



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

**Ανάπτυξη Ερμηνεύσιμων Μοντέλων Βαθιάς Μάθησης για
την Ταξινόμηση Ιστοπαθολογικών Εικόνων Καρκίνου του
Γαστρεντερικού Συστήματος**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Αικατερίνη Ε. Μαρτάκου Γαλιατσάτου

Επιβλέπων : Κωνσταντίνα Νικήτα
Καθηγήτρια Ε.Μ.Π.

Αθήνα, Ιούνιος 2025



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

**Ανάπτυξη Ερμηνεύσιμων Μοντέλων Βαθιάς Μάθησης για
την Ταξινόμηση Ιστοπαθολογικών Εικόνων Καρκίνου του
Γαστρεντερικού Συστήματος**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Αικατερίνη Ε. Μαρτάκου Γαλιατσάτου

Επιβλέπων : Κωνσταντίνα Νικήτα
Καθηγήτρια Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 4η Ιουλίου 2025.

.....
Κωνσταντίνα Νικήτα
Καθηγήτρια Ε.Μ.Π

.....
Γεώργιος Στάμου
Καθηγητής Ε.Μ.Π.

.....
Αθανάσιος Βουλόδημος
Επίκουρος Καθηγητής Ε.Μ.Π

Αθήνα, Ιούνιος 2025

.....
Αικατερίνη Ε. Μαρτάκου Γαλιατσάτου

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

Copyright © Αικατερίνη Μαρτάκου Γαλιατσάτου, 2025

Με επιφύλαξη παντός δικαιώματος. All rights reserved.

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

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

Περίληψη

Ο καρκίνος αποτελεί μία από τις κυριότερες αιτίες θανάτου παγκοσμίως, με εκατομμύρια νέα περιστατικά να διαγιγνώσκονται κάθε έτος. Η έγκαιρη και ακριβής διάγνωση είναι κρίσιμη για την επιτυχή θεραπευτική αντιμετώπιση και την πρόγνωση των ασθενών. Η ιστοπαθολογική εξέταση, η οποία βασίζεται στη μικροσκοπική ανάλυση δειγμάτων ιστών που έχουν υποστεί χρώση, θεωρείται το βέλτιστο πρότυπο (gold standard) για τη διάγνωση πλήθους μορφών καρκίνου.

Ωστόσο, η παρατήρηση και η ανάλυση των εικόνων που προκύπτουν από την εξέταση αυτή, παρουσιάζουν σημαντικές προκλήσεις. Η διαδικασία είναι εξαιρετικά χρονοβόρα, απαιτεί υψηλή εξειδίκευση, ενώ επηρεάζεται από την υποκειμενικότητα και τη διαγνωστική μεταβλητότητα μεταξύ παθολογοανατόμων.

Η εξέλιξη της Ψηφιακής Παθολογίας (Digital Pathology), δηλαδή η σάρωση και αποθήκευση ιστολογικών δειγμάτων σε ψηφιακή μορφή υψηλής ανάλυσης, επιτρέπει την αξιοποίηση τεχνικών τεχνητής νοημοσύνης για την ερμηνεία αυτών των εικόνων. Ιδιαίτερα, τα Συνελκτικά Νευρωνικά Δίκτυα (Convolutional Neural Networks - CNNs) έχουν αναδειχθεί ως ένα ισχυρό εργαλείο για την αναγνώριση μορφολογικών χαρακτηριστικών σε ιστοπαθολογικές εικόνες. Πλέον, σε συγκεκριμένες διεργασίες όπως η ταξινόμηση ιστών, τα CNNs επιτυγχάνουν ακρίβεια συγκρίσιμη με αυτή ειδικών ιατρών.

Στο πλαίσιο αυτής της διπλωματικής εργασίας, αναπτύσσεται και αξιολογείται ένα μοντέλο βαθιάς μάθησης για την ταξινόμηση ιστοπαθολογικών εικόνων σε καλοήθειες και κακοήθειες κατηγορίες. Το μοντέλο βασίζεται σε αρχιτεκτονική συνελκτικού νευρωνικού δικτύου και εκπαιδεύεται, αρχικά, στο σύνολο δεδομένων GasHisSDB, το οποίο περιλαμβάνει ιστοπαθολογικές εικόνες γαστρικού καρκίνου. Εμπλουτίζεται με το εργαλείο ερμηνευσιμότητας Grad-CAM, το οποίο παράγει θερμικούς χάρτες που αναδεικνύουν τις περιοχές της εικόνας που επηρέασαν την τελική πρόβλεψη. Παράλληλα, εφαρμόζεται η μέθοδος Monte Carlo Dropout για την αποτίμηση της αβεβαιότητας των προβλέψεων του μοντέλου. Η προτεινόμενη προσέγγιση συγκρίνεται με ήδη υπάρχοντα state-of-the-art μοντέλα και επιτυγχάνει ακρίβεια 94% στο σύνολο αξιολόγησης. Σε δεύτερο στάδιο, το μοντέλο επανεκπαιδεύεται στο σύνολο δεδομένων MHIST που αποτελείται από ιστοπαθολογικές εικόνες καρκίνου του παχέος εντέρου, με σκοπό να αξιολογηθεί η γενικευσιμότητά του.

Λέξεις-Κλειδιά

Βαθιά Μάθηση, Συνελκτικά Νευρωνικά Δίκτυα, Γαστρικός Καρκίνος, Ιστοπαθολογία, Ταξινόμηση Εικόνων, Grad-CAM, Monte Carlo Dropout

Abstract

Cancer is one of the leading causes of death worldwide, with millions of new cases diagnosed each year. Early and accurate diagnosis is critical for effective treatment and patient prognosis. Histopathological examination, which involves the microscopic analysis of stained tissue samples, is considered the gold standard for diagnosing many forms of cancer.

However, the observation and analysis of the resulting images from such examinations present significant challenges. The process is extremely time-consuming, requires a high level of expertise, and is affected by subjectivity and diagnostic variability among pathologists.

The advancement of Digital Pathology, meaning the scanning and storage of histological samples in high-resolution digital format, enables the application of artificial intelligence techniques for interpreting these images. In particular, Convolutional Neural Networks (CNNs) have emerged as a powerful tool for identifying morphological features in histopathological images. In specific tasks such as tissue classification, CNNs now achieve accuracy comparable to that of medical experts.

Within the framework of this thesis, a deep learning model is developed and evaluated for classifying histopathological images into benign and malignant categories. The model is based on a CNN architecture and is trained on the GasHisSDB dataset, which contains histopathological images of gastric cancer. It is enhanced with the interpretability tool Grad-CAM, which generates heatmaps highlighting the regions of the image that influenced the final prediction. Additionally, the Monte Carlo Dropout method is applied to estimate the uncertainty of the model's predictions. The proposed approach is compared with existing state-of-the-art models and achieves an accuracy of 94% on the evaluation set. In a second stage, the model is retrained on the MHIST dataset, which consists of histopathological images of colorectal cancer, in order to evaluate its ability to generalize across datasets.

Keywords

Deep Learning, Convolutional Neural Networks, Gastric Cancer, Histopathology, Image Classification, Grad-CAM, Monte Carlo Dropout

Ευχαριστίες

Θα ήθελα να εκφράσω τις θερμές μου ευχαριστίες στην Καθηγήτρια του Εθνικού Μετσόβιου Πολυτεχνείου, κα Κωνσταντίνα Νικήτα, για την ανάθεση και την επίβλεψη της παρούσας διπλωματικής εργασίας, καθώς και για την πολύτιμη ευκαιρία που μου έδωσε να ασχοληθώ με ένα αντικείμενο σε έναν τόσο ραγδαία αναπτυσσόμενο τομέα.

Επίσης, ευχαριστώ ιδιαίτερα τη διδάκτορα Μαρία Αθανασίου για την καθοδήγησή της, την υποστήριξή της, καθώς και για τις ουσιαστικές συμβουλές της κατά την πρώτη μου ενασχόληση με τις τεχνικές που εφαρμόστηκαν στο πλαίσιο της εργασίας. Θα ήθελα ακόμη να ευχαριστήσω τα λοιπά μέλη της τριμελούς εξεταστικής επιτροπής, κ. Γεώργιο Στάμου και κ. Αθανάσιο Βουλόδημο, για τον χρόνο και την προσοχή που αφιέρωσαν στην αξιολόγηση της εργασίας μου.

Θα ήθελα ακόμα να ευχαριστήσω θερμά τον Θεοφάνη Γαντιδίδη για την πολύτιμη συμβολή του στη επίλυση προβλημάτων σύνδεσης με τον server του εργαστηρίου, καθώς και την θεία μου, Αγγελική Καζνέση, φιλόλογο, για την υποστήριξη και την επιμέλειά της κατά την αποτύπωση του κειμένου της παρούσας εργασίας.

Τέλος, ένα βαθύ ευχαριστώ στους γονείς μου και τον αδερφό μου για την στήριξη και την κατανόησή τους όλα αυτά τα χρόνια.

Περιεχόμενα

Περίληψη.....	5
Abstract	7
Ευχαριστίες	9
Περιεχόμενα	11
1. Εισαγωγή	19
2. Καρκίνος	20
2.1.Ορισμός.....	20
2.1.1.Θεραπευτικές προσεγγίσεις	21
2.1.2.Επιπλοκές.....	22
2.2.Γαστρικός καρκίνος (καρκίνος στομάχου)	23
2.2.1.Στομάχι.....	24
2.2.2.Ανάπτυξη και τύποι Γαστρικού Καρκίνου	25
2.2.3.Συμπτώματα και θεραπεία	28
2.2.4.Ιστοπαθολογία.....	28
2.3.Καρκίνος του παχέος εντέρου	29
2.3.1.Παχύ έντερο	29
2.3.2.Ανάπτυξη καρκίνου του παχέος εντέρου	30
2.3.3.Θεραπεία και συμπτώματα	33
2.3.4.Ιστοπαθολογία.....	33
2.4.Ιατρική απεικόνιση καρκίνου.....	34
2.4.1.Ακτινολογικές Μέθοδοι - X-ray και Υπολογιστική Τομογραφία (CT)	34
2.4.2.Μαγνητική Τομογραφία (MRI).....	35
2.4.3.Τομογραφία Εκπομπής Ποζιτρονίων – PET	35
2.4.4.Υπερηχογράφημα (Ultrasound)	36
2.4.5.Ενδοσκόπηση και Ενδοσκοπικός Υπέρηχος (EUS).....	37
2.4.6.Ιστοπαθολογικές εικόνες.....	37
3. Μηχανική Μάθηση	40
3.1.Βαθιά μάθηση	40

3.1.1.Επιβλεπόμενη μάθηση	42
3.1.2.Μη επιβλεπόμενη μάθηση	42
3.1.3.Ενισχυτική μάθηση	43
3.2.Τεχνητά Νευρωνικά Δίκτυα	44
3.2.1.Νευρωνικά δίκτυα.....	44
3.2.2.Τεχνητά Νευρωνικά Δίκτυα (ΤΝΔ)	44
3.2.3.Βιολογική έμπνευση	44
3.2.4.Δομή ΤΝΔ και Επίπεδα.....	45
3.3.Συνελκτικά Νευρωνικά Δίκτυα.....	47
3.3.1.Συνελκτικό Επίπεδο (Convolutional Layer)	48
3.3.2.Συνάρτηση Ενεργοποίησης (Activation Function).....	49
3.3.3.Επίπεδο pooling (Pooling Layer).....	49
3.3.4.Πλήρως Συνδεδεμένο Επίπεδο (Fully Connected Layer-FC).....	50
3.4.Προεκπαιδευμένα Συνελκτικά Νευρωνικά Δίκτυα (Pretrained CNN)	51
3.4.1.VGG	52
3.4.2.ResNet.....	53
3.4.3.Inception (GoogLeNet).....	55
3.4.4.EfficientNet.....	57
3.5.Ερμηνευσιμότητα.....	58
3.5.1.Χάρτες Σαφήνειας (Saliency Maps)	58
3.5.2.Ενσωματωμένες Κλίσεις (Integrated Gradients)	59
3.5.3.Grad-CAM	59
3.6.Αβεβαιότητα	62
3.6.1.Bayesian Νευρωνικά Δίκτυα (Bayesian Neural Networks - BNNs).....	62
3.6.2.Μοντέλα Ensemble (Ensemble Models).....	62
3.6.3.Monte Carlo Dropout.....	62
4. Σύγχρονες Μέθοδοι Ταξινόμησης Ιστοπαθολογικών Εικόνων Καρκίνου του Γαστρεντερικού Συστήματος.....	65
4.1.Γαστρικός Καρκίνος.....	65
4.1.1.Προσαρμοσμένα Συνελκτικά Νευρωνικά Δίκτυα (Custom CNNs)	65
4.1.2.Προεκπαιδευμένα CNNs.....	66

4.1.3.Μηχανισμοί Προσοχής (Attention Mechanisms)	66
4.1.4.Οπτικοί Μετασχηματιστές	67
4.2.Καρκίνος του Παχέος Εντέρου	67
4.2.1.Προεκπαιδευμένα CNNs.....	67
4.2.2.Στρατηγικές Stacking.....	68
4.2.3.Μετασχηματιστές.....	68
5. Μεθοδολογία.....	69
5.1.Σύνολα Δεδομένων - Datasets.....	69
5.1.1.GasHisSDB	69
5.1.2.MHIST	73
5.1.3.Προεπεξεργασία Δεδομένων.....	74
5.2.Ανάπτυξη Μοντέλων Βαθιάς Μάθησης.....	75
5.2.1.Προσέγγιση Μεταφοράς Μάθησης.....	75
5.2.2.Ανάπτυξη Προσαρμοσμένων Αρχιτεκτονικών CNN.....	76
5.3.Εκπαίδευση και Αξιολόγηση	79
5.3.1.Stratified k-fold.....	79
5.3.2.Μετρικές Αξιολόγησης	81
5.4.Ερμηνευσιμότητα και οπτικοποίηση.....	83
5.4.1.Grad-CAM	83
5.4.2.HistomicsTK.....	83
5.5.Αβεβαιότητα	86
6. Αποτελέσματα.....	87
6.1.Αξιολόγηση Διακριτικής Ικανότητας.....	87
6.1.1.GasHisSDB	87
6.1.2.MHIST	88
6.1.3.Σύγκριση Επίδοσης με Μοντέλα στη Βιβλιογραφία	89
6.2.Ποσοτικοποίηση Αβεβαιότητας μοντέλου στο GasHisSDB	90
6.3.Αποτελέσματα Grad-CAM στο GasHisSDB	92
6.3.1.Αληθώς Θετικά	93
6.3.2.Αληθώς Αρνητικά	94

6.3.3.Ψευδώς Θετικά.....	95
6.3.4.Ψευδώς Αρνητικά.....	96
6.4.Εξαγωγή χαρακτηριστικών Αληθώς Θετικών Προβλέψεων με HistomicsTK	97
6.4.1.Απεικόνιση Ενεργοποίησης Πυρήνων	98
6.4.2.Εξαγωγή Μορφολογικών, Γεωμετρικών Χαρακτηριστικών και Χαρακτηριστικών Υφής και Έντασης	99
7. Συμπεράσματα και Μελλοντική Εργασία.....	105
7.1.Σύνοψη	105
7.2.Συμπεράσματα	105
7.3.Μελλοντική Εργασία	106
Βιβλιογραφία.....	107

Ευρετήριο Σχημάτων

Σχήμα 2.1 Ακτινοθεραπεία	22
Σχήμα 2.2 Γαστρικός Καρκίνος	23
Σχήμα 2.3 Ανατομία στομάχου	24
Σχήμα 2.4 Στρώσεις στομάχου	25
Σχήμα 2.5 Αδενοκαρκίνωμα σε σύγκριση με υγιή ιστό στομάχου	26
Σχήμα 2.6 Στάδια Γαστρικού Καρκίνου	27
Σχήμα 2.7 Καρκίνος Παχέος Εντέρου	29
Σχήμα 2.8 Ανατομία εντέρου	30
Σχήμα 2.9 Πολύποδες εντέρου	30
Σχήμα 2.10 Υγιής ιστός σε σύγκριση με ύπαρξη υπερπλαστικών πολυπόδων	31
Σχήμα 2.11 Ιστολογικά χαρακτηριστικά Ακτινωτών Αδενωμάτων (SSA)	32
Σχήμα 2.12 Αξονική Τομογραφία Στομάχου	34
Σχήμα 2.13 Μαγνητική Τομογραφία Εγκεφάλου	35
Σχήμα 2.14 PET-scan	36
Σχήμα 2.15 Ενδοσκοπικός Υπέρηχος Στομάχου	37
Σχήμα 2.16 Ιστοπαθολογικές Εικόνες Καρκινικών Ιστών	39
Σχήμα 2.17 Ιστοπαθολογικές Εικόνες Υγιούς Ιστού Στομάχου	39
Σχήμα 3.1 Ροή Ανάπτυξης Μοντέλου Μηχανικής Μάθησης	40
Σχήμα 3.2 Σύγκριση Προσέγγισης Μηχανικής και Βαθιάς Μάθησης	41
Σχήμα 3.3 Επιβλεπόμενη Μάθηση	42
Σχήμα 3.4 Μη Επιβλεπόμενη Μάθηση	43
Σχήμα 3.5 Ενισχυτική Μάθηση	44
Σχήμα 3.6 Βιολογικός Νευρώνας	45
Σχήμα 3.7 Τεχνητός Νευρώνας	46
Σχήμα 3.8 Τεχνητό Νευρωνικό Δίκτυο	47
Σχήμα 3.9 Βαθύ Νευρωνικό Δίκτυο	47
Σχήμα 3.10 Ακολουθία Βημάτων για Τυπική Ταξινόμηση Εικόνων	48
Σχήμα 3.11 Συνελκτικό Επίπεδο	48

Σχήμα 3.12 Επίπεδο Pooling	50
Σχήμα 3.13 Πλήρως Συνδεδεμένο Επίπεδο	50
Σχήμα 3.14 Αρχιτεκτονική Συνελκτικού Νευρωνικού Δικτύου	51
Σχήμα 3.15 Αρχιτεκτονική VGG	53
Σχήμα 3.16 Αρχιτεκτονική ResNet	55
Σχήμα 3.17 Αρχιτεκτονική Inception	56
Σχήμα 3.18 Αρχιτεκτονική EfficientNet	57
Σχήμα 3.19 Διάγραμμα Ροής Μεθόδου Grad-CAM	61
Σχήμα 3.20 Monte-Carlo Dropout	63
Σχήμα 5.1 Διαδικασία Δημιουργίας Sub-Datasets	70
Σχήμα 5.2 Παραδείγματα Εικόνων GasHisSDB	70
Σχήμα 5.3 Sub-Datasets GasHisSDB	71
Σχήμα 5.4 Χρωματική Διαφορά	72
Σχήμα 5.5 Παραδείγματα Εικόνων MHIST	73
Σχήμα 5.6 Αρχιτεκτονική Δικτύου EfficientNetB0	76
Σχήμα 5.7 Βασική Αρχιτεκτονική Προσαρμοσμένου CNN	76
Σχήμα 5.7 5-Fold Cross-Validation	80
Σχήμα 6.6 Ιστόγραμμα Αβεβαιότητας	92
Σχήμα 6.1 Abnormal εικόνες με τα αντίστοιχα heatmaps και overlays	94
Σχήμα 6.2 Normal εικόνες με τα αντίστοιχα heatmaps και overlays	95
Σχήμα 6.3 Normal εικόνες με τα αντίστοιχα heatmaps και overlay	96
Σχήμα 6.5 Abnormal εικόνες, τμηματοποιημένοι πυρήνες, heatmaps	99

Ευρετήριο Πινάκων

Πίνακας 6.1 Αποτελέσματα μετρικών αξιολόγησης του προσαρμοσμένου CNN μοντέλου στο σύνολο δεδομένων GasHisDB, 5Fold	87
Πίνακας 6.2 Αποτελέσματα μετρικών αξιολόγησης του προσαρμοσμένου CNN μοντέλου στο σύνολο δεδομένων GasHisDB, 90-10	87
Πίνακας 6.3 Αποτελέσματα μετρικών αξιολόγησης του προσαρμοσμένου CNN μοντέλου στο σύνολο δεδομένων MHIST	88
Πίνακας 6.4 Συγκριτική αξιολόγηση του προσαρμοσμένου CNN μοντέλου και state-of-the-art μοντέλων για Γαστρικό Καρκίνο	89
Πίνακας 6.5 Συγκριτική αξιολόγηση του προσαρμοσμένου CNN μοντέλου και state-of-the-art μοντέλων για Καρκίνο του Παχέος Εντέρου	90
Πίνακας 6.6 Τιμές αβεβαιότητας για μικρό αριθμό δειγμάτων	91
Πίνακας 6.7 Μετρικές αξιολόγησης του προσαρμοσμένου CNN μοντέλου στο πλαίσιο της ποσοτικοποίησης αβεβαιότητας με MC Dropout	91
Πίνακας 6.8 Αριθμών συνολικών και ενεργών πυρήνων	98
Πίνακας 6.9 FSD χαρακτηριστικά με διαφορά τιμών μεταξύ ενεργών και ανενεργών πυρήνων	100
Πίνακας 6.10 Μορφολογικά χαρακτηριστικά με διαφορά τιμών μεταξύ ενεργών και ανενεργών πυρήνων	101
Πίνακας 6.11 Χαρακτηριστικά διαβάθμισης με διαφορά τιμών μεταξύ ενεργών και ανενεργών πυρήνων	102
Πίνακας 6.12 Χαρακτηριστικά Haralick με διαφορά τιμών μεταξύ ενεργών και ανενεργών πυρήνων	103

1. Εισαγωγή

Η ιστοπαθολογία είναι η διάγνωση και μελέτη των νόσων των ιστών και περιλαμβάνει τη μικροσκοπική εξέταση ιστών και κυττάρων. Η διαχείριση και η ταξινόμηση εικόνων στην ιστοπαθολογία αποτελεί μια σημαντική διαδικασία στη διάγνωση και πρόγνωση ασθενειών, όπως ο καρκίνος. Η διαδικασία βασίζεται στην ανάλυση ψηφιακών μικροσκοπικών εικόνων ιστών.

Η ταξινόμηση τους πλέον πραγματοποιείται όλο και περισσότερο μέσω τεχνικών μηχανικής και βαθιάς μάθησης, όπως τα συνελκτικά νευρωνικά δίκτυα (CNNs). Τα μοντέλα εκπαιδεύονται να αναγνωρίζουν μορφολογικά χαρακτηριστικά και πρότυπα, τα οποία σχετίζονται με φυσιολογικές ή παθολογικές καταστάσεις. Αυτή η προσέγγιση μειώνει την υποκειμενικότητα της ανθρώπινης αξιολόγησης και επιταχύνει τη διαδικασία διάγνωσης. Οι λεπτές διαφορές σε κυτταρικά ή ιστολογικά χαρακτηριστικά, που ενδέχεται να διαφύγουν της ανθρώπινης παρατήρησης, εντοπίζονται ευκολότερα, συμβάλλοντας έτσι στην έγκαιρη ανίχνευση του καρκίνου.

Ο γαστρικός καρκίνος αποτελεί συχνό είδος κακοήθειας, κατά βάση ασυμπτωματικής, στοιχείο που την καθιστά σοβαρή απειλή για την παγκόσμια υγεία. Σε παγκόσμιο επίπεδο, ο γαστρικός είναι ο πέμπτος σε συχνότητα διαγνωσμένος τύπος καρκίνου και ευθύνεται για το 18,0% των συνολικών θανάτων από καρκίνο. Η ιστοπαθολογική εξέταση του γαστρικού καρκίνου περιλαμβάνει μικροσκοπική ανάλυση τομών, οι οποίες προέρχονται από ιστούς που έχουν ληφθεί από πιθανές καρκινικές περιοχές, και πραγματοποιείται από έμπειρους παθολόγους.

Η ιατρική απεικόνιση θεωρείται ένα από τα πιο σημαντικά στάδια της υγειονομικής περίθαλψης. Η επιτυχής ενσωμάτωση των τεχνικών Βαθιάς Μάθησης (DL) στην κλινική πρακτική έχει ακριβώς ως στόχο την επίτευξη διαγνωστικής ακρίβειας που να μην υπολείπεται αυτής των επαγγελματιών, διασφαλίζοντας ταυτόχρονα τη μείωση του χρόνου και του κόστους διάγνωσης.

Στην παρούσα διπλωματική εργασία, αναπτύχθηκε ένα μοντέλο βασισμένο στα Βαθιά Νευρωνικά Δίκτυα για τη ταξινόμηση εικόνων ιστοπαθολογίας, αφού προηγήθηκε έρευνα προϋπαρχόντων μοντέλων βαθιάς μάθησης για τη διαχείριση δεδομένων ιατρικής απεικόνισης καρκίνου. Ειδικότερα, σκοπός ήταν η εκπαίδευση ενός μοντέλου πάνω σε ένα συγκεκριμένο dataset γαστρικού καρκίνου, το GasHisSDB, το οποίο περιέχει εικόνες που ανήκουν σε δύο κλάσεις με απώτερο στόχο την κατά το δυνατό ακριβέστερη δυαδική ταξινόμηση.

2. Καρκίνος

Ο καρκίνος παραμένει η δεύτερη πιο συχνή αιτία θανάτου παγκοσμίως, με τη σταθερή αυξητική τάση τις τελευταίες δεκαετίες[1]. Το 2022 καταγράφηκαν περίπου 20 εκατομμύρια νέα περιστατικά καρκίνου και 9,7 εκατομμύρια θάνατοι που σχετίζονται με τον καρκίνο παγκοσμίως. Τα στοιχεία αυτά μεταφράζονται στο ότι περίπου ένας στους πέντε άνδρες ή γυναίκες θα αναπτύξει καρκίνο κατά τη διάρκεια της ζωής του/της, με τα ποσοστά θνησιμότητας να ανέρχονται σε έναν στους εννέα για τους άνδρες και μία στις δώδεκα για τις γυναίκες. Μεταξύ 2020 και 2050, οι καρκίνοι θα επιφέρουν κόστος 25,2 τρισεκατομμυρίων διεθνών δολαρίων στην παγκόσμια οικονομία, το οποίο ισοδυναμεί με ετήσια επιβάρυνση ύψους 0,55% επί του παγκόσμιου ακαθάριστου εγχώριου προϊόντος (ΑΕΠ).

Αυτά τα στατιστικά στοιχεία αντικατοπτρίζουν τις σοβαρές προκλήσεις που θέτει ο καρκίνος για τα συστήματα υγειονομικής περίθαλψης και τις κοινωνίες όχι μόνο λόγω των άμεσων επιπτώσεων του στην υγεία, αλλά και εξαιτίας των ευρύτερων κοινωνικοοικονομικών συνεπειών του.

Κάθε χρόνο, ο καρκίνος σκοτώνει περισσότερους ανθρώπους από ότι το AIDS, η ελονοσία και η φυματίωση μαζί[2]. Προβάλλεται συχνά ως «σύγχρονη μάστιγα» επειδή αφενός τα αίτια εμφάνισης του είναι ακόμη άγνωστα και αφετέρου επειδή συνδυάζει χαρακτηριστικά επιδημιών του παρελθόντος.

2.1.Ορισμός

Καρκίνος είναι το αποτέλεσμα γενετικών και επιγενετικών μεταβολών που οδηγούν σε δυσλειτουργία των μηχανισμών ελέγχου του κυτταρικού κύκλου, με αποτέλεσμα την ανεξέλεγκτη κυτταρική διαίρεση, την αποφυγή του προγραμματισμένου κυτταρικού θανάτου και την ικανότητα διήθησης και μετάστασης[3]. Τα πλεονάζοντα κύτταρα σχηματίζουν μάζες, που καλούνται όγκοι. Σε μερικές περιπτώσεις αυτά τα παθολογικά κύτταρα κάνουν μετάσταση, δηλαδή εξαπλώνονται και σε άλλα μέρη του σώματος δημιουργώντας δευτερεύοντες όγκους (μεταστατικούς όγκους) παρόμοιους με αυτούς του αρχικού καρκίνου.

Οι περισσότερες μορφές καρκίνου είναι στην ουσία όγκοι εκτός από ορισμένους τύπους καρκίνου όπως η λευχαιμία, των οποίων τα κύτταρα κυκλοφορούν μέσα στο αίμα και κατ' επέκταση στα διάφορα όργανα και ιστούς. Υπάρχουν πάνω από 200 είδη καρκίνου για καθένα από τα οποία υιοθετείται διαφορετική θεραπευτική προσέγγιση[4].

Το στάδιο της νόσου περιγράφει την έκταση και την εξάπλωση του καρκίνου κατά το χρόνο της διάγνωσης και καθορίζεται βάσει του μεγέθους του όγκου, της εμπλοκής ή μη των παραπλήσιων λεμφαδένων και της παρουσίας ή μη μεταστάσεων σε άλλα σημεία του σώματος. Από τη στιγμή που θα προσδιοριστούν αυτοί οι παράγοντες, ο καρκίνος ταξινομείται ως σταδίου I, II, III ή IV, με το στάδιο I να είναι το πιο πρώιμο και το IV το πιο προχωρημένο[2].

Ο καρκίνος συνιστά ιάσιμη νόσο, εάν αυτή διαγνωσθεί έγκαιρα. Σύμφωνα με τον

Παγκόσμιο Οργανισμό Υγείας (WHO), εκτιμάται ότι ένα ποσοστό άνω του 30% των καρκίνων μπορεί να προληφθεί, ενώ η έγκαιρη διάγνωση διασφαλίζει τη θεραπεία σε ποσοστό άνω του 40%. Συνήθως προσβάλλει ανθρώπους μεγάλης ηλικίας, υπάρχουν όμως και μορφές καρκίνου που εμφανίζονται σε νεαρής ηλικίας άτομα, ακόμη και σε παιδιά.

2.1.1.Θεραπευτικές προσεγγίσεις

Η θεραπεία του καρκίνου εξαρτάται από τον τύπο, το στάδιο, τη γενική κλινική εικόνα του ασθενούς και συχνά απαιτεί ένα συνδυασμό προσεγγίσεων για να είναι αποτελεσματική. Οι πιο συνήθεις μορφές θεραπείας είναι η χειρουργική επέμβαση, η χημειοθεραπεία και η ακτινοθεραπεία, καθώς χρησιμοποιούνται για την αντιμετώπιση πολλών ειδών καρκίνου και έχουν τεκμηριωμένα αποτελέσματα.

Η χειρουργική επέμβαση είναι η πιο συχνή και η πιο αποτελεσματική θεραπεία για καρκίνους που δεν έχουν εξαπλωθεί, στοχεύοντας στην αφαίρεση όλων των καρκινικών κυττάρων καθώς και μέρους του περιβάλλοντος υγιούς ιστού. Ωστόσο, η επιτυχία της εξαρτάται από τη δυνατότητα ασφαλούς αφαίρεσης ολόκληρου του όγκου. Για παράδειγμα, αν ο όγκος βρίσκεται μέσα σε ζωτικό όργανο, η χειρουργική επέμβαση μπορεί να μην είναι ασφαλής επιλογή. Επιπλέον, αυτή η θεραπεία δεν μπορεί να εξαλείψει καρκίνους του αίματος, όπως η λευχαιμία[5].

Η ακτινοθεραπεία χρησιμοποιεί υψηλές δόσεις ακτινοβολίας για να καταστρέψει καρκινικούς όγκους, προκαλώντας βλάβη στα καρκινικά κύτταρα και οδηγώντας τα σε αποδόμηση και θάνατο. Οι τύποι ακτινοθεραπείας περιλαμβάνουν και απεικονίζονται στην εικόνα 1:

- Εξωτερική ακτινοβολία (External Beam Radiation Therapy - EBRT): Χρησιμοποιείται εξωτερική συσκευή για να κατευθύνει την ακτινοβολία στον όγκο.
- Εσωτερική ακτινοβολία (brachytherapy): Τοποθετούνται ραδιενεργά εμφυτεύματα πάνω ή κοντά στον όγκο. Η ακτινοβολία μπορεί επίσης να εφαρμοστεί και κατά τη διάρκεια της χειρουργικής επέμβασης.
- Θερμική θεραπεία (Hyperthermia): Χρησιμοποιείται θερμότητα για να καταστρέψει τα καρκινικά κύτταρα και να συρρικνώσει τους όγκους. Μπορεί επίσης να ενισχύσει την αποτελεσματικότητα άλλων θεραπειών.

Η ακτινοθεραπεία μπορεί να χρησιμοποιηθεί μόνη της ή σε συνδυασμό με άλλες θεραπείες, όπως η χημειοθεραπεία. Συχνά εφαρμόζεται μετά τη χειρουργική επέμβαση για να εξαλειφθούν υπολειπόμενα καρκινικά κύτταρα.

Η χημειοθεραπεία χρησιμοποιεί φάρμακα για να σκοτώσει τα καρκινικά κύτταρα. Τα φάρμακα αυτά δρουν μέσω της κυκλοφορίας του αίματος καταστρέφοντας ταχύτατα αναπτυσσόμενα καρκινικά κύτταρα κατά την πορεία τους. Η χημειοθεραπεία είναι μία από τις πιο συχνές και αποτελεσματικές θεραπείες για καρκίνους που έχουν κάνει μετάσταση ή έχουν υποτροπιάσει. Μπορεί να χορηγηθεί ως μονοθεραπεία (με ένα μόνο φάρμακο) ή να απαιτείται συνδυασμός διαφορετικών φαρμάκων[6].

Η ανοσοθεραπεία και η στοχευμένη θεραπεία είναι νεότερες μέθοδοι με σημαντική πρόοδο τα τελευταία χρόνια, ιδιαίτερα για συγκεκριμένους τύπους καρκίνου, όπως το

μελάνωμα ή ο καρκίνος του πνεύμονα, επειδή εκμεταλλεύονται τον ίδιο τον οργανισμό για να καταπολεμήσουν τον όγκο με λιγότερες παρενέργειες. Κάθε θεραπεία επιλέγεται με βάση το μοριακό και ιστολογικό προφίλ του όγκου, τις πιθανότητες ίασης ή ύφεσης και την ποιότητα ζωής του ασθενούς.

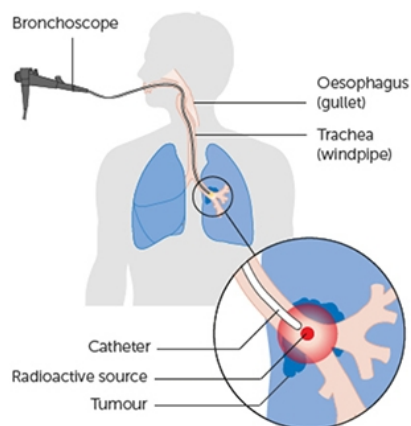
Μερικές επιπλέον θεραπείες που δεν χρησιμοποιούνται ευρέως είναι οι εξής:

- Ορμονοθεραπεία
- Αφαίρεση όγκου με θερμική ή χημική κατάλυση (ablation)
- Μεταμόσχευση μυελού των οστών

HOW DOES RADIOTHERAPY WORK?



EXTERNAL BEAM RADIOTHERAPY



BRACHY THERAPY

Σχήμα 2.1 Ακτινοθεραπεία [7]

2.1.2.Επιπλοκές

Ο καρκίνος και οι θεραπείες του μπορούν να προκαλέσουν διάφορες επιπλοκές που επηρεάζουν τη ζωή των ασθενών, όπως πόνο, κόπωση, δυσκολία στην αναπνοή, ναυτία καθώς και δυσλειτουργία του εντέρου προκαλώντας διάρροια ή δυσκοιλιότητα. Επιπλέον, η απώλεια βάρους είναι σύνηθες σύμπτωμα, καθώς τα καρκινικά κύτταρα καταναλώνουν τα θρεπτικά συστατικά του σώματος. Η τεχνητή διατροφή έχει περιορισμένη αποτελεσματικότητα στην ανάκτηση βάρους.

Επίσης, μπορεί να προκύψουν χημικές διαταραχές του οργανισμού, που επηρεάζουν την ισορροπία του και προκαλούν συμπτώματα όπως δίψα, σύγχυση και συχνή ούρηση.

Στην περίπτωση που ο καρκίνος επηρεάσει το νευρικό σύστημα, μπορεί να προκαλέσει αδυναμία, πονοκεφάλους ή νευρολογικά συμπτώματα.

Σε σπάνιες περιπτώσεις, το ανοσοποιητικό σύστημα αντιδρά υπερβολικά, προσβάλλοντας υγιή κύτταρα. Αυτά τα παρανεοπλασματικά σύνδρομα μπορεί να εκδηλωθούν με σπασμούς ή διαταραχές στη βάδιση. Επιπλέον, σε προχωρημένα στάδια, ο καρκίνος μπορεί να κάνει μετάσταση σε άλλα σημεία και όργανα του σώματος. Τέλος, υπάρχει ο κίνδυνος υποτροπής μετά τη θεραπεία, γεγονός που καθιστά απαραίτητη τη συστηματική ιατρική παρακολούθηση με εξετάσεις και απεικονίσεις.

2.2.Γαστρικός καρκίνος (καρκίνος στομάχου)

Ο καρκίνος του στομάχου είναι μια ασθένεια στην οποία καρκινικά κύτταρα, συνήθως του βλεννογόνου (90-95%) πολλαπλασιάζονται με μη φυσιολογικούς ρυθμούς[8]. Μπορεί να αναπτυχθεί σε οποιοδήποτε μέρος του στομάχου και να εξαπλωθεί σε άλλα όργανα μέσω:

- της ανάπτυξης του όγκου και του μεγέθους του στο τοίχωμα του στομάχου
- της κυκλοφορίας του αίματος
- του λεμφικού συστήματος

Καθώς αναπτύσσεται, εισβάλλει βαθύτερα στο τοίχωμα του στομάχου έως και τον μυϊκό ή ορογόνο χιτώνα. Συνήθως, οι καρκινικές αλλοιώσεις υφίστανται για αρκετά χρόνια προτού αναπτυχθούν σε κακοήθεια.

Ο γαστρικός καρκίνος είναι ένας από τους συχνότερους τύπους καρκίνου σε παγκόσμια κλίμακα. Με βάση τα επιδημιολογικά δεδομένα, είναι πιο συχνός σε Ανατολική Ευρώπη, Λατινική Αμερική και Άπω Ανατολή. Σημειώνεται, επίσης, ότι ο συγκεκριμένος τύπος πλήττει περισσότερο τον ανδρικό πληθυσμό και η συχνότητα είναι ακόμα μεγαλύτερη μετά τα 65 έτη. Παρά ταύτα, τα τελευταία χρόνια, η θνησιμότητα μειώνεται κατά πολύ, καθώς πρόκειται για μία από τις εύκολα θεραπεύσιμες μορφές καρκίνου.



Σχήμα 2.2 Γαστρικός Καρκίνος
[9]

2.2.1.Στομάχι

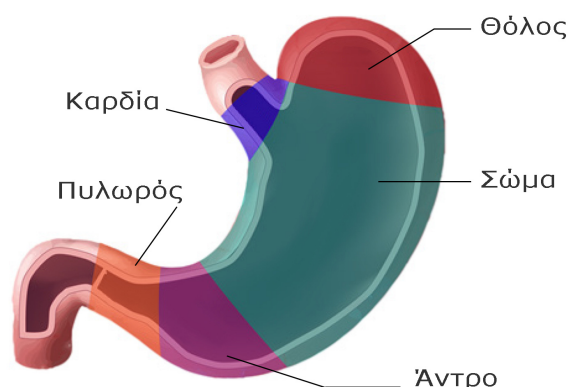
Η κατανόηση του γαστρικού καρκίνου απαιτεί καλή γνώση της φυσιολογικής δομής και της λειτουργίας του στομάχου.

Το στομάχι αποτελείται από 5 μέρη. Τα πρώτα 3 μέρη συνιστούν το εγγύς (πάνω) στομάχι:

- Καρδία: είναι το πρώτο σημείο, κοντά στον οισοφάγο.
- Θόλος (fundus): το άνω μέρος του στομαχιού, δίπλα στην καρδία.
- Σώμα (corpus): το κύριο, μεσαίο τμήμα του στομαχιού.

Τα δύο χαμηλότερα μέρη αποτελούν το άπω (κάτω) στομάχι:

- Άντρο: το χαμηλό σημείο του στομαχιού (κοντά στο λεπτό έντερο), όπου η τροφή αναμειγνύεται με τον γαστρικό χυμό.
- Πυλωρός: το τελευταίο κομμάτι του στομαχιού, που λειτουργεί σαν βαλβίδα και ελέγχει πότε περνά η τροφή στο λεπτό έντερο.

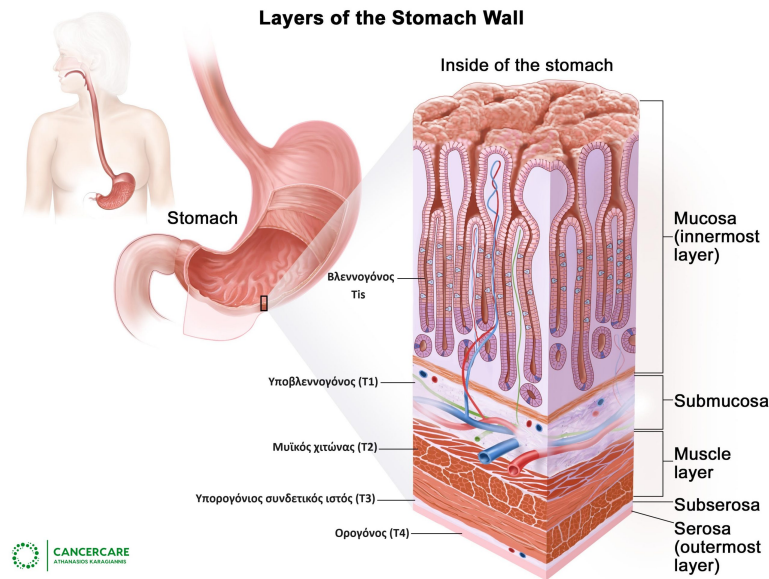


Σχήμα 2.3 Ανατομία στομάχου
[10]

Το τοίχωμα του στομάχου έχει 5 στρώσεις:

- Η εσωτερική στρώση λέγεται βλεννογόνος. Εκεί παράγεται το γαστρικό οξύ και τα πεπτικά ένζυμα. Οι περισσότεροι καρκίνοι του στομάχου ξεκινούν από αυτή τη στρώση.
- Κάτω από τον βλεννογόνο βρίσκεται ο υποβλεννογόνιος χιτώνας, ένα υποστηρικτικό στρώμα.
- Εξωτερικά από αυτή βρίσκεται η ιδίως μυϊκή (muscularis propria), μια παχιά στρώση μυών που βοηθά στην ανάμειξη και μετακίνηση της τροφής.

- Οι δύο εξωτερικές στρώσεις είναι η υποορογόνος και η ορογόνος, που περιβάλλουν το στομάχι σαν κάλυμμα.



Σχήμα 2.4 Στρώσεις στομάχου

Οι στρώσεις αυτές είναι πολύ σημαντικές για τον καθορισμό του σταδίου του καρκίνου, κάτι που επηρεάζει τις θεραπευτικές επιλογές και την πρόγνωση. Όσο βαθύτερα εισχωρεί ο καρκίνος από τον βλεννογόνο προς τις εξωτερικές στρώσεις, τόσο πιο προχωρημένο είναι το στάδιο και τόσο πιο εντατική μπορεί να χρειάζεται να είναι η θεραπεία.

2.2.2. Ανάπτυξη και τύποι Γαστρικού Καρκίνου

Οι καρκίνοι του στομάχου συνήθως αναπτύσσονται αργά στο διάστημα αρκετών χρόνων. Πριν εμφανιστεί ο καρκίνος, συχνά προηγούνται προκαρκινικές αλλοιώσεις στο εσωτερικό τοίχωμα του στομαχιού. Αυτές οι πρώιμες αλλαγές σπάνια προκαλούν συμπτώματα, για αυτό και τις περισσότερες φορές παραμένουν αδιάγνωστες.

