

AKADEMIA GÓRNICZO-HUTNICZA

WYDZIAŁ ELEKTROTECHNIKI, AUTOMATYKI, INFORMATYKI I ELEKTRONIKI
KIERUNEK INFORMATYKA



SYSTEMY INFORMATYCZNE W MEDYCYNIE

System detekcji raka piersi

Dokumentacja

Grzegorz Bylina, Agata Paciorek, Piotr Knop

Kraków, 2014

1 Wstęp teoretyczny

2 Koncepcja algorytmów

2.1 Oparte na extreme learning machines (ELM)

Wstępne przygotowanie obrazu

Pierwszym krokiem tej metody jest usunięcie szumu z obrazu. Uzyskuje się to poprzez zastosowanie filtru medianowego na oknie 3 x 3 piksele. Następnie należy uwydatnić szczegóły obrazu, gdyż są one zamazane. Pierwszym krokiem jest transformacja oryginalnego obrazu w taki sposób, aby podzielić przedział szarości na dwie części, a następnie zredukować zakres przedziału tak, aby uwydatnić przedział kontrastu oryginalnego obrazu.

Wykrycie nowotworu

Kolejnym krokiem jest segmentacja krawędzi. W pierwszej kolejności użyty WMM, czyli "wavelet transformation of local modulus maxima", który zwraca nam wszystkie krawędzie, które podejrzane są o należenie do raka, którego staramy się wykryć. Aby wśród nich znaleźć te, które rzeczywiście należą do krawędzi nowotworu, użyto operacji morfologicznych: erozji, a następnie otwarcia. Mając prawdopodobne krawędzie szukanego obiektu, wykorzystać należy metodę zwaną "Region growing". Polega ona na tym, że wybiera się jeden z pikseli oryginalnego obrazu, który znajduje się w wyznaczonym obszarze przez krawędzie, a następnie sprawdza się jego sąsiadów. Jeżeli są spełnione warunki, wtedy kolejne piksele są dołączane, aż wypełniony zostanie cały obszar ograniczony krawędziami. Otrzymany w ten sposób obszar jest nowotworem, który należało znaleźć.

Klasyfikacja nowotworu

Ponadto opracowane zostały metody do rozpoznawania rodzaju nowotworu. Wykorzystywane są do tego wektory, które charakteryzują dane cechy. Analizowany jest również kształt, wielkość obszaru, czy gładkość krawędzi. Do osiągnięcia należytych wyników wykorzystywane jest tytułowy ELM, czyli "extreme learning machine". Jest to algorytm oparty na sieciach neuronowych, którego cechą jest zdolność bardzo szybkiej nauki. Na podstawie zdefiniowanych wcześniej rodzajów, stara się on odpowiednio dopasować dany obraz po przekształceniach, które zostały wcześniej dokonane.

2.2 Kombinacja dwóch znanych metod detekcji

Wykrywanie tą metodą można podzielić na dwie równoległe ścieżki: Detekcję bez nadzoru ("Unsupervised detection") oraz detekcję z nadzorem ("Supervised detection").

Unsupervised detection

Zadaniem tej części jest wyznaczenie krawędzi obszaru, którym jest nowotwór. W tym celu wykorzystywana jest kombinacja wynikowego obrazu z metody "local statistical measurements" (LSM) oraz wynikowego obrazu z metody "sliding band filtering" (SBF). Następnie przeprowadzana jest detekcja krawędzi na którą składają się następujące kroki: konstrukcja izo-konturowej mapy, formowanie drzewa integracji oraz obliczenie głębokości zagnieżdżenia.

Supervised detection

Zadaniem tej części jest nauka i klasyfikacja charakterystyk obiektu. Wykorzystywany jest algorytm "wavelet model-based detection algorithm" do wykrywania krawędzi, a następnie za pomocą pierwszego klasyfikatora Support Vectors Machines (SVM) dokonywana jest klasyfikacja.

Rezultat

Następnie dokonywana jest kombinacja powyższych wyników, której efektem jest wyznaczenie i sklasyfikowanie obszaru zawierającego nowotwór. Dodatkowo w celu zminimalizowania błędów wykorzystywana jest redukcja tzw. "false-positive" czyli wyników określonych jako podejrzane, które w rzeczywistości okazały się fałszywym alarmem. Uzyskuje się to za pomocą klasyfikatorów, na podstawie których dokonany jest trening algorytmów klasyfikacji.

2.3 Za pomocą multiscale spatial Weber law

Weber law descriptor (WLD)

WLD, czyli Weber law descriptor reprezentuje obraz jako histogram wzbudzeń różniczkowych, zgodnie z odpowiadającymi kierunkami gradientu. Bazuje na prawie Webera, które mówi, że stosunek przyrostu progu do intensywności tła jest stały. Pierwszym krokiem jest obliczenie różnicy wzbudzenia (DE - differential excitation) dla każdego piksela. Następnie wyliczenie kierunku gradientu (GO - gradient orientation). Na tej podstawie obliczany jest histogram. W ten sposób zostały otrzymane wszystkie krawędzie na obrazie.

Support vector machine (SVM)

W tym przypadku również używany jest SVM. Jest to jeden z najbardziej zaawansowanych klasyfikatorów oraz bardzo dobrze znana metoda do klasyfikacji problemu dwu-klasowego. Dodatkowo wykorzystując radial basis function" (RBF) można w skuteczny sposób zredukować "false-positives" by w jak największym stopniu wyeliminować potencjalne błędy programu.

3 API