

Łódź, 15 sierpnia 2022 r.

prof. dr hab. inż. Michał Strzelecki
Instytut Elektroniki Politechniki Łódzkiej
ul. Wólczańska 211/215, 90-924 Łódź
michal.strzelecki@p.lodz.pl

Recenzja rozprawy doktorskiej mgr inż. Dominiki Sułot
**„Development of machine learning algorithms operating on small datasets to assist
in the diagnosis of eye diseases”**
 (“Opracowanie metod uczenia maszynowego operujących na małych zbiorach danych
w celu wspomagania diagnozy chorób oczu”),
promotorzy prof. dr hab. inż. D. Robert Iskander, dr David Alonso Caneiro

Podstawą niniejszej recenzji jest pismo przewodniczącej Rady Dyscypliny Naukowej Inżynieria Biomedyczna Politechniki Wrocławskiej prof. dr hab. inż. lek. Haliny Podbielskiej z dn. 1 lipca 2022 r. powołujące mnie na recenzenta w postępowaniu o nadanie stopnia doktora mgr inż. Dominice Sułot, prowadzonym w dyscyplinie inżynieria biomedyczna.

Rozwój algorytmów analizy danych medycznych, w tym danych pochodzących z urządzeń obrazowych, spowodował możliwość pozyskania dodatkowych ilościowych informacji o badanych narządach i organach. Informacje te, uzyskane dzięki tzw. badaniom dodatkowym (obejmującym m.in. analizę biosygnatów lub obrazów biomedycznych) stanowią uzupełniające, niemniej niezwykle cenne, źródło informacji dla lekarzy umożliwiając częstokroć postawienie bardziej wiarygodnej, obiektywnej i powtarzalnej diagnozy. Jednak do właściwego działania takich algorytmów potrzebne są duże zbiory danych, nie zawsze dostępne w wielu dziedzinach medycyny. Potrzeba posiadania takich zbiorów dotyczy szczególnie metod uczenia maszynowego, a tym bardziej ich podzbioru – metod uczenia głębokiego. Duża ilość danych jest niezbędna do poprawnego wytrenowania algorytmów klasyfikacji, segmentacji, rozpoznawania danych, aby mogły one uzyskać wymagany poziom generalizacji. Jednakże w przypadku analizy obrazów biomedycznych pozyskanie odpowiednio licznych zbiorów danych nie zawsze jest możliwe. Doktorantka w swojej rozprawie zajęła się

poszukiwaniem takich algorytmów analizy danych (uzyskanych głównie z obrazów ze skaningowego oftalmoskopu laserowego SLO oraz optycznego tomografu koherentnego OCT), które mimo ograniczonych zasobów danych obrazowych umożliwią uzyskanie wiarygodnych informacji diagnostycznych. Tematykę rozprawy uważam zatem za bardzo ważną i aktualną, ponieważ opracowane metody mogą znaleźć potencjalne zastosowanie w praktyce klinicznej we wspomaganiu diagnostyki jaskry. Autorka analizowała różne algorytmy, związane z doбором właściwych parametrów w celu wykrywania jaskry, wyboru optymalnego zestawu takich parametrów oraz segmentacji wybranych struktur w obrazach OCT. W efekcie w sposób poprawny sformułowano 4 tezy pracy:

1. Obrazy laserowej oftalmoskopii skaningowej o niskiej rozdzielczości zawierają istotne informacje, które mogą wspomagać diagnozę jaskry, zwłaszcza w połączeniu z wykorzystaniem algorytmów głębokiego uczenia.
2. Ilość danych podawanych na wejście sieci neuronowej może mieć wpływ na jej wydajność, szczególnie w przypadku segmentacji błony Brucha, rodzaj danych wejściowych może mieć wpływ na ostateczne wyniki działania wytrenowanego modelu sieci.
3. Algorytmy uczenia maszynowego mogą być z powodzeniem stosowane w przypadkach, kiedy ilość danych jest ograniczona pod warunkiem zastosowania dodatkowych metod uwzględniających niewielkie zbiory danych.
4. Zastosowanie ortogonalizacji w sieciach splotowych może mieć pozytywny wpływ na ich wydajność, szczególnie gdy model jest trenowany dla niewielkich zbiorów danych.

