών η απώλεια μειώνεται.

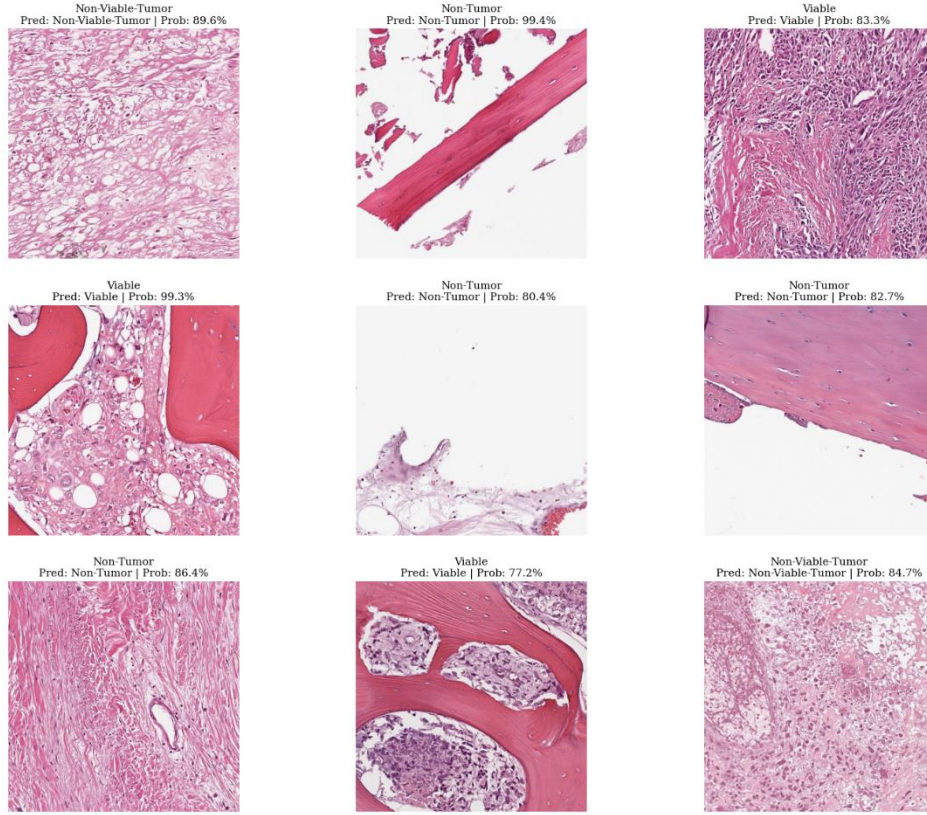
Στη δεξιά πλευρά, ένα άλλο γραμμικό γράφημα παρουσιάζει την ακρίβεια ανά εποχή, καταδεικνύοντας την ικανότητα του μοντέλου να κάνει σωστές προβλέψεις. Η ακρίβεια αντιπροσωπεύει το ποσοστό των σωστά ταξινομημένων δειγμάτων επί του συνόλου των δειγμάτων που αξιολογήθηκαν. Το γράφημα αποκαλύπτει πώς εξελίσσεται η ακρίβεια του μοντέλου με την πάροδο του χρόνου, δείχνοντας ενδεχομένως μια ανοδική τάση καθώς το μοντέλο γίνεται πιο ικανό στις προβλέψεις του. Παρατηρούμε ότι, όπως ήταν επιθυμητό, με την αύξηση των εποχών η ακρίβεια αυξάνεται.

### **3.1.2 Test Set Analysis**

Η εικόνα 13 έχει αντιπροσωπευτικά δείγματα από το σύνολο δεδομένων και έχουν παραμεριστεί ειδικά για τον σκοπό της τελειοποίησης και της αξιολόγησης της απόδοσης ενός μοντέλου μηχανικής μάθησης που εκπαιδεύεται στις υπόλοιπες εικόνες. Αναλύοντας τις προβλέψεις του μοντέλου στο σύνολο επικύρωσης, μπορούμε να αποκτήσουμε πληροφορίες σχετικά με τις παραμέτρους αξιολόγησης, ακρίβεια, F1\_score κτλ., επιτρέποντάς μας να βελτιστοποιήσουμε την απόδοση του μοντέλου και να διασφαλίσουμε την αποτελεσματικότητά του στη διάγνωση του οστεοσαρκώματος.



