

prof. dr hab. inż. Tomasz Trzcíński  
Instytut Informatyki Politechniki Warszawskiej

**RECENZJA ROZPRAWY DOKTORSKIEJ DLA RADY DYSCYPLINY NAUKOWEJ  
INFORMATYKA TECHNICZNA I TELEKOMUNIKACJA  
POLITECHNIKI ŚLĄSKIEJ**

**Tytuł rozprawy:** Unmixing of Hyperspectral Images: Where Artificial Intelligence Meets Earth Observation

**Autor rozprawy:** Łukasz Tulczyjew

## 1 Analiza strony merytorycznej rozprawy

### 1.1 Obszar problemowy

Recenzowana rozprawa doktorska dotyczy analizy i rozwiązywania problemu rozdzielania (unmixingu) obrazów hiperspektralnych w zastosowaniach obserwacji Ziemi. Obrazy hiperspektralne rejestrują dla każdego piksela dziesiątki lub setki wąskich pasm widmowych, dzięki czemu umożliwiają identyfikację materiałów na podstawie ich „podpisów spektralnych”. W praktyce sygnał widziany w pojedynczym pikselu jest często mieszkanką sygnałów wielu materiałów, dlatego badacze zajmują się tzw. rozdzielaniem pikselowym (ang. *unmixing*), polegającym na estymacji udziałów poszczególnych klas materiałów na podstawie danych spektralnych i przestrzennych.

W ostatnich latach pojawia się coraz więcej metod opartych na głębokich sieciach neuronowych, które pozwalają poprawić dokładność rozdzielania, ale równocześnie wymagają starannego podziału danych na zbiory treningowe i testowe. Ze względu na silne korelacje przestrzenne w obrazach satelitarnych łatwo o „przeciek informacji” – jeżeli piksele z tej samej sceny lub sąsiadujących fragmentów trafią jednocześnie do zbioru uczącego i walidacyjnego, to miara dokładności jest przeszacowana. Autor podjął się systematycznego przeanalizowania tego problemu i zaproponował rygorystyczne strategie podziału danych, które minimalizują przeciek informacji między zbiorami treningowymi i testowymi oraz pozwalają na uczciwą ocenę modeli.

Doktorant bada też wpływ liczebności zestawu uczącego na jakość szacowanych udziałów materiałów. Na podstawie obszernych eksperymentów pokazuje, że małe zbiory treningowe (np. poniżej 6% dostępnych danych) prowadzą do zaniżonej dokładności estymacji, natomiast powiększanie zbioru uczącego powyżej 33% nie przynosi już istotnych korzyści. Równocześnie wskazuje, że modele korzystające z informacji przestrzennej (tzw. *cube-based models*) przewyższają klasyczne podejścia pikselowe.

Na tej podstawie Autor projektuje i ocenia nowe architektury sieci neuronowych przeznaczone do unmixingu. W pracy pojawiają się m.in. zespoły głębokich modeli neuronowych, sieci grafowe łączące cechy spektralne i przestrzenne, modele wykorzystujące moduł uwagi oraz elastyczna procedura selekcji pasm, pozwalająca zredukować wymiar danych bez pogorszenia dokładności. Ostatnia część pracy dotyczy kwantyzacji modeli, co umożliwia implementację zaawansowanych architektur na urządzeniach o ograniczonych zasobach, takich jak pokładowe komputery satelitów. Tak zarysowany obszar badawczy jest aktualny i ma duże znaczenie zarówno z perspektywy dyscypliny informatyka techniczna i telekomunikacja, jak i zastosowań w teledetekcji, ochronie środowiska oraz geologii.

Wyraźnie zdefiniowany problem badawczy pozwolił Doktorantowi sformułować spójną tezę pracy, według której możliwe jest uzyskanie znacznie lepszych wyników rozdzielania obrazów hiperspektralnych dzięki połączeniu starannie opracowanego podziału danych, adekwatnych architektur głębokiego uczenia i metod kompresji modeli. Rozprawa w naturalny sposób wpisuje się w nurt badań nad uczeniem maszynowym w analizie obrazów satelitarnych i stanowi znaczący wkład w tę dziedzinę.