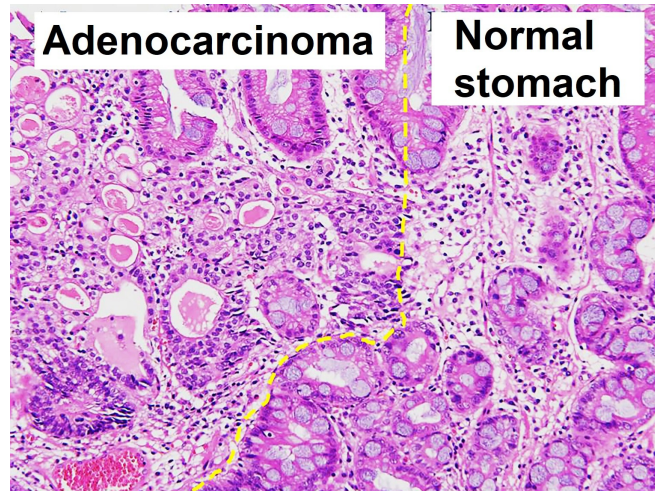
Αδενοκαρκινώματα

Οι περισσότεροι καρκίνοι του στομάχου (περίπου 90% έως 95%) είναι αδενοκαρκινώματα. Αυτοί οι καρκίνοι ξεκινούν από τα αδενικά κύτταρα της εσωτερικής επένδυσης του στομάχου (βλεννογόνος). Τα περισσότερα δεδομένα που αναφέρονται στον καρκίνο του στομάχου αφορούν αυτόν τον τύπο.

Υπάρχουν δύο βασικοί τύποι αδενοκαρκινώματος στομάχου:

- Εντερικός τύπος (intestinal type): Τείνει να έχει καλύτερη πρόγνωση. Τα καρκινικά κύτταρα εμφανίζουν συχνά συγκεκριμένες γενετικές αλλαγές, που επιτρέπουν στοχευμένες θεραπείες με φάρμακα.

- Διάχυτος τύπος (diffuse type): Τείνει να εξαπλώνεται πιο γρήγορα, είναι λιγότερο συχνός και πιο δύσκολος στη θεραπεία.



Σχήμα 2.5 Αδενοκαρκίνωμα σε σύγκριση με υγιή ιστό στομάχου [10]

Γαστρεντερικοί στρωματικοί όγκοι (GISTs)

Αυτοί οι σπάνιοι όγκοι ξεκινούν από πολύ πρώιμα κύτταρα στο τοίχωμα του στομάχου, που ονομάζονται διηθητικά κύτταρα του Cajal. Παρόλο που μπορεί να εμφανιστούν σε οποιοδήποτε σημείο του πεπτικού σωλήνα, οι περισσότεροι ξεκινούν στο στομάχι.

Νευροενδοκρινικοί όγκοι (NETs / καρκινοειδή)

Ξεκινούν από κύτταρα που έχουν χαρακτηριστικά τόσο νευρικών κυττάρων όσο και ορμονοπαραγωγών (ενδοκρινών) κυττάρων. Οι περισσότεροι NETs είναι αργά αναπτυσσόμενοι και δεν εξαπλώνονται εύκολα, αν και κάποιοι είναι πιο επιθετικοί.

Λεμφώματα

Αυτοί οι καρκίνοι ξεκινούν από λεμφοκύτταρα του ανοσοποιητικού συστήματος. Συνήθως εκδηλώνονται σε άλλα σημεία του σώματος, αλλά μπορούν να ξεκινήσουν και στο τοίχωμα του στομάχου. Η θεραπεία και η πρόγνωση εξαρτώνται από τον τύπο του λεμφώματος και άλλους παράγοντες.

Μερικοί άλλοι τύποι που εμφανίζονται πολύ σπάνια είναι το πλακώδες καρκίνωμα, το καρκίνωμα μικροκυττάρων και τα λειομυοσαρκώματα.

Στάδια [11]

Στάδιο 0 (καρκίνωμα in situ):

Ανώμαλα κύτταρα εντοπίζονται μόνο στον βλεννογόνο (εσωτερική επένδυση). Είναι το πιο πρώιμο στάδιο και δεν συνιστά ακόμα καρκίνο.

Στάδιο 1:

Διακρίνεται σε 1A και 1B.

- 1A: Ο καρκίνος είναι στον βλεννογόνο ή στον υποβλεννογόνο χιτώνα.
- 1B: Έχει φτάσει μέχρι τη μυϊκή στιβάδα ή έχει επεκταθεί σε 1-2 λεμφαδένες.

Στάδιο 2:

Διακρίνεται σε 2A και 2B.

- 2A: Ο καρκίνος έχει επεκταθεί βαθύτερα και σε 3–6 λεμφαδένες ή μέχρι την υποορογόνια.
- 2B: Έχει επεκταθεί πιο εκτεταμένα (π.χ. σε 7–15 λεμφαδένες ή έως και την ορογόνο).

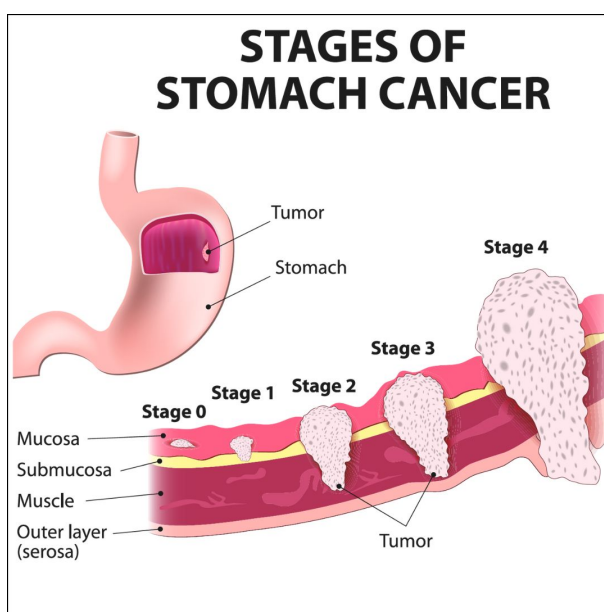
Στάδιο 3:

Διακρίνεται σε 3A, 3B και 3C.

Σε αυτά τα στάδια, ο καρκίνος έχει φτάσει βαθύτερες στιβάδες και πολλούς λεμφαδένες (πάνω από 16) ή έχει εξαπλωθεί σε γειτονικά όργανα όπως σπλήνα, ήπαρ, πάγκρεας, παχύ έντερο κ.ά.

Στάδιο 4 (μεταστατικός καρκίνος):

Ο καρκίνος έχει εξαπλωθεί σε απομακρυσμένα σημεία όπως πνεύμονες, ήπαρ, μακρινούς λεμφαδένες ή το περιτόναιο. Ονομάζεται μεταστατικός καρκίνος στομάχου, ακόμα κι αν εντοπίζεται σε άλλα όργανα.



Σχήμα 2.6 Στάδια Γαστρικού Καρκίνου

2.2.3.Συμπτώματα και θεραπεία

Τα συμπτώματα του γαστρικού καρκίνου εμφανίζονται αρκετά αργότερα από την ανάπτυξη καρκινικής αλλοίωσης και έχουν άμεση σχέση με τη λειτουργία του στομάχου. Συχνότερα είναι τα εξής:

- Ανορεξία
- Απώλεια βάρους
- Κοιλιακό άλγος
- Φούσκωμα, Δυσπεψία
- Αδυναμία
- Έμετοι, Αιματέμεση

Οι πιο συχνές θεραπείες για τον καρκίνο του στομάχου είναι, όπως αναφέρθηκαν και προηγουμένως γενικότερα για τον καρκίνο, η χημειοθεραπεία, η ακτινοθεραπεία και η ανοσοθεραπεία.

2.2.4.Ιστοπαθολογία

Ο ιστολογικός τύπος του γαστρικού όγκου αποτελεί σημαντικό παράγοντα που επηρεάζει την πρόγνωση, αλλά και βασικό στοιχείο για τον καθορισμό της έκτασης της χειρουργικής εκτομής και τη διαμόρφωση κατάλληλου χειρουργικού σχεδίου. Αντίστοιχα, ο βαθμός διαφοροποίησης των μικτών γαστρικών όγκων αποτελούν επίσης σημαντικό προγνωστικό παράγοντα[12].

Η ιστοπαθολογία επομένως αποτελεί θεμέλιο λίθο για τη διάγνωση και πρόγνωση του γαστρικού καρκίνου, καθώς επιτρέπει την άμεση μικροσκοπική αξιολόγηση του ιστού που λαμβάνεται μέσω βιοψίας. Μέσω της ιστοπαθολογικής ανάλυσης προσδιορίζεται ο τύπος του καρκίνου, το βάθος διήθησης στις στιβάδες του στομάχου, καθώς και η ύπαρξη ή μη λεμφαγγειακής και νευρικής διήθησης.

Επιπλέον, παρέχει πληροφορίες για το βαθμό διαφοροποίησης και για τα μοριακά χαρακτηριστικά, που επηρεάζουν άμεσα την επιλογή θεραπείας και την εκτίμηση πρόγνωσης. Η ακριβής ιστοπαθολογική ταξινόμηση είναι επομένως απαραίτητη για τη σωστή σταδιοποίηση της νόσου και τη χάραξη εξατομικευμένου θεραπευτικού πρωτοκόλλου.

Σημειώνεται ότι η παρακολούθηση ιστοπαθολογικών εικόνων μικτών γαστρικών όγκων παρουσιάζει αρκετές δυσκολίες δεδομένου ότι υφίστανται σημαντικές διαφορές στα βιολογικά χαρακτηριστικά, τον βαθμό κακοήθειας και τα παθολογοανατομικά τους γνωρίσματα.

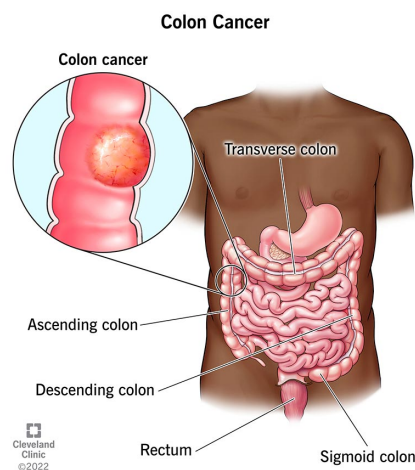
Επί του παρόντος, δεν υπάρχει ενιαίο σύστημα παθολογικής ταξινόμησης για τους μικτούς γαστρικούς όγκους, γεγονός που επηρεάζει την κλινική αντιμετώπιση και την πρόγνωση. Επιπλέον, με την πρόοδο της βιολογικής έρευνας και την ανάπτυξη νέων στοχευμένων μοριακών θεραπειών, η ακριβής παθολογική διάγνωση των γαστρικών όγκων με μεγάλη ετερογένεια καθίσταται ακόμα πιο κρίσιμη[12].

2.3.Καρκίνος του παχέος εντέρου

Ο καρκίνος του παχέος εντέρου (colon cancer) είναι μια κακοήθης νόσος η οποία παρουσιάζει σημαντική αύξηση συχνότητας εμφάνισης στον πληθυσμό τα τελευταία χρόνια, ειδικά στις δυτικές κοινωνίες[13]. Απαντάται στατιστικά σε όλο και νεαρότερης ηλικίας ασθενείς χωρίς απαραίτητα να έχουν κληρονομικό ιστορικό.

Είναι πολύ σημαντικό να αναφερθεί ότι η πρόγνωση του είναι πολύ καλή σε σχέση με άλλα κακοήθη νοσήματα και η ίαση του είναι σε πολύ καλά ποσοστά αν επιτευχθεί έγκαιρη διάγνωση. Ο καρκίνος του παχέος εντέρου εμφανίζεται στους άνδρες με ελαφρά υψηλότερη συχνότητα σε σχέση με τις γυναίκες. Επίσης, αποτελεί την τρίτη συχνότερη αιτία θανάτων από καρκίνο διεθνώς.

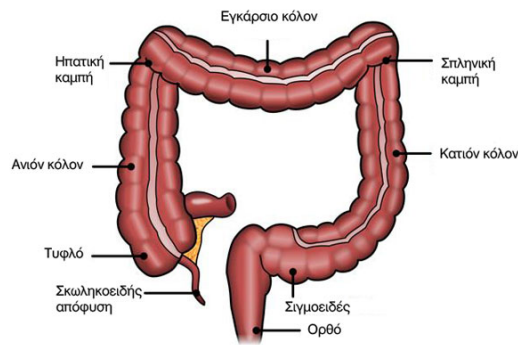
Ο καρκίνος του παχέος εντέρου αναπτύσσεται από ορισμένους πολύποδες ή ανάπτυξεις στον εσωτερικό βλεννογόνο του παχέος εντέρου. Οι πάροχοι υγειονομικής περίθαλψης διαθέτουν εξετάσεις screening που ανιχνεύουν προκαρκινώδεις πολύποδες πριν γίνουν καρκινώδεις όγκοι[14]. Ο καρκίνος του παχέος εντέρου που δεν ανιχνεύεται ή δεν αντιμετωπίζεται μπορεί να εξαπλωθεί σε άλλα μέρη του σώματος.



Σχήμα 2.7 Καρκίνος Παχέος Εντέρου

2.3.1.Παχύ έντερο

Το παχύ έντερο αποτελεί το τελευταίο τμήμα του γαστρεντερικού συστήματος και είναι υπεύθυνο για την απορρόφηση κυρίως του νερού και των ηλεκτρολυτών. Εκεί γίνεται ο σχηματισμός των κοπράνων τα οποία στη συνέχεια θα αποβληθούν[15]. Το μήκος του παχέος εντέρου είναι περίπου 1,5 μέτρο. Το τελικό τμήμα του που έχει μήκος περίπου 15 εκατοστά αποτελεί το ορθό. Το ορθό είναι ουσιαστικά ο αποθηκευτικός χώρος του παχέος εντέρου και έχει χωρητικότητα 0,6-1,2 λίτρα. Ο βασικός ρόλος του ορθού είναι η αποβολή των κοπράνων (αφόδευση).



Σχήμα 2.8 Ανατομία εντέρου

2.3.2. Ανάπτυξη καρκίνου του παχέος εντέρου

Η αυξημένη συχνότητα εμφάνισης καρκίνου εντέρου οφείλεται στο ότι η ύπαρξη πολυπόδων είναι αρκετά συχνή και ένα μεγάλο ποσοστό από αυτούς θα εξελιχθούν σε καρκίνο εντέρου.

Οι πολύποδες του παχέος εντέρου αυξάνονται σε συχνότητα όσο αυξάνεται η ηλικία και αποτελούν το πρώτο καλόηθες στάδιο πριν τον καρκίνο παχέος εντέρου. Το ένα τρίτο του πληθυσμού μετά τα 50 έχει κάποιο πολύποδα και το ποσοστό αυτό αυξάνεται στο 50% μετά την ηλικία των 70 ετών. Όπως έχει αποδειχθεί οι αδενωματώδεις πολύποδες είναι αυτοί που θα εξελιχθούν σε κακοήθεια.

Η μετατροπή ενός πολύποδα σε κακοήθη γίνεται με αρκετά βραδύ ρυθμό οπότε προλαβαίνουμε με την κολonosκόπηση να τον ανακαλύψουμε εγκαίρως και να τον αφαιρέσουμε. Χρειάζονται 5-10 έτη για να εξελιχθεί ένας πολύποδας σε καρκίνο εντέρου συνεπώς κάνοντας την πρώτη κολonosκόπηση στην ηλικία των 50 ετών και επαναλαμβάνοντας ανά 5ετία προλαμβάνουμε τη νόσο και αποφεύγουμε την ανάγκη για χειρουργική επέμβαση.

Το μέγεθος του πολύποδα εντέρου, η μορφολογία και η σύσταση του επηρεάζουν την πιθανότητα ανάπτυξης κακοήθειας. Οι πολύποδες που είναι μικρότεροι από 1 εκ. σπάνια κρύβουν κακοήθεια (1,5%), ενώ πάνω από 2 εκ το ποσοστό κακοήθειας φτάνει σχεδόν στο 50%. Όταν κατά την κολonosκόπηση βρεθεί κάποιος πολύποδας αυτός θα πρέπει να αφαιρείται και να αποστέλλεται για βιοψία. Εφόσον δεν υπάρχει κακοήθεια στη βιοψία και η αφαίρεση είναι ολική δεν χρειάζεται περαιτέρω θεραπεία.



Σχήμα 2.9 Πολύποδες εντέρου
[16]

Υπερπλαστικοί πολύποδες (*Hyperplastic Polyps*)

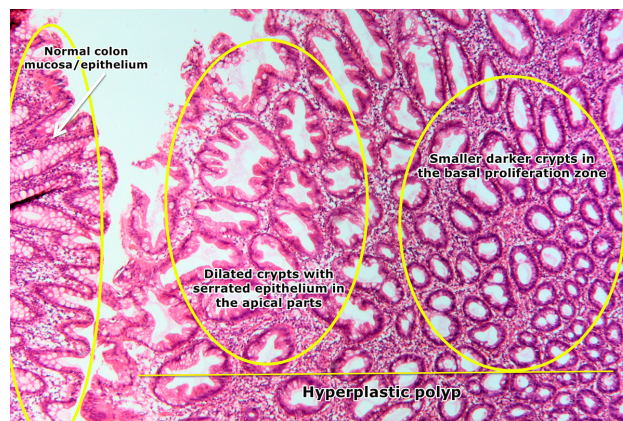
Οι υπερπλαστικοί πολύποδες είναι καλοήθειες αναπτύξεις που δημιουργούνται κατά μήκος του εσωτερικού βλεννογόνου του παχέος εντέρου και του ορθού. Είναι πιο συνηθισμένοι στο λημμικό μέρος του παχέος εντέρου, το οποίο είναι το τελευταίο τμήμα του παχέος εντέρου[17]. Οι υπερπλαστικοί πολύποδες στο παχύ έντερο δεν είναι συνήθως καρκινογόνοι ή εγγενώς επικίνδυνοι.

Υπάρχουν τρία είδη υπερπλαστικών πολυπόδων[18] :

- Υπερπλαστικοί πολύποδες με φτωχή παρουσία βλέννης: οδοντωτός πολύποδας που περιέχει μικρά κύτταρα χωρίς βλέννα.
- Υπερπλαστικοί πολύποδες πλούσιοι σε καλυκοειδή κύτταρα: οδοντωτός πολύποδας που περιέχει καλυκοειδή κύτταρα, τα οποία εκκρίνουν βλέννα.
- Οδοντωτοί πολύποδες με μικροκυστικές φυσαλίδες: οδοντωτός πολύποδας που έχει έναν συνδυασμό κυλινδρικών και καλυκοειδών κυττάρων.

Οι υπερπλαστικοί πολύποδες κατά κανόνα θεωρούνται καλοήθειες και χαμηλού κινδύνου για καρκινογένεση, ειδικά όταν είναι μικροί και εντοπίζονται στο ορθό ή το απώτερο παχύ έντερο. Ωστόσο, ανάλογα με τον τύπο, έχει αποδειχθεί ότι μπορούν να εξελιχθούν σε καρκινικές αλλοιώσεις του εντέρου.

Επομένως, η διάκριση μεταξύ των διαφόρων τύπων οδοντωτών πολυπόδων είναι σημαντική γιατί οι περισσότεροι έχουν προκαρκινική δυναμική. Για αυτό τον λόγο, η ιστολογική εξέταση είναι απαραίτητη για τη σωστή αξιολόγηση και παρακολούθηση.



Σχήμα 2.10 Υγιής ιστός σε σύγκριση με ύπαρξη υπερπλαστικών πολυπόδων

Ακτινωτά αδενώματα (*Sessile Serrated Adenomas - SSA*)

Οι κύριες προκαρκινικές βλάβες είναι οι καθιστικοί ακτινωτοί αδενωματώδεις πολύποδες (SSA)[19]. Σε σύγκριση με τα συμβατικά αδενώματα, οι SSA έχουν πιο διακριτικά μορφολογικά χαρακτηριστικά, με ασαφή όρια, και μπορεί να είναι δύσκολο να εντοπιστούν ενδοσκοπικά. Αυτές οι αλλοιώσεις έχουν μία οδοντωτή ιστολογική εμφάνιση των επιθηλιακών κρύπτων τους, όπως και κάποιι τύποι των υπερπλαστικών πολυπόδων, με

κάθε τύπο να ορίζεται από συγκεκριμένα αρχιτεκτονικά χαρακτηριστικά, την ανατομική του θέση και την έκταση της ζώνης πολλαπλασιασμού των κυττάρων.

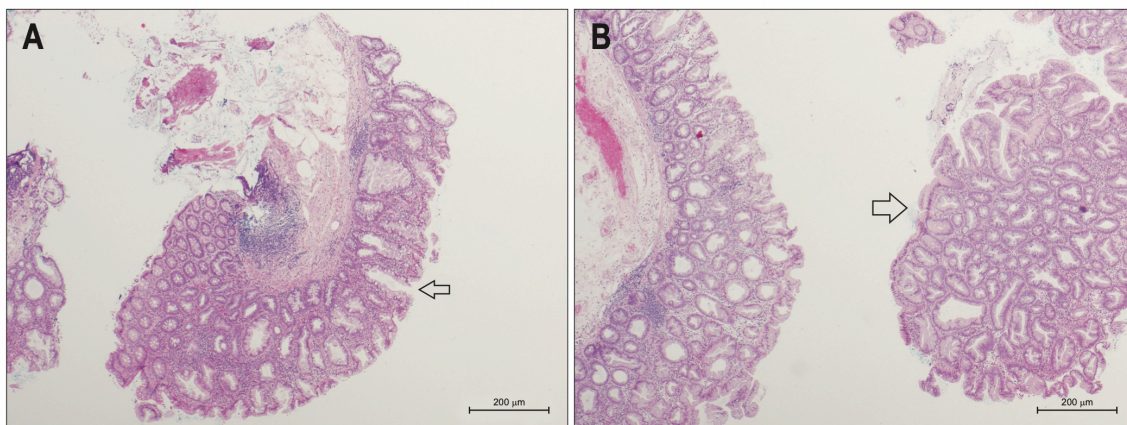
Μέχρι και τον προηγούμενο αιώνα, σχεδόν όλοι οι οδοντωτοί πολύποδες χαρακτηρίζονταν ως υπερπλαστικοί πολύποδες (HPs), καθώς τότε πιστευόταν ότι δεν ενέχουν κίνδυνο κακοήθειας και επομένως είχαν μικρή κλινική σημασία. Πλέον γνωρίζουμε ότι οι καθιστικοί οδοντωτοί πολύποδες/αδενώματα (SSA/Ps) και τα παραδοσιακά οδοντωτά αδενώματα (TSAs) έχουν τη δυνατότητα ανάπτυξης δυσπλασίας και μετέπειτα κακοήθους μετασχηματισμού, και ευθύνονται για έως και το 30% όλων των περιπτώσεων καρκίνου του παχέος εντέρου (CRC).

Η καρκινογένεση στα SSA θεωρείται ότι εξελίσσεται μέσω μιας ιδιαίτερης επιγενετικής "οδού" (οδός νεοπλασίας), μια διαδικασία κατά την οποία τα γονίδια δεν αλλοιώνονται με μεταλλάξεις, αλλά απενεργοποιούνται μέσω "χημικών σημαδιών" στο DNA. Ένα από τα βασικά "σημάδια" είναι η λεγόμενη μεθυλίωση, που συμβαίνει σε συγκεκριμένα σημεία του γενετικού υλικού (στις λεγόμενες CpG νησίδες). Όταν πολλές από αυτές τις περιοχές τροποποιηθούν (μια κατάσταση που λέγεται CIMP-high), σημαντικά γονίδια που προστατεύουν το κύτταρο από τον καρκίνο απενεργοποιούνται, και έτσι το κύτταρο γίνεται πιο ευάλωτο και μπορεί να εξελιχθεί σε καρκινικό.

Ιστοπαθολογικά χαρακτηριστικά SSA

Τα SSA χαρακτηρίζονται από την οδοντωτή αρχιτεκτονική του επιθηλίου που καλύπτει τις κρύπτες του παχέος εντέρου, χαρακτηριστικό που πιστεύεται ότι οφείλεται σε μειωμένη απόπτωση των επιθηλιακών κυττάρων. Οι τύποι των οδοντωτών αλλοιώσεων μπορούν να διαφοροποιηθούν με βάση τη θέση και την έκταση της ζώνης πολλαπλασιασμού. Οι συγκεκριμένες αιτίες αυτών των αλλαγών θεωρείται ότι σχετίζονται με επιγενετικές μεταβολές σε γονίδια υπεύθυνα για τον κυτταρικό πολλαπλασιασμό και τη διαφοροποίηση.

Τα ακτινωτά αδενώματα χαρακτηρίζονται από παραμορφωμένη ανάπτυξη των εντερικών κρύπτων (μικρών σωληνοειδών δομών μέσα στον ιστό) και διεύρυνση στη βάση τους. Αυτή η παραμόρφωση οδηγεί στη δημιουργία κρύπτων με σχήματα που θυμίζουν "μπότα", "L" ή "άγκυρα".



Σχήμα 2.11 Ιστολογικά χαρακτηριστικά Ακτινωτών Αδενωμάτων (SSA)
(A) SSA χωρίς δυσπλασία. Διακρίνονται τα κλασικά μορφολογικά χαρακτηριστικά, όπως ευρείες βάσεις και διατεταμένοι κρύπτες (βέλος)
(B) SSA με ήπια δυσπλασία στο δεξί τμήμα του δείγματος (βέλος)

2.3.3.Θεραπεία και συμπτώματα

Η κύρια θεραπεία τόσο για υπερπλαστικούς πολύποδες όσο και για οδοντωτά αδενώματα είναι η ενδοσκοπική αφαίρεση (πολυεκτομή). Ειδικότερα, για τους HP αρκεί μία απλή ενδοσκοπική αφαίρεση, κυρίως για λόγους επιβεβαίωσης της διάγνωσης μέσω ιστολογικής εξέτασης και δεν απαιτείται στενή παρακολούθηση εκτός αν υπάρχουν πολλαπλοί πολύποδες ή ιστορικό CRC.

Αντίθετα για τα SSA,εφόσον θεωρούνται προκαρκινικές αλλοιώσεις, απαιτείται πλήρης ενδοσκοπική εκτομή. Ανάλογα με το μέγεθος, την εντόπιση και την παρουσία δυσπλασίας, μπορεί να γίνει ψυχρή αφαίρεση, θερμική εκτομή ή, για τις πιο μεγάλες βλάβες, Ενδοσκοπική Βλεννογονική Εκτομή (EMR).

Τα βασικά συμπτώματα του καρκίνου του παχέος εντέρου είναι αίμα στα κόπρανα, κοιλιακός πόνος, αλλαγές στη λειτουργία του εντέρου (πχ δυσκοιλιότητα), κούραση, εμετοί και ανεξήγητη απώλεια βάρους.

2.3.4.Ιστοπαθολογία

Η χρήση ιστοπαθολογικών εικόνων για την αξιολόγηση των υπερπλαστικών πολυπόδων και των οδοντωτών αδενωμάτων (όπως οι καθιστικοί οδοντωτοί πολύποδες – SSA/Ps και τα παραδοσιακά οδοντωτά αδενώματα – TSAs) αποτελεί βασικό εργαλείο για την ακριβή διάγνωση και την εκτίμηση του κινδύνου κακοήθους εξαλλαγής. Οι δύο αυτοί τύποι εντερικών αλλοιώσεων παρουσιάζουν συχνά παρόμοια ενδοσκοπική εμφάνιση, αλλά έχουν πολύ διαφορετική βιολογική συμπεριφορά και καρκινογόνο δυναμική. Οι υπερπλαστικοί πολύποδες είναι συνήθως καλοήθεις και δεν απαιτούν στενή παρακολούθηση, ενώ οι οδοντωτοί αδενωματοειδείς πολύποδες σχετίζονται με τον καρκίνο του παχέος εντέρου μέσω της οδοντωτής νεοπλασιακής οδού και, επομένως, απαιτούν διαφορετική διαχείριση.

Η ιστοπαθολογική ανάλυση συμβάλλει επίσης ουσιαστικά στην πρόγνωση. Η παρουσία υψηλόβαθμης δυσπλασίας ή ευρημάτων όπως η εκτεταμένη υπερμεθυλίωση αποτελεί ένδειξη πιο επιθετικής βιολογικής συμπεριφοράς της βλάβης και αυξημένης πιθανότητας εξέλιξης σε καρκίνο. Αντίθετα, η απουσία αυτών των χαρακτηριστικών συσχετίζεται με πιο ευνοϊκή πρόγνωση. Με αυτόν τον τρόπο, η ιστοπαθολογία δεν καθορίζει μόνο τι είναι η βλάβη, αλλά και πόσο επικίνδυνη μπορεί να είναι μακροπρόθεσμα για τον ασθενή.

Όπως και στο στομάχι, όπου η παρουσία προκαρκινικών αλλοιώσεων επιτρέπει την πρόληψη και πρόληψη του γαστρικού καρκίνου, έτσι και στο παχύ έντερο, οι οδοντωτοί πολύποδες – ιδίως οι καθιστικοί οδοντωτοί αδενωματοειδείς πολύποδες (SSA) – αποτελούν προκαρκινικές βλάβες που μπορούν να εξελιχθούν σε κακοήθεια[20].

Η ιστοπαθολογική αξιολόγηση αυτών των αλλοιώσεων μέσω μικροσκοπικών εικόνων επιτρέπει την ακριβή διάγνωση, τον εντοπισμό δυσπλαστικών χαρακτηριστικών και την έγκαιρη θεραπευτική παρέμβαση. Κατά συνέπεια, η ιστοπαθολογία δεν συμβάλλει μόνο στην κατανόηση της φύσης των βλαβών, αλλά και στην πρόληψη της ανάπτυξης και εξέλιξης καρκινικών όγκων, ενισχύοντας τον ρόλο της πρόωμης διάγνωσης τόσο στο πεπτικό σύστημα όσο και γενικότερα στην ογκολογική πρόγνωση.

2.4.Ιατρική απεικόνιση καρκίνου

Η απεικόνιση παίζει σημαντικό ρόλο στην ανίχνευση του καρκίνου. Η σύγχρονη ιατρική τεχνολογία προσφέρει ένα ευρύ φάσμα μεθόδων απεικόνισης στους ακτινολόγους, οι οποίες παρέχουν λεπτομερείς εικόνες για την ακριβή θέση και την έκταση της νόσου. Γνωστές μέθοδοι που χρησιμοποιούνται για την ανίχνευση του καρκίνου είναι ο υπέρηχος (US), η αξονική τομογραφία (CT), η μαγνητική τομογραφία (MRI) και η τομογραφία εκπομπής ποζιτρονίων (PET)[21].

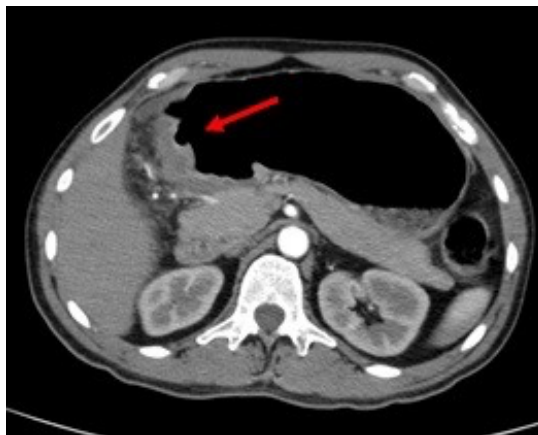
2.4.1.Ακτινολογικές Μέθοδοι - X-ray και Υπολογιστική Τομογραφία (CT)

Η ακτινογραφία είναι από τις παλαιότερες και πιο προσβάσιμες μεθόδους, χρήσιμη για την ανίχνευση καρκινικών βλαβών σε περιοχές όπως οι πνεύμονες ή τα οστά. Ωστόσο, παρουσιάζει περιορισμένη δυνατότητα διάκρισης μαλακών ιστών.

Η αξονική τομογραφία (CT) προσφέρει τρισδιάστατες εικόνες υψηλής ανάλυσης. Οι πολυτομικές CT (multi-slice CT) χρησιμοποιούνται ευρέως για σταδιοποίηση και απεικόνιση μεταστάσεων, καθώς και για την καθοδήγηση βιοψιών και επεμβάσεων[22]. Παρέχουν πληροφορίες για το μέγεθος και τη θέση του όγκου, και σε συνδυασμό με σκιαγραφικά μέσα, μπορούν να διαφοροποιήσουν ενεργούς όγκους από ουλές ή νεκρωμένους ιστούς. Ειδικές εφαρμογές της CT περιλαμβάνουν:

- CT αγγειογραφία, για την απεικόνιση της αιμάτωσης όγκων.
- CT θώρακος για την παρακολούθηση πνευμονικών όζων και μεταστάσεων.
- CT καθοδηγούμενη βιοψία ή θερμική κατάλυση όγκων.

Το στομάχι μπορεί να απεικονιστεί αρκετά καλά μέσω της αξονικής τομογραφίας (CT), η οποία συχνά μπορεί να επιβεβαιώσει τη θέση του καρκίνου. Η αξονική τομογραφία μπορεί επίσης να δείξει γειτονικά όργανα, όπως το ήπαρ, τους λεμφαδένες και απομακρυσμένα όργανα στα οποία ενδέχεται να έχει εξαπλωθεί ο καρκίνος. Η CT μπορεί να βοηθήσει στον προσδιορισμό της έκτασης ή του σταδίου του καρκίνου και να υποδείξει εάν η χειρουργική επέμβαση αποτελεί καλή επιλογή θεραπείας.



Σχήμα 2.12 Αξονική Τομογραφία Στομάχου [23]

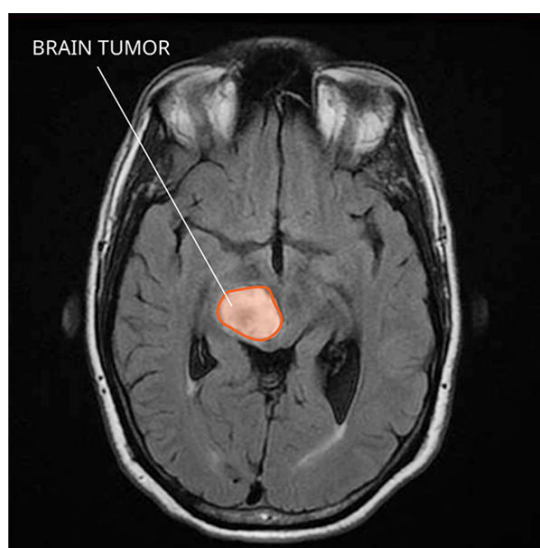
2.4.2.Μαγνητική Τομογραφία (MRI)

Η MRI δεν χρησιμοποιεί ιονίζουσα ακτινοβολία και παρέχει εξαιρετική απεικόνιση μαλακών ιστών. Είναι πολύ σημαντικό εργαλείο απεικόνισης σε περιπτώσεις καρκίνου του εγκεφάλου, προστάτη, ήπατος και μαστού.

Οι τεχνικές της:

- Diffusion-Weighted Imaging (DWI): Ανιχνεύει μεταβολές στην κυτταρική πυκνότητα και χρησιμοποιείται ευρέως στην παρακολούθηση της ανταπόκρισης στη θεραπεία.
- Dynamic Contrast Enhanced MRI (DCE-MRI): Απεικονίζει την αγγειογένεση του όγκου.
- Magnetic Resonance Spectroscopy (MRS): Αναλύει χημικές αλλαγές εντός του όγκου.
- MR Elastography: Εκτιμά τη σκληρότητα των ιστών, χρήσιμο σε όργανα όπως το ήπαρ και ο προστάτης.

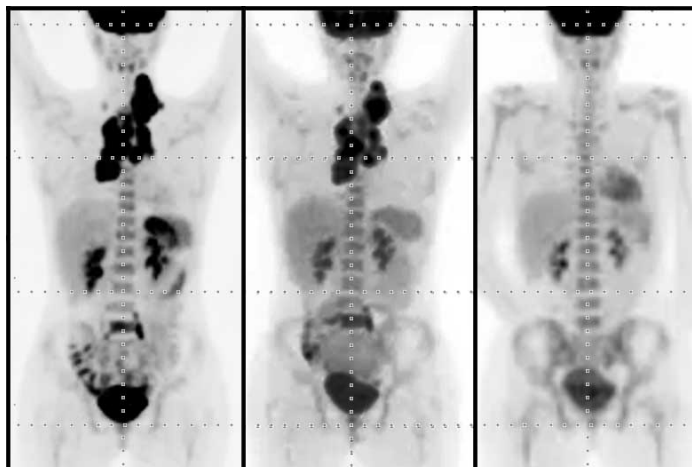
Η MRI είναι βασική για ακριβή ανατομική απεικόνιση και συχνά προτιμάται για παιδιατρικούς ασθενείς λόγω της απουσίας ακτινοβολίας.



Σχήμα 2.13 Μαγνητική Τομογραφία Εγκεφάλου [24]

2.4.3.Τομογραφία Εκτομής Ποζιτρονίων – PET

Η PET παρέχει πληροφορίες για τη μεταβολική δραστηριότητα του όγκου χρησιμοποιώντας έναν ραδιοανιχνευτή. Το πιο κοινό ραδιοφάρμακο είναι το ^{18}F -FDG, ένα ανάλογο της γλυκόζης, που συσσωρεύεται σε κύτταρα με υψηλό μεταβολισμό όπως τα καρκινικά. Η PET συνήθως συνδυάζεται με άλλες γνωστές μεθόδους ιατρικής απεικόνισης για υψηλότερη απεικονιστική ποιότητα.



Σχήμα 2.14 PET-scan

Η PET/CT συνδυάζει λειτουργικές και ανατομικές πληροφορίες και είναι βασικό εργαλείο για σταδιοποίηση, ανίχνευση υποτροπών και αξιολόγηση ανταπόκρισης στη θεραπεία.

Η PET/MRI είναι μια νεότερη υβριδική μέθοδος με υψηλή απεικονιστική ικανότητα για το νευρικό σύστημα και την πύελο.

Η PET/CT είναι πιο διαδεδομένη. Είναι αποτελεσματική στην επανασταδιοποίηση και στην ανίχνευση υποτροπής σε ασθενείς που έχουν ήδη υποβληθεί σε θεραπεία. Η PET/CT συμβάλλει επίσης στον ακριβή σχεδιασμό της ακτινοθεραπείας, ενώ χρησιμοποιείται για την αξιολόγηση της ανταπόκρισης του όγκου σε διάφορες θεραπευτικές παρεμβάσεις, επιτρέποντας έτσι την εξατομίκευση της θεραπείας με βάση τα μεταβολικά χαρακτηριστικά του όγκου[25].

Όλες οι εφαρμογές PET/CT χρησιμοποιούν έναν ραδιοανιχνευτή που εκπέμπει ποζιτρόνια, ωστόσο διαφορετικοί ανιχνευτές στοχεύουν σε διαφορετικές διεργασίες ή πρωτεΐνες στο σώμα. Χάρη στην ικανότητα στοχευμένης απεικόνισης συγκεκριμένων κυτταρικών δομών μέσω της επιλογής κατάλληλου ραδιοανιχνευτή, η PET/CT είναι γνωστή και ως «Μοριακή Απεικόνιση».

2.4.4.Υπερηχογράφημα (Ultrasound)

Το υπερηχογράφημα είναι οικονομικό, γρήγορο και ευρέως διαθέσιμο. Παρέχει πραγματικού χρόνου εικόνες και χρησιμοποιείται σε διαγνωστικούς και επεμβατικούς σκοπούς (π.χ. βιοψίες με καθοδήγηση υπερήχου).

Χρησιμοποιείται κυρίως για θυρεοειδικούς όζους, προστάτη, μαστογραφία σε γυναίκες με πυκνό μαστό, καθώς και ήπαρ και γυναικολογικούς όγκους. Επιπλέον, η ελαστογραφία υπερήχων επιτρέπει την εκτίμηση της σκληρότητας των ιστών, πολύ χρήσιμη στον καρκίνο του ήπατος και του μαστού.

2.4.5.Ενδοσκόπηση και Ενδοσκοπικός Υπέρηχος (EUS)

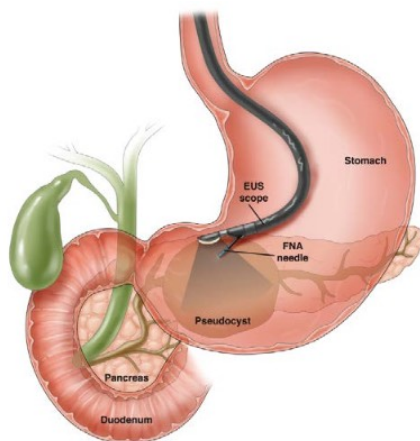
Η ενδοσκόπηση και ο ενδοσκοπικός υπέρηχος (EUS - Endoscopic Ultrasound) είναι δύο διαγνωστικές μέθοδοι που χρησιμοποιούνται κυρίως για την εξέταση του γαστρεντερικού σωλήνα.

Η ενδοσκόπηση είναι μία απεικονιστική εξέταση κατά την οποία ένας λεπτός σωλήνας με κάμερα (ενδοσκόπιο) εισάγεται μέσω του στόματος ή του πρωκτού για την άμεση οπτική παρατήρηση του εσωτερικού των κοίλων οργάνων, όπως ο οισοφάγος, το στομάχι, το δωδεκαδάκτυλο ή το παχύ έντερο. Χρησιμοποιείται για την ανίχνευση φλεγμονών, ελκών, πολυπόδων ή καρκίνων. Παρέχει οπτική εικόνα του βλεννογόνου και επιτρέπει τη λήψη βιοψιών[26].

Κατά τη διάρκεια ενός ενδοσκοπικού υπερήχου (EUS), ένας μικρός υπερηχογραφικός μετατροπέας τοποθετείται στην άκρη του ενδοσκοπίου και το ενδοσκόπιο περνάει από τον λαιμό προς το στομάχι. Ο μετατροπέας τοποθετείται απευθείας πάνω στο τοίχωμα του στομάχου, στο σημείο όπου εντοπίζεται ο καρκίνος.

Μέσω του EUS μπορούν να απεικονιστούν με ακρίβεια τα επιμέρους στρώματα του τοιχώματος του στομάχου, καθώς και οι γειτονικοί λεμφαδένες και άλλες δομές που βρίσκονται ακριβώς έξω από το στομάχι. Η ποιότητα της εικόνας που παρέχει ο EUS είναι συνήθως ανώτερη από αυτή ενός συμβατικού υπερήχου, λόγω της μικρότερης απόστασης που διανύουν τα ηχητικά κύματα.

Ο ενδοσκοπικός υπέρηχος είναι ιδιαίτερα χρήσιμος για τον προσδιορισμό του πόσο βαθιά έχει επεκταθεί ο καρκίνος στο τοίχωμα του στομάχου, στους γύρω ιστούς ή στους κοντινούς λεμφαδένες.



Σχήμα 2.15 Ενδοσκοπικός Υπέρηχος Στομάχου

2.4.6.Ιστοπαθολογικές εικόνες

Οι παραπάνω διαγνωστικές μέθοδοι παράγουν ποικίλα είδη ιατρικών δεδομένων, τα οποία διαφέρουν ως προς τη μορφή, την ανάλυση και το διαγνωστικό βάθος. Η αξονική και μαγνητική τομογραφία παράγουν τρισδιάστατες ανατομικές εικόνες, οι οποίες συνήθως αποθηκεύονται σε μορφή DICOM και περιέχουν πληροφορίες για την πυκνότητα των ιστών.

Οι τεχνικές PET/CT και PET/MRI δημιουργούν υβριδικά δεδομένα, τα οποία συνδυάζουν ανατομική και λειτουργική πληροφορία, όπως μεταβολική δραστηριότητα, πρόσληψη ραδιοανιχνευτών ή υποδοχέων κυττάρων. Οι υπέρηχοι, συμπεριλαμβανομένου και του EUS, παράγουν δυναμικά υπερηχογραφήματα, τα οποία συχνά καταγράφονται και αξιολογούνται σε πραγματικό χρόνο.

Τα δεδομένα αυτά χρησιμοποιούνται για τη διάγνωση, σταδιοποίηση και παρακολούθηση της θεραπευτικής ανταπόκρισης, ωστόσο παρέχουν περιορισμένη πληροφορία σε μικροσκοπικό κυτταρικό επίπεδο.

Η ιστοπαθολογία καλείται να καλύψει αυτό το κενό παρέχοντας ένα πληρέστερο απεικονιστικά επίπεδο. Οι εικόνες που προκύπτουν από την ιστοπαθολογική ανάλυση αποτελούν ακριβείς ψηφιοποιημένες μικροσκοπικές απεικονίσεις βιοψιών, που έχουν χρωσθεί συνήθως με αιματοξυλίνη και ηωσίνη (H&E staining), καθιστώντας την ιστοπαθολογική απεικόνιση ιδιαίτερης σημασίας.

Στην κλινική πρακτική, η ιστοπαθολογία αποτελεί μορφή ανατομικής παθολογίας που επικεντρώνεται ειδικά στους ιστούς και τα όργανα που έχουν ληφθεί μέσω βιοψίας ή χειρουργικής αφαίρεσης. Τα δείγματα ιστών υφίστανται επεξεργασία με την δημιουργία λεπτών τομών και την τοποθέτησή τους σε γυάλινες πλάκες. Στη συνέχεια, μπορούν να υποβληθούν σε χρώσεις ή άλλες τεχνικές, ώστε να επισημανθούν, να διαχωριστούν ή να απομακρυνθούν συγκεκριμένα κύτταρα, ώστε να είναι πιο ευδιάκριτα στο μικροσκόπιο[27].

