Prof. dr hab. inż. Andrzej Czyżewski
Politechnika Gdańska, Wydział ETI
Katedra Systemów Multimedialnych

02. 08. 2023 r.

Opinia o rozprawie doktorskiej mgr inż. **Marcina Platy**
pt.: „Zastosowanie zaawansowanych metod sztucznej inteligencji do wybranych problemów bezpieczeństwa informacji"
przygotowanej w Instytucie Informatyki PAN pod kier. prof. dr hab. inż. Marka Klonowskiego
(promotor pomocniczy dr inż. Piotr Syga)

Rozprawa obejmuje 6 zagadnień. Dwa z nich dotycząc znakowania wodnego, jedno rozpoznawania twarzy, kolejne - zapobieganiu atakom w procesie rozpoznawania mówcy i ostatnie - identyfikacji osób na podstawie konturu dłoni. Ponieważ są to różne zagadnienia, budowę rozprawy należy uznać za nietypową i nie w pełni spójną.

Rozprawa obejmuje 158 stron tekstu wraz z rysunkami i tablicami, jednak treści dotyczące poszczególnych zagadnień mieszczą się w zakresie 20-30 stron, są więc opisane stosunkowo krótko. Również wykaz bibliografii, obejmujący 199 pozycji, dotyczy sześciu w znacznym stopniu odrębnych zagadnień, więc dzieli się on ilościowo pomiędzy poszczególne tematy, zatem żadne z zagadnień nie zostało zaprezentowane w oparciu o przegląd literatury, który można by uznać za gruntowny. Opisy badań związanych z poszczególnymi tematami nie zawierają wyraźnie sformułowanej tezy, mają raczej formę sprawozdania z wykonanych eksperymentów cząstkowych, poprzedzonego krótkimi przeglądami dotychczasowych rozwiązań i zakończonymi krótkimi podsumowaniami sformułowanymi w formie autorecenzji. Całą pracę poprzedza streszczenie oraz 3,5 stronicowy rozdział, w którym Autor omówił strukturę pracy. W kończącym rozprawę rozdziale 7 Autor próbuje powiązać przedstawione w poprzednich rozdziałach osobne tematy, argumentując, że ich spoiwem są możliwości zastosowania metod zaliczanych do sztucznej inteligencji w dwóch szerszych zagadnieniach, jakimi są znakowanie wodne i wybrane metody biometryczne.

Układ pracy mógłby sugerować, że praca ma postać przewodnika po publikacjach napisanych z udziałem Autora, ale narracja na ogół istotnie odbiega od ujęcia treści w tych publikacjach. Autor powołuje się na 3 współautorskie referaty, oraz na artykuł opublikowany w czasopiśmie Pattern Recognition i na materiał mający formę preprintu. Jak, zatem, napisano już wcześniej, praca jest skonstruowana nietypowo, nie jest ona ani spójną monografią, ani formalnym przewodnikiem po publikacjach Autora, a raczej podsumowaniem rozwijanych z jego udziałem kilku wątków badawczych, z których część, dotycząca znakowania wodnego, jest owocem prowadzonych projektów OPUS (NCN) i POIR (NCBR). Brak wyodrębnienia tez i ich formalnego udowadniania dodatkowo oddala sposób ujęcia rozprawy od powszechnie przyjętych zasad.

Pomimo nietypowego zamysłu i układu, rozprawa doktorska jest interesująca i dowodzi obszernej wiedzy Doktoranta i jego biegłości w rozwiązywaniu problemów naukowych o potencjalnym znaczeniu praktycznym, może być więc poddana ocenie jako praca kwalifikacyjna.

Rozprawa została napisana w liczbie mnogiej. Czytelnik, wobec tego, stale odnosi wrażenie istnienie współautorstwa pracy. Ta maniera redakcyjna nie tylko jest myląca w odniesieniu do wielu fragmentów, ale także miejscami wygląda pretensjonalnie. Typowa dla języka pisanych po polsku prac doktorskich w dziedzinie nauk technicznych jest forma bezosobowa, ze zmianą jej na osobową we fragmentach, które szczególnie wymagają wskazania na autora, bądź autorów danej koncepcji, czy wyniku. Autor również samodzielnie i arbitralnie określa swoją rolę w publikacjach, które wszystkie mają charakter wieloautorski, przypisując sobie w ich powstaniu wiodącą rolę, co jednak nie zostało potwierdzone oświadczeniami współautorów, gdyż recenzent nie otrzymał ich kopii. W tym miejscu czytelnik rozprawy upewnia się, że rozprawa nie została pomyślana jako ścisły przewodnik po publikacjach, bo omawianych w pracy zagadnień jest więcej, niż w wymienionych publikacjach, których zresztą również nie dostarczono w formie załączników.

