



ΕΘΝΙΚΟ ΚΑΙ ΚΑΠΟΔΙΣΤΡΙΑΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

**ΣΧΟΛΗ ΘΕΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΜΗΜΑ ΠΛΗΡΟΦΟΡΙΚΗΣ ΚΑΙ ΤΗΛΕΠΙΚΟΙΝΩΝΙΩΝ**

**ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
ΕΠΕΞΕΡΓΑΣΙΑ ΣΗΜΑΤΟΣ ΓΙΑ ΕΠΙΚΟΙΝΩΝΙΕΣ ΚΑΙ ΠΟΛΥΜΕΣΑ**

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

**Ανάπτυξη Συστήματος Αυτόματης Ταξινόμησης Όγκων
Εγκεφάλου σε Πλατφόρμα Τηλεπαθολογίας με Βάση
Ιστοπαθολογικές Εικόνες**

Σπυριδούλα Δ. Ξενάκη

Επιβλέποντες: Διονύσιος Κάβουρας, Καθηγητής

ΑΘΗΝΑ

ΝΟΕΜΒΡΙΟΣ 2012

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

Ανάπτυξη Συστήματος Αυτόματης Ταξινόμησης Όγκων Εγκεφάλου σε Πλατφόρμα
Τηλεπαθολογίας με Βάση Ιστοπαθολογικές Εικόνες

Σπυριδούλα Δ. Ξενάκη
Α.Μ.: 1132

ΕΠΙΒΛΕΠΟΝΤΕΣ: Διονύσιος Κάβουρας, Καθηγητής

ΕΞΕΤΑΣΤΙΚΗ ΕΠΙΤΡΟΠΗ: Διονύσιος Κάβουρας, Καθηγητής
Εμμανουήλ Σαγκριώτης, Αναπληρωτής Καθηγητής

Νοέμβριος 2012

ΠΕΡΙΛΗΨΗ

Στα πλαίσια της παρούσας εργασίας αναπτύχθηκε σύστημα τηλεπαθολογίας και υποβοήθησης της διάγνωσης για την βελτίωση της ακρίβειας ταξινόμησης καρκίνων εγκεφάλου σε βαθμούς κακοήθειας. Το κλινικό υλικό που χρησιμοποιήθηκε ήταν δείγματα ιστοπαθολογικού ιστού, που συλλέχτηκαν από 35 ασθενείς από το Πανεπιστημιακό Νοσοκομείο της Πάτρας, οι οποίοι είχαν διαγνωστεί με αστροκυτταρικό όγκο. Πραγματοποιήθηκε προετοιμασία των δεδομένων κατά την οποία τα δείγματα ιστού χρωματίστηκαν με H&E (Hematoxylin & Eosin), ώστε να φανερωθούν τα διάφορα συστατικά των κυττάρων και των πυρήνων αλλά και για να διακριθούν μεταξύ τους. Στη συνέχεια, ένας ιστοπαθολόγος εξέτασε τις χρωματισμένες εικόνες των δειγμάτων ιστού και πραγματοποίησε διάγνωση με βάση ιστολογικά κριτήρια σε τρεις βαθμούς κακοήθειας (I, II, III ή IV) σύμφωνα με τον Παγκόσμιο Οργανισμό Υγείας (World Health Organization - WHO). Ταυτόχρονα, ο ίδιος ιστοπαθολόγος σημείωσε πάνω στις πλάκες των δειγμάτων την πιο αντιπροσωπευτική περιοχή (Region Of Interest - ROI) για την εξέταση αυτών. Έτσι, ακολούθησε η ψηφιοποίηση της εικόνας, όπου από κάθε δείγμα ψηφιοποιήθηκαν εικόνες, η οποίες πάρθηκαν από το προκαθορισμένο ROI και εφαρμόστηκε τμηματοποίηση πυρήνων. Το σημαντικότερο στοιχείο μιας τέτοιας εικόνας, από το οποίο μπορούν να εξαχθούν συμπεράσματα για την διάγνωση, είναι οι πυρήνες. Επομένως, δημιουργήθηκαν αλγόριθμοι, οι οποίοι απομάκρυναν όλες τις υπόλοιπες περιοχές της εικόνας και τμηματοποίησαν τους πυρήνες για περισσότερη ανάλυση αυτών. Από τα αποτελέσματα της τμηματοποίησης των εικόνων προέκυψε ότι για τον υπολογισμό των χαρακτηριστικών της παρούσας εργασίας απομονώθηκαν και μελετήθηκαν συνολικά κατά μέσο όρο 588 τμηματοποιημένοι πυρήνες για κάθε δείγμα-ασθενή, ενόσω έχει αποδειχθεί ότι ακόμη και 200 ορθά τμηματοποιημένοι πυρήνες είναι επαρκείς για την εξαγωγή χαρακτηριστικών. Κατόπιν, εξήχθησαν χαρακτηριστικά μορφολογίας, υψής και αρχιτεκτονικής από τους τμηματοποιημένους πυρήνες για να περιγράψουν τον βαθμό κακοήθειας του κάθε δείγματος-ασθενή. Τα χαρακτηριστικά αυτά αποτέλεσαν την είσοδο σε ένα σύστημα αναγνώρισης προτύπων που σχεδιάστηκε, έτσι ώστε να προβλέπει την επικινδυνότητα του κάθε όγκου. Το σύστημα αυτό δομείται με αλγορίθμους supervised, semi-supervised και unsupervised. Ο SVM Supervised Classifier (Polynomial ή Quadratic kernel) αποτέλεσε μια λύση στην αυτοματοποιημένη ταξινόμηση αστροκυττωμάτων, καθώς έδωσε ποσοστό ακρίβειας 94.29% στην ταξινόμηση των δεδομένων. Οι unsupervised clustering k-Means και Fuzzy έδωσαν ποσοστό ακρίβειας 74.29% και ο Semi-Supervised with Co-Training ταξινομητής έδωσε μέγιστο ποσοστό ακρίβειας 88.57%. Τέλος, τα χαρακτηριστικά Mean (χαρακτηριστικό υψής 1^{ης} τάξης), mean Correlation, range Correlation, mean Gray Level Non-Uniformity και mean Run Length Non-Uniformity (χαρακτηριστικά υψής 2^{ης} τάξης) αναδείχθηκαν τα χαρακτηριστικά εκείνα με το μεγαλύτερο ποσοστό πιθανότητας εμφάνισης στο διάνυσμα χαρακτηριστικών εκείνο, το οποίο δίνει το μέγιστο (καλύτερο) ποσοστό ακρίβειας ενός ταξινομητή.

ΘΕΜΑΤΙΚΗ ΠΕΡΙΟΧΗ: Επεξεργασία Εικόνας, Αναγνώριση Προτύπων

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ: μικροσκόπιο, αστροκύττωμα, supervised / unsupervised / semi-supervised αλγόριθμοι, τμηματοποίηση εικόνας, τηλεπαθολογία

ABSTRACT

In this thesis, a telepathology system has been developed which assists in improving diagnostic accuracy of brain cancer classification into grades of malignancy. The clinical material was histopathological tissue samples, which were collected from 35 patients from the University Hospital of Patra, who were diagnosed with astrocytic tumor. During data preparation the tissue samples were stained with H & E (Hematoxylin & Eosin), in order to expose the various components of the cells and nuclei. Then, a histopathologist examined the stained histopathological images of tissue samples and performed diagnosis based on histological criteria of malignancy in three grades (I, II, III or IV) according to the World Health Organization (WHO). At the same time, the histopathologist noted on the samples the most representative region (Region of Interest - ROI) for the examination. Thus, each sample image was digitized, taken from the predefined ROI and applied nuclei segmentation. The most important element of such a picture, from which conclusions can be drawn for the diagnosis, is the nucleus. Therefore, algorithms were created, which removed all remaining areas of the image and segmented nuclei for further analysis of these. From the results of image segmentation revealed that overall average 588 segmented nuclei were isolated and studied for each patient sample, while it has been shown that even 200 correctly segmented nuclei are sufficient to feature extraction. Then morphological characteristics, texture and architecture of their segmented nuclei were exported to describe the degree of malignancy of each sample-patient relationship. These characteristics were entering a pattern recognition system designed to provide the dangerousness of each tumor. This system was constructed by supervised, semi-supervised and unsupervised algorithms. The SVM Supervised Classifier (Polynomial or Quadratic kernel) has been proved a solution in the automated classification of astrocytomas as it had an accuracy rate of 94.29% in classifying data. The unsupervised clustering k-Means and Fuzzy gave an accuracy rate of 74.29% and Semi-Supervised with Co-Training classifier gave maximum accuracy rate of 88.57%. Finally, the characteristics Mean (first class texture feature), mean Correlation, range Correlation, mean Gray Level Non-Uniformity and mean Run Length Non-Uniformity (second class textural) highlighted as those features with the highest percentage likelihood to appear in feature vector that gives the maximum (best) accuracy rate of a classifier.

SUBJECT AREA: Image Processing, Pattern Recognition

KEYWORDS: microscope, astrocytoma, supervised / unsupervised / semi-supervised algorithms, image segmentation, telepathology

*Στους γονείς μου Δημήτρη και Ειρήνη,
για την αμέριστη υποστήριξη και συμπαράσταση*

ΕΥΧΑΡΙΣΤΙΕΣ

Από τη θέση αυτή θα ήθελα να ευχαριστήσω θερμά έναν αριθμό ανθρώπων για τη συμβολή και την καθοδήγησή τους στην εργασία αυτή. Αρχικά θα ήθελα να ευχαριστήσω τον επιβλέπων της διπλωματικής μου εργασίας, κ. Διονύσιο Κάβουρα, καθηγητή του τμήματος Τεχνολογίας Ιατρικών Οργάνων του ΤΕΙ Αθηνών, που μου επέτρεψε να διεξάγω τη μελέτη μου στα πλαίσια του εργαστηρίου του, καθώς και για την πολύτιμη βοήθεια και καθοδήγησή του κατά τη διάρκεια της δουλειάς μου.

Ολόθερμες ευχαριστίες οφείλω στον Δημήτριο Γκλώτσο, καθηγητή εφαρμογών του τμήματος Τεχνολογίας Ιατρικών Οργάνων του ΤΕΙ Αθηνών, αλλά και στον Σπύρο Κωστόπουλο, μεταδιδακτορικού ερευνητή, τόσο για την συνδρομή τους στο γνωστικό πεδίο της διπλωματικής μου εργασίας, όσο και για την κατανόηση και την υποστήριξή τους.

Για την έμπρακτη και ουσιαστική υποστήριξή τους ευχαριστώ την οικογένειά μου και τους φίλους μου, που στάθηκαν δίπλα μου και με στήριξαν καθ' όλη τη διάρκεια της προσπάθειάς μου, ειδικά σε περιόδους μεγάλης πίεσης.

ΠΕΡΙΕΧΟΜΕΝΑ

ΕΙΣΑΓΩΓΗ	13
1.1 Κλινικό Πρόβλημα – Όγκοι εγκεφάλου	14
1.1.1 Γενικά	14
1.1.2 Αιτιολογία	14
1.1.3 Συμπτώματα	15
1.1.4 Τύποι.....	16
1.1.5 Παθολογική ανατομία γλοιωμάτων	16
1.2 Αστροκυττώματα	17
1.2.1 Γενικά	17
1.2.2 Διάγνωση	18
1.2.3 Ταξινόμηση σε βαθμίδες (Grades)	19
1.2.4 Θεραπεία	20
1.2.5 Επανεμφάνιση και εξέλιξη.....	21
2. ΟΠΤΙΚΟ ΜΙΚΡΟΣΚΟΠΙΟ	22
2.1 Βασικές Αρχές Λειτουργίας.....	22
2.2 Τμήματα και όργανα του μικροσκοπίου	25
3. ΥΛΙΚΟ ΚΑΙ ΜΕΘΟΔΟΙ	27
3.1 Κλινικό υλικό	27
3.2 Αναγνώριση προτύπων	27
3.3 Τμηματοποίηση εικόνας	27
3.4 Εξαγωγή χαρακτηριστικών	28
3.4.1 Χαρακτηριστικά 1 ^{ης} τάξης	28
3.4.2 Χαρακτηριστικά 2 ^{ης} τάξης	29
3.4.3 Μορφολογικά χαρακτηριστικά.....	33
3.5 Επιλογή χαρακτηριστικών	34
3.5.1 Ταξινόμηση προτύπων - Supervised αλγόριθμοι	35
3.5.2 Ομαδοποίηση προτύπων - Unsupervised αλγόριθμοι	40
3.5.3 Semi-supervised αλγόριθμοι	41
4. ΑΠΟΤΕΛΕΣΜΑΤΑ	43
4.1 Γενικά	43
4.2 Κλινικό υλικό	45
4.3 Επεξεργασία και τμηματοποίηση εικόνας	45
4.3.1 Ανασκόπηση της βιβλιογραφίας	45
4.3.2 Μέθοδος τμηματοποίησης εικόνας.....	46
4.3.3 Σύνοψη και συμπεράσματα.....	51
4.4 Αναγνώριση προτύπων	51
4.4.1 Ανασκόπηση της βιβλιογραφίας	52

4.4.2	Αποτελέσματα supervised ταξινομητών.....	53
4.4.3	Αποτελέσματα unsupervised ταξινομητών	60
4.4.4	Αποτελέσματα semi-supervised ταξινομητών.....	63
4.4.5	Σύνοψη και συμπεράσματα.....	64
5.	ΤΗΛΕΠΑΘΟΛΟΓΙΑ	66
5.1	Γενικά	66
5.2	Σύστημα εικονικού μικροσκοπίου (Virtual microscope system)	66
6.	ΣΥΝΟΨΗ ΚΑΙ ΣΥΜΠΕΡΑΣΜΑΤΑ.....	70
6.1	Σύνοψη και συμπεράσματα.....	70
6.2	Περιορισμοί της έρευνας	71
7.	ΜΕΛΛΟΝΤΙΚΕΣ ΠΡΟΟΠΤΙΚΕΣ.....	72
	ΠΙΝΑΚΑΣ ΟΡΟΛΟΓΙΑΣ	73
	ΣΥΝΤΜΗΣΕΙΣ – ΑΡΚΤΙΚΟΛΕΞΑ – ΑΚΡΩΝΥΜΙΑ	74
	ΑΝΑΦΟΡΕΣ.....	75

ΚΑΤΑΛΟΓΟΣ ΣΧΗΜΑΤΩΝ

Σχήμα 1: Διάγραμμα απεικόνισης του καλύτερου ποσοστού επιτυχίας του MDC για κάθε πλήθος χαρακτηριστικών.....	53
Σχήμα 2: Διάγραμμα απεικόνισης του καλύτερου ποσοστού επιτυχίας του LSMD για κάθε πλήθος χαρακτηριστικών	54
Σχήμα 3: Διάγραμμα απεικόνισης του καλύτερου ποσοστού επιτυχίας του Bayes Linear και Bayes Quadratic για κάθε πλήθος χαρακτηριστικών.....	55
Σχήμα 4: Διάγραμμα απεικόνισης του καλύτερου ποσοστού επιτυχίας του k-NN για κάθε πλήθος χαρακτηριστικών για k=3 και k=7.....	56
Σχήμα 5: Διάγραμμα απεικόνισης του καλύτερου ποσοστού επιτυχίας του PNN Gaussian για κάθε πλήθος χαρακτηριστικών για $\sigma=0.5$, $\sigma=0.2$ και $\sigma=5$	57
Σχήμα 6: Διάγραμμα απεικόνισης του καλύτερου ποσοστού επιτυχίας του PNN Reciprocal για κάθε πλήθος χαρακτηριστικών για $\sigma=0.5$ και $\sigma=5$	57
Σχήμα 7: Διάγραμμα απεικόνισης του καλύτερου ποσοστού επιτυχίας του PNN Exponential για κάθε πλήθος χαρακτηριστικών για $\sigma=0.5$ και $\sigma=5$	58
Σχήμα 8: Διάγραμμα απεικόνισης των καλύτερων χαρακτηριστικών που έδωσαν το μέγιστο ποσοστό επιτυχίας για τον PNN Exponential Classifier για $\sigma=5$ (85.71% - συνδυασμός χαρακτηριστικών ανά 3).....	59
Σχήμα 9: Διάγραμμα απεικόνισης του καλύτερου ποσοστού επιτυχίας του SVM Gaussian, Polynomial και Quadratic για κάθε πλήθος χαρακτηριστικών.....	60
Σχήμα 10: Διάγραμμα απεικόνισης του καλύτερου ποσοστού επιτυχίας του k-Means Clustering για κάθε πλήθος χαρακτηριστικών.....	61
Σχήμα 11: Διάγραμμα απεικόνισης των καλύτερων χαρακτηριστικών που έδωσαν το μέγιστο ποσοστό επιτυχίας για τον k-Means Clustering (74.29% - συνδυασμός χαρακτηριστικών ανά 3).....	62
Σχήμα 12: Διάγραμμα απεικόνισης του καλύτερου ποσοστού επιτυχίας του k-Means Clustering για κάθε πλήθος χαρακτηριστικών.....	63
Σχήμα 13: Διάγραμμα απεικόνισης του καλύτερου ποσοστού επιτυχίας του k-NN with Co-Training για κάθε πλήθος χαρακτηριστικών	64

ΚΑΤΑΛΟΓΟΣ ΕΙΚΟΝΩΝ

Εικόνα 1: Οι διαφορετικές δομές του εγκεφάλου	14
Εικόνα 2: Οι περιοχές του εγκεφάλου.....	16
Εικόνα 3: Σχηματική αναπαράσταση της πορείας των ακτινών στο οπτικό μικροσκόπιο[14].....	22
Εικόνα 4: Σχηματική αναπαράσταση της γωνίας ευκρινούς οράσεως	23
Εικόνα 5: Δύο σημεία του δείγματος (α) στο διακριτικό τους όριο και (β) όταν τα κεντρικά τους σημεία επικαλύπτονται[14]	25
Εικόνα 6: Τμήματα του μικροσκοπίου: (1) Προσοφθάλμιοι φακοί, (2) Αντικειμενοφόρος τράπεζα, (3) Συμπυκνωτής, (4) Ίρις, (5) Αντικειμενικός φακός, (6) Φωτεινή πηγή[17] ..	25
Εικόνα 7: Παράδειγμα υπολογισμού co-occurrence μήτρας: (α) η εικόνα, (β) η θέση για κάθε ζευγάρι στην co-occurrence μήτρα, (γ) η co-occurrence μήτρα για κάθε φ	30
Εικόνα 8: Παράδειγμα υπολογισμού run-length μήτρας: (α) η εικόνα, (β) η run-length μήτρα σχηματικά, (γ) η run-length μήτρα για κάθε γωνία φ	32
Εικόνα 9: Μορφολογικά χαρακτηριστικά πυρήνων	33
Εικόνα 10: Δύο διαχωριστικές ευθείες σε δισδιάστατο χώρο. Εικονίζεται το διαχωριστικό υπερεπίπεδο που δίνει το μέγιστο δυνατό περιθώριο[31]	38
Εικόνα 11: Συλλογή και προετοιμασία δεδομένων.....	44
Εικόνα 12: Αυτόματη Διάγνωση.....	44
Εικόνα 13: Ψηφιοποιημένη εικόνα μετά από χρώση H&E	47
Εικόνα 14: Ασπρόμαυρη ψηφιοποιημένη εικόνα μετά από χρώση H&E	47
Εικόνα 15: Τμηματοποιημένη εικόνα μετά από εφαρμογή του k-means color based clustering, που απομονώνει τους πυρήνες.....	48
Εικόνα 16: Μάσκα πυρήνων μετά από αυτόματη επεξεργασία	49
Εικόνα 17: Χειροκίνητη επιλογή πυρήνων για αφαίρεση από την μάσκα, λόγω λανθασμένης τμηματοποίησης αυτών	49
Εικόνα 18: Τελική τμηματοποιημένη εικόνα πυρήνων	50
Εικόνα 19: Η τελική μάσκα πυρήνων εφαρμοσμένη στην αρχική εικόνα	50
Εικόνα 20: Παράδειγμα 6 εικόνων που δέχεται ως είσοδο το σύστημα	67

Εικόνα 21: Το τελικό μωσαϊκό που δημιουργείται από τις 6 εικόνες εισόδου.....	68
Εικόνα 22: Ο χρήστης επιλέγει εικόνα από το μωσαϊκό.....	68
Εικόνα 23: Ο χρήστης κατοχύρωσε με το ποντίκι την επιλογή του	69
Εικόνα 24: Με το πάτημα του κουμπιού <i>enter</i> εμφανίζεται η εικόνα που ο χρήστης επέλεξε	69

ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ

Πίνακας 1: Βάση δεδομένων για το υλικό της βιοψίας	45
Πίνακας 2: Μορφολογικά χαρακτηριστικά και χαρακτηριστικά υφής	52

ΠΡΟΛΟΓΟΣ

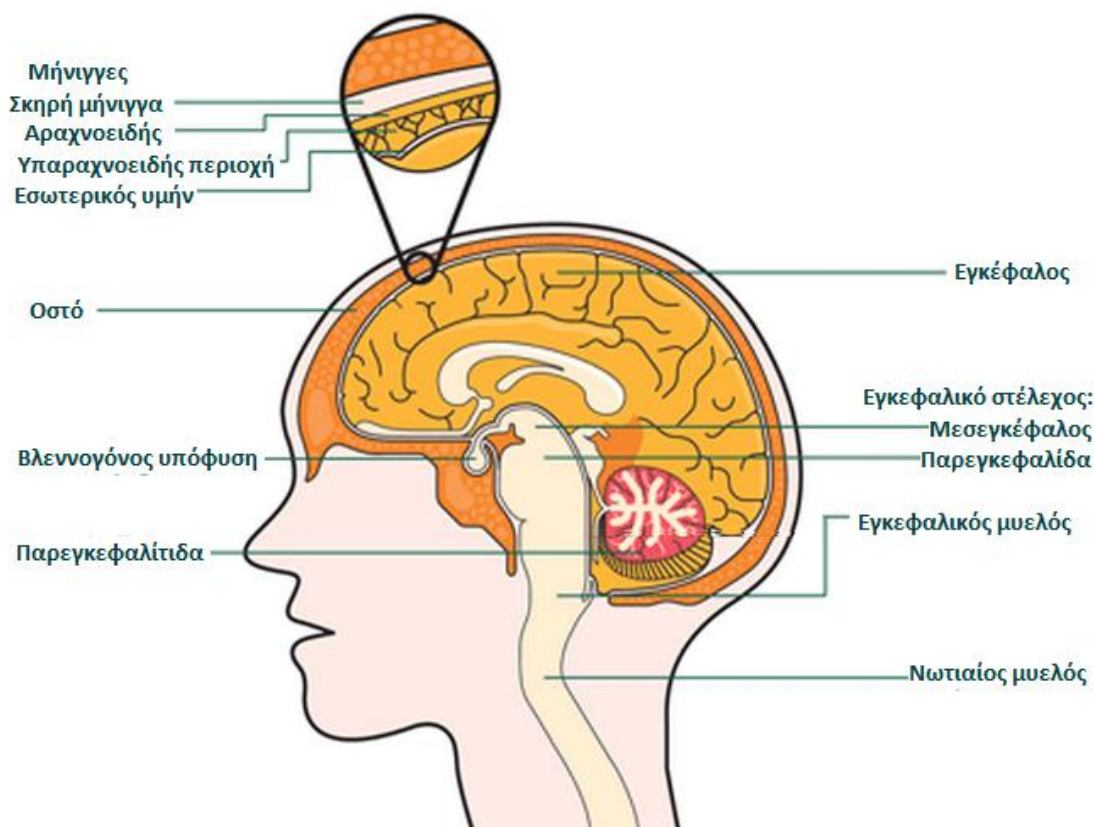
Η παρούσα διπλωματική εργασία πραγματεύεται την ανάπτυξη και υλοποίηση μεθόδων επεξεργασίας και ανάλυσης ιατρικών εικόνων, καθώς και την ανάπτυξη αυτόματων τεχνικών για την βελτίωση της ακρίβειας στην διάγνωση των αστροκυτταρικών όγκων. Οι ιστοπαθολογικές εικόνες δημιουργούνται από τμήμα ιστού που προέρχεται από βιοψία. Οι ιστοπαθολόγοι καλούνται να εξετάσουν τις εικόνες αυτές και να αξιολογήσουν την επιθετικότητα του καρκινικού ιστού. Αυτό επιτυγχάνεται εξετάζοντας τους πυρήνες των κυττάρων, οι οποίοι φέρουν σημαντικές πληροφορίες που οδηγούν στην τελική διάγνωση και κατά επέκταση πρόγνωση του όγκου. Ο προσδιορισμός του βαθμού του όγκου είναι βασικότατο στάδιο της διαδικασίας της διάγνωσης, καθώς είναι καθοριστικός για το είδος της θεραπείας που θα πρέπει ο ασθενής να ακολουθήσει. Σύμφωνα με τον Παγκόσμιο Οργανισμό Υγείας (WHO), οι βαθμοί καθορίζονται με βάση ιστολογικών κριτηρίων σε χαμηλού βαθμού (I, II) και υψηλού βαθμού (III, IV). Στην εργασία αυτή αναπτύχθηκαν και υλοποιήθηκαν αλγόριθμοι, οι οποίοι τμηματοποιούν τους όγκους εγκεφάλου και τους ταξινομούν σε χαμηλού και σε υψηλού βαθμού κακοήθειας, εξετάζοντας το ποσοστό ακρίβειάς τους.

ΕΙΣΑΓΩΓΗ

1.1 Κλινικό Πρόβλημα – Όγκοι εγκεφάλου

1.1.1 Γενικά

Οι εγκεφαλικοί όγκοι θεωρούνται ως μια μορφή καρκίνου με μεγάλη θνησιμότητα, καθώς συγκαταλέγονται στις πιο δύσκολα θεραπεύσιμες ασθένειες και αντιπροσωπεύουν σχεδόν το 90% των όγκων του κεντρικού νευρικού συστήματος (ΚΝΣ)[1]. Κάθε χρόνο καταγράφονται κατά μέσο όρο 7 περιπτώσεις ανά 100.000 άτομα, ενώ η θνησιμότητα ανέρχεται ετησίως στα 4-5 άτομα ανά 100.000. Μελέτες [2] έδειξαν ότι η ηλικία των ασθενών έχει μεγάλη συσχέτιση με την συχνότητα εμφάνισης καρκίνου του εγκεφάλου. Αναφορικά, στην ηλικιακή ομάδα των 0-4 ετών, η συχνότητα εμφάνισης ανέρχεται στα 3 άτομα στα 100.000, ενώ στην ηλικιακή ομάδα των 15-24 ετών παρουσιάζει μείωση. Τέλος, στην ηλικιακή ομάδα των 65-79 ετών η συχνότητα εμφάνισης παρουσιάζει σημαντική κορύφωση, όπου εμφανίζονται χιλιάδες νέες περιπτώσεις καρκίνου του εγκεφάλου ή άλλων μορφών καρκίνου του κεντρικού νευρικού συστήματος (ΚΣΝ), ετησίως[2,3]. Στην εικόνα 1 παρουσιάζονται οι διαφορετικές δομές του εγκεφάλου, ενώ στην εικόνα 2 οι περιοχές αυτού.



Εικόνα 1: Οι διαφορετικές δομές του εγκεφάλου

1.1.2 Αιτιολογία

Ένα καίριο ερώτημα αποτελεί το ποια είναι η αιτιολογία των όγκων του εγκεφάλου και το αν υπάρχουν κάποιες συγκεκριμένες συνθήκες κάτω από τις οποίες αυτοί μπορούν ευκολότερα να δημιουργηθούν. Δυστυχώς, δεν υπάρχουν γνωστοί χημικοί ή περιβαλλοντικοί παράγοντες που να σχετίζονται με την εμφάνιση όγκων στον εγκέφαλο[4]. Ωστόσο, γνωστή είναι η συσχέτιση που υπάρχει με την έκθεση σε

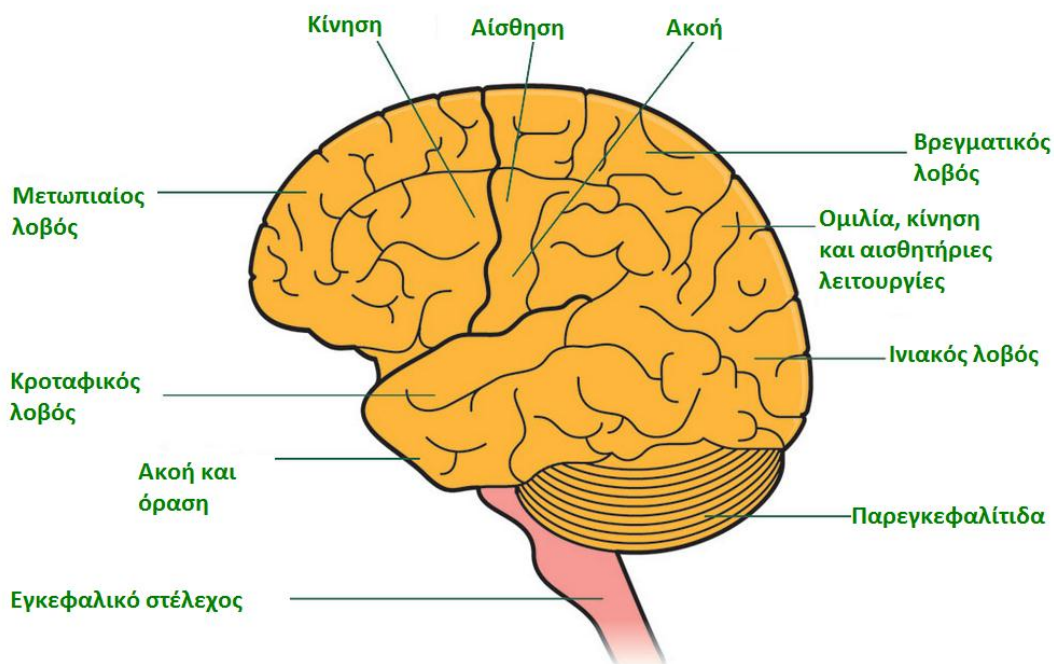
βινυλοχλωρίδιο. Πιθανοί, επίσης, συσχετισμοί με την εμφάνιση όγκων εγκεφάλου θεωρούνται η χρήση κινητών τηλεφώνων και η γενικότερη τεχνολογία που χρησιμοποιεί ραδιοκύματα, χωρίς ωστόσο να υπάρχει κάποια απόδειξη αυτού[5]. Τέλος, υπάρχουν αρκετά κληρονομικά γενετικά σύνδρομα που μπορεί να δημιουργούν προδιάθεση στην ανάπτυξη καρκινικών όγκων στον εγκέφαλο. Στα γενετικά αυτά σύνδρομα συγκαταλέγονται το von Hippel-Lindau σύνδρομο, η νευροϊνωμάτωση και η οζώδης σκλήρυνση[6].

1.1.3 Συμπτώματα

Ο πονοκέφαλος είναι ένα από τα κυριότερα συμπτώματα του όγκου στον εγκέφαλο και είναι πιθανό να συνοδεύεται από εμετό. Οι κακοήθεις μάζες προκαλούν πίεση στον εγκέφαλο και μπορεί να οδηγήσουν σε θολή όραση, διπλή όραση ή ακόμα και απώλεια όρασης. Επιπλέον συμπτώματα μπορεί να είναι οι επιληπτικές κρίσεις, η αδυναμία, το μούδιασμα της μιας πλευράς του σώματος ή μέρος αυτού, καθώς επίσης και αλλαγές στη διάθεση ή και τη σκέψη. Ο ασθενής είναι πιθανό να βρίσκεται σε μια γενική κατάσταση ευημερίας[4].

Τα συμπτώματα του καρκίνου του εγκεφάλου σχετίζονται άμεσα και εξαρτώνται από το μέγεθος του όγκου και από την θέση του[7,8]. Συγκεκριμένα:

- **Μετωπιαίος λοβός:** Συνοδεύεται από επιληπτικές κρίσεις, διαταραχές της κριτικής ικανότητας και της μνήμης, αλλαγές στην προσωπικότητα ή στις διανοητικές ικανότητες, παράλυση στην μία πλευρά του σώματος.
- **Βρεγματικός λοβός:** Οι ασθενείς συνήθως χάνουν την ικανότητα να γράφουν και κυρίως, όταν ο όγκος εντοπίζεται στο αριστερό ημισφαίριο, εμφανίζουν διαταραχές στον λόγο τους καθώς και επιληπτικές κρίσεις.
- **Ινιακός λοβός:** Σχετίζεται άμεσα με την όραση, επομένως μπορεί να οδηγήσει σε τύφλωση ή και επιληπτικές κρίσεις.
- **Κροταφικός λοβός:** Συνήθως δεν εμφανίζονται συμπτώματα όταν ο όγκος βρίσκεται σε αυτή την θέση του εγκεφάλου. Ωστόσο, ο ασθενής μπορεί περιστασιακά να εμφανίσει διαταραχές στην ομιλία ή και επιληπτικές κρίσεις.
- **Μήνιγγες:** Τα συμπτώματα που εμφανίζονται εξαρτώνται από την περιοχή την οποία πιέζει ο όγκος.
- **Παρεγκεφαλίδα:** Παρατηρείται έλλειψη συντονισμού, που επηρεάζει τα πόδια και την ομιλία, καθώς επίσης και αστάθεια, τρεμάμενες κινήσεις των οφθαλμών, εμετός και δυσκαμψία του αυχένα.
- **Επένδυμα (Ependyma):** Κύριο σύμπτωμα είναι η υδροκεφαλία, η οποία είναι μια παθολογική κατάσταση του εγκεφάλου που χαρακτηρίζεται από διαστολή ή διεύρυνση των κοιλιών του εγκεφάλου και υπερβολική συγκέντρωση εγκεφαλονωτιαίου υγρού (ENY).
- **Δευτερογενείς όγκοι:** Πρόκειται για τους μεταστατικούς όγκους, οι οποίοι μπορεί να προκαλέσουν οίδημα, το οποίο με την σειρά του προκαλεί πονοκέφαλο, εμετό και ναυτία[7,8].



Εικόνα 2: Οι περιοχές του εγκεφάλου

1.1.4 Τύποι

Υπάρχουν δύο κύριοι τύποι όγκων εγκεφάλου, οι οποίοι είναι οι πρωτογενείς και οι δευτερογενείς όγκοι. Οι πρωτογενείς όγκοι αποτελούνται είτε από νευρώνες (εγκεφαλικά κύτταρα) είτε από νευροεπιθηλιακά κύτταρα (υποστηρικτικά κύτταρα) και χαρακτηρίζονται ως καλοήθεις (μη καρκινικοί) ή κακοήθεις (καρκινικοί) όγκοι[1,2]. Οι δευτερογενείς όγκοι αποτελούν μεταστατικούς όγκους, οι οποίοι προέρχονται από καρκινικά κύτταρα που έχουν μεταναστεύσει από την αρχική τους θέση και έχουν εισέλθει στο κεντρικό νευρικό σύστημα, μέσω του αιματοεγκεφαλικού φραγμού (blood-brain barrier).

Τα κύτταρα των καλοηθών όγκων έχουν αργή ανάπτυξη και δημιουργούν καλά ορισμένους σχηματισμούς. Οι καλοήθεις όγκοι έχουν αποτελεσματική θεραπεία όταν ο όγκος δεν βρίσκεται σε ευαίσθητη περιοχή ζωτικής σημασίας και μπορεί να αφαιρεθεί. Ωστόσο, η διάγνωσή τους είναι αρκετά δύσκολη, διότι στο μικροσκόπιο μοιάζουν κατά πολύ με τα κανονικά κύτταρα. Σημαντικό αποτελεί το γεγονός ότι το 40% των πρωτογενών όγκων είναι καλοήθεις[3].