Η ιστοπαθολογία διαφέρει από την κυτταροπαθολογία, η οποία εξετάζει μεμονωμένα κύτταρα. Αντιθέτως, η ιστοπαθολογία αξιολογεί όχι μόνο τη δομή των κυττάρων αλλά και τον τρόπο με τον οποίο είναι οργανωμένα μεταξύ τους.

Στο πλαίσιο της σύγχρονης ιατρικής, η ψηφιοποίηση των ιστοπαθολογικών διαφανειών και η ανάλυσή τους με υπολογιστικά εργαλεία και αλγορίθμους μηχανικής μάθησης (π.χ. CNNs, Vision Transformers) έχει οδηγήσει στην ανάπτυξη Υποβοηθούμενων Συστημάτων Διάγνωσης (CAD) που συμβάλλουν στην ακριβέστερη και ταχύτερη αξιολόγηση καρκινικών αλλοιώσεων.

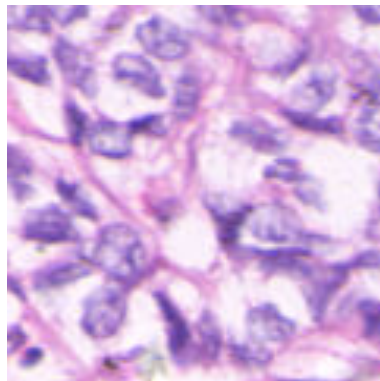
Από τα σύνολα δεδομένων που προκύπτουν από την ιστοπαθολογία, αναλύονται:

- Το σχήμα, μέγεθος και διάταξη των κυττάρων και πυρήνων
- Η αρχιτεκτονική των ιστών
- Η πυκνότητα, η χρώση και η υφή
- Η ύπαρξη φλεγμονώδους διήθησης, αγγείων ή νέκρωσης

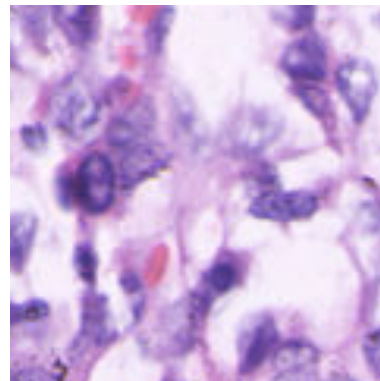
Αυτού του είδους τα δεδομένα είναι πλούσια σε πληροφορία και ιδιαίτερα κατάλληλα για μοντέλα βαθιάς μάθησης, καθιστώντας την ιστοπαθολογική εικόνα το επίκεντρο πολλών ερευνητικών προσπαθειών για την ταξινόμηση, πρόγνωση και εξατομίκευση της θεραπείας του καρκίνου.

Οι ιστοπαθολογικές εικόνες είναι συνήθως πολύ υψηλής ανάλυσης (Whole Slide Images - WSIs) και περιλαμβάνουν πληροφορία σε πολλαπλά επίπεδα μεγέθυνσης. Συνήθως χρωματίζονται με ειδικές βαφές όπως η Hematoxylin & Eosin (H&E), οι οποίες δίνουν χαρακτηριστικούς χρωματισμούς: η αιματοξυλίνη βάφει τους πυρήνες των κυττάρων μπλε/μωβ και η ηωσίνη βάφει το κυτταρόπλασμα και άλλες δομές σε ροζ/κόκκινο. Οι χρωματικές παραλλαγές μεταξύ εργαστηρίων καθιστούν ενίοτε την τυποποίηση δύσκολη.

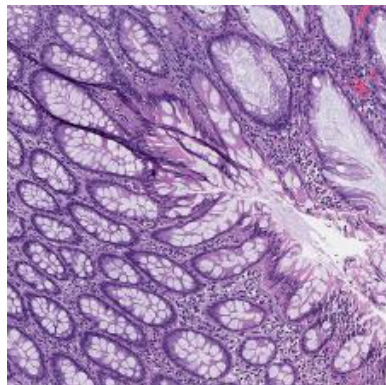
Μεταξύ των δειγμάτων από διαφορετικούς ασθενείς ή από διαφορετικά όργανα σημειώνεται σημαντική ετερογένεια. Αυτή συνίσταται σε διαφορές ως προς τη μορφολογία των κυττάρων, την πυκνότητα του ιστού και την παρουσία φλεγμονής ή νεκρωτικών περιοχών. Η ερμηνεία απαιτεί κατανόηση πολύπλοκων σχέσεων μεταξύ κυτταρικών και ιστικών χαρακτηριστικών. Ως εκ τούτου δεν είναι πάντα προφανές αν μια αλλοίωση είναι καλοήθης, προκαρκινική ή κακοήθης χωρίς προηγούμενη συσχέτιση με ευρύτερα μορφολογικά μοτίβα.



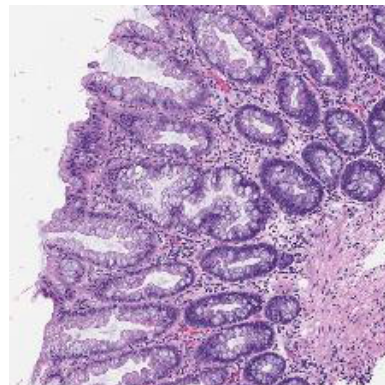
A1



A2

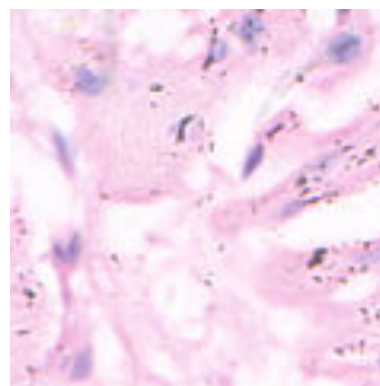
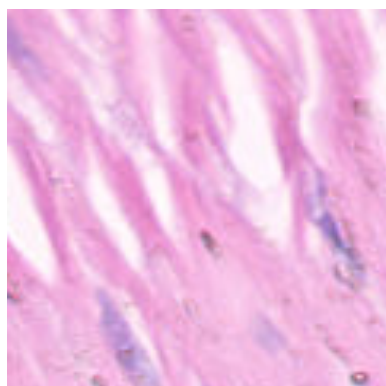


B1



B2

Σχήμα 2.16 Ιστοπαθολογικές Εικόνες Καρκινικών Ιστών
(A1-2) Στομάχου (B1-2) Παχέος Εντέρου



Σχήμα 2.17 Ιστοπαθολογικές Εικόνες Υγιούς Ιστού Στομάχου

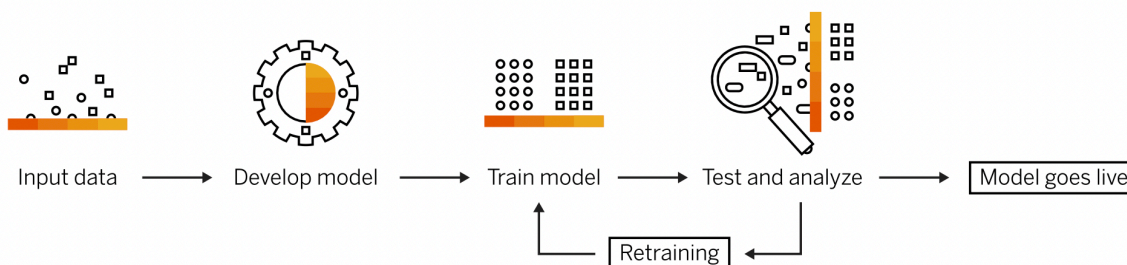
3. Μηχανική Μάθηση

Η μηχανική μάθηση είναι ένα υποσύνολο της τεχνητής νοημοσύνης (AI). Επικεντρώνεται στη διδασκαλία των υπολογιστών να μαθαίνουν από τα δεδομένα και να βελτιώνονται με την εμπειρία, αντί να είναι ρητά προγραμματισμένοι να το κάνουν. Στη μηχανική μάθηση, οι αλγόριθμοι εκπαιδεύονται για να βρουν μοτίβα και συσχετίσεις σε μεγάλα σύνολα δεδομένων και να λαμβάνουν τις καλύτερες αποφάσεις και προβλέψεις με βάση αυτή την ανάλυση[28].

Η μηχανική μάθηση χρησιμοποιείται για να κάνει τους υπολογιστές να εκτελούν δραστηριότητες, που οι άνθρωποι μπορούν να πραγματοποιήσουν, πιο αποτελεσματικά. Έχει απλοποιήσει τη ζωή και έχει καταστεί απαραίτητο εργαλείο σε πολλούς τομείς.

Στη μηχανική μάθηση, ένα πρόγραμμα υπολογιστή λαμβάνει ένα σύνολο από καθήκοντα προς εκτέλεση. Θεωρείται ότι το μηχάνημα έχει «μάθει» από την εμπειρία του όταν η μετρούμενη απόδοσή του σε αυτά τα καθήκοντα βελτιώνεται με την πάροδο του χρόνου, καθώς αποκτά όλο και περισσότερη εξάσκηση στην εκτέλεσή τους. Αυτό σημαίνει ότι το μηχάνημα λαμβάνει αποφάσεις και κάνει προβλέψεις βασισμένο σε ιστορικά δεδομένα[29].

Οι αλγόριθμοι μηχανικής μάθησης είναι βασικά σχεδιασμένοι για να ταξινομούν, να βρίσκουν μοτίβα, να προβλέπουν αποτελέσματα και να λαμβάνουν τεκμηριωμένες αποφάσεις. Οι αλγόριθμοι μπορούν να χρησιμοποιηθούν ένας κάθε φορά ή να συνδυαστούν για να επιτευχθεί η καλύτερη δυνατή ακρίβεια όταν εμπλέκονται σύνθετα και πιο απρόβλεπτα δεδομένα[28].



Σχήμα 3.1 Ροή Ανάπτυξης Μοντέλου Μηχανικής Μάθησης

3.1.Βαθιά μάθηση

Η Βαθιά Μάθηση (Deep Learning) είναι μία τεχνική μηχανικής μάθησης στην οποία αρκετά «επίπεδα» απλών μονάδων επεξεργασίας συνδέονται σε ένα δίκτυο, το ένα πίσω από το άλλο, με αποτέλεσμα η είσοδος στο σύστημα να διέρχεται διαδοχικά μέσα από κάθε ένα από αυτά.

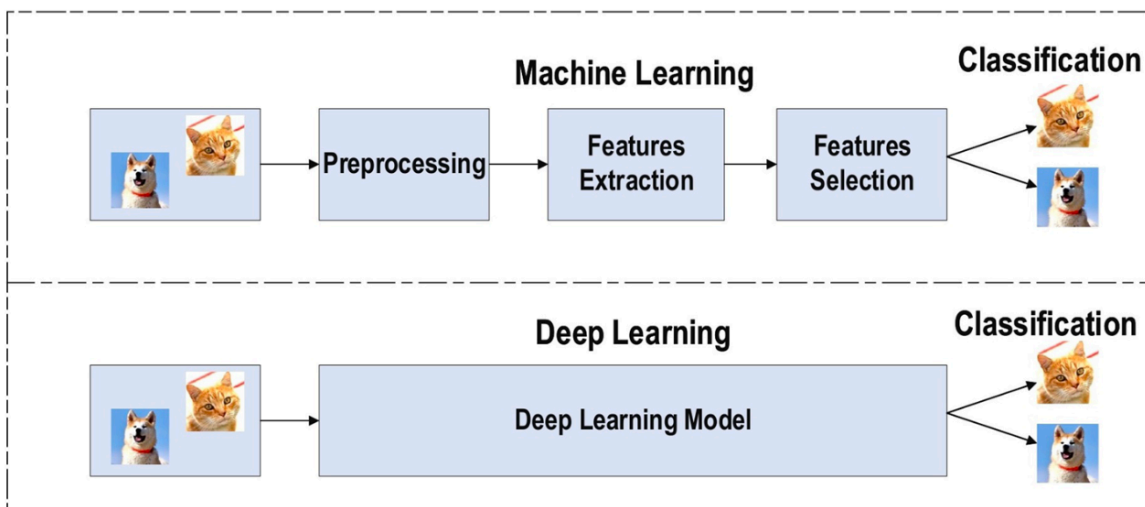
Η συγκεκριμένη αρχιτεκτονική επινοήθηκε με πρότυπο την επεξεργασία των οπτικών πληροφοριών στον εγκέφαλο, δηλ. των πληροφοριών που εισέρχονται μέσω των ματιών και συλλαμβάνονται από τον αμφιβληστροειδή χιτώνα, προχωρούν μέσα στο οπτικό

νεύρο και φτάνουν στον εγκέφαλο[30].

Στην τεχνητή νοημοσύνη (AI), η μηχανική μάθηση είναι η ικανότητα αυτόματης προσαρμογής με ελάχιστη ή και καθόλου ανθρώπινη παρέμβαση, ενώ η βαθιά μάθηση αποτελεί υποσύνολο της μηχανικής μάθησης και χρησιμοποιεί νευρωνικά δίκτυα για να προσομοιώσει τη διαδικασία μάθησης του ανθρώπινου εγκεφάλου. Υπάρχει σημαντική διαφορά ανάμεσα σε αυτές τις δύο έννοιες. Αν και η βαθιά μάθηση απαιτεί μεγαλύτερο όγκο δεδομένων για την εκπαίδευση, μπορεί να προσαρμόζεται σε νέες συνθήκες και να διορθώνει τα δικά της σφάλματα[29].

Η βαθιά μάθηση προσφέρει αυτόνομη εκμάθηση χαρακτηριστικών και την ιεραρχική αναπαράστασή τους σε πολλαπλά επίπεδα. Ολόκληρη η αρχιτεκτονική της χρησιμοποιείται για την εξαγωγή και τροποποίηση χαρακτηριστικών. Τα πρώτα επίπεδα πραγματοποιούν βασική επεξεργασία των εισερχόμενων δεδομένων ή μαθαίνουν απλά χαρακτηριστικά, και το αποτέλεσμα μεταφέρεται στα ανώτερα επίπεδα, τα οποία είναι υπεύθυνα για την εκμάθηση πιο σύνθετων χαρακτηριστικών. Ως εκ τούτου, η βαθιά μάθηση είναι κατάλληλη για την επεξεργασία μεγαλύτερων συνόλων δεδομένων και μεγαλύτερης πολυπλοκότητας.

Τα μοντέλα βαθιάς μάθησης υπερέχουν σε σχέση με τα παραδοσιακά μοντέλα μηχανικής μάθησης λόγω του αυξημένου αριθμού επιπέδων μάθησης και του υψηλότερου επιπέδου αφαίρεσης. Ένας ακόμη λόγος για αυτό το πλεονέκτημα είναι ότι η εκμάθηση όλων των στοιχείων του μοντέλου βασίζεται απευθείας στα δεδομένα, χωρίς την ανάγκη για χειροκίνητη εξαγωγή χαρακτηριστικών.



Σχήμα 3.2 Σύγκριση Προσέγγισης Μηχανικής και Βαθιάς Μάθησης

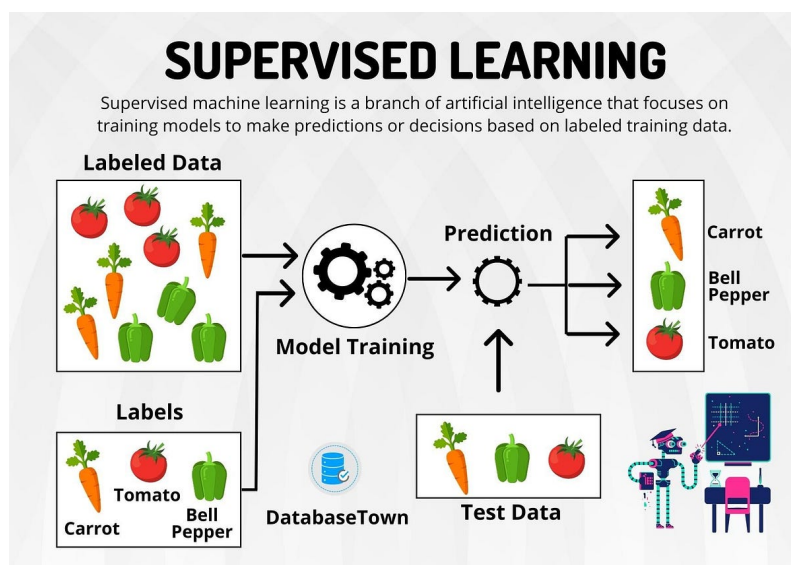
Οι τεχνικές βαθιάς μάθησης ταξινομούνται σε τρεις κύριες κατηγορίες: μη επιβλεπόμενη (unsupervised), μερικώς επιβλεπόμενη (semi-supervised) και επιβλεπόμενη (supervised) μάθηση. Επιπλέον, η βαθιά ενισχυτική μάθηση (Deep Reinforcement Learning - DRL), γνωστή και ως ενισχυτική μάθηση (Reinforcement Learning - RL), αποτελεί μια ακόμη τεχνική μάθησης, η οποία συνήθως θεωρείται ότι εμπίπτει στην κατηγορία των μερικώς επιβλεπόμενων, και περιστασιακά, των μη επιβλεπόμενων μεθόδων μάθησης.

3.1.1.Επιβλεπόμενη μάθηση

Αυτή η τεχνική βασίζεται σε επισημασμένα δεδομένα (labeled data), δηλαδή γνωρίζουμε τόσο τις εισόδους όσο και τις αντίστοιχες σωστές εξόδους. Στο πλαίσιο αυτό, το σύστημα προσπαθεί να προβλέψει τη σωστή έξοδο για μια δεδομένη είσοδο και αξιολογεί την ακρίβεια της πρόβλεψης μέσω μιας συνάρτησης απώλειας (loss function). Στη συνέχεια, τροποποιεί επαναληπτικά τις παραμέτρους του, ώστε να βελτιώσει τις προβλέψεις του. Όσο προχωρά η εκπαίδευση και το μοντέλο αποκτά εμπειρία, γίνεται ικανό να απαντά σωστά σε νέα δεδομένα που προέρχονται από το ίδιο περιβάλλον.

Στη βαθιά μάθηση (DL), υπάρχουν πολλές επιβλεπόμενες τεχνικές μάθησης, όπως τα επαναληπτικά νευρωνικά δίκτυα (RNNs), τα συνελκτικά νευρωνικά δίκτυα (CNNs) και τα βαθιά νευρωνικά δίκτυα (DNNs). Επιπλέον, η κατηγορία των RNN περιλαμβάνει πιο εξελιγμένες προσεγγίσεις, όπως οι μονάδες gated recurrent units (GRUs) και οι δομές long short-term memory (LSTM).

Το κύριο πλεονέκτημα αυτής της τεχνικής είναι η ικανότητα συλλογής δεδομένων ή παραγωγής εξόδου βασισμένης σε προηγούμενη γνώση. Ωστόσο, ένα μειονέκτημα είναι ότι το όριο απόφασης μπορεί να «υπερτεντωθεί» (overstrained) όταν το σύνολο εκπαίδευσης δεν περιλαμβάνει παραδείγματα που θα έπρεπε να ανήκουν σε κάποια κλάση. Συνολικά, αυτή η τεχνική θεωρείται απλούστερη από άλλες, προσφέροντας παράλληλα υψηλή απόδοση κατά τη διαδικασία μάθησης.



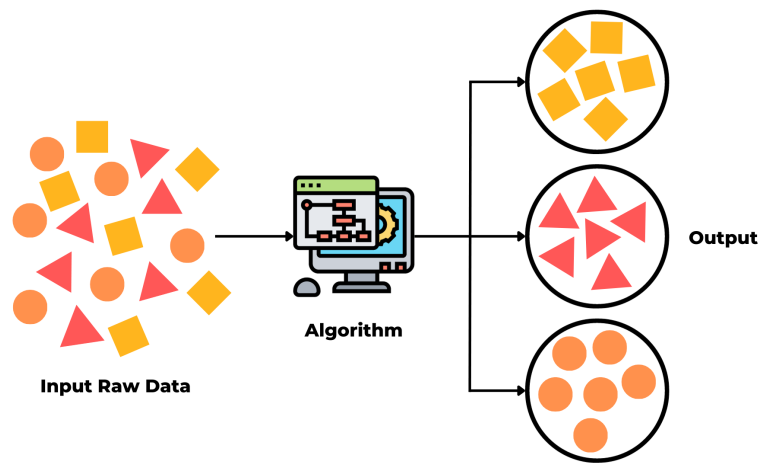
Σχήμα 3.3 Επιβλεπόμενη Μάθηση

3.1.2.Μη επιβλεπόμενη μάθηση

Αυτή η τεχνική επιτρέπει την εκπαίδευση ενός μοντέλου χωρίς να απαιτούνται επισημασμένα δεδομένα, δηλαδή χωρίς την ανάγκη ύπαρξης ετικετών. Σε αυτό το πλαίσιο, το σύστημα μαθαίνει να αναγνωρίζει τα σημαντικά χαρακτηριστικά που είναι απαραίτητα για να ανακαλύψει συσχετίσεις μέσα στα δεδομένα εισόδου.

Ορισμένες από τις πιο αποτελεσματικές μεθόδους της βαθιάς μάθησης για τέτοιου είδους καθήκοντα οι αυτόματοι κωδικοποιητές (autoencoders) και τα GANs, που αποτελούν νεότερες και πιο εξελιγμένες προσεγγίσεις. Μία από τις πιο γνωστές και ευρέως χρησιμοποιούμενες προσεγγίσεις μη επιβλεπόμενης μάθησης είναι η ομαδοποίηση (clustering).

Ωστόσο, η μη επιβλεπόμενη μάθηση παρουσιάζει μειονεκτήματα, όπως η αδυναμία παροχής ακριβούς πληροφορίας για την ταξινόμηση των δεδομένων και η αυξημένη υπολογιστική πολυπλοκότητα.



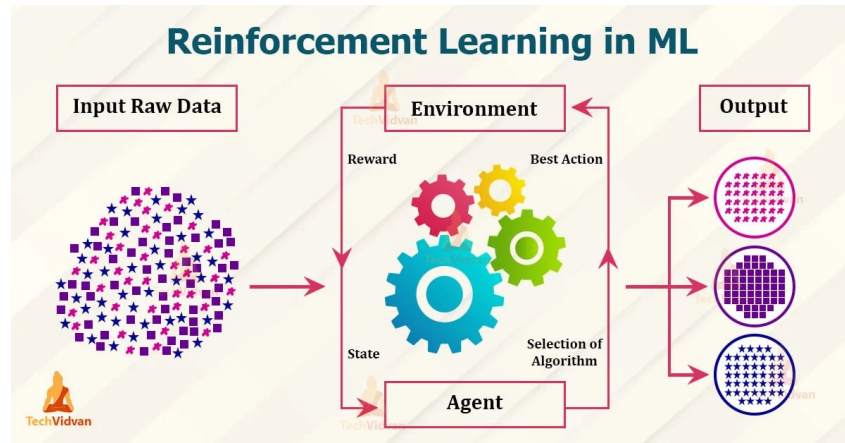
Σχήμα 3.4 Μη Επιβλεπόμενη Μάθηση

3.1.3. Ενισχυτική μάθηση

Η Ενισχυτική Μάθηση βασίζεται στην αλληλεπίδραση με το περιβάλλον, σε αντίθεση με την Επιβλεπόμενη Μάθηση, η οποία λειτουργεί πάνω σε ήδη δομημένα και επισημασμένα δεδομένα.

Επειδή η κατανομή πιθανοτήτων που διέπει το περιβάλλον είναι άγνωστη, η αλληλεπίδραση βασίζεται σε ερωτήσεις που θέτει το περιβάλλον και σε απαντήσεις που δίνει το σύστημα, οι οποίες συνοδεύονται από έναν "θορυβώδη" δείκτη απόδοσης. Λόγω αυτής της ιδιαιτερότητας, η ενισχυτική μάθηση συχνά θεωρείται μια μορφή ημιεπιβλεπόμενης μάθησης.

Βασιζόμενοι σε αυτό το πλαίσιο, έχουν αναπτυχθεί διάφορες τεχνικές, τόσο επιβλεπόμενες όσο και μη επιβλεπόμενες. Σε σύγκριση με την επιβλεπόμενη μάθηση, η ενισχυτική μάθηση είναι πιο απαιτητική, καθώς δεν υπάρχει άμεση συνάρτηση απώλειας για να καθοδηγήσει τη μάθηση. Αντίθετα, η βελτιστοποίηση γίνεται μέσω συνεχούς αλληλεπίδρασης με το περιβάλλον.



Σχήμα 3.5 Ενισχυτική Μάθηση

3.2.Τεχνητά Νευρωνικά Δίκτυα

3.2.1.Νευρωνικά δίκτυα

Με τον όρο Νευρωνικά Δίκτυα αναφερόμαστε σε ένα κύκλωμα συνδεδεμένων υπολογιστικών κόμβων, οι οποίοι ονομάζονται νευρώνες. Αυτοί διαχωρίζονται σε βιολογικούς και τεχνητούς. Στους βιολογικούς αναφερόμαστε σε ένα μέρος του νευρικού ιστού, ενώ στους τεχνητούς σε αλγοριθμικό δημιούργημα[31].

Τα τεχνητά νευρωνικά δίκτυα αποτελούν ένα είδος τεχνολογικής εκμάθησης που εμπίπτει στον τομέα της Τεχνητής Νοημοσύνης. Κέντρισαν το ενδιαφέρον πολλών επιστημών, καθώς η δημιουργία ενός προγράμματος το οποίο εκτελεί διεργασίες όπως ο ανθρώπινος εγκέφαλος θεωρήθηκε πολύ σημαντική.

3.2.2.Τεχνητά Νευρωνικά Δίκτυα (ΤΝΔ)

Τα Τεχνητά Νευρωνικά Δίκτυα (ΤΝΔ) αποτελούν μια κατηγορία υπολογιστικών μοντέλων τα οποία έχουν εμπνευστεί από τον τρόπο λειτουργίας του ανθρώπινου εγκεφάλου. Στόχος τους είναι η μίμηση της διαδικασίας μάθησης, της γενίκευσης και της λήψης αποφάσεων μέσω της επεξεργασίας δεδομένων. Η βασική αρχή λειτουργίας ενός ΤΝΔ είναι η μετατροπή εισόδων σε εξόδους μέσω πολλαπλών επιπέδων (layers) υπολογιστικών νευρώνων, οι οποίοι εφαρμόζουν συναρτήσεις ενεργοποίησης και μαθαίνουν από παραδείγματα μέσω εκπαίδευσης[32].

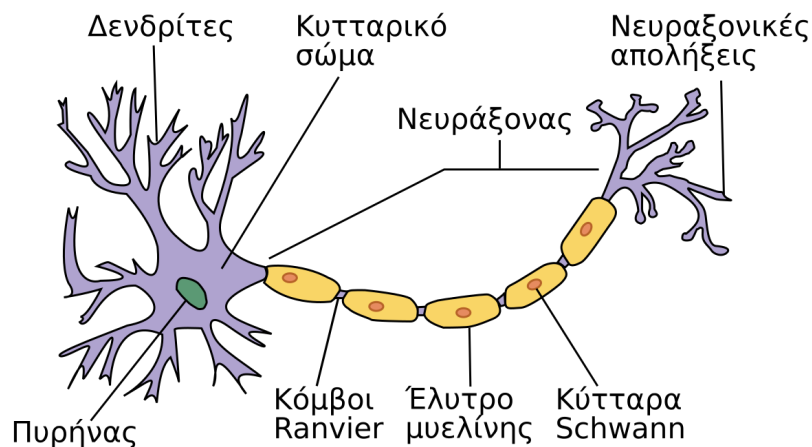
3.2.3.Βιολογική έμπνευση

Η λειτουργία των ΤΝΔ αντλεί έμπνευση από τα βιολογικά νευρωνικά δίκτυα του εγκεφάλου. Σε ένα βιολογικό νευρώνα, η πληροφορία φτάνει μέσω των δενδριτών, επεξεργάζεται στο σώμα του κυττάρου και αποστέλλεται μέσω του άξονα. Οι συνάψεις επιτρέπουν την επικοινωνία μεταξύ νευρώνων μέσω χημικών σημάτων[33]. Αντίστοιχα, σε

ένα ΤΝΔ, οι τεχνητοί νευρώνες λαμβάνουν αριθμητικές εισόδους, εφαρμόζουν βάρη και κλίσεις, και στη συνέχεια περνούν το αποτέλεσμα από μια μη γραμμική συνάρτηση ενεργοποίησης (π.χ. ReLU ή sigmoid).

Το νευρικό κύτταρο ή νευρώνας είναι το βασικό δομικό στοιχείο του εγκεφάλου τόσο στον άνθρωπο όσο και στα ζώα. Ο νευρώνας είναι ένα μεγάλο σε μέγεθος κύτταρο το οποίο, ανατομικά, αποτελείται από τα εξής τμήματα: το σώμα, τους δένδριτες, τον άξονα και τις συνάψεις που συνδέουν τις διακλαδώσεις του άξονα με τους δένδριτες άλλων νευρώνων δημιουργώντας έτσι ένα νευρωνικό δίκτυο[34].

Οι δένδριτες αποτελούν τις πύλες εισόδου του νευρώνα, ενώ ο άξονας είναι η μοναδική πύλη εξόδου του. Στέλνει σήματα προς άλλους νευρώνες υπό μορφή ηλεκτρικών παλμών σταθερού πλάτους αλλά μεταβλητής συχνότητας. Οι συνάψεις, από την άλλη, είναι τα σημεία ένωσης μεταξύ των άξονα ενός νευρώνα και των δένδριτων άλλων νευρώνων. Το ποσοστό της ηλεκτρικής δραστηριότητας που μεταδίδεται τελικά στον κάθε δεδνρίτη ονομάζεται συναπτικό βάρος.

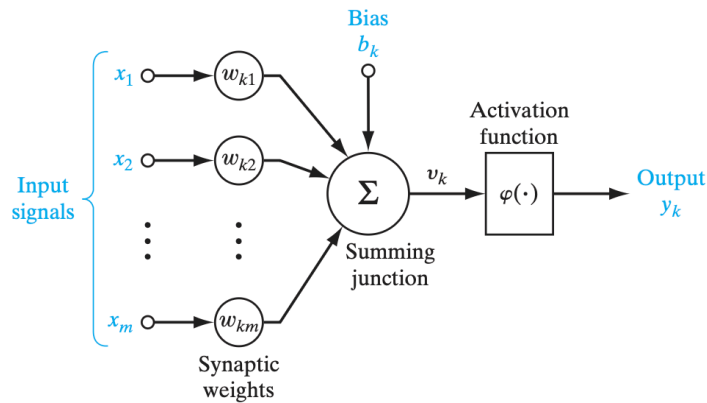


Σχήμα 3.6 Βιολογικός Νευρώνας

3.2.4. Δομή ΤΝΔ και Επίπεδα

Έχοντας ως έμπνευση το μοντέλο αυτό, οι Αμερικανοί επιστήμονες McCulloch και Pitts[35] ήταν οι πρώτοι, το 1943, που περιέγραψαν ένα απλό μοντέλο δραστηριότητας του νευρώνα.

Στον τεχνητό νευρώνα, οι συνάψεις περιγράφονται από τα συναπτικά βάρη w_{ki} , που είναι πραγματικοί αριθμοί. Τα βάρη w_{ki} μπορούμε να τα δούμε σαν παραμέτρους που ορίζουν την αντίδραση του νευρώνα για ένα δεδομένο σύνολο εισόδων, και οι τιμές τους μπορούν να προσαρμοστούν με σκοπό να γίνει εκμάθηση μιας προσέγγισης του επιθυμητού σήματος εξόδου.



Σχήμα 3.7 Τεχνητός Νευρώνας [31]

Οι εισοδοί x_i του νευρώνα συνδυάζονται για να παραχθεί το άθροισμα u_k του φορτίου που δέχεται ο νευρώνας.

$$u_k = \sum_{i=1}^n w_{ki}x_i + b_k \quad (3.2.1)$$

όπου b_k είναι το bias (όρος μετατόπισης) που μετατοπίζει τη συνάρτηση ώστε να μπορεί να μοντελοποιήσει πιο σύνθετες εξαρτήσεις. Η έξοδος u_k του αθροίσματος περνά από μία μη γραμμική συνάρτηση $\varphi(\cdot)$, όπως ReLU, sigmoid, tanh κ.ά., για να προκύψει η τελική έξοδος:

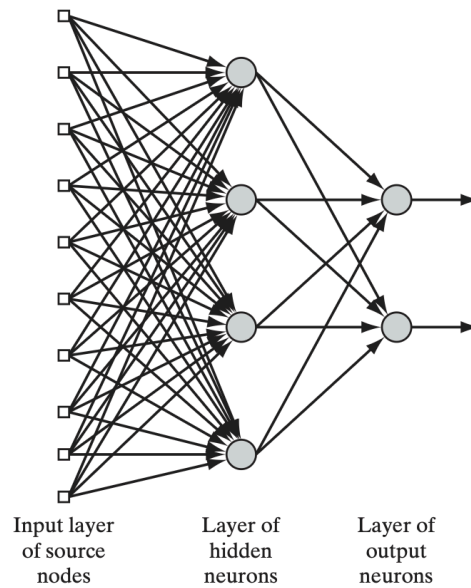
$$y_k = \varphi(u_k) \quad (3.2.2)$$

Το αποτέλεσμα y_k είναι η έξοδος του νευρώνα, που είτε πηγαίνει στην επόμενη στρώση, είτε χρησιμοποιείται ως τελική πρόβλεψη (π.χ. πιθανότητα κατηγορίας).

Ένα τυπικό ΤΝΔ περιλαμβάνει[36]:

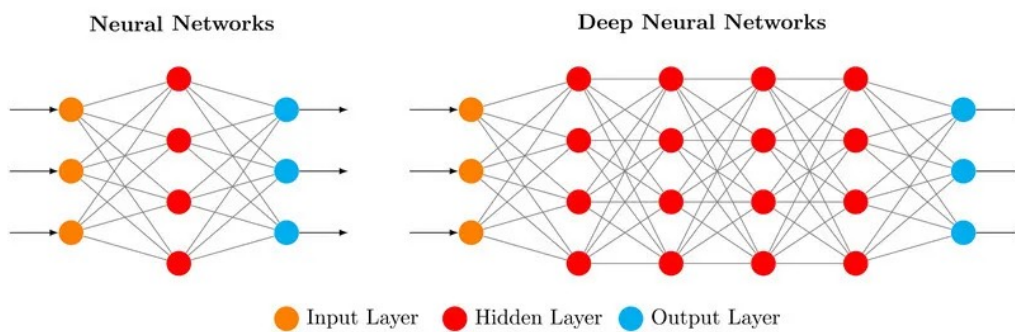
- Επίπεδο εισόδου (Input Layer) : Πρόκειται για το πρώτο επίπεδο σε ένα νευρωνικό δίκτυο που λαμβάνει τα αρχικά δεδομένα εισόδου. Οι κόμβοι εισόδου επεξεργάζονται τα δεδομένα αυτά, τα αναλύουν και τα μεταβιβάζουν στο επόμενο επίπεδο.
- Κρυφά επίπεδα (Hidden Layers) : Τα κρυμμένα στρώματα αποτελούν τα ενδιάμεσα επίπεδα ανάμεσα στο επίπεδο εισόδου και εξόδου και επιτελούν το μεγαλύτερο μέρος του υπολογισμού. Ενδέχεται να υπάρχουν πολλά κρυμμένα επίπεδα σε ένα νευρωνικό δίκτυο.
- Επίπεδο εξόδου (Output Layer) : Το επίπεδο εξόδου είναι το τελικό επίπεδο σε ένα νευρωνικό δίκτυο και οδηγεί στην έξοδο του δικτύου. Ο αριθμός των νευρώνων στο επίπεδο αυτό εξαρτάται από το εκάστοτε πρόβλημα που λύνεται.

Κάθε επίπεδο αποτελείται από νευρώνες οι οποίοι λαμβάνουν εισόδους από τους προηγούμενους και εφαρμόζουν ενεργοποιήσεις.



Σχήμα 3.8 Τεχνητό Νευρωνικό Δίκτυο

Σε βαθιά δίκτυα (DNN), ο αριθμός των κρυφών επιπέδων αυξάνεται, επιτρέποντας στο δίκτυο να μάθει πολύπλοκες συναρτήσεις και δομές στα δεδομένα.



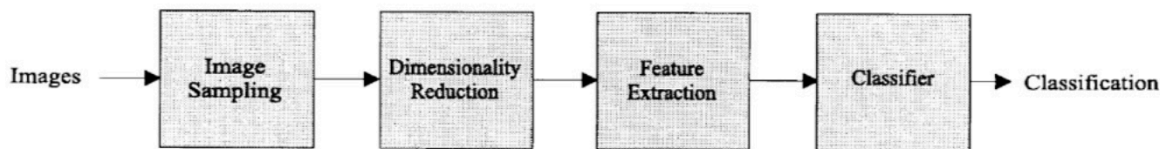
Σχήμα 3.9 Βαθύ Νευρωνικό Δίκτυο [37]

3.3. Συνελκτικά Νευρωνικά Δίκτυα

Τα Συνελκτικά Νευρωνικά Δίκτυα (CNNs) είναι μια μορφή βαθιάς μάθησης με αρχιτεκτονική προώθησης (feedforward), ιδανικά για επεξεργασία εικόνων και σημάτων. Σε αντίθεση με τα παραδοσιακά νευρωνικά δίκτυα, δεν απαιτούν χειροκίνητη εξαγωγή χαρακτηριστικών, αλλά αντλούν αυτόματα τα σημαντικότερα πρότυπα από τα δεδομένα εισόδου[38].

Χάρη στην κοινή χρήση βαρών (weight sharing) και τη δομή τοπικών υποπεδίων (receptive fields), μειώνεται δραστικά ο αριθμός παραμέτρων προς εκπαίδευση, ενισχύοντας τη γενίκευση και αποτρέποντας το overfitting. Τα συνελκτικά νευρωνικά δίκτυα έχουν σχεδιαστεί για την αναγνώριση και ταξινόμηση εικόνων. Χρησιμοποιούν αρχές της γραμμικής άλγεβρας για την εξαγωγή χαρακτηριστικών και μοτίβων από μια εικόνα.

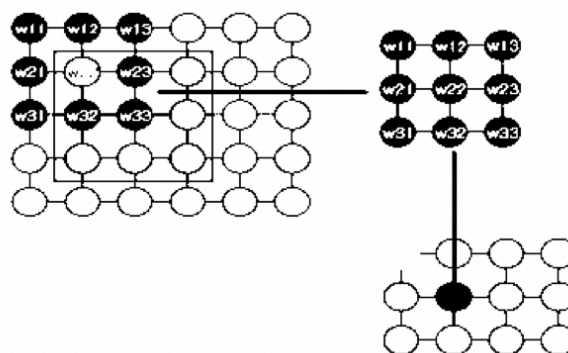
Ένα γενικό μοντέλο συνελκτικού νευρωνικού δικτύου αποτελείται από τέσσερα βασικά στάδια : το συνελκτικό επίπεδο, το επίπεδο υποδειγματοληψίας (pooling), τη συνάρτηση ενεργοποίησης και το πλήρως συνδεδεμένο επίπεδο.



Σχήμα 3.10 Ακολουθία Βημάτων για Τυπική Ταξινόμηση Εικόνων

3.3.1. Συνελκτικό Επίπεδο (Convolutional Layer)

Η εικόνα που πρόκειται να ταξινομηθεί παρέχεται στο επίπεδο εισόδου, και η έξοδος είναι η προβλεπόμενη ετικέτα κατηγορίας, η οποία υπολογίζεται βάσει των χαρακτηριστικών που έχουν εξαχθεί από την εικόνα. Σε αυτό το επίπεδο, το δίκτυο εφαρμόζει φίλτρα (kernels) σε περιοχές της εικόνας για να εξάγει χαρακτηριστικά (feature maps). Κάθε φίλτρο ανιχνεύει συγκεκριμένα πρότυπα (π.χ. ακμές, γωνίες). Όλοι οι νευρώνες χρησιμοποιούν τα ίδια βάρη (weight sharing), διευκολύνοντας την ανίχνευση ίδιων μοτίβων σε διαφορετικά σημεία.



Σχήμα 3.11 Συνελκτικό Επίπεδο
[38]

3.3.2. Συνάρτηση Ενεργοποίησης (Activation Function)

Μόλις δοθούν τα ακατέργαστα δεδομένα εισόδου, αυτά προωθούνται μέσα από το δίκτυο επίπεδο προς επίπεδο (feedforward). Σε κάθε επίπεδο, οι νευρώνες εφαρμόζουν τις προκαθορισμένες συναρτήσεις ενεργοποίησης και μεταβιβάζουν την επεξεργασμένη έξοδο στο επόμενο επίπεδο. Με αυτόν τον τρόπο, το δίκτυο μετασχηματίζει σταδιακά τα αρχικά δεδομένα σε πιο αφηρημένες και χρήσιμες αναπαραστάσεις μέχρι να παραχθεί η τελική έξοδος[37].

Η αντιστοίχιση της εισόδου με την έξοδο αποτελεί τον βασικό ρόλο κάθε συνάρτησης ενεργοποίησης σε κάθε τύπο νευρωνικού δικτύου. Η τιμή εισόδου καθορίζεται με τον υπολογισμό του σταθμισμένου άθροισματος των εισόδων ενός νευρώνα, μαζί με τον όρο μετατόπισης (bias), εφόσον υπάρχει. Με βάση αυτή την τιμή, η συνάρτηση ενεργοποίησης αποφασίζει εάν ο νευρώνας θα «ενεργοποιηθεί» ή όχι, δημιουργώντας την αντίστοιχη έξοδο. Με άλλα λόγια, η συνάρτηση ενεργοποίησης καθορίζει τη «συμπεριφορά» του νευρώνα ως απόκριση σε συγκεκριμένο ερέθισμα[39].

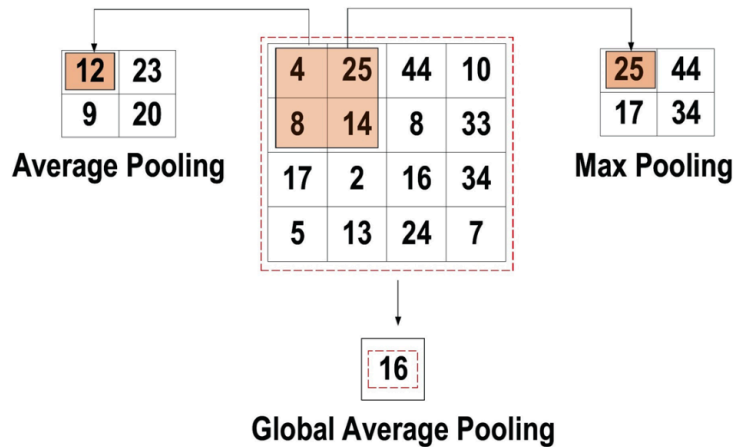
Συχνές Συναρτήσεις Ενεργοποίησης :

- Sigmoid : Η συνάρτηση sigmoid συμπιέζει τις τιμές εισόδου στο διάστημα $[0, 1]$, καθιστώντας την κατάλληλη για προβλήματα δυαδικής ταξινόμησης.
- Softmax : Η συνάρτηση softmax μετατρέπει τους αριθμούς εξόδου σε πιθανότητες, εξασφαλίζοντας ότι το άθροισμα των εξόδων ισούται με 1. Χρησιμοποιείται κυρίως στο τελικό επίπεδο ταξινόμησης πολλαπλών κατηγοριών.
- ReLU (Ορθογωνική Μονάδα Γραμμικής Ενεργοποίησης) : Η ReLU μηδενίζει τις αρνητικές τιμές εισόδου και αφήνει αναλλοίωτες τις θετικές. Είναι απλή και ιδιαίτερα αποδοτική για βαθιά δίκτυα. Η ReLU χρησιμοποιείται πιο συχνά από όλες, όσον αφορά τα συνελκτικά δίκτυα και έχει το χαμηλότερο υπολογιστικό φορτίο[38].

