

Dr hab. inż. Marek Stanisław Mróz, prof. UWM  
Katedra Geodezji  
Instytut Geodezji i Budownictwa  
Wydział Geoinżynierii  
Uniwersytet Warmińsko-Mazurski w Olsztynie

Olsztyn, 14 kwietnia 2025 r.



### Recenzja rozprawy doktorskiej Pani mgr inż. Pauliny Zachar

sporządzona na zamówienie Przewodniczącej Rady Naukowej Dyscypliny Inżynieria Lądowa, Geodezja i Transport dr hab. inż. Katarzyny Osińskiej - Skotak, prof. Politechniki Warszawskiej (pismo nr RND.ILGiT.06.02.2025 z dnia 20.02.2025 r.), na podstawie art. 190 ust. 2 ustawy z dnia 20 lipca 2018r. - Prawo o szkolnictwie wyższym i nauce (t.j. Dz.U. z 2018 r., poz. 1668 z późn. zm.).

Temat rozprawy: **Wykorzystanie sztucznej inteligencji do wykrywania obiektów na danych fotogrametrycznych.**

Promotor: **prof. dr hab. inż. Zdzisław Kurczyński**

#### 1. Ocena formalna i metodyczna

Przedstawiona mi do recenzji rozprawa ma formę monografii składającej się ze wstępu, części teoretycznej zawierającej dwa rozdziały z podrozdziałami oraz części eksperymentalnej podzielonej na rozdziałów dziewięć. Rozprawę kończy bibliografia oraz spis tabel i rysunków. Bibliografia zawiera 147 pozycji wykorzystanych w różnym stopniu, ale zawsze związanych z jakimś konkretnym fragmentem rozprawy. Konstrukcja rozprawy jest dosyć oryginalna, ponieważ w piętnastostronicowym rozdziale p.t. „Wstęp”, podzielonym na cztery części znajdujemy w skondensowanej formie informacje nt. problemu badawczego, celu i pytań badawczych, tezę pracy oraz opis metodyki badań. To bardzo ułatwia percepcję całości dzieła. Tytuł rozprawy jest zgodny z jej treścią, choć w moim odczuciu nieco zbyt ogólny.

Rozprawa rozpoczyna się opisem problemu badawczego, jakim jest dążenie do skutecznego wykorzystania metod głębokiego uczenia maszynowego (*ang. deep learning*) w procesie pozyskiwania geoinformacji z fotogrametrycznych zasobów obrazowych i „wiedzy zewnętrznej”. Jasno wyartykułowany cel rozprawy to „opracowanie metodyki wykrywania obiektów na ukośnych zdjęciach lotniczych z wykorzystaniem metod głębokiego uczenia, a dokładniej konwolucyjnych sieci neuronowych w celu zasilania miejskich baz danych”.

Cel rozprawy, od strony stricte naukowej, został ujęty w formie tezy brzmiącej „Zastosowanie metod głębokiego uczenia wykazuje przydatność w wykrywaniu obiektów miejskich, a wykorzystanie potencjału lotniczych zdjęć ukośnych umożliwi określenie położenia obiektów w terenowym układzie współrzędnych i może być źródłem zasilania baz danych o obiektach miejskich”, konsekwentnie udowodnianej w kolejnych krokach badawczych.

W pytaniach badawczych zawiera się wartość metodyczna rozprawy, m.in.:

- Jak rozwiązać problem braku danych uczących w algorytmach sztucznej inteligencji?
- Jak poprawnie ocenić efekty działania tych algorytmów?

- Jak połączyć aspekt przestrzenny, inherentny danym fotogrametrycznym, z radiometrycznie „zakodowanym”, płaskim obrazem poszczególnych obiektów świata rzeczywistego, aby ekstrakcja informacji do baz topograficznych odbywała się szybko, poprawnie i „hurtowo”?

