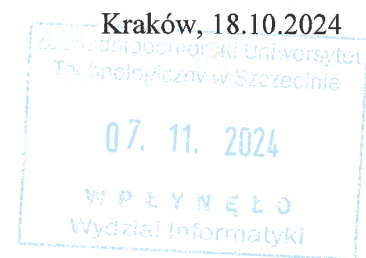


Prof. dr hab. inż. Bogusław Cyganek
Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie
Wydział Informatyki, Elektroniki i Telekomunikacji
Instytut Elektroniki
cyganek@agh.edu.pl



**Recenzja rozprawy doktorskiej
Pana magistra inżyniera Antona Smolińskiego
z tytułowanej:**

***Przetwarzanie i integracja wielomodalnych danych obrazowych
wspomagające wykrywanie zachowań związanych z obniżeniem
poziomu koncentracji użytkownika pojazdu mechanicznego***

Wstęp

Recenzja dotyczy rozprawy doktorskiej Pana magistra inżyniera Antona Smolińskiego pt. „Przetwarzanie i integracja wielomodalnych danych obrazowych wspomagające wykrywanie zachowań związanych z obniżeniem poziomu koncentracji użytkownika pojazdu mechanicznego”, która powstała w roku 2024 w Zachodniopomorskim Uniwersytecie Technologicznym w Szczecinie na Wydziale Informatyki. Promotorem pracy jest Pan dr hab. inż. Paweł Forczmański, prof. ZUT, natomiast promotorem pomocniczym jest Pan dr inż. Adam Nowosielski. Przygotowanie recenzji zostało wykonane na zlecenie Dziekana Wydziału Informatyki ZUT Pana dra hab. inż. Krzysztofa Małeckiego. Praca została napisana w języku polskim. Recenzja sporządzona została w postaci odpowiedzi na pytania dotyczące rozprawy doktorskiej Pana magistra inżyniera Antona Smolińskiego.

1. Problem badawczy i jego znaczenie

Jaki jest najważniejszy problem rozwiązany w rozprawie?

Jednym z istotniejszych problemów cywilizacyjnych jest zapewnienie bezpieczeństwa w wielu aspektach życia, a w szczególności bezpieczeństwa związanego z ruchem drogowym. Niestety bardzo często słyszymy o kolejnych wypadkach drogowych, w których wiele osób odnosi uszczerbek na zdrowiu, a nierzadko też traci życie. W przeważającej liczbie przypadków powodem jest uczestnik lub uczestnicy ruchu drogowego, w tym – w znaczącej liczbie – są to kierowcy pojazdów. Główne czynniki prowadzące do wypadków to nadmierna prędkość, wymuszenie pierwszeństwa, ale również zmęczenie, czy wręcz zaśnięcie kierowcy. Niestety też, ciągle borykamy się z problemem pijanych kierowców, czy też kierowców będących pod wpływem środków psychoaktywnych.

Nasuwa się więc pytanie, jak można zapobiec, czy też ograniczyć ryzyko zaistnienia wypadku, a szczególnie takiego którego powodem jest nieprawidłowy, czy też wręcz zły stan psychofizyczny kierującego pojazdem? Z pomocą przychodzą nowoczesne technologie – odpowiedni zestaw kamer

umieszczony w kabinie pojazdu wraz z jednostką obliczeniową, a przede wszystkim z odpowiednimi algorytmami mogą obserwować stan kierowcy i sygnalizować stan zagrożenia.

Praca naukowa Pana magistra inżyniera Antona Smolińskiego wpisuje się właśnie w nurt badań nad zaawansowanymi systemami wspomagania kierowcy ADAS (ang. *Advanced Driver Assistance Systems*). Doktorant postawił sobie ambitny plan opracowania środowiska laboratoryjnego służącego do zgromadzenia odpowiedniej ilości danych wizyjnych pochodzących z różnych tzw. modalności, czyli rodzajów pomiarów fizycznych, takich jak obrazy spektrum widzialnego RGB, obrazy termowizyjne oraz dane pomiaru głębi sceny, jak również – a nawet przede wszystkim – algorytmów do przetwarzania tego typu danych w celu komputerowej oceny stanu kierowcy w celu wykrywania zachowań związanych z obniżeniem poziomu koncentracji kierowców.

