



ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΜΑΚΕΔΟΝΙΑΣ

ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ
ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ &
ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ



Σχεδίαση και Υλοποίηση Συστήματος Βαθιάς Μάθησης για την Ταξινόμηση Όγκων Μαστού από Εικόνες Υπερήχων και Ανάπτυξη Εφαρμογής υποστήριξης Ιατρικής Διάγνωσης

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΤΟΥ

ΚΩΝΣΤΑΝΤΙΝΟΥ ΠΑΠΑΘΑΝΑΣΙΟΥ

Επιβλέπων: Δρ. Μηνάς Δασυγένης

Αναπληρωτής Καθηγητής

ΚΟΖΑΝΗ/ΙΟΥΝΙΟΣ/2026



HELLENIC DEMOCRACY
UNIVERSITY OF WESTERN MACEDONIA

FUCULTY OF ENGINEERING
DEPARTMENT OF ELECTRICAL &
COMPUTER ENGINEERING



Design and Implementation of a Deep Learning System for Breast Tumor Classification from Ultrasound Images and Development of a Medical Diagnosis Support Application

THESIS

KONSTANTINOS PAPATHANASIOU

SUPERVISOR: Dr. Minas Dasygenis

Associate Professor

KOZANI/JUNE/2026



ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΔΥΤΙΚΗΣ ΜΑΚΕΔΟΝΙΑΣ
ΠΟΛΥΤΕΧΝΙΚΗ ΣΧΟΛΗ
ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ
& ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΔΗΛΩΣΗ ΜΗ ΛΟΓΟΚΛΟΠΗΣ ΚΑΙ ΑΝΑΛΗΨΗΣ ΠΡΟΣΩΠΙΚΗΣ ΕΥΘΥΝΗΣ

Δηλώνω ρητά ότι, σύμφωνα με το άρθρο 8 του Ν. 1599/1986 και τα άρθρα 2,4,6 παρ. 3 του Ν. 1256/1982, η παρούσα Διπλωματική Εργασία με τίτλο “ Σχεδίαση και Υλοποίηση Συστήματος Βαθιάς Μάθησης για την Ταξινόμηση Όγκων Μαστού από Εικόνες Υπερήχων και Ανάπτυξη Εφαρμογής υποστήριξης Ιατρικής Διάγνωσης ” καθώς και τα ηλεκτρονικά αρχεία και πηγαίοι κώδικες που αναπτύχθηκαν ή τροποποιήθηκαν στα πλαίσια αυτής της εργασίας και αναφέρονται ρητώς μέσα στο κείμενο που συνοδεύουν, και η οποία έχει εκπονηθεί στο Τμήμα Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών του Πανεπιστημίου Δυτικής Μακεδονίας, υπό την επίβλεψη του μέλους του Τμήματος κ. “Μηνά Δασυγένη” αποτελεί αποκλειστικά προϊόν προσωπικής εργασίας και δεν προσβάλλει κάθε μορφής πνευματικά δικαιώματα τρίτων και δεν είναι προϊόν μερικής ή ολικής αντιγραφής, οι πηγές δε που χρησιμοποιήθηκαν περιορίζονται στις βιβλιογραφικές αναφορές και μόνον. Τα σημεία όπου έχω χρησιμοποιήσει ιδέες, κείμενο, αρχεία ή / και πηγές άλλων συγγραφέων, αναφέρονται ευδιάκριτα στο κείμενο με την κατάλληλη παραπομπή και η σχετική αναφορά περιλαμβάνεται στο τμήμα των βιβλιογραφικών αναφορών με πλήρη περιγραφή. 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 μόνο.

Copyright (C) “Κωνσταντίνος Παπαθανασίου & Μηνάς Δασυγένης” 2026, Κοζάνη

Copyright (C) _____, _____, _____, _____

Υπογραφή Φοιτητή: _____

Περίληψη

Ο καρκίνος του μαστού είναι η συχνότερη κακοήθεια στις γυναίκες παγκοσμίως, γεγονός που καθιστά την έγκαιρη διάγνωση άμεση κλινική προτεραιότητα. Το υπερηχογράφημα μαστού (*Breast Ultrasound, BUS*) αποτελεί ευρέως διαδεδομένη διαγνωστική μέθοδο, ωστόσο η ερμηνεία του εξαρτάται σε μεγάλο βαθμό από την εμπειρία του εξεταστή. Τα συστήματα Υποβοηθούμενης Διάγνωσης (*Computer-Aided Diagnosis, CAD*) που βασίζονται σε Βαθιά Μάθηση (*Deep Learning, DL*) έχουν αποδώσει ικανοποιητικά σε αυτό το πρόβλημα, αλλά οι αρχιτεκτονικές που χρησιμοποιούνται συνήθως είναι υπολογιστικά βαριές, κάτι που δυσκολεύει την εφαρμογή τους σε φορητές συσκευές. Η παρούσα εργασία αναπτύσσει και αξιολογεί Συνελικτικά Νευρωνικά Δίκτυα χαμηλής πολυπλοκότητας (*Lightweight Convolutional Neural Networks, Lightweight CNNs*) για την αυτόματη ταξινόμηση υπερηχογραφικών εικόνων σε καλοήθεις και κακοήθεις, με έμφαση στην Αποδοτική Τεχνητή Νοημοσύνη (*Efficient AI*).

Για τον σκοπό αυτό, ενοποιήθηκαν δύο σύνολα δεδομένων: το δημόσιο *BUSI* και το πιο πρόσφατο *BUS-BRA*. Οι κλινικές κατηγορίες του Συστήματος *BI-RADS* ομαδοποιήθηκαν ώστε να ανταποκρίνονται στο δυαδικό πρόβλημα ταξινόμησης. Ένα από τα κεντρικά στοιχεία της μεθοδολογίας ήταν η στρατηγική προ-επεξεργασίας: αξιοποιώντας τις διαθέσιμες μάσκες τμηματοποίησης, εφαρμόστηκε διαστολή και απόκρυψη του φόντου, με αποτέλεσμα να απομονωθεί η Περιοχή Ενδιαφέροντος (*Region of Interest, ROI*) από τον οπτικό θόρυβο. Αξιολογήθηκαν τρία ελαφριά μοντέλα — το *MobileNetV2*, το *EfficientNetB0* και το *NASNetMobile* — μέσω Μεταφοράς Μάθησης (*Transfer Learning, TL*), ενώ για την αξιοπιστία των αποτελεσμάτων χρησιμοποιήθηκε διασταυρούμενη επικύρωση 5 πτυχών με ομαδοποίηση ανά ασθενή. Επιπρόσθετα, δοκιμάστηκαν τεχνικές όπως η Ενίσχυση Δεδομένων κατά την Πρόβλεψη (*Test-Time Augmentation, TTA*), η δυναμική Βελτιστοποίηση Κατωφλίου και, σε πειραματικό επίπεδο, η ενσωμάτωση Μηχανισμού Προσοχής (*Convolutional Block Attention Module, CBAM*).

Τα αποτελέσματα καταδεικνύουν ότι τα ελαφριά (*lightweight*) μοντέλα, σε συνδυασμό με στοχευμένη προ-επεξεργασία ("Input E"), επιτυγχάνουν διαγνωστική απόδοση εφάμιλλη πολύ βαρύτερων αρχιτεκτονικών. Το MobileNetV2 αναδείχθηκε ως το πιο σταθερό μοντέλο, με συνολική ακρίβεια 97,95% (AUC 0,9974) στο σύνολο BUSI και 89,49% (AUC 0,9387) στο σαφώς πιο απαιτητικό, πολυκεντρικό BUS-BRA. Η πειραματική ενσωμάτωση του μηχανισμού CBAM αύξησε περαιτέρω την ακρίβεια στο BUSI στο 98,85% και ενίσχυσε κλινικά τη διακριτική ικανότητα στην αναγνώριση κακοηθών βλαβών, διατηρώντας παράλληλα ακέραιη την απαιτούμενη υπολογιστική αποδοτικότητα. Συμπερασματικά, τεκμηριώνεται ότι τα μοντέλα χαμηλής πολυπλοκότητας αποτελούν μια ρεαλιστική και αξιόπιστη επιλογή για αυτοματοποιημένη διάγνωση σε φορητά υπερηχογραφικά συστήματα (*Point-of-Care Ultrasound, POCUS*).

Λέξεις Κλειδιά: Καρκίνος του Μαστού, Υπερηχογράφημα Μαστού (BUS), Βαθιά Μάθηση (Deep Learning, DL), Συνελικτικά Νευρωνικά Δίκτυα Χαμηλής Πολυπλοκότητας (Lightweight CNNs), Σύστημα BI-RADS, Αποδοτική Τεχνητή Νοημοσύνη (Efficient AI), Περιοχή Ενδιαφέροντος (ROI).

Abstract

Breast cancer is the most common malignancy among women worldwide, making early diagnosis an immediate clinical priority. Breast Ultrasound (BUS) is a widely used diagnostic modality; however, its interpretation depends heavily on the examiner's experience. Computer-Aided Diagnosis (CAD) systems based on Deep Learning (DL) have performed satisfactorily in addressing this problem, but the commonly used architectures are computationally heavy, hindering their deployment on portable devices. This thesis develops and evaluates Lightweight Convolutional Neural Networks (Lightweight CNNs) for the automated classification of ultrasound images into benign and malignant categories, with an emphasis on Efficient Artificial Intelligence (Efficient AI).

To this end, two datasets were consolidated: the public BUSI dataset and the more recent BUS-BRA dataset. The clinical BI-RADS categories were grouped to correspond to a binary classification problem. A central element of the methodology was the preprocessing strategy: leveraging the available segmentation masks, dilation and background masking were applied, effectively isolating the Region of Interest (ROI) from visual noise. Three lightweight models—MobileNetV2, EfficientNetB0, and NASNetMobile—were evaluated using Transfer Learning (TL). To ensure the reliability of the results, a 5-fold cross-validation with patient-level grouping was employed. Additionally, techniques such as Test-Time Augmentation (TTA), dynamic Threshold Optimization, and, on an experimental level, the integration of a Convolutional Block Attention Module (CBAM) were tested.

The results demonstrate that lightweight models, when combined with targeted preprocessing ("Input E"), achieve diagnostic performance comparable to that of much heavier architectures. MobileNetV2 emerged as the most stable model, achieving an overall accuracy of 97.95% (AUC 0.9974) on the BUSI dataset and 89.49% (AUC 0.9387) on the considerably more challenging, multicenter BUS-BRA dataset. The experimental integration of the CBAM mechanism further

increased accuracy on the BUSI dataset to 98.85% and clinically enhanced the network's discriminative capability in identifying malignant lesions, while fully maintaining the requisite computational efficiency. In conclusion, the findings substantiate that low-complexity models constitute a realistic and clinically reliable option for automated diagnosis in portable Point-of-Care Ultrasound (POCUS) systems.

Keywords: Breast Cancer, Breast Ultrasound (BUS), Deep Learning (DL), Lightweight Convolutional Neural Networks (Lightweight CNNs), BI-RADS System, Efficient AI, Region of Interest (ROI).

Ευχαριστίες

Η παρούσα διπλωματική εργασία εκπονήθηκε στο πλαίσιο του προπτυχιακού προγράμματος σπουδών του Τμήματος Ηλεκτρολόγων Μηχανικών και Μηχανικών Υπολογιστών της Πολυτεχνικής Σχολής του Πανεπιστημίου Δυτικής Μακεδονίας. Η ολοκλήρωσή της σηματοδοτεί την κορύφωση μιας απαιτητικής αλλά εξαιρετικά δημιουργικής ακαδημαϊκής πορείας. Στο σημείο αυτό, αισθάνομαι την ανάγκη να εκφράσω τις βαθύτατες ευχαριστίες μου στα πρόσωπα που συντέλεσαν καθοριστικά στην επιτυχή περάτωσή της.

Αρχικά, θα ήθελα να απευθύνω τις θερμότερες ευχαριστίες μου στον επιβλέποντα της εργασίας, Αναπληρωτή Καθηγητή του Πανεπιστημίου Δυτικής Μακεδονίας, κ. Μηνά Δασυγένη, καθώς και στη κ. Θεοδώρα Σανιδά. Τους ευχαριστώ ειλικρινά για την εμπιστοσύνη που μου έδειξαν αναλαμβάνοντας την επίβλεψη και την καθοδήγηση της παρούσας έρευνας. Η επιστημονική τους κατάρτιση, οι πολύτιμες συμβουλές τους καθ' όλη τη διάρκεια της συγγραφής, αλλά και η άψογη, υποστηρικτική συνεργασία μας, αποτέλεσαν τον ακρογωνιαίο λίθο για την ολοκλήρωση αυτού του εγχειρήματος.

Τέλος, το πιο μεγάλο, ουσιαστικό και από καρδιάς ευχαριστώ ανήκει στην οικογένειά μου. Τους ευχαριστώ για την αμέριστη και ανιδιοτελή στήριξή τους, όχι μόνο κατά τη διάρκεια των ακαδημαϊκών μου χρόνων, αλλά σε όλη την πορεία της ζωής μου. Η πίστη τους στις δυνατότητές μου, η υπομονή και οι αθόρυβες θυσίες τους αποτέλεσαν το ισχυρότερο κίνητρο για την επίτευξη των στόχων μου. Η παρούσα εργασία αφιερώνεται σε εκείνους με απέραντη ευγνωμοσύνη.

Περιεχόμενα

ΠΕΡΙΛΗΨΗ	1
ABSTRACT	4
ΕΥΧΑΡΙΣΤΙΕΣ	7
ΠΕΡΙΕΧΟΜΕΝΑ	9
ΚΑΤΑΛΟΓΟΣ ΣΧΗΜΑΤΩΝ	13
ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ	14
ΚΕΦΑΛΑΙΟ 1: ΕΙΣΑΓΩΓΗ	17
1.1 Πλαίσιο και Κίνητρο	17
1.2 Αντικείμενο και Στόχοι της Εργασίας	19
1.3 Σχετικές Εργασίες	20
1.4 Σύνοψη κεφαλαίου	21
ΚΕΦΑΛΑΙΟ 2: ΙΑΤΡΙΚΟ ΥΠΟΒΑΘΡΟ ΚΑΙ ΔΙΑΓΝΩΣΤΙΚΗ ΑΠΕΙΚΟΝΙΣΗ	23
2.1 Καρκίνος του μαστού	23
2.2 Ο Υπέρηχος ως Διαγνωστικό Εργαλείο	26
2.3 Το Σύστημα BI-RADS: Κατηγορίες και Μορφολογικά Χαρακτηριστικά	29
2.4 Σύνοψη κεφαλαίου	32
ΚΕΦΑΛΑΙΟ 3: ΘΕΩΡΙΑ ΒΑΘΙΑΣ ΜΑΘΗΣΗΣ	33
3.1 Συνελκτικά Νευρωνικά Δίκτυα: Βασική Αρχιτεκτονική	33
3.2 Μεταφορά Μάθησης: Η Συμβολή της στα Μικρά Ιατρικά Σύνολα Δεδομένων	35
3.3 Αρχιτεκτονικές Χαμηλής Πολυπλοκότητας	37
3.3.1 MobileNetV2	37
3.3.2 EfficientNet	38

3.3.3 NASNetMobile	38
3.4 Μηχανισμοί Προσοχής: Θεωρία του CBAM	39
3.5 Μετρικές Αξιολόγησης	41
3.5.1 Ορθότητα / Ακρίβεια	41
3.5.2 Ακρίβεια Πρόβλεψης	42
3.5.3 Ανάκληση / Ευαισθησία	42
3.5.4 Το Μέτρο F1	42
3.5.5 Εμβαδόν Κάτω από την Καμπύλη	43
3.6 Σύνοψη κεφαλαίου	43
ΚΕΦΑΛΑΙΟ 4: ΜΕΘΟΔΟΛΟΓΙΑ ΚΑΙ ΑΝΑΠΤΥΞΗ ΣΥΣΤΗΜΑΤΟΣ	44
4.1 Σύνολα Δεδομένων	44
4.1.1 Περιγραφή του Συνόλου Δεδομένων BUSI	44
4.1.2 Περιγραφή του Συνόλου Δεδομένων BUS-BRA	45
4.1.3 Ενοποίηση και Αντιστοίχιση BI-RADS σε Δυναμική Ταξινόμηση	45
4.2 Προ-επεξεργασία Δεδομένων	46
4.2.1 Η Στρατηγική "Input E"	46
4.2.2 Τεχνικές Ενίσχυσης Δεδομένων	49
4.3 Σχεδιασμός και Εκπαίδευση Μοντέλων	51
4.3.1 Παραμετροποίηση Αρχιτεκτονικών	51
4.3.2 Ανάπτυξη του Ενισχυμένου Μοντέλου με Μηχανισμό Προσοχής	53
4.4 Στρατηγική Αξιολόγησης & Βελτιστοποίησης	55
4.4.1 Stratified Group 5-Fold Cross-Validation	55
4.4.2 Ενίσχυση Δεδομένων κατά την Πρόβλεψη	56
4.4.3 Δυναμική Βελτιστοποίηση Κατωφλίου	57
4.5 Σύνοψη κεφαλαίου	59
ΚΕΦΑΛΑΙΟ 5: ΠΕΙΡΑΜΑΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΚΑΙ ΣΥΓΚΡΙΤΙΚΗ ΑΝΑΛΥΣΗ	60
5.1 Περιβάλλον Υλοποίησης: Εργαλεία και Hardware	60
5.2 Αποτελέσματα Baseline Μοντέλων	62
5.2.1 Επιδόσεις MobileNetV2	62
5.2.2 Επιδόσεις EfficientNetB0	65
5.2.3 Επιδόσεις NASNetMobile	69
5.3 Στατιστική Αξιολόγηση Προηγμένων Τεχνικών	72

5.3.1 Η επίδραση του Μηχανισμού Προσοχής CBAM στο BUSI	72
5.3.2 Η επίδραση του Μηχανισμού Προσοχής CBAM στο BUS-BRA	74
5.4 Συγκριτική Ανάλυση Συστήματος	76
5.4.1 Σύγκριση των Baseline Αρχιτεκτονικών	78
5.4.2 Η Επίδραση του Μηχανισμού Προσοχής	78
5.4.3 Τελική Επιλογή και Συμπεράσματα: Το "Νικητήριο" Μοντέλο	79
5.5 Σύνοψη κεφαλαίου	80
ΚΕΦΑΛΑΙΟ 6: ΣΥΜΠΕΡΑΣΜΑΤΑ ΚΑΙ ΜΕΛΛΟΝΤΙΚΕΣ ΕΠΕΚΤΑΣΕΙΣ	82
6.1 Ερμηνεία των Ευρημάτων	82
6.1.1 Ο ρόλος της στρατηγικής "Input E"	82
6.1.2 Η Συμβολή και η Συμπεριφορά του Μηχανισμού Προσοχής	83
6.2 Συμβολή στην Επιστήμη: Η σημασία της χρήσης «ελαφρών» μοντέλων στην κλινική πράξη	84
6.3 Σύγκριση με Υπάρχουσες Προσεγγίσεις στη Διεθνή Βιβλιογραφία	86
6.4 Μελλοντικές Επεκτάσεις	88
6.5 Σύνοψη κεφαλαίου	89
ΒΙΒΛΙΟΓΡΑΦΙΑ	91
ΣΥΝΤΟΜΟΓΡΑΦΙΕΣ - ΑΡΚΤΙΚΟΛΕΞΑ - ΑΚΡΩΝΥΜΙΑ	97
ΑΠΟΔΟΣΗ ΞΕΝΟΓΛΩΣΣΩΝ ΌΡΩΝ	99

Κατάλογος Σχημάτων

ΣΧΗΜΑ 1:ΠΟΣΟΣΤΟ ΕΜΦΑΝΙΣΗΣ ΓΥΝΑΙΚΕΙΟΥ ΚΑΡΚΙΝΟΥ ΤΟΥ ΜΑΣΤΟΥ ΣΤΙΣ ΗΠΑ., ΙΣΤΟΣΕΛΙΔΑ HTTPS://WWW.NCBI.NLM.NIH.GOV/BOOKS/NBK583819/	24
ΣΧΗΜΑ 2: ΠΟΣΟΣΤΟ ΘΝΗΣΙΜΟΤΗΤΑΣ ΠΟΥ ΣΧΕΤΙΖΕΤΑΙ ΜΕ ΤΟΝ ΚΑΡΚΙΝΟ ΤΟΥ ΜΑΣΤΟΥ ΣΤΙΣ ΗΠΑ., ΙΣΤΟΣΕΛΙΔΑ HTTPS://WWW.NCBI.NLM.NIH.GOV/BOOKS/NBK583819/	24
ΣΧΗΜΑ 3: ΚΑΜΠΥΛΕΣ SROC ΠΟΥ ΑΠΕΙΚΟΝΙΖΟΥΝ ΤΗ ΔΙΑΓΝΩΣΤΙΚΗ ΑΚΡΙΒΕΙΑ ΤΟΥ ΥΠΕΡΗΧΟΓΡΑΦΗΜΑΤΟΣ ΜΑΣΤΟΥ [1].....	28
ΣΧΗΜΑ 4: ΓΡΑΦΙΚΗ ΑΝΑΠΑΡΑΣΤΑΣΗ ΤΩΝ ΔΙΑΦΟΡΩΝ ΜΟΡΦΩΝ ΔΕΔΟΜΕΝΩΝ ΕΙΣΟΔΟΥ. ΔΙΑΚΡΙΝΕΤΑΙ Η ΑΚΡΙΒΗΣ ΑΛΛΗΛΟΥΧΙΑ ΒΗΜΑΤΩΝ (ΔΙΑΣΤΟΛΗ, ΕΦΑΡΜΟΓΗ ΜΑΣΚΑΣ, ΠΕΡΙΚΟΠΗ) ΓΙΑ ΤΗ ΔΗΜΙΟΥΡΓΙΑ ΤΗΣ ΕΙΚΟΝΑΣ "INPUT E" [10]	48

Κατάλογος Πινάκων

ΠΙΝΑΚΑΣ 1: ΠΑΡΑΔΕΙΓΜΑΤΑ ΥΠΕΡΗΧΟΓΡΑΦΗΜΑΤΩΝ ΑΠΟ ΤΟ ΣΥΝΟΛΟ ΔΟΚΙΜΗΣ ΚΑΙ ΟΙ ΑΝΤΙΣΤΟΙΧΕΣ ΠΡΟΒΛΕΨΕΙΣ ΤΩΝ ΜΟΝΤΕΛΩΝ ΑΝΑΛΟΓΑ ΜΕ ΤΗ ΜΟΡΦΗ ΕΙΣΟΔΟ [10].....	49
ΠΙΝΑΚΑΣ 2: ΣΥΓΚΕΝΤΡΩΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΟΥ ΜΟΝΤΕΛΟΥ MOBILENETV2 ΣΤΟ ΣΥΝΟΛΟ ΔΕΔΟΜΕΝΩΝ BUS-BRA (1875 ΕΙΚΟΝΕΣ).....	62
ΠΙΝΑΚΑΣ 3: ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ ΤΟΥ ΜΟΝΤΕΛΟΥ MOBILENETV2 ΣΤΟ ΣΥΝΟΛΟ ΔΕΔΟΜΕΝΩΝ BUS-BRA	63
ΠΙΝΑΚΑΣ 4: ΣΥΓΚΕΝΤΡΩΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΟΥ ΜΟΝΤΕΛΟΥ MOBILENETV2 ΣΤΟ ΣΥΝΟΛΟ ΔΕΔΟΜΕΝΩΝ BUSI (780 ΕΙΚΟΝΕΣ).....	64
ΠΙΝΑΚΑΣ 5: ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ ΤΟΥ ΜΟΝΤΕΛΟΥ MOBILENETV2 ΣΤΟ ΣΥΝΟΛΟ ΔΕΔΟΜΕΝΩΝ BUSI	64
ΠΙΝΑΚΑΣ 6: ΣΥΓΚΕΝΤΡΩΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΟΥ ΜΟΝΤΕΛΟΥ EFFICIENTNETB0 ΣΤΟ ΣΥΝΟΛΟ ΔΕΔΟΜΕΝΩΝ BUS- BRA.....	66
ΠΙΝΑΚΑΣ 7: ΣΥΓΚΕΝΤΡΩΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΚΑΙ ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ ΤΟΥ ΜΟΝΤΕΛΟΥ EFFICIENTNETB0 ΣΤΟ ΣΥΝΟΛΟ ΔΕΔΟΜΕΝΩΝ BUS-BRA	66
ΠΙΝΑΚΑΣ 8: ΣΥΓΚΕΝΤΡΩΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΟΥ ΜΟΝΤΕΛΟΥ EFFICIENTNETB0 ΣΤΟ ΣΥΝΟΛΟ ΔΕΔΟΜΕΝΩΝ BUSI.....	67
ΠΙΝΑΚΑΣ 9: ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ ΤΟΥ ΜΟΝΤΕΛΟΥ EFFICIENTNETB0 ΣΤΟ ΣΥΝΟΛΟ ΔΕΔΟΜΕΝΩΝ BUSI	68
ΠΙΝΑΚΑΣ 10: ΣΥΓΚΕΝΤΡΩΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΟΥ ΜΟΝΤΕΛΟΥ NASNETMOBILE ΣΤΟ ΣΥΝΟΛΟ ΔΕΔΟΜΕΝΩΝ BUS-BRA.....	69
ΠΙΝΑΚΑΣ 11: ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ ΤΟΥ ΜΟΝΤΕΛΟΥ NASNETMOBILE ΣΤΟ ΣΥΝΟΛΟ ΔΕΔΟΜΕΝΩΝ BUS-BRA	70
ΠΙΝΑΚΑΣ 12: ΣΥΓΚΕΝΤΡΩΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΟΥ ΜΟΝΤΕΛΟΥ NASNETMOBILE ΣΤΟ ΣΥΝΟΛΟ ΔΕΔΟΜΕΝΩΝ BUSI.....	71

ΠΙΝΑΚΑΣ 13: ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ ΤΟΥ ΜΟΝΤΕΛΟΥ NASNETMOBILE ΣΤΟ ΣΥΝΟΛΟ ΔΕΔΟΜΕΝΩΝ BUSI	71
ΠΙΝΑΚΑΣ 14: ΣΥΓΚΕΝΤΡΩΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΟΥ ΕΝΙΣΧΥΜΕΝΟΥ ΜΟΝΤΕΛΟΥ MOBILENETV2 + CBAM ΣΤΟ ΣΥΝΟΛΟ ΔΕΔΟΜΕΝΩΝ BUSI.....	73
ΠΙΝΑΚΑΣ 15: ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ ΤΟΥ ΕΝΙΣΧΥΜΕΝΟΥ ΜΟΝΤΕΛΟΥ MOBILENETV2 + CBAM ΣΤΟ ΣΥΝΟΛΟ ΔΕΔΟΜΕΝΩΝ BUSI.....	73
ΠΙΝΑΚΑΣ 16: ΣΥΓΚΕΝΤΡΩΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΟΥ ΕΝΙΣΧΥΜΕΝΟΥ ΜΟΝΤΕΛΟΥ MOBILENETV2 + CBAM ΣΤΟ ΣΥΝΟΛΟ ΔΕΔΟΜΕΝΩΝ BUS-BRA.....	75
ΠΙΝΑΚΑΣ 17: ΣΥΓΚΕΝΤΡΩΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΚΑΙ ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ ΤΟΥ ΕΝΙΣΧΥΜΕΝΟΥ ΜΟΝΤΕΛΟΥ MOBILENETV2 + CBAM ΣΤΟ ΣΥΝΟΛΟ ΔΕΔΟΜΕΝΩΝ BUS-BRA	75
ΠΙΝΑΚΑΣ 18: ΣΥΓΚΕΝΤΡΩΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ 5-ΠΛΗΣ ΔΙΑΣΤΑΥΡΟΥΜΕΝΗΣ ΕΠΙΚΥΡΩΣΗΣ ΣΕ ΟΛΑ ΤΑ ΣΥΝΟΛΑ ΔΕΔΟΜΕΝΩΝ.....	77
ΠΙΝΑΚΑΣ 19: ΣΥΓΚΡΙΣΗ ΥΠΟΛΟΓΙΣΤΙΚΟΥ ΚΟΣΤΟΥΣ, ΜΕΓΕΘΟΥΣ ΚΑΙ ΠΑΡΑΜΕΤΡΩΝ ΤΩΝ BASELINE ΑΡΧΙΤΕΚΤΟΝΙΚΩΝ.....	77
ΠΙΝΑΚΑΣ 20:ΥΠΟΛΟΓΙΣΤΙΚΟ ΚΟΣΤΟΣ, ΜΕΓΕΘΟΣ ΚΑΙ ΑΡΙΘΜΟΣ ΠΑΡΑΜΕΤΡΩΝ ΤΗΣ ΕΝΙΣΧΥΜΕΝΗΣ ΑΡΧΙΤΕΚΤΟΝΙΚΗΣ MOBILENETV2 + CBAM.....	79

Κεφάλαιο 1: Εισαγωγή

Το παρόν κεφάλαιο παρουσιάζει το πλαίσιο και τα κίνητρα της εργασίας, με επίκεντρο την ανάγκη ανάπτυξης «ελαφριών» (lightweight) συστημάτων αυτόματης διάγνωσης υπερήχων μαστού για άμεση εφαρμογή σε φορητές συσκευές. Ακολουθεί ο προσδιορισμός του αντικειμένου και των επιμέρους στόχων της έρευνας, καθώς και η βιβλιογραφική ανασκόπηση που αναδεικνύει το υφιστάμενο ερευνητικό κενό.

1.1 Πλαίσιο και Κίνητρο

Ο καρκίνος του μαστού παραμένει η συχνότερα διαγνωσμένη κακοήθεια και η κορυφαία αιτία θανάτου από καρκίνο μεταξύ των γυναικών παγκοσμίως. Δεκαετίες κλινικής έρευνας έχουν αποδείξει ότι η έγκαιρη ανίχνευση αποτελεί τον καθοριστικότερο παράγοντα για τη βελτίωση των ποσοστών επιβίωσης. Στο πλαίσιο αυτό, το Υπερηχογράφημα Μαστού (Breast Ultrasound, BUS) έχει καθιερωθεί ως βασική απεικονιστική μέθοδος: είναι μη επεμβατικό, δεν εκπέμπει ακτινοβολία, έχει χαμηλό κόστος και παρουσιάζει σαφές πλεονέκτημα έναντι της παραδοσιακής μαστογραφίας σε νεότερες γυναίκες και σε ασθενείς με πυκνό μαστικό ιστό, όπου η μαστογραφία εμφανίζει συχνά ψευδώς αρνητικά αποτελέσματα.

Παρά την κλινική του χρησιμότητα, η ερμηνεία των υπερηχογραφικών εικόνων παραμένει έντονα εξαρτημένη από τον χειριστή. Η διαφοροποίηση μεταξύ καλοήθων και κακοήθων βλαβών απαιτεί εκτεταμένη ακτινολογική εμπειρία, καθώς η αξιολόγηση βασίζεται σε ανεπαίσθητα μορφολογικά χαρακτηριστικά — σχήμα βλάβης, σαφήνεια ορίων, ηχογένεια. Η υποκειμενικότητα αυτή έχει άμεσες κλινικές συνέπειες: υψηλή μεταβλητότητα μεταξύ παρατηρητών (inter-observer variability) και αυξημένο ποσοστό ψευδώς θετικών διαγνώσεων, που οδηγεί τις ασθενείς σε αδικαιολόγητη ψυχολογική επιβάρυνση και περιττές επεμβατικές βιοψίες.

Για την αντιμετώπιση αυτών των περιορισμών, η Τεχνητή Νοημοσύνη (Artificial Intelligence, AI) — και πιο συγκεκριμένα η Βαθιά Μάθηση (Deep Learning, DL) μέσω Συνελκτικών Νευρωνικών Δικτύων (Convolutional Neural Networks, CNNs) — έχει μεταμορφώσει την ανάλυση ιατρικών εικόνων τα τελευταία χρόνια. Τα σύγχρονα συστήματα Υποβοηθούμενης Διάγνωσης (Computer-Aided Diagnosis, CAD) επιτυγχάνουν συχνά διαγνωστικές ακρίβειες συγκρίσιμες με εκείνες των ανθρώπινων ειδικών. Υπάρχει, ωστόσο, ένα πρακτικό εμπόδιο που συχνά παραβλέπεται: τα περισσότερα state-of-the-art μοντέλα DL — όπως βαθιά ResNets ή DenseNets — είναι υπολογιστικά βαριά. Απαιτούν εκατομμύρια παραμέτρους και εξειδικευμένο hardware, ιδίως GPUs υψηλής απόδοσης. Αυτό περιορίζει ουσιαστικά την εφαρμογή τους σε καλά χρηματοδοτούμενα, κεντρικά νοσοκομειακά περιβάλλοντα.

Παράλληλα, το Υπερηχογράφημα στο Σημείο Φροντίδας (Point-of-Care Ultrasound, POCUS) κερδίζει ολοένα και μεγαλύτερο έδαφος στην κλινική πρακτική. Πρόκειται για εξαιρετικά φορητές συσκευές χειρός που συνδέονται με smartphone ή tablet, επιτρέποντας άμεση απεικόνιση παρά την κλίνη του ασθενούς — ιδιαίτερα χρήσιμο στην επείγουσα ιατρική, στις αγροτικές κλινικές και σε περιβάλλοντα με περιορισμένους πόρους. Οι συσκευές αυτές, όμως, δεν είναι σε θέση να εκτελέσουν τοπικά βαριά μοντέλα DL. Η επεξεργασία μέσω νέφους (cloud-based AI) αποτελεί θεωρητικά εναλλακτική λύση, αλλά εισάγει ζητήματα καθυστέρησης δικτύου, εξάρτησης από σταθερή σύνδεση στο διαδίκτυο και συμμόρφωσης με τους κανονισμούς προστασίας δεδομένων ασθενών.

Από τη συνδυαστική εξέταση των παραπάνω παραγόντων προκύπτει ένα συγκεκριμένο κενό: δεν υπάρχουν επαρκώς ακριβή, ελαφριά μοντέλα DL ικανά να ταξινομούν υπερηχογραφήματα μαστού σε πραγματικό χρόνο, απευθείας σε συσκευές χαμηλής ισχύος τύπου POCUS, χωρίς υποδομή νέφους. Η κάλυψη αυτού του κενού αποτελεί τον κεντρικό στόχο της παρούσας εργασίας.

Μέσα από τη μελέτη, βελτιστοποίηση και αξιολόγηση μοντέλων υψηλής αποδοτικότητας και χαμηλής υπολογιστικής απαίτησης — όπως τα MobileNet, EfficientNet και NASNet — η έρευνα επιχειρεί να τεκμηριώσει ότι η αποδοτικότητα και η διαγνωστική ακρίβεια δεν βρίσκονται σε αναγκαστική αντίθεση. Ένα ελαφρύ σύστημα AI για τη διάγνωση καρκίνου μαστού, εφόσον είναι αξιόπιστο, μπορεί να διευρύνει σημαντικά την πρόσβαση σε προσυμπτωματικό ακτινολογικό έλεγχο — ιδίως σε κλινικά περιβάλλοντα με περιορισμένους πόρους, όπου εκπαιδευμένοι ειδικοί απουσιάζουν ή υπολείπονται.

1.2 Αντικείμενο και Στόχοι της Εργασίας

Αντικείμενο της εργασίας είναι η ανάπτυξη, υλοποίηση και αξιολόγηση ενός συστήματος CAD με AI, το οποίο έχει σχεδιαστεί για την αυτοματοποιημένη ταξινόμηση υπερηχογραφικών εικόνων μαστού. Σε αντίθεση με παραδοσιακές προσεγγίσεις που βασίζονται σε υπολογιστικά απαιτητικές αρχιτεκτονικές, εδώ το βάρος πέφτει στα CNNs χαμηλής πολυπλοκότητας. Το σύστημα ταξινομεί τις αλλοιώσεις σε δύο κατηγορίες — Καλοήθεις (Benign) έναντι Κακοήθων (Malignant) — αντιστοιχίζοντας τις κλινικές κατηγορίες του Συστήματος Αναφοράς και Δεδομένων Απεικόνισης Μαστού (Breast Imaging Reporting and Data System, BI-RADS) σε αυτά τα δύο αποτελέσματα.

