

ROITT - upr. 24.01.2023
M. Skon



Prof. dr hab. inż. Khalid Saeed
Wydział Informatyki
Politechnika Białostocka
ul. Wiejska 45A, 15-351 Białystok
Tel. (+48-85) 746 9196
Fax: (+48-85) 746 9057
k.saeed@pb.edu.pl

Białystok, 12.01.2023 r.

RECENZJA rozprawy doktorskiej
mgr. inż. Kamila Książka
z Wydziału Automatyki, Elektroniki i Informatyki
Politechniki Śląskiej

zatytułowanej "Optimization of deep learning network architectures for
hyperspectral data classification"

Promotor:

dr hab. inż. Przemysław Głomb
Wydział Automatyki, Elektroniki i Informatyki
Politechnika Śląska

Niniejszą recenzję przygotowałem na zlecenie zawarte w piśmie z dnia 21.11.2022, które otrzymałem od profesora Andrzeja Polańskiego przewodniczącego Rady Dyscypliny Informatyka Techniczna i Telekomunikacja Politechniki Śląskiej na podstawie uchwały Rady, podjętej dnia 25 października 2022 r.

I. Omówienie zawartości rozprawy

Praca doktorska mgr. inż. Kamila Książka miała na celu udowodnienie tezy, że optymalizacja architektur sieci głębokiego uczenia i metody reinicjalizacji wag zwiększają wydajność sieci w zastosowaniu do danych nadwiedmowych (hiperspektralnych) - *Hyperspectral data*. Wyniki badań doktoranta pozwoliły udowodnić tezę i uzasadniły, że optymalizacja architektur sieci głębokiego uczenia zwiększa wydajność sieci w zastosowaniu do danych hiperspektralnych oraz że metody reinicjalizacji wag także zwiększają wydajność. Wiodący cel rozprawy został dogłębnie przeanalizowany i przedstawiony wraz z przeprowadzonymi eksperymentami. Rozprawa doktorska napisana jest w języku angielskim. Zawiera 144 strony tekstu, rysunków i ilustracji. Składa się z pięciu rozdziałów oraz załącznika i bibliografii.

Rozdział 1. „Wstęp”, gdzie autor przedstawia wprowadzenie do tematyki rozprawy – omawia tezy rozprawy, które udowadnia w swojej pracy. Wskazuje architektury stosowane w dalszych badaniach oraz podstawowe pojęcia związane z obrazowaniem hiperspektralnym.

W rozdziale 2. „Klasyfikacja plam krwi na podstawie obrazów hiperspektralnych z użyciem sieci głębokiego uczenia” doktorant przeprowadza optymalizację sieci głębokiego uczenia na przykładzie zbioru danych do klasyfikacji plam krwi oraz innych wizualnie podobnych do krwi substancji. Autor wykonał szereg eksperymentów, wliczając w to scenariusze transduktywne i induktywne (*transductive and inductive learning*) skupiając się na mieszaninach widm pikseli, z których wnioskuje, iż zmiana warunków akwizycji obrazu może negatywnie wpływać na pracę klasyfikatora.

Rozdział 3. „Badanie metod inicjalizacji wag dla autokoderów” skupia się na badaniu liniowych autokoderów do rozmieszczania spektralnego w celu odzyskania oryginalnych widm, które są mieszaninami różnych substancji oraz proporcji, z jakimi występują w mieszaninie. Doktorant przeprowadził wiele eksperymentów z użyciem różnych architektur, zbiorów danych, funkcji straty oraz metod inicjalizacji wag w sieciach i zoptymalizował hiperparametry sieci w kierunku badania stabilności autokoderów. Autor wnioskował, że inicjalizacja wag w sieciach neuronowych istotnie wpływa na błąd rekonstrukcji pikseli po zakończeniu treningu sieci.

Ostatnim rozdziałem merytorycznym jest Rozdział 4. „Metody reinicjalizacji sieci dla poprawy wydajności autokoderów”. Doktorant opisuje próbę rozwiązania problemu związanego z identyfikacją zjawiska martwych aktywacji i neuronów, co powoduje niestabilność sieci. Autor zaprezentował trzy metody reinicjalizacji wag sieci, które miały minimalizować negatywny skutek martwych aktywacji na wydajność sieci. W każdej z nich badana była proporcja martwych aktywacji w stosunku do niezerowych aktywacji w danych warstwach lub dla danych neuronów. Doktorant przeprowadził eksperymenty i potwierdził skuteczność swoich rozwiązań w danych hiperspektralnych.

