

Bydgoszcz, 19.02.2021r.

dr hab. inż. Tomasz Andrysiak  
Profesor uczelni  
Wydział Telekomunikacji, Informatyki i Elektrotechniki  
Uniwersytet Technologiczno-Przyrodniczy w Bydgoszczy  
e-mail: [andrys@utp.edu.pl](mailto:andrys@utp.edu.pl)  
tel: +48 602 605 500

## RECENZJA ROZPRAWY DOKTORSKIEJ

dla Rady Dyscypliny Naukowej Informatyka Techniczna i Telekomunikacja  
Uniwersytetu Zielonogórskiego

Niniejsza recenzja została przygotowana w wyniku powołania na recenzenta rozprawy doktorskiej pt. „Przetwarzanie sekwencji wideo przy zastosowaniu podejścia deep learning do automatycznego generowania instrukcji stanowiskowych” przez Senat Uniwersytetu Zielonogórskiego na posiedzeniu w dniu 16 grudnia 2020 roku oraz na podstawie pisma Przewodniczącego Rady Naukowej Dyscypliny „Informatyka Techniczna i Telekomunikacja” Uniwersytetu Zielonogórskiego prof. dr hab. inż. Józefa Korbicza z dnia 21.12.2020 roku.

**Tytuł rozprawy:** Przetwarzanie sekwencji wideo przy zastosowaniu podejścia deep learning do automatycznego generowania instrukcji stanowiskowych.

**Autor rozprawy:** mgr inż. Daniel Halikowski.

**Promotor rozprawy:** dr hab. inż. Justyna Patalas Maliszewska, prof. UZ.

### Podstawa opracowania recenzji:

- otrzymany egzemplarz rozprawy doktorskiej,
- autoreferat doktoranta zawierający m. in. życiorys zawodowy oraz dorobek naukowy,
- pismo Przewodniczącego Rady Naukowej Dyscypliny „Informatyka Techniczna i Telekomunikacja” Uniwersytetu Zielonogórskiego prof. dr hab. inż. Józefa Korbicza z dnia 21.12.2020 roku.

### Zawartość treściowa rozprawy oraz analiza zagadnienia naukowego

Recenzowana rozprawa dotyczy zagadnień związanych z przetwarzaniem sekwencji wideo przy zastosowaniu podejścia związanego z uczeniem głębokim (deep learning) w celu automatycznej oceny aktywności pracownika na stanowisku roboczym. Głównym celem rozprawy jest opracowanie nowego modelu dla automatycznego generowania instrukcji stanowiskowych. Proponowane rozwiązanie poszukiwane jest dla działu utrzymania ruchu w przedsiębiorstwie, w którym prowadzone są szkolenia dla osób podejmujących zatrudnienie na danym stanowisku pracy.

Praca składa się z siedmiu rozdziałów, bibliografii oraz spisu tabel i rysunków. W rozdziale pierwszym przedstawiono wprowadzenie do tematyki rozprawy zawierające genezę, cel i zakres pracy oraz zarysowano problem badawczy. Rozdział drugi zawiera omówienie metod i technik uczenia głębokiego w odniesieniu do analizy i rozpoznawania obrazów. W trzecim rozdziale zaprezentowane zostały algorytmy klasyfikacji oraz przedstawiono metody dyskretyzacji szeregów czasowych jak również metody redukcji cech. Rozdział czwarty prezentuje autorską metodę automatycznego generowania instrukcji stanowiskowych, w szczególności opisuje elementy zaprojektowanego modelu oraz opracowane algorytmy. Rozdział piąty zawiera ocenę efektywności zaproponowanego modelu wraz z opisem trzech eksperymentów badawczych. W rozdziale szóstym przedstawiono zakres, funkcjonalność oraz działanie autorskiego systemu informatycznego zaprojektowanego do wspomaganie procesów szkoleń pracowniczych. Ostatni rozdział zawiera wnioski, podsumowanie pracy oraz kierunki dalszych prac rozwojowych.

