
ΑΣΚΗΣΗ

8

Ανάλυση και επεξεργασία εικόνων DICOM με τη χρήση Python

1. Σκοπός

Η ανάλυση και επεξεργασία ιατρικών εικόνων αποτελεί ένα ιδιαίτερος δημοφιλές πεδίο στο χώρο της βιοϊατρικής, με χιλιάδες άρθρα να δημοσιεύονται κάθε χρόνο¹. Η εξέλιξή του συμβάλλει στην ανάπτυξη εργαλείων που αποσκοπούν να βοηθήσουν επαγγελματίες υγείας στην κλινική ανάλυση, διάγνωση, και επέμβαση. Η αυξανόμενη χρήση απεικονιστικών τεχνικών και ηλεκτρονικών φακέλων υγείας, αλλά και η αλματώδης ανάπτυξη τεχνικών μηχανικής μάθησης για την ανάλυση εικόνων, έχουν συμβάλλει στην ολοένα και ταχύτερη ανάπτυξη του πεδίου, η οποία προβλέπεται να συνεχίσει με αμείωτους ρυθμούς και στο μέλλον. Στο πλαίσιο της άσκησης αυτής, αποσκοπείται η εξοικείωση των φοιτητών με το πεδίο της ανάλυσης και επεξεργασίας ιατρικών εικόνων, κάνοντας μια εισαγωγή στο πρότυπο DICOM που χρησιμοποιείται για την αποθήκευση των εικόνων αυτών, τα βασικά δομικά στοιχεία μιας ψηφιακής εικόνας, τις RGB διαστάσεις της και βασικές τεχνικές επεξεργασίας της.

2. Προσδοκώμενα Αποτελέσματα

Όταν θα έχετε ολοκληρώσει τη μελέτη της άσκησης θα γνωρίζετε:

- Τι είναι το πρότυπο DICOM το οποίο χρησιμοποιείται για την αποθήκευση ιατρικών εικόνων (πχ. Απεικονιστικών εξετάσεων όπως η αξονική τομογραφία).
- Τι είναι μια ψηφιακή εικόνα.
- Τι είναι οι διαστάσεις RGB.
- Βασικές τεχνικές ανάλυσης και επεξεργασίας ψηφιακών εικόνων (γραμμικό

¹ L. Maier-Hein *et al.*, ‘Why rankings of biomedical image analysis competitions should be interpreted with care’, *Nature Communications*, vol. 9, no. 1, p. 5217, Dec. 2018, doi: 10.1038/s41467-018-07619-7.

φιλτράρισμα, ανίχνευση ακμών, ανάλυση σχήματος και υφής).

- Τη τεχνικής μηχανικής μάθησης των k- κοντινότερων γειτόνων (k- Nearest Neighbor).

3. Περιγραφή του προτύπου DICOM

Η ψηφιακή επεξεργασία ιατρικής εικόνας ξεκίνησε παράλληλα με την ανάπτυξη ενός προτύπου για τη μεταφορά ψηφιακών εικόνων, ώστε να επιτρέψει στους χρήστες την ανάκτηση εικόνων και σχετιζομένων πληροφοριών από απεικονιστικά μηχανήματα με προτυποποιημένο τρόπο που θα είναι ο ίδιος για όλα τα μηχανήματα, ανεξαρτήτως κατασκευαστή. Το πρώτο αποτέλεσμα προς αυτήν την κατεύθυνση ήταν το πρότυπο που αφορούσε εικόνες ακτινολογίας από τον αμερικανικό οργανισμό ACR-NEMA (American College of Radiology – National Electrical Manufacturers' Association). Το πρότυπο αυτό αναφερόταν σε συνδέσεις από σημείο σε σημείο (point-to-point). Η ραγδαία ανάπτυξη της τεχνολογίας δικτύων περιόρισε τη χρησιμότητα αυτού του προτύπου. Σαν συνέπεια, το πρότυπο επανασχεδιάστηκε παίρνοντας υπόψη υπάρχοντα πρότυπα δικτύωσης. Το αποτέλεσμα ήταν το πρότυπο DICOM (Digital Imaging and Communications in Medicine), το οποίο αναπτύχθηκε με σκοπό την κάλυψη της ανάγκης διασύνδεσης διαφόρων ιατρικών απεικονιστικών και υπολογιστικών συστημάτων, είτε ανά δύο είτε σε δίκτυο, με σκοπό τη διαχείριση (καταχώρηση, διατήρηση και ανάκτηση) ιατρικών εικόνων και σχετιζόμενων πληροφοριών. Αρχικά το πρότυπο αναφερόταν σε ακτινολογικές εικόνες, λόγω όμως του γεγονότος ότι είναι εύκολα προσαρμόσιμο, γρήγορα άρχισε να χρησιμοποιείται και για ιατρικές εικόνες από άλλες απεικονιστικές συσκευές. Σήμερα, το DICOM είναι εξαιρετικά διαδεδομένο και οι περισσότεροι κατασκευαστές ιατρικών συσκευών το υποστηρίζουν. Στην Ευρώπη, ο Ευρωπαϊκός Οργανισμός Προτυποποίησης (CEN) χρησιμοποίησε το DICOM ως βάση για το πλήρως συμβατό πρότυπο MEDICOM. Το DICOM βρίσκεται αυτήν τη στιγμή στην έκδοση 3.0.

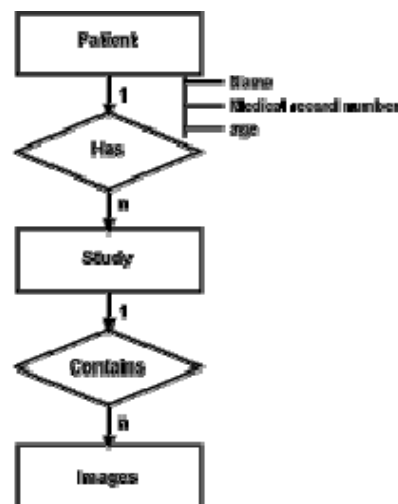
3.1 Ανταλλαγή στοιχείων με DICOM

Η ουσία του προτύπου DICOM είναι ότι ορίζει ένα σύνολο κοινών κανόνων για την ανταλλαγή και μεταφορά ψηφιακών εικόνων και των συνοδευτικών τους πληροφοριών. Το DICOM ακολουθεί το μοντέλο επικοινωνίας κατά ISO-OSI, πάνω στο οποίο στηρίζεται όλη η σύγχρονη τεχνολογία δικτύων. Το μοντέλο αυτό αποτελείται από επτά ανεξάρτητα επίπεδα ή στρώματα (layers) και καθορίζει τις λειτουργίες επικοινωνίας του κάθε επιπέδου και τις σχέσεις ανάμεσά τους. Αυτό γίνεται με ένα σύνολο κανόνων, τα πρωτόκολλα, τα οποία καθορίζουν τον τρόπο λειτουργίας των δικτύων για την επικοινωνία και ανταλλαγή πληροφοριών.

Όταν δύο συσκευές ή υπολογιστικά συστήματα θέλουν να επικοινωνήσουν και να ανταλλάξουν δεδομένα με χρήση του προτύπου ακολουθείται η παρακάτω διαδικασία: Αρχικά γίνεται προσπάθεια έναρξης επικοινωνίας πάνω από το δίκτυο. Το πρωτόκολλο δικτύου ενημερώνει για τη διαθεσιμότητα του δικτύου. Εάν το δίκτυο είναι διαθέσιμο, τότε το DICOM αρχίζει μια σειρά ενεργειών για να πραγματοποιηθεί η σύνδεση. Η συσκευή που αιτείται την επικοινωνία ενημερώνει για το είδος των ενεργειών που πρέπει να πραγματοποιηθούν και η συσκευή που λαμβάνει την αίτηση ενημερώνει με τη σειρά της για τις δυνατότητές της. Στην πραγματικότητα δεν είναι οι ίδιες οι συσκευές που υλοποιούν την παραπάνω διαδικασία, αλλά οι εφαρμογές που εκτελούνται σε αυτές και υλοποιούν το πρότυπο DICOM. Με την αρχική αυτή διαπραγμάτευση καθορίζονται οι δυνατότητες κάθε συσκευής και ο τρόπος με τον οποίο πρέπει να ανταλλαχθούν τα δεδομένα, ανάλογα με τις δυνατότητες της κάθε συσκευής και τα κοινά τους χαρακτηριστικά.

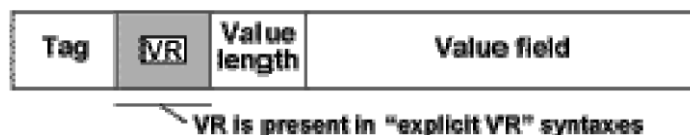
Πέραν των παραπάνω, το DICOM ορίζει και τη μορφή που πρέπει να έχει κάθε πληροφορία που ανταλλάσσεται. Στον τομέα αυτό χρησιμοποιήθηκαν και στοιχεία από άλλα πρότυπα. Για παράδειγμα, ο τρόπος καταγραφής του ονόματος του ασθενούς ακολουθεί τη μεθοδολογία που προκύπτει από το πρότυπο Health Level 7 (HL7). Έτσι, το όνομα του ασθενούς χωρίζεται σε «Όνομα», «Επίθετο» και «Πρόθεμα», όπου ως πρόθεμα εννοείται η προσφώνηση που μπορεί να χρησιμοποιείται (π.χ. Δόκτωρ ή κάτι παρόμοιο).

Το DICOM ορίζει επίσης μοντέλα οντοτήτων-συσχετίσεων (Entity-Relationship models) που δείχνουν τη σχέση μεταξύ των διαφόρων στοιχείων (οντοτήτων) που χρησιμοποιούνται στο πρότυπο. Για παράδειγμα, στην Εικόνα 1 απεικονίζεται ένα τέτοιο μοντέλο που δείχνει τη σχέση μεταξύ ενός ασθενούς, της εξέτασης (Study) στην οποία έχει υποβληθεί και των εικόνων που έχουν παραχθεί. Ο ασθενής, η εξέταση και οι εικόνες αποτελούν ορισμένες από τις οντότητες που χρησιμοποιούνται στο πρότυπο.



Εικόνα 1. Μοντέλο οντοτήτων-συσχετίσεων

Επιπλέον, κάθε οντότητα έχει και κάποια χαρακτηριστικά (attributes), τα οποία επίσης ορίζονται στο DICOM. Για παράδειγμα, ένας ασθενής περιγράφεται με χαρακτηριστικά όπως όνομα, φύλο, ηλικία, κωδικό ιατρικού φακέλου κ.ά. Στην πράξη, τα στοιχεία δεδομένων που ανταλλάσσονται στο DICOM, και τα οποία αναφέρθηκαν παραπάνω, είναι τα χαρακτηριστικά κάθε οντότητας. Η Εικόνα 2 δείχνει τη δομή ενός χαρακτηριστικού, όπως χρησιμοποιείται στο DICOM.



Εικόνα 2. Η δομή ενός χαρακτηριστικού (attribute) DICOM. Περιλαμβάνει την ονομασία (tag) του χαρακτηριστικού, το μήκος του και την τιμή του

Όπως φαίνεται από την παραπάνω περιγραφή, το DICOM ακολουθεί τις αρχές της αντικειμενοστραφούς σχεδίασης (object oriented design). Κάθε οντότητα (ο ασθενής, η εικόνα, η εξέταση κ.λπ.) είναι ένα αντικείμενο με κάποια χαρακτηριστικά. Τα αντικείμενα στο DICOM ονομάζονται αντικείμενα πληροφορίας (information objects). Η περιγραφή για κάθε αντικείμενο πληροφορίας στο DICOM καθορίζει ποια από τα χαρακτηριστικά είναι υποχρεωτικά, ποια προαιρετικά και ποια χρησιμοποιούνται κατά περίπτωση.

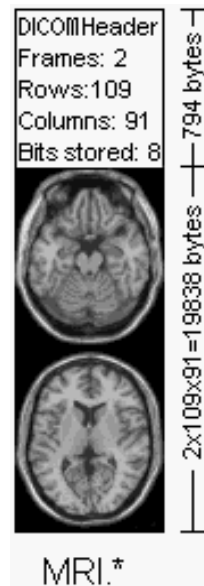
3.2 Υπηρεσίες DICOM

Τα αντικείμενα πληροφορίας του DICOM χρησιμοποιούνται για τη μεταφορά των εικόνων και των σχετιζόμενων πληροφοριών μεταξύ συσκευών. Αυτή η διακίνηση πληροφορίας δεν είναι όμως το μόνο που πρέπει να γίνει. Πέραν της ανταλλαγής δεδομένων πρέπει να πραγματοποιηθούν και κάποιες ενέργειες με αυτά τα δεδομένα. Για παράδειγμα, όταν στέλνεται μια εικόνα σε έναν εκτυπωτή, αυτός θα πρέπει να την τυπώσει, ή, εάν αποστέλλεται σε ένα σύστημα αποθήκευσης, τότε θα πρέπει να αποθηκευτεί. Για το σκοπό αυτό το DICOM ορίζει υπηρεσίες (services) οι οποίες χρησιμοποιούνται με τα αντικείμενα πληροφορίας. Υπάρχει ένας αριθμός βασικών υπηρεσιών με βάση τις οποίες προκύπτουν οι πιο περίπλοκες. Οι βασικές αυτές υπηρεσίες ονομάζονται message service elements (DIMSE) και χωρίζονται σε δύο κατηγορίες: στις λειτουργίες (operations), όπως «αποθήκευση» (store) η οποία αποθηκεύει τα δεδομένα, και στις ανακοινώσεις (notifications), όπως «αναφορά συμβάντος» (event report) που γνωστοποιεί ότι κάτι συνέβη. Αυτές οι βασικές υπηρεσίες χρησιμοποιούνται, όπως προαναφέρθηκε, για την υλοποίηση πιο πολύπλοκων υπηρεσιών. Για παράδειγμα, η υπηρεσία «αναζήτηση- ανάκτηση» (query-retrieve) εκτελεί μια αναζήτηση σε μια συσκευή αποθήκευσης και επιστρέφει τα αποτελέσματα. Η υπηρεσία αυτή προκύπτει με χρήση των βασικών υπηρεσιών «βρες» (find), «πάρε» (get) και «μετακίνηση» (move).

Κάθε συσκευή μπορεί να παρέχει ή να χρησιμοποιεί μια ή περισσότερες υπηρεσίες. Ο ρόλος κάθε συσκευής πρέπει να διευκρινιστεί πριν από κάθε επικοινωνία.

3.3 Η μορφή ενός αρχείου DICOM

Ένα αρχείο DICOM αποτελείται από μια επικεφαλίδα (header), η οποία περιέχει συνοδευτικές πληροφορίες (π.χ. τα στοιχεία του ασθενούς) και ακολουθείται από τα δεδομένα της εικόνας. Στην Εικόνα 3 φαίνεται αυτή η δομή.



Εικόνα 3. Η δομή ενός αρχείου DICOM

Όπως φαίνεται στην εικόνα, στο υποθετικό αυτό αρχείο η επικεφαλίδα έχει μήκος 794 bytes και περιέχει πληροφορίες σε μορφή κειμένου που αφορούν διάφορες οντότητες και τα χαρακτηριστικά τους, όπως οι διαστάσεις και ο τύπος της αποθηκευμένης εικόνας, πληροφορίες σχετικά με τον ασθενή και την εξέταση κ.ά. Το μέγεθος της επικεφαλίδας μπορεί να αλλάζει ανάλογα με τον όγκο της πληροφορίας που αποθηκεύεται. Παρακάτω φαίνονται ορισμένα από τα περιεχόμενα μιας επικεφαλίδας ενός αρχείου DICOM, όπως παρουσιάζονται από ένα πρόγραμμα ανάγνωσης αρχείων DICOM (στο Παράρτημα δίνεται ένα παράδειγμα στο οποίο παρουσιάζονται όλα τα δεδομένα που περιέχει μια επικεφαλίδα DICOM).

0008,0020, Study Date: 25-01-1999
 0010,0010, Patient's Name: Jack Kerok
 0010,0040, Patient Sex: Male
 0020,0013, Image Number: 1
 0028,0010, Rows: 576
 0028,0011, Columns: 1012

Κάθε χαρακτηριστικό που περιλαμβάνεται στην επικεφαλίδα χαρακτηρίζεται από δύο δεκαεξαδικούς αριθμούς. Ο πρώτος καθορίζει την οντότητα ή αντικείμενο στο οποίο ανήκει το χαρακτηριστικό, ενώ ο δεύτερος είναι ο κωδικός αναγνώρισης του συγκεκριμένου χαρακτηριστικού. Για παράδειγμα, το όνομα του ασθενούς έχει τους κωδικούς 0010, 0010. Ο πρώτος λέει ότι το χαρακτηριστικό αυτό αναφέρεται στην οντότητα «ασθενής», ενώ ο δεύτερος ότι πρόκειται για το χαρακτηριστικό «όνομα ασθενούς». Στο συγκεκριμένο παράδειγμα η τιμή του χαρακτηριστικού αυτού είναι «Jack Kerok». Οι κωδικοί 0010, 0040 αναφέρονται στο φύλο του ασθενούς. Ο πρώτος κωδικός είναι πάλι ο 0010 καθώς, όπως και το όνομα, έτσι και το φύλο χαρακτηρίζει την οντότητα «ασθενής». Ο δεύτερος κωδικός (0040) αναφέρεται στο χαρακτηριστικό «φύλο ασθενούς». Με παρόμοιο τρόπο ορίζονται όλα τα αντικείμενα πληροφορίας και τα χαρακτηριστικά τους, που μπορούμε να αποθηκεύσουμε στην επικεφαλίδα ενός αρχείου DICOM.

Μετά την επικεφαλίδα ακολουθούν τα δυαδικά δεδομένα της ιατρικής εικόνας. Η ιατρική εικόνα μπορεί να αποθηκεύεται ασυμπίεστη (raw μορφή) ή να είναι συμπίεσμένη με χρήση διαφόρων αλγορίθμων, όπως π.χ να αποθηκεύεται σε JPEG format.

4. Βασικά στοιχεία ψηφιακών εικόνων

Για να επεξεργαστεί ένας ηλεκτρονικός υπολογιστής μια εικόνα θα πρέπει αυτή να είναι ψηφιοποιημένη, δηλαδή να αποτελείται από ένα σύνολο διακριτών στοιχείων. Τα στοιχεία αυτά ονομάζονται εικονοστοιχεία (pixels). Ουσιαστικά μια ψηφιακή εικόνα είναι ένας πίνακας, κάθε στοιχείο του οποίου αντιστοιχεί σε ένα pixel της εικόνας. Η τιμή κάθε στοιχείου του πίνακα αυτού αντιπροσωπεύει το «χρώμα» του αντίστοιχου pixel. Για μια grayscale εικόνα, η τιμή του pixel αντιπροσωπεύει τη διαβάθμιση του γκρι. Ο πιο διαδεδομένος τύπος pixel είναι το byte, το οποίο είναι ένας αριθμός αποθηκευμένος ως 8-bit ακέραιος με τιμές από 0-255. Το 0 αντιπροσωπεύει το μαύρο και το 255 το λευκό. Για την αναπαράσταση έγχρωμων εικόνων (RGB) το κάθε pixel αντιπροσωπεύεται από 3 συνισταμένες, την ένταση του κόκκινου (Red), του πράσινου (Green) και του μπλε (Blue) χρώματος. Η τιμή του κάθε pixel επομένως στην περίπτωση αυτή είναι ένα διάνυσμα των τιμών των 3 χρωμάτων.

4.1 Οι τύποι ψηφιακών εικόνων και η δομή τους στη Python

Υπάρχουν 3 βασικοί τύποι εικόνων που μπορεί κανείς να αναγνωρίσει:

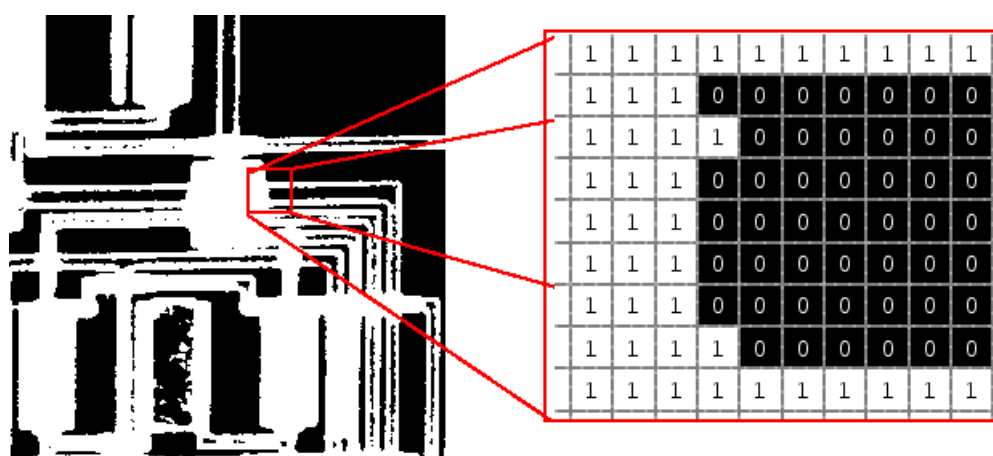
- Δυαδικές εικόνες
- Ασπρόμαυρες εικόνες
- Έγχρωμες εικόνες (RGB)

Δυαδικές (binary) εικόνες

Σε μια δυαδική εικόνα, κάθε pixel μπορεί να πάρει μια από τις δύο διακριτές τιμές (0 ή 1). Η τιμή 0 αντιστοιχεί στο μαύρο και η τιμή 1 στο λευκό. Μια δυαδική εικόνα αποθηκεύεται ως ένας δισδιάστατος πίνακας μηδενικών και μονάδων.

Μια δυαδική εικόνα μπορεί να θεωρηθεί ως ένας ειδικός τύπος ασπρόμαυρης εικόνας, η οποία όμως περιέχει μόνο μαύρο και άσπρο. Συνήθως τις δυαδικές εικόνες τις χρησιμοποιούμε για δεικτοδότηση (indexing) περιοχών ενδιαφέροντος άλλων εικόνων (τα pixels με τιμή 1 υποδεικνύουν τα pixels μιας περιοχής ενδιαφέροντος). Για το λόγο αυτό, μια δυαδική εικόνα ονομάζεται επίσης και «μάσκα».

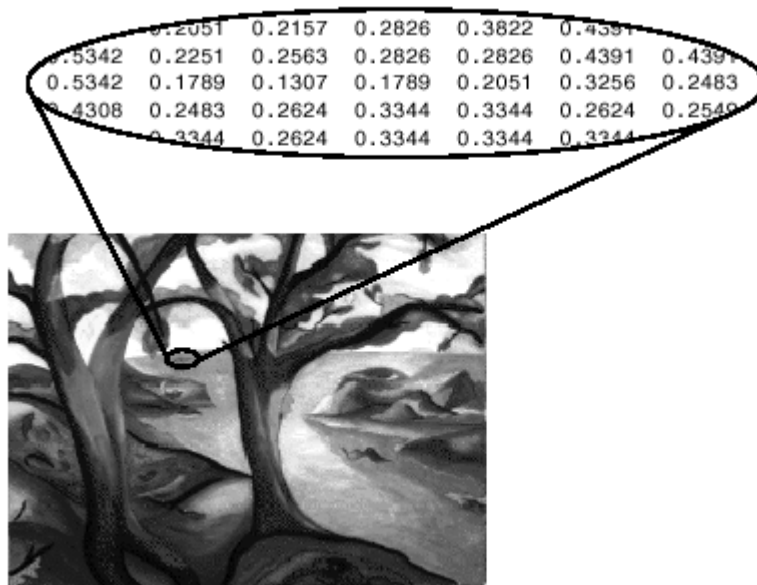
Παρακάτω βλέπουμε ένα παράδειγμα δυαδικής εικόνας.