3.3.3. Επίπεδο pooling (Pooling Layer)

Η ακριβής θέση ενός χαρακτηριστικού παύει να είναι σημαντική μόλις αυτό εντοπιστεί. Για τον λόγο αυτό, το συνελκτικό επίπεδο ακολουθείται από ένα επίπεδο υποδειγματοληψίας. Το βασικό πλεονέκτημα της τεχνικής pooling είναι ότι μειώνει σημαντικά τον αριθμό των παραμέτρων που πρέπει να εκπαιδευτούν επιτρέποντας στο δίκτυο να αναγνωρίζει χαρακτηριστικά ανεξάρτητα από τη θέση τους στην εικόνα. Η λειτουργία pooling πραγματοποιείται επιλέγοντας ένα παράθυρο και εφαρμόζοντας μια συναρτησιακή πράξη στα στοιχεία της εισόδου που βρίσκονται μέσα σε αυτό το παράθυρο.

Η συνάρτηση pooling παράγει ένα νέο διανυσματικό σήμα εξόδου. Υπάρχουν διάφορες τεχνικές pooling, όπως η μέση υποδειγματοληψία (average pooling) και η υποδειγματοληψία μέγιστης τιμής (max-pooling), με τη δεύτερη να είναι η πιο διαδεδομένη, καθώς μειώνει σημαντικά το μέγεθος των χαρακτηριστικών χαρτών.



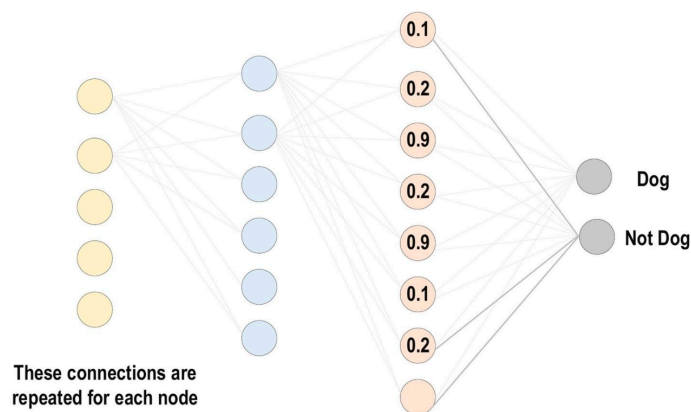
Σχήμα 3.12 Επίπεδο Pooling [39]

3.3.4. Πλήρως Συνδεδεμένο Επίπεδο (Fully Connected Layer-FC)

Το πλήρως συνδεδεμένο επίπεδο συνήθως βρίσκεται στο τελικό στάδιο κάθε αρχιτεκτονικής CNN. Σε αυτό το επίπεδο, κάθε νευρώνας συνδέεται με όλους τους νευρώνες του προηγούμενου επιπέδου.

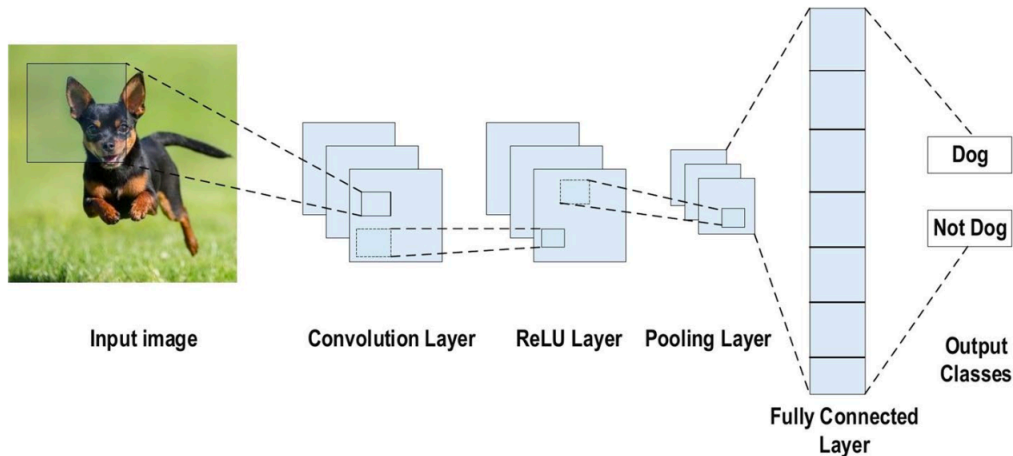
Αποτελεί τον ταξινομητή του CNN, ακολουθώντας τη βασική αρχή του πολυεπίπεδου νευρώνα, καθώς είναι τύπος προωθητικού τεχνητού νευρωνικού δικτύου (feed-forward ANN). Η είσοδος του FC επιπέδου προέρχεται από το τελευταίο επίπεδο pooling ή συνελκτικό επίπεδο και έχει τη μορφή διανύσματος, το οποίο δημιουργείται από τους χάρτες χαρακτηριστικών μέσω της διαδικασίας επιπέδωσης (flattening).

Η έξοδος του πλήρως συνδεδεμένου επιπέδου αποτελεί και την τελική έξοδο του CNN, δηλαδή την προβλεπόμενη κλάση ή πιθανότητα [39].



Σχήμα 3.13 Πλήρως Συνδεδεμένο Επίπεδο

Η εκπαίδευση πραγματοποιείται με τη μέθοδο της καθοδικής κλίσης (gradient descent), κατά την οποία η συνάρτηση κόστους εκτιμάται για ολόκληρο το σύνολο εκπαίδευσης, και τα βάρη ενημερώνονται μόνο στο τέλος κάθε εποχής (epoch). Η μέθοδος αυτή είναι ιδιαίτερα αργή όταν το σύνολο εκπαίδευσης είναι μεγάλο. Για να ξεπεραστεί αυτό το πρόβλημα, χρησιμοποιείται συχνά η στοχαστική καθοδική κλίση (stochastic gradient descent), η οποία ενημερώνει τα βάρη συχνότερα και με μικρότερο υπολογιστικό κόστος.



Σχήμα 3.14 Αρχιτεκτονική Συνελκτικού Νευρωνικού Δικτύου

3.4. Προεκπαιδευμένα Συνελκτικά Νευρωνικά Δίκτυα (Pretrained CNN)

Στο πλαίσιο της μηχανικής όρασης και ειδικά στην ταξινόμηση ιατρικών εικόνων, τα προεκπαιδευμένα συνελκτικά νευρωνικά δίκτυα (CNNs) αποτελούν σημαντικό εργαλείο που προσφέρει ακρίβεια, ταχύτητα και σταθερότητα. Τα μοντέλα αυτά έχουν εκπαιδευτεί προηγουμένως σε μεγάλα γενικά σύνολα δεδομένων, όπως το ImageNet, το οποίο περιλαμβάνει πάνω από 1 εκατομμύριο εικόνες από 1000 διαφορετικές κατηγορίες. Μέσω αυτής της εκπαίδευσης, τα μοντέλα μαθαίνουν να εξάγουν χρήσιμα και επαναχρησιμοποιήσιμα χαρακτηριστικά (features) από εικόνες, όπως άκρα, υφές, μοτίβα και δομές.

Αυτή η γνώση μπορεί να μεταφερθεί και να επαναχρησιμοποιηθεί σε άλλες εφαρμογές με λιγότερα δεδομένα, όπως η διάγνωση καρκινικών ιστών σε ιστοπαθολογικές εικόνες. Η διαδικασία αυτή είναι γνωστή ως μεταφορά μάθησης (transfer learning), και περιλαμβάνει είτε τη χρήση των προεκπαιδευμένων layers ως εξαγωγείς χαρακτηριστικών, είτε το fine-tuning, δηλαδή την επανεκπαίδευση μέρους ή όλου του μοντέλου πάνω στο νέο σύνολο δεδομένων.

Η αρχιτεκτονική των προεκπαιδευμένων μοντέλων διαφέρει, αλλά μοιράζονται κοινά χαρακτηριστικά. Αποτελούνται από πολλαπλά επίπεδα που εξάγουν προοδευτικά χαρακτηριστικά από τις εισόδους εικόνας. Τα αρχικά επίπεδα ανιχνεύουν χαρακτηριστικά χαμηλού επιπέδου, ενώ τα βαθύτερα επίπεδα αναγνωρίζουν πιο σύνθετα πρότυπα. Τα προεκπαιδευμένα μοντέλα μπορούν να προσαρμοστούν σε διάφορους τομείς, από την ιατρική απεικόνιση έως την αυτόνομη οδήγηση. Η ευελιξία και η αποτελεσματικότητά τους τα καθιστούν ανεκτίμητα εργαλεία στον τομέα της υπολογιστικής όρασης[40].

3.4.1.VGG

Το VGG είναι συντομογραφία του Visual Geometry Group και αποτελεί μια κλασική αρχιτεκτονική βαθιού συνελκτικού νευρωνικού δικτύου (CNN) με πολλαπλά επίπεδα. Ο όρος "βαθύ" αναφέρεται στον αριθμό των στρώσεων, καθώς τα μοντέλα VGG-16 και VGG-19 περιλαμβάνουν αντίστοιχα 16 και 19 συνελκτικά επίπεδα[41].

Η αρχιτεκτονική VGG αποτέλεσε τη βάση για επαναστατικά μοντέλα αναγνώρισης αντικειμένων. Ως βαθύ νευρωνικό δίκτυο, το VGGNet υπερέβη τις βασικές επιδόσεις (baselines) σε πολλές εργασίες και σύνολα δεδομένων, όχι μόνο στο ImageNet, αλλά και πέρα από αυτό. Παρά την εμφάνιση πιο σύγχρονων μοντέλων, το VGG παραμένει ένα από τα πιο δημοφιλή μοντέλα αναγνώρισης εικόνων, χάρη στην απλότητά του και τη σταθερή του απόδοση.

Το μοντέλο VGG, ή VGGNet, που υποστηρίζει 16 επίπεδα, είναι γνωστό και ως VGG16. Πρόκειται για ένα συνελκτικό νευρωνικό δίκτυο (CNN) που προτάθηκε από τους A. Zisserman και K. Simonyan του Πανεπιστημίου της Οξφόρδης. Το μοντέλο αυτό παρουσιάστηκε στο ερευνητικό άρθρο με τίτλο "Very Deep Convolutional Networks for Large-Scale Image Recognition"[42]. Το VGG16 είναι ένα από τα πρώτα μοντέλα συνελκτικού νευρωνικού δικτύου (CNN) που πέτυχε υψηλή ακρίβεια στο σύνολο δεδομένων ImageNet.

Αποτελείται από 16 επίπεδα: 13 συνελκτικά (convolutional) και 3 πλήρως συνδεδεμένα (fully connected). Οι συνελκτικές στρώσεις είναι οργανωμένες σε blocks, καθένα από τα οποία περιλαμβάνει προκαθορισμένο αριθμό στρώσεων, συνήθως δύο ή τρεις. Αυτή η δομημένη και επαναλαμβανόμενη αρχιτεκτονική καθιστά το VGG16 απλό στη σύλληψη αλλά ισχυρό σε απόδοση, ειδικά σε προβλήματα ταξινόμησης εικόνας μεγάλης κλίμακας[43].

Μια σύντομη περιγραφή της αρχιτεκτονικής VGG:

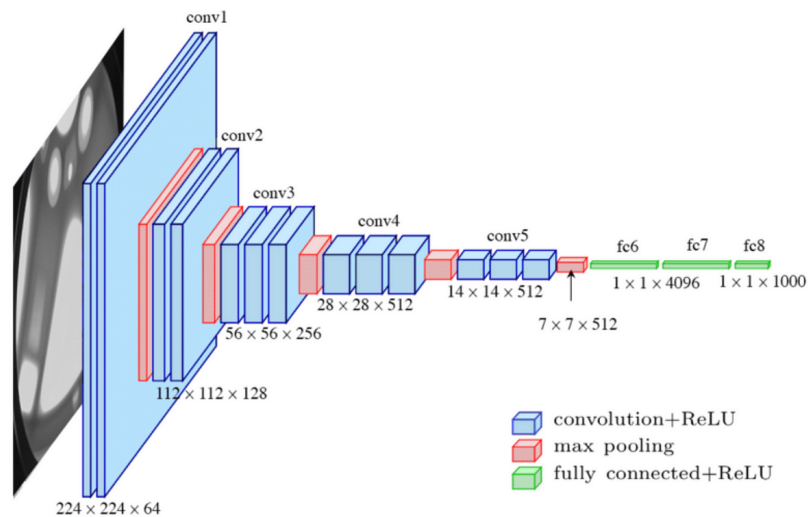
Είσοδος (Input): Το VGGNet δέχεται ως είσοδο εικόνες διαστάσεων 224×224 pixels.

Συνελκτικά Επίπεδα (Convolutional Layers): Οι συνελκτικές στρώσεις του VGG χρησιμοποιούν πολύ μικρό πεδίο υποδοχής (receptive field), δηλαδή φίλτρα μεγέθους 3×3, το μικρότερο δυνατό μέγεθος που εξακολουθεί να συλλαμβάνει πληροφορίες προς όλες τις κατευθύνσεις (πάνω, κάτω, αριστερά, δεξιά).

Υπάρχουν επίσης φίλτρα 1×1, τα οποία λειτουργούν ως γραμμικός μετασχηματισμός της εισόδου. Κάθε συνελκτικό επίπεδο ακολουθείται από τη συνάρτηση ενεργοποίησης ReLU (Rectified Linear Unit), η οποία αποτέλεσε σημαντική καινοτομία από το AlexNet, μειώνοντας σημαντικά τον χρόνο εκπαίδευσης. Η μετατόπιση (stride) του convolution είναι 1 pixel, ώστε να διατηρείται η χωρική ανάλυση της εικόνας.

Κρυφά Επίπεδα (Hidden Layers): Όλα τα κρυφά επίπεδα του VGG χρησιμοποιούν ReLU. Το VGG δεν χρησιμοποιεί Local Response Normalization (LRN), καθώς αυτό αυξάνει την κατανάλωση μνήμης και τον χρόνο εκπαίδευσης, χωρίς να βελτιώνει την ακρίβεια.

Πλήρως Συνδεδεμένα Επίπεδα (Fully Connected Layers): Το VGGNet περιλαμβάνει 3 πλήρως συνδεδεμένα επίπεδα. Τα δύο πρώτα έχουν από 4096 κανάλια, ενώ το τρίτο έχει 1000 κανάλια, ένα για κάθε κατηγορία της ImageNet.



Σχήμα 3.15 Αρχιτεκτονική VGG

3.4.2. ResNet

Το ResNet είναι συντομογραφία του Residual Network (Δικτύου Υπολοίπου)[44]. Πρόκειται για μια καινοτόμα αρχιτεκτονική νευρωνικού δικτύου που παρουσιάστηκε για πρώτη φορά από τους Kaiming He, Xiangyu Zhang, Shaoqing Ren και Jian Sun στο ερευνητικό τους άρθρο του 2015 με τίτλο “Deep Residual Learning for Image Recognition” [45].

Το μοντέλο σημείωσε τεράστια επιτυχία, γεγονός που επιβεβαιώνεται από τη νίκη του στον διαγωνισμό ταξινόμησης του ILSVRC 2015 (ImageNet Large Scale Visual Recognition Challenge), όπου το σύνολο των μοντέλων του πέτυχε ποσοστό σφάλματος στην εκπαίδευση μόλις 3.57%.

Το ResNet διαθέτει πολλές παραλλαγές που βασίζονται στην ίδια βασική αρχή αλλά διαφέρουν στον αριθμό των επιπέδων και στη χρήση στρώσεων pooling. Η παραλλαγή ResNet50, για παράδειγμα, περιλαμβάνει 50 επίπεδα νευρωνικού δικτύου.

Το ResNet50 είναι η πιο ισορροπημένη και διαδομένη παραλλαγή της οικογένειας ResNet, καθιστώντας το ιδανική επιλογή για μεταφορά μάθησης και γενικά προβλήματα ταξινόμησης. Δεν είναι το "καλύτερο" σε κάθε μέτρηση, αλλά συχνά αποτελεί την πιο έξυπνη προεπιλεγμένη επιλογή.

Η πρώτη αρχιτεκτονική ResNet ήταν το ResNet-34, το οποίο εισήγαγε τις λεγόμενες συνδέσεις συντόμευσης (shortcut connections), μετατρέποντας ένα απλό συνελκτικό δίκτυο (plain network) στο αντίστοιχο υπολοίπου (residual) δίκτυο. Το βασικό δίκτυο πάνω στο οποίο βασίστηκε το ResNet-34 ήταν εμπνευσμένο από τα VGG δίκτυα, με χρήση φίλτρων 3×3 στις συνελκτικές στρώσεις. Ωστόσο, σε σύγκριση με τα VGG δίκτυα, τα ResNet διαθέτουν λιγότερα φίλτρα και μικρότερη υπολογιστική πολυπλοκότητα.

Το ResNet-34 αποτελείται από 34 στρώσεις με βάρη (weighted layers).

Η σχεδίαση του ResNet-34 βασίστηκε σε δύο απλούς κανόνες:

- Οι στρώσεις που παρήγαγαν χάρτες χαρακτηριστικών ίδιου μεγέθους χρησιμοποιούσαν τον ίδιο αριθμό φίλτρων.
- Όταν το μέγεθος του χάρτη χαρακτηριστικών μειωνόταν στο μισό, ο αριθμός των φίλτρων διπλασιαζόταν, ώστε να διατηρείται η χρονική πολυπλοκότητα ανά στρώση.

Οι συνδέσεις συντόμευσης (shortcuts) προστέθηκαν πάνω σε αυτό το βασικό δίκτυο. Όταν οι διαστάσεις εισόδου και εξόδου ήταν ίδιες, γινόταν άμεση χρήση της συντόμευσης ταυτότητας (identity shortcut). Όταν οι διαστάσεις αυξάνονταν, υπήρχαν δύο επιλογές:

- είτε η συντόμευση συνέχιζε ως ταυτότητα και προστίθεντο μηδενικές τιμές (zero-padding) για την αντιστοίχιση των διαστάσεων,
- είτε χρησιμοποιούταν μια συντόμευση μέσω προβολής (projection shortcut) για να γίνει προσαρμογή των διαστάσεων μέσω γραμμικής μετασχηματιστικής στρώσης.

Αυτή η προσέγγιση με τα residual μπλοκ και τις συνδέσεις συντόμευσης ήταν το στοιχείο που έκανε τα ResNet μοντέλα τόσο βαθιά και ταυτόχρονα εύκολα να εκπαιδευτούν.

Στην περίπτωση του ResNet50, το βασικό δομικό στοιχείο (building block) της αρχιτεκτονικής τροποποιήθηκε σε σχεδίαση bottleneck λόγω ανησυχιών για τον χρόνο εκπαίδευσης των επιπέδων.

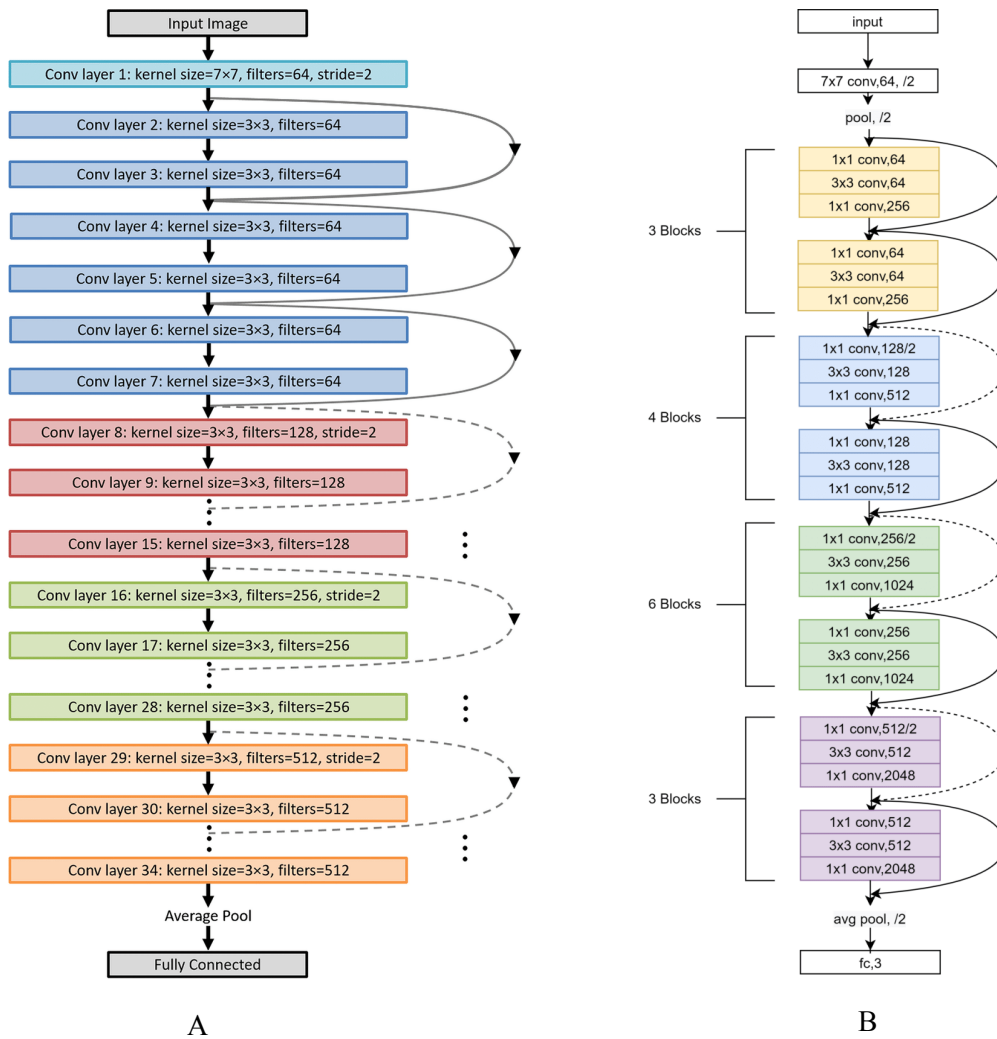
Συγκεκριμένα, αντί για τα 2 επίπεδα που υπήρχαν στα blocks του ResNet-34, στο ResNet-50 κάθε block αποτελείται από στοίβα 3 επιπέδων.

Έτσι, κάθε 2-layer block του ResNet-34 αντικαταστάθηκε με ένα 3-layer bottleneck block, σχηματίζοντας την αρχιτεκτονική του ResNet-50. Αυτό το μοντέλο παρουσίασε σημαντικά υψηλότερη ακρίβεια σε σύγκριση με το 34-layer ResNet, διατηρώντας παράλληλα διαχειρίσιμο χρόνο εκπαίδευσης χάρη στη bottleneck προσέγγιση.

Το ResNet50 αποτελείται από 50 συνελκτικά επίπεδα και ενσωματώνει residual συνδέσεις με shortcuts, οι οποίες βοηθούν το μοντέλο να διαχειρίζεται πιο αποτελεσματικά το πρόβλημα της εξασθένισης του βαθμίδωτου (vanishing gradient) και να εκπαιδεύει βαθύτερες αρχιτεκτονικές με επιτυχία.

Η αρχιτεκτονική του ResNet50 διαιρείται σε στάδια, κάθε ένα από τα οποία περιλαμβάνει μια ακολουθία από convolutional blocks και identity blocks. Κάθε convolutional block περιλαμβάνει τρεις συνελκτικές στρώσεις, ενώ κάθε identity block περιλαμβάνει μόνο μία βασική σύνδεση με ταυτότητα (identity mapping) χωρίς αλλαγή διαστάσεων.

Το τελευταίο στρώμα του ResNet50 είναι ένα fully connected layer, το οποίο αναλαμβάνει την τελική ταξινόμηση των χαρακτηριστικών που έχουν εξαχθεί μέσω των προηγούμενων επιπέδων του δικτύου. Αυτή η δομή καθιστά το ResNet50 ένα ιδιαίτερα αποτελεσματικό εργαλείο για εργασίες ταξινόμησης εικόνων[43].



Σχήμα 3.16 Αρχιτεκτονική ResNet
 (A) ResNet-34 (B) ResNet-50

3.4.3. Inception (GoogLeNet)

Το GoogLeNet, που κυκλοφόρησε το 2014, έθεσε νέα πρότυπα στην ταξινόμηση και ανίχνευση αντικειμένων, χάρη στην καινοτόμα προσέγγισή του. Είναι ιδιαίτερα γνωστό για τη χρήση των λεγόμενων Inception modules, τα οποία αποτελούν τα δομικά του στοιχεία.

Κάθε Inception module περιλαμβάνει παράλληλες συνελκτικές διεργασίες με φίλτρα διαφορετικών μεγεθών (1x1, 3x3 και 5x5) μέσα στο ίδιο επίπεδο. Στη συνέχεια, τα εξαγόμενα χαρακτηριστικά (outputs) από τα διαφορετικά φίλτρα συνενώνονται (concatenated).

Αυτός ο συνδυασμός εξόδων από φίλτρα διαφορετικών κλιμάκων δημιουργεί πλουσιότερες αναπαραστάσεις των χαρακτηριστικών, ενισχύοντας την ικανότητα του δικτύου να κατανοεί πληροφορίες σε διαφορετικά επίπεδα λεπτομέρειας. Η προσέγγιση αυτή επέτρεψε στο GoogLeNet να επιτύχει εξαιρετική ακρίβεια με σχετικά περιορισμένο αριθμό παραμέτρων, καθιστώντας το ένα από τα πιο σημαντικά μοντέλα στην ιστορία της υπολογιστικής όρασης[46].

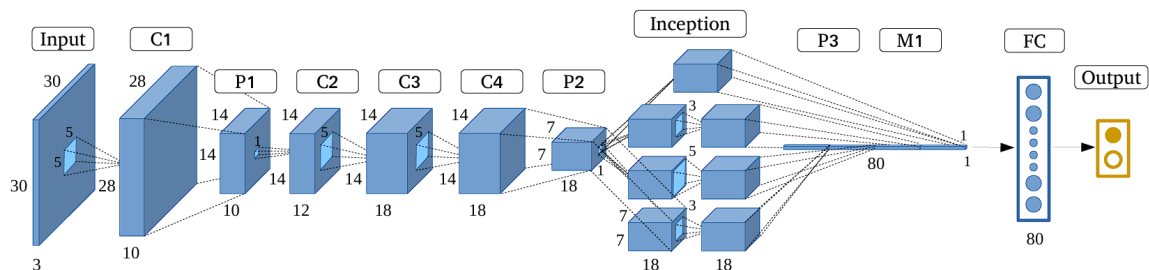
Μετά την επιτυχία του GoogLeNet (Inception v1), αναπτύχθηκαν διάφορες παραλλαγές της αρχιτεκτονικής Inception, με στόχο τη βελτίωση της ακρίβειας, της αποδοτικότητας και της ευκολίας εκπαίδευσης. Η σημαντικότερη, από άποψη ακρίβειας σε συνδυασμό με υπολογιστικό κόστος, είναι το InceptionV3

Το InceptionV3 εστιάζει κυρίως στη μείωση της υπολογιστικής κατανάλωσης, μέσω τροποποιήσεων στις προηγούμενες αρχιτεκτονικές Inception[47]. Η ιδέα αυτή παρουσιάστηκε στο άρθρο “Rethinking the Inception Architecture for Computer Vision”, που δημοσιεύθηκε το 2015 και συνυπογράφηκε από τους Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe και Jonathon Shlens[48].

Σε σύγκριση με το VGGNet, τα Inception Networks (όπως το GoogLeNet/Inception v1) αποδείχθηκαν πολύ πιο αποδοτικά όσον αφορά τον αριθμό των παραμέτρων και το συνολικό υπολογιστικό κόστος (μήμη, πόροι, χρόνος).

Βασικά Στοιχεία της Αρχιτεκτονικής:

- Inception modules: Κάθε module περιλαμβάνει παράλληλες διαδρομές με διαφορετικά είδη συνελκτικών και pooling επιπέδων. Αυτό ενισχύει την πληρότητα των χαρακτηριστικών που εξάγονται.
- 1×1 Convolutions: Χρησιμοποιούνται για μείωση της διάστασης (dimensionality reduction), μειώνοντας τις παραμέτρους και το υπολογιστικό κόστος.
- Auxiliary classifiers: Πρόσθετοι βοηθητικοί ταξινομητές ενσωματώνονται στο ενδιάμεσο του δικτύου κατά την εκπαίδευση, ώστε να βελτιώσουν τη ροή του gradient.
- Factorization: Αντί για μεγάλα φίλτρα (π.χ. 5×5), χρησιμοποιούνται συνδυασμοί μικρότερων (π.χ. 1×5 και 5×1), ώστε να μειώνεται το κόστος υπολογισμού διατηρώντας την αποτελεσματικότητα.
- Global Average Pooling: Στο τέλος του δικτύου αντί για πλήρως συνδεδεμένα layers, εφαρμόζεται παγκόσμια μέση τιμή (average pooling), μειώνοντας τον αριθμό παραμέτρων και περιορίζοντας το overfitting[46].



Σχήμα 3.17 Αρχιτεκτονική Inception

3.4.4. EfficientNet

Το EfficientNet είναι ένα συνελκτικό νευρωνικό δίκτυο (CNN) που βασίζεται στην έννοια της "συνδυαστικής κλιμάκωσης" (compound scaling). Η προσέγγιση αυτή στοχεύει στην αντιμετώπιση του διαχρονικού συμβιβασμού μεταξύ μεγέθους μοντέλου, ακρίβειας και υπολογιστικής απόδοσης.

Η βασική ιδέα πίσω από το compound scaling είναι ότι κλιμακώνει ταυτόχρονα και συντονισμένα τις τρεις θεμελιώδεις διαστάσεις ενός νευρωνικού δικτύου:

- Depth (βάθος): Πόσο βαθύ είναι το δίκτυο – πόσες στρώσεις περιέχει.
- Width (πλάτος): Πόσα κανάλια/φίλτρα έχει κάθε συνελκτικό επίπεδο.
- Resolution (ανάλυση): Το μέγεθος (σε pixels) της εικόνας εισόδου.

Η διαδικασία στο EfficientNet ξεκινά με ένα βασικό μοντέλο (baseline model), το οποίο λειτουργεί ως σημείο εκκίνησης. Πρόκειται για ένα μέτριο μεγέθους νευρωνικό δίκτυο που έχει καλή απόδοση σε ένα συγκεκριμένο πρόβλημα, αλλά δεν είναι βέλτιστο από άποψη υπολογιστικής αποδοτικότητας.

Στη συνέχεια, εισάγεται ένας συνδυαστικός συντελεστής (compound coefficient), ο οποίος είναι μια παραμετροποιήσιμη τιμή από τον χρήστη. Πρόκειται για έναν μοναδικό βαθμωτό παράγοντα (ϕ), που καθορίζει πόσο θα κλιμακωθούν τα βασικά χαρακτηριστικά του δικτύου: το πλάτος (width), το βάθος (depth) και η ανάλυση εισόδου (resolution). Με τη ρύθμιση του ϕ μπορούμε να ελέγξουμε το συνολικό μέγεθος, την πολυπλοκότητα και τις απαιτήσεις σε πόρους του μοντέλου.

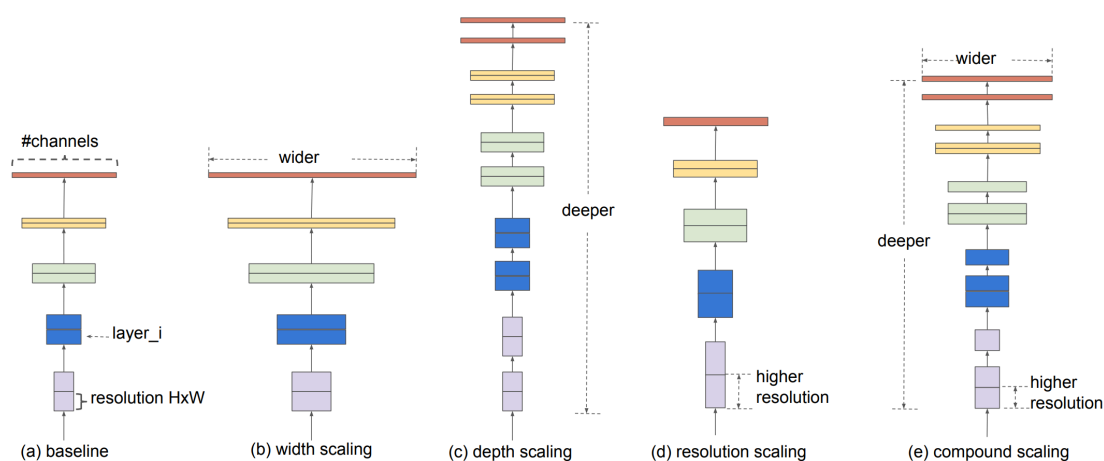


Figure 2. **Model Scaling.** (a) is a baseline network example; (b)-(d) are conventional scaling that only increases one dimension of network width, depth, or resolution. (e) is our proposed compound scaling method that uniformly scales all three dimensions with a fixed ratio.

Σχήμα 3.18 Αρχιτεκτονική EfficientNet
[50]

Το EfficientNet, μέσω της συνδυαστικής κλιμάκωσης (compound scaling), επηρέασε τον τρόπο με τον οποίο κατανοούμε τη σχέση μεταξύ αποδοτικότητας και ακρίβειας στη βαθιά μάθηση. Αντί να αυξάνει μόνο το βάθος ή το πλάτος του δικτύου, το μοντέλο αυτό κλιμακώνει ισορροπημένα τις τρεις βασικές διαστάσεις.

Αυτή η προσέγγιση αποδεικνύει ότι είναι δυνατό να επιτευχθεί υψηλή ακρίβεια χωρίς τεράστιο υπολογιστικό κόστος, κάτι που καθιστά το EfficientNet ιδανική επιλογή για εφαρμογές όπου η υπολογιστική αποδοτικότητα είναι εξίσου σημαντική με την προβλεπτική ισχύ.

3.5.Ερμηνευσιμότητα

Τα μοντέλα μηχανικής μάθησης, ειδικά τα βαθιά νευρωνικά δίκτυα, έχουν υψηλή ακρίβεια αλλά είναι «μαύρα κουτιά». Αυτό δημιουργεί δυσπιστία, ιδιαίτερα σε κρίσιμους τομείς όπως η υγεία. Έτσι, αναδύεται η ανάγκη για Explainable AI (XAI), δηλαδή μεθόδους που επιτρέπουν να κατανοήσουμε πώς και γιατί το μοντέλο έβγαλε μια απόφαση[51].

Η ερμηνευσιμότητα (interpretability) αποτελεί κρίσιμη παράμετρο στην ανάπτυξη μοντέλων βαθιάς μάθησης, ιδιαίτερα όταν αυτά χρησιμοποιούνται σε ευαίσθητες εφαρμογές όπως η διάγνωση ασθενειών από ιστοπαθολογικές εικόνες. Ερμηνεύσιμο θεωρείται ένα μοντέλο όταν μπορούμε να κατανοήσουμε πώς καταλήγει σε μία πρόβλεψη ή απόφαση. Η ανάγκη αυτή προκύπτει τόσο για λόγους εμπιστοσύνης του τελικού χρήστη, όσο και για την αναγνώριση πιθανών λαθών στο σύστημα. Οι τεχνικές ερμηνευσιμότητας μοντέλων ειδικά στην ταξινόμηση εικόνων αποτελούν απαραίτητο εργαλείο.

3.5.1. Χάρτες Σαφήνειας (Saliency Maps)

Οι Saliency Maps αποτελούν μία από τις πρώτες και πιο απλές τεχνικές ερμηνευσιμότητας για βαθιά νευρωνικά δίκτυα που εφαρμόζονται σε εικόνες. Η βασική ιδέα είναι να υπολογιστούν οι παραγώγοι της εξόδου της κλάσης ενδιαφέροντος ως προς κάθε pixel της εικόνας εισόδου. Αυτή η διαδικασία αποκαλύπτει πόσο ευαίσθητη είναι η τελική πρόβλεψη στις μεταβολές κάθε pixel. Το αποτέλεσμα οπτικοποιείται ως ένας χάρτης σαφήνειας (saliency map), στον οποίο τα pixel με υψηλές τιμές υποδεικνύουν περιοχές σημαντικές για την απόφαση του μοντέλου.

Το κύριο πλεονέκτημα αυτής της μεθόδου είναι η ταχύτητα και απλότητα της υλοποίησής της. Δεν απαιτεί κάποια ειδική τροποποίηση στο δίκτυο και μπορεί να εφαρμοστεί άμεσα σε οποιοδήποτε προκαταρτισμένο μοντέλο. Ωστόσο, οι χάρτες σαφήνειας συχνά είναι θορυβώδεις και δυσνόητοι, ιδιαίτερα όταν χρησιμοποιούνται σε πολύπλοκα δίκτυα. Επίσης, η εξάρτηση από τις πρώτες παραγώγους καθιστά τη μέθοδο ευαίσθητη σε μικρές αλλαγές στα δεδομένα εισόδου και μπορεί να οδηγήσει σε ασυνεπή αποτελέσματα.

Η μέθοδος παρουσιάστηκε από τον Simonyan et al. (2014) στο άρθρο “Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps”, και αποτέλεσε τη βάση για πολλές επόμενες ερμηνευτικές τεχνικές που αναπτύχθηκαν στη συνέχεια[52].

3.5.2. Ενσωματωμένες Κλίσεις (Integrated Gradients)

Οι Integrated Gradients αποτελούν μία πιο εξελιγμένη και θεωρητικά θεμελιωμένη μέθοδο ερμηνευσιμότητας, που απαντά σε αδυναμίες των απλών παραγώγων. Η βασική ιδέα είναι ότι η σημασία κάθε εισόδου (pixel) δεν πρέπει να υπολογίζεται μόνο στο σημείο της εικόνας, αλλά κατά μήκος μιας ολόκληρης "διαδρομής" από μια βασική (baseline) εικόνα προς την πραγματική εικόνα εισόδου. Καθώς μεταβαίνουμε σταδιακά από την baseline προς την εικόνα, υπολογίζονται και ενσωματώνονται οι παραγώγοι της εξόδου ως προς τα pixel.

Αυτή η προσέγγιση οδηγεί σε πιο σταθερές και συνεπείς ερμηνείες, που είναι λιγότερο ευαίσθητες στο θόρυβο και αντανακλούν καλύτερα τη συνολική συμβολή κάθε pixel στην τελική απόφαση. Ένα από τα βασικά της πλεονεκτήματα είναι ότι ικανοποιεί σημαντικές αξιωματικές ιδιότητες, όπως Ευαισθησία και implementation invariance. Παρόλα αυτά, έχει και κάποιες προκλήσεις: απαιτεί πολλαπλούς υπολογισμούς για κάθε εικόνα, και τα αποτελέσματα εξαρτώνται σε κάποιο βαθμό από την επιλογή της baseline εικόνας.

Η μέθοδος εισήχθη από τους Sundararajan et al. (2017) στο άρθρο "Axiomatic Attribution for Deep Networks"[53], το οποίο απέδειξε θεωρητικά ότι οι Integrated Gradients είναι η μοναδική μέθοδος που ικανοποιεί τα βασικά αιτήματα ερμηνευσιμότητας χωρίς να παραβιάζει τη συμπεριφορά του μοντέλου.

3.5.3. Grad-CAM

Συγκεκριμένα στην ταξινόμηση εικόνων, η πιο ευρέως χρησιμοποιημένη και πρακτική μέθοδος οπτικοποίησης είναι οι Gradient-weighted Class Activation Maps (Grad-CAM). Σε ένα συνελκτικό νευρωνικό δίκτυο, το Grad-CAM αξιοποιεί τους χάρτες χαρακτηριστικών του τελευταίου συνελκτικού επιπέδου και υπολογίζει τη συμβολή κάθε χωρικής τοποθεσίας στην πρόβλεψη μιας συγκεκριμένης κατηγορίας. Βασίζεται στις παραγώγους (gradients) που υπολογίζονται μέσω της οπισθοδιάδοσης (backpropagation).

Ειδικότερα, στο τελευταίο συνελκτικό στρώμα, υπολογίζονται τα gradients των χαρακτηριστικών για την κλάση με την υψηλότερη πρόβλεψη. Στη συνέχεια εφαρμόζεται μέσος όρος (global average) σε αυτά, ώστε να προκύψουν βάρη (weights). Το εσωτερικό γινόμενο αυτών των weights με τους αντίστοιχους χάρτες χαρακτηριστικών δίνει την έξοδο του Grad-CAM. Για να εστιάσουμε μόνο στα θετικά στοιχεία που συμβάλλουν στην απόφαση του μοντέλου, εφαρμόζεται συνάρτηση ενεργοποίησης ReLU (Rectified Linear Unit), αποκλείοντας τις αρνητικές τιμές.

Τέλος, οι διαστάσεις του χάρτη Grad-CAM που προκύπτει αλλάζουν έτσι ώστε να ταιριάζει με αυτές της αρχικής εικόνας, ώστε να μπορεί να υπερτεθεί σε αυτήν, προσφέροντας έτσι οπτική ερμηνεία της απόφασης του μοντέλου.

Η μέθοδος Grad-CAM, όπως περιγράφεται από τους Selvaraju et al. (2017)[54], στηρίζεται στον υπολογισμό της παραγώγου του σκορ y^c μιας κατηγορίας c ως προς τους χάρτες χαρακτηριστικών του τελευταίου συνελκτικού επιπέδου A^k .

Η βασική ιδέα είναι ότι η χωρική συνεισφορά κάθε καναλιού του συνελκτικού επιπέδου προς την τελική πρόβλεψη μπορεί να εκτιμηθεί με τη βοήθεια των παραγώγων:

$$\frac{\partial y^c}{\partial A^k} \quad (3.5.1)$$

Στη γενική περίπτωση ενός πολυκατηγορικού προβλήματος, όπου η έξοδος του μοντέλου αποτελείται από λογικές τιμές (logits) που κανονικοποιούνται μέσω softmax, το y^c θεωρείται το score πριν από την εφαρμογή της συνάρτησης softmax, ώστε οι παράγωγοι να εκφράζουν την άμεση συμβολή των χαρακτηριστικών στην απόφαση, χωρίς παρεμβολή της κανονικοποίησης.

Οπτικοποίηση

Η διαδικασία οπτικοποίησης περιλαμβάνει τα εξής στάδια:

1. Υπολογισμός παραγώγων: Για μια επιλεγμένη κλάση ενδιαφέροντος, υπολογίζονται οι παράγωγοι της εξόδου ως προς τους χάρτες ενεργοποίησης του τελευταίου συνελκτικού επιπέδου του δικτύου.
2. Global Average Pooling: Οι υπολογισμένες παράγωγοι υφίστανται καθολικό μέσο όρο (Global Average Pooling) σε κάθε κανάλι ξεχωριστά, ώστε να εξαχθούν οι αντίστοιχοι συντελεστές σημασίας (weights) που αποδίδουν τη σχετική συμβολή κάθε χάρτη ενεργοποίησης στην τελική πρόβλεψη.

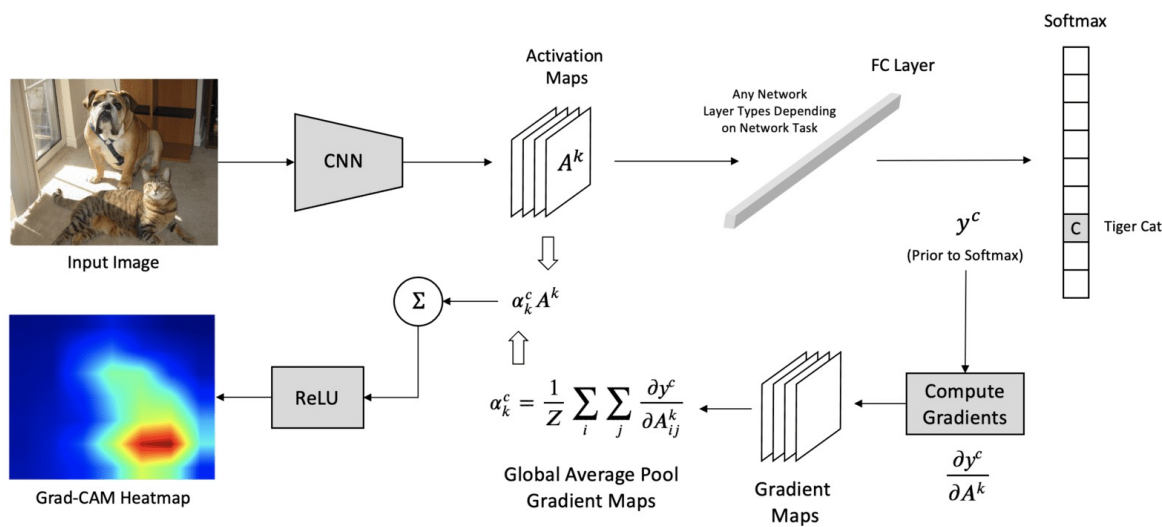
$$\alpha_k^c = \frac{1}{Z} \sum_i \sum_j \frac{\partial y^c}{\partial A_{ij}^k} \quad (3.5.2.)$$

Όπου α_k^c οι συντελεστές σημασίας ανά κανάλι.