Αντίθετα με τους καλοήθεις, οι κακοήθεις όγκοι έχουν επιθετική φύση, εισβάλλουν και έχουν ανεξέλεγκτη αύξηση της μάζας τους. Οι κακοήθεις όγκοι του εγκεφάλου, ωστόσο, σπάνια εξαπλώνονται σε άλλα μέρη του σώματος (πνεύμονας, ήπαρ, μαστός, κτλ.), σε αντίθεση με άλλα είδη κακοηθών όγκων. Ο πιο κοινός τύπος κακοήθους πρωτογενούς όγκου εγκεφάλου είναι το πολύμορφο γλοιοβλάστωμα (grade IV), το οποίο αποτελεί το 20% του συνόλου των πρωτογενών όγκων, ενώ παράλληλα η χειρουργική τους επέμβαση είναι εξαιρετικά επικίνδυνη λόγω της ευαισθησίας των γύρω ιστών[9].

1.1.5 Παθολογική ανατομία γλοιωμάτων

Τα νεοπλάσματα του νευροεπιθηλιακού ιστού αποτελούν τη μεγαλύτερη ομάδα πρωτοπαθών νεοπλασμάτων του ΚΝΣ, με συχνότερα τα γλοιώματα, που προέρχονται

από τα νευρογλοιακά κύτταρα. Τα γλοιώματα είναι οι πιο συχνοί πρωτοπαθείς όγκοι του εγκεφάλου με μέγιστη συχνότητα εμφάνισή τους την Πέμπτη και έκτη δεκαετία της ζωής και αποτελούν το 40% έως 50% του συνόλου των ενδοκρανιακών όγκων[10]. Ανάλογα με την προέλευση και την μορφολογία τους, τα νεοπλάσματα χαρακτηρίζονται ως αστροκυτταρικά, ολιγοδενδρογλοιακά και επενδυματικά, ενώ υπάρχουν και μεικτές μορφές (πχ. ολιγοαστροκυττώματα)[10].

Τα αστροκυτταρικά, τα συχνότερα πρωτοπαθή νεοπλάσματα του ΚΣΝ, αποτελούν μια ετερογενή ομάδα από περιγεγραμμένες βλάβες με βραδεία εξέλιξη, έως διηθητικά νεοπλάσματα υψηλής κακοήθειας. Στους ενήλικες αναπτύσσονται κυρίως στα εγκεφαλικά ημισφαίρια ενώ στην παιδική ηλικία συχνότερα στο εγκεφαλικό στέλεχος ή το θάλαμο. Λιγότερο συχνά αναπτύσσονται στο νωτιαίο μυελό και την παρεγκεφαλίδα.

Τα αστροκυτταρικά νεοπλάσματα διακρίνονται σε[10]:

- α) Διάχυτα διηθητικά (διάχυτο αστροκύττωμα, αναπλαστικό αστροκύττωμα, γλοιοβλάστωμα) και
- β) Καλά περιγεγραμμένα (πιλοκυτταρικό, πολύμορφο ξανθοαστροκύττωμα, υποεπενδυματικό γιγαντοκυτταρικό αστροκύττωμα),

ενώ ανάλογα με την βιολογική τους συμπεριφορά διακρίνονται σε:

- α) Χαμηλής κακοήθειας (πιλοκυτταρικό, διάχυτο αστροκύττωμα, πολύμορφο ξανθοαστροκύττωμα) και
- β) Υψηλής κακοήθειας (αναπλαστικό, γλοιοβλάστωμα).

Τα ολιγοδενδρογλοιακά διακρίνονται στα ολιγοδενδρογλοιώματα και στα αναπλαστικά ολιγοδενδρογλοιώματα και αναπτύσσονται σε όλο το νευράξονα, κυρίως όμως στα εγκεφαλικά ημισφαίρια και συνήθως στους ενήλικες[10].

Τα επενδυματικά αναπτύσσονται σε όλο το νευράξονα σε στενή σχέση με το επένδυμα ή υπολειμματικά στοιχεία. Στα παιδιά αναπτύσσονται συχνότερα στην τέταρτη κοιλία, ενώ στους ενήλικους υπερσκληνίδια και στον νωτιαίο μυελό. Τα επενδυματικά διακρίνονται σε υποεπενδυμώματα, μυξοθελώδη επενδυμώματα, επενδυμώματα και αναπλαστικά επενδυμώματα[10].

1.2 Αστροκυττώματα

Τα νευρογλοιακά νεοπλάσματα αποτελούν πρωτοπαθείς καρκινικούς όγκους του εγκεφάλου και αποτελούνται από τα νευρογλοιακά κύτταρα, τα οποία παρέχουν σημαντική ή διαρθρωτική στήριξη για τα νευρικά κύτταρα στον εγκέφαλο. Ο χαρακτηρισμός των νευρογλοιακών όγκων βασίζεται στον κυτταρικό τύπο και την προέλευσή του. Συγκεκριμένα, οι όγκοι που προέρχονται από τα αστροκύτταρα ονομάζονται αστροκυττώματα, κοκ[11].

1.2.1 Γενικά

Τα αστροκύτταρα έχουν σχήμα αστεριού και κύριο καθήκον τους είναι να μεταφέρουν τις θρεπτικές ουσίες από τα αιμοφόρα αγγεία στους νευρώνες. Οι νευρώνες καταναλώνουν μια μεγάλη ποσότητα τροφής από την οποία παράγουν ενέργεια. Μετά από αυτήν την μεταβολική λειτουργία, οι νευρώνες πρέπει να υποβάλλουν τα υποπροϊόντα της πέψης. Τα αστροκύτταρα συλλέγουν τα υποπροϊόντα αυτά στα αιμοφόρα αγγεία και τα απομακρύνουν[1,2]. Τα αστροκύτταρα μπορούν να

προκαλέσουν τη διαμόρφωση των μαζών που αποτελούνται από ανώμαλα κύτταρα, δημιουργώντας το σχήμα των αστροκυτταρικών όγκων, των αστροκυττωμάτων.

Τα ανώμαλα αστροκυτταρικά κύτταρα δεν εκτελούν τις χρήσιμες λειτουργίες για τον οργανισμό και επιπρόσθετα καταναλώνουν τα θρεπτικά συστατικά και το οξυγόνο που χρειάζονται τα φυσιολογικά κύτταρα και εκκρίνουν τα υποπροϊόντα στο εξωκυττάριο υγρό. Επιπλέον, παρουσιάζουν ανωμαλίες στο ποσοστό μίτωσης, που είναι υψηλότερο από αυτό των φυσιολογικών κυττάρων. Μερικά από αυτά είναι σε θέση να μετακινούνται από το ένα μέρος στο άλλο και φαίνεται πως μπορούν να αποφύγουν τον εντοπισμό τους από το ανοσολογικό σύστημα, το οποίο θα μπορούσε να τα εντοπίσει και να τα σκοτώσει[1-3].

1.2.2 Διάγνωση

Ο εντοπισμός ενός όγκου του εγκεφάλου συνήθως περιλαμβάνει νευρολογική εξέταση από τον γιατρό, τομογραφίες εγκεφάλου ή και ανάλυση του εγκεφαλικού ιστού. Οι γιατροί χρησιμοποιούν τις διαγνωστικές πληροφορίες για την ταξινόμηση του όγκου από λιγότερο επιθετική (καλοήθεια) σε πιο επιθετική (κακοήθεια). Διάφορες ανωμαλίες σε ορισμένες νευρολογικές λειτουργίες, όπως κίνηση των ματιών ή αντανακλαστικά, αποτελούν ενδείξεις σχετικά με την ύπαρξη όγκου. Ο προσδιορισμός του τύπου του όγκου βοηθά τους γιατρούς να καθορίσουν την πλέον ενδεδειγμένη πορεία της θεραπείας.

Υπάρχουν διάφορες εξειδικευμένες τεχνικές απεικόνισης και εργαστηριακές εξετάσεις που μπορούν να ανιχνεύσουν την παρουσία όγκου και να παρέχουν πληροφορίες σχετικά με την θέση αυτού, τον τύπο του και την έκταση της εξάπλωσης. Η νευρολογική εξέταση είναι μια σειρά από τεστ για την μέτρηση της λειτουργίας του νευρικού συστήματος του ασθενούς και της σωματικής και πνευματικής του εγρήγορσης.

Η σάρωση του εγκεφάλου είναι η εικόνα των εσωτερικών δομών του εγκεφάλου. Ένα ειδικό μηχάνημα παίρνει μια σάρωση με τον ίδιο τρόπο που μια φωτογραφική μηχανή παίρνει μια φωτογραφία. Χρησιμοποιώντας, την τεχνολογία των υπολογιστών, η σάρωση συγκεντρώνει μια εικόνα του εγκεφάλου με φωτογράφιση από διάφορες οπτικές γωνίες. Ορισμένοι τύποι σαρώσεων χρησιμοποιούν σκιαγραφικά μέσα, τα οποία βοηθούν τον γιατρό να δει την διαφορά μεταξύ παθολογικού και φυσιολογικού ιστού. Το σκιαγραφικό υλικό εγχέεται σε μια φλέβα και διαρρέεται στον εγκεφαλικό ιστό. Ο ανώμαλος ή παθολογικός ιστός του εγκεφάλου απορροφά περισσότερη χρωστική ουσία σε σχέση με τους υγιείς ιστούς. Οι πιο συνηθισμένες σαρώσεις που χρησιμοποιούνται για την διάγνωση είναι οι εξής:

- **Μαγνητική τομογραφία (MRI):** Η μαγνητική τομογραφία είναι μια συσκευή σάρωσης που χρησιμοποιεί μαγνητικά πεδία και τους υπολογιστές για να συλλάβει τις εικόνες του εγκεφάλου και να τις τοποθετήσει σε ταινία. Ο τομογράφος δεν χρησιμοποιεί ακτίνες Χ αλλά τα τοιχώματά του αποτελούνται από μαγνήτες, που δημιουργούν ένα μόνιμο μαγνητικό πεδίο. Ο ασθενής εκτίθεται σε ισχυρό μαγνητικό πεδίο έως και 1000 φορές μεγαλύτερο από το μαγνητικό πεδίο της Γης. Το μαγνητικό πεδίο προσανατολίζει τους πυρήνες σε κανονική διάταξη με τη βοήθεια μιας ακολουθίας ηλεκτρομαγνητικών κυμάτων. Στη συνέχεια, τα άτομα επανέρχονται στην αρχική τους κατάσταση, εκπέμποντας ακτινοβολία. Η ακτινοβολία συλλαμβάνεται από ειδικούς δέκτες και στέλνεται υπό μορφή ηλεκτρικών ώσεων σε ηλεκτρονικό υπολογιστή. Τέλος, γίνεται επεξεργασία του σήματος για τον σχηματισμό της εικόνας. Ο μαγνητικός

τομογράφος παρέχει εικόνες από διάφορα επίπεδα, τα οποία επιτρέπουν στους γιατρούς να δημιουργήσουν μια τρισδιάστατη εικόνα όγκου[7].

- **Αξονική τομογραφία (CT):** Η αξονική τομογραφία συνδυάζει ακτίνες Χ, που απεικονίζουν εγκάρσιες τομές του ανθρώπινου σώματος, και τεχνολογία υπολογιστών και μπορεί να δείξει ένα συνδυασμό των μαλακών ιστών, οστών και των αιμοφόρων αγγείων. Οι εικόνες της αξονικής τομογραφίας μπορούν να καθορίσουν ορισμένους τύπους όγκων, καθώς και να βοηθήσουν στον εντοπισμό οιδήματος, αιμορραγίας και ασβεστοποίησης οστών και ιστών. Μερικές φορές γίνεται χρήση σκιαγραφικού υλικού για την ευκολότερη εντόπιση του όγκου και συνήθως το ιώδιο είναι η ουσία που χρησιμοποιείται κατά την διάρκεια μιας αξονικής τομογραφίας[7,12].

Η αξονική και μαγνητική τομογραφία παρέχουν πληροφορίες σχετικά με την θέση, την έκταση και την σύνθεση των κακοηθών νεοπλασιών. Η αξονική τομογραφία είναι καλύτερη για την ανίχνευση ασβεστοποίησης, βλάβες κρανίου και υπεροξεία αιμορραγία, ενώ με την μαγνητική τομογραφία μπορούν να ανιχνευτούν καλύτερα οιδήματα, τένοντες, σύνδεσμοι και βλάβες. Άλλοι τρόποι απεικόνισης του εγκεφάλου είναι οι εξής:

- **Τομογραφία εκπομπής ποζιτρονίων (PET):** Η σάρωση αυτή παρέχει μια εικόνα δραστηριότητας του εγκεφάλου, παρά την δομή του, μετρώντας τον ρυθμό με τον οποίο ένας όγκος απορροφά γλυκόζη. Ο τομογράφος μετρά τη δραστηριότητα του εγκεφάλου και στέλνει πληροφορίες σε έναν υπολογιστή, στον οποίο δημιουργεί μια ζωντανή εικόνα.
- **Αξονική τομογραφία εκπομπής φωτονίων (SPECT):** Η σάρωση αυτή είναι παρόμοια με μία PET σάρωση. Μπορεί να βοηθήσει για την ανίχνευση ανωμαλιών στον αιματοεγκεφαλικό φραγμό, καθώς δείχνει την ροή του αίματος στον εγκέφαλο.

Το PET και το SPECT χρησιμοποιούνται κυρίως για μετεγχειρητικούς σκοπούς, όπως για την διαφοροποίηση της νέκρωσης από ακτινοβολία και την επανεμφάνιση όγκου[7].

Το τελικό βήμα για την ολοκλήρωση της διαγνωστικής διαδικασίας είναι η **βιοψία**. Η βιοψία είναι μια χειρουργική διαδικασία κατά την οποία ένα λεπτό δείγμα ιστού, που ανήκει στον ύποπτο όγκο, αφαιρείται με χειρουργική επέμβαση, προκειμένου να εξεταστούν κάτω από το μικροσκόπιο. Η βιοψία παρέχει πληροφορίες για τους τύπους των ανώμαλων κυττάρων που παρουσίασαν τον όγκο. Ο σκοπός της βιοψίας είναι να ανακαλύψουμε τον τύπο και τον βαθμό του όγκου και είναι η πιο ακριβής μέθοδος απόκτησης διάγνωσης, ενώ η θεραπεία είναι έντονα συνυφασμένη με το συμπέρασμα σχετικά με τα κύτταρα του δειγματικού ιστού[1,8].

1.2.3 Ταξινόμηση σε βαθμίδες (Grades)

Η πρόγνωση των αστροκυττωμάτων ποικίλει σημαντικά και μπορεί να εκτιμηθεί χρησιμοποιώντας την ταξινόμηση κατά την οποία ελέγχεται η ιστολογική προέλευση των όγκων, η ανατομική εντόπισή τους και η κλινική τους πορεία[1,2]. Σύμφωνα με το σύστημα ταξινόμησης του Παγκόσμιου Οργανισμού Υγείας (WHO)[11], τα αστροκυττώματα κατατάσσονται σε 4 κατηγορίες (I,II,III,IV) με βάση ιστολογικών χαρακτηριστικών. Τα κριτήρια αυτά είναι η κυτταροβρίθεια, ο κυτταρικός πλειομορφισμός, η πυρηνική ατυπία, ο αριθμός μιτώσεων, η νέκρωση και ο μικροαγγειακός πολλαπλασιασμός. Επίσης, τα αστροκυττώματα ταξινομούνται αναφορικά με την ικανότητα διήθησης σε γειτονικό ιστό και την εξελικτική τους πορεία.

Χαμηλής κακοήθειας[10]:

- α) **Μη διηθητικό διάχυτο** αστροκύττωμα (Grade I): Πρόκειται για αστροκύτταρα χωρίς αναπλασία και εδώ συγκαταλέγεται το πιλοκυτταρικό αστροκύττωμα, τα πολυμυξοειδικά, το πλειομορφικό ξανθοαστροκύττωμα και το υποεπενδυματικό γιγαντοκυτταρικό αστροκύττωμα. Το πιλοκυτταρικό αστροκύττωμα απεικονιστικά εμφανίζεται ως σαφώς ορισμένος όγκος και η κλασική εικόνα είναι αυτή ενός κυστικού όγκου με εμπλουτισμένο τοιχωματικό οζίδιο. Το υποεπενδυματικό γιγαντοκυτταρικό εμφανίζεται πάντα στα πλαίσια της οζώδους σκλήρυνσης, προκαλώντας απόφραξη και τελικά υδροκέφαλο. Στην αξονική τομογραφία απεικονίζονται υπέρπυκνα με παρουσία απασβεστώσεων.
- β) **Διάχυτο διηθητικό** αστροκύττωμα (Grade II): Πρόκειται για καλώς διαφοροποιημένο αστροκύττωμα και ο συνηθέστερος τύπος είναι το ινώδες αστροκύττωμα, ενώ άλλα είδη είναι τα πρωτοπλασματικά και τα γεμιστοκυτταρικά αστροκυττώματα. Βασικά χαρακτηριστικά της κατηγορίας αυτής είναι η βραδεία ανάπτυξη, η χαμηλή κυτταροβρίθεια και η πυρηνική ατυπία, ενώ δεν παρατηρούνται μιτώσεις. Απεικονιστικά οι όγκοι αυτοί είναι ανομοιογενείς. Στην αξονική τομογραφία μπορεί να είναι δύσκολο να εντοπιστούν και συνήθως δεν εμπλουτίζονται. Στην μαγνητική τομογραφία εμφανίζουν χαμηλό σήμα στις T1 εικόνες και υψηλό σήμα στις T2. Μετά την χορήγηση παραμαγνητικής ουσίας συνήθως δεν εμπλουτίζονται ή εμπλουτίζονται σε μικρό βαθμό.

Υψηλής κακοήθειας[10]:

- α) **Αναπλαστικό** αστροκύττωμα (Grade III): Το αναπλαστικό αστροκύττωμα είναι διηθητική βλάβη με βιολογική συμπεριφορά μεταξύ του αστροκυττώματος και του γλοιοβλαστώματος. Βασικά χαρακτηριστικά αυτής της κατηγορίας είναι η μέτρια αναπλασία της τάξεως του 50%, πολλές μιτώσεις, εστίες έντονης κυτταροβρίθειας, πιθανές νεκρώσεις, αγγειακή υπερπλασία και πυρηνική ατυπία. Τα αναπλαστικά αστροκυττώματα είναι πολύ επιθετικοί όγκοι, διεισδύουν σε παρακείμενες δομές του εγκεφάλου και έχουν κακή πρόγνωση. Συνήθως προκύπτουν από την εξέλιξη χαμηλότερου βαθμού όγκων. Στην αξονική και μαγνητική τομογραφία απεικονίζεται ως ασαφής ανομοιογενής βλάβη με περιεστιακό οίδημα. Μετά τη χορήγηση σκιαστικού μπορεί να μην εμπλουτίζεται ή μπορεί να εμφανίζει εστιακό ή οζώδη ανομοιογενή εμπλουτισμό.
- β) **Πολύμορφο γλοιοβλάστωμα** (Grade IV): Το πολύμορφο γλοιοβλάστωμα είναι το συχνότερο νεόπλασμα του εγκεφάλου. Η κεντρική νέκρωση είναι το σήμα κατατεθέν του γλοιοβλαστώματος. Άλλα χαρακτηριστικά είναι η εκσεσημασμένη αναπλασία των περισσότερων κυττάρων, άφθονες μιτώσεις, πυρηνική ατυπία, αγγειακή θρόμβωση, αγγειακή υπερπλασία και πολλά παράξενα και πολυστηριζόμενα γιγαντιαία κύτταρα. Είναι οι ταχύτερα αναπτυσσόμενοι και επιθετικοί όγκοι. Στην αξονική και μαγνητική τομογραφία απεικονίζεται ως ανομοιογενής βλάβη λόγω της νέκρωσης, της αιμορραγίας και της κυτταροβρίθειας. Ένα, ακόμα, χαρακτηριστικό στοιχείο είναι η παρουσία περιεστιακού οιδήματος, που προκαλεί σημαντικά πιεστικά φαινόμενα.

1.2.4 Θεραπεία

Ο γενικός κανόνας για την θεραπεία του όγκου είναι η μέγιστη μείωση του καρκινικού ιστού, όσο το δυνατόν νωρίτερα. Ωστόσο, τις περισσότερες φορές δεν εφαρμόζεται λόγω της πολυπλοκότητας της θέσης και του είδους του όγκου του εγκεφάλου.

Στο πιλοκυτταρικό αστροκύτωμα (βαθμός I), το οποίο εμφανίζεται κυρίως στην παιδική και εφηβική ηλικία, είναι δυνατή η ριζική θεραπεία του εάν γίνει ολική αφαίρεση της κύστης και του οζιδίου[10]. Εάν ο όγκος αυτός δεν αφαιρεθεί ολικά τότε μπορεί να υποτροπιάσει και στην περίπτωση αυτή απαιτείται εκ νέου χειρουργική εξαίρεση ή ακτινοθεραπεία. Η θεραπευτική αντιμετώπιση του χαμηλού βαθμού αστροκυτώματος (βαθμός II), το οποίο εμφανίζεται επίσης στην παιδική ή εφηβική ηλικία και σε νεαρούς ενήλικες, περιλαμβάνει χειρουργική εξαίρεση και σε υποτροπή επανάληψη της χειρουργικής εξαίρεσης και ακτινοθεραπεία. Μερικές φορές, όταν υποτροπιάσουν μεταλλάσσονται σε πιο επιθετικά νεοπλάσματα (βαθμού III ή IV)[10]. Το αναπλαστικό αστροκύτωμα (βαθμός III) προσβάλλει ενήλικες και αποτελεί ένα επιθετικό κακοήθες νεόπλασμα. Η θεραπευτική του αντιμετώπιση περιλαμβάνει χειρουργική εξαίρεση ακολουθούμενη από ακτινοθεραπεία και χημειοθεραπεία[10]. Τέλος, το πολύμορφο γλοιοβλάστωμα (βαθμός IV) είναι ένας πολύ κακοήθης όγκος με πολύ λίγο χρόνο επιβίωσης του ασθενή (περίπου 12 μήνες) ανεξάρτητα από την θεραπεία. Η αντιμετώπισή του περιλαμβάνει επίσης χειρουργική αφαίρεση, ακτινοβολία και χημειοθεραπεία σε ορισμένες περιπτώσεις[10]. Ωστόσο, λόγω του ότι η συστηματική χορήγηση χημειοθεραπείας έχει πενιχρά αποτελέσματα, διάφορες νέες πειραματικές θεραπευτικές μέθοδοι περιλαμβάνουν τοπικές εγχύσεις φαρμάκων, διάφορες ανοσολογικές προσεγγίσεις όπως εμβόλια, γονιδιακή θεραπεία και χορήγηση εμβρυϊκών κυττάρων[10].

1.2.5 Επανεμφάνιση και εξέλιξη

Δυστυχώς, δεν υπάρχει προγνωστική μέθοδος που να ενδεικνύει πότε ένας όγκος είναι πιθανό να εμφανιστεί. Παρόλο που σε επιτυχημένες θεραπείες ο όγκος τίθεται σε πλήρη έλεγχο, τα καρκινικά κύτταρα συχνά επανεμφανίζονται μετά από μικρό χρονικό διάστημα. Για χαμηλού βαθμού αστροκυτώματα (βαθμός II) υπολογίζεται περίπου σε 5 χρόνια, για αναπλαστικά αστροκυτώματα (βαθμός III) σε 3 χρόνια και για γλοιοβλαστώματα (βαθμός IV) σε λίγους μήνες. Η υποτροπή του όγκου ακολουθούμενη από την εξέλιξή του συχνά εξελίσσεται σε υψηλότερου βαθμού κακοήθειας και η είναι η σημαντικότερη αιτία θανάτου[13]. Ωστόσο, μερικοί προγνωστικοί παράγοντες για την εξέλιξη θεωρούνται η μικρή ηλικία, η καλή κατάσταση του ασθενούς, η απουσία νευρολογικών εκδηλώσεων, η θέση του όγκου και η μακροσκοπικά πλήρης αφαίρεση του όγκου.

2. ΟΠΤΙΚΟ ΜΙΚΡΟΣΚΟΠΙΟ

2.1 Βασικές Αρχές Λειτουργίας

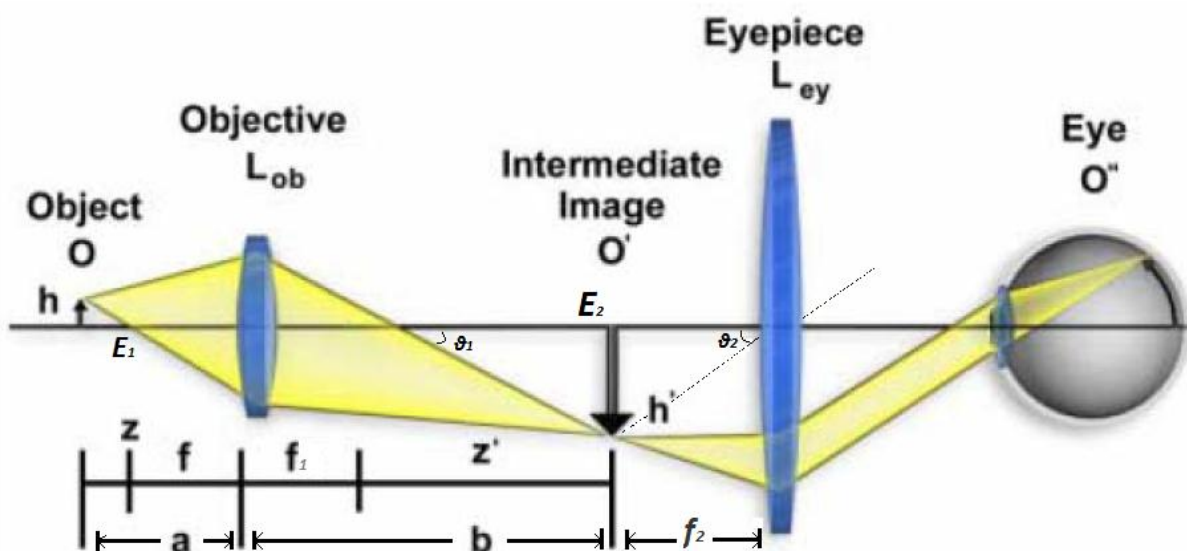
Η οπτική μικροσκοπία αποτελεί την απλούστερη μέθοδο μεγέθυνσης μικρών αντικειμένων και μελέτης του μικρόκοσμου. Στην κλασική μικροσκοπία και στην περίπτωση της μελέτης μικρο-οργανισμών, για να αναδειχθούν και να φανούν τα προς μελέτη κύτταρα ή οι υποκυτταρικές δομές ενός δείγματος, έχουν αναπτυχθεί διάφορες τεχνικές στερέωσης και χρώσης, επιλεκτικές για ορισμένα κύτταρα ή οργανίδια κυττάρων.

Το οπτικό μικροσκόπιο περιλαμβάνει ένα σύστημα συγκλινόντων φακών: τον προσοφθάλμιο φακό και τον αντικειμενικό φακό, με εστιακή απόσταση (f_1) μικρότερη από αυτήν του προσοφθάλμιου (f_2). Οι φακοί αυτοί είναι κατάλληλα επιλεγμένοι για την ελαχιστοποίηση σφαλμάτων, όπως η χρωματική ή η σφαιρική εκτροπή.

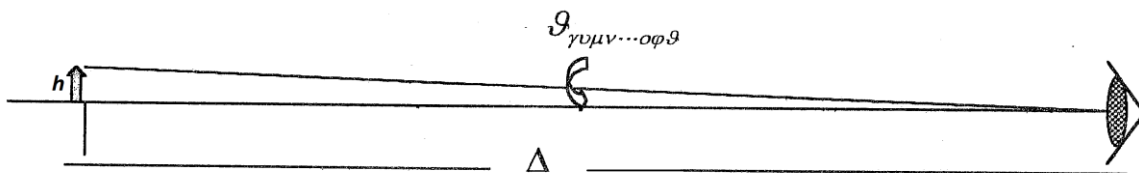
Το προς παρατήρηση αντικείμενο h τοποθετείται πριν από την κύρια εστία E_1 του αντικειμενικού φακού L_{ob} (εικόνα 3). Αρχικά σχηματίζεται ένα πραγματικό, μεγαλύτερο του αντικειμένου είδωλο h' που αποτελεί το αντικείμενο για το προσοφθάλμιο σύστημα φακών και βρίσκεται κοντά στην εστία E_2 του προσοφθάλμιου φακού L_{ey} . Ο φακός L_{ey} λειτουργεί ως μεγεθυντικός φακός και δίδει φανταστικό είδωλο ανεστραμμένο. Το τελικό είδωλο h'' βρίσκεται στην ελάχιστη απόσταση ευκρινούς οράσεως Δ .

Μεγέθυνση (M) ενός μικροσκοπίου καλείται ο λόγος της γωνίας οράσεως θ_2 υπό την οποία βλέπουμε το φανταστικό είδωλο από το μικροσκόπιο προς τη γωνία οράσεως $\theta_{\gamma\mu.οφθ}$ υπό την οποία βλέπουμε το αντικείμενο h δια γυμνού οφθαλμού, όταν αυτό βρίσκεται στην ελάχιστη απόσταση ευκρινούς οράσεως Δ (εικόνα 4)[14].

$$M = \frac{\theta_2}{\theta_{\gamma\mu.οφθ}} \tag{2.1}$$



Εικόνα 3: Σχηματική αναπαράσταση της πορείας των ακτίνων στο οπτικό μικροσκόπιο[14]



Εικόνα 4: Σχηματική αναπαράσταση της γωνίας ευκρινούς οράσεως

Όταν αυτές οι γωνίες είναι μικρές ($\tan\theta \approx \theta$), η μεγέθυνση δίνεται από τη σχέση[15]:

$$M = \frac{\tan \theta_2}{\tan \theta_{\gamma\mu.\sigma\phi\theta}} \quad (2.2)$$

όπου οι $\tan\theta_2$ και η $\tan\theta_{\gamma\mu.\sigma\phi\theta}$ είναι αντίστοιχα:

$$\tan \theta_2 = \frac{h'}{f_2} \quad \tan \theta_{\gamma\mu.\sigma\phi\theta} = \frac{h}{\Delta} \quad (2.3),(2.4)$$

Έτσι έχουμε τη μεγέθυνση[15,16]:

$$M = \frac{h'}{h} \frac{\Delta}{f_2} \quad (2.5)$$

Επειδή η εγκάρσια μεγέθυνση, m_1 , του αντικειμενικού φακού είναι, εξ ορισμού, ίση με[15]:

$$m_1 = \frac{h'}{h} \quad (2.6)$$

η δε γωνιακή μεγέθυνση, M_2 , του προσοφθάλμιου είναι ίση με[15]:

$$M_2 = \frac{\Delta}{f_2} \quad (2.7)$$

η ολική μεγέθυνση M μπορεί να γραφτεί ως το γινόμενο της μεγέθυνσης του αντικειμενικού φακού (εγκάρσια μεγέθυνση), m_1 , επί τη γωνιακή μεγέθυνση του προσοφθάλμιου φακού, M_2 ,

$$M = m_1 \cdot M_2 \quad (2.8)$$

Η μεγέθυνση μπορεί να εκφραστεί και μέσω των εστιακών αποστάσεων f_1 και f_2 των δυο φακών. Στο παραπάνω σχήμα συμβολίζουμε με z' την απόσταση των εστιών E_1 και E_2 , η οποία καλείται οπτικό μήκος του μικροσκοπίου. Αν ονομάζουμε a την απόσταση του αντικειμένου από τον αντικειμενικό φακό και β την απόσταση f_1+z' , βάσει του γνωστού τύπου των κατασκευαστών των φακών έχουμε την σχέση[15,16]:

$$\frac{1}{a} + \frac{1}{\beta} = \frac{1}{f_1} \quad (2.9)$$

και επειδή, από τα όμοια τρίγωνα που σχηματίζονται, $m_1 = \frac{\beta}{\alpha}$, (2.10)

έπεται ότι: $m_1 = \frac{z'}{f_1}$ (2.11)

οπότε τελικά η μεγέθυνση παίρνει την μορφή: $M = \frac{\Delta z'}{f_1 f_2}$ (2.12)

Τόσο στους αντικειμενικούς φακούς, όσο και στους προσοφθάλμιους των μικροσκοπίων, η μεγέθυνση αναγράφεται συμβολικά ως Mx, π.χ. 40x. Συνεπώς μπορούμε να βρούμε την μεγέθυνση ενός μικροσκοπίου με απλό πολλαπλασιασμό των τιμών που αναγράφονται στους δυο φακούς.

Επειδή ο φωτισμός του δείγματος κατανέμεται κατά την μεγέθυνσή του σε πολύ μεγαλύτερη επιφάνεια, η τελική εικόνα είναι αμυδρή και επομένως απαιτείται ισχυρός φωτισμός του αντικειμένου, με την βοήθεια μιας φωτεινής πηγής και του συλλέκτη φακού[14]. Ο φωτισμός πρέπει να προσαρμόζεται κατάλληλα όταν μεταβάλλεται η μεγέθυνση (αύξηση της μεγέθυνσης – αύξηση φωτισμού).