Czy praca ma on charakter naukowy?

Praca Pana magistra inżyniera Antona Smolińskiego ma charakter zarówno naukowy, jak i posiada aspekty prac technologicznych. Mimo licznych firm samochodowych, jak i zespołów naukowych oraz prac badawczych na świecie ciągle nie istnieją systemy, które w pełni spełniałyby wymagane założenia. Dlatego istnieje potrzeba kontynuacji badań w tym zakresie, a dostarczenie nowych rozwiązań wymaga zgromadzenia danych, przeprowadzenia licznych pomiarów, a następnie opracowania oryginalnych algorytmów i wykonania z ich użyciem eksperymentów. Te ostatnie, aby mogły odpowiedzieć na pytania naukowe, wymagają od badaczy pokonania bardzo wielu przeszkód natury technologicznej. Zarówno zagadnienia naukowe, jak i wyzwania technologiczne, a co ważniejsze, sposoby ich rozwiązania zaprezentowane są w rozprawie doktorskiej Pana mgr inż. Antona Smolińskiego.

Czy ma on znaczenie praktyczne?

Podjęty przez Pana mgr inż. Antona Smolińskiego problem badawczy ma bardzo istotne znaczenie zarówno w aspekcie teoretycznym, jak i praktycznym. Zaproponowane rozwiązania są dość uniwersalne i wieloaspektowe. Mogą więc prowadzić nie tylko do aplikacji związanych z systemami ADAS, ale również do rozwiązań ogólniejszych problemów związanych z obserwacją osób w różnych warunkach oświetlenia celem oceny ich stanu, czy też zachowania, ale również mogą prowadzić do licznych aplikacji w innych pojazdach i systemach komunikacyjnych.

2. Wkład autora

Jaki jest najważniejszy wkład autora opisywane w rozprawie?

Główne osiągnięcia Doktoranta Pana mgr inż. Antona Smolińskiego są następujące.

1. *Współpraca przy opracowaniu stanowiska do badania zachowań kierowców oraz akwizycji danych wielomodalnych.*

Współczesne metody przetwarzania obrazów bazują na danych, znacznie bardziej niż starsze algorytmy, które w znacznym stopniu opierały się na modelach i wiedzy eksperckiej. Stąd też zgromadzenie, a następnie poetykietowanie odpowiednio dużych zbiorów danych jest kluczowym, ale też i bardzo wymagającym, zadaniem naukowym. Zdobywanie odpowiednich danych obrazujących różnorodne warunki prowadzenia pojazdów mechanicznych w warunkach

rzeczywistego ruchu drogowego jest niezmiernie trudne. Stąd też doktorant Pan mgr inż. Anton Smoliński wraz z zespołem naukowym pod kierownictwem Promotora, Pana dra hab. Pawła Forczmańskiego, prof. ZUT, zdecydował się na zbudowanie oryginalnego stanowiska laboratoryjnego, w którym w kontrolowanych warunkach możliwa jest akwizycja sygnałów pochodzących z różnych czujników, takich jak: kamery spektrum widzialnego RGB, kamera termowizyjna oraz kamera głębi sceny. Niemniej istotne było zorganizowanie całości eksperymentów dzięki zaproszeniu 38 uczestników oraz przygotowanie szczegółowych scenariuszy zachowań. Dzięki temu możliwe było przeprowadzenie wielogodzinnych eksperymentów, podczas których osoby uczestniczące odgrywały prowadzenie pojazdu wraz z różnymi stanami zmęczenia, typu mrużenie i pocieranie oczu, ziewanie, czy też opadanie głowy. Zgromadzona została pokaźna baza danych wideo, która następnie – dzięki bardzo pracochłonnemu procesowi obróbki oraz opisu – została przekształcona w naukowo cenną bazę wielomodalnych sekwencji prowadzenia pojazdu.

