RDITT – wpł. 20.01.2023
M. Kłosy

Kraków, 09.01.2023

Prof. dr hab. inż. Bogusław Cyganek
Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie
Wydział Informatyki, Elektroniki i Telekomunikacji
Instytut Elektroniki

Recenzja rozprawy doktorskiej *„Optimization of deep learning network architectures for hyperspectral data classification"* Pana magistra Kamila Książka

Wstęp

Recenzja rozprawy doktorskiej Pana magistra Kamila Książka zatytułowanej *„Optimization of deep learning network architectures for hyperspectral data classification*[1]". Praca powstała w roku 2022 na Wydziale Automatyki Elektroniki i Informatyki Politechniki Śląskiej. Promotorem pracy jest Pan dr hab. inż. Przemysław Głomb, natomiast promotorem pomocniczym jest Pan dr Krisztián Búza. Przygotowanie recenzji zostało wykonane na zlecenie Przewodniczącego Rady Dyscypliny Naukowej Informatyka Techniczna i Telekomunikacja, Pana prof. dr hab. inż. Andrzeja Polańskiego. Praca została napisana w języku angielskim. Recenzja sporządzona została w postaci odpowiedzi na pytania dotyczące rozprawy doktorskiej Pana magistra Kamila Książka.

1. *Jakie zagadnienie naukowe jest rozpatrzone w pracy i czy zostało ono dostatecznie jasno sformułowane prze autora? Jaki charakter ma rozprawa?*

Główne zagadnienie naukowe podjęte w pracy doktorskiej Pana magistra Kamila Książka dotyczy *opracowania oraz weryfikacji naukowej metod optymalizacji architektur głębokich sieci neuronowych do zagadnienia klasyfikacji danych wielospektralnych.*

Chyba najpowszechniejszym i najlepiej znanym typem obrazów cyfrowych są obrazy kolorowe składające się z trzech mapowań spektralnych, czyli składowej czerwonej (R), zielonej (G) oraz niebieskiej (B) – obserwujemy je codzienne na ekranach naszych smartfonów, monitorów, displejów samochodowych, czy też przydrożnych reklam. Tym niemniej, istnieją dziedziny które wymagają akwizycji obrazów obejmujących znacznie większe spektrum niż wspomniane RGB. Obrazy tego typu, zwane obrazami wielospektralnymi, czy też hiperspektralnymi, są stosowane głównie do analizy pokrycia powierzchni Ziemi, analizy upraw, czy też identyfikacji formacji geologicznych. Obrazy te

[1] Optymalizacja architektur sieci głębokiego uczenia na potrzeby klasyfikacji danych hiperspektralnych

wymagają zastosowania kamery wyposażonej w specjalny sensor, który umożliwia jednoczesną akwizycję nawet kilkuset wąskich zakresów promieniowana. Dzięki dużej rozdzielczości spektralnej zdjęcia otrzymane w powyższy sposób umożliwiają identyfikację materiałów składowych obserwowanych obiektów. Analizę tego typu oczywiście mogą przeprowadzać eksperci i specjaliści np. od analizy obrazów satelitarnych, ale we współczesnym świecie cyfrowym jest to możliwe za pomocą odpowiednio zaprojektowanych algorytmów, takich jak głębokie sieci neuronowe, będące prawdziwym przełomem technologicznych ostatnich lat. Projektowanie optymalnych architektur sieci tego typu w zastosowaniach do komputerowej analizy obrazów wielospektralnych jest jednym z zagadnień naukowych podjętych w rozprawie Pana magistra Kamila Książka, o czym będzie jeszcze mowa w dalszej części tej recenzji.