Część teoretyczna rozprawy jest ujęta w bardzo przemyślany ciąg – od ogółu do szczegółu- od opisu metod sztucznej inteligencji, tj. metod głębokiego uczenia, przez transfer wiedzy w modelach, cechy geometryczne i semantyczne danych fotogrametrycznych, aż do dwóch kategorii obiektów „quasi-topograficznych” tj. latarni ulicznych i okien w budynkach będących obiektem zainteresowania w wymiarze bardzo praktycznych zagadnień urbanistycznych. W podrozdziale 1.1.2. autorka zebrała sporo informacji o dostępnych modelach głębokiego uczenia, w tym o sieciach neuronowych YOLO i SAM. To jest walor edukacyjny rozprawy. W tym duchu na str. 25 rozszyfrowałbym literki B, L, H zaszyte w nazwach ViT-B SAM, ViT-L SAM, ViT-H SAM ku pożytkowi mniej zaawansowanych czytelników.

## 2. Ocena merytoryczna i kwalifikacja rozprawy

Na tle oceny metodycznej chciałbym w kolejnej części recenzji rozprawy odnieść się do meritum zagadnień. Przebudowałbym tezę rozprawy, która w nowym brzmieniu „nabrałaby odpowiedniego kolorytu” i podkreśliła należne miejsce problematyce rozprawy i ogólnie pozycji fotogrametrii w obecnej dyscyplinie naukowej ILGiT: „Lotnicze zdjęcia ukośne wytwarzają efekt synergiczny między cyfrowym pomiarem fotogrametrycznym i automatycznym rozpoznaniem obrazowym”. Do tego w gruncie rzeczy sprowadza się cel, metoda i utylitarne wyniki rozprawy. Powyższa moja sugestia znajduje nawet odzwierciedlenie na str.17 rozprawy, gdzie pojawia się to słowo „synergia”. Za jedną z większych wartości rozprawy uważam zdefiniowanie luki badawczej i jej kompleksowe wypełnienie.

Sporo uwagi autorka poświęca zagadnieniom transferu wiedzy między modelami. Uważam to za bardzo ważny fragment rozprawy. Dowodzi to wiedzy autorki wynikającej zarówno z doświadczenia jak i wycucia zagrożenia przy wyborze „drogi na skróty”. I kolejne stwierdzenie wynikające z przeprowadzonych badań - „Zapewne w niektórych aplikacjach rozwiązanie z tworzeniem sztucznych danych może być wystarczające, jednak należy rozważyć czy automatyzacja manualnej pracy związanej z etykietowaniem nie przyniosłaby lepszych efektów w konkretnym zadaniu.”

Za ważne osiągnięcie uważam posługiwanie się chmurami punktów DIM i ALS w generowaniu zbiorów uczących. Cytat z rozprawy - „W tej części wykorzystano opracowaną wcześniej metodykę, w której wykorzystuje się fragmenty wyciętych chmur punktów dla obiektów zainteresowania jako dane wejściowe do tworzenia zbioru treningowego do detekcji na zdjęciach ukośnych. Współrzędne chmury punktów z układu terenowego przeliczone zostały do układu pikselowego zdjęcia i w ten sposób powstał zbiór uczący składający się ze zdjęć i informacji o ramach ograniczających.” Bardzo wysoko oceniam zawartość rozdziału 6.3. - „Kompletność chmur punktów”, natomiast zastanawia mnie, dlaczego przy ocenie wyników detekcji latarni z wykorzystaniem modelu YOLO (str.85), w odróżnieniu od oceny działania modelu SAM (str.115) – nie obliczono indeksu Jaccarda?

Osobiście uważam, że jedną z bardzo ważnych konkluzji rozprawy jest zdanie „Biorąc pod uwagę, że model (SAM – przyp. rec.) osiągnął dokładność 50-60% przy domyślnych ustawieniach bez ponownego trenowania i uaktualniania wag, można to uznać za całkiem dobry wynik” (str.104).

Rozdział „Wnioski” bardzo ładnie podsumowuje to przenikanie się „fotogrametrii pomiarowej” (*metric photogrammetry*) z „fotogrametrią interpretacyjną” (*interpretative photogrammetry*), jak je nazywa prof. Paul Wolf w podręczniku „*Elements of Photogrammetry with Applications in GIS*”. Uczenie, trenowanie, dotrenowanie na danych fotogrametrycznych z inherentnym atrybutem jakim jest georeferencja wchodzi do domeny automatycznego, inteligentnego rozpoznania obrazowego i powraca jako nowy produkt – teraz produkt geodezyjny i topograficzny do wektorowej bazy danych. Wyniki osiągnięte w pracy nie są jeszcze „*production ready*”, ale droga została już, mówiąc językiem geodezyjnym, „wytyczona”.