2. *Opracowanie nowatorskich algorytmów rozpoznawania zachowań kierowców w sekwencjach multimodalnych z wykorzystaniem głębokich sieci neuronowych.*

Do najistotniejszych osiągnięć Doktoranta Pana mgr inż. Antona Smolińskiego zaliczam opracowanie nowatorskich metod detekcji zdarzeń związanych z kierowaniem pojazdów mechanicznych z użyciem sekwencji wideo. W przeciwieństwie do pojedynczych obrazów, sekwencje tego typu, składające się z wielu, często setek czy tysięcy następujących po sobie klatek obrazowych, umożliwiają uwzględnienie dynamiki sceny, czyli zmiennej temporalnej. Jest to podejście najbardziej właściwe do skutecznej oceny trwających w czasie, a nie występujących sporadycznie zdarzeń. W tym celu Autor zaproponował, zaimplementował, a następnie zweryfikował empirycznie liczne warianty systemów składające się z różnorodnie połączonych sieci neuronowych o różnych parametrach. W szczególności, do ekstrakcji cech w różnych strumieniach danych Autor przetestował głębokie sieci splotowe typu VGGNet-16, VGGNet-19, ResNet-50 oraz Inception-V3. Sieci te nie były dodatkowo dotrenowywane, lecz pracowały w trybie ekstrakcji uniwersalnych cech obrazowych. Miało to swoje przełożenie na detekcje w innych modalnościach niż obrazy RGB. Do klasyfikacji sekwencji czasowych Doktorant Pan mgr inż. Anton Smoliński zaproponował użycie sieci typu LSTM (ang. *Long Short-Term Memory*), a w szczególności Bi-LSTM (ang. *Bidirectional LSTM*). Do tego zbadał możliwość oraz zweryfikował eksperymentalnie różne tryby połączenia strumieni przetwarzania informacji – są to tzw. moduły fuzji danych, gdzie Pan mgr inż. Anton Smoliński wypróbował tzw. fuzję „wczesną” oraz „późną”, w zależności od miejsca łączenia danych w torze systemu. Powyższe architektury, składające się z wielu sieci neuronowych i różnych modułów fuzji, zostały użyte do przeprowadzenia licznych eksperymentów. Oceniano głównie dokładność odpowiedzi w kontekście rozpoznawania wielu różnorodnych charakterystyk zachowań świadczących o potencjalnym zmęczeniu kierującego pojazdem, takich jak ziewanie, mruganie, pocieranie oczu, opadanie głowy, jak również przetestowano system w nieco prostszym, lecz być może bardziej użytecznym z praktycznego punktu widzenia trybie binarnym, w którym zadaniem systemu było podanie odpowiedzi czy kierujący wykazuje, czy też nie wykazuje oznaki zmęczenia. Szczególnie w pierwszej grupie badań, tzn. uwzględniającej wszystkie rodzaje zachowań, otrzymane rezultaty nie są co prawda perfekcyjne, ale są na wystarczająco wysokim poziomie przekraczającym 83% (szczegóły zamieszczono w pracy w rozdziale 5). W drugim przypadku, tzn. binarnym, dokładności są już znacząco wyższe osiągając pułap 98-99% na danych testowych. Doktorant Pan mgr inż. Anton

Smoliński wykazał, że połączenie wielu cech oraz wielu klasyfikatorów głębokich umożliwia realizację tego zadania z wysokim współczynnikiem dokładności odpowiedzi. W ten sposób wykazał postawioną tezę pracy.