Wracając do specyfiki obrazów wielospektralnych należy wspomnieć, że duża liczba składowych spektralnych – między kilkadziesiąt a kilkaset zakresów – otrzymywana jest kosztem znacznie zubożonej rozdzielczości przestrzennej, co jest bezpośrednią konsekwencją fizycznych ograniczeń samego sensora obrazowego, choćby wynikających z możliwości przetwarzania energii o określonej wartości. W konsekwencji, każdy piksel obrazu wielospektralnego odpowiada nie jednemu, lecz wielu *zmieszanym* reflektancjom różnych materiałów w obserwowanej scenie, podczas gdy my przeważnie zainteresowani jesteśmy określeniem udziału tzw. składowych głównych (ang. *endmembers*), odpowiadającym poszczególnym i precyzyjnie określonym typom materiałów, takim jak woda, drzewa, ziemia uprawna, szosa itd. Stąd też potrzeba opracowania metod „od-miksowania" tych składowych, tj. identyfikacji rodzaju składowych głównych oraz parametrów modelu ich zmieszania. Jest to bardzo ważne zagadnienie naukowe w dziedzinie przetwarzania obrazów wielospektralnych, do którego też można zastosować głębokie sieci neuronowe. Zagadnienie to zostało również podjęte przez Pana Kamila Książka i zaprezentowane w Jego rozprawie doktorskiej. Obydwa wyżej wspomniane zagadnienia naukowe są niezmiernie trudne, między innymi ze względu na złożoność podjętych tematów oraz brak danych trenujących.

Podsumowując należy stwierdzić, że problemy badawcze podjęte przez Doktoranta Pana magistra Kamila Książka mają znaczenie zarówno w obszarze badań teoretycznych, jak i eksperymentalno aplikacyjnych. Praca ma więc charakter teoretyczno-doświadczalny.

2. *Czy w rozprawie przeprowadzono w sposób właściwy analizę źródeł, w tym literatury światowej, stanu wiedzy i zastosowań w przemyśle, czy świadczą one o dostatecznej wiedzy autora? Czy wnioski z przeglądu źródeł sformułowano w sposób jasny i przekonywujący?*

Zagadnienia badawcze podjęte w pracy Pana magistra Kamila Książka są niezwykle obszerne i obejmują wiele zagadnień. Każde z nich znajduje olbrzymie zainteresowanie w środowisku naukowym i wpływa na liczbę publikacji w periodykach światowych. Chodzi tu między innymi o zagadnienia przetwarzania obrazów, ze szczególnym uwzględnieniem obrazów wielospektralnych, jak również nowoczesne metody klasyfikacji i przetwarzania danych z wykorzystaniem wielorakich struktur głębokich sieci neuronowych. Szczególnie ta ostatnia dziedzina jest bardzo dynamicznie rozwijana, czego rezultatem są wręcz setki publikacji zaledwie w przeciągu ostatnich kilku lat. Choćby pobieżne śledzenie osiągnięć i prac, jak również próba usystematyzowania tak wielkiego strumienia publikacyjnego wymaga od badaczy olbrzymiej pracy. Doktorant, Pan Kamil Książek, dobrze radzi sobie z tym zadaniem – w swojej bardzo obszernej rozprawie odwołuje się on aż do 177 pozycji

literaturowych, z których w kilku przypadkach jest jednym z autorów. Świadczy to o zaangażowaniu Doktoranta, jak również o jego dobrym warsztacie naukowym.

3. *Czy autor rozwiązał postawione zagadnienia, czy użył właściwej do tego metody i czy przyjęte założenia są uzasadnione?*

Główne zagadnienia naukowe zaprezentowane przez Pana magistra Kamila Książka w Jego rozprawie doktorskiej dotyczą opracowania nowych metod optymalizacji architektur głębokich sieci neuronowych do klasyfikacji danych wielospektralnych. Teza pracy została sformułowana następująco (str. 3, w tłumaczeniu):

Optymalizacja architektur głębokich sieci uczących i metody reinicjalizacji wag poprawiają wydajność sieci neuronowych dla danych hiperspektralnych.

