



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ & ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ  
ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ  
ΕΡΓΑΣΤΗΡΙΟ ΒΙΟΪΑΤΡΙΚΗΣ ΤΕΧΝΟΛΟΓΙΑΣ

**Αυτοματοποιημένη Διάγνωση και Πρόγνωση  
Παθολογιών μέσω Απεικονιστικών Εξετάσεων:  
Προσεγγίσεις Μηχανικής Μάθησης  
στο Οστεοσάρκωμα,  
στην Αλλεργική Δερματίτιδα Εξ Επαφής,  
και στο Αδενοκαρκίνωμα του Παγκρέατος**

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ  
ΤΟΥ  
**Ιωάννη Α. Βεζάκη**

Αθήνα, Νοέμβριος 2023





ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ  
ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ & ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ  
ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ  
ΕΡΓΑΣΤΗΡΙΟ ΒΙΟΪΑΤΡΙΚΗΣ ΤΕΧΝΟΛΟΓΙΑΣ

**Αυτοματοποιημένη Διάγνωση και Πρόγνωση  
Παθολογιών μέσω Απεικονιστικών Εξετάσεων:  
Προσεγγίσεις Μηχανικής Μάθησης  
στο Οστεοσάρκωμα,  
στην Αλλεργική Δερματίτιδα Εξ Επαφής,  
και στο Αδενοκαρκίνωμα του Παγκρέατος**

ΔΙΔΑΚΤΟΡΙΚΗ ΔΙΑΤΡΙΒΗ  
ΤΟΥ  
**Ιωάννη Α. Βεζάκη**

Συμβουλευτική Επιτροπή: Δημήτριος-Διονύσιος Κουτσούρης  
Γεώργιος Ματσόπουλος  
Παναγιώτης Τσανάκας

Εγκρίθηκε από την επταμελή εξεταστική επιτροπή την 21<sup>η</sup> Νοεμβρίου 2023.

---

Δ.-Δ. Κουτσούρης  
Ομ. Καθηγητής Ε.Μ.Π.

---

Γ. Ματσόπουλος  
Καθηγητής Ε.Μ.Π.

---

Π. Τσανάκας  
Καθηγητής Ε.Μ.Π.

---

Μ. Κωνσταντουλάκης  
Καθηγητής Ε.Κ.Π.Α.

---

Ε. Χριστοφόρου  
Καθηγητής Ε.Μ.Π.

---

Α. Παναγόπουλος  
Καθηγητής Ε.Μ.Π.

---

Π. Ασβεστάς  
Καθηγητής ΠΑ.Δ.Α.

---

Ιωάννης Α. Βεζάκης

Διδάκτωρ Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Copyright © Ιωάννης Α. Βεζάκης, 2023

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

*Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.*

*Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Εθνικού Μετσόβιου Πολυτεχνείου.*

# Περίληψη

Σε μια εποχή που χαρακτηρίζεται από πρωτοφανείς εξελίξεις στην τεχνολογία και την υγειονομική περίθαλψη, η συνεργασία της μηχανικής μάθησης και της ιατρικής αποτελεί την πλέον σύγχρονη πρόκληση για το επόμενο μεγάλο βήμα που θα φέρει την επανάσταση στον τρόπο με τον οποίο πραγματοποιείται η περίθαλψη των ασθενών. Στο επίκεντρο αυτής της τομής βρίσκεται ο τομέας της ανάλυσης ιατρικών εικόνων, όπου η συνεργασία της υπολογιστικής ικανότητας και της κλινικής εμπειρογνωμοσύνης κρύβει τεράστιες δυνατότητες.

Καθώς η ζήτηση και οι τρέχουσες ερευνητικές προσπάθειες επικεντρώνονται σε ολοένα πιο ακριβή και αποτελεσματικά διαγνωστικά και προγνωστικά εργαλεία, η παρούσα διδακτορική διατριβή προσπαθεί να πρωτοπορήσει σε νέες προσεγγίσεις που αξιοποιούν τη δύναμη της μηχανικής μάθησης με σκοπό τη διάγνωση και την πρόγνωση παθολογιών που εμφανίζονται σε τρεις διαφορετικές απεικονιστικές εξετάσεις. Πιο συγκεκριμένα, η έρευνα εστιάζει στην ανάλυση ιστολογικών εικόνων με παρουσία οστεοσαρκώματος, πολυφασματικών δερματοσκοπικών εικόνων με παρουσία αλλεργικής δερματίτιδας εξ επαφής, και προεγχειρητικών αξονικών τομογραφιών ασθενών με αδενοκαρκίνωμα του παγκρέατος.

Στην πρώτη μελέτη, η έρευνα επικεντρώνεται στην διάγνωση ιστολογικών εικόνων οστεοσαρκώματος. Συγκεκριμένα, υιοθετείται μια συγκριτική μεθοδολογική προσέγγιση για την αξιολόγηση σύγχρονων βαθιών νευρωνικών δικτύων τα οποία εκπαιδεύτηκαν σε ένα δημόσια διαθέσιμο σύνολο δεδομένων, με σκοπό την αυτοματοποιημένη ανίχνευση βιώσιμων και νεκρωτικών όγκων, καθώς και υγιούς ιστού. Η έρευνα εστιάζει στη σύγκριση διαφορετικών αρχιτεκτονικών δικτύων, του βάθους τους, καθώς και των διαστάσεων των εικόνων εισόδου, υπογραμμίζοντας την σημασία της προσεκτικής επιλογής τους, και πετυχαίνοντας βέλτιστη συνολική ακρίβεια 91% με τη χρήση του δικτύου MobileNetV2 με μέγεθος εικόνας εισόδου 256×256 εικονοστοιχεία.

Η δεύτερη μελέτη επικεντρώνεται στην ανάλυση δερματοσκοπικών εικόνων επιδερμικών δοκιμασιών επιθέματος με σκοπό τη διάγνωση Αλλεργικής Δερματίτιδας εξ Επαφής (ΑΔΕ) και τον εντοπισμό των υπεύθυνων αλλεργιογόνων, χρησιμοποιώντας μοντέλα βαθιάς μάθησης, διαφορετικές τεχνικές προεπεξεργασίας και εικονομορφές. Προτείνεται και αξιολογείται μια προσέγγιση βαθιάς μάθησης που χρησιμοποιεί μια τεχνική προεπεξεργασίας για την ανάδειξη περιοχών ενδιαφέροντος διατηρώντας παράλληλα τη συνολική πληροφορία των εικόνων. Επιπλέον,

αξιολογείται η χρησιμότητα διαφορετικών εικονομορφών, καθώς και του συνδυασμού τους, με σκοπό την επίτευξη της βέλτιστης διαγνωστικής ακρίβειας. Τα αποτελέσματα της μελέτης έδειξαν πως ο συνδυασμός της έγχρωμης εικόνας και του ερυθήματος μπορεί να επιφέρει τα βέλτιστα αποτελέσματα, πετυχαίνοντας ακρίβεια μεγαλύτερη του 91% στην ανίχνευση των αλλεργιογόνων στα οποία εμφάνισε ευαισθησία ο ασθενής. Η έρευνα αυτή αποτελεί μια από τις πρώτες ερευνητικές προσπάθειες για την αυτοματοποιημένη διάγνωση ΑΔΕ, ανοίγοντας έτσι τον δρόμο για την ανάπτυξη προηγμένων εργαλείων και τεχνικών που θα βοηθήσουν στην ταχύτερη και ακριβέστερη διάγνωση, μειώνοντας ταυτόχρονα τον φόρτο εργασίας των κλινικών ιατρών.

Στην τρίτη μελέτη, διερευνάται μιας πλήρως αυτοματοποιημένη προσέγγιση για την πρόγνωση του αδενοκαρκινώματος του παγκρέατος, με τη χρήση προεγχειρητικών αξονικών τομογραφιών ασθενών που πρόκειται να υποβληθούν σε παγκρεατεκτομή. Συγκεκριμένα, προτείνεται η χρήση τεχνικών βαθιάς μάθησης για την κατάτμηση του όγκου και του παγκρέατος, καθώς και η εξαγωγή ραδιομικών και κλινικών χαρακτηριστικών για τον υπολογισμό της συνολικής επιβίωσης με τη χρήση κλασικών τεχνικών μηχανικής μάθησης. Η προτεινόμενη προσέγγιση πέτυχε δείκτη συμφωνίας του Harrell (C-Index) 0,731, ξεπερνώντας τις υπάρχουσες μεθόδους που βασίζονται σε κλινικές μεταβλητές, όπως η σταδιοποίηση κατά TNM, αναδεικνύοντας παράλληλα την κρίσιμη σημασία συγκεκριμένων ραδιομικών χαρακτηριστικών, ιδίως του LAHGLE.

**Λέξεις-κλειδιά:** Μηχανική μάθηση, Βαθιά Μάθηση, Νευρωνικά Δίκτυα, Ιατρικές εικόνες, Οστεοσάρκωμα, Αλλεργική δερματίτιδα, Αδενοκαρκίνωμα, Πάγκρεας, Διάγνωση, Πρόγνωση, Ραδιομική, Ιστοπαθολογία, Δερματοσκόπηση, Επιδερμίδες Δοκιμασίες Επιθέματος, Αξονική Τομογραφία, Ακτινοδιαγνωστική

# Abstract

In an era characterized by groundbreaking advancements in technology and healthcare, the convergence of machine learning and medicine has emerged as a contemporary frontier poised to reshape the landscape of healthcare. Central to this convergence is the domain of medical image analysis, where the synergy between computational prowess and clinical expertise holds immense promise.

Amid escalating demands for precise and effective diagnostic and prognostic tools, this doctoral thesis endeavors to lead the way with novel approaches harnessing the potential of machine learning to diagnose and prognose pathologies across three distinct imaging modalities. Specifically, the research focuses on the analysis of histological images featuring osteosarcoma, multispectral dermoscopy images indicative of contact allergic dermatitis, and preoperative CT scans of patients with pancreatic ductal adenocarcinoma.

In the first study, the research delves into the diagnosis of histological osteosarcoma images. A comparative methodological approach is adopted to evaluate modern deep neural networks, trained on a publicly available dataset, for the automated detection of viable and necrotic tumors, as well as healthy tissue. The research emphasizes the importance of careful selection of different network architectures, their depth, and the dimensions of the input images, achieving optimal overall accuracy of 91% using the MobileNetV2 network and an input image size of 256×256 pixels.

The second study focuses on the diagnosis of Allergic Contact Dermatitis (ACD) using deep learning models, different pre-processing schemes and image modalities. A deep learning approach is employed, incorporating a pre-processing technique to highlight regions of interest while preserving the overall information of the images. Additionally, the study explores the utility of different image modalities and the combination thereof to achieve optimal diagnostic accuracy. The results showed that the combination of the color and redness image modalities produce optimal results, achieving an accuracy greater than 91% in detecting the allergens to which the patient manifested sensitivity to. This research represents one of the earliest efforts in automated diagnosis of ACD, paving the way for the development of advanced tools and techniques that expedite diagnosis while reducing clinicians' workload.

In the third study, a fully automated approach is proposed for the prog-

nostication of pancreatic adenocarcinoma, utilizing preoperative CT scans of patients undergoing pancreatectomy. This approach entails the use of deep learning techniques for tumor and pancreas segmentation, as well as the extraction of radiomics and clinical features to calculate overall survival using traditional machine learning techniques. The proposed approach achieved a Harrell's Concordance Index (C-Index) of 0.731, outperforming existing methods based on clinical parameters, such as TNM staging, while highlighting the importance of specific radiomics features, especially the LAHGLE.

**Keywords:** Machine learning, Deep learning, Neural Networks, Medical images, Osteosarcoma, Allergic dermatitis, Adenocarcinoma, Pancreas, Diagnosis, Prognosis, Radiomics, Histopathology, Dermoscopy, Patch Tests, Computed Tomography, Radiology



# Ευχαριστίες

Καθώς κλείνει ένα ακόμη σημαντικό κεφάλαιο του ακαδημαϊκού μου ταξιδιού, η παρούσα διδακτορική διατριβή δεν συμβολίζει μόνο το αποκορύφωμα μιας προσπάθειας επί σειράς ετών, αλλά και την υποστήριξη και καθοδήγηση που έλαβα κατά τη διάρκεια αυτής της πορείας.

Θα ήθελα να ξεκινήσω εκφράζοντας τις θερμές ευχαριστίες μου στον επιβλέποντα Καθηγητή μου, κ. Δημήτριο Κουτσούρη, για την ευκαιρία που μου έδωσε να εκπονήσω την παρούσα εργασία στο Εργαστήριο Βιοϊατρικής Τεχνολογίας, την εμπιστοσύνη του και την πολύτιμη καθοδήγησή του όλα αυτά τα χρόνια. Επιπλέον, ευχαριστώ θερμά τον Καθηγητή κ. Γεώργιο Ματσόπουλο, νέο Διευθυντή του Εργαστηρίου Βιοϊατρικής Τεχνολογίας, για τη συνεχή του παρουσία, καθοδήγηση, και υποστήριξη, οι οποίες υπήρξαν καθοριστικοί παράγοντες στη διαμόρφωση και ολοκλήρωση της διατριβής μου.

Ευχαριστώ επίσης την υπόλοιπη επιταμελή εξεταστική επιτροπή, τους Καθηγητές Ε.Μ.Π. κ. Παναγιώτη Τσανάκα, κ. Ευάγγελο Χριστοφόρου, και κ. Αθανάσιο Παναγόπουλο, τον Καθηγητή Ε.Κ.Π.Α. κ. Μανούσο Κωνσταντουλάκη, και τον Καθηγητή ΠΑ.Δ.Α. κ. Παντελεήμων Ασβεστά.

Θα ήθελα επιπλέον να ευχαριστήσω τον κ. Γεώργιο Λάμπρου, για την πολύτιμη βοήθεια και καθοδήγησή του στη συγγραφή των πρώτων μου επιστημονικών δημοσιεύσεων σε διεθνή περιοδικά. Τις κυρίες Άννα Τάγκα και Αικατερίνη Κυρίτση από την Α' Πανεπιστημιακή Κλινική Δερματικών και Αφροδίσιων Νοσημάτων του Νοσοκομείου « Ανδρέας Συγγρός » για την άψογη συνεργασία και συνεισφορά στη μελέτη της Αλλεργικής Δερματίτιδας Εξ Επαφής. Ειδική μνεία οφείλω στην Β' Χειρουργική Κλινική του Αρεταίειου Νοσοκομείου, και ιδιαίτερα στον πατέρα μου, Καθ. Αντώνιο Βεζάκη, ο οποίος με βοήθησε σε κάθε πτυχή της μελέτης μου στο αδενοκαρκίνωμα του παγκρέατος — από τη συγγραφή του κλινικού πρωτοκόλλου για την κατάθεση στην επιτροπή ηθικής και δεοντολογίας, έως την συλλογή, ανάλυση, και αξιολόγηση των δεδομένων. Επιπρόσθετα, ευχαριστώ τον Καθ. Ανδρέα Πολυδώρου, που συνέβαλε στον καθορισμό του θέματος της μελέτης από την πρώτη στιγμή, και που πάντα προσέφερε τη βοήθειά του σε οτιδήποτε χρειαζόμουν.

Τέλος, θα ήθελα να εκφράσω τη βαθύτατη ευγνωμοσύνη μου στην οικογένειά μου, τους γονείς μου Αντώνη-Γιούλα και τον αδελφό μου Ανδρέα, των οποίων η ακλόνητη πίστη στις δυνατότητές μου, η διαρκής ενθάρρυνση και η αδιάκοπη

υποστήριξή τους αποτέλεσαν τον πυρήνα της δύναμής μου, και οδηγό σε κάθε δύσκολη στιγμή αυτού του ταξιδιού.

# Περιεχόμενα

<b>1</b>	<b>Εισαγωγή</b>	<b>20</b>
1.1	Ιατρικά Απεικονιστικά Δεδομένα . . . . .	20
1.1.1	Ακτινοδιαγνωστική . . . . .	21
1.1.2	Ιστοπαθολογία . . . . .	24
1.1.3	Δερματοσκόπηση . . . . .	26
1.2	Μηχανική Μάθηση . . . . .	27
1.2.1	Το Πρόβλημα της Μάθησης . . . . .	27
1.2.2	Τρόποι Μάθησης . . . . .	27
1.2.3	Αλγόριθμοι Μάθησης . . . . .	28
1.3	Η Μηχανική Μάθηση ως Εργαλείο στην Ιατρική . . . . .	31
1.3.1	Διάγνωση . . . . .	32
1.3.2	Πρόγνωση . . . . .	35
1.4	Αντικείμενο και Συνεισφορές της Παρούσας Διδακτορικής Διατριβής	36
<b>2</b>	<b>Διάγνωση και Ταξινόμηση Οστεοσαρκώματος σε Ιστοπαθολογικές Εξετάσεις</b>	<b>38</b>
2.1	Οστεοσάρκωμα . . . . .	38
2.1.1	Παθογένεση . . . . .	38
2.1.2	Διάγνωση . . . . .	39
2.1.3	Πρόγνωση . . . . .	40
2.1.4	Θεραπεία . . . . .	40
2.2	Αντικείμενο της Μελέτης . . . . .	40
2.3	Υλικό και Μέθοδοι . . . . .	41
2.3.1	Σύνολο δεδομένων . . . . .	42
2.3.2	Συνελικτικά Νευρωνικά Δίκτυα . . . . .	42
2.3.3	Πειραματική διάταξη . . . . .	46
2.3.4	Αξιολόγηση των δικτύων . . . . .	49
2.3.5	Επακόλουθο πείραμα . . . . .	50
2.4	Αποτελέσματα . . . . .	51
2.4.1	Σύγκριση δικτύων . . . . .	51
2.4.2	Επακόλουθο πείραμα . . . . .	53
2.5	Συζήτηση . . . . .	55
2.5.1	Σύγκριση νευρωνικών δικτύων . . . . .	56

2.5.2	Περιορισμοί και μελλοντικές προοπτικές . . . . .	58
<b>3</b>	<b>Ανίχνευση Αλλεργικής Δερματίτιδας Εξ Επαφής σε Επιδερμидικές Δοκιμασίες Επιθέματος</b>	<b>60</b>
3.1	Αλλεργική Δερματίτιδα Εξ Επαφής . . . . .	60
3.1.1	Παθογένεση . . . . .	60
3.1.2	Διάγνωση . . . . .	61
3.2	Αντικείμενο της Μελέτης . . . . .	61
3.3	Υλικό και Μέθοδοι . . . . .	62
3.3.1	Σύνολο δεδομένων . . . . .	62
3.3.2	Πειραματική διάταξη . . . . .	66
3.3.3	Πρώτο πείραμα: Σύγκριση τεχνικών προεπεξεργασίας . . . . .	68
3.3.4	Δεύτερο πείραμα: Σύγκριση εικονομορφών . . . . .	70
3.3.5	Τρίτο πείραμα: Σύνηξη εικονομορφών . . . . .	73
3.3.6	Στατιστικός έλεγχος McNemar . . . . .	74
3.3.7	F-test 5×2 . . . . .	76
3.4	Αποτελέσματα . . . . .	77
3.4.1	Πρώτο πείραμα: Σύγκριση τεχνικών προεπεξεργασίας . . . . .	77
3.4.2	Δεύτερο πείραμα: Σύγκριση εικονομορφών . . . . .	78
3.4.3	Τρίτο πείραμα: Σύνηξη εικονομορφών . . . . .	79
3.5	Συζήτηση . . . . .	80
3.5.1	Περιορισμοί . . . . .	82
<b>4</b>	<b>Πρόγνωση Ασθενών με Αδενοκαρκίνωμα του Παγκρέατος σε Αξονι- κές Τομογραφίες</b>	<b>83</b>
4.1	Αδενοκαρκίνωμα του Παγκρέατος . . . . .	83
4.1.1	Διάγνωση . . . . .	83
4.1.2	Θεραπεία . . . . .	84
4.2	Αντικείμενο της Μελέτης . . . . .	84
4.3	Υλικό και Μέθοδοι . . . . .	85
4.3.1	Σύνολο Δεδομένων . . . . .	85
4.3.2	Αρχιτεκτονική Επισκόπηση . . . . .	86
4.3.3	Κατάτμηση του Παγκρέατος και του Όγκου . . . . .	87
4.3.4	Εξαγωγή Χαρακτηριστικών . . . . .	88
4.3.5	Επιλογή Χαρακτηριστικών . . . . .	89
4.3.6	Μοντελοποίηση Επιβίωσης . . . . .	90
4.3.7	Ανάλυση Σημαντικότητας Χαρακτηριστικών . . . . .	92
4.3.8	Ταξινόμηση Διαιτούς Επιβίωσης . . . . .	93
4.4	Αποτελέσματα . . . . .	93
4.4.1	Μοντελοποίηση Επιβίωσης . . . . .	93
4.4.2	Ανάλυση Σημαντικότητας Χαρακτηριστικών . . . . .	93
4.4.3	Ταξινόμηση Διαιτούς Επιβίωσης . . . . .	93

4.5	Συζήτηση . . . . .	95
4.5.1	Μοντελοποίηση Επιβίωσης . . . . .	95
4.5.2	Ανάλυση Σημαντικότητας Χαρακτηριστικών . . . . .	96
4.5.3	Ταξινόμηση Διαιτούς Επιβίωσης . . . . .	97
4.5.4	Περιορισμοί . . . . .	97
4.5.5	Μελλοντικές Προοπτικές . . . . .	98
<b>5</b>	<b>Συμπεράσματα</b>	<b>99</b>

# Κατάλογος σχημάτων

1.1	Η πρώτη ακτινογραφία που λήφθηκε ποτέ. Εικονίζεται το χέρι της συζύγου του W.K. Röntgen, φορώντας τη βέρα της. « Radiograph of hand by W.K. Roentgen » by Wellcome Library, London is licensed under CC BY 4.0 . . . . .	22
1.2	Αξονική τομογραφία κεφαλής, σε οβελιαίο, στεφανιαίο, και εγκάρσιο επίπεδο. . . . .	23
1.3	Ιστοπαθολογική εικόνα διαυγοκυτταρικού νεφρικού καρκινώματος, με χρώση H&E σε μεγέθυνση $\times 400$ . . . . .	25
1.4	Δερματοσκοπική εικόνα βασικοκυτταρικού καρκινώματος με τηλαγγειεκτατικά αγγεία. Image by Wikipedia user David.moreno72 is licensed under CC BY-SA 4.0 . . . . .	26
1.5	Γραφική αναπαράσταση αποτελέσματος μάθησης. Στην πρώτη εικόνα, το μοντέλο δεν έχει μάθει να προσαρμόζεται στα δεδομένα. Στη δεύτερη εικόνα, το μοντέλο προσεγγίζει την πραγματική λύση. Στην τρίτη εικόνα, το μοντέλο έχει προσαρμοστεί υπερβολικά στα διαθέσιμα δεδομένα, με κίνδυνο να μην μπορεί να γενικεύσει τη γνώση του σε νέα δείγματα. . . . .	28
2.1	Παράδειγμα εικόνων από το σύνολο δεδομένων για το οστεοσάρκωμα (μεγέθυνση $\times 10$ ). (A) Μη όγκος, (B) βιώσιμος όγκος, (C) νέκρωση. . . . .	43
2.2	Επισκόπηση της μεθοδολογίας για την συγκριτική μελέτη διαφόρων αρχιτεκτονικών νευρωνικών δικτύων στην ταξινόμηση του οστεοσαρκώματος. Οι ιστολογικές εικόνες μετασχηματίζονται στις επιθυμητές διαστάσεις και κανονικοποιούνται πριν τη χρήση τους για την εκπαίδευση των δικτύων. Στο διάγραμμα, με τον όρο EfficientNets, MobileNets κλπ, εννοείται όλη η οικογένεια των δικτύων αυτών, όπως φαίνεται στον πίνακα 2.6. Η βέλτιστη αρχιτεκτονική επιλέγεται και επανεκπαιδεύεται στο ίδιο σύνολο δεδομένων, με τη χρήση πενταπλής διασταυρούμενης επικύρωσης. . . . .	46
2.3	Καμπύλες ROC για: (A) Μη όγκο, (B) βιώσιμο όγκο, (C) νέκρωση. . . . .	55

3.1 Ένα ενιαίο επίθεμα χωρίζεται σε δύο περιοχές, την Περιοχή 1 και την Περιοχή 2. Κάθε περιοχή καταγράφεται ανεξάρτητα με την κάμερα λόγω του ότι το οπτικό της πεδίο περιορίζεται σε 56 × 56 mm. Κάθε αριθμός αντιστοιχεί σε έναν θαλαμίσκο επιδερμικού επιθέματος. . . . .	66
3.2 Δείγματα των διαφορετικών εικονομορφών για την ίδια περιοχή του δέρματος ενός ασθενούς. <b>(Α)</b> Έγχρωμη εικόνα, <b>(Β)</b> Ερύθημα, <b>(Γ)</b> Υφή, <b>(Δ)</b> Λεπτές Γραμμές, <b>(Ε)</b> Πτυχώσεις, <b>(ΣΤ)</b> Ανισοϋψώσεις. . . . .	67
3.3 Ένα παράδειγμα που αναδεικνύει γιατί η συνολική πληροφορία της εικόνας είναι σημαντική. Στο αλλεργιογόνο με την επισήμανση <b>A</b> έχει εκδηλωθεί πολύ έντονη αντίδραση. Δεδομένων των συνθηκών, εάν το αλλεργιογόνο <b>B</b> εξεταστεί μεμονωμένα, είναι πιθανό να προκύψει ψευδώς θετικό αποτέλεσμα αφού φαίνεται να έχει αυξημένη ερυθρότητα. . . . .	69
3.4 Το αποτέλεσμα που προκύπτει για κάθε μία από τις τρεις διαφορετικές τεχνικές προεπεξεργασίας. <b>(Α)</b> η αρχική εικόνα με επισημασμένη την περιοχή ενδιαφέροντος (ROI), <b>(Β)</b> η εικόνα περικόπτεται στην ROI, <b>(Γ)</b> MA - η αρχική εικόνα με εφαρμογή Μάσκας <i>a</i> <b>(Δ)</b> EMA - η εικόνα με Μάσκα <i>a</i> και επιπλέον προσθήκη πρόσθετων πληροφοριών από την παρακείμενη περιοχή του επιδερμικού επιθέματος. . . . .	71
3.5 Διαγραμματική αναπαράσταση της πειραματικής διαδικασίας για την αξιολόγηση των τριών διαφορετικών τεχνικών προεπεξεργασίας των εικόνων. . . . .	72
3.6 Διαγραμματική αναπαράσταση της πειραματικής διαδικασίας του δεύτερου πειράματος. Σε αυτό το πείραμα πραγματοποιείται σύγκριση της επίδοσης που μπορεί να επιτευχθεί με χρήση κάθε εικονομορφής ξεχωριστά. . . . .	73
3.7 Διαγραμματική αναπαράσταση της πειραματικής διαδικασίας του τρίτου πειράματος. Σε αυτό το πείραμα πραγματοποιείται σύγκριση της επίδοσης που μπορεί να επιτευχθεί με κάθε συνδυασμό ανά δύο όλων των εικονομορφών. . . . .	75
3.8 Ραβδόγραμμα αποτελεσμάτων για τις τρεις διαφορετικές τεχνικές προεπεξεργασίας. . . . .	77
3.9 Σύγκριση μετρικών για όλες τις διαφορετικές μορφές εικόνας. Κάθε τιμή υπολογίζεται ως ο μέσος όρος σε μια πενταπλή διασταυρούμενη επικύρωση. Η γραμμή σφάλματος αντιστοιχεί στην τυπική απόκλιση. . . . .	78
4.1 Καμπύλη επιβίωσης Kaplan-Meier για 40 ασθενείς με αδενοκαρκίνωμα του παγκρέατος, έπειτα από παγκρεατεκτομή στο Αρεταίειο Νοσοκομείο. . . . .	85
4.2 Αρχιτεκτονική επισκόπηση της προτεινόμενης μεθοδολογίας. . . . .	87

4.3	Παράδειγμα που δείχνει την κατάτμηση του αδενοκαρκινώματος (κόκκινο) και του υπόλοιπου παγκρέατος (πράσινο). Η εικόνα <b>(A)</b> απεικονίζει την αρχική εικόνα, και η εικόνα <b>(B)</b> απεικονίζει την εικόνα με τις τμηματοποιήσεις του παγκρέατος και του όγκου. . . . .	88
4.4	Τα 20 κορυφαία χαρακτηριστικά βάσει της μέσης σημαντικότητάς τους. . . . .	94



# Κατάλογος πινάκων

2.1	Ο αριθμός των παραμέτρων για κάθε παραλλαγή δικτύου. . . . .	48
2.2	Βαθμολογία $F1$ για κάθε συνδυασμό αρχιτεκτονικής δικτύου και διαστάσεων των εικόνων εισόδου. Οι κατηγορίες MO, BO, και NK αναφέρονται σε μη όγκο, βιώσιμο όγκο, και νέκρωση, αντίστοιχα. .	52
2.3	Μέσοι όροι και τυπικές αποκλίσεις διαφόρων μετρικών για την απόδοση του MobileNetV2 υπολογισμένοι με χρήση πενταπλής διασταυρούμενης επικύρωσης. . . . .	53
2.4	Πίνακας σύγχυσης του MobileNetV2 με συγκεντρωτικά αποτελέσματα για την πενταπλή διασταυρούμενη επικύρωση. . . . .	54
2.5	Μέσοι όροι και τυπικές αποκλίσεις των μετρικών απόδοσης του MobileNetV2 για πέντε διαμερίσεις, εξαιρουμένων των διφορούμενων εικόνων από το υπό εξέταση σύνολο δεδομένων (ως διφορούμενες εικόνες θεωρήθηκαν εκείνες που περιείχαν ταυτόχρονα NK και BO). .	54
2.6	Σύγκριση της συνολικής ακρίβειας με σχετική εργασία στο ίδιο σύνολο δεδομένων. . . . .	56
3.1	Σειρά αλλεργιογόνων που χρησιμοποιήθηκε στις επιδερμικές δοκιμασίες. . . . .	64
3.2	Παράδειγμα πίνακα ενδεχομένων για τον στατιστικό έλεγχο McNemar. .	74
3.3	Μετρικές των αποτελεσμάτων για κάθε τεχνική προεπεξεργασίας. . .	77
3.4	Τιμές $p$ που αντιστοιχούν στη σύγκριση ανά ζεύγη μεταξύ των διαφορετικών τεχνικών προεπεξεργασίας, υπολογισμένες με χρήση του τεστ McNemar. . . . .	78
3.5	Μετρικές αποτελεσμάτων για κάθε μορφή εικόνας. . . . .	78
3.6	Τιμές $p$ που αντιστοιχούν στη σύγκριση κατά ζεύγη των διαφορετικών μορφών απεικόνισης, υπολογισμένες με τη χρήση του τεστ McNemar. Οι λαμβανόμενες τιμές $p$ αντιστοιχούν στον πρόσθετο διαχωρισμό 70/30 και όχι στα αποτελέσματα της διασταυρούμενης επικύρωσης που απεικονίζονται στο Σχήμα 3.9. . . . .	79
3.7	Βαθμολογία $F1$ όπως υπολογίστηκε από τη διασταυρούμενη επικύρωση $5 \times 2$ σε κάθε συνδυασμό εικονομορφών. . . . .	80

3.8	Τιμές $p$ που αντιστοιχούν στη σύγκριση κατά ζεύγη της ταξινόμησης των χαρακτηριστικών που μαθαίνονται αποκλειστικά από την έγχρωμη μορφή, και του συνδυασμού των χαρακτηριστικών που μαθαίνονται από την σύντηξη όλων των άλλων μορφών, όπως υπολογίζεται με τη χρήση του $f$ -τεστ $5 \times 2$ . . . . .	81
4.1	Το πρωτόκολλο παγκρέατος για τη λήψη αξονικών τομογραφιών άνω-κάτω κοιλίας σε ασθενείς με αδενοκαρκίνωμα του παγκρέατος. . . .	86
4.2	Πλήρης κατάλογος των 37 χαρακτηριστικών έπειτα από επιλογή χαρακτηριστικών με τη χρήση του συντελεστή διόγκωσης της διακύμανσης (VIF). . . . .	91
4.3	Μετρικές επιδόσεων για το πρόβλημα δυαδικής ταξινόμησης. Οι ασθενείς που επέζησαν πέραν των 24 μηνών θεωρήθηκαν θετικά δείγματα. . . . .	94

# Κατάλογος Ακρωνυμίων

**AUC** Area Under the Receiver Operating Curve

**CI** Concordance Index

**CT** Computed Tomography

**MRI** Magnetic Resonance Imaging

**RF** Random Forests

**RNN** Recurrent Neural Network

**SVM** Support Vector Machine

**VIF** Variance Inflation Factor

**WSI** Whole Slide Images

**k-NN** k-Nearest Neighbors

**ΑΔΕ** Αλλεργική Δερματίτιδα εξ Επαφής

**BNN** Βαθύ Νευρωνικό Δίκτυο

**MM** Μηχανική Μάθηση

**ΣNN** Συνελικτικό Νευρωνικό Δίκτυο

**TN** Τεχνητή Νοημοσύνη

**ΧΜΘ** Χημειοθεραπεία

# Κεφάλαιο 1

## Εισαγωγή

Μπορεί η Μηχανική Μάθηση (MM) να αυτοματοποιήσει την αξιολόγηση απεικονιστικών ιατρικών εξετάσεων, συμβάλλοντας σε μια ταχύτερη, ακριβέστερη, και πιο αξιόπιστη διαδικασία λήψης αποφάσεων; Η Εθνική Ακαδημία Ιατρικής των ΗΠΑ υπολογίζει πως σχεδόν κάθε ασθενής θα εκτεθεί σε τουλάχιστον ένα διαγνωστικό σφάλμα κατά τη διάρκεια της ζωής του [1], γεγονός που καθιστά τη βελτίωση της διαγνωστικής διαδικασίας ως ηθική, επαγγελματική, και υγειονομική επιταγή.

Χαρακτηριστικό των τελευταίων ετών αποτελεί η αλματώδης ανάπτυξη της MM, η οποία κατέστη εφικτή από την ολοένα αυξανόμενη επεξεργαστική ισχύ, μνήμη, και αποθηκευτική ικανότητα των ηλεκτρονικών υπολογιστών. Αν και σήμερα το ενδιαφέρον της κοινής γνώμης σχεδόν μονοπωλείται από τα εντυπωσιακά επιτεύγματα της MM στην Επεξεργασία Φυσικής Γλώσσας (αγγλ. Natural Language Processing – NLP) με τη χρήση μετασχηματισμών [2] (βλ. ChatGPT), η Επεξεργασία Εικόνας αποτελεί ένα πάγια σημαντικό και συνάμα ταχύτατα εξελισσόμενο πεδίο.

### 1.1 Ιατρικά Απεικονιστικά Δεδομένα

Στο διαρκώς εξελισσόμενο τοπίο της σύγχρονης ιατρικής, η σημασία των οπτικών πληροφοριών είναι υψίστης σημασίας. Τα ιατρικά απεικονιστικά δεδομένα, με την ικανότητά τους να αποτυπώνουν περίπλοκες λεπτομέρειες του ανθρώπινου σώματος, αποτελούν απαραίτητο εργαλείο για τη διάγνωση, τη θεραπεία και την κατανόηση πολλών ιατρικών καταστάσεων. Οι εικόνες αυτές προέρχονται από διάφορες πηγές, συμπεριλαμβανομένων των ακτινολογικών εξετάσεων, των παθολογοανατομικών δειγμάτων, των ενδοσκοπήσεων, κ.ά. Είτε εξερευνώντας τις λεπτομέρειες των εσωτερικών οργάνων, είτε εξετάζοντας κυτταρικές ανωμαλίες, είτε παρακολουθώντας την πρόοδο μιας θεραπευτικής αγωγής, οι ιατρικές εικόνες δραματίζουν καθοριστικό ρόλο σχεδόν σε κάθε πτυχή της υγειονομικής περίθαλψης.

Για τους παραπάνω λόγους, τα ιατρικά απεικονιστικά δεδομένα αποτελούν συνήθως τον μεγαλύτερο όγκο ιατρικών δεδομένων σε έναν ηλεκτρονικό φάκελο

υγείας ασθενή. Η ακτινογραφία, η αξονική τομογραφία, η μαγνητική τομογραφία, το υπερηχογράφημα, αποτελούν κάποιες από τις συνηθέστερες ιατρικές εξετάσεις, το αποτέλεσμα των οποίων είναι εικόνες. Κατά συνέπεια, γίνεται αντιληπτό πως πρωτίστως η επεξεργασία εικόνας αποτελεί βασικό εργαλείο για την βελτίωση και αυτοματοποίηση της αξιολόγησης ιατρικών ευρημάτων.

### **1.1.1 Ακτινοδιαγνωστική**

Η ακτινοδιαγνωστική, αναφερόμενη ως *medical imaging* στα αγγλικά, αποτελεί ακρογωνιαίο λίθο της σύγχρονης ιατρικής, έχοντας φέρει μια πραγματική επανάσταση στον τρόπο με τον οποίο πραγματοποιείται η διάγνωση και διαχείριση ασθενειών, τραυματισμών και παθήσεων. Ο τομέας αυτός έχει υποστεί σημαντική μεταμόρφωση από την ίδρυσή του, με πρωτοποριακές τεχνολογίες και καινοτομίες που διαμορφώνουν συνεχώς το τοπίο του.

Ο όρος « ακτινοδιάγνωση » αναφέρεται στη μη επεμβατική απεικόνιση των εσωτερικών δομών και λειτουργιών του ανθρώπινου σώματος, η οποία επιτρέπει στους επαγγελματίες υγείας να λαμβάνουν κρίσιμες πληροφορίες σχετικά με την κατάσταση ενός ασθενούς, βοηθώντας στη διάγνωση, τον προγραμματισμό της θεραπείας, και την παρακολούθηση. Καταγράφοντας λεπτομερείς εικόνες οργάνων, ιστών και βιολογικών διεργασιών, η ακτινοδιαγνωστική βρίσκεται στο επίκεντρο της σύγχρονης υγειονομικής περίθαλψης.

Η πορεία του τομέα της ακτινοδιαγνωστικής ξεκίνησε το 1895, με μια τυχαία ανακάλυψη του Γερμανού καθηγητή φυσικής Wilhelm Conrad Röntgen [3]. Κατά τη διάρκεια ενός πειράματος με δέσμες ηλεκτρονίων, ένα απροσδόκητο φαινόμενο εκτυλίχθηκε στο εργαστήριο του Röntgen; παρατήρησε ότι μια φθορίζουσα οθόνη άρχισε να λάμπει παρόλο που δεν βρισκόταν στην άμεση πορεία των καθοδικών ακτίνων. Αρχικά, ο Röntgen υπέθεσε ότι η λάμψη των καθοδικών ακτίνων ήταν υπεύθυνη για αυτόν τον φθορισμό, γεγονός που τον ώθησε να καλύψει τον εξοπλισμό του για να περιορίσει το φως. Ωστόσο, προς έκπληξή του, ο φθορισμός της οθόνης συνέχισε να προκαλείται.

Μετά από περαιτέρω έρευνα, ο Röntgen αποκάλυψε την ύπαρξη μιας αόρατης μορφής ακτινοβολίας, ικανής να διαπερνά διάφορα υλικά, συμπεριλαμβανομένου του ανθρώπινου σώματος. Επιπλέον, αυτές οι μυστηριώδεις ακτίνες άφηναν πίσω τους σκιάδες εικόνες στερεών αντικειμένων όταν κατευθύνονταν σε φωτογραφικές πλάκες. Ακόμη πιο ενδιαφέρον ήταν ότι, όταν οι ακτίνες διαπερνούσαν την ανθρώπινη σάρκα, δημιουργούσαν σκιάδες εικόνες των οστών. Έτσι, η πρώτη « ακτινογραφία » που καταγράφηκε ποτέ, απεικονίζει το χέρι της συζύγου του Röntgen, απεικονίζοντας τα οστά της και τη βέρα της (Σχήμα 1.1). Λόγω της αινιγματικής και άγνωστης φύσης αυτών των ακτίνων, ο Röntgen τις ονόμασε ακτίνες « X » [3].

Η μνημειώδης ανακάλυψη του Röntgen πυροδότησε μια επαναστατική αλλαγή στον τομέα της ιατρικής. Για πρώτη φορά στην ιστορία, η μη επεμβατική απεικόνιση των εσωτερικών δομών του ανθρώπινου σώματος έγινε πραγματικότητα. Αυτή



Σχήμα 1.1: Η πρώτη ακτινογραφία που λήφθηκε ποτέ. Εικονίζεται το χέρι της συζύγου του W.K. Röntgen, φορώντας τη βέρα της. « Radiograph of hand by W.K. Roentgen » by Wellcome Library, London is licensed under CC BY 4.0



Σχήμα 1.2: Αξονική τομογραφία κεφαλής, σε οβελιαίο, στεφανιαίο, και εγκάρσιο επίπεδο.

η ανακάλυψη σηματοδότησε την έναρξη της ακτινοδιαγνωστικής, ενός τομέα που θα άλλαζε για πάντα το τοπίο της υγειονομικής περίθαλψης. Το πρωτοποριακό έργο του Röntgen, του χάρισε το βραβείο Νόμπελ Φυσικής το 1901, και έκτοτε η τεχνολογία των ακτίνων X έχει γίνει απαραίτητο εργαλείο στην ιατρική αλλά και άλλους τομείς.

Ωστόσο, παρά τον αρχικό ενθουσιασμό, οι περιορισμοί έγιναν γρήγορα αντιληπτοί, καθώς οι λαμβανόμενες ακτινογραφίες περιορίζονταν στις δύο διαστάσεις. Έναν χρόνο μετά την ανακάλυψη των ακτίνων X, ο E. Thompson προσπαθούσε ήδη να παράξει μια τρισδιάστατη ακτινογραφία, χρησιμοποιώντας στεροσκοπικές τεχνικές [4]. Η τεχνολογία όμως δεν επέτρεψε την ολική επίτευξη του στόχου αυτού παρά μόνο αρκετές δεκαετίες αργότερα. Έπρεπε να επέλθει η ανάπτυξη των ηλεκτρονικών υπολογιστών για να γίνει εφικτή η εφεύρεση της «υπολογιστικής» ή «αξονικής» τομογραφίας, αγγλ. Computed Tomography (CT), τη δεκαετία του 1970 από τους Allan Cormack και Godfrey Hounsfield, για την οποία κέρδισαν το βραβείο Νόμπελ το 1979. Η εφεύρεση του αξονικού τομογράφου έφερε εκ νέου μια επανάσταση στη διαγνωστική τεχνολογία, λόγω της δυνατότητάς του να απεικονίζει το εσωτερικό ενός ασθενούς ως εγκάρσιες τομές υψηλής ανάλυσης, παρέχοντας έτσι μια λεπτομερή τρισδιάστατη ανατομική εικόνα με μη-επεμβατικό τρόπο (Σχήμα 1.2).

Ωστόσο, δεν ήταν μόνο ο αξονικός τομογράφος που μπορούσε, τότε, να απεικονίσει τομές του σώματος. Το υπερηχογράφημα είχε ήδη εφευρεθεί από τη δεκαετία του 1940 [5], όμως ήταν και αυτό περιορισμένο στις δύο διαστάσεις. Επιπλέον, η αξονική τομογραφία υπερτερούσε τόσο όσον αφορά την ανάλυση της εικόνας, όσο και την ικανότητα αποτύπωσης λεπτών ανατομικών λεπτομερειών.

Τη δεκαετία του 1980, η εφεύρεση της μαγνητικής τομογραφίας, αγγλ. Magnetic Resonance Imaging (MRI), αποτέλεσε μία ακόμη επανάσταση [4], [5]. Πλέον, η απεικόνιση δεν περιοριζόταν από τις ανησυχίες σχετικά με την έκθεση στην ακτινοβολία και την ασφάλεια, επιτρέποντας τη διενέργεια περισσότερων εξετάσεων, εντός ενός μεγαλύτερου χρονικού πλαισίου. Επιπλέον, το ευρύ φάσμα πληροφοριών που μπορούσε να λάβει η MRI επέτρεψε την μέτρηση ενός πλήθους φυσιολογικών διεργασιών όπως της ροής του αίματος, της νευρικής δραστηριότητας, της εναπόθεσης σιδήρου κ.ά.

Έκτοτε, η ακτινοδιαγνωστική συνεχίζει να εξελίσσεται. Οι παραγόμενες εικόνες αποκτούν ολοένα και μεγαλύτερη ανάλυση, λεπτομέρεια και ευκρίνεια, ενώ συνεχώς περισσότερα ιατρικά κέντρα αποκτούν σύγχρονα ακτινοδιαγνωστικά μηχανήματα, με τις ακτινοδιαγνωστικές εξετάσεις να αποτελούν σήμερα ένα απαραίτητο και αναντικατάστατο εργαλείο στην ιατρική. Καθώς τα παραγόμενα ιατρικά απεικονιστικά δεδομένα συνεχώς αυξάνονται, η βελτίωση της διαδικασίας ανάγνωσής τους, αποτελεί μία επιτακτική ανάγκη για το άμεσο μέλλον.

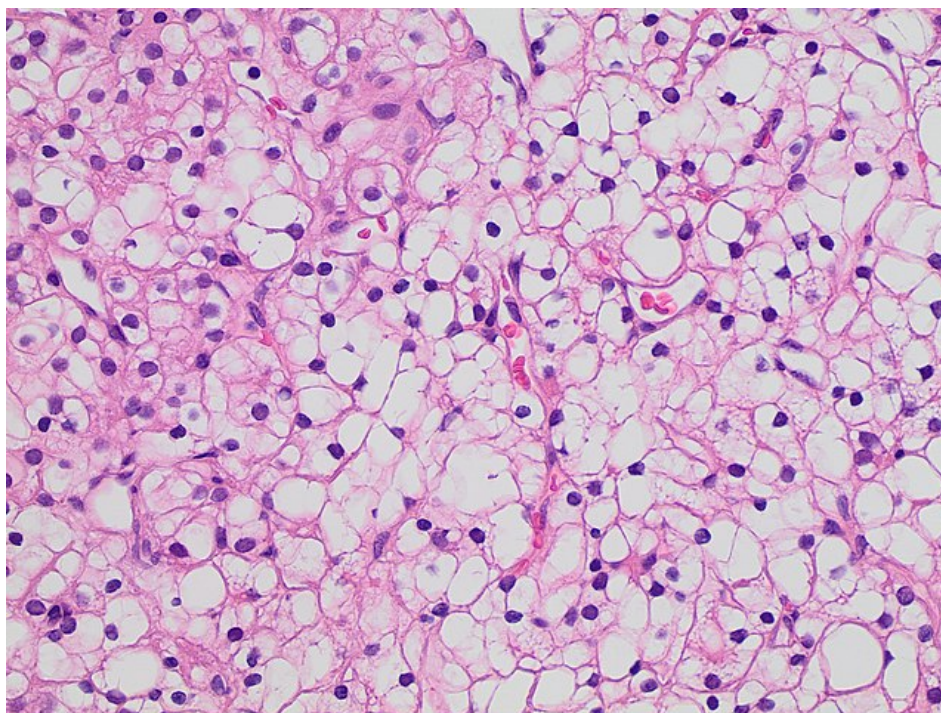
### **1.1.2 Ιστοπαθολογία**

Αν και, όπως αναφέρθηκε προηγουμένως, η ακτινοδιαγνωστική έχει φέρει μια επανάσταση στη διάγνωση και παρακολούθηση πολλών ιατρικών καταστάσεων, είναι σημαντικό να αναγνωριστεί πως το πεδίο των ιατρικών εικόνων εκτείνεται πέρα από την ακτινολογία. Η ιστοπαθολογία, ή αλλιώς παθολογική ανατομία, αποτελεί μια κρίσιμη πτυχή της ιατρικής πρακτικής, αποκαλύπτοντας μια εντελώς διαφορετική διάσταση στην απεικόνιση ασθενειών, παθήσεων και κυτταρικών ανωμαλιών.

Ενώ η ακτινοδιαγνωστική παρέχει πληροφορίες σχετικά με τις δομικές πτυχές του ανθρώπινου σώματος, η ιστοπαθολογική απεικόνιση εμβαθύνει περισσότερο, εξερευνώντας τον μικροσκοπικό κόσμο των ιστών και των κυττάρων (Σχήμα 1.3). Πιο συγκεκριμένα, η ιστοπαθολογία αφορά την εξέταση μιας «βιοψίας», δηλαδή ενός δείγματος ιστού, υπό το μικροσκόπιο, αφού το δείγμα έχει υποστεί επεξεργασία και οι ιστολογικές τομές έχουν τοποθετηθεί σε γυάλινες αντικειμενοφόρες πλάκες.