W rozdziale drugim oraz trzecim Autor porusza problem znakowania wodnego w kontekście podatności na różnego rodzaju ataki, takie jak kompresja stratna obrazu. Proponuje w rozdziale drugim zastosowanie sieci neuronowej do znakowania wodnego obrazów.
W rozdziale 3 omawiane jest zagadnienie znakowania wodnego rozszerzone o schemat dyskryminator-detektor.

Uwagi ogólne: porównanie jakościowe obrazów pomiędzy metodami opisanymi w rozdziałach drugim i trzecim mogłoby zostać pokazane na jednakowych obrazach, co dałoby możliwość oceny wpływu każdej z metod na jakość obrazu. Komentarz do tabeli 2.3: wydaje się oczywiste, że najlepsze wyniki sieci uzyskują na podobnych danych które zostały zastosowane do treningu. Obserwacja Autora tylko potwierdza tę sytuację - skuteczność najlepsza przy uczeniu sieci jedną grupą ataków daje dobre wyniki w testach.

Uwagi redakcyjne do rozdziałów drugiego i trzeciego:
Str. 17 rysunek dwa jeden: brak objaśnienia symbolu metryki LE oraz LC
W dalszej części pracy wydają się one nieużywane. <- jest, na str. 29
Str. 21 wygładzanie gaussowiskie —> gaussowskie
Str. 23 takich składający się —> składających się
Strona 24 w trakcie eksperymentach —> w trakcie eksperymentów
Strona 29 wzór 2.10: Objaśnienie symbolu lambda znajduje się dopiero przy wzorze 2.11
Strona 34 niewykorzystanego w procesie treningowy —> w procesie treningowym
Strona 35 stenograficznych —> steganograficznych
Strona 43 rysunek 3.1: brak czerwonych strzałek wymienionych w podpisie rysunku. Na rysunku występują wyłącznie strzałki koloru czarnego.

Rozdział 4
W rozdziale czwartym Autor porusza zagadnienie prywatności w biometrii wizerunku twarzy.
Dokonuje arbitralnego przeglądu literatury pod kątem zagrożeń prywatności w biometrii twarzy.
Wskazuje potencjalne zagrożenie metod biometrycznych wynikające chociażby z możliwości odtworzenia wizerunku twarzy z reprezentacji wektorów cech.
Autor proponuje metodę rozpoznawania twarzy wykorzystującą mimikę twarzy wyrażoną poprzez analizę przemieszczenia punktów charakterystycznych twarzy. W tym celu konstruuje bazę danych nagrań osób których zadaniem jest przeczytanie krótkiego tekstu. Nagrania Autor przeprowadza za pomocą popularnej kamery stosowanej w większości telefonów komórkowych.
Następnie za pomocą opisanych w literaturze algorytmów wykrywania twarzy oraz wykrywania 68 punktów charakterystycznych twarzy oblicza trajektorię punktów charakterystycznych podczas wypowiadania przez poszczególne osoby zdefiniowanych zdań.
W celu zwiększenia ochrony prywatności Autor dokonuje normalizacji współrzędnych punktów charakterystycznych względem cechy osobniczej każdego badanego.
Podjęto próby wyboru najbardziej znaczących, w kontekście skuteczności rozpoznawania osób, punktów charakterystycznych ze zbioru 68 punktów, jednakże należy uznać te badania jako wstępne ze względu na małą liczność autorskiego zbioru nagrań wynoszącą 34 osoby.
Tym niemniej, uzyskane wyniki w kontekście ochrony prywatności próbek biometrycznych wydają się zadowalające. Autor przeprowadza dyskusję odnośnie możliwości ataku polegającego na rekonstrukcji wizerunku twarzy. Wykazuje niskie prawdopodobieństwo rekonstrukcji ze względu na niską korelację występującą pomiędzy wyznaczonymi proponowaną metodą cechami.

