

ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ
ΘΕΣΣΑΛΟΝΙΚΗΣ



ΤΜΗΜΑ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΗΛΕΚΤΡΟΝΙΚΩΝ
ΤΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΗΛΕΚΤΡΟΝΙΚΗΣ ΚΑΙ ΥΠΟΛΟΓΙΣΤΩΝ

**MLKL μοντέλο ημι-επιβλεπόμενης μάθησης για
κατάτμηση ιατρικών εικόνων**

Διπλωματική εργασία των:
Υπατία Δάμη
Δημήτρη Μουτζόγλου

Επιβλέπων: Ιωάννης Θεοχάρης

4 Απριλίου 2023

Περίληψη

Μία από τις προκλήσεις της μηχανικής μάθησης και των εφαρμογών της στη βιοιατρική είναι η αυτοματοποιημένη κατάτμηση ιατρικών εικόνων. Με την δημιουργία διαγνωστικών μοντέλων πρόβλεψης καθίσταται δυνατή η παράλληλη παραχολούμενη της πορείας των νοσημάτων ασθενών στο πέρασμα του χρόνου τόσο από υπολογιστές όσο και από ιατρούς. Στην παρούσα εργασία υλοποιείται ένας αλγόριθμος αυτόματης κατάτμησης ιατρικών εικόνων εγκεφάλων και γονάτων και παράλληλα διερευνώνται πιθανές παραλλαγές τους οι οποίες και θα αποδώσουν το βέλτιστο αποτέλεσμα. Ο αλγόριθμος εντάσσεται στο πεδίο της ημιεπιβλεπόμενης μάθησης όπου με περιορισμένο αριθμό ταξινομημένων δεδομένων είναι ικανός να κλιμακώσει την ικνότητα πρόβλεψής του στο σύνολο των αταξινόμητων δεδομένων. Παράλληλα εξετάζεται η συσχέτηση της απόδοσης του αλγορίθμου με τις τεχνικές προεπεξεργασίας των εικόνων, τις τεχνικές επιλογής χαρακτηριστικών και τον αριθμό των χαρακτηριστικών που εξάγονται για την εκπαίδευση του μοντέλου. Η εργασία περιλαμβάνει τη θεωρητική ανάλυση του προτεινόμενου μοντέλου καθώς και την υλοποίησή του για την αναγνώριση των ιστών που εντοπίζονται σε μαγνητικές εικόνες εγκεφάλων και γονάτων.

Abstract

One of the challenges of machine learning and its applications on biomedical engineering is the automated segmentation of medical images. The creation of diagnostic prediction models enables the monitoring of the progress of patients' ailments in the passage of time both by computers and doctors of medicine. This dissertation implements an algorithm of automated segmentation on brain and knee images along with possible variations in order to achieve optimal results. The algorithm belongs to the field of semi-supervised learning, where with a limited number of classified data it is able to scale its prediction ability to the set of unclassified data. At the same time, an examination of the correlation of the algorithm's performance between image preprocessing techniques, feature selection techniques and the number of features extracted to train the model, is conducted. The dissertation includes both the theoretical analysis of the proposed model as well as its implementation for the identification of tissues located in magnetic images of brains and knees.

Ευχαριστίες

Θα θέλαμε να ευχαριστήσουμε θερμά τον καθήγητή μας Ιωάννη Θεοχάρη και τον διδάκτωρα Χρήστο Χαδουλό για την βοήθεια και την καθοδήγησή τους κατά την εκπόνηση της εργασίας.

Περιεχόμενα

1 Εισαγωγή	6
1.1 Ορισμός της Μηχανικής Μάθησης	6
1.2 Επιβλεπόμενη Μηχανική Μάθηση	6
1.3 Μη Επιβλεπόμενη Μηχανική Μάθηση	6
1.4 Ημι-Επιβλεπόμενη Μηχανική Μάθηση	7
1.5 Περιγραφή του προβλήματος	7
2 Θεωρία των SVM	7
2.1 Ορισμός SVM	7
2.2 Γραμμική Ταξινόμηση	8
2.3 Soft-Margin SVMs	12
2.4 Μη γραμμική ταξινόμηση με SVM	15
2.4.1 Kernel trick	15
2.4.2 Τύποι πυρήνων	17
2.5 Multi-class ταξινόμηση με SVM	18
2.6 Multiple Kernel Learning (MKL)	20
3 Multi-Atlas Segmentation	22
3.1 Εγγραφή Ατλάντων	22
3.2 Επιλογή Ομοίων Ατλάντων	24
3.3 Συγχώνευση ετικετών	27
4 Ανάλυση Προτεινόμενου Μοντέλου	29
4.1 Θεωρητικό υπόβαθρο αλγορίθμου	29
4.2 Εφαρμογή σε ιατρικές εικόνες	34
5 Εξαγωγή Χαρακτηριστικών	37
5.1 Οικογένειες Χαρακτηριστικών Εικόνας	37
5.1.1 Στατιστικά Χαρακτηριστικά	37
5.1.2 Γεωγραφικά Χαρακτηριστικά	42
5.1.3 Χαρακτηριστικά Ύφης	42
5.2 Το πρόβλημα της πολυδιαστασιμότητας	67
5.2.1 Τεχνικές Μείωσης Διαστασιμότητας	68
5.2.2 Τεχνικές Επιλογής Χαρακτηριστικών	69

6 Πειραματικά αποτελέσματα στον εγκέφαλο	70
6.1 Πρόβλημα	70
6.2 Σετ Δεδομένων	71
6.3 Προεπεξεργασία Δεδομένων	73
6.4 Επιλογή Χαρακτηριστικών	80
6.5 Εκπαίδευση μοντέλου	84
6.6 Αποτελέσματα	88
7 Πειραματικά αποτελέσματα στο γόνατο	94
7.1 Πρόβλημα	94
7.2 Σετ Δεδομένων	95
7.3 Προεπεξεργασία Δεδομένων	95
7.4 Επιλογή Χαρακτηριστικών	100
7.5 Εκπαίδευση μοντέλου	102
7.6 Αποτελέσματα	106
8 Μελλοντικές Επεκτάσεις	110
9 Βιβλιογραφία	111

1 Εισαγωγή

1.1 Ορισμός της Μηχανικής Μάθησης

Μηχανική μάθηση είναι μία τεχνική της τεχνητής νοημοσύνης που εκπαιδεύει τους υπολογιστές να ‘μαθαίνουν’ μέσω της εμπειρίας. Οι αλγόριθμοι μηχανικής μάθησης χρησιμοποιούν υπολογιστικές μεθόδους για να ‘μάθουν’ απευθείας από τα δεδομένα χωρίς να στηρίζονται σε κάποια προκαθορισμένη εξίσωση ως μοντέλο ενώ παράλληλα προσαρμόζονται και βελτιώνουν την απόδοση τους όσο ο αριθμός των διαθέσιμων δειγμάτων για εκπαίδευση αυξάνεται. Ο στόχος της μηχανικής μάθησης είναι να δημιουργήσει υπολογιστικά μοντέλα που παρουσιάζουν ‘έξυπνες’ ανθρώπινες συμπεριφορές όπως η αναγνώριση εικόνας, κατανόηση ήχου/κειμένου, αυτοματοποιημένη λήψη αποφάσεων κ.α.[1]. Η μηχανική μάθηση χωρίζεται σε τρεις βασικές υποκατηγορίες ανάλογα με τον τρόπο εκπαίδευσης του μοντέλου και την διαθεσιμότητα των δεδομένων:

- Επιβλεπόμενη Μάθηση
- Μη Επιβλεπόμενη Μάθηση
- Ημι-Επιβλεπόμενη Μάθηση

1.2 Επιβλεπόμενη Μηχανική Μάθηση

Η επιβλεπόμενη μαθήση είναι ο τύπος μηχανικής μάθησης που επιλύει προβλήματα εξάγοντας μία συνάρτηση που απεικονίζει μία είσοδο ανεξάρτητων μεταβλητών με μία έξοδο-εξαρτημένη μεταβλητή, βασιζόμενη σε διαθέσιμα δεδομένα που αποτελούνται από κατηγοριοποιημένες εγγραφές. Η ταξινόμηση των εγγραφών σε κατηγορίες να γίνεται εκ των προτέρων από ειδικούς[2] και συνέχει μεγάλο κόστος και την πιθανότητα ανθρωπίνου λάθους. Συνήθως για την δημιουργία ενός συνοπτικού μοντέλου που ‘γνωρίζει’ την κατανομή των κλάσεων συναρτήσει των χαρακτηριστικών απαιτείται η εκπαίδευση του με εκτεταμένα σετ γνωστών δεδομένων ενώ στη συνέχεια το μοντέλο καλείται να γενικεύσει την ικανότητα πρόβλεψης του πάνω σε άγνωστα δεδομένα.

1.3 Μη Επιβλεπόμενη Μηχανική Μάθηση

Η μη-επιβλεπόμενη μάθηση αποτελείται από αλγόριθμους που εξάγουν μοτίβα από ένα σετ αταξινόμητων δεδομένων. Σε αντίθεση με την επιβλεπόμενη μάθηση,

οι μέθοδοι αυτοί δεν μπορούν να εφαρμοστούν σε κάποιο πρόβλημα ταξινόμησης ή παλινδρόμησης επειδή δεν υπάρχει καμία γνώση γύρω από τις τιμές των κλάσεων. Η μη-επιβλεπόμενη μάθηση μπορεί να χρησιμοποιηθεί αντιθέτως για την ανακάλυψη της υποκείμενης δομής των δεδομένων, την ομαδοποίηση και την μείωση των διαστάσεων των χαρακτηριστικών.

1.4 Ήμι-Επιβλεπόμενη Μηχανική Μάθηση

Η ήμι-επιβλεπόμενη μάθηση αποτελεί ένα υβρίδιο της επιβλεπόμενης και της μη-επιβλεπόμενης μάθησης καθώς καλείται να λύσει προβλήματα όπου τα κατηγοριοποιημένα δεδομένα αποτελούν ένα μικρό μέρος το συνολικού σετ δεδομένων, με την κλάση των υπόλοιπων δεδομένων να παραμένει άγνωστη. Οι Ήμι-επιβλεπόμενοι εκτιμητές προτιμώνται σε περιπτώσεις όπου η απόκτηση ταξινομημένων δεδομένων είναι δυσχερής καθώς είναι ικανοί να εκμεταλλευτούν την πληροφορία που παρέχουν τα άγνωστα (αταξινόμητα) δεδομένα, να καταγράψουν την υποκείμενη κατανομή τους και να γενικεύσουν τις προβλέψεις τους σε νέα δεδομένα κατά τρόπο αποτελεσματικότερο των άλλων δύο μεθόδων.

1.5 Περιγραφή του προβλήματος

Η παρούσα εργασία συγκαταλέγεται στην τρίτη κατηγορία μηχανικής μάθησης, αυτή της ήμι-επιβλεπόμενης μάθησης. Αφορά την υλοποίηση ενός αλγορίθμου που καλείται να κατατμήσει ιατρικές μαγνητικές εικόνες (MRI) εγκεφάλων και γονάτων στα είδη των ιστών που τις αποτελούν, συνδυάζοντας τεχνικές SVM, MLKL και Multi-Atlas Segmentation. Στο πλαίσιο της εργασίας αναπτύσσεται τόσο ο κώδικας του αλγορίθμου κατάτμησης όσο και των υπολοίπων απαραίτητων εργασιών όπως η προεπεξεργασία των εικόνων και η εξαγωγή των χαρακτηριστικών.

2 Θεωρία των SVM

2.1 Ορισμός SVM

“Support Vector Machine” (SVM) είναι ένας αλγόριθμος επιβλεπόμενης μηχανικής μάθησης που μπορεί να χρησιμοποιηθεί εξίσου για προβλήματα ταξινόμησης και παλινδρόμησης[3]. Κυρίως χρησιμοποιείται σε προβλήματα ταξινόμησης και ψεωρείται από τους πιο αποδοτικούς αλγορίθμους πρόβλεψης. Στον

SVM αλγόριθμο δυαδικής γραμμικής ταξινόμησης, απεικονίζουμε το κάθε δείγμα δεδομένων ως ένα σημείο σε έναν N-διάστατο χώρο(όπου N ο αριθμός των χαρακτηριστικών των δεδομένων που διαθέτουμε) με την τιμή του κάθε χαρακτηριστικού να αντιστοιχεί σε μία συγκεκριμένη συντεταγμένη του χώρου. Έπειτα βρίσκουμε το υπερεπίπεδο που διαχωρίζει βέλτιστα τις δύο κλάσεις και ταξινομούμε νέα σημεία βάση της θέσης τους σε σχέση με το υπερεπίπεδο.

2.2 Γραμμική Ταξινόμηση

Η Δυαδική ταξινόμηση υλοποιείται συχνά με την χρήση μίας συνάρτησης f πραγματικών τιμών, όπου $f : X \subseteq \mathbb{R}^k \rightarrow \mathbb{R}$, με τον ακόλουθο τρόπο: η είσοδος $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)'$ κατανέμεται στην θετική κλάση εάν, $f(\mathbf{x}) \geq 0$, διαφορετικά κατανέμεται στην αρνητική κλάση. Θεωρούμε την περίπτωση όπου $f(\mathbf{x})$ είναι μία γραμμική συνάρτηση του $\mathbf{x} \in \mathbf{X}$, τέτοια ώστε να μπορεί να οριστεί ως:

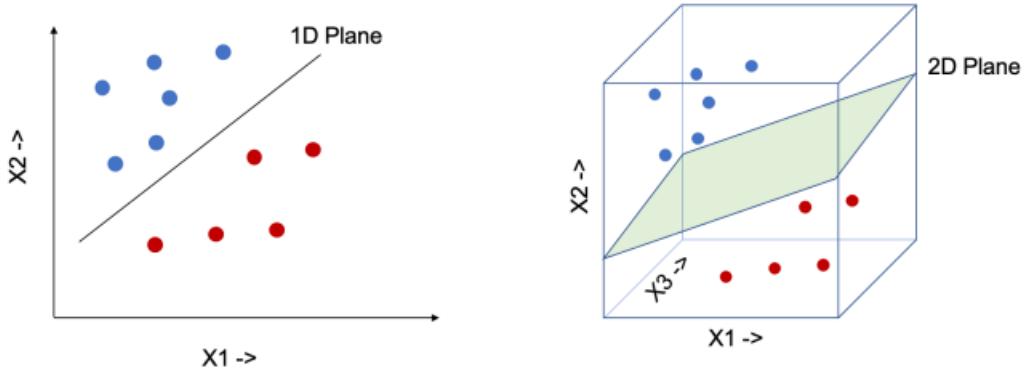
$$f(\mathbf{x}) = \mathbf{w}^T \mathbf{x} + b \quad (2.1)$$

$$= \sum_{i=1}^n \mathbf{w}_i, \mathbf{x}_i + b \quad (2.2)$$

όπου $(w, b) \in R^n \times R$ είναι οι παράμετροι βαρος(weight) και συστηματικό σφάλμα(bias) που ελέγχουν την συνάρτηση και ο κανόνας απόφασης δίνεται από την $\text{sign}(f(\mathbf{x}))$, παραραδεχόμενοι την σύμβαση πως $\text{sign}(0) = 1$. Η μεθοδολογία της μάθησης προϋπονθέτει πως αυτές οι παράμετροι μπορούν να γίνουν γνωστές από τα δεδομένα. Μια γεωμετρική ερμηνεία αυτής της υπόθεσης είναι πως το πεδίο τιμών X διαχωρίζεται σε δύο μέρη από ένα υπερεπίπεδο που ορίζεται από την εξίσωση $f(\mathbf{x}) = w^T \mathbf{x}_i + b = c$ (Σχήμα 2.1). Το υπερεπίπεδο είναι ένας ομοπαράλληλος υποχώρος διάστασης $n-1$ που διχοτομεί τον χώρο σε δύο μέρη που αντιστοιχούν στις εισόδους των δύο ξεχωριστών κλάσεων.

Η απόσταση μεταξύ του διαχωριστικού υπερεπιπέδου και του κοντινότερου δείγματος του σετ εκπαίδευσης στο υπερεπίπεδο ονομάζεται περιθώριο. Αναζητώντας το βέλτιστο διαχωριστικό υπερεπίπεδο, η ευκλείδια απόσταση ανάμεσα σε ένα δείγμα του σετ εκπαίδευσης και του υπερεπιπέδου δίνεται από τη σχέση $\frac{|f(\mathbf{x})|}{\|\mathbf{w}\|}$. Επομένως όλα τα δείγματα του σετ πρέπει να ικανοποιούν τη σχέση

$$\frac{y_i f(\mathbf{x}_i)}{\|\mathbf{w}\|} \geq \delta \text{ για } i = 1, \dots, M \quad (2.3)$$



Σχήμα 2.1: Διαχωρισμός των δύο κλάσεων από το βέλτιστο υπερεπίπεδο

όπου δείναι το περιθώριο. Εάν η (w, b) είναι μια λύση, τότε η (aw, ab) είναι επίσης μια λύση, όπου a ένα βαθμωτό μέγεθος. Επομένως ορίζουμε τον παρακάτω περιορισμό:

$$\delta||\mathbf{w}|| = 1 \quad (2.4)$$

Από τις (2.3) και (2.4), για να βρούμε το βέλτιστο υπερεπίπεδο, πρέπει να ελαχιστοποιηθεί το $||w||$, αρκεί λοιπόν να ελαχιστοποιήσουμε την εξίσωση:

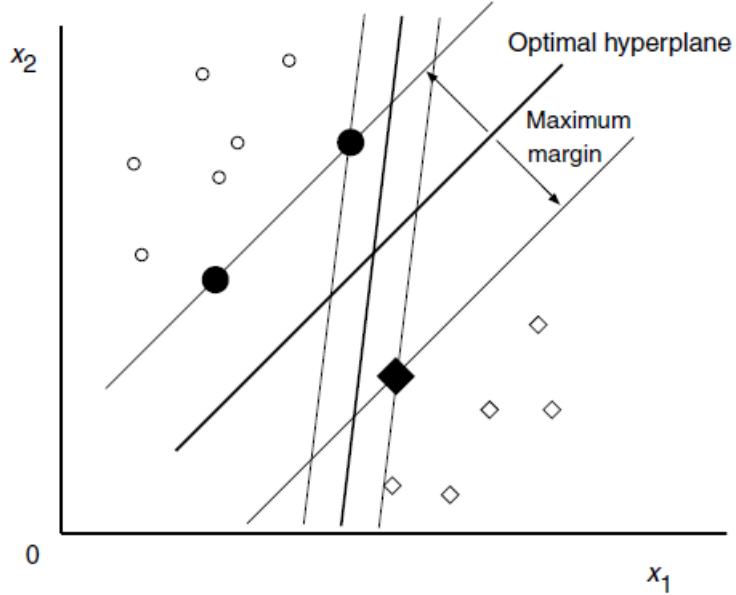
$$Q(w) = \frac{1}{2}||\mathbf{w}||^2 \quad (2.5)$$

με τα w και b να υπόκεινται στον περιορισμό:

$$y_i(\mathbf{w}^T \mathbf{x}_i + b) \geq 1 \text{ για } i = 1, \dots, M \quad (2.6)$$

Επειδή μπορούμε να αποκτήσουμε το βέλτιστο υπερεπίπεδο ακόμα και αν αφαιρέσουμε όλα τα δείγματα που ικανοποιούν τις αυστηρές ανισότητες, τα δείγματα που ικανοποιούν τις ισότητες ονομάζονται support vectors (Σχήμα 2.2).

Όταν ο αριθμός των μεταβλητών εισόδου είναι μικρός, μπορούμε να λύσουμε τις (2.5) και (2.6) με την χρήση τετραγωνικού προγραμματισμού. Όμως, στις περισσότερες περιπτώσεις επειδή απεικονίζουμε τον χώρο της εισόδου σε ένα χώρο πολλών διαστάσεων, ενίστε απείρων, μετατρέπουμε τις σχέσεις (2.5), (2.6) στο ισοδύναμο δυαδικό πρόβλημα του οποίου ο αριθμός των μεταβλητών ισούται με τον αριθμό των δειγμάτων. Προκειμένου να λύσουμε το δυαδικό πρόβλημα, πρέπει πρώτα να μετατρέψουμε το υπό περιορισμούς πρόβλημα της (2.5) στο



Σχήμα 2.2: Βέλτιστο υπερεπίπεδο και support vectors

άνευ περιορισμών πρόβλημα:

$$Q(\mathbf{w}, b, \boldsymbol{\alpha}) = \frac{1}{2} \mathbf{w}^T \mathbf{w} - \sum_{i=1}^M a_i \{y_i(\mathbf{w}^T \mathbf{x}_i + b) - 1\}, \quad (2.7)$$

όπου $a = (a_1, \dots, a_M)^T$ και a_i είναι οι μη-αρνητικοί παράγοντες Lagrange. Η βέλτιστη λύση της (2.7) δίνεται από το σαγματικό σημείο, όπου η (2.7) ελαχιστοποιείται σε σχέση με τα \mathbf{w} και b και μεγιστοποιείται όσον αφορά τα $a_i (\geq 0)$, ενώ παράλληλα ικανοποιεί τις ακόλουθες Karush-Kuhn-Tucker (KKT) συνθήκες.

$$\frac{\partial Q(\mathbf{w}, b, \boldsymbol{\alpha})}{\partial \mathbf{w}} = 0 \quad (2.8)$$

$$\frac{\partial Q(\mathbf{w}, b, \boldsymbol{\alpha})}{\partial b} = 0 \quad (2.9)$$

$$a_i y_i (\mathbf{w}^T \mathbf{x}_i + b) - 1 = 0 \quad \text{για } i = 1, \dots, M \quad (2.10)$$

$$a_i \geq 0 \quad \text{για } i = 1, \dots, M \quad (2.11)$$

Αντικαθιστώντας την σχέση (2.7), στις (2.8) και (2.9) καταλήγουμε σε :

$$\mathbf{w} = \sum_{i=1}^M a_i y_i \mathbf{x}_i \quad (2.12)$$

και

$$\sum_{i=1}^M a_i y_i = 0 \quad (2.13)$$

Τώρα, αντικαθιστώντας τις (2.12), (2.13) στην (2.7), καταλήγουμε στο ακόλουθο δυαδικό πρόβλημα. Μεγιστοποίησε την:

$$Q(\boldsymbol{\alpha}) = \sum_{i=1}^M a_i - \frac{1}{2} \sum_{i,j=1}^M a_i a_j y_i y_j \mathbf{x}_i^T \mathbf{x}_j \quad (2.14)$$

με το a_i να υπόκειται στους περιορισμούς:

$$\sum_{i=1}^M a_i y_i = 0, \quad a_i \geq 0 \text{ για } i = 1, \dots, M. \quad (2.15)$$

Αν υπάρχει λύση ή διαφορετικά αν το πρόβλημα ταξινόμησης είναι γραμμικά διαχωρίσιμο, το ολικό βέλτιστο a_i ($i=1, \dots, M$) υπάρχει. Για τον τετραγωνικό προγραμματισμό, οι τιμές του πρωτεύοντος και του δυαδικού προβλήματος συμπέφτουν στις βέλτιστες λύσεις αν αυτές υπάρχουν. Αυτό ονομάζεται μηδενικό κενό δυαδικότητας. Εντέλει, αντικαθιστώντας τα α στην (2.12) και έπειτα τα ω στην (2.1), η συνάρτηση απόφασης καταλήγει στη μορφή:

$$f(\mathbf{x}) = \sum_{i \in S} a_i y_i \mathbf{x}_i^T \mathbf{x} + b \quad (2.16)$$

όπου Σ είναι το σύνολο των δεικτών των support vectors, και από τις συνθήκες KKT, το b δίνεται ως:

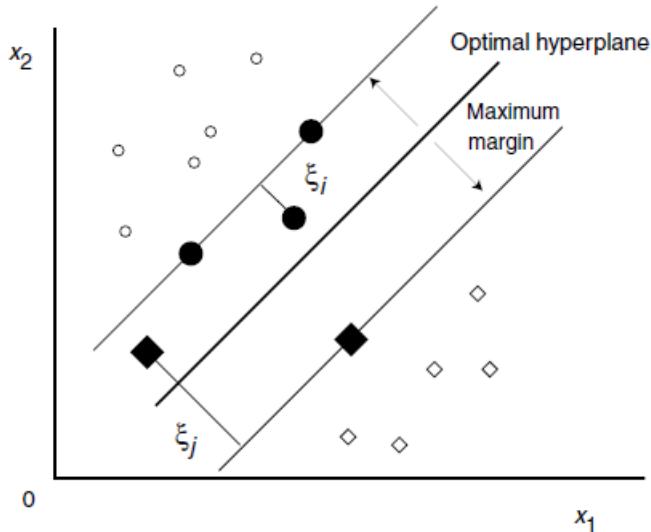
$$b = y_i - \mathbf{w}^T \mathbf{x}_i \quad (2.17)$$

όπου \mathbf{x}_i είναι support vector.

2.3 Soft-Margin SVMs

Μέχρι στιγμής έχουμε ορίσει τα hard-margin SVMs στα οποία η λύση είναι εφικτή αν και μόνο αν τα δεδομένα είναι γραμμικά διαχωρίσιμα. Στη συνέχεια θα επεκταθούμε στις περιπτώσεις όπου τα δεδομένα δεν είναι διαχωρίσιμα, με τον ορισμό των soft-margin SVMs. Για να επιτρέψουμε την αδιαφετότητα στα δεδομένα μας, εισάγουμε τις μη-αρνητικές μεταβλητές $\xi_i \geq 0$ όπως φαίνεται στην (2.18):

$$y_i(\mathbf{w}^T \mathbf{x}_i + b) \geq 1 - \xi_i \quad \text{για } i = 1, \dots, M \quad (2.18)$$



Σχήμα 2.3: Γραμμικά μη-διαχωρίσιμα δεδομένα σε δισδιάστατο χώρο και soft-margin SVMs

Με τις μεταβλητές ξ_i , υπάρχουν πάντα εφικτές λύσεις. Για το σετ εκπαίδευσης \mathbf{x}_i , αν $0 < \xi_i < 1$, τα δεδομένα δεν έχουν το μέγιστο δυνατό περιθώριο, πάραυτα ταξινομούνται σωστά. Αντίθετα, αν το $\xi_i \geq 1$, τα δεδομένα ταξινομούνται λανθασμένα από το βέλτιστο υπερεπίπεδο. Για να πετύχουμε το βέλτιστο υπερεπίπεδο στο οποίο τα δεδομένα εκπαίδευσης έχουν ελαχιστοποιηθεί χωρίς το μέγιστο περιθώριο, χρειάζεται να ελαχιστοποιήσουμε την:

$$Q(\mathbf{w}, b, \xi) = \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^M \xi_i^p \quad (2.19)$$

που υπόκειται στους περιορισμούς της (2.18) όπου $\xi = (\xi_1, \dots, \xi_M)^T$ και C είναι η παράμετρος του περιθωρίου που καθορίζει τον συμβιβασμό ανάμεσα στην μεγιστοποίηση του περιθωρίου και την ελαχιστοποίηση του λάθους ταξινόμησης. Η μεταβλητή p στην σχέση (2.19) ισούται είτε με 1 είτε με 2. Αν το $p = 1$ ονομάζουμε το SVM, L1 soft-margin support vector machine ή L1 SVM. Κατά αντιστοιχία ονομάζεται και το L2 SVM, για $p = 2$. Παρακάτω αναλύεται το L1 SVM. Αντίστοιχα με την γραμμικώς διαχωρίσιμη περίπτωση, εισάγουμε τους μη-αρνητικούς παράγοντες Lagrange a_i και b_i και έχουμε:

$$Q(\mathbf{w}, b, \xi, \alpha, \beta) = \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^M \xi_i - \sum_{i=1}^M a_i (y_i (\mathbf{w}^T \mathbf{x}_i + b) - 1 + \xi_i) - \sum_{i=1}^M \beta_i \xi_i \quad (2.20)$$

όπου $a = (a_1, \dots, a_M)^T$
και $\beta = (\beta_1, \dots, \beta_M)^T$.

Για την βέλτιστη λύση, οι ακόλουθες KKT συνθήκες πρέπει να ικανοποιούνται:

$$\frac{\partial Q(\mathbf{w}, b, \xi, \alpha, \beta)}{\partial \mathbf{w}} = 0 \quad (2.21)$$

$$\frac{\partial Q(\mathbf{w}, b, \xi, \alpha, \beta)}{\partial b} = 0 \quad (2.22)$$

$$\frac{\partial Q(\mathbf{w}, b, \xi, \alpha, \beta)}{\partial \xi} = 0 \quad (2.23)$$

$$a_i (y_i (\mathbf{w}^T \mathbf{x}_i + b) - 1 + \xi_i) = 0 \quad \text{για } i = 1, \dots, M \quad (2.24)$$

$$\beta_i \xi_i = 0 \quad \text{για } i = 1, \dots, M \quad (2.25)$$

$$a_i \geq 0, \beta_i \geq 0, \xi_i \geq 0 \quad \text{για } i = 1, \dots, M \quad (2.26)$$

Αντικαθιστώντας την (2.20) στις (2.21-23) αντίστοιχα, έχουμε:

$$\mathbf{w} = \sum_{i=1}^M a_i y_i \mathbf{x}_i, \quad (2.27)$$

$$\sum_{i=1}^M a_i y_i = 0, \quad (2.28)$$

$$a_i + \beta_i = C \quad \text{για } i = 1, \dots, M \quad (2.29)$$

Έπειτα αντικαθιστώντας τις (2.27-29) στην (2.20) καταλήγουμε στο ακόλουθο δυαδικό πρόβλημα. Μεγιστόποιησε την:

$$Q(\boldsymbol{\alpha}) = \sum_{i=1}^M a_i - \frac{1}{2} \sum_{i,j=1}^M a_i a_j y_i y_j \mathbf{x}_i^T \mathbf{x}_j \quad (2.30)$$

που υπόκειται στους περιορισμούς:

$$\sum_{i=1}^M y_i a_i = 0, \quad C \geq a_i \geq 0 \quad \text{για } i = 1, \dots, M \quad (2.31)$$

Η μοναδική διαφορά ανάμεσα στα L1 soft-margin support vector machines και τα hard-margin support vector machines είναι πως τα a_i δεν μπορούν να ξεπεράσουν την τιμή C . Η συνάρτηση απόφασης είναι η ίδια με αυτήν των hard-margin support vector machines και δίνεται:

$$f(\mathbf{x}) = \sum_{i \in S} a_i y_i \mathbf{x}_i^T \mathbf{x} + b \quad (2.32)$$

όπου S είναι το σετ των δεικτών των support vectors.

2.4 Μη γραμμική ταξινόμηση με SVM

2.4.1 Kernel trick

Σε ένα SVM, το βέλτιστο υπερεπίπεδο καθορίζεται με στόχο την μεγιστοποίηση της ικανότητας γενίκευσης. Όμως, αν το σετ εκπαίδευσης δεν είναι γραμμικά διαχωρίσιμο, ο ταξινομητής δεν θα διαθέτει υψηλή ικανότητα γενίκευσης ακόμα και αν το υπερεπίπεδο καθορίζεται κατά βέλτιστο τρόπο. Προκειμένου να βελτιώσουμε την γραμμική διαχωρισμότητα, ο αρχικός χώρος εισόδου απεικονίζεται/μετασχήματιζεται σε έναν υψηλών-διαστάσεων feature space μέσω του εσωτερικού γινομένου. Χρησιμοποιώντας λοιπόν την διανυσματική συνάρτηση $g(\mathbf{x}) = (g_1(\mathbf{x}), \dots, g_l(\mathbf{x}))^T$ που απεικονίζει το διάνυσμα εισόδου \mathbf{x} και μήκους m στον feature space l διαστάσεων, η γραμμική συνάρτηση απόφασης δίνεται ως:

$$f(\mathbf{x}) = \mathbf{w}^T g(\mathbf{x}) + b \quad (2.33)$$

όπου το \mathbf{w} είναι ένα διάνυσμα l διαστάσεων και b το *bias*.

Σύμφωνα με τη θεωρία των Hilbert-Schmidt, αν μία συμμετρική συνάρτηση $H(\mathbf{x}, \mathbf{x}')$ ικανοποιεί την ανισότητα:

$$\sum_{i,j=1}^M h_i h_j H(\mathbf{x}_i, \mathbf{x}_j) \geq 0 \quad (2.34)$$

για όλα τα M , \mathbf{x}_i , και h_i , όπου το M παίρνει τιμές φυσικών αριθμών και τα h_i πραγματικών αριθμών, τότε υπάρχει μια συνάρτηση $g(\mathbf{x})$, που αποικονίζει τα \mathbf{x} στο feature space εσωτερικών γινομένων και ικανοποιεί τη σχέση:

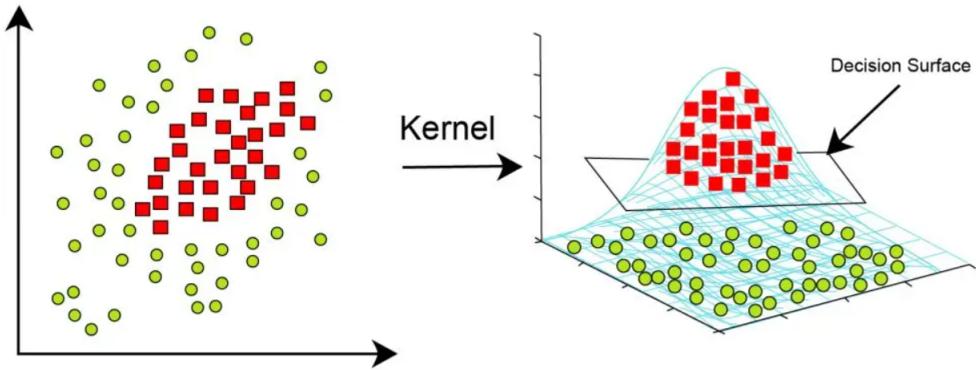
$$H(\mathbf{x}, \mathbf{x}') = g^T(\mathbf{x})g(\mathbf{x}') \quad (2.35)$$

Η συνάρτηση που ικανοποιεί την (2.34) ονομάζεται συνάρτηση πυρήνα (kernel function). Το πλεονέκτημα στη χρήση των συναρτήσεων πυρήνων έγκειται στο ότι δεν χρειάζεται να συμπεριφερθούμε αναλυτικά στον υψηλής διάστασης χώρο. Αυτή η τεχνική ονομάζεται kernel trick.