Στα τέλη του 1960, εμφανίζεται η πρώτη καταγεγραμμένη περίπτωση «τηλεπαθολογίας», αναφερόμενη σε μία υπηρεσία *τηλεοπτικής μικροσκοπίας* σε πραγματικό χρόνο, η οποία δημιουργήθηκε μεταξύ του Γενικού Νοσοκομείου της Μασαχουσέτης, και του ιατρικού σταθμού του αεροδρομίου Logan στη Βοστώνη [6]. Έκτοτε, οι παθολογοανατομικές εξετάσεις εξ αποστάσεως έγιναν ένα συχνότερο φαινόμενο, αν και ο παθολογοανατόμος που πραγματοποιούσε την εξέταση δεν





Σχήμα 1.3: Ιστοπαθολογική εικόνα διαυγοκυτταρικού νεφρικού καρκινώματος, με χρώση H&E σε μεγέθυνση  $\times 400$ .

μπορούσε να χειρίζεται ο ίδιος τον εξοπλισμό. Ωστόσο, αυτός ο περιορισμός σταδιακά εξαλείφθηκε, αρχικά με τον ρομποτικό τηλεχειρισμό του εξοπλισμού (δυναμική τηλεπαθολογία), και αργότερα με τις ψηφιακές Εικόνες Ολόκληρων Αντικειμενοφόρων Πλακών, αγγλ. Whole Slide Images (WSI). Μια WSI εικόνα είναι μία ψηφιακή φωτογραφία πολύ υψηλής ανάλυσης που απεικονίζει ολόκληρο το δείγμα ιστού (παρασκευάσμα), όπως αυτό έχει ληφθεί από έναν, ειδικά σχεδιασμένο για αυτό τον σκοπό, ψηφιακό σαρωτή. Επομένως, πλέον έχουμε εισέλθει στην εποχή της *ψηφιακής ιστοπαθολογίας*, όπου η εξέταση των δειγμάτων πραγματοποιείται μέσω του ηλεκτρονικού υπολογιστή, και όχι απευθείας από το μικροσκόπιο.

Σήμερα, η εξέλιξη της ιστοπαθολογίας έχει επιφέρει νέες προκλήσεις. Για παράδειγμα, ο φόρτος εργασίας στα εργαστήρια διαγνωστικής ιστοπαθολογίας αυξάνεται παγκοσμίως χρόνο με το χρόνο. Στο Ηνωμένο Βασίλειο, η αύξηση του αριθμού των αιτημάτων ιστοπαθολογικής εξέτασης υπολογίζεται σε περίπου 4,5% κάθε έτος [7]. Επιπλέον, η προσπάθεια εντοπισμού προκαρκινικών καταστάσεων και καρκίνων σε πρώιμο στάδιο αυξάνει την πολυπλοκότητα της ιστοπαθολογικής αξιολόγησης, αφού οι παθολογοανατόμοι καλούνται να εντοπίζουν ανεπαίσθητες λεπτομέρειες σε μικρότερα διαγνωστικά δείγματα [7]. Εύκολα γίνεται αντιληπτό, λοιπόν, πως η ψηφιακή επεξεργασία και ανάλυση εικόνας μπορεί να διαδραματίσει σημαντικό ρόλο και στην ιστοπαθολογία, βελτιώνοντας τη διαδικασία της αξιολόγησης.



Σχήμα 1.4: Δερματοσκοπική εικόνα βασικοκυτταρικού καρκινώματος με τηλαγγειεκτατικά αγγεία. Image by Wikipedia user David.moreno72 is licensed under CC BY-SA 4.0

### 1.1.3 Δερματοσκόπηση

Η δερματοσκόπηση αντιπροσωπεύει μια σημαντική μεθοδολογική πρωτοπορία στον τομέα της δερματολογίας. Αυτή η προηγμένη τεχνική επιτρέπει την *in vivo* απεικόνιση των μελαγχρωματικών δερματικών βλαβών, σε βάθος που φτάνει μέχρι και το χόριο, δηλαδή το δίκτυο κολλαγόνων ινών που βρίσκεται κάτω από την επιδερμίδα.

Η εν λόγω μέθοδος χρησιμοποιεί μια εξειδικευμένη συσκευή, παρόμοια στη μορφή με το ωτοσκόπιο, αλλά εξοπλισμένη με έναν ειδικό φακό επαφής, που είναι γνωστός ως δερματοσκόπιο. Το δερματοσκόπιο αυτό εκπέμπει μια στοχευμένη δέσμη φωτός υπό ακριβή γωνία 20° στην επιφάνεια του δέρματος. Με την εφαρμογή ενός υγρού, όπως λάδι, νερό, τζελ, ή γλυκερίνης, στη διεπιφάνεια μεταξύ της επιδερμίδας και της γυάλινης αντικειμενοφόρου πλάκας της συσκευής, εξουδετερώνονται αποτελεσματικά οι ανακλάσεις του φωτός. Κατά συνέπεια, αυτή η εξάλειψη των ανακλάσεων επιτρέπει τη σαφή απεικόνιση των δερματοσκοπικών χαρακτηριστικών, τα οποία προέρχονται από την παρουσία μελανίνης και αιμοσφαιρίνης στα διάφορα στρώματα του δέρματος (Σχήμα 1.4). Πιο πρόσφατα, τα σύγχρονα δερματοσκόπια πολωμένου φωτός έχουν εξαλείψει την ανάγκη της χρήσης υγρού. Συνήθως, τα δερματοσκόπια προσφέρουν 10πλάσια μεγέθυνση, αλλά η έλευση των ψηφιακών δερματοσκοπιών έχει εισαγάγει μεγεθύνσεις έως και 70πλάσιες, διατηρώντας παράλληλα τη σαφήνεια και την ευκρίνεια της εικόνας.

Σε αντίθεση με τις εικόνες που λαμβάνονται στην ιστοπαθολογία, οι εικόνες του δερματοσκοπίου δεν επικεντρώνονται μόνο σε ένα επίπεδο του δέρματος, αλλά αποτελούνται από την υπέρθεση πολλαπλών δερματικών επιπέδων. Οι παραγόμενες εικόνες, επομένως, διαφέρουν τόσο από μια απλή κλινική εικόνα, όσο και από μια ιστολογική εικόνα. Συνεπώς, η δερματοσκόπηση αποτελεί μια χρήσιμη διεπαφή μεταξύ των δύο αυτών τομέων.

## 1.2 Μηχανική Μάθηση

Η ΜΜ αποτελεί ένα υποπεδίο της Τεχνητής Νοημοσύνης (ΤΝ). Η ΤΝ αναφέρεται σε υπολογιστικά συστήματα τα οποία μπορούν να μιμηθούν ανθρώπινες γνωστικές ικανότητες όπως η ανάλυση δεδομένων, η επίλυση προβλημάτων, και τελικά η λήψη αποφάσεων. Οι λειτουργίες αυτές πραγματοποιούνται σύμφωνα με ένα προκαθορισμένο σύνολο κανόνων το οποίο έχει ορίσει προηγουμένως ο προγραμματιστής του συστήματος. Ωστόσο, η ΜΜ δεν απαιτεί την εισαγωγή των κανόνων από τον προγραμματιστή. Αντιθέτως, ένας αλγόριθμος ο οποίος διαθέτει την ικανότητα μάθησης, πρέπει να μπορεί να είναι σε θέση να «μάθει», ή να εξάγει, μόνος του, τους απαραίτητους κανόνες αναλύοντας κάποιο σύνολο δεδομένων.

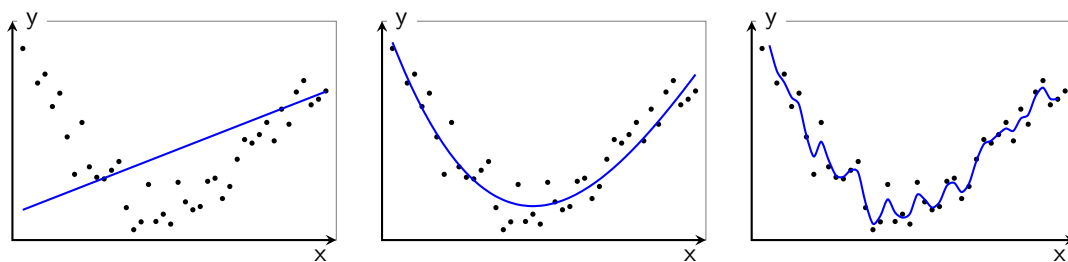
### 1.2.1 Το Πρόβλημα της Μάθησης

Η μάθηση αποτελεί μια διαδικασία κατά την οποία αποκτάται κάποια καινούργια γνώση, δεξιότητα, ή συμπεριφορά. Επομένως, κάθε υπολογιστικό σύστημα το οποίο δύναται να αλλάξει τον τρόπο λειτουργίας του και να προσαρμοστεί βάσει της λήψης νέων δεδομένων, θεωρείται πως διαθέτει την ικανότητα της μάθησης. Στην πιο απλή του μορφή, ένα τέτοιο σύστημα απομνημονεύει κάθε πρόβλημα που συναντά και ανακαλεί τη λύση του όποτε αυτή χρειάζεται. Ωστόσο, στις περισσότερες των περιπτώσεων η λειτουργία αυτή δεν επαρκεί: το πραγματικό ζητούμενο δεν είναι η λύση προβλημάτων που έχουν ήδη παρουσιαστεί στο παρελθόν, αλλά η λύση νέων, άγνωστων, προβλημάτων που ενδέχεται να εμφανιστούν στο μέλλον. Επομένως, η ικανότητα ενός συστήματος να *γενικεύσει* την πληροφορία που έχει λάβει, και να οδηγηθεί *επαγωγικά* σε μία λύση για κάποιο άγνωστο πρόβλημα, είναι το πραγματικό ζητούμενο της ΜΜ. Το Σχήμα 1.5 απεικονίζει τρία διαφορετικά μοντέλα τα οποία προσπαθούν να μάθουν τη συνάρτηση που ορίζει το σύνολο δεδομένων. Στην πρώτη εικόνα, το μοντέλο έχει αποτύχει να προσαρμοστεί στα δεδομένα, αφού δεν έχει καταφέρει να προσεγγίσει την καμπύλη που δημιουργούν. Αντιθέτως, στην δεύτερη εικόνα, το μοντέλο έχει καταφέρει μια καλή προσέγγιση, αφού κάθε σημείο είναι κοντά στα πραγματικά δεδομένα. Στην τρίτη εικόνα, το μοντέλο έχει προσαρμοστεί υπερβολικά στα διαθέσιμα δεδομένα (δεδομένα εκπαίδευσης), και ως εκ τούτου, είναι πολύ πιθανό να αποτύχει να προβλέψει αξιόπιστα μελλοντικές παρατηρήσεις (δείγματα).

### 1.2.2 Τρόποι Μάθησης

Ανάλογα με τον τρόπο μάθησης, η ΜΜ μπορεί να χωριστεί σε τρεις επιμέρους κατηγορίες:

- **Επιβλεπόμενη μάθηση:** Το σύστημα δέχεται τα δεδομένα ενός συνόλου παρόμοιων προβλημάτων, καθώς και τις τελικές λύσεις τους, με σκοπό την εύρεση γενικών κανόνων οι οποίοι να μπορούν να προβλέψουν την τελική λύση



Σχήμα 1.5: Γραφική αναπαράσταση αποτελέσματος μάθησης. Στην πρώτη εικόνα, το μοντέλο δεν έχει μάθει να προσαρμόζεται στα δεδομένα. Στη δεύτερη εικόνα, το μοντέλο προσεγγίζει την πραγματική λύση. Στην τρίτη εικόνα, το μοντέλο έχει προσαρμοστεί υπερβολικά στα διαθέσιμα δεδομένα, με κίνδυνο να μην μπορεί να γενικεύσει τη γνώση του σε νέα δείγματα.

ενός προβλήματος παρόμοιας φύσης. Επομένως, για ένα σύνολο δεδομένων εκπαίδευσης  $S = ((x_1, y_1), (x_2, y_2), \dots, (x_N, y_N))$  όπου  $x_n$  τα δεδομένα ενός προβλήματος  $n$ , και  $y_n$  η λύση του, ένας αλγόριθμος επιβλεπόμενης μάθησης μπορεί να εξάγει μια υπόθεση (ή ένα σύνολο κανόνων)  $h$ , σύμφωνα με την οποία για ένα νέο πρόβλημα  $x_i$  να βρει μία λύση  $\hat{y}_i$  τέτοια ώστε να είναι όσο το δυνατόν πιο κοντά στην πραγματική  $y_i$ , δηλαδή  $h(x_i) \approx y_i$ .

- Μη επιβλεπόμενη μάθηση: Χρησιμοποιείται όταν δεν υπάρχει κάποια συγκεκριμένη επιθυμητή λύση για ένα πρόβλημα. Αντίθετα με την επιβλεπόμενη μάθηση, στη μη επιβλεπόμενη μάθηση το σύστημα μαθαίνει μόνο του να αναλύει τα δεδομένα και να εξάγει συμπεράσματα, τα οποία ωστόσο μπορεί να μην είναι πάντα ξεκάθαρα, ή ακόμη και χρήσιμα. Συνήθως, το ζητούμενο είναι η ομαδοποίηση των ομοιογενών δεδομένων με στόχο την ανακάλυψη κρυφών συσχετίσεων στα δεδομένα, στην μείωση της πολυπλοκότητάς τους, κ.ά.
- Ενισχυτική μάθηση: Πρόκειται για μια ειδική περίπτωση ΜΜ, όπου ένας ευφυής πράκτορας αλληλεπιδρά με το περιβάλλον του και παρακολουθεί το αποτέλεσμα της αλληλεπίδρασης. Το περιβάλλον ανατροφοδοτεί τον πράκτορα με ένα θετικό ή αρνητικό σήμα, και ο πράκτορας προσαρμόζεται αναλόγως.

Στην ιατρική, η επιβλεπόμενη και μη επιβλεπόμενη μάθηση αποτελούν τις συνηθέστερες κατηγορίες ΜΜ.

### 1.2.3 Αλγόριθμοι Μάθησης

#### Κλασικές Τεχνικές

Οι « κλασικές » τεχνικές ΜΜ αναφέρονται σε θεμελιώδεις μεθόδους που συνήθως έχουν αναπτυχθεί και χρησιμοποιούνται εδώ και αρκετές δεκαετίες. Ωστόσο

η απλότητα, η ερμηνευσιμότητα, και η αποδεδειγμένη αποτελεσματικότητά τους, τις καθιστά ως μία πολύ καλή εναλλακτική πιο σύγχρονων μεθόδων, η οποία μάλιστα μπορεί να οδηγήσει και σε καλύτερα αποτελέσματα, ιδίως όταν τα δεδομένα εκπαίδευσης είναι περιορισμένα. Χαρακτηριστικό των κλασικών τεχνικών αποτελεί η εξάρτησή τους από τα χαρακτηριστικά τα οποία χρησιμοποιούνται ως είσοδοι στα μοντέλα, και τα οποία επιλέγει, μετασχηματίζει, ή κατασκευάζει ο ανθρώπινος παράγοντας από ακατέργαστα δεδομένα. Για παράδειγμα, στην περίπτωση ανάλυσης μίας εικόνας, δεν δίνεται η εικόνα αυτούσια ως είσοδος σε ένα κλασικό μοντέλο MM, αλλά πρώτα εξάγονται συγκεκριμένα χαρακτηριστικά όπως ποσοτικοποιήσεις της φωτεινότητας, της χρωματικής απόχρωσης, της υφής, κ.ά. Στη συνέχεια, το μοντέλο MM μαθαίνει να συσχετίζει τα χαρακτηριστικά αυτά με τέτοιο τρόπο, ώστε να παράγει το επιθυμητό αποτέλεσμα. Βασική πρόκληση αποτελεί η προσεκτική επιλογή των χαρακτηριστικών ούτως ώστε να επιτευχθούν τα βέλτιστα αποτελέσματα για ένα συγκεκριμένο πρόβλημα. Συνηθέστερες κλασικές τεχνικές MM αποτελούν οι Μηχανές Διανυσμάτων Υποστήριξης, αγγλ. Support Vector Machine (SVM), τα Τυχαία Δάση, αγγλ. Random Forests (RF), και οι k-κοντινότεροι γείτονες, αγγλ. k-Nearest Neighbors (k-NN).

**Μηχανές Διανυσμάτων Υποστήριξης** Οι Μηχανές Διανυσμάτων Υποστήριξης, συχνά αναφερόμενες απλά ως SVM, αποτελούν μία από τις εξέχοντες μεθόδους της κλασικής μηχανικής μάθησης. Η βασική λειτουργία τους στηρίζεται στην έννοια της γραμμικής διαχωρισιμότητας. Πιο συγκεκριμένα, μια SVM προσπαθεί να υπολογίσει ένα υπερεπίπεδο στον χώρο των χαρακτηριστικών, που να διαχωρίζει με τον βέλτιστο τρόπο τα πολυδιάστατα δείγματα ενός συνόλου δεδομένων σε  $n$  κατηγορίες [8]. Ένα υπερεπίπεδο είναι μια επίπεδη, πολυδιάστατη επιφάνεια που διαχωρίζει τα σημεία δεδομένων. Σε έναν δισδιάστατο χώρο, ένα υπερεπίπεδο είναι μια ευθεία γραμμή, στις τρεις διαστάσεις μια δισδιάστατη επιφάνεια, κ.ο.κ. Στόχος των SVM είναι η εύρεση του υπερεπιπέδου που μεγιστοποιεί το περιθώριο μεταξύ των διαφορετικών κατηγοριών, ούτως ώστε το μοντέλο να είναι πιο ανθεκτικό και γενικεύσιμο σε νέα δεδομένα. Ωστόσο, τα SVM δεν είναι κατάλληλα μόνο για γραμμικά δεδομένα. Αν και κάποιο σύνολο δεδομένων ενδεχομένως να μην είναι γραμμικά διαχωρίσιμο στον αρχικό χώρο χαρακτηριστικών του, θα μπορούσε να είναι γραμμικά διαχωρίσιμο σε έναν χώρο χαρακτηριστικών υψηλότερων διαστάσεων. Επομένως, με τη χρήση κατάλληλων συναρτήσεων, τα SVM μπορούν να χειριστούν και μη-γραμμικά δεδομένα, με την αντιστοίχισή τους σε έναν χώρο χαρακτηριστικών υψηλότερων διαστάσεων.

**Τυχαία Δάση** Τα τυχαία δάση αποτελούν μια ισχυρή, συνδυαστική μέθοδο μάθησης. Βάση της αποτελούν τα δέντρα αποφάσεων (αγγλ. decision trees), η λειτουργία των οποίων συνίσταται στην ανάλυση της διαδικασίας λήψης αποφάσεων σε μία ακολουθία δυαδικών επιλογών, δηλαδή σε ένα διάγραμμα ροής [9]. Βασικό πλεονέκτημα των δέντρων αποφάσεων αποτελεί η απλή και εύκολα ερμηνεύσιμη

προσέγγισή τους. Ωστόσο, ένα δέντρο αποφάσεων από μόνο του είναι επιρρεπές στην υπερπροσαρμογή στα δεδομένα, και τελικά στην αστάθεια. Επομένως, τα τυχαία δάση συνδυάζουν πολλαπλά δέντρα αποφάσεων, και αξιοποιούν τη συλλογική ευφυΐα τους για να βελτιώσουν την ακρίβεια. Αυτό που κάνει τα Τυχαία Δάση να ξεχωρίζουν είναι η ευρωστία, η επεκτασιμότητα και η εφαρμογή τους τόσο σε εργασίες ταξινόμησης όσο και σε εργασίες παλινδρόμησης.

**k-Κοντινότεροι Γείτονες** Ο αλγόριθμος k-Κοντινότερων γειτόνων (k-NN) είναι μια απλή αλλά αποτελεσματική μέθοδος μάθησης. Λειτουργεί με βάση την αρχή της εγγύτητας, ταξινομώντας τα σημεία δεδομένων σύμφωνα με την πλειοψηφική κλάση μεταξύ των k-κοντινότερων γειτόνων τους στο χώρο χαρακτηριστικών [10]. Βασίζεται, δηλαδή, στην ιδέα ότι παρόμοια σημεία δεδομένων (δείγματα) που ανήκουν στην ίδια κλάση, παρουσιάζουν παρόμοιες ιδιότητες, και άρα καταλαμβάνουν γειτονικές θέσεις στον χώρο χαρακτηριστικών. Ο k-NN είναι ιδιαίτερα χρήσιμος για εργασίες που περιλαμβάνουν ομαδοποίηση, ταξινόμηση και παλινδρόμηση.

### **Βαθιά Νευρωνικά Δίκτυα**

Σε αντίθεση με τις κλασικές μεθόδους ΜΜ, τα Βαθιά Νευρωνικά Δίκτυα (BNN), έχουν τη δυνατότητα να μαθαίνουν αυτόματα τα απαραίτητα χαρακτηριστικά από τα δεδομένα εισόδου, χωρίς να στηρίζονται στον ανθρώπινο παράγοντα. Χαρακτηριστικό τους αποτελεί η «βαθιά» αρχιτεκτονική τους, η οποία περιλαμβάνει πολλαπλά στρώματα διασυνδεδεμένων νευρώνων, δηλαδή μη-γραμμικών υπολογιστικών δομών. Ωστόσο, προϋπόθεση για την εκπαίδευση BNN είναι η ύπαρξη ενός επαρκώς μεγάλου συνόλου δεδομένων. Όσο βαθύτερο, και γενικότερα μεγαλύτερο, είναι ένα δίκτυο, τόσο μεγαλύτερος είναι ο όγκος δεδομένων που απαιτείται για την εκπαίδευσή του. Ενδεικτικά, τα σημερινά BNN που χρησιμοποιούνται για την ταξινόμηση εικόνων εκπαιδεύονται σε σύνολα δεδομένων όπως το ImageNet, το οποίο περιέχει εκατομμύρια εικόνες [11], [12]. Ιδιαίτερα στην ιατρική, όπου τα διαθέσιμα δεδομένα είναι συνήθως περιορισμένα, η εκπαίδευση των BNN αποτελεί μια σημαντική πρόκληση.

Τεχνικές όπως η μεταφορά μάθησης δύναται να αντιμετωπίσουν το πρόβλημα έλλειψης δεδομένων. Αυτή η μέθοδος περιλαμβάνει την προ-εκπαίδευση ενός BNN σε ένα μεγάλο, συνήθως δημόσια διαθέσιμο, σύνολο δεδομένων, και στη συνέχεια την επανεκπαίδευσή του, με χρήση μικρότερου ρυθμού μάθησης, σε ένα μικρότερο σύνολο δεδομένων. Με αυτό τον τρόπο αξιοποιείται η γνώση που αποκτήθηκε κατά την προ-εκπαίδευση και συχνά οδηγεί σε βελτιωμένη απόδοση σε περιορισμένα δεδομένα [13].

**Συνελικτικά Νευρωνικά Δίκτυα** Τα Συνελικτικά Νευρωνικά Δίκτυα (ΣΝΝ), αποτελούν μια από τις πιο καθοριστικές καινοτομίες στο πεδίο της βαθιάς μάθησης.

Το 2012, οι Krizhevsky et al. [12] δημοσίευσαν την αρχιτεκτονική ενός ΣΝΝ με την ονομασία AlexNet. Το δίκτυο αυτό έκανε χρήση πολλαπλών συνελκτικών επιπέδων, αρχιτεκτονική πολύ βαθύτερη από τα προηγούμενα νευρωνικά δίκτυα που είχαν χρησιμοποιηθεί για την ταξινόμηση εικόνων. Με αυτό τον τρόπο το AlexNet νίκησε τον διαγωνισμό του ImageNet ξεπερνώντας σημαντικά όλες τις υπόλοιπες συμμετοχές. Η επιτυχία του AlexNet στον διαγωνισμό αποτέλεσε σημείο καμπής στον τομέα της όρασης υπολογιστών και της βαθιάς μάθησης, αφού κατέδειξε την αποτελεσματικότητα των βαθιών ΣΝΝ στην ανάλυση και ταξινόμηση εικόνων, πυροδοτώντας έτσι ένα κύμα έρευνας και εξελίξεων στη βαθιά μάθηση. Έκτοτε, αρχιτεκτονικές ΣΝΝ όπως τα VGG [14], ResNet [15], και EfficientNet [16] έχουν ξεπεράσει την απόδοση του AlexNet, και χρησιμοποιούνται ευρέως σε ποικίλες εφαρμογές στον τομέα της όρασης υπολογιστών.

### **1.3 Η Μηχανική Μάθηση ως Εργαλείο στην Ιατρική**

Αν και το πεδίο της ΤΝ θεμελιώθηκε τη δεκαετία του 1950, με το Logic Theorist (1956) να αποτελεί το πρώτο πρόγραμμα αυτοματοποιημένου συλλογισμού [17], ήδη από το 1959 είχαν πραγματοποιηθεί προσπάθειες ένταξης της ΤΝ στην ιατρική μέσω της αυτοματοποιημένης αξιολόγησης συμπτωμάτων από τους Brodman et al. [18]. Συνεπώς, η αξία της ΤΝ στην ιατρική είχε γίνει αντιληπτή από πολύ νωρίς. Ωστόσο, οι περιορισμοί στην επεξεργαστική ισχύ και μνήμη των υπολογιστών αποτέλεσαν εμπόδιο στην εξέλιξη του πεδίου με συνέπεια οι εφαρμογές ΤΝ να μην βρουν εφαρμογή στην κλινική πρακτική κατά τις δεκαετίες του 1970 και 1980 [19].

Σταδιακά, από τη δεκαετία του 1990, η αύξηση της επεξεργαστικής ισχύος επέτρεψε την είσοδο της ΤΝ στην ιατρική. Η αυτοματοποιημένη αξιολόγηση του ηλεκτροκαρδιογραφήματος (ΗΚΓ), η μέτρηση των λευκών αιμοσφαιρίων σε δείγματα αίματος, η ανάλυση φωτογραφιών του αμφιβληστροειδούς και δερματικών παθήσεων, αποτελούν πλέον κομμάτι της κλινικής πρακτικής [19]. Ωστόσο, παρά τη χρησιμότητά τους, τα συστήματα αυτά δεν προορίζονται για την αντικατάσταση των ιατρών, αλλά για την υποβοήθησή τους. Καθώς η τεχνολογία εξελίσσεται, η κλινική πρακτική μεταβαίνει ολοένα και περισσότερο σε έναν συνδυασμό εμπειρογνωμοσύνης, λήψης αποφάσεων βάσει τεκμηρίων (αγγλ. evidence-based decision making), και εργαλείων ΤΝ [20].

Τα τελευταία έτη, οι τεχνολογικές εξελίξεις στην ΤΝ, και πιο συγκεκριμένα στην ΜΜ, έχουν αυξήσει κατακόρυφα το ερευνητικό ενδιαφέρον σε όλο το ευρύ πεδίο εφαρμογών τους, ιδιαίτερος στην ιατρική. Συνήθεις εφαρμογές αποτελούν η διάγνωση, η πρόγνωση, η εκλογή θεραπείας, αλλά και η βελτίωση της εργασιακής ροής των ιατρών. Η παρούσα διδακτορική διατριβή εστιάζει στην αυτοματοποιημένη διάγνωση και την πρόγνωση μέσω ανάλυσης απεικονιστικών εξετάσεων.

### 1.3.1 Διάγνωση

Η διάγνωση αποτελεί τον ακρογωνιαίο λίθο της ιατρικής πρακτικής, αποτελώντας τη διαδικασία με την οποία οι επαγγελματίες υγείας εντοπίζουν και καθορίζουν τη φύση της ασθένειας ή της κατάστασης ενός ασθενούς. Επομένως, η ακριβής και έγκαιρη διάγνωση είναι υψίστης σημασίας καθώς αποτελεί τη βάση για τις αποφάσεις θεραπειάς, την εκτίμηση της πρόγνωσης και, τελικά, την ευημερία των ασθενών.

Η διάγνωση βασίζεται σε πολύ μεγάλο βαθμό στις γνώσεις, την κλινική εμπειρία, και την κρίση των ιατρών, οι οποίοι αναλύουν σχολαστικά τα συμπτώματα, το ιατρικό και οικογενειακό ιστορικό, τα ευρήματα της κλινικής εξέτασης και των διαφόρων διαγνωστικών εξετάσεων. Ενώ ο ανθρώπινος παράγοντας παραμένει αναντικατάστατος, η ενσωμάτωση της TN και της MM στη διαγνωστική διαδικασία, μπορεί να επιφέρει ακριβέστερα και ταχύτερα αποτελέσματα. Συνεπώς, πολλές έρευνες έχουν επικεντρωθεί ανά τα έτη στην αυτοματοποιημένη διάγνωση παθολογιών με ιδιαίτερη έμφαση στην αξιολόγηση δεδομένων μη-επεμβατικών απεικονιστικών εξετάσεων, αλλά και ιστολογικών εικόνων. Παραδείγματα μη-επεμβατικών εξετάσεων αποτελούν η δερματοσκόπηση, η αξονική τομογραφία, αγγλ. Computed Tomography (CT), η μαγνητική τομογραφία, αγγλ. Magnetic Resonance Imaging (MRI), το υπερηχογράφημα, κ.ά.

Στη δερματοσκόπηση ιδιαίτερο ενδιαφέρον παρουσιάζει η διάγνωση του μελανώματος, μία από τις πιο θανατηφόρες μορφές καρκίνου του δέρματος, με την έγκαιρη ανίχνευσή του να αυξάνει σημαντικά τις πιθανότητες επιβίωσης του ασθενούς. Ωστόσο, το μελάνωμα μπορεί συχνά να μιμείται καλοήθεις δερματικές αλλοιώσεις, καθιστώντας δύσκολη τη διάκριση μεταξύ αβλαβών δερματικών ανωμαλιών και δυνητικά κακοήθους μελανώματος [21]. Συνεπώς, η μηχανική μάθηση μπορεί να διαδραματίσει καθοριστικό ρόλο στην έγκαιρη και ακριβή ανίχνευσή του. Προς αυτή την κατεύθυνση, οι Codella et al. (2017) [22], πρότειναν ένα σύστημα το οποίο συνδύαζε βαθιά ΣNN με κλασικές τεχνικές MM για την ανίχνευση και κατάτμηση του μελανώματος, πετυχαίνοντας ακρίβεια 76%. Στη συγκεκριμένη προσέγγιση, ένα ΣNN χρησιμοποιήθηκε αρχικά για την κατάτμηση της δερματικής βλάβης, και στη συνέχεια ένας συνδυασμός ΣNN και κλασικών τεχνικών MM, για την εξαγωγή χαρακτηριστικών από τη βλάβη και την ταξινόμησή τους. Το ίδιο έτος, οι Yu et al. [23] πρότειναν τη χρήση μιας αρχιτεκτονικής ΣNN δύο σταδίων, η οποία περιλάμβανε δύο βαθιά ΣNN στη σειρά, τα οποία εκτελούσαν διαδοχικά κατάτμηση της βλάβης, και στη συνέχεια τη διάγνωση, πετυχαίνοντας ακρίβεια 95%. Πιο πρόσφατα (2022), οι Nawaz et al. [24] πρότειναν και αυτοί μια πλήρως αυτοματοποιημένη μέθοδο για τη διάγνωση του μελανώματος, η οποία αξιοποιούσε ΣNN με βάση την περιοχή (RCNN), σε συνδυασμό με ασαφή ομαδοποίηση k-means (FKM). Η μέθοδος περιλάμβανε την προεπεξεργασία εικόνων για την ενίσχυση της οπτικής πληροφορίας, τη χρήση RCNN για την εξαγωγή χαρακτηριστικών, και τη χρήση FKM για την κατάτμηση, πετυχαίνοντας μέση ακρίβεια



95,4%, 93,1% και 95,6% σε τρία διαφορετικά σύνολα δεδομένων.

Η αξονική τομογραφία είναι μια από τις πιο σημαντικές και συχνά χρησιμοποιούμενες μη-επεμβατικές απεικονιστικές μεθόδους στον τομέα της ακτινοδιαγνωστικής. Τα τελευταία χρόνια, ο συνδυασμός της αξονικής τομογραφίας με προηγμένες τεχνικές TN και MM έχει επεκτείνει περαιτέρω τις δυνατότητές της, επιτρέποντας την αυτοματοποιημένη διάγνωση και την ακριβή ανάλυση ιατρικών παθήσεων. Ένας από τους βασικούς παράγοντες αυτής της σύγκλισης της τεχνολογίας και της ιατρικής απεικόνισης είναι η «ραδιομική» ανάλυση (radiomics). Η ραδιομική είναι ένας αναδυόμενος τομέας που εξάγει ένα εξαιρετικά ευρύ φάσμα ποσοτικών χαρακτηριστικών από ιατρικές εικόνες, συμπεριλαμβανομένων των αξονικών αλλά και μαγνητικών τομογραφιών. Τα χαρακτηριστικά αυτά περιλαμβάνουν όχι μόνο τα παραδοσιακά οπτικά χαρακτηριστικά, αλλά και περίπλοκα δεδομένα σχετικά με τις υφές, τα σχήματα και τις χωρικές σχέσεις εντός των εικόνων. Ιδιαίτερα στην έρευνα για τον καρκίνο, τα ραδιομικά χαρακτηριστικά αποκτούν ολοένα και μεγαλύτερη σημασία καθώς, σε συνδυασμό με τη μηχανική μάθηση, μπορούν αποδεδειγμένα να συμβάλλουν στην ανάπτυξη επιτυχημένων κλινικών συστημάτων υποστήριξης αποφάσεων [25]. Συνεπώς, μία κοινή τεχνική αποτελεί η εξαγωγή εκατοντάδων χαρακτηριστικών με τη χρήση ραδιομικής ανάλυσης, και στη συνέχεια, την ταξινόμηση της εξέτασης βάσει αυτών των χαρακτηριστικών με χρήση κλασικών τεχνικών MM. Για παράδειγμα, το 2019, οι Chu et al. [26] αξιολόγησαν τη χρησιμότητα των ραδιομικών χαρακτηριστικών στη διάκριση αδενοκαρκινώματος του παγκρέατος από το φυσιολογικό πάγκρεας, σε εικόνες αξονικής τομογραφίας. Για κάθε εικόνα, εξήχθησαν 478 ραδιομικά χαρακτηριστικά από το φυσιολογικό πάγκρεας και το αδενοκαρκίνωμα ξεχωριστά, 40 από τα οποία επιλέχθηκαν ως τα πιο σχετικά. Ένας ταξινομητής τυχαίου δάσους χρησιμοποιήθηκε για τη δυαδική ταξινόμηση, επιτυγχάνοντας συνολική ακρίβεια 99,2% στη διάκριση μεταξύ φυσιολογικού παγκρέατος και καρκίνου. Το 2020, οι Park et al. [27], χρησιμοποίησαν ραδιομική ανάλυση για τη διάκριση μεταξύ αυτοάνοσης παγκρεατίτιδας και αδενοκαρκινώματος. Οι ερευνητές ανέλυσαν αναδρομικά τις αξονικές τομογραφίες ασθενών, εξάγοντας ραδιομικά χαρακτηριστικά τα οποία ταξινόμησαν με τη χρήση ενός ταξινομητή τυχαίου δάσους. Τα αποτελέσματα έδειξαν ότι η MM μπορεί να επιτύχει ακρίβεια 95,2% στη διάκριση μεταξύ καρκίνου και παγκρεατίτιδας. Ωστόσο, όπως έχει προαναφερθεί, η χειροκίνητη εξαγωγή χαρακτηριστικών όπως με τη χρήση ραδιομικής ανάλυσης, δεν είναι μονόδρομος. Τεχνικές βαθιάς μάθησης μπορούν να εξάγουν και αυτές επιτυχώς, πλήρως αυτοματοποιημένα, τα δικά τους χαρακτηριστικά. Για παράδειγμα, στην έρευνα των Alves et al. (2022) [28], χρησιμοποιήθηκε το βαθύ ΣNN nnU-Net για την ανίχνευση και κατάτμηση του παγκρεατικού αδενοκαρκινώματος, του παγκρέατος, καθώς και των αγγείων, του παγκρεατικού πόρου, και του χοληδόχου πόρου. Η προτεινόμενη μέθοδος πέτυχε πολύ υψηλά αποτελέσματα, με το εμβαδό κάτω από την καμπύλη λειτουργικών χαρακτηριστικών, αγγλ. Area Under the Receiver Op-

erating Curve (AUC), να ανέρχεται σε 0,91.

Η ανάλυση ιστοπαθολογικών εικόνων επί του παρόντος απαιτεί την προσεκτική εξέταση τεράστιων, σε διαστάσεις, εικόνων, από ιατρούς ειδικευμένους στην « παθολογοανατομία ». Η χρήση της TN για την υποβοήθηση της διαδικασίας αυτής έχει απασχολήσει από νωρίς τον τομέα της βιοϊατρικής τεχνολογίας, με την πρώτη εφαρμογή αλγορίθμων ανάλυσης εικόνων κυττάρων να εμφανίζεται τη δεκαετία του 1960 [29]. Έκτοτε, ο τομέας της *υπολογιστικής ιστοπαθολογίας* (αγγλ. Computational Pathology – CPATH) συνεχίζει να εξελίσσεται με επιταχυνόμενους ρυθμούς, τόσο λόγω της βελτίωσης των μεθόδων λήψης WSI, όσο και της εξέλιξης της TN και MM. Το 2016, για πρώτη φορά, δημοσιοποιήθηκε ένα πολύ μεγάλο σύνολο δεδομένων (1.399) ιστοπαθολογικών εικόνων, με σκοπό τη δημιουργία του διαγωνισμού CAMELYON [30] για την αυτοματοποιημένη διάγνωση μεταστάσεων του καρκίνου του μαστού σε λεμφαδένες φρουρούς. Το συγκεκριμένο σύνολο δεδομένων αποτελεί ένα ορόσημο στον τομέα της υπολογιστικής ιστοπαθολογίας, αφού δημιούργησε την απαιτούμενη ώθηση για την ανάπτυξή της [29]. Βασική πρόκληση αποτελεί το πολύ μεγάλο μέγεθος των εικόνων προς ανάλυση, σε συνδυασμό με τα συχνά ελλιπή δεδομένα, οδηγώντας σε νέους τρόπους διαχείρισης των εικόνων. Για παράδειγμα, το 2020, οι Wang et al. [31] δημοσίευσαν μία προσέγγιση χρησιμοποιώντας ασθενώς επιβλεπόμενη MM για τη διάγνωση καρκίνου στους πνεύμονες. Πιο συγκεκριμένα, η μέθοδος ξεκινά με τη χρήση ενός ΣNN το οποίο αξιολογεί ξεχωριστά επιμέρους τμήματα της εικόνας WSI. Στη συνέχεια, ο αλγόριθμος επιλέγει τα τμήματα της εικόνας που θεωρεί πιο πιθανό να περιέχουν σημαντική πληροφορία, εξάγει από αυτά χαρακτηριστικά, και τα ταξινομεί με χρήση ενός ταξινομητή τυχαίου δάσους, πετυχαίνοντας ακρίβεια 97,3%. Αυτή η τεχνική διαίρεσης της αρχικής εικόνας σε επιμέρους κομμάτια αποτελεί μία πολύ δημοφιλή προσέγγιση στην υπολογιστική ιστοπαθολογία. Ένα ακόμη παράδειγμα αποτελεί η έρευνα των Kwak et al. (2021) [32], στην οποία διαίρεσαν τις εικόνες WSI σε επιμέρους τμήματα διαστάσεων 224×224 εικονοστοιχείων, ταξινομώντας κάθε ένα από αυτά μεταξύ επτά κατηγοριών: φυσιολογικό βλεννογόνο, στρώμα, λεμφοκύτταρα, βλέννα, λιπώδη ιστό, λείο μυ, και καρκίνο. Το μοντέλο MM στο οποίο βασίστηκε η μέθοδος ήταν ένα ΣNN τύπου U-Net, επιτυγχάνοντας σκορ dice 0,892. Πιο πρόσφατα, το 2023, οι Wagner et al. [33] διαφοροποιήθηκαν από τις υπόλοιπες έρευνες που χρησιμοποιούσαν ΣNN για την ανάλυση ιστοπαθολογικών εξετάσεων, χρησιμοποιώντας μετασχηματιστές, οι οποίοι, όπως προαναφέρθηκε, αναδείχθηκαν λόγω της χρήσης τους, με μεγάλη επιτυχία, στην επεξεργασία φυσικής γλώσσας. Στην έρευνα των Wagner et al., οι μετασχηματιστές χρησιμοποιούνται για την αξιολόγηση WSI και τη διάγνωση καρκίνου του παχέος εντέρου. Χρησιμοποιώντας ένα τεράστιο σύνολο δεδομένων 13,000 ασθενών, οι συγγραφείς πέτυχαν ευαισθησία 99% στη διάγνωση του καρκίνου, με αρνητική προγνωστική αξία επίσης 99%.

### 1.3.2 Πρόγνωση

Στο πλαίσιο της ιατρικής πρακτικής, η « πρόγνωση » αναφέρεται στην πρόβλεψη ή την εκτίμηση της πιθανής πορείας και έκβασης μιας νόσου σε έναν μεμονωμένο ασθενή. Περιλαμβάνει την εκτίμηση της μελλοντικής εξέλιξης της πάθησης, των πιθανών επιπλοκών και της πιθανότητας ανάρρωσης, επιδείνωσης, ή κατάληξης του ασθενή. Οι προγνωστικές πληροφορίες είναι ζωτικής σημασίας τόσο για τους ασθενείς όσο και για τους επαγγελματίες υγείας, καθώς καθοδηγούν τις θεραπευτικές αποφάσεις με σκοπό την καλύτερη διαχείριση της νόσου.

Προς αυτή την κατεύθυνση, οι Kakino et al. (2020) [34] χρησιμοποίησαν ραδιομικά χαρακτηριστικά σε συνδυασμό με κλινικές παραμέτρους, για την πρόβλεψη της τοπικής υποτροπής και της απομακρυσμένης μετάστασης σε ασθενείς με πρώιμο μη μικροκυτταρικό καρκίνο του πνεύμονα, μετά από στερεοτακτική ακτινοθεραπεία σώματος. Πιο συγκεκριμένα, οι συγγραφείς χρησιμοποίησαν τις αξονικές τομογραφίες των ασθενών, στις απεικονίσεις των οποίων οριοθετήθηκαν οι όγκοι από ακτινολόγους ιατρούς. Χρησιμοποιώντας τα ακριβή όρια των όγκων, οι συγγραφείς εξήγαγαν 944 ραδιομικά χαρακτηριστικά αποκλειστικά από αυτές τις περιοχές, και τα συνδύασαν με κλινικές παραμέτρους. Οι κλινικές αυτές παράμετροι περιελάμβαναν, μεταξύ άλλων, την ηλικία, το φύλλο, το δείκτη μάζας σώματος, ιστορικό καπνίσματος, και τα αποτελέσματα της ιστολογικής εξέτασης. Στη συνέχεια, έγινε χρήση του τυχαίου δάσους επιβίωσης (random survival forest), μιας μεθόδου MM η οποία εκτιμά την πιθανότητα ενός γεγονότος συναρτήσει του χρόνου. Οι συγγραφείς πέτυχαν Concordance Index (CI) (ελλ. δείκτης συμφωνίας) ίσο με 0,61 για την πρόβλεψη της τοπικής υποτροπής, και 0,68 για την απομακρυσμένη μετάσταση. Σε μια παρόμοια εφαρμογή, οι Xu et al. (2019) [35] ανέλυσαν αξονικές τομογραφίες ασθενών με μη μικροκυτταρικό καρκίνο του πνεύμονα, προβλέποντας την επιβίωση των ασθενών. Η διαφοροποίηση της συγκεκριμένης μελέτης έγκειται στο ότι οι συγγραφείς εκμεταλλεύτηκαν το γεγονός ότι η αξονική τομογραφία μπορεί εύκολα να επαναληφθεί με την πάροδο του χρόνου, και ανέπτυξαν ένα μοντέλο βασισμένο σε ΣNN που ενημέρωνε τις προβλέψεις επιβίωσης κατά τη διάρκεια της θεραπείας. Το 2021, οι Lu et al. [36] χρησιμοποίησαν μοντέλα βαθιάς μάθησης για την πρόβλεψη της μορφολογικής αλλαγής του όγκου στον μεταστατικό καρκίνο του παχέος εντέρου από αξονικές τομογραφίες, γεγονός που μπορεί να επιτρέψει έγκαιρες προσαρμογές κατά τη διάρκεια της θεραπείας. Πιο συγκεκριμένα, η μελέτη χρησιμοποίησε ΣNN για την εξαγωγή χαρακτηριστικών σε πολλαπλά χρονικά σημεία κατά τη διάρκεια της θεραπείας, και στη συνέχεια ένα επαναλαμβανόμενο νευρωνικό δίκτυο, αγγλ. Recurrent Neural Network (RNN), για το συνδυασμό τους και την εκτίμηση του τελικού αποτελέσματος. Τελικά, οι συγγραφείς πέτυχαν ένα CI ίσο με 0,694 στην πρόβλεψη της μορφολογικής μεταβολής του όγκου.

Όπως και με τη διάγνωση, η πρόγνωση μπορεί να πραγματοποιηθεί με χρήση ποικίλων απεικονιστικών μεθόδων. Τα προηγούμενα παραδείγματα έκαναν χρήση

ακτινοδιαγνωστικών εξετάσεων; σε αυτή την υποενότητα παρουσιάζονται παραδείγματα πρόγνωσης με χρήση MM σε ιστοπαθολογικές εικόνες. Το 2019, οι Courtiol et al. [37] ανέπτυξαν μια νέα προσέγγιση βασισμένη σε ΣNN για την πρόβλεψη της επιβίωσης ασθενών με μεσοθηλίωμα, αναλύοντας ιστολογικές WSI εικόνες. Η προσέγγισή τους βασίστηκε σε πολλαπλά στάδια: πρώτα εκπαιδευσαν ένα ΣNN U-Net για την κατάτμηση των ιστών, και στη συνέχεια, διαίρεσαν την WSI εικόνα σε επιμέρους τμήματα. Στη συνέχεια, ένα ΣNN ResNet, προεκπαιδευμένο στο σύνολο δεδομένων ImageNet χρησιμοποιήθηκε για την εξαγωγή χαρακτηριστικών από τμήματα που το περιεχόμενό τους εκτιμήθηκε πως απεικόνιζε ιστούς σε ποσοστό μεγαλύτερο του 20%. Τέλος, τα χαρακτηριστικά αυτά αξιολογήθηκαν από μια απλή αρχιτεκτονική νευρωνικού δικτύου για την εκτίμηση του χρόνου επιβίωσης, πετυχαίνοντας CI ίσο με 0,643. Σε μία άλλη έρευνα, από τους Bychkov et al. [38], χρησιμοποιήθηκε ένας συνδυασμός ΣNN και RNN για την πρόγνωση του καρκίνου του παχέος εντέρου, με χρήση ιστοπαθολογικών WSI εικόνων. Πιο συγκεκριμένα, το προεκπαιδευμένο ΣNN χρησιμοποιήθηκε για την εξαγωγή χαρακτηριστικών από επιμέρους τμήματα των WSI, και στη συνέχεια το RNN ανέλυε τα χαρακτηριστικά και εξήγαγε την πιθανότητα πενταετούς επιβίωσης του ασθενή. Το 2022, οι Esteva et al. [39] πρότειναν μια προσέγγιση η οποία συνδύαζε πληροφορίες τόσο από τις ιστοπαθολογικές εικόνες, όσο και από κλινικά δεδομένα για τη πρόβλεψη της πιθανότητας πενταετούς και δεκαετούς μετάστασης, καθώς και δεκαετούς επιβίωσης, σε ασθενείς με καρκίνο του προστάτη. Οι συγγραφείς πέτυχαν AUC ίσο με 0,84 για την πρόβλεψη πενταετούς μετάστασης, 0,78 για τη δεκαετή μετάσταση, και 0,65 για τη δεκαετή επιβίωση.

## 1.4 Αντικείμενο και Συνεισφορές της Παρούσας Διδακτορικής Διατριβής

Στην παρούσα διδακτορική διατριβή, αναπτύσσονται και αξιολογούνται μοντέλα MM για την αυτοματοποιημένη διάγνωση και πρόγνωση παθολογιών, σε τρεις διαφορετικές απεικονιστικές τεχνικές.

Στο **Κεφάλαιο 2** ερευνάται η αυτοματοποιημένη αξιολόγηση ιστοπαθολογικών εξετάσεων για την διάγνωση οστεοσαρκώματος, καθώς και για την εκτίμηση της νέκρωσης του όγκου έπειτα από νεοεπικουρική χημειοθεραπεία. Προς αυτή την κατεύθυνση πραγματοποιείται μία συγκριτική προσέγγιση διαφόρων τύπων βαθιών ΣNN. Η προσέγγιση αυτή αξιολογεί τόσο τις διαφορετικές αρχιτεκτονικές των δικτύων όσο και της επίδρασης του βάθους του δικτύου και του μεγέθους των εικόνων εισόδου, στην αποτελεσματικότητα της διάγνωσης.