3. *Opracowanie nowatorskich algorytmów rozpoznawania zdarzeń w pojedynczych obrazach multimodalnych z wykorzystaniem głębokich sieci neuronowych.*

Do ciekawych osiągnięć Doktoranta Pana mgr inż. Antona Smolińskiego zaliczyć można przeprowadzenie szeregu eksperymentów dotyczących możliwości skutecznej detekcji zdarzeń z wykorzystaniem pojedynczych obrazów. W szczególności Pan mgr inż. Anton Smoliński zbadał możliwość detekcji cech twarzy, takich jak obrys głowy, usta, oczy, jak również zjawisk ziewania, zakrywania ust oraz pocierania oczu. Do tego celu Autor użył wyłącznie jednej architektury splotowej sieci neuronowej MobileNet-V1. Mimo iż nie jest to najnowsza architektura głęboka, a trenowanie wykonano wyłącznie z użyciem procesora a nie karty graficznej, to uzyskane rezultaty w postaci dokładności detekcji nie są złe i w zależności od rodzaju sygnału przekraczają 84%. Są to ciekawe badania o nieco ograniczonych możliwościach aplikacyjnych choćby ze względu na dostępne nowsze i bardziej wydajne architektury typu YOLO. Jednakże pokazują one kierunki rozwoju oraz poszukiwań Doktoranta w przeciągu ostatnich lat.

4. *Opracowanie nowatorskich algorytmów rozpoznawania zachowań kierowców w obrazach multimodalnych z wykorzystaniem cech obrazowych.*

Doktorant Pan mgr inż. Anton Smoliński opracował szereg metod detekcji oraz klasyfikacji charakterystycznych cech twarzy z wykorzystaniem starszych metod takich jak histogramy orientacji gradientów (ang. HOG – *Histogram of Oriented Gradients*), falki (ang. *wavelets*), lokalne wzorce binarne (ang. LBP – *Local Binary Patterns*), jak również klasyfikatorów z wektorami wspierającymi (ang. SVM – *Support Vector Machine*), najbliższy sąsiad (ang. k-NN *k-Nearest Neighbor*), KStar oraz szereg dalszych, które jednakże dostarczały gorsze wyniki. Są to metody nieco starsze i znacząco gorsze w kryterium dokładności odpowiedzi, ale które mogą znaleźć zastosowanie jako metody pomocnicze, bądź też mogą wykazywać się dużo mniejszymi wymaganiami dotyczącym mocy obliczeniowych platform, na których są uruchomione. W tym ostatnim kontekście aż zabrakło mi przysłowiowej „kropki nad i” i zweryfikowanie tak różnorodnych klasyfikatorów we współpracującym zespole (ang. *ensemble of classifiers*). Już dawno zostało wykazane, że zespoły tego typu – choćby ze względu na dywersyfikację swoich obszarów kompetencji, jak również podział przestrzeni danych – mogą osiągać rezultaty przekraczające możliwości jednego i to często bardziej złożonego klasyfikatora. Byłoby to bardzo ciekawe skonfrontowanie metod starszego typu z metodami wykorzystującymi głębokie sieci neuronowe, czy też transformery wizyjne.

Podsumowując, teza pracy pt. *„Odpowiednia ekstrakcja, przetwarzanie oraz integracja cech obrazów wielomodalnych twarzy pozwalają na skuteczniejszą detekcję zachowań charakterystycznych dla obniżenia poziomu koncentracji użytkownika pojazdu mechanicznego w porównaniu do metod bazujących na obrazach unimodalnych”* rozprawy doktorskiej Pana mgr inż. Antona Smolińskiego została wykazana. Wszystkie opisane zagadnienia badawcze, jak i ich oryginalne rozwiązania naukowe zaproponowane przez Pana mgr inż. Antona Smolińskiego świadczą o dużej wiedzy oraz dojrzałości naukowej.