W rozprawie zdefiniowano następujące pytania badawcze: *”Czy możliwe jest zbudowanie metody, której zastosowanie umożliwi monitorowanie i weryfikację czynności wykonywanych przez pracownika na stanowisku roboczym w przedsiębiorstwie produkcyjnym?”* oraz *„Czy możliwe jest zbudowanie takich algorytmów, które w wyniku integracji z CNN+SVM, CNN+R-CNN, YOLOv3 umożliwią wygenerowanie scenariusza postępowania dla pracownika szkolonego na danym stanowisku pracy?”*. Oba pytania badawcze zostały sformułowane poprawnie oraz adekwatnie od postawionego celu głównego oraz celów szczegółowych rozprawy.

Autor rozprawy w oparciu o postawione pytania badawcze sformułował następujące tezy pracy: *„Wykorzystanie podejścia deep learning w oparciu o CNN, CNN-SVM, R-CNN/YOLOv3 pozwoli na zbudowanie modelu przetwarzania sekwencji wideo do automatycznego generowania instrukcji stanowiskowych w dziale utrzymania ruchu”* oraz *„Zastosowanie autorskich algorytmów oraz ich implementacja w komputerowym systemie wspomaganie procesu szkoleń pozwoli na efektywną realizację szkoleń w dziale utrzymania ruchu”*. Obie tezy zostały sformułowane w sposób jasny, rzeczowy i odpowiedni do rozpatrywanego problemu badawczego. A także zostały udowodnione w rozprawie w oparciu o otrzymane wyniki eksperymentów komputerowych.

### **Przydatność rozprawy dla gospodarki narodowej**

Tematyka rozprawy mieści się w dyscyplinie naukowej Informatyka Techniczna i Telekomunikacja. Z punktu widzenia nauk technicznych przydatność rozprawy dotyczy przede wszystkim opracowania nowego modelu dla automatycznego generowania instrukcji stanowiskowych oraz autorskich algorytmów, jak również walidacji proponowanego rozwiązania w środowisku rzeczywistym. Opracowane przez Autora metody zostały zweryfikowane na drodze eksperymenty komputerowego w oparciu o stworzony system informatyczny wspomagający procesy szkoleniowe. Założenia przyjęte przez Autora rozprawy są uzasadnione i w szczególności dotyczą to takich aspektów jak: wykrywanie działań zarówno na podstawie podobieństw wyekstrahowanych cech obrazów sekwencji wideo, jak również zdefiniowanych obiektów pojawiających się na określonych obrazach. Wymienione założenia dotyczą z jednej strony aktualnych problemów badawczych, a z drugiej znajdują odzwierciedlenie w realnych potrzebach rozwiązywania problemów praktycznych. Autor zauważył również, że należy wykluczyć sytuację w których pomimo, że zbiory etykiet klas obiektów na obrazach referencyjnych i testowych badanych sekwencji wideo pokrywały się, to cechy klatek nie były wystarczająco podobne. Spostrzeżenie to pozwoliło w dużej mierze wyeliminować prawdopodobne niestabilności proponowanego rozwiązania związane z możliwą niejednoznacznością procesów rozpoznawania i klasyfikacji.

Tematyka rozprawy jest interesująca i w pełni uzasadniona a także wykazuje duży potencjał wykorzystania w rzeczywistych problemach dotyczących automatyzacji generowania instrukcji stanowiskowych w celu optymalizacji procesów szkoleniowych.