Στο **Κεφάλαιο 3** παρουσιάζεται μια νέα προσέγγιση για την αυτοματοποιημένη διάγνωση αλλεργικής δερματίτιδας εξ επαφής με τη χρήση επιδερμικών δοκιμασιών επιθέματος (patch tests) και δερματοσκοπικής κάμερας. Στο πλαίσιο αυτό, σε συνεργασία με την Α' Πανεπιστημιακή Κλινική Δερματικών και Αφροδίσων

Νόσων του νοσοκομείου « Ανδρέας Συγγρός », συλλέχθηκε ένα σύνολο δεδομένων 200 ασθενών με την υποψία αλλεργικής δερματίτιδας εξ επαφής. Με τη χρήση αυτών των δειγμάτων, παρουσιάζεται και αξιολογείται μία προσέγγιση βαθιάς μάθησης για την ταυτοποίηση των αλλεργιογόνων στα οποία παρουσιάζει ευαισθησία ο κάθε ασθενής, χρησιμοποιώντας μια νέα, αλλά απλή, τεχνική προεπεξεργασίας των εικόνων η οποία αναδεικνύει την περιοχή ενδιαφέροντος, διατηρώντας ταυτόχρονα τη συνολική πληροφορία της εικόνας. Παράλληλα, διερευνάται η πρόσθετη αξία που παρέχουν διαφορετικοί τύποι εικόνων, συμπεριλαμβανομένων της απλής έγχρωμης εικόνας, του ερυθρήματος, της υψής, των λεπτών γραμμών, των πτυχώσεων, και των ανισοϋψώσεων της επιδερμίδας.

Στο **Κεφάλαιο 4**, παρουσιάζεται μία αυτοματοποιημένη προσέγγιση για την πρόγνωση ασθενών με αδενοκαρκίνωμα του παγκρέατος που πρόκειται να υποβληθούν σε παγκρεατεκτομή, μέσω της προεγχειρητικής αξονικής τομογραφίας άνω και κάτω κοιλίας. Για το σκοπό αυτό, συλλέχθηκε ένα σύνολο δεδομένων 40 ασθενών, σε συνεργασία με την Β' Χειρουργική Κλινική, και το Α' Εργαστήριο Ακτινολογίας του Αρεταίειου Νοσοκομείου. Χρησιμοποιώντας ένα βαθύ ΣΝΝ, εκπαιδευμένο σε ένα εξωτερικό, δημόσια διαθέσιμο, σύνολο δεδομένων, πραγματοποιείται πλήρως αυτοματοποιημένα ο εντοπισμός της θέσης του όγκου. Στη συνέχεια, διερευνάται η χρησιμότητα των ραδιομικών χαρακτηριστικών, σε συνδυασμό με κλινικά χαρακτηριστικά, για την πρόγνωση των ασθενών, σε βάθος διαιτίας και πενταετίας.

## Κεφάλαιο 2

# Διάγνωση και Ταξινόμηση Οστεοσαρκώματος σε Ιστοπαθολογικές Εξετάσεις

### 2.1 Οστεοσάρκωμα

Το οστεοσάρκωμα (ΟΣ), επίσης γνωστό ως οστεογενές σάρκωμα, είναι ένας καρκίνος των οστών ο οποίος αναπτύσσεται από κύτταρα μεσεγχυματικής προέλευσης που διαφοροποιούνται σε οστεοβλάστες [40], [41]. Είναι ο πιο διαδεδομένος πρωτοπαθής καρκίνος των οστών και αποτελεί ένα εξαιρετικά επιθετικό είδος κακοήθειας. Το ΟΣ αποτελεί τον συχνότερο πρωτοπαθή σκελετικό καρκίνο σε παιδιά και εφήβους. Αντιπροσωπεύει το 2% όλων των νεοπλασιών και αναπτύσσεται κυρίως μεταξύ των ηλικιών 10 και 14 ετών, κατά τη διάρκεια της εφηβικής αυξητικής περιόδου [42]. Παρά το γεγονός ότι πρόκειται για έναν σχετικά σπάνιο τύπο καρκίνου, μόνο στις ΗΠΑ περίπου 1000 παιδιά διαγιγνώσκονται κάθε χρόνο με ΟΣ.

#### 2.1.1 Παθογένεση

Στο ΟΣ, τα μεσεγχυματικά κύτταρα, ιδιαίτερα στις θέσεις σχηματισμού των οστών, παράγουν οστεοειδές και ανώριμο οστικό ιστό. Η ανάπτυξη οστικού ιστού μέσω πολλαπλασιασμού των οστεοβλαστών είναι πιθανό να εμφανίσει τόσο γονιδιωματικές όσο και επιγενετικές μεταβολές.

Λόγω του σημαντικού βαθμού ποικιλομορφίας του ΟΣ, η κατανόηση των μοριακών μηχανισμών πίσω από την παθογένειά του αποτελεί μία σημαντική πρόκληση. Διάφορες έρευνες έχουν εξετάσει τους μοριακούς παθογενετικούς μηχανισμούς του.

Τα γονίδια (ή οι πρωτεΐνες) *Rb* [43], *TP53* [44], *Grim-19* [45], [46], *P21/RAS* [47], [48] και *NF-κB* [49] έχουν προταθεί ως βασικοί παράγοντες του ΟΣ. Αρκετοί καρκίνοι έχουν ανώμαλη έκφραση του ογκοκατασταλτικού γονιδίου *Rb*. Το γονί-

διο αυτό είναι γνωστό για τη λειτουργία του στον έλεγχο της μετάβασης από τη φάση G1 στη φάση S του κυτταρικού κύκλου. Η πρωτεΐνη Rb, η οποία φωσφορυλιώνεται στην pRb, έχει την ικανότητα να αποτρέπει την υπερβολική ανάπτυξη των κυττάρων. Το γεγονός ότι οι πρωτεΐνες Rb και p53 (TP53) είναι και οι δύο ελαττωματικές, είναι ένα από τα βασικά ευρήματα σχετικά με το ΟΣ [50]. Η πρωτεΐνη TP53, είναι ένας «Μεταγραφικός παράγοντας που ρυθμίζει κρίσιμα γονίδια στην απόκριση σε βλάβες του DNA, στην εξέλιξη του κυτταρικού κύκλου και στα μονοπάτια της απόπτωσης» [44]. Σε κάθε μορφή όγκου, η TP53 λειτουργεί ως ογκοκατασταλτικός παράγοντας. Τα επίπεδα της TP53 είναι χαμηλά στα φυσιολογικά κύτταρα, ενώ είναι είτε μεταλλαγμένα είτε υπορυθμισμένα στις κακοήθειες [51]. Η πρωτεΐνη MDM2 είναι ο βασικός αρνητικός ρυθμιστής της TP53 και έχει την ικανότητα να προκαλεί την αποικοδόμηση του συμπλέγματος p53-ουμπικουϊνίνης [52]. Μεταλλάξεις της TP53 έχουν επίσης συσχετιστεί με το πρώιμο παιδικό ΟΣ. Το *GRIM-19* είναι ένα άλλο ογκοκατασταλτικό γονίδιο που αναγνωρίζεται για τη λειτουργία του στο ΟΣ. Η πρωταρχική του εργασία είναι να διευκολύνει την απόπτωση. Πρόσφατες έρευνες έδειξαν ότι το *GRIM-19* είναι υπορυθμισμένο στο ΟΣ, ενώ η επαγόμενη από την ακτινοβολία απόπτωση συσχετίζεται στενά με την υπερέκφραση της TP53 στο ΟΣ η οποία με της σειρά της συνεπάγεται αντίσταση στην επαγόμενη από την ακτινοβολία απόπτωση [45].

Ο μεταγραφικός παράγοντας NF-B είναι ένα ακόμη ενδιαφέρον σημαντικό συστατικό στη βιολογία του ΟΣ [53]. Η λειτουργία αυτού του μεταγραφικού παράγοντα στη φλεγμονή [54], στην ανάπτυξη του όγκου [55], στην αντίσταση στη χημειοθεραπεία (και στην ακτινοβολία) [56], και ειδικότερα για τη συμμετοχή του στην αντίσταση στην ακτινοβολία που σχετίζεται με το ΟΣ [49], [57], [58], τον καθιστούν έναν από τους μεταγραφικούς παράγοντες που έχουν λάβει τη μεγαλύτερη προσοχή. Πρόσφατες έρευνες έδειξαν ότι η αντίσταση των όγκων στην ακτινοβολία και τη χημειοθεραπεία σχετίζονται με την ενεργοποίηση του NF-B [57], [58]. Παρόλο που έχει αποκαλυφθεί ότι η NF-B δρα μέσω της οδού Akt/NF-B, ο μηχανισμός που διέπει την αντίσταση του ΟΣ στη χημειοθεραπεία και την ακτινοθεραπεία εξακολουθεί να είναι σε μεγάλο βαθμό άγνωστος [59], [60].

### **2.1.2 Διάγνωση**

Η ακριβής διάγνωση και ο προσδιορισμός του σταδίου της νόσου αποτελούν βασικές προϋποθέσεις για τον αποτελεσματικό σχεδιασμό της θεραπείας. Η ακτινοδιάγνωση αποτελεί ένα σύννηθες εργαλείο για τη διάγνωση του ΟΣ. Απεικονιστικές, μη επεμβατικές, εξετάσεις όπως η αξονική τομογραφία, η μαγνητική τομογραφία, η τομογραφία εκπομπής ποζιτρονίων, και το σπινθηρογράφημα οστών, όλες αποτελούν πιθανά διαγνωστικά εργαλεία για την διάγνωση του ΟΣ καθώς και για την αναγνώριση των μεταστάσεων [61]. Ωστόσο, η οριστική επιβεβαίωση του καρκίνου παρέχεται από την ιστοπαθολογική εξέταση δειγμάτων του ιστού.

### 2.1.3 Πρόγνωση

Γενικώς, το ΟΣ θεωρείται ως μια ιδιαίτερα θανατηφόρα ασθένεια. Οι μεταστάσεις αποτελούν το μεγαλύτερο εμπόδιο στην πρόγνωση του ασθενούς, παρά το γεγονός ότι οι σύγχρονες θεραπευτικές δυνατότητες έχουν αυξήσει σημαντικά την επιβίωση των ασθενών [40]. Τα τελευταία χρόνια, η πενταετής επιβίωση των ασθενών έχει αυξηθεί στο 60-65%, αν και η πρόγνωση παραμένει γενικότερα δυσμενής [40]. Δυστυχώς, η συνολική επιβίωση των ασθενών με μεταστατική νόσο είναι μόλις 27% [62], [63].

### 2.1.4 Θεραπεία

Στο παρελθόν, την πρώτη γραμμή θεραπείας αποτελούσαν οι ακρωτηριασμοί με στόχο την πλήρη αφαίρεση του όγκου. Ωστόσο, η πρόοδος στις τεχνικές απεικόνισης, καθώς και η προεγχειρητική (νεοεπικουρική) και μετεγχειρητική (επικουρική) χημειοθεραπεία (ΧΜΘ) διευκόλυναν τη μετάβαση στη χειρουργική επέμβαση διάσωσης των άκρων και αύξησαν το ποσοστό πενταετούς επιβίωσης από <20% στη δεκαετία του 1970 σε 65% το 2015 [64], [65]. Τα υψηλά ποσοστά νέκρωσης του όγκου (>90%) έπειτα από νεοεπικουρική ΧΜΘ συνδέονται με καλύτερη πρόγνωση και σημαντικά μεγαλύτερο ποσοστό επιβίωσης [65], [66]. Παρά την πρόσφατη πρόοδο των διαγνωστικών μεθόδων, τόσο η μοριακή όσο και η απεικονιστική ιστοπαθολογική αξιολόγηση παραμένουν ο σημερινός χρυσός κανόνας για την αξιολόγηση της θεραπείας.

## 2.2 Αντικείμενο της Μελέτης

Στην περίπτωση του ΟΣ, η ιστοπαθολογική εξέταση διευκολύνει τον προσδιορισμό του βαθμού εισβολής του όγκου, της διαφοροποίησης και της νέκρωσης. Συνεπώς, αποτελεί πολύτιμο εργαλείο τόσο για τη διάγνωση, όσο και για την αξιολόγηση της αποτελεσματικότητας της ΧΜΘ, και, τελικά, την πρόγνωση των ασθενών. Ωστόσο, αυτή η εξέταση είναι μια χρονοβόρα και επίπονη διαδικασία για τον κλινικό ιατρό που μπορεί να οδηγήσει σε λάθη και υποκειμενικά συμπεράσματα (observer bias) [67], [68]. Η διαδικασία αυτή περιπλέκεται περαιτέρω λόγω της σημαντικής ετερογένειας του ΟΣ. Ως εκ τούτου, η αυτοματοποίηση της ιστολογικής αξιολόγησης της νόσου θα μπορούσε να οδηγήσει σε εξετάσεις που είναι ταχύτερες, αποτελεσματικότερες και πιο οικονομικές [68].

Επί του παρόντος, οι πιο σύγχρονες μέθοδοι (State-of-the-Art) για την αξιολόγηση και ταξινόμηση ιστολογικών εικόνων χρησιμοποιούν τεχνικές Μηχανικής Μάθησης (MM). Στο παρελθόν, τεχνικές MM όπως οι Μηχανές Διανυσμάτων Υποστήριξης (SVM) και τα Τυχαία Δάση (RF) ήταν πολύ διαδεδομένες στην ταξινόμηση εικόνων [69]–[72]. Ωστόσο, η διαδικασία της εξαγωγής κατάλληλων χαρακτηριστικών αποτελεί ένα δύσκολο εγχείρημα. Παλαιότερα, οι πιο διαδεδομένες τεχνικές



χρησιμοποιούσαν ένα γνωστό σύνολο γενικών χαρακτηριστικών, όπως στατιστικά χαρακτηριστικά (μέση τιμή και τυπική απόκλιση φωτεινότητας, φασματικό περιεχόμενο εικόνας, κ.ά). Πρόσφατα, τεχνικές βαθιάς μάθησης οι οποίες μαθαίνουν να πραγματοποιούν μόνες τους την εξαγωγή χαρακτηριστικών, όπως τα Συνελκτικά Νευρωνικά Δίκτυα (ΣΝΝ) (αγγλ. Convolutional Neural Networks - CNN) και οι Μετασχηματιστές Ορασης (Με.Ο.) (αγγλ. Vision Transformers - ViT), έχουν δείξει εξαιρετικά αποτελέσματα που είναι εφάμιλλα των ανθρώπινων επιδόσεων σε πολλές εργασίες αναγνώρισης και ταξινόμησης εικόνων. Ομοίως, πρόσφατες μελέτες έχουν καταδείξει την αξία της βαθιάς μάθησης στην ιστοπαθολογία και τα πλεονεκτήματά της έναντι των παραδοσιακών τεχνικών μηχανικής μάθησης [73]–[75]. Ωστόσο, η αποτελεσματικότητά τους εξαρτάται σε μεγάλο βαθμό από τον όγκο και την ποικιλομορφία του συνόλου δεδομένων εκπαίδευσης, καθώς και από τον συνολικό αρχιτεκτονικό σχεδιασμό του δικτύου. Ως αποτέλεσμα, νέες τοπολογίες δικτύων εμφανίζονται διαρκώς.

Η εφαρμογή της ΜΜ στην ιστοπαθολογία έχει σχετικά σύντομη ιστορία, με τις πρώτες σχετικές δημοσιεύσεις να χρονολογούνται από τη δεκαετία του 1990. Αυτές οι πρώιμες έρευνες επικεντρώθηκαν κυρίως στην εφαρμογή βασικών τεχνικών ανάλυσης εικόνας, όπως η ανίχνευση ακμών και η κατωφλίωση, για τον εντοπισμό βασικών δομών εντός των εικόνων [76]. Με την εξέλιξη της τεχνολογίας, οι ερευνητές άρχισαν να διερευνούν όλο και πιο σύνθετες τεχνικές, όπως η ανάλυση υψής και η αναγνώριση προτύπων, για να βελτιώσουν την ταξινόμηση των Εικόνων Ολόκληρων Αντικειμενοφόρων Πλακών (αγγλ. Whole Slide Images - WSI). Μια WSI εικόνα είναι μία ψηφιακή φωτογραφία πολύ υψηλής ανάλυσης που απεικονίζει ολόκληρο το δείγμα ιστού (παρασκευάσμα), όπως αυτό έχει ληφθεί από έναν ψηφιακό σαρωτή. Οι εικόνες αυτές είναι τεράστιες σε μέγεθος, το οποίο ανέρχεται συνήθως σε αρκετά gigabytes, αφού αποτελούνται από εκατομμύρια εικονοστοιχεία. Επομένως, η ανάλυση μιας εικόνας αυτού του μεγέθους αποτελεί σημαντική πρόκληση, καθώς οι τρέχουσες δυνατότητες της τεχνολογίας δεν επαρκούν για την επεξεργασία δυνητικά χιλιάδων παρόμοιων εικόνων.

Στην παρούσα μελέτη, εκπαιδεύτηκαν σύγχρονα νευρωνικά δίκτυα για την ανίχνευση βιώσιμων και νεκρωτικών όγκων, καθώς και υγιούς ιστού, σε ιστολογικές εξετάσεις ασθενών με ΟΣ. Προς αυτή την κατεύθυνση, χρησιμοποιήθηκε ένα δημόσια διαθέσιμο σύνολο δεδομένων για την εκπαίδευση και σύγκριση διαφόρων ΣΝΝ, με σκοπό τον προσδιορισμό της βέλτιστης αρχιτεκτονικής, βάθους δικτύου, και μεγέθους εικόνας εισόδου.

## **2.3 Υλικό και Μέθοδοι**

Τα σύγχρονα ΣΝΝ έχουν δημιουργηθεί ως επί το πλείστον για να ταξινομήσουν εικόνες από βάσεις δεδομένων όπως το ImageNet, όπου η μέση ανάλυση των εικόνων είναι 469×387 εικονοστοιχεία [11], [13]. Προκειμένου να εξοικονομηθεί

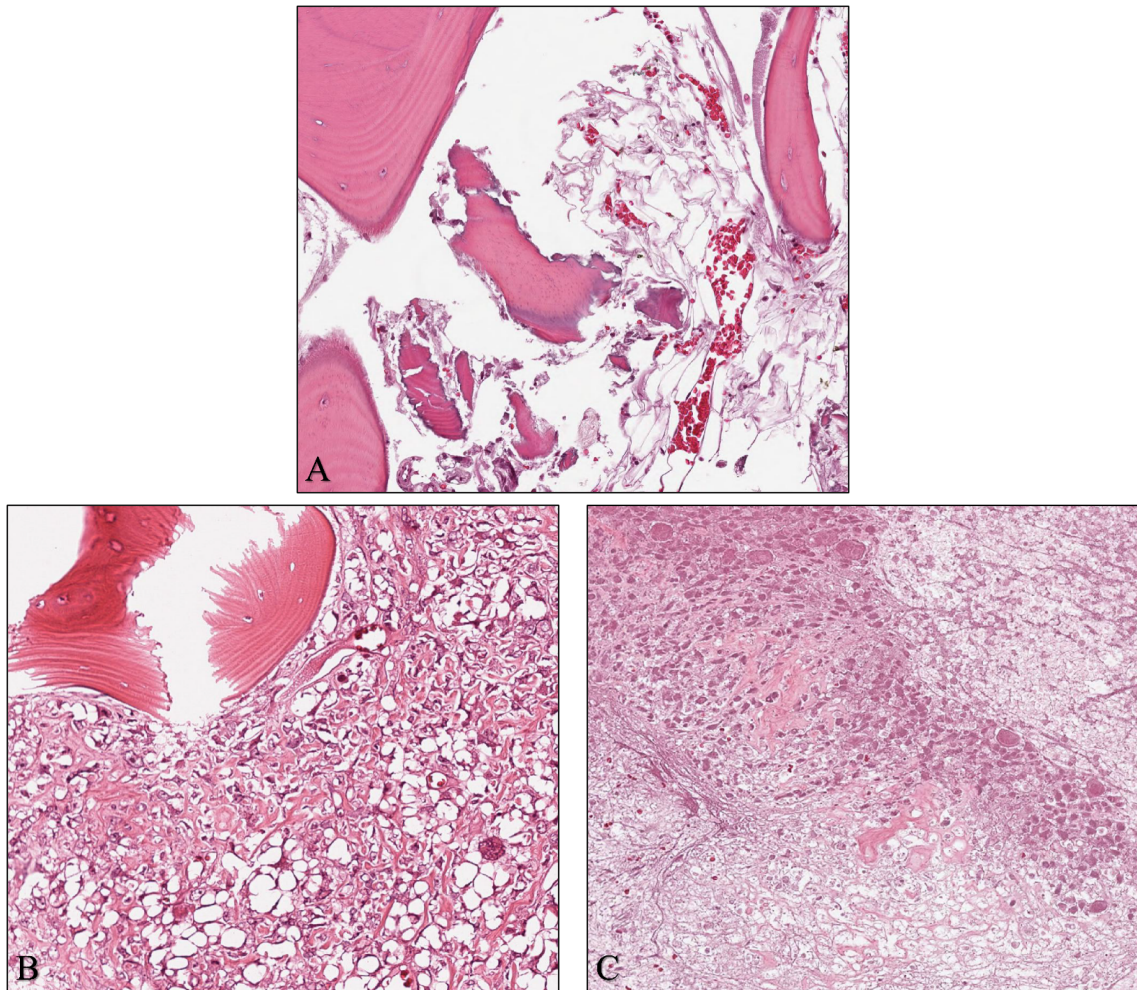
μνήμη και να αυξηθεί η υπολογιστική απόδοση, οι εικόνες αυτές συνήθως περικόπτονται ή συρρικνώνονται σε  $224 \times 224$  ή  $256 \times 256$  εικονοστοιχεία [12]. Ωστόσο, οι WSI εικόνες, οι οποίες μπορεί να αποτελούνται από έως και  $100.000 \times 100.000$  εικονοστοιχεία [77], έχουν σημαντικά μεγαλύτερες διαστάσεις. Προκειμένου να παρακαμφθούν οι σημερινοί τεχνολογικοί περιορισμοί, οι ερευνητές συχνά εξετάζουν τοπικές περιοχές που έχουν περικυβεί από τις WSI, με την κάθε περιοχή να ταξινομείται ξεχωριστά. Με την αύξηση της υπολογιστικής ισχύος και της μνήμης, καθίστανται δυνατά προοδευτικά μεγαλύτερα μεγέθη περιοχών προς ανάλυση. Θεωρητικά, όταν η ανάλυση πραγματοποιείται χρησιμοποιώντας περιοχές μεγαλύτερου μεγέθους, τα αποτελέσματα θα πρέπει να είναι ακριβέστερα καθώς τα δεδομένα ενσωματώνουν ένα πολύ μεγαλύτερο οπτικό πλαίσιο και περισσότερα σημεία δεδομένων (εικονοστοιχεία). Ωστόσο, στην παρούσα περίπτωση, η επεξεργασία μεγάλων εικόνων παρουσίασε αρκετές προκλήσεις λόγω περιορισμών στις δυνατότητες μνήμης και επεξεργαστικής ισχύος.

### **2.3.1 Σύνολο δεδομένων**

Το σύνολο δεδομένων που χρησιμοποιήθηκε στην παρούσα μελέτη ήταν τα, δημόσια προσβάσιμα, δεδομένα για το οστεοσάρκωμα από το UT Southwestern/UT Dallas για την αξιολόγηση βιώσιμων και νεκρωτικών όγκων. Οι ασθενείς που συμπεριλήφθηκαν στο σύνολο δεδομένων υποβλήθηκαν σε χειρουργική εκτομή στο Ιατρικό Κέντρο Παιδών του Dallas, μεταξύ του 1995 και 2015. Τα οστά που αφαιρέθηκαν κόπηκαν σε κομμάτια, αποσβεστώθηκαν, υποβλήθηκαν σε επεξεργασία με χρώση αιματοξυλίνης και ηωσίνης (H&E), και μεταφέρθηκαν σε αντικειμενοφόρες πλάκες. Οι πλάκες σαρώθηκαν σε ψηφιακές WSI εικόνες, 40 από τις οποίες επιλέχθηκαν από δύο παθολογοανατόμους με βάση την ετερογένεια του όγκου και τα χαρακτηριστικά ανταπόκρισης. Από κάθε μία εικόνα επιλέχθηκαν τυχαία 30 υποπεριοχές διαστάσεων  $1024 \times 1024$  εικονοστοιχείων, με αποτέλεσμα να προκύψουν 1200, συνολικά, εικόνες. Από αυτές τις εικόνες απορρίφθηκαν όσες δεν απεικόνιζαν ιστό, ή απεικόνιζαν αποκλειστικά μελάνι, ή ήταν θολές. Τελικά, 1144 εικόνες παρέμειναν για ανάλυση. Κάθε μία από τις εικόνες αυτές ταξινομήθηκε έπειτα από προσεκτική εξέταση από τους παθολογοανατόμους μεταξύ τριών κατηγοριών: μη όγκος (MO), βιώσιμος όγκος (BO), νέκρωση (NK). Η τελική κατανομή των εικόνων ήταν 536 (47%) MO, 263 (23%) NK, και 345 (30%) BO. Είναι σημαντικό να αναφερθεί πως 53 από τις 263 (20%) εικόνες NK απεικόνιζαν επίσης τμήματα BO. Στο Σχήμα 2.1 παρουσιάζεται ένα παράδειγμα των εικόνων που περιέχει το σύνολο δεδομένων.

### **2.3.2 Συνελικτικά Νευρωνικά Δίκτυα**

Τα ΣΝΝ αποτελούν μία ειδική κατηγορία των νευρωνικών δικτύων, καθώς το βασικό δομικό τους στοιχείο αποτελεί ο νευρώνας. Όπως και στη βιολογία, στα τεχνητά νευρωνικά δίκτυα οι νευρώνες συνδέονται μεταξύ τους μέσω των συνάψεων,



Σχήμα 2.1: Παράδειγμα εικόνων από το σύνολο δεδομένων για το οστεοσάρκωμα (μεγέθυνση  $\times 10$ ). (A) Μη όγκος, (B) βιώσιμος όγκος, (C) νέκρωση.

κάθε μία από τις οποίες χαρακτηρίζεται από μία τιμή βάρους  $w$ . Στις αρχιτεκτονικές πλήρως συνδεδεμένων νευρώνων, οι νευρώνες διατάσσονται σε επίπεδα. Όσο περισσότερα τα επίπεδα, τόσο βαθύτερο θεωρείται το νευρωνικό δίκτυο. Στα δίκτυα αυτά, ο κάθε νευρώνας υπολογίζει το σταθμισμένο άθροισμα όλων των συνάψεων, και στη συνέχεια περνά το αποτέλεσμα μέσα από μία μη-γραμμική συνάρτηση ενεργοποίησης  $\phi(x)$ . Η λειτουργία αυτή συνοψίζεται με την εξίσωση:

$$o_i^k = \phi \left( \sum_{j=1}^{r_{k-1}} w_{ji}^k o_j^{k-1} + b_i^k \right) \quad (2.1)$$

όπου  $o_j^{k-1}$  είναι το σήμα εξόδου του νευρώνα  $j$  στο επίπεδο  $k-1$ ,  $w_{ji}^k$  το βάρος της σύναψης που συνδέει τον νευρώνα  $j$  του επιπέδου  $k-1$  με τον νευρώνα  $i$  του επιπέδου  $k$ ,  $\phi$  η συνάρτηση ενεργοποίησης,  $r_{k-1}$  το πλήθος των νευρώνων του επιπέδου  $k-1$ , και  $b_i^k$  η τιμή πόλωσης του νευρώνα  $i$  στο επίπεδο  $k$ .

Ωστόσο, η συγκεκριμένη αρχιτεκτονική αντιμετωπίζει προκλήσεις όσον αφορά την εφαρμογή της σε εικόνες. Το σημαντικότερο μειονέκτημα αποτελεί το γεγονός ότι τα πλήρως συνδεδεμένα δίκτυα δεν επηρεάζονται από τον προσανατολισμό ενός διανύσματος εισόδου, δηλαδή, η σειρά των συνιστωσών του μπορεί να αναδιαταχθεί χωρίς να επηρεασθεί το τελικό αποτέλεσμα. Συμπερασματικά, εάν σε ένα πλήρως συνδεδεμένο νευρωνικό δίκτυο δίνονταν ως είσοδοι τα εικονοστοιχεία μίας εικόνας, η διάταξη των στοιχείων αυτών δε θα έπαιζε κανένα ρόλο! Το δεύτερο μειονέκτημα αφορά τις απαιτήσεις επεξεργαστικής ισχύος. Για παράδειγμα, μια μικρή, έγχρωμη εικόνα διαστάσεων  $256 \times 256 \times 3$  αποτελείται από 196608 εικονοστοιχεία. Ευθύς εξ αρχής, ένα επίπεδο με δύο, μόνο, νευρώνες, απαιτεί  $2 \cdot 196608 = 393216$ , δηλαδή 400.000, σχεδόν, βάρη. Γίνεται, λοιπόν, αμέσως αντιληπτό ότι οι απαιτήσεις σε μνήμη αυξάνονται σε υπερβολικά υψηλές τιμές καθώς το δίκτυο μεγαλώνει και βαθαίνει.

Τα παραπάνω προβλήματα μπορούν να λυθούν με τη χρήση ΣΝΝ, το όνομα των οποίων προκύπτει από το γεγονός ότι τα επίπεδα του δικτύου αποτελούνται από εκπαιδευσιμα, συνελκτικά, φίλτρα. Δοθείσας μιας εικόνας  $I$ , και ενός συνελκτικού φίλτρου (πυρήνα)  $K$ , εξάγεται μία νέα εικόνα χαρακτηριστικών  $O$  από τη συνέλιξη τους:

$$O(x, y) = f(I * W)(x, y) = f \left( \sum_{i=0}^{H-1} \sum_{j=0}^{W-1} I(x+i, y+j) \cdot K(i, j) + b \right) \quad (2.2)$$

όπου  $f$  η μη-γραμμική συνάρτηση (η οποία στην περίπτωση των ΣΝΝ ορίζεται συνήθως ως η συνάρτηση ReLU:  $\max(0, x)$ ),  $H$  και  $W$  το ύψος και πλάτος της εικόνας, και  $b$  η τιμή πόλωσης.

Καθώς το δίκτυο βαθαίνει, συχνά γίνεται χρήση συγκεντρωτικών επιπέδων (pooling layers), τα οποία χρησιμοποιούνται για τη μείωση των διαστάσεων των εικόνων χαρακτηριστικών. Αν και υπάρχουν πολλοί τύποι τέτοιων εικόνων, συνήθως χρησιμοποιούνται επίπεδα max pooling. Δοθέντος ενός παραθύρου μεγέθους  $p \times p$ , η

έξοδος  $O$  ορίζεται ως:

$$O(x, y) = \max_{i=0}^{p-1} \max_{j=0}^{p-1} I(x \cdot p + i, y \cdot p + j) \quad (2.3)$$

Αυτή η λειτουργία μειώνει τις χωρικές διαστάσεις της εικόνας χαρακτηριστικών, επιλέγοντας τη μέγιστη τιμή σε κάθε συγκεντρωτικό παράθυρο, γεγονός που επιτρέπει τη σύλληψη των πιο σημαντικών χαρακτηριστικών, μειώνοντας ταυτόχρονα την υπολογιστική πολυπλοκότητα.

Η εκπαίδευση ενός ΣΝΝ περιλαμβάνει τη βελτιστοποίηση των παραμέτρων του δικτύου, δηλαδή τα βάρη των συνελκτικών φίλτρων (τιμές των πυρήνων), και τις τιμές πόλωσης, για την ελαχιστοποίηση μιας συνάρτησης απώλειας (loss function). Στα νευρωνικά δίκτυα, η λειτουργία αυτή επιτελείται με τη χρήση του αλγορίθμου καθόδου κλίσης (gradient descent) και της οπισθοδιάδοσης (backpropagation). Συνοπτικά, η εκπαίδευση περιλαμβάνει τα εξής βήματα:

1. Το πρόσθιο πέρασμα, όπου το δίκτυο υπολογίζει τις προβλέψεις και την αντίστοιχη απώλεια. Η έξοδος  $\hat{y}$  υπολογίζεται ως εξής:

$$\hat{y} = N(x; w) \quad (2.4)$$

όπου  $N$  είναι το ΣΝΝ,  $x$  τα δεδομένα εισόδου, και  $w$  οι παράμετροι (βάρη και πόλωσης) του δικτύου. Η συνάρτηση απώλειας  $\mathcal{L}$  μετρά την απόκλιση μεταξύ της πραγματικής και της επιθυμητής εξόδου:

$$\mathcal{L} = \text{Loss}(\hat{y}, y) \quad (2.5)$$

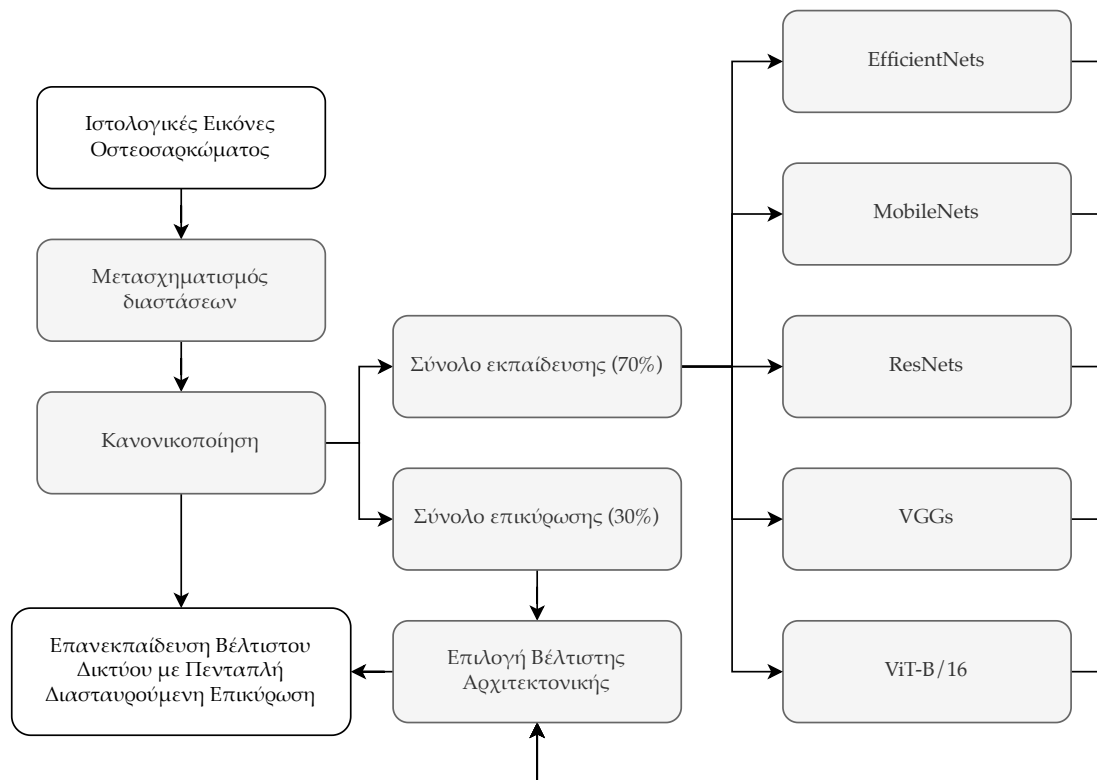
2. Την οπισθοδιάδοση, όπου πραγματοποιείται ο υπολογισμός των κλίσεων της απώλειας ως προς την απώλεια του δικτύου. Ο κανόνας της αλυσίδας εφαρμόζεται για τη διάδοση των κλίσεων προς τα πίσω μέσω του δικτύου. Επομένως, για μια παράμετρο  $\theta$  του δικτύου, η κλίση υπολογίζεται ως:

$$\frac{\partial \mathcal{L}}{\partial \theta} \quad (2.6)$$

3. Τον αλγόριθμο βελτιστοποίησης, όπου τα βάρη του δικτύου ενημερώνονται σύμφωνα με κάποιον κανόνα. Για παράδειγμα, ο αλγόριθμος της καθόδου κλίσης ορίζει την ενημέρωση των παραμέτρων ως:

$$\theta \leftarrow \theta - \eta \frac{\partial \mathcal{L}}{\partial \theta} \quad (2.7)$$

όπου  $\theta$  είναι η παράμετρος προς ενημέρωση, και  $\eta$  ο ρυθμός μάθησης (learning rate).



Σχήμα 2.2: Επισκόπηση της μεθοδολογίας για την συγκριτική μελέτη διαφόρων αρχιτεκτονικών νευρωνικών δικτύων στην ταξινόμηση του οστεοσαρκώματος. Οι ιστολογικές εικόνες μετασχηματίζονται στις επιθυμητές διαστάσεις και κανονικοποιούνται πριν τη χρήση τους για την εκπαίδευση των δικτύων. Στο διάγραμμα, με τον όρο EfficientNets, MobileNets κλπ, εννοείται όλη η οικογένεια των δικτύων αυτών, όπως φαίνεται στον πίνακα 2.6. Η βέλτιστη αρχιτεκτονική επιλέγεται και επανεκπαιδεύεται στο ίδιο σύνολο δεδομένων, με τη χρήση πενταπλής διασταυρούμενης επικύρωσης.

### 2.3.3 Πειραματική διάταξη

Όπως απεικονίζεται στο σχήμα 2.2, πραγματοποιήθηκε σύγκριση των επιδόσεων διαφόρων σύγχρονων αρχιτεκτονικών βαθιάς μάθησης στο σύνολο δεδομένων για το ΟΣ. Για την εκπαίδευση των μοντέλων έγινε χρήση «μεταφοράς μάθησης», μία τεχνική που περιλαμβάνει την προεκπαίδευση των μοντέλων σε ένα διαφορετικό, επαρκώς μεγάλο σύνολο δεδομένων (συνήθως το ImageNet), και στη συνέχεια την επανεκπαίδευση του μοντέλου στο σύνολο δεδομένων προς διερεύνηση. Αυτή η τεχνική επιτρέπει την αξιοποίηση της γνώσης που έλαβε το μοντέλο από την αρχική του εκπαίδευση, και την ευκολότερη προσαρμογή του στο νέο, αλλά σαφώς πολύ μικρότερο, σύνολο δεδομένων.

Όλοι οι υπολογισμοί πραγματοποιήθηκαν με τη χρήση του πακέτου PyTorch. Τα δίκτυα εκπαιδεύτηκαν σε μία κάρτα γραφικών NVIDIA Titan Xp με 12 GB μνήμη. Ο πηγαίος κώδικας για την εκπαίδευση και αξιολόγηση των μοντέλων είναι δημόσια διαθέσιμος στη σελίδα Zenodo (<https://doi.org/10.5281/zenodo.7765031>, τελευταία πρόσβαση στις 23 Μαρτίου 2023).

Οι αρχιτεκτονικές βαθιάς μάθησης που επιλέχθηκαν περιλάμβαναν το δίκτυο Visual Geometry Group (VGG) [14], το Residual Network (ResNet) [15], το MobileNetV2 [78], την οικογένεια δικτύων EfficientNet [16] και το Vision Transformer [79]. Οι αρχιτεκτονικές επιλέχθηκαν επειδή είναι καθιερωμένες με αποδεδειγμένη επιτυχία στην ταξινόμηση εικόνων, καθώς όλες έχουν επιτύχει στο παρελθόν κορυφαία αποτελέσματα στο ImageNet.

Ανάλογα με τις ανάγκες του εκάστοτε πειράματος, μία συγκεκριμένη αρχιτεκτονική νευρωνικού δικτύου μπορεί συνήθως να κλιμακωθεί προς τα πάνω ή προς τα κάτω όσον αφορά το μέγεθος (δηλαδή τον αριθμό των παραμέτρων). Τα μεγαλύτερα (και συνάμα βαθύτερα) δίκτυα είναι πιο ικανά για μάθηση, αλλά είναι επίσης πιο επιρρεπή στην υπερπροσαρμογή, ιδίως όταν εργάζονται με περιορισμένα σύνολα δεδομένων. Ως εκ τούτου, στην παρούσα έρευνα έγινε επιλογή μιας σειράς παραλλαγών κάθε αρχιτεκτονικής δικτύου, με μέγεθος που κυμαίνεται από πολύ μικρά (~2,2 εκατ. παράμετροι) έως μεγάλα (~86 εκατ. παράμετροι) μοντέλα. Στον πίνακα 2.1 απεικονίζεται ο αριθμός των παραμέτρων για κάθε επιλεγμένη παραλλαγή δικτύου. Λόγω των αλλαγών που έγιναν στα τελικά επίπεδά τους, ο αριθμός των παραμέτρων για κάθε επιλεγμένη εκδοχή δικτύου (ResNet [15], VGG [14], MobileNet [78] και EfficientNet [16]) μπορεί να διαφέρει από αυτούς που αναφέρονται στις αρχικές τους δημοσιεύσεις. Για παράδειγμα, ένα VGG16 διαθέτει κανονικά 138 εκατομ. παραμέτρους. Ωστόσο, στην παρούσα μελέτη, ακολουθήθηκε η εργασία των Anisuzzaman et al. (2021) [80], όπου αντικατέστησαν το τελευταίο πλήρως συνδεδεμένο επίπεδο μεγέθους 4096 νευρώνων (το οποίο αποτελεί έναν ταξινομητή που περιέχει ~120 εκατομ. παραμέτρους), με δύο πλήρως συνδεδεμένα επίπεδα μεγέθους 512 και 1024 νευρώνων, που περιέχουν μόνο ~13 εκατομ. παραμέτρους. Αυτή η τροποποίηση είχε ως αποτέλεσμα ένα πολύ μικρότερο δίκτυο χωρίς σημαντικές αλλαγές στην αρχιτεκτονική του. Η τροποποίηση του τελευταίου επιπέδου των δικτύων είναι απαραίτητη κατά τη μεταφορά μάθησης σε ένα διαφορετικό σύνολο δεδομένων. Στη συγκεκριμένη περίπτωση, ενώ τα δίκτυα είχαν αρχικά σχεδιαστεί για την ταξινόμηση εικόνων μεταξύ 1000 κατηγοριών (σύνολο δεδομένων ImageNet), ήταν επιθυμητή η ταξινόμηση μόνο μεταξύ τριών κατηγοριών (MO, BO, NK).

Για τη σύγκριση των επιδόσεων των διαφορετικών αρχιτεκτονικών βαθιάς μάθησης, το σύνολο δεδομένων χωρίστηκε σε ένα σύνολο εκπαίδευσης και ένα σύνολο επικύρωσης, με διαχωρισμό 70/30. Διασφαλίστηκε ότι κάθε δίκτυο εκπαιδεύτηκε και αξιολογήθηκε στο ίδιο σύνολο εικόνων, χρησιμοποιώντας κάθε φορά τον ίδιο διαχωρισμό, παρέχοντας έτσι μια δίκαιη σύγκριση. Για όλα τα πειράματα, χρησιμοποιήθηκε ο βελτιστοποιητής Adam [81] με αποσυνδεδεμένη μείωση βάρους [82] και ρυθμό μάθησης (LR) 0,0003. Προκειμένου να επιτευχθεί μια σταθερότερη ελάττωση του σφάλματος κατά τη διάρκεια της εκπαίδευσης, ο ρυθμός μάθησης μειωνόταν σταδιακά καθόλη τη διάρκεια της εκπαίδευσης από  $3 \times 10^{-4}$  σε  $1 \times 10^{-5}$ . Παρόλο που ο Adam μπορεί να μην είναι ο καλύτερος βελτιστοποιητής για όλες

<b>Μοντέλο</b>	<b>Αριθμός παραμέτρων</b>
EfficientNetB0	4.0 M
EfficientNetB1	6.5 M
EfficientNetB3	11 M
EfficientNetB5	28 M
EfficientNetB7	64 M
MobileNetV2	2.2 M
ResNet18	11 M
ResNet34	21 M
ResNet50	24 M
VGG16	28 M
VGG19	33 M
ViT-B/16	86 M

Πίνακας 2.1: Ο αριθμός των παραμέτρων για κάθε παραλλαγή δικτύου.

τις εργασίες, είναι μια ευρέως χρησιμοποιούμενη προεπιλεγμένη επιλογή στη βιβλιογραφία [2], [80], [81], [83]. Χρησιμοποιήθηκε επειδή ο κύριος στόχος ήταν η σύγκριση της απόδοσης διαφορετικών αρχιτεκτονικών και όχι η βελτιστοποίηση των υπερπαραμέτρων. Το επιλεγμένο μέγεθος παρτίδας (batch size) ποίκιλλε ανάλογα με την αρχιτεκτονική του δικτύου και το μέγεθος των εικόνων εισόδου. Για το μέγιστο μέγεθος εικόνας  $1024 \times 1024$  εικονοστοιχείων, το μέγεθος παρτίδας ορίστηκε σε 2, καθώς για τα περισσότερα δίκτυα η μνήμη της κάρτας γραφικών δεν επαρκούσε κατά την επεξεργασία μεγαλύτερων παρτίδων δεδομένων. Για τα δίκτυα στα οποία ακόμη και οι παρτίδες των 2 εικόνων δεν χωρούσαν στη μνήμη της κάρτας γραφικών, εκπαιδεύτηκαν χρησιμοποιώντας ένα μικρότερο μέγεθος εικόνας  $896 \times 896$  εικονοστοιχείων. Μια αξιοσημείωτη εξαίρεση ήταν το EfficientNetB7, το οποίο ήταν πολύ μεγάλο για να χωρέσει στη μνήμη ακόμη και με το μειωμένο μέγεθος εικόνας.

Εκτός από την εκπαίδευση στις εικόνες πλήρους ανάλυσης, τα δίκτυα εκπαιδεύτηκαν επίσης σε μικρότερες εκδοχές των εικόνων, με ανάλυση  $512 \times 512$  και  $256 \times 256$  εικονοστοιχεία. Οι διαστάσεις των εικόνων μειώθηκαν με τη μέθοδο της υποδειγματοληψίας χρησιμοποιώντας διγραμμική παρεμβολή. Ταυτόχρονα, το μέγεθος παρτίδας ορίστηκε ως αντιστρόφως ανάλογο του μεγέθους των εικόνων, διπλασιάζοντάς το για κάθε υποδιπλασιασμό των διαστάσεων. Η υποδειγματοληψία πραγματοποιείται συχνά όταν ο στόχος είναι η ταξινόμηση εικόνων με σκοπό τη μείωση του υπολογιστικού κόστους και των απαιτήσεων μνήμης, επιτυγχάνοντας εξαιρετικά αποτελέσματα. Παρόλο που η υποδειγματοληψία είναι μια καταστροφική διαδικασία που αφαιρεί πληροφορία από την εικόνα, στην πράξη αυτό δεν αποτελεί σημαντικό πρόβλημα στην ταξινόμηση εικόνων, καθώς το δίκτυο ενδιαφέρεται μόνο για την γενική κατηγορία της εκάστοτε εικόνας και όχι για την ακριβή θέση ενός αντικειμένου. Επιπλέον, η υποδειγματοληψία μπορεί να αποδειχθεί ακόμη και επωφελής, καθώς διευρύνει αποτελεσματικά το δεκτικό πεδίο των συνελκτικών στρωμάτων του ΣΝΝ [84], επιτρέποντας στο δίκτυο να μά-



θει περισσότερα γενικά, και όχι τοπικά, χαρακτηριστικά. Ωστόσο, υπάρχει ένας συμβιβασμός μεταξύ των πλεονεκτημάτων του διευρυμένου δεκτικού πεδίου και της απώλειας πληροφοριών λόγω της υποδειγματοληψίας. Ως εκ τούτου, δεν είναι πάντα σαφές ποιά διάσταση εικόνας εισόδου είναι η βέλτιστη.

Για την προεπεξεργασία των εικόνων, εκτός από την αλλαγή των διαστάσεων, πραγματοποιήθηκε ανεξάρτητα η κανονικοποίηση κάθε καναλιού RGB της εικόνας εισόδου, αφαιρώντας τη μέση τιμή και διαιρώντας με την τυπική απόκλιση. Οι μέσες τιμές που χρησιμοποιήθηκαν ήταν 0,485, 0,456 και 0,406 και οι τιμές τυπικής απόκλισης ήταν 0,229, 0,224 και 0,225 για τα κανάλια R, G και B, αντίστοιχα. Επιπλέον, κατά την εκπαίδευση εφαρμόστηκαν τεχνικές επαύξησης δεδομένων για να αυξηθεί η ποικιλομορφία του συνόλου εκπαίδευσης. Συγκεκριμένα, σε κάθε εποχή εκπαίδευσης οι εικόνες αναστρέφονταν τυχαία οριζόντια και κάθετα, και στη συνέχεια περιστρέφονταν έως και 20 μοίρες γύρω από το κέντρο τους.

