



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

# Πολυτροπική μάθηση για την εκτίμηση της επικινδυνότητας των αθηρωματικών πλακών σε ασθενείς με καρωτιδική νόσο

*Μελέτη και υλοποίηση*

---

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

της

**ΕΛΕΝΗΣ-ΕΛΠΙΔΑΣ Κ. ΚΑΨΑΛΗ**

**Επιβλέπουσα:** Κωνσταντίνα Νικήτα  
Καθηγήτρια Ε.Μ.Π.

Αθήνα, Νοέμβριος 2022

---





ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

**Πολυτροπική μάθηση για την εκτίμηση της  
επικινδυνότητας των αθρωματικών πλακών σε  
ασθενείς με καρωτιδική νόσο**

*Μελέτη και υλοποίηση*

---

**ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ**

της

**ΕΛΕΝΗΣ-ΕΛΠΙΔΑΣ Κ. ΚΑΨΑΛΗ**

**Επιβλέπουσα:** Κωνσταντίνα Νικήτα  
Καθηγήτρια Ε.Μ.Π.

Εγκρίθηκε από την τριμελή εξεταστική επιτροπή την 9η Νοεμβρίου 2022.

*(Υπογραφή)*

*(Υπογραφή)*

*(Υπογραφή)*

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

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

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

Αθήνα, Νοέμβριος 2022





ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

(Υπογραφή)

.....  
Ελένη-Ελπίδα Κ. Καψάλη

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

Copyright © Ελένη-Ελπίδα Κ. Καψάλη, 2022.

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

---

Σύμφωνα με τον Παγκόσμιο Οργανισμό Υγείας (2022 Heart Disease Stroke Statistical Update Fact Sheet Global Burden of Disease) τα καρδιαγγειακά νοσήματα αποτελούν τη 1η αιτία θανάτων παγκοσμίως. Η αθηροσκλήρωση αποτελεί καρδιαγγειακή ασθένεια η οποία οφείλεται στην δημιουργία αθηρωματικών πλακών στα τοιχώματα των αρτηριών. Σε περίπτωση ρήξης ή/και αποκόλλησης των πλακών δημιουργούνται θρόμβοι οι οποίοι προκαλούν απόφραξη των αγγείων. Έχει εκτιμηθεί ότι έως και 20-25% των ισχαιμικών εγκεφαλικών προκαλούνται από αθηροσκλήρωση κύριας αρτηρίας και η αθηροσκλήρωση της καρωτιδικής αρτηρίας πιστεύεται ότι ευθύνεται σε ποσοστό μεταξύ 10% και 20%. Με την εξέλιξη των απεικονιστικών τεχνολογιών σε συνδυασμό με ιστολογικές εξετάσεις έχει διευρυνθεί το πλήθος και το είδος των παραμέτρων που χαρακτηρίζουν μία αθηρωματική πλάκα και την πιθανή συμπεριφορά της. Οι παράμετροι αυτοί μαζί με τις κλινικές παρατηρήσεις και τις άλλες εργαστηριακές εξετάσεις αποτελούν ένα μεγάλο σύνολο δεδομένων που με κατάλληλο συνδυασμό, επεξεργασία και γνώση διαμορφώνουν πολύ αξιόπιστα προγνωστικά εργαλεία.

Στόχος της παρούσας εργασίας είναι η δημιουργία πολυτροπικών μοντέλων Βαθιάς Μάθησης (Multimodal Learning) τα οποία θα αξιοποιούν τις εικόνες υπερήχων καρωτίδας και τα κλινικά και εργαστηριακά δεδομένα ασθενών με καρωτιδική νόσο και θα κατατάσσουν τους ασθενείς σε δύο κατηγορίες, Υψηλού και Χαμηλού Κινδύνου. Για το σκοπό αυτό αξιοποιήθηκαν ευρέως χρησιμοποιούμενα μοντέλα βαθιάς μάθησης όπως τα ResNet18, AlexNet και VGG16 και αναπτύχθηκαν δύο πολυτροπικά μοντέλα, ένα με τη μέθοδο Feature-Level Fusion και ένα με τη Decision-Level Fusion. Το σύνολο δεδομένων αποτελείται από 64 ασθενείς Υψηλού Κινδύνου και 11 Χαμηλού και 268 εικόνες υπερηχογραφημάτων. Τελικά, τα μοντέλα που αναπτύχθηκαν αξιολογήθηκαν ως προς αρκετές μετρικές, συγκρίθηκαν μεταξύ τους και εξάχθηκαν συμπεράσματα ως προς την αποδοτικότητά τους. Η μέθοδος Feature-Level Fusion έδωσε τα καλύτερα αποτελέσματα πετυχαίνοντας Balanced Accuracy 77%, AUC 87%, Precision 76%, F1 Score 75% και Matthews Correlation Coefficient (MCC) 0.533. Τα αποτελέσματα κρίνονται ικανοποιητικά λόγω του περιορισμένου πλήθους των δεδομένων και η χρήση πολυτροπικών μοντέλων στην καρωτιδική νόσο φαίνεται να έχει προγνωστικές ικανότητες που μπορούν να βελτιώσουν την αποτίμηση της κλινικής κατάστασης του ασθενούς και να οδηγήσουν σε βέλτιστες θεραπευτικές αποφάσεις.

## Λέξεις Κλειδιά

Καρωτιδική νόσος, Αθηρωματική πλάκα, Σύστημα υποστήριξης αποφάσεων, Βαθιά Μάθηση, Πολυτροπική Μάθηση, εξισορρόπηση δεδομένων, επαύξηση δεδομένων, ResNet18, ResNet50, AlexNet, VGG16, VGG19





# Abstract

---

According to the World Health Organization (2022 Heart Disease Stroke Statistical Update Fact Sheet Global Burden of Disease) cardiovascular diseases are the 1st cause of death worldwide. Atherosclerosis is a cardiovascular disease caused by the formation of atheromatous plaques in the walls of the arteries. In case of rupture and/or detachment of the plaques, clots are created which cause blockage of the vessels. It has been estimated that up to 20-25% of ischemic strokes are caused by main artery atherosclerosis, and carotid artery atherosclerosis is thought to account for between 10% and 20%. With the development of imaging technologies combined with histological examinations, the number and type of parameters that characterize an atherosclerotic plaque and its possible behavior have expanded. These parameters, together with clinical observations and other laboratory tests, form a large set of data that, with appropriate combination, processing and knowledge, form very reliable prognostic tools.

The aim of this work is to create Multimodal Deep Learning models that utilize carotid ultrasound images and clinical and laboratory data of patients with carotid disease and classify the patients into two categories, High and Low Risk. For this purpose, widely used deep learning models such as ResNet18, AlexNet and VGG16 were exploited and two multimodal models were developed, one with the Feature-Level Fusion method and one with the Decision-Level Fusion. The dataset consists of 64 High and 11 Low Risk patients and 268 ultrasound images. Finally, the developed models were evaluated in terms of several metrics, compared with each other and conclusions were drawn regarding their efficiency. The Feature-Level Fusion method had the best results in achieving Balanced Accuracy 77%, AUC 87%, Precision 76%, F1 Score 75% and Matthews Correlation Coefficient (MCC) 0.533. The results are satisfactory due to the limited amount of data and the use of multimodal models in carotid disease appears to have predictive capabilities that can improve the assessment of the patient's clinical condition and lead to optimal treatment decisions.

## Keywords

Carotid disease, Atherosclerotic plaque, Decision Support System, Deep Learning, Multimodal Learning, data balancing, data augmentation, ResNet18, ResNet50, AlexNet, VGG16 VGG19



*στους γονείς μου*



## Ευχαριστίες

---

Θα ήθελα να ευχαριστήσω πολύ την καθηγήτρια κα. Κωνσταντίνα Νικήτα για την ανάθεση αυτού του πολύ ενδιαφέροντος θέματος, την επίβλεψη και την καθοδήγησή της. Ευχαριστώ ιδιαίτερα την Δρα. Καλλιόπη Δαλακλείδη για το χρόνο που αφιέρωσε στην επίβλεψη αυτής της εργασίας και τις εύστοχες συμβουλές και παρατηρήσεις της.

Ένα μεγάλο ευχαριστώ οφείλω στην οικογένειά μου για τη στήριξη και την αγάπη της. Ξεχωριστά θα ήθελα να ευχαριστήσω τους γονείς μου Κυριάκο και Ασημίνα που είναι πάντα εκεί για εμένα σε όλες τις όμορφες και τις δύσκολες στιγμές, το δίδυμο αδερφό μου και συνάδελφο Γιώργο που υπήρξε ο πολυτιμότερος συνοδοιπόρος των σχολικών και φοιτητικών μου χρόνων και τον αδερφό μου Δημοσθένη που με κάνει πάντα να χαμογελώ. Τέλος, ένα τεράστιο ευχαριστώ στους φίλους και συμφοιτητές μου που μοιραστήκαμε αυτό το όμορφο ταξίδι μαζί.

Αθήνα, Νοέμβριος 2022

*Ελένη-Ελπίδα Κ. Καψάλη*



# Περιεχόμενα

---

<b>Περίληψη</b>	<b>1</b>
<b>Abstract</b>	<b>3</b>
<b>Ευχαριστίες</b>	<b>7</b>
<b>1 Εισαγωγή</b>	<b>17</b>
1.1 Εισαγωγικά Στοιχεία . . . . .	17
1.2 Αντικείμενο της διπλωματικής . . . . .	18
1.3 Δομή της εργασίας . . . . .	19
<b>I Θεωρητικό Μέρος</b>	<b>21</b>
<b>2 Καρωτιδική νόσος</b>	<b>23</b>
2.1 Καρδιαγγειακές παθήσεις, καρωτιδικές αρτηρίες και καρωτιδική νόσος . . . . .	23
2.1.1 Οι καρωτιδικές αρτηρίες . . . . .	23
2.1.2 Αθηρωματική πλάκα-Μηχανισμός αθηρογένεσης . . . . .	25
2.1.3 Κυτταρικός μηχανισμός αθηρογένεσης . . . . .	26
2.1.4 Καρωτιδική νόσος . . . . .	27
2.1.5 Συμπτωματικός και ασυμπτωματικός καρωτιδικός ασθενής . . . . .	28
2.1.6 Διάγνωση της καρωτιδικής νόσου . . . . .	28
2.1.7 Αντιμετώπιση της καρωτιδικής νόσου . . . . .	29
2.2 Μέθοδοι απεικόνισης των αθηρωματικών πλακών . . . . .	31
<b>3 Τεχνητή Νοημοσύνη</b>	<b>35</b>
3.1 Εισαγωγή . . . . .	35
3.2 Τεχνητά Νευρωνικά Δίκτυα και Βαθιά Μάθηση . . . . .	36
3.3 Συνελκτικά Νευρωνικά Δίκτυα (CNNs) . . . . .	37
3.4 Συναρτήσεις Ενεργοποίησης . . . . .	38
3.5 Επίπεδα Υποδειγματοληψίας (Pooling Layers) . . . . .	40
3.6 Επίπεδα Batch Normalization . . . . .	40
3.7 Επίπεδα Dropout . . . . .	40
3.8 Τεχνικές μείωσης διαστατικότητας των δεδομένων . . . . .	41
3.8.1 Ανάλυση Κυρίων Συνιστωσών (Principal Components Analysis) . . . . .	41
3.9 Προβλήματα των δεδομένων . . . . .	42
3.9.1 Synthetic Minority Oversampling Technique -SMOTE . . . . .	42

3.9.2	CTGAN	43
3.9.3	Adaptive Synthetic Sampling-ADASYN	47
3.9.4	Bootstrap aggregating-BAGGING	47
3.9.5	Μεταφορά Μάθησης-Transfer Learning	48
<b>4</b>	<b>Πολυτροπική Μάθηση</b>	<b>49</b>
4.1	Εισαγωγή	49
4.2	Κατηγορίες της Πολυτροπικής Μάθησης	49
4.3	Σχετικές Εργασίες	52
4.3.1	Πολυτροπική Μάθηση στη Βιοϊατρική	52
4.3.2	Τεχνητή Νοημοσύνη στην ιατρική απεικόνιση και στα καρδιαγγειακά νοσήματα	58
<b>II</b>	<b>Πρακτικό Μέρος</b>	<b>61</b>
<b>5</b>	<b>Σύνολο Δεδομένων και Υλοποίηση</b>	<b>63</b>
5.1	Στόχος της εργασίας	63
5.2	Παρουσίαση Συνόλου Δεδομένων	63
5.2.1	Το υποσύνολο των εικόνων	64
5.2.2	Το υποσύνολο των κλινικών και βιοχημικών δεδομένων	64
5.3	Υλοποίηση	70
5.3.1	Περιβάλλον ανάπτυξης	70
5.3.2	Μεθοδολογία εκπαίδευσης και αξιολόγησης των μοντέλων	70
5.3.3	Προεπεξεργασία των υπερηχογραφικών εικόνων	72
5.3.4	Προεπεξεργασία των δεδομένων πίνακα	73
5.4	Περιγραφή των μοντέλων	76
5.4.1	Μοντέλα για την ταξινόμηση των ασθενών με βάση τις εικόνες	76
5.4.2	Μοντέλο για την ταξινόμηση των ασθενών με βάση τα κλινικά και εργαστηριακά χαρακτηριστικά	81
5.4.3	Πολυτροπικά Μοντέλα	81
5.4.4	Χειρισμός της ανισορροπίας των δεδομένων	82
5.5	Μετρικές αξιολόγησης των μοντέλων	84
5.6	Μετατροπή δεδομένων πίνακα σε εικόνες	86
<b>6</b>	<b>Αποτελέσματα</b>	<b>89</b>
6.1	Αποτελέσματα unimodal μοντέλων για τις εικόνες μετά από προεκπαίδευση	89
6.2	Αποτελέσματα unimodal μοντέλων για τις εικόνες έπειτα από εξισορρόπηση και επαύξηση	91
6.3	Αποτελέσματα πολυτροπικών μοντέλων	93
6.3.1	Μοντέλο Feature-Level Fusion	93
6.3.2	Μοντέλο Decision-Level Fusion	93
6.4	Σύνοψη αποτελεσμάτων	94



---

<b>7 Συμπεράσματα και Συζήτηση</b>	<b>97</b>
7.1 Συμπεράσματα . . . . .	97
7.2 Μελλοντικές προεκτάσεις . . . . .	98
<b>Παραρτήματα</b>	<b>101</b>
<b>Α' Συμπληρωματικοί πίνακες</b>	<b>103</b>
Α.1 Απεικονιστικές μέθοδοι για την καρωτίδα . . . . .	103
Α.2 Κατηγοριοποίηση συστάσεων της ESVS . . . . .	103
<b>Βιβλιογραφία</b>	<b>115</b>
<b>Συντομογραφίες - Αρκτικόλεξα - Ακρωνύμια</b>	<b>117</b>



## Κατάλογος Σχημάτων

---

2.1	Η δεξιά καρωτιδική αρτηρία η οποία διακλαδίζεται στην έσω και στην έξω αρτηρία [1]. . . . .	24
2.2	Η δομή της αρτηρίας [2] . . . . .	25
2.3	Στάδια δημιουργία και ρήξης της αθηρωματικής πλάκας [3] . . . . .	26
2.4	Σημεία καρωτιδικών αρτηριών όπου εμφανίζονται πλάκες [4] . . . . .	27
2.5	Μέθοδοι υπολογισμού στένωσης έσω καρωτίδας [5, 6] . . . . .	30
2.6	Αλγόριθμος στρατηγικών διαχείρισης ασθενών με συμπτωματική και ασυμπτωματική καρωτιδική νόσο, a=στένωση έσω καρωτίδας στο σημείο σχεδόν απόφραξης, b= ύπαρξη κλινικών/απεικονιστικών χαρακτηριστικών που μπορεί να συσχετίζονται με αυξημένο εγκεφαλικού επεισοδίου σε BMT σε ασυμπτωματικούς ασθενείς,c= κλινικά/απεικονιστικά χαρακτηριστικά που μπορεί να καταστήσουν έναν ασθενή «υψηλού κινδύνου για CEA»,* σύσταση για CAS σε συμπτωματικούς ασθενείς με στενώσεις 70-99% που θεωρούνται «υψηλού κινδύνου για CEA»,** σύσταση για CAS σε συμπτωματικούς ασθενείς με 70-99% στενώσεις που θεωρούνται «μεσαίου κινδύνου για CEA» [7] . . . . .	32
2.7	Κατηγοριοποίηση επιστημονικής βιβλιογραφίας με θέμα τον χαρακτηρισμό της καρωτιδικής νόσου με βάση τη χρησιμοποιούμενη απεικονιστική μέθοδο [8]	33
3.1	Δομή τεχνητού νευρώνα και τεχνητού νευρωνικού δικτύου [9] . . . . .	37
3.2	Παράδειγμα αρχιτεκτονικής DCNN [10] . . . . .	39
3.3	Βασικές συναρτήσεις ενεργοποίησης . . . . .	40
3.4	Τεχνικές Resampling . . . . .	43
3.5	Παραγωγικά Ανταγωνιστικά Δίκτυα (GANs) [11] . . . . .	43
3.6	Παράδειγμα πολυτροπικής κατανομής . . . . .	45
3.7	Mode-specific Normalization . . . . .	45
3.8	Training-by-Sampling . . . . .	46
3.9	Τεχνική Bagging [12] . . . . .	48
4.1	Κατηγορίες της Πολυτροπικής Μάθησης . . . . .	50
4.2	Αρχιτεκτονικές Πολυτροπικής Μάθησης [13] . . . . .	50
4.3	Τρόποι συγχώνευσης των δεδομένων . . . . .	52
4.4	Εξέλιξη των δημοσιεύσεων στη βαθιά μάθηση για την ιατρική απεικόνιση [13]	53
4.5	Προτεινόμενη αρχιτεκτονική στο [14] . . . . .	54
4.6	Προτεινόμενη αρχιτεκτονική στο [15] . . . . .	55
4.7	Προτεινόμενο μοντέλο GCN [16] . . . . .	56

4.8	Προτεινόμενη αρχιτεκτονική με μηχανισμό προσοχής στο [17]	57
5.1	Πλήθος εικόνων ανά ασθενή	65
5.2	Εικόνες αθηρωματικής πλάκας όπως προέκυψαν από τα frames ενός B-Mode Ultrasound βίντεο (αριστερά η εικόνα κατά τη συστολή της αρτηρίας και δεξιά η εικόνα της ίδιας πλάκας κατά τη διαστολή)	65
5.3	Πλήθος ασθενών ανά ηλικιακή ομάδα	66
5.4	Διάγραμμα μέσω τιμών των 11 αιματολογικών χαρακτηριστικών για κάθε κατηγορία ασθενών	69
5.5	3-Fold Cross Validation στο αρχικό σύνολο	71
5.6	Εκπαίδευση και αξιολόγηση με Nested Cross Validation	71
5.7	Μετατροπή μεγέθους εικόνων σε $224 \times 224$	74
5.8	PCA στα συνεχή κλινικά και εργαστηριακά χαρακτηριστικά	75
5.9	Αρχιτεκτονική VGG16 [18]	77
5.10	Αρχιτεκτονική VGG19 [19]	77
5.11	Αρχιτεκτονική Residual Block [20]	78
5.12	Αρχιτεκτονικές ResNet [20]	79
5.13	Αρχιτεκτονική AlexNet [21]	80
5.14	Overlapping Pooling: Στην πάνω εικόνα με $\text{stride}=2$ δεν έχουμε overlapping ενώ στην κάτω εικόνα με $\text{stride}=1$ έχουμε overlapping.	80
5.15	Αρχιτεκτονική μοντέλου Feature-Level Fusion	82
5.16	Αρχιτεκτονική μοντέλου Decision-Level Fusion	83
5.17	Καμπύλη ROC	87
6.1	Confusion Matrices για τα μοντέλα μόνο με προεκπαίδευση	91
6.2	Confusion Matrices για τα μοντέλα μόνο με προεκπαίδευση	92
6.3	Επίδοση ResNet18 με τρεις τρόπους επαύξησης δεδομένων συνόλου εκπαίδευσης	93
6.4	Ενδεικτικά διαγράμματα ROC με Feature-Level Fusion	95
6.5	Ενδεικτικά διαγράμματα ROC με Feature-Level Fusion	95
A.1	Πλεονεκτήματα και μειονεκτήματα απεικονιστικών μεθόδων καρωτίδας	104
A.2	ESVS: Κατηγοριοποίηση των συστάσεων	105
A.3	Συστάσεις των American Heart Association και American Stroke Association σε σχέση με τη ενδαρτηρεκτομή για την αντιμετώπιση της συμπτωματικής στένωσης της καρωτίδας.	105

## Κατάλογος Πινάκων

---

5.1	Ρυθμίσεις υπερηχογραφικής συσκευής. . . . .	65
5.2	One-hot κωδικοποίηση του διακριτού χαρακτηριστικού Smoker . . . . .	74
5.3	Κλινικά και εργαστηριακά χαρακτηριστικά . . . . .	75
6.1	Επίπεδα που εκπαιδεύονται κατά τη μεταφορά μάθησης . . . . .	89
6.2	Επίδοση unimodal μοντέλων ResNet18, ResNet50, AlexNet για εικόνες μόνο με προεκπαίδευση . . . . .	90
6.3	Επίδοση unimodal μοντέλων VGG19, VGG16 για τις εικόνες μόνο με προεκπαίδευση . . . . .	90
6.4	Επίδοση unimodal μοντέλων ResNet18, ResNet50, AlexNet για τις εικόνες μετά από μεταφορά μάθησης και εξισορρόπηση train set με επαύξηση . . . . .	91
6.5	Επίδοση unimodal μοντέλων VGG19, VGG16 για τις εικόνες μετά από μεταφορά μάθησης και εξισορρόπηση train set με επαύξηση . . . . .	92
6.6	Επίδοση Feature Fusion μοντέλων ResNet18, ResNet50, AlexNet . . . . .	93
6.7	Επίδοση Feature Fusion μοντέλων με VGG19, VGG16 . . . . .	94
6.8	Επίδοση Decision Fusion μοντέλων ResNet18, VGG16, AlexNet . . . . .	94



# Κεφάλαιο **1**

## Εισαγωγή

---

### 1.1 Εισαγωγικά Στοιχεία

Σύμφωνα με τον Παγκόσμιο Οργανισμό Υγείας (2022 Heart Disease Stroke Statistical Update Fact Sheet Global Burden of Disease) [22] τα καρδιαγγειακά νοσήματα αποτελούν τη 1η αιτία θανάτων παγκοσμίως. Το 2020, περίπου 19,1 εκατομμύρια θάνατοι αποδόθηκαν σε καρδιαγγειακή νόσο. Το ηλικιακά προσαρμοσμένο ποσοστό θνησιμότητας ανά 100.000 πληθυσμού ήταν 239,8. Το ποσοστό επιπολασμού προσαρμοσμένο με την ηλικία ήταν 7.354,1 ανά 100.000.

Με βάση τα στοιχεία του 2019, υπολογίζεται ότι το 32% όλων των θανάτων παγκοσμίως οφείλεται σε καρδιαγγειακά νοσήματα. Το ποσοστό αυτό ανεβαίνει στο 38% για τις ηλικίες μικρότερες των 70 ετών. Από αυτούς τους θανάτους, το 85% οφείλονταν σε έμφραγμα του μυοκαρδίου και εγκεφαλικό.

Σχεδόν 60.000 αγγειακά εγκεφαλικά επεισόδια σημειώνονται κάθε χρόνο στην Ελλάδα και η αναλογία εγκεφαλικών στη χώρα μας ανά 100.000 πληθυσμού κυμαίνεται γύρω στα 500 ετησίως, ενώ στην Ευρώπη ο αριθμός είναι περίπου 350-400 [23].

Οι καρδιακές προσβολές και τα εγκεφαλικά είναι συνήθως οξεία συμβάντα και προκαλούνται κυρίως από απόφραξη που εμποδίζει τη ροή του αίματος προς την καρδιά ή τον εγκέφαλο. Η πιο συνηθισμένη αιτία για αυτό είναι η αθηροσκλήρωση, δηλαδή η συσσώρευση λιπιδίων στα εσωτερικά τοιχώματα των αιμοφόρων αγγείων που τροφοδοτούν την καρδιά ή τον εγκέφαλο. Τα εγκεφαλικά μπορεί να προκληθούν από αιμορραγία σε κάποιο αιμοφόρο αγγείο στον εγκέφαλο ή από θρόμβους αίματος. [24].

Για τη διάγνωση των αγγειακών νόσων και την πρόγνωση της εξέλιξης και των πιθανών επιπτώσεων χρησιμοποιούνται α) κλινικές παρατηρήσεις, β) εργαστηριακές κυρίως αιματολογικές εξετάσεις και γ) απεικονιστικές μέθοδοι.

Οι σημαντικότεροι συμπεριφορικοί παράγοντες κινδύνου για καρδιακές παθήσεις, εγκεφαλικό και αθηροσκλήρωση είναι η ανθυγιεινή διατροφή, η έλλειψη σωματικής δραστηριότητας, το κάπνισμα και το αλκοόλ. Οι συνήθειες αυτές καταλήγουν σε παθολογικά ευρήματα αυξημένη αρτηριακή πίεση, αυξημένη γλυκόζη αίματος, αυξημένα λιπίδια αίματος, αυξημένο δείκτη μάζας σώματος και παχυσαρκία. Οι περισσότερες καρδιαγγειακές παθήσεις μπορούν να προληφθούν με την αντιμετώπιση των συμπεριφορικών παραγόντων κινδύνου. Η διακοπή του καπνίσματος, η μείωση του αλατιού στη διατροφή, η κατανάλωση περισσότερων φρούτων και λαχανικών, η τακτική σωματική δραστηριότητα και η αποφυγή του αλκοόλ

έχουν αποδειχθεί ότι μειώνουν τον κίνδυνο καρδιαγγειακών παθήσεων. Είναι σημαντικό να εντοπιστεί η καρδιαγγειακή νόσος όσο το δυνατόν νωρίτερα, ώστε να μπορεί να ξεκινήσει η διαχείριση με συμβουλευτική και φάρμακα [24].

Επίσης έχει εκτιμηθεί ότι το είδος και η συμπεριφορά της αθηρωματικής πλάκας στην καρωτίδα αποτελεί ένα πολύ σημαντικό δείκτη για την εξέλιξη της αθηροσκλήρωσης γενικότερα στα αγγεία ενός οργανισμού και κατά βάση στα στεφανιαία και κατά συνέπεια έχει σημαντική προγνωστική αξία για καρδιαγγειακές ασθένειες [25, 7, 26, 7]. Όμοια είναι τα ευρήματα των μελετών που συγκρίνουν τα αποτελέσματα της αγγειογραφίας των στεφανιαίων μ' αυτά της υπερηχογραφικής αξιολόγησης των καρωτίδων [27]. Η αθηρωμάτωση των καρωτίδων σχετίζεται όχι μόνο με την διαγνωσμένη αλλά επίσης και με την τυχαία στεφανιαία νόσο [28].

Θεμελιώδη συμβολή στη διάγνωση και πρόγνωση των αγγειακών νοσημάτων έχουν διαδραματίσει οι απεικονιστικές μέθοδοι. Από τις τρεις κύριες απεικονιστικές μεθόδους, πιο συγκεκριμένα, μαγνητική τομογραφία (MPI), υπολογιστική (αξονική) τομογραφία (CT) και υπερηχογραφία (US), ο υπερήχος B-mode (BUS) προσφέρει πολλά πλεονεκτήματα, συγκεκριμένα, χαμηλό κόστος, φιλικότητα προς τον χρήστη και ικανότητα απόκτησης υψηλής ανάλυσης εικόνων (μέσω των σύνθεσης χαρακτηριστικών από αρμονικές και ελεγχόμενο κέρδος). Έτσι, το B-US προτιμάται ως διαγνωστικό εργαλείο στις περισσότερες των περιπτώσεων [29, 30, 31].

Με την εξέλιξη των απεικονιστικών τεχνολογιών σε συνδυασμό την διεύρυνση των ιστολογικών εξετάσεων έχει διευρυνθεί το πλήθος και το είδος των παραμέτρων που χαρακτηρίζουν μία αθηρωματική πλάκα και την πιθανή συμπεριφορά της. Οι παράμετροι αυτοί περιλαμβάνουν γεωμετρικά χαρακτηριστικά (πάχος, επιφάνεια, όγκος), σύσταση, υφή, νεοαγγειώσεις στο εσωτερικό, φλεγμονές, αιμοδυναμικές φορτίσεις, κινητική συμπεριφορά της πλάκας κ.α. [32]. Οι παράμετροι αυτοί μαζί με τις κλινικές παρατηρήσεις και τις άλλες εργαστηριακές εξετάσεις αποτελούν ένα μεγάλο σύνολο δεδομένων που με κατάλληλο συνδυασμό, επεξεργασία και γνώση διαμορφώνουν πολύ αξιόπιστα προγνωστικά εργαλεία.

Λαμβάνοντας υπόψη ότι τα δεδομένα αυτά προέρχονται από διάφορες πηγές, έχουν ποικίλες μορφές όπως εικόνες, βίντεο, ήχοι, αριθμητικά, κείμενο κλπ. και τη διάδοση των ψηφιακών τεχνολογιών και των εφαρμογών Τεχνητής Νοημοσύνη, μία από τις σύγχρονες προκλήσεις είναι η συνδυαστική, ολοκληρωμένη και αυτοματοποιημένη διαχείρισή τους με σκοπό την ανάπτυξη και λειτουργία ιδιαίτερα αξιόπιστων μοντέλων/εργαλείων εκτίμησης των σχετικών κινδύνων για έγκαιρη πρόληψη και αντιμετώπιση.

Γίνεται σημαντική πρόοδος στον τομέα αυτό και η παρούσα εργασία φιλοδοξεί να αποτελέσει μία απειροελάχιστη συμβολή σε αυτήν την προσπάθεια.

## 1.2 Αντικείμενο της διπλωματικής

Στόχος της παρούσας διπλωματικής εργασίας είναι να προταθεί ένα σύστημα υποστήριξης αποφάσεων το οποίο θα υποστηρίζει τις κλινικές αποφάσεις των γιατρών για τους ασθενείς με καρωτιδική νόσο. Η εργασία αυτή εστιάζει στην κατηγοριοποίηση των ασθενών με καρωτιδική νόσο σε δύο κατηγορίες. Στην πρώτη κατηγορία κατατάσσονται οι ασθενείς υψηλού κινδύνου στους οποίους προτείνεται η χειρουργική αντιμετώπιση της νόσου και στη δεύτερη



κατηγορία οι ασθενείς χαμηλού κινδύνου για τους οποίους προτείνεται η συντηρητική αντιμετώπιση. Μια επιτυχημένη κατηγοριοποίηση των ασθενών από το σύστημα θα προσφέρει στους ειδικούς ένα ακόμη εργαλείο και θα συμβάλει στο να αποφεύγονται περιττές χειρουργικές επεμβάσεις. Το σύστημα θα βασίζεται σε μεθόδους τεχνητής νοημοσύνης και πιο συγκεκριμένα σε Πολυτροπική Μηχανική Μάθηση (Multimodal Machine Learning). Τα δύο είδη δεδομένων (modalities) τα οποία χρησιμοποιούνται στην παρούσα εργασία είναι κλινικά δεδομένα που προέρχονται από βιοχημικές εξετάσεις και εικόνες οι οποίες προέρχονται από βίντεο υπερήχων β-σάρωσης καρωτίδων καρωτιδικών ασθενών.

### **1.3 Δομή της εργασίας**

Η εργασία είναι δομημένη σε δύο κύρια μέρη: στο Θεωρητικό Μέρος και στο Πρακτικό Μέρος. Στο Θεωρητικό Μέρος, το οποίο διαιρείται σε επιμέρους κεφάλαια, παρουσιάζονται όλες οι απαραίτητες έννοιες για την κατανόηση του θέματος οι οποίες αφορούν τόσο στο κομμάτι της ιατρικής όσο και στο κομμάτι της Μηχανικής Μάθησης. Πιο συγκεκριμένα, στο Θεωρητικό Μέρος παρουσιάζεται το βασικό ιατρικό υπόβαθρο και τα χαρακτηριστικά της καρωτιδικής νόσου, ο αλγόριθμος διάκρισης των ασθενών στις δύο προαναφερθέντες κατηγορίες, οι τρόποι διάγνωσης και αντιμετώπισης της νόσου αλλά και έννοιες της Μηχανικής Μάθησης. Τέλος, στο Πρακτικό Μέρος παρουσιάζεται το υπό μελέτη σύνολο δεδομένων και οι ιδιαιτερότητές του, περιγράφονται τα μοντέλα και οι τεχνικές που χρησιμοποιήθηκαν και παρουσιάζονται και σχολιάζονται τα πειραματικά αποτελέσματα. Τέλος, γίνεται μια σύνοψη του συνόλου της εργασίας όπου παρουσιάζονται τα βασικά συμπεράσματα που προέκυψαν από τη μελέτη και προτείνονται μελλοντικές προεκτάσεις της.



## **Μέρος I**

### **Θεωρητικό Μέρος**

---



## Κεφάλαιο **2**

# Καρωτιδική νόσος

---

### 2.1 Καρδιαγγειακές παθήσεις, καρωτιδικές αρτηρίες και καρωτιδική νόσος

Σύμφωνα με τον Παγκόσμιο Οργανισμό Υγείας (2022 Heart Disease & Stroke Statistical Update Fact Sheet Global Burden of Disease) [33] τα καρδιαγγειακά νοσήματα αποτελούν τη 1η αιτία θανάτων παγκοσμίως. Το 2020, περίπου 19,1 εκατομμύρια θάνατοι αποδόθηκαν σε καρδιαγγειακή νόσο. Το ηλικιακά προσαρμοσμένο ποσοστό θνησιμότητας ανά 100.000 πληθυσμού ήταν 239,8. Το ποσοστό επιπολασμού προσαρμοσμένο με την ηλικία ήταν 7.354,1 ανά 100.000. Με βάση τα στοιχεία του 2019, υπολογίζεται ότι το 32% όλων των θανάτων παγκοσμίως οφείλεται σε καρδιαγγειακά νοσήματα. Το ποσοστό αυτό ανεβαίνει στο 38% για τις ηλικίες κάτω των 70 ετών. Από αυτούς τους θανάτους, το 85% οφείλονταν σε έμφραγμα του μυοκαρδίου και εγκεφαλικό.[34] Η καρωτιδική νόσος, η οποία αποτελεί και το θέμα της παρούσας εργασίας, αποτελεί μία αρκετά συχνή καρδιαγγειακή πάθηση η οποία οφείλεται στην δημιουργία αθηρωματικών πλακών στα τοιχώματα των καρωτιδικών αρτηριών. Όπως θα δούμε και στη συνέχεια στην περίπτωση μιας ασταθούς πλάκας, το ινώδες περίβλημα πολλές φορές σπάει προκαλώντας μιας μορφής θρόμβωση. Η θρόμβωση προκαλεί εμβολή της αρτηρίας, η οποία, με τη σειρά της, οδηγεί σε απόφραξη της ροής του αίματος και τελικά σε εγκεφαλικό ή σε έμφραγμα του μυοκαρδίου.

#### 2.1.1 Οι καρωτιδικές αρτηρίες

Οι καρωτίδες είναι αιμοφόρα αγγεία που ανήκουν στην κατηγορία των αρτηριών. Αρτηρίες ονομάζονται τα αγγεία του ανθρώπινου οργανισμού που μεταφέρουν οξυγονωμένο αίμα από την καρδιά προς τα υπόλοιπα μέρη του σώματος. Η κύρια αρτηρία του ανθρώπινου σώματος είναι η αορτή η οποία λαμβάνει απευθείας αίμα από την αριστερή κοιλία της καρδιάς μέσω της αορτικής βαλβίδας. Καθώς απομακρυνόμαστε από την καρδιά η αορτή διακλαδίζεται σε μικρότερες αρτηρίες οι οποίες διακλαδίζονται κι αυτές με τη σειρά τους σχηματίζοντας όλο και μικρότερα αγγεία και τελικά καταλήγουν στα τριχοειδή αγγεία. Στα τριχοειδή αγγεία τα αρτηρίδια, οι μικρότερες αρτηρίες, συναντώνται με τα φλεβίδια, τις μικρότερες φλέβες, και πραγματοποιείται η ανταλλαγή των αερίων της αναπνοής. Οι πρώτοι κλάδοι της αορτής είναι οι στεφανιαίες αρτηρίες, οι οποίες τροφοδοτούν με αίμα τον καρδιακό μυ και ακολουθούν οι κλάδοι του αορτικού τόξου. Οι κλάδοι του αορτικού τόξου είναι:

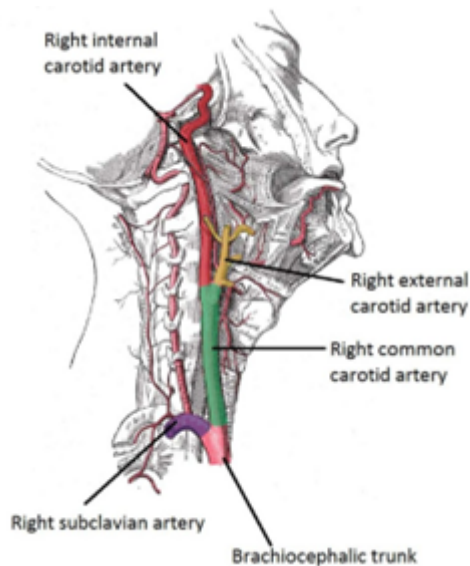
η βραχιοκεφαλική αρτηρία, η αριστερή κοινή καρωτίδα και η αριστερή υποκλείδια αρτηρία. Η βραχιοκεφαλική αρτηρία με τη σειρά της διακλαδίζεται σε δύο κλάδους στην δεξιά κοινή καρωτίδα και στη δεξιά υποκλείδια αρτηρία.

Οι καρωτιδικές αρτηρίες ανήκουν στις μεγαλύτερες και σημαντικότερες αρτηρίες του ανθρώπινου σώματος και είναι υπεύθυνες για την κύρια αιμάτωση της κεφαλής. Είναι τα κύρια αιμοφόρα αγγεία στο λαιμό που παρέχουν αίμα στον εγκέφαλο, το λαιμό και το πρόσωπο. Όπως είδαμε και προηγουμένως υπάρχουν δύο καρωτίδες, η δεξιά και η αριστερή.

Η αριστερή και δεξιά κοινή καρωτιδική αρτηρία ανεβαίνουν στον αυχένα, πλάγια προς την τραχεία και τον οισοφάγο. Στο επίπεδο του άνω περιθωρίου του θυρεοειδούς χόνδρου (C4), οι καρωτιδικές αρτηρίες χωρίζονται σε δύο τμήματα (Εικόνα 2.1):

- Στην έσω καρωτίδα, η οποία τροφοδοτεί με αίμα τον εγκέφαλο.
- Στην έξω καρωτίδα, η οποία τροφοδοτεί με αίμα το πρόσωπο και το λαιμό.

Αυτή η διχοτόμηση εμφανίζεται σε μια ανατομική περιοχή γνωστή ως καρωτιδικό τρίγωνο. Στο κύριο σημείο της διακλάδωσης η κοινή καρωτίδα και η εσωτερική καρωτίδα είναι ελαφρώς διασταλμένες. Η περιοχή αυτή ονομάζεται καρωτιδικός κόλπος ή καρωτιδικός βολβός και περιέχει αισθητήρες που βοηθούν στη ρύθμιση της αρτηριακής πίεσης. Ο παλμός της καρωτιδικής αρτηρίας μπορεί κανονικά να γίνει αισθητός στον λαιμό πιέζοντας τα άκρα των δακτύλων στο πλάι της τραχείας.

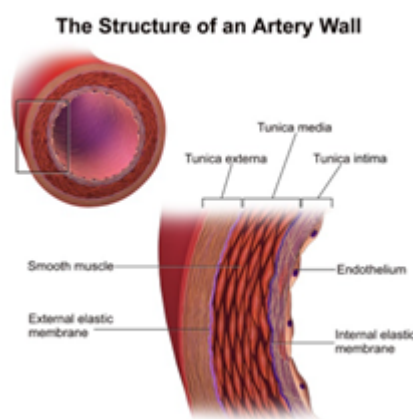


Σχήμα 2.1: Η δεξιά καρωτιδική αρτηρία η οποία διακλαδίζεται στην έσω και στην έξω αρτηρία [1].

Όπως όλες οι αρτηρίες έτσι και οι καρωτιδικές αποτελούνται από τρία στρώματα ιστού (Σχήμα 2.2):

- Τον **έσω χιτώνα (Intima)**: το λείο εσωτερικό στρώμα το οποίο αποτελείται από μια στιβάδα ενδοθηλιακών κυττάρων, λείους μύες και ελασίνη. Περιβάλλει τον αυλό του αγγείου μέσα στον οποίο κυκλοφορεί το πλούσιο σε οξυγόνο αίμα.

- Τον **μέσο χιτώνα (Media)**: το μυϊκό μεσαίο στρώμα το οποίο είναι συνήθως το παχύτερο στρώμα. Παρέχει υποστήριξη στο αγγείο και αλλάζει τη διάμετρό του για να ρυθμίσει τη ροή του αίματος και την αρτηριακή πίεση. Αποτελείται από περισσότερο λείο μυ, σε σχέση με τον έσω χιτώνα, που μπορεί να διαστέλλεται ή να συστέλλεται γεγονός που προσαρμόζει τις πιέσεις που γίνονται αισθητές στα αρτηριακά τοιχώματα κατά τη συστολική άντληση. Καθώς οι μυς συστέλλονται, τα τοιχώματα θα αισθάνονται μεγαλύτερη πίεση από την αριστερή κοιλία και ομοίως, καθώς διαστέλλονται τα αγγεία, η παρατηρούμενη πίεση θα μειωθεί.
- Τον **έξω χιτώνα (Adventitia)**: το εξωτερικό στρώμα το οποίο συνδέει το αγγείο με τον περιβάλλοντα ιστό, συμπεριλαμβανομένων των αγγειακών νεύρων τα οποία ελέγχουν τους λείους μύες στις αρτηρίες. Αποτελείται από συνδετικό ιστό με ποικίλες ποσότητες ελαστικών και κολλαγόνων ινών. Ο συνδετικός ιστός σε αυτό το στρώμα είναι αρκετά πυκνός στο μέσο του χιτώνα αλλά μετατρέπεται σε χαλαρό συνδετικό ιστό κοντά στην περιφέρεια του αγγείου.