3. Σύνθεση Χάρτη Θερμότητας (Heatmap): Οι χάρτες ενεργοποίησης πολλαπλασιάζονται με τα αντίστοιχα βάρη και συνδυάζονται γραμμικά. Στο τελικό αποτέλεσμα εφαρμόζεται η συνάρτηση ReLU, ώστε να διατηρηθούν μόνο οι θετικές συνεισφορές, δηλαδή τα χαρακτηριστικά που ενισχύουν την παρουσία της υπό εξέταση κλάσης.

$$L_{\text{Grad-CAM}}^c = \text{ReLU} \left(\sum_k \alpha_k^c A^k \right) \quad (3.5.3)$$

4. Οπτικοποίηση: Ο προκύπτων θερμικός χάρτης επαναδειγματοληπτείται (resampled) ώστε να συμφωνεί με τις διαστάσεις της αρχικής εικόνας και επικαλύπτεται επάνω της, δημιουργώντας μια οπτικοποίηση που τονίζει τις περιοχές της εικόνας που επηρέασαν περισσότερο την απόφαση του μοντέλου.



Σχήμα 3.19 Διάγραμμα Ροής Μεθόδου Grad-CAM

Ένα από τα σημαντικότερα πλεονεκτήματα της μεθόδου Grad-CAM είναι η δυνατότητα εφαρμογής της σε ήδη εκπαιδευμένα συνελκτικά νευρωνικά δίκτυα χωρίς την ανάγκη τροποποίησης της αρχιτεκτονικής ή επανεκπαίδευσης του μοντέλου. Παρά την εύκολη του εφαρμογή, το Grad-CAM παρουσιάζει ορισμένους σημαντικούς περιορισμούς. Αρχικά, η ποιότητα των παραγόμενων heatmaps εξαρτάται σε μεγάλο βαθμό από την επιλογή του συνελκτικού επιπέδου που χρησιμοποιείται ως βάση για την ερμηνεία. Εάν το επιλεγμένο επίπεδο είναι πολύ πρώιμο, οι ενεργοποιήσεις είναι γενικές και ασαφείς. Εάν είναι πολύ βαθύ, η χωρική πληροφορία ενδέχεται να έχει περιοριστεί σημαντικά. Επιπλέον, η μέθοδος τείνει να εντοπίζει κυρίως περιοχές με σχετικώς απλή συσχέτιση με την τελική πρόβλεψη, γεγονός που περιορίζει την ερμηνευτική της ισχύ σε περιπτώσεις όπου η απόφαση του δικτύου βασίζεται σε πολύπλοκες ή μη γραμμικές αλληλεπιδράσεις.

Η ιατρική απεικόνιση αποτελεί έναν από τους σημαντικότερους τομείς εφαρμογής της μεθόδου Grad-CAM, καθώς απαντά στην επιτακτική ανάγκη για ερμηνευσιμότητα των προβλέψεων σε κλινικά κρίσιμα περιβάλλοντα. Συστήματα βαθιάς μάθησης έχουν αναδειχθεί ως εξαιρετικά αποτελεσματικά στην αυτόματη ανάλυση ιατρικών εικόνων, προσφέροντας ακρίβεια που ανταγωνίζεται ή και ξεπερνά την ανθρώπινη απόδοση σε συγκεκριμένες διαγνώσεις.

Επιπλέον, έχει φανεί χρήσιμο στην ανίχνευση όγκων, φλεγμονών ή ανωμαλιών σε εγκεφαλικές MRI, ενισχύοντας τη διαγνωστική αξιοπιστία και προσφέροντας ένα επίπεδο διαφανούς συνεργασίας μεταξύ ανθρώπου και αλγορίθμου. Με αυτόν τον τρόπο, το Grad-CAM δεν ενισχύει μόνο την αποδοχή των τεχνολογιών τεχνητής νοημοσύνης από τους επαγγελματίες υγείας, αλλά συμβάλλει και στη δημιουργία συστημάτων υποστήριξης αποφάσεων που είναι πραγματικά αξιόπιστα και κλινικά εφαρμόσιμα.

3.6.Αβεβαιότητα

Η αβεβαιότητα στη βαθιά μάθηση αναφέρεται στην ικανότητα ενός μοντέλου να εκτιμά πόσο σίγουρο είναι για τις προβλέψεις του. Σε αντίθεση με τις απλές προβλέψεις που προκύπτουν από την έξοδο του softmax, η οποία μπορεί να δώσει υπερβολικά υψηλή εμπιστοσύνη, η εκτίμηση της αβεβαιότητας επιτρέπει στο μοντέλο να αναγνωρίζει πότε δεν είναι βέβαιο για μια απόφαση.

3.6.1. Bayesian Νευρωνικά Δίκτυα (Bayesian Neural Networks - BNNs)

Τα Bayesian Neural Networks αποτελούν μία προσέγγιση εκτίμησης αβεβαιότητας που εισάγει πιθανοκρατική θεώρηση στις παραμέτρους του μοντέλου. Αντί να μαθαίνονται σταθερές τιμές για τα βάρη ενός νευρωνικού δικτύου, το δίκτυο αυτό αναπαριστά τα βάρη ως κατανομές πιθανότητας. Κατά τη φάση της πρόβλεψης, το δίκτυο πραγματοποιεί δειγματοληψίες από αυτές τις κατανομές και υπολογίζει τόσο τη μέση πρόβλεψη όσο και την τυπική απόκλιση, η οποία χρησιμοποιείται ως ένδειξη της αβεβαιότητας του μοντέλου[55].

Με αυτόν τον τρόπο, οι Bayesian μέθοδοι επιτρέπουν την εκτίμηση γνωσιολογικής (epistemic) αβεβαιότητας, δηλαδή της αβεβαιότητας που προκύπτει από άγνοια ή ανεπαρκή γνώση στα δεδομένα εκπαίδευσης. Ωστόσο, τα BNNs είναι υπολογιστικά απαιτητικά και η εκπαίδευσή τους προϋποθέτει προσεγγίσεις όπως η παραμετρική προσέγγιση ολοκλήρωσης (variational inference) ή η προσέγγιση Laplace, γεγονός που καθιστά τη χρήση τους τεχνικά πιο σύνθετη.

3.6.2. Μοντέλα Ensemble (Ensemble Models)

Η μέθοδος των Ensemble μοντέλων βασίζεται στην εκπαίδευση πολλών ανεξάρτητων μοντέλων, καθένα από τα οποία έχει διαφορετική αρχικοποίηση ή εκπαιδεύεται σε διαφορετικό υποσύνολο των δεδομένων. Η τελική πρόβλεψη προκύπτει από τη μέση των προβλέψεων όλων των μοντέλων, ενώ η διασπορά των εξόδων χρησιμοποιείται για την εκτίμηση της γνωσιολογικής αβεβαιότητας.

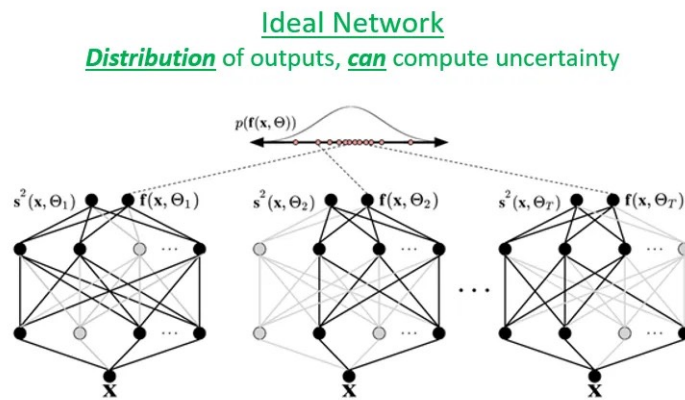
Πρόκειται για μία εμπειρικά αποδοτική προσέγγιση, η οποία έχει αποδειχθεί ιδιαίτερα αξιόπιστη στην πράξη. Ωστόσο, η βασική της αδυναμία είναι ο αυξημένος υπολογιστικός φόρτος, καθώς απαιτείται η συντήρηση και χρήση πολλών δικτύων ταυτόχρονα. Παρά ταύτα, παραμένει μία από τις πιο ισχυρές εναλλακτικές για εκτίμηση αβεβαιότητας σε προβλήματα ταξινόμησης εικόνας[56].

3.6.3. Monte Carlo Dropout

Η μέθοδος Monte Carlo Dropout (MC Dropout) αποτελεί μια απλή αλλά ισχυρή τεχνική για την προσέγγιση της Βαϋζιανής εκτίμησης αβεβαιότητας σε βαθιά νευρωνικά δίκτυα, η οποία παρουσιάστηκε από τους Yarin Gal και Zoubin Ghahramani (2016)[57]. Η βασική ιδέα πίσω από τη μέθοδο είναι ότι το dropout, όταν χρησιμοποιείται τόσο κατά την εκπαίδευση όσο και κατά τη φάση της πρόβλεψης, μπορεί να λειτουργήσει ως στοχαστικός μηχανισμός δειγματοληψίας από μια κατανομή πιθανών μοντέλων.

Κατά την εκπαίδευση, η MC Dropout δεν διαφέρει από τις κλασικές τεχνικές regularization. Σε κάθε forward pass απενεργοποιούνται τυχαία νευρώνες, περιορίζοντας την υπερπροσαρμογή και ενισχύοντας τη γενίκευση του μοντέλου. Ωστόσο, το καινοτόμο στοιχείο της μεθόδου έγκειται στο ότι κατά τη φάση της πρόβλεψης (test time), το dropout διατηρείται ενεργό, αντί να απενεργοποιείται όπως γίνεται συνήθως. Αυτό επιτρέπει στο μοντέλο να παράγει πολλαπλές στοχαστικές προβλέψεις για την ίδια είσοδο, μέσω διαδοχικών forward passes με διαφορετικά dropout masks.

Οι προβλέψεις που προκύπτουν από αυτές τις επαναλαμβανόμενες διεργασίες συνδυάζονται για την εκτίμηση της τελικής πρόβλεψης μέσω του μέσου όρου. Παράλληλα, η διακύμανση των προβλέψεων αξιοποιείται για την ποσοτικοποίηση της αβεβαιότητας του μοντέλου. Έτσι, η MC Dropout παρέχει ένα πρακτικό μέσο εκτίμησης της λεγόμενης επιστημικής αβεβαιότητας (epistemic uncertainty), η οποία σχετίζεται με την αβεβαιότητα για τις παραμέτρους του ίδιου του μοντέλου.



Σχήμα 3.20 Monte-Carlo Dropout

Μέση Πιθανότητα Πρόβλεψης (mean prediction)

Πρόκειται για τον μέσο όρο των T προβλέψεων που παράγει το μοντέλο για μια συγκεκριμένη είσοδο. Στην περίπτωση ταξινόμησης, το μοντέλο δίνει διάφορες πιθανότητες για μία συγκεκριμένη κλάση και ο μέσος όρος αυτών εκφράζει την τελική εκτίμηση του μοντέλου για την πιθανότητα να ανήκει η εικόνα σε αυτήν την κατηγορία. Αυτή η μέση τιμή είναι πιο σταθερή και ανθεκτική στον θόρυβο σε σχέση με μία μόνο πρόβλεψη.

Η τιμή της μέσης πρόβλεψης προκύπτει από την παρακάτω σχέση :

$$\mathbb{E}_{q(y^*|x^*)}(y^*) \approx \frac{1}{T} \sum_{t=1}^T \hat{y}^*(x^*, W_1^{(t)}, \dots, W_L^{(t)}) \quad (3.6.1)$$

Τυπική Απόκλιση των Προβλέψεων (prediction standard deviation)

Η τυπική απόκλιση αυτών των T τιμών μάς δίνει ένα μέτρο της διασποράς τους, δηλαδή του πόσο “σίγουρο” είναι το μοντέλο για αυτήν την απόφαση. Αν η τυπική απόκλιση είναι χαμηλή, σημαίνει ότι όλες οι προβλέψεις είναι παρόμοιες και το μοντέλο έχει υψηλή εμπιστοσύνη. Αν είναι υψηλή, σημαίνει ότι οι προβλέψεις διαφέρουν πολύ μεταξύ τους και υπάρχει μεγάλη αβεβαιότητα.

Η τιμή της αβεβαιότητας προκύπτει από την σχέση :

$$\text{Var}_{q(y^*|x^*)}(y^*) \approx \tau^{-1}I + \frac{1}{T} \sum_{t=1}^T \hat{y}_t^* \hat{y}_t^{*T} - \left(\frac{1}{T} \sum_{t=1}^T \hat{y}_t^* \right) \left(\frac{1}{T} \sum_{t=1}^T \hat{y}_t^* \right)^T \quad (3.6.2)$$

Και στην περίπτωση δυαδικής ταξινόμησης όπου το \hat{y}_t^* είναι απλός αριθμός και όχι διάνυσμα :

$$\text{Var}(y^*) \approx \frac{1}{T} \sum_{t=1}^T (\hat{y}_t^*)^2 - \left(\frac{1}{T} \sum_{t=1}^T \hat{y}_t^* \right)^2 \quad (3.6.1)$$

Η πρακτικότητα της MC Dropout έγκειται στο ότι μπορεί να εφαρμοστεί χωρίς τροποποιήσεις στην υπάρχουσα αρχιτεκτονική ενός νευρωνικού δικτύου, προσφέροντας μία "drop-in" λύση για την εκτίμηση αβεβαιότητας. Το μόνο που απαιτείται είναι η εκτέλεση πολλαπλών forward passes με ενεργό dropout, κάτι που είναι εύκολα εφαρμόσιμο σε υπάρχουσες υποδομές.

Συνοψίζοντας, η MC Dropout αποτελεί ένα υπολογιστικά ελαφρύ και εύχρηστο εργαλείο για την εκτίμηση αβεβαιότητας σε deep learning. Έχει ήδη αξιοποιηθεί σε κρίσιμους τομείς όπως η ιατρική διάγνωση, η ανάλυση εικόνας, και τα αυτόνομα συστήματα, αποδεικνύοντας ότι η απλότητα και η αποτελεσματικότητα μπορούν να συνυπάρχουν σε μια σύγχρονη μεθοδολογία μηχανικής μάθησης.

4. Σύγχρονες Μέθοδοι Ταξινόμησης Ιστοπαθολογικών Εικόνων Καρκίνου του Γαστρεντερικού Συστήματος

Η ταξινόμηση ιστοπαθολογικών εικόνων αποτελεί κομβικό εργαλείο για την πρόωμη διάγνωση και την εξατομικευμένη θεραπεία του καρκίνου του γαστρεντερικού συστήματος. Στην πρόσφατη ερευνητική βιβλιογραφία έχουν παρουσιαστεί ποικίλες μέθοδοι που αξιοποιούν τη δύναμη της βαθιάς μάθησης, των μηχανισμών προσοχής και των μετασχηματιστών για να βελτιώσουν την ακρίβεια και τη γενικευσιμότητα των μοντέλων.

Κοινό χαρακτηριστικό των περισσότερων προσεγγίσεων αποτελεί η αξιοποίηση προϋπαρχόντων συνελκτικών νευρωνικών δικτύων (CNNs), όπως τα VGG και ResNet είτε μέσω μεταφοράς μάθησης είτε ως βάση για την ανάπτυξη συνδυαστικών αρχιτεκτονικών. Πολλές από τις μελέτες αυτές εστιάζουν στην ενίσχυση της ακρίβειας μέσω υβριδικών μοντέλων, προσθέτοντας εξειδικευμένα επίπεδα ή εφαρμόζοντας τεχνικές ενισχυμένης μάθησης, οπτικής ερμηνευσιμότητας και εξαγωγής χειροποίητων χαρακτηριστικών.

Η ποιότητα και ποικιλομορφία των βάσεων δεδομένων αποτελούν κρίσιμη παράμετρο για την εκπαίδευση και αξιολόγηση μοντέλων βαθιάς μάθησης στην ταξινόμηση ιστοπαθολογικών εικόνων. Στην περίπτωση του καρκίνου του γαστρεντερικού συστήματος, χρησιμοποιούνται διάφορες δημόσιες και ιδιωτικές βάσεις, οι οποίες διαφέρουν ως προς τη μορφολογία των δεδομένων, το μέγεθος, την ανάλυση, τον τρόπο επισημείωσης και το στάδιο της ασθένειας που περιλαμβάνουν.

Η καταγραφή των μεθοδολογιών αυτών συμβάλλει στην καλύτερη κατανόηση των υφιστάμενων προσεγγίσεων, στην ανάδειξη των πλεονεκτημάτων και των περιορισμών τους, και θέτει τις βάσεις για τη συγκριτική αξιολόγηση της παρούσας εργασίας ως προς την απόδοσή της.

4.1.Γαστρικός Καρκίνος

4.1.1.Προσαρμοσμένα Συνελκτικά Νευρωνικά Δίκτυα (Custom CNNs)

Τα custom CNNs είναι δίκτυα που έχουν σχεδιαστεί ειδικά για συγκεκριμένες ιατρικές εφαρμογές, όπως η ταξινόμηση ιστοπαθολογικών εικόνων. Σε αντίθεση με τα γενικά μοντέλα που προεκπαιδεύονται σε μεγάλα datasets όπως το ImageNet, τα προσαρμοσμένα CNNs μπορούν να έχουν ελαφριά αρχιτεκτονική και να εκπαιδεύονται από το μηδέν πάνω σε ιατρικά δεδομένα, εκμεταλλευόμενα πλήρως τα ιδιαίτερα χαρακτηριστικά των ιστολογικών εικόνων. Τα custom CNNs σχεδιάζονται εξ αρχής για τις ιδιαιτερότητες των H&E εικόνων του γαστρικού ιστού, όπου τα ογκολογικά μορφολογικά σήματα (πυκνότητα πυρήνων, αδενική αρχιτεκτονική, ποικιλία χρώσης) εμφανίζονται σε πολύ μικρή κλίμακα.

4.1.2. Προεκπαιδευμένα CNNs

Η μεταφορά μάθησης (transfer learning) με χρήση προεκπαιδευμένων CNNs αποτελεί μια από τις πιο διαδεδομένες και αποτελεσματικές προσεγγίσεις στην ταξινόμηση ιστοπαθολογικών εικόνων γαστρικού καρκίνου. Τα δίκτυα αυτά έχουν αρχικά εκπαιδευτεί σε μεγάλα datasets γενικών εικόνων, όπως ImageNet και έπειτα προσαρμόζονται (fine-tuned) πάνω σε ιατρικά δεδομένα, αποφεύγοντας την ανάγκη για εκπαίδευση από το μηδέν.

Τα πιο συχνά χρησιμοποιούμενα προεκπαιδευμένα δίκτυα στην ανάλυση ιστοπαθολογικών εικόνων γαστρικού καρκίνου περιλαμβάνουν τα VGG16 και VGG19, τα οποία διαθέτουν απλή και βαθιά αρχιτεκτονική με διαδοχικά συνελκτικά και max-pooling επίπεδα. Χρησιμοποιούνται ευρέως λόγω της σταθερής και προβλέψιμης συμπεριφοράς τους. Το ResNet50, γνωστό για τις residual συνδέσεις (skip connections), επιτρέπει την εκπαίδευση βαθύτερων δικτύων χωρίς πρόβλημα εξαφάνισης του σφάλματος, και χρησιμοποιείται συχνά ως feature extractor. Το DenseNet121 διακρίνεται για τη δυνατότητα επαναχρησιμοποίησης χαρακτηριστικών μέσω των πυκνών συνδέσεων (dense blocks), προσφέροντας πλούσια ροή πληροφορίας μεταξύ των επιπέδων. Τέλος, τα InceptionV3 και Xception ενσωματώνουν πολυκλαδική συνελκτική επεξεργασία (multi-branch) και depthwise-separable convolutions, εξισορροπώντας την υπολογιστική αποδοτικότητα με την ικανότητα εξαγωγής πολύπλοκων χαρακτηριστικών. Όλα αυτά τα μοντέλα αξιοποιούνται είτε ως σταθεροί extractors είτε με fine-tuning, επιτυγχάνοντας υψηλές επιδόσεις σε συνδυασμό με κατάλληλη επεξεργασία των δεδομένων εισόδου.

4.1.3. Μηχανισμοί Προσοχής (Attention Mechanisms)

Οι μηχανισμοί προσοχής (attention mechanisms) έχουν καθιερωθεί ως μια κρίσιμη επέκταση των CNNs στην ανάλυση ιστοπαθολογικών εικόνων, επιτρέποντας στο μοντέλο να «επικεντρώνεται» στις πιο διαγνωστικά σημαντικές περιοχές μιας εικόνας και να φιλτράρει το περιττό υπόβαθρο. Εμπνευσμένοι από τον τρόπο που το ανθρώπινο οπτικό σύστημα δίνει έμφαση σε συγκεκριμένα ερεθίσματα, οι attention μηχανισμοί υλοποιούνται είτε στο χωρικό επίπεδο (spatial attention), είτε σε επίπεδο καναλιών (channel attention), είτε σε συνδυασμό και των δύο. Η ενσωμάτωσή τους σε συνελκτικά δίκτυα ενισχύει σημαντικά την ικανότητα του μοντέλου να ανιχνεύει πολύπλοκες ή διασκορπισμένες παθολογικές αλλοιώσεις.

Ένα από τα πιο προηγμένα συστήματα προσοχής που έχει εφαρμοστεί στον γαστρικό καρκίνο είναι το MCAM (Multi-Channel Attention Mechanism) framework[58]. Το μοντέλο αυτό συνδυάζει τρεις ξεχωριστούς μηχανισμούς προσοχής σε ένα συνελκτικό δίκτυο: το multi-scale global information channel, το spatial information channel, και το multi-scale spatial information channel. Κάθε κανάλι εστιάζει σε διαφορετική πτυχή της εικόνας, από ευρεία αρχιτεκτονική δομή έως τοπικά μοτίβα κυττάρων και υφής.

Οι έξοδοι των καναλιών συνδυάζονται μέσω voting scheme και το τελικό μοντέλο είναι ικανό να αναγνωρίζει κακοήθειες με ακρίβεια άνω του 99% σε datasets όπως το GasHisSDB και το HCRF. Επιπλέον, ενσωματώνει Grad-CAM visualizations, προσφέροντας διαφάνεια στην απόφαση του δικτύου και διευκολύνοντας τη χρήση του στην κλινική πράξη. Η μελέτη των Zubair et al. (2025) επιβεβαιώνει την υπεροχή του MCAM έναντι κλασικών προεκπαιδευμένων CNNs όπως το VGG16 και το Xception, καθιστώντας το μια από τις πλέον αξιόπιστες αρχιτεκτονικές στον τομέα.

4.1.4.Οπτικοί Μετασχηματιστές

Τα Vision Transformers (ViTs) αποτελούν μια σύγχρονη εναλλακτική στα CNNs, εισάγοντας την έννοια της παγκόσμιας προσοχής (global self-attention) στην επεξεργασία εικόνων. Σε αντίθεση με τα CNNs που περιορίζονται σε τοπικά receptive fields, οι transformers μπορούν να μοντελοποιήσουν εξαρτήσεις μεταξύ απομακρυσμένων περιοχών της εικόνας, κάτι ιδιαίτερα χρήσιμο για τις ιστοπαθολογικές εικόνες όπου η παθολογική πληροφορία είναι συχνά διασκορπισμένη. Οι transformers έχουν το πλεονέκτημα της μηχανιστικής γενίκευσης και σταθερότητας σε μεταβολές χρώσης ή υφής, γεγονός που τους καθιστά ιδιαίτερα χρήσιμους σε πραγματικές, ετερογενείς ιατρικές συνθήκες.

Ένα χαρακτηριστικό παράδειγμα εφαρμογής transformer αρχιτεκτονικής στον γαστρικό καρκίνο είναι το GasHis-Transformer, που προτάθηκε από τους Chen et al. (2022) [59]. Το μοντέλο υλοποιεί έναν υβριδικό σχεδιασμό, συνδυάζοντας τοπική πληροφορία μέσω ενός CNN (Local Information Module – LIM) και παγκόσμια πληροφορία μέσω ενός position-encoded transformer encoder (Global Information Module – GIM). Η επεξεργασία γίνεται σε παράλληλους κλάδους, ενώ η τελική έξοδος συγχωνεύεται για τη λήψη της τελικής απόφασης.

Το μοντέλο αξιολογήθηκε σε H&E patches από τη βάση δεδομένων HE-GHI-DS και επέδειξε υψηλή ακρίβεια, 98.55%, και ανθεκτικότητα, με προσθήκη και lightweight παραλλαγής μέσω DropConnect για μείωση του υπολογιστικού κόστους. Η καινοτομία του GasHis-Transformer έγκειται στην πολυκλιμακική ανάλυση και συγχώνευση τοπικών-παγκόσμιων χαρακτηριστικών, κάτι που δεν προσφέρεται από κλασικά CNNs ή καθαρούς transformers, και καθιστά το μοντέλο αυτό ιδανικό για εφαρμογές με ποικίλες μορφολογικές προκλήσεις.

4.2.Καρκίνος του Παχέος Εντέρου

4.2.1.Προεκπαιδευμένα CNNs

Στην ταξινόμηση ιστοπαθολογικών εικόνων από καρκίνο του παχέος εντέρου, όπως και από γαστρικό καρκίνο, τα προεκπαιδευμένα CNNs αποτελούν σταθερή και αποτελεσματική επιλογή, ειδικά όταν χρησιμοποιούνται σε συνδυασμό με δεδομένα patch-level. Τα πιο διαδεδομένα μοντέλα είναι τα VGG16, InceptionV3, ResNet50 και DenseNet121, λόγω της ικανότητάς τους να εξάγουν πλούσια μορφολογικά χαρακτηριστικά. Αυτά τα δίκτυα εφαρμόζονται είτε ως σταθεροί extractors, είτε με fine-tuning σε ένα ή περισσότερα blocks, ανάλογα με το μέγεθος και την ομοιογένεια του dataset.

Η χρήση τους συνοδεύεται συνήθως από τεχνικές βελτίωσης όπως data augmentation, color jittering, normalization με βάση τις χρωστικές H&E, και επιλογή κατάλληλων loss functions ώστε να αντιμετωπιστεί η ανισορροπία των κλάσεων. Τα CNNs αυτά προσφέρουν καλές baseline επιδόσεις και μπορούν εύκολα να ενσωματωθούν σε pipelines με ερμηνευσιμότητα και μετα-εκπαίδευση σε νέα δεδομένα από διαφορετικά νοσοκομεία.

4.2.2.Στρατηγικές Stacking

Η τεχνική stacking συνδυάζει τις προβλέψεις από διαφορετικά μοντέλα πρώτου επιπέδου ώστε να ενισχυθεί η τελική απόδοση μέσω ενός meta-classifier. Στην πράξη, κάθε CNN εκπαιδεύεται ανεξάρτητα και οι τελικές πιθανότητες από κάθε μοντέλο συγκεντρώνονται ως χαρακτηριστικά για έναν δεύτερο ταξινομητή, όπως ένας Support Vector Machine (SVM) ή λογιστική παλινδρόμηση.

Ένα χαρακτηριστικό παράδειγμα αυτής της προσέγγισης είναι η μελέτη των Gabralla et al. (2023)[43], στην οποία εφαρμόστηκε ένα στοιβαγμένο σχήμα χρησιμοποιώντας VGG16, InceptionV3, ResNet50, και DenseNet121 ως βασικούς “μαθητές”. Οι τελικές προβλέψεις από κάθε μοντέλο εισήχθησαν σε έναν SVM ταξινομητή, ο οποίος εκτελούσε την τελική απόφαση. Το σύστημα δοκιμάστηκε στο σύνολο δεδομένων LC25000, ένα από τα πιο ευρέως χρησιμοποιούμενα ιστοπαθολογικά σύνολα δεδομένων για την ανίχνευση καρκίνου του παχέος εντέρου, επιτυγχάνοντας 100% ακρίβεια υπερβαίνοντας κάθε μεμονωμένο μοντέλο. Το stacking αποδείχθηκε ιδιαίτερα αποτελεσματικό για την επίτευξη robust γενικευσιμότητας μεταξύ διαφορετικών μορφών καρκινικών ιστών και χρώσεων.

4.2.3.Μετασχηματιστές

Οι Vision Transformers και τα υβριδικά δίκτυα CNN-Transformer έχουν επίσης αρχίσει να παίζουν σημαντικό ρόλο στον καρκίνο του παχέος εντέρου, κυρίως λόγω της ικανότητάς τους να αποτυπώνουν μακροπρόθεσμες εξαρτήσεις και παγκόσμια χωρικά μοτίβα στα δεδομένα. Οι εικόνες ιστοπαθολογίας του εντέρου, που περιλαμβάνουν σύνθετες δομές όπως αδένες, πυρήνες και στρώμα, επωφελούνται ιδιαίτερα από την ικανότητα των transformers να ενσωματώνουν πληροφορία από ολόκληρη την εικόνα ταυτόχρονα.

Ένα εντυπωσιακό παράδειγμα αποτελεί το μοντέλο CCDNet (Coordinate Attention with Cross-shaped window Swin Transformer and Atrous Convolution)[60]. Πρόκειται για ένα υβριδικό δίκτυο που συνδυάζει συνελκτικά επίπεδα με atrous (dilated) convolution για την εξαγωγή πολυκλιμακικών χαρακτηριστικών, και ταυτόχρονα χρησιμοποιεί Swin Transformer με cross-shaped παράθυρα για την παγκόσμια κατανόηση της εικόνας. Η coordinate attention μονάδα επιτρέπει στο μοντέλο να εντοπίζει και να εστιάζει σε τοπικά κρίσιμες περιοχές με μεγαλύτερη ακρίβεια. Το μοντέλο αξιολογήθηκε σε μεγάλες βάσεις δεδομένων όπως το NCT-CRC-HE-100K και παρουσίασε εξαιρετικές επιδόσεις με ακρίβεια που ξεπέρασε το 98%, υποδεικνύοντας τη δυνατότητα των υβριδικών μοντέλων να συνδυάσουν την ακρίβεια των CNNs με την παγκόσμια κατανόηση των transformers.

5. Μεθοδολογία

5.1.Σύνολα Δεδομένων - Datasets

5.1.1.GasHisSDB

Το πρώτο dataset, στο οποίο έγινε το training του μοντέλου, είναι το GasHisSDB. Αυτό αποτελείται από 245.196 υπο-εικόνες (sub-size images) καταναμημένες σε δύο κατηγορίες, 97.076 “μη φυσιολογικών” εικόνων (Abnormal) και 148,120 φυσιολογικών (Normal), οι οποίες προέκυψαν από δύο στάδια επεξεργασίας[61].

Αρχικά, από το Νοσοκομείο Longhua του Πανεπιστημίου Κινέζικης Ιατρικής (Longhua Hospital Shanghai University of Traditional Chinese Medicine), τέσσερις παθολόγοι παρείχαν 600 εικόνες ιστοπαθολογίας γαστρικού καρκίνου μεγέθους 2048×2048 pixel και με βάση τον ιστό, ταξινόμησαν τις εικόνες σε Normal και Abnormal.

Στο δεύτερο στάδιο, βάσει των δεδομένων του πρώτου, πέντε ερευνητές βιοϊατρικής από το Northeastern University ετοίμασαν 245.196 υπο-εικόνες, χωρίζοντας τις 600 αρχικές σε μικρότερα κομμάτια. Έπειτα, δύο έμπειροι παθολόγοι από το Νοσοκομείο και Ινστιτούτο Καρκίνου της Liaoning πραγματοποίησαν τη βαθμονόμηση.

Προετοιμασία εικόνων

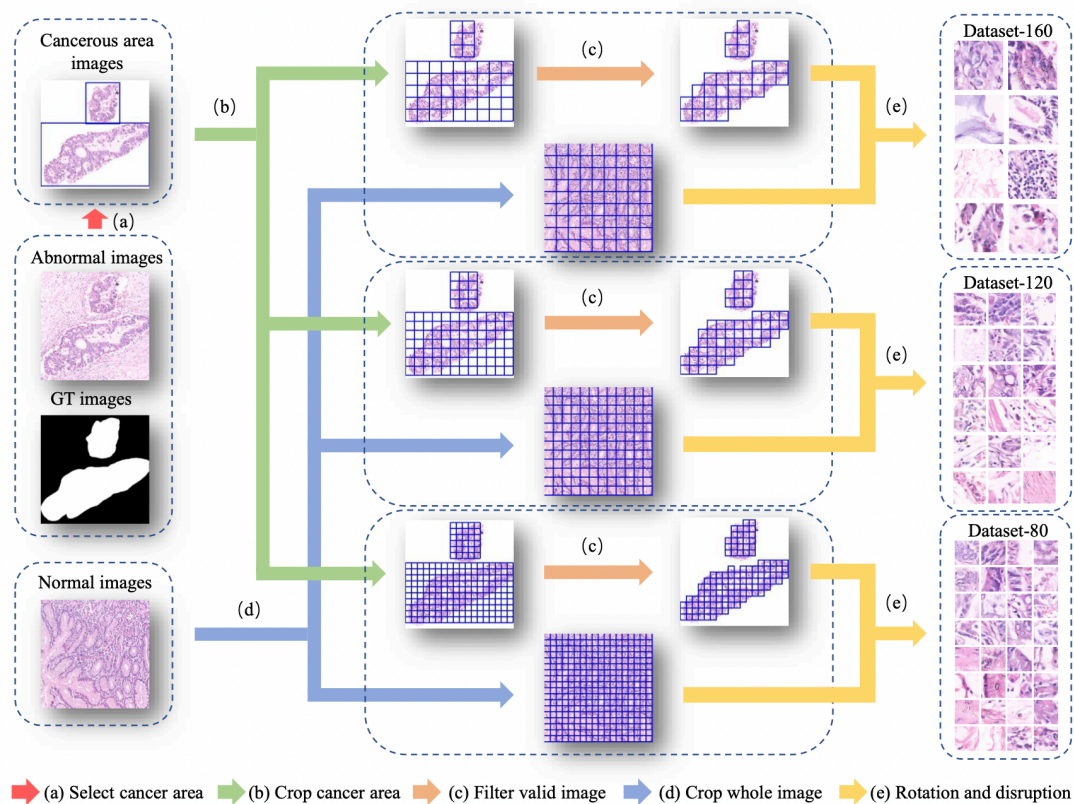
Οι αρχικές εικόνες χωρίστηκαν στο δεύτερο στάδιο σε 3 sub-datasets, σύμφωνα με τις διαστάσεις τους.

- 160 × 160 pixels - 33.284 εικόνες
- 120 × 120 pixels - 65.261 εικόνες
- 80 × 80 pixels - 146.651 εικόνες

Από τις φυσιολογικές ιστοπαθολογικές τομές, αποκόπτονται άμεσα οι εικόνες των παραπάνω διαστάσεων. Για τις καρκινικές τομές (Abnormal), είναι απαραίτητο να επιλέγεται η καρκινική περιοχή ως περιοχή ενδιαφέροντος (region of interest).

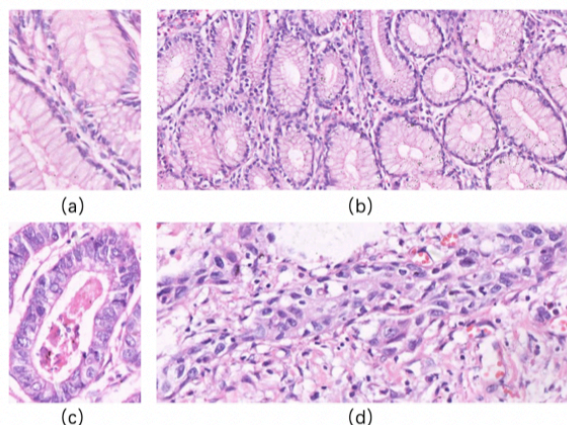
Η περιοχή ενδιαφέροντος και η ground truth επισημείωση γίνονται ταυτόχρονα. Αυτή η διαδικασία βοηθά στον αποκλεισμό εικόνων που περιέχουν ελάχιστες καρκινικές περιοχές, οι οποίες ενδέχεται να είναι μη αντιπροσωπευτικές.

Για τη μείωση της συσχέτισης μεταξύ υπο-εικόνων που προέρχονται από την ίδια αρχική εικόνα, κάθε υπο-εικόνα υφίσταται τυχαία περιστροφή, ενώ η σειρά των εικόνων στο σύνολο δεδομένων αναδιατάσσεται τυχαία (shuffling), ώστε να διασφαλιστεί η ανεξαρτησία μεταξύ των δειγμάτων κατά την εκπαίδευση.



Σχήμα 5.1 Διαδικασία Δημιουργίας Sub-Datasets

Κάθε Normal εικόνα δεν περιέχει καρκινικές περιοχές. Τα κύτταρα παρουσιάζουν καμία ή ελάχιστη άτυπη μορφολογία (Σχήμα 5.2 (a)). Επιπλέον, οι πυρήνες των κυττάρων δεν διασπώνται και είναι κανονικά διατεταγμένοι σε μία μόνο στοιβάδα (Σχήμα 5.2 (b)). Συνεπώς, κατά την παρατήρηση με οπτικό μικροσκόπιο, εφόσον δεν εντοπίζονται αλλοιώσεις κυττάρων ή ιστών και πληρούνται τα χαρακτηριστικά της φυσιολογικής εικόνας, μπορούμε να συμπεράνουμε ότι πρόκειται για φυσιολογικό δείγμα.



Σχήμα 5.2 Παραδείγματα Εικόνων GasHisSDB
 (a)(b) Φυσιολογικός ιστός (Normal)
 (c)(d) Μη φυσιολογικός (Abnormal)

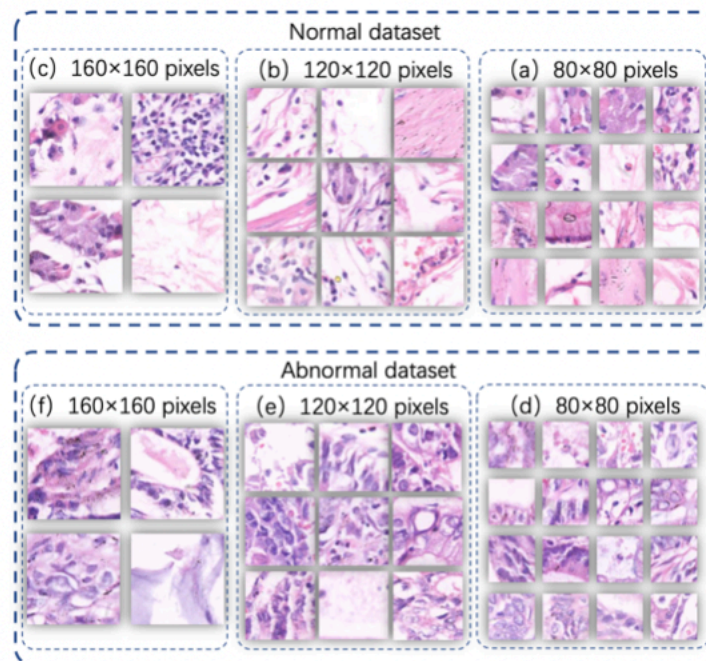
Κατά την προετοιμασία του συνόλου δεδομένων φυσιολογικών εικόνων, οι εικόνες αποκόπτονται απευθείας από ολόκληρη την αρχική εικόνα, ακριβώς λόγω αυτών των χαρακτηριστικών. Ενδεικτικά παραδείγματα των παραγόμενων εικόνων παρουσιάζονται στα μέρη (a), (b), (c) του Σχήματος 5.3.

Κάθε Abnormal εικόνα απεικονίζει γαστρικό καρκίνο. Η γενική μορφολογία του γαστρικού καρκίνου είναι κυρίως ελκωτικού τύπου (ulcer). Καθώς εξελίσσεται η νόσος, οι καρκινικές εστίες εισχωρούν από τον βλεννογόνο (mucosal) μέχρι τον μυϊκό και τον ορογόνο (serosal) χιτώνα. Ο όγκος έχει σκληρή υφή και η τομή του είναι συχνά γκρίζου-λευκού χρώματος.

Στο μικροσκόπιο, τα καρκινικά κύτταρα είναι σαφώς οριοθετημένα από το στρώμα και διατεταγμένα σε μορφή φωλεών, αδενίων, σωληναρίων ή νηματοειδών δομών (Σχήμα 5.2 (c)). Όταν όμως τα καρκινικά κύτταρα διεισδύσουν στο στρώμα, το όριο ανάμεσά τους γίνεται ασαφές (Σχήμα 5.2(d)).

Βάσει των παραπάνω, όταν παρατηρούνται αδενικές ή αδενοειδείς δομές με ανόμοιο μέγεθος, διαφορετικά σχήματα και ακανόνιστη διάταξη, τότε η παθολογική εικόνα κρίνεται ως μη φυσιολογική. Στις εικόνες αυτές, τα καρκινικά κύτταρα εμφανίζονται συχνά πολυεπίπεδα και ακανόνιστα διατεταγμένα, ενώ οι πυρήνες τους διαφέρουν σε μέγεθος και εμφανίζουν φαινόμενα μίτωσης.

Κατά τη διαδικασία δημιουργίας του Abnormal dataset, αποκόπτεται κάθε επιλεγμένη περιοχή καρκινικής αλλοίωσης και στη συνέχεια οι εικόνες φιλτράρονται, ώστε η καρκινική περιοχή να καταλαμβάνει συνήθως τουλάχιστον το 50% της εικόνας. Τα τελικά δείγματα παρουσιάζονται στα Σχήματα 5.3-(d), (e), (f).



Σχήμα 5.3 Sub-Datasets GasHisSDB

Για την εκπαίδευση και την τελική επιλογή παραμέτρων του μοντέλου επιλέχθηκε το υποσύνολο του GasHisSDB που περιλαμβάνει εικόνες διαστάσεων 120×120 pixels, το οποίο περιλαμβάνει συνολικά 65.261 δείγματα. Η επιλογή αυτή έγινε ώστε να επιτευχθεί μια ισορροπία μεταξύ επαρκούς πληροφορίας ανά εικόνα και υπολογιστικής αποδοτικότητας, καθώς οι εικόνες μεσαίου μεγέθους παρέχουν ικανοποιητική μορφολογική λεπτομέρεια ενώ παραμένουν διαχειρίσιμες από πλευράς μνήμης και χρόνου επεξεργασίας. Αν και οι εικόνες των 160×160 pixels προσφέρουν μεγαλύτερη ανάλυση, το αντίστοιχο υποσύνολο περιλαμβάνει σχεδόν τις μισές εικόνες (33.284), γεγονός που θα περιόριζε το εύρος των διαθέσιμων δεδομένων για εκπαίδευση.

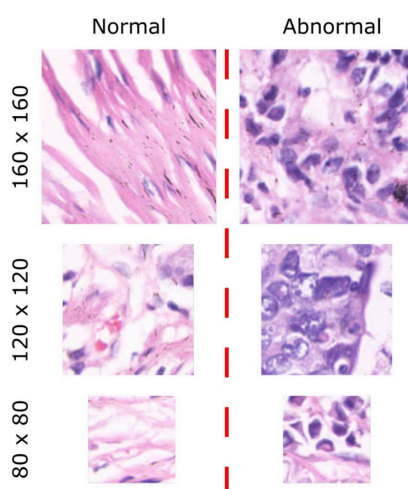
Μέθοδος Χρώσης (Staining method)

Η μέθοδος χρώσης H&E (αιματοξυλίνη και ηωσίνη) είναι μία από τις πιο διαδεδομένες τεχνικές χρώσης στην τεχνολογία τομής με παραφίνη και θεωρείται συχνά ως το gold standard. Η αιματοξυλίνη είναι αλκαλική και χρωματίζει τους κυτταρικούς πυρήνες σε ιώδες-μπλε, ενώ η ηωσίνη είναι όξινη και χρωματίζει το κυτταρόπλασμα και τη θεμέλια ουσία του ιστού σε ροζ χρώμα, με άλλες δομές να εμφανίζουν διαφορετικές αποχρώσεις και συνδυασμούς των δύο χρωμάτων.

Χάρη σε αυτή τη χρωματική αντίθεση, οι παθολόγοι μπορούν εύκολα να διακρίνουν το πυρηνικό και το κυτταροπλασματικό τμήμα των κυττάρων. Επιπλέον, το συνολικό πρότυπο χρώσης παρέχει πληροφορίες για τη διάταξη και την κατανομή των κυττάρων, καθώς και μία γενική επισκόπηση της δομής του ιστοικού δείγματος.

