

Prof. dr hab. inż. Krzysztof Krawiec
Instytut Informatyki
Politechnika Poznańska
ul. Piotrowo 2
60-965 Poznań

Poznań, 1.12.2025

Recenzja rozprawy doktorskiej mgr inż. Łukasza Tulczyjewa p.t. “Unmixing of Hyperspectral Images: Where Artificial Intelligence Meets Earth Observation”

1. Tematyka rozprawy

Rozprawa doktorska mgr inż. Łukasza Tulczyjewa dotyczy analizy i modelowania danych pochodzących z teledetekcji (remote sensing), a dokładniej obrazowania hiperspektralnego (hyperspectral imaging), w którym każdy piksel obrazu opisany jest przy pomocy spektrum, obejmującym zazwyczaj relatywnie dużą liczbą pomiarów. W rozprawie zaproponowano i przebadano empirycznie szereg podejść uczenia maszynowego i architektur uczenia głębokiego dedykowanych do analizy takich danych i rozwiązywania przy ich pomocy konkretnych problemów predykcyjnych. To pozwala mi stwierdzić iż przedłożona rozprawa wpisuje się w obszar dziedziny informatyka techniczna i telekomunikacja.

2. Zawartość rozprawy

Rozprawa składa się z sześciu rozdziałów, z których pierwsze trzy motywują opisane w niej badania i wprowadzają wybrane pojęcia podstawowe. Rozdział 4 zawiera obszerny przegląd powiązanej literatury. Oryginalne przyczynki Autora zaprezentowane są w kolejnych sekcjach rozdziału 5.

Sekcja 5.1 porównuje empirycznie cztery metody predykcji komponentów spektralnych, różniące się sposobem agregacji cech obrazowych i spektralnych, na dwóch problemach praktycznych (zbiory danych Jasper Ridge i Urban). Głównym celem porównania było zbadanie skuteczności predykcyjnej otrzymanych modeli w funkcji rozmiaru zbioru uczącego. Otrzymana konkluzja wskazuje na mniejsze od spodziewanego zapotrzebowanie rozważanych architektur na dane uczące.

Sekcja 5.2 koncentruje się na technikach podziału danych na zbiór uczący i testujący, lub innymi słowy próbkowania zbioru uczącego i testującego. Autor słusznie wskazuje na wady

konwencjonalnego losowego próbkowania pikseli z obrazów, które prowadzi do przecieków informacyjnych (information leakage) i w efekcie zawyżonych wartości metryk. Eksperymenty Doktoranta wskazują na ryzyko nawet 5-krotnego zaniżenia błędów predykcji. W sekcji zaproponowano też algorytm podziału GSV adresujący te problemy. Także w tym rozdziale Autor badał wpływ rozmiaru zbioru uczącego na zdolność predykcyjną modeli.

Sekcja 5.3 bada skuteczność architektur angażujących fuzję cech ekstrahowanych z obrazu splotami 1, 2 i 3-wymiarowymi (gdzie trzecim wymiarem jest wymiar spektralny) a także architektury wyposażone w wiele potoków (gałęzi przetwarzania) oraz kilka wariantów grafowych sieci neuronowych (GCNNs). Autor przeprowadził eksperymenty ablacyjne i wykazał w nich że proponowane rozwiązania architektoniczne (w szczególności wykorzystanie wielu gałęzi przetwarzania i fuzja obliczanych przez nie cech) przekłada się na najwyższą jakość modeli w kontekście rozważanych instancji problemów (szczególnie skuteczna okazała się gałąź 3D, integrująca agregację przestrzenną i spektralną obserwowanych sygnałów; s. 110). Modele grafowe osiągały zauważalnie gorsze wyniki (choć nie najgorsze), niemniej były znacznie bardziej kompaktowe w sensie liczby parametrów. Sekcja zawiera także podsumowanie wyników badań dotyczących odporności modeli na obecność szumu.