Σχήμα 2.2: Η δομή της αρτηρίας [2]

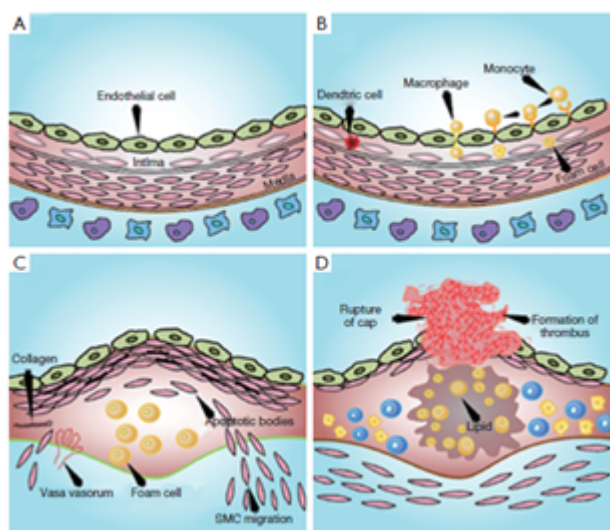
### 2.1.2 Αθηρωματική πλάκα-Μηχανισμός αθηρογένεσης

Η δημιουργία των αθηρωματικών πλακών, η αθηρογένεση, χαρακτηρίζεται από τη συσσώρευση λιπιδίων και ινών στοιχείων και την ασβεστοποίηση εντός των μεγάλων αρτηριών. Η διαδικασία του σχηματισμού των πλακών ξεκινά με την ενεργοποίηση του ενδοθηλίου, ακολουθούμενη από μία σειρά γεγονότων και την ενεργοποίηση διαφόρων κυτταρικών μηχανισμών που έχουν ως αποτέλεσμα τη στένωση του αγγείου και την πυροδότηση φλεγμονωδών αντιδράσεων. Όπως είδαμε και προηγουμένως ο έσω χιτώνας των αρτηριών αποτελείται από ένα ενιαίο στρώμα ενδοθηλιακών κυττάρων, το ενδοθήλιο, το οποίο μαζί με το κολλαγόνο και τις ελαστικές ίνες σχηματίζουν τον αυλό στον οποίο ρέει το αίμα. Επομένως το ενδοθήλιο του έσω χιτώνα αποτελεί το πρώτο φράγμα για μόρια και κύτταρα που επιχειρούν να εισέλθουν στην ή να εξέλθουν από την κυκλοφορία του αίματος καθώς βρίσκεται μεταξύ του κυκλοφορούντος αίματος και των ιστών. Επιπλέον το ενδοθήλιο λειτουργεί ως αισθητήρας και μετατροπέας σημάτων μέσω της παραγωγής βιολογικά δραστικών ουσιών. Αλλαγές όπως

μηχανικές δυνάμεις μεταξύ του αίματος και των τοιχωμάτων του αγγείου (Wall Shear Stress (WSS)) και αλλαγές στην συγκέντρωση ουσιών ανιχνεύονται από το ενδοθήλιο, το οποίο στη συνέχεια λειτουργεί ως μεταγωγέας των σημάτων στα άλλα στρώματα του αγγειακού τοιχώματος [35].

### 2.1.3 Κυτταρικός μηχανισμός αθηρογένεσης

Έχει παρατηρηθεί ότι οι βλάβες των αρτηριών εντοπίζονται κυρίως σε περιοχές που χαρακτηρίζονται από χαμηλό WSS και διαχωρισμό της ροής του αίματος όπως είναι τα σημεία διακλάδωσής τους. Στα σημεία αυτά δημιουργούνται ανωμαλίες στη δομή του ενδοθηλίου με αποτέλεσμα να διεισδύει στον έσω χιτώνα χοληστερόλη η οποία περιέχει λιποπρωτεΐνες. Οι πρωτεογλυκάνες στον χιτώνα συμβάλλουν ώστε η λιποπρωτεΐνη χαμηλής πυκνότητας (LDL) να παγιδευτεί στο αγγειακό τοίχωμα όπου και οξειδώνεται. Η οξείδωση της κατακρατούμενης LDL με τη σειρά της πυροδοτεί φλεγμονώδεις αντιδράσεις του ενδοθηλίου. Επομένως στο ενδοθήλιο δημιουργούνται μόρια προσκόλλησης (adhesive molecules) για τα λευκοκύτταρα. Τα μονοκύτταρα και τα βοηθητικά Τ-λεμφοκύτταρα (κατηγορίες λευκοκυττάρων) προσδένονται στα μόρια προσκόλλησης του ενδοθηλίου κι έπειτα διεισδύουν στον έσω χιτώνα του αγγείου. Τότε τα μονοκύτταρα διαφοροποιούνται (ή εξειδικεύονται) σε μακροφάγα τα οποία δεσμεύουν την οξειδωμένη LDL δημιουργώντας τα foam cells. Η δημιουργία των foam cells αποτελεί κομβικό σημείο για την δημιουργία των αθηρωματικών πλάκων καθώς αυτά προκαλούν τη μετανάστευση λείων μυϊκών κυττάρων (SMC) του μέσου χιτώνα στον έσω χιτώνα της αρτηρίας και τον πολλαπλασιασμό τους. Η αθηρωματική πλάκα δημιουργείται από τα λιπίδια κι άλλες ουσίες που ελευθερώνονται στον έσω χιτώνα από την καταστροφή των foam cells ενώ ο πολλαπλασιασμός των SMC έχει ως αποτέλεσμα την αύξηση της έκκρισης κολλαγόνου το οποίο σκληραίνει τα τοιχώματα των αρτηριών [36, 37, 35]. Στο Σχήμα 2.3 φαίνονται τα βασικά στάδια της δημιουργία των αθηρωματικών πλάκων.

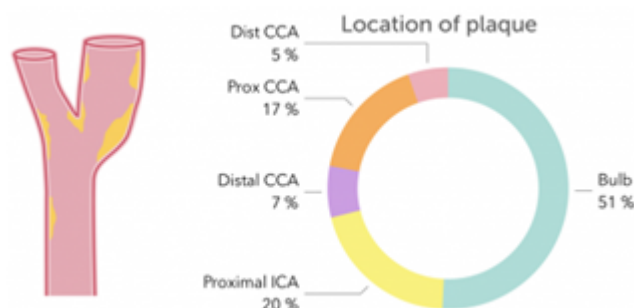


Σχήμα 2.3: Στάδια δημιουργία και ρήξης της αθηρωματικής πλάκας [3]



### 2.1.4 Καρωτιδική νόσος

Η καρωτιδική νόσος είναι η στένωση των καρωτιδικών αρτηριών. Η στένωση οφείλεται στις αθηρωματικές πλάκες οι οποίες μπορούν να διαρραγούν προκαλώντας θρομβώσεις οι οποίες μπορεί να οδηγήσουν ακόμη και στο θάνατο. Οι θρομβώσεις αυτές μπορεί να διακόψουν την παροχή αίματος σε σημεία του εγκεφάλου προκαλώντας ισχαιμικά εγκεφαλικά επεισόδια. Σημαντικό να σημειώσουμε ότι τα εγκεφαλικά επεισόδια διακρίνονται σε ισχαιμικά και αιμορραγικά αγγειακά εγκεφαλικά επεισόδια (ΑΕΕ). Τα ισχαιμικά, τα οποία μας απασχολούν και στην παρούσα εργασία, προκαλούνται από θρόμβους σε αιμοφόρα αγγεία που τροφοδοτούν τον εγκέφαλο, ενώ τα αιμορραγικά οφείλονται σε ρήξη αρτηριών στον εγκέφαλο. Το εγκεφαλικό είναι η τρίτη κύρια αιτία θνησιμότητας στις ΗΠΑ ενώ περισσότερα από 15 εκατομμύρια άνθρωποι παθαίνουν εγκεφαλικά κάθε χρόνο παγκοσμίως. Οι άνδρες τείνουν να εμφανίζουν συχνότερα εγκεφαλικά σε μικρότερη ηλικία (<75 ετών) σε σχέση με τις γυναίκες, αλλά συνολικά οι γυναίκες εμφανίζουν τα περισσότερα. Επιπλέον, αποτελεί καθοριστικής σημασίας για τον ασθενή εάν μια στένωση ή μια πλάκα εντοπίζεται στην έσω ή στην έξω καρωτίδα καθώς, όπως αναφέραμε, η έσω καρωτίδα τροφοδοτεί τον εγκέφαλο ενώ η έξω καρωτίδα τις εξωκράνιες δομές της κεφαλής και του λαιμού. Επομένως ισχαιμικό επεισόδιο θα συμβεί μόνο εάν εμπλέκεται η έσω καρωτίδα. Μια στένωση της έξω καρωτιδικής αρτηρίας συνήθως έχει ελάχιστες συνέπειες [10]. Στο Σχήμα 2.4 παρουσιάζονται τα σημεία των καρωτιδικών αρτηριών όπου εμφανίζονται συνήθως αθηρωματικές πλάκες. Ο καρωτιδικός βολβός (bulb) είναι το κύριο σημείο όπου εμφανίζονται αθηρώματα.



Σχήμα 2.4: Σημεία καρωτιδικών αρτηριών όπου εμφανίζονται πλάκες [4]

Όταν μια πλάκα δημιουργήσει θρόμβωση η οποία οδηγήσει σε απόφραξη αρτηριών ο ασθενής εμφανίζει συμπτώματα. Τα συμπτώματα της νόσου οφείλονται, όπως έχουμε δει, στα εγκεφαλικά επεισόδια τα οποία μπορεί να είναι είτε παροδικά (ΠΠΕ) είτε σοβαρά αγγειακά εγκεφαλικά επεισόδια και τα κυριότερα από αυτά είναι [38]:

- **Παροδική απώλεια της όρασης από το ένα μάτι**, διάρκειας λίγων δευτερολέπτων έως και λίγων λεπτών. Συνήθως ο ασθενής αναφέρει κάτι σαν θάμπωμα, σαν να βλέπει πίσω από μια κουρτίνα.
- **Παροδική απώλεια της δύναμης του ενός χεριού ή/και του ομόπλευρου ποδιού**, διάρκειας λίγων λεπτών έως λίγων ωρών ή και για περισσότερο από 24 ώρες. Ο ασθενής μπορεί να αναφέρει ότι, το χέρι του δεν «υπακούει» στις εντολές του.

- **Δυσκολία στην ομιλία**
- **Ζάλη, αστάθεια, ίλιγγος** είναι μη ειδικά συμπτώματα που μπορεί να συνυπάρχουν.
- **Μόνιμες βλάβες σε περιοχές του εγκεφάλου**, οφείλονται στα σοβαρά αγγειακά εγκεφαλικά επεισόδια.

### 2.1.5 Συμπτωματικός και ασυμπτωματικός καρωτιδικός ασθενής

Τουλάχιστον 15-20% όλων των ισχαιμικών εγκεφαλικών επεισοδίων αποδίδονται στην καρωτική νόσο και υπολογίζεται ότι το 5-10% των ατόμων άνω των 65 ετών πάσχουν από ασυμπτωματική αθηροσκλήρωση της καρωτίδας με τουλάχιστον 50% στένωση αρτηρίας. Συνεπώς η διάκριση της καρωτιδικής νόσου σε συμπτωματική και ασυμπτωματική αποτελεί καθοριστικό παράγοντα στην επιλογή της κατάλληλης θεραπείας [39]. Σύμφωνα με την Ευρωπαϊκή Ένωση για την Αγγειοχειρουργική (European Society for Vascular Surgery), οι ασθενείς θεωρούνται συμπτωματικοί εάν έχουν εμφανίσει συμπτώματα της καρωτιδικής νόσου τους τελευταίους 6 μήνες, δηλαδή έχουν πάθει παροδικό ισχαιμικό επεισόδιο (ΠΙΕ) ή εγκεφαλικό τους τελευταίους 6 μήνες [7]. Επιπλέον, ευάλωτες (vulnerable) χαρακτηρίζονται οι αθηρωματικές πλάκες οι οποίες είναι επιρρεπείς σε ρήξη. Οι ευάλωτες πλάκες, οι οποίες χαρακτηρίζονται επίσης ως ασταθείς (unstable) και συμπτωματικές (symptomatic), διαθέτουν πλούσιο λιπώδη πυρήνα (lipid core), λεπτό ινώδες περίβλημα και χαμηλές ποσότητες κολλαγόνου. Αντίθετα οι πλάκες που χαρακτηρίζονται από φτωχό λιπώδη πυρήνα, παχύ ινώδες περίβλημα και μεγάλη ποσότητα κολλαγόνου τείνουν να μην σπάνε και να μην προκαλούν συμπτώματα γι'αυτό χαρακτηρίζονται ως σταθερές (stable) και ασυμπτωματικές (asymptomatic) [40].

### 2.1.6 Διάγνωση της καρωτιδικής νόσου

Η διάγνωση της καρωτιδικής νόσου και της αθηροσκλήρωσης γενικότερα πραγματοποιείται με συνδυασμό των παρακάτω μεθόδων:

- **Εξετάσεις αίματος:** Συνήθως γίνονται εξετάσεις αίματος για τον έλεγχο των επιπέδων σακχάρου και χοληστερόλης στο αίμα. Τα υψηλά επίπεδα σακχάρου και χοληστερόλης στο αίμα αυξάνουν τον κίνδυνο αθηροσκλήρωσης. Επίσης κάποιες φορές πραγματοποιείται και C-reactive protein (CRP) τεστ ώστε να εντοπιστεί μια πρωτεΐνη η οποία συνδέεται με τη φλεγμονή των αρτηριών.
- **Ηλεκτροκαρδιογράφημα (ΗΚΓ):** Με το ΗΚΓ μπορεί να προσδιοριστεί εάν υπάρχει μειωμένη ροή αίματος στην καρδιά. Κατά την εξέταση Αισθητήρες (ηλεκτρόδια) συνδέονται στο στήθος και μερικές φορές στα χέρια ή τα πόδια. Τα καλώδια συνδέουν τους αισθητήρες σε ένα μηχάνημα, το οποίο εμφανίζει ή εκτυπώνει τα αποτελέσματα.
- **Τεστ κοπώσεως:** Κατά την εξέταση παρακολουθείται ο καρδιακός ρυθμός και η αρτηριακή πίεση του εξεταζόμενου ενώ ασκείται σε διάδρομο ή στατικό ποδήλατο. Επειδή η άσκηση κάνει την καρδιά να αντλεί το αίμα με διαφορετικό ρυθμό από ό,τι στις περισσότερες καθημερινές δραστηριότητες το τεστ κοπώσεως μπορεί να εντοπίσει καρδιακά προβλήματα που ίσως διαφορετικά να αγνοούνταν.

- **Ηχοκαρδιογράφημα:** Το ηχοκαρδιογράφημα δείχνει τη ροή του αίματος μέσα από την καρδιά χρησιμοποιώντας ηχητικά κύματα. Μερικές φορές, η εξέταση πραγματοποιείται ταυτόχρονα με το τεστ κοπώσεως.
- **Έγχρωμο υπερηχογράφημα καρωτίδων και σπονδυλικών αρτηριών (triplex):** Χρησιμοποιεί ηχητικά κύματα για να δημιουργήσει μια εικόνα της αρτηρίας η οποία δείχνει εάν υπάρχει απόφραξη.
- **Ankle-brachial index (ABI):** Σε αυτή την εξέταση συγκρίνεται η αρτηριακή πίεση στον αστράγαλο με αυτή στο χέρι. Μια διαφορά μεταξύ των μετρήσεων του αστραγάλου και του βραχίονα μπορεί να οφείλεται σε περιφερική αγγειακή νόσο, η οποία συνήθως προκαλείται από αθηροσκλήρωση.
- **Αγγειογράφημα καρδιάς:** είναι ένας τύπος ακτινογραφίας θώρακα που λαμβάνεται μετά την έγχυση ραδιενεργού χρώματος στις αρτηρίες.

### 2.1.7 Αντιμετώπιση της καρωτιδικής νόσου

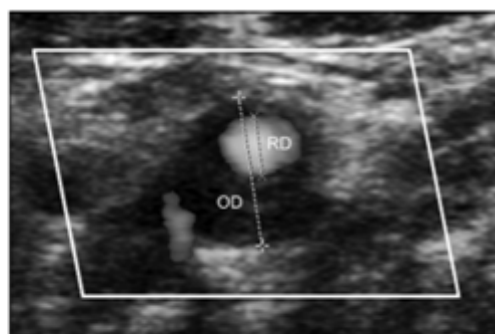
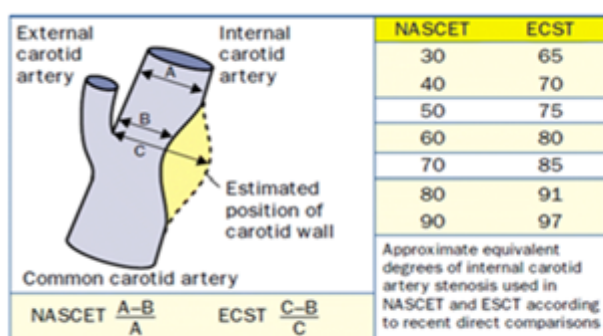
Η θεραπεία-αντιμετώπιση της καρωτιδικής νόσου μπορεί να είναι είτε αμιγώς συντηρητική, η οποία περιλαμβάνει κατάλληλη φαρμακευτική αγωγή ή/και αλλαγές στον τρόπο ζωής, είτε συντηρητική σε συνδυασμό με κάποια επέμβαση. Επίσης μεταξύ συμπτωματικών και ασυμπτωματικών ασθενών υπάρχει διαφοροποίηση στην αντιμετώπιση της νόσου.

- **Συντηρητική αντιμετώπιση:** Περιλαμβάνει φαρμακευτική αγωγή για μείωση της υπέρτασης με στόχο ρύθμιση της αρτηριακής πίεσης σε τιμές  $\leq 140/90$  mmHg. Επίσης περιλαμβάνει φαρμακευτική αγωγή, συνήθως με στατίνες, για μείωση της LDL-χοληστερόλης σε τιμές με φάρμακα όπως οι στατίνες την προσθήκη παράγοντες μείωσης των λιπιδίων όπως οι στατίνες με στόχο τη μείωση της τιμής της LDL-χοληστερόλης σε  $\leq 110$  mg/dl, μείωση των τριγλυκεριδίων σε λιγότερο από 150 mg/dl και αύξηση των επιπέδων HDL-χοληστερόλης σε περισσότερο από 40 mg/dl. Υπάρχουν καλές ενδείξεις ότι η θεραπεία με στατίνες μειώνει τον κίνδυνο εγκεφαλικού κατά περίπου 30%. Παράλληλα συστήνονται η συστηματική άσκηση, η διακοπή του καπνίσματος, η απώλεια βάρους και η υγιεινή διατροφή.
- **Επεμβατική αντιμετώπιση:** Περιλαμβάνει είτε το stenting καρωτίδας (carotid artery stenting (CAS)) είτε την ενδοαρτηριοτομή καρωτίδας (Carotid endarterectomy (CEA)), μεθόδους τις οποίες θα περιγράψουμε στη συνέχεια. Σύμφωνα με την εγκεκριμένη από τον FDA έρευνα Carotid Revascularization Endarterectomy Versus Stenting Trial (2010) επεμβατική αντιμετώπιση επιλέγεται στις περιπτώσεις που ο βαθμός στένωσης είναι μεγαλύτερος από 50% για τους συμπτωματικούς ασθενείς και μεγαλύτερος από 70% για τους ασυμπτωματικούς, με προτίμηση στη CEA έναντι της CAS καθώς ο κίνδυνος εγκεφαλικού επεισοδίου, εμφράγματος του μυοκαρδίου ή θάνατος είναι έναντι 6,8% στην περίπτωση της CEA και 7,2% στην περίπτωση της μεθόδου CAS.

Όπως θα δούμε και βασικός παράγοντας για την αξιολόγηση των ασθενών με καρωτιδικών νόσου αποτελεί ο βαθμός στένωσης. Για τη μέτρηση του βαθμού στένωσης υπάρχουν διαφορετικές μέθοδοι [7, 5].

- Η μέθοδος **European Carotid Surgery Trial (ECST)** υπολογίζει τη στένωση ως το πηλίκο της ελάχιστης υπολειπόμενης διαμέτρου του αυλού προς την εκτιμώμενη διάμετρο του αγγείου (συνήθως ο βολβός της καρωτίδας).
- Η μέθοδος **North American Symptomatic Carotid Endarterectomy Trial (NASCET)** υπολογίζει τη στένωση ως το πηλίκο της ελάχιστης υπολειπόμενης διαμέτρου του αυλού προς την διάμετρο ενός υγειούς τμήματος της έσω καρωτίδας πάνω από τη στένωση, όπου τα τοιχώματα των αγγείων είναι περίπου παράλληλα. Μια στένωση 50% κατά NASCET ισοδυναμεί με στένωση 75% κατά ECST. Μια στένωση 70% κατά NASCET ισοδυναμεί με στένωση 85% κατά ECST. Οι δύο μέθοδοι φαίνονται και στο Σχήμα 2.5.
- Η μέτρηση του βαθμού στένωσης σε ένα **B-mode υπερηχογράφημα ή σε ένα έγχρωμο Doppler** γίνεται στο τμήμα της έσω καρωτίδας (ICA) με τη μεγαλύτερη στένωση. Αρχικά σε μια διαμήκη σάρωση μετράται η μέγιστη συστολική ταχύτητα (PSV) και έπειτα σε μια εγκάρσια σάρωση του τμήματος αυτού υπολογίζονται η αρχική διάμετρος (OD) και η υπολειπόμενη διάμετρος (RD) της αρτηρίας. Η RD ορίζεται ως η μικρότερη διάμετρος του υπολειπόμενου αυλού στο σημείο της ICA με τη μεγαλύτερη στένωση. Το OD ορίστηκε ως η μετρηθείσα διάμετρος από το εξωτερικό μέσο στο εξωτερικό μέσο της πάσχουσας αρτηρίας στο ίδιο επίπεδο και στην ίδια κατεύθυνση με το RD. Ο βαθμός στένωσης της καρωτίδας (CS%) υπολογίζεται από την παρακάτω σχέση (βλ. και Σχήμα 2.5).

$$CS\% = \left(1 - \frac{RD}{OD}\right) \times 100\% \quad (2.1)$$



(a)

(b)

Σχήμα 2.5: Μέθοδοι υπολογισμού στένωσης έσω καρωτίδας [5, 6]

Οι σύγχρονες μέθοδοι αφαίρεσης της πλάκας είναι δύο:

- **Καρωτιδική ενδαρτηρεκτομή:** Αποτελεί την πιο συνηθισμένη μέθοδο για τη θεραπεία της σοβαρής καρωτιδικής νόσου. Ο χειρουργός κάνει μια τομή στο μπροστινό μέρος του λαιμού, ανοίγει την προσβεβλημένη καρωτίδα και αφαιρεί τις πλάκες. Τέλος, κλείνει την τομή είτε με ράμματα είτε με μόσχευμα. Στον Πίνακα του Σχήματος Α'3 στο Παράρτημα παρουσιάζονται οι συστάσεις των American Heart Association και American Stroke Association σε σχέση με το πότε πρέπει ένας ασθενής να υποβληθεί σε ενδαρτηρεκτομή και πότε να την αποφύγει ακολουθώντας συντηρητική αγωγή [41].

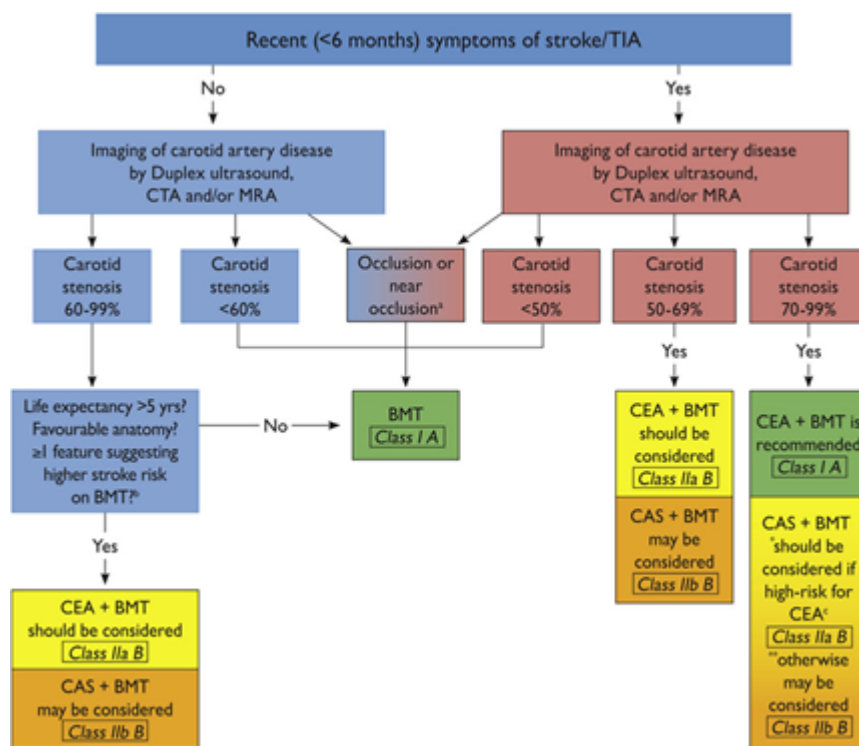
- **Αγγειοπλαστική καρωτίδας και stenting** Η μέθοδος αυτή επιλέγεται όταν η απόφραξη της καρωτίδας είναι πολύ δύσκολο να επιτευχθεί με την προηγούμενη μέθοδο ή εάν η γενική κατάσταση της υγείας του ασθενούς καθιστά τη χειρουργική επέμβαση πολύ επικίνδυνη. Σε αυτή τη μέθοδο χορηγείται τοπική αναισθησία και ένας μακρύς, κοίλος σωλήνας (καθετήρας) περνά μέσα από τις αρτηρίες στη φραγμένη καρωτίδα στον λαιμό. Στη συνέχεια τοποθετείται ένα φίλτρο για να συλλάβει τυχόν υπολείμματα που μπορεί να σπάσουν κατά τη διαδικασία και ένα μικροσκοπικό μπαλόνι στο τέλος του καθετήρα διογκώνεται για να διευρύνει την αρτηρία. Τέλος, ένας μεταλλικός διχτυωτός σωλήνας (stent) εισάγεται στο αγγείο και χρησιμεύσει ως ικρίωμα που βοηθά στην αποτροπή της εκ νέου στένωσης της αρτηρίας [41].

Η Ευρωπαϊκή Ένωση για την Αγγειοχειρουργική (ESVS) δημοσιεύει οδηγίες (guidelines) με συστάσεις για το είδος της θεραπείας που πρέπει να ακολουθηθεί ανάλογα με τα χαρακτηριστικά των ασθενών [7]. Στην τελευταία σχετική δημοσιευμένη οδηγία το 2018 προτείνει τον αλγόριθμο που παρουσιάζεται στο Σχήμα 2.6 και περιγράφει λεπτομερείς στρατηγικές διαχείρισης για ασθενείς με συμπτωματική και ασυμπτωματική αθηροσκλήρωση. Το χρώμα των πλαισίων αναφέρεται στα επίπεδα των συστάσεων της ESVS όπως αυτά παρουσιάζονται στην εικόνα Α.2 του Παραρτήματος. Ο αλγόριθμος βασίζεται στα ευρήματα απεικονιστικών μεθόδων, στον βαθμό στένωσης και στην ηλικία του ασθενούς. Σημαντικό να σημειώσουμε ότι στον αλγόριθμο αυτό ο βαθμός στένωσης αφορά στη μέθοδο NASCET. Οι συντομογραφίες που χρησιμοποιούνται στο διάγραμμα επεξηγούνται στο Γλωσσάρι. Τα γενικά σημεία του αλγορίθμου μπορούν να συνοψιστούν στα εξής:

- Ασθενείς με μεγάλη στένωση (>60% για ασυμπτωματικούς και >50% για συμπτωματικούς) αντιμετωπίζονται επεμβατικά με την ενδαρτηρεκτομή να υπερισχύει συνήθως έναντι του stenting.
- Ασυμπτωματικοί ασθενείς μεγάλης ηλικίας (η ηλικία τους είναι τουλάχιστον 5 χρόνια μικρότερη από το προσδόκιμο ζωής) δεν υποβάλλονται σε επεμβάσεις αλλά αντιμετωπίζονται χειρουργικά.
- Ασθενείς, είτε συμπτωματικοί είτε ασυμπτωματικοί, με πλήρη έμφραξη αρτηρίας αντιμετωπίζονται φαρμακευτικά καθώς έχουν αυξημένο κίνδυνο επιπλοκών κατά την επέμβαση.

## 2.2 Μέθοδοι απεικόνισης των αθηρωματικών πλακών

Η αθηρωματική πλάκα μπορεί να ανιχνευθεί χρησιμοποιώντας απεικονιστικές μεθόδους όπως ο μαγνητικός συντονισμός (MR), η υπολογιστική τομογραφία (CT) και το υπερηχογράφημα (US). Παρά την καθοριστική συμβολή των απεικονιστικών μεθόδων στη διάγνωση και παρακολούθηση της καρωτιδικής νόσου ο προσδιορισμός της φύσης των πλακών μπορεί να αποτελεί μεγάλη πρόκληση με κάθε είδος απεικονιστικής μεθόδου να παρουσιάζει τόσο πλεονεκτήματα όσο και μειονεκτήματα. Ενδεικτικά κάποιες από τις προκλήσεις των προαναφερθέντων απεικονιστικών μεθόδων είναι το partial volume effect στην απεικόνιση μαγνητικού συντονισμού (MRI), οι μεταβαλλόμενες τιμές Hausdorff στις περιοχές της πλάκας στην



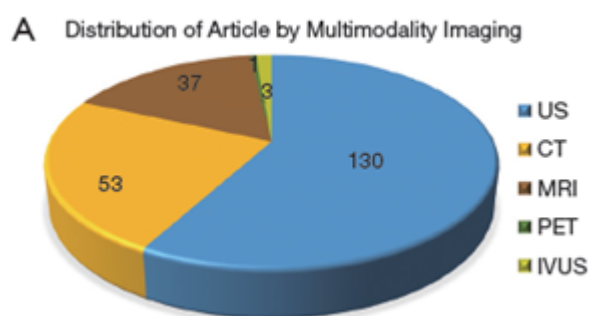
Σχήμα 2.6: Αλγόριθμος στρατηγικών διαχείρισης ασθενών με συμπτωματική και ασυμπτωματική καρωτιδική νόσο, α=στένωση έσω καρωτίδας στο σημείο σχεδόν απόφραξης, b= ύπαρξη κλινικών/απεικονιστικών χαρακτηριστικών που μπορεί να συσχετίζονται με αυξημένο εγκεφαλικού επεισοδίου σε BMT σε ασυμπτωματικούς ασθενείς, c= κλινικά/απεικονιστικά χαρακτηριστικά που μπορεί να καταστήσουν έναν ασθενή «υψηλού κινδύνου για CEA»,\* σύσταση για CAS σε συμπτωματικούς ασθενείς με στενώσεις 70-99% που θεωρούνται «υψηλού κινδύνου για CEA»,\*\* σύσταση για CAS σε συμπτωματικούς ασθενείς με 70-99% στενώσεις που θεωρούνται «μεσαίου κινδύνου για CEA» [7]

CT και η εξασθένηση των κυμάτων που ανακλώνται από την πλάκα κατά το υπερηχογράφημα με αποτέλεσμα το φαινόμενο της ακουστικής σκίασης (acoustic shadowing) [8].

Η διείσδυση της τεχνητής νοημοσύνης στην ανάλυση και στην επεξεργασία της ιατρικής εικόνας κατάφερε να προσφέρει μεγαλύτερη ακρίβεια στον χαρακτηρισμό και στην ταξινόμηση των ιστών για την απεικόνιση της πλάκας της καρωτίδας και στις τρεις μεθόδους απεικόνισης. Στο άρθρο "Multimodality carotid plaque tissue characterization and classification in the artificial intelligence paradigm: a narrative review for stroke application, Luca Saba et al." [8] παρουσιάζονται τα αποτελέσματα έρευνας στην επιστημονική βιβλιογραφία για άρθρα που χρησιμοποιούν απεικονιστικές μεθόδους για τον χαρακτηρισμό και την ταξινόμηση των καρωτιδικών πλακών. Προέκυψε ότι το υπερηχογράφημα αποτελεί την πιο ευρέως χρησιμοποιούμενη μέθοδο για αυτό τον σκοπό ενώ ακολουθεί η αξονική (υπολογιστική) τομογραφία. (Σχήμα 2.7).

### Ultrasound (US) για την απεικόνιση των καρωτιδικών πλακών

Το υπερηχογράφημα καρωτίδας (Carotid Ultrasound ή Carotid Duplex Sonography) είναι μια από τις πιο σημαντικές μεθόδους απεικόνισης υπερήχων στην ιατρική. Παρέχει εικόνες υψηλής ανάλυσης που μας επιτρέπουν να μελετήσουμε το πάχος του έσω χιτώνα, να ανιχνεύσουμε τις αθηρωματικές πλάκες και να αξιολογήσουμε το βαθμό στένωσης των



Σχήμα 2.7: Κατηγοριοποίηση επιστημονικής βιβλιογραφίας με θέμα τον χαρακτηρισμό της καρωτιδικής νόσου με βάση τη χρησιμοποιούμενη απεικονιστική μέθοδο [8]

αρτηριών.

Οι διαγνωστικές συσκευές υπερήχων διαθέτουν δύο κύρια μέρη: έναν σαρωτή διπλής σάρωσης και έναν μετατροπέα (transducer) με γραμμικό εύρος ζώνης πλάτους 7–12 MHz και ανάλυση 20 pixels/mm συνήθως καθώς τα χαρακτηριστικά κάθε συσκευής διαφέρουν. Οι υπέρηχοι λειτουργούν με βάση την αρχή του πιεζοηλεκτρικού φαινομένου. Ο μετατροπέας περιέχει πιεζοηλεκτρικούς κρυστάλλους, οι οποίοι λειτουργούν ως αποστολείς και δέκτες σημάτων υπερήχων [4]. Οι πιεζοηλεκτρικοί κρύσταλλοι παράγουν ηχητικά κύματα όταν μέσα τους ρέει ρεύμα και παράγουν τάση όταν δέχονται υπερηχητικό κύμα. Όταν τα κύματα υπερήχων προσπίπτουν στον ιστό, ο ιστός τα αντανακλά (ηχώ) και ο μετατροπέας συλλέγει τα ανακλώμενα σήματα. Στη συνέχεια δημιουργεί και εμφανίζει μια εικόνα στην οθόνη με βάση αυτά τα σήματα. Ανάλογα με τον τρόπο με τον οποίο τα ανακλώμενα κύματα απεικονίζονται διακρίνονται διάφορες μέθοδοι υπερηχογραφήματος. Στη μέθοδο A-mode (Amplitude Mode) οι ανακλάσεις απεικονίζονται στην οθόνη ως συνάρτηση του βάθους. Στη B-mode (Brightness Mode), η οποία είναι και πιο διαδεδομένη, τα ανακλώμενα σήματα απεικονίζονται με κουκίδες ανάλογης φωτεινότητας στην αντίστοιχη γεωμετρική τους θέση.

Με το υπερηχογράφημα καρωτίδας μπορούν να διακριθούν οι «ευάλωτες» (vulnerable) και δυνητικά συμπτωματικές πλάκες από τις σταθερές και ασυμπτωματικές. Οι μεν πρώτες επειδή είναι ηχογενείς απεικονίζονται με φωτεινά σημεία, οι δε δεύτερες επειδή είναι υποηχογενείς απεικονίζονται με σκούρα σημεία [42]. Επίσης, κατά τη διάρκεια του υπερηχογραφήματος γίνεται και εκτίμηση του πάχους του έσω χιτώνα (IMT) το οποίο χρησιμοποιείται για την αξιολόγηση του κινδύνου ρήξης των πλακών. Ο έσω χιτώνας είναι τυπικά φωτεινός (ηχογενής), ενώ τα περιβάλλοντα μέσα είναι πιο σκούρα (υποηχογενή) [37].

Εστίασαμε στη μέθοδο των υπερήχων καθώς θα μας απασχολήσει και στη συνέχεια της εργασίας. Οι άλλες μέθοδοι παρουσιάζονται συνοπτικά στον Πίνακα ;; του Παραρτήματος όπου παρατίθενται τα πλεονεκτήματα και τα μειονεκτήματά τους καθώς και οι βασικές τους εφαρμογές.





# Τεχνητή Νοημοσύνη

---

## 3.1 Εισαγωγή

Η Τεχνητή Νοημοσύνη είναι η προσομοίωση διαδικασιών ανθρώπινης νοημοσύνης από μηχανές. Ως όρος αφορά, επίσης, στον κλάδο της πληροφορικής ο οποίος ασχολείται με τη μελέτη και ανάπτυξη συστημάτων με ικανότητες όπως η προσαρμοστικότητα, η μάθηση από εμπειρίες του παρελθόντος, η γενίκευση και η εξαγωγή συμπερασμάτων [43]. Η Μηχανική Μάθηση αποτελεί το πεδίο της Τεχνητής Νοημοσύνης το οποίο ασχολείται με την ανάπτυξη και εφαρμογή μεθόδων που αξιοποιούν δεδομένα για τη βελτίωση της απόκρισής ενός υπολογιστικού συστήματος σε ορισμένες εργασίες [44]. Με άλλα λόγια, αντί να κωδικοποιήσουμε τη γνώση μέσα σε μια μηχανή, η Μηχανική Μάθηση ασχολείται με τα εργαλεία και τις μεθόδους με τα οποία διδάσκουμε μια μηχανή να μαθαίνει από την εμπειρία, για παράδειγμα μέσω της αναγνώρισης διάφορων προτύπων [43]. Ο όρος χρησιμοποιήθηκε για πρώτη φορά από τον Άρθουρ Σάμιουελ το 1959, όταν εργαζόταν στον τομέα των computer games και της τεχνητής νοημοσύνης στην IBM [44].

Η Μηχανική Μάθηση μπορεί να διαιρεθεί σε τέσσερις (4) μεγάλες κατηγορίες [45, 46]:

- **Εποπτευόμενη Μάθηση (Supervised Learning):** Στην εποπτευόμενη μάθηση η μηχανή μαθαίνει από παραδείγματα. Τα δεδομένα εκπαίδευσης πρέπει να συνοδεύονται πάντα από την ετικέτα τους. Κατά την εκπαίδευση ο αλγόριθμος κάνει παρατηρήσεις και προσδιορίζει μοτίβα στα δεδομένα ώστε να κάνει προβλέψεις. Γνωρίζοντας τις επιθυμητές τιμές εξόδου (ετικέτες των δεδομένων) αξιολογεί τις προβλέψεις του και μαθαίνει από τα λάθη του έως ότου επιτύχει υψηλή απόδοση. Σε αυτήν την κατηγορία ανήκουν οι αλγόριθμοι Ταξινόμησης (Classification), Παλινδρόμησης (Regression) και Πρόβλεψης (Forecasting). Ενώ η Ταξινόμηση στοχεύει στο να διακρίνει το σύστημα δεδομένα που ανήκουν σε διαφορετικές κατηγορίες, η Παλινδρόμηση στοχεύει στο να εκτιμήσει και να μελετήσει τις σχέσεις μεταξύ μεταβλητών. Τέλος, στην κατηγορία της Πρόβλεψης πραγματοποιούνται προβλέψεις για το μέλλον χρησιμοποιώντας τα υπάρχοντα δεδομένα. Κάποια γνωστοί αλγόριθμοι εποπτευόμενης μάθησης είναι οι Naive Bayes Classifier, Support Vector Machine, Linear Regression, Decision Trees, Random Forests και K-Nearest Neighbors.
- **Μη-εποπτευόμενη Μάθηση (Unsupervised Learning):** Στη μη-εποπτευόμενη μάθηση τα δεδομένα εκπαίδευσης είναι μη χαρακτηρισμένα, δηλαδή δεν παρέχονται οι ε-

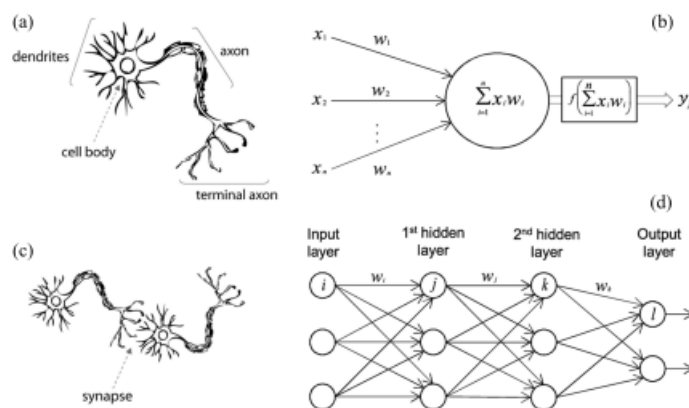
τικέτες τους. Ο αλγόριθμος μάθησης αναλύει και ερμηνεύει τα δεδομένα με στόχο να αναγνωρίσει μοτίβα, να τα οργανώσει και να κατανοήσει τη δομή τους. Σε αντίθεση με την εποπτευόμενη μάθηση δεν υπάρχει κανένα σφάλμα ή ανταμοιβή για να ενημερώσει τον αλγόριθμο εάν είναι κοντά ή μακριά από τη σωστή απάντηση αλλά κατά την εκπαίδευση το μοντέλο βελτιώνει η ικανότητά του να λαμβάνει αποφάσεις σχετικά με τα συγκεκριμένα δεδομένα. Σε αυτή την κατηγορία μάθησης ανήκουν τα προβλήματα Ομαδοποίησης (Clustering) και Μείωσης Διαστατικότητας (Dimension Reduction). Η Ομαδοποίηση όπως λέει και η λέξη αφορά στην εξαγωγή ομάδων από ένα μεγάλο σύνολο δεδομένων με βάση καθορισμένα κριτήρια. Οι αλγόριθμοι Μείωσης Διαστάσεων μετασχηματίζουν τα δεδομένα από έναν χώρο σε έναν χώρο μικρότερης διάστασης, με τρόπο ώστε και στον χώρο μικρότερης διάστασης να διατηρούνται κάποιες σημαντικές ιδιότητες των αρχικών δεδομένων. Οι αλγόριθμοι Μείωσης Διαστάσεων εφαρμόζονται συνήθως στα δεδομένα πριν χρησιμοποιηθούν ως είσοδοι σε αλγορίθμους ταξινόμησης καθώς όσο μεγαλύτερος είναι ο αριθμός των χαρακτηριστικών τόσο δυσκολότερη είναι και η διαδικασία της ταξινόμησης. Γνωστοί αλγόριθμοι μη-εποπτευόμενης μάθησης είναι οι K-means clustering και Principal Components Analysis.

- **Ενισχυτική Μάθηση (Reinforcement Learning):** Σε ένα σύστημα ενισχυτικής μάθησης περιγράφουμε το περιβάλλον του συστήματος, την τρέχουσα κατάστασή του και καθορίζουμε την κατάσταση ή τις καταστάσεις στόχου και μια λίστα επιτρεπόμενων ενεργειών μεταξύ των καταστάσεων. Το σύστημα μαθαίνει να επιλέγει τις βέλτιστες ενέργειες που το οδηγούν στο στόχο μέσα από τη δοκιμή και το σφάλμα (trial and error principle) με στόχο τη μεγιστοποίηση της ανταμοιβής και την ελαχιστοποίηση της ποινής. Βασικοί αλγόριθμοι ενισχυτικής μάθησης είναι οι Value Function, Monte Carlo Methods, Temporal Difference και Q-Learning.
- **Ημι-εποπτευόμενη Μάθηση (Semi-supervised Learning):** Στην ημι-εποπτευόμενη μάθηση μόνο ένα μικρό μέρος των δεδομένων εκπαίδευσης συνοδεύεται από ετικέτα και η εκπαίδευση αποτελεί έναν συνδυασμό εποπτευόμενης και μη-εποπτευόμενης μάθησης.

## 3.2 Τεχνητά Νευρωνικά Δίκτυα και Βαθιά Μάθηση

Στους αλγορίθμους και στις μεθόδους της Μηχανικής Μάθησης ανήκουν διαφορετικές κατηγορίες αλγορίθμων όπως τα Μοντέλα Παλινδρόμησης, τα Δέντρα Αποφάσεων και τα Τεχνητά Νευρωνικά Δίκτυα (ΤΝΔ). Τα ΤΝΔ είναι εμπνευσμένα από τους βιολογικούς νευρώνες και αποτελούνται από μονάδες επεξεργασίας μεταξύ των οποίων διαδίδονται σήματα, όπως συμβαίνει και με τις συνάψεις των βιολογικών. Στο Σχήμα 3.1 φαίνεται η δομή ενός τεχνητού νευρώνα και ένα μικρό νευρωνικό δίκτυο δύο επιπέδων. Ο νευρώνας αποτελεί τη δομική μονάδα των νευρωνικών δικτύων και η λειτουργία του περιλαμβάνει τα επόμενα στάδια: Στα σήματα που φτάνουν στην είσοδο κάθε νευρώνα από τα προηγούμενα επίπεδα εφαρμόζονται **βάρη** τα οποία ενισχύουν ή εξασθενούν τα σήματα. Έπειτα όλα τα σήματα στην είσοδο του νευρώνα προστίθενται μέσω μιας **γραμμικής συνάρτησης** και το συνολικό σήμα αποτελεί

την έξοδο του νευρώνα. Στην έξοδο του νευρώνα εφαρμόζεται μία **συνάρτηση ενεργοποίησης** και η έξοδος της συνάρτησης ενεργοποίησης προωθείται στα επόμενα επίπεδα του δικτύου. Ο στόχος της συνάρτησης ενεργοποίησης είναι να μηδενίζει, πρακτικά να μην προωθή στο επόμενο επίπεδο του δικτύου, σήματα τα οποία υπολείπονται ενός κατωφλίου το οποίο καθορίζεται από την εκάστοτε συνάρτηση ενεργοποίησης.