## **Charakterystyka rozwiązanie postawionego zagadnienia**

Autor rozwiązał zagadnienie zdefiniowane w tezach rozprawy. W tym celu sformułował własne autorskie koncepcje dotyczące: wyznaczania czynności i etapów czynności na podstawie wyodrębnionych klatek referencyjnych, wyznaczania cech i etykiet klas obiektów klatek referencyjnych oraz testowych, jak również wyznaczania zbioru instrukcji graficznych oraz kontroli przebiegu procesu serwisowania. Zastosowana metodologia sprowadza się do następującej sekwencji: zdefiniowania problemu badawczego, propozycji rozwiązania postawionego problemu, zdefiniowania środowiska eksperymentowania z wykorzystaniem adekwatnych do postawionego problemu metod i technik jak również testowych oraz produkcyjnych baz danych, analizy otrzymanych wyników z wykorzystaniem odpowiednich wskaźników jakości rozpoznawania i klasyfikacji. Tego typu postępowanie jest typowym podejściem stosowanym przez społeczność naukową w analizie zagadnień związanych z uczeniem maszynowym, tak więc zaprezentowane w dysertacji podejście jest prawidłowe i właściwe w kontekście tematyki rozprawy.

## **Charakter rozprawy**

Recenzowana praca ma charakter koncepcyjno-eksperymentalny. Autor zaproponował nowy model dla automatycznego generowania instrukcji stanowiskowych z wykorzystaniem podejścia związanego z uczeniem głębokim w celu automatycznej oceny aktywności pracownika na stanowisku roboczym. Uzyskane przez Autora wyniki potwierdzają postawione na wstępie pracy tezy badawcze, które zostały udowodnione w sposób eksperymentalny. Dysertacja zawiera niezbędne elementy eksperymentu komputerowego, które pozwalają na ocenę jakości wykorzystywanych metod uczenia maszynowego, a poruszana tematyka jest aktualna co do stanu wiedzy dotyczącego wykorzystania metod sztucznej inteligencji w zagadnieniach rozpoznawania i klasyfikacji obrazów.

## **Oryginalność rozprawy**

Wkład autora w rozwój metod i technik związanych z analizą, rozpoznawaniem i klasyfikacją obrazów w kontekście automatycznego generowania instrukcji stanowiskowych w celu podniesienia efektywności procesów szkoleniowych polega na opracowaniu i implementacji algorytmów (dwuetapowego modelu identyfikacji aktualnie wykonywanej czynności procedury serwisowej) tj.:

- wyznaczania cechy klatek referencyjnych oraz testowych,
- identyfikacji obiektów na klatkach referencyjnych oraz testowych,
- porównywania cechy i obiekty zidentyfikowanych na klatkach referencyjnych i testowych,
- generowania scenariuszy postępowania na podstawie zidentyfikowanego etapu czynności serwisowej,
- weryfikacji poprawności realizowanych czynności serwisowej na podstawie czasu ich trwania.

Ponadto do oryginalnych osiągnięć można zaliczyć zaprojektowanie i implementację systemu informatycznego wspomagającego realizację procedur serwisowych dla przedsiębiorstwa produkcyjnego. Oryginalnym osiągnięciem jest również wykazanie skuteczności zastosowania przedmiotowego systemu informatycznego w celu efektywnej realizacji szkoleń w dziale utrzymania ruchu.

## **Słabe strony rozprawy**

Pomimo osiągnięcia postawionych celów rozprawy, udowodnienia sformułowanych tez oraz przedstawienia odpowiedzi na postawione pytania badawcze w kontekście proponowanego rozwiązania oraz otrzymanych wyników eksperymentów, w recenzowanej rozprawie można zauważyć pewne braki, zarówno natury merytorycznej, wymagające wyjaśnienia i doprecyzowania, jak również mniej znaczące usterki natury językowej i edycyjnej.

