

09-09-2025

WPŁYNĘŁO

PRZEWODNICZĄCY RADY DYSCYPLINY
Automatyka, Elektronika, Elektrotechnika
i Technologia Kosmiczne

prof. dr hab. inż. Wojciech Szelaż

dr hab. inż. Jakub Nalepa, prof. PŚ
Katedra Algorytmiki i Oprogramowania
Politechnika Śląska
Jakub.Nalepa@polsl.pl

Gliwice, 3 września 2025 r.

Recenzja rozprawy doktorskiej

Tytuł rozprawy: Point-oriented object localization and tracking in low-altitude aerial imagery

Autor: Mgr inż. Bartosz Ptak

Promotor: Dr hab. inż. Paweł Drapikowski, prof. PP

Promotor pomocniczy: Dr inż. Marek Kraft

1 Tematyka rozprawy

Recenzowana rozprawa doktorska Pana mgr. inż. Bartosza Ptaka dotyczy zagadnień związanych z punktową lokalizacją i śledzeniem bardzo małych obiektów w danych obrazowych pozyskanych z wykorzystaniem dronów. Zagadnienia te są istotne w szeregu praktycznych zastosowań związanych z monitoringiem i analizą tłumu, imprez sportowych i kulturalnych, i wielu innych wydarzeń, których monitoring naziemny byłby uciążliwy lub w praktyce niemożliwy. W pracy Doktorant skupił się na opracowaniu algorytmów i technik punktowej lokalizacji obiektów oraz ich późniejszego śledzenia. W obu przypadkach, mgr inż. Bartosz Ptak opracował szereg usprawnień algorytmów wyjściowych, które zostały wszechstronnie zweryfikowane z wykorzystaniem istniejącego i nowych (opracowanych w ramach rozprawy: UP-COUNT oraz UP-COUNT-TRACK) zbiorów danych rzeczywistych. Doktorant precyzyjnie zidentyfikował zalety i wady istniejących metod lokalizacji i śledzenia małych obiektów, a także jasno przedstawił hipotezy badawcze i motywację, która przyświecała opracowywaniu poszczególnych mechanizmów i technik w obu algorytmach.

Uważam, że tematyka rozprawy jest ciekawa i niezwykle aktualna. Tematyka jest zgodna z obecnym nurtem badań związanych z monitorowaniem obiektów z wykorzystaniem dronów (opracowane metody mogą być w przyszłości aplikowane do zagadnień związanych z obserwacją Ziemi z wykorzystaniem kamer umieszczonych na pokładzie samolotów lub satelitów) i dotyczy istotnych aspektów praktycznego wykorzystywania technik wizji komputerowej oraz uczenia maszynowego w rzeczywistych problemach związanych z analizą danych obrazowych. W przyszłości – jak zauważył Doktorant – opracowane metody mogą zostać wdrożone na pokładzie urządzeń brzegowych wyposażonych w różnego rodzaju jednostki przetwarzania danych. Takie podejście mogłoby istotnie skrócić czas niezbędny do ekstrakcji istotnych informacji z danych obrazowych i sekwencji wideo, które mogą wpłynąć na decyzje podejmowane na Ziemi.

2 Ocena strony merytorycznej rozprawy doktorskiej

Rozprawa doktorska jest napisana w języku angielskim i składa się z sześciu numerowanych rozdziałów oraz bibliografii. **Rozdział pierwszy** stanowi krótkie i przejrzyste wprowadzenie do tematyki rozprawy – Doktorant krótko omawia najistotniejsze wyzwania, które są związane z analizą danych pozyskanych z wykorzystaniem dronów i płynnie przechodzi do omówienia problemów punktowej lokalizacji i śledzenia obiektów. W podrozdziale 1.2 (*Research hypothesis and objectives*), Autor zaprezentował hipotezę badawczą, która pozwoliła na sformułowanie celów badawczych niniejszej

rozprawy. Warto jednak podkreślić, że sformułowanie "...to evolve computer vision and neural network methods for point-oriented object localization..." może być mylące i może błędnie sugerować, że celem rozprawy będzie opracowanie metod optymalizacji ewolucyjnej do automatycznego tworzenia systemów wizji komputerowej wykorzystujących algorytmy uczenia maszynowego – unikałbym tego rodzaju niejasnych sformułowań. Podrozdział 1.2 został zwieńczony krótkim podsumowaniem najistotniejszego wkładu Doktoranta, który powstał w ramach rozprawy, a w kolejnych podrozdziałach krótko omówiono wpływ najistotniejszych osiągnięć Autora na naukę i praktyczne zastosowania opracowanych metod. W kolejnych podrozdziałach mgr inż. Bartosz Ptak omawia granty i projekty, które pozyskał i realizował, a także przedstawia listę najistotniejszych artykułów, których jest współautorem (publikacje podzielono na te bezpośrednio związane z tematyką rozprawy oraz na pozostałe artykuły). Należy podkreślić, że prace Doktoranta zostały opublikowane w bardzo dobrych i dobrych czasopiśmie, takich jak m.in. *Neurocomputing*, *Engineering Applications of Artificial Intelligence*, *Remote Sensing*, czy *SoftwareX* (prace były również prezentowane na międzynarodowych konferencjach, np. IEEE IGARSS). Rozdział 1. został zwieńczony krótkim omówieniem struktury pracy.

W rozdziale 2. Autor zawarł przegląd i omówienie metod znanych z literatury związanej z wizją komputerową, zliczaniem obiektów w danych obrazowych oraz detekcją, lokalizacją i śledzeniem małych obiektów (podrozdziały 2.1–2.4). W kolejnym podrozdziale, 2.5 *Datasets*, Doktorant przedstawia zbiory danych, które zostaną wykorzystane w pracy i porównuje znany z literatury zbiór DroneCrowd z nowymi zbiorami, które zostały opracowane w ramach prowadzonych badań (UP-COUNT i UP-COUNT-TRACK). Tworzenie nowych zbiorów benchmarkowych, dla których zdefiniowane są procedury walidacyjne jest niezwykle ważne, ponieważ pozwala na zapewnienie pełnej powtarzalności eksperymentalnej prowadzonych badań – jest to szczególnie istotne w kontekście tzw. „kryzysu reprodukowalności”, z którym musimy się obecnie mierzyć [1]. W kolejnym podrozdziale omówione zostały metryki, które będą wykorzystane w ramach badań eksperymentalnych.

Rozdziały 3–4 stanowią, moim zdaniem, najważniejszą część rozprawy – w rozdziale 3. Autor przedstawia opracowany system do punktowej lokalizacji obiektów, a w rozdziale 4., do ich śledzenia. Omówienie opracowanych algorytmów rozpoczyna się od wprowadzenia do metody wejściowej, która – w kolejnych etapach badań – jest rozszerzana o nowe elementy i mechanizmy analizy danych opracowane przez Doktoranta. W przypadku każdego z takich elementów, przedstawiono motywację (która, w obecnej formie jest moim zdaniem raczej hipotezą badawczą – zamiast np. „*Generating a synthetic dataset using a game engine and incorporating it into the neural network training process enhances performance metrics in the drone-based people-counting task.*”, Autor powinien raczej napisać „*To generate a synthetic dataset (...) in order to (...)*”), opis metody oraz wyniki badań eksperymentalnych. W przypadku systemu lokalizacji małych obiektów, Autor zauważył, że informacja zawarta w obrazie może zostać wzbogacona o informację dotyczącą ruchu obiektów i wyznaczoną z wykorzystaniem przepływu optycznego (z ang. *optical flow*) (Podrozdział 3.1). W kolejnych podrozdziałach (3.3 – 3.7), Doktorant kolejno analizuje wpływ (i) rozdzielczości danych obrazowych na jakość działania algorytmów lokalizacji obiektów, (ii) wykorzystania syntetycznie wygenerowanych danych treningowych na jakość generalizacji modeli uczenia głębokiego w powyższym zadaniu, (iii) wykorzystania dodatkowych danych i „wstrzyknięcia” ich do modelu uczenia maszynowego, możliwych do pozyskania na pokładzie drona (np. dotyczących wysokości przelotu), (iv) wykorzystania danych obrazowych w pełnej dostępnej rozdzielczości w opracowanym modelu sieci neuronowej, a także (v) użycia nowej funkcji straty, dedykowanej do treningu modeli w zadaniu punktowej lokalizacji obiektów. W podrozdziale 3.8 przedstawiono wyniki końcowej ewaluacji opracowanego systemu, a w kolejny podrozdział jest krótkim podsumowaniem badań związanych z punktową lokalizacją obiektów. Struktura rozdziału 5., w którym mgr inż. Bartosz Ptak omawia opracowane metody śledzenia obiektów jest bardzo podobna – w kolejnych podrozdziałach przedstawiono i omówiono algorytm wyjściowy oraz opracowane techniki kompensacji ruchu kamery (problem ten jest niezwykle istotny w przypadku urządzeń brzegowych, takich jak drony), wykorzystania informacji pozyskanej na pokładzie drona (wysokość lotu) do adaptacyjnej aktualizacji hiperparametrów algorytmu śledzenia obiektów, redukcji błędów (zwłaszcza tzw. fałszywych pozytywnów) oraz zapew-

nienia ciągłości trajektorii ślędzonych obiektów. Analizę pełnego systemu, wraz z inkrementacyjnie włączanymi ulepszeniami, omówiono w podrozdziale 4.6, a następnie Autor przeprowadził analizę statystyczną, która miała zweryfikowanie, czy jakość działania opracowanych metod jest zależna od czasu ślędzenia obiektów i charakterystyki sceny, np. liczby i „gęstości” obiektów oraz ich wielkości. Uważam, że algorytmy i metody opracowane przez Autora zostały rzetelnie zweryfikowane, a sam opis układu się w logiczną i ciekawą całość.

W **rozdziale 5.** przedstawiono przykłady możliwych zastosowań opracowanych metod (w zadaniach związanych ze zliczaniem i ślędeniem obiektów w monitorowaniu tłumu). Warto podkreślić, że pokazanie praktycznego aspektu wykorzystania rozwijanych systemów jest bardzo ważne i jest ich dodatkową walidacją (i eksperymentalnym wykazaniem możliwości wdrażania takich systemów w rzeczywistych i społecznie istotnych zastosowaniach). W ostatnim **rozdziale 6.** Autor przedstawia wnioski, odnosząc się do hipotez badawczych, omawia ograniczenia zaproponowanych algorytmów i możliwe kierunki dalszych badań.

Opracowane algorytmy oraz otrzymane wyniki eksperymentalne przedstawione w rozprawie pozwalają z pełnym przekonaniem stwierdzić, że Doktorant jest świetnie przygotowany do tego, żeby prowadzić dalsze badania związane z opracowywaniem systemów wizji komputerowej wykorzystujących techniki uczenia maszynowego w zagadnieniach związanych (nie tylko) z detekcją, lokalizacją i ślędeniem obiektów w szeregu rzeczywistych zastosowań. Należy podkreślić, że zbiory danych i implementacje zostały upublicznione – jak wspomniano we wcześniejszej części recenzji – jest to bardzo ważne i cenne. Rozprawa doktorska mgr. inż. Bartosza Ptaka prezentuje wiedzę teoretyczną i praktyczną Doktoranta w dyscyplinie *Automatyka, elektronika, elektrotechnika i technologie kosmiczne*, a jej stronę merytoryczną oceniam bardzo wysoko.

2.1 Pytania i uwagi

W trakcie lektury rozprawy nasunęły mi się następujące pytania i uwagi:

1. Niektóre elementy rozprawy zostały opublikowane w recenzowanych artykułach czasopismowych i konferencyjnych – warto byłoby to jasno podkreślić w tekście rozprawy (Doktorant pisze np. „*In the research [112]...*” – warto byłoby uwypuklić, że chodzi o pracę Autora).
2. Analizując wyniki badań eksperymentalnych możemy zauważyć, że metody opracowane przez Doktoranta pozwalają na polepszenie wartości metryk dla zbiorów testowych. W niektórych przypadkach poprawa jest raczej niewielka – czy Autor rozważał przeprowadzenie testów statystycznych, które pomogłyby zweryfikować, czy wprowadzenie zaproponowanych ulepszeń prowadzi do uzyskania statystycznie istotnie lepszych wyników? Jakie testy można byłoby tu przeprowadzić? Czy takie niewielkie różnice mogą jakkolwiek wpłynąć na użyteczność opracowanych systemów w praktycznych zastosowaniach?
3. Pewien niedosyt pozostawia analiza jakościowa otrzymywanych wyników. Przykładowo, w podrozdziale 3.2.3 czytamy: „*When the video source, such as a drone camera, is in motion, the generated motion matrices become noisy and fail to accurately capture object movement.*”, a w podrozdziale 4.1: „*These settings make the algorithm robust enough to reduce both false positive and false negative trajectories for long and dynamic sequences.*” – takie (i inne) obserwacje mogłyby zostać wzmocnione poprzez umieszczenie rzeczywistych przykładów i odpowiednich wyników numerycznych w rozprawie.
4. W opisie badań eksperymentalnych (wszystkich wprowadzanych komponentów) warto byłoby podkreślić, że podział danych jest zawsze spójny pomiędzy eksperymentami – czy jest to zapewnione dla wszystkich eksperymentów?
5. W sekcji 3.4 warto byłoby zaprezentować ilościowe charakterystyki (i więcej przykładów wizualnych) wygenerowanych danych syntetycznych.

6. W pracy zabrakło mi analizy czułościowej hiperparametrów opracowywanych algorytmów – Autor wspomina, że niektóre z nich zostały dobrane eksperymentalnie, ale czasami niejasny jest wpływ (i istotność) tego wyboru na jakość działania opracowanych metod (np. wagi w opracowanej funkcji straty, str. 51).
7. Opis implementacji udostępnionej na platformie GitHub jest niepełny (np. w repozytorium DronePeopleCounting) – warto byłoby nad tym popracować.
8. Czy Doktorant rozważał zorganizowanie konkursu z wykorzystaniem opracowanych zbiorów danych, aby zweryfikować jakość metod detekcji, lokalizacji i śledzenia obiektów opracowywanych przez inne grupy badawcze? Jeśli tak, to jak taki konkurs mógłby wyglądać?
9. Jakie wyzwania, według Doktoranta, byłyby najistotniejsze do rozwiązania w przypadku wdrażania opracowanych metod dla innych modalności danych obrazowych i w przypadku innych platform sprzętowych (np. na pokładzie satelitów obrazujących powierzchnię Ziemi)?

3 Ocena strony formalnej rozprawy doktorskiej

3.1 Ocena układu pracy

Układ rozprawy jest logiczny, a kolejne rozdziały wprowadzają czytelnika w najistotniejsze aspekty prac Doktoranta.

3.2 Ocena zastosowanego piśmiennictwa

Bibliografia zawiera około 140 pozycji („około”, ponieważ zauważyłem, że np. oryginalna praca dotycząca architektury U-Net jest zduplikowana). Dobór prac jest poprawny i wskazuje na to, że Autor świetnie orientuje się w aktualnym nurcie badań związanych z tematyką rozprawy. W trakcie analizy piśmiennictwa zauważyłem pewne drobne uchybienia:

1. W bibliografii możemy znaleźć niekompletne wpisy, np. [45], [82], [124]. Czasami brakuje spójności zapisu nazw konferencji i czasopism (w bibliografii możemy znaleźć wpisy bez i ze skrótami nazwy konferencji i czasopism).
2. Czy na pewno praca [65] jest obecnie „*submitted for review*” (od 2014 roku)?...
3. W niektórych wpisach w bibliografii możemy zauważyć niepoprawne wykorzystanie wielkich i małych liter, np. „u-net” i „U-Net”.

3.3 Uwagi redakcyjne

Rozprawa jest napisana bardzo starannie, a tekst jest poprawny pod względem językowym, stylistycznym i interpunkcyjnym. W zasadzie dopatrzyłem się tylko kilku drobnych uchybień: czasami skróty nie są zdefiniowane przy ich pierwszym użyciu (np. na str. 17), najlepsze wyniki (dla wszystkich metryk) powinny być spójnie pogrubione we wszystkich tabelach, czcionka w niektórych rysunkach jest zbyt mała, by móc odczytać tekst (np. na Rys. 4.2) i tylko niektóre rozwinięcia skrótów w tabeli na str. 1 są napisane kursywą.

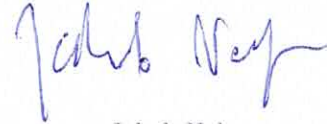
4 Konkluzja

Z pełnym przekonaniem stwierdzam, że przedmiotem rozprawy doktorskiej jest oryginalne rozwiązanie poprawnie zdefiniowanego problemu naukowego, a recenzowana dysertacja Pana mgr. inż. Bartosza Ptaka **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 *Automatyka, elektronika, elektrotechnika i technologie kosmiczne* oraz umiejętność samodzielnego prowadzenia pracy naukowej.

W związku z powyższym, **wnioskuje o przyjęcie rozprawy doktorskiej oraz o dopuszczenie mgr. inż. Bartosza Ptaka do publicznej obrony.**

Ponadto, w związku z bardzo wysokim poziomem merytorycznym badań, których wyniki zostały przedstawione w niniejszej rozprawie oraz z bogatym dorobkiem publikacyjnym Doktoranta, **wnioskuje o wyróżnienie rozprawy doktorskiej** Pana mgr. inż. Bartosza Ptaka.



Jakub Nalepa

Literatura

- [1] S. Kapoor, A. Narayanan, Leakage and the reproducibility crisis in machine-learning-based science, *Patterns* (Aug. 2023). doi:10.1016/j.patter.2023.100804.
URL [https://www.cell.com/patterns/abstract/S2666-3899\(23\)00159-9](https://www.cell.com/patterns/abstract/S2666-3899(23)00159-9)