Εικόνα 4. Δυαδική (binary) εικόνα

Ασπρόμαυρες (grayscale) εικόνες

Μια ασπρόμαυρη εικόνα (εικόνα έντασης φωτεινότητας) είναι ένας πίνακας δεδομένων I του οποίου οι τιμές των στοιχείων αναπαριστούν την ένταση της φωτεινότητας κάθε pixel. Η Python (με τη χρήση Numpy) αποθηκεύει μια εικόνα έντασης σε ένα διδιάστατο πίνακα, του οποίου κάθε στοιχείο αναφέρεται σε ένα και μοναδικό pixel. Οι τιμές των pixels αναπαριστούν διάφορες εντάσεις φωτεινότητας (επίπεδα του γκρι), όπου εάν ο τύπος της εικόνας είναι uint8 (δηλαδή 8-bit unsigned integer) η τιμή 0 αναπαριστά το μαύρο και η τιμή 255 αναπαριστά το λευκό. Το διάστημα $[0, 255]$ μέσα στο οποίο μπορεί να πάρει τιμή φωτεινότητας ένα pixel, πολλές φορές κανονικοποιείται σε $[0, 1]$, όπως φαίνεται και στην παρακάτω εικόνα.

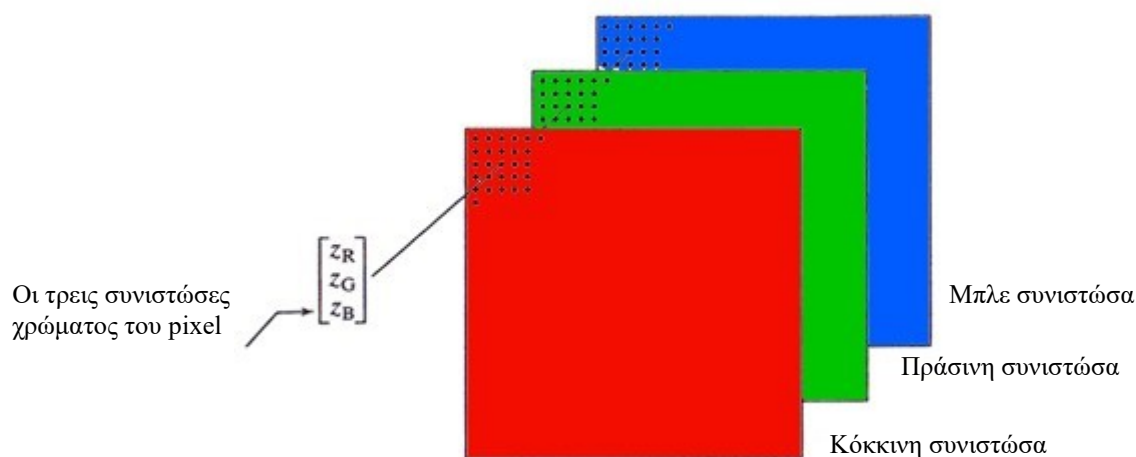


Εικόνα 5. Ασπρόμαυρη (grayscale) εικόνα με τις τιμές φωτεινότητας κανονικοποιημένες στο διάστημα $[0, 1]$

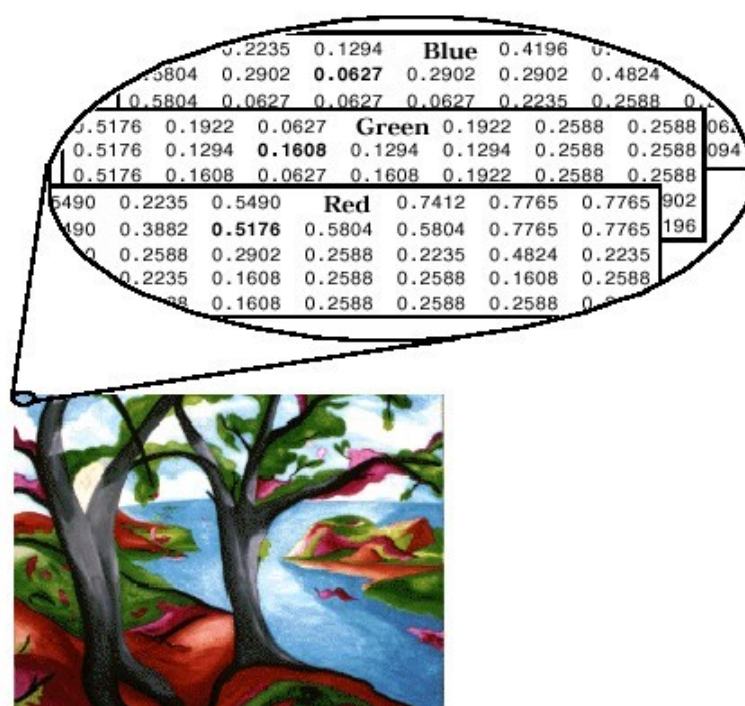
Έγχρωμες εικόνες (RGB)

Μια RGB εικόνα αποθηκεύεται ως ένας πίνακας δεδομένων $M \times N \times 3$, ο οποίος καθορίζει τα κόκκινα, πράσινα και μπλε χρωματικά στοιχεία κάθε ανεξάρτητου pixel. Κάθε pixel αποτελείται από 3 συνιστώσες (κόκκινη, πράσινη, μπλε), και η τιμή κάθε μιας από αυτές περιγράφει την «ποσότητα» (ένταση) του κόκκινου, πράσινου και μπλε αντίστοιχα που χρησιμοποιείται για τη δημιουργία του τελικού χρώματος. Πρακτικά, μια εικόνα RGB δημιουργείται από τη σύνθεση τριών grayscale εικόνων, κάθε μια από τις οποίες αναπαριστά την ένταση της φωτεινότητας των pixels για το κόκκινο (Red εικόνα), το πράσινο (Green εικόνα) και το μπλε (Blue εικόνα) χρώμα. Ένα pixel το οποίο έχει τις τιμές (0,0,0) παρουσιάζει το μαύρο, ένα pixel το οποίο έχει τιμές (1,1,1) παρουσιάζει το λευκό, ενώ το pixel με τιμές (1,0,0) παρουσιάζει το κόκκινο χρώμα.

Στην Εικόνα 6 παρουσιάζεται μια σχηματική επεξήγηση των συνιστωσών του pixel για μια RGB εικόνα, ενώ στην Εικόνα 7 φαίνονται οι πίνακες Red, Green και Blue μιας αποθηκευμένης εικόνας (κανονικοποιημένες τιμές στο διάστημα $[0, 1]$).



Εικόνα 6: Σχηματική αναπαράσταση των συνιστωσών εικόνας RGB



Εικόνα 7. Οι 3 πίνακες Red, Green και Blue που συνθέτουν μια έγχρωμη εικόνα RGB

4.2 Ιστόγραμμα ψηφιακής εικόνας

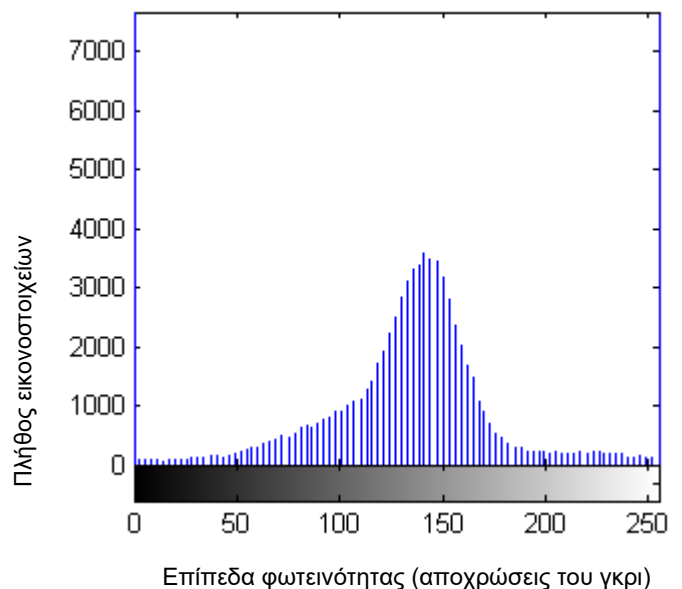
Το ιστόγραμμα (histogram) μίας εικόνας αποχρώσεων του γκρι (grayscale) περιέχει σημαντικές πληροφορίες για την εικόνα και για το λόγο αυτό είναι ένα από τα σημαντικότερα εργαλεία στην επεξεργασία ψηφιακών εικόνων. Μπορεί να χρησιμοποιηθεί για τη βελτιστοποίηση της εικόνας, την τμηματοποίησή της σε επιμέρους περιοχές (segmentation), την τροποποίησή της, την εξαγωγή χαρακτηριστικών κ.ά.

Το ιστόγραμμα μιας ψηφιακής εικόνας είναι ένα γράφημα στο οποίο ο οριζόντιος άξονας αναπαριστά τα επίπεδα φωτεινότητας από 0-255 και ο κατακόρυφος το πλήθος των pixels που έχουν κάθε επίπεδο φωτεινότητας. Είναι ένας τρόπος αναπαράστασης της κατανομής των διαφορετικών επιπέδων του γκρι μιας grayscale εικόνας. Το γράφημα αυτό παρέχει γενικές πληροφορίες για τη μορφή της εικόνας και το περιεχόμενό της, όπως το δυναμικό εύρος και την αντίθεση της εικόνας (σκοτεινή, φωτεινή, υψηλής ή χαμηλής αντίθεσης). Ωστόσο, το ιστόγραμμα δεν δίνει πληροφορίες για τη θέση των pixels μέσα στην εικόνα.

Ανάλογα με την εφαρμογή, ο κατακόρυφος άξονας μπορεί να κανονικοποιηθεί με βάση τη μέγιστη τιμή του ιστογράμματος. Στην περίπτωση που η μέγιστη τιμή του ιστογράμματος αντιστοιχηθεί στο 1, τότε το κανονικοποιημένο ιστόγραμμα που προκύπτει δίνει ουσιαστικά την κατανομή πυκνότητας πιθανότητας των επιπέδων του γκρι στην εικόνα.

Για μια εικόνα RGB έχουμε 3 ιστογράμματα, ένα για κάθε χρωματική συνιστώσα. Έτσι, μια έγχρωμη εικόνα περιγράφεται από το ιστόγραμμα R, το ιστόγραμμα G, και το ιστόγραμμα B, κάθε ένα από τα οποία εκφράζει την κατανομή των επιπέδων φωτεινότητας του κόκκινου, του πράσινου και του μπλε χρώματος της εικόνας αντίστοιχα.

Στην παρακάτω εικόνα παρουσιάζεται μια εικόνα Υπολογιστικής Τομογραφίας (CT) εγκεφάλου (αριστερά) και το ιστόγραμμά της (δεξιά). Κάθε κάθετη γραμμή του ιστογράμματος αντιστοιχεί στο πλήθος των pixels στην εικόνα που έχουν μια συγκεκριμένη τιμή φωτεινότητας.



Εικόνα 8. Εγκάρσια εγκεφαλική τομή Υπολογιστικής Τομογραφίας (CT) (αριστερά) και το ιστόγραμμα της εικόνας CT (δεξιά)

5. Τεχνικές επεξεργασίας ψηφιακών εικόνων

Οι τεχνικές επεξεργασίας ψηφιακών εικόνων αποσκοπούν στην επεξεργασία μιας εικόνας (ή ενός πλήθους εικόνων, π.χ. βίντεο) με χρήση αλγορίθμων ή μετασχηματισμών για την επίτευξη στόχων όπως:

- Βελτίωση της ποιότητας της εικόνας (αύξηση αντίθεσης, αποθορυβοποίηση κ.ά.)
- Ανίχνευση διαφόρων χαρακτηριστικών της (ανίχνευση ακμών, κ.ά.)
- Τμηματοποίηση της εικόνας – Εξαγωγή και διαχωρισμός των αντικειμένων που υπάρχουν σε αυτή
- Συμπίεση της εικόνας
- Μετασχηματισμός της γεωμετρίας της εικόνας ή των αντικειμένων της (περιστροφή, κ.ά.)
- Μετασχηματισμός σε άλλα πεδία, π.χ. στο πεδίο των συχνοτήτων, για αποκάλυψη πληροφορίας που δεν είναι εμφανής στο αρχικό πεδίο

Οι τεχνικές αυτές έχουν ως είσοδο μια εικόνα (εικόνα εισόδου – αρχική εικόνα) και μέσω ενός μετασχηματισμού ή αλγορίθμου παράγουν μια άλλη εικόνα (εικόνα εξόδου), όπως φαίνεται και στο παρακάτω σχήμα. Κατά την ψηφιακή επεξεργασία μιας εικόνας, στόχος μας είναι να δημιουργήσουμε μια εικόνα εξόδου που θα είναι πιο κατάλληλη από την αρχική για μια συγκεκριμένη εφαρμογή. Η ψηφιακή επεξεργασία μιας ιατρικής εικόνας εξυπηρετεί στην ανάδειξη πληροφορίας η οποία είναι κλινικά χρήσιμη και ίσως αθέατη διά γυμνού οφθαλμού.

Ανάλογα με την εφαρμογή, η εικόνα εξόδου μπορεί να είναι ίδιων διαστάσεων με την αρχική (π.χ. σε εφαρμογές βελτίωσης της ποιότητας, ανίχνευσης ακμών κ.λπ.) ή διαφορετικών (π.χ. αλλαγή μεγέθους της εικόνας, εξαγωγή μιας περιοχής κ.λπ.).



Εικόνα 9. Τυπική διαδικασία εφαρμογής μιας τεχνικής ψηφιακής επεξεργασίας εικόνας

Η ψηφιακή ανάλυση μιας εικόνας αφορά την περιγραφή και αναγνώριση του περιεχομένου της. Οι τεχνικές ψηφιακής ανάλυσης έχουν ως είσοδο είτε την αρχική εικόνα είτε εικόνα που έχει υποστεί κατάλληλη επεξεργασία, και επιστρέφουν πληροφορίες (δεδομένα) σχετικά με την εικόνα ή τα αντικείμενα που περιέχει αυτή. Περισσότερα σχετικά με την ψηφιακή ανάλυση θα παρουσιαστούν στην επόμενη ενότητα.

Η ψηφιακή επεξεργασία εικόνων είναι μια σύνθετη διαδικασία, η οποία περιλαμβάνει στη γενικότερη περίπτωση πολλά και διαφορετικά στάδια ανάλογα με τη μορφή και την αρχική κατάσταση της ψηφιακής εικόνας, καθώς και το επιθυμητό τελικό αποτέλεσμα. Έτσι, ανάλογα με την εφαρμογή, μπορεί να χρειαστεί να εφαρμόσουμε σταδιακά διάφορες τεχνικές επεξεργασίας και ανάλυσης προκειμένου να καταλήξουμε στο τελικό αποτέλεσμα, όπως θα δούμε και στη συνέχεια.

Οι βασικότερες τεχνικές επεξεργασίας μιας ψηφιακής εικόνας μπορούν να χωριστούν σε δύο κατηγορίες:

- Τεχνικές επεξεργασίας στο πεδίο του χώρου (spatial domain)
- Τεχνικές επεξεργασίας στο πεδίο της συχνότητας (frequency domain)

Η επεξεργασία στο πεδίο της συχνότητας συνίσταται στην ανάλυση και την επεξεργασία του φασματικού περιεχομένου μιας εικόνας και στοχεύει στην ανάλυση και τη βελτίωσή της λαμβάνοντας υπόψη την κατανομή των συχνοτήτων της. Στις ψηφιακές εικόνες, οι συχνότητες αντιπροσωπεύουν την ταχύτητα και την κατεύθυνση μεταβολής της φωτεινότητας ή του χρώματος. Στα πλαίσια της παρούσας άσκησης δεν θα ασχοληθούμε με τεχνικές της κατηγορίας αυτής.

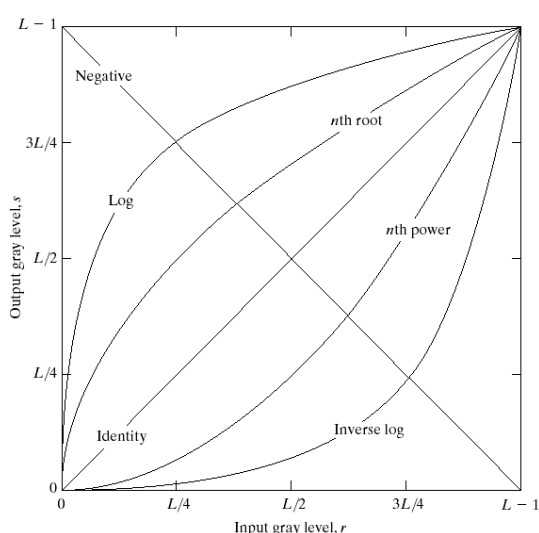
Οι τεχνικές επεξεργασίας στο πεδίο του χώρου αφορούν τεχνικές που εφαρμόζονται απευθείας πάνω στα pixels μιας εικόνας. Οι τεχνικές αυτές μπορούν να χωριστούν στις εξής κατηγορίες:

- Σημειακή επεξεργασία (Point Processing): Η τιμή κάθε pixel στην εικόνα εξόδου (pixel εξόδου) υπολογίζεται χρησιμοποιώντας την τιμή του pixel στην αντίστοιχη θέση της εικόνας εισόδου (pixel εισόδου). Δηλαδή, κάθε pixel εξόδου εξαρτάται από την τιμή φωτεινότητας του αντίστοιχου pixel εισόδου και του εκάστοτε μετασχηματισμού που εφαρμόζεται. Οι τεχνικές βελτίωσης εικόνας με χρήση μετασχηματισμών φωτεινότητας που παρουσιάζονται στη συνέχεια, αφορούν τεχνικές σημειακής επεξεργασίας.
- Τοπική επεξεργασία (Local Processing): Η τιμή κάθε pixel στην εικόνα εξόδου υπολογίζεται χρησιμοποιώντας την τιμή του pixel στην αντίστοιχη θέση της εικόνας εισόδου, καθώς επίσης και τις τιμές των γειτονικών pixels. Επομένως, στις τεχνικές αυτές, η τιμή ενός pixel εξόδου δεν εξαρτάται μόνο από την τιμή του αντίστοιχου pixel εισόδου και του εφαρμοζόμενου μετασχηματισμού, αλλά και από τις τιμές των γειτονικών pixels γύρω από το pixel εισόδου. Χαρακτηριστικά παραδείγματα τεχνικών τοπικής επεξεργασίας αποτελούν τα χωρικά φίλτρα τα οποία θα παρουσιαστούν στη συνέχεια. Τέτοια φίλτρα χρησιμοποιούνται για διάφορες εφαρμογές, όπως εξάλειψη του θορύβου, ανίχνευση ακμών κ.ά.
- Καθολική επεξεργασία (Global Processing): Για να υπολογιστεί η τιμή ενός pixel της εικόνας εξόδου χρησιμοποιούνται όλα τα pixels της εικόνας εισόδου. Χαρακτηριστικό παράδειγμα καθολικής επεξεργασίας αποτελεί η διαδικασία ολικής εξισορρόπησης ιστογράμματος (global histogram equalization). Η τεχνική αυτή μετασχηματίζει τα επίπεδα φωτεινότητας μιας εικόνας έτσι ώστε αυτά να κατανέμονται ομοιόμορφα σε όλη την κλίμακα φωτεινοτήτων (δηλαδή το ιστόγραμμα της αρχικής εικόνας να μετατραπεί σε ομοιόμορφο). Επομένως, η τιμή φωτεινότητας κάθε pixel εξόδου εξαρτάται τόσο από την τιμή που έχει το αντίστοιχο pixel εισόδου όσο και από τις τιμές φωτεινότητας όλων των υπολοίπων pixels της εικόνας εισόδου. Η εικόνα εξόδου που προκύπτει από την διαδικασία εξισορρόπησης ιστογράμματος είναι αυξημένης αντίθεσης σε σχέση με την αρχική.
- Γεωμετρικοί μετασχηματισμοί: Στην πιο απλή περίπτωση, τα pixels της εικόνας εξόδου υπολογίζονται από το μετασχηματισμό των συντεταγμένων των pixels της εικόνας εισόδου. Χαρακτηριστικό παράδειγμα αποτελεί η περιστροφή μιας εικόνας κατά 90° .

5.1 Βελτίωση εικόνας με χρήση μετασχηματισμών φωτεινότητας

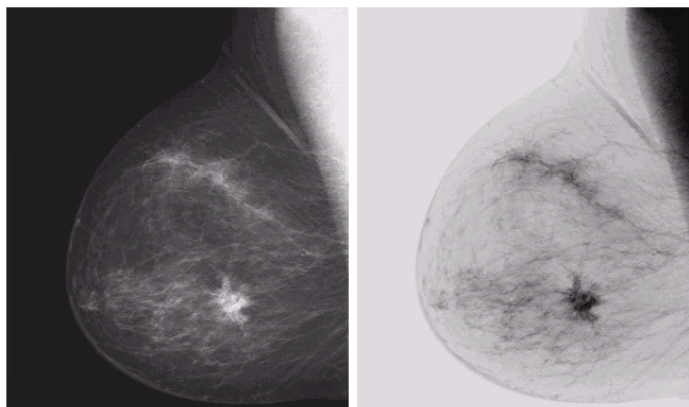
Σκοπός των τεχνικών βελτίωσης εικόνας με χρήση μετασχηματισμών φωτεινότητας είναι η κατάλληλη ρύθμιση των επιπέδων φωτεινότητας προκειμένου να προκύψει μια βελτιωμένη εικόνα, στην οποία περιοχές φωτεινότητας τονίζονται ή εξομαλύνονται.

Οι τεχνικές αυτές είναι τεχνικές σημειακής επεξεργασίας, έτσι κάθε pixel της εικόνας εξόδου υπολογίζεται από ένα γραμμικό ή μη γραμμικό μετασχηματισμό της τιμής του αντίστοιχου pixel της εικόνας εισόδου. Χρησιμοποιώντας κατάλληλες συναρτήσεις μετασχηματισμού μπορούμε να πετύχουμε ένα πλήθος μετατροπών που βρίσκουν εφαρμογή σε διάφορες περιπτώσεις. Ανάλογα με το πρόβλημα, έχουν προταθεί διάφορες συναρτήσεις μετασχηματισμού, γραμμικές (όπως π.χ. στην περίπτωση του αρνητικού μίας εικόνας) είτε μη γραμμικές (λογαριθμικές, εκθετικές κ.λπ). Ορισμένες βασικές συναρτήσεις μετασχηματισμού παρουσιάζονται στην παρακάτω εικόνα.



Εικόνα 10. Βασικές συναρτήσεις μετασχηματισμού φωτεινότητας που χρησιμοποιούνται για τη βελτίωση της εικόνας. Για τυπικές grayscale εικόνες είναι $L=256$.