Σχήμα 3.1: Δομή τεχνητού νευρώνα και τεχνητού νευρωνικού δικτύου [9]

Επομένως, πρακτικά, ένας τεχνητός νευρώνας είναι μια μαθηματική συνάρτηση. Τα νευρωνικά δίκτυα αποτελούνται από πολλούς νευρώνες οι οποίοι οργανώνονται σε επίπεδα (layers). Το πρώτο επίπεδο λαμβάνει τα δεδομένα εισόδου και ονομάζεται επίπεδο εισόδου (input layer), το τελευταίο επίπεδο ονομάζεται επίπεδο εξόδου (output layer) και παράγει το τελικό αποτέλεσμα και τα ενδιάμεσα επίπεδα ονομάζονται κρυφά επίπεδα (hidden layers). Ένα τεχνητό νευρωνικό δίκτυο με περισσότερα από ένα κρυφά επίπεδα ονομάζεται Βαθιά Νευρωνικό Δίκτυο. Οι μέθοδοι μηχανικής μάθησης που βασίζονται σε Βαθιά Νευρωνικά Δίκτυα συνθέτουν την Βαθιά Μάθηση (Deep Learning).

### 3.3 Συνελκτικά Νευρωνικά Δίκτυα (CNNs)

Η ταχεία εξέλιξη στον τομέα της Όρασης Υπολογιστών και των εφαρμογών με εικόνες και βίντεο οφείλεται σε σημαντικό βαθμό στα CNNs καθώς απαιτούν σημαντικά μικρότερο υπολογιστικό κόστος από τα Πλήρως Συνδεδεμένα Δίκτυα (Fully Connected Networks-FCNs). Στα FCNs κάθε νευρώνας συνδέεται με όλους τους νευρώνες του προηγούμενου και του επόμενου επιπέδου εφόσον υπάρχουν. Αυτό δε συμβαίνει στα CNNs με αποτέλεσμα να απαιτούν μικρότερο αριθμό παραμέτρων και επιπλέον επαναχρησιμοποιούν τα βάρη τους.

Οι ψηφιακές εικόνες αποθηκεύονται στους ηλεκτρονικούς υπολογιστές ως πίνακες αριθμών, δηλαδή ως πίνακες από pixels. Οι αριθμοί αντιστοιχούν στη φωτεινότητα (brightness) των pixels και συνήθως είναι στην κλίμακα 0 (άσπρο) έως και 255 (μαύρο). Οι 256 αυτές τιμές είναι αρκετές ώστε να απεικονίζουν τις αποχρώσεις στις οποίες είναι ευαίσθητο το ανθρώπινο μάτι και επιπλέον χωρούν σε 1 byte γεγονός το οποίο διευκολύνει την αποθήκευσή τους. Για την αποθήκευση των ασπρόμαυρων εικόνων ένας τέτοιος πίνακας αρκεί ενώ για τις έγχρωμες RGB εικόνες απαιτούνται τρεις τέτοιοι πίνακες, ένας για κάθε ένα από τα κανάλια Red, Green και Blue. Στο Red κανάλι, για παράδειγμα, η τιμή 0 αντιστοιχεί στην

απουσία κόκκινου χρώματος από το pixel, η τιμή 255 στην παρουσία και οι ενδιάμεσες τιμές αντιστοιχούν σε ενδιάμεσες αποχρώσεις.

Όπως λέει και το όνομά τους τα CNNs βασίζονται στην πράξη της συνέλιξης. Η συνέλιξη μεταξύ δύο πινάκων  $f, g$  ορίζεται ως εξής:

$$(f * g)[x, y] = f[x, y] * g[x, y] = \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} f[n_1, n_2]g[x - n_1, y - n_2]$$

Τα CNN εξάγουν χαρακτηριστικά από τις εικόνες με κατάλληλες διαδοχικές συνέλιξεις μεταξύ των πινάκων εισόδου και άλλων μικρότερων πινάκων οι οποίοι ονομάζονται πυρήνες ή φίλτρα (kernels ή filters). Οι παράμετροι των πυρήνων μαθαίνονται κατά την εκπαίδευση του δικτύου. Τα αποτελέσματα των συνέλιξεων ονομάζονται feature maps και μπορούν να χρησιμοποιηθούν σε διάφορες εφαρμογές όπως για παράδειγμα στην Όραση Υπολογιστών και στην κατηγοριοποίηση ιατρικών εικόνων για διαγνωστικούς σκοπούς.

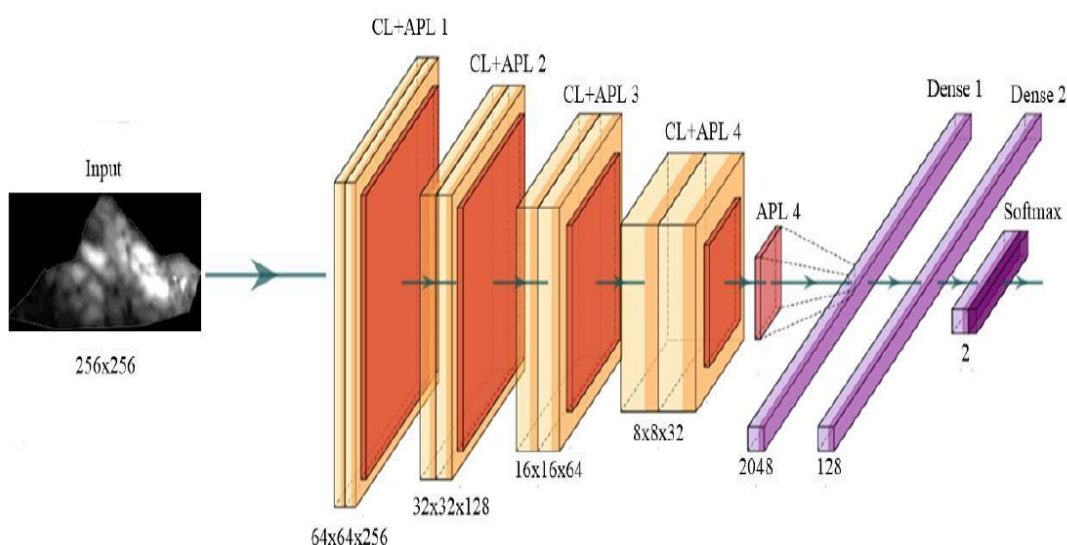
Η γενική αρχιτεκτονική ενός Βαθιού Συνελικτικού Νευρωνικού Δικτύου (Deep Convolutional Neural Network-DCNN) χωρίζεται σε δύο μέρη. Το πρώτο μέρος αποτελείται από διαδοχικά συνελικτικά και Pooling επίπεδα μέσω των οποίων γίνεται η εξαγωγή των χαρακτηριστικών ενώ το δεύτερο μέρος μέσω κατάλληλων επιπέδων παράγει από τα χαρακτηριστικά αυτά το τελικό αποτέλεσμα ανάλογα με το πρόβλημα, για παράδειγμα, μια κατηγοριοποίηση της εικόνας σε δύο κλάσεις. Για παράδειγμα το DCNN που προτείνεται σε δημοσίευση για την κατηγοριοποίηση εικόνων υπερήχων καρδιάς σε συμπτωματικές και ασυμπτωματικές αποτελείται από τέσσερα συνελικτικά επίπεδα ακολουθούμενα από ένα επίπεδο Average Pooling. Τα εννέα (9) αυτά επίπεδα με τη σειρά τους ακολουθούνται από ένα επίπεδο Flatten το οποίο μετατρέπει τον διδιάστατο feature map σε μονοδιάστατο. Το Flatten επίπεδο ακολουθούν δύο Dense κρυφά επίπεδα. Το τελικό αποτέλεσμα προκύπτει από την εφαρμογή ενός Softmax επιπέδου το οποίο παράγει τις πιθανότητες η εικόνα εισόδου να ανήκει σε κάθε μία από τις κλάσεις του προβλήματος. Ως συνάρτηση ενεργοποίησης έχει επιλεγεί η ReLU. Η αρχιτεκτονική φαίνεται στο Σχήμα 3.2 όπου με πορτοκαλί χρώμα διακρίνεται ο εξαγωγέας χαρακτηριστικών και με μωβ ο ταξινομητής [10].

### 3.4 Συναρτήσεις Ενεργοποίησης

Όπως είδαμε και όταν περιγράψαμε τη δομή ενός τεχνητού νευρώνα, οι συναρτήσεις ενεργοποίησης καθορίζουν τον τρόπο με τον οποίον ο νευρώνας θα επηρεάσει το επόμενο επίπεδο του νευρωνικού δικτύου μέσω της εξόδου του. Δηλαδή αποφασίζουν αν ένας νευρώνας θα ενεργοποιηθεί ή όχι και συνεπώς διαδραματίζουν καθοριστικό ρόλο στην εκπαίδευση και επίδοση του δικτύου. Συνεπώς η επιλογή των συναρτήσεων ενεργοποίησης για κάθε επίπεδο του δικτύου αποτελεί σημαντικό στάδιο του σχεδιασμού και της ανάπτυξης ενός μοντέλου ΤΝΔ. Οι κυριότερες συναρτήσεις ενεργοποίησης είναι οι:

- **Binary Step:** Καθώς η έξοδος της είναι δυαδική (0 ή 1) επιτρέπει μόνο δυαδική ταξινόμηση.

$$f(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}$$



Σχήμα 3.2: Παράδειγμα αρχιτεκτονικής DCNN [10]

- **Linear:** Χρησιμοποιείται στα απλά μοντέλα παλινδρόμησης. Είναι ακατάλληλη για εκπαίδευση μοντέλων με backpropagation διότι η παράγωγος της είναι μία σταθερά.

$$f(x) = x$$

- **ReLU:** Είναι η πιο ευρέως χρησιμοποιούμενη συνάρτηση ενεργοποίησης. Υπολογίζεται εύκολα με αποτέλεσμα η εκπαίδευση του νευρωνικού να συγκλίνει πολύ γρήγορα. Επίσης είναι επιρρεπής στο vanishing gradient problem αλλά μόνο για αρνητικές τιμές. Η Leaky ReLU δεν παρουσιάζει vanishing gradient.

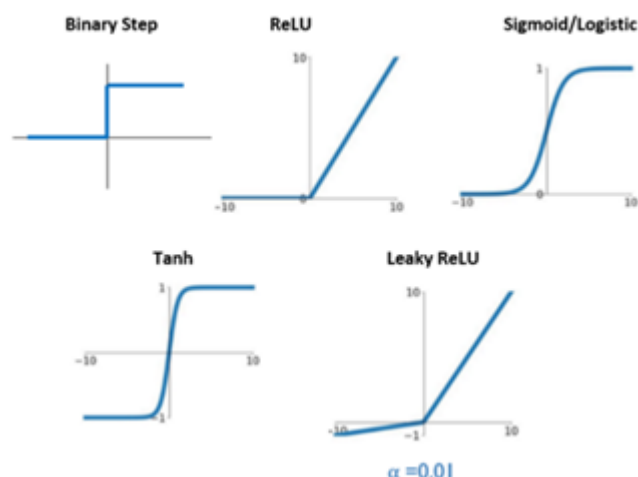
$$f(x) = \begin{cases} x & x \geq 0 \\ 0 & x < 0 \end{cases}$$

- **Sigmoid:** Είναι χρήσιμη στα προβλήματα ταξινόμησης καθώς κανονικοποιεί την έξοδο του νευρώνα στο 0-1, δίνοντας την πιθανότητα της τιμής εισόδου. Ο υπολογισμός της είναι απαιτητικός υπολογιστικά λόγω του εκθετικού και η συνάρτηση συγκλίνει αργά. Επίσης είναι επιρρεπής στο vanishing gradient problem.

$$f(x) = \frac{1}{1 + e^{-x}}$$

- **Tanh:** Κανονικοποιεί την έξοδο του νευρώνα στο εύρος μεταξύ -1 και 1 και σε αντίθεση με τη Sigmoid είναι κεντραρισμένη γύρω από το μηδέν με αποτέλεσμα την ευκολότερη βελτιστοποίηση της συνάρτησης απώλειας. Κι αυτή η συνάρτηση είναι επιρρεπής στο vanishing gradient problem.

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$



Σχήμα 3.3: Βασικές συναρτήσεις ενεργοποίησης

### 3.5 Επίπεδα Υποδειγματοληψίας (Pooling Layers)

Τα επίπεδα υποδειγματοληψίας χρησιμοποιούνται για τη μείωση των διαστάσεων των χαρτών χαρακτηριστικών. Με αυτόν τον τρόπο μειώνεται το πλήθος των παραμέτρων για εκμάθηση και η υπολογιστική πολυπλοκότητα. Το επίπεδο υποδειγματοληψίας συνοψίζει τα χαρακτηριστικά μιας περιοχής του χάρτη χαρακτηριστικών που προκύπτει από ένα συνεκτικό επίπεδο.

### 3.6 Επίπεδα Batch Normalization

Η κανονικοποίηση παρτίδας χρησιμοποιείται για να επιταχύνει και να κάνει πιο σταθερή την εκπαίδευση των νευρωνικών δικτύων. Εφαρμόζει κανονικοποίηση στις εισόδους των επιπέδων. Για μικρή παρτίδα mini batch  $B$  του συνόλου εκπαίδευσης με μέγεθος  $m$  η κανονικοποίηση για είσοδο  $(x^{(1)}, \dots, x^{(d)})$  γίνεται ως εξής:

$$\hat{x}_i^{(k)} = \frac{x_i^{(k)} - \mu_B^{(k)}}{\sqrt{(\sigma_B^{(k)})^2 + \epsilon}}$$

, όπου  $k \in [1, d]$  και  $\mu_B^{(k)}, \sigma_B^{(k)}$  η μέση τιμή και η τυπική απόκλιση στη διάσταση  $k$ .

### 3.7 Επίπεδα Dropout

Ο όρος dropout αναφέρεται στην απομάκρυνση νευρώνων ενός νευρωνικού δικτύου από το επίπεδο εισόδου και από τα κρυφά επίπεδα. Όλες οι συνδέσεις των νευρώνων που αφαιρέθηκαν αφαιρούνται και αυτές προσωρινά. Κατά τη διάρκεια της εκπαίδευσης, σε κάθε εποχή, τα επίπεδα Dropout αφαιρούν κάποιο κλάσμα των κόμβων πριν από τον υπολογισμό του επόμενου επιπέδου. Τα επίπεδα Dropout έχουν ως παράμετρο μία πιθανότητα  $p$  που αντιστοιχεί στην πιθανότητα ένας νευρώνας να απομακρυνθεί. Τα επίπεδα αυτά χρησιμο-

ποιούνται συνήθως για την αντιμετώπιση της υπερεκπαίδευσης διότι μειώνουν τις αλληλεπιδράσεις των νευρώνων κι απλουστεύουν τη δομή του δικτύου [47].

### 3.8 Τεχνικές μείωσης διαστατικότητας των δεδομένων

Σε αρκετές περιπτώσεις απαιτείται η μείωση των διαστάσεων των δεδομένων με τρόπο ώστε να χάνεται όσο το δυνατόν λιγότερη πληροφορία (Dimensionality Reduction). Μία βασική τεχνική μείωσης διαστατικότητας είναι η Ανάλυση Κύριων Συνιστωσών (PCA-Principal Components Analysis).

#### 3.8.1 Ανάλυση Κυρίων Συνιστωσών (Principal Components Analysis)

Έστω σύνολο  $n$  δειγμάτων και  $p$  χαρακτηριστικών. Το σύνολο μπορεί να παρασταθεί ως πίνακας διαστάσεων  $n \times p$  και μπορεί να είναι για παράδειγμα ένα πίνακας tabular δεδομένων ή μία εικόνα. Η μέθοδος PCA προσπαθεί να μειώσει τη διάσταση  $p$  διατηρώντας όσο το δυνατόν περισσότερο τη στατιστική πληροφορία του αρχικού πίνακα. Αυτό το πετυχαίνει δημιουργώντας νέες μεταβλητές χαρακτηριστικών οι οποίες είναι γραμμικές συναρτήσεις των αρχικών, μεγιστοποιούν τη διασπορά και είναι ασυσχέτιστες μεταξύ τους. Αυτές οι νέες μεταβλητές ονομάζονται κύριες συνιστώσες. Ο αλγόριθμος προσπαθεί να συμπίεσει όσο το δυνατόν περισσότερη πληροφορία στην πρώτη κύρια συνιστώσα, όσο το δυνατόν περισσότερη από την υπολειπόμενη πληροφορία στη δεύτερη κύρια συνιστώσα κ.ο.κ. Τα στάδια του αλγορίθμου περιγράφονται στη συνέχεια.

Αρχικά γίνεται κανονικοποίηση των τιμών του πίνακα στο διάστημα  $[0,1]$  ώστε όλες οι μεταβλητές (χαρακτηριστικά) να συνεισφέρουν το ίδιο στην ανάλυση. Η κανονικοποίηση (standardization) γίνεται με βάση τον τύπο:  $x' = \frac{(x-\mu)}{\sigma}$ , όπου  $x$  η τιμή του χαρακτηριστικού πριν την κανονικοποίηση και  $x'$  η τιμή του μετά,  $\mu$  η μέση τιμή του χαρακτηριστικού στα διαθέσιμα δείγματα και  $\sigma$  η τυπική απόκλιση.

Στη συνέχεια ο αλγόριθμος προσπαθεί να εντοπίσει συσχετίσεις μεταξύ των μεταβλητών εισόδου οι οποίες μαρτυρούν επανάληψη πληροφορίας. Γι'αυτό το σκοπό υπολογίζεται ο πίνακας συνδιακύμανσης (covariance matrix) των μεταβλητών. Όταν η συνδιακύμανση  $\text{Cov}(a,b)$  δύο μεταβλητών  $a,b$  είναι θετική τότε οι δύο μεταβλητές αυξάνουν μαζί ενώ όταν είναι αρνητική όταν αυξάνεται η μία μειώνεται η άλλη. Έπειτα, για τον προσδιορισμό των κύριων συνιστωσών, υπολογίζονται οι  $p$  ιδιοτιμές και τα  $p$  ιδιοδιανύσματα του πίνακα συνδιακύμανσης. Αποδεικνύεται ότι τα ιδιοδιανύσματα του πίνακα συνδιακύμανσης δείχνουν τις κατευθύνσεις όπου εντοπίζεται η μεγαλύτερη διασπορά και οι αντίστοιχες ιδιοτιμές δείχνουν την μέγεθος της διασποράς στη συνιστώσα. Διατάσσοντας τα ιδιοδιανύσματα σε φθίνουσα σειρά με βάση την ιδιοτιμή τους, ο αλγόριθμος επιστρέφει τις κύριες συνιστώσες με σειρά σημαντικότητας. Το πλήθος των συνιστωσών που ο αλγόριθμος θα επιλέξει να κρατήσει δίνεται ως παράμετρος. Έστω ότι η παράμετρος αυτή έχει τιμή  $k$ . Ο αλγόριθμος θα κατασκευάσει ένα πίνακα  $k$  στηλών από τα  $k$  πρώτα ιδιοδιανύσματα, έστω  $F$ , ο οποίος ονομάζεται Διάνυσμα Χαρακτηριστικών (Feature Vector). Τελικά, ο νέος πίνακας δεδομένων  $X'$  διαστάσεων  $n \times k$

προκύπτει από τον  $X$  με βάση τον μετασχηματισμό:

$$X' = F^T X^T$$

Για την επιλογή της παραμέτρου  $k$  μπορούμε να χρησιμοποιήσουμε ένα κατώφλι όπως 0.9 ή 0.95.

$$\frac{\sum_{i=1}^k \hat{\rho}_i}{\sum_{i=1}^p \hat{\rho}_i} > threshold$$

Η πολυπλοκότητα του αλγορίθμου για πίνακα εισόδου  $n \times p$  είναι  $O(n \cdot p \cdot \min(n, p) + p^3)$ ,  $O(n \cdot p \cdot \min(n, p))$  για τον υπολογισμό του πίνακα συνδυαστικής και  $O(p^3)$  για τον προσδιορισμό των ιδιοτιμών. Καθώς η πολυπλοκότητα του αλγορίθμου αυξάνεται σημαντικά καθώς αυξάνεται το μέγεθος του πίνακα εισόδου, βιβλιοθήκες, όπως για παράδειγμα η *Sckit-Learn*, χρησιμοποιούν τον αλγόριθμο Singular Value Decomposition (SVD) για τον προσδιορισμό των ιδιοτιμών με πολυπλοκότητα  $O(n \cdot p \cdot \min(n, p))$ . Τέλος, να σημειώσουμε ότι η μέθοδος PCA εφαρμόζεται σε συνεχή δεδομένα. Για τα κατηγορικά χαρακτηριστικά όπως είναι για παράδειγμα το φύλο ή αν το άτομο είναι καπνιστής ή όχι χρησιμοποιούνται άλλες μέθοδοι όπως για παράδειγμα η Multiple Correspondence Analysis (MCA) [48].

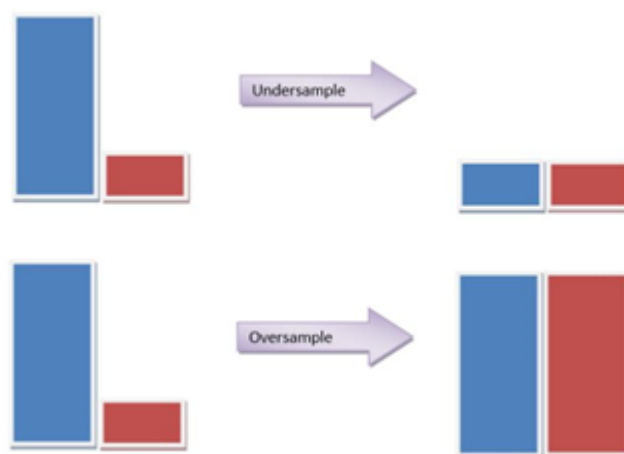
### 3.9 Προβλήματα των δεδομένων

Δύο σημαντικές δυσκολίες που αντιμετωπίζουμε όταν χρησιμοποιούμε μεθόδους Μηχανικής Μάθησης στα προβλήματα ταξινόμησης είναι το **περιορισμένο πλήθος δεδομένων** για την εκπαίδευση των μοντέλων και η **ανισορροπία των δεδομένων**. Ανισορροπία στα δεδομένα έχουμε όταν διαθέτουμε σημαντικά μικρότερο αριθμό δεδομένων από ορισμένες κατηγορίες του προβλήματος σε σχέση με τις υπόλοιπες. Χαρακτηριστικό παράδειγμα ανισορροπίας αποτελεί το δικό μας πρόβλημα όπου όπως θα δούμε διαθέτουμε σχεδόν εξαπλάσια δείγματα από τη μία κλάση εκ των δύο. Η μεν έλλειψη δεδομένων σε πολλές περιπτώσεις δεν επιτρέπει την επαρκή εκπαίδευση των μοντέλων και η δε ανισορροπία οδηγεί στη μονόπλευρη εκπαίδευση των ταξινομητών στις κλάσεις πλειοψηφίας. Για να αντιμετωπιστούν αυτά τα προβλήματα εφαρμόζονται διάφορες τεχνικές επαύξησης και εξισορρόπησης των δεδομένων. Οι τεχνικές επαύξησης των δεδομένων μπορούν να διακριθούν σε τεχνικές **Υπερδειγματοληψίας** (Oversampling) με τις οποίες αυξάνουμε τα δείγματα της κλάσης μειοψηφίας και σε τεχνικές **Υποδειγματοληψίας** (Undersampling) με τις οποίες μειώνουμε τα δείγματα της κλάσης πλειοψηφίας ώστε το σύνολο να γίνει ισορροπημένο (Σχήμα 3.4). Ας δούμε πιο αναλυτικά κάποιες από τις τεχνικές επαύξησης και χειρισμού της ανισορροπίας των δεδομένων.

#### 3.9.1 Synthetic Minority Oversampling Technique -SMOTE

Είναι τεχνική για την υπερδειγματοληψία της κατηγορίας μειοψηφίας. Η απλή επανάληψη δειγμάτων της κατηγορίας μειοψηφίας (διπλότυπες εγγραφές) συχνά δεν οδηγεί σε καλύτερη εκπαίδευση του μοντέλου καθώς δεν προσθέτει νέα πληροφορία. Με τον αλγόριθμο SMOTE συντίθενται από τα υπάρχοντα δεδομένα νέα δείγματα της κατηγορίας μειοψηφίας. Ο αλγόριθμος επιλέγει τυχαία ένα στιγμιότυπο, έστω διάνυσμα  $X$ , της κλάσης μειοψηφίας.



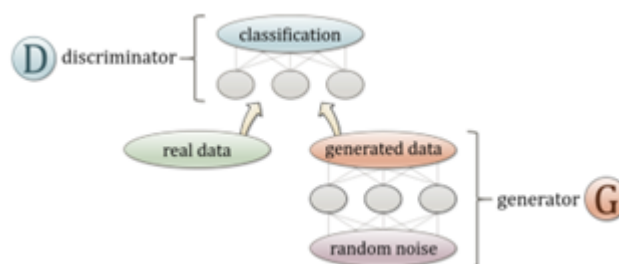


Σχήμα 3.4: Τεχνικές Resampling

Για αυτό το στιγμιότυπο βρίσκει τους  $k$  πλησιέστερους γείτονές του, επιλέγει έναν από αυτούς, έστω διάνυσμα  $Y$ , και υπολογίζει τη διαφορά των δύο διανυσμάτων  $X, Y$ . Τέλος, πολλαπλασιάζει τη διαφορά με έναν τυχαίο αριθμό στο  $[0, 1]$  και έστω  $Z$  το διάνυσμα που προκύπτει. Το συνθετικό δείγμα προκύπτει από το άθροισμα των διανυσμάτων  $X$  και  $Z$  [49].

### 3.9.2 CTGAN

Η τεχνική αυτή προτείνεται στην εργασία Modeling Tabular Data using Conditional GAN των Lei Xu et al. που δημοσιεύτηκε το 2019 [50]. Το μοντέλο που προτείνεται ονομάζεται CTGAN, ανήκει στην κατηγορία των Παραγωγικών Ανταγωνιστικών Δικτύων (Generative Adversarial Networks (GANs)) και έχει στόχο να δημιουργεί συνθετικά δεδομένα πίνακα (synthetic tabular data). Τα βασικά στοιχεία της αρχιτεκτονικής GAN είναι δύο βαθιά νευρωνικά δίκτυα, ο Γεννήτορας (Generator)  $G$  και ο Διευκρινιστής (Discriminator)  $D$ . Ο Γεννήτορας, όπως συνάγεται και από το όνομά του είναι ένα παραγωγικό μοντέλο που εξάγει δεδομένα από μια είσοδο θορύβου και ο Διευκρινιστής είναι ένα μοντέλο ταξινόμησης που ταξινομεί τα δεδομένα που λαμβάνει στην είσοδό του σε πραγματικά και συνθετικά. Στόχος του Γεννήτορα είναι να ξεγελάσει τον Διευκρινιστή, ενώ στόχος του Διευκρινιστή είναι να μάθει να αναγνωρίζει τα πραγματικά από τα συνθετικά δεδομένα. Ακριβώς επειδή τα δύο δίκτυα κατά την εκπαίδευσή τους έχουν αντίθετους στόχους ονομάζονται ανταγωνιστικά. Οι



Σχήμα 3.5: Παραγωγικά Ανταγωνιστικά Δίκτυα (GANs) [11]

Ian J. Goodfellow et al. στο [51] περιγράφουν το Ανταγωνιστικό Δίκτυο όπου ο Γεννήτορας

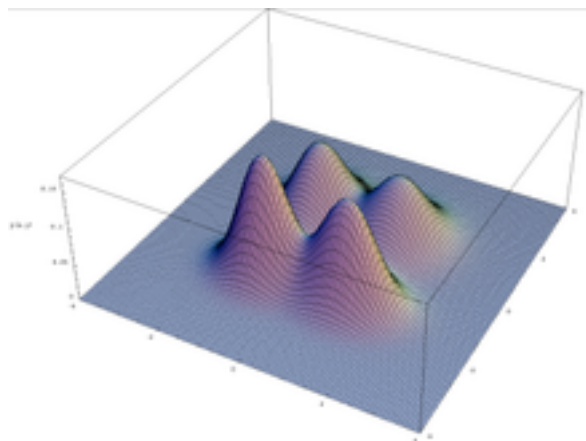
και ο Διευκρινιστής είναι δύο MLPs. Για την εκμάθηση της κατανομής του  $G$   $p_g(x)$  για τα δεδομένα  $x$  ορίζεται prior  $p_z(z)$  των μεταβλητών θορύβου εισόδου και απεικόνιση  $G(z; \theta_g)$  στο χώρο των δεδομένων, όπου  $G$  διαφορίσιμη συνάρτηση που αντιστοιχεί σε ένα MLP με παραμέτρους  $\theta_g$ . Το δεύτερο MLP αντιστοιχεί σε συνάρτηση  $D(x; \theta_d)$  όπου η  $D(x)$  δίνει την πιθανότητα το  $x$  να προέρχεται από τα πραγματικά δεδομένα και όχι από την  $p_g(x)$ . Ο  $D$  εκπαιδεύεται ώστε να μεγιστοποιηθεί η πιθανότητα εκχώρησης της σωστής ετικέτας τόσο στα πραγματικά δείγματα όσο και στα δείγματα από τον  $G$ . Ταυτόχρονα εκπαιδεύεται ο  $G$  ώστε να ελαχιστοποιηθεί το  $\log(1 - D(G(z)))$ . Δηλαδή Γεννήτορας και Διευκρινιστής παίζουν ένα παιχνίδι minimax δύο παικτών με συνάρτηση τιμής (value function)  $V(G,D)$ :

$$\min_G \max_D V(D, G) = \mathbb{E}_{x \sim p_{data}(x)} [\log D(x)] + \mathbb{E}_{z \sim p_z(z)} [1 - \log D(G(z))]$$

Το CTGAN είναι ένα μοντέλο που βασίζεται στα GANs για τη μοντελοποίηση της κατανομής των δεδομένων πίνακα και για τη δειγματοληψία σειρών από την κατανομή οι οποίες θα χρησιμοποιηθούν για την παραγωγή των συνθετικών δεδομένων. Στη συνέχεια θα περιγράψουμε κάποια βασικά σημεία του τρόπου λειτουργίας του CTGAN. Οι προκλήσεις των δεδομένων πίνακα που καλείται να αντιμετωπίσει το CTGAN είναι:

- **Μικτά είδη δεδομένων:** Τα δεδομένα πίνακα αποτελούνται από διακριτές ή συνεχείς στήλες ή και τα δύο. Για να δημιουργηθεί ταυτόχρονα ένας συνδυασμός διακριτών και συνεχών στηλών, τα GANs πρέπει να εφαρμόζουν τόσο softmax όσο και tanh στην έξοδο.
- **Μη Gaussian κατανομές:** Οι συνεχείς τιμές σε δεδομένα πίνακα είναι συνήθως μη Gaussian με αποτέλεσμα να είναι επιρρεπή στο vanishing gradient problem.
- **Πολυτροπικές κατανομές:** Μια πολυτροπική κατανομή είναι μια κατανομή πιθανότητας με περισσότερες από μία κορυφές (peak ή mode στη βιβλιογραφία). Οι Srivastava et al. [52] έδειξαν ότι το vanilla GAN δεν μπορεί να μοντελοποιήσει όλα τα mode σε ένα απλό 2D σύνολο δεδομένων. Παράδειγμα πολυτροπικής κατανομής στο Σχήμα 3.6.
- **Μάθηση από one-hot-encoded διανύσματα**
- **Έντονα μη-ισορροπημένες διακριτές μεταβλητές:** Για παράδειγμα σε σύνολο κλινικών δεδομένων ασθενών έχουμε την κατηγορία “Diabetes” όπου σημειώνουμε αν ο/η ασθενής είναι διαβητικός/ή. Έστω ότι το 90% των ασθενών του συνόλου δεδομένων είναι διαβητικοί και το υπόλοιπο 10% μη διαβητικοί, δηλαδή η κατηγορία “Diabetes” είναι μη-ισορροπημένη. Εάν κάνουμε απευθείας δειγματοληψία σε αυτό το σύνολο δεδομένων, η κατηγορία μειονότητας θα αγνοηθεί λόγω της μικρής αναλογίας των δειγμάτων και η στατιστική της κατανομή δεν θα έχει μεγάλη επίδραση. Αυτό θα οδηγήσει σε απώλεια των σημαντικών πληροφοριών που υπάρχουν στις τιμές της κατηγορίας αυτής.

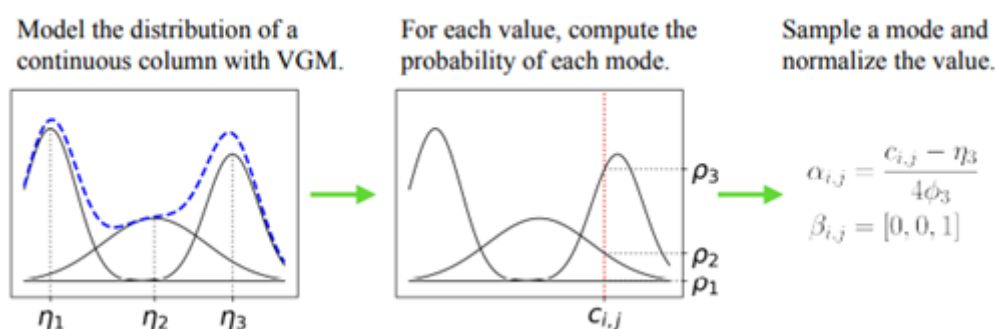
Για να αντιμετωπίσει τις προκλήσεις των μη-Gaussian και των πολυτροπικών κατανομών το CTGAN χρησιμοποιεί τη μέθοδο Mode-specific Normalization. Συνοπτικά, τα τρία στάδια της μεθόδου είναι:



Σχήμα 3.6: Παράδειγμα πολυτροπικής κατανομής

- Για κάθε συνεχή στήλη  $C_i$  χρησιμοποιείται ένα VGM μοντέλο (Variational Gaussian mixture model) για να εκτιμήσει το πλήθος των modes  $m_i$  της κατανομής της.
- Έπειτα, για κάθε τιμή  $c_{i,j}$  της στήλης  $C_i$  υπολογίζεται η πυκνότητα πιθανότητας  $\rho_k(c_{i,j}) = \mu_k N(c_{i,j}; \eta_k, \varphi_k)$  για κάθε mode  $k = 1, \dots, m_i$ , όπου  $\mu_k$  το βάρος και  $\varphi_k$  η τυπική απόκλιση.
- Δεδομένων των  $\rho_k$  επιλέγεται ένα mode και με βάση αυτό κανονικοποιείται η τιμή  $c_{i,j}$ . Η τιμή  $c_{i,j}$  απεικονίζεται με ένα one-hot-encoded διάνυσμα  $\beta_{i,j}$  και την κανονικοποιημένη τιμή  $a_{i,j}$ .

Στο Σχήμα 3.7 φαίνεται ένα παράδειγμα της μεθόδου όπου το πλήθος των modes είναι τρία (3) και επιλέγεται το τρίτο για την τιμή  $c_{i,j}$ . Το one-hot-encoded διάνυσμα  $\beta_{i,j}$  έχει 1 στη θέση που αντιστοιχεί στο τρίτο mode. Τελικά, κάθε σειρά  $j$  του πίνακα απεικονίζεται ως:



Σχήμα 3.7: Mode-specific Normalization

$$r_j = a_{1,j} \oplus \beta_{1,j} \oplus \dots \oplus a_{N_c,j} \oplus \beta_{N_c,j} \oplus d_{1,j} \dots \oplus d_{N_d,j}$$

Όπου  $d$  one-hot-encoded διανύσματα των διακριτών στηλών και  $\oplus$  η απλή συνένωση των διανυσμάτων.  $N_c$  και  $N_d$  το πλήθος των συνεχών και των διακριτών στηλών αντίστοιχα.

Για να αντιμετωπίσει τα προβλήματα από τις μη-ισορροπημένες διακριτές στήλες χρησιμοποιεί έναν Conditional Generator και τη μέθοδο Training-by-Sampling.

### Conditional Generator και Training-by-Sampling

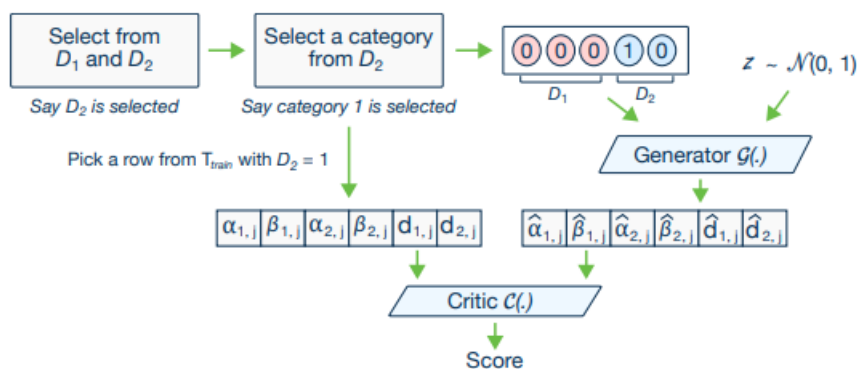
Ο στόχος είναι να γίνει αποτελεσματική επαναδειγματοληψία (resampling) των μη ισοροπημένων δεδομένων με τρόπο ώστε όλες οι κατηγορίες από διακριτά χαρακτηριστικά να δειγματοληπτούνται ισότιμα (αλλά όχι απαραίτητα με βάση την ομοιόμορφη κατανομή) κατά τη διάρκεια της εκπαίδευσης και να ανακτηθεί η κατανομή των πραγματικών δεδομένων κατά τη διάρκεια της δοκιμής (test) με βάση τον τύπο:

$$\mathbb{P}(row) = \sum_{k \in D_{i^*}} \mathbb{P}_{\mathbb{G}}(row | D_{i^*} = k^*) \mathbb{P}(D_{i^*} = k)$$

Στον παραπάνω τύπο  $k^*$  η τιμή της  $i^*$ -οστής διακριτής στήλης  $D(i^*)$  που πρέπει να αντιστοιχιστεί με τα παραγόμενα δείγματα  $r$  και ο Γεννήτορας εκφράζεται ως η υπό συνθήκη κατανομή των σειρών που δίνουν την  $k^*$  στην  $i^*$ -οστή στήλη, δηλαδή  $\mathbb{P}_{\mathbb{G}}(row | D_{i^*} = k^*)$ . Το conditional vector ορίζεται ως  $cond = m_1 \oplus \dots \oplus m_{N_d}$  όπου

$$m_i^{(k)} = \begin{cases} 1 & \text{if } i = i^* \text{ and } k = k^* \\ 0 & \text{otherwise} \end{cases}$$

Τελικά, ο Conditional Generator δημιουργεί συνθετικές σειρές με βάση μία από τις διακριτές στήλες. Στη μέθοδο training-by-sampling, το conditional vector και τα δείγματα εκπαίδευσης δειγματοληπτούνται σύμφωνα με τη log-frequency κάθε κατηγορίας, έτσι το CTGAN μπορεί να εξερευνήσει ομοιόμορφα όλες τις πιθανές διακριτές τιμές. Η έξοδος που παράγεται από τον Conditional Generator πρέπει να αξιολογηθεί από τον Διευκρινιστή (ή Κριτή), ο οποίος εκτιμά την απόσταση μεταξύ της κατανομής  $\mathbb{P}_{\mathbb{G}}(row | cond)$  και της κατανομής  $\mathbb{P}(row | cond)$  των πραγματικών δεδομένων. Η δειγματοληψία των πραγματικών δεδομένων εκπαίδευσης και η κατασκευή του conditional vector γίνεται με τρόπο ώστε να διευκολύνουν τον Κριτή να εκτιμήσει την απόσταση. Γι'αυτό χρησιμοποιείται η μέθοδος Training-by-Sampling (Σχήμα 3.8).



Σχήμα 3.8: Training-by-Sampling

Η αρχιτεκτονική του Conditional Generator είναι:

$$G(z, cond) = \begin{cases} h0 = z \oplus cond \\ h1 = h0 \oplus ReLU(BN(FC_{|cond|+|z| \rightarrow 256}(h0))) \\ h2 = h1 \oplus ReLU(BN(FC_{|cond|+|z|+256 \rightarrow 256}(h1))) \\ \hat{a}_i = \tanh(FC_{|cond|+|z|+512 \rightarrow 1}(h2)) \\ \hat{\beta}_i = \text{gumbel}_{0.2}(FC_{|cond|+|z|+512 \rightarrow m_i}(h2)) \\ \hat{d}_i = \text{gumbel}_{0.2}(FC_{|cond|+|z|+512 \rightarrow |D_i|}(h2)) \end{cases}$$

Η αρχιτεκτονική του Κριτή χρησιμοποιεί το το PacGAN framework με pac size 10 και είναι:

$$C(r_1, \dots, r_10, cond_1, \dots, cond_10) = \begin{cases} h0 = r_1 \oplus \dots \oplus r_10 \oplus cond_1 \oplus \dots \oplus cond_10 \\ h1 = \text{drop}(\text{leaky}_{0.2} FC_{10|r|+10|cond| \rightarrow 256}(h0)) \\ h2 = \text{drop}(\text{leaky}_{0.2} FC_{256 \rightarrow 256}(h1)) \\ C(\cdot) = FC_{256 \rightarrow 1}(h2) \end{cases}$$

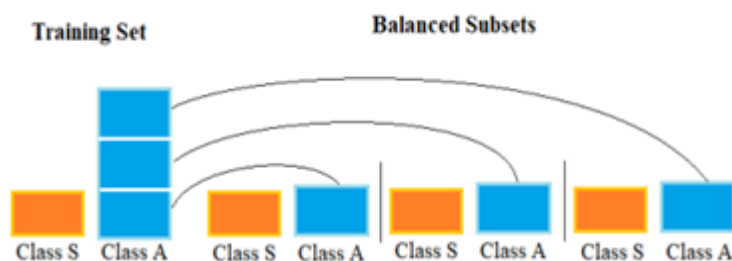
### 3.9.3 Adaptive Synthetic Sampling-ADASYN

Η ιδέα του αλγορίθμου ADASYN είναι η προσαρμοστική δημιουργία δειγμάτων της κλάσης μειοψηφίας. Πιο συγκεκριμένα ο αλγόριθμος παράγει συνθετικά δεδομένα για τα πιο «δύσκολα» δείγματα της κλάσης μειοψηφίας και με τον όρο «δύσκολα» εννοούμε εκείνα τα δείγματα στα οποία το μοντέλο παρουσιάζει μεγαλύτερη δυσκολία μάθησης. Για να αποφασίσει τον αριθμό των συνθετικών δειγμάτων που θα δημιουργηθούν για κάθε δείγμα της κλάσης μειοψηφίας χρησιμοποιήσει μια κατανομή πυκνότητας [53]. Η κατανομή πυκνότητας  $r^i$  για κάθε δείγμα  $x_i$  της κλάσης μειοψηφίας είναι  $r^i = \frac{\Delta_i}{K}$ , όπου  $\Delta_i$  το πλήθος των  $K$  κοντινότερων γειτόνων του  $x_i$  που ανήκουν στην κλάση πλειοψηφίας. Το πλήθος των συνθετικών δειγμάτων που θα δημιουργηθούν για το δείγμα  $x_i$  είναι  $g_i = \hat{r}_i \times G$ , όπου  $\hat{r}_i$  η κανονικοποιημένη τιμή του  $r^i$  και  $G = (m_i - m_s)\beta$  με  $m_i, m_s$  το πλήθος των δειγμάτων των κλάσεων πλειοψηφίας και μειοψηφίας αντίστοιχα και  $\beta \in [0, 1]$  παράμετρος που προσδιορίζει το μέγεθος της ισορροπίας του συνόλου των δεδομένων μετά την επαύξηση. Για  $\beta=1$  προκύπτει πλήρως ισορροπημένο σύνολο.