Tak postawiona teza została wykazana, zarówno od strony teoretycznej, jak również zweryfikowana licznymi eksperymentami z wykorzystaniem przez siebie stworzonych oraz dostępnych z innych źródeł danych testowych. W następnej kolejności, wyniki tych eksperymentów zostały opisanie oraz przeanalizowane i omówione w rozprawie. Ponadto, opracowane metody zostały opublikowane w znanych materiałach konferencyjnych oraz periodykach naukowych, o czym będzie jeszcze mowa w kolejnym punkcie tej recenzji.

Realizując wyżej wspomniane zagadnienia naukowe, Pan magister Kamil Książek, wykazał się głęboką wiedzą, dotyczącą przede wszystkim przetwarzania danych wielospektralnych oraz projektowania i optymalizacji różnorodnych architektur głębokich sieci neuronowych. Użyte przez niego metody i przyjęte założenia, które zostały przedstawione w rozprawie uważam za uzasadnione oraz w pełni odpowiadające poziomowi oraz osiągnięciom światowym w tej dziedzinie. Świadczą one o głębokiej wiedzy Doktoranta w dyscyplinie Informatyka Techniczna i Telekomunikacja.

4. *Na czym polega oryginalność rozprawy, co stanowi samodzielny i oryginalny dorobek Autora, jaka jest pozycja rozprawy w stosunku do stanu wiedzy czy poziomu techniki reprezentowanych przez literaturę światową?*

Jak już wspomniano, główne zagadnienia naukowe zaprezentowane przez Pana magistra Kamila Książka w Jego rozprawie doktorskiej dotyczą opracowania nowych metod optymalizacji architektur głębokich sieci neuronowych do (i) klasyfikacji danych wielospektralnych oraz (ii) odkodowywania danych wielospektralnych, czyli identyfikacji rodzaju składowych głównych oraz parametrów modelu w zjawisku mieszania obserwowanych sygnałów.

Główne osiągnięcia Doktoranta Pana Kamila Książka w tych dziedzinach i zaprezentowane w obszernej rozprawie są następujące.

1. Opracowanie metod klasyfikacji śladów krwi w obrazach wielospektralnych za pomocą architektur głębokich sieci neuronowych (Rozdział 2).

   Jako punkt wyjścia został tu wybrany stosunkowo nowy zbiór danych wielospektralnych zawierających różnorodne ślady krwi pozostawione na różnych materiałach. Zbiór ten został przygotowany i zaprezentowany w 2021 roku przez zespół badawczy (M. Romaszewski, P.

Głomb, A. Sochan, M. Cholewa), z którym związany jest Doktorant. Celem przygotowania tego zbioru danych było umożliwienie wytrenowania skutecznych klasyfikatorów do późniejszego wykorzystania przez służby śledcze, policję, prokuraturę itd. Główna zaleta tego typu zbioru danych zastosowanego w dziedzinie wielospektralnej to możliwość określenia rodzaju śladu oraz czasu od jego powstania. Doktorant, Pan magister Kamil Książek użył tego zbioru do wytrenowania szeregu różnych głębokich sieci neuronowych typu splotowe 1D, 2D oraz 3D, jak również sieci rekurencyjne typu GRU. Wykorzystał do tego celu znaną bibliotekę *Deep HyperX*, jednakże z własnymi modyfikacjami i usprawnieniami. Następnie, na podstawie wielu testów określił struktury optymalne w realizacji tego zadania. Dodatkowo zweryfikował eksperymentalnie problem używania danych testowych pochodzących z tego samego, jak również różnego rozkładu statystycznego na działanie klasyfikatorów neuronowych. Oczywiście, z praktycznego punktu widzenia to drugie podejście jest bardziej interesujące.

2. Przeprowadzenie analizy wpływu sposobu inicjalizacji wag głębokich autoenkoderów neuronowych do identyfikacji składowych głównych w obrazach wielospektralnych oraz opracowanie metod ich reinicjalizacji w celu poprawy działania (Rozdziały 3 oraz 4).