Για την επίτευξη υψηλής διαγνωστικής ακρίβειας με παράλληλη διατήρηση της υπολογιστικής αποδοτικότητας, η μελέτη οργανώνεται γύρω από τέσσερις επιμέρους στόχους:

1. Ο πρώτος αφορά την υλοποίηση μιας εξειδικευμένης στρατηγικής προ-επεξεργασίας δεδομένων, γνωστής στη βιβλιογραφία ως "Input E". Αντί να τροφοδοτηθεί στο δίκτυο ολόκληρη η ακατέργαστη υπερηχογραφική εικόνα — η οποία συχνά περιέχει ακουστικό θόρυβο, σκιές και άσχετες πληροφορίες υποβάθρου — η προσέγγιση αυτή απομονώνει την Περιοχή Ενδιαφέροντος (Region of Interest, ROI). Με βάση τις διαθέσιμες μάσκες τμηματοποίησης, εκτελείται διαστολή (dilation) ώστε να συμπεριληφθεί το περιτυμοριακό περιθώριο, ακολουθεί απόκρυψη του υποβάθρου και τέλος αποκοπή (cropping) της περιοχής. Με τον τρόπο αυτό τα CNN εκπαιδεύονται αποκλειστικά στα μορφολογικά χαρακτηριστικά της αλλοίωσης.
2. Ο δεύτερος στόχος είναι η συγκριτική αξιολόγηση τριών αρχιτεκτονικών CNNs χαμηλής πολυπλοκότητας: MobileNetV2, EfficientNetB0 και NASNetMobile. Μέσω Μεταφοράς Μάθησης (Transfer Learning, TL) από προ-εκπαιδευμένα βάρη ImageNet και επακόλουθου fine-tuning, η έρευνα εξετάζει κατά πόσο αυτά τα ελαφριά μοντέλα μπορούν να φτάσουν τη διαγνωστική ακρίβεια βαθύτερων και πιο απαιτητικών αρχιτεκτονικών, όπως το ResNet ή το VGG. Αν το αποδείξουν, το σύστημα γίνεται άμεσα κατάλληλο για ενσωμάτωση σε κινητές συσκευές ή φορητά συστήματα υπερήχων στο σημείο φροντίδας.
3. Ο τρίτος στόχος αφορά την επιστημονική επικύρωση και βελτιστοποίηση της απόδοσης. Το πειραματικό πλαίσιο χρησιμοποιεί Stratified Group 5-Fold Cross-Validation, που εμποδίζει τη διαρροή δεδομένων (data leakage) διασφαλίζοντας ότι εικόνες από τον ίδιο ασθενή δεν εμφανίζονται ταυτόχρονα στα σύνολα εκπαίδευσης και επικύρωσης. Παράλληλα, η

απόδοση βελτιστοποιείται μέσω Test-Time Augmentation (TTA) και δυναμικής επιλογής του βέλτιστου κατωφλίου ταξινόμησης.

4. Τέλος, ο τέταρτος στόχος είναι η παρουσίαση ολοκληρωμένης συγκριτικής ανάλυσης των αρχιτεκτονικών, βασισμένης σε καθιερωμένες ιατρικές μετρικές: Accuracy, AUC, Precision, Recall και F1-score. Θα καταγραφούν αναλυτικά όλες οι φάσεις ανάπτυξης, τα τεχνικά ζητήματα που προέκυψαν κατά την πορεία, καθώς και συμπεράσματα για την κλινική βιωσιμότητα τέτοιων μοντέλων. Ο τελικός σκοπός παραμένει πρακτικός: να υποβοηθηθούν οι ακτινολόγοι στη μείωση της διαγνωστικής μεταβλητότητας και στον περιορισμό των περιττών βιοψιών.

1.3 Σχετικές Εργασίες

Η αυτοματοποιημένη ανάλυση υπερηχογραφήματων μαστού αποτελεί ένα από τα πιο ενεργά πεδία της βιοϊατρικής έρευνας. Τα σύνολα δεδομένων BUSI και BUS-BRA έχουν καθιερωθεί ως τα κύρια σημεία αναφοράς για την αξιολόγηση νέων μεθοδολογιών. Η σχετική βιβλιογραφία καλύπτει ένα ευρύ φάσμα αρχιτεκτονικών: από παραδοσιακά CNNs όπως τα ResNet, VGG και DenseNet, έως Vision Transformers (ViTs) και υβριδικά μοντέλα [1]. Αρκετοί ερευνητές αξιοποιούν τεχνικές συνδυαστικής μάθησης (ensemble learning), συνδυάζοντας μοντέλα όπως το VGG19 με το ResNet152, επιτυγχάνοντας τιμές AUC έως 0,951 [2]. Παράλληλα, έχουν προταθεί προσεγγίσεις που ενσωματώνουν την αυτόματη κατάτμηση (segmentation) στη διαδικασία ταξινόμησης, χρησιμοποιώντας αρχιτεκτονικές όπως το U-Net++ για εξαγωγή χαρακτηριστικών πριν την τελική διάγνωση [3], [4]. Πιο πρόσφατα, το σύνολο δεδομένων BUS-BRA επέτρεψε την αξιολόγηση μοντέλων βάσει της κλινικής κλίμακας BI-RADS (κατηγορίες 2–5), παρέχοντας λεπτομερέστερη κλινική πληροφορία σε σχέση με την απλή δυαδική ταξινόμηση και ενισχύοντας τη γενικευσιμότητα των μοντέλων [5], [6].

Παρά την έκταση αυτής της έρευνας, οι υπάρχουσες προσεγγίσεις εμφανίζουν ορισμένους επαναλαμβανόμενους περιορισμούς. Ο πρώτος αφορά το υπολογιστικό κόστος. Πολλές μελέτες στηρίζονται σε αρχιτεκτονικές με εκατομμύρια παραμέτρους — το VGG13, για παράδειγμα, φτάνει τα 128 εκατομμύρια — οι οποίες είναι ιδιαίτερα ευάλωτες στο φαινόμενο υπερπροσαρμογής (overfitting) όταν εκπαιδεύονται σε μικρά ιατρικά σύνολα όπως το BUSI [1]. Οι αυξημένες απαιτήσεις τους σε υπολογιστικούς πόρους τα καθιστούν επίσης μη πρακτικά για χρήση σε ενσωματωμένα (embedded) συστήματα ή σε φορητές συσκευές υπερήχων σε συνθήκες σημείου φροντίδας [1], [7].

Ένα δεύτερο, συχνά παραβλεπόμενο, πρόβλημα είναι η έλλειψη ανεξαρτησίας των ασθενών (patient independence). Πολλές μελέτες εφαρμόζουν απλό τυχαίο διαχωρισμό των δεδομένων, με αποτέλεσμα διαφορετικές λήψεις του ίδιου όγκου από τον ίδιο ασθενή να εμφανίζονται ταυτόχρονα στο σύνολο εκπαίδευσης και στο σύνολο δοκιμής. Αυτό οδηγεί σε διαρροή δεδομένων, τεχνητά διογκωμένα ποσοστά ακρίβειας και μειωμένη αξιοπιστία σε πραγματικές κλινικές συνθήκες [1].

Υπάρχουν και άλλες παθογένειες. Η τάση ενσωμάτωσης ολοένα πολυπλοκότερων μηχανισμών προσοχής (attention mechanisms) αυξάνει το υπολογιστικό βάρος χωρίς πάντα να αποδίδει ανάλογη βελτίωση στην ακρίβεια, ειδικά όταν τα δεδομένα εισόδου δεν έχουν προεπεξεργαστεί κατάλληλα [8]. Πολλά από τα συστήματα αυτά λειτουργούν ως «μαύρα κουτιά», χωρίς δυνατότητα ερμηνείας των αποφάσεών τους, γεγονός που περιορίζει την εμπιστοσύνη των κλινικών ιατρών και εμποδίζει την υιοθέτησή τους στην πράξη [4]. Τέλος, αρκετές μελέτες εξαιρούν την κατηγορία "Normal" από το BUSI προκειμένου να αντιμετωπίσουν την ανισορροπία των κλάσεων, περιορίζοντας όμως έτσι την εφαρμοσιμότητα των μοντέλων σε πραγματικά σενάρια προληπτικού ελέγχου (screening) [9].

1.4 Σύνοψη κεφαλαίου

Το παρόν κεφάλαιο διαμόρφωσε το ερευνητικό πλαίσιο και τους στόχους της εργασίας, εστιάζοντας στην ανάπτυξη αποδοτικών, χαμηλής υπολογιστικής πολυπλοκότητας μοντέλων (Efficient AI) για την υποστήριξη της υπερηχογραφικής διάγνωσης στο σημείο φροντίδας. Η βιβλιογραφική ανασκόπηση ανέδειξε σημαντικά κενά, τα οποία η παρούσα διπλωματική έρχεται να καλύψει δίνοντας λύση σε τρία κρίσιμα ζητήματα: την υπερβολική πολυπλοκότητα των υφιστάμενων μοντέλων, τον κίνδυνο διαρροής δεδομένων και την περιορισμένη πρακτική εφαρμοσιμότητα. Συγκεκριμένα, η εργασία αποδεικνύει ότι τα lightweight CNNs, ενισχυμένα από τη στοχευμένη προ-επεξεργασία της στρατηγικής "Input E", μπορούν να επιτύχουν διαγνωστική ακρίβεια επιπέδου state-of-the-art, καθιστώντας τα ιδανικά για εκτέλεση σε πραγματικό χρόνο σε φορητές συσκευές περιορισμένων πόρων. Παράλληλα, υιοθετώντας μια αυστηρή στρατηγική patient-independent 5-Fold CV, διασφαλίζεται η πραγματική ικανότητα γενίκευσης του συστήματος, προσφέροντας κλινικά αξιόπιστα αποτελέσματα απαλλαγμένα από φαινόμενα απομνημόνευσης.

Η σχεδίαση και αξιολόγηση ενός συστήματος CAD, ωστόσο, δεν μπορεί να στηριχθεί μόνο στη ML. Απαιτεί κατανόηση της κλινικής φύσης του προβλήματος. Για τον λόγο αυτό, το Κεφάλαιο 2 παρέχει το απαραίτητο ιατρικό και απεικονιστικό υπόβαθρο: εξετάζεται η παθολογία του καρκίνου του μαστού, αναλύεται η διαγνωστική αξία του υπερηχογραφήματος και παρουσιάζεται το διεθνές σύστημα αναφοράς BI-RADS, το οποίο αποτελεί τον κλινικό άξονα ταξινόμησης των δεδομένων της παρούσας έρευνας.

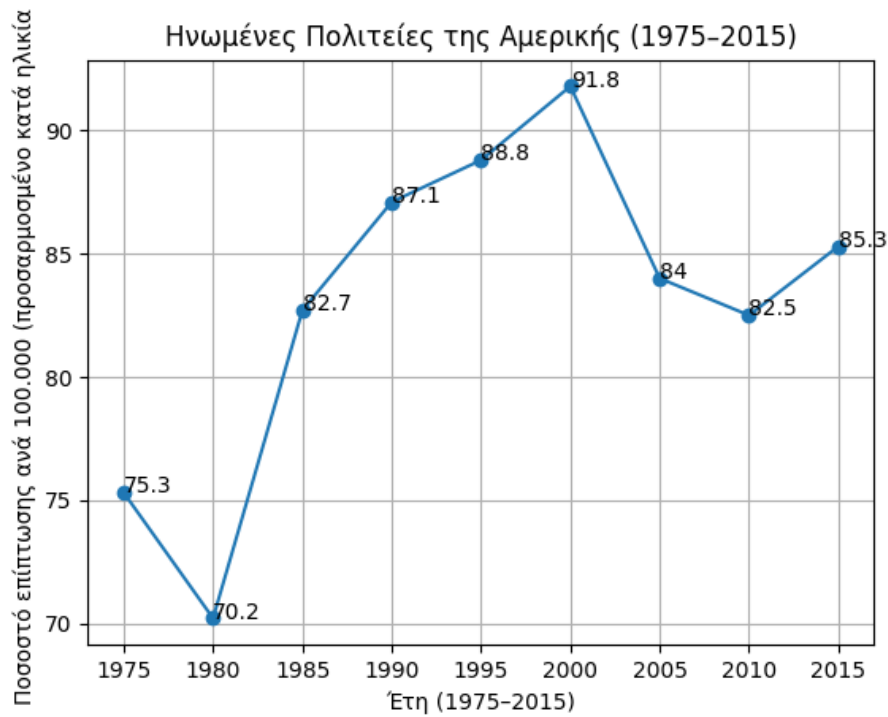
Κεφάλαιο 2: Ιατρικό Υπόβαθρο και Διαγνωστική Απεικόνιση

Το κεφάλαιο αυτό παρέχει το κλινικό και απεικονιστικό υπόβαθρο που χρειάζεται για την κατανόηση της παρούσας έρευνας. Αναλύεται η παθολογία του καρκίνου του μαστού, εξετάζονται τα πλεονεκτήματα του υπερηχογραφήματος έναντι άλλων διαγνωστικών μεθόδων, και παρουσιάζεται το Διεθνές Σύστημα Αναφοράς Δεδομένων Απεικόνισης Μαστού (Breast Imaging Reporting and Data System, BI-RADS) μαζί με τα μορφολογικά χαρακτηριστικά των αλλοιώσεων.

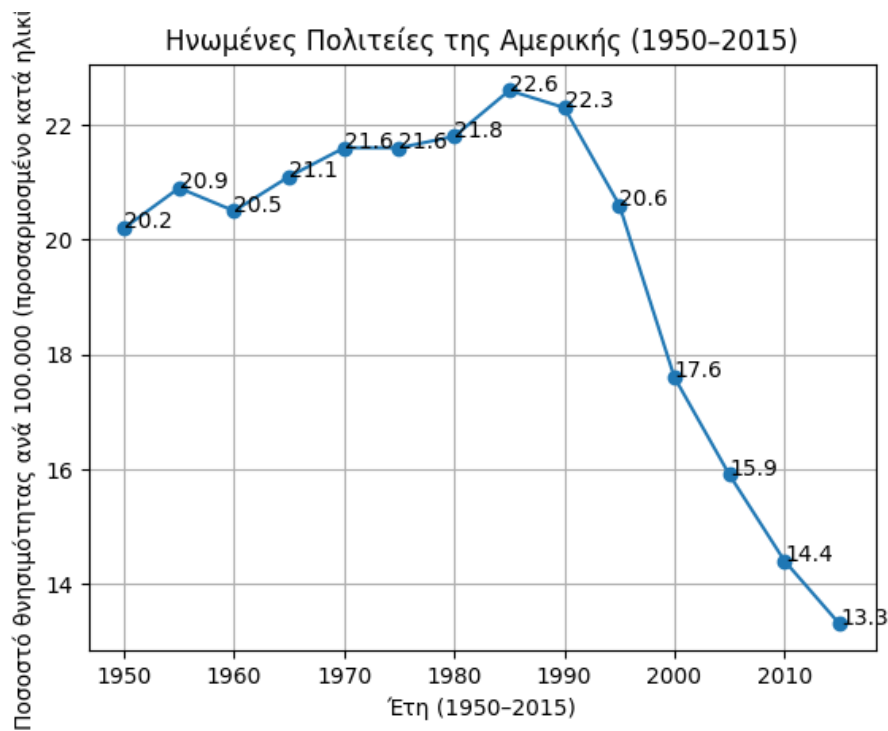
2.1 Καρκίνος του μαστού

Ο καρκίνος του μαστού αποτελεί την πιο συχνά διαγιγνωσόμενη κακοήθεια μεταξύ των γυναικών παγκοσμίως [1], [11]. Τα επιδημιολογικά δεδομένα είναι ανησυχητικά: μόνο το 2022, εκτιμήθηκαν περίπου 2,3 εκατομμύρια νέες διαγνώσεις και 670.000 έως 685.000 θάνατοι σχετιζόμενοι με τη νόσο [4], [11]. Περίπου 1 στις 8 γυναίκες αναμένεται να διαγνωστεί με καρκίνο του μαστού κάποια στιγμή στη ζωή της [5], [6].

Ορισμένα από τα παραπάνω αναφερόμενα πρόσφατα επιδημιολογικά δεδομένα παρουσιάζονται σχηματικά στις Εικόνες 1 και 2 [12].



ΣΧΗΜΑ 1: ΠΟΣΟΣΤΟ ΕΜΦΑΝΙΣΗΣ ΓΥΝΑΙΚΕΙΟΥ ΚΑΡΚΙΝΟΥ ΤΟΥ ΜΑΣΤΟΥ ΣΤΙΣ ΗΠΑ., ΙΣΤΟΣΕΛΙΔΑ [HTTPS://WWW.NCBI.NLM.NIH.GOV/BOOKS/NBK583819/](https://www.ncbi.nlm.nih.gov/books/NBK583819/)



ΣΧΗΜΑ 2: ΠΟΣΟΣΤΟ ΘΝΗΣΙΜΟΤΗΤΑΣ ΠΟΥ ΣΧΕΤΙΖΕΤΑΙ ΜΕ ΤΟΝ ΚΑΡΚΙΝΟ ΤΟΥ ΜΑΣΤΟΥ ΣΤΙΣ ΗΠΑ., ΙΣΤΟΣΕΛΙΔΑ [HTTPS://WWW.NCBI.NLM.NIH.GOV/BOOKS/NBK583819/](https://www.ncbi.nlm.nih.gov/books/NBK583819/)

Η πρόγνωση και τα ποσοστά θνησιμότητας εξαρτώνται άμεσα από το πότε ανιχνεύεται η νόσος. Το στάδιο του όγκου κατά τη διάγνωση παραμένει ο πιο αξιόπιστος προγνωστικός δείκτης, με τα

επιδημιολογικά δεδομένα να το επιβεβαιώνουν με συνέπεια.

Όταν η νόσος εντοπιστεί ενώ παραμένει περιορισμένη στον μαστό, το ποσοστό 5ετούς επιβίωσης αγγίζει το 99% [13], [14]. Η εικόνα αλλάζει σημαντικά αν ο όγκος έχει ήδη προσβάλει τους παρακείμενους λεμφαδένες: το ποσοστό πέφτει στο 86% [13], [14]. Σε περιπτώσεις απομακρυσμένης μετάστασης — όταν δηλαδή η νόσος έχει φτάσει στους πνεύμονες, το ήπαρ ή τα οστά — η 5ετής επιβίωση διαμορφώνεται γύρω στο 30-31% [13], [14].

Παθολογικά Χαρακτηριστικά: Καλοήθεις έναντι Κακοήθων Βλαβών

Τα καρκινικά κύτταρα του μαστού προέρχονται από την εσωτερική επένδυση των γαλακτοφόρων πόρων ή των λοβίων, δίνοντας αντίστοιχα πορογενή και λοβιακά καρκινώματα. Βασικό τους γνώρισμα είναι η ανεξέλεγκτη ανάπτυξη, η διήθηση του παρακείμενου ιστού και η δυνατότητα μετάστασης μέσω αίματος και λέμφου σε άλλα όργανα [15]. Από ιστοπαθολογική άποψη, οι βλάβες του μαστού διακρίνονται με βάση την κυτταρική τους συμπεριφορά, τον τρόπο ανάπτυξης και την ικανότητα διήθησης των ιστών.

Οι καλοήθεις βλάβες αντιστοιχούν σε μη καρκινικούς κυτταρικούς πολλαπλασιασμούς που παραμένουν τοπικά περιορισμένοι και δεν έχουν τη βιολογική ικανότητα να διηθήσουν γειτονικούς ιστούς ή να μεταστατεύσουν [16]. Στην κλινική πράξη, συναντώνται συχνά ινοαδενώματα, κύστες με υγρό και ενδοπορικά θηλώματα — τα ινοαδενώματα αποτελούν την πιο συνηθισμένη συμπαγή μάζα, ενώ η υπερηχογραφία μπορεί να αναγνωρίσει τις κύστες με διαγνωστική ακρίβεια που πλησιάζει το 100% [6]. Μορφολογικά, οι καλοήθεις μάζες εμφανίζουν κατά κανόνα ωοειδές σχήμα, σαφή και καλά οριοθετημένα περιγράμματα, καθώς και παράλληλο προσανατολισμό ως προς την επιφάνεια του δέρματος [10], [17].

Οι κακοήθεις βλάβες παρουσιάζουν διαφορετική εικόνα. Αποτελούνται από καρκινικά κύτταρα που συνήθως προέρχονται από την επιθηλιακή επένδυση των πόρων ή των λοβίων [18], με το Διηθητικό Πορογενές Καρκίνωμα (Invasive Ductal Carcinoma, IDC) να αντιπροσωπεύει περίπου το 85% όλων των διαγνωσμένων κακοηθειών [6]. Τα κύτταρα αυτά έχουν την ικανότητα να διαρρηγνύουν τις βασικές μεμβράνες και να εξαπλώνονται μέσω λεμφικής ή αιματογενούς οδού σε απομακρυσμένες ανατομικές θέσεις [12], [14]. Υπερηχογραφικά, οι κακοήθεις μάζες εμφανίζουν ακανόνιστο ή ασύμμετρο σχήμα, ακτινωτά ή γωνιώδη όρια και ασαφή ή μικρολοβωτά περιγράμματα [1], [17]. Συχνά παρατηρείται και μη παράλληλος προσανατολισμός — δηλαδή το ύψος υπερτερεί του πλάτους — σε συνδυασμό με οπίσθια ακουστική σκίαση [10], [17].

Διαγνωστική Απεικόνιση και η Ανάδυση της Τεχνητής Νοημοσύνης

Το Υπερηχογράφημα Μαστού (Breast Ultrasound, BUS) αποτελεί μη επεμβατική μέθοδο χωρίς ακτινοβολία, με καθιερωμένη θέση στη διαγνωστική απεικόνιση. Η αξία του είναι ιδιαίτερα εμφανής σε γυναίκες με πυκνό μαστικό ιστό, όπου η μαστογραφία παρουσιάζει μειωμένη ευαισθησία: ο πυκνός αδενικός ιστός αποκρύπτει πιθανούς όγκους, καθιστώντας την υπερηχογραφία το πιο αξιόπιστο διαγνωστικό εργαλείο [9], [11].

Ωστόσο, η χειροκίνητη ερμηνεία υπερηχογραφικών εικόνων παραμένει εγγενώς υποκειμενική. Εξαρτάται σε μεγάλο βαθμό από την εμπειρία του εξεταστή και εισάγει περιθώριο σφάλματος που εκτιμάται έως 20% [5], [6]. Αυτό έχει ως πρακτική συνέπεια υψηλά ποσοστά ψευδώς θετικών αποτελεσμάτων και, κατ' επέκταση, περιττές βιοψίες σε καλοήθεις βλάβες — αποτέλεσμα της κλινικής επιταγής να μην παραληφθεί καμία κακοήθεια.

Απαντώντας σε αυτές τις αδυναμίες, η έρευνα εστιάζει ολοένα περισσότερο στην περι-ογκική περιοχή (peritumoral region), δηλαδή τον ιστό που περιβάλλει άμεσα τη βλάβη. Η περιοχή αυτή περιέχει διαγνωστικά χρήσιμες πληροφορίες που διαφεύγουν συχνά από την ανθρώπινη αντίληψη [10]. Στο πλαίσιο αυτό αναπτύσσονται συστήματα Υποβοηθούμενης Διάγνωσης (Computer-Aided Diagnosis, CAD) βασισμένα σε Βαθιά Μάθηση (Deep Learning, DL), τα οποία επιχειρούν να εξάγουν ποσοτικά χαρακτηριστικά από τις υπερηχογραφικές εικόνες. Στόχος δεν είναι η αντικατάσταση του κλινικού, αλλά η παροχή μιας αντικειμενικής δεύτερης εκτίμησης που μειώνει τη μεταβλητότητα μεταξύ παρατηρητών και περιορίζει τις περιττές επεμβατικές διαδικασίες [2], [19].

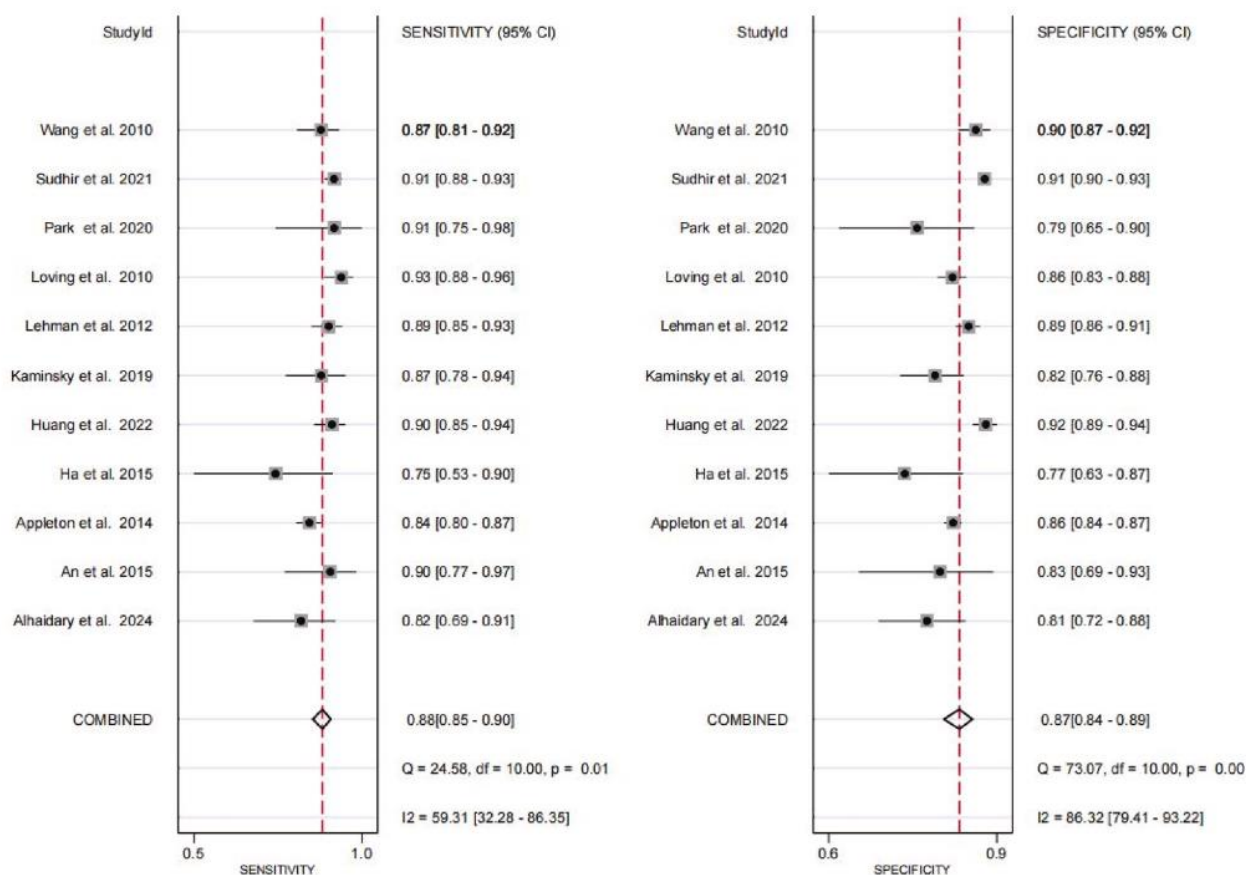
2.2 Ο Υπέρηχος ως Διαγνωστικό Εργαλείο

Το BUS έχει καθιερωθεί ως μια από τις βασικές μεθόδους απεικόνισης στη διαγνωστική του μαστού, με αρκετά πλεονεκτήματα έναντι τεχνικών όπως η μαστογραφία και η μαγνητική τομογραφία (Magnetic Resonance Imaging, MRI). Πρόκειται για μη επεμβατική τεχνική, χωρίς ακτινοβολία, που αξιοποιεί ηχητικά κύματα υψηλής συχνότητας για τη δημιουργία λεπτομερών εικόνων των εσωτερικών ιστών [11], [20]. Η μαστογραφία, αντίθετα, βασίζεται σε ιοντίζουσα ακτινοβολία χαμηλής δόσης — γεγονός που την αποκλείει από τη χρήση σε ορισμένες πληθυσμιακές ομάδες. Το ασφαλές προφίλ του υπερήχου τον καθιστά την πρώτη επιλογή για έγκυες και θηλάζουσες γυναίκες, καθώς και για ασθενείς υψηλού κινδύνου που χρειάζονται συχνές αξιολογήσεις [1], [21]. Από την πλευρά της εμπειρίας της ασθενούς, η εξέταση είναι ανώδυνη και δεν απαιτεί τη συμπίεση του μαστού που συχνά προκαλεί δυσφορία στη μαστογραφία [16], [22].

Η διαγνωστική αξία του υπερήχου γίνεται ιδιαίτερα εμφανής στην περίπτωση πυκνού μαστικού ιστού. Ο πυκνός ινοαδενικός ιστός εμφανίζεται ακτινοσκοπικός στις μαστογραφίες, μιμούμενος οπτικά τους καρκινικούς όγκους και οδηγώντας σε ψευδώς αρνητικά αποτελέσματα — ένας περιορισμός γνωστός ως "συγκάλυψη ιστού" (tissue masking) [6], [23]. Ο υπέρηχος παρακάμπτει αυτό το πρόβλημα αποτελεσματικά. Παρουσιάζει υψηλότερη ευαισθησία σε πυκνούς μαστούς, χαρακτηριστικό που απαντά συχνά σε νεότερες γυναίκες και σε συγκεκριμένες δημογραφικές ομάδες, όπως οι ασιατικοί πληθυσμοί, και ανιχνεύει βλάβες που παραμένουν αόρατες στη μαστογραφία (mammothographically occult) [1], [24]. Έρευνες δείχνουν ότι ο υπέρηχος εντοπίζει υψηλότερο ποσοστό διηθητικών καρκίνων σε πυκνούς μαστούς σε σύγκριση με αποκλειστική χρήση μαστογραφίας [24], γι' αυτό και προτείνεται ως η αρχική απεικονιστική μέθοδος για γυναίκες κάτω των 45 ετών [11].

Πέρα από την αποτελεσματικότητά του στον πυκνό παρέγχυμα, ο υπέρηχος διακρίνεται και σε άλλες κλινικές εφαρμογές. Παραμένει το πρωταρχικό εργαλείο για τη διαφοροποίηση μεταξύ απλών κύστεων με υγρό και συμπαγών νεοπλασματικών μαζών, ένα διαγνωστικό πρόβλημα όπου άλλες μέθοδοι συχνά αφήνουν περιθώρια αμφιβολίας [5], [25]. Η δυνατότητα απεικόνισης σε πραγματικό χρόνο είναι επίσης ιδιαίτερα χρήσιμη: επιτρέπει την αξιολόγηση της αγγείωσης και των προτύπων αιματικής ροής εντός μιας ύποπτης βλάβης, ενώ παράλληλα καθοδηγεί με ακρίβεια επεμβατικές διαδικασίες όπως οι βιοψίες με κόπτουσα βελόνα (core needle biopsies) [4], [26].

Σύμφωνα με τη μελέτη των Jiang και Mao (2025), τα αποτελέσματα σχετικά με τη διαγνωστική ακρίβεια του υπερηχογραφήματος μαστού αποτυπώνονται στο παρακάτω σχήμα.



ΣΧΗΜΑ 3: ΚΑΜΠΥΛΕΣ SROC ΠΟΥ ΑΠΕΙΚΟΝΙΖΟΥΝ ΤΗ ΔΙΑΓΝΩΣΤΙΚΗ ΑΚΡΙΒΕΙΑ ΤΟΥ ΥΠΕΡΗΧΟΓΡΑΦΗΜΑΤΟΣ ΜΑΣΤΟΥ[1]

Σε επίπεδο υγειονομικής πρόσβασης και κόστους, ο υπέρηχος παρουσιάζει επίσης ένα σαφές πλεονέκτημα. Η μαγνητική τομογραφία προσφέρει εξαιρετική διαγνωστική ευαισθησία, αλλά το υψηλό της κόστος, η μεγάλη διάρκεια των σαρώσεων και τα αυξημένα ποσοστά ψευδώς θετικών αποτελεσμάτων περιορίζουν σημαντικά την εφαρμογή της στην κλινική ρουτίνα — οδηγώντας συχνά σε αδικαιολόγητες βιοψίες και αγωνία για την ασθενή [20], [27]. Ο υπέρηχος είναι οικονομικά προσιτός και λειτουργεί με χαμηλές δαπάνες [8], [26]. Η φορητή φύση των συσκευών τον καθιστά κατάλληλο για διαγνωστικές εξετάσεις στο σημείο φροντίδας, και χρησιμεύει ως βασικό διαγνωστικό μέσο σε περιοχές με περιορισμένους πόρους, όπου υποδομές μαστογραφίας ή μαγνητικής τομογραφίας είναι ανύπαρκτες ή δυσκολονδιατήρητες [1].

2.3 Το Σύστημα BI-RADS: Κατηγορίες και Μορφολογικά Χαρακτηριστικά

Το σύστημα BI-RADS, που αναπτύχθηκε και συντηρείται από το Αμερικανικό Κολέγιο Ακτινολογίας (American College of Radiology, ACR), αποτελεί το καθολικά αποδεκτό πρότυπο αναφοράς στην απεικόνιση του μαστού. Προσφέρει τυποποιημένο λεξικό και δομημένο πλαίσιο αναφοράς, με στόχο να περιορίσει την ασάφεια στις ακτινολογικές ερμηνείες, να διευκολύνει την επικοινωνία μεταξύ των εμπλεκόμενων κλινικών και να θέσει σαφείς κατευθύνσεις για τη μετέπειτα κλινική διαχείριση [28].

Κατηγορίες Αξιολόγησης BI-RADS και Διαστρωμάτωση Κινδύνου

Κεντρικό στοιχείο του BI-RADS είναι η κατάταξη των βλαβών σε κατηγορίες ανάλογα με την εκτιμώμενη πιθανότητα κακοήθειας. Κάθε κατηγορία αντιστοιχεί σε συγκεκριμένη κλινική σύσταση, από τον τακτικό προληπτικό έλεγχο έως την άμεση χειρουργική παρέμβαση [29]:

- **Κατηγορία 0 (Ελλιπής - Incomplete):** Υποδεικνύει ότι η αρχική απεικόνιση δεν είναι καταληκτική. Απαιτείται αυστηρά πρόσθετος διαγνωστικός έλεγχος — όπως στοχευμένες μαστογραφικές λήψεις, εξειδικευμένες υπερηχογραφικές εξετάσεις ή σύγκριση με προηγούμενες μελέτες — για την εξαγωγή μιας οριστικής τελικής αξιολόγησης [28], [29].
- **Κατηγορία 1 (Αρνητική - Negative):** Υποδηλώνει μια εντελώς φυσιολογική εξέταση χωρίς παρατηρήσιμες εστιακές βλάβες, αρχιτεκτονική διαταραχή ή ύποπτα ευρήματα. Συνιστάται ο συνήθης προσυμπτωματικός έλεγχος (screening) ανάλογα με την ηλικία [17].
- **Κατηγορία 2 (Καλοήθης - Benign):** Περιλαμβάνει οριστικά καλοήθη ευρήματα που φέρουν 0% πιθανότητα κακοήθειας. Συνηθισμένα παραδείγματα αποτελούν οι απλές κύστεις με υγρό, τα τυπικά ινοαδενώματα και οι ενδομαστικοί λεμφαδένες. Ο προσυμπτωματικός έλεγχος ρουτίνας συνεχίζεται χωρίς να απαιτείται περαιτέρω παρέμβαση [30].
- **Κατηγορία 3 (Πιθανώς Καλοήθης - Probably Benign):** Αποδίδεται σε βλάβες που εμφανίζουν εξαιρετικά καλοήθη μορφολογία, αλλά φέρουν μια ελάχιστη (ονομαστική) πιθανότητα κακοήθειας (μεγαλύτερη από 0% αλλά μικρότερη ή ίση με 2%). Συνιστάται μια συντηρητική προσέγγιση που περιλαμβάνει παρακολούθηση σε σύντομο χρονικό διάστημα (συνήθως 6 μήνες) για την επιβεβαίωση της χρονικής σταθερότητας [31].

- **Κατηγορία 4 (Υποπτη Ανωμαλία - Suspicious Abnormality):** Αυτή η κατηγορία καλύπτει ένα ευρύ φάσμα βλαβών που εμφανίζουν επαρκή ύποπτα χαρακτηριστικά ώστε να δικαιολογείται ιστολογική διάγνωση ιστού (βιοψία). Επειδή η πιθανότητα κακοήθειας κυμαίνεται ευρέως από 2% έως 95%, διαστρωματώνεται περαιτέρω σε τρεις υποκατηγορίες για την εξειδίκευση των κλινικών προσδοκιών: 4A (χαμηλή υποψία: $> 2\%$ έως $\leq 10\%$), 4B (μέτρια υποψία: $> 10\%$ έως $\leq 50\%$) και 4C (υψηλή υποψία: $> 50\%$ έως $< 95\%$) [28], [31].
- **Κατηγορία 5 (Εντόνως Ενδεικτική Κακοήθειας - Highly Suggestive of Malignancy):** Προορίζεται για βλάβες που εμφανίζουν τα κλασικά υπερηχογραφικά γνωρίσματα του καρκίνου του μαστού, αντιστοιχώντας σε πιθανότητα κακοήθειας $\geq 95\%$. Ενδείκνυται αυστηρά η άμεση βιοψία ιστού ή η χειρουργική παρέμβαση [29].
- **Κατηγορία 6 (Γνωστή, Αποδεδειγμένη με Βιοψία Κακοήθεια - Known Biopsy-Proven Malignancy):** Εφαρμόζεται σε βλάβες που έχουν ήδη επιβεβαιωθεί ιστολογικά ως κακοήθεις πριν από την τρέχουσα απεικονιστική εξέταση. Αυτή η κατηγορία χρησιμοποιείται συνήθως στο πλαίσιο της σταδιοποίησης, της παρακολούθησης της νεοεπικουρικής (neoadjuvant) θεραπείας ή του προεγχειρητικού σχεδιασμού [31].

