

dr hab. inż. Maciej Majewski, prof. PG
Politechnika Gdańska
Wydział Inżynierii Mechanicznej i Okrętownictwa
Instytut Technologii Maszyn i Materiałów
Zakład Technologii Maszyn i Automatykacji Produkcji

RECENZJA ROZPRAWY DOKTORSKIEJ

mgr inż. Piotra Pawła SKROBKA
z Wydziału Mechanicznego Technologicznego
Politechniki Warszawskiej

nt.: METODA WIZYJNEJ IDENTYFIKACJI OBIEKTÓW
DLA ROBOTÓW STEROWANYCH GŁOSOWO

Promotor: dr hab. inż. Adam Rogowski
Profesor PW

dr hab. inż. Maciej Majewski, prof. PG
Politechnika Gdańska
Wydział Inżynierii Mechanicznej i Okrętownictwa
Instytut Technologii Maszyn i Materiałów
Zakład Technologii Maszyn i Automatykacji Produkcji

Gdańsk, dn. 22.12.2022r.

RECENZJA ROZPRAWY DOKTORSKIEJ

mgr inż. Piotra Pawła SKROBKA
z Wydziału Mechanicznego Technologicznego Politechniki Warszawskiej

nt.: METODA WIZYJNEJ IDENTYFIKACJI OBIEKTÓW
DLA ROBOTÓW STEROWANYCH GŁOSOWO

Promotor: dr hab. inż. Adam Rogowski
Profesor PW

Podstawa opracowania recenzji: pismo Dziekana Wydziału Mechanicznego Technologicznego Politechniki Warszawskiej z dnia 11 października 2022 r. (MT.521.7.2022), do którego dołączono egzemplarz rozprawy doktorskiej.

1. Zakres i charakterystyka rozprawy

Przedstawiona do recenzji rozprawa zawiera 138 stron, a w tym: 4 strony streszczenia, 1 strona słownika ważniejszych akronimów, 2 strony spisu treści, 6 stron bibliografii (108 pozycji), 15 stron załącznika oraz 2 strony spisu rysunków. Rozprawa składa się z 7 rozdziałów oraz mieści 33 rysunki i 12 tabel. Praca obejmuje obszerną i dobrze udokumentowaną analizę literatury. Rozprawa zawiera starannie opracowane i przedstawione zbiory oraz struktury danych stanowiące zasoby eksperymentalne dla opracowanych metod i przeprowadzanych analiz.

We wprowadzeniu – rozdział 1 (str. 13 – str. 19) – Autor przedstawia genezę problemu i krótką charakterystykę aktualnie prowadzonych badań z zakresu rozpoznawania mowy oraz analizy obrazu w robotyce i zastosowań robotów współpracujących. Możliwość głębszego poznania wymagań dotyczących metod

służących do automatycznej analizy obrazu uwzględniających najlepiej specyfikę współpracujących robotów przemysłowych stała się główną przesłanką do podjęcia przez Doktoranta własnych badań i eksperymentów. Pozytywne wyniki wstępnych testów wskazywały na zasadność poszerzenia zakresu prac z ukierunkowaniem na opracowanie metod wykorzystujących sieci neuronowe i wymagających minimalnej ilości danych, a którego efektem powinien być rozwój zastosowań systemów sterowania głosowego robotów przemysłowych posiadających możliwość automatycznej analizy obrazu. Zaproponowanie metodyki do opracowywania uniwersalnych sygnatur zarysu dla narzędzi i innych obiektów w celu rozpoznawania ich przez sieci neuronowe otwiera dalsze możliwości rozwoju systemów wizyjnych i wspomagania pracy operatora w przemyśle.

Biorąc pod uwagę poznawcze aspekty i potencjalny zakres wdrożenia kompleksowego rozwiązania do wizyjnej identyfikacji obiektów dla współpracujących robotów przemysłowych sterowanych głosowo, wpisującego się w innowacyjne technologie według koncepcji Przemysłu 4.0, wybór tematu pracy doktorskiej uznaję za trafny i w pełni uzasadniony do realizacji.