Do głównych wad rozprawy, aczkolwiek mających w większości charakter dyskusyjny oraz nie umniejszających wartości pracy w znaczący sposób, zaliczyć należy kwestie wymienione poniżej (stanowią one zagadnienia polemiczne do wyjaśnienia lub dyskusji podczas obrony rozprawy doktorskiej):

- Rozważania zawarte w rozdziałach 2 i 3 zapewne istotne i potrzebne z punktu widzenia wyjaśnienia metod i technik wykorzystywanych w pracy, stanowią jednakże znaczącą część rozprawy tj. liczą około 40 stron. Bardziej syntetyczne ujęcie tych zagadnień wraz z znaczącym uzupełnieniem części badawczej z pewnością wzbogaciłoby merytorycznie i pozwoliło wnikliwiej ocenić recenzowaną rozprawę doktorską.
- Zawartość rozdziału 4 opisuje metodę automatycznego generowania instrukcji stanowiskowych w oparciu o sieci konwolucyjne wykorzystywane do zadań wykrywania, segmentacji oraz rozpoznawania obiektów oraz regionów na obrazach. Pewien niedosyt budzi jednak brak jakichkolwiek rozważań dotyczących wpływu zniekształceń obrazu tj. szumu czy rodzaju oświetlenia sceny na jakość realizowanych zadań detekcji.
- Nie przedstawiono wyczerpujących rozważań dotyczących wpływu architektury sieci neuronowej (tzn. złożoności, ilości warstw, rodzaje funkcji aktywacji czy też wartości początkowych wag) na osiągnięte wyniki detekcji. Samo określenie skuteczności zawarte w rozdziale 5 nie jest wystarczające do obiektywnej weryfikacji przyjętej metody badawczej w kontekście rozważanych typów sieci neuronowych.
- Brak jest jednoznacznego i precyzyjnego określenia parametru dotyczącego efektywności systemu informatycznego (tj. realizacji szkoleń przy zastosowaniu systemu generowania instrukcji stanowiskowych). Sformułowanie zawarte w podrozdziale 6.4, że „analiza efektywności systemu informatycznego polegała na ocenie skuteczności wykrywania czynności, etapów czynności serwisowych oraz generowanego scenariusza postępowania (wraz z instrukcjami graficznymi) na danym stanowisku roboczym” nie definiuje w sposób ścisły przedmiotowego parametru. Istotnym mankamentem jest również fakt braku definicji pojęcia skuteczności w odniesieniu do analizowanego rozwiązania.

Powyższe uwagi wynikają przede wszystkim (prawdopodobnie) z braku prezentacji wyników eksperymentów, które były zapewne przeprowadzone podczas opracowywania poszczególnych algorytmów. Ich prezentacja pozwoliłaby potwierdzić pełną efektywność przyjętych rozwiązań, a być może pozwoliłaby na dalsze udoskonalenie, zwłaszcza metod wykorzystywanych do analizy obrazów i klasyfikacji obiektów.

Recenzowana rozprawa napisana jest poprawnie pod względem językowym, stylistycznym oraz redakcyjnym. Niemniej jednak w pracy można znaleźć błędy językowe oraz redakcyjne, takie jak:

- str. 2 niepoprawna struktura w spisie treści tj. podrozdział 2.4.1 jest jedynym elementem podrozdziału 2.4,
- str. 8 wykazuje niejednolite stosowanie myślników tj. zamienne stosowanie łącznika i półpauzy w wykazie skrótów,
- str. 25 błąd interpunkcyjny w tekście „(regulacja wartości progu aktywacji.)”,
- str. 26, 37, 38, 53, 54, 56, 57-59 błędy interpunkcyjne (brak przecinka lub kropki na końcu linii wzoru matematycznego),
- str. 39 błąd interpunkcyjny, przecinek zamiast kropki w tekście „klasy obiektu”,
- str. 43 użycie dwukropka (:) jako separatora dziesiętnego zamiast przecinka (.),
- str. 43, 44, 46, 47, 65, 67, 69-71, 73-74 niewystarczająca czytelność rysunku 13-15, 18-19, 28, 29, 30-32, 33-34,
- str. 52 powinno być z dużej litery w tekście „gdzie:”,
- str. 54 kropka na początku zdania „Dla ....”,
- str. 107-115 brak jednolitego standardu zapisu pozycji literaturowych oraz liczne błędy interpunkcyjne,

- str.116-117 błędy edycyjne związane z formatowaniem tekstu.