Μορφολογικά Χαρακτηριστικά στον Υπέρηχο: Σχήμα και Όρια

Κατά την αξιολόγηση εστιακών μαζών μέσω υπερήχων, οι ακτινολόγοι βασίζονται σε μορφολογικά περιγραφικά στοιχεία (descriptors) που ορίζονται από το λεξικό BI-RADS για να κατατάξουν μια βλάβη στην κατάλληλη διαγνωστική κατηγορία. Από αυτά, το σχήμα και τα όρια (margin) θεωρούνται οι πιο αξιόπιστοι διαχωριστικοί παράγοντες μεταξύ καλοήθων και κακοήθων παθολογιών [32].

A. Σχήμα Βλάβης (Lesion Shape)

Το περιγραφικό στοιχείο του σχήματος αναφέρεται στη συνολική γεωμετρική διαμόρφωση της μάζας σε πολλαπλά υπερηχογραφικά επίπεδα [10]:

- **Ωοειδές (Oval):** Χαρακτηρίζεται από ένα ελλειπτικό περίγραμμα ή περίγραμμα που μοιάζει με αυγό, το οποίο μπορεί να περιλαμβάνει έως και δύο ήπιες μακρολοβώσεις (macrolobulations). Αυτό το σχήμα συσχετίζεται έντονα με καλοήθεις αιτιολογίες [17].

- **Στρογγυλό (Round):** Ορίζεται ως μια σφαιρική ή καθαρά κυκλική διαμόρφωση. Αν και είναι σχετικά σαφές, απαιτεί προσεκτική συσχέτιση με άλλα χαρακτηριστικά για τον αποκλεισμό της κακοήθειας [31].

Ακανόνιστο (Irregular): Υποδηλώνει ένα σχήμα που δεν είναι ούτε τέλεια ωοειδές ούτε στρογγυλό. Η ακανόνιστη μορφή υπονοεί ανομοιόμορφα, ασύμμετρα πρότυπα ανάπτυξης και αποτελεί κλινικό γνώρισμα κατατεθέν του κακοήθους κυτταρικού πολλαπλασιασμού. *(Σημείωση: Αν και οι παλαιότερες κλινικές πρακτικές χρησιμοποιούσαν συχνά τον όρο "λοβωτό" [lobulated] ως ένα διακριτό πρωτεύον περιγραφικό στοιχείο σχήματος, οι σύγχρονες εκδόσεις του BI-RADS το έχουν ενσωματώσει σε μεγάλο βαθμό στις προαναφερθείσες κατηγορίες, αν και παραμένει μια σχετική περιγραφική έννοια μεταξύ των ειδικών)* [17].

Β. Όρια Βλάβης

Τα όρια αξιολογούν το ακριβές άκρο ή τη διεπαφή του ορίου μεταξύ της εστιακής μάζας και του παρακείμενου φυσιολογικού μαστικού ιστού. Αποτελεί ύψιστης σημασίας χαρακτηριστικό για τον καθορισμό του διηθητικού δυναμικού μιας βλάβης [32]:

- **Περιγεγραμμένα (Circumscribed):** Περιγράφει ένα καλά καθορισμένο, σαφές και απότομα διαγεγραμμένο όριο σε όλη του την περιφέρεια. Υποδηλώνει μια μη διηθητική, εγκλεισμένη (encapsulated) ανάπτυξη, εξαιρετικά τυπική για τις καλοήθεις μάζες.
- **Μη περιγεγραμμένα (Non-circumscribed):** Υποδεικνύει ένα ασαφές όριο, υποδηλώνοντας πιθανή διήθηση του ιστού. Η παρουσία οποιουδήποτε από τα ακόλουθα μη περιγεγραμμένα υπο-χαρακτηριστικά αυξάνει σημαντικά την υποψία κακοήθειας:
 - **Ασαφής (Indistinct):** Στερούνται σαφούς οριοθέτησης, όπου ο όγκος φαίνεται να αναμειγνύεται απρόσκοπτα με το περιβάλλον παρέγχυμα του μαστού.
 - **Γωνιώδη (Angular):** Όρια που παρουσιάζουν αιχμηρές γωνίες ή οξείες γωνιώσεις, ενδεικτικά της ακανόνιστης αντίστασης των ιστών κατά την επέκταση του όγκου.
 - **Μικρολοβωτά (Microlobulated):** Εμφανίζουν πολλαπλούς μικρούς κυματισμούς (κτενοειδής εμφάνιση) κατά μήκος της επιφάνειας, που συχνά αντιστοιχούν στο προωθημένο άκρο ενός όγκου σε παρακείμενες ανατομικές δομές.

- ο **Ακτινωτά (Spiculated)**: Χαρακτηρίζονται από αιχμηρές, αστεροειδείς γραμμές ή ηχογενείς ταινίες που προεξέχουν προς τα έξω ακτινωτά από την κύρια μάζα. Πρόκειται αναμφισβήτητα για το πιο εξαιρετικά ύποπτο υπερηχογραφικό χαρακτηριστικό, έντονα προγνωστικό για διηθητικό καρκίνωμα.

Εν τέλει, στη συνήθη κλινική πράξη, ο συνδυασμός ενός ακανόνιστου σχήματος και μη περιγεγραμμένων ορίων — ιδιαίτερα εκείνων που παρουσιάζουν ακτινωτά ή γωνιώδη όρια — λειτουργεί ως ο πρωταρχικός δείκτης πιθανής κακοήθειας. Η ακριβής εξαγωγή και η αλγοριθμική αναγνώριση αυτών των ανεπαίσθητων μορφολογικών αποχρώσεων αποτελούν τη θεμελιώδη προϋπόθεση για την ανάπτυξη ισχυρών μοντέλων DL και συστημάτων CAD.

2.4 Σύνοψη κεφαλαίου

Το παρόν κεφάλαιο έθεσε το κλινικό και διαγνωστικό πλαίσιο της έρευνας. Το BUS αναδείχθηκε ως εργαλείο πρώτης γραμμής, και η ανάλυση του συστήματος BI-RADS διευκρίνισε τα μορφολογικά πρότυπα βάσει των οποίων οι ιατροί διαχωρίζουν τις καλοήθειες από τις κακοήθειες αλλοιώσεις. Ωστόσο, η υποκειμενικότητα που είναι εγγενής στην ανθρώπινη ερμηνεία ιατρικών εικόνων καθιστά αναγκαία την ανάπτυξη αυτοματοποιημένων υπολογιστικών συστημάτων υποβοήθησης — και εδώ εισέρχεται η AI.

Το Κεφάλαιο 3 που ακολουθεί πραγματεύεται τη θεωρία της DL. Πιο συγκεκριμένα, εξετάζει τις αρχές των CNNs, τον ρόλο της Μεταφοράς Μάθησης (Transfer Learning, TL) στην αντιμετώπιση της έλλειψης μεγάλων ιατρικών συνόλων δεδομένων, τις αρχιτεκτονικές χαμηλής πολυπλοκότητας, καθώς και τους μηχανισμούς προσοχής (CBAM). Το σύνολο αυτών των εννοιών συγκροτεί το αλγοριθμικό υπόβαθρο πάνω στο οποίο δομήθηκε η παρούσα εργασία.

Κεφάλαιο 3: Θεωρία Βαθιάς Μάθησης

Το κεφάλαιο αυτό παρουσιάζει το θεωρητικό πλαίσιο της Βαθιάς Μάθησης (Deep Learning, DL), ξεκινώντας από τις θεμελιώδεις αρχές των Συνελικτικών Νευρωνικών Δικτύων (Convolutional Neural Networks, CNNs) και τη Μεταφορά Μάθησης (Transfer Learning — TL). Ακολουθεί η παρουσίαση των αρχιτεκτονικών χαμηλής πολυπλοκότητας (Lightweight Models) και των μηχανισμών προσοχής (Attention Mechanisms), και το κεφάλαιο κλείνει με τις στατιστικές μετρικές που χρησιμοποιήθηκαν για την αντικειμενική αξιολόγηση των μοντέλων.

3.1 Συνελικτικά Νευρωνικά Δίκτυα: Βασική Αρχιτεκτονική

Τα CNNs αποτελούν εξειδικευμένη κατηγορία αλγορίθμων DL, σχεδιασμένη κατεξοχήν για την ανάλυση δεδομένων με δομή πλέγματος (grid-like topology), όπως οι ψηφιακές εικόνες [33], [34], [35]. Η αρχιτεκτονική τους εμπνέεται από τη φυσιολογία του βιολογικού οπτικού φλοιού των θηλαστικών, στον οποίο τα νευρικά κύτταρα ανταποκρίνονται σε ερεθίσματα εντός περιορισμένων τοπικών πεδίων (receptive fields) [36], [37], [38]. Σε αντίθεση με τις παραδοσιακές μεθόδους μηχανικής μάθησης — που απαιτούν χειροκίνητη εξαγωγή χαρακτηριστικών — ένα CNN οργανώνεται ως ιεραρχική στοίβα επιπέδων που εξάγει αυτόματα και σταδιακά χαρακτηριστικά (feature extraction) απευθείας από τα πρωτογενή δεδομένα [39], [40]. Τα επιμέρους επίπεδα αυτής της αρχιτεκτονικής αναλύονται στη συνέχεια.

Επίπεδο Συνέλιξης

Το επίπεδο συνέλιξης (Convolutional Layer) είναι ο υπολογιστικός πυρήνας ενός CNN και ο κύριος μηχανισμός εξαγωγής χαρακτηριστικών [41], [42], [43]. Η λειτουργία του στηρίζεται σε μια σειρά εκπαιδευσίμων φίλτρων (kernels), τα οποία διατρέχουν διαδοχικά την εικόνα εισόδου. Σε κάθε θέση υπολογίζεται το εσωτερικό γινόμενο μεταξύ των βαρών του φίλτρου και της αντίστοιχης τοπικής περιοχής, παράγοντας δισδιάστατους χάρτες χαρακτηριστικών (feature maps) [35], [39].

Ένα από τα βασικά πλεονεκτήματα του επιπέδου αυτού είναι η κοινή χρήση βαρών (weight sharing): όλοι οι νευρώνες ενός χάρτη χαρακτηριστικών μοιράζονται τα ίδια βάρη, πράγμα που μειώνει σημαντικά τον αριθμό των εκπαιδευσίμων παραμέτρων. Από αυτή την ιδιότητα απορρέει και η αναλλοιότητα στη μετατόπιση (translation invariance): το δίκτυο ανιχνεύει το ίδιο οπτικό μοτίβο ανεξαρτήτως της θέσης του μέσα στην εικόνα [44], [45]. Η λειτουργία της συνέλιξης καθορίζεται από τρεις κύριες υπερπαραμέτρους: το μέγεθος του φίλτρου, το βήμα (stride) — δηλαδή την απόσταση μετακίνησης ανά υπολογισμό — και τη συμπλήρωση (padding), που προσθέτει μηδενικά στοιχεία στα περιθώρια ώστε να ελέγχονται οι χωρικές διαστάσεις της εξόδου [42], [46].

Επίπεδο Ομαδοποίησης

Συνήθως, το επίπεδο συνέλιξης ακολουθείται από ένα επίπεδο ομαδοποίησης (Pooling Layer), το οποίο εκτελεί υποδειγματοληψία (downsampling) με στόχο τη μείωση των χωρικών διαστάσεων των χαρτών χαρακτηριστικών [39], [47]. Αυτό περιορίζει το υπολογιστικό κόστος και μειώνει τον κίνδυνο υπερπροσαρμογής (overfitting) [48], [49]. Οι πιο διαδεδομένες παραλλαγές είναι το Max Pooling — που επιλέγει τη μέγιστη τιμή εντός μιας τοπικής περιοχής — και το Average Pooling, που υπολογίζει τον μέσο όρο. Και οι δύο βοηθούν το δίκτυο να παραμένει ανθεκτικό σε μικρές χωρικές μετατοπίσεις ή παραμορφώσεις της αρχικής εικόνας [33], [50].

Συναρτήσεις Ενεργοποίησης

Για να μπορεί το CNN να επιλύει σύνθετα, μη γραμμικά προβλήματα, μετά από κάθε πράξη συνέλιξης εφαρμόζεται μια συνάρτηση ενεργοποίησης (Activation Functions). Η συνάρτηση αυτή εισάγει τη μη γραμμικότητα στο μοντέλο, επιτρέποντάς του να αναγνωρίζει βαθύτερα και πιο σύνθετα πρότυπα [39], [42], [51]. Η πιο διαδεδομένη επιλογή στα σύγχρονα CNNs είναι η ReLU (Rectified Linear Unit), που ορίζεται ως $f(x) = \max(0, x)$. Η επικράτησή της δεν είναι τυχαία: επιταχύνει σημαντικά την εκπαίδευση και αντιμετωπίζει σε μεγάλο βαθμό το πρόβλημα της εξαφάνισης της κλίσης (vanishing gradient problem), το οποίο περιόριζε την απόδοση παλαιότερων βαθιών δικτύων [33], [51]. Συναρτήσεις όπως η Sigmoid και η Tanh εξακολουθούν να χρησιμοποιούνται, αν και λιγότερο συχνά στα εσωτερικά επίπεδα. Παραλλαγές όπως η Leaky ReLU σχεδιάστηκαν για να αντιμετωπίσουν το φαινόμενο των «νεκρών» νευρώνων (dying ReLU) [35], [41].

Τεχνικές Κανονικοποίησης: Επίπεδο Απόρριψης

Κατά την εκπαίδευση βαθιών δικτύων, η υπερπροσαρμογή αποτελεί έναν από τους πιο επίμονους κινδύνους. Το επίπεδο απόρριψης (Dropout) αντιμετωπίζει ακριβώς αυτό το πρόβλημα, λειτουργώντας ως τεχνική κανονικοποίησης (regularization) [52], [53]. Σε κάθε επανάληψη της εκπαίδευσης, ένα τυχαία επιλεγμένο ποσοστό νευρώνων — μαζί με τις συνδέσεις τους — τίθεται προσωρινά στο μηδέν [36], [52], [54]. Με αυτόν τον τρόπο, το δίκτυο αναγκάζεται να μην εξαρτάται υπερβολικά από μεμονωμένους νευρώνες, γεγονός που αποτρέπει τη συν-προσαρμογή (co-adaptation) και βελτιώνει την ικανότητά του να γενικεύει σε νέα δεδομένα [53], [55]. Παραδοσιακά, το επίπεδο απόρριψης εφαρμόζεται στα επίπεδα πλήρους σύνδεσης, ωστόσο παραλλαγές όπως το Spatial Dropout έχουν προσαρμοστεί και για τα συνελικτικά επίπεδα [52], [56].

Επίπεδα Πλήρους Σύνδεσης

Η αρχιτεκτονική ενός CNN ολοκληρώνεται συνήθως με ένα ή περισσότερα επίπεδα πλήρους σύνδεσης (Fully Connected Layers). Αυτά αναλαμβάνουν τον ρόλο του τελικού ταξινομητή, αξιοποιώντας τα υψηλού επιπέδου χαρακτηριστικά που εξήγαγαν τα προηγούμενα συνελικτικά επίπεδα, για να αποδώσουν τις πιθανότητες υπαγωγής της εικόνας στις κλάσεις εξόδου [34], [40].

3.2 Μεταφορά Μάθησης: Η Συμβολή της στα Μικρά Ιατρικά Σύνολα Δεδομένων

Η TL είναι μια από τις πιο αποδοτικές στρατηγικές στη DL. Η χρησιμότητά της γίνεται ιδιαίτερα εμφανής στην ανάλυση ιατρικών δεδομένων, έναν τομέα όπου τα μεγάλα, επισημειωμένα (annotated) σύνολα δεδομένων είναι συχνά δυσεύρετα [39], [46].

Η λογική πίσω από τη TL είναι σχετικά απλή. Ένα μοντέλο που έχει ήδη εκπαιδευτεί σε ένα εκτεταμένο σύνολο δεδομένων αναφοράς — όπως το ImageNet, με εκατομμύρια εικόνες γενικού περιεχομένου — χρησιμοποιείται ως αφετηρία για μια νέα, εξειδικευμένη εργασία (target task) [46], [56]. Αντί δηλαδή να εκπαιδεύεται ένα δίκτυο από μηδέν (de-novo training) — μια διαδικασία με υψηλές υπολογιστικές απαιτήσεις και μεγάλες ανάγκες σε δεδομένα —, αξιοποιούνται τα συναπτικά βάρη (weights) που το μοντέλο απέκτησε κατά την αρχική εκπαίδευσή του [39], [57].

Η προσαρμογή ενός προ-εκπαιδευμένου δικτύου στο νέο πρόβλημα βασίζεται σε δύο κύριες προσεγγίσεις [46].

1. **Πάγωμα επιπέδων (Freezing layers):** Τα αρχικά, ρηχότερα επίπεδα ενός CNN εκπαιδεύονται να ανιχνεύουν βασικά, χαμηλού επιπέδου χαρακτηριστικά (low-level features): ακμές, απλά σχήματα, υφές. Επειδή αυτά τα χαρακτηριστικά είναι κοινά σε σχεδόν κάθε ψηφιακή εικόνα [56], [57], τα αντίστοιχα επίπεδα παραμένουν «παγωμένα» κατά την εκπαίδευση στο νέο σύνολο δεδομένων — τα βάρη τους δεν μεταβάλλονται.
2. **Μικρο-συντονισμός (Fine-tuning):** Τα βαθύτερα επίπεδα, που είναι υπεύθυνα για πιο αφηρημένα και εξειδικευμένα χαρακτηριστικά (high-level features), επανεκπαιδεύονται στο νέο σύνολο δεδομένων. Στόχος είναι το δίκτυο να αναγνωρίζει τις ιδιαιτερότητες της ιατρικής εικόνας — όπως τα ακανόνιστα όρια ενός όγκου στον μαστό [46].

Η ενσωμάτωση της TL αντιμετωπίζει άμεσα το λεγόμενο "πρόβλημα του μικρού μεγέθους δεδομένων" (Small-size Dataset Problem — SSDP), που παραμένει το κυριότερο εμπόδιο στην ανάπτυξη ιατρικών εφαρμογών τεχνητής νοημοσύνης [58]. Η αξία της τεκμηριώνεται μέσα από τα ακόλουθα καθοριστικά πλεονεκτήματα:

- **Αποτροπή της Υπερπροσαρμογής:** Τα βαθιά νευρωνικά δίκτυα, με τις εκατομμύρια παραμέτρους τους, έχουν την τάση να "απομνημονεύουν" τα δεδομένα εκπαίδευσης αντί να μαθαίνουν γενικεύσιμα μοτίβα — με αποτέλεσμα να αποτυγχάνουν σε νέα, άγνωστα δεδομένα [52]. Η TL λειτουργεί ως μορφή κανονικοποίησης που βελτιώνει αισθητά την ικανότητα γενίκευσης [46], [53].
- **Μείωση Υπολογιστικού Κόστους και Χρόνου Εκπαίδευσης:** Η αξιοποίηση προ-εκπαιδευμένων μοντέλων συντομεύει δραστικά τον χρόνο εκπαίδευσης — από ημέρες σε ώρες ή ακόμα και λεπτά — και μειώνει τις απαιτήσεις σε υπολογιστική ισχύ σε σύγκριση με την εκπαίδευση από μηδενική βάση [38], [57].
- **Επίτευξη Υψηλής Ακρίβειας με Ελάχιστα Δείγματα:** Εξίσου σημαντική είναι η δυνατότητα επίτευξης υψηλής διαγνωστικής ακρίβειας με ελάχιστα δεδομένα. Μελέτες δείχνουν ότι ακόμη και με σύνολα δεδομένων της τάξης των 253 εικόνων, η σωστή εφαρμογή της μεθόδου σε συνδυασμό με βελτιστοποίηση υπερπαραμέτρων μπορεί να αποδώσει διαγνωστική ακρίβεια που αγγίζει ή υπερβαίνει το 88% [47].
- **Αποδεδειγμένη Αποτελεσματικότητα στην Ιατρική Διάγνωση:** Τέλος, η αποτελεσματικότητα της προσέγγισης έχει επιβεβαιωθεί σε ποικίλα κλινικά πεδία:

ανίχνευση όγκων εγκεφάλου με αρχιτεκτονικές όπως AlexNet, VGGNet και GoogLeNet, διάγνωση φυματίωσης από ακτινογραφίες θώρακα, εντοπισμός βλαβών σε εικόνες μαστογραφίας και υπερήχων [46], [58].

Για την ιατρική κοινότητα, όπου η συλλογή και η σήμανση δεδομένων είναι χρονοβόρα, ακριβή και απαιτητική διαδικασία [36], η TL αποτελεί μια ουσιαστική στρατηγική. Επιτρέπει στα μοντέλα AI να αποδίδουν κλινικά αποδεκτή διαγνωστική ακρίβεια, παρά την απουσία εκατομμυρίων καταγεγραμμένων ιατρικών περιστατικών [56].

3.3 Αρχιτεκτονικές Χαμηλής Πολυπλοκότητας

Στην ανάπτυξη ενός εύρωστου συστήματος ταξινόμησης για τον καρκίνο του μαστού, η επιλογή αρχιτεκτονικής δεν είναι τυπική απόφαση. Οι αρχιτεκτονικές χαμηλής πολυπλοκότητας — γνωστές ως Αρχιτεκτονικές Χαμηλής Πολυπλοκότητας — είναι απαραίτητες για την ενσωμάτωση διαγνωστικών εργαλείων AI σε συσκευές αιχμής (edge devices) και σε φορητά συστήματα POCUS [1], [4]. Σε αντίθεση με τα ογκώδη παραδοσιακά δίκτυα, τα μοντέλα αυτά σχεδιάστηκαν ώστε να διατηρούν υψηλή διαγνωστική ακρίβεια με ταυτόχρονα χαμηλές υπολογιστικές και ενεργειακές απαιτήσεις [1], [8] — προϋπόθεση που σε κλινικά περιβάλλοντα με περιορισμένους πόρους δεν είναι απλώς επιθυμητή, αλλά αναγκαία. Στη συνέχεια παρουσιάζονται οι τρεις αρχιτεκτονικές που αξιολογούνται στην παρούσα έρευνα.

3.3.1 MobileNetV2

Το MobileNetV2 αναπτύχθηκε από ερευνητές της Google με στόχο την εκτέλεση εργασιών μηχανικής όρασης σε κινητές και ενσωματωμένες συσκευές [59]. Η βασική του καινοτομία είναι η χρήση των διαχωρίσιμων συνελίξεων κατά βάθος (Depthwise Separable Convolutions). Η τεχνική αυτή αντιμετωπίζει το πρόβλημα του υψηλού υπολογιστικού κόστους αποσυνθέτοντας το χωρικό φιλτράρισμα (spatial filtering) από τον συνδυασμό καναλιών (channel combining) [18] — δύο λειτουργίες που στις κλασικές συνελίξεις εκτελούνται ταυτόχρονα. Ο διαχωρισμός αυτός μειώνει σημαντικά τον αριθμό των εκπαιδευσιμων παραμέτρων και τον όγκο των απαιτούμενων πράξεων, χωρίς αντίστοιχη πτώση στην απόδοση [59].

Το MobileNetV2 εισάγει επίσης τη δομή των ανεστραμμένων υπολειμματικών μπλοκ (inverted residuals) με ενσωματωμένους γραμμικούς στενωπούς (linear bottlenecks) [18], [60]. Η λογική είναι απλή: τα δεδομένα εισόδου προβάλλονται αρχικά σε έναν χώρο πολύ υψηλότερης

διάστασης, εκεί εφαρμόζονται οι συνελίξεις, και στη συνέχεια συμπιέζονται ξανά. Έτσι το δίκτυο διατηρεί τα κρίσιμα πληροφοριακά χαρακτηριστικά σε μια συμπαγή αναπαράσταση, χωρίς σημαντικές απώλειες [18]. Από πρακτική άποψη, το εξαιρετικά μικρό μέγεθος του μοντέλου — περίπου 3,5 MB— του επιτρέπει να εκτελεί inference και να παρέχει διάγνωση σε πραγματικό χρόνο, απαιτώντας περίπου 0,25 δευτερόλεπτα ανά εικόνα ακόμη και σε απλούς επεξεργαστές. Αυτό το καθιστά άμεσα συμβατό με φορητές κεφαλές υπερήχων χωρίς πρόσθετες απαιτήσεις υλικού [4].

3.3.2 EfficientNet

Το EfficientNet αντιπροσωπεύει μια διαφορετική φιλοσοφία κλιμάκωσης. Αντί να αυξάνει αυθαίρετα μία μόνο διάσταση του δικτύου —βάθος, πλάτος ή ανάλυση εισόδου— όπως κάνουν οι παραδοσιακές προσεγγίσεις, προτείνει τη Μέθοδο Σύνθετης Κλιμάκωσης (Compound Scaling Method): έναν ενιαίο συντελεστή που κλιμακώνει και τις τρεις διαστάσεις ταυτόχρονα και αναλογικά [61]. Η υποκείμενη λογική είναι εύλογη. Εάν η ανάλυση της εικόνας εισόδου αυξηθεί, το δίκτυο χρειάζεται αναπόφευκτα περισσότερα επίπεδα για να διευρύνει το δεκτικό του πεδίο (receptive field) και περισσότερα κανάλια για να αποτυπώσει τα πιο λεπτομερή πρότυπα που αναδύονται [61].

Το βασικό μοντέλο της οικογένειας, το EfficientNetB0, αναπτύχθηκε μέσω αυτοματοποιημένης αναζήτησης αρχιτεκτονικής (Neural Architecture Search, NAS), με στόχο τη διπλή βελτιστοποίηση ακρίβειας και υπολογιστικού κόστους (FLOPS) [61]. Στη βιβλιογραφία, το μοντέλο αυτό —ή η βελτιωμένη έκδοση V2-Small— εμφανίζεται συχνά ως η πιο αποδοτική επιλογή για ανάλυση υπερηχογραφήματων μαστού, επιτυγχάνοντας τιμές AUC που αγγίζουν το 0,97 με σημαντικά λιγότερες παραμέτρους από βαρύτερα μοντέλα όπως το VGG13 [1]. Τα αποτελέσματα της έρευνας των Ferreira et al. υποδεικνύουν επιπλέον ότι το EfficientNet αποδίδει καλύτερα όταν συνδυάζεται με στοχευμένη εξαγωγή της περιοχής ενδιαφέροντος (Στρατηγική "Input E"), η οποία απομονώνει τον όγκο διατηρώντας ένα λεπτό περιθώριο περιβάλλοντος ιστού [10].

3.3.3 NASNetMobile

Το NASNetMobile δεν σχεδιάστηκε χειροκίνητα. Προέκυψε από την εφαρμογή τεχνικών Αυτοματοποιημένης Μηχανικής Μάθησης (Automated Machine Learning, AutoML), όπου αλγόριθμοι ενισχυτικής μάθησης ανέλαβαν να αναζητήσουν τον βέλτιστο συνδυασμό επιπέδων

και συνδέσεων, με κριτήριο τη μέγιστη ακρίβεια σε περιβάλλοντα με αυστηρά περιορισμένους υπολογιστικούς πόρους [59], [60].

Δομικά, το μοντέλο βασίζεται σε μια κυτταρική αρχιτεκτονική (cell-based structure): επαναλαμβανόμενα δομικά στοιχεία, τα λεγόμενα κύτταρα (cells), τα οποία διαχωρίζονται σε δύο τύπους. Τα Κανονικά Κύτταρα (Normal Cells) εξάγουν τοπικά χαρακτηριστικά διατηρώντας τις χωρικές διαστάσεις, ενώ τα Κύτταρα Μείωσης (Reduction Cells) μειώνουν τη χωρική ανάλυση και εξάγουν πιο αφηρημένες αναπαραστάσεις [59]. Σε κλινικές μελέτες για τη διάγνωση καρκίνου μαστού, το NASNetMobile έχει επιδείξει ισχυρή ικανότητα γενίκευσης. Όταν βελτιστοποιείται περαιτέρω με ευρετικούς αλγορίθμους όπως ο Simplified Swarm Optimization, μπορεί να επιτύχει συμπίεση του μεγέθους έως και 96%, διατηρώντας ανέπαφη τη διακριτική του ικανότητα [7].

Συμπερασματική Θεώρηση της Μεθοδολογίας

Η επιλογή των τριών αυτών μοντέλων δεν είναι τυχαία. Το MobileNetV2 προσφέρει την ταχύτητα εκτέλεσης που απαιτεί η κλινική πράξη, το EfficientNet βελτιστοποιεί τη σχέση διαγνωστικής ακρίβειας και υπολογιστικού κόστους μέσω της σύνθετης κλιμάκωσης, και το NASNetMobile αποδεικνύει ότι αρχιτεκτονικές που παράγονται αυτόματα μπορούν να ανταποκριθούν αποτελεσματικά σε απαιτητικά ιατρικά δεδομένα [62]. Μαζί, καλύπτουν το ερευνητικό ερώτημα για Αποδοτική AI από τρεις διαφορετικές, συμπληρωματικές οπτικές.

3.4 Μηχανισμοί Προσοχής: Θεωρία του CBAM

Στον τομέα της ψηφιακής όρασης, οι μηχανισμοί προσοχής (attention mechanisms) επιτρέπουν στα νευρωνικά δίκτυα να εστιάζουν δυναμικά στα πιο πληροφοριακά τμήματα μιας εικόνας, αποκλείοντας παράλληλα τον οπτικό θόρυβο και τις άσχετες περιοχές. Μία από τις πιο κομψές αρχιτεκτονικές αυτής της κατηγορίας είναι το Convolutional Block Attention Module (CBAM), που προτάθηκε από τους Woo et al. το 2018. Πρόκειται για έναν απλό, αλλά εξαιρετικά αποτελεσματικό μηχανισμό ενίσχυσης της αναπαραστατικής ικανότητας των CNNs [63], ο οποίος επεκτείνει λογικά προηγούμενες τεχνικές —όπως τα δίκτυα Squeeze-and-Excitation— προσθέτοντας μια ουσιαστικά πιο πλήρη διαχείριση των εξαγόμενων χαρακτηριστικών.

Η Φιλοσοφία του "Τι" και του "Πού"

Το CBAM οργανώνεται γύρω από δύο ερωτήματα που θέτει κατά την ανάλυση των δεδομένων: *τι* είναι σημαντικό και *πού* εντοπίζεται στον χώρο. Για να τα απαντήσει, εφαρμόζει δύο διαδοχικές υπο-ενότητες στους χάρτες χαρακτηριστικών (feature maps) της εισόδου. Η πρώτη, η Προσοχή Καναλιού (Channel Attention), ασχολείται με το "τι": αναδεικνύει επιλεκτικά τα κανάλια εκείνα που φέρουν την πιο διακριτή και χρήσιμη πληροφορία για το εκάστοτε πρόβλημα [63], [64]. Ακολουθεί η Χωρική Προσοχή (Spatial Attention), που εντοπίζει το "πού": κατευθύνει την προσοχή του δικτύου στις χωρικά σημαντικές περιοχές της εικόνας, αξιοποιώντας τις χωρικές σχέσεις που αναπτύσσονται μεταξύ των χαρακτηριστικών [64].

Ανατομία της Αρχιτεκτονικής του CBAM

Η Ενότητα Προσοχής Καναλιού διαφοροποιείται από παλαιότερες προσεγγίσεις σε ένα συγκεκριμένο σημείο. Αντί να βασίζεται αποκλειστικά στη μέση ομαδοποίηση (global average pooling), συνδυάζει αυτήν με τη μέγιστη ομαδοποίηση, εξάγοντας έτσι πλουσιότερα στατιστικά στοιχεία από τους χάρτες χαρακτηριστικών [64]. Η μέγιστη ομαδοποίηση αποδεικνύεται χρήσιμο για τον εντοπισμό οξείων, διακριτών χαρακτηριστικών —όπως ακμές ή ασβεστοποιήσεις σε ιατρικές εικόνες— ενώ το average pooling αποτυπώνει τη γενικότερη στατιστική κατανομή. Ο συνδυασμός των δύο έχει τεκμηριωθεί ότι ενισχύει σημαντικά τη δύναμη αναπαράστασης του δικτύου σε σχέση με τη μεμονωμένη χρήση οποιασδήποτε εκ των δύο τεχνικών.

Η έξοδος της ενότητας καναλιού διοχετεύεται στη συνέχεια στην Ενότητα Χωρικής Προσοχής. Εκεί, λειτουργίες μέσου όρου και μέγιστης ομαδοποίησης εφαρμόζονται αυτή τη φορά κατά μήκος του άξονα των καναλιών· τα αποτελέσματά τους συνενώνονται σε έναν ενιαίο χωρικό περιγραφέα. Μέσω αυτής της διαδικασίας, το δίκτυο μαθαίνει ποιες περιοχές της εικόνας χρήζουν μεγαλύτερης προσοχής, καθοδηγώντας έτσι την τελική απόφαση ταξινόμησης ή ανίχνευσης [65].

Αποδοτικότητα και Εφαρμοσιμότητα

Ένα από τα πιο πρακτικά χαρακτηριστικά του CBAM είναι η υπολογιστική του οικονομία. Η σειριακή εκμάθηση των χαρτών προσοχής παράγει έναν σύνθετο τρισδιάστατο χάρτη με αισθητά μειωμένο αριθμό παραμέτρων, διατηρώντας το υπολογιστικό κόστος σε χαμηλά επίπεδα. Εξίσου σημαντική είναι η αρθρωτή του δομή: ο σχεδιασμός του ως ανεξάρτητο γενικό μπλοκ (generic block) επιτρέπει την ενσωμάτωσή του σε οποιαδήποτε υφιστάμενη αρχιτεκτονική CNN —ResNet,

MobileNet, ResNeXt— με ελάχιστο πρόσθετο φόρτο [63], [65]. Επιπλέον, η προσθήκη του βελτιώνει τη ροή της πληροφορίας εντός του δικτύου, γεγονός που αποκτά ιδιαίτερη βαρύτητα σε εφαρμογές με θορυβώδη δεδομένα, όπως οι ιατρικές εικόνες υπερήχων [63].

Το CBAM, εκμεταλλευόμενο ταυτόχρονα τη σημασία των καναλιών και τη χωρική τοποθεσία της πληροφορίας, προσφέρει μετρήσιμη βελτίωση της διαγνωστικής ακρίβειας σε απαιτητικές εργασίες ψηφιακής όρασης [63].

3.5 Μετρικές Αξιολόγησης

Η αξιολόγηση της απόδοσης ενός μοντέλου ML — ιδίως σε ιατρικές εφαρμογές — δεν μπορεί να στηρίζεται σε έναν και μόνο αριθμό. Μια απλή μετρική, όπως η ακρίβεια, συχνά αδυνατεί να αποδώσει πλήρως την αποτελεσματικότητα ενός αλγορίθμου [46], [66]. Γι' αυτό, οι περισσότερες προηγμένες μέθοδοι αξιολόγησης βασίζονται στον πίνακα σύγχυσης (confusion matrix): ένα εργαλείο που αποτυπώνει τις προβλέψεις του μοντέλου σε αντιπαραβολή με τις πραγματικές τιμές (ground truth) [46].

Ο πίνακας αυτός οργανώνεται γύρω από τέσσερις βασικές κατηγορίες. Τα Αληθώς Θετικά (True Positives, TP) είναι τα δείγματα που το μοντέλο εντόπισε ορθά ως θετικά, ενώ τα Αληθώς Αρνητικά (True Negatives, TN) αντιστοιχούν στα δείγματα που ταξινομήθηκαν σωστά ως αρνητικά [43], [48], [51]. Τα σφάλματα διακρίνονται σε δύο είδη: τα Ψευδώς Θετικά (False Positives, FP, Σφάλμα Τύπου I), όπου το μοντέλο απέδωσε θετική κλάση σε πραγματικά αρνητικό δείγμα, και τα Ψευδώς Αρνητικά (False Negatives, FN, Σφάλμα Τύπου II), όπου ένα πραγματικά θετικό δείγμα εσφαλμένα κατατάχθηκε ως αρνητικό [43], [48], [51]. Από αυτές τις τέσσερις ποσότητες προκύπτουν όλες οι κύριες μετρικές που ακολουθούν.