### 2.3.4 Αξιολόγηση των δικτύων

Οι ακόλουθες αρχιτεκτονικές δικτύων εκπαιδεύτηκαν στο σύνολο δεδομένων του ΟΣ: EfficientNetB0, EfficientNetB1, EfficientNetB3, EfficientNetB5, EfficientNetB7, MobileNetV2, ResNet18, ResNet34, ResNet50, VGG16, VGG19 και ViT-B/16. Η αρχιτεκτονική του ViT έχει σχεδιαστεί έτσι ώστε να αντιμετωπίζει την εικόνα εισόδου ως μια ακολουθία 14 επιμέρους εικόνων με μέγεθος  $16 \times 16$  εικονοστοιχεία, περιορίζοντας έτσι το μέγεθος της εικόνας σε ακριβώς  $224 \times 224$ . Ως αποτέλεσμα, δεν ήταν δυνατή η εκπαίδευση με χρήση μεταφοράς μάθησης για αυτό το μοντέλο. Η απόδοση κάθε εκπαιδευμένου μοντέλου αξιολογήθηκε υπολογίζοντας τη βαθμολογία  $F1$ , ακολουθώντας τη στρατηγική πολλαπλών κατηγοριών One-vs-Rest (OnR). Πιο συγκεκριμένα:

$$F1 = 2 \frac{\text{αξιοπιστία} \cdot \text{ευαισθησία}}{\text{αξιοπιστία} + \text{ευαισθησία}} \quad (2.8)$$

$$\text{αξιοπιστία} = \frac{A_{\Theta}}{A_{\Theta} + \Psi_{\Theta}} \quad (2.9)$$

$$\text{ευαισθησία} = 2 \frac{A_{\Theta}}{A_{\Theta} + \Psi_A} \quad (2.10)$$

όπου  $A_{\Theta}$ ,  $A_A$ ,  $\Psi_{\Theta}$  και  $\Psi_A$  είναι οι αληθώς θετικές, οι αληθώς αρνητικές, οι ψευδώς θετικές και οι ψευδώς αρνητικές προβλέψεις, αντίστοιχα. Σύμφωνα με τη στρατηγική OnR, κάθε κλάση αντιμετωπίζεται ως ένα δυαδικό πρόβλημα ταξινόμησης, με τη θετική κλάση να είναι η κλάση ενδιαφέροντος και την αρνητική κλάση να περιλαμβάνει όλες τις άλλες κλάσεις. Ως εκ τούτου, οι όροι  $A_{\Theta}$  και  $\Psi_{\Theta}$  αναφέρονται στον αριθμό των εικόνων που ταξινομήθηκαν ως ανήκουσες στην κλάση ενδιαφέροντος, με  $A_{\Theta}$  τον αριθμό των σωστών ταξινομήσεων και  $\Psi_{\Theta}$  τον αριθμό των εσφαλμένων ταξινομήσεων. Ομοίως, οι όροι  $A_A$ ,  $\Psi_A$  αναφέρονται στον αριθμό

των ορθών και εσφαλμένων ταξινομήσεων των εικόνων που ταξινομήθηκαν ως μη ανήκουσες στην κλάση ενδιαφέροντος.

### 2.3.5 Επακόλουθο πείραμα

Κατόπιν της αξιολόγησης των δικτύων και της ανάλυσης των αποτελεσμάτων, επιλέχθηκε ο φερόμενος ως καλύτερος συνδυασμός δικτύου και μεγέθους εισόδου εικόνας, ο οποίος και επανεκπαιδεύτηκε χρησιμοποιώντας πενταπλή διασταυρούμενη επικύρωση (αγγλ. 5-fold cross validation). Με αυτόν τον τρόπο είναι δυνατός ο υπολογισμός μιας ακριβέστερης εκτίμησης της απόδοσης του μοντέλου, η οποία δεν θα είναι συνδεδεμένη με ένα συγκεκριμένο διαχωρισμό των δεδομένων μεταξύ εκπαίδευσης και επικύρωσης. Πιο συγκεκριμένα, για τον εκ νέου υπολογισμό της απόδοσης, το σύνολο δεδομένων χωρίστηκε σε πέντε διαμερίσεις (folds), εκ των οποίων οι τέσσερις χρησιμοποιούνται για την εκπαίδευση του δικτύου, και η μία εναπομένουσα για την επικύρωσή του. Η διαδικασία αυτή επαναλήφθηκε πεντάκις, με κάθε μία εκ των πέντε διαμερίσεων να χρησιμοποιήθηκε μία φορά ως σύνολο επικύρωσης. Στη συνέχεια, η τελική απόδοση υπολογίστηκε ως ο μέσος όρος των τιμών που προέκυψαν από κάθε επικύρωση. Οι υπερπαραμέτροι της διαδικασίας εκπαίδευσης παρέμειναν οι ίδιες όπως και στο προηγούμενο πείραμα, δηλαδή χρησιμοποιήθηκε ο ίδιος αριθμός εποχών, ο ίδιος βελτιστοποιητής, ο ρυθμός μάθησης και η στρατηγική επαύξησης δεδομένων. Επιπλέον, σε αυτό το πείραμα υπολογίστηκαν πρόσθετα μέτρα για την επίδοση του μοντέλου, προκειμένου να πραγματοποιηθεί εκτενέστερη διερεύνηση της επίδοσης του δικτύου, και να ερμηνευθεί η χρησιμότητά του σε ένα κλινικό περιβάλλον. Για τον σκοπό αυτό, υπολογίστηκε ο μέσος όρος και η τυπική απόκλιση του  $F1$  σκορ, της ακρίβειας (accuracy), της ειδικότητας (specificity), της ευαισθησίας (recall) και της αξιοπιστίας (precision) για κάθε ένα από τα πέντε μέρη, καθώς και ο συνδυασμός των πινάκων σύγχυσης μέσω άθροισης, και η καμπύλη ROC (Receiver Operating Characteristic). Συνεχίζοντας από τις προηγούμενες εξισώσεις, η αξιοπιστία (εξίσωση 2.9) και η ειδικότητα (εξίσωση 2.12) του ταξινομητή υπολογίστηκαν ως εξής:

$$\text{ακρίβεια} = \frac{A\Theta + AA}{A\Theta + \Psi\Theta + AA + \Psi A} \quad (2.11)$$

$$\text{ειδικότητα} = \frac{AA}{AA + \Psi\Theta} \quad (2.12)$$

Ο πίνακας σύγχυσης είναι ένας πίνακας που συνοψίζει την απόδοση ενός ταξινομητή συγκρίνοντας τις προβλέψεις του με τις πραγματικές κλάσεις ενός συνόλου δεδομένων επικύρωσης. Παρέχει μια ανάλυση του αριθμού των  $A\Theta$ ,  $AA$ ,  $\Psi\Theta$  και  $\Psi A$  για κάθε κλάση, επιτρέποντας μια πιο λεπτομερή αξιολόγηση της απόδοσης ενός μοντέλου. Η καμπύλη ROC (Receiver Operating Characteristic) είναι μια γραφική αναπαράσταση της απόδοσης ενός δυαδικού ταξινομητή κα-

θώς μεταβάλλεται το κατώφλι διάκρισης. Παρουσιάζει το ποσοστό αληθώς θετικών αποτελεσμάτων (ευαισθησία) έναντι του ποσοστού ψευδώς θετικών αποτελεσμάτων για διαφορετικές τιμές κατωφλίου. Η περιοχή κάτω από την καμπύλη ROC (AUC) είναι μια ευρέως χρησιμοποιούμενη μετρική για την αξιολόγηση της συνολικής απόδοσης του ταξινομητή, με υψηλότερη AUC να υποδηλώνει καλύτερη απόδοση.

## 2.4 Αποτελέσματα

### 2.4.1 Σύγκριση δικτύων

Η απόδοση των διαφόρων αρχιτεκτονικών πραγματοποιήθηκε χρησιμοποιώντας την βαθμολογία  $F1$ , η οποία είναι ένα ισχυρό μέτρο της απόδοσης ταξινόμησης που λαμβάνει υπόψη τόσο την αξιοπιστία όσο και την ευαισθησία. Η αξιοπιστία μετρά το ποσοστό των αληθώς θετικών προβλέψεων επί όλων των θετικών προβλέψεων που πραγματοποιεί το μοντέλο (εξίσωση 2.9), ενώ η ευαισθησία μετρά το ποσοστό των αληθώς θετικών προβλέψεων επί όλων των πραγματικών θετικών δειγμάτων (εξίσωση 2.10). Η βαθμολογία  $F1$  υπολογίζεται ως ο αρμονικός μέσος όρος των δύο τιμών (εξίσωση 2.8), απαιτώντας και οι δύο να συνεισφέρουν ομοιόμορφα προκειμένου να επιτευχθεί υψηλή τιμή βαθμολογίας. Ο Πίνακας 2.2 παρουσιάζει τις βαθμολογίες των δικτύων, παρέχοντας μια συγκριτική ανάλυση της απόδοσής τους σε καθεμία από τις τρεις κατηγορίες (MO, BO και NK).

Τα αποτελέσματα οδηγούν στη διαπίστωση ότι οι βαθύτερες, και μεγαλύτερες σε αριθμό παραμέτρων, αρχιτεκτονικές νευρωνικών δικτύων δεν παρουσιάζουν απαραίτητα καλύτερες επιδόσεις από τις ρηχότερες (Πίνακας 2.2). Συγκεκριμένα, το MobileNetV2, το οποίο είναι μακράν το μικρότερο δίκτυο που χρησιμοποιήθηκε στα πειράματα, απέδωσε την υψηλότερη μέση βαθμολογία  $F1$ . Ομοίως, το EfficientNetB0 είχε την ίδια ή καλύτερη απόδοση από τις μεγαλύτερες παραλλαγές του B1, B3, B5 και B7. Αντίστοιχα, το ResNet34 ξεπέρασε το μεγαλύτερο αντίστοιχό του, το ResNet50. Αυτή η συμπεριφορά οφείλεται στην υπερπροσαρμογή που παρατηρήθηκε με σχεδόν κάθε αρχιτεκτονική, όπου η ακρίβεια στο σύνολο εκπαίδευσης πλησίαζε το 100%. Αξιοσημείωτες εξαιρέσεις αποτελούν τα δίκτυα VGG, η εκπαίδευση των οποίων απέτυχε σε κάθε υποπαραλλαγή, εκτός από το VGG16, το οποίο πέτυχε καλά αποτελέσματα μόνο με εικόνες διαστάσεων  $256 \times 256$ . Αυτό οφείλεται στο γεγονός ότι ο αρχικός ρυθμός μάθησης  $3 \times 10^{-4}$  για τον βελτιστοποιητή Adam ήταν πολύ μεγάλος για το δίκτυο, προκαλώντας μεγάλες ενημερώσεις στα βάρη των δικτύων και συγκλίνοντας σε μια μη βέλτιστη λύση όπου κάθε δείγμα ταξινομήθηκε ως MO. Τα μεταγενέστερα πειράματα με ρυθμό μάθησης  $1 \times 10^{-5}$  έδωσαν αποτελέσματα συγκρίσιμα με δίκτυα παρόμοιου μεγέθους. Τα αποτελέσματα αυτά δεν συμπεριλήφθηκαν στην παρούσα μελέτη επειδή (α) δεν επηρέασαν τα τελικά συμπεράσματα και (β) δεν πραγματοποιήθηκε βελτιστοποίηση υπερπαραμέτρων για τα υπόλοιπα δίκτυα.

Πίνακας 2.2: Βαθμολογία  $F1$  για κάθε συνδυασμό αρχιτεκτονικής δικτύου και διαστάσεων των εικόνων εισόδου. Οι κατηγορίες MO, BO, και NK αναφέρονται σε μη όγκο, βιώσιμο όγκο, και νέκρωση, αντίστοιχα.

Δίκτυο	Διαστ. Εικόν.	Βαθμολογία $F1$			
		MO	BO	NK	Μέση Τιμή
EfficientNetB0	1024×1024	0.93	0.89	0.84	0.887
EfficientNetB0	512×512	0.93	0.88	0.83	0.880
EfficientNetB0	256×256	0.95	0.87	0.85	0.890
EfficientNetB1	1024×1024	0.93	0.85	0.82	0.867
EfficientNetB1	512×512	0.94	0.88	0.82	0.880
EfficientNetB1	256×256	0.95	0.86	0.84	0.883
EfficientNetB3	1024×1024	0.94	0.89	0.84	0.890
EfficientNetB3	512×512	0.93	0.86	0.81	0.867
EfficientNetB3	256×256	0.93	0.87	0.81	0.870
EfficientNetB5	896×896	0.92	0.89	0.81	0.873
EfficientNetB5	512×512	0.93	0.87	0.82	0.873
EfficientNetB5	256×256	0.94	0.84	0.80	0.860
EfficientNetB7	512×512	0.94	0.88	0.84	0.887
EfficientNetB7	256×256	0.95	0.87	0.83	0.883
MobileNetV2	1024×1024	0.82	0.84	0.66	0.773
MobileNetV2	512×512	0.92	0.85	0.81	0.860
MobileNetV2	256×256	0.94	0.89	0.85	0.893
ResNet18	1024×1024	0.83	0.86	0.72	0.803
ResNet18	512×512	0.92	0.85	0.78	0.850
ResNet18	256×256	0.92	0.88	0.81	0.870
ResNet34	1024×1024	0.82	0.87	0.70	0.797
ResNet34	512×512	0.93	0.92	0.82	0.890
ResNet34	256×256	0.92	0.92	0.82	0.887
ResNet50	896×896	0.90	0.89	0.77	0.853
ResNet50	512×512	0.92	0.88	0.82	0.873
ResNet50	256×256	0.94	0.89	0.82	0.883
VGG16	1024×1024	0.63	-	-	0.630
VGG16	512×512	0.63	-	-	0.630
VGG16	256×256	0.93	0.89	0.81	0.877
VGG19	896×896	0.63	-	-	0.630
VGG19	512×512	0.63	-	-	0.630
VGG19	256×256	0.63	-	-	0.630
ViT-B/16	256×256	0.88	0.83	0.72	0.810

Πίνακας 2.3: Μέσοι όροι και τυπικές αποκλίσεις διαφόρων μετρικών για την απόδοση του MobileNetV2 υπολογισμένοι με χρήση πενταπλής διασταυρούμενης επικύρωσης.

Μέτρο	Μη Όγκος	Βιώσιμος Όγκος	Νέκρωση
Βαθμολογία $F1$	$0.95 \pm 0.02$	$0.90 \pm 0.04$	$0.85 \pm 0.03$
Ακρίβεια	$0.95 \pm 0.02$	$0.95 \pm 0.02$	$0.92 \pm 0.02$
Ειδικότητα	$0.96 \pm 0.03$	$0.96 \pm 0.02$	$0.96 \pm 0.02$
Ευαισθησία	$0.95 \pm 0.03$	$0.93 \pm 0.05$	$0.83 \pm 0.05$
Αξιοπιστία	$0.95 \pm 0.03$	$0.88 \pm 0.05$	$0.88 \pm 0.05$

Η αύξηση των διαστάσεων της εικόνας δεν φάνηκε να προσφέρει σημαντικά οφέλη, καθώς εισήγαγε νέα σημεία δεδομένων που θα μπορούσαν να προκαλέσουν υπερπροσαρμογή του μοντέλου. Αυτή η τάση παρατηρήθηκε στα δίκτυα (EfficientNetB0, EfficientNetB1, ResNet18, ResNet34, ResNet50 και MobileNetV2). Ορισμένα δίκτυα πέτυχαν ελαφρώς υψηλότερα αποτελέσματα όταν εκπαιδεύτηκαν σε εικόνες εισόδου μεγαλύτερων διαστάσεων, συμπεριλαμβανομένων των EfficientNetB3, EfficientNetB5 και EfficientNetB7. Αυτή η μικρή αύξηση των επιδόσεων που παρατηρήθηκε θα μπορούσε να αποδοθεί στον τρόπο με τον οποίο τα EfficientNets έχουν σχεδιαστεί. Πιο συγκεκριμένα, αυτή η οικογένεια δικτύων έχει σχεδιαστεί και εκπαιδευτεί κάνοντας χρήση μιας ομοιόμορφης κλιμάκωσης του βάθους, του πλάτους, και των διαστάσεων των εικόνων εισόδου. Κατά συνέπεια, όσο μεγαλύτερο, σε αριθμό παραμέτρων, είναι ένα EfficientNet, τότε τόσο μεγαλύτερες και οι διαστάσεις των εικόνων στο οποίο έχει προ-εκπαιδευτεί. Αυτό το γεγονός θα μπορούσε να τους δώσει ένα μικρό πλεονέκτημα κατά τη μεταφορά μάθησης σε μεγαλύτερες εικόνες.

#### 2.4.2 Επακόλουθο πείραμα

Με βάση τη μέση βαθμολογία  $F1$ , το MobileNetV2, με διατάξεις εικόνας εισόδου  $256 \times 256$  εικονοστοιχεία, επιλέχθηκε ως ο βέλτιστος συνδυασμός. Στη συνέχεια, το δίκτυο εκπαιδεύτηκε εκ νέου (όπως περιγράφεται στην ενότητα 2.5). Ο μέσος όρος  $\pm$  τυπική απόκλιση της βαθμολογίας  $F1$ , της ακρίβειας, της ειδικότητας, της ευαισθησίας και της αξιοπιστίας που επιτεύχθηκαν συνοψίζονται στον πίνακα 2.3. Επιπλέον, για την παροχή μιας καλύτερης επισκόπησης της ολικής απόδοσης του δικτύου, υπολογίστηκαν και αθροίστηκαν οι πίνακες σύγχυσης για κάθε διαμέριση από την πενταπλή διασταυρούμενη επικύρωση. Με αυτό τον τρόπο ο προκύπτων πίνακας σύγχυσης (πίνακας 2.4) περιέχει και τα 1144 δείγματα του συνόλου δεδομένων.

Το MobileNetV2 εκπαιδεύτηκε εκ νέου με τη χρήση πενταπλής διασταυρούμενης επικύρωσης για την επιβεβαίωση της συνέπειας των αποτελεσμάτων. Η μέση τιμή της βαθμολογίας  $F1$  ήταν 0,95 και 0,90 για MO και BO αντίστοιχα, υποδεικνύοντας ότι το μοντέλο είχε καλή απόδοση στον εντοπισμό αυτών των κατηγοριών.

Πίνακας 2.4: Πίνακας σύγχυσης του MobileNetV2 με συγκεντρωτικά αποτελέσματα για την πενταπλή διασταυρούμενη επικύρωση.

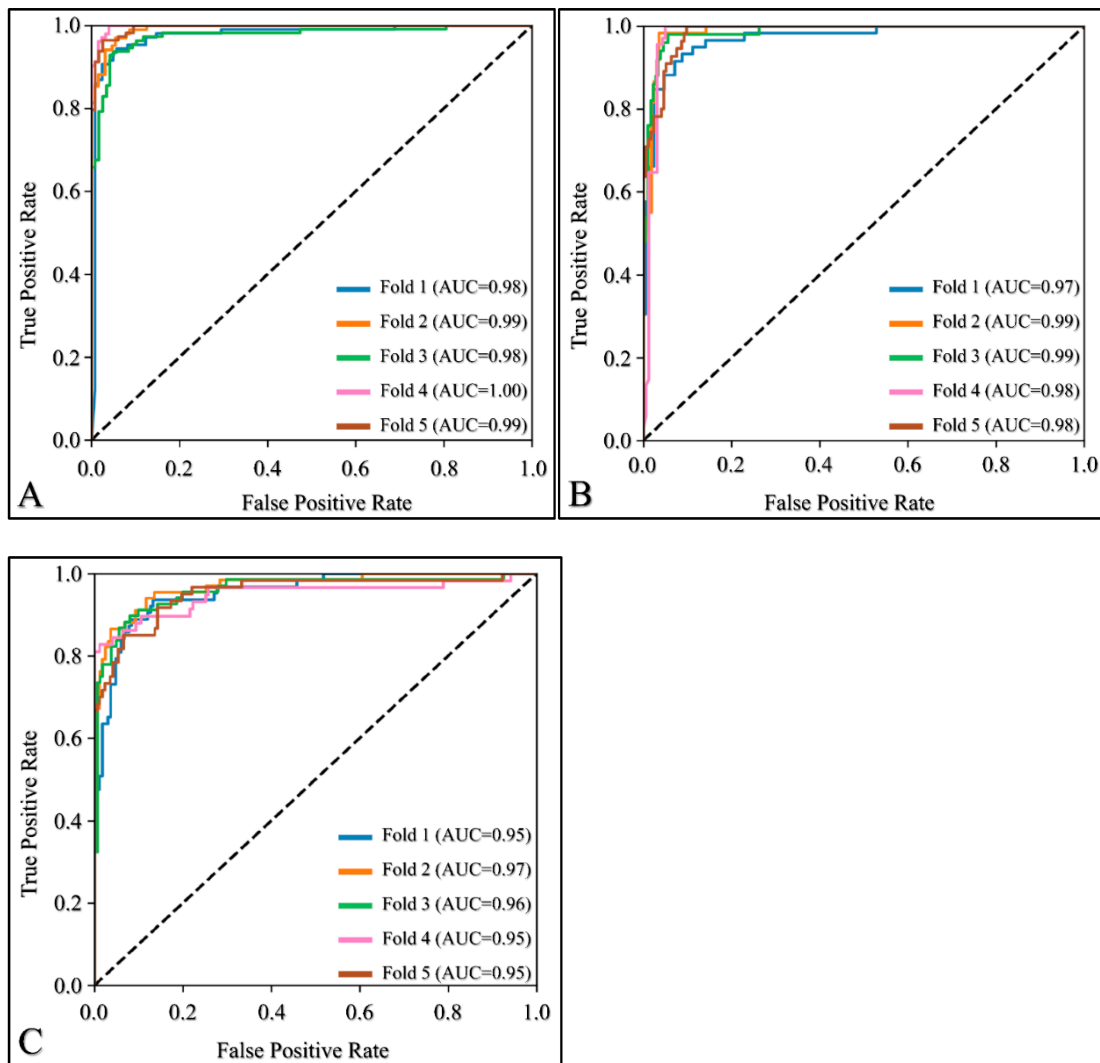
		Πρόβλεψη		
		ΜΟ	ΒΟ	ΝΚ
Πραγματικότητα	Μη Όγκος	510	7	19
	Βιώσιμος Όγκος	3	272	17
	Νέκρωση	24	30	262

Πίνακας 2.5: Μέσοι όροι και τυπικές αποκλίσεις των μετρικών απόδοσης του MobileNetV2 για πέντε διαμερίσεις, εξαιρουμένων των διαφορούμενων εικόνων από το υπό εξέταση σύνολο δεδομένων (ως διαφορούμενες εικόνες θεωρήθηκαν εκείνες που περιείχαν ταυτόχρονα ΝΚ και ΒΟ).

Μέτρο	Μη Όγκος	Βιώσιμος Όγκος	Νέκρωση
Βαθμολογία $F1$	$0.96 \pm 0.03$	$0.97 \pm 0.02$	$0.93 \pm 0.03$
Ακρίβεια	$0.96 \pm 0.03$	$0.99 \pm 0.01$	$0.97 \pm 0.02$
Ειδικότητα	$0.97 \pm 0.02$	$0.99 \pm 0.01$	$0.97 \pm 0.04$
Ευαισθησία	$0.95 \pm 0.06$	$0.98 \pm 0.05$	$0.93 \pm 0.09$
Αξιοπιστία	$0.97 \pm 0.02$	$0.97 \pm 0.03$	$0.93 \pm 0.08$

Ωστόσο, η βαθμολογία  $F1$  για την ΝΚ ήταν χαμηλότερη, με τιμή 0,85. Οι συγκρίσιμες τιμές ακρίβειας για την ΝΚ και τον ΒΟ υποδεικνύουν πως το μοντέλο ήταν εξίσου σίγουρο κατά την πρόβλεψη οποιασδήποτε από τις δύο κατηγορίες. Ωστόσο, η χαμηλότερη τιμή ευαισθησίας που παρατηρήθηκε για την ΝΚ υποδηλώνει πως το μοντέλο δυσκολεύτηκε να διακρίνει αυτή την κατηγορία από τις υπόλοιπες. Κατά την εξέταση του πίνακα 2.4, διαπιστώνεται ότι από τις 316, συνολικά, εικόνες ΝΚ, οι 24 ταξινομήθηκαν ως ΜΟ και οι 30 ως ΒΟ. Παρατηρείται επίσης ότι όταν οι εικόνες ΜΟ και ΒΟ ταξινομούσαν εσφαλμένα, σπάνια συγχέονταν μεταξύ τους, αλλά μάλλον χαρακτηρίζονταν ως ΝΚ. Αυτό υποδηλώνει ότι η παρατηρούμενη χαμηλότερη ακρίβεια για την ΝΚ δεν οφειλόταν σε ανισορροπία κλάσεων εντός του συνόλου δεδομένων, καθώς το δίκτυο θα ήταν προκατειλημμένο προς την ταξινόμηση περισσότερων δειγμάτων ως ΜΟ, εάν συνέβαινε κάτι τέτοιο. Αντίθετα, τα ευρήματα αυτά υποδηλώνουν ότι η ΝΚ μοιράζεται οπτικά χαρακτηριστικά τόσο με εικόνες ΜΟ όσο και με ΒΟ, οδηγώντας σε λανθασμένες εκτιμήσεις από το δίκτυο. Αυτή η συμπεριφορά παρατηρήθηκε στο σύνολο των πειραμάτων (πίνακας 2.2). Πράγματι, όπως εξηγείται στην ενότητα 2.2, περίπου το 20% των εικόνων ΝΚ απεικονίζαν επίσης ΒΟ. Η επανεκπαίδευση του δικτύου χωρίς αυτές τις διαφορούμενες εικόνες αύξησε σημαντικά την ακρίβεια του δικτύου (πίνακας 2.5).

Για την πρόσθετη επισκόπηση της απόδοσης του MobileNetV2 σε κάθε διαμέριση της πενταπλής διασταυρούμενης επικύρωσης, χρησιμοποιήθηκε η ανάλυση ROC με τα αποτελέσματα να απεικονίζονται στο Σχήμα 2. Οι καμπύλες ROC που υπολογίστηκαν για κάθε διαμέριση απεικονίζονται επάλληλα για την διευκόλυνση



Σχήμα 2.3: Καμπύλες ROC για: (A) Μη όγκο, (B) βιώσιμο όγκο, (C) νέκρωση.

της σύγκρισης των αποτελεσμάτων. Η απόδοση σε κάθε διαμέριση ήταν συνεπής, με πολύ υψηλές τιμές του AUC που κυμάνθηκαν από 0,98 έως 1 για τον ΜΟ (Σχήμα 2.3Α), 0,97 έως 0,99 για τον ΒΟ (Σχήμα 2.3Β) και 0,95 έως 0,97 για την ΝΚ (Σχήμα 2.3Γ). Οι παρατηρούμενες ομοιότητες σε όλες τις τιμές του AUC για όλες τις κατηγορίες υποδηλώνουν ότι το επιλεγμένο νευρωνικό δίκτυο είναι πιθανό να επιτύχει παρόμοια απόδοση ταξινόμησης σε διαφορετικά, αλλά συγκρίσιμα, σύνολα δεδομένων.

## 2.5 Συζήτηση

Στην παρούσα μελέτη, διερευνήθηκαν διάφορες αρχιτεκτονικές δικτύων με σκοπό την αξιολόγησή τους στην ταξινόμηση ιστολογικών εικόνων ΟΣ.

Πίνακας 2.6: Σύγκριση της συνολικής ακρίβειας με σχετική εργασία στο ίδιο σύνολο δεδομένων.

Μελέτη	Μέθοδος	Στρατ. Επικύρ.	Ακρίβεια
Arunachalam et al. (2019)	ΣNN	Παρακράτηση	0.910
Anisuzzaman et al. (2021)	VGG19	Παρακράτηση	0.940
Bansal et al. (2022)	Συνδ. χαρακτ. ΧΠ και BM	Παρακράτηση	0.995
Παρούσα μελέτη	MobileNetV2	Διασταυρούμενη	0.910

### 2.5.1 Σύγκριση νευρωνικών δικτύων

Στο πλαίσιο της καλύτερης κατανόησης της αποτελεσματικότητας των δικτύων, πραγματοποιήθηκε σύγκριση των αποτελεσμάτων της επανεκπαίδευσης του MobileNetV2 με άλλες μελέτες που χρησιμοποίησαν το ίδιο σύνολο δεδομένων ΟΣ (Πίνακας 6). Προκειμένου να διασφαλιστεί μια δίκαιη σύγκριση, αποκλείστηκαν οι μελέτες που πραγματοποίησαν περαιτέρω τροποποιήσεις στις εικόνες, όπως την περικοπή υποπεριοχών για την τεχνητή επέκταση του συνόλου δεδομένων. Παρότι τέτοιου είδους τροποποιήσεις έχουν παρατηρηθεί στο παρελθόν να βελτιώνουν την απόδοση ταξινόμησης [73], είναι πιθανό να οδηγήσουν σε διαρροή δεδομένων σε περίπτωση μη προσεκτικού χειρισμού. Για παράδειγμα, δύο υποπεριοχές από την ίδια αντικειμενοφόρα πλάκα οι οποίες βρίσκονται σε κοντινή απόσταση, μπορεί να εισέλθουν και στο σύνολο εκπαίδευσης και ελέγχου. Ως αποτέλεσμα, η απόδοση του δικτύου θα επικυρωθεί σε εικόνες που είναι παρόμοιες με το σύνολο εκπαίδευσης, με αποτέλεσμα οι ακρίβειες να είναι μάλλον υπερβολικά αισιόδοξες [85].

Οι Arunachalam et al. (2019) [73] ανέπτυξαν μια προσαρμοσμένη αρχιτεκτονική ΣNN για την ταξινόμηση του ίδιου συνόλου δεδομένων ΟΣ. Υπολόγισαν τα αποτελέσματά τους με χρήση ενός περιορισμένου υποσυνόλου 230 εικόνων, καθώς πραγματοποίησαν έναν διαχωρισμό 80-20 μεταξύ του συνόλου εκπαίδευσης και του συνόλου επικύρωσης (στρατηγική επικύρωσης με παρακράτηση, αγγλ. holdout). Πέτυχαν συνολική ακρίβεια 0,91 και ευαισθησία 89,5, 92,6 και 91,5 για MO, BO και NK αντίστοιχα. Παρόλο που οι συγγραφείς δεν ανέφεραν συγκεκριμένες λεπτομέρειες εφαρμογής της διαδικασίας επανεκπαίδευσης, όπως το μέγεθος εισόδου εικόνας όταν χρησιμοποιούν WSIs, μια βασική διαφορά φαίνεται να είναι ότι η προσέγγισή τους για την ταξινόμηση αποτελείται από δύο στάδια. Το πρώτο στάδιο περιελάμβανε την ταξινόμηση της εικόνας είτε ως όγκου είτε ως μη όγκου. Το δεύτερο στάδιο εκτελούνταν υπό όρους: εάν η εικόνα ταξινομούταν ως όγκος, στη συνέχεια ταξινομούταν περαιτέρω ως BO ή NK. Με τη χρήση αυτής της προσέγγισης, οι συγγραφείς πέτυχαν αποτελέσματα παρόμοια με της παρούσας εργασίας, παρά το γεγονός ότι χρησιμοποίησαν ένα πολύ απλούστερο δίκτυο με μόλις τρία επίπεδα συνελίξεων που εκπαιδεύτηκε από την αρχή (σε αντίθεση με τη παρούσα προσέγγιση μεταφοράς μάθησης). Αυτό υποδηλώνει ότι η ιεραρχική τους προσέγγιση ενδεχομένως να βελτιώνει περαιτέρω τα αποτελέσματα της παρούσας μελέτης.



Οι Anisuzzaman et al. (2021) [80] εκπαιδύσαν διάφορα ΣΝΝ, πετυχαίνοντας το καλύτερο αποτέλεσμα με το VGG19, επιτυγχάνοντας συνολική ακρίβεια 0,94 με χρήση του βελτιστοποιητή Adam και ρυθμό μάθησης 0,01. Ομοίως με την μελέτη των Arunachalam et al. (2019) [73], ακολούθησαν μια στρατηγική επικύρωσης με παρακράτηση και υπολόγισαν τα αποτελέσματά τους σε ένα μικρό υποσύνολο 230 εικόνων. Ωστόσο, τα αποτελέσματά τους διαφέρουν σημαντικά από την παρούσα μελέτη, καθώς παρατηρήθηκε ότι το VGG19 δεν σύγκλινε σε κάποια λύση λόγω του ότι ο ρυθμός μάθησης  $3 \times 10^{-4}$  για τον βελτιστοποιητή Adam ήταν πολύ υψηλός. Κατά την επανεκπαίδευση του VGG19 με ρυθμό μάθησης 0,01, διαπιστώθηκε και πάλι δεν πως δεν σύγκλινε σε λύση (τα δεδομένα δεν παρουσιάζονται). Επιπλέον, οι συγγραφείς ανέφεραν μια πολύ χαμηλή μέση βαθμολογία  $F1$  με το ResNet50, ενώ αντίθετα στην παρούσα μελέτη διαπιστώθηκε πως το ResNet50 είχε καλή απόδοση σε όλες τις περιπτώσεις. Παρόλο που έγινε προσπάθεια να ακολουθηθεί όσο το δυνατόν πιο πιστά η υλοποίηση των συγγραφέων, προσθέτοντας στο τέλος του δικτύου δύο πλήρως συνδεδεμένα επίπεδα που περιείχαν 512 και 1024 νευρώνες, υπήρχαν και πάλι κάποιες διαφορές που θα μπορούσαν να επηρεάσουν το αποτέλεσμα. Πρώτον, οι συγγραφείς χρησιμοποίησαν την βιβλιοθήκη Keras για την δημιουργία του μοντέλου VGG19, ενώ αντίθετα στην παρούσα μελέτη χρησιμοποιήθηκε η υλοποίηση που παρέχει η βιβλιοθήκη Torchvision της PyTorch. Δεύτερον, χρησιμοποιήθηκε μέγεθος παρτίδας 4 για εικόνες εισόδου διαστάσεων  $512 \times 512$ , και 8 για εικόνες διαστάσεων  $256 \times 256$ , το οποίο διέφερε από το μέγεθος παρτίδας 80 των συγγραφέων. Τρίτον, οι συγγραφείς έκαναν υποδειγματοληψία όλων των εικόνων σε διαστάσεις  $375 \times 375$  για την εκπαίδευση και την αξιολόγηση. Τέλος, στην παρούσα μελέτη χρησιμοποιήθηκε μια ελαφρώς διαφορετική υλοποίηση του βελτιστοποιητή Adam, με την ονομασία AdamW, η οποία διορθώνει τον τρόπο με τον οποίο υλοποιείται η μείωση του βάρους [82]).

Οι Bansal et al. (2022) [86] χρησιμοποίησαν έναν συνδυασμό χειροποίητων (ΧΠ) χαρακτηριστικών και χαρακτηριστικών βαθιάς μάθησης (BM) που εξάγονται από το δίκτυο Xception για να εκπαιδύσουν μία Μηχανή Διανυσμάτων Υποστήριξης (ΜΔΥ, αγγλ. Support Vector Machine - SVM) με πυρήνα την Ακτινική Συνάρτηση (Radial Basis Function - RBF). Ανέφεραν μια εξαιρετικά υψηλή συνολική ακρίβεια 0,995 έπειτα από επικύρωση σε ένα μικρό υποσύνολο 219 εικόνων, αποδεικνύοντας ένα σαφές πλεονέκτημα όταν συνδυάζονται χαρακτηριστικά από διαφορετικές μεθόδους. Στην ίδια εργασία, αναφέρθηκε συνολική ακρίβεια 0,968 με το EfficientNetB0, η οποία είναι υψηλότερη σε σύγκριση με τα ευρήματά της παρούσας εργασίας. Από τη μία πλευρά, αυτή η απόκλιση είναι αναμενόμενη κατά την επικύρωση των δικτύων σε διαφορετικά σύνολα δειγμάτων. Για παράδειγμα, κατά τα πειράματα της παρούσας μελέτης παρατηρήθηκαν διαμερίσεις όπου το MobileNetV2 πέτυχε συνολική ακρίβεια μεγαλύτερη από 0,96, αλλά αυτή δεν ήταν μια ρεαλιστική εκτίμηση, όπως αποδεικνύεται από τα αποτελέσματα της διασταυρούμενης επικύρωσης. Από την άλλη πλευρά, η αυξημένη αναφερόμενη ακρίβεια

μπορεί επίσης να οφείλεται σε βελτιστοποίηση των συγγραφέων. Πιο συγκεκριμένα, οι συγγραφείς τροποποίησαν το δίκτυο μετά την εκπαίδευση αφαιρώντας το τελευταίο στρώμα FC για να εξάγουν ένα σύνολο χαρακτηριστικών, τα οποία στη συνέχεια φιλτράρονταν με τη χρήση ενός αλγορίθμου δυαδικής αριθμητικής βελτιστοποίησης και ταξινομούσαν με τη χρήση ενός RBF-SVM. Απαιτείται περαιτέρω αξιολόγηση με χρήση διασταυρούμενης επικύρωσης και με τις δύο προσεγγίσεις για να εκτιμηθεί κατά πόσον η τεχνική αυτή μπορεί να βελτιώσει περαιτέρω την απόδοση του δικτύου.

## **2.5.2 Περιορισμοί και μελλοντικές προοπτικές**

Παρόλο που η τρέχουσα μελέτη παρήγαγε πολύτιμες γνώσεις σχετικά με τη χρήση δικτύων βαθιάς μάθησης για την ταξινόμηση ιστολογικών εικόνων ΟΣ, υπάρχουν αρκετοί περιορισμοί που πρέπει να ληφθούν υπόψη. Καταρχάς, η μελέτη συγκρίνει την απόδοση διαφορετικών αρχιτεκτονικών δικτύων χρησιμοποιώντας προκαθορισμένες υπερπαραμέτρους. Αν και τα αποτελέσματα παρέχουν ισχυρές ενδείξεις ότι τα μικρότερα δίκτυα μπορεί να έχουν καλύτερη απόδοση από τα μεγαλύτερα και ότι τα MobileNetV2 και EfficientNetB0 είναι τα πιο αποτελεσματικά μοντέλα, είναι πιθανό ότι η απόδοση άλλων δικτύων θα μπορούσε να βελτιωθεί με τη βελτιστοποίηση των υπερπαραμέτρων τους. Συνεπώς, μελλοντικές έρευνες θα μπορούσαν να εξετάσουν τον αντίκτυπο της βελτιστοποίησης των υπερπαραμέτρων στην απόδοση των αξιολογηθέντων δικτύων.

Επιπρόσθετα, το σύνολο δεδομένων ΟΣ περιείχε μόνο περιορισμένο αριθμό εικόνων που απεικόνιζαν πολλαπλές κατηγορίες ιστών, ενώ η πλειονότητα των εικόνων περιείχε μόνο ΜΟ, ΒΟ ή ΝΚ. Παρατηρήθηκε ότι μετά την αφαίρεση των διφορούμενων εικόνων που απεικόνιζαν ταυτόχρονα ΒΟ και ΝΚ, η απόδοση του δικτύου βελτιώθηκε. Συνεπώς, αναμένεται πως η ακρίβεια του δικτύου θα μειωθεί σε ένα σενάριο όπου οι εικόνες περιέχουν οποιονδήποτε συνδυασμό αυτών των κατηγοριών. Η αύξηση του συνόλου δεδομένων, ώστε να περιλαμβάνει περισσότερα διφορούμενα δείγματα ή η επέκταση των ετικετών ώστε να περιλαμβάνουν ταξινόμηση σε επίπεδο εικονοστοιχείων για προσεγγίσεις τμηματοποίησης, θα μπορούσαν να οδηγήσουν σε πιο αξιόπιστα αποτελέσματα.

Τέλος, η μελέτη περιορίστηκε από το μέγεθος του συνόλου δεδομένων ΟΣ, το οποίο περιείχε εικόνες από μόλις τέσσερις ασθενείς που επιλέχθηκαν από τους παθολόγους βάσει της ποικιλίας των όγκων. Διαφορετικοί τύποι όγκων, στάδια ή ακόμη και δημογραφικά χαρακτηριστικά μπορεί να έχουν ως αποτέλεσμα τη σημαντική διαφοροποίηση των εικόνων. Αυτοί οι παράγοντες θα μπορούσαν να επηρεάσουν τη γενικευσιμότητα των ευρημάτων, καθώς το σύνολο δεδομένων δεν θα αντιπροσωπεύει επαρκώς τον ευρύτερο πληθυσμό ασθενών με ΟΣ. Συνεπώς, απαιτούνται μεγαλύτερα και ποικιλόμορφα σύνολα δεδομένων ΟΣ για να επιβεβαιωθεί η αποτελεσματικότητα των δικτύων που εντοπίστηκαν και να διασφαλιστεί η δυνατότητα γενίκευσής τους σε νέα σύνολα δεδομένων και πληθυσμούς. Η

συμπερίληψη δεδομένων από πολλαπλά κέντρα θα μπορούσε να συμβάλει στην περαιτέρω αύξηση της ποικιλομορφίας των συνόλων δεδομένων.

## **Κεφάλαιο 3**

# **Ανίχνευση Αλλεργικής Δερματίτιδας Εξ Επαφής σε Επιδερμίδικές Δοκιμασίες Επιθέματος**

### **3.1 Αλλεργική Δερματίτιδα Εξ Επαφής**

Η Αλλεργική Δερματίτιδα εξ Επαφής (ΑΔΕ) (αγγλ. contact dermatitis) είναι μια ευρέως διαδεδομένη ανοσολογική αντίδραση του δέρματος που σχετίζεται με την έκθεση σε περιβαλλοντικούς χημικούς παράγοντες, επηρεάζοντας όλες τις ηλικιακές ομάδες και φύλα [87].

#### **3.1.1 Παθογένεση**

Οι μηχανισμοί που διέπουν την ευαισθητοποίηση του δέρματος και την ΑΔΕ περιλαμβάνουν πολύπλοκες ανοσολογικές διεργασίες. Η ΑΔΕ θεωρείται μια μορφή υπερευαισθησίας καθυστερημένου τύπου με τη μεσολάβηση Τ λεμφοκυττάρων. Τα βασικά στοιχεία αυτής της διαδικασίας περιλαμβάνουν χημικά αλλεργιογόνα, έκθεση σε αλλεργιογόνα, επεξεργασία, μεταφορά και ενεργοποίηση των Τ λεμφοκυττάρων [87].

Τα χημικά αλλεργιογόνα είναι συνήθως απτένια, τα οποία δεν μπορούν να διεγείρουν από μόνα τους ανοσολογική απόκριση. Γίνονται ανοσογόνα σχηματίζοντας σταθερές ενώσεις με πρωτεΐνες, δημιουργώντας συζεύγματα απτενίου-πρωτεΐνης. Αυτά τα συζεύγματα αναγνωρίζονται και επεξεργάζονται από το ανοσοποιητικό σύστημα, οδηγώντας σε ευαισθητοποίηση.

Μέχρι σήμερα, περισσότερες από 3.000 χημικές ουσίες έχουν αναγνωριστεί ως αλλεργιογόνα που δύναται να προκαλέσουν ΑΔΕ [88]. Εκδηλώνεται συνηθέστερα σε εκτεθειμένα μέρη του σώματος, όπως τα χέρια και το πρόσωπο. Τα συμπτώματα εκδηλώνονται ως εκζεματική δερματίτιδα και περιλαμβάνουν ξηρότητα του

δέρματος, κνησμό και φλεγμονή.

### **3.1.2 Διάγνωση**

Ο εντοπισμός του/των αλλεργιογόνου/ων που προκαλούν την ΑΔΕ αποτελεί προϋπόθεση για τη σωστή αντιμετώπισή της. Η μέθοδος της επιδερμидικής δοκιμασίας επιθέματος (patch test), αν και χρονολογείται από το 1895, παραμένει μέχρι και σήμερα το χρυσό πρότυπο για τον εντοπισμό των αλλεργιογόνων που μπορούν να προκαλέσουν ΑΔΕ σε έναν ασθενή [89], [90]. Η διαδικασία της εξέτασης περιλαμβάνει την επιλογή μιας κατάλληλης σειράς αλλεργιογόνων που είναι πιθανότερο να προκαλέσουν ΑΔΕ με βάση το ιστορικό του ασθενούς και τη φύση της δερματικής πάθησής του, και στη συνέχεια την επανέκθεσή του στα επιλεγμένα αλλεργιογόνα υπό ελεγχόμενες συνθήκες για την επαλήθευση της διάγνωσης [91]. Τα κοινά αλλεργιογόνα περιλαμβάνουν μέταλλα (όπως το νικέλιο), αρώματα, συντηρητικά και χημικές ουσίες όπως το καουτσούκ. Το επιδερμидικό δοκιμαστικό επίθεμα είναι μια ειδικά κατασκευασμένη ταινία εφοδιασμένη με « θαλαμίσκους » (αγγλ. chambers), οι οποίοι είναι υποδοχές κατασκευασμένες από αλουμίνιο ή πλαστικό στις οποίες τοποθετούνται τα αλλεργιογόνα. Τα δοκιμαστικά επιθέματα επικολλώνται απευθείας στο δέρμα του ασθενούς, συνήθως στο άνω μέρος της πλάτης. Παραμένουν στη θέση τους για 48-72 ώρες, οπότε και αφαιρούνται για μια πρώτη αξιολόγηση. Κατά την ανάγνωση της δοκιμασίας, η περιοχή επικόλλησης εξετάζεται για ενδείξεις δερματικής αντίδρασης, όπως ερυθρότητα, οίδημα ή κνησμό. Ο ασθενής επανέρχεται για δεύτερη αξιολόγηση έπειτα από επιπλέον 24-48 ώρες. Η εξέταση είναι σχετικά ασφαλής και οι αντιδράσεις είναι συνήθως ήπιες και υποχωρούν μέσα σε λίγες ημέρες. Ωστόσο, σε σπάνιες περιπτώσεις μπορεί να εμφανιστούν σοβαρές αντιδράσεις [90]. Τα αποτελέσματα της επιδερμидικής δοκιμασίας βοηθούν στον προσδιορισμό των αλλεργιογόνων τα οποία προκαλούν αλλεργική αντίδραση στον ασθενή, μειώνοντας τον κίνδυνο μελλοντικών επεισοδίων.

## **3.2 Αντικείμενο της Μελέτης**

Τα κριτήρια ICDRG (International Contact Dermatitis Research Group) για την ανάγνωση των δοκιμασιών επιθέματος είναι παγκοσμίως αναγνωρισμένα και βασίζονται στην οπτική εξέταση του δέρματος και την ψηλάφηση της μορφολογίας της αντίδρασης. Ωστόσο, παρά την αποτελεσματικότητά της ως διαγνωστικού εργαλείου, η δοκιμασία επιθέματος υπόκειται στο φαινόμενο « μεροληψίας του παρατηρητή », καθώς η ερμηνεία των αποτελεσμάτων μπορεί να διαφέρει μεταξύ των κλινικών. Για παράδειγμα, έχει διαπιστωθεί μεταβλητότητα μεταξύ των παρατηρητών στη διάκριση μεταξύ αμφίβολων και ερεθιστικών αντιδράσεων, καθώς και στη διάκριση μεταξύ αμφίβολων και ασθενών θετικών αντιδράσεων [92]. Για τον μετριασμό αυτού του φαινομένου συνιστάται η συνεχής προτυποποίηση και η

εκπαίδευση του κλινικού προσωπικού. Οι Johansen et al. αναφέρουν επίσης ότι για να διασφαλιστεί μια ομοιόμορφη τεχνική ανάγνωσης της δοκιμασίας επιθέματος μεταξύ των ατόμων, είναι απαραίτητη η συνεχής εκπαίδευση σε επιστημονικές συναντήσεις [92]. Αυτό υπογραμμίζει περαιτέρω την υποκειμενικότητα που ενέχει η ανάγνωση των δοκιμασιών επιθέματος, καθώς απαιτείται ένα ορισμένο επίπεδο εμπειρογνωμοσύνης και κατάρτισης για την ακριβή ερμηνεία των αποτελεσμάτων. Πέρα από τη μεροληψία του παρατηρητή, οι δοκιμασίες επιθέματος απαιτούν επίσης πολύτιμους ανθρώπινους πόρους, μια πρόκληση που μπορεί να είναι ιδιαίτερα έντονη σε μεγάλες κλινικές.

Η χρήση μοντέλων βαθιάς μηχανικής μάθησης για την αντιμετώπιση παρόμοιων προκλήσεων σε ιατρικά περιβάλλοντα είναι ένα συνεχώς μεταβαλλόμενο και ταχύτατα εξελισσόμενο πεδίο [93]–[95]. Ωστόσο, η χρήση της βαθιάς μάθησης για την αυτοματοποίηση της ανίχνευσης δερματικών αντιδράσεων σε επιδερμидικές δοκιμασίες είναι, προς το παρόν, ένα θέμα που δεν έχει διερευνηθεί επαρκώς. Στην παρούσα έρευνα, διερευνάται η δυνατότητα ενός μοντέλου βαθιάς μάθησης να αυτοματοποιήσει την ανάγνωση των επιδερμидικών δοκιμασιών, καθώς και την επίδραση των διαφόρων τεχνικών προεπεξεργασίας και τρόπων λήψης εικόνας στην ακρίβεια της αυτοματοποιημένης διάγνωσης.