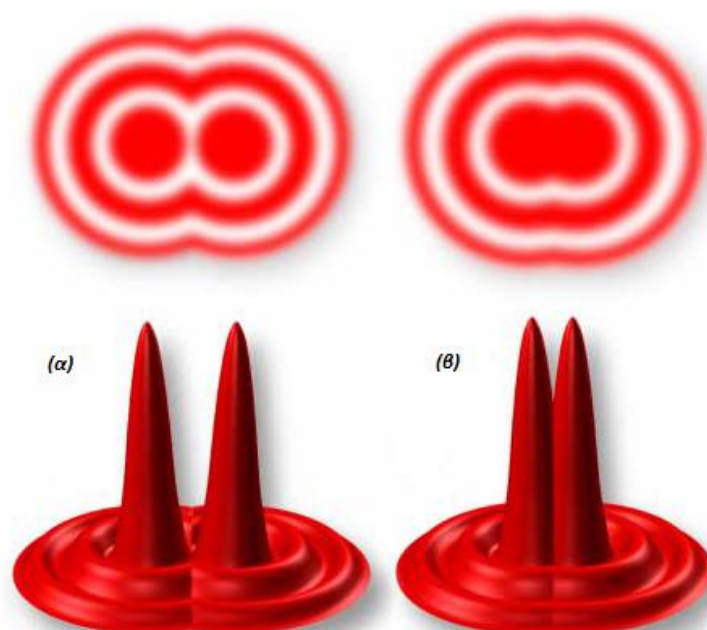
Στους αντικειμενικούς φακούς αναγράφεται επίσης το αριθμητικό άνοιγμα A αυτών, το οποίο εξαρτάται από το δείκτη διάθλασης n του μέσου εντός του οποίου ευρίσκεται το αντικείμενο και από το ημίτονο του ημίσεως του γωνιακού ανοίγματος Φ του φακού,

σύμφωνα με τη σχέση: $A = n \sin \frac{\Phi}{2}$ (2.13)

Μέσω του μικροσκοπίου μπορούμε θεωρητικά να πετύχουμε οποιαδήποτε μεγέθυνση θέλουμε, με κατάλληλη εκλογή της ολικής μεγέθυνσης $m_1 m_2$. Αυτό όμως δεν έχει καμία πρακτική σημασία αν δεν μπορούμε να διακρίνουμε λεπτομέρειες του αντικειμένου. Η ικανότητα μεγέθυνσης μικροδομών στο οπτικό μικροσκόπιο περιορίζεται από το διακριτικό του όριο. Ως διακριτικό όριο του μικροσκοπίου ορίζεται η ελάχιστη απόσταση δύο σημείων του δείγματος, τα οποία δίνουν είδωλα διακριτά ως δυο διαφορετικά σημεία. Το διακριτικό όριο του μικροσκοπίου, R, δίνεται από τη σχέση[14]:

$$R = 1.22 \left(\frac{\lambda}{2A} \right) \quad (2.14)$$

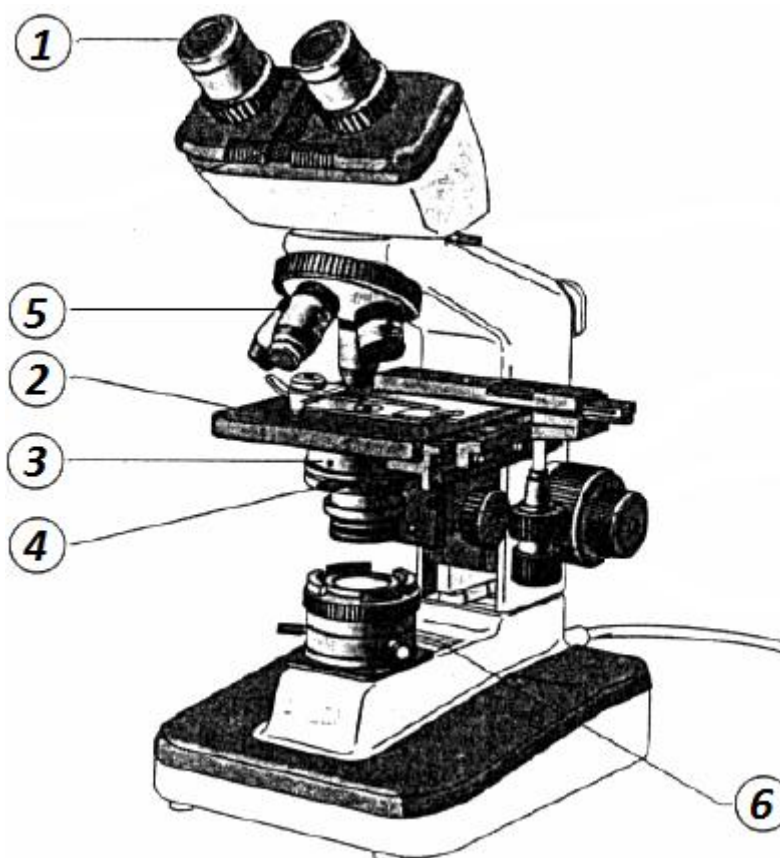
όπου λ είναι το μήκος κύματος του χρησιμοποιούμενου φωτός στο κενό και A το αριθμητικό άνοιγμα του φακού. Ο περιορισμός στην τιμή του διακριτικού ορίου του μικροσκοπίου οφείλεται στο φαινόμενο της περίθλασης του φωτός. Το αντίστροφο του διακριτικού ορίου του μικροσκοπίου λέγεται διακριτική ικανότητα, δ.



Εικόνα 5: Δύο σημεία του δείγματος (α) στο διακριτικό τους όριο και (β) όταν τα κεντρικά τους σημεία επικαλύπτονται[14]

2.2 Τμήματα και όργανα του μικροσκοπίου

Το κύριο τμήμα του οπτικού μικροσκοπίου είναι το οπτικό σύστημα, η λειτουργία του οποίου υποστηρίζεται από το μηχανικό σύστημα του οργάνου.



Εικόνα 6: Τμήματα του μικροσκοπίου: (1) Προσοφθάλμιοι φακοί, (2) Αντικειμενοφόρος τράπεζα, (3) Συμπυκνωτής, (4) Ίρις, (5) Αντικειμενικός φακός, (6) Φωτεινή πηγή[17]

Το οπτικό σύστημα αποτελείται από δυο συγκλίνοντες ομοαξονικούς φακούς, τον αντικειμενικό και τον προσοφθάλμιο (εικόνα 6 ,(5,1))[17].

Το μικροσκοπικό παρασκεύασμα τοποθετείται σε μικρή απόσταση από την εστία του αντικειμενικού φακού σχηματίζοντας είδωλο πραγματικό, ανεστραμμένο και με μέγεθος αντιστρόφως ανάλογο με την απόσταση ανάμεσα στο αντικείμενο και την εστία[17]. Ο προσοφθάλμιος φακός, μέσω του οποίου γίνεται η μικροσκοπική παρατήρηση, μεγεθύνει το παραπάνω είδωλο. Το γυάλινο πλακίδιο τοποθετείται στην αντικειμενοφόρο τράπεζα μικροσκόπησης (εικόνα 6, (2)) και συγκρατείται με το άγκιστρο του οδηγητή.

Ο φωτισμός του αντικειμένου εξασφαλίζεται από ένα λαμπτήρα ενσωματωμένο στη βάση του μικροσκοπίου (εικόνα 6, (6)). Κάτω από το κυκλικό άνοιγμα της τράπεζας, μέσω του οποίου το φως φτάνει στο παρασκεύασμα, βρίσκεται ένα σύνθετο σύστημα που συγκροτείται από το συμπυκνωτή και το διάφραγμα - ίριδα και συνήθως από έναν φορέα φίλτρων[17].

Το διάφραγμα - ίρις (εικόνα 6, (4)), διάφραγμα με μεταβαλλόμενη διάμετρο, επιτρέπει την αυξομείωση του φωτισμού ανάλογα με το χρησιμοποιούμενο αντικειμενικό φακό, το πάχος του υλικού παρατήρησης κ.α., εξασφαλίζοντας έτσι τον ομοιόμορφο φωτισμό του παρασκευάσματος[17].

Ο συμπυκνωτής (εικόνα 6, (3)) αποτελείται από ένα σύστημα φακών με διαφορετικό δείκτη διάθλασης και σχετική ανεξαρτησία κινήσεων, έτσι ώστε ο κατάλληλος στην κάθε περίπτωση συνδυασμός τους να εξασφαλίζει τη διέλευση όσο το δυνατόν περισσότερων φωτεινών ακτινών μέσα από το υλικό μικροσκόπησης[17]. Απουσία συμπυκνωτή ένα μεγάλο μέρος των ακτινών εκτρέπεται με αποτέλεσμα ο φωτισμός του αντικειμένου να είναι μειωμένος και το είδωλο λιγότερο ευκρινές.

3. ΥΛΙΚΟ ΚΑΙ ΜΕΘΟΔΟΙ

3.1 Κλινικό υλικό

Το κλινικό υλικό που χρησιμοποιήθηκε προήλθε από 35 ασθενείς που είχαν υποβληθεί σε χειρουργική επέμβαση στο Πανεπιστημιακό Νοσοκομείο της Πάτρας. Εξετάζοντας μικροσκοπικά τις βιοψίες είναι δυνατή η απόκτηση πληροφοριών σχετικά με τον ιστολογικό τύπο και τον βαθμό του όγκου, τα οποία αποτελούν απαραίτητα στοιχεία για τον προσδιορισμό της κατάλληλης θεραπευτικής αγωγής.

Οι περισσότεροι ιστοί είναι άχρωμοι με αποτέλεσμα να είναι δύσκολη η εξέτασή τους στο μικροσκόπιο. Για τον λόγο αυτό γίνεται χρώση των ιστών για να φανερωθούν τα διάφορα συστατικά των κυττάρων και των πυρήνων αλλά και για να διακριθούν μεταξύ τους. Αυτό γίνεται χρησιμοποιώντας μίγματα διαφόρων χρωστικών ουσιών, με το πιο πολυχρησιμοποιημένο στην περίπτωση των αστροκυττωμάτων να είναι το μίγμα Αιματοξυλίνης - Ηωσίνης (Hematoxylin & Eosin - H&E). Η αιματοξυλίνη χρωματίζει μπλε τους πυρήνες των κυττάρων, ενώ η ηωσίνη προσδίδει στο κυτταρόπλασμα ένα κόκκινο χρώμα.

Τα αστροκυττώματα χαρακτηρίζονται από μεγάλη πολυμορφία και ετερογένεια και κατηγοριοποιούνται με βάση ιστολογικά κριτήρια σε τρεις βαθμούς κακοήθειας (I, II, III ή IV) σύμφωνα με τον Παγκόσμιο Οργανισμό Υγείας (World Health Organization - WHO). Από τις 35 βιοψίες, 19 χαρακτηρίστηκαν από ιστοπαθολόγο ιατρό ως χαμηλής κακοήθειας (βαθμός I ή II) και 16 ως υψηλής κακοήθειας (βαθμός III ή IV) και σημειώθηκαν οι περιοχές ενδιαφέροντος από τις οποίες, μέσω οπτικού μικροσκοπίου, εξήχθησαν πολλαπλές εικόνες για κάθε ασθενή με μεγέθυνση 40x για περαιτέρω επεξεργασία και απομόνωση των πυρήνων.

3.2 Αναγνώριση προτύπων

Τα συστήματα αναγνώρισης προτύπων κατηγοριοποιούνται σε supervised και unsupervised αναγνώριση προτύπων. Στην supervised αναγνώριση προτύπων (ταξινόμηση) το μοντέλο είναι σχεδιασμένο ώστε να υπολογίζει τον βαθμό συσχέτισης μεταξύ των δεδομένων εισόδου (πχ. εικόνες μικροσκοπίου) και των επιθυμητών εξόδων (πχ. διάγνωση). Στην unsupervised αναγνώριση προτύπων (ομαδοποίηση) το μοντέλο είναι σχεδιασμένο ώστε να μπορεί να διερευνήσει τη δημιουργία ομάδων δεδομένων, χωρίς ωστόσο να έχει πρότερη γνώση των δεδομένων. Κάθε σύστημα αναγνώρισης προτύπων αποτελείται από τρία σημαντικά στάδια:

- α) Τμηματοποίηση εικόνας,
- β) Εξαγωγή χαρακτηριστικών,
- γ) Επιλογή χαρακτηριστικών και
- δ) Ταξινόμηση ή Ομαδοποίηση.

3.3 Τμηματοποίηση εικόνας

Οι κυτταρολογικές εικόνες μικροσκοπίου χαρακτηρίζονται από πολυμορφία φόντου και αποτελείται από τις κυτταρικές δομές, τις μεμβράνες και τον περιβάλλοντα ιστό. Σε αυτές τις εικόνες, οι πιο σημαντικές περιοχές είναι οι πυρήνες, οι οποίοι είναι αναγνωρισμένοι ως σημαντικοί δείκτες της κακοήθειας των όγκων. Προκειμένου να εξαχθούν χαρακτηριστικά γνωρίσματα από αυτές τις περιοχές ενδιαφέροντος, είναι

βασικό να διαιρεθεί η αρχική εικόνα στις χρήσιμες περιοχές (πυρήνες) και στις άχρηστες περιοχές (ιστός). Αυτή η διαδικασία ονομάζεται τμηματοποίηση της εικόνας και περιγράφεται αναλυτικά στο υποκεφάλαιο §4.3.

3.4 Εξαγωγή χαρακτηριστικών

Έχοντας τις τμηματοποιημένες περιοχές ενδιαφέροντος (πυρήνες), είναι εφικτός ο υπολογισμός χρήσιμων διαγνωστικών και προγνωστικών πληροφοριών. Για την περιγραφή του βαθμού κακοήθειας των πυρήνων έχουν αποδειχθεί αποτελεσματικές δυο κατηγορίες χαρακτηριστικών: (α) τα χαρακτηριστικά υφής και (β) τα μορφολογικά[18-21].

3.4.1 Χαρακτηριστικά 1^{ης} τάξης

Τα χαρακτηριστικά 1^{ης} τάξης χρησιμοποιούνται για να αποδώσουν την συχνότητα εμφάνισης κάθε επιπέδου του γκρι στους πυρήνες[22]. Τα χαρακτηριστικά αυτής της κατηγορίας που χρησιμοποιήθηκαν στην παρούσα εργασία περιγράφονται παρακάτω.

Η Μέση Τιμή (mean value):

(3.1)	<i>Μέση Τιμή (μ)</i>	$\mu = \frac{\sum_i \sum_j g(i, j)}{N}$
-------	----------------------	---

όπου $g(i,j)$ είναι η ένταση του εικονοστοιχείου στην θέση (i,j) και N είναι ο συνολικός αριθμός των εικονοστοιχείων.

Η Διακύμανση (variance), η οποία περιγράφει την διακύμανση του επιπέδου του γκρι της εικόνας από την μέση τιμή μ :

(3.2)	<i>Διακύμανση (σ^2)</i>	$\sigma^2 = \frac{\sum_i \sum_j (g(i, j) - \mu)^2}{N}$
-------	---	--

Η Ασυμμετρία (skewness), η οποία περιγράφει την κατανομή ασυμμετρίας γύρω από την μέση τιμή:

(3.3)	<i>Skewness (sk)</i>	$sk = \frac{\sum_i \sum_j (g(i, j) - \mu)^3}{N \cdot \sigma^3}$
-------	----------------------	---

Η Κύρτωση (kurtosis), η οποία περιγράφει την κατανομή ευκρίνειας εν συγκρίσει με την κανονική κατανομή:

(3.4)	<i>Κύρτωση (k)</i>	$k = \frac{\sum_i \sum_j (g(i, j) - \mu)^4}{N \cdot \sigma^4}$
-------	--------------------	--

3.4.2 Χαρακτηριστικά 2^{ης} τάξης

Τα χαρακτηριστικά 2^{ης} τάξης δεν δίνουν καμία πληροφορία σχετικά με τα επίπεδα του γκρι των πυρήνων, εν αντιθέσει με αυτά της 1^{ης} τάξης. Σε αυτή τη κατηγορία, η πληροφορία εξάγεται από της μήτρες co-occurrence και run-length[23]. Για κάθε χαρακτηριστικό που υπολογίζεται από αυτές τις δυο μήτρες, λήφθηκαν υπόψιν η μέση τιμή και το εύρος αυτού.

3.4.2.1 Εξαγωγή χαρακτηριστικών από την co-occurrence μήτρα

Η co-occurrence μήτρα εκφράζει τον τρόπο με τον οποίο εμφανίζονται σε μια εικόνα τα δυνατά επίπεδα γκρι. Είναι ένας πίνακας χωρικά εξαρτωμένων συχνοτήτων εμφάνισης επιπέδων φωτεινότητας και είναι συμμετρικός. Με βάση τον πίνακα αυτόν υπολογίζονται τα χαρακτηριστικά 2^{ης} τάξης, τα οποία εξαρτώνται από τη συχνότητα με την οποία δυο επίπεδα φωτεινότητας εμφανίζονται σε δυο εικονοστοιχεία που απέχουν συγκεκριμένη απόσταση d και έχουν συγκεκριμένη γωνιακή σχέση φ (0°, 45°, 90°, 135°) στην επιφάνεια της εικόνας[23].

Συγκεκριμένα, ο co-occurrence πίνακας αποτελείται από στοιχεία p_{ij} τα οποία εκφράζουν τις συχνότητες με τις οποίες εμφανίζονται δυο εικονοστοιχεία με τιμές φωτεινότητας i και j. Το παρακάτω παράδειγμα παρουσιάζει τον τρόπο, με τον οποίο υπολογίζεται ο co-occurrence πίνακας. Έστω ότι υπάρχει μια εικόνα (εικόνα 7α) με τέσσερα επίπεδα του γκρι. Στην μήτρα αυτή, το ζευγάρι (2,3), στην προς τα μπρος και προς τα πίσω σάρωση με κατεύθυνση φ και απόσταση d=1, θα έχει τιμή ίση με την συχνότητα της εμφάνισης δυο συνεχόμενων τόνων του γκρι με τιμές 2 και 3. Στην οριζόντια κατεύθυνση (0°), το ζευγάρι (2,3) εμφανίζεται δύο φορές, στην διαγώνια (45°) εμφανίζεται μια φορά, στην κάθετη κατεύθυνση (90°) εμφανίζεται καμία φορά και στην αντι-διαγώνια κατεύθυνση (135°) εμφανίζεται δύο φορές. Τα αποτελέσματα της δοθείσας εικόνας για την κατασκευή της co-occurrence μήτρας δίνονται στην εικόνα 7.

0	1	0	1
3	2	1	1
0	2	1	2
2	2	3	1

(α)

		Τόνοι του γκρι			
		0	1	2	3
Τόνοι του γκρι	0	(0,0)	(0,1)	(0,2)	(0,3)
	1	(1,0)	(1,1)	(1,2)	(1,3)
	2	(2,0)	(2,1)	(2,2)	(2,3)
	3	(3,0)	(3,1)	(3,2)	(3,3)

(β)

Co-occurrence 0°				Co-occurrence 45°				Co-occurrence 90°				Co-occurrence 135°			
0	3	1	0	0	0	2	0	0	1	1	2	0	1	2	0
3	2	3	1	0	4	2	1	1	4	3	1	1	4	2	0
1	3	2	2	2	2	2	1	1	3	4	0	2	2	0	2
0	1	2	0	0	1	1	0	2	1	0	0	0	0	2	0

(γ)

Εικόνα 7: Παράδειγμα υπολογισμού co-occurrence μήτρας: (α) η εικόνα, (β) η θέση για κάθε ζευγάρι στην co-occurrence μήτρα, (γ) η co-occurrence μήτρα για κάθε φ

Τα χαρακτηριστικά που μπορούν να εξαχθούν από την co-occurrence μήτρα περιγράφονται παρακάτω[23].

Η Γωνιακή Δεύτερη Στιγμή (Angular Second Moment - ASM), η οποία περιγράφει την ομαλότητα της εικόνας:

$$(3.5) \quad \text{Angular Second Moment (ASM)} \quad ASM = \sum_{i=0}^{N_x-1} \sum_{j=0}^{N_y-1} p(i, j)^2$$

όπου N_g είναι το πλήθος των επιπέδων του γκρι.

Η Αντίθεση (Contrast - CON), η οποία αυξάνει για υψηλής αντίθεσης πυρήνες:

$$(3.6) \quad \text{Contrast (CON)} \quad CON = \sum_{i=0}^{N_x-1} \sum_{j=0}^{N_y-1} (i-j)^2 p(i, j)$$

Η Συσχέτιση (Correlation - COR), η οποία πληροφορεί σχετικά με την συσχέτιση των γκρι τόνων των πυρήνων:

$$(3.7) \quad \text{Correlation (COR)} \quad COR = \frac{\sum_{i=0}^{N_x-1} \sum_{j=0}^{N_y-1} i j p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y}$$

όπου μ_x , μ_y , σ_x και σ_y οι αντίστοιχες μέσες τιμές και τυπικές αποκλίσεις των c_x και c_y , που δίνονται από τους τύπους:

$$(3.8) \quad c_x(i) = \sum_j p(i, j) \quad c_y(j) = \sum_i p(i, j)$$

Η Διακύμανση (Variance - VAR):

(3.9)

<i>Variance (VAR)</i>	$VAR = \sum_i \sum_j (i - \mu)^2 p(i, j)$
-----------------------	---

Η Αντίστροφη Διαφορετική Στιγμή (Inverse Different Moment - IDM), η οποία αυξάνει για χαμηλής αντίθεσης πυρήνες:

(3.10)

<i>Inverse Different Moment (IDM)</i>	$IDM = \frac{\sum_i \sum_j p(i, j)}{1 + (i - j)^2}$
---------------------------------------	---

Η Εντροπία (Entropy - ENT), η οποία πληροφορεί σχετικά με την τυχαιότητα και λαμβάνει χαμηλές τιμές για ομαλούς πυρήνες κυττάρων:

(3.11)

<i>Entropy (ENT)</i>	$ENT = - \sum_i \sum_j p(i, j) \log \{p(i, j)\}$
----------------------	--

3.4.2.2 Εξαγωγή χαρακτηριστικών από την run-length μήτρα

Η μήτρα run-length αποτελείται από στοιχεία r_{ij} τα οποία περιγράφουν την συχνότητα εμφάνισης μιας σειράς διαδοχικών εικονοστοιχείων, όπου έχουν το ίδιο επίπεδο γκρι[24]. Το παρακάτω παράδειγμα παρουσιάζει τον τρόπο, με τον οποίο υπολογίζεται ο run-length πίνακας. Έστω ότι υπάρχει μια εικόνα (εικόνα 8α) με τέσσερα επίπεδα του γκρι. Οι πιθανές γωνιακές σχέσεις μεταξύ των εικονοστοιχείων είναι η οριζόντια (0°), η διαγώνια (45°), η κάθετη (90°) και η αντι-διαγώνια (135°). Σαρώνοντας την εικόνα μόνο προς τα μπρος, παρατηρείται ότι το σύνολο των δυο συνεχόμενων εικονοστοιχείων που έχουν την τιμή 2 – για παράδειγμα – εμφανίζεται μια φορά στην οριζόντια κατεύθυνση (0°), μια φορά στην διαγώνια κατεύθυνση (45°), καμία φορά στην κάθετη κατεύθυνση (90°) και καμία φορά στην αντι-διαγώνια κατεύθυνση (135°) (σχήμα 8γ).

0	1	0	1
3	2	1	1
0	2	1	2
2	2	3	1

(α)

		Run Lengths			
		1	2	3	4
Τόνοι του γκρι	0	X	X	X	X
	1	X	X	X	X
	2	X	X	X	X
	3	X	X	X	X

(β)

Run-Length 0°				Run-Length 45°				Run-Length 90°				Run-Length 135°			
3	0	0	0	3	0	0	0	3	0	0	0	3	0	0	0
4	1	0	0	2	2	0	0	2	2	0	0	2	2	0	0
3	1	0	0	3	1	0	0	2	0	1	0	5	0	0	0
2	0	0	0	2	0	0	0	2	0	0	0	2	0	0	0

(γ)

Εικόνα 8: Παράδειγμα υπολογισμού run-length μήτρας: (α) η εικόνα, (β) η run-length μήτρα σχηματικά, (γ) η run-length μήτρα για κάθε γωνία φ

Τα χαρακτηριστικά που μπορούν να εξαχθούν από την run-length μήτρα περιγράφονται παρακάτω[24].

Η Έμφαση Μικρής Διαδρομής (Short Run Emphasis - SRE), η οποία δίνει έμφαση στις μικρές διαδρομές που οφείλεται στην ύπαρξη του j^2 στον παρονομαστή:

$$(3.12) \quad \begin{array}{l} \text{Short Run Emphasis} \\ \text{(SRE)} \end{array} \quad SRE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} \frac{r(i, j)}{j^2}}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} r(i, j)}$$

όπου N_g είναι το πλήθος των επιπέδων του γκρι και N_r είναι η μέγιστη πιθανή διαδρομή.

Η Έμφαση Μεγάλης Διαδρομής (Long Run Emphasis - LRE), η οποία δίνει έμφαση στις μεγάλες διαδρομές που οφείλεται στην ύπαρξη του j^2 στον αριθμητή:

$$(3.13) \quad \begin{array}{l} \text{Long Run Emphasis (LRE)} \end{array} \quad LRE = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} j^2 r(i, j)}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} r(i, j)}$$

Η Ανομοιογένεια Γκρι Επιπέδου (Gray Level Non-Uniformity - GLNU), η οποία αυξάνει για τους πυρήνες που έχουν πολλαπλές διαδρομές για τις ίδιες τιμές γκρι επιπέδου:

$$(3.14) \quad \begin{array}{l} \text{Gray Level Non-Uniformity} \\ \text{(GLNU)} \end{array} \quad GLNU = \frac{\sum_{i=1}^{N_g} \left(\sum_{j=1}^{N_r} r(i, j) \right)^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} r(i, j)}$$

Η Ανομοιογένεια Διαδρομής (Run Length Non-Uniformity - RLNU), η οποία παίρνει χαμηλές τιμές για πυρήνες με ομοιογενή κατανομή διαδρομής:

(3.15)	<i>Run Length Non-Uniformity (RLNU)</i>	$RLNU = \frac{\sum_{j=1}^{N_r} \left(\sum_{i=1}^{N_g} r(i, j) \right)^2}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} r(i, j)}$
--------	---	---

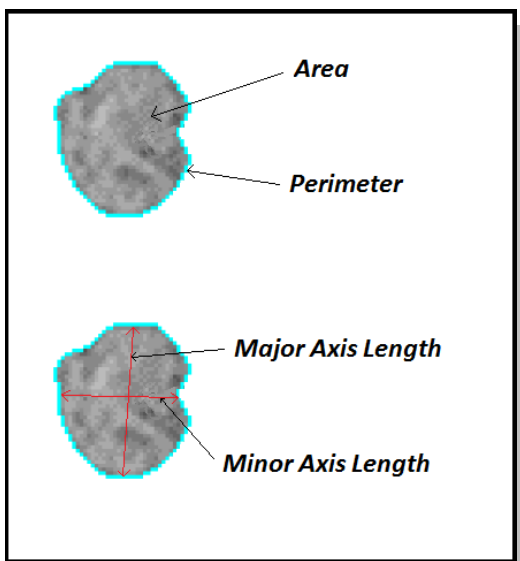
Το Ποσοστό Διαδρομών (Run Percentage - RP), το οποίο παίρνει χαμηλότερες τιμές όταν οι πυρήνες έχουν γραμμικές δομές:

(3.16)	<i>Run Percentage (RP)</i>	$RP = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_r} r(i, j)}{P}$
--------	----------------------------	--

όπου P είναι ο πιθανός συνολικός αριθμός διαδρομών στην εικόνα του πυρήνα.

3.4.3 Μορφολογικά χαρακτηριστικά

Η ποσοτικοποίηση των μορφολογικών χαρακτηριστικών των πυρήνων μπορεί να περιέχει σημαντικές διαγνωστικές πληροφορίες σχετικά με τον βαθμό του όγκου, καθώς στους υψηλού βαθμού όγκους συνήθως υπάρχουν γιγάντιοι πυρήνες με ακανόνιστο σχήμα και κυρτότητα, ενώ οι χαμηλού βαθμού όγκοι είναι συνήθως μικρότεροι και πιο στρογγυλοί. Τα μορφολογικά χαρακτηριστικά που χρησιμοποιήθηκαν στην παρούσα εργασία είναι οι μετρήσεις περιοχής (area), η περίμετρος (perimeter), το μήκος ελάχιστου άξονα (minor axis length), το μήκος μέγιστου άξονα (major axis length) και η εκκεντρικότητα (eccentricity).



Εικόνα 9: Μορφολογικά χαρακτηριστικά πυρήνων

Η μέτρηση της περιοχής (area) επιτυγχάνεται με την καταμέτρηση των εικονοστοιχείων που βρίσκονται μέσα στα όρια του πυρήνα[25].

Η περίμετρος (perimeter) υπολογίζεται ως το μήκος του περιγράμματος του πυρήνα.[26]

Το μήκος ελάχιστου άξονα (minor axis length) υπολογίζεται ως το μήκος του ελάχιστου άξονα της έλλειψης που έχει την ίδια κανονικοποιημένη δεύτερη κεντρική ροπή (second central moment) με την περιοχή του πυρήνα.[26]

Αντίστοιχα, το μήκος μέγιστου άξονα (major axis length) υπολογίζεται ως το μήκος του μέγιστου άξονα της έλλειψης που έχει την ίδια κανονικοποιημένη δεύτερη κεντρική ροπή (second central moment) με την περιοχή του πυρήνα.[26]

Η εκκεντρικότητα (eccentricity) είναι ο λόγος της απόστασης μεταξύ των εστιών της έλλειψης και του μήκους μέγιστου άξονα.[26]

3.5 Επιλογή χαρακτηριστικών

Κατά την διαδικασία της ταξινόμησης, η ποσοτική είσοδος, δηλαδή τα χαρακτηριστικά, μετατρέπεται σε ποιοτική έξοδο, δηλαδή τη διάγνωση, την πρόγνωση κλπ. Η έξοδος του ταξινομητή μπορεί να είναι είτε μια ξεχωριστή τιμή, η οποία αντιστοιχεί σε μια από τις ήδη προκαθορισμένες κλάσεις, είτε ένα διάνυσμα πραγματικών τιμών, το οποίο αντικατοπτρίζει την πιθανότητα το πρότυπο να προέρχεται από μια συγκεκριμένη κλάση. Στην προσπάθεια μεγιστοποίησης της απόδοσης του ταξινομητή, ο καλύτερος συντονισμός της κατάτμησης (segmentation), της εξαγωγής χαρακτηριστικών και της επιλογής αυτών, αποτελεί βασικότερη προϋπόθεση. Η δημιουργία ενός ταξινομητή απαιτεί τρία στάδια:

- Το στάδιο εκμάθησης (training stage),
- Το στάδιο της αξιολόγησης επιδόσεων,
- Το στάδιο των δοκιμών (testing stage).

Στο στάδιο εκμάθησης, ο ταξινομητής εκπαιδεύεται για να μάθει τις διαφορές μεταξύ των γνωστών κλάσεων βάσει των διανυσμάτων των χαρακτηριστικών γνωρισμάτων γνωστής ταξινόμησης, ως είσοδο. Βασιζόμενος σε αυτά τα γνωστά διανύσματα, ο ταξινομητής δημιουργεί μια συγκεκριμένη περιγραφή για κάθε κλάση. Όταν η εκμάθηση ολοκληρωθεί, ο ταξινομητής είναι έτοιμος να ταξινομήσει νέα δεδομένα.

Παρόλο που με μια πρώτη σκέψη, φαίνεται πιο λογικό ότι ένας μεγαλύτερος αριθμός χαρακτηριστικών θα είναι κατατοπιστικός από όσο ένας μικρότερος αριθμός χαρακτηριστικών, ωστόσο, αυτό δεν συμβαίνει για τρεις κυρίως λόγους[22,27]. Αρχικά, το υπολογιστικό κόστος και η πολυπλοκότητα του ταξινομητή αυξάνεται σημαντικά. Έπειτα, έχει αποδειχθεί ότι μετά από ένα αριθμό συνδυασμών χαρακτηριστικών, η ακρίβεια του ταξινομητή μειώνεται[27]. Τέλος, σε περίπτωση περιορισμένου αριθμού των διαθέσιμων άγνωστων δεδομένων, είναι πιθανό κάποια χαρακτηριστικά με μικρή διακριτική δύναμη, να προκαλέσουν αύξηση του θορύβου και να μειώσουν την γενικευτική ικανότητα του ταξινομητή[27]. Κατά επέκταση, η επιλογή των χαρακτηριστικών αποτελεί ένα πολύ σημαντικό βήμα για την ανάδειξη των πιο κατατοπιστικών χαρακτηριστικών για βέλτιστη ρύθμιση της ικανότητας του ταξινομητή στην ταξινόμηση άγνωστων δεδομένων.

Η βέλτιστη διαδικασία επιλογής χαρακτηριστικών είναι ο αλγόριθμος πλήρους αναζήτησης (exhaustive search algorithm). Σύμφωνα με την μέθοδο αυτή, τα χαρακτηριστικά συνδυάζονται με όλους τους δυνατούς τρόπους, δηλαδή συνδυασμοί

ανά 2, 3 και μέχρι τον συνολικό αριθμό των χαρακτηριστικών. Το καλύτερο διάνυσμα χαρακτηριστικών θεωρείται εκείνο, το οποίο οδηγεί στο ελάχιστο σφάλμα ταξινόμησης. Αυτό το σφάλμα ταξινόμησης μπορεί να επικυρωθεί με την χρήση μεθόδων, όπως η self-consistency, η cross-validation ή η leave-some-out, τα οποία αποτελούν κλασικές μέθοδοι για την αξιολόγηση της απόδοσης ενός ταξινομητή[22,27].

Η self-consistency μέθοδος εκτιμά την αξιοπιστία του ταξινομητή, χρησιμοποιώντας όλα τα διαθέσιμα δεδομένα εκπαίδευσης τόσο για την εκπαίδευση όσο και για την δοκιμή και τον έλεγχο του ταξινομητή. Αυτό έχει σαν αποτέλεσμα η μέθοδος αυτή να αποτελεί μια προκατειλημμένη εκτίμηση του ταξινομητή, καθώς τα χαρακτηριστικά που χρησιμοποιούνται για την δοκιμή του ταξινομητή είναι τα ίδια με αυτά που χρησιμοποιήθηκαν στην εκπαίδευση αυτού.

Η k-fold cross-validation είναι μια αρκετά δημοφιλής μέθοδος, σύμφωνα με την οποία τα δεδομένα χωρίζονται σε k τυχαία training σετ – συνήθως k=10, τα οποία χρησιμοποιούνται για την εκπαίδευση του ταξινομητή, ενώ τα υπόλοιπα χρησιμοποιούνται για την δοκιμή/έλεγχο του ταξινομητή. Αυτή η διαδικασία επαναλαμβάνεται k φορές[28].

3.5.1 Ταξινόμηση προτύπων - Supervised αλγόριθμοι

Ως supervised αλγόριθμοι ταξινόμησης ορίζονται οι αλγόριθμοι, οι οποίοι για την εκπαίδευσή τους χρησιμοποιούν πρότυπα των οποίων η κλάση είναι γνωστή εκ των προτέρων.

Οι supervised αλγόριθμοι ταξινόμησης που θα παρουσιαστούν σε αυτήν την εργασία είναι οι εξής:

- α) Minimum Distance (MDC)
- β) Least Squares Minimum Distance (LSMD)
- γ) Bayesian
- δ) K-Nearest Neighbors (k-NN)
- ε) Probabilistic Neural Network (PNN)
- στ) Support Vector Machine (SVM)

3.5.1.1 Minimum Distance (MDC)

Ο Minimum Distance ταξινομητής είναι ένας από τους πιο απλούς ταξινομητές. Με βάση την ελάχιστη απόσταση από τα κέντρα των ήδη υπαρχόντων κλάσεων αποφασίζεται το πού θα ταξινομηθεί το νέο χαρακτηριστικό. Συνήθως, ως κέντρο μιας κλάσης χρησιμοποιείται το διάνυσμα της μέσης τιμής της κλάσης αυτής. Οι δημοφιλέστεροι τρόποι μέτρησης της απόστασης ενός διανύσματος από μια κλάση είναι η Ευκλείδεια απόσταση και η απόσταση Mahalanobis. Η συνάρτηση διάκρισης του ταξινομητή της ελάχιστης απόστασης με χρήση της Ευκλείδειας απόστασης δίνεται από τον τύπο [22]:

$$(3.17) \quad g_j(x) = x^T m_j - \frac{1}{2} (m_j m_j^T)$$

Αντίστοιχα, η συνάρτηση διάκρισης του ταξινομητή της ελάχιστης απόστασης με χρήση της απόστασης Mahalanobis δίνεται από τον τύπο:

$$(3.18) \quad g_j(x) = \left((\underline{x} - \underline{m}_j)^T \Sigma^{-1} (\underline{x} - \underline{m}_j) \right)^{\frac{1}{2}}$$

όπου x είναι το διάνυσμα του νέου προτύπου και m_j η μέση τιμή της κλάσης j .