### 3.9.4 Bootstrap aggregating-BAGGING

Η τεχνική αυτή ανήκει στην κατηγορία της συλλογικής μάθησης και σχεδιάστηκε για να βελτιώνει την επίδοση των αλγορίθμων μηχανικής μάθησης σε προβλήματα ταξινόμησης και παλινδρόμησης [54]. Συνήθως εφαρμόζεται σε μεθόδους δένδρων αποφάσεων αλλά μπορεί να χρησιμοποιηθεί και με άλλα είδη μοντέλων καθώς επιπλέον συνεισφέρει στην αποφυγή της υπερεκπαίδευσης. Όπως παρουσιάζεται και στο Σχήμα 3.9 η μέθοδος αυτή διαμερίζει το σύνολο των δεδομένων σε  $m$  ισοδύναμα και ισορροπημένα υποσύνολα. Για να γίνει αυτό μπορεί να χρειαστεί κάποιο δείγμα να ανήκει σε περισσότερα από δύο υποσύνολα. Στη συνέχεια εκπαιδεύουμε σε κάθε ένα από τα  $m$  υποσύνολα κι από ένα διαφορετικό μοντέλο. Για παράδειγμα σε πρόβλημα ταξινόμησης εκπαιδεύουμε  $m$  διαφορετικούς ταξινομητές. Τέλος, οι προβλέψεις των  $m$  μοντέλων συναθροίζονται με κατάλληλες μεθόδους ώστε να προκύψει το τελικό αποτέλεσμα. Ο αλγόριθμος Random Forest αποτελεί ίσως τον γνωστότερο αλγόριθμο

συλλογικής μάθησης και Bagging. Η τεχνική Bagging μπορεί να χρησιμοποιηθεί και για την αντιμετώπιση της ανισορροπίας του συνόλου των δεδομένων δημιουργώντας ισορροπημένα υποσύνολα από το αρχικό όπως παρουσιάζεται σε εργασία για την ταξινόμηση υπερήχων καρωτίδας (Σχήμα 3.9) [12].



Σχήμα 3.9: Τεχνική Bagging [12]

Επιπλέον, καθοριστικό παράγοντα στην ορθή αξιολόγηση των μοντέλων που έχουν εκπαιδευτεί σε μη ισορροπημένα σύνολα δεδομένων αποτελεί η προσεκτική επιλογή της μετρικής αξιολόγησης. Σε αυτή την περίπτωση οι πιο κατάλληλες μετρικές είναι η F1-Score, η Balanced Accuracy, η ROC/AUC και η Matthew Correlation Coefficient (MCC)

### 3.9.5 Μεταφορά Μάθησης-Transfer Learning

Η Μεταφορά Μάθησης είναι η επαναχρησιμοποίηση ενός προεκπαιδευμένου μοντέλου σε ένα νέο πρόβλημα. Αυτή τη στιγμή είναι πολύ δημοφιλές στη βαθιά μάθηση, επειδή μπορεί να εκπαιδεύσει βαθιά νευρωνικά δίκτυα με σχετικά λίγα δεδομένα. Επιλύει τις υπολογιστικές προκλήσεις υψηλής απόδοσης που απαιτούνται για εικόνες πλούσιες σε πληροφορία, όπως οι ιατρικές, και μειώνει και τον χρόνο που απαιτείται για την εκπαίδευση του μοντέλου σε σύγκριση με το DL. Αυτή η εξοικονόμηση χρόνου μπορεί να είναι ζωτικής σημασίας σε ιατρικές εφαρμογές [10].

## Κεφάλαιο 4

# Πολυτροπική Μάθηση

---

### 4.1 Εισαγωγή

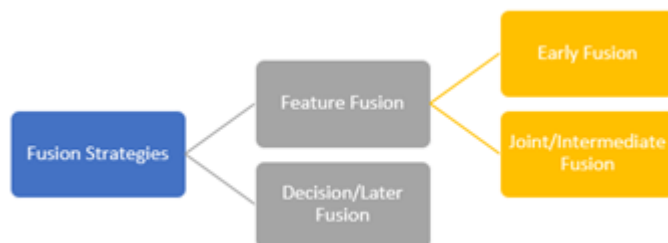
Πολλές σύγχρονες εφαρμογές όπως όπως η οπτικοακουστική αναγνώριση ομιλίας (AVSR), ο υποτιτλισμός εικόνων και βίντεο, η οπτική απάντηση ερωτήσεων (VQA), η δημιουργία εικόνων από κείμενο (text-to-image) και πολλές ιατρικές εφαρμογές περιλαμβάνουν περισσότερους από έναν τύπους δεδομένων εισόδου. Κατά την ιατρική διάγνωση σπάνια λαμβάνεται υπόψη μόνο ένα είδος δεδομένων αλλά εξετάζεται μία ποικιλία από αυτά όπως απεικονιστικές καταγραφές, βιοχημικοί δείκτες, κλινικά χαρακτηριστικά, ηχητικές καταγραφές κ.α. Η Πολυτροπική Μάθηση (Multimodal Learning) υλοποιεί ακριβώς αυτό και ασχολείται με τη δημιουργία μοντέλων Τεχνητής Νοημοσύνης τα οποία αξιοποιούν μια ποικιλία διαφορετικών δεδομένων. Τα διαφορετικά δεδομένα ονομάζονται modalities που στα ελληνικά ο όρος μπορεί να μεταφραστεί ως "τρόποι". Στη συνέχεια της εργασίας προτιμήσαμε να χρησιμοποιούμε το όρο modalities.

### 4.2 Κατηγορίες της Πολυτροπικής Μάθησης

Στην παρούσα εργασία θα εστιάσουμε στις μεθόδους πολυτροπικής μάθησης στο πεδίο της Βαθιάς Μάθησης. Οι μέθοδοι συνδυασμού διαφορετικών ειδών δεδομένων ως προς τον τρόπο που γίνεται ο συνδυασμός διακρίνονται στις παρακάτω κατηγορίες [13, 55]:

- **Feature-Level Fusion:** Στη συγχώνευση σε επίπεδο χαρακτηριστικών τα χαρακτηριστικά των διαφορετικών modalities συνενώνονται σε ένα κοινό διάνυσμα το οποίο χρησιμοποιείται ως είσοδος στο πολυτροπικό μοντέλο. Το κοινό διάνυσμα των modalities μπορεί να προκύψει είτε από συνένωση των αρχικών χαρακτηριστικών είτε από συνένωση χαρακτηριστικών που έχουν εξαχθεί από τα modalities μέσω άλλων μοντέλων όπως για παράδειγμα τα CNNs για τις εικόνες.
- **Decision-Level/Late Fusion:** Στη συγχώνευση σε επίπεδο απόφασης, αρχικά, χρησιμοποιούνται απλά μοντέλα, ένα για κάθε modality. Τα αποτελέσματα των προβλέψεων αυτών των μοντέλων συγχωνεύονται με διάφορες μεθόδους όπως η επιλογή της πλειοψηφίας, το σταθμισμένο άθροισμα ή τον μέσο όρο των επιμέρους προβλέψεων, για να προκύψει το τελικό αποτέλεσμα.

Για την κατηγορία Feature-Level Fusion μπορούμε να διακρίνουμε δύο περιπτώσεις [13, 55]:

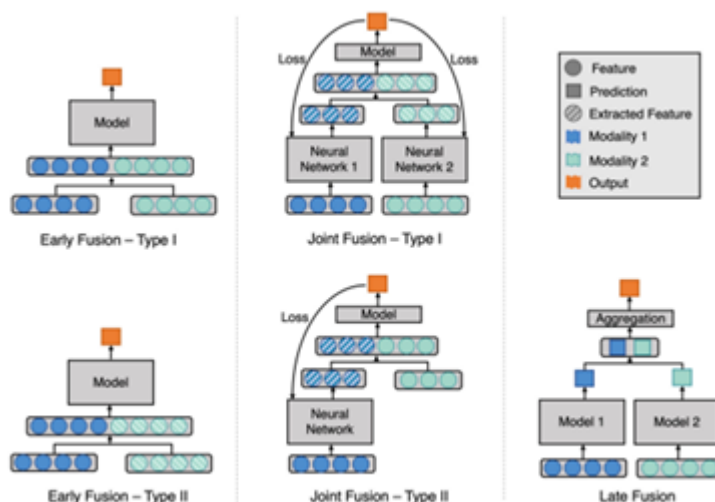


Σχήμα 4.1: Κατηγορίες της Πολυτροπικής Μάθησης

- **Early Fusion:** Ανήκει στην κατηγορία Feature-Level Fusion. Τα χαρακτηριστικά που συνενώνονται για να προκύψει το ενιαίο διάνυσμα χαρακτηριστικών που τροφοδοτεί τελικά το πολυτροπικό μοντέλο είναι είτε τα αρχικά χαρακτηριστικά των modalities είτε αυτά που έχουν εξαχθεί από ήδη εκπαιδευμένους εξαγωγείς χαρακτηριστικών.
- **Intermediate/Joint Fusion:** Ανήκει στην κατηγορία Feature-Level Fusion. Σε αντίθεση με τη μέθοδο Early Fusion, κατά την εκπαίδευση του πολυτροπικού μοντέλου το σφάλμα της τελικής πρόβλεψης μεταδίδεται πίσω στο μοντέλο εξαγωγής χαρακτηριστικών δημιουργώντας έτσι καλύτερες αναπαραστάσεις χαρακτηριστικών σε κάθε εποχή. Δηλαδή το μοντέλο εξαγωγής χαρακτηριστικών εκπαιδεύεται ταυτόχρονα με το πολυτροπικό.

Στο Σχήμα 4.2 παρουσιάζονται οι διάφορες στρατηγικές δημιουργίας πολυτροπικών μοντέλων όπως τις περιγράψαμε παραπάνω.

Βασικό πλεονέκτημα της μεθόδου Decision-Level Fusion είναι το ότι επιτρέπει τη χρήση όλων των διαθέσιμων δεδομένων ακόμη και σε περιπτώσεις που για κάποιο δείγμα διαθέτουμε δεδομένα μόνο από το ένα modality καθώς για κάθε κατηγορία δεδομένων χρησιμοποιεί διαφορετικό μοντέλο. Όσον αφορά στις μεθόδους Feature-Level Fusion ξεχωρίζουν για τη δυνατότητα μοντελοποίησης αλληλεπιδράσεων μεταξύ των χαρακτηριστικών με διαφορετικούς τρόπους [13].



Σχήμα 4.2: Αρχιτεκτονικές Πολυτροπικής Μάθησης [13]

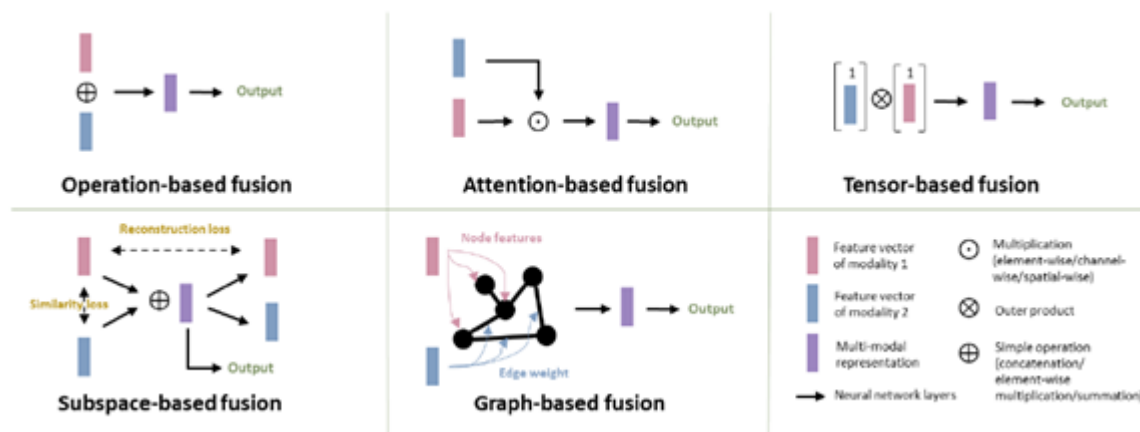
Το Feature-Level Fusion δηλαδή η συνένωση των χαρακτηριστικών όλων των modalities



μπορεί να γίνει με διάφορους τρόπους οι οποίοι αποτελούν και πεδίο έρευνας και επέκτασης. Αρκετοί από αυτούς που χρησιμοποιούνται στις μέρες μας στις εφαρμογές βαθιάς μάθησης παρουσιάζονται στο Σχήμα 4.3 και περιγράφονται στη συνέχεια [55]:

- **Συγχώνευση βασισμένη σε απλές πράξεις διανυσμάτων-Operation-based fusion methods:** Για να συνδυαστούν διαφορετικά διανύσματα χαρακτηριστικών συνήθως χρησιμοποιούνται απλές πράξεις μεταξύ διανυσμάτων όπως η συνένωση (concatenation), η πρόσθεση κατά στοιχείο (element-wise summation) και ο πολλαπλασιασμός κατά στοιχείο (element-wise multiplication). Αυτοί οι τρόποι συγχώνευσης ξεχωρίζουν για την απλότητά τους και αρκετές εργασίες ανέδειξαν την αποτελεσματικότητά τους σε σχέση με τα unimodal μοντέλα, από την άλλη μεριά, όμως, απαιτούν τη μετατροπή των διανυσμάτων χαρακτηριστικών στην ίδια διάσταση και δεν εκμεταλλεύονται πλήρως τους πολύπλοκους συσχετισμούς μεταξύ των ετερογενών δεδομένων [55]. Επιπλέον, η συγχώνευση δημιουργεί μεγάλα διανύσματα χαρακτηριστικών τα οποία μπορεί να οδηγήσουν σε υπερεκπαίδευση όταν η ποσότητα των δεδομένων δεν είναι επαρκής [13]. Ένα ακόμη μειονέκτημα αυτής της μεθόδου συγχώνευσης είναι το γεγονός ότι τα υψηλών διαστάσεων χαρακτηριστικά των εικόνων μπορεί να υπερισχύουν σε βάρος των σημαντικά λιγότερων κλινικών χαρακτηριστικών [55]. Για να αντιμετωπίσουν αυτό το πρόβλημα οι Yan et al. πρότειναν τη χρήση ενός αυτόματου κωδικοποιητή απενεργοποίησης θορύβου για να αυξήσουν τη διάσταση των κλινικών χαρακτηριστικών [56].
- **Συγχώνευση βασισμένη σε πράξεις τανυστών-Tensor-based fusion methods:** Οι μέθοδοι συγχώνευσης που βασίζονται σε τανυστές εφαρμόζουν εξωτερικά γινόμενα μεταξύ διανυσμάτων χαρακτηριστικών διαφορετικών modalities για να σχηματίσουν έναν πίνακα συν-συμβάντος (co-occurrence matrix) υψηλότερης τάξης. Οι αλληλεπιδράσεις υψηλής τάξης τείνουν να παρέχουν περισσότερες προγνωστικές πληροφορίες πέρα από αυτές που μπορούν να παρέχουν τα μεμονωμένα χαρακτηριστικά. Για παράδειγμα η τιμή της αρτηριακής πίεσης από μόνη της δίνει λιγότερη πληροφορία από ότι αν συνδυαστεί, για παράδειγμα, με το γεγονός ότι το άτομο λαμβάνει κάποιο αντιυπερτασικό φάρμακο.
- **Συγχώνευση βασισμένη σε υπόχωρους-Subspace-based fusion methods:** Οι μέθοδοι αυτοί στοχεύουν στην εκμάθηση ενός κοινού υπόχωρου πολυτροπικότητας. Μια δημοφιλής στρατηγική είναι η ενίσχυση της συσχέτισης ή της ομοιότητας των χαρακτηριστικών από διαφορετικά modalities.
- **Συγχώνευση βασισμένη σε μηχανισμούς προσοχής-Attention-based fusion:** Οι μέθοδοι αυτοί υπολογίζουν και ενσωματώνουν τους βαθμούς σημασίας (βαρύτητα προσοχής) των χαρακτηριστικών των διαφορετικών modalities κατά τη συγχώνευση. Αυτό προσομοιώνει την καθημερινή κλινική πρακτική καθώς, για παράδειγμα, οι πληροφορίες από τις κλινικές αναφορές ενός ασθενούς μπορεί να αποτελούν ένδειξη για τους γιατρούς να δώσουν μεγαλύτερη προσοχή σε μια συγκεκριμένη περιοχή μιας εικόνα MRI.
- **Συγχώνευση βασισμένη σε γράφους-Graph-based fusion method:** Η συγχώνευση βασισμένη σε γράφους υλοποιεί την ιδέα ότι οι αλληλεπιδράσεις και οι συσχετίσεις

μεταξύ των χαρακτηριστικών μπορούν να αναπαρασταθούν με τη μορφή γραφημάτων. Πάνω σε αυτό έχουν γίνει εργασίες όπως η [16] όπου το προτεινόμενο GCN μοντέλο ξεπέρασε τις επιδόσεις του MLP με συγχώνευση διανυσμάτων χαρακτηριστικών.



Σχήμα 4.3: Τρόποι συγχώνευσης των δεδομένων

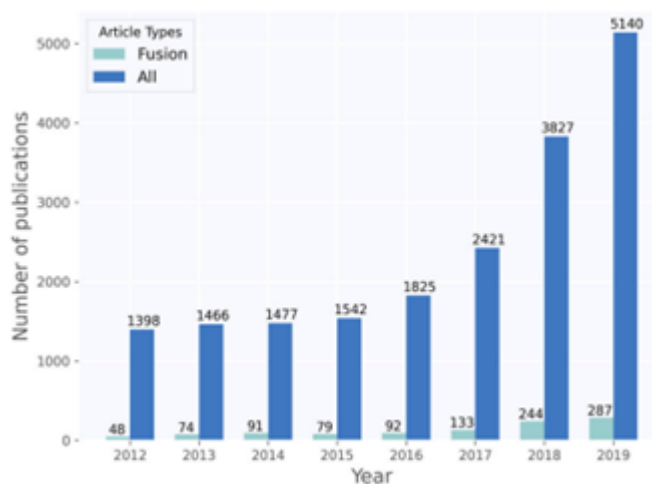
### 4.3 Σχετικές Εργασίες

Σε αυτό το κομμάτι της εργασίας θα αναφερθούμε σε κάποιες σχετικές εργασίες οι οποίες παρουσιάζουν εφαρμογές της πολυτροπικής μάθησης στον τομέα της βιοϊατρικής. Κάποιες από αυτές τις εργασίες θα τις παρουσιάσουμε αναλυτικά αρχικά διότι η κατανόηση των προτεινόμενων μοντέλων αποτελεί βήμα προς τη δημιουργία του δικού μας μοντέλου, επιπλέον θίγουν τεχνικές βελτίωσης της εκπαίδευσης των μοντέλων όπως τεχνικές επαύξησης και εξισορρόπησης των δεδομένων και τέλος δίνουν υλικό για μελλοντικές προεκτάσεις της παρούσας εργασίας. Έπειτα θα αναφερθούμε σε εργασίες που αναδεικνύουν εφαρμογές της Τεχνητής Νοημοσύνης στην ιατρική απεικόνιση και στο ευρύτερο πεδίο έρευνας για την αθηροσκλήρωση και την καρωτιδική νόσο.

#### 4.3.1 Πολυτροπική Μάθηση στη Βιοϊατρική

Οι Shih-Cheng Huang et al. [13] το 2020 δημοσίευσαν συστηματική ανασκόπηση των μεθόδων συνδυασμού απεικονιστικών ιατρικών δεδομένων και καταγραφών ηλεκτρονικής υγείας με χρήση βαθιάς μάθησης. Ενδιαφέρον παρουσιάζει το διάγραμμα που παρουσιάζεται στη δημοσίευσή τους και παρατίθεται στο Σχήμα 4.4 που δείχνει την εξέλιξη του πλήθους των δημοσιεύσεων στη βαθιά μάθηση για την ιατρική απεικόνιση τα τελευταία χρόνια και αναδεικνύει ότι οι μέθοδοι συνδυασμού των δεδομένων αποτελούν ένα μικρό αλλά αυξανόμενο υποσύνολο των δημοσιεύσεων αυτών. Στα συμπεράσματα της εργασίας αυτής διαπιστώνεται ότι η χρήση πολυτροπικών μοντέλων οδήγησε γενικά σε αύξηση της ακρίβεια των προβλέψεων (1,2-27,7%) και του AUROC score (0,02-0,16) ενώ μια ποικιλία μεθόδων συνδυασμού των δεδομένων χρησιμοποιήθηκαν. Επίσης παρατηρήθηκε ότι στις περισσότερες εργασίες χρησιμοποιήθηκαν μικρής πολυπλοκότητας κλινικά δεδομένα όπως δημογραφικές πληροφορίες (ηλικία, φύλο κ.α.), περιορισμένο εύρος κατηγορικών ιατρικών δεδομένων (υπέρταση,

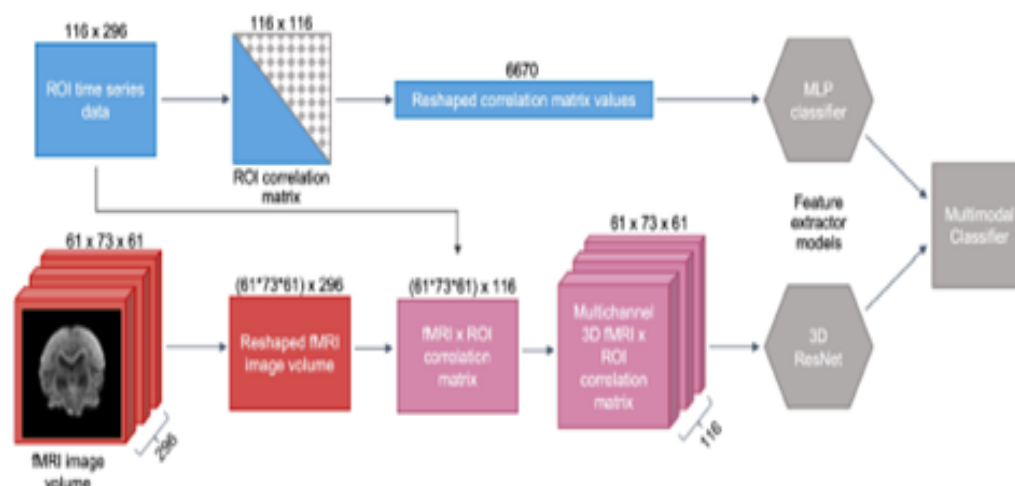
κάπνισμα) και σε ορισμένες περιπτώσεις απλές αιματολογικές και πιο εξειδικευμένες εξετάσεις. Στις περισσότερες περιπτώσεις που μελετήθηκαν χρησιμοποιήθηκε η μέθοδος Early Fusion, ενώ η μέθοδος Late Fusion αποδείχθηκε πολύ χρήσιμη στις περιπτώσεις που για κάποια δείγματα διαθέτουμε μόνο το ένα είδος δεδομένου, για παράδειγμα για έναν ασθενή διαθέτουμε μόνο κλινικά δεδομένα αλλά όχι και εικόνες.



Σχήμα 4.4: Εξέλιξη των δημοσιεύσεων στη βαθιά μάθηση για την ιατρική απεικόνιση [13]

Η εργασία των Michelle Tang et al. η οποία δημοσιεύτηκε το 2020 προτείνει ένα μοντέλο βαθιάς πολυτροπικής μάθησης για τη διάγνωση της Διαταραχής Αυτιστικού Φάσματος (ΔΑΦ) [14]. Χρησιμοποιεί fMRI εικόνες και χρονοσειρές ROI 1035 ατόμων εκ των οποίων οι 505 διαγνώστηκαν με ΔΑΦ. Το μοντέλο εξαγεί αναπαραστάσεις τόσο για τη σάρωση fMRI όσο και για την είσοδο ROI, συνενώνει τα διανύσματα χαρακτηριστικών χρησιμοποιώντας απλή συνένωση διανυσμάτων και καταλήγει σε μια διάγνωση χρησιμοποιώντας τα συνδυασμένα χαρακτηριστικά, τα οποία διαδίδονται μέσω τεσσάρων πλήρως συνδεδεμένων στρωμάτων πριν ταξινομηθούν σε μία από τις δύο κατηγορίες. Για την εξαγωγή των χαρακτηριστικών χρησιμοποιεί ένα MLP για τις σειρές ROI και το ResNet18 3D για τον πίνακα συσχέτισης (correlation matrix) μεταξύ fMRI και χρονοσειρών ROI. Τελικά, το δίκτυο ResNet-18, εκπαιδευμένο στους χάρτες ενεργοποίησης ROIxfMRI πέτυχε ακρίβεια ταξινόμησης 73,1%, ο ταξινομητής MLP με τις χρονοσειρές ROI πέτυχε ακρίβεια 70,8% και το πολυτροπικό μοντέλο πέτυχε ακρίβεια 74%, recall 95% και F1-score 0,805. Η σχετικά χαμηλή επίδοση των μοντέλων αποδόθηκε στο μικρό μέγεθος του συνόλου των δεδομένων. Η αρχιτεκτονική φαίνεται στο Σχήμα 4.5

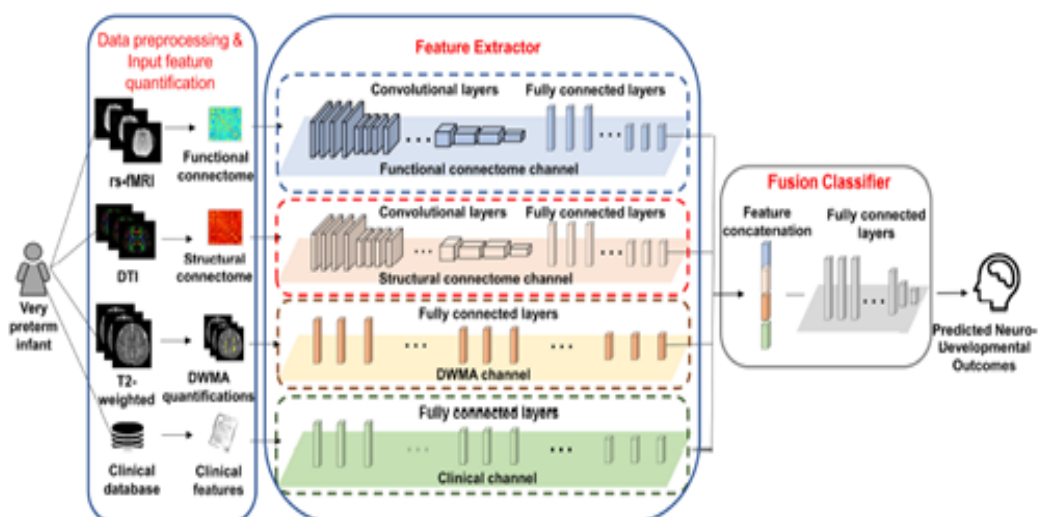
Στην εργασία των Lili He et. al, που δημοσιεύτηκε το 2021, προτείνεται μοντέλο βαθιάς πολυτροπικής μάθησης για την έγκαιρη πρόβλεψη νευροαναπτυξιακών διαταραχών σε πρόωρα βρέφη χρησιμοποιώντας απεικονίσεις μαγνητικής τομογραφίας (MRI) και κλινικά δεδομένα, όπως για παράδειγμα πληροφορίες για τη μητέρα, την κύηση και τη γέννα [15]. Χρησιμοποιήθηκαν τρεις κατηγορίες απεικονίσεων MRI, απεικόνιση των δομικών συνδέσεων του εγκεφάλου με απεικόνιση τανυστή διάχυσης (DTI), λειτουργικά δεδομένα των συνδέσεων του εγκεφάλου από δεδομένα σύνδεσης λειτουργικής μαγνητικής τομογραφίας σε κατάσταση ηρεμίας (rs-fMRI) και DWMA ποσοτικοποιημένο από ανατομικές εικόνες με στάθμιση T2. Η αρχιτεκτονική του πολυτροπικού μοντέλου παρουσιάζεται στο Σχήμα 4.6 και αποτελείται από



Σχήμα 4.5: Προτεινόμενη αρχιτεκτονική στο [14]

έναν feature extractor για κάθε είδος δεδομένων. Ως features extractors για τα δεδομένα DTI και rs-fMRI χρησιμοποιήθηκαν 16 συνελκτικά επίπεδα και 5 επίπεδα υποδειγματοληψίας που υιοθετήθηκαν από το προεκπαιδευμένο μοντέλο VGG-19 ακολουθούμενα από πλήρως συνδεδεμένα επίπεδα. Καθώς οι διαστάσεις του DWMA και των κλινικών δεδομένων είναι μικρές, για αυτά χρησιμοποιήθηκαν μόνο πλήρως συνδεδεμένα μπλοκ. Στη συνέχεια έγινε συνένωση των τεσσάρων κωδικοποιημένων διανυσμάτων χαρακτηριστικών και το διάνυσμα που προέκυψε χρησιμοποιήθηκε ως είσοδος σε πλήρως συνδεδεμένα επίπεδα πριν το δίκτυο εξάγει την πρόβλεψή του. Οι επιδόσεις των πολυτροπικών μοντέλων σε όλες τις δοκιμές που έγιναν ήταν καλύτερες από τις αντίστοιχες των απλών μοντέλων και η επιτευχθείσα ακρίβεια, έπειτα από εξισορρόπηση και επαύξηση των δεδομένων, ήταν κοντά στο 85%. Για την επαύξηση και εξισορρόπηση του συνόλου εκπαίδευσης χρησιμοποιήθηκε η μέθοδος που περιγράφεται στο [57] η οποία δημιουργεί συνθετικά δείγματα της κλάσης μειοψηφίας με βάση τους  $k$ -κοντινότερους γείτονες. Πιο συγκεκριμένα, για το τυχαίο επιλεγμένο δείγμα της κλάσης μειοψηφίας  $x_0$  το συνθετικό δείγμα που προκύπτει είναι  $x_s = \beta_0 x_0 + \beta_k x_k$ , όπου  $[x_1, \dots, x_k]$  οι  $k$ -κοντινότεροι, με βάση την Ευκλείδεια απόσταση, γείτονες του  $x_0$  και  $\beta_i$  βάρη για τα οποία  $\sum_{i=0}^k \beta_i = 1$ .

Στη δημοσίευση του 2019 [56] οι Rui Yan et al. αναπτύσσουν πολυτροπικά μοντέλα για την κατηγοριοποίηση των όγκων του μαστού σε καλοήθεις και κακοήθεις χρησιμοποιώντας εικόνες και δεδομένα από τον ηλεκτρονικό ιατρικό φάκελο ασθενών. Το σύνολο των δεδομένων αποτελείται από 185 ασθενείς εκ των οποίων 82 ήταν καλοήθεις και οι 103 κακοήθεις. Η μέθοδος που εφαρμόζεται είναι η Feature-Level Fusion. Όπως έχουμε ήδη αναφέρει σημαντικό μειονέκτημα της μεθόδου αποτελεί το γεγονός η συγχώνευση των χαμηλών διαστάσεων δεδομένων του ιατρικού φακέλου με τα υψηλών διαστάσεων χαρακτηριστικά από τις εικόνες δεν είναι πάντα αποτελεσματική καθώς τα δεύτερα υπερिशύουν. Για να αντιμετωπιστεί αυτό το πρόβλημα η δημοσίευση προτείνει τη χρήση ενός αυτοκωδικοποιητή αποθορυβοποίησης με στόχο την αύξηση των διαστάσεων των δεδομένων ιατρικού φακέλου πριν τη συγχώνευση των δεδομένων. Έτσι η αρχική διάσταση των 29 κλινικών χαρακτηριστικών αυξήθηκε σε 580. Για την εξαγωγή των χαρακτηριστικών από τις εικόνες χρησιμοποιείται



Σχήμα 4.6: Προτεινόμενη αρχιτεκτονική στο [15]

το VGG16 ως εξής. Αρχικά γίνεται εξαγωγή των χαρτών χαρακτηριστικών (feature map) από το τρίτο, τέταρτο και πέμπτο συνελκτικό επίπεδο του VGG16. Έπειτα οι χάρτες χαρακτηριστικών διαστάσεων  $56 \times 56 \times 256$ ,  $28 \times 28 \times 512$  και  $14 \times 14 \times 512$  αντίστοιχα συμπιέζονται με επίπεδα average pooling σε διανύσματα διαστάσεων  $1 \times 256$ ,  $1 \times 512$  και  $1 \times 512$  και τα διανύσματα αυτά συνενώνονται σε ένα διάνυσμα διάστασης 1280. Τελικά τα διανύσματα διάστασης 1280 και τα αντίστοιχα διανύσματα από τα κλινικά δεδομένα όπως προέκυψαν από τον αυτοκωδικοποιητή συγχωνεύτηκαν σε ένα διάνυσμα διάστασης 1860. Τέλος το διάνυσμα διάστασης 1860 τροφοδότησε ένα μπλοκ τριών FC επιπέδων το οποίο έδωσε την τελική κατηγοριοποίηση σε καλοήθεις και κακοήθεις. Όσον αφορά στη διαδικασία εκπαίδευσης ο αυτοκωδικοποιητής αποθρομβοποίησης και το VGG16 εκπαιδεύονται σε κάθε modality ξεχωριστά και έπειτα οι παράμετροί τους «παγώνουν» και εκπαιδεύεται το FC νευρωνικό στα συγχωνευμένα δεδομένα.

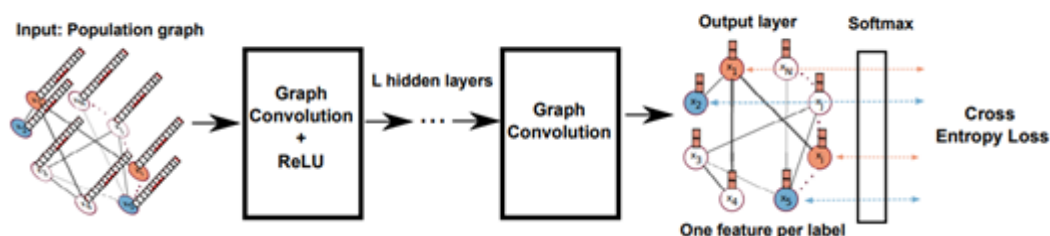
Οι Parisot et al. στο [16] πρότειναν την κατασκευή ενός αραιού, σταθμισμένου γραφήματος από τα φαινοτυπικά, δημογραφικά κ.α. χαρακτηριστικά και τα χαρακτηριστικά εικόνας για την πρόβλεψη της διαταραχής του φάσματος του αυτισμού και της νόσου του Αλτσχάιμερ. Πιο συγκεκριμένα, οι κόμβοι του γραφήματος αντιστοιχούν στα διανύσματα χαρακτηριστικών που εξήχθησαν από τις εικόνες fMRI και οι φαινοτυπικές πληροφορίες ενσωματώθηκαν στα βάρη των ακμών. Για να μειώσουν τις διαστάσεις των διανυσμάτων χαρακτηριστικών που εξήχθησαν από τις εικόνες πειραματίστηκαν με τέσσερις τεχνικές τις οποίες θα παραθέσουμε συνοπτικά καθώς ενδεχομένως μπορούν να χρησιμοποιηθούν σε περιπτώσεις Feature-Level Fusion στις οποίες τα υψηλών διαστάσεων διανύσματα χαρακτηριστικών των εικόνων υπερισχύουν σε βάρος των μικρών διαστάσεων κλινικών, δημογραφικών κ.α. χαρακτηριστικών. Οι μέθοδοι επιλογής χαρακτηριστικών (feature selection) που προτάθηκαν στην [16] είναι οι παρακάτω όπου με  $C$  συμβολίζεται το πλήθος των επιθυμητών χαρακτηριστικών του διανύσματος μετά το feature extraction:

- **Recursive Feature Elimination (RFE) με χρήση ταξινομητή Ridge Classifier:**

Είναι μια επαναληπτική διαδικασία όπου σε κάθε επανάληψη ο ταξινομητής Ridge εκ-

παιδεύεται στο σύνολο εκπαίδευσης με τα τρέχοντα διανύσματα χαρακτηριστικών. Στο τέλος κάθε επανάληψης, με βάση τους συντελεστές του ταξινομητή, τα χαρακτηριστικά διατάσσονται σε φθίνουσα σειρά και αυτά με τον μικρότερο συντελεστή αφαιρούνται από το διάνυσμα χαρακτηριστικών. Ο αλγόριθμος σταματάει όταν τα τρέχοντα διανύσματα έχουν τις επιθυμητές διαστάσεις.

- **Principal Component Analysis:** Έχει περιγραφεί αναλυτικά σε προηγούμενο κεφάλαιο. Στην [16] δεν λειτούργησε καλά λόγω του μεγάλου πλήθους των χαρακτηριστικών (6105) σε σχέση με το πλήθος των διαθέσιμων δεδομένων-δειγμάτων (871).
- **Χρήση MLP:** Σε αυτή την προσέγγιση ένα απλό feedforward MLP που αποτελείται από ένα κρυφό επίπεδο διάστασης  $C$  εκπαιδεύεται στο σύνολο εκπαίδευσης. Το τελικό διάνυσμα χαρακτηριστικών λαμβάνεται εξάγοντας το διάνυσμα  $C$  διαστάσεων που δημιουργείται στο κρυφό επίπεδο του MLP.
- **Αυτοκωδικοποιητές (Autoencoders):** Οι αυτοκωδικοποιητές προσπαθούν να μάθουν μια αναπαράσταση χαμηλότερων διαστάσεων (έναν κώδικα) των δεδομένων εισόδου. Οι Parisot et al. δημιούργησαν έναν αυτοκωδικοποιητή που αποτελείται από τρία πλήρως συνδεδεμένα επίπεδα. Ένα επίπεδο εισόδου, ένα κρυφό επίπεδο μεγέθους  $C$  με σιγμοειδή συνάρτηση ενεργοποίησης και ένα επίπεδο εξόδου με ενεργοποίηση  $\tanh$ . Ο αυτοκωδικοποιητής εκπαιδεύτηκε με συνάρτηση απώλειας το μέσο τετραγωνικό σφάλμα. Τα διανύσματα χαρακτηριστικών μεγέθους  $C$  εξάχθηκαν από το κρυφό επίπεδο.

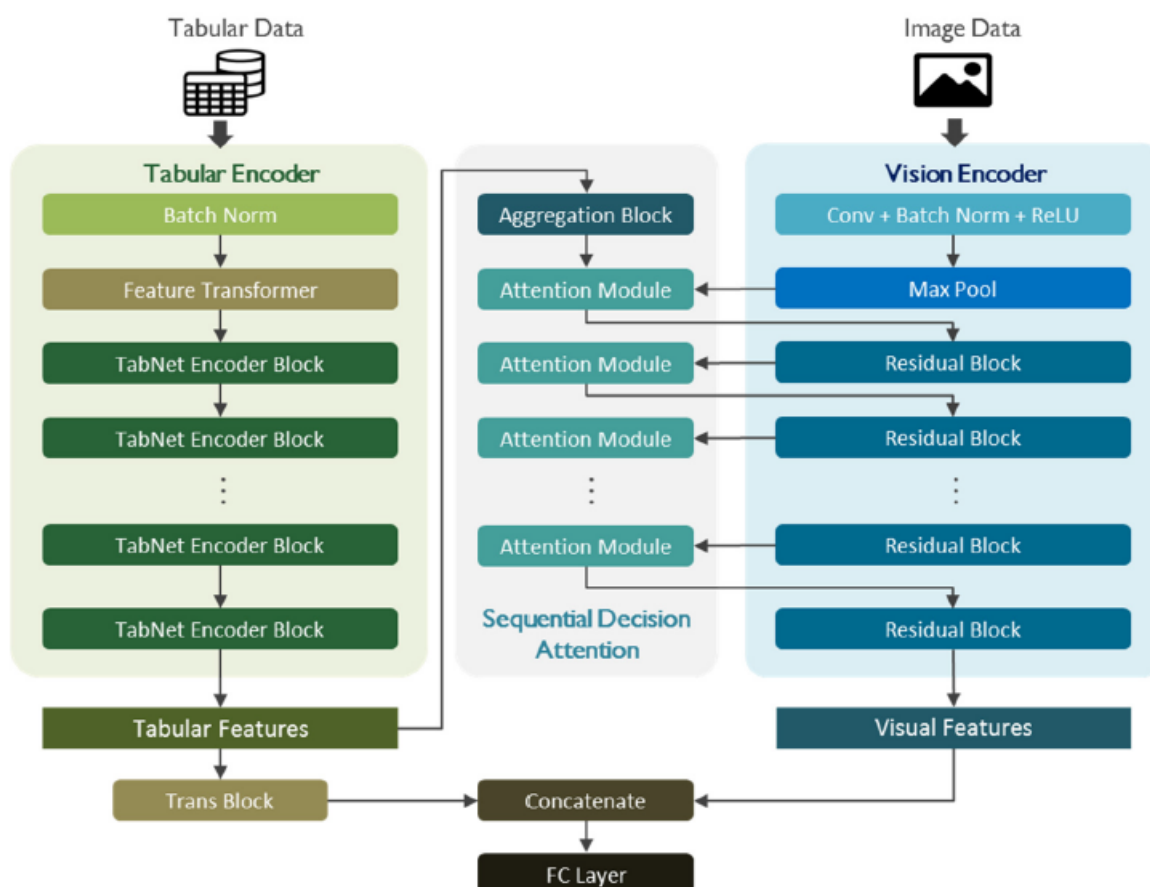


Σχήμα 4.7: Προτεινόμενο μοντέλο GCN [16]

Σε δημοσίευση του 2022 οι Yi Liu et al. [17] προτείνουν μια πολυτροπική προσέγγιση μάθησης το TabVisionNet, το οποίο μοντελοποιείται χρησιμοποιώντας τις πληροφορίες τόσο από δεδομένα πίνακα όσο και από δεδομένα εικόνας. Στο TabVisionNet ενσωματώθηκε ένας νέος μηχανισμός προσοχής που ονομάζεται Sequential Decision Attention. Η αρχιτεκτονική του συστήματος φαίνεται στο Σχήμα 4.8 και περιλαμβάνει τα εξής τρία κύρια στοιχεία :

- Κωδικοποιητή TabNet που χρησιμοποιείται ως κωδικοποιητής πίνακα για την εξαγωγή χαρακτηριστικών από τα δεδομένα πίνακα
- Συνελκτικά νευρωνικά δίκτυα (CNNs) που χρησιμοποιούνται ως κωδικοποιητές για την εξαγωγή των οπτικών χαρακτηριστικών από τις εικόνες.
- Μηχανισμός προσοχής για τον έλεγχο του κωδικοποιητή όρασης ώστε να εστιάζει στα κατάλληλα οπτικά χαρακτηριστικά από τις δεδομένες γενικές πληροφορίες περιβάλλοντος που εξάγονται τόσο από την εικόνα όσο και από τα δεδομένα πίνακα.

Ως συνάρτηση σφάλματος για την εκπαίδευση των μοντέλων χρησιμοποιήθηκε η  $L_{total} = L_{cls} + \beta \cdot L_{sparse}$ , όπου  $L_{cls}$  η συνάρτηση Cross Entropy Loss για τη δυαδική ταξινόμηση και  $L_{sparse}$  η ποινή κανονικοποίησης (regularization penalty) για τον έλεγχο της αραιότητας των επιλεγμένων χαρακτηριστικών στο TabNet. Το βασικό συμπέρασμα της δημοσίευσης ήταν ότι ο προτεινόμενος μηχανισμός προσοχής καταγράφει τις πολύπλοκες σχέσεις μεταξύ των δύο modalities με μια σύνθετη στρατηγική μοντελοποίησης, η οποία παράγει καλύτερες κοινές αναπαραστάσεις προς αξιοποίηση από το πολυτροπικό μοντέλο. Αν και η δημοσίευση εφαρμόζει το μοντέλο σε δεδομένα της βιομηχανίας κατασκευής οθονών θα μπορούσε να χρησιμοποιηθεί και σε βιοϊατρικές εφαρμογές. Ένας μηχανισμός προσοχής για την κατάτμηση των αθηρωματικών πλακών στις εικόνες υπερήχων περιγράφεται και στο [58].



