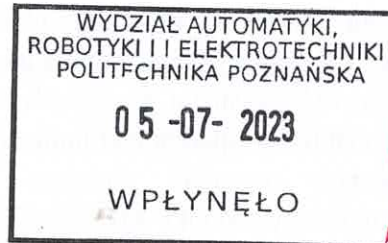


Tytuł, stopień, imię i nazwisko
prof. dr hab. inż. Bogdan Kwolek



data 27.06.2023 r.

PRZEWODNICZĄCY RADY DYSCYPLINY
Automatyka, Elektronika, Elektrotechnika
i Technologia Kosmiczne
prof. dr hab. inż. Wojciech Szelaąg

RECENZJA ROZPRAWY DOKTORSKIEJ
przygotowana dla Rady dyscypliny Automatyki, Elektroniki, Elektrotechniki
i Technologii Kosmicznych Politechniki Poznańskiej

Tytuł rozprawy: Combining interaction and perception to determine the physical properties of the robot environment

Autor rozprawy: mgr inż. Michał Bednarek

Promotorzy: dr hab. inż. Paweł Drapikowski, prof. PP, dr inż. Krzysztof Walas

1. Przedstawiona mi do recenzji rozprawa doktorska dotyczy percepcji na potrzeby przemieszczania się robotów oraz manipulacji obiektami w oparciu o dane sensoryczne. Prace badawczo-eksperymentalne obejmowały projektowanie oraz budowę krzepkich modułów pokładowych dla robotów operujących w środowiskach niedeterministycznych. Zasadniczą część prac dotyczyła klasyfikacji typu powierzchni kontaktu z robotem w oparciu o metody uczenia nadzorowanego, regresji parametrów fizycznych dotykanego obiektu, w tym jego sztywności. Znaczną część uwagi poświęcono rozwinięciu i udoskonaleniu technik uczenia nienadzorowanego dla lepszego rozumienia danych sensorycznych, a także wykorzystania ich bez nadzoru, m.in. w kontekście osiągnięcia wymaganej krzepkości systemu pokładowego oraz szeroko rozumianej odporności torów percepcyjnych operujących na rzeczywistych danych sensorycznych. Przedmiotem badań były sieci neuronowe, w tym także mechanizmy uwagi maszynowej oraz techniki fuzji wielomodalnych danych sensorycznych. Dane pozyskiwano w rzeczywistych scenariuszach z robotami kroczącymi oraz scenariuszach z ramionami robota wchodzącymi w interakcje ze środowiskiem. Podjęte w pracy badania nad grupowaniem danych haptycznych są bardzo perspektywiczne. Autorskie metody klasyfikacji typu powierzchni w trakcie kontaktu oraz metody regresji wybranych parametrów fizycznych dorównują skuteczności najbardziej konkurencyjnych metod znanych z literatury. Wspomniane metody są ogólne i mogą być użyteczne dla szerszej klasy problemów wymagających szybkich i efektywnych rozwiązań dla systemów operujących na danych sensorycznych. Autorska warstwa uwagi modalności MAL jest oryginalnym rozwiązaniem, której to istotą jest ważenie modalności w zależności od ich istotności, pozwala na zachowanie krzepkości systemu pokładowego w razie degradacji informacji sensorycznej pozyskiwanej w oparciu o jedną z modalności. Na dorobek Autora składają się także metody fuzji danych wielomodalnych, w tym także na bieżąco napływających niejednorodnych danych sensorycznych typu siły/momenty oraz obrazy. Weryfikację skuteczności rozwiązań do estymacji sztywności realizowano na syntetycznym zbiorze danych, który rozszerzono dodatkowo o rzeczywiste dane. Dla omawianego zagadnienia zaproponowano metodę opartą o techniki uczenia