Rozdział 2 (str. 20 – str. 65) obejmuje obszerną (46 stron) analizę stanu badań dotyczących metod rozpoznawania mowy i analizy obrazu w robotyce, wykorzystania kontekstu w komunikacji głosowej z maszynami, zastosowań sygnatur w rozpoznawaniu obiektów i klasyfikacji ich na obrazie oraz sposobów wykorzystania robotów współpracujących. Analiza literatury obejmuje również wybrany zakres metod uczenia sieci neuronowych, analizy typów tych sieci, narzędzi służących do realizacji eksperymentów z nimi oraz niektóre zadania optymalizacji modeli sieci neuronowych.

Analizując postępy i wyniki badań w publikacjach, Doktorant dochodzi do słusznej konkluzji stwierdzając, że w komunikacji z przemysłowymi robotami współpracującymi za pośrednictwem głosu poprzez wydawane komendy przez operatora, obiekty występujące w kontekście wydawanych poleceń, można identyfikować za pomocą systemu wizyjnego na podstawie specjalnie generowanych zestawów sygnatur, korzystnie z wykorzystaniem metod i technik sztucznej inteligencji, a w szczególności sieci neuronowych. Jednym z zasadniczych problemów naukowych jest opracowanie metody wizyjnej identyfikacji obiektów manipulowanych przez sterowane głosowo roboty przemysłowe, na podstawie automatycznie generowanych sygnatur elastycznych konturów wzorcowych odpowiadających pojęciom ogólnym, do których odwołują się komendy.

Opracowana metoda może przynależeć swoją funkcjonalnością do przyszłościowych systemów interakcji człowiek-maszyna (a w tym człowiek-robot), mogących przetwarzać polecenia głosowe operatora wydawane za pośrednictwem języka naturalnego oraz budować wiedzę o realizowanych procesach z dokonywaną analizą ich

stanu i oceny bezpieczeństwa, a ponadto wspomagać operatora w podejmowaniu decyzji w warunkach niepewności i niepowtarzalności procesów.

Doktorant ujął podjętą problematykę jako „integracja systemu komunikacji głosowej człowiek-robot przemysłowy, z systemem automatycznej analizy obrazu, zapewniającym kontekst dla komend operatora”. Cel użyteczny zakłada, że realizacja systemu automatycznej analizy kontekstu komend głosowych dla przemysłowych robotów współpracujących, na podstawie danych z systemu analizy obrazu będzie miała korzystny wpływ na produkcję małoseryjną. Główny wkład naukowy w postępy w rozwiązaniach z tego zakresu stanowi opracowanie algorytmu, wykorzystującego sieć neuronową, umożliwiającego identyfikację obiektów na podstawie porównywania zarysów ich obrazów z elastycznymi edytowalnymi wzorcami konturów (FECT). Pan mgr inż. Piotr Skrobek przedstawił również te zagadnienia na podstawie starannej analizy badań wykonywanych w różnych ośrodkach naukowych.

Syntezę stanu dotychczasowych badań dotyczących zagadnień integracji systemów sterowania głosowego dla robotów przemysłowych z systemami analizy obrazu doktorant ujął w opracowanych wnioskach z analizy stanu wiedzy na końcu rozdziału. Wykazał, że dla rozwoju tych rozwiązań konieczne są dalsze badania – w tym badania eksperymentalne.

Analiza tej części opracowania potwierdza, że Doktorant opanował na wymaganym i wysokim poziomie umiejętność analizy oraz syntezy złożonych problemów badawczych. Wobec tego, jest odpowiednio przygotowany do podjęcia zaproponowanego tematu rozprawy doktorskiej.

W rozdziale 3 (str. 66 – str. 67) zostały przedstawione **teza, cel badawczy i użyteczny oraz zakres pracy**. Główny cel pracy został sformułowany następująco: „opracowanie metody wizyjnej identyfikacji obiektów manipulowanych przez sterowane głosowo roboty przemysłowe, na podstawie automatycznie generowanych sygnałów elastycznych konturów wzorcowych odpowiadających pojęciom ogólnym, do których odwołują się komendy”.

Uwaga dyskusyjna:

- Według mnie warto dodać, że opracowana metoda z zastosowaniem sieci neuronowej umożliwia uogólnianie wiedzy dotyczącej sygnałów konturów i ułatwianie komunikacji w języku naturalnym wykorzystującej podobne znaczeniowo polecenia mające różne cechy składniowe uwarunkowane leksykalnie, co wymaga większej tolerancji i czułości metod wykorzystujących wzorce.