Για το σκοπό αυτό, χρησιμοποιήθηκε η κάμερα Antera 3D® για την συλλογή ενός συνόλου δεδομένων 1579 πολυμορφικών εικόνων από 200 ασθενείς, που απεικονίζουν δερματικές αντιδράσεις σε μια σειρά αλλεργιογόνων. Τα αποτελέσματα αναδεικνύουν μια τεχνική προεπεξεργασίας που εστιάζει στην υπό εξέταση περιοχή διατηρώντας ταυτόχρονα όλη την πληροφορία της εικόνας, καθώς και τον συνδυασμό της έγχρωμης εικόνας και του ψευδοχρωματικού χάρτη συγκέντρωσης της αιμοσφαιρίνης για τη βελτίωση της διαγνωστικής ακρίβειας. Τα αποτελέσματα αυτής της μελέτης θα μπορούσαν να έχουν σημαντικές επιπτώσεις στην αυτοματοποιημένη διάγνωση της ΑΔΕ, επιτρέποντας, ενδεχομένως, την ταχύτερη και ακριβέστερη διάγνωση, καθώς και τη μείωση του φόρτου εργασίας των κλινικών ιατρών. Κατά τη συγγραφή της παρούσας διδακτορικής διατριβής, μόνο ένα άρθρο βρέθηκε να έχει δημοσιευθεί το οποίο να διερευνά το θέμα της αυτοματοποιημένης διάγνωσης επιδερμидικών δοκιμασιών [96]. Συνεπώς, η παρούσα εργασία αποτελεί μία αφετηρία για μελλοντικές προσπάθειες.

### **3.3 Υλικό και Μέθοδοι**

#### **3.3.1 Σύνολο δεδομένων**

Το σύνολο δεδομένων που χρησιμοποιήθηκε στην παρούσα μελέτη περιλαμβάνει εικόνες από 200 ασθενείς, οι οποίες συλλέχθηκαν στο Εργαστήριο Επιδερμидικών Δοκιμασιών και Κέντρο Αναφοράς Επαγγελματικών Δερματοπαθειών, του Πανεπιστημιακού Νοσοκομείου «Ανδρέας Συγγρός». Το λεπτομερές κλινικό πρωτόκολλο περιγράφεται στην ενότητα 3.3.1.

## **Κριτήρια ένταξης**

Όλοι οι ασθενείς εισήχθησαν στο νοσοκομείο με την υποψία ΑΔΕ [97], [98]. Από το σύνολο των ασθενών επιλέχθηκαν τυχαία 200, προκειμένου να αποφευχθεί η μεροληψία λόγω επιλεκτικής συμπερίληψης.

## **Κριτήρια αποκλεισμού**

Αν και γενικότερα δεν υπάρχει οριστική συναίνεση για τα κριτήρια αποκλεισμού από μια διαδικασία επιδερμικής δοκιμασίας, οι περισσότερες μελέτες συμφωνούν ότι η χρήση κορτικοστεροειδών [99], ανοσοκατασταλτικών (όπως η κυκλοσπορίνη και τα χημειοθεραπευτικά) [100], αντιφλεγμονώδεις θεραπείες [92], άλλες χρόνιες δερματοπάθειες [99] και η υψηλή έκθεση στην υπεριώδη ακτινοβολία [101] (όπως για παράδειγμα κατά τη διάρκεια του καλοκαιριού) είναι παράγοντες που ενδέχεται να προκαλέσουν ψευδώς αρνητικά ή ψευδώς θετικά αποτελέσματα και συνεπώς πρέπει να αποφεύγονται. Ως εκ τούτου, από την παρούσα μελέτη αποκλείστηκαν οι ασθενείς που εισήχθησαν και τελούσαν υπό αντιφλεγμονώδη αγωγή, υπό θεραπεία με κυκλοσπορίνη, υπό χρόνια χρήση κορτικοστεροειδών, υπό χημειοθεραπευτικές αγωγές, καθώς και οι πάσχοντες από άλλες χρόνιες δερματοπάθειες.

## **Επιδερμική δοκιμασία επιθέματος και κλινική αξιολόγηση**

Οι περισσότερες χώρες διαθέτουν μια βασική σειρά που περιέχει τα συνηθέστερα αλλεργιογόνα στα οποία εμφανίζει ευαισθησία ο γενικός πληθυσμός. Η σειρά αυτή δύναται να αλλάζει με την πάροδο του χρόνου. Ειδικότερα για την Ελλάδα, χρησιμοποιείται η ευρωπαϊκή βασική σειρά, η οποία περιέχει διάφορες κατηγορίες αλλεργιογόνων, όπως μέταλλα, αρώματα, συντηρητικά, ελαστικά, τοπικά θεραπευτικά και έκδοχα [102]. Στην παρούσα μελέτη, η ευαισθητοποίηση αξιολογήθηκε χρησιμοποιώντας μια σειρά 28 αλλεργιογόνων από την ευρωπαϊκή σειρά, στην οποία προστέθηκαν επιπλέον αλλεργιογόνα με σκοπό την ενημέρωση της εθνικής βασικής σειράς αλλεργιογόνων. Ο πίνακας 3.1 περιέχει τον πλήρη κατάλογο των αλλεργιογόνων που χρησιμοποιήθηκαν.

Οι επιδερμικές δοκιμασίες πραγματοποιήθηκαν σύμφωνα με τις κατευθυντήριες γραμμές της Ευρωπαϊκής Εταιρείας Δερματίτιδας εξ Επαφής (European Society of Contact Dermatitis) [92]. Σε κάθε ασθενή επικολλήθηκαν τρία επιθέματα 5 × 2 θαλαμίσκων στο μέσο της ράχως, σε δέρμα χωρίς τρίχες ή βλάβες. Ο βέλτιστος χρόνος έκθεσης ορίστηκε σε 48 ώρες (δύο ημέρες). Σύμφωνα με τις αντίστοιχες κατευθυντήριες γραμμές, η ανάγνωση της επιδερμικής δοκιμασίας πρέπει να πραγματοποιείται δύο φορές, με την πρώτη αμέσως μετά την αφαίρεση του επιθέματος (48 ώρες έπειτα από την πρώτη έκθεση) και τη δεύτερη 2-4 ημέρες αργότερα. Στην πλειονότητα των ασθενών, η μέγιστη αντίδραση παρατηρήθηκε μεταξύ της 3ης και της 4ης ημέρας έπειτα από την εφαρμογή των επιθεμάτων. Το

Πίνακας 3.1: Σειρά αλλεργιογόνων που χρησιμοποιήθηκε στις επιδερμидικές δοκιμασίες.

Αλλεργιογόνο	Συγκέντρωση (%)
Διχρωμικό κάλιο	0,5
Θειική νεομυκίνη	20
Μίγμα θειουράμων	1
Παραφαινυλενοδιαμίνη	1
Χλωριούχο κοβάλτιο	1
Βενζοκαΐνη	7
Φορμαλδεΐδη	2
Κολοφώνιο	20
Μεθακρυλικό 2-υδροξυαιθυλο	2
Βάλσαμο του Περού	25
Μίγμα μαύρου πλαστικού	0,1
Αλκοόλες μαλλιού	30
Μερκαπτοβενζολθειαζόνη	2
Εποξική Ρητίνη	1
Μίγμα παραβενίου	16
Ρητίνη π. βουτυλφενόλη φορμαλδεΐδη	1
Μίγμα αρωμάτων	8
Διυδροχλωρική αιθυλοδιαμίνη	1
Quaternium	1
Θειικό νικέλιο	5
Kathon	0,02
Μερκαπτοβενζοθειαζόλη	2
Πρόπολη	10
Μίγμα χρωστικών για υφάσματα	6,6
Γέλη πετρελαίου	100
Βουδεσονίδη	0,01
Θειομερσάλη	0,1
Μεθυλδιβρωμογλουταρονιτρίλιο	0,5
Μίγμα σεσκιτερπενικής λακτόνης	0,1
Μείγμα αρωμάτων	14

αποτέλεσμα της επιδερμидικής δοκιμασίας αξιολογήθηκε με βάση τη μορφολογία των βλαβών, με τις αντιδράσεις να ταξινομούνται ως θετικές/αλλεργικές σε κλίμακα (i) + ασθενής, (ii) ++ έντονη, ή (iii) +++ πολύ έντονη, σύμφωνα με τα κριτήρια της Διεθνούς Ομάδας Έρευνας για την Δερματίτιδα εξ Επαφής (International Contact Dermatitis Research Group - ICDRG). Όλες οι υπόλοιπες αντιδράσεις, συμπεριλαμβανομένων των ερεθιστικών, ταξινομήθηκαν ως αρνητικές.

### Ανάλυση δέρματος

Το δέρμα αξιολογήθηκε με μια κάμερα Antera 3D® (Miravex Limited, Δουβλίνο, Ιρλανδία) 0 ώρες, 48 ώρες και 72 ώρες μετά την πρώτη έκθεση στα αλλεργιογόνα. Η ιδιαιτερότητα της συγκεκριμένης κάμερας έγκειται στο ότι χρησιμοποιεί πολυκατευθυντικό φωτισμό για τη λήψη εικόνων από διαφορετικές γωνίες,



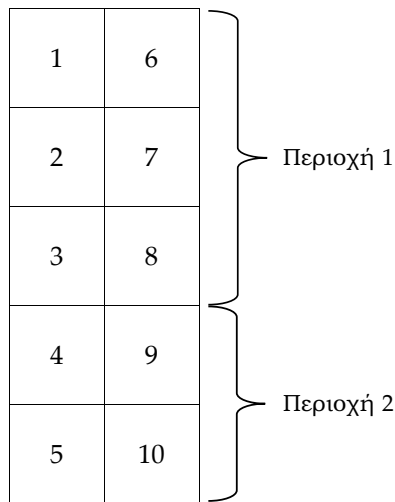
οι οποίες στη συνέχεια αναλύονται από το συνοδευτικό λογισμικό της για την ανακατασκευή της επιφάνειας του δέρματος σε δύο και τρεις διαστάσεις. Επιπλέον, η χρήση LED φωτισμού ικανού να εκπέμπει διαφορετικά μήκη κύματος φωτός επιτρέπει στη συσκευή να εκτελεί τόσο χωρική όσο και πολυφασματική ανάλυση, αποτυπώνοντας έτσι την τοπογραφία του δέρματος και τη συγκέντρωση χρωμοφόρων. Επιπλέον, μέσω των φίλτρων πόλωσης της κάμερας και ειδικών τεχνικών, οι εικόνες που λαμβάνονται είναι ανεξάρτητες από τις συνθήκες φωτισμού, γεγονός που εγγυάται την ακρίβεια και την αναπαραγωγικότητα των αποτελεσμάτων σε κάθε μέτρηση [103]–[105].

Με τη χρήση της κάμερας είναι δυνατή η λήψη μιας περιοχής επιφανείας 3136 τετραγωνικών χιλιοστών (56 × 56 mm). Καθώς αυτό το μέγεθος περιοχής αποδείχθηκε πολύ μικρό ακόμη και για την πλήρη απεικόνιση ενός, μόνο, επιθέματος, απαιτήθηκαν αντ' αυτού δύο λήψεις για κάθε ένα από τα συνολικά τρία επιθέματα. Επομένως, κάθε επίθεμα χωρίστηκε σε δύο περιοχές οι οποίες αριθμήθηκαν ως 1 και 2 (Σχήμα 3.1).

Για κάθε μεμονωμένη λήψη, το λογισμικό της κάμερας παρήγαγε έξι διαφορετικές εικόνες, οι οποίες εφεξής αναφέρονται ως « μορφές » για την διάκρισή τους. Οι εικόνες αυτές αντιστοιχούν:

1. Στην αρχική έγχρωμη εικόνα, η οποία απεικονίζει το δέρμα όπως φαίνεται με γυμνό μάτι.
2. Στο ερύθημα, το οποίο απεικονίζεται ως ένας ψευδοχρωματικός χάρτης των περιοχών υψηλότερης συγκέντρωσης αιμοσφαιρίνης.
3. Στην υφή, μια εικόνα που απεικονίζει την ποσοτικοποίηση της τραχύτητας. Το λογισμικό της κάμερας ποσοτικοποιεί την τραχύτητα ως την κάθετη απόκλιση μιας πραγματικής επιφάνειας από την ιδανική της μορφή και την υπολογίζει ως ένα μέτρο χωρίς κάποια μονάδα μέτρησης.
4. Στις « Λεπτές γραμμές », ένας χάρτης ψευδοχρωμάτων όπου το σκούρο κόκκινο χρώμα υποδεικνύει τα βαθύτερα χαρακτηριστικά και το λευκό αντιστοιχεί στην κορυφή της επιδερμίδικης επιφάνειας.
5. Στις « Πτυχώσεις », ένας χάρτης παρόμοιος με τις Λεπτές Γραμμές, αποτυπώνοντας τα ίδια χαρακτηριστικά αλλά σε μεγαλύτερη κλίμακα.
6. Στις « Ανισοϋψώσεις », ένας χάρτης που απεικονίζει επάρματα ή κοιλώματα στο δέρμα. Η χαρτογράφηση ψευδοχρωμάτων κυμαίνεται από σκούρο κόκκινο για τις κοιλότητες, έως σκούρο βιολετί για τα επάρματα.

Το σχήμα 3.2 απεικονίζει και τις έξι μορφές εικόνων για την ίδια περιοχή του δέρματος. Όλες οι εικόνες εξήχθησαν χρησιμοποιώντας το λογισμικό της κάμερας, το οποίο παρήγαγε μια εικόνα ανάλυσης 480 × 480 εικονοστοιχείων, συμπιεσμένη σε μορφή JPEG.



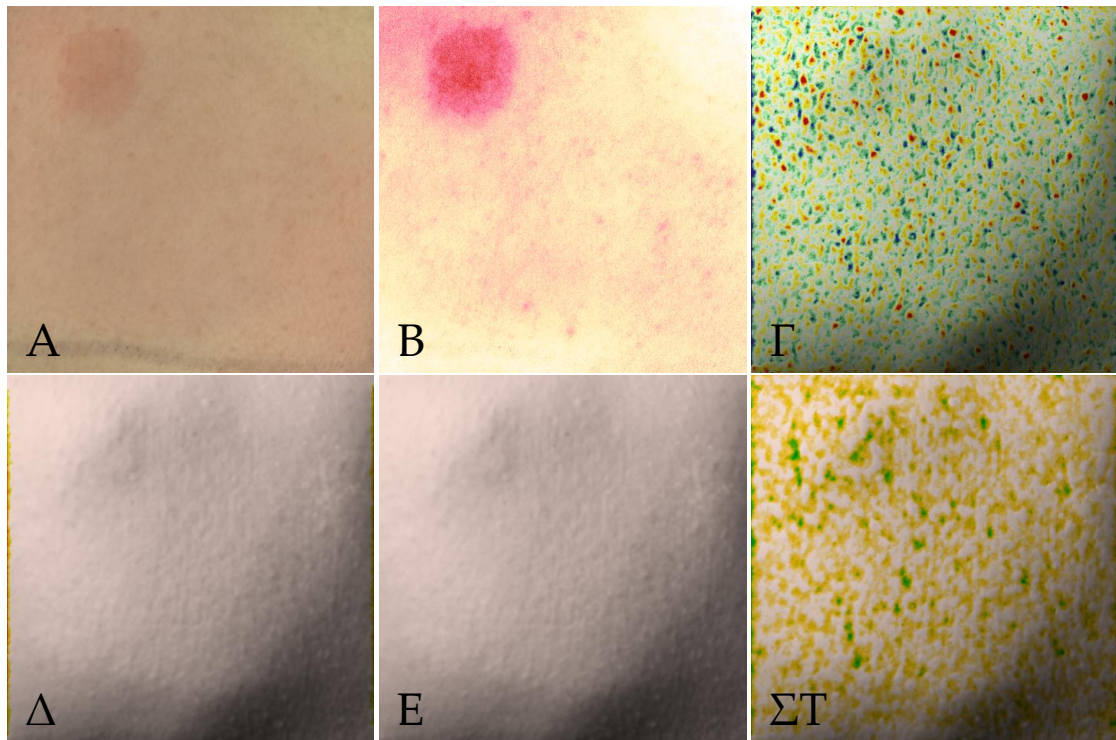
Σχήμα 3.1: Ένα ενιαίο επίθεμα χωρίζεται σε δύο περιοχές, την Περιοχή 1 και την Περιοχή 2. Κάθε περιοχή καταγράφεται ανεξάρτητα με την κάμερα λόγω του ότι το οπτικό της πεδίο περιορίζεται σε  $56 \times 56$  mm. Κάθε αριθμός αντιστοιχεί σε έναν θαλαμίσκο επιδερμικού επιθέματος.

### Επισημάνση εικόνων

Μετά τη διάγνωση από τον κλινικό ιατρό, εντοπίστηκαν τα αλλεργιογόνα στα οποία ο ασθενής βρέθηκε θετικός, και με χρήση του λογισμικού της κάμερας, οριοθετήθηκαν ψηφιακά τα σημεία του δέρματος τα οποία βρίσκονταν σε επαφή με το αλλεργιογόνο. Η διαδικασία αυτή πραγματοποιήθηκε με τη χρήση εικόνων της μορφής « Λεπτές Γραμμές » στις 48 ώρες, όπου τα όρια κάθε θαλαμίσκου ήταν ορατά στην επιφάνεια του επιδερμίδας. Το λογισμικό μετέφερε αυτόματα το ίδιο πλαίσιο οριοθέτησης σε όλα τα χρονικά σημεία (δηλ. 0 ώρες και 72 ώρες). Για να αποφευχθεί η σοβαρή ανισορροπία μεταξύ αρνητικών και θετικών δειγμάτων στο σύνολο δεδομένων, ο κλινικός ιατρός ανίχνευσε και οριοθέτησε επιπλέον ίσο αριθμό αλλεργιογόνων στα οποία ο ασθενής είχε αρνητική εξέταση. Τέλος, εντοπίστηκαν και οριοθετήθηκαν επίσης οι θαλαμίσκοι που περιείχαν βαζελίνη (δείγμα ελέγχου), καθώς και παράγοντες που έβαφαν το δέρμα, δηλαδή η παραφαινυλενοδιαμίνη και το μίγμα χρωστικών υφασμάτων. Συνολικά, για κάθε χρονική στιγμή ανιχνεύθηκαν και οριοθετήθηκαν 1579 αλλεργιογόνα, 501 από τα οποία διαγνώστηκαν ως θετικά.

### 3.3.2 Πειραματική διάταξη

Στην παρούσα μελέτη πραγματοποιήθηκαν τρία διαφορετικά πειράματα για τη διερεύνηση της ακρίβειας που μπορεί να επιτύχει ένα μοντέλο βαθιάς μάθησης στη δυαδική ταξινόμηση (αρνητικό/θετικό αποτέλεσμα) εικόνων επιδερμικών δοκιμασιών, με χρήση διαφορετικών τεχνικών προεπεξεργασίας και μορφών εικόνας. Στο πρώτο πείραμα (που περιγράφεται λεπτομερώς στην ενότητα 3.3.3), συγκρίθηκε η διαφορά στην απόδοση κατά τη χρήση τριών διαφορετικών μεθόδων.



Σχήμα 3.2: Δείγματα των διαφορετικών εικονομορφών για την ίδια περιοχή του δέρματος ενός ασθενούς. **(Α)** Έγχρωμη εικόνα, **(Β)** Ερύθημα, **(Γ)** Υφή, **(Δ)** Λεπτές Γραμμές, **(Ε)** Πτυχώσεις, **(ΣΤ)** Ανισοϋψώσεις.

δων προεπεξεργασίας στην απεικόνιση του ερυθήματος. Στο δεύτερο πείραμα (που περιγράφεται λεπτομερώς στην ενότητα 3.3.4), έγινε σύγκριση της επίδοσης που μπορεί να επιτευχθεί με χρήση κάθε μορφής εικόνας ξεχωριστά. Στο τρίτο και τελευταίο πείραμα (που περιγράφεται λεπτομερώς στην ενότητα 3.3.5), έγινε συνδυασμός των μορφών εικόνας για την αξιολόγηση της προστιθέμενης αξίας.

### Διαμόρφωση μοντέλου

Για όλα τα πειράματα, έγινε χρήση ενός Συνελικτικού Νευρωνικού Δικτύου (ΣΝΝ) EfficientNetB0 [16], προεκπαιδευμένο στο σύνολο δεδομένων ImageNet. Κάθε εκπαίδευση πραγματοποιήθηκε για 250 εποχές, χρησιμοποιώντας μέγεθος παρτίδας 8. Επιλέχθηκε ο βελτιστοποιητής AdamW [82], και ο ρυθμός μάθησής του ορίστηκε να μειώνεται με συνημιτονικό ρυθμό ανόπτωσης από  $3 \times 10^{-4}$  σε  $1 \times 10^{-5}$  κατά τη διάρκεια της πορείας εκπαίδευσης. Για να την επαύξηση της ποικιλομορφίας του συνόλου δεδομένων και για την αποφυγή της μεροληψίας του δικτύου όσον αφορά τη θέση του αλλεργιογόνου, χρησιμοποιήθηκε επαύξηση των δεδομένων με τη μορφή τυχαίας οριζόντιας και κάθετης αναστροφής, καθώς και τυχαίας περιστροφής εντός 20 μοιρών γύρω από το κέντρο των εικόνων. Όλες οι εικόνες κανονικοποιήθηκαν σε κάθε ένα από τα κανάλια RGB χρησιμοποιώντας τη μέση τιμή και την τυπική απόκλιση του συνόλου δεδομένων ImageNet, ενώ ταυτόχρονα έγινε χρήση της διγραμμικής παρεμβολής για την μετατροπή των

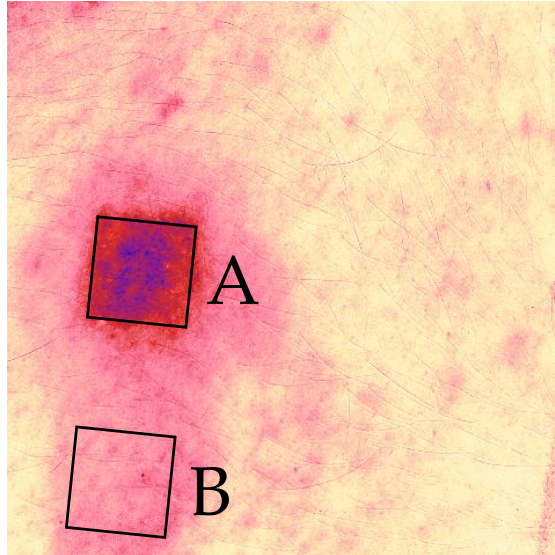
διαστάσεών τους σε  $224 \times 224$  εικονοστοιχεία.

### 3.3.3 Πρώτο πείραμα: Σύγκριση τεχνικών προεπεξεργασίας

Η προεπεξεργασία δεδομένων είναι πάντοτε ένα πολύ κρίσιμο πρώτο βήμα στις εργασίες βαθιάς μάθησης [106]. Αποσκοπεί στη μετατροπή των ακατέργαστων δεδομένων σε μια μορφή που είναι ευκολότερο για έναν αλγόριθμο βαθιάς μάθησης να επεξεργαστεί. Για παράδειγμα, είναι σύνηθες σε προβλήματα αναγνώρισης εικόνων, οι εικόνες να προσαρμόζονται σε ένα τυπικό μέγεθος και να μειώνεται η υπολογιστική πολυπλοκότητα μετατρέποντάς τις από έγχρωμες σε ασπρόμαυρες. Με αυτό τον τρόπο διατηρούνται σημαντικά χαρακτηριστικά όπως οι ακμές και οι υφές [107]. Η προεπεξεργασία δεδομένων μπορεί να έχει σημαντικό αντίκτυπο στην ακρίβεια και την απόδοση ενός μοντέλου βαθιάς μάθησης.

Κάθε εικόνα δέρματος στην παρούσα μελέτη απεικονίζει είτε τέσσερα είτε έξι μεμονωμένα σημεία εφαρμογής αλλεργιογόνου (Σχήμα 3.1). Ο απλούστερος τρόπος για τον περιορισμό της ταξινόμησης σε ένα μόνο αλλεργιογόνο τη φορά, είναι η περικοπή ή η κάλυψη της εικόνας ώστε να διατηρηθούν μόνο τα εικονοστοιχεία που αντιστοιχούν στην περιοχή ενδιαφέροντος (αγγλ. Region of Interest - ROI), στην προκειμένη περίπτωση την περιοχή ενός μεμονωμένου αλλεργιογόνου (Σχήμα 3.4 **B**). Πράγματι, οι Chan, Srivastava, Damaraju, *et al.* [96] επίσης περικόπουν τις εικόνες τους σε ένα μόνο αλλεργιογόνο στην πρόσφατη μελέτη τους για την αυτοματοποιημένη ανίχνευση δερματικών αντιδράσεων σε επιδερμιακές δοκιμασίες. Ωστόσο, είναι πιθανό ότι αυτή η προσέγγιση περικοπής ή κάλυψης σε μια συγκεκριμένη ROI απορρίπτει πολύτιμες πληροφορίες. Για παράδειγμα, η διαφορά στη μελάγχρωση μεταξύ του σημείου εφαρμογής του αλλεργιογόνου και της περιοχής του δέρματος που το περιβάλλει, αποτελεί καλύτερο δείκτη αντίδρασης από την απόλυτη τιμή της μελάγχρωσης στο αλλεργιογόνο. Επιπλέον, οι αντιδράσεις που προκαλούνται σε γειτονικούς θαλαμίσκους μπορεί να επηρεάσουν την ROI, με αποτέλεσμα να παραχθούν ψευδώς θετικά αποτελέσματα. Το σχήμα 3.3 απεικονίζει ένα τέτοιο παράδειγμα, όπου η αντίδραση στο αλλεργιογόνο A είναι τόσο ισχυρή που εξαπλώνεται και στο αλλεργιογόνο B. Εάν το δέρμα στο αλλεργιογόνο B εξεταζόταν μεμονωμένα, ένας αλγόριθμος ταξινόμησης θα ήταν πολύ πιθανό να προβλέψει εσφαλμένα ότι η εκδηλωθείσα μελάγχρωση είναι αποτέλεσμα αντίδρασης σε αυτό το αλλεργιογόνο.

Επομένως, για τη βελτίωση της απόδοσης ταξινόμησης προτείνεται η διατήρηση του συνόλου πληροφορίας της αρχικής εικόνας. Για το σκοπό αυτό, αντί για περικοπή της εικόνας στην ROI, εφαρμόστηκε μια ειδική μάσκα, για την οποία επινοήθηκε ο όρος «μάσκα  $\alpha$ » (αγγλ. alpha mask). Στα γραφικά υπολογιστών, η τιμή του  $\alpha$  αναφέρεται στο επίπεδο αδιαφάνειας ή διαφάνειας μιας εικόνας, με το 0%  $\alpha$  να είναι εντελώς διαφανές και το 100%  $\alpha$  να είναι εντελώς αδιαφανές. Μια μάσκα  $\alpha$  είναι μια εικόνα  $M$  εκφρασμένη σε κλίμακα του γκρι όπου κάθε τιμή εικονοστοιχείου αντιστοιχεί στην τιμή  $\alpha$  και κυμαίνεται από 0 έως 1. Η μάσκα



Σχήμα 3.3: Ένα παράδειγμα που αναδεικνύει γιατί η συνολική πληροφορία της εικόνας είναι σημαντική. Στο αλλεργιογόνο με την επισήμανση **A** έχει εκδηλωθεί πολύ έντονη αντίδραση. Δεδομένων των συνθηκών, εάν το αλλεργιογόνο **B** εξεταστεί μεμονωμένα, είναι πιθανό να προκύψει ψευδώς θετικό αποτέλεσμα αφού φαίνεται να έχει αυξημένη ερυθρότητα.

εφαρμόζεται με την εξίσωση 3.1:

$$I' = I \odot M \quad (3.1)$$

όπου  $I$  είναι η αρχική δερματική εικόνα και  $\odot$  είναι ένας τελεστής για τον πολλαπλασιασμό πινάκων στοιχείο προς στοιχείο. Όλες τις τιμές των εικονοστοιχείων στη μάσκα  $a$  τέθηκαν στο 0,5, εκτός από εκείνες που ανήκουν στην ROI, οι οποίες τέθηκαν στην τιμή 1. Με αυτόν τον τρόπο, αμβλύνεται η φωτεινότητα της εικόνας σε όλα τα σημεία εκτός από την ROI, τονίζοντας ουσιαστικά την περιοχή που πρόκειται να ταξινομηθεί, αλλά διατηρώντας την συνολική πληροφορία της εικόνας (Σχήμα 3.4 **Γ**). Στο εξής, αυτή η μορφή προεπεξεργασίας θα αναφέρεται ως «MA» (Μάσκα  $a$ ).

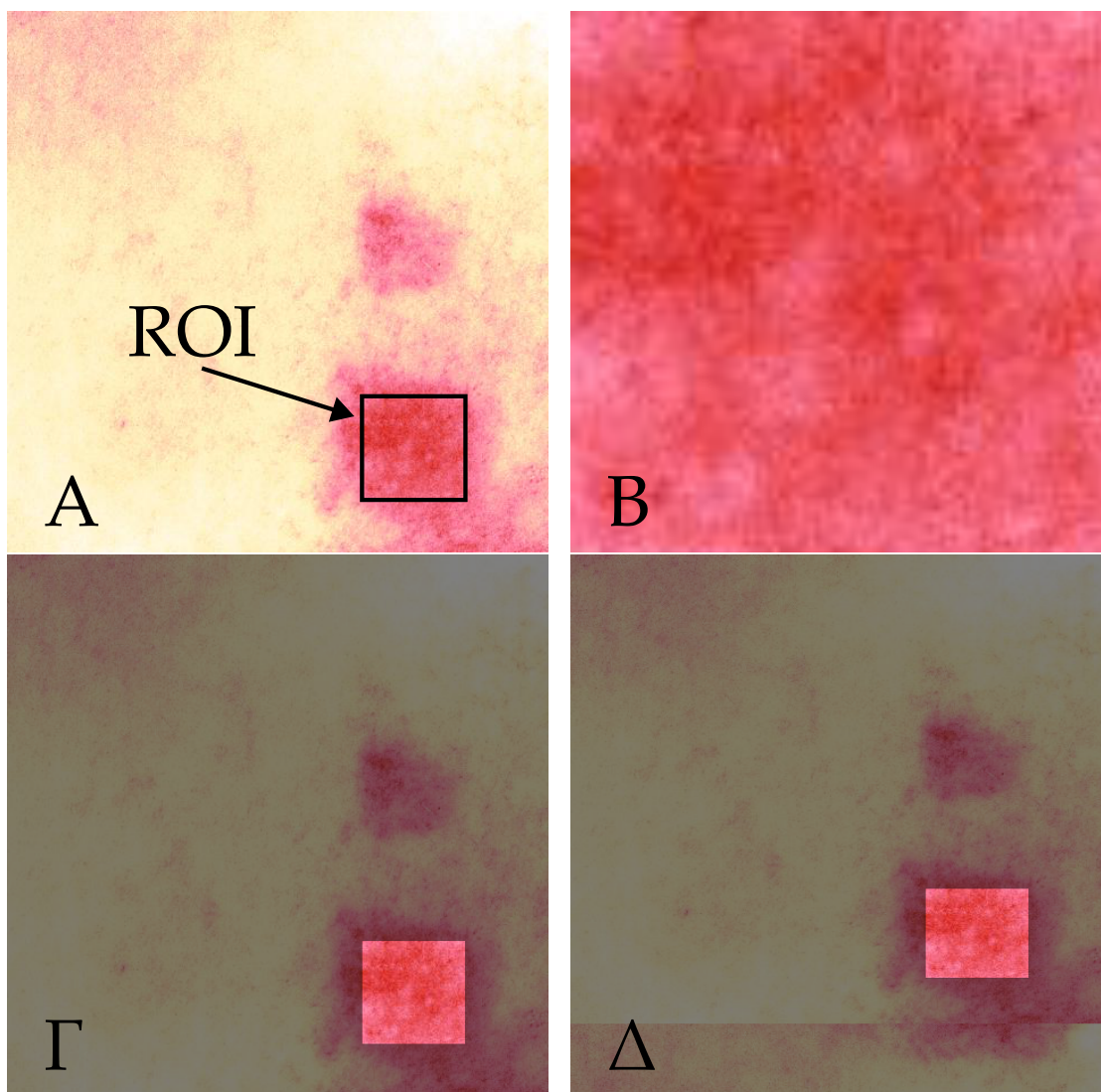
Εκτός από τη διατήρηση της πληροφορίας της συνολικής εικόνας με χρήση MA, μπορεί κανείς να παρατηρήσει ότι τα αλλεργιογόνα της τελευταίας σειράς στην περιοχή 1, είναι γειτονικά με τα αλλεργιογόνα της πρώτης σειράς στην περιοχή 2 (Σχήμα 3.1). Ωστόσο, δεδομένου ότι κάθε περιοχή αντιστοιχεί σε διαφορετική εικόνα, αυτή η πληροφορία γεινίασης χάνεται. Για την αντιμετώπιση αυτού του ζητήματος, η πληροφορία κάθε εξεταζόμενης εικόνας επεκτείνεται ώστε ενσωματώσει εικονοστοιχεία από τη γειτονική της περιοχή. Για τις εικόνες που αντιστοιχούν στην περιοχή 1, περικόπτονται 75 εικονοστοιχεία από την αντίστοιχη περιοχή 2 και ενώνονται στο κάτω μέρος της εικόνας. Για τις εικόνες που αντιστοιχούν στην περιοχή 2, περικόπτονται 75 εικονοστοιχεία από την περιοχή 1 και ενώνονται στην κορυφή της εικόνας. Σε αντιστοιχία με την προηγούμενη τεχνική, για την προεπεξεργασία αυτή επινοήθηκε ο όρος «Επαυξημένη Μάσκα  $a$ » (EMA, αγγλ.

Contextually Enhanced Alpha Mask - CEAM). Το σχήμα 3.4  $\Delta$  απεικονίζει τις εικόνες που προέκυψαν με τη χρήση αυτής της τεχνικής προεπεξεργασίας.

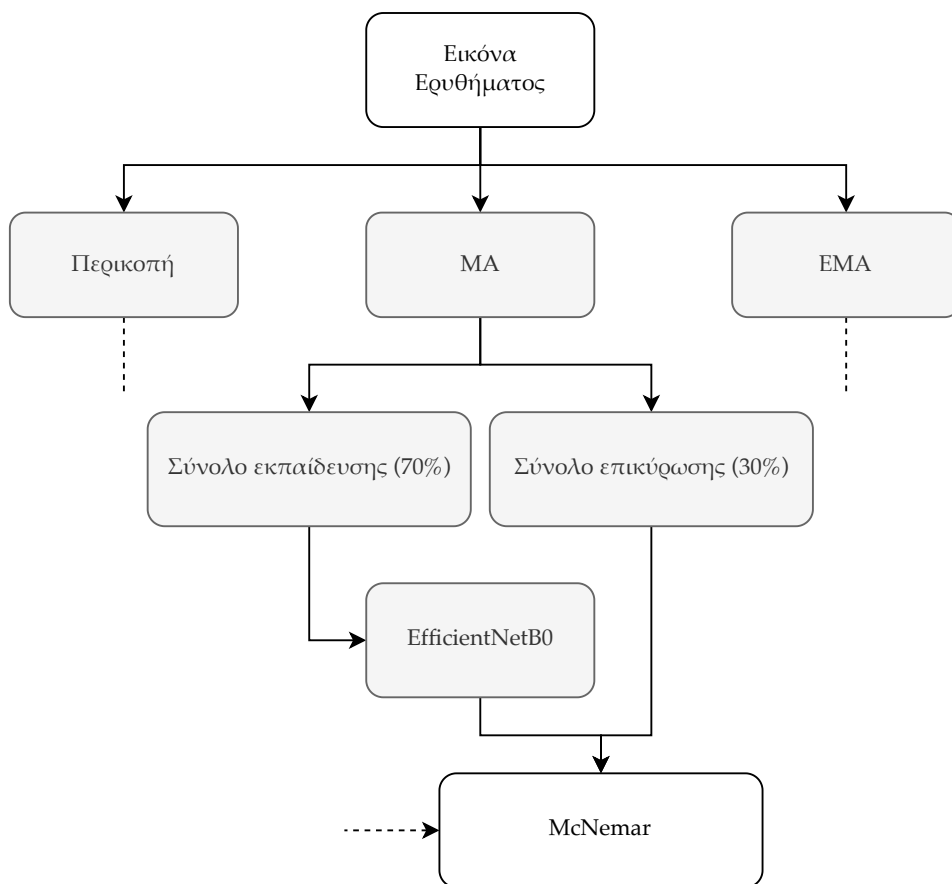
Η αξιολόγηση των τριών συστημάτων προεπεξεργασίας περιλάμβανε την εκπαίδευση ενός SNN EfficientNetB0 με τη διαμόρφωση που περιγράφεται στην ενότητα 3.3.2. Το σύνολο δεδομένων διαχωρίστηκε σε σύνολα εκπαίδευσης και επικύρωσης χρησιμοποιώντας διαχωρισμό 70/30, διασφαλίζοντας ταυτοχρόνως ότι δεν υπήρχε αλληλοεπικάλυψη μεταξύ των ασθενών. Χρησιμοποιήθηκαν μόνο εικόνες που ελήφθησαν 72 ώρες μετά την έκθεση στα αλλεργιογόνα. Στη συνέχεια, το δίκτυο εκπαιδεύτηκε και αξιολογήθηκε τρεις φορές, μία φορά για κάθε σύστημα προεπεξεργασίας, χρησιμοποιώντας ακριβώς τον ίδιο διαχωρισμό δεδομένων. Αν και το πρόβλημα είναι στην πραγματικότητα πρόβλημα ταξινόμησης πολλαπλών κατηγοριών λόγω του ότι σε κάθε αλλεργιογόνο αποδίδεται μια ετικέτα ανάλογα με την αντίστοιχη αντίδραση (καμία, ασθενής, έντονη, πολύ έντονη), για την απλούστευση του προβλήματος επιλέχθηκε η μετατροπή του σε δυαδική ταξινόμηση. Έτσι, οι ετικέτες συμπιύχθηκαν σε δυαδικές τιμές θεωρώντας την ετικέτα « καμία » ως την « αρνητική » ετικέτα και τις υπόλοιπες ως την « θετική » ετικέτα. Για κάθε σύστημα, υπολογίστηκε η βαθμολογία  $F1$ , η ακρίβεια, η ειδικότητα, η ανάκληση και η ακρίβεια, όπως ορίστηκαν στις ενότητες 2.3.4 και 2.3.5. Για την διερεύνηση ύπαρξης στατιστικά σημαντικής διαφοράς μεταξύ των τριών τεχνικών προεπεξεργασίας, εφαρμόστηκε το τεστ McNemar στα αποτελέσματα επικύρωσης (Σχήμα 3.5).

### 3.3.4 Δεύτερο πείραμα: Σύγκριση εικονομορφών

Το δεύτερο πείραμα επικεντρώνεται στη σύγκριση των επιδόσεων δυαδικής ταξινόμησης του EfficientNetB0 σε κάθε μία διαφορετική μορφή εικόνας, όπως αυτές αναλύθηκαν στην ενότητα 3.3.1. Για την προεπεξεργασία των εικόνων χρησιμοποιήθηκε η τεχνική της EMA. Για την απόκτηση μιας ακριβέστερης εκτίμησης των μετρικών απόδοσης και την περαιτέρω αξιολόγηση της σταθερότητας σε διαφορετικά σύνολα δεδομένων, το μοντέλο εκπαιδεύτηκε και αξιολογήθηκε χρησιμοποιώντας πενταπλή διασταυρούμενη επικύρωση. Ωστόσο, δεν πραγματοποιήθηκαν στατιστικές δοκιμές στα αποτελέσματα της διασταυρούμενης επικύρωσης, καθώς όλες οι δοκιμές απαιτούν ανεξάρτητα και πανομοιότυπα καταναμημένα (αγγλ. Independent and Identically Distributed - IID) δείγματα. Αυτό δεν ισχύει στην περίπτωση της διασταυρούμενης επικύρωσης, καθώς τα σύνολα εκπαίδευσης επικαλύπτονται μεταξύ των διαμερίσεων. Συγκεκριμένα, κάθε ζεύγος διαμερίσεων από το σύνολο εκπαίδευσης μοιράζεται  $1 - \frac{2}{K}$  των δεδομένων, όπου  $K$  ο αριθμός των διαμερίσεων. Επομένως, αυτή η επικάλυψη εμποδίζει το στατιστικό τεστ από το να αποκτήσει μια καλή εκτίμηση της διακύμανσης [108]. Ενώ η δοκιμή διασταυρούμενης επικύρωσης  $5 \times 2$  έχει προταθεί ως μια καλή εναλλακτική λύση [108], αυτή περιλαμβάνει πέντε διασταυρούμενες επικυρώσεις 2 φορές, και είναι μη βέλτιστη για την επίτευξη αξιόπιστων εκτιμήσεων επιδόσεων σε μικρά σύνολα δεδομένων

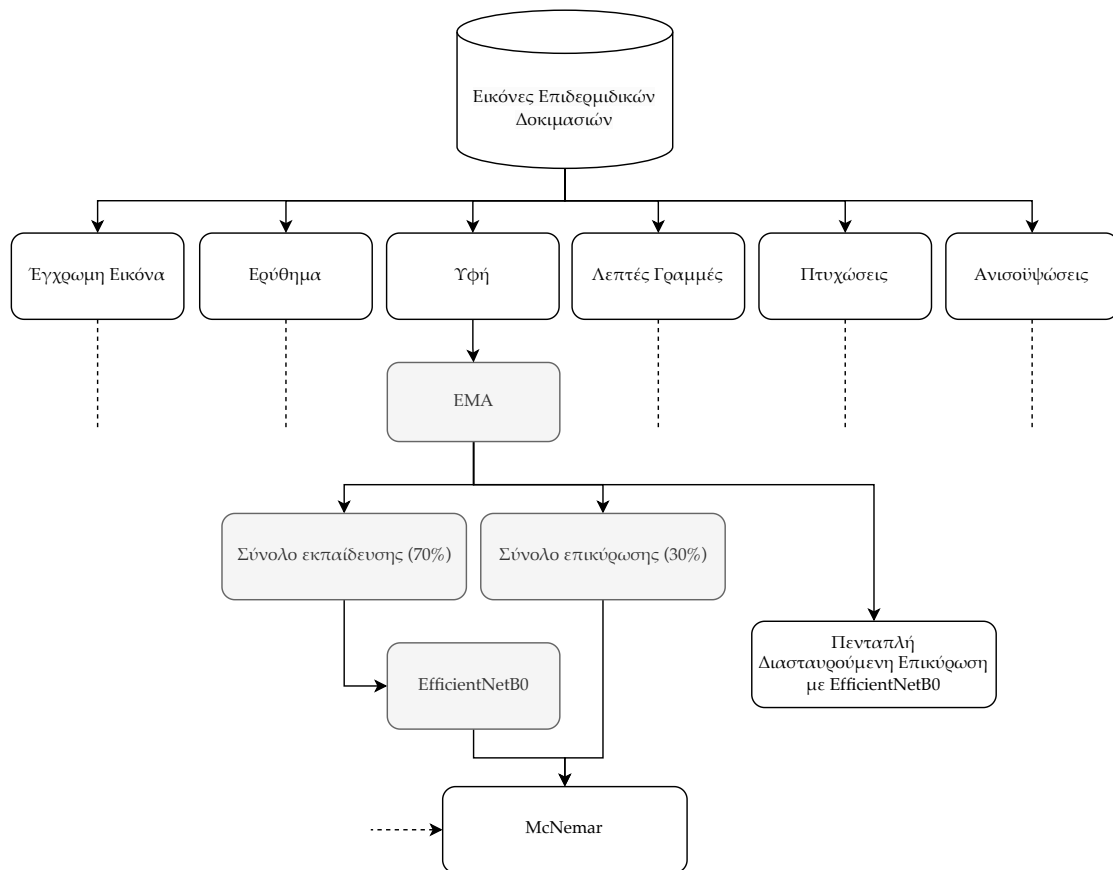


Σχήμα 3.4: Το αποτέλεσμα που προκύπτει για κάθε μία από τις τρεις διαφορετικές τεχνικές προεπεξεργασίας. **(Α)** η αρχική εικόνα με επισημασμένη την περιοχή ενδιαφέροντος (ROI), **(Β)** η εικόνα περικόπεται στην ROI, **(Γ)** MA - η αρχική εικόνα με εφαρμογή Μάσκας  $a$  **(Δ)** EMA - η εικόνα με Μάσκα  $a$  και επιπλέον προσθήκη πρόσθετων πληροφοριών από την παρακείμενη περιοχή του επιδερμικού επιθέματος.



Σχήμα 3.5: Διαγραμματική αναπαράσταση της πειραματικής διαδικασίας για την αξιολόγηση των τριών διαφορετικών τεχνικών προεπεξεργασίας των εικόνων.





Σχήμα 3.6: Διαγραμματική αναπαράσταση της πειραματικής διαδικασίας του δεύτερου πειράματος. Σε αυτό το πείραμα πραγματοποιείται σύγκριση της επίδοσης που μπορεί να επιτευχθεί με χρήση κάθε εικονομορφής ξεχωριστά.

λόγω i) της μείωσης του συνόλου εκπαίδευσης περισσότερο από όσο χρειάζεται, ii) του ότι τα δεδομένα εκπαίδευσης πιθανόν να διαφέρουν σημαντικά από τα δεδομένα επικύρωσης. Ως εκ τούτου, επιλέχθηκε και πάλι το τεστ McNemar για την αξιολόγηση της στατιστικής σημαντικότητας, ενώ έγινε χρήση πενταπλών διασταυρούμενων επικυρώσεων για την αξιολόγηση της απόδοσης του μοντέλου. Για την εκτέλεση του τεστ McNemar χρησιμοποιήθηκε διαχωρισμός 70/30 μεταξύ εκπαίδευσης και επικύρωσης. Στη συνέχεια το μοντέλο εκπαιδεύτηκε και αξιολογήθηκε εκ νέου για τον προσδιορισμό της στατιστικής σημαντικότητας (Σχήμα 3.6).

### 3.3.5 Τρίτο πείραμα: Σύντηξη εικονομορφών

Στο τρίτο πείραμα διερευνήθηκε κατά πόσον ο συνδυασμός διαφορετικών μορφών εικόνων θα μπορούσε να βελτιώσει τη συνολική απόδοση ταξινόμησης. Διαισθητικά, παρόλο που δύο μορφές εικόνων μπορεί να παρουσιάζουν παρόμοιες επιδόσεις στα προηγούμενα πειράματα, το ΣNN ενδέχεται να μαθαίνει χαρακτηριστικά τα οποία να είναι συμπληρωματικά μεταξύ τους, και τελικά να δύναται να αυξήσουν περαιτέρω την ακρίβεια ταξινόμησης όταν αυτά συνδυάζονται.

Για την αποφυγή αλλαγής της δομής του προ-εκπαιδευμένου ΣNN Efficient-

		Προβλέψεις ταξινομητή 2	
		Αληθείς	Ψευδείς
Προβλέψεις ταξινομητή 1	Αληθείς	$a$	$b$
	Ψευδείς	$c$	$d$

Πίνακας 3.2: Παράδειγμα πίνακα ενδεχομένων για τον στατιστικό έλεγχο McNemar.

Net, με αποτέλεσμα την εξαγωγή διαφορετικών χαρακτηριστικών, σε αυτό το πείραμα έγινε επαναχρησιμοποίηση των εκπαιδευμένων μοντέλων από την ενότητα 3.3.4. Για τον σκοπό αυτό, για κάθε εικόνα στο σύνολο επικύρωσης, έγινε εξαγωγή των χαρακτηριστικών που το ΣNN έμαθε να υπολογίζει, απομονώνοντας τις εξόδους του επιπέδου το οποίο προηγείται του τελικού πλήρως συνδεδεμένου επιπέδου του δικτύου. Γενικά, στα ΣNN που επιτελούν ταξινόμηση, το τελικό πλήρως συνδεδεμένο επίπεδο είναι ο ταξινομητής. Τα επίπεδα που προηγούνται του ταξινομητή μαθαίνουν τα βέλτιστα χαρακτηριστικά που θα χρησιμοποιηθούν στην τελική ταξινόμηση. Επομένως, παραλείποντας το τελευταίο επίπεδο, η έξοδος του δικτύου παίρνει τη μορφή των χαρακτηριστικών αυτών αντί για την τελική πρόβλεψη. Στη συνέχεια, το σύνολο των χαρακτηριστικών από κάθε διαφορετική εικονομορφή οργανώνονται σε ζεύγη και εκπαιδεύονται με τη χρήση μιας Μηχανής Υποστήριξης Διανύσματος (αγγλ. Support Vector Machine - SVM) με πυρήνα την ακτινική συνάρτηση (αγγλ. Radial Basis Function - RBF), θέτοντας ευρετικά την παράμετρο C στο 3,5. Δεδομένου ότι το σύνολο επικύρωσης που αποτελείται από το 30% των δεδομένων είναι πολύ μικρό την εφαρμογή ενός επιπλέον διαχωρισμού 70/30 σε αυτό, επιλέχθηκε αντ' αυτού η διασταυρούμενη επικύρωση. Για τον προσδιορισμό της στατιστικής σημαντικότητας των αποτελεσμάτων, ακολούθηθηκε η συνδυασμένη δοκιμή διασταυρούμενης επικύρωσης F-test  $5 \times 2$  που περιγράφεται από Alpaydın [109] (Σχήμα 3.7).