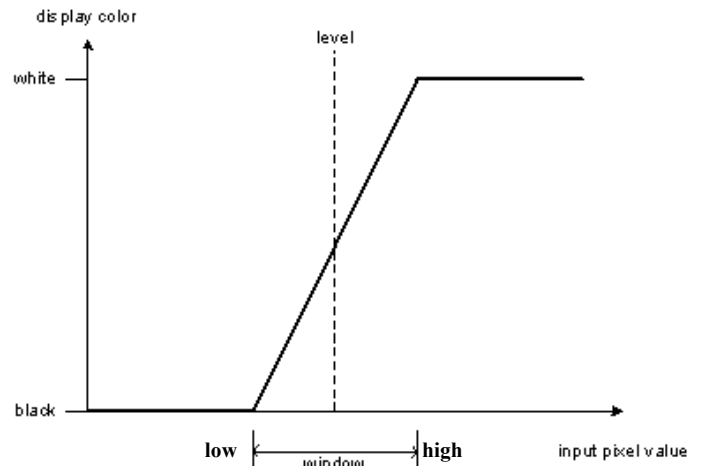
Το πιο απλό παράδειγμα αυτής της κατηγορίας τεχνικών αφορά το αρνητικό μιας εικόνας, όπου η τιμή φωτεινότητας κάθε pixel της εικόνας εξόδου προκύπτει από τη συνάρτηση μετασχηματισμού Negative: $T(g) = 255 - g$ (όπου $T(g)$ η τιμή του pixel στην εικόνα εξόδου, g η τιμή του pixel στην αντίστοιχη θέση της εικόνας εισόδου). Το αρνητικό μιας εικόνας είναι χρήσιμο σε πολλές εφαρμογές ιατρικών εικόνων, όπως για παράδειγμα στην ψηφιακή μαστογραφία (Εικόνα 11).



Εικόνα 11. Εικόνα ψηφιακής μαστογραφίας (αριστερά) και το αρνητικό της (δεξιά)

Μια άλλη εφαρμογή των τεχνικών αυτών είναι η βελτίωση της αντίθεσης μιας εικόνας (Contrast Enhancement), διαδικασία η οποία οδηγεί στην αύξηση της ορατότητας μιας περιοχής ή μιας δομής της εικόνας, έτσι ώστε να είναι εγκυρότερες οι εξαγόμενες πληροφορίες και ευκολότερη η εφαρμογή άλλων τεχνικών. Ανάλογα με το πρόβλημα, για τη βελτίωση της αντίθεσης μιας εικόνας χρησιμοποιούνται κατάλληλες συναρτήσεις μετασχηματισμού ώστε να αυξάνεται (ή να μειώνεται) η αντίθεση (ή η φωτεινότητα) σε προκαθορισμένες περιοχές φωτεινότητας. Τυπικές συναρτήσεις μετασχηματισμού που χρησιμοποιούνται για τη βελτίωση της αντίθεσης ή της φωτεινότητας μιας εικόνας είναι η λογαριθμική συνάρτηση, η εκθετική συνάρτηση και οι συναρτήσεις δύναμης (Εικόνα 10).

Μια πιο απλή περίπτωση είναι η γραμμική συνάρτηση που ορίζεται σε μια υποπεριοχή (παράθυρο) του εύρους φωτεινότητας, δηλαδή μεταξύ κάποιων ορίων φωτεινότητας [low, high], όπως απεικονίζεται στο διπλανό σχήμα. Στην περίπτωση αυτή, τα pixels εξόδου που αντιστοιχούν σε pixels εισόδου που έχουν τιμές μικρότερες της τιμής low, λαμβάνουν τιμή 0, ενώ τα pixels εξόδου που αντιστοιχούν σε pixels εισόδου που έχουν τιμές μεγαλύτερες της τιμής high, λαμβάνουν τιμή 255. Με τον τρόπο αυτό πετυχαίνουμε προσαρμογή της αντίθεσης σε ένα εύρος φωτεινότητας.



Στην Εικόνα 12 παρουσιάζεται το αποτέλεσμα της εφαρμογής προσαρμογής της αντίθεσης σε εύρος φωτεινότητας με κάτω όριο low=77 και πάνω όριο high=180, στην Εικόνα 8. Παρατηρώντας την παρακάτω εικόνα και την αρχική (Εικόνα 8), διαπιστώνουμε πως με την εφαρμογή της συγκεκριμένη διαδικασίας αναδεικνύονται και τονίζονται λεπτομέρειες από το εσωτερικό του εγκεφάλου.

Η βελτίωση της αντίθεσης μιας εικόνας μπορεί να πραγματοποιηθεί και με άλλες τεχνικές πέραν των προαναφερθέντων. Η τεχνική εξισορρόπησης ιστογράμματος, που αναφέρθηκε στις τεχνικές καθολικής επεξεργασίας, αποτελεί μια από τις πιο χρήσιμες τεχνικές για τη βελτίωση της αντίθεσης εικόνας.



Εικόνα 12. Αύξηση της αντίθεσης της εικόνας 8 προκειμένου να αναδειχθούν λεπτομέρειες από την περιοχή του εγκεφάλου

5.2 Χωρικά Φίλτρα (Spatial Filters)

Το χωρικό φιλτράρισμα μιας εικόνας αποτελεί τεχνική τοπικής επεξεργασίας. Τα χωρικά φίλτρα χρησιμοποιούνται για διάφορες εφαρμογές, όπως εξάλειψη του θορύβου, ανίχνευση ακμών κ.ά., κάποιες από τις οποίες θα παρουσιαστούν στη συνέχεια.

Για την εφαρμογή των τεχνικών τοπικής επεξεργασίας χρησιμοποιείται συνήθως μια διαδικασία σάρωσης όλης της εικόνας με χρήση κυλιόμενων παραθύρων.

3.2.1 Λειτουργίες με χρήση κυλιόμενων παραθύρων

Η διαδικασία κυλιόμενων παραθύρων-γειτονιών (Sliding Neighborhood Operation) χρησιμοποιείται ευρέως στην επεξεργασία ψηφιακής εικόνας για την εφαρμογή διαφόρων τεχνικών, όπως:

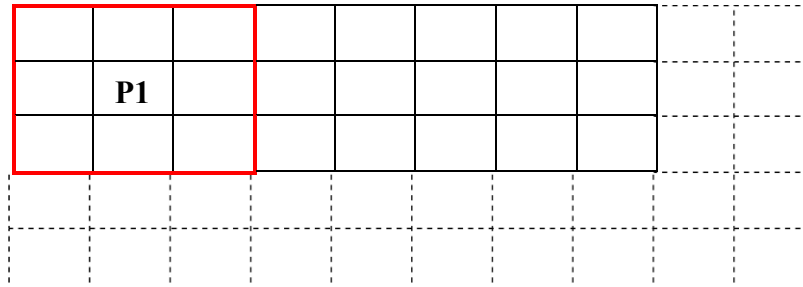
- εφαρμογή χωρικών φίλτρων
- επεξεργασία εικόνας με αλγορίθμους που βασίζονται σε τοπικές τεχνικές
- εξαγωγή τοπικών χαρακτηριστικών
- εντοπισμός περιοχών ενδιαφέροντος (περιοχών που παρουσιάζουν συγκεκριμένα χαρακτηριστικά)

Κατά τη διαδικασία αυτή, η τιμή ενός pixel της εικόνας εξόδου καθορίζεται από την εφαρμογή ενός αλγορίθμου στη γειτονιά του αντίστοιχου pixel της εικόνας εισόδου.

Έστω ότι στην Εικόνα 13 απεικονίζεται μια ψηφιακή εικόνα μεγέθους $M \times N$ όπου κάθε στοιχείο του πίνακα αντιστοιχεί σε ένα pixel της. Κάθε pixel μιας εικόνας γειτονεύει άμεσα με οκτώ άλλα εικονοστοιχεία, 4 με τα οποία έχει κοινά όρια και 4 με τα οποία έχει κοινές γωνίες. Στην Εικόνα 13, τα 8 pixels που βρίσκονται γύρω από το pixel P1 (το οποίο είναι το κεντρικό της γειτονιάς) αποτελούν την άμεση γειτονιά του.

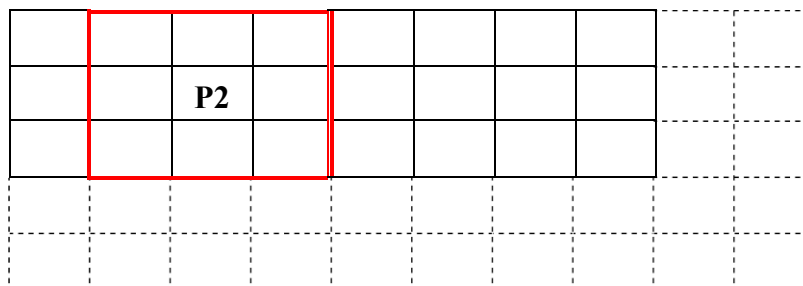
Κατά την επεξεργασία με κυλιόμενο παράθυρο αρχικά πρέπει να οριστεί το μέγεθος του κυλιόμενου παραθύρου. Η διαδικασία αυτή ονομάζεται και διαδικασία κυλιόμενης γειτονιάς, καθώς το κάθε παράθυρο ορίζεται ως προς τη γειτονιά ενός κεντρικού pixel. Η άμεση γειτονιά ενός pixel είναι τα 8 γειτονικά του pixels, επομένως το πιο μικρό παράθυρο είναι παράθυρο μεγέθους 3×3 (Εικόνα 13). Στις περισσότερες περιπτώσεις χωρικών φίλτρων χρησιμοποιείται το παράθυρο γειτονιάς 3×3 . Ωστόσο, ανάλογα με το πρόβλημα, μπορούμε να ορίσουμε και ως γειτονικά τα pixels που συνορεύουν με τα άμεσα γειτονικά pixels του κεντρικού pixel, και να προκύψουν έτσι παράθυρα μεγέθους 5×5 , 7×7 κ.λπ. Καθώς η διαδικασία λαμβάνει χώρα γύρω από ένα κεντρικό pixel, οι γειτονιές θα πρέπει να είναι μονών διαστάσεων. Ωστόσο, μπορούμε να ορίσουμε και γειτονιές μεγέθους $m \times n$, στις οποίες ως κεντρικό pixel ορίζεται αυτό που βρίσκεται στη θέση $\lfloor ([m - n] + 1) / 2 \rfloor$.

Έστω ότι επιθυμούμε να επεξεργαστούμε την παρακάτω εικόνα ανά περιοχές μεγέθους 3×3 pixels. Το κόκκινο πλαίσιο στην Εικόνα 13 απεικονίζει την πρώτη περιοχή (παράθυρο-γειτονιά) 3×3 της εικόνας αυτής (το κεντρικό pixel P1 βρίσκεται στη θέση (2,2)). Κατά τη διαδικασία επεξεργασίας με κυλιόμενο παράθυρο, επεξεργαζόμαστε την υποπεριοχή αυτή ως μια νέα εικόνα και το αποτέλεσμα της επεξεργασίας διαμορφώνει την τιμή του pixel της εικόνας εξόδου που βρίσκεται στην αντίστοιχη θέση με το κεντρικό pixel του παραθύρου αυτού (στην προκειμένη περίπτωση διαμορφώνει την τιμή του pixel εξόδου που βρίσκεται στη θέση (2,2)).



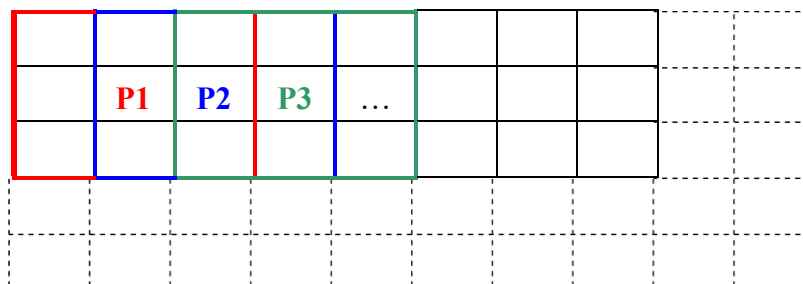
Εικόνα 13. Γειτονιά 3x3 του pixel P1

Κατά αντιστοιχία, το pixel της εικόνας εξόδου που βρίσκεται στη θέση (3,2) θα λάβει τιμή σύμφωνα με το αποτελέσματα της επεξεργασίας του παραθύρου 3x3 γύρω από το pixel εισόδου P2 (Εικόνα 14).



Εικόνα 14. Γειτονιά 3x3 του pixel P2

Η διαδικασία ονομάζεται επεξεργασία με κυλιόμενο παράθυρο καθώς – όπως φαίνεται στην Εικόνα 15 – πραγματοποιείται σάρωση όλης της εικόνας με ένα κυλιόμενο παράθυρο προκειμένου όλα τα pixels της εικόνας εξόδου να λάβουν τιμή με βάση τον αλγόριθμο που εφαρμόζεται στο κάθε παράθυρο.



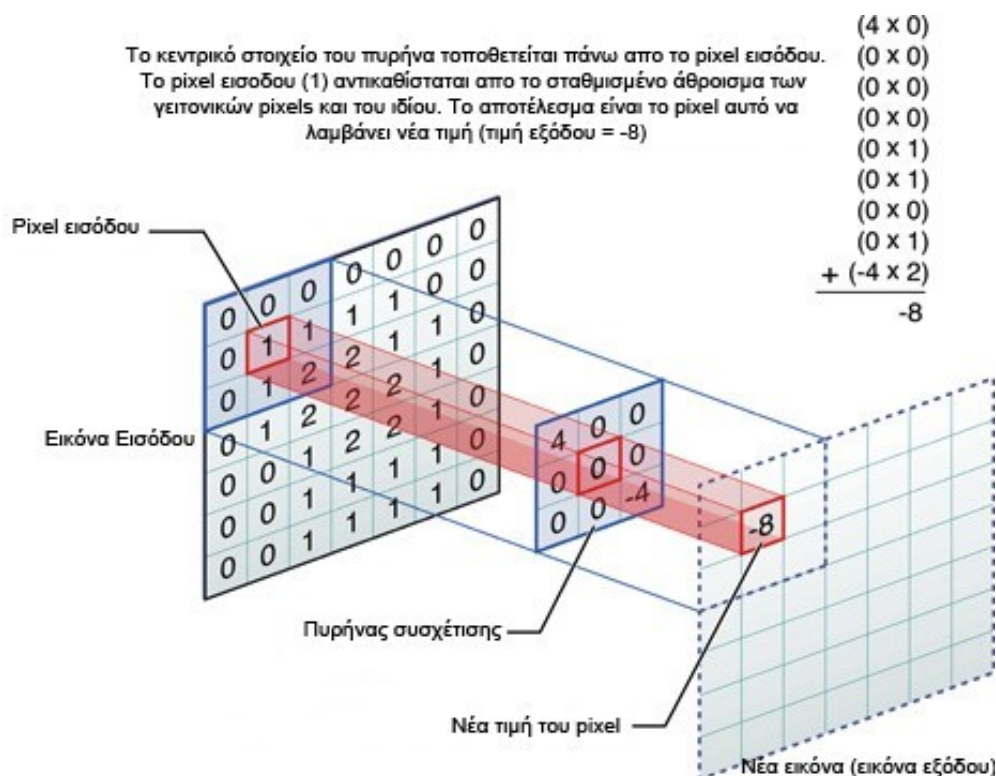
Εικόνα 15. Διαδικασία επεξεργασίας με χρήση κυλιόμενου παραθύρου

Όπως φαίνεται και παραπάνω, με τη συγκεκριμένη διαδικασία υπάρχει περιορισμός όσον αφορά τα pixels που βρίσκονται στα άκρα της εικόνας, (το pixel P1 που αναλύεται πρώτο είναι το pixel στη θέση (2,2)), δηλαδή τα pixels που βρίσκονται στις θέσεις (i,1) και (1,j). Για αυτά πραγματοποιείται ξεχωριστή διαδικασία κυλιόμενου παραθύρου, σύμφωνα με κάποια από τις προτεινόμενες προσεγγίσεις (π.χ. προστίθενται μηδενικά pixels στα όρια της εικόνας ώστε να είναι εφικτή η διαδικασία – zero padding, ή αντιγράφονται τα pixels των άκρων προκειμένου να γεμίσουν τα στοιχεία των παραθύρων – border replication).

3.2.2 Γραμμικό χωρικό φιλτράρισμα

Τα χωρικά φίλτρα χρησιμοποιούνται σε πολλές εφαρμογές της ψηφιακής επεξεργασίας εικόνας, όπως για παράδειγμα για τη βελτίωση μιας εικόνας, για αποθορυβοποίηση, για ανίχνευση ακμών κ.λπ. Τα χωρικά φίλτρα χωρίζονται σε δύο κατηγορίες, στα γραμμικά και στα μη γραμμικά. Κατά την επεξεργασία με χρήση γραμμικών φίλτρων, η τιμή ενός pixel εξόδου αποτελεί γραμμικό συνδυασμό των τιμών των pixels που βρίσκονται στη γειτονιά του αντίστοιχου pixel εισόδου. Το χωρικό φιλτράρισμα με χρήση γραμμικών φίλτρων πραγματοποιείται με χρήση είτε δισδιάστατης συνέλιξης (convolution), είτε δισδιάστατης συσχέτισης (correlation).

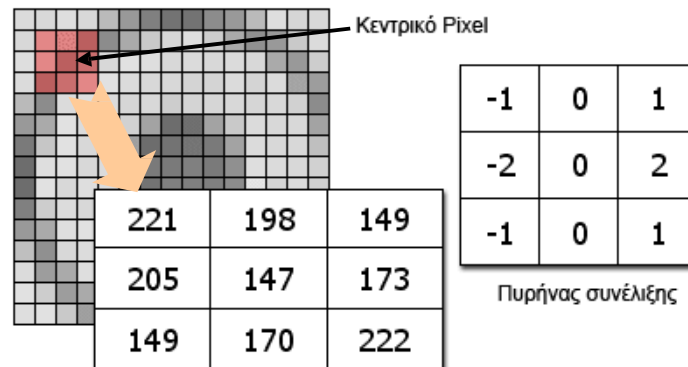
Η συνέλιξη και η συσχέτιση αποτελούν διαδικασίες κυλιόμενου παραθύρου κατά τις οποίες η τιμή ενός pixel στην εικόνα εξόδου είναι το σταθμισμένο άθροισμα των τιμών των pixels της γειτονιάς του pixel που βρίσκεται στην αντίστοιχη θέση της εικόνας εισόδου. Στη συσχέτιση, το σταθμισμένο άθροισμα είναι το άθροισμα των γινομένων των pixels της γειτονιάς με τους αντίστοιχους συντελεστές βάρους που ορίζονται από τον πυρήνα συσχέτισης, ο οποίος καθορίζει και το μέγεθος της γειτονιάς. Τυπικά ένας πυρήνας συσχέτισης (ή φίλτρο ή μάσκα) είναι ένας πίνακας διαστάσεων 3×3 . Στην παρακάτω εικόνα παρουσιάζεται η διαδικασία της συσχέτισης. Σημειώνεται ότι εφόσον δεν ορίζεται αρνητική φωτεινότητα, όταν προκύπτει αρνητικός αριθμός από τις λειτουργίες αυτές, αντικαθίσταται από την τιμή 0.



Εικόνα 16. Παράδειγμα συσχέτισης μιας εικόνας με έναν πυρήνα συσχέτισης

Η συνέλιξη είναι παρόμοια διαδικασία με τη διαφορά ότι ο πυρήνας συνέλιξης περιστρέφεται 180° πριν υπολογιστεί το σταθμισμένο άθροισμα. Έτσι, στη συνέλιξη το σταθμισμένο άθροισμα είναι το άθροισμα των γινομένων των pixels της γειτονιάς εισόδου με τους απέναντι συντελεστές βάρους του πυρήνα συνέλιξης. Ομοίως με τη συσχέτιση,

ένας πυρήνας συνέλιξης είναι συνήθως ένας πίνακας διαστάσεων 3x3. Παρακάτω παρουσιάζεται ένα παράδειγμα συνέλιξης.



Εικόνα 17. Παράδειγμα συνέλιξης μιας εικόνας με έναν πυρήνα συνέλιξης

Η τιμή του pixel της εικόνας εξόδου που βρίσκεται στην αντίστοιχη θέση με το κεντρικό pixel της εικόνας εισόδου, θα είναι:

$$P_{out} = (-1 * 222) + (0 * 170) + (1 * 149) + (-2 * 173) + (0 * 147) + (2 * 205) + (-1 * 149) + (0 * 198) + (1 * 221) = 63$$

Έτσι, ενώ στην αρχική εικόνα το κεντρικό pixel έχει τιμή 147, στην εικόνα εξόδου το αντίστοιχο pixel έχει τιμή 63.

Ανάλογα με τη διαδικασία (συνέλιξη ή συσχέτιση) καθώς και με τον πυρήνα που θα επιλέξουμε, μπορούμε να πραγματοποιήσουμε διάφορα είδη χωρικού φιλτραρίσματος, όπως αποθορυβοποίηση, ανίχνευση ακμών κ.λπ.

Ένα από τα πλέον απλά γραμμικά φίλτρα είναι το φίλτρο της μέσης τιμής (Mean filter). Χρησιμοποιείται για την εξομάλυνση και μείωση του θορύβου των εικόνων. Η λειτουργία του φίλτρου αυτού συνίσταται στην αντικατάσταση της φωτεινότητας κάθε pixel με τη μέση φωτεινότητα της γειτονιάς του. Ένα τυπικό φίλτρο μέσης τιμής πραγματοποιείται με τη διαδικασία συνέλιξης (ή συσχέτισης) και με χρήση του πυρήνα:

$$\frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

(πολλαπλασιάζουμε με συντελεστή 1/9 προκειμένου να προκύψει από τη συνέλιξη η μέση τιμή του αθροίσματος των τιμών των 9 pixels της γειτονιάς)

Η μέθοδος αυτή, καθώς και όλες οι γραμμικές μέθοδοι φιλτραρίσματος για την αποθορυβοποίηση εικόνας, εκτός από την εξομάλυνση του θορύβου προκαλούν θάμπωμα της εικόνας και αποκοπή των οξείων λεπτομερειών της. Για το λόγο αυτό δεν είναι ιδανικές για την αποκοπή θορύβου σε ψηφιακές εικόνες.

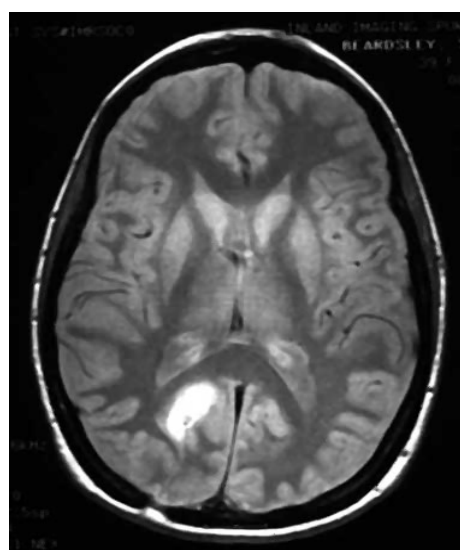
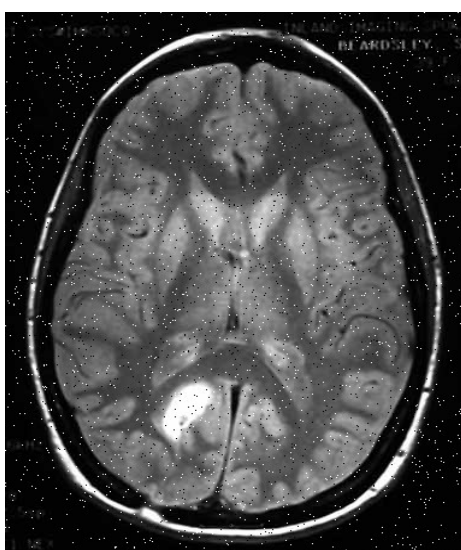
3.2.3 Μη γραμμικό χωρικό φιλτράρισμα

Οι μη γραμμικές μέθοδοι φιλτραρίσματος αποτελούν μεθόδους που λειτουργούν – όπως και οι γραμμικές – με χρήση κυλιόμενων παραθύρων, αλλά για την εφαρμογή τους χρησιμοποιείται κάποιος μη γραμμικός αλγόριθμος και όχι συνέλιξη ή συσχέτιση.