Rozdział 5. „Wnioski” zamyka część merytoryczną i podsumowuje wszystkie aspekty rozpatrywane w pracy. Doktorant konkluduje, że wyniki otrzymane w Rozdziałach 2. i 3. potwierdziły pierwszą część tezy rozprawy - *optymalizacja architektur sieci głębokiego uczenia zwiększa wydajność sieci w zastosowaniu do danych hiperspektralnych*. Ponadto doktorant wnioskuje, że zaproponowane trzy metody reinicjalizacji wag zmniejszają negatywny wpływ problemu martwych aktywacji. W ten sposób potwierdza to drugą część tezy rozprawy - *metody reinicjalizacji wag zwiększające wydajność sieci w zastosowaniu do danych hiperspektralnych*.

Bibliografia zawiera 177 pozycji wybranych referatów i artykułów z literatury światowej pokazującej stan wiedzy i odzwierciedlającej dostateczną wiedzę doktoranta. Dodatkowo został opracowany załącznik z materiałami, które moim zdaniem mogły być umieszczone w głównej części pracy.

Interesującym aspektem redakcyjnym pracy jest to, że każdy rozdział merytoryczny kończy się podsumowaniem istotnych osiągnięć wymienionych w danym rozdziale.

II. Opinia o rozprawie doktorskiej

Rozprawę doktorską mgr. inż. Kamila Książka ocenię w dwóch płaszczyznach: technicznej i merytorycznej oraz klarowności i czytelności rozprawy.

A. Techniczne brzmienie i merytoryczna kompletność rozprawy

Praca badawcza doktoranta ma charakter eksperymentalno-statystyczny. Autor wykazał w swojej pracy umiejętność przekonywującego przedstawienia uzyskanych przez siebie wyników. Rozprawa stanowi oryginalne rozwiązanie postawionych zadań oraz osobisty wkład doktoranta w rozwój metod optymalizacji architektur sieci głębokiego uczenia. Autor zaprezentował statystyczne opracowania nowych metod w tej dziedzinie. Zagadnienia z ilustrującymi przykładami są podane szczegółowo. Algorytmy, rysunki i tabele są czytelne i starannie opracowane. Przykładowe podejścia innych badaczy naukowych zostały wybrane poprawnie, przedstawiające aktualny stan wiedzy z zakresu metod reprezentacji danych wielowymiarowych, chociaż zabrakło analiz i podejść pewnych zespołów światowych i krajowych – wymienię niektóre poniżej. Istotne dla tematyki zagadnienia zostały omówione w sposób czytelny i przejrzysty. Nie znalazłem błędów merytorycznych, chociaż są błędy językowe.

Uważam, że mgr inż. Kamil Książek osiągnął wyznaczony mu cel rozprawy doktorskiej, która wnosi nowe aspekty do nauk technicznych w zakresie informatyki.

Pomimo wysiłków autora, aby praca prezentowała jego osiągnięcia w sposób klarowny, praktycznie niemożliwe do uniknięcia są drobne usterki, których przykłady zostały wymienione poniżej.

(1) Cytowana literatura jest interesująca, ale niestety nie zawiera bardzo ważnych pozycji takich jak:

- Suju Rajan, Joydeep Ghosh, and Melba M Crawford, *An Active Learning Approach to Hyperspectral Data Classification*, IEEE Trans. Geoscience and Remote Sensing, 46(4):1231–1242, 2008.
- Paweł Ksieniewicz, Bartosz Krawczyk, Michał Woźniak, *Ensemble of Extreme Learning Machines with Trained Classifier Combination and Statistical Features for Hyperspectral Data*, Neurocomputing, 2017, DOI: 10.1016/j.neucom.2016.04.076.
- Rozprawa doktorska Pawła Ksieniewicza z Politechniki Wrocławskiej "*Multidimensional data representation and analysis*", gdzie zaprezentowano szereg algorytmów, przykładów i zastosowań. Autor przedstawił metody przetwarzania danych z obrazowania nadwiedmowego (*Hyperpspectral Imaging*) w obszarze nienadzorowanego, jak i nadzorowanego uczenia maszynowego.

