

Opinia

o rozprawie doktorskiej mgr inż. Marcina Batora pt.

„Automatyczna detekcja zmian nowotworowych w obrazach mammograficznych z wykorzystaniem dopasowania wzorców i wybranych narzędzi sztucznej inteligencji”
IPPT PAN, Warszawa, 2008 r.

Rozprawa doktorska mgr inż. Marcina Batora dotyczy komputerowej metody wspomagania diagnostyki medycznej zmian nowotworowych piersi na podstawie analizy obrazów radiologicznych. Można ją zatem zakwalifikować do inżynierii biomedycznej, w tym szczególnie do komputerowej analizy obrazów medycznych. Jest to tematyka aktualna, intensywnie rozwijana na świecie, mająca walory poznawcze i duże znaczenie społeczne. Rozprawa stanowi też kolejny etap w ciągu prac poświęconych komputerowemu wspomaganiu diagnozowania nowotworów piersi, prowadzonych od szeregu lat pod kierunkiem promotora rozprawy, prof. dra hab. inż. Mariusza Nieniewskiego.

Rozprawa składa się z części wstępnej, zawierającej sformułowanie tezy, celów i zakresu pracy, 5 rozdziałów (w tym Wprowadzenia i Podsumowania), 3 Dodatków (w tym Spisu skrótów i symboli) oraz (niewymienionej w Spisie treści) Bibliografii obejmującej 107 pozycji (w tym 2 pozycje z udziałem autora rozprawy). Do rozprawy dołączono też płytę CD zawierającą: zbiory obrazów referencyjnych z bazy danych MIAS wraz z wynikami dopasowania wzorców półsferycznych, obrazów uczących i testujących pochodzących z bazy danych DDSM wraz z wynikami dopasowania wzorców uzyskanych przy pomocy algorytmu ewolucyjnego, programy obliczeniowe opracowane przez Autora rozprawy i tekst rozprawy doktorskiej.

Główna teza rozprawy brzmi następująco: „Efektywna detekcja zmian w postaci mas nowotworowych w obrazach mammograficznych jest możliwa za pomocą metody dopasowania wzorców w wielu skalach i w piramidzie rozdzielczości – poprzez generację obszarów zainteresowania ROI (*Region of Interest*). Optymalizacja wykorzystywanych wzorców kołowo-symetrycznych jest możliwa za pomocą algorytmu ewolucyjnego” (s. 5). Jako rozwinięcie powyższej tezy w rozprawie sformułowano 6 celów szczegółowych określających konstrukcyjno-doświadczalny zakres i charakter pracy.

We Wprowadzeniu (Rozdz. 1), opartym w znacznej części na przeglądzie literatury, przedstawiono ogólnie społeczne i medyczne znaczenie diagnozowania nowotworów sutka, budowę anatomiczną sutka, rodzaje obserwowanych w nim zmian nowotworowych, scharakteryzowano bazy referencyjnych danych obrazowych związanych z wykrywaniem i rozpoznawaniem zmian nowotworowych w sutku, a wreszcie – przedstawiono metody komputerowego wspomagania diagnozy nowotworów sutka. Powyższa problematyka została tu przedstawiona bardzo skrótowo; opisu budowy anatomicznej sutka nie zilustrowano żadnym rysunkiem, opis metod diagnostycznych ograniczono tylko do badań rentgenowskich, pomijając metody badań radiografii izotopowej (SPECT) lub ultrasonograficznych, choćby w celu uzasadniająco skoncentrowanie uwagi na badaniach rentgenowskich. Nie mniej jednak, we Wprowadzeniu zostały przedstawione te podstawowe pojęcia, z których korzysta się w dalszych rozdziałach rozprawy. Do badań doświadczalnych wykorzystano bazy danych MIAS oraz DDSM, dostępne poprzez Internet, nie korzystano natomiast z żadnych klinicznych danych obrazowych pochodzących ze źródeł polskich.

W Rozdziale 2 zdefiniowano szereg kolejnych pojęć oraz przedstawiono założenia, na których oparto badania doświadczalne stanowiące treść następnych rozdziałów. Stwierdzono,