## Test Set Predictions



*Εικόνα 13: Προβλέψεις σε τυχαία δείγματα του μοντέλου*

### 3.2 Δείκτες αξιολόγησης

Η αξιολόγηση ενός νευρωνικού δικτύου γίνεται με τους ακόλουθες πέντε δείκτες: Ευαισθησία (Recall), Ειδικότητα (Specificity), Ακρίβεια (Accuracy), Θετική Διαγνωστική Αξία (Precision) και F1\_score, με τύπους<sup>9</sup>:

$$recall = \frac{TP}{P} = \frac{TP}{TP + FN} \quad (4)$$

$$specificity = \frac{TN}{N} = \frac{TN}{TN + FP} \quad (5)$$

$$precision = \frac{TP}{TP + FP} \quad (6)$$

<sup>9</sup> <https://www.analyticsvidhya.com/blog/2021/07/metrics-to-evaluate-your-classification-model-to-take-the-right-decisions/>



$$accuracy = \frac{TP + TN}{P + N} = \frac{TP + TN}{TP + TN + FP + FN} \quad (7)$$

$$F1\_score = 2 \cdot \frac{precision \cdot recall}{precision + recall} \quad (8)$$

, όπου TP, TN, FP, FN είναι τα true positives, true negatives, false positives, και false negatives. Αυτή η αξιολόγηση μπορεί να γίνει εάν αντιμετωπίσουμε την ταξινόμηση σε κλάση ως ένα πρόβλημα δυαδικής ταξινόμησης. Σε αυτή την προσέγγιση, η θετική κλάση αντιστοιχεί στη συγκεκριμένη κλάση ενδιαφέροντος, ενώ η αρνητική κλάση περιλαμβάνει όλες τις υπόλοιπες κλάσεις στο πεδίο του προβλήματος. Ακολουθώντας, τα TP και FP υποδηλώνουν τον αριθμό των εικόνων που ταξινομούνται ως ανήκουσες στη θετική κλάση, όπου το TP υποδηλώνει τις σωστές ταξινομήσεις και το FP τις λανθασμένες ταξινομήσεις. Ομοίως, τα TN και FN αντιπροσωπεύουν τον αριθμό των εικόνων που ταξινομούνται ως μη ανήκουσες στη θετική κλάση, όπου TN σημαίνει σωστές ταξινομήσεις και FN σημαίνει λανθασμένες ταξινομήσεις.

Η απώλεια (loss) αναφέρεται στη διαφορά μεταξύ των προβλεπόμενων εξόδων του μοντέλου και των πραγματικών ετικετών των εικόνων στο σύνολο εκπαίδευσης.

Η ακρίβεια (accuracy) είναι ένας δείκτης απόδοσης που μετράει πόσο καλά το CNN προβλέπει τις σωστές ετικέτες κλάσης για τις εικόνες. Αντιπροσωπεύει το ποσοστό των σωστά ταξινομημένων εικόνων προς το συνολικό αριθμό εικόνων στο σύνολο εκπαίδευσης για μια συγκεκριμένη εποχή.

**Ευαισθησία (Recall):** Απαντά στο ερώτημα: τι ποσοστό των πραγματικά θετικών(TP) κατηγοριοποιείται σωστά.

**Ειδικότητα (Specificity):** Αντιπροσωπεύει το ποσοστό των πραγματικών αρνητικών (TN) επί του συνόλου των αρνητικών δεδομένων(N).

**Θετική Διαγνωστική Αξία (Precision):** Δείχνει το πόσες από τις προβλεπόμενες θετικές είναι πράγματι θετικές.

**F1\_score:** Το F1\_score κυμαίνεται από 0 έως 1, όπου το 1 υποδηλώνει την καλύτερη δυνατή απόδοση. Μια υψηλότερη βαθμολογία F1 υποδηλώνει καλύτερη ισορροπία μεταξύ θετικής διαγνωστικής αξίας (precision) και ευαισθησίας (recall).

	Non-Tumor	Non-Viable Tumor	Viable
<b>Accuracy</b>	0.987	0.972	0.957
<b>Specificity</b>	0.975	0.992	0.991
<b>Precision</b>	0.975	0.972	0.978
<b>Recall</b>	0.987	0.972	0.957
<b>F1 Score</b>	0.981	0.972	0.967

