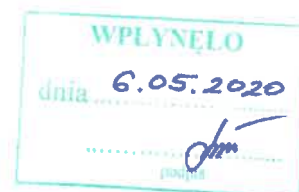


Prof. dr hab. inż. Marek Kurzyński
Politechnika Wrocławska
Wydział Elektroniki
Katedra Systemów i Sieci Komputerowych

Wrocław, 21.04.2020 r.



RECENZJA

rozprawy doktorskiej mgr. inż. Pawła KŁECZKA

pt. *Computer vision and machine learning algorithms for dermatopathological image analysis (Algorytmy wizji komputerowej i uczenia maszynowego w analizie obrazów dermatopatologicznych)*

1. Zakres i cel rozprawy

Opiniowana praca doktorska jest rozprawą z zakresu automatycznego wspomaganie diagnozy medycznej poprzez przetwarzanie, analizę i rozpoznawanie informacji obrazowej, ukierunkowanej na diagnozowanie skórnych zmian melanocytowych na podstawie obrazów histopatologicznych barwionych preparatów skóry. Obejmuje złożoną problematykę analizy zadania diagnostycznego w celu zdefiniowania cech różnicujących normę od nowotworu z uwzględnieniem stopnia złośliwości, pozyskiwania preparatów mikroskopowych, ich przetwarzania w celu poprawy jakości, wydobycia obiektów i wyznaczenia cech morfometrycznych, a także metody i algorytmy klasyfikacji zmierzające do podjęcia stosownej decyzji diagnostycznej. Zawiera również wyniki badań eksperymentalnych weryfikujących praktyczną przydatność proponowanych koncepcji oraz dyskusję rezultatów opracowanych metod.

Medyczna diagnostyka obrazowa znajduje – w miarę rozwoju technik obrazowania – coraz większe zastosowanie w codziennej praktyce klinicznej. Mimo dużej na ogół zawartości informacyjnej obrazu, która ułatwia określenie stanu pacjenta i podjęcie stosownej decyzji diagnostycznej bądź terapeutycznej, diagnostyka obrazowa dostarcza także wielu problemów. Wynikają one przede wszystkim z trudności w oddzieleniu obiektu istotnego od tła, odróżnieniu informacji bardziej od mniej istotnej, czy na wyodrębnieniu określonych klas obiektów istniejących w obrazie. Trudności przysparza także liczbowa ocena niektórych cech charakteryzujących analizowane obiekty, która często jest wręcz niemożliwa do przeprowadzenia na drodze wzrokowego oglądu. Zastosowanie w tym zakresie metod cyfrowego (komputerowego) przetwarzania i analizy obrazów stanowi niewątpliwe ułatwienie w ilościowej ocenie obiektów zawartych na obrazie, która może być dalej wykorzystana w automatycznym procesie komputerowego wspomaganie podejmowania decyzji.

Niniejsza praca mieści się w tym właśnie bardzo aktualnym nurcie informatyki medycznej, dlatego tematykę podjętą w rozprawie uważam za oryginalną i w pełni uzasadnioną, a samą pracę za wartościowy i istotny wkład we współczesną problematykę automatycznego przetwarzania i analizy medycznej informacji obrazowej oraz wspomagania procesów decyzyjnych w medycynie. Pozytywna ocena tematu rozprawy wynika również z faktu, iż każda praca z tego zakresu, której wynikiem są konstruktywne algorytmy przetwarzania i klasyfikacji informacji obrazowej ukierunkowane na konkretny medyczny problem diagnostyczny, jest celowa i ważna z praktycznego punktu widzenia i za taką właśnie uważam opiniowaną rozprawę.