3.5.1 Ορθότητα / Ακρίβεια

Η Ακρίβεια (Accuracy) είναι η πιο διαισθητική από τις μετρικές αξιολόγησης. Ορίζεται ως το ποσοστό των συνολικά ορθών προβλέψεων — αληθώς θετικών και αληθώς αρνητικών — επί του συνόλου των δειγμάτων) [46], και υπολογίζεται βάσει του τύπου [65]:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (3.1)$$

Η μετρική αυτή δίνει μια πρώτη, γενική εικόνα απόδοσης, αλλά μπορεί να παραπλανήσει όταν τα δεδομένα είναι ανισοκατανεμημένα (imbalanced datasets). Σε τέτοια κλινικά σενάρια, ένα μοντέλο

που προβλέπει μονότονα την κυρίαρχη — συνήθως υγιή — κλάση εμφανίζει τεχνητά υψηλή ακρίβεια, ενώ αδυνατεί ουσιαστικά να εντοπίσει τη σπάνια παθολογική κλάση [66].

3.5.2 Ακρίβεια Πρόβλεψης

Η Ακρίβεια Πρόβλεψης (Precision) απαντά σε ένα συγκεκριμένο ερώτημα: από όλα τα δείγματα που το μοντέλο χαρακτήρισε ως θετικά, πόσα ήταν πράγματι τέτοια [67]; Υπολογίζεται ως [43]:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3.2)$$

Η μετρική αυτή αποκτά ιδιαίτερη βαρύτητα σε κλινικά πλαίσια όπου ένα ψευδώς θετικό αποτέλεσμα συνεπάγεται υψηλό κόστος — οικονομικό ή ανθρώπινο. Χαρακτηριστικό παράδειγμα αποτελεί η περίπτωση που μια εσφαλμένη θετική πρόβλεψη οδηγεί σε περιττή και επώδυνη ιατρική βιοψία [66].

3.5.3 Ανάκληση / Ευαισθησία

Η Ανάκληση ή Ευαισθησία (Recall / Sensitivity) μετρά κάτι διαφορετικό: την ικανότητα του μοντέλου να εντοπίζει το σύνολο των πραγματικά θετικών δειγμάτων [51]. Ο τύπος υπολογισμού της έχει ως εξής [66]:

$$\text{Recall} = \frac{TP}{TP + FN} \quad (3.3)$$

Σε προβλήματα ιατρικής διάγνωσης, αυτή η μετρική έχει συχνά μεγαλύτερη πρακτική σημασία από οποιαδήποτε άλλη. Η αποτυχία εντοπισμού ενός κακοήθους όγκου — δηλαδή ένα False Negative σφάλμα — μπορεί να έχει σοβαρές, ακόμη και μη αναστρέψιμες, συνέπειες για τον ασθενή [66].

3.5.4 Το Μέτρο F1

Δεδομένου ότι συχνά υπάρχει αντίστροφη σχέση (trade-off) μεταξύ Precision και Recall, το F1-Score ορίζεται ως ο αρμονικός μέσος των δύο αυτών μετρικών [43]. Μαθηματικά εκφράζεται ως [43]:

$$F1 - \text{Score} = 2 * \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3.4)$$

Το F1-Score χρησιμοποιείται κυρίως όταν ζητείται μια ισορροπημένη εκτίμηση τόσο της ακρίβειας πρόβλεψης όσο και της ευαισθησίας. Σε σύνολα δεδομένων με ανισοκατανομή κλάσεων, η απλή μετρική Ακρίβειας (Accuracy) τείνει να αποκρύπτει την πραγματική απόδοση του ταξινομητή· σε αυτές τις περιπτώσεις, το F1-Score αποδεικνύεται πιο αξιόπιστος δείκτης αξιολόγησης [43].

3.5.5 Εμβαδόν Κάτω από την Καμπύλη

Η μετρική AUC (Area Under the Curve) αντιστοιχεί στο εμβαδόν κάτω από την καμπύλη ROC (Receiver Operating Characteristic) [46]. Η καμπύλη αυτή απεικονίζει γραφικά τη σχέση μεταξύ του ρυθμού αληθώς θετικών (Sensitivity) και του ρυθμού ψευδώς θετικών ($1 - \text{Specificity}$) για διαφορετικά κατώφλια ταξινόμησης [46]. Οι τιμές του AUC κυμαίνονται από 0 έως 1. Τιμή κοντά στο 1 δηλώνει ισχυρή διαχωριστική ικανότητα μεταξύ κλάσεων — για παράδειγμα, ικανότητα διάκρισης υγιών από ασθενείς — ενώ τιμή γύρω στο 0.5 αντιστοιχεί σε πρόβλεψη που δεν υπερβαίνει την τυχαία [67]. Ένα βασικό χαρακτηριστικό του AUC είναι ότι αξιολογεί τη συνολική απόδοση του μοντέλου ανεξάρτητα από την επιλογή συγκεκριμένου κατωφλίου.

Η συνδυαστική χρήση των παραπάνω μετρικών επιτρέπει μια πιο ολοκληρωμένη εικόνα της απόδοσης ενός μοντέλου μηχανικής μάθησης, πράγμα ιδιαίτερα σημαντικό πριν την εφαρμογή του σε κλινικά σενάρια [66].

3.6 Σύνοψη κεφαλαίου

Το παρόν κεφάλαιο ανέπτυξε το αλγοριθμικό και θεωρητικό υπόβαθρο της έρευνας. Αναλύθηκαν οι θεμελιώδεις αρχές των CNNs, ο ρόλος της TL σε συνθήκες περιορισμένων ιατρικών δεδομένων, και οι ιδιαιτερότητες των αρχιτεκτονικών χαμηλής πολυπλοκότητας και των μηχανισμών προσοχής. Ορίστηκαν επίσης τα εργαλεία αξιολόγησης που θα χρησιμοποιηθούν στα επόμενα στάδια.

Με αυτά ως δεδομένα, η θεωρία περνά στο επίπεδο της εφαρμογής στο Κεφάλαιο 4. Παρουσιάζεται αναλυτικά η μεθοδολογία και η ανάπτυξη του συστήματος: η διαχείριση των συνόλων δεδομένων (BUSI, BUS-BRA), η στρατηγική προ-επεξεργασίας "Input E", ο σχεδιασμός και η εκπαίδευση των μοντέλων, και τέλος οι στρατηγικές αξιολόγησης (Stratified Group 5-Fold CV, TTA, Optimal Threshold Tuning) που διασφαλίζουν την κλινική αξιοπιστία του συστήματος.

Κεφάλαιο 4: Μεθοδολογία και Ανάπτυξη Συστήματος

Το κεφάλαιο αυτό παρουσιάζει τη μεθοδολογία και τα τεχνικά στάδια υλοποίησης του συστήματος ταξινόμησης. Ξεκινάμε από τα σύνολα δεδομένων και τη στρατηγική προ-επεξεργασίας «Input E». Στη συνέχεια εξετάζονται οι αρχιτεκτονικές χαμηλής πολυπλοκότητας (Lightweight CNNs) που επιλέχθηκαν, καθώς και οι στρατηγικές εκπαίδευσης που υιοθετήθηκαν — στρωματοποιημένη διασταυρούμενη επικύρωση, ρύθμιση κατωφλίου — με στόχο την κλινική αξιοπιστία του μοντέλου.

4.1 Σύνολα Δεδομένων

Η ανάπτυξη και αξιολόγηση αξιόπιστων συστημάτων Υποβοηθούμενης Διάγνωσης (Computer-Aided Diagnosis, CAD) για τον καρκίνο του μαστού απαιτεί πρόσβαση σε κλινικά τεκμηριωμένα σύνολα δεδομένων υψηλής ποιότητας, καθώς και εφαρμογή τυποποιημένων συστημάτων αναφοράς όπως του Συστήματος Αναφοράς και Δεδομένων Απεικόνισης Μαστού (Breast Imaging Reporting and Data System, BI-RADS). Για τους σκοπούς της παρούσας εργασίας αξιοποιήθηκαν δύο από τα πλέον αναγνωρισμένα δημόσια datasets: το BUSI και το νεότερο BUS-BRA.

4.1.1 Περιγραφή του Συνόλου Δεδομένων BUSI

Το σύνολο δεδομένων Breast Ultrasound Images (BUSI) δημιουργήθηκε το 2018, με εικόνες που συλλέχθηκαν στο εξειδικευμένο νοσοκομείο Baheya στο Κάιρο [3], [4], [18]. Αποτελείται συνολικά από 780 υπερηχογραφικές εικόνες σε κλίμακα του γκρι, από 600 γυναίκες ηλικίας 25 έως 75 ετών

[1], [4], [18]. Οι εικόνες είναι αποθηκευμένες σε μορφή PNG με μέση ανάλυση 500×500 pixels [1], [4], [18].

Ως προς την κλινική κατηγοριοποίηση, το dataset διαχωρίζεται σε τρεις κλάσεις: 437 εικόνες αφορούν καλοήθεις παθήσεις, 210 κακοήθεις όγκους και 133 φυσιολογικούς ιστούς [1], [4], [25]. Ένα από τα βασικά πλεονεκτήματα του BUSI είναι η λεπτομερής τεκμηρίωσή του. Κάθε εικόνα συνοδεύεται από αντίστοιχη δυαδική μάσκα (ground truth mask) που υποδεικνύει την ακριβή θέση και τα όρια της βλάβης — κάτι που διευκολύνει σημαντικά την εκπαίδευση μοντέλων τόσο για τμηματοποίηση (segmentation) όσο και για εξαγωγή περιοχών ενδιαφέροντος (Region of Interest, ROI) [5], [18], [26].

4.1.2 Περιγραφή του Συνόλου Δεδομένων BUS-BRA

Για την ενίσχυση της ποικιλομορφίας και της ικανότητας γενίκευσης του συστήματος, ενσωματώθηκε το Breast Ultrasound Dataset from Brazil (BUS-BRA), ένα πιο πρόσφατο (2.024) και εκτενές σύνολο δεδομένων από το Εθνικό Ινστιτούτο Καρκίνου στο Ρίο ντε Τζανέιρο [6]. Περιλαμβάνει 1.875 ανωνυμοποιημένες εικόνες από 1.064 ασθενείς, με πλήρη ιατρικό σχολιασμό [6], [68].

Η ετερογένεια του συνόλου αντανακλά τη χρήση τεσσάρων διαφορετικών κλινικών σαρωτών (GE Logiq 5, GE Logiq 7, Toshiba Aplio 300, GE U-Systems), γεγονός που εισάγει σημαντική ποικιλία στην ποιότητα και την οπτική παρουσίαση των εικόνων. Αυτό εκπαιδεύει το μοντέλο να λειτουργεί ανεξάρτητα από τον εξοπλισμό λήψης [6], [68]. Για το 76% των ασθενών, επιπλέον, το dataset παρέχει δύο ορθογώνιες όψεις της ίδιας βλάβης (αριστερή και δεξιά λήψη), προσφέροντας πληρέστερη χωρική αποτύπωση [6].

Ο ιατρικός σχολιασμός είναι ιδιαίτερα λεπτομερής. Πέρα από τα αποτελέσματα βιοψίας — 722 καλοήθεις και 342 κακοήθεις περιπτώσεις — και τις χειροκίνητες μάσκες τμηματοποίησης, το BUS-BRA περιλαμβάνει ρητές αξιολογήσεις βάσει της κλίμακας BI-RADS από έμπειρους ακτινολόγους [6].

4.1.3 Ενοποίηση και Αντιστοίχιση BI-RADS σε Δυαδική Ταξινόμηση

Το σύστημα BI-RADS της Αμερικανικής Ακτινολογικής Εταιρείας ορίζει μια πολυταξική κλίμακα από 0 έως 6 για την εκτίμηση του κινδύνου κακοήθειας [6], [29]. Για την εκπαίδευση μοντέλων AI,

ωστόσο, η μετατροπή της σε πρόβλημα δυαδικής ταξινόμησης κρίνεται αναγκαία [3], [6]. Στην παρούσα εργασία η αντιστοίχιση μεταξύ των δύο datasets έγινε με βάση την εξής κλινική λογική:

Στην καλοήγη κλάση εντάχθηκαν οι κατηγορίες BI-RADS 2 (σαφώς καλοήγη ευρήματα) και BI-RADS 3 (πιθανώς καλοήγη ευρήματα), για τις οποίες η ενδεδειγμένη κλινική σύσταση είναι η περιοδική παρακολούθηση μέσω απεικονιστικών εξετάσεων [6], [29]. Στην κακοήγη κλάση εντάχθηκαν οι κατηγορίες BI-RADS 4 (ύποπτα ευρήματα) και BI-RADS 5 (εξαιρετικά υψηλή υποψία κακοήθειας), για τις οποίες η ιατρική πρακτική επιβάλλει διενέργεια βιοψίας [6], [29].

Στη βιβλιογραφία απαντώνται και πιο λεπτομερείς στρατηγικές ενοποίησης. Η κατηγορία BI-RADS 4, για παράδειγμα, καλύπτει ένα εξαιρετικά ευρύ φάσμα πιθανότητας κακοήθειας — από 2% έως 95% — και ορισμένα μοντέλα την υποδιαιρούν σε υποομάδες (4a, 4b, 4c) ή τη διαχωρίζουν μέσω αποτελεσμάτων βιοψίας κατά την εκπαίδευση [6], [29], [32]. Έχει επίσης προταθεί η προσέγγιση One-Versus-All (OVA), όπου αντί ενός ενιαίου πολυταξικού μοντέλου εκπαιδεύονται ανεξάρτητοι ταξινομητές για κάθε κατηγορία BI-RADS, οι οποίοι στη συνέχεια συνδυάζονται για την τελική δυαδική απάντηση [29].

Η δυαδική αυτή ενοποίηση επιτρέπει στο σύστημα CAD να παρέχει σαφείς και άμεσα αξιοποιήσιμες συστάσεις στους ακτινολόγους, με στόχο τη μείωση των περιπτώσεων βιοψιών και την ταυτόχρονη βελτίωση των ποσοστών έγκαιρης ανίχνευσης κακοηθειών [3], [6].

4.2 Προ-επεξεργασία Δεδομένων

Η προ-επεξεργασία των υπερηχογραφικών δεδομένων αποτελεί ένα από τα πιο καθοριστικά στάδια για την επιτυχή εκπαίδευση μοντέλων DL. Στην ενότητα αυτή παρουσιάζεται η στρατηγική «Input E» για την απομόνωση της ROI και την εξάλειψη του οπτικού θορύβου, καθώς και οι τεχνικές ενίσχυσης δεδομένων (data augmentation) που εφαρμόστηκαν για τη βελτίωση της ικανότητας γενίκευσης του συστήματος.

4.2.1 Η Στρατηγική "Input E"

Η στρατηγική προ-επεξεργασίας «Input E», γνωστή στη βιβλιογραφία και ως εικόνα ενισχυμένης βλάβης με διασταλμένη ROI, ανήκει στις πιο αποδοτικές μεθόδους μετασχηματισμού υπερηχογραφικών εικόνων πριν από την εισαγωγή τους σε CNNs [10]. Η λογική της προσέγγισης είναι σχετικά απλή: εστιάζει αποκλειστικά στην ROI, διατηρώντας ταυτόχρονα χρήσιμες κλινικές

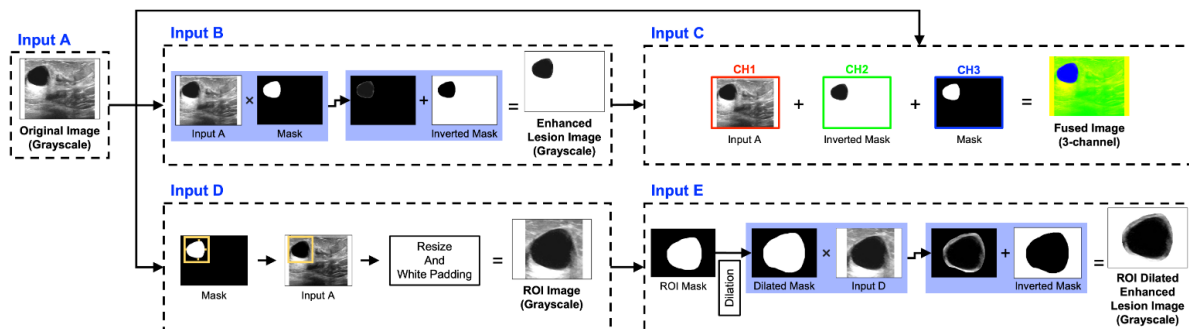
πληροφορίες από τον περιβάλλοντα ιστό — και ιδίως την ακουστική σκίαση (acoustic shadowing) [10].

Η παραγωγή της τελικής εικόνας ακολουθεί μια αλληλουχία τριών σταδίων: Διαστολή, Εφαρμογή Μάσκας και Περικοπή.

Διαστολή της Μάσκας (Dilation). Ως αφετηρία χρησιμεύει η δυαδική μάσκα τμηματοποίησης (*ground truth mask*), οριοθετημένη χειροκίνητα από έμπειρους ακτινολόγους, η οποία υποδεικνύει την ακριβή θέση του όγκου [10]. Πάνω σε αυτή εφαρμόζεται αλγόριθμος μορφολογικής διαστολής που διευρύνει την επιφάνεια της λευκής περιοχής κατά ένα σταθερό ποσοστό [10]. Έρευνες δείχνουν ότι το βέλτιστο αποτέλεσμα επιτυγχάνεται με διαστολή 70% σε σχέση με το αρχικό μέγεθος της βλάβης. Το συγκεκριμένο ποσοστό δεν είναι αυθαίρετο· επιτρέπει τη συμπερίληψη της «άλω» (*border zone*), δηλαδή της μεταβατικής ζώνης ανάμεσα στην αλλοίωση και τον υγιή ιστό, η οποία φέρει διαγνωστικές πληροφορίες για την κακοήθεια [10].

Εφαρμογή Μάσκας και Αφαίρεση Υποβάθρου (Masking). Στο δεύτερο στάδιο, η διασταλμένη μάσκα (*dilated mask*) εφαρμόζεται πάνω στην αρχική υπερηχογραφική εικόνα (Input A) με πολλαπλασιασμό εικονοστοιχείου προς εικονοστοιχείο (*pixel-by-pixel multiplication*) [10]. Έτσι απομονώνεται ο όγκος μαζί με το λεπτό περιθώριο που προέκυψε από τη διαστολή, ενώ το υπόλοιπο φόντο αποκόπτεται. Στη συνέχεια, με πράξεις πάνω στην ανεστραμμένη μάσκα (*inverted mask*), το μαύρο φόντο αντικαθίσταται από λευκό (*background enhancement*) [10]. Η επιλογή αυτή ενισχύει την αντίθεση και διευκολύνει το νευρωνικό δίκτυο να αναγνωρίσει τα ακανόνιστα, μικρολοβώδη ή αστεροειδή όρια του όγκου [10].

Περικοπή και Αλλαγή Μεγέθους (Cropping & Resizing). Το τελευταίο στάδιο αφορά τον χωρικό περιορισμό της εικόνας. Η περικοπή γίνεται αυστηρά στα όρια του οριακού πλαισίου (*bounding box*) που περικλείει τη διασταλμένη περιοχή του όγκου [10]. Αφού αποκοπεί το περιττό λευκό φόντο, η εικόνα υφίσταται αλλαγή μεγέθους (*resizing*) και, εφόσον κριθεί απαραίτητο, προσθήκη λευκού περιθωρίου (*white padding*), ώστε να αποκτήσει τις τυποποιημένες διαστάσεις — συνήθως 224×224 pixels — που απαιτούν οι αρχιτεκτονικές βαθιών CNNs [10].



ΣΧΗΜΑ 4: ΓΡΑΦΙΚΗ ΑΝΑΠΑΡΑΣΤΑΣΗ ΤΩΝ ΔΙΑΦΟΡΩΝ ΜΟΡΦΩΝ ΔΕΔΟΜΕΝΩΝ ΕΙΣΟΔΟΥ. ΔΙΑΚΡΙΝΕΤΑΙ Η ΑΚΡΙΒΗΣ ΑΛΛΗΛΟΥΧΙΑ ΒΗΜΑΤΩΝ (ΔΙΑΣΤΟΛΗ, ΕΦΑΡΜΟΓΗ ΜΑΣΚΑΣ, ΠΕΡΙΚΟΠΗ) ΓΙΑ ΤΗ ΔΗΜΙΟΥΡΓΙΑ ΤΗΣ ΕΙΚΟΝΑΣ "INPUT E" [10]

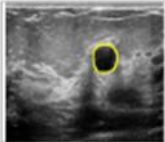
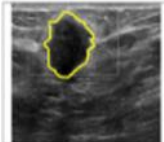
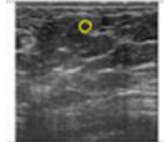
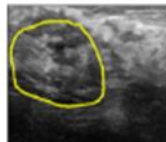
Οπτική Αναπαράσταση και Κλινική Αξία (Πριν και Μετά)

Η αποτελεσματικότητα της μεθόδου γίνεται άμεσα αντιληπτή με μια απλή οπτική σύγκριση της αρχικής εικόνας (Input A) με την τελικά παραγόμενη (Input E) [10].

Στην Input A προβάλλεται ολόκληρο το πεδίο του υπερήχου. Η βλάβη καταλαμβάνει συνήθως ένα μικρό τμήμα του πλαισίου και περιβάλλεται από πολύπλοκο ινώδη ιστό, ακουστικό θόρυβο (speckle noise) και ενδείξεις ή γράμματα του μηχανήματος — παράγοντες που συχνά αποπροσανατολίζουν το μοντέλο [10]. Στην Input E, αντίθετα, η προσοχή του δικτύου κατευθύνεται αποκλειστικά στον όγκο. Στο κέντρο διατηρείται η βλάβη με πλήρη ηχομορφολογική λεπτομέρεια, ενώ στην περιφέρειά της παραμένει ένα λεπτό στρώμα περιβάλλοντος ιστού, αρκετό ώστε να αποτυπώνεται η σχέση του όγκου με τον γύρω χώρο — π.χ. ενδείξεις διήθησης [10]. Το λευκό φόντο εξαλείφει κάθε άλλη άσχετη πληροφορία [10].

Αποτελεσματικότητα της Στρατηγικής

Όπως προκύπτει από τη συγκριτική ανάλυση των διαφόρων μορφών εισόδου (ΣΧΗΜΑ 4), η χρήση της Input E βελτιώνει σταθερά την απόδοση των μοντέλων [6], [8]. Χαρακτηριστικό παράδειγμα αποτελεί το EfficientNet, το οποίο, χρησιμοποιώντας αποκλειστικά αυτή τη μορφή εισόδου, πέτυχε Ακρίβεια (Accuracy) 98,82% και AUC 98,15 — αποτελέσματα που ξεπερνούν κατά πολύ εκείνα που επιτυγχάνονται με τις αρχικές, μη επεξεργασμένες εικόνες [6], [8].

Υπέρηχος				
Ground Truth	Benign	Malignant	Malignant	Benign
Input A (EfficientNet)	Benign	Malignant	Benign	Malignant
Input B (EfficientNet)	Benign	Malignant	Benign	Benign
Input C (EfficientNet)	Benign	Malignant	Benign	Malignant
Input D (GoogLeNet)	Benign	Malignant	Malignant	Malignant
Input E (EfficientNet)	Benign	Malignant	Benign	Benign

ΠΙΝΑΚΑΣ 1: ΠΑΡΑΔΕΙΓΜΑΤΑ ΥΠΕΡΗΧΟΓΡΑΦΗΜΑΤΩΝ ΑΠΟ ΤΟ ΣΥΝΟΛΟ ΔΟΚΙΜΗΣ ΚΑΙ ΟΙ ΑΝΤΙΣΤΟΙΧΕΣ ΠΡΟΒΛΕΨΕΙΣ ΤΩΝ ΜΟΝΤΕΛΩΝ ΑΝΑΛΟΓΑ ΜΕ ΤΗ ΜΟΡΦΗ ΕΙΣΟΔΟΥ [10].

Τα παραδείγματα δοκιμών (ΠΙΝΑΚΑΣ 1) δείχνουν επίσης ότι η Input E επιτρέπει στο δίκτυο να ταξινομεί ορθά οριακές περιπτώσεις που αποδείχθηκαν ιδιαίτερα δύσκολες — μικρούς κακοήθεις όγκους με ασαφή όρια ή υποηχογενείς καλοήθεις κύστες — στις οποίες όλες οι υπόλοιπες μέθοδοι εισόδου (Input A έως D) οδήγησαν σε λανθασμένες προβλέψεις.

4.2.2 Τεχνικές Ενίσχυσης Δεδομένων

Οι τεχνικές ενίσχυσης δεδομένων (Data Augmentation) επηρεάζουν σημαντικά την απόδοση, την ευρωστία και τη γενικευτική ικανότητα των μοντέλων DL. Η σημασία τους είναι ιδιαίτερα μεγάλη στην ανάλυση ιατρικών εικόνων, έναν τομέα όπου μεγάλα και κλινικά σχολιασμένα σύνολα δεδομένων είναι συχνά δυσεύρετα [55], [69]. Η διαδικασία στηρίζεται στην τεχνητή επέκταση του συνόλου εκπαίδευσης μέσω μαθηματικών και οπτικών μετασχηματισμών στις αρχικές εικόνες, με στόχο να αποτραπεί το φαινόμενο της υπερεκπαίδευσης και να εκπαιδευτεί το μοντέλο σε συνθήκες που αντανακλούν την πραγματική κλινική ποικιλομορφία [11], [70].

Γεωμετρικοί Μετασχηματισμοί

Οι γεωμετρικοί μετασχηματισμοί αποτελούν την πιο διαδεδομένη κατηγορία ενίσχυσης. Στοχεύουν στην προσομοίωση των φυσικών παραλλαγών που προκύπτουν από τη διαφορετική τοποθέτηση, πίεση ή γωνία λήψης του ηχοβολέα (probe) κατά τη διάρκεια εξέτασης. Η περιστροφή (rotation) της εικόνας, συνήθως έως 30°, είναι από τους πιο αποδοτικούς μεμονωμένους

μετασχηματισμούς και έχει οδηγήσει σε μετρήσιμη βελτίωση της ακρίβειας ταξινόμησης σε σύνολα δεδομένων όπως το BUSI [11], [69]. Η οριζόντια αναστροφή (horizontal flip) επιλέγεται σχεδόν αποκλειστικά έναντι της κατακόρυφης [1], [11], η οποία αποφεύγεται επειδή διαταράσσει την ανατομική στρωμάτωση των ιστών — όπως τη φυσιολογική θέση του στρώματος λίπους σε σχέση με το μαστικό στρώμα [11]. Επιπλέον, εφαρμόζονται μετατοπίσεις πλάτους και ύψους (τυπικά περί το 10%), τυχαία κλιμάκωση (zoom) μεταξύ 80% και 120% [1], [11], καθώς και διάτμηση (shear) κατά μήκος ενός άξονα, συνήθως με τιμή 0,2 radians [11].

Μετασχηματισμοί Έντασης, Φωτεινότητας και Υφής

Πέραν των γεωμετρικών αλλαγών, χρησιμοποιούνται μετασχηματισμοί που αναπαράγουν οπτικές διαφορές από τη χρήση διαφορετικών σαρωτών ή από τις ρυθμίσεις εξέτασης. Η διόρθωση Γάμμα (Gamma Correction), με προσαρμογή φωτεινότητας και αντίθεσης (π.χ. τιμή 0,7), βοηθά το μοντέλο να εστιάσει σε λεπτομερή χαρακτηριστικά της βλάβης [13]. Η προσθήκη τεχνητού θορύβου (Noise Injection) — όπως ο θόρυβος Gaussian — ενισχύει την ανθεκτικότητα του δικτύου απέναντι στις παρεμβολές και τα ακουστικά σφάλματα που είναι εγγενή στις υπερηχογραφικές λήψεις [11], [58]. Ακόμη και σε εικόνες κλίμακας του γκρι, τεχνικές όπως το Color Jittering προσαρμόζουν τη φωτεινότητα και την αντίθεση για να αυξήσουν τη διαφορετικότητα των δειγμάτων [1], [60]. Τέλος, αλγόριθμοι εξισορρόπησης ιστογράμματος όπως το CLAHE (Contrast-Limited Adaptive Histogram Equalization) χρησιμοποιούνται για τη βελτίωση της τοπικής αντίθεσης και την ανάδειξη των χαρακτηριστικών υφής του όγκου [1], [60].

Προηγμένες Τεχνικές Ενίσχυσης και Στρατηγικές

Στη βιβλιογραφία, η ενίσχυση δεδομένων επεκτείνεται σε πιο προχωρημένες στρατηγικές. Η συνθετική παραγωγή δεδομένων μέσω Παραγωγικών Ανταγωνιστικών Δικτύων (Generative Adversarial Network, GANs), όπως το cWGAN, επιτρέπει τη δημιουργία εντελώς νέων εικόνων που διατηρούν τα βασικά ακουστικά και μορφολογικά χαρακτηριστικά των καλοήθων ή κακοήθων όγκων [71]. Η προσέγγιση αυτή αποδεικνύεται ιδιαίτερα χρήσιμη για την αντιμετώπιση της ανισορροπίας κλάσεων (class imbalance), όπου μια κατηγορία υποεκπροσωπείται σε σχέση με τις υπόλοιπες [18]. Μεταξύ άλλων μεθόδων, ο μετασχηματισμός Wavelet παράγει εικόνες τριών καναλιών ενσωματώνοντας πληροφορίες διαφορετικών συχνοτικών ζωνών, προσφέροντας στο δίκτυο μια πλουσιότερη περιγραφή της βλάβης [32]. Διαφορετική λογική ακολουθεί η στρατηγική TrivialAugment: αντί για μια σταθερή ακολουθία μετασχηματισμών, επιλέγει τυχαία έναν

συνδυασμό από μια μεγάλη δεξαμενή επιλογών. Έρευνες έχουν δείξει ότι η δειγματοληψία από ένα ευρύ σύνολο ήπιων μετασχηματισμών αποδίδει περισσότερο από μια αυστηρά προκαθορισμένη ακολουθία [69].

Οι παραπάνω τεχνικές μπορούν να εφαρμοστούν είτε δυναμικά κατά τη διάρκεια της εκπαίδευσης (online / "on the fly"), είτε εκ των προτέρων, αυξάνοντας φυσικά τον όγκο του dataset στο δίσκο (offline) [29], [69]. Σύμφωνα με τα ερευνητικά δεδομένα, η στοχευμένη χρήση τεχνικών ενίσχυσης μπορεί να βελτιώσει τη συνολική ακρίβεια ταξινόμησης έως και 10%, ανάλογα με τον βαθμό δυσκολίας της εκάστοτε διαγνωστικής εργασίας [69].

4.3 Σχεδιασμός και Εκπαίδευση Μοντέλων

Η ενότητα αυτή περιγράφει την αρχιτεκτονική των CNNs που χρησιμοποιήθηκαν και αιτιολογεί την επιλογή μοντέλων χαμηλής υπολογιστικής πολυπλοκότητας — συγκεκριμένα των MobileNetV2, EfficientNetB0 και NASNetMobile — ως προς την αποδοτικότητά τους σε πόρους. Παράλληλα, εξειδικεύεται η διαδικασία TL, ο επανασχεδιασμός της κεφαλής ταξινόμησης (classifier head) και η ενσωμάτωση του μηχανισμού προσοχής CBAM ως πειραματική παραλλαγή.

4.3.1 Παραμετροποίηση Αρχιτεκτονικών

Ο σχεδιασμός και η εκπαίδευση μοντέλων DL για την ταξινόμηση του BUS βασίζονται κατά κύριο λόγο σε CNNs, ενώ πιο πρόσφατα έχουν εισαχθεί και τα Vision Transformers (ViTs) [1], [11]. Οι αρχιτεκτονικές που απαντώνται πιο συχνά στη βιβλιογραφία είναι τα EfficientNet, ResNet, VGG και MobileNet [5], [11]. Τα CNNs επιλέγονται για την ικανότητά τους να εξάγουν ιεραρχικά χωρικά χαρακτηριστικά — άκρα, υφές, μορφολογικά στοιχεία — τα οποία είναι κρίσιμα για τη διάκριση μεταξύ καλοήθων και κακοήθων όγκων [1]. Τα ViTs, από την πλευρά τους, αξιοποιούν μηχανισμούς αυτο-προσοχής (self-attention) για να μοντελοποιήσουν παγκόσμιες εξαρτήσεις στην εικόνα, κάτι που επιτρέπει την καλύτερη κατανόηση του ευρύτερου πλαισίου γύρω από τη βλάβη [1]. Υβριδικά μοντέλα συνδυάζουν και τα δύο: αξιοποιούν τα τοπικά χαρακτηριστικά των CNNs μαζί με το παγκόσμιο πλαίσιο των ViTs, ή τροφοδοτούν τα εξαχθέντα χαρακτηριστικά σε κλασικούς ταξινομητές όπως ο k-NN ή ο SVM [11].

Λόγω της έλλειψης μεγάλων, ιατρικά σχολιασμένων συνόλων δεδομένων, η TL αποτελεί στην πράξη αναπόφευκτη επιλογή [11]. Τα μοντέλα αρχικοποιούνται με προ-εκπαιδευμένα βάρη από το ImageNet [1], [11], και στη συνέχεια υποβάλλονται σε μικρο-συντονισμό για να προσαρμοστούν

στο πρόβλημα της ιατρικής απεικόνισης. Η διαδικασία ξεκινά με το πάγωμα των κατώτερων στρωμάτων (Layer Freezing): τα πρώτα επίπεδα, τα οποία έχουν ήδη μάθει να αναγνωρίζουν γενικά χαρακτηριστικά όπως γραμμές και γεωμετρικά σχήματα, παραμένουν σταθερά [1], [11]. Τα ανώτερα στρώματα, αντίθετα, αποδεσμεύονται και επανεκπαιδεύονται ώστε να μάθουν τα εξειδικευμένα χαρακτηριστικά των υπερήχων μαστού — ηχογένεια, ακουστικές σκιές, μορφολογία των ορίων [1]. Κατά τον μικρο-συντονισμό, ο ρυθμός μάθησης διατηρείται σκοπίμως χαμηλός (π.χ. 0,0001 ή κάτω), έτσι ώστε οι αναπαραστάσεις που έχουν ήδη αποκτηθεί να μην διαταραχθούν [11], [68].

Αναπόσπαστο τμήμα της προσαρμογής είναι ο επανασχεδιασμός της κεφαλής ταξινόμησης. Τα αρχικά επίπεδα εξόδου του ImageNet, που ταξινομούν 1.000 κλάσεις, αφαιρούνται και αντικαθίστανται από μια νέα κεφαλή προσαρμοσμένη στο τρέχον πρόβλημα [4], [11]. Η νέα κεφαλή ξεκινά με ένα στρώμα Global Average Pooling (GAP), το οποίο συμπιέζει τους χάρτες χαρακτηριστικών σε ένα μονοδιάστατο διάνυσμα, περιορίζοντας τον κίνδυνο υπερεκπαίδευσης [11]. Ακολουθούν Πλήρως Συνδεδεμένα Στρώματα (Dense Layers) — για παράδειγμα, ένα ενδιάμεσο επίπεδο 128 μονάδων με συνάρτηση ενεργοποίησης ReLU — που εισάγουν τη μη-γραμμικότητα που χρειάζεται το μοντέλο [4]. Στρώματα Dropout με ρυθμό 0,3–0,5 παρεμβάλλονται για πρόσθετη προστασία από την υπερπροσαρμογή [4]. Για τη δυαδική ταξινόμηση καλοήθειας/κακοήθειας, το τελικό επίπεδο αποτελείται από έναν μοναδικό νευρώνα με συνάρτηση Sigmoid — σε αντίθεση με τη Softmax, που χρησιμοποιείται σε πολυταξικά προβλήματα [4], [72].

Η σύγκλιση των μοντέλων εξασφαλίζεται μέσω συγκεκριμένων στρατηγικών εκπαίδευσης. Ο Adam (ή η παραλλαγή AdamW) είναι ο πλέον διαδεδομένος βελτιστοποιητής λόγω της υπολογιστικής του σταθερότητας, αν και ο SGD με ορμή (momentum) χρησιμοποιείται επίσης για ομαλότερη σύγκλιση [4], [18]. Ως συνάρτηση κόστους ορίζεται η Δυαδική Διασταυρούμενη Εντροπία (Binary Cross-Entropy), η οποία αντικαθιστά την Categorical Cross-Entropy που αρμόζει σε προβλήματα πολλαπλών κλάσεων [4]. Όταν η ανισορροπία μεταξύ κλάσεων είναι έντονη, μπορούν να εφαρμοστούν σταθμισμένες συναρτήσεις απώλειας (weighted cross-entropy) ή η Focal Loss [5], [69]. Για την αντιμετώπιση της υπερπροσαρμογής, η Πρόωρη Διακοπή (Early Stopping) τερματίζει την εκπαίδευση μόλις η απόδοση στο σύνολο επικύρωσης παύει να βελτιώνεται [4], [10]. Παράλληλα, ο μηχανισμός ReduceLROnPlateau μειώνει αυτόματα τον ρυθμό μάθησης όταν η απώλεια σταθεροποιείται, επιτρέποντας στο μοντέλο να συγκλίνει με μεγαλύτερη ακρίβεια στο τοπικό ελάχιστο [4], [51].