*Πίνακας 3: Πίνακας Αξιολόγησης Μοντέλου*

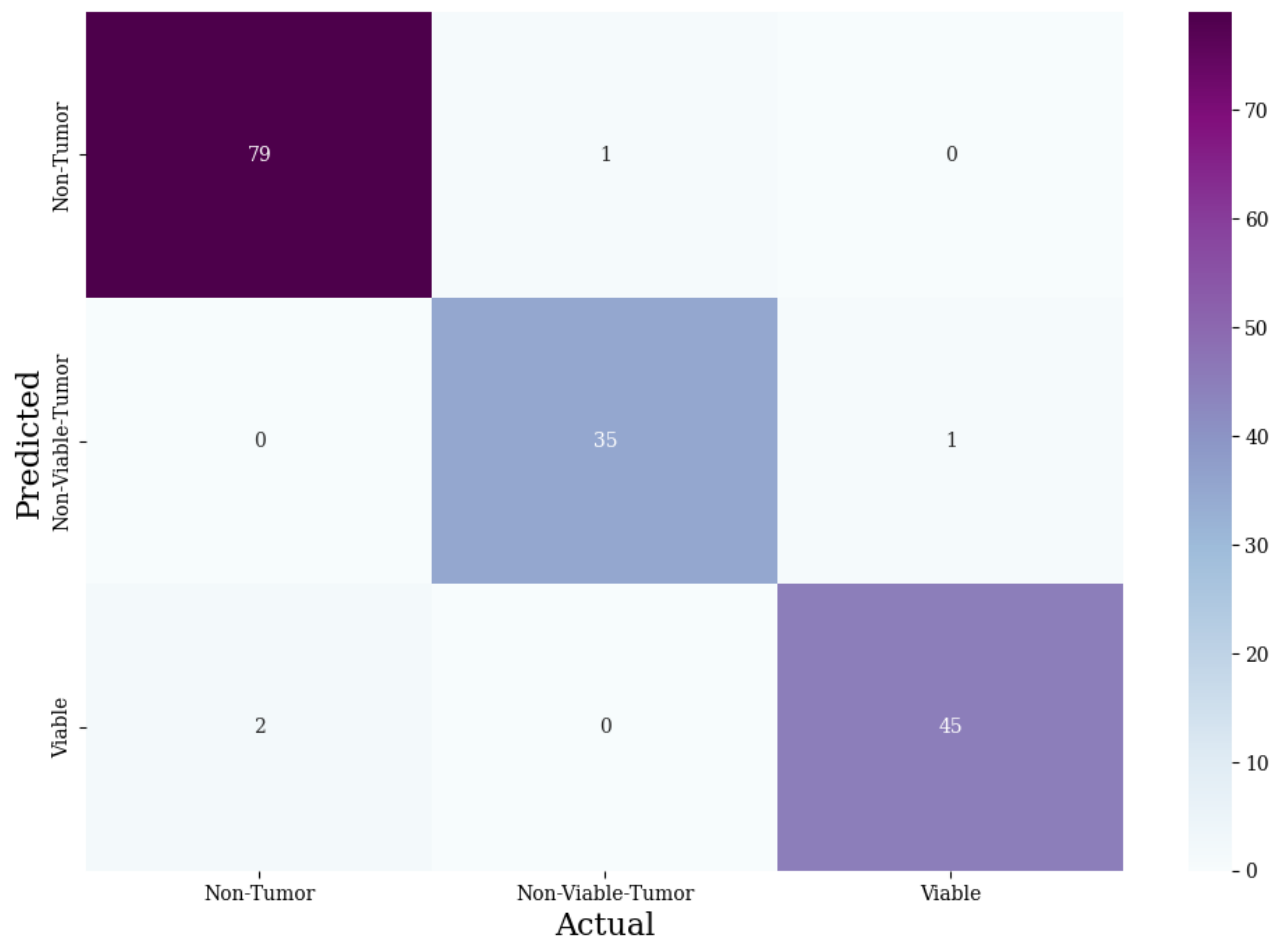
### 3.2.1 Ανάλυση Πίνακα Σύγχυσης

Επιπλέον, υπολογίζουμε και απεικονίζουμε τον πίνακα σύγχυσης, έναν πίνακα 3x3 που είναι οργανωμένος σε γραμμές και στήλες, με κάθε στήλη να αντιπροσωπεύει τις πραγματικές ετικέτες του συνόλου δεδομένων και κάθε γραμμή να αντιπροσωπεύει τις προβλεπόμενες ετικέτες από το μοντέλο. Οι γραμμές και οι στήλες αντιστοιχούν στις διάφορες κλάσεις που εμπλέκονται στην εργασία ταξινόμησης.

Ο πίνακας σύγχυσης, ένα θεμελιώδες εργαλείο για την αξιολόγηση της απόδοσης της ταξινόμησης των εικόνων οστεοσαρκώματος, παρέχει μια ολοκληρωμένη ανάλυση των προβλέψεων του μοντέλου στις τρεις διαφορετικές κατηγορίες: Μη όγκος, μη βιώσιμος όγκος και βιώσιμος όγκος. Ο πίνακας σύγχυσης που αναπαρίσταται ως πίνακας 3x3, προσφέρει λεπτομερή ανάλυση των αποτελεσμάτων ταξινόμησης του μοντέλου. Κάθε γραμμή του πίνακα αντιστοιχεί στις προβλεπόμενες ετικέτες των δειγμάτων από το μοντέλο, ενώ κάθε στήλη αντιπροσωπεύει τις πραγματικές ετικέτες των εικόνων που ανήκουν σε μια συγκεκριμένη κλάση. Τα διαγώνια στοιχεία του πίνακα υποδεικνύουν τα σωστά ταξινομημένα δείγματα, όπου η προβλεπόμενη κλάση ευθυγραμμίζεται με την πραγματική κλάση, αντικατοπτρίζοντας έτσι την ακρίβεια του μοντέλου. Από την άλλη πλευρά, τα εκτός διαγωνίου στοιχεία υπογραμμίζουν τις λανθασμένες ταξινομήσεις, αντιπροσωπεύοντας περιπτώσεις όπου το μοντέλο προέβλεψε λανθασμένα την κλάση ενός δείγματος.