2. Zawartość rozprawy

Recenzowana praca składa się z 8 rozdziałów poprzedzonych streszczeniami w języku polskim i angielskim i spisem stosownych symboli oraz bibliografii zawierającej 151 pozycji i obejmuje 145 stron maszynopisu. We wstępie (*Introduction*) wprowadzającym w „klimat” pracy przedstawiono charakterystykę zawartości poszczególnych rozdziałów pracy i wyszczególniono projekty badawcze związane z realizacją rozprawy. Rozdział zawiera również jednoznacznie i szczegółowo sformułowane tezy pracy, wyszczególniony zestaw uzyskanych wyników ze szczególnym uwzględnieniem ich innowacyjności oraz schemat blokowy całościowego procesu przetwarzania i analizy informacji obrazowej od wejściowego obrazu preparatu histopatologicznego do wyniku końcowego w formie diagnozy. Rozdział drugi (*Medical background*) zawiera medyczny kontekst problemu diagnostycznego rozpatrywanego w pracy. Przedstawiono tu w szczególności podstawy histologii skóry i dermatopatologii, proces przygotowania preparatów histologicznych, digitalizację przezroczy oraz cechy widoczne na obrazie, kluczowe w zadaniu różnicowania nowotworu jako łagodnego lub złośliwego. Informacje te, stanowiąc wiedzę zorientowaną dziedzinowo, mają istotne znaczenie dla proponowanych dalej algorytmów segmentacji i przyjętych metod automatycznej klasyfikacji. Autor zwrócił dodatkową uwagę na trudności diagnostyczne skórnych zmian melanocytowych przy użyciu popularnego badania dermoskopowego, gdyż według doniesień literaturowych nawet doświadczony lekarz nie osiąga wyższych wartości czułości i specyficzności swoich rozstrzygnięć niż odpowiednio 90% i 60%. Powoduje to, iż złotym standardem są badania histopatologiczne, ale i tutaj trudności interpretacyjne sprawiają, iż opinie histopatologów bywają rozbieżne nawet w 30% – 40% ocenianych preparatów. W rozdziale 3 (*Preface to the proposed algorithms*) przedstawiono układ opracowanych algorytmów składających się na całościowy system diagnostyki zmian skórnych na podstawie obrazów dermatopatologicznych. System ten zawiera następujące etapy: (1) segmentacja obszaru preparatu histopatologicznego (tkanki), (2) segmentacja naskórka, (3) segmentacja sopli naskórkowych (*rete ridges*) i wyznaczenie ich cech morfometrycznych, (4) ekstrakcja cech i klasyfikacja. Rozdział zawiera także wyniki studiów literaturowych dotyczących istniejących rozwiązań algorytmicznych etapów przetwarzania preparatów oraz charakterystykę zbiorów obrazów wykorzystywanych w badaniach eksperymentalnych. Cztery dalsze rozdziały (Rozdział 4: *Tissue segmentation*, Rozdział 5: *Epidermis segmentation*, Rozdział 6: *Epidermis morphometry*, Rozdział 7: *Feature extraction*

and classification) zawierają zasadnicze wyniki pracy w podziale na wymienione etapy. W każdym z rozdziałów zawarto szczegółową prezentację opracowanych algorytmów oraz wyniki przeprowadzonych analiz i badań eksperymentalnych na zgromadzonych zbiorach zdjęć preparatów skórnych zmian melanocytowych. Rozdział 8 (*Summary*) podsumowuje pracę przedstawiając między innymi zestawienie oryginalnych wyników oraz propozycje dalszych badań w rozpatrywanym zakresie.

Przedstawiona rozprawa prezentuje w sposób precyzyjny i szczegółowy najważniejsze rezultaty objęte jej tematyką w logicznym i przejrzystym układzie oraz w formie, w której widoczna jest duża dbałość doktoranta o należyty poziom edytorski, ze szczególnym uwzględnieniem bardzo bogatej i trafnie skomponowanej sfery ilustracyjnej. Umożliwiło to zaprezentowanie uzyskanych wyników w sposób szczegółowy i kompletny, co zostało w ocenianym tekście starannie wykorzystane. Ogólnie zatem objętość pracy, jej zakres, sposób ujęcia materiału oraz redakcję całości oceniam pozytywnie.