Χρησιμοποιώντας την συνάρτηση πυρήνα $H(\mathbf{x}, \mathbf{x}')$, το δυαδικό πρόβλημα στον feature space προκύπτει ως εξής.

Μεγιστοποίησε την:

$$Q(\boldsymbol{\alpha}) = \sum_{i=1}^M a_i - \frac{1}{2} \sum_{i,j=1}^M a_i a_j y_i y_j H(\mathbf{x}_i, \mathbf{x}_j) \quad (2.36)$$



Σχήμα 2.4: Μετασχηματισμός σε υψηλότερη διάσταση με χρήση συνάρτησης πυρήνα

που υπόκειται στους περιορισμούς

$$\sum_{i=1}^M y_i a_i = 0, \quad C \geq a_i \geq 0 \quad \text{για } i = 1, \dots, M \quad (2.37)$$

Επειδή ο $H(\mathbf{x}, \mathbf{x}')$ είναι ένας θετικός ημι-ορισμένος πυρήνας, το πρόβλημα της βελτιωποίησης είναι ένα κοίλο πρόβλημα τετραγωνικού προγραμματισμού. Παράλληλα, επειδή $a = 0$ είναι μία εφικτή λύση, το πρόβλημα έχει ολική βέλτιστη λύση. Οι KKT συνθήκες δίνονται ως εξής:

$$a_i \left(y_i \left(\sum_{j=1}^M y_j a_j H(\mathbf{x}_i, \mathbf{x}_j) + b \right) - 1 + \xi_i \right) = 0 \quad \text{για } i = 1, \dots, M \quad (2.38)$$

$$(C - a_i)x_i = 0, \quad a_i \geq 0, \quad , \xi_i \geq 0 \quad \text{για } i = 1, \dots, M \quad (2.39)$$

Η συνάρτηση απόφασης ορίζεται ως:

$$f(\mathbf{x}) = \sum_{i \in S} a_i y_i H(\mathbf{x}_i, \mathbf{x}) + b \quad (2.40)$$

όπου το b δίνεται ως:

$$b = y_j - \sum_{i \in S} a_i y_i H(\mathbf{x}_i, \mathbf{x}_j) \quad (2.41)$$

2.4.2 Τύποι πυρήνων

Παρακάτω αναλύονται ορισμένοι τύποι πυρήνων που χρησιμοποιούνται στα SVM.

Γραμμικοί πυρήνες

Αν το πρόβλημα ταξινόμησης είναι γραμμικά διαχωρίσιμο στο χώρο εισόδου, δεν χρειάζεται να απεικονίσουμε τον χώρο εισόδου σε χώρο υψηλών διαστάσεων. Σε αυτή την περίπτωση χρησιμοποιούμε γραμμικούς πυρήνες:

$$H(\mathbf{x}, \mathbf{x}') = \mathbf{x}^T \mathbf{x}' \quad (2.42)$$

Πολυωνυμικοί πυρήνες

Ο πολυωνυμικός πυρήνας με βαθμό d , όπου d είναι ένας φυσικός αριθμός, δίνεται ως:

$$H(\mathbf{x}, \mathbf{x}') = (\mathbf{x}^T \mathbf{x}' + 1)^d \quad (2.43)$$

Εδώ, το 1 προστίθεται προκειμένου οι χιαστί όροι με βαθμό ίσο ή μικρότερο του d να συμπεριληφθούν.

Όταν $d = 1$, ο πυρήνας είναι ένας γραμμικός πυρήνας συν 1. Έτσι, μεταβάλλοντας το d στη συνάρτηση απόφασης είναι ισοδύναμος με το γραμμικό πυρήνα. Για $d = 2$ και $m = 2$ ο πολυωνυμικός πυρήνας γίνεται:

$$\begin{aligned} H(\mathbf{x}, \mathbf{x}') &= 1 + 2x_1x'_1 + 2x_2x'_2 + 2x_1x'_1x_2x'_2 + x_1^2x'^2_1 + x_2^2x'^2_2 \\ &= g(\mathbf{x})^T g(\mathbf{x}') \end{aligned} \quad (2.44)$$

όπου $g(\mathbf{x}) = (1, \sqrt{2}x_1, \sqrt{2}x_2, \sqrt{2}x_1x_2, x_1^2, x_2^2)^T$.

Radial Basis Function πυρήνες

Ο Radial Basis Function (RBF) πυρήνας δίνεται από τη σχέση:

$$H(\mathbf{x}, \mathbf{x}') = \exp(-\gamma \|\mathbf{x} - \mathbf{x}'\|^2) \quad (2.45)$$

όπου γ μία θετική παράμετρος για τον έλεγχο της ακτίνας. Αναδιατυπώνοντας:

$$H(\mathbf{x}, \mathbf{x}') = \exp(-\gamma \|\mathbf{x}\|^2) \exp(-\gamma \|\mathbf{x}'\|^2) \exp(2\gamma \|\mathbf{x}^T \mathbf{x}'\|^2) \quad (2.46)$$

επειδή ισχύει:

$$\exp(2\gamma \mathbf{x}^T \mathbf{x}') = 1 + 2\gamma \mathbf{x}^T \mathbf{x}' + 2\gamma^2 (\mathbf{x}^T \mathbf{x}')^2 + \frac{(2\gamma)^3}{3!} (\mathbf{x}^T \mathbf{x}')^3 + \dots, \quad (2.47)$$

$\exp(2\gamma \mathbf{x}^T \mathbf{x}')$ είναι ένα άπειρο άθροισμα πολυωνύμων. Για αυτό και είναι πυρήνας. Επιπρόσθετα $\exp(-\gamma \|\mathbf{x}\|^2)$ και $\exp(-\gamma \|\mathbf{x}'\|^2)$ αποδεικνύονται να είναι πυρήνες και το γινόμενο πυρήνων είναι επίσης πυρήνας. Επομένως η (2.45) είναι πυρήνας.

Από την (2.40), η συνάρτηση απόφασης δίνεται από τη σχέση:

$$f(\mathbf{x}) = \sum_{i \in S} a_i y_i \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}'\|^2) + b \quad (2.48)$$

Εδώ, τα support vectors είναι τα κέντρα των radial basis functions.

2.5 Multi-class ταξινόμηση με SVM

Όπως αναλύθηκε στην προηγούμενη ενότητα, τα SVMs δομούνται για δυαδικά προβλήματα. Επειδή όμως τα SVMs χρησιμοποιούν μία συνάρτηση απόφασης, η επέκταση της σε προβλήματα πολλών κλάσεων δεν μπορεί να γίνει απευθείας. Υπάρχουν πρακτικά 4 τύποι SVMs που χειρίζονται προβλήματα πολλών κλάσεων και είναι οι εξής:

- one-against-all SVMs,
- pairwise SVMs,
- error-correcting output code (ECOC) SVMs,
- all-at-once SVMs,

Στην παρούσα εργασία ωστε να υλοποιηθεί η κατηγορία one-against-all SVMs η οποία και αναλύεται παρακάτω.

One-against-all SVMs

Θεωρούμε ένα πρόβλημα N κλάσεων. Για ένα one-against-all SVM, θεωρούμε n συναρτήσεις απόφασης που διαχωρίζουν μία κλάση από τις εναπομέίνουσες κλάσεις. Θεωρούμε την i -οστή συνάρτηση απόφασης, με τό μέγιστο περιθώριο που διαχωρίζει την i -οστή κλάση από τις υπόλοιπες κλάσεις:

$$D_i(\mathbf{x}) = \mathbf{w}_i^T g(\mathbf{x}) + b_i \quad (2.49)$$

όπου w_i είναι το l -διάστατο διάνυσμα, $g(\mathbf{x})$ είναι η συνάρτηση που απεικονίζει το \mathbf{x} στον l -διάστατο feature space και b_i ο όρος bias.

Το υπερεπίπεδο $D(\mathbf{x}) = 0$ διαμορφώνει το ιδανικό διαχωριστικό υπερεπίπεδο, και αν το πρόβλημα ταξινόμησης είναι διαχωρίσιμο, τα δεδομένα εκπαίδευσης που ανήκουν στην κλάση i ικανοποιούν την $D_i(\mathbf{x}) \geq 1$, ενώ αυτά που ανήκουν στις υπόλοιπες κλάσεις ικανοποιούν την $D_i(\mathbf{x}) \leq 1$. Ειδικά τα support vectors ικανοποιούν $D_i(\mathbf{x}) = 1$. Αν το πρόβλημα είναι μη διαχωρίσιμο, τα μη φραγμένα support vectors ικανοποιούν $D_i(\mathbf{x}) = 1$ και τα φραγμένα ικανοποιούν $y_i D_i(\mathbf{x}) \leq 1$. Τα υπόλοιπα δεδομένα εκπαίδευσης ικανοποιούν $y_i D_i(\mathbf{x}) \geq 1$.

Στο πρόβλημα ταξινόμησης, αν για το διάνυσμα εισόδου \mathbf{x} ισχύει η ανισότητα:

$$D_i(\mathbf{x}) > 0 \quad (2.50)$$

για ένα i , το \mathbf{x} ταξινομείται στην i -οστή τάξη. Επειδή μόνο το πρόσημο της συνάρτησης απόφασης χρησιμοποιείται, η απόφαση είναι διαχριτή.

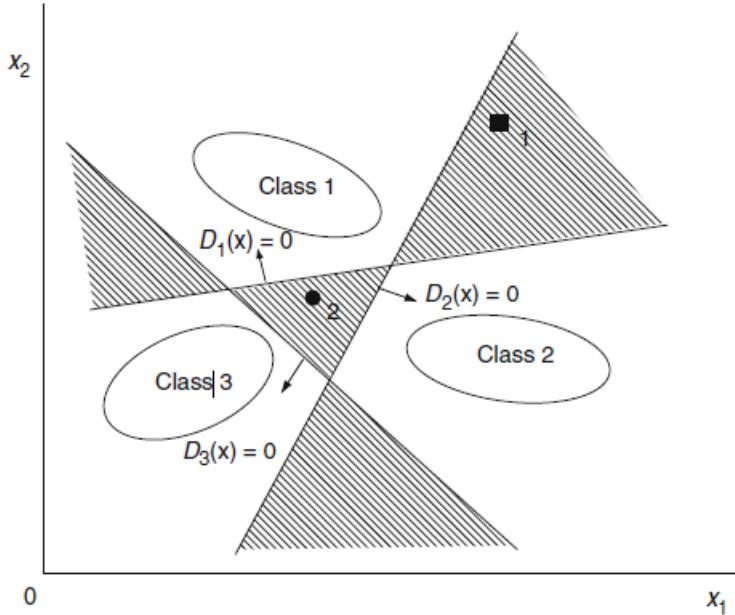
Αν η (2.51) ικανοποιείται για πολλά i ή για κανένα i , τότε το \mathbf{x} είναι μη ταξινομήσιμο. Θεωρούμε το πρόβλημα τριών κλάσεων με διάνυσμα εισόδου δύο διαστάσεων, όπου τα βέλη δείχνουν την θετική πλευρά των υπερεπιπέδων. Για το δείγμα \mathbf{x}_1 , οι τρεις συναρτήσεις απόφασης είναι:

$$D_1(\mathbf{x}_1) > 0 \quad D_2(\mathbf{x}_1) > 0 \quad D_3(\mathbf{x}_1) < 0. \quad (2.51)$$

Επειδή το x_1 ανήκει και στην κλάση 1 και στην κλάση 2 είναι μη ταξινομήσιμο. Παρόμοια, για το δείγμα \mathbf{x}_2 οι τρεις συναρτήσεις απόφασης είναι:

$$D_1(\mathbf{x}_2) < 0 \quad D_2(\mathbf{x}_2) < 0 \quad D_3(\mathbf{x}_2) < 0. \quad (2.52)$$

Επομένως το \mathbf{x}_2 είναι μη ταξινομήσιμο.



Σχήμα 2.5: Μη-ταξινομήσιμες περιοχές από την μέθοδο one vs all

Για να το αποφύγουμε αυτό, αντί για διακριτή συνάρτηση απόφασης, επιλέγουμε συνεχείς συναρτήσεις για την ταξινόμηση. Επομένως το δείγμα x θα ταξινομηθεί στην κλάση:

$$\arg \max_{i=1, \dots, n} f_i(x)$$

Έτσι το δείγμα \mathbf{x}_1 ταξινομείται στην κλάση 1 επειδή η $D_1(\mathbf{x}_1)$ έχει την μεγαλύτερη τιμή από τις άλλες δύο. Παρομοίως το δείγμα x_2 ταξινομείται στην κλάση 1.

2.6 Multiple Kernel Learning (MKL)

Κάποια πρόβλημα μηχανικής μάθησης μπορεί να απαιτούν πιο ευέλικτα μοντέλα. Πρόσφατες δημοσιεύσεις [4] έδειξαν πως αν χρησιμοποιηθούν πολλοί πυρήνες αντί για έναν μπορεί να ενισχυθεί η ερμηνευτικότητα της συνάρτησης απόφασης και να βελτιωθεί η απόδοση του μοντέλου. Σε τέτοιες περιπτώσεις, μια βολική προσέγγιση είναι να θεωρήσουμε πως ο πυρήνας $K(x, x')$ είναι ένας κυρτός συνδυασμός πυρήνων βάσης:

$$K(x, x') = \sum_{m=1}^M d_m K_m(x, x') \quad (2.53)$$

με

$$d_m \geq 0, \sum_{m=1}^M d_m = 1 \quad (2.54)$$

όπου M ο συνολικός αριθμός των πυρήνων. Κάθε πυρήνας βάσης μπορεί να χρησιμοποιεί είτε το πλήρες σετ των μεταβλητών που περιγράφουν το x , είτε υποσύνολα μεταβλητών που προκύπτουν από διαφορετικές πηγές δεδομένων. Εναλλακτικά, οι πυρήνες K_m μπορούν απλώς να είναι κλασικοί πυρήνες, (πχ γκαουσιανοί) με διαφορετικές παραμέτρους. Έτσι το πρόβλημα ανάγεται στην επιλογή των βαρών d_m . Το να εκπαιδευτεί το μοντέλο για να μάθει τις μεταβλητές α_i και d_m στο ίδιο πρόβλημα βελτιστοποίησης, είναι γνωστό ως το *multiple kernel learning (MKL)* πρόβλημα[5].

Στο πλαίσιο του *MKL*, η συνάρτηση απόφασης παίρνει τη μορφή

$$f(x) + b = \sum_m f_m(x) + b \quad (2.55)$$

όπου κάθε συνάρτηση f_m ανήκει σε διαφορετικό RKHS H_m μέσω ενός πυρήνα K_m . Σύμφωνα με το παραπάνω πλαίσιο, προκύπτει το primal πρόβλημα του MKL όπου καλούμαστε να ελαχιστοποιήσουμε την:

$$Q(\mathbf{w}, b, \xi, d) = \frac{1}{2} \sum_m \|\mathbf{w}\|^2 + C \sum_{i=1}^n \xi_i \quad (2.56)$$

$$y_i \sum_m f_m(x_i) + y_i b \geq 1 - \xi_i \quad (2.57)$$

$$\xi_i \geq 0 \quad (2.58)$$

$$\sum_m d_m = 1, d_m \geq 0 \quad (2.59)$$

όπου κάθε d_m ελέγχει την τετραγωνική νόρμα της w στην συνάρτηση Q .

3 Multi-Atlas Segmentation

Η Multi-Atlas κατάτμηση εικόνων είναι μια τεχνική στην ανάλυση ιατρικών εικόνων που περιέχει τη χρήση πολλών ατλάντων με σκοπό την κατάτμηση μιας μη κατηγοριοποιημένης εικόνας. Ένας άτλαντας αποτελεί ουσιαστικά μια κατηγοριοποιημένη εικόνα από ειδικούς που χρησιμοποιείται ως εικόνα αναφοράς για την κατάτμηση μη κατηγοριοποιημένων εικόνων [6]. Στόχος είναι η χρήση όλης της πληροφορίας του σετ ατλάντων για να παραχθεί μια αποτελερματική κατάτμηση. Η Multi-Atlas κατάτμηση αποτελείται από τις εξής διαδικασίες:

1. Εγγραφή ενός σετ ατλάντων στην εικόνα στόχο.
2. Επιλογή των πιο όμοιων ατλάντων με την εικόνα στόχο.
3. Συγχώνευση ετικετών με στόχο την τελική πρόβλεψη.

3.1 Εγγραφή Ατλάντων

Η εγγραφή ατλάντων είναι μια διαδικασία στην ανάλυση ιατρικών εικόνων που αποτελεί ουσιαστικά την ευθυγράμμιση ή ταύτιση ενός ή παραπόνω ατλάντων στην εικόνα προς κατάτμηση. Ο σκοπός της εγγραφής είναι η δημιουργία μιας αντιστοιχίας μεταξύ των δομικών στοιχείων των ατλάντων και της εικόνας στόχου έτσι ώστε να καταστεί δυνατή η μεταφορά όλης της χρήσιμης πληροφορίας των ατλάντων και να επιτευχθεί η βέλτιστη απόδοση του μοντέλου κατάτμησης. Η διαδικασία περιέχει τα εξής βήματα:

1. Προεπεξεργασία ατλάντων

Πριν την εγγραφή είναι αναγκαία η προεπεξεργασία των δεδομένων όπως η μείωση θορύβου, κανονικοποίηση φωτεινότητας και διόρθωση πόλωσης, έτσι ώστε να βελτιωθεί η ακρίβεια της διαδικασίας

2. Επιλογή μοντέλου μετασχηματισμού

Η επιλογή του μοντέλου μετασχηματισμού πρέπει να γίνει με βάση την επιθυμητή σχέση μεταξύ των ατλάντων και της εικόνας στόχου[7]. Μερικά από τα είδη μετασχηματισμών είναι τα εξής:

Affine

Ο affine μετασχηματισμός προϋποθέτει ότι οι εικόνες διαφέρουν στην κλίμακα, την θέση και τον προσανατολισμό. Εφαρμόζει επαναποθέτηση του αντικειμένου στόχου στους άξονες x, y, z , περιστροφή, κλιμάκωσή και "κούρεμά" του. Χρησιμοποιείται κυρίως όταν οι εικόνες δεν είναι ταυτόσημες, αλλά έχουν παρόμοιο περιεχόμενο και μορφή του σχήματος προς εγγραφή.

Rigid

Ο rigid (δύσκαμπτος) μετασχηματισμός προϋποθέτει πως στις δύο εικόνες η σχετική θέση, ο προσανατολισμός και η κλίμακα είναι ίδια. Εφαρμόζει μόνο περιστροφή και επαναποθέτηση στους άξονες x, y, z και όχι κλιμάκωση και "κούρεμα". Χρησιμοποιείται όταν οι εικόνες είναι σχεδόν ταυτόσημες, αλλά δεν είναι απόλυτα ευθυγράμμισμένες.

Deformable

Ο Deformable μετασχηματισμός προϋποθέτει πως δύο εικόνες μπορεί να είναι σημαντικά διαφορετικές στο σχήμα λόγω τοπικών παραμορφώσεων. Περιέχει εύκαμπτη παραμόρφωση των εικόνων για να εντοπιστούν περίπλοκες παραμορφώσεις[8]. Οι πιο γνωστοί Deformable μετασχηματισμοί είναι οι B-spline και οι Demons.

Μη γραμμικός μετασχηματισμός

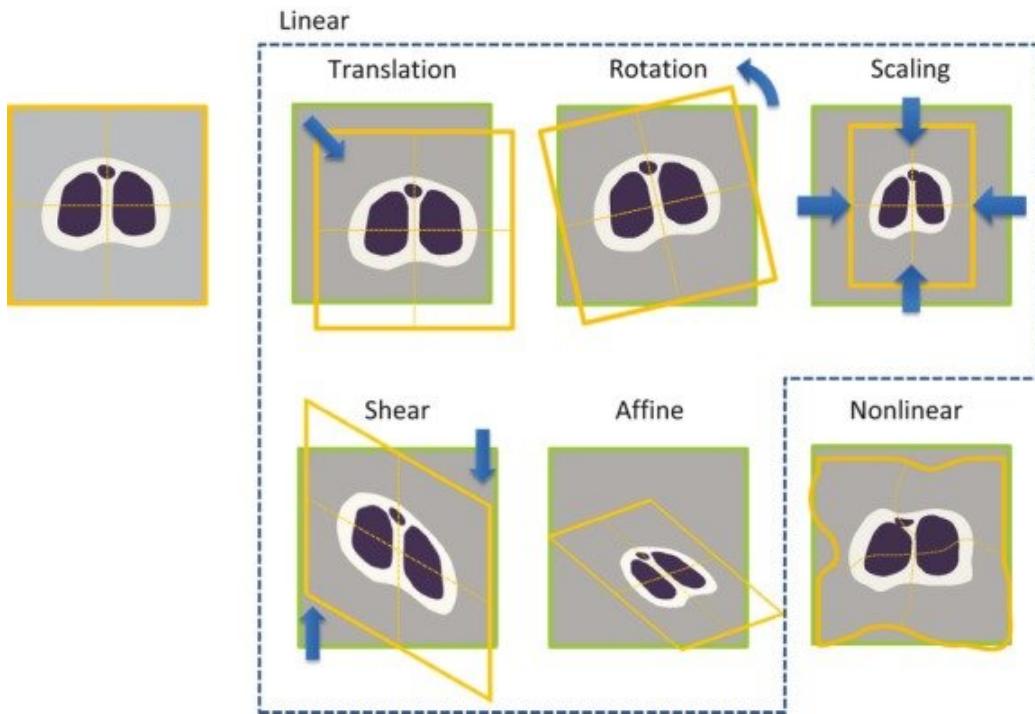
Ο μη γραμμικός μετασχηματισμός είναι ένα είδος Deformable μετασχηματισμού που μπορεί να εντοπίσει πιο περίπλοκες παραμορφώσεις από τους γραμμικούς μετασχηματισμούς. Αυτές οι παραμορφώσεις εκφράζονται ουσιαστικά ως πολυωνυμικές ή spline συναρτήσεις.

3. Υπολογισμός παραμέτρων μετασχηματισμού

Οι παράμετροι του μετασχηματισμού υπολογίζονται βελτιστοποιώντας μια μετρική ομοιότητας μεταξύ των ατλάντων και της εικόνας στόχου. Ο αλγόριθμος βελτιστοποίησης μπορεί να ποικείλει βάση του μοντέλου μετασχηματισμού και της μετρικής ομοιότητας που χρησιμοποιούνται.

4. Εφαρμογή του μετασχηματισμού

Αφού έχουν υπολογιστεί οι βέλτιστες παράμετροι μετασχηματισμού, χρησιμοποιούνται για να προσαρμόσουν τους άτλαντες στην εικόνα στόχο.



Σχήμα 3.1: Γραμμικός και μη γραμμικός μετασχηματισμός εικόνας κατά τη διαδικασία εγγραφής στην εικόνα στόχο

3.2 Επιλογή Ομοίων Ατλάντων

Έχοντας εφαρμόσει εγγραφή του σετ ατλάντων στην εικόνα προς κατάτμηση, ακολουθεί η διαδικασία της επιλογής ενός υποσετ που αποτελείται από τους ομοιότερους άτλαντες. Σκοπός της διαδικασίας αυτής είναι να μειώσει το υπολογιστικό κόστος μειώνοντας το μέγευθος του σετ των διαθέσιμων ατλάντων και να βελτιώσει την ακρίβεια της τελικής κατάτμησης, επιλέγοντας τους ομοιότερους άτλαντες με βάση την καθολική τους ομοιότητα με την εικόνα προς κατάτμηση[9]. Υπάρχουν διάφορες μετρικές ομοιότητας φωτεινότητας που χρησιμοποιούνται ευρέως για τη σύγχριση εικόνων, με τις καταλληλότερες για σύγχριση ιατρικών εικόνων να είναι οι παρακάτω:

Mean Squared Error(MSE)

Η μετρική MSE μετράει τη μέση τετραγωνική διαφορά μεταξύ των φωτεινοτήτων των στοιχείων δύο εικόνων.

Mean Absolut Error(MAE)

Η μετρική MAE μετράει τη μέση απόλυτη διαφορά μεταξύ των φωτεινοτήτων των στοιχείων δύο εικόνων.

Mutual information

Η μετρική Mutual Information υπολογίζει το ποσό κοινής πληροφορίας φωτεινότητας που μοιράζονται δύο εικόνες.

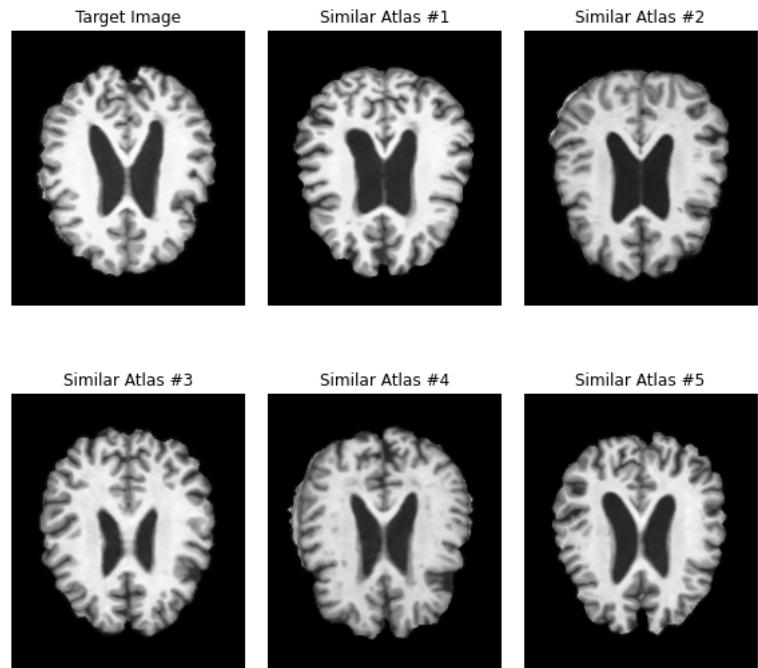
Correlation coefficient

Η μετρική Correlation υπολογίζει τη γραμμική σχέση μεταξύ των φωτεινοτήτων δύο εικόνων.

Structural similarity index (SSIM)

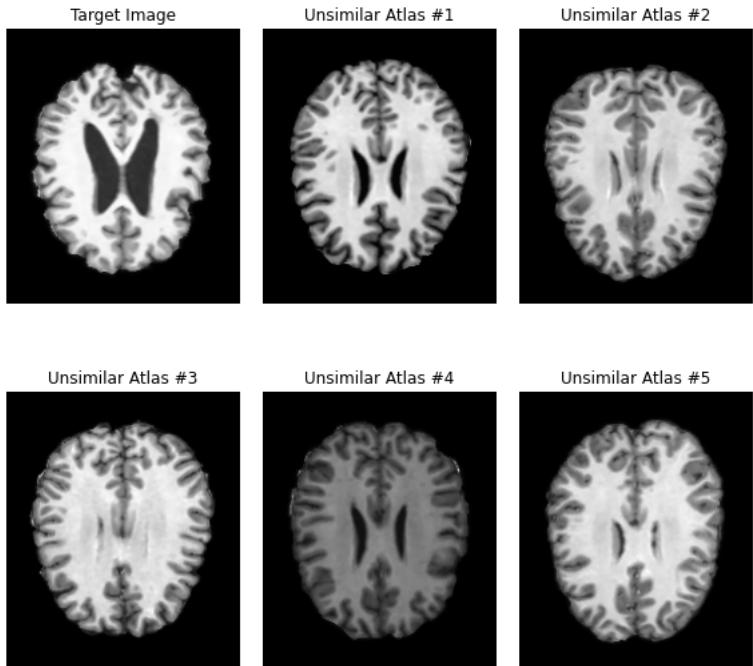
Η μετρική SSIM υπολογίζει την ομοιότητα των δομικών μοτίβων μεταξύ δύο εικόνων.

Πολλές φορές είναι αναγκαίος ο συνδυασμός μετρικών ομοιότητας για την επίτευξη βέλτιστου αποτελέσματος, διαμορφώνοντας το σετ εκπαίδευσης με τον συνδυασμό ατλάντων που επιλέχθηκαν από τη χρήση διαφορετικών μετρικών. Στο Σχήμα 3.2 βλέπουμε στο πάνω αριστερά μέρος μια τομή της εικόνας προς κατάτμηση. Οι υπόλοιπες εικόνες αναπαριστούν την ίδια τομή των ατλάντων που επιλέχθηκαν ως οι ομοιότεροι με την εικόνα προς κατάτμηση, σύμφωνα με τη μετρική MAE.



Σχήμα 3.2: Η εικόνα στόχος και οι ομοιότεροι άτλαντες του σετ σύμφωνα με τη μετρική MAE.

Στο Σχήμα 3.3 βλέπουμε για την ίδια εικόνα προς κατάτμηση και τομή, τις πιο ανόμοιες εικόνες του σετ, σύμφωνα με τη μετρική MAE. Παρατηρούμε πως υπάρχει μεγάλη διαφορά στη δομή των εγκεφάλων και την κατανομή φωτεινότητας που θα αποτρέψει το μοντέλο από τη βέλτιστη εκπαίδευσή του.



Σχήμα 3.3: Η εικόνα στόχος και οι πέντε ανόμοιοι άτλαντες του σετ σύμφωνα με τη μετρική MAE.

3.3 Συγχώνευση ετικετών

Τελευταίο στάδιο της Multi-Atlas κατάτμησης, μετά την εγγραφή των ατλάντων στην εικόνα προς κατάτμηση και την επιλογή των ομοιότερων εξ αυτών ως προς την εικόνα στόχο με τη χρήση των μετρικών που αναφέρθηκαν προηγουμένως, αποτελεί η συγχώνευση των ετικετών τους ώστε να προκύψει η τελική κατάτμηση. Η βασική ιδέα είναι να χρησιμοποιηθεί η ομοιότητα μεταξύ των μη κατηγοριοποιημένων τοπικών περιοχών της εικόνας στόχου και των αντίστοιχων κατηγοριοποιημένων τοπικών περιοχών των ατλάντων, έτσι ώστε να διαδοθούν οι ετικέτες στην εικόνα στόχο. Η διαδικασία ακολουθεί τα παρακάτω βήματα:

1. Υπολογισμός τοπικής ομοιότητας

Σε αυτό το βήμα υπολογίζεται η τοπική ομοιότητα υπο-περιοχών των ατλάντων και των αντίστοιχων υπο-περιοχών της εικόνας στόχου χρησιμοποιώντας κάποια

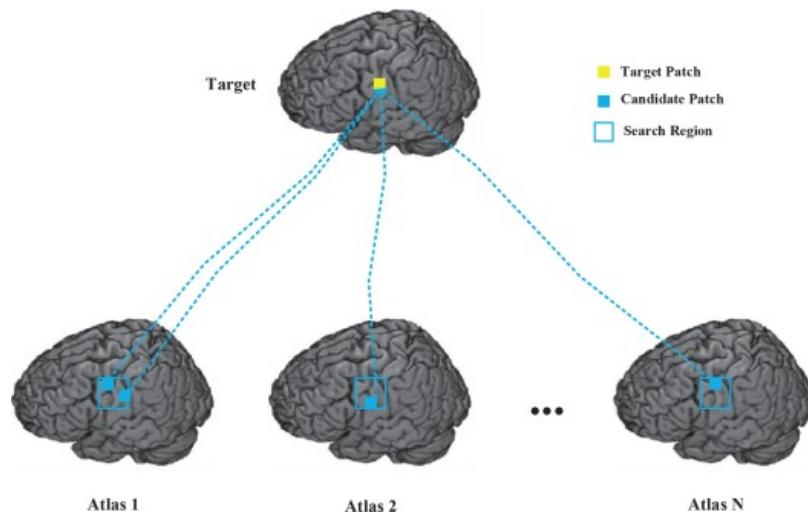
μετρική ομοιότητας όπως cross correlation ή mutual information.

2. Διάδοση ετικετών

Σε αυτό το βήμα οι ετικέτες των κατηγοριοποιημένων υπο-περιοχών των ατλάντων διαδίδονται στις αντίστοιχες μη κατηγοριοποιημένες περιοχές της εικόνας προς κατάτμηση. Η διάδοση των ετικετών αποτελεί ουσιαστικά τον υπολογισμό μιας σταθμισμένης ετικέτας για την εκάστοτε υποπεριοχή του κάθε άτλαντα[10]. Τα βάρη που δίνονται στην κάθε ετικέτα είναι σχετικά με την ομοιότητα μεταξύ της κατηγοριοποιημένης υπο-περιοχής του κάθε άτλαντα και της μη κατηγοριοποιημένης υπο-περιοχής της εικόνας προς κατάτμηση.

3. Συγχώνευση ετικετών

Οι ετικέτες που διαδόθηκαν από τον κάθε άτλαντα συνδυάζονται για την κάθε υποπεριοχή της εικόνας, χρησιμοποιώντας κάποια τεχνική συγχώνευσης ετικετών, όπως majority voting ή weighted voting για να προκύψει η τελική ετικέτα της κάθε υπο-περιοχής.



Σχήμα 3.4: Διάδοση σταθμισμένων ετικετών των ατλάντων προς την εικόνα στόχου σύμφωνα με την τοπική ομοιότητα των υπο-περιοχών των ατάντων με την υπο-περιοχή της εικόνας στόχου.

