

Szczecin, 13.12.2022

dr hab. inż. Jacek Piskorowski, prof. ZUT
Katedra Inżynierii Systemów, Sygnałów i Elektroniki
Wydział Elektryczny
Zachodniopomorski Uniwersytet Technologiczny w Szczecinie

POLITECHNIKA ŚLĄSKA
Biuro Rady Dyscypliny
Automatyka, Elektronika, Elektrotechnika
i Technologie Kosmiczne
wzłynęto dnia 29.12.2022
nr 37 zat.

RECENZJA

Rozprawy doktorskiej mgr inż. Pauliny Bugiel pt.: "Automatyczna analiza dużych zbiorów nagrań wideo w celu wskazania w nich obrazów wybranych obiektów i określenia warunków pogodowych".

Podstawą formalną opracowania recenzji jest Uchwała nr 77/2022 Rady Dyscypliny Automatyka, Elektronika i Elektrotechnika Politechniki Śląskiej z dnia 25 października 2022 roku, a także pismo o sygnaturach RDAEE/87/2022, RDAEE.512.11.2022 Przewodniczącej Rady Dyscypliny Automatyka, Elektronika i Elektrotechnika Politechniki Śląskiej dr hab. inż. Moniki Kwoki, prof. PŚ z dnia 9 listopada 2022 roku. Promotorem opiniowanej rozprawy doktorskiej jest prof. dr hab. inż. Jacek Izydorczyk, a promotorem pomocniczym dr inż. Marcin Szelest.

1. Znaczenie podjętej tematyki

Opiniowana rozprawa doktorska dotyczy aspektów aktywnego bezpieczeństwa w branży motoryzacyjnej. Przeprowadzone przez Doktorantkę badania dotyczą w szczególności anotacji oraz segregacji nagrań wideo z samochodowych jazd testowych. Obecnie obserwuje się dynamiczny rozwój systemów bezpieczeństwa w samochodach, których zadaniem jest wsparcie kierowców w wymagających warunkach drogowych, a także podniesienie poziomu ochrony zdrowia i życia pasażerów oraz pieszych. Rozwój aktywnych systemów bezpieczeństwa to ogólnoswiatowy trend wymuszony w dużym stopniu przez Parlament

Europejski, który wytyczył kierunek rozwoju motoryzacji stawiający na bezpieczeństwo. Od 2020 roku na terenie Unii Europejskiej obowiązują przepisy mające na celu wyposażenie produkowanych pojazdów w systemy zwiększające ochronę życia ich użytkowników. Zgodnie z wprowadzonymi przepisami wszystkie produkowane i sprowadzane od 2022 roku pojazdy muszą być wyposażone m.in. w takie funkcje, jak inteligentny asystent prędkości, wykrywanie pieszych i rowerzystów, automatyczne hamowanie awaryjne, awaryjne utrzymanie auta na pasie ruchu, a także monitorowanie koncentracji kierowcy. Wprowadzenie nowych przepisów ma przede wszystkim na celu zmniejszenie liczby ofiar wypadków drogowych. W branży motoryzacyjnej wykorzystywane są obecnie zaawansowane systemy, które informują kierowców z wyprzedzeniem o sytuacji na drodze oraz wspomagają w trakcie manewrów. Systemy te wyposażone są w rozmaite sensory, takie jak kamery, LiDARy, a także radary, które monitorują otoczenie pojazdów w trakcie jazdy oraz czytają znaki drogowe i ograniczenia prędkości.

Podjęta przez Doktorantkę tematyka wpisuje się wprost w zakres dyscypliny Automatyka, Elektronika i Elektrotechnika. Należy również podkreślić, że tematyka rozprawy doktorskiej jest niezwykle ważna i aktualna z zarówno z teoretycznego, jak i praktycznego punktu widzenia. Opracowane metody w zakresie analizy i augmentacji danych mogą znaleźć zastosowanie w systemach wspomagania kierowcy.