3. Poprawność

Czy stwierdzenia zawarte w rozprawie są godne zaufania? Czy uzasadnienia są poprawne? Wskaż zauważone słabości i błędy. Wskaż także te aspekty dotyczące poprawności, które są najbardziej wartościowe.

Doktorant Pan mgr inż. Anton Smoliński prowadził badania naukowe w sposób systematyczny. Po określeniu problemów badawczych i postawieniu hipotez naukowych proponował swoje i w dużej mierze nowatorskie rozwiązania, implementował je, a w końcu dokonywał ich weryfikacji eksperymentalnej. Na podstawie otrzymanych wyników dokonywał modyfikacji i powtarzał proces badawczy. W podejściu tym brakuje mi nieco większego uwzględnienia nowszych rozwiązań, ale być może Doktorant podejmie sugerowane kierunki badań w swoich przyszłych działaniach.

Należy tu dodać, że poprawność zaproponowanych rozwiązań jest też potwierdzona poprzez publikacje naukowe z udziałem Pana mgra inż. Antona Smolińskiego, w recenzowanych materiałach w przeważającej mierze o zasięgu światowym. W tym kontekście Doktorant przytacza 7 swoich prac. W jednej z nich Doktorant jest pierwszym z autorów – jest to publikacja z czasopiśmie MDPI Electronics (IF=2,6), wydana w roku 2024.

Mimo, że w pracy przedstawiono wiele rozwiązań to kilka kwestii wymaga pogłębionej analizy. Poniższa lista zawiera zauważane kwestie wymagające lepszego uzasadnienia lub dodatkowej dyskusji i wyjaśnienia.

1. Wyjaśnienia i omówienia wymaga wpływ wyboru środowiska eksperymentalnego na możliwość zastosowania otrzymanych rezultatów w warunkach rzeczywistego pojazdu mechanicznego. Zespół naukowy, do którego należy Doktorant Pan mgr inż. Anton Smoliński zdecydował się na budowę stacjonarnego stanowiska laboratoryjnego wyposażonego w kamery, jak również prosty symulator jazdy samochodem. Podejście takie ma niewątpliwe zalety, a przede wszystkim jest bezpieczne. Jednakże biorący udział ochotnicy nie tylko operują w sztucznym i nieruchomym środowisku, ale dodatkowo odgrywają przygotowane w scenariuszu sceny ziewania, mrugania, pocierania oczu itd. co może powodować pewne „odchylenie” od wariantu gdyby sceny takie w rzeczywistości były nagrywane w realnych warunkach poruszającego się pojazdu. Oczywiście to drugie podejście byłoby o wiele droższe i trudniejsze do realizacji. Tym niemniej brakuje mi w pracy choćby pobieżnej dyskusji tego problemu. Dlaczego Autorzy nawet nie spróbowali nagrać choćby kilku sekwencji w prawdziwym pojeździe? Można było też skorzystać z sekwencji tego typu nagranych przez inne zespoły badawcze i w ten sposób zweryfikować swoje podejście.
2. Oddzielnej analizy wymaga problem użyteczności każdego z badanych rodzajów sygnałów wejściowych, czyli spektrum widzialnego RGB, spektrum dalekiej podczerwieni, tzw. termowizji oraz mapy odległości. Po analizie wielu prac i rozwiązań jedynie podejścia wykorzystujące termowizję lub też bliską podczerwień wydają się być adekwatne do użycia w warunkach rzeczywistego pojazdu. Co prawda Autor wspomina o tym problemie, jednak brakuje mi głębszego omówienia tego zagadnienia.
3. Bardzo zdziwiło mnie podejście Autora zaprezentowane w rozdziale 5.3 i polegające na wykorzystaniu stosunkowo starej architektury MobileNet-V1 (jest już V2 i kolejne) do problemu detekcji i klasyfikacji charakterystycznych zjawisk występujących w pojedynczych obrazach. Zdziwiony jestem, że Autor nawet nie wypróbował innych architektur, znacznie lepiej pasujących do problemu. Mam tu na myśli choćby jedną z wersji sieci YOLO (ang. *You Look Only Once*), które nie dość że umożliwiają bardzo wysokie współczynniki dokładności

detekcji i to dla wielu klas, ale również należą do najszybszych wśród całej gamy sieci dedykowanych do detekcji obiektów w obrazach. Istnieje też olbrzymia bibliografia na temat tych sieci, jak i ich wielu modyfikacji domenowych.