4 Ανάλυση Προτεινόμενου Μοντέλου

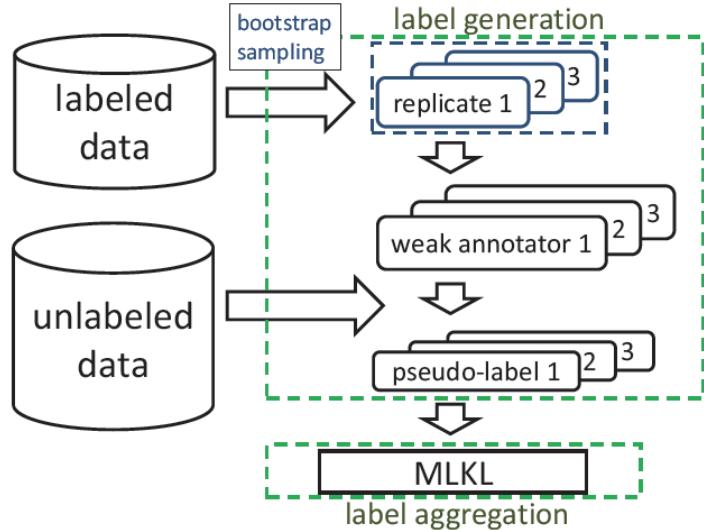
Τα τελευταία χρόνια οι εφαρμογές διαπραγματεύονται μεγάλο όγκο δεδομένων με τη συλλογή τους να γίνεται όλο και πιο προσιτή ενώ η διαδικασία της κατηγοριοποίησης τους ενέχει μεγάλο υπολογιστικό και χρονικό κόστος, καθώς απαιτεί εκτεταμένους υπολογιστικούς πόρους αλλά και εξειδικευμένο ανθρώπινο δυναμικό. Για τον λόγο αυτό, είναι αναγκαία η χρήση της ημιεπιβλεπόμενης μάθησης σε αλγορίθμους ταξινόμησης ή κατάταξης. Όπως αναφέρθηκε παραπάνω, η εκπαίδευση μοντέλων ημιεπιβλεπόμενης μάθησης στηρίζεται στην χρήση μικρού μέρους ταξινομημένων δεδομένων συνδυαστικά με μη-ταξινομημένα δεδομένα, σε αντίθεση με την συμβατική επιβλεπόμενη μάθηση που χρησιμοποιεί μεγάλο όγκο διαθέσιμων ταξινομημένων δεδομένων. Ένα πεδίο στο οποίο παρατηρείται αυξανόμενη συλλογή δεδομένων και παράλληλα αποτελεί πρόσφορο έδαφος για την εφαρμογή πλήθους αλγορίθμων μηχανικής μάθησης είναι ο τομέας της βιοιατρικής, στον οποίο επαφίεται και η παρούσα διπλωματική.

4.1 Θεωρητικό υπόβαθρο αλγορίθμου

Το προτεινόμενο μοντέλο βασίζεται στην χρήση τεχνικών Support Vector Machines (SVM), την εκμετάλλευση τόσο ταξινομημένων όσο και αταξινόμητων δεδομένων και την σύνθεση προβλέψεων ενός πλήθους αδύναμων ταξινομητών όπως περιγράφηκε από τον Yan[11]. Σε πρώτη φάση, ο κάθε ταξινονητής εκπαιδεύεται σε μία bootstrap ρέπλικα του συνολικού σετ ταξινομημένων δεδομένων. Οι ταξινομητές χαρακτηρίζονται αδύναμοι διότι εκπαιδεύονται σε περιορισμένο αριθμό ταξινομημένων εγγραφών, γεγονός που κατατάσσει την υλοποίηση αυτή στο φάσμα της ημιεπιβλεπόμενης μάθησης. Στην συνέχεια, ο κάθε ταξινομητής καλείται να προβλέψει πάνω σε σετ αταξινόμητων δεδομένων, παράγοντας κατά αυτόν τον τρόπο από ένα διάνυσμα ψευδοπροβλέψεων. Ακολούθως, με την χρήση της τεχνικής label aggregation(συγχώνευση ετικετών), όπως προτάθηκε από τον Li[12], οι ψευδοπροβλέψεις των εκτιμητών συνδυάζονται για να προκύψει ένα διάνυσμα τελικής εκτίμησης.

Έστω ένα πρόβλημα δυαδικής ταξινόμησης, με το $D_L = \{\mathbf{x}_i, y_i\}_{i=1}^L$ να δηλώνει το σετ των ταξινομημένων δεδομένων,
όπου

$$\xi_i \in \mathbb{R}^d$$



Σχήμα 4.1: Δημιουργία *bootstrap* δειγμάτων από το σετ κατηγοριοποιημένων δεδομένων και εκπαίδευση Μ αδύναμων SVM ταξινομητών. Διάδοση των Μ ψευδοετικετών που παρήχθησαν για την εξαγωγή της τελικής πρόβλεψης μέσω MLKL

και

$$y_i \in \{-1, +1\}$$

δηλώνουν το διάνυσμα χαρακτηριστικών και την πρόβλεψη του i -ου δεήματος ανίστοιχα. Η γενική αντικειμενική συνάρτηση ορίζεται ως ακολούθως:

$$\min_w \Omega(\mathbf{w}) + C l(\mathbf{w}) \quad (4.1)$$

όπου

$\mathbf{w} \in \mathbb{R}^d$ το διάνυσμα των βαρών,

$\Omega(\mathbf{w})$ ο ορος κανονικοποίησης,

$l(\mathbf{w})$ η συνάρτηση κόστους

και

C η παράμετρος κανονικοποίησης

Σε αυτή την ανάλυση θα εστιάσουμε στην l_2 -κανονικοποιημένη μορφή της συνάρτησης κόστους.

Η αντικειμενική συνάρτηση του *hinge loss* μπορεί να γραφτεί με τον εξής τρόπο:

$$\min_w \frac{1}{2} \|\mathbf{w}\|_2^2 + C \sum_{i=1}^l \xi_i \quad (4.2)$$

$$\text{s.t. } y_i \mathbf{w}^T \mathbf{x}_i \geq 1 - \xi_i, \xi_i \geq 0, i = 1, \dots, l. \quad (4.3)$$

όπου ξ_i η μεταβλητή χαλάρωσης του i -οστου δείγματος. Αξιοποιώντας την πληθύρα των μη κατηγοριοποιημένων δεδομένων, ορίζουμε ως $D_U = \{\mathbf{x}_i\}_{i=l+1}^n$ το σετ των μη κατηγοριοποιημένων δεδομένων και εισάγουμε την πληροφορία τους στην αντικειμενική συνάρτηση, η οποία μπορεί να γραφτεί ως εξής:

$$\min_{\bar{y} \in Y} \min_w \frac{1}{2} \|\mathbf{w}\|_2^2 + C_1 \sum_{i=1}^l \xi_i + C_2 \sum_{i=l+1}^n \xi_i \quad (4.4)$$

$$\text{s.t. } \bar{y}_i \mathbf{w}^T \mathbf{x}_i \geq 1 - \xi_i, \xi_i \geq 0, i = 1, \dots, n, \quad (4.5)$$

όπου C_1, C_2 οι παράμετροι κανονικοποίησης που ελέγχουν τον συμβιβασμό μεταξύ της πολυπλοκότητας του μοντέλου, του κόστους που προκύπτει από τα κατηγοριοποιημένα δεδομένα και του κόστους που προκύπτει από τα μη κατηγοριοποιημένα δεδομένα και

$$Y = \{\bar{y} | \bar{y} = [\mathbf{y}_L; \tilde{\mathbf{y}}_U], \tilde{\mathbf{y}}_U \in \{-1, +1\}^{n-l}\}$$

όπου $y_L \in \mathbb{R}^l$ αναπαριστά το διάνυσμα προβλέψεων για τα κατηγοριοποιημένα δεδομένα, και το $\tilde{\mathbf{y}}_U$ αναπαριστά οποιεσδήποτε προβλέψεις για τα μη κατηγοριοποιημένα δεδομένα. Επομένως, υπάρχουν εκθετικά πιθανές τιμές για τα y_U , δηλαδή τις προβλέψεις των μη κατηγοριοποιημένων δεδομένων, πράγμα που καθιστά το πρόβλημα βελτιστοποίησης δυσεπίλυτο.

Κάνοντας χρήση των δυαδικών μεταβλητών $\alpha \in \mathbb{R}^n$, η Λανγκραντζιανή της (4.4) δίνεται από τον τύπο:

$$\begin{aligned} L(\mathbf{w}, \boldsymbol{\alpha}) &= \frac{1}{2} \|\mathbf{w}\|_2^2 + C_1 \sum_{i=1}^l \xi_i + C_2 \sum_{i=l+1}^n \xi_i \\ &\quad + \sum_{i=1}^n \alpha_i (1 - \xi_i - \bar{y}_i \mathbf{w}^T \mathbf{x}_i). \end{aligned} \quad (4.6)$$

Θέτοντας τις μερικές παραγώγους της L προς τα w και ξ_i 0, η Λανγκρατζιανή προκύπτει ως:

$$L = -\frac{1}{2}\alpha^T \left((XX^T) \cdot \left(\sum_{m=1}^M \mu_m \tilde{y}_m \tilde{y}_m^T \right) \right) \alpha + 1^T \alpha, \quad (4.7)$$

όπου

$$\alpha \in \mathbb{A}$$

και

$$\mathbb{A} = \{\alpha | 0 \leq \alpha_i \leq C_1, 0 \leq \alpha_j \leq C_2, 1 \leq i \leq l, l+1 \leq j \leq n\}$$

Μπορούμε τώρα να αντικαταστήσουμε το πρωταρχικό primal πρόβλημα της (4.4) με το δυαδικό dual:

$$\min_{\tilde{y} \in Y} \max_{\alpha \in A} -\frac{1}{2}\alpha^T \left((XX^T) \cdot \tilde{y} \tilde{y}^T \right) \alpha + 1^T \alpha, \quad (4.8)$$

όπου

$$X = [x_1, x_2, \dots, x_n]^T$$

Και σε αυτή τη μορφή του προβλήματος είναι συνήθως δύσκολο να βελτιστοποιηθεί το υ λόγω του αυξημένου αριθμού πιθανών τιμών.

Σε αυτό το σημείο καλούμαστε να λύσουμε το πρόβλημα μέσω MLKL, εκπαιδεύοντας ένα σετ από αδύναμους εκτιμητές.

Δεδομένου ενός σετ προβλέψεων μεγέθους M

$$\{\tilde{y}_1, \tilde{y}_2, \dots, \tilde{y}_M\}$$

μέσω του Multi Label Kernel Learning (MLKL) το πρόβλημα μετασχηματίζεται σε:

$$\min_{\mu \in U} \max_{\alpha \in A} -\frac{1}{2}\alpha^T \left((XX^T) \cdot \sum_{m=1}^M \mu_m \tilde{y}_m \tilde{y}_m^T \right) \alpha + 1^T \alpha, \quad (4.9)$$

που στοχεύει στην εύρεση ενός σταθμισμένου συνδυασμού πυρήνων των διανυσμάτων προβλέψεων $\sum_{m=1}^M \mu_m \tilde{y}_m \tilde{y}_m^T$ προσεγγίζοντας έτσι τις πραγματικές προβλέψεις $\tilde{y}^* \tilde{y}^{*T}$, όπου

$$U = \{\mu \sum_{m=1}^M \mu_m = 1, \mu_m \geq 0\},$$

$$\mathbb{A} = \{\boldsymbol{\alpha} | 0 \leq \alpha_1 \leq C_1, 0 \leq \alpha_j \leq C_2, 1 \leq i \leq l, l+1 \leq j \leq n\},$$

και

$$\mu = [\mu_1, \mu_2, \dots, \mu_M]^T$$

το διάνυσμα των βαρών των πυρήνων βάσης. Σε πρώτη φάση, δημιουργείται ένα σετ από ψευδοεκτιμήσεις. Κάθε διάνυσμα ψευδοεκτιμήσεων προέρχεται από έναν αδύναμο εκτιμητή που εκπαιδεύεται με μία bootstrap ρεπλικα του σετ κατηγοριοποιημένων δεδομένων. Για την εκπαίδευση τους οι εκτημητές λύνουν το δυαδικό κλασικό πρόβλημα SVM.

Σε δεύτερη φάση δημιουργείται ένας MLKL ταξινομητής. Έστω ότι από το προηγούμενο βήμα έχουν προκύψει M ψευδοεκτιμήσεις, ονόματι

$$Y_m = \{\tilde{\mathbf{y}}_1, \tilde{\mathbf{y}}_2, \dots, \tilde{\mathbf{y}}_M\}$$

, τότε ανακτούμε το πρωταρχικό του προβλήματος (4.9) ως,

$$\begin{aligned} & \min_{\mu \in U, w, m, \xi} \frac{1}{2} \sum_{m=1}^M \frac{1}{\mu_m} \|\mathbf{w}_m\|_2^2 + C_1 \sum_{i=1}^l \xi_i + C_2 \sum_{i=l+1}^n \xi_i \\ & \text{s.t. } \sum_{m=1}^M \hat{y}_{mi} \mathbf{w}_m x_i \geq 1 - \xi_i, \xi_i \geq 0, i = 1, \dots, n \end{aligned} \quad (4.10)$$

όπου \hat{y}_{mi} η πρόβλεψη για το i -οστό δείγμα του διανύσματος \mathbf{y}_m .

Θέτοντας

$$\begin{aligned} \hat{\mathbf{w}} &= \left[\frac{\mathbf{w}_1}{\sqrt{\mu_1}}, \dots, \frac{\mathbf{w}_M}{\sqrt{\mu_M}} \right]^T, \\ \hat{\mathbf{x}}_i &= [\sqrt{\mu_1} \mathbf{x}_i, \sqrt{\mu_2} \tilde{y}_1, \tilde{y}_2 \mathbf{x}_i, \dots, \sqrt{\mu_T} \tilde{y}_1, \tilde{y}_{M_i} \mathbf{x}_i]^T, \end{aligned}$$

και

$$\mathbf{y} = \hat{\mathbf{y}}_1$$

το πρωταρχικό πρόβλημα MLKL γίνεται:

$$\begin{aligned} & \min_{\hat{\mathbf{w}}, \xi} \frac{1}{2} \|\mathbf{w}\|_F^2 + C_1 \sum_{i=1}^l \xi_i + C_2 \sum_{i=l+1}^n \xi_i \\ & \text{s.t. } \hat{y}_i \hat{\mathbf{w}}^T \hat{\mathbf{x}}_i \geq 1 - \xi_i, \xi_i \geq 0, i = 1, \dots, n \end{aligned} \quad (4.11)$$

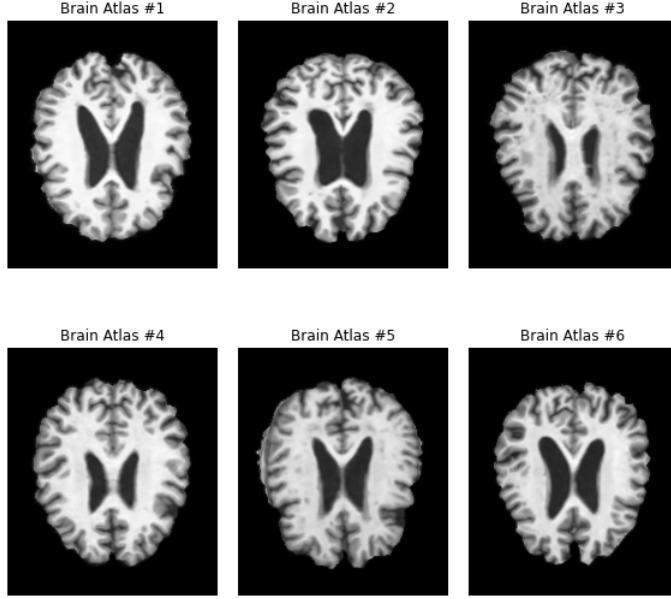
Το πρόβλημα (4.11) είναι παρόμοιο με ένα κλασσικό πρωταρχικό πρόβλημα SVM και μπορεί να λυθεί με υπάρχοντες αλγορίθμους.

Αλγόριθμος 1 Αλγόριθμος ημι-επιβλεπόμενης μάθησης πολλαπλών ταξινομητών.

- 1: Αρχικοποίησε M , αριθμό αδύναμων ταξινομητών
 - 2: **για** $k = 1$ μέχρι K **επανάλαβε**
 - 3: $\Delta_{\text{ημιούργηση}} \quad M \quad \text{bootstrap} \quad \text{δείγματα}$
 $\{(\bar{\mathbf{X}}_1, \bar{y}_{k1}), (\bar{\mathbf{X}}_2, \bar{y}_{k2}), \dots, (\bar{\mathbf{X}}_T, \bar{y}_{kM})\}$ από το σετ κατηγοριοποιημένων δεδομένων \mathcal{D}_L .
 - 4: **για** $m = 1$ μέχρι M **επανάλαβε**
 - 5: Προπόνησε ένα SVM μοντέλο \mathcal{M}_{km} πάνω στα $\bar{\mathbf{X}}_m$ και \bar{y}_{km} .
 - 6: Πάραξε \hat{y}_{km} προβλέποντας πάνω στο σετ μη κατηγοριοποιημένων δεδομένων X_U χρησιμοποιώντας το \mathcal{M}_{km} .
 - 7: Πρόσθεσε το \hat{y}_{km} στο σετ \mathcal{Y}_{km} .
 - 8: **τέλος για**
 - 9: Υπολόγισε $\{\mathbf{w}_{k1}, \mathbf{w}_{k2}, \dots, \mathbf{w}_{kM}\}$ και $\boldsymbol{\mu}_k$ λύνοντας το πρόβλημα (4.10).
 - 10: Υπολόγισε πρόβλεψη $p_{jk} = \sum_{m=1}^M \mu_{km} \mathbf{w}_{km}^T \mathbf{x}_j$ για τεστ δεδομένο x_j .
 - 11: **τέλος για**
 - 12: Επίλεξε την κλάση για x_j από $\text{argmax}_k \{p_{jk}\}_{k=1}^K$.
-

4.2 Εφαρμογή σε ιατρικές εικόνες

Το προτεινόμενο μοντέλο προσαρμόζει τον αλγόριθμό που αναλύθηκε στην προηγούμενη υποενότητα έτσι ώστε να βρει εφαρμογή στην κατάτμηση ιατρικών εικόνων, συνδυάζοντας τον με την ιδέα της Multi Atlas κατάτμησης που αναλύθηκε στην ενότητα 3. Η βασική ιδέα είναι να εκπαιδευτεί ο κάθε αδύναμος SVM ταξινομητής του Αλγορίθμου 1 σε δείγματα διαφορετικών ατλάντων του σετ δεδομένων, και να παραχθεί η τελική κατάτμηση της εικόνας στόχου μέσω label aggregation των προβλέψεων του κάθε ταξινομητή, αντίστοιχα με το βήμα 10 του Αλγορίθμου 1.



Σχήμα 4.2: Σετ ατλάντων εκπαίδευσης. Κάθε SVM ταξινομητής δειγματοληπτεί από έναν άτλαντα του σετ.

Θεωρώντας ένα αρχικό σετ N ατλάντων \mathcal{D}_{A0} υέτουμε έναν τυχαίο άτλαντα ως εικόνα προς κατάτμηση και τους υπόλοιπους $N - 1$ ως σετ εκπαίδευσης \mathcal{D}_{A0} . Αντίστοιχα με το βήμα 1 της Multi-Atlas κατάτμησης (Ενότητα 3.1), γίνεται εγγραφή των ατλάντων πάνω στην εικόνα προς κατάτμηση. Έπειτα, αντίστοιχα με το βήμα 2 της Multi-Atlas κατάτμησης (Ενότητα 3.2), επιλέγονται από το σετ εκπαίδευσης οι M πιο όμοιοι άτλαντες με την εικόνα στόχο, όπου η βέλστιση της M προκύπτει πειραματικά διαμορφώνοντας το νέο σετ εκπαίδευσης \mathcal{D}_{AL} , το οποίο αποτελεί το σετ κατηγοριοποιημένων δεδομένων, αντιστοιχα με το σετ \mathcal{D}_L του αλγορίθμου 1. Από την εικόνα στόχο έπειτα από δειγματοληψία διαμορφώνεται το σετ μη κατηγοριοποιημένων δεδομένων \mathcal{D}_{AU} , το οποίο χωρίζεται στη συνέχεια στο σετ εκπαίδευσης μη κατηγοριοποιημένων δεδομένων \mathcal{D}_{AUTr} , αντίστοιχο με το σετ \mathcal{Q}_U του αλγορίθμου 1, και στο τεστ σετ \mathcal{D}_{AUTe} . Στη συνέχεια, ακολουθείται το βήμα 3 της Multi-Atlas κατάτμησης (Ενοτητα 3.3), εφαρμόζοντας όμως τον αλγόριθμο 1 ως εξής: Εφαρμόζουμε την τεχνική one-vs-all που αναπτύχθηκε στην ενότητα 2.4, ανατρέχουμε την εξωτερική λούπα (2) του αλγορίθμου 1 K φορές, όσες και οι

κλάσεις του προβλήματος, δημιουργώντας ένα δυαδικό πρόβλημα που αποτελέιται από δύο κλάσεις, την εξεταζόμενη από τη λούπα κλάση και τις υπόλοιπες κλάσεις ως μια. Ο κάθε αδύναμος ταξινομητής αντλεί ένα δείγμα $(\bar{\mathbf{X}}_m, \bar{y}_{km})$ από έναν άτλαντα του σετ εκπαίδευσης \mathcal{D}_{AL} . Έτσι ο κάθε ταξινομητής εκπαιδεύεται πάνω στα δεδομένα ενός διαφορετικού άτλαντα του σετ και παράγει ένα διάνυσμα φευτοπροβλέψεων \hat{y}_{km} , προβλέποντας πάνω στο σετ μη κατηγοριοποιημένων δεδομένων εκπαίδευσης \mathcal{D}_{AUTr} , το οποίο προέρχεται από την εικόνα προς κατάτμηση. Στη συνέχεια, όπως στο βήμα 2 της Ενότητας 3.3, γίνεται διάδοση των διανυσμάτων φευτοεπικετών \hat{y}_{km} που έχουν προκύψει από τους SVM ταξινομητές και υπολογίζονται τα βάρη τους μ_k και το διάνυσμα $\mathbf{w}_{k1}, \mathbf{w}_{k2}, \dots, \mathbf{w}_{kM} \}$ λύνοντας το πρόβλημα (4.10) του αλγορίθμου 1. Σε αντίθεση με την Multi-Atlas κατάτμηση που ακολουθεί επαναληπτική διαδικασία για την εύρεση των βαρών, ο αλγόριθμος 1 απαιτεί μόνο 1 επανάληψη. Στη συνέχεια όπως στο βήμα 3 της Ενότητας 3.3 γίνεται συγχώνευση των προβλέψεων του κάθε SVM ταξινομητή και παράγεται μια πρόβλεψη p_{jk} για το τεστ δεδομένο x_j του τεστ σετ \mathcal{D}_{AUTe} . Τέλος, αφού η εξωτερική λούπα (2) τρέξει για όλες τις κλάσεις, επιλέγουμε την πρόβλεψη της κλάσης με τη μεγαλύτερη τιμή και την αποδίδουμε ως ετικέτα στο τεστ δεδομένο x_j .

Αλγόριθμος 2 Προτεινόμενος Αλγόριθμος.

- 1: Αρχικοποίησε M , αριθμό αδύναμων ταξινομητών
 - 2: **για** $k = 1$ μέχρι K **επανάλαβε**
 - 3: Δημιουργησε M δείγματα $\{(\bar{\mathbf{X}}_1, \bar{y}_{k1}), (\bar{\mathbf{X}}_2, \bar{y}_{k2}), \dots, (\bar{\mathbf{X}}_T, \bar{y}_{kM})\}$ από το σετ κατηγοριοποιημένων δεδομένων των όμοιων ατλαντών \mathcal{D}_{AL} .
 - 4: **για** $m = 1$ μέχρι M **επανάλαβε**
 - 5: Προπόνησε ένα SVM μοντέλο \mathcal{M}_{km} πάνω στα $\bar{\mathbf{X}}_m$ και \bar{y}_{km} .
 - 6: Πάραξε \hat{y}_{km} προβλέποντας πάνω στο σετ εκπαίδευσης μη κατηγοριοποιημένων δεδομένων ατλαντών \mathcal{D}_{AUTr} χρησιμοποιώντας το \mathcal{M}_{km} .
 - 7: Πρόσθεσε το \hat{y}_{km} στο σετ \mathcal{Y}_{kM} .
 - 8: **τέλος για**
 - 9: Υπολόγισε $\{\mathbf{w}_{k1}, \mathbf{w}_{k2}, \dots, \mathbf{w}_{kM}\}$ και μ_k λύνοντας το πρόβλημα (4.10).
 - 10: Υπολόγισε πρόβλεψη $p_{jk} = \sum_{m=1}^M \mu_{km} \mathbf{w}_{km}^T x_j$ για τεστ δεδομένο x_j του τεστ σετ \mathcal{D}_{AUTe} .
 - 11: **τέλος για**
 - 12: Επίλεξε την κλάση για x_j από $\text{argmax}_k \{p_{jk}\}_{k=1}^K$.
-

5 Εξαγωγή Χαρακτηριστικών

5.1 Οικογένειες Χαρακτηριστικών Εικόνας

Στην αναγνώριση εικονας τα δεδομένα προς επεξεργασία είναι ουσιαστικά ένας πίνακας των pixels(εικονοστοιχείων) (2D για εικόνες και 3D για όγκους) με το περιεχόμενο των pixels να είναι η φωτεινότητά τους. Για τη δημιουργία μοντέλων μηχανικής μάθησης, είναι απαραίτητη η εξαγωγή περαιτέρω χαρακτηριστικών που περιγράφουν επαρκώς το περιεχόμενο της εικόνας. Τα χαρακτηριστικά μπορούν να εξαχθούν τόσο από 2D όσο και από 3D εικόνες. Στην περίπτωση των 2D εικόνων, υπολογίζονται χαρακτηριστικά για ένα pixel λαμβάνοντας υπόψη το περιβάλλον τετράγωνο με κέντρο αυτό, ενώ για 3D εικόνες υπολογίζονται χαρακτηριστικά για ένα voxel(ογκοστοιχείο), λαμβάνοντας υπόψη τον περιβάλλοντα κυβο. Οι κυριότερες κατηγορίες χαρακτηριστικών είκόνων είναι:

- Στατιστικά χαρακτηριστικά πρώτης τάξης
- Χαρακτηριστικά υφής
- Τοπολογικά χαρακτηριστικά

και αναλύονται παρακάτω.

5.1.1 Στατιστικά Χαρακτηριστικά

Τα στατιστικά χαρακτηριστικά ή στατιστικά πρώτης τάξης είναι τα πιο ευνόητα και ενστικτώδη προς υπολογισμό χαρακτηριστικά που χρησιμοποιούνται σε εφαρμογές ανάλυσης εικόνας. Υπολογίζονται από ιστογράμματα που προκύπτουν από τις τιμές φωτεινότητας που απαντώνται στην εικόνα. Περιέχουν πληροφορία που προκύπτει από τις φωτεινότητες των μεμονωμένων pixels αλλά αγνοούν την χωρική αλληλεπίδραση μεταξύ τους [13]. Τα χαρακτηριστικά ιστογραμμάτων δεν αντανακλούν αντικείμενα ή μοτίβα στην εικόνα, μόνο την κατανομή των τιμών της φωτεινότητας. Αυτό το μειονέκτημα τους τα καθιστά ένα ακατάλληλο εργαλείο για την ποσοτικοποίηση αλλαγών στις εικόνες, ή οποιασδήποτε μεταβολή στη χωρική κατανομή των τιμών φωτεινότητας αλλά ο υπολογισμός τους ενέχει μικρό υπολογιστικό κόστος και ταχύτητα, δημιουργώντας μια εντύπωση για την κατανομή της φωτεινότητας στην εικόνα.

Θεωρούμε:

-
- \mathbf{X} ένα σετ που αποτελείται από N_p voxels που συμπεριλαμβάνονται στο πεδίο ενδιαφέροντος.
 - $\mathbf{P}(i)$ το ιστόγραμμα πρώτης τάξης με N_g διαχριτές τιμές φωτεινότητας, όπου N_g είναι ο αριθμός των μη μηδενικών bins.
 - $p(i)$ το κανονικοποιημένο ιστόγραμμα πρώτης τάξης, ίσο με $\frac{\mathbf{P}(i)}{N_p}$

Τα στατιστικά που προκύπτουν από το ιστόγραμμα πρώτης τάξης περιγράφονται παρακάτω.

1. Energy

$$Energy = \sum_{i=1}^{N_p} (\mathbf{X}(i) + c)^2 \quad (5.1)$$

Η Energy μετράει το πλάτος των τιμών φωτεινότητων των voxels της εικόνας. Μια μεγάλη τιμή καταδεικνύει ένα μεγάλο άθροισμα των τετραγώνων αυτών των τιμών.

2. Total Energy

$$Total\ Energy = V_{voxel} \sum_{i=1}^{N_p} (\mathbf{X}(i) + c)^2 \quad (5.2)$$

Η Total Energy είναι η μετρική Energy κλιμακωμένη από τον όγκο του voxel σε κυβικά mm.

3. Entropy

$$Entropy = - \sum_{i=1}^{N_g} p(i) \log_2 (p(i) + \epsilon) \quad (5.3)$$

Η Entropy ορίζει την αβεβαιότητα στις τιμές της εικόνας. Μετράει το μέσο ποσό της πληροφορίας που χρειάζεται για να κωδικοποιηθεί η εικόνα.

4. Minimum

$$Minimum = \min(\mathbf{X}) \quad (5.4)$$

Η τιμή της ελάχιστης φωτεινότητας στο πεδίο ενδιαφέροντος.

5. 10th percentile

Το δέκατο εκατοστημόριο της X.

6. 90th percentile

Το ενενηκοστό εκατοστημόριο της X.

7. Maximum

$$\text{Maximum} = \max(\mathbf{X}) \quad (5.5)$$

Η τιμή της μέγιστης φωτεινότητας στο πεδίο ενδιαφέροντος.

8. Mean

$$\text{Mean} = \frac{1}{N_p} \sum_{i=1}^{N_p} \mathbf{X}(i) \quad (5.6)$$

Η Mean υπολογίζει τη μέση τιμή των φωτεινοτήτων της εικόνας.

9. Median

Η διάμεση τιμή των φωτεινοτήτων.

10. Interquartile Range

$$\text{interquartile range} = \mathbf{P}_{75} - \mathbf{P}_{25} \quad (5.7)$$

όπου το \mathbf{P}_{75} και το \mathbf{P}_{25} είναι το 75ο και το 25ο εκατοστημόριο της εικόνας.

11. Range

$$\text{Range} = \max(\mathbf{X}) - \min(\mathbf{X}) \quad (5.8)$$

Το εύρος των τιμών φωτεινοτήτων στην εικόνα.

12. Mean Absolute Deviation (MAD)

$$MAD = \frac{1}{N_p} \sum_{i=1}^{N_p} |\mathbf{X}(i) - \bar{X}| \quad (5.9)$$

Η *MAD* υπολογίζει τη μέση απόσταση των φωτεινοτήτων της εικόνας από τη μέση τιμή της εικόνας.

13. Robust Mean Absolute Deviation (rMAD)

$$rMAD = \frac{1}{N_{10-90}} \sum_{i=1}^{N_{10-90}} |\mathbf{X}(i)_{10-90} - \bar{X}_{10-90}| \quad (5.10)$$

Η *rMAD* υπολογίζει τη μέση απόσταση των φωτεινοτήτων από τη μέση τιμή, υπολογισμένη στο υποσύνολο τιμών της εικόνας που βρίσκονται μεταξύ η ίσες του 10ου και 90ου εκατοστημορίου της εικόνας.

14. Root Mean Squared (RMS)

$$RMS = \sqrt{\frac{1}{N_p} \sum_{i=1}^{N_p} (\mathbf{X}(i) + c)^2} \quad (5.11)$$

Η *RMS* είναι η ρίζα της μέσης τιμής των τετραγώνων των τιμών φωτεινοτήτων της εικόνας. Είναι ένα ακόμη μέτρο του πλάτους των φωτεινοτήτων της εικόνας.

15. Standard Deviation

$$Standard Deviation = \sqrt{\frac{1}{N_p} \sum_{i=1}^{N_p} (\mathbf{X}(i) - \bar{X})^2} \quad (5.12)$$

Η (standard deviation μετρά τη διακύμανση των τιμών φωτεινοτήτων γύρω από τη μέση τιμή τους. Εξ' ορισμού ισχύει standard deviation = $\sqrt{variance}$.

16. Skewness

$$Skewness = \frac{\mu_3}{\sigma^3} = \frac{\frac{1}{N_p} \sum_{i=1}^{N_p} (\mathbf{X}(i) - \bar{X})^3}{\sqrt{\frac{1}{N_p} \sum_{i=1}^{N_p} (\mathbf{X}(i) - \bar{X})^2}}^3 \quad (5.13)$$

όπου το μ_3 είναι η 3η κεντρική ροπή.

Η *skewness* μετράει την ασυμετρία της κατανομής των τιμών της εικόνας γύρω από την μέση τιμή τους. Αυτή η τιμή μπορεί να είναι θετική ή αρνητική, αναλόγως το που φτάνει η ουρά της κατανομής και το πού είναι συγκεντρωμένη η μάζα της.