Cel użyteczny został również sformułowany, który stanowi „ułatwienie – w warunkach produkcji małoseryjnej – realizacji systemu automatycznej analizy

kontekstu komend głosowych dla przemysłowych robotów współpracujących, na podstawie danych z systemu analizy obrazu". Za ważne etapy rozwiązania podjętej pracy uznają opracowanie struktur i istotnych zestawów danych oraz wykonanie prototypu w pełni funkcjonalnej aplikacji wizyjnej identyfikacji obiektów znajdujących się w polu widzenia kamery.

W badaniach eksperymentalnych zostało przeprowadzone porównanie efektywności opracowanej metody klasyfikacji obiektów z efektywnością metod wykorzystujących sieci konwolucyjne. Wyniki tych prac posłużyły do sformułowania tezy (str. 66). Zawarta teza nie jest szczegółowo opisana, aczkolwiek oddaje istotę formułowanej tezy, a istotne szczegóły zawarte są w zakresie pracy.

Uwaga dyskusyjna:

- Czy dla zaproponowanej metody warto byłoby opracować warianty procedur przetwarzania danych dla zastosowań w identyfikacji obiektów tylko określonych rodzajów ze względu na wybrany proces, aby poprawić jej efektywność i zwiększyć różne możliwości zastosowań? ponieważ niektóre obiekty nie posiadają wielu znaczących cech w swoim obrysie lecz na różnych częściach swojej powierzchni, a cechy różnią się m.in. znaczeniem i istotnością.

Zakres pracy został jednoznacznie sprecyzowany i ujęty w ośmiu punktach. Doktorant zaproponował przeprowadzenie badań obejmujących analizę stosowalności opracowanego rozwiązania według złożonej koncepcji, zaprojektowanie sieci neuronowej służącej do klasyfikacji, opracowanie danych wejściowych do sieci i wzorców konturów dla szeregu typowych obiektów, przeprowadzenie eksperymentów uczenia sieci na podstawie wzorców, przeprowadzenie klasyfikacji obiektów z wykorzystaniem zdjęć z kamery, porównanie efektywności opracowanej metody klasyfikacji obiektów z innymi metodami stosującymi sieci konwolucyjne oraz zakończenie wnioskami odnośnie zastosowań metody.

Rozdział 4 (str. 68 – str. 72) zawierający charakterystykę elastycznych edytowalnych wzorców konturów i analizę możliwości ich zastosowania do generowania sygnatur opisywanych kształtów. W rozdziale zawarto wybrane opisy elastycznego szablonu w formacie FCD oraz przykładowe kontury obiektów opisanych za pomocą wspólnego wzorca elastycznego.

Wskazano również współautorską publikację (wyk. bib. poz. nr 79), w której przedstawiono strukturę i zasady tworzenia wzorców FECT oraz omówiono format FCD służący do opisu takich wzorców. Autorzy publikacji podkreślili, że klasyczne metody bazujące na analizie sygnatury pozwalają na prawidłowe przypisanie różniących się od siebie kształtów do jednej klasy tylko w przypadku, gdy kształty te mają ściśle zachowane proporcje. Natomiast zaproponowana metoda wykorzystująca FECT nie

zawiera takich ograniczeń i ma dodatkowo inne zalety umożliwiające wykorzystanie kognitywnych metod.

Opracowane rozwiązanie w postaci elastycznych edytowalnych wzorców konturów (FECT) oceniam jako nowatorskie i rozwojowe oraz istotne dla postępów w digitalizacji procesów i tworzenia reprezentacji różnych elementów dla potrzeb wirtualnych modeli.

Pytanie:

- Cel utylitarny rozprawy stanowi "ułatwienie rozwiązania zagadnienia integracji przemysłowo zorientowanego systemu komunikacji głosowej człowiek-robot z systemem rozpoznawania obrazu będącym źródłem kontekstu dla komend głosowych". Głównie oznacza to, że system wizyjny identyfikuje obiekty do których odnoszą się polecenia głosowe. Czy jest przewidywane również budowanie szerszego kontekstu służącego analizie procesu i wnioskowaniu o jego przebiegu, na podstawie zbioru kolejno zidentyfikowanych obiektów w sekwencji poleceń wydanych przez operatora, a co za tym idzie np. możliwość rozpoznawania niektórych procesów i ewentualna automatyczna konfiguracja urządzenia technologicznego np. obrabiarki, w różnych celach wspomaganie operatora, a ponadto zwiększania jakości i dokładności procesu ?