Sekcja 5.4 prezentuje autorską, nieparametryczną metodę selekcji prążków pasm hiperspektralnych CANNIBAL i poddaje ją eksperymentalnej walidacji. Metoda angażuje elementy algorytmów ewolucyjnych, w szczególności tzw. linkage learning, i wykorzystuje postprocessing poprzez klastrowanie powstałych grup zmiennych. Doktorant wykazał przewagę CANNIBALA nad algorytmami referencyjnymi w kategoriach błędu regresji (estymacji komponentów spektralnych HU) i innych metryk, a także w sensie frakcji eliminowanych cech.

Sekcja 5.5 opisuje proponowane przez Autora komitety modeli predykcyjnych i ich zastosowanie do zadania segmentacji obrazu oraz rozplatania komponentów spektralnych (hyperspectral unmixing, HU). Modele wchodzące w skład komitetów różnicowane były poprzez wstrzykiwanie szumu podczas ich uczenia (co jednocześnie może stanowić pewną formę augmentacji danych i regularyzacji modelu) a ich predykcje agregowane były kilkoma alternatywnymi technikami. Eksperymenty obliczeniowe wykazały przewagę heterogenicznych komitetów nad pojedynczymi modelami, a także wysoką skuteczność adaptacyjnej metody agregacji opartej na klasyfikatorze będącym lasem drzew losowych (random forest). Inną konkluzją prowadzonych badań była zdolność konstruowanych modeli do skutecznej generalizacji nawet gdy uczone były one jedynie na frakcji dostępnych danych uczących, oraz pozytywny wpływ iniekcji szumu na zdolność generalizacyjną modeli.

Sekcja 5.6 prezentuje próby kwantyzacji wybranych modeli proponowanych przez Doktoranta, tj. ich konwersji do reprezentacji całkowitoliczbowych, celem ich posadowienia na substratach obliczeniowych o niskim poborze mocy (w szczególności FPGA). W efekcie przeprowadzonych badań i eksperymentów, udało się uzyskać modele o jakości zbliżonej do oryginalnej (floating-point) przy jednoczesnej znacznej redukcji zajętości pamięci, nawet o rząd wielkości.

Rozdział 6 podsumowuje rozprawę, dyskutuje jej przyczynki i oczekiwany impakt, oraz wyznacza kierunki dalszych prac.

3. Ocena treści rozprawy i wkładu oryginalnego

3.1 Ocena treści rozprawy

Rozprawa mgr Tulczyjewa zawiera szereg wartościowych przyczynków do badań i praktyki obrazowania hiperspektralnego, obejmujących architektury predykcyjne, metody selekcji cech spektralnych, oraz metody adaptacji modeli i potoków przetwarzania do architektur obliczeniowych o ograniczonej mocy obliczeniowej pamięci operacyjnej. Autor wykorzystał adekwatne narzędzia badawcze, angażujące m.in. uczenie wielu modeli celem uzyskania odpornych estymatów ich skuteczności, odpowiednie testy statystyczne i weryfikację hipotez, oraz eksperymenty ablacyjne. Eksperymenty obliczeniowe zostały przeprowadzone rzetelnie i obejmowały porównanie proponowanych architektur z wieloma algorytmami referencyjnymi (np. w sekcji 5.3). Autor zadbał o możliwość replikacji eksperymentów udostępniając kody źródłowe proponowanych algorytmów.

Szczególnie doceniam następujące przyczynki/aspekty pracy:

1. Próbę zaadresowania aktualnych i dotkliwych wyzwań stojących przed teledetekcją, w szczególności ograniczonej dostępności danych nadzorowanych, tj. danych obrazowych którym towarzyszą wiarygodne pomiary w terenie (tzw. ground truth). Doktorant adresował to wyzwanie badając, w niemal każdym eksperymencie, wpływ rozmiaru zbioru uczącego na zdolność predykcyjną modeli. Wyniki większości z tych eksperymentów wskazują na niższe (od powszechnie domniemanego) zapotrzebowanie proponowanych metod na dane etykietowane.
2. Dbłość o zachowanie dobrych praktyk uczenia maszynowego, w szczególności niezależności próbkowania zbiorów uczących i testujących (sekcja 5.2) oraz estymowanie wariancji metryk architektur poprzez uczenie i ocenę wielu modeli.
3. Utrzymanie praktycznej perspektywy badań poprzez uwzględnianie ograniczeń technologicznych związanych z teledetekcją, w szczególności konieczności osadzania modeli na platformach sprzętowych o ograniczonej mocy i ograniczonej dostępności pamięci operacyjnej/dyskowej (sekcja 5.6). W podobny sposób wydają się być umotywowane wątki pracy dotyczące selekcji kanałów hiperspektralnych (sekcja 5.4).