Przed wyrażeniem jednoznacznie pozytywnej konkluzji co do wartości rozprawy i spełnienia kryteriów dalszego procedowania w przewodzie doktorskim pozwalam sobie na zamieszczenie serii uwag szczegółowych o różnym charakterze, które pojawiły się w trakcie lektury. Mają one na celu wywołanie dyskusji podczas obrony w szerszym gronie, ponieważ taka jest istota dyskursu naukowego, w którym „nie broda czyni filozofem” a oryginalność myślenia. Oto niektóre z nich:

Str.8 i dalej – „metodologia”. Osobiście, co podkreślam w wielu recenzjach i przy wielu okazjach, uważam, że nie powinno się używać tego terminu jako synonimu słowa „metodyka”, tylko w znaczeniu głównym, jako „nauka o metodach badań naukowych”. „Metodologia” zapożyczana z angielskiego na użytek „metodyki lub metody” brzmi niepotrzebnie zbyt patetycznie.

Str.8 – „naturalne sceny naziemne” brzmi nie tylko zagadkowo, ale i nienaturalnie ... choć po angielsku brzmi pewnie nieźle. Może po prostu „zdjęcia naziemne”, „typowe otoczenie ulicy”, „krajobraz miejski” ..

Str. 14 i dalej - Unikałbym terminu „szkolenia, szkoleniowego, szkolącego” na rzecz utrwalonych o wiele dawniej w podręcznikach terminów „uczenie, uczyć,..” , lub ostatecznie „treningowy”. W moim odczuciu choć można je uznać za synonimy, to jednak „szkolenie” ma wydźwięk „instruktażowy”, w odróżnieniu od „uczenia”, które ma silniejsze konotacje z rozumem i inteligencją, tak naturalną, jak i sztuczną...

Str.21 – „unikalny” jako rusycyzm zamieniłbym na „unikatowy” lub „szczególny” / „niepowtarzalny” ...

Str.26 – za wszelką cenę unikałbym słów „mapowanie, mapowania, mapować”. W kontekście pomiarów i geodezji zamieniłbym je na „kartowanie, kartografia”, natomiast w kontekście informatyczno-programistycznym na „odwzorowanie”, np. zestawu cech w zestaw innych cech lub parametrów...

Str.28 – „flip – odbicie lustrzane wokół osi x lub y „ – poprawnie jest „względem osi”, nie „wokół osi”.

Str.28 - „Zasadniczo jest to generowanie prawdziwie wyglądających fałszywych danych.” Ta kwestia powinna być szerzej wyjaśniona.

Str.43 - fragment ostatniego zdania (przed tabelą) .. „czy też w innych wysokopoziomowych zadaniach związanych z szeroko pojętym rozumieniem sceny”. Jak rozumiem chodzi o

„skomplikowane zadania związane z postrzeganiem otoczenia obserwatora”. Tekst w rozprawie brzmi jak wygenerowany przez ChatGPT lub tłumaczony zbyt dosłownie z jęz. angielskiego.

Str.56 – „hybrydowe systemy mapujące” zamieniłbym na „hybrydowe” lub „zintegrowane” systemy fotogrametryczne. Byłby to również krok w uznaniu technologii ALS za technologię fotogrametryczną.

Str. 66 – „Dane zostały przetworzone przy użyciu fotopunktów (GCP - ang. Ground Control Point), w tym dziesięciu fotopunktów (GCP) i czterech fotopunktów kontrolnych (ang. check points)”. Zdanie niejednoznaczne. Prosiłbym o wyjaśnienie podczas obrony.

Str.67 – pkt. 6.1. – zbyt mało szczegółowy opis. Jaka firma wykonała zdjęcia i produkty „orto” oraz chmury punktów? W jakim oprogramowaniu i przy jakich parametrach?

Str.68 – autorka pisze „Aby mieć pewność, że wszystkie punkty chmury należące do obiektu zostaną uwzględnione w wyciętych fragmentach zdefiniowano warunek, że wszystkie punkty leżące w promieniu 2m od punktu wskazanego na gruncie należą do obiektu.” Dlaczego 2m? Co stało za tą decyzją?