17. Kyrtosis

$$Kyrtosis = \frac{\mu_4}{\sigma^4} = \frac{\frac{1}{N_p} \sum_{i=1}^{N_p} (\mathbf{X}(i) - \bar{X})^4}{\left(\frac{1}{N_p} \sum_{i=1}^{N_p} (\mathbf{X}(i) - \bar{X})^2 \right)^2} \quad (5.14)$$

όπου το μ_4 είναι η 4η κεντρική ροπή.

Η *Kyrtosis* είναι μια μετρική της *peakedness* της κατανομής των τιμών της εικόνας. Μία μεγάλη τιμή καταδεικνύει πως η μάζα της κατανομής είνει συγκεντρωμένη στις ουρές της κατανομής και όχι γύρω από τη μέση τιμή.

18. Variance

$$Variance = \frac{1}{N_p} \sum_{i=1}^{N_p} (\mathbf{X}(i) - \bar{X})^2 \quad (5.15)$$

Η *variance* μετρά τη διασπορά των τιμών φωτεινοτήτων γύρω από τη μέση τιμή τους. Εξ' ορισμού ισχύει $variance = \sigma^2$.

19. Uniformity

$$Uniformity = \sum_{i=1}^{N_p} p(i)^2 \quad (5.16)$$

Η *Uniformity* μετρά το άθροισμα των τετραγώνων των τιμών φωτεινοτήτων της εικόνας. Είναι μια μετρική που εκφράζει την ομοιογένεια ενός πίνακα εικόνας, με μια μεγάλη τιμή να καταδεικνύει μεγάλη ομοιογένεια, ή ένα μικρότερο εύρος διαχριτών τιμών φωτεινότητας.

5.1.2 Γεωγραφικά Χαρακτηριστικά

Ως κατηγορία γεωγραφικών χαρακτηριστικών ορίζουμε την ψέση του κάθε voxel προς εξέταση στους άξονες x, y, z . Οι τιμές της ψέσης μπορούν να λειτουργήσουν μεμονωμένα ως χαρακτηριστικά ή να συνδυαστούν και να αποτελέσουν ένα ενιαίο χαρακτηριστικό.

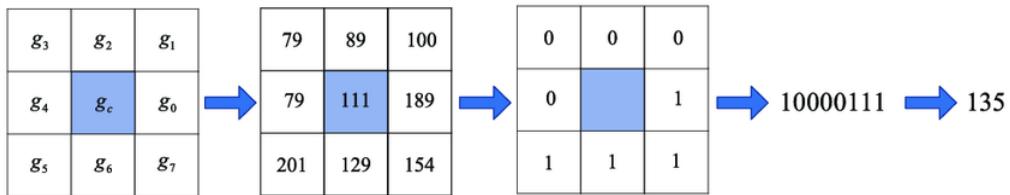
5.1.3 Χαρακτηριστικά Υφής

Η υφή είναι ένα από τα πιο σημαντικά χαρακτηριστικά που χρησιμοποιούνται για την αναγνώριση στοιχείων ή περιοχών που συγκεντρώνουν ενδιαφέρον σε μία εικόνα, είτε ή εικόνα αφορά μια μικροφωτογραφία, μια αεροφωτογραφία ή μια εικόνα μαγνητικής τομογραφίας. Η υφή μπορεί να χαρακτηριστεί ως λεπτή, τραχιά ή απαλή όπως και κυματισμένη, πορώδης, ακανόνιστη ή γραμμωμένη. Περιέχει ουσιώδη πληροφορία για την δομική οργάνωση των επιφανειών και της σχέσης τους με το περιβάλλον τους. Αποτελεί αντίστοιχης σημασίας πληροφορία και για τρισδιάστατες ιατρικές εικόνες που εξάγονται με την τεχνική της μαγνητικής τομογραφίας. Η βασική αντίληψη είναι πως χρωματικός τόνος και υφή έχουν μία πολύπλοκη και άρρηκτη σύνδεση μεταξύ τους. Όταν μια περιοχή της εικόνας παρουσιάζει χαρακτηριστικά με μικρή διακύμανση χρωματικού τόνου, η ιδιότητα που περιγράφει την εικόνα είναι ο τόνος[14]. Αντίθετα, όταν μια περιοχή παρουσιάζει μια ευρεία διακύμανση χαρακτηριστικών με διαφορετικό χρωματικό τόνο, αναδύεται ως κυρίαρχη ιδιότητα η υφή. Τα χαρακτηριστικά υφής περιγράφουν κάθε pixel με βάση τα πλησιέστερα του και κατά αυτό τον τρόπο περικλείουν περισσότερη πληροφορία από τις υπόλοιπες οικογένειες χαρακτηριστικών με αντίβαρο μεγαλύτερο υπολογιστικό κόστος.

Local Binary Patterns (LBP)

Το LBP είναι χαρακτηριστικό υφής ευρέως διαδεδομένο στην ανάλυση και επεξεργασία εικόνας. Αρχικά υλοποιήθηκε για την ανάλυση δισδιάστατων εικόνων

αλλά η εφαρμογή του επεκτάθηκε και στον τρισδιάστατο χώρο [15]. Για κάθε προς εξέταση pixel της εικόνας δημιουργείται ένας πίνακας διαστάσεων 3×3 , με τις θέσεις του πίνακα να αντιστοιχούν στα γειτνιάζοντα του κεντρικού pixel. Σε κάθε κελί του παραπάνω πίνακα αποθηκεύεται η τιμή 0 σε περίπτωση που η φωτεινότητα του αντίστοιχου pixel είναι μικρότερη από την τιμή του κεντρικού, που αποτελεί και το κατώφλι των συγχρίσεων ή η τιμή 1 αν είναι μεγαλύτερη από την τιμή του κεντρικού. Στη συνέχεια, παραθέτονται με συγχεκριμένη σειρά τις περιεχόμενες τιμές του πίνακα δημιουργείται ένας δυαδικός αριθμός, ο οποίος στην συνέχεια μεταφράζεται στο δεκαδικό σύστημα. Ο δεκαδικός αριθμός αποτελεί και τον τελικό LBP περιγραφέα του κεντρικού pixel.



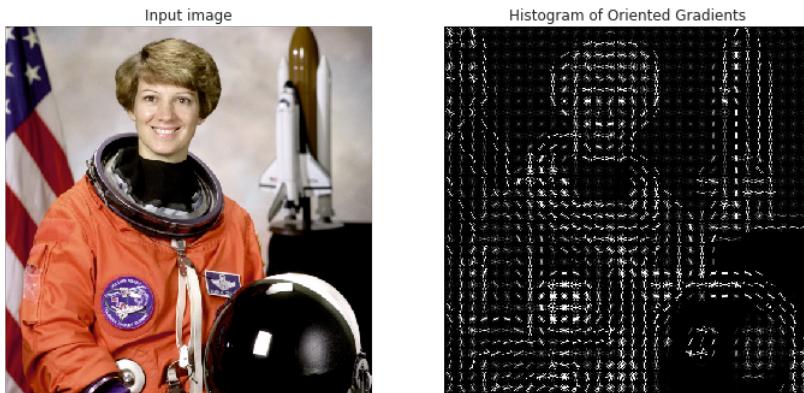
Σχήμα 5.1: Εξαγωγή χαρακτηριστικού LBP για ένα pixel σε γειτονιά ακτίνας 1.

Αντίστοιχα, για τον τρισδιάστατο χώρο δημιουργείται ένας πίνακας διαστάσεων $3 \times 3 \times 3$ τα κελιά του οποίου ταυτίζονται με τα voxels του κύβου $3 \times 3 \times 3$ που περιέχουν στο κέντρο τους το voxel προς εξέταση. Κατ' αντιστοιχία, ο πίνακας συμπληρώνεται με τα αποτελέσματα των συγχρίσεων των φωτεινοτήτων και παράγει τον τελικό περιγραφέα. Επέκταση των υλοποιήσεων αποτελεί και η αύξηση των διαστάσεων του τετραγώνου ή του κύβου που περιέχει το προς εξέταση voxel για την πιο λεπτομερή περιγραφή της ευρύτερης γειτονιάς.

Histogram of Oriented Gradients (HOG)

Τα χαρακτηριστικά HOG [16] είναι περιγραφές που χρησιμοποιούνται κατά βάση στην αναγνώριση εικόνας, όπως στην αναγνώριση αντικειμένων και την αναγνώριση προσώπου. Βασίζονται στην ιδέα ότι η εμφάνιση ή το σχήμα ενός τοπικού αντικειμένου ή όγκου μπορεί να προσδιοριστεί επαρκώς από την κατανομή των τοπικών κλίσεων των φωτεινοτήτων ή των κατευθύνσεων των ακμών

του αντικειμένου. Η διαδικασία που ακολουθεί είναι να υπολογιστεί το μέτρο και η κατεύθυνση της κλίσης της φωτεινότητας για κάθε pixel της εικόνας. Ακολούθως, τα μέτρα και οι κλίσεις χρησιμοποιούνται για τη δημιουργία ενός ιστογράμματος κατευθυνόμενων κλίσεων για κάθε pixel. Στη συνέχεια αυτά τα ιστογράμματα συγχωνεύονται για να σχηματίσουν ένα μοναδικό διάνυσμα που περιγράφει την συνολική δομή μεταβόλων φωτεινοτήτων της εικόνας. Τα ιστογράμματα κατευθυνόμενων κλίσεων έχουν αποδειχθεί αρκετά αποτελεσματικοί περιγραφείς για μεγάλο εύρος υπολογιστικών προβλημάτων, ειδικότερα όταν συνδυάζονται με αλγορίθμους μηχανικής μάθησης όπως τα SVMs ή τα CNNs.



Σχήμα 5.2: Περιγραφή εικόνας μέσω του ιστογράμματος κατευθυνόμενων κλίσεων

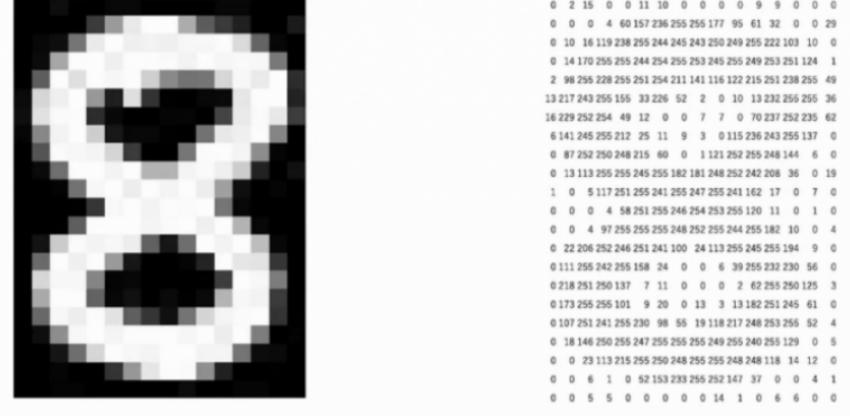
Χαρακτηριστικά Αναγνώρισης Ακμών

Μία άλλη ευρέως χρησιμοποιούμενη ομάδα χαρακτηριστικών είναι η αναγνώριση των ακμών μιας εικόνας, η οποία βρίσκει εφαρμογή και σε πολλά προβλήματα κατάτμησης ιατρικών εικόνων. Η λογική σε αυτή την κατηγορία χαρακτηριστικών έγκειται στο ότι το περιεχόμενο μιας εικόνας είναι δυνατόν να αναγνωριστεί με την ίδια ευκολία, αν η εικόνα περιγράφεται μόνο από τις ακμές των αντικειμένων που την απαρτίζουν. Με τον τρόπο αυτό, απομακρύνουμε την περιττή πληροφορία και το ύδρυβο από την εικόνα απλοποιώντας την περιγραφή της.



Σχήμα 5.3: Περιγραφή εικόνας μέσω των ακμών της

Οι ακμές μιας εικόνας ουσιαστικά περιγράφονται μέσω των σημαντικών και απότομων διαφορών στις τιμές φωτεινότητας. Όπως φαίνεται στο Σχήμα (5.3), το αντικείμενο αποτελείται από pixel υψηλών φωτεινοτήτων και ο χώρος που το περιβάλλει από pixel χαμηλών φωτεινοτήτων. Τα pixels με μεγάλες τιμές φωτεινότητας που συνορεύουν με pixels πολύ χαμηλότερης φωτεινότητας αποτελούν τις ακμές, τα ορια δηλαδή του αντικειμένου.



Σχήμα 5.4: Εύρεση ακμών μιας εικόνας. Τα pixels που παρουσιάζουν έντονες διαφορές φωτεινότητας με τα γειτονικά τους αποτελούν τις ακμές του σχήματος της εικόνας.

Gray Level Co-occurrence Matrix (GLCM)

Τα χαρακτηριστικά GLCM υπολογίζονται στηριζόμενα τόσο στην κατανομή των τιμών φωτεινότητας όσο και στην σχετική θέση των των pixels στον δισδιάστατο χώρο [17]. Ο GLCM είναι ένας πίνακας $L \times L$, όπου L τα διακριτά επίπεδα φωτεινότητας της εικόνας. Το $(i, j)^o$ στοιχείο του πίνακα αναπαριστά τον αριθμό των φορών που ο συνδυασμός των επιπέδων φωτεινότητας i και j απαντώνται σε δύο pixels/voxels στην εικόνα και χωρίζονται από μια απόσταση δ και υπό γωνία θ . Έστω ότι, το $G_p(i, j, \delta, \theta)$ αναπαριστά τον δισδιάστατο πίνακα GLCM του pixel p με απόσταση δ και γωνία θ από το κεντρικό pixel (i, j) είναι ο αριθμός της γραμμής και της στήλης του πίνακα. Έστω επίσης, P^2 το σύνολο των ζευγών pixel της υπό εξέταση 2D εικόνας, το GLCM μπορεί να περιγραφεί με την παρακάτω εξίσωση:

$$G_p(i, j, \delta, \theta) = \# \left\{ (p_1, p_2) \in P^2 \middle| \begin{array}{l} I(p_1) = i, I(p_2) = j \\ Che_dis(p_1, p_2) = \delta \\ \Theta(p_1, p_2) = \theta \end{array} \right\} \quad (5.17)$$

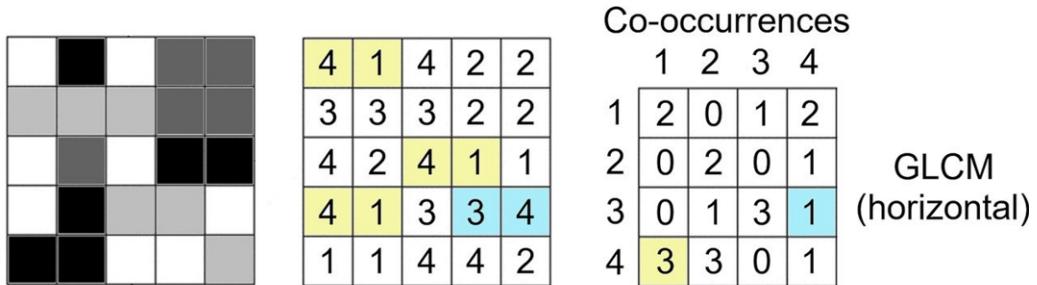
όπου $I(p)$ η φωτεινότητα του pixel p , $Che_dis(p_1, p_2)$ και $\Theta(p_1, p_2)$ είναι η απόσταση Chebyshev [18] και η επίπεδη κατεύθυνση αντίστοιχα και $\#$ ο τε-

λεστής που υπολογίζει τον αριθμό εμφανίσεων. Η απόσταση *Chebyshev* είναι μία μετρική στο διανυσματικό χώρο όπου η απόσταση μεταξύ δύο διανυσμάτων ισούται με τη μέγιστη διαφορά τους προς οποιαδήποτε κατεύθυνση. Τα ζεύγη pixel με συγκεκριμένη απόσταση και γωνία μπορούν να αναπαρασταθούν από ένα διάνυσμα μετατόπισης.

Κατεύθυνση(θ)	Διάνυσμα Μετατόπισης
0°	$(\delta, 0)$
45°	(δ, δ)
90°	$(0, \delta)$
135°	$(-\delta, \delta)$

Πίνακας 1: Διανύσματα μετατόπισης σε διάφορες κατευθύνσεις στον δισδιάστατο χώρο

Για παράδειγμα, όπως βλέπουμε στο Σχήμα (5.5), για ένα παράθυρο τεσσάρων διακριτών φωτεινοτήτων παράγεται ο πίνακας GLCM για απόσταση $\delta = 1$ και $\theta = 0$.



Σχήμα 5.5: Εξαγωγή πίνακα GLCM για 4 στάδιμες φωτεινοτήτων

Η επέκταση στον υπολογισμό GLCM στις τρεις διαστάσεις ακολουθεί την ίδια διαδικασία με τη διαφορά την αύξηση των κατευθύνσεων από 4 σε 13 όπως απεικονίζονται στον πίνακα 2.

Κατεύθυνση(ϑ, φ)	Διάνυσμα Μετατόπισης
$0^\circ, 90^\circ$	$(\delta, 0, 0)$
$45^\circ, 90^\circ$	$(\delta, \delta, 0)$
$90^\circ, 90^\circ$	$(0, \delta, 0)$
$135^\circ, 90^\circ$	$(-\delta, \delta, 0)$
$\sim, 0^\circ$	$(0, 0, \delta)$
$0^\circ, 45^\circ$	$(\delta, 0, \delta)$
$45^\circ, 90^\circ - \arcsin \sqrt{\frac{1}{3}}$	(δ, δ, δ)
$90^\circ, 45^\circ$	$(0, \delta, \delta)$
$135^\circ, 90^\circ - \arcsin \sqrt{\frac{1}{3}}$	$(-\delta, \delta, \delta)$
$0^\circ, 135^\circ$	$(\delta, 0, -\delta)$
$45^\circ, 90^\circ + \arcsin \sqrt{\frac{1}{3}}$	$(\delta, \delta, -\delta)$
$90^\circ, 135^\circ$	$(0, \delta, -\delta)$
$135^\circ, 90^\circ + \arcsin \sqrt{\frac{1}{3}}$	$(-\delta, \delta, -\delta)$

Πίνακας 2: Διανύσματα μετατόπισης σε διάφορες κατευθύνσεις στον τρισδιάστατο χώρο

Θεωρούμε:

- ϵ μια αυθαίρετη πολύ μικρή σταθερά ($\approx 22 \times 10^{-16}$)
- $\mathbf{P}(i, j)$ ο πίνακας *co – occurrence* για ορισμένο δ και ϑ .
- $p(i, j)$ ο κανονικοποιήμενος πίνακας *co – occurrence* και ίσος με $\frac{\mathbf{P}(i, j)}{\sum \mathbf{P}(i, j)}$
- \mathbf{X} ένα σετ που αποτελείται από N_p voxels που συμπεριλαμβάνονται στο πεδίο ενδιαφέροντος.
- $\mathbf{P}(i)$ το ιστόγραμμα πρώτης τάξης με N_g διαχριτές τιμές φωτεινότητας, όπου N_g είναι ο αριθμός των μη μηδενικών *bins*.
- $p(i)$ το κανονικοποιημένο ιστόγραμμα πρώτης τάξης, ίσο με $\frac{\mathbf{P}(i)}{N_p}$
- N_g ο αριθμός των διαχριτών επιπέδων φωτεινότητας στην εικόνα.

-
- $p_x(i) = \sum_{j=1}^{N_g} p(i, j)$ οι οριακές πιθανότητες ανά γραμμή
 - $p_y(j) = \sum_{i=1}^{N_g} p(i, j)$ οι οριακές πιθανότητες ανά στήλη
 - μ_x η μέση τιμή φωτεινότητας της p_x που ορίζεται ως $\mu_x = \sum_{i=1}^{N_g} p_x(i)i$
 - μ_y η μέση τιμή φωτεινότητας της p_y που ορίζεται ως $\mu_y = \sum_{i=1}^{N_g} p_y(j)j$
 - σ_x η τυπική απόκλιση της p_x
 - σ_y η τυπική απόκλιση της p_y
 - $p_{x+y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j)$, όπου $i + j = k$, και $k = 2, 3, \dots, 2N_g$
 - $p_{x-y}(k) = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j)$, όπου $|i - j| = k$, και $k = 0, 1, \dots, N_g - 1$
 - $HX = -\sum_{i=1}^{N_g} p_x(i) \log_2(p_x(i) + \epsilon)$ η εντροπία της p_x
 - $HY = -\sum_{j=1}^{N_g} p_y(j) \log_2(p_y(j) + \epsilon)$ η εντροπία της p_y
 - $HXY = -\sum_{i=1}^{N_g} p(i, j) \log_2(p(i, j) + \epsilon)$ η εντροπία της $p(i, j)$
 - $HXY1 = -\sum_{i=1}^{N_g} p(i, j) \log_2(p_x(i)p_y(j) + \epsilon)$ η εντροπία της $p(i, j)$
 - $HXY2 = -\sum_{i=1}^{N_g} p_x(i)p_y(j) \log_2(p_x(i)p_y(j) + \epsilon)$ η εντροπία της $p(i, j)$

1. Autocorrelation

$$\text{autocorrelation} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j)ij \quad (5.18)$$

H Autocorrelation είναι μία μετρική του μεγέθους της απαλότητας και της τραχύτητας της υφής.

2. Joint Average

$$\text{joint average} = \mu_x = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j)i \quad (5.19)$$

To JointAverage είναι μέση τιμή φωτεινότητας της i -οστής κατανομής.

3. Cluster Prominence

$$\text{cluster prominence} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i + j - \mu_x - \mu_y)^4 p(i, j) \quad (5.20)$$

To *ClusterProminence* είναι ένα μέτρο της λοξότητας και της ασυμμετρίας του (*GLCM*). Υψηλότερες τιμές καταδεικνύουν μεγαλύτερη ασυμμετρία γύρω από την μέση τιμή ενώ χαμηλότερες τιμές καταδεικνύουν ένα μέγιστο κοντά στην μέση τιμή και μικρότερη διακύμανση γύρω από αυτή.

4. Cluster Shade

$$\text{cluster shade} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i + j - \mu_x - \mu_y)^3 p(i, j) \quad (5.21)$$

To *ClusterShade* είναι ένα μέτρο της λοξότητας και της ομοιομορφίας του (*GLCM*). Υψηλότερες τιμές καταδεικνύουν μεγαλύτερη ασυμμετρία γύρω από την μέση τιμή.

5. Cluster Tendency

$$\text{cluster tendency} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i + j - \mu_x - \mu_y)^2 p(i, j) \quad (5.22)$$

To *ClusterTendency* είναι ένα μέτρο της ομαδοποίησης voxels με παρόμοιες τιμές φωτεινότητας.

6. Contrast

$$\text{contrast} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - j)^2 p(i, j) \quad (5.23)$$

To *Contrast* είναι ένα μέτρο της τοπικής διακύμανσης της μέσης τιμής, ευνοώντας τιμές μακριά από την διαγώνιο $i = j$. Υψηλότερη τιμή καταδεικνύει μεγαλύτερη ανισότητα ανάμεσα σε γειτνιάζοντα voxels.

7. Correlation

$$\text{correlation} = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j)ij - \mu_x\mu_y}{\sigma_x(i)\sigma_y(j)} \quad (5.24)$$

H *Correlation* είναι μία μετρική με τιμή ανάμεσα στο 0 (ασυσχέτιστα) και το 1 (τέλεια συσχέτιση) που δείχνει την γραμμική εξάρτηση των τιμών φωτεινότητας των αντίστοιχων voxels στον GLCM.

8. Difference Average

$$\text{difference average} = \sum_{k=0}^{N_g-1} kp_{x-y}(k) \quad (5.25)$$

H *DifferenceAverage* μετράει την σχέση μεταξύ εμφανίσεων ζευγών με παρόμοιες τιμές φωτεινότητας και εμφανίσεων ζευγών με άνισων τιμών.

9. Difference Entropy

$$\text{difference entropy} = \sum_{k=0}^{N_g-1} p_{x-y}(k) \log_2(p_{x-y}(k) + \epsilon) \quad (5.26)$$

H *DifferenceEntropy* είναι μία μετρική της τυχαιότητας στις διαφορές τιμών φωτεινότητας σε μία γειτονιά voxels.

10. Difference Variance

$$\text{difference variance} = \sum_{k=0}^{N_g-1} (k - DA)^2 p_{x-y}(k) \quad (5.27)$$

H *DifferenceVariance* είναι μία μετρική της ετερογένειας που εναποθέτει μεγαλύτερα βάρη στα ζεύγη διαφορετικών τιμών φωτεινότητας που απέχουν από την μέση.

11. Joint Energy

$$\text{joint energy} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (p(i,j))^2 \quad (5.28)$$

H *JointEnergy* είναι μία μετρική ομοιγενών μοτίβων στην εικόνα. Υψηλότερη τιμή δείχνει πως υπάρχουν περισσότερες περιπτώσεις ζευγών με όμοια φωτεινότητα που γειτνιάζουν με μεγαλύτερη συχνότητα.

12. Joint Entropy

$$\text{joint entropy} = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) \log_2(p(i, j) + \epsilon) \quad (5.29)$$

H *JointEntropy* είναι μία μετρική της τυχαιότητας στις τιμές φωτεινότητας σε μία γειτονιά voxels.

13. Informational Measure of Correlation(IMC) 1

$$IMC\ 1 = \frac{HXY - HXY1}{\max\{HX, HY\}} \quad (5.30)$$

14. Informational Measure of Correlation(IMC) 2

$$IMC\ 2 = \sqrt{1 - e^{-2(HXY2 - HXY)}} \quad (5.31)$$

15. Inverse Difference Moment (IDM)

$$IDM = \sum_{k=0}^{N_g-1} \frac{p_{x-y}(k)}{1+k^2} \quad (5.32)$$

H *IDM* είναι μία μετρική της τοπικής ομοιογένειας σε μία εικόνα. Τα βάρη της *IDM* είναι αντίστροφα των βαρών της *Contrast* (μειώνονται εκθετικά της απόστασης από την διαγώνιο $i = j$ του GLCM).

16. Maximal Correlation Coefficient (MCC)

$$MCC = \sqrt{\text{second largest eigenvalue of } Q}$$

H *MCC* είναι μία μετρική της περιπλοκότητας της υφής.

$$Q(i, j) = \sum_{k=0}^{N_g} \frac{p(i, k)p(j, k)}{p_x(i)p_y(k)} \quad (5.33)$$

17. Inverse Difference Moment Normalized (IDMN)

$$IDMN = \sum_{k=0}^{N_g-1} \frac{p_{x-y}(k)}{1 + \left(\frac{k^2}{N_g^2}\right)} \quad (5.34)$$

H *IDMN* είναι μία μετρική της τοπικής ομοιογένειας της εικόνας. Τα βάρη της *IDMN* είναι αντίστροφα των βαρών της *Contrast* (μειώνονται εκθετικά της απόστασης από την διαγώνιο $i = j$ του GLCM). Παράλληλα, η *IDMN* κανονικοποιεί το τετράγωνο της διαφοράς τιμών φωτεινότητας μεταξύ γειτνιάζοντων voxels διαιρόντας με το τετράγωνο του αριθμού των διαχριτών τιμών φωτεινότητας.

18. Inverse Difference (ID)

$$ID = \sum_{k=0}^{N_g-1} \frac{p_{x-y}(k)}{1 + k} \quad (5.35)$$

H *ID* είναι άλλη μία μετρική της τοπικής ομοιογένειας της εικόνας. Με περισσότερα όμοια επίπεδα φωτεινότητας, ο παρανομαστής θα παραμείνει μικρός με αποτέλεσμα υψηλότερη συνολική τιμή για την μετρική.

19. Inverse Difference Normalized (IDN)

$$IDN = \sum_{k=0}^{N_g-1} \frac{p_{x-y}(k)}{1 + \left(\frac{k}{N_g}\right)} \quad (5.36)$$

H *IDN* είναι μία μετρική της τοπικής ομοιογένειας της εικόνας. Σε αντίθεση με την *ID*, η *IDN* κανονικοποιεί την διαφορά μεταξύ γειτνιάζοντων τιμών φωτεινότητας, διαιρώντας με τον αριθμό των διαχριτών επιπέδων φωτεινότητας.

20. Inverse Variance (IV)

$$inverse\ variance = \sum_{k=0}^{N_g-1} \frac{p_{x-y}(k)}{k^2} \quad (5.37)$$

H *InverseVariance* είναι μία μετρική αντίστροφη της μεταβλητότητας.

21. Maximum Probability

$$\text{maximum probability} = \max(p(i, j)) \quad (5.38)$$

Η *MaximumProbability* μετράει των αριθμό των περιπτώσεων των ζευγών γειτνιάζοντων φωτεινοτήτων που εμφανίζονται με την μεγαλύτερη συχνότητα.

22. Sum Average

$$\text{sum average} = \sum_{k=2}^{2N_g} 2N_g p_{x+y}(k) k \quad (5.39)$$

Το *SumAverage* είναι ένα μέτρο της σχέσης μεταξύ των περιπτώσεων εμφάνισης ζευγών με χαμηλότερες τιμές φωτεινότητας και των περιπτώσεων εμφάνισης ζευγών με υψηλότερες τιμές φωτεινότητας.

23. Sum Entropy

$$\text{sum entropy} = \sum_{k=2}^{2N_g} p_{x+y}(k) \log_2(p_{x+y}(k) + \epsilon) \quad (5.40)$$

Το *SumEntropy* είναι το άθροισμα των διαφορών τιμών φωτεινότητας σε μία γειτονιά.

24. Sum of Squares

$$\text{sum squares} = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i - \mu_x)^2 p(i, j) \quad (5.41)$$

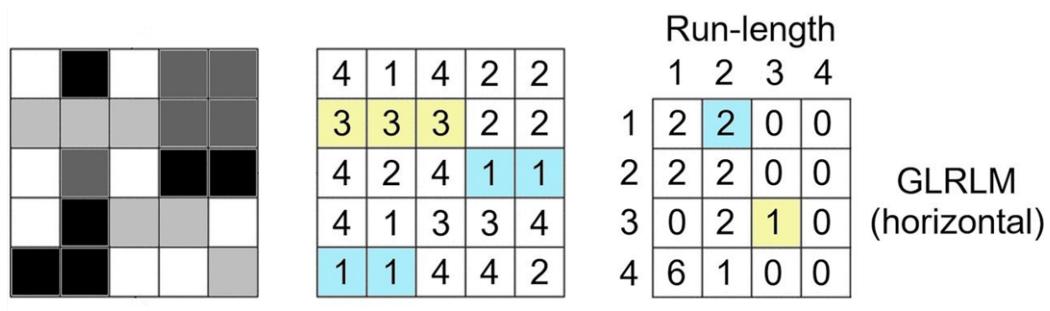
Το *SumofSquares* είναι ένα μέτρο της κατανομής γειτνιάζοντων ζευγών τιμών φωτεινότητας γύρω από την μέση τιμή στον GLCM.

Gray Level Run Length Matrix (GLRLM)

Ένας πίνακας P *glrlm* ορίζεται ως, κάθε στοιχείο $P(i, j)$ αναπαριστά τον αριθμό της αλληλουχίας των pixel με φωτεινότητα i και μήκος αλληλουχίας ίσο με j πάνω σε μία συγκεκριμένη κατεύθυνση. Το μέγεθος του πίνακα P είναι $L \times k$, όπου L ο αριθμός των διαχριτών φωτεινοτήτων της εικόνας και k ισούται

με το μέγιστο δυνατό μήκος αλληλουχίας ίσων φωτεινοτήτων στην εικόνα[19]. Η κατεύθυνση ορίζεται από ένα διάνυσμα μετατόπισης όπως και στον πίνακα GLCM. Οι πιθανές μετατοπίσεις απεικονίζονται αντίστοιχα στον πίνακα 1 για τον δισδιάστατο χώρο.

Για παράδειγμα, όπως βλέπουμε στο Σχήμα (5.6), για ένα παράθυρο τεσσάρων διακριτών φωτεινοτήτων παράγεται ο πίνακας $glrlm$ για απόσταση $d = 1$ και $\theta = 0$. Ο οριζόντιος άξονας αναπαριστά το μήκος των αλληλουχίων και ο κάθετος τις διακριτές τιμές φωτεινότητας από την χαμηλότερη στην υψηλότερη.



Σχήμα 5.6: Εξαγωγή πίνακα $GLRLM$ για 4 στάθμες φωτεινότητας

Ο υπολογισμός των $glrlm$ για την υφή τρισδιάστατων όγκων, ακολουθεί τη λογική επέκτασης του GLCM με την αύξηση των πιθανών κατευθύνσεων όπως απεικονίζονται στον πίνακα 2. Από τον πίνακα $glrlm$ εξάγονται τα παρακάτω χαρακτηριστικά. Θεωρούμε:

- N_g τον αριθμό των διακριτών φωτεινοτήτων στην εικόνα.
- N_r τον αριθμό των διακριτών *run lengths* στην εικόνα.
- N_p τον αριθμό των voxels στην εικόνα.
- $N_r(\theta)$ τον αριθμό των *runs* στην εικόνα κατά μήκος της γωνίας θ , που ισούται με $\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \mathbf{P}(i, j | \theta)$ και ισχύει $1 \leq N_r(\theta) \leq N_p$.

-
- $\mathbf{P}(i, j|\theta)$ ο πίνακας των *run lengths* για μια κατεύθυνση θ .
 - $\mathbf{p}(i, j|\theta)$ ο κανονικοποιημένος πίνακας των *run lengths* ορισμένος ως $\mathbf{p}(i, j|\theta) = \frac{\mathbf{P}(i, j|\theta)}{N_r(\theta)}$

Τα χαρακτηριστικά που προκύπτουν από τον πίνακα P περιγράφονται παρακάτω.

1. Short Run Emphasis(SRE)

$$SRE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{\mathbf{P}(i, j|\theta)}{j^2}}{N_r(\theta)} \quad (5.42)$$

Η *SRE* είναι ένα μέτρο της κατανομής των *short run lengths*. Υψηλότερες τιμές καταδεικνύουν κοντύτερα *run lengths* και πιο λεπτές υφές.