Σχήμα 4.8: Προτεινόμενη αρχιτεκτονική με μηχανισμό προσοχής στο [17]

Τέλος, στα [59, 60, 61] περιγράφονται εφαρμογές της πολυτροπικής μάθησης στην ανίχνευση της πνευμονικής εμβολής, στην πρόβλεψη της έκβασης των εγκεφαλικών και στην κατηγοριοποίηση του καρκίνου το μαστού από MRI αντίστοιχα. Από αυτές τις τρεις εργασίες αντλούμε αρκετά στοιχεία για τη δική μας εφαρμογή.

### 4.3.2 Τεχνητή Νοημοσύνη στην ιατρική απεικόνιση και στα καρδιαγγειακά νοσήματα

Σε αυτό το κομμάτι της εργασίας θα αναφερθούμε περιεκτικά σε εργασίες που καταδεικνύουν την ποικιλότητα εφαρμογή και συνεισφορά της τεχνητή νοημοσύνη στην περαιτέρω αξιοποίηση των απεικονιστικών και άλλων ειδών δεδομένων για την υποστήριξη των ιατρικών αποφάσεων.

Στην [62] περιλαμβάνεται μία περιεκτική ανασκόπηση των ΑΙ μοντέλων για την ταξινόμηση ιατρικών εικόνων. Δοκιμάστηκαν σε ακτινογραφίες θώρακος για την ανίχνευση της πνευμονίας ένα μοντέλο SVM, δύο μοντέλα CNN (VGG16 και Inception V3) και ένα μοντέλο CapsuleNet. Μετά τις κατάλληλες προσαρμογές και το Transfer Learning το VGG16 πέτυχε τα καλύτερα αποτελέσματα (ακρίβεια >92%). Στην [63] γίνεται ανασκόπηση της πρόσφατης βιβλιογραφίας σε πάνω από 120 ερευνητικές εργασίες ιατρικής απεικόνισης και στην [64] εξετάζονται οι κύριες έννοιες βαθιάς μάθησης που σχετίζονται με την ανάλυση της ιατρικής εικόνας και συνοψίζονται περισσότερες από 300 συνεισφορές στον τομέα.

Στην εργασία [65] παρουσιάζονται οι διαφορές των CapsNets με τα CNN και αξιολογείται η συμπεριφορά τους στην ταξινόμηση ιατρικών εικόνων όπου συνήθως τα δεδομένα είναι λίγα και μη-ισορροπημένα.

Στην εργασία των Saba et Al. [8] γίνεται εκτεταμένη ανασκόπηση μοντέλων τεχνητής νοημοσύνης για τον πολυτροπικό χαρακτηρισμό ιστού και την ταξινόμηση της καρωτιδικής αρτηρίας. Πάνω από το 95% των μελετών υιοθέτησαν 2-D απεικόνιση κατά τη διάρκεια του σχεδιασμού του μοντέλου ΑΙ. Η μελέτη ταξινόμησε τα μοντέλα τεχνητής νοημοσύνης σε πέντε κατηγορίες: χειροκίνητη, στατιστική, μηχανική μάθηση, βαθιά μάθηση και μάθηση μεταφοράς. Στη βαθιά μάθηση η αρχιτεκτονική του CNN υιοθετήθηκε περισσότερο για ταξινόμηση και το U-Net για τμηματοποίηση. Το VGG16 ήταν το πιο δημοφιλές στην μεταφορά μάθηση.

Η [42] είναι η πρώτη πολυκεντρική μελέτη για την κατηγοριοποίηση της καρωτιδικής αθηρωματικής πλάκας από εικόνες US. Χρησιμοποιήθηκαν δεδομένα από Λονδίνο, Λισαβόνα και συνδυαστικά. Δοκιμάστηκαν με τις κατάλληλες προσαρμογές και βελτιστοποιήσεις 6 είδη μοντέλων ΑΙ, 4 μοντέλα ML (k-NN, SVM, Decision Tree και Random Forest), ένα σύστημα TL (VGG-19) και ένα σύστημα DL (DCNN με 5, 7,9,11,15 και 19 στρώματα).

Στην [66] για την εκτίμηση του κινδύνου πρόκλησης εγκεφαλικού επεισοδίου από αθηρωματική πλάκα στην καρωτίδα χρησιμοποιήθηκαν 361 εικόνες από US. Δοκιμάστηκαν διαφορετικά μοντέλα Μηχανικής Μάθησης (CART decision tree, Random Forest και Logistic Regression), Βαθιάς Μάθησης (CNN) και Μεταφοράς Μάθησης (Mobilenet και Capsulenet). Το Random Forest έδωσε τα καλύτερα αποτελέσματα (ακρίβεια 91,41%). Το μοντέλο CNN εξαιτίας των λίγων δεδομένων έδωσε καλά αποτελέσματα στην εκπαίδευση αλλά όχι στην επικύρωση. Με τη διαδικασία της μεταφερόμενης μάθησης το μοντέλο Capsulenet πέτυχε ακρίβεια 95,7%.

Οι Kang MK et al. στο [67] επεξεργάστηκαν με τη μέθοδο της ανάλυσης παλινδρόμησης Cox μίας μεταβλητής πολυτροπικά δεδομένα (κλινικές παράμετροι, συμβατικοί δείκτες κινδύνου όπως ο Framingham, υπερηχογραφικές παράμετροι όπως υπερηχογράφημα καρωτίδας διπλής όψεως κ.α.) για 150 ασθενείς. Το συμπέρασμα ήταν ότι η εφαρμογή των



δεικτών κινδύνου FSRS και ASCVD είναι εφικτή σε ασθενείς με ασυμπτωματική καρωτιδική νόσο για την πρόβλεψη του ισχαιμικού εγκεφαλικού και της στεφανιαίας νόσου και επιπλέον η προσθήκη υπερηχογραφικών παραμέτρων βελτιώνει ικανοποιητικά την προγνωστική δύναμη των συμβατικών δεικτών.

Στο εργαστήριο Βιοϊατρικών Προσομοιώσεων και Απεικονιστικής Τεχνολογίας του Εθνικού Μετσοβίου Πολυτεχνείου έχει γίνει σημαντική έρευνα στον τομέα της ανάλυσης και επεξεργασίας ιατρικής εικόνας με τεχνικές Μηχανικής Μάθησης. Στο [68] οι Dalakleidi Kalliopi et al. μελετούν τις διαφορές στις ταχύτητες και επιταχύνσεις της αθηρωματικής πλάκας με βάση τους υπερήχους μεταξύ ασυμπτωματικών ασθενών με μέτρια και σοβαρή στένωση. Στο [69] γίνεται μια έρευνα των μεθοδολογιών που χρησιμοποιούνται για την εκτίμηση της κίνησης του καρδιαγγειακού ιστού από αλληλουχίες εικόνων υπερήχων. Στο [70] γίνεται μελέτη της ηχογένεια υπερηχογραφικών εικόνων B-mode της καρωτίδας, στο [71] εκτίμηση με fractals της διάστασης των αθηρωματικών πλακών καρωτίδας, στο [72] διάγνωση της αθηροσκλήρωσης με βάση εικόνες υπερήχων και νευρωνικά δίκτυα, στα [73, 74, 75] πραγματοποιείται ταξινόμηση υψής της καρωτιδικής αθηροσκλήρωσης, στα [76, 77, 78, 79, 80, 81, 82] αναλύονται τεχνικές μελέτης της κίνησης των αρτηριών και εξάγονται συμπεράσματα από αυτήν. Εφαρμογές της τεχνητής νοημοσύνης στην καρωτιδική νόσο και στην ιατρική απεικόνιση περιγράφονται επίσης στα [83, 84, 85] και στα [86, 87] αναδύκενυεται η σημασία των EHR στη βελτίωση της προβλεπτικής ικανότητας των μοντέλων ML.

Τέλος, στο [88] γίνεται μια συστηματική έρευνα για τα χρησιμοποιούμενα μοντέλα βαθιάς μάθησης στο τομέα του Multimodal Data Fusion.



**Μέρος **

**Πρακτικό Μέρος**

---



## Κεφάλαιο **5**

# Σύνολο Δεδομένων και Υλοποίηση

---

### 5.1 Στόχος της εργασίας

Στόχος της εργασίας είναι η δημιουργία ενός πολυτροπικού μοντέλου Βαθιάς Μάθησης (Multimodal Learning) το οποίο θα αξιοποιεί τις εικόνες υπερήχων καρωτίδας και τα κλινικά και εργαστηριακά δεδομένα ασθενών με καρωτιδική νόσο και θα κατατάσσει τους ασθενείς σε δύο κατηγορίες. Στην πρώτη κατηγορία ανήκουν οι ασθενείς για τους οποίους συνιστάται επεμβατική αντιμετώπιση της νόσου και οι οποίοι χαρακτηρίζονται ως Υψηλού Κινδύνου, ενώ η δεύτερη κατηγορία αφορά στους ασθενείς για τους οποίους προτείνεται η συντηρητική αντιμετώπιση και χαρακτηρίζονται ως Χαμηλού Κινδύνου.

### 5.2 Παρουσίαση Συνόλου Δεδομένων

Το σύνολο δεδομένων αποτελείται από υπερηχογραφήματα καρωτίδας και άλλα κλινικά και εργαστηριακά χαρακτηριστικά ασθενών με καρωτιδική νόσο. Για κάθε ασθενή γνωρίζουμε αν είναι συμπτωματικός ή ασυμπτωματικός σύμφωνα με τον ορισμό που δώσαμε στο Κεφάλαιο 2 όπου ως συμπτωματικός ορίζεται ο ασθενής ο οποίος έχει εμφανίσει συμπτώματα της καρωτιδικής νόσου τους τελευταίους 6 μήνες. Η διάκριση των ασθενών σε Υψηλού και Χαμηλού Κινδύνου γίνεται ως εξής:

- Στην κατηγορία **Υψηλού Κινδύνου** ανήκουν οι συμπτωματικοί ασθενείς με βαθμό στένωσης από 50% και πάνω και οι ασυμπτωματικοί ασθενείς με βαθμό στένωσης από 70% και άνω. Σε αυτή την περίπτωση κρίνεται απαραίτητη η επεμβατική αντιμετώπιση της νόσου με τις μεθόδους που περιγράψαμε στο Κεφάλαιο 2 σε συνδυασμό με κατάλληλη φαρμακευτική αγωγή.
- Στην κατηγορία **Χαμηλού Κινδύνου** ανήκουν όλες οι υπόλοιπες περιπτώσεις. Σε αυτήν την κατηγορία ασθενών δεν συνιστάται η επεμβατική αντιμετώπιση της νόσου καθώς μπορεί να αντιμετωπιστεί αποτελεσματικά αποκλειστικά με φαρμακευτική αγωγή.

Το σύνολο δεδομένων αποτελείται από 272 εικόνες, οι οποίες προέρχονται από υπερηχογραφικές απεικονιστικές καταγραφές β-σάρωσης (B-Mode Ultrasounds) της καρωτίδας 76 ασθενών, και τέσσερα (4) csv αρχεία με κλινικά και εργαστηριακά δεδομένα. Τα δεδομένα αντιστοιχούν σε 76 ασθενείς με καρωτιδική νόσο εκ των οποίων οι 65 έχουν κατηγοριοποι-

ηθεί στην κλάση Υψηλού Κινδύνου-High Risk (ετικέτα Risk=1, ποσοστό 85.52%) και οι 11 στην κλάση Χαμηλού Κινδύνου-Low Risk (Risk=0, ποσοστό 14.47%).

Σε κάθε έναν ασθενή αντιστοιχεί μία σειρά σε κάθε ένα csv αρχείο. Το πρώτο csv αρχείο με τίτλο “blood.csv” περιέχει τις τιμές των αιματολογικών εξετάσεων των ασθενών. Το δεύτερο csv με τίτλο “clinical.csv” περιέχει δεδομένα όπως η ηλικία, το φύλο, το αν ο/η ασθενής είναι καπνιστής/καπνίστρια και το αν έχει διαβήτη. Το τρίτο csv με τίτλο “protein\_markers.csv” περιέχει τις τιμές κάποιων πρωτεϊνικών δεικτών, όπως για παράδειγμα των τιμών ινσουλίνης και ινωδογόνου στο αίμα, που προέρχονται από αναλύσεις του αίματος. Το τέταρτο csv με τίτλο “images.csv” περιέχει την αντιστοίχιση των εικόνων των πλακών με τον ασθενή και τον αντίστοιχο βαθμό στένωσης. Τα δεδομένα προέρχονται από τη βάση δεδομένων του ερευνητικού έργου με κωδικό 09ΣΥΝ-12-1054 και τίτλο «Η αθηρωματική πλάκα στην καρωτίδα: μία πολυεπιστημονική προσέγγιση με στόχο τη βέλτιστη προσέγγιση συμπτωματικών και ασυμπτωματικών ασθενών» που χρηματοδοτήθηκε από το πρόγραμμα «Ανταγωνιστικότητα και Επιχειρηματικότητα», ΕΣΠΑ 2007-2013. Τα δεδομένα αφορούν σε ασθενείς που υποβλήθηκαν σε εξετάσεις στο Αγγειοχειρουργικό Τμήμα του Πανεπιστημιακού Γενικού Νοσοκομείου Αθηνών «ΑΤΤΙΚΟΝ» στο διάστημα 2012-2015. Στην παρούσα εργασία δεν θα χρησιμοποιηθούν τα δεδομένα του αρχείου “protein\_markers.csv”. Επίσης, για τον ασθενή με ID 155, που ανήκει στους High Risk, δεν διαθέτουμε δεδομένα για τις αιματολογικές εξετάσεις του και άρα θα εργαστούμε με τους 75 από τους 76 ασθενείς.

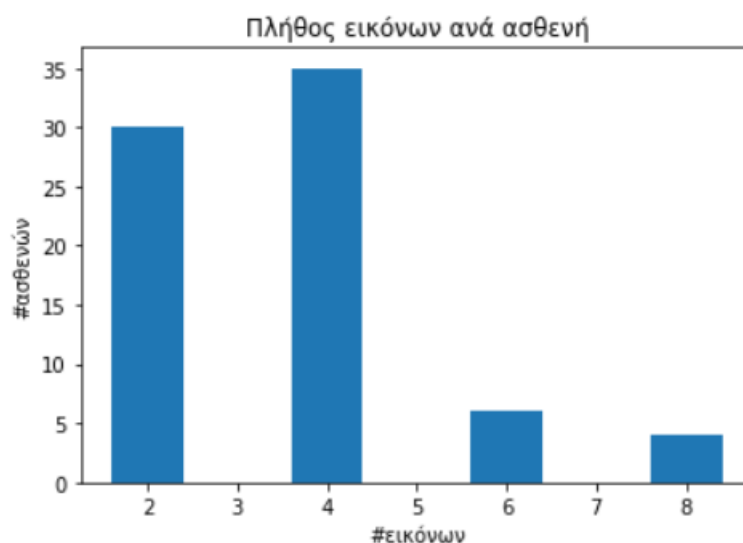
### 5.2.1 Το υποσύνολο των εικόνων

Οι εικόνες του συνόλου δεδομένων αποτελούν στιγμιότυπα (frames) των υπερηχογραφικών βίντεο. Τα βίντεο απεικονίζουν τις καρωτίδες. Έπειτα από επεξεργασία των βίντεο επιλέχθηκαν δύο (2) frames από κάθε βίντεο το ένα εκ των οποίων αντιστοιχεί στη συστολή της αρτηρίας και το άλλο στη διαστολή. Επιπλέον, ορισμένοι ασθενείς εμφανίζουν περισσότερες από μία αθηρωματικές πλάκες στις καρωτιδικές αρτηρίες και συνεπώς σε κάποιους ασθενείς αντιστοιχούν περισσότερα του ενός βίντεο. Στο διάγραμμα φαίνεται ο αριθμός των εικόνων ανά ασθενή ο οποίος είναι πάντα ζυγός αριθμός καθώς από κάθε πλάκα κρατήσαμε ένα frame από τη διαστολή και ένα από τη συστολή της αρτηρίας. Συνεπώς οι 272 εικόνες προέρχονται από 136 βίντεο υπερηχογραφήματων καρωτίδων 76 ασθενών. (Σχήμα 5.1).

Από τα frames με τεχνικές κατάτμησης εικόνων έχουν εξαχθεί τα τμήματα της εικόνας που αντιστοιχούν στις αθηρωματικές πλάκες. Οι τελικές εικόνες, οι οποίες χρησιμοποιούνται σαν είσοδο στα μοντέλα, απεικονίζουν μόνο τις αθηρωματικές πλάκες, είναι grayscale στην κλίμακα [0,255] και διαστάσεων 120X334. Στην Εικόνα 5.2 φαίνεται μία από αυτές τις εικόνες. Οι ρυθμίσεις της διάταξης υπερήχων που χρησιμοποιείται επηρεάζουν τα χαρακτηριστικά και την ποιότητα των υπερηχογραφικών απεικονιστικών καταγραφών β-σάρωσης. Σε όλες τις καταγραφές του συνόλου δεδομένων έχουν χρησιμοποιηθεί οι ίδιες ρυθμίσεις οι οποίες προβλέπονται από τα σχετικά πρωτόκολλα και παρουσιάζονται στον Πίνακα 5.1.

### 5.2.2 Το υποσύνολο των κλινικών και βιοχημικών δεδομένων

Από τα διαθέσιμα κλινικά και βιοχημικά δεδομένα για κάθε ασθενή επιλέξαμε να χρησιμοποιήσουμε κάποια από αυτά για τα οποία από τη βιβλιογραφία προέκυψε ότι είναι πιο



Σχήμα 5.1: Πλήθος εικόνων ανά ασθενή



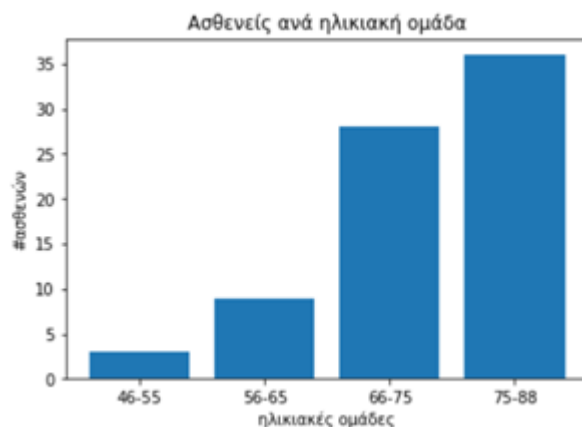
Σχήμα 5.2: Εικόνες αθηρωματικής πλάκας όπως προέκυψαν από τα frames ενός B-Mode Ultrasound βίντεο (αριστερά η εικόνα κατά τη συστολή της αρτηρίας και δεξιά η εικόνα της ίδιας πλάκας κατά τη διαστολή)

Παράμετροι	Ρύθμιση
Κεφαλή	Γραμμική 3-12 MHz, General Electric
Είδος εικόνων	B-mode
Βάθος	3.5-4 cm
Persistence	0
Κέρδος	60 dB ή 75 dB
Χρονική ανάλυση	≥ 25καρνανδευτερλεπτο

Πίνακας 5.1: Ρυθμίσεις υπερηχογραφικής συσκευής.

στενά συνδεδεμένα με την καρωτιδική νόσο. Σε μελλοντική επέκταση της παρούσας εργασίας θα μπορούσαν να χρησιμοποιηθούν περισσότερα από τα διαθέσιμα δεδομένα και να μελετηθεί ο βαθμός που συνεισφέρουν στη διάκριση των ασθενών σε υψηλού και χαμηλού κινδύνου. Τα δεδομένα τα οποία επιλέξαμε να χρησιμοποιήσουμε είναι συνολικά 15. Παρακάτω περιγράφουμε κάθε ένα από αυτά τα χαρακτηριστικά ώστε να δώσουμε μια πλήρη εικόνα η οποία θα διευκολύνει την ερμηνεία των πειραματικών αποτελεσμάτων. Επίσης, για κάθε ένα από τα χαρακτηριστικά προσπαθούμε να δώσουμε τη σχέση του με την καρωτιδική νόσο. Σε αυτό φάνηκε πολύ χρήσιμο το πόνημα των Vinood B. Patel και Victor R. Preedy Biomarkers in Cardiovascular Disease (μτφρ. Βιοδείκτες στα καρδιαγγειακά νοσήματα) [89].

- **Ηλικία ασθενούς:** Ο νεότερος ασθενής είναι 46 ετών και ο μεγαλύτερος 88 ετών. Στο διάγραμμα (Σχήμα 5.3) παρουσιάζεται το πλήθος των ασθενών ανά ηλικιακή κατηγορία. Παρατηρούμε ότι οι περισσότεροι ασθενείς ανήκουν στην ηλικιακή ομάδα 75-88 ετών.



Σχήμα 5.3: Πλήθος ασθενών ανά ηλικιακή ομάδα

- **Φύλο ασθενούς:** Το σύνολο δεδομένων αποτελείται από 57 άνδρες και 19 γυναίκες.
- **Καπνιστής/Καπνίστρια**
- **Διαβήτης:** Το σύνολο δεδομένων περιέχει 30 ασθενείς διαγνωσμένους με σακχαρώδη διαβήτη και 46 χωρίς. Ο σακχαρώδης διαβήτης έχει συνδεθεί με επιτάχυνση της εξέλιξης της καρωτιδικής νόσου καθώς και με άλλες αγγειακές παθήσεις [90, 91, 92].
- **Αιμοπετάλια:** Τα αιμοπετάλια είναι κυτταρικά θραύσματα (δηλ. κύτταρα χωρίς πυρήνα). Ζουν πολύ λίγο, περίπου 5-9 μέρες, και αποτελούν βασικό συστατικό του αίματος καθώς χρησιμεύουν στην πήξη του και συνεπώς στην επούλωση των τραυμάτων. Όταν ο αριθμός τους είναι πολύ χαμηλός υπάρχουν αυξημένες πιθανότητες αιμορραγίας ενώ όταν είναι πολύ υψηλός μπορεί να σχηματιστούν θρόμβοι στο αίμα οι οποίοι να φράξουν αγγεία. Ο φυσιολογικός αριθμός αιμοπεταλίων κυμαίνεται από 150.000-450.000 αιμοπετάλια/μL αίματος. Επίσης, τα αιμοπετάλια διαδραματίζουν καθοριστικό ρόλο στη ρήξη της αθηρωματικής πλάκας ενώ σύγχρονες μελέτες υποστηρίζουν ότι τα αιμοπετάλια έχουν έναν φλεγμονώδη ρόλο ως κύτταρα του ανοσοποιητικού και συμμετέχουν στην ανάπτυξη της αθηροσκλήρωσης [93, 94, 95].
- **Ουρία:** Η ουρία παράγεται στο ήπαρ ως απόβλητο της πέψης. Από τη διάσπαση των πρωτεϊνών κατά την πέψη, στο ήπαρ παράγεται αμμωνία η οποία περιέχει άζωτο. Το άζωτο συνδυάζεται με άλλα στοιχεία όπως άνθρακα, υδρογόνο και οξυγόνο και σχηματίζει την ουρία. Η ουρία μέσω της κυκλοφορίας του αίματος καταλήγει στους νεφρούς όπου φιλτράρεται και τελικά αποβάλλεται από τον οργανισμό μέσω της ούρησης. Με την εξέταση αίματος μετράται το άζωτο της ουρίας του αίματος (BUN-Blood Urea Nitrogen) και υψηλές τιμές BUN αποτελούν ένδειξη πιθανής δυσλειτουργίας των νεφρών. Οι φυσιολογικές τιμές BUN κυμαίνονται στο διάστημα 6-24 mg/dL. Μελέτη έδειξε ότι



υψηλά επίπεδο BUN (13,51 mg/dL) σχετίζονται με αυξημένη πιθανότητα εμφάνισης καρδιακής ανεπάρκειας στις γυναίκες ηλικίας μεγαλύτερης των 60 ετών [96].

- **Κρεατινίνη:** Αποτελεί μεταβολικό προϊόν της διάσπασης της φωσφορικής κρεατίνης των μυών και αποβάλλεται μέσω των νεφρών ενώ ένα μικρό μέρος της απορροφάται ξανά από τον οργανισμό. Αυξημένα επίπεδα κρεατινίνης μπορεί να είναι σημάδι νεφρικής ανεπάρκειας. Οι φυσιολογικές τιμές της στο αίμα είναι 0,74-1,35 mg/dL (65,4-119,3 micromoles/L) για τους ενήλικες άντρες και 0,59-1,04 mg/dL (52,2-91,9 micromoles/L) για τις ενήλικες γυναίκες.
- **Οξαλοξική Τρανσαμινάση:** Η ασπαρτική αμινοτρανσφεράση (AST) ή αλλιώς οξαλοξική τρανσαμινάση (SGOT) είναι ένα ένζυμο που βρίσκεται κυρίως στην καρδιά, το συκώτι και τους μύες. Απελευθερώνεται στην κυκλοφορία του αίματος μετά από τραυματισμό ή θάνατο των κυττάρων. Τα επίπεδα της AST συνήθως αυξάνονται εντός 12 ωρών από τον τραυματισμό και παραμένουν αυξημένα για 5 ημέρες. Έτσι, η εξέταση αυτή είναι μία από τις πολλές που πραγματοποιούνται όταν υπάρχει βλάβη στο μυοκάρδιο, όπως στο έμφραγμα του μυοκαρδίου καθώς και κατά την αξιολόγηση ηπατικής βλάβης. Οι φυσιολογικές τιμές είναι 10-40 IU/L για τους άνδρες και 9-32 IU/L για τις γυναίκες.
- **Πυροσταφυλική τρανσαμινάση:** Η πυροσταφυλική τρανσαμινάση (SGPT) ή αλλιώς αμινοτρανσφεράση της αλανίνης (ALT) είναι ένα ένζυμο που βρίσκεται στους νεφρούς, την καρδιά και στους σκελετικούς μύες, αλλά κυρίως στο ήπαρ. Λειτουργεί ως καταλύτης στην αντίδραση για την παραγωγή αμινοξέων. Η SGPT αξιολογείται μαζί με την Οξαλοξική Τρανσαμινάση (SGOT/AST) για τον εντοπισμό πιθανής ηπατικής βλάβης. Οι τιμές δύο τρανσαμινάσες συνυπάρχουν συνήθως σε αναλογία περίπου 1:1. Οι φυσιολογικές τιμές είναι 10-55 IU/L για τους άνδρες και 7-30 IU/L για τις γυναίκες.
- **γ-Γλουταμυλοτρανσφεράση (γ-GT):** Είναι ένζυμο που βρίσκεται κυρίως στα χοληφόρα και στο ήπαρ και σε μικρότερο βαθμό στην καρδιά, στους νεφρούς, στο πάγκρεας, στον εγκέφαλο, στους σιελογόνους αδένες, στον προστάτη και στον σπλήνα. Η λειτουργία της γ-γλουταμυλοτρανσφεράσης έχει σχέση με τη μεταφορά των αμινοξέων διαμέσου των κυτταρικών μεμβρανών. Η γGT συχνά μετράται σε συνδυασμό με την αλκαλική φωσφατάση (ALP) για να προσδιοριστεί αν η ALP αυξάνεται λόγω ηπατικής νόσου καθώς ενώ η ALP αυξάνεται είτε σε διαταραχές του ήπατος και της χολής είτε στον οστών, η γGT αυξάνεται μόνο στις διαταραχές του ήπατος και της χολής. Οι φυσιολογικές τιμές της γ-GT διαφέρουν από εργαστήριο σε εργαστήριο ανάλογα με το αντιδραστήριο που χρησιμοποιείται. Ενδεικτικά, οι φυσιολογικές τιμές κυμαίνονται από 5 έως 40 IU/L. Επιδημιολογικές και παθολογικές μελέτες έχουν εξερευνήσει τον ρόλο της στην παθογένεση και την κλινική εξέλιξη των καρδιαγγειακών παθήσεων που προκαλούνται από την αθηροσκλήρωση [97].
- **Αλκαλική Φωσφατάση (ALP):** Είναι ένα ένζυμο που εντοπίζεται στο ήπαρ, τα οστά, τον πλάκουντα, το έντερο και τα νεφρά, αλλά κυρίως στα επιθηλιακά κύτταρα των χοληφόρων σωληναρίων καθώς και στους οστεοβλάστες που συμμετέχουν στο σχηματισμό

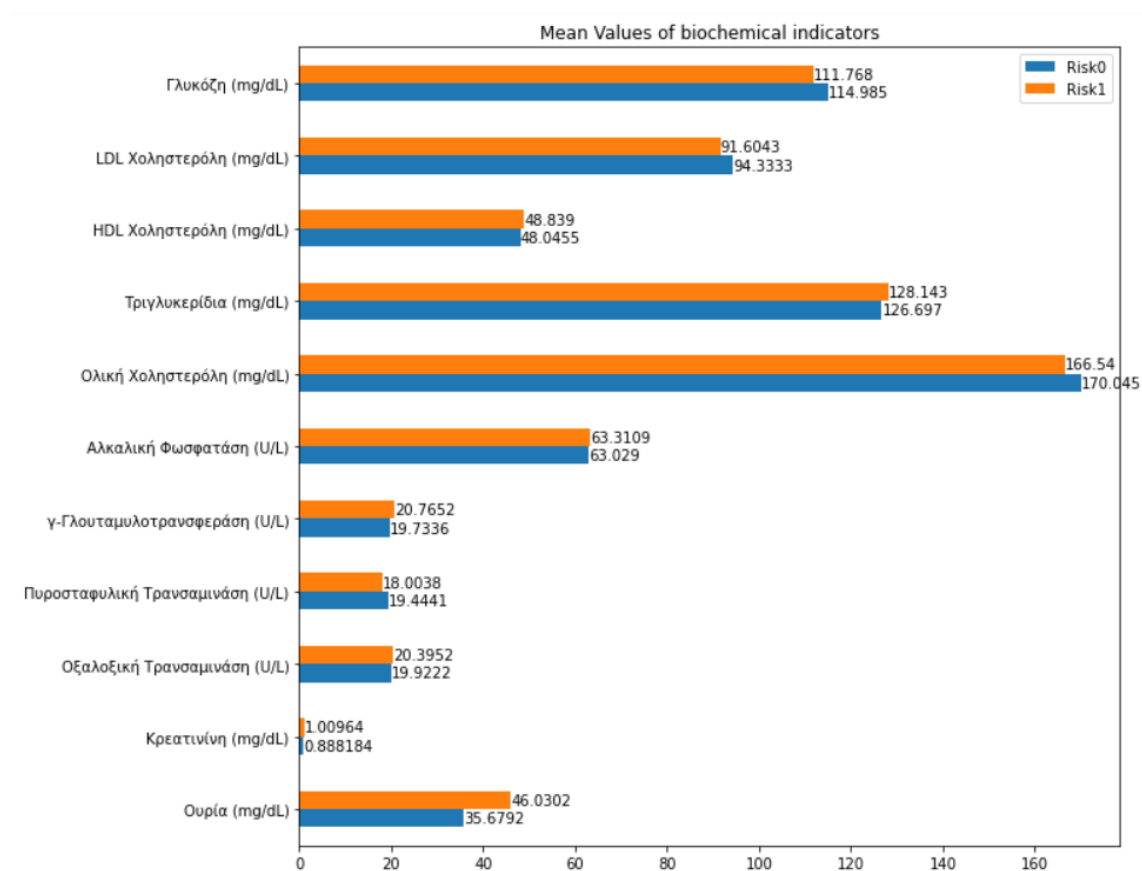
νέου οστού. Η αλκαλική φωσφατάση φυσιολογικά εκκρίνεται από το ήπαρ στη χολή. Αυξημένα επίπεδα αλκαλικής φωσφατάσης παρατηρούνται τις περιόδους ανάπτυξης των οστών και επίσης αποτελούν δείκτη διαφόρων νόσων του ήπατος και των χοληφόρων. Θεωρείται επίσης δείκτης νεοπλασματικής δραστηριότητας επειδή αυξάνεται στην περίπτωση οστεοσαρκώματος καθώς και στον καρκίνο του μαστού ή του προστάτη με μεταστάσεις στα οστά.

- **Ολική Χοληστερόλη:** Η χοληστερόλη μαζί με τα τριγλυκερίδια αποτελεί δομικό στοιχείο των κυτταρικών μεμβρανών. Οι διαμοριακές δυνάμεις που αναπτύσσονται μεταξύ των μορίων της χοληστερόλης και του νερού που βρίσκεται εντός και εκτός κυττάρου συμβάλλει στη ρευστή αλλά αδιάσπαστη δομή της μεμβράνης, και της προσδίδει λειτουργικότητα. Επιπλέον, έχει καθοριστικό ρόλο στην ανάπτυξη των νευρικών συνδέσμων του εγκεφάλου και συμβάλλει στο σχηματισμό και την προστασία των νευρώνων και στη δημιουργία των ιστών. Δεσμεύεται με πρωτεΐνες σχηματίζοντας τις λιποπρωτεΐνες και βρίσκεται σε μεγαλύτερη συγκέντρωση σε ιστούς που είτε την παράγουν είτε έχουν πυκνές μεμβράνες, για παράδειγμα στο συκώτι, τη σπονδυλική στήλη και τον εγκέφαλο, όπως επίσης και στα αθηρώματα. Υπολογίζεται προσθέτοντας την HDL, την LDL και 20% των τριγλυκεριδίων. Η τιμή της δεν πρέπει να ξεπερνά τα 200 mg/dL.
- **Τριγλυκερίδια:** Αποτελούν ένα είδος λιπιδίου του αίματος όπως και οι χοληστερόλη. Η διαφορά τους είναι ότι τα τριγλυκερίδια τροφοδοτούν τον οργανισμό με ενέργεια, ενώ η χοληστερόλη χρησιμοποιείται στη δομή των κυττάρων και για την παραγωγή ορισμένων ορμονών. Οι φυσιολογικές τιμές τριγλυκεριδίων είναι οι μικρότερες από 150 mg/dL. Έχει δειχθεί ότι τα τριγλυκερίδια και οι μεταβολίτες τους μπορούν να προάγουν την αθηροσκλήρωση μέσω της πυροδότησης φλεγμονής, του οξειδωτικού στρες και του σχηματισμού foam cells [98].
- **HDL (High-density lipoprotein) Χοληστερόλη:** Είναι γνωστή ως «καλή» χοληστερόλη και συλλέγει την περίσσεια χοληστερόλης στο αίμα την μεταφέρει πίσω στο συκώτι όπου διασπάται και τελικά απομακρύνεται από το σώμα. Οι επιθυμητές τιμές στον άνθρωπο είναι οι μεγαλύτερες από 60 mg/dL για τους άνδρες και για τις γυναίκες ενώ τιμές μικρότερες των μικρότερες των 40 mg/dL για τους άνδρες και μικρότερες 50 ενμγ/δλ για τις γυναίκες αποτελούν το κατώφλι κινδύνου.
- **LDL (Low-density lipoprotein) Χοληστερόλη:** Είναι γνωστή ως «κακή» χοληστερόλη και αποτελεί μία από τις πέντε κύριες ομάδες λιποπρωτεϊνών που μεταφέρουν όλα τα μόρια λίπους σε όλο το σώμα. Επίσης αποτελεί το μεγαλύτερο ποσοστό χοληστερόλης στον ανθρώπινο οργανισμό. Όπως είδαμε και στον μηχανισμό αθηρογένεσης η LDL χοληστερόλη διαδραματίζει καθοριστικό ρόλο στον σχηματισμό των αθηρωματικών πλακών. Οι ιδανικές τιμές για τον οργανισμό είναι οι μικρότερες των 100 mg/dL, οι τιμές στο διάστημα 100-129 mg/dL θεωρούνται ανεκτές και οι τιμές 130-150 mg/dL είναι το όριο των υψηλών.
- **Γλυκόζη:** Είναι ένας μονοσακχαρίτης και σχηματίζεται κατά την πέψη των υδατανθράκων και τη μετατροπή του γλυκογόνου από το ήπαρ. Αποτελεί την κύρια πηγή

κυτταρικής ενέργειας στον οργανισμό και είναι απαραίτητη για τη λειτουργία του εγκεφάλου και των ερυθροκυττάρων. Η περίσσεια γλυκόζης αποθηκεύεται ως γλυκογόνο στο ήπαρ και στα μυϊκά κύτταρα. Οι φυσιολογικές τιμές της γλυκόζης καθορίζονται στις 80 έως 130 mg/dL (4,5 έως 7,0 mmol/l) αμέσως μετά την πρωινή έγερση και πριν από τα γεύματα και στις μικρότερες από 180 mg/dL (κάτω από 10,0 mmol/l) στις δύο ώρες μετά από γεύμα.

Επίσης, αξίζει να σημειωθεί ότι από τους 76 ασθενείς οι 44 λάμβαναν το φάρμακο Ezetimibe κατά την περίοδο λήψης των δειγμάτων τους. Το φάρμακο αυτό χρησιμοποιείται για τη μείωση των επιπέδων της χοληστερόλης στο αίμα, η οποία όπως είδαμε αποτελεί σημαντικό παράγοντα της καρωτιδικής νόσου, και στη θεραπεία ορισμένων άλλων ανωμαλιών των λιπιδίων.

Στο παρακάτω διάγραμμα (Σχήμα 5.4) παρουσιάζουμε για κάθε κατηγορία ασθενών τον μέσο όρο των τιμών των 11 υπό μελέτη βιοχημικών δεικτών έχοντας παραλείψει τα αιμοπετάλια για λόγους ευκρίνειας του διαγράμματος. Η μέση τιμή των αιμοπεταλίων είναι 186,39 K/μL για τους ασθενείς Χαμηλού Κινδύνου και 1356,85 K/μL για τους ασθενείς Υψηλού Κινδύνου.



Σχήμα 5.4: Διάγραμμα μέσω τιμών των 11 αιματολογικών χαρακτηριστικών για κάθε κατηγορία ασθενών

## 5.3 Υλοποίηση

### 5.3.1 Περιβάλλον ανάπτυξης

Χρησιμοποιήθηκε το Colab το οποίο είναι περιβάλλον για ανάπτυξη σημειωματάρων Jupyter (Jupyter Notebooks) το οποίο εκτελείται στο cloud. Διατίθεται από τη Google και υποστηρίζει πολλές δημοφιλείς βιβλιοθήκες μηχανικής μάθησης που μπορούν εύκολα να φορτωθούν στα σημειωματάρια. Παρέχει επίσης τη δυνατότητα χρήσης GPU η οποία κρίνεται απαραίτητη στις περισσότερες εφαρμογές της μηχανικής μάθησης. Στις περισσότερες περιπτώσεις της δωρεάν έκδοσης του Colab χρησιμοποιούνται K80 GPUs και RAM των 12 GB. Κατά την ανάπτυξη του κώδικα της παρούσας εργασίας χρησιμοποιήθηκε κυρίως η βιβλιοθήκη Pytorch καθώς και βοηθητικές συναρτήσεις από τις Tensorflow και Sckit-Learn.

### 5.3.2 Μεθοδολογία εκπαίδευσης και αξιολόγησης των μοντέλων

Για την εκπαίδευση (train) και την αξιολόγηση (evaluation) των μοντέλων που θα παρουσιάσουμε στη συνέχεια απαιτείται η διαμέριση του συνόλου των δεδομένων σε τρία μέρη:

- σύνολο εκπαίδευσης (train set)
- σύνολο επικύρωσης (validation set)
- σύνολο ελέγχου (test set)

Το σύνολο των δεδομένων μας όπως έχουμε δει αποτελείται από 64 ασθενείς Υψηλού κινδύνου και 11 Χαμηλού συνεπώς είναι έντονα μη ισορροπημένο με κλάση μειονότητας την Χαμηλού κινδύνου. Η έντονη αυτή ανισορροπία των δεδομένων θα μας απασχολήσει πολύ κατά τη διαδικασία ανάπτυξης των μοντέλων καθώς επηρεάζει την ποιότητά της εκπαίδευσής τους. Λόγω της ανισορροπίας μια απλή διαμέριση των δεδομένων σε σύνολα εκπαίδευσης, επικύρωσης και ελέγχου, με έναν κανόνα 70%-30% για παράδειγμα, οδηγεί σε μη επαρκή αξιολόγηση των μοντέλων. Συνεπώς ακολουθήσαμε μία πιο σύνθετη διαδικασία εκπαίδευσης και αξιολόγησης την **Εμφωλευμένη Μέθοδο Διασταυρούμενης Επικύρωσης (Nested Cross-Validation)** η οποία περιγράφεται στη συνέχεια.

Για την Εμφωλευμένη Μέθοδο Διασταυρούμενης Επικύρωσης συνδυάζουμε 3-Fold Cross Validation στο train-test split και 2-Fold Cross Validation στο train-validation split. Η επιλογή του πλήθους των folds έγινε με αυτόν τον τρόπο ώστε σε κάθε fold να υπάρχουν περισσότερα από ένα δείγματα της κλάσης μειονότητας. Επιπλέον, είναι σημαντικό να διατηρήσουμε την κατανομή των δειγμάτων του αρχικού συνόλου σε κάθε fold. Συνεπώς σε κάθε στάδιο διαμέρισης των δεδομένων επιλέχθηκε συγκεκριμένος αριθμός δειγμάτων από κάθε κλάση ώστε κάθε fold να διαθέτει ίδιο ή σχεδόν ίδιο πλήθος δειγμάτων κάθε κατηγορίας ασθενών (Stratified k-Fold Cross Validation).

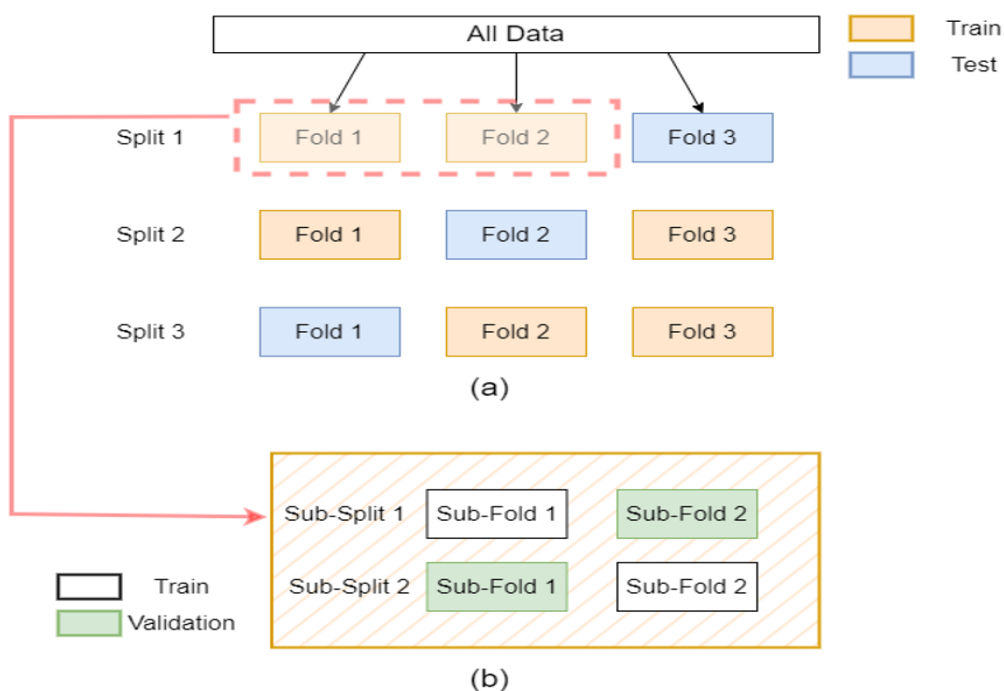
Επίσης, η **διαμέριση των δειγμάτων** στα σύνολα εκπαίδευσης, επικύρωσης και ελέγχου έγινε **με βάση τους ασθενείς** κι όχι τις εικόνες αφενός ώστε να αποφύγουμε εικόνες από τις πλάκες του ίδιου ασθενούς να βρίσκονται σε διαφορετικά σύνολα πράγμα που θα σήμαινε τη διαρροή πληροφορίας από το σύνολο ελέγχου στο σύνολο εκπαίδευσης και αφετέρου ώστε να συμβαδίζει με το στόχο της παρούσας εργασίας που είναι η κατηγοριοποίηση των ασθενών

και όχι απλώς των εικόνων. Επομένως για 3-Fold Cross Validation στο σύνολο εκπαίδευσης και ελέγχου με τους 75 ασθενείς η διαμέριση φαίνεται στο Σχήμα 5.14. Το μπλε πλαίσιο αντιστοιχεί στο σύνολο που θα χρησιμοποιηθεί για έλεγχο (test) και τα κόκκινα αντιστοιχούν στο σύνολο εκπαίδευσης και επικύρωσης.