   Doktorant podjął się realizacji trudnego zadania identyfikacji składowych głównych w obrazach wielospektralnych, którego rozwiązanie ma duże znaczenie teoretyczne, gdyż może posłużyć w innych dziedzinach w których bada się problem odtwarzania składowych z mikstury sygnałów, jak również praktyczne, gdyż w tym aspekcie umożliwia identyfikację materiałów składowych obserwowanych w wielu spektrach obiektów rzeczywistych, takich jak np. skład osadów, czy też udział roślinności danego typu itd. Problem ten Doktorant rozwiązywał bazując na rozwiązaniach już publikowanych przez inne zespoły badawcze (np. Pallson B. i in. *Hyperspectral Unmixing Using a Neural Network Autoencoder*, itd.), zauważając problem obserwowalnej w pewnych przypadkach niestabilności głębokich sieci neuronowych użytych do tego zadania – w tym przypadku chodzi tu o tzw. sieci autoenkoderów, czyli struktur neuronowych złożonych z tzw. enkodera i znajdującego się za nim dekodera, a uczonych próbkami tych samych obiektów. Doktorant przeprowadził szereg badań za pomocą testów statystycznych aby wykazać wpływ sposobu inicjalizacji wag sieci, bądź brak tego wpływu, na dokładność odpowiedzi tej sieci. Badania te potwierdziły hipotezę o istotnym wpływie procesu inicjalizacji wag, co nasunęło kolejny problem do rozwiązania, czyli opracowanie sposobu takiej reinicjalizacji używanych sieci, aby zapobiec, bądź też istotnie zmniejszyć wspomnianą niestabilność sieci neuronowych. Niemniej istotne w tym kontekście było też dokładne przeanalizowanie problemu zanikającego gradientu w procesie uczenia sieci neuronowych.

   Pan magister Kamil Książek opracował i zbadał szereg sposobów reinicjalizacji wag sieci neuronowej podczas jej uczenia w celu eliminacji, lub też maksymalnej możliwej neutralizacji niekorzystnych zjawisk niestabilności oraz tzw. umierających neuronów (ang. *dying neuronons*) na skutek zanikającego sygnału aktywacji w procesie ich uczenia. Co więcej, opracowane metody zostały zweryfikowane nie tylko na danych wielospektralnych, ale również na zbiorze referencyjnym MNIST.

Wszystkie wyżej wymienione osiągnięcia świadczą o dojrzałości naukowej Pana magistra Kamila Książka, który swobodnie porusza się różnorodnych zagadnieniach analizy i przetwarzania obrazów wielospektralnych, jak również w aspektach budowy i modyfikacji wielu typów architektur głębokich sieci neuronowych.

3. *Jaka jest umiejętność poprawnego i przekonywującego przedstawienia uzyskanych wyników oraz jaki jest dorobek publikacyjny Doktoranta*

Rozprawa doktorska Pana magistra Kamila Książka jest bardzo obszerna i liczy 177 strony. Napisana jest w języku angielskim. Praca podzielona jest na 6 rozdziałów głównych i dodatki. Bibliografia liczy 177 pozycje literaturowe, w tym publikacje autora. W grupie tych ostatnich wymienić należy dwie, które stanowią bazę całej rozprawy doktorskiej Pana Kamila Książka – są to:

1. Kamil Książek, M. Romaszewski, P. Głomb, B. Grabowski, M. Cholewa, *Blood Stain Classification with Hyperspectral Imaging and Deep Neural Networks*, MDPI Sensors, 2020, Vol. 20, Issue 22, No. 6666, pp. 1-24, ISSN 1424-8220.
2. Kamil Książek , P. Głomb, M. Romaszewski, M. Cholewa, B. Grabowski, K. Buza, *Improving Autoencoder Training Performance for Hyperspectral Unmixing with Network Reinitialisation*, ICIAP 2022, In: S. Sclaroff, C. Distante, M. Leo, G.M. Farinella, F. Tombari, (eds) Image Analysis and Processing – ICIAP 2022, Lecture Notes in Computer Science, 2022, Vol. 13231, Springer, Cham, pp. 391-403, ISBN: 978-3-031-06426-5.