(2) Zastosowane bazy danych są interesujące. Zabrakło jednak istotnych innych baz, które warto byłoby wykorzystać do swoich badań, na przykład:

- *Salinas Dataset*: collected over the Valley of Salinas, Southern California, 1998.
- *Indian Pines*: The scene gathered by Aviris over the Indian Pines test site in North-Western Indiana.
- *Pavia University*: acquired by the Rosis sensor during a flight campaign over Pavia, Northern Italy.
- *Kennedy Space Centre*: Acquired by the Nasa Aviris spectrometer, over the Kennedy Space Center (ksc), Florida, 1996.

Wszystkie te i inne znane pozycje to konwencjonalne zbiory danych i testy porównawcze często spotykane w literaturze (jako przykład patrz *Weles* - <https://github.com/w4k2/weles>, *A Machine Learning library*). Ich zastosowanie wzbogaciłoby pracę doktoranta i dałoby szerokie aspekty do porównania ze stanem wiedzy w tej dziedzinie.

Powyżej wymienione uwagi mają charakter dyskusyjny. Nie obniża to wartości rozprawy, jednakże chciałbym, żeby autor ustosunkował się do tych kwestii na obronie.

B. Klarowność i czytelność rozprawy

Rozprawa napisana w języku angielskim, jest przejrzysta i czytelna, chociaż nie brakuje błędów gramatycznych i edytorskich (na przykład na str. 1. „*AI is an rapidly developing field*” lub „*The Algorithm 1, which presented ...*” na str. 102). Pewnym usprawiedliwieniem jest to, że nie jest to język ojczysty doktoranta. Algorytmy, twierdzenia, wymagane rysunki i tabele są prawidłowo opracowane, chociaż nie wszędzie są dokładnie zaprezentowane. Aktualny stan wiedzy dotyczący tematyki optymalizacji architektur sieci głębokiego uczenia dla danych hiperspektralnych w celu ich klasyfikacji został przedstawiony poprawnie.

Mam wrażenie, że autor nie ustrzegł się pewnej liczby nieścisłości, mogących mieć wpływ na zrozumienie tekstu rozprawy. Niektórych szczegółów nie przedstawiono standardowo, podczas gdy inne zagadnienia można było sformułować trochę inaczej lub w ogóle nie umieszczać w pracy. Uważam, że jako recenzent powinienem wskazać doktorantowi możliwość dalszego doskonalenia pracy. Oto moje uwagi – raczej redakcyjne:

- Figure 1.1: Rysunek reflektancji powinien mieć zaznaczone na osi Y wartość 1.0, a na osi X wartość B jako maksymalną liczbę bandów, żeby miał charakter bardziej poglądowy. Na rysunku dobrze byłoby wskazać, że dla roślinności jest inna reflektancja niż np. dla zbiornika wodnego. Brak wyjaśnienia w tekście, że sensor mierzy nie tylko faktyczną reflektancję, ale również promieniowanie rozproszone, które dociera do obiektu. Takie wyjaśnienie znajduje się dopiero na stronie 13.