3.5.1.2 Least Squares Minimum Distance (LSMD)

Ο Minimum Distance ταξινομητής υποθέτει ότι όλα τα δείγματα που ανήκουν σε κάποια κατηγορία είναι δυνατόν να αντιπροσωπευτούν από την μέση τιμή αυτής της τάξης. Ωστόσο, αυτή η υπόθεση δεν είναι πάντα αντιπροσωπευτική, διότι συχνά συμβαίνει διαφορετικές κατηγορίες να επικαλύπτονται. Μια μέθοδος για να αντιμετωπιστεί αυτός ο περιορισμός είναι η χρήση του Least Squares Minimum Distance ταξινομητή. Με τον ταξινομητή αυτόν όλα τα πρότυπα μιας κλάσης ωθούνται να ομαδοποιηθούν γύρω από ένα προκαθορισμένο σημείο σε έναν χώρο «απόφασης». Η ταξινόμηση του άγνωστου προτύπου βασίζεται στην ελάχιστη του απόσταση από όλα τα προκαθορισμένα σημεία. Ο Least Squares μετασχηματισμός στον χώρο «απόφασης» δίνεται από τον τύπο:

$$(3.19) \quad A = \left[\sum_{j=1}^{N_i} V_i (x_{i,j}^T) \right] \left[\sum_{j=1}^{N_i} (x_{i,j} x_{i,j}^T) \right]^{-1}$$

και πρέπει

$$(3.20) \quad e_j = \tilde{x}_{i,j} - V_i = Ax_{i,j} - V_i$$

όπου V είναι το προκαθορισμένο σημείο γύρω από το οποίο τα δεδομένα ομαδοποιούνται, i είναι ο αριθμός της κλάσης, $x_{i,j}$ είναι τα χαρακτηριστικά διανύσματα εισόδου της αντίστοιχης διαστατικότητας $i=1, \dots, N_i$, $\tilde{x}_{i,j}$ είναι το μετασχηματισμένο χαρακτηριστικό διάνυσμα και e_i είναι το λάθος. Η συνάρτηση διάκρισης του LSMD δίνεται από τον τύπο:

$$(3.21) \quad g_i(x) = \sum_{i=1}^M a_{ci} x_i - b_c$$

όπου M είναι ο αριθμός των χαρακτηριστικών, a_c και b_c είναι οι συντελεστές βάρους που προκύπτουν από την συνάρτηση μετασχηματισμού A .

3.5.1.3 Bayesian

Ο Bayesian ταξινομητής δίνει την ελάχιστη πιθανότητα σφάλματος για τα δεδομένα, ακολουθώντας μια συγκεκριμένη κατανομή, η οποία συνήθως επιλέγεται να είναι η Gaussian κατανομή. Το άγνωστο πρότυπο x^T ταξινομείται στην κλάση j με την μέγιστη τιμή r_j . Η συνάρτηση διάκρισης του Bayesian ταξινομητή με Gaussian κατανομή για ισοπίθανα δεδομένα δίνεται από τον τύπο[29]:

$$(3.22) \quad g_j = x^T \Sigma^{-1} m_j - \frac{1}{2} m_j^T \Sigma^{-1} m_j$$

όπου m_j είναι η μέση τιμή και Σ είναι η μήτρα συνδιακύμανσης.

3.5.1.4 K-Nearest Neighbors (k-NN)

Ο K-Nearest Neighbors (k-NN) ταξινομητής ακολουθεί την παρακάτω διαδικασία: Αρχικά, εντοπίζονται οι k πλησιέστεροι γείτονες του άγνωστου προτύπου και στην συνέχεια το πρότυπο ταξινομείται στην κλάση j , αν η πλειοψηφία των k γειτόνων ανήκουν στην κλάση j [29].

3.5.1.5 Probabilistic Neural Network (PNN)

Μια από τις πιο δημοφιλείς προσεγγίσεις για την εκτίμηση της συνάρτησης πυκνότητας πιθανότητας (PDF) έχει προταθεί από τον Parzen[30]. Η εκτίμηση αυτή δίνεται από την σχέση:

$$(3.23) \quad g_j(x) = \frac{1}{(2\pi)^{p/2} \sigma^p N_j} \sum_{i=1}^{N_j} w(y_i)$$

όπου N είναι ο αριθμός των δειγμάτων x_1, \dots, x_N και $w(y_i)$ είναι η συνάρτηση kernel. Οι συνήθεις επιλογές για την συνάρτηση kernel είναι η Gaussian, η Exponential και η Reciprocal.

Η Gaussian δίνεται από την σχέση:

$$w(y_i) = e^{-\frac{\|x - x_i\|^2}{2\sigma^2}} \quad (3.24)$$

η οποία οδηγεί στην PDF:

$$g_j(x) = \frac{1}{(2\pi)^{p/2} \sigma^p N_j} \sum_{i=1}^{N_j} e^{-\frac{\|x - x_i\|^2}{2\sigma^2}} \quad (3.25)$$

Η Exponential δίνεται από την σχέση:

$$w(y_i) = e^{-\frac{\|x - x_i\|}{\sigma}} \quad (3.26)$$

η οποία οδηγεί στην PDF:

$$g_j(x) = \frac{1}{(2\pi)^{p/2} \sigma^p N_j} \sum_{i=1}^{N_j} e^{-\frac{\|x-x_i\|^2}{\sigma^2}} \quad (3.27)$$

Η Reciprocal δίνεται από την σχέση:

$$w(y_i) = \frac{1}{1 + \|x-x_i\|^2/\sigma^2} \quad (3.28)$$

η οποία οδηγεί στην PDF:

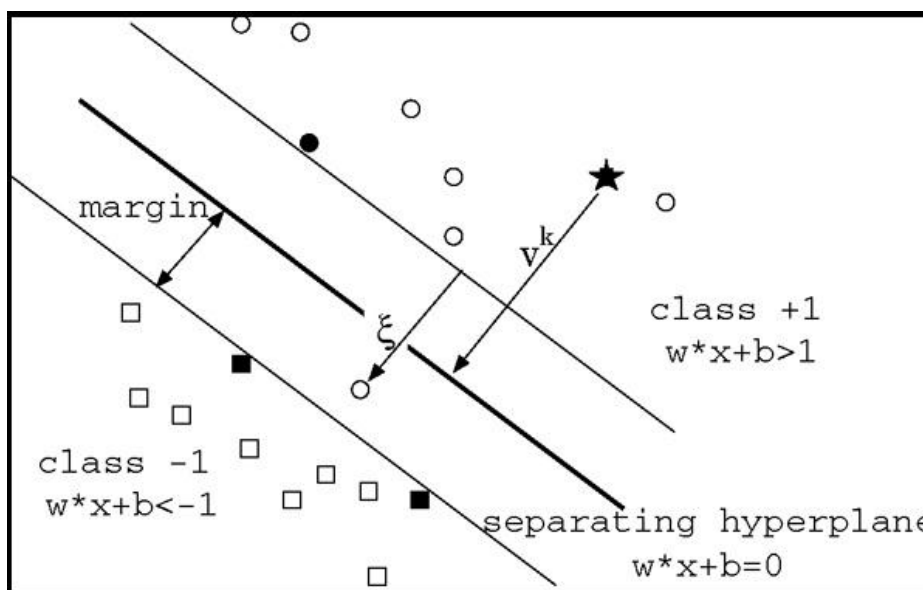
$$g_j(x) = \frac{1}{(2\pi)^{p/2} \sigma^p N_j} \sum_{i=1}^{N_j} \frac{1}{1 + \|x-x_i\|^2/\sigma^2} \quad (3.29)$$

όπου x είναι το άγνωστο πρότυπο, x_i είναι το i -ιστό διάνυσμα του προτύπου εκπαίδευσης, N_j είναι ο αριθμός των προτύπων της κλάσης j , σ είναι μια παράμετρος εξομάλυνσης και p είναι ο αριθμός των χαρακτηριστικών των προτύπων.

Η παράμετρος σ επηρεάζει την εκτίμηση με τον εξής τρόπο: Οι πολύ μεγάλες τιμές του σ εξομαλύνουν σε μεγάλο βαθμό τις λεπτομέρειες, ενώ οι πολύ μικρές το άκρως αντίθετο.

3.5.1.6 Support Vector Machine (SVM)

Στην περίπτωση των γραμμικά διαχωρίσιμων προτύπων, ο γραμμικός SVM βρίσκει, ανάμεσα σε όλα τα υπερεπίπεδα που ελαχιστοποιούν το λάθος της εκπαίδευσης, το υπερεπίπεδο εκείνο που διαχωρίζει πλήρως τα πρότυπα εκπαίδευσης, έτσι ώστε να απέχει ίσες αποστάσεις (περιθώρια-margin) από τα πλησιέστερα πρότυπα των κλάσεων και παράλληλα το περιθώριο αυτό να είναι το μέγιστο δυνατό. Τα πρότυπα των κλάσεων που απέχουν απόσταση ίση με το περιθώριο από το διαχωριστικό υπερεπίπεδο λέγονται διανύσματα στήριξης (support vectors).



Εικόνα 10: Δύο διαχωριστικές ευθείες σε δισδιάστατο χώρο. Εικονίζεται το διαχωριστικό υπερεπίπεδο που δίνει το μέγιστο δυνατό περιθώριο[31]

Το διαχωριστικό αυτό υπερεπίπεδο δίνεται από την σχέση:

$$(3.30) \quad y_i [w^T x_i + b] \geq 1, \quad i=1, \dots, n$$

όπου x_i είναι το πρότυπο εκπαίδευσης που ανήκει στην κλάση $y_i = \{-1, +1\}$, w είναι ένα κάθετο διάνυσμα στα υπερεπίπεδα και b το όριο που περιγράφει την απόσταση του υπερεπιπέδου από το $O(0,0)$, η οποία ισούται με $b/||w||$.

Το διαχωριστικό υπερεπίπεδο ελαχιστοποιεί το $||w||^2$. Πρόκειται για ένα πρόβλημα βελτιστοποίησης με περιορισμούς ανισότητας, το οποίο μπορεί να λυθεί με χρήση της Lagrangian μεθόδου, μεγιστοποιώντας την $L_d(\alpha)$ που δίνεται από την παρακάτω σχέση:

$$(3.31) \quad L_d(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n y_i y_j \alpha_i \alpha_j x_i^T x_j$$

Υπό τις συνθήκες:

$$(3.32) \quad \sum_{i=1}^n \alpha_i y_i = 0 \quad \alpha_i \geq 0$$

Το πρόβλημα θα μπορούσε να διατυπωθεί και με την χρήση μήτρας, ως εξής:

$$(3.33) \quad L_d(\alpha) = -0.5 \alpha^T H \alpha + f^T \alpha$$

υπό τους παρακάτω περιορισμούς:

$$(3.34) \quad y^T \alpha = 0 \quad \alpha \geq 0$$

όπου $(\alpha)_i = \alpha_i$, H είναι η Hessian μήτρα (η οποία υπολογίζεται ως: $H_{ij} = y_i y_j (x_i^T x_j)$) και f είναι το μοναδιαίο διάνυσμα με μέγεθος ίσο με το μέγεθος των δεδομένων εκπαίδευσης.

Ο μη γραμμικός SVM ταξινομητής έχει σχεδιαστεί για μη γραμμικά διαχωρίσιμα δεδομένα. Ο τρόπος, με τον οποίο λειτουργεί ο ταξινομητής αυτός, ακολουθεί δύο βασικά βήματα. Αρχικά, σχεδιάζει τα χαρακτηριστικά εισόδου σε ένα υψηλότερου βαθμού διαστατικότητας χώρο, χρησιμοποιώντας μια μη γραμμική συνάρτηση μεταφοράς (kernel). Στόχος είναι τα χαρακτηριστικά να σχεδιαστούν σε ένα χώρο, όπου θα μπορούν να είναι γραμμικά διαχωρίσιμα. Τέλος, σε αυτόν τον νέο χώρο αναζητείται το υπερεπίπεδο που θα μεγιστοποιεί το περιθώριο από τα πλησιέστερα πρότυπα των κλάσεων, όπως προηγουμένως. Η συνάρτηση του μη γραμμικού SVM ταξινομητή 2-κλάσεων δίνεται από την σχέση:

$$(3.35) \quad L_d(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n y_i y_j \alpha_i \alpha_j K(x_i, x_j)$$

όπου K είναι η συνάρτηση μεταφοράς (kernel), η οποία μπορεί να πάρει την μορφή ενός πολυωνυμικού πυρήνα ή ενός Gaussian πυρήνα, όπως φαίνεται παρακάτω.

(3.36)	Πολυωνυμικός πυρήνας	$K(x, y) = (x \cdot y + 1)^D$
--------	----------------------	-------------------------------

3.5.2 Ομαδοποίηση προτύπων - Unsupervised αλγόριθμοι

Ως unsupervised αλγόριθμοι ταξινόμησης ορίζονται οι αλγόριθμοι στους οποίους δεν είναι γνωστό το πλήθος των κλάσεων και παράλληλα δεν υπάρχουν γνωστά πρότυπα για την εκπαίδευση του ταξινομητή.

Οι unsupervised αλγόριθμοι ταξινόμησης που θα παρουσιαστούν σε αυτήν την εργασία είναι οι εξής:

- α) k-means clustering
- β) Fuzzy c-means clustering

3.5.2.1 K-means clustering

Ο αλγόριθμος k-means ομαδοποιεί τα δεδομένα εισόδου σε k διαφορετικές ομάδες, ελαχιστοποιώντας το ακόλουθο κριτήριο:

(3.37)	$J = \sum_{j=1}^k \sum \ x_n - m_j\ ^2$
--------	---

όπου x_n είναι το n -ιοστό διάνυσμα εισόδου στον ταξινομητή και m_j είναι η μέση τιμή για κάθε μια από τις k κλάσεις. Ο αλγόριθμος ξεκινά θεωρώντας τυχαία έναν εκπρόσωπο από τα πρότυπα εισόδου για κάθε κλάση. Αφού τα υπόλοιπα πρότυπα ταξινομηθούν στις k κλάσεις σύμφωνα με το παραπάνω κριτήριο, ο εκπρόσωπος κάθε κλάσης παίρνει την τιμή της μέσης τιμής των προτύπων της κλάσης αυτής και όλα τα πρότυπα ξαναταξινομούνται στις k κλάσεις με τον ίδιο τρόπο για τους νέους εκπροσώπους των κλάσεων. Η διαδικασία αυτή επαναλαμβάνεται έως ότου δεν θα υπάρχουν πλέον αλλαγές σε καμία από τις k κλάσεις[32].

3.5.2.2 Fuzzy c-means clustering

Στον αλγόριθμο Fuzzy c-means, κάθε πρότυπο έχει έναν βαθμό συσχέτισης για κάθε κλάση, που δείχνει το κατά πόσο ανήκει σε αυτήν ή όχι. Κατά επέκταση, τα πρότυπα που βρίσκονται στα όρια της κλάσης έχουν μικρότερο βαθμό συσχέτισης για αυτήν την κλάση σε σχέση με τα πρότυπα που βρίσκονται στο κέντρο της κλάσης αυτής. Κάθε πρότυπο έχει ένα σύνολο συντελεστών που ορίζουν τον βαθμό συσχέτισής του με την κλάση k . Ως κέντρο μιας κλάσης ορίζεται η μέση τιμή όλων των προτύπων της κλάσης αυτής, σταθμισμένη από τον βαθμό συσχέτισης με την κλάση αυτή, σύμφωνα με την παρακάτω σχέση[32]:

(3.38)

$$c_j = \frac{\sum_{i=1}^N u_{ij}^m \cdot x_i}{\sum_{i=1}^N u_{ij}^m}$$

όπου u_{ij} είναι ο βαθμός συσχέτισης με την κλάση j και $1 \leq m < \infty$.

Ο βαθμός συσχέτισης με την κλάση j σχετίζεται άμεσα με την απόσταση του προτύπου από το κέντρο της κλάσης, όπως φαίνεται παρακάτω:

(3.39)

$$u_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{\frac{2}{m-1}}}$$

όπου m μια παράμετρος, η οποία ελέγχει το βάρος που δίνεται για το κοντινότερο κέντρο.

3.5.3 Semi-supervised αλγόριθμοι

Η εκπαίδευση ενός semi-supervised αλγορίθμου αποτελείται από έναν συνδυασμό γνωστών και άγνωστων δεδομένων, αξιοποιώντας τα πλεονεκτήματα τόσο των supervised όσο και των unsupervised αλγορίθμων. Ο αλγόριθμος που υλοποιήθηκε και χρησιμοποιήθηκε στα πλαίσια αυτής της εργασίας είναι ο k-nearest neighbors με Co-Training (k-NN with Co-Training).

3.5.3.1 K-nearest neighbors με Co-Training

Σύμφωνα με τον αλγόριθμο αυτόν, υπάρχει ένα σύνολο γνωστών δεδομένων και ένα σύνολο άγνωστων δεδομένων U , τα οποία ζητείται να ταξινομηθούν. Δημιουργείται ένα μικρότερο τυχαίο υποσύνολο U' του συνόλου U , το οποίο περιέχει u άγνωστα δεδομένα. Αρχικά, χρησιμοποιείται το σύνολο L για την εκπαίδευση δυο k-NN ταξινομητών: h_1 και h_2 . Ο h_1 ταξινομητής χρησιμοποιεί κάποια από τα χαρακτηριστικά των προτύπων και ο h_2 ταξινομητής χρησιμοποιεί τα υπόλοιπα. Οι δύο αυτοί ταξινομητές εξετάζουν τα άγνωστα δεδομένα του U' συνόλου και επιλέγονται p δεδομένα, τα οποία χαρακτηρίζονται με την περισσότερη βεβαιότητα ως θετικά (positive) και n δεδομένα, τα οποία χαρακτηρίζονται με την περισσότερη βεβαιότητα ως αρνητικά (negative). Κάθε πρότυπο, που επιλέγεται με αυτόν τον τρόπο ($p+n$), προστίθεται στο σύνολο L με τα γνωστά δεδομένα. Τέλος, το σύνολο U' ανανεώνεται τοποθετώντας σε αυτόν $2p+2n$ δεδομένα από το σύνολο των άγνωστων δεδομένων U και η διαδικασία επαναλαμβάνεται[33-36].

Ο k-NN Co-Training αλγόριθμος παρουσιάζεται στο παρακάτω πλαίσιο:

Δίνεται:

- Σύνολο L γνωστών προτύπων
- Σύνολο U άγνωστων προτύπων

Δημιουργείται U' , επιλέγοντας τυχαία u πρότυπα από το U .

Για k επαναλήψεις:

- ✓ Χρήση του L για την εκπαίδευση του h_1 ταξινομητή, χρησιμοποιώντας x_1 από τα χαρακτηριστικά
- ✓ Χρήση του L για την εκπαίδευση του h_2 ταξινομητή, χρησιμοποιώντας x_2 από τα χαρακτηριστικά
- ✓ Ο h_1 και ο h_2 ταξινομούν p θετικά και n αρνητικά πρότυπα από το U'
- ✓ Αυτά προσθέτονται στον L ως πλέον γνωστά πρότυπα
- ✓ Τυχαία επιλέγονται $2p+2n$ πρότυπα από τον U και ανανεώνουν τον U' .

4. ΑΠΟΤΕΛΕΣΜΑΤΑ

4.1 Γενικά

Το πειραματικό μέρος της παρούσας εργασίας χωρίζεται σε επτά κύρια μέρη.

Αρχικά, πραγματοποιήθηκε *συλλογή δεδομένων*, όπου δείγματα ιστοπαθολογικού ιστού συλλέχτηκαν από 35 ασθενείς από το Πανεπιστημιακό Νοσοκομείο της Πάτρας, οι οποίοι είχαν διαγνωστεί με αστροκυτταρικό όγκο.

Δεύτερον, έγινε η *προετοιμασία των δεδομένων* κατά την οποία τα δείγματα ιστού χρωματίστηκαν με H&E (Hematoxylin & Eosin), ώστε να φανερωθούν τα διάφορα συστατικά των κυττάρων και των πυρήνων αλλά και για να διακριθούν μεταξύ τους.

Τρίτον, πραγματοποιήθηκε *διάγνωση δεδομένων* κατά την οποία ένας ιστοπαθολόγος εξέτασε τις χρωματισμένες εικόνες των δειγμάτων ιστού και πραγματοποίησε διάγνωση με βάση ιστολογικά κριτήρια σε δύο βαθμούς κακοήθειας (I, II ή III, IV) σύμφωνα με τον Παγκόσμιο Οργανισμό Υγείας (World Health Organization - WHO). Επίσης, ο ίδιος ιστοπαθολόγος σημείωσε πάνω στις πλάκες των δειγμάτων την πιο αντιπροσωπευτική περιοχή (Region Of Interest - ROI) για την εξέταση αυτών.

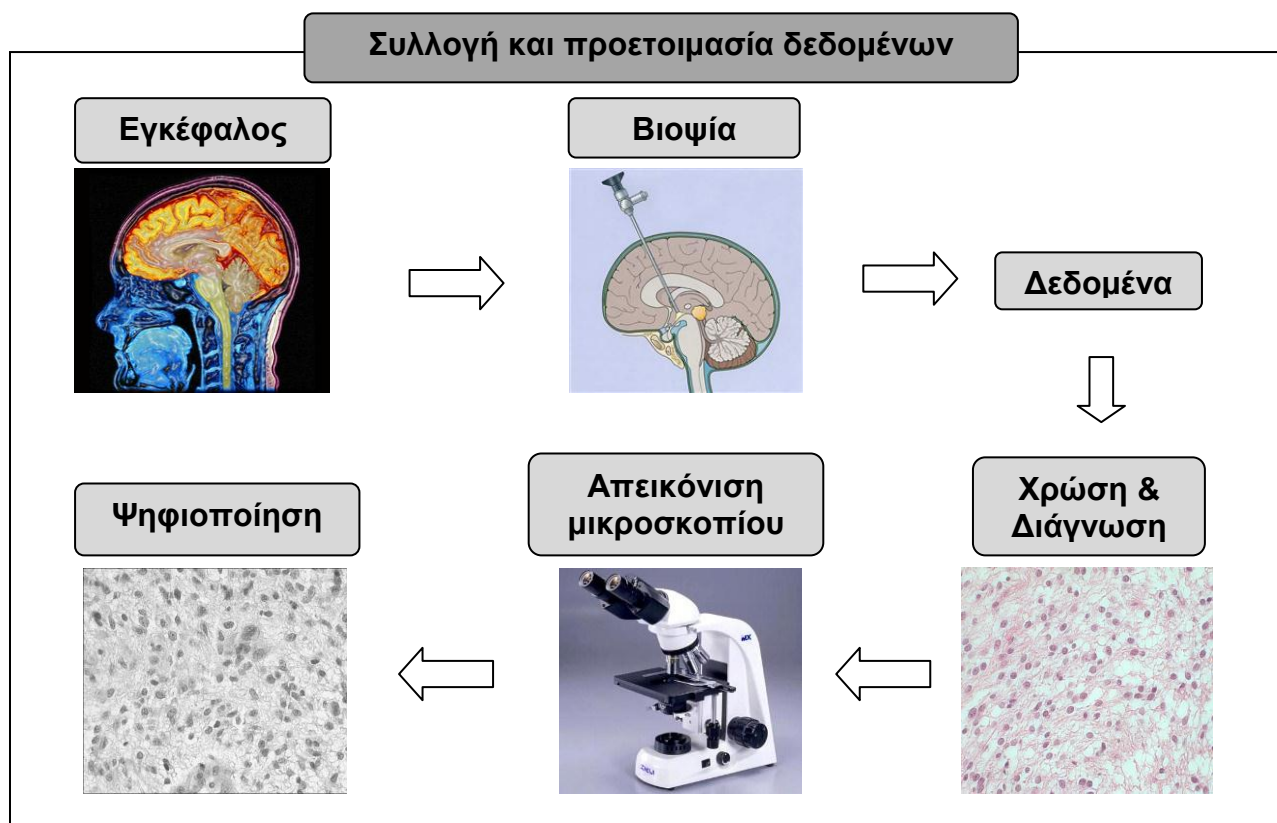
Τέταρτον, ακολούθησε η *ψηφιοποίηση εικόνας*, όπου από κάθε δείγμα ψηφιοποιήθηκαν 6 εικόνες, η οποίες πάρθηκαν από το προκαθορισμένο ROI.

Πέμπτον, εφαρμόστηκε *τμηματοποίηση πυρήνων*. Το σημαντικότερο στοιχείο μιας τέτοιας εικόνας, από το οποίο μπορούν να εξαχθούν συμπεράσματα για την διάγνωση, είναι οι πυρήνες. Έτσι, δημιουργήθηκαν αλγόριθμοι, οι οποίοι απομάκρυναν όλες τις υπόλοιπες περιοχές της εικόνας και τμηματοποίησαν τους πυρήνες για περισσότερη ανάλυση αυτών.

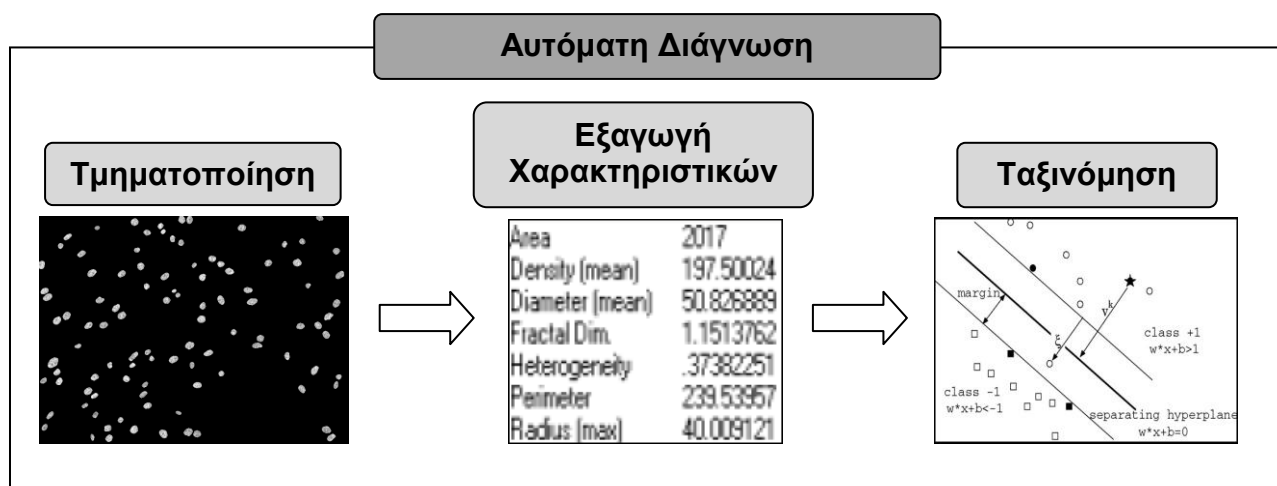
Έκτον, πραγματοποιήθηκε *εξαγωγή χαρακτηριστικών* από τους τμηματοποιημένους πυρήνες. Οι βιολογικές πληροφορίες που υπάρχουν στους πυρήνες κωδικοποιήθηκαν σε ένα σύνολο μορφολογικών χαρακτηριστικών και χαρακτηριστικών υφής, τα οποία έχουν αποδειχθεί αποτελεσματικές κατηγορίες για την περιγραφή του βαθμού κακοήθειας των πυρήνων.

Τέλος, χρησιμοποιήθηκε *αναγνώριση προτύπων και ταξινόμηση* αυτών για αυτόματη διάγνωση. Με βάση τα παραπάνω χαρακτηριστικά, δημιουργήθηκαν και εκπαιδεύτηκαν Supervised, Unsupervised και Semi-supervised συστήματα ταξινόμησης προτύπων, ώστε να μάθουν τις διαφορές μεταξύ των βαθμών κακοήθειας I, II, III και IV.

Όλοι οι αλγόριθμοι που δημιουργήθηκαν στα πλαίσια της παρούσας εργασίας εφαρμόστηκαν στο λογισμικό MATLAB (The MathWorks Inc., Version R2008a).



Εικόνα 11: Συλλογή και προετοιμασία δεδομένων



Εικόνα 12: Αυτόματη Διάγνωση

4.2 Κλινικό υλικό

Το κλινικό υλικό που χρησιμοποιήθηκε προήλθε από 35 ασθενείς που είχαν υποβληθεί σε χειρουργική επέμβαση στο Πανεπιστημιακό Νοσοκομείο της Πάτρας. Ακολουθήθηκε χρώση των ιστών με το μίγμα Αιματοξυλίνης - Ηωσίνης (Hematoxylin & Eosin - H&E). Οι βιοψίες κατηγοριοποιήθηκαν σε τρεις βαθμούς κακοήθειας (I, II, III ή IV) σύμφωνα με τον Παγκόσμιο Οργανισμό Υγείας (World Health Organization - WHO), από ιστοπαθολόγο ιατρό (Ιατρός Παναγιώτα Ραβαζούλα). Από τις 35 βιοψίες, 19 χαρακτηρίστηκαν από ιστοπαθολόγο ιατρό ως χαμηλής κακοήθειας (βαθμός I ή II) και 16 ως υψηλής κακοήθειας (βαθμός III ή IV) και σημειώθηκαν οι περιοχές ενδιαφέροντος από τις οποίες, μέσω οπτικού μικροσκοπίου, εξήχθησαν πολλαπλές εικόνες για κάθε ασθενή με μεγέθυνση 40x για περαιτέρω επεξεργασία και απομόνωση των πυρήνων.

Πίνακας 1: Βάση δεδομένων για το υλικό της βιοψίας

Βάση Δεδομένων	Χαμηλός Βαθμός	Υψηλός Βαθμός
	Grade I ή II	Grade III ή IV
Πανεπιστημιακό Νοσοκομείο Πάτρας	19	16

Ύστερα από πραγματοποίηση επανάληψης της διάγνωσης από συνεργάτη ιστοπαθολόγο παρατηρήθηκε σε μερικές περιπτώσεις μικρή διακύμανση στα αποτελέσματα της διάγνωσης. Αυτό πιθανόν να οφείλεται στο γεγονός ότι τα κριτήρια κατηγοριοποίησης του Παγκόσμιου Οργανισμού Υγείας εφαρμόζονται υποκειμενικά από τους ειδικούς για τον ακριβή ορισμό συνόρων μεταξύ χαμηλής και υψηλής κακοήθειας. Στην παρούσα εργασία χρησιμοποιήθηκε ως κλινικό υλικό για περαιτέρω ανάλυση οι 35 περιπτώσεις, στις οποίες υπήρξε ομοφωνία μεταξύ της επίσημης διάγνωσης του Πανεπιστημιακού Νοσοκομείου της Πάτρας και της νέας διάγνωσης του συνεργάτη ιστοπαθολόγου, σχετικά με την κατηγοριοποίηση των δεδομένων.

4.3 Επεξεργασία και τμηματοποίηση εικόνας

Είναι γνωστό ότι ο πυρήνας φέρει σημαντικές πληροφορίες, οι οποίες οδηγούν στην διάγνωση και ενδεχομένως στην πρόβλεψη της πορείας της ασθένειας. Ο ακριβής ορισμός, λοιπόν, των ορίων και η σωστή τμηματοποίηση των πυρήνων παίζουν βασικότατο ρόλο στην διάγνωση και πρόγνωση.

Η τμηματοποίηση των πυρήνων αποτελεί ένα όχι εύκολα αντιμετωπίσιμο πρόβλημα, καθώς η χρώση στις ιστοπαθολογικές εικόνες θεωρείται δύσκολη στην επεξεργασία. Αυτή η πολυπλοκότητα στην επεξεργασία τέτοιων εικόνων δικαιολογείται στην ύπαρξη υψηλής διακύμανσης μεταξύ των χρωματισμένων πυρήνων και του περιβάλλοντα ιστού.

4.3.1 Ανασκόπηση της βιβλιογραφίας

Η τμηματοποίηση του πυρήνα πρόκειται για ένα σημαντικό πεδίο έρευνας από την δεκαετία του '70. Μεταξύ κάποιων μεθόδων προσέγγισης του προβλήματος περιλαμβάνονται το ιστόγραμμα και τεχνικές εντοπισμού του περιγράμματος, που βασίζονται στην ιδέα ότι οι πυρήνες παρουσιάζουν σημαντικά υψηλότερη αντίθεση σε σύγκριση με τον περιβάλλοντα ιστό. Επίσης, έχουν εφαρμοστεί προσαρμοστικοί

αλγόριθμοι κατωφλίου με βάση τις πληροφορίες ακμών, χρησιμοποιώντας παραδοχές για την καμπυλότητα των πυρήνων. Οι τεχνικές αυτές κατωφλίου, ωστόσο, δεν είναι κατάλληλες σε περιπτώσεις όπου οι πυρήνες έχουν ακανόνιστο σχήμα και δεν φωτίζονται ομοιόμορφα οι επικαλυπτόμενες περιοχές[37,38].

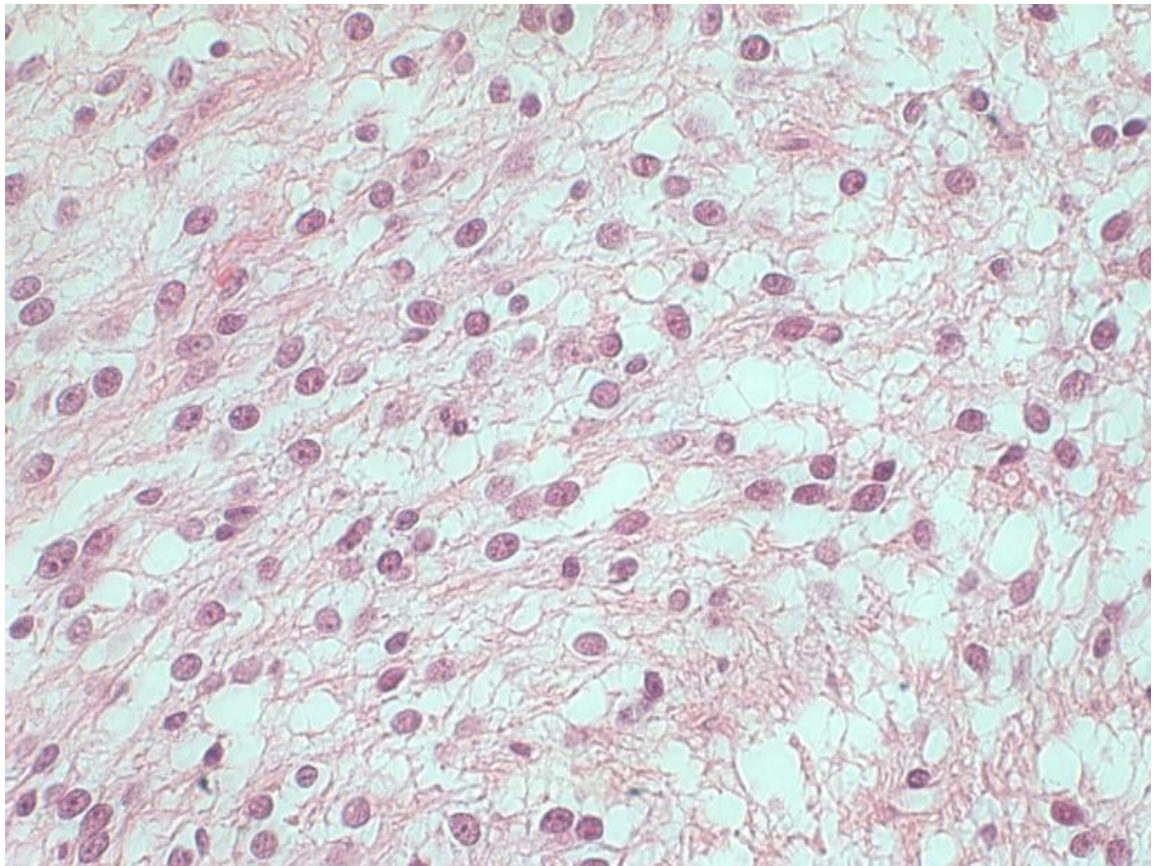
Στις σύγχρονες προσεγγίσεις μπορούν να ταξινομηθούν οι ημι-αυτόματες και οι αυτόματες κατηγοριοποιήσεις. Στις ημι-αυτόματες προσεγγίσεις απαιτείται η διαδραστική συμμετοχή του χρήστη. Τέτοιες προσεγγίσεις σχεδιάστηκαν για να λειτουργούν σε συνδυασμό με τον ανθρώπινο έλεγχο. Στις αυτόματες μεθόδους έχουν ερευνηθεί τεχνικές με βάση το ιστόγραμμα, μέθοδοι ανίχνευσης ακμών και προσεγγίσεις αναγνώρισης προτύπων[39,40]. Οι τεχνικές με βάση το ιστόγραμμα αποδείχθηκαν επιτυχημένες για εικόνες με ομοιόμορφο φωτισμό[41]. Οι μέθοδοι ανίχνευσης ακμών στηρίζονται στην ιδέα της ανίχνευσης ασυνεχειών στην ένταση των pixel στα όρια των διαφόρων αντικειμένων. Οι προσεγγίσεις αναγνώρισης προτύπων έχουν δημιουργηθεί με βάση ενός supervised ταξινομητή εκπαίδευσης, ώστε να διακριθούν οι πυρήνες από τα pixel του φόντου. Ωστόσο, οι μέθοδοι ανίχνευσης ακμών και οι τεχνικές με βάση το ιστόγραμμα είναι πιο ευαίσθητοι στην παρουσία θορύβου[42].