Η ανάλυση του πίνακα σύγχυσης επιτρέπει τη λεπτομερή εξέταση των δυνατών και αδύνατων σημείων του μοντέλου στην ταξινόμηση εικόνων οστεοσαρκώματος. Τα μοτίβα που προκύπτουν από τις λανθασμένες ταξινομήσεις μπορούν να αποκαλύψουν πιθανές προκλήσεις, όπως η διάκριση μεταξύ των κατηγοριών μη βιώσιμου όγκου και βιώσιμου όγκου ή η διαφοροποίηση των δειγμάτων μη όγκου από τις κατηγορίες που σχετίζονται με τον όγκο. Εντοπίζοντας συγκεκριμένες περιοχές εσφαλμένης ταξινόμησης, οι ερευνητές και οι κλινικοί γιατροί μπορούν να επικεντρωθούν στη βελτίωση της απόδοσης του μοντέλου σε αυτές τις συγκεκριμένες πτυχές, οδηγώντας ενδεχομένως σε αυξημένη διαγνωστική ακρίβεια και φροντίδα των ασθενών. Έτσι, ο πίνακας σύγχυσης χρησιμεύει ως κρίσιμο διαγνωστικό εργαλείο, βοηθώντας στην αξιολόγηση, τη βελτίωση και τη βελτιστοποίηση του μοντέλου ταξινόμησης εικόνων οστεοσαρκώματος.

Έτσι λοιπόν παρατηρώντας τον πίνακα σύγχυσης του μοντέλου μας, στην εικόνα 14, διαπιστώνουμε την ότι η ταξινόμηση των εικόνων σε μεγάλο ποσοστό έγινε ορθώς, αφού σε συντριπτική πλειοψηφία τα δεδομένα βρίσκονται πάνω στη διαγώνιο. Συγκεκριμένα ακολουθούν και τα ποσοστά αποτυχίας:



*Εικόνα 14: Πίνακας σύγχυσης του μοντέλου μας*

<b>Non-Tumor</b>	2.47%
<b>Non-Viable Tumor</b>	2.78%
<b>Viable Tumor</b>	2.17%

*Πίνακας 4: Ποσοστά Αποτυχίας στον Πίνακα Σύγχυσης*

# 4

## *Συμπεράσματα - Επίλογος*

### *4.1 Συμπεράσματα της Εργασίας*

Εν κατακλείδι, η παρούσα εργασία κατέδειξε την αποτελεσματικότητα της χρήσης του CNN EfficientNet-B7 για την ταξινόμηση εικόνων οστεοσαρκώματος. Με την εκπαίδευση και την αξιολόγηση του μοντέλου σε ένα αρκετά περιορισμένο σύνολο δεδομένων εικόνων οστεοσαρκώματος, πετύχαμε υποσχόμενα αποτελέσματα όσον αφορά την ακρίβεια, την ακρίβεια, την ανάκληση και το F1-score και γενικότερα σε όλους τους δείκτες αποδοτικότητας. Συγκεκριμένα, οι δείκτες αξιολόγησης του μοντέλου μας βρίσκονται ως επί των πλείστων πάνω από 90%, με ορισμένους όπως η ειδικότητα της κλάσης του μη-βιώσιμου όγκου να φτάνουν το 99.2%. Ο πίνακας σύγχυσης δε, είναι σχεδόν διαγώνιος, με τις τιμές των μη διαγώνιων στοιχείων να είναι το πολύ 2 (δηλαδή  $2/81 = 2.47\%$  των πραγματικών τιμών).

Το μοντέλο που χρησιμοποιήθηκε, μέσω του transfer learning, έδειξε μεγάλες δυνατότητες για την υποβοήθηση της διάγνωσης και της θεραπείας του οστεοσαρκώματος. Η ακριβής ταξινόμηση των εικόνων οστεοσαρκώματος μπορεί να συμβάλει στην έγκαιρη ανίχνευση, επιτρέποντας την έγκαιρη παρέμβαση και τη βελτίωση των αποτελεσμάτων των ασθενών. Η ικανότητα διάκρισης μεταξύ των κατηγοριών μη όγκου, μη βιώσιμου όγκου και βιώσιμου όγκου παρέχει πολύτιμες πληροφορίες για τους κλινικούς γιατρούς στον καθορισμό των κατάλληλων στρατηγικών θεραπείας.