Podczas lektury rozprawy napotkałem jedynie pojedyncze fragmenty które wzbudziły moje wątpliwości merytoryczne. Operacja Flatten i przejście na warstwy gęste w Rys. 5.1-5.4 (a także kolejne) wydają się rozbiegać z opisem tych modeli w tekście: jeśli model prowadzi wnioskowanie end-to-end na pojedynczych pikselach, to użyte tam warstwy gęste są de facto splotami 1x1. Wzór (5.7) powinien po prawej stronie znaku równości zawierać $p^2\lambda$ a nie $p^2 \times \lambda$, bo wynikowy tensor jest jednowymiarowym wektorem (1D). Dalej, autor definiuje (standardowo) RMSE jako pierwiastek kwadratowy z MSE. Jednak wydaje się że wyniki raportowane w Tabelach 5.18 i 5.19, 5.20 nie wydają się być spójne z tą konwencją: albo błędy MSE są zbyt duże względem RMSE, lub vice versa; innymi słowy, proste podniesienie do kwadratu liczb z wierszy 'RMSE' nie daje wartości z

odpowiednich wierszy 'MSE'. Podział metod linkage learning na statystyczne i empiryczne (s. 117) wydaje się nieco nieadekwatny, bo metody statystyczne też są empiryczne, ponieważ bazują na konkretnych obserwacjach. Na Rys. 5.4 model przyjmuje na wejściu tensor $N \times N \times \lambda$, a na wyjściu produkuje $n \times n \times \lambda$ – ale jest to autoenkoder uczony przez autoasocjację, a zatem rozmiary tensora wyjściowego powinny być takie same jak wejściowego.

Na s. 114 pada stwierdzenie "GCNNs lend themselves well to quantization techniques, allowing further compression of model size and acceleration of inference without significant loss of accuracy." Czy Autor chciał w ten sposób zasugerować że architektury typu GCNN są *szczególnie* podatne na kwantyzację, w szczególności bardziej podatne niż np. architektury splotowe? Podobna wątpliwość nasuwa się dalej na tej samej stronie, gdy Autor stwierdza też że "GC-NNs offer significant advantages in terms of simplicity and interpretability." – jakie są przesłanki aby twierdzić że architektury GCNN są bardziej interpretowalne od innych architektur neuronowych?

Na s. 119 pada stwierdzenie o złożoności czasowej metody CANNIBAL, ale prezentowana formuła wydaje się abstrahować od liczby pokoleń algorytmu ewolucyjnego czy rozmiaru populacji, co jest zaskakujące.

Uwagi te mają jednak stanowić informację zwrotną i nie wpływają na moją ogólnie pozytywną ocenę przedłożonej rozprawy. Z innych obserwacji (bardziej w charakterze wskazówki/sugestii niż wątpliwości), wydaje się że metoda podziału pikseli na zbiory uczące i testujące opisywana w Sec. 5.2.1 mogłaby bazować na losowaniu wyłącznie współrzędnej Y (height), bo zmiany współrzędnej X (W) w porównaniu ze współrzędną Y są mało znaczące.

3.2 Ocena redakcji rozprawy

Rozprawa mgr Tulczyjewa ma dobrze przemyślaną strukturę i została napisana w języku angielskim. Tok argumentacji i organizacja tekstu są w większości części pracy logiczne i czytelne. Jest bogato ilustrowana i zawiera wiele dopracowanych diagramów (np. Rys. 5.11).