Należy jednak zauważyć, że teza nr 1 (sformułowana w dość ogólny sposób) została już częściowo udowodniona, ponieważ istnieją publikacje opisujące zastosowanie sieci konwolucyjnych do wspomagania wykrywania jaskry w obrazach SLO.

Rozprawa została podzielona na przedmowę, gdzie przedstawiono motywację do przeprowadzonych badań i tezy pracy oraz 4 rozdziały. Pierwszy rozdział ma charakter wprowadzający w tematykę rozprawy i stanowi przegląd podstawowych metod uczenia maszynowego oraz uczenia głębokiego, opisuje także problemy związane ze stosowaniem tych metod w przypadku ograniczonej ilości danych wejściowych oraz podejścia umożliwiające częściowe uporanie się takimi problemami. Przedstawiono również przegląd literatury dotyczący wykorzystania metod uczenia maszynowego do wspomagania diagnostyki chorób oczu. Rozdział 2 definiuje cztery problemy badawcze, którymi Autorka zajmowała się podczas prowadzonych badań. W rozdziale 3 przedstawiono i omówiono wyniki tych badań. Rozprawa byłaby niewątpliwie bardziej czytelna, gdyby każde zagadnienie badawcze zostało opisane w osobnym rozdziale, który zawierałby definicję problemu, zastosowane metody, uzyskane wyniki oraz ich dyskusję. Obecnie czytelnik zagłębiający się w opisane w rozdziale 3 wyniki, musi często wracać do rozdziału 2, aby przypomnieć sobie tło danego problemu. Rozdział 4 zawiera wnioski z przeprowadzonych badań oraz planowane kierunki dalszych prac Doktorantki.

Pierwsze omawiane zagadnienie badawcze dotyczy metod detekcji jaskry pierwotnej otwartego kąta (jaskry prostej przewlekłej) w obrazach SLO, która jest najpopularniejszym typem jaskry występującym w Polsce. Do klasyfikacji obrazów wykorzystano różne parametry, zarówno anatomiczne jak grubość warstwy włókien nerwowych siatkówki jak i wyznaczone na podstawie analizy tekstury obrazu (wybrane cechy obliczone dla macierzy zdarzeń) czy wartości składowych głównych wyznaczone dla obrazu przekształconego do postaci

wektorowej. Zastosowano typowe klasyfikatory, jak wielowarstwowe perceptronowe sieci neuronowe (MLP), drzewa decyzyjne, klasyfikator najbliższego sąsiada (k-NN) czy maszynę wektorów nośnych (SVM). Zastosowano również metody uczenia głębokiego, w szczególności sieć o architekturze Inception-v3, z augmentacją oraz bez augmentacji danych, również badano efekty uzyskane z użyciem uczenia transferowego. Zgodnie z oczekiwaniami najlepsze wyniki uzyskano dla sieci głębokich – zarówno dla sieci Inception-v3 jak i dla zmodyfikowanej przez Doktorantkę architektury (ograniczono liczbę parametrów sieci). Ta ostatnia metoda umożliwiła uzyskanie dokładności o ok. 0.1 większej niż dla najlepszej metody „klasycznego” uczenia maszynowego. Wszystkie stosowane metody zostały poprawnie zwalidowane przy użyciu 5- i 10-krotnej walidacji krzyżowej (oraz bardziej wymagającej wersji tej metody określanej jako cross-validation ensemble użytej dla sieci głębokich).