Οι εικόνες με H&E χρώση θεωρούνται από τις πιο πολύπλοκες και θορυβώδεις εικόνες και χαρακτηρίζονται από μεγάλη μεταβλητότητα στην ένταση του χρώματος. Έτσι, οι προαναφερόμενες τεχνικές κατάτμησης εικόνας είναι δύσκολο να εφαρμοστούν στις H&E εικόνες.

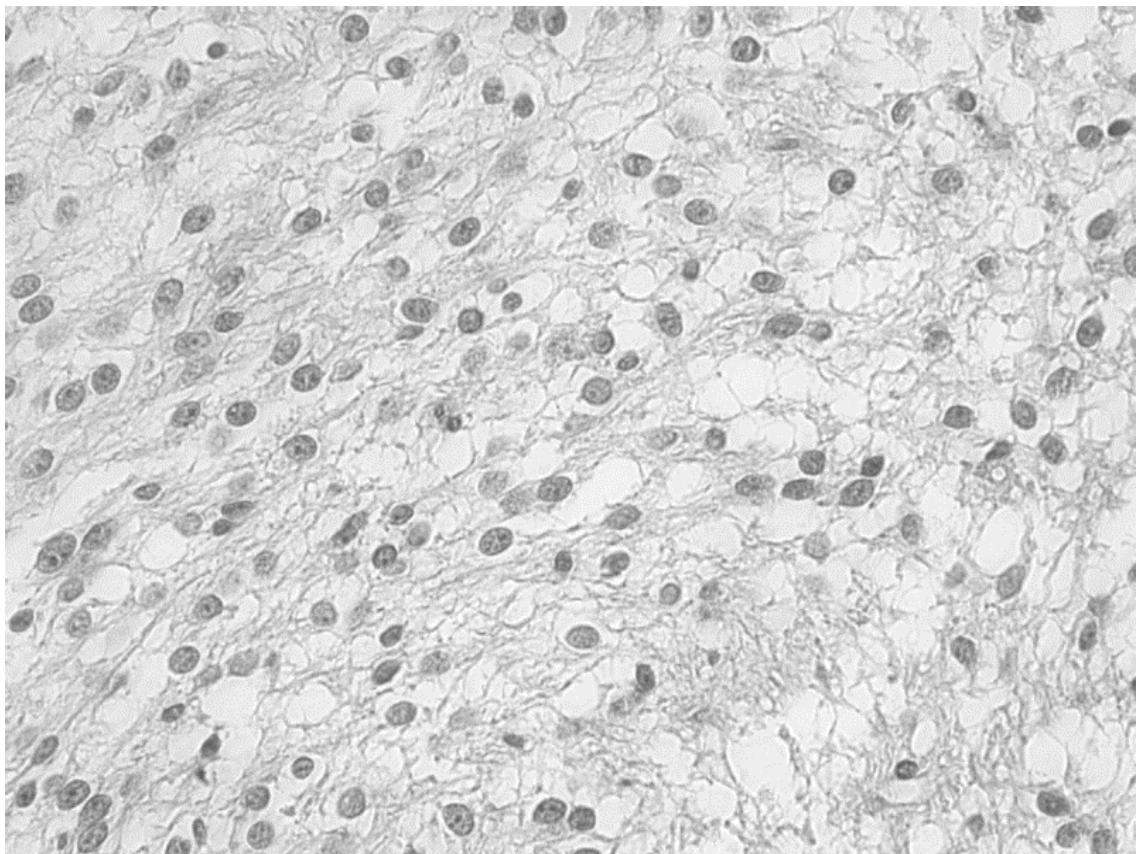
4.3.2 Μέθοδος τμηματοποίησης εικόνας

Στην παρούσα εργασία χρησιμοποιήθηκε ιστοπαθολογικό υλικό από 35 ασθενείς. Πραγματοποιήθηκε ψηφιοποίηση της πλάκας και πάρθηκαν 6 συνολικά εικόνες από τα ROI (Region Of Interest) για κάθε δείγμα-ασθενή. Από αυτές εξήχθησαν τιμές για κάθε χαρακτηριστικό που υπολογίστηκε και ως τελικό αποτέλεσμα για τον συγκεκριμένο ασθενή δόθηκε ο μέσος όρος των έξι αυτών εικόνων.

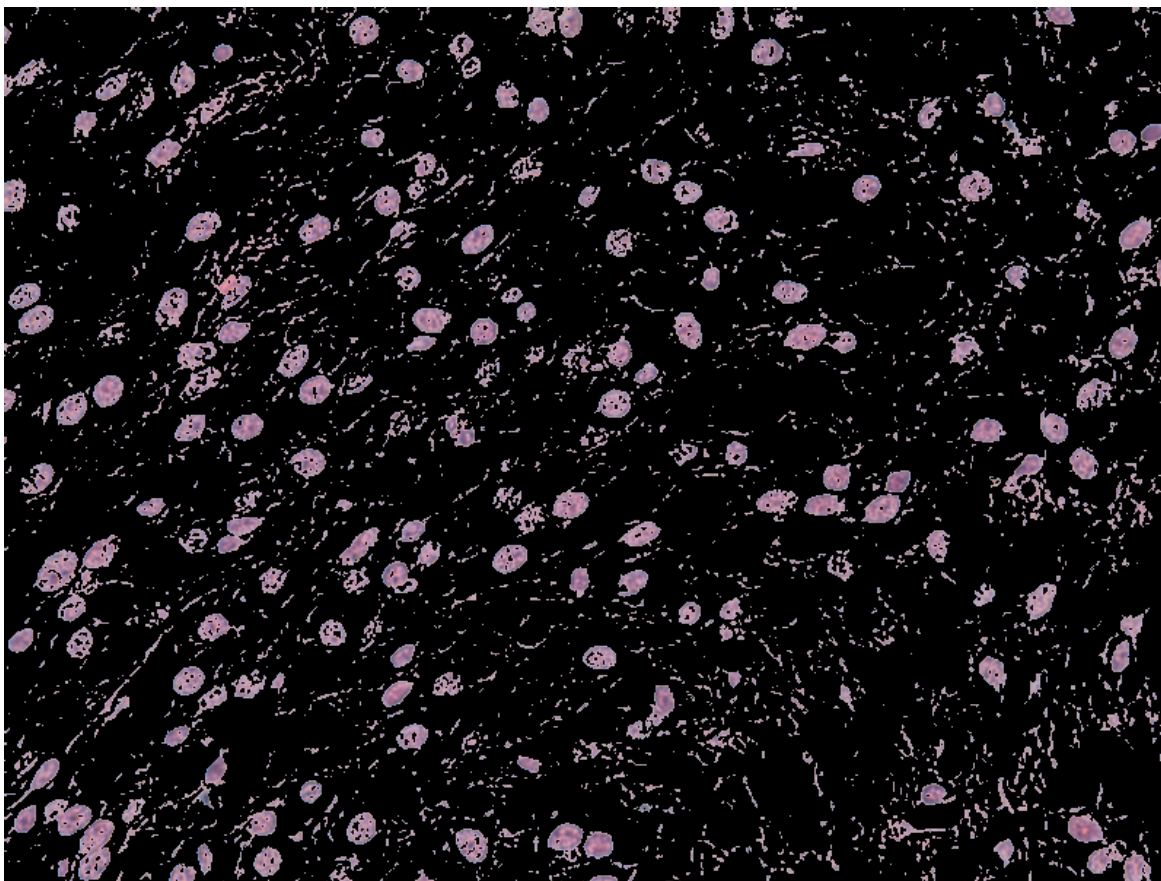
Ο αλγόριθμος αυτόματης τμηματοποίησης των εικόνων ακολούθησε τα εξής στάδια: Αρχικά, εφαρμόστηκε ομαδοποίηση των pixel της εικόνας σε δύο κατηγορίες με βάση το χρώμα αυτών με στόχο να απομονωθεί ο πυρήνας, ο οποίος μετά την χρώση παίρνει μια μπλε-μωβ απόχρωση, από τον περιβάλλοντα ιστό, ο οποίος μετά την χρώση έχει κόκκινη απόχρωση (εικόνα 13). Η ομαδοποίηση αυτή με βάση το χρώμα (clustering by color), έγινε με χρήση του k-means unsupervised ταξινομητή.



Εικόνα 13: Ψηφιοποιημένη εικόνα μετά από χρώση Η&Ε

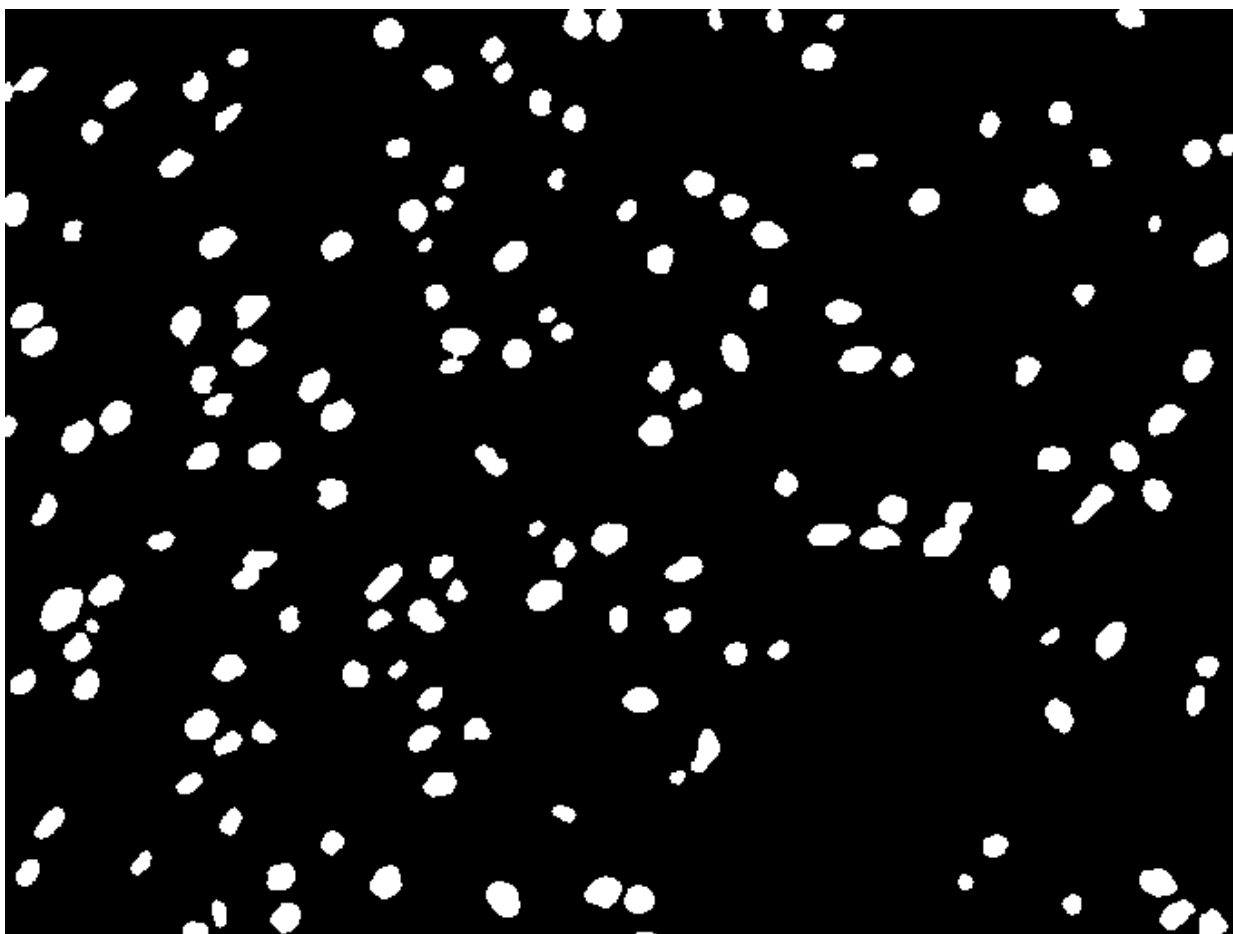


Εικόνα 14: Ασπρόμαυρη ψηφιοποιημένη εικόνα μετά από χρώση Η&Ε



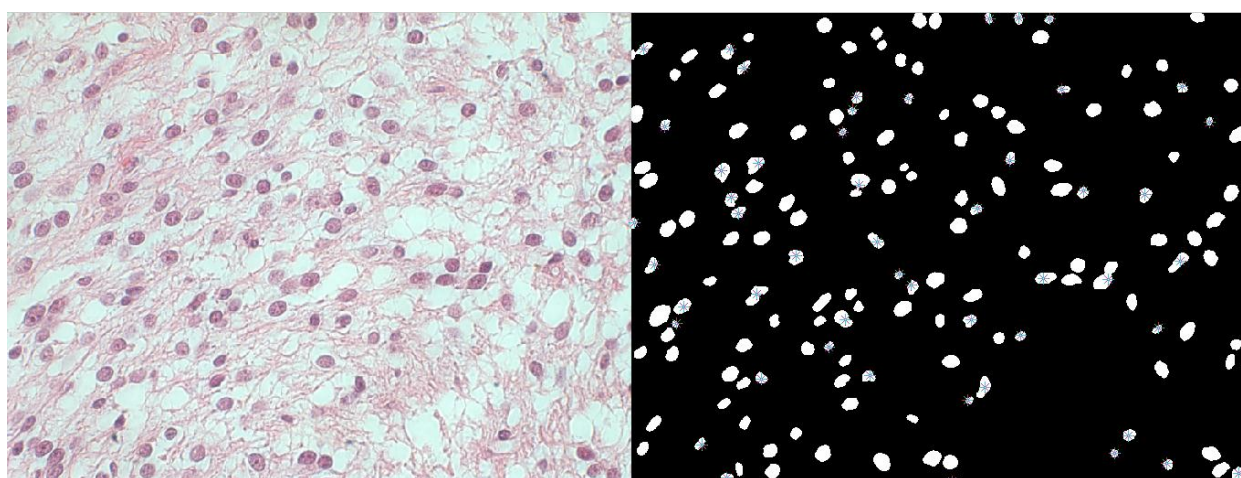
Εικόνα 15: Τμηματοποιημένη εικόνα μετά από εφαρμογή του k-means color based clustering, που απομονώνει τους πυρήνες

Έπειτα, η εικόνα που περιλαμβάνει μόνο τους πυρήνες (εικόνα 15) υπέστη περαιτέρω επεξεργασία. Σε πρώτη φάση, έγινε μετατροπή της εικόνας από έγχρωμη σε ασπρόμαυρη, δημιουργήθηκε η μάσκα αυτής (binary image) και έπειτα εφαρμόστηκαν τεχνικές επεξεργασίας εικόνας, όπως filling holes, closing και opening, για να επιτυγχανθεί η καλύτερη απομόνωση των πυρήνων. Στην συνέχεια του αλγορίθμου εφαρμόστηκε αυτόματη αφαίρεση των πολύ μικρών και των πολύ μεγάλων πυρήνων, που τυχαίνει να δημιουργηθούν μετά από την επεξεργασία της εικόνας (εικόνα 16). Αυτό γίνεται, καθώς πολλές φορές συμβαίνει κάποιοι πυρήνες να επικαλύπτονται και ο αλγόριθμος να αναγνωρίσει αυτούς ως έναν εξωπραγματικά μεγάλο πυρήνα.



Εικόνα 16: Μάσκα πυρήνων μετά από αυτόματη επεξεργασία

Για να εξασφαλισθεί η ακριβής και σωστή απομόνωση των πυρήνων ζητείται, στο τέλος, από τον χρήστη να αφαιρέσει χειροκίνητα τους πυρήνες που θεωρεί ότι έχουν τμηματοποιηθεί λανθασμένα, έχοντας στην οθόνη του την αρχική εικόνα και την μάσκα με τους τμηματοποιημένους πυρήνες (εικόνα 17).

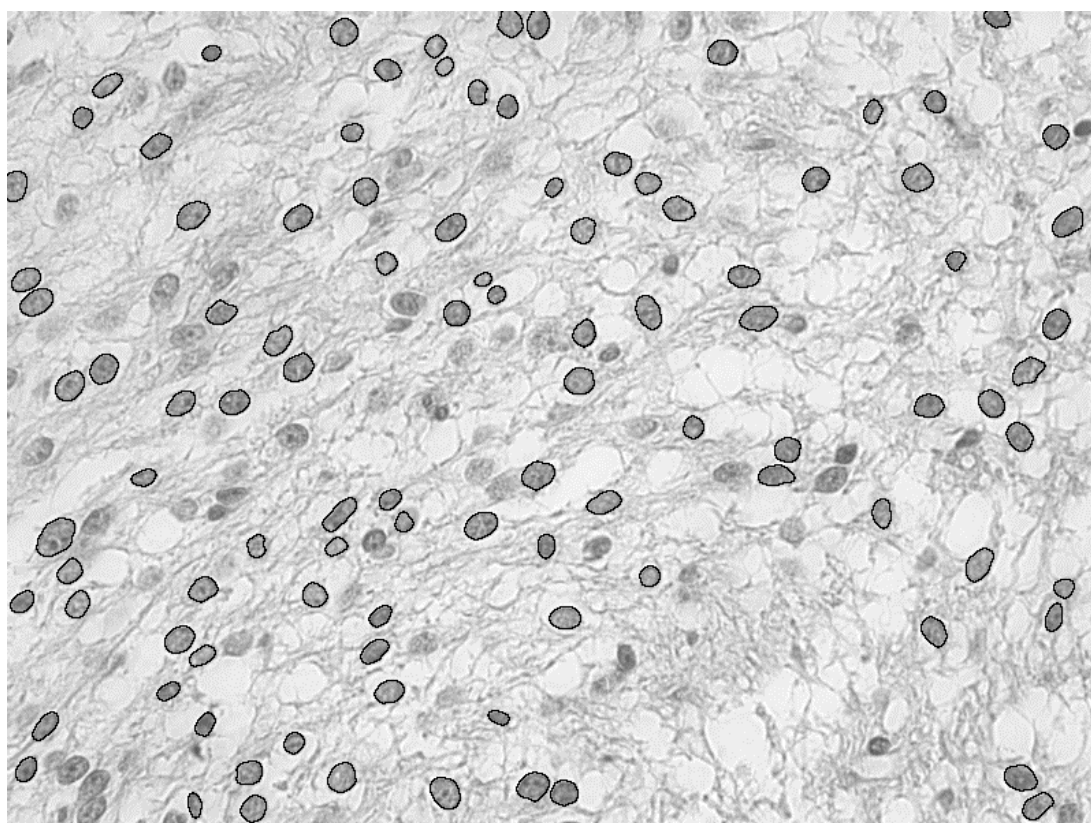


Εικόνα 17: Χειροκίνητη επιλογή πυρήνων για αφαίρεση από την μάσκα, λόγω λανθασμένης τμηματοποίησης αυτών

Τέλος, είναι βασικό να αναφερθεί ότι πραγματοποιήθηκε αφαίρεση του φόντου, που προέρχεται από την εικόνα που έχει ο φακός του μικροσκοπίου, χωρίς να υπάρχει κάποια πλάκα-δείγμα σε αυτόν.



Εικόνα 18: Τελική τμηματοποιημένη εικόνα πυρήνων



Εικόνα 19: Η τελική μάσκα πυρήνων εφαρμοσμένη στην αρχική εικόνα

4.3.3 Σύνοψη και συμπεράσματα

Τα αποτελέσματα της τμηματοποίησης των εικόνων ήταν ενθαρρυντικά δεδομένου ότι η χρώση H&E δεν είναι ένα αρκετά αυστηρό πρωτόκολλο χρώσης, ώστε να ξεχωρίζουν έντονα οι πυρήνες από τον περιβάλλοντα ιστό. Θα πρέπει να τονιστεί ότι σκοπός της διαδικασίας της τμηματοποίησης δεν ήταν να απομονωθούν όλοι οι πυρήνες από κάθε εικόνα του καθενός δείγματος-ασθενή. Αντίθετα, σκοπός της διαδικασίας ήταν να εξάγει ένα αντιπροσωπευτικό αριθμό ορθά τμηματοποιημένων πυρήνων από κάθε σύνολο εικόνων που περιγράφουν κάθε περίπτωση (ασθενή), προκειμένου να υπολογισθούν χαρακτηριστικά. Στην παρούσα εργασία απομονώθηκαν κατά μέσο όρο περίπου 98 πυρήνες από κάθε εικόνα και για κάθε δείγμα-ασθενή υπήρχαν 6 εικόνες ψηφιοποιημένες από τα δεδομένα της βιοψίας. Επομένως, για τον υπολογισμό των χαρακτηριστικών μελετήθηκαν συνολικά κατά μέσο όρο 588 τμηματοποιημένοι πυρήνες για κάθε δείγμα-ασθενή. Επομένως, αν ληφθεί υπόψιν το παραπάνω και δεδομένου ότι έχει αποδειχθεί ότι ακόμη και 200 ορθά τμηματοποιημένοι πυρήνες είναι επαρκείς για την εξαγωγή χαρακτηριστικών, ο μικρός αριθμός πυρήνων που δεν τμηματοποιήθηκαν σωστά και δεν ληφθήκαν υπόψιν στον υπολογισμό χαρακτηριστικών είναι μικρής σημασίας. Σημαντικό αποτελεί το γεγονός, ότι με αυτήν την προσέγγιση τμηματοποίησης ανιχνεύονται τα όρια των πυρήνων διατηρώντας σημαντικές πληροφορίες του μεγέθους και του σχήματος αυτών.

4.4 Αναγνώριση προτύπων

Ο προσδιορισμός του βαθμού του καρκινικού όγκου είναι το πιο κρίσιμο στάδιο κατά τη διάγνωση, επειδή καθορίζει την διαχείριση του ασθενή και την θεραπεία του. Οι ιστοπαθολόγοι αξιολογούν την επιθετικότητα των όγκων με εξέταση ενός τμήματος ιστού από την βιοψία με το μικροσκόπιο. Με βάση τις κατευθυντήριες γραμμές που δημοσιεύθηκαν από τον Παγκόσμιο Οργανισμό Υγείας (WHO), οι βαθμοί καθορίζονται με βάση ιστολογικών κριτηρίων σε χαμηλού βαθμού (I, II) και υψηλού βαθμού (III, IV). Ο WHO αποτελεί το πιο δημοφιλές σύστημα βαθμολόγησης για την ταξινόμηση των αστροκυτωμάτων, ωστόσο, περιορίζεται από την ασαφή περιγραφή για τον ακριβή καθορισμό του κάθε βαθμού, ενώ σε μερικές περιπτώσεις αυτό επιτυγχάνεται μόνο από την εμπειρία του ιστοπαθολόγου. Για τον παραπάνω λόγο, πολλές φορές προτείνεται επαναληψιμότητα στην διάγνωση από δεύτερο εξεταστή, η οποία μπορεί να οδηγήσει σε πολύ χαμηλά ποσοστά της τάξης του 36-51%.

Το κλινικό υλικό που χρησιμοποιήθηκε προήλθε από 35 ασθενείς που είχαν υποβληθεί σε χειρουργική επέμβαση στο Πανεπιστημιακό Νοσοκομείο της Πάτρας. Οι βιοψίες κατηγοριοποιήθηκαν σε τρεις βαθμούς κακοήθειας (I, II, III ή IV) σύμφωνα με τον Παγκόσμιο Οργανισμό Υγείας (World Health Organization - WHO), από ιστοπαθολόγο ιατρό (Ιατρός Παναγιώτα Ραβαζούλα). Από τις 35 βιοψίες, 19 χαρακτηρίστηκαν από ιστοπαθολόγο ιατρό ως χαμηλής κακοήθειας (βαθμός I ή II) και 16 ως υψηλής κακοήθειας (βαθμός III ή IV). Μετά την τμηματοποίηση των έξι (6) εικόνων για κάθε ασθενή, πραγματοποιήθηκε εξαγωγή τριάντα ένα (31) μορφολογικών χαρακτηριστικών και χαρακτηριστικών υφής.

Πίνακας 2: Μορφολογικά χαρακτηριστικά και χαρακτηριστικά υψής

Χαρακτηριστικά υψής		Μορφολογικά χαρακτηριστικά	
Χαρακτηριστικά 1 ^{ης} τάξης	Χαρακτηριστικά 2 ^{ης} τάξης		
	<i>Co-occurrence matrix</i>	<i>Run Length matrix</i>	
Mean	<i>ASM**</i>	<i>SRE**</i>	Area
Variance	<i>CON**</i>	<i>LRE**</i>	Perimeter
Skewness	<i>COR**</i>	<i>GLNU**</i>	Minor Axis Length
Kurtosis	<i>Variance**</i>	<i>RLNU**</i>	Major Axis Length
	<i>IDM**</i>	<i>RP**</i>	Eccentricity
	<i>Entropy**</i>		

**Για κάθε χαρακτηριστικό υπολογίστηκε η μέση τιμή και το εύρος του

4.4.1 Ανασκόπηση της βιβλιογραφίας

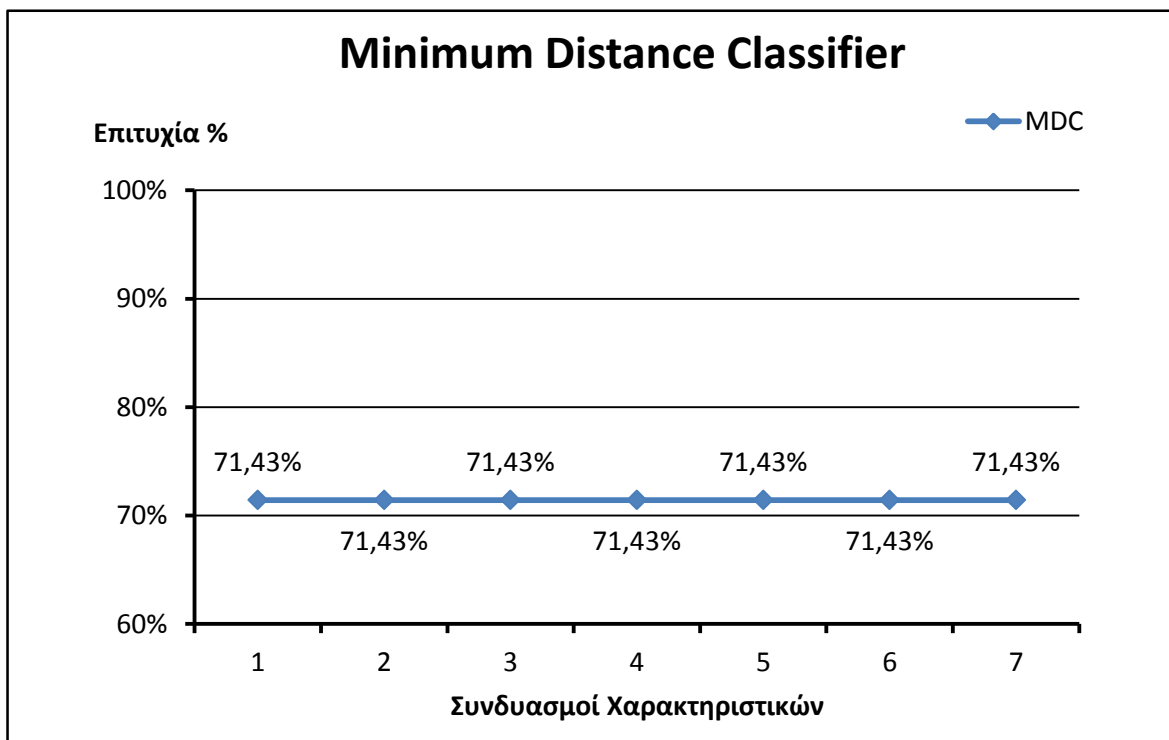
Στον τομέα της ανάπτυξης αυτόματων τεχνικών για την βελτίωση της ακρίβειας στην διάγνωση των αστροκυτταρικών όγκων έχουν εξαχθεί αρκετές μελέτες. Στην έρευνα του Schad[18] ερευνήθηκαν ημιποσοτικά χαρακτηριστικά πυρήνων, τα οποία εξήχθησαν από χρώση Feulgen. Τα αποτελέσματα έδειξαν 94% συμφωνία με τα αποτελέσματα του Kernohan[43]. Ο Decaestecker[44], εξετάζοντας ποσοτικά χαρακτηριστικά πυρήνων βαμμένα με χρώση Feulgen, διέκρινε τα αστροκυττώματα σε χαμηλού και υψηλού βαθμού με ακρίβεια 55%. Ο Kolles[45] ανέπτυξε ένα νέο σύστημα βαθμολόγησης (HOM), το οποίο φάνηκε να είναι πιο αποτελεσματικό στην αυτόματη ταξινόμηση γλοιωμάτων εν συγκρίσει με το WHO σύστημα βαθμολόγησης. Στην έρευνα του Kolles διακρίθηκαν διαφορετικού βαθμού γλοιώματα με ακρίβεια 97% χρησιμοποιώντας το HOM σύστημα, 49% χρησιμοποιώντας το Daumas-Duport σύστημα και 61% ακρίβεια χρησιμοποιώντας το WHO σύστημα βαθμολόγησης. Ο Belacel[32] ανέπτυξε μια μέθοδο fuzzy για την ταξινόμηση των αστροκυττωμάτων, η οποία ταξινομούσε αυτόματα διαφορετικού βαθμού δείγματα με ακρίβεια 66%.

Στην εργασία αυτή υλοποιήθηκε και αναπτύχθηκε αλγόριθμος, ο οποίος διαχωρίζει τους όγκους εγκεφάλου σε χαμηλού βαθμού και σε υψηλού βαθμού κακοήθειας. Αρχικά, πραγματοποιήθηκε η τμηματοποίηση εικόνων μικροσκοπίου, όπως περιγράφηκε στο υποκεφάλαιο §4.3. Στη συνέχεια, πραγματοποιήθηκε εξαγωγή μορφολογικών χαρακτηριστικών και χαρακτηριστικών υψής από τους τμηματοποιημένους πυρήνες των εικόνων. Τέλος, εφαρμόστηκαν διαδοχικά supervised, unsupervised και semi-supervised ταξινομητές χρησιμοποιώντας συνδυασμούς χαρακτηριστικών ανά ένα, ανά δύο έως και ανά επτά, ώστε να υπολογιστεί το ποσοστό ακρίβειας αυτών. Στο σύστημα που αναπτύχθηκε δεν εφαρμόστηκαν όλοι οι συνδυασμοί των χαρακτηριστικών (δηλαδή ανά οκτώ έως και ανά τριάντα ένα – που ήταν και το συνολικό πλήθος των χαρακτηριστικών), διότι εκτός από την αύξηση του υπολογιστικού κόστους και της πολυπλοκότητας έχει αποδειχθεί ότι μετά από ένα αριθμό συνδυασμών των χαρακτηριστικών, η ακρίβεια του ταξινομητή μειώνεται.

4.4.2 Αποτελέσματα supervised ταξινομητών

4.4.2.1 Minimum Distance (MDC)

Ο Minimum Distance ταξινομητής είναι ένας από τους πιο απλούς ταξινομητές. Με βάση την ελάχιστη απόσταση από τα κέντρα των ήδη υπαρχόντων κλάσεων αποφασίζεται το πού θα ταξινομηθεί το νέο χαρακτηριστικό. Ως κέντρο μιας κλάσης χρησιμοποιήθηκε το διάνυσμα της μέσης τιμής της κλάσης αυτής και ως τρόπος μέτρησης της απόστασης ενός διανύσματος από μια κλάση ορίστηκε η Ευκλείδεια απόσταση.

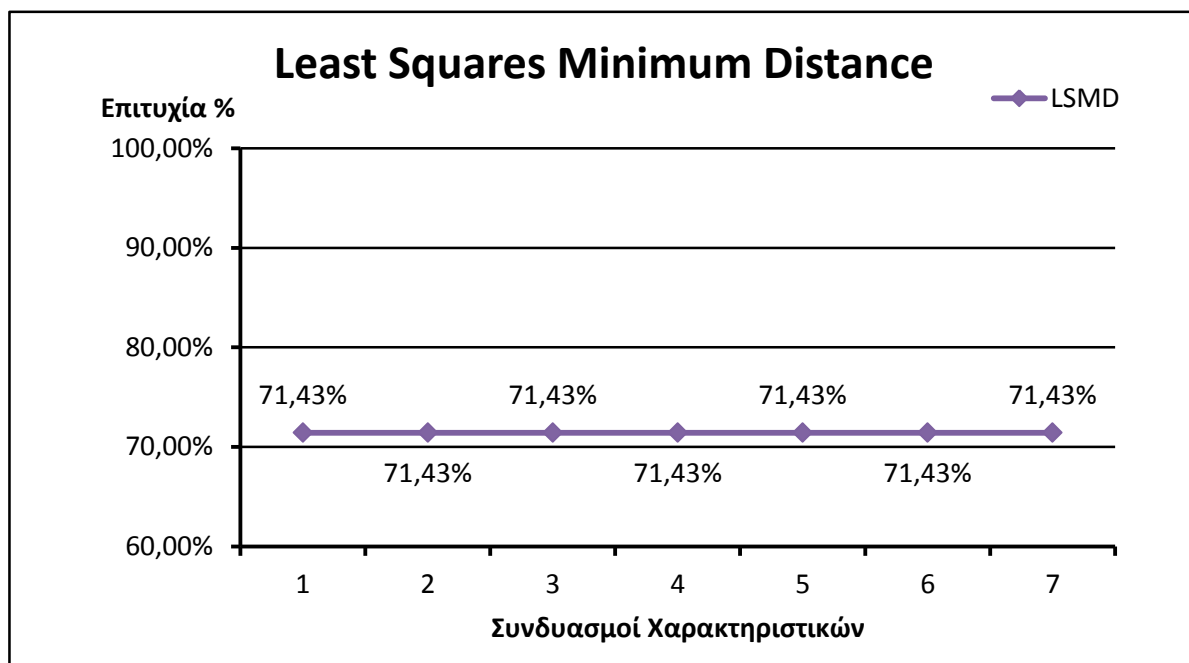


Σχήμα 1: Διάγραμμα απεικόνισης του καλύτερου ποσοστού επιτυχίας του MDC για κάθε πλήθος χαρακτηριστικών

Το παραπάνω διάγραμμα απεικονίζει τα καλύτερα ποσοστά επιτυχίας ανά πλήθος χαρακτηριστικών, όπως αυτά παρουσιάστηκαν μετά την εκτέλεση του MDC classifier αλγορίθμου. Σε αυτόν τον ταξινομητή καλύτερο χαρακτηριστικό διαπιστώθηκε να είναι το Mean (χαρακτηριστικό υψής 1^{ης} τάξης), καθώς έδωσε 71,43% ποσοστό επιτυχίας, τόσο όταν λήφθηκε υπόψιν ως μοναδικό χαρακτηριστικό, όσο και όταν συνδυάστηκε με πληθώρα άλλων χαρακτηριστικών.