Ενώ τα αποτελέσματα είναι ελπιδοφόρα, είναι σημαντικό να αναγνωριστούν ορισμένοι περιορισμοί της παρούσας μελέτης. Μία από τις κύριες προκλήσεις για την ανάπτυξη του μοντέλου ταξινόμησης εικόνων οστεοσαρκώματος ήταν η περιορισμένη διαθεσιμότητα δεδομένων με ετικέτες, ενώ πρέπει να ληφθεί κι υπόψιν, ότι οι εικόνες που χρησιμοποιήθηκαν, φέρουν διάγνωση από μόνον έναν ογκολόγο. Το οστεοσάρκωμα είναι μια σχετικά σπάνια πάθηση και η απόκτηση ενός μεγάλου και ποικίλου συνόλου δεδομένων με επακριβώς σχολιασμένες

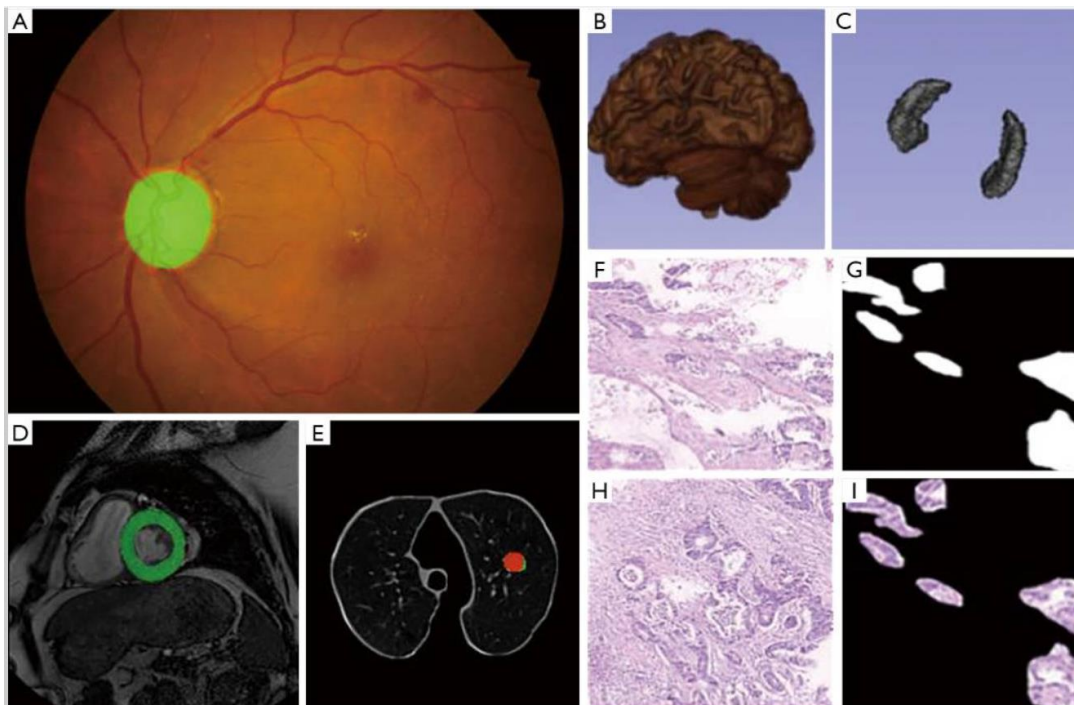
εικόνες μπορεί να αποτελέσει πρόκληση. Η έλλειψη δεδομένων μπορεί να οδηγήσει σε δυσκολίες στην εκπαίδευση ενός ισχυρού και γενικεύσιμου μοντέλου, καθώς μπορεί να δυσκολεύεται να συλλάβει την πλήρη πολυπλοκότητα και μεταβλητότητα της νόσου. Επιπλέον, η δυνατότητα γενίκευσης του μοντέλου σε διάφορους πληθυσμούς ασθενών χρήζει περαιτέρω διερεύνησης.