Jedyną zauważalną słabą stroną redakcji rozprawy jest (miejscami) rozwlekłość i redundancja. Na przykład architektury sieci neuronowych przedstawiane są w postaci diagramów *oraz* opisu tekstowego, który zasadniczo powtarza te informacje 1:1. Równanie (5.10, s. 118) wydaje się nadmiarowe względem (5.9), tj. prezentuje to samo pojęcie (różnice wartości funkcji przystosowania) z podstawieniem zmiennej (a poprzedzający tekst sugeruje liczbą mnogą ("differences") wprowadzenie więcej niż jednego pojęcia). Dalej, nawet jeśli po części wynika to z zamierzonej przez Autora struktury pracy, to metryki byłoby jednak lepiej zaprezentować raz (a nie wielokrotnie, w każdej sekcji rozdziału 5), tym bardziej że większość z nich jest szeroko znana. Wiele tabel które zajmują całe strony (np. 5.8) dałoby się umieścić razem z tekstem. Równania poprzedzone są nadmiernie dużymi odstępami. Taka prezentacja tekstu, w połączeniu z dużym rozmiarem czcionki, zmusza czytelnika do częstego wertowania pracy.

Autor wydaje się mieć tendencję do rozpoczynania opisu metod od wymieniania ich zalet, a dopiero dalej ich detalicznego przedstawiania (np. w 5.4.1 i 5.5), podczas gdy bardziej naturalna (i przystępna dla czytelnika) jest moim zdaniem odwrotna kolejność: rzeczowe zaprezentowanie metody po którym następuje omówienie jej właściwości. Generalnie opis metody CANNIBAL wydaje się nieco chaotyczny i mógł być zaprezentowany lepiej. Np. na s. 121 pojawia się odwołanie do akronimu (metody) BOMBS, która nie została wcześniej wprowadzona; domyślam się że jest to jedna z metod wymienionych w sekcji 5.4.2.

Nie uważam aby we współczesnym kontekście konieczne było wprowadzanie w pracy doktorskiej podstawowych pojęć dot. sieci neuronowych (rozdział 3), choć nie traktuję obecności tego rozdziału jako błędu w sztuce.

Zaznaczam jednak że wyżej wymienione uwagi redakcyjne mają pewien odcień subiektywny, i nie oczekuję od Doktoranta ustosunkowania się do nich podczas obrony.

3. Konkluzja końcowa

Rozprawa doktorska mgr inż. Łukasza Tulczyjewa zawiera szereg wartościowych wyników wpisujących się w badania podstawowe i stosowane w obszarze teledetekcji. Znaczna część tych przyczynków ma unikalny charakter, adresuje aktualne wyzwania badawcze i ma znaczące przełożenie praktyczne. Mgr Tulczyjew jest pierwszym autorem czterech z pięciu publikacji które stanowiły podstawę przedłożonej rozprawy, co po części dowodzi jego kluczowej roli w opisywanych badaniach (i znajduje potwierdzenie w zadeklarowanych w rozprawie udziałach). Publikacje te ukazały się w renomowanych czasopismach o zasięgu globalnym i materiałach konferencyjnych dobrych konferencji. Wobec powyższego stwierdzam, że **rozprawa doktorska mgr inż. Łukasza Tulczyjewa spełnia z nawiązką warunki stawiane przez ustawę o tytule naukowym i stopniach naukowych w odniesieniu do rozpraw doktorskich, a zatem powinna być dopuszczona do publicznej obrony, o co wnoszę do Rady Dyscypliny Informatyka Techniczna i Telekomunikacja Politechniki Śląskiej.**

Ponadto, biorąc pod uwagę emanujący z rozprawy nacisk na poprawność metodyki badań, znaczne przełożenie praktyczne wyników, oraz liczne inne publikacje autora w czasopismach i materiałach konferencyjnych (łącznie 28 publikacji wg Google Scholar cytowanych 377 razy na moment przygotowania recenzji), **wnioskuję o jej wyróżnienie.**