2. Ogólna charakterystyka, zakres i cel rozprawy

Rozprawa doktorska dotyczy automatycznej analizy dużych zbiorów nagrań wideo z samochodowych jazd testowych. Przeprowadzone badania oraz analizy miały na celu usprawnienie wizyjnych systemów wspomagania kierowcy. Zasadniczym celem pracy jest opracowanie systemu usprawniającego anotację oraz segregację nagrań wideo z samochodowych jazd testowych. Do realizacji postawionego celu projektu wykorzystano metody przetwarzania obrazów oraz metody głębokiego uczenia. W opisie celu rozprawy Doktorantka zasygnalizowała, że obecnie wykorzystywane wizyjne systemy wspomagania, bazujące na sieciach neuronowych uczonych w sposób nadzorowany, wymagają dużego zbioru anotowanych nagrań wideo z jazdy samochodem. W efekcie konieczne jest usprawnienie anotacji, która jest bardzo czasochłonnym procesem. Istotnym celem rozprawy doktorskiej było również określenie minimalnego, ale dostatecznie dużego zbioru danych,

niezbędnego do walidacji i uczenia detektora poprzez automatyczne dodanie do nagrań metadanych. Tak postawione cele pracy doprowadziły do sformułowania następujących tez rozprawy doktorskiej:

- Zestaw obrazów służących do uczenia sieci konwolucyjnej może zostać uzupełniony o nowe dane bez konieczności ich anotacji poprzez przetwarzanie już oznaczonych obrazów, a aktualnie istniejące do tego techniki nie wyczerpują wszystkich możliwości i jest możliwe znalezienie nowych skutecznych metod augmentacji obrazów.
- Degradacje obrazów z sensora optycznego, jakie mogą być spotykane w systemach wspomagania kierowcy, mają wpływ na jakość detekcji i zależnie od rodzaju zniekształcenia, w różnym stopniu pogarszają rozpoznawalność obiektów.

Cele oraz tezy pracy zdefiniowano w precyzyjny sposób. Można stwierdzić, że postawione cele pracy są istotne i aktualne na tle obecnego stanu wiedzy w zakresie wizyjnych systemów wspomagania pojazdów. Cele oraz tezy pracy odpowiadają zakresowi i tematyce rozprawy oraz determinują zakres przeprowadzonych badań. Uważam, że podjęcie tematu rozprawy doktorskiej było celowe zarówno ze względów poznawczych, teoretycznych oraz praktycznych. Tezy pracy zostały prawidłowo postawione, natomiast cele główne rozprawy doktorskiej zostały zrealizowane.

3. Struktura rozprawy doktorskiej

Rozprawa doktorska mgr inż. Pauliny Bugiel pt.: "Automatyczna analiza dużych zbiorów nagrań wideo w celu wskazania w nich obrazów wybranych obiektów i określenia warunków pogodowych" jest podzielona na 4 zasadnicze rozdziały, które zostały uzupełnione wykazem 86 pozycji literaturowych.

W rozdziale pierwszym, będącym jednocześnie wprowadzeniem do rozprawy doktorskiej, sformułowano obszar tematyczny pracy, a także przedstawiono motywację podjęcia tematyki rozprawy. Ponadto, zdefiniowano cele pracy oraz tezy pracy, a także przedstawiono charakterystykę poszczególnych rozdziałów rozprawy doktorskiej.