W obydwu tych publikacjach Doktorant, Pan magister Kamil Książek jest głównym autorem. Należy też wspomnieć, że pierwsza z publikacji posiada tzw. współczynnik wpływu IF=3,576 (rok 2020).

Imponująca jest również druga lista obejmująca 14 publikacji, w których Pan Kamil Książek jest współautorem, a aż w 6 przypadkach jest autorem głównym.

Zarówno liczba, która znacznie przekracza minimalną wartość wymaganą w przewodach doktorskich, jak i jakość publikacji świadczą o aktywności i dojrzałości naukowej Doktoranta, jak również o jego rozpoznawalności w środowisku naukowym.

4. *Jakie są słabe strony rozprawy i jej główne wady?*

Rozprawa doktorska Pana Kamila Książka zawiera opis głównych metod opracowanych przez Doktoranta, o których była już mowa w poprzednich punktach. Mimo znacznej liczby opisanych metod i wynikającej z tego objętości pracy, nie znalazłem poważniejszych błędów, czy też uchybień merytorycznych. Należy jednak jasno zaznaczyć, że cała rozprawa w znacznym stopniu bazuje na wcześniejszych publikacjach Doktoranta wraz z innymi współautorami, jak również na innych publikacjach oraz zaprezentowanych tam metodach i oprogramowaniu, o których była mowa w poprzednim punkcie.

Poniżej przedstawiam uwagi i pytania, zaczynając od listy pytań zasadniczych, po których następuje lista pytań bardziej szczegółowych.

1. W przypadku opracowanych metod klasyfikacji śladów krwi w obrazach wielospektralnych za pomocą architektur głębokich sieci neuronowych (Rozdział 2), opis przedstawiony w rozprawie bazuje na artykule naukowym, wymieniony w poprzedni punkcie jako nr 2 (jest to pozycja [80] w spisie bibliograficznym zawartym w rozprawie). Mimo iż na str. XV rozprawy jest pewien opis wkładu Doktoranta, to jednak nie ma opisu wkładu pozostałych czterech autorów. Co więcej, bardzo trudno dostrzec istotne różnice naukowe w obydwu publikacjach, tzn. w rozprawie oraz we wspomnianym artykule naukowym. Wiele to wręcz kopie – dla przykładu, Rysunek 2.3 z rozprawy to kopia Rysunku 2 ze wspomnianej pracy, o czym jednak nie ma

żadnej zmianki w opisie Rysunku 2.3 w rozprawie doktorskiej. Dobrze byłoby gdyby Doktorant dokładnie wyliczył te różnice i je krótko opisał. W innym przypadku nasuwa się pytanie o cel tworzenia tak rozbudowanej i trudnej dla czytelnika rozprawy doktorskiej, skoro cały istotny opis znajduje się w już zrecenzowanej publikacji (tylko 24 strony), którą wystarczyłoby zacytować?

Podobną prośbę można też skierować w celu wyjaśnienia różnic pomiędzy pracą nr [79], a opisem znajdującym się w Rozdziale 3 rozprawy (jest ona wspomniana w Rozdziale 3.1.3 na str. 57).