że celem pracy jest opracowanie metody wykrywania złośliwych zmian nowotworowych w szerokim zakresie rozmiarów od 3 do 50 mm, a także – wykrywanie ich na obrzeżach obrazów. Na podstawie danych literaturowych zdefiniowano tu takie miary oceny jakości detekcji i klasyfikacji zmian nowotworowych jak dokładność, czułość i swoistość. Przez „klasyfikację”, tak tu, jak i w całej rozprawie, rozumie się rozpoznawanie; jest to często, przez wielu autorów popełniana nieścisłość terminologiczna (w rzeczywistości *klasyfikacja* polega na podziale zbioru obiektów na rozłączne klasy podobieństwa, natomiast *rozpoznawanie* polega na przypisaniu nowego obiektu do jednej z tak utworzonych klas). W Rozdziale 2 zdefiniowano też znane z literatury metody oceny algorytmów detekcji w postaci krzywych *FROC* i *ROC* oraz miarę pola powierzchni pod krzywą *ROC*, czyli A_{UC} . W dalszej części rozprawy do oceny wyników własnych badań zostały wykorzystane głównie krzywe *FROC*, czyli zależności funkcyjne frakcji prawidłowych decyzji pozytywnych (tj. wykryć) *TPF* od uśrednionej liczby nieprawidłowych decyzji pozytywnych w pojedynczym obrazie *FPI*. W tej części rozprawy występują drobne błędy formalne: w prawej części wzoru (2.4) pominięto nawias, wskutek czego procentowa dokładność A_C może przyjmować wartości ujemne, a w opisie wzoru (2.8) podano błędnie zmienną ρ_i zamiast $\Delta\rho_i$. Interesującą częścią Rozdziału 2 jest analiza trudności oceny zgodności pokrycia obszaru zmiany nowotworowej wykrytego przez testowany algorytm z obszarem wskazanym w opisie danych referencyjnych. Podstawą badań w dalszych rozdziałach są też scharakteryzowane w końcu Rozdziału 2 cechy mas nowotworowych, takie jak: masa centralna, spikule i przejaśnienia. Scharakteryzowano je tu jednak tylko opisowo, bez zilustrowania obrazami rzeczywistymi lub choćby rysunkami.

W końcowej części Rozdziału 2 zostały przedstawione wyniki badań Autora, w których korzystając z bazy danych MIAS porównano efektywność detekcji zmian nowotworowych typu spikularnego poprzez cechę *ALOE*, związaną z histogramem gradientu jasności, w obszarze okrągłym i pierścieniowym. Badania te wskazały, że cecha *ALOE* w małym stopniu wpływa na wykrywalność form spikularnych w peryferyjnych częściach badanej struktury. Wynik ten wpłynął w dalszej części rozprawy na skoncentrowanie uwagi na wykrywanie zmian nowotworowych o regularnym obrysie kołowym. Wydaje się jednak, że rezygnacja z badania wykrywalności zmian o nieregularnym kształcie na podstawie innych cech niż *ALOE* nie jest na tej podstawie w pełni uzasadniona. Zauważmy też, że na prawym wykresie 2.5 błędnie oznaczono jedną z osi, co jednak nie ma wpływu na tok rozważań.

W Rozdziale 3 przedstawiono opracowaną przez Autora metodę detekcji mas nowotworowych opartą na dopasowaniu wzorców w wielu skalach i w piramidzie rozdzielczości. Jako wzorzec masy nowotworowej przyjęto tu jego postać półsferyczną, co w odniesieniu do rzeczywistości jest oczywiście przybliżeniem, nie biorącym pod uwagę nie tylko kształtów spikularnych, lecz także obłych (elipsoidalnych). Zaletą takiego założenia jest natomiast jego względna prostota. Dopasowanie wzorca (poprzez dobór promienia półsfery) do struktury widocznej na podobrazie jest oparte na mierze podobieństwa, za którą przyjęto współczynnik korelacji wzajemnej pary funkcji. Takie założenie z kolei uniezależnia wynik od liniowych transformacji skali jasności obrazów. Zbadano też wpływ kwantyzacji jasności na wartość współczynnika korelacji stwierdzając, że już przy kilkunastu poziomach kwantyzacji staje się on pomijalnie mały. Wzór na współczynnik korelacji został poddany przekształceniu (opisanemu dokładniej w jednym z Dodatków), które zmniejszyło koszty obliczeń współczynnika. Ważną cechą opracowanej metody dopasowania wzorców jest jej wielkoskalowość i wielorozdzielczość, która umożliwia wykrywanie mas nowotworowych w szerokim zakresie ich rozmiarów. Ponieważ, jak wspomniano, Autor zakłada półsferyczną

postać wzorców, zmusiło Go to do zaproponowania metody postępowania z masami nowotworowymi o kształtach wydłużonych lub występujących na tle innych obiektów wydłużonych. Zaproponowano tu wstępne usuwanie takich obiektów z obrazu przy pomocy innych, odpowiednio dobranych metod. Takie postępowanie, jak pokazały przeprowadzone badania, okazało się jednak nie dość efektywne i w pewnej mierze problem ten w rozprawie pozostawiono otwartym, choć Autor nie odważył się wyciągnąć stąd wniosku, że założona w rozprawie klasa wzorców powinna być odpowiednio rozszerzona. Zaproponowaną metodę dopasowania wzorców sprawdzono na obrazach ze zbioru MIAS i DDSM, wyniki przedstawiono w postaci krzywych *FROC* (na lewym wykresie 3.13 zabrakło oznaczenia osi pionowej).

