

Ανάπτυξη ολοκληρωμένου περιβάλλοντος ανάλυσης και ταξινόμησης μαστογραφικών εικόνων

Κωνσταντίνα Λεβέντη

ΠΕΡΙΛΗΨΗ

Ο καρκίνος του μαστού είναι μια από τις συχνότερες μορφές καρκίνου στον γυναικείο πληθυσμό. Η μαστογραφία είναι μια αποτελεσματική και ασφαλής μέθοδος για την ανίχνευση και διάγνωση αυτής της ασθένειας. Ωστόσο, η ερμηνεία των μαστογραφιών ενέχει συχνά δυσκολίες για τους ακτινολόγους που αφορούν παράγοντες όπως η κακή ποιότητα της εικόνας ή η απροσεξία των ίδιων των γιατρών. Έτσι, τις τελευταίες δεκαετίες έχουν αναπτυχθεί συστήματα υποβοηθούμενης διάγνωσης (CAD) που παρέχουν στους ακτινολόγους μια δεύτερη γνώμη για να διαμορφώσουν την τελική τους διάγνωση.

Ένα από τα σημαντικότερα ευρήματα που σχετίζεται με τον καρκίνο του μαστού είναι οι μικροασβεστώσεις (MCs), μικρές δομές που δημιουργούνται από συγκέντρωση αλάτων ασβεστίου και εμφανίζονται συνήθως με την μορφή συμπλεγμάτων. Τα ευρήματα αυτά μπορεί να είναι τόσο καλοήθη όσο και κακοήθη και έχουν γίνει πολλές προσπάθειες για να ταξινομηθούν ανάλογα με τα χαρακτηριστικά τους.

Στην παρούσα διπλωματική εργασία παρουσιάζουμε συνολικά τη λειτουργία του συστήματος υποβοηθούμενης διάγνωσης με την ονομασία Ιπποκράτης-μστ που αναπτύχθηκε στο Εργαστήριο Πληροφορικής της Ακαδημίας Αθηνών. Το σύστημα στηρίζεται στην ανάλυση και αξιολόγηση μεμονωμένων μικροασβεστώσεων, καθώς και αντίστοιχων συμπλεγμάτων. Η υλοποίηση γίνεται σε τέσσερα στάδια: α) αρχειοθέτηση ασθενών, β) χρήση τεχνικών ανάλυσης εικόνας για την εξέταση των μαστογραφιών, γ) ανίχνευση και ανάλυση των μικροασβεστώσεων και δ) εξαγωγή διάγνωσης (καλοήθεια ή κακοήθεια). Η φάση της διάγνωσης στηρίζεται στον αλγόριθμο μηχανικής μάθησης των support vector machines (SVM). Η μεθοδολογία SVM επιλέχθηκε καθώς έχει καλύτερα αποτελέσματα από άλλους αλγορίθμους μηχανικής μάθησης όταν εφαρμόζεται σε δεδομένα που δεν ανήκουν στο σύνολο εκπαίδευσης, συνεπώς έχει μεγαλύτερη ικανότητα γενίκευσης και βρίσκει ταχύτερα την βέλτιστη λύση.

Στα πλαίσια της εργασίας αυτής αναπτύχθηκε ένα συνδυαστικό σχήμα ταξινόμησης μαστογραφικών εικόνων ως καλοήθεις ή κακοήθεις, το οποίο

αποτελείται από έναν SVM ταξινομητή και έναν νέο ταξινομητή που δημιουργήσαμε. Ο SVM εκπαιδεύεται με ένα μικρό σύνολο χαρακτηριστικών των μικροασβεστώσεων που επελέγησαν μετά από υπολογισμούς. Ο νέος ταξινομητής προσφέρει δυνατότητα κατηγοριοποίησης νέων μαστογραφικών εικόνων με βάση το περιεχόμενό τους και στηρίζεται στον υπολογισμό αποστάσεων ανάμεσα στο διάνυσμα χαρακτηριστικών της άγνωστης εικόνας και των γνωστών εικόνων. Η απόφαση προκύπτει από τις ψήφους των κοντινότερων γνωστών εικόνων. Η τελική πρόβλεψη της άγνωστης εικόνας, προκύπτει από τον συνδυασμό των προβλέψεων των δυο ταξινομητών, με εφαρμογή ενός απλού κανόνα.

Επίσης, επικαιροποιήθηκε η βάση MIRACLE ώστε να είναι δυνατή η διασύνδεση με το μετασχηματισμένο σύστημα του Ιπποκράτη-μστ που λειτουργεί σε περιβάλλον Java για να είναι προσιτό ως προς τη χρήση του.

Στο τέλος της εργασίας, παρουσιάζεται μια ολοκληρωμένη λύση για την υποβοήθηση της μαστογραφικής διάγνωσης, μαζί με την αξιολόγησή της, καθώς και τα συμπεράσματα που προέκυψαν από όλη αυτή τη μελέτη.

ΘΕΜΑΤΙΚΗ ΠΕΡΙΟΧΗ: Επεξεργασία Εικόνας, Ανάλυση Εικόνας, Ταξινόμηση μαστογραφικής εικόνας, Υπολογιστική Υποβοηθούμενη Διάγνωση

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ: Ταξινόμηση, Απεικονιστικά Ευρήματα, Αλγόριθμος SVM, Content-based Αλγόριθμος, DDSM Database, Καλοήθεια, Κακοήθεια, Ιπποκράτης-μστ, Σύστημα CAD