Rozdział drugi poświęcony jest omówieniu podstawowych pojęć dotyczących sieci neuronowych. Do budowy narzędzia służącego do automatycznego wykrywania obiektów na filmach lub obrazach, które mogłyby dawać podpowiedzi dla użytkownika, Doktorantka zdecydowała się wykorzystać mechanizmy uczenia maszynowego. W ramach rozdziału wprowadzono pojęcie perceptronu oraz sieci neuronowej. Scharakteryzowano jednokierunkową sieć neuronową, funkcje kosztu oraz metody ich minimalizacji, a także funkcje aktywacji. Omówiono uczenie nienadzorowane, uczenie nadzorowane oraz uczenie przez wzmacnianie. Przedstawiono ponadto problematykę doboru danych uczących. Odniesiono się w tym zakresie przede wszystkim do wielkości zbioru uczącego, dokładności anotacji, a także podziału danych. Scharakteryzowano parametry odnoszące się do sterowania procesem uczenia sieci. W dalszej kolejności omówiono augmentację danych, a w szczególności w odniesieniu do klasyfikacji oraz detekcji obrazów. Wprowadzono pojęcie sieci konwolucyjnej oraz omówiono zasady jej działania. Następnie omówiono problematykę klasyfikacji i detekcji obiektów na obrazach. Odniesiono się m.in. do sieci R-CNN, sieci Fast R-CNN, sieci Faster R-CNN, a także sieci YOLO. W dalszej kolejności scharakteryzowano metodyki i narzędzia stosowane do anotacji obrazów (w szczególności w branży motoryzacyjnej). W ramach rozdziału omówiono również współzawodniczące sieci generatywne GAN. Przedstawiono zasadę działania tego typu sieci, omówiono architekturę dyskryminatora i generatora, opisano mechanizmy trenowania sieci GAN, a także scharakteryzowano popularne architektury sieci GAN. Następnie omówiono narzędzia i miary wykorzystywane do ewaluacji sieci.

Rozdział trzeci jest zasadniczą częścią rozprawy doktorskiej, w której zaprezentowane zostały wyniki przeprowadzonych badań i analiz. Prace badawcze dotyczące augmentacji danych w kontekście detekcji obrazów były prowadzone w oparciu o bazę danych udostępnioną przez pracodawcę Doktorantki, a także z wykorzystaniem ogólnodostępnych danych. W rozprawie scharakteryzowano zbiory danych, które zostały wykorzystane do procesu uczenia sieci neuronowej. Zbiory danych zostały opisane zarówno w jakościowy, jak i ilościowy sposób. Scharakteryzowano zbiór znaków drogowych, zbiór filmów, zbiór Berkeley DeepDrive, zbiór Cityscapes, a także dane z symulatora jazdy Carla. Następnie przedstawiono i omówiono metodykę badań. Odniesiono się m.in. do procesu przygotowania danych, przygotowania środowiska programistycznego, procesu trenowania sieci, wyznaczenia metryk oceny wytrenowania sieci, a także analizy uzyskanych wyników. Omówiono ponadto stanowisko, które zostało wykorzystane do badań, a także platformę

programistyczną Darknet, która została użyta do trenowania sieci neuronowych. W środowisku Darknet została zaimplementowana architektura sieci YOLO. W dalszej kolejności przedstawiono wyniki badań dotyczących augmentacji obrazów. Scharakteryzowano wykorzystane podczas badań metody augmentacji. Wykorzystane metody zostały zaimplementowane w postaci autorskiego kodu wykorzystującego biblioteki OpenCV. Opracowano również skrypty do transformacji anotacji. Wykorzystano m.in. operacje w postaci rotacji, ścinania, korekcji barwy, a także wstawiania próbek. Operacja wstawiania próbek to autorska metoda Doktorantki, która polega na umieszczaniu obiektów na obrazach tła (obrazach na których nie było żadnych interesujących obiektów). Następnie przedstawiono i omówiono wyniki badań dotyczących augmentacji obrazów. W trakcie badań stworzono 17 zbiorów treningowych i wytrenowano 34 sieci. W kolejnej części rozdziału przedstawiono i omówiono wyniki badań dotyczących augmentacji obrazów z wykorzystaniem symulatora jazdy. Dane z symulatora stanowiły źródło dodatkowych danych do trenowania detektora. Doktorantka zauważyła, że dodanie do danych prezentujących rzeczywiste obiekty próbek z symulatora nie poprawia, a nawet pogarsza detekcję na danych realnych. W dalszej kolejności zaprezentowano wyniki badań dotyczących augmentacji z wykorzystaniem sieci generatywnych GAN. Doktorantka zaobserwowała, że poprawę wyników przy zastosowaniu danych augmentowanych można osiągnąć jedynie dla sieci OASIS. Z kolei sieć CycleGAN dla części klas daje poprawę wyników, a dla innych klas skutkuje pogorszeniem. W ogólności potwierdzono, że sieci GAN nadają się do wykorzystania w celu augmentacji danych. W dalszej części rozdziału omówiono wpływ degradacji obrazu na jakość detekcji. Przeanalizowano m.in. rozmycie gaussowskie, wpływ cząsteczek obecnych na soczewce kamery, degradację głębi bitowej skali szarości, cykliczne oślepienie, a także okluzję fragmentu obrazu. Omówiono również wykorzystany algorytm klastrowania. W ramach badań przeprowadzono eksperymenty odzwierciedlające wpływ sztucznie wprowadzonych zaburzeń obrazu na jakość działania modułów rozpoznających inne samochody na drodze podczas symulacji jazdy testowej autonomicznego pojazdu. Uzyskane wyniki wyznaczyły granice odporności algorytmów wizyjnych pracujących w różnych warunkach podczas jednostajnie zwiększającego się poziomu degradacji danych wejściowych.