2. Również w przypadku metod opisanych w Rozdziale 2 Doktorant mocno bazuje na rozwiązaniach zaprezentowanych w pracy Audebert N. i in. *Deep Learning for Classification of Hyperspectral Data – A comparative review*, IEEE 2019 (praca oznaczona [13] w spisie literatury). Zostało również użyte oprogramowanie z tej publikacji, tj. wersja biblioteki *Deep HyperX*. Dobrze byłoby aby Doktorant jasno określił jakie istotne modyfikacje wprowadził do zaprezentowanych w tej publikacji metod, jak również do wspomnianej biblioteki, oprócz użycia innego zbioru danych wielospektralnych.
W Rozdziale 2.4 Doktorant wspomina użycie *cross-entropy* jako funkcji kosztu dla wszystkich przypadków uczonych sieci neuronowych – ale czemu tylko ta funkcja? Być może warto by się zastanowić nad innymi funkcjami kosztu, np. *focal loss*, która lepiej radzi sobie z problemem niezbalansowania zbioru uczącego. Przy okazji ciekawe byłoby też przedyskutowanie wpływu tego zjawiska, tj. *niezbalansowania danych*, na eksperymenty przedstawione w rozprawie.

3. Odnośnie wyników zaprezentowanych w Rozdziale 2.7.3 może warto byłoby wyjść poza przestrzeń publikacji [13] z 2019 roku, czy też [109] z roku 2017(!) i zadać sobie bardziej ambitne zadania, dla przykładu używając transformery wizyjne?

4. Sama analiza dwóch przypadków użycia różnych rodzajów danych do testowania klasyfikatorów, tj. transduktywna oraz indukcyjna, jak to określa Doktorant – *Hyperspectral Transductive Classification* (HTC) oraz *Hyperspectral Inductive Classification* (HIC) – wymaga pogłębionej dyskusji. Pierwszy jest wymieniany wielokrotnie, ale jakie ma on znaczenie i kiedy jest uzasadniony? Jakie są ograniczenia statystyczne w przypadku dobrze generalizujących głębokich sieci neuronowych? Czy skoncentrowanie się wyłącznie na scenariuszu "inductive" nie byłoby lepszym rozwiązaniem? Czy może się tak zdarzyć, że jakiś algorytm działa dobrze i radzi sobie w trybie "inductive", a nie najlepiej z "transductive"? W wielu publikacjach podejście "transductive" uważane byłoby jako niewłaściwe, bądź wymagałoby bardzo dobrego uzasadnienia, np. chronicznego braku danych itd. (np. w często cytowanej przez Doktoranta pracy [13] te pojęcia w ogóle nie padają).
A jak wygląda sprawa tzw. *augmentacji danych* – czy to jest strategia "transductive", czy "inductive"? Dlaczego Doktorant praktycznie w swoich rozwiązaniach pominął możliwość augmentacji danych, czyli generowania obrazów o podobnym rozkładzie do obrazów trenujących, jednakże z pewnym modyfikacjami typu dodatkowe obroty, dodany szum itd.

W tym kontekście również wspomniany został przez Doktoranta tzw. *transfer learning*, ale temat ten nie został rozszerzony. Nasuwają się pytania, na ile może on być w ogóle użyty w przypadku obrazów wielospektralnych? Czy na przykład wagi sieci neuronowej wytrenowanej innym zbiorem, o innym zakresie spektralnym, mogłyby zostać użyte przy trenowaniu innej sieci? Jeśli tak, to być może możliwy byłby transfer z przestrzeni RGB?

5. W analizie czynników wpływających na powstawanie tzw. martwych neuronów (Rozdział 4) brakuje mi pogłębionej analizy potencjalnych czynników wpływających na to zjawisko, takich jak:
   - Wpływ tzw. sygnału biasu (np. o dużej wartości);
   - Wpływ zbyt dużego współczynnika uczenia (tzw. „przerzucenie" wag w kierunku ujemnym co powoduje ujemny argument ReLu i 0 na wyjściu);
   - Analizy innych modyfikacji funkcji aktywacyjnej ReLu, np Leaky ReLU, parametric ReLU (PReLU), exponential (ELU), Gaussian error linear units (GELU) i kilka podobnych, czemu nic o nich nie ma w rozprawie?
   - A być może są to problemy czysto obliczeniowe? Czy Doktorant zna zjawisko "*catastrophic cancellation*" w przypadku obliczania gradientów i odejmowania wartości zmiennoprzecinkowych (ang. *floating-point*) o zbliżonej wartości? Czy zbadał to zjawisko w swoim kodzie?