Kolejny problem badawczy dotyczył segmentacji w obrazach OCT przerwy membrany Brucha w obszarze przejścia nerwu optycznego przez siatkówkę. Szerokość tej przerwy jest jednym z istotnych parametrów diagnostycznych w ocenie jaskry. Przetestowano tu 3 rodzaje danych wejściowych: pojedyncze kolumny obrazu OCT, zestawy takich kolumn oraz całe obrazy. Pierwsze dwa podejścia są oryginalnym wkładem Autorki w metodę przygotowania danych wejściowych. Zbadano również wersję „trójwymiarową” dzięki wykorzystaniu kolumn (lub ich zestawów) z kolejnych dwóch sąsiednich skanów OCT. Do segmentacji, w zależności od rodzaju danych wejściowych, wykorzystano odpowiednio wielowarstwowy perceptron, sieć konwolucyjną oraz dla całych obrazów typową dla takich zastosowań sieć o architekturze U-Net. Uzyskane wyniki są ciekawe – w przypadku najbardziej ograniczonej ilości danych wejściowych uzyskana wartość współczynnika Dice’a była zaledwie o 0.04 mniejsza niż dokładność segmentacji uzyskana dla pełnych obrazów z użyciem sieci U-Net (model ten okazał się najlepszy w stosunku do testowanych jego różnych modyfikacji, m.in. z użyciem sieci rekurencyjnych). Bardzo dobre wyniki segmentacji (błąd lokalizacji membrany w zakresie 1 piksela) potwierdziły również inne uzyskane wartości błędów.

Następne zagadnienie analizowane przez Doktorantkę dotyczy problemu redukcji liczby parametrów opisujących badane obiekty, w tym przypadku byli to pacjenci z jaskrą pierwotną, podejrzeniem jaskry i zdrowi. Rozpatrywano problem klasyfikacji pacjentów na podstawie 48 cech wykorzystywanych w diagnostyce jaskry (m.in. parametry morfologiczne dysku optycznego, grubość warstwy włókien nerwowych siatkówki, wartości ciśnienia wewnątrzgałkowego). Do wyboru cech zaproponowano oryginalną metodę należąca do grupy metod selekcji parametrów, która losuje poszczególne cechy na podstawie prawdopodobieństwa proporcjonalnego do wartości statystyki testu Fishera uzyskanej dzięki jednoczynnikowej analizie wariancji. Zaletą tej metody jest brak eliminacji jakiegokolwiek z cech, choć ich wykorzystanie zależy od przydatności określonej wartością testu F. Do klasyfikacji wykorzystano zestaw klasyfikatorów (uczonych dla podzbiorów losowanych cech), zaś jako klasyfikatory bazowe użyto podobnie jak w pierwszym problemie badawczym MLP, k-NN, SVM oraz odmianę drzewa decyzyjnego. Wykazano, że zaproponowana metoda zapewnia najdokładniejsze wyniki klasyfikacji w porównaniu do dwóch metod referencyjnych (k-best i Random Subspace), niezależnie od liczby uwzględnionych cech. Pokazano również, że najbardziej skutecznym klasyfikatorem bazowym w zestawie klasyfikatorów jest SVM, szczególnie dla ograniczonej liczby cech.

Ostatni problem poruszany w pracy dotyczył zastosowania ortogonalnych sieci konwolucyjnych (OCNN), które umożliwiają ograniczenie liczby warstw splotowych przez pominięcie warstw o zbliżonych parametrach. Sieci takie zastosowano do klasyfikacji obrazów