Nie wiem kiedy były wykonywane te eksperymenty, ale kolejnym zaskoczeniem było dla mnie wielogodzinne trenowanie powyższych sieci wynikające z użycia wyłącznie procesora CPU, podczas gdy już standardem w tego typu treningach jest użycie choćby najprostszej karty graficznej GPU.

4. Niejako skutkiem użycia moim zdaniem nieoptymalnych do tego problemu architektur sieci splotowych są bardzo niekorzystne czasy działania systemu, zaprezentowane w rozdziale 5.2.5. Tym niemniej, doceniam że Autor wypróbował 3 platformy obliczeniowe: komputer PC, kartę graficzną NVIDIA Xavier oraz moduł wbudowany Raspberry Pi 4. W tym kontekście dziwne jest jednak, że NVIDIA Xavier o potężnej mocy obliczeniowej ma również stosunkowo niski współczynnik FPS (ang. *frames per second*). Prawdopodobnie jest to wynik użycia niezbyt optymalnej implementacji. Na przyszłość sugeruję zapoznać się ze środowiskiem obliczeniowym SDK DeepStream firmy NVIDIA, dedykowanym właśnie na te platformy. Jego użycie na pewno przyczyniłoby się do znacznego przyspieszenia obliczeń.
5. Nie wykorzystano danych testowych, są jedynie treningowe i walidacyjne, które odgrywają rolę testowych, jednakże są używane w procesie trenowania. Co więcej, zupełnie brak jest danych testowych z innych dystrybucji niż te otrzymane w laboratorium Doktoranta. A można było chociaż spróbować, gdyż sekwencje tego typu, choćby z wybranych modalności, są dostępne w Internecie. Należało też użyć wielokrotnego podziału danych na tzw. foldy, po którym powinno nastąpić kolejne uczenie oraz weryfikacja klasyfikatorów. Niemniej istotna i uzasadniona naukowo byłaby też weryfikacja klasyfikatorów na podstawie analizy hipotez statystycznych.
6. W wielu miejscach Autor porównuje otrzymane dokładności sugerując się wartościami na trzecim miejscu po przecinku wartości dokładności klasyfikacji. Uważam, że jest to działanie zbyt pochopne i słabo uzasadnione.
7. Brakuje mi próby połączenia wielu klasyfikatorów w zespół współpracujących klasyfikatorów, o których była już mowa we wcześniejszej części tej recenzji.
8. Zupełnie brakuje analizy czasów odpowiedzi opracowanych systemów, które Autor zaprezentował w rozdziale 5.4. Czy umożliwiają one działanie w warunkach rzeczywistego pojazdu?

Drobniejsze problemy i usterki zauważone w pracy są następujące:

9. Tabela 2.4 – zacytowane metody zapewne bazują na różnych zbiorach danych, więc trudno jest je porównywać (kolumna: dokładność).
10. Rozdział 3.3.1 – opisana jest metoda pozyskiwania mapy głębi za pomocą technologii czasu przelotu (ang. *Time-of-Flight*). Jednakże brakuje mi analizy na ile dane tego typu mogą być użyte w warunkach rzeczywistego pojazdu.
11. Rozdział 4 – dobrze byłoby się dowiedzieć jaki jest indywidualny wkład Doktoranta, zarówno w budowę stanowiska, jak również w akwizycję i obróbkę danych oraz opracowanie każdego z algorytmów. Dotyczy to również procesu manualnego oznaczania akcji, opisanego na str. 67.
12. Rozdział 4.3.1 – brakuje mi pogłębionej analizy problemu synchronizacji wielu kamer.