Należy podkreślić, że przytoczone powyżej błędy językowe oraz redakcyjne nie pomniejszają wartości naukowej oraz oryginalności rozprawy.

### **Ocena wiedzy Autora oraz znajomości współczesnej literatury z zakresy dyscypliny**

Rozdziały nr 2 oraz 3 opisują kolejno stan wiedzy dotyczący uczenia głębokiego (które polega na wytrenowaniu wielowarstwowych sztucznych sieciach neuronowych w celu tworzenia złożonych koncepcji na podstawie prostszych reprezentacji) w analizie obrazu oraz problematyki klasyfikacji polegającej na podejmowaniu decyzji o przynależności do danej klasy, w wyniku obserwacji aktualnego wektora cech i wiedzy nabytej (w procesie uczenia) na podstawie określonego zbioru uczącego. Przedstawione w wymienionych rozdziałach treści wskazują, że Autor rozprawy posiada wiedzę teoretyczną która dotyczy omawianej w pracy problematyki i mieści się w aktualnym nurcie badań związanych z przetwarzaniem obrazów, w szczególności z zagadnieniami wykorzystania uczenia maszynowego. Treści tych rozdziałów odnoszą się do aktualnego stanu wiedzy i odpowiednio wprowadzają czytelnika do poszczególnych zagadnień, które w dysertacji stanowią oryginalny wkład Autora. Spis literatury liczy 139 pozycji. Cytowane prace dobrane są prawidłowo i odnoszą się do omawianych w pracy problemów oraz świadczą o umiejętności korzystania z istniejącej literatury.

### **Ocena umiejętności poprawnego i przekonującego przedstawienia wyników**

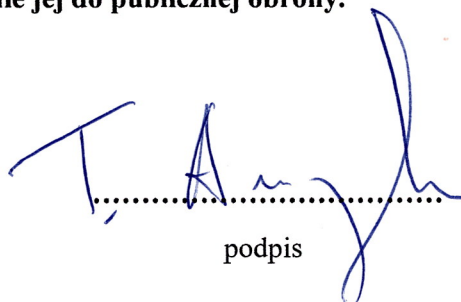
Autor wykazał umiejętności poprawnego i przekonującego przedstawienia wyników działania opracowanych autorskich algorytmów oraz ich eksperymentalnej weryfikacji z wykorzystaniem referencyjnych zbiorów danych. Styl i forma prezentacji jest poprawna i zrozumiała. Dodatkowo wyniki badań eksperymentalnych zostały szeroko skomentowane w treści rozprawy.

### **Podsumowanie**

Przedstawiona do recenzji rozprawa doktorska mgr inż. Daniela Halikowskiego pt. „*Przetwarzanie sekwencji wideo przy zastosowaniu podejścia deep learning do automatycznego generowania instrukcji stanowiskowych*”, której promotorem jest dr hab. inż. Justyna Patalas Maliszewska, prof. UZ, spełnia wszystkie wymagania stawiane rozprawom doktorskim przez *Ustawę o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki* z dnia 14 marca 2003 roku (Dz. U. nr 65 poz. 595, z późn. zm.).

Recenzowana rozprawa doktorska stanowi samodzielne rozwiązanie problemu badawczego mieszczącego się w zakresie dyscypliny naukowej *informatyka techniczna i telekomunikacja* w dziedzinie nauk inżynieryjno-technicznych, zgodnie z aktualną klasyfikacją (obejmując m. in. zagadnienia wcześniejszej dyscypliny *informatyka* w dziedzinie nauk technicznych) i potwierdza umiejętność rozwiązywania problemów naukowych zgodnych z jej zakresem przez Autora rozprawy.

**Wnioskuje o przyjęcie rozprawy i dopuszczenie jej do publicznej obrony.**



podpis