Ένα απλό παράδειγμα μη γραμμικού φιλτραρίσματος είναι η αποκοπή θορύβου salt and pepper με χρήση φίλτρου Median. Ο θόρυβος salt and pepper χαρακτηρίζεται από τυχαία κατανεμημένα λευκά και μαύρα pixels σε μια εικόνα (Εικόνα 18). Παρατηρώντας την Εικόνα 18 είναι προφανές πως το φίλτρο μέσης τιμής δεν είναι κατάλληλο για την αποκοπή του συγκεκριμένου θορύβου, καθώς τα λευκά pixels που αποτελούν το θόρυβο δεν θα εξαλειφθούν πλήρως, αντιθέτως θα αναμειχθούν με τα γειτονικά τους και θα προκαλέσουν αύξηση της μέσης τιμής φωτεινότητας κάθε γειτονιάς.

Το φίλτρο Median αντικαθιστά την τιμή κάθε pixel με την ενδιάμεση τιμή από ένα παράθυρο pixels στη γειτονιά του. Για παράδειγμα, ας θεωρήσουμε ένα παράθυρο 3x3 στη γειτονιά ενός pixel που έχει τιμή 250, με τιμές των γειτονικών pixels {10, 20, 20, 20, 250, 20, 20, 25, 15} όπως φαίνεται παρακάτω. Καθώς οι τιμές των γειτονικών pixels κυμαίνονται μεταξύ 10 και 25, το pixel με την τιμή 250 αντιστοιχεί σε θόρυβο salt and pepper. Το φίλτρο Median αναδιατάσσει τις τιμές κατά αύξουσα διάταξη {10, 15, 20, 20, 20, 20, 20, 20, 25} και στη συνέχεια επιλέγει την ενδιάμεση τιμή, δηλαδή την τιμή 20, για να αντικαταστήσει την τιμή 250. Με τον τρόπο αυτό ο συγκεκριμένος θόρυβος εξαλείφεται πλήρως. Η διαδικασία επαναλαμβάνεται για όλα τα pixels της εικόνας. Το αποτέλεσμα της αποθορυβοποίησης με χρήση φίλτρου Median φαίνεται στην Εικόνα 18.

10	20	20	20	10	...		
20	250	20	20	10	...		
20	25	15	10	10	...		
10	20	10	20	10	...		



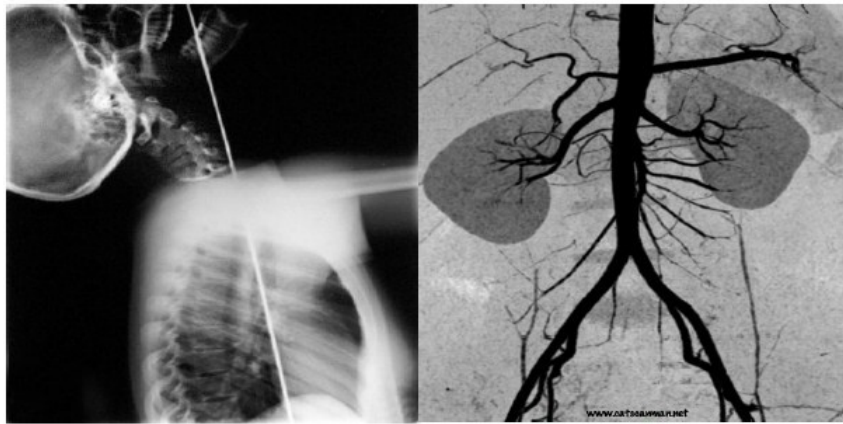
Εικόνα 18. MRI εικόνα εγκεφάλου με “salt and pepper” θόρυβο (αριστερά) και η ίδια εικόνα με εξάλειψη του θορύβου (δεξιά)

3.3 Ανίχνευση ακμών (Edge Detection)

Η ανίχνευση ακμών αποτελεί θεμελιώδες εργαλείο στην επεξεργασία εικόνας και την όραση υπολογιστών, ιδιαίτερα στους τομείς της ανίχνευσης και εξαγωγής χαρακτηριστικών, οι οποίοι αποσκοπούν στον εντοπισμό σημείων σε μια εικόνα στα οποία η φωτεινότητά της αλλάζει απότομα (παρουσιάζει ασυνέχειες).

Ως ακμή σε μία gray-scale εικόνα ορίζεται η σχετική ασυνέχεια μεταξύ δύο διαβαθμίσεων του γκρι. Στην πράξη μια ακμή θεωρείται το σύνορο μεταξύ δύο ομοιογενών περιοχών με διαφορετική φωτεινότητα ή αλλιώς το περίγραμμα αυτών. Είναι απαραίτητο να υιοθετηθούν διάφορες παραδοχές προκειμένου να ξεχωρίσουμε τις ασυνέχειες που έχουν νόημα από εκείνες που δεν έχουν. Για παράδειγμα, οι δύο παρακάτω εικόνες αναπαριστούν μία εικόνα ακτινογραφίας (αριστερά) και μία εικόνα αγγειογραφίας (δεξιά). Η ανίχνευση των ακμών στην πρώτη εικόνα είναι πολύ πιο εύκολη υπόθεση από ό,τι στη δεύτερη, καθώς είναι πιο εύκολο να οριστούν οι ασυνέχειες, η μετάβαση δηλαδή από pixels με τιμή 0 (μαύρο) – τα οποία αντιστοιχούν στο φόντο της εικόνας – σε pixels με τιμές μεγαλύτερες του 150 – τα οποία αντιστοιχούν στο σκελετό. Αντίθετα, στη δεύτερη εικόνα θα πρέπει να βρούμε μια φόρμουλα που να ορίζει ποιες μεταβάσεις αποτελούν ασυνέχειες στα gray levels, π.χ. η μετάβαση από pixels με τιμή x1 σε pixels με τιμή x2.

Στην ιδανική περίπτωση, το αποτέλεσμα της εφαρμογής ενός ανιχνευτή ακμών σε μια εικόνα οδηγεί σε ένα σύνολο συνδεδεμένων καμπυλών που δείχνουν τα όρια (περιγράμματα) των διαφόρων αντικειμένων της εικόνας. Ωστόσο, καθώς δεν υπάρχει γενικό μαθηματικό μοντέλο που να καθορίζει τις ακμές σε μια εικόνα, έχουν κατά καιρούς προταθεί διάφορες προσεγγίσεις ανίχνευσης ακμών που καταλήγουν σε διαφορετικά αποτελέσματα ανίχνευσης.



Εικόνα 19.

Αριστερά: Ακτινογραφία της άνω περιοχής του σώματος

Δεξιά: Εικόνα αγγειογραφίας στην περιοχή των νεφρών

Οι δύο πιο γνωστές τεχνικές ανίχνευσης ακμών είναι οι τεχνικές που προτάθηκαν από τους Sobel και Canny.

Η τεχνική Sobel δίνει έμφαση σε περιοχές υψηλής χωρικής συχνότητας, οι οποίες αντιστοιχούν σε ακμές. Η τεχνική εφαρμόζεται με χρήση ενός τελεστή (τελεστής Sobel) ο οποίος αποτελείται από δύο φίλτρα (μάσκες ακμών), ένα για να ανιχνεύσει τις αλλαγές στην κάθετη κατεύθυνση (G_y) και το άλλο για να ανιχνεύσει τις αλλαγές στην οριζόντια (G_x). Οι δύο μάσκες του τελεστή Sobel είναι οι δύο πυρήνες συνέλιξης 3×3 που παρουσιάζονται παρακάτω. Κατά τη διαδικασία ανίχνευσης ακμών πραγματοποιείται

συνέλιξη μεταξύ της εικόνας και των δύο αυτών μασκών. Από τον κατάλληλο συνδυασμό των δύο εικόνων που προκύπτουν αναδεικνύονται οι ακμές των αντικειμένων της εικόνας. Στην πιο απλή περίπτωση υπολογίζεται ένα μέγεθος κλίσης για κάθε pixel, π.χ $|G| = |G_x| + |G_y|$, και στην περίπτωση που αυτό υπερβεί ένα κατώφλι τότε ορίζεται πως το pixel αποτελεί σημείο ακμής και λαμβάνει τιμή 255 (αλλιώς λαμβάνει τιμή 0). Με τη διαδικασία αυτή σαρώνεται όλη η εικόνα και όσα σημεία ανιχνευθούν ότι αποτελούν σημεία ακμών αναδεικνύονται στο προσκήνιο, όπως φαίνεται και στο παράδειγμα της Εικόνας 21. Ο τελεστής Sobel υπολογίζει την απόκλιση της έντασης κάθε pixel σε σχέση με τα γειτονικά του και δίνει ως αποτέλεσμα την πιο πιθανή μέγιστη αύξηση από τα σκούρα gray levels στα πιο ανοικτά, καθώς και τη μεταβολή αυτής. Ουσιαστικά δείχνει πώς μεταβάλλεται η τιμή φωτεινότητας στις οριζόντιες και κάθετες κατευθύνσεις, και γι' αυτό τα σημεία μεγάλης μεταβολής είναι πολύ πιθανό να είναι ακμές. Ανήκει στις τεχνικές ανίχνευσης ακμών με χρήση πρώτης παραγώγου.

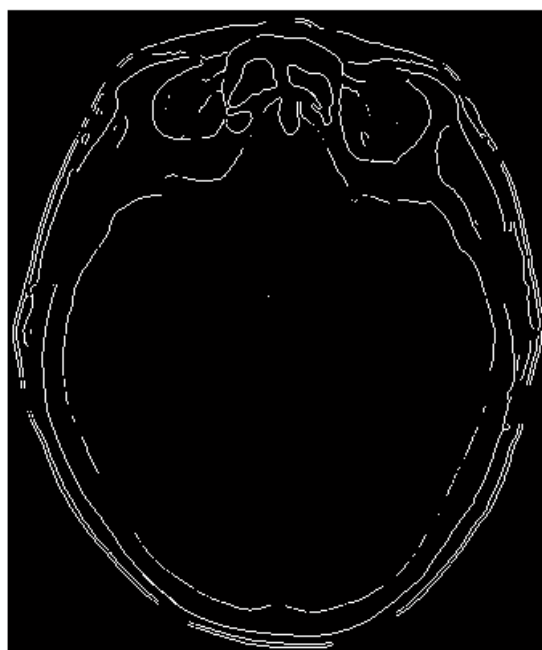
-1	0	+1
-2	0	+2
-1	0	+1

G_x

+1	+2	+1
0	0	0
-1	-2	-1

G_y

Εικόνα 20. Μάσκες Sobel μεγέθους 3x3



Εικόνα 21. Ανίχνευση των ακμών της αρχικής CT εικόνας εγκεφάλου (εικόνα 8) με χρήση της τεχνικής Sobel

Ο αλγόριθμος Canny αποτελεί μια από τις καλύτερες τεχνικές ανίχνευσης ακμών. Είναι ένας σύνθετος πολυσταδιακός αλγόριθμος που αναπτύχθηκε το 1986 από τον John F. Canny με στόχο τη βέλτιστη και ευρεία ανίχνευση ακμών. Ο αλγόριθμος εφαρμόζει σταδιακά διάφορες τεχνικές, όπως εξομάλυνση της εικόνας με χρήση Gaussian συνέλιξης, ανίχνευση των περιοχών με υψηλή χωρική πρώτη παράγωγο (συνήθως με χρήση τελεστή Sobel), κατωφλίωση κ.ά., προκειμένου να ανιχνευθεί και αναδειχθεί ένα μεγάλο εύρος ακμών στην εικόνα. Η τεχνική ανίχνευσης ακμών Canny υπερτερεί των τεχνικών ανίχνευσης ακμών με χρήση πρώτης παραγώγου (όπως ο Sobel) στο ότι ανιχνεύει επιπλέον αδύνατες ακμές που ενδεχομένως να περιέχουν σημαντική πληροφορία.



Εικόνα 22 Ανίχνευση των ακμών της αρχικής CT εικόνας εγκεφάλου με χρήση του αλγόριθμου Canny

3.4 Τμηματοποίηση εικόνας (Image Segmentation)

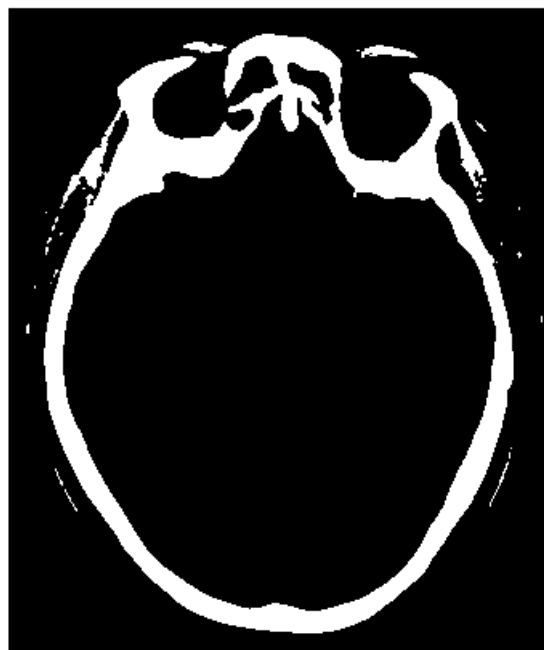
Η τμηματοποίηση (κατάτμηση) εικόνας έχει ως στόχο την υποδιαίρεση μιας εικόνας στις περιοχές ή στα αντικείμενα που την αποτελούν. Είναι μια διαδικασία με την οποία διαχωρίζεται μια εικόνα σε κατάλληλες περιοχές ή αντικείμενα. Βασίζεται στο γεγονός πως μια περιοχή ή ένα αντικείμενο σε μια εικόνα αναμένεται να έχει ομοιογενή χαρακτηριστικά, όπως ένταση, υφή κ.ά. Αφορά μεθόδους που έχουν ως εισόδους ολόκληρες εικόνες και ως εξόδους στοιχεία που εξάγουμε από αυτές τις εικόνες. Οι αλγόριθμοι που εφαρμόζονται στηρίζονται στην ασυνέχεια -απότομη μεταβολή- των τιμών των pixels (ανίχνευση μεμονωμένων σημείων, γραμμών ή ακμών) ή στην ομοιότητά τους (π.χ. thresholding). Η απλούστερη μέθοδος τμηματοποίησης είναι αυτή της κατωφλίωσης.

- **Κατωφλίωση (Thresholding)**: Η εφαρμογή της τεχνικής κατωφλίωσης σε μια ψηφιακή εικόνα βασίζεται στην υπόθεση ότι τα pixels του προς μελέτη αντικειμένου (προσκήνιο) μπορούν να διαχωριστούν από τα pixels των άλλων αντικειμένων και του φόντου (παρασκήνιο) με βάση τις τιμές φωτεινότητάς τους. Ένας απλός τρόπος εφαρμογής είναι ο καθορισμός μιας κλίμακας τιμών φωτεινότητας (κατώφλι) σύμφωνα με τις τιμές που εμφανίζουν τα pixels που

ανήκουν στην περιοχή ενδιαφέροντος, και στη συνέχεια η εφαρμογή του κατωφλίου σε όλη την εικόνα προκειμένου τα pixels που βρίσκονται εκτός των ορίων του κατωφλίου να τοποθετηθούν στο παρασκήνιο. Στη δημιουργηθείσα εικόνα τα pixels που παρουσιάζουν φωτεινότητα εντός του εύρους τιμών φωτεινότητας του κατωφλίου παραμένουν αμετάβλητα, ενώ αυτά που βρίσκονται εκτός του εύρους μηδενίζονται. Με τον τρόπο αυτό, η νέα εικόνα περιλαμβάνει μόνο το αντικείμενο ενδιαφέροντος. Ένας τρόπος επιλογής του κατωφλίου είναι με οπτική εξέταση του ιστογράμματος της εικόνας. Μια άλλη μέθοδος επιλογής τιμής του κατωφλίου είναι μέσω δοκιμής και σφάλματος. Στην περίπτωση αυτή επιλέγουμε διάφορες τιμές του κατωφλίου έως ότου κάποια παράγει καλά αποτελέσματα με βάση δεδομένα κριτήρια του παρατηρητή. Όπως είναι προφανές, η τεχνική της κατωφλίωσης αποτελεί τεχνική σημειακής επεξεργασίας.

- **Region Growing**: Είναι μια διαδικασία που ομαδοποιεί pixels σε μεγαλύτερες περιοχές. Η πιο απλή εκδοχή της είναι η pixel aggregation, η οποία ξεκινά με μια ομάδα «σπόρων» (seeds) και προσδιορίζει μια περιοχή προσαρτώντας σε κάθε «σπόρο» εκείνα τα γειτονικά pixels που έχουν παρόμοιες ιδιότητες (διαβάθμιση του γκρι, χρώμα, υφή). Βασικές δυσκολίες της τεχνικής αυτής είναι η επιλογή των αρχικών «σπόρων» και των κριτηρίων ομοιότητας, καθώς και η διαμόρφωση ενός stopping rule με τη χρήση εννοιών όπως το μέγεθος και το σχήμα της περιοχής, η ομοιότητα του υπό εξέταση pixel με την ήδη δημιουργηθείσα περιοχή κ.ά.

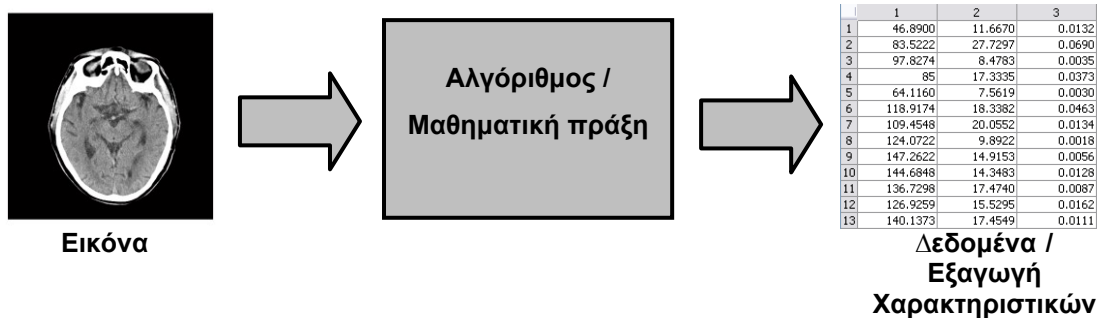
Για παράδειγμα, σε μια CT εικόνα εγκεφάλου ενδεχομένως να μας ενδιαφέρει η κατακράτηση του κρανιακού οστού. Αυτό επιτυγχάνεται διακρίνοντας τις τιμές φωτεινότητας που έχει το οστό στην αρχική εικόνα. Μελετώντας την αρχική εικόνα (Εικόνα 8) είναι εύκολο να διαπιστώσουμε ότι τα pixels που αναπαριστούν το κρανίο παρουσιάζουν τιμές φωτεινότητας πολύ κοντά στο 255 (λευκό). Επομένως, ένας απλός τρόπος να δημιουργήσουμε μια εικόνα στην οποία θα απεικονίζεται μόνο το κρανίο είναι να εφαρμόσουμε κατωφλίωση με εύρος τιμών κατωφλίου [250, 255]. Με αυτό τον τρόπο θα κρατήσουμε όλα τα pixels που παρουσιάζουν φωτεινότητα εντός του διαστήματος [250, 255] και θα μηδενίσουμε όλα τα υπόλοιπα. Το αποτέλεσμα αυτής της κατωφλίωσης φαίνεται στην Εικόνα 23 όπου διακρίνεται ξεκάθαρα μόνο το κρανίο από την εικόνα CT.



Εικόνα 23. Κατωφλιωμένη (thresholded) CT εικόνα εγκεφάλου (ανάδειξη κρανιακού οστού)

6. Ανάλυση ψηφιακής εικόνας

Η ανάλυση μιας ψηφιακής εικόνας αποσκοπεί στην εξαγωγή χρήσιμης πληροφορίας (ποσοτικής και ποιοτικής) από αυτή, καθώς και στην κατανόηση της πληροφορίας αυτής. Η ανάλυση ψηφιακής εικόνας αφορά διαδικασίες κατά τις οποίες είσοδο αποτελεί μια εικόνα και έξοδο αποτελεί κάποιο είδος πληροφορίας, όπως δεδομένα τα οποία αντιστοιχούν σε ορισμένα χαρακτηριστικά της εικόνας. Τα χαρακτηριστικά χρησιμοποιούνται για να περιγράψουν ή να χαρακτηρίσουν αντικείμενα που περιέχονται στις εικόνες και πρέπει επομένως να επιλέγονται κατάλληλα και ανάλογα με τις ιδιαίτερες απαιτήσεις της κάθε εφαρμογής.



Εικόνα 24. Εξαγωγή χαρακτηριστικών από την ανάλυση μιας εικόνας

Στις περισσότερες περιπτώσεις, κατά το στάδιο της ανάλυσης, η εικόνα εισόδου είναι μια εικόνα που έχει προηγουμένως υποστεί κατάλληλη επεξεργασία προκειμένου να βελτιωθεί η ποιότητα της αρχικής εικόνας ή να αναδειχθούν (απομονωθούν) οι περιοχές ή τα αντικείμενα που επιθυμούμε να μελετήσουμε. Για παράδειγμα, μια εικόνα που περιέχει θόρυβο, πρέπει αρχικά να υποστεί κατάλληλη αποθορυβοποίηση προκειμένου να αναλυθεί σωστά. Είναι επομένως σημαντικό, η εικόνα εξόδου που δημιουργείται κατά την προεπεξεργασία (και η οποία αποτελεί την είσοδο του συστήματος ανάλυσης) να είναι κατάλληλη προκειμένου να εξαχθεί χρήσιμη και ποιοτική πληροφορία. Επίσης, απαιτείται πολλές φορές η περαιτέρω ανάλυση (π.χ. στατιστική ανάλυση) της εξαγόμενης πληροφορίας, προκειμένου να καταλήξουμε σε κάποιο χρήσιμο συμπέρασμα για την εικόνα ή τα αντικείμενα που μελετούμε. Η εξαγωγή συμπερασμάτων για ένα αντικείμενο που περιγράφεται από ένα σύνολο χαρακτηριστικών μπορεί να πραγματοποιηθεί με χρήση τεχνικών ταξινόμησης προτύπων, όπως θα δούμε παρακάτω.

Από την ανάλυση μιας εικόνας μπορούν να εξαχθούν διάφορα χαρακτηριστικά που αντιπροσωπεύουν τα αντικείμενα που περιέχει αυτή, όπως χρώμα, σχήμα, υφή, συχνοτικό περιεχόμενο (σε περίπτωση που η ανάλυση λαμβάνει χώρα στο πεδίο των συχνοτήτων), κ.ά. Παρακάτω θα αναλυθούν ορισμένα από αυτά.