13. Str. 69 – Autor używa pojęcia „mapy cieplnej”, jest to trochę niefortunna nazwa bo sugeruje użycie wyłącznie kamer termowizyjnych, a zapewne tak nie jest.
14. Rozdział 5.1 – podana jest analiza dokładności lokalizacji punktów twarzy i podane są dokładności, ale nie jest jasne jak były mierzone?
15. Rozdział 5.2 – Autor wymienia tu pracę wielo-autorską o numerze [49], ale jaki jest indywidualny wkład pracy Doktoranta.
16. Rozdział 5.2.2 – Autor wspominał, że najlepszym rozwiązaniem okazał się klasyfikator 1-NN, ale czy były badane inne wersje k-NN, dla $k > 1$?
17. Rozdział 5.3 – przetwarzanie pojedynczych obrazów w stosunku do analizy strumienia wideo wymaga uściślenia, gdyż poprzednie modele również działają na poszczególnych klatkach, z których składa się strumień wideo – choćby obliczanie HOG czy LBB itd. Z drugiej strony śledzenie (ang. *tracking*) jest możliwe również przy zastosowaniu CNN. Przecież już nawet w schemacie na Rys. 5.2 pokazany jest krok "pobierz ramkę"...
18. Rozdział 5.3.1 – Doktorant napisał że „Modele zostały zoptymalizowane pod kątem minimalizacji funkcji straty” – a jest jakaś inna możliwość?
19. Rozdział 5.4 – Autor używa „gotowych” sieci aby uniknąć konieczności ich uczenia; jednakże sieci nie musi się trenować od podstaw, gdyż jak są dostępne dane domenowe, tak jak miał to Doktorant, to stosunkowo prosto można dotrenować już istniejący model (jest to tzw. *transfer learning*). Szczególnie istotne byłoby zastosowanie tego procesu w przypadku obrazów termowizyjnych, których cechy mogą być bardziej „odległe” od cech uzyskanych z wykorzystaniem sieci trenowanych wyłącznie na danych spektrum widzialnego.
20. Rozdział 5.4.2 – Doktorant wspomina o konieczności wyboru interwału, ale nie bardzo wiadomo jakby ten parametr wybierać w praktyce.
21. Rozdział 5.4.2 – tu Autor używa sieci ResNet-50, ale już w rozdziale 5.4.3 następuje przeskok do architektury Inception-V3, dlaczego? Autor też tutaj używa stałej wartości 25 klatek, nie bardzo tłumacząc czemu akurat ta wartość.
22. Rozdział 5.4.5 – dokładności, ale czego? Przecież w Tabeli 5.13 mamy zdefiniowane 6 rodzajów zachowań, to które są omawiane w tabeli 5.17?
23. Rozdział 5.4.5 – nie jest jednak jasne jak te cechy zostały połączone na wejściu sieci Bi-LSTM.
24. Tabela 5.21 – wymaga wyjaśnienia; dokładność połączonych modalności wynosi tutaj 0,864, ale jest to niższa wartość niż nawet jednej modalności RBG1, czyli 0,891. Podobnie ostatnia konkluzja na str. 99 – jest tak ale tylko w przypadku VGGNet19 i to tylko 0,825, co wynosi jednak mniej niż jedna modalność RGB z ResNet-50, gdzie mamy 0,891. Problem ten należy głębiej wyjaśnić.
25. Tabela 5.22 – skoro ResNet-50 jest najlepszy dla RGB to po co tu jest np. VGG?
26. Str. 101 – Autor wymienia „wysokie koszty obliczeniowe”, ale jakie rzeczywiście one są, bo nie podano ani architektury ani czasów trenowania, nie znany jest też czas predykcji systemu.