Str.71 – „Zdarzało się również tak, że latarnie znajdujące się w środku nie odwzorowały się na chmurze punktów DIM pomimo pełnego pokrycia zdjęciami z wielu kierunków” i cd. na str. 72 „analiza tych latarni, ...(),..., nie pozwoliła na znalezienie odpowiedzi, dlatego niektóre zostały odwzorowane w chmurze punktów, a inne nie.” To zagadnienie chciałbym omówić podczas obrony pracy.

Str.88 – „Do najczęstszych błędów, które występowały, jeśli chodzi o detekcję latarni, było wykrywanie cieni latarni zamiast samej latarni, wykrywanie zarówno cienia, jak i latarni jako dwóch różnych obiektów” ... Czy opisaną sytuację uznać za porażkę, czy może zgodnie ze starym powiedzeniem należy „porażkę przekuć w sukces”? Jak się Pani zapatruje na pomysł dodania cienia rzucanego jako wyróżnika latarni w procesie uczenia modelu?

Str.106 – „... ESRI opracowała model Text SAM integrujący SAM z Grounding Dino ...” Czy model TexSAM i Grounding Dino są opisane w literaturze? Nie zauważyłem odniesienia do bibliografii.

Str.107 – dlaczego Bordeaux?

Str.119 – „Na potrzeby badań zdecydowano się na wykorzystanie zamknięcia, które polega na następujących po sobie dylatacji i erozji.” Czy morfologiczną operację zamknięcia wykonano jako jednokrotne użycie po sobie dylatacji i erozji, czy powtórzono to w kilku iteracjach?

#### **Konkluzje recenzji:**

Uwzględniając wymogi ustawowe oraz dobre praktyki i zalecenia Rady Doskonałości Naukowej, po zapoznaniu się z treścią rozprawy, wyrażam swoją opinię w trzech głównych zakresach oceny dysertacji doktorskiej mgr inż. Pauliny Zachar:

1) ocena wraz z uzasadnieniem, czy rozprawa doktorska prezentuje ogólną wiedzę teoretyczną osoby ubiegającej się o nadanie stopnia doktora w dyscyplinie albo dyscyplinach:

Z rozprawy wynika, że kandydatka do stopnia doktora ma szeroką wiedzę teoretyczną w dyscyplinie inżynieria lądowa, geodezja i transport, a w szczególności w specjalności fotogrametria i głębokie uczenie. Zna techniki fotogrametryczne, rozumie podstawy matematyczne opracowania obserwacji fotogrametrycznych.

2) ocena wraz z uzasadnieniem, czy rozprawa doktorska wykazuje umiejętność samodzielnego prowadzenia pracy naukowej przez osobę ubiegającą się o nadanie stopnia doktora:

Rozprawa doktorska prezentuje wyniki eksperymentów obliczeniowych prowadzonych na specyficznych zdjęciach fotogrametrycznych – zdjęciach ukośnych. Kandydatka do stopnia doktora potrafi zaprezentować koncepcje, metodykę, uzasadnić założenia i poprawnie skomentować wyniki. Ma świadomość wartości rzetelności naukowej eksperymentów, powtarzalności wyników i skrupulatności w obliczeniach.

3) ocena wraz z uzasadnieniem, czy rozprawa doktorska stanowi oryginalne rozwiązanie problemu naukowego, oryginalne rozwiązanie w zakresie zastosowania wyników własnych badań naukowych:

Rozprawa doktorska stanowi oryginalne rozwiązanie problemu naukowego, daje wkład do rozwoju dyscypliny i poszerza jej warsztat badawczy.

#### **Kwalifikacja rozprawy**

Zgodnie z wymogami formalnymi stwierdzam, że rozprawa spełnia warunki stawiane rozprawom doktorskim określone w art. 190 ust. 1 ustawy z dnia 20 lipca 2018 r. - Prawo o szkolnictwie wyższym i nauce (t.j. Dz.U. z 2022 r., poz. 574 z późn. zm.). Na tej podstawie wnoszę o dopuszczenie Kandydatki do dalszych etapów procedury doktoryzowania przez Radę Naukową Dyscypliny Inżynieria Lądowa, Geodezja i Transport Politechniki Warszawskiej.

Olsztyn, 14 kwietnia 2025 r.

Digitally signed  
by Marek Mróz  
Date:  
2025.04.15  
18:57:22  
+02'00'