6.1 Ανάλυση χρώματος (Colour Analysis)

Η ανάλυση χρώματος αποσκοπεί στην εξαγωγή πληροφορίας σχετικής με το χρώμα που παρουσιάζει μια εικόνα ή τα αντικείμενα που περιέχει. Όπως παρουσιάστηκε στην Ενότητα 2.1, μια έγχρωμη εικόνα RGB αποτελείται από τρεις συνιστώσες, κάθε μια από τις οποίες εκφράζει την ένταση του κόκκινου, του πράσινου και του μπλε χρώματος στην εικόνα. Κατά τη χρωματική ανάλυση εξάγουμε πληροφορίες σχετικά με το χρώμα των υπό μελέτη αντικειμένων, όπως για παράδειγμα τη μέση τιμή των τριών συνιστωσών R, G, και

Β. Στόχος μας είναι να μετατρέψουμε το χρώμα σε ένα μετρήσιμο μέγεθος, ούτως ώστε να το χρησιμοποιήσουμε στη συνέχεια για διάφορους σκοπούς, όπως για τη σύγκριση δύο αντικειμένων που παρουσιάζουν παρόμοιο χρώμα. Για το σκοπό αυτό έχουν οριστεί τα λεγόμενα χρωματικά μοντέλα (ή χρωματικοί χώροι) που χρησιμοποιούνται για να περιγράψουν την χρωματική πληροφορία μιας εικόνας. Το πιο γνωστό χρωματικό μοντέλο είναι το RGB. Πέρα όμως από το RGB, υπάρχουν και άλλα χρωματικά μοντέλα που χρησιμοποιούνται για να περιγράψουν το χρώμα μιας εικόνας, όπως το μοντέλο HSV (Hue, Saturation, Value), το μοντέλο L^*a^*b κ.ά. Τα μοντέλα αυτά έχουν δημιουργηθεί προκειμένου να περιγράψουν το χρώμα μιας εικόνας με διαφορετικό τρόπο από ό,τι το RGB, ούτως ώστε να δίνονται επιπλέον πληροφορίες για αυτό. Για παράδειγμα, η απόχρωση ενός αντικειμένου στο χρωματικό χώρο RGB περιγράφεται από τρεις συνιστώσες (R, G, B). Στο χρωματικό χώρο HSV η απόχρωση περιγράφεται από μία συνιστώσα, τη συνιστώσα Hue (απόχρωση). Έτσι, εάν θέλουμε να συγκρίνουμε την απόχρωση δύο αντικειμένων είναι πολύ πιο εύκολο να αναλύσουμε την εικόνα με το χρωματικό μοντέλο HSV παρά με το RGB. Για να το πετύχουμε αυτό, μετασχηματίζουμε την εικόνα από το χώρο RGB στο χώρο HSV μέσω κατάλληλων μετασχηματισμών. Ως παράδειγμα ας θεωρήσουμε μια μωβ απόχρωση, η οποία στο χώρο RGB αναπαρίσταται από τις τιμές $[R, G, B] = [150, 20, 150]$, ενώ στο χώρο HSV αναπαρίσταται από την τιμή $H = 300^\circ$ (στον HSV η συνιστώσα Hue μετριέται σε μοίρες, από 0 έως 360).

6.2 Ανάλυση σχήματος (Shape Analysis)

Η ανάλυση σχήματος αποσκοπεί στην εξαγωγή πληροφορίας σχετικής με το σχήμα που έχουν τα αντικείμενα μιας εικόνας. Έχουν προταθεί διάφορες τεχνικές για την ανάλυση του σχήματος ενός αντικειμένου, οι οποίες εξάγουν διάφορα χαρακτηριστικά σχήματος. Τα πιο απλά χαρακτηριστικά σχήματος είναι τα γεωμετρικά χαρακτηριστικά, τα οποία προκύπτουν άμεσα από τη γεωμετρική μορφή των αντικειμένων. Ορισμένα από τα σημαντικότερα γεωμετρικά χαρακτηριστικά είναι τα εξής:

- Περίμετρος
- Επιφάνεια
- Ακτίνα
- Εκκεντρότητα
- Καμπυλότητα

Η περίμετρος ενός αντικειμένου ορίζεται ως το μήκος του εξωτερικού περιγράμματός του. Σε μια ψηφιακή εικόνα η περίμετρος ενός αντικειμένου μπορεί να ισούται απλά με το πλήθος των pixels του περιγράμματός του. Αντίστοιχα, η επιφάνεια (εμβαδό) ενός αντικειμένου ισούται με το πλήθος των pixels που περιέχει αυτό (μέσα στην περίμετρό του). Με παρόμοιο τρόπο ορίζεται και η ακτίνα ενός αντικειμένου (σε περίπτωση που αυτό δεν είναι κύκλος ορίζονται 2 ακτίνες, η μέγιστη και η ελάχιστη απόσταση του κέντρου μάζας του από την περίμετρο). Η εκκεντρότητα ορίζεται από το λόγο της μέγιστης ακτίνας προς την ελάχιστη και είναι χαρακτηριστικό της κυκλικότητας του αντικειμένου.

Αυτά τα γεωμετρικά χαρακτηριστικά (και πολλά άλλα) είναι πολύ χρήσιμα στην περιγραφή και ανάλυση ενός αντικειμένου μιας εικόνας. Στην παρακάτω εικόνα μαγνητικής τομογραφίας (τομή από εγκέφαλο), η λευκή περιοχή που βρίσκεται εντός του εγκεφάλου αντιστοιχεί σε όγκο (σημειωμένη με πλαίσιο). Στην εικόνα αυτή είναι πολύ σημαντικό να γνωρίζουμε τα γεωμετρικά χαρακτηριστικά της υπό μελέτη περιοχής. Με κατάλληλη επεξεργασία της εικόνας MRI (πολυκατωφλίωση και εξαγωγή των περιοχών

ενδιαφέροντος) μπορούμε να εξάγουμε την περιοχή όγκου προκειμένου να την αναλύσουμε ξεχωριστά από τις υπόλοιπες περιοχές. Το αποτέλεσμα της επεξεργασίας οδηγεί σε μια νέα εικόνα η οποία αντιπροσωπεύει την περιοχή του όγκου (Εικόνα 25, δεξιά). Η ανάλυση σχήματος της νέας αυτής εικόνας οδηγεί στα εξής γεωμετρικά χαρακτηριστικά του όγκου:

- Περίμετρος = 146 pixels
- Επιφάνεια = 1074 pixels
- Μέγιστη ακτίνα = 56 pixels
- Ελάχιστη ακτίνα = 25 pixels
- Εκκεντρότητα = 0.893

Λαμβάνοντας διάφορες τομές του εγκεφάλου και εφαρμόζοντας τεχνικές ανακατασκευής ψηφιακής εικόνας, μπορούμε να δημιουργήσουμε ένα τρισδιάστατο μοντέλο του εγκεφάλου (3D Reconstruction). Από αυτό το τρισδιάστατο μοντέλο μπορούμε να υπολογίσουμε τον όγκο της περιοχής που καταλαμβάνει ο καρκίνος, δεδομένο πολύ χρήσιμο για τους γιατρούς.



Εικόνα 25: MRI εικόνα εγκεφάλου (αριστερά) και εξαγωγή της περιοχής ενδιαφέροντος που αντιστοιχεί στον όγκο (δεξιά) προκειμένου να υπολογιστούν τα γεωμετρικά χαρακτηριστικά του

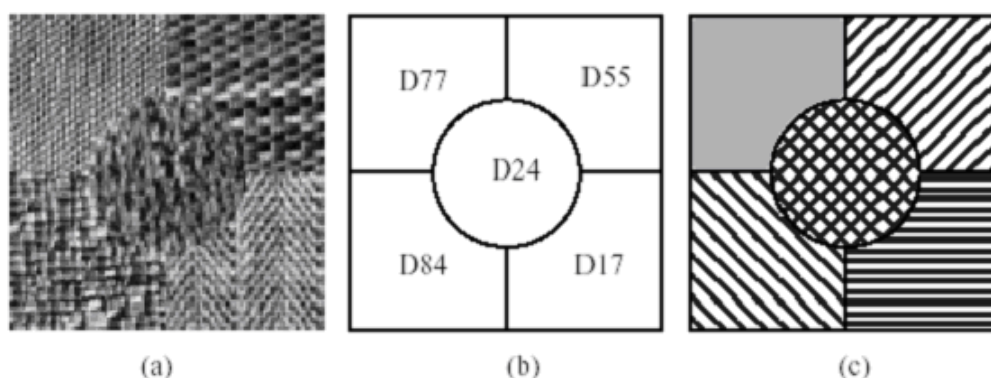
6.3 Ανάλυση υφής (Texture Analysis)

Η έννοια της υφής αναφέρεται στην αίσθηση που δημιουργείται κατά την επαφή του ανθρώπου με μια επιφάνεια και εξαρτάται από τη δομή της επιφάνειας. Στις ψηφιακές εικόνες δεν υπάρχει αυστηρός ορισμός της υφής. Οι εικόνες είναι πίνακες pixels, κάθε ένα από τα οποία περιγράφεται από τη χρωματικότητα του. Στην περίπτωση των ασπρόμαυρων εικόνων κάθε στοιχείο περιγράφεται από έναν και μοναδικό αριθμό, την ένταση της φωτεινότητας στο συγκεκριμένο σημείο. Η υφή σχετίζεται με τα επίπεδα φωτεινότητας (gray levels) που εμφανίζονται στην εικόνα και τον τρόπο που αυτά κατανέμονται, ώστε να δίνεται η αίσθηση του τρισδιάστατου και του ανάγλυφου. Μεταβολές στην ένταση φωτεινότητας των pixels μιας εικόνας, εμφανιζόμενες με κάποιο συστηματικό τρόπο, χαρακτηρίζουν την υφή της.

Η υφή μπορεί να προσδιορισθεί από τις παρακάτω παραμέτρους :

- Αδρότητα (Coarseness)
- Αντίθεση (Contrast)
- Τραχύτητα (Roughness)
- Ομαλότητα (Regularity)
- Περιοδικότητα (Periodicity)
- Ινώδες (Line-Lightness)
- Κατευθυντικότητα (Directionality)

Η ανάλυση της υφής στις εικόνες είναι χρήσιμη σε ένα μεγάλο αριθμό εφαρμογών και είναι αντικείμενο έρευνας εδώ και πολλά χρόνια. Η πιο άμεση εφαρμογή είναι η αναγνώριση περιοχών μιας εικόνας με βάση την υφή. Για παράδειγμα, στην παρακάτω εικόνα μπορούμε να αναγνωρίσουμε πέντε περιοχές με διαφορετική υφή.



Εικόνα 26. α) Εικόνα με 5 διαφορετικής υφής περιοχές. β) Χάρτης ταξινόμησης της εικόνας με βάση την κατηγορία υφής. γ) Τμηματοποιημένη εικόνα με βάση την υφή

Πολλές φορές η υφή που έχουν κάποια αντικείμενα μπορεί να φανερώσει την κατάσταση στην οποία βρίσκονται (π.χ. ένα καρκινικό κύτταρο μπορεί να έχει διαφορετική υφή από ένα υγιές) ή άλλες φορές είναι μοναδική και μάς βοηθά να το ξεχωρίζουμε από άλλα (π.χ η υφή του ξύλου ή η υφή του τούβλινου τοίχου). Σε τέτοιες περιπτώσεις μπορούμε να εκμεταλλευτούμε την υφή που απεικονίζεται σε ψηφιακές εικόνες ώστε να αναγνωρίσουμε τα αντικείμενα που υπάρχουν σε αυτήν. Για να γίνει αυτό πρέπει να έχουμε μια βάση δεδομένων με υφές οι οποίες αντιστοιχούν σε αντικείμενα. Η νέα υφή που δεχόμαστε ως είσοδο πρέπει να συγκριθεί με όλες τις υφές της βάσης και να καταταχθεί σε μία από τις υπάρχουσες κατηγορίες. Τελικά μπορούμε να αποφανθούμε για το είδος ή τη φύση του αντικειμένου στο οποίο ανήκει η υφή εισόδου. Το πρόβλημα αυτό ονομάζεται κατηγοριοποίηση ή *ταξινόμηση υφής (texture classification)* και με αυτό θα ασχοληθούμε στην παρούσα άσκηση.

Σε κάποιες περιπτώσεις, η υφή που δίνεται ως είσοδος είτε δεν υπάρχει στη βάση δεδομένων, είτε για κάποιους λόγους δεν μπορεί να κατηγοριοποιηθεί σε μια από τις υπάρχουσες κατηγορίες. Επίσης πολύ συχνά μας ενδιαφέρει η εύρεση ομοιογενών περιοχών, δηλαδή περιοχών με την ίδια υφή. Σε αυτές τις περιπτώσεις το ζητούμενο είναι η κατηγοριοποίηση των περιοχών (περιοχή 1 με συγκεκριμένα σύνορα, περιοχή 2 πάλι με συγκεκριμένα σύνορα κ.ο.κ). Το πρόβλημα αυτό ονομάζεται *τμηματοποίηση περιοχών υφής (texture segmentation)*.

4.3.1 Μέθοδοι ανάλυσης υφής

Τα τελευταία χρόνια πολλές μέθοδοι, μικρής ή μεγάλης πολυπλοκότητας, έχουν αναπτυχθεί για να αναλύσουν και να περιγράψουν την υφή εικόνας. Οι κυριότερες μέθοδοι ανάλυσης και προσδιορισμού χαρακτηριστικών υφής είναι οι εξής:

- Στατιστικά στοιχεία υφής (Texture Statistics)
- Πίνακες συνεμφάνισης – Στατιστικά στοιχεία πινάκων συνεμφάνισης (Co-occurrence matrices Statistics)
- Μέθοδος μήκους διαδρομής (Primitive or Run Length method)
- Αυτοσυσχέτιση (Autocorrelation)
- Φάσμα Ισχύος Fourier (Fourier Power Spectrum)
- Φάσμα Υφής (Texture Spectrum)
- Χαρακτηριστικά ενέργειας υφής (Texture Energy)
- Χαρακτηριστικά υφής βασιζόμενα σε fractals
- Χαρακτηριστικά υφής βασιζόμενα σε κυματίδια (wavelets)

Κάθε μια από τις παραπάνω μεθόδους ανάλυσης υφής εμφανίζει συγκεκριμένα πλεονεκτήματα και μειονεκτήματα και είναι βέλτιστη για διαφορετικές περιπτώσεις ανάλυσης. Στην παρούσα άσκηση θα ασχοληθούμε με την πιο απλή μέθοδο, δηλαδή τα στατιστικά στοιχεία υφής.

Στατιστικά στοιχεία υφής

Τα στατιστικά χαρακτηριστικά υφής (texture statistics) είναι απλά στατιστικά μεγέθη όπως η μέση τιμή, η ενδιάμεση τιμή, η τυπική απόκλιση και οι ροπές 3^{ης} και 4^{ης} τάξης των επιπέδων φωτεινότητας της εικόνας. Βασίζονται στην κατανομή των επιπέδων του γκρι (gray-levels) των pixels, δηλαδή στο ιστόγραμμα της εικόνας.

Έστω g η τιμή της έντασης της φωτεινότητας ενός pixel. Στην περίπτωση των ασπρόμαυρων εικόνων, η φωτεινότητα g παίρνει ακέραιες τιμές στο διάστημα $[0, 255]$ για την αναπαράσταση των διάφορων αποχρώσεων του γκρι. Αν $h(g)$ είναι το συνολικό πλήθος των pixels της εικόνας με φωτεινότητα g , και $\sum_{g=0}^{255} h(g) = N$ το πλήθος όλων των pixels της

εικόνας, τότε το κανονικοποιημένο ιστόγραμμα, $H(g) = h(g) / N$, αποτελεί μια εκτίμηση της πιθανότητας εμφάνισης της φωτεινότητας με τιμή g . Τα στατιστικά στοιχεία περιγράφουν τη συνολική κατανομή των επιπέδων φωτεινότητας μιας εικόνας και υπολογίζονται από το κανονικοποιημένο gray-level ιστόγραμμα. Οι μαθηματικές εκφράσεις που χρησιμοποιούνται για τον υπολογισμό ορισμένων από αυτά τα στατιστικά στοιχεία υφής μιας εικόνας παρουσιάζονται παρακάτω:

Μέση τιμή φωτεινότητας (Average Gray Level)

$$AVG = \mu = \sum_g gH(g)$$

Τυπική απόκλιση φωτεινότητας (Standard Deviation)

$$STD = \sigma = \sqrt{\sum_g (g - \mu)^2 H(g)}$$

Λοξότητα επιπέδων φωτεινότητας (Skewness)

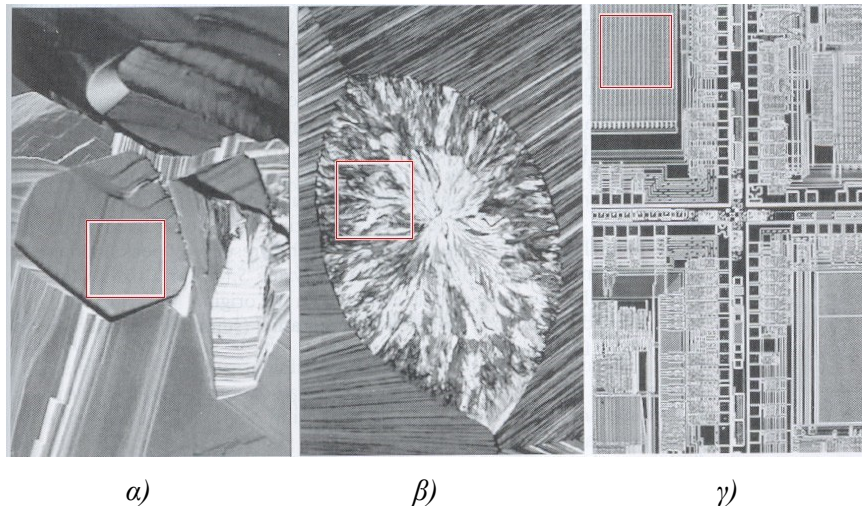
$$SKN = \frac{1}{\sigma^3} \sum_g (g - \mu)^3 H(g)$$

Κύρτωση επιπέδων φωτεινότητας (Kurtosis)

$$KRT = \left[\frac{1}{\sigma^4} \sum_g (g - \mu)^4 H(g) \right] - 3$$

Η μέση τιμή μιας εικόνας είναι η μέση τιμή των εντάσεων των pixels της. Αυτή μπορεί να υπολογιστεί είτε σε όλη την εικόνα είτε σε μέρος της. Η τυπική απόκλιση μιας εικόνας είναι ένα μέτρο της μεταβλητότητας των εντάσεων των pixels της. Η λοξότητα των επιπέδων φωτεινότητας (skewness) είναι ένα μέτρο της ασυμμετρίας των τιμών των pixels σε σχέση με τη μέση τιμή. Αν η ασυμμετρία είναι αρνητική, τότε οι εντάσεις των pixels είναι περισσότερο διαμοιρασμένες πιο αριστερά από τη μέση τιμή. Αν η ασυμμετρία είναι θετική, τότε οι εντάσεις είναι μοιρασμένες περισσότερο προς τα δεξιά της μέσης τιμής. Η κύρτωση είναι ένα μέτρο του κατά πόσο η κατανομή των εντάσεων στην εικόνα έχει κορύφωση ή είναι επίπεδη σε σχέση με μια φυσιολογική κατανομή. Η κατανομή των εντάσεων σε μια εικόνα με υψηλή κύρτωση τείνει να έχει μια κορυφή, η οποία μειώνεται σχετικά απότομα. Η κατανομή των εντάσεων μιας εικόνας με χαμηλή κύρτωση τείνει να έχει μια επίπεδη κορυφή κοντά στη μέση τιμή.

Ως παράδειγμα παρατίθενται οι παρακάτω εικόνες και ορισμένα στατιστικά στοιχεία που υπολογίστηκαν για μια περιοχή τους. Η υφή της πρώτης περιοχής είναι λεία (smooth), της δεύτερης τραχεία (coarse) και της τρίτης περιοδική (periodic). Αυτά τα αποτελέσματα είναι συμβατά με τις τιμές των στατιστικών στοιχείων που παρουσιάζονται στον Πίνακα 1. Για παράδειγμα η εντροπία, η τυπική απόκλιση και η λοξότητα στη δεύτερη εικόνα, όπου η υφή είναι τραχεία, είναι μεγαλύτερες από τις αντίστοιχες τιμές των άλλων δύο εικόνων.

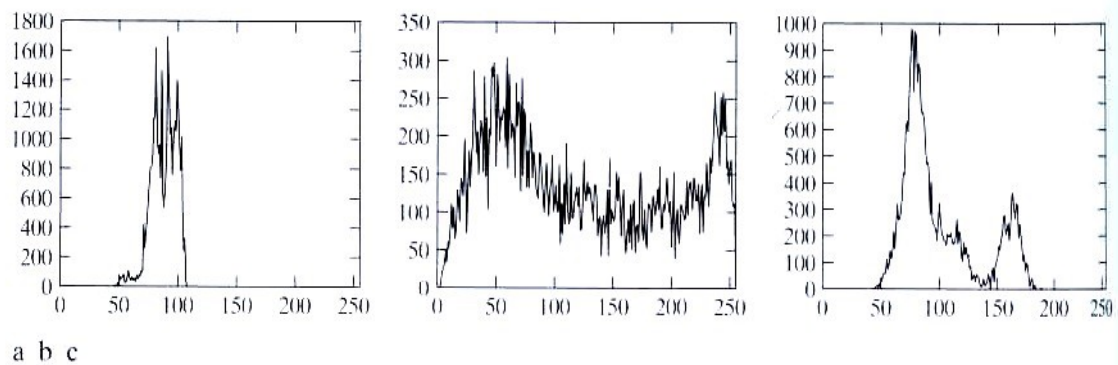


Εικόνα 27. α) Μικρογραφία ημιαγωγού (Επιλεγμένη λεία επιφάνεια). β) Μικρογραφία ανθρώπινης χοληστερόλης (Επιλεγμένη τραχεία επιφάνεια). γ) Μικρογραφία μικροεπεξεργαστή (Επιλεγμένη περιοδική περιοχή)

Πίνακας 1. Χαρακτηριστικά υφής για τις περιοχές που έχουν επιλεγθεί στην Εικόνα 27

Υφή	Μέση τιμή φωτεινότητας	Τυπική απόκλιση	Skewness	Entropy
Λεία	87,02	11.17	-0.011	5.367
Τραχεία	119.93	73.89	2.074	7.842
Περιοδική	98.48	33.50	0.557	6.517

Τα ιστογράμματα των επιλεγμένων περιοχών παρουσιάζονται στην Εικόνα 28.



Εικόνα 28. Ιστογράμματα λείας επιφάνειας (a), τραχείας επιφάνειας (b), περιοδικής επιφάνειας (c)

4.4 Ταξινόμηση προτύπων (Pattern Classification)

Η ταξινόμηση προτύπων αποτελεί αντικείμενο μελέτης για πολλούς επιστημονικούς τομείς, ενώ οι εφαρμογές της αφορούν την επεξεργασία εικόνας, την τεχνητή νοημοσύνη, την όραση υπολογιστών αλλά και τη βιολογία, την ιατρική, την οικονομία κ.λπ. Είναι ένα πρόβλημα *μηχανικής μάθησης*, δηλαδή ένα πρόβλημα που αφορά τη δημιουργία, την εκπαίδευση και τη βελτίωση ενός γνωστικού μοντέλου από ένα εκπαιδευτικό σύνολο δεδομένων. Η ταξινόμηση προτύπων συνίσταται στην εξαγωγή συμπεράσματος για ένα δεδομένο (δείγμα) που περιγράφεται από ένα σύνολο χαρακτηριστικών. Για παράδειγμα, πρόβλημα ταξινόμησης είναι η εξαγωγή συμπερασμάτων για την υγεία ενός ασθενούς όταν γνωρίζουμε διάφορα χαρακτηριστικά του (επιδημιολογικά χαρακτηριστικά, εξετάσεις κ.ά.). Στην παρούσα άσκηση, τεχνικές ταξινόμησης θα χρησιμοποιηθούν προκειμένου να κατατάξουμε κύτταρα σε υγιή ή καρκινικά, ανάλογα με ορισμένα χαρακτηριστικά τους.

Οι σημαντικότερες τεχνικές ταξινόμησης προτύπων βασίζονται σε στατιστικές μεθόδους και στα τεχνητά νευρωνικά δίκτυα. Στην παρούσα άσκηση θα μελετηθεί μονάχα μια στατιστική μέθοδος, η Ταξινόμηση Κοντινότερου Γείτονα (Nearest Neighbor).