W rozdziale 5 (str. 73 – str. 89) dotyczącym identyfikacji obiektów na podstawie FECT zostały zawarte opisy algorytmu dopasowywania segmentów rozpoznawanych konturów do segmentów FECT wraz z jego istniejącymi ograniczeniami oraz przedstawiona koncepcja zastosowania sieci neuronowej w tej identyfikacji.

Należy nadmienić, że ww. algorytm dopasowywania segmentów został szczegółowo przedstawiony w publikacji dr hab. inż. Adama Rogowskiego prof. PW i mgr inż. Piotra Skrobka oraz umieszczony na wykazie bibliograficznym rozprawy pod nr 79. Praca ta ukazuje dwa istotne algorytmy do klasyfikacji kształtu ze zgrubnym i szczegółowym dopasowaniem. W publikacji tej zawarto również strukturę danych dla FECT i analizę jej wykorzystania, a ponadto pokazano stanowisko badawcze określone jako „zrobotyzowana komórka obróbcza wyposażona w systemy rozpoznawania mowy i obrazu”, w tym również na filmie dostępnym jako materiały dodatkowe. Na opublikowanym filmie przedstawiono kolejne kroki działania systemu głosowej komunikacji wspieranego działaniem systemu wizyjnego sprzężonego z przemysłowym robotem wspomagającym operatora, a w tym wykorzystanie algorytmu do identyfikacji opisu konturu w formacie FCD. Film ukazuje etapy rozpoznawania charakterystycznych punktów konturu w celu wyizolowania jego pojedynczych segmentów, a następnie odnajdywania punktów charakterystycznych z użyciem wykresu sygnatury konturu i ich pochodnych. W następnym kroku wszystkie kontury są dopasowywane do opisu obiektu w formacie VCD. Dokumentacja eksperymentu zawiera również pliki ukazujące następujące opracowane zasoby:

- elastyczne szablony konturów wybranych typów narzędzi w formacie FCD;
- przykładowe obrazy i odpowiadające im wykresy sygnatur;
- obrazy narzędzi użytych w eksperymentach, pozyskanych z kamery i po binaryzacji jako dane wejściowe do przetwarzania.

Opublikowane materiały jako dodatek do publikacji związanej z tematem rozprawy oceniam jako przydatne i wartościowe, ponieważ stanowią dobrze dobrane i starannie opracowane zasoby pozwalające na przeprowadzenie istotnych eksperymentów i prób oraz wyciągnięcie konkluzji. Ponadto te zasoby umożliwiają zaczerpnięcie z doświadczeń i powtórzenie eksperymentów poprzez wykorzystanie próbek i opracowanych struktur danych.

W ramach badań doświadczalnych Autor stawia cel przeprowadzenia eksperymentów z wykorzystaniem dwóch metod:

1. metody z automatycznym wygenerowaniem wielkiej liczby obiektów każdej klasy na podstawie opisu ich zarysów w postaci elastycznych opisów konturu w formacie FCD, a następnie potraktowanie tych obrazów jako danych wejściowych do konwolucyjnych sieci neuronowych do klasyfikacji obiektów na obrazie;
2. metody z wykorzystaniem opracowanego algorytmu generującego automatycznie szereg sygnatur na podstawie elastycznych edytowalnych szablonów konturów (FECT) jako danych uczących, umożliwiającego uwzględnienie znacznej zmienności parametrów konturów obiektów tej samej klasy i wykorzystanie minimalnej liczby niepowtarzalnych danych wejściowych dla sieci neuronowych do klasyfikacji obiektów na obrazie.