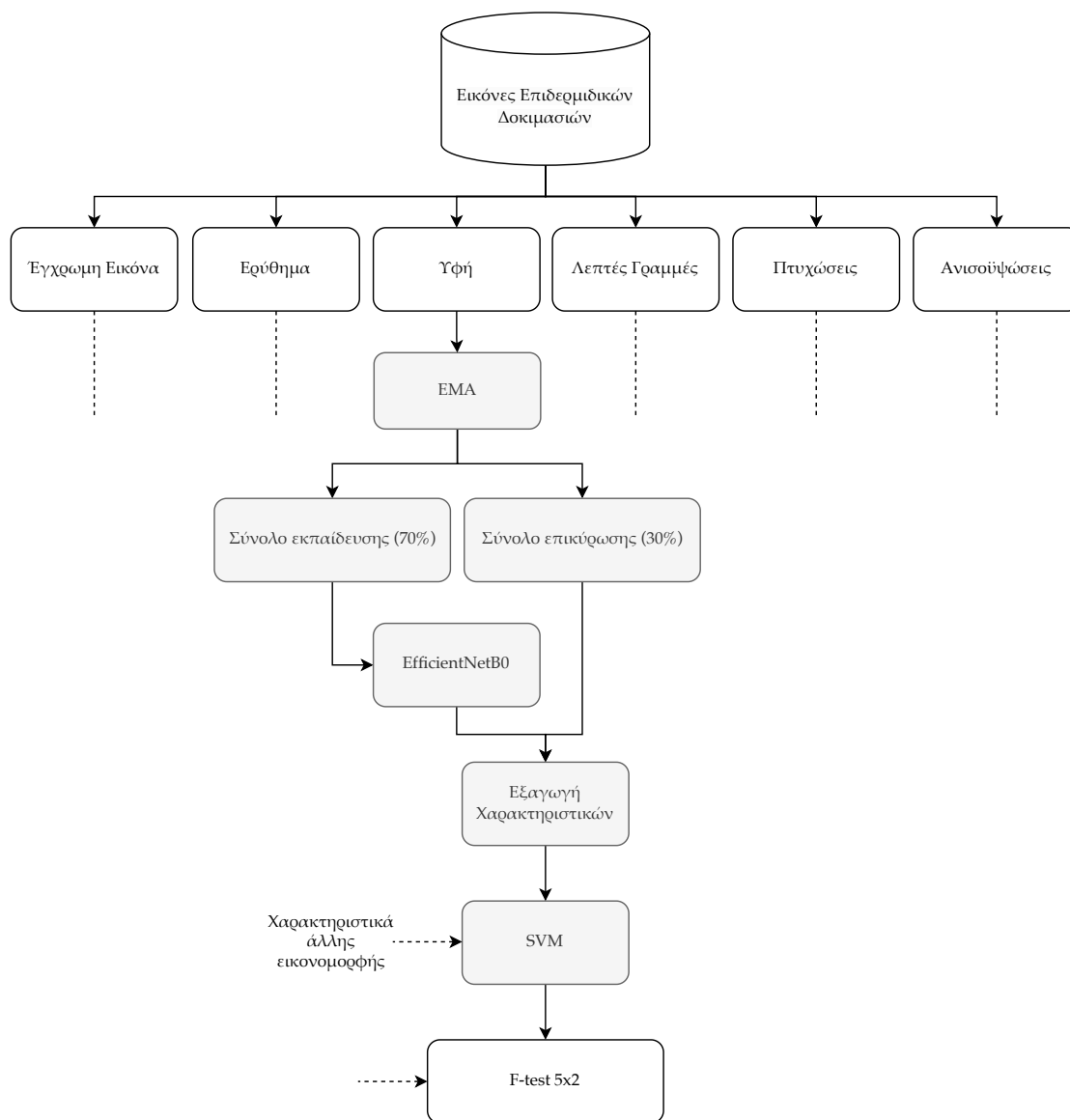
### 3.3.6 Στατιστικός έλεγχος McNemar

Το τεστ του McNemar είναι ένας έλεγχος της μηδενικής υπόθεσης ότι το ποσοστό των ζευγών προβλέψεων με ασύμφωνο αποτέλεσμα από δύο ταξινομητές (δηλαδή, η μία είναι αληθής και η άλλη ψευδής), είναι ίσο με 0,5 (δηλαδή, οι εν λόγω ταξινομητές είναι εξίσου ακριβείς). Ο έλεγχος βασίζεται στην κατανομή  $\chi^2$  και απαιτεί έναν πίνακα ενδεχομένων των ζευγών δεδομένων, όπως φαίνεται στον πίνακα 3.2.

Για τον υπολογισμό του στατιστικού ελέγχου McNemar υπολογίζεται το εξής:

$$\chi^2 = \frac{(b - c)^2}{(b + c)} \quad (3.2)$$

Υπό τη μηδενική υπόθεση της όμοιας ακρίβειας των ταξινομητών, αυτός ο στα-



Σχήμα 3.7: Διαγραμματική αναπαράσταση της πειραματικής διαδικασίας του τρίτου πειράματος. Σε αυτό το πείραμα πραγματοποιείται σύγκριση της επίδοσης που μπορεί να επιτευχθεί με κάθε συνδυασμό ανά δύο όλων των εικονομορφών.

τιστικός έλεγχος ακολουθεί κατανομή  $\chi^2$  με έναν βαθμό ελευθερίας. Αυτή η κατανομή μπορεί να χρησιμοποιηθεί για τον υπολογισμό μιας τιμής  $p$  για τον έλεγχο, με στόχο την εκτίμηση εάν η διαφορά στην ακρίβεια μεταξύ των ταξινομητών είναι στατιστικά σημαντική. Συνήθως, απαιτείται  $p < 0,05$ .

### 3.3.7 F-test 5×2

Το 5×2  $F$ -test, το οποίο προτάθηκε από τον E. Alpaydin [109], είναι μια στατιστική διαδικασία που έχει σχεδιαστεί για να παρέχει μια πιο ισχυρή εναλλακτική λύση στη διαδικασία 5×2  $t$ -test του Dietterich [108]. Ο στατιστικός αυτός έλεγχος έχει αναπτυχθεί ειδικά για την αξιολόγηση και τη σύγκριση της απόδοσης αλγορίθμων μηχανικής μάθησης.

Στην προσέγγιση του Dietterich, η απόδοση διαφορετικών αλγορίθμων αξιολογείται με την εκτέλεση ενός  $t$ -test στα ζεύγη αποτελεσμάτων όπως προκύπτουν από την 5×2 διασταυρούμενη επικύρωση. Ωστόσο, αυτή η μέθοδος εξαρτάται από τη σειρά των αναδιπλώσεων - εφόσον αυτή αλλάξει, τότε αλλάζουν και τα αποτελέσματα της δοκιμής. Για το λόγο αυτό, στο [109] προτάθηκε η συνδυασμένη 5×2 δοκιμή που χρησιμοποιεί το στατιστικό  $F$ -δοκιμής αντί του στατιστικού  $t$ -test. Πιο συγκεκριμένα:

Ας θεωρήσουμε δύο ταξινομητές,  $A$ , και  $B$ . Πραγματοποιούμε διπλή διασταυρούμενη επικύρωση 5 φορές. Σε κάθε επανάληψη, τα δεδομένα χωρίζονται σε δύο σύνολα ίσου μεγέθους,  $S_1$  και  $S_2$ . Κάθε ταξινομητής εκπαιδεύεται σε κάθε σύνολο, και επικυρώνεται στο άλλο. Ως εκ τούτου, λαμβάνονται οι εξής εκτιμήσεις σφάλματος:  $p_A^{(1)}$ ,  $p_B^{(1)}$ ,  $p_A^{(2)}$ , και  $p_B^{(2)}$ . Η διαφορά απόδοσης υπολογίζεται ως:

$$p^{(1)} = p_A^{(1)} - p_B^{(1)} \quad (3.3)$$

$$p^{(2)} = p_A^{(2)} - p_B^{(2)} \quad (3.4)$$

Στη συνέχεια υπολογίζεται η μέση τιμή  $\bar{p}$  και η διακύμανση  $s^2$  των διαφορών:

$$\bar{p} = \frac{p^{(1)} + p^{(2)}}{2} \quad (3.5)$$

$$s^2 = (p^{(1)} - \bar{p})^2 + (p^{(2)} - \bar{p})^2 \quad (3.6)$$

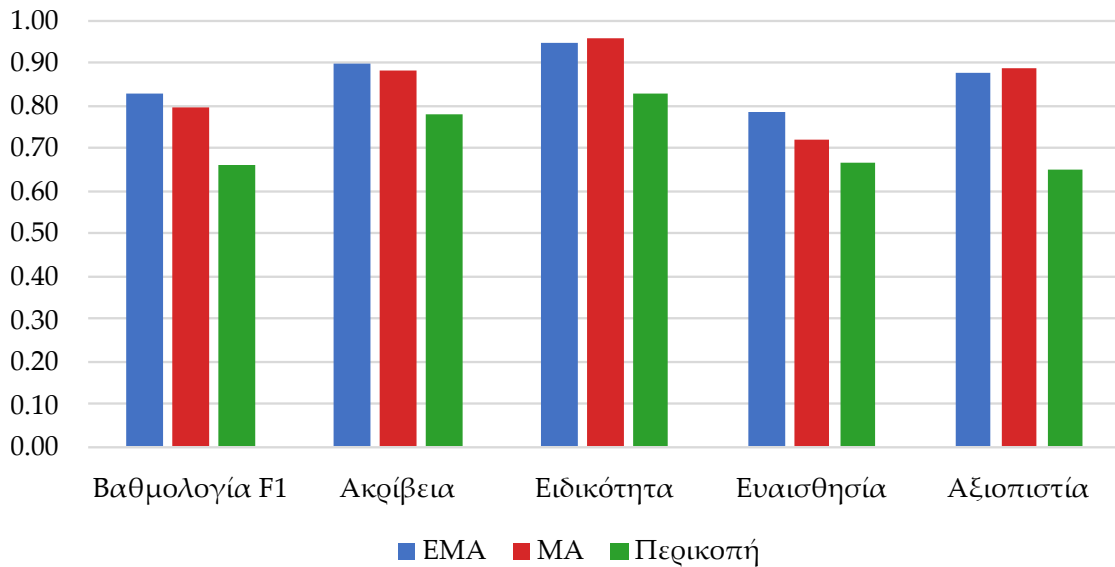
Το στατιστικό  $F$  υπολογίζεται ως:

$$f = \frac{\sum_{i=1}^5 \sum_{j=1}^2 (p_i^{(j)})^2}{2 \sum_{i=1}^5 s_i^2} \quad (3.7)$$

το οποίο είναι κατά προσέγγιση  $F$  κατανομή με 10 και 5 βαθμούς ελευθερίας.

Πίνακας 3.3: Μετρικές των αποτελεσμάτων για κάθε τεχνική προεπεξεργασίας.

Τεχνική προπεξ.	Βαθμ. F1	Ακρίβεια	Ειδικότητα	Ευαισθησία	Αξιοπιστία
EMA	0,83	0,90	0,95	0,79	0,87
MA	0,79	0,88	0,96	0,72	0,89
Περικοπή	0,66	0,78	0,83	0,67	0,65



Σχήμα 3.8: Ραβδόγραμμα αποτελεσμάτων για τις τρεις διαφορετικές τεχνικές προεπεξεργασίας.

### 3.4 Αποτελέσματα

#### 3.4.1 Πρώτο πείραμα: Σύγκριση τεχνικών προεπεξεργασίας

Ο πίνακας 3.3 παρουσιάζει τη βαθμολογία F1, την ακρίβεια, την ειδικότητα, την ευαισθησία και την αξιοπιστία που επιτεύχθηκαν για κάθε τεχνική προεπεξεργασίας. Το σχήμα 3.8 απεικονίζει ένα ραβδόγραμμα με τις ίδιες μετρικές για ευκολότερη οπτική σύγκριση.

Τα αποτελέσματα δείχνουν ότι η υψηλότερη βαθμολογία F1, ακρίβεια και ευαισθησία, επιτεύχθηκε με την τεχνική προεπεξεργασίας της EMA. Αντίθετα, η υψηλότερη ειδικότητα και αξιοπιστία επιτεύχθηκε με την τεχνική προεπεξεργασίας της απλής MA. Η απλή περικοπή της ROI της εικόνας είχε σημαντικά χειρότερες επιδόσεις σε όλες τις μετρικές. Το τεστ McNemar (Πίνακας 3.4) επιβεβαίωσε ότι υπήρχε στατιστικά σημαντική διαφορά ( $p < 0,05$ ) μεταξύ των δύο τεχνικών προεπεξεργασίας EMA - Περικοπή και MA - Περικοπή. Ωστόσο, δεν ισχύει το ίδιο για την EMA - MA, καθώς η τιμή  $p$  0,281 δεν επιτρέπει την ασφαλή υπόθεση ότι η μία τεχνική υπερισχύει της άλλης.

Πίνακας 3.4: Τιμές  $p$  που αντιστοιχούν στη σύγκριση ανά ζεύγη μεταξύ των διαφορετικών τεχνικών προπεξεργασίας, υπολογισμένες με χρήση του τεστ McNemar.

Υπόθεση	Τιμή $p$
EMA - MA	0,281
EMA - Περικοπή	0,0
MA - Περικοπή	0,0

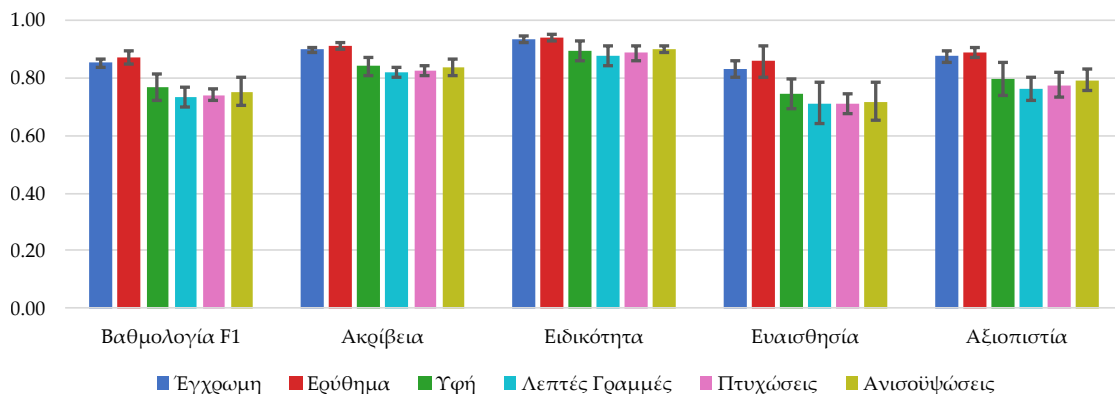
Πίνακας 3.5: Μετρικές αποτελεσμάτων για κάθε μορφή εικόνας.

Εικονομορφή	Βαθμ. F1	Ακρίβεια	Ειδικότητα	Ευαισθησία	Αξιοπιστία
Έγχρωμη	0,85 ± 0,01	0,90 ± 0,01	0,94 ± 0,01	0,83 ± 0,03	0,88 ± 0,02
Ερύθημα	0,87 ± 0,02	0,91 ± 0,01	0,94 ± 0,01	0,86 ± 0,06	0,89 ± 0,02
Υφή	0,77 ± 0,04	0,84 ± 0,03	0,89 ± 0,04	0,74 ± 0,05	0,79 ± 0,06
Λεπτές Γραμμές	0,73 ± 0,03	0,82 ± 0,02	0,88 ± 0,03	0,71 ± 0,07	0,76 ± 0,04
Πτυχώσεις	0,74 ± 0,02	0,83 ± 0,01	0,89 ± 0,03	0,71 ± 0,03	0,78 ± 0,04
Ανισοϋψώσεις	0,75 ± 0,05	0,84 ± 0,03	0,90 ± 0,01	0,72 ± 0,07	0,79 ± 0,04

### 3.4.2 Δεύτερο πείραμα: Σύγκριση εικονομορφών

Ο πίνακας 3.5 παρουσιάζει τη μέση τιμή ± τυπική απόκλιση της βαθμολογίας F1, της ακρίβειας, της ειδικότητας, της ευαισθησίας και της αξιοπιστίας που επιτεύχθηκαν για κάθε μορφή εικόνας κατά την πενταπλή διασταυρούμενη επικύρωση. Συνολικά, τα αποτελέσματα υποδεικνύουν ότι η μορφή Ερυθήματος υπερτερεί έναντι όλων των άλλων εικονομορφών όσον αφορά τη μέση τιμή σε όλες τις μετρικές. Η μορφή της απλής Έγχρωμης εικόνας πέτυχε επίσης υψηλές επιδόσεις και βρέθηκε να είναι παρόμοια με το Ερύθημα. Όλες οι υπόλοιπες λειτουργίες οδήγησαν σε σημαντικά χαμηλότερες επιδόσεις. Το σχήμα 3.9 παρέχει μια οπτική αναπαράσταση των αποτελεσμάτων.

Η μεθοδολογία όπως της ενότητας 3.3.3 χρησιμοποιήθηκε εκ νέου για τον εντοπισμό στατιστικά σημαντικών διαφορών μεταξύ των επιδόσεων που επιτυγχάνονται κατά τη χρήση κάθε μορφής εικόνας. Όπως φαίνεται στον πίνακα 3.6, η



Σχήμα 3.9: Σύγκριση μετρικών για όλες τις διαφορετικές μορφές εικόνας. Κάθε τιμή υπολογίζεται ως ο μέσος όρος σε μια πενταπλή διασταυρούμενη επικύρωση. Η γραμμή σφάλματος αντιστοιχεί στην τυπική απόκλιση.

Πίνακας 3.6: Τιμές  $p$  που αντιστοιχούν στη σύγκριση κατά ζεύγη των διαφορετικών μορφών απεικόνισης, υπολογισμένες με τη χρήση του τεστ McNemar. Οι λαμβανόμενες τιμές  $p$  αντιστοιχούν στον πρόσθετο διαχωρισμό 70/30 και όχι στα αποτελέσματα της διασταυρούμενης επικύρωσης που απεικονίζονται στο Σχήμα 3.9.

Υπόθεση	Τιμή $p$
Έγχρωμη - Ερύθημα	0,760
Έγχρωμη - Υφή	<b>0,001</b>
Έγχρωμη - Ανισοϋψώσεις	<b>0,002</b>
Έγχρωμη - Πτυχώσεις	<b>0,000</b>
Έγχρωμη - Λεπτές Γραμμές	<b>0,001</b>
Ερύθημα - Υφή	<b>0,000</b>
Ερύθημα - Ανισοϋψώσεις	<b>0,001</b>
Ερύθημα - Πτυχώσεις	<b>0,000</b>
Ερύθημα - Λεπτές Γραμμές	<b>0,000</b>
Υφή - Ανισοϋψώσεις	0,890
Υφή - Πτυχώσεις	0,061
Υφή - Λεπτές Γραμμές	0,327
Ανισοϋψώσεις - Πτυχώσεις	<b>0,045</b>
Ανισοϋψώσεις - Λεπτές Γραμμές	0,255
Πτυχώσεις - Λεπτές Γραμμές	0,942

εύρεση στατιστικά σημαντικής διαφοράς μεταξύ των μορφών Έγχρωμης εικόνας - Ερυθήματος, καθώς και Υφής - Ανισοϋψώσεων, Υφής - Πτυχώσεων, Υφής - Λεπτών Γραμμών, Ανισοϋψώσεων - Λεπτών Γραμμών, και Πτυχώσεων - Λεπτών Γραμμών, δεν ήταν εφικτή ( $p \geq 0,05$ ).

### 3.4.3 Τρίτο πείραμα: Σύντηξη εικονομορφών

Στο τελευταίο πείραμα, διερευνήθηκε εάν το μοντέλο βαθιάς μάθησης μάθαινε συμπληρωματικά χαρακτηριστικά μεταξύ δύο διαφορετικών μορφών εικόνας. Για τον σκοπό αυτό, χρησιμοποιήθηκε ένα σχήμα διασταυρούμενης επικύρωσης  $5 \times 2$  και υπολογίστηκε η βαθμολογία  $F1$  για κάθε εικονομορφή στον πίνακα 3.7. Πραγματοποιήθηκε επίσης ένας έλεγχος  $F$ -test όλων των πιθανών ζευγών χαρακτηριστικών μεταξύ της έγχρωμης μορφής εικόνας και των υπολοίπων, με τα αποτελέσματα να παρουσιάζονται στον Πίνακα 3.8.

Τα ευρήματά αναδεικνύουν ότι τα καλύτερα αποτελέσματα επιτεύχθηκαν με το συνδυασμό των χαρακτηριστικών που εξήχθησαν από την απλή Έγχρωμη εικόνα και το Ερύθημα. Αυτή η προσέγγιση παρουσίασε στατιστικά σημαντική διαφορά σε σύγκριση με τη χρήση μόνο της Έγχρωμης εικόνας ( $p < 0,05$ ). Ωστόσο, ο συνδυασμός είτε της Έγχρωμης εικόνας είτε του Ερυθήματος με άλλες λειτουργίες δεν απέδωσε καλύτερα αποτελέσματα από τη χρήση αποκλειστικά της Έγχρωμης εικόνας. Επιπλέον, τα αποτελέσματα ήταν κατώτερα για το συνδυασμό χαρακτηριστικών που περιλάμβαναν μόνο τις Ανισοϋψώσεις, την Υφή, τις Λεπτές Γραμμές ή τις Πτυχώσεις, καθώς παρήγαγαν συγκρίσιμα αποτελέσματα με την απλή Έγ-

Πίνακας 3.7: Βαθμολογία  $F1$  όπως υπολογίστηκε από τη διασταυρούμενη επικύρωση  $5 \times 2$  σε κάθε συνδυασμό εικονομορφών.

<b>Εικονομορφή</b>	<b>Βαθμολογία <math>F1</math></b>
Έγχρωμη	$0,70 \pm 0,04$
Έγχρωμη + Ερύθημα	$0,74 \pm 0,04$
Έγχρωμη + Ανισοϋψώσεις	$0,70 \pm 0,05$
Έγχρωμη + Υφή	$0,68 \pm 0,05$
Έγχρωμη + Λεπτές Γραμμές	$0,69 \pm 0,04$
Έγχρωμη + Πτυχώσεις	$0,69 \pm 0,04$
Ερύθημα + Ανισοϋψώσεις	$0,68 \pm 0,03$
Ερύθημα + Υφή	$0,71 \pm 0,02$
Ερύθημα + Λεπτές Γραμμές	$0,68 \pm 0,03$
Ερύθημα + Πτυχώσεις	$0,69 \pm 0,02$
Ανισοϋψώσεις + Υφή	$0,56 \pm 0,02$
Ανισοϋψώσεις + Λεπτές Γραμμές	$0,50 \pm 0,03$
Ανισοϋψώσεις + Πτυχώσεις	$0,53 \pm 0,03$
Υφή + Λεπτές Γραμμές	$0,53 \pm 0,04$
Υφή + Πτυχώσεις	$0,52 \pm 0,04$
Λεπτές Γραμμές + Πτυχώσεις	$0,47 \pm 0,05$

χρωμη εικονομορφή.

### 3.5 Συζήτηση

Στην παρούσα μελέτη, έγινε διερεύνηση της χρήσης βαθιάς μάθησης για τη διάγνωση δερματικών αντιδράσεων σε επιδερμικές δοκιμασίες επιθέματος. Τα αποτελέσματα έδειξαν ότι η χρήση μιας μάσκας  $a$  αντί της απλής περικοπής της εικόνας εισόδου, καθώς και ο συνδυασμός των εικονομορφών Χρώματος και Ερυθήματος, όπως παράγονται από την κάμερα Antera 3D®, είναι οι πιο αποτελεσματικές στη δυαδική ταξινόμηση.

Επαληθεύτηκε ότι η περικοπή των εικόνων εισόδου σε μια συγκεκριμένη ROI απορρίπτει πολύτιμες πληροφορίες και, ως εκ τούτου, αποτελεί μια μη βέλτιστη τεχνική προεπεξεργασίας. Για το σκοπό αυτό, προτάθηκε η χρήση μιας μάσκας  $a$  (MA) για την ανάδειξη της ROI διατηρώντας παράλληλα τη συνολική πληροφορία της εικόνας. Επιπλέον, έγινε επέκταση όλων των εικόνων ώστε να περιλαμβάνουν επιπλέον πληροφορίες από παρακείμενες περιοχές. Αυτή η μέθοδος, η οποία ονομάστηκε Επαυξημένη Μάσκα  $a$  (EMA), βελτίωσε τη βαθμολογία  $F1$  σε σχέση με τις περικομμένες εικόνες, από 0,66 σε 0,83. Το τεστ McNemar επαλήθευσε περαιτέρω τη στατιστική σημαντικότητα της προτεινόμενης μεθόδου προεπεξεργασίας. Από την άλλη πλευρά, δεν διαπιστώθηκε κάποια στατιστικά σημαντική διαφορά χρησιμοποιώντας το τεστ McNemar μεταξύ EMA και MA. Αυτό θα μπορούσε να αποδοθεί στο περιορισμένο δείγμα δεδομένων και στη μη ύπαρξη αρκετών περιπτώσεων όπου η γειτονική περιοχή επιθέματος επηρέαζε μια περιοχή αλλεργιογόνου στην οποία ο ασθενής ήταν αρνητικός.



Πίνακας 3.8: Τιμές  $p$  που αντιστοιχούν στη σύγκριση κατά ζεύγη της ταξινόμησης των χαρακτηριστικών που μαθαίνονται αποκλειστικά από την έγχρωμη μορφή, και του συνδυασμού των χαρακτηριστικών που μαθαίνονται από την σύντηξη όλων των άλλων μορφών, όπως υπολογίζεται με τη χρήση του  $f$ -τεστ  $5 \times 2$ .

Υπόθεση	Τιμή $p$
Έγχρωμη - Έγχρωμη + Ερύθημα	0,016
Έγχρωμη - Έγχρωμη + Ανισοϋψώσεις	0,373
Έγχρωμη - Έγχρωμη + Υφή	0,353
Έγχρωμη - Έγχρωμη + Λεπτές Γραμμές	0,418
Έγχρωμη - Έγχρωμη + Πτυχώσεις	0,463
Έγχρωμη - Ερύθημα + Ανισοϋψώσεις	0,306
Έγχρωμη - Ερύθημα + Υφή	0,583
Έγχρωμη - Ερύθημα + Λεπτές Γραμμές	0,430
Έγχρωμη - Ερύθημα + Πτυχώσεις	0,724
Έγχρωμη - Ανισοϋψώσεις + Υφή	0,014
Έγχρωμη - Ανισοϋψώσεις + Λεπτές Γραμμές	0,002
Έγχρωμη - Ανισοϋψώσεις + Πτυχώσεις	0,003
Έγχρωμη - Υφή + Λεπτές Γραμμές	0,030
Έγχρωμη - Υφή + Πτυχώσεις	0,010
Έγχρωμη - Λεπτές Γραμμές + Πτυχώσεις	0,007

Τα αποτελέσματα υποδηλώνουν επίσης ότι οι πληροφορίες υφής που περιέχονται στις εικονομορφές Υφή, Ανισοϋψώσεις, Λεπτές γραμμές και Πτυχώσεις, δεν επαρκούν για την ταξινόμηση μιας αντίδρασης. Η ανάλυσή έδειξε ότι η συγκέντρωση αιμοσφαιρίνης, η οποία αντιπροσωπεύεται από την Εικόνα Ερυθήματος, ήταν ο ισχυρότερος προγνωστικός παράγοντας θετικής διάγνωσης, επιτυγχάνοντας μέση βαθμολογία  $F1$  0,87 με χρήση πενταπλής διασταυρωμένης επικύρωσης. Επιπλέον, τα αποτελέσματα έδειξαν ότι η ταξινόμηση σε απλές έγχρωμες εικόνες απέδωσε εξίσου καλά με το ερύθημα. Συγκεκριμένα, η τιμή  $p$  βρέθηκε να είναι 0,760 κατά τη σύγκριση των δύο, χρησιμοποιώντας το τεστ McNemar.

Από την άλλη πλευρά, κατά τον συνδυασμό διαφορετικών εικονομορφών, διαπιστώθηκε ότι το ΣΝΝ έμαθε συμπληρωματικά χαρακτηριστικά μεταξύ της έγχρωμης μορφής εικόνας και της ερυθρότητας εικόνας, επιτυγχάνοντας μεγαλύτερα αποτελέσματα σε σχέση με την απλή έγχρωμη εικόνα. Η σημαντικότητα των αποτελεσμάτων επαληθεύτηκε από την τιμή  $p$  του στατιστικού ελέγχου (0,016). Αυτός ήταν ο μόνος συνδυασμός χαρακτηριστικών στον οποίο βρέθηκε προστιθέμενη αξία. Συμπερασματικά, η προηγμένη κάμερα που χρησιμοποιήθηκε μπορεί να συμβάλει σε ακριβέστερη ταξινόμηση, αλλά μόνο όταν πραγματοποιείται σύντηξη χαρακτηριστικών μέσω διαφορετικών μορφών εικόνας.

Εξ όσων γνωρίζουμε, μέχρι στιγμής μόνο μία άλλη εργασία έχει δημοσιευθεί που να ασχολείται με το πρόβλημα της αυτοματοποιημένης διάγνωσης αντιδράσεων σε επιδερμικές δοκιμασίες επιθήματος. Οι Chan, Srivastava, Damaraju, *et al.* [96] συνέλεξαν ένα σύνολο δεδομένων 3695 έγχρωμων εικόνων, καθεμία από

τις οποίες απεικόνιζε μεμονωμένα αλλεργιογόνα (παρόμοια με την παρούσα τεχνική προεπεξεργασίας με περικοπή). Μόνο 118 από αυτές τις εικόνες χαρακτηρίστηκαν ως θετικές. Οι συγγραφείς χρησιμοποίησαν το προ-εκπαιδευμένο ΣΝΝ Xception της Google, επιτυγχάνοντας βαθμολογία  $F1$  0,89, η οποία συμφωνεί με τα αποτελέσματα της παρούσας μελέτης.

### 3.5.1 Περιορισμοί

Η παρούσα μελέτη έχει αρκετούς περιορισμούς που πρέπει να αναγνωριστούν.

Ο πρώτος περιορισμός που πρέπει να επισημανθεί είναι η πιθανότητα μεροληψίας του παρατηρητή. Η διάγνωση περιλαμβάνει οπτική επιθεώρηση του δέρματος με γυμνό μάτι, η οποία, σε συνδυασμό με μεγάλο αριθμό κλινικών περιστατικών κατά τη διάρκεια μιας τυπικής ημέρας, θα μπορούσε να οδηγήσει σε υποκειμενικά αποτελέσματα.

Δεύτερον, το μέγεθος του δείγματος του συνόλου δεδομένων περιορίστηκε σε 1579 δείγματα αλλεργιογόνων από 200 ασθενείς, γεγονός που ενδέχεται να εμποδίζει τη δυνατότητα γενίκευσης των αποτελεσμάτων σε μεγαλύτερους πληθυσμούς. Επιπλέον, η μελέτη επικεντρώθηκε σε ένα συγκεκριμένο σύνολο αλλεργιογόνων, που παρουσιάζονται στον πίνακα 3.1, και δεν συμπεριέλαβε άλλα πιθανά αλλεργιογόνα που μπορεί να είναι σημαντικά σε διαφορετικές χώρες ή πληθυσμούς.

Τρίτον, η μελέτη χρησιμοποίησε μία μόνο αρχιτεκτονική ΣΝΝ, συγκεκριμένα το EfficientNetB0. Ωστόσο, οι επιδόσεις διαφόρων αρχιτεκτονικών ΣΝΝ σε ένα παρόμοιο σε μέγεθος σύνολο ιατρικών δεδομένων διερευνήθηκαν ήδη στο κεφάλαιο 2 και το EfficientNetB0 βρέθηκε να είναι ένα από τα δίκτυα με τις καλύτερες επιδόσεις.

## **Κεφάλαιο 4**

# **Πρόγνωση Ασθενών με Αδενοκαρκίνωμα του Παγκρέατος σε Αξονικές Τομογραφίες**

### **4.1 Αδενοκαρκίνωμα του Παγκρέατος**

Μία από τις κύριες αιτίες θνησιμότητας που σχετίζεται με τον καρκίνο παγκρομιάως είναι το αδενοκαρκίνωμα του παγκρέατος, γνωστό απλά και ως καρκίνος του παγκρέατος. Από το 2023, κατατάσσεται ως η τρίτη κύρια αιτία θανάτου από καρκίνο στις Ηνωμένες Πολιτείες της Αμερικής, με εκτιμώμενες 64.050 νέες διαγνώσεις και 50.550 θανάτους [110]. Παρά τις βελτιώσεις στις μεθόδους ανίχνευσης και θεραπείας, το ποσοστό πενταετούς επιβίωσης εξακολουθεί να είναι εξαιρετικά χαμηλό, περίπου στο 12,5% [110]. Το ποσοστό αυτό αναδεικνύει την ιδιαίτερως επιθετική φύση της νόσου, και υπογραμμίζει την επιτακτική ανάγκη για ακριβέστερα και έγκαιρα προγνωστικά μοντέλα.

Τα κλινικά χαρακτηριστικά του καρκίνου του παγκρέατος περιλαμβάνουν συχνά απώλεια βάρους, κνησμό, ίκτερο, πόνο, δυσπεψία, και νεοεμφανιζόμενο διαβήτη, μεταξύ άλλων [111]. Ωστόσο, η έλλειψη πρώιμων προειδοποιητικών συμπτωμάτων και η συχνά καθυστερημένη διάγνωση συμβάλλουν στην κακή πρόγνωση που σχετίζεται με τη νόσο [112], [113].

#### **4.1.1 Διάγνωση**

Η αξονική τομογραφία (CT) αποτελεί σήμερα τη βασική απεικονιστική τεχνική για την αρχική αξιολόγηση του καρκίνου του παγκρέατος, με την ευαισθησία της να έχει βελτιωθεί έως και 95% χάρη στην βελτίωση της τεχνολογίας των μηχανημάτων, των πρωτοκόλλων εξέτασης και των τεχνικών μετεπεξεργασίας [114], [115]. Παρά τις προόδους αυτές, η αξονική τομογραφία εξακολουθεί να αντιμετωπίζει

σημαντικούς περιορισμούς, ιδίως όσον αφορά την πρόγνωση. Συγκεκριμένα, οι αξονικές τομογραφίες μπορούν να παρέχουν πληροφορίες για τη σταδιοποίηση, αλλά δεν προσφέρουν σημαντικές προγνωστικές πληροφορίες, ούτε μπορούν να κατευθύνουν τη θεραπεία πέραν της αρχικής σταδιοποίησης [116].

#### **4.1.2 Θεραπεία**

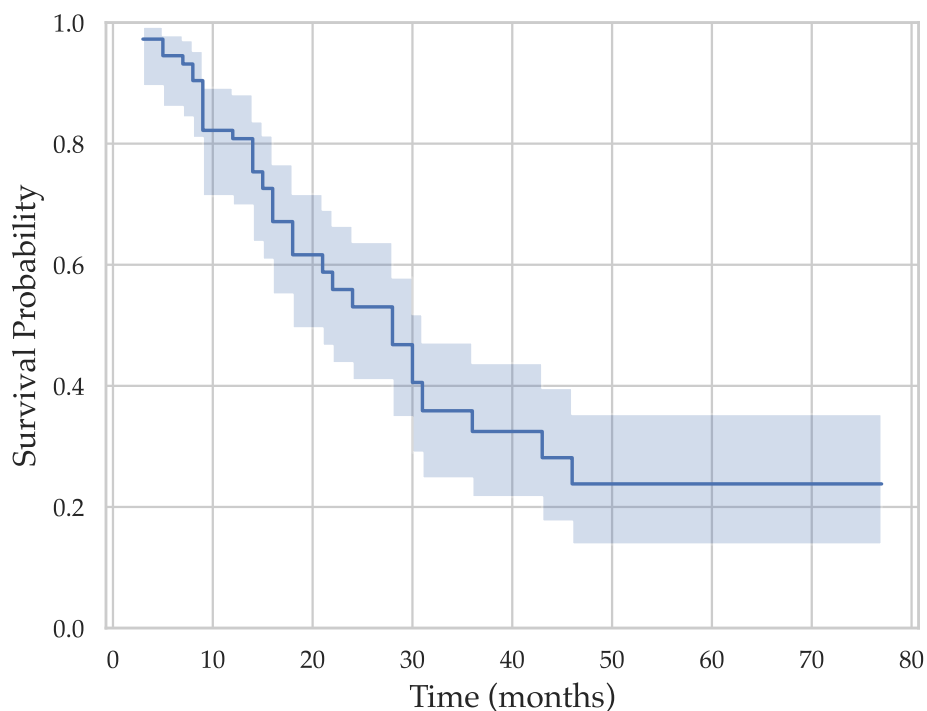
Η θεραπεία του καρκίνου του παγκρέατος περιλαμβάνει χειρουργική επέμβαση, χημειοθεραπεία, ακτινοθεραπεία και παρηγορητική φροντίδα. Η βέλτιστη θεραπευτική μέθοδος θα πρέπει να καθορίζεται σύμφωνα με τη σταδιοποίηση της νόσου.

Η χειρουργική εκτομή θεωρείται η μόνη θεραπεία για ίαση και μπορεί να οδηγήσει σε σημαντικά μεγαλύτερη επιβίωση σε σύγκριση με άλλες θεραπευτικές επιλογές. Η χειρουργική επέμβαση για το αδενοκαρκίνωμα του παγκρέατος περιλαμβάνει χειρουργική εκτομή του πρωτοπαθούς όγκου και των περιφερικών λεμφαδένων. Πρέπει να εκτελείται μόνο εφόσον έχουν αποκλειστεί οι απομακρυσμένες μεταστάσεις. Για το αδενοκαρκίνωμα της κεφαλής του παγκρέατος πραγματοποιείται μερική παγκρεατοδωδεκαδακτυλεκτομή (επέμβαση Whipple), και για το σώμα και την ουρά περιφερική παγκρεατεκτομή με σπληνεκτομή. Για την επίτευξη ριζικής εκτομής μπορεί να απαιτηθεί και αγγειακή εκτομή με ανακατασκευή. Σε ελάχιστες περιπτώσεις μπορεί να απαιτηθεί ολική παγκρεατεκτομή. Σε εξειδικευμένα κέντρα η θνητότητα μετά από παγκρεατεκτομή παραμένει κάτω από 5%.

## **4.2 Αντικείμενο της Μελέτης**

Τα τελευταία χρόνια, αρκετές μελέτες έχουν διερευνήσει τη χρήση ραδιομικών χαρακτηριστικών για την πρόγνωση του καρκίνου του παγκρέατος, επιδεικνύοντας ενθαρρυντικά αποτελέσματα [117]. Αυτά τα χαρακτηριστικά, που εξάγονται, συνήθως, από τις ακτινοδιαγνωστικές εξετάσεις, ποσοτικοποιούν χαρακτηριστικά του όγκου που δεν εκτιμώνται εύκολα από το ανθρώπινο μάτι, παρέχοντας έτσι δυνητικά πρόσθετη προγνωστική αξία. Ωστόσο, η εξαγωγή αυτών των χαρακτηριστικών από τους όγκους του παγκρέατος απαιτεί την λεπτομερή οριοθέτησή τους από εκπαιδευμένους κλινικούς ιατρούς, κάτι το οποίο αποτελεί μια επίπονη και χρονοβόρα διαδικασία. Η απαίτηση αυτή αποτελεί σημαντικό εμπόδιο για την ευρύτερη υιοθέτηση των εν λόγω τεχνικών στην κλινική πρακτική.

Στην παρούσα μελέτη, οι προκλήσεις αυτές αντιμετωπίζονται προτείνοντας ένα αυτοματοποιημένο προγνωστικό μοντέλο για το αδενοκαρκίνωμα του παγκρέατος, το οποίο χρησιμοποιεί προεγχειρητικές αξονικές τομογραφίες σε συνδυασμό με αλγορίθμους μηχανικής μάθησης. Αυτό το μοντέλο αυτοματοποιεί την κατάτμηση του παγκρέατος και των συναφών όγκων, οδηγώντας σε μια επακόλουθη πρόβλεψη της επιβίωσης των ασθενών βάσει ενός συνόλου τόσο ραδιομικών όσο και κλινικών χαρακτηριστικών που προέρχονται από τον όγκο και την υπόλοιπη



Σχήμα 4.1: Καμπύλη επιβίωσης Kaplan-Meier για 40 ασθενείς με αδενοκαρκίνωμα του παγκρέατος, έπειτα από παγκρεατεκτομή στο Αρεταίειο Νοσοκομείο.

παγκρεατική περιοχή. Αυτή η προσέγγιση απλοποιεί την εξαγωγή ραδιομικών χαρακτηριστικών, παρακάμπτοντας την απαίτηση για χειροκίνητη οριοθέτηση των περιοχών από τον ιατρό.

## 4.3 Υλικό και Μέθοδοι

### 4.3.1 Σύνολο Δεδομένων

Το σύνολο δεδομένων που χρησιμοποιήθηκε στην παρούσα μελέτη περιλαμβάνει προεγχειρητικές αξονικές τομογραφίες 40 ασθενών που υποβλήθηκαν σε παγκρεατεκτομή για καρκίνο του παγκρέατος στο Αρεταίειο Νοσοκομείο. Σε όλους τους ασθενείς η αξονική τομογραφία πραγματοποιήθηκε με το ειδικό πρωτόκολλο του παγκρέατος (βλ. Πίνακα 4.1). Η κατανομή των ασθενών ήταν 24 άνδρες και 16 γυναίκες, με διάμεση ηλικία 70 ετών (εύρος 41-82 έτη). Η ιστολογική ανάλυση που πραγματοποιήθηκε μετά την επέμβαση επιβεβαίωσε την παρουσία αδενοκαρκινώματος σε όλους τους ασθενείς. Σε δύο ασθενείς χορηγήθηκε νεοεπικουρική χημειοθεραπεία, ενώ όλοι οι ασθενείς έλαβαν επικουρική χημειοθεραπεία. Από το σύνολο των ασθενών, 32 (80%) υποβλήθηκαν σε παγκρεατοδωδεκαδακτυλεκτομή (επέμβαση Whipple), 6 ασθενείς (15%) σε περιφερική παγκρεατεκτομή με σπληνεκτομή, και 2 ασθενείς (5%) σε ολική παγκρεατεκτομή. 11 ασθενείς (27,5%) ήταν εν ζωή στο τέλος της περιόδου μελέτης. Στο σχήμα 4.1 απεικονίζεται η καμπύλη επιβίωσης Kaplan-Meier. Ο διάμεσος χρόνος επιβίωσης ήταν 24 μήνες.

Πίνακας 4.1: Το πρωτόκολλο παγκρέατος για τη λήψη αξονικών τομογραφιών άνω-κάτω κοιλίας σε ασθενείς με αδενοκαρκίνωμα του παγκρέατος.

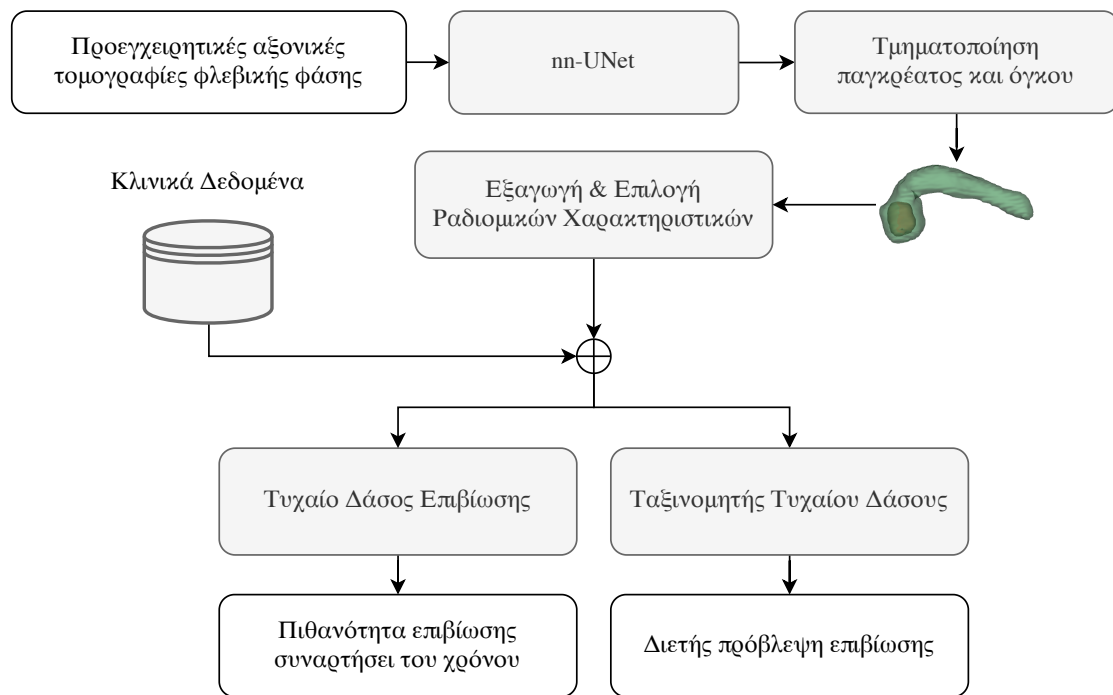
<b>Παράμετρος</b>	<b>Λεπτομέρειες</b>
Τύπος σάρωσης	Ελικοειδής
Πάχος τομής	Κατά προτίμηση κάτω του χιλιοστού (0.5–1.0 mm)
Διάστημα	Ίδιο με πάχος τομής
Στοματικός σκιαγραφικός παράγοντας	Ουδέτερος ή χαμηλές τιμές Hounsfield
Ενδοφλέβιος σκιαγραφικός παράγοντας	Κατά προτίμηση με υψηλή συγκέντρωση ιοδίου (>300 mg I/mL) με ρυθμό έγχυσης 3–5 mL/s
Λήψη σάρωσης	Παγκρεατική παρεγχυματική φάση στα 40–50", φλεβική φάση στα 65–70"
Ανακατασκευή εικόνας	Πάχος 2–5 mm σε εγκάρσιο επίπεδο, 2–3 mm πάχος σε στεφανιαίο επίπεδο Προβολές μέγιστης έντασης ή τρισδιάστατες ογκομετρικές παχιές τομές για αγγειακή αξιολόγηση

#### 4.3.2 Αρχιτεκτονική Επισκόπηση

Η αρχιτεκτονική επισκόπηση της μεθοδολογίας απεικονίζεται στο σχήμα 4.2. Πρώτον, για κάθε μεμονωμένο ασθενή, η προεγχειρητική αξονική τομογραφία φλεβικής φάσης επεξεργάζεται από το βαθύ ΣΝΝ nnU-Net, μια μέθοδος που έχει επιδείξει ανώτερες επιδόσεις σε εργασίες τμηματοποίησης ιατρικών εικόνων [118]. Το nnU-Net αναγνωρίζει και οριοθετεί αυτόματα το πάγκρεας και τον όγκο, παράγοντας έναν χάρτη τμηματοποίησης που περιλαμβάνει και τις δύο περιοχές. Για τον σκοπό αυτό, το nnU-Net εκπαιδεύτηκε αρχικά σε ένα εξωτερικό, δημόσια διαθέσιμο, σύνολο δεδομένων ασθενών με καρκίνο του παγκρέατος [119].

Μετά την τμηματοποίηση, εξάγονται ραδιομικά χαρακτηριστικά ξεχωριστά για το πάγκρεας και τις περιοχές του όγκου χρησιμοποιώντας τη βιβλιοθήκη PyRadiomics [120]. Αυτό πραγματοποιείται με την επιλογή του συνόλου των voxels από την αρχική αξονική τομογραφία, τα οποία έχουν επισημανθεί είτε ως «πάγκρεας», είτε ως «όγκος» κατά την τμηματοποίηση. Στη συνέχεια, υπολογίζονται τα ραδιομικά χαρακτηριστικά σε κάθε σύνολο voxel ξεχωριστά, αποτυπώνοντας ένα ευρύ φάσμα χαρακτηριστικών που σχετίζονται με την ένταση, την υφή και το σχήμα.