## **4.2 Μελλοντικές Κατευθύνσεις Έρευνας**

Υπάρχουν αρκετοί βασικοί τομείς για μελλοντική έρευνα στον τομέα της ταξινόμησης εικόνων οστεοσαρκώματος. Πρώτον, η επέκταση του συνόλου δεδομένων με τη συλλογή ενός μεγαλύτερου και πιο διαφορετικού συνόλου εικόνων οστεοσαρκώματος θα βοηθούσε στη βελτίωση της γενικευσιμότητας και της ευρωστίας των μοντέλων ταξινόμησης. Επιπλέον, η αντιμετώπιση του ζητήματος της ανισορροπίας των κλάσεων, όπου ορισμένες κλάσεις μπορεί να υποεκπροσωπούνται, είναι ζωτικής σημασίας για την επίτευξη ακριβών και ισορροπημένων προβλέψεων. Επιπλέον, η ενίσχυση της επεξηγηματικότητας των μοντέλων, με την ενσωμάτωση τεχνικών ερμηνευσιμότητας, όπως μηχανισμοί προσοχής, θα παρείχε πολύτιμες πληροφορίες για τη διαδικασία λήψης αποφάσεων των ιατρών, η οποία θα αύξανε την εμπιστοσύνη στις προβλέψεις του μοντέλου ταξινόμησης, δίνοντας βάση στα σημεία που εστιάζει και ένας ειδικός. Η διεξαγωγή μελετών κλινικής επικύρωσης για την αξιολόγηση της δυνατότητας εφαρμογής και της αποτελεσματικότητας των μοντέλων στον πραγματικό κόσμο είναι απαραίτητη για την υιοθέτησή τους σε κλινικές ρυθμίσεις. Τέλος, η υιοθέτηση πολυτροπικών προσεγγίσεων που συνδυάζουν διαφορετικές απεικονιστικές μεθόδους ή ενσωματώνουν κλινικά δεδομένα μπορεί να οδηγήσει σε πιο ολοκληρωμένη και ακριβή ταξινόμηση του οστεοσαρκώματος. Με την αντιμετώπιση αυτών των προκλήσεων και την επιδίωξη περαιτέρω έρευνας σε αυτούς τους τομείς, μπορούμε να προωθήσουμε τον τομέα της ταξινόμησης εικόνων του οστεοσαρκώματος και τελικά να βελτιώσουμε τη διάγνωση, το σχεδιασμό της θεραπείας και τα συνολικά αποτελέσματα για τους ασθενείς με αυτή τη νόσο.

Ένας άλλος τομέας προς πειραματισμό και βελτίωση είναι η χρήση διαφορετικών νευρωνικών δικτύων ή ακόμα και διαφορετικών μοντέλων αρχιτεκτονικής των συνελκτικών νευρωνικών δικτύων, όπως τα MobileNetV2, ResNet, VGG και άλλα. Στο άρθρο τους οι Vezakis et al. (2023) συγκρίνουν την απόδοση διαφορετικών μοντέλων και κρίνουν τα MobileNetV2 και EfficientNetB0, με το MobileNetV2 να επιτυγχάνει συνολική ακρίβεια (accuracy) 0.91 και μέση ευαισθησία (mean recall) 0.95, 0.93, 0.83 για μη-βιώσιμους όγκους, βιώσιμους όγκους και περιοχές νέκρωσης, αντίστοιχα. Από την άλλη, οι Loraksa et al. (2021) και Anisuzzaman et al. (2020) αξιολογούν τα VGG ως αποδοτικότερα, με τους πρώτους να κλίνουν στο VGG-16 και τους δεύτερους στο VGG-19, όπου επέτυχαν απόδοση 0.96. Οι αξιοσημείωτες αποδόσεις αυτών των μοντέλων, αλλά και ο αενάως καινοτόμος και εξελισσόμενος τομέας των νευρωνικών δικτύων δείχνουν πως υπάρχει εκτενές ερευνητικό ενδιαφέρον και χώρος για βελτίωση στην αυτόματη ταξινόμηση απεικονίσεων οστεοσαρκώματος [13],[19],[14].

Επιπλέον, ως φυσικό επόμενο της υψηλής επίδοσης του αλγορίθμου μας σε απεικονίσεις τόσο περίπλοκες όσο αυτές του οστεοσαρκώματος είναι η εφαρμογή του σε εικόνες άλλων ιστολογικών απεικονίσεων. Μετά την επιτυχία των Krizhevsky *et al.* στον διαγωνισμό ImageNet image classification με το συνελκτικό μοντέλο AlexNet, η χρήση αλγορίθμων μηχανικής μάθησης για ταξινόμηση εικόνων παρατήρησε έξαρση και βρήκε διάφορες εφαρμογές στον ιατρικό τομέα (βλ. εικόνα 15). Έτσι αντίστοιχα, μπορούμε να εφαρμόσουμε τον αλγόριθμό μας σε ιατρικές προκλήσεις όπως η ταξινόμηση άλλων ιστολογικών απεικονίσεων όγκων, η ανίχνευση διαβητικής αμφιβληστροειδοπάθειας, η εξέταση πνευμονικών όζων και πολλές άλλες.[20]



**Εικόνα 15:** Εφαρμογές της βαθιάς μάθησης στην ανάλυση ιατρικών εικόνων. (Α) Ανίχνευση του βυθού- (Β,Γ) κατάτμηση του ιππόκαμπου- (Δ) κατάτμηση της αριστερής κοιλίας- (Ε) ταξινόμηση πνευμονικών όζων- (F,Γ,H,I) κατάτμηση της παθολογίας του γαστρικού καρκίνου. [20]