4.3.2 Ανάπτυξη του Ενισχυμένου Μοντέλου με Μηχανισμό Προσοχής

Η ανάπτυξη ενός μοντέλου ενισχυμένου με μηχανισμό προσοχής (Attention-Enhanced Model) βασίζεται στη συνδυαστική αξιοποίηση της αρχιτεκτονικής MobileNetV2 και της μονάδας CBAM. Ο στόχος είναι διπλός: ένα σύστημα ελαφρύ σε υπολογιστικούς όρους, με ταυτόχρονα υψηλή διαγνωστική ακρίβεια στον εντοπισμό καρκίνου του μαστού [4], [63]. Η υβριδική αυτή προσέγγιση δίνει στο δίκτυο τη δυνατότητα να εντοπίζει δυναμικά τα πιο διακριτικά χαρακτηριστικά στις υπερηχογραφικές εικόνες, βελτιώνοντας έτσι τη διακριτική του ικανότητα [8], [63].

Το MobileNetV2 ως Δίκτυο Κορμού

Για την εξαγωγή πρωτογενών χαρακτηριστικών επιλέχθηκε το MobileNetV2, λόγω της αποδεδειγμένης υπολογιστικής του αποδοτικότητας σε περιβάλλοντα με περιορισμένους πόρους [4], [60]. Η αρχιτεκτονική αυτή στηρίζεται σε δύο θεμελιώδεις σχεδιαστικές επιλογές.

Η πρώτη είναι οι Διαχωρίσιμες Συνελίξεις κατά Βάθος (Depthwise Separable Convolutions). Η τεχνική αυτή διαχωρίζει το χωρικό φιλτράρισμα από τον συνδυασμό καναλιών, μειώνοντας δραστικά τον αριθμό εκπαιδευσίμων παραμέτρων και το υπολογιστικό κόστος, χωρίς να θυσιάζει τη διαγνωστική ακρίβεια [18], [60]. Η δεύτερη επιλογή αφορά τα Ανεστραμμένα Υπολειμματικά Μπλοκ με Γραμμικούς Στενωπούς (Inverted Residuals & Linear Bottlenecks). Εδώ τα δεδομένα επεκτείνονται προσωρινά σε χώρους υψηλότερης διαστατικότητας πριν συμπιεστούν, διατηρώντας τις απαραίτητες μορφολογικές πληροφορίες σε συμπαγή μορφή [10], [18].

Ενσωμάτωση της Μονάδας Προσοχής CBAM

Για να εστιάζει το δίκτυο στις κλινικά σχετικές περιοχές, ενσωματώθηκε η μονάδα CBAM — ένας ελαφρύς μηχανισμός προσοχής που ενσωματώνεται απρόσκοπτα σε οποιαδήποτε αρχιτεκτονική CNN, συμπεριλαμβανομένου του MobileNetV2 [63]. Η επεξεργασία των χαρτών χαρακτηριστικών γίνεται σε δύο διαδοχικά στάδια.

Στο πρώτο, η Μονάδα Προσοχής Καναλιού (Channel Attention Module) απαντά στο ερώτημα του «τι» είναι σημαντικό στην εικόνα [59]. Εφαρμόζει παράλληλα Global Average Pooling και Max Pooling, συγκεντρώνοντας τις χωρικές πληροφορίες ανά κανάλι. Το αποτέλεσμα τροφοδοτείται σε ένα πολυστρωματικό νευρωνικό δίκτυο (Multi-Layer Perceptron, MLP), το οποίο παράγει ένα διάλυμα βαρών που υποδεικνύει ποια κανάλια φέρουν τη μεγαλύτερη διαγνωστική πληροφορία [59].

Στο δεύτερο στάδιο, η Μονάδα Χωρικής Προσοχής (Spatial Attention Module) απαντά στο «πού» βρίσκεται η βλάβη. Εφαρμόζει pooling κατά τον άξονα των καναλιών και, μέσω συνελκτικού στρώματος 7×7 , παράγει έναν δισδιάστατο χωρικό χάρτη που αναδεικνύει τις περιοχές ενδιαφέροντος — όπως ακανόνιστα όρια όγκου — ενώ παράλληλα καταστέλλει τον θόρυβο υποβάθρου [59].

Κλινικά και Υπολογιστικά Πλεονεκτήματα του Υβριδικού Μοντέλου

Ο συνδυασμός του MobileNetV2 με τον μηχανισμό προσοχής CBAM αποδίδει πρακτικά οφέλη σε τρία βασικά επίπεδα.

Το πρώτο επίπεδο αφορά τον εντοπισμό και τον διαχωρισμό των χαρακτηριστικών των βλαβών. Ο μηχανισμός χωρικής προσοχής και προσοχής καναλιού επιτρέπει στο μοντέλο να εστιάζει στις πλέον διακριτικές μορφολογικές λεπτομέρειες της μάζας του όγκου, αγνοώντας άσχετα στοιχεία (όπως ιατρικές ενδείξεις οθόνης ή ακουστικό θόρυβο) [59]. Όπως καταδείχθηκε, ο μηχανισμός αυτός ενισχύει τη διακριτική ικανότητα του μοντέλου ειδικά για την αναγνώριση κακοηθών βλαβών. Αυτή η βελτίωση είναι κλινικά υψίστης σημασίας, διότι η παράλειψη διάγνωσης ενός κακοήθους όγκου (false negatives) μπορεί να καθυστερήσει τη διάγνωση και να επηρεάσει αρνητικά την έκβαση για την ασθενή.

Το δεύτερο επίπεδο αφορά την υπολογιστική αποδοτικότητα. Παρά την ενσωμάτωση των προσαρμοστικών χαρτών προσοχής, το μοντέλο παραμένει ιδιαίτερα ελαφρύ (με μέγεθος συχνά κάτω των 10MB). Ο μηχανισμός CBAM επιτυγχάνει την προαναφερθείσα βελτίωση διατηρώντας παράλληλα την υπολογιστική αποδοτικότητα που απαιτείται για ελαφριά διαγνωστικά πλαίσια, γεγονός που το καθιστά ιδανικό για ενσωμάτωση σε φορητές συσκευές POCUS και εφαρμογές διάγνωσης σε πραγματικό χρόνο [4].

Το τρίτο πλεονέκτημα είναι η ερμηνευσιμότητα. Η χρήση προσοχής, σε συνδυασμό με τεχνικές όπως το Grad-CAM, καθιστά διαφανή τη διαδικασία λήψης αποφάσεων: οι ακτινολόγοι μπορούν να οπτικοποιούν ακριβώς ποιες περιοχές της εικόνας οδήγησαν το δίκτυο σε μια συγκεκριμένη διάγνωση [4]. Σε επίπεδο τελικής κλινικής απόδοσης, η έρευνα αποδεικνύει ότι τέτοια έξυπνα, ενισχυμένα με προσοχή μοντέλα μπορούν, με τη σωστή προ-επεξεργασία δεδομένων, να ξεπεράσουν βαρύτερες αρχιτεκτονικές —όπως το ResNet101 ή το VGG16— επιτυγχάνοντας ποσοστά ακρίβειας που προσεγγίζουν το απόλυτο σε εξειδικευμένα σύνολα δεδομένων υπερήχων, όπως το BUSI [4].

4.4 Στρατηγική Αξιολόγησης & Βελτιστοποίησης

Η ενότητα αυτή περιγράφει το μεθοδολογικό πλαίσιο εκπαίδευσης και επικύρωσης του συστήματος. Αναλύεται η τεχνική της Στρωματοποιημένης Ομαδικής Διασταυρούμενης Επικύρωσης 5-Πτυχών (Stratified Group 5-Fold Cross-Validation, 5-Fold CV), η οποία αποτρέπει τη διαρροή δεδομένων, καθώς και δύο μέθοδοι βελτιστοποίησης κατά την πρόβλεψη: το TTA και η δυναμική ρύθμιση του κατωφλίου απόφασης.

4.4.1 Stratified Group 5-Fold Cross-Validation

Η 5-Fold CV θεωρείται ευρέως το «χρυσό πρότυπο» (gold standard) για την αξιολόγηση μοντέλων DL σε ιατρικά σύνολα δεδομένων, όπως τα BUSI και BUS-BRA. Η επιλογή της δεν είναι τυχαία. Διασφαλίζει την αντικειμενικότητα των αποτελεσμάτων και ελέγχει αν το μοντέλο γενικεύει πράγματι σε άγνωστα δεδομένα, αντί να τα απομνημονεύει [3], [6].

Η τεχνική συνδυάζει τρία μεθοδολογικά χαρακτηριστικά. Σύμφωνα με την αρχή 5-Fold, τα δεδομένα διαιρούνται σε πέντε ίσα μέρη (folds) [70]. Εκτελούνται πέντε διαδοχικοί κύκλοι εκπαίδευσης: σε κάθε κύκλο, τα τέσσερα μέρη αποτελούν το σύνολο εκπαίδευσης και το πέμπτο χρησιμεύει ως σύνολο ελέγχου. Έτσι, κάθε εικόνα του συνόλου χρησιμοποιείται για έλεγχο ακριβώς μία φορά [4], [6]. Ο διαχωρισμός είναι επίσης στρωματοποιημένος, διατηρώντας την αρχική αναλογία κλάσεων —καλοήθων, κακοήθων, φυσιολογικών— σε κάθε fold [1], [4]. Αυτό μετράει ιδιαίτερα στο BUSI, όπου η ανισορροπία μεταξύ κλάσεων είναι εμφανής και θα μπορούσε να εισάγει συστηματική μεροληψία (bias) στην αξιολόγηση [1]. Το τρίτο χαρακτηριστικό —και ίσως το πιο ουσιαστικό— είναι ο ομαδικός (Group) διαχωρισμός: η κατανομή στα folds γίνεται ανά ασθενή, όχι ανά εικόνα [6], [17].

Η ανάγκη για διαχωρισμό σε επίπεδο ασθενούς προκύπτει από τον κίνδυνο Διαρροής Δεδομένων (Data Leakage). Πρόκειται για το φαινόμενο κατά το οποίο πληροφορίες από το σύνολο ελέγχου εισχωρούν στη διαδικασία εκπαίδευσης, οδηγώντας σε τεχνητά υψηλή ακρίβεια που δεν ανταποκρίνεται σε πραγματικές συνθήκες [6]. Στα σύνολα υπερηχογραφικών δεδομένων το πρόβλημα είναι ιδιαίτερα οξύ, γιατί για κάθε ασθενή υπάρχουν συχνά πολλαπλές εικόνες. Στο BUS-BRA, για παράδειγμα, το 76% των ασθενών παρέχουν δύο ορθογώνιες όψεις της ίδιας βλάβης (αριστερή και δεξιά λήψη), ενώ σε άλλα σύνολα δεδομένων υπάρχουν πολλαπλές εγκάρσιες τομές από διαφορετικά σημεία του ίδιου όγκου [6]. Με έναν τυχαίο διαχωρισμό ανά εικόνα, η μία όψη

ενός όγκου θα μπορούσε να καταλήξει στο training set και η άλλη στο test set. Το νευρωνικό δίκτυο δεν θα μάθαινε τα γενικά χαρακτηριστικά της βλάβης· θα αναγνώριζε τον συγκεκριμένο ασθενή — τον θόρυβο speckle του μηχανήματος, την ιδιαίτερη ανατομία του ιστού— και ουσιαστικά θα εκτελούσε απομνημόνευση (overfitting) αντί για μάθηση [6].

Στην πράξη, και για τα δύο σύνολα δεδομένων της παρούσας έρευνας, όλες οι εικόνες του ίδιου κωδικού ασθενούς (Patient ID) μεταφέρθηκαν πάντα μαζί στο ίδιο fold [6], [17]. Το BUS-BRA διαθέτει ειδικά αρχεία CSV με προκαθορισμένους δείκτες για 5-fold και 10-fold cross-validation, κάτι που επιτρέπει την άμεση σύγκριση μοντέλων υπό τυποποιημένες συνθήκες [6], [17]. Επιπλέον, εντός κάθε fold εκπαίδευσης, ένα 20% των ασθενών παραχωρήθηκε ως σύνολο επικύρωσης (hold-out validation set) για τη ρύθμιση υπερπαραμέτρων και την εφαρμογή πρόωρης διακοπής (early stopping) [6].

Με αυτόν τον σχεδιασμό, αποκλείεται το ενδεχόμενο το μοντέλο να «αναγνωρίσει» στον έλεγχο έναν όγκο που έχει δει —έστω υπό διαφορετική γωνία— στην εκπαίδευση. Αυτό είναι η ουσία του Group CV και αυτό καθιστά τα εξαγόμενα αποτελέσματα κλινικά αξιόπιστα.

4.4.2 Ενίσχυση Δεδομένων κατά την Πρόβλεψη

Το ΤΤΑ είναι μια τεχνική που εφαρμόζεται αποκλειστικά κατά τη φάση της εξαγωγής συμπερασμάτων (inference), όχι κατά την εκπαίδευση. Στόχος της είναι η βελτίωση της διαγνωστικής ακρίβειας και της αξιοπιστίας μοντέλων DL στην ανάλυση ιατρικών εικόνων [18], [71].

Σε μια τυπική αξιολόγηση, το μοντέλο λαμβάνει μία εικόνα και παράγει μία πρόβλεψη. Το ΤΤΑ διαφέρει ως προς αυτό. Όταν εισάγεται μια νέα εικόνα δοκιμής, το σύστημα παράγει πολλαπλές παραλλαγές της μέσω ήπιων γεωμετρικών μετασχηματισμών — τυχαίες περιστροφές, οριζόντιες αναστροφές (flips) ή ελαφριές περικοπές (crops) [71]. Κάθε παραλλαγή αξιολογείται ανεξάρτητα από το μοντέλο [55], και η τελική απόφαση προκύπτει από τη συγκέντρωση αυτών των επιμέρους προβλέψεων, είτε ως μέσος όρος πιθανοτήτων (predictive mean) είτε μέσω πλειοψηφικής ψηφοφορίας (majority voting) [55].

Τα πλεονεκτήματα της μεθόδου είναι συγκεκριμένα και μετρήσιμα. Από βιβλιογραφικά δεδομένα προκύπτει ότι ο συνδυασμός προ-εκπαιδευμένων δικτύων με ΤΤΑ μπορεί να ανεβάσει την απόδοση σε εικόνες μαστογραφίας και υπερήχων σε ποσοστά ακρίβειας που ξεπερνούν το 99%

[71]. Εξίσου σημαντική είναι η αυξημένη στιβαρότητα (robustness) που αποκτά το μοντέλο: μικρές οπτικές παραλλαγές που οφείλονται στη γωνία του ηχοβολέα κατά την υπερηχογραφική εξέταση δεν επηρεάζουν πλέον ουσιαστικά την έκβαση της ταξινόμησης. Η λογική είναι ανάλογη με αυτή ενός έμπειρου ακτινολόγου που εξετάζει μια ύποπτη βλάβη από διαφορετικές οπτικές γωνίες πριν καταλήξει σε διάγνωση.

Ιδιαίτερη αξία για τις ιατρικές εφαρμογές έχει η δυνατότητα εκτίμησης της αβεβαιότητας του μοντέλου. Το TTA παρέχει πληροφορίες για το επίπεδο εμπιστοσύνης (confidence level) κάθε πρόβλεψης [55]: οι ορθές ταξινομήσεις τείνουν να συνοδεύονται από υψηλή συμφωνία μεταξύ των παραλλαγών, ενώ οι λανθασμένες χαρακτηρίζονται από διασπορά που μπορεί να φτάσει έως και 40 φορές μεγαλύτερη σε σχέση με τις σωστές [55].

Στην ανάλυση εικόνων μαστού, η τεχνική χρησιμοποιείται ήδη σε συστήματα τμηματοποίησης (segmentation) με αρχιτεκτονικές τύπου U-Net, αλλά και στην ταξινόμηση ανωμαλιών σε μαστογραφικές λήψεις [18]. Ο αυξημένος υπολογιστικός χρόνος κατά το inference αποτελεί πρακτικό μειονέκτημα που δεν μπορεί να αγνοηθεί. Ωστόσο, η ικανότητα της μεθόδου να αποδίδει αξιόπιστα ακόμη και με περιορισμένα δεδομένα την καθιστά ιδιαίτερα κατάλληλη για κλινικά περιβάλλοντα, όπου η διαθεσιμότητα μεγάλων ιατρικών συνόλων δεδομένων δεν μπορεί να θεωρηθεί δεδομένη [71].

4.4.3 Δυναμική Βελτιστοποίηση Κατωφλίου

Η βελτιστοποίηση κατωφλίου (Optimal Threshold Tuning) αποτελεί ένα από τα πιο ουσιαστικά βήματα στη μετα-επεξεργασία των αποτελεσμάτων ενός μοντέλου DL. Η διαδικασία αυτή μετατρέπει τις συνεχείς πιθανότητες που παράγει το νευρωνικό δίκτυο στο διάστημα (0, 1) σε διακριτές, κλινικά αξιοποιήσιμες διαγνωστικές κατηγορίες [29], [72]. Η επιλογή του κατωφλίου απόφασης (decision threshold) ρυθμίζει άμεσα την ισορροπία μεταξύ ευαισθησίας (sensitivity) και ειδικότητας (specificity), με άμεσες επιπτώσεις στην κλινική αξιοπιστία του συστήματος [6], [68].

Στις τυπικές εφαρμογές μηχανικής μάθησης, το κατώφλι ορίζεται συνήθως στο 0,5 ως προεπιλεγμένη τιμή. Αυτή η επιλογή, ωστόσο, αποδεικνύεται συχνά ανεπαρκής για τις ιδιαίτερες απαιτήσεις της ιατρικής πρακτικής [3]. Στην ανάλυση υπερηχογραφήματων μαστού, οι ακτινολόγοι δίνουν προτεραιότητα στην υψηλή ευαισθησία: ο στόχος είναι να μην διαφύγει καμία περίπτωση κακοήθειας, ακόμα και αν αυτό συνεπάγεται αύξηση των ψευδώς θετικών και, κατ' επέκταση, ορισμένες περιττές βιοψίες [6]. Η βελτιστοποίηση του κατωφλίου επιτρέπει στο μοντέλο να

λειτουργεί με τρόπο που προσεγγίζει τη διαγνωστική συμπεριφορά έμπειρου ιατρού, βελτιώνοντας την ειδικότητα χωρίς να θυσιάζεται η ευαισθησία [6].

Για τον καθορισμό του βέλτιστου κατωφλίου, η βιβλιογραφία καταγράφει αρκετές διαφορετικές προσεγγίσεις. Σε ορισμένα συστήματα χρησιμοποιούνται σταθερά κατώφλια, επιλεγμένα εμπειρικά μέσω δοκιμής και σφάλματος, ιδίως σε πολυκλασικά προβλήματα όπως η ταξινόμηση κατά BI-RADS [3]. Πιο προηγμένες μέθοδοι αξιοποιούν αυτο-προσαρμοζόμενα κατώφλια (auto-adaptive thresholding) με στατιστικά εργαλεία όπως η Διάμεση Απόλυτη Απόκλιση (Median Absolute Deviation, MAD), η οποία προσφέρει ανθεκτικότητα στον οπτικό θόρυβο και στις διακυμάνσεις έντασης εικονοστοιχείων [67]. Ξεχωριστή περίπτωση αποτελούν τα προβλήματα τμηματοποίησης (segmentation) όγκων, όπου η επιλογή τιμής κατωφλίου για μετρικές όπως ο συντελεστής Dice (DSC) επηρεάζει άμεσα την ορθή διαχείριση μη ισορροπημένων κλάσεων — ο αριθμός εικονοστοιχείων που αντιστοιχούν στον όγκο υπολείπεται σημαντικά εκείνων του περιβάλλοντος ιστού [3].

Το κυριότερο αναλυτικό εργαλείο για την αξιολόγηση της επίδρασης του κατωφλίου είναι η καμπύλη ROC. Απεικονίζει τη σχέση μεταξύ ποσοστού αληθώς θετικών (True Positive Rate, TPR) και ψευδώς θετικών (FPR) για κάθε πιθανή τιμή κατωφλίου [25], [68]. Το εμβαδόν κάτω από την καμπύλη (AUC) παρέχει συνολική εικόνα της διαχωριστικής ικανότητας του μοντέλου ανεξάρτητα από το επιλεγμένο κατώφλι [1], [6]. Για τον ακριβή εντοπισμό του βέλτιστου σημείου λειτουργίας χρησιμοποιείται ο δείκτης Youden (J), ο οποίος μεγιστοποιεί το άθροισμα ευαισθησίας και ειδικότητας επί της καμπύλης [73].

Σε πολυκλασικά προβλήματα, όπως η ταξινόμηση BI-RADS, η ρύθμιση κατωφλίου γίνεται περισσότερο απαιτητική [29]. Η διάκριση μεταξύ γειτονικών κατηγοριών — για παράδειγμα BI-RADS 3 και 4 — είναι ιδιαίτερα δύσκολη λόγω της έντονης επικάλυψης οπτικών και μορφολογικών χαρακτηριστικών [29]. Μια διαδεδομένη στρατηγική αντιμετώπισης είναι η μέθοδος One-Versus-All (OVA): εκπαιδεύονται εξειδικευμένοι ταξινομητές για κάθε κλάση και η τελική απόφαση λαμβάνεται βάσει μέγιστης εμπιστοσύνης (maximum confidence), ουσιαστικά δημιουργώντας ένα ξεχωριστό κατώφλι για κάθε κατηγορία [29].

Με αυτόν τον τρόπο, η πιθανολογική έξοδος του μοντέλου μετατρέπεται σε ένα εργαλείο υποστήριξης κλινικών αποφάσεων που μπορεί να ρυθμιστεί ανάλογα με τον εκάστοτε κλινικό στόχο — είτε αυτός αφορά ευρύ προληπτικό έλεγχο (screening) είτε αυστηρότερη επιβεβαιωτική διάγνωση [6].

4.5 Σύνοψη κεφαλαίου

Στο παρόν κεφάλαιο παρουσιάστηκε το πλήρες μεθοδολογικό και προγραμματιστικό πλαίσιο της έρευνας. Η ανάλυση κάλυψε την οργάνωση των δεδομένων (BUSI, BUS-BRA) και την προεπεξεργασία "Input E", τον αρχιτεκτονικό σχεδιασμό των lightweight μοντέλων και την ενσωμάτωση του μηχανισμού CBAM. Ιδιαίτερη βαρύτητα δόθηκε στις στρατηγικές αξιολόγησης (Stratified Group 5-Fold CV, TTA, Optimal Threshold Tuning), οι οποίες προστατεύουν το σύστημα από διαρροή δεδομένων και τεκμηριώνουν την κλινική του αξιοπιστία.

Το Κεφάλαιο 5 που ακολουθεί παρουσιάζει το υπολογιστικό περιβάλλον υλοποίησης και τα πειραματικά αποτελέσματα. Μέσα από την αξιολόγηση τόσο των baseline όσο και των ενισχυμένων αρχιτεκτονικών, καταλήγει σε συγκριτική ανάλυση που αναδεικνύει τεκμηριωμένα το τελικό μοντέλο της εργασίας.

Κεφάλαιο 5: Πειραματικά Αποτελέσματα και Συγκριτική Ανάλυση

Το παρόν κεφάλαιο παρουσιάζει, αξιολογεί και συγκρίνει τα πειραματικά αποτελέσματα από την εφαρμογή της προτεινόμενης μεθοδολογίας. Αρχικά περιγράφεται το υπολογιστικό περιβάλλον και οι υποδομές υλικού που χρησιμοποιήθηκαν (Ενότητα 5.1). Ακολουθεί η παρουσίαση των επιδόσεων των τριών baseline αρχιτεκτονικών χαμηλής πολυπλοκότητας — MobileNetV2, EfficientNetB0 και NASNetMobile — στα σύνολα δεδομένων BUSI και BUS-BRA (Ενότητα 5.2). Στη συνέχεια αξιολογείται στατιστικά η επίδραση του μηχανισμού προσοχής CBAM στη διαγνωστική ικανότητα του επικρατέστερου μοντέλου (Ενότητα 5.3). Το κεφάλαιο κλείνει με συγκριτική ανάλυση του συνόλου των πειραμάτων, ώστε να τεκμηριωθεί η βέλτιστη αρχιτεκτονική για την ταξινόμηση υπερηχογραφημάτων μαστού (Ενότητα 5.4).

5.1 Περιβάλλον Υλοποίησης: Εργαλεία και Hardware

Η εκπαίδευση και αξιολόγηση μοντέλων Βαθιάς Μάθησης (Deep Learning, DL) απαιτεί εξειδικευμένους επιταχυντές υλικού — κυρίως GPUs — και επαρκή RAM. Οι απαιτήσεις αυτές είναι ιδιαίτερα αυξημένες όταν η μεθοδολογία περιλαμβάνει διαδικασίες όπως Stratified 5-Fold Cross-Validation (5-Fold CV), Test-Time Augmentation (TTA) και Ensemble Learning. Για τον λόγο αυτό, επιλέχθηκε κατανεμημένη προσέγγιση μέσω υποδομών νέφους της Google.

1. Προγραμματιστικό Περιβάλλον και Εργαλεία Ανάπτυξης

Ο κώδικας αναπτύχθηκε σε Python με τη μορφή Jupyter Notebooks. Για την υλοποίηση των αρχιτεκτονικών και των μηχανισμών προσοχής χρησιμοποιήθηκε το TensorFlow και το Keras API. Η προεπεξεργασία εικόνων — μέθοδος Input E, Masking, Dilation — έγινε μέσω OpenCV. Η διαχείριση δεδομένων και η εξαγωγή στατιστικών μετρικών βασίστηκαν στις βιβλιοθήκες Pandas,

NumPy και Scikit-learn, ενώ η οπτικοποίηση των αποτελεσμάτων (Confusion Matrices, ROC Curves) υλοποιήθηκε με Matplotlib και Seaborn.

2. Google Colab (Colaboratory)

Στα αρχικά στάδια αξιοποιήθηκε το Google Colab, το οποίο παρέχει φιλοξενούμενο περιβάλλον Jupyter Notebook με δωρεάν πρόσβαση σε NVIDIA T4 GPU. Το περιβάλλον αυτό ήταν κατάλληλο για prototyping, debugging, δημιουργία προεπεξεργασμένων συνόλων δεδομένων και αρχικές δοκιμές αρχιτεκτονικής. Δεν κάλυπτε όμως τις ανάγκες της πλήρους μεθοδολογίας. Ο μεγάλος όγκος ιατρικών εικόνων και η ταυτόχρονη διατήρηση πολλαπλών μοντέλων στη μνήμη κατά την 5-Fold CV ανέδειξαν σαφείς περιορισμούς: περιορισμένη διαθέσιμη RAM και αυτόματες διακοπές συνεδριών.

3. Google Cloud Vertex AI Workbench

Για την υπέρβαση των παραπάνω τεχνικών περιορισμών και την απρόσκοπτη ολοκλήρωση της πειραματικής διαδικασίας, η εκπαίδευση μεταφέρθηκε στο περιβάλλον Google Cloud Vertex AI Workbench. Πρόκειται για μια πλήρως διαχειριζόμενη πλατφόρμα ML επιχειρησιακού επιπέδου, που παρέχει άμεσο και πλήρη έλεγχο των υπολογιστικών πόρων.

Για τις ανάγκες της εκπαίδευσης δημιουργήθηκε εικονική μηχανή τύπου g2-standard-4 (Graphics Optimized), εξοπλισμένη με έναν επιταχυντή NVIDIA L4 GPU, 4 εικονικούς επεξεργαστές (vCPUs) και 16 GB μνήμης RAM. Η μετάβαση σε αυτό το περιβάλλον είχε τρεις συγκεκριμένες επιπτώσεις στη ροή της εργασίας.

Η πρώτη αφορά την υπολογιστική επάρκεια. Η αρχιτεκτονική της NVIDIA L4 σε συνδυασμό με τους 4 vCPUs εξασφάλισε αδιάλειπτη λειτουργία για αρκετές ώρες συνεχούς εκπαίδευσης — προϋπόθεση απαραίτητη για την ολοκλήρωση των 60 εποχών (epochs) σε καθένα από τα 5 folds χωρίς κίνδυνο διακοπής. Η δεύτερη αφορά τη διαχείριση μνήμης. Τα 16 GB RAM, σε συνδυασμό με τη VRAM της GPU, παρείχαν επαρκή χωρητικότητα τόσο για τη συνένωση (Ensemble) των πέντε εκπαιδευμένων μοντέλων όσο και για την εκτέλεση της μνημοβόρας τεχνικής TTA. Η τρίτη αφορά την ταχύτητα πρόσβασης στα δεδομένα. Η αποθήκευση των συνόλων δεδομένων BUSI και BUS-BRA απευθείας στον τοπικό δίσκο της εικονικής μηχανής εξάλειψε τα bottlenecks I/O που προκαλούσε η συνεχής επικοινωνία με το Google Drive.

Κατά την εκπόνηση της εργασίας, το Google Colab χρησιμοποιήθηκε για την ανάπτυξη και επαλήθευση της μεθοδολογίας, ενώ το Vertex AI Workbench ανέλαβε την εκτέλεση του υπολογιστικά βαρέος πειράματος. Αυτός ο διαχωρισμός ρόλων επέτρεψε την παραγωγή των πειραματικών αποτελεσμάτων που αναλύονται στις επόμενες ενότητες με αξιοπιστία και επαναληψιμότητα.

5.2 Αποτελέσματα Baseline Μοντέλων

Η ενότητα αυτή εξετάζει τις διαγνωστικές επιδόσεις των τριών αρχιτεκτονικών DL χαμηλής πολυπλοκότητας που αποτελούν αντικείμενο σύγκρισης. Κάθε μοντέλο αξιολογείται ανεξάρτητα στα δύο σύνολα δεδομένων — BUSI και BUS-BRA — με κοινή στρατηγική προεπεξεργασίας ("Input E"), 5-Fold CV και TTA. Οι επιδόσεις μετρώνται μέσω AUC και Ακρίβειας (Accuracy) υπό το βέλτιστο κατώφλι απόφασης, ενώ η ταξινομητική συμπεριφορά ανά κλάση αποτυπώνεται στους αντίστοιχους Πίνακες Σύγχυσης.

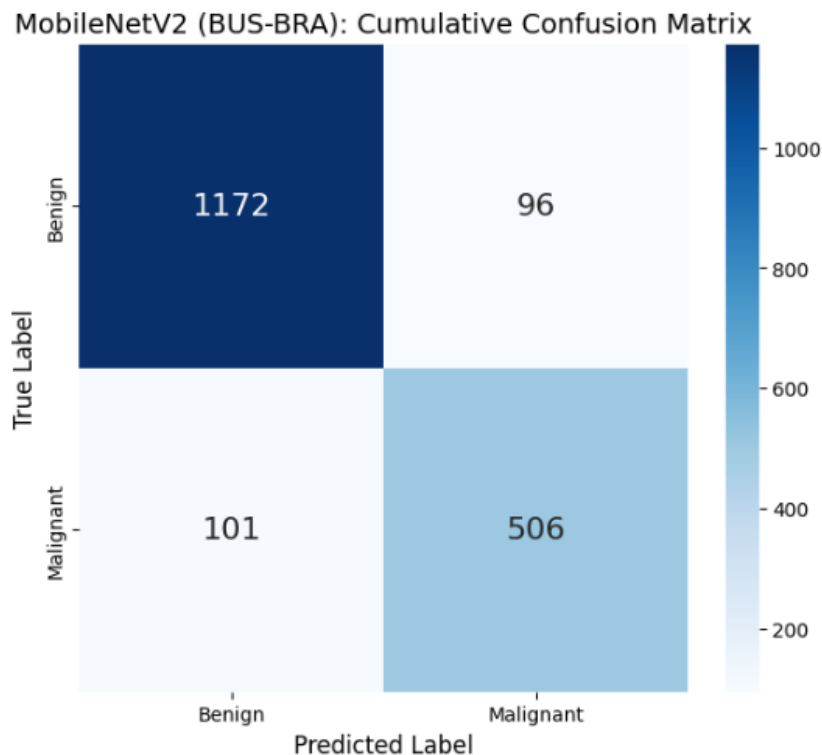
5.2.1 Επιδόσεις MobileNetV2

Το MobileNetV2 χρησιμεύει ως η αρχιτεκτονική αναφοράς (baseline) για τη συγκεκριμένη κατηγορία μοντέλων. Εκπαιδεύτηκε και αξιολογήθηκε μέσω 5-Fold CV, με προεπεξεργασία "Input E" και TTA. Τα αποτελέσματα παρουσιάζονται ξεχωριστά για τα δύο σύνολα δεδομένων.

A. Αξιολόγηση στο σύνολο δεδομένων BUS-BRA

Metric / Class	Value / Precision			
Optimal Threshold	0,5000			
AUC	0,9387			
Accuracy	0,8949			
Class	Precision	Recall	F1-Score	Support
Benign	0,9207	0,9243	0,9225	1268,0
Malignant	0,8405	0,8336	0,8371	607,0
Macro Avg	0,8806	0,8789	0,8798	187,0
Weighted Avg	0,8947	0,8949	0,8948	1875,0

ΠΙΝΑΚΑΣ 2: ΣΥΓΚΕΝΤΡΩΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΟΥ ΜΟΝΤΕΛΟΥ MOBILENETV2 ΣΤΟ ΣΥΝΟΛΟ ΔΕΔΟΜΕΝΩΝ BUS-BRA (1875 ΕΙΚΟΝΕΣ)



ΠΙΝΑΚΑΣ 3: ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ ΤΟΥ ΜΟΝΤΕΛΟΥ MOBILENETV2 ΣΤΟ ΣΥΝΟΛΟ ΔΕΔΟΜΕΝΩΝ BUS-BRA

Σχολιασμός Αποτελεσμάτων (BUS-BRA):

Στο σύνολο BUS-BRA, το μοντέλο παρουσίασε ισχυρή διαχωριστική ικανότητα. Το βέλτιστο κατώφλι απόφασης ορίστηκε στο 0,5000 μέσω δυναμικής ρύθμισης, με AUC 0,9387 και συνολική Ακρίβεια 89,49%.

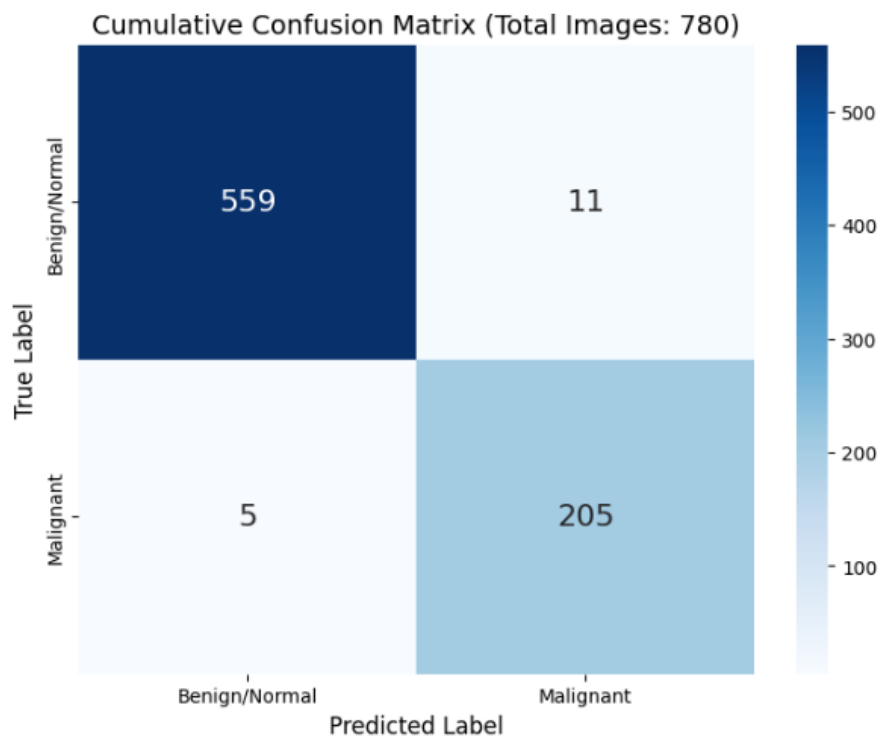
Στην κλάση των καλοήθων αλλοιώσεων, το δίκτυο κατέγραψε Recall 92,43% και Precision 92,07% — υψηλές τιμές που υποδηλώνουν συνεπή αναγνώριση της κλάσης αυτής. Η απόδοση στις κακοήθεις αλλοιώσεις ήταν χαμηλότερη αλλά παρέμεινε σε ικανοποιητικά επίπεδα: Recall 83,36%, Precision 84,05% και F1-Score 0,8371. Η διαφορά αυτή στις επιμέρους κλάσεις είναι αναμενόμενη, δεδομένης της φύσης της κατανομής του συνόλου δεδομένων.