nienadzorowanego, którą oparto o algorytm głębokiej klasteryzacji DEC (ang. *Deep Embeddings Clustering*). Oryginalnym i nowatorskim rozwiązaniem jest transformer HAPTR, który poprzez umiejętne wykorzystanie mechanizmów uwagi umożliwia nie tylko bardziej skuteczną klasyfikację podłoża w oparciu o strumienie danych sensorycznych IMU/haptic, ale co najważniejsze posiada znacznie mniejsze wymagania obliczeniowe. Pokazano eksperymentalnie, że uzyskane wyniki klasyfikacji są konkurencyjne w porównaniu do wyników uzyskiwanych przez najbardziej konkurencyjne algorytmy. Wyniki prac eksperymentalnych w obszarze fuzji wielomodalnych danych sensorycznych potwierdzają nie tylko znaczący potencjał fuzji pomiarów i danych w robotyce, ale także wskazują na ich znaczący potencjał w zakresie utrzymania krzepkości całego systemu w razie degradacji informacji sensorycznej dla jednej z modalności. Autor zaproponował oryginalne rozwiązania oraz zbudował kompletne systemy pokładowe dla percepcji środowiska przez robot i wyposażenia go w zdolności haptyczne. Teza wyraża się w przekonaniu, że percepcję robota, zdefiniowaną jako zdolność agenta robota do rozpoznawania złożonych wzorców w danych kinestetycznych oraz reprezentujących dotyk, można osiągnąć za pomocą głębokich metod uczenia sieci neuronowych z uwagi na to, że uzyskują one dobrą skuteczność i wydajność w kluczowych zadaniach szeroko rozumianej robotyki. Autor postawił tezę, że wspomnianą zdolność można dodatkowo rozszerzyć o metody uczenia bez nadzoru, które pozwalają analizować napływające dane i mierzyć ich podobieństwo do wcześniej pozyskanych danych sensorycznych, co poprawia ogólne zrozumienie środowiska robota. Tezy pracy są konsekwentnie rozwijane w rozprawie. Prawdziwość tez została wykazana empirycznie. Tematyka podjęta w rozprawie jest w pełni uzasadniona pod względem poznawczym, a na płynące z niej wnioski istnieje zapotrzebowanie w obszarze robotyki i sztucznej inteligencji. Wyniki badań eksperymentalnych są rzetelne i przekonujące.

2. Rozprawa napisana jest w języku angielskim. Praca składa się z czterech zasadniczych części w których kolejno wprowadzono w problematykę pracy, omówiono metody, zaprezentowano wyniki badań eksperymentalnych oraz podsumowano rozprawę. Rozprawa liczy łącznie 127 stron i jej treść pracy podzielono na 12 rozdziałów. Bibliografia liczy łącznie 156 pozycji. Streszczenia pracy przedstawiono zarówno w języku polskim jak i angielskim. Na wstępie omówiono dość szczegółowo układ i zawartość pracy. W rozdziale pierwszym zawarto ogólne wprowadzenie do tematyki rozprawy, przedstawiono znaczenie problematyki badawczej, a następnie postawiono tezy badawcze. W kolejnych trzech rozdziałach składających się na drugą część pracy omówiono zadanie rozpoznawania materiału, rozpoznawania podłoża, a następnie fuzję wielomodalnych danych sensorycznych. W rozdziale otwierającym tę część pracy zaprezentowano pokrewne prace w zakresie pomiarów oraz estymacji sztywności, rozpoznawania materiałów, a następnie nakreślono tło dla badań związanych z metodami nienadzorowanymi. Po zaprezentowaniu proponowanego rozwiązania szczegółowo omówiono estymację sztywności w podejściu nadzorowanym, a następnie rozpoznawania dotyku w podejściu nienadzorowanym. Zarówno dla podejścia nadzorowanego jak i nienadzorowanego scharakteryzowano zbiory danych. W rozdziale poświęconym rozpoznawaniu podłoża, po prezentacji pokrewnych prac omówiono proponowane rozwiązanie, a następnie szczegółowo zaprezentowano Autorskie rozwiązanie oparte o transformer, w tym także zgromadzone dane oraz konkurencyjne metody. W rozdziale zamykającym drugą część pracy szczegółowo omówiono fuzję wielomodalnych danych sensorycznych w kontekście pokrewnych prac, w świetle proponowanego rozwiązania, w tym także z perspektywy robotyki. Prezentacja