Ένας ακόμη τομέας διερεύνησης είναι ο εντοπισμός οστεοσαρκώματος σε εικόνες πέρα από τις ιστολογικές απεικονίσεις, οι οποίες απαιτούν επεμβατικές και χρονοβόρες διαδικασίες για να δημιουργηθούν. Για παράδειγμα, οι Asito, et al (2021) υλοποίησαν ένα σύστημα διάγνωσης που, βασισμένο σε συνελκτικά νευρωνικά δίκτυα (CNNs), πετυχαίνει τον εντοπισμό του όγκου οστεοσαρκώματος σε ακτινογραφίες οστών [21], αλλά και υποδεικνύει τις περιοχές της εικόνας που ενδέχεται να περιέχουν όγκους αυτούς. Για την τμηματοποίηση των περιοχών συνέκριναν δύο CNN μοντέλα, το πρώτο χωρίς προ-εκπαίδευση (pre-training) ενώ το δεύτερο είναι το προ-εκπαιδευμένο VGG16. Τα δύο μοντέλα αυτά, συγκρίθηκαν με τέσσερα μοντέλα μηχανικής μάθησης που χρησιμοποιούν χαρακτηριστικά που εξάγονται από τα παράθυρα εικόνας ως είσοδοι

και είναι τα εξής: το MLP (Multilayer Perceptron), το Random Forest, το Decision Tree και το MLP με επιλογή χαρακτηριστικών. Από την σύγκριση των δύο CNN, συμπέραναν πως το προ-εκπαιδευμένο μοντέλο είναι το πιο αποδοτικό, καθώς με αυτό πέτυχαν ακρίβεια της τάξης του 0.77 και ευαισθησία της τάξης του 0.8.

Ένα σημείο με αξία επισήμανσης είναι η ποιότητα και διαφορετική μορφή των εικόνων (BMP, JPG, PNG). Όπως εξετάζουν οι Loraksa et al. (2021) διαφορετικές μορφές εικόνων μπορεί να έχουν σημαντικό αντίκτυπο στην ακρίβεια των μοντέλων CNN [19]. Συγκεκριμένα, στο άρθρο τους τα διάφορα είδη εικόνων συγκρίνονται με τρία από τα πιο γνωστά μοντέλα CNN για ανίχνευση βιώσιμων όγκων οστεοσαρκώματος, τα VGG-16, ResNet-50 και MobileNet-V2. Έτσι, με τον υπολογισμό του F1 Score αλλά και της ακρίβειας οι συντάκτες καταλήγουν στο ότι η προτιμώμενη μορφή εικόνας για την ταξινόμηση εικόνων οστεοσαρκώματος είναι η PNG.

Τέλος, άξιο διερεύνησης είναι και ο επιλογή μοντέλου με μικρότερους υπολογιστικές απαιτήσεις. Αν και εμείς χρησιμοποιήσαμε μία Nvidia A100 GPU με 40GB vRAM, με το μοντέλο να απαιτεί περίπου 9GB μνήμης. Ένας μικρότερο μοντέλο, όπως το ίδιας αρχιτεκτονικής EfficientNet B0 χρησιμοποιείται από τους Vezakis et al. (2023) [13], θα μπορούσε να έχει παρόμοια ακρίβεια με μικρότερο χρόνο διεκπεραίωσης και υπολογιστικές απαιτήσεις, πράγμα χρήσιμο για μαζική χρήση, αλλά ακόμα και φορητές συσκευές που διαθέτουν περιορισμένους πόρους.

# 5

## Βιβλιογραφία

- [1] A. Luetke, P. A. Meyers, I. Lewis, and H. Juergens, “Osteosarcoma treatment – Where do we stand? A state of the art review,” *Cancer Treatment Reviews*, vol. 40, no. 4, pp. 523–532, May 2014, doi: <https://doi.org/10.1016/j.ctrv.2013.11.006>.
- [2] T. J. Fuchs, P. J. Wild, H. Moch, and J. M. Buhmann, “Computational pathology analysis of tissue microarrays predicts survival of renal clear cell carcinoma patients,” *Medical image computing and computer-assisted intervention: MICCAI ... International Conference on Medical Image Computing and Computer-Assisted Intervention*, vol. 11, no. Pt 2, pp. 1–8, 2008, doi: [https://doi.org/10.1007/978-3-540-85990-1\\_1](https://doi.org/10.1007/978-3-540-85990-1_1).
- [3] “Osteosarcoma data from UT Southwestern/UT Dallas for Viable and Necrotic Tumor Assessment (Osteosarcoma Tumor Assessment) - The Cancer Imaging Archive (TCIA) Public Access - Cancer Imaging Archive Wiki,” *wiki.cancerimagingarchive.net*. <https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=52756935#527569353f16e82c40944b93b979189b7a94689d> (accessed Jun. 05, 2023).
- [4] L. Mirabello, R. J. Troisi, and S. A. Savage, “Osteosarcoma incidence and survival rates from 1973 to 2004,” *Cancer*, vol. 115, no. 7, pp. 1531–1543, Apr. 2009, doi: <https://doi.org/10.1002/cncr.24121>.
- [5] G. Koutsomplia and G. I. Lambrou, “Resistance mechanisms in the radiation therapy of osteosarcoma: a brief review,” *Journal of Research and Practice on the Musculoskeletal System*, vol. 04, no. 01, pp. 15–19, Mar. 2020, doi: <https://doi.org/10.22540/jrpms-04-015>.
- [6] A. Misaghi, A. Goldin, M. Awad, and A. A. Kulidjian, “Osteosarcoma: a comprehensive review,” *SICOT-J*, vol. 4, no. 12, p. 12, 2018, doi: <https://doi.org/10.1051/sicotj/2017028>.
- [7] D. A. Levison and R. Muir, *Muir’s textbook of pathology*. London: Hodder Arnold, 2008.
- [8] M. J. Klein and G. P. Siegal, “Osteosarcoma,” *American Journal of Clinical Pathology*, vol. 125, no. 4, pp. 555–581, Apr. 2006, doi: <https://doi.org/10.1309/uc6kqhld9lv2kenn>.