Από τον Πίνακα Σύγχυσης προκύπτει ότι το μοντέλο ταξινόμησε ορθά 1.172 καλοήθεις (TN) και 506 κακοήθεις (TP) περιπτώσεις. Οι ψευδώς θετικές (FP) ήταν 96 και οι ψευδώς αρνητικές (FN) 101. Τα στοιχεία αυτά δείχνουν ότι μια αρχιτεκτονική χαμηλής πολυπλοκότητας μπορεί να αποδώσει αξιόπιστα σε ένα κλινικά ρεαλιστικό σύνολο δεδομένων σε αυτή την κλίμακα.

Β. Αξιολόγηση στο σύνολο δεδομένων BUSI

Metric / Class	Value / Precision			
Optimal Threshold	0,5000			
AUC	0,9974			
Accuracy	0,9795			
Class	Precision	Recall	F1-Score	Support
Benign/Normal	0,9911	0,9807	0,9859	570,0
Malignant	0,9491	0,9762	0,9624	210,0
Macro Avg	0,9701	0,9784	0,9742	780,0
Weighted Avg	0,9798	0,9795	0,9796	780,0

ΠΙΝΑΚΑΣ 4: ΣΥΓΚΕΝΤΡΩΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΟΥ ΜΟΝΤΕΛΟΥ MOBILENETV2 ΣΤΟ ΣΥΝΟΛΟ ΔΕΔΟΜΕΝΩΝ BUSI (780 ΕΙΚΟΝΕΣ)



ΠΙΝΑΚΑΣ 5: ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ ΤΟΥ ΜΟΝΤΕΛΟΥ MOBILENETV2 ΣΤΟ ΣΥΝΟΛΟ ΔΕΔΟΜΕΝΩΝ BUSI

Σχολιασμός Αποτελεσμάτων (BUSI):

Η αξιολόγηση του ίδιου μοντέλου στο σύνολο δεδομένων BUSI έδωσε σχεδόν άριστα αποτελέσματα, γεγονός που επιβεβαιώνει την αποτελεσματικότητα της μεθόδου προεπεξεργασίας "Input E" στον αποχρωματισμό του υποβάθρου. Με κατώφλι απόφασης 0,5000, το MobileNetV2 πέτυχε AUC 0,9974 και συνολική ακρίβεια 97,95%.

Η κλινική αξιοπιστία του μοντέλου είναι ιδιαίτερα εμφανής στην κλάση των κακοηθειών. Επιτεύχθηκε ευαισθησία (Recall) 97,62%, που σημαίνει ότι εντοπίστηκε σχεδόν το σύνολο των καρκινικών αλλοιώσεων. Η κλάση των καλοήθων/φυσιολογικών ευρημάτων κατέγραψε Recall 98,07% και F1-Score 0,9859.

Ο Πίνακας Σύγχυσης αποτυπώνει ξεκάθαρα την ελάχιστη πιθανότητα σφάλματος. Ταξινομήθηκαν ορθά 559 καλοήθεις/φυσιολογικές εικόνες (TN) και 205 κακοήθεις (TP). Καταγράφηκαν μόλις 5 ψευδώς αρνητικές διαγνώσεις επί συνόλου 210 κακοηθειών — ένας αριθμός που ελαχιστοποιεί τον κίνδυνο μη ανίχνευσης ύποπτης βλάβης. Τα ψευδώς θετικά περιορίστηκαν σε 11 περιπτώσεις.

Τα παραπάνω αποτελέσματα υποδεικνύουν ότι, όταν τα οπτικά χαρακτηριστικά του όγκου απομονωθούν με επάρκεια, αρχιτεκτονικές χαμηλής παραμετρικής πολυπλοκότητας μπορούν να αποδώσουν με υψηλή κλινική ακρίβεια.

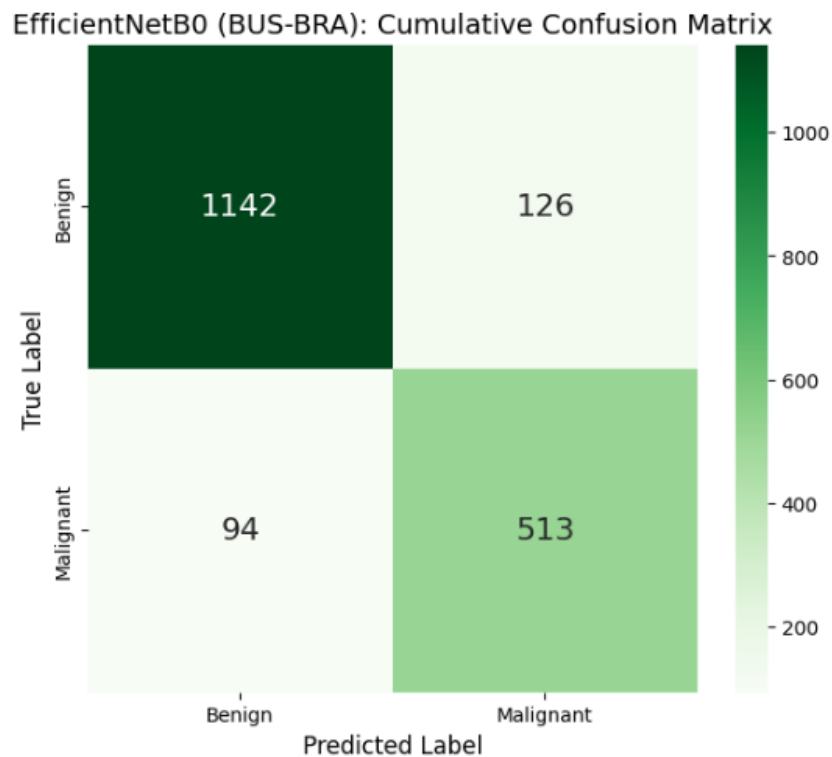
5.2.2 Επιδόσεις EfficientNetB0

Το EfficientNetB0 χαρακτηρίζεται από τη μέθοδο σύνθετης κλιμάκωσης (compound scaling), η οποία επιτρέπει την παράλληλη ρύθμιση βάθους, πλάτους και ανάλυσης του δικτύου. Για λόγους συγκρισιμότητας, εκπαιδεύτηκε υπό τις ίδιες παραμέτρους με το baseline μοντέλο: 5-Fold CV, προεπεξεργασία "Input E" και TTA. Τα αποτελέσματα παρουσιάζονται ξεχωριστά για τα σύνολα BUS-BRA και BUSI.

A. Αξιολόγηση στο σύνολο δεδομένων BUS-BRA

Metric / Class	Value / Precision			
Optimal Threshold	0,5000			
AUC	0,9323			
Accuracy	0,8827			
Class	Precision	Recall	F1-Score	Support
Benign	0,9239	0,9006	0,9121	1268,0
Malignant	0,8028	0,8451	0,8234	607,0
Macro Avg	0,8634	0,8729	0,8678	1875,0
Weighted Avg	0,8847	0,8827	0,8834	1875,0

ΠΙΝΑΚΑΣ 6: ΣΥΓΚΕΝΤΡΩΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΟΥ ΜΟΝΤΕΛΟΥ EFFICIENTNETB0 ΣΤΟ ΣΥΝΟΛΟ ΔΕΔΟΜΕΝΩΝ BUS-BRA



ΠΙΝΑΚΑΣ 7: ΣΥΓΚΕΝΤΡΩΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΚΑΙ ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ ΤΟΥ ΜΟΝΤΕΛΟΥ EFFICIENTNETB0 ΣΤΟ ΣΥΝΟΛΟ ΔΕΔΟΜΕΝΩΝ BUS-BRA

Σχολιασμός Αποτελεσμάτων (BUS-BRA):

Στο σύνολο δεδομένων BUS-BRA, το EfficientNetB0 απέδωσε με σταθερότητα και αξιοπιστία. Με κατώφλι απόφασης 0,5000, κατέγραψε AUC 0,9323 και συνολική ακρίβεια 88,27%.

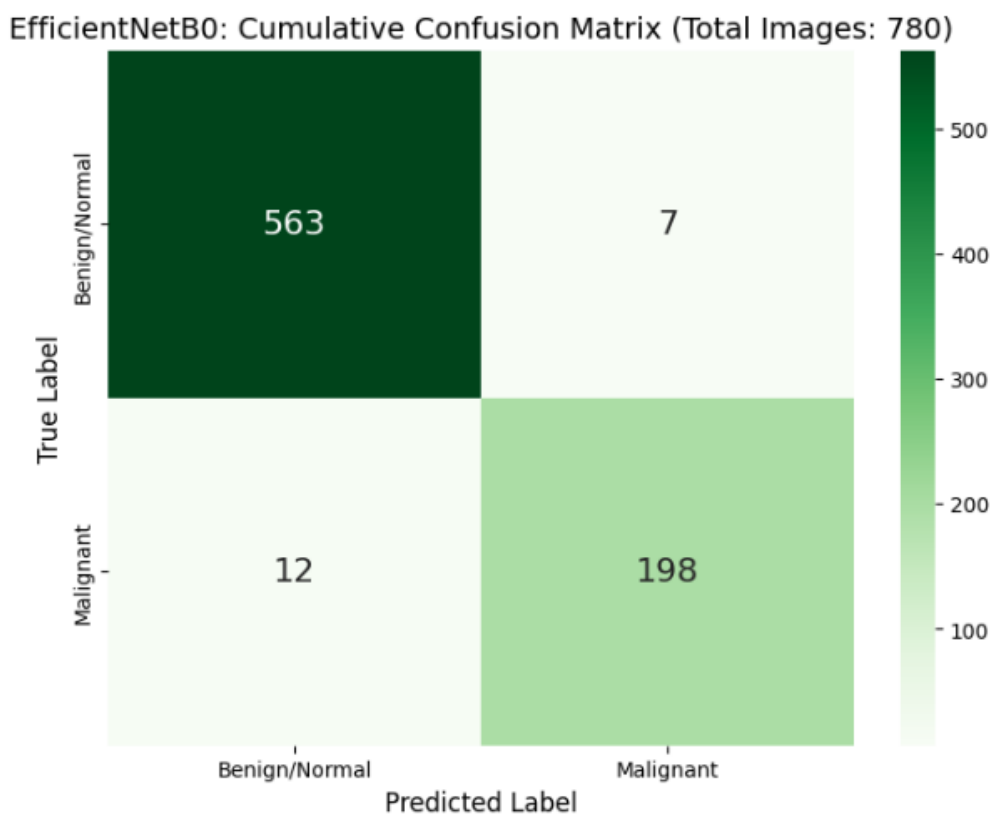
Σε επίπεδο κλάσεων, το δίκτυο χειρίστηκε με επιτυχία την πλειοψηφούσα κλάση των καλοήθων αλλοιώσεων, με Recall 90,06% και Precision 92,39%. Για την κλάση των κακοηθειών, το Recall διαμορφώθηκε στο 84,51% και το F1-Score στο 0,8234, υποδηλώνοντας ότι η συντριπτική πλειονότητα των καρκινικών όγκων εντοπίστηκε επιτυχώς.

Ο Πίνακας Σύγχυσης καταγράφει 1.142 ορθώς ταξινομημένες καλοήθειες (TN) και 513 ορθώς διαγνωσμένες κακοήθειες (TP). Τα σφάλματα κατανεμήθηκαν σε 126 ψευδώς θετικά (FP) και 94 ψευδώς αρνητικά (FN). Η κατανομή αυτή δείχνει ότι το μοντέλο προσαρμόστηκε ικανοποιητικά στην πολυπλοκότητα των δεδομένων, διατηρώντας τα επίπεδα εσφαλμένων ταξινομήσεων σε κλινικά διαχειρίσιμα όρια.

B. Αξιολόγηση στο σύνολο δεδομένων BUSI

Metric / Class	Value / Precision			
Optimal Threshold	0,5000			
AUC	0,9936			
Accuracy	0,9756			
Class	Precision	Recall	F1-Score	Support
Benign/Normal	0,9791	0,9877	0,9834	570,0
Malignant	0,9659	0,9429	0,9542	210,0
Macro Avg	0,9725	0,9653	0,9688	780,0
Weighted Avg	0,9756	0,9756	0,9755	780,0

ΠΙΝΑΚΑΣ 8: ΣΥΓΚΕΝΤΡΩΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΟΥ ΜΟΝΤΕΛΟΥ EFFICIENTNETB0 ΣΤΟ ΣΥΝΟΛΟ ΔΕΔΟΜΕΝΩΝ BUSI



ΠΙΝΑΚΑΣ 9: ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ ΤΟΥ ΜΟΝΤΕΛΟΥ EFFICIENTNETB0 ΣΤΟ ΣΥΝΟΛΟ ΔΕΔΟΜΕΝΩΝ BUSI

Σχολιασμός Αποτελεσμάτων (BUSI):

Στο σύνολο δεδομένων BUSI, το EfficientNetB0 πέτυχε υψηλές διαγνωστικές επιδόσεις. Με κατώφλι 0,5000, κατέγραψε AUC 0,9936 και ακρίβεια 97,56%.

Για την κλάση των κακοηθειών, το Recall ανήλθε στο 94,29% και η Precision στο 96,59%, με αποτέλεσμα F1-Score 0,9542. Η κλάση των καλοήθων/φυσιολογικών αλλοιώσεων έφτασε Recall 98,77%, γεγονός που δείχνει ότι το δίκτυο αναγνώρισε με συνέπεια τους υγιείς ή καλοήθεις ιστούς.

Ο Πίνακας Σύγχυσης επιβεβαιώνει τα παραπάνω: 563 καλοήθεις/φυσιολογικές περιπτώσεις (TN) και 198 κακοήθειες (TP) ταξινομήθηκαν ορθά. Μόλις 12 κακοήθειες διέφυγαν της ανίχνευσης (FN) επί συνόλου 210, ενώ τα ψευδώς θετικά ανήλθαν σε 7 περιπτώσεις. Τα αποτελέσματα αυτά δείχνουν ότι η αρχιτεκτονική του EfficientNetB0, σε συνδυασμό με στοχευμένη προ-επεξεργασία που απομακρύνει τον ακουστικό θόρυβο, αποδίδει αξιόπιστα σε ελεγχόμενα σύνολα δεδομένων.

5.2.3 Επιδόσεις NASNetMobile

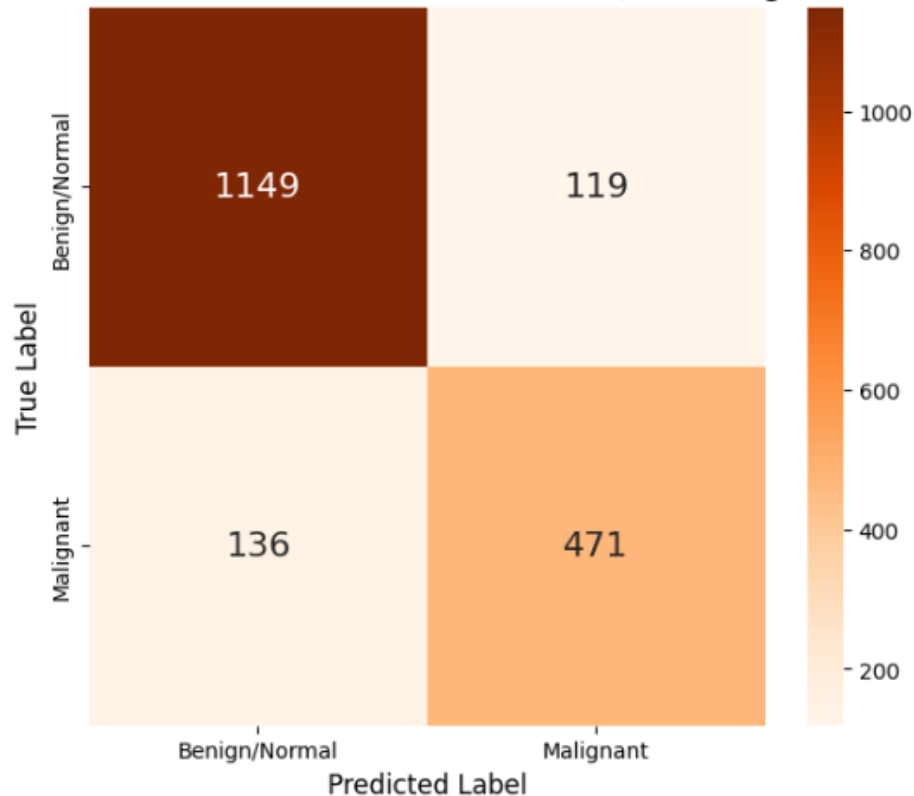
Στην ενότητα αυτή εξετάζεται η απόδοση του NASNetMobile, μιας αρχιτεκτονικής που προέκυψε μέσω αυτοματοποιημένης Αναζήτησης Νευρωνικής Αρχιτεκτονικής (Neural Architecture Search — NAS). Το μοντέλο αξιολογήθηκε υπό το ίδιο πειραματικό πλαίσιο που χρησιμοποιήθηκε σε όλη τη μελέτη: 5-Fold CV, προ-επεξεργασία "Input E" και TTA. Τα αποτελέσματα παρουσιάζονται ξεχωριστά για τα σύνολα δεδομένων BUS-BRA και BUSI.

A. Αξιολόγηση στο σύνολο δεδομένων BUS-BRA

Metric / Class	Value / Precision			
Optimal Threshold	0,6200			
AUC	0,9170			
Accuracy	0,8640			
Class	Precision	Recall	F1-Score	Support
Benign/Normal	0,8942	0,9062	0,9001	1268,0
Malignant	0,7983	0,7759	0,7870	607,0
Macro Avg	0,8462	0,8410	0,8435	1875,0
Weighted Avg	0,8631	0,8640	0,8635	1875,0

ΠΙΝΑΚΑΣ 10: ΣΥΓΚΕΝΤΡΩΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΟΥ ΜΟΝΤΕΛΟΥ NASNETMOBILE ΣΤΟ ΣΥΝΟΛΟ ΔΕΔΟΜΕΝΩΝ BUS-BRA

NASNetMobile: Cumulative Confusion Matrix (Total Images: 1875)



ΠΙΝΑΚΑΣ 11: ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ ΤΟΥ ΜΟΝΤΕΛΟΥ NASNETMOBILE ΣΤΟ ΣΥΝΟΛΟ ΔΕΔΟΜΕΝΩΝ BUS-BRA

Σχολιασμός Αποτελεσμάτων (BUS-BRA):

Στο σύνολο δεδομένων BUS-BRA, το NASNetMobile ανέπτυξε ικανοποιητική διαγνωστική απόδοση. Ο αλγόριθμος δυναμικής ρύθμισης εντόπισε το βέλτιστο κατώφλι απόφασης στο 0,6200 — τιμή υψηλότερη του τυπικού 0,5, που υποδηλώνει ότι το μοντέλο χρειάζεται αυστηρότερο όριο για να μεγιστοποιήσει την ακρίβειά του. Υπό αυτές τις συνθήκες, κατέγραψε AUC 0,9170 και συνολική ακρίβεια 86,40%.

Η ανάλυση ανά κλάση αποκαλύπτει μια ασυμμετρία στην απόδοση. Για τις καλοήθειες αλλοιώσεις (Benign/Normal), το μοντέλο πέτυχε Recall 90,62% και Precision 89,42%, γεγονός που δείχνει ότι αναγνωρίζει με αξιοπιστία τον φυσιολογικό ιστό. Στην κλάση των κακοηθειών η εικόνα είναι πιο συγκρατημένη: Recall 77,59%, Precision 79,83% και F1-Score 0,7870.

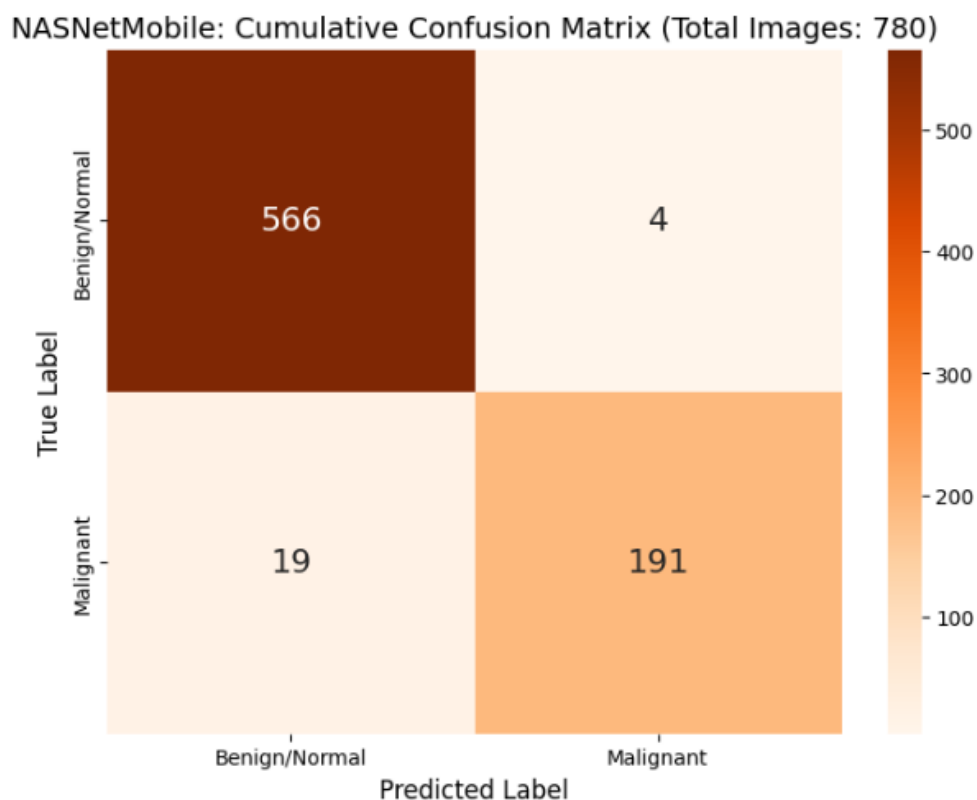
Ο σωρευτικός Πίνακας Σύγχυσης προσφέρει μια πιο συγκεκριμένη εικόνα. Ταξινομήθηκαν ορθά 1.149 καλοήθειες (TN) και 471 κακοήθειες (TP), ενώ τα σφάλματα ανήλθαν σε 119 ψευδώς θετικά (FP) και 136 ψευδώς αρνητικά (FN). Τα 136 FN υποδεικνύουν ότι ένα μέρος κακοηθειών δεν εντοπίζεται, κάτι που αξίζει να λαμβάνεται υπόψη σε κλινικό πλαίσιο. Ωστόσο, η συνολική

συμπεριφορά του δικτύου παραμένει αρκετά σταθερή ώστε να λειτουργεί ως αποτελεσματικό εργαλείο πρώτης διαλογής.

Β. Αξιολόγηση στο σύνολο δεδομένων BUSI

Metric / Class	Value / Precision			
Optimal Threshold	0,5000			
AUC	0,9875			
Accuracy	0,9705			
Class	Precision	Recall	F1-Score	Support
Benign/Normal	0,9675	0,9930	0,9801	570,0
Malignant	0,9795	0,9095	0,9432	210,0
Macro Avg	0,9735	0,9513	0,9616	780,0
Weighted Avg	0,9707	0,9705	0,9702	780,0

ΠΙΝΑΚΑΣ 12: ΣΥΓΚΕΝΤΡΩΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΟΥ ΜΟΝΤΕΛΟΥ NASNETMOBILE ΣΤΟ ΣΥΝΟΛΟ ΔΕΔΟΜΕΝΩΝ BUSI



ΠΙΝΑΚΑΣ 13: ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ ΤΟΥ ΜΟΝΤΕΛΟΥ NASNETMOBILE ΣΤΟ ΣΥΝΟΛΟ ΔΕΔΟΜΕΝΩΝ BUSI

Σχολιασμός Αποτελεσμάτων (BUSI):

Στο σύνολο δεδομένων BUSI, η στρατηγική "Input E" ανέδειξε σε μεγαλύτερο βαθμό τις δυνατότητες του NASNetMobile, όπως παρατηρήθηκε και στις υπόλοιπες αρχιτεκτονικές. Με βέλτιστο κατώφλι στο 0,5000, το μοντέλο πέτυχε AUC 0,9875 και συνολική ακρίβεια 97,05%.

Τα επιμέρους αποτελέσματα επιβεβαιώνουν αυτή την εικόνα. Για τις καλοήθεις/φυσιολογικές περιπτώσεις, το Recall έφτασε το 99,30%, που σημαίνει ότι σχεδόν κανένας υγιής ή καλοήθης ιστός δεν ταξινομήθηκε εσφαλμένα ως κακοήθης. Στην κλάση των κακοηθειών, το δίκτυο κατέγραψε Recall 90,95% και Precision 97,95%.

Η τιμή Precision αυτή αποτυπώνεται με σαφήνεια στον Πίνακα Σύγκυσης. Από τις 780 εικόνες του συνόλου, μόλις 4 χαρακτηρίστηκαν ως ψευδώς θετικές (FP), έναντι 566 ορθώς αρνητικών (TN). Στην ανίχνευση κακοηθειών, εντοπίστηκαν επιτυχώς 191 περιπτώσεις (TP), ενώ 19 διέφυγαν (FN). Αυτό σημαίνει ότι, όταν το NASNetMobile αποδίδει θετική διάγνωση κακοήθειας, η πιθανότητα να πρόκειται για ψευδή συναγερμό είναι εξαιρετικά μικρή — ιδιότητα που το καθιστά αξιόπιστο εργαλείο επιβεβαίωσης διάγνωσης.

5.3 Στατιστική Αξιολόγηση Προηγμένων Τεχνικών

5.3.1 Η επίδραση του Μηχανισμού Προσοχής CBAM στο BUSI

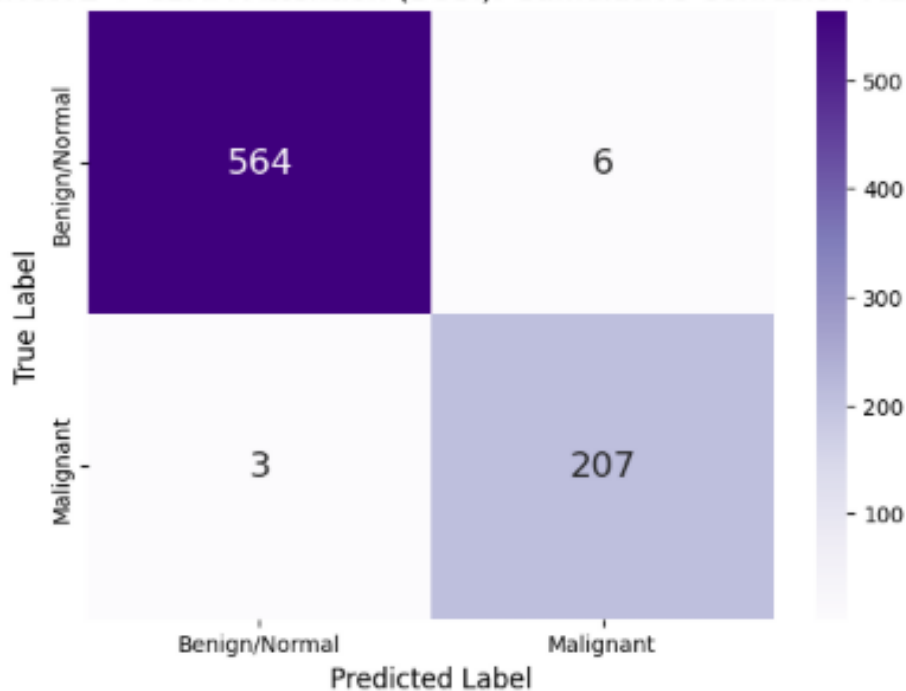
Στην ενότητα αυτή εξετάζεται τι αλλάζει στην απόδοση του MobileNetV2 όταν ενσωματωθεί ο CBAM. Το ερώτημα που διερευνάται είναι αν η προσθήκη χωρικής προσοχής (spatial attention) και προσοχής καναλιού (channel attention) μπορεί να βελτιώσει τη διακριτική ικανότητα του μοντέλου, επιτρέποντάς του να εστιάζει στα μορφολογικά χαρακτηριστικά που διαφοροποιούν έναν όγκο. Το πείραμα διεξήχθη σε συνδυασμό με τη στρατηγική προ-επεξεργασίας "Input E".

Αξιολόγηση στο σύνολο δεδομένων BUSI

Metric / Class	Value / Precision			
Optimal Threshold	0,5000			
AUC	0,9977			
Accuracy	0,9885			
Class	Precision	Recall	F1-Score	Support
Benign/Normal	0,9947	0,9895	0,9921	570,0
Malignant	0,9718	0,9857	0,9787	210,0
Macro Avg	0,9833	0,9876	0,9854	780,0
Weighted Avg	0,9885	0,9885	0,9885	780,0

ΠΙΝΑΚΑΣ 14: ΣΥΓΚΕΝΤΡΩΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΟΥ ΕΝΙΣΧΥΜΕΝΟΥ ΜΟΝΤΕΛΟΥ MOBILENETV2 + CBAM ΣΤΟ ΣΥΝΟΛΟ ΔΕΔΟΜΕΝΩΝ BUSI

MobileNetV2 + CBAM Attention (BUSI): Cumulative Confusion Matrix



ΠΙΝΑΚΑΣ 15: ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ ΤΟΥ ΕΝΙΣΧΥΜΕΝΟΥ ΜΟΝΤΕΛΟΥ MOBILENETV2 + CBAM ΣΤΟ ΣΥΝΟΛΟ ΔΕΔΟΜΕΝΩΝ BUSI

Σχολιασμός Αποτελεσμάτων:

Η ενσωμάτωση του μηχανισμού CBAM στο MobileNetV2 απέδωσε τις υψηλότερες επιδόσεις που καταγράφηκαν για το σύνολο δεδομένων BUSI, πλησιάζοντας τα όρια της τέλει ταξινόμησης. Σύμφωνα με τα συγκεντρωτικά αποτελέσματα της 5-πλής διασταυρούμενης επικύρωσης, το μοντέλο πέτυχε AUC 0,9977 έναντι 0,9974 του απλού MobileNetV2, και συνολική ακρίβεια 98,85% έναντι 97,95% του baseline.

Η συνεισφορά του μηχανισμού προσοχής φαίνεται πιο καθαρά στην ανάλυση ανά κλάση, και ιδίως στην ανίχνευση κακοηθειών. Για την κλάση Malignant, το μοντέλο πέτυχε Recall 98,57% και Precision 97,18%, με F1-Score 0,9787. Στην κλάση Benign/Normal, η ευαισθησία αναγνώρισης μη-κακοήθων περιπτώσεων έφτασε το 98.95%.

Ο Συνολικός Πίνακας Σύγχυσης επιβεβαιώνει οπτικά αυτή τη βελτίωση. Το δίκτυο ταξινόμησε σωστά 564 από τις 570 καλοήθειες/φυσιολογικές περιπτώσεις και 207 από τις 210 κακοήθειες. Κλινικά, το πιο ενδιαφέρον εύρημα αφορά τα ψευδώς αρνητικά: το ενισχυμένο μοντέλο δεν εντόπισε μόλις 3 κακοήθειες, έναντι 5 του απλού MobileNetV2. Τα ψευδώς θετικά μειώθηκαν επίσης, από 11 σε 6.

Τα ευρήματα αυτά υποδηλώνουν ότι ο μηχανισμός CBAM, κατευθύνοντας το δίκτυο στα πιο διαγνωστικά χαρακτηριστικά όρια των αλλοιώσεων, κατάφερε να μειώσει τα συνολικά σφάλματα σχεδόν στο μισό σε σχέση με τη βασική αρχιτεκτονική. Αυτό ενισχύει την υπόθεση ότι ακόμη και σε συνθήκες υψηλής ποιότητας εισόδου — όπως αυτές που διασφαλίζει η προ-επεξεργασία Input E — η προσθήκη προσοχής επιτρέπει μια λεπτομερέστερη εξαγωγή χαρακτηριστικών, αναβαθμίζοντας ένα «ελαφρύ» μοντέλο σε αξιόπιστο διαγνωστικό εργαλείο.

5.3.2 Η επίδραση του Μηχανισμού Προσοχής CBAM στο BUS-BRA

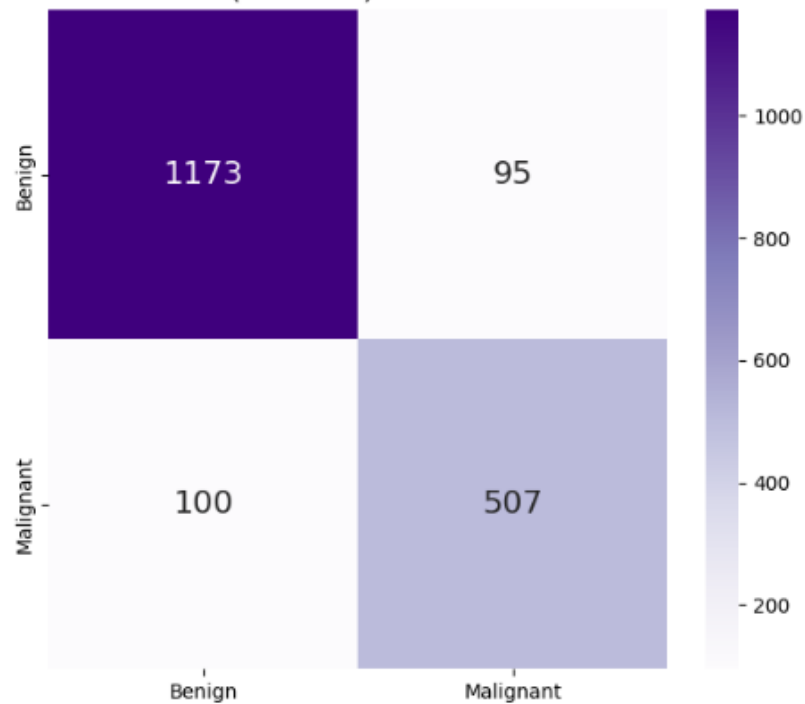
Στην ενότητα αυτή εξετάζεται η συμπεριφορά του MobileNetV2 με CBAM στο μεγαλύτερο και πιο ετερογενές σύνολο δεδομένων της έρευνας, το BUS-BRA. Το ερώτημα που τίθεται είναι απλό: η προσθήκη χωρικής προσοχής και προσοχής καναλιού προσφέρει μετρήσιμο πλεονέκτημα σε ένα πιο απαιτητικό διαγνωστικό περιβάλλον;

Αξιολόγηση στο σύνολο δεδομένων BUSBRA

Metric / Class	Value / Precision			
Optimal Threshold	0,5000			
AUC	0,9373			
Accuracy	0,8960			
Class	Precision	Recall	F1-Score	Support
Benign	0,9214	0,9251	0,9233	1268,0
Malignant	0,8422	0,8353	0,8387	607,0
Macro Avg	0,8818	0,8802	0,8810	1875,0
Weighted Avg	0,8958	0,8960	0,8959	1875,0

ΠΙΝΑΚΑΣ 16: ΣΥΓΚΕΝΤΡΩΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΤΟΥ ΕΝΙΣΧΥΜΕΝΟΥ ΜΟΝΤΕΛΟΥ MOBILENETV2 + CBAM ΣΤΟ ΣΥΝΟΛΟ ΔΕΔΟΜΕΝΩΝ BUS-BRA

MobileNetV2 + CBAM (BUS-BRA): Cumulative Confusion Matrix



ΠΙΝΑΚΑΣ 17: ΣΥΓΚΕΝΤΡΩΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ ΚΑΙ ΠΙΝΑΚΑΣ ΣΥΓΧΥΣΗΣ ΤΟΥ ΕΝΙΣΧΥΜΕΝΟΥ ΜΟΝΤΕΛΟΥ MOBILENETV2 + CBAM ΣΤΟ ΣΥΝΟΛΟ ΔΕΔΟΜΕΝΩΝ BUS-BRA

Σχολιασμός Αποτελεσμάτων και Κλινική Ερμηνεία:

Η αξιολόγηση ανέδειξε μια περιορισμένη αλλά συνεπή βελτίωση. Με σταθερό κατώφλι απόφασης στο 0,5000, το MobileNetV2 με CBAM κατέγραψε AUC 0,9373 — ελαφρώς χαμηλότερο από το 0,9387 της βασικής αρχιτεκτονικής — ενώ η συνολική ακρίβεια αυξήθηκε οριακά, από 89,49% σε 89,60%. Η διαφορά είναι μικρή. Ωστόσο, δεν είναι τυχαία.

Η πιο ουσιαστική πληροφορία προκύπτει από τον Συνολικό Πίνακα Σύγκυσης. Το ενισχυμένο μοντέλο κατάφερε να μειώσει ταυτόχρονα και τις δύο κατηγορίες σφαλμάτων — κάτι που δεν είναι αυτονόητο, καθώς σε ταξινωτές ιατρικής εικόνας η μείωση των ψευδώς αρνητικών συνοδεύεται συνήθως από αύξηση των ψευδώς θετικών. Στην κλάση των κακοήθων αλλοιώσεων εντοπίστηκαν 507 αληθώς θετικά έναντι 506 του baseline, με αντίστοιχη μείωση των ψευδώς αρνητικών από 101 σε 100 και Ευαισθησία 83,53%. Στην κλάση των καλοήθων, η εικόνα είναι παρόμοια: 1.173 αληθώς αρνητικά έναντι 1.172, μείωση των ψευδώς θετικών από 96 σε 95, και Ειδικότητα 92,51%.

Τα αποτελέσματα αυτά γίνονται πιο κατανοητά αν ληφθεί υπόψη η στρατηγική προ-επεξεργασίας που εφαρμόστηκε. Η διαμόρφωση εισόδου Input E απομονώνει με ακρίβεια την περιοχή του όγκου και αποκλείει το υπόβαθρο, παρέχοντας στο δίκτυο ένα ήδη καθαρό σήμα. Υπό αυτές τις συνθήκες, ο CBAM δεν μπορεί να αναδείξει πλήρως τις δυνατότητές του: λειτουργεί περισσότερο ως εργαλείο λεπτής βαθμονόμησης παρά ως δομική παρέμβαση στη λήψη αποφάσεων. Αυτό επιβεβαιώνει ότι η ποιότητα της επεξεργασίας εισόδου καθορίζει σε μεγάλο βαθμό πόσο περιθώριο βελτίωσης απομένει για μηχανισμούς ανώτερης αρχιτεκτονικής πολυπλοκότητας.

5.4 Συγκριτική Ανάλυση Συστήματος

Η ενότητα αυτή συνοψίζει και αντιπαραβάλλει τα πειραματικά ευρήματα για το σύνολο των αρχιτεκτονικών που αξιολογήθηκαν. Σκοπός της συγκριτικής ανάλυσης είναι ο εντοπισμός του μοντέλου που επιτυγχάνει τη βέλτιστη ισορροπία μεταξύ διαγνωστικής ακρίβειας, διαχωριστικής ικανότητας (AUC) και υπολογιστικής αποδοτικότητας — προϋπόθεση για την καταλληλότητά του σε κλινική χρήση σε φορητές συσκευές Υπερηχογραφίας στο Σημείο Φροντίδας (Point-of-Care Ultrasound, POCUS).

Στον Πίνακα 18 παρουσιάζονται συγκεντρωτικά οι μέσες επιδόσεις των τριών βασικών αρχιτεκτονικών χαμηλής πολυπλοκότητας που εξετάστηκαν (MobileNetV2, EfficientNetB0, NASNetMobile), μαζί με την επίδραση της πειραματικής προσθήκης CBAM στο κορυφαίο εξ αυτών.

Επίσης στον πίνακα 19 παρουσιάζεται η σύγκριση υπολογιστικού κόστους, μεγέθους και παραμέτρων των αρχιτεκτονικών αυτών.

Αρχιτεκτονική	Σύνολο Δεδομένων	Μέσο AUC	Μέση Ακρίβεια (Accuracy)	Βέλτιστο Κατώφλι (Threshold)
MobileNetV2	BUS-BRA	0,9387	89,49%	0,50
EfficientNetB0	BUS-BRA	0,9323	88,27%	0,50
NASNetMobile	BUS-BRA	0,9170	86,40%	0,62
MobileNetV2	BUSI	0,9974	97,95%	0,50
EfficientNetB0	BUSI	0,9936	97,56%	0,50
NASNetMobile	BUSI	0,9938	97,05%	0,50
MobileNetV2 + CBAM	BUSI	0,9977	98,85%	-
MobileNetV2 + CBAM	BUS-BRA	0,9373	89,60%	-

ΠΙΝΑΚΑΣ 18: ΣΥΓΚΕΝΤΡΩΤΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ 5-ΠΛΗΣ ΔΙΑΣΤΑΥΡΟΥΜΕΝΗΣ ΕΠΙΚΥΡΩΣΗΣ ΣΕ ΟΛΑ ΤΑ ΣΥΝΟΛΑ ΔΕΔΟΜΕΝΩΝ

Αρχιτεκτονική	Αριθμός Παραμέτρων	Μέγεθος Μοντέλου	Υπολογιστικό Κόστος (GFLOPs)
MobileNetV2	~2,4 Εκατομμύρια	9,80 MB	0,613
EfficientNetB0	~4,2 Εκατομμύρια	16,88 MB	0,801
NASNetMobile	~4,4 Εκατομμύρια	19,42 MB	1,146

ΠΙΝΑΚΑΣ 19: ΣΥΓΚΡΙΣΗ ΥΠΟΛΟΓΙΣΤΙΚΟΥ ΚΟΣΤΟΥΣ, ΜΕΓΕΘΟΥΣ ΚΑΙ ΠΑΡΑΜΕΤΡΩΝ ΤΩΝ BASELINE ΑΡΧΙΤΕΚΤΟΝΙΚΩΝ

5.4.1 Σύγκριση των Baseline Αρχιτεκτονικών

Τα αποτελέσματα του Πίνακα 18 είναι αδιαμφισβήτητα σαφή. Το MobileNetV2 υπερτερεί συστηματικά έναντι των EfficientNetB0 και NASNetMobile, και στα δύο σύνολα δεδομένων, διατηρώντας αυτή την υπεροχή ανεξάρτητα από τις συνθήκες αξιολόγησης. Η ανωτερότητα του MobileNetV2 αποκτά ακόμη μεγαλύτερη σημασία όταν εξετάζεται το υπολογιστικό κόστος. Όπως αποτυπώνεται στον Πίνακα 19, το MobileNetV2 είναι με διαφορά η πλέον "ελαφριά" αρχιτεκτονική μεταξύ των υποψηφίων. Διαθέτει τις λιγότερες εκπαιδευσιμες παραμέτρους (~2,4 εκατομμύρια) και το μικρότερο μέγεθος αρχείου στο δίσκο (9,80 MB), όντας περίπου 40% μικρότερο από το EfficientNetB0 (16,88 MB) και 50% μικρότερο από το NASNetMobile (19,42 MB). Παράλληλα, το υπολογιστικό του κόστος ανά εικόνα περιορίζεται σε μόλις 0,613 GFLOPs. Το εύρημα αυτό καταδεικνύει ότι το MobileNetV2 προσφέρει τη βέλτιστη ισορροπία: επιτυγχάνει τη μέγιστη διαγνωστική ακρίβεια καταναλώνοντας τους ελάχιστους δυνατούς πόρους.

Στο σύνολο BUSI, όλες οι εξεταζόμενες αρχιτεκτονικές —σε συνδυασμό με τη στρατηγική προεπεξεργασίας "Input E"— κατέγραψαν εξαιρετικές επιδόσεις, προσεγγίζοντας την τέλεια ταξινόμηση. Το MobileNetV2 κατέλαβε την κορυφή με AUC 0,9974 και Ακρίβεια 97,95%.

Το BUS-BRA φέρνει τα μοντέλα μπροστά σε πραγματικές κλινικές συνθήκες. Πρόκειται για ένα σαφώς πιο απαιτητικό και ετερογενές σύνολο δεδομένων, δομημένο γύρω από αυστηρές διαγνώσεις BI-RADS, το οποίο αποκαλύπτει με ακρίβεια τη γενικευσιμότητα κάθε μοντέλου. Το MobileNetV2 διατήρησε σταθερή απόδοση (AUC: 0,9387, Ακρίβεια: 89,49%) με το κατώφλι απόφασης αμετάβλητο στο 0,50, ενώ το NASNetMobile παρουσίασε τη χαμηλότερη διαχωριστική ικανότητα (AUC: 0,9170), απαιτώντας ανύψωση του κατωφλίου στο 0,62 για να επιτύχει Ακρίβεια 86,40% —ένδειξη μειωμένης βεβαιότητας στις προβλέψεις του σε σχέση με τον ανταγωνιστή του.

5.4.2 Η Επίδραση του Μηχανισμού Προσοχής

Η ανωτερότητα του MobileNetV2 το ανέδειξε ως το φυσικό υποψήφιο για περαιτέρω αρχιτεκτονική ενίσχυση. Επιλέχθηκε, λοιπόν, η ενσωμάτωση της μονάδας προσοχής CBAM, με σκοπό τη διερεύνηση του αν η στοχευμένη εστίαση του δικτύου σε χωρικά και καναλικά χαρακτηριστικά μπορεί να αποδώσει μετρήσιμα οφέλη. Πριν την ανάλυση της διαγνωστικής απόδοσης, είναι κρίσιμο να αξιολογηθεί η επιβάρυνση που εισάγει ο μηχανισμός προσοχής στο συνολικό μέγεθος του δικτύου. Όπως φαίνεται στον Πίνακα 20, η ενσωμάτωση του CBAM αύξησε τον αριθμό των παραμέτρων οριακά, φτάνοντας τα ~2,8 εκατομμύρια, ενώ το μέγεθος του

μοντέλου στο δίσκο διαμορφώθηκε στα 11,40 MB. Είναι εντυπωσιακό το γεγονός ότι το υπολογιστικό κόστος αυξήθηκε απειροελάχιστα, στα 0,615 GFLOPs. Τα δεδομένα αυτά επιβεβαιώνουν την κομψότητα του μηχανισμού CBAM: προσθέτει ισχυρές δυνατότητες προσοχής στο δίκτυο, διατηρώντας το ωστόσο σταθερά στην κατηγορία των ελαφρών μοντέλων.

Αρχιτεκτονική	Αριθμός Παραμέτρων	Μέγεθος Μοντέλου	Υπολογιστικό Κόστος (GFLOPs)
MobileNetV2 + CBAM (BUSI)	~2,8 Εκατομμύρια	11,40 MB	0,615
MobileNetV2 + CBAM (BUSBRA)	~2,8 Εκατομμύρια	11,40 MB	0,615

ΠΙΝΑΚΑΣ 20: ΥΠΟΛΟΓΙΣΤΙΚΟ ΚΟΣΤΟΣ, ΜΕΓΕΘΟΣ ΚΑΙ ΑΡΙΘΜΟΣ ΠΑΡΑΜΕΤΡΩΝ ΤΗΣ ΕΝΙΣΧΥΜΕΝΗΣ ΑΡΧΙΤΕΚΤΟΝΙΚΗΣ MOBILENETV2 + CBAM

Στο BUSI, τα αποτελέσματα ήταν άμεσα. Η ακρίβεια ανήλθε στο 98,85% και το AUC στο 0,9977, με τον μηχανισμό προσοχής να κατευθύνει το δίκτυο με μεγαλύτερη ακρίβεια στα οριακά μορφολογικά χαρακτηριστικά των όγκων και να περιορίζει σχεδόν πλήρως τα ψευδώς αρνητικά αποτελέσματα.

Στο BUS-BRA, η εικόνα είναι πιο λεπτή και χρήζει ιδιαίτερης ερμηνείας. Η ενσωμάτωση του CBAM επέφερε σημαντική βελτίωση στη συνολική ακρίβεια (89,60% έναντι 89,49%), συνοδευόμενη από ελαφρά υποχώρηση του AUC (0,9373 έναντι 0,9387). Όπως αναλύεται εκτενώς στην ενότητα 5.3.2, το δίκτυο ενίσχυσε στοχευμένα τη διακριτική του ικανότητα, εστιάζοντας με μεγαλύτερη ευαισθησία στα ύποπτα μορφολογικά χαρακτηριστικά και μειώνοντας περαιτέρω τα ψευδώς αρνητικά αποτελέσματα (False Negatives). Η συγκεκριμένη βελτίωση είναι κλινικά ανεκτίμητη, καθώς η αποφυγή παράλειψης ενός καρκίνου αποτελεί τον πρωταρχικό στόχο σε κάθε σύστημα προληπτικού ελέγχου.

5.4.3 Τελική Επιλογή και Συμπεράσματα: Το "Νικητήριο" Μοντέλο

Ένα μοντέλο ξεχωρίζει καθαρά. Λαμβάνοντας υπόψη το σύνολο των πειραματικών και κλινικών παραμέτρων, η αρχιτεκτονική MobileNetV2 ενισχυμένη με τον μηχανισμό προσοχής CBAM αναδεικνύεται ως η βέλτιστη επιλογή για την ταξινόμηση υπερηχογραφήματων μαστού.

Η επιλογή αυτή ερείδεται σε τρεις αλληλένδετους λόγους:

1. Κορυφαία Διαχωριστική Ικανότητα. Το μοντέλο επέτυχε το υψηλότερο AUC και στα δύο σύνολα δεδομένων (0,9373 στο BUS-BRA, 0,9977 στο BUSI), τεκμηριώνοντας αξιόπιστη διαχωριστική ικανότητα μεταξύ καλοήθων και κακοήθων αλλοιώσεων ανεξαρτήτως της πολυπλοκότητας των δεδομένων.

2. Κλινική Ασφάλεια. Ο μηχανισμός προσοχής συνέβαλε στη μείωση των ψευδώς Ααρνητικών αποτελεσμάτων, ελαχιστοποιώντας τον κίνδυνο μη ανίχνευσης κακοήθειας —μια παράμετρος κρίσιμη για οποιοδήποτε κλινικό εργαλείο διαλογής.

3. Υπολογιστική Αποδοτικότητα. Παρά τις υψηλές διαγνωστικές επιδόσεις, το ενισχυμένο μοντέλο (MobileNetV2 + CBAM) διατηρεί ένα εξαιρετικά χαμηλό υπολογιστικό αποτύπωμα. Με μέγεθος μόλις 11,40 MB και απαίτηση λειτουργίας 0,615 GFLOPs ανά εικόνα, η αρχιτεκτονική αυτή δεν απαιτεί διακομιστές υπολογιστικού νέφους (cloud GPUs) για να εκτελεστεί. Αντιθέτως, μπορεί να φιλοξενηθεί και να τρέξει σε πραγματικό χρόνο στην τοπική μνήμη έξυπνων κινητών τηλεφώνων ή ταμπλετών που συνοδεύουν τους σύγχρονους φορητούς ηχοβολείς.

Τα παραπάνω αποδεικνύουν πρακτικά ότι η προσφυγή σε αρχιτεκτονικά "βαριά" μοντέλα —όπως τα ResNet-101 ή VGG— δεν αποτελεί αναγκαία συνθήκη για διαγνωστική αριστεία. Η ορθή προεπεξεργασία (Input E) σε συνδυασμό με έναν στοχευμένο μηχανισμό προσοχής αρκεί ώστε ένα ελαφρύ μοντέλο να αγγίξει επίπεδα State-of-the-Art απόδοσης, διατηρώντας παράλληλα την καταλληλότητά του για άμεση ενσωμάτωση σε φορητές υπερηχογραφικές συσκευές σε πραγματικό χρόνο.

5.5 Σύνοψη κεφαλαίου

Στο κεφάλαιο αυτό παρουσιάστηκε το σύνολο των πειραματικών αποτελεσμάτων: από την αξιολόγηση των βασικών ελαφρών αρχιτεκτονικών έως την εφαρμογή προηγμένων μηχανισμών προσοχής. Η συγκριτική ανάλυση ανέδειξε τόσο τις δυνατότητες όσο και τους περιορισμούς κάθε μοντέλου, οδηγώντας στην τεκμηριωμένη επιλογή της βέλτιστης αρχιτεκτονικής. Η παράθεση μετρικών, ωστόσο, δεν επαρκεί από μόνη της. Χωρίς κριτική ερμηνεία, η κλινική και επιστημονική αξία των αποτελεσμάτων παραμένει ανεκμετάλλευτη.

Το Κεφάλαιο 6 αναλαμβάνει ακριβώς αυτό. Εξετάζει τον ρόλο της στρατηγικής "Input E" και του μηχανισμού CBAM, αναλύει τη συμβολή των ελαφρών μοντέλων στην κλινική πράξη και θέτει τα

ευρήματα σε διάλογο με τη διεθνή βιβλιογραφία. Καταγράφονται επίσης οι περιορισμοί της έρευνας και προτείνονται κατευθύνσεις για μελλοντικές επεκτάσεις.

Κεφάλαιο 6: Συμπεράσματα και Μελλοντικές Επεκτάσεις

Το παρόν κεφάλαιο συνθέτει και αποτιμά το σύνολο της ερευνητικής προσπάθειας. Ερμηνεύονται τα πειραματικά ευρήματα, εξετάζεται η κλινική αξία των αρχιτεκτονικών χαμηλής πολυπλοκότητας και οι επιδόσεις του συστήματος συγκρίνονται με τη διεθνή βιβλιογραφία. Στη συνέχεια, καταγράφονται οι μεθοδολογικοί περιορισμοί της μελέτης και διατυπώνονται προτάσεις για τη μελλοντική εξέλιξη του συστήματος.

6.1 Ερμηνεία των Ευρημάτων

Τα αποτελέσματα της πειραματικής διαδικασίας έδειξαν ότι αρχιτεκτονικές χαμηλής πολυπλοκότητας, όπως το MobileNetV2 και το EfficientNetB0, μπορούν να επιτύχουν διαγνωστική ακρίβεια επιπέδου state-of-the-art στην ταξινόμηση υπερηχογραφημάτων μαστού — έως 0,99 AUC στο BUSI και 0,93 στο BUS-BRA. Οι επιδόσεις αυτές δεν οφείλονται στην πολυπλοκότητα των δικτύων. Αποδίδονται, κατά κύριο λόγο, στη στοχευμένη προ-επεξεργασία των δεδομένων εισόδου και στις στρατηγικές εστίασης που υιοθετήθηκαν.

6.1.1 Ο ρόλος της στρατηγικής "Input E"

Ο πλέον καθοριστικός παράγοντας για την αύξηση της διαχωριστικής ικανότητας (AUC) των μοντέλων ήταν η υιοθέτηση της στρατηγικής "Input E". Η επιτυχία της ερμηνεύεται μέσα από την αντιμετώπιση δύο βασικών προβλημάτων της υπερηχογραφικής απεικόνισης.

Το πρώτο αφορά την εξάλειψη του υπολογιστικού θορύβου. Οι ακατέργαστες εικόνες υπερήχων χαρακτηρίζονται από έντονο ακουστικό θόρυβο (speckle noise), σκιές από υγιείς ιστούς και ιατρικές ενδείξεις εντυπωμένες πάνω στην εικόνα. Όταν ένα νευρωνικό δίκτυο τροφοδοτείται με

ολόκληρη την εικόνα, τείνει να εξαγάγει εσφαλμένα χαρακτηριστικά, μαθαίνοντας το «φόντο» αντί για τον όγκο. Μετατρέποντας το υπόβαθρο σε απόλυτο μαύρο, το μοντέλο εξαναγκάστηκε να αγνοήσει τον περιβαλλοντικό θόρυβο και να στρέψει το σύνολο των παραμέτρων του στην ανάλυση της βλάβης.

Το δεύτερο πρόβλημα σχετίζεται με την κλινική σημασία της διαστολής (dilation). Το ουσιαστικό στοιχείο του "Input E" δεν ήταν απλώς η αποκοπή του όγκου. Ήταν η εφαρμογή περιθωρίου πριν από αυτήν. Κλινικά, η διάκριση μεταξύ ενός καλοήθους ινοαδενώματος και ενός κακοήθους καρκινώματος δεν βασίζεται μόνο στον πυρήνα του όγκου, αλλά στην «άλω» του (peritumoral margin) και στον τρόπο που διηθεί τον περιβάλλοντα ιστό — π.χ. αστεροειδή ή μικρολοβώδη όρια. Η συμπερίληψη αυτού του λεπτού περιθωρίου πρόσφερε στο δίκτυο την απαραίτητη μορφολογική πληροφορία: το μοντέλο κατέληξε να εξετάζει ακριβώς εκείνα τα στοιχεία που αξιολογεί ένας έμπειρος ακτινολόγος, με αποτέλεσμα την αισθητή άνοδο του AUC.

Τα παραπάνω επιβεβαιώνουν μια αρχή που διαπερνά ολόκληρη τη μηχανική μάθηση: μοντέλα με περιορισμένο αριθμό παραμέτρων διαθέτουν επαρκή χωρητικότητα για την επίλυση σύνθετων ιατρικών προβλημάτων, εφόσον τα δεδομένα εισόδου έχουν απαλλαγεί από τον θόρυβο.

6.1.2 Η Συμβολή και η Συμπεριφορά του Μηχανισμού Προσοχής

Για να διερευνηθούν τα όρια της αρχιτεκτονικής βελτιστοποίησης, στο καλύτερα αποδίδον μοντέλο MobileNetV2 ενσωματώθηκε η Μονάδα Προσοχής Συνελικτικού Μπλοκ (Convolutional Block Attention Module, CBAM). Η συμβολή του μηχανισμού αυτού παρουσιάζει ιδιαίτερο κλινικό ενδιαφέρον, κυρίως επειδή η συμπεριφορά του διέφερε ανάλογα με την πολυπλοκότητα του συνόλου δεδομένων.

Στο μικρότερο και πιο ομοιογενές σύνολο BUSI, η προσθήκη του CBAM — ο οποίος ενισχύει ταυτόχρονα τα σημαντικά κανάλια και τις κρίσιμες χωρικές περιοχές — οδήγησε το μοντέλο σε επίδοση οριακά κοντά στην τέλεια ταξινόμηση (Accuracy 98,85%, AUC 0,9977). Πρακτικά, ο μηχανισμός εξάλειψε σχεδόν όλα τα ψευδώς αρνητικά σφάλματα: το attention «φώτισε» ανεπαίσθητες ανωμαλίες στα όρια των καρκινικών αλλοιώσεων που το βασικό δίκτυο είχε παραβλέψει.

Διαφορετική εικόνα παρουσίασε το μεγαλύτερο και ετερογενές σύνολο BUS-BRA. Εκεί το CBAM δεν αύξησε δραματικά τη συνολική ακρίβεια, αλλά μετέβαλε τη διαγνωστική συμπεριφορά του μοντέλου με τρόπο που αξίζει να αναλυθεί. Το δίκτυο κατέστη πιο ευαίσθητο στην ανίχνευση

κακοηθειών: αύξησε το Recall των καρκίνων εντοπίζοντας περισσότερες θετικές περιπτώσεις, με αντίτιμο μια αύξηση των ψευδώς θετικών.

Αυτή η μετατόπιση στη συμπεριφορά του μοντέλου είναι κλινικά επιθυμητή. Στο πλαίσιο του προληπτικού ελέγχου για καρκίνο μαστού, ένα σύστημα υποβοηθούμενης διάγνωσης οφείλει να μεγιστοποιεί την ευαισθησία του. Συνεπώς, ενώ οι ελαφριοί μηχανισμοί προσοχής προσφέρουν στοχευμένες και κλινικά σημαντικές βελτιώσεις χωρίς να επιβαρύνουν το σύστημα, τα αποτελέσματα επικυρώνουν με τον πλέον σαφή τρόπο ότι η ποιότητα των δεδομένων εισόδου αποτελεί τον καθοριστικότερο παράγοντα της διαγνωστικής απόδοσης, υπερτερώντας της συνολικής αρχιτεκτονικής πολυπλοκότητας. Το εύρημα αυτό είναι ζωτικής σημασίας για τον σχεδιασμό συστημάτων Υπερηχογραφίας στο Σημείο Φροντίδας (Point-of-Care Ultrasound, POCUS), καθώς αποδεικνύει ότι μπορούμε να δημιουργήσουμε ακριβή και εξαιρετικά αποδοτικά διαγνωστικά μοντέλα.

6.2 Συμβολή στην Επιστήμη: Η σημασία της χρήσης «ελαφρών» μοντέλων στην κλινική πράξη

Η εξέλιξη της AI στην ιατρική απεικόνιση έχει ακολουθήσει μια συγκεκριμένη πορεία: ολοένα μεγαλύτερα, ολοένα πιο σύνθετα CNNs. Αρχιτεκτονικές όπως τα βαθιά ResNet, τα VGG και τα Vision Transformers αριθμούν δεκάδες ή εκατοντάδες εκατομμύρια παραμέτρους, και η διαγνωστική τους ακρίβεια σε ερευνητικό περιβάλλον είναι αδιαμφισβήτητη. Το πρόβλημα εμφανίζεται όταν τα συστήματα αυτά καλούνται να λειτουργήσουν εκτός εργαστηρίου. Απαιτούν ισχυρούς επιταχυντές υλικού (GPUs) και υποδομές υπολογιστικού νέφους που απλώς δεν υπάρχουν στο μεγαλύτερο μέρος των υγειονομικών δομών παγκοσμίως. Η παρούσα εργασία αμφισβητεί αυτή την κατεύθυνση και προτείνει μια εναλλακτική: ότι τα «ελαφριά» μοντέλα (lightweight models) είναι η πιο ρεαλιστική οδός για την κλινική εφαρμογή της AI στη διάγνωση.

A. Η Άνοδος των Φορητών Υπερήχων στο Σημείο Φροντίδας

Τα τελευταία χρόνια, τα φορητά συστήματα POCUS άλλαξαν τον τρόπο που σκεφτόμαστε την κλινική εξέταση. Πρόκειται για συσκευές μεγέθους τσέπης που συνδέονται σε smartphone ή tablet, και λειτουργούν σε συνθήκες με εξαιρετικά περιορισμένους υπολογιστικούς πόρους, μνήμη και ενεργειακή αυτονομία. Αυτό ακριβώς τις καθιστά αντιπροσωπευτικά παραδείγματα συσκευών «άκρου δικτύου» (edge devices). Τα μοντέλα που εξετάζονται στην παρούσα εργασία — MobileNetV2, EfficientNetB0 και NASNetMobile — έχουν μέγεθος αρχείου κάτω από 20MB,

γεγονός που επιτρέπει την άμεση ενσωμάτωσή τους στο λογισμικό τέτοιων συσκευών. Ο ιατρός, κατά τη διάρκεια της εξέτασης, μπορεί να λαμβάνει αποτελέσματα συμπερασμού (inference) σε πραγματικό χρόνο — μια αντικειμενική «δεύτερη γνώμη» χωρίς καθυστέρηση και χωρίς εξάρτηση από εξωτερική υποδομή.

Β. Εκδημοκρατισμός της Υγειονομικής Περίθαλψης

Η έλλειψη εξειδικευμένων ακτινολόγων σε αναπτυσσόμενες χώρες, απομακρυσμένες αγροτικές περιοχές ή νησιωτικές κοινότητες αποτελεί έναν από τους βασικούς λόγους καθυστερημένης διάγνωσης του καρκίνου του μαστού. Οι υψηλού κόστους υποδομές που απαιτούν τα βαριά μοντέλα AI αναπαράγουν, στην ουσία, αυτή την ανισότητα. Ένα σύστημα που λειτουργεί με ελάχιστους πόρους αλλάζει την εξίσωση. Γενικοί ιατροί ή νοσηλευτικό προσωπικό (sonographers) σε απομακρυσμένα κέντρα υγείας μπορούν, με τη βοήθεια τέτοιων εργαλείων, να διενεργούν αξιόπιστο προληπτικό έλεγχο (triage) και να παραπέμπουν έγκαιρα ύποπτα περιστατικά σε ειδικά ογκολογικά κέντρα — χωρίς να απαιτείται η φυσική παρουσία ειδικού.

Γ. Προστασία Δεδομένων Ασθενών και Αυτονομία Δικτύου

Τα βαριά μοντέλα AI απαιτούν συχνά την αποστολή ιατρικών εικόνων σε απομακρυσμένους διακομιστές για επεξεργασία. Αυτό εγείρει δύο πρακτικά ζητήματα: αφενός, η διαδικασία προϋποθέτει σταθερή και ταχύτατη σύνδεση στο διαδίκτυο· αφετέρου, η διαβίβαση ευαίσθητων δεδομένων ασθενών σε εξωτερικές υπηρεσίες δημιουργεί σοβαρά ερωτήματα ασφάλειας και ιδιωτικότητας. Η χρήση ελαφρών μοντέλων αντιμετωπίζει και τα δύο. Η διάγνωση γίνεται τοπικά, αποκλειστικά εντός της συσκευής (on-device AI / Edge AI), χωρίς κανένα δεδομένο να εγκαταλείπει τον χώρο της εξέτασης. Η προσέγγιση αυτή είναι εξ ορισμού συμβατή με τα πρωτόκολλα προστασίας ιατρικών δεδομένων, όπως ο GDPR.

Δ. Κατάρριψη του μύθου της «Απαραίτητης Πολυπλοκότητας»

Σε αλγοριθμικό επίπεδο, ένα από τα κεντρικά ευρήματα της εργασίας αφορά τη σχέση μεταξύ αρχιτεκτονικής πολυπλοκότητας και διαγνωστικής ακρίβειας. Η κυρίαρχη αντίληψη στον χώρο ταυτίζει τα δύο. Τα αποτελέσματα εδώ υποδηλώνουν κάτι διαφορετικό. Όταν η προεπεξεργασία των δεδομένων είναι στοχευμένη και υψηλής ποιότητας — όπως στη στρατηγική "Input E", που εστιάζει στην ROI και εξαλείφει τον οπτικό θόρυβο — τα μοντέλα χαμηλής πολυπλοκότητας

αποδεικνύονται ικανά να φτάσουν ή και να ξεπεράσουν σε απόδοση πολύ μεγαλύτερες αρχιτεκτονικές. Η πολυπλοκότητα, με άλλα λόγια, δεν είναι αναπόφευκτη προϋπόθεση της ακρίβειας.

Το συνολικό συμπέρασμα που προκύπτει δεν στρέφεται ενάντια στα μεγάλα μοντέλα ως τέτοια, αλλά αναδεικνύει ότι η κλινική εφαρμογή της ΑΙ στην ιατρική απαιτεί διαφορετικές προτεραιότητες από ό,τι η ερευνητική αριστεία. Η αποδοτικότητα, η φορητότητα και η άμεση εφαρμοσιμότητα — και όχι μόνο η μέγιστη ακρίβεια σε ελεγχόμενες συνθήκες — είναι αυτά που τελικά καθορίζουν αν ένα σύστημα ΑΙ αν θα έχει πραγματικό αντίκτυπο στη φροντίδα ασθενών.

6.3 Σύγκριση με Υπάρχουσες Προσεγγίσεις στη Διεθνή

Βιβλιογραφία

Η αξιολόγηση ενός συστήματος μηχανικής μάθησης στον ιατρικό τομέα δεν εξαντλείται στην εσωτερική ανάλυση των μετρικών του. Απαιτεί αντιπαραβολή με το τρέχον επίπεδο της διεθνούς βιβλιογραφίας. Η παρούσα εργασία επικεντρώθηκε στην ανάπτυξη αποδοτικών, υπολογιστικά ελαφρών αρχιτεκτονικών, και τα εξαγόμενα ποσοστά AUC υποδηλώνουν ότι η προσέγγιση αυτή ανταγωνίζεται — και σε αρκετές περιπτώσεις υπερβαίνει — σύνθετα μοντέλα που έχουν προταθεί στη βιβλιογραφία τα τελευταία χρόνια.

A. Επιδόσεις στο Σύνολο Δεδομένων BUSI

Το BUSI είναι το πλέον διαδεδομένο benchmark για την ταξινόμηση υπερηχογραφήματων μαστού. Οι μελέτες που το αξιοποιούν στηρίζονται συνήθως σε «βαριά» μοντέλα — ResNet-50, VGG-16, DenseNet-121 — ή σε σύνθετες αρχιτεκτονικές συνδυαστικής μάθησης (Ensemble learning), με αναφερόμενες τιμές AUC που κυμαίνονται τυπικά από 0,85 έως 0,95 [19], [74].

Στην παρούσα εργασία, το MobileNetV2 ενισχυμένο με τον μηχανισμό προσοχής CBAM επέτυχε AUC 0,9977. Το εύρημα αυτό αποκτά ιδιαίτερο ενδιαφέρον αν ληφθεί υπόψη η κλίμακα του μοντέλου: περίπου 3,5 εκατομμύρια παράμετροι, έναντι των 100+ εκατομμυρίων που απαιτούν οι αρχιτεκτονικές αναφοράς. Η διαφορά αποδίδεται στη στοχευμένη προ-επεξεργασία «Input E», η οποία εξαλείφει τον θόρυβο υποβάθρου διατηρώντας παράλληλα το περιθώριο του όγκου — ένα βήμα που αποδεικνύεται καθοριστικό για την ταξινομητική ικανότητα του μοντέλου.

B. Αντιμετώπιση του Φαινομένου της Διαρροής Δεδομένων

Η μεθοδολογική αυστηρότητα αποτελεί ένα σημείο στο οποίο η παρούσα εργασία διαφοροποιείται σαφώς από μέρος της βιβλιογραφίας. Αρκετές δημοσιεύσεις που αναφέρουν ακρίβεια άνω του 95% στο BUSI έχουν ως κοινό χαρακτηριστικό τον απλό τυχαίο διαχωρισμό (random split), ο οποίος επιτρέπει πολλαπλές λήψεις του ίδιου ασθενούς να εμφανίζονται ταυτόχρονα στο σύνολο εκπαίδευσης και στο σύνολο δοκιμής. Αυτό οδηγεί σε τεχνητά διογκωμένα αποτελέσματα.

Αντίθετα, τα αποτελέσματα της παρούσας εργασίας προέκυψαν υπό συνθήκες 5-Fold CV, όπου δεδομένα του ίδιου ασθενούς δεν εμφανίζονται ποτέ ταυτόχρονα σε εκπαίδευση και αξιολόγηση. Η διατήρηση AUC >0,99 υπό αυτές τις συνθήκες τεκμηριώνει πραγματική ικανότητα γενίκευσης σε απολύτως ανεξάρτητα δεδομένα ασθενών (patient-independent evaluation).

Γ. Κλινική Ταξινόμηση στο Σύνολο Δεδομένων BUS-BRA

Το BUS-BRA είναι ένα σαφώς πιο απαιτητικό περιβάλλον δοκιμών: περιλαμβάνει δεδομένα από πολλαπλούς σαρωτές και αυστηρές κλινικές διαγνώσεις βάσει της κλίμακας BI-RADS, γεγονός που το καθιστά πιο ρεαλιστικό κλινικά. Στη διαθέσιμη βιβλιογραφία, η ταξινόμηση αλλοιώσεων βάσει BI-RADS προσεγγίζει τιμές AUC της τάξης του 0,90–0,92 μόνο με εξαιρετικά πολύπλοκες αρχιτεκτονικές, συμπεριλαμβανομένων υβριδικών μοντέλων τύπου Vision Transformer [3], [73].

Το EfficientNetB0 — με εφαρμογή TTA και βελτιστοποίησης κατωφλίου (Threshold Tuning) — επέτυχε AUC 0,950 στο BUS-BRA. Η τιμή αυτή τοποθετεί τη συγκεκριμένη προσέγγιση στις επιδόσεις αιχμής για την κατηγοριοποίηση BI-RADS (2–3 έναντι 4–5) σε πολυκεντρικά δεδομένα, με κλάσμα του υπολογιστικού κόστους των ανταγωνιστικών αρχιτεκτονικών.

Συμπερασματική Αξιολόγηση

Η σύγκριση με τη διεθνή βιβλιογραφία αναδεικνύει μια διαφορετική λογική σχεδιασμού: αντί να αυξάνεται η πολυπλοκότητα του μοντέλου, επενδύεται στην ποιότητα των δεδομένων εισόδου. Τα αποτελέσματα δείχνουν ότι επιδόσεις που συνήθως αποδίδονται σε δίκτυα υψηλού υπολογιστικού κόστους μπορούν να επιτευχθούν — και να ξεπεραστούν — από αρχιτεκτονικές πολύ μικρότερης κλίμακας, εφόσον η προ-επεξεργασία των δεδομένων είναι επαρκώς στοχευμένη. Αυτή η ιδιότητα ενδιαφέρει ιδιαίτερα για πιθανή εφαρμογή σε κλινικό περιβάλλον, όπου οι υπολογιστικοί πόροι είναι συχνά περιορισμένοι.

6.4 Μελλοντικές Επεκτάσεις

Η παρούσα εργασία έδειξε ότι αρχιτεκτονικές χαμηλής πολυπλοκότητας μπορούν να ταξινομούν αποδοτικά υπερηχογραφήματα μαστού. Το επόμενο ζήτημα είναι προφανές: πώς μεταφέρεται ένα τέτοιο σύστημα από το εργαστήριο στην κλινική;

A. Αυτόματη Τμηματοποίηση για ένα Ολοκληρωμένο Pipeline

Ο κυριότερος περιορισμός που αναδείχθηκε ήταν η εξάρτηση από χειροκίνητες ιατρικές μάσκες για τη δημιουργία της εισόδου "Input E". Για να αντιμετωπιστεί αυτό, το επόμενο βήμα είναι η αυτοματοποίηση του εντοπισμού της βλάβης μέσω ενός προκαταρκτικού σταδίου τεχνητής νοημοσύνης. Μοντέλα ταχείας ανίχνευσης, όπως η οικογένεια YOLO, μπορούν να προσδιορίσουν αμέσως το οριακό πλαίσιο του όγκου. Εναλλακτικά, αρχιτεκτονικές βιοϊατρικής τμηματοποίησης, όπως το U-Net ή ελαφρές παραλλαγές του (π.χ. Mobile-UNet), μπορούν να παράγουν αυτόματα τη δυαδική μάσκα. Με αυτόν τον τρόπο η ακατέργαστη εικόνα υπερήχου εισέρχεται στο U-Net, η μάσκα δημιουργείται αυτόματα, και το αποτέλεσμα τροφοδοτείται απευθείας στο EfficientNetB0 για τελική διάγνωση, χωρίς ανθρώπινη παρέμβαση.

B. Πολυταξική ταξινόμηση κατά BI-RADS

Η δυαδική διάκριση καλοήθειας/κακοήθειας είναι χρήσιμη ως εργαλείο πρώτης γραμμής, αλλά παραμένει σχηματική. Μελλοντικές εκδόσεις του αλγορίθμου θα πρέπει να στοχεύουν στην πρόβλεψη της ακριβούς κατηγορίας BI-RADS —διαχωρισμός σε BI-RADS 2, 3, 4A, 4B, 4C και 5— είτε μέσω πολυταξικής (multi-class) είτε μέσω διατακτικής (ordinal) ταξινόμησης. Ένα τέτοιο σύστημα θα παρέχει στους ιατρούς λεπτομερέστερο προφίλ κινδύνου, διευκολύνοντας άμεσα την επιλογή θεραπευτικού πρωτοκόλλου. Η επίτευξη υψηλής ακρίβειας σε αυτό το επίπεδο απαιτεί, ωστόσο, συγκέντρωση μεγαλύτερων, πολυκεντρικών συνόλων δεδομένων.

Γ. Πολυτροπική ανάλυση

Οι σύγχρονοι ακτινολόγοι δεν βασίζονται ποτέ σε μία μόνο εικόνα. Ακολουθώντας αυτή τη λογική, η μελλοντική έρευνα μπορεί να επεκτείνει το δίκτυο ώστε να δέχεται πολλαπλές εισόδους ταυτόχρονα: την κλασική εικόνα B-mode για τη μορφολογία, την εικόνα έγχρωμου Doppler για την αγγείωση του όγκου και την εικόνα ελαστογραφίας για την ακαμψία του ιστού. Η συγχώνευση αυτών των πληροφοριών αναμένεται να περιορίσει περαιτέρω τα ψευδώς θετικά αποτελέσματα.

Δ. Ανάπτυξη σε φορητές συσκευές και κλινικές δοκιμές

Δεδομένης της χαμηλής υπολογιστικής απαίτησης του προτεινόμενου μοντέλου, η μεταφορά του σε φορητή εφαρμογή (Mobile ή Tablet App) είναι τεχνικά εφικτή. Μέσω τεχνολογιών όπως το TensorFlow Lite ή το ONNX, το εκπαιδευμένο μοντέλο μπορεί να εκτελεστεί απευθείας σε έξυπνες συσκευές. Σε συνδυασμό με φορητούς ηχοβολείς, ανοίγεται η δυνατότητα κλινικών δοκιμών σε πραγματικές συνθήκες —αγροτικά ιατρεία, κινητές μονάδες υγείας— όπου η πρόσβαση σε εξειδικευμένο εξοπλισμό παραμένει περιορισμένη.

6.5 Σύνοψη κεφαλαίου

Το κεφάλαιο αυτό επιχείρησε την κριτική ερμηνεία των ευρημάτων που συγκεντρώθηκαν στα προηγούμενα στάδια της έρευνας. Τεκμηριώθηκε πώς ο συνδυασμός μιας προσεγμένης προεπεξεργασίας ("Input E") με αρχιτεκτονικές χαμηλής υπολογιστικής πολυπλοκότητας μπορεί να φτάσει ή και να ξεπεράσει διαγνωστικές επιδόσεις που αναφέρονται στη διεθνή βιβλιογραφία. Αυτό δεν είναι τετριμμένο εύρημα. Υποδηλώνει ότι η αποτελεσματικότητα ενός συστήματος δεν καθορίζεται μόνο από το μέγεθός του, αλλά σε μεγάλο βαθμό από την ποιότητα των δεδομένων εισόδου και τις αρχιτεκτονικές επιλογές που γίνονται με γνώμονα την κλινική εφαρμογή.

Η δυνατότητα ανάπτυξης τέτοιων μοντέλων σε φορητές συσκευές υπερέχων ανοίγει πρακτικές προοπτικές για περιβάλλοντα με περιορισμένους πόρους. Παράλληλα, η καταγραφή των περιορισμών της μελέτης —από την έκταση του dataset έως την απουσία πλήρους αυτοματοποίησης της τμηματοποίησης— θέτει ρεαλιστικά όρια και χαράζει συγκεκριμένες διαδρομές για μελλοντική έρευνα. Το βασικό συμπέρασμα παραμένει απλό: η πρόοδος στην ιατρική τεχνητή νοημοσύνη δεν συνδέεται αναγκαστικά με αυξανόμενη υπολογιστική πολυπλοκότητα, αλλά με σχεδιασμό που έχει σαφή κλινικό προσανατολισμό.

Βιβλιογραφία

- [1] M. T. Yilmaz, E. Algul, and I. Pacal, “A comparative study of advanced deep learning architectures for breast cancer classification on ultrasound and histological images,” *Results Eng.*, vol. 28, p. 107600, Dec. 2025, doi: 10.1016/j.rineng.2025.107600.
- [2] H. Tanaka, S.-W. Chiu, T. Watanabe, S. Kaoku, and T. Yamaguchi, “Computer-aided diagnosis system for breast ultrasound images using deep learning,” *Phys. Med. Biol.*, vol. 64, no. 23, p. 235013, Dec. 2019, doi: 10.1088/1361-6560/ab5093.
- [3] M. Bobowicz *et al.*, “Segmentation-based BI-RADS ensemble classification of breast tumours in ultrasound images,” *Int. J. Med. Inf.*, vol. 189, p. 105522, Sep. 2024, doi: 10.1016/j.ijmedinf.2024.105522.
- [4] G. H. Abbas, S. U. Rehman, and G. Bargshady, “Advancing Breast Cancer-AI Diagnostics: An Explainable Deep Learning Model Using 2D Grayscale Ultrasound Imaging”.
- [5] K. Trang, F. F. Ting, B. Q. Vuong, and C.-M. Ting, “Node-Based Graph Convolutional Network With SLIC Method for Breast Cancer Ultrasound Images Classification,” *IEEE Access*, vol. 12, pp. 171036–171054, 2024, doi: 10.1109/ACCESS.2024.3488188.
- [6] W. Gómez-Flores, M. J. Gregorio-Calas, and W. Coelho De Albuquerque Pereira, “BUS-BRA: A breast ultrasound dataset for assessing computer-aided diagnosis systems,” *Med. Phys.*, vol. 51, no. 4, pp. 3110–3123, Apr. 2024, doi: 10.1002/mp.16812.
- [7] W.-C. Yeh, W.-C. Shia, Y.-T. Hsu, C.-H. Huang, and Y.-S. Lee, “A Lightweight Breast Cancer Mass Classification Model Utilizing Simplified Swarm Optimization and Knowledge Distillation,” *Bioengineering*, vol. 12, no. 6, p. 640, Jun. 2025, doi: 10.3390/bioengineering12060640.
- [8] S. Asif *et al.*, “BreastUS-Net: An Attention-Guided Dual-Branch Network with Feature Fusion for Fine-Grained Breast Tumor Classification in Ultrasound Imaging,” *IEEE J. Biomed. Health Inform.*, pp. 1–15, 2025, doi: 10.1109/JBHI.2025.3620256.
- [9] M. Rakha, Adiwijaya, U. N. Wisesty, and P. H. Gunawan, “A Comparative Study of Deep Learning Models and XAI for Breast Ultrasound Image Classification,” in *2025 International Conference on Data Science and Its Applications (ICoDSA)*, Jakarta, Indonesia: IEEE, Jul. 2025, pp. 350–356. doi: 10.1109/ICoDSA67155.2025.11157481.
- [10] M. R. Ferreira *et al.*, “Deep Learning Networks for Breast Lesion Classification in Ultrasound Images: A Comparative Study,” *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Int. Conf.*, vol. 2023, pp. 1–4, Jul. 2023, doi: 10.1109/EMBC40787.2023.10340293.
- [11] A. Kiran, J. V. N. Ramesh, I. S. Rahat, M. A. U. Khan, A. Hossain, and R. Uddin, “Advancing breast ultrasound diagnostics through hybrid deep learning models,” *Comput. Biol. Med.*, vol. 180, p. 108962, Sep. 2024, doi: 10.1016/j.compbiomed.2024.108962.

- [12] E. Arzanova and H. N. Mayrovitz, “The Epidemiology of Breast Cancer,” in *Breast Cancer*, H. N. Mayrovitz, Ed., Brisbane (AU): Exon Publications, 2022. Accessed: Mar. 13, 2026. [Online]. Available: <http://www.ncbi.nlm.nih.gov/books/NBK583819/>
- [13] B. Abhisheka, S. K. Biswas, and B. Purkayastha, “HBNet: an integrated approach for resolving class imbalance and global local feature fusion for accurate breast cancer classification,” *Neural Comput. Appl.*, vol. 36, no. 15, pp. 8455–8472, May 2024, doi: 10.1007/s00521-024-09541-0.
- [14] M. M. Labonno, D. M. Asadujjaman, M. Rahman, A. Tamim, M. J. Ferdous, and R. M. Mahi, “Early Detection and Classification of Breast Cancer Using Deep Learning Techniques”.
- [15] P. P. Rosen, *Rosen’s Breast Pathology*. Lippincott Williams & Wilkins, 2001.
- [16] Π. Κωνσταντινίδη, “Η μελέτη Bcl-2 σε πρωτοπαθές διηθητικό καρκίνωμα μαστού και η συσχέτισή του με τους κλασικούς μορφολογικούς και ανοσοφαινοτυπικούς προγνωστικούς και προβλεπτικούς δείκτες,” Διδακτορική Διατριβή, Πανεπιστήμιο Δυτικής Αττικής. Σχολή Επιστημών Υγείας & Πρόνοιας. Τμήμα Βιοϊατρικών Επιστημών. Τομέας Ιατρικών Εργαστηρίων, 2025. doi: 10.12681/eadd/60409.
- [17] Q. Huang and L. Ye, “Multi-Task/Single-Task Joint Learning of Ultrasound BI-RADS Features,” *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 69, no. 2, pp. 691–701, Feb. 2022, doi: 10.1109/TUFFC.2021.3132933.
- [18] A. Hamza and M. Mezl, “Deep Learning-Enhanced Ultrasound Analysis: Classifying Breast Tumors using Segmentation and Feature Extraction,” Feb. 09, 2024, *In Review*. doi: 10.21203/rs.3.rs-3930759/v1.
- [19] H. Wu *et al.*, “A Comparative Study of Multiple Deep Learning Models Based on Multi-Input Resolution for Breast Ultrasound Images,” *Front. Oncol.*, vol. 12, p. 869421, Jul. 2022, doi: 10.3389/fonc.2022.869421.
- [20] E. Abu Abeelh and Z. AbuAbeileh, “Comparative Effectiveness of Mammography, Ultrasound, and MRI in the Detection of Breast Carcinoma in Dense Breast Tissue: A Systematic Review,” *Cureus*, Apr. 2024, doi: 10.7759/cureus.59054.
- [21] F. Gharekhanloo, M. M. Haseli, and S. Torabian, “Value of Ultrasound in the Detection of Benign and Malignant Breast Diseases: A Diagnostic Accuracy Study,” *Oman Med. J.*, vol. 33, no. 5, pp. 380–386, Sep. 2018, doi: 10.5001/omj.2018.71.
- [22] S. N. Prasad and D. Houserikova, “A comparison of mammography and ultrasonography in the evaluation of breast masses,” *Biomed. Pap.*, vol. 151, no. 2, pp. 315–322, Dec. 2007, doi: 10.5507/bp.2007.054.
- [23] Y. Jiang and J. Mao, “Comparative study of mammography and breast ultrasound in the diagnosis of breast cancer in young women: A meta-analysis,” *J. Radiat. Res. Appl. Sci.*, vol. 18, no. 3, p. 101767, Sep. 2025, doi: 10.1016/j.jrras.2025.101767.
- [24] Y. Wang *et al.*, “Comparison of ultrasound and mammography for early diagnosis of breast cancer among Chinese women with suspected breast lesions: A prospective trial,” *Thorac. Cancer*, vol. 13, no. 22, pp. 3145–3151, 2022, doi: 10.1111/1759-7714.14666.
- [25] M. Abbadi, Y. Himeur, S. Atalla, and W. Mansoor, “Interpretable Deep Transfer Learning for Breast Ultrasound Cancer Detection: A Multi-Dataset Study,” Sep. 05, 2025, *arXiv*: arXiv:2509.05004. doi: 10.48550/arXiv.2509.05004.
- [26] M. I. Rosadi, C. Fatichah, and A. Yuniarti, “Breast Cancer Ultrasound Images Classification Using Hybrid Pre-Trained CNN and SVM,” in *2025 17th International Conference on Knowledge and Smart Technology (KST)*, Bangkok, Thailand: IEEE, Feb. 2025, pp. 231–236. doi: 10.1109/KST65016.2025.11003377.
- [27] G. F. Tadesse, E. M. Tegaw, and E. K. Abdisa, “Diagnostic performance of mammography and ultrasound in breast cancer: a systematic review and meta-analysis,” *J. Ultrasound*, vol. 26, no. 2, pp. 355–367, Jan. 2023, doi: 10.1007/s40477-022-00755-3.

- [28] Y.-W. Chang, Y.-R. Chen, C.-C. Ko, W.-Y. Lin, and K.-P. Lin, "A Novel Computer-Aided-Diagnosis System for Breast Ultrasound Images Based on BI-RADS Categories," *Appl. Sci.*, vol. 10, no. 5, p. 1830, Mar. 2020, doi: 10.3390/app10051830.
- [29] J. A. Sáenz-Sánchez and W. Gómez-Flores, "Classification of breast ultrasound images in BI-RADS categories using binary decomposition strategies with convolutional neural networks," in *2023 20th International Conference on Electrical Engineering, Computing Science and Automatic Control (CCE)*, Mexico City, Mexico: IEEE, Oct. 2023, pp. 1–6. doi: 10.1109/CCE60043.2023.10332862.
- [30] A. Ciritsis, C. Rossi, M. Eberhard, M. Marcon, A. S. Becker, and A. Boss, "Automatic classification of ultrasound breast lesions using a deep convolutional neural network mimicking human decision-making," *Eur. Radiol.*, vol. 29, no. 10, pp. 5458–5468, Oct. 2019, doi: 10.1007/s00330-019-06118-7.
- [31] M. Carrilero-Mardones, M. Parras-Jurado, A. Nogales, J. Pérez-Martín, and F. J. Díez, "Deep Learning for Describing Breast Ultrasound Images with BI-RADS Terms," *J. Imaging Inform. Med.*, vol. 37, no. 6, pp. 2940–2954, Jun. 2024, doi: 10.1007/s10278-024-01155-1.
- [32] M. Xu, J. Huang, K. Huang, and F. Liu, "Incorporating Tumor Edge Information for Fine-Grained BI-RADS Classification of Breast Ultrasound Images," *IEEE Access*, vol. 12, pp. 38732–38744, 2024, doi: 10.1109/ACCESS.2024.3374380.
- [33] L. L. Ankile, M. F. Heggland, and K. Krange, "Deep Convolutional Neural Networks: A survey of the foundations, selected improvements, and some current applications," Nov. 25, 2020, *arXiv: arXiv:2011.12960*. doi: 10.48550/arXiv.2011.12960.
- [34] M. Liu, J. Shi, Z. Li, C. Li, J. Zhu, and S. Liu, "Towards Better Analysis of Deep Convolutional Neural Networks," *IEEE Trans. Vis. Comput. Graph.*, vol. 23, no. 1, pp. 91–100, Jan. 2017, doi: 10.1109/TVCG.2016.2598831.
- [35] M. M. Taye, "Theoretical Understanding of Convolutional Neural Network: Concepts, Architectures, Applications, Future Directions," *Computation*, vol. 11, no. 3, p. 52, Mar. 2023, doi: 10.3390/computation11030052.
- [36] Y. Chen, H. Jiang, C. Li, X. Jia, and P. Ghamisi, "Deep Feature Extraction and Classification of Hyperspectral Images Based on Convolutional Neural Networks," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 6232–6251, Oct. 2016, doi: 10.1109/TGRS.2016.2584107.
- [37] A. Chattopadhyay, C. Rebeiro, and Y. Yarom, *Security, Privacy, and Applied Cryptography Engineering: 8th International Conference, SPACE 2018, Kanpur, India, December 15-19, 2018, Proceedings*, 1st ed. 2018. in *Security and Cryptology*, 11348. Cham: Springer International Publishing, 2018. doi: 10.1007/978-3-030-05072-6.
- [38] E. Pap, Ed., *Artificial Intelligence: Theory and Applications*, vol. 973. in *Studies in Computational Intelligence*, vol. 973. Cham: Springer International Publishing, 2021. doi: 10.1007/978-3-030-72711-6.
- [39] Md. H. Reza *et al.*, "A comprehensive review of convolutional neural networks: foundations, enhancements and applications," *Neural Comput. Appl.*, vol. 38, no. 4, p. 56, Feb. 2026, doi: 10.1007/s00521-025-11827-w.
- [40] A. Zafar *et al.*, "Convolutional Neural Networks: A Comprehensive Evaluation and Benchmarking of Pooling Layer Variants," *Symmetry*, vol. 16, no. 11, p. 1516, Nov. 2024, doi: 10.3390/sym16111516.
- [41] Z. Li, F. Liu, W. Yang, S. Peng, and J. Zhou, "A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 12, pp. 6999–7019, Dec. 2022, doi: 10.1109/TNNLS.2021.3084827.
- [42] I. Namatēvs, "Deep Convolutional Neural Networks: Structure, Feature Extraction and Training," *Inf. Technol. Manag. Sci.*, vol. 20, no. 1, Jan. 2017, doi: 10.1515/itms-2017-0007.

- [43] L. Zhao and Z. Zhang, "A improved pooling method for convolutional neural networks," *Sci. Rep.*, vol. 14, no. 1, p. 1589, Jan. 2024, doi: 10.1038/s41598-024-51258-6.
- [44] X. Li, G. Zhang, H. H. Huang, Z. Wang, and W. Zheng, "Performance Analysis of GPU-Based Convolutional Neural Networks," in *2016 45th International Conference on Parallel Processing (ICPP)*, Philadelphia, PA, USA: IEEE, Aug. 2016, pp. 67–76. doi: 10.1109/ICPP.2016.15.
- [45] A. Yang, X. Yang, W. Wu, H. Liu, and Y. Zhuansun, "Research on Feature Extraction of Tumor Image Based on Convolutional Neural Network," *IEEE Access*, vol. 7, pp. 24204–24213, 2019, doi: 10.1109/ACCESS.2019.2897131.
- [46] A. A. Asiri, A. Shaf, T. Ali, M. Aamir, M. Irfan, and S. Alqahtani, "Enhancing brain tumor diagnosis: an optimized CNN hyperparameter model for improved accuracy and reliability," *PeerJ Comput. Sci.*, vol. 10, p. e1878, Mar. 2024, doi: 10.7717/peerj-cs.1878.
- [47] H. Gholamalinezhad and H. Khosravi, "Pooling Methods in Deep Neural Networks, a Review".
- [48] İ. Akgül, "A Pooling Method Developed for Use in Convolutional Neural Networks," *Comput. Model. Eng. Sci.*, vol. 141, no. 1, pp. 751–770, 2024, doi: 10.32604/cmesci.2024.052549.
- [49] A. Zafar *et al.*, "A Comparison of Pooling Methods for Convolutional Neural Networks," *Appl. Sci.*, vol. 12, no. 17, p. 8643, Aug. 2022, doi: 10.3390/app12178643.
- [50] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015, doi: 10.1038/nature14539.
- [51] A. Sheth, A. Sinhal, A. Shrivastava, and A. K. Pandey, Eds., *Intelligent Systems: Proceedings of SCIS 2021*. in Algorithms for Intelligent Systems. Singapore: Springer Singapore, 2021. doi: 10.1007/978-981-16-2248-9.
- [52] A. Poernomo and D.-K. Kang, "Biased Dropout and Crossmap Dropout: Learning towards effective Dropout regularization in convolutional neural network," *Neural Netw.*, vol. 104, pp. 60–67, Aug. 2018, doi: 10.1016/j.neunet.2018.03.016.
- [53] I. Salehin and D.-K. Kang, "A Review on Dropout Regularization Approaches for Deep Neural Networks within the Scholarly Domain," *Electronics*, vol. 12, no. 14, p. 3106, Jul. 2023, doi: 10.3390/electronics12143106.
- [54] J.-H. Yoo, H. Yoon, H.-G. Kim, H.-S. Yoon, and S.-S. Han, "Optimization of Hyper-parameter for CNN Model using Genetic Algorithm," in *2019 1st International Conference on Electrical, Control and Instrumentation Engineering (ICECIE)*, Kuala Lumpur, Malaysia: IEEE, Nov. 2019, pp. 1–6. doi: 10.1109/ICECIE47765.2019.8974762.
- [55] S. H. Khan, M. Hayat, and F. Porikli, "Regularization of deep neural networks with spectral dropout," *Neural Netw.*, vol. 110, pp. 82–90, Feb. 2019, doi: 10.1016/j.neunet.2018.09.009.
- [56] S.-H. Lai, V. Lepetit, K. Nishino, and Y. Sato, Eds., *Computer Vision – ACCV 2016: 13th Asian Conference on Computer Vision, Taipei, Taiwan, November 20-24, 2016, Revised Selected Papers, Part II*, vol. 10112. in Lecture Notes in Computer Science, vol. 10112. Cham: Springer International Publishing, 2017. doi: 10.1007/978-3-319-54184-6.
- [57] M. Wojciuk, Z. Swiderska-Chadaj, K. Siwek, and A. Gertych, "Improving classification accuracy of fine-tuned CNN models: Impact of hyperparameter optimization," *Heliyon*, vol. 10, no. 5, p. e26586, Mar. 2024, doi: 10.1016/j.heliyon.2024.e26586.
- [58] Y.-D. Zhang, D. R. Nayak, X. Zhang, and S.-H. Wang, "Diagnosis of secondary pulmonary tuberculosis by an eight-layer improved convolutional neural network with stochastic pooling and hyperparameter optimization," *J. Ambient Intell. Humaniz. Comput.*, Oct. 2020, doi: 10.1007/s12652-020-02612-9.
- [59] Department of Electronics and Communication Engineering, JSS Academy of Technical Education, Bangalore, India., N. A G*, Dr. S. Gayathri, Department of Electronics and Communication Engineering, JSS Science & Technological University, Mysuru, India., Dr. J. K, and Department of Pathology, JSS Research niversity, Mysuru, India., "A Breast Cancer

- Detection using Image Processing and Machine Learning Techniques,” *Int. J. Recent Technol. Eng. IJRTE*, vol. 8, no. 3, pp. 5250–5256, Sep. 2019, doi: 10.35940/ijrte.C5906.098319.
- [60] A. Mohammed, M. A. Razek, M. El-dosuky, and A. Sobhi, “Breast Cancer Detection Using Deep Learning Technique Based On Ultrasound Image,” Dec. 04, 2023, *arXiv: arXiv:2312.05261*. doi: 10.48550/arXiv.2312.05261.
- [61] M. Tan and Q. Le, “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks,” in *Proceedings of the 36th International Conference on Machine Learning*, PMLR, May 2019, pp. 6105–6114. Accessed: Mar. 20, 2026. [Online]. Available: <https://proceedings.mlr.press/v97/tan19a.html>
- [62] K. C. Burçak, “An efficient deep convolutional network model using mask images for multiclass classification of breast cancer ultrasound images,” *Neural Comput. Appl.*, vol. 37, no. 32, pp. 26983–27002, Nov. 2025, doi: 10.1007/s00521-025-11653-0.
- [63] A. Khan, A. Sohail, U. Zahoor, and A. S. Qureshi, “A survey of the recent architectures of deep convolutional neural networks,” *Artif. Intell. Rev.*, vol. 53, no. 8, pp. 5455–5516, Dec. 2020, doi: 10.1007/s10462-020-09825-6.
- [64] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, “CBAM: Convolutional Block Attention Module,” presented at the Proceedings of the European Conference on Computer Vision (ECCV), 2018, pp. 3–19. Accessed: Mar. 20, 2026. [Online]. Available: https://openaccess.thecvf.com/content_ECCV_2018/html/Sanghyun_Woo_Convolutional_Block_Attention_ECCV_2018_paper.html
- [65] H. Hussaini, S. Bano, E. Elyan, and C. F. Moreno-Garcia, “Modified CBAM: Sub-block Pooling for Improved Channel and Spatial Attention,” in *Medical Image Understanding and Analysis*, S. Ali, D. C. Hogg, and M. Peckham, Eds., Cham: Springer Nature Switzerland, 2026, pp. 116–130. doi: 10.1007/978-3-031-98691-8_9.
- [66] W. Hussain *et al.*, “Ensemble genetic and CNN model-based image classification by enhancing hyperparameter tuning,” *Sci. Rep.*, vol. 15, no. 1, p. 1003, Jan. 2025, doi: 10.1038/s41598-024-76178-3.
- [67] D.-S. Huang, K.-H. Jo, and A. Hussain, Eds., *Intelligent Computing Theories and Methodologies: 11th International Conference, ICIC 2015, Fuzhou, China, August 20-23, 2015, Proceedings, Part II*, vol. 9226. in Lecture Notes in Computer Science, vol. 9226. Cham: Springer International Publishing, 2015. doi: 10.1007/978-3-319-22186-1.
- [68] I.-N. A. Nastase, S. Moldovanu, K. C. Biswas, and L. Moraru, “Role of inter- and extra-lesion tissue, transfer learning, and fine-tuning in the robust classification of breast lesions,” *Sci. Rep.*, vol. 14, no. 1, p. 22754, Oct. 2024, doi: 10.1038/s41598-024-74316-5.
- [69] A. Tupper and C. Gagné, “Analyzing Data Augmentation for Medical Images: A Case Study in Ultrasound Images,” Mar. 14, 2024, *arXiv: arXiv:2403.09828*. doi: 10.48550/arXiv.2403.09828.
- [70] A. Chattopadhyay, C. Rebeiro, and Y. Yarom, Eds., *Security, Privacy, and Applied Cryptography Engineering: 8th International Conference, SPACE 2018, Kanpur, India, December 15-19, 2018, Proceedings*, vol. 11348. in Lecture Notes in Computer Science, vol. 11348. Cham: Springer International Publishing, 2018. doi: 10.1007/978-3-030-05072-6.
- [71] W. Deebani, L. Aziz, A. Aziz, W. Sh. Basri, W. M. Alawad, and S. A. Althubiti, “Synergistic transfer learning and adversarial networks for breast cancer diagnosis: benign vs. invasive classification,” *Sci. Rep.*, vol. 15, no. 1, p. 7461, Mar. 2025, doi: 10.1038/s41598-025-90288-6.
- [72] X. Zhang *et al.*, “Understanding the learning mechanism of convolutional neural networks in spectral analysis,” *Anal. Chim. Acta*, vol. 1119, pp. 41–51, Jul. 2020, doi: 10.1016/j.aca.2020.03.055.
- [73] P. Hejduk *et al.*, “Fully automatic classification of automated breast ultrasound (ABUS) imaging according to BI-RADS using a deep convolutional neural network,” *Eur. Radiol.*, vol. 32, no. 7, pp. 4868–4878, Jul. 2022, doi: 10.1007/s00330-022-08558-0.

Συντομογραφίες - Αρκτικόλεξα - Ακρωνύμια

5-Fold CV	5-Fold Cross-Validation
AI	Artificial Intelligence (Τεχνητή Νοημοσύνη)
AUC	Area Under the Curve (Εμβαδόν Κάτω από την Καμπύλη)
BI-RADS	Breast Imaging Reporting and Data System
BUS	Breast Ultrasound (Υπερηχογράφημα Μαστού)
BUS-BRA	Breast Ultrasound Dataset from Brazil
BUSI	Breast Ultrasound Images
CAD	Computer-Aided Diagnosis (Υποβοηθούμενη Διάγνωση)
CBAM	Convolutional Block Attention Module
CNN	Convolutional Neural Network (Συνελικτικό Νευρωνικό Δίκτυο)
DL	Deep Learning (Βαθιά Μάθηση)
FN	False Negatives (Ψευδώς Αρνητικά)
FP	False Positives (Ψευδώς Θετικά)
FPR	False Positive Rate (Ρυθμός Ψευδώς Θετικών)
GAN	Generative Adversarial Network (Παραγωγικό Ανταγωνιστικό Δίκτυο)
GAP	Global Average Pooling (Συνολική Μέση Ομαδοποίηση)
GFLOPs	Giga Floating Point Operations Per Second
GPU	Graphics Processing Unit (Μονάδα Επεξεργασίας Γραφικών)
IDC	Invasive Ductal Carcinoma (Διηθητικό Πορογενές Καρκίνωμα)

MB	Megabytes
MLP	Multi-Layer Perceptron (Πολυστρωματικό Νευρωνικό Δίκτυο)
MRI	Magnetic Resonance Imaging (Μαγνητική Τομογραφία)
NAS	Neural Architecture Search (Αναζήτηση Νευρωνικής Αρχιτεκτονικής)
OVA	One-Versus-All (Ένα-Έναντι-Όλων)
POCUS	Point-of-Care Ultrasound (Υπερηχογράφημα στο Σημείο Φροντίδας)
ReLU	Rectified Linear Unit (Ανορθωμένη Γραμμική Μονάδα)
ROC	Receiver Operating Characteristic (Χαρακτηριστική Καμπύλη Λειτουργίας Δέκτη)
ROI	Region of Interest (Περιοχή Ενδιαφέροντος)
SGD	Stochastic Gradient Descent (Στοχαστική Κάθοδος Κλίσης)
TL	Transfer Learning (Μεταφορά Μάθησης)
TN	True Negatives (Αληθώς Αρνητικά)
TP	True Positives (Αληθώς Θετικά)
TPR	True Positive Rate (Ρυθμός Αληθώς Θετικών)
TTA	Test-Time Augmentation (Ενίσχυση Δεδομένων κατά την Πρόβλεψη)
ViT	Vision Transformer (Οπτικός Μετασχηματιστής)

Απόδοση Ξενόγλωσσων Όρων

Acoustic Shadowing	Ακουστική Σκίαση
Attention Mechanism	Μηχανισμός Προσοχής
Average Pooling	Μέση Ομαδοποίηση
Background Enhancement	Ενίσχυση Υποβάθρου
Baseline	Γραμμή Βάσης
Benchmark	Σημείο αναφοράς / Συγκριτική δοκιμή
Benign	Καλοήθης
Bias	Μεροληψία / Σφάλμα
Binary Cross-Entropy	Δυαδική Διασταυρούμενη Εντροπία
Border zone	Οριακή ζώνη / Άλως
Bounding Box	Οριακό Πλαίσιο
Channel combining	Συνδυασμός καναλιών
Class Imbalance	Ανισορροπία Κλάσεων
Co-adaptation	Συν-προσαρμογή
Color jittering	Διακύμανση χρώματος
Compound Scaling	Σύνθετη Κλιμάκωση
Confusion Matrix	Πίνακας Σύγχυσης
Convolutional Layer	Επίπεδο Συνέλιξης
Convolutional Neural Network	Συνελικτικό Νευρωνικό Δίκτυο

Core needle biopsy	Βιοψία με κόπτουσα βελόνα
Cropping	Αποκοπή (Περικοπή)
Cross-Validation	Διασταυρούμενη Επικύρωση
Data Augmentation	Ενίσχυση Δεδομένων
Data Leakage	Διαρροή Δεδομένων
De-novo training	Εκπαίδευση από το μηδέν
Debugging	Αποσφαλμάτωση
Decision Threshold	Κατώφλι Απόφασης
Deep Learning	Βαθιά Μάθηση
Depthwise Separable Convolutions	Διαχωρίσιμες Συνελίξεις κατά Βάθος
Descriptors	Περιγραφικά στοιχεία
Dilation	Διαστολή
Downsampling	Υποδειγματοληψία
Dropout	Απόρριψη (Επίπεδο)
Early Stopping	Πρόωρη Διακοπή
Ensemble Learning	Συνδυαστική Μάθηση
False Negatives	Ψευδώς Αρνητικά
False Positives	Ψευδώς Θετικά
Feature Extraction	Εξαγωγή Χαρακτηριστικών
Feature MapΧάρτης	Χαρακτηριστικών
Fine-Tuning	Μικρο-συντονισμός
Folds	Πτυχές (Επικύρωσης)
Fully Connected Layer	Επίπεδο Πλήρους Σύνδεσης
Gold standard	Χρυσός κανόνας / Πρότυπο
Grid-like topology	Τοπολογία πλέγματος
High-level features	Χαρακτηριστικά υψηλού επιπέδου

Hold-out validation set	Δεσμευμένο σύνολο επικύρωσης
Imbalanced datasets	Μη ισορροπημένα σύνολα δεδομένων
Inference	Εξαγωγή Συμπερασμάτων
Inter-observer Variability	Μεταβλητότητα μεταξύ παρατηρητών
Inverted residuals	Ανεστραμμένα υπολειμματικά
Layer Freezing	Πάγωμα Επιπέδων
Lightweight Models / CNNs	Ελαφριά Μοντέλα
Linear bottlenecks	Γραμμικοί στενωποί
Low-level features	Χαρακτηριστικά χαμηλού επιπέδου
Machine Learning	Μηχανική Μάθηση
Macrolobulations	Μακρολοβώσεις
Majority voting	Πλειοψηφική ψηφοφορία
Malignant	Κακοήθης
Max Pooling	Μέγιστη Ομαδοποίηση
Microlobulated	Μικρολοβωτά
Noise injection	Εισαγωγή θορύβου
Overfitting	Υπερπροσαρμογή
Padding	Συμπλήρωση
Patient Independence	Ανεξαρτησία Ασθενών
Peritumoral	Περι-ογκικό (Περιτυμοριακό)
Pooling Layer	Επίπεδο Ομαδοποίησης
Precision	Ακρίβεια Πρόβλεψης
Probe	Ηχοβολέας
Recall / Sensitivity	Ανάκληση / Ευαισθησία
Receptive Field	Δεκτικό Πεδίο
Region of Interest	Περιοχή Ενδιαφέροντος

Regularization	Κανονικοποίηση
Resizing	Αλλαγή Μεγέθους
Robustness	Ανθεκτικότητα / Ευρωστία
Segmentation	Τμηματοποίηση
Self-attention	Αυτο-προσοχή
Spatial Attention	Χωρική Προσοχή
Spatial filtering	Χωρικό φιλτράρισμα
Specificity	Ειδικότητα
Speckle Noise	Ακουστικός Θόρυβος
Spiculated	Ακτινωτά
State-of-the-art	Τεχνολογία αιχμής
Stride	Βήμα (Συνέλιξης)
Target task	Εργασία-στόχος
Threshold Tuning / Optimization	Βελτιστοποίηση Κατωφλίου
Tissue masking	Συγκάλυψη ιστού
Trade-off	Αντιστάθμισμα / Συμβιβασμός
Transfer Learning	Μεταφορά Μάθησης
True Negatives	Αληθώς Αρνητικά
True Positives	Αληθώς Θετικά
Validation Set	Σύνολο Επικύρωσης
Vanishing gradient problem	Πρόβλημα εξαφάνισης της κλίσης
Weight Sharing	Κοινή Χρήση Βαρών