4.4.2.2 Least Squares Minimum Distance (LSMD)

Με τον Least Squares Minimum Distance ταξινομητή όλα τα πρότυπα μιας κλάσης ωθούνται να ομαδοποιηθούν γύρω από ένα προκαθορισμένο σημείο σε έναν χώρο «απόφασης». Η ταξινόμηση του άγνωστου προτύπου βασίζεται στην ελάχιστη του απόσταση από όλα τα προκαθορισμένα σημεία.

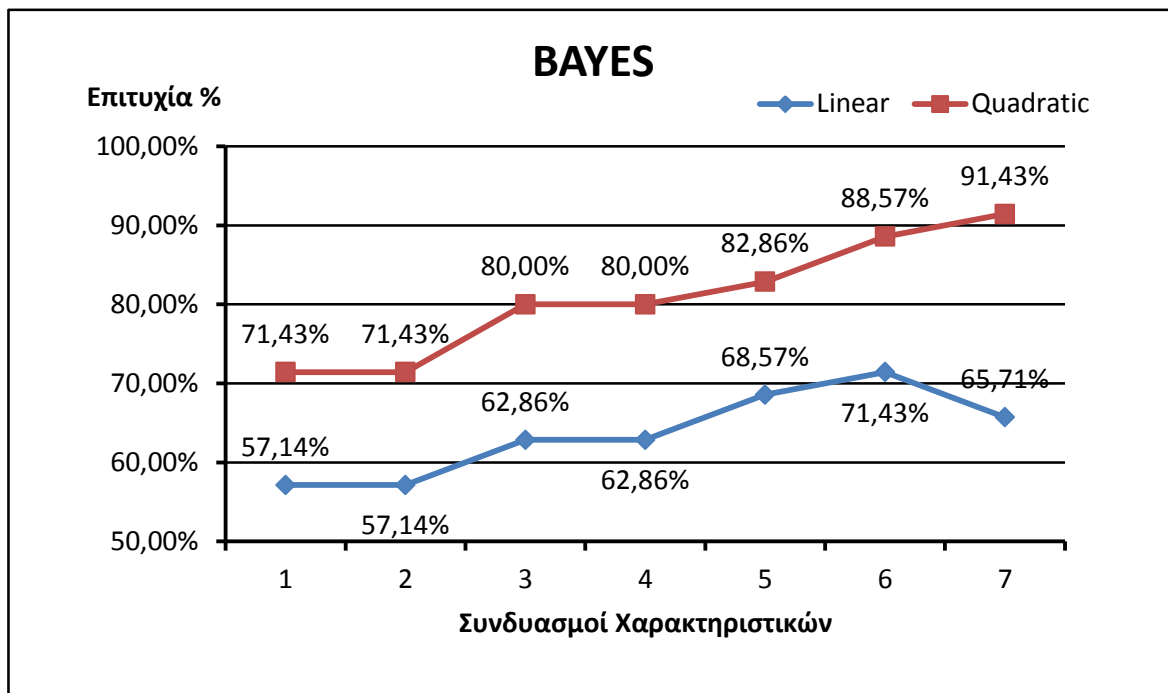


Σχήμα 2: Διάγραμμα απεικόνισης του καλύτερου ποσοστού επιτυχίας του LSMD για κάθε πλήθος χαρακτηριστικών

Το παραπάνω διάγραμμα απεικονίζει τα καλύτερα ποσοστά επιτυχίας ανά πλήθος χαρακτηριστικών, όπως αυτά παρουσιάστηκαν μετά την εκτέλεση του LSMD classifier αλγορίθμου. Σε αυτόν τον ταξινομητή, όπως και στον MDC, καλύτερο χαρακτηριστικό διαπιστώθηκε να είναι το Mean (χαρακτηριστικό υψής 1^{ης} τάξης), καθώς έδωσε 71,43% ποσοστό επιτυχίας, τόσο όταν λήφθηκε υπόψιν ως μοναδικό χαρακτηριστικό, όσο και όταν συνδυάστηκε με πληθώρα άλλων χαρακτηριστικών.

4.4.2.3 Bayesian

Ο Bayesian ταξινομητής δίνει την ελάχιστη πιθανότητα σφάλματος για τα δεδομένα, ακολουθώντας μια συγκεκριμένη κατανομή, η οποία συνήθως επιλέγεται να είναι η Gaussian κατανομή. Στην παρούσα εργασία εκτελέστηκαν τόσο ο Bayes Linear (Gaussian κατανομή) όσο και ο Bayes Quadratic.

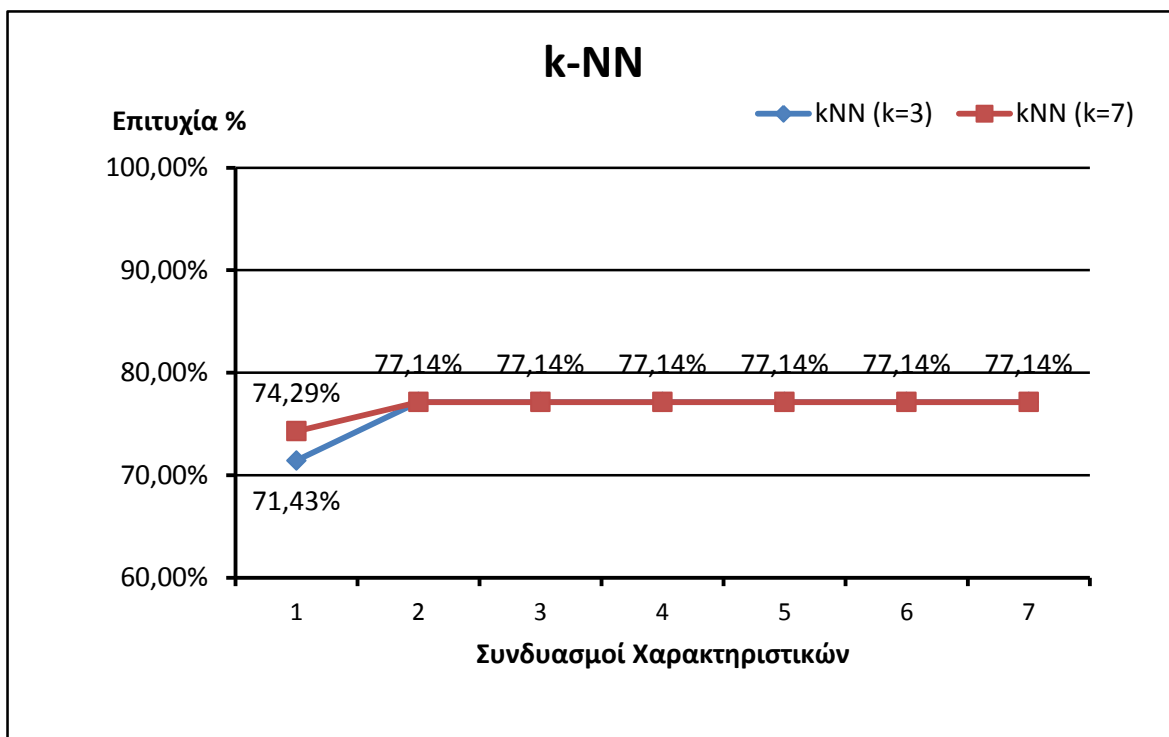


Σχήμα 3: Διάγραμμα απεικόνισης του καλύτερου ποσοστού επιτυχίας του Bayes Linear και Bayes Quadratic για κάθε πλήθος χαρακτηριστικών

Το παραπάνω διάγραμμα απεικονίζει τα καλύτερα ποσοστά επιτυχίας ανά πλήθος χαρακτηριστικών, όπως αυτά παρουσιάστηκαν μετά την εκτέλεση του Bayes Linear classifier και του Bayes Quadratic classifier αλγορίθμου. Σε αυτό το σημείο αξίζει να σημειωθεί ότι το υψηλό ποσοστό επιτυχίας της τάξης του 91.43% επετεύχθη με τον Bayes Quadratic classifier αλγόριθμο με τον συνδυασμό των επτά χαρακτηριστικών Mean και Variance (από τα χαρακτηριστικά υψής 1^{ης} τάξης), mean Contrast, mean Correlation, mean Entropy και mean Gray Level Non-Uniformity (από τα χαρακτηριστικά υψής 2^{ης} τάξης) και Perimeter (από τα μορφολογικά χαρακτηριστικά). Επίσης, το υψηλότερο ποσοστό επιτυχίας (71.43%) του Bayes Linear classifier αλγορίθμου επετεύχθη με τον συνδυασμό των έξι χαρακτηριστικών Skewness (από τα χαρακτηριστικά υψής 1^{ης} τάξης), mean Short Run Emphasis, range Long Run Emphasis, mean Gray Level Non-Uniformity και range Gray Level Non-Uniformity (από τα χαρακτηριστικά υψής 2^{ης} τάξης) και Minor Axis Length (από τα μορφολογικά χαρακτηριστικά).

4.4.2.4 K-Nearest Neighbors (k-NN)

Στον K-Nearest Neighbors (k-NN) ταξινομητή, αρχικά εντοπίζονται οι k πλησιέστεροι γείτονες του άγνωστου προτύπου και στην συνέχεια το πρότυπο ταξινομείται στην κλάση j, αν η πλειοψηφία των k γειτόνων ανήκουν στην κλάση j.

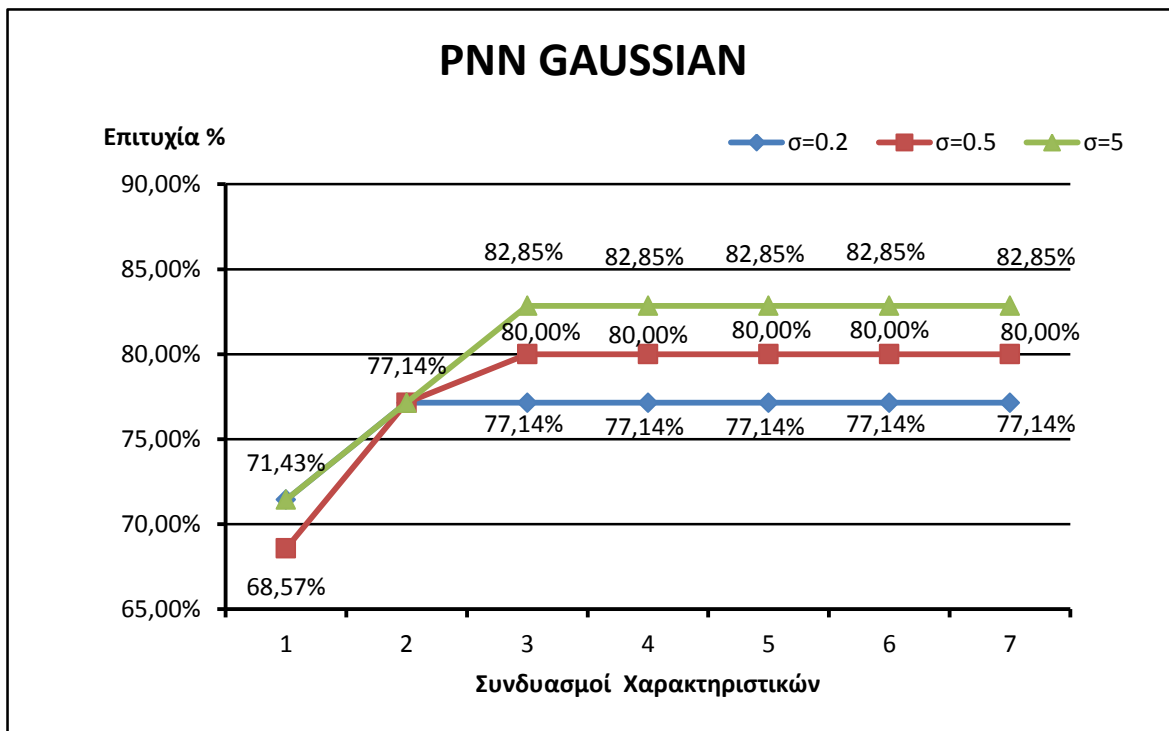


Σχήμα 4: Διάγραμμα απεικόνισης του καλύτερου ποσοστού επιτυχίας του k-NN για κάθε πλήθος χαρακτηριστικών για k=3 και k=7

Το παραπάνω διάγραμμα απεικονίζει τα καλύτερα ποσοστά επιτυχίας ανά πλήθος χαρακτηριστικών, όπως αυτά παρουσιάστηκαν μετά την εκτέλεση του k-NN classifier αλγορίθμου. Για k=3 παρουσιάστηκε μέγιστο ποσοστό επιτυχίας το 77.14%, το οποίο σημειώθηκε με τον συνδυασμό των δύο χαρακτηριστικών mean Correlation και range Entropy, ενώ για k=7 το μέγιστο ποσοστό επιτυχίας είναι πάλι το 77.14%, το οποίο σημειώθηκε από τα χαρακτηριστικά Mean (χαρακτηριστικό υψής 1^{ης} τάξης) και mean Correlation ή mean Correlation και mean Inverse Difference Moment (από τα χαρακτηριστικά υψής 2^{ης} τάξης).

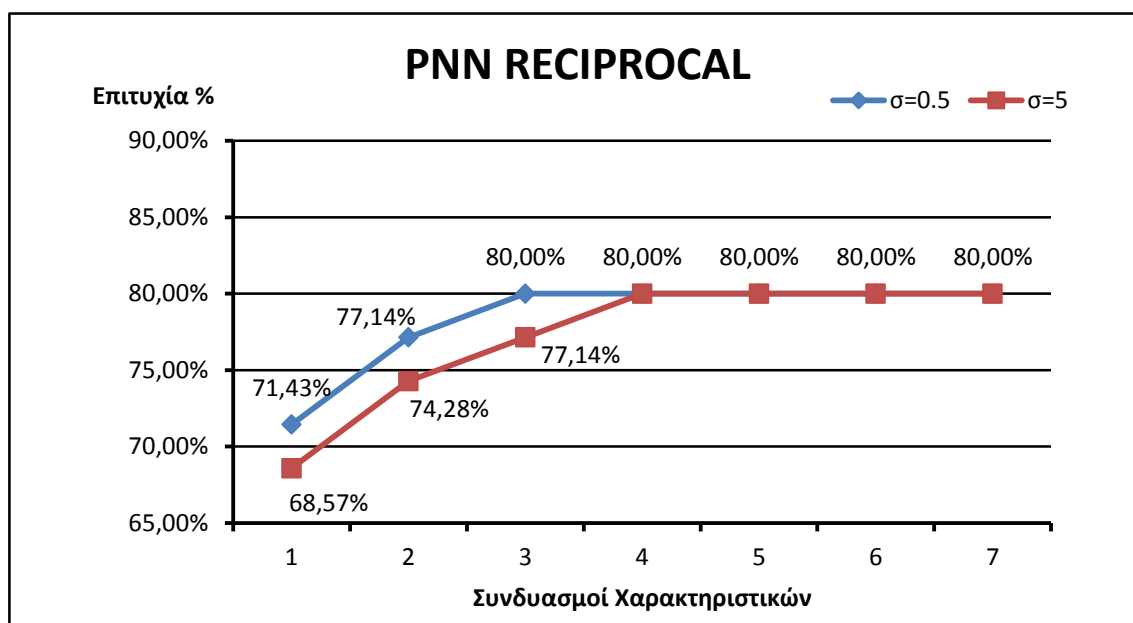
4.4.2.5 Probabilistic Neural Network (PNN)

Μια από τις πιο δημοφιλείς προσεγγίσεις για την εκτίμηση της συνάρτησης πυκνότητας πιθανότητας (PDF) έχει προταθεί από τον Parzen, όπως αναφέρθηκε στην υποενότητα §3.5.1.5. Στην παρούσα εργασία οι επιλογές για την συνάρτηση kernel ήταν η Gaussian, η Exponential και η Reciprocal για διάφορες τιμές της σ παραμέτρου εξομάλυνσης.



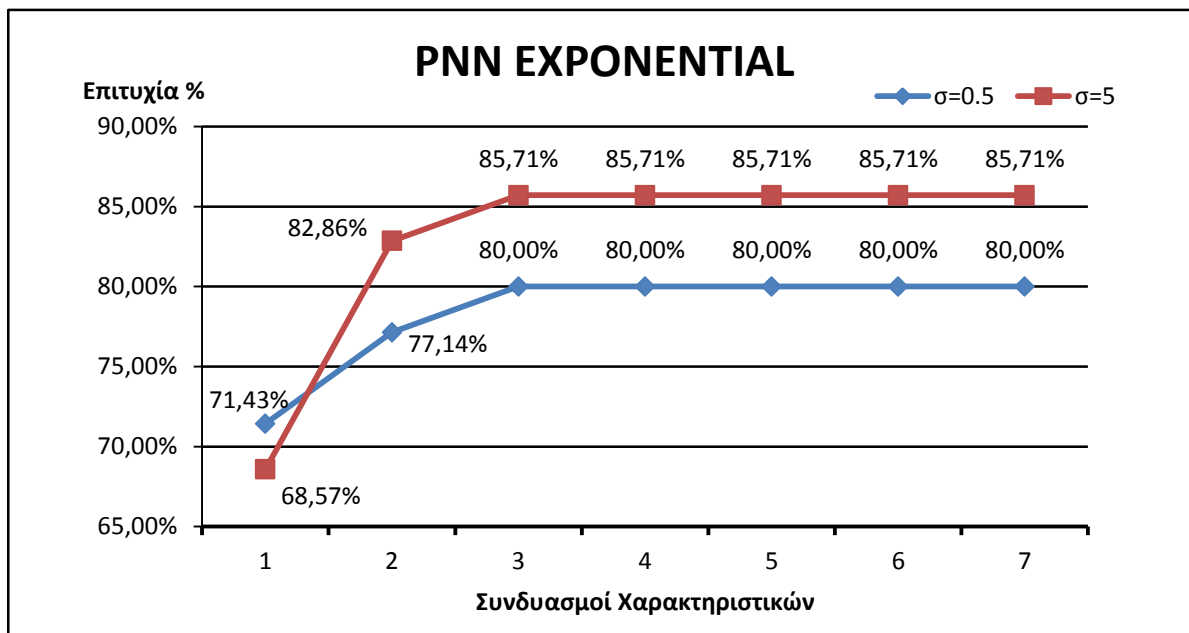
Σχήμα 5: Διάγραμμα απεικόνισης του καλύτερου ποσοστού επιτυχίας του PNN Gaussian για κάθε πλήθος χαρακτηριστικών για $\sigma=0.5$, $\sigma=0.2$ και $\sigma=5$

Το παραπάνω διάγραμμα απεικονίζει τα καλύτερα ποσοστά επιτυχίας ανά πλήθος χαρακτηριστικών, όπως αυτά παρουσιάστηκαν μετά την εκτέλεση του PNN classifier αλγορίθμου με συνάρτηση kernel την Gaussian και για τιμές $\sigma=0.2$, $\sigma=0.5$ και $\sigma=5$. Όπως παρατηρείται το μέγιστο ποσοστό ακρίβειας ήταν 82.85%, το οποίο παρατηρήθηκε για $\sigma=5$ με τον συνδυασμό των τριών χαρακτηριστικών Mean (χαρακτηριστικό υψής 1^{ης} τάξης), range Inverse Difference Moment (από τα χαρακτηριστικά υψής 2^{ης} τάξης) και Area (από τα μορφολογικά χαρακτηριστικά).



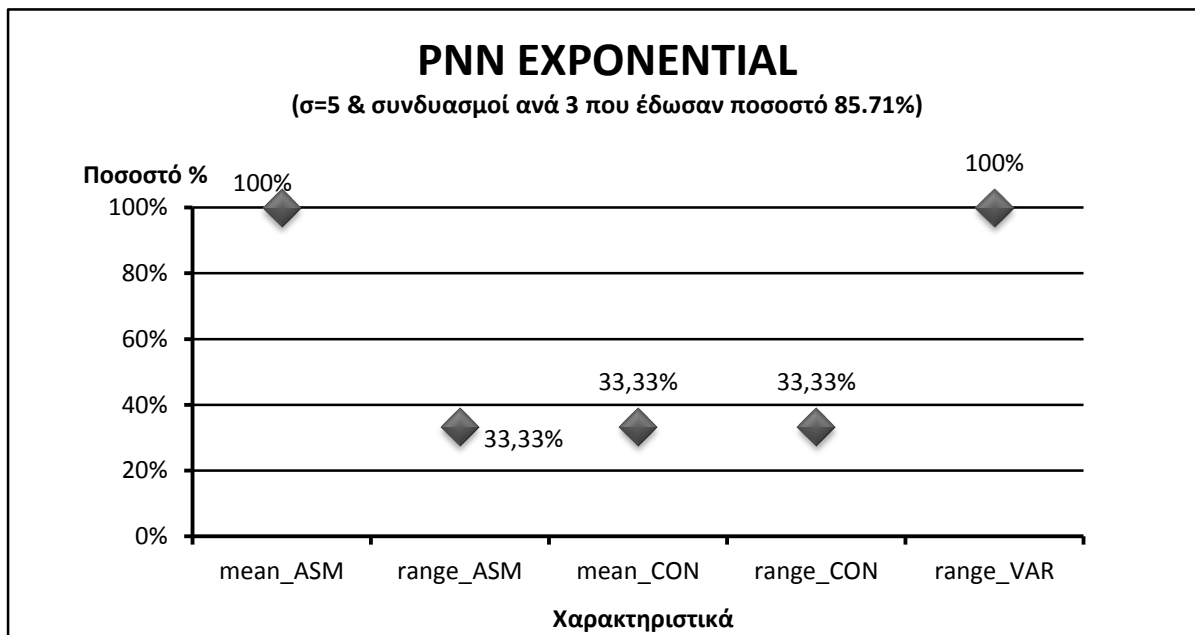
Σχήμα 6: Διάγραμμα απεικόνισης του καλύτερου ποσοστού επιτυχίας του PNN Reciprocal για κάθε πλήθος χαρακτηριστικών για $\sigma=0.5$ και $\sigma=5$

Το διάγραμμα του σχήματος 6 απεικονίζει τα καλύτερα ποσοστά επιτυχίας ανά πλήθος χαρακτηριστικών, όπως αυτά παρουσιάστηκαν μετά την εκτέλεση του PNN classifier αλγορίθμου με συνάρτηση kernel την Reciprocal και για τιμές $\sigma=0.5$ και $\sigma=5$. Το μέγιστο ποσοστό ακρίβειας που παρατηρήθηκε ήταν 80%, το οποίο παρατηρήθηκε για $\sigma=0.5$ με τον συνδυασμό των τριών χαρακτηριστικών Mean (χαρακτηριστικό υψής 1^{ης} τάξης), range Inverse Difference Moment και range Entropy (από τα χαρακτηριστικά υψής 2^{ης} τάξης).



Σχήμα 7: Διάγραμμα απεικόνισης του καλύτερου ποσοστού επιτυχίας του PNN Exponential για κάθε πλήθος χαρακτηριστικών για $\sigma=0.5$ και $\sigma=5$

Το διάγραμμα του σχήματος 7 απεικονίζει τα καλύτερα ποσοστά επιτυχίας ανά πλήθος χαρακτηριστικών, όπως αυτά παρουσιάστηκαν μετά την εκτέλεση του PNN classifier αλγορίθμου με συνάρτηση kernel την Exponential και για τιμές $\sigma=0.5$ και $\sigma=5$. Το μέγιστο ποσοστό ακρίβειας που παρατηρήθηκε ήταν 85.71%, το οποίο παρατηρήθηκε για $\sigma=5$ με τον συνδυασμό των τριών χαρακτηριστικών mean Angular Second Moment, range Angular Second Moment και range Variance (από τα χαρακτηριστικά υψής 2^{ης} τάξης) ή των τριών χαρακτηριστικών mean Angular Second Moment, mean Contrast και range Variance (από τα χαρακτηριστικά υψής 2^{ης} τάξης) ή των τριών χαρακτηριστικών mean Angular Second Moment, range Contrast και range Variance (από τα χαρακτηριστικά υψής 2^{ης} τάξης).

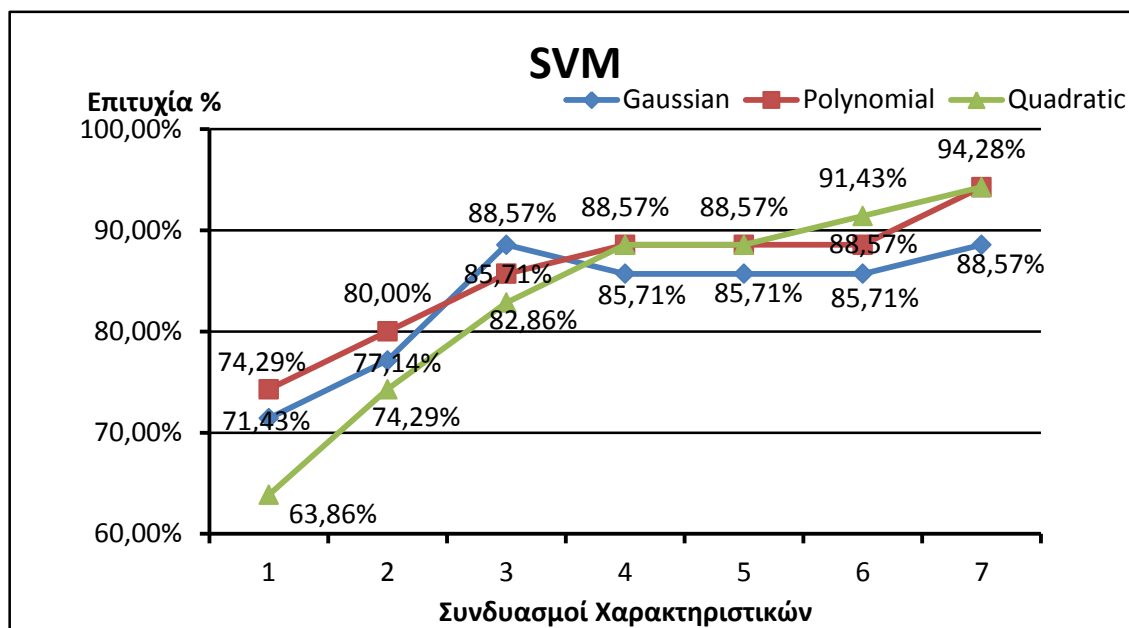


Σχήμα 8: Διάγραμμα απεικόνισης των καλύτερων χαρακτηριστικών που έδωσαν το μέγιστο ποσοστό επιτυχίας για τον PNN Exponential Classifier για $\sigma=5$ (85.71% - συνδυασμός χαρακτηριστικών ανά 3)

Στο ανωτέρω διάγραμμα (σχήμα 8) απεικονίζονται τα χαρακτηριστικά εκείνα, των οποίων ο συνδυασμός ανά τρία, έδωσαν το καλύτερο ποσοστό ακρίβειας για τον PNN Exponential με $\sigma=5$, όπως προαναφέρθηκε. Όπως παρατηρείται, το χαρακτηριστικό mean Angular Second Moment και το range Variance (χαρακτηριστικά υψής 2^{ης} τάξης), υπήρχαν σε όλους τους συνδυασμούς που έδωσαν ποσοστό ακρίβειας 85.71%, ενώ τα χαρακτηριστικά range Angular Second Moment, mean Contrast και range Contrast (χαρακτηριστικά υψής 2^{ης} τάξης) υπήρχαν στο ένα τρίτο των συνδυασμών που έδωσαν το μέγιστο ποσοστό ακρίβειας που σημειώθηκε (85.71%).

4.4.2.6 Support Vector Machine (SVM)

Στην περίπτωση των γραμμικά διαχωρίσιμων προτύπων, ο γραμμικός SVM βρίσκει, ανάμεσα σε όλα τα υπερεπίπεδα που ελαχιστοποιούν το λάθος της εκπαίδευσης, το υπερεπίπεδο εκείνο που διαχωρίζει πλήρως τα πρότυπα εκπαίδευσης, έτσι ώστε να απέχει ίσες αποστάσεις (περιθώρια-margin) από τα πλησιέστερα πρότυπα των κλάσεων και παράλληλα το περιθώριο αυτό να είναι το μέγιστο δυνατό. Τα πρότυπα των κλάσεων που απέχουν απόσταση ίση με το περιθώριο από το διαχωριστικό υπερεπίπεδο λέγονται διανύσματα στήριξης (support vectors).



Σχήμα 9: Διάγραμμα απεικόνισης του καλύτερου ποσοστού επιτυχίας του SVM Gaussian, Polynomial και Quadratic για κάθε πλήθος χαρακτηριστικών

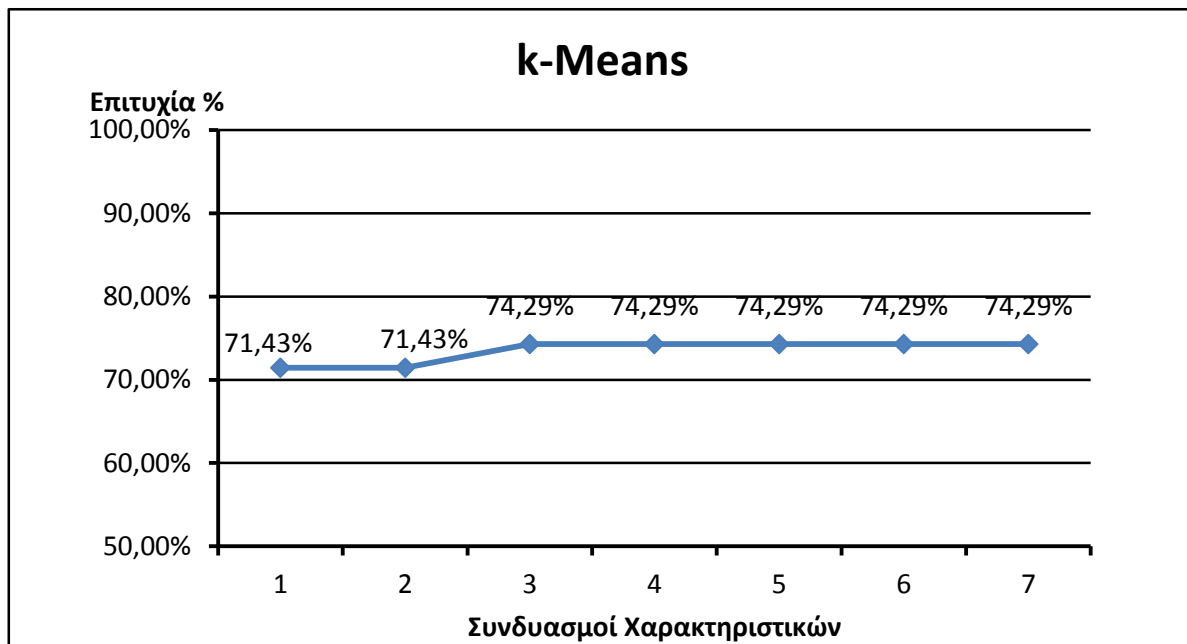
Το παραπάνω διάγραμμα απεικονίζει τα καλύτερα ποσοστά επιτυχίας ανά πλήθος χαρακτηριστικών, όπως αυτά παρουσιάστηκαν μετά την εκτέλεση του SVM classifier αλγορίθμου με kernel συνάρτηση την Gaussian, την Polynomial και την Quadratic. Όπως παρατηρείται το μέγιστο ποσοστό ακρίβειας ήταν 94.28%, το οποίο παρατηρήθηκε τόσο με την Polynomial kernel συνάρτηση όσο και με την Quadratic. Τα καλύτερα χαρακτηριστικά, τα οποία οδήγησαν σε αυτό το ποσοστό ακρίβειας, στην περίπτωση της Polynomial kernel συνάρτησης ήταν συνδυασμός επτά χαρακτηριστικών και ήταν τα Mean (χαρακτηριστικό υψής 1^{ης} τάξης), range Correlation, mean Inverse Difference Moment, mean Entropy, range Entropy, mean Gray Level Non-Uniformity και range Gray Level Non-Uniformity (από τα χαρακτηριστικά υψής 2^{ης} τάξης). Αντίστοιχα, τα καλύτερα χαρακτηριστικά, τα οποία οδήγησαν στο ίδιο ποσοστό ακρίβειας, στην περίπτωση της Quadratic kernel συνάρτησης ήταν και πάλι συνδυασμός επτά χαρακτηριστικών και ήταν τα Variance (χαρακτηριστικό υψής 1^{ης} τάξης), range Contrast, mean Correlation, range Correlation, mean Short Run Emphasis, mean Run Percentage (από τα χαρακτηριστικά υψής 2^{ης} τάξης) και Eccentricity (από τα μορφολογικά χαρακτηριστικά).

4.4.3 Αποτελέσματα unsupervised ταξινομητών

4.4.3.1 K-means clustering

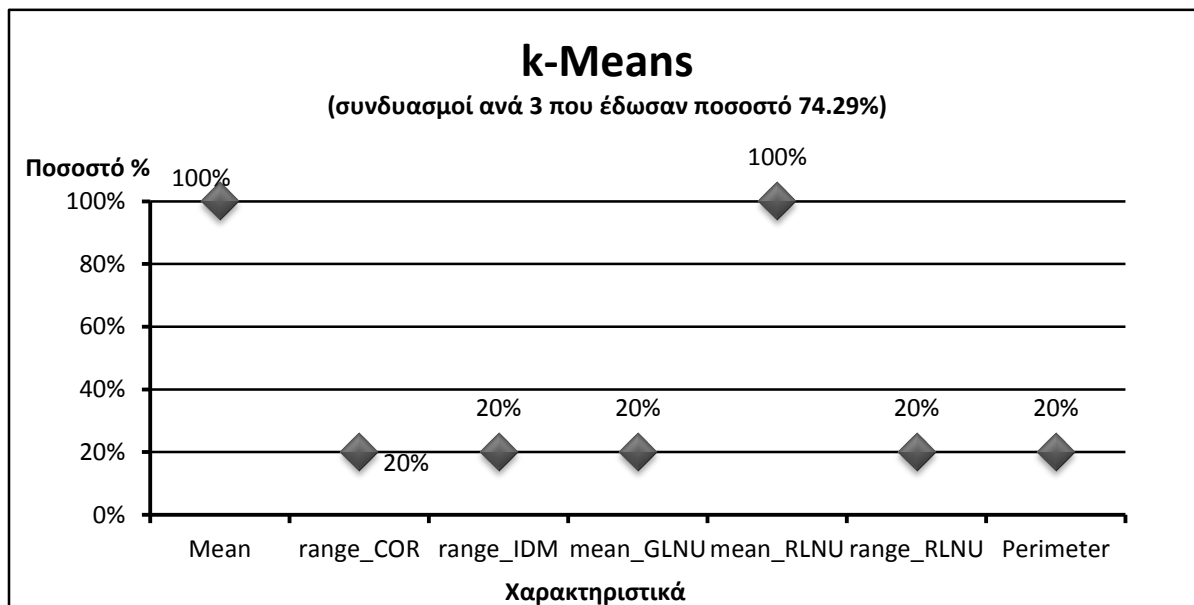
Ο αλγόριθμος k-means, στην παρούσα εργασία ομαδοποιεί τα δεδομένα εισόδου σε 2 διαφορετικές ομάδες (χαμηλού βαθμού και υψηλού βαθμού κακοήθεια) και ξεκινά θεωρώντας τυχαία έναν εκπρόσωπο για κάθε κλάση από τα πρότυπα εισόδου. Αφού τα υπόλοιπα πρότυπα ταξινομηθούν στις 2 κλάσεις σύμφωνα με το παραπάνω κριτήριο, όπως αυτό περιγράφηκε στην υποενότητα §3.5.2.1, ο εκπρόσωπος κάθε κλάσης παίρνει την τιμή της μέσης τιμής των προτύπων της κλάσης αυτής και όλα τα πρότυπα ξανα-ταξινομούνται στις 2 κλάσεις με τον ίδιο τρόπο για τους νέους εκπροσώπους των

κλάσεων. Η διαδικασία αυτή επαναλαμβάνεται έως ότου δεν θα υπάρχουν πλέον αλλαγές σε καμία από τις 2 κλάσεις.