Η μέθοδος H&E αποτελεί τη βασικότερη και ευρύτερα χρησιμοποιούμενη τεχνική σε πεδία όπως η ιστολογία, η εμβρυολογία, η παθολογία και η επιστημονική έρευνα. Όσον αφορά την οπτική διαφοροποίηση:

- Οι φυσιολογικές εικόνες μετά τη χρώση H&E εμφανίζουν περισσότερες ροζ και λευκές περιοχές.
- Αντίθετα, στις μη φυσιολογικές, οι ιώδεις-μπλε περιοχές (που αντιστοιχούν σε πυρήνες καρκινικών κυττάρων) είναι περισσότερες, ακανόνιστα κατανεμημένες και πιο έντονες.



Σχήμα 5.4 Χρωματική Διαφορά
[58]

5.1.2.MHIST

Το MHIST είναι ένα dataset δυαδικής ταξινόμησης που αποτελείται από 3,152 ιστοπαθολογικές εικόνες πολύπιδων του παχέος εντέρου, διαστάσεων 224 x 224 pixels, οι οποίες προέρχονται από το τμήμα Παθολογίας και Εργαστηριακής Ιατρικής του Ιατρικού Κέντρου Dartmouth-Hitchcock (DHMC)[62].

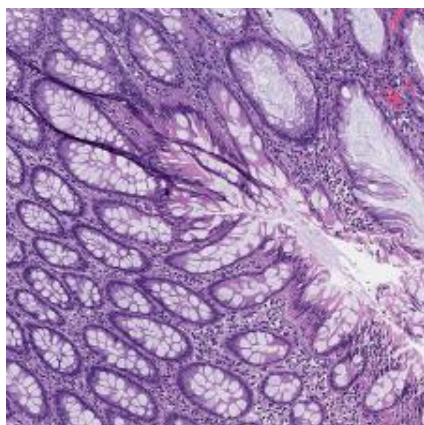
Κάθε εικόνα φέρει ετικέτα που έχει καθοριστεί από την πλειοψηφική γνώμη επτά πιστοποιημένων γαστρεντερολόγους.

Οι δύο κατηγορίες αυτού του συνόλου δεδομένων είναι :

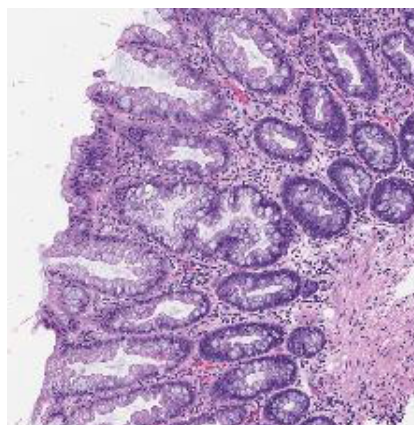
- Υπερπλαστικοί πολύποδες (Hyperplastic Polyps) – Καλοήθειες αλλοιώσεις του παχέος εντέρου, οι οποίες δεν θεωρούνται προδρομικές του καρκίνου.
- Ακτινωτά αδενώματα (Sessile Serrated Adenomas - SSA) – Προκαρκινικές αλλοιώσεις, που σχετίζονται με την ανάπτυξη καρκίνου του παχέος εντέρου μέσω του λεγόμενου "serrated pathway".

Η διάγνωση μεταξύ αυτών των δύο τύπων είναι συχνά προκλητική, ακόμη και για έμπειρους παθολογοανατόμους, λόγω των μορφολογικών ομοιοτήτων και της διαβαθμισμένης μετάβασης από καλοήθειες σε προκαρκινικές μορφές.

Ο όγκος αυτού του dataset είναι αρκετά μικρός. Για αυτό το λόγο χρησιμοποιήθηκε για transfer learning με βασικό μοντέλο, αυτό που εκπαιδεύτηκε στο προηγούμενο dataset (GasHisSDB). Συγκεκριμένα, έχουμε 2,162 εικόνες HP και 990 εικόνες SSA.



HP



SSA

Σχήμα 5.5 Παραδείγματα Εικόνων MHIST

Παρόλο που τόσο το MHIST όσο και το GasHisSDB χρησιμοποιούνται ευρέως για την εκπαίδευση και αξιολόγηση αλγορίθμων βαθιάς μάθησης σε ιστοπαθολογικές εικόνες, διαφέρουν ουσιωδώς ως προς τον τύπο ιστού, τις κατηγορίες ταξινόμησης και το φάσμα των παθολογιών που απεικονίζουν. Το MHIST περιλαμβάνει αποκλειστικά παθολογικές εικόνες του παχέος εντέρου, εστιάζοντας σε δύο τύπους προκαρκινικών βλαβών: υπερπλαστικοί πολύποδες και ακτινωτά αδενώματα, χωρίς την παρουσία φυσιολογικού ιστού. Συνεπώς, το πρόβλημα που προσεγγίζει είναι λεπτό και εστιάζει στη διαφορική διάγνωση μεταξύ μορφολογικά παρόμοιων αλλοιώσεων, με κλινικό ενδιαφέρον στον έγκαιρο εντοπισμό SSA ως πρόδρομης κατάστασης καρκίνου.

Αντίθετα, το GasHisSDB αφορά το στομάχι και περιλαμβάνει φυσιολογικές και μη-φυσιολογικές εικόνες, δίνοντας έμφαση στη γενικότερη διάκριση κακοήθειας έναντι υγιούς ιστού. Οι καρκινικές εικόνες προέρχονται από διάφορους τύπους καρκίνου του γαστρικού ιστού, ενώ το μεγάλο μέγεθος του dataset και η παρουσία σαφούς ground truth το καθιστούν κατάλληλο για εκπαίδευση γενικευτικών μοντέλων ταξινόμησης. Επομένως, ενώ το MHIST εξειδικεύεται στη διαγνωστική ακρίβεια εντός παθολογικού ιστού, το GasHisSDB καλύπτει ένα ευρύτερο διαγνωστικό φάσμα, ενσωματώνοντας τη βασική διάκριση μεταξύ υγιούς και καρκινικού ιστού και επιτρέποντας την ανάπτυξη μοντέλων εντοπισμού κακοήθειας σε ευρύτερη κλίμακα.

5.1.3. Προεπεξεργασία Δεδομένων

Το μοντέλο δέχτηκε ως είσοδο έγχρωμες εικόνες διαστάσεων $128 \times 128 \times 3$, οι οποίες αντιστοιχούν σε τυπικές RGB ψηφιοποιημένες ιστοπαθολογικές εικόνες. Το μέγεθος των εικόνων προέκυψε από κατάλληλη επαναδειγματοληψία (resizing) με σκοπό την ομοιομορφία εισόδου και την αποδοτική εκπαίδευση.

Κάθε εικόνα αναπαριστά μια περιοχή ιστού, προερχόμενη από Whole Slide Images (WSIs), και περιέχει είτε φυσιολογικά είτε καρκινικά μορφολογικά χαρακτηριστικά, ανάλογα με το εκάστοτε σύνολο δεδομένων.

Πριν την είσοδο στο δίκτυο, εφαρμόζεται κανονικοποίηση τιμών εικόνων στην κλίμακα $[0, 1]$, μέσω μετασχηματισμού $\text{rescale}=1.0/255$ στον μηχανισμό παραγωγής δεδομένων (ImageDataGenerator). Αυτό βελτιώνει τη σταθερότητα της εκπαίδευσης και επιταχύνει τη σύγκλιση του μοντέλου. Δεν εφαρμόζεται επιπλέον τυπική τυποποίηση, δεδομένου ότι η απλή κανονικοποίηση είναι συνήθως επαρκής για έγχρωμες εικόνες μικρής κλίμακας, ενώ δεν χρησιμοποιήθηκαν αλγοριθμικές τεχνικές data augmentation σε αυτή τη φάση.

Η επιλογή του συγκεκριμένου μεγέθους (128×128) επιτυγχάνει ισορροπία μεταξύ χωρικής πληροφορίας και υπολογιστικής αποδοτικότητας, επιτρέποντας την ανάλυση μικρομορφολογικών μοτίβων χωρίς υπερβολική αύξηση παραμέτρων ή υπερφόρτωση της μνήμης. Ενδείκνυται για εισαγωγή σε δίκτυα με πολλαπλά συνελκτικά επίπεδα, όπως αυτό που υλοποιήθηκε στην παρούσα εργασία.

5.2. Ανάπτυξη Μοντέλων Βαθιάς Μάθησης

Σκοπός ήταν ο προσδιορισμός της βέλτιστης προσέγγισης αρχιτεκτονικής συνελκτικού νευρωνικού δικτύου (CNN) για την ταξινόμηση ιστοπαθολογικών εικόνων, με στόχο την επίτευξη υψηλής ακρίβειας και γενικευσιμότητας. Στο πλαίσιο αυτό, διερευνήθηκαν τόσο προεκπαιδευμένα μοντέλα μεταφοράς μάθησης, αξιοποιώντας CNNs όπως το EfficientNet, όσο και η ανάπτυξη μίας προσαρμοσμένης αρχιτεκτονικής CNN, ειδικά σχεδιασμένης και βελτιστοποιημένης για τα χαρακτηριστικά του υπό μελέτη συνόλου δεδομένων. Μέσα από συγκριτική αξιολόγηση, εξετάστηκαν τα πλεονεκτήματα και οι περιορισμοί κάθε προσέγγισης, με στόχο την επιλογή της καταλληλότερης μεθόδου για το πρόβλημα της διάκρισης παθολογικών από μη παθολογικές περιοχές.

5.2.1. Προσέγγιση Μεταφοράς Μάθησης

Κατά τα πρώτα στάδια της μελέτης, εξετάστηκε η προσέγγιση της μεταφοράς μάθησης (transfer learning), με στόχο την επιτάχυνση της εκπαίδευσης και τη βελτίωση της απόδοσης του μοντέλου. Για τον σκοπό αυτό χρησιμοποιήθηκε το προεκπαιδευμένο μοντέλο EfficientNetB0, το οποίο έχει εκπαιδευτεί στο σύνολο δεδομένων ImageNet. Το EfficientNetB0 βασίζεται σε μια αρχή ομοιόμορφης κλιμάκωσης, όπου το εύρος, το βάθος και η ανάλυση εισόδου του δικτύου αυξάνονται με ισορροπημένο τρόπο, αποφεύγοντας την υπερβολική αύξηση παραμέτρων. Η επιλογή του συγκεκριμένου μοντέλου έγινε λόγω της ικανότητάς του να επιτυγχάνει υψηλές επιδόσεις σε πλήθος προβλημάτων ταξινόμησης με σημαντικά μειωμένο αριθμό παραμέτρων και μικρότερο χρόνο εκπαίδευσης συγκριτικά με άλλες βαθύτερες και πιο απαιτητικές αρχιτεκτονικές.

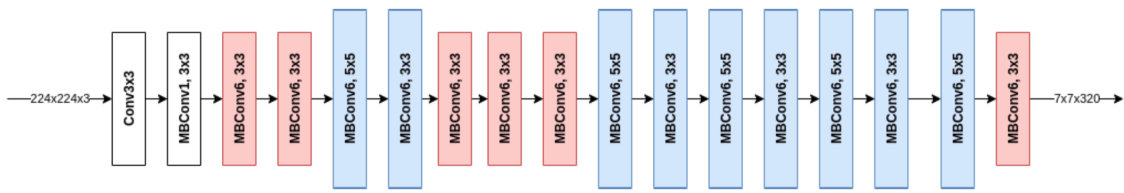
Η τελική ταξινόμηση επιτεύχθηκε μέσω προσαρμοσμένων πλήρως συνδεδεμένων στρωμάτων στην έξοδο του μοντέλου. Το EfficientNetB0 εκπαιδεύτηκε σε training set με κατανομή 90%-10%, ωστόσο τα αποτελέσματα στο test set υπήρξαν μη ικανοποιητικά.

Συγκεκριμένα, επιτεύχθηκαν:

- Test Sensitivity: 58.61%
- Test Specificity: 56.72%

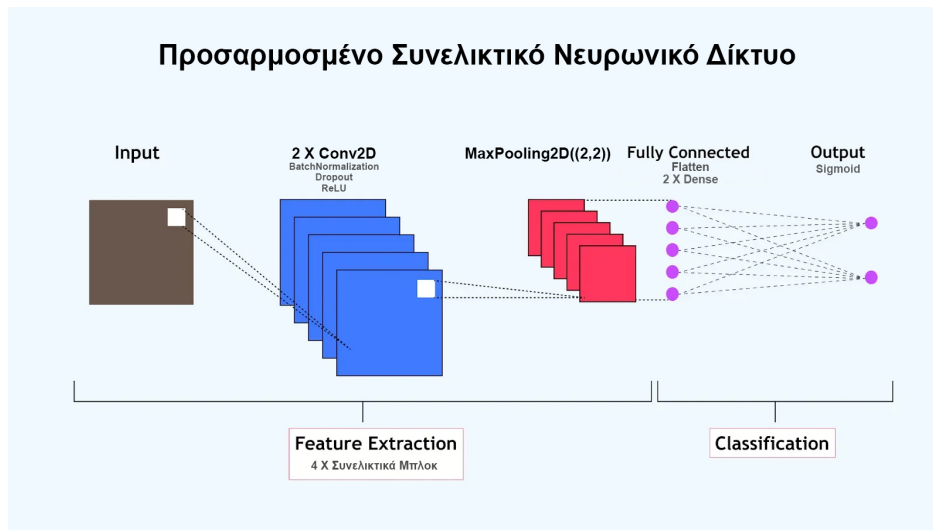
Η σύγκριση μεταξύ των δύο κατηγοριών (confusion matrix) υποδήλωσε έντονη αστάθεια του μοντέλου στην ταξινόμηση των εικόνων, με σχεδόν ισοδύναμο ποσοστό σφαλμάτων και για τις δύο κλάσεις. Η απόδοση αυτή πιθανότατα οφείλεται στη δομική απόκλιση μεταξύ των φυσικών εικόνων του ImageNet και των ιστοπαθολογικών εικόνων, με αποτέλεσμα τα προεκπαιδευμένα φίλτρα να μην είναι επαρκώς εξειδικευμένα για την ανάλυση των μορφολογικών χαρακτηριστικών που σχετίζονται με καρκινικές αλλοιώσεις.

Ως εκ τούτου, αποφασίστηκε η εγκατάλειψη της στρατηγικής μεταφοράς μάθησης υπέρ της σχεδίασης μιας custom συνελκτικής αρχιτεκτονικής CNN, ειδικά προσαρμοσμένης στο πρόβλημα.



Σχήμα 5.6 Αρχιτεκτονική Δικτύου EfficientNetB0 [63]

5.2.2. Ανάπτυξη Προσαρμοσμένων Αρχιτεκτονικών CNN



Σχήμα 5.7 Βασική Αρχιτεκτονική Προσαρμοσμένου CNN [64]

Δομή Συνελκτικών Επιπέδων (Conv Blocks)

Η βασική αρχιτεκτονική του δικτύου αποτελείται από τέσσερα διαδοχικά συνελκτικά μπλοκ (convolutional blocks), καθένα από τα οποία περιλαμβάνει δύο συνελκτικά επίπεδα Conv2D, ενδιάμεση ή τελική κανονικοποίηση μέσω Batch Normalization, και τεχνικές υποδειγματοληψίας MaxPooling2D, καθώς και κανονικοποίησης (regularization) μέσω Dropout.

Συγκεκριμένα, το πρώτο μπλοκ του μοντέλου περιλαμβάνει δύο συνελκτικά επίπεδα με N φίλτρα μεγέθους 3×3 και συνάρτηση ενεργοποίησης ReLU. Η επιλογή φίλτρων 3×3 , αν και μικρότερη σε σχέση με τα 5×5 , αποτελεί βέλτιστη λύση για την αποτελεσματική εξαγωγή τοπικών μορφολογικών χαρακτηριστικών στις ιστοπαθολογικές εικόνες, ενώ ταυτόχρονα προσφέρει χαμηλότερο υπολογιστικό κόστος. Η είσοδος κάθε Conv2D είναι μηδενισμένη στα άκρα (padding="same") για διατήρηση των διαστάσεων.

Ακολουθεί BatchNormalization, το οποίο σταθεροποιεί την κατανομή των ενεργοποιήσεων, μειώνει την εσωτερική μεταβλητότητα και επιταχύνει τη σύγκλιση του δικτύου. Στη συνέχεια, εφαρμόζεται MaxPooling2D(2×2) για μείωση των διαστάσεων και εξαγωγή των κυρίαρχων χαρακτηριστικών, και Dropout με ποσοστό 25% για αποτροπή υπερεκπαίδευσης.

Σε κάθε επόμενο μπλοκ, το πλήθος των φίλτρων διπλασιάζεται ($N \rightarrow 2N \rightarrow 4N \rightarrow 8N$), επιτρέποντας στο δίκτυο να αυξήσει τη χωρητικότητά του και να μάθει πιο αφηρημένα και σύνθετα χαρακτηριστικά σε βαθύτερα επίπεδα. Ο αριθμός των παραμέτρων αυξάνεται σταδιακά με ελεγχόμενο τρόπο, διατηρώντας τη σταθερότητα μέσω της επαναλαμβανόμενης χρήσης BatchNormalization. Η τιμή του Dropout αυξάνεται ανά στάδιο (0.25, 0.3, 0.4) καθώς βαθαίνει το δίκτυο, για να διατηρείται η ικανότητα γενίκευσης.

Η πολυεπίπεδη αυτή συνελικτική δομή επιτρέπει στο μοντέλο να ενσωματώνει πληθώρα χωρικών και δομικών χαρακτηριστικών από τις εισερχόμενες εικόνες, χτίζοντας ιεραρχικά από απλές ακμές και υφές έως σύνθετα καρκινικά μοτίβα.

Πλήρως Συνδεδεμένα Επίπεδα (Dense Layers)

Μετά την εξαγωγή χαρακτηριστικών από τα συνελικτικά επίπεδα, εφαρμόζεται ένα επίπεδο Flatten, το οποίο μετατρέπει τον τετραδιάστατο χάρτη χαρακτηριστικών (feature map) σε μονοδιάστατο διάνυσμα. Αυτή η μετατροπή επιτρέπει τη σύνδεση της εξαγόμενης πληροφορίας με τα τελικά επίπεδα ταξινόμησης, όπου οι ενεργοποιήσεις όλων των νευρώνων συμμετέχουν στην τελική απόφαση.

Η ταξινόμηση βασίζεται σε δύο διαδοχικά πλήρως συνδεδεμένα επίπεδα με 512 και 256 νευρώνες αντίστοιχα, καθένα από τα οποία χρησιμοποιεί τη συνάρτηση ενεργοποίησης ReLU.

Αυτά τα επίπεδα λειτουργούν ως μετασχηματιστές υψηλού επιπέδου (high-level transformation layers), που συνδυάζουν μη γραμμικά τα χαρακτηριστικά εξόδου από τα συνελικτικά στάδια και ενισχύουν την εκφραστική δύναμη του μοντέλου.

Μετά από κάθε Dense επίπεδο εφαρμόζεται Batch Normalization, το οποίο ομαλοποιεί την έξοδο του κάθε επιπέδου και συμβάλλει σε σταθερότερη εκπαίδευση, και Dropout, με ποσοστό που ορίζεται δυναμικά μέσω του υπερπαραμέτρου dropout_rate. Ο συνδυασμός αυτός επιτρέπει την αποφυγή υπερεκπαίδευσης, ειδικά στα dense layers, τα οποία περιέχουν μεγάλο αριθμό παραμέτρων.

Το τελικό επίπεδο του μοντέλου είναι ένας μόνο νευρώνας Dense με sigmoid ενεργοποίηση, κατάλληλος για δυαδική ταξινόμηση. Επιστρέφει τιμές στο διάστημα $[0, 1]$, οι οποίες ερμηνεύονται ως πιθανότητες να ανήκει η εικόνα στην θετική κατηγορία (π.χ. παθολογική εικόνα).

Η χρήση αυτής της συμμετρικής και προοδευτικής δομής πλήρως συνδεδεμένων επιπέδων ενισχύει τη δυνατότητα του μοντέλου να συνδυάζει την τοπική πληροφορία των συνελικτικών επιπέδων σε μια τελική, συνολική διάγνωση.

Regularization και Σταθεροποίηση Εκπαίδευσης

Το μοντέλο ενσωματώνει πολλαπλές τεχνικές regularization και σταθεροποίησης, με σκοπό τη βελτίωση της γενίκευσης και την αποτροπή φαινομένων υπερεκπαίδευσης, τα οποία είναι ιδιαίτερα πιθανά σε ιατρικά σύνολα δεδομένων περιορισμένου όγκου.

Πρωταρχικό μέσο regularization αποτελεί η χρήση της τεχνικής Dropout, η οποία εφαρμόζεται σε όλα τα συνελκτικά και πλήρως συνδεδεμένα στάδια του μοντέλου. Πιο συγκεκριμένα, χρησιμοποιούνται ποσοστά από 0.25 έως 0.4, τα οποία αυξάνονται προοδευτικά όσο βαθιάει το δίκτυο, ώστε να περιοριστεί η συσσώρευση εξαρτήσεων σε συγκεκριμένους νευρώνες.

Η στοχαστική απενεργοποίηση νευρώνων κατά τη διάρκεια της εκπαίδευσης συμβάλλει στην αποτροπή overfitting και ενισχύει την ικανότητα του μοντέλου να γενικεύει σε άγνωστα δείγματα.

Παράλληλα, σε όλα τα συνελκτικά και πλήρως συνδεδεμένα μπλοκ εφαρμόζεται Batch Normalization, το οποίο κανονικοποιεί τις ενδιάμεσες κατανομές ενεργοποιήσεων. Η τεχνική αυτή βελτιώνει τη σταθερότητα της εκπαίδευσης, επιταχύνει τη σύγκλιση και επιτρέπει τη χρήση υψηλότερων learning rates, ενώ δρα έμμεσα και ως μορφή regularization.

Για την εκπαίδευση του μοντέλου χρησιμοποιείται ο προσαρμοστικός βελτιστοποιητής Adam, ο οποίος επιλέχθηκε λόγω της αποτελεσματικότητάς του σε σύνθετα προβλήματα ταξινόμησης με noisy ή ετερογενή δεδομένα, όπως οι ιστοπαθολογικές εικόνες. Ο ρυθμός μάθησης καθορίζεται ως υπερπαραμέτρος (learning_rate) κατά τη διαδικασία πειραματικού προσδιορισμού των βέλτιστων τιμών (tuning), και προσφέρει ευελιξία προσαρμογής του βήματος ενημέρωσης βαρών σε κάθε φάση της εκπαίδευσης.

Ο συνδυασμός των παραπάνω τεχνικών εξασφαλίζει ότι το δίκτυο μπορεί να μάθει σταθερά, αποδοτικά και γενικεύσιμα, ακόμα και όταν εφαρμόζεται σε datasets με ποικίλες παραλλαγές στη χρώση, την ανάλυση και τη μορφολογία των ιστών.

Grid Search

Η επιλογή των υπερπαραμέτρων του μοντέλου πραγματοποιήθηκε μέσω Grid Search, δηλαδή εξαντλητικής αναζήτησης σε προκαθορισμένο χώρο τιμών, με στόχο τη βελτιστοποίηση της επίδοσης σε σύνολο επικύρωσης. Η διαδικασία βασίστηκε σε συγκεκριμένο πλέγμα παραμέτρων, όπου μεταβλήθηκαν οι εξής υπερπαραμέτροι:

- num_filters: [32,64]
- kernel_size: [(3, 3), (5, 5)]
- dropout_rate: [0.3, 0.5]
- learning_rate: [0.001, 0.0001]

Ο αριθμός των φίλτρων πήρε τιμές 32 και 64, ενώ για κάθε τιμή έγινε εξερεύνηση δύο μεγεθών συνελκτικού πυρήνα (3×3 και 5×5), δύο ποσοστών dropout (30% και 50%) και δύο τιμών ρυθμού μάθησης (0.001 και 0.0001). Έτσι, το συνολικό πλέγμα περιλάμβανε $2 \times 2 \times 2 \times 2 = 16$ διαφορετικούς συνδυασμούς υπερπαραμέτρων.

Ως κριτήριο βελτιστοποίησης επιλέχθηκε η ακρίβεια στο validation set, ενώ η αξιολόγηση κάθε συνδυασμού έγινε με βάση σταθερό αρχιτεκτονικό πλαίσιο και ίδια δεδομένα.

Μετά την ολοκλήρωση της διαδικασίας, οι βέλτιστες υπερπαραμέτροι που αναδείχθηκαν ήταν:

- num_filters = 64
- kernel_size = 3x3
- dropout_rate = 0.5
- learning_rate = 0.0001

Αυτός ο συνδυασμός προσέφερε βελτιστοποιημένη σταθερότητα εκπαίδευσης, με περιορισμένο overfitting, σταδιακή σύγκλιση, και υψηλή ακρίβεια στην ταξινόμηση των ιστοπαθολογικών εικόνων. Η χρήση μικρότερου πυρήνα συνέλιξης (3×3) αποδείχθηκε αποτελεσματικότερη, καθώς προσέφερε καλύτερη γενίκευση και ταχύτερη σύγκλιση, χωρίς σημαντική επιβάρυνση στον υπολογιστικό χρόνο. Ο συνδυασμός αυξημένου dropout (0.5) με χαμηλό learning rate συνέβαλε στη σταθερότερη εκπαίδευση και στην αποτροπή υπερπροσαρμογής (overfitting).

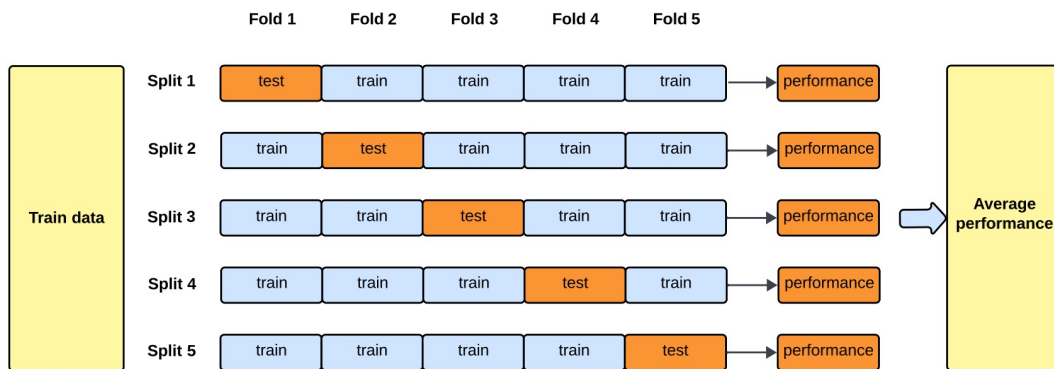
5.3. Εκπαίδευση και Αξιολόγηση

5.3.1. Stratified k-fold

Για την αξιόπιστη αξιολόγηση της γενικευσιμότητας του CNN, εφαρμόστηκε η τεχνική της διαστρωματωμένη διασταυρούμενη επικύρωση (Stratified 5-Fold Cross-Validation). Η διαδικασία αυτή διασφαλίζει ότι κάθε παρατήρηση του συνόλου δεδομένων συμμετέχει ακριβώς μία φορά στο test set και τέσσερις φορές στο training set, διατηρώντας παράλληλα την αναλογία των κλάσεων (stratification) σε κάθε fold.

Διαδικασία Ανάλυσης ανά Fold

- Το αρχικό πλήθος δεδομένων διαχωρίστηκε σε 5 μη επικαλυπτόμενα υποσύνολα με χρήση της StratifiedKFold.
- Σε κάθε fold, τέσσερα υποσύνολα χρησιμοποιήθηκαν για εκπαίδευση και το πέμπτο για τελικό έλεγχο απόδοσης.
- Από τα τέσσερα υποσύνολα εκπαίδευσης, πραγματοποιήθηκε περαιτέρω διαχωρισμός (75%-25%) ώστε να δημιουργηθεί το validation set με stratified split, διατηρώντας και εκεί την κατανομή των κλάσεων.
- Η επεξεργασία εικόνας έγινε μέσω ImageDataGenerator με ανακατασκευή των pixel (rescale=1.0/255) και δημιουργία κατάλληλων generators για training, validation και test.



Σχήμα 5.7 5-Fold Cross-Validation [65]

Το κάθε fold εκπαιδεύτηκε από την αρχή με το 8-layer CNN, για έως 5 εποχές με χρήση EarlyStopping που παρακολουθεί την val_loss. Οι παράμετροι ήταν σταθερές σε όλα τα folds για δίκαιη σύγκριση.

Μετά από κάθε εκπαίδευση, το μοντέλο αξιολογήθηκε στο αντίστοιχο test set του fold. Υπολογίστηκαν οι εξής δείκτες απόδοσης, οι οποίοι αναλύονται σε επόμενη υποενότητα:

- Ευαισθησία (Sensitivity / Recall)
- Ειδικότητα (Specificity)
- Ακρίβεια (Accuracy)
- Ακρίβεια Θετικών (Precision)
- F1-Score
- Εμβαδόν κάτω από την καμπύλη ROC (AUC)
- Συντελεστής Matthew (MCC)
- Δείκτης Youden (Youden's J)
- Πίνακας Σύγχυσης (Confusion Matrix)

Όλες οι τιμές καταγράφηκαν ανά fold και στο τέλος υπολογίστηκε ο μέσος όρος όλων μεταξύ των 5 folds για μια συνολική εκτίμηση της απόδοσης του μοντέλου.

5.3.2.Μετρικές Αξιολόγησης

Για την αξιολόγηση ενός μοντέλου μηχανικής μάθησης, ειδικά σε ιατρικά δεδομένα όπου η ακρίβεια της διάγνωσης είναι κρίσιμη, απαιτείται η χρήση κατάλληλων μετρήσεων (metrics) που αποτυπώνουν διαφορετικές πτυχές της απόδοσής του. Παρακάτω παρουσιάζονται αναλυτικά κάθε μία από αυτές, εξηγώντας τη σημασία τους και τον τρόπο υπολογισμού τους.

- **Ευαισθησία (Sensitivity / Recall)**

Η ευαισθησία μετρά το ποσοστό των πραγματικά θετικών περιπτώσεων που αναγνωρίζονται σωστά από το μοντέλο. Δηλώνει πόσο καλά το μοντέλο εντοπίζει τις περιπτώσεις της θετικής κλάσης. Στο dataset με το οποίο ασχοληθήκαμε η ευαισθησία αφορά την Abnormal κλάση.

$$Sensitivity = \frac{TP}{TP + FN}$$

Όπου TP: True Positives, FN: False Negatives.

- **Ειδικότητα (Specificity)**

Η ειδικότητα μετρά το ποσοστό των πραγματικά αρνητικών περιπτώσεων που αναγνωρίζονται σωστά, δηλαδή την ικανότητα του μοντέλου να βρίσκει σωστά την Normal κλάση. Αντικατοπτρίζει την ικανότητα του μοντέλου να αποφεύγει ψευδώς θετικά αποτελέσματα.

$$Specificity = \frac{TN}{TN + FP}$$

Όπου TN: True Negatives, FP: False Positives.

- **Ακρίβεια (Accuracy)**

Η ακρίβεια δείχνει το ποσοστό των συνολικών σωστών προβλέψεων, είτε θετικών είτε αρνητικών, επί του συνόλου των δειγμάτων. Είναι μια συνολική μέτρηση επιτυχίας.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

- **Ακρίβεια Θετικών (Precision)**

Η ακρίβεια θετικών δείχνει το ποσοστό των προβλέψεων ως θετικές που είναι πράγματι θετικές. Δηλαδή, από όσα το μοντέλο προβλέπει ως Abnormal, πόσα είναι όντως.

$$Precision = \frac{TP}{TP + FP}$$

- **F1-Score**

Το F1-score είναι το αρμονικό μέσο της ευαισθησίας και της ακρίβειας. Χρησιμοποιείται όταν θέλουμε να έχουμε μια ισορροπημένη εικόνα μεταξύ ψευδώς θετικών και ψευδώς αρνητικών.

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

- **Εμβαδόν κάτω από την καμπύλη ROC (AUC)**

Το AUC (Area Under the Curve) μετρά την ικανότητα του μοντέλου να διακρίνει ανάμεσα στις δύο κλάσεις. Όσο πιο κοντά στο 1, τόσο καλύτερα διαχωρίζει τις θετικές από τις αρνητικές περιπτώσεις. Προκύπτει από την καμπύλη ROC (Receiver Operating Characteristic).

- **Συντελεστής Matthew (MCC)**

Ο συντελεστής Matthew είναι μια ισορροπημένη μετρική που λαμβάνει υπόψη όλες τις τιμές του confusion matrix (TP, TN, FP, FN). Είναι ιδιαίτερα χρήσιμο όταν τα δεδομένα είναι μη ισορροπημένα.

$$\text{Τύπος: } MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

- **Δείκτης Youden (Youden's J)**

Ο δείκτης Youden αξιολογεί την αποτελεσματικότητα ενός binary classifier και δίνεται από το άθροισμα της ευαισθησίας και της ειδικότητας μείον 1. Χρησιμοποιείται για την επιλογή του βέλτιστου threshold.

$$J = \text{Sensitivity} + \text{Specificity} - 1$$

Στον παρόντα αλγόριθμο, η διαδικασία λήψης απόφασης για τη δυαδική ταξινόμηση των ιστοπαθολογικών εικόνων δε βασίζεται σε ένα προκαθορισμένο κατώφλι, αλλά σε ένα δυναμικά υπολογιζόμενο κατώφλι μέσω του δείκτη Youden (Youden's J).

Ο δείκτης Youden ορίζεται ως η διαφορά TPR – FPR, και το κατώφλι που τον μεγιστοποιεί θεωρείται το βέλτιστο. Αυτό το κατώφλι χρησιμοποιείται για να μετατραπούν οι συνεχείς πιθανότητες σε τελικές δυαδικές προβλέψεις, εξασφαλίζοντας έτσι μια ισορροπία ανάμεσα στην ευαισθησία και την ειδικότητα του μοντέλου. Με αυτόν τον τρόπο, η αξιολόγηση του ταξινομητή βασίζεται σε μια προσαρμοστική στρατηγική που βελτιστοποιεί την απόδοσή του με βάση τα χαρακτηριστικά του εκάστοτε dataset.

5.4.Ερμηνευσιμότητα και οπτικοποίηση

5.4.1.Grad-CAM

Για τη βελτίωση της ερμηνευσιμότητας του μοντέλου, εφαρμόστηκε η τεχνική Gradient-weighted Class Activation Mapping (Grad-CAM), η οποία επιτρέπει την οπτική επισήμανση των περιοχών της εικόνας που συνέβαλαν περισσότερο στη ληφθείσα απόφαση.

Όπως αναφέρθηκε και προηγουμένως, το Grad-CAM χρησιμοποιεί την τιμή ενεργοποίησης της επιλεγμένης κλάσης πριν από την κανονικοποίηση, όπου εφαρμόζεται η συνάρτηση softmax.

Στην παρούσα εργασία, το μοντέλο είναι δυαδικός ταξινομητής με σιγμοειδούς συνάρτηση ενεργοποίησης:

$$\text{output} = \sigma(z) = \frac{1}{1 + e^{-z}} \quad (5.4.1)$$

όπου z είναι το λογιστικό σκορ (logit). Επομένως, το Grad-CAM εφαρμόζεται απευθείας πάνω στην έξοδο της sigmoid, χωρίς χρήση softmax, και το y^c αντιστοιχεί στο συνεχές σκορ πρόβλεψης της κατηγορίας “Abnormal”.

Η έξοδος αυτή καταγράφεται και αποθηκεύεται ανά fold, το οποίο, επιτρέπει να αξιολογήσουμε όχι μόνο τη συνολική ακρίβεια του μοντέλου, αλλά και τη συνεκτικότητα της προσοχής του σε κάθε είδος πρόβλεψης.

Με τη χρήση του Grad-CAM είναι εμφανές ότι έχουμε μία καθαρή εικόνα των σημείων που στόχευσε το μοντέλο ώστε να προβλέψει στη συγκεκριμένη περίπτωση ορθά. Το heatmap, σε συνδυασμό με την αρχική εικόνα του ιστού, είναι ένα σημαντικό εργαλείο για παθολογοανατόμους, οι οποίοι είναι ικανοί να αξιολογήσουν τα κριτήρια με τα οποία ο αλγόριθμος έκανε την ταξινόμηση. Το ζήτημα δεν είναι μόνο αν η εικόνα ταξινομήθηκε ορθά, αλλά και αν η απόφαση πάρθηκε βάσει σωστών χαρακτηριστικών, στη συγκεκριμένη περίπτωση, καρκινικών πυρήνων.

5.4.2.HistomicsTK

Για την επεξεργασία των ιστοπαθολογικών εικόνων και την εξαγωγή μορφολογικών χαρακτηριστικών και χαρακτηριστικών υφής από τους πυρήνες, αξιοποιήθηκε η βιβλιοθήκη HistomicsTK[66]. Η HistomicsTK είναι μία ανοικτού κώδικα βιβλιοθήκη επεξεργασίας εικόνας, ειδικά σχεδιασμένη για την ανάλυση ψηφιακών ιστοπαθολογικών εικόνων (WSIs). Αναπτύχθηκε από το Kitware και αποτελεί μέρος της ευρύτερης πλατφόρμας Digital Slide Archive (DSA), η οποία υποστηρίζει την αποθήκευση, διαχείριση και ανάλυση ψηφιοποιημένων βιοϊατρικών εικόνων μεγάλης κλίμακας.

Η HistomicsTK παρέχει ένα πλήρες σύνολο εργαλείων για:

- Προεπεξεργασία χρωματισμένων εικόνων (π.χ. color normalization, stain deconvolution),

- Τμηματοποίηση (segmentation) κυτταρικών δομών όπως οι πυρήνες (π.χ. με μεθόδους LoG και watershed),
- Εξαγωγή χαρακτηριστικών από περιοχές ενδιαφέροντος, συμπεριλαμβανομένων μορφολογικών, φασματικών, εντατικών και υφικών (GLCM/Haralick) περιγραφών.

Ιδιαίτερα σημαντική είναι η δυνατότητά της να εφαρμόζει αποσύνθεση χρώματος (color deconvolution), απομονώνοντας τα χρωστικά συστατικά (όπως η αιματοξυλίνη και η ηωσίνη), καθώς και η εφαρμογή σύνθετων μεθόδων εξαγωγής χαρακτηριστικών όπως οι Haralick υφικοί δείκτες, οι οποίοι επιτρέπουν την ποσοτική αποτύπωση της υφής εντός των κυτταρικών περιοχών.

Η βιβλιοθήκη διατίθεται μέσω του GitHub και μπορεί να εγκατασταθεί και να χρησιμοποιηθεί τοπικά ή ενσωματωμένα σε pipelines μέσω Python[67]. Χρησιμοποιείται σε πληθώρα βιοϊατρικών ερευνητικών εργασιών, ιδιαίτερα σε μελέτες όπου απαιτείται ποσοτική κυτταρική ανάλυση και συσχέτιση μορφολογικών προτύπων με διαγνωστικά ή προβλεπτικά μοντέλα. Η δυνατότητά της να συνεργάζεται με Whole Slide Image viewers και να κλιμακώνεται για μεγάλες εικόνες την καθιστά ιδιαίτερα κατάλληλη για εφαρμογές στην ψηφιακή παθολογία.

Στην εργασία η βιβλιοθήκη αυτή χρησιμοποιήθηκε για την κατανόηση των αποτελεσμάτων της μεθόδου Grad-CAM και για την γενική εξαγωγή χαρακτηριστικών από τις ιστοπαθολογικές εικόνες. Ειδικότερα, με βάση τα αποτελέσματα της ερμηνευσιμότητας, επιλέχθηκαν όλες οι TP (true positive) εικόνες, για τις οποίες αποκτήθηκαν τα heatmaps, και υπέστησαν μια σειρά διαδικασιών.

Αρχικά, εφαρμόστηκε κανονικοποίηση χρώματος με τη μέθοδο Reinhard, ώστε να μειωθεί η διακύμανση στις χρώσεις (H&E) μεταξύ των εικόνων. Στη συνέχεια, μέσω αποχρωματισμού, απομονώθηκε το χρωματικό κανάλι της αιματοξυλίνης (hematoxylin), το οποίο περιέχει κυρίως την πληροφορία των πυρηνικών δομών.

Για την τμηματοποίηση των πυρήνων, εφαρμόστηκε LoG-based πυρηνική ανίχνευση με χρήση πυρήνων διαφόρων μεγεθών (multi-scale Laplacian of Gaussian), ακολουθούμενη από μέγιστη ομαδοποίηση (max clustering). Το τελικό αποτέλεσμα ήταν μία ετικετοποιημένη μάσκα πυρήνων, στην οποία κάθε αντικείμενο αντιστοιχεί σε έναν μεμονωμένο πυρήνα.

Στο σύνολο των τμηματοποιημένων πυρήνων εφαρμόστηκε εξαγωγή χαρακτηριστικών με τις ακόλουθες συναρτήσεις του HistomicsTK:

- **compute_morphometry_features** : Γεωμετρικά και μορφολογικά χαρακτηριστικά

Η συνάρτηση αυτή υπολογίζει βασικές μορφολογικές ιδιότητες κάθε πυρήνα, οι οποίες είναι κρίσιμες για την κατανόηση της κυτταρικής μορφολογίας όπως έκταση, περιφέρεια, εκκεντρότητα και επιπλέον δείκτες περιγραφής του σχήματος του πυρήνα.

- **compute_fsd_features** : Fourier Shape Descriptors (FSD)

Οι Fourier Shape Descriptors είναι χαρακτηριστικά που βασίζονται στην παραμετροποίηση του περιγράμματος κάθε πυρήνα μέσω μετασχηματισμού Fourier.

Η συνάρτηση αυτή αντιπροσωπεύει το σχήμα του πυρήνα ως μια σειρά συνημιτονικών και ημιτονικών συνιστωσών και επιτρέπει την ανάλυση της μορφής ανεξάρτητα από το μέγεθος ή τον προσανατολισμό, καθιστώντας τα FSDs χρήσιμα σε περιβάλλοντα με έντονες παραλλαγές κυτταρικής γεωμετρίας.

- **compute_gradient_features** : Τοπικές Διαβαθμίσεις Έντασης

Η συνάρτηση αυτή εξάγει πληροφορία από τη μεταβολή της έντασης των pixel εντός του πυρήνα, δηλαδή πληροφορία για την κατεύθυνση και την ομαλότητα των μεταβολών έντασης, η οποία σχετίζεται με τη χρωματική ομοιογένεια. Περιγράφει την οξύτητα των ακμών εντός κάθε πυρήνα (edge sharpness), άρα πόσο απότομα μεταβάλλεται η ένταση στα όρια των δομών μέσα στον πυρήνα.

- **compute_haralick_features** : Χαρακτηριστικά Υφής με βάση GLCM

Η συνάρτηση αυτή εξάγει χαρακτηριστικά υφής με βάση τον Gray-Level Co-occurrence Matrix (GLCM), που εφαρμόζεται στο κανάλι αιματοξυλίνης της εικόνας (δηλαδή την πληροφορία του πυρήνα). Η μέθοδος GLCM καταγράφει τη συχνότητα εμφάνισης ζευγών τιμών έντασης σε γειτονικές θέσεις εντός κάθε πυρήνα και παράγει υφικά descriptors όπως:

- contrast (αντίθεση),
- correlation (συσχέτιση),
- energy (ομοιομορφία),
- homogeneity (ομοιογένεια), που αποτυπώνουν την εσωτερική υφή κάθε κυττάρου.

Αυτά τα χαρακτηριστικά χρησιμοποιήθηκαν για να μελετηθεί η συσχέτισή τους με τις ενεργές περιοχές Grad-CAM και την προβλεπτική συμπεριφορά του μοντέλου.

Για να μελετηθεί η σχέση ανάμεσα στη μορφολογία των πυρήνων και στη συμπεριφορά του μοντέλου, πραγματοποιήθηκε συνδυαστική ανάλυση μεταξύ των χαρακτηριστικών των τμηματοποιημένων πυρήνων και των θερμικών χαρτών Grad-CAM. Συγκεκριμένα, για κάθε εικόνα που ταξινομήθηκε σωστά ως "Abnormal" (True Positive), προσδιορίστηκαν ποιοι από τους ανιχνευμένους πυρήνες βρίσκονταν εντός των περιοχών υψηλής ενεργοποίησης του Grad-CAM, χρησιμοποιώντας κατάλληλο κατώφλι.

Οι πυρήνες χωρίστηκαν σε δύο κατηγορίες:

- Ενεργοί πυρήνες (active nuclei): όσοι τέμνονταν με περιοχές ενεργοποίησης του Grad-CAM.
- Ανενεργοί πυρήνες (inactive nuclei): όσοι βρίσκονταν εκτός αυτών των περιοχών.