zawierających cyfry (baza CIFAR-10). Badano tu wpływ redukcji rozmiaru zbioru treningowego na dokładność klasyfikacji przy użyciu sieci ortogonalnych i klasycznych. Potwierdzono, że sieć OCNN zapewnia lepsze wyniki klasyfikacji dla mniejszych zbiorów danych, przy czym różnica ta maleje wraz ze wzrostem liczebności zbioru treningowego. Ten fragment rozprawy jest nieco niespójny z resztą pracy. Uzyskanych wniosków nie można rozszerzyć dla przypadku analizy danych stosowanych w okulistyce (lub innych dziedzinach medycyny), ponieważ w eksperymentach nie wykorzystano żadnych obrazów biomedycznych (lub danych uzyskanych dla takich obrazów). Kolejny eksperyment polegał na porównaniu działania ortogonalnej i klasycznej sieci U-Net do segmentacji wewnętrznych warstw siatkówki w obrazach OCT zwyrodnienia plamki żółtej spowodowanej wiekiem. Również w tym przypadku wykazano, że dla mniejszych zbiorów obrazów treningowych sieć OU-Net zapewnia dokładniejsze wyniki segmentacji niż jej klasyczny odpowiednik.

Liczący 123 pozycje wykaz literatury obejmuje najważniejsze pozycje literatury światowej dotyczące tematyki związanej z rozprawą. Wykaz ten zawiera również 5 publikacji współautorskich zamieszczonych w materiałach konferencyjnych, w monografii (jako rozdział) oraz w czasopismach OSA Continuum (IF = 0.67) i PloS One (IF = 3.75).

Podsumowując merytoryczną ocenę rozprawy stwierdzam, że Doktorantka osiągnęła założone cele badawcze oraz udowodniła postawione tezy dzięki zbadaniu właściwości wielu metod uczenia maszynowego, w tym własnych, w warunkach ograniczonej liczby danych wejściowych. Umożliwiło to określenie optymalnych metod analizy danych dla takich przypadków, głównie w zagadnieniach wspomagania diagnostyki jaskry. Należy podkreślić, że osiągnięcie celów pracy wymagało wykazania się przez mgr Dominikę Sułot szeroką wiedzą nie tylko z dyscypliny inżynieria biomedyczna (implementacja różnych metod uczenia maszynowego w analizie danych obrazowych stosowanych w okulistyce) ale i z dyscypliny informatyka techniczna i telekomunikacja (opracowanie, implementacja, modyfikacja i walidacja algorytmów selekcji cech i uczenia głębokiego). Do najważniejszych osiągnięć Autorki pracy zaliczam:

1. Opracowanie i implementację nowych metod segmentacji przerwy membrany Brucha w obszarze przejścia nerwu optycznego przez siatkówkę w obrazach OCT z użyciem pojedynczych A-skanów (wybranych kolumn obrazu OCT lub ich kilku zestawów). Dokładność tej metody okazała się zbliżona do klasycznej, wykorzystującej cały obraz, natomiast zastosowane do segmentacji sieci neuronowe zawierają mniej parametrów niż sieć U-Net użyta do segmentacji pełnoobrazowej.
2. Opracowanie oryginalnej metody selekcji cech umożliwiającej ponadto wykorzystanie wszystkich dostępnych cech z odpowiednio przypisanymi wagami i uczenie za pomocą ich losowanych podzbiorów zestawu klasyfikatorów. Wykazano, że metoda jest skuteczna przy klasyfikacji jaskry na podstawie szeregu różnych parametrów. Proponowana metoda zapewniła najdokładniejszą klasyfikację w stosunku do dwóch metod referencyjnych przy zastosowaniu zestawu klasyfikatorów wykorzystujących różne rodzaje klasycznych klasyfikatorów bazowych.
3. Implementację i analizę wielu metod uczenia maszynowego (w tym uczenia głębokiego) oraz metod wstępnego przetwarzania danych wykazując, że odpowiednie połączenie takich metod stanowi skuteczne narzędzie wspomagania diagnostyki chorób oczu nawet w przypadku ograniczonej ilości danych wejściowych.

Praca jest napisana poprawnym językiem angielskim, nie budzi również zastrzeżeń od strony redakcyjnej. Drobne uwagi dotyczą używania skrótu SVC zamiast SVM dla klasyfikatora support vector machines, zamiany opisu górnej i dolnej części tabeli 3.1 oraz użycia słowa „layers” zamiast „neurons” na str. 63 przy opisie klasyfikatora MLP.