Σχήμα 5.5: 3-Fold Cross Validation στο αρχικό σύνολο

Για τη διαμέριση του συνόλου εκπαίδευσης και επικύρωσης εφαρμόζουμε 2-Fold Cross Validation όπως φαίνεται στο Σχήμα 5.6.



Σχήμα 5.6: Εκπαίδευση και αξιολόγηση με Nested Cross Validation

Να σημειώσουμε ότι όποιες τεχνικές επαύξησης των δεδομένων επιλέχθηκαν, εφαρμόστηκαν αποκλειστικά και μόνο στο σύνολο εκπαίδευσης. Ούτε στο σύνολο επικύρωσης, ούτε στο

σύνολο ελέγχου. Για την επιλογή των βέλτιστων υπερπαραμέτρων κάθε μοντέλο χρησιμοποιήθηκε η μέθοδος GridSearch. Ο αλγόριθμος εκπαίδευσης και αξιολόγησης των μοντέλων για την επιλογή των βέλτιστων υπερπαραμέτρων περιγράφεται στη συνέχεια.

1. Αρχικά χωρίζουμε το συνολικό σύνολο δεδομένων σε 3 folds ώστε να εφαρμόσουμε 3-Fold Cross Validation όπως φαίνεται στην Σχήμα 5.14 και περιγράψαμε παραπάνω.
2. Για κάθε split  $i$ :
  - Κάθε training set το χωρίζουμε σε 2 subfolds όπως φαίνεται και στο (Σχήμα 5.6). Τα λευκά κουτιά αφορούν στα folds που θα χρησιμοποιηθούν για training και τα πράσινα σε αυτά που θα χρησιμοποιηθούν για validation.
  - Για κάθε ένα από τα υποψήφια μοντέλα και εκπαιδεύουμε και αξιολογούμε στα κατάλληλα subfolds διάφορες παραλλαγές του. Οι παραλλαγές προκύπτουν από τη μεταβολή κάποιων παραμέτρων του μοντέλου. Για κάθε παραλλαγή κρατάμε το μέσο όρο των μετρικών ενδιαφέροντος από το evaluation στο subfold1 και στο subfold0. Στο τέλος κρατάμε την παραλλαγή με τον υψηλότερο μέσο όρο μετρικών.
  - Για κάθε μοντέλο τρέχουμε στο test fold του split  $i$  και σημειώνουμε τις τιμές των μετρικών της παραλλαγής με το βέλτιστο συνδυασμό παραμέτρων όπως προέκυψε από το προηγούμενο βήμα.
3. Από το προηγούμενο βήμα για κάθε split  $i$  έχουμε από έναν καλύτερο συνδυασμό παραμέτρων για κάθε μοντέλο που δοκιμάσαμε καθώς και την επίδοσή του στα test folds. Επιλέγουμε μεταξύ των μοντέλων με βάση τη μέση τιμή και την τυπική απόκλιση των μετρικών κάθε μοντέλου στα τρία splits.

Επιπλέον, κατά την εκπαίδευση όλων των μοντέλων βαθιάς μάθησης πραγματοποιείται έλεγχος **Πρόωρου Τερματισμού (Early Stopping)**. Ο έλεγχος Early Stopping σε κάθε εποχή συγκρίνει το validation loss με το καλύτερο (μικρότερο) validation loss έως εκείνη τη στιγμή. Σε περίπτωση που το νέο validation loss είναι καλύτερο από το τρέχον καλύτερο αποθηκεύεται με checkpoint η νέα κατάσταση του μοντέλου. Αν για συγκεκριμένο αριθμό, οριζόμενο από τον χρήστη, συνεχόμενων εποχών το validation loss δεν βελτιώνεται η εκπαίδευση σταματά και στο μοντέλο φορτώνεται η κατάσταση του τελευταίου checkpoint δηλαδή οι βέλτιστες τιμές των παραμέτρων με βάση το υπό εξέταση validation set. Το Early Stopping αποτρέπει την υπερεκπαίδευση του δικτύου η οποία περιορίζει την ικανότητα γενίκευσης του μοντέλου.

### 5.3.3 Προεπεξεργασία των υπερηχογραφικών εικόνων

Αρχικά προεπεξεργαστήκαμε τις υπερηχογραφικές εικόνες ώστε να ταιριάζουν με τις RGB εικόνες διαστάσεων  $224 \times 224$  του ImageNet, στο οποίο είναι προεκπαιδευμένα τα DCNNs που θα χρησιμοποιήσουμε.. Οι διαθέσιμες εικόνες έχουν διαστάσεις  $120 \times 334$  και οι τιμές των pixels είναι από 0 έως 225 σε κλίμακα grayscale. Δοκιμάσαμε δύο μεθόδους προεπεξεργασίας των εικόνων.

#### 1η μέθοδος προεπεξεργασίας των εικόνων

Καθώς η μία διάσταση των εικόνων είναι μικρότερη από την επιθυμητή και το `resizing` σε  $224 \times 224$  έχει επίδραση στην ποιότητα της εικόνας για να μετατρέψουμε τις εικόνες σε διαστάσεις  $224 \times 224$  επιλέξαμε να ακολουθήσουμε τα εξής βήματα :

- Αρχικά περικόψαμε τις εικόνες στο κέντρο (`center crop`) διατηρώντας τη διάσταση του ύψους σταθερή (120) και επιλέγοντας την άλλη διάσταση ίση με 224. Παρατηρήσαμε ότι σε κάποιες εικόνες χάσαμε πληροφορία από τις άκρες της πλάκας ενώ διατηρήσαμε την αρχική ανάλυση και δεν παραμορφώσαμε τα όρια των πλακών. Επομένως μετά από αυτό το βήμα οι εικόνες είχαν διαστάσεις  $120 \times 224$ .
- Έπειτα κάναμε `zero padding` (μαύρο χρώμα) τη διάσταση του ύψους ώστε να προκύψουν εικόνες  $224 \times 224$ .
- Τέλος, για τις δοκιμές με τα προεκπαιδευμένα μοντέλα, κανονικοποιήσαμε τις τιμές των pixels των εικόνων με βάση τη μέση τιμή και την τυπική απόκλιση των εικόνων του ImageNet όπως συστήνεται.

Να σημειώσουμε ότι για την ανάγνωση των εικόνων χρησιμοποιήσαμε τη συνάρτηση `imread` της Skimage και μετατρέψαμε τη μονοκαναλική grayscale εικόνα σε RGB επαναλαμβάνοντας σε κάθε κανάλι την αρχική.

### 2η μέθοδος προεπεξεργασίας των εικόνων

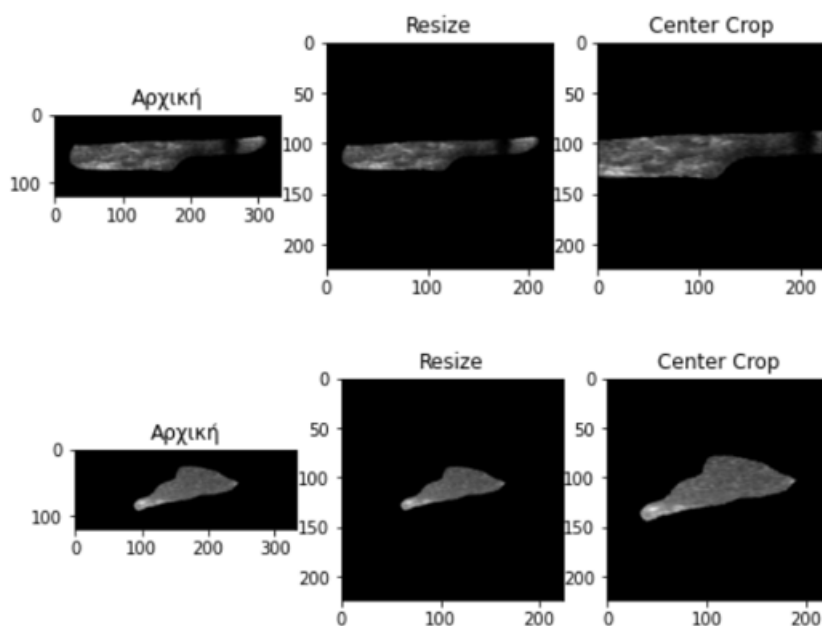
Σε αυτήν την περίπτωση μετατρέψαμε τις ορθογώνιες εικόνες σε τετράγωνα  $224 \times 224$  κάνοντας τις `resize` διατηρώντας όμως τις αναλογίες της.

- Αρχικά μετατρέψαμε τις εικόνες σε διαστάσεις  $80 \times 224$ . Το 80 προκύπτει για τη διατήρηση των αναλογιών των εικόνων.  $\frac{120}{334} \approx \frac{80}{224}$ .
- Έπειτα κάναμε `zero padding` (μαύρο χρώμα) στη διάσταση του ύψους ώστε να προκύψουν εικόνες  $224 \times 224$ .
- Τέλος κι εδώ, για τα προεκπαιδευμένα μοντέλα, κανονικοποιήσαμε τις τιμές των pixels των εικόνων με βάση το ImageNet.

Έπειτα από δοκιμές παρατηρήσαμε ότι η 1η μέθοδος έδωσε καλύτερα αποτελέσματα στα νευρωνικά μοντέλα για τις εικόνες καθώς θυσιάζοντας πληροφορία από τις άκρες των πλακών σε λίγες εικόνες δεν χάσαμε περισσότερη πληροφορία συμπιέζοντας όλες τις διαθέσιμες εικόνες. Στο Σχήμα 5.16 βλέπουμε αριστερά την αρχική εικόνα, στη μέση την εικόνα έπειτα από `resize` και δεξιά την εικόνα μετά από `center crop`. Στην πρώτη σειρά δείχνουμε μία εικόνα όπου με την `center crop` χάνουμε κομμάτι της εικόνας και στη δεύτερη σειρά μια εικόνα όπου και με τις δύο μεθόδους δεν χάνεται κομμάτι της. Υπενθυμίζουμε ότι οι εικόνες σαν την πρώτη είναι λίγες στο σύνολο δεδομένων μας.

## 5.3.4 Προεπεξεργασία των δεδομένων πίνακα

Αρχικά δημιουργήσαμε έναν κοινό πίνακα δεδομένων συνενώνοντας κατάλληλα τα δεδομένα από τα διαφορετικά αρχεία. Διαθέτουμε τόσο συνεχή όσο και διακριτά (κατηγορικά) χαρακτηριστικά (Πίνακας 5.3) . Στη συνέχεια για όλα τα διακριτά χαρακτηριστικά που δεν είναι ήδη `one-hot-encoded` τα κωδικοποιήσαμε κατάλληλα. Το χαρακτηριστικό “Smoker”



Σχήμα 5.7: Μετατροπή μεγέθους εικόνων σε 224 × 224

που λαμβάνει τις τιμές 1,2,3 κωδικοποιείται με τεχνική one-hot-encoded και προκύπτει η κατηγοριοποίηση του Πίνακα 5.2.

Smoker	Current_Smoker	Old_Smoker
1	0	0
2	0	1
3	1	0

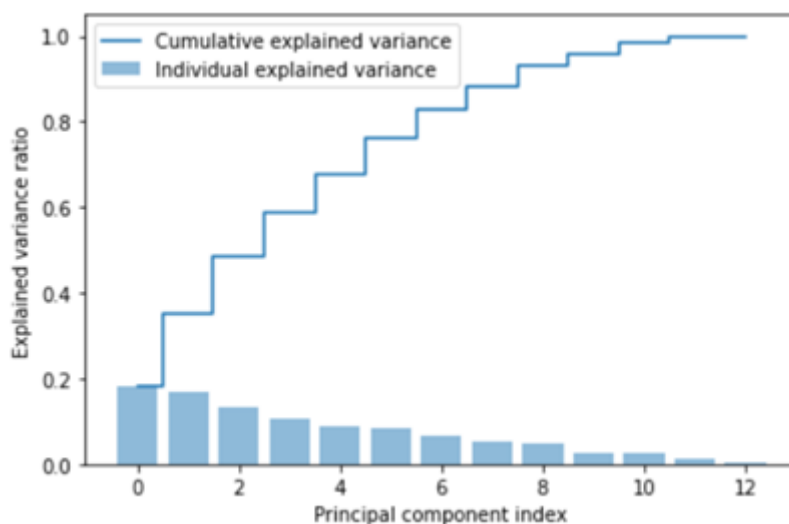
Πίνακας 5.2: One-hot κωδικοποίηση του διακριτού χαρακτηριστικού Smoker

Στη συνέχεια εφαρμόζουμε τον αλγόριθμο PCA, μετά από κανονικοποίηση, στα 13 συνεχή χαρακτηριστικά. Από το Σχήμα 5.8 παρατηρούμε ότι για να μην χάσουμε στατιστική πληροφορία χρειαζόμαστε τουλάχιστον 9 κύριες συνιστώσες. Κρατώντας τις 9 πρώτες συνιστώσες της Ανάλυσης Κύριων Συνιστωσών έχουμε διατηρήσει το 93.34% της συνολικής διακύμανσης η οποία κρίνεται αρκετή.

Για την εκπαίδευση των πολυτροπικών μοντέλων έπειτα από πειραματισμούς καταλήξαμε σε:

- Πλήθος εποχών: 20 εποχές με παράμετρο patience του Early Stopping ίση με 5, δηλαδή τερματισμό της εκπαίδευσης εάν το validation loss δεν παρουσιάζει βελτίωση σε 5 διαδοχικές εποχές.
- Batch Size: Επιλέχθηκε batch size 10 καθώς έδωσε καλύτερες επιδόσεις συγκριτικά με μεγαλύτερες και μικρότερες τιμές.
- Συνάρτηση σφάλματος: Επιλέχθηκε η Binary Cross Entropy Loss with Logits που παρέχεται από τη βιβλιοθήκη της PyTorch. Συνδυάζει ένα Sigmoid layer και το BCEloss σε μία μόνο κατηγορία γι'αυτό και χρησιμοποιείται στα προβλήματα δυαδικής ταξι-





Σχήμα 5.8: PCA στα συνεχή κλινικά και εργαστηριακά χαρακτηριστικά

Συνεχή χαρακτηριστικά	Διακριτά χαρακτηριστικά
Αιμοπετάλια Ουρία Κρεατινίνη Οξαλοξική Τρανσαμινάση Πυροσταφυλική Τρανσαμινάση γ-Γλουταμυλοτρανσφεράση Αλκαλική Φωσφατάση Ολική Χοληστερόλη Τριγλυκερίδια HDL Χοληστερόλη LDL Χοληστερόλη Γλυκόζη Ηλικία	Φύλο (Gender) Διαβήτης (Diabetes) Καπνιστής στο παρελθόν Καπνιστής

Πίνακας 5.3: Κλινικά και εργαστηριακά χαρακτηριστικά

νόμησης όπως το δικό μας. Σύμφωνα με το documentation της PyTorch ο συνδυασμός αυτός είναι πιο σταθερός αριθμητικά από τη χρήση ενός απλού Sigmoid που ακολουθείται από ένα BCEloss. Δίνεται από τον τύπο:

$$l(x, y) = L = \{l_1, l_2, \dots, l_N\}^T, l_n = -w_n [y_n \log(p(x_n)) + (1 - y_n) \log(1 - p(x_n))]$$

, όπου N το batch size.

- Βελτιστοποιητής: Μετά από δοκιμές επιλέχθηκε ο Adam με learning rate 0.001 έναντι του SGD.

## 5.4 Περιγραφή των μοντέλων

### 5.4.1 Μοντέλα για την ταξινόμηση των ασθενών με βάση τις εικόνες

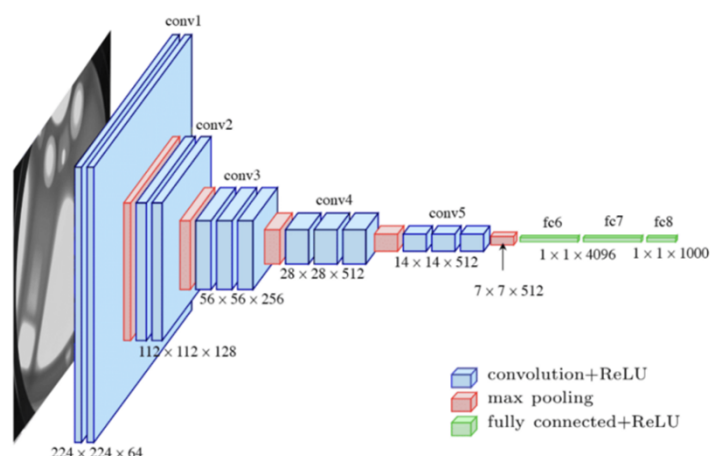
Για την ταξινόμηση των ασθενών στις δύο κατηγορίες, Υψηλού και Χαμηλού κινδύνου, με βάση τις υπερηχογραφικές εικόνες τους δοκιμάσαμε σύγχρονα συνελκτικά μοντέλα βαθιάς μάθησης. Τα μοντέλα που επιλέξαμε να ασχοληθούμε είναι τα ResNet18, ResNet50, AlexNet, VGG16 και VGG19.

#### Μοντέλο VGG16

Το μοντέλο VGG16 προτάθηκε από τους Karen Simonyan και Andrew Zisserman [99] και κέρδισε το πρώτο βραβείο στην κατηγορία ανίχνευσης αντικειμένων (object detection) και το δεύτερο βραβείο στην κατηγορία ταξινόμησης (classification) στο γνωστό διαγωνισμό ImageNet Large-Scale Visual Recognition Challenge (ILSVRC) ImageNet Challenge το 2014. Βασικό στοιχείο της δημιουργίας του VGG16 υπήρξε η αύξηση του βάθους του δικτύου με την προσθήκη περισσότερων συνελκτικών επιπέδων, κάτι που έγινε εφικτό λόγω της χρήσης πολύ μικρών ( $3 \times 3$ ) φίλτρων συνέλιξης σε όλα τα επίπεδα. Αποτελείται από δεκατρία (13) συνελκτικά επίπεδα, πέντε (5) επίπεδα max pooling και 3 πλήρως συνδεδεμένα επίπεδα δηλαδή συνολικά 21 επίπεδα από τα οποία μόνο τα 16 διαθέτουν βάρη δηλαδή εκπαιδευσιμες παραμέτρους. Συνολικά οι εκπαιδευσιμες παράμετροι του δικτύου είναι 138 χιλιάδες.

Η αρχιτεκτονική του VGG16 φαίνεται στο Σχήμα και περιγράφεται στη συνέχεια. Αποτελείται από 5 συνελκτικά τμήματα και μετά ακολουθούν τα πλήρως συνδεδεμένα επίπεδα. Τα συνελκτικά τμήματα αποτελούνται από διαδοχικά συνελκτικά στρώματα με συνελκτικά φίλτρα διαστάσεων  $3 \times 3$ , stride και padding ίσα με 1 και συναρτήσεις ενεργοποίησης ReLU και ακολουθεί ένα επίπεδο max pooling με πυρήνα  $2 \times 2$  και stride 2. Πιο συγκεκριμένα :

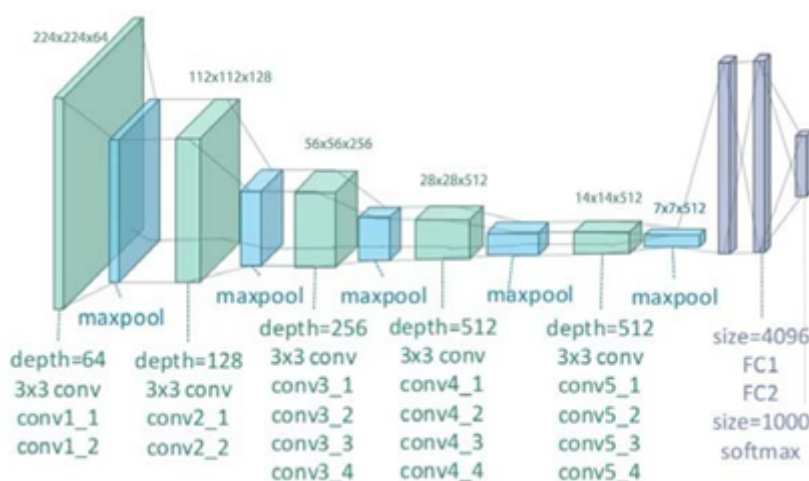
- Το δίκτυο λαμβάνει ως είσοδο εικόνες RGB διαστάσεων  $224 \times 224$ . Η μόνη προεπεξεργασία που γίνεται είναι να αφαιρεθεί η μέση τιμή RGB, που υπολογίζεται στο σύνολο εκπαίδευσης, από κάθε pixel.
- Τα συνελκτικά επίπεδα του πρώτου συνελκτικού τμήματος είναι δύο των 64 συνελκτικών φίλτρων το καθένα.
- Τα συνελκτικά επίπεδα του δεύτερου συνελκτικού τμήματος είναι δύο των 128 συνελκτικών φίλτρων το καθένα.
- Τα συνελκτικά επίπεδα του τρίτου συνελκτικού τμήματος είναι τρία των 256 συνελκτικών φίλτρων το καθένα.
- Τα συνελκτικά επίπεδα του τέταρτου και πέμπτου συνελκτικού τμήματος είναι τρία των 512 συνελκτικών φίλτρων το καθένα.
- Τα συνελκτικά τμήματα ακολουθεί ένα επίπεδο flatten και έπειτα 1 πλήρως συνδεδεμένα επίπεδα με 4096 νευρώνες.
- Τέλος το πλήρως συνδεδεμένο επίπεδο εξόδου αποτελείται από 1000 νευρώνες, όσες και οι κλάσεις στο ILSVRC ImageNet Challenge, και ενεργοποίηση Softmax.



Σχήμα 5.9: Αρχιτεκτονική VGG16 [18]

### Μοντέλο VGG19

Το μοντέλο VGG19 προτάθηκε κι αυτό από τους Karen Simonyan και Andrew Zisserman και περιγράφεται στην ίδια δημοσίευση με το VGG16 [99]. Στην πραγματικότητα αποτελεί παραλλαγή του VGG16 που αποτελείται από τρία επιπλέον επίπεδα εκπαιδευσιμων παραμέτρων. Η αρχιτεκτονική του φαίνεται στο Σχήμα και παρατηρούμε ότι οι αλλαγές σε σχέση με το VGG16 είναι ότι στο τρίτο, στο τέταρτο και στο πέμπτο συνελκτικα τμήματα έχει προστεθεί από ένα ακόμη συνελκτικό επίπεδα. Οι εκπαιδευσιμες παράμετροι του δικτύου είναι 144 χιλιάδες.



Σχήμα 5.10: Αρχιτεκτονική VGG19 [19]

### Μοντέλα ResNet

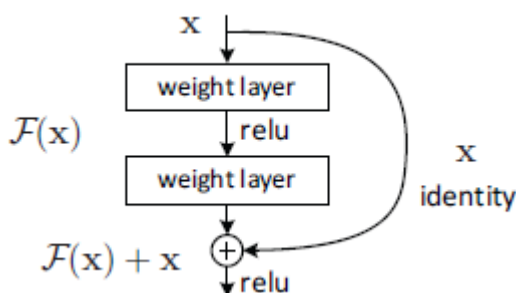
Τα μοντέλα ResNet αναπτύχθηκαν ως απάντηση στα VGG τα οποία δεν μπορούσαν να αυξήσουν άλλο το βάθος τους χωρίς να χάνουν την ικανότητά γενίκευσής τους και συνεπώς την ιδιότητά τους να λειτουργούν καλά ως ταξινομητές. Παρατηρήθηκε ότι καθώς το βάθος του δικτύου αυξανόταν, η ακρίβειά του έφτανε σε κορεσμό και στη συνέχεια υποβαθμιζόταν με γρήγορο ρυθμό ενώ επίσης αυξανόταν και το σφάλμα εκπαίδευσης [20]. Η κύρια ιδέα

που εισήγαγαν οι δημιουργοί του ResNet για να αντιμετωπίσει αυτό το πρόβλημα ήταν η χρήση συνδέσεων άλματος που συναντώνται στη βιβλιογραφία ως *jumping connections* ή *shortcut connections* ή *identity connections*. Οι συνδέσεις άλματος υλοποιήθηκαν με τα *residual blocks* (Σχήμα ). Το *residual block* ορίζεται ως:

$$y = F(x, W_i) + x$$

Με  $x$  συμβολίζεται το διάνυσμα εισόδου του block και με  $y$  το διάνυσμα εξόδου. Η συνάρτηση  $F(x, W_i)$  είναι το *residual mapping* που πρέπει να μάθει το block. Για παράδειγμα για το *residual block* δύο επιπέδων του Σχήματος 5.11 έχουμε  $F = W_2\sigma(W_1x)$ , όπου ο  $\sigma$  η συνάρτηση ενεργοποίησης ReLU κι επιπλέον για απλότητα έχουν παραλειφτεί τα *biases*. Στην περίπτωση που οι διαστάσεις των  $F$  και  $x$  δεν είναι οι ίδιες χρησιμοποιείται μια γραμμική προβολή  $W_s$  και τελικά:

$$y = F(x, W_i) + W_sx$$



Σχήμα 5.11: Αρχιτεκτονική *Residual Block* [20]

Πιο συγκεκριμένα, αυτό που κάνουν τα *shortcut connections* είναι να δημιουργούν συντομεύσεις μεταξύ των επιπέδων και να χρησιμοποιούν ξανά τις ενεργοποιήσεις προηγούμενων επιπέδων αποφεύγοντας το *gradient descent*. Επιπλέον οι μεταπηδήσεις συμπιέζουν το δίκτυο με αποτέλεσμα να μαθαίνει πιο γρήγορα. Τα βασικά στοιχεία της αρχιτεκτονικής των ResNets είναι τα παρακάτω:

- Χρήση κυρίως φίλτρων  $3 \times 3$ .
- Για το ίδιο μέγεθος χάρτη χαρακτηριστικών εξόδου, τα επίπεδα έχουν τον ίδιο αριθμό φίλτρων.
- Όταν το μέγεθος του χάρτη χαρακτηριστικών μειώνεται στο μισό, ο αριθμός των φίλτρων διπλασιάζεται έτσι ώστε να διατηρηθεί η χρονική πολυπλοκότητα ανά επίπεδο.
- Η υποδειγματοληψία πραγματοποιείται απευθείας με συνελκτικά επίπεδα με *stride* 2.
- Τα τελευταία επίπεδα του δικτύου αποτελούν ένα επίπεδο *average pooling* ακολουθούμενο από ένα πλήρως συνδεδεμένο επίπεδο μεγέθους 1000 με ενεργοποίηση *Soft-max*.

Στο Σχήμα 5.12 παρουσιάζονται οι αρχιτεκτονικές ResNet18 και ResNet50 όπου οι αριθμοί 18 και 50 αναφέρονται στο πλήθος των εκπαιδευσιμων επιπέδων του κάθε μοντέλου.

layer name	output size	18-layer	34-layer	50-layer	101-layer	152-layer
conv1	112×112	7×7, 64, stride 2				
		3×3 max pool, stride 2				
conv2_x	56×56	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
conv3_x	28×28	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 8$
conv4_x	14×14	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 23$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 36$
conv5_x	7×7	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	1×1	average pool, 1000-d fc, softmax				
FLOPs		$1.8 \times 10^9$	$3.6 \times 10^9$	$3.8 \times 10^9$	$7.6 \times 10^9$	$11.3 \times 10^9$

Σχήμα 5.12: Αρχιτεκτονικές ResNet [20]

### Μοντέλο AlexNet

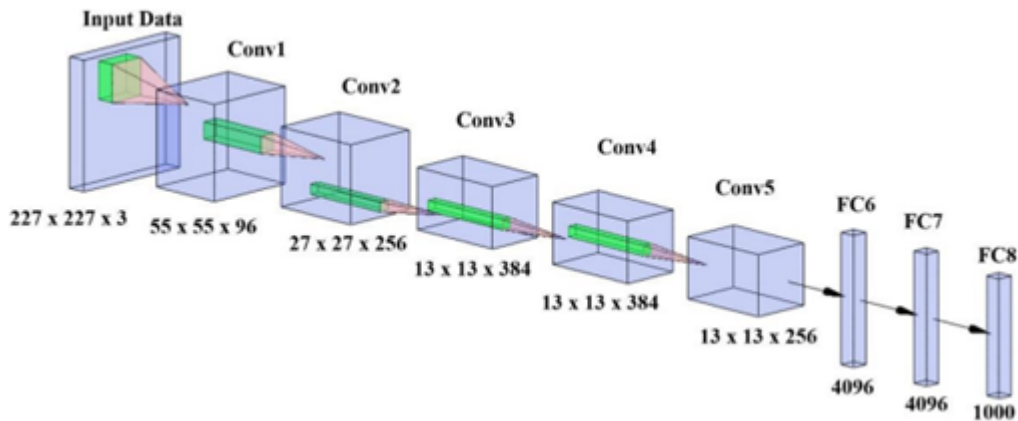
Πρωτοεμφανίστηκε στο διαγωνισμό ILSVRC το 2012 [100]. Αποτελείται από οχτώ επίπεδα και διαθέτει 60 χιλιάδες παραμέτρους εκπαίδευσης. Από τα οχτώ (8) επίπεδα τα πέντε (5) είναι συνελκτικά και τα υπόλοιπα πλήρως συνδεδεμένα. Οι νέες προσεγγίσεις που εισήγαγε το AlexNet είναι:

- Η χρήση της ReLU ως συνάρτηση ενεργοποίησης αντί της tanh που χρησιμοποιούταν τις περισσότερες φορές ως τότε. Το μεγαλύτερο πλεονέκτημα της ReLU ήταν η συνεισφορά της στη μείωση του χρόνου εκπαίδευσης.
- Ο σχεδιασμός του ώστε να μπορεί να τρέξει σε πολλαπλές GPUs.
- Εισήγαγε το Overlapping Pooling. Οι δημιουργοί του, με χρήση του Overlapping Pooling, παρατήρησαν, αρχικά, μείωση του σφάλματος από 0.3 έως 0.4% και έπειτα ότι η τεχνική συνεισέφερε στην αποφυγή της υπερκπαίδευσης. Η έννοια του Overlapping Pooling επεξηγείται στο Σχήμα 5.14 χρησιμοποιώντας ως παράδειγμα το Max Pooling.

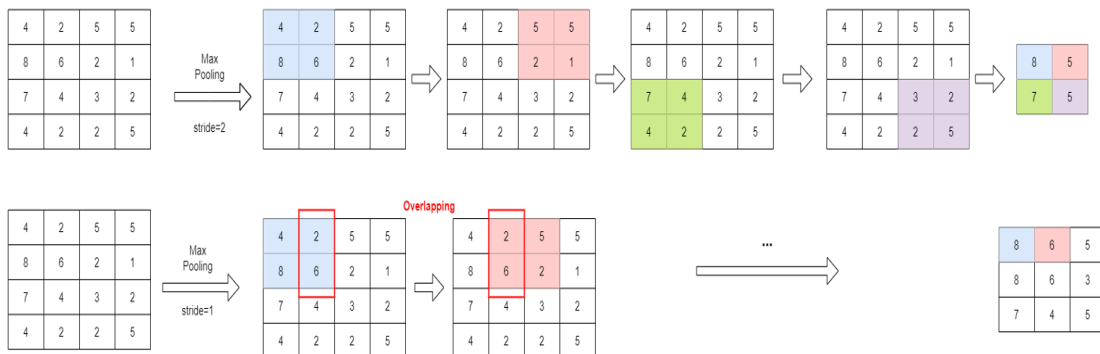
Το AlexNet περιλαμβάνει οκτώ επίπεδα με βάρη από τα οποία τα πρώτα πέντε είναι συνελκτικά και τα υπόλοιπα τρία πλήρως συνδεδεμένα. Η έξοδος του τελευταίου πλήρως συνδεδεμένου στρώματος τροφοδοτείται σε ένα επίπεδο softmax. Το δίκτυο μεγιστοποιεί την πολυωνυμική λογιστική παλινδρόμηση (Multinomial Logistic Regression objective), το οποίο ισοδυναμεί με τη μεγιστοποίηση του μέσου όρου της λογαριθμικής πιθανότητας καταγραφής της σωστής ετικέτας υπό την κατανομή πρόβλεψης σε όλες τις περιπτώσεις εκπαίδευσης. Οι πυρήνες του δεύτερου, του τέταρτου και του πέμπτου συνελκτικού επιπέδου συνδέονται μόνο με εκείνους τους χάρτες πυρήνα του προηγούμενου επιπέδου που βρίσκονται στην ίδια GPU ενώ οι πυρήνες του τρίτου συνελκτικού επιπέδου συνδέονται με όλους τους χάρτες πυρήνα του δεύτερου επιπέδου. Οι νευρώνες στα πλήρως συνδεδεμένα στρώματα συνδέονται με όλους τους νευρώνες στο προηγούμενο στρώμα. Επίσης το πρώτο και το δεύτερο επίπεδο συνέλιξης ακολουθούνται από επίπεδα κανονικοποίησης. Τα στρώματα Max-pooling ακολουθούν τόσο τα επίπεδα κανονικοποίησης όσο και το πέμπτο συνελκτικό στρώμα. Η συνάρτηση ReLU εφαρμόζεται στην έξοδο κάθε συνελκτικού και κάθε πλήρως συνδεδεμένου στρώματος.

Στη συνέχεια θα περιγράψουμε την κανονικοποίηση που χρησιμοποιείται στο AlexNet. Με  $a_{x,y}^i$  συμβολίζεται η απόκριση ενός νευρώνα μετά την εφαρμογή του πυρήνα  $i$  στη θέση  $(x, y)$  και την εφαρμογή της ReLU. Η κανονικοποιημένη απόκριση  $b_{x,y}^i$  δίνεται από τον παρακάτω τύπου όπου το άθροισμα "τρέχει" πάνω από  $n$  «γειτονικούς» χάρτες πυρήνα στην ίδια χωρική θέση. Με  $N$  συμβολίζεται ο συνολικός αριθμός των πυρήνων στο στρώμα. Η σειρά των χαρτών του πυρήνα είναι φυσικά αυθαίρετη και καθορίζεται πριν ξεκινήσει η εκπαίδευση. Οι σταθερές  $k, n, a$  και  $\beta$  είναι υπερπαραμέτροι των οποίων οι τιμές προσδιορίζονται χρησιμοποιώντας το σύνολο επικύρωσης. Στο προτότυπο paper οι συγγραφείς χρησιμοποίησαν  $k = 2, n = 5, a = 10^{-4}$  και  $\beta = 0.75$ .

$$b_{x,y}^i = \frac{a_{x,y}^i}{(k + a \sum_{j=\max(0, i-n/2)}^{\min(N-1, i+n)} (a_{x,y}^j)^2)^\beta}$$



Σχήμα 5.13: Αρχιτεκτονική AlexNet [21]



Σχήμα 5.14: *Overlapping Pooling*: Στην πάνω εικόνα με  $stride=2$  δεν έχουμε *overlapping* ενώ στην κάτω εικόνα με  $stride=1$  έχουμε *overlapping*.

### 5.4.2 Μοντέλο για την ταξινόμηση των ασθενών με βάση τα κλινικά και εργαστηριακά χαρακτηριστικά

Χρησιμοποιήσαμε ένα MLP τριών επιπέδων. Αναπτύχθηκε με τη βιβλιοθήκη της PyTorch και περιλαμβάνει τρία πλήρως συνδεδεμένα επίπεδα. Τα δύο πρώτα έχουν διάσταση εξόδου 64 ακολουθούνται από συνάρτηση ενεργοποίησης ReLU και επίπεδο Batch Normalization. Το δεύτερο επίπεδο ακολουθείται επιπλέον από επίπεδο Dropout με  $p=0.1$ . Το τρίτο επίπεδο είναι το επίπεδο εξόδου.

### 5.4.3 Πολυτροπικά Μοντέλα

Αναπτύχθηκαν δύο μοντέλα ανάλογα με το επίπεδο στο οποίο γίνεται η συγχώνευση των δύο modalities. Τα μοντέλα αναπτύχθηκαν με τη βιβλιοθήκη PyTorch.

#### Πολυτροπικό μοντέλο Feature-Level Fusion

Το πολυτροπικό μοντέλο Feature-Level Fusion υλοποιείται σε τρεις φάσεις:

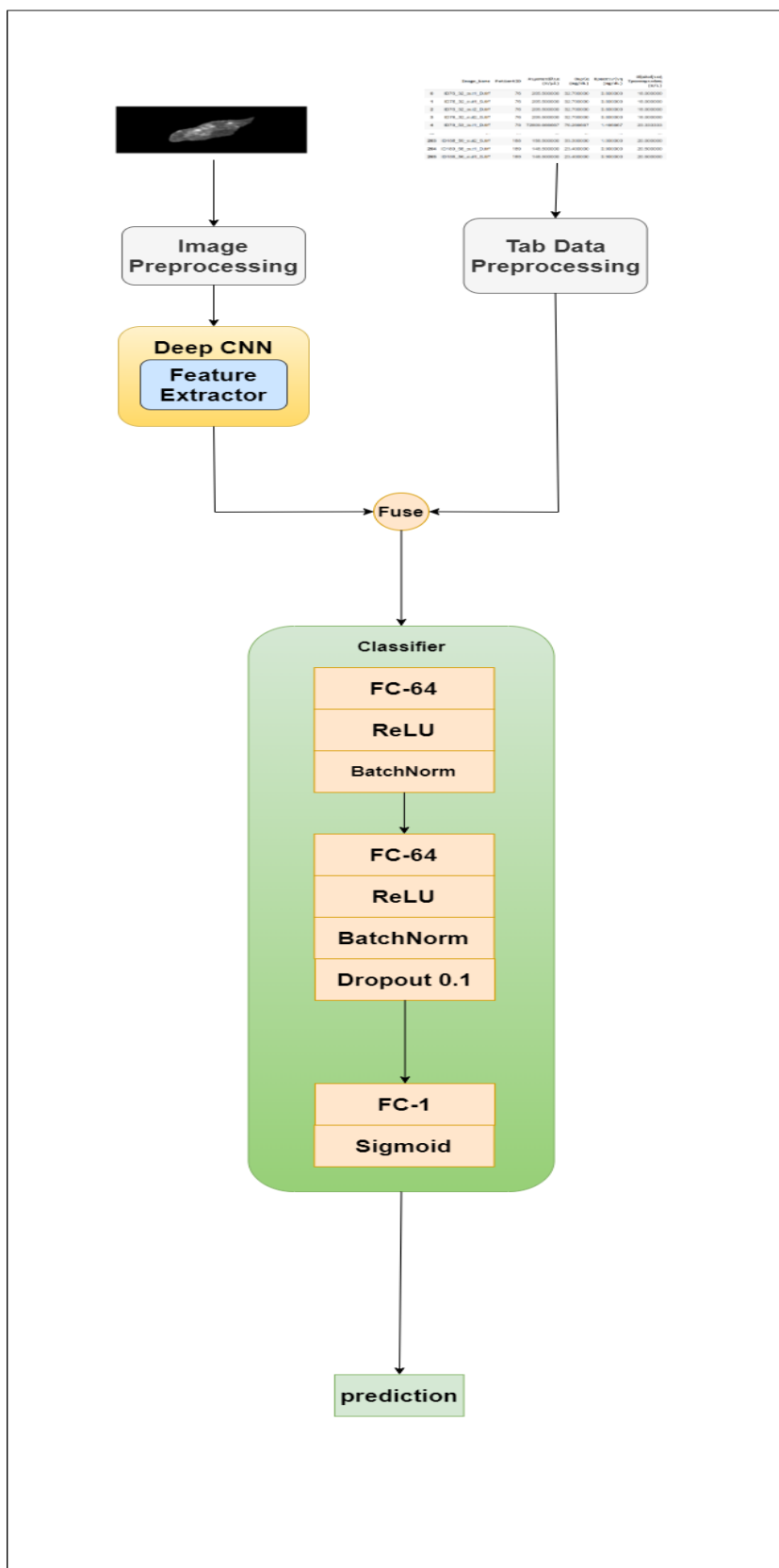
- Εξαγωγή των χαρακτηριστικών των υπερηχογραφικών εικόνων. Ως Features Extractors χρησιμοποιούμε τα μοντέλα που χρησιμοποιήθηκαν για την ταξινόμηση των εικόνων, δηλ. τα ResNet18, ResNet50, AlexNet, VGG16 και VGG19 αφαιρώντας τους ταξινομητές τους.
- Συνένωση του διανύσματος χαρακτηριστικών που εξάγεται από τις εικόνες με τα διανύσματα κλινικών και εργαστηριακών χαρακτηριστικών. Το διάνυσμα χαρακτηριστικών κάθε εικόνας συνενώνεται με το διανύσματα κλινικών και εργαστηριακών χαρακτηριστικών του ασθενή στον οποίον αντιστοιχεί η εικόνα.
- Το νέο διάνυσμα χαρακτηριστικών ταξινομείται στις δύο κλάσεις του προβλήματος από ταξινομητή τριών πλήρως συνδεδεμένων επιπέδων. Η αρχιτεκτονική του ταξινομητή παρουσιάζεται στο Σχήμα 5.15.

#### Πολυτροπικό μοντέλο Decision-Level Fusion

Το πολυτροπικό μοντέλο Decision-Level Fusion υλοποιείται στα παρακάτω στάδια:

- Ένα εκπαιδευμένο μοντέλο εικόνων ταξινομεί την εικόνα στις κλάσεις του προβλήματος.
- Ο εκπαιδευμένος ταξινομητής που χρησιμοποιήσαμε για τα δεδομένα πίνακα ταξινομεί το διάνυσμα κλινικών και εργαστηριακών χαρακτηριστικών στις κλάσεις του προβλήματος.
- Οι προβλέψεις των δύο μοντέλων συγχωνεύονται σε ένα κοινό διάνυσμα. Το διάνυσμα αυτό χρησιμοποιείται ως είσοδος σε νευρωνικό δίκτυο τριών επιπέδων από το οποίο προκύπτει η τελική πρόβλεψη.

Η αρχιτεκτονική του φαίνεται στο Σχήμα 5.16.

