



PRZEWODNICĄCY RADY DISCYPLINY
Automatyka, Elektronika, Elektrotechnika
i Technologie Kosmiczne
prof. dr hab. inż. Wojciech Szelaąg

Warszawa, 29 września 2025

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

RECENZJA ROZPRAWY DOKTORSKIEJ DLA RADY DISCYPLINY NAUKOWEJ AUTOMATYKA, ELEKTRONIKA, ELEKTROTECHNIKA I TECHNOLOGIE KOSMICZNE POLITECHNIKI POZNAŃSKIEJ

Tytuł rozprawy: A Multi-Module Object Detection System Based on Deep Neural Networks

Autor rozprawy: Aleksandra Kos

1 Analiza strony merytorycznej

1.1 Obszar problemowy

Rozprawa Aleksandry Kos jest poświęcona detekcji bardzo małych obiektów na obrazach o wysokiej rozdzielczości z zastosowaniami w systemach pracujących w czasie rzeczywistym, takich jak roboty autonomiczne czy bezzałogowe statki powietrzne. Autorka trafnie identyfikuje ograniczenia tradycyjnych detektorów: proste podejścia jednoujęciowe z konieczności przeskalowują obraz, co prowadzi do utraty szczegółów i obniżenia dokładności, z kolei metody „sliding-window” są zbyt obciążające obliczeniowo, aby stosować je w warunkach czasu rzeczywistego. W odpowiedzi na te ograniczenia zaproponowany został modułowy system SegTrackDetect, łączący lekką sieć detekcyjną z trzema modułami: (1) estymacją regionów zainteresowania (ROI) z wykorzystaniem segmentacji semantycznej, (2) predykcją ROI na podstawie śledzenia obiektów w sekwencjach wideo oraz (3) globalnym filtrowaniem wyników detekcji przy użyciu autorskich algorytmów Overlapping Box Suppression (OBS) i Overlapping Box Merging (OBM). Tak zbudowany system umożliwia selektywne wykonywanie detekcji w pełnej rozdzielczości tylko w istotnych obszarach obrazu, zmniejszając liczbę wywołań detektora i zużycie zasobów, a jednocześnie zachowując kontekst przestrzenny niezbędny do wykrycia bardzo małych obiektów.

1.2 Ocena wyników oraz stopnia ich oryginalności

Praca ma wyraźnie zarysowaną hipotezę badawczą i trzy tezy pomocnicze, które odnoszą się do poszczególnych komponentów systemu. Autorka pokazuje, że zastosowanie segmentacji do wyznaczania okien detekcji (Teza 1) pozwala znacząco zwiększyć prędkość przetwarzania i poprawia dokładność w porównaniu do naiwnej metody okien przesuwanych. Włączenie dodatkowego źródła ROI w postaci predykcji opartej na śledzeniu (Teza 2) jeszcze bardziej zwiększa recall dla obiektów szybko znikających z detekcji i pozwala zmniejszyć rozmiar wejść segmentera. Wreszcie, autorskie algorytmy OBS i OBM (Teza 3) skutecznie redukują częściowe detekcje i łączą poszczególne wyniki w pełne obiekty, co podnosi precyzję i pozwala na wygenerowanie spójnych wyników.

Ocena eksperymentalna została przeprowadzona na trzech zestawach danych (MTSD, SeaDrones-See, ZebraFish) przy użyciu porównywalnych konfiguracji detektorów. Wyniki wskazują, że system SegTrackDetect osiąga wyższą miarę AP i AR oraz znacząco większą liczbę klatek na sekundę niż metody okna przesuwanego i biblioteka SAHI. Dla przykładów z SeaDronesSee autorka raportuje $AP \approx 52\%$, $AR \approx 61\%$ przy prędkości ponad 44 kl./s, co oznacza ponad 25-krotne przyspieszenie względem SAHI. Takie rezultaty świadczą o wysokiej oryginalności oraz dużym znaczeniu zaproponowanej metody dla praktyki, szczególnie w zastosowaniach robotycznych. Otwarty kod źródłowy systemu czyni eksperymenty powtarzalnymi i umożliwia dalszy rozwój. Na dodatkowe uznanie zasługuje fakt, iż opisane w rozprawie prace były realizowane w ramach tzw. doktoratu wdrożeniowego, co często ogranicza możliwości skupienia się na pracy badawczej, w tym publikowaniu rezultatów prac, natomiast Doktorantka opublikowała szereg artykułów, w tym m.in. w wysokopunktowanych czasopiśmie *SoftwareX* oraz *Engineering Applications of Artificial Intelligence*.

1.3 Zagadnienia dyskusyjne

Choć zaproponowany system prezentuje wyraźne korzyści, pewne aspekty wymagają dyskusji. *Po pierwsze*, omawiane zestawy danych obejmują głównie sceny drogowe i morskie, w szczególności pochodzące z dronów; ciekawi mnie, jak system radziłby sobie w innych domenach, np. z obrazami medycznymi czy zdjęciami satelitarnymi. *Po drugie*, praca bazuje na publikacjach, których współautorką była Doktorantka, jednak brakuje w rozprawie jasnego wskazania, które fragmenty publikacji (a co za tym idzie, które elementy proponowanych metod) były wkładem badawczym Doktorantki. Jest to o tyle istotne, że ocenie powinna podlegać właśnie ta część proponowanych metod, którą zaproponowała Doktorantka. W mojej recenzji przyjąłem, iż opisane w sekcji 1.4 a także w odpowiadających sekcjach kolejnych rozdziałów kontrybucje są w pełni autorstwa Doktorantki, jednak z chęcią uzyskałbym dodatkowe informacje w tym zakresie. *Po trzecie*, jedną z głównych kontrybucji tezy jest mechanizm wykorzystujący wyniki trackingu w ramach pętli zwrotnej, w celu polepszenia detekcji. Mechanizm ten bazuje na filtrowaniu Kalmana, który jest doskonale znany i często używany w zastosowaniach związanych z teorią kontroli. W widzeniu maszynowym, natomiast, w celu trackingu obiektów często stosuje się metody bazujące na przepływach optycznych (*ang.* optical flow). Ciekawi mnie, czy uwzględnienie tej perspektywy w ewentualnym rozszerzeniu proponowanej metody mogłoby przynieść korzyści jakościowe (wydajnościowe zapewne nie, ze względu na wymagania obliczeniowe tych podejść).

2 Analiza strony formalnej

2.1 Ocena układu pracy i redakcji manuskryptu

Rozprawa jest obszerna (ponad 130 stron) i składa się z rozdziału wprowadzającego, przeglądu literatury, opisu proponowanych modułów, eksperymentów, analiz porównawczych oraz aneksów. Układ jest logiczny, każdy rozdział rozpoczyna się omówieniem problemu i kończy syntezą wyników. Autorka konsekwentnie używa notacji matematycznej i tabel do prezentacji wyników. Wyróżnić można liczne ilustracje oraz schematy blokowe, które ułatwiają zrozumienie zaproponowanej architektury. Jednakże brak osobnego podsumowującego rozdziału nieco utrudnia syntetyczne ujęcie wyników i wniosków. Ponadto w niektórych miejscach rozdziały nawiązują do sekcji z innych części (np. „patrz rys. 4.2”), co wymaga ciągłego przewijania tekstu; dobrym rozwiązaniem mogłoby być rozszerzenie prezentowanych w rozprawie grafik (np. Rys. 3.1) o odnośniki do poszczególnych rozdziałów i/lub sekcji.

2.2 Uwagi szczególne

Praca została napisana po angielsku i charakteryzuje się wysoką jakością językową; poniżej zebrano jednak kilka drobnych uwag stylistycznych i redakcyjnych:

- Wstęp zawiera bardzo długie zdania (np. „Ground-truth masks are generated directly from the detection labels, enabling support for both dense and sparse object distributions, in contrast to density- or clustering-based ROI selection methods that favor high-density regions”), które utrudniają odbiór; warto je podzielić na krótsze i uprościć składnię.
- W niektórych miejscach pojawiają się powtórzenia tego samego stwierdzenia w różnych sekcjach, np. opis zalet estymacji ROI i predykcji ROI powtarza się w rozdziałach 1 i 4. Można by je skrócić lub przenieść do jednego, zwartego fragmentu.
- Nazwy modułów i skróty zapisywane są różnie – raz jako „ROI Estimation Module”, innym razem „Estymacja ROI” lub „Estimator”; zalecane jest ujednolicenie nazewnictwa w całej pracy.
- W uzasadnieniu wyboru wersji YOLOv7 można wyjaśnić, co oznacza termin „architectural stability”.
- W streszczeniu po polsku w zdaniu „została publicznie udostępniona umożliwiając zastosowanie w rzeczywistych problemach zarówno naukowych jak i wdrożeniowych” brakuje przecinków oddzielających wtrącenia; powinno być „... została publicznie udostępniona, umożliwiając zastosowanie w rzeczywistych problemach, zarówno naukowych, jak i wdrożeniowych”.

3 Konkluzja

Przedstawiona rozprawa doktorska Aleksandry Kos jest wartościowym i oryginalnym wkładem w rozwój detekcji małych obiektów w obrazach wysokiej rozdzielczości. Autorka zaproponowała modułowy system SegTrackDetect, który w przemyślany sposób łączy segmentację semantyczną, śledzenie obiektów i zaawansowane filtrowanie w celu zwiększenia szybkości i jakości detekcji. Bogata analiza eksperymentalna wykazała przewagę zaproponowanej metody nad licznymi podejściami z literatury, a otwartość kodu ułatwia reprodukowalność i dalsze badania. Nieliczne uchybienia redakcyjne i brak wyczerpującego podsumowania nie umniejszają merytorycznej wartości pracy.

Z pełnym przekonaniem stwierdzam, że przedmiotem rozprawy doktorskiej jest oryginalne rozwiązanie poprawnie zdefiniowanego problemu naukowego, a recenzowana dysertacja Pani mgr. inż. Aleksandry Kos **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.



Tomasz Trzcíński