Strona 83 rysunek 4.7: na rysunku pokazano linie w dwóch kolorach niebieskim i zielonym natomiast w podpisie rysunku jest wymieniona dwukrotnie linia zielona. Poza tym podpis osi poziomej w postaci „liczba punktów" jest mylący i raczej powinien brzmieć „liczba cech". Dodatkowo kolejność liczb na osi poziomej mogłaby być pokazana nietypowo, w porządku malejącym, co lepiej odzwierciedliłoby pokazywaną redukcję liczby cech.

W opinii recenzenta należałoby dodać informację o współczynnikach identyfikacji (IR) osiąganych przez współcześnie stosowane systemy biometrii twarzy (state-of-the-art).

Pewnym ograniczeniem zaproponowanego rozwiązania jest konieczność wypowiedzenia przez użytkownika zdefiniowanej frazy, co może prowadzić do błędów w przypadku pomyłki podczas wypowiadania zdania. Ponadto można zauważyć pewną nieścisłość w rozdziale 4.3.1 przy opisywaniu procedury porównywania wektorów przesunięć. Jak napisano, „algorytm DTW liczy podobieństwo dla każdego z 136 punktów charakterystycznych". Wydaje się, że punktów charakterystycznych dla jednej reprezentacji twarzy jest 68, natomiast wykonywana operacja spłaszczenia tensora powoduje uzyskanie 136 nie punktów charakterystycznych, lecz cech. Sformułowanie „cech" pojawia się w dalszej części pracy.

Uwaga do rysunku 4.8 b: podzbiór cech wskazanych jako osiągających największą skuteczność identyfikacji jest raczej kwestią zależną od bazy danych. Przykładowo, dlaczego punkt charakterystyczny nr 8 jest reprezentatywny zarówno w przemieszczeniu poziomym jak i pionowym, a punkt nr 9 lub 10 tylko w przemieszczeniu pionowym? Czy „symetryczny" do

punktu 8 punkt 10 nie powinien „zachowywać się” podobnie? Może to wpływ warunków oświetleniowych?
Ponadto Autor nie poruszył kwestii występowania znaczących ruchów całej twarzy podczas wypowiedzi, które mogą wprowadzać zakłócenia do liczonych cech.

W literaturze można znaleźć podobne prace:
https://cedar.buffalo.edu/~govind/CSE666/fall2007/face_expression_biometrics.pdf
https://ieeexplore.ieee.org/document/4270392
Różnica polega głównie na tym, że wyniki uzyskano nie na podstawie czytania tekstu, ale ekspresji emocji.

Rozdział 5.
W rozdziale piątym Autor porusza zagadnienie związane z zapobieganiem atakowi polegającemu na podstawianiu głosu w torze fonicznym w systemie rozpoznawania mówców.
Ataki tego typu polegające na nagraniu oryginalnego głosu a następnie odtworzeniu go na wejściu systemu biometrycznego są dość łatwe w realizacji i jednocześnie cały czas trudne do wykrycia.
Podrozdział 5.4 zawiera rozbudowany, sformalizowany matematycznie, opis systemu rozpoznawania mówcy z zastosowaniem algorytmu wykrywania podstawienia głosu.
Autor podejmuje próbę zastosowania sieci neuronowej do wykrycia próby podstawienia głosu. Wykorzystuje zbiór danych udostępniony w ramach konkursu ASVspoof challenge 2019. W pierwszej fazie badań Autor próbuje wyizolować grupy cech reprezentowanych przez wektory otrzymane na wyjściu sieci neuronowej, składającej się z trzech warstw splotowych oraz czterech warstw liniowych z zastosowaniem opisywanej w literaturze funkcji kosztu n-pair angular loss bazującej na funkcji triplet loss. Docelowo grupy wektorów cech reprezentujące próbki dźwięku odtwarzanego z różnych odległości i różnej jakości sprzętem powinny być odseparowane od próbek dźwięku bezpośredniego. Po 200 epokach treningu udało się jedynie wyodrębnić niezależny klaster zawierający tylko próbki autentyczne i jednak nie ma rozdzielenia próbek nagranych z różnych odległości. Dalsze eksperymenty mające na celu wykrycie ataku podejmowane są w oparciu o opracowaną przez Autora technikę Attack-Out Cross-Validation, która jest adaptacją standardowej k-krotnej walidacji krzyżowej. Technika zrealizowana jest jako potrójna walidacja krzyżowa, gdzie zdefiniowany jest każdorazowo zbiór treningowy, testowy i walidacyjny. Dane do poszczególnych zbiorów dobierane są w taki sposób, że metody ataku o tych samych parametrach nie występują w różnych zbiorach w ramach jednego przebiegu, kroku (fold). Przeprowadzany jest trening dwóch modeli: bayesowskiej sieci neuronowej oraz lekkiej splotowej sieci neuronowej. W celu próby zniwelowania problemu nadmiernego dopasowania danych Autor stosuje techniki regularyzacji, takie jak uczenie wielozadaniowe, kombinacja liniowa elementów wejściowych, powielanie danych. Lepsze wyniki osiągnięte zostały za pomocą sieci lekkiej splotowej. Dalsze eksperymenty pokazują wyniki powstałe poprzez ważone połączenie wyników dwóch modeli, które, jak można się spodziewać, wykazuje poprawę skuteczności działania.