Στη στατιστική προσέγγιση κάθε δείγμα (αντικείμενο) περιγράφεται από ένα διάνυσμα n -χαρακτηριστικών και θεωρείται ότι αποτελεί σημείο ενός n -διάστατου χώρου. Το κάθε δείγμα προς ταξινόμηση αντιστοιχίζεται σε μία από τις κατηγορίες του προβλήματος με βάση τη στατιστική ανάλυση των χαρακτηριστικών του ως προς τα χαρακτηριστικά γνωστών δειγμάτων.

Η Ταξινόμηση Κοντινότερου Γείτονα είναι μια στατιστική μέθοδος επιβλεπόμενης ταξινόμησης, δηλαδή είναι γνωστές οι κατηγορίες του προβλήματος και ένας αριθμός

δειγμάτων που ανήκουν σε αυτές. Ο αλγόριθμος της μεθόδου ταξινομεί ένα νέο άγνωστο δείγμα σε εκείνη την κατηγορία από την οποία απέχει λιγότερο, με βάση μία απόσταση.

Η μέθοδος ταξινόμησης Κοντινότερου Γείτονα περιγράφεται παρακάτω για το πρόβλημα ταξινόμησης σε δύο κατηγορίες (παρόμοια βέβαια επεκτείνεται και σε περισσότερες από δύο κατηγορίες).

Ας θεωρήσουμε δύο κατηγορίες K1, K2 και το προς ταξινόμηση δείγμα p. Το δείγμα p θα ταξινομηθεί στην κατηγορία η οποία έχει κάποιο δείγμα που να απέχει το λιγότερο δυνατό από αυτό. Μπορεί, δηλαδή, να οριστεί μια συνάρτηση απόφασης f(p) ως εξής:

$$f(p) = (\text{Μικρότερη απόσταση από K1}) - (\text{Μικρότερη απόσταση από K2})$$

Η απόσταση μεταξύ δύο διανυσμάτων x, y διάστασης n μπορεί να είναι:

- Ευκλείδεια Απόσταση: $E = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$
- Τετραγωνική Απόσταση $S = \max_i |x_i - y_i|$
- ή και κάποια άλλη, όπως απόσταση Hamming, απόσταση Mahalanobis, κ.λπ.

Ο αλγόριθμος υπολογίζει τις αποστάσεις μεταξύ του προς ταξινόμηση δείγματος με όλα τα γνωστά δείγματα και για κάθε κατηγορία κρατά τη μικρότερη. Στη συνέχεια ελέγχει την απόφαση f(p):

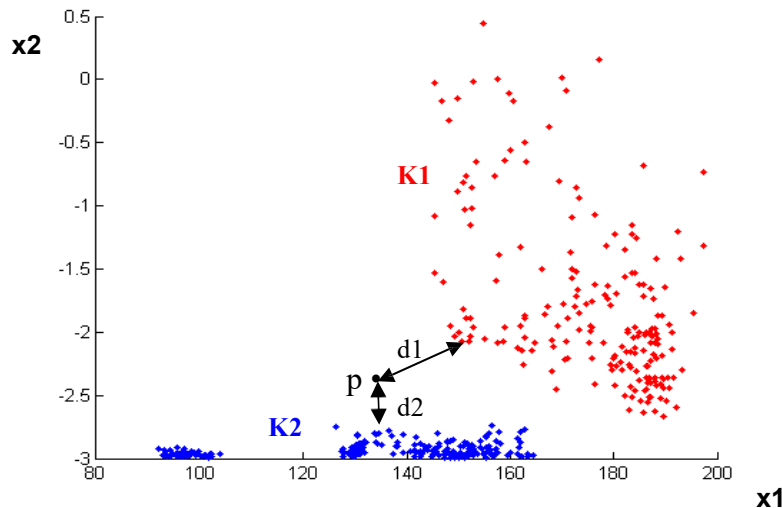
- Εάν $f(p) < 0$ τότε το p ανήκει στην κατηγορία K1
- Εάν $f(p) > 0$ τότε το p ανήκει στην κατηγορία K2

Στη συνέχεια παρουσιάζεται ένα παράδειγμα Ταξινόμησης Κοντινού Γείτονα όπου κατηγοριοποιείται μια άγνωστη περιοχή ιστού (π.χ. δέρμα) σε υγιή ή κακοήγη. Για την καλύτερη κατανόηση του προβλήματος παρουσιάζεται το 2-διάστατο πρόβλημα (n=2), δηλαδή χρησιμοποιούνται μόνο 2 χαρακτηριστικά για να περιγράψουμε τα δείγματα.

Ας θεωρήσουμε λοιπόν ότι έχουμε ένα πρόβλημα ταξινόμησης ιστού σε 2 κατηγορίες, την κατηγορία K1 που αφορά κακοήθεια και την κατηγορία K2 που αφορά υγιή ιστό. Για να μπορέσουμε να ταξινομήσουμε έναν άγνωστο ιστό σε μια από τις 2 κατηγορίες με τη μέθοδο του Κοντινότερου Γείτονα, καταρχάς θα πρέπει να έχουμε ένα πλήθος από γνωστά δείγματα (ιστούς που γνωρίζουμε την κατάστασή τους). Ας υποθέσουμε ότι έχουμε 200 δείγματα από κακοήθεις ιστούς και 200 δείγματα από υγιείς ιστούς. Στη συνέχεια πρέπει να αναλύσουμε τα γνωστά αυτά δείγματα ούτως ώστε να εξαγάγουμε διάφορα χαρακτηριστικά τους, τα οποία μπορούν να χρησιμοποιηθούν για τη διάκριση κακοήθης και υγιούς ιστού. Στο υποτιθέμενο παράδειγμα, έστω ότι βρήκαμε μετά από ανάλυση πως ο κακοήθης ιστός διαφέρει από τον υγιή ως προς τη μέση τιμή φωτεινότητας και την κυρτότητα των επιπέδων φωτεινότητας, και επομένως για κάθε ένα γνωστό δείγμα υπολογίζουμε τις τιμές των δύο αυτών χαρακτηριστικών. Οι τιμές των χαρακτηριστικών όλων των γνωστών δειγμάτων (400) συγκεντρώνονται σε μια βάση δεδομένων, η οποία αποτελεί τη βάση γνώσης του ταξινομητή.

Στην Εικόνα 29 φαίνεται η ταξινόμηση του άγνωστου δείγματος p σε μία από τις κατηγορίες σύμφωνα με τη μέθοδο του Κοντινότερου Γείτονα. Ο άξονας x1 είναι η μέση τιμή φωτεινότητας μιας περιοχής και ο x2 η κυρτότητα των επιπέδων φωτεινότητάς της. Οι μπλε κουκίδες (200) αφορούν τις περιοχές υγιούς ιστού και οι κόκκινες (200) τις περιοχές κακοήθειας που έχουμε στη βάση δεδομένων που έχουμε δημιουργήσει (βάση γνώσης). Κάθε γνωστή περιοχή (από τη βάση γνώσης) χαρτογραφείται πάνω στο παρακάτω

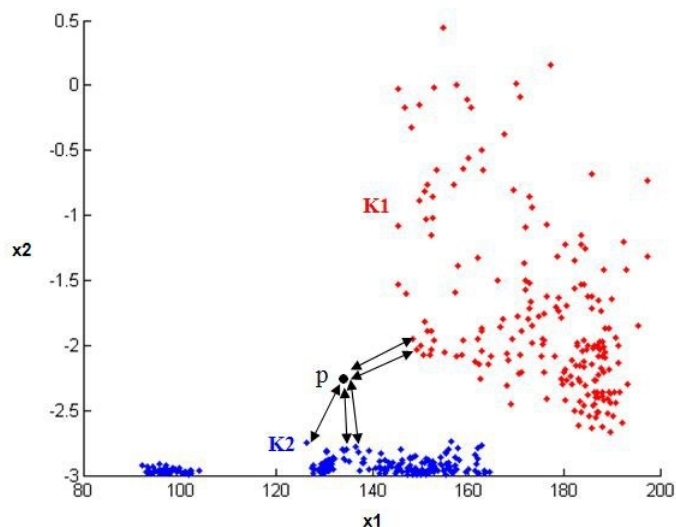
διάγραμμα σύμφωνα με τις (x_1, x_2) μετρήσεις της (*scatter διάγραμμα*). Έστω ότι έχουμε μια καινούρια περιοχή ιστού που δεν την έχουμε κατηγοριοποιήσει και θέλουμε να δούμε πού ανήκει (δείγμα p). Μετρώντας τη μέση τιμή φωτεινότητας της περιοχής και την κυρτότητα των επιπέδων φωτεινότητάς της, τη χαρτογραφούμε πάνω στο scatter διάγραμμα. Με τον αλγόριθμο Κοντινού Γείτονα το σύστημα ταξινόμησης κατηγοριοποιεί την άγνωστη περιοχή (Εικόνα 19).



Εικόνα 29. Ταξινόμηση Κοντινότερου Γείτονα. Το πρότυπο p ταξινομείται στην κατηγορία K2, γιατί ισχύει $d_2 < d_1$

Στην πράξη το πρόβλημα δεν είναι τόσο απλό, καθώς οι κατηγορίες K1 και K2 δεν είναι τόσο ξεκάθαρα διαχωρίσιμες, όπως προϋποθέτει το παραπάνω κριτήριο διαχωρισμού. Σε πολλές περιπτώσεις ένα δείγμα μπορεί να ανήκει σε μια κατηγορία αλλά να βρίσκεται πλησιέστερα σε μια άλλη. Προκειμένου να αποφευχθεί η δυσκολία αυτή, μετράται η απόσταση του p από πολλά δείγματα κάθε κατηγορίας, έτσι ώστε η επίδραση οποιουδήποτε διφορούμενου προτύπου να εξομαλυνθεί. Αυτός ο τρόπος ταξινόμησης ονομάζεται ταξινόμηση k -κοντινών γειτόνων (*k-nearest neighbor* ή *k-nn*), όπου k είναι το πλήθος των γειτονικών προτύπων, ως προς τα οποία μετράται η απόσταση και με βάση τα οποία γίνεται η ταξινόμηση του νέου προτύπου.

Στην Εικόνα 30 παρουσιάζεται η διαδικασία ταξινόμησης του δείγματος p του προηγούμενου παραδείγματος, με τη μέθοδο των k κοντινών γειτόνων, για $k=5$. Όπως βλέπουμε υπολογίζονται οι αποστάσεις του άγνωστου δείγματος από τα 5 πιο κοντινά γνωστά δείγματα. Τα 3 πιο κοντινά ανήκουν στην κατηγορία K2 (υγιής ιστός) ενώ τα άλλα 2 στην κατηγορία K1 (κακοήθεια). Επομένως το δείγμα ταξινομείται στην κατηγορία K2.



Εικόνα 30. Ταξινόμηση με $k=5$ κοντινούς γείτονες. Οι 3 κοντινότεροι γείτονες ανήκουν στην κατηγορία K2, επομένως το δείγμα p ταξινομείται σε αυτήν

5. Εργαστηριακό μέρος

5.1 Άνοιγμα και επεξεργασία αρχείου DICOM

Αρχικά, πρέπει να ανεβάσετε το υλικό που παρέχεται μαζί με την άσκηση στο Google Colab, ή σε οποιαδήποτε άλλη υπηρεσία επιτρέπει την επεξεργασία jupyter notebooks. Πιο συγκεκριμένα, είναι απαραίτητα τα αρχεία:

- ctdicom.dcm
- cytimage.jpg
- groups_train.mat

Εάν χρησιμοποιείτε Jupyter Notebook τοπικά στον υπολογιστή σας, θα πρέπει απλώς να πλοηγηθείτε στον φάκελο ο οποίος περιέχει τα αρχεία αυτά, ώστε να μπορείτε να τα χρησιμοποιήσετε.

Εάν χρησιμοποιείτε Google Colab, είναι απαραίτητο να εγκαταστήσετε το module pydicom με την εντολή:

```
!pip install pydicom
```

Εάν χρησιμοποιείτε Jupyter Notebook πρέπει **να μη χρησιμοποιήσετε** την παραπάνω εντολή. Θα πρέπει να έχετε πραγματοποιήσει την εγκατάσταση όλων των modules που θα χρειαστείτε από το terminal, σύμφωνα με τις οδηγίες που θα βρείτε στο mycourses.

Ας κάνουμε import το pydicom και ας διαβάσουμε το αρχείο ctdicom.dcm.

```
import pydicom

ds = pydicom.dcmread('ctdicom.dcm')
print(ds)
```

Βλέπουμε, λοιπόν, πως το αρχείο αυτό περιέχει διάφορες πληροφορίες όπως το όνομα του ασθενή, το φύλο, την ηλικία, το όνομα του νοσοκομείου, κ.ά.

Μπορούμε να χρησιμοποιήσουμε αυτές τις πληροφορίες, χρησιμοποιώντας την αντίστοιχη λέξη κλειδί. Για παράδειγμα, εάν θέλουμε να τυπώσουμε την ηλικία του ασθενή:

```
print(ds.PatientAge)
>>> 073Y
```

Παρατηρούμε πως το αρχείο dicom που τυπώσαμε προηγουμένως, περιέχει την εξής πληροφορία:

(7fe0, 0010) Pixel Data OW: Array of 262144 elements

Αυτό είναι το πεδίο το οποίο περιέχει την εικόνα μας. Ωστόσο, εάν προσπαθήσουμε να φορτώσουμε την εικόνα από εκεί, θα παρατηρήσουμε ότι μας επιστρέφει πληροφορία τύπου "Bytes", την οποία δεν μπορούμε να επεξεργαστούμε απευθείας.

```
img = ds.PixelData
print(type(img))
```

```
>>> <class 'bytes'>
```

Ωστόσο, pydicom μας προσφέρει έναν έτοιμο τρόπο για τη φόρτωση της εικόνας σε μορφή που να μπορούμε να επεξεργαστούμε.

```
img = ds.pixel_array
print(type(img))
>>> <class 'numpy.ndarray'>
```

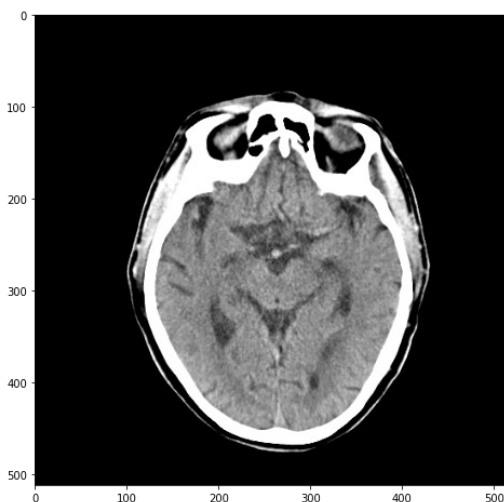
Ας εμφανίσουμε την εικόνα που φορτώσαμε, για να δούμε τί περιέχει.

```
print("Image dimensions: {}\nImage type: {}".format(img.shape,
img.dtype))
```

```
import matplotlib.pyplot as plt
```

```
plt.figure(figsize=(8, 8))
plt.imshow(img, cmap='gray')
plt.show()
```

```
>>> Image dimensions: (512, 512)
>>> Image type: uint8
```



Βλέπουμε πως πρόκειται για μια εικόνα μεγέθους 512x512, τύπου uint8. Αυτό σημαίνει πως κάθε pixel αποτελείται από έναν unsigned integer (δηλαδή έναν ακέραιο χωρίς πρόσημο), μεγέθους 8 bit. $2^8 = 256$, άρα περιμένουμε πως κάθε pixel θα κυμαίνεται μεταξύ των τιμών [0, 255].

Στη συνέχεια, θα πραγματοποιήσουμε κατωφλίωση της εικόνας. Αυτό μπορεί να είναι επιθυμητό εάν θέλουμε, για παράδειγμα, να απομονώσουμε τη δομή των οστών.

Το μόνο που έχουμε να κάνουμε είναι να θέσουμε ίσα με το 0, όλα τα pixels τα οποία δεν ανήκουν μεταξύ δύο ακραίων τιμών. Σε αυτή την περίπτωση, ως ακραίες τιμές επιλέγουμε το 250 και το 255.

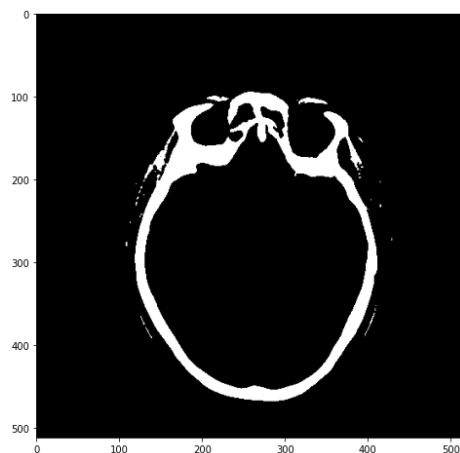
```
import numpy as np
```

```
new_img = np.clip(img, 250, 255)
```

```
plt.figure(figsize=(8, 8))
plt.imshow(new_img, cmap='gray')
```

Εμφάνισε την εικόνα που μόλις διαβάστηκε

```
plt.show()
```

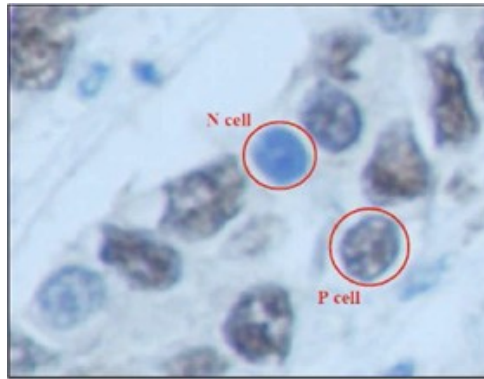


5.2 Ανάλυση και επεξεργασία κυτταρολογικών εικόνων για τον αυτόματο εντοπισμό καρκινικών κυττάρων

Στην εφαρμογή αυτή παρουσιάζεται ένα αλγόριθμος για τον εντοπισμό καρκινικών κυττάρων σε μια κυτταρολογική εικόνα μικροσκοπίου.

Για την ακριβή διάγνωση οποιουδήποτε καρκίνου απαιτείται κυτταρολογική ή ιστολογική επιβεβαίωση με λήψη δείγματος (π.χ. βιοψία) από την ύποπτη περιοχή. Ένα από τα αντικείμενα της κυτταρολογίας είναι ο χαρακτηρισμός και η μέτρηση των κυττάρων και των κυτταρικών συστατικών. Μπορεί οι μετρήσεις αυτές να αφορούν φυσικές (π.χ. μήκος, όγκο, χρώμα) ή βιοχημικές ιδιότητες (π.χ. πρωτεϊνικό ή λιπιδικό περιεχόμενο) ή συνδυασμό αυτών των δύο (π.χ. κατανομή κάποιων κυτταρικών περιεχομένων). Η απεικονιστική κυτταρολογία έχει ως αντικείμενο τη μέτρηση των ιδιοτήτων του κυττάρου με χρήση εικόνων λαμβανομένων από μικροσκόπιο. Οι εικόνες μπορούν να αναλυθούν οπτικά, δηλαδή μετρώντας το μέγεθος των κυττάρων ή μετρώντας τον αριθμό των χρωματισμένων κηλίδων σε κάθε κύτταρο. Το βασικό μειονέκτημα αυτής της μεθόδου είναι η ανιαρή διαδικασία της μέτρησης και της επισήμανσης, ενώ η μη αυτοματοποιημένη διαδικασία είναι πολύ επιρρεπής σε σφάλματα και στα προβλήματα που δημιουργούνται από έναν υποκειμενικό παρατηρητή. Είναι δηλαδή πολύ πιθανό, αν η μη αυτοματοποιημένη ανάλυση γίνει δεύτερη φορά από το ίδιο ή διαφορετικό πρόσωπο, τα αποτελέσματα που θα προκύψουν να είναι εντελώς διαφορετικά.

Οι εικόνες που χρησιμοποιούνται για τη συγκεκριμένη εφαρμογή προέρχονται από μια ανοσοϊστολογική εξέταση για καρκίνο του πνεύμονα, κατά την οποία πραγματοποιήθηκε έγχυση με συγκεκριμένη χημική ουσία σε δείγμα ιστού από βιοψία. Χρησιμοποιώντας κάποιες χημικές ουσίες, είναι δυνατό να δοθεί συγκεκριμένο χρώμα σε βιολογικά συστατικά. Με αυτόν τον τρόπο δίνεται η ευκαιρία να μελετηθούν όχι μόνο ο αριθμός, το σχήμα και το μέγεθος των κυττάρων, αλλά και τμήματά τους και η μεταφορά κάποιων ουσιών διαμέσου των κυττάρων. Στην υπό μελέτη περίπτωση, το αποτέλεσμα της έγχυσης είναι ο χρωματισμός των καρκινικών κυττάρων με καφέ χρώμα. Η παρακάτω παρουσιάζει ένα παράδειγμα από χρωματισμένα κύτταρα μεγεθυμένα κατά 40 φορές. Τα υγιή κύτταρα (αρνητικά (N)) απεικονίζονται σε αποχρώσεις του μπλε και τα καρκινικά (θετικά (P)) απεικονίζονται σε αποχρώσεις του καφέ.



Το πλήθος των θετικών κυττάρων σε αναλογία με το συνολικό πλήθος των κυττάρων της εικόνας (ή αναλογία θετικών κυττάρων προς αρνητικά κύτταρα) είναι μια ιδιαίτερα σημαντική παράμετρος, καθώς καθορίζει την ύπαρξη καρκίνου, το στάδιο της καρκινογένεσης, τον τύπο του καρκίνου κ.ά. και χρησιμοποιείται από το γιατρό τόσο για τη διάγνωση όσο και για το σχεδιασμό κατάλληλης θεραπείας.

Η αξιολόγηση κυτταρολογικών/ιστολογικών εικόνων είναι μια εργασία που απαιτεί κόπο και μπορούν να την αναλάβουν μόνο πολύ καλά εκπαιδευμένοι κυτταρολόγοι/παθολογοανατόμοι. Καθώς η ερμηνεία τους επηρεάζεται από τον υποκειμενικό παράγοντα και είναι αντικείμενο διαγνωστικού λάθους, είναι προφανές ότι υπάρχει τεράστιο ενδιαφέρον για την ανάπτυξη μιας αυτοματοποιημένης διαδικασίας ανάλυσης των εικόνων αυτών. Τα τελευταία χρόνια έχουν παρουσιαστεί διάφορα συστήματα ψηφιακής επεξεργασίας ιστολογικών/κυτταρολογικών εικόνων με σκοπό την ανάλυση και επεξεργασία των εικόνων αυτών και την εξαγωγή πληροφορίας που μπορεί να βοηθήσει τον κυτταρολόγο στη λήψη απόφασης. Τέτοιου είδους συστήματα τα οποία παρέχουν υποστήριξη στη λήψη απόφασης για τη διάγνωση ή την επιλογή σχημάτων παρακολούθησης και θεραπείας, ονομάζονται Συστήματα Υποστήριξης Διάγνωσης (Computer Aided Diagnosis) ή Συστήματα Υποστήριξης Λήψης Κλινικής Απόφασης (Clinical Decision Support Systems).

Στη συνέχεια παρουσιάζεται ένα σύστημα υποστήριξης διάγνωσης, που βασίζεται σε αλγορίθμους ψηφιακής επεξεργασίας εικόνας, για τον αυτόματο εντοπισμό και την κατάλληλη πλαισίωση των καρκινικών κυττάρων στην ιστολογική εικόνα. Σκοπός μας είναι να βοηθήσουμε το γιατρό ώστε να καταλήξει σε μια πιο γρήγορη, και κυρίως, πιο ακριβή μέτρηση, επισημαίνοντας ακόμα και καρκινικά κύτταρα τα οποία δεν είναι ευδιάκριτα και μπορούν να διαφύγουν από την παρατήρησή του.

5.2.1 Χρωματική ανάλυση και επεξεργασία

Ας φορτώσουμε την κυτταρολογική εικόνα της άσκησης.