2. Long Run Emphasis(LRE)

$$LRE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \mathbf{P}(i, j|\theta) j^2}{N_r(\theta)} \quad (5.43)$$

Η *SRE* είναι ένα μέτρο της κατανομής των *long run lengths*. Υψηλότερες τιμές καταδεικνύουν μακρύτερα *run lengths* και πιο τραχείες υφές.

3. Gray Level Non-Uniformity(GLN)

$$GLN = \frac{\sum_{i=1}^{N_g} \left(\sum_{j=1}^{N_r} \mathbf{P}(i, j|\theta) \right)^2}{N_r(\theta)} \quad (5.44)$$

Η *GLN* είναι ένα μέτρο της ομοιότητας των τιμών φωτεινοτήτων της εικόνας, όπου μία μικρή τιμή σχετίζεται με μεγαλύτερη ομοιότητα των τιμών φωτεινοτήτων.

4. Run Length Non-Uniformity(RLN)

$$RLN = \frac{\sum_{i=1}^{N_r} \left(\sum_{j=1}^{N_g} \mathbf{P}(i, j|\theta) \right)^2}{N_r(\theta)} \quad (5.45)$$

Η *GLNN* μετράει την ομοιότητα των *run lengths* σε μία εικόνα, με χαμηλότερες τιμές να καταδεικνύουν μεγαλύτερη ομοιογένεια μεταξύ των *run lengths*.

στην εικόνα.

5. Run Percentage(RP)

$$RP = \frac{N_r(\theta)}{N_p} \quad (5.46)$$

To RP μετράει την τραχύτητα της υφής υπολογίζοντας τον λόγο του αριθμού των *runs* και του αριθμού των voxels της εικόνας σε ένα πεδίο ενδιαφέροντος της εικόνας.

Οι τιμές του βρίσκονται στο εύρος $\frac{1}{N_p} \leq RP \leq 1$, με υψηλές τιμές να καταδεικνύουν ότι ένα μεγαλύτερο μέρος του πεδίου ενδιαφέροντος αποτελέται από *short runs*(πιο λεπτή υφή).

6. Gray Level Variance(GLV)

$$GLV = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i, j | \theta) (i - \mu)^2 \quad (5.47)$$

όπου $\mu = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i, j | \theta) i$.

To GLV μετράει την διακύμανση των τιμών φωτεινοτήτων για τα *runs*.

7. Run Variance(RV)

$$RV = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i, j | \theta) (j - \mu)^2 \quad (5.48)$$

όπου $\mu = \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i, j | \theta) j$.

To RV μετράει την διακύμανση των *runs* για τα *run lengths*.

8. Run Entropy(RE)

$$RE = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_r} p(i, j | \theta) \log_2 (p(i, j | \theta) + \epsilon) \quad (5.49)$$

όπου ϵ ένας πολύ μικρός αριθμός ($\approx 2.2 \times 10^{-16}$).

To RE μετράει την τυχαιότητα στην κατανομή των *run lengths* και των επιπέδων φωτεινότητας. Υψηλότερες τιμές καταδυκνύουν περισσότερη ετερογένεια

στα μοτίβα υφής.

9. Low Gray Level Run Emphasis(LGLRE)

$$LGLRE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{\mathbf{P}(i,j|\theta)}{i^2}}{N_r(\theta)} \quad (5.50)$$

Η *LGLRE* είναι ένα μέτρο της κατανομής των χαμηλών επιπέδων φωτεινότητας. Υψηλότερες τιμές καταδυκνύουν μεγαλύτερη συγκέντρωση των χαμηλών επιπέδων φωτεινότητας στην εικόνα.

10. High Gray Level Run Emphasis(HGLRE)

$$HGLRE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \mathbf{P}(i,j|\theta) i^2}{N_r(\theta)} \quad (5.51)$$

Η *HGLRE* είναι ένα μέτρο της κατανομής των υψηλών επιπέδων φωτεινότητας. Υψηλότερες τιμές καταδυκνύουν μεγαλύτερη συγκέντρωση των υψηλών επιπέδων φωτεινότητας στην εικόνα.

11. Short Run Low Gray Level Emphasis(SRLGLE)

$$SRLGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{\mathbf{P}(i,j|\theta)}{i^2 j^2}}{N_r(\theta)} \quad (5.52)$$

Η *SRLGLE* είναι ένα μέτρο της από κοινού κατανομής των κοντύτερων *run lengths* με χαμηλές τιμές φωτεινοτήτων στην εικόνα.

12. Short Run High Gray Level Emphasis(SRHGLE)

$$SRHGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{\mathbf{P}(i,j|\theta) i^2}{j^2}}{N_r(\theta)} \quad (5.53)$$

Η *SRHGLE* είναι ένα μέτρο της από κοινού κατανομής των κοντύτερων *run lengths* με υψηλές τιμές φωτεινοτήτων στην εικόνα.

13. Long Run Low Gray Level Emphasis(LRHGLE)

$$LRHGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{\mathbf{P}(i,j|\theta) j^2}{i^2}}{N_r(\theta)} \quad (5.54)$$

Η *LRHGLE* είναι ένα μέτρο της από κοινού κατανομής των μακρύτερων *run lengths* με χαμηλές τιμές φωτεινοτήτων στην εικόνα.

14. Long Run High Gray Level Emphasis(LRLGLE)

$$LRLGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \mathbf{P}(i, j | \theta) i^2 j^2}{N_r(\theta)} \quad (5.55)$$

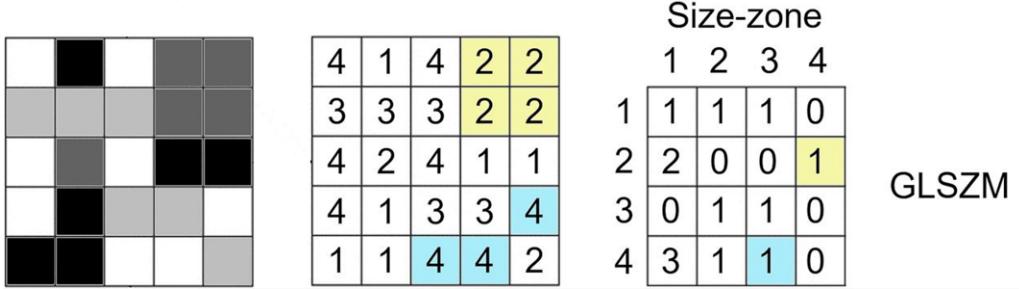
Η *LRLGLE* είναι ένα μέτρο της από κοινού κατανομής των μακρύτερων *run lengths* με υψηλές τιμές φωτεινοτήτων στην εικόνα.

Gray Level Size Zone Matrix (GLSZM)

Ο Gray Level Size Zone Matrix ποσοτικοποιεί τις ζώνες φωτειότητας στην εικόνα. Ως ζώνη φωτεινότητας ορίζεται ένας αριθμός διαδοχικών voxels που μοιράζονται την ίδια τιμής φωτεινότητας. Ένα voxel θεωρείται διαδοχικό ενός δεύτερου αν η απόσταση d μεταξύ τους είναι ίση με 1 προς οποιαδήποτε κατεύθυνση. Έτσι έχουμε 8 διαδοχικά voxels γύρω από ένα κεντρικό σε μία εικόνα δύο διαστάσεων και 26 διαδοχικά voxels γύρω από ένα κεντρικό σε μία εικόνα τριών διαστάσεων. Ο *glszm* αποτελεί ουσιαστικά την επέκταση του *glrlm* από την μέτρηση αλληλουχιών ίδιων φωτεινοτήτων κατά μία κατεύθυνση, υπολογίζοντας δηλαδή μήκος, στην μέτρηση κατά οποιαδήποτε κατεύθυνση, υπολογίζοντας εμβαδό (ή όγκο σε εικόνες τριών διαστάσεων). Για παράδειγμα, όπως βλέπουμε στο Σχήμα (5.7), για ένα παράθυρο τεσσάρων διακριτών φωτεινοτήτων παράγεται ο πίνακας *glszm*. Ο οριζόντιος άξονας αναπαριστά το εμβαδό των αλληλουχιών και ο κάθετος τις διακριτές τιμές φωτεινότητας από την χαμηλότερη στην υψηλότερη.

Θεωρούμε:

- N_g τον αριθμό των διακριτών φωτεινοτήτων στην εικόνα.
- N_s τον αριθμό των διακριτών μεγεθών ζωνών φωτεινοτήτων στην εικόνα.
- N_p τον αριθμό των voxels στην εικόνα.
- N_z τον αριθμό των ζωνών φωτεινότητας στο *ROI*, που ισούται με $\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \mathbf{P}(i, j)$ και ισχύει $1 \leq N_z \leq N_p$.
- $\mathbf{P}(i, j)$ ο πίνακας των μεγεθών των ζωνών φωτεινότητας.



Σχήμα 5.7: Εξαγωγή πίνακα $GLSZM$ για 4 στάθμες φωτεινότητας

- $\mathbf{P}(i, j)$ ο κανονικοποιημένος πίνακας των μεγεθών των ζωνών φωτεινότητας ορισμένος ως $\mathbf{P}(i, j) = \frac{\mathbf{P}_{(i,j)}}{N_z}$

Τα χαρακτηριστικά που προκύπτουν από τον πίνακα P περιγράφονται παρακάτω.

1. Small Area Emphasis

To SAE είναι ένα μέτρο της κατανομής ζωνών μεγέθους μήκους, με τις υψηλότερες τιμές να είναι ενδεικτικές περισσότερων ζωνών μικρού μεγέθους και απολών υφών.

$$SAE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \frac{\mathbf{P}(i,j)}{j^2}}{N_z} \quad (5.56)$$

2. Large Area Emphasis(LAE)

$$LAE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \mathbf{P}(i,j)j^2}{N_z} \quad (5.57)$$

To LAE είναι ένα μέτρο της κατανομής ζωνών μεγάλου μεγέθους, με τις υψηλότερες τιμές να είναι ενδεικτικές περισσότερων ζωνών μεγάλου μεγέθους και τραχείων υφών.

3. Gray Level Non-Uniformity (GLN)

$$GLN = \frac{\sum_{i=1}^{N_g} \left(\sum_{j=1}^{N_s} \mathbf{P}(i,j) \right)^2}{N_z} \quad (5.58)$$

Το GLN είναι ένα μέτρο της μεταβλητότητας των τιμών φωτεινότητας στην εικόνα ή τον όγκο, με τις χαμηλότερες τιμές να είναι ενδεικτικές μεγαλύτερης ομοιογένειας στις τιμές φωτεινότητας.

4. Gray Level Non-Uniformity Normalized(GLNN)

$$GLNN = \frac{\sum_{i=1}^{N_g} \left(\sum_{j=1}^{N_s} \mathbf{P}(i, j) \right)^2}{N_z^2} \quad (5.59)$$

Το $GLNN$ είναι ένα μέτρο της μεταβλητότητας των τιμών φωτεινότητας στην εικόνα ή τον όγκο, με τις χαμηλότερες τιμές να είναι ενδεικτικές μεγαλύτερης ομοιότητας στις τιμές φωτεινότητας. Αποτελεί την κανονικοποιημένη μορφή του GLN .

5. Size-Zone Non-Uniformity(SZN)

$$SZN = \frac{\sum_{j=1}^{N_s} \left(\sum_{i=1}^{N_g} \mathbf{P}(i, j) \right)^2}{N_z} \quad (5.60)$$

Το SZN είναι ένα μέτρο της μεταβλητότητας του μεγέθους των όγκων στην εικόνα, με τις χαμηλότερες τιμές να είναι ενδεικτικές μεγαλύτερης ομοιογένειας στο μέγεθος των όγκων.

6. Size-Zone Non-Uniformity Normalized(SZNN)

$$SZNN = \frac{\sum_{j=1}^{N_s} \left(\sum_{i=1}^{N_g} \mathbf{P}(i, j) \right)^2}{N_z^2} \quad (5.61)$$

Το SZN είναι ένα μέτρο της μεταβλητότητας του μεγέθους των όγκων στην εικόνα, με τις χαμηλότερες τιμές να είναι ενδεικτικές μεγαλύτερης ομοιογένειας στο μέγεθος των όγκων. Αποτελεί την κανονικοποιημένη μορφή του SZN .

7. Zone Percentage(ZP)

$$ZP = \frac{N_z}{N_p} \quad (5.62)$$

Το ZP είναι ένα μέτρο της τραχύτητας της υφής, όντας ο λόγος του αριθμού των ζωνών προς τον αριθμό των voxels στην περιοχή ενδιαφέροντος.

8. Gray Level Variance(GLV)

$$GLV = \sum_{i=1}^{N_g} \sum_{j=1}^{N_s} p(i,j)(i - \mu)^2 \quad (5.63)$$

To GLV είναι ένα μέτρο της διακύμανσης των τιμών φωτεινότητας στις ζώνες.

9. Zone Variance(ZV)

$$ZV = \sum_{i=1}^{N_g} \sum_{j=1}^{N_s} p(i,j)(j - \mu)^2 \quad (5.64)$$

To ZV είνι ένα μέτρο της διακύμανσης των όγκων των ζωνών.

10. Zone Entropy(ZE)

$$ZE = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_s} p(i,j) \log_2(p(i,j)) + \epsilon \quad (5.65)$$

To ZE είναι ένα μέτρο της αβεβαιότητας/τυχαιότητας στην κατανομή των μεγεθών των ζωνών και των τιμών φωτεινότητας. Υψηλότερη τιμή καταδεικνύει μεγαλύτερη ετερογένεια στα μοτίβα της υφής.

11. Low Gray Level Zone Emphasis(LGLZE)

$$LGLZE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \frac{\mathbf{P}(i,j)}{i^2}}{N_z} \quad (5.66)$$

To $LGLZE$ είναι ένα μέτρο της κατανομής των μεγεθών των ζωνών χαμηλότερης φωτεινότητας, με τις υψηλότερες τιμές να καταδεικνύουν ισχυρότερη αναλογία χαμηλότερων τιμών φωτεινότητας και ζώνες υψηλότερης φωτεινότητας στην εικόνα.

12. High Gray Level Zone Emphasis(HGLZE)

$$HGLZE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \mathbf{P}(i,j)i^2}{N_z} \quad (5.67)$$

To $HGLZE$ είναι ένα μέτρο της κατανομής των μεγεθών των ζωνών υψηλότερης φωτεινότητας, με τις υψηλότερες τιμές να καταδεικνύουν ισχυρότερη

αναλογία υψηλότερων τιμών φωτεινότητας και ζώνες υψηλότερης φωτεινότητας στην εικόνα.

13. Small Area Low Gray Level Emphasis(SALGLE)

$$SALGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \frac{\mathbf{P}(i,j)}{i^2 j^2}}{N_z} \quad (5.68)$$

Το *SALGLE* είναι ένα μέτρο της αναλογίας της από κοινού κατανομής ζωνών μικρότερου μεγέθους με χαμηλότερες τιμές φωτεινότητας στην εικόνα.

14. Small Area High Gray Level Emphasis(SAHGLE)

$$SAHGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \frac{\mathbf{P}(i,j) i^2}{j^2}}{N_z} \quad (5.69)$$

Το *SAHGLE* είναι ένα μέτρο της αναλογίας της από κοινού κατανομής ζωνών μικρότερου μεγέθους με υψηλότερες τιμές φωτεινότητας στην εικόνα.

15. Large Area Low Gray Level Emphasis(LALGLE)

$$LALGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \frac{\mathbf{P}(i,j) j^2}{i^2}}{N_z} \quad (5.70)$$

Το *LALGLE* είναι ένα μέτρο της αναλογίας της από κοινού κατανομής ζωνών μεγαλύτερου μεγέθους με χαμηλότερες τιμές φωτεινότητας στην εικόνα.

16. Large Area High Gray Level Emphasis(LAHGLE)

$$LAHGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_s} \mathbf{P}(i,j) i^2 j^2}{N_z} \quad (5.71)$$

Το *LAHGLE* είναι ένα μέτρο της αναλογίας της από κοινού κατανομής ζωνών μεγαλύτερου μεγέθους με υψηλότερες τιμές φωτεινότητας στην εικόνα.

Gray Level Dependence Matrix (GLDM)

Ο *glcm* είναι ένας πίνακας με διαστάσεις $L \times L$ όπου L είναι οι διαχριτές τιμές φωτεινότητας της εικόνας. Στον *glcm*, το $(i, j)^o$ στοιχείο του πίνακα περιέχει το σύνολο των φορών που παρατηρείται το ζέυγος φωτεινότητων i, j σε

απόσταση $d = 1$ και προς μία συγκεκριμένη κατεύθυνση και την αντιδιαμετρική της[20]. Οι πίνακες *gldm* είναι συμμετρικοί και εξάγονται απευθείας από τον πίνακα GLCM.

Θεωρούμε:

- N_g τον αριθμό των διαχριτών φωτεινοτήτων στην εικόνα.
- N_d τον αριθμό των διαχριτών μεγεθών εξαρτήσεων στην εικόνα.
- N_z τον αριθμό των ζωνών εξαρτήσεων στην εικόνα, που ισούται με $\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} \mathbf{P}(i, j)$.
- $\mathbf{P}(i, j)$ ο πίνακας των εξαρτήσεων.
- $\mathbf{p}(i, j)$ ο κανονικοποιημένος πίνακας των εξαρτήσεων ορισμένος ως $\mathbf{p}(i, j|\theta) = \frac{\mathbf{P}_{(i,j)}}{N_z}$

Τα χαρακτηριστικά που προκύπτουν από τον πίνακα P περιγράφονται παρακάτω.

1. Small Dependence Emphasis (SDE)

$$SDE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} \frac{\mathbf{P}_{(i,j)}}{i^2}}{N_z} \quad (5.72)$$

Η *SDE* είναι μια μετρική της κατανομής των μικρών εξαρτήσεων με υψηλότερες τιμές να καταδεικνύουν μικρότερη εξάρτηση και λιγότερο ομοιογενείς υφές.

2. Large Dependence Emphasis (LDE)

$$SDE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} \mathbf{P}(i, j) j^2}{N_z} \quad (5.73)$$

Η *LDE* είναι μια μετρική της κατανομής των μεγάλων εξαρτήσεων με υψηλότερες τιμές να καταδεικνύουν μεγαλύτερη εξάρτηση και περισσότερο ομοιογενείς υφές.

3. Gray Level Non-Uniformity (GLN)

$$GLN = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} (\mathbf{P}(i, j))^2}{N_z} \quad (5.74)$$

Η GLN είναι μια μετρική της ομοιότητας των επιπέδων φωτεινοτήτων της εικόνας, με μια χαμηλότερες τιμές να συσχετίζονται με μεγαλύτερη ομοιότητα στα επίπεδα φωτεινοτήτων.

4. Dependence Non-Uniformity (DN)

$$DN = \frac{\sum_{i=1}^{N_d} \sum_{j=1}^{N_g} (\mathbf{P}(i, j))^2}{N_z} \quad (5.75)$$

Η DN είναι μια μετρική της εξάρτησης των τιμών φωτεινοτήτων της εικόνας, με χαμηλότερες τιμές να καταδεικνύουν μεγαλύτερη εξάρτηση στα επίπεδα φωτεινοτήτων.

5. Gray Level Variance (GLV)

$$GLV = \sum_{i=1}^{N_g} \sum_{j=1}^{N_d} p(i, j)(i - \mu)^2 \quad (5.76)$$

όπου $\mu = \sum_{i=1}^{N_g} \sum_{j=1}^{N_d} ip(i, j)$.

Η GLV μετρά την διασπορά των επιπέδων φωτεινοτητας της εικόνας.

6. Dependence Variance (DV)

$$DV = \sum_{i=1}^{N_g} \sum_{j=1}^{N_d} p(i, j)(j - \mu)^2 \quad (5.77)$$

όπου $\mu = \sum_{i=1}^{N_g} \sum_{j=1}^{N_d} jp(i, j)$.

Η DV μετρά την διασπορά του μεγέθους εξάρτησης της εικόνας.

7. Dependence Entropy (DE)

$$DE = - \sum_{i=1}^{N_g} \sum_{j=1}^{N_d} p(i, j) \log_2(p(i, j) + \epsilon) \quad (5.78)$$

8. Low Gray Level Emphasis (LGE)

$$LGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} \frac{\mathbf{P}(i,j)}{i^2}}{N_z} \quad (5.79)$$

Η *LGLE* είναι μια μετρική της κατανομής των χαμηλών φωτεινοτήτων στην εικόνα, με υψηλότερες τιμές να καταδεικνύουν υψηλή συγκέντρωση χαμηλών φωτεινοτήτων.

9. High Gray Level Emphasis (HGLE)

$$HGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} \mathbf{P}(i,j) i^2}{N_z} \quad (5.80)$$

Η *HGLE* είναι μια μετρική της κατανομής των υψηλών φωτεινοτήτων στην εικόνα, με υψηλότερες τιμές να καταδεικνύουν υψηλή συγκέντρωση υψηλών φωτεινοτήτων.

10. Small Dependence Low Gray Level Emphasis (SDLGLE)

$$SDLGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} \frac{\mathbf{P}(i,j)}{i^2 j^2}}{N_z} \quad (5.81)$$

Η *SDLGLE* είναι μια μετρική της από κοινού κατανομής των μικρών εξαρτήσεων των χαμηλών φωτεινοτήτων της εικόνας.

11. Small Dependence High Gray Level Emphasis (SDHGLE)

$$SDHGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} \frac{\mathbf{P}(i,j)}{i^2 j^2}}{N_z} \quad (5.82)$$

Η *SDHGLE* είναι μια μετρική της από κοινού κατανομής των μικρών εξαρτήσεων των χαμηλών φωτεινοτήτων της εικόνας (;;; λειπει σχεση).

12. Large Dependence Low Gray Level Emphasis (LDLGLE)

$$LDLGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} \frac{\mathbf{P}(i,j) j^2}{i^2}}{N_z} \quad (5.83)$$

Η *LDLGLE* είναι μια μετρική της από κοινού κατανομής των μεγάλων εξαρτήσεων των χαμηλών φωτεινοτήτων της εικόνας .

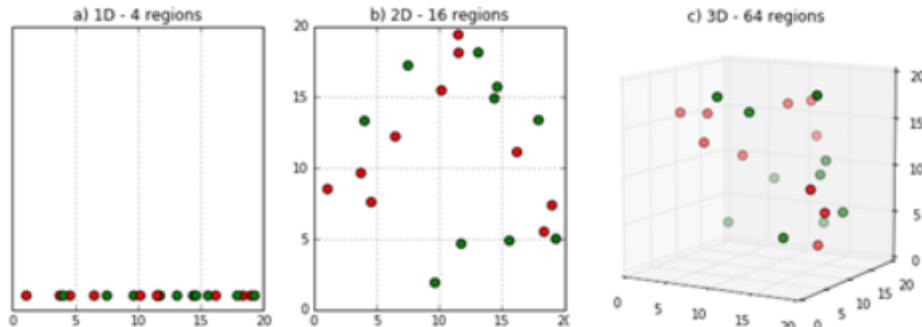
13. Large Dependence High Gray Level Emphasis (LDHGLE)

$$LDHGLE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_d} \mathbf{P}(i, j) i^2 j^2}{N_z} \quad (5.84)$$

Η *LDHGLE* είναι μια μετρική της από κοινού κατανομής των μεγάλων εξαρτήσεων των υψηλών φωτεινοτήτων της εικόνας.

5.2 Το πρόβλημα της πολυδιαστασιμότητας

Το πρόβλημα της πολυδιαστασιμότητας αναφέρεται σε μια κατάσταση όπου το μοντέλο προς εκπαίδευση έχει πολύ μεγάλο αριθμό χαρακτηριστικών, ή το σύνολο αυτών των χαρακτηριστικών των δεδομένων περιέχουν παράγοντες που ενέχουν υψηλή συσχέτιση, καθιστώντας δύσκολη την απομόνωση της χρήσημης πληροφορίας που προσφέρει το κάθε ένα από αυτά. Αυτό μπορεί να οδηγήσει στην υπερπροσαρμογή του μοντέλου στα δεδομένα εκπαίδευσης, πετυχαίνοντας σε αυτά υψηλές τιμές ακρίβειας, αλλά έχει πολύ κακή απόδοση σε νέα δεδομένα. Για να λυθεί το πρόβλημα της πολυδιαστασιμότητας, εφαρμόζονται τεχνικές μείωσης της διαστασιμότητας των δεδομένων ή τεχνικές επιλογής χαρακτηριστικών.



Σχήμα 5.8: Δεδομένα στους 1D, 2D και 3D χώρους. Όσο αυξάνονται οι διαστάσεις του χώρου, αυξάνονται οι πιθανές περιοχές των δεδομένων, καθιστώντας πολύ δυσκολότερη την κατηγοριοποίησή τους.

5.2.1 Τεχνικές Μείωσης Διαστασιμότητας

Οι τεχνικές μείωσης της διαστασιμότητας των δεδομένων αφορούν τον μετασχηματισμό των χαρακτηριστικών των δεδομένων σε έναν χώρο χαμηλότερης διάστασης που περιέχει την σημαντικότερη πληροφορία, με σκοπό την μείωση της υπολογιστικής πολυπλοκότητας του μοντέλου, την ενίσχυση της ερμηνευσιμότητάς του και την αποφυγή υπερπροσαρμογής του στα δεδομένα [21]. Κάποιες από τις γνωστότερες τεχνικές μείωσης διαστασιμότητας είναι οι εξής:

1. Principal Component Analysis (PCA)

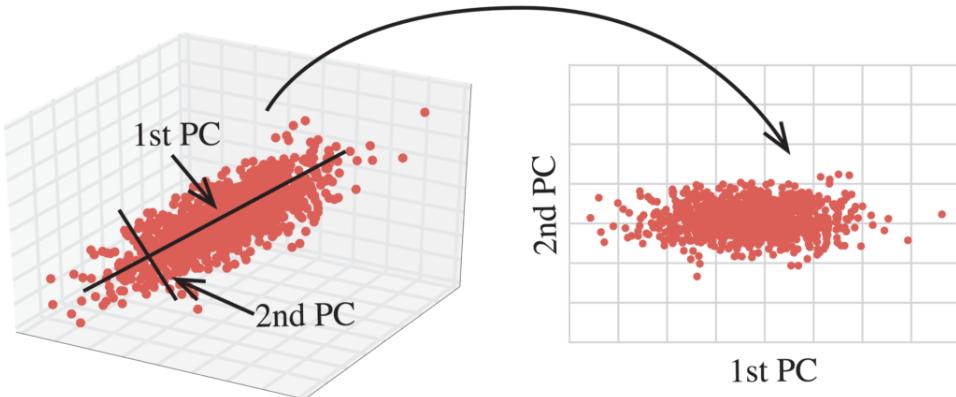
Η τεχνική PCA είναι μια ευρέως χρησιμοποιούμενη τεχνική που εφαρμόζει ένα μετασχηματισμό των αρχικών χαρακτηριστικών σε ένα νέο σετ ορθογώνιων χαρακτηριστικών, που ονομάζονται principal components και περιέχουν το μέγιστο ποσό πληροφορίας των δεδομένων.

2. Linear Discriminant Analysis (LDA)

Η τεχνική LDA είναι μια επιβλεπόμενη τεχνική μείωσης διαστασιμότητας που στοχεύει στο να βρει έναν γραμμικό συνδυασμό των χαρακτηριστικών που να μεγιστοποιεί τον διαιχωρισμό των κλάσεων[22]. Χρησιμοποιείται συνήθως σε προβλήματα που ο αριθμός των χαρακτηριστικών είναι πολύ μεγαλύτερος από τον αριθμό των δειγμάτων.

3. t-distributed Stochastic Neighbor Embedding(t-SNE)

Η t-SNE είναι μια μη γραμμική τεχνική μείωσης διαστασιμότητας που χρησιμοποιείται συνήθως για την οπτικοποίηση δεδομένων πολύ υψηλών διαστάσεων[23]. Αντιστοιχίζει τα πρωταρχικά χαρακτηριστικά σε ένα χώρο χαμηλότερης διάστασης ενώ διατηρεί την τοπική πληροφορία των δεδομένων.



Σχήμα 5.9: Μετασχηματισμός από τον 3D χώρο στον 2D χώρο με τη χρήση της τεχνικής PCA εκφράζοντας βέλτιστα με τα πρώτα 2 principal components τη μεταβλητότητα των δεδομένων.

5.2.2 Τεχνικές Επιλογής Χαρακτηριστικών

Οι τεχνικές επιλογής χαρακτηριστικών βασίζονται στη διαδικασία επιλογής των πιο σχετικών με το πρόβλημα χαρακτηριστικών από ένα ευρύτερο σύνολο χαρακτηριστικών που περιγράφουν ένα σετ δεδομένων[24]. Όπως και οι τεχνικές μείωσης διαστασιμότητας, βοηθάνε στην βελτίωση της ακρίβειας του μοντέλου, στην αποφυγή υπερπροσαρμογής του μοντέλου και στην μείωση του υπολογιστικού κόστους. Μερικές από τις πιο γνωστές τεχνικές επιλογής χαρακτηριστικών είναι οι εξής:

1. Μέθοδοι φίλτρου

Οι μέθοδοι φίλτρου αξιολογούν την σχετικότητα του κάθε χαρακτηριστικού ανεξάρτητα από το μοντέλο. Τα βαθμολογούν με τη χρήση στατιστικών μετρικών όπως συσχέτιση, κοινή πληροφορία, *chi-square* τεστ. Τα χαρακτηριστικά με τον υψηλότερο βαθμό επιλέγονται για την εκπαίδευση του μοντέλου.

2. Wrapper Μέθοδοι

Οι μέθοδοι Wrapper αξιολογούν τη σχετικότητα των χαρακτηριστικών στα πλαίσια ενός συγκεκριμένου μοντέλου. Χρησιμοποιούν την απόδοση του μοντέλου ως κριτήριο για την επιλογή τους. Αυτές οι μέθοδοι έχουν συνήθως υψηλό

υπολογιστικό κόστος καθώς απαιτούν τη συνεχή εκπαίδευση του μοντέλου με διαφορετικά υποσετ χαρακτηριστικών.

3. Recursive Feature Elimination (RFE)

Η μέθοδος RFE είναι μια μέθοδος Wrapper που αφαιρεί επαναληπτικά τα λιγότερο σημαντικά χαρακτηριστικά από το σετ δεδομένων ώσπου να καταλήξει στον επιθυμητό αριθμό χαρακτηριστικών[25]. Χρησιμοποιεί την απόδοση του μοντέλου ως κριτήριο επίλογής των χαρακτηριστικών.

6 Πειραματικά αποτελέσματα στον εγκέφαλο

Ο εγκέφαλος αποτελεί το σπουδαιότερο τμήμα του κεντρικού νευρικού συστήματος. Περιβάλλεται από το εγκεφαλικό χρανίο ενώ στην βάση του εκφύονται οι εγκεφαλικές συζυγίες ή νεύρα που αποτελούν την σύνδεση με το νωτιαίο μυελό. Ο εγκέφαλος διαμερίζεται ανατομικά σε τρεις περιοχές οι οποίες είναι τα εγκεφαλικά ημισφαίρια που χωρίζονται από την επιμήκη σχισμή, το στέλεχος και την παρεγκεφαλίδα. Ο εγκέφαλος συντίθεται από νευρώνες, οι οποίοι επιφορτίζονται με την λήψη, την επεξεργασία και την μεταβίβαση ερευνημάτων. Εντός της μάζας του εντοπίζονται εξειδικευμένες περιοχές ή κέντρα, υπεύθυνα για τις αισθήσεις, την αντίληψη, τον έλεγχο και το συντονισμό των μυικών κινήσεων και τις ανώτερες πνευματικές λειτουργίες. Τα τελευταία χρόνια, σημειώνεται ραγδαία ανάπτυξη στον τομέα της βιοιατρικής, με την μηχανική μάθηση να βρίσκεται σε όλο και αυξανόμενες εφαρμογές, μία εκ των οποίων είναι τα διαγνωστικά μοντέλα ή τα μοντέλα κατάτμησης για την παρακολούθηση της υγείας των ασθενών και την εξαγωγή συμπερασμάτων για την λειτουργία του ανθρωπίνου σώματος.

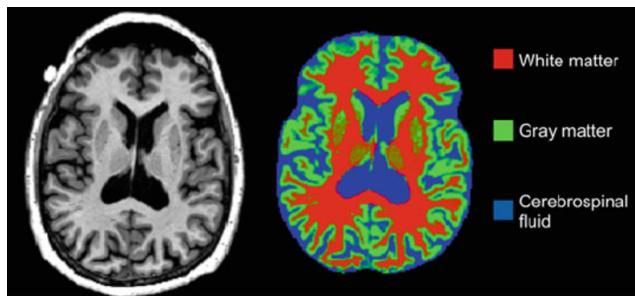
6.1 Πρόβλημα

Στο πρώτο σκέλος της παρούσας εργασίας καλούμαστε να λύσουμε το πρόβλημα της κατάτμησης ιατρικών εικόνων εγκεφάλων. Πιο συγκεκριμένα το μοντέλο ταξινόμησης εκπαιδεύεται με στόχο τον βέλτιστο διαχωρισμό των τριών βασικών ιστών που αποτελούν τον ανθρώπινο εγκέφαλο:

- **Φαιά ουσία :** Αποτελεί μέρος του κεντρικού νευρικού συστήματος και

περιέχει μεταξύ άλλων κυττάρων, τους νευρώνες, κυτταρικά σώματα και μεγάλο τμήμα νευραξόνων. Οφείλει το χαρακτηριστικό γκρι της χρώμα στην μεγάλη συγκέντρωση και τα σώματα των παραπάνω κυττάρων. Εντοπίζεται στο εξωτερικό μέρος του εγκεφάλου, σχηματίζοντας τις έλικες και τις αύλακες της επιφάνειας του. Παράλληλα, μικρότερα μέρη φαιάς ουσίας βρίσκονται στο εσωτερικό του εγκεφάλου, δημιουργώντας δομές εντός της λευκής ουσίας που ονομάζονται εγκεφαλικοί πυρήνες.