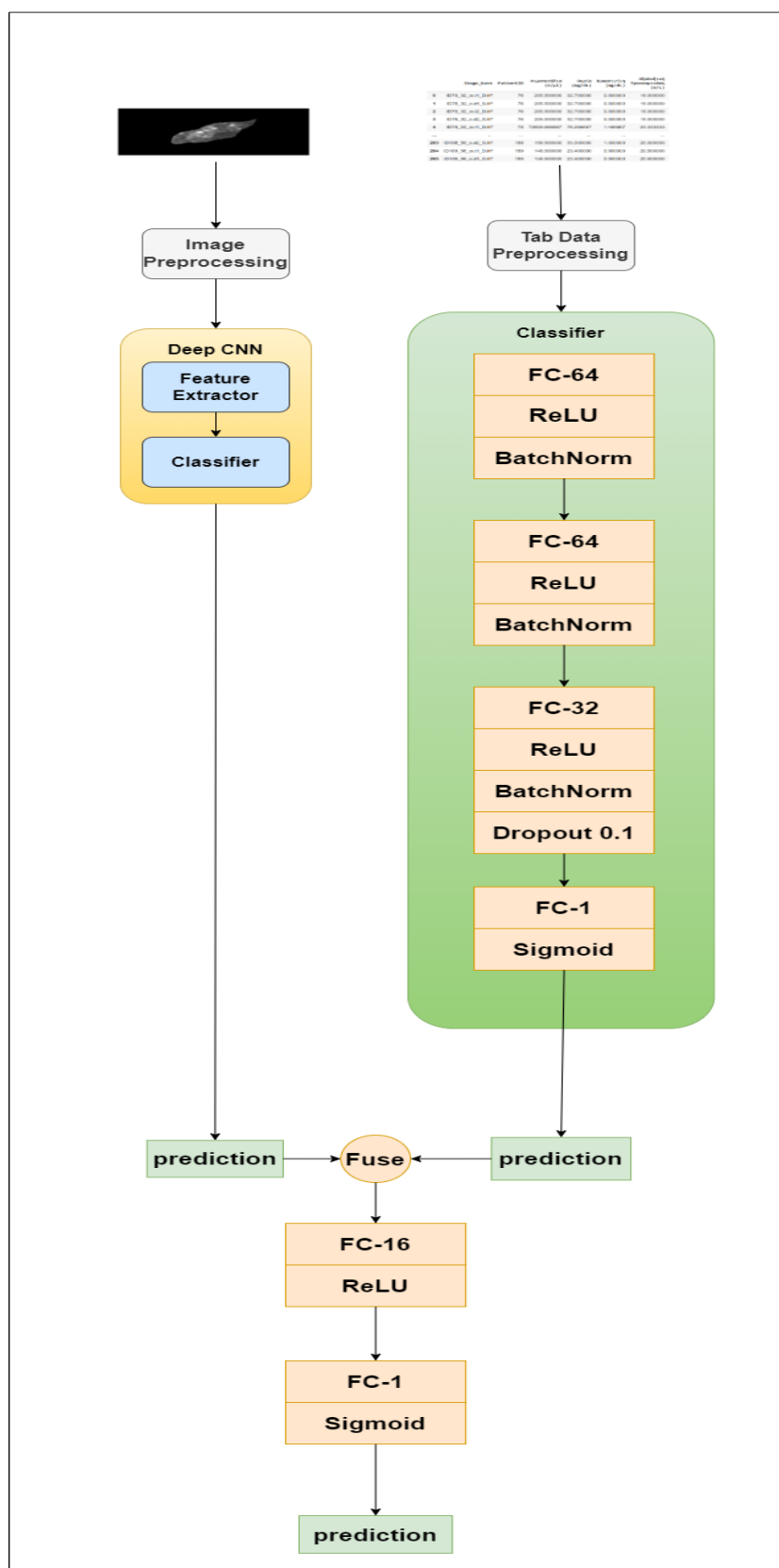


Σχήμα 5.15: Αρχιτεκτονική μοντέλου Feature-Level Fusion

#### 5.4.4 Χειρισμός της ανισορροπίας των δεδομένων

Για το χειρισμό της έντονης ανισορροπίας των δεδομένων πραγματοποιήσαμε **επαύξηση** των δεδομένων στο σύνολο εκπαίδευσης. Πιο συγκεκριμένα η επαύξηση πραγματοποιείται με *Διπλωματική Εργασία*





Σχήμα 5.16: Αρχιτεκτονική μοντέλου *Decision-Level Fusion*

τις μεθόδους SMOTE και CTGAN για τα δεδομένα πίνακα και με μετασχηματισμούς για τις εικόνες. Οι μετασχηματισμοί που χρησιμοποιήθηκαν για τις εικόνες είναι: περιστροφή κατά

90 μοίρες, οριζόντια αναστροφή (flip), τυχαίος μετασχηματισμός προοπτικής με πιθανότητα  $p=0.5$  και τυχαίος αφινικός μετασχηματισμός όπου το εύρος των μοιρών του κυμαίνεται από 30 έως 70 μοίρες και ο συντελεστής της κλίμακας από 0.7 έως 0.8. Γενικά αφινικός είναι ένας μετασχηματισμός στην Ευκλείδεια γεωμετρία ο οποίος διατηρεί τις γραμμές και τις παραλληλίες, αλλά όχι απαραίτητα τις αποστάσεις και τις γωνίες. Δοκιμάσαμε, αρχικά, να επαυξήσουμε μόνο την κατηγορία μειονότητας (Χαμηλού Κινδύνου) ώστε το σύνολο εκπαίδευσης να είναι ισορροπημένο και στη συνέχεια να επαυξήσουμε τα δεδομένα της κλάσης πλειονότητας και αντίστοιχα και της μειονότητας. Η επαύξηση μόνο της κλάσης μειονότητας, όπως θα φανεί και στην παρουσίαση των αποτελεσμάτων, βελτίωσε σημαντικά την απόδοση των μοντέλων. Αντίθετα ο πενταπλασιασμός σε συνδυασμό με εξισορρόπηση του συνόλου δεν προσέφερε κάποια βελτίωση ενώ ο διπλασιασμός του προσέφερε μικρή βελτίωση.

Ο άλλος τρόπος χειρισμού της ανισορροπίας που δοκιμάσαμε ήταν η **προσθήκη κατάλληλου βάρους στη συνάρτηση σφάλματος**. Η συνάρτηση σφάλματος Binary Cross Entropy Loss with logits που παρέχεται από την PyTorch διαθέτει την παράμετρο `pos_weight` που αφορά στο βάρος της θετικής κλάσης. Τροποποιώντας την παράμετρο αυτή η συνάρτηση σφάλματος δίνει διαφορετική βαρύτητα στα σφάλματα που αφορούν στα δείγματα μειονότητας

Επιπλέον δοκιμάσαμε να μεταβάλλουμε το κατώφλι απόφασης των ταξινομητών από 0.5 σε 0.4 και σε 0.7.

## 5.5 Μετρικές αξιολόγησης των μοντέλων

Η αξιολόγηση των μοντέλων βασίζεται σε επιλεγμένες μετρικές οι οποίες παρουσιάζονται παρακάτω. Μεγαλύτερη βαρύτητα δόθηκε στις μετρικές Balanced Accuracy, F1-Score και MCC διότι δίνουν πιο αντιπροσωπευτικές τιμές σε μη ισορροπημένα σύνολα. Προτού ορίσουμε τις μετρικές δίνουμε τους παρακάτω ορισμούς.

- **True Positive (TP)** αποτέλεσμα: το αποτέλεσμα του μοντέλου όταν κατηγοριοποιεί ορθά ένα δείγμα που ανήκει στη θετική κλάση
- **False Positive (FP)** αποτέλεσμα: το αποτέλεσμα του μοντέλου όταν κατηγοριοποιεί ένα δείγμα στη θετική κλάση ενώ το δείγμα ανήκει στην πραγματικότητα στην αρνητική
- **True Negative (TN)** αποτέλεσμα: το αποτέλεσμα του μοντέλου όταν κατηγοριοποιεί ορθά ένα δείγμα που ανήκει στη αρνητική κλάση
- **False Negative (FN)** αποτέλεσμα: το αποτέλεσμα του μοντέλου όταν κατηγοριοποιεί ένα δείγμα στην αρνητική κλάση ενώ το δείγμα ανήκει στην πραγματικότητα στη θετική
- **False Negative Rate (FNR)**: ορίζεται ως ο λόγος του πλήθους των False Negative αποτελεσμάτων (FN) προς το πλήθος των θετικών δειγμάτων (TP+FN). Τα πλήθος των θετικών δειγμάτων είναι ίσο με τα δείγματα που ο ταξινομητής κατέταξε ορθά ως θετικά (TP) και επιπλέον τα δείγματα που κατέταξε εσφαλμένα ως αρνητικά (FN).

- **False Positive Rate (FPR):** ορίζεται ως ο λόγος του πλήθους των False Positive αποτελεσμάτων (FP) προς το σύνολο των αρνητικών δειγμάτων (TN+FP).

Να υπενθυμίσουμε, επίσης, ότι στο πρόβλημά μας ορίζουμε ως αρνητική κλάση την κλάση Χαμηλού Κινδύνου (Low Risk) και ως θετική κλάση την Υψηλού Κινδύνου (High Risk). Για την αξιολόγηση των μοντέλων χρησιμοποιούμε τις μετρικές ROC/AUC, Accuracy, Balanced Accuracy, Sensitivity, Specificity, Precision και F1-Score οι οποίες είναι οι πλέον κατάλληλες για τα προβλήματα ταξινόμησης όπως το δικό μας.

- **Accuracy:** Δίνει το ποσοστό των δειγμάτων, τόσο της θετικής όσο και της αρνητικής κλάσης, που έχουν κατηγοριοποιηθεί ορθά. Στις περιπτώσεις μη ισορροπημένων συνόλων δεδομένων η μετρική αυτή δεν δίνει αντιπροσωπευτικά αποτελέσματα καθώς για παράδειγμα στην περίπτωση που όλα τα δείγματα, θετικά και αρνητικά, κατηγοριοποιηθούν στην κλάση της πλειοψηφίας η τιμή της μετρικής θα είναι υψηλή.

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

- **Balanced Accuracy:** Όπως και η Accuracy αποτελεί μέτρο του κατά πόσο έχουν κατηγοριοποιηθεί ορθά τα δείγματα. Χρησιμοποιείται στις περιπτώσεις των μη ισορροπημένων συνόλων δεδομένων της Accuracy καθώς δίνει πιο αντιπροσωπευτικά αποτελέσματα.

$$\frac{TPR + TNR}{2}$$

- **Sensitivity ή Recall ή True Positive Rate (TPR):** Ορίζεται ως ο λόγος του πλήθους των True Positive αποτελεσμάτων (TP) προς το σύνολο των θετικών αποτελεσμάτων του ταξινομητή (TP+FN). Με τον όρο θετικά αποτελέσματα εννοούμε το πλήθος των δειγμάτων που ταξινομήθηκαν στη θετική κλάση.

$$\frac{TP}{TP + FN} = 1 - FNR$$

- **Specificity ή True Negative Rate (TNR):** Ορίζεται ως ο λόγος του πλήθους των True Negative αποτελεσμάτων (TN) προς το σύνολο των αρνητικών αποτελεσμάτων του ταξινομητή (TN+FP). Με τον όρο αρνητικά αποτελέσματα εννοούμε το πλήθος των δειγμάτων που ταξινομήθηκαν στη θετική κλάση.

$$\frac{TN}{TN + FP} = 1 - FPR$$

- **Precision:** Ποσοτικοποιεί τον αριθμό των σωστών θετικών προβλέψεων που έγιναν.

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

- **F1-Score:** Είναι ο αρμονικός μέσος όρος των Precision και Recall

$$2 \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} = \frac{2TP}{2TP + FP + FN}$$

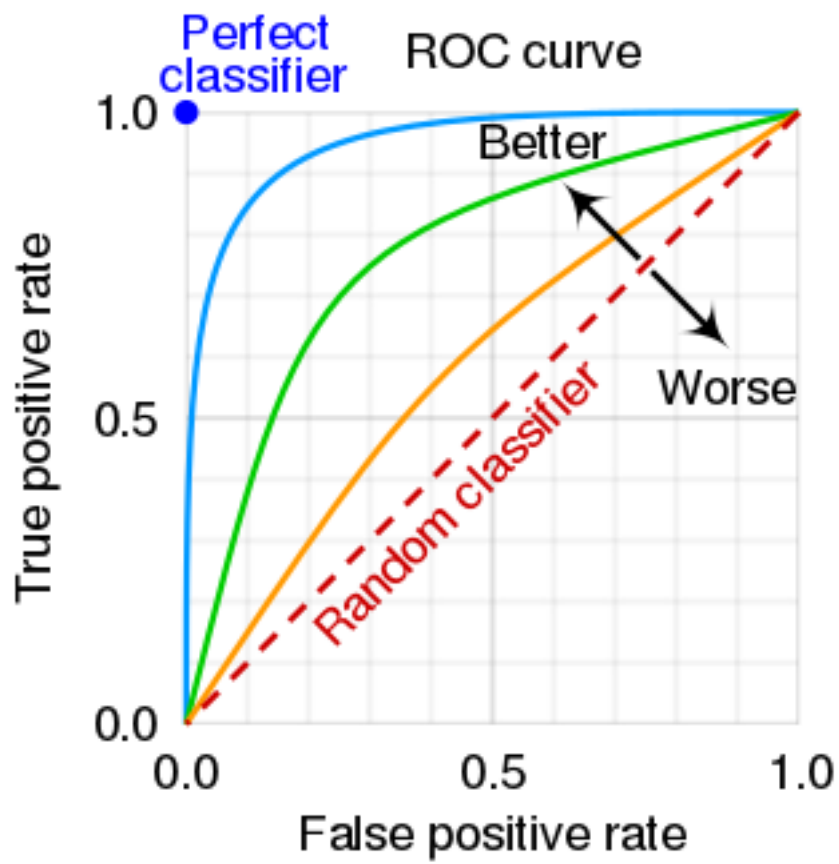
- **ROC/AUC:** Αποτελεί συντομογραφία του όρου Area Under the Curve (AUC) of Receiver Characteristic Operator (ROC). Η μετρική αυτή χρησιμοποιείται ευρέως σε προβλήματα δυαδικής ταξινόμησης ως δείκτης του πόσο καλά καταφέρνει το μοντέλο ταξινομήσει τα δείγματα στις δύο κλάσεις. Για τον υπολογισμό της μετρικής AUC χρησιμοποιούμε την καμπύλη ROC. Όταν σχεδιάζουμε την καμπύλη ROC βάζουμε στον x-άξονα τον FPR και στον ψ-άξονα τον TPR όπως φαίνεται στο παράδειγμα του Σχήματος 5.17. Η καμπυλη ROC απεικονίζει γραφικά τη διαγνωστική ικανότητα ενός δυαδικού ταξινομητή για όλες τις τιμές (0 έως 1) του κατωφλίου απόφασης του ταξινομητή (Σχήμα 5.17). Η μετρική AUC ισούται με το εμβαδό της επιφάνειας κάτω από την καμπύλη ROC. Αποδεικνύεται μαθηματικά ότι η μετρική AUC είναι ίση με την πιθανότητα ο ένας ταξινομητής να κατατάξει μια τυχαία επιλεγμένη θετική περίπτωση υψηλότερα από μια τυχαία επιλεγμένη αρνητική. Όταν  $AUC=1$  το μοντέλο καταφέρνει να διαχωρίσει πλήρως τις δύο κλάσεις και συνεπώς ταξινομεί σωστά όλες τις εισόδους και το αντίστροφο ισχύει για  $AUC=0$ . Αξίζει να σημειώσουμε ότι η καμπύλη ROC μπορεί να μας βοηθήσει στο να αποφασίσουμε το κατώφλι του ταξινομητή μας το οποίο κρίνεται απαραίτητο ειδικά στις περιπτώσεις μη ισορροπημένων συνόλων δεδομένων.
- **Matthews correlation coefficient (MCC):** Ορίζεται ως

$$\frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

Σύμφωνα με δημοσιεύσεις [101] είναι η καταλληλότερη μετρική για την αξιολόγηση σε μη ισορροπημένα σύνολα καθώς ένα υψηλό MCC σκορ συνεπάγεται ότι ο ταξινομητής κάνει σωστές προβλέψεις τόσο για την πλειονότητα των αρνητικών περιπτώσεων όσο και για την πλειονότητα των θετικών. Οι τιμές της μετρικής κυμαίνονται στο διάστημα  $[-1, +1]$ , με τις ακραίες τιμές  $-1$  και  $+1$  να επιτυγχάνονται στην περίπτωση πλήρους εσφαλμένης και πλήρους ορθής ταξινόμησης, αντίστοιχα, ενώ  $MCC = 0$  είναι η αναμενόμενη τιμή για έναν τυχαίο ταξινομητή.

## 5.6 Μετατροπή δεδομένων πίνακα σε εικόνες

Παρατηρήσαμε ότι τα συνελκτικά δίκτυα βαθιάς μάθησης που χρησιμοποιήσαμε για τα υπερηχογραφήματα έδωσαν ικανοποιητικά αποτελέσματα. Για να μπορέσουμε να αξιοποιήσουμε τις δυνατότες αυτών των μοντέλων στα δεδομένα πίνακα δοκιμάσαμε να μετατρέψουμε τα δεδομένα πίνακα σε εικόνες. Χρησιμοποιήσαμε το πακέτο `tab2img` που παρέχεται από την Python. Αν και τα αποτελέσματα, έπειτα από εξισορρόπηση, στο σύνολο εκπαίδευσης ήταν ενθαρρυντικά τα αποτελέσματα στα σύνολα επικύρωσης και ελέγχου δεν ήταν ικανοποιητικά. Προτείνεται η περαιτέρω διερεύνηση του θέματος.



Σχήμα 5.17: Καμπύλη ROC



## Κεφάλαιο 6

### Αποτελέσματα

---

Σε αυτό το κεφάλαιο θα παρουσιάσουμε τα πειραματικά αποτελέσματα της εργασίας. Θα προσπαθήσουμε να δώσουμε αποτελέσματα από τα ενδιάμεσα βήματα που ακολουθήσαμε μέχρι να καταλήξουμε στα τελικά μοντέλα ώστε η παρουσίαση να είναι κατά το δυνατόν πλήρης και αναλυτική.

*Σημείωση:* Στους πίνακες που παρουσιάζονται στη συνέχεια οι συντομογραφίες είναι ως εξής: ACC για Accuracy, B-Acc για Balanced Accuracy, Spec για Specificity, Sens για Sensitivity, F1 για F1-Score και MCC για Matthews Correlation Coefficient.

#### 6.1 Αποτελέσματα unimodal μοντέλων για τις εικόνες μετά από προεκπαίδευση

Αρχικά, έπειτα από την προεπεξεργασία των δεδομένων, δοκιμάσαμε τα μοντέλα με προεκπαίδευση αλλά όχι κάποια άλλη βελτιστοποίηση. Παρατηρήσαμε ότι τα αποτελέσματα ήταν πολύ χαμηλά. Η επίδραση του μη-ισορροπημένου συνόλου εκπαίδευσης ήταν έντονη καθώς τα μοντέλα αποτύγχαναν σχεδόν πλήρως να ταξινομήσουν ορθά τα δείγματα της αρνητικής κλάσης που είναι και η κλάση μειονότητας. Στη συνέχεια εφαρμόσαμε μεταφορά μάθησης με χρήση των προεκπαιδευμένων στο ImageNet μοντέλων. Δοκιμάσαμε αρχικά να "παγώσουμε" όλα τα επίπεδα των μοντέλων για τη μεταφορά μάθησης, μετά να "ξεπαγώσουμε" τα 2 τελευταία κι έπειτα τα 3 τελευταία επίπεδα από κάθε μοντέλο. Τα βέλτιστα αποτελέσματα πήραμε όταν επιτρέψαμε την εκπαίδευση των 2 τελευταίων επιπέδων κάθε μοντέλου. Πιο συγκεκριμένα για κάθε μοντέλο στον Πίνακα 6.1

Μοντέλο	Επίπεδα που εκπαιδεύονται
ResNet18, ResNet50	fc, layer4
AlexNet	classifier(4), classifier(6)
VGG16, VGG19	classifier(3), classifier(6)

Πίνακας 6.1: Επίπεδα που εκπαιδεύονται κατά τη μεταφορά μάθησης

Στους Πίνακες 6.2, 6.3 παρουσιάζονται οι μέσοι όροι των 3 splits του Nested Cross Validation των μετρικών στο σύνολο ελέγχου (test set) για τα unimodal μοντέλα των εικόνων με μόνη βελτιστοποίηση την προεκπαίδευση, 20 εποχές, Early-Stopping patience 5 εποχές, Adam Optimizer με learning rate 0.001 και συνάρτηση σφάλματος Binary Cross Entropy

## Loss with Logits .

	ResNet18	ResNet50	AlexNet
ACC(%)	82.21 ±2.22	78.55 ±9.05	78.20 ±8.77
B-ACC(%)	74.33 ±5.39	71.06 ±14.42	<b>74.47</b> ±4.50
AUC(%)	82.19 ±7.53	81.09 ±9.97	<b>83.19</b> ±6.01
Spec(%)	61.11 ±15.71	58.33 ±20.67	<b>69.52</b> ±9.42
Sens(%)	87.55 ±5.62	83.79 ±5.59	79.43 ±12.13
Prec(%)	72.00 ±4.02	66.39 ±9.54	69.28 ±5.19
F1(%)	71.55 ±1.91	68.03 ±11.34	69.80 ±6.50
MCC	<b>0.4562</b> ±0.0454	0.3710 ±0.2353	0.4341 ±0.0981

Πίνακας 6.2: Επίδοση unimodal μοντέλων ResNet18, ResNet50, AlexNet για εικόνες μόνο με προεκπαίδευση

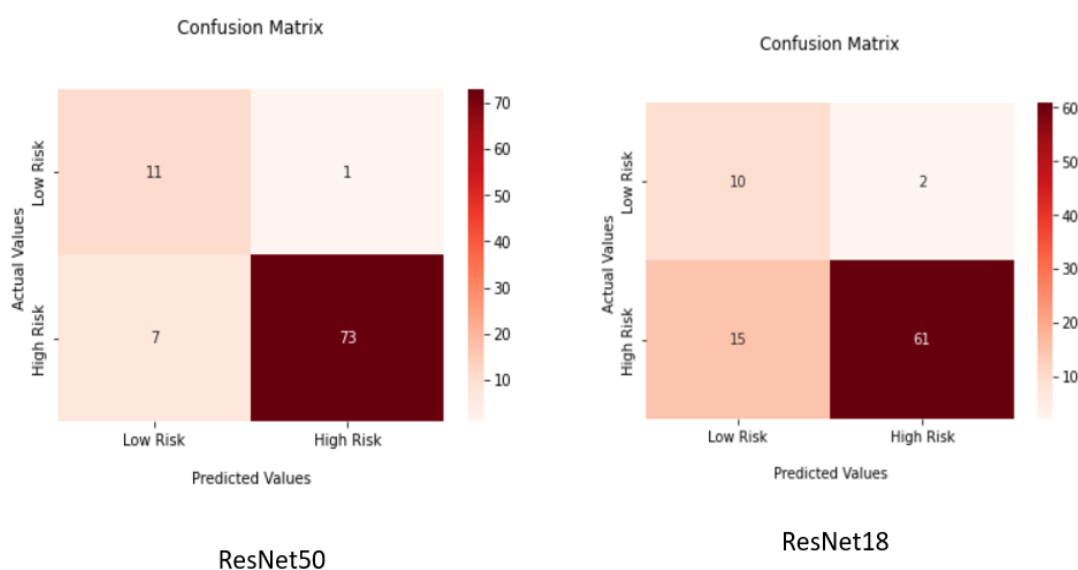
	VGG16	VGG19
ACC(%)	78.86 ±8.39	<b>84.06</b> ±6.88
B-ACC(%)	64.67 ±9.34	61.07 ±10.15
AUC(%)	79.27 ±2.69	75.97 ±9.11
Spec(%)	41.76 ±32.98	26.73 ±20.47
Sens(%)	<b>87.59</b> ±16.4	95.40 ±3.38
Prec(%)	<b>74.44</b> ±13.63	63.40 ±12.8
F1(%)	61.24 ±8.83	60.67 ±11.25
MCC	0.3456 ±0.1721	0.2526 ±0.2062

Πίνακας 6.3: Επίδοση unimodal μοντέλων VGG19, VGG16 για τις εικόνες μόνο με προεκπαίδευση

Παρατηρούμε ότι η μετρική Accuracy είναι σχετικά υψηλή σε σχέση με τις υπόλοιπες μετρικές. Επιβεβαιώνουμε ότι η μετρική Accuracy είναι αρκετά παραπλανητική σε ένα μη-ισορροπημένο σύνολο δεδομένων. Όλα τα μοντέλα παρουσίασαν ικανοποιητικές επιδόσεις σε όλες τις μετρικές εκτός από την Specificity. Αυτό σημαίνει ότι δεν προβλέπουν με επιτυχία την αρνητική κλάση η οποία είναι και η κλάση μειονότητας. Επίσης, σε αρκετές περιπτώσεις παρατηρείται μεγάλη τυπική απόκλιση. Ο μέσος όρος υπολογίζεται επί των 3 splits και σε ορισμένες περιπτώσεις η επίδοση του μοντέλου στα 3 διαφορετικά test sets παρουσιάζει διακυμάνσεις. Αυτό, πιθανότατα, οφείλεται στο γεγονός ότι σε κάποια folds οι ασθενείς Χαμηλού Κινδύνου που περιέχονται δεν αποτελούν αντιπροσωπευτικά δείγματα της κλάσης τους. Υπενθυμίζουμε ότι σε κάθε test fold οι ασθενείς Χαμηλού κινδύνου είναι 3 ή 4 και οι ασθενείς Υψηλού Κινδύνου 21 ή 22.

Στη συνέχεια παρατίθενται κάποιοι από τους confusion matrices που προέκυψαν από τα σύνολα ελέγχου. (Σχήματα 6.1, 6.2).





Σχήμα 6.1: Confusion Matrices για τα μοντέλα μόνο με προεκπαίδευση

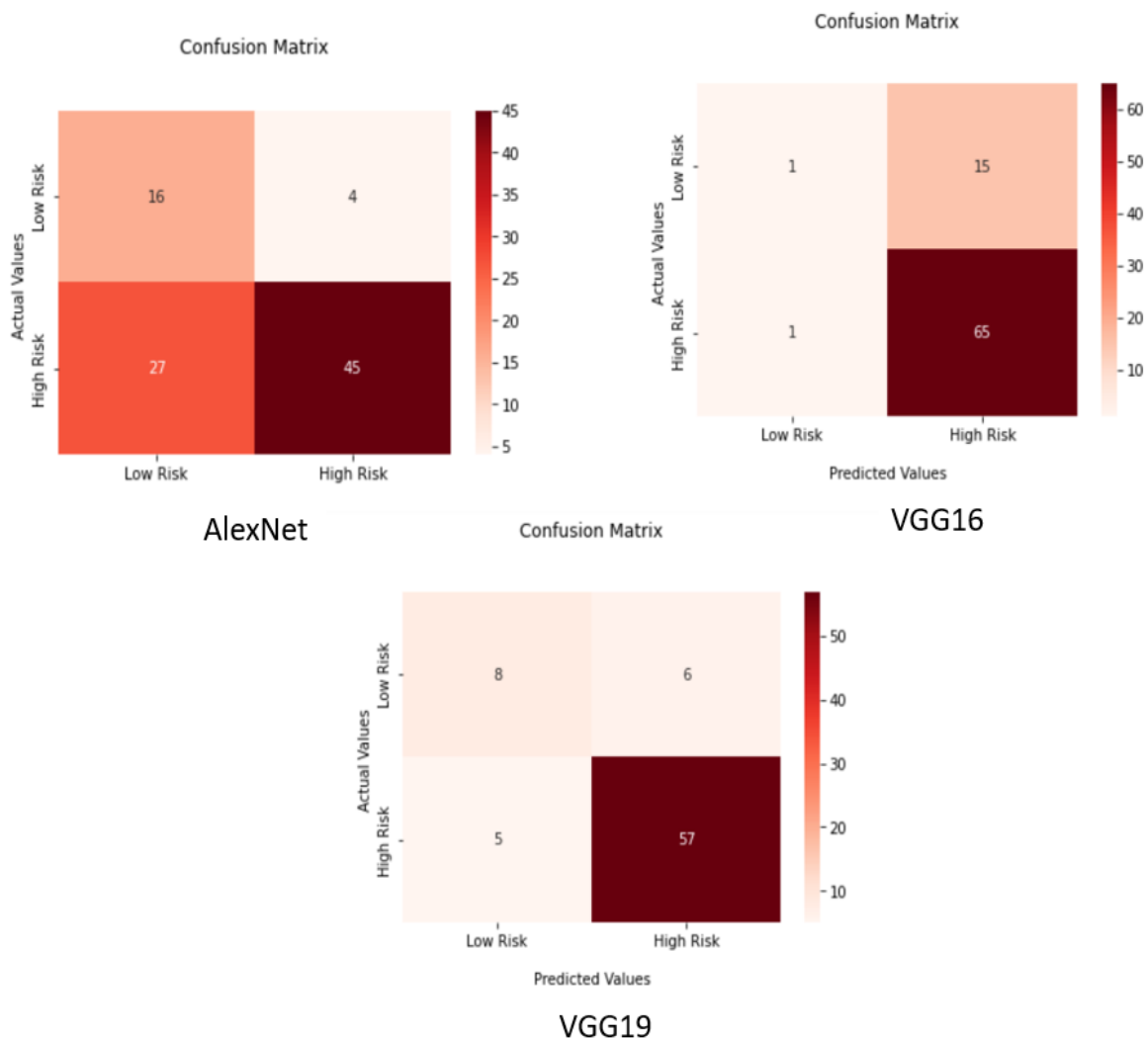
## 6.2 Αποτελέσματα unimodal μοντέλων για τις εικόνες έπειτα από εξισορρόπηση και επαύξηση

Στον Πίνακα 5.1 παρουσιάζονται οι μέσες τιμές και η τυπική απόκλιση των μετρικών στα 3 test folds για όλα τα μοντέλα εικόνων έπειτα από εξισορρόπηση του συνόλου δεδομένων με επαύξηση.

	ResNet18	ResNet50	AlexNet
ACC(%)	<b>86.52</b> ±5.50	84.42 ±4.76	80.55 ±3.92
B-ACC(%)	<b>74.37</b> ±9.72	69.09 ±10.41	71.79 ±5.51
AUC(%)	<b>90.24</b> ±5.13	88.33 ±7.86	87.22 ±3.51
Spec(%)	56.48 ±22.71	43.33 ±23.21	56.69 ±19.34
Sens(%)	92.27 ±7.59	94.85 ±3.98	86.64 ±9.17
Prec(%)	83.33 ±1.039	76.56 ±11.80	70.85 ±5.70
F1(%)	74.75 ±8.33	68.37 ±7.65	68.77 ±3.84
MCC	<b>0.5437</b> ±0.1300	0.4290 ±0.1555	0.4169 ±0.6610

Πίνακας 6.4: Επίδοση unimodal μοντέλων ResNet18, ResNet50, AlexNet για τις εικόνες μετά από μεταφορά μάθησης και εξισορρόπηση train set με επαύξηση

Στο Σχήμα 6.3 παρουσιάζονται, ενδεικτικά, 3 περιπτώσεις επαύξησης δεδομένων για το ResNet18. Επαύξηση μόνο της κλάσης μειονότητας ώστε να εξισορροπηθεί το train set (πράσινο), εξισορρόπηση του συνόλου και στη συνέχεια πενταπλασιασμός των δεδομένων (μπλε) και εξισορρόπηση του συνόλου και στη συνέχεια διπλασιασμός των δεδομένων (πορτοκαλί). Παρατηρούμε ότι μεγαλύτερη επαύξηση του συνόλου μειώνει σημαντικά τις μετρικές Specificity και MCC. Αυτό δείχνει ότι το σύστημα δεν συμπεριφέρεται καλά στην κλάση μειονότητας. Η μετρική MCC που αποτελεί έναν πολύ καλό δείκτη της συμπεριφοράς του

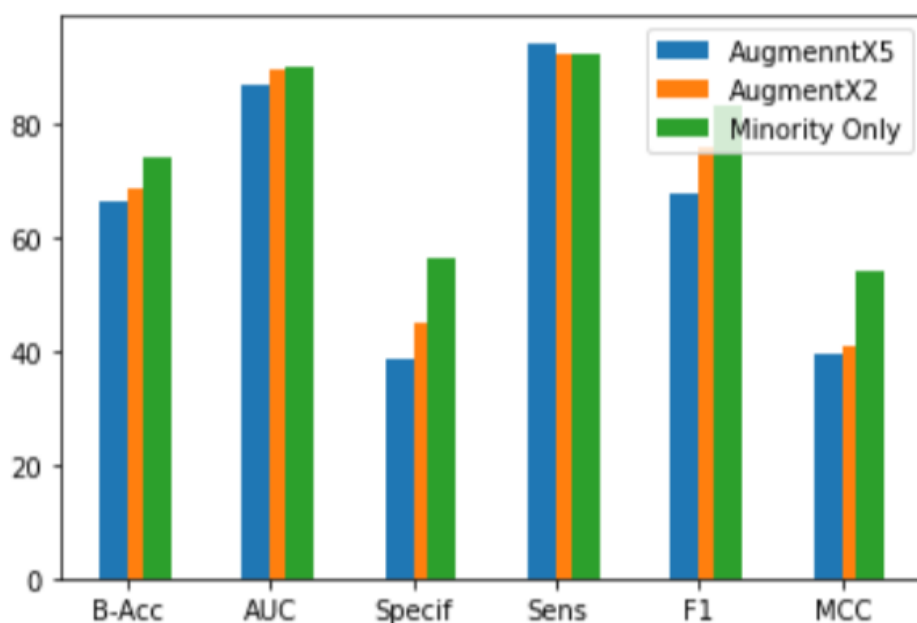


Σχήμα 6.2: Confusion Matrices για τα μοντέλα μόνο με προεκπαίδευση

	VGG16	VGG19
ACC(%)	75.86 ±8.39	83.92 ±6.10
B-ACC(%)	74.23 ±9.34	58.06 ±4.98
AUC(%)	81.89 ±9.69	77.47 ±2.07
Spec(%)	<b>71.30</b> ±7.14	17.53 ±9.13
Sens(%)	77.16 ±8.4	<b>98.58</b> ±1.20
Prec(%)	69.07 ±9.02	8.36 ±12.46
F1(%)	68.68 ±12.38	59.31 ±8.48
MCC	0.4285 ±0.1858	0.2993 ±0.1529

Πίνακας 6.5: Επίδοση unimodal μοντέλων VGG19, VGG16 για τις εικόνες μετά από μεταφορά μάθησης και εξισορρόπηση train set με επαύξηση

μοντέλου είναι σημαντικά καλύτερη στην περίπτωση της απλής εξισορρόπησης.



Σχήμα 6.3: Επίδοση ResNet18 με τρεις τρόπους επαύξησης δεδομένων συνόλου εκπαίδευσης

## 6.3 Αποτελέσματα πολυτροπικών μοντέλων

### 6.3.1 Μοντέλο Feature-Level Fusion

Τα αποτελέσματα των μοντέλων με τη μέθοδο Feature-Level Fusion παρουσιάζονται στους Πίνακες 6.8, 6.7. Παρατηρούμε ότι οι περισσότερες μετρικές έχουν παρουσιάσει βελτίωση σε σχέση με το προηγούμενο βήμα όπως επίσης και οι τυπικές αποκλίσεις, στις περισσότερες περιπτώσεις, έχουν μειωθεί.

	ResNet18	ResNet50	AlexNet
ACC(%)	<b>84.76</b> ±4.47	82.78 ±2.88	83.03 ±30.48
B-ACC(%)	69.09 ±8.07	68.95 ±3.27	73.32 ±9.27
AUC(%)	<b>88.33</b> ±7.86	85.24 ±2.33	86.25 ±6.67
Spec(%)	43.33 ±20.21	46.85 ±8.68	58.33 ±2.56
Sens(%)	<b>97.28</b> ±3.8	91.05 ±2.28	88.31 ±8.26
Prec(%)	76.56 ±11.4	70.50 ±2.18	<b>78.34</b> ±11.07
F1(%)	68.37 ±7.65	68.90 ±2.05	70.87 ±5.56
MCC	0.4290 ±0.1559	0.3908 ±0.0275	0.4852 ±0.0944

Πίνακας 6.6: Επίδοση Feature Fusion μοντέλων ResNet18, ResNet50, AlexNet

### 6.3.2 Μοντέλο Decision-Level Fusion

Η πρώτη παρατήρηση για το μοντέλο Decision Fusion είναι ο μεγάλος χρόνος που χρειάζεται για να τρέξει σε σχέση με το Feature Fusion. Επιπλέον οι επιδόσεις δεν είναι ικανοποιητικές σε σχέση με το Feature Fusion. Ιδιαίτερα η μετρική Specificity κινείται σε

	VGG16	VGG19
ACC(%)	84.54 ±4.41	78.40 ±4.71
B-ACC(%)	<b>77.49</b> ±4.98	65.18 ±7.98
AUC(%)	87.85 ±5.94	78.52 ±6.67
Spec(%)	<b>65.55</b> ±12.22	44.07 ±21.51
Sens(%)	89.42 ±7.29	86.30 ±6.22
Prec(%)	76.90 ±9.15	63.53 ±3.31
F1(%)	<b>75.12</b> ±4.63	62.93 ±5.59
MCC	<b>0.533</b> ±0.0960	0.2830 ±0.1123

Πίνακας 6.7: Επίδοση Feature Fusion μοντέλων με VGG19, VGG16

χαμηλά επίπεδα. Τα unimodal μοντέλα έχουν υψηλότερες επιδόσεις στις εικόνες από ότι στα κλινικά χαρακτηριστικά καθώς οι εικόνες παρέχουν τις πληροφορίες για τις ίδιες τις πλάκες. Επίσης, η κατηγοριοποίηση των ασθενών σε Υψηλού και Χαμηλού κινδύνου εξαρτάται κατά κύριο λόγο από το μέγεθος, τη μορφή και την υφή της πλάκας που προκύπτουν από τα υπερηχογραφήματα. Επομένως, επειδή στο μοντέλο Decision Fusion χρησιμοποιούμε με τον ίδιο βαθμό συμμετοχής τις προβλέψεις των δύο μοντέλων, οι θετικές επιδόσεις του μοντέλου από τις εικόνες φαίνεται ότι επισκιάζονται από τις χαμηλότερες επιδόσεις στα κλινικά χαρακτηριστικά. Τα αποτελέσματα παρουσιάζονται στους παρακάτω πίνακες.

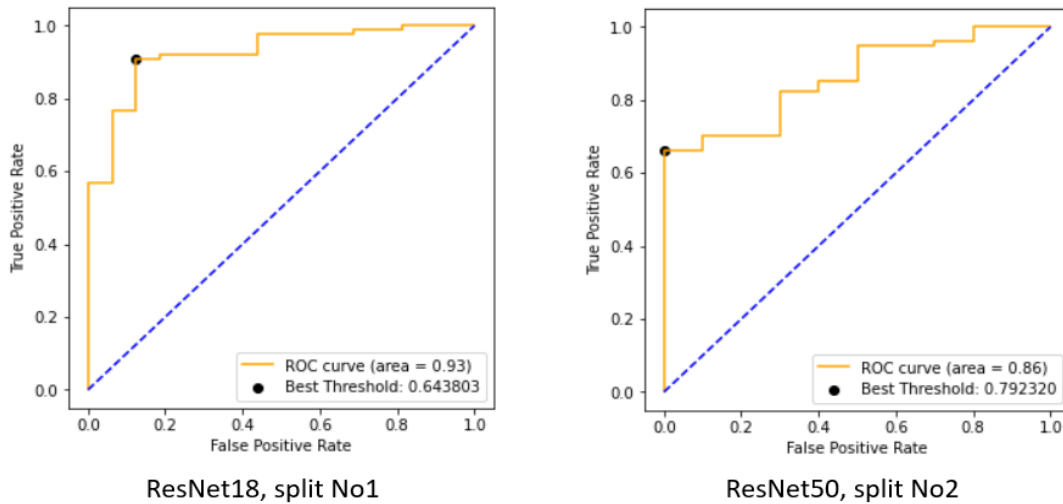
	ResNet18	VGG16	AlexNet
ACC(%)	<b>81.43</b> ±1.31	65.69 ±12.22	59.76 ±29.63
B-ACC(%)	57.93 ±4.82	<b>59.41</b> ±10.75	48.84 ±2.14
AUC(%)	<b>65.88</b> ±8.49	71.61 ±10.11	41.76 ±6.26
Spec(%)	21.38 ±6.87	<b>48.41</b> ±23.45	33.33 ±4.71
Sens(%)	<b>94.44</b> ±2.27	70.42 ±17.08	63.63 ±4.58
Prec(%)	<b>64.64</b> ±3.58	55.85 ±13.63	30.71 ±15.09
F1(%)	<b>58.28</b> ±5.76	53.28 ±14.76	34.96 ±13.25
MCC	<b>0.2094</b> ±0.0857	0.1489 ±0.2830	-0.0520 ±0.0736

Πίνακας 6.8: Επίδοση Decision Fusion μοντέλων ResNet18, VGG16, AlexNet

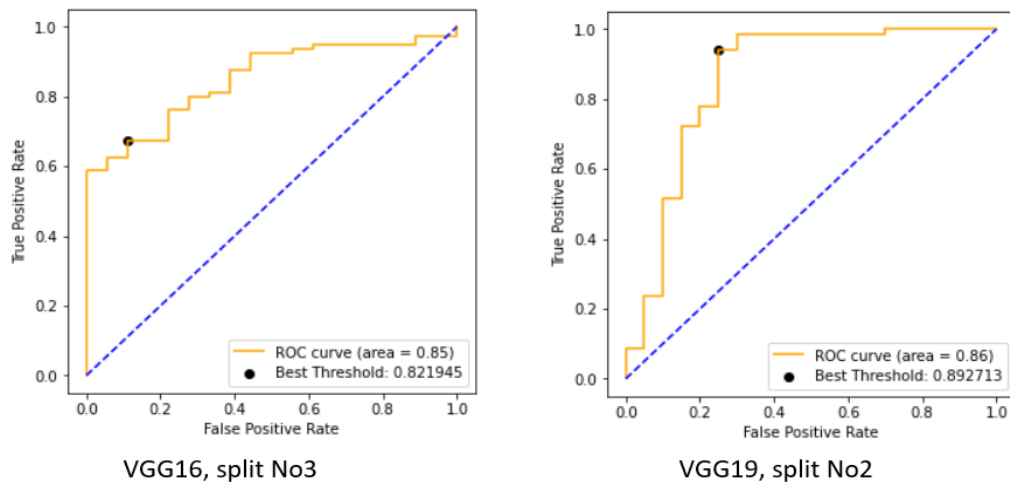
Ενδεικτικά στο Σχήμα 6.5 παραθέτουμε και κάποια από τα διαγράμματα της καμπύλης ROC σε διάφορα splits με τη μέθοδο Feature-Level Fusion.

## 6.4 Σύνοψη αποτελεσμάτων

Στο Feature-Level Fusion η συγχώνευση των κλινικών χαρακτηριστικών με τα χαρακτηριστικά που εξάχθηκαν από τα υπερηχογραφήματα ενίσχυσαν τις προβλέψεις που γίνονταν μόνο με τις εικόνες. Καθώς οι απεικονιστικές μέθοδοι είναι πιο καθοριστικές για την κατηγο-



Σχήμα 6.4: Ενδεικτικά διαγράμματα ROC με Feature-Level Fusion



Σχήμα 6.5: Ενδεικτικά διαγράμματα ROC με Feature-Level Fusion

ριοποίηση των ασθενών σε υψηλού και χαμηλού κινδύνου, διότι παρέχουν την πληροφορία για τη γεωμετρία και την υφή της αθηρωματικής πλάκας καθώς και για το βαθμό στένωσης, ο ισότιμος συμψηφισμός των προβλέψεων από τα δύο modalities στο Decision Fusion υποβαθμίζει τη συνολική επίδοση του συστήματος. Ακόμη, παρατηρούμε ότι σε αρκετές περιπτώσεις η τυπική απόκλιση των μετρικών είναι σχετικά μεγάλη το οποίο εξηγείται από το γεγονός ότι σε ορισμένα splits η ποιότητα των δειγμάτων σε συνδυασμό με το μικρό τους πλήθος αποτρέπει την καλή εκπαίδευση του μοντέλου.