Στη συνέχεια, για κάθε ομάδα πυρήνων, αναλύθηκαν στατιστικά τα μορφολογικά και τα χαρακτηριστικά υφής που είχαν εξαχθεί με HistomicsTK, όπως έκταση, εκκεντρότητα, αντίθεση και ομοιογένεια. Αυτή η ανάλυση επέτρεψε την ερμηνεία των προβλεπτικών περιοχών του μοντέλου με βάση βιολογικά σημαντικά χαρακτηριστικά των κυττάρων, και ανέδειξε κατά πόσο το δίκτυο εστιάζει πράγματι σε μορφολογικά ύποπτες περιοχές για τη λήψη των αποφάσεών του.

5.5.Αβεβαιότητα

Για την εκτίμηση της αβεβαιότητας στις προβλέψεις του συνελκτικού νευρωνικού δικτύου, αξιοποιήθηκε η μέθοδος Monte Carlo Dropout (MC Dropout). Η μεθοδολογία εφαρμόστηκε στον τελικό κώδικα με το 90-10 split του dataset.

Αφού ολοκληρώθηκε η εκπαίδευση, για κάθε εικόνα του test set πραγματοποιήθηκαν 50 στοχαστικές προωθήσεις (forward passes) με ενεργό dropout. Από αυτές υπολογίστηκε:

- η μέση πιθανότητα ταξινόμησης, η οποία χρησιμοποιήθηκε ως τελική πρόβλεψη, και
- η τυπική απόκλιση των προβλέψεων, η οποία χρησιμοποιήθηκε ως μέτρο αβεβαιότητας.

Τα αποτελέσματα καταγράφηκαν σε αρχείο .csv, περιλαμβάνοντας για κάθε δείγμα το όνομα της εικόνας, την πραγματική ετικέτα, την ετικέτα που προέβλεψε το μοντέλο, τη μέση πρόβλεψη και την αβεβαιότητα (τυπική απόκλιση). Επιπλέον, δημιουργήθηκε και αποθηκεύτηκε ιστόγραμμα της κατανομής των αβεβαιοτήτων, το οποίο αποτυπώνει την εμπιστοσύνη του μοντέλου στις αποφάσεις του σε όλο το test set.

Η εκτίμηση των τελικών μετρικών πραγματοποιήθηκε με βάση τις μέσες τιμές των στοχαστικών προβλέψεων. Με αυτόν τον τρόπο, όχι μόνο αξιολογήθηκε η απόδοση του μοντέλου, αλλά αποκτήθηκε και μια ποσοτική εικόνα της αβεβαιότητας που το συνοδεύει, γεγονός κρίσιμο για εφαρμογές υψηλής ευθύνης, όπως η ανάλυση ιατρικών εικόνων.

Η χρήση του Monte Carlo Dropout συμβάλλει ουσιαστικά στη βελτίωση της αξιοπιστίας των τιμών αξιολόγησης, καθώς προσφέρει μια πιο σταθερή και αντιπροσωπευτική εκτίμηση της συμπεριφοράς του μοντέλου μέσω στοχαστικών προβλέψεων. Με τον τρόπο αυτό μειώνεται η πιθανότητα να καταγραφούν υψηλές επιδόσεις "κατά τύχη", όπως μπορεί να συμβεί με μία και μόνο ντετερμινιστική πρόβλεψη.

Παράλληλα, η μεθοδολογία αυτή προσθέτει μια επιπλέον διάσταση αξιολόγησης, την ποσοτικοποιημένη αβεβαιότητα, η οποία απουσιάζει πλήρως από τα κλασικά μοντέλα. Η πληροφορία αυτή είναι ιδιαίτερα χρήσιμη σε εφαρμογές υψηλής ευθύνης, καθώς επιτρέπει την αναγνώριση περιπτώσεων στις οποίες το μοντέλο εμφανίζει δισταγμό ή έλλειψη εμπιστοσύνης στην απόφασή του, υποστηρίζοντας έτσι διαδικασίες όπως η επιλεκτική ανθρώπινη επιθεώρηση ή η ενίσχυση της εκπαίδευσης με δύσκολα παραδείγματα μέσω ενεργούς μάθησης (active learning).

6. Αποτελέσματα

6.1. Αξιολόγηση Διακριτικής Ικανότητας

6.1.1. GasHisSDB

5-Fold Cross-Validation (Διασταυρούμενη Επικύρωση)

Για την αρχική αξιολόγηση του μοντέλου εφαρμόστηκε 5-Fold Cross-Validation, προκειμένου να εκτιμηθεί η σταθερότητα και η γενικευσιμότητα του CNN ταξινομητή. Το σύνολο δεδομένων χωρίστηκε σε πέντε ίσα μέρη και σε κάθε επανάληψη χρησιμοποιήθηκε ένα διαφορετικό fold ως test set, ενώ τα υπόλοιπα τέσσερα συνέθεσαν το training set.

Ο μέσος όρος των metrics που προέκυψαν από τις πέντε επαναλήψεις συνοψίζεται παρακάτω.

Πίνακας 6.1 Αποτελέσματα μετρικών αξιολόγησης του προσαρμοσμένου CNN μοντέλου στο σύνολο δεδομένων GasHisDB, 5Fold

Ακρίβεια	Ευαισθησία	Ειδικότητα	Ακρίβεια Θετικών	F1-Score	AUC	MCC
90.08%	89.15%	90.65%	85.44%	87.25%	0.9605	0.7919

Τα παραπάνω αποτελέσματα υποδηλώνουν ότι το μοντέλο παρουσιάζει ισορροπημένη συμπεριφορά ως προς την αναγνώριση και των δύο κατηγοριών, με υψηλή διακριτική ικανότητα ($AUC > 0.96$) και σημαντικά θετική συσχέτιση ανάμεσα στις προβλέψεις και τις πραγματικές τιμές ($MCC \approx 0.79$). Η χρήση του δείκτη Youden's J επέτρεψε την επιλογή ενός κατάλληλου κατωφλίου που ενισχύει την ισορροπία μεταξύ ευαισθησίας και ειδικότητας σε κάθε fold.

Κατανομή 90-10

Μετά τη διαδικασία βελτιστοποίησης υπερπαραμέτρων και την αξιολόγηση μέσω 5-Fold Cross-Validation, το τελικό μοντέλο εκπαιδεύτηκε εξ' ολοκλήρου στο 90% των δεδομένων, και αξιολογήθηκε στο υπόλοιπο 10%. Η αξιολόγηση πραγματοποιήθηκε με βάση τις ίδιες τιμές απόδοσης.

Τα αποτελέσματα του μοντέλου στο test set παρουσιάζονται παρακάτω.

Πίνακας 6.2 Αποτελέσματα μετρικών αξιολόγησης του προσαρμοσμένου CNN μοντέλου στο σύνολο δεδομένων GasHisDB, 90-10

Model	Ακρίβεια	Ευαισθησία	Ειδικότητα	Ακρίβεια Θετικών	F1-Score	AUC	MCC
Custom CNN	94.12%	93.15%	94.71%	91.52%	92.33%	0.9860	0.8757

Το μοντέλο παρουσιάζει ιδιαίτερα υψηλή ικανότητα διάκρισης μεταξύ των δύο κατηγοριών. Στη εφαρμογή 90-10, παρατηρείται αύξηση όλων των μετρήσεων απόδοσης συγκριτικά με την αξιολόγηση 5-fold που έδωσε το μέσο όρο των 5 επαναλήψεων.

6.1.2.MHIST

Μετά την αρχική εκπαίδευση του μοντέλου στο GasHisSDB, εφαρμόστηκε ο ίδιος ακριβώς κώδικας, χωρίς καμία τροποποίηση στη δομή του μοντέλου ή στη διαδικασία εκπαίδευσης, στο MHIST. Στόχος αυτής της προσέγγισης ήταν η αξιολόγηση της γενικευσιμότητας του μοντέλου, δηλαδή της ικανότητάς του να διατηρεί υψηλές επιδόσεις όταν αντιμετωπίζει νέα δεδομένα με διαφορετικά χαρακτηριστικά. Η διαδικασία περιελάμβανε εκ νέου προεπεξεργασία, διαχωρισμό σε train-val-test σύνολα και εκπαίδευση από την αρχή χρησιμοποιώντας τα δεδομένα του MHIST.

Οι μετρήσεις απόδοσης του μοντέλου παρουσιάζονται παρακάτω. Αν και χαμηλότερες σε σχέση με την απόδοση στο αρχικό σύνολο δεδομένων, θεωρούνται ικανοποιητικές με βάση τη φύση του προβλήματος και τις προκλήσεις του MHIST.

Πίνακας 6.3 Αποτελέσματα μετρικών αξιολόγησης του προσαρμοσμένου CNN μοντέλου στο σύνολο δεδομένων MHIST

Model	Ακρίβεια	Ευαισθησία	Ειδικότητα	Ακρίβεια Θετικών	F1-Score	AUC	MCC
Custom CNN	73.78%	72.73%	72.81%	54.96%	62.62%	0.7943	0.4287

Η πτώση στην απόδοση μπορεί να αποδοθεί σε πολλούς παράγοντες, όπως η διαφορετική φύση των κατηγοριών. Σε αντίθεση με το GasHisSDB, όπου υπάρχει σαφής διαχωρισμός φυσιολογικού και παθολογικού ιστού, το MHIST περιλαμβάνει δύο τύπους καρκινικών αλλοιώσεων, με παρόμοια μορφολογία και μικρές διαφορές σε υφή και σχήμα, γεγονός που δυσχεραίνει τη διάκριση. Επιπλέον, το MHIST περιλαμβάνει μικρότερο αριθμό δειγμάτων σε σχέση με το αρχικό subset των 120×120 του GasHisSDB, περιορίζοντας την ικανότητα του μοντέλου να μάθει ποικίλα πρότυπα. Ακόμα, η αρχιτεκτονική είναι ένα απλό CNN με 8 layers, χωρίς προεκπαίδευση, γεγονός που το καθιστά ευαίσθητο σε αλλαγές στη στατιστική κατανομή των δεδομένων.

Τέλος, το MHIST περιλαμβάνει αποκλειστικά καρκινικές αλλοιώσεις. Ως αποτέλεσμα, οι εικόνες έχουν πολύ παρόμοια μορφολογία και παρόμοιο χρωματισμό, με κυρίαρχους τόνους σκούρου μωβ, χαρακτηριστικούς της αιματοξυλίνης. Αντιθέτως, στο GasHisSDB οι φυσιολογικές εικόνες τείνουν να έχουν ροζ απόχρωση, καθώς απουσιάζουν υπερχρωματισμοί από κυτταρικούς πυρήνες. Αυτό προσφέρει στο μοντέλο πιο σαφή και έντονα οπτικά χαρακτηριστικά για να διακρίνει μεταξύ των δύο κατηγοριών.

Αυτό το χρωματικό contrast στο GasHisSDB διευκολύνει την ταξινόμηση, ενώ στο MHIST η οπτική ομοιομορφία μεταξύ HP και SSA αυξάνει σημαντικά τη δυσκολία του προβλήματος. Επομένως, η χαμηλότερη απόδοση του μοντέλου στο MHIST μπορεί να οφείλεται στη χρωματική και μορφολογική ομοιότητα των δειγμάτων. Αυτό το καθιστά ένα πιο "λεπτό" πρόβλημα ταξινόμησης και απαιτεί πιο εξειδικευμένα ή υψηλής απόδοσης μοντέλα για να εντοπίσουν subtle διαφοροποιήσεις.

6.1.3. Σύγκριση Επίδοσης με Μοντέλα στη Βιβλιογραφία

Στις μελέτες που διερευνήθηκαν, παρατηρήθηκε ο ίδιος στόχος, η γενίκευση σε διαφορετικού τύπου ιστοπαθολογικές εικόνες. Όμως τα αποτελέσματα που πέτυχαν ήταν αισθητά καλύτερα. Η βασική αιτία για αυτή τη διαφορά δεν έγκειται μόνο στο είδος των δεδομένων, αλλά κυρίως στις τεχνικές που χρησιμοποιήθηκαν.

Στις εν λόγω μελέτες, αξιοποιήθηκαν προεκπαιδευμένα δίκτυα CNN (π.χ. ResNet, EfficientNet), τα οποία φέρουν γνώση από πολύ μεγάλα και γενικά datasets, όπως το ImageNet. Αυτή η γνώση μεταφέρεται και επιτρέπει στο μοντέλο να αναγνωρίσει βασικά μορφολογικά πρότυπα με λιγότερα δεδομένα και με μεγαλύτερη ακρίβεια. Επίσης, χρησιμοποιήθηκαν μετασχηματιστές (Transformers), οι οποίοι προσφέρουν ισχυρές δυνατότητες κατανόησης συσχετίσεων στην εικόνα, δηλαδή όχι μόνο τοπικά χαρακτηριστικά, αλλά και δομικά μοτίβα που εμφανίζονται σε διαφορετικές περιοχές της εικόνας.

Η προσέγγισή μου είναι εντελώς διαφορετική ως προς την υπολογιστική πολυπλοκότητα και τη στρατηγική μάθησης. Η απουσία προεκπαίδευσης, καθιστά το έργο της γενίκευσης πολύ πιο δύσκολο. Αν και το μοντέλο πέτυχε ικανοποιητικές επιδόσεις τόσο στο αρχικό σύνολο δεδομένων, GasHisSDB (περίπου 95%), όσο και στο μικρότερο και απαιτητικό MHIST (περίπου 70%).

Πίνακας 6.4 Συγκριτική αξιολόγηση του προσαρμοσμένου CNN μοντέλου και state-of-the-art μοντέλων για Γαστρικό Καρκίνο

Model	Ακρίβεια	Ευαισθησία	Ειδικότητα	Ακρίβεια Θετικών	F1-Score	AUC	MCC
Custom CNN	94.12%	93.15%	94.71%	91.52%	92.33%	0.9860	0.8757
MCAM	99.60%	-	-	-	98.5%	0.9866	-
GasHis-Transformer	97.97%	97.38%	-	98.55%	97.97%	-	-

Ο παραπάνω πίνακας συνοψίζει τις επιδόσεις του προτεινόμενου Custom CNN πάνω στο σύνολο δεδομένων GasHisSDB σε σύγκριση με δύο σύγχρονα state-of-the-art μοντέλα που έχουν χρησιμοποιηθεί για την ανάλυση ιστοπαθολογικών εικόνων γαστρικού καρκίνου. Το Custom CNN καθώς και το MCAM εκπαιδεύτηκαν και αξιολογήθηκαν στο ίδιο dataset, επιτρέποντας άμεση σύγκριση των αποτελεσμάτων τους υπό τις ίδιες συνθήκες. Αντίθετα, το μοντέλο GasHis-Transformer, αξιολογήθηκε σε διαφορετικό dataset του ίδιου κλινικού πεδίου, γεγονός που επιτρέπει ενδεικτική αλλά όχι απόλυτα ισοδύναμη σύγκριση. Παρ' όλα αυτά, τα αποτελέσματα του Custom CNN εμφανίζονται ανταγωνιστικά, παρουσιάζοντας υψηλές τιμές ευαισθησίας, ειδικότητας και ακρίβειας. Αυτή η απόδοση καταδεικνύει τη δυναμική του μοντέλου σε πραγματικά κλινικά σενάρια, και ενισχύει την αξιοπιστία του συγκριτικά με ήδη καθιερωμένες προσεγγίσεις στον χώρο.

Πίνακας 6.5 Συγκριτική αξιολόγηση του προσαρμοσμένου CNN μοντέλου και state-of-the-art μοντέλων για Καρκίνο του Παχέος Εντέρου

Model	Ακρίβεια	Ευαισθησία	Ειδικότητα	Ακρίβεια Θετικών	F1-Score	AUC	MCC
Custom CNN	73.78%	72.73%	72.81%	54.96%	62.62%	0.7943	0.4287
SVM	-	100%	-	100%	100%	-	-
CCDNet	98.96%	98.80%	99.73%	-	98.64%	0.9912	-

Ο παραπάνω πίνακας συγκρίνει την απόδοση τριών μοντέλων ταξινόμησης σε ιστοπαθολογικές εικόνες καρκίνου του παχέος εντέρου, με χρήση διαφορετικών αρχιτεκτονικών και συνόλων δεδομένων. Το πρώτο μοντέλο, το Custom CNN, αποτελεί τη δική μου πρόταση και αξιολογήθηκε στο MHIST dataset, το οποίο, όπως έχει ήδη αναλυθεί, παρουσιάζει ιδιαιτερότητες. Περιλαμβάνει μόνο καρκινικές αλλοιώσεις, με έντονο σκούρο μωβ χρωματισμό και χωρίς κανονικές εικόνες, γεγονός που πιθανώς περιόρισε τη γενικευσιμότητα του μοντέλου μας, το οποίο είχε εκπαιδευτεί σε πιο ποικιλόμορφα δεδομένα. Παρόλα αυτά, κατέγραψε ικανοποιητικές τιμές με ακρίβεια 73.78%, ευαισθησία 72.73%, και AUC 0.7943.

Αντίθετα, το SVM μοντέλο που παρουσιάζεται στη βιβλιογραφία, παρότι επιτυγχάνει τέλεια ευαισθησία και ακρίβεια θετικών (100%), δεν παρέχει πλήρη εικόνα, καθώς απουσιάζουν σημαντικές μετρικές όπως η ακρίβεια, το AUC και ο MCC, οι οποίες είναι κρίσιμες για μια ολοκληρωμένη αξιολόγηση.

Το CCDNet, ένα πιο εξελιγμένο deep learning μοντέλο, κατέγραψε εξαιρετικές επιδόσεις, με ακρίβεια σχεδόν 99% και AUC 0.9912.

Αξίζει επίσης να σημειωθεί ότι τα περισσότερα από τα συγκρινόμενα μοντέλα βασίζονται σε σύνθετες αρχιτεκτονικές και αξιοποιούν προεκπαιδευμένα μοντέλα για τη βελτιστοποίηση των επιδόσεών τους. Αντίθετα, το προτεινόμενο μοντέλο αποτελεί ένα απλό 8-layer CNN χωρίς προεκπαίδευση ή επιπλέον τεχνικές ενίσχυσης, γεγονός που καθιστά τις επιδόσεις του ακόμα πιο εντυπωσιακές.

Παρότι δεν ξεπερνά απόλυτα τα state-of-the-art μοντέλα στις μετρήσεις, επιτυγχάνει πολύ ικανοποιητικά και σταθερά αποτελέσματα, αποδεικνύοντας ότι ένα καλά σχεδιασμένο και απλοποιημένο μοντέλο μπορεί να προσφέρει αξιόπιστες λύσεις χωρίς την ανάγκη πολύπλοκων επεμβάσεων ή βαριών υπολογιστικών πόρων.

6.2. Ποσοτικοποίηση Αβεβαιότητας μοντέλου στο GasHisSDB

Μετά την εφαρμογή της μεθόδου Monte Carlo Dropout στο test set, καταγράφηκαν τα αποτελέσματα σε αρχείο .csv, όπου για κάθε εικόνα αποθηκεύτηκαν η πραγματική ετικέτα (true_label), η μέση πιθανότητα ταξινόμησης ως “Abnormal” (predicted_prob), η τελική απόφαση του μοντέλου (predicted_label) βάσει του βέλτιστου κατωφλίου, και η αβεβαιότητα της πρόβλεψης με τη μορφή της τυπικής απόκλισης (uncertainty_std). Παρακάτω, προστίθεται ένα κομμάτι του .csv αρχείου για την κατανόηση των αποτελεσμάτων ανά εικόνα.

Πίνακας 6.6 Τιμές αβεβαιότητας για μικρό αριθμό δειγμάτων

filename	true_label	predicted_prob	predicted_label	uncertainty_std
Normal-09138.png	0	0.5196140250563621	0	0.23626159439660174
Abnormal-19787.png	1	0.977973827123642	1	0.02243582295570054
Normal-02041.png	0	0.0003702644078657613	0	0.000619076429417348
Normal-29255.png	0	0.6289201551675796	1	0.1676283625763863
Abnormal-09138.png	1	0.9864827513694763	1	0.027186363040557684
Normal-39475.png	0	0.00027435315038019324	0	0.000221410535094807

Στον παραπάνω πίνακα παρουσιάζονται ενδεικτικά κάποια δείγματα από το test set. Παρατηρούμε ότι η εικόνα "Normal-09138.png", παρότι ανήκει στην κατηγορία "Normal" (true_label = 0), έλαβε σχετικά οριακή πιθανότητα 0.5196 αλλά χαμηλή εμπιστοσύνη, όπως αποτυπώνεται από την υψηλή αβεβαιότητα 0.2363. Αντίθετα, η εικόνα "Abnormal-19787.png" ταξινομήθηκε σωστά ως "Abnormal" με υψηλή πιθανότητα 0.9779 και χαμηλή αβεβαιότητα 0.0224, γεγονός που υποδηλώνει υψηλή εμπιστοσύνη του μοντέλου στην απόφασή του.

Ακολουθως, υπολογίστηκαν οι τελικές μετρήσεις αξιολόγησης χρησιμοποιώντας τις μέσες τιμές των στοχαστικών προβλέψεων. Με αυτόν τον τρόπο επιτυγχάνεται πιο σταθερή και αξιόπιστη αποτίμηση της απόδοσης του μοντέλου, σε αντίθεση με την κλασική ντετερμινιστική πρόβλεψη, η οποία μπορεί να επηρεαστεί από μεμονωμένες αποκλίσεις.

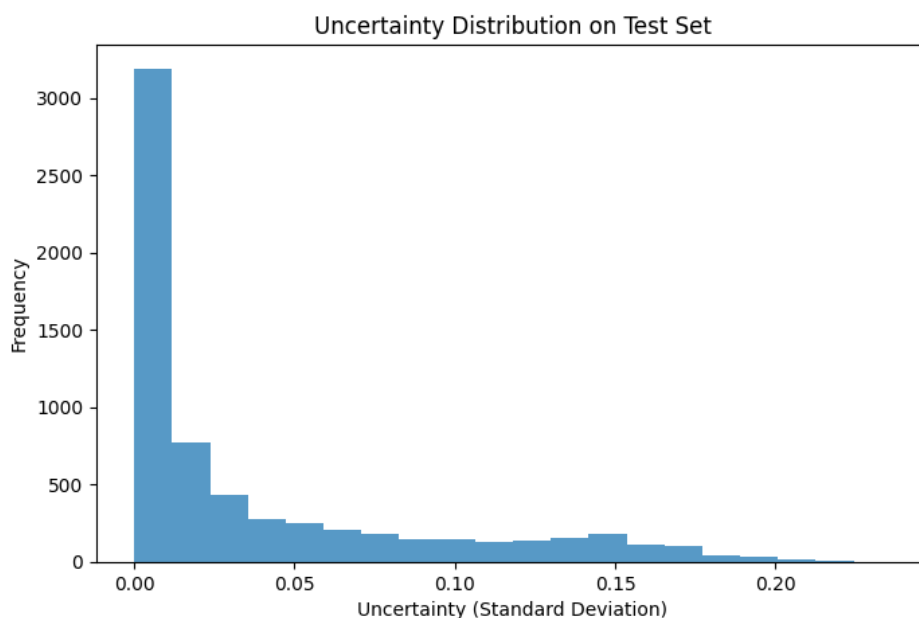
Τα αποτελέσματα συνοψίζονται στον παρακάτω πίνακα κάτω από τα αντίστοιχα αποτελέσματα του αρχικού μοντέλου, πριν τις στοχαστικές επαναλήψεις:

Πίνακας 6.7 Μετρικές αξιολόγησης του προσαρμοσμένου CNN μοντέλου στο πλαίσιο της ποσοτικοποίησης αβεβαιότητας με MC Dropout

Model	Ακρίβεια	Εναισθησία	Ειδικότητα	Ακρίβεια Θετικών	F1-Score	AUC	MCC
CNN-MC	94.84%	93.59%	95.60%	92.88%	93.23%	0.9865	0.8906

Η εφαρμογή της μεθόδου Monte Carlo Dropout (CNN-MC) οδήγησε σε γενική βελτίωση της απόδοσης του μοντέλου σε σχέση με την απλή εκδοχή του Custom CNN. Συγκεκριμένα, παρατηρείται αύξηση σε όλα τα κρίσιμα μετρικά, γεγονός που δείχνει ότι το μοντέλο όχι μόνο αναγνωρίζει πιο σωστά τα παθολογικά δείγματα, αλλά παράλληλα μειώνει και τα false positives. Παρότι η διαφορά είναι μικρή, η χρήση MC Dropout προσφέρει επιπλέον το πλεονέκτημα της ποσοτικοποιημένης αβεβαιότητας, ενισχύοντας την εμπιστοσύνη και διαφάνεια της τελικής απόφασης, στοιχείο καθοριστικής σημασίας σε εφαρμογές ιατρικής απεικόνισης.

Για την οπτική αποτύπωση της αβεβαιότητας του μοντέλου στο test set, δημιουργήθηκε ιστόγραμμα των τιμών τυπικής απόκλισης. Η κατανομή δείχνει ότι στη συντριπτική πλειοψηφία των περιπτώσεων το μοντέλο είχε χαμηλή αβεβαιότητα, γεγονός που υποδηλώνει υψηλή εμπιστοσύνη στις αποφάσεις του.



Σχήμα 6.6 Ιστόγραμμα Αβεβαιότητας

Η δυνατότητα να εντοπίζονται όχι μόνο οι προβλέψεις, αλλά και η βαθμίδα εμπιστοσύνης του μοντέλου είναι ιδιαίτερα σημαντική σε εφαρμογές όπως η ιατρική απεικόνιση. Η ύπαρξη πρόβλεψης συνοδευόμενης από υψηλή αβεβαιότητα μπορεί να λειτουργήσει ως σηματοδότης για ανθρώπινη επιθεώρηση, ενώ παράλληλα επιτρέπει την ταυτοποίηση δύσκολων περιπτώσεων που μπορούν να αξιοποιηθούν σε μελλοντικές φάσεις ενεργής μάθησης (active learning), ενισχύοντας τη συνολική ασφάλεια και απόδοση του συστήματος.

6.3.Αποτελέσματα Grad-CAM στο GasHisSDB

Όπως αναφέρθηκε και στην ενότητα της μεθοδολογίας, για την ενίσχυση της ερμηνευσιμότητας του μοντέλου και την κατανόηση των αποφάσεών του, εφαρμόστηκε η τεχνική Grad-CAM (Gradient-weighted Class Activation Mapping) τόσο κατά τη διάρκεια της διαδικασίας 5-Fold Cross-Validation, όσο και στο τελικό μοντέλο με 90-10 split. Για κάθε εικόνα του test set σε κάθε fold, δημιουργούνται δύο τύποι εξόδου:

- Το κανονικοποιημένο heatmap Grad-CAM, το οποίο αντιπροσωπεύει την ακατέργαστη ενεργοποίηση του τελευταίου συνελκτικού επιπέδου μετά από ReLU και συσσώρευση.

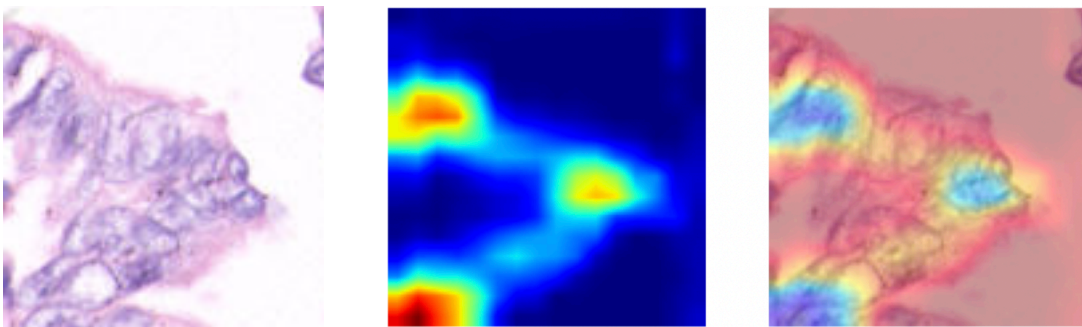
- Η τελική επεξεργασμένη εικόνα (overlayed image), στην οποία το heatmap έχει χρωματιστεί με colormap τύπου JET και επικολληθεί ημι-διαφανώς στην αρχική εικόνα εισόδου.

Ειδικότερα, στο cross validation έγινε εξαγωγή πληροφορίας για 4 περιπτώσεις ανά fold. Επομένως προέκυψαν 20 ζευγάρια heatmap-overlayed image, 5 απο κάθε κατηγορία :

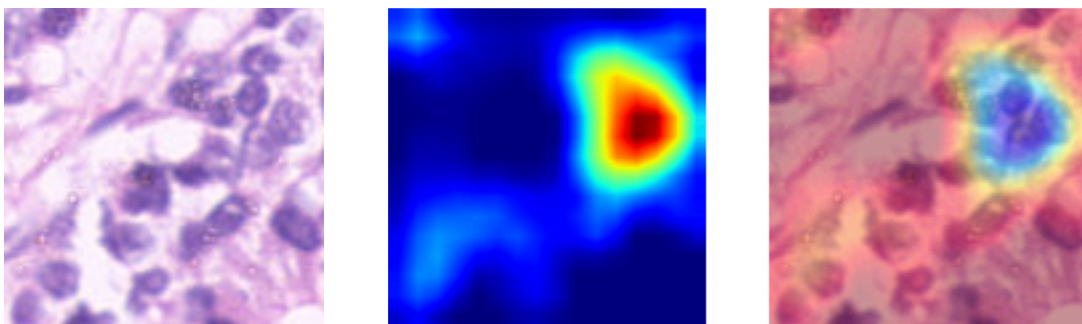
- Αληθώς Θετικά (True Positive - TP)
- Αληθώς Αρνητικά (True Negative - TN)
- Ψευδώς Θετικά (False Positive - FP)
- Ψευδώς Αρνητικά (False Negative - FN)

6.3.1.Αληθώς Θετικά

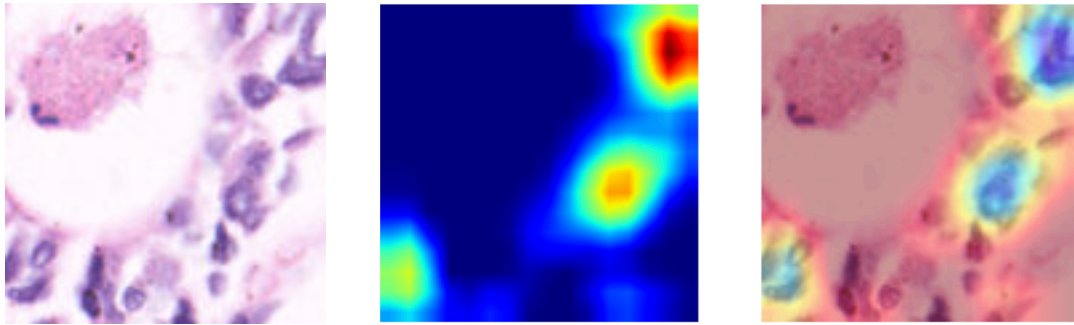
Κατά την ανάλυση των TP, διαπιστώθηκε ότι το μοντέλο εστιάζει σε μορφολογικά χαρακτηριστικά όπως πυρηνική ατυπία, κυτταρική πυκνότητα και πυρηνική υπερχρωμία. Παρακάτω απεικονίζονται μερικές TP περιπτώσεις και τα heatmaps που προέκυψαν.



A



B



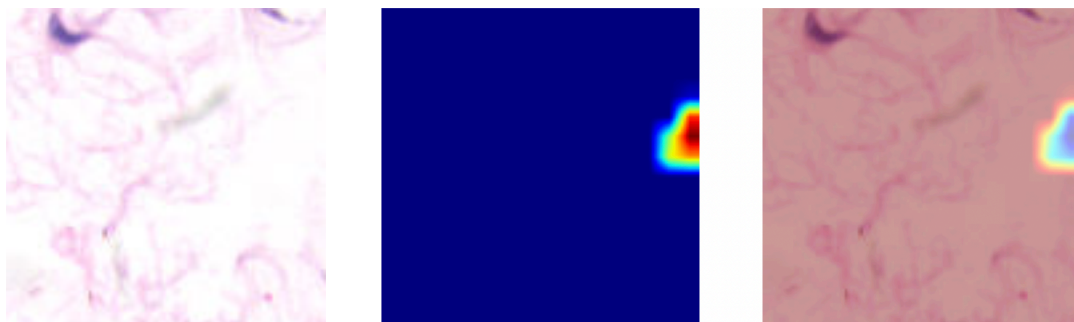
C

Σχήμα 6.1 Abnormal εικόνες με τα αντίστοιχα heatmaps και overlays
 (A)Abnormal δείγμα 23815
 (B)Abnormal δείγμα 12287
 (C)Abnormal δείγμα 12012

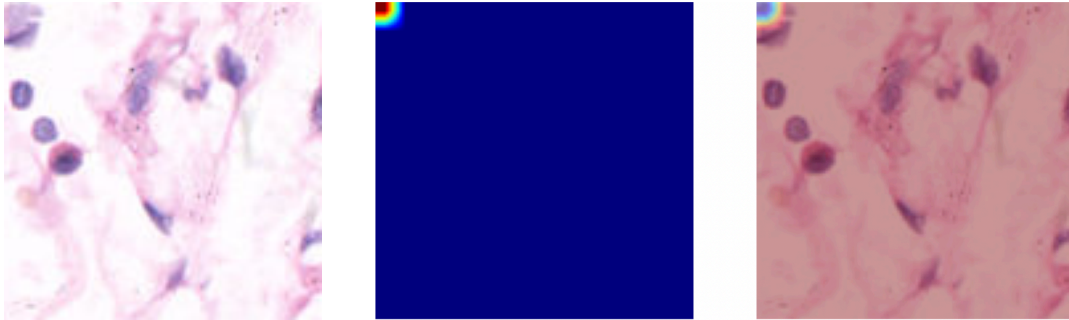
Η οπτική παρατήρηση των θερμικών χαρτών Grad-CAM δείχνει ότι η ενεργοποίηση του μοντέλου είναι συστηματικά εντοπισμένη σε κυτταρικές δομές με παθολογικά χαρακτηριστικά. Πιο συγκεκριμένα, παρατηρείται ότι τα ενεργά σημεία του Grad-CAM επικεντρώνονται γύρω από μεγάλους, σκούρους ή μορφολογικά άτυπους πυρήνες, στοιχείο που υποδηλώνει ότι το μοντέλο δεν αντιδρά τυχαία, αλλά έχει μάθει να εντοπίζει περιοχές με διαγνωστικό ενδιαφέρον.

Επιπλέον, οι περιοχές υψηλής ενεργοποίησης αντιστοιχούν συχνά σε ομάδες πυρήνων με μεγάλη πυκνότητα. Αντίθετα, σε περιοχές χωρίς κυτταρικές δομές, όπως background ή φυσιολογικό ιστό, η ενεργοποίηση είναι αμελητέα, γεγονός που δείχνει ότι το μοντέλο έχει μάθει να φιλτράρει το “θόρυβο” και να εστιάζει σε βιολογικά σημαντικές περιοχές.

6.3.2.Αληθώς Αρνητικά



A

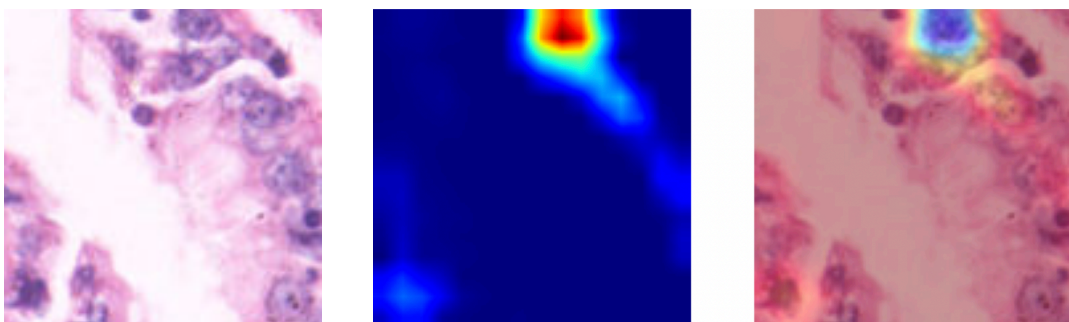


B

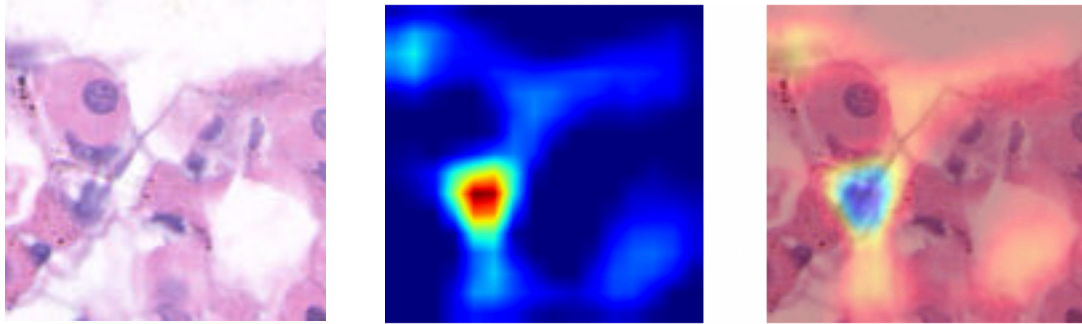
Σχήμα 6.2 Normal εικόνες με τα αντίστοιχα heatmaps και overlays
 (A)Normal δείγμα 31629
 (B)Normal δείγμα 38507

Στις True Negative περιπτώσεις, τα Grad-CAM heatmaps παρουσιάζουν ελάχιστη ή καθόλου ενεργοποίηση σε κυτταρικές δομές, γεγονός που καταδεικνύει την ικανότητα του μοντέλου να αγνοεί φυσιολογικά πρότυπα ιστού. Ακόμα και στις περιπτώσεις όπου εμφανίζεται κάποιο τοπικό hotspot, αυτό είναι απομονωμένο και δεν σχετίζεται με ομάδες κυττάρων ή μορφολογικά ύποπτους πυρήνες. Οι περιοχές αυτές είτε αντιστοιχούν σε καλά σχηματισμένους πυρήνες είτε σε άκρα του patch που πιθανώς περιέχουν τεχνητά οπτικά χαρακτηριστικά. Η συμπεριφορά αυτή ενισχύει την αξιοπιστία του μοντέλου, υποδεικνύοντας ότι η Grad-CAM ενεργοποίηση σε φυσιολογικά δείγματα περιορίζεται και δεν οδηγεί σε αποπροσανατολισμό της απόφασης.

6.3.3.Ψευδώς Θετικά



A

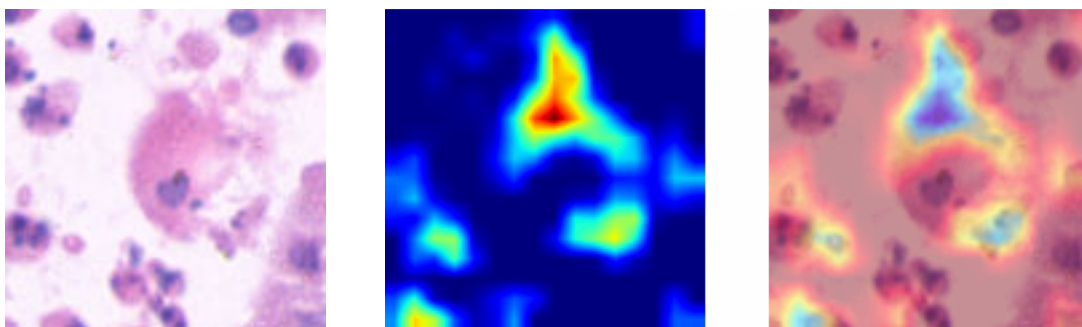


B

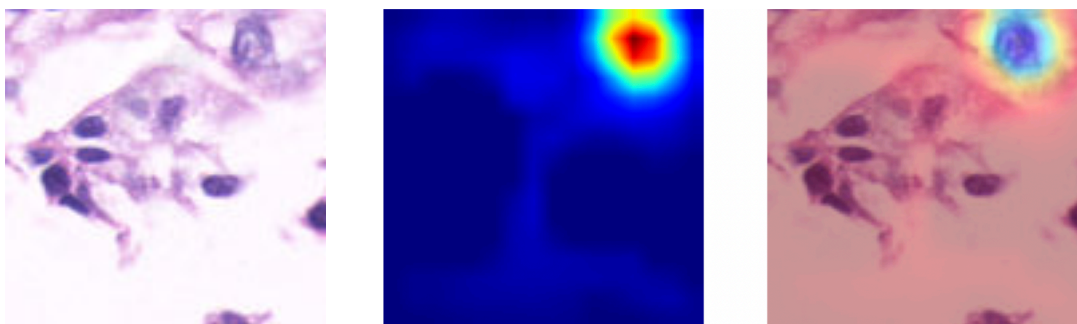
Σχήμα 6.3 Normal εικόνες με τα αντίστοιχα heatmaps και overlay
 (A)Normal δείγμα 10595
 (B)Normal δείγμα 28556

Στις False Positive (FP) περιπτώσεις, παρατηρείται ενεργοποίηση σε περιοχές που επιφανειακά μοιάζουν με παθολογικές, αλλά στην πραγματικότητα είναι φυσιολογικές. Το Grad-CAM εντοπίζει περιοχές με πυκνή συγκέντρωση κυττάρων, ή ίσως κάποια ασυνήθιστη χρώση ή υφή, αλλά χωρίς σαφείς αλλοιώσεις. Αυτό δείχνει μια τάση του μοντέλου να υπεραντιδρά σε "ύποπτα" αλλά μη παθολογικά χαρακτηριστικά, οδηγώντας σε FP. Επιπλέον, η ενεργοποίηση σε αυτές τις περιπτώσεις τείνει να είναι πιο διάχυτη ή να εμφανίζεται σε άκρες της εικόνας, γεγονός που μπορεί να υποδεικνύει σύγχυση λόγω context.

6.3.4.Ψευδώς Αρνητικά



A



B

Σχήμα 6.4 Abnormal εικόνες με τα αντίστοιχα heatmaps και overlay
(A)Abnormal δείγμα 08688
(B)Abnormal δείγμα 02045

Στις False Negative (FN) περιπτώσεις, όπου η εικόνα είναι παθολογική αλλά το μοντέλο την κατηγοριοποιεί ως Normal, το Grad-CAM heatmap παρουσιάζει ελάχιστη ή καθόλου ενεργοποίηση σε περιοχές όπου υπάρχουν εμφανώς άτυποι ή ύποπτοι πυρήνες. Η απουσία ενεργοποίησης σε αυτές τις περιπτώσεις δείχνει είτε αποτυχία γενίκευσης είτε ότι οι πυρήνες αυτοί αποκλίνουν από τα μορφολογικά πρότυπα που “έμαθε” το μοντέλο ως παθολογικά κατά την εκπαίδευση. Πρόκειται για σημαντική διαγνωστική αδυναμία, καθώς τέτοια λάθη μπορεί να οδηγήσουν σε μη έγκαιρη ανίχνευση.

Συνολικά, η σύγκριση Grad-CAM ενεργοποίησης ανάμεσα στις τέσσερις κατηγορίες (TP, TN, FP, FN) προσφέρει κρίσιμη εικόνα για το πώς και γιατί το μοντέλο λαμβάνει τις αποφάσεις του. Ιδιαίτερη προσοχή πρέπει να δοθεί στις FP και FN περιπτώσεις κατά τη βελτίωση ή ανατροφοδότηση του μοντέλου με πιο ισορροπημένο training ή ανασχεδιασμό του loss function με βάση ενεργοποιήσεις.

6.4. Εξαγωγή χαρακτηριστικών Αληθώς Θετικών Προβλέψεων με HistomicsTK

Το εργαλείο HistomicsTK αξιοποιήθηκε, όπως αναφέρθηκε προηγουμένως, για την εξαγωγή χαρακτηριστικών και την τμηματοποίηση των καρκινικών πυρήνων. Ειδικότερα, χρησιμοποιήθηκε για την απομόνωση και ανάλυση πυρήνων στις Abnormal εικόνες. Η βασική ιδέα ήταν να εντοπιστούν οι πυρήνες, με βάση τη χρώση αιματοξυλίνης, να εξαχθούν από αυτούς μορφολογικά χαρακτηριστικά και χαρακτηριστικά υφής και να αποθηκευτούν για περαιτέρω ανάλυση. Αυτό δημιούργησε έναν πλούσιο ποσοτικό χάρτη που μπορεί να συνδυαστεί με τους Grad-CAM heatmaps του μοντέλου, για βαθύτερη κατανόηση της απόφασης του μοντέλου.