Pytania dyskusyjne:

- Czy według Autora w pracy przydatny dla szerszego odbiorcy byłby schemat blokowy ukazujący tzw. potok przetwarzania danych (ang. *data processing pipeline*), nawet dla takiego przypadku małych danych (ang. *small data*), stanowiący jeden ze sposobów sekwencyjnego przetwarzania danych w cyklu w odrębnych blokach połączonych ze sobą, a mogących wnieść swoją procedurę? Schemat ukazałby poszczególne etapy przetwarzania danych w systemie i określałby ich format, a w tym przykładowo zawierałby bloki m.in. podziału danych wejściowych do sieci na zbiór uczący i zbiór testowy (opisany na str. 94), procesu i metody uczenia oraz oceny modelu.
- Jaki sposób okazał się najlepszym rozwiązaniem na dokonanie (według metody nr 1) „automatycznego wygenerowania wielkiej liczby obiektów każdej klasy na podstawie opisu ich zarysów w postaci elastycznych opisów konturu w formacie FCD”, co w założeniu odpowiada warunkom, w którym pozyskana jest wielka liczba obrazów zróżnicowanych obiektów tej samej klasy (np. wielu zdjęć

różnych kluczy tego samego rodzaju) z użyciem kamery ? W jaki sposób następuje zróżnicowanie obiektów danej klasy w sposób automatyczny ? To z reguły wymaga wprowadzenia automatycznie generowanych modyfikacji kształtu zarysu (np. zmian w wymiarach i krzywiznach fragmentów zarysu), w celu zwiększenia tolerancji metody klasyfikacji. Dążę w tym pytaniu do podkreślenia kolejnego osiągnięcia rozprawy.

Ekstrakcja wybranych cech sygnatury (inaczej istotnych elementów zarysu obiektu na obrazie) i zapisanie tych cech we wzorcach uczących, a następnie nauczanie sieci tymi wzorcami reprezentacji cech (ang. representation learning), stanowi sposób uogólnienia wiedzy o sygnaturze zakodowanej w wagach sieci.

W uczeniu głębokim przetwarzanie odbywa się przez strukturę sieci składającą się z wielu warstw wejściowych, wyjściowych i ukrytych, a dane wejściowe przekształcane są w wybrane informacje, które następne warstwy wykorzystują do wykonania zadań predykcyjnych. Wobec tego, struktura umożliwia uczenie się za pomocą własnego przetwarzania danych, a istniejąca różnorodność głębokich struktur sieciowych pozwala na wykorzystywanie niektórych sieci głębokiego uczenia do definiowania klas w celu ich porównania lub do rekonstrukcji uogólnionych obiektów (warstwy konwolucyjne i dekonwolucyjne). Uczenie głębokie zatem pozwala na uczenie sieci różnych kategorii i hierarchii cech w celu poprawienia wyników działania sieci, przykładowo dokładności wynikowej klasyfikacji obrazów występujących na bogatym i zróżnicowanym tle z wieloma elementami i szczegółami. Sieci głębokie poprawiają wyniki uczenia się dzięki stopniowemu i hierarchicznemu procesowi ekstrakcji cech z surowych danych, w tym przypadku obrazów pochodzących ze zdjęć. Przede wszystkim szukane są cechy: różnicujące, pewne lub niezmiennie. Algorytmy uczenia głębokiego umożliwiają m.in.: reprezentację podstawowych, wtórnych i wyższych rzędów cech oraz wykorzystanie różnych nadzorowanych i nienadzorowanych strategii uczenia poszczególnych warstw, a ponadto uzyskanie różnych poziomów abstrakcji dzięki podziałowi przetwarzania obrazu pomiędzy warstwy i ich podsieci. **Niestety proces ten stanowi wykorzystanie danych na wielką skalę.** Głębokie sieci konwolucyjne potrafią stopniowo filtrować różne części danych uczących i wyodrębiać ważne cechy w procesie dyskryminacji wykorzystanym do rozpoznawania lub klasyfikacji wzorców. Wykorzystanie metod głębokiego uczenia wymaga pozyskania dużych ilości danych tzn. dużej liczby obrazów ze zdjęć, a następnie progresywnej redukcji rozmiaru przestrzennego do zredukowania ilości cech i złożoności obliczeniowej sieci, przy czym możliwe jest również zaistnienie problemu przeuczenia sieci i utraty wrażliwości modelu.

W tym świetle metoda opracowana przez Doktoranta ma istotne znaczenie w aspekcie uzyskania specjalizacji systemu wizyjnego oraz wydajnego i jakościowego rozpoznawania specyficznych obiektów w powiązaniu z typem realizowanych procesów, a ponadto z pominięciem kłopotliwego tła obrazu.

Obszar zastosowań analizy obrazu ma istotne znaczenie przy opracowywaniu rozwiązań, ponieważ dedykowane metody mogą dawać nowe możliwości rozwoju systemów interakcji, w związku z wykorzystaniem kognitywnych technik.