Στη συνέχεια, για τη μείωση της διαστατικότητας των δεδομένων και τη βελτίωση της απόδοσης του μοντέλου, πραγματοποιείται επιλογή ενός υποσυνόλου των χαρακτηριστικών. Για το σκοπό αυτό, η πολυσυγγραμμικότητα μεταξύ των υπολογισμένων ραδιομικών χαρακτηριστικών αξιολογείται κατά τη φάση της εκπαίδευσης για τη μείωση του, απορρίπτοντας τα χαρακτηριστικά τα οποία εμφα-



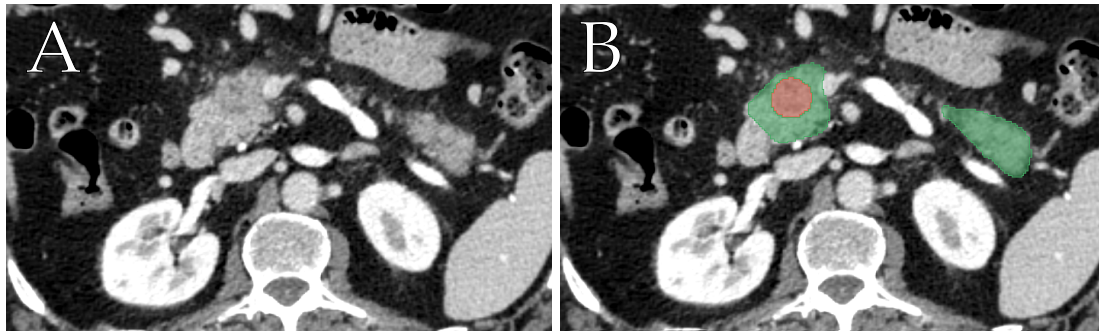
Σχήμα 4.2: Αρχιτεκτονική επισκόπηση της προτεινόμενης μεθοδολογίας.

νίζουν υψηλή γραμμική σχέση με άλλες μεταβλητές. Έτσι, διατηρούνται μόνο τα θεμελιώδη χαρακτηριστικά του συνόλου εκπαίδευσης. Στη συνέχεια, το επιλεγμένο υποσύνολο ραδιομικών χαρακτηριστικών συνδυάζεται με κλινικά δεδομένα, συγκεκριμένα τη σταδιοποίηση του όγκου και την ηλικία του ασθενούς.

Τέλος, για την ανάλυση των χαρακτηριστικών, χρησιμοποιούνται δύο διαφορετικά μοντέλα μηχανικής μάθησης. Για τη μοντελοποίηση της συνάρτησης πιθανότητας επιβίωσης του ασθενούς με την πάροδο του χρόνου, χρησιμοποιείται η μέθοδος των τυχαίων δασών επιβίωσης [121]. Από την άλλη πλευρά, για την πρόβλεψη του δυαδικού προβλήματος της διετούς επιβίωσης του ασθενούς, χρησιμοποιείται ένας ταξινομητής τυχαίου δάσους.

### 4.3.3 Κατάτμηση του Παγκρέατος και του Όγκου

Για την αυτοματοποιημένη οριοθέτηση των ορίων του παγκρέατος και του όγκου, χρησιμοποιήθηκε το δίκτυο nnU-Net [118]. Το δίκτυο αυτό επιλέχθηκε λόγω της απλότητάς του καθώς και της ανώτερης απόδοσής του, σε σύγκριση με άλλες μεθόδους, όπως αποδεικνύεται από τα κορυφαία αποτελέσματά του στο Medical Segmentation Decathlon [119], το οποίο περιλαμβάνει ποικίλες προκλήσεις ανάλυσης ιατρικών εικόνων, συμπεριλαμβανομένων αξονικών τομογραφιών ασθενών με καρκίνο στο πάγκρεας. Ειδικότερα, το nnU-Net ήταν η μέθοδος με τις καλύτερες επιδόσεις, επιτυγχάνοντας συντελεστή ομοιότητας Dice 0,80 και 0,52 για την τμηματοποίηση του παγκρέατος, και του όγκου, αντίστοιχα. Το βασικό πλεονέκτημα του nnU-Net είναι ότι δεν είναι απαραίτητη κάποια νέα αρχιτεκτονική δικτύου, καθώς αντί της ανάπτυξης ενός ειδικού σχεδιασμού δικτύου, μιας νέας συνάρτη-



Σχήμα 4.3: Παράδειγμα που δείχνει την κατάτμηση του αδενοκαρκινώματος (κόκκινο) και του υπόλοιπου παγκρέατος (πράσινο). Η εικόνα (A) απεικονίζει την αρχική εικόνα, και η εικόνα (B) απεικονίζει την εικόνα με τις τμηματοποιήσεις του παγκρέατος και του όγκου.

σης απωλειών ή ενός νέου σχήματος εκπαίδευσης, το nnU-Net αξιοποιεί ένα απλό U-Net (εξ ου και το όνομα “no new net”), και προσεκτικά επιλεγμένα σχήματα προεπεξεργασίας.

Το nnU-Net εκπαιδεύτηκε να τμηματοποιεί το πάγκρεας και τους όγκους χρησιμοποιώντας τα 282 δείγματα εκπαίδευσης που παρείχε το Medical Segmentation Decathlon. Στη συνέχεια, το εκπαιδευμένο δίκτυο εφαρμόστηκε για την ανάλυση και τμηματοποίηση του συνόλου δεδομένων ασθενών από το Αρεταίειο Νοσοκομείο, και οι παραγόμενες μάσκες παγκρέατος-όγκου αποθηκεύτηκαν για να χρησιμοποιηθούν για την εξαγωγή χαρακτηριστικών. Στο Σχήμα 4.3 απεικονίζεται ένα παράδειγμα αποτελέσματος τμηματοποίησης σε μία μόνο τομή.

Για τη διασφάλιση της βέλτιστης ποιότητας και συνοχής των δεδομένων, ακολουθήθηκε η μέθοδος προεπεξεργασίας που περιγράφεται στο [118]. Αρχικά, υπολογίστηκε η διάμεση απόσταση (σε χιλιοστά) σε όλες τις αξονικές τομογραφίες και όλα τα δεδομένα μετασχηματίστηκαν αναλόγως με χρήση γραμμικής παρεμβολής. Ακολούθως, πραγματοποιήθηκε μεταβολή της αντίθεσης με κατωφλίωση της έντασης όλων των εικονοστοιχείων της εικόνας εντός του κατώτερου 0,5% και του ανώτερου 99,5%. Τέλος, όλες οι τομές κανονικοποιήθηκαν αφαιρώντας τη μέση τιμή και διαιρώντας με την τυπική απόκλιση, οι οποίες υπολογίστηκαν και πάλι με βάση ολόκληρο το σύνολο δεδομένων.

#### 4.3.4 Εξαγωγή Χαρακτηριστικών

Στο πλαίσιο της παρούσας έρευνας, χρησιμοποιήθηκαν ραδιομικά χαρακτηριστικά για την εξαγωγή πληροφοριών από τμήματα όγκου και του παγκρέατος. Αυτά τα χαρακτηριστικά καταγράφουν διάφορες ποσοτικοποιήσεις που σχετίζονται με την ένταση, την υφή και το σχήμα του οριοθετημένου ιστού. Προηγούμενες μελέτες έχουν αποδείξει πως αυτά τα χαρακτηριστικά δύναται να χρησιμοποιηθούν για την πρόβλεψη της ανταπόκρισης και των αποτελεσμάτων της θεραπείας, τη σταδιοποίηση του όγκου και την ταυτοποίηση ιστών, μεταξύ άλλων [25], [122], [123].



Για τον υπολογισμό των ραδιομικών χαρακτηριστικών, χρησιμοποιήθηκε η βιβλιοθήκη PyRadiomics, ένα πακέτο python ανοικτού κώδικα που παρέχει έναν τυποποιημένο και αποτελεσματικό τρόπο υπολογισμού ραδιομικών χαρακτηριστικών από ιατρικές εικόνες. Το πακέτο αυτό αναλαμβάνει τη φόρτωση τόσο της εικόνας όσο και του αντίστοιχου χάρτη κατάτμησης, την εφαρμογή φίλτρων (όπως τα φίλτρα wavelet), και τον υπολογισμό των χαρακτηριστικών, προσφέροντας μια βελτιωμένη προσέγγιση που χρησιμεύει ως πρότυπο αναφοράς για την ακτινοδιαγνωστική, εξασφαλίζοντας την αναπαραγωγικότητα και τη συγκρισιμότητα.

Συνολικά, εξήχθησαν 854 ραδιομικά χαρακτηριστικά από κάθε τμηματοποίηση παγκρέατος και όγκου. Τα χαρακτηριστικά αυτά περιλάμβαναν χαρακτηριστικά πρώτης τάξης, χαρακτηριστικά σχήματος, χαρακτηριστικά Matrix Co-occurrence Matrix (GLCM), χαρακτηριστικά Grey Level Size Zone Matrix (GLSZM), χαρακτηριστικά Grey Level Run Length Matrix (GLRLM), χαρακτηριστικά Neighbouring Grey Tone Difference Matrix (NGTDM) και χαρακτηριστικά Gray Level Dependence Matrix (GLDM), τα οποία υπολογίστηκαν στις αρχικές εικόνες και στις εικόνες με φίλτρο wavelet.

#### **4.3.5 Επιλογή Χαρακτηριστικών**

Η εξαγωγή 854 ραδιομικών χαρακτηριστικών, τόσο από την κατάτμηση του παγκρέατος όσο και από την κατάτμηση του όγκου, δημιουργεί έναν υπερβολικά εκτεταμένο και υψηλών διαστάσεων χώρο χαρακτηριστικών, που υπερβαίνει κατά πολύ τον συνολικό αριθμό των δειγμάτων στο σύνολο δεδομένων. Αυτή η κατάσταση, που συχνά αποκαλείται ως «κατάρτα της διάστασης», προκαλεί προκλήσεις στην ανάπτυξη ακριβών μοντέλων πρόβλεψης λόγω της σπανιότητας των δεδομένων. Η υπερβολική πληθώρα χαρακτηριστικών δημιουργεί τον κίνδυνο το μοντέλο να συλλάβει θόρυβο και μη σημαντικά πρότυπα αντί για πραγματικές συσχετίσεις, με αποτέλεσμα να παρουσιάζει προβλήματα υπερπροσαρμογής στα δεδομένα και ανεπαρκούς γενίκευσης.

Για την αντιμετώπιση αυτών των προκλήσεων και τη βελτίωση της απόδοσης του μοντέλου, ένα ζωτικής σημασίας βήμα προεπεξεργασίας περιλαμβάνει την επιλογή χαρακτηριστικών. Η διαδικασία αυτή αποσκοπεί στον εντοπισμό των πιο σχετικών χαρακτηριστικών που συμβάλλουν σημαντικά στην προγνωστική ισχύ του μοντέλου, ενώ παράλληλα απορρίπτονται τα μη σημαντικά ή περιττά χαρακτηριστικά.

Για το σκοπό αυτό, χρησιμοποιήθηκε η τεχνική του συντελεστή διόγκωσης της διακύμανσης (Variance Inflation Factor — VIF) για την αξιολόγηση της πολυσυγγραμμικότητας μεταξύ των ακτινολογικών χαρακτηριστικών που εξήχθησαν. Συγκεκριμένα, ο VIF μετρά πόσο αυξάνεται η διακύμανση ενός εκτιμώμενου συντελεστή παλινδρόμησης όταν μια συγκεκριμένη μεταβλητή πρόβλεψης περιλαμβάνεται σε ένα μοντέλο γραμμικής παλινδρόμησης, σε σύγκριση με την περίπτωση που η εν λόγω μεταβλητή πρόβλεψης αποκλείεται. Μια υψηλή τιμή VIF για μια συγκεκριμένη μεταβλητή υποδηλώνει υψηλή γραμμική συσχέτιση με άλλες μετα-

βλητές.

Οι τιμές VIF υπολογίστηκαν επαναληπτικά για όλα τα ραδιομικά χαρακτηριστικά. Μετά από κάθε επανάληψη, το χαρακτηριστικό με την υψηλότερη VIF τιμή αφαιρούνταν από το σύνολο δεδομένων και η ανάλυση επαναλαμβάνονταν. Η επαναληπτική διαδικασία συνεχίστηκε έως ότου η υψηλότερη παρατηρούμενη τιμή VIF μειωθεί κάτω από το όριο των 10, γεγονός που σηματοδοτεί μειωμένη πολυσυγγραμμικότητα. Με την εφαρμογή αυτής της τεχνικής, ο αριθμός των χαρακτηριστικών περιορίστηκε αποτελεσματικά σε ένα πιο εύχρηστο σύνολο των 29. Ο κατάλογος των επιλεγμένων ακτινολογικών χαρακτηριστικών βρίσκεται στον Πίνακα 4.2. Αυτή η εξορθολογισμένη προσέγγιση επιλογής χαρακτηριστικών διασφαλίζει ότι τα επιλεγμένα ραδιομικά χαρακτηριστικά συλλαμβάνουν τις πιο σχετικές πληροφορίες για το προγνωστικό μοντέλο.

Εκτός από τα επιλεγμένα ραδιομικά χαρακτηριστικά, προστέθηκαν κλινικά χαρακτηριστικά, δηλαδή η ταξινόμηση TNM του όγκου [124] εκφρασμένη ως τρεις διαφορετικούς ακέραιους αριθμούς με τιμές μεταξύ 0-3 και η ηλικία των ασθενών σε έτη. Έτσι, ο συνολικός αριθμός των χαρακτηριστικών ανήλθε σε 33.

#### **4.3.6 Μοντελοποίηση Επιβίωσης**

Στην παρούσα μελέτη, χρησιμοποιήθηκε η μέθοδος των τυχαίων δασών επιβίωσης που προτάθηκε από τους Ishwaran et al. [121] για τη μοντελοποίηση της συνάρτησης πιθανότητας επιβίωσης για κάθε ασθενή με βάση τα αντίστοιχα ραδιομικά και κλινικά χαρακτηριστικά. Η μέθοδος των τυχαίων δασών επιβίωσης είναι μια μέθοδος που βασίζεται σε σύνολα δέντρων και έχει σχεδιαστεί για την ανάλυση δεδομένων επιβίωσης. Η λογική πίσω από την επιλογή των τυχαίων δασών επιβίωσης έναντι πιο παραδοσιακών μεθόδων, όπως το μοντέλο αναλογικών κινδύνων του Cox [125], έγκειται στην ικανότητά του να αποτυπώνει μη γραμμικές σχέσεις μεταξύ των προβλεπτικών παραγόντων και των αποτελεσμάτων επιβίωσης, ενώ δεν κάνει την υπόθεση των αναλογικών κινδύνων. Αυτή η ιδιότητα το καθιστά ιδιαίτερα κατάλληλο για σενάρια που περιλαμβάνουν σύνθετες αλληλεπιδράσεις μεταξύ χαρακτηριστικών και χρόνων επιβίωσης.

Προκειμένου να αξιολογηθεί η προγνωστική αξία των εξαχθέντων χαρακτηριστικών, το σύνολο δεδομένων χωρίστηκε σε δύο διαφορετικές ομάδες: η μία αποτελείται από χαρακτηριστικά του όγκου και η άλλη περιλαμβάνει χαρακτηριστικά από το υπόλοιπο τμήμα του παγκρέατος. Κατά συνέπεια, πραγματοποιήθηκαν ξεχωριστές έρευνες σχετικά με την προγνωστική ικανότητα του μοντέλου για κάθε ομάδα χαρακτηριστικών, καθώς και για τον συνδυασμό τους. Για τον υπολογισμό μιας ακριβέστερης αξιολόγησης στο περιορισμένο σύνολο δεδομένων, υιοθετήθηκε μια διαδικασία πενταπλής διασταυρούμενης επικύρωσης. Στη συνέχεια, συγκεντρώθηκαν όλες οι προβλέψεις όπως υπολογίστηκαν σε κάθε αναδίπλωση και υπολογίστηκε ο δείκτης συμφωνίας του Harrell (C-index) [126]. Για τη λήψη μιας εκτίμησης του μεγέθους της μεταβλητότητας που οφείλεται στην ανακατα-

Πίνακας 4.2: Πλήρης κατάλογος των 37 χαρακτηριστικών έπειτα από επιλογή χαρακτηριστικών με τη χρήση του συντελεστή διόγκωσης της διακύμανσης (VIF).

<b>Τύπος εικόνας</b>	<b>Τάξη χαρακτηριστικού</b>	<b>Χαρακτηριστικό</b>
Original	GLDM	Small Dependence Emphasis
Original	GLRLM	High Gray Level Run Emphasis
Original	GLSZM	Large Area High Gray Level Emphasis
Original	GLSZM	Small Area Low Gray Level Emphasis
Wavelet (LLH)	First Order	Total Energy
Wavelet (LLH)	GLSZM	Gray Level Non-Uniformity
Wavelet (LLH)	GLSZM	High Gray Level Zone Emphasis
Wavelet (LLH)	GLSZM	Small Area Emphasis
Wavelet (LLH)	GLSZM	Zone Entropy
Wavelet (LHL)	GLCM	Cluster Shade
Wavelet (LHL)	GLCM	Imc2
Wavelet (LHL)	GLSZM	Gray Level Non-Uniformity
Wavelet (LHH)	GLCM	Difference Entropy
Wavelet (LHH)	GLDM	Low Gray Level Emphasis
Wavelet (LHH)	GLRLM	Gray Level Variance
Wavelet (LHH)	GLSZM	Large Area Emphasis
Wavelet (LHH)	GLSZM	Small Area Low Gray Level Emphasis
Wavelet (HLL)	First Order	Kurtosis
Wavelet (HLL)	GLSZM	Size Zone Non-Uniformity
Wavelet (HLL)	GLSZM	Small Area Low Gray Level Emphasis
Wavelet (HLH)	GLRLM	Long Run High Gray Level Emphasis
Wavelet (HLH)	GLSZM	Gray Level Variance
Wavelet (HLH)	GLSZM	Small Area Emphasis
Wavelet (HLH)	GLSZM	Small Area Low Gray Level Emphasis
Wavelet (HHL)	GLSZM	Size Zone Non-Uniformity
Wavelet (HHL)	GLSZM	Small Area Low Gray Level Emphasis
Wavelet (HHH)	First Order	Median
Wavelet (LLL)	GLRLM	Long Run Low Gray Level Emphasis
Wavelet (LLL)	GLSZM	Gray Level Non-Uniformity

νομή των δεδομένων, η διαδικασία επαναλήφθηκε 10 φορές, κάθε φορά με τυχαία επανακατανομή των δεδομένων στις πέντε διαμερίσεις. Ο τελικός δείκτης CI υπολογίστηκε ως προς τη μέση τιμή και την τυπική απόκλιση για τις 10 επαναλήψεις. Οι υπερπαραμέτροι του μοντέλου ορίστηκαν ευρετικά ως εξής: ο αριθμός των δέντρων ορίστηκε σε 100, ο μέγιστος αριθμός χαρακτηριστικών που λαμβάνεται υπόψιν ορίστηκε ως ο λογάριθμος της βάσης 2 του συνολικού αριθμού των χαρακτηριστικών, ο ελάχιστος αριθμός δειγμάτων για έναν κόμβο-φύλλο σε 5, και ο ελάχιστος αριθμός δειγμάτων που απαιτούνται για τη διάσπαση ενός κόμβου σε 2. Όσον αφορά το βάθος του δέντρου, οι κόμβοι επεκτάθηκαν μέχρι όλα τα φύλλα να περιέχουν λιγότερα από 2 δείγματα.

#### **4.3.7 Ανάλυση Σημαντικότητας Χαρακτηριστικών**

Για την αξιολόγηση της σημαντικότητας των επιλεγμένων ραδιομικών και κλινικών χαρακτηριστικών στην πρόγνωση των ασθενών, πραγματοποιήθηκε μια ανάλυση σημαντικότητας των χαρακτηριστικών με τη μέθοδο της αντιμετάθεσης [9]. Στόχος ήταν να αξιολογηθεί η επίδραση των μεμονωμένων χαρακτηριστικών στην προγνωστική απόδοση του μοντέλου τυχαίου δάσους επιβίωσης.

Η σημαντικότητα ενός χαρακτηριστικού με τη χρήση αντιμετάθεσης υπολογίζεται ως η μείωση της απόδοσης ενός μοντέλου όταν οι τιμές ενός μεμονωμένου χαρακτηριστικού ανακατεύονται τυχαία εντός του συνόλου δεδομένων, ενώ όλα τα άλλα χαρακτηριστικά παραμένουν αμετάβλητα [9]. Με την αντιμετάθεση των τιμών ενός συγκεκριμένου χαρακτηριστικού, καταστρέφεται η σχέση μεταξύ του εν λόγω χαρακτηριστικού και της μεταβλητής-στόχου. Η προκύπτουσα πτώση της απόδοσης του μοντέλου παρέχει πληροφορίες για τη σημαντικότητα του χαρακτηριστικού.

Προκειμένου να υπάρχει μια βασική γραμμή αναφοράς, ένα χαρακτηριστικό θορύβου τυχαία αντλημένο από μια Γκαουσιανή κατανομή με μέση τιμή 0 και τυπική απόκλιση 1 εισήχθη στο σύνολο δεδομένων. Αξιολογώντας τη σημασία άλλων χαρακτηριστικών σε σχέση με το χαρακτηριστικό θορύβου, είναι δυνατή η ποσοτικοποίηση της σχετικής συμβολής τους στην προβλεπτική απόδοση του μοντέλου.

Με την προσθήκη του χαρακτηριστικού θορύβου στο σύνολο δεδομένων, το μοντέλο τυχαίου δάσους επιβίωσης εκπαιδεύτηκε εκ νέου στα τροποποιημένα δεδομένα. Για κάθε χαρακτηριστικό στο σύνολο δεδομένων, υπολογίστηκε η μείωση της απόδοσης κατά την αντιμετάθεσή του. Στη συνέχεια, η απόδοση του μοντέλου αξιολογήθηκε σε κάθε αναδίπλωση με τη χρήση του C-index. Η σημαντικότητα του κάθε χαρακτηριστικού υπολογίστηκε βάσει της διαφοράς μεταξύ του αρχικού δείκτη CI και του δείκτη CI έπειτα τυχαίας ανακατανομής του χαρακτηριστικού υπό διερεύνηση. Για τη λήψη μιας αξιόπιστης εκτίμησης της σημαντικότητας του κάθε χαρακτηριστικού, υπολογίστηκε η μέση τιμή και η τυπική απόκλιση της σημαντικότητας κάθε χαρακτηριστικού σε όλες τις διαμερίσεις.

### **4.3.8 Ταξινόμηση Διετούς Επιβίωσης**

Εκτός από την εκτίμηση της καμπύλης πιθανότητας επιβίωσης με την πάροδο του χρόνου, διερευνήθηκε η ικανότητα ενός ταξινομητή τυχαίου δάσους να προβλέψει τη διετή επιβίωση των ασθενών. Για τον σκοπό αυτό, εκπαιδεύτηκαν δύο ταξινομητές, ο καθένας από τους οποίους χρησιμοποίησε τα χαρακτηριστικά που προέρχονταν από τον όγκο και το υπόλοιπο πάγκρεας, αντίστοιχα, χρησιμοποιώντας πενταπλή διασταυρούμενη επικύρωση. Τα τελικά αποτελέσματα υπολογίστηκαν ως ο μέσος όρος και η τυπική απόκλιση των μετρικών καθόλο το σύνολο διαμερίσεων.

## **4.4 Αποτελέσματα**

### **4.4.1 Μοντελοποίηση Επιβίωσης**

Ένα τυχαίο δάσος επιβίωσης εκπαιδεύτηκε ξεχωριστά σε χαρακτηριστικά που εξήχθησαν από τον όγκο και το υπόλοιπο πάγκρεας, υπολογίζοντας τον δείκτη CI σε όλες τις προβλέψεις με τη χρήση πενταπλής διασταυρούμενης επικύρωσης, επαναλαμβάνοντας τη διαδικασία 10 φορές. Για τα χαρακτηριστικά του όγκου, το CI υπολογίστηκε ως  $0,731 \pm 0,015$ . Για τα χαρακτηριστικά από το υπόλοιπο πάγκρεας, ο επιτευχθείς δείκτης CI ήταν  $0,485 \pm 0,041$ .

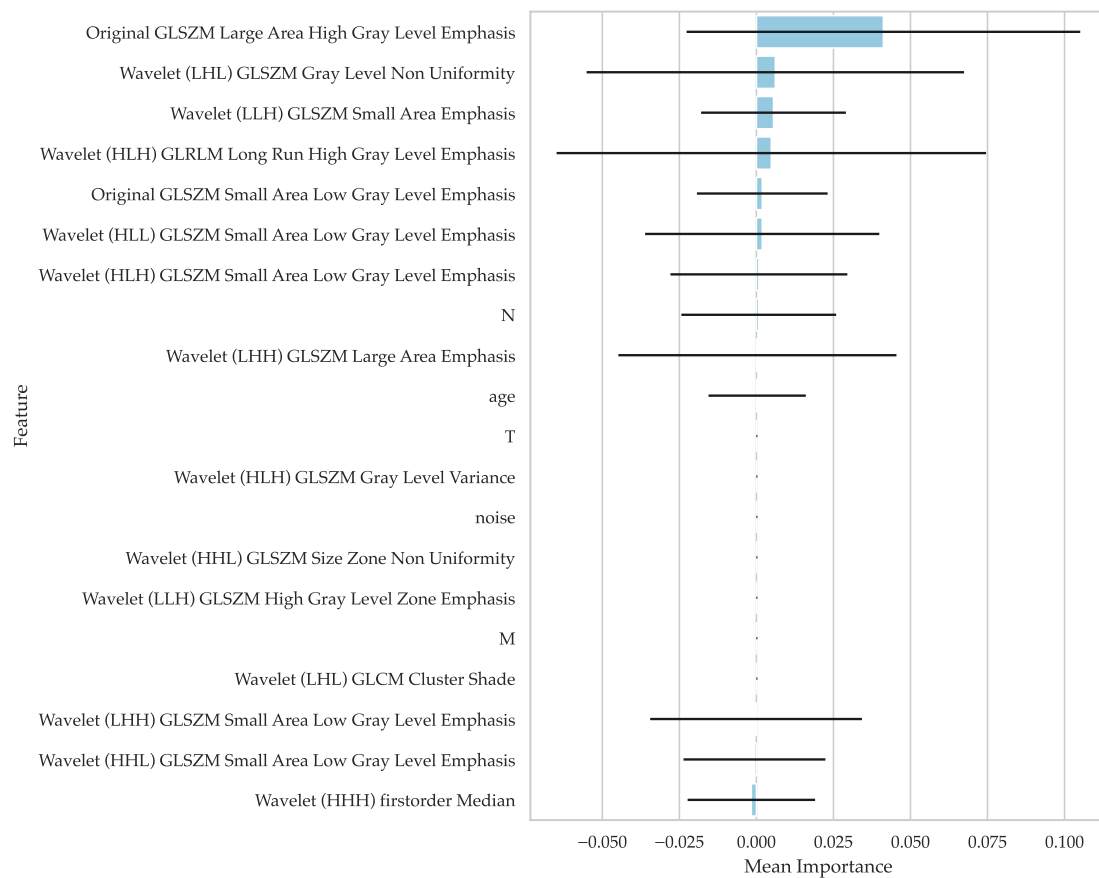
### **4.4.2 Ανάλυση Σημαντικότητας Χαρακτηριστικών**

Το Σχήμα 4.4 απεικονίζει μια οπτική σύγκριση της σημαντικότητας των 20 κορυφαίων χαρακτηριστικών με βάση τη μέση σημαντικότητα μετατροπής τους, η οποία εκφράζεται ως η μείωση του δείκτη CI κάθε φορά που ένα συγκεκριμένο χαρακτηριστικό ανακατεύεται τυχαία εντός του συνόλου δεδομένων.

Συνολικά, το χαρακτηριστικό Large Area High Gray Level Emphasis (LAH-GLE) στον αρχικό τύπο εικόνας ξεπέρασε κατά πολύ τα υπόλοιπα χαρακτηριστικά σε σημαντικότητα, με τιμή 0,041. Συγκριτικά, το Gray Level Non Uniformity στην εικόνα wavelet (LHL) ήρθε στη δεύτερη θέση, με τιμή σημαντικότητας 0,006. Ο θόρυβος, ο οποίος εισήχθη ως τιμή αναφοράς, εμφάνισε σημασία μηδέν.

### **4.4.3 Ταξινόμηση Διετούς Επιβίωσης**

Ο Πίνακας 4.3 απεικονίζει την ακρίβεια, την ευαισθησία, την ειδικότητα, την αξιοπιστία και τη βαθμολογία F1, όπως υπολογίστηκαν κατά τη διάρκεια της 10 φορές επαναλαμβανόμενης πενταπλής διασταυρούμενης επικύρωσης. Για αυτό το δυαδικό πρόβλημα ταξινόμησης, ένας ασθενής που επέζησε πέραν των 24 μηνών θεωρήθηκε ως « θετικό » δείγμα.



Σχήμα 4.4: Τα 20 κορυφαία χαρακτηριστικά βάσει της μέσης σημαντικότητάς τους.

Πίνακας 4.3: Μετρικές επιδόσεων για το πρόβλημα δυαδικής ταξινόμησης. Οι ασθενείς που επέζησαν πέραν των 24 μηνών θεωρήθηκαν θετικά δείγματα.

Προέλ. χαρακτ.	Βαθμ. F1	Ακρίβεια	Ειδικότητα	Ευαισθησία	Αξιοπιστία
Όγκος	0,74 ± 0,04	0,76 ± 0,04	0,78 ± 0,04	0,74 ± 0,05	0,75 ± 0,04
Πάγκρεας	0,48 ± 0,07	0,51 ± 0,06	0,54 ± 0,08	0,47 ± 0,07	0,50 ± 0,06

## 4.5 Συζήτηση

### 4.5.1 Μοντελοποίηση Επιβίωσης

Στο πλαίσιο της μοντελοποίησης της συνολικής επιβίωσης, οι δοκιμές στα τυχαία δάση επιβίωσης πραγματοποιήθηκαν χωριστά σε χαρακτηριστικά που προέρχονταν από τον όγκο και το υπόλοιπο πάγκρεας. Η μεθοδολογία περιελάμβανε τον υπολογισμό του CI σε όλες τις προβλέψεις με τη χρήση πενταπλής διασταυρούμενης επικύρωσης, η οποία επαναλήφθηκε 10 φορές. Τα χαρακτηριστικά του όγκου συνέβαλαν στην επίτευξη CI με μέση τιμή 0,731, σημαντικά υψηλότερη από το 0,485 που προέκυψε από τα χαρακτηριστικά του υπόλοιπου παγκρέατος. Αυτή η απόκλιση υποδεικνύει τη σχετική σημασία και την προγνωστική ισχύ των χαρακτηριστικών του όγκου για την πρόγνωση του καρκίνου. Επιπλέον, το επιτευχθέν CI 0,731 υποδεικνύει ότι η πλήρως αυτοματοποιημένη προσέγγιση που προτείνεται στην παρούσα μελέτη θα μπορούσε ενδεχομένως να παρέχει μια ανώτερη, ως προς την προγνωστική αξία, μέθοδο, σε σύγκριση με τη σταδιοποίηση TNM. Αυτό είναι προφανές εάν ληφθεί υπόψιν η έρευνα που διεξήχθη από τους Chen et al. (2016) [116], όπου αξιολογήθηκε η προγνωστική αποτελεσματικότητα των χαρακτηριστικών σταδιοποίησης από την 8η έκδοση του εγχειριδίου σταδιοποίησης καρκίνου AJCC [127], όπου ο αντίστοιχος δείκτης CI υπολογίστηκε ως 0,585. Ομοίως, σε μελέτη που διεξήχθη από τους Mohammad et al. (2023) [128], ο παρατηρούμενος δείκτης CI βρέθηκε να είναι 0,633.

Στη σχετική βιβλιογραφία, παρατηρείται μια επικράτηση μελετών που αναφέρουν βελτιωμένη απόδοση με τη χρήση της ραδιομικών χαρακτηριστικών έναντι των χαρακτηριστικών σταδιοποίησης, ένα εύρημα που ευθυγραμμίζεται με τα αποτελέσματα της παρούσας μελέτης. Για παράδειγμα, σε μια μελέτη των Park et al. (2021) [27], αναφέρθηκε ότι η προσθήκη της ραδιομικής στα κλινικά χαρακτηριστικά βελτίωσε τον δείκτη CI από 0,679 σε 0,741. Παρόμοια αποτελέσματα αναφέρθηκαν από τους Xie et al. (2020) [129], οι οποίοι πέτυχαν δείκτη CI 0,726, μέσω της χρήσης ραδιομικών χαρακτηριστικών. Αντιθέτως, άλλες έρευνες κατέγραψαν χαμηλότερους δείκτες CI με τη χρήση ραδιομικής ανάλυσης, παρόλο που αυτοί παρέμεναν ανώτεροι από τα αποτελέσματα που επιτεύχθηκαν αποκλειστικά μέσω κλινικών χαρακτηριστικών. Οι Healy et al. (2022) [130] ανέφεραν δείκτη CI 0,545 με χρήση κλινικών-ραδιομικών χαρακτηριστικών, ενώ οι Zhang et al. (2020) [131] ανέφεραν δείκτη CI 0,491 με ραδιομικά χαρακτηριστικά μόνο, ο οποίος βελτιώθηκε αργότερα σε 0,651 όταν χρησιμοποιήθηκαν χαρακτηριστικά που εξήχθησαν αυτόματα με χρήση ενός ΣΝΝ.

Οι αποκλίσεις μεταξύ των αποτελεσμάτων αυτών των μελετών μπορούν να αποδοθούν σε διάφορες αιτίες, συμπεριλαμβανομένης της ποικιλομορφίας στις μεθόδους επιλογής χαρακτηριστικών, των διαφοροποιήσεων στις τεχνικές προεπεξεργασίας των χαρακτηριστικών και της εγγενούς ετερογένειας των χρησιμοποιούμενων συνόλων δεδομένων. Ωστόσο, η ευθυγράμμιση των αποτελεσμάτων μας με τις

κορυφαίες μελέτες προσφέρει αξιοσημείωτα στοιχεία για τη δυνατότητα πλήρους αυτοματοποίησης της προγνωστικής διαδικασίας για τον καρκίνο. Αυτό είναι ιδιαίτερα σημαντικό αν αναλογιστεί κανείς ότι όλες οι μελέτες μέχρι στιγμής εξαρτώνται από τη χειροκίνητη τμηματοποίηση του όγκου και του παγκρέατος.

#### **4.5.2 Ανάλυση Σημαντικότητας Χαρακτηριστικών**

Η γραφική απεικόνιση της σημαντικότητας των χαρακτηριστικών στο Σχήμα 4.4 αποκάλυψε ότι μόνο ένας περιορισμένος αριθμός χαρακτηριστικών είναι σημαντικό για την εκτίμηση της επιβίωσης. Το LAHGLE στον αρχικό τύπο εικόνας ξεχώρισε ως το πιο σημαντικό χαρακτηριστικό, με τιμή σημαντικότητας 0,041. Αυτό το μέτρο ξεπέρασε κατά πολύ το δεύτερο πιο σημαντικό χαρακτηριστικό, το Gray Level Non Uniformity στην εικόνα wavelet (LHL), το οποίο επέδειξε χαμηλή σημαντικότητα 0,006.

Το LAHGLE μετρά την αναλογία στην εικόνα της κοινής κατανομής των μεγαλύτερων ζωνών με υψηλότερες τιμές στάθμης του γκρι [132]. Έτσι, φαίνεται ότι η πιθανότητα επιβίωσης μπορεί, τουλάχιστον εν μέρει, να σχετίζεται με την έκταση του όγκου που αποτελείται από μεγάλες ζώνες συνεχόμενων εικονοστοιχείων υψηλής έντασης. Η συγκεκριμένη κατευθυντικότητα αυτής της συσχέτισης παραμένει απροσδιόριστη, δεδομένου ότι η υπολογιζόμενη σημασία του χαρακτηριστικού δεν περιγράφει αν μια υψηλότερη πιθανότητα επιβίωσης συσχετίζεται με μια υψηλότερη ή χαμηλότερη τιμή LAHGLE.

Η εφαρμογή του μοντέλου αναλογικών κινδύνων Cox [125] χρησιμοποιώντας μόνο το χαρακτηριστικό LAHGLE έδωσε δείκτη CI  $0,682 \pm 0,014$ . Παρόλο που ο δείκτης αυτός δεν ταιριάζει με αυτόν που προέκυψε με τη χρήση όλων των διαθέσιμων χαρακτηριστικών, υπογραμμίζει περαιτέρω την προγνωστική σημασία του συγκεκριμένου χαρακτηριστικού.

Επιπλέον, η ανάλυση σημαντικότητας φαίνεται να υποδηλώνει ότι η ταξινόμηση TNM έχει περιορισμένη σημασία για την πρόβλεψη. Ωστόσο, θα πρέπει να σημειωθεί ότι αυτή η συνάφεια εξαρτάται από τα ειδικά χαρακτηριστικά του συνόλου δεδομένων, το οποίο στην παρούσα μελέτη επέδειξε ελάχιστη ετερογένεια όσον αφορά την ταξινόμηση TNM. Για παράδειγμα, κανένας ασθενής της ομάδας δεν παρουσίασε μετάσταση (το M ήταν 0). Επιπλέον, είναι ζωτικής σημασίας να αναγνωριστεί ότι η διαπιστωμένη σημασία των χαρακτηριστικών δεν αποτελεί απόλυτο μέτρο της συσχέτισής τους με την επιβίωση. Αντίθετα, είναι συγκεκριμένη για το μοντέλο πρόβλεψης που χρησιμοποιήθηκε στην παρούσα μελέτη, τα τυχαία δάση επιβίωσης, καθώς και το χρησιμοποιούμενο δίκτυο τμηματοποίησης. Ως εκ τούτου, η προσδιορισθείσα σημασία των χαρακτηριστικών μπορεί να διαφέρει με τη χρήση άλλων προγνωστικών μοντέλων ή προσεγγίσεων τμηματοποίησης.



### 4.5.3 Ταξινόμηση Διαιτούς Επιβίωσης

Όπως απεικονίζεται στον Πίνακα 4.3, τα χαρακτηριστικά που προέκυψαν από τον όγκο υπερέιχαν σημαντικά εκείνων που προέκυψαν από το υπόλοιπο πάγκρεας. Τα χαρακτηριστικά του όγκου παρουσίασαν βαθμολογία F1  $0,74 \pm 0,04$ , ακρίβεια  $0,76 \pm 0,04$  και αξιοπιστία  $0,75 \pm 0,04$ . Αντίθετα, οι μετρικές απόδοσης για τα χαρακτηριστικά που προέκυψαν από το υπόλοιπο πάγκρεας ήταν σημαντικά χαμηλότερες. Το εύρημα αυτό υπογραμμίζει τη σημασία της ιεράρχησης των ειδικών για τον όγκο χαρακτηριστικών, τα οποία φαίνεται να έχουν αυξημένη προγνωστική αξία σε σύγκριση με τα χαρακτηριστικά που προέρχονται από το υπόλοιπο πάγκρεας.

Τα αποτελέσματα συμφωνούν με ανάλογες μελέτες που εξετάζουν την πρόβλεψη της διαιτούς επιβίωσης με τη χρήση ραδιομικών χαρακτηριστικών που εξάγονται από χειροκίνητες τμηματοποιήσεις. Συγκεκριμένα, η μελέτη που διεξήχθη από τον Osman (2019) [133] ανέφερε ευαισθησία 0,79 και περιοχή κάτω από την καμπύλη (AUC) 0,82. Αντίστοιχα, η έρευνα των Chakraborty et al. (2017) [134] ανέφερε ακρίβεια 0,75, ευαισθησία 0,68 και ειδικότητα 0,80. Αυτή η ευθυγράμμιση των αποτελεσμάτων επαληθεύει περαιτέρω την προτεινόμενη προσέγγιση της παρούσας έρευνας και ενισχύει την εγκυρότητα της χρήσης μιας αυτοματοποιημένης προσέγγισης για την πρόγνωση.

### 4.5.4 Περιορισμοί

Η παρούσα μελέτη έχει δύο περιορισμούς που πρέπει να επισημανθούν.

Πρώτον, η επικύρωση της αυτοματοποιημένης προσέγγισής μας πραγματοποιήθηκε σε ένα περιορισμένο σύνολο δεδομένων 40 ασθενών από ένα μόνο κέντρο. Παρόλο που πραγματοποιήθηκε εκτενής δοκιμή με τη χρήση επαναλαμβανόμενης διασταυρούμενης επικύρωσης και υπολογίζοντας τα αποτελέσματα σε ολόκληρο το μέγεθος του συνόλου δεδομένων, είναι πιθανό ότι η ικανότητα γενίκευσης των ευρημάτων είναι περιορισμένη σε διαφορετικές κλινικές ρυθμίσεις και πληθυσμούς ασθενών. Η μελλοντική έρευνα θα πρέπει να στοχεύει στην επικύρωση αυτών των ευρημάτων σε ένα μεγαλύτερο, πολυκεντρικό σύνολο δεδομένων, το οποίο θα παρέχει μια πιο ισχυρή δοκιμή της αποτελεσματικότητας του συστήματος.

Δεύτερον, ο πλήρως αυτοματοποιημένος χαρακτήρας της προσέγγισης συνεπάγεται ότι η ποιότητα της τμηματοποίησης επηρεάζει άμεσα την απόδοση της μετέπειτα ραδιομικής ανάλυσης. Παρόλο που το nnU-Net έχει επιδειξει ικανοποιητικά αποτελέσματα τμηματοποίησης, θα πρέπει να σημειωθεί ότι τυχόν ανακρίβειες στη διαδικασία τμηματοποίησης μπορεί να οδηγήσουν σε μη βέλτιστη εξαγωγή χαρακτηριστικών, η οποία μπορεί, με τη σειρά της, να επηρεάσει την απόδοση των ταξινομητών επιβίωσης και δυαδικών ταξινομητών.

#### **4.5.5 Μελλοντικές Προοπτικές**

Μελλοντικές μελέτες θα μπορούσαν να αξιοποιήσουν αυτή την αυτοματοποιημένη προσέγγιση και τα αναγνωρισμένα προγνωστικά χαρακτηριστικά για να καθοδηγήσουν τις αποφάσεις θεραπείας σύμφωνα με τα προβλεπόμενα αποτελέσματα των ασθενών. Αυτό θα μπορούσε να κυμαίνεται από μια πιο επιθετική προσέγγιση σε ασθενείς που προβλέπεται να έχουν πιο ευνοϊκή πρόγνωση έως μια πιο συντηρητική προσέγγιση σε εκείνους που προβλέπεται να έχουν λιγότερο ευνοϊκή έκβαση (ή και το αντίστροφο).

## Κεφάλαιο 5

# Συμπεράσματα

Στην παρούσα διδακτορική διατριβή, αναπτύχθηκαν και αξιολογήθηκαν μοντέλα MM για την αυτοματοποιημένη διάγνωση και πρόγνωση τριών διαφορετικών παθολογιών, σε τρεις διαφορετικές απεικονιστικές τεχνικές.

Στην πρώτη μελέτη της διατριβής, αξιολογήθηκαν διάφορα δίκτυα βαθιάς μάθησης για την ταξινόμηση ιστολογικών εικόνων οστεοσαρκώματος. Τα αποτελέσματα έδειξαν ότι τα κοινώς χρησιμοποιούμενα δίκτυα βαθιάς μάθησης παρουσιάζουν πιθανώς υπερπροσαρμογή στο σύνολο δεδομένων, με αποτέλεσμα να υπερτερούν κυρίως τα μικρότερα σε μέγεθος δίκτυα. Συγκεκριμένα, τα μοντέλα MobileNetV2 και EfficientNetB0 οδήγησαν στην αποτελεσματικότερη συνολική ταξινόμηση των μη όγκων, των βιώσιμων όγκων και της νέκρωσης όταν οι διαστάσεις των αρχικών εικόνων μειώθηκαν από  $1024 \times 1024$  εικονοστοιχεία σε  $256 \times 256$ . Η επανεκπαίδευση του MobileNetV2 χρησιμοποιώντας πενταπλή διασταυρούμενη επικύρωση έδειξε συνεπή αποτελέσματα σε όλες τις διαμερίσεις, επιτυγχάνοντας συνολική ακρίβεια 0,91 και μέση ευαισθησία 0,95, 0,93 και 0,83 για μη όγκους, βιώσιμους όγκους και νέκρωση, αντίστοιχα. Η αφαίρεση των εικόνων που περιείχαν ταυτόχρονα τόσο βιώσιμους όγκους, όσο και νέκρωση, βελτίωσε περαιτέρω τα αποτελέσματα, πετυχαίνοντας μέση τιμή ευαισθησίας 0,95, 0,98 και 0,93 και συνολική ακρίβεια 0,96. Τα ευρήματα αυτά παρέχουν ισχυρές ενδείξεις ότι τα μικρότερα και οικονομικότερα (όσον αφορά τους υπολογιστικούς πόρους) δίκτυα δύνανται να χρησιμοποιηθούν για τη βελτίωση των αποτελεσμάτων στο σύνολο δεδομένων του ΟΣ χωρίς να χρειάζεται η προσφυγή σε ολοένα μεγαλύτερα και πιο πολύπλοκα μοντέλα. Ως εκ τούτου, μελλοντικές έρευνες θα μπορούσαν να επικεντρωθούν στην αξιολόγηση των αποτελεσμάτων πιο δραστικών τεχνικών κανονικοποίησης (regularization), όπως η προ-εκπαίδευση των μοντέλων σε παρόμοια αλλά μεγαλύτερα σύνολα δεδομένων, η χρήση πιο δημιουργικών τεχνικών επαύξησης, η μείωση της διαστατικότητας εισόδου και του μεγέθους παρτίδας κατά την εκπαίδευση, καθώς και η προσθήκη πρώιμης διακοπής της εκπαίδευσης [135]. Αυτές οι τεχνικές θα μπορούσαν να επιτύχουν καλύτερα αποτελέσματα στο σύνολο δεδομένων ΟΣ και να χρησιμεύσουν ως ένα ανεκτίμητο εργαλείο που, όταν χρησιμοποιείται σε συνδυασμό με την τεχνογνωσία και την κλινική εμπειρία των

παθολογοανατόμων, θα μπορούσε τελικά να οδηγήσει σε μια βελτιωμένη έκβαση της νόσου για τους ασθενείς.

Στη δεύτερη μελέτη, διερευνήθηκε η ακρίβεια δυαδικής ταξινόμησης ενός μοντέλου βαθιάς μάθησης για την ανίχνευση δερματικών αντιδράσεων σε επιδερμικές δοκιμασίες επιθέματος, χρησιμοποιώντας διαφορετικές τεχνικές προεπεξεργασίας και μορφές εικόνας. Πραγματοποιήθηκαν τρία πειράματα που περιελάμβαναν τη σύγκριση διαφορετικών τεχνικών προεπεξεργασίας, τη σύγκριση των διαφορετικών εικονομορφών που μπορούν να ληφθούν με την κάμερα, και τη διερεύνηση του συνδυασμού των διαφορετικών αυτών εικονομορφών. Τα αποτελέσματα έδειξαν ότι η χρήση μιας τεχνικής προεπεξεργασίας που διατηρεί τη συνολική πληροφορία της εικόνας μπορεί να οδηγήσει σε βελτιωμένη ακρίβεια ταξινόμησης, ενώ ο συνδυασμός της απλής έγχρωμης εικόνας και του ερυθρήματος μπορεί επίσης να προσφέρει προστιθέμενη αξία στην αυτοματοποιημένη αξιολόγηση, και να επιτύχει ακρίβεια μεγαλύτερη του 91%. Συνολικά, η μελέτη υποδηλώνει ότι τα μοντέλα βαθιάς μάθησης έχουν τη δυνατότητα να ανιχνεύουν και να ταξινομούν με ακρίβεια τις δερματικές αντιδράσεις σε δοκιμασίες επιδερμικού επιθέματος. Ωστόσο, είναι σημαντικό να επισημανθούν οι περιορισμοί της μελέτης, συμπεριλαμβανομένου του μικρού μεγέθους του δείγματος και της πιθανής μεροληψίας του παρατηρητή. Απαιτούνται περαιτέρω μελέτες με μεγαλύτερα και πιο διαφορετικά σύνολα δεδομένων για την επικύρωση των ευρημάτων και τη βελτίωση της γενικευσιμότητας του μοντέλου. Επιπλέον, η μελέτη επικεντρώθηκε στη δυαδική ταξινόμηση, δηλαδή στη διάκριση μεταξύ θετικών και αρνητικών αντιδράσεων στα αλλεργιογόνα. Μελλοντικές μελέτες θα πρέπει να διερευνήσουν τη δυνατότητα πρόβλεψης και της σοβαρότητας της αντίδρασης. Τέλος, ενώ η μελέτη δεν έλαβε υπόψη το κλινικό ιστορικό του ασθενή, αλλά και πληροφορίες όπως η ηλικία και το φύλο, η συμπερίληψη τέτοιων πληροφοριών σε μελλοντικές μελέτες θα μπορούσε να συμβάλει στην βελτιστοποίηση και εξατομίκευση της διαγνωστικής διαδικασίας. Μειώνοντας την ανάγκη για αξιολόγηση από έναν κλινικό ιατρό, οι αυτοματοποιημένες τεχνικές έχουν τη δυνατότητα να βελτιώσουν σημαντικά την αποτελεσματικότητα και να μειώσουν τον φόρτο εργασίας, επιτρέποντας στο ιατρικό προσωπικό να επικεντρωθεί σε πιο σύνθετες περιπτώσεις. Επιπλέον, η αυτοματοποιημένη ανάγνωση των επιδερμικών δοκιμασιών επιθέματος μπορεί να παρέχει πιο αντικειμενικά και συνεπή αποτελέσματα, ελαχιστοποιώντας την επίδραση της μεροληψίας του παρατηρητή.

