

3 kwietnia 2025 r.

prof. dr hab. inż. Rafał Scherer  
Katedra Sztucznej Inteligencji  
Wydział Informatyki i Sztucznej Inteligencji  
Politechnika Częstochowska  
al. Armii Krajowej 36  
42-200 Częstochowa

### **Recenzja**

rozprawy doktorskiej mgra Adama Surówki, pt.: Synteza zaawansowanych technik wizyjnych w inteligentnym systemie nadzoru obiektów użytku publicznego.

Niniejszą recenzję opracowano zgodnie z uchwałą nr 2/2025 Rady Naukowej dyscypliny Automatyka, elektronika, elektrotechnika i technologie kosmiczne PK. Promotorem jest prof. dr hab. inż. Janusz Gołdasz, a promotorem pomocniczym dr inż. Grzegorz Litawa.

## **1. Charakterystyka tematu, celu i tezy badawczej rozprawy**

Nowoczesne metody analizy wideo i scen oferują zaawansowane narzędzia do przetwarzania, interpretacji i wykorzystania danych wizualnych w czasie rzeczywistym lub po nagraniu. Dzięki technologiom takim jak uczenie maszynowe przetwarzanie obrazów i analiza behawioralna, możliwe jest automatyczne wykrywanie obiektów, śledzenie ruchu, rozpoznawanie twarzy, identyfikacja emocji, a nawet przewidywanie zachowań. Systemy te wykorzystują również głębokie uczenie do klasyfikacji scen, wykrywania anomalii czy analizy tłumu. Metody te są niezbędne w wielu dziedzinach, takich jak bezpieczeństwo publiczne (monitorowanie zagrożeń, identyfikacja przestępców), transport (analiza ruchu drogowego, autonomiczne pojazdy), medycyna (diagnostyka obrazowa, monitorowanie pacjentów), marketing, sport oraz rozrywka. Pozwalają one na szybsze i dokładniejsze podejmowanie decyzji, redukcję kosztów oraz poprawę efektywności w różnych sektorach gospodarki. W erze cyfrowej transformacji analiza wideo i scen staje się kluczowym narzędziem do przetwarzania ogromnych ilości danych wizualnych i wyciągania z nich wartościowych wniosków. Praca dotyczy Monitoringu wizyjnego, a mianowicie stworzenia metod które wykrywają automatycznie zdarzenia najważniejsze z punktu widzenia bezpieczeństwa.

## **2. Zawartość rozprawy**

Recenzowana praca mgra Adama Surówki składa się z siedmiu rozdziałów, wykazu oznaczeń i skrótów, zmiennych, spisu rysunków, tabel oraz bibliografii. Dokument liczy 96 stron.

Pierwszy rozdział jest krótkim wprowadzeniem do tematyki, zawiera dwie tezy pracy oraz krótkie omówienie zawartości.

Rozdział drugi omawia techniki analizy sygnału wideo, poczynając od podstaw rejestracji obrazu, poprzez biblioteki programistyczne, takie jak OpenCV, Scikit-image, MATLAB IPT. Omówiono rodzaje kamer i ich kalibrację pod kątem różnych zniekształceń. Przedstawiono również ujęcie historyczne sposobów detekcji i klasyfikacji obiektów i wymieniono kilka zbiorów danych. Dalej omówiono techniki rozpoznawania tekstów ze szczególnym naciskiem na detekcję i odczyt tablic rejestracyjnych pojazdów, porównano otwarte systemy softwarowe rozpoznawania znaków i wybrano Tesseract OCR jako stosowany w badaniach. Omówiono metody estymacji sylwetki człowieka oraz najpopularniejsze biblioteki implementujące te metody.

Rozdział 3 omawia inteligentne systemy nadzoru wizyjnego i ich obszary zastosowań. Dokonano klasyfikacji takich systemów, omówiono ich architektury, dokonano również przeglądu metod używanych w takich systemach. Ogromną liczbę takich systemów zebrano w tabeli 3.1 i porównano je pod względem różnych kryteriów. Omówiono algorytm analizy przestrzennej sceny, algorytm śledzenia oraz różne zbiory danych. Ciekawy jest również wykaz detekcji zdarzeń alarmowych wykrywanych przez różne systemy nadzoru. Podano również luki badawcze tak jak brak uniwersalności skalowalności w systemach, ograniczenia w kalibracji i estymacji przestrzeni sceny, nie do końca dobrze opracowane śledzenie obiektów, brak rozpoznawania pewnych zdarzeń alarmowych, brak publicznie dostępnych zbiorów danych do uczenia i testowania takich systemów oraz brak metod, które adaptują się do zmieniających się warunków środowiska.

Rozdział 4 omawia projektowanie systemu nadzoru wideo z wykorzystaniem modeli neuronowych. Oraz stworzoną zbiór nagrań wideo podzielony na aktywności, które mogą być wykrywane przez system nadzoru. Część badań wykonano również na modelach trójwymiarowych wykonanych w systemie blender oraz tablic rejestracyjnych generowanych przez oprogramowanie PlatesMania. System nadzoru składa się z wielu elementów takich jak metody korekcji kamery, kalibracji kamery, metody korekty jasności obrazów, itd.