Σχήμα 10: Διάγραμμα απεικόνισης του καλύτερου ποσοστού επιτυχίας του k-Means Clustering για κάθε πλήθος χαρακτηριστικών

Το διάγραμμα του σχήματος 10 απεικονίζει τα καλύτερα ποσοστά επιτυχίας ανά πλήθος χαρακτηριστικών, όπως αυτά παρουσιάστηκαν μετά την εκτέλεση του k-Means Clustering αλγορίθμου. Το μέγιστο ποσοστό ακρίβειας που παρατηρήθηκε ήταν 74.29%, το οποίο παρατηρήθηκε με τον συνδυασμό των τριών χαρακτηριστικών Mean (από τα χαρακτηριστικά υψής 1^{ης} τάξης), range Correlation και mean Run Length Non-Uniformity (από τα χαρακτηριστικά υψής 2^{ης} τάξης) ή των τριών χαρακτηριστικών Mean (από τα χαρακτηριστικά υψής 1^{ης} τάξης), range Inverse Different Moment και mean Run Length Non-Uniformity (από τα χαρακτηριστικά υψής 2^{ης} τάξης) ή των τριών χαρακτηριστικών Mean (από τα χαρακτηριστικά υψής 1^{ης} τάξης), mean Gray Level Non-Uniformity και mean Run Length Non-Uniformity (από τα χαρακτηριστικά υψής 2^{ης} τάξης) ή των τριών χαρακτηριστικών Mean (από τα χαρακτηριστικά υψής 1^{ης} τάξης), mean Run Length Non-Uniformity και range Run Length Non-Uniformity (από τα χαρακτηριστικά υψής 2^{ης} τάξης) ή των τριών χαρακτηριστικών Mean (από τα χαρακτηριστικά υψής 1^{ης} τάξης), mean Run Length Non-Uniformity (από τα χαρακτηριστικά υψής 2^{ης} τάξης) και Perimeter (από τα μορφολογικά χαρακτηριστικά).

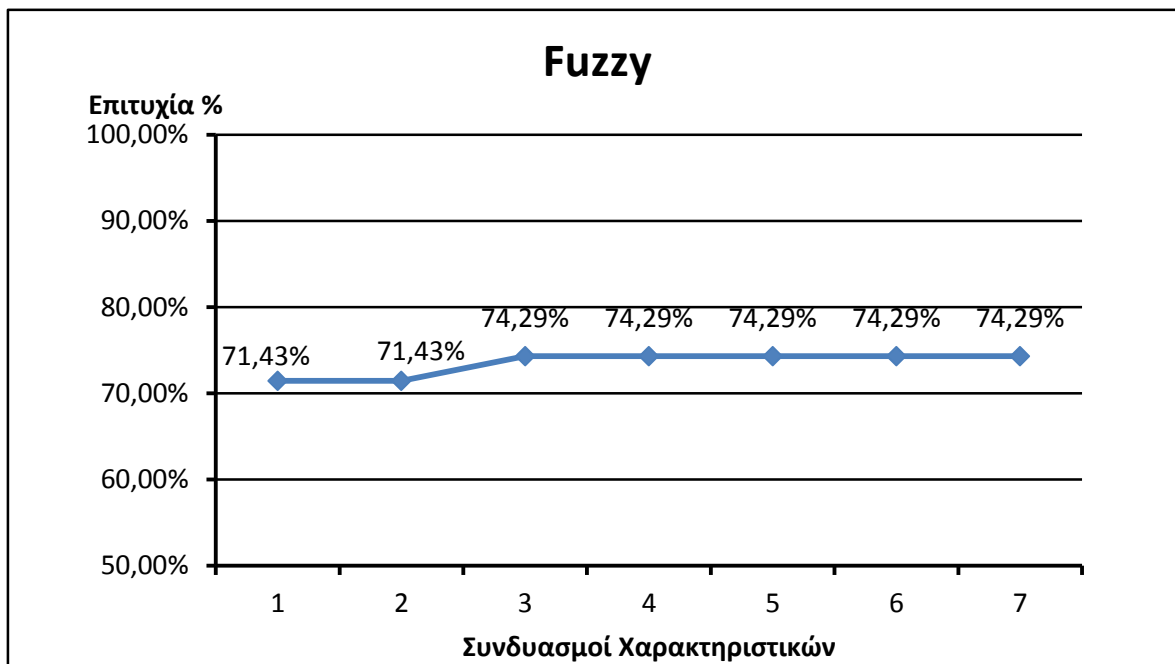


Σχήμα 11: Διάγραμμα απεικόνισης των καλύτερων χαρακτηριστικών που έδωσαν το μέγιστο ποσοστό επιτυχίας για τον k-Means Clustering (74.29% - συνδυασμός χαρακτηριστικών ανά 3)

Στο διάγραμμα του σχήματος 11 απεικονίζονται τα χαρακτηριστικά εκείνα, των οποίων ο συνδυασμός ανά τρία, έδωσαν το καλύτερο ποσοστό ακρίβειας για τον k-Means Clustering, όπως προαναφέρθηκε. Παρατηρείται ότι το χαρακτηριστικό Mean (χαρακτηριστικό υφής 1^{ης} τάξης) και το χαρακτηριστικό mean Run Length Non-Uniformity (χαρακτηριστικό υφής 2^{ης} τάξης), υπήρχαν σε όλους τους συνδυασμούς που έδωσαν ποσοστό ακρίβειας 74.29%, ενώ τα χαρακτηριστικά range Correlation, range Inverse Different Moment, mean Gray Level Non-Uniformity, range Run Length Non-Uniformity (χαρακτηριστικά υφής 2^{ης} τάξης) και το χαρακτηριστικό Perimeter (μορφολογικό χαρακτηριστικό) υπήρχαν στο ένα πέμπτο των συνδυασμών που έδωσαν το μέγιστο ποσοστό ακρίβειας που σημειώθηκε (74.29%).

4.4.3.2 Fuzzy c-means clustering

Στον αλγόριθμο Fuzzy c-means, κάθε πρότυπο έχει έναν βαθμό συσχέτισης για κάθε κλάση, που δείχνει το κατά πόσο ανήκει σε αυτήν ή όχι. Κατά επέκταση, τα πρότυπα που βρίσκονται στα όρια της κλάσης έχουν μικρότερο βαθμό συσχέτισης για αυτήν την κλάση σε σχέση με τα πρότυπα που βρίσκονται στο κέντρο της κλάσης αυτής. Κάθε πρότυπο έχει ένα σύνολο συντελεστών που ορίζουν τον βαθμό συσχέτισής του με την κλάση k. Ως κέντρο μιας κλάσης ορίζεται η μέση τιμή όλων των προτύπων της κλάσης αυτής, σταθμισμένη από τον βαθμό συσχέτισης με την κλάση αυτή. Στην παρούσα εργασία, ο αλγόριθμος Fuzzy c-means clustering εκτελέστηκε για m=2, όπου m μια παράμετρος, η οποία ελέγχει το βάρος που δίνεται για το κοντινότερο κέντρο.



Σχήμα 12: Διάγραμμα απεικόνισης του καλύτερου ποσοστού επιτυχίας του k-Means Clustering για κάθε πλήθος χαρακτηριστικών

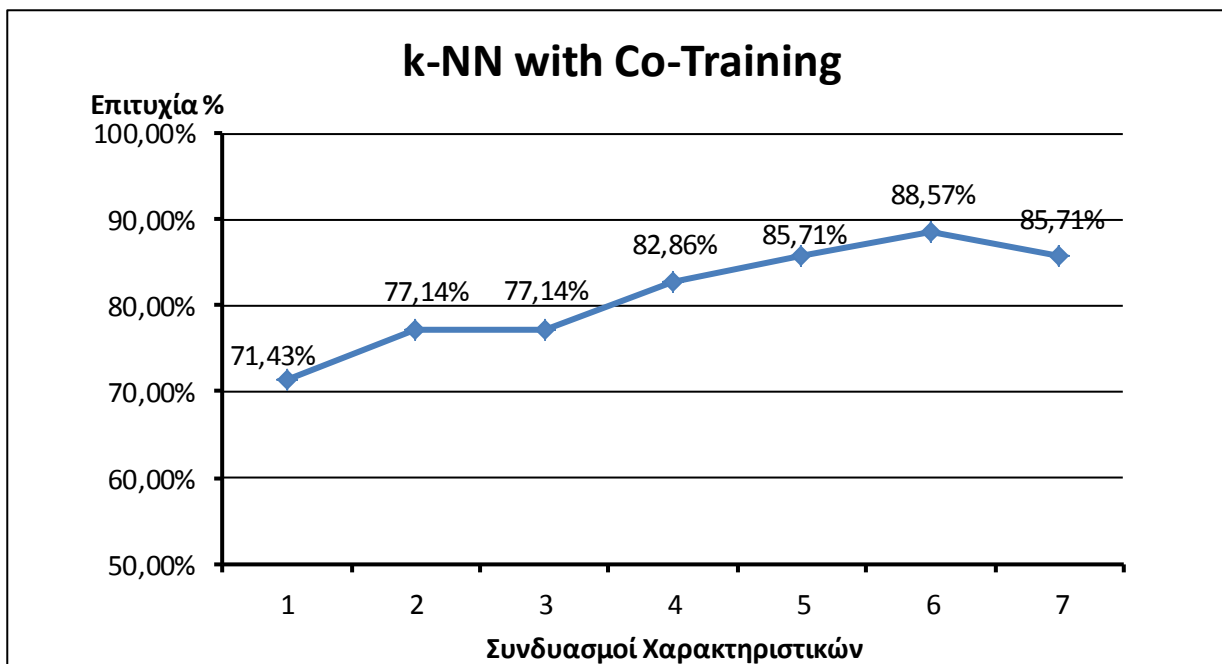
Το διάγραμμα του σχήματος 12 απεικονίζει τα καλύτερα ποσοστά επιτυχίας ανά πλήθος χαρακτηριστικών, όπως αυτά παρουσιάστηκαν μετά την εκτέλεση του Fuzzy Clustering αλγορίθμου. Το μέγιστο ποσοστό ακρίβειας που παρατηρήθηκε ήταν 74.29%, το οποίο παρατηρήθηκε με τον συνδυασμό των τριών χαρακτηριστικών Mean (από τα χαρακτηριστικά υφής 1^{ης} τάξης), range Correlation και mean Gray Level Non-Uniformity (από τα χαρακτηριστικά υφής 2^{ης} τάξης).

4.4.4 Αποτελέσματα semi-supervised ταξινομητών

Η εκπαίδευση ενός semi-supervised αλγορίθμου αποτελείται από έναν συνδυασμό γνωστών και άγνωστων δεδομένων, αξιοποιώντας τα πλεονεκτήματα τόσο των supervised όσο και των unsupervised αλγορίθμων. Ο αλγόριθμος που υλοποιήθηκε και χρησιμοποιήθηκε στα πλαίσια αυτής της εργασίας είναι ο k-nearest neighbors με Co-Training (k-NN with Co-Training).

4.4.4.1 K-nearest neighbors με Co-Training

Στην παρούσα εργασία, ο k-NN Co-Training αλγόριθμος, ο οποίος παρουσιάστηκε στην υποενότητα §3.5.3.1, εκτελέστηκε χρησιμοποιώντας τον k-NN classification αλγόριθμο για την ταξινόμηση των προτύπων σε κλάσεις με $k=5$. Επίσης, θεωρήθηκε $p=1$ και $n=1$, όπου p τα δεδομένα, τα οποία χαρακτηρίζονται με την περισσότερη βεβαιότητα ως θετικά (positive) και n τα δεδομένα, τα οποία χαρακτηρίζονται με την περισσότερη βεβαιότητα ως αρνητικά (negative). Τέλος, πρέπει να αναφερθεί ότι έγινε η παραδοχή ότι το σύνολο L των γνωστών δεδομένων αποτελούνταν αρχικά από $l=5$ δεδομένα και επομένως το σύνολο U των άγνωστων δεδομένων αποτελούνταν από τα υπόλοιπα $l'=35-5=30$ δεδομένα.



Σχήμα 13: Διάγραμμα απεικόνισης του καλύτερου ποσοστού επιτυχίας του k-NN with Co-Training για κάθε πλήθος χαρακτηριστικών

Το διάγραμμα του σχήματος 13 απεικονίζει τα καλύτερα ποσοστά επιτυχίας ανά πλήθος χαρακτηριστικών, όπως αυτά παρουσιάστηκαν μετά την εκτέλεση του k-NN with Co-Training αλγορίθμου. Το μέγιστο ποσοστό ακρίβειας που παρατηρήθηκε ήταν 88.57%, το οποίο παρατηρήθηκε με τον συνδυασμό των έξι χαρακτηριστικών Mean (από τα χαρακτηριστικά υψής 1^{ης} τάξης), mean Entropy, mean Long Run Emphasis, mean Gray Level Non-Uniformity (από τα χαρακτηριστικά υψής 2^{ης} τάξης), Perimeter και Major Axis Length (από τα μορφολογικά χαρακτηριστικά).

4.4.5 Σύνοψη και συμπεράσματα

Προκειμένου να υπολογισθεί η υποκειμενικότητα των ιατρών στην ταξινόμηση των αστροκυττωμάτων, έχουν προταθεί διάφορα υπολογιστικά συστήματα. Στην παρούσα εργασία, πραγματοποιήθηκε ψηφιοποίηση και τμηματοποίηση ιστοπαθολογικών εικόνων. Το σημαντικότερο στοιχείο μιας τέτοιας εικόνας, από το οποίο μπορούν να εξαχθούν συμπεράσματα για την διάγνωση, είναι οι πυρήνες. Έτσι, δημιουργήθηκαν αλγόριθμοι, οι οποίοι απομάκρυναν όλες τις υπόλοιπες περιοχές της εικόνας και τμηματοποίησαν τους πυρήνες για περισσότερη ανάλυση αυτών. Στην συνέχεια, πραγματοποιήθηκε εξαγωγή χαρακτηριστικών από τους τμηματοποιημένους πυρήνες. Οι βιολογικές πληροφορίες που υπάρχουν στους πυρήνες κωδικοποιήθηκαν σε ένα σύνολο μορφολογικών χαρακτηριστικών και χαρακτηριστικών υψής, τα οποία έχουν αποδειχθεί αποτελεσματικές κατηγορίες για την περιγραφή του βαθμού κακοήθειας των πυρήνων. Τέλος, δημιουργήθηκαν και εκπαιδεύτηκαν Supervised, Unsupervised και Semi-supervised συστήματα ταξινόμησης προτύπων, ώστε να μάθουν τις διαφορές μεταξύ των βαθμών κακοήθειας – χαμηλού βαθμού (I & II) και υψηλού βαθμού (III & IV) - και να ταξινομήσουν τα δεδομένα, υλοποιώντας αυτόματη διάγνωση.

Από τα αποτελέσματα των ταξινομητών, προκύπτει ότι καλύτερος supervised ταξινομητής ήταν ο SVM Classifier τόσο με την Polynomial kernel συνάρτηση όσο και με την Quadratic, καθώς έδωσε 94.29% ποσοστό ακρίβειας στην ταξινόμηση των δεδομένων, το οποίο θεωρείται ένα αρκετά ενθαρρυντικό αποτέλεσμα, αν ληφθούν

υπόψη και προηγούμενες μελέτες. Επίσης, οι unsupervised clustering (k-Means και Fuzzy) έδωσαν ποσοστό ακρίβειας 74.29%, το οποίο μπορεί να θεωρηθεί ικανοποιητικό αν ληφθεί υπόψιν ο μικρός αριθμός δεδομένων-περιστατικών. Είναι αντιληπτό πως όσο περισσότερα δεδομένα-πρότυπα υπάρχουν ως είσοδος στο σύστημα μιας unsupervised ομαδοποίησης, τόσο πιο ακριβής θα είναι η ομαδοποίηση αυτών, ελέγχοντας όμως την απότομη αύξηση της πολυπλοκότητας των αλγορίθμων. Μεγάλο ενδιαφέρον θα παρουσίαζε μια επανεξέταση αυτών σε ένα αυτόματο σύστημα ομαδοποίησης δεδομένων, όπου την είσοδο θα αποτελούσε ένα αρκετά μεγαλύτερο πλήθος προτύπων σε σχέση με αυτά που ήταν διαθέσιμα στην παρούσα εργασία. Τέλος, μέσα από την μελέτη αυτή προέκυψε ότι ο Semi-Supervised with Co-Training ταξινομητής, σημείωσε μέγιστο ποσοστό ακρίβειας 88.57%, που κρίνεται αρκετά ικανοποιητικό.

Ένα ακόμη σημαντικό σχόλιο, που κρίνεται αναγκαίο να αναφερθεί, αποτελεί το γεγονός ότι κατά την διάρκεια της επιλογής χαρακτηριστικών, κάποια χαρακτηριστικά κρίνονταν σημαντικότερα κάποιων άλλων στην απόδοση μέγιστης ακρίβειας των ταξινομητών. Συγκεκριμένα, το χαρακτηριστικό Mean (χαρακτηριστικό υψής 1^{ης} τάξης) αποτελεί βασικό χαρακτηριστικό για τους περισσότερους ταξινομητές που χρησιμοποιήθηκαν, καθώς ήταν ένα χαρακτηριστικό, το οποίο βρίσκονταν συχνότατα στην ομάδα εκείνων των χαρακτηριστικών που έδιναν το μέγιστο ποσοστό ακρίβειας. Επιπλέον, τα χαρακτηριστικά mean Correlation, range Correlation, mean Gray Level Non-Uniformity και mean Run Length Non-Uniformity (χαρακτηριστικά υψής 2^{ης} τάξης) υπήρξαν τα βασικότερα χαρακτηριστικά, καθώς σε συνδυασμό με άλλα χαρακτηριστικά σημείωσαν τα καλύτερα ποσοστά επιτυχίας για τους ταξινομητές, που χρησιμοποιήθηκαν στην παρούσα εργασία. Το χαρακτηριστικό Correlation (Συσχέτιση) πληροφορεί σχετικά με την συσχέτιση των γκρι τόνων των πυρήνων, το χαρακτηριστικό Gray Level Non-Uniformity (Ανομοιογένεια του γκρι επιπέδου) αυξάνει για τους πυρήνες που έχουν πολλαπλές διαδρομές για τις ίδιες τιμές του γκρι επιπέδου και το χαρακτηριστικό Run Length Non-Uniformity (Ανομοιογενής διαδρομή) μειώνεται για τους πυρήνες με ομοιογενή κατανομή διαδρομής. Αυτό αποτελεί ένα σημαντικό στοιχείο για το ότι η ανώμαλη και μη ομοιογενής διάσταση ενός πυρήνα προδίδει μια υψηλού βαθμού κακοήθεια.

5. ΤΗΛΕΠΑΘΟΛΟΓΙΑ

5.1 Γενικά

Η ιατρική ψηφιακή ανάλυση εικόνας, ακολουθώντας τις ραγδαίες εξελίξεις στον τομέα της τεχνολογίας των πληροφοριών, γίνεται μια όλο και περισσότερο αποδεκτή μέθοδος[46]. Νέα αυτόνομα αυτοματοποιημένα συστήματα μικροσκοπίου και πλήρως αυτοματοποιημένοι ψηφιακοί σαρωτές αναπτύχθηκαν για εξέταση ιστοκυτταρολογικών εικόνων κατά την τελευταία δεκαετία.

Αρκετά νέα ημιαυτόματα μικροσκόπια αναπτύχθηκαν, τα οποία κοινοποιούν την περιοχή σάρωσης και τις εικόνες και παράλληλα υλοποιούν την ηλεκτρονική καταγραφή των επιλεγμένων εικόνων. Η αυτοματοποιημένη ιστοπαθολογική ανάλυση αποτελεί ένα σημαντικό ερευνητικό έργο εδώ και αρκετά χρόνια αλλά τα αποτελέσματα δεν ήταν ικανοποιητικά για εφαρμογές ρουτίνας[47,48]. Ωστόσο, η ιστοπαθολογική διάγνωση μπορεί να υποστηριχτεί ήδη από νέες ηλεκτρονικές τεχνικές[49].

Την τελευταία δεκαετία δημιουργήθηκαν τηλεπαθολογικές υπηρεσίες σε δύο τεχνολογικές φόρμες. Αρχικά, η δυναμική τηλεπαθολογία περιλαμβάνει απομακρυσμένο έλεγχο του μικροσκοπικού ελέγχου, χρησιμοποιώντας απευθείας σύνδεση για μεταφορά της εικόνας σε πραγματικό χρόνο[50-52]. Η μέθοδος αυτή έχει το πλεονέκτημα της πρόσβασης σε ολόκληρη την σάρωση και όχι των μεμονωμένα επιλεγμένων εικόνων, αποφεύγοντας το πιθανό λάθος μιας λανθασμένα προεπιλεγμένης εικόνας μικροσκοπίου. Ωστόσο, τόσο το κόστος του συστήματος αυτού όσο και της εργασίας πάνω σε αυτό είναι υψηλό[53]. Η στατικές προεπιλεγμένες εικόνες, από την άλλη πλευρά, χρειάζονται αρκετά μικρότερη επένδυση υλικού, ωστόσο είναι πιθανό να συμβεί λανθασμένη επιλογή μικροσκοπικής εικόνας[54,55]. Χρησιμοποιώντας το διαδίκτυο ως τρόπο τηλεπικοινωνίας, οι στατικές εικόνες είναι μια λύση χαμηλού κόστους και ευρέως διαθέσιμη.

Η εφαρμογή ενός ψηφιακού σαρωτή αποτελεί σημαντικό πεδίο της τηλεπαθολογίας για την αυτοματοποιημένη ιστοπαθολογική ανάλυση. Δουλεύοντας πάνω στις ψηφιακές σαρώσεις, απαιτείται μια ψηφιακή ή εικονική μικροσκοπία (virtual microscopy). Παράλληλα, έχουν αναφερθεί τα θετικά αποτελέσματα πάνω σε ένα περιορισμένο αριθμό μωσαϊκών μικροσκοπικών εικόνων (mosaic image)[56].

Η απομακρυσμένη διάγνωση δεν είναι ο μόνος τομέας που μπορεί να ωφεληθεί από την τηλεπαθολογία. Η τηλεπαθολογία, έχοντας την ικανότητα της ηλεκτρονικής καταγραφής και την αποθήκευση και ανάκτηση δεδομένων, μπορεί να προσφέρει εξαιρετικές υπηρεσίες σε τομείς, όπως η εξ αποστάσεως εκπαίδευση και έρευνα. Τέλος, η τηλεπαθολογία μπορεί να είναι ένα καλό εργαλείο για υλοποίηση πολυμεσικών βάσεων δεδομένων για την διενέργεια τηλεδιασκέψεων[52].

5.2 Σύστημα εικονικού μικροσκοπίου (Virtual microscope system)

Το σύστημα εικονικού μικροσκοπίου παρέχει την δυνατότητα πρόσβασης σε υψηλής ανάλυσης ψηφιακές εικόνες από ολόκληρες ιστοπαθολογικές σαρώσεις. Οι λειτουργίες του περιλαμβάνουν την συγκέντρωση μεμονωμένων τμημάτων εικόνων σε μια ενιαία σύνθετη εικόνα, την αποθήκευση των δεδομένων της εικόνας σε μια ειδικά διαμορφωμένη βάση δεδομένων και την δοκιμή του συνόλου των δεδομένων για απεικόνιση στην οθόνη ενός τοπικού ή απομακρυσμένου χρήστη[57]. Το σύστημα απεικόνισης είναι σχεδιασμένο, ώστε να μιμείται την πραγματική δραστηριότητα μιας σάρωσης κατά μήκος μιας επιφάνειας, συμπεριλαμβανομένης της κίνησης και της εστίασης. Μια εφαρμογή ενός εικονικού μικροσκοπίου μιμείται τη συνήθη συμπεριφορά

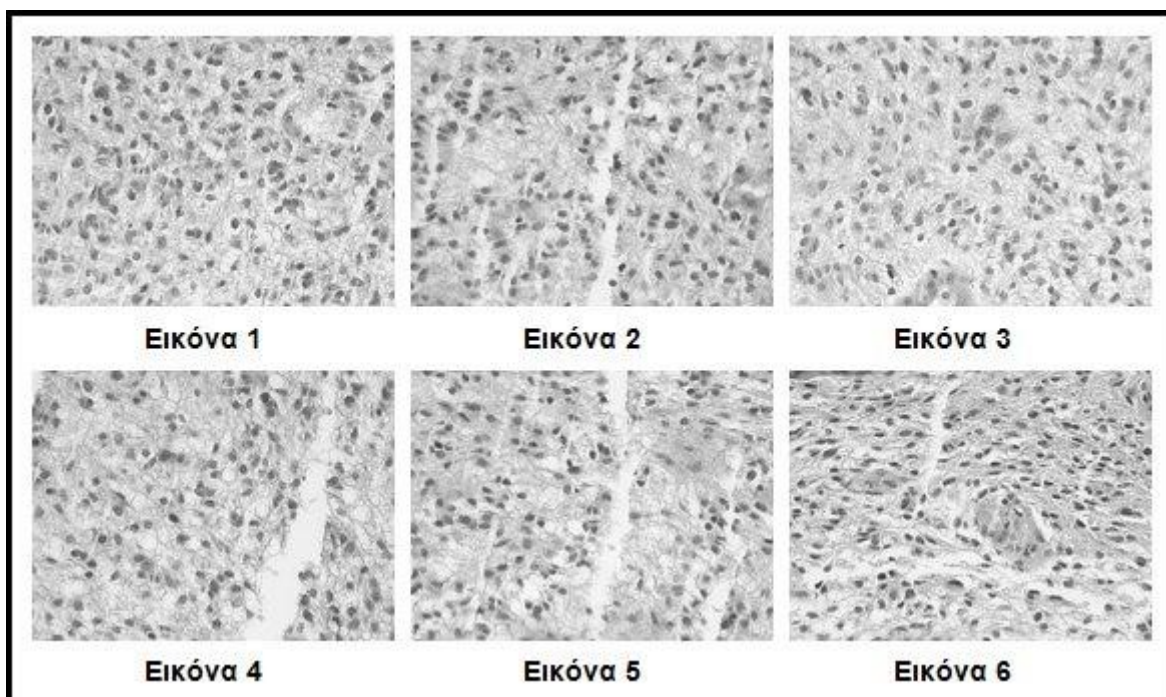
ενός φυσικού μικροσκοπίου. Η ανάκτηση μιας σάρωσης, δημιουργεί την ανάγκη για πρόσβαση στα δεδομένα της εικόνας, χωρίς όμως να απαιτείται η φυσική πρόσβαση σε αυτήν. Ωστόσο, το εικονικό μικροσκόπιο παρέχει λειτουργικότητες που ένα φυσικό μικροσκόπιο δεν μπορεί, όπως για παράδειγμα η απεικόνιση της ίδιας σάρωσης σε διαφορετικούς χρήστες την ίδια στιγμή[57].

Στην παρούσα εργασία υλοποιήθηκε εφαρμογή, η οποία χωρίζεται σε δύο τμήματα. Στο πρώτο κομμάτι, η εφαρμογή λαμβάνει ως είσοδο 6 διαφορετικές ιστοπαθολογικές εικόνες από τις περιοχές ενδιαφέροντος (ROIs) του πλακιδίου του κάθε ασθενή και με αυτές ανασυνθέτει μια πλήρη εικόνα (μωσαϊκό). Στο δεύτερο κομμάτι, παρουσιάζεται στον χρήστη το μωσαϊκό του περιστατικού που θέλει να μελετήσει και με χειροκίνητη επιλογή επιλέγει μια από τις εικόνες που συνθέτουν το μωσαϊκό για περαιτέρω μελέτη. Στην συνέχεια, η εικόνα που επιλέχθηκε εισάγεται στο σύστημα αναγνώρισης προτύπων, το οποίο παρουσιάστηκε στην υποπαράγραφο §3.5, για δημιουργία χαρακτηριστικών και ταξινόμηση.

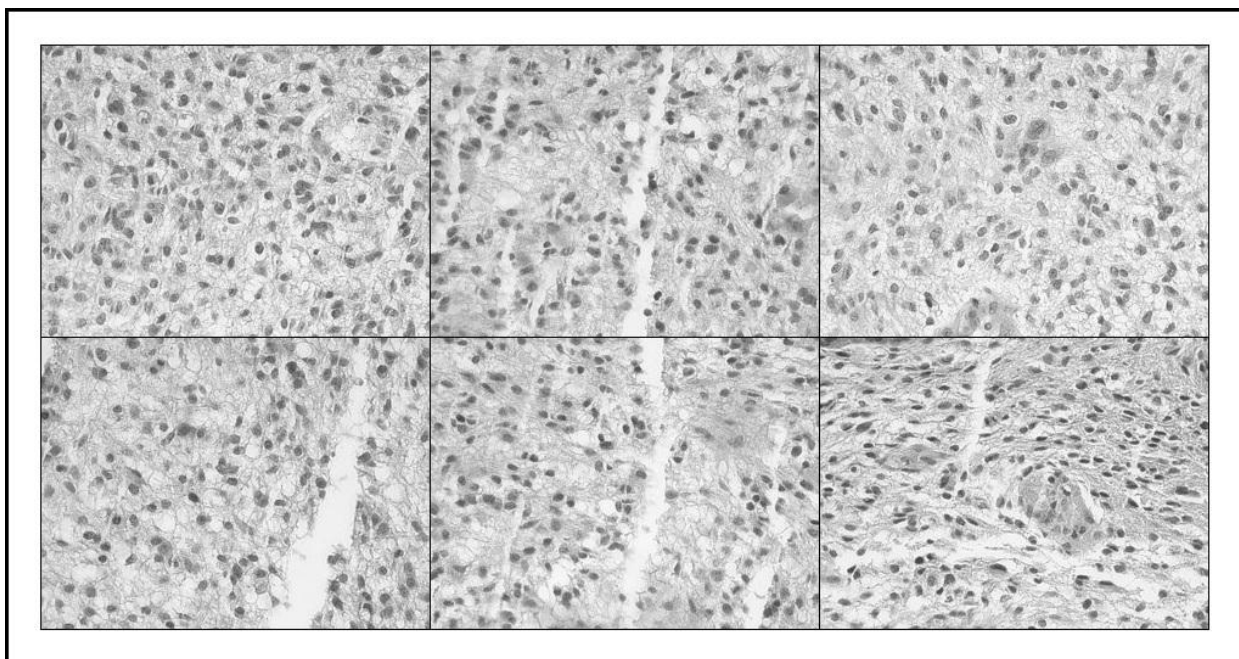
Για την ανάπτυξη της εφαρμογής αυτής χρησιμοποιήθηκε το λογισμικό MATLAB (The MathWorks Inc., Version R2008a).

✚ Δημιουργία μωσαϊκού

Η εφαρμογή αυτή δέχεται ως είσοδο 6 ιστοπαθολογικές εικόνες (εικόνα 20), η οποίες πάρθηκαν από το πλακίδιο ενός ασθενή/περιστατικού και τις συνθέτει σε μια ενιαία εικόνα (μωσαϊκό - εικόνα 21).