3. Uzyskane wyniki

Celem pracy było wykazanie następujących tez (cyt.):

Teza główna: *Zaawansowane metody przetwarzania obrazów oraz uczenia maszynowego umożliwiają szczegółową analizę morfometrii naskórka zmiany skórnej na obrazach histopatologicznych.*

Teza I: *Przy wykorzystaniu zaproponowanych metod segmentacji i analizy obrazu możliwa jest ocena stopnia złośliwości zmiany skórnej.*

Teza II: *Metody uczenia maszynowego pozwalają na sformułowanie podstawowych kryteriów diagnostycznych, możliwych do zastosowania podczas badań histopatologicznych.*

Realizacja tak nakreślonych tez doprowadziła do szeregu szczegółowych i konkretnych osiągnięć. Do szczególnie ważnych i wartościowych wyników pracy – zarówno w aspekcie teoretycznym, jak i praktycznym – zaliczam:

1. Opracowanie algorytmu do segmentacji obszaru preparatu histopatologicznego (tkanki), który bazuje na statystycznej analizie rozkładu koloru, zarówno w całym obrazie, jak i w lokalnych otoczeniach jego wybranych fragmentów. Po dokonaniu preprocesingu obrazu z wykorzystaniem filtru gaussowskiego, którego celem jest eliminacja szumów matrycy CCD następuje wyznaczenie średniego koloru tła. Przyjmuje się tutaj założenie, iż piksele tła są zbliżone do „kolorów achromatycznych” czyli odcieni szarości. Pozwoliło to na opracowanie oryginalnej procedury, w której wyznaczenie średniego koloru tła odbywa się na podstawie histogramu obrazu w przestrzeni CIELab w zakresie pikseli o małej saturacji. Parametry procedury są wyznaczane poprzez rozwiązanie problemu optymalizacyjnego. Następnie odbywa się wstępna segmentacja tła jako pikseli, których odległość od wartości średniej w przestrzeni CIELab nie przekracza zadanego progu. Końcowy wynik segmentacji jest uzyskany na etapie poprawy segmentacji wstępnej przez uwzględnienie fragmentów pomiędzy strukturami tkanki wypełnionych parafiną. W badaniach eksperymentalnych porównano wynik segmentacji z wykorzystaniem zaproponowanej metody z segmentacją ręczną traktowaną jako bezbłędną oraz z dwoma metodami literaturowymi (FESI i GrabCut). Do badań wykorzystano zbiór 60 preparatów,

a do porównania wyników zastosowano wskaźnik Jaccarda w trzech formach. Uzyskane rezultaty pokazują dużą zgodność autorskiej metody z ręczną metodą segmentacji oraz wyższość nad metodami literaturowymi.

Opracowany algorytm jest łatwy do implementacji i mało czuły na powiększenie obrazu preparatu, warunki oświetlenia, artefakty i zakłócenia.

2. Opracowanie algorytmu segmentacji warstwy naskórka wykorzystującego analizę rozmieszczenia i kształtów obszarów tła obrazu w obrębie tkanki oraz histogramów stężeń hematoksyliny i eozyny w preparacie. Zaproponowany algorytm składa się z 3 kroków. W pierwszej kolejności wysegmentowana na etapie poprzednim tkanka zostaje poddana procedurze nazwanej analizą porowatości. Wykorzystuje się tu fakt, iż puste przestrzenie (pory) są typowe dla niektórych obszarów występujących w preparacie (np. w tkance tłuszczowej, pomiędzy wiązkami włókien kolagenu, w pobliżu jąder komórkowych i połączeń międzykomórkowych), co jest konsekwencją złożonych procesów przygotowania preparatu. W odróżnieniu od tych obszarów, naskórek jest warstwą stosunkowo zwartą i pozbawioną przejrzystości. To spostrzeżenie pozwoliło na zaproponowanie metody, w której dokonuje się właśnie analizy porowatości poprzez identyfikację pustych przestrzeni, a następnie usuwa się obszary z dużą porowatością traktując je jako głębsze warstwy skóry. Uzyskany rezultat segmentacji zawiera jednak wciąż jeszcze regiony fałszywie pozytywne, takie jak ciasno upakowane włókna kolagenowe, skupiska komórek pigmentowych, czy też grubsze fragmenty warstwy rogowej naskórka. Jako rozwiązanie tego problemu zastosowano analizę koncentracji barwników (hematoksyliny i eozyny) uzasadnioną faktem, iż włókna kolagenowe i komórki rogowate naskórka w niewielkim stopniu absorbują oba barwniki.