## Κεφάλαιο **7**

# Συμπεράσματα και Συζήτηση

---

### 7.1 Συμπεράσματα

Η καρωτιδική νόσος, δηλαδή η αθηροσκλήρωση των καρωτιδικών αρτηριών από μόνη της αλλά και σε συνδυασμό με τη στεφανιαία νόσο και τις άλλες αγγειακές παθήσεις συνιστούν ένα σημαντικό παγκόσμιο υγειονομικό πρόβλημα. Με την εξέλιξη της τεχνολογίας έχει γίνει σημαντική πρόοδος στις απεικονιστικές μεθόδους (MRI, CT, PET) και μεταξύ αυτών σημαντική θέση κατέχουν τα υπερηχογραφήματα (US, CEUS, INVUS) τα οποία έχουν τη δυνατότητα απεικόνισης πολλών παραμέτρων που αφορούν τον σχηματισμό, την εξέλιξη και τις επιπτώσεις των αθηρωματικών πλακών στις αρτηρίες και μεταξύ αυτών στις καρωτίδες.

Τα τελευταία χρόνια οι εφαρμογές AI γνωρίζουν σημαντική ανάπτυξη στο τομέα της υγείας. Ένας από τους πλέον εξελισσόμενους τομείς είναι η βαθιά μάθηση (deep learning) και τα συνελκτικά νευρωνικά δίκτυα. Η ανάπτυξη εφαρμογών AI και βαθιάς μάθησης στην ταξινόμηση του είδους και της επικινδυνότητας της αθηρωματικής πλάκας στις καρωτίδες απασχολεί ικανό αριθμό επιστημονικών ομάδων και έχουν παρουσιαστεί αρκετές σχετικές εργασίες. Μία από τις σχετικά πρόσφατε εξελίξεις στην ταξινόμηση αυτή είναι ο συνδυασμός πολυτροπικών δεδομένων δηλαδή δεδομένων από διαφορετικές πηγές και διαφορετικών μορφών.

Στην παρούσα εργασία αναπτύχθηκε κώδικας για τη χρήση πέντε δημοφιλών μοντέλων Deep Learning σε συνδυασμό με transfer learning για την ταξινόμηση σε δύο βαθμίδες (Υψηλή, Χαμηλή), της επικινδυνότητας των αθηρωματικών πλακών σε δείγμα 75 ασθενών και 268 εικόνων. Για κάθε ασθενή αξιοποιήθηκαν και τα διαθέσιμα κλινικά δεδομένα και δεδομένα εργαστηριακών εξετάσεων. Επίσης αναπτύχθηκε κώδικας στη γλώσσα Python για δύο (2) μοντέλα βαθιάς μάθησης που υλοποιούν την ταξινόμηση με συνένωση χαρακτηριστικών από τις εικόνες, τις κλινικές παραμέτρους και τις εξετάσεις. Στο 1ο μοντέλο η συνένωση γίνεται με τη μέθοδο Feature-Level Fusion και στο 2ο μοντέλο με την Decision-Level Fusion. Ως εξαγωγείς χαρακτηριστικών από τις εικόνες στα δύο αυτά μοντέλα δοκιμάστηκαν τα συνελκτικά μοντέλα βαθιάς μάθησης ResNet18, ResNet50, AlexNet, VGG16 και VGG19. Τα πέντε(5) μοντέλα, ResNet18, ResNet50, AlexNet, VGG16 και VGG19 με μόνο προ-εκπαίδευση από την ανοικτή βάση εικόνων ImageNet έδωσαν σχετικά ικανοποιητικά αποτελέσματα αναλογικά με το μικρό αριθμό δεδομένων και το γεγονός ότι ήταν μη-ισορροπημένα (11 περιπτώσεις χαμηλού κινδύνου και 64 υψηλού). Ειδικότερα καταγράφηκαν καλές επιδόσεις στον εντοπισμό των επιδόσεων που ανήκουν στην πλειοψηφούσα τάξη. Τελικά, το 1ο πολυτροπικό

μοντέλο υπερίσχυσε έναντι του 2ου αλλά και των αντίστοιχων μονοτροπικών πετυχαίνοντας Balanced Accuracy 77%, AUC 87%, Precision 76%, F1 Score 75% και Matthews Correlation Coefficient (MCC) 0.533.

Τα μοντέλα με τη συνένωση δεδομένων έδωσαν αποτελέσματα μικτής ανάγνωσης. Σε κάποια σημεία βελτιώθηκαν επιδόσεις (Ακρίβεια, ευαισθησία, κλπ) και σε άλλα όχι. Διαφάνεται ότι είναι καθοριστική η βαρύτητα της κάθε παραμέτρου στην απόφαση ταξινόμησης καθώς και η διακύμανση των διαφόρων μεγεθών στο δείγμα εκπαίδευσης, το μέγεθος του δείγματος και η ισορροπία του δείγματος. Σαν συνέπεια καθιστά πολύ σημαντική τη δυνατότητα ερμηνείας της κάθε απόφασης και την ανάπτυξη και χρήση σχετικών εργαλείων που ήδη χρησιμοποιούνται σε κάποιες περιπτώσεις.

Συμπερασματικά, η χρήση πολυτροπικών μοντέλων στην καρωτιδική νόσο φαίνεται να έχει δυναμικό προγνωστικής αξίας και χρήζει περαιτέρω διερεύνησης και επεξεργασίας.

## 7.2 Μελλοντικές προεκτάσεις

Καθώς το θέμα είναι σύγχρονο και αφορά σε ένα πολύ σημαντικό κομμάτι της καθημερινότητας του ανθρώπου, αυτό της υγείας, οι επεκτάσεις της παρούσας εργασίας είναι πολλές. Θα περιγράψουμε κάποιες από αυτές.

Αρχικά, όπως είδαμε, μία σημαντική δυσκολία που συναντήσαμε ήταν το περιορισμένο πλήθος των διαθέσιμων δεδομένων, ειδικά από την κατηγορία ασθενών χαμηλού κινδύνου. Γι'αυτό κρίνεται απαραίτητο να αναζητηθούν και άλλα δεδομένα. Για αυτό το σκοπό μπορούν να αξιοποιηθούν, έπειτα από σχετικό αίτημα και κατάλληλη προ επεξεργασία, και δεδομένα όπως αυτά που χρησιμοποιούνται στην [102] και περιλαμβάνουν πλήρεις υπερηχογραφικές εξετάσεις B-mode και χαρακτηριστικά όπως ηλικία και φύλο, υπέρταση, διαβήτης, ιστορικό καπνίσματος, γλυκόζη, HDL χοληστερόλη, γ-GLT, περίμετρος της μέσης, καρδιακός ρυθμός. Παράλληλα θα ήταν ενδιαφέρον να διερευνηθούν και άλλες τεχνικές επαύξησης και εξισορρόπησης των δεδομένων οι οποίες και να συγκριθούν με αυτές που δοκιμάστηκαν στην παρούσα εργασία.

Μία δεύτερη πιθανή επέκταση είναι η χρήση hand-crafted χαρακτηριστικών από τις εικόνες σαν επιπλέον είσοδο στο πολυτροπικό μοντέλο. Τέτοια χαρακτηριστικά είναι δείκτες που εξάγονται από τις υπερηχογραφικές εικόνες όπως για παράδειγμα ο δείκτης JBA (Juxtaluminal Black Area) και δείκτες που αφορούν στην υφή των πλάκων και της καρωτίδας όπως ο IMT (intima media thickness). Από τα υπερηχογραφήματα που περιλαμβάνονται στο σύνολο δεδομένων της παρούσας εργασίας έχουν εξαχθεί τέτοια χαρακτηριστικά σε προηγούμενες εργασίες του εργαστηρίου [103], [104]. Στην [105] οι Min Kyoung KangI et al. ανέπτυξαν μοντέλα πρόβλεψης του κινδύνου για ισχαιμικό εγκεφαλικό επεισόδιο συνδυάζοντας υπερηχογραφικές παραμέτρους και συμβατικές βαθμολογίες αξιολόγησης του καρδιαγγειακού κινδύνου (Framingham risk score (FRS) και Atherosclerotic Cardiovascular Disease (ASCVD) score).

Έχει δειχθεί ότι η κινητικότητα και η δυναμική της πλάκας και της ροής του αίματος στην περιοχή διαδραματίζουν καθοριστικό ρόλο στο πόσο πιθανόν είναι η πλάκα να διαρραγεί προκαλώντας θρομβώσεις. Συνεπώς θα είχε ενδιαφέρον αντί για τις εικόνες υπερήχων που απομονώθηκαν από τα υπερηχογραφήματα να χρησιμοποιηθούν ως είσοδο στα μοντέλα



ολόκληρα τα βίντεο υπερήχων στα οποία αποτυπώνεται και η κίνηση. Τα απλά (δισδιάστατα) συνελικτικά επίπεδα που χρησιμοποιήσαμε στην παρούσα εργασία θα αντικατασταθούν με τρισδιάστατα. Επιπλέον νέα μοντέλα της όρασης υπολογιστών, πιο κατάλληλα για τα βίντεο, θα μπορούσαν να αναζητηθούν και να δοκιμαστούν.

Παράλληλα, καθώς το πλήθος των διαθέσιμων συμπεριφορικών, εργαστηριακών και κλινικών χαρακτηριστικών είναι μικρό σε σχέση με το πλήθος των εξαγόμενων χαρακτηριστικών από τις εικόνες υπερήχων κι αυτό όπως είδαμε πιθανώς να επηρεάζει αρνητικά την επίδοση των Feature-Level Fusion μοντέλων θα ήταν χρήσιμο να διερευνηθούν μέθοδοι επαύξησης των κλινικών χαρακτηριστικών όπως αυτή του [56] την οποία σκιαγραφήσαμε σε προηγούμενο κεφάλαιο αλλά και άλλες μέθοδοι χειρισμού αυτού του θέματος όπως η επιλογή βαρών για κάθε modality. Στο [61] υλοποιείται μέθοδος όπου κατά την εκπαίδευση του πολυτροπικού νευρωνικού δικτύου υπολογίζονται και χρησιμοποιούνται ξεχωριστά τα σφάλματα για το κάθε modality.

Κατόπιν, θα ήταν ενδιαφέρον να διερευνηθούν τεχνικές απεικόνισης των δεδομένων της παρούσας εργασίας και των αλληλεπιδράσεών τους με τη μορφή γραφήματος ώστε να χρησιμοποιηθούν μοντέλα GCN όπως στο [16] που είδαμε, τα οποία έχουν κεντρίσει έντονα το ενδιαφέρον της ερευνητικής κοινότητα τα τελευταία χρόνια.

Τέλος, θα είχε πολύ ενδιαφέρον να διερευνηθεί περαιτέρω ο μετασχηματισμός των δεδομένων πίνακα (εργαστηριακά και κλινικά χαρακτηριστικά) σε εικόνες και έπειτα οι εικόνες αυτές να χρησιμοποιηθούν από το πολυτροπικό μοντέλο σε συνδυασμό με τα υπερηχογραφήματα. Πολυτροπικά μοντέλα με modalities διαφορετικές κατηγορίες εικόνων έχουν παρουσιαστεί σε αρκετές εργασίες όπως για παράδειγμα στην [106]. Στην παρούσα εργασία έγινε προσπάθεια προς αυτή την κατεύθυνση η οποία όμως δεν είχε τα αναμενόμενα αποτελέσματα ενδεχομένως λόγω του μικρού πλήθους των χαρακτηριστικών των δεδομένων πίνακα. Προς αυτή την κατεύθυνση θα μπορούσαν να δοκιμαστούν κι άλλοι αλγόριθμοι μετατροπής των δεδομένων πίνακα σε εικόνες όπως αυτός που περιγράφεται στο [107].



# Παραρτήματα

---



## Συμπληρωματικοί πίνακες

---

### **A'.1 Απεικονιστικές μέθοδοι για την καρωτίδα**

Στον πίνακα του Σχήματος **A'.1** του παρουσιάζονται τα πλεονεκτήματα και τα μειονεκτήματα των βασικών απεικονιστικών μεθόδων για την καρωτίδα.

### **A'.2 Κατηγοριοποίηση συστάσεων της ESVS**

Η Ευρωπαϊκή Ένωση για την Αγγειοχειρουργική (European Society for Vascular Surgery) κατηγοριοποιεί τις συστάσεις της με βάση τον παρακάτω πίνακα Σχήμα **A'.2**.

- Κλάση I : Αποδεικτικά στοιχεία ή/και γενική συμφωνία ότι μια δεδομένη θεραπεία ή διαδικασία είναι ωφέλιμη, χρήσιμη, αποτελεσματική
- Κλάση II : Αντικρουόμενα στοιχεία και/ή διάσταση απόψεων σχετικά με τη χρησιμότητα/αποτελεσματικότητα της συγκεκριμένης θεραπείας ή διαδικασίας
- Κλάση IIa : Το βάρος των αποδεικτικών στοιχείων/γνωμοδότηση είναι υπέρ της χρησιμότητας/αποτελεσματικότητας
- Κλάση IIβ: Η χρησιμότητα/αποτελεσματικότητα είναι λιγότερο καλά τεκμηριωμένη από στοιχεία/γνωμοδότηση
- Κλάση III : Αποδεικτικά στοιχεία ή γενική συμφωνία ότι η δεδομένη θεραπεία ή διαδικασία δεν είναι χρήσιμη/αποτελεσματική και σε ορισμένες περιπτώσεις μπορεί να είναι επιβλαβής

Μέθοδος	Εφαρμογή	Πλεονεκτήματα	Μειονεκτήματα
Υπερηχογράφημα - Ultrasound	<ul style="list-style-type: none"> <li>• προσυμπτωματικός έλεγχος</li> <li>• αναγνώριση μορφολογίας πλάκας</li> <li>• ποσοτικοποίηση στένωσης</li> <li>• follow-up</li> </ul>	<ul style="list-style-type: none"> <li>• μη-επεμβατική</li> <li>• χωρίς ακτινοβολία</li> <li>• φθηνή</li> <li>• φορητότητα συσκευών</li> <li>• παρέχει αιμοδυναμικές πληροφορίες (π.χ. έγχρωμο Doppler)</li> </ul>	<ul style="list-style-type: none"> <li>• εξαρτάται από τον εξεταστή</li> <li>• acoustic shadowing</li> </ul>
Υπολογιστική αγγειογραφία - (CT- Angiography)	<ul style="list-style-type: none"> <li>• επιβεβαίωση ύπαρξης στένωσης</li> <li>• ποσοτικοποίηση της στένωσης σε περιπτώσεις που αποτυγχάνει το υπερηχογράφημα (π.χ. ασβεστοποιημένες βλάβες)</li> <li>• σχεδιασμός επεμβατικών και χειρουργικών επεμβάσεων</li> <li>• οπτικοποίηση ενδοκρανιακών αγγείων</li> </ul>	<ul style="list-style-type: none"> <li>• γρήγορη λήψη εικόνας</li> <li>• επιτρέπει την αξιολόγηση μη αγγειακών δομών</li> <li>• απεριόριστες γωνίες θέασης</li> <li>• οι σπονδυλικές αρτηρίες μπορούν πάντα να οπτικοποιηθούν</li> <li>• η ποιότητα της απεικόνισης δεν εξαρτάται από τον εξεταστή</li> </ul>	<ul style="list-style-type: none"> <li>• έκθεση σε ακτινοβολία</li> <li>• πιθανές παρενέργειες από σκιαγραφικό</li> <li>• ποσοτικοποίηση της στένωσης μόνο με βάση τη μορφολογία</li> <li>• οι εκτεταμένες ασβεστώσεις δημιουργούν πρόβλημα</li> </ul>
Μαγνητική αγγειογραφία - (MRI- Angiography)	<ul style="list-style-type: none"> <li>• επιβεβαίωση ύπαρξης στένωσης</li> <li>• ποσοτικοποίηση της στένωσης σε περιπτώσεις που αποτυγχάνει το υπερηχογράφημα (π.χ. ασβεστοποιημένες βλάβες)</li> <li>• σχεδιασμός επεμβατικών και χειρουργικών επεμβάσεων</li> <li>• οπτικοποίηση ενδοκρανιακών αγγείων</li> </ul>	<ul style="list-style-type: none"> <li>• χωρίς ακτινοβολία</li> <li>• απεικόνιση και μη αγγειακών δομών</li> <li>• υψηλή ανάλυση εικόνων</li> </ul>	<ul style="list-style-type: none"> <li>• μικρή διαθεσιμότητα</li> <li>• υψηλό κόστος</li> <li>• μη φορητή συσκευή</li> <li>• χρονοβόρα εξέταση</li> <li>• ευαισθησία στα artifacts</li> <li>• πιθανές παρενέργειες από σκιαγραφικό</li> </ul>
Ψηφιακή Αγγειογραφία Αφαίρεσης - Digital subtraction Angiography	<ul style="list-style-type: none"> <li>• χρησιμοποιείται κατά τη διάρκεια των επεμβάσεων</li> </ul>	<ul style="list-style-type: none"> <li>• απεικόνιση ολόκληρου του συστήματος των καρωτιδικών αρτηριών,</li> <li>• υψηλή ανάλυση</li> <li>• απεικόνιση μικρών αγγείων</li> </ul>	<ul style="list-style-type: none"> <li>• περιορισμένη χωρική ανάλυση</li> <li>• επεμβατική (κίνδυνος εγκεφαλικού)</li> <li>• πιθανές παρενέργειες από σκιαγραφικό</li> <li>• έκθεση σε ακτινοβολία (υψηλότερη από την CT αγγειογραφία)</li> </ul>

Σχήμα Α.1: Πλεονεκτήματα και μειονεκτήματα απεικονιστικών μεθόδων καρωτίδας

<b>Class I</b>	Evidence and/or general agreement that a given treatment or procedure is beneficial, useful, effective
<b>Class II</b>	Conflicting evidence and/or a divergence of opinion about the usefulness/efficacy of the given treatment or procedure
<b>Class IIa</b>	<i>Weight of evidence/opinion is in favour of usefulness/efficacy</i>
<b>Class IIb</b>	<i>Usefulness/efficacy is less well established by evidence/opinion</i>
<b>Class III</b>	Evidence or general agreement that the given treatment or procedure is not useful/effective, and in some cases may be harmful

Σχήμα Α.2: ESVS: Κατηγοριοποίηση των συστάσεων

Stenosis (%)	Recommendations	Level of recommendations
<i>Symptomatic stenosis</i>		
High grade (≥70%)	CEA by a surgeon with perioperative mortality rate <6%	Class I Level of evidence A
Moderate (≥50 and <70%)	CEA, depending on patient-specific factors such as age, sex, comorbidities and severity of symptoms	Class I Level of evidence A
Mild (<50%)	No indications for CEA	Class I Level of evidence A
<i>Asymptomatic stenosis</i>		
High grade (≥60%)	CEA when performed by a surgeon with a perioperative mortality rate of <3%	Class I Level of evidence A

Σχήμα Α.3: Συστάσεις των American Heart Association και American Stroke Association σε σχέση με τη ενδαρτηρεκτομή για την αντιμετώπιση της συμπτωματικής στένωσης της καρωτίδας.





## Βιβλιογραφία

---

- [1] *Major Arteries of the Head and Neck*. <https://teachmeanatomy.info/neck/vessels/arterial-supply/>. Ημερομηνία πρόσβασης: 31-10-2022.
- [2] <https://en.wikipedia.org/wiki/Artery>. Ημερομηνία πρόσβασης: 31-10-2022.
- [3] Virani SS, Alonso A., Benjamin EJ και et al. *Heart disease and stroke statistics-2020 update: a report from the American Heart Association*. *Circulation*, 2020.
- [4] *CAROTID ULTRASOUND WEBBOOK FREE WIKI*. <https://123sonography.com/carotid-ultrasound-webbook>. Ημερομηνία πρόσβασης: 31-10-2022.
- [5] Siegel Axel D., Daub K., Seizer P., Lindemann S. και Gawaz M. *Platelet lipoprotein interplay: Trigger of foam cell formation and driver of atherosclerosis*. *Cardiovasc. Res.*, 2008.
- [6] Donnan GA, Davis SM, Chambers BR και Gates PC. *Surgery for prevention of stroke*. *Lancet*, 1998.
- [7] A.R. Nayloret al. *Management of Atherosclerotic Carotid and Vertebral Artery Disease: 2017 Clinical Practice Guidelines of the European Society for Vascular Surgery (ESVS)*. *Ultrasound Med Biol.*, 2017.
- [8] Saba L, Sanagala SS, Gupta SK, Koppula VK, Johri AM, Khanna NN, Mavrogeni S, Laird JR, Pareek G, Miner M, Sfrikakis PP, Protogerou A, Misra DP, Agarwal V, Sharma AM, Viswanathan V, Rathore VS, Turk M, Kolluri R, Viskovic K, Cuadrado Godia E, Kitas GD, Sharma N, Nicolaidis A και Suri JS. *Multimodality carotid plaque tissue characterization and classification in the artificial intelligence paradigm: a narrative review for stroke application*. *Ann Transl Med* 2021, 2021.
- [9] Zhenzhu Meng, Yating Hu και Christophe Ancey. *Approach to Predict Waves Generated by Gravity Driven Mass Flows*. *Water*, 2020.
- [10] Sanagala SS, Nicolaidis A, Gupta SK, Koppula VK, Saba L, Agarwal S, Johri AM, Kalra MS και Suri JS. *Ten Fast Transfer Learning Models for Carotid Ultrasound Plaque Tissue Characterization in Augmentation Framework Embedded with Heatmaps for Stroke Risk Stratification*. *Diagnostics*, 2019.
- [11] *A Leap into the Future: Generative Adversarial Networks*. <https://medium.datadriveninvestor.com/a-leap-into-the-future-generative-adversarial-networks-96a780ed8ee6>. Ημερομηνία πρόσβασης: 31-10-2022.

- [12] Theofanis Ganitidis, Maria Athanasiou, Kalliopi Dalakleidi, Nikos Melanitis, Spyretta Golemati και Konstantina S Nikita. *Stratification of carotid atheromatous plaque using interpretable deep learning methods on B-mode ultrasound images*. 2021.
- [13] SC. Huang, A. Pareek και S. et al. Seyyedi. *Fusion of medical imaging and electronic health records using deep learning: a systematic review and implementation guidelines*. 2020.
- [14] Michelle Tang, Pulkit Kumar, Hao Chen και Abhinav Shrivastava. *Deep Multimodal Learning for the Diagnosis of Autism Spectrum Disorder*. *Journal of Imaging*, 2020.
- [15] Chen M Wang J Altaye M Dillman JR Parikh NA He L, Li H. *Deep Multimodal Learning From MRI and Clinical Data for Early Prediction of Neurodevelopmental Deficits in Very Preterm Infants*. *Front Neurosci*, 2021.
- [16] S. Parisot et al. *Disease prediction using graph convolutional networks: Application to Autism Spectrum Disorder and Alzheimer's disease*. *Medical Image Analysis*, 2021.
- [17] Yi Liu, Hsueh Ping Lu και Ching Hao Lai. *A Novel Attention-Based Multi-Modal Modeling Technique on Mixed Type Data for Improving TFT-LCD Repair Process*. *BMC Medical Informatics and Decision Making*, 2022.
- [18] *An overview of VGG16 and NiN models*. <https://medium.com/mlearning-ai/an-overview-of-vgg16-and-nin-models-96e4bf398484>. Ημερομηνία πρόσβασης: 31-10-2022.
- [19] Yufeng Zhenget al. *Breast cancer screening using convolutional neural network and follow-up digital mammography*. 2018.
- [20] Kaiming He, Xiangyu Zhang, Shaoqing Ren και Jian Sun. *Deep Residual Learning for Image Recognition*. 2015.
- [21] Faheem Khan et al. *IR-UWB Radar-Based Robust Heart Rate Detection Using a Deep Learning Technique Intended for Vehicular Applications*. 2022.
- [22] Valery L Feigin, Michael Brainin, Bo Norrving, Sheila Martins, Ralph L Sacco, Werner Hacke, Marc Fisher, Jeyaraj Pandian και Patrice Lindsay. *World Stroke Organization (WSO): Global Stroke Fact Sheet 2022*. *International Journal of Stroke* 2022, 2022.
- [23] Κωνσταντίνος Βαδικόλιας: καθηγητής Νευρολογίας . *Ελληνική Συμμαχία- Συνεργασία για τα Αγγειακά Εγκεφαλικά Επεισόδια*, συνέδριο Ευρωπαϊκής Συμμαχίας για το Εγκεφαλικό . Μεταπτυχιακή διπλωματική εργασία, Θεσσαλονίκη , 2022.
- [24] *World Health Organization: Cardiovascular diseases (CVDs)*. [https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)). Ημερομηνία πρόσβασης: 31-10-2022.

- [25] Σπύρος Σαμαράς, Ιατρός Καρδιολόγος . *Η θέση της υπερηχοτομογραφικής μεθόδου στη λήψη απόφασης για στεντυγ των καρωτίδων*. Μεταπτυχιακή διπλωματική εργασία, Πανεπιστήμιο Θεσσαλίας, 2017.
- [26] Solberg LA και Strong JP. *Risk factors and atherosclerotic lesions. A review of autopsy studies*. *Arteriosclerosis*. 1983.
- [27] Wei Young, John W. Gofman, Robert Tandy, Nathan Malamud και Eunice S.G. Waters. *The quantitation of atherosclerosis: III. The extent of correlation of degrees of atherosclerosis within and between the coronary and cerebral vascular beds*. *The American Journal of Cardiology*, σελίδες 300–308, 1960.
- [28] Khoury Zet al. *Relation of coronary artery disease to atherosclerotic disease in the aorta, carotid, and femoral arteries evaluated by ultrasound*. 1997.
- [29] Rothwell PM, Villagra R, Gibson R, Donders RC και Warlow CP. *Evidence of a chronic systemic cause of instability of atherosclerotic plaques*. 2000.
- [30] Naghavi Met al. *From vulnerable plaque to vulnerable patient: A call for new definitions and risk assessment strategies: Part I*. 2003.
- [31] M et al. Zamani. *Advanced ultrasound methods in assessment of carotid plaque instability: a prospective multimodal study*. 2020.
- [32] A. et al. Gastounioti. *A novel computerized tool to stratify risk in carotid atherosclerosis using kinematic features of the arterial wall*. 2014.
- [33] Tsao CWet al. on behalf of the American Heart Association Council on Epidemiology, Prevention Statistics Committee και Stroke Statistics Subcommittee. H. *Heart disease and stroke statistics—2022 update: a report from the American Heart Association*. *Circulation*, 2022.
- [34] Tsao CWet al. *World Health Organization: Cardiovascular diseases (CVDs)*. 2021.
- [35] S. Jebari-Benslaiman, U. Galicia-García, A. Larrea-Sebal, J.R. Olaetxea, I. Alloza, K. Vandenbroeck και C Benito-Vicente, A.and Martín. *Pathophysiology of Atherosclerosis*. *Int. J. Mol. Sci*, 2022.
- [36] Magnus Bäck και Goran K. Hansson. *Basic Mechanisms of Atherosclerosis*. *Int. J. Mol. Sci*, 2022.
- [37] G. Cismaru, T. Serban και Tirpe A. *Ultrasound Methods in the Evaluation of Atherosclerosis: From Pathophysiology to Clinic*. *Biomedicines*, 2021.
- [38] Π.Γ. Βακαλόπουλος “Στένωση Καρωτίδας”. <https://aggeiologia.gr/stenosi-karotidas/>. Ημερομηνία πρόσβασης: 31-10-2022.
- [39] Wang X, Jackson DC, Mitchell CC, Varghese T, Wilbrand SM, Rocque BG, Hermann BP και Dempsey RJ. *Classification of Symptomatic and Asymptomatic Patients with*

- and without Cognitive Decline Using Noninvasive Carotid Plaque Strain Indices as Biomarkers. Ultrasound Med Biol.*, 2016.
- [40] Bentzon JF, Otsuka F, Virmani R και Falk E. *Mechanisms of plaque formation and rupture. Circ Res.*, 2014.
- [41] “Carotid angioplasty and stenting, Mayo Clinic”. <https://www.mayoclinic.org/tests-procedures/carotid-angioplasty-and-stenting/about/pac-20385111>. Ημερομηνία πρόσβασης: 29-09-2022.
- [42] L. Sabaet al. *A Multicenter Study on Carotid Ultrasound Plaque Tissue Characterization and Classification Using Six Deep Artificial Intelligence Models: A Stroke Application. IEEE Transactions on Instrumentation and Measurement*, 70, 2021.
- [43] *Britanica*. <https://www.britannica.com/technology/artificial-intelligence>. Ημερομηνία πρόσβασης: 31-10-2022.
- [44] [https://en.wikipedia.org/wiki/Machine\\_learning#cite\\_note-cyberthreat-13](https://en.wikipedia.org/wiki/Machine_learning#cite_note-cyberthreat-13). Ημερομηνία πρόσβασης: 29-09-2022.
- [45] C. Janiesch, P. Zschech και K. Heinrich. *Machine learning and deep learning. Electron Markets*, 2021.
- [46] [https://www.sas.com/en\\_gb/insights/articles/analytics/machine-learning-algorithms.html](https://www.sas.com/en_gb/insights/articles/analytics/machine-learning-algorithms.html). Ημερομηνία πρόσβασης: 29-09-2022.
- [47] [https://classic.d2l.ai/chapter\\_multilayer-perceptrons/dropout.html](https://classic.d2l.ai/chapter_multilayer-perceptrons/dropout.html). Ημερομηνία πρόσβασης: 29-09-2022.
- [48] Costa P, Santos NC, Cunha P, Cotter J και Sousa N. *The Use of Multiple Correspondence Analysis to Explore Associations between Categories of Qualitative Variables in Healthy Ageing. J Aging Res*, 2013.
- [49] W. Satriaaji και R. Kusumaningrum. *Effect of Synthetic Minority Oversampling Technique (SMOTE), Feature Representation, and Classification Algorithm on Imbalanced Sentiment Analysis. ICICoS*, 2018.
- [50] Lei Xu, Maria Skoularidou, Alfredo Cuesta-Infante και Kalyan Veeramachaneni. *Modeling Tabular data using Conditional GAN. NeurIPS*, 2019.
- [51] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville και Yoshua Bengio. *Generative Adversarial Nets. arXiv:1406.2661*, 2014.
- [52] Chris Russell Michael U Gutmann Akash Srivastava, Lazar Valkov και Charles Sutton. *Veegan: Reducing mode collapse in gans using implicit variational learning. Neural Information Processing Systems*, 2017.

- [53] Chris Russell Michael U Gutmann Akash Srivastava, Lazar Valkov και Charles Sutton. *ADASYN: Adaptive synthetic sampling approach for imbalanced learning. IEEE World Congress on Computational Intelligence*, 2008.
- [54] David Opitz και Richard Maclin. *Popular Ensemble Methods: An Empirical Study*. 2011.
- [55] Can Cui, Haichun Yang, Yaohong Wang, Shilin Zhao, Zuhayr Asad, Lori A. Coburn, Keith T. Wilson, Bennett A. Landman και Yuankai Huo. *Deep Multi-modal Fusion of Image and Non-image Data in Disease Diagnosis and Prognosis: A Review*. 2022.
- [56] R. Yan et al. *Richer fusion network for breast cancer classification based on multimodal data. BMC Medical Informatics and Decision Making*, 2021.
- [57] J. Kawahara, C. J. Brown, S. P. Miller, Chau V. Booth, B. G., R. E. Grunau και et al. *BrainNetCNN: convolutional neural networks for brain networks; towards predicting neurodevelopment*. 2017.
- [58] Pankaj K Jain, Abhishek Dubey, Luca Saba, Narender N Khanna, John R Laird, Andrew Nicolaides, Mostafa M Fouda, Jasjit S Suri και Neeraj Sharma. *Attention-Based UNet Deep Learning Model for Plaque Segmentation in Carotid Ultrasound for Stroke Risk Stratification: An Artificial Intelligence Paradigm. Journal of Cardiovascular Development and Disease*, 2022.
- [59] Shih-Cheng Huang, Anuj Pareek, Roham Zamanian, Imon Banerjee και Matthew P. Lungren. *Multimodal fusion with deep neural networks for leveraging CT imaging and electronic health record: a case-study in pulmonary embolism detection*. 2020.
- [60] Esra Zihniet et al. *Multimodal Fusion Strategies for Outcome Prediction in Stroke*. 2020.
- [61] Gregory Holste, Savannah C. Partridge, Habib Rahbar, Debosmita Biswas, Christoph I. Lee και Adam M. Alessio. *End-to-End Learning of Geoffrey E. Hinton Fused Image and Non-Image Features for Improved Breast Cancer Classification from MRI*. 2021.
- [62] S.S. Yadav και S.M. Jadhav. *Deep convolutional neural network based medical image classification for disease diagnosis. J Big Data*, 2019.
- [63] S. Suganyadevi, V. Seethalakshmi και K. Balasamy. *A review on deep learning in medical image analysis. International Journal of Multimedia Information Retrieval*, 2022.
- [64] G.J.S. et al. Litjens. *A survey on deep learning in medical image analysis*. 2017.
- [65] Jimenez Sanchez A, Albarqouni S και Mateus D. *Capsule networks against medical imaging data challenges. Intravascular imaging and computer assisted stenting and large-scale annotation of biomedical data and expert label synthesis*. 2018.

- [66] Latha S, Muthu P, Lai KW, Khalil A και Dhanalakshmi S. *Performance Analysis of Machine Learning and Deep Learning Architectures on Early Stroke Detection Using Carotid Artery Ultrasound Images*. *Front Aging Neurosci*, 2022.
- [67] Kang MK, Nam K-W, Shin JH, Kwon H-M και Lee Y-S. *Cardiovascular risk scores in asymptomatic carotid stenosis: A validation study with ultrasonographic parameters*. *PLoS One*, 2022.
- [68] Dalakleidi Kalliopi, Golemati Spyretta, Gastounioli Aimilia, Liapis Christos και Nikita Konstantina. *Ultrasound-Based Velocity and Acceleration of the Carotid Atheromatous Plaque in Asymptomatic Patients with Moderate and Severe Stenosis*. 2020.
- [69] Spyretta Golemati, Aimilia Gastounioli και Konstantina S Nikita. *Ultrasound-Image-Based Cardiovascular Tissue Motion Estimation*. *IEEE Reviews in Biomedical Engineering*, 2016.
- [70] S. Golemati, T. J. Tegos, A. Sassano, K.S. Nikita και A. N. Nicolaidis. *Echogenicity of B-mode sonographic images of the carotid artery, Work-in-progress*. 2004.
- [71] P Asvestas, S Golemati, GK Matsopoulos, KS Nikita και AN Nicolaidis. *Fractal dimension estimation of carotid atherosclerotic plaques from B-mode ultrasound: a pilot study*. 2002.
- [72] J. Stoitsis, I. Valavanis, S.G. Mougiakakou, T. Golemati, A. Nikita και K.S. Nikita. *Computer Assisted Diagnosis based on Medical Image Processing and Artificial Intelligence Methods*. 2002.
- [73] S. Mougiakakou, S. Golemati, I. Gousias, A.N. Nikolaidis και K.S. Nikita. *Computer-Aided Diagnosis of Carotid Atherosclerosis Based on Ultrasound Image Statistics, Laws' Texture and Neural Networks*. 2007.
- [74] N. Tsiaparas, S. Golemati, I. Andreadis, J. Stoitsis, I. Valavanis και K.S. Nikita. *Comparison of multiresolution features for texture classification of carotid atherosclerosis from B-mode mltrasound*. 2011.
- [75] N.N. Tsiaparas, S. Golemati, I. Andreadis, J.S. Stoitsis, I. Valavanis και K.S. Nikita. *Assessment of carotid atherosclerosis from B-mode ultrasound images using directional multiscale texture features*. 2012.
- [76] S. Golemati, J.S. Stoitsis, A. Gastounioli, A.C. Dimopoulos, V. Koropouli και K.S. Nikita. *Comparison of block-matching and differential methods for motion analysis of the carotid artery wall from ultrasound images*. 2012.
- [77] A. Gastounioli, S. Golemati, J.S. Stoitis και K.S. Nikita. *Carotid artery wall motion analysis from B-mode ultrasound using adaptive block matching: in silico evaluation and in vivo application*. 2013.

- [78] A. Gastouniotti, S. Golemati, J. Stoitsis και K.S. Nikita. *Comparison of Kalman-filter-based approaches for block matching in arterial wall motion analysis from B-mode ultrasound*. 2011.
- [79] S. Golemati, A. Gastouniotti και K.S. Nikita. *Ultrasound-image-based cardiovascular tissue motion estimation*. 2016.
- [80] F.Y. Rizi, H. Yli Ollila J. Au, S. Golemati, M. Makunaite, M. Orkisz, N. Navab, M. MacDonald, T. M. Laitinen, H. Behnam, Z. Gao, A. Gastouniotti, R. Jurkonis, D. Vray, A. Serusclat T. Laitinen, K. S. Nikita και G. Zahnd. *Carotid wall longitudinal motion in ultrasound imaging: An expert consensus review*. *Ultrasound in Medicine and Biology*, 2020.
- [81] S. Golemati, E. Patelaki, A. Gastouniotti, I.S. Andreadis, C. D. Liapis και K. S. Nikita. *Motion synchronization patterns of the carotid atheromatous plaque from B-mode ultrasound*. 2020.
- [82] S. Golemati, E. Patelaki, A. Gastouniotti, I.S. Andreadis, C. D. Liapis και K. S. Nikita. *Motion synchronization patterns of the carotid atheromatous plaque from B-mode ultrasound*. 2020.
- [83] J. Stoitsis, S. Golemati και K.S. Nikita. *A modular software system to assist interpretation of medical images - Application to vascular ultrasound images*. 2006.
- [84] A. Gastouniotti, S. Golemati, J.S. Stoitis και K.S. Nikita. *CAROTID- A web-based platform for optimal personalized management of atherosclerotic patients*. 2014.
- [85] A. S. Panayides, A. Amini, N. D. Filipovic, A. Sharma, S. A. Tsaftaris, A. Young, D.Foran, N. Do, S. Golemati, T. Kurc, K. Huang, K. S. Nikita, B. P. Veasey, M. Zervakis, J. H. Saltz και C. S. Pattichis. *AI in medical imaging informatics: Current challenges and future directions*. *IEEE Journal of Biomedical and Health Informatics*, 2020.
- [86] Marios Skevofilakas, Konstantia Zarkogianni, Basil G Karamanos και Konstantina S Nikita. *A hybrid Decision Support System for the risk assessment of retinopathy development as a long term complication of Type 1 Diabetes Mellitus*. 2010.
- [87] I. Valavanis, S.G. Mougiakakou, K.A. Grimaldi και K.S. Nikita. *A multifactorial analysis of obesity as CVD risk factor: use of neural network based methods in a nutrigenetics context*. 2010.
- [88] Jing Gao, Peng Li, Zhikui Chen και Jianing Zhang. *A Survey on Deep Learning for Multimodal Data Fusion*. 2020.
- [89] Ryan Bradley. *Biomarkers in Cardiovascular Disease*. Springer Netherlands, Dordrecht, 2016.
- [90] Katsiki N και Mikhailidis DP. *Diabetes and carotid artery disease: a narrative review*. *Annals of Translational Medicine*, 2020.

- [91] Wagenknecht LE, Zaccaro D, Espeland MA, Karter AJ, O’Leary DH και Haffner SM. *Diabetes and progression of carotid atherosclerosis: the insulin resistance atherosclerosis study. Arteriosclerosis, Thrombosis, and Vascular Biology*, 2003.
- [92] Tzu Wei Wu, Chao Liang Chou, Chun Fang Cheng, Shu Xin Lu και Li Yu Wang. *Prevalences of diabetes mellitus and carotid atherosclerosis and their relationships in middle-aged adults and elders: a community-based study. Journal of the Formosan Medical Association*, 121:1133-1140, 2022.
- [93] Wang L και Tang C. *Targeting Platelet in Atherosclerosis Plaque Formation: Current Knowledge and Future Perspectives. International Journal of Molecular Sciences*, 2020.
- [94] Von Hundelshausen P. και Weber C. *Platelets as Immune Cells Bridging Inflammation and Cardiovascular Disease. Circulation Research*, 2007.
- [95] VSiegel Axel D., Daub K., Seizer P., Lindemann S. και Gawaz M. *Platelet lipoprotein interplay: Trigger of foam cell formation and driver of atherosclerosis. Cardiovascular Research*, 2008.
- [96] Lan Q, Zheng L, Zhou X, Wu H, Buys N, Liu Z, Sun J και Fan H. *The Value of Blood Urea Nitrogen in the Prediction of Risks of Cardiovascular Disease in an Older Population. Frontiers in Cardiovascular Medicine*, 2021.
- [97] Michele Emdin, Alfonso Pompella και Aldo Paolicchi. *Gamma-glutamyltransferase, atherosclerosis, and cardiovascular disease: triggering oxidative stress within the plaque. Circulation*, 112(14):2078-2080, 2005.
- [98] Bai Hui Zhang, Fan Yin, Ya Nan Qiao και Shou Dong Guo. *Triglyceride and Triglyceride-Rich Lipoproteins in Atherosclerosis. Frontiers in Molecular Biosciences*, 9, 2022.
- [99] Karen Simonyan και Andrew Zisserman. *Very Deep Convolutional Networks for Large-Scale Image Recognition*. 2015.
- [100] Alex Krizhevsky, Ilya Sutskever και Geoffrey E. Hinton. *ImageNet Classification with Deep Convolutional Neural Networks. Circulation Research*, 2012.
- [101] Davide Chicco και Giuseppe Jurman. *The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genomics*, 2005.
- [102] Fan J, Chen M, Luo J, Yang S, Shi J, Yao Q, Zhang X, Du S, Qu H, Cheng Y, Ma S, Zhang M, Xu X, Wang Q και Zhan S. *The prediction of asymptomatic carotid atherosclerosis with electronic health records: a comparative study of six machine learning models. BMC Medical Informatics and Decision Making*, 2021.



- [103] Κ. Λέκκας. *Εκτίμηση της επικινδυνότητας αθηρωματικής πλάκας της καρωτίδας: Προσδιορισμός του δείκτη Θυξιαβιμυναβί Βλίαςκ Αρεα - ΘΒΑ σε εικόνες υπερήχου Β-σάρωσης*. Μεταπτυχιακή διπλωματική εργασία, Εθνικό Μετσόβιο Πολυτεχνείο, 2016.
- [104] Γ. Μαυράκης. *Αυτόματη ανίχνευση αθηρωματικής πλάκας σε εικόνες Β-μοδε υπερήχων μέσω ανάλυσης υφής*. Μεταπτυχιακή διπλωματική εργασία, Εθνικό Μετσόβιο Πολυτεχνείο, 2017.
- [105] Min Kyoung Kang, Ki Woong Nam, Jung Hwan Shin, Hyung Min Kwon και Yong Seok Lee. *Cardiovascular risk scores in asymptomatic carotid stenosis: A validation study with ultrasonographic parameters*. *PLOS One*, 2021.
- [106] Yi Li.et al. *Multimodal Medical Supervised Image Fusion Method by CNN*. 2021.
- [107] Yitan Zhuet al. *Converting tabular data into images for deep learning with convolutional neural networks*. *Scientific Reports*, 2021.



## Συντομογραφίες - Αρκτικόλεξα - Ακρωνύμια

---

ΤΙΑ (αγγλ.)	Transient Ischemic Attack
ΠΙΕ (ελλ.)	Παροδικό Ισχαιμικό Επεισόδιο
ΑΕΕ (ελλ.)	Αγγειακό Εγκεφαλικό Επεισόδιο
ESVS (αγγλ.)	European Society for Vascular Surgery
BMT	Best Medical Therapy
CAS	Carotid Artery Stenting
CEA	Carotid Endarterectomy
CTA	Computed Tomography Angiography
MRA	Magnetic Resonance Angiography
ANN (αγγλ.)	Artificial Neural Networks
TNN (ελλ.)	Τεχνητά Νευρωνικά Δίκτυα
EHR (αγγλ.)	Electronic Health Records