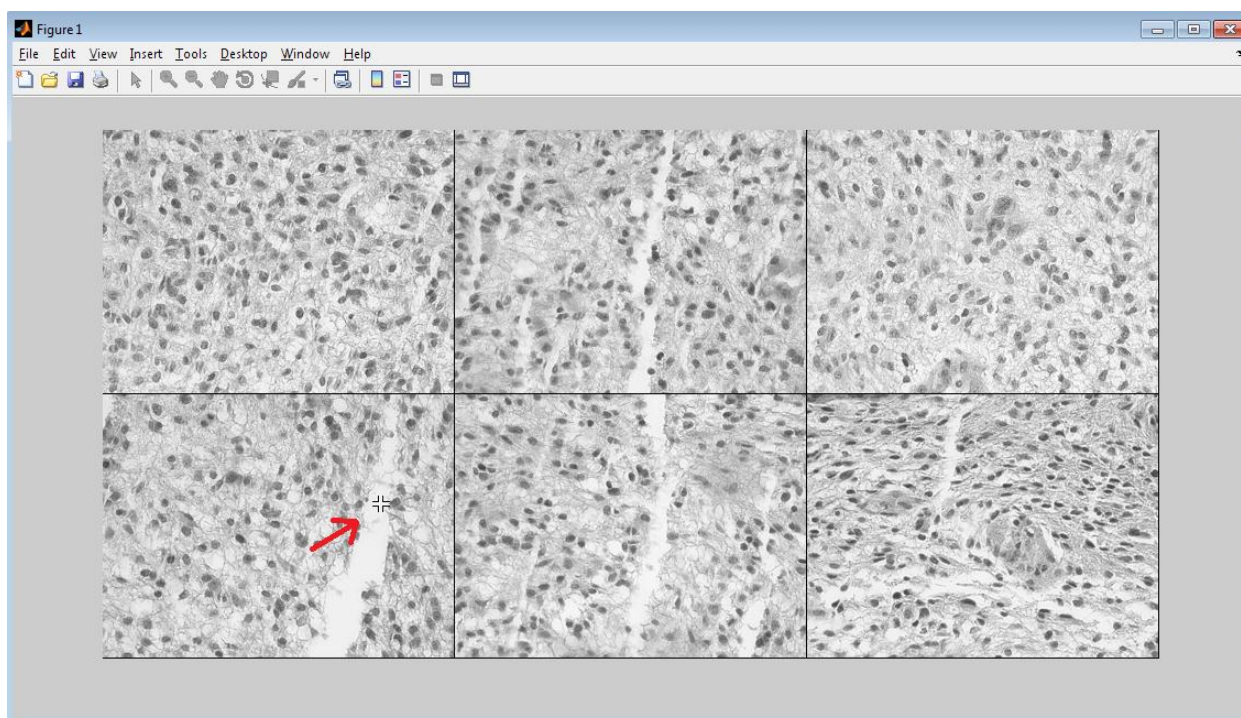
Εικόνα 20: Παράδειγμα 6 εικόνων που δέχεται ως είσοδο το σύστημα



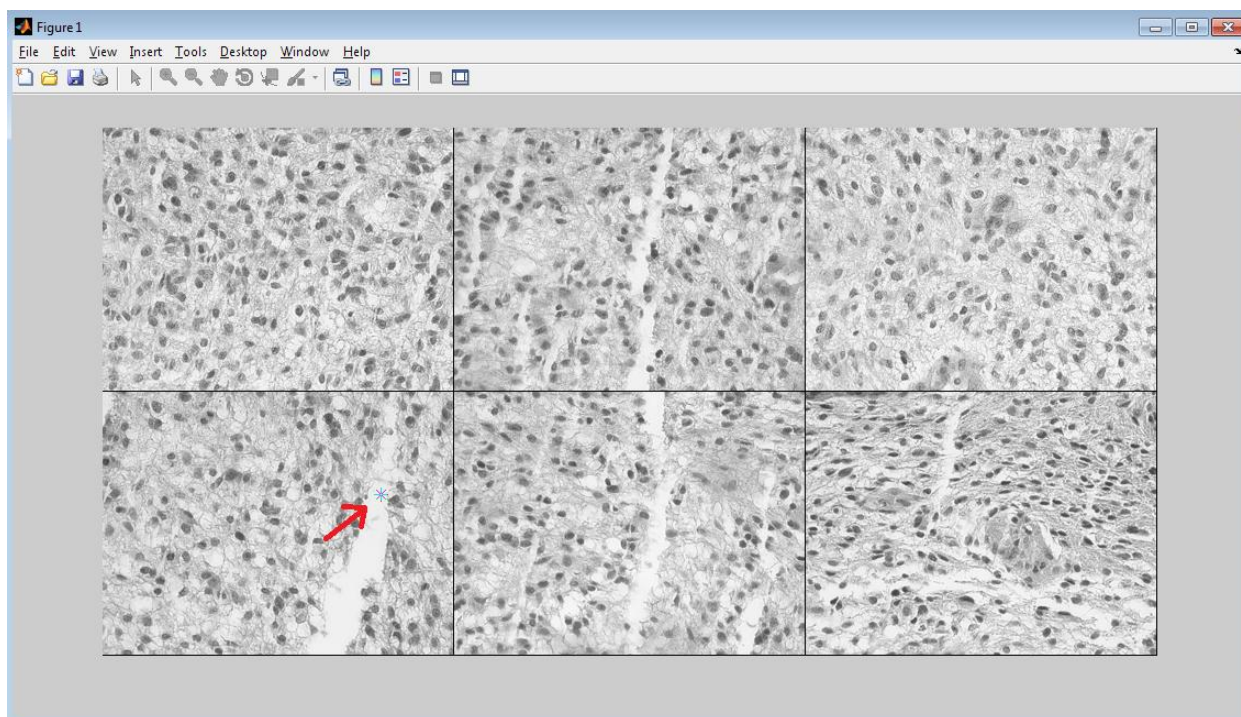
Εικόνα 21: Το τελικό μωσαϊκό που δημιουργείται από τις 6 εικόνες εισόδου

✚ Επιλογή εικόνας από το μωσαϊκό

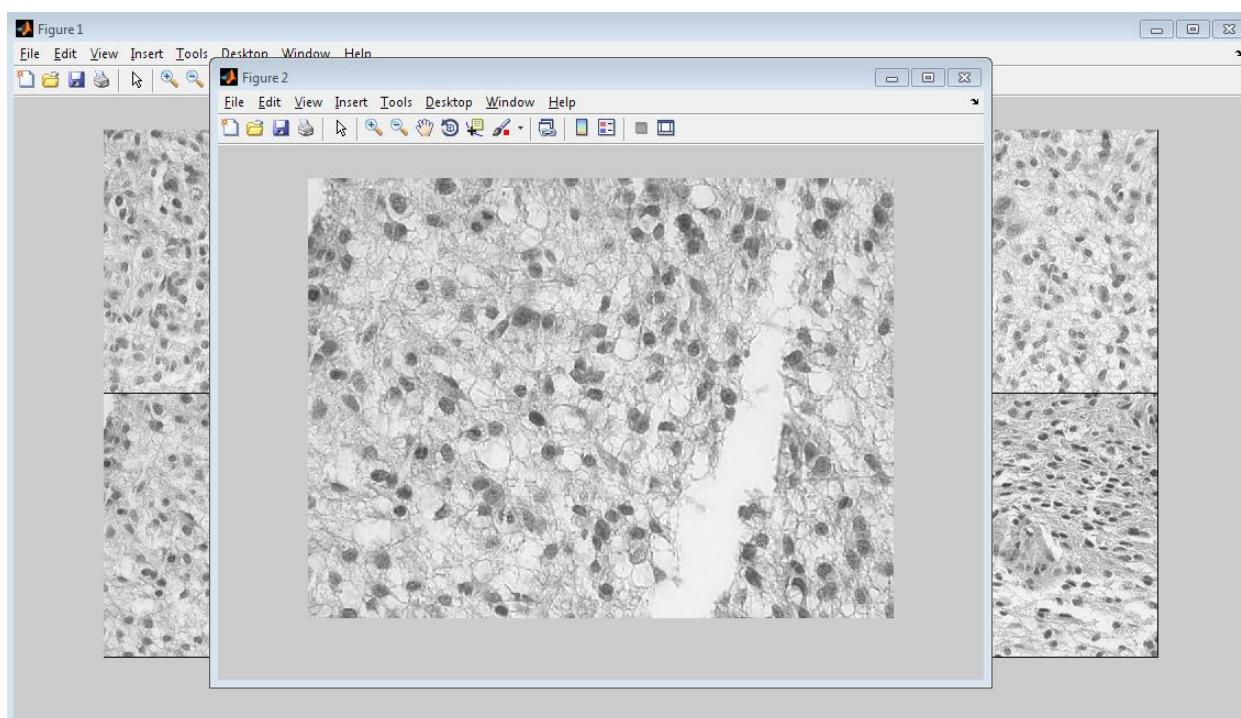
Ο χρήστης, έχοντας στην κατοχή του το μωσαϊκό, καλείται να επιλέξει ποια από τις εικόνες που το αποτελούν επιθυμεί να μελετήσει περαιτέρω και κατά επέκταση να εισάγει στο σύστημα δημιουργίας χαρακτηριστικών και αναγνώρισης προτύπων. Στην εικόνα 22 φαίνεται η πλατφόρμα εμφάνισης του μωσαϊκού, στην οποία ο χρήστης μπορεί να επιλέξει με χρήση του ποντικιού την εικόνα που επιθυμεί (εικόνα 23). Με το πάτημα του κουμπιού *enter*, η εικόνα αυτή εμφανίζεται στον χρήστη (εικόνα 24) και στέλνεται στο σύστημα αναγνώρισης προτύπων.



Εικόνα 22: Ο χρήστης επιλέγει εικόνα από το μωσαϊκό



Εικόνα 23: Ο χρήστης κατοχύρωσε με το ποντίκι την επιλογή του



Εικόνα 24: Με το πάτημα του κουμπιού *enter* εμφανίζεται η εικόνα που ο χρήστης επέλεξε

6. ΣΥΝΟΨΗ ΚΑΙ ΣΥΜΠΕΡΑΣΜΑΤΑ

Η ταξινόμηση των αστροκυττωμάτων είναι στενά συνδεδεμένη με την σοβαρότητα της νόσου του καρκίνου του εγκεφάλου και χρησιμεύει στην διαχείριση και θεραπεία του ασθενή. Η πρόγνωση των αστροκυττωμάτων ποικίλει σημαντικά και μπορεί να εκτιμηθεί χρησιμοποιώντας την ταξινόμηση κατά την οποία ελέγχεται η ιστολογική προέλευση των όγκων, η ανατομική εντόπισή τους και η κλινική τους πορεία. Μια λανθασμένη διάγνωση θα οδηγούσε σε ανεπαρκή θεραπεία σε υψηλού βαθμού όγκους ή επιβλαβή θεραπεία σε χαμηλού βαθμού όγκους.

6.1 Σύνοψη και συμπεράσματα

Στο πλαίσιο της παρούσας εργασίας διερευνάται αρχικά πραγματοποιήθηκε συλλογή δεδομένων, όπου δείγματα ιστοπαθολογικού ιστού συλλέχτηκαν από 35 ασθενείς από το Πανεπιστημιακό Νοσοκομείο της Πάτρας, οι οποίοι είχαν διαγνωστεί με αστροκυτταρικό όγκο και έπειτα η προετοιμασία των δεδομένων κατά την οποία τα δείγματα ιστού χρωματίστηκαν με H&E (Hematoxylin & Eosin), ώστε να φανερωθούν τα διάφορα συστατικά των κυττάρων και των πυρήνων αλλά και για να διακριθούν μεταξύ τους. Στην συνέχεια, πραγματοποιήθηκε διάγνωση δεδομένων κατά την οποία ένας ιστοπαθολόγος εξέτασε τις χρωματισμένες εικόνες των δειγμάτων ιστού και πραγματοποίησε διάγνωση με βάση ιστολογικά κριτήρια σε τρεις βαθμούς κακοήθειας (I, II, III ή IV) σύμφωνα με τον Παγκόσμιο Οργανισμό Υγείας (World Health Organization - WHO). Επίσης, ο ίδιος ιστοπαθολόγος σημείωσε πάνω στις πλάκες των δειγμάτων την πιο αντιπροσωπευτική περιοχή (Region Of Interest - ROI) για την εξέταση αυτών. Έτσι, ακολούθησε η ψηφιοποίηση εικόνας, όπου από κάθε δείγμα ψηφιοποιήθηκαν εικόνες, η οποίες πάρθηκαν από το προκαθορισμένο ROI και εφαρμόστηκε τμηματοποίηση πυρήνων. Το σημαντικότερο στοιχείο μιας τέτοιας εικόνας, από το οποίο μπορούν να εξαχθούν συμπεράσματα για την διάγνωση, είναι οι πυρήνες. Έτσι, δημιουργήθηκαν αλγόριθμοι, οι οποίοι απομάκρυναν όλες τις υπόλοιπες περιοχές της εικόνας και τμηματοποίησαν τους πυρήνες για περισσότερη ανάλυση αυτών. Στην συνέχεια, πραγματοποιήθηκε εξαγωγή χαρακτηριστικών από τους τμηματοποιημένους πυρήνες και οι βιολογικές πληροφορίες που υπάρχουν στους πυρήνες κωδικοποιήθηκαν σε ένα σύνολο μορφολογικών χαρακτηριστικών και χαρακτηριστικών υψής, τα οποία έχουν αποδειχθεί αποτελεσματικές κατηγορίες για την περιγραφή του βαθμού κακοήθειας των πυρήνων. Τέλος, χρησιμοποιήθηκε αναγνώριση προτύπων και ταξινόμηση αυτών για αυτόματη διάγνωση. Με βάση τα παραπάνω χαρακτηριστικά, δημιουργήθηκαν και εκπαιδεύτηκαν Supervised, Unsupervised και Semi-supervised συστήματα ταξινόμησης προτύπων, ώστε να μάθουν τις διαφορές μεταξύ των βαθμών κακοήθειας I, II, III και IV.

Βασικά συμπεράσματα, που προκύπτουν από την παρούσα εργασία, αποτελούν τα παρακάτω:

- Τα αποτελέσματα της τμηματοποίησης των εικόνων ήταν ενθαρρυντικά δεδομένου ότι η χρώση H&E δεν είναι ένα αρκετά αυστηρό πρωτόκολλο χρώσης, ώστε να ξεχωρίζουν έντονα οι πυρήνες από τον περιβάλλοντα ιστό. Απομονώθηκαν κατά μέσο όρο περίπου 98 πυρήνες από κάθε εικόνα και για κάθε δείγμα-ασθενή υπήρχαν 6 εικόνες ψηφιοποιημένες από τα δεδομένα της βιοψίας. Επομένως, για τον υπολογισμό των χαρακτηριστικών μελετήθηκαν συνολικά κατά μέσο όρο 588 τμηματοποιημένοι πυρήνες για κάθε δείγμα-ασθενή, ενόσω έχει αποδειχθεί ότι ακόμη και 200 ορθά τμηματοποιημένοι πυρήνες είναι επαρκείς για την εξαγωγή χαρακτηριστικών.

- Ο SVM Supervised Classifier (Polynomial ή Quadratic kernel) αποτελεί μια λύση στην αυτοματοποιημένη ταξινόμηση αστροκυττωμάτων, καθώς έδωσε 94.29% ποσοστό ακρίβειας στην ταξινόμηση των δεδομένων, το οποίο θεωρείται ένα αρκετά ενθαρρυντικό αποτέλεσμα.
- Οι unsupervised clustering k-Means και Fuzzy έδωσαν ποσοστό ακρίβειας 74.29%, το οποίο μπορεί να θεωρηθεί ικανοποιητικό αν ληφθεί υπόψιν ο μικρός αριθμός δεδομένων-περιστατικών.
- Ο Semi-Supervised with Co-Training ταξινομητής έδωσε μέγιστο ποσοστό ακρίβειας 88.57%.
- Τα χαρακτηριστικά Mean (χαρακτηριστικό υψής 1^{ης} τάξης), mean Correlation, range Correlation, mean Gray Level Non-Uniformity και mean Run Length Non-Uniformity (χαρακτηριστικά υψής 2^{ης} τάξης) αναδείχθηκαν τα χαρακτηριστικά εκείνα με το μεγαλύτερο ποσοστό πιθανότητας εμφάνισης στο διάγραμμα χαρακτηριστικών εκείνο, το οποίο δίνει το μέγιστο (καλύτερο) ποσοστό ακρίβειας ενός ταξινομητή. Αυτό αποτελεί ένα σημαντικό στοιχείο για το ότι η ανώμαλη και μη ομοιογενής διάσταση ενός πυρήνα προδίδει μια υψηλού βαθμού κακοήθεια.

6.2 Περιορισμοί της έρευνας

Σημαντικό περιορισμό της παρούσας έρευνας αποτέλεσε ο παράγοντας χρόνος, ο οποίος έθεσε περιορισμούς στην συλλογή της πληθώρας δεδομένων. Όπως έχει ήδη αναφερθεί, όσο περισσότερα δεδομένα-πρότυπα υπάρχουν ως είσοδος στο σύστημα μιας supervised ταξινόμησης ή unsupervised ομαδοποίησης, τόσο πιο ακριβής θα είναι η ταξινόμηση ή η ομαδοποίηση αυτών, ελέγχοντας όμως την απότομη αύξηση της πολυπλοκότητας των αλγορίθμων. Θα ήταν, επίσης, επιθυμητή μια εξονυχιστική αναζήτηση των αποδόσεων των ταξινομητών εκείνων, οι οποίοι εξαρτώνται έως κάποιο βαθμό από τις τιμές διαφόρων μεταβλητών, εκτελώντας τους για μεγαλύτερο πλήθος διαφορετικών τιμών των μεταβλητών αυτών και ελέγχοντας την απόδοσή τους σε κάθε περίπτωση.

7. ΜΕΛΛΟΝΤΙΚΕΣ ΠΡΟΟΠΤΙΚΕΣ

Η εξαγωγή χαρακτηριστικών είναι ένα πολύ κρίσιμο βήμα και οι επιπτώσεις της είναι στενά συνδεδεμένες με τα αποτελέσματα της ταξινόμησης. Όπως και στην παρούσα εργασία, στις περιπτώσεις που υπάρχει περιορισμένος αριθμός διαθέσιμων δεδομένων και μεγάλος όγκος διαθέσιμων χαρακτηριστικών, είναι πιθανό τα χαρακτηριστικά με χαμηλή ή μηδενική διακριτική ισχύ να αυξάνουν τον θόρυβο μειώνοντας την γενικευτική ικανότητα του ταξινομητή σε άγνωστα δεδομένα. Μελλοντικά θα παρουσιάζε μεγάλο ενδιαφέρον να γίνει προσπάθεια δημιουργίας ενός αποδοτικού αλγορίθμου μείωσης του θορύβου αυτού, κάτι που θα έκανε ευκολότερο τον διαχωρισμό ενός χαμηλού βαθμού αστροκυττώματος από ένα υψηλού βαθμού.

Αφού εξαχθούν τα χαρακτηριστικά ακολουθεί η διαδικασία επιλογής των αποδοτικότερων από αυτά, επιλέγοντας το διάνυσμα εκείνο με το μικρότερο πλήθος χαρακτηριστικών και παράλληλα με την μέγιστη ακρίβεια ταξινόμησης. Η επιλογή κατάλληλων χαρακτηριστικών μπορεί να βοηθήσει στη μείωση του χρόνου εκτέλεσης ενός αλγορίθμου και να οδηγήσει στην αποφυγή υπερβολικής εκπαίδευσης του ταξινομητή. Δυστυχώς, δεν υπάρχει καμία εγγύηση ότι η διαδικασία επιλογής χαρακτηριστικών θα αποδώσει καλύτερο ποσοστό ακρίβειας του ταξινομητή, επομένως, τα επιλεγμένα χαρακτηριστικά για την ταξινόμηση απαιτούν εγκυρότερη επικύρωση, μέσα από μια μελλοντική έρευνα.

Η επιλογή των χαρακτηριστικών και του ταξινομητή που θα χρησιμοποιηθεί για την ταξινόμηση των δεδομένων πρέπει να γίνουν πολύ προσεκτικά. Ο στόχος μιας επιτυχημένης ταξινόμησης είναι να δημιουργηθεί ένας αξιόπιστος ταξινομητής, ο οποίος θα έχει την ικανότητα να ταξινομεί άγνωστα δείγματα με πιθανό σφάλμα σε εύλογο οριοθετημένο εύρος. Η απόδοση ενός αλγορίθμου ταξινόμησης μπορεί να μεταβάλλεται με την επιλογή της μεθόδου της επιλογής χαρακτηριστικών. Έτσι, είναι απαραίτητο να προσδιοριστεί το καλύτερο ζευγάρι ενός ταξινομητή και μιας μεθόδου επιλογής χαρακτηριστικών.

ΠΙΝΑΚΑΣ ΟΡΟΛΟΓΙΑΣ

Ξενόγλωσσος όρος	Ελληνικός Όρος
Ependyma	Επένδυμα
Blood-brain barrier	Αιματοεγκεφαλικός φραγμός
Grade	Βαθμός
Hematoxylin & Eosin	Αιματοξυλίνη - Ηωσίνη
World Health Organization	Παγκόσμιο Οργανισμό Υγείας
Mean value	Μέση τιμή
Variance	Διακύμανση
Skewness	Ασυμμετρία
Kurtosis	Κύρτωση
Angular Second Moment	Γωνιακή Δεύτερη Στιγμή
Contrast	Αντίθεση
Correlation	Συσχέτιση
Inverse Different Moment	Αντίστροφη Διαφορετική Στιγμή
Entropy	Εντροπία
Short Run Emphasis	Έμφαση Μικρής Διαδρομής
Long Run Emphasis	Έμφαση Μεγάλης Διαδρομής
Gray Level Non-Uniformity	Ανομοιογένεια Γκρι Επιπέδου
Run Length Non-Uniformity	Ανομοιογένεια Διαδρομής
Run Percentage	Ποσοστό Διαδρομών
Area	Περιοχή
Perimeter	Περίμετρος
Minor axis length	Μήκος ελάχιστου άξονα
Major axis length	Μήκος μέγιστου άξονα
Eccentricity	Εκκεντρικότητα
Segmentation	Κατάτμηση
Training stage	Στάδιο εκμάθησης
Testing stage	Στάδιο δοκιμών
Exhaustive search algorithm	Αλγόριθμος πλήρους αναζήτησης
Clustering	Ομαδοποίηση
Classification	Ταξινόμηση
Region Of Interest	Περιοχή ενδιαφέροντος
Mosaic Image	Μωσαϊκό εικόνας
Virtual Microscopy	Εικονικό μικροσκόπιο

ΣΥΝΤΜΗΣΕΙΣ – ΑΡΚΤΙΚΟΛΕΞΑ – ΑΚΡΩΝΥΜΙΑ

ΚΝΣ	Κεντρικό Νευρικό Σύστημα
ΕΝΥ	Εγκεφαλονωτιαίο Υγρό
MRI	Μαγνητική τομογραφία
CT	Αξονική τομογραφία
PET	Τομογραφία εκπομπής ποζιτρονίων
WHO	Παγκόσμιος Οργανισμός Υγείας
H&E	Αιματοξυλίνη - Ηωσίνη
ASM	Angular Second Moment
CON	Contrast
COR	Correlation
VAR	Variance
IDM	Inverse Different Moment
ENT	Entropy
SRE	Short Run Emphasis
LRE	Long Run Emphasis
GLNU	Gray Level Non-Uniformity
RLNU	Run Length Non-Uniformity
RP	Run Percentage
MDC	Minimum Distance Classifier
LSMD	Least Squares Minimum Distance
k-NN	K-Nearest Neighbors
PNN	Probabilistic Neural Network
SVM	Support Vector Machine
RBF	Radial Basis Function
ROI	Region Of Interest

ΑΝΑΦΟΡΕΣ

- [1] L. DeAngelis, "Medical progress: Brain Tumours", *New England Journal of Medicine*, vol. 344, pp. 114-123, 2001
- [2] W. Shapiro and J. Shapiro, "Biology and treatment of malignant gliomas," *Oncology*, vol. 12, pp. 233-240, 1998.
- [3] D. Parkin, F. Bray, J. Ferlay, and P. Pisani, "Estimating the world cancer burden: GLOBOCAN 2000," *International Journal of Cancer*, vol. 94, pp. 153-156, 2001.
- [4] V. Levin, S. Leibel, and P. Gutin, "Neoplasms of the central nervous system," in *Cancer: Principles and Practice of Oncology*, 6th ed: Lippincott Williams & Wilkins, 2001, pp. 2100-60.
- [5] A. R. Moss, "Occupational exposure and brain tumors," *Journal of Toxicology and Environmental Health*, vol. 16, pp. 703-11, 1985.
- [6] F. H. Hochberg, G. Miller, R. T. Schooley, M. S. Hirsch, P. Feorino, and W. Henle, "Central-nervous-system lymphoma related to Epstein-Barr virus," *New England Journal of Medicine*, vol. 309, pp. 745-8, 1983.
- [7] A. Hutter, K. E. Schwetye, A. J. Bierhals, and R. C. McKinstry, "Brain neoplasms: epidemiology, diagnosis, and prospects for cost-effective imaging," *Neuroimaging Clinics in North America*, vol. 13, pp. 237-50, 2003.
- [8] C. Haskell, "Brain," in *Cancer Treatment*, 5th ed. Philadelphia: WB Saunders Co, 2001, pp. 1106-42.
- [9] G. J. Kitange, K. L. Templeton, and R. B. Jenkins, "Recent advances in the molecular genetics of primary gliomas," *Current Opinions in Oncology*, vol. 15, pp. 197-203, 2003.
- [10] Γ. Καρκαβέλας, "Η παθολογική ανατομική των γλοιωμάτων", Ιατρική Σχολή ΑΠΘ.
- [11] W. H. Organization, *World Health Organization international histological classification of tumors: Histological typing of tumors of the central nervous system*, 2nd edition ed. Berlin: Springer-Verlag, 1993.
- [12] P. E. Ricci, "Imaging of adult brain tumors," *Neuroimaging Clinics in North America*, vol. 9, pp. 651-69, 1999.
- [13] M. Westphal, D. C. Hilt, E. Bortey, P. Delavault, R. Olivares, P. C. Warnke, I. R. Whittle, J. Jaaskelainen, and Z. Ram, "A phase 3 trial of local chemotherapy with biodegradable carmustine (BCNU) wafers (Gliadel wafers) in patients with primary malignant glioma," *Journal of Neuro-Oncology*, vol. 5, pp. 79-88, 2003.
- [14] M. W. Davidson and M. Abramowitz, "Optical Microscopy", <http://micro.magnet.fsu.edu/primer/opticalmicroscopy.html>.
- [15] http://www.physics.ntua.gr/~micronano/ERG%20TECHNIKES/Simeioseis_TEM.pdf
- [16] <http://www.routias-lab.gr/files/Download/%CE%9C%CE%B9%CE%BA%CF%81%CE%BF%CF%83%CE%BA%CE%BF%CF%80%CE%AF%CE%B1.pdf>
- [17] http://www.biology.uoc.gr/courses/BIO102_zoologia/For%20website/LAB01.%20Introduction%20to%20the%20lab.pdf
- [18] L. R. Schad, H. P. Schmitt, C. Oberwittler, and W. J. Lorenz, "Numerical grading of astrocytomas," *Medical Informatics*, vol. 12, pp. 11-22, 1987.
- [19] D. Glotsos, P. Spyridonos, P. Petalas, D. Cavouras, P. Ravazoula, S. Dadioti, E. Lekka, G. Nikiforidis, "Computer-based malignancy grading of astrocytomas employing a Support Vector Machine classifier, the WHO grading system and the regular Haematoxylin-eosin diagnostic staining procedure, *Analytical and Quantitative Cytology and Histology*", 26(2): 77-83, 2004.
- [20] P. Spyridonos, D. Cavouras, P. Ravazoula, and G. Nikiforidis, "A computer-based diagnostic and prognostic system for assessing urinary bladder tumour grade and predicting cancer recurrence," *Medical Informatics and the Internet in Medicine*, vol. 27, pp. 111-22, 2002.
- [21] C. Decaestecker, I. Camby, N. Nagy, J. Brotchi, R. Kiss, and I. Salmon, "Improving morphology-based malignancy grading schemes in astrocytic tumors by means of computer-assisted techniques," *Brain Pathology*, vol. 8, pp. 29-38, 1998.
- [22] S. Theodoridis and K. Koutroubas, *Pattern Recognition*: Academic Press, 1999.
- [23] R. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 3, pp. 610-621, 1973.
- [24] M. M. Galloway, "Texture Analysis Using Grey Level Run Lengths," *Computer Graphics and Image Processing*, vol. 4, pp. 172-179, 1975.
- [25] N. Street, "Cancer diagnosis and prognosis via linear programming based machine learning," P.h.D, Madison: University of Wisconsin, 1994.
- [26] <http://www.mathworks.com/help/images/ref/regionprops.html>
- [27] G. J. McLachlan, *Discriminant analysis and Statistical Pattern Recognition*. New York: Wiley, 1992.

- [28] C. Ambrose and G. J. McLachlan, "Selection bias in gene extraction on the basis of microarray gene-expression data," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, pp. 6562-6566, 2002.
- [29] S. Theodoridis and K. Koutroumbas, *Pattern recognition*, 2nd ed. Amsterdam; Boston: Academic Press, 2003.
- [30] E. Parzen, "On the estimation of a probability density function and the mode," *Annals of Mathematical Statistics*, vol. 33, pp. 1962, 1962
- [31] <http://www.jvrb.org/past-issues/3.2006/760>
- [32] N. Belacel and M. Boulassel, "Multicriteria fuzzy assignment method: a useful tool to assist medical diagnosis," *Artificial intelligence in medicine*, vol. 21, pp. 201-207, 2001.
- [33] D. Cai, X. He, J. Han, "Semi-supervised Discriminant analysis".
- [34] A. Blum, T. Mitchell, "Combining Labeled and Unlabeled Data with Co-Training", pp. 92-100, 1998.
- [35] X. Zhu, "Semi-supervised learning literature survey", 2008.
- [36] X. Zhu, "Semi-supervised learning with graphs", 2005.
- [37] J. F. Brenner, T. F. Necheles, I. A. Bonacossa, R. Fristensky, B. A. Weintraub, and P. W. Neurath, "Scene segmentation techniques for the analysis of routine bone marrow smears from acute lymphoblastic leukemia patients," *Journal of Histochemistry & Cytochemistry*, vol. 25, pp. 601-13, 1977.
- [38] J. Brenner, J. Lester, and W. Selles, "Scene segmentation in automated histopathology: techniques evolved from cytology automatation," *Pattern Recognition*, vol. 13, pp. 65-77, 1981.
- [39] H.-S. Wu and J. Barba, "An efficient semi-automatic algorithm for cell contour extraction," *Journal of microscopy*, vol. 179, pp. 270-276, 1995.
- [40] D. L. Rosenthal, D. Acosta, and R. K. Peters, "Computed-Assisted Rescreening of Clinically Important False Negative Cervical Smears Using the PAPNET Testing System," *Acta Cytologica*, vol. 40, pp. 120-126, 1996.
- [41] W. Lie, "Automatic target segmentation by locally adaptive image thresholding," *IEEE Transactions on Image Processing*, vol. 4, pp. 1036-1041, 1995.
- [42] J. Canny, "Computational approach to edge detection," *IEEE Transactions of Pattern Analysis and Machine Intelligence*, vol. PAMI-8, pp. 679-698, 1986.
- [43] W. Kernohan, R. F. Mabon, and H. J. Svien, "A simplified classification of gliomas," presented at Proc Staff Meet Mayo Clin, 1949.
- [44] C. Decaestecker, I. Salmon, O. Dewitte, I. Camby, P. Van Ham, J. L. Pasteels, J. Brotchi, and R. Kiss, "Nearest-neighbor classification for identification of aggressive versus nonaggressive low-grade astrocytic tumors by means of image cytometry-generated variables," *Journal of Neurosurgery*, vol. 86, pp. 532-7, 1997.
- [45] H. Kolles, A. von Wangenheim, J. Rahmel, I. Niedermayer, and W. Feiden, "Data-driven approaches to decision making in automated tumor grading. An example of astrocytoma grading," *Analytical and Quantitative Cytology and Histology*, vol. 18, pp. 298-304, 1996.
- [46] Shotton DM (1995) Robert Feulgen Prize Lecture 1995. Electronic light microscopy: present capabilities and future prospects. *Histochem Cell Biol* 104:97-137
- [47] Belhomme P, Elmoataz A, Herlin P, Bloyet D (1996) Generalised region growing operator with optimal scanning: application to segmentation of breast cancer images. *J Microsc* 886:41-50
- [48] Ong SH, Jin XC, Sinniah R (1996) Image analysis of tissue sections. *Comput Biol Med* 26:269-279
- [49] Dictor M (1997) The surgical pathologist in a client/server computer network: work support, quality assurance, and the graphical user interface. *Mod Pathol* 10:259-266
- [50] Dunn BE, Almagro UA, Choi H, et al (1997) Dynamic-robotic telepathology: department of Veteran Affairs feasibility. *Hum Pathol* 28:8-12
- [51] Weinstein RS, Battacharayya AK, Graham AR, et al (1997) Telepathology a ten-year progress report. *Hum Pathol* 28:1-7
- [52] F. Petropoulos, K. Sidiropoulos, S. Kostopoulos, D. Glotsos, D. Cavouras, "Web-Based Telepathology Platform for Assisting Diagnosis in Histopathology", 4th IC-SCCE, 2010.
- [53] S. Kumar, B. E. Dunn, "Telepathology", 2009.
- [54] Weinberg DS, Allaert FA, Dusserre P, et al (1996) Telepathology diagnosis by means of digital still images: an international validation study. *Hum Pathol* 27:111-118
- [55] Weinstein RS (1996) Prospects for telepathology. *Hum Pathol* 17:433-434
- [56] Singson RPC, Natarajan S, Greenson JK, Marchevsky AM (1999) Virtual microscopy and the Internet as telepathology consultation tools. A study of gastrointestinal biopsy specimens. *Am J Pathol* 111:792-795
- [57] A. Afework, M. D. Beynon, F. Bustamante, S. Cho, A. Demarzo, R. Ferreira, R. Miller, M. Silberman, J. Saltz, A. Sussman, H. Tsang, "Digital Dynamic Telepathology – the Virtual Microscope", pp. 912-916, 1998 AMIA, Inc.

- [58] P. Spyridonos, D. Cavouras, P. Ravazoula, and G. Nikiforidis, "Neural network-based segmentat ion and classification system for automated grading of histologic sections of bladder carcinoma," *Analytical Quantitative Cytology and Histology*, vol. 24, pp. 317-24, 2002.
- [59] P. Bougioukos, "Functional classification of proteins using mass spectrometry data and exploration of their frequency of identification in proteomic analysis", PhD Thesis, 2009.
- [60] W. Coons, P. Jhonson, B. Sceithauer, A. Yates, and D. Pearl, "Improving diagnostic accuracy and interobserver concordance in the classification and grading of Primary Gliomas," *Cancer*, vol. 79, pp. 1381-93, 1997.
- [61] D. Glotsos, "Development of new machine learning methods for medical image processing and analysis", PhD Thesis, 2006.
- [62] D. N. Louis, H. Ohgaki, O. D. Wiestler, W. K. Cavenee, P. C. Burger, A. Jouvett, B. W. Scheithauer, P. Kleihues, "The 2007 WHO Classification of Tumours of the Central Nervous System", 2007.
- [63] J. C. Buckner, P. D. Brown, B. P. O'Neill, F. B. Meyer, C. J. Wetmore, J. H. Uhm, "Central Nervous System Tumors", 2007.
- [64] C. R. Miller, A. Perry, "Glioblastoma, Morphologic and Molecular Genetic Diversity", 2007.
- [65] D. Glotsos, I. Kalatzis, P. Spyridonos, S. Kostopoulos, A. Daskalakis, E. Athanasiadis, P. Ravazoula, G. Nikiforidis, D. Cavouras, "Improving accuracy in astrocytomas grading by integrating a robust least squares mapping driven support vector machine classifier into a two level grade classification scheme", 2008.
- [66] <http://www.mdguidelines.com/astrocytoma>
- [67] <http://www.cancer.gov>
- [68] <http://brain.mgh.harvard.edu/patientguide.htm>
- [69] www.astrocytoma.org
- [70] Γ. Κληρονόμος, "Μελέτη μοριακών μηχανισμών της διηθητικής ικανότητας των πρωτοπαθών όγκων του Κ.Ν.Σ.", PhD Thesis, 2009.
- [71] G. Reifenberger, V. P. Collins, "Pathology and molecular genetics of astrocytic gliomas", 2004.
- [72] P. Sallinen, "Quantitative Pathology in Astrocytic Tumors", pp. 1-24, 1999.
- [73] Γ. Δ. Κακκάβας, "Κατηγοριοποίηση καρκινικών ιστών προστάτη με υπολογισμό χαρακτηριστικών ενέργειας υψής και αλγορίθμους ταξινόμησης υπολογιστικής μάθησης", 2010.
- [74] Q. Wu, F. A. Merchant, K. R. Castleman, "Microscope Image Processing", 2008.
- [75] P. Spyridonos, E. I. Papageorgiou, P. P. Groumpos, G. N. Nikiforidis, "Integration of Expert Knowledge and image analysis Techniques for medical Diagnosis", pp. 110-121, 2006.
- [76] R. Prayson, D. Agamanolis, M. Cohen, M. Estes, B. Kleinschmidt-DeMasters, F. Abdul-Karim, S. McClure, B. Sebek, and R. Vinay, "Interobserver reproducibility among neuropathologists and surgical pathologists in fibrillary astrocytoma grading," *Journal of the Neurological Sciences*, vol. 175, pp. 33-39, 2000.
- [77] E. Maneas, C. Konstantinou, S. Kostopoulos, P. Ravazoula, D. Cavouras, D. Glotsos, "Prediction of Brain tumours grade using minimum spanning trees", 2011.
- [78] D. A. Clausi, Y. Zhao, "Grey level co-occurrence integrated algorithm (GLCIA): a superior computational method to rapidly determine co-occurrence probability texture features", 2003.
- [79] X. Tang, "Texture Information in Run-Length Matrices", 1998.