Jak poprzednio, skuteczność opracowanej metody segmentacji naskórka poddano eksperymentalnej weryfikacji. Po eksperymentalnym doborze parametrów metody porównywano wyniki segmentacji warstwy naskórka z segmentacją uzyskaną literaturowymi metodami (CET, GTSA i THM) na zbiorze 75 preparatów. Jako kryteria oceny przyjęto zdefiniowane względem segmentacji referencyjnej przeprowadzonej ręcznie przez doświadczonego patologa 3 miary jakości: czułość, specyficzność i dokładność. Uzyskane rezultaty pokazują, iż zaproponowana metoda jest skuteczna i efektywna: we wszystkich analizowanych preparatach została wysegmentowana warstwa naskórka przy porównywalnych lub lepszych wartościach kryteriów.

3. Opracowanie algorytmu segmentacji tzw. sopli naskórkowych poprzez analizę krzywizny konturu obszaru oraz szkieletu obszaru w reprezentacji grafowej oraz wyznaczenie cech opisujących morfometrię naskórka i sopli naskórkowych, istotnych z punktu widzenia rozpatrywanego zadania diagnostycznego. Jedną z cech wykorzystywanych w klasyfikacji jest długość głównej osi podstawy (bez sopli) naskórka. Wymaga to wysegmentowania podstawy naskórka, co z uwagi na występujące wgłębienia (kieszenie) pomiędzy naskórkiem a warstwą rogową naskórka oraz wspomniane już sople naskórkowe wnikaające w głębsze warstwy skóry, nie jest zadaniem prostym. Autor zastosował tu szereg procedur pozwalających na zlokalizowanie tzw. „zewnątrznej” krawędzi naskórka, jej wygładzenie z zastosowaniem odpowiednich operacji morfologicznych oraz odcięcie sopli naskórkowych poprzez zlokalizowanie ich podstaw. Zlokalizowanie podstawy sopli

naskórkowych umożliwiło detekcję samych soplei oraz wyznaczenie kolejnych cech związanych z ich morfometrią (szerokość, długość i wysokość).

Przeprowadzone badania eksperymentalne miały na celu odpowiedzieć na pytania, ile soplei naskórkowych zostało zlokalizowanych prawidłowo, czy istnieje problem nadwrażliwości metody poprzez generowanie nadmiarowych soplei i na ile wyznaczone parametry morfometryczne soplei naskórkowych są zbliżone do parametrów wyznaczonych w sposób ręczny przez doświadczonych dermatopatologów. Jako wskaźnik oceny skuteczności opracowanej metody przyjęto miarę Jaccarda w dwóch wariantach, aby uwzględnić dość częste przypadki występowania soplei o nietypowym kształcie, które można dwójako interpretować (dwa sople albo jeden soplel z podwójnym zakończeniem).

Badania dotyczące oceny skuteczności wykrywania soplei przeprowadzono na zbiorze 25 zdjęć preparatów, które zawierały łącznie 992 ręcznie wysegmentowanych soplei naskórkowych. Oceniana metoda wykryła 894 sople, z czego 825 zostało zlokalizowanych prawidłowo. W kolejnym eksperymencie przeprowadzonym na zbiorze 75 obrazów zawierających łącznie 3709 ręcznie wysegmentowanych soplei, opracowany algorytm wykrył 2613 soplei, w tym 1571 prawidłowo. Przeprowadzona szczegółowa analiza popełnianych błędów pokazała silną zależność niewykrywania soplei od ich długości. Z kolei badania dotyczące poprawności wyznaczania cech morfometrycznych soplei (szerokość i długość) pokazały, iż wyniki mają rozrzut normalny o niewielkiej wariancji i nie wykazują błędu systematycznego.