Powyższa lista pytań dotyczy dodatkowych zagadnień naukowych lub też prośby o głębsze potraktowanie pewnych tematów poruszonych w rozprawie doktorskiej Pana mgr inż. Antona Smolińskiego. Tym niemniej, mają one charakter dyskusyjny i nie umniejszają wcześniej opisanego przeze mnie wkładu naukowego Doktoranta Pana mgr inż. Antona Smolińskiego.

4. Wiedza kandydata

Które z rozdziałów (lub sekcji w rozdziałach) rozprawy omawiają istniejący stan wiedzy i dzięki temu potwierdzają ogólny stan wiedzy kandydata w zakresie Informatyki? Jakie obszary tych dyscyplin zostały omówione w tych rozdziałach/sekcjach? Jaka jest opinia recenzenta o jakości tych rozdziałów sekcji? Jaka jest opinia recenzenta o bibliografii? Na ile bibliografia jest kompletna? Prosimy o podanie innych argumentów za lub przeciw stwierdzeniu, że kandydat posiada ogólną wiedzę w dyscyplinie Informatyka techniczna i telekomunikacja.

Rozprawa doktorska Pana mgr inż. Antona Smolińskiego jest dość obszerna. Praca liczy 120 stron, na które składa się 6 rozdziałów oraz spis literatury, w których Autor zamieścił 180 prac naukowych oraz stron internetowych. Bibliografia ta jest dość kompletna i uwzględnia różnorodne publikacje w tematyce opracowywania systemów ADAS.

Niestety nie wszystkie rozdziały i zawarte w nich opisy w ogóle były konieczne. Nie bardzo rozumiem dlaczego Autor zdecydował poświęcić aż cały rozdział 3 na dość powierzchowne opisy stosunkowo leciwych już metod i algorytmów ekstrakcji cech oraz klasyfikacji, takich jak metody Viola-Jones, czy też filtry Gabora? Również zawarte tam pobieżne opisy sieci neuronowych uważam za zupełnie nietrafione i zbędne – wszystkie informacje można znacznie szybciej, a przede wszystkim znacznie precyzyjniej opisać znaleźć choćby na stronach Wikipedii, czy też w wielu łatwo dostępnych artykułach naukowych. Nieco lepiej przedstawia się rozdział 2, w którym Autor Pan mgr inż. Anton Smoliński opisał przegląd problematyki związanej z opracowaniem skutecznych metod monitorowania kierowców celem wykrywania stanu zmęczenia oraz obniżenia poziomu koncentracji. W szczególności, bardzo ciekawy jest tu opis prac prowadzonych przez firmy motoryzacyjne.

Najistotniejsze są 2 rozdziały rozprawy Pana mgr inż. Antona Smolińskiego – są to rozdziały 4 oraz 5, w których Autor opisał metodykę pozyskiwania i przygotowywania danych (rozdział 4) oraz metody ich klasyfikacji (rozdział 5).

Podsumowując, opracowanie nowych metod do ambitnego zadania stworzenia skutecznych metod analizy stanu kierowców wymagało od Pana mgr inż. Antona Smolińskiego zarówno dogłębnej wiedzy w zakresie przetwarzania obrazów, jak i budowy złożonych systemów klasyfikacji z użyciem różnorodnych architektur głębokich sieci neuronowych. Nie mniej istotne były umiejętności programistyczne Pana mgr inż. Antona Smolińskiego, który stworzył oprogramowanie, które następnie posłużyło do weryfikacji eksperymentalnej zaproponowanych rozwiązań.

5. Podsumowanie

Stwierdzam, że recenzowana rozprawa doktorska Pana magistra inżyniera Antona Smolińskiego spełnia warunki określone w art. 13.1. Ustaw z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki (z późn. zm.). Wnioskuje o dopuszczenie Pana magistra inżyniera Antona Smolińskiego do dalszych etapów przewodu doktorskiego.


Podpis