Kolejne, bardziej szczegółowe pytania i uwagi są następujące:

6. Rozdział 2.2 – część opisanych tu prac jest już „leciwa"; trochę też za mało jest tu dyskusji na temat relacji zaproponowanej metody z tym co jest w literaturze (przeważnie po to się robi analizę zbliżonych prac (ang. *related works*), aby pokazać pewne miejsca, czy też możliwości udoskonalenia/rozszerzenia tak przytoczonych metod innych autorów).
7. W Rozdziale 2.8 Doktorant pisze „*We prepared an optimization of deep neural networks* [...]" – czy Doktorant mógłby przedstawić te oryginalne optymalizacje w bardziej czytelny i syntetyczny sposób?
8. Rysunek 2.12 (str. 47) – bardzo dziwna macierz konfuzji na rysunku (i), czy Doktorant mógłby uzasadnić taki jej kształt?
9. W opisanych w Rozdziale 3.3 przypadkach metod inicjalizacji wag sieci neuronowych, chodzi o metodę wg He (publikacja [60]) oraz Glorot (publikacja [50]), zasadniczo odnoszą się do obrazów „wąskospektralnych" RGB takich jak w ImageNet. Nasuwa się więc pytanie jak mogą one zostać wprost użyte w trenowaniu sieci obrazami wielospektralnymi?
10. W Rozdziale 3.4.2 Doktorant w ogóle nie wyjaśnił specyficznej „krótkiej" architektury dekodera, o czym jest mowa w pracy Palsson B. i in. (nr [114]), a co wynika z przyjętej liczby składowych głównych w procesie mieszania i związaną z tym liczbą neuronów w warstwie ukrytej, jak również modelem liniowym. Czy ta architektura jest więc taka sama jak w pracy [114], czy też się różnią, a jeśli tak to czym?
    Czym dokładnie różnią się wersje oznaczone w rozprawie jako „original" oraz „basic" (str. 66)? Jaka jest ich relacja do implementacji z pracy [114] (jest o tym mowa również w Rozdziale 3.5.4)?
    Dlaczego Doktorant nie użył innego rodzaju autoenkodera, np. wariacyjnego?
11. W tym samym rozdziale (str. 66) – wspomniana jest funkcja aktywacji ReLU, ale wcześniej była wymieniona „*the sigmoid activation function*" (str. 65), to która jest użyta w wersji przedstawionej w rozprawie?
12. Rozdział 3.5.2 (str. 68), Autor wspomina o 2500 modelach, co jest wyzwaniem, tym niemniej znajdowanie optymalnej struktury jest jeszcze ważniejsze niż optymalnego modelu dla danej struktury, więc warto by się zastanowić jaka jest relacja między tymi dwoma „optimami"?
13. Wzór (3.25), nie do końca jest jasne czemu endmembers error też nie jest RMSE? Jakie są przewagi stosowania SAD w tym przypadku (i nie jest to *sum of absolute differences*).