4. Zaimplementowanie 4 klasyfikatorów i przeprowadzenie eksperymentalnych badań porównawczych z analizą statystyczną uzyskanych wyników.

Biorąc pod uwagę wiedzę kliniczną doświadczonych dermatopatologów zaproponowano ostatecznie następujący zestaw cech, istotnych z punktu widzenia klasyfikacji skórnych zmian melanocytowych: (1) długość zmiany patologicznej (długość głównej osi naskórka), (2) gęstość soplei naskórkowych (ilość/długość zmiany) (3) gęstość długich soplei naskórka, (4) mediana grubości podstawy naskórka, (5) liczba przerostowych obszarów naskórka (liczba obszarów naskórka 2 razy grubszych niż mediana), (6) gęstość przerostowych obszarów naskórka (liczba obszarów odniesiona do długości zmiany patologicznej), (7) stosunek długości sekcji z brakującymi soplami naskórkowymi do całej długości zmiany patologicznej.

Wyznaczone cechy były podstawą działania zaimplementowanych następujących klasyfikatorów: (1) wielomianowa regresja logistyczna (REG), (2) Naiwny klasyfikator bayesowski (NB), (3) drzewo decyzyjne (DT) i (4) metoda SVM. Skuteczność zastosowanych klasyfikatorów została oceniona eksperymentalnie z użyciem zbioru 75 zdjęć preparatów skórnych zmian melanocytowych. W procedurze trenowania klasyfikatorów i ich testowania zastosowano 1000 krotnie powtórzoną metodę 5CV. Przeprowadzono 2 eksperymenty: w pierwszym przyjęto 0-1 funkcję strat, a w drugim koszt zaklasyfikowania złośliwego nowotworu jako łagodnego jest dwa razy droższy. Wyniki badań ujęte w formie wykresów pudełkowych, w postaci tabel ujmujących wartość średnią, medianę, dyspersję, kwartyle Q1 i Q3 dla częstości błędnej klasyfikacji oraz macierz pomyłek dla wybranych przypadków, potwierdzają tezy rozprawy.

Należy podkreślić, że opracowane algorytmy segmentacji i wyznaczania cech

morfometrycznych zostały oparte nie tylko na uniwersalnych podstawach metodologicznych, ale także (a może przede wszystkim) wykorzystują gruntowną dziedzinową wiedzę medyczną z zakresu morfometrii i właściwości biochemicznych skórnych struktur tkankowych. Powoduje to, iż zaproponowane metody są także dziedzinowo-zorientowane, a tym samym w pełnym stopniu wykorzystują specyfikę rozwiązywanego problemu diagnostycznego.

4. Uwagi szczegółowe

Obok niewątpliwie wartościowych wyników, które wymieniłem i które przesądzą o pozytywnej ocenie całości, dostrzegam w recenzowanej rozprawie również niedostatki. Można je ująć w następujące punkty:

1. W badaniach eksperymentalnych zabrakło analizy statystycznej uzyskanych wyników porównawczych. Sama różnica w wartości kryterium jakości klasyfikacji nie oznacza, że jest to różnica istotna statystycznie. Można tutaj skorzystać z gotowych procedur (sprowadzają się one do odpowiednich testów istotności dla weryfikacji hipotez) – np. przedstawionych w pracy: T.G.Dietterich, Approximate statistical tests for comparing supervised classification learning algorithms, *Neural Computation* 10: 1895-1923, 1998.
2. Dla eksperymentalnych badań porównawczych w ujęciu dwuklasowym (zmiana złośliwa i łagodna) interesująca byłaby analiza ROC (*Receiver Operating Characteristic*) przedstawiająca wykres działania klasyfikatora dychotomicznego dla różnych jego punktów pracy (np. wartości progów dla porównania funkcji klasyfikujących) w układzie współrzędnych FP (*False positive*) i TP (*True positive*).
3. Brak informacji o ilości zdjęć preparatów w poszczególnych bazach, także w podziale na klasy.
4. Nieprecyzyjny opis badań eksperymentalnych: ilo klasowy problem był rozpatrywany (tabele 7.3 i 7.4, rysunki 7.3 i 7.4), jakie medyczne znaczenie miały poszczególne klasy, jak trenowano klasyfikatory (np. NB) dla 0-1 funkcji strat, a jak dla funkcji strat zależnej od rodzaju błędu (str. 123), co to znaczy, że w algorytmie NB estymowano prawdopodobieństwa a priori i warunkowe rozkłady cech w klasach metodą MAP (str. 119), jakie jest uzasadnienie na wyznaczenie kwartyli Q1, mediany, Q3 dla stopy błędu i jakie wnioski można wyciągnąć na podstawie uzyskanych wartości?
5. Co prawda, uzyskane rezultaty poprawnej klasyfikacji (70%-75%) potwierdzają tezę I, która została sformułowana dość ostrożnie (jest możliwa ocena stopnia złośliwości), ale czy można zarekomendować zaproponowaną metodę do praktycznych zastosowań klinicznych?
6. Przyjęta forma spisu literatury (według kolejności odwołań) jest mało wygodna, gdyż bardzo trudno jest znaleźć w spisie interesującą publikację.

Podkreślić należy, iż wymienione usterki nie podważają zasadniczych tez i wyników pracy, a są jedynie niedoskonałościami w sposobie ich prezentacji.

5. Podsumowanie recenzji - ocena rozprawy

Reasumując stwierdzam, iż mgr inż. Paweł Kłeczek wykazał się dużą wiedzą z zakresu obrazowania medycznego oraz cyfrowego przetwarzania informacji obrazowej, a także opanowaniem i właściwym posługiwaniem się odpowiednim warsztatem naukowym.

Przedstawiona praca zawiera poprawnie sformułowany i rozwiązany problem badawczy oraz stanowi istotny i wartościowy wkład w dziedzinę automatycznego wspomaganie medycznej diagnostyki obrazowej. Zawarte w niej rezultaty obejmujące wszechstronną analizę problemu diagnozowania zmian skórnych na podstawie obrazów preparatów histopatologicznych, oryginalne algorytmy segmentacji obszarów zainteresowania, wyznaczenie cech morfometrycznych oraz ich wykorzystanie w zaimplementowanych klasyfikatorach, a także przeprowadzenie bogatych badań eksperymentalnych i analiz porównawczych na obrazach rzeczywistych, są oryginalne i zostały przedstawione w sposób wyczerpujący na wysokim poziomie formalnego opisu. Uważam, że praca doktorska Pana mgr. inż. Pawła Kłeczka lokująca się w dyscyplinie Automatyka, Elektronika i Elektrotechnika spełnia wymogi stawiane pracom doktorskim przez odpowiednią Ustawę i wnioskuję o dopuszczenie jej do publicznej obrony.

Dodatkowo, z uwagi na:

1. wyjątkowo kompleksowe podejście do rozwiązywanego problemu, w którym poprzez zgłębienie wiedzy medycznej obejmującej histopatologię skóry wraz z analizą morfologii skórnych struktur tkankowych opracowany został dziedzinowo-zorientowany zestaw oryginalnych algorytmów do automatycznej segmentacji i wyznaczania cech morfometrycznych o kluczowym znaczeniu dla diagnostyki stopnia złośliwości nowotworu,
2. przeprowadzenie bogatych i różnorodnych badań eksperymentalnych na obrazach rzeczywistych,
3. wysoki poziom edytorski pracy ze szczególnie starannie przygotowaną sferą ilustracyjną,
4. opublikowanie wyników rozprawy w 4 pracach, w tym w dwóch artykułach w czasopiśmie z listy JCR,

wnioskuję o wyróżnienie pracy doktorskiej Pana mgr. inż. Pawła Kłeczka.