Strona 123 uczenie wielozadaniowe (ang. multi-task learninig —> learning

Rozdział szósty.
W rozdziale szóstym Autor opisuje zagadnienie identyfikacji opartej na konturze dłoni. Omówione zostały wcześniejsze prace w tej dziedzinie opisane w literaturze. W kolejnych podrozdziałach opisano sposób pobierania próbek konturu dłoni za pomocą skanera biurowego oraz metody wyznaczania punktów charakterystycznych dłoni i ekstrakcja cech biometrycznych. Wątpliwości budzi opis algorytmu wyznaczania punktów wgłębienia pomiędzy przedramieniem i kciukiem w okolicy nadgarstka. Czy możliwe jest przeprowadzenie tego procesu w sposób pełni automatyczny, to znaczy bez określania manualnego dodatkowych danych wejściowych odnośnie np. położenia kciuka?. Wyodrębniane są łącznie 62 cechy biometryczne konturu dłoni, przy czym 12 z nich to autorskie współczynniki określające krzywizny 4 palców (oprócz kciuka). Podobieństwo konturów dłoni obliczane jest jako pewna odległość pomiędzy wektorem uśrednionych cech (od jednego do trzech skanów dłoni), zapisanych w bazie z wektorem złożonym przez użytkownika podlegającego weryfikacji lub identyfikacji.
Dodatkowo Autor przeprowadza analizę znaczenia poszczególnych cech dla procesu rozpoznawania dłoni i wyeliminował te, które wprowadzały zakłócenia do miary odległości. Określił też przydatność, połączonych z podstawowymi, dodatkowych współczynników krzywizny palców poprzez obliczenie miar skuteczności weryfikacji oraz identyfikacji osób na podstawie 50 współczynników bazowych osobno i 12 dodatkowych osobno (zebrane w tab. 6.5, w porównaniu do tab. 6.4). Autor dowiódł wysokiej skuteczności działania opracowanej metody, osiągając współczynnik fałszywej akceptacji równe zero przy współczynniku fałszywego odrzucenia wynoszącym 1,19% dla własnej, stosunkowo niewielkiej, bazy danych zawierającej skany konturów dłoni 60 osób.
Po lekturze poprzednich rozdziałów czytelnik mógłby się spodziewać dyskusji na temat ataku na powyższy system polegającego chociażby na podstawieniu obrazu zeskanowanej dłoni wydrukowanego na papierze. W tym kontekście można odczuwać pewien niedosyt i niespójność z pozostałymi rozdziałami pracy.

Uwagi do bibliografii.
Pozycja [MK14] - brak autorów.
Według źródeł to: Nasrin M. Makbol, Bee Ee Khoo, A new robust and secure digital image watermarking scheme based on the integer wavelet transform and singular value decomposition, Digital Signal Processing, Volume 33, 2014, Pages 134-147, ISSN 1051-2004, https://doi.org/10.1016/j.dsp.2014.06.012

Wniosek

Rozprawa pana mgr inż. Marcina Platy została zrealizowana w sposób odzwierciedlający wymagane kwalifikacje jej Autora, wystarczający nakład pracy badawczej, implementacyjnej i eksperymentalnej, jak również jej część jej wyników została opublikowana na konferencjach i w treści dwóch artykułów. W mojej opinii treść rozprawy mgr inż. Marcina Platy, pomimo jej nietypowej konstrukcji formalnej, spełnia zatem wymogi Prawa o Szkolnictwie Wyższym i Nauce, z dnia 20 lipca 2018 r. (Dz. U. 30. 08. 2018 r. Poz. 1668), stawiane kandydatom do stopnia naukowego doktora.