- **Λευκή ουσία :** Αποτελεί κατά βάση την συνέχεια των νευραξόνων της φαιάς ουσίας, σχηματίζοντας δέσμες νευρικών αποφυάδων. Εντοπίζεται στο εσωτερικό των εγκεφαλικών ημισφαιρίων και περικλείεται από την μάζα της φαιάς ουσίας.
- **Εγκεφαλονωτιαίο υγρό :** Είναι διαυγές, άχρωμο σωματικό υγρό που βρίσκεται μέσα στον ιστό που περιβάλλει τον εγκέφαλο και τον νωτιαίο μυελό όλων των σπονδυλωτών. Αντικαθιστά το σωματικό υγρό που βρίσκεται έξω από τα κύτταρα όλων των αμφίπλευρων ζώων. Παρατηρείται να συμπληρώνει τον χώρο στις κοιλίες, τις δεξαμενές και τον κεντρικό θόλο του εγκεφάλου.



Σχήμα 6.1: Brain-Tissue κατάτμηση στους κυριότερους ιστούς του εγκεφάλου, τη λευκή ουσία, τη φαιά ουσία και το εγκεφαλονωτιαίο υγρό

6.2 Σετ Δεδομένων

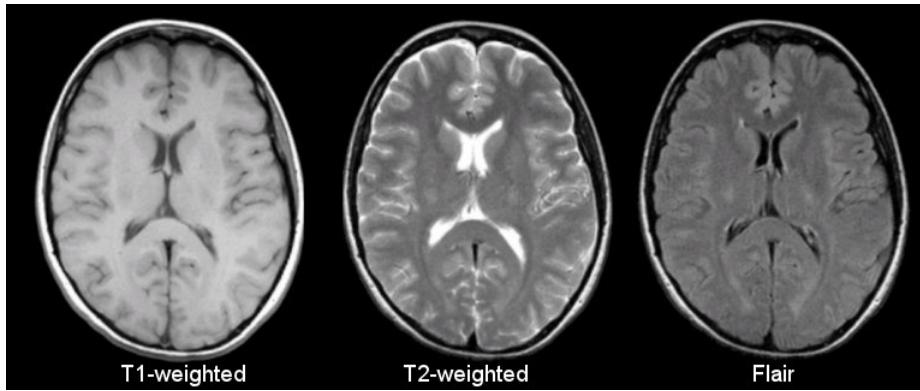
Η μαγνητική τομογραφία ή απεικόνιση μαγνητικού συντονοσισμού (MRI), είναι μία τεχνική για την απεικόνιση της ανατομίας και των φυσιολογικών διεργασι-

ών του ανθρώπινου σώματος. Οι μαγνητικοί τομογράφοι εφαρμόζουν μαγνητικά πεδία ισχυρής έντασης και ραδιοκύματα για να δημιουργήσουν εικόνες των οργάνων και των ιστών. Η μαγνητική τομογραφία δεν περιλαμβάνει την χρήση ιονίζουσας ακτινοβολίας ή ακτίνων X, γεγονός που την διαχωρίζει από την αξονική και την ποζιτρονική τομογραφία (CT/PET). Χρησιμοποιείται ευρέως για διαγνωστικούς σκοπούς και την παρακολούθηση της πορείας ασθενειών. Συγχριτικά με την αξονική/ποζιτρονική τομογραφία παράγει καλύτερη αντίθεση στην απεικόνιση μαλακών ιστών όπως για παράδειγμα στον εγκέφαλο.

Τηράχουν τρεις κατηγορίες τεχνικών απεικόνισης που διαφοροποιούνται από τις τιμές που δίνουμε στις μεταβλητές TR και TE, όπου TR είναι το χρονικό διάστημα ανάμεσα στις αλληλουχίες διαδοχικών παλμών που εφαρμόζονται στην ίδια τομή και TE το χρονικό διάστημα που μεσολαβεί ανάμεσα στην αποστολή του ραδιοκύματος του παλμού και την λήψη του σήματος. Αναλυτικότερα:

- **T1** : Στην απεικόνιση τύπου T1, οι λιπώδεις ιστοί λαμβάνουν υψηλότερες τιμές φωτεινότητας, ενώ οι υγροί ιστοί όπως το εγκεφαλονωτιαίο υγρό και τα κόκκαλα χαμηλότερες. Φαιά και λευκή ουσία λαμβάνουν ενδιάμεσες τιμές φωτεινότητας με την φαιά ουσία να παρουσιάζεται σκοτεινότερη της λευκής.
- **T2** : Στην απεικόνιση τύπου T2, τα μέρη του σώματος που έχουν υψηλή περιεκτικότητα σε νερό όπως τα οιδήματα και οι όγκοι παρουσιάζουν υψηλότερες τιμές φωτεινότητας. Φαιά και λευκή ουσία λαμβάνουν ενδιάμεσες τιμές φωτεινότητας με την λευκή ουσία να παρουσιάζεται σκοτεινότερη της φαιάς.
- **Flair** : Στην απεικόνιση τύπου FLAIR (*Fluid-attenuated inversion recovery*) εξασφαλίζεται η εκμηδένιση των υγρών ιστών του απεικονιζόμενου όγκου, δίνοντας προτεραιότητα στην εμφάνιση των λιπωδών ιστών όπως η φαιά και η λευκή ουσία.

Για την εκπαίδευση του μοντέλου κατάτμησης του εγκεφάλου στους βασικούς του ιστούς χρησιμοποιήθηκαν εικόνες του σετ δεδομένων OASIS3. Στο σετ περιέχονται εικόνες εγκεφάλων ασθενών απεικονισμένες με την T1 τεχνική καθώς και οι αντίστοιχες μάσκες που διαχωρίζουν τα στοιχεία της κάθε εικόνας στις τρεις επιμέρους κλάσεις. Η διαδικασία του διαχωρισμού πραγματοποιήθηκε



Σχήμα 6.2: T1,T2 και FLAIR τεχνικές απεικόνισης

από ιατρικούς ειδικούς. Το εγκεφαλονωτιαίο υγρό απεικονίζεται με την χαμηλότερη φωτεινότητα, η φαιά ουσία με ενδιάμεσες τιμές φωτεινότητας και η λευκή ουσία με τις υψηλότερες. Όλες οι εικόνες του σετ είναι εγγεγραμμένες σε έναν άτλαντα πρότυπο, οπότε στο στάδιο της προεπεξεργασίας που θα αναλυθεί παρακάτω δεν είναι απαραίτητη η εγγραφή των εικόνων.

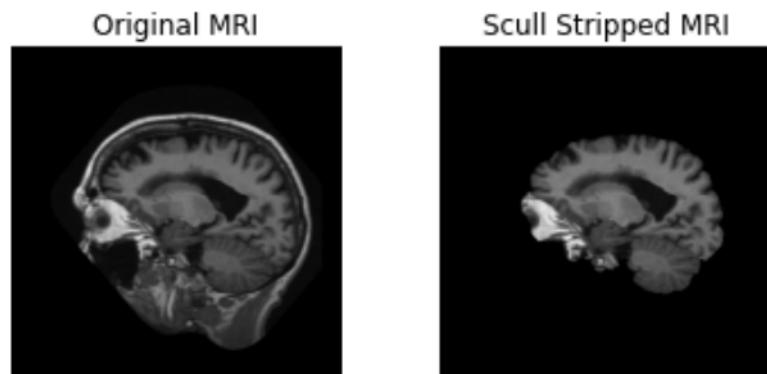
6.3 Προεπεξεργασία Δεδομένων

Οι άτλαντες που αποτελούν το σετ δεδομένων πριν τη χρήση τους για την εκπαίδευση του μοντέλου υποβλήθηκαν σε μία σειρά διαδικασιών προεπεξεργασίας τους, με σκοπό την απομάκρυνση του θορύβου και της περιττής πληροφορίας και την εξομάλυνση των δεδομένων.

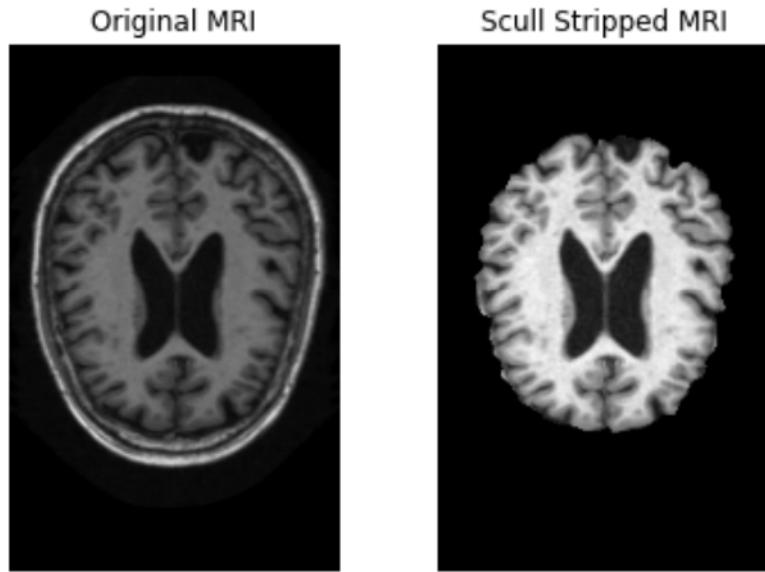
1. Skull Stripping

Οι άτλαντες του σετ στην πρωταρχική τους μορφή απεικονίζουν τον ανθρώπινο εγκέφαλο και τους περιβάλλοντες ιστούς όπως το κρανίο και τα μάτια. Σε πρώτη φάση της προεπεξεργασίας χρίθηκε αναγκαίο να απομονωθεί ο εγκέφαλος, που αποτελεί το σημείο ενδιαφέροντος της κατάτμησης, από το κρανίο και τους περιβάλλοντες ιστούς που δεν αφορούν το πρόβλημα, έτσι ώστε να πραγματοποιηθεί με ακρίβεια το επόμενο βήμα προεπεξεργασίας, η εγγραφή της κάθε εικόνας στην αντίστοιχη μάσκα της. Για την αφαίρεση του κρανίου

από την εικόνα χρησιμοποιήθηκαν εργαλεία της βιβλιοθήκης FSL, που χρησιμοποιείται ευρέως στην ανάλυση και προεπεξεργασία ιατρικών εικόνων MRI εγκεφάλων[26]. Τα Σχήματα (6.3) και (6.4) απεικονίζουν δύο διαφορέτικές όψεις ενός άτλαντα του σετ. Στο αριστερό μέρος των εικόνων βλέπουμε έναν από τους άτλαντες του σετ στην πρωταρχική του μορφή και στο δεξί μέρος τον άτλαντα μετά την αφαίρεση του κρανίου.



Σχήμα 6.3: Πρόσοψη MRI εικόνας πριν και μετά την αφαίρεση του κρανίου



Σχήμα 6.4: Κάτοψη MRI εικόνας πριν και μετά την αφαίρεση του κρανίου

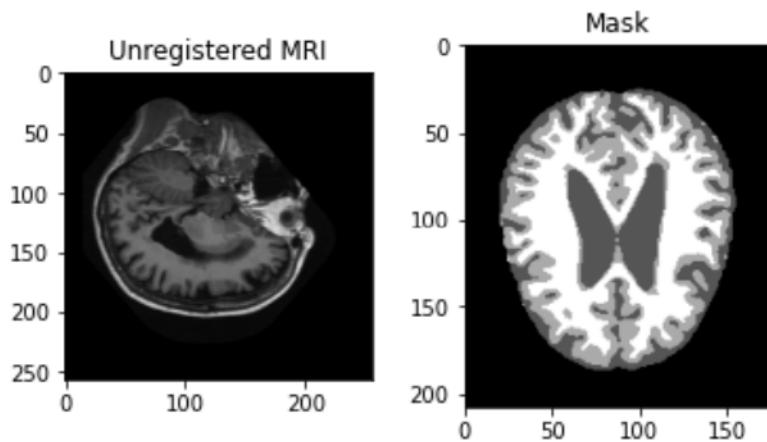
2. Mask-Image Registration

Σε επόμενη φάση της προεπεξεργασίας, εφαρμόστηκε εγγραφή του κάθε άτλαντα στην αντίστοιχη μάσκα του. Η διαδικασία αυτή ήταν αναγκαία καθώς οι διαστάσεις των εικόνων είναι διαφορετικές από τις διαστάσεις των μασκών [27]. Η εγγραφή της εικόνας στην μάσκα της όταν έχουν διαφορετικές διαστάσεις αποτελεί ουσιαστικά τον μετασχηματισμό της εικόνας έτσι ώστε να ταυτίζεται με τον χωρικό προσανατολισμό και την κλίμακα της μάσκας. Η διαδικασία αποτελέεται από τα εξής βήματα:

1. Δειγματοληψία της εικόνας έτσι ώστε να αποκτήσει το ίδιο μέγεθος και ανάλυση με την μάσκα. Με χρήση παρεμβολής, η εικόνα που προκύπτει έχει pixels ίδιων διαστάσεων και κατεύθυνσης με την μάσκα.
2. Εγγραφή της μάσκας στη δειγματοληπτούμενη εικόνα με τη χρήση κάποιου μετασχηματισμού.

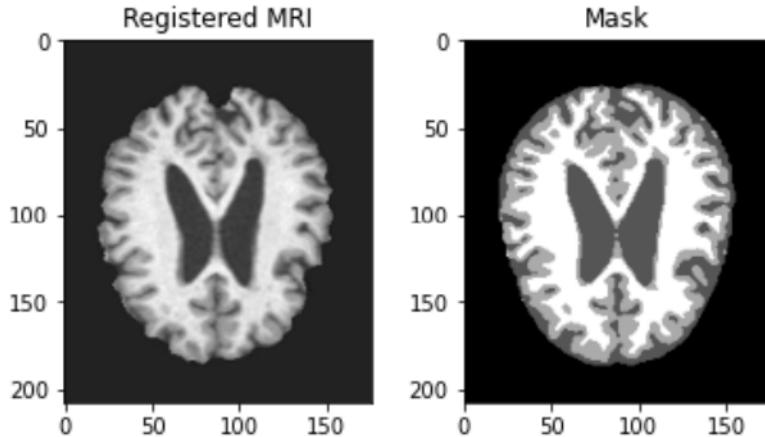
Για την εγγραφή των εικόνων στις μάσκες τους έγινε χρήση της βιβλιοθήκης Simple Elastix, που χρησιμοποιείται για την εγγραφή ιατρικών εικόνων και α-

νήκει στην ευρύτερη βιβλιοθήκη ανάλυσης και επεξεργασίας ιατρικών εικόνων SimpleITK[28]. Η εγγραφή εφαρμόστηκε με χρήση γραμμικής παρεμβολής της εικόνας και affine μετασχηματισμού. Στο Σχήμα (6.5) φαίνεται ένας άτλαντας και η μάσκα του πριν την εγγραφή. Οι διαστάσεις της MRI εικόνας είναι διαφορετικές και οι δύο εικόνες έχουν διαφορετικό προσανατολισμό.



Σχήμα 6.5: MRI εικόνα πριν την εγγραφή στην μάσκα της. Ο προσανατολισμός του εγκέφαλου και οι διαστάσεις της MRI εικόνας είναι διαφορετικές από της μάσκας

Στο Σχήμα (6.6) φαίνεται ο ίδιος άτλαντας και η μάσκα του μετά από την εγγραφή.



Σχήμα 6.6: MRI εικόνα μετά την εγγραφή της στη μάσκα της

3.Denoisation

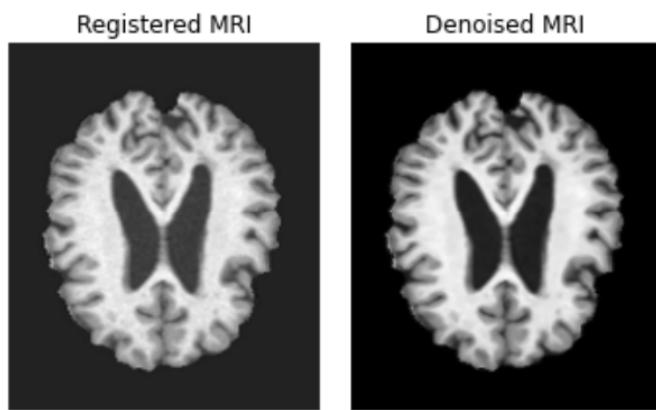
Η αποθορυβοποίηση είναι μια πολύ σημαντική διαδικασία στην επεξεργασία εικόνας, καθώς οι εικόνες σχεδόν πάντα έχουν υποστεί αλλοίωση από διάφορες πηγές θορύβου. Υπάρχουν αρκετές τεχνικές αποθορυβοποίησης, μερικές εκ των οποίων είναι:

1.Gaussian Smoothing: Εφαρμόζεται ένα φίλτρο εξομάλυνσης στην εικόνα που μετριάζει τις τιμές των γειτονικών pixels, με αποτέλεσμα να απομακρυνθεί ο θόρυβος υψηλών συχνοτήτων, με το τίμημα να χαθεί ένα μέρος πληροφορίας.

2.Median φίλτρο: Σε αυτή την τεχνική κάθε pixel της εικόνας αντικαθίσταται από τη μέση τιμή των γειτονικών του pixels. Απομακρύνει αποτελεσματικά τον θόρυβο, ενώ διατηρεί την πληροφορία των ακμών και σημαντικά χαρακτηριστικά της εικόνας.

3.Non Local Means (NLM) φίλτρο : Εφαρμόζει όπως και το γκαουσιανό φίλτρο εξομάλυνση των τιμών των γειτονικών pixels, συνυπολογίζοντας όμως την ομοιότητα των γειτονικών pixels. Με αυτόν τον τρόπο τα pixels με όμοια γειτονικά pixels παίρνουν υψηλότερες τιμές βαρών στη διαδικασία εξομάλυνσης, ενώ αυτά με ανόμοια γειτονικά pixels παίρνουν χαμηλότερες τιμές βαρών.

Η κύρια πηγή θορύβου στις εικόνες εγκεφάλων που εξετάζουμε είναι γκαουσιανής προέλευσης. Έτσι εφαρμόστηκε για την αποθορυβοποίησή τους η τεχνική Non local Means (NLM), που είναι αρκετά αποτελεσματική σε τέτοιου είδους θόρυβο. Τα εργαλεία που χρησιμοποιήθηκαν για την αποθορυβοποίηση βρίσκονται στην βιβλιοθήκη της Python DIPY που χρησιμοποιείται για ανάλυση και προεπεξεργασία ιατρικών εικόνων MRI[29]. Στο αριστερό μέρος του Σχήματος (6.7) φαίνεται ένας άτλαντας του σετ πριν την αποθορυβοποίηση και στο δεξί μέρος μετά την αποθορυβοποίηση.

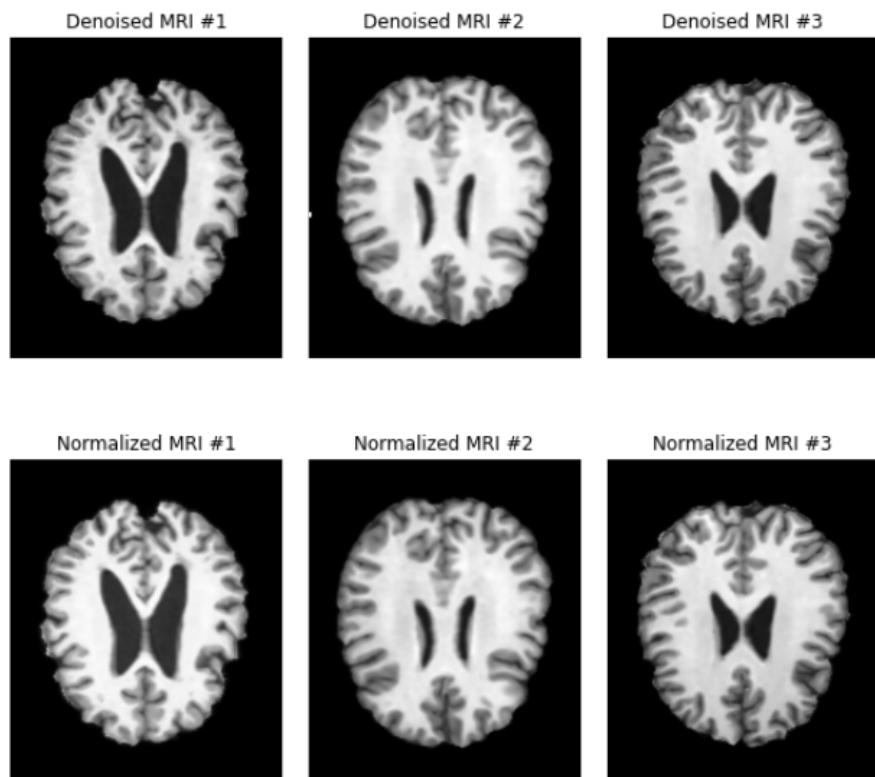


Σχήμα 6.7: MRI εικόνα πριν και μετά την αποθορυβοποίηση με χρήση *non-local – means* φίλτρου

4. Κανονικοποίηση φωτεινοτήτων

Η κανονικοποίηση φωτεινοτήτων είναι μια διαδικασία όπου οι τιμές φωτεινοτήτων της εικόνας μετασχηματίζονται σε ένα συγκεχριμένο έυρος τιμών. Αυτό συμβαίνει σε περίπτωση που υπάρχει ένα σετ εικόνων προς σύγκριση έτσι ώστε να έχουν τιμές στο ίδιο εύρος και να γίνει η διαδικασία σύγκρισης ή η εκπαίδευση ενός μοντέλου πιο αποτελεσματική. Η διαδικασία στοχεύει σε εξομάλυνση ανομοιοτήτων φωτεινοτήτων που οφείλονται σε πιθανές διαφορές παραμέτρων κατά τη διαδικασία άλλων προεπεξεργαστικών μεθόδων ή στις διαφορές μεταξύ των scanner που πάραξαν τις εικόνες.

Για το σετ εικόνων εγκεφάλων εφαρμόστηκε η μέθοδος κανονικοποίησης Nyul [30]. Η μέθοδος αυτή αντιστοιχίζει τις φωτεινότητες κάθε εικόνας σε ένα έυρος φωτεινοτήτων αναφοράς. Στο πάνω μέρος του Σχήματος (6.8) φαίνονται 3 εικόνες του σετ πριν την κανονικοποίηση και στο κάτω μέρος του Σχήματος (6.8) οι ίδιες εικόνες μετά την κανονικοποίηση.



Σχήμα 6.8: MRI εικόνες πριν την κανονικοποίηση Nyul, (πάνω μέρος της εικόνας) και MRI εικόνες μετά την κανονικοποίηση Nyul (κάτω μέρος της εικόνας)

6.4 Επιλογή Χαρακτηριστικών

Για την περιγραφή του σετ ατλάντων εξετάστηκαν όλα τα χαρακτηριστικά από τις διάφορες κατηγορίες που αναπτύχθηκαν στην ενότητα 5, που είναι ανακεφαλαιώνοντας τα εξής :

- Φωτεινότητα
- Τοπολογικά Χαρακτηριστικά
 - Συντεταγμένες x, y, z
 - Χωρική θέση
- Χαρακτηριστικά Ακμών
- Στατιστικά Χαρακτηριστικά
- Χαρακτηριστικά Υφής
 - *HOG*
 - *LBP*
 - *GLCM*
 - *GLRLM*
 - *GLSZM*
 - *GLDM*
 - *NGTDM*

Με στόχο τη μείωση της διαστασιμότητας του μοντέλου και την επιλογή των χαρακτηριστικών που περιέχουν την περισσότερη πληροφορία, εφαρμόστηκαν διάφορες τεχνικές επιλογής χαρακτηριστικών κατηγορίας φίλτρου και wrapper, όπως αναπτύχθηκε στην ενότητα 5.2 :

- **kBest**

Η τεχνική φίλτρου kBest χρησιμοποιεί την στατιστική μετρική Mutual Information για να εξετάσει τη συσχέτιση και την πληροφορία των χαρακτηριστικών, επιλέγοντας τα N καλύτερα χαρακτηριστικά που θα ορίσει ο χρήστης.

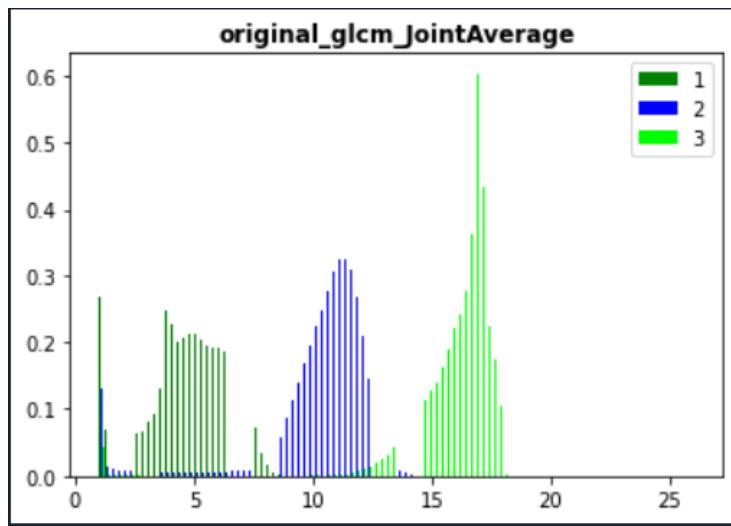
- **Recursive Feature Elimination(RFE)**

H Wrapper τύπου τεχνική RFE επιστρέφει τα N καλύτερα χαρακτηριστικά που όρισε ο χρήστης εκπαιδεύοντας ένα μοντέλο και απορρίπτοντας σε κάθε εκπαίδευση το λιγότερο χρήσιμο χαρακτηριστικό. Ως μοντέλο εκτίμησης επιλέχθηκε ένας γραμμικός SVM ταξινομητής.

- **Random Forest Recursive Feature Elimination(RFE)**

Στην τεχνική αυτή εφαρμόζεται η μέθοδος RFE με τη χρήση ενός Random Forest Classifier, ώστε να προκύψουν τα επιλεγμένα χαρακτηριστικά με σειρά σημαντικότητας.

Επίσης σε κάποια από τα χαρακτηριστικά, η συνεισφορά τους τους στο μοντέλο μπορεί να γίνει αντιληπτή από την παρατήρηση της διαχωρισμότητας των κατανομών των τιμών των τριών κλάσεων της εικόνας όπως φαίνεται στο Σχήμα (6.9).



Σχήμα 6.9: MRI Κατανομές φωτεινοτήτων των τριών κλάσεων για το χαρακτηριστικό *GLCM*

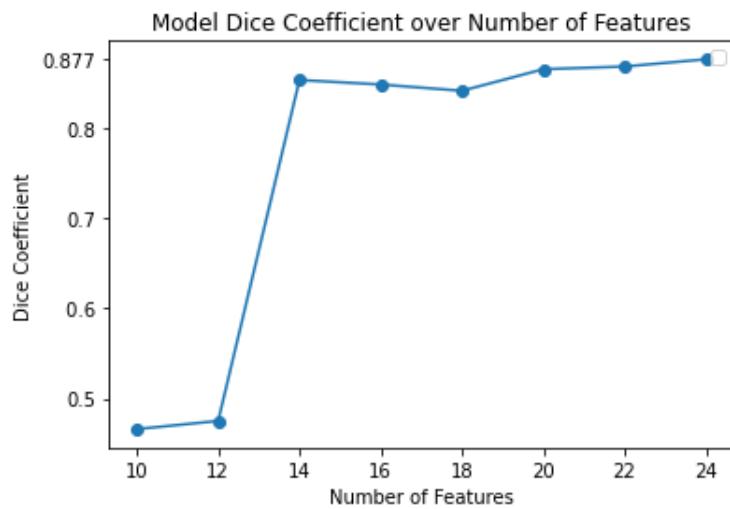
Χαρακτηριστικά που δεν παρουσιάζουν αυτή την ολική διαχωρισμότητα δε σημαίνει πως δεν είναι χρήσιμα, διότι μπορεί σε τοπικό επίπεδο να διαχωρίζουν τις κλάσεις και να μην γίνεται αντιληπτό σε ολικό επίπεδο. Συμπερασματικά για

την τελική επιλογή συνδυάζουμε την επίβλεψη των κατανομών, την εμπειρική μας γνώση και την εφαρμογή της μεθόδου RFE με τη χρήση ενός Random Forest Classifier. Αναλυτικότερα, η μέθοδος εφαρμόστηκε σε μία τυχαία εικόνα του σετ, για όλες τις κατηγορίες χαρακτηριστικών που αναφέρθηκαν, με σκοπό την ιεράρχηση των 25 πιο πλούσιων σε πληροφορία χαρακτηριστικών. Στη συνέχεια, διεξάχθηκε σειρά πειραμάτων στο σύνολο των 25 εικόνων του σετ εκπαίδευσης για την εύρεση του βέλτιστου αριθμού χαρακτηριστικών με κριτήριο τη μετρική Dice Coefficient. Τα αποτελέσματα παρουσιάζονται στο σχήμα (6.10). Στη σειρά των πειραμάτων μοναδική μεταβλητή αποτελεί ο αριθμός των χαρακτηριστικών ο οποίος αυξάνεται κάθε φορά με βήμα 2, ξεκινώντας από 10 χαρακτηριστικά και καταλήγοντας σε 24. Η μετρική παρουσιάζει απότομη αύξηση στα 14 χαρακτηριστικά και έπειτα παρουσιάζει μικρές αυξομειώσεις μέχρι την τελική και μέγιστη τιμή που λαμβάνει στα 24 χαρακτηριστικά. Παρατηρείται ότι έπειτα από τα 20 σημαντικότερα χαρακτηριστικά, η καμπύλη παρουσιάζει οριακά θετική κλίση, με εν δυνάμει αύξηση των επιλεγόμενων χαρακτηριστικών να συνεπάγεται ισχνή βελτίωση του μοντέλου αλλά και αυξημένο υπολογιστικό κόστος.

Για την εξαγωγή των στατιστικών χαρακτηριστικών και των χαρακτηριστικών υφής χρησιμοποιήθηκε η βιβλιοθήκη Pyradomics, που είναι ευρέως γνωστή για την εξαγωγή χαρακτηριστικών σε ιατρικές εικόνες [31]. Όλα τα χαρακτηριστικά εξήχθησαν για κάθε voxel της εικόνας (Voxel-based εξαγωγή) και κανονικοποιήθηκαν στο εύρος τιμών $[0, 1]$ για να συμβάλλουν ισότιμα στην εκπαίδευση του μοντέλου. Για την εξαγωγή των χαρακτηριστικών υφής του καθε voxel εξετάζεται μια περιοχή γειτονικών voxels της οποίας η ακτίνα ορίστηκε ίση με 3, σε απόσταση δηλαδή τριών γειτονικών voxels προς κάθε κατεύθυνση. Εντέλει για την εκπαίδευση του μοντέλου επιλέχθηκαν τα εξής χαρακτηριστικά:

1. Φωτεινότητα
2. Συντεταγμένη X
3. Συντεταγμένη Y
4. Συντεταγμένη Z
5. Χωρική Θέση
6. LBP
7. Mean(Στατιστικά)
8. Median(Στατιστικά)
9. Root Mean Squared(Στατιστικά)
10. 10 Percentile(Στατιστικά)
11. 90 Percentile(Στατιστικά)

-
12. $\text{Minimum}(\Sigma\tau\alpha\tau\iota\sigma\tau\iota\chi\acute{\alpha})$
 13. $\text{Maximum}(\Sigma\tau\alpha\tau\iota\sigma\tau\iota\chi\acute{\alpha})$
 14. $\text{Energy}(\Sigma\tau\alpha\tau\iota\sigma\tau\iota\chi\acute{\alpha})$
 15. $\text{Total Energy}(\Sigma\tau\alpha\tau\iota\sigma\tau\iota\chi\acute{\alpha})$
 16. $\text{Autocorrelation(GLCM)}$
 17. $Idn(GLCM)$
 18. High Gray Level Emphasis(GLRLM)
 19. Low Gray Level Emphasis(GLRLM)
 20. Short Run High Gray Level Emphasis(GLRLM)
 21. Long Run High Gray Level Emphasis(GLRLM)
 22. High Gray Level Emphasis(GLDM)
 23. Low Gray Level Emphasis(GLDM)
 24. Large Dependence High Gray Level Emphasis(GLDM)



Σχήμα 6.10: Η μετρική Dice για αυξανόμενο αριθμό χαρακτηριστικών

6.5 Εκπαίδευση μοντέλου

Μετά την διαδικασία της προεπεξεργασίας των εικόνων και της εξαγωγής χαρακτηριστικών που να τις περιγράφουν ακολούθησε η εκπαίδευση του προτεινόμενου μοντέλου. Θεωρώντας αρχικά μια εικόνα του σετ ως εικόνα προς κατάτμηση και τις υπόλοιπες εικόνες του σετ ως άτλαντες εκπαίδευσης εφαρμόστηκαν τα εξής βήματα.

- Επιλογή των N πιο όμοιων άτλαντων με την εικόνα προς κατάτμηση μέσω της μετρικής MAE, ώστε οι αδύναμοι SVM ταξινομητές να εκπαιδευτούν πάνω σε αυτές.
- Εφαρμογή του αλγόριθμου 2 της ενότητας 4.2.
- Αξιολόγηση των αποτελέσματων και διορθωση των παραμέτρων του μοντέλου.
- Εκπαίδευση του μοντέλου με τις νέες παραμέτρους μέχρι να φτάσει σε ικανοποιητικό αποτέλεσμα.