- Na str. 3 autor jedynie wspomina o błędach wynikających z atmosfery. Brakuje krótkiego scharakteryzowania, czym jest *Radiometric Correction*, *Geometric Correction* oraz *Atmospheric Correction*.
- W pracy brak odniesień do klasycznych metod, jak *Spectral Angle Mapper*, które pojawia się dopiero na stronie 69.
- Autor stosuje niestandardowe sposoby walidacji (str. 22) "*The Authors proposed an improvement of the cross-validation approach in which the validation error of samples is weighted by the ratio of test and training distributions*". Proszę o odniesienie się do tej kwestii.
- Figure 2.1: Rysunek w tej formie niewiele mówi o rozkładzie badanych próbek. Dodawanie + std oraz -std dałoby większe wyjaśnienie.
- DeepHyperX: Informacja o własnej implementacji sieci i kontrybucji w otwarte oprogramowanie wspomniane jest na stronie 37. Przy czym różnica między frameworkami nie jest dokładnie opisana. Na str. 26 wspomniano o DHX, a na str. 37 jest mowa o modyfikacji implementacji, jednak bez szczegółów.
- W podrozdziale 2.5.1 autor pisze "*The spatial resolution of the output images is 696×520 while the number of spectral bands is equal to 128. According to suggestions from [133], some spectral bands, i.e. 0-4, 48-50 and 122-128, were removed due to noise, and finally, 113 bands were left*". Jakie jest uzasadnienie na wyrzucenie 15 bandów tj. około 11% zbioru danych?
- Podrozdział 2.5.4 *Evaluation metrics*: Doktorant powołuje się na trzy metryki – OA, AA oraz Cohen's Coefficient. Metryki takie jak czułość bądź selektywność, albo chociażby miara łącząca F1-score dodałyby bogatsze wyniki badań.
- Brak wartości numerycznych, co przyjęto za $n_{samples}^{min}$ powoduje niejasność w zrozumieniu, co autor chciał wyjaśnić (str. 36). Ponadto doktorant pisze "*Finally, other pixels that are not within k-neighborhood of pixels chosen for the training set or near the image borders, are assigned to the test set. We assumed that $k = 2$ because the maximum patch size for the architectures used is equal to 5×5* ". Podział na zbiór treningowy i testowy nie jest klarowny. Brak wykorzystania pikseli sąsiadujących wygląda jak celowe zmniejszenie liczebności zbioru testowego. Brak zestawienia, ile pikseli wykorzystano jako testowe lub ile jest tablic o określonym wymiarze. Grafiki z macierzami pomyłek na rysunkach 2.11-2.13 są znormalizowane. Dane te także nie znajdują się w załączniku ani w tekście dotyczącym informacji o tych rysunkach.
- Figure 1.1: Dobrze byłoby, aby obok rysunku dla HTC dodać również przypadek HIC. Wyniki na stronie 50 pokazują, że w tym przypadku występować może dość spora wariancja wyników (około 6%). Nie wiadomo, czy wynikają one z tzw. elementów odstających.
- Doktorant pisze o zaobserwowaniu zjawiska „*Vanishing Gradient* (str. 82). Jednak brakuje informacji, w jaki sposób autor przeciwdziałał temu zjawisku.
- Figure 3.14: Rysunek opisany w sposób niezrozumiały. Brak skali *colorbar*.

- Przy optymalizacji z wykorzystaniem funkcji kwadratu tj. MSE (str. 85) nie obserwuje się dużych różnic, które potęgują wartość funkcji kosztu. Warto byłoby dodać taką adnotację.

- Zawartość podrozdziału 4.7 (wykorzystanie próbek podpisu ręcznego - baza danych MNIST) jest raczej poza zakresem rozprawy. W ten sposób autor próbuje zweryfikować skuteczność swoich metod reinicjalizacji sieci przy innych zastosowaniach poza obrazami hiperspektralnymi. Uważam, że tego typu uogólnienie nie jest potrzebne, a więc podrozdział jest zbędny.

III. Merytoryczne osiągnięcia doktoranta

Mgr inż. Kamil Książek osiągnął wyznaczony cel rozprawy doktorskiej, która wnosi nowe aspekty do nauk technicznych w zakresie informatyki. Temat rozważań autora jest aktualnym kierunkiem badań.

Pan mgr Książek jest współautorem 11 artykułów w czasopismach recenzowanych oraz 5 recenzowanych referatów opublikowanych w materiałach konferencyjnych. Świadczy to o uznaniu osiągniętych wyników pracy naukowej doktoranta w dziedzinie reprezentacji danych wielowymiarowych w obszarze uczenia maszynowego. Ponadto doktorant ma inne sukcesy naukowe:

- Laureat nagrody *Student's Nobel Award* 2019 w dziedzinie IT.
- Stypendysta Ministerstwa Nauki i Edukacji w roku 2018/19.
- Laureat nagrody *Omnium Studiosorum Optimo Award* w roku 2019.
- Laureat nagrody Rektora przez 5 semestrów.

IV. Wnioski końcowe

Wystawiam pozytywną ocenę rozprawie doktorskiej mgr. inż. Kamila Książka pt. "*Optimization of deep learning network architectures for hyperspectral data classification*". Rozprawa stanowi oryginalne rozwiązanie problemu naukowego oraz osobisty wkład doktoranta w rozwój metod optymalizacji architektur sieci głębokiego uczenia oraz metod reinicjalizacji wag zwiększających wydajność sieci w zastosowaniu do danych hiperspektralnych.

Stwierdzam, że praca spełnia wymagania i warunki nakładane przez ustawę o stopniach naukowych.

Na tej podstawie wnioskuję o dopuszczenie autora rozprawy doktorskiej mgr. inż. Kamila Książka do jej obrony.