## 1.2 Ocena wyników oraz stopnia ich oryginalności

Najważniejsze wyniki pracy mają wymiar zarówno metodologiczny, jak i empiryczny. Po pierwsze, Autor przedstawia oryginalną i dobrze ugruntowaną procedurę podziału scen hiperspektralnych, eliminującą przeciek informacji między zbiorami treningowym i testowym. Podkreśla, że proste podziały losowe są niewystarczające, a zastosowanie odpowiednich metod próbkowania jest warunkiem poprawnej walidacji. Po drugie, Doktorant w przekonujący sposób analizuje zależność między rozmiarem zbioru uczącego a jakością estymacji udziałów materiałów. Wykazuje, że zwiększanie liczby próbek początkowo znacząco poprawia dokładność, ale powyżej pewnego progu nie przynosi on już istotnych korzyści. Taka analiza jest niezwykle przydatna dla praktyków, którzy do implementacji modeli w rzeczywistych zastosowaniach potrzebują tego typu wskazówek.

Oryginalność rozprawy przejawia się także w konstrukcji nowych architektur sieci neuronowych. Autor projektuje zespoły modeli, które łączą wyniki kilku odmiennych architektur, w tym sieci grafowych i tych opartych na modułach uwagi, dzięki czemu uzyskuje dokładniejsze i stabilniejsze rezultaty. W pracy pojawiają się również techniki selekcji pasm oraz lekkie modele kwantyzowane przeznaczone do implementacji na urządzeniach końcowych (tzw. *edge devices*). Wyniki eksperymentalne potwierdzają, że zaproponowane rozwiązania poszerzają aktualny stan wiedzy, a uzyskane przez Doktoranta rezultaty pozwalają na publikację w najważniejszy dla teledetekcji miejscach. Pewien niedosyt pozostawia brak na liście publikacji doktoranta artykułów opublikowanych na najważniejszych konferencjach informatycznych w dziedzinie uczenia maszynowego, takich jak NeurIPS czy CVPR, choć konkretna i relatywnie niszowa tematyka badań zapewne istotnie ograniczają taką możliwość. Ważnym atutem pracy jest klarowne określenie wkładu autora w powstanie poszczególnych publikacji oraz fakt, że w większości z nich występuje on jako pierwszy autor.

Do niewielkich niedociągnięć należy zaliczyć miejscami powtarzający się opis tych samych zbiorów danych w różnych rozdziałach, co utrudnia czytelnikowi odnalezienie wszystkich informacji w jednym miejscu. Rozumiem intencję pozostawienia rozdziałów jako zamkniętych części, jednak nieliczne odwołania np. do metryk przedstawionych w oddzielnym rozdziale nie utrudniłyby, moim zdaniem, znacząco lektury rozprawy, a pozwoliłyby jednak skoncentrować się na meritum poszczególnych rozdziałów. Poza tym w pracy brakuje rozdziału syntetyzującego wszystkie przedstawione metody i podsumowującego ich zalety względem konkurencyjnych podejść – czytelnik zmuszony jest sam porównywać wyniki zamieszczone w różnych częściach rozprawy.

## 1.3 Zagadnienia dyskusyjne

Poniższe pytania i uwagi mają na celu pogłębienie dyskusji nad zaprezentowanymi w pracy wynikami i ich szerszymi konsekwencjami:

- **Strategie podziału danych.** Autor proponuje rygorystyczny sposób dzielenia scen na zbiory treningowe i testowe. Jakie są praktyczne koszty obliczeniowe czy czasowe takiego podziału przy analizie większych obrazów i czy zalecenia dotyczące rozmiarów tak tworzonych zbiorów pozostają niezależne od typu rozwiązywanego problemu (np. klasyfikacja *vs* segmentacja *vs* rekonstrukcja)?
- **Etykietowanie próbek danych.** Rozprawa wskazuje na niezwykle istotnym w praktycznych zastosowaniach uczenia maszynowego problem, a mianowicie kłopot z uzyskaniem etykiet (czy też adnotacji) próbek danych. W ramach swoich badań Autor wskazuje różne podejścia do tego problemu, w tym metody nienadzorowane, jednak zabrakło mi dyskusji na temat syntetycznego tworzenia zbiorów danych czy poszerzania rzeczywistych zbiorów w sposób syntetyczny poprzez tzw. *augmentation*, choćby z wykorzystaniem najnowszych modeli generatywnych takich jak sieci dyfuzyjne. Czy Autor może skomentować wykorzystanie tych metod do rozwiązania stawianych w tezie problemów?
- **Uogólnienie metod.** Zaproponowane architektury skupiają się na problemach unmixingu i klasyfikacji. Czy Autor rozważał zastosowanie ich do innych problemów obrazowania hiperspektralnego, takich jak detekcja anomalii czy rekonstrukcja obrazu? Jakie przeszkody przewiduje w adaptacji proponowanych modeli do takich zadań?