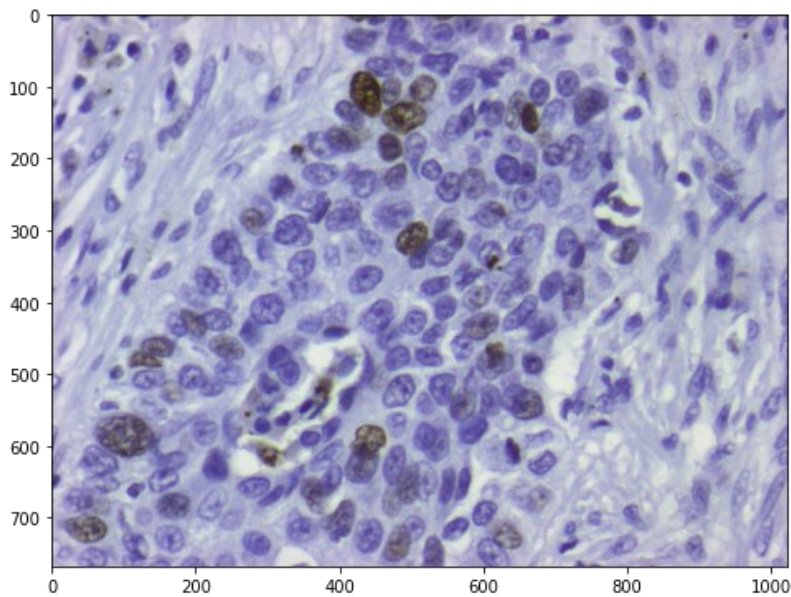
```
import skimage.io
```

```
img = skimage.io.imread('cytimage.jpg')
```

```
plt.figure(figsize=(8, 8))
```

```
plt.imshow(img)
```

```
plt.show()
```



Παρατηρώντας την εικόνα αυτή μπορούμε εύκολα να διαπιστώσουμε πως υπάρχουν αρκετά θετικά κύτταρα τα οποία αναγνωρίζονται εύκολα (έντονα καφέ κύτταρα). Ωστόσο, υπάρχουν και πολλά κύτταρα τα οποία δε γίνεται να κατηγοριοποιηθούν με την πρώτη ματιά ως θετικά (κύτταρα με ανοιχτό καφέ χρώμα), και ορίζονται ως ύποπτα κύτταρα. Τα τελευταία πρέπει να μελετηθούν με ιδιαίτερη προσοχή από τον κυτταρολόγο, ούτως ώστε να καταταχθούν σωστά ως καρκινικά ή μη. Επιπλέον, καθώς το πλήθος των κυττάρων στην παραπάνω εικόνα είναι μεγάλο, υπάρχει πάντα η πιθανότητα ορισμένα ύποπτα κύτταρα, τα οποία είναι θετικά, να διαφύγουν από την παρατήρηση του κυτταρολόγου. Είναι αντιληπτό πως τόσο η αξιολόγηση των ύποπτων κυττάρων όσο και η συνολική μέτρηση των θετικών κυττάρων είναι μια ιδιαίτερα χρονοβόρα διαδικασία και απαιτεί πολύ κόπο και μεγάλη εμπειρία.

Αρχικά, η διαθέσιμη πληροφορία που έχουμε είναι η διαφορά στο χρώμα που παρουσιάζουν τα θετικά με τα αρνητικά κύτταρα. Έτσι, ως πρώτο βήμα αναλύουμε την εικόνα σε χρωματικό επίπεδο, ώστε να ελέγξουμε ποιες χρωματικές συνιστώσες προσφέρουν σημαντική πληροφορία.

Κατά τη χρωματική ανάλυση, αρχικώς απεικονίζουμε τις συνιστώσες Red, Green, Blue του χώρου RGB ξεχωριστά, ούτως ώστε να ελέγξουμε σε ποιες από αυτές τα καρκινικά κύτταρα διακρίνονται καλύτερα. Στην πράξη απεικονίζουμε ξεχωριστά τις τρεις grayscale εικόνες R, G και B οι οποίες συνθέτουν την αρχική εικόνα.

Παράδειγμα: Προκειμένου να απεικονίσουμε και να μελετήσουμε τη συνιστώσα R μιας εικόνας I, την αποσπάμε από αυτήν ορίζοντας ως εξής: `I[:, :, 0]`

```
print(img.shape)

labels = ['Red', 'Green', 'Blue']

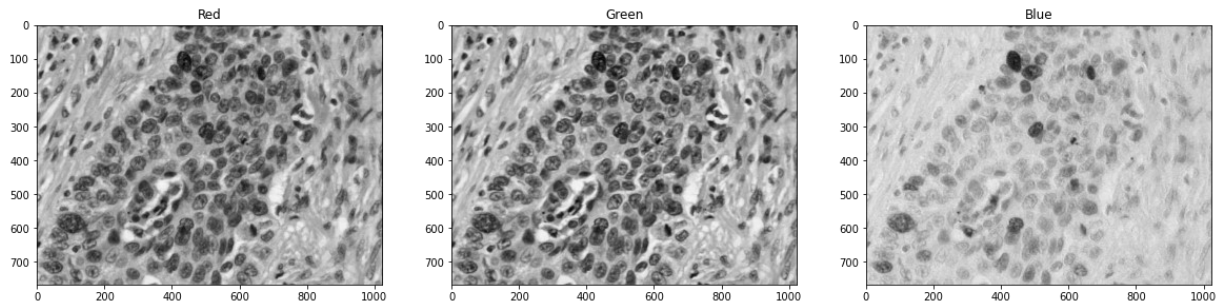
plt.figure(figsize=(20,20))

for i in range(3):
    plt.subplot(1,3,i+1)
    plt.title(labels[i])

    plt.imshow(img[:, :, i], cmap='gray')

plt.show()
```

```
>>> (768, 1024, 3)
```



Μελετώντας τις παραπάνω νέες εικόνες, διαπιστώνουμε πως η συνιστώσα η οποία παρέχει την πιο χρήσιμη πληροφορία για το διαχωρισμό θετικών και αρνητικών κυττάρων είναι η συνιστώσα B του χώρου RGB. Η διαπίστωση αυτή προήλθε αναλύοντας τη φωτεινότητα που παρουσιάζουν τα καρκινικά και τα υγιή κύτταρα στις παραπάνω εικόνες. Κατά την ανάλυση των εικόνων αυτών προέκυψε πως τα καρκινικά κύτταρα παρουσιάζουν αρκετά χαμηλότερη φωτεινότητα σε σχέση με τα υγιή στην εικόνα B. Πρακτικά αυτό σημαίνει πως τα καρκινικά κύτταρα περιέχουν λιγότερο μπλε χρώμα (κάτι το οποίο το γνωρίζαμε εκ των προτέρων). Παρατηρώντας την αρχική εικόνα και την εικόνα B διαπιστώνουμε πως στην τελευταία διακρίνονται καλύτερα οι περιοχές ενδιαφέροντος. Το γεγονός αυτό μας οδηγεί στην περαιτέρω επεξεργασία της εικόνας B προκειμένου να οδηγηθούμε σε μια εικόνα στην οποία τα καρκινικά κύτταρα θα απεικονίζονται ευκρινέστερα.

Μια διαδικασία που μπορούμε να εφαρμόσουμε προκειμένου να αυξήσουμε την ευκρίνεια των καρκινικών κυττάρων, είναι η κατάλληλη αύξηση της αντίθεσης της εικόνας B, τέτοια ώστε τα καρκινικά κύτταρα να αναδειχθούν σε σχέση με τα υγιή κύτταρα και τα υπόλοιπα στοιχεία της εικόνας.

Η διαδικασία της προσαρμογής της αντίθεσης περιλαμβάνει δύο ενέργειες.

1. Την επιλογή των δύο ακραίων τιμών (thresholds). Όλα τα pixels με τιμή κάτω του κατώτατου ορίου θα τεθούν ίσα με το 0, ενώ αντίστοιχα όλα τα pixels με τιμή πάνω του ανωτάτου ορίου θα τεθούν ίσα με το 255.
2. Τη γραμμική απεικόνιση των υπόλοιπων pixels μεταξύ του 0 και 255.

Στο παρακάτω κομμάτι κώδικα η συνάρτηση της κατωφλίωσης δίνεται έτοιμη.

Μετά από διάφορες δοκιμές, καταλήξαμε πως τα καρκινικά κύτταρα αναδεικνύονται ικανοποιητικά με την εφαρμογή διαδικασίας προσαρμογής της αντίθεσης της εικόνας B στο εύρος φωτεινότητων [128, 165].

```
def imadjust(img, lower, upper):
    out = (img.astype(float) - lower) * (255. / (upper - lower))
    return out.clip(0, 255).round().astype(np.uint8)

blue_ch = img[:, :, 2]

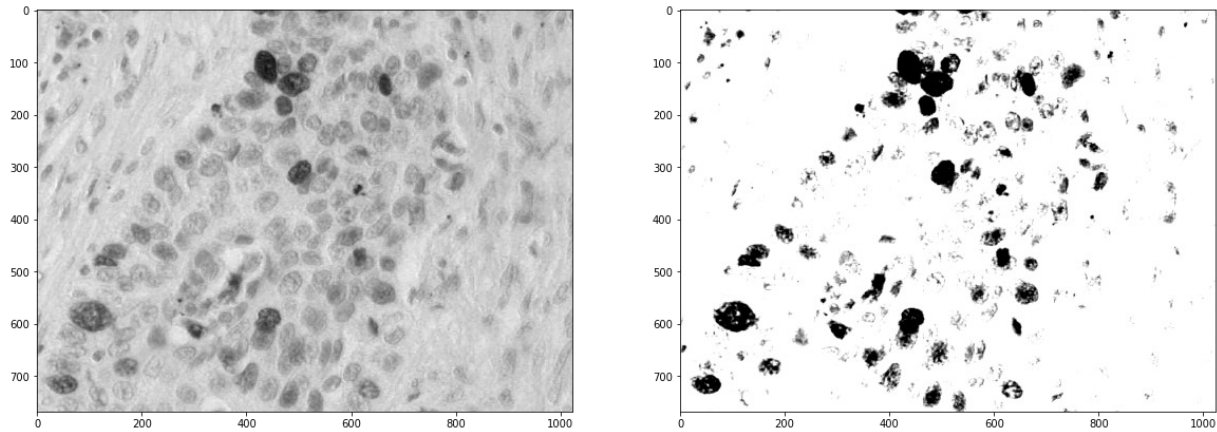
plt.figure(figsize=(20, 20))

plt.subplot(1, 2, 1)
plt.imshow(blue_ch, cmap='gray')
```

```
blue_ch_adj = imadjust(blue_ch, 128, 165)

plt.subplot(1,2,2)
plt.imshow(blue_ch_adj, cmap='gray')

plt.show()
```



Η παραπάνω εικόνα, αν και αναδεικνύει τα καρκινικά κύτταρα, δεν είναι άμεσα χρήσιμη στο γιατρό καθώς έχουν χαθεί άλλες χρήσιμες πληροφορίες, όπως τα υγιή κύτταρα ή τα κύτταρα που συνορεύουν ή επικαλύπτονται με τα καρκινικά.

Προκειμένου λοιπόν να δημιουργήσουμε μια εικόνα η οποία θα είναι χρήσιμη για το γιατρό και ταυτόχρονα θα αναδεικνύει τα καρκινικά κύτταρα, μια τεχνική που μπορούμε να εφαρμόσουμε είναι να χρησιμοποιήσουμε την επεξεργασμένη εικόνα `blue_ch_adj` ως τη συνιστώσα B της έγχρωμης εικόνας RGB (δηλαδή να αντικαταστήσουμε τη B (ή `blue_ch`) με τη `blue_ch_adj`). Για να το κάνουμε αυτό πρέπει να ανακατασκευάσουμε την εικόνα RGB, χρησιμοποιώντας τις εικόνες R, G, και `blue_ch_adj`.

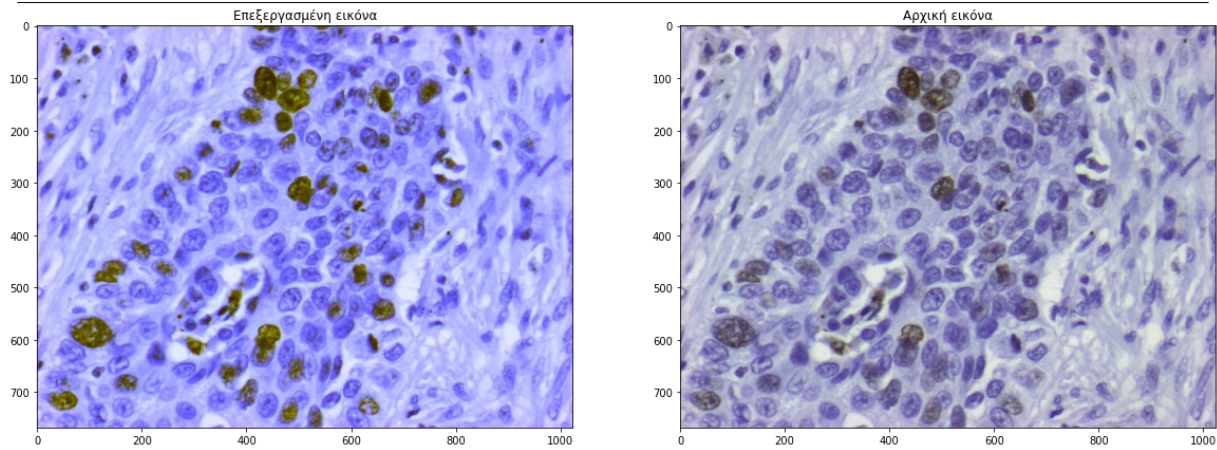
```
new_img = np.dstack((img[:, :, 0], img[:, :, 1], blue_ch_adj))

plt.figure(figsize=(20,20))

plt.subplot(1,2,1)
plt.title("Επεξεργασμένη εικόνα")
plt.imshow(new_img)

plt.subplot(1,2,2)
plt.title("Αρχική εικόνα")
plt.imshow(img)

plt.show()
```

Συγκρίνοντας το αποτέλεσμα με την αρχική εικόνα, καθίσταται προφανές πως με τη διαδικασία που ακολουθήσαμε καταφέραμε να αυξήσουμε την ευκρίνεια των καρκινικών κυττάρων και επομένως να τα αναδείξουμε.

Με τη διαδικασία αυτή, αν και τα θετικά κύτταρα έχουν αναδειχθεί, το χρωματικό περιεχόμενο της εικόνας έχει μεταβληθεί, γεγονός το οποίο πολλές φορές δεν είναι επιθυμητό από τους γιατρούς. Για το λόγο αυτό εφαρμόζουμε στη συνέχεια ορισμένες πιο σύνθετες τεχνικές προκειμένου να ανιχνεύσουμε και αναδείξουμε τα θετικά κύτταρα, και ταυτόχρονα να παραμείνουν τα χρώματα και η ποιότητα της αρχικής εικόνας αμετάβλητα.

Περαιτέρω μελέτη της εικόνας B ανέδειξε πως σε αυτήν τα θετικά κύτταρα διαφέρουν και ως προς την υφή τους σε σχέση με τα αρνητικά. Αναλύοντας όπως και πριν τις τιμές φωτεινότητας των καρκινικών κυττάρων, διαπιστώνουμε πως εκτός του ότι παρουσιάζουν χαμηλότερη φωτεινότητα σε σχέση με τα υγιή, παρουσιάζουν και μεγαλύτερη ανομοιογένεια στις τιμές φωτεινότητάς τους. Η διαπίστωση αυτή μας οδήγησε στο να εφαρμόσουμε τεχνικές ανάλυσης της υφής στην εικόνα B.

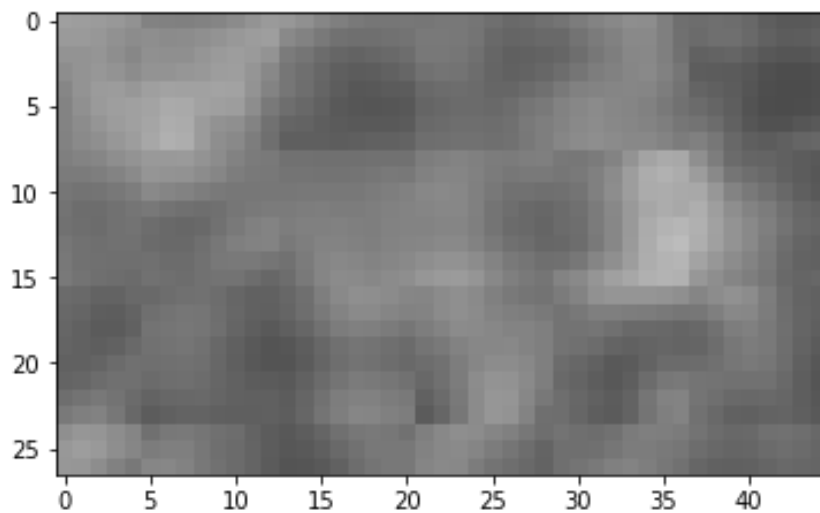
5.2.2 Ανάλυση υφής

Όπως αναφέρθηκε και στο θεωρητικό μέρος, για την ανάλυση υφής χρησιμοποιούμε την πιο απλή τεχνική, δηλαδή τα στατιστικά στοιχεία υφής.

Κατά την ανάλυση υφής, αρχικά υπολογίζουμε ορισμένα στατιστικά στοιχεία υφής για διάφορες περιοχές ενδιαφέροντος (Region Of Interest – ROI) της εικόνας. Αρχικά λοιπόν πρέπει να αποκόψουμε αρκετές περιοχές ενδιαφέροντος από την εικόνα που μελετούμε (B) προκειμένου να τις αναλύσουμε.

```
cropped_img = blue_ch[568:595, 75:120]
```

```
plt.imshow(cropped_img, cmap='gray', vmin=0, vmax=255)
plt.show()
```

Για το ROI που επιλέξαμε, υπολογίζουμε 4 στατιστικά στοιχεία:

1. Μέση τιμή φωτεινότητας (mean grey level)
2. Τυπική απόκλιση των επιπέδων φωτεινότητας (standard deviation)
3. Λοξότητα των επιπέδων φωτεινότητας (skewness)
4. Κύρτωση των επιπέδων φωτεινότητας (kurtosis)

```
from scipy.stats import skew, entropy, kurtosis

avg = np.mean(cropped_img)
std = np.std(cropped_img)
skn = skew(cropped_img, axis=None)
krt = kurtosis(cropped_img, axis=None)

print('Average gray level: {}\nStandard deviation:
{}\nSkewness: {}\nKurtosis: {}'.format(avg, std, skn, krt))

>>> Average gray level: 119.99259259259259
>>> Standard deviation: 20.359544598085353
>>> Skewness: 0.525375880419298
>>> Kurtosis: 0.06853685809741927
```

5.2.3 Ταξινόμηση περιοχών βάσει της υφής, με χρήση μηχανικής μάθησης

Για τη συγκεκριμένη εφαρμογή αποφασίσαμε να κατασκευάσουμε έναν από τους πιο απλούς ταξινομητές, τον ταξινομητή k κοντινών γειτόνων.

Όπως αναφέρθηκε και στο θεωρητικό μέρος, η πρώτη φάση για την κατασκευή ενός ταξινομητή είναι η συγκέντρωση και ανάλυση γνωστών δεδομένων τα οποία θα χρησιμοποιηθούν για την εκπαίδευσή του. Ο ταξινομητής k κοντινών γειτόνων δεν ακολουθεί κάποιο ειδικό βήμα εκπαίδευσης, αλλά χρησιμοποιεί τα γνωστά δείγματα προκειμένου να συγκρίνει το άγνωστο δείγμα με αυτά και επομένως να το κατατάξει σε κάποια κατηγορία.

Σε αυτή τη φάση της κατασκευής του συστήματος, οι γιατροί μάς τροφοδότησαν με διάφορες κυτταρολογικές εικόνες, στις οποίες φαίνονται ξεκάθαρα θετικά και αρνητικά κύτταρα. Από αυτές τις εικόνες, και υπό την καθοδήγηση των γιατρών, αποκόπηκαν 20 επιβεβαιωμένα καρκινικά κύτταρα, 20 υγιή κύτταρα και 20 περιοχές που αντιστοιχούν στο φόντο της εικόνας (background). Το επόμενο βήμα είναι η εξαγωγή χαρακτηριστικών των γνωστών δειγμάτων προκειμένου να κατασκευαστεί η βάση γνώσης του συστήματος. Όπως είδαμε και προηγουμένως, 4 χαρακτηριστικά υφής εξάγονται από τις υπό διερεύνηση περιοχές.

Συνοψίζοντας, για τη βάση γνώσης του ταξινομητή έχουμε συγκεντρώσει 240 μετρήσεις που προκύπτουν από 4 χαρακτηριστικά υφής για τη συνιστώσα B, 60 γνωστών δειγμάτων, τα οποία χωρίζονται σε 3 κατηγορίες (θετικά, αρνητικά, φόντο).

Ας φορτώσουμε τις μετρήσεις αυτές.

```
import scipy.io

mat = scipy.io.loadmat('groups_train.mat')
```

Στη συνέχεια, εκπαιδεύουμε τον ταξινομητή με τη χρήση των δεδομένων μας.

```
from sklearn.neighbors import KNeighborsClassifier

neigh = KNeighborsClassifier(n_neighbors=3)
neigh.fit(mat["TRAIN_B"], mat["GROUP"].ravel())
```

```
>>> KNeighborsClassifier(algorithm='auto', leaf_size=30,
metric='minkowski', metric_params=None, n_jobs=None,
n_neighbors=3, p=2, weights='uniform')
```

Πώς μπορούμε να χρησιμοποιήσουμε τον παραπάνω, εκπαιδευμένο πλέον, ταξινομητή για την ταξινόμηση κάθε περιοχής της εικόνας;

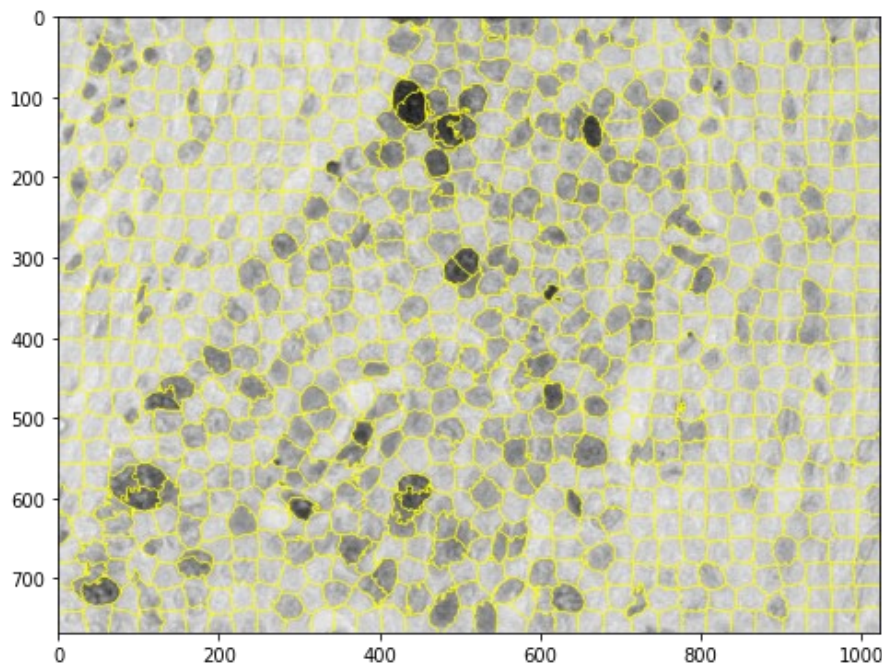
Ένας απλός, αλλά και αργός τρόπος, θα ήταν να ταξινομήσουμε κάθε pixel ξεχωριστά. Δηλαδή, με τη χρήση ενός κυλίομενου παραθύρου, σε κάθε pixel της εικόνας να υπολογίζουμε τα στατιστικά χαρακτηριστικά μιας μικρής περιοχής (πχ 11x11 pixels), στο κέντρο της οποίας βρίσκεται το pixel υπό διερεύνηση. Όμως με τη χρήση αυτού του τρόπου θα πρέπει να διερευνηθούν $768 \times 1024 = 786432$ περιοχές.