Zaproponowana metoda umożliwia definiowanie nowych parametrów i cech elementów obrazów z systemów wizyjnych, które można szeroko konfigurować w systemach wykorzystujących komunikację głosową w języku naturalnym, ponieważ można nadawać wagi różnym cechom obiektów stanowiących narzędzia lub komponenty w związku z realizowanymi procesami i zadaniami.

W ramach rozprawy doktorskiej, zaproponowano zatem rozwiązanie problemu w sposób zautomatyzowany z minimalną ilością danych i małą złożonością obliczeniową. Metoda umożliwia ekstrakcję cech pod kontrolą i tworzenie klas obiektów w zależności od potrzeb związanych z komunikacją głosową w języku naturalnym i ewentualnym definiowaniem słowników. Wykorzystanie sygnatur może mieć również znaczenie w selekcji informacji identyfikującej dotyczącej wybranych cech obiektów i wykrywaniu zmian ich położenia, a także przy wykorzystaniu w technologii rozszerzonej rzeczywistości (ang. augmented reality), zwłaszcza w warunkach eksplozji danych i informacji oraz przeciążenia nimi operatora (ang. data and information explosion and overload).

Rozdział 6 (str. 90 – str. 109) dotyczy przebiegu i wyników badań eksperymentalnych. Metodę badawczą dla wieloetapowych badań Doktorant zawarł w 8 punktach (str. 90).

Opracowane koncepcje i metody przetwarzania danych w zaprojektowanych strukturach, a w tym dotyczące m.in. podziału konturów na sekcje dla poszczególnych wejść sieci, różnicowania podziału konturu na regularny i proporcjonalny, uwzględniania wyznaczonej krzywizny maksymalnej i średniej, różnicowania ze względu na charakter i istotność określonych cech, należy ocenić jako niezbędne i istotne w realizacji badań eksperymentalnych oraz niosące wkład w definiowanie zadań dla inżynierii cech z takiego zakresu wykorzystania danych dla metod i technik sztucznej inteligencji, a w tym sieci neuronowych.

Projektowanie zbioru uczącego i przebieg badań ukazuje wysokie umiejętności badawcze. W ramach optymalizacji danych wejściowych do sieci neuronowej przeprowadzono badania niezawodności rozpoznawania obiektów w dwóch wariantach: dla regularnego i proporcjonalnego podziału konturu oraz danych wejściowych w postaci krzywizny średniej i maksymalnej. Ponadto dokonano zbadania wpływu architektury sieci na wyniki rozpoznawania obiektów oraz porównania rezultatów opracowanej metody z tradycyjną metodą zastosowania sieci neuronowych w rozpoznawaniu obrazów. Badania eksperymentalne podsumowano analizą wyników i ich stosownym komentarzem.

Pytania dyskusyjne:

- W przypadku testów mających zweryfikować poprawność koncepcji i sprawdzić podstawowe algorytmy (rys. 27 na str. 94), które zostały zrealizowane przy użyciu środowiska Microsoft Visual Studio, warto podać wybrane elementy i cechy konfiguracji takiego środowiska, do których mogą należeć m.in. język programowania, biblioteki służące do wykorzystywania uczenia maszynowego lub sieci neuronowych, rodzaje wybranych sieci dla klasyfikacji i metod uczenia nadzorowanego.
- W przypadku testów mających na celu porównanie czasu potrzebnego do przeprowadzenia obliczeń i uczenia sieci dla opracowanej metody oraz standardowo stosowanych metod rozpoznawania obrazów wykorzystujących sieci konwolucyjne, wykorzystano środowisko Google Colab i język Python oraz biblioteki TensorFlow i Keras, nie wspomniano natomiast o zapewne wykorzystywanych plikach roboczych tego środowiska tzw. notatnikach Jupytera (ang. Jupyter Notebooks) ? Czy Autor rozpatrywał wykorzystanie w testach innych platform uczenia maszynowego w chmurze, również opartych o wykorzystanie Jupyter Notebooks jak wybrany Google Colab, a w szczególności Amazon SageMaker ?