- **Kwantyzacja modeli.** Praca omawia kompresję modeli wykorzystującą kwantyzację. Brakuje mi jednak bardziej szczegółowych wyników dotyczących czasu działania i zużycia energii na rzeczywistych platformach. Czy Autor może przeanalizować te aspekty kwantyzowanych modeli w kontekście urządzeń o ograniczonych mocach obliczeniowych, takich jak roboty czy urządzenia mobilne? Czy zaproponowane podejście jest w stanie rzeczywiście zredukować wymagania energetyczne modeli?

## 2 Analiza strony formalnej rozprawy

### 2.1 Ocena układu pracy i redakcji manuskryptu

Rozprawa została napisana w języku angielskim, a poprzedzają ją streszczenia w języku polskim i angielskim. Praca składa się z rozbudowanego wprowadzenia, rozdziału opisującego wkład i powiązane publikacje, części teoretycznej omawiającej podstawy głębokich sieci neuronowych oraz rozdziałów poświęconych eksperymentom i wynikom autorskich metod. Całość wieńczy podsumowanie zawierające wnioski oraz propozycje dalszych kierunków badań. Bibliografia jest obszerna i aktualna; obejmuje zarówno klasyczne prace z zakresu teledetekcji, jak i najnowsze artykuły dotyczące uczenia głębokiego.

Praca jest starannie zredagowana i zawiera liczne wykresy oraz wizualizacje pomagające zrozumieć działanie proponowanych metod. Układ rozdziałów jest logiczny, choć warto byłoby dodać oddzielny rozdział syntetyzujący wyniki i porównujący wszystkie architektury. W kilku miejscach powtarza się opis tych samych zestawów danych, co można by przenieść do wspólnej sekcji. Merytoryczna jakość bibliografii i cytowań nie budzi zastrzeżeń.

### 2.2 Uwagi szczególne

Tekst rozprawy jest w przeważającej większości poprawny pod względem językowym i stylistycznym. Natrafiłem jednak na kilka błędów, które warto poprawić w ewentualnej publikacji książkowej:

- W jednym z opisów architektur na stronie 13 użyto sformułowania *responsible of transforming the activation maps*; poprawne jest *responsible for transforming the activation maps*.
- W sekcji dotyczącej detekcji szczegółowych cech pojawia się błąd ortograficzny: *fine-grainted features* powinno brzmieć *fine-grained features*, a następnie *difficult to recognized* należy zastąpić *difficult to recognize* (str. 75).
- W rozdziale o podziale danych występuje sformułowanie *such has DL models validation*; powinno być *such as DL models validation* (str. 81).
- Zdanie *To prevent from the information leak phenomenon* można skrócić do *To prevent the information leak phenomenon* (również str. 81).
- W zdaniu *There are numerous benefits of this algorithms* błędnie użyto liczby mnogiej; poprawnie: *benefits of this algorithm* (również str. 81).
- Sformułowanie *Our approach strives in its simplicity* jest nienaturalne; można je zastąpić np. *Our approach excels due to its simplicity* (str. 82).
- W jednym zdaniu na stronie 81 brakuje podmiotu: zdanie *Hence, may lead to overoptimistic results* powinno brzmieć *Hence, this may lead to over-optimistic results*.

## 3 Konkluzja

Z pełnym przekonaniem stwierdzam, że przedmiotem rozprawy doktorskiej jest oryginalne rozwiązanie poprawnie zdefiniowanego problemu naukowego, a recenzowana dysertacja Pana mgr. inż. Łukasza Tulczyjewa **spełnia** wymagania stawiane rozprawom doktorskim przez Ustawę *Prawo o szkolnictwie wyższym i nauce, Dz. U. 2023, poz. 742 z późn. zm.* – praca prezentuje ogólną wiedzę teoretyczną i praktyczną Doktoranta w dyscyplinie *Informatyka techniczna i telekomunikacja* oraz umiejętność samodzielnego prowadzenia pracy naukowej. Autor podejmuje aktualny i istotny problem rozdzielania

obrazów hiperspektralnych, proponując szereg oryginalnych metod, które poprawiają jakość estymacji, zapewniają poprawną walidację oraz umożliwiają implementację modeli na urządzeniach o ograniczonych zasobach. Wyniki eksperymentalne dokumentują skuteczność zaproponowanych rozwiązań, a liczne publikacje – w tym artykuły w wysoko punktowanych czasopismach i materiałach konferencyjnych – potwierdzają wagę wkładu Doktoranta do stanu wiedzy.

Uwagi zawarte w niniejszej recenzji mają charakter w dużej mierze dyskusyjny i nie wpływają na pozytywną ocenę pracy. Na tej podstawie **wnioskuje o dopuszczenie mgr. inż. Łukasza Tulczyjewa do dalszych kroków procedury uzyskania stopnia doktora nauk technicznych.**

Tomasz Trzcíński