



POLITECHNIKA ŚLĄSKA
WYDZIAŁ AUTOMATYKI, ELEKTRONIKI I INFORMATYKI

Praca doktorska

Automatyczna analiza dużych zbiorów nagrań wideo w celu wskazania
w nich obrazów wybranych obiektów i określenia warunków
pogodowych

Autor: mgr inż. Paulina Bugiel

Promotor: prof. dr hab. inż. Jacek Izydorczyk

Promotor pomocniczy: dr inż. Marcin Szelest

Gliwice, wrzesień 2022

Streszczenie rozprawy doktorskiej

Automatyczna analiza dużych zbiorów nagrań wideo w celu wskazania w nich obrazów wybranych obiektów i określenia warunków pogodowych

Celem niniejszej pracy doktorskiej było opracowanie systemu usprawniającego anotację oraz segregację nagrań wideo z samochodowych jazd testowych. Do realizacji tego celu zostały wykorzystane metody przetwarzania obrazów i głębokiego uczenia.

Anotowanie nagrań wideo jest czasochłonnym zajęciem wymagającym zaangażowania dużej grupy ludzi, którzy zaznaczają wszystkie obiekty na kolejnych klatkach filmu. Z punktu widzenia firmy zlecającej takie zadanie jest to spory koszt poniesiony za wynagrodzenia, wynajęcie przestrzeni na laboratorium do anotacji oraz zakup potrzebnego sprzętu. Usprawnienie procesu anotacji bezpośrednio przenosi się na redukcję kosztów. Uczenie nadzorowane wymaga bazy danych przygotowanej przez człowieka, tak aby docelowa sieć neuronowa mogła odtwarzać ludzki punkt widzenia. Nie można zatem całkowicie wyeliminować tu udziału ludzi w procesie anotacji, można jednak ten proces usprawnić, co było głównym celem rozpoczęcia niniejszej pracy doktorskiej.

Ułatwienie anotacji nagrań miało się odbywać dzięki systemowi podpowiedzi generowanych przez pomocniczą sieć wykrywającą obiekty na obrazach. Badania skupiały się głównie na sieciach R-CNN oraz ich wariantach, które są obecnie najwydajniejsze i najdokładniejsze do tego typu zadań. Sieci te potrzebują jednak bardzo dużych zbiorów danych uczących, więc spora część badań oparto na tzw. metodach augmentacji danych – przekształceniom już istniejących danych tak, aby powiększyć liczebność zbioru uczącego bez ingerencji użytkownika. Każda metoda augmentacji danych daje potencjalny wzrost liczebności zbioru o 100 procent co może skutkować poprawą jakości detektora. W praktyce jednak możliwy jest spadek jakości, jeśli wykorzystana metoda augmentacji jest nieodpowiednia. Badania pozwoliły pokazać, iż aktualnie stosowane metody augmentacji nie wyczerpują wszystkich potencjalnych algorytmów poprzez wyłonienie nowych metod augmentacji.

Inny, powiązany tor badań, skupiał się na przekształceniach obrazów uczących, które mogą wprowadzić spadek jakości rozpoznań przez sieć neuronową. W badaniach tych przetestowano kilka metod degradacji obrazów oraz ich wpływ na skuteczność detektora. Brano również to w jakim stopniu pogarsza się jakość rozpoznań, zależnie od charakteru samych obrazów, czyli tego w jakich warunkach zostały zarejestrowane. Udało się dzięki temu wykazać, że Degradacje obrazów z sensora optycznego, jakie mogą być spotykane w systemach wspomaganie kierowcy, mają wpływ na jakość detekcji i zależnie od rodzaju zniekształcenia, w różnym stopniu pogarszają rozpoznawalność obiektów.