Στην τρίτη και τελευταία μελέτη της παρούσας διδακτορικής διατριβής, διερευνήθηκε η σκοπιμότητα της χρήσης μιας πλήρως αυτοματοποιημένης μεθόδου για την πρόγνωση του καρκίνου του παγκρέατος. Αυτή η μέθοδος έκανε χρήση ενός τρισδιάστατου συνελκτικού νευρωνικού δικτύου, συγκεκριμένα του nnU-Net, αρχικά για την κατάτμηση του παγκρέατος και του όγκου. Με βάση αυτές τις τμηματοποιήσεις, εξήχθησαν ραδιομικά χαρακτηριστικά και στη συνέχεια συνδυάστηκαν με κλινικά δεδομένα, συγκεκριμένα με την ηλικία των ασθενών και

την ταξινόμηση του όγκου κατά TNM. Χρησιμοποιώντας τυχαία δάση επιβίωσης, πραγματοποιήθηκε εκτίμηση της πιθανότητας επιβίωσης των ασθενών με την πάροδο του χρόνου. Επιπλέον, χρησιμοποιήθηκε ένας ταξινομητής τυχαίου δάσους για την αντιμετώπιση του ζητήματος της δυαδικής ταξινόμησης για την πρόβλεψη της διετούς επιβίωσης. Τα αποτελέσματα της μελέτης έδειξαν ότι η προτεινόμενη πλήρως αυτοματοποιημένη μέθοδος παρέχει αρκετά υποσχόμενες επιδόσεις για την πρόγνωση του καρκίνου. Συγκεκριμένα, η προσέγγιση παρουσίασε ανώτερες μετρήσεις απόδοσης σε σύγκριση με τις υπάρχουσες μεθόδους που βασίζονται αποκλειστικά σε κλινικές μεταβλητές, όπως η σταδιοποίηση TNM, με δείκτη CI 0,731. Επιπλέον, τα ευρήματα αυτής της μελέτης ήταν συνεπή με παρόμοιες μελέτες στις οποίες χρησιμοποιήθηκαν χειροκίνητες τμηματοποιήσεις, παρέχοντας μια ένδειξη της σκοπιμότητας και των πιθανών κερδών απόδοσης μιας πλήρως αυτοματοποιημένης προσέγγισης. Τα αποτελέσματα ανέδειξαν επίσης την κρίσιμη σημασία συγκεκριμένων χαρακτηριστικών, ιδίως του LAHGLE, υποδηλώνοντας περαιτέρω την προγνωστική του αξία για την εκτίμηση της επιβίωσης. Συμπερασματικά, η έρευνά παρείχε στοιχεία για τη σκοπιμότητα και τη δυνητική αποτελεσματικότητα μιας πλήρως αυτοματοποιημένης μεθόδου για την πρόγνωση του αδενοκαρκινώματος του παγκρέατος. Ωστόσο, απαιτείται περαιτέρω διερεύνηση και επικύρωση σε μεγαλύτερα και πιο διαφορετικά σύνολα δεδομένων για να τεκμηριωθεί η γενικευσιμότητα και η ευρωστία της, προκειμένου αυτή η προσέγγιση να προωθήσει μια πιο ακριβή, αποτελεσματική και εξατομικευμένη φροντίδα των ασθενών.

# Βιβλιογραφία

- [1] Committee on Diagnostic Error in Health Care, Board on Health Care Services, Institute of Medicine, and The National Academies of Sciences, Engineering, and Medicine, *Improving Diagnosis in Health Care*, E. P. Balogh, B. T. Miller, and J. R. Ball, Eds. Washington, D.C.: National Academies Press, Dec. 29, 2015, p. 21 794, ISBN: 978-0-309-37769-0. DOI: 10.17226/21794.
- [2] A. Vaswani, N. Shazeer, N. Parmar, *et al.* “Attention Is All You Need.” arXiv: 1706.03762 [cs]. (Dec. 5, 2017), [Online]. Available: <http://arxiv.org/abs/1706.03762> (visited on 07/05/2023), preprint.
- [3] D. J. Nolan, “100 years of x rays,” *BMJ (Clinical research ed.)*, vol. 310, no. 6980, pp. 614–615, Mar. 11, 1995, ISSN: 0959-8138. DOI: 10.1136/bmj.310.6980.614. pmid: 7703738.
- [4] R. Cierniak, *X-Ray Computed Tomography in Biomedical Engineering*. London: Springer London, 2011, ISBN: 978-0-85729-026-7 978-0-85729-027-4. DOI: 10.1007/978-0-85729-027-4.
- [5] J. H. Scatliff and P. J. Morris, “From Röntgen to Magnetic Resonance Imaging: The History of Medical Imaging,” *North Carolina Medical Journal*, vol. 75, no. 2, pp. 111–113, Mar. 2014, ISSN: 0029-2559, 0029-2559. DOI: 10.18043/ncm.75.2.111.
- [6] N. Farahani and L. Pantanowitz, “Overview of Telepathology,” *Surgical Pathology Clinics*, vol. 8, no. 2, pp. 223–231, Jun. 2015, ISSN: 18759181. DOI: 10.1016/j.path.2015.02.018.
- [7] B. J. Williams, D. Bottoms, and D. Treanor, “Future-proofing pathology: The case for clinical adoption of digital pathology,” *Journal of Clinical Pathology*, vol. 70, no. 12, pp. 1010–1018, Dec. 2017, ISSN: 0021-9746, 1472-4146. DOI: 10.1136/jclinpath-2017-204644.
- [8] W. S. Noble, “What is a support vector machine?” *Nature Biotechnology*, vol. 24, no. 12, pp. 1565–1567, Dec. 2006, ISSN: 1087-0156, 1546-1696. DOI: 10.1038/nbt1206-1565.
- [9] L. Breiman, “Random Forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, Oct. 1, 2001, ISSN: 1573-0565. DOI: 10.1023/A:1010933404324.

- [10] O. Kramer, “K-Nearest Neighbors,” in *Dimensionality Reduction with Un-supervised Nearest Neighbors*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, vol. 51, pp. 13–23, ISBN: 978-3-642-38651-0 978-3-642-38652-7. DOI: 10.1007/978-3-642-38652-7\_2.
- [11] J. Deng, W. Dong, R. Socher, L.-J. Li, Kai Li, and Li Fei-Fei, “ImageNet: A large-scale hierarchical image database,” in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, Miami, FL: IEEE, Jun. 2009, pp. 248–255, ISBN: 978-1-4244-3992-8. DOI: 10.1109/CVPR.2009.5206848.
- [12] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, May 24, 2017, ISSN: 0001-0782, 1557-7317. DOI: 10.1145/3065386.
- [13] H.-C. Shin, H. R. Roth, M. Gao, *et al.*, “Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning,” *IEEE Transactions on Medical Imaging*, vol. 35, no. 5, pp. 1285–1298, May 2016, ISSN: 0278-0062, 1558-254X. DOI: 10.1109/TMI.2016.2528162.
- [14] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” presented at the International Conference on Learning Representations, 2015.
- [15] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” Dec. 10, 2015. arXiv: 1512.03385 [cs].
- [16] M. Tan and Q. V. Le, “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks,” in *36th International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, vol. 97, Long Beach, California, USA, 2019, pp. 6105–6114.
- [17] S. J. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach* (Pearson Series in Artificial Intelligence), Fourth Edition, in collab. with M.-w. Chang, J. Devlin, A. Dragan, *et al.* Hoboken, NJ: Pearson, 2021, 1115 pp., ISBN: 978-0-13-461099-3.
- [18] K. Brodman, A. J. Van Woerkom, A. J. Erdmann, and L. S. Goldstein, “Interpretation of symptoms with a data-processing machine,” *A.M.A. Archives of Internal Medicine*, vol. 103, no. 5, pp. 776–782, May 1959, ISSN: 0888-2479. DOI: 10.1001/archinte.1959.00270050098015. pmid: 13636498.

- [19] C. J. Haug and J. M. Drazen, "Artificial Intelligence and Machine Learning in Clinical Medicine, 2023," *New England Journal of Medicine*, vol. 388, no. 13, J. M. Drazen, I. S. Kohane, and T.-Y. Leong, Eds., pp. 1201–1208, Mar. 30, 2023, ISSN: 0028-4793, 1533-4406.  
DOI: 10.1056/NEJMra2302038.
- [20] G. Letterie, "Artificial intelligence and assisted reproductive technologies: 2023. Ready for prime time? Or not," *Fertility and Sterility*, vol. 120, no. 1, pp. 32–37, Jul. 2023, ISSN: 00150282.  
DOI: 10.1016/j.fertnstert.2023.05.146.
- [21] J. N. Bishop, V. Bataille, A. Gavin, *et al.*, "The prevention, diagnosis, referral and management of melanoma of the skin: Concise guidelines," *Clinical Medicine (London, England)*, vol. 7, no. 3, pp. 283–290, Jun. 2007, ISSN: 1470-2118. DOI: 10.7861/clinmedicine.7-3-283. pmid: 17633952.
- [22] N. C. F. Codella, Q.-B. Nguyen, S. Pankanti, *et al.*, "Deep learning ensembles for melanoma recognition in dermoscopy images," *IBM Journal of Research and Development*, vol. 61, no. 4/5, 5:1–5:15, Jul. 1, 2017, ISSN: 0018-8646, 0018-8646. DOI: 10.1147/JRD.2017.2708299.
- [23] L. Yu, H. Chen, Q. Dou, J. Qin, and P.-A. Heng, "Automated Melanoma Recognition in Dermoscopy Images via Very Deep Residual Networks," *IEEE Transactions on Medical Imaging*, vol. 36, no. 4, pp. 994–1004, Apr. 2017, ISSN: 0278-0062, 1558-254X. DOI: 10.1109/TMI.2016.2642839.
- [24] M. Nawaz, Z. Mehmood, T. Nazir, *et al.*, "Skin cancer detection from dermoscopic images using deep learning and fuzzy  $k$ -means clustering," *Microscopy Research and Technique*, vol. 85, no. 1, pp. 339–351, Jan. 2022, ISSN: 1059-910X, 1097-0029. DOI: 10.1002/jemt.23908.
- [25] P. Lambin, R. T. Leijenaar, T. M. Deist, *et al.*, "Radiomics: The bridge between medical imaging and personalized medicine," *Nature Reviews Clinical Oncology*, vol. 14, no. 12, pp. 749–762, Dec. 2017, ISSN: 1759-4774, 1759-4782. DOI: 10.1038/nrclinonc.2017.141.
- [26] L. C. Chu, S. Park, S. Kawamoto, *et al.*, "Utility of CT Radiomics Features in Differentiation of Pancreatic Ductal Adenocarcinoma From Normal Pancreatic Tissue," *American Journal of Roentgenology*, vol. 213, no. 2, pp. 349–357, Aug. 2019, ISSN: 0361-803X, 1546-3141.  
DOI: 10.2214/AJR.18.20901.
- [27] S. Park, J. G. Sham, S. Kawamoto, *et al.*, "CT Radiomics-Based Preoperative Survival Prediction in Patients With Pancreatic Ductal Adenocarcinoma," *American Journal of Roentgenology*, vol. 217, no. 5, pp. 1104–1112, Nov. 2021, ISSN: 0361-803X, 1546-3141.  
DOI: 10.2214/AJR.20.23490.



- [28] N. Alves, M. Schuurmans, G. Litjens, J. S. Bosma, J. Hermans, and H. Huisman, “Fully Automatic Deep Learning Framework for Pancreatic Ductal Adenocarcinoma Detection on Computed Tomography,” *Cancers*, vol. 14, no. 2, p. 376, Jan. 13, 2022, ISSN: 2072-6694. DOI: 10.3390/cancers14020376. pmid: 35053538.
- [29] J. Van Der Laak, G. Litjens, and F. Ciompi, “Deep learning in histopathology: The path to the clinic,” *Nature Medicine*, vol. 27, no. 5, pp. 775–784, May 2021, ISSN: 1078-8956, 1546-170X. DOI: 10.1038/s41591-021-01343-4.
- [30] G. Litjens, C. I. Sánchez, N. Timofeeva, *et al.*, “Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis,” *Scientific Reports*, vol. 6, no. 1, p. 26286, May 23, 2016, ISSN: 2045-2322. DOI: 10.1038/srep26286.
- [31] X. Wang, H. Chen, C. Gan, *et al.*, “Weakly Supervised Deep Learning for Whole Slide Lung Cancer Image Analysis,” *IEEE transactions on cybernetics*, vol. 50, no. 9, pp. 3950–3962, Sep. 2020, ISSN: 2168-2275. DOI: 10.1109/TCYB.2019.2935141. pmid: 31484154.
- [32] M. S. Kwak, H. H. Lee, J. M. Yang, *et al.*, “Deep Convolutional Neural Network-Based Lymph Node Metastasis Prediction for Colon Cancer Using Histopathological Images,” *Frontiers in Oncology*, vol. 10, p. 619803, 2020, ISSN: 2234-943X. DOI: 10.3389/fonc.2020.619803. pmid: 33520727.
- [33] S. J. Wagner, D. Reisenbüchler, N. P. West, *et al.*, “Transformer-based biomarker prediction from colorectal cancer histology: A large-scale multicentric study,” *Cancer Cell*, vol. 41, no. 9, 1650–1661.e4, Sep. 2023, ISSN: 15356108. DOI: 10.1016/j.ccell.2023.08.002.
- [34] R. Kakino, M. Nakamura, T. Mitsuyoshi, *et al.*, “Application and limitation of radiomics approach to prognostic prediction for lung stereotactic body radiotherapy using breath-hold CT images with random survival forest: A multi-institutional study,” *Medical Physics*, vol. 47, no. 9, pp. 4634–4643, Sep. 2020, ISSN: 0094-2405, 2473-4209. DOI: 10.1002/mp.14380.
- [35] Y. Xu, A. Hosny, R. Zeleznik, *et al.*, “Deep Learning Predicts Lung Cancer Treatment Response from Serial Medical Imaging,” *Clinical Cancer Research*, vol. 25, no. 11, pp. 3266–3275, Jun. 1, 2019, ISSN: 1078-0432, 1557-3265. DOI: 10.1158/1078-0432.CCR-18-2495.
- [36] L. Lu, L. Dercle, B. Zhao, and L. H. Schwartz, “Deep learning for the prediction of early on-treatment response in metastatic colorectal cancer from serial medical imaging,” *Nature Communications*, vol. 12, no. 1,

- p. 6654, Nov. 17, 2021, ISSN: 2041-1723.  
DOI: 10.1038/s41467-021-26990-6.
- [37] P. Courtiol, C. Maussion, M. Moarii, *et al.*, “Deep learning-based classification of mesothelioma improves prediction of patient outcome,” *Nature Medicine*, vol. 25, no. 10, pp. 1519–1525, Oct. 2019, ISSN: 1078-8956, 1546-170X. DOI: 10.1038/s41591-019-0583-3.
- [38] D. Bychkov, N. Linder, R. Turkki, *et al.*, “Deep learning based tissue analysis predicts outcome in colorectal cancer,” *Scientific Reports*, vol. 8, no. 1, p. 3395, Feb. 21, 2018, ISSN: 2045-2322.  
DOI: 10.1038/s41598-018-21758-3.
- [39] A. Esteva, J. Feng, D. Van Der Wal, *et al.*, “Prostate cancer therapy personalization via multi-modal deep learning on randomized phase III clinical trials,” *npj Digital Medicine*, vol. 5, no. 1, p. 71, Jun. 8, 2022, ISSN: 2398-6352. DOI: 10.1038/s41746-022-00613-w.
- [40] 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, ISSN: 03057372.  
DOI: 10.1016/j.ctrv.2013.11.006.
- [41] 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. 1, 2020, ISSN: 25852787. DOI: 10.22540/JRPMS-04-015.
- [42] S. J. Taran, R. Taran, and N. B. Malipatil, “Pediatric Osteosarcoma: An Updated Review,” *Indian Journal of Medical and Paediatric Oncology*, vol. 38, no. 01, pp. 33–43, Jan. 2017, ISSN: 0971-5851, 0975-2129.  
DOI: 10.4103/0971-5851.203513.
- [43] N. A. Thomas, R. G. Abraham, B. Dedi, and N. A. Krucher, “Targeting retinoblastoma protein phosphorylation in combination with EGFR inhibition in pancreatic cancer cells,” *International Journal of Oncology*, vol. 54, no. 2, pp. 527–536, Feb. 2019, ISSN: 1791-2423.  
DOI: 10.3892/ijo.2018.4658. pmid: 30535494.
- [44] G. Matlashewski, P. Lamb, D. Pim, J. Peacock, L. Crawford, and S. Benchimol, “Isolation and characterization of a human p53 cDNA clone: Expression of the human p53 gene.,” *The EMBO Journal*, vol. 3, no. 13, pp. 3257–3262, Dec. 1984, ISSN: 02614189.  
DOI: 10.1002/j.1460-2075.1984.tb02287.x.
- [45] W. Chen, Q. Liu, B. Fu, K. Liu, and W. Jiang, “Overexpression of GRIM-19 accelerates radiation-induced osteosarcoma cells apoptosis by p53 stabilization,” *Life Sciences*, vol. 208, pp. 232–238, Sep. 2018, ISSN: 00243205.  
DOI: 10.1016/j.lfs.2018.07.015.

- [46] G. I. Lambrou, S. Vlahopoulos, C. Papathanasiou, *et al.*, “Prednisolone exerts late mitogenic and biphasic effects on resistant acute lymphoblastic leukemia cells: Relation to early gene expression,” *Leukemia Research*, vol. 33, no. 12, pp. 1684–1695, Dec. 2009, ISSN: 01452126. DOI: 10.1016/j.leukres.2009.04.018.
- [47] A. C. Miller, K. Kariko, C. E. Myers, E. P. Clark, and D. Samid, “Increased radioresistance of ejras-transformed human osteosarcoma cells and its modulation by lovastatin, an inhibitor of p21ras isoprenylation,” *International Journal of Cancer*, vol. 53, no. 2, pp. 302–307, Jan. 21, 1993, ISSN: 00207136, 10970215. DOI: 10.1002/ijc.2910530222.
- [48] A. Miller, J. Gafner, E. Clark, and D. Samid, “Differences in Radiation-induced Micronuclei Yields of Human Cells: Influence of *Ras* Gene Expression and Protein Localization,” *International Journal of Radiation Biology*, vol. 64, no. 5, pp. 547–554, Jan. 1993, ISSN: 0955-3002, 1362-3095. DOI: 10.1080/09553009314551761.
- [49] K. J. Campbell, N. R. Chapman, and N. D. Perkins, “UV stimulation induces nuclear factor  $\kappa$ B (NF- $\kappa$ B) DNA-binding activity but not transcriptional activation,” *Biochemical Society Transactions*, vol. 29, no. 6, pp. 688–691, Nov. 1, 2001, ISSN: 0300-5127, 1470-8752. DOI: 10.1042/bst0290688.
- [50] A. Chaussade, G. Millot, C. Wells, *et al.*, “Correlation between RB1 germline mutations and second primary malignancies in hereditary retinoblastoma patients treated with external beam radiotherapy,” *European Journal of Medical Genetics*, vol. 62, no. 3, pp. 217–223, Mar. 2019, ISSN: 17697212. DOI: 10.1016/j.ejmg.2018.07.017.
- [51] S. Surget, M. P. Khoury, and J.-C. Bourdon, “Uncovering the role of p53 splice variants in human malignancy: A clinical perspective,” *Oncotargets and Therapy*, vol. 7, pp. 57–68, Dec. 19, 2013, ISSN: 1178-6930. DOI: 10.2147/OTT.S53876. pmid: 24379683.
- [52] D. E. Park, J. Cheng, C. Berrios, *et al.*, “Dual inhibition of MDM2 and MDM4 in virus-positive Merkel cell carcinoma enhances the p53 response,” *Proceedings of the National Academy of Sciences*, vol. 116, no. 3, pp. 1027–1032, Jan. 15, 2019, ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1818798116.
- [53] T. D. Gilmore, “Introduction to NF- $\kappa$ B: Players, pathways, perspectives,” *Oncogene*, vol. 25, no. 51, pp. 6680–6684, Oct. 30, 2006, ISSN: 0950-9232, 1476-5594. DOI: 10.1038/sj.onc.1209954.

- [54] K. Taniguchi and M. Karin, "NF- $\kappa$ B, inflammation, immunity and cancer: Coming of age," *Nature Reviews Immunology*, vol. 18, no. 5, pp. 309–324, May 2018, ISSN: 1474-1733, 1474-1741.  
DOI: 10.1038/nri.2017.142.
- [55] S. A. Vlahopoulos, O. Cen, N. Hengen, *et al.*, "Dynamic aberrant NF- $\kappa$ B spurs tumorigenesis: A new model encompassing the microenvironment," *Cytokine & Growth Factor Reviews*, vol. 26, no. 4, pp. 389–403, Aug. 2015, ISSN: 13596101. DOI: 10.1016/j.cytogfr.2015.06.001.
- [56] M. Nouri, S. Massah, J. Caradec, *et al.*, "Transient Sox9 Expression Facilitates Resistance to Androgen-Targeted Therapy in Prostate Cancer," *Clinical Cancer Research*, vol. 26, no. 7, pp. 1678–1689, Apr. 1, 2020, ISSN: 1078-0432, 1557-3265. DOI: 10.1158/1078-0432.CCR-19-0098.
- [57] B. Zhang, Z.-L. Shi, B. Liu, X.-B. Yan, J. Feng, and H.-M. Tao, "Enhanced anticancer effect of gemcitabine by genistein in osteosarcoma: The role of Akt and nuclear factor- $\kappa$ B," *Anti-Cancer Drugs*, vol. 21, no. 3, pp. 288–296, Mar. 2010, ISSN: 0959-4973.  
DOI: 10.1097/CAD.0b013e328334da17.
- [58] I. Tsagaraki, E. C. Tsilibary, and A. K. Tzinia, "TIMP-1 interaction with  $\alpha$ v $\beta$ 3 integrin confers resistance to human osteosarcoma cell line MG-63 against TNF- $\alpha$ -induced apoptosis," *Cell and Tissue Research*, vol. 342, no. 1, pp. 87–96, Oct. 2010, ISSN: 0302-766X, 1432-0878.  
DOI: 10.1007/s00441-010-1025-1.
- [59] K. Li, X. Li, J. Tian, H. Wang, J. Pan, and J. Li, "Downregulation of DNA-PKcs suppresses P-gp expression via inhibition of the Akt/NF- $\kappa$ B pathway in CD133-positive osteosarcoma MG-63 cells," *Oncology Reports*, vol. 36, no. 4, pp. 1973–1980, Oct. 2016, ISSN: 1021-335X, 1791-2431.  
DOI: 10.3892/or.2016.4991.
- [60] M. Yan, J. Ni, D. Song, M. Ding, and J. Huang, "Activation of unfolded protein response protects osteosarcoma cells from cisplatin-induced apoptosis through NF- $\kappa$ B pathway," *International Journal of Clinical and Experimental Pathology*, vol. 8, no. 9, pp. 10204–10215, 2015, ISSN: 1936-2625. PMID: 26617729.
- [61] F. Jafari, S. Javdansirat, S. Sanaie, *et al.*, "Osteosarcoma: A comprehensive review of management and treatment strategies," *Annals of Diagnostic Pathology*, vol. 49, p. 151654, Dec. 2020, ISSN: 10929134.  
DOI: 10.1016/j.anndiagpath.2020.151654.
- [62] R. L. Siegel, K. D. Miller, H. E. Fuchs, and A. Jemal, "Cancer statistics, 2022," *CA: A Cancer Journal for Clinicians*, vol. 72, no. 1, pp. 7–33, Jan. 2022, ISSN: 0007-9235, 1542-4863. DOI: 10.3322/caac.21708.

- [63] S. Smeland, S. S. Bielack, J. Whelan, *et al.*, “Survival and prognosis with osteosarcoma: Outcomes in more than 2000 patients in the EURAMOS-1 (European and American Osteosarcoma Study) cohort,” *European Journal of Cancer*, vol. 109, pp. 36–50, Mar. 2019, ISSN: 09598049. DOI: 10.1016/j.ejca.2018.11.027.
- [64] J. C. Friebele, J. Peck, X. Pan, M. Abdel-Rasoul, and J. L. Mayerson, “Osteosarcoma: A Meta-Analysis and Review of the Literature,” *American Journal of Orthopedics (Belle Mead, N.J.)*, vol. 44, no. 12, pp. 547–553, Dec. 2015, ISSN: 1934-3418. PMID: 26665241.
- [65] Y. Jiang, X. Wang, Y. Cheng, *et al.*, “Associations between inflammatory gene polymorphisms (TNF- $\alpha$  308G/A, TNF- $\alpha$  238G/A, TNF- $\beta$  252A/G, TGF- $\beta$ 1 29T/C, IL-6 174G/C and IL-10 1082A/G) and susceptibility to osteosarcoma: A meta-analysis and literature review,” *Oncotarget*, vol. 8, no. 57, pp. 97 571–97 583, Nov. 14, 2017, ISSN: 1949-2553. DOI: 10.18632/oncotarget.18813.
- [66] J. Bajpai, R. Kumar, V. Sreenivas, *et al.*, “Prediction of Chemotherapy Response by PET-CT in Osteosarcoma: Correlation With Histologic Necrosis,” *Journal of Pediatric Hematology/Oncology*, vol. 33, no. 7, e271–e278, Oct. 2011, ISSN: 1077-4114. DOI: 10.1097/MPH.0b013e31820ff29e.
- [67] J. L. Emerson, “Observer Bias in Histopathological Examinations,” in *Carcinogenicity*, H. C. Grice and J. L. Ciminera, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 1988, pp. 137–147, ISBN: 978-3-642-64810-6 978-3-642-61364-7. DOI: 10.1007/978-3-642-61364-7\_17.
- [68] A. Asilian Bidgoli, S. Rahnamayan, T. Dehkharghanian, A. Grami, and H. Tizhoosh, “Bias reduction in representation of histopathology images using deep feature selection,” *Scientific Reports*, vol. 12, no. 1, p. 19994, Nov. 21, 2022, ISSN: 2045-2322. DOI: 10.1038/s41598-022-24317-z.
- [69] C. Bührenmann, S. Li, H. Yu, *et al.*, “Quantification of the Heterogeneity of Prognostic Cellular Biomarkers in Ewing Sarcoma Using Automated Image and Random Survival Forest Analysis,” *PLoS ONE*, vol. 9, no. 9, K. Metze, Ed., e107105, Sep. 22, 2014, ISSN: 1932-6203. DOI: 10.1371/journal.pone.0107105.
- [70] E. Essa, Xianghua Xie, R. J. Errington, and N. White, “A multi-stage random forest classifier for phase contrast cell segmentation,” in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Milan: IEEE, Aug. 2015, pp. 3865–3868, ISBN: 978-1-4244-9271-8. DOI: 10.1109/EMBC.2015.7319237.

- [71] W. Li, Y. Dong, W. Liu, *et al.*, “A deep belief network-based clinical decision system for patients with osteosarcoma,” *Frontiers in Immunology*, vol. 13, p. 1003347, Nov. 18, 2022, ISSN: 1664-3224. DOI: 10.3389/fimmu.2022.1003347.
- [72] R. Shen, Z. Li, L. Zhang, *et al.*, “Osteosarcoma Patients Classification Using Plain X-Rays and Metabolomic Data,” in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Honolulu, HI: IEEE, Jul. 2018, pp. 690–693, ISBN: 978-1-5386-3646-6. DOI: 10.1109/EMBC.2018.8512338.
- [73] H. B. Arunachalam, R. Mishra, O. Daescu, *et al.*, “Viable and necrotic tumor assessment from whole slide images of osteosarcoma using machine-learning and deep-learning models,” *PLOS ONE*, vol. 14, no. 4, K. Li, Ed., e0210706, Apr. 17, 2019, ISSN: 1932-6203. DOI: 10.1371/journal.pone.0210706.
- [74] D. Komura and S. Ishikawa, “Machine Learning Methods for Histopathological Image Analysis,” *Computational and Structural Biotechnology Journal*, vol. 16, pp. 34–42, 2018, ISSN: 20010370. DOI: 10.1016/j.csbj.2018.01.001.
- [75] S. K. Melanthota, D. Gopal, S. Chakrabarti, A. A. Kashyap, R. Radhakrishnan, and N. Mazumder, “Deep learning-based image processing in optical microscopy,” *Biophysical Reviews*, vol. 14, no. 2, pp. 463–481, Apr. 2022, ISSN: 1867-2450, 1867-2469. DOI: 10.1007/s12551-022-00949-3.
- [76] S. Ong, X. Jin, Jayasooriah, and R. Sinniah, “Image analysis of tissue sections,” *Computers in Biology and Medicine*, vol. 26, no. 3, pp. 269–279, May 1996, ISSN: 00104825. DOI: 10.1016/0010-4825(96)00004-2.
- [77] F. Wang, T. W. Oh, C. Vergara-Niedermayr, T. Kurc, and J. Saltz, “Managing and Querying Whole Slide Images,” *Proceedings of SPIE-the International Society for Optical Engineering*, vol. 8319, 83190J, Feb. 16, 2012, ISSN: 0277-786X. DOI: 10.1117/12.912388. pmid: 22844574.
- [78] A. G. Howard, M. Zhu, B. Chen, *et al.*, “MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications,” version 1, 2017. DOI: 10.48550/ARXIV.1704.04861.
- [79] A. Dosovitskiy, L. Beyer, A. Kolesnikov, *et al.* “An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale.” arXiv: 2010.11929 [cs]. (Jun. 3, 2021), [Online]. Available: <http://arxiv.org/abs/2010.11929> (visited on 11/25/2022), preprint.

- [80] D. Anisuzzaman, H. Barzekar, L. Tong, J. Luo, and Z. Yu, "A deep learning study on osteosarcoma detection from histological images," *Biomedical Signal Processing and Control*, vol. 69, p. 102931, Aug. 2021, ISSN: 17468094. DOI: 10.1016/j.bspc.2021.102931.
- [81] D. P. Kingma and J. Ba. "Adam: A Method for Stochastic Optimization." arXiv: 1412.6980 [cs]. (Jan. 29, 2017), [Online]. Available: <http://arxiv.org/abs/1412.6980> (visited on 07/05/2023), preprint.
- [82] I. Loshchilov and F. Hutter, "Decoupled Weight Decay Regularization," presented at the International Conference on Learning Representations, ser. International Conference on Learning Representations, 2019.
- [83] R. Mishra, O. Daescu, P. Leavey, D. Rakheja, and A. Sengupta, "Convolutional Neural Network for Histopathological Analysis of Osteosarcoma," *Journal of Computational Biology*, vol. 25, no. 3, pp. 313–325, Mar. 2018, ISSN: 1557-8666. DOI: 10.1089/cmb.2017.0153.
- [84] W. Luo, Y. Li, R. Urtasun, and R. Zemel, "Understanding the Effective Receptive Field in Deep Convolutional Neural Networks," *Advances in neural information processing systems*, vol. 29, 2016. arXiv: 1701.04128 [cs].
- [85] I. E. Tampu, A. Eklund, and N. Haj-Hosseini, "Inflation of test accuracy due to data leakage in deep learning-based classification of OCT images," *Scientific Data*, vol. 9, no. 1, p. 580, Sep. 22, 2022, ISSN: 2052-4463. DOI: 10.1038/s41597-022-01618-6.
- [86] P. Bansal, K. Gehlot, A. Singhal, and A. Gupta, "Automatic detection of osteosarcoma based on integrated features and feature selection using binary arithmetic optimization algorithm," *Multimedia Tools and Applications*, vol. 81, no. 6, pp. 8807–8834, Mar. 2022, ISSN: 1380-7501, 1573-7721. DOI: 10.1007/s11042-022-11949-6.
- [87] I. Kimber, D. A. Basketter, G. Gerberick, and R. J. Dearman, "Allergic contact dermatitis," *International Immunopharmacology*, vol. 2, no. 2-3, pp. 201–211, Feb. 2002, ISSN: 15675769. DOI: 10.1016/S1567-5769(01)00173-4.
- [88] J. L. Nelson and C. M. Mowad, "Allergic Contact Dermatitis: Patch Testing Beyond the TRUE Test," *The Journal of Clinical and Aesthetic Dermatology*, vol. 3, no. 10, pp. 36–41, Oct. 2010, ISSN: 2689-9175. pmid: 20967194.
- [89] D. E. Cohen, "Contact dermatitis: A quarter century perspective," *Journal of the American Academy of Dermatology*, vol. 51, no. 1, pp. 60–63, Jul. 2004, ISSN: 01909622. DOI: 10.1016/j.jaad.2003.01.002.

- [90] C. M. Mowad, B. Anderson, P. Scheinman, S. Pootongkam, S. Nedorost, and B. Brod, "Allergic contact dermatitis," *Journal of the American Academy of Dermatology*, vol. 74, no. 6, pp. 1043–1054, Jun. 2016, ISSN: 01909622. DOI: 10.1016/j.jaad.2015.02.1144.
- [91] M. Lindberg and M. Matura, "Patch Testing," in *Contact Dermatitis*, J. D. Johansen, P. J. Frosch, and J.-P. Lepoittevin, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, pp. 439–464, ISBN: 978-3-642-03826-6 978-3-642-03827-3. DOI: 10.1007/978-3-642-03827-3\_24.
- [92] J. D. Johansen, K. Aalto-Korte, T. Agner, *et al.*, "European Society of Contact Dermatitis guideline for diagnostic patch testing - recommendations on best practice: ESCD PATCH TEST GUIDELINE," *Contact Dermatitis*, vol. 73, no. 4, pp. 195–221, Oct. 2015, ISSN: 01051873. DOI: 10.1111/cod.12432.
- [93] S. Suganyadevi, V. Seethalakshmi, and K. Balasamy, "A review on deep learning in medical image analysis," *International Journal of Multimedia Information Retrieval*, vol. 11, no. 1, pp. 19–38, Mar. 2022, ISSN: 2192-6611, 2192-662X. DOI: 10.1007/s13735-021-00218-1.
- [94] S. K. Zhou, H. Greenspan, C. Davatzikos, *et al.*, "A Review of Deep Learning in Medical Imaging: Imaging Traits, Technology Trends, Case Studies With Progress Highlights, and Future Promises," *Proceedings of the IEEE*, vol. 109, no. 5, pp. 820–838, May 2021, ISSN: 0018-9219, 1558-2256. DOI: 10.1109/JPROC.2021.3054390.
- [95] Y. Yang, Y.-X. Li, R.-Q. Yao, X.-H. Du, and C. Ren, "Artificial intelligence in small intestinal diseases: Application and prospects," *World Journal of Gastroenterology*, vol. 27, no. 25, pp. 3734–3747, Jul. 7, 2021, ISSN: 1007-9327. DOI: 10.3748/wjg.v27.i25.3734.
- [96] W. Chan, R. Srivastava, N. Damaraju, *et al.*, "Automated detection of skin reactions in epicutaneous patch testing using machine learning," *British Journal of Dermatology*, vol. 185, no. 2, pp. 456–458, Aug. 2021, ISSN: 0007-0963, 1365-2133. DOI: 10.1111/bjd.20141.
- [97] D. TENNSTEDT, « Tests épicutanés : indications ou quand les réaliser ? », *Annales de Dermatologie et de Vénérologie*, t. 136, n° 8-9, p. 579-583, août 2009, ISSN : 01519638. DOI : 10.1016/j.annder.2009.06.012.
- [98] R. Lazzarini, I. Duarte, and A. L. Ferreira, "Patch tests," *Anais Brasileiros de Dermatologia*, vol. 88, no. 6, pp. 879–888, Dec. 2013, ISSN: 0365-0596. DOI: 10.1590/abd1806-4841.20132323.
- [99] N. Vigneshkarthik, "Patch Test as a Diagnostic Tool in Hand Eczema," *JOURNAL OF CLINICAL AND DIAGNOSTIC RESEARCH*, 2016, ISSN: 2249782X. DOI: 10.7860/JCDR/2016/23994.8884.



- [100] D. Rosmarin, A. B. Gottlieb, A. Asarch, and P. L. Scheinman, "Patch-Testing While on Systemic Immunosuppressants:" *Dermatitis*, vol. 20, no. 5, pp. 265–270, Sep. 2009, ISSN: 1710-3568.  
DOI: 10.2310/6620.2009.09005.
- [101] S. Seite, H. Zucchi, D. Moyal, *et al.*, "Alterations in human epidermal Langerhans cells by ultraviolet radiation: Quantitative and morphological study," *British Journal of Dermatology*, vol. 148, no. 2, pp. 291–299, Feb. 2003, ISSN: 0007-0963, 1365-2133.  
DOI: 10.1046/j.1365-2133.2003.05112.x.
- [102] M. Wilkinson, M. Gonçalo, O. Aerts, *et al.*, "The European baseline series and recommended additions: 2019," *Contact Dermatitis*, vol. 80, no. 1, pp. 1–4, Jan. 2019, ISSN: 01051873. DOI: 10.1111/cod.13155.
- [103] F. Linming, H. Wei, L. Anqi, *et al.*, "Comparison of two skin imaging analysis instruments: The VISIA<sup>®</sup> from Canfield vs the ANTERA 3D<sup>®</sup> CS from Miravex," *Skin Research and Technology*, vol. 24, no. 1, pp. 3–8, Feb. 2018, ISSN: 0909752X. DOI: 10.1111/srt.12381.
- [104] A. Kyritsi, S. Kikionis, A. Tagka, *et al.*, "Management of Acute Radio-dermatitis in Non-Melanoma Skin Cancer Patients Using Electrospun Nanofibrous Patches Loaded with Pinus halepensis Bark Extract," *Cancers*, vol. 13, no. 11, p. 2596, May 26, 2021, ISSN: 2072-6694.  
DOI: 10.3390/cancers13112596.
- [105] A. R. Matias, M. Ferreira, P. Costa, and P. Neto, "Skin colour, skin redness and melanin biometric measurements: Comparison study between Antera<sup>®</sup> 3D, Mexameter<sup>®</sup> and Colorimeter<sup>®</sup>," *Skin Research and Technology*, vol. 21, no. 3, pp. 346–362, Aug. 2015, ISSN: 0909752X.  
DOI: 10.1111/srt.12199.
- [106] N. M. Nawi, W. H. Atomi, and M. Rehman, "The Effect of Data Pre-processing on Optimized Training of Artificial Neural Networks," *Procedia Technology*, vol. 11, pp. 32–39, 2013, ISSN: 22120173.  
DOI: 10.1016/j.protcy.2013.12.159.
- [107] H. M. Bui, M. Lech, E. Cheng, K. Neville, and I. S. Burnett, "Using grayscale images for object recognition with convolutional-recursive neural network," in *2016 IEEE Sixth International Conference on Communications and Electronics (ICCE)*, Ha-Long City, Quang Ninh Province, Vietnam: IEEE, Jul. 2016, pp. 321–325, ISBN: 978-1-5090-1801-7.  
DOI: 10.1109/CCE.2016.7562656.
- [108] T. G. Dietterich, "Approximate Statistical Tests for Comparing Supervised Classification Learning Algorithms," *Neural Computation*, vol. 10, no. 7, pp. 1895–1923, Oct. 1, 1998, ISSN: 0899-7667, 1530-888X.  
DOI: 10.1162/089976698300017197.

- [109] E. Alpaydın, “Combined  $5 \times 2$  cv F Test for Comparing Supervised Classification Learning Algorithms,” *Neural Computation*, vol. 11, no. 8, pp. 1885–1892, Nov. 1, 1999, ISSN: 0899-7667, 1530-888X. DOI: 10.1162/089976699300016007.
- [110] *SEER Cancer Stat Facts: Pancreatic Cancer, 2023*.
- [111] J. D. Mizrahi, R. Surana, J. W. Valle, and R. T. Shroff, “Pancreatic cancer,” *The Lancet*, vol. 395, no. 10242, pp. 2008–2020, Jun. 2020, ISSN: 01406736. DOI: 10.1016/S0140-6736(20)30974-0.
- [112] S. Kaur, M. J. Baine, M. Jain, A. R. Sasson, and S. K. Batra, “Early diagnosis of pancreatic cancer: Challenges and new developments,” *Biomarkers in Medicine*, vol. 6, no. 5, pp. 597–612, Oct. 2012, ISSN: 1752-0371. DOI: 10.2217/bmm.12.69. pmid: 23075238.
- [113] N Howlader, g.-i. family=Noone given=AM, M Krapcho, *et al.*, *SEER Cancer Statistics Review, 1975-2014*, National Cancer Institute, Apr. 2017.
- [114] A. Haj-Mirzaian, S. Kawamoto, A. Zaheer, R. H. Hruban, E. K. Fishman, and L. C. Chu, “Pitfalls in the MDCT of pancreatic cancer: Strategies for minimizing errors,” *Abdominal Radiology*, vol. 45, no. 2, pp. 457–478, Feb. 2020, ISSN: 2366-004X, 2366-0058. DOI: 10.1007/s00261-019-02390-9.
- [115] D. V. Sahani, Z. K. Shah, O. A. Catalano, G. W. Boland, and W. R. Brugge, “Radiology of pancreatic adenocarcinoma: Current status of imaging: Diagnostic imaging: Pancreatic adenocarcinoma,” *Journal of Gastroenterology and Hepatology*, vol. 23, no. 1, pp. 23–33, Dec. 13, 2007, ISSN: 08159319. DOI: 10.1111/j.1440-1746.2007.05117.x.
- [116] Y.-T. Chen, Z.-P. Huang, Z.-W. Zhou, and M.-M. He, “Equipping the American Joint Committee on Cancer staging for resectable pancreatic ductal adenocarcinoma with tumor grade: A recursive partitioning analysis,” *Medical Oncology*, vol. 33, no. 11, p. 122, Nov. 2016, ISSN: 1357-0560, 1559-131X. DOI: 10.1007/s12032-016-0839-4.
- [117] W. Bakasa and S. Viriri, “Pancreatic Cancer Survival Prediction: A Survey of the State-of-the-Art,” *Computational and Mathematical Methods in Medicine*, vol. 2021, H. Chen, Ed., pp. 1–17, Sep. 30, 2021, ISSN: 1748-6718, 1748-670X. DOI: 10.1155/2021/1188414.
- [118] F. Isensee, P. F. Jaeger, S. A. A. Kohl, J. Petersen, and K. H. Maier-Hein, “nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation,” *Nature Methods*, vol. 18, no. 2, pp. 203–211, Feb. 2021, ISSN: 1548-7091, 1548-7105. DOI: 10.1038/s41592-020-01008-z.

- [119] M. Antonelli, A. Reinke, S. Bakas, *et al.*, “The Medical Segmentation Decathlon,” *Nature Communications*, vol. 13, no. 1, p. 4128, Jul. 15, 2022, ISSN: 2041-1723. DOI: 10.1038/s41467-022-30695-9.
- [120] J. J. Van Griethuysen, A. Fedorov, C. Parmar, *et al.*, “Computational Radiomics System to Decode the Radiographic Phenotype,” *Cancer Research*, vol. 77, no. 21, e104–e107, Nov. 1, 2017, ISSN: 0008-5472, 1538-7445. DOI: 10.1158/0008-5472.CAN-17-0339.
- [121] H. Ishwaran, U. B. Kogalur, E. H. Blackstone, and M. S. Lauer, “Random survival forests,” *The Annals of Applied Statistics*, vol. 2, no. 3, Sep. 1, 2008, ISSN: 1932-6157. DOI: 10.1214/08-AOAS169.
- [122] S. S. F. Yip and H. J. W. L. Aerts, “Applications and limitations of radiomics,” *Physics in Medicine and Biology*, vol. 61, no. 13, R150–R166, Jul. 7, 2016, ISSN: 0031-9155, 1361-6560. DOI: 10.1088/0031-9155/61/13/R150.
- [123] R. J. Gillies, P. E. Kinahan, and H. Hricak, “Radiomics: Images Are More than Pictures, They Are Data,” *Radiology*, vol. 278, no. 2, pp. 563–577, Feb. 2016, ISSN: 0033-8419, 1527-1315. DOI: 10.1148/radiol.2015151169.
- [124] S. B. Edge και A. J. C. on Cancer, επιμελητές, *AJCC Cancer Staging Manual*, 7th ed. New York: Springer, 2010, 648 **pagetotals**, ISBN: 978-0-387-88440-0 978-0-387-88442-4.
- [125] P. K. Andersen, “Fifty Years with the Cox Proportional Hazards Regression Model,” *Journal of the Indian Institute of Science*, vol. 102, no. 4, pp. 1135–1144, Oct. 2022, ISSN: 0970-4140, 0019-4964. DOI: 10.1007/s41745-021-00283-9.
- [126] F. E. Harrell, “Evaluating the Yield of Medical Tests,” *JAMA: The Journal of the American Medical Association*, vol. 247, no. 18, p. 2543, May 14, 1982, ISSN: 0098-7484. DOI: 10.1001/jama.1982.03320430047030.
- [127] M. B. Amin, F. L. Greene, S. B. Edge, *et al.*, “The Eighth Edition AJCC Cancer Staging Manual: Continuing to build a bridge from a population-based to a more “personalized” approach to cancer staging: The Eighth Edition AJCC Cancer Staging Manual,” *CA: A Cancer Journal for Clinicians*, vol. 67, no. 2, pp. 93–99, Mar. 2017, ISSN: 00079235. DOI: 10.3322/caac.21388.
- [128] F. Mohammad, M. Hanks, K. Kandiah, *et al.*, “O151 A comparison of pathological TNM 7 with TNM 8 staging for pancreatic ductal adenocarcinoma and distal cholangiocarcinoma,” *British Journal of Surgery*, vol. 110, znad101.151, Supplement\_3 May 10, 2023, ISSN: 0007-1323, 1365-2168. DOI: 10.1093/bjs/znad101.151.

- [129] T. Xie, X. Wang, M. Li, T. Tong, X. Yu, and Z. Zhou, “Pancreatic ductal adenocarcinoma: A radiomics nomogram outperforms clinical model and TNM staging for survival estimation after curative resection,” *European Radiology*, vol. 30, no. 5, pp. 2513–2524, May 2020, ISSN: 1432-1084. DOI: 10.1007/s00330-019-06600-2. PMID: 32006171.
- [130] G. M. Healy, E. Salinas-Miranda, R. Jain, *et al.*, “Pre-operative radiomics model for prognostication in resectable pancreatic adenocarcinoma with external validation,” *European Radiology*, vol. 32, no. 4, pp. 2492–2505, Apr. 2022, ISSN: 0938-7994, 1432-1084. DOI: 10.1007/s00330-021-08314-w.
- [131] Y. Zhang, E. M. Lobo-Mueller, P. Karanicolas, S. Gallinger, M. A. Haider, and F. Khalvati, “CNN-based survival model for pancreatic ductal adenocarcinoma in medical imaging,” *BMC Medical Imaging*, vol. 20, no. 1, p. 11, Dec. 2020, ISSN: 1471-2342. DOI: 10.1186/s12880-020-0418-1.
- [132] C. Scapicchio, M. Gabelloni, A. Barucci, D. Cioni, L. Saba, and E. Neri, “A deep look into radiomics,” *La Radiologia Medica*, vol. 126, no. 10, pp. 1296–1311, Oct. 2021, ISSN: 1826-6983. DOI: 10.1007/s11547-021-01389-x. PMID: 34213702.
- [133] M. H. Osman, “Abstract 1644: Pancreatic cancer survival prediction using machine learning and comparing its performance with TNM staging system and prognostic nomograms,” *Cancer Research*, vol. 79, pp. 1644–1644, 13\_Supplement Jul. 1, 2019, ISSN: 0008-5472, 1538-7445. DOI: 10.1158/1538-7445.AM2019-1644.
- [134] J. Chakraborty, L. Langdon-Embry, K. M. Cunanan, *et al.*, “Preliminary study of tumor heterogeneity in imaging predicts two year survival in pancreatic cancer patients,” *PLOS ONE*, vol. 12, no. 12, S. K. Batra, Ed., e0188022, Dec. 7, 2017, ISSN: 1932-6203. DOI: 10.1371/journal.pone.0188022.
- [135] X. Ying, “An Overview of Overfitting and its Solutions,” *Journal of Physics: Conference Series*, vol. 1168, p. 022022, Feb. 2019, ISSN: 1742-6588, 1742-6596. DOI: 10.1088/1742-6596/1168/2/022022.