W rozdziale czwartym przedstawiono podsumowanie wyników badań. Odniesiono się do badań w zakresie augmentacji danych, a także degradacji obrazów. Doktorantka podkreśliła, że co do zasady augmentacja danych poprawia jakość detekcji. Niemniej jednak nie wszystkie metody są odpowiednie dla danego zestawu danych. Analizując wyniki badań

dotyczących wpływu degradacji obrazów na jakość detekcji Doktorantka podkreśliła, że niektóre degradacje powodują różny spadek jakości detekcji, zależnie od tego do jakich danych zostały zastosowane.

Bibliografia zawiera 86 pozycji. W spisie literatury wyodrębniono również 5 prac współautorstwa Doktorantki. Literaturę dobrano w staranny sposób. Odniesiono się do reprezentatywnych prac z systemów wizyjnych, systemów aktywnego bezpieczeństwa w branży motoryzacyjnej, metod uczenia maszynowego, augmentacji danych oraz degradacji obrazów.

4. Ogólna ocena rozprawy

Autorka rozprawy doktorskiej zrealizowała postawione cele w sposób adekwatny, używając do tego właściwej metodyki badań. Przyjęte założenia są uzasadnione. W sposób przejrzysty odniesiono się do źródeł. Do najistotniejszych osiągnięć Doktorantki należy zaliczyć przede wszystkim:

- weryfikację dostępnych metod augmentacji danych w celu poprawy sprawności automatycznego detektora obiektów na obrazach;
- opracowanie autorskiej metody wstawiania próbek ;
- weryfikację wykorzystania sieci generatywnych do augmentacji obrazów;
- weryfikację wpływu degradacji obrazów na detekcję obiektów w systemach wspomagania kierowcy.

Badania przedstawione w rozprawie doktorskiej dotyczyły głównie analizy wpływu danych uczących na sprawność splotowych sieci neuronowych wykorzystywanych do rozpoznawania obiektów. Wykazano, że augmentacja danych poprawia jakość detekcji. Zaproponowana przez Doktorantkę metoda wstawiania próbek umożliwia poprawę jakości detekcji w porównaniu do powszechnie wykorzystywanych rozwiązań. Z przeprowadzonych przez Doktorantkę badań wynika, że wpływ degradacji obrazów na detekcję obiektów jest mocno uzależniony od rodzaju zmian w obrazie, ich umiejscowienia, zmienności w czasie, a także od warunków pogodowych oraz otoczenia pojazdu.

