

dr hab. inż. Piotr Lipiński
Politechnika Łódzka
Instytut Informatyki
Wydział Fizyki Technicznej Informatyki i Matematyki Stosowanej
ul. Wólczańska 215
93-005 Łódź
e-mail: piotr.lipinski@p.lodz.pl

PRZEWODNICZĄCY RADY DISCYPLINY
Automatyka, Elektronika, Elektrotechnika
i Technologie Kosmiczne
prof. dr hab. inż. Wojciech Szeląg

Łódź, dnia 27.01.2025



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

Tytuł rozprawy:

**DATA-EFFICIENT AND EXPLAINABLE MACHINE LEARNING IN VISUAL
PERCEPTION FOR AUTONOMOUS VEHICLES**

Autor rozprawy: mgr inż. Tomasz Nowak.

Promotor: prof. dr hab. inż. Piotr Skrzypczyński

1. Wstęp

Niniejsza recenzja została sporządzona na podstawie pisma Pana Przewodniczącego Rady Dyscypliny Automatyka, Elektronika, Elektrotechnika i Technologie Kosmiczne Politechniki Poznańskiej prof. dr hab. inż. Wojciecha Szeląga z dn. 27.11.2024 r., w którym zawarta jest informacja o powołaniu mnie Uchwałą nr 14/2024-2028 z dnia 19.11.2024, podjętą przez wyżej wymienioną Radę na recenzenta rozprawy doktorskiej mgr inż. Tomasza Nowaka. Do pisma dołączono tekst rozprawy.

Celem recenzji jest weryfikacja spełnienia przez rozprawę doktorską warunków określonych w art. 187 ust. 1 i 2 Ustawy z dnia 20.07.2018 *Prawo o szkolnictwie wyższym i nauce* (Dz. U. 2024 poz. 1571). Recenzję przygotowano zgodnie z Uchwałą Nr 143/2020-2024 Senatu Akademickiego Politechniki Poznańskiej z dnia 5 lipca 2023 r., Uchwałą nr 42/2023-2024 z dnia 18.06.2024 Rady Dyscypliny Automatyka, Elektronika, Elektrotechnika i Technologie Kosmiczne Politechniki Poznańskiej oraz Dobrymi Praktykami w zakresie wyróżniania prac doktorskich obowiązującymi w ramach wyżej wymienionej dyscypliny na Politechnice Poznańskiej.

2. Przedmiot i zakres rozprawy

Przedłożona do recenzji rozprawa doktorska dotyczy zagadnień związanych z opracowaniem efektywnych i wyjaśnialnych modeli uczenia maszynowego dla percepcji wizualnej w mobilnych systemach autonomicznych. W szczególności koncentruje się ona na zwiększeniu dokładności i określeniu niepewności pomiaru lokalizacji na podstawie danych wizyjnych poprzez rozwój ulepszonych architektur sieci neuronowych, zastosowanie nowych metod nauki oraz wykorzystywanie ograniczeń geometrycznych. Zagadnienia te wpisują się w najnowsze kierunki badań w tym obszarze. Autor podejmuje jedno z najistotniejszych problemów, na które napotyka rozwój transportu autonomicznego, takie jak ograniczenie zasobów sprzętowych, wpływ zmiennych warunków zewnętrznych oraz ograniczony dostęp do wysokojakościowych danych uczących. Badania te mogą

znaleźć zastosowanie nie tylko w przedstawionych w rozprawie przykładach, ale także przez ich uniwersalność mogą mieć znacznie szerszy zakres zastosowań.

3. Układ, struktura rozprawy

Rozprawa doktorska mgr inż. Tomasza Nowaka pt. „Data-efficient and explainable machine learning in visual perception for autonomous vehicles” została wydana w formie maszynopisu w Instytucie Robotyki i Inteligencji Maszynowej na Wydziale Automatyki, Robotyki i Elektrotechniki Politechniki Poznańskiej w 2024 r. Praca zawiera 112 stron, w tym 102 strony tekstu, streszczenie w języku polskim i angielskim, spis treści, spis skrótów i oznaczeń oraz bibliografię. Tekst rozprawy został podzielony na 7 rozdziałów. W rozdziale 1 autor zawarł krótkie wprowadzenie, nakreślił kontekst prowadzonych badań, sformułował problem badawczy jak również zawarł listę publikacji, w których przedstawione zostały uzyskane przez niego wyniki. W rozdziale 2 autor skoncentrował się na metodach wykorzystania map uwagi w procesie doboru danych uczących. W rozdziale 3 przedstawił trzy nowe architektury sieci neuronowych, zaprojektowane w celu estymacji położenia punktów charakterystycznych na obrazach stacji ładowania autobusów elektrycznych i wprowadził nową funkcję straty, która wspomaga proces trenowania sieci. W rozdziale 4 przedstawił architekturę sieci neuronowej do estymacji współrzędnych punktów charakterystycznych w przestrzeni trójwymiarowej na podstawie pojedynczego obrazu, trenowaną przy użyciu funkcji straty bazującej na reprojektacji, a także zaproponował nową metodę automatycznego pozyskiwania rzeczywistych lokalizacji kluczowych punktów 3D z zestawu danych zawierającego obrazy, oznaczenia kluczowych punktów 2D oraz modele CAD pojazdów. Dodatkowo zaprezentował dwie architektury sieci neuronowych do kodowania i estymacji kształtu samochodów. Rozdział 5 opisuje potok przetwarzania do estymacji lokalizacji samochodu, rozpoczynający się od detekcji kluczowych punktów 2D na obrazach za pomocą sieci neuronowej zaprezentowanej w Rozdziale 3 oraz estymacji odpowiadających im współrzędnych 3D za pomocą algorytmu przedstawionego w Rozdziale 4. Estymacja położenia realizowana została przy użyciu niestandardowego algorytmu PnP, który wykorzystuje korelacje między punktami 2D i 3D. Dodatkowo potok ten obejmuje mechanizm estymacji niepewności, który propaguje niepewność z wykrytych kluczowych punktów 2D i oszacowanych punktów 3D, dostarczając miarę niepewności związaną z wynikami estymacji pozycji. Rozdział 6 prezentuje praktyczne zastosowania technik opisanych w Rozdziałach 2–5, ilustrując ich skuteczność za pomocą studium przypadków i wdrożeń w warunkach zbliżonych do rzeczywistych. W rozdziale uwzględniono także badanie, będące systematyczną oceną wpływu różnych komponentów zaproponowanych metod. Autor określił zarówno indywidualne, jak i skumulowane efekty tych komponentów na wydajność całego systemu. Ponadto rozdział zawiera kompleksowe porównanie wyników z istniejącymi technologiami, wykazując ulepszenia wprowadzone przez proponowane metody w dziedzinie systemów percepcji dla pojazdów autonomicznych. Ostatni rozdział zawiera podsumowanie pracy, wnioski oraz potencjalne kierunki dalszych badań.

Podział rozprawy na rozdziały jest poprawny pod względem merytorycznym, natomiast w ocenie recenzenta jest nietypowy. Podczas czytania rozprawy po raz pierwszy czytelnik ma wrażenie, że każdy z rozdziałów 2 - 5 dotyczy niezwiązanych ze sobą zagadnień, ponieważ składa się ze wstępu, przeglądu literatury i zaproponowanych rozwiązań, stanowiąc niejako autonomiczną część rozprawy. Dopiero lektura rozdziału 6 pozwala na znalezienie wspólnego mianownika dla wszystkich wcześniejszych analiz i zaproponowanych rozwiązań. Choć jest to podejście nietypowe nie wpływa na negatywnie na odbiór pracy. Warto w tym miejscu wspomnieć, że zapewne jest ono spowodowane bardzo szerokim zakresem przeprowadzonych prac oraz liczbą opracowanych, nowych algorytmów, która znacznie wykracza poza typowy zakres rozprawy doktorskiej.

4. Analiza źródeł i stanu wiedzy

Analiza źródeł i stanu wiedzy dotyczącego zakresu rozprawy została przeprowadzona w rozdziałach 2 – 6, oddzielnie dla rozpatrywanego w każdym z rozdziałów zagadnienia. W rozdziale 2 autor przedstawił analizę literatury dla zagadnień związanych z detekcją obiektów, w rozdziale 3 dla zagadnień związanych z detekcją punktów charakterystycznych, w rozdziale 4 do określania kształtów obiektów, natomiast w rozdziale 5 do określenia położenia obiektów. Rozdział 6 również zawiera przegląd literatury dotyczący z kolei systemów automatycznego dokowania. W zasadzie w każdym rozpatrywanym obszarze tematycznym autor odnosi się do nowych lub najnowszych algorytmów dostępnych w literaturze. W zakresie zagadnień związanych z detekcją obiektów przytacza między innymi metody publikowane na wiodących konferencjach dotyczących tej tematyki, takich jak np. CVPR, jak również najnowsze algorytmy publikowane w roku złożenia pracy doktorskiej. W zakresie metod poszukiwania punktów kluczowych odwołuje się do najnowszych publikacji wykorzystujących sieci typu transformer. W przypadku algorytmów estymacji kształtu, dobór źródeł jest właściwy, natomiast nie zawiera części najnowszych osiągnięć z tego zakresu.

Praktyczny aspekt przeprowadzonych prac został przez autora zamieszczony w rozdziale 6, który również zawiera przegląd dostępnych rozwiązań w zakresie dokowania (w rozpatrywanym przypadku autobusu elektrycznego w stacji ładowania). Analiza literatury w tym zakresie jest wystarczająca, choć mogłaby być uzupełniona o najnowsze rozwiązania.

Podsumowując, należy uznać, że autor przeprowadził analizę stanu wiedzy w sposób prawidłowy, co potwierdza dostateczną wiedzę autora i znajomość literatury w dyscyplinie naukowej, której dotyczy praca.

5. Wartość naukowa rozprawy

Zakres tematyczny pracy wpisuje się w bardzo szeroką tematyka jaką jest mobilność autonomiczna pojazdów. Zagadnienie to badane jest już od kilku dekad i jest dobrze ugruntowane obszary badań naukowych. Rozwój algorytmów sztucznej inteligencji oraz zwiększenie mocy obliczeniowej komputerów sprawił, że zakres zastosowań mobilności autonomicznej znacznie się rozszerzył, co sprawia, że tematyka ta nadal stanowi obszar bardzo intensywnych badań naukowych. Dzieje się tak również ze względu na liczne zastosowania praktyczne, w tym w szczególności te związane z budową autonomicznych samochodów, które bez wątpienia stanowią element tak zwanej czwartej rewolucji przemysłowej. Fakt, że prowadzone przez autora badania wpisują się w ten obszar sprawia, że **tematyka rozprawy jest bardzo istotna z praktycznego punktu widzenia.**

Aby poruszanie się pojazdów autonomicznych było możliwe i bezpieczne, konieczna jest ich lokalizacja w przestrzeni oraz sprawienie, że „rozumieją” one otoczenie zewnętrzne. Zapewnienie tych, wydawałoby się, dwóch prostych funkcjonalności okazało się zadaniem bardziej złożonym niż pierwotnie oczekiwano, czego najlepszym dowodem jest fakt, iż pomimo szumnych zapowiedzi wielu producentów samochodów, te w pełni autonomiczne nadal nie powstały, a częściowa autonomia przemieszczania się tych już istniejących nadal jest mocno ograniczona.

Badania przedstawione w pracy przedłożonej do recenzji dotyczą zarówno zagadnień związanych z lokalizacją pojazdu w przestrzeni, jak i stanowią pewien wkład w próbę zrozumienia otoczenia pojazdu. Autor w swojej pracy zaproponował cztery następujące tezy:

A1. Opracowanie architektur głębokich sieci neuronowych, które umożliwiają ekstrakcję i wizualizację istotnych cech pośrednich, pozwalających na kierowanie procesem uczenia poprzez wzbogacenie istniejących zbiorów danych.

A2 Opracowanie architektur głębokich sieci neuronowych, które umożliwiają opisanie niepewności cech geometrycznych generowanych przez sieć.

A3 Opracowanie metod, które dzięki wykorzystaniu dostępnych modeli 3D obserwowanych obiektów umożliwiają uczenie cech geometrycznych na podstawie obrazów 2D tych obiektów bez potrzeby dokładnego ich oznaczania, jednocześnie poprawiając detekcję cech geometrycznych na podstawie obrazów 2D tych obiektów.

A4 Opracowanie metod, które dzięki znajomości ograniczeń geometrycznych wynikających ze znanych modeli obiektów pozwalają architekturze głębokiego uczenia na zmniejszenie liczby fałszywie wykrytych cech i zwiększenie dokładności lokalizacji wykrytych cech.

Poniżej przedstawiono analizę każdego z nich.

Ad 1. Niezmiennie istotnym zagadnieniem w procesie uczenia sztucznych sieci neuronowych jest odpowiedni dobór zbioru uczącego. Wpływa on nie tylko na szybkość procesu uczenia, ale również na jakość uzyskanych wyników. Problem ten staje się jeszcze bardziej istotny, gdy dostęp do danych uczących jest ograniczony. Jednym ze stosowanych rozwiązań w takim przypadku jest augmentacja danych. W zaproponowanej przez autora metodzie rozszerza on zbiór uczący tylko o takie dane, które są istotne z punktu widzenia map uwagi zaproponowanych w algorytmie Faster R-CNN. Takie podejście bez wątpienia pozwala na przyspieszenie procesu nauki i uzyskanie lepszych wyników, co autor demonstruje zarówno na rzeczywistych danych pomiarowych pochodzących z systemu lokalizacji stacji ładowania autobusów (PUT dataset) jak i na danych ogólnodostępnych (KITTI). Uzyskane przez autora wyniki potwierdzają słuszność zastosowanych nowych rozwiązań w rozważanym zastosowaniu.

Ad 2. W przypadku drugiego osiągnięcia autor porusza niezmiennie istotne zagadnienie jakim jest niepewność lokalizacji cech geometrycznych generowanych przez sieć neuronową. Jak wiadomo, każdy proces pomiarowy związany jest z niepewnością pomiarową. O ile w przypadku klasycznych algorytmów lokalizacyjnych kwestia określenia niepewności pomiaru jest zagadnieniem dobrze znanym i udokumentowanym, o tyle w przypadku zastosowania do tego celu sztucznych sieci neuronowych staje się ono znacznie bardziej złożone. Aby określić położenie cech geometrycznych w rozpatrywanym przypadku autor zdecydował się rozwinąć koncepcje wykorzystane w szacowaniu pozycji człowieka. Wykorzystał do tego celu biblioteki MMPose z repozytorium OpenMMLab oraz HRNet dostępne na Github. Rozwinął dostępne w tych bibliotekach narzędzia o dodatkowe funkcjonalności pozwalające na estymację niepewności lokalizacji cech geometrycznych, które wykorzystują ograniczenia nakładane przez kształt lokalizowanej stacji ładowania. Zaproponowane podejście może nie jest wynikiem przełomowym, ale pokazuje, że autor swobodnie posługuje się dostępnymi narzędziami i potrafi świadomie dostosowywać je do swoich potrzeb. Biorąc pod uwagę złożoność rozpatrywanych zagadnień zdecydowanie należy pozytywnie ocenić zarówno takie podejście i uzyskane wyniki.

Ad 3. Zadanie określenia oryginalnego kształtu przedmiotu na podstawie jego dwuwymiarowego obrazu uzyskanego z kamery jest niewątpliwie jednym z istotnych zagadnień we współczesnej robotyce i ze względu na swoje szerokie zastosowania praktyczne skupia na sobie zainteresowanie środowiska naukowego. Do rozwiązania tego zadania autor wykorzystał znane architektury sieci neuronowych High-Resolution Network i Vision Transformer w celu znalezienia punktów charakterystycznych. Następnie wprowadził oryginalne modyfikacje polegające na wykorzystaniu modeli CAD pojazdów do tworzenia zbioru testowego. Ze względu na fakt, iż w takim podejściu, aby prawidłowo zrzutować punkty charakterystyczne z przestrzeni 3D na lokalizację punktów charakterystycznych w przestrzeni 2D konieczne było uwzględnienie perspektywy oraz kierunku obserwacji kamery doktorant zaproponował wykorzystanie modelu w postaci siatki trójkątów. Dodatkowo, ze względu na wykorzystanie w procesie nauki modeli pojazdów pochodzących z różnych baz danych (ApolloCar3D

i VUMO) zaproponował metodę unifikacji siatki dla obu tych zbiorów danych oraz modyfikację funkcji straty wykorzystywanej w procesie uczenia sieci.

Ad 4. Bardzo zbliżonym koncepcyjnie problemem badawczym do przedstawionego w punkcie Ad 3 jest uwzględnienie w procesie nauki sieci neuronowej, służącej do wykrywania obiektów, ograniczeń geometrycznych wynikających ze znanych a priori modeli tych obiektów. W tezie opisanej w rozdziale 1.2 autor stwierdza, że fakt uwzględnienia tych informacji w procesie budowy architektury i nauki sieci neuronowej pozwala na obniżenie liczby nieprawidłowych detekcji oraz zwiększenie dokładności ich lokalizacji. Zaproponowane przez autora metody oraz ich weryfikacja praktyczna przedstawiona w rozdziale 7 potwierdzają prawdziwość tak sformułowanej tezy.

W ocenie recenzenta zaproponowane przez autora tezy rozprawy są zbyt ogólne i nie do końca oddają rzeczywisty wkład autora w rozwój dyscypliny. W zasadzie należałoby uznać, że główną tezę rozprawy jest teza A4, a pozostałe tezy zarówno pod względem formy jak i treści stanowią jedynie punkty mające na celu realizację tezy A4. Warto w tym miejscu podkreślić, że oryginalny wkład autora wybiega bardzo daleko poza jedynie zaproponowanie architektury sieci do augmentacji danych oraz do określania niepewności położenia punktów charakterystycznych wykrytych obiektów opisane w tezach rozprawy. Autor bowiem:

1. Opracował metodę poprawy wydajności modeli uczenia maszynowego dla niezwykle często występującej sytuacji w której rozmiar zbioru treningowego jest niewystarczający. W zaproponowanym podejściu wykorzystał ukierunkowaną informację zwrotną z metody wizualizacji w celu zidentyfikowania błędów w procesie uczenia w procesie augmentacji danych, rozszerzając zbiór uczący jedynie o próbki niosące największą wartość informacyjną.
2. Dostosował znane architektury sieci neuronowych do wykrywania punktów charakterystycznych do wykrywania ładowarki autobusu elektrycznego poprzez uwzględnienie w procesie uczenia kształtu ładowarki. Modyfikacja ta została zrealizowana przez rozbudowę funkcji straty o dodatkowe parametry wynikające z geometrii ładowarki.
3. Rozbudował architekturę sieci neuronowej o dodatkową funkcjonalność związaną z określeniem niepewności lokalizacji punktów charakterystycznych. Modyfikacja ta pozwoliła na bezpośrednie wyznaczenie lokalizacji punktu charakterystycznego wraz wartością określającą niepewność lokalizacji takiego punktu.
4. Zaproponował uniwersalną, automatyczną metodę oznaczania punktów charakterystycznych pojazdów w przestrzeni trójwymiarowej na podstawie modeli samochodów, która może być stosowana dla różnych trójwymiarowych modeli pojazdów. Aby uniezależnić metodę od sposobu opisu modelu zbudował dedykowaną sztuczną sieć neuronową do tworzenia uniwersalnej siatki opisującej powierzchnię samochodu.
5. Opracował metodę określania pozycji pojazdów na pojedynczym obrazie dwuwymiarowym wykorzystującą autorski algorytm usuwania anomalii.
6. Wykorzystał filtr bezwonny do propagacji błędu wyznaczania poszczególnych punktów charakterystycznych w celu określenia położenia pojazdu.

Warto w tym miejscu podkreślić, że zaproponowane przez autora rozwiązania nie są przypadkowe. W wielu publikacjach pod hasłem „opracowania architektury sieci neuronowej”, użytym w tezach pracy, kryje się podejście siłowe, polegające na dostosowywaniu architektury i parametrów sieci neuronowej tak długo, aż uzyskane wyniki będą lepsze od istniejących. W przypadku autora rozprawy taka sytuacja nie miała miejsca. Dla każdego z nowych, wymienionych powyżej rozwiązań autor dokonywał przeglądu literatury, w sposób świadomy wybierał właściwe metody i modele sieci neuronowych dostosowane do danego zadania, a następnie wprowadzał ich modyfikacje, które, co należy podkreślić, miały uzasadnienie logiczne i teoretyczne. Następnie rozwiązania te implementował i testował

warunkach zbliżonych do rzeczywistych, co w ocenie recenzenta znacznie zwiększa wartość przeprowadzonych badań. Warto wspomnieć, że do tego aby w sposób świadomy modyfikować istniejące metody i ich implementacje należy dogłębnie zrozumieć zasadę ich działania oraz biegle programować.

W ocenie recenzenta praca przedstawiona do oceny w pełni wyczerpuje omawiane zagadnienie. Autor rozwiązał postawione przed nim problemy naukowe, które pojawiły się podczas realizacji praktycznych projektów. Uzyskane podczas prowadzonych badań wyniki były publikowane na szeregu konferencji oraz w czasopiśmie punktowanym, co świadczy o innowacyjności zastosowanych rozwiązań. Praca zawiera więc wszystkie elementy, od praktycznego projektu i wynikającego z niego problemu badawczego, poprzez odpowiedni dobór rozwiązań, kontekst matematyczny, po oryginalne wyniki, pokazujące wyższość zastosowanych rozwiązań nad wynikami uzyskiwanymi przy użyciu innych znanych metod. Warto wspomnieć, że autor nie zaproponował jednego, ale aż sześć nowych rozwiązań, co w znacznym stopniu wykracza poza zakres typowej rozprawy doktorskiej.

Podsumowując analizę osiągnięć doktoranta można stwierdzić, że wszystkie koncepcje przedstawione w rozdziale 1.2 jako autorskie, dalej rozwinięte w rozdziałach 3-5 i poparte wynikami praktycznymi opisanymi w rozdziale 7 są nowe lub zawierają elementy nowości. **Należy zatem uznać, że uwzględniając stan wiedzy i techniki reprezentowanych przez literaturę światową, tematyka rozprawy jest aktualna, rozprawa jest oryginalna i stanowi samodzielny i oryginalny dorobek autora, co potwierdza umiejętność samodzielnego prowadzenia pracy naukowej przez jej autora.**

W kontekście wyżej przedstawionych faktów należy również stwierdzić, że **tezy rozprawy zostały sformułowane i zweryfikowane w sposób prawidłowy**, natomiast w ocenie recenzenta ich doprecyzowanie lepiej uwypukliłoby rzeczywisty wkład autora w rozwój dyscypliny.

6. Uwagi krytyczne i pytania

Ogólnie praca prezentuje bardzo wysoki poziom zarówno pod względem naukowym, praktycznym i edytorskim. Największym zarzutem, który można postawić autorowi jest zbyt mało klarowne opisanie swojego wkładu w rozwój dyscypliny, co stoi w sprzeczności z obecnymi trendami, w których nawet niewielka poprawa prostego algorytmu jest opisywana jak epokowe osiągnięcie. Do zrozumienia istoty wkładu naukowego autora konieczna jest bardzo wnikliwa analiza przeprowadzonych prac oraz bardzo dobra znajomość tematu i, co należy podkreślić, wkład autora jest znacznie większy niż na to wskazuje pobieżna lektura rozprawy.

W ocenie recenzenta na trudny w percepcji układ pracy wpłynął z dużym prawdopodobieństwem sposób prowadzenia badań, który wynikał bezpośrednio z konkretnego praktycznego projektu implementacyjnego. Autor prawdopodobnie próbował rozwiązać postawione przed nim zadanie przy użyciu dostępnych metod, po czym udoskonalał je aby uzyskać pożądane wyniki. Następnie postawione pierwotnie przed nim zadanie zostało odwrócone i stanął on przed koniecznością opisania oryginalnego wkładu w dyscyplinę, w której ubiega się o stopień doktora, a sam proces implementacyjny stał się jedynie przykładem zastosowania zaproponowanej, nowej metody. To typowy problem badań mających kontekst praktyczny, który z jednej strony pozytywnie wpływa na ocenę rozprawy, ponieważ doktorant mierzy się z rzeczywistymi problemami natury technicznej, a nie tylko rozważaniami teoretycznymi, z drugiej, może negatywnie wpływać ona na czytelność rozprawy.

Recenzent ma pewne wątpliwości dotyczące sposobu sformułowania tez rozprawy, co zostało już podniesione w poprzednim rozdziale. Wydaje się, że bardziej czytelnym byłoby uwypuklenie konkretnych modyfikacji metod i algorytmów zaproponowanych przez doktoranta w miejsce sformułowań ogólnych. Na przykład, teza pierwsza w brzmieniu: *“Deep learning architectures that*

allow us to extract and visualise meaningful intermediate features make it possible to guide learning by augmenting the existing data sets” w ocenie recenzenta znacznie lepiej oddawałaby wkład autora w brzmieniu na przykład *“The refinement of deep learning architectures to enable the extraction and visualization of meaningful intermediate features improves training by selectively augmenting training-relevant data only, thereby achieving significant enhancements in localization accuracy”*. Analogicznie można byłoby doprecyzować pozostałe tezy. Należy podkreślić, że wyżej opisane wątpliwości dotyczą raczej stylu formułowania tez, a nie ich zasadniczej istoty.

Uwagi szczegółowe i pytania:

1. Str. 46, równanie (4.2) – autor wprowadza nową funkcję straty. Z punktu widzenia klarowności zaproponowanej miary cenna byłaby analiza w zakresie charakterystyki każdego z elementów składowych w pełnym zakresie zmienności miary.
2. Str. 47, równanie (4.3) – autor w równaniu oznaczył normę euklidesową jako $\|\cdot\|_2$, podczas gdy często stosowanym jej oznaczeniem jest $|\cdot|$. Nasuwa się zatem pytanie, czy autor stosował inne normy w rozpatrywanym zastosowaniu.
3. Str. 59, oznaczenia indeksów górnych użytych w równaniu (5.5) są niespójne z umieszczonym poniżej opisem oraz notacją umieszczoną na str. ix.
4. W części praktycznej (rozdział 6) autor przedstawia różne metody lokalizacji stacji ładowania w przestrzeni trójwymiarowej na podstawie punktów charakterystycznych wyznaczanych przy użyciu algorytmów sztucznej inteligencji. Czy łatwiejszym rozwiązaniem nie byłoby zastosowanie znaczników lokalizacyjnych, np. ArUCO, dla których dostępny jest bardzo szeroki zakres dobrze przetestowanego oprogramowania, zamiast wykorzystywać do tego punkty charakterystyczne?
5. Czy autor prowadził badania wpływu warunków atmosferycznych na uzyskiwane wyniki? Jak na działanie algorytmu lokalizacji stacji ładowania wpłynął by na przykład śnieg zalegający na stacji ładowania? Analogicznie, jak zachowywałby się algorytm lokalizacji pojazdów, w przypadku śniegu zalegającego na pojeździe, bagażnika dachowego czy też przyczepy dołączonej do samochodu.

7. Podsumowanie i ocena końcowa

W ocenie recenzenta rozprawa doktorska mgr inż. Tomasza Nowaka zatytułowana „Data-Efficient And Explainable Machine Learning in Visual Perception for Autonomous Vehicles” **spełnia wymagania stawiane rozprawom doktorskim** zgodnie z art. 187 ust. 1 i 2 „Ustawy z dnia 20 lipca 2018 Prawo o szkolnictwie wyższym i nauce” (Dz. U. 2024 poz. 1571). Recenzent wnioskuję do Rady Dyscypliny Automatyka, Elektronika, Elektrotechnika i Technologie Kosmiczne Politechniki Poznańskiej o jej przyjęcie i dopuszczenie mgr inż. Tomasza Nowaka do dalszych etapów przewodu doktorskiego.

Praca prezentuje bardzo wysoki poziom zarówno naukowy, techniczny jak i edytorski. Doktorant opracował szereg nowatorskich metod i algorytmów, stanowiących rozwinięcie najnowszych metod dostępnych w literaturze, które zostały dodatkowo poparte analizą matematyczną jak i wynikami praktycznymi wykonanymi w warunkach zbliżonych do rzeczywistych. Dodatkowo doktorant jest:

- autorem publikacji w renomowanym czasopiśmie naukowym (Q2),
- autorem publikacji artykułu na konferencji znajdującej się w wykazie czasopism naukowych i recenzowanych materiałów z konferencji międzynarodowych (CORE B),
- wykonawcą w grantie finansowanym przez NCBR.

W ocenie recenzenta rozprawa doktorska przekracza istotnie średni poziom rozpraw doktorskich w zakresie rangi rozwiązywanego problemu badawczego oraz jakości i zakresu badań jak również ma wyjątkowe walory wdrożeniowe. Biorąc powyższe pod uwagę **recenzent stawia wniosek o wyróżnienie rozprawy doktorskiej.**

Piotr Lipiński