Ένας άλλος τρόπος, είναι με χρήση superpixels. Πρόκειται για έναν clustering αλγόριθμο (αλγόριθμος ομαδοποίησης), ο οποίος τμηματοποιεί την εικόνα σε έναν προεπιλεγμένο αριθμό περιοχών. Με αυτόν τον τρόπο, αντί να ελέγχουμε κάθε pixel ξεχωριστά, μπορούμε να ελέγχουμε κάθε superpixel, τα οποία ούτως ή άλλως περιέχουν γειτονικά pixels που εμφανίζουν παρόμοια χαρακτηριστικά. Επιλέγοντας τη δημιουργία 800 superpixels, επιταχύνουμε τον αλγόριθμο εξετάζοντας μόνο 800 περιοχές, αντί για 786432.

```
from skimage.segmentation import slic, mark_boundaries

segm = slic(blue_ch, n_segments=800, compactness=0.2)

plt.figure(figsize=(8,8))
plt.imshow(mark_boundaries(blue_ch, segm))
plt.show()
```



Έχοντας τμηματοποιήσει την εικόνα, ας ταξινομήσουμε κάθε superpixel ξεχωριστά.

Καθώς δεν εφαρμόζουμε κάποιο γραμμικό φίλτρο (ούτε συνέλιξη ή συσχέτιση), ο αλγόριθμος που θα δώσει την τιμή του κάθε pixel στην εικόνα εξόδου καθορίζεται ανάλογα με το τελικό αποτέλεσμα που επιθυμούμε να δημιουργήσουμε. Στη συγκεκριμένη εφαρμογή μας ενδιαφέρει ο εντοπισμός των περιοχών της εικόνας που αντιστοιχούν σε θετικά κύτταρα. Για το σκοπό αυτό, ορίζουμε τον εξής αλγόριθμο:

- Εάν ένα superpixel ταξινομηθεί στην κατηγορία ‘1’ (θετικά κύτταρα), τότε όλα τα pixels που περιέχονται στο superpixel αυτό, θα λάβουν τιμή 1.
- Αντίστοιχα, εάν ένα superpixel ταξινομηθεί στην κατηγορία ‘2’, τότε όλα τα pixels που περιέχονται στο superpixel αυτό, θα λάβουν τιμή 0.

Με την παραπάνω διαδικασία δημιουργούμε στην πράξη μια δυαδική εικόνα δεικτοδότησης (indexing) στην οποία οι λευκές περιοχές (pixels με τιμή 1) αντιστοιχούν σε περιοχές θετικών κυττάρων.

```
from tqdm import tqdm

mask = np.zeros_like(blue_ch)

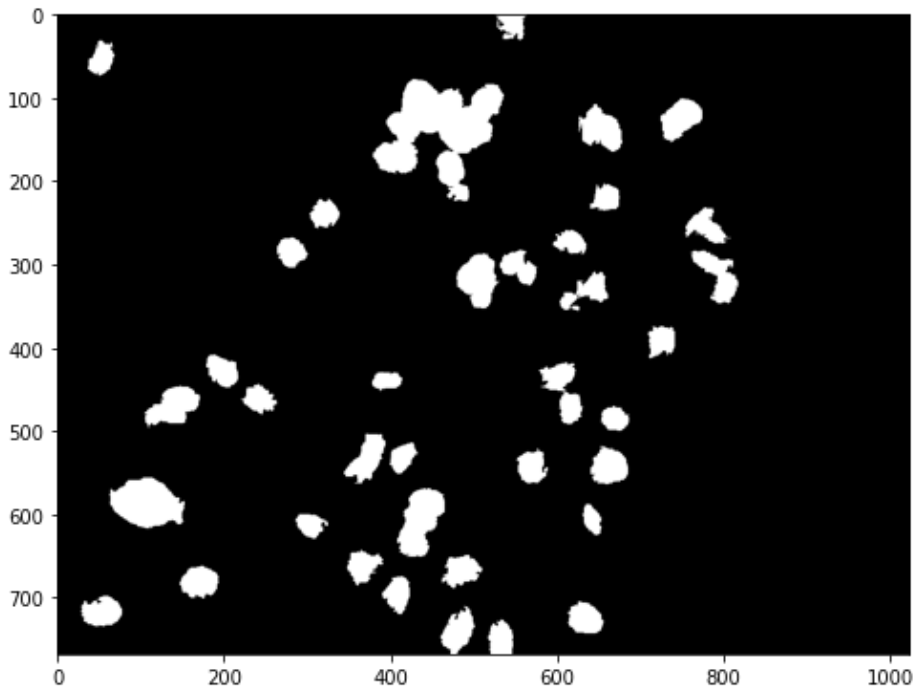
for i in tqdm(range(np.max(seg))):
    roi = blue_ch[seg==i]

    statistics = [[
        np.mean(roi),
        np.std(roi),
        skew(roi, axis=None),
        kurtosis(roi, axis=None)
    ]]

    pred = neigh.predict(statistics)

    if (pred == 1):
        mask[seg==i] = 1
```

```
plt.figure(figsize=(8,8))
plt.imshow(mask, cmap='gray')
plt.show()
```



5.2.4 Ανίχνευση ακμών ενδιαφέροντος

Η παραπάνω εικόνα είναι μια μάσκα η οποία δείχνει σε ποιες περιοχές της αρχικής εικόνας ανιχνεύθηκαν από τον ταξινομητή θετικά κύτταρα. Σε αυτό το σημείο μπορούμε να εκμεταλλευτούμε την πληροφορία που δίνει η εικόνα αυτή, με τρόπο τέτοιο ώστε να είναι χρήσιμη για το γιατρό. Μια απλή ιδέα είναι να σχεδιάσουμε περιγράμματα στην αρχική εικόνα, γύρω από τις περιοχές που μας δείχνει η μάσκα (λευκές περιοχές).

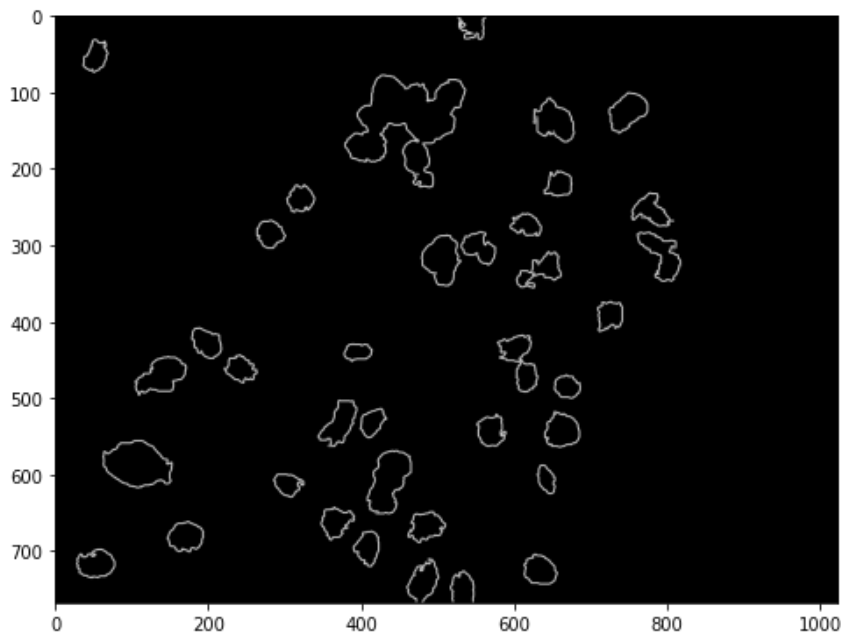
Για να το πετύχουμε αυτό, το πρώτο βήμα που πρέπει να γίνει είναι να ανιχνεύσουμε τα περιγράμματα των λευκών περιοχών στη μάσκα. Αυτό μπορούμε να το πετύχουμε με εφαρμογή τεχνικών ανίχνευσης ακμών.

Ας χρησιμοποιήσουμε τον τελεστή Sobel για την ανίχνευση των ακμών:

```
from skimage import filters

mask2 = filters.sobel(mask)

plt.figure(figsize=(8,8))
plt.imshow(mask2, cmap='gray')
plt.show()
```



Πλέον, με κατάλληλη επικάλυψη (overlay) της εικόνας πάνω στην αρχική, μπορούμε να εφαρμόσουμε τα περιγράμματα των θετικών κυττάρων.

5.2.5 Τελική επεξεργασία της κυτταρολογικής εικόνας και συμπεράσματα

Προκειμένου να επισημάνουμε στον κυτταρολόγο τις περιοχές όπου ανιχνεύθηκαν θετικά κύτταρα, σχεδιάζουμε τα περιγράμματα της τελευταίας εικόνας με κόκκινο χρώμα πάνω στην αρχική εικόνα.

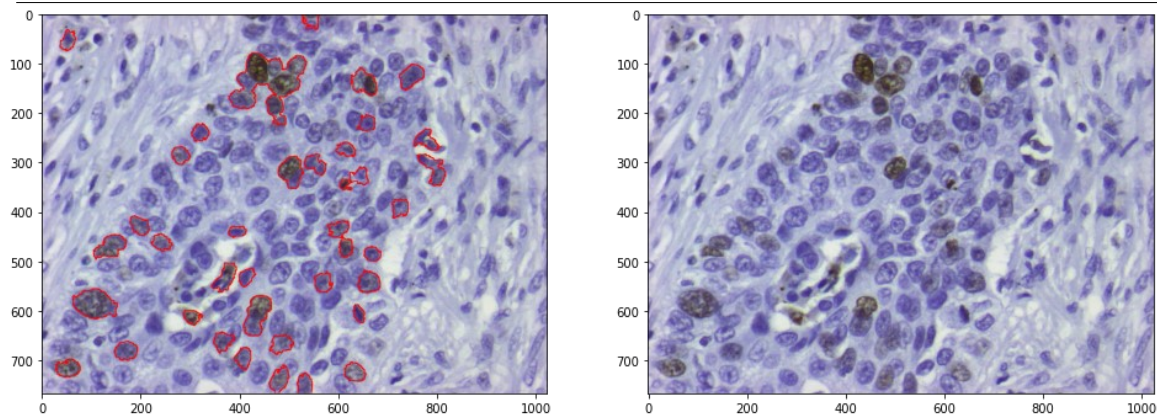
Για να το πετύχουμε αυτό αρκεί απλώς να επεξεργαστούμε την αρχική εικόνα, ώστε τα pixels τα οποία συμπίπτουν με τα pixels του περιγράμματος να πάρουν την τιμή $[255, 0, 0]$, δηλαδή καθαρό κόκκινο χρώμα.

```
final_img = np.copy(img)
final_img[mask2>0] = [255, 0, 0]
```

```
plt.figure(figsize=(25,25))
plt.subplot(1,2,1)
plt.imshow(final_img)
```

```
plt.subplot(1,2,2)
plt.imshow(img)
```

```
plt.show()
```



Πλαισιώνοντας κατάλληλα τα κύτταρα που ανιχνεύθηκαν ως θετικά από τον ταξινομητή, βοηθούμε το γιατρό σε μεγάλο βαθμό να κάνει τη μέτρηση αρκετά πιο γρήγορα. Με την παραπάνω επεξεργασία, τον βοηθούμε επίσης να κάνει λιγότερα λάθη, επισημαίνοντάς του και κύτταρα τα οποία μπορεί να διαφύγουν από την παρατήρησή του. Παρατηρώντας την αρχική εικόνα και την τελική επεξεργασμένη εικόνα είναι εύκολο να διαπιστώσει κανείς πως το σύστημα ανίχνευσε ως θετικά καρκινικά κύτταρα τα οποία δεν φαίνονται ξεκάθαρα στην αρχική εικόνα.

Χρησιμοποιώντας εξελιγμένες τεχνικές επεξεργασίας εικόνας, όπως ειδικές τεχνικές ανάλυσης υφής, και εξελιγμένες τεχνικές ταξινόμησης, όπως τεχνητά νευρωνικά δίκτυα, μπορούμε να βελτιώσουμε το σύστημα αυτό και να καταλήξουμε σε ακόμα καλύτερα αποτελέσματα. Για παράδειγμα, εφαρμόζοντας εξελιγμένες τεχνικές επεξεργασίας εικόνας μπορούμε να διαχωρίσουμε τα κύτταρα που είναι ενωμένα μεταξύ τους και τελικώς να υπολογίσουμε αυτόματα το πλήθος των θετικών κύτταρων. Αλλάζοντας τον αλγόριθμο ταξινόμησης ώστε να ανιχνεύονται τα αρνητικά κύτταρα, μπορούμε τελικά να υπολογίσουμε και το πλήθος των αρνητικών κυττάρων της εικόνας. Έτσι, μια ολοκληρωμένη εφαρμογή υποστήριξης διάγνωσης θα περιλαμβάνει τον αυτόματο υπολογισμό του πλήθους των θετικών και των αρνητικών κυττάρων της εικόνας και θα επιστρέφει την αναλογία αυτών, παράμετρος που όπως αναφέρθηκε αποτελεί πολύ σημαντικό στοιχείο για τη διάγνωση.

Συνοψίζοντας, στη συγκεκριμένη εφαρμογή παρουσιάσαμε ένα σύστημα υποστήριξης διάγνωσης για την ανίχνευση καρκινικών κυττάρων σε κυτταρολογικές/ιστολογικές εικόνες, χρησιμοποιώντας ορισμένες βασικές τεχνικές επεξεργασίας εικόνας και αναγνώρισης προτύπων. Σύμφωνα και με τη γνώμη των γιατρών, το σύστημα αυτό αποδίδει αρκετά ικανοποιητικά, αποτελώντας ένα χρήσιμο εργαλείο για την κυτταρολογία. Αξίζει να σημειωθεί ότι από τη στιγμή που έχουμε υλοποιήσει την παραπάνω εφαρμογή, μπορούμε ανά πάσα στιγμή να αναλύσουμε νέες κυτταρολογικές εικόνες.

Θερμές ευχαριστίες στον Αναπληρωτή Καθηγητή κ. Πέτρο Καρακίτσο, Εργαστήριο Διαγνωστικής Κυτταρολογίας της Ιατρικής Σχολής του Εθνικού και Καποδιστριακού Πανεπιστημίου Αθηνών, για την παραχώρηση των εικόνων και τη βοήθειά του στα πλαίσια της παρούσας άσκησης.

Παράρτημα

Τα κείμενα του προτύπου DICOM

- Part 1: introduction and overview
- Part 2: conformance
- Part 3: information object definitions

- Part 4: service class specifications
- Part 5: data structures and encoding
- Part 6: data dictionary
- Part 7: message exchange
- Part 8: network communication support for message exchange
- Part 9: point-to point communication support for message exchange
- Part 10: media storage and file format for media interchange
- Part 11: media storage application profiles
- Part 12: media formats and physical media for media interchange
- Part 13: print management point-to-point communication support
- Part 14: grayscale standard display function

Επικεφαλίδα ενός αρχείου DICOM

Κάθε DICOM αρχείο περιέχει metadata στην επικεφαλίδα του. Παρακάτω παρατίθεται ένα παράδειγμα των δεδομένων που περιέχει η επικεφαλίδα ενός αρχείου DICOM.

Filename: 'test16.dcm'	TransferSyntaxUID: '1.2.840.10008.1.2.1'
FileModDate: '11-Mar-2003 15:25:25'	ImplementationClassUID:
FileSize: 264144	'1.2.756.777.001'
Format: 'DICOM'	SourceApplicationEntityTitle:
FormatVersion: 3	'PAPYRUS 3.0'
Width: 512	IdentifyingGroupLength: 674
Height: 512	SpecificCharacterSet: 'ISO_IR 100'
BitDepth: 8	ImageType:
ColorType: 'grayscale'	'ORIGINAL\PRIMARY\AXIAL\CT_
SelectedFrames: []	SOM5 SEQ'
FileStruct: [1x1 struct]	SOPClassUID:
StartOfPixelData: 1988	'1.2.840.10008.5.1.4.1.1.2'
MetaElementGroupLength: 144	SOPInstanceUID: '64.572.218.916'
FileMetaInformationVersion: 1	StudyDate: '20021126'
MediaStorageSOPClassUID:	SeriesDate: '20021126'
'1.2.840.10008.5.1.4.1.1.2'	AcquisitionDate: '2002.11.26'
MediaStorageSOPInstanceUID:	ImageDate: '20021126'
'64.572.218.916.2'	

StudyTime: '181334.609000'	Exposure: 350
SeriesTime: '181429.406000'	FilterType: '0'
AcquisitionTime: '181443.312489'	GeneratorPower: 42
ImageTime: '181443.312489'	FocalSpots: 1.2000
Modality: 'CT'	DateOfLastCalibration: '20021126'
Manufacturer: 'SIEMENS'	TimeOfLastCalibration:
InstitutionName: 'HYGEIA'	'081846.000000'
InstitutionAddress: [1x40 char]	ConvolutionKernel: 'H40s'
StationName: 'CTNAVI'	PositionerType: "
StudyDescription: 'Head^HeadSeq'	PositionerPrimaryAngle: "
SeriesDescription: 'Brain5mmSeq 5.0 H40s'	PositionerSecondaryAngle: "
ManufacturersModelName: 'Volume Zoom'	PatientPosition: 'HFS'
ReferencedImageSequence: [1x1 struct]	ImageGroupLength: 252
SourceImageSequence: [1x1 struct]	StudyInstanceUID: [1x46 char]
PatientGroupLength: 122	SeriesInstanceUID:
PatientsName: [1x1 struct]	'1.2.756.9999.999.99.9'
PatientID: [1x47 char]	StudyID: '1'
PatientsBirthDate: '1929.01.01'	SeriesNumber: 2
PatientsSex: 'M'	AcquisitionNumber: 4
PatientsAge: '073Y'	InstanceNumber: 1
AcquisitionGroupLength: 314	ImagePositionPatient: [3x1 double]
BodyPartExamined: 'HEAD'	ImageOrientationPatient: [6x1 double]
InterventionalStatus: "	FrameOfReferenceUID: [1x46 char]
SliceThickness: 5	SliceLocation: -1.2535e+003
KVP: 120	ImagePresentationGroupLength: 214
SoftwareVersions: 'VA40C'	SamplesPerPixel: 1
ProtocolName: 'HeadSeq'	PhotometricInterpretation:
ReconstructionDiameter: 248	'MONOCHROME2'
DistanceSourceToDetector: 1040	Rows: 512
DistanceSourceToPatient: 570	Columns: 512
GantryDetectorTilt: 0	PixelSpacing: [2x1 double]
TableHeight: 185	BitsAllocated: 8
TableMotion: "	BitsStored: 8
RotationDirection: 'CW'	HighBit: 7
ExposureTime: 1000	PixelRepresentation: 0
XRayTubeCurrent: 350	SmallestImagePixelValue: 0
RescaleIntercept: 0	LargestImagePixelValue: 255
RescaleSlope: 1	WindowCenter: 128
	WindowWidth: 256
WindowCenterWidthExplanation: 'WINDOW1\WINDOW2'	
RecommendedViewingMode: "	
OverlayGroupLength_0: 0	
Private_6001_GroupLength: 28	
Private_6001_10xx_Creator: 'PAPYRUS 3.0'	
Private_6001_1001: "	
PixelDataGroupLength: 262156	

Όπως βλέπουμε υπάρχουν πολλές metadata τιμές. Οι τιμές των “Rows”, “Columns”, και “BitsStored” είναι χρήσιμες πληροφορίες για την εικόνα μας. Επίσης πρέπει να λάβουμε υπόψιν ότι η εικόνα περιέχει signed data - μία “PixelPresentation” τιμή “1”, η οποία υπονοεί το είδος των δεδομένων. Γενικά, ένα DICOM αρχείο εκτός από την εικόνα περιέχει και πληροφορίες για τον ασθενή, την εξέταση, το απεικονιστικό μηχάνημα κ.λπ. Όλες αυτές οι πληροφορίες συνιστούν ένα Πληροφοριακό Αντικείμενο (Information Object). Οι προδιαγραφές DICOM περιέχουν πολλαπλές Διευκρινίσεις Πληροφοριακών Αντικειμένων (Information Object Definitions-IODs), όπως MR εικόνα, Σχέδιο Ακτινοθεραπείας.

Τα IODs ορίζονται σε σχέση με τις λειτουργικές μονάδες οι οποίες αντιστοιχούν σε πραγματικά αντικείμενα, όπως ασθενείς, απεικονιστικό εξοπλισμό, κ.λπ. Για παράδειγμα, ο επόμενος πίνακας καταγράφει τις λειτουργικές μονάδες οι οποίες συνθέτουν ένα IOD μιας MR εικόνας:

Μονάδα	Περιγραφή
Patient	Details about the patient
General Study	General information about the study, a set of imaging procedures
Patient Study (U)	Information about the patient at the time of the study
General Series	General information about one particular imaging procedure
Frame of Reference	Information to spatially relate images in a series
General Equipment	General information about the image modality that produced the series
General Image	Details that identify and describe an image within the series
Image Plane	Slice thickness/location/spacing and other orientation details
Image Pixel	The actual image pixels and information on how to interpret them
Contrast/Bolus (C)	Details about medication given at the time of imaging
MR Image	Values specific to magnetic resonance imaging
Overlay Plane (U)	Graphics or bitmapped text to display with a particular image
VOI LUT (U)	Information to transform the contrast or intensity of the image
SOP Common	Date/time of message creation and other generic DICOM details

Βιβλιογραφία

- [1] ACR-NEMA, *Digital Imaging and Communications in Medicine (DICOM) Version 3.x*, Final Text, NEMA Standards Publication PS3.x, 1993
- [2] DICOM Homepage: <http://medical.nema.org/>
- [3] http://en.wikipedia.org/wiki/Digital_Imaging_and_Communications_in_Medicine
- [4] R. Gonzalez, R. Woods, *Digital Image Processing (Second Edition)*, Prentice Hall, New Jersey, 2002
- [5] R. Gonzalez, R. Woods, S. Eddins, *Digital Image Processing using Matlab*, Prentice Hall, New Jersey, 2004
- [6] Κ. Νικήτα, *Στοιχεία Ψηφιακής Επεξεργασίας Ιατρικών Εικόνων*, Σημειώσεις του μαθήματος της Σχολής ΗΜΜΥ του ΕΜΠ «Ιατρική Απεικόνιση & Ψηφιακή Επεξεργασία Ιατρικής Εικόνας»
- [7] Mathworks.com - Image Processing Toolbox online help documentation: <http://www.mathworks.com/access/helpdesk/help/toolbox/images/>
- [8] R. O. Duda, P. E. Hart, D. G. Stork, *Pattern Classification (Second Edition)*, Wiley Publications, 2000
- [9] S. Theodoridis, K. Koutroumbas, *Pattern Recognition*, Elsevier Academic Press, 2003
- [10] D. G. Stork, E. Yom-Tov, *Computer Manual in MATLAB to accompany Pattern Classification*, Wiley Publications, 2001
- [11] P. Phukpattaranont and P. Boonyaphiphat, “Color Based Segmentation of Nuclear Stained Breast Cancer Cell Images”, *ECTI Trans. of Electrical, Electronics, and Communications*, vol.5, no.2, Aug. 2007