W ramach rozwiązania problemu postawionego w rozprawie, zaproponowano podejście poprzez metodę z całokształtem jej procedur, która stwarza możliwości dla operatora-projektanta procesu obróbki, montażu lub innego, korzystającego z nowoczesnego systemu wytwarzania wyposażonego w roboty współpracujące i komunikację głosową wspomaganą systemem wizyjnym, w stosunku do istniejących już i stosowalnych metod (a w tym głębokiego uczenia), wykorzystania szerszego potencjału konfiguracji systemu wytwarzania.

Uzyskane efekty stanowią: umożliwienie uwzględniania specyfiki procesu i stosowanych urządzeń na poziomie wykorzystującym wszechstronne doświadczenie operatora-projektanta, poprzez zdefiniowane dostosowanych zbiorów danych sterujących i uczących, a tym samym budowanie neuronowych modeli wiedzy definiujących różne zależności w procesie, które występują w sposób jawny i niejawny.

Reasumując można powiedzieć, że aktualnie znane metody głębokiego uczenia dla systemów zautomatyzowanych stanowiące konkurencyjne rozwiązanie w stosunku do proponowanego, nie spełniają założonych wymagań jakościowych dla potrzeb procesów wytwarzania ze specjalistycznymi narzędziami albo komponentami. Tym samym można uznać, że metoda opracowana przez Doktoranta pod kierownictwem Promotora posiada istotne walory nowości i otwiera nowe znaczące drogi rozwoju dla systemów wytwarzania, a w tym inteligentnego wytwarzania (ang. smart manufacturing).

W rozdziale 7 (str. 100 – str. 113) zawierającym podsumowanie i wyznaczającym kierunki dalszych badań, zawarto również podkreślenie naukowego i aplikacyjnego znaczenia pracy. W ramach kierunków dalszych badań wskazano na dalsze innowacyjne i przyszłościowe prace badawcze.

W ramach **oceny poziomu redakcyjnego rozprawy** należy stwierdzić, że rozprawa została napisana bardzo ładnym i poprawnym językiem technicznym oraz prawidłowo zredagowana. W niektórych jej fragmentach występują drobne i nieliczne literówki, które zaznaczyłem bezpośrednio w pracy. Zapewne wynikają one z pośpiechu w pracach nad ukończeniem niektórych części rozprawy. Drobne uwagi dotyczące strony edytorskiej pracy są następujące:

Lp.	Strona:	Linijka:	Jest:	Powinno być:
1	9	22	ROSA	ROS
2	50	8	samo uczenia	samouczenia
3	56	6	ewolucyjnych	konwolucyjnych
4	60	6	puring	pruning
5	61	19	early stoping	early stopping
6	86	21	Lsi	l _{si}

W rozprawie występują również inne nieistotne uchybienia, które stanowią:

- Brak odnośników do: rys. 1 (str. 22), rys. 2 (str. 24), rys. 6 (str. 30), rys. 9 (str. 36), rys. 11 (str. 37), rys. 12 (str. 38);
- W podpisie rys. 16 (str. 47), odnosząc się do neuronów lepiej byłoby napisać „węzły” zamiast „okręgi”.

Należy jednak podkreślić, że Autor ustrzegł się bardzo wielu powszechnie popełnianych błędów redakcyjnych, a wykazane uchybienia są nieliczne w odniesieniu do objętości rozprawy. W rezultacie nie wpływają one na czytelność przekazywanych treści, a tym samym na poziom merytoryczny rozprawy i nie pomniejszają jej wartości.

Przedstawiona przeze mnie skrótowa charakterystyka opiniowanej rozprawy doktorskiej w pełni potwierdza, że Pan mgr inż. Piotr Paweł Skrobek gruntownie i kompleksowo przemyślał koncepcję podjętej rozprawy doktorskiej.

2. Ocena metodologicznej i metodycznej koncepcji rozprawy doktorskiej

Na podstawie przedstawionej analizy rozprawy doktorskiej i procedury rozwiązywania postawionych zadań badawczych, **metodologiczną i metodyczną koncepcję rozprawy doktorskiej oceniam w pełni pozytywnie**, ponieważ zawiera ona spójną merytorycznie analizę stanu techniki i rozwiązań służących do wizyjnej identyfikacji obiektów dla robotów przemysłowych sterowanych głosowo oraz opracowanie własnej koncepcji rozwiązania postawionego problemu badawczego, przeprowadzenie kompleksowych prac badawczych i eksperymentalnych, jako podstawy dla rozwiązań technicznych i zaleceń metodycznych.