5. Uwagi krytyczne

Uwagi merytoryczne:

- W badaniach dotyczących augmentacji obrazów wykorzystano m.in. operację rotacji. W rozprawie doktorskiej zadeklarowano, że obrazy poddano rotacji o losowy kąt z zakresu $[-15, -4]$ u $[4, 15]$ stopni. Czy przeprowadzono analizę wpływu tego kąta na wyniki badań?
- W badaniach dotyczących augmentacji obrazów wykorzystano m.in. operację ścinania. W rozprawie doktorskiej zadeklarowano, że kąt ścinania wybierano losowo z zakresu $[-8,5; 8,5]$ stopnia. Z czego wynika przyjęty przedział kątowy? Czy przeprowadzono analizę wpływu tego kąta na wyniki badań?
- W badaniach dotyczących augmentacji obrazów wykorzystano m.in. korekcję barwy. W rozprawie doktorskiej zadeklarowano, że wartość odcienia zmieniono o losową wartość z zakresu $[-50, 50]$. Z czego wynika przyjęty przedział? Czy przeprowadzono analizę wpływu wartości odcienia na wyniki badań?
- Jaka jest geneza opracowania autorskiej metody wstawiania próbek? Czy były konkretne powody opracowania tego typu metody, czy też powstała ona w wyniku różnych manipulacji zbiorem danych?
- W sekcji 3.3.5 zadeklarowano, że „uczenie odbywało się na tyle długo, aby wartość funkcji strat nie spadała już zbyt gwałtownie i można było przeprowadzić badania porównawcze”. Jak należy rozumieć pojęcie „zbyt gwałtownego spadku”? Czy można to pojęcie wyrazić w mierzalny sposób?
- W przypadku badań nad wpływem degradacji obrazów na jakość detekcji dokonano analizy 5 typów zaburzeń. Dlaczego ograniczono się do 5 typów zaburzeń. Jakie było kryterium wyboru tych zaburzeń?

Uwagi redakcyjne:

- Rozprawa doktorska zawiera liczne usterki językowe w postaci błędów stylistycznych oraz interpunkcyjnych.
- Rozprawa doktorska została podzielona za zbyt wiele sekcji niskiego poziomu. Taka struktura rozprawy utrudnia jej lekturę.

- Treść rozprawy doktorskiej zawiera za dużo ogólnie znanej wiedzy w zakresie sieci neuronowych.

Wskazane uwagi krytyczne i komentarze mają charakter dyskusyjny i nie wpływają na ogólnie pozytywną ocenę wyników zawartych w opiniowanej rozprawie doktorskiej.

6. Wniosek końcowy

Podsumowując stwierdzam, że mgr inż. Paulina Bugiel wykazała się dużą wiedzą z zakresu analizy i przetwarzania obrazów, a także metod sztucznej inteligencji. Rozprawa doktorska mgr inż. Pauliny Bugiel pt.: "Automatyczna analiza dużych zbiorów nagrań wideo w celu wskazania w nich obrazów wybranych obiektów i określenia warunków pogodowych" jest oryginalnym, interesująco przedstawionym, uzasadnionym i twórczym wkładem w dyscyplinę Automatyka, Elektronika i Elektrotechnika. Niniejsza rozprawa doktorska zawiera poprawnie sformułowany i rozwiązany problem badawczy oraz posiada bardzo duży aspekt praktyczny. Stanowi zatem oryginalne rozwiązanie problemu naukowego.

W efekcie stwierdzam, że mgr inż. Paulina Bugiel spełnia wymagania określone w art. 13 ust. 1 Ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz stopniach i tytule w zakresie sztuki (Dz. U. 2003 poz. 595 z późn. zm.) oraz w art. 179 ust. 1 Ustawy z dnia 3 lipca 2018 r. - Przepisy wprowadzające ustawę - Prawo o szkolnictwie wyższym i nauce (Dz. U. 2018 poz. 1669 z późn. zm.). W związku z powyższym wnioskuję o dopuszczenie mgr inż. Pauliny Bugiel do publicznej obrony rozprawy doktorskiej przed Radą Dyscypliny Automatyka, Elektronika i Elektrotechnika Politechniki Śląskiej.