Για την επιβεβαίωση πως η διαδικασία της μέτρησης ομοιότητας ανάμεσα στις εικόνες βελτιώνει την ικανότητα πρόβλεψης του μοντέλου, εφαρμόστηκε σειρά πειραμάτων στις 25 εικόνες του σετ, όπου για κάθε εικόνα προς κατάτμηση επιλέχθησαν ως σετ εκπαίδευσης εναλλακτικά οι 3 πιο ανόμοιες και οι 3 πιο όμοιες εικόνες. Το αποτέλεσμα της σύγκρισης παρουσιάζεται στον Πίνακα 3 και καταδεικνύει την σημασία της επιλογής ομοίων εικόνων ως σετ εκπαίδευσης.

Metric	Unsimilar	Similar
Dice Coefficient	0.515	0.877
Accuracy	0.513	0.885
F1-Score	0.492	0.881
Precision	0.627	0.895
Recall	0.513	0.885
Precision Class 1	0.886	0.904
Precision Class 2	0.379	0.947
Precision Class 3	0.544	0.944
Recall Class 1	0.506	0.834
Recall Class 2	0.531	0.889
Recall Class 3	0.530	0.908

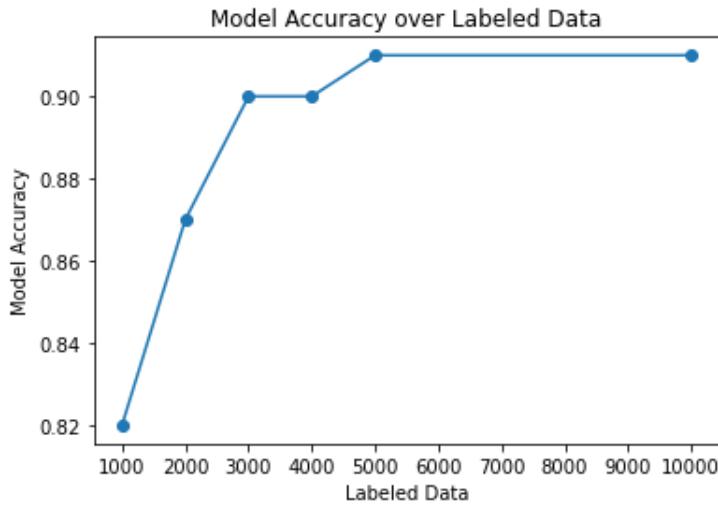
Πίνακας 3: Μετρικές απόδοσης βέλτιστου μοντέλου για τις πιο ανόμοιες (2η στήλη) και τις πιο όμοιες εικόνες (3η στήλη) για $W=3$ και $F=24$

Η παραμετροποίηση του μοντέλου αφορά τα παραχώτω μεγέθη:

- Αριθμός κατηγοριοποιημένων δεδομένων που δειγματολειπτεί ο κάθε SVM ταξινομητής από έναν άτλαντα εκπαίδευσης.
- Αριθμός μη κατηγοριοποιημένων δεδομένων της εικόνας προς κατάτμηση για την επίλυση του προβλήματος (4.10) του αλγορίθμου 2.
- Υπερπαράμετροι C των SVM ταξινομητών του μοντέλου.
- Αριθμός SVM ταξινομητών.

Για την εκπαίδευση των αδύναμων SVM ταξινομητών και για την επίλυση του προβλήματος (4.10) χρησιμοποιήθηκε ο γραμμικός SVM αλγόριθμος LIBLINEAR ο οποίος χρησιμοποιεί γραμμικό πυρήνα, πράγμα που τον καθιστά αρκετά ταχύτερο από άλλους SVM ταξινομητές. Μετά από ρύθμιση η βέλτιστη υπερπαράμετρος για τον ταξινομητή των αδύναμων SVM ταξινομητών επιλέχθηκε η $c = 1$, και για τον επιλυτή του προβλήματος (4.10) επιλέχθηκε πάλι η $c = 1$. Αρκετά ενδιαφέρον κομμάτι αποτελεί ο αριθμός των κατηγοριοποιημένων δεδομένων που χρησιμοποιούνται για την εκπαίδευση των αδύναμων SVM ταξινομητών καθώς ορίζει τον βαθμό της ημιεπιβλεπόμενης μάθησης που χαρακτηρίζει το μοντέλο. Στο Σχήμα (6.11) φαίνεται η απόδοση του μοντέλου για μια εικόνα του σετ, για διάφορες τιμές των κατηγοριοποιημένων δεδομένων. Παρατηρούμε

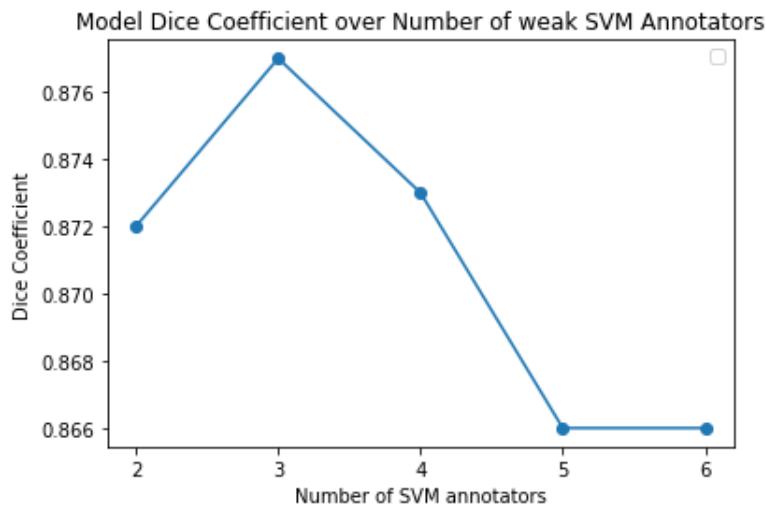
πως για τιμές άνω των 5000 η ακρίβεια του μοντέλου παραμένει η ίδια. Μία εικόνα περιέχει περίπου 1500000 δεδομένα (pixels) επομένως βλέπουμε πως μπορούμε να εκπαιδεύσουμε αποτελεσματικά το μοντέλο με μόλις 0.33% των συνολικών δεδομένων του κάθε άτλαντα εκπαίδευσης.



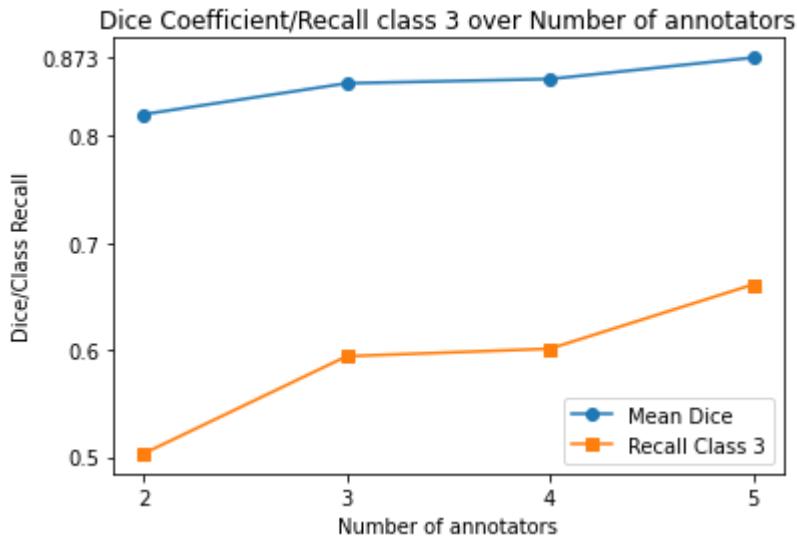
Σχήμα 6.11: Η ακρίβεια του μοντέλου αναλογικά με τον αριθμό κατηγοριοποιημένων δεδομένων που χρησιμοποιούνται για την εκπαίδευση του

Η επόμενη σειρά πειραμάτων αφορά τον βέλτιστο αριθμό αδύναμων ταξινομητών W . Για παράδειγμα, για $W = 2$ επιλέγονται οι 2 πιο όμοιοι άτλαντες με την εικόνα στόχο από το σετ εκπαίδευσης, με τον κάθε αδύναμο ταξινομητή να εκπαιδεύεται σε μία εικόνα, για $W = 3$ οι 3 πιο όμοιες κ.ο.κ.. Στο διάγραμμα του Σχήματος (6.12) φαίνεται ο μέσος όρος της μετρικής Dice για τις διάφορες τιμές του W . Παρατηρούμε πως το μοντέλο παρουσιάζει βέλτιστη απόδοση για $W = 3$ και στη συνέχεια όσο προσθέτουμε ταξινομητές η τιμή της μετρικής μειώνεται. Αυτό μπορεί να ερμηνευτεί ως την ύπαρξη κατά μέσο όρο τριών πιο όμοιων ατλάντων μέσα στο σετ που περιέχουν χρήσιμη πληροφορία για την κατάτμηση της εικόνας στόχου, ενώ όσο προσθέτουμε πιο ανόμοιες εικόνες οι αδύναμοι ταξινομητές που συμβάλλουν στην εκπαίδευση του μοντέλου δεν είναι τόσο αποδοτικοί ζημιώνοντας τη συνολική απόδοση του. Παράλληλα παρατηρήθηκε πως μεμονωμένες εικόνες του σετ για τον βέλτιστο αριθμό αδύναμων

ταξινομητών παρουσιάζουν απόδοση υπό του μέσου όρου, ενώ για αριθμό αδύναμων ταξινομητών μεγαλύτερου του 3 παρατηρήθηκε σημαντική βελτίωση της απόδοσής τους. Η υποβέλτιστη απόδοση των συγκεχριμένων εικόνων οφείλεται κατά βάση στην αδυναμία του βέλτιστου αριθμού ταξινομητών να πρόβλεψει επαρκώς μια από τις 3 κλάσεις, επομένως η προσθήκη ταξινομητών συμβάλλει στην ενίσχυση της πρόβλεψης επί της υπολειπόμενης κλάσης, όπως φαίνεται στο σχήμα (6.13). Από αυτή τη συμπεριφορά του μοντέλου φαίνεται πως ο βέλτιστος αριθμός των αδύναμων ταξινομητών εξαρτάται άμεσα από το μέγεθος του σετ εκπαίδευσης και τον αριθμό των όμοιων εικόνων που υπάρχουν σε αυτό με την εικόνα στόχο.



Σχήμα 6.12: Η μετρική Dice για διάφορες τιμές αδύναμων SVM ταξινομητών



Σχήμα 6.13: Οι μετρικές Dice-Recall σε μία εικόνα με μειωμένη επιτυχία πρόβλεψης για διάφορες τιμές αδύναμων SVM ταξινομητών

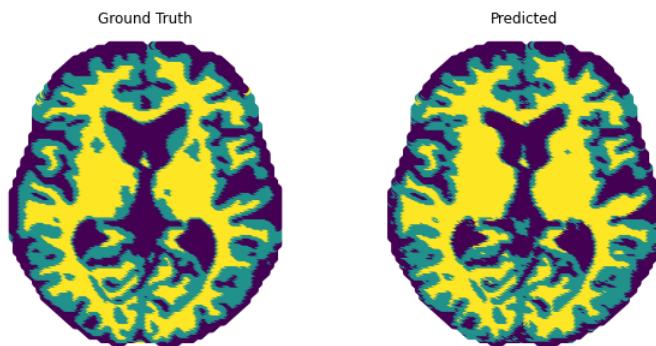
6.6 Αποτελέσματα

Σαν τελική αξιολόγηση του μοντέλου πραγματοποιήθηκε σε σετ 25 εικόνων η τεχνική k-fold Cross Validation για $k = 1$. Για 25 επαναλήψεις θεωρούμε κάθε φορά μια εικόνα προς κατάτμηση από το σετ και τις υπόλοιπες ως σετ εκπαίδευσης. Στον πίνακα 4 φαίνονται οι μετρικές απόδοσης του βέλτιστου μοντέλου για $W = 3$ και $F = 24$. Στα Σχήματα (6.14-23) φαίνονται οι προβλέψεις του μοντέλου για διάφορες εικόνες του σετ. Με μωβ απεικονίζεται η κλάση 1 (εγκεφαλονωτιαίο υγρό), με μπλε η κλάση 2 (γκρι ουσία) και με κίτρινο η κλάση 3 (λευκή ουσία). Όπως φαίνεται στις χαρακτηριστικές τομές, το μοντέλο αποδίδει σε σημαντικό βαθμό τη γεωγραφία του ανθρώπινου εγκεφάλου προβλεποντας με ακρίβεια τις ραβδώσεις και τις αυλακώσεις του. Εντοπίζονται βέβαια και περιοχές που οι κλάσεις 2 και 3 συγχέονται, πιο συγκεκριμένα σε εσωτερικά σημεία του εγκεφάλου όπου η κλάση 2 εμφανίζεται με τη μορφή νησίδων που περικλείονται από την κλάση 3 όπως φαίνεται στα σχήματα (6.16,6.14,6.17). Το ίδιο φαινόμενο παρατηρείται στην περιοχή του στελέχους του εγκεφάλου που ανήκει στην κλάση 3, και ταξινομείται μερικώς στην κλάση 2, όπως φαίνεται

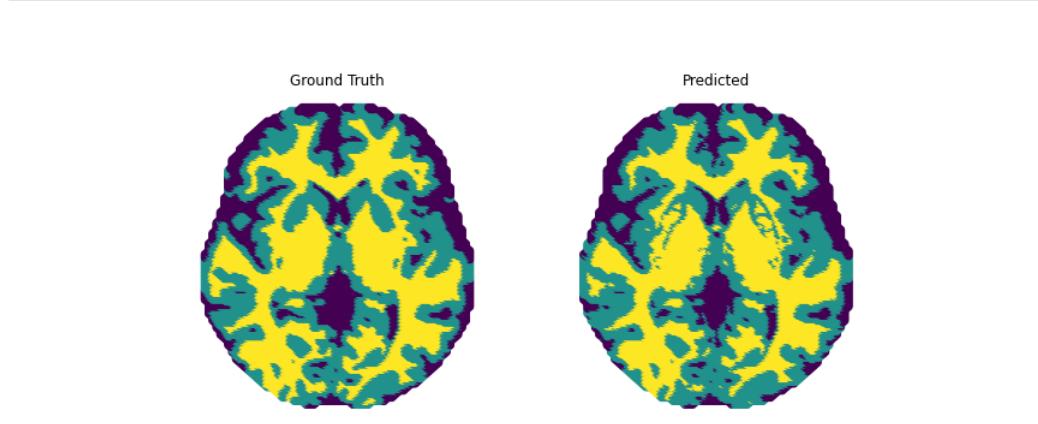
στο σχήμα (6.19). Παρόλα αυτά, αυτές οι αποτυχημένες προβλέψεις είναι περιορισμένες και δεν επιδρούν σημαντικά στην ικανότητα πρόβλεψης πάνω στις κλάσεις 2 και 3 όπως μαρτυρά ο πίνακας 4. Τη χαμηλότερη ικανότητα πρόβλεψης παφουσιάζει η κλάση 1, η οποία τείνει να συγχέεται με την κλάση 2, όταν στην τελευταία το μοντέλο αποτυγχάνει να περιγράψει τις αυλακώσεις της, στις οποίες και εισρέει η κλάση 1, δημιουργώντας συμπαγείς δομές όπως φαίνεται στα σχήματα (6.22 , 6.20).

Dice Coefficient	0.877
Accuracy	0.885
F1-Score	0.881
Precision	0.895
Recall	0.885
Precision Class 1	0.904
Precision Class 2	0.947
Precision Class 3	0.944
Recall Class 1	0.834
Recall Class 2	0.889
Recall Class 3	0.908

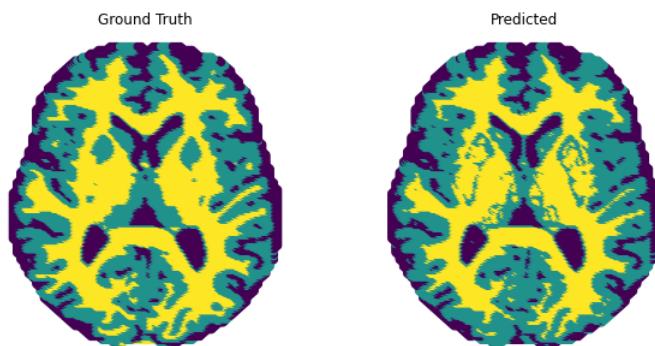
Πίνακας 4: Μετρικές απόδοσης βέλτιστου μοντέλου για $W=3$ και $F=24$



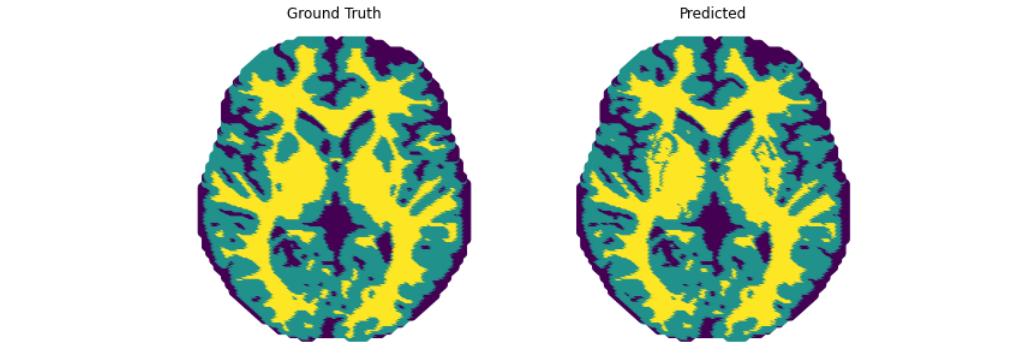
Σχήμα 6.14: Dice=0.91 Αριστερά η εικόνα προς κατάτμηση και δεξιά η πρόβλεψη του μοντέλου



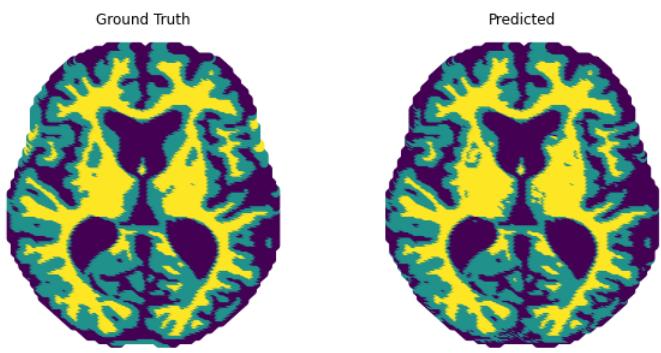
Σχήμα 6.15: Dice=0.93 Αριστερά η εικόνα προς κατάτμηση και δεξιά η πρόβλεψη του μοντέλου



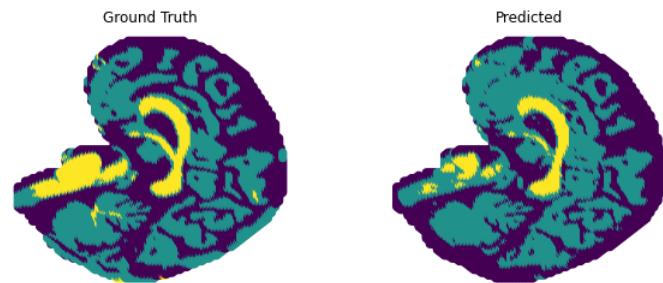
Σχήμα 6.16: Dice=0.92 Αριστερά η εικόνα προς κατάτμηση και δεξιά η πρόβλεψη του μοντέλου



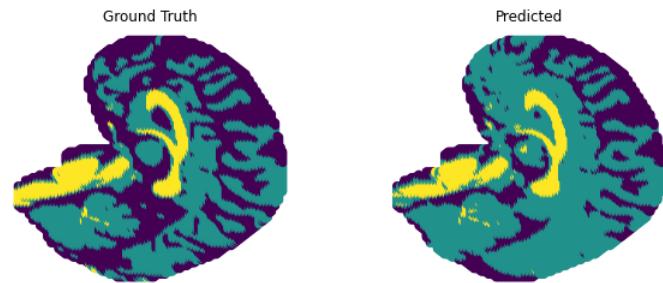
Σχήμα 6.17: Dice=0.92 Αριστερά η εικόνα προς κατάτυπη και δεξιά η πρόβλεψη του μοντέλου



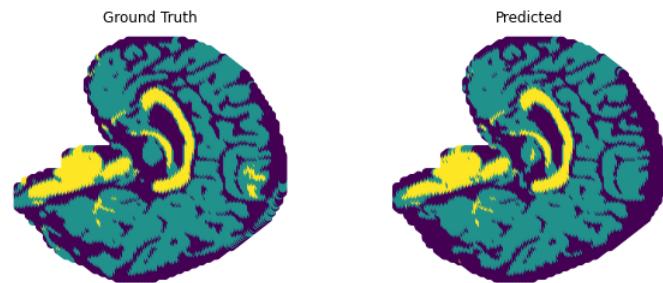
Σχήμα 6.18: Dice=0.95 Αριστερά η εικόνα προς κατάτυπη και δεξιά η πρόβλεψη του μοντέλου



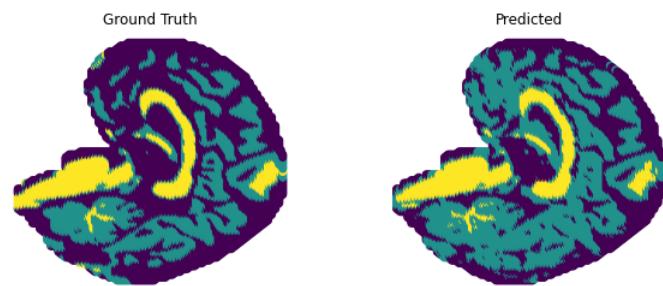
Σχήμα 6.19: Dice=0.88 Αριστερά η εικόνα προς κατάτμηση και δεξιά η πρόβλεψη του μοντέλου



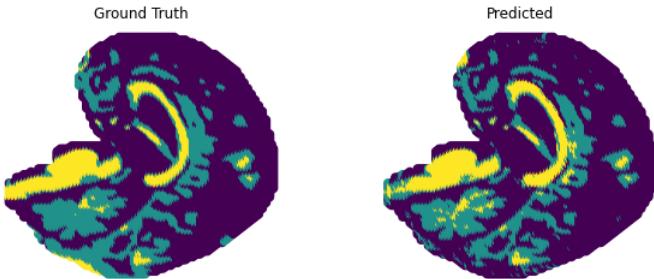
Σχήμα 6.20: Dice=0.86 Αριστερά η εικόνα προς κατάτμηση και δεξιά η πρόβλεψη του μοντέλου



Σχήμα 6.21: Dice=0.95 Αριστερά η εικόνα προς κατάτυπη και δεξιά η πρόβλεψη του μοντέλου



Σχήμα 6.22: Dice=0.86 Αριστερά η εικόνα προς κατάτυπη και δεξιά η πρόβλεψη του μοντέλου



Σχήμα 6.23: Dice=0.94 Αριστερά η εικόνα προς κατάτμηση και δεξιά η πρόβλεψη του μοντέλου

7 Πειραματικά αποτελέσματα στο γόνατο

Το γόνατο είναι η άρθρωση που ενώνει 4 από τα υψηλότερα και τα χαμηλότερα οστά του ποδιού, που ονομάζονται μηριαίο οστό, επιγονατίδα, περόνη και κνήμη. Η άρθρωση του γονάτου είναι η μεγαλύτερη στο ανθρώπινο σώμα και αποτελεί ένα εκτεταμένο δίκτυο μυών και συνδέσμων. Το γόνατο στηρίζεται στους συνδέσμους που ενώνουν τα κόκκαλα και παράγουν το υγρό που λιπάνει την κίνηση των οστών. Ο μηνίσκος είναι ένα είδος χόνδρου που εντοπίζεται στις εξωτερικές και εσωτερικές επιφάνειες της κνήμης. Απορροφάει τους κραδασμούς και παρέχει την κατάλληλη κατανομή του βάρους.

7.1 Πρόβλημα

Στο δεύτερο σκέλος της εργασίας καλούμαστε να λύσουμε το πρόβλημα κατάτμησης ιατρικών εικόνων γονάτων. Αναλυτικότερα, το μοντέλο που έχει υλοποιηθεί επικεντρώνεται στην ταξινόμηση των παρακάτω χόνδρων και οστών:

- **Μηριαίο οστό:** Το μηριαίο οστό αποτελεί το σκελετικό υπόβαθρο του μηρού. Είναι το μακρύτερο οστό του ανθρώπου. Η διάφυση του μηριαίου οστού έχει τριγωνικό σχήμα. Το κάτω μέρος του έχει δύο επικόνδυλους και αρθρώνεται με την επιγονατίδα ενώ με την άρθρωση του γονάτου αρθρώνεται με την κνήμη.
- **Κνημιαίο οστό:** Η κνήμη είναι το μεγαλύτερο και πιο εσωτερικό οστό του κάτω άκρου υπό του γονάτου και μεταφέρει το βάρος του σώματος.

Είναι το δεύτερο μεγαλύτερο οστό του ανθρώπου. Είναι το μόνο οστό της κνήμης που συμμετέχει στην άρθρωση του γονάτου. Το άνω μέρος της κνήμης είναι πλατύ και έχει δύο κονδύλους. Η διάφυση της κνήμης είναι τριγωνική.

- **Μηριαίος χόνδρος:** Ο μηριαίος χόνδρος βρίσκεται περιμετρικά του κάτω μέρους του μηριαίου οστού. Είναι ένας ιστός μαλακότερος των οστών. Η λειτουργία του είναι να αποτρέπει τις τριβές μεταξύ των οστών και να προσδίδει ευλυγισία στην κίνηση.
- **Κνημιαίος χόνδρος:** Ο κνημιαίος χόνδρος βρίσκεται περιμετρικά του άνω μέρους του κνημιαίου οστού. Γειτνιάζει του μηριαίου χόνδρου.

7.2 Σετ Δεδομένων

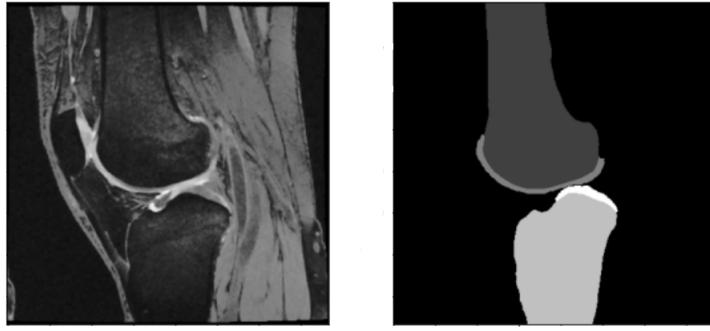
Για την εκπαίδευση του μοντέλου κατάτμησης του γονάτου στους 4 παραπάνω ιστούς αντλήθηκαν εικόνες από το ευρύτερο σετ δεδομένων OAI-ZIB που δημιουργήθηκε με σκοπό την διάγνωση της οστεοαρθρίτιδας. Οι εικόνες είναι παράγωγα μαγνητικής τομογραφίας τύπου T2 ενώ στο σετ περιέχονται και οι αντίστοιχες μάσκες στις οποίες τα γόνατα έχουν διαμεριστεί στις 4 επιμέρους κλάσεις ενδιαφέροντος από ιατρικούς ειδικούς. Στο αριστερό μέρος του Σχήματος (7.1) φαίνεται η τομή μιας τρισδιάστασης εικόνας του σετ, ενώ δεξιά η αντίστοιχη μάσκα. Με μάύρο απεικονίζονται ιστοί και μύες που δεν παρουσιάζουν ενδιαφέρον και χαρακτηρίζονται ως background και από την χαμηλότερη στην υψηλότερη φωτεινότητα απεικονίζονται το μηριαίο οστό, ο μηριαίος χόνδρος, το κνημιαίο οστό και ο κνημιαίος χόνδρος.

7.3 Προεπεξεργασία Δεδομένων

Όπως και στο σετ ατλάντων εγκεφάλων, ακολουθήθηκε διαδικασία προεπεξεργασίας και για το σετ ατλάντων των γονάτων, με σκοπό την απομάκρυνση του θορύβου και της περιττής πληροφορίας, την εξομάλυνση των δεδομένων αλλά και την χωρική ευθυγράμμιση των εικόνων για να επιτευχθούν τα βέλτιστα αποτελέσματα με την εφαρμογή του αλγορίθμου κατάτμησης.

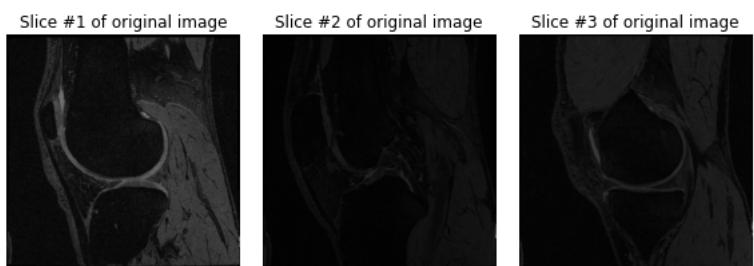
2.Denoisation

Για την αποθορυβοποίησή των MRI εικόνων γονάτων εφαρμόστηκε ένα Median



Σχήμα 7.1: Τομή T2-MRI εικόνας από το σετ OAI-ZIB

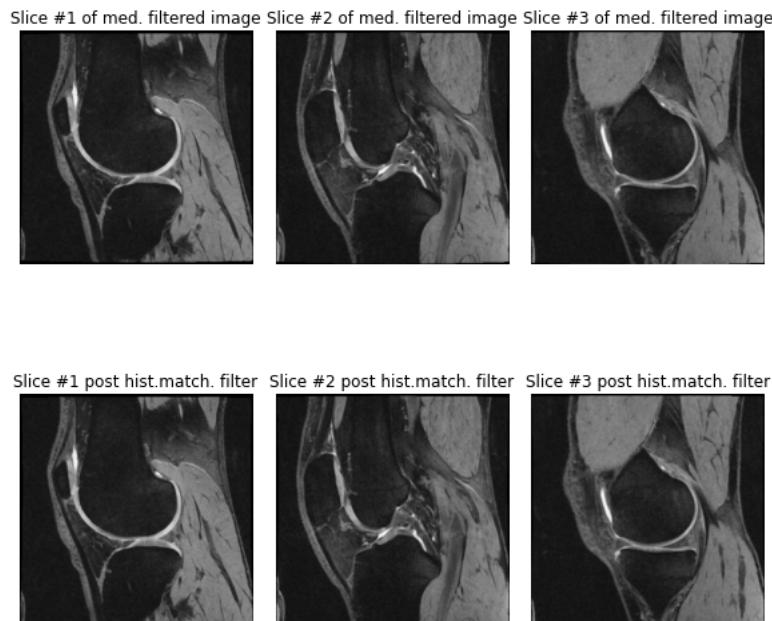
φίλτρο που είναι αρχετά αποτελεσματικό σε τέτοιου είδους εικόνες. Όπως αναλύθηκε στην ενότητα 6.3, το φίλτρο αυτό αντικαθιστά την τιμή του κάθε pixel με τη μέση τιμή των γειτονικών του, διατηρώντας τη χρήσιμη πληροφορία των στοιχείων της εικόνας και των ακμών της. Τα εργαλεία που χρησιμοποιήθηκαν για την αποθορυβοποίηση βρίσκονται στην βιβλιοθήκη της Python SimpleITK που χρησιμοποιείται για ανάλυση και προεπεξεργασία ιατρικών εικόνων MRI. Στο άνω μέρος του Σχήματος (7.2) φαίνεται ένας άτλαντας του σετ πριν την αποθορυβοποίηση και στο κάτω μέρος μετά την αποθορυβοποίηση.



Σχήμα 7.2: Τομές εικόνας πριν και μετά το median filtering

3. Intensity Histogram Matching

Σε επόμενο στάδιο της προεπεξεργασίας, λόγω των μεγάλων διαφορών μεταξύ των φωτεινοτήτων των εικόνων στο σετ (π.χ κάποιες εικόνες έχουν σκοτεινότερο περιεχόμενο από άλλες), εφαρμόζεται η τεχνική Intensity Histogram Matching, που ουσιαστικά προσαρμόζει το εύρος φωτεινοτήτων της κάθε εικόνας σε ένα προκαθορισμένο εύρος φωτεινοτήτων για να είναι ευκολότερη η σύγκρισή τους. Τα εργαλεία που χρησιμοποιήθηκαν βρίσκονται στη βιβλιοθήκη της Python, SimpleITK. Στο πάνω μέρος του Σχήματος (7.3) φαίνονται 3 εικόνες του σετ πριν το Histogram Matching και στο κάτω μέρος της εικόνας οι ίδιες εικόνες μετά το Histogram Matching.



Σχήμα 7.3: Τομές εικόνων πριν και μετά το histogram matching

3. Standardization

Επόμενο στάδιο της προεπεξεργασίας της εικόνας αποτελεί η κανονικοποίηση των τιμών φωτεινοτήτων της εικόνας. Στη διαδικασία αυτή οι τιμές της φωτεινότητας μετασχηματίζονται σε ένα προκαθορισμένο εύρος τιμών. Στόχος της κανονικοποίησης είναι η εξομάλυνση των διαφορών των τιμών φωτεινότητας

ανάμεσα στις εικόνες, και η ένταξη των κατανομών της φωτεινότητας σε ένα κοινό πλαίσιο με αποτέλεσμα την βελτίωση της ικανότητας σύγχρισης ανάμεσα στις εικόνες και την αποτελεσματικότητα της εκπαίδευσης ενός μοντέλου.