Podtrzymując nadal założenie o kołowym obrysie masy nowotworowej w Rozdziale 4 rozprawy Autor rozszerza klasę wzorców dopuszczając w nich bardziej zróżnicowane postaci jasności w funkcji odległości od centrum wzorca, w tym zwłaszcza uwzględniając występowanie efektu halo na obwodzie masy nowotworowej. W tym celu kołowo-symetryczną postać tej funkcji dobiera się optymalnie przy pomocy algorytmu ewolucyjnego. Jako genotyp funkcji przyjmuje się stabelaryzowaną dyskretną postać funkcji, a jako funkcję dopasowania wspomnianą już wartość współczynnika korelacji wzajemnej podobrazu testowego i wzorca. W procesie optymalizacji wykorzystuje się standardowe operacje krzyżowania i mutacji genotypów. Autor dostosował algorytm ewolucyjny do swych potrzeb w taki sposób, ażeby eliminować z procesu optymalizacji rozwiązania niespełniające ograniczeń (malejący gradient jasności w środkowej części wzorca i rosnący w części otaczającej). Wyniki optymalizacji doboru wzorców zostały przedstawione w postaci krzywych *FROC*, na których porównano je z wynikami dopasowania wzorców półsferycznych. Wyniki te potwierdziły poprawność i przydatność zaproponowanej metody, choć procentowa poprawa jakości dopasowania uzyskana dzięki zastosowaniu algorytmu ewolucyjnego na Rys. 4.8 została przedstawiona bez oceny statystycznej wielkości błędu. Wykresy przedstawione na Rys. 4.5 nie mają oznaczonych osi odciętych i rzędnych.

Na uzyskane wyniki warto jednak spojrzeć także z medycznego punktu widzenia. Krzywe *FROC* na Rys. 4.8 u dołu przedstawiają zależności *TPF* od *FPI*, czyli frakcji rozpoznań prawidłowych od średniej liczby rozpoznań nieprawidłowych przypadającej na jeden obraz, przy czym wykresy z lewej strony odnoszą się do przypadku, w którym obrazy rozpatrywano oddzielnie, a prawe – do przypadku, w którym rozpatrywano je łącznie. Otóż łączne rozpatrywanie obrazów daje wprawdzie bardziej „optymistyczną” postać wykresów, ale z lekarskiego punktu widzenia jest mniej wiarogodną oceną algorytmu rozpoznawania, gdyż maskuje ono przypadki rozpoznań nieprawidłowych w sytuacjach „trudnych” poprzez rozpoznania prawidłowe uzyskane w sytuacjach „łatwych”. Z kolei, ocena algorytmu na podstawie oddzielnego rozpatrywania obrazów dała wyniki dochodzące maksymalnie do $TPF = 0.85$, co oznacza, że ok. 15% przypadków zmian nowotworowych nie zostało wykrytych. Uwaga ta nie podważa wartości omawianej tu rozprawy, wskazuje raczej na trudność zagadnienia i konieczność prowadzenia dalszych badań. Porównania uzyskanych wyników z opisanymi w literaturze przez innych autorów dokonano w Rozdziale 5. Oceny takie z zasady mają charakter wieloaspektowy. Można ogólnie stwierdzić, że metody przedstawione w rozprawie są porównywalne z literaturowymi w tym sensie, że w niektórych przypadkach przewyższają one wyniki opublikowane przez innych autorów, a w niektórych przypadkach ustępując im nieznacznie ilościowo są jednocześnie oparte na lepiej udokumentowanych podstawach. Autorowi pracy udało się też opracować programy w środowisku Matlab działające stosunkowo wydajnie, umożliwiające przeanalizowanie typowego mammografu w ciągu kilku minut, co kwalifikuje je do szerszego stosowania.

Rozprawa jest napisana językiem wystarczająco poprawnym. Liczba usterek formalnych lub edytorskich jest w niej stosunkowo niewielka i nie mają one istotnego wpływu na merytoryczną wartość pracy. Pewne fragmenty rozprawy zostały przedstawione zbyt powierzchownie (dotyczy to części wstępnej, lecz także pominięcia w części doświadczalnej oceny błędów standardowych estymowanych wartości). Zdarzają się też w niej niezbyt fortunate skróty myślowe, na przykład na str. 35, gdzie stwierdza się, że „Współczynniki korelacji między wzorcem a kolejnymi podobrazami tworzą obraz prawdopodobieństwa obecności poszukiwanego obiektu”. Jest to o tyle nieprawdziwe, że korelacja może przyjmować zarówno wartości dodatnie, jak i ujemne, a prawdopodobieństwo – jedynie wartości nieujemne, a przy tym w rozprawie nie przytoczono żadnej zależności między wartością współczynnika korelacji i wspomnianym prawdopodobieństwem.

Mimo wymienionych w niniejszej Opinii uwag krytycznych uważam, że w rozprawie uzasadniono postawioną na wstępie tezę oraz zostały osiągnięte założone cele badań. Rozprawa odznacza się ogólnym poziomem naukowym wymaganym od rozpraw doktorskich w dziedzinie nauk technicznych, a także prezentuje oryginalne i ważne wyniki badań naukowych. W związku z powyższym uważam, że rozprawa mgra inż. Marcina Batora w pełni zasługuje na dopuszczenie do jej obrony publicznej.