14. W konkluzjach w Rozdziale 3.8 opisany został wpływ wielu czynników, ale jakie są wnioski konstruktywne? Tzn. co używać w realnych zadaniach? Jakie są potencjalne nowe kierunki badań i rozwoju w tym zadaniu?
15. Rozdział 4.2, kilka pytań:
    W definicji 4.2.1 czy zawsze mówimy o $g_{l,i}=0$, czy też o wartościach bliskich 0, czyli $g_{l,i}<t$, gdzie t jest wartością progową (tzw. *thresholdem*) związanym np. z precyzją reprezentacji zmiennoprzecinkowej?
    Dowód str. 92 – wystarczyłoby chyba prześledzić wykres funkcji cosinus, która dla wartości $\geq\pi/2$, które przyjmują wartości ujemne; jednakże Doktorant w ogóle nie uwzględnił istnienia składowej addytywny w postaci tzw. biasu b_l. Poza tym jeszcze jest obliczany gradient (splot, pochodne), to tzw. zanikający (ang. *vanishing*) gradient może być związany z tym że funkcja nie zmienia swojej wartości, ale niekoniecznie jest 0 (czyli jest stała, choćby ze względu na wspomniany niezerowy bias).
    Opisany na str. 94 problem nieaktywnych neuronów i usuwania nieistotnych wag (ang. *pruning*) jak się ma do zjawiska tzw. drop-outu, czyli celowego usuwania niektórych połączeń neuronowych w niektórych sieciach w celu podniesienia właściwości generalizacyjnych tych sieci? Obydwa zjawiska, tj. dying neurons oraz dropout, są dość podobne, choć zapewne w swej naturze i celu/przyczynie różne, więc warto by im się bliżej przyjrzeć.
16. W analizie korelacji (Rozdział 4.6.1) czy zamiast współczynnika korelacji Spearmana nie możemy użyć jakiejś metody wprost bazującej na wartościach gradientów w sieci?
17. Czy użycie tak prostego zbioru danych jak MNIST (Rozdział 4.7), a w dodatku monochromatycznego, jest uzasadnione w badaniu tak złożonego zjawiska jak scenariusz inicjalizacji wag? Jakie wnioski możemy wyciągnąć z tego typu eksperymentów, w których nawet nie zmodyfikowaliśmy w żaden sposób architektury bazowej, liczności warstw itd?
    Ponadto, nie jest jasne jak zostały otrzymane histogramy przedstawione na Rysunku 4.7, jak została obliczona wartość „Density"?

Powyższe uwagi wymagają dalszego wyjaśnienia ze strony Doktoranta, a być może też rozszerzenia pewnych zagadnień natury naukowej poruszonych w rozprawie.

5. *Jaka jest przydatność rozprawy dla nauk technicznych?*

Metody oraz algorytmy optymalizacji architektur głębokich sieci neuronowych do zagadnień przetwarzania obrazów wielospektralnych, opracowane oraz opisane w rozprawie doktorskiej Pana magistra Kamila Książka, mają istotne znaczenie dla nauk technicznych, zarówno w aspekcie osiągnięć teoretycznych, jak i możliwości aplikacyjnych. Wartość ta wynika z głównych i samodzielnych osiągnięć Doktoranta, o których była mowa w poprzednich rozdziałach recenzji. Metody te zostały opublikowane w recenzowanych periodykach naukowych o zasięgu międzynarodowym, co na pewno przyczyni się do ich widoczności w świecie naukowym. Co więcej, wszystkie z przedstawionych metod autorstwa Pana magistra Kamila Książka mogą być dalej rozwijane w celu zwiększenia ich możliwości w dziedzinie analizy wielospektralnej, a być może mogą one też być użyteczne w przypadku budowy głębokich sieci neuronowych do innych typów obrazów, takich jak „klasyczne" spektrum widzialne RGB, ale również na przykład obrazowanie termiczne, czy też obrazy sonarowe.

Wnioski końcowe

Postawione przez Autora Pana magistra Kamila Książka zadania badawcze zostały rozwiązane. Uzyskane rezultaty stanowią oryginalny własny wkład Doktoranta Pana magistra Kamila Książka w rozwój dyscypliny naukowej Informatyka Techniczna i Telekomunikacja. Uzyskane wyniki świadczą, że Pan Kamil Książek wykazał się dogłębną znajomością najnowszych metod i algorytmów przetwarzania danych wielospektralnych za pomocą zoptymalizowanych architektur głębokich sieci neuronowych, jak również dobrym opanowaniem warsztatu badawczego, a w rezultacie również wysoką dojrzałością naukową.

Recenzowaną pracę oceniam jako spełniającą wymagania stawiane rozprawom doktorskim. Wnioskuję o jej przyjęcie oraz o dopuszczenie Pana magistra Kamila Książka do publicznej obrony.

B. Gry