4. Image Registration

Το τελευταίο στάδιο της προεπεξεργασίας των ατλάντων γονάτων περιέχει την εγγραφή των εικόνων σε έναν κοινό όπλαντα, έτσι ώστε να υπάρξει μια χωρική ταύτιση των ιστών που απαρτίζουν τις εικόνες. Όπως αναλύθηκε στην ενότητα 3.1, η διαδικασία της εγγραφής μπορεί να περιέχει περιστροφή, μετακίνηση, κλιμάκωση και περικοπή της εικόνας που εγγράφεται έτσι ώστε να ευθυγραμμιστεί με την εικόνα πρότυπο. Ως μετασχηματισμός εγγραφής επιλέχθηκε ο Affine και γραμμικός παρεμβολέας για δειγματολειψία. Στο πάνω μέρος του Σχήματος (7.4) παρουσιάζεται η εικόνα στόχος και 2 εικόνες του σετ πριν την εγγραφή τους σε αυτή και στο κάτω μέρος του σχήματος, η εικόνα στόχος και οι αντίστοιχες 2 εικόνες του σετ μετά την εγγραφή.



Σχήμα 7.4: Τομές εικόνας πριν και μετά το registration

7.4 Επιλογή Χαρακτηριστικών

Για την περιγραφή του σετ ατλάντων εξετάστηκαν όλα τα χαρακτηριστικά από τις διάφορες κατηγορίες που αναπτύχθηκαν στην ενότητα 5, όπως αναφέρθηκαν και στην ενότητα 6.4. Με στόχο τη μείωση της διαστασιμότητας του μοντέλου και την επιλογή των χαρακτηριστικών που περιέχουν την περισσότερη πληροφορία, εφαρμόστηκε μία παραλλαγή της διαδικασίας που περιγράφηκε στην ενότητα 6.4. Πιο συγκεκριμένα, διεξήχθη το ίδιο πείραμα όμως δόθηκε έμφαση στην διακύμανση των τιμών των μετρικών που αφορούν την ακρίβεια πρόβλεψης των κλάσεων ενδιαφέροντος, δηλαδή των κλάσεων 2 και 4 που αφορούν τους χόνδρους μεταξύ των οστών. Όπως παρουσιάζεται στα σχήματα (7.5-7.7) η βέλτιστη ικανότητα πρόβλεψης επί των χόνδρων επιτυγχάνεται για αριθμό χαρακτηριστικών ίσο με 22. Τα επιλεγόμενα χαρακτηριστικά είναι τα εξής:

1. Φωτεινότητα
2. Συντεταγμένη X
3. Συντεταγμένη Y
4. Συντεταγμένη Z
5. Ευκλείδια απόσταση voxel από το επίπεδο $z=0$
6. Ευκλείδια απόσταση voxel από το επίπεδο $z=z_{max}$
7. LBP
8. $Gradient_x$ (HOG)
9. $Gradient_y$ (HOG)
10. $Gradient_z$ (HOG)
11. $Gradient_{xy}$ (HOG)
12. $Gradient_{xz}$ (HOG)
13. $Gradient_{yz}$ (HOG)
14. $Gradient_{xyz}$ (HOG)
15. Magnitude(HOG)
16. Orientation(HOG)
17. Ευκλείδια απόσταση voxel από την γωνία του χώρου $(x_{max}, 0, 0)$
18. Ευκλείδια απόσταση voxel από την γωνία του χώρου $(x_{max}, y_{max}, 0)$
19. Ευκλείδια απόσταση voxel από την γωνία του χώρου $(x_{max}, 0, z_{max})$
20. Ευκλείδια απόσταση voxel από την γωνία του χώρου $(x_{max}, y_{max}, z_{max})$
21. High Gray Level Emphasis(GLRLM)
22. Short Run High Gray Level Emphasis(GLRLM)

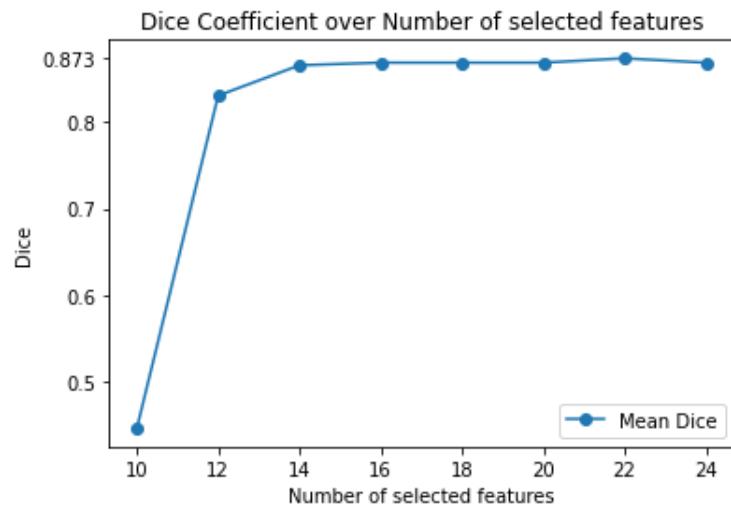


Figure 7.5: Η μετρική Dice για μεταβλητό αριθμό χαρακτηριστικών

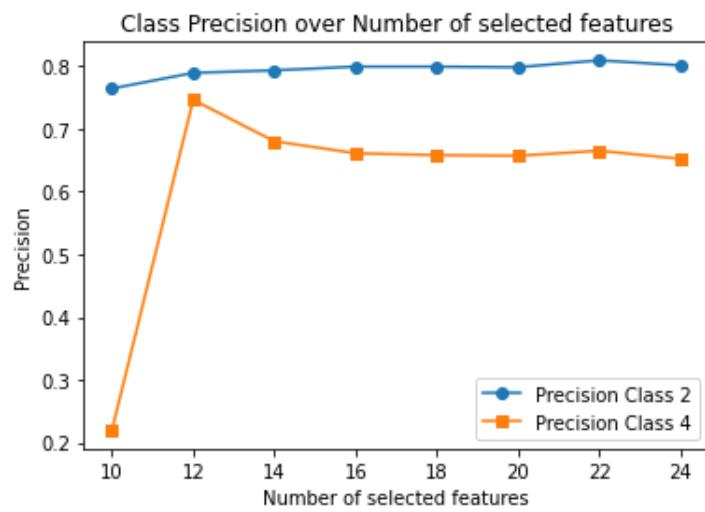


Figure 7.6: Η μετρική Precision για μεταβλητό αριθμό χαρακτηριστικών

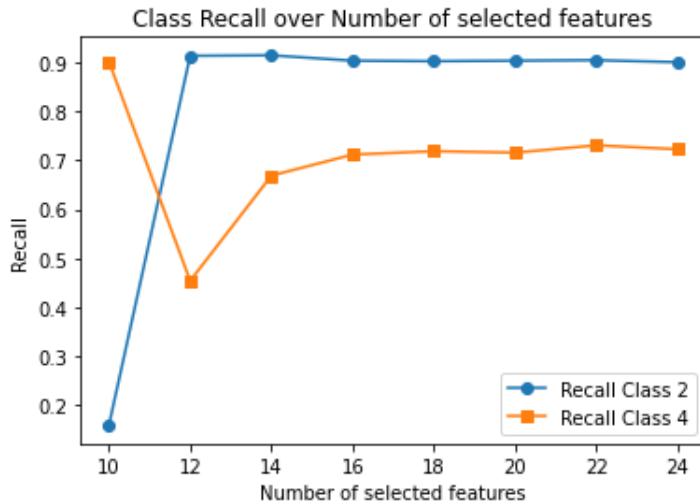


Figure 7.7: Η μετρική Recall για μεταβλητό αριθμό χαρακτηριστικών

7.5 Εκπαίδευση μοντέλου

Αντίστοιχα με την διαδικασία της εκπαίδευσης του μοντέλου πρόβλεψης σε εικόνες εγκεφάλων με το πέρας της διαδικασίας της προεπεξεργασίας των εικόνων και της εξαγωγής χαρακτηριστικών που να τις περιγράφουν ακολούθησε η εκπαίδευση του προτεινόμενου μοντέλου πρόβλεψης των γονάτων. Ορίζοντας μια εικόνα του σετ ως εικόνα προς κατάτμηση και τις υπόλοιπες εικόνες του σετ ως άτλαντες εκπαίδευσης εφαρμόστηκαν τα εξής βήματα.

- Επιλογή των N πιο όμοιων άτλαντων με την εικόνα προς κατάτμηση μέσω της μετρικής MAE, ώστε οι αδύναμοι SVM ταξινομητές να εκπαιδευτούν πάνω σε αυτές.
- Εφαρμογή του αλγόριθμου 2 της ενότητας 4.2.
- Αξιολόγηση των αποτελέσματων και διορθωση των παραμέτρων του μοντέλου.

-
- Εκπαίδευση του μοντέλου με τις νέες παραμέτρους μέχρι να φτάσει σε ικανοποιητικό αποτέλεσμα.

Η παραμετροποίηση του μοντέλου αφορά τα παρακάτω μεγέθη:

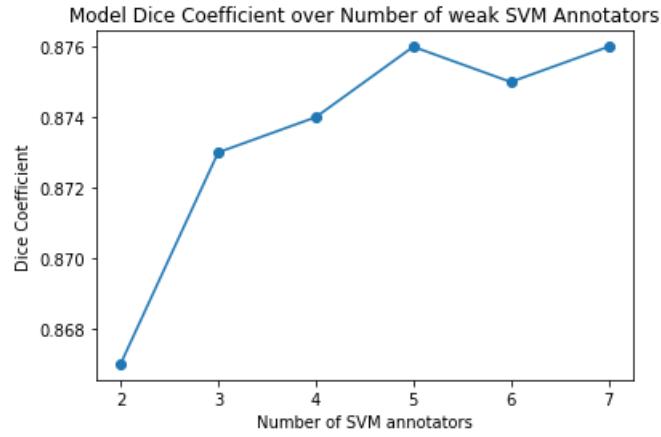
- Χρήση βαρών για την εξισορρόπηση των κλάσεων των κατηγοριοποιημένων δεδομένων που δειγματολειπτεί ο κάθε SVM ταξινομητής από έναν άτλαντα εκπαίδευσης.
- Υπερπαράμετροι C των SVM ταξινομητών του μοντέλου.
- Αριθμός SVM ταξινομητών.

Για την εκπαίδευση κάθε αδύναμου ταξινομητή εξασφαλίζουμε πως το σετ δεδομένων πάνω στο οποίο προπονείται περιλαμβάνει ιστότιμα ως ένα βαθμό, τόσο δεδομένα των κλάσεων που χυριαρχούν στην εικόνα όσο και δεδομένα των κλάσεων που υποεκπροσωπούνται. Κατά αυτόν τον τρόπο, το σετ δεδομένων που δίνουμε ως είσοδο στους ταξινομητές παραμένει ένα bootstrap κλάσμα του συνολικού σετ χωρίς να παραγνωρίζει την ύπαρξη των κλάσεων που αφορούν τους χόνδρους.

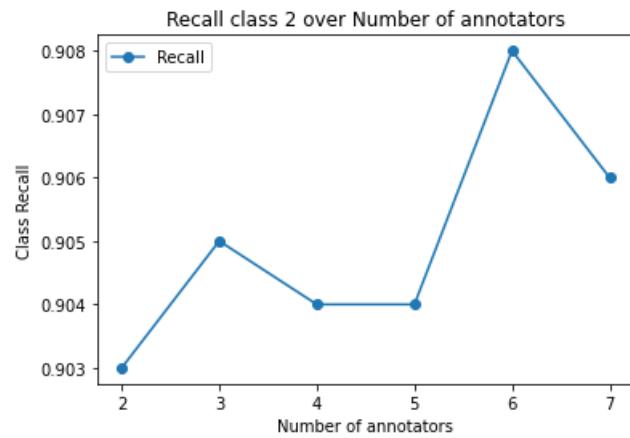
Για την εκπαίδευση των αδύναμων SVM ταξινομητών και για την επίλυση του προβλήματος (4.10) χρησιμοποιήθηκε ο γραμμικός SVM αλγόριθμος LIBLINEAR ο οποίος χρησιμοποιεί γραμμικό πυρήνα, πράγμα που τον καθιστά αρκετά ταχύτερο από άλλους SVM ταξινομητές.

Μετά από διαδικασίες επαναλαμβανόμενων ρυθμίσεων η βέλτιστη υπερπαράμετρος για τον ταξινομητή των αδύναμων SVM ταξινομητών επιλέχθηκε η $c=1$, και για τον επιλυτή του προβλήματος (4.10) επιλέχθηκε η $c=2$.

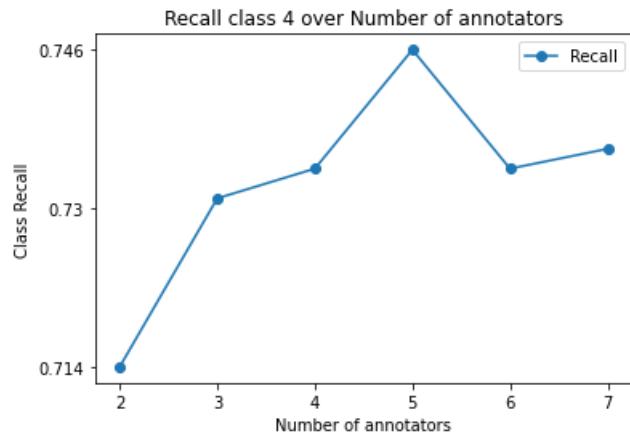
Για την εύρεση του βέλτιστου αριθμού ταξινομητών ακολουθήθηκε η ίδια διαδικασία με της ενότητας 6.5. Αντίστοιχα λοιπόν, διεξήχθη πείραμα για μεταβαλλόμενο αριθμό w αδύναμων SVM ταξινομητών. Στο διάγραμμα του Σχήματος (7.8) φαίνεται ο μέσος όρος της μετρικής Dice για τις διάφορες τιμές του w. Παρατηρούμε πως το μοντέλο παρουσιάζει βέλτιστη απόδοση για $w=5$ όπως μαρτυρούν οι τιμές των μετρικών Recall, Precision επί των κλάσεων 2 και 4 αλλά και η συνολική μετρική Dice. Από τα σχήματα (7.8-12) παρατηρείται με την αύξηση του αριθμού των ταξινομητών μια γενικότερη αυξητική τάση στις τιμές των μετρικών των κλάσεων ενδιαφέροντος. Ενδεχόμενη κλιμάκωση του αριθμού των ταξινομητών πιθανώς θα συνεπάγονταν βελτίωση της απόδοσης του μοντέλου. Παράλληλα, η επέκταση του συνολικού σετ εκπαίδευσης μπορεί να συμβάλλει στην αύξηση της απόδοσης του μοντέλου.



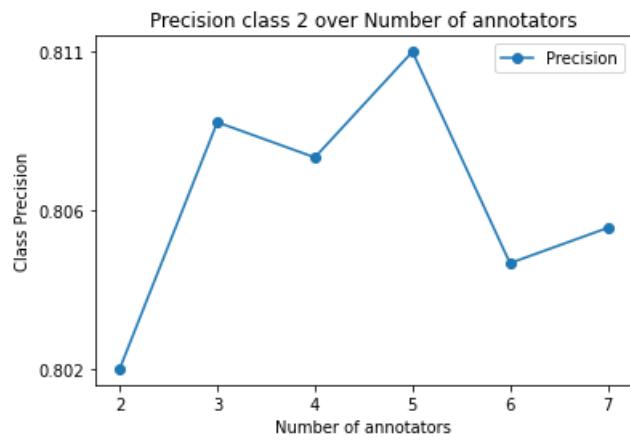
Σχήμα 7.8: Η μετρική Dice για διάφορες τιμές αδύναμων SVM ταξινομητών.



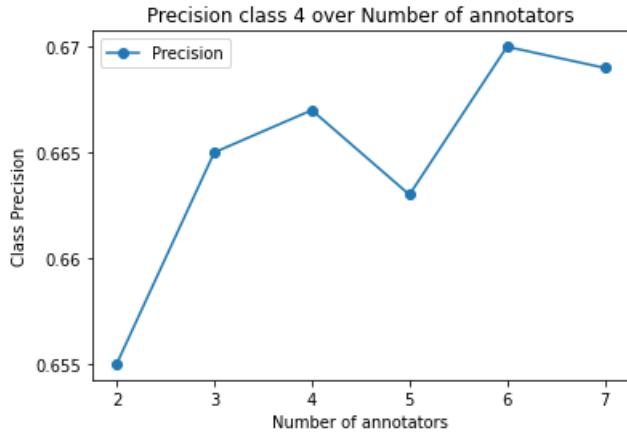
Σχήμα 7.9: Η μετρική Recall επί της κλάσης 2 για διάφορες τιμές αδύναμων SVM ταξινομητών.



Σχήμα 7.10: Η μετρική Recall επί της κλάσης 4 για διάφορες τιμές αδύναμων SVM ταξινομητών.



Σχήμα 7.11: Η μετρική Precision επί της κλάσης 2 για διάφορες τιμές αδύναμων SVM ταξινομητών.



Σχήμα 7.12: Η μετρική Precision επί της κλάσης 4 για διάφορες τιμές αδύναμων SVM ταξινομητών.

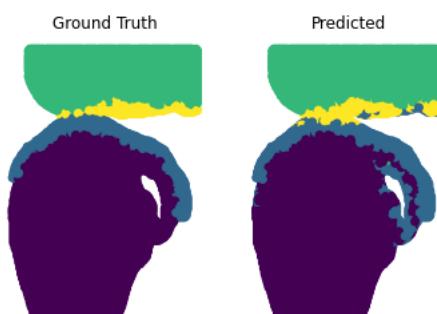
7.6 Αποτελέσματα

Ως τελική αξιολόγηση του μοντέλου πραγματοποιήθηκε σε σετ 25 εικόνων η τεχνική k-fold Cross Validation για $k=1$. Για 25 επαναλήψεις θεωρούμε κάθε φορά μια εικόνα προς κατάτμηση από το σετ και τις υπόλοιπες ως σετ εκπαίδευσης. Με μία ανάγνωση του πίνακα 4 όπου φάνονται οι μετρικές απόδοσης του μοντέλου για $w=5$, συμπεραίνεται πως ως επί το πλείστον το μοντέλο ταξινομεί αρκετά καλά τις κλάσεις των οστών, ενώ παρουσιάζει περιθώριο βελτίωσης στις κλάσεις των χόνδρων που γειτνιάζουν. Το παραπάνω αποτέλεσμα μπορεί να αποδοθεί στο γεγονός ότι το επιλεγόμενο διάνυσμα χαρακακτηριστικών αποτελεί μία επαρκή βάση για την πρόβλεψη των κλάσεων των χόνδρων αλλά όχι απαραίτητη βέλτιστη για να αποδώσει πλήρως την γεωγραφία των χόνδρων. Παράλληλα, το δεδομένο σετ των γονάτων παρουσιάζει σημαντικότατη υποεκπρόσωπηση στις κλάσεις ενδιαφέροντος η οποία δεν είναι δυνατό να αντιμετωπιστεί στο στάδιο δειγματοληψίας των δεδομένων που χρησιμοποιούνται στο δεύτερο επίπεδο εκπαίδευσης του μοντέλου (label aggregation), καθότι αυτά είναι αταξινόμητα. Επομένως δεν μπορεί να εφαρμοστεί κάποια διαδικασία εξισορρόπησης του μίγματος όλων των κλάσεων, όπως στο στάδιο της δειγματοληψίας για την εκπαίδευση των αδύναμων ταξινομητών, με συνέπεια το μοντέλο να απολλύει την δυνατότητα εύρεσης του ιδανικού διανύσματος βαρών για τις προβλέψεις

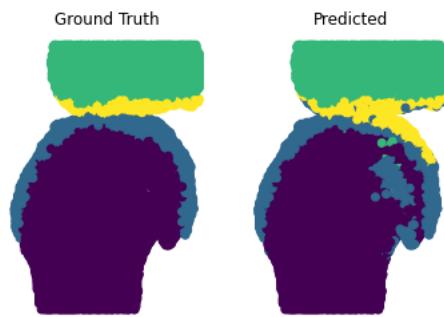
των ταξινομητών. Στα σχήματα (7.13-7.19) φαίνεται η πρόβλεψη του μοντέλου για δεδομένη εικόνα γονάτου. Με πράσινο απεικονίζεται η κλάση 1 (μηριαίο οστό), με κίτρινο η κλάση 2 (μηριαίος χόνδρος), με μωβ η κλάση 3 (κνημιαίο οστό) και με μπλε η κλάση 4 (κνημιαίος χόνδρος).

Dice Coefficient	0.87
Accuracy	0.96
F1-Score	0.96
Precision	0.96
Recall	0.96
Precision Class 1	0.97
Precision Class 2	0.90
Precision Class 3	0.98
Precision Class 4	0.74
Recall Class 1	0.99
Recall Class 2	0.81
Recall Class 3	0.98
Recall Class 4	0.66

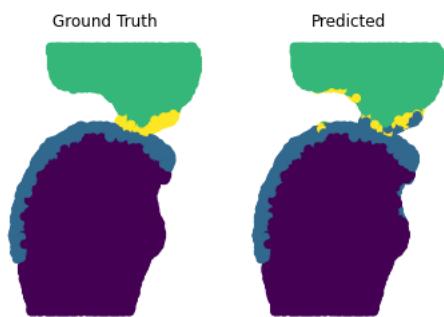
Πίνακας 5: Μετρικές απόδοσης μοντέλου για $W=5$



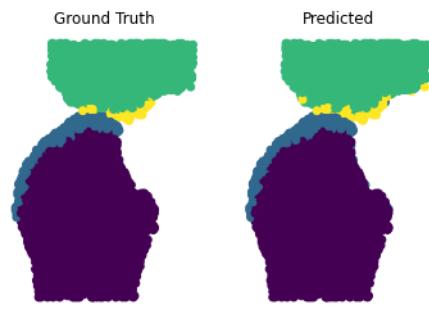
Σχήμα 7.13: Dice=0.89 Αριστερά η εικόνα προς κατάτυπη και δεξιά η πρόβλεψη του μοντέλου



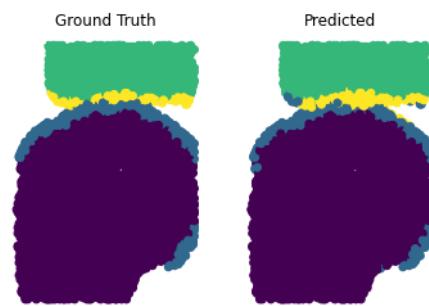
Σχήμα 7.14: Dice=0.88 Αριστερά η εικόνα προς κατάτυπηση και δεξιά η πρόβλεψη του μοντέλου



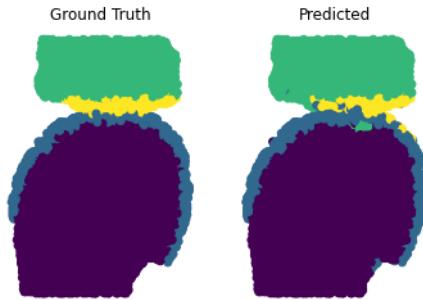
Σχήμα 7.15: Dice=0.90 Αριστερά η εικόνα προς κατάτυπηση και δεξιά η πρόβλεψη του μοντέλου



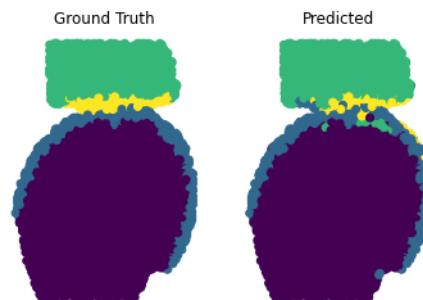
Σχήμα 7.16: Dice=0.90 Αριστερά η εικόνα προς κατάτυπη και δεξιά η πρόβλεψη του μοντέλου



Σχήμα 7.17: Dice=0.88 Αριστερά η εικόνα προς κατάτυπη και δεξιά η πρόβλεψη του μοντέλου



Σχήμα 7.18: Dice=0.88 Αριστερά η εικόνα προς κατάτυπη και δεξιά η πρόβλεψη του μοντέλου



Σχήμα 7.19: Dice=0.87 Αριστερά η εικόνα προς κατάτυπη και δεξιά η πρόβλεψη του μοντέλου

8 Μελλοντικές Επεκτάσεις

Δεδομένης της σχετικής επιτυχίας του μοντέλου στην πρόβλεψη επί των δύο σετ δεδομένων, εμφανίζεται σημαντικό περιθώριο επιπλέον βελτίωσης με την εφαρμογή εναλλακτικών προσεγγίσεων και τροποποιήσεων σε αρκετά στάδια της διαδικασίας.

Αναλυτικότερα, στο στάδιο της εξαγωγής των χαρακτηριστικών από κάθε εικόνα, περιθώρια εντοπίζονται στην αύξηση του αριθμού και της ποικιλομορφίας των χαρακτηριστικών που χρησιμοποιούνται για την περιγραφή και επομένως και την κατάτμηση τους. Συνάρτηση της αύξησης των χαρακτηριστικών και των δοκιμών που πραγματοποιούνται για την εξακρίβωση της αξίας τους αποτελεί η διαθεσιμότητα σημαντικών υπολογιστικών πόρων, καθώς η διαδικασία της εξαγωγής είναι η πλέον χρονοβόρα. Πρόσθετα, μένει να διερευνηθεί η συμβολή τεχνικών μείωσης διαστασιμότητας όπως η PCA.

Στο πλαίσιο της επιλογής των ομοίων ατλάντων, έχει εφαρμοστεί μία σχετικά απλοική μετρική (MAE), δίνοντας ικανοποιητικά αλλά όχι απαραίτητα βέλτιστα αποτελέσματα. Η διερευνησή εναλλακτικών και σύνθετων μετρικών αλλά και ο συνδυαστική χρήση τους ενδεχομένως να απέδιδε στην επιλογή των ομοιότερων ατλάντων για εκπαίδευση. Παράλληλα, σημαντικά αποτελέσματα πιθανώς να επέφερε ένα διευρυμένο σετ εικόνων με ενισχυμένη ποικιλομορφία. Με δεδομένες τις παραπάνω συνθήκες και μία εξακριβώμενα αποδοτική μετρική ομοιότητας, το μοντέλο θα μπορούσε να έχει μεταβλητό αριθμό ταξινομητών για κάθε εικόνα προς κατάτμηση ο οποίος θα καθορίζεται από τις τιμές ομοιότητας των εικόνων του σετ εκπαίδευσης που υπερβαίνουν ένα ορισμένο κατώφλι.

Στο στάδιο της εκπαίδευσης του μοντέλου μπορούν να γίνουν σημαντικές διαφοροποιήσεις και πειραματισμοί. Χαρακτηριστικά, μπορεί αντί για γραμμικό ταξινομητή SVM-Liblinear να χρησιμοποιηθεί ο ταξινομητής LibSVM που εφαρμόζει μη-γραμμικούς πηρύνες, με σημαντικά αυξημένο υπολογιστικό κόστος αλλά πιθανώς και βελτίωση της προβλεπτικής ικανότητας.

Σαν τελευταία πρόταση προς υλοποίηση, θα μπορούσε να εφαρμοστεί τμηματική κατάτμηση της εικόνας διευκολύνοντας τον αλγόριθμο να διαχρίνει τις τοπικές μορφολογίες και ιδιεταιρότητες των εικόνων.

9 Βιβλιογραφία

References

- [1] Mahesh, B. (2020). Machine Learning Algorithms - A Review. International Journal of Science and Research, 9(1), 1054-1058. Retrieved from <https://www.ijsr.net/archive/v9i1/ART20204233.pdf>.
- [2] Kotsiantis, S. B. (2007). Supervised Machine Learning: A Review of Classification Techniques. Informatica, 31(3), 249-268.

-
- [3] Cristianini, N. & Shawe-Taylor, J. (2000). *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press.
 - [4] Lanckriet, G.R.G., Cristianini, N., Bartlett, P., El Ghaoui, L., & Jordan, M.I. (2004). A Statistical Framework for Genomic Data Fusion using Multiple Kernel Learning. *Journal of Machine Learning Research*, 5, 1-25.
 - [5] Rakotomamonjy, A., Bach, F., Canu, S., & Grandvalet, Y. (2008). SimpleMKL. *The Journal of Machine Learning Research*, 9, 2491-2521.
 - [6] Aertsen, M., van Hoof, S., Gonnissen, A., de Senneville, B. D., Öllers, M., Crijns, S. P., & Tijssen, R. H. (2017). Multi-atlas based image registration with automated quality assurance for localisation-based radiotherapy planning of head and neck cancer. *Computerized Medical Imaging and Graphics*, 55, 33-42.
 - [7] Ardekani, B. A., Guckemus, S., Bachman, A., Hoptman, M. J., Wojtaszek, M., & Nierenberg, J. (2005). Quantitative comparison of algorithms for inter-subject registration of 3D volumetric brain MRI scans. *Journal of neuroscience methods*, 142(1), 67-76.
 - [8] Rueckert, D., Sonoda, L. I., Hayes, C., Hill, D. L., Leach, M. O., & Hawkes, D. J. (1999). Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Transactions on Medical Imaging*, 18(8), 712-721.
 - [9] Avants, B. B., Tustison, N. J., Song, G., Cook, P. A., Klein, A., & Gee, J. C. (2011). A Reproducible Evaluation of ANTs Similarity Metric Performance in Brain Image Registration. *NeuroImage*, 54(3), 2033-2044.
 - [10] Rousseau, F., Habas, P. A., & Studholme, C. (2006). Registration-based approach for automatic infant brain segmentation and atlas formation. *Medical Image Analysis*, 10(5), 769-782.
 - [11] Yan, Y., Xu, D., Li, L., & Zhang, H. (2016). Robust Semi-Supervised Learning through Label Aggregation and Interpolation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 1-9). Las Vegas, NV, USA.

-
- [12] Li, L., Zhao, Z., & Chellappa, R. (2011). Convex and Scalable Weakly Labeled SVMs. In Proceedings of the IEEE International Conference on Computer Vision (ICCV) (pp. 2015-2022). Barcelona, Spain.
 - [13] Ortiz, A., Palacio, A. A., Górriz, J. M., Ramírez, J., & Salas-González, D. (2013). Segmentation of Brain MRI Using SOM-FCM-Based Method and 3D Statistical Descriptors. Computational and Mathematical Methods in Medicine, 2013.
 - [14] R. M. Haralick, K. Shanmugam, and I. H. Dinstein, "Textural Features for Image Classification," IEEE Transactions on Systems, Man, and Cybernetics, vol. 3, no. 6, pp. 610-621, Nov. 1973.
 - [15] Ojala, T., Pietikainen, M., & Harwood, D. (1994). Performance Evaluation of Texture Measures with Classification Based on Kullback Discrimination of Distributions. In Proceedings of the 12th International Conference on Pattern Recognition (ICPR), 2, 582-585, Jerusalem, Israel, October 1994.
 - [16] Dalal, N., & Triggs, B. (2005). Histograms of oriented gradients for human detection. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), vol. 1, pp. 886-893, San Diego, CA, USA, June 2005.
 - [17] Yan, L., & Xia, W. (2014). A modified three-dimensional gray-level co-occurrence matrix for image classification with digital surface model. ISPRS International Journal of Geo-Information, 3(4), 1342-1356.
 - [18] Kloke, T., Lin, T. T., Tsai, S. C., & Tzeng, W. G. (2010). Permutation Arrays Under the Chebyshev Distance: Construction and Properties. IEEE Transactions on Information Theory, 56(6), 2611-2617.
 - [19] Xu, D., Kurani, A. S., Furst, J. D., & Raicu, D. S. (2005). Run-length encoding for volumetric texture. International Journal of Biomedical Imaging, 2005(1), 1-8.
 - [20] Ahmadi, M., Mirniaharikandehei, S., & Mahdavi, S. R. (2017). Automatic diagnosis of breast tumors using SVM and GLDM features in mammography images. Journal of X-Ray Science and Technology, 25(1), 55–64.

-
- [21] Lu, Y., Zheng, Y., & Yang, J. (2015). An automatic method of MRI brain image feature extraction based on PCA and ICA. *Journal of Medical Systems*, 39(9), 100.
 - [22] Shyu, C. R., Chen, S. C., & Sarinnapakorn, K. (2003). A novel approach to extracting features from spatial datasets using the fisher linear discriminant. *International Journal of Geographical Information Science*, 17(2), 131-146.
 - [23] Kobak, D., & Berens, P. (2019). The art of using t-SNE for single-cell transcriptomics. *Nature Communications*, 10(1), 1-14.
 - [24] Chandrashekhar, G., & Sahin, F. (2014). A survey on feature selection methods. *Computers & Electrical Engineering*, 40(1), 16-28.
 - [25] Golino, H., & Demetriou, A. (2017). Using recursive feature elimination and support vector machines to identify potential biomarkers of schizophrenia. *Journal of Psychiatric Research*, 85, 140-147.
 - [26] Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., et al. (2004). FMRIB Software Library (FSL). *NeuroImage*, 23(S1), S208–S219.
 - [27] Kim, J., Lee, J., Kim, B. H., & Lee, S. K. (2018). Multi-atlas-based automatic segmentation of the brain using local affine registration and robust label fusion: Application to basal ganglia segmentation. *NeuroImage: Clinical*, 18, 439-449.
 - [28] Lowekamp, B., Chen, D. T., Ibáñez, L., & Blezek, D. (2013). The design of SimpleITK. *Frontiers in Neuroinformatics*, 7, 45.
 - [29] Garyfallidis E, Brett M, Amirbekian B, et al. Dipy, a library for the analysis of diffusion MRI data. *Front Neuroinform*. 2014;8:8.
 - [30] Nyul LG, Udupa JK, Zhang X. New variants of a method of MRI scale standardization. *IEEE Trans Med Imaging*. 2000;19(2):143-150.
 - [31] van Griethuysen, J. J. M., Fedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V., Beets-Tan, R. G. H., Fillion-Robin, J. C., Pieper, S., Aerts, H. J. W. L. (2017). Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Research*, 77(21), e104–e107.