- [9] F. N and M. LC, "Primary malignant giant cell tumour of the proximal tibia: a case report," *Journal of Cancer Research & Therapy*, vol. 1, no. 6, pp. 174–177, Aug. 2013, doi: <https://doi.org/10.14312/2052-4994.2013-26>.
- [10] Z. Ling, S. Yang, F. Gou, Z. Dai and J. Wu, "Intelligent Assistant Diagnosis System of Osteosarcoma MRI Image Based on Transformer and Convolution in Developing Countries," in *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 11, pp. 5563-5574, Nov. 2022, doi: 10.1109/JBHI.2022.3196043.
- [11] H. Babu Arunachalam *et al.*, "COMPUTER AIDED IMAGE SEGMENTATION AND CLASSIFICATION FOR VIABLE AND NON-VIABLE TUMOR IDENTIFICATION IN OSTEOSARCOMA," Jan. 2017, doi: [https://doi.org/10.1142/9789813207813\\_0020](https://doi.org/10.1142/9789813207813_0020).
- [12] S. Gawade, A. Bhansali, K. Patil, and D. Shaikh, "Application of the convolutional neural networks and supervised deep-learning methods for osteosarcoma bone cancer detection," *Healthcare Analytics*, vol. 3, p. 100153, Nov. 2023, doi: <https://doi.org/10.1016/j.health.2023.100153>.
- [13] I. A. Vezakis, G. I. Lambrou, and G. K. Matsopoulos, "Deep Learning Approaches to Osteosarcoma Diagnosis and Classification: A Comparative Methodological Approach," *Cancers (Basel)*, vol. 15, no. 8, Apr. 2023, doi: 10.3390/cancers15082290.
- [14] D. M. Anisuzzaman, H. Barzekar, L. Tong, J. Luo, and Z. Yu, "A Deep Learning Study on Osteosarcoma Detection from Histological Images," Nov. 2020, doi: 10.1016/j.bspc.2021.102931.
- [15] Z. Cai, O. Daescu, and M. Li, Eds., *Bioinformatics Research and Applications*, vol. 10330. in *Lecture Notes in Computer Science*, vol. 10330. Cham: Springer International Publishing, 2017. doi: 10.1007/978-3-319-59575-7.
- [16] K. O'Shea and R. Nash, "An Introduction to Convolutional Neural Networks," *ArXiv*, 2015, Accessed: Oct. 14, 2022. [Online]. Available: <https://www.semanticscholar.org/paper/An-Introduction-to-Convolutional-Neural-Networks-O%E2%80%99Shea-Nash/f46714d200d69eb9cb5cce176297b89a3f5e3a2c>
- [17] F. Zhuang *et al.*, "A Comprehensive Survey on Transfer Learning," Nov. 2019, [Online]. Available: <http://arxiv.org/abs/1911.02685>
- [18] M. Tan and Q. V Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks."
- [19] Loraksa, C., Mongkolsomlit, S., Nimsuk, N., Uscharapong, M., & Kiatisevi, P. (2021, December 27). *Effectiveness of Learning Systems from Common Image File Types to Detect Osteosarcoma Based on Convolutional Neural Networks (CNNs) Models*. MDPI. <https://doi.org/10.3390/jimaging8010002>

- [20] L. Cai, J. Gao, and D. Zhao, “A review of the application of deep learning in medical image classification and segmentation,” *Annals of Translational Medicine*, vol. 8, no. 11, pp. 713–713, Jun. 2020, doi: <https://doi.org/10.21037/atm.2020.02.44>.
- [21] Asito, Larissa Y., et al. "Detection of Osteosarcoma on Bone Radiographs Using Convolutional Neural Networks." *Proceedings of the 2021 Congresso Brasileiro de Inteligência Computacional, Joinville, Brasil*. 2021.