Lektura pracy nasuwa również kilka przedstawionych poniżej uwag polemicznych i dyskusyjnych.

1. Omówienie metod redukcji liczby cech (str. 25 oraz rys. 1.15) nie jest w pełni precyzyjne, ponieważ jako sposób ograniczenia wymiarowości zbioru cech (projekcji) podano jedynie metodę analizy składowych głównych (PCA). Tymczasem nie jest jedyną metodą projekcji cech. Ponadto, metoda PCA, mimo generowania nieskorelowanego zestawu cech, nie zawsze prowadzi do uzyskaniu podzbioru parametrów zapewniających najlepszą dyskryminację danych. W zależności od rozkładu wartości tych cech kierunek największych zmian zbioru danych nie musi odpowiadać kierunkowi zapewniającemu ich najlepszą separowalność. Dlatego z punktu widzenia rozdzielania danych często lepsze wyniki generują inne metody redukcji wymiarowości, jak liniowa lub nieliniowa analiza dyskryminacyjna. Często również dobre wyniki zapewnia łączne zastosowanie metod selekcji i projekcji cech.
2. W przypadku klasyfikacji obrazów SLO jaskry zastosowano kilka wybranych parametrów tekstury wyznaczonych dla macierzy zdarzeń. Zabrakło uzasadnienia, dlaczego wybrano akurat takie parametry z bardzo szerokiego zakresu cech tekstury obrazu oraz w jaki sposób skonstruowano macierze GLCM.
3. W pracy zabrakło porównania uzyskanych wyników klasyfikacji i segmentacji analizowanych obrazów z opublikowanymi wynikami uzyskanymi przez innych badaczy. Zdają sobie sprawę, że pełne porównanie nie jest możliwe choćby ze względu na różne zbiory danych, dla których przeprowadzono badania. Jednak zawsze należy taką analizę wykonać, aby porównać skuteczność stosowanych metod z innymi opisanymi w literaturze.
4. W pracy brakuje też podania czasów uczenia poszczególnych klasyfikatorów, w tym sieci głębokich. Jest to istotny parametr, który również służy do oceny stosowanych metod analizy danych.
5. Ponadto zabrakło szerszego wykorzystania stosowanych w pracy metod do posiadanych danych. Przykładowo, Doktorantka badając sieci ortogonalne nie zastosowała ich do klasyfikacji obrazów OCT jaskry (pierwszy problem badawczy). Byłoby bardzo interesujące porównanie wyników działania tych sieci z użytą siecią Inception-v3 (i jej modyfikacjami). Podobnie, zaproponowana metoda selekcji cech powinna zostać również użyta w tym samym problemie klasyfikacji jaskry. Należało dokonać fuzji wszystkich parametrów (teksturalnych, grubości RNFL, uzyskanych dzięki PCA) i zastosować metodę omówioną w rozdziale 2.3. Również w tym przypadku porównanie uzyskanych dokładności klasyfikacji byłoby niezwykle ciekawe i pouczające.

Wszystkie moje uwagi krytyczne i dyskusyjne w żadnym stopniu nie wpływają na jednoznacznie pozytywną ocenę recenzowanej pracy, która ma charakter interdyscyplinarny, co zasługuje na szczególne uznanie (Autorka wykazała osiągnięcia w dyscyplinach inżynieria

biomedyczna oraz informatyka techniczna i telekomunikacja). Stwierdzam, że praca „Development of machine learning algorithms operating on small datasets to assist in the diagnosis of eye diseases” spełnia wymagania stawiane rozprawom doktorskim zgodnie z ustawą Prawo o szkolnictwie wyższym i nauce z dnia 20 lipca 2018 r. z późniejszymi zmianami, oraz wnioskuję o przyjęcie tej rozprawy i dopuszczenie mgr inż. Dominiki Sułot do publicznej obrony.

Michał Stalke