Rozdział 5 omawia wstępne przetwarzanie obrazu. Na początku opisano szczegółowo kalibrację obrazu kamery za pomocą specjalnych znaczników. Pozwala to na korektę obrazu oraz pomiar rzeczywistych rozmiarów obiektów. Następnie opisano autorską metodę korekcji jasności służącą do korekty miejsc o niskiej ekspozycji (ciemnych). Metoda pracuje w oparciu o histogram kanału L przestrzeni LAB tworząc nową tablicę LUT służącą do wyliczenia zmiany jasności. Zwiększenie jasności ciemnych obszarów poprawiło działanie właściwych metod wykrywających obiekty. Następnie opisano użycie detektorów obiektów YOLO i Tesseract OCR do wykrywania ludzi, pojazdów, zwierząt i różnego rodzaju zdarzeń oraz metody śledzenia obiektów za pomocą śledzenia regionów bounding box za pomocą filtra Kalmana. Rozdział kończy omówienie sposobu kalibracji kamer.

Rozdział 6 omawia metody detekcji zdarzeń alarmowych poczynając od identyfikacji pojazdów. Wystąpienia obiektów pojazd i tablica są identyfikowane przez sieć YOLO 4. Wykryte obszary obrazu są podawane różnym operacją i następnie rozpoznawany jest tekst na

tablicy rejestracyjnej a znaki dopasowany są do znaków ze słownika możliwych znaków występujących na tablicach rejestracyjnych rozpoznany tekst porównywany jest z tekstem identyfikatorów pojazdów w bazie. Wstępne testy to zostały wykonane z użyciem modeli Blender z wygenerowanymi syntetycznymi tablicami rejestracyjnymi. Następnym systemem był system detekcji zwierząt również oparty, tym razem bezpośrednio, na sieci YOLO 4. Bardziej rozbudowaną metodą było wykrywanie sabotażu kamery, czyli zasłonięcia, obrotu czy oślepienia kamery. Był to dwuetapowy algorytm opracowany przez Doktoranta. Metoda uzyskała bardzo wysokie wyniki dokładności detekcji. Następnymi opracowanymi metodami były wykrywanie nielegalnego parkowania czyli wykrywania przekroczenia czasu postoju w zadanej strefie oraz naruszenie wirtualnej granicy. Następnym algorytmem było wykrywanie przeludnienia, czyli przekroczenia zadanej liczby osób na monitorowanym obszarze. W ramach pracy opracowano również metodę detekcji aktywności, włóczęgostwa oraz naruszenia wirtualnej strefy, pożaru, upadku, kradzieży czy przemocy. Działanie stworzonych metod opisanych w rozdziale 6 zestawione jest w tabeli 6.18. Większość z nich uzyskało bardzo wysoką jakość działania. Porównanie z konkurencyjnymi systemami również pokazało bardzo dobrą jakość działania. System jest bardzo wszechstronny, pozwala na analizę całej sceny oraz wykrywanie statycznych i dynamicznych zdarzeń. Dzięki użyciu sieci neuronowych możliwe będzie tworzenie nowych funkcji systemu przez douczanie sieci lub wykorzystanie innych klas obiektów.

Rozdział 7 podsumowuje rozprawę oraz proponuje dalsze możliwości rozbudowy zaproponowanych metod. Dalej następuje zestawienie dorobku naukowego Doktoranta, spis ilustracji i tabel oraz bibliografia składająca się ze 143 pozycji.

### 3. Ocena rozprawy

W ramach rozprawy doktorskiej Pan mgr inż. Adam Surówka zaproponował oryginalne rozwiązanie systemu analizy obrazu. Tematyka pracy jest bardzo aktualna i potrzebna, oryginalny dorobek autora polega na stworzeniu:

- modułowego systemu wykrywania zdarzeń w oparciu o wiele cech obrazu wideo,
- zbioru danych ujęć wideo,
- metody automatycznej korekty jasności obrazu, uniwersalnego algorytmu kalibracji sceny,
- rozlicznych metod wykrywania zdarzeń obejmujących większość potrzeb systemów nadzoru wideo ogólnego przeznaczenia.

Rozprawa doktorska uwidacznia wysoką ogólną wiedzę teoretyczną i praktyczną oraz umiejętność prowadzenia pracy naukowej mgr inż. Adama Surówki. Opracował wprowadzenie do tematyki i przegląd literatury na temat metod wizji komputerowej związanych z tematyką pracy. Rozprawa doktorska stanowi oryginalne rozwiązanie problemu naukowego. Zaproponowane metody mają duże znaczenie dla nauk technicznych oraz

przemysłu, zarówno teoretyczne, jak i aplikacyjne. System może być łatwo rozbudowany o nowe metody wykrywające specyficzne zdarzenia zależnie od potrzeb danego środowiska.

Niezależnie od mojej dobrej oceny pracy, nasunęło mi się pytanie, czemu zdecydowano się na użycie akurat modelu YOLO 4, gdy istnieje wiele innych, np. następne wersje YOLO, Faster R-CNN, MobileNet, EfficientNet, EfficientViT, model opisany w pracy SlowFast Networks for Video Recognition, 2019 i wiele innych.

#### 4. Wnioski końcowe recenzji

Podsumowując recenzję stwierdzam, że rozprawa doktorska „Synteza zaawansowanych technik wizyjnych w inteligentnym systemie nadzoru obiektów użytku publicznego” prezentuje oryginalne rezultaty stanowiące rozwiązanie problemu naukowego oraz wkład w rozwój dyscypliny automatyka, elektronika, elektrotechnika i technologie kosmiczne. Pan Adam Surówka wykazał się umiejętnością samodzielnej pracy badawczej, znajomością literatury światowej i wiedzą w zakresie wizji komputerowej. Rozprawa doktorska prezentuje ogólną wiedzę teoretyczną kandydata. Recenzowana praca spełnia wymagania Ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce (Dz.U. z 2022 r. poz. 574 z późn. zm.). w dyscyplinie naukowej Automatyka, elektronika, elektrotechnika i technologie kosmiczne. Wnoszę o jej przyjęcie i dopuszczenie do dalszych etapów postępowania doktorskiego.