Doktorant potwierdził bardzo dobre przygotowanie do prowadzenia prac badawczych o czym świadczą: rzeczowa, merytoryczna analiza i synteza literatury, szerokie wykorzystanie wiedzy z zakresu projektowania układów technicznych i algorytmów, modelowania danych, budowania systemów, opracowywania wyspecjalizowanych modułów i modeli uczących się oraz planowanie organizacji badań, weryfikacja w eksperymentach wytypowanych założeń teoretycznych, umiejętne i na wysokim poziomie opracowanie wyników badań.

3. Ocena końcowa rozprawy doktorskiej

W świetle dokonanej analizy i sformułowanych ocen rozprawy doktorskiej mgr inż. Piotra Pawła SKROBKA z Politechniki Warszawskiej nt.: METODY WIZYJNEJ IDENTYFIKACJI OBIEKTÓW DLA ROBOTÓW STEROWANYCH GŁOSOWO, której promotorem jest dr hab. inż. Adam Rogowski prof. PW, mogę stwierdzić, iż Autor:

- wybrał nowoczesną tematykę rozprawy, ważną dla rozwoju i wdrażania nowych metod wspomagania pracy operatora przy zastosowaniu systemu komunikacji głosowej zintegrowanego z systemem wizyjnym;
- opracował nowatorską metodę niespotykaną w komercyjnych systemach i przeprowadził poprawne badania - rozwiązał wiele pracochłonnych zadań oraz ważnych problemów badawczych, a opracowane analizy i otrzymane wyniki są na wysokim poziomie;
- opracował innowacyjne elastyczne edytowalne wzorce konturów (FECT) do wykorzystania przez metody sztucznej inteligencji;
- przyjął szeroki i spójny zakres zagadnień rozpatrywanych w rozprawie;
- wykazał się umiejętnościami w zakresie planowania badań eksperymentalnych i obiektywnością naukową w ocenie uzyskanych wyników;


- uzyskał wyniki, które mogą być wykorzystane w innych pracach badawczych i rozwojowych;
- w prawidłowy i metodyczny sposób przeprowadził proces dowodzenia tezy rozprawy doktorskiej;
- stworzył podstawy do dalszych badań i zastosowań.

Przedstawiona rozprawa doktorska należy do aktualnego i ważnego obszaru badawczego, związanego z integracją przemysłowo zorientowanych systemów głosowej komunikacji człowiek – robot współpracujący z systemami rozpoznawania obrazów, które wykorzystują metody i techniki sztucznej inteligencji. Autor wykonał wartościową pracę badawczą oraz wykazał się dogłębną znajomością warsztatu naukowego.

Pracę oceniam jako bardzo przydatną dla celów praktycznych. Według mnie jest to szczególnie ważne w obecnych czasach, kiedy oczekuje się konkretnych rozwiązań związanych z koncepcją Przemysłu 4.0. Autor w sposób jasny i jednoznaczny wykazał czego dokonał w trakcie realizacji zadań badawczych oraz doświadczalnych, potrafił wyciągnąć z nich poprawne i logiczne wnioski sformułowane w oparciu o odpowiednią podbudowę teoretyczną. Pan mgr inż. Piotr Paweł Skrobek opanował na wysokim poziomie współczesne metody organizacji badań i właściwe dla nich narzędzia informacyjne, służące do rozwiązywania złożonych, wielowariantowych problemów badawczych.

Opiniowana rozprawa doktorska, mieszcząca się obecnie w dyscyplinie „Inżynieria mechaniczna” (również „Budowa i eksploatacja maszyn”), posiada oryginalne cechy nowości. **Koncepcję rozwiązania problemów naukowych oceniam bardzo wysoko.** Zakres zagadnień ujętych w rozprawie jest komplementarny, a **uzyskane wyniki mają dużą wartość naukową i utylitarną.**

Na podstawie przedstawionej analizy stwierdzam, że rozprawa doktorska mgr inż. Piotra Pawła Skrobka **w bardzo dobrym stopniu spełnia wymagania stawiane pracom doktorskim.** Wnoszę o jej przyjęcie i dopuszczenie jej Autora do publicznej obrony.


.....
dr hab. inż. Maciej Majewski