6.4.1.Απεικόνιση Ενεργοποίησης Πυρήνων

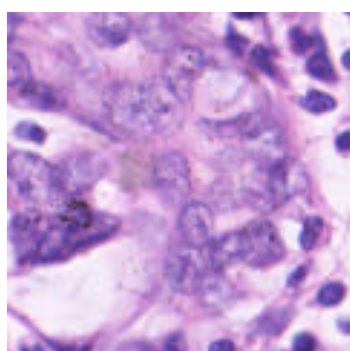
Για τις εικόνες που απεικονίζουν τα heatmaps και για τις εικόνες στις οποίες σημειώθηκαν οι πυρήνες, μέσω segmentation, αποθηκεύτηκαν τα numpy arrays, (μάσκες). Έτσι, πέρα από την οπτική σύγκρισή τους, είχαμε και χωρική συσχέτιση η οποία έφερε πιο ειδική πληροφορία.

Αυτή η διαδικασία έγινε για τα πέντε True Positive δείγματα που προέκυψαν από το Grad-CAM. Σκοπός ήταν η δομική σύγκριση, προκειμένου να εντοπιστούν πόσοι και ποιοι πυρήνες ενεργοποιήθηκαν από το μοντέλο σε κάθε εικόνα. Συγκεκριμένα, για κάθε εικόνα φορτώθηκε το αντίστοιχο heatmap, σε μορφή .npy και η μάσκα πυρήνων προηγουμένως αποθηκευμένη ως .npy από HistomicsTK.

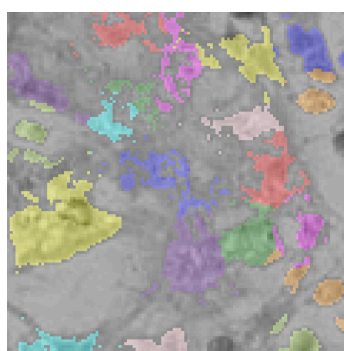
Πίνακας 6.8 Αριθμών συνολικών και ενεργών πυρήνων

fold	image	total_nuclei	active_nuclei	percent_in_GradCAM
1	Abnormal-01902	20	9	45.0
2	Abnormal-23815	11	6	54.55
3	Abnormal-12012	15	9	60.0
4	Abnormal-16643	18	7	38.89
5	Abnormal-12287	20	6	30.0

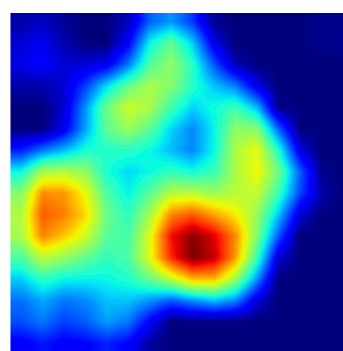
Παρακάτω απεικονίζονται για μερικές εικόνες, η αρχική της μορφή χωρίς επεξεργασία, η τμηματοποίηση των πυρήνων της και το heatmap που προέκυψε από το gradcam.



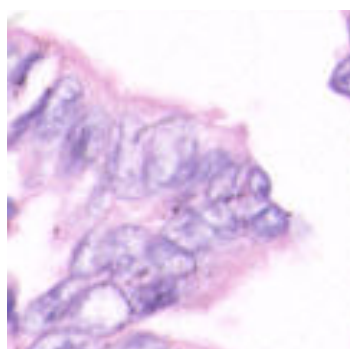
Abnormal-01902



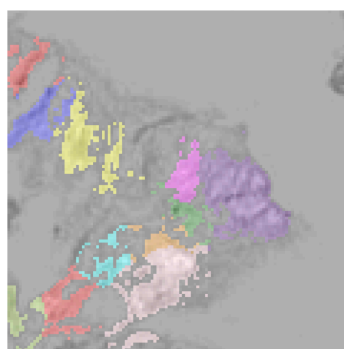
01902-Nuclei



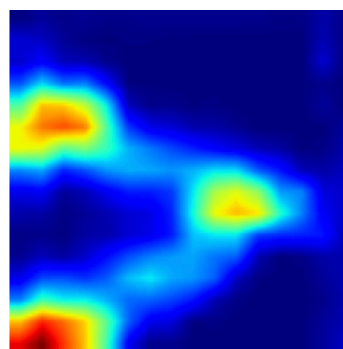
01902-Heatmap



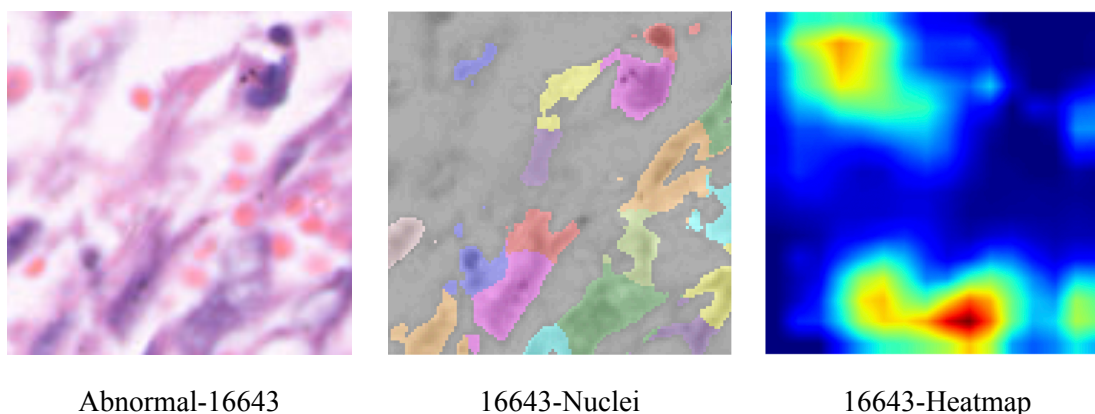
Abnormal-23815



23815-Nuclei



23815-Heatmap



Σχήμα 6.5 Abnormal εικόνες, τμηματοποιημένοι πυρήνες, heatmaps

Έπειτα, το heatmap μετατρέπεται σε δυαδική μάσκα ενεργοποίησης με κατώφλι (threshold=0.5). Για κάθε πυρήνα που έχει μοναδικό label μέσα στη μάσκα, ελέγχεται αν επικαλύπτεται με ενεργοποιημένη περιοχή του Grad-CAM. Αν υπάρχει έστω ένα κοινό pixel, θεωρείται ότι αυτός ο πυρήνας είναι "ενεργοποιημένος". Καταγράφτηκε ο συνολικός αριθμός πυρήνων, ο αριθμός των ενεργών και το ποσοστό ενεργών πυρήνων επί του συνόλου.

Αυτό το βήμα προσφέρει ποσοτική σύνδεση μεταξύ των Grad-CAM ενεργοποιήσεων και των βιολογικών μονάδων, δηλαδή των πυρήνων. Αντί να εξετάζουμε μόνο οπτικά "που κοιτάει" το δίκτυο, μπορούμε πλέον να ποσοτικοποιήσουμε τα δεδομένα. Αυτός ο συνδυασμός Grad-CAM και masks αποτελεί ισχυρό εργαλείο ερμηνευσιμότητας (explainability) και ενισχύει τη βιολογική τεκμηρίωση των αποφάσεων ενός deep learning μοντέλου στην παθολογία.

6.4.2.Εξαγωγή Μορφολογικών, Γεωμετρικών Χαρακτηριστικών και Χαρακτηριστικών Υφής και Έντασης

Σε δεύτερο στάδιο, συνδυάστηκαν τα ποσοτικά χαρακτηριστικά που εξήχθησαν για κάθε πυρήνα μέσω του HistomicsTK με την πληροφορία που παρέχεται από τα Grad-CAM heatmaps σχετικά με το αν ένας πυρήνας συνέβαλε στην τελική απόφαση του μοντέλου. Για κάθε πυρήνα σε κάθε εικόνα, προστίθεται μια δυαδική στήλη, η οποία παίρνει τιμή 1 αν ο πυρήνας επικαλύπτεται χωρικά με ενεργή περιοχή του Grad-CAM και 0 διαφορετικά. Αυτή η διεργασία έγινε ξεχωριστά για το είδος χαρακτηριστικών κάθε εικόνας.

Αυτός ο εμπλουτισμός των δεδομένων επιτρέπει τη διασύνδεση της ερμηνευσιμότητας του μοντέλου (μέσω Grad-CAM) με τα πραγματικά βιομορφολογικά χαρακτηριστικά των κυττάρων. Κατ' επέκταση, διευκολύνεται η διερεύνηση για την διάκριση των χαρακτηριστικών και των τιμών τους στους ενεργοποιημένους πυρήνες.

Προκειμένου να ελεγχθεί αν τα χαρακτηριστικά των πυρήνων που ενεργοποιούνται από το Grad-CAM διαφέρουν στατιστικά από εκείνα που δεν ενεργοποιούνται, εφαρμόστηκε το μη παραμετρικό τεστ Mann-Whitney U[63]. Η ανάλυση εκτελέστηκε ξεχωριστά για κάθε εικόνα και για κάθε τύπο χαρακτηριστικών (μορφομετρικά, FSD, υφής και gradient), συγκρίνοντας τις τιμές μεταξύ ενεργών και ανενεργών πυρήνων.

Το Mann-Whitney U είναι ένα μη παραμετρικό στατιστικό τεστ που χρησιμοποιείται για να συγκρίνει ανεξάρτητα δείγματα ως προς τη διαφορά στην κατανομή τους. Δεν προϋποθέτει κανονική κατανομή των δεδομένων, γεγονός που το καθιστά ιδανικό για βιολογικά δεδομένα όπως χαρακτηριστικά κυττάρων, τα οποία συχνά παρουσιάζουν ασυμμετρία, outliers ή skewness.

Στο πλαίσιο αυτής της ανάλυσης, το τεστ εφαρμόζεται για κάθε χαρακτηριστικό πυρήνα (π.χ. έκταση, περιφέρεια, υφή κ.λπ.), συγκρίνοντας:

- Την ομάδα των πυρήνων που συμμετείχαν στην απόφαση του μοντέλου (gradcam_active = 1),
- Με την ομάδα πυρήνων που αγνοήθηκαν από το μοντέλο (gradcam_active = 0).

Το αποτέλεσμα του τεστ είναι:

- Η στατιστική U, που μετρά το overlap των τάξεων μεταξύ των ομάδων,
- Και η p-value, που δείχνει πόσο πιθανό είναι να παρατηρήσουμε τέτοια διαφορά αν οι πυρήνες είχαν όντως ίδια κατανομή.

Μια χαμηλή p-value (π.χ. < 0.05) υποδηλώνει ότι το συγκεκριμένο χαρακτηριστικό διαφέρει σημαντικά μεταξύ ενεργών και ανενεργών πυρήνων και πιθανόν υπάρχει ουσιαστική βιολογική αιτία, και άρα σχετίζεται πιθανώς με τη λειτουργία του Grad-CAM και την τελική απόφαση του μοντέλου.

Για κάθε εικόνα, υπολογίστηκε ο μέσος όρος της τιμής του χαρακτηριστικού ξεχωριστά για την ομάδα των ενεργοποιημένων και των μη ενεργοποιημένων πυρήνων. Παράλληλα, εφαρμόστηκε το Mann-Whitney U ώστε να αξιολογηθεί αν οι διαφορές αυτές είναι στατιστικά σημαντικές σε επίπεδο κατανομής. Οι μέσοι όροι παρέχουν μια συνοπτική εικόνα της τάσης, ενώ η p-value τεκμηριώνει τη στατιστική αξιοπιστία της διαφοράς.

Παρακάτω παρουσιάζονται τα αποτελέσματα από αυτήν την ανάλυση για κάθε τύπο χαρακτηριστικών.

Γεωμετρικά Χαρακτηριστικά (FSD)

Πίνακας 6.9 FSD χαρακτηριστικά με διαφορά τιμών μεταξύ ενεργών και ανενεργών πυρήνων

image_id	feature	p-value
Abnormal-12012	Shape.FSD6	0.025574
Abnormal-16643	Shape.FSD3	0.020425
Abnormal-16643	Shape.FSD4	0.008296
Abnormal-16643	Shape.FSD5	0.000440

Η στατιστική ανάλυση έδειξε ότι στις εικόνες Abnormal-12012 και Abnormal-16643, ορισμένα FSD χαρακτηριστικά παρουσίασαν σημαντικές διαφορές μεταξύ των ενεργών και μη ενεργών πυρήνων. Συγκεκριμένα, τα Shape.FSD3, FSD4, FSD5 και FSD6 εμφάνισαν p-values κάτω από το 0.05, υποδεικνύοντας ότι σχετίζονται με την ενεργοποίηση του μοντέλου. Αυτό ενισχύει την υπόθεση ότι το μοντέλο λαμβάνει υπόψη συγκεκριμένες όψεις του σχήματος και της περιφέρειας των πυρήνων κατά τη λήψη απόφασης.

Μορφολογικά Χαρακτηριστικά (Morphometric)

Πίνακας 6.10 Μορφολογικά χαρακτηριστικά με διαφορά τιμών μεταξύ ενεργών και ανενεργών πυρήνων

image_id	feature	p-value
Abnormal-12012	Size.Area	0.038995
Abnormal-12012	Size.Perimeter	0.017582
Abnormal-12012	Shape.EquivalentDiameter	0.038995
Abnormal-12287	Size.Area	0.040764
Abnormal-12287	Size.ConvexHullArea	0.032559
Abnormal-12287	Size.MinorAxisLength	0.025645
Abnormal-12287	Size.Perimeter	0.011713
Abnormal-12287	Shape.EquivalentDiameter	0.040764
Abnormal-12287	Shape.FractalDimension	0.025841

Στο σετ μορφολογικών χαρακτηριστικών (morpho), η στατιστική ανάλυση έδειξε ότι στις εικόνες Abnormal-12012 και Abnormal-12287, αρκετά μεγέθη που σχετίζονται με την έκταση και τη γεωμετρία των πυρήνων διαφοροποιούνται σημαντικά μεταξύ των ενεργών και μη ενεργών πυρήνων. Συγκεκριμένα, χαρακτηριστικά όπως Size.Area, Perimeter, EquivalentDiameter και FractalDimension παρουσίασαν χαμηλά p-values, γεγονός που υποδηλώνει ότι το μοντέλο ενεργοποιείται κατά κύριο λόγο σε πυρήνες με μεγαλύτερο μέγεθος ή αυξημένη μορφολογική πολυπλοκότητα. Ειδικότερα, το χαρακτηριστικό Fractal Dimension, που εκφράζει το βαθμό τραχύτητας του περιγράμματος ενός πυρήνα, επιτρέπει τον εντοπισμό ακανόνιστων ή ανώμαλων σχημάτων τα οποία συναντώνται συχνά σε δυσπλαστικούς ή κακοήθεις ιστούς.

Αυτή η ανάλυση ενισχύει την ερμηνευσιμότητα του Grad-CAM, καθώς οι συγκεκριμένοι μορφολογικοί δείκτες είναι συχνά ενδεικτικοί παθολογικών καταστάσεων, τέτοιοι πυρήνες σχετίζονται συχνά με δυσπλασία ή παθολογικές αλλοιώσεις.

Χαρακτηριστικά Έντασης (Gradient-based)

Πίνακας 6.11 Χαρακτηριστικά διαβάθμισης με διαφορά τιμών μεταξύ ενεργών και ανενεργών πυρήνων

image_id	feature	p-value
Abnormal-01902	Gradient.Mag.Mean	0.001840
Abnormal-01902	Gradient.Mag.Std	0.006237
Abnormal-01902	Gradient.Mag.Kurtosis	0.040239
Abnormal-01902	Gradient.Mag.HistEntropy	0.040239
Abnormal-01902	Gradient.Canny.Sum	0.022654
Abnormal-01902	Gradient.Canny.Mean	0.009792
Abnormal-12287	Gradient.Canny.Sum	0.020021

Η ανάλυση gradient-based χαρακτηριστικών έδειξε σημαντικές διαφορές μεταξύ ενεργών και μη ενεργών πυρήνων, ιδίως στην εικόνα Abnormal-01902. Στατιστικά σημαντική διαφορά παρατηρήθηκε σε τιμές όπως το Gradient.Mag.Mean, που εκφράζει τη μέση ένταση των μεταβολών φωτεινότητας, καθώς και την τυπική απόκλιση, Gradient.Mag.Std, η οποία καταγράφει την ετερογένεια των μεταβολών εντός κάθε πυρήνα. Επιπλέον, το χαρακτηριστικό Kurtosis ανέδειξε διαφορές στην κυρτότητα της κατανομής των μεταβολών, με υψηλότερες τιμές να υποδηλώνουν πιο αιχμηρές ή έντονες μεταβολές στην εικόνα. Τέλος, τα χαρακτηριστικά Canny.Mean και Canny.Sum, που προέρχονται από τον αλγόριθμο ανίχνευσης ακμών Canny, αποτυπώνουν την παρουσία και ένταση των ορίων των πυρήνων, δείχνοντας ότι οι ενεργοποιημένοι πυρήνες παρουσιάζουν πιο έντονες ακμές. Συνολικά, τα αποτελέσματα υποδεικνύουν ότι οι ενεργοποιημένοι πυρήνες εμφανίζουν εντονότερες μεταβολές έντασης και μεγαλύτερη πολυπλοκότητα περιγράμματος. Το μοντέλο εντοπίζει παθολογικά χαρακτηριστικά μέσω της ανάλυσης τοπικής δομής και ορίων των πυρήνων.

Χαρακτηριστικά Υφής (Haralick)

Τα χαρακτηριστικά Haralick ενσωματώνουν ιδιότητες υφής και έντασης των πυρήνων, και η στατιστική σύγκριση ανέδειξε ισχυρές διαφορές μεταξύ Grad-CAM ενεργών και μη ενεργών πυρήνων. Ειδικά στην εικόνα Abnormal-01902, καταγράφηκαν στατιστικά σημαντικές διαφορές σε contrast, correlation, entropy, IMC και άλλες παραμέτρους ($p < 0.01$), γεγονός που υποδηλώνει ότι το μοντέλο αξιοποιεί πληροφορία σχετική με τη μικροδομή και τη χρωματική ποικιλία των πυρήνων. Αντίστοιχα, σε Abnormal-12012 και Abnormal-12287, παρατηρούνται διαφοροποιήσεις που εδραιώνουν τη σημασία της υφής στη διαδικασία ενεργοποίησης του μοντέλου.

Πίνακας 6.12 Χαρακτηριστικά Haralick με διαφορά τιμών μεταξύ ενεργών και ανενεργών πυρήνων

image_id	feature	p-value
Abnormal-01902	Haralick.Contrast.Mean	0.002374
Abnormal-01902	Haralick.Contrast.Range	0.003047
Abnormal-01902	Haralick.Correlation.Mean	0.004938
Abnormal-01902	Haralick.SumOfSquares.Mean	0.002374
Abnormal-01902	Haralick.SumOfSquares.Range	0.040239
Abnormal-01902	Haralick.SumAverage.Mean	0.027578
Abnormal-01902	Haralick.SumVariance.Mean	0.002374
Abnormal-01902	Haralick.SumVariance.Range	0.003889
Abnormal-01902	Haralick.SumEntropy.Mean	0.003889
Abnormal-01902	Haralick.DifferenceVariance.Mean	0.012172
Abnormal-01902	Haralick.DifferenceEntropy.Mean	0.003889
Abnormal-01902	Haralick.IMC1.Mean	0.000829
Abnormal-01902	Haralick.IMC2.Mean	0.001087
Abnormal-01902	Haralick.IMC2.Range	0.001087
Abnormal-12012	Haralick.SumAverage.Range	0.049550
Abnormal-12012	Haralick.SumEntropy.Range	0.049550
Abnormal-12012	Haralick.Entropy.Mean	0.049550
Abnormal-12287	Haralick.ASM.Mean	0.040764
Abnormal-12287	Haralick.SumAverage.Range	0.020021
Abnormal-12287	Haralick.Entropy.Mean	0.025645
Abnormal-12287	Haralick.DifferenceVariance.Mean	0.040764
Abnormal-12287	Haralick.DifferenceVariance.Range	0.040764
Abnormal-12287	Haralick.DifferenceEntropy.Mean	0.025645
Abnormal-12287	Haralick.IMC1.Mean	0.025645

Η συσχέτιση των ενεργοποιημένων περιοχών του Grad-CAM με μεμονωμένους πυρήνες, αναδεικνύει μια σημαντική διάσταση στην ερμηνευσιμότητα των deep learning μοντέλων στη διαγνωστική παθολογία. Μέσω της ανάλυσης, παρατηρήθηκε ότι οι περιοχές που το μοντέλο θεωρεί σημαντικές συχνά αντιστοιχούν σε πυρήνες με αυξημένη υφή, έντονη ακτογραφία ή μορφολογικές ιδιαιτερότητες, χαρακτηριστικά που είναι ήδη γνωστά στον παθολογοανατόμο ως ενδείξεις δυσπλασίας ή κακοήθειας. Αν ένα τέτοιο εργαλείο

ενσωματωθεί στην καθημερινή πρακτική, θα μπορούσε να λειτουργήσει ως συμπληρωματικός “ψηφιακός βοηθός”, επισημαίνοντας περιοχές ενδιαφέροντος με τρόπο τεκμηριωμένο και διαφανή. Ο παθολογοανατόμος, έχοντας πρόσβαση τόσο στην τελική πρόβλεψη όσο και στα χαρακτηριστικά που οδήγησαν σε αυτή, θα μπορεί όχι μόνο να επιβεβαιώσει ή να απορρίψει την εκτίμηση, αλλά και να ενισχύσει την τεκμηρίωση της γνωμάτευσης, συνδυάζοντας τη βαθιά του εμπειρία με τη στοχευμένη υποστήριξη της τεχνητής νοημοσύνης.

7. Συμπεράσματα και Μελλοντική Εργασία

7.1. Σύνοψη

Η παρούσα διπλωματική εργασία επικεντρώθηκε στην ανάπτυξη, εκπαίδευση και αξιολόγηση ενός συνελκτικού νευρωνικού δικτύου (CNN) για την ταξινόμηση ιστοπαθολογικών εικόνων καρκίνου, με κύρια εφαρμογή στη διάγνωση γαστρικού και παχέος εντέρου. Αρχικά, παρουσιάστηκε το θεωρητικό υπόβαθρο γύρω από τον καρκίνο, τις ιστοπαθολογικές τεχνικές, καθώς και τις βασικές αρχές της βαθιάς μάθησης και των CNNs. Εξετάστηκαν σύγχρονες μελέτες και μοντέλα state-of-the-art, τα οποία συνέβαλαν στην καθοδήγηση της μεθοδολογίας.

Το κύριο πείραμα βασίστηκε στο dataset GasHisSDB, με στόχο την κατηγοριοποίηση μεταξύ φυσιολογικού και παθολογικού ιστού του στομάχου. Σχεδιάστηκε ένα απλό αλλά αποτελεσματικό CNN 8 επιπέδων, το οποίο εκπαιδεύτηκε και αξιολογήθηκε σε περίπου 60.000 εικόνες διαστάσεων 120x120 pixels. Η βελτιστοποίηση πραγματοποιήθηκε με Stratified k-fold cross-validation και Grid Search για την επιλογή υπερπαραμέτρων. Η απόδοση αξιολογήθηκε με τις απαραίτητες μετρήσεις, ενώ πραγματοποιήθηκε και τελική εκπαίδευση με μοίρασμα δεδομένων 90%-10%.

Για τη ερμηνεία των αποτελεσμάτων, εφαρμόστηκε η τεχνική Grad-CAM, προκειμένου να απεικονιστούν οι περιοχές ενεργοποίησης του μοντέλου. Επιπλέον, χρησιμοποιήθηκε το εργαλείο HistomicsTK για την εξαγωγή μορφολογικών χαρακτηριστικών των κυτταρικών πυρήνων εντός των περιοχών ενεργοποίησης, ενισχύοντας τη βιολογική ερμηνεία. Η αβεβαιότητα των προβλέψεων ποσοτικοποιήθηκε με χρήση Monte Carlo Dropout.

Για την αξιολόγηση της γενικευσιμότητας και την ανάδειξη της δυναμικής του εφαρμογής σε διαφορετικά ιστολογικά περιβάλλοντα, το ίδιο μοντέλο επανεκπαιδεύτηκε από την αρχή στο dataset MHIST, το οποίο αφορά καρκίνο του παχέος εντέρου, με στόχο τη διάκριση μεταξύ HP και SSA.

7.2. Συμπεράσματα

Η εργασία ανέδειξε τη δυνατότητα ενός σχετικά απλού συνελκτικού νευρωνικού δικτύου να επιτύχει αξιόλογη απόδοση στην ταξινόμηση ιστοπαθολογικών εικόνων καρκίνου του στομάχου. Παρά την απουσία ιδιαίτερα βαθιών ή προεκπαιδευμένων αρχιτεκτονικών, η απόδοση του μοντέλου ήταν ανταγωνιστική σε σχέση με τα state-of-the-art μοντέλα, επιβεβαιώνοντας ότι η προσεκτική σχεδίαση, η κατάλληλη προεπεξεργασία και η συστηματική αξιολόγηση μπορούν να αντισταθμίσουν την πολυπλοκότητα.

Η εφαρμογή τεχνικών όπως η 5-fold cross-validation και το grid search συνέβαλαν στη σταθερότητα και αξιοπιστία των αποτελεσμάτων, ενώ η χρήση Monte Carlo Dropout επέτρεψε τη μέτρηση αβεβαιότητας, ενισχύοντας τη διαφάνεια του μοντέλου. Επιπλέον, η ποσοτική σύγκριση με μορφολογικά χαρακτηριστικά πυρήνων πρόσφερε ένα πρώτο βήμα

βελτίωσης ερμηνείας της βιοϊατρικής εικόνας σε συνδυασμό με την ερμηνευσιμότητα του μοντέλου.

Αξιοσημείωτο εύρημα αποτελεί η δυνατότητα του ίδιου αρχιτεκτονικού σχήματος να αποδώσει ικανοποιητικά και σε ένα εντελώς διαφορετικό dataset (MHIST), με ανεξάρτητη εκπαίδευση. Αυτό το αποτέλεσμα αναδεικνύει τη γενικευσιμότητα της προσέγγισης και υποδεικνύει ότι η μεταφορά τεχνογνωσίας μεταξύ διαφορετικών τύπων ιστοπαθολογικών δεδομένων είναι εφικτή, ακόμη και χωρίς ειδικά προσαρμοσμένες τεχνικές.

7.3.Μελλοντική Εργασία

Σε κάθε περίπτωση είναι σαφές ότι είναι δυνατή η περαιτέρω βελτίωση. Η χρήση ήδη εκπαιδευμένων CNNs και συνδυαστικών τεχνικών θα μπορούσε να ενισχύσει τη σταθερότητα και γενικευσιμότητα του μοντέλου, ενώ η ενσωμάτωση τεχνικών ημειπιβλεπόμενης μάθησης ενδέχεται να αξιοποιήσει καλύτερα τις μεγάλες ποσότητες μη επισημασμένων ιατρικών δεδομένων. Ειδικά σε ό,τι αφορά το dataset MHIST, μελλοντική εργασία θα μπορούσε να επικεντρωθεί στην εφαρμογή τεχνικών μεταφοράς μάθησης (transfer learning), ενεργού μάθησης (active learning) και προσαρμογής περιοχής (domain adaptation), προκειμένου να αξιοποιηθεί η γνώση από το αρχικό σύνολο δεδομένων και να μειωθεί η ανάγκη για πλήρη εκπαίδευση από την αρχή σε κάθε νέα ταξινομητική εργασία.

Οι μελλοντικές κατευθύνσεις στην εφαρμογή της βαθιάς μάθησης στη βιοϊατρική περιλαμβάνουν τη δημιουργία ενοποιημένων συστημάτων που συνδυάζουν εικόνες, κλινικά δεδομένα και γενετικές πληροφορίες, καθώς και τη μετάβαση προς πιο διαφανή και αξιόπιστα μοντέλα, ικανά να ενσωματωθούν στην καθημερινή ιατρική πρακτική ως εργαλεία υποβοήθησης στη διάγνωση και πρόγνωση. Μέσα από τέτοιες προσεγγίσεις, η βαθιά μάθηση μπορεί να συμβάλει ουσιαστικά στη βελτίωση της εξατομικευμένης φροντίδας ασθενών και της ορθότητας των ιατρικών αποφάσεων.

Βιβλιογραφία

- [1] Z. Qiu *et al.*, “Decomposing the global burden of cancers: quantifying the contributions of disease severity changes across regions and time (1990–2021),” *Journal of Global Health*, vol. 15, Jun. 2025, doi: <https://doi.org/10.7189/jogh.15.04161>.
- [2] Moh.gov.gr, “Παγκόσμια Ημέρα κατά του Καρκίνου,” Υπουργείο Υγείας, Feb. 03, 2013. <https://www.moh.gov.gr/articles/news/1574-pagkosmia-hmera-kata-toy-karkinoy>
- [3] Weinberg, R. A., *The Biology of Cancer*, 2nd edition, 2013
- [4] “Καρκίνος,” Εθνικός Οργανισμός Δημόσιας Υγείας. <https://eody.gov.gr/disease/karkinos/>
- [5] C. Clinic, “Cancer treatments kill cancer cells and stop cancer spread. Common treatment options include surgery, chemotherapy, radiation, immunotherapy and targeted therapy,” Cleveland Clinic, Sep. 27, 2024. <https://my.clevelandclinic.org/health/treatments/cancer-treatment>
- [6] “Θεραπεία του Καρκίνου – Contemporary Oncology Team,” Oncologists.gr, 2025. <https://oncologists.gr/cancer-treatment-2/> (accessed Jun. 30, 2025).
- [7] “World Cancer Day: Safe Cancer Treatment With Radiotherapy | IAEA,” [iaea.org](https://www.iaea.org/newscenter/multimedia/photoessays/world-cancer-day-safe-cancer-treatment-with-radiotherapy), Feb. 04, 2017. <https://www.iaea.org/newscenter/multimedia/photoessays/world-cancer-day-safe-cancer-treatment-with-radiotherapy>
- [8] “Καρκίνος Στομάχου: Τύποι, παράγοντες κινδύνου, σταδια, θεραπεία,” Δημήτρης Γιάλβαλης, Aug. 2024. <https://gialvalis.gr/karkinos-stomaxou/> (accessed Jun. 30, 2025).
- [9] GIS, “Stomach Cancer,” Gastrointestinal Society. <https://badgut.org/information-centre/a-z-digestive-topics/stomach-cancer/>
- [10] StoryMD, “Adenocarcinoma of the Stomach,” StoryMD.com, 2025. <https://storymd.com/journal/mqle8qaij-adenocarcinoma/page/xpv8391dx6ne-adenocarcinoma-of-the-stomach#iss=https%3A%2F%2Fauth.storymd.com%2Fauth%2Frealms%2Fstorymd> (accessed Jun. 30, 2025).
- [11] National Cancer Institute, “Stages of Stomach Cancer - NCI,” [www.cancer.gov](https://www.cancer.gov/types/stomach/stages), Apr. 12, 2023. <https://www.cancer.gov/types/stomach/stages>
- [12] F.-H. Zhu *et al.*, “The Histopathological Types and Distribution Characteristics of Gastric Mixed Tumors,” *Frontiers in Oncology*, vol. 12, Jun. 2022, doi: <https://doi.org/10.3389/fonc.2022.873005>.
- [13] H. Team, “Καρκίνος παχέος εντέρου: Νεότερα δεδομένα στην έγκαιρη διάγνωση & θεραπεία | Νοσοκομείο ΥΓΕΙΑ,” [hygeia.gr](https://www.hygeia.gr/karkinos-pacheos-enteroy-neoteradedomena-stin-egkairi-diagnosi-amp-therapeia/), Sep. 29, 2021. <https://www.hygeia.gr/karkinos-pacheos-enteroy-neoteradedomena-stin-egkairi-diagnosi-amp-therapeia/> (accessed Jan. 05, 2025).
- [14] Cleveland Clinic, “Colorectal (colon) cancer,” Cleveland Clinic, Nov. 14, 2022. <https://my.clevelandclinic.org/health/diseases/14501-colorectal-colon-cancer>
- [15] “Καρκίνος παχέος εντέρου συμπτώματα διάγνωση θεραπεία - Ι. Μπολάνης,” Ιωάννης Μπολάνης Χειρουργός, Nov. 20, 2023. <https://bolanis.gr/karkinos-paxeos-enterou/>
- [16] “Πολύποδες εντέρου | Συμπτώματα & αντιμετώπιση | Proctolab Χειρουργικό - Πρωκτολόγος Θεσσαλονίκη,” Proctolab Χειρουργικό - Πρωκτολογικό Κέντρο Β. Ελλάδος, Jul. 31, 2023. <https://www.proctolab.gr/diseases/polypodes-enterou/> (accessed Jun. 30, 2025).
- [17] M. Rees, “What is a hyperplastic polyp in the colon?,” [Medicalnewstoday.com](https://www.medicalnewstoday.com/articles/hyperplastic-polyp-in-colon), Aug. 02, 2023. <https://www.medicalnewstoday.com/articles/hyperplastic-polyp-in-colon> (accessed Jun. 30, 2025).
- [18] M. Unlu, E. Uzun, G. Bengi, O. Sagol, and S. Sarioglu, “Molecular Characteristics of Colorectal Hyperplastic Polyp Subgroups,” *The Turkish Journal of Gastroenterology*, May 2020, doi: <https://doi.org/10.5152/tjg.2020.19322>.

- [19] M. X. Ma and M. J. Bourke, “Sessile Serrated Adenomas: How to Detect, Characterize and Resect,” *Gut and Liver*, vol. 11, no. 6, pp. 747–760, May 2017, doi: <https://doi.org/10.5009/gnl16523>.
- [20] D. K. Rex et al., “Serrated Lesions of the Colorectum: Review and Recommendations From an Expert Panel,” *The American journal of gastroenterology*, vol. 107, no. 9, p. 1315, Sep. 2012, doi: <https://doi.org/10.1038/ajg.2012.161>.
- [21] Affidea, Affidea, Feb. 02, 2022. <https://affidea.gr/el/news/o-rolos-tis-iatrikhs-apeikonisis-sthn-ogkologia/> (accessed Jun. 30, 2025).
- [22] L. Fass, “Imaging and cancer: A review,” *Molecular Oncology*, vol. 2, no. 2, pp. 115–152, May 2008, doi: <https://doi.org/10.1016/j.molonc.2008.04.001>.
- [23] “Gastric Cancer (Stomach Cancer) - What It Is, Signs/Symptoms, Diagnosis, Treatment - NUHS,” NUHS, 2023. <https://www.nuhs.edu.sg/patient-care/find-a-condition/gastric-cancer-stomach-cancer>
- [24] National Cancer Institute, “Diffuse Midline Gliomas,” National Cancer Institute, Jul. 25, 2019. <https://www.cancer.gov/rare-brain-spine-tumor/tumors/diffuse-midline-gliomas>
- [25] “PET CT Scan For Oncology | MIC Medical Imaging,” MIC Medical Imaging, Dec. 18, 2023. <https://www.mic.ca/for-patients/procedure-information/pet-ct-scan/oncology-pet/>
- [26] “Gastric Cancer (Stomach Cancer) - What It Is, Signs/Symptoms, Diagnosis, Treatment - NUHS,” NUHS, 2023. <https://www.nuhs.edu.sg/patient-care/find-a-condition/gastric-cancer-stomach-cancer>
- [27] I. Mallick, “What Is Histopathology?,” *Verywell Health*, Feb. 21, 2022. <https://www.verywellhealth.com/histopathology-2252152>
- [28] “Τι είναι η μηχανική μάθηση; | Ορισμός, τύποι και παραδείγματα | SAP Insights,” SAP. <https://www.sap.com/greece/products/artificial-intelligence/what-is-machine-learning.html>
- [29] M. M. Taye, “Understanding of Machine Learning with Deep Learning: Architectures, Workflow, Applications and Future Directions,” *Computers*, vol. 12, no. 5, Apr. 2023, doi: <https://doi.org/10.3390/computers12050091>.
- [30] Δημήτρης, “Τι είναι και πως λειτουργεί η βαθιά μάθηση,” *iGuRu.gr*, Sep. 30, 2021. <https://iguru.gr/einai-kai-pos-leitourgei-vathia-mathisi/>
- [31] S. Haykin, *Neural Networks and Learning Machines*. Upper Saddle River, N.J.: Pearson Education, 2009.
- [32] J. Heaton, “Ian Goodfellow, Yoshua Bengio, and Aaron Courville: Deep learning,” *Genetic Programming and Evolvable Machines*, vol. 19, no. 1–2, pp. 305–307, Oct. 2017, doi: <https://doi.org/10.1007/s10710-017-9314-z>.
- [33] Gurney Gevin, *An introduction to neural networks*. London: Ucl Press, Cop, 1997.
- [34] Δ. Κ. Τεχνητά Νευρωνικά Δίκτυα, Κλειδάριθμος, 2017
- [35] W. McCulloch και W. Pitts, «A logical calculus and the ideas immanent in the nervous activity,» *Bulletin of Mathematical Biophysics*, τόμ. 5, pp. 115-133, 1943.
- [36] developNET, “Νευρωνικά Δίκτυα (Neural Networks): Ορισμός & Εφαρμογές,” *Big Blue Data Academy*. <https://bigblue.academy/gr/neuronika-diktua>
- [37] ZEMIM, “Deep Neural Network (DNN) Explained - ZEMIM - Medium,” *Medium*, Apr. 14, 2024. <https://medium.com/@zemim/deep-neural-network-dnn-explained-0f7311a0e869>
- [38] S. Indolia, A. K. Goswami, S. P. Mishra, and P. Asopa, “Conceptual Understanding of Convolutional Neural Network- A Deep Learning Approach,” *Procedia Computer Science*, vol. 132, pp. 679–688, 2018, doi: <https://doi.org/10.1016/j.procs.2018.05.069>.
- [39] L. Alzubaidi et al., “Review of deep learning: concepts, CNN architectures, challenges, applications, future directions,” *Journal of Big Data*, vol. 8, no. 1, Mar. 2021, doi: <https://doi.org/10.1186/s40537-021-00444-8>.

- [40] GeeksforGeeks, “Top PreTrained Models for Image Classification,” GeeksforGeeks, Jul. 03, 2024. <https://www.geeksforgeeks.org/computer-vision/top-pre-trained-models-for-image-classification/> (accessed Jun. 30, 2025).
- [41] G. Boesch, “VGG Very Deep Convolutional Networks (VGGNet) - What you need to know,” viso.ai, Oct. 06, 2021. <https://viso.ai/deep-learning/vgg-very-deep-convolutional-networks/>
- [42] K. Simonyan and A. Zisserman, “VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION,” Apr. 2015. Available: <https://arxiv.org/pdf/1409.1556>
- [43] Lubna Abdelkareim Gabralla et al., “Automated Diagnosis for Colon Cancer Diseases Using Stacking Transformer Models and Explainable Artificial Intelligence,” *Diagnostics*, vol. 13, no. 18, pp. 2939–2939, Sep. 2023, doi: <https://doi.org/10.3390/diagnostics13182939>.
- [44] G. Boesch, “Deep Residual Networks (ResNet, ResNet50) - Guide in 2021,” viso.ai, Aug. 29, 2021. <https://viso.ai/deep-learning/resnet-residual-neural-network/>
- [45] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition,” Dec. 2015. Available: <https://arxiv.org/pdf/1512.03385>
- [46] G. Boesch, “GoogLeNet Explained: The Inception Model that Won ImageNet,” viso.ai, May 07, 2024. <https://viso.ai/deep-learning/googlenet-explained-the-inception-model-that-won-imagenet/>
- [47] “A Review of Popular Deep Learning Architectures: ResNet, InceptionV3, and SqueezeNet | DigitalOcean,” Digitalocean.com, 2015. <https://www.digitalocean.com/community/tutorials/popular-deep-learning-architectures-resnet-inceptionv3-squeezenet>
- [48] C. Szegedy, V. Vanhoucke, S. Ioffe, and J. Shlens, “Rethinking the Inception Architecture for Computer Vision.” Available: <https://arxiv.org/pdf/1512.00567>
- [49] P. Potrimba, “What is EfficientNet? The Ultimate Guide.,” Roboflow Blog, Aug. 09, 2023. <https://blog.roboflow.com/what-is-efficientnet/>
- [50] “Papers with Code - EfficientNet Explained,” paperswithcode.com. <https://paperswithcode.com/method/efficientnet>
- [51] P. Linardatos, V. Papastefanopoulos, and S. Kotsiantis, “Explainable AI: A Review of Machine Learning Interpretability Methods,” *Entropy*, vol. 23, no. 1, p. 18, Dec. 2020, doi: <https://doi.org/10.3390/e23010018>.
- [52] K. Simonyan, A. Vedaldi, and A. Zisserman, “Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps,” arXiv.org, Apr. 19, 2014. <https://arxiv.org/abs/1312.6034v2>
- [53] M. Sundararajan, A. Taly, and Q. Yan, “Axiomatic Attribution for Deep Networks,” arXiv.org, Jun. 12, 2017. <https://arxiv.org/abs/1703.01365v2>
- [54] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, “Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization,” 2017 IEEE International Conference on Computer Vision (ICCV), pp. 618–626, Oct. 2017, doi: <https://doi.org/10.1109/iccv.2017.74>.
- [55] C. Blundell, J. Cornebise, K. Kavukcuoglu, and D. Wierstra, “Weight Uncertainty in Neural Networks,” arXiv.org, May 21, 2015. <https://arxiv.org/abs/1505.05424v2>
- [56] Balaji Lakshminarayanan, A. Pritzel, and C. Blundell, “Simple and Scalable Predictive Uncertainty Estimation using Deep Ensembles,” arXiv (Cornell University), Jan. 2016, doi: <https://doi.org/10.48550/arxiv.1612.01474>.
- [57] Y. Gal and Z. Ghahramani, “Dropout as a Bayesian Approximation: Appendix,” arXiv (Cornell University), Jan. 2015, doi: <https://doi.org/10.48550/arxiv.1506.02157>.

- [58] M. Zubair et al., “An interpretable framework for gastric cancer classification using multi-channel attention mechanisms and transfer learning approach on histopathology images,” *Scientific Reports*, vol. 15, no. 1, Apr. 2025, doi: <https://doi.org/10.1038/s41598-025-97256-0>.
- [59] H. Chen et al., “GasHis-Transformer: A multi-scale visual transformer approach for gastric histopathological image detection,” *Pattern Recognition*, vol. 130, p. 108827, Oct. 2022, doi: <https://doi.org/10.1016/j.patcog.2022.108827>.
- [60] M. Khalid, Sugitha Deivasigamani, Sathiya V, and S. Rajendran, “An efficient colorectal cancer detection network using atrous convolution with coordinate attention transformer and histopathological images,” *Scientific Reports*, vol. 14, no. 1, Aug. 2024, doi: <https://doi.org/10.1038/s41598-024-70117-y>.
- [61] W. Hu et al., “GasHisSDB: A new gastric histopathology image dataset for computer aided diagnosis of gastric cancer,” *Computers in Biology and Medicine*, vol. 142, pp. 105207–105207, Mar. 2022, doi: <https://doi.org/10.1016/j.compbiomed.2021.105207>.
- [62] J. Wei et al., “A Petri Dish for Histopathology Image Analysis,” *arXiv (Cornell University)*, Jan. 2021, doi: <https://doi.org/10.48550/arxiv.2101.12355>.
- [63] H. Amin, A. Darwish, A. E. Hassanien, and M. Soliman, “End-to-End Deep Learning Model for Corn Leaf Disease Classification,” *IEEE Access*, vol. 10, pp. 31103–31115, 2022, doi: <https://doi.org/10.1109/access.2022.3159678>.
- [64] M. Gurucharan, “Basic CNN Architecture: Explaining 5 Layers of Convolutional Neural Network,” *upGrad blog*, Dec. 07, 2020. <https://www.upgrad.com/blog/basic-cnn-architecture/>
- [65] H. Pelletier, “How-To: Cross Validation with Time Series Data | Towards Data Science,” *Towards Data Science*, Dec. 29, 2023. <https://towardsdatascience.com/how-to-cross-validation-with-time-series-data-9802a06272c6/>
- [66] “Navigation — HistomicsTK documentation,” *Github.io*, 2025. <https://digitalslidearchive.github.io/HistomicsTK/#> (accessed Jul. 01, 2025).
- [67] DigitalSlideArchive, “GitHub - DigitalSlideArchive/HistomicsTK: A Python toolkit for pathology image analysis algorithms,” *GitHub*, Feb. 07, 2025. <https://github.com/DigitalSlideArchive/HistomicsTK> (accessed Jul. 01, 2025).
- [68] H. B. Mann and D. R. Whitney, “On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other,” *The Annals of Mathematical Statistics*, vol. 18, no. 1, pp. 50–60, Mar. 1947, doi: <https://doi.org/10.1214/aoms/1177730491>.