zagadnień i własnego wkładu według konsekwentnie stosowanego schematu: wprowadzenie, pokrewne prace, proponowane rozwiązanie, omówienie metod z naciskiem na klarowne sformułowanie problemu i omówienie własnej metody, charakterystyka zbiorów danych, a następnie uzasadnienie wyboru metod do porównania i ewaluacji systematyzuje przekaz, pokazuje potencjał opracowanych i zweryfikowanych praktycznie rozwiązań, a w szczególności wskazuje oryginalne aspekty rozprawy w kontekście aktualnego stanu wiedzy. Prezentacja wyników uzyskanych w pracach eksperymentalno-badawczych ma miejsce w trzeciej części pracy, w której każdy z trzech rozdziałów nawiązuje tematycznie do rozdziału w części drugiej rozprawy. W rozdziale związanym z rozpoznawaniem materiału, po omówieniu metryk MAE i MAPE zaprezentowano wyniki uzyskane przez sieci CNN, CNN-LSTM oraz CNN-BiLSTM na sygnałach symulujących ściskanie obiektu. Następnie dla sieci neuronowej, która uzyskała najlepsze wyniki w regresji trenowano jedną sieć dla sygnałów inercyjnych z pomiarów w trakcie ściskania obiektów o trzech różnych kształtach (piłki, pudełka i cylindra), a następnie wyznaczano i porównywano uzyskane wartości metryk MAE i MAPE. W końcowej fazie tej części eksperymentów wyznaczano wartości wspomnianych metryk dla sieci trenowanych na pomiarach z syntetycznego zbioru danych wzbogaconych dodatkowo o wartości sygnałów z pomiarów rzeczywistych. W dalszej części rozdziału zaprezentowano wyniki klasteryzacji danych na zbiorze Touching zarejestrowanym w niniejszej pracy, który zawiera optyczne pomiary siły uzyskane za pomocą trzyosiowego sensora siły OptoForce. W końcowej części omawianego rozdziału zamieszczono wyniki prac, które zrealizowano celem zweryfikowania prawdziwości hipotezy o wymiarowości dotykowej przestrzeni tekstury. Wyniki badań eksperymentalnych zaprezentowano w czterech tabelach oraz na kilku wykresach. W rozdziale związanym tematycznie z klasyfikacją podłoża zaprezentowano wyniki uzyskane przez proponowany HAPTR odnosząc je przy tym do wyników uzyskiwanych przez konkurencyjne metody. Ewaluację oraz porównanie skuteczności i efektywności realizowano na repozytorium danych PUT ANYmal. W dalszej części rozdziału zamieszczono wyniki uzyskane w badaniach eksperymentalnych nad krzepkością percepcji. W oparciu o Autorski MAL analizowano wyniki uzyskiwane przez HAPTR w symulowanych scenariuszach z przenoszeniem ładunku oraz symulowanej degradacji pomiarów sensorycznych. W końcowej części rozdziału uzyskane wyniki odniesiono do wyników uzyskiwanych przez konkurencyjne metody. Wyniki zaprezentowano w czterech tabelach, sześciu wykresach oraz na jednej macierzy pomyłek. W rozdziale związanym tematycznie z fuzją danych zaprezentowano wyniki badań eksperymentalnych uzyskane na zbiorach BiGS, HaTT oraz PHAC-2. Przedmiotem prac eksperymentalnych były cztery metody fuzji danych, które analizowano w kontekście krzepkości systemu oraz augmentacji danych. W analizach porównawczych porównywano dokładności klasyfikacji, a także wartości metryk AUC oraz AUC-ROC. Następnie ze wsparciem map ciepła (ang. *heatmaps*) wizualizowano dokładności uzyskane przez każdą z czterech metod w zależności od stopnia degradacji danych i modalności podlegających fuzji. Zaobserwowano, że w analizowanych zbiorach danych da się wyróżnić dominującą modalność, która ma najbardziej decydujący wpływ na separację między klasami. Zaobserwowano również, że metody fuzji Mid oparte na łączeniu cech wyznaczonych przez sieci neuronowe typu enkoder umożliwiają uzyskanie najlepszych wyników. Na potrzeby ewaluacji metod wyznaczono wartości dokładności klasyfikacji i AUC-ROC. Wyniki zaprezentowano w sześciu tabelach oraz na dziesięciu mapach ciepła. Wnioski podsumowujące wyniki badań eksperymentalnych zaprezentowano w trzech rozdziałach. Od strony formalnej rozprawa zredagowana jest poprawnie, z właściwą systematyką podjętych problemów naukowych. Następstwo rozdziałów

jest właściwe, tytuły rozdziałów i podrozdziałów dają syntetyczny pogląd na zawartą w nich treść. Prezentacja podjętej problematyki badawczej jest wyczerpująca. Autor precyzyjnie i poprawnie formuje wnioski wynikające ze zrealizowanych prac eksperymentalnych. Wnioski podsumowujące rozprawę mają oparcie w proponowanych rozwiązaniach oraz uzyskanych wynikach empirycznych. Nowoczesny i bogaty warsztat badawczy oraz kreatywne spojrzenie na zagadnienie percepcji robotów zaowocowały wartościowym opracowaniem naukowym. Zagadnienia podjęte w recenzowanej rozprawie są ważne nie tylko w aspekcie naukowym, ale przede wszystkim w aspekcie praktycznym. Proponowane rozwiązania mają potencjał, który może sprawić, że będą one stymulować dalszy rozwój technologii haptycznych.

3. Tematyka podjęta w pracy jest istotna. Percepcja środowiska przez autonomiczne maszyny i roboty jest intensywnie rozwijającą się dziedziną nauki. Ze względu na dynamiczny rozwój jest domeną wielu prac badawczych i zaawansowanych projektów inżynierskich. Jednym z kluczowych zadań w tym obszarze, w szczególności ze względu na kluczową rolę w pojazdach autonomicznych jest fuzja wielomodalnych danych i pomiarów sensorycznych. Obecne znane rozwiązania dla systemów percepcyjnych umożliwiają w dość ograniczonym stopniu rozpoznawanie i rozumienie wzorców i zależności w danych kinestetycznych oraz reprezentujących dotyk lub szerzej ujmując - właściwości taktylnych otoczenia agenta. Wspomniany kontekst pracy znacząco podnosi wagę uzyskanych wyników, a w szczególności umiejscawia niniejszą pracę w nurcie nowoczesnych badań. Metody i rozwiązania zaproponowane w niniejszej pracy oparte są o uczenie głębokie. Zaproponowano oryginalne rozwiązania w obszarze percepcji dynamicznego otoczenia, których efektywność i skuteczność rozpoznawania lub klasyfikacji nie ustępuje lub przewyższa rozwiązania konkurencyjne. Ewaluacji dokonywano na ogólnie dostępnych zbiorach danych, także na wymagających zbiorach danych zarejestrowanych w niniejszej pracy, które są ogólnie dostępne. Warto wspomnieć, że Autor udostępnił także kody źródłowe.

Najważniejszymi osiągnięciami Doktoranta są zarejestrowane i upublicznione zbiory danych, które wypełniają nie tylko lukę i są odpowiedzią na istniejące zapotrzebowanie gwałtownie rozwijającej się dyscypliny, ale wnoszą nową jakość w obszarze percepcji dynamicznego otoczenia agenta, m.in. poprzez możliwość szacowania sztywności materiałów w oparciu o sensory intercyjne oraz interakcję z obiektami. Istotne są rezultaty prac nad wymiarami dotykowymi w danych haptycznych uzyskane na upubliczniczonych zbiorach z uwagi na potencjalne możliwości rozwoju systemów rozpoznawania dotyku w sposób nienadzorowany, co jest stosunkowo słabo rozwiniętym obszarem badawczym. Godne uwagi są także zaproponowane warstwy uwagi do klasyfikacji terenu dla robotów kroczących, które umożliwiają poprawę skuteczności klasyfikacji w porównaniu z konkurencyjnymi metodami, w szczególności dla zaszumianych pomiarów, w dodatku przy znacząco niższym zapotrzebowaniu na moc obliczeniową. Co więcej, omawiane mechanizmy mogą być użyteczne do określenia istotności strumieni w wielomodalnych pomiarach sensorycznych, w tym także na potrzeby fuzji danych. Namacalnym wkładem w rozwój dyscypliny są także uzyskane wyniki w zakresie oceny typu szumów i czynników degradujących jakość klasyfikacji szeregów czasowych. Na wysoką ocenę pracy składa się także wysoka jakość prezentacji uzyskanych wyników, mnogość przeprowadzonych eksperymentów badawczych, a także bogaty materiał ilustracyjny.

Pojęta w pracy problematyka jest istotna, nie tylko w aspekcie naukowym, ale przede wszystkim z uwagi na upublicznione repozytoria danych, które przyczynią się do realizacji

zawansowanych systemów percepcji otoczenia agenta. Zaprezentowane wyniki badań oraz ich wnikliwa analiza potwierdziły prawdziwość stawianych w pracy hipotez. Należy docenić znaczący nakład pracy jaki wykonał Autor rozprawy oraz to, że proponowane rozwiązania są oryginalne. Rozprawa jest ważnym przyczynkiem w zakresie wiedzy dotyczącej percepcji dynamicznego otoczenia robota, a przedstawione w niej podejście może mieć znaczący wpływ na dziedzinę. Autorskie propozycje rozwiązań wskazują na dobrą znajomość zagadnień związanych z przedmiotem rozprawy. Wyniki badań eksperymentalnych są spójne, przekonujące i należyście udokumentowane. Wyniki prac opublikowano w czternastu współautorskich publikacjach oraz jednej autorskiej pracy, m.in. *Robotics and Autonomous Systems* z IF=3.7, *Sensors* z IF=3.8, *Electronics* z IF=2.7, a także materiałach flagowych konferencji, w tym także ICRA. Na podkreślenie zasługuje współpraca z bardzo dobrymi ośrodkami badawczymi i odbyte staże, m.in. EPFL, Oxford University, a także zaangażowanie w realizację projektów naukowych H-2020 (THING, REMODEL), w tym także grantu LIDER.

4. Poniższe uwagi odnoszą się do bardziej szczegółowych kwestii związanych z realizacją celu rozprawy i postawionej tezy, nie są zarzutami i mają na ogół polemiczny, a nie krytyczny charakter, nie ujmujący jednak wartości pracy. Pracę wzbogaciłaby dyskusja nt. złożoności obliczeniowej modeli sieci opartych na mechanizmach uwagi, w tym także sieci typu transformer, także zapotrzebowania na pamięć oraz liczby parametrów trenowanych modeli, w szczególności w razie konieczności operowania na wielowymiarowych szeregach czasowych. Interesujące byłoby przedyskutowanie zagadnień i możliwości odkrywania związków czasowo-przestrzennych, w szczególności dla pomiarów z wielu elektrod pomiarowych. Interesujące byłoby także zobrazowanie związków odkrytych przez mechanizmy uwagi, w tym także sygnałów z wykorzystywanych sensorów i elektrod pomiarowych. Od strony poznawczej pracę wzbogaciłaby dyskusja na temat mechanizmów wykorzystywanych przez organizmy żywe do krzepkiej i niezawodnej percepcji środowiska, w szczególności w kontekście zagadnień podjętych w niniejszej rozprawie, np. pamięci sensorycznej, w tym także pamięci zmysłowej, haptycznej pamięci sensorycznej, sprzężenia haptycznego, itp. Pracę wzbogaciłoby niewątpliwie szersze, niż to ma w niej miejsce, odniesienie się do wymagań stawianych przez systemy czasu rzeczywistego, w szczególności w kontekście sterowania, także ze sprzężeniem zwrotnym. W pracy brak jest doprecyzowania, czy powodem użycia metryki AUC-ROC dla ewaluacji skuteczności klasyfikacji, np. na zbiorze PHAC-2, było niezbalansowanie klasyfikowanych danych, czy też klasyfikacja danych „multi-label”. Rozprawie z pewnością nie zaszkodziłoby przybliżenie wyznaczania metryki AUC-ROC na wspomnianych danych, w tym także uśredniania typu „macro” i „weighted”. Pracy nie zaszkodziłoby z pewnością staranniejsza korekta edytorska (np. Bag-of-SFA Smbols -> Bag-of-SFA Symbols (str. 8 i 44), Chapter 3 presents the multi-modal -> Chapter 4 ..., Atomated Guided Vehicle -> Automated ..., dodanie opisów odciętych, np. na Rys. 2.2, niepoprawne użycie przedimków a, an, the, np. na str. 32 w odniesieniu do stiffness coefficient k , używania symboli matematycznych przed ich zdefiniowaniem, np. $I_{\{KL\}}$, brakujących cytowań, np. Kalman Filter (str. 57)). Wartość pracy z pewnością wzrosłaby po zamieszczeniu przykładowych zdjęć z wykorzystywanych repozytoriów danych, a wręcz nawet z korespondującymi sygnałami, np. w kontekście zdania: „Each photo received a corresponding haptic signal f ”.

Powyższe uwagi mają przede wszystkim charakter dyskusyjny i nie umniejszają w najmniejszym stopniu wartości pracy, która stanowi oryginalny wkład w dziedzinę robotyki i inteligencji maszynowej.

5. Przewieszona mi do recenzji rozprawa doktorska mgr. inż. Michała Bednarka proponuje skuteczne i efektywne rozwiązania ważnych zagadnień w obszarze percepcji otoczenia agenta. Rozprawa potwierdza znaczącą wiedzę Doktoranta w obszarze robotyki oraz szeroko rozumianej sztucznej inteligencji, umiejętność samodzielnego prowadzenia badań naukowych, a także dużą sprawność w zakresie projektowania, realizacji i ewaluacji złożonych systemów robotycznych. Rozprawa jest ważnym przyczynkiem w zakresie wiedzy dotyczącej percepcji dynamicznego otoczenia robota, a przedstawione w niej rozwiązania oraz udostępnione pomiary sensoryczne mogą mieć znaczący wpływ na dziedzinę. W podsumowaniu stwierdzam, że recenzowana rozprawa mgra inż. Michała Bednarka spełnia wszystkie wymagania ustawowe zwyczajowe stawiane rozprawom doktorskim. Stwierdzam niniejszym, że przedstawiona mi do recenzji rozprawa doktorska spełnia kryteria określone w art. 13 ust. 1 Ustawy. Wnoszę o dopuszczenie Pana mgr. inż. Michała Bednarka do kolejnych etapów postępowania o nadanie stopnia doktora nauk inżynieryjno-technicznych w dyscyplinie automatyka, elektronika i elektrotechnika. Ponadto, wysoki poziom merytoryczny i potencjał praktyczny zaprojektowanych i zrealizowanych rozwiązań są przesłanką do wnioskowania do Rady Dyscypliny Automatyki, Elektroniki, Elektrotechniki i Technologii Kosmicznych Politechniki Poznańskiej o wyróżnienie rozprawy doktorskiej Pana mgr. inż